Assimilating realistically simulated wide-swath altimeter observations in a high-resolution shelf-seas forecasting system

Robert R. King[1] and Matthew J. Martin[1]

[1]Met Office, Exeter, UK

Correspondence: Robert R. King (robert.r.king@metoffice.gov.uk)

Abstract.

The impact of assimilating simulated wide-swath altimetry observations from the upcoming SWOT mission is assessed using Observing System Simulation Experiments (OSSEs). These experiments use the Met Office 1.5 km resolution North-West European Shelf analysis and forecasting system. In an effort to understand the importance of future work to account for

5   correlated errors in the data assimilation scheme, we simulate SWOT observations with and without realistic correlated errors. These are assimilated in OSSEs along with simulated observations of the standard observing network, also with realistic errors added.

It was found that while the assimilation of SWOT observations without correlated errors reduced the RMSE in sea surface height (SSH) and surface current speeds by up to 20%, the inclusion of correlated errors in the observations degraded both

10 the SSH and surface currents, introduced an erroneous increase in the mean surface currents, and degraded the sub-surface temperature and salinity. While restricting the SWOT data to the inner half of the swath and applying observation averaging with a 5 km radius negated most of the negative impacts, it also severely limited the positive impacts. To realise the full benefits in the prediction of the ocean mesoscale offered by wide-swath altimetry missions it is crucial that methods to ameliorate the effects of correlated errors in the processing of the SWOT observations and to account for the correlated errors in the

15 assimilation are implemented.

20 # 1 Introduction

Satellite altimeter observations of the sea surface height (SSH) have been available for over 25 years and are routinely assimilated in operational global ocean analysis and forecasting systems (e.g. Waters et al., 2015; Oke et al., 2015; Lellouche et al., 2018). The global coverage of a constellation of altimeters provides constraints on the modelled SSH and geostrophic velocity

field, including the position and strength of mesoscale eddies. The SSH data can also be used to adjust the sub-surface density

25 structure (e.g. Weaver et al., 2005). The complementarity of satellite altimetry and in situ measurements of the temperature
and salinity was demonstrated by Lea et al. (2014) where it was shown that assimilating SSH observations improves the model
forecast of other variables even where more direct measurements are available.

Due to the large gaps between tracks and long repeat cycles (10–35 days), existing along-track altimeter observations provide limited spatial and temporal resolution compared with the scales of the ocean mesoscale and sub-mesoscale dynamics.

30 Various studies have attempted to quantify the impact of different numbers of standard along-track (nadir) satellite altimeters
on ocean analysis and forecasting systems. For instance, Jacobs et al. (2014) showed that even with 4 nadir altimeters, while
the SSH itself can be reasonably well constrained, other parameters related to the prediction of frontal positions were not well
constrained. Although along-track observations can have a sampling frequency of ∼7 km, various sources of noise limit the
feature resolution to ∼100 km (Xu and Fu, 2013). This is already sub-optimal for the initialisation of mesoscale features in

35 the current generation of operational ocean prediction systems. With the move to higher-resolutions, this disparity between the
observed and modelled scales will increase.

The operational shelf-seas short-range prediction system at the Met Office uses an eddy-resolving 1.5 km model in the
North-West European Shelf (NWS; Tonani et al., 2019) and delivers products to various users, including the Copernicus
Marine Service (https://marine.copernicus.eu). The current operational system assimilates altimeter observations only in the

40 deep water (>700 m depth) regions of the domain (King et al., 2018). Satellite and in situ sea surface temperature (SST) data,
and in situ profiles of temperature and salinity, are assimilated throughout the domain. Given the sparseness of the in situ
observations over this region (De Mey-Frémaux et al., 2019), satellite altimeters could provide potentially useful information
on the vertically-integrated water column across the domain. However, if there are insufficient in situ observations to constrain
the sub-surface density structure, the assimilation of SSH observations can lead to unrealistic changes to the sub-surface density

45 (Lea et al., 2014) and an overall degradation in forecast quality. As well as inferring corrections to sub-surface properties, SSH
data from altimeters could help to constrain processes such as tides and storm surges which are represented in the model and
are observed by altimeters. However, the sampling of these processes by existing nadir altimeters is not currently sufficient to
adequately constrain them in the NWS region.

Upcoming wide-swath altimetry missions, such as SWOT (Durand et al., 2010) and COMPIRA (Uematsu et al., 2013),

50 will provide a step-change in the ability to observe ocean mesoscale features by measuring the SSH along a wide swath and
providing 2D measurements of the ocean surface topography. The SWOT mission will have a 120 km wide swath with a 20 km
central gap combined with a nadir altimeter and aims to provide an effective resolution of ∼15 km (Morrow et al., 2018), but
with sampling of ∼2 km within the swath. As with the current nadir altimeters, the long repeat cycle of SWOT (21 days, 11 day
average repeat time) means that although high-resolution 2D maps of the SSH will be available, there will be long periods

55 where each location is unobserved.

Wide-swath altimetry observations will be subject to instrument noise and large correlated geophysical and instrumental
errors (Gaultier et al., 2016). Although strategies to reduce the magnitude of some of these errors may be developed before
routine dissemination of the observations, this presents a challenge for data assimilation schemes. Most operational data as-

simulation schemes assume observation errors to be uncorrelated. Common methods to ameliorate the effects of correlated
60 observation errors are to thin or average nearby observations, or to increase the observation error variances to avoid over-fitting
the data (e.g., Fowler et al., 2018). All these common methods suffer from the fact that they do not allow the data assimilation
to make best use of the small-scale information contained in the observations. There has been work to extend data assimilation
schemes to explicitly account for correlated observation errors. For example, Bédard and Buehner (2020) used a 1D model to
demonstrate the combination of direct and spatial difference observations to account for observation error correlations, while
65 Ruggiero et al. (2016) and Guillet et al. (2019) parameterised the observation error covariances using the diffusion equation
to avoid the need to invert a non-diagonal covariance matrix. On the other hand, Metref et al. (2020) used the correlated-
error reduction (CER) method to account for correlated errors when assimilating SWOT observations in a one-and-a-half layer
quasi-geostrophic model. These methods are fairly immature in terms of implementation in operational systems, but should be
a focus for implementation over the coming years.

A number of studies have also examined the potential impact of assimilating SWOT observations in operational-like ocean
70 analysis and forecasting systems. Carrier et al. (2016) demonstrated positive impacts from assimilating simulated SWOT
observations with the full correlated errors into a 6 km resolution regional model, where the SWOT observations were thinned
to ∼20 km separation between observations. D'Addezio et al. (2019) found that assimilating simulated SWOT observations
with only uncorrelated errors in a 1 km model decreased the minimum constrained length-scale, but did not significantly
improve area-averaged errors. On the other hand, Bonaduce et al. (2018) and Li et al. (2019) find significant improvements in
75 regional models from assimilating SWOT observations, but in both cases the simulated observations did not include correlated
errors. The limitations of ignoring the effects of one of the expected sources of correlated errors was explored by Le Henaff
et al. (2008) using a barotropic model of the North Sea.

The aim of this work is to assess the benefit of assimilating realistic SWOT-like observations within our current shelf-
80 seas forecasting system and to examine some simple strategies to address the impact of the large correlated errors expected
in realistic wide-swath altimeter observations. Our intention is initially not to adapt the data assimilation scheme to account
for correlated errors. Rather we aim to quantify the impact that might be expected when mitigating for their effects and to
understand the issues that the full set of expected correlated errors bring about when used in the assimilation. Here we have
investigated the impact of restricting the SWOT data assimilated to the inner half of the swath (to discard the largest errors)
85 and the impact of applying median-averaging of the observations with different averaging radii.

We use Observing System Simulation Experiments (OSSEs), as described in Section 2, to perform idealised tests of the
assimilation of SWOT observations alongside the existing network of satellite and in situ observations in the 1.5 km model
of the North-West European Shelf mentioned earlier. This system uses the NEMO model together with the NEMOVAR data
assimilation system which are described in Section 3. The details of the simulation of the observations assimilated in the OSSE
90 are presented in Section 4. The results of the experiments are presented in Section 5 and a discussion of the main points to
come out of the results is given in Section 6. Finally our conclusions are presented in Section 7.

Table 1. Experimental set-up

Experiment	Standard Observations	Swath width	SWOT errors	Superob radius	Period		Surface forcing
Nature Run	-	-	-	-	01 Jan 2017	31 Dec 2018	MetUM
Free Run	-	-	-	-	01 Jul 2017	31 Dec 2018	ECMWF IFS
Control	Y	-	-	-	01 Jul 2017	30 Jun 2018	ECMWF IFS
HalfSWOT	Y	Half width	All	-	01 Jul 2017	30 Jun 2018	ECMWF IFS
SWOT	Y	Full width	All	-	01 Jul 2017	30 Jun 2018	ECMWF IFS
LowErrSWOT	Y	Full width	Uncorrelated	-	01 Jul 2017	30 Jun 2018	ECMWF IFS
HalfSWOT_5km	Y	Half width	All	5 km	01 Jul 2017	30 Jun 2018	ECMWF IFS
HalfSWOT_20km	Y	Half width	All	20 km	01 Jul 2017	30 Jun 2018	ECMWF IFS
SWOT_5km	Y	Full width	All	5 km	01 Jul 2017	30 Jun 2018	ECMWF IFS
SWOT_20km	Y	Full width	All	20 km	01 Jul 2017	30 Jun 2018	ECMWF IFS

2 Experiment Design

An OSSE is used to assess the impact on ocean analyses and forecasts of potential changes to the current observing system and consists of several components: a Nature Run, observations simulated from the Nature Run, and additional OSSE runs using a different model set-up into which the simulated observations are assimilated. Examples of best practices for running OSSEs include Hoffman and Atlas (2016) and Halliwell et al. (2017).

The Nature Run provides the truth against which other experiments are assessed and from which we sample observations. To draw useful conclusions about changes to the observing system, the Nature Run needs to be a realistic representation of the real ocean. This is generally accomplished by using the highest resolution model available. To emulate an operational ocean analysis system, realistic errors must be added to the simulated observations sampled from the Nature Run. In Section 4, we detail how we have simulated the observations for each type assimilated in these experiments.

For the OSSE runs, the ocean model must differ in enough respects that there is sufficient error growth between the Nature Run and OSSEs to emulate the differences between an operational system and the real ocean. Ideally, an OSSE would use a separate lower resolution ocean model with different parameterisations, surface and lateral boundaries, and be initialised from a different (though realistic) state (Halliwell et al., 2014). A Control Run, in which some standard network of observations is assimilated, then provides a baseline against which to determine the additional benefit (or detriment) of introducing assimilation of additional observations. Further runs of this model assimilating extra observations in addition to the standard set are used to assess the impact of the additional observations.

For this work, we have employed the same high resolution shelf-seas model in all experiments: a lower resolution model would be unable to resolve the scales which are resolved in the higher resolution model and are expected to be observed with wide-swath altimeter missions. Also, we do not have access to a long run of a realistic higher resolution model over the

domain of interest. To introduce realistic errors, we have forced the Nature Run and the OSSEs with different atmospheric fields (detailed in Section 3.1) and have selected initial conditions valid at the correct time of year, but taken from different existing runs. In Section 3.1.1, we will compare the differences in free-running versions of the Nature Run and OSSE models. A

115   Free Run of the OSSE model, although not strictly necessary, allows an assessment of the errors introduced by the differences between the two models before introducing assimilation. This gives a baseline for the error reduction that assimilation can provide.

Here we have performed a Nature Run, a Free Run, a Control Run, and 7 further experiments (as summarised in Table 1). The Nature Run covered the period 01 Jan 2017 – 31 Dec 2018, while the Free Run started from 01 July 2017 and ran until

120   31 Dec 2018. The additional 6 months at the start of the Nature Run was to allow the effects of the differing surface forcing to spin-up as all experiments were initialised from runs which had been forced by ECMWF (European Centre for Medium Range Weather Forecasts) fluxes (as used in the OSSE runs). The 8 OSSE runs all cover the period 01 July 2017 – 30 June 2018. The Nature and Free runs do not assimilate any observations, but the Control Run assimilates simulated Sea Surface Temperature (SST) observations, in situ temperature and salinity (T/S) profiles, and nadir Sea Level Anomaly (SLA) observations from 4

125   satellite altimeters representative of the current altimeter constellation (Jason-3, AltiKa, Sentinal-3A, and Sentinal-3B). The SWOT satellite, which will have a 120 km wide swath with a central 20 km gap, will also carry a nadir altimeter which will observe along the track in the centre of this swath. All of the OSSEs which include SWOT observations here also include the SWOT nadir altimeter.

In addition to the Control experiment, seven further OSSEs (detailed in Table 1) were performed assimilating the standard

130   set of observations assimilated in the Control run plus additional simulated SWOT observations. In the SWOT experiment, we assimilated the full SWOT data with the full correlated errors, while in the HalfSWOT experiment we assimilated only the inner half of the SWOT swath (i.e., 10–35 km from the nadir point on both sides of the swath). In the LowErrSWOT experiment we again assimilated the full swath, but in this case applied only the uncorrelated KaRIn (Ka-band Radar Interferometer) errors to the simulated observations, not the full correlated errors included in the SWOT and HalfSWOT experiments. The

135   various components of the errors in SWOT observations are detailed further in Section 4.2. Finally, we repeated the SWOT and HalfSWOT experiments with median-averaging (super-obbing) of the SWOT observations using a 5 km and a 20 km averaging radius.

3   Model and Assimilation Scheme

3.1   Ocean Models

140   The operational Met Office North-West European Shelf ocean forecasting system (NWS, Tonani et al., 2019) which delivers forecast products to the Copernicus Marine Environment Monitoring Service (CMEMS) is used as the basis for all experiments in this work. This system couples the NEMO ocean model with the WAVEWATCH III spectral wave model (Saulter et al., 2017; Lewis et al., 2019). The domain (shown in Figure 1) extends from approximately 45°N, 20°W to 63°N,11°E on a rotated pole grid, covering the North-West European Shelf and part of the North-East Atlantic ocean with a horizontal resolution of
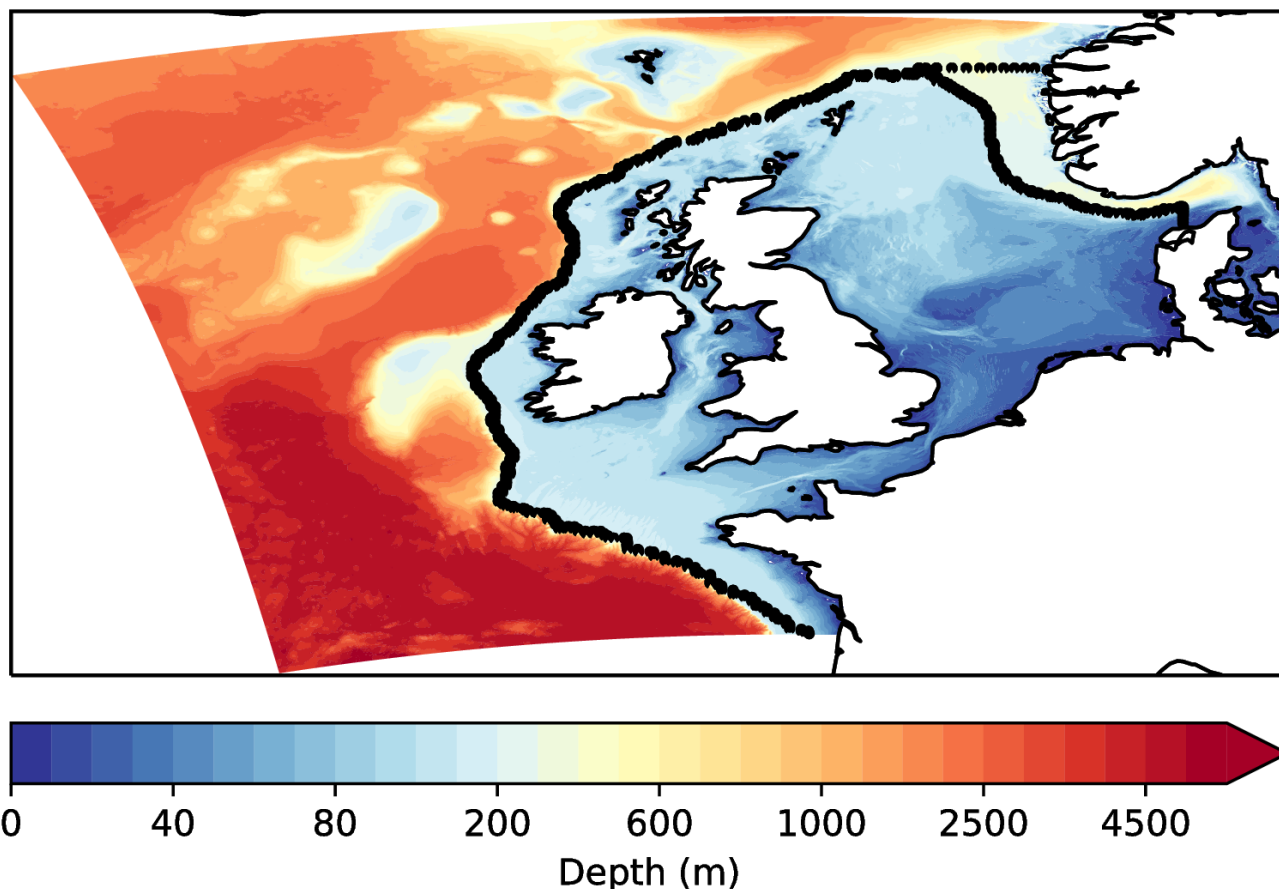
Figure 1. The bathymetry of the 1.5 km resolution North-West European shelf system. The boundaries between the off-shelf, on-shelf, and Norwegian Trench regions are marked by a black line. Note that a non-linear colour-sale has been used to highlight the complex bathymetry.

145 ~1.5 km. This domain has a complex bathymetry, containing regions of shallow water, a continental shelf, and regions of deep ocean (depth$\gtrsim$5000 m). The resulting dynamics produce mesoscale eddies along with seasonally-stratified and well-mixed regions with marked boundaries. The on-shelf and off-shelf regions used in the analysis here are shown in Figure 1. Broadly, the on-shelf region is defined by the seas around the UK, Ireland and continental Europe where the depth is less than 700 m and excluding the Norwegian Trench and Baltic boundary region.

150 The ocean model component of the NWS system uses the eddy-resolving 1.5 km Atlantic Margin Model configuration (AMM15, Graham et al., 2018) of NEMO version 3.6 (Madec, 2008). The model employs a non-linear free surface, 51 vertical terrain-following levels and includes 11 tidal constituents, specified on the open boundary via the Flather radiation condition (Flather, 1981). The AMM15 is nested within the Met Office 1/12° North Atlantic model (Storkey et al., 2010; Blockley et al., 2014) which provides the lateral boundary conditions, except for the Baltic boundary which is taken from the CMEMS Baltic

155   Sea model (Berg and Poulsen, 2012). The ocean and wave components are coupled hourly using the OASIS3-MCT coupler
(Craig et al., 2017).

The operational system is forced at the surface using 3-hourly wind and atmospheric pressure fields, radiative and E-P
(evaporation-minus-precipitation) fluxes from the ECMWF operational Integrated Forecasting System. For the OSSE runs we
have used the same surface forcing as used in the operational NWS system, but the Nature Run is forced at the surface by the

160   Met Office Unified Model (MetUM) operational global atmospheric model.

3.1.1   Assessment of the Nature Run and Free Run

Graham et al. (2018) shows that the AMM15 configuration on which our experiments are based provides a realistic represen-
tation of the mean state and variability of the North-West European Shelf. It is able to reproduce the large-scale circulation as
well as dynamical features such as mesoscale eddies, frontal jets and internal tides. Although biases are present in the mean

165   surface temperatures and depth of the mixed layer, the model reproduces the variability in the observed long-term average
surface temperatures, and the seasonal stratification across the domain.

As mentioned above, the Free Run uses the same ocean model configuration as the Nature Run, but with different atmo-
spheric forcing and a different initial condition. As expected, the different surface forcing affects the ocean mixed layer, with
the off-shelf average mixed layer being up to $\sim$100 m deeper in the Free Run in winter, but with little difference on-shelf in

170   any season. The mean and standard deviation of the SSH over the year is very similar between the Free Run and the Nature
Run, with the SSH standard deviation being $\sim$3% lower on- and off-shelf in the Free Run (compared to the NR). However, as
is shown in Figure 2, the bias in the daily mean SSH highlights errors introduced by the differences between the simulations
(and which assimilation aims to reduce). The impact of the different initial conditions and surface forcing on the Free Run 3D
temperature and salinity biases is shown in Figure 3 for the full 18-month period of these runs. The deeper mixed layer depth

175   (MLD) in the Free Run results in cooler temperatures within the mixed layer (due to broadly similar irradiative fluxes between
the atmospheric forcing datasets), but with little difference in the off-shelf temperature below the mixed layer. The large initial
salinity bias, particularly off-shelf, remains until near the end of the 18-month runs. The eventual convergence here may be a
result of using the same lateral boundaries and river inputs for both runs.

3.2   Data Assimilation Scheme

180   The NEMOVAR data assimilation scheme is used in all configurations of the operational Met Office ocean analysis and fore-
casting systems. The Met Office NWS ocean forecasting system uses NEMOVAR in its incremental 3D-VAR FGAT (First
Guess at Appropriate Time) configuration. This is a multi-variate, multi-lengthscale assimilation scheme developed collabora-
tively by the Met Office, Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS), ECMWF,
and Institut National de Recherche en Informatique et en Automatique (INRIA) (Mogensen et al., 2009, 2012; Waters et al.,

185   2015). The implementation is based on that used in the lower resolution AMM7 system (7 km resolution) as described by King
et al. (2018) with modifications as described by Tonani et al. (2019). Two background error correlation length-scales are used
in the 1.5 km NWS system: 100 km for the long length-scale and the Rossby radius of deformation for the short length-scale
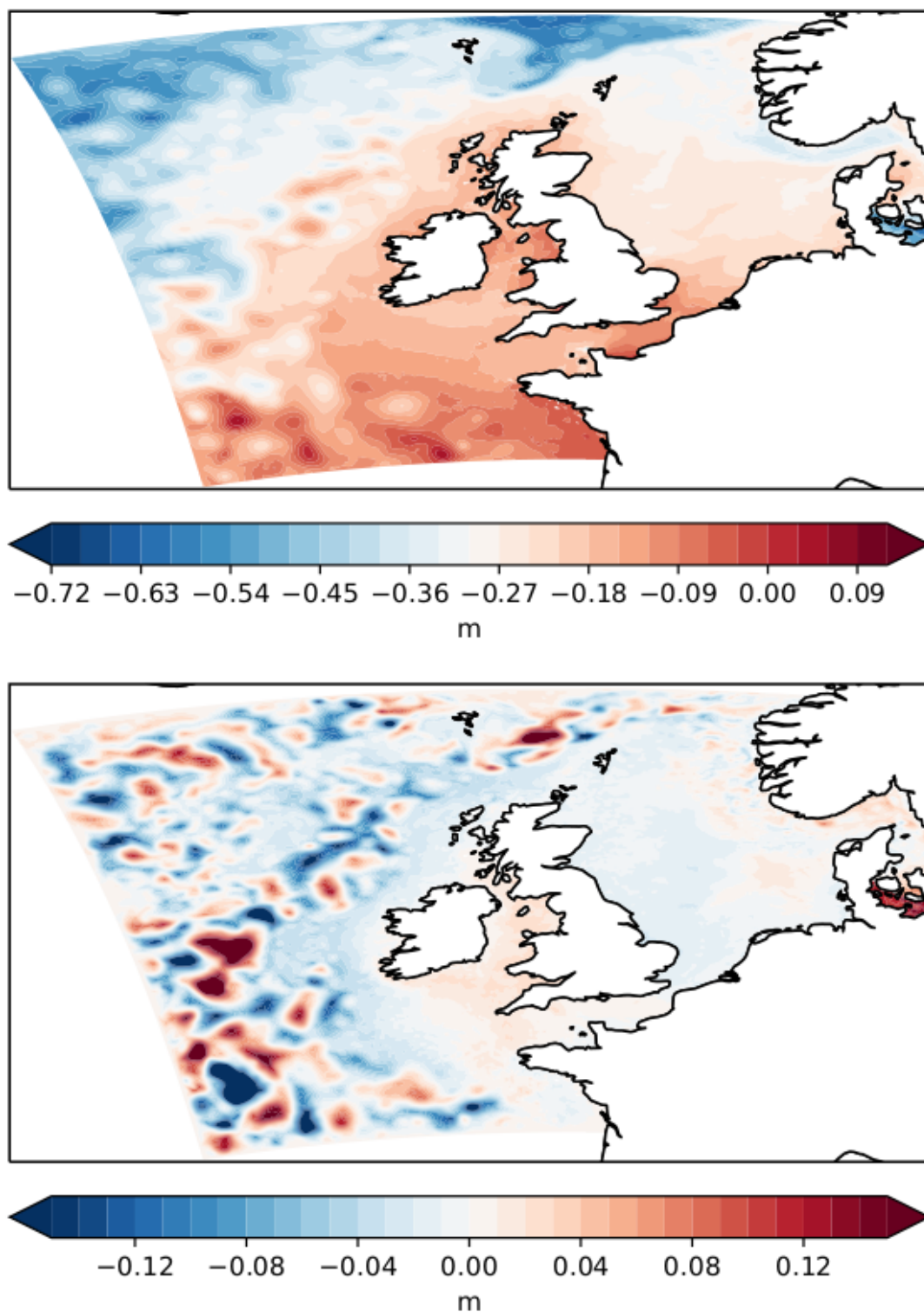
Figure 2. Nature Run 25-hour mean SSH for an example day (30 June 2018, top) and the difference (bias) in the same field from the Free Run.
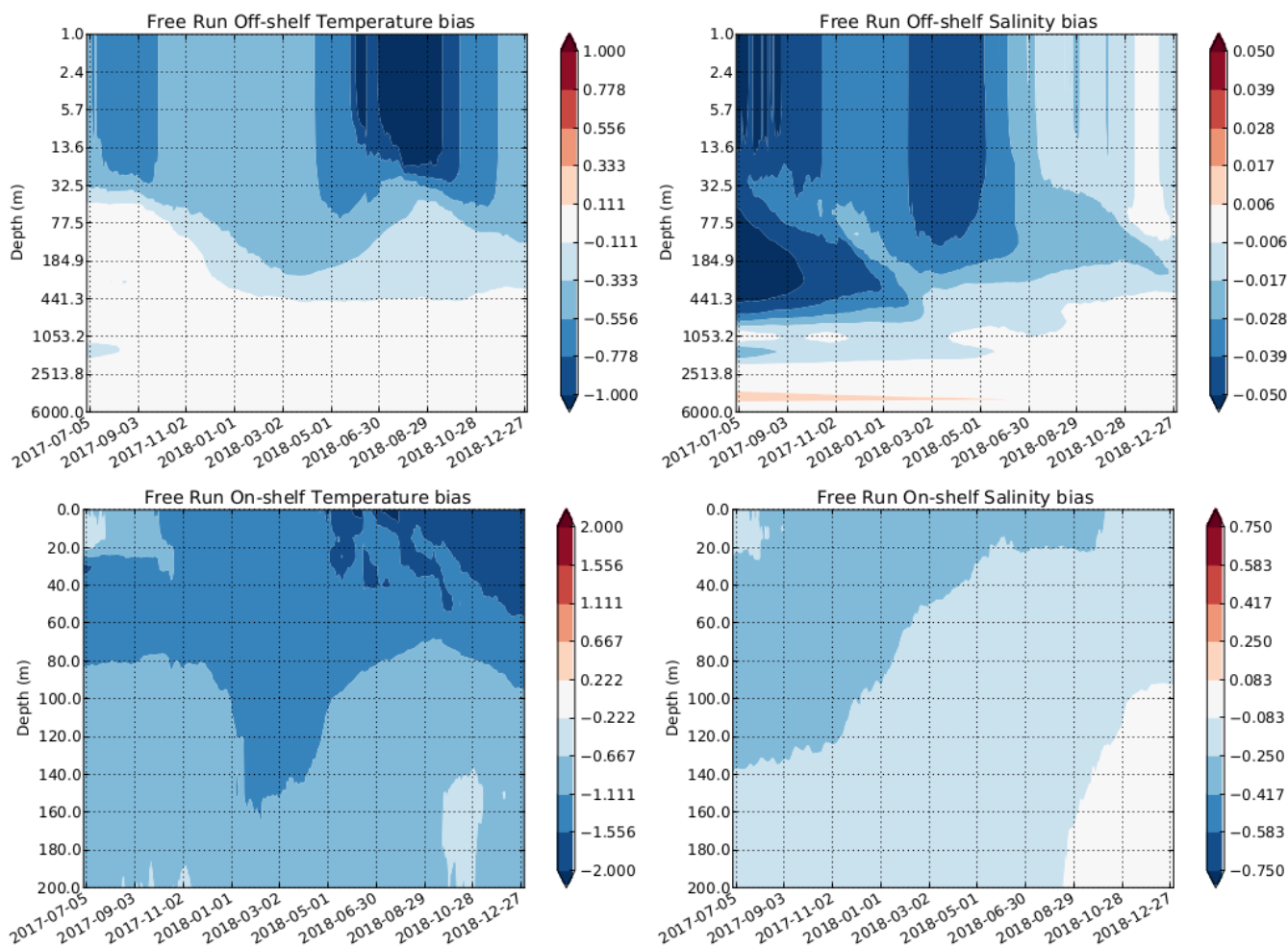
Figure 3. Free Run temperature (left) and salinity (right) biases relative to the Nature Run for the off-shelf (top) and on-shelf (bottom) regions. Note that the off-shelf plots are shown on a log depth scale.

(limited to a minimum of 3 times the grid scale to avoid numerical issues). Each background error scale component has an associated set of spatially and seasonally varying error variances.

190   An assimilation window of 24 hours is used to assimilate observations including satellite and in situ SST, satellite altimeter observations of the sea level anomaly (SLA), and in situ temperature and salinity profiles from a variety of platforms. The operational system includes a bias correction scheme for SST (While and Martin, 2019) and SLA (Lea et al., 2008) observations, but this is not used here as the observations are simulated directly from the Nature Run. Although we have aimed to add realistic errors to the observations, we have not included any deliberate biases. The assimilation increments are added

195   incrementally over 24 hours through an analysis update cycle (Bloom et al., 1996).

NEMOVAR uses linearised balance relationships to account for correlations between ocean variables as defined by Weaver et al. (2005). The linearised balance relationships between increments to temperature (T), salinity (S), sea surface height (η), and the u- and v-components of velocity (u,v) can be summarised as follows:

$$\delta T = \delta T,$$
$$\delta S = K_{S,T}\delta T + \delta S_U,$$
$$\delta \eta = K_{\eta,\rho}\delta \rho + \delta \eta_U,$$
$$\delta u = K_{u,p}\delta p + \delta u_U,$$
$$\delta v = K_{v,p}\delta p + \delta v_U. \tag{1}$$

200   where the subscript U denotes the unbalanced component of the variable. The operators K_{ij} are linearised balance transforms from variable j to i. The density (ρ) and pressure (p) increments are then defined as:

$$\delta \rho = K_{\rho,T}\delta T + K_{\rho,S}\delta S,$$
$$\delta p = K_{p,\rho}\delta \rho + K_{p,\eta}\delta \eta. \tag{2}$$

The temperature-salinity balance operator ($K_{S,T}$) is a linearised version of the water-mass property conservation scheme of Troccoli and Haines (1999). The SSH balance operator ($K_{\eta,\rho}$) is a linearised version of the Cooper and Haines (1996) scheme

205   in which SSH changes are associated with the lifting and lowering of the entire water column. The velocity balance operators ($K_{uv,p}$) are based on the geostrophic relation, density balance is the linearised equation of state, and the pressure balance is the vertically-integrated hydrostatic equation.

The real SLA observations assimilated in the operational Met Office systems (obtained from CMEMS) are provided with various corrections to remove the effect of tides, atmospheric pressure and winds. These are used in our global model which

210   does not include tides or atmospheric pressure effects. However, our shelf-seas model does include these processes and so the corrections supplied with the SLA observations are not used. However for this work, we have adapted the system to remove tides from the SLA innovations before assimilation which was found in testing to result in slightly reduced errors compared to

including the tides in the innovations. This is done by using a 24hour 50minute mean field for the background SSH model field
and by using the 25-hour mean fields in the simulation of the SLA observations (as described below in section 4.2).

215   We have also removed the previous restrictions on SLA assimilation at high latitudes and in shallow water (depth<700 m)
as detailed in King et al. (2018) so that altimeter observations are assimilated throughout the domain. The SSH is adjusted by
applying barotropic (unbalanced) increments to the sea surface height and a baroclinic (balanced) component which acts on
the density structure. To avoid making unphysical corrections to the sub-surface temperature and salinity, the full SSH balance
(as detailed in equations 1 and 2) is only applied in regions where the top-to-bottom temperature difference exceeds 2°C with
220   no balance applied where the difference is less than 1°C and with a linear transition between these thresholds.

In the experiments reported here, we have used the known observation error variances for the simulated SWOT observations,
but have not yet introduced any specific adaptions for the assimilation of wide-swath altimeter observations. In the near future,
we will investigate the impact of a number of approaches. Initially, we may test the impact of inflating observation errors
for wide-swath altimetry observations to avoid over-fitting the data. However, it will be necessary to investigate methods to
225   account for the correlations directly in the assimilation scheme.

4   Simulation of Observations

4.1   SST and T/S profiles

To simulate the in situ and level-2 satellite SST observations and in situ T/S profiles assimilated in the operational NWS
system, we have used the positions and times of the real observations for the period of the experiments. These include T/S
230   profiles from Argo, XBTs, CTDs, marine mammals, ships and moorings taken from the EN4 database (Good et al., 2013);
in situ SST observations from ships, moorings and surface drifters; and satellite SST observations from NOAA-AVHRR,
METOP-AVHRR, MSG-SEVIRI, Sentinal3-SLSTR, AMSR-2, and VIIRS. This allowed us to define the position and time
of the pseudo-observations to ensure a realistic spatial and temporal coverage. An example of the spatial coverage of daily
SST and monthly T/S profile observations is shown in Figure 4. This demonstrates the excellent temporal and spatial sampling
235   of SST observations in contrast to the sparse nature of the available sub-surface observations. Note that the unusually dense
coverage of T/S profiles in the English Channel during this month was due to an Ifremer research cruise.

These observations were passed through our observation operator to extract the model equivalent value from the Nature
Run at the correct position and at the nearest model time-step to the observation time. To make these perfect 'observations'
of the Nature Run's truth realistic, measurement and representation error must be added. For both SST and T/S profile ob-
240   servations, we added random Gaussian noise to each observation with platform-specific errors. SST measurement errors
are defined according to the errors supplied with the observations with values from 0.3–0.6 K depending on the platform.
T/S measurement error standard deviations are specified as follows: XBTs 0.15 K; Argo 0.002 K, 0.01 PSU; CTDs 0.002 K,
0.005 PSU; moored buoys 0.05 K, 0.02 PSU; and other types/unknown 0.1 K, 0.05 PSU (taken from Ingleby and Huddleston
(2007), https://www.pmel.noaa.gov/gtmba/sensor-specifications, and from http://www.argo.ucsd.edu/FAQ.html). In addition to
245   random measurement errors, to include some simulated representation error in the observations we have added a perturbation
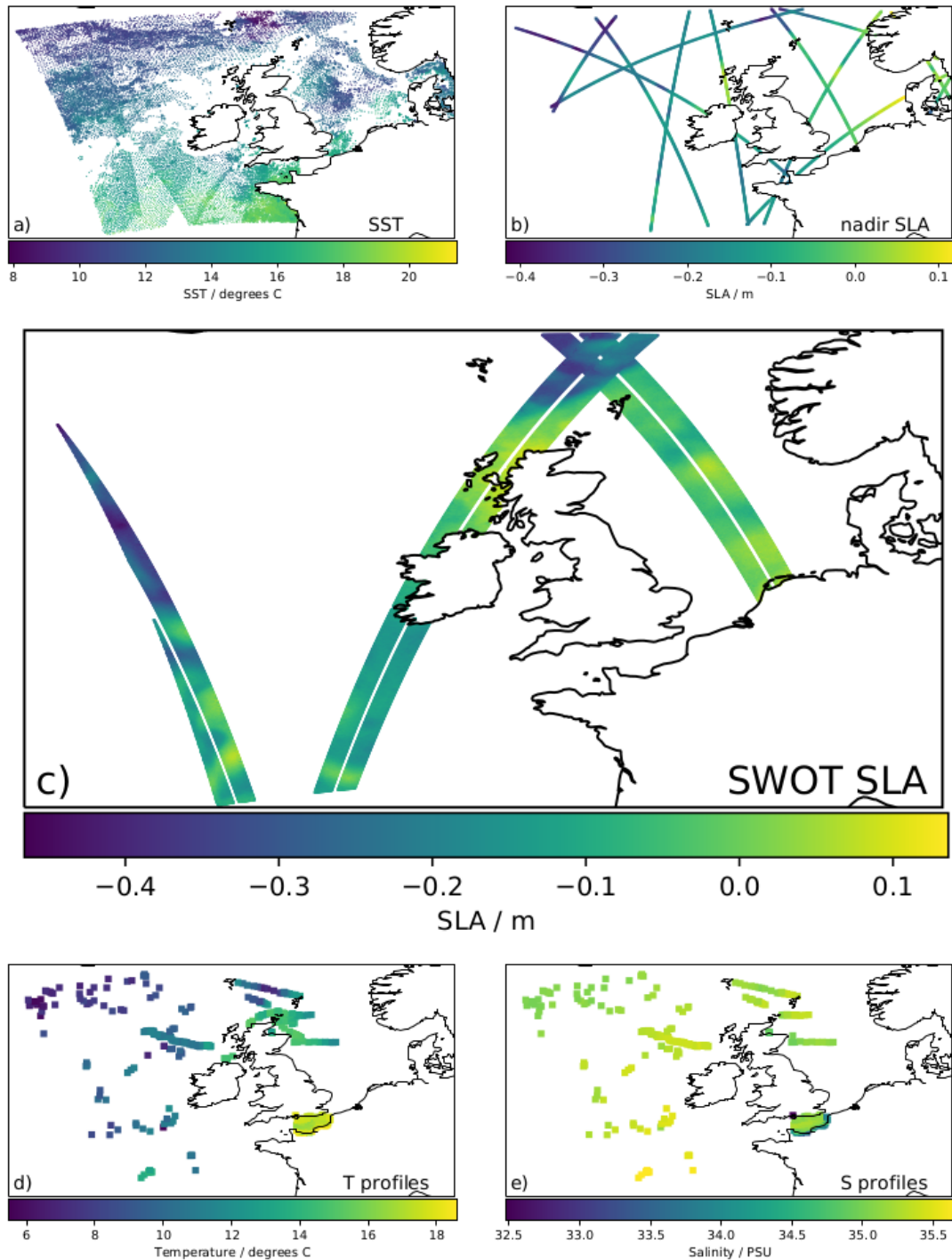
Figure 4. Example observation coverage for 1-day (31 Oct 2017) for a) SST, b) nadir SLA, and c) SWOT SLA observations, and 1-month (Oct 2017) for d) temperature and e) salinity profile observations.

to the location of each observation (as sampled from the Nature Run gridded fields). This is done by perturbing the real latitude
and longitude of each observation by a random amount sampled from a Gaussian distribution with standard deviation of 0.03°.

Table 2. SWOT swath and all nadir error statistics (RMSE and extrema) for 1-month (July 2017) of simulated observations. Each component
of the simulated errors is given for our experimental set-ups with the full swath (SWOT), half-swath (HalfSWOT), and the uncorrelated-error
experiment (LowErrSWOT).

Error (cm)	RMSE	Extr.	RMSE	Extr.	RMSE	Extr.
			SWOT swath			
	SWOT		HalfSWOT		LowErrSWOT	
All	6.2	39	4.2	24	1.2	7
Phase	4.9	26	3.0	15	-	-
Roll	3.1	16	1.9	9	-	-
Timing	1.8	6	1.8	6	-	-
KaRIn	1.2	7	1.2	7	1.2	7
Baseline Dilation	0.6	4	0.2	1	-	-
Residual Path Delay	0.5	3	0.5	2	-	-
			All nadir			
All	1.4	6	1.4	6	1.4	6

4.2 Altimeter observations

To simulate the nadir and swath altimeter measurements of SLA, we have used the SWOTsimulator of Gaultier et al. (2016).

250 This tool uses user-supplied model SSH fields and produces SWOT-like swath and nadir observations with random realisations
of instrumental and geophysical errors. The SWOTsimulator was also used to simulate along-track nadir observations from
Jason-3, Sentinal-3A, Sentinal-3B, and AltiKa to emulate the existing altimeter constellation: each of these observations has
the same along-track resolution and spatial and temporal sampling as the real observations. An example of the spatial coverage
of daily SLA from the existing nadir altimeters and from SWOT is shown in Figure 4. This demonstrates the 2D nature of

255 SWOT SLA observations in comparison to the current network of nadir altimeters (note that the point-size of the nadir SLA
observations is larger than the actual cross-track resolution to allow the observations to be seen on this scale), but also highlights
the significant gaps between tracks on daily time-scales caused by the long repeat cycle of the SWOT satellite (21 days). Future
constellations of swath altimeters could address this mismatch in the spatial and temporal sampling.

To produce the simulated observations, we supplied the SWOTsimulator with 25-hour mean fields from the Nature Run to

260 approximately remove the effect of tides. SWOT observations are expected to be provided on a 1–2 km grid (Morrow et al.,
2018), so the default SWOTsimulator settings were used to produce the simulated SWOT observations on a 2x2 km grid within
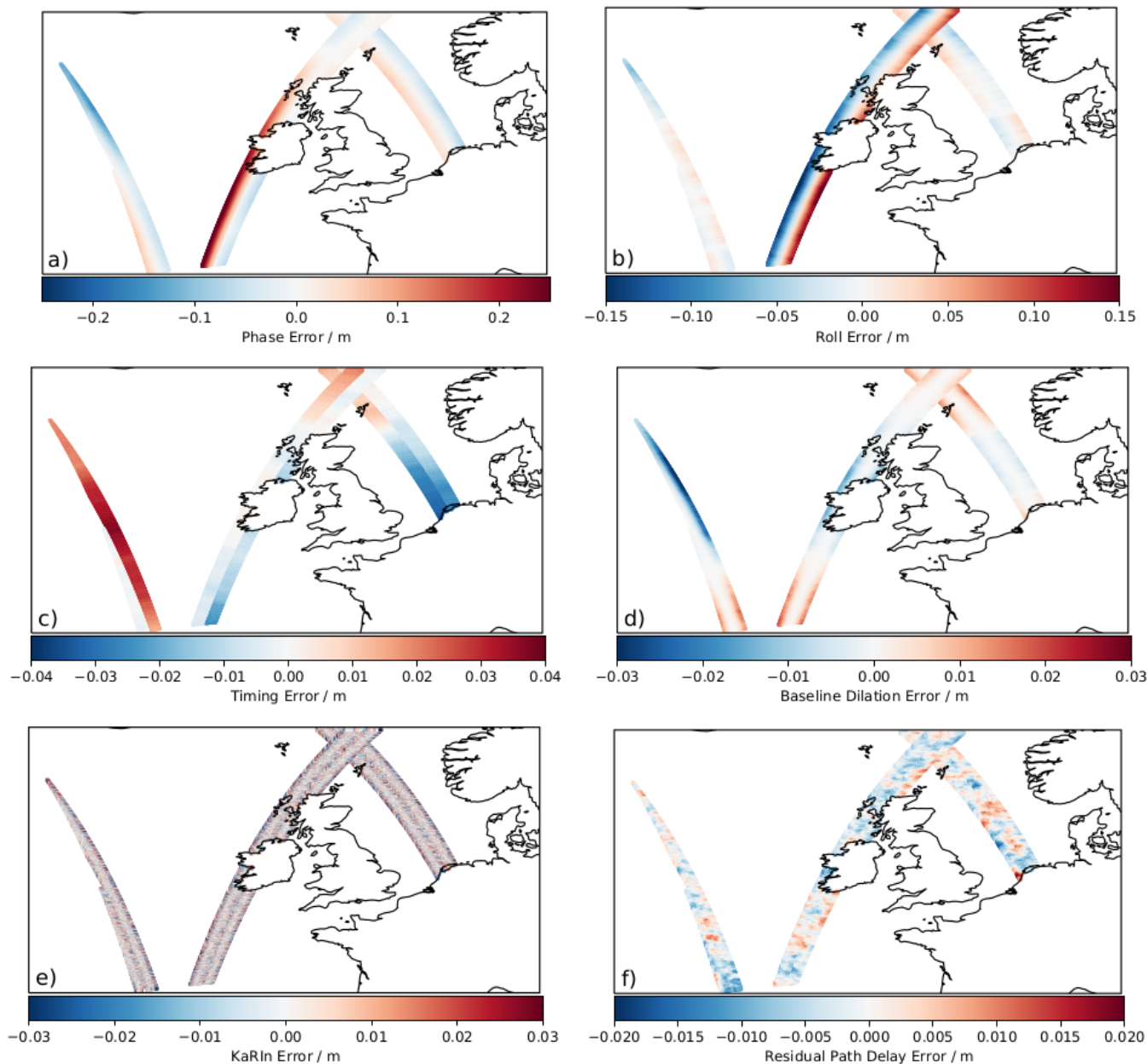
Figure 5. Individual components of the SWOT errors for an example day (31 Oct 2017): a) phase error, b) roll error, c) timing error, d) baseline dilation error, e) KaRIn error and f) residual path delay error. Note the difference in the scales for each error component.

the swath. Since we are using the same model for the Nature Run and OSSEs, the mean dynamic topography (MDT) is known perfectly which introduces a necessary idealisation in our experiments. While the real MDT is expected to be reasonably well specified off-shelf, the sparseness of observations limits its accuracy on-shelf.

Table 3. SSH, surface current, and total column temperature and salinity RMSE for the on- and off-shelf regions from each experiment. The percentage change in RMSE compared to the Control Run is shown in parentheses.

	SSH (m)		Temperature (K)		Salinity (PSU)		Surface Current Speed (m/s)	
Off-shelf								
Free Run	0.061	(+79%)	0.510	(+57%)	0.069	(+30%)	0.120	(+15%)
Control	0.034		0.324		0.053		0.104	
HalfSWOT	0.037	(+9%)	0.352	(+9%)	0.052	(-2%)	0.111	(+7%)
SWOT	0.041	(+21%)	0.391	(+21%)	0.054	(+2%)	0.114	(+10%)
LowErrSWOT	0.027	(-21%)	0.340	(+5%)	0.051	(-4%)	0.091	(-13%)
HalfSWOT_5km	0.032	(-6%)	0.324	(0%)	0.052	(-2%)	0.104	(0%)
HalfSWOT_20km	0.036	(+6%)	0.328	(+1%)	0.053	(0%)	0.106	(+2%)
SWOT_5km	0.036	(+6%)	0.334	(+3%)	0.052	(-2%)	0.106	(+2%)
SWOT_20km	0.037	(+9%)	0.326	(+1%)	0.053	(0%)	0.107	(+3%)
On-shelf								
Free Run	0.021	(+24%)	0.612	(+90%)	0.137	(+20%)	0.035	(+3%)
Control	0.017		0.322		0.114		0.034	
HalfSWOT	0.018	(+6%)	0.328	(+2%)	0.108	(-5%)	0.036	(+6%)
SWOT	0.021	(+24%)	0.366	(+14%)	0.115	(+1%)	0.039	(+15%)
LowErrSWOT	0.017	(0%)	0.314	(-2%)	0.108	(-5%)	0.034	(0%)
HalfSWOT_5km	0.017	(0%)	0.325	(+1%)	0.111	(-3%)	0.035	(+3%)
HalfSWOT_20km	0.017	(0%)	0.324	(+1%)	0.113	(-1%)	0.035	(+3%)
SWOT_5km	0.018	(+6%)	0.333	(+3%)	0.110	(-4%)	0.036	(+6%)
SWOT_20km	0.017	(0%)	0.325	(+1%)	0.112	(-2%)	0.035	(+3%)

265 The wide-swath altimeter observations are subject to a number of sources of error, including phase error, roll error, timing error, KaRIn noise, baseline dilation error, and residual path delay error (see Gaultier et al., 2016). These errors are significantly larger than that associated with current nadir SLA observations. Each of the noise components is shown in Figure 5 for an example day and the standard deviation and extreme values from 1-month of simulated observations are summarised in Table 2. The phase and roll errors in particular can introduce spatially correlated errors in excess of 10 cm along or across a swath. The

270 length-scale of these correlations can also be of the same order as the size of the domain.

To test the limitations imposed by these large correlated errors on the assimilation of SWOT observations, we have run the OSSEs detailed in Table 1 and so have created simulated SWOT observations for each experiment: the full swath with all error components (for the SWOT runs); the inner half of the swath with all error components (for the HalfSWOT runs); and the full swath with only the uncorrelated KaRIn errors (for the LowErrSWOT run). The effect of these limitations on the

275 total observation error (and on each component of the error) is shown in Table 2. Reducing the width of the swath by half results in a drop of >30% in the RMSE (root mean square error) and removes the most extreme values. The LowErrSWOT observations include only the KaRIn errors and have an RMSE closer to that of the current nadir altimeter observations. In all of the experiments the SLA observations have been quality-controlled before assimilation. The observation-model background differences and the observation and background errors are used to assess the probability of gross error. Those observations with

280 a probability of gross error greater than 50% are not assimilated. This however only removes ~0.3% of the SLA observations in the SWOT experiment. More stringent QC could remove more of the observations which are affected by large correlated errors, but at the expense of valid observations where the model most needs to be constrained.

5 Results

In this section we will detail the impact of assimilating the simulated SWOT observations by comparing each experiment with

285 the truth provided by the full 4D fields from the Nature Run. Since the SSH signal is dominated by barotropic processes in shallow water and baroclinic processes in deeper water, we present these comparisons separately for the on-shelf and off-shelf regions as marked in Figure 1. The impact of each experiment on the SSH, surface current speed, temperature and salinity RMSE is summarised in Table 3.

5.1 Impact on SSH

290 The impact of assimilation is clearly positive in the off-shelf SSH RMSE in all experiments, each of which reach a stable RMSE within the first month of the runs (see Figure 6), but with considerable seasonal variation. However, the Control is superior to the HalfSWOT and SWOT experiments which assimilate the SWOT observations with the full correlated errors. The HalfSWOT and SWOT experiments increase the off-shelf SSH RMSE (calculated over the full run excluding the first month, see Table 3) by 9% and 21% respectively relative to the Control, while the LowErrSWOT experiment which did not

295 include the large correlated errors shows a reduction of 21% in the RMSE relative to the Control. The impact on the off-shelf bias is less clear. When the bias in the mixed layer depth increases into winter the Free Run SSH bias increases. However, all assimilative runs show similar slightly negative biases, suggesting that the bias is relatively well-constrained by profile assimilation.

The modelled SSH variability on-shelf is lower than in the off-shelf regions. However, as with the off-shelf region, the

300 Free Run shows a drift in the on-shelf SSH bias, which is well-constrained in all of the assimilative runs. In absolute terms, there is little difference in the on-shelf SSH RMSE between the assimilative experiments. However, in percentage terms the impact in the HalfSWOT and SWOT experiments is similar to that seen off-shelf, with an increase in the on-shelf SSH RMSE
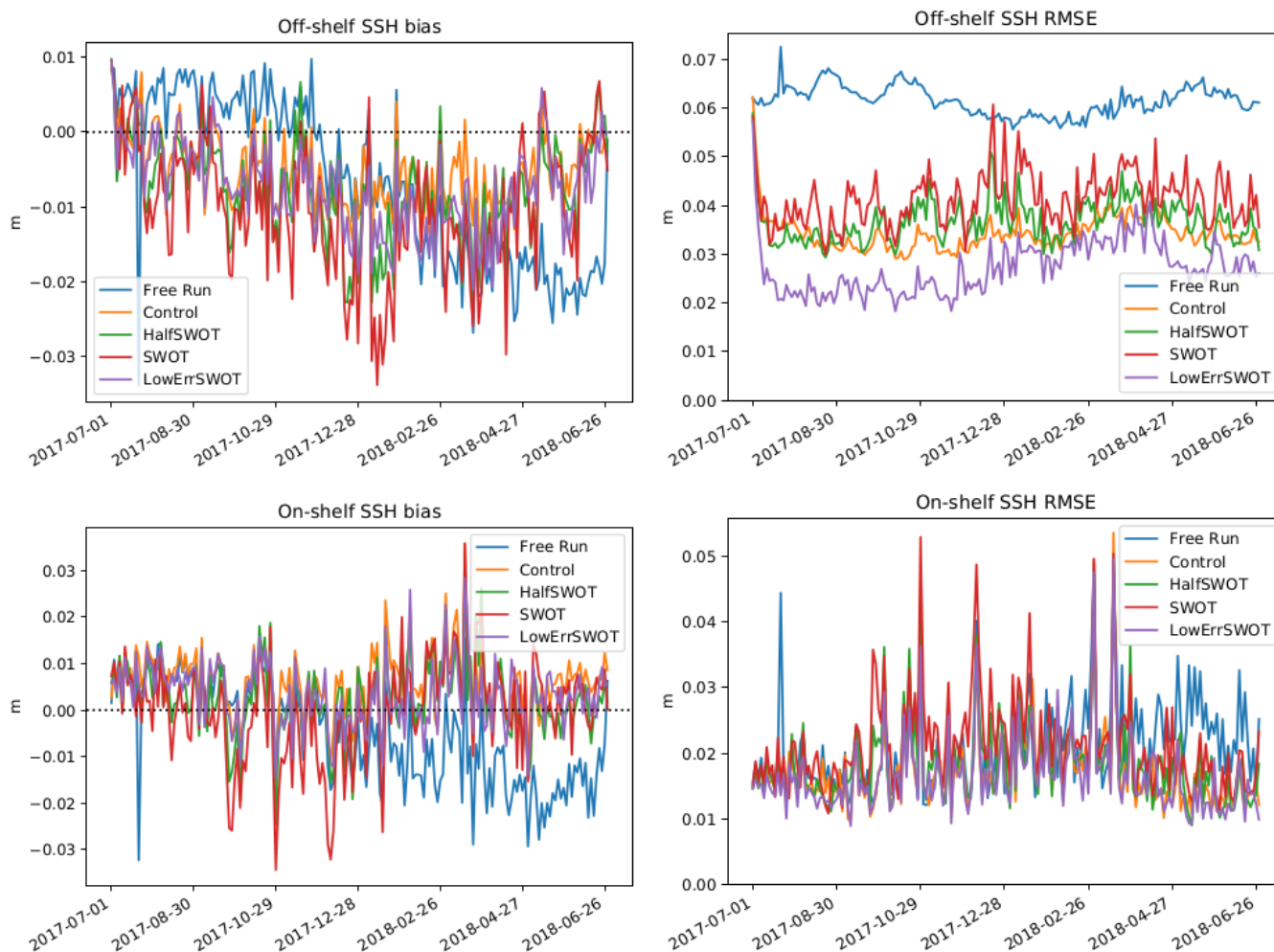
Figure 6. SSH bias (m, left) and RMSE (right) for the off-shelf (top) and on-shelf (bottom) regions.

(calculated over the full run excluding the first month) of 6% and 24% respectively. However, the LowErrSWOT experiment shows a negligible difference in the RMSE relative to the Control.

305   One notable feature is the frequent large spikes in the on-shelf SSH RMSE, which are present in the Free Run and the assimilative runs. These appear relatively unaffected by the assimilation of the standard network of observations, or by the additional SWOT observations. These are likely due to errors in the surface forcing (i.e., differences between the ECMWF IFS and MetUM surface fluxes) which can have a large impact on the SSH during surge events. This reflects similar RMSE spikes seen in our operational system and is an issue we aim to address in the near future. However, the experimental set-up

310   here where the SLA observations were simulated using 25-hour model mean fields is sub-optimal for constraining these short period events.
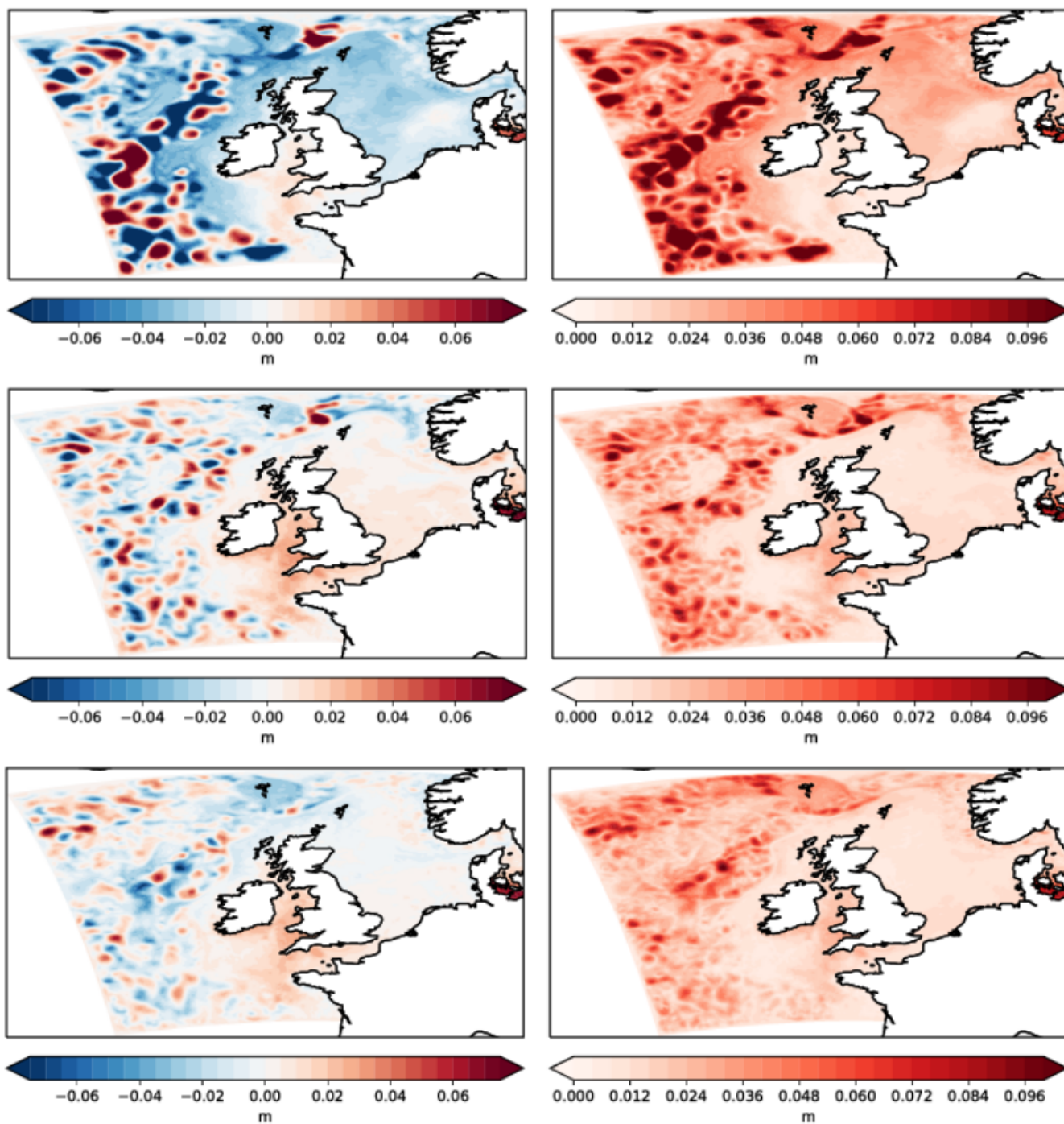
Figure 7. June 2018 average SSH bias (left) and RMSE (right) for Free Run (top), Control (middle), LowErrSWOT experiments (bottom).
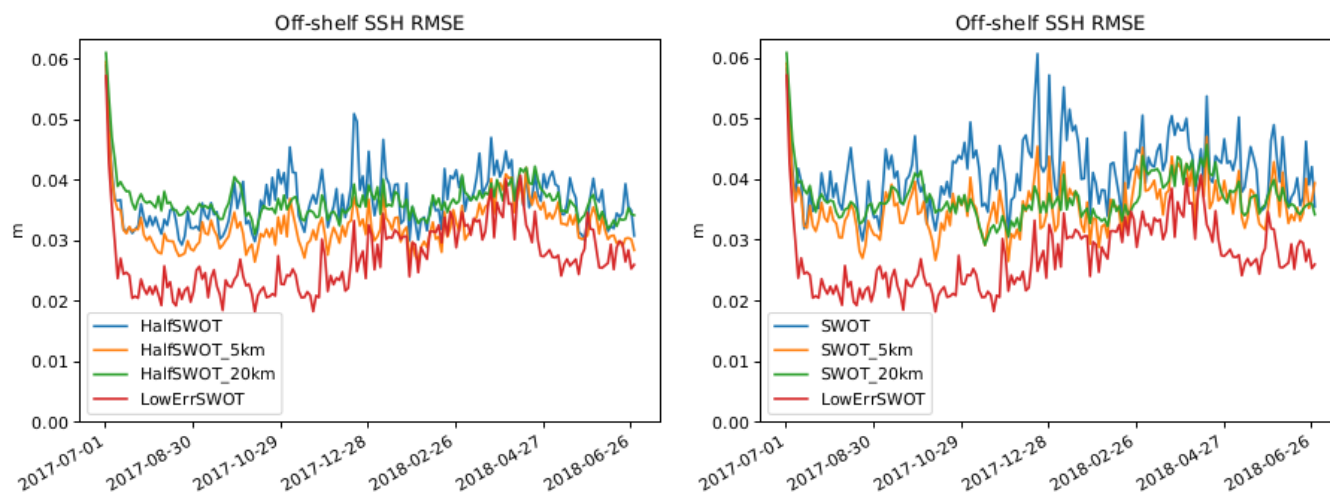
Figure 8. Off-shelf SSH RMSE for the HalfSWOT (left) and SWOT (right) experiments with the corresponding experiments where SWOT observations have been averaged over 5 km and 20 km radii. The LowErrSWOT experiment is also show in both cases.

Spatial maps of the mean bias and RMSE over the final month of the experiments (June 2018) further demonstrate the improvement both on- and off-shelf when assimilating SWOT observations without correlated errors. Figure 7 shows the progression of the bias and RMSE from the Free Run to the Control and then to the LowErrSWOT experiment. The large SSH
315 bias across the domain and the increased RMSE off-shelf seen in the Free Run are clearly improved when assimilating the standard network of observations, with a significant further improvement when assimilating SWOT observations without correlated errors. Although there is a change in the sign of the bias in the Irish and Celtic Seas from the Free Run to the Control, the assimilation of uncorrelated SWOT observations improves the bias and RMSE.

5.1.1 Impact of Observation Averaging

320 Assimilating SWOT observations without correlated errors can clearly have a large positive impact. By limiting the SWOT data to the inner half of the swath, we have shown that some of the degradation caused by correlated errors can be reduced. We next considered whether observation-averaging within the SWOT swath could further reduce the problems associated with assimilating SWOT observations with correlated errors. Averaging radii of 5 km and 20 km were applied to SWOT data from both the full and half swath widths. We chose these radii to test a minimal level of averaging (5 km) and a larger radius similar
325 to that used in other studies (20 km, e.g., Carrier et al., 2016). Although the length-scales of the correlations apparent in the simulated SWOT observations can be considerably longer than 20 km (see Fig. 5), we chose not to test larger averaging radii as this would severely degrade the unique ability of SWOT observations to resolve the surface topography of the ocean.

The impact of observation averaging on the SSH RMSE is shown in Figure 8. Both levels of observation averaging improve the results compared to no averaging, for the full and half swath width cases. In the experiments assimilating the full swath
330 width, it is clear that the lower the level of SWOT observation averaging the greater is the day-to-day variability in the SSH
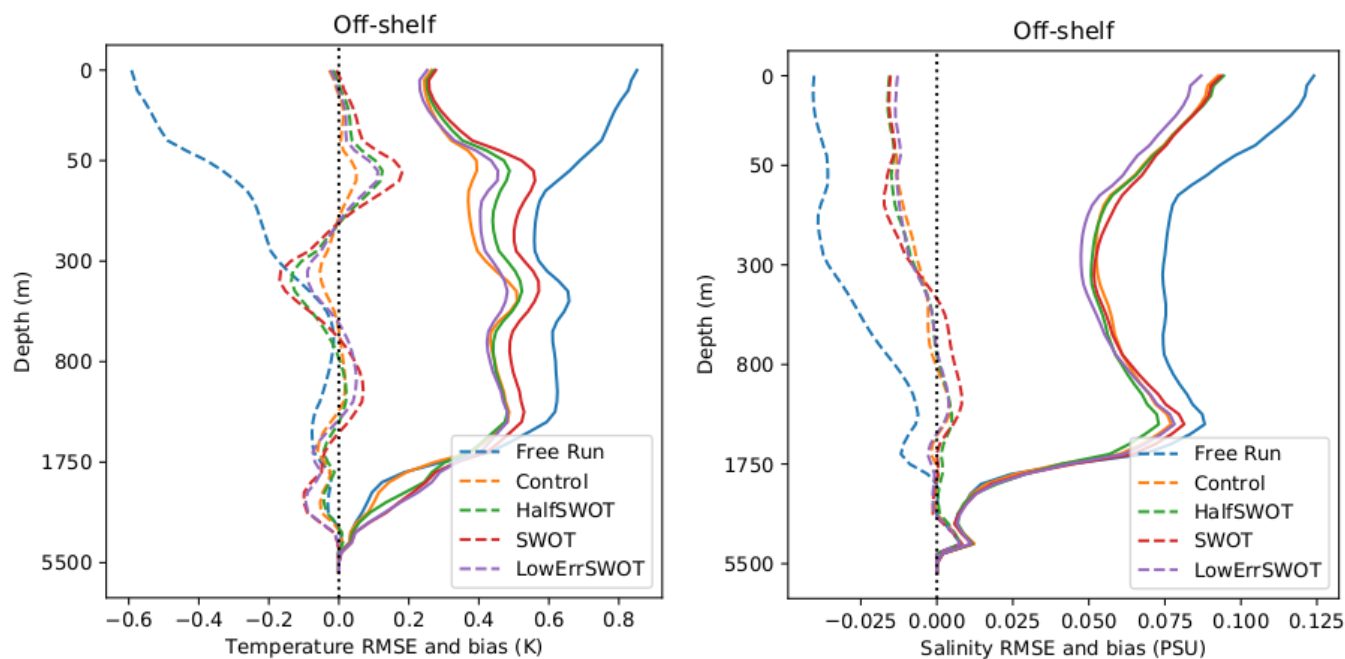
Figure 9. Off-shelf temperature and salinity RMSE (solid) and bias (dashed) averaged over the full year of the experiments. Note the log-scale used for depth to highlight the differences in the upper ocean.

RMSE. This is presumed to be caused by the assimilation of observations with the largest errors. Overall the best results when assimilating SWOT observations with correlated errors are found when using half the SWOT swath and 5 km observation averaging (HalfSWOT_5km).

5.2 Impact on Sub-surface Temperature and Salinity

335 ### 5.2.1 Off-shelf T/S

In the off-shelf region, the near-surface temperature bias found in the Free Run (see Fig. 3) is very quickly removed and the RMSE significantly reduced in all assimilative runs. This is not surprising given the large quantity of SST observations assimilated in all OSSE runs. However, the average off-shelf temperature bias and RMSE for all experiments over the full year (shown in Fig. 9) shows a marginal additional reduction in the near-surface temperature RMSE in the LowErrSWOT

340 experiment.

The assimilation of the standard network of observations appears to introduce a sub-surface bias around 30–80 m depth in the Control run, which is exacerbated in the SWOT, HalfSWOT and LowErrSWOT experiments. This effect, due to the method used to spread SST information in the vertical, was noted in King et al. (2018) and methods to avoid or reduce it are currently being investigated. It is known that profile observations can constrain this increase in the thermocline bias and RMSE, but

345 additional altimeter observations appear to exacerbate the issue.
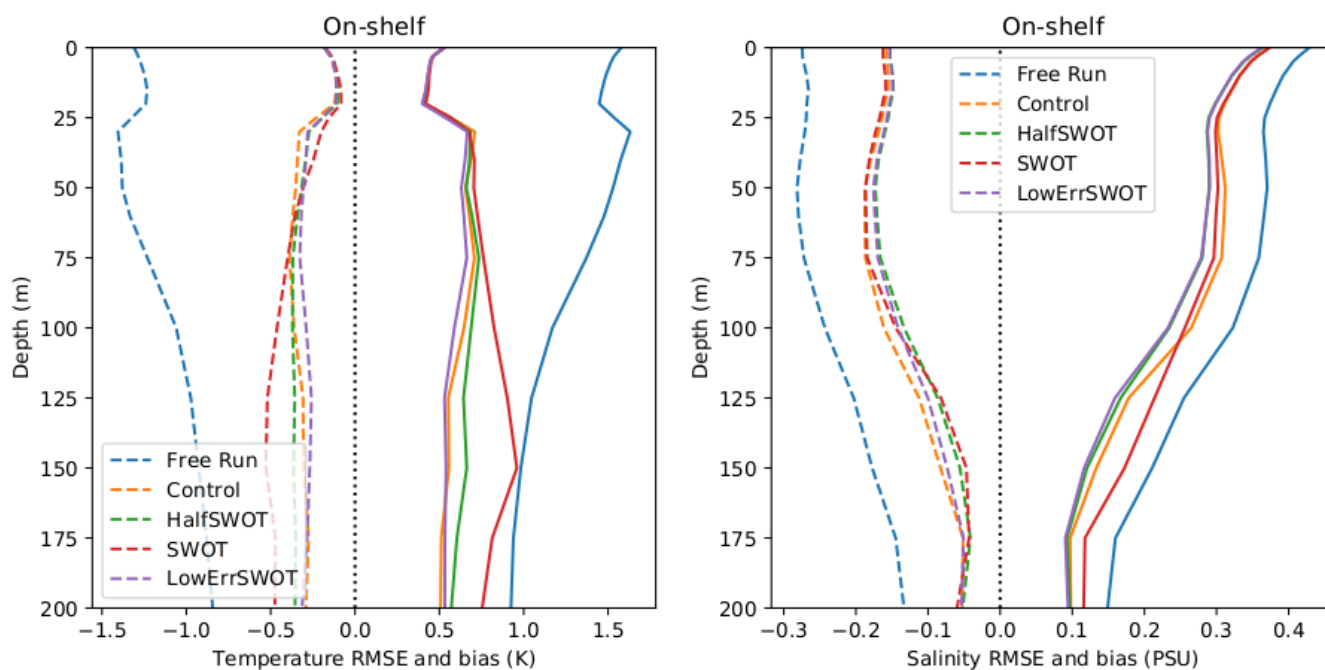
Figure 10. On-shelf temperature and salinity RMSE (solid) and bias (dashed) averaged over the full year of the experiments.

Below $\sim 500\,\mathrm{m}$, the LowErrSWOT experiment again shows a marginal improvement in RMSE compared to the Control, while the HalfSWOT experiment shows little change and the SWOT experiment a degradation. Overall, apart from the known issues near the thermocline, the additional SWOT observations can reduce the off-shelf temperature errors, but the full correlated SWOT errors lead to a degradation at almost all depths.

350   Off-shelf salinity biases are also rapidly reduced by assimilation in the Control run, with a quicker decrease in bias seen in the LowErrSWOT experiment (not shown). Near-surface salinity RMSE is also reduced in the Control relative to the Free Run, but not to the extent as for temperature due to the relative sparseness of the in situ salinity observations. The salinity is not affected by the known issue discussed above where SST assimilation leads to inflated biases near the thermocline.

The average off-shelf salinity bias and RMSE over the full year of the experiments shown in Figure 9 demonstrates that all
355   assimilative runs have improved salinity errors compared to the Free Run. Overall, the additional SWOT observations in the LowErrSWOT experiment decrease the salinity biases and RMSE at all depths, but assimilating SWOT observations with the full correlated errors leads to a slight degradation compared to the Control, especially when assimilating the full swath.

5.2.2 On-shelf T/S

As for the off-shelf region, the near-surface on-shelf temperature bias and RMSE is constrained by the assimilation of the
360   available SST observations (see Figure 10). In the upper 25 m there is almost no difference in the on-shelf temperature RMSE or bias between all of the assimilative runs. The average on-shelf temperature bias and RMSE over the full year of the exper-

iments shown in Figure 10 show an improvement in the temperature errors at all depths below ∼25 m for the LowErrSWOT experiment. On the other hand, the SWOT experiment with correlated errors shows significantly larger RMSE below ∼75 m, though this RMSE rise is limited when assimilating only the inner half of the swath in the HalfSWOT experiment.

365   In contrast, the on-shelf salinity bias and RMSE is marginally improved by the assimilation of SWOT observations in the LowErrSWOT and HalfSWOT experiments (see Figure 10). Only below ∼100 m is there a degradation (compared to the Control) to the salinity RMSE in the SWOT experiment with the full correlated errors and full swath width.

In summary, on-shelf temperature and salinity errors are improved by the assimilation of SWOT observations without correlated errors. Restricting the SWOT data with the full correlated errors to the inner half of the swath limits the negative impacts.

370   However, when the full swath width SWOT observations contain realistic correlated errors, there is a degradation in both temperature and salinity at almost all depths.

5.2.3   Impact of Observation Averaging

In general, observation averaging was found to improve the off-shelf temperature bias and RMSE in experiments with both the full and half-swath width (compared to no observation averaging, see Table 3). Near the thermocline, where there is a known

375   issue caused by the SST assimilation (discussed above), the larger averaging radius leads to a lower bias and RMSE due to the lower quantity of observations assimilated (as noted above, the effect is exacerbated by additional observations). However, elsewhere the 5 km averaging radius is marginally superior. The on-shelf temperature bias and RMSE when assimilating SWOT with correlated errors are also improved by observation averaging, but in this case the 20 km radius is superior.

For salinity, there is a smaller overall impact from observation averaging. Off-shelf, while the 5 km averaging is marginally

380   better than no averaging, using the 20 km averaging radius gives mixed results with small degradations at some depths. On-shelf, when using the full swath observation averaging improves the RMSE and bias, with the 5 km radius giving best results, but when using the half-swath observation averaging degrades the results, with the larger radius showing the largest degradation. Overall, a 5 km observation averaging radius appears to give most improvements to the temperature and salinity statistics across the domain with the fewest and smallest degradations.

385   ### 5.3   Impact on surface currents

A comparison of daily and monthly mean surface current speeds (not shown) highlights that the major surface current features of the Nature Run are also present in the Free Run. However, with no observations to constrain the surface currents many of the eddies and current meanders are misplaced. The Control run assimilating the standard network of observations improves the positions of individual eddies and straightens the current through the the Faroe-Shetland channel. With the addition of

390   the uncorrelated SWOT observations, the LowErrSWOT experiment further improves the simulation of the surface current features.

This improvement in the positioning of surface current features is clearly seen in the maps of monthly surface current errors shown in Figure 11. In particular, there is a clear additional benefit of assimilating SWOT observations with uncorrelated errors. However, when assimilating SWOT observations with correlated errors, the daily difference fields show correlated
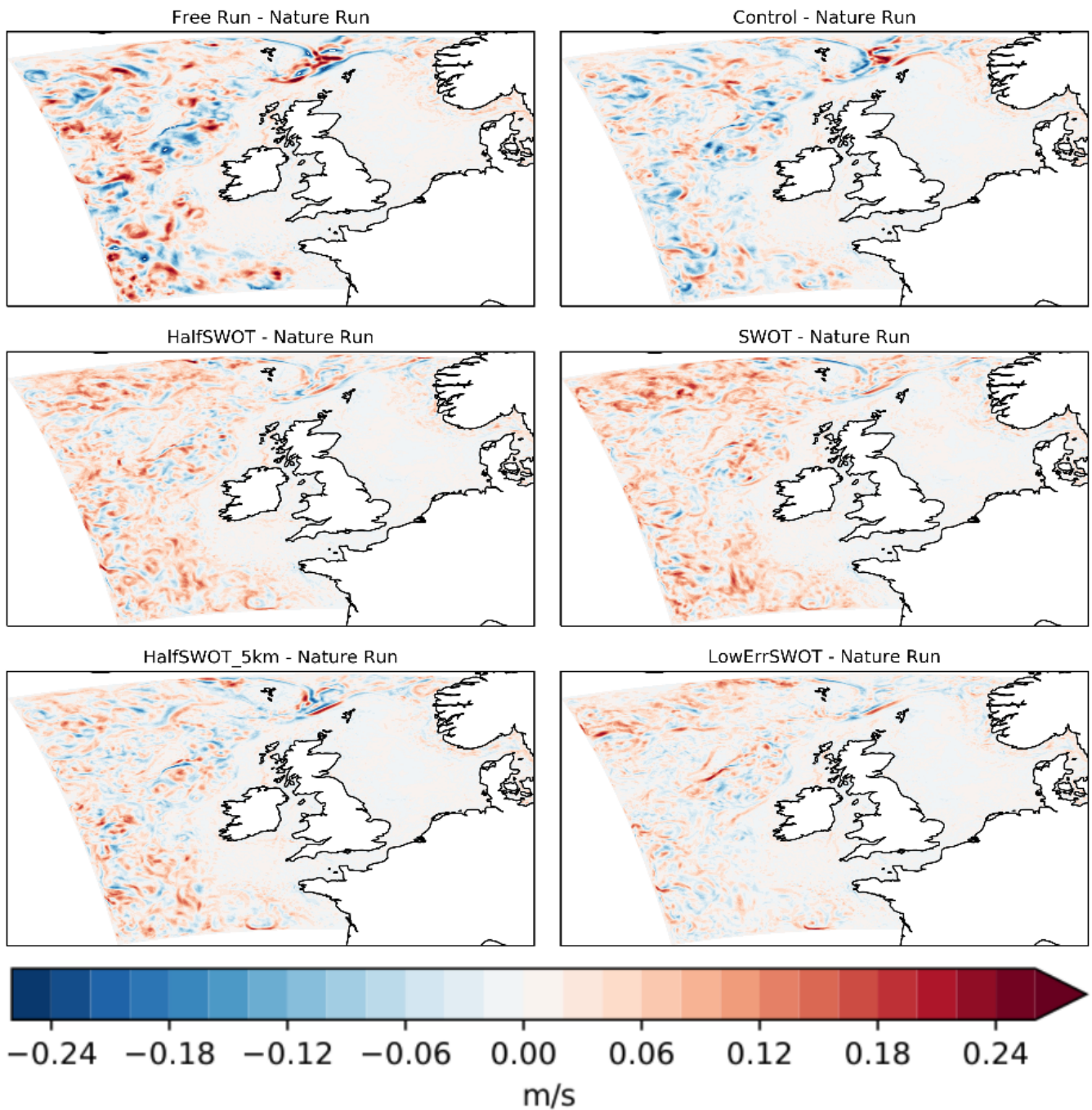
Figure 11. Monthly (June 2018) mean surface current speed error (compared to Nature Run) from the Free Run (top left), Control, HalfS-WOT, HalfSWOT_5km, SWOT and LowErrSWOT experiments (bottom).
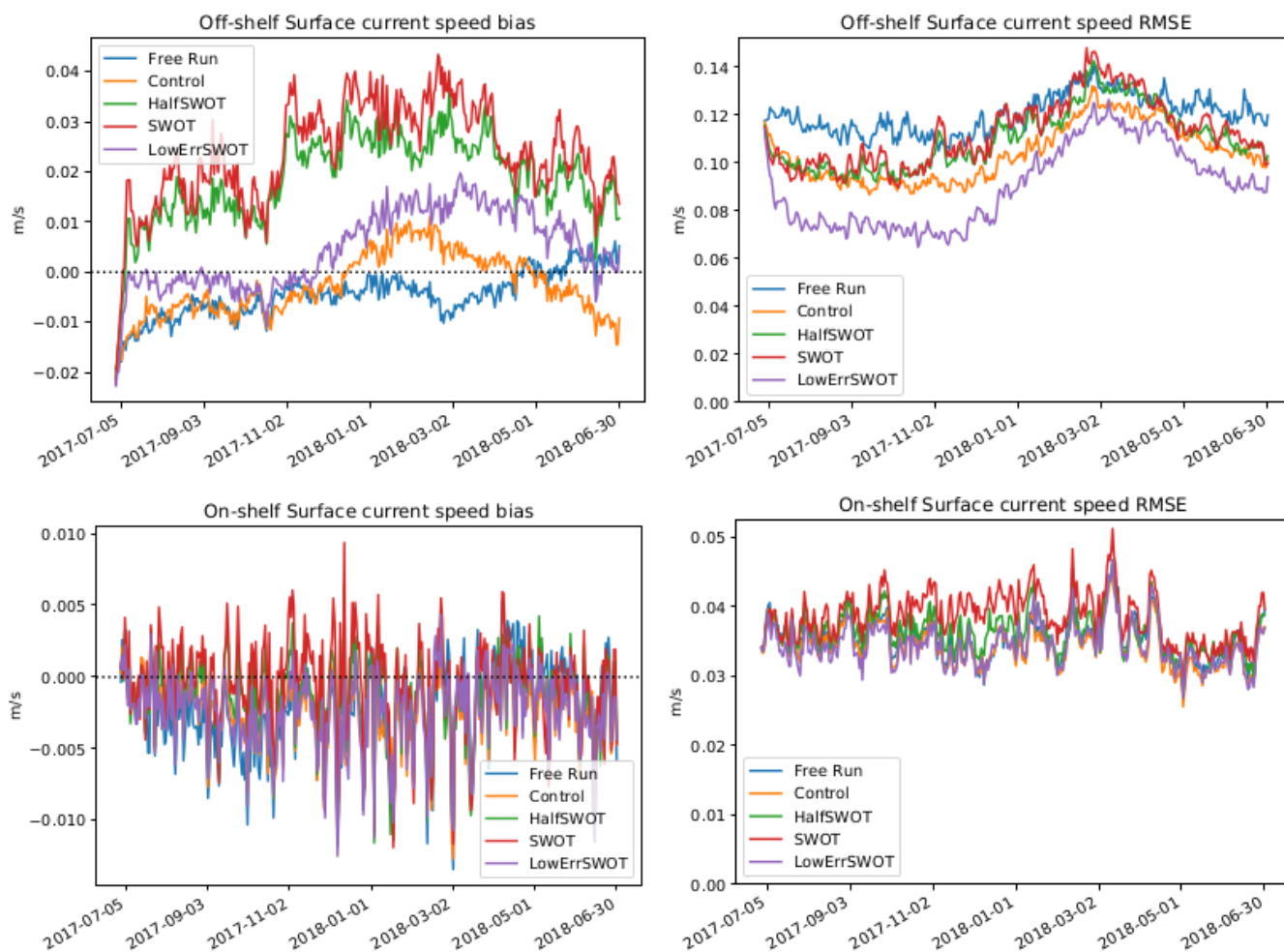
Figure 12. Surface current speed bias (left) and RMSE (right) for the off-shelf (top) and on-shelf (bottom) regions.

features aligned with the SWOT swaths which manifest in the monthly errors as a generally positive bias - that is, the average surface current speed is erroneously increased when assimilating SWOT observations with large correlated errors.

The area-averaged impact of assimilation in each of the experiments is summarised in the time-series of bias and RMSE for the on- and off-shelf regions in Figure 12. As with the SSH RMSE, the surface current reaches a stable RMSE within the first month of the runs. The HalfSWOT and SWOT experiments increase the off-shelf surface current speed errors (calculated over the full run excluding the first month, see Table 3) by 7% and 10%, respectively, while the LowErrSWOT experiment which did not include the large correlated errors shows a reduction of 13% in the RMSE relative to the Control (a similar absolute change in RMSE as from the Free Run to the Control). There is also significant seasonal variation in the error reduction achieved in the LowErrSWOT experiment, with a peak RMSE reduction of ∼20%. As with the SSH RMSE, the on-shelf differences are smaller in absolute terms, but the HalfSWOT and SWOT experiments increase the on-shelf surface current speed errors by 6%

Figure 13. Off-shelf surface current speed bias (left) and RMSE (right) for the HalfSWOT (top) and SWOT (bottom) experiments showing the impact of observation averaging. The LowErrSWOT experiment is also show in both cases.

405 and 15%, respectively, while the LowErrSWOT experiment has a negligible impact on the surface current speed errors relative
to the Control.

The effect of assimilating SWOT observations on the surface current speed bias is less straight-forward. While initially the
off-shelf bias is little changed from the Free Run to the Control and is significantly reduced in the LowErrSWOT experiment,
in the latter half of the experiments the bias drifts in all runs. However, both the experiments assimilating SWOT observations
410 with correlated errors show an increase in the overall RMSE and a large increase in the bias (as seen in the earlier spatial maps
of surface current errors).

Overall, the assimilation of SWOT observations can better initialise the position and strength of eddies and significantly
reduce the surface current RMSE, but correlated SWOT errors lead to a bias toward faster surface currents.

5.3.1 Impact of Observation Averaging

415 While we have shown that assimilating correlated SWOT observations degrades the surface current speed RMSE and bias,
when observation averaging is used we find a significant decrease in the bias and RMSE compared to the experiments without
averaging. The reduced bias is apparent in the maps of monthly surface current errors shown in Figure 11 where the HalfS-
WOT_5km experiment shows a clear reduction in the bias relative to the Control and HalfSWOT experiments, but still larger
bias than in the LowErrSWOT experiment. Figure 13 demonstrates that the 5 km and 20 km observation averaging reduces the
420 overall RMSE by similar amounts in the HalfSWOT and SWOT experiments, but with lower variability in the RMSE over the
year when averaging over the larger radius. In experiments using the full and half-swath widths, the larger averaging radius
results in a lower bias over the year. However, the trend is the same in all the experiments with correlated errors, so the same
underlying issue affects all of the experiments.

6 Discussion

425 Overall, the correlated errors expected in real SWOT observations hinder the full exploitation of this step change in our ability to
observe the ocean mesoscale dynamics. The off-shelf SSH, surface currents, and 3D temperature and salinity can be improved
by the assimilation of SWOT observations with uncorrelated errors. However, when the full correlated errors are included in
the simulated SWOT observations, the SSH RMSE increases relative to the Control, surface currents are erroneously increased
and there is a significant temperature and salinity degradation at almost all depths.
430 The assimilation of the standard network of observations introduces a sub-surface bias near the base of the mixed layer in
the Control experiment, which is exacerbated by additional SWOT observations. This was noted in King et al. (2018) and may
be due to the assimilation of SST observations in our configuration of NEMOVAR. Solutions to this problem are currently
being developed.

On-shelf, the temperature below the mixed layer and salinity at all depths is improved by the assimilation of SWOT obser-
435 vations with uncorrelated errors, but again is degraded when using the full correlated errors, and is exacerbated when using
the full swath width. Due to the lower variability on-shelf, the impact of assimilating SWOT observations on SSH and the

surface currents is less clear, but follows a similar pattern: the assimilation of observations with uncorrelated errors marginally decreases errors, while assimilation of observations with correlated errors leads to a degradation.

440 More generally, the multivariate balances used in our assimilation scheme were originally implemented for an open ocean system. It is therefore likely that the specification of the balances are sub-optimal for on-shelf assimilation. For instance, the water mass balance used may be less appropriate on the shelf, particularly where there is significant influence of rivers. Additionally, the vertical correlation length-scales which were originally parameterised for on open ocean location could be adapted for the shelf-seas. This requires further assessment to understand the effects of each observation type and with simplified balances to understand how they operate in practice in this regime.

445 From the experiments reported here, restricting the SWOT data to the inner half of the swath and applying observation averaging with a 5 km radius can reduce some of the problems caused by assimilating observations with large correlated errors, while retaining some of the benefits from the increased horizontal resolution. To make full use of the upcoming high spatial-resolution SSH observations from wide-swath altimeters, it will be necessary to develop and test techniques to directly account for correlated errors within our data assimilation scheme.

7 Conclusions

450

The impact of assimilating simulated wide-swath altimetry observations from the upcoming SWOT mission was assessed using Observing System Simulation Experiments (OSSEs) with a high-resolution shelf-seas model. A Nature Run using different initial conditions and atmospheric forcing to the OSSEs provided the truth against which the OSSEs were assessed. Simulated observations with realistic errors were then sampled from the Nature Run and assimilated in the OSSEs. A Control run was

455 performed assimilating the standard network of observations, including SST, T/S profiles and nadir altimeter SLA. Further experiments were performed assimilating the standard network of observations in addition to simulated SWOT observations: one with the full swath width and the full correlated errors; another with only half the width of the swath and the full correlated errors; and another experiment assimilating the full-width swath, but which included only uncorrelated errors in the simulated SWOT observations. Four further experiments were performed with the full and half swath data, but employing observation

460 averaging with 5 km and 20 km averaging radii.

The aim of this work was to assess the impact of assimilating SWOT observations with and without the expected correlated errors and to assess simple methods to reduce problems associated with correlated errors. In the off-shelf regions, the SSH, surface currents, and subsurface temperature and salinity were improved to varying degrees by the assimilation of SWOT observations with uncorrelated errors, with both SSH and surface current RMS errors reduced by up to 20%. Similarly, the

465 on-shelf temperature below the mixed layer and salinity at all depths were also improved. However, when correlated errors are included in the full swath SWOT observations, there is a degradation in the sub-surface temperature and salinity, and the SSH and surface currents are degraded with a clear increase in the mean surface currents. While restricting the SWOT data to the inner half of the swath and applying observation averaging with a 5 km radius negated most of the negative impacts, it also severely limited the positive impacts.
470   This work was a first assessment of the potential of assimilating wide-swath altimeter observations in the Met Office shelf-seas analysis and forecasting system. By quantifying the impact of assimilating SWOT observations without correlated errors and the problems encountered when assimilating the observations with correlated errors in our current assimilation scheme, we hope to highlight the potential of swath altimeter observations and the challenge in their use within a current operational system.

475   The experiments presented here have shown that the assimilation of wide-swath altimetry observations has the potential to significantly improve the sea surface height and surface current errors in our operational analyses and forecasts. The assimilation of sub-surface profiles and altimetry can be complementary and so improve forecasts of the vertical structure of the ocean. This has the potential to bring benefits to our services to the marine industry and defence users, search and recovery operations, and others. With further investigation, the increased observation coverage offered by missions such as SWOT could

480   also improve surge forecasting capabilities, potentially improving prediction during high impact weather events. To maximise the benefits from upcoming wide-swath altimetry missions, we aim to further explore methods to ameliorate the effects of correlated errors in the data assimilation and explore methods to explicitly model the correlated errors within our data assimilation scheme.

Data availability. The nature of the 4-D data generated in running the model experiments requires a large tape storage facility. These data

485   are of the order of 10 TB (terabytes) for each of the ten simulations. The model data can be made available upon request from the authors.

Author contributions. RRK designed the experiments and their validation strategy. Both authors contributed to the scientific evaluation and analysis of the results.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Funding support from the Copernicus Marine Environment Monitoring Service is gratefully acknowledged.

490 **References**

Bédard, J. and Buehner, M.: A practical assimilation approach to extract smaller-scale information from observations with spatially correlated
errors: An idealized study, Quarterly Journal of the Royal Meteorological Society, 146, 468–482, 2020.

Berg, P. and Poulsen, J. W.: Implementation details for HBM, DMI Technical Report No. 12-11, 2012.

Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., Mirouze, I., Peterson, K. A., Sellar, A., and Storkey,

495 D.: Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM
forecasts, Geoscientific Model Development, 7, 2613–2638, https://doi.org/10.5194/gmd-7-2613-2014, 2014.

Bloom, S. C., Takacs, L. L., da Silva, A. M., and Ledvina, D.: Data Assimilation Using Incremental Analysis Updates, Mon. Wea. Rev., 124,
1256–1271, https://doi.org/10.1175/1520-0493(1996)124%3C1256:dauiau%3E2.0.co;2, 1996.

Bonaduce, A., Benkiran, M., Remy, E., Le Traon, P. Y., and Garric, G.: Contribution of future wide-swath altimetry missions to ocean

500 analysis and forecasting, Ocean Science, 14, 1405–1421, https://doi.org/10.5194/os-14-1405-2018, 2018.

Carrier, M. J., Ngodock, H. E., Smith, S. R., Souopgui, I., and Bartels, B.: Examining the Potential Impact of SWOT Observations in an
Ocean Analysisž2013Forecasting Systema, Monthly Weather Review, 144, 3767–3782, https://doi.org/10.1175/MWR-D-15-0361.1, 2016.

Cooper, M. and Haines, K.: Altimetric assimilation with water property conservation, Journal of Geophysical Research: Oceans, 101, 1059–
1077, 1996.

505 Craig, A., Valcke, S., and Coquart, L.: Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3. 0., Geosci-
entific Model Development, 10, 2017.

D'Addezio, J. M., Smith, S., Jacobs, G. A., Helber, R. W., Rowley, C., Souopgui, I., and Carrier, M. J.: Quantifying wavelengths con-
strained by simulated SWOT observations in a submesoscale resolving ocean analysis/forecasting system, Ocean Modelling, 135, 40 –
55, https://doi.org/https://doi.org/10.1016/j.ocemod.2019.02.001, 2019.

510 De Mey-Frémaux, P., Ayoub, N., Barth, A., Brewin, R., Charria, G., Campuzano, F., Ciavatta, S., Cirano, M., Edwards, C. A., Federico, I.,
Gao, S., Garcia Hermosa, I., Garcia Sotillo, M., Hewitt, H., Hole, L. R., Holt, J., King, R., Kourafalou, V., Lu, Y., Mourre, B., Pascual, A.,
Staneva, J., Stanev, E. V., Wang, H., and Zhu, X.: Model-Observations Synergy in the Coastal Ocean, Frontiers in Marine Science, 6, 436,
https://doi.org/10.3389/fmars.2019.00436, 2019.

Durand, M., Fu, L.-L., Lettenmaier, D. P., Alsdorf, D. E., Rodriguez, E., and Esteban-Fernandez, D.: The surface water and ocean topography

515 mission: Observing terrestrial surface water and oceanic submesoscale eddies, Proceedings of the IEEE, 98, 766–779, 2010.

Flather, R.: Results from a model of the north-east Atlantic relating to the Norwegian Coastal Current In The Norwegian Coastal Current,
eds, R. Sætre and M. Mork, University of Bergen, 2, 458, 1981.

Fowler, A. M., Dance, S. L., and Waller, J. A.: On the interaction of observation and prior error correlations in data assimilation, Quarterly
Journal of the Royal Meteorological Society, 144, 48–62, https://doi.org/https://doi.org/10.1002/qj.3183, 2018.

520 Gaultier, L., Ubelmann, C., and Fu, L.-L.: The challenge of using future SWOT data for oceanic field reconstruction, Journal of Atmospheric
and Oceanic Technology, 33, 119–126, 2016.

Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses
with uncertainty estimates, Journal of Geophysical Research: Oceans, 118, 6704–6716, 2013.

Graham, J. A., O'Dea, E., Holt, J., Polton, J., Hewitt, H. T., Furner, R., Guihou, K., Brereton, A., Arnold, A., Wakelin, S., et al.: AMM15: a

525 new high-resolution NEMO configuration for operational simulation of the European north-west shelf, Geoscientific Model Development,
11, 681–696, 2018.

Guillet, O., Weaver, A. T., Vasseur, X., Michel, Y., Gratton, S., and Gürol, S.: Modelling spatially correlated observation errors in variational data assimilation using a diffusion operator on an unstructured mesh, Quarterly Journal of the Royal Meteorological Society, 145, 1947–1967, 2019.

530 Halliwell, G. R., Mehari, M. F., Le Hénaff, M., Kourafalou, V. H., Androulidakis, I. S., Kang, H. S., and Atlas, R.: North Atlantic Ocean OSSE system: Evaluation of operational ocean observing system components and supplemental seasonal observations for potentially improving tropical cyclone prediction in coupled systems, Journal of Operational Oceanography, 10, 154–175, 2017.

Halliwell, G. R., J., Srinivasan, A., Kourafalou, V., Yang, H., Willey, D., Le Henaff, M., and Atlas, R.: Rigorous Evaluation of a Fraternal Twin Ocean OSSE System for the Open Gulf of Mexico, Journal of Atmospheric and Oceanic Technology, 31, 105–130, https://doi.org/10.1175/JTECH-D-13-00011.1, 2014.

535 Hoffman, R. N. and Atlas, R.: Future observing system simulation experiments, Bulletin of the American Meteorological Society, 97, 1601–1616, 2016.

Ingleby, B. and Huddleston, M.: Quality control of ocean temperature and salinity profiles̆2014Historical and real-time data, Journal of Marine Systems, 65, 158–175, 2007.

540 Jacobs, G. A., Richman, J. G., Doyle, J. D., Spence, P. L., Bartels, B. P., Barron, C. N., Helber, R. W., and Bub, F. L.: Simulating conditional deterministic predictability within ocean frontogenesis, Ocean Modelling, 78, 1–16, 2014.

King, R. R., While, J., Martin, M. J., Lea, D. J., Lemieux-Dudon, B., Waters, J., and O'Dea, E.: Improving the initialisation of the Met Office operational shelf-seas model, Ocean Modelling, 130, 1 – 14, https://doi.org/https://doi.org/10.1016/j.ocemod.2018.07.004, 2018.

Le Henaff, M., De Mey, P., Mourre, B., and Le Traon, P.-Y.: Contribution of a Wide-Swath Altimeter in a Shelf Seas As-
545 similation System: Impact of the Satellite Roll Errors, Journal of Atmospheric and Oceanic Technology, 25, 2133 – 2144, https://doi.org/10.1175/2008JTECHO576.1, 2008.

Lea, D., Drecourt, J.-P., Haines, K., and Martin, M.: Ocean altimeter assimilation with observational-and model-bias correction, Quarterly Journal of the Royal Meteorological Society, 134, 1761–1774, 2008.

Lea, D., Martin, M., and Oke, P.: Demonstrating the complementarity of observations in an operational ocean forecasting system, Quarterly
550 Journal of the Royal Meteorological Society, 140, 2037–2049, 2014.

Lellouche, J.-M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C.-E., Bourdalle-Badie, R., Gasparin, F., et al.: Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1/12 degree high-resolution system, Ocean Science, 14, 1093–1126, 2018.

Lewis, H. W., Castillo Sanchez, J. M., Siddorn, J., King, R. R., Tonani, M., Saulter, A., Sykes, P., Pequignet, A.-C., Weedon, G. P., Palmer, T.,
555 et al.: Can wave coupling improve operational regional ocean forecasts for the north-west European Shelf?, Ocean Science, 15, 669–690, 2019.

Li, Z., Wang, J., and Fu, L.-L.: An Observing System Simulation Experiment for Ocean State Estimation to Assess the Performance of the SWOT Mission: Part 1̆2014A Twin Experiment, Journal of Geophysical Research: Oceans, 124, 4838–4855, https://doi.org/10.1029/2018JC014869, 2019.

560 Madec, G.: NEMO ocean engine, Note du Pole de modèlisation, France, no 27, ISSN No 1288-1619, 2008.

Metref, S., Cosme, E., Le Guillou, F., Le Sommer, J., Brankart, J.-M., and Verron, J.: Wide-Swath Altimetric Satellite Data Assimilation With Correlated-Error Reduction, Frontiers in Marine Science, 6, 822, https://doi.org/10.3389/fmars.2019.00822, 2020.

Mogensen, K., Balmaseda, M., Weaver, A., Martin, M., and Vidard, A.: NEMOVAR: A variational data assimilation system for the NEMO ocean model, ECMWF newsletter, 120, 17–22, 2009.

565   Mogensen, K., Balmaseda, M. A., and Weaver, A.: The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4, European Centre for Medium-Range Weather Forecasts, 2012.

Morrow, R., Blurmstein, D., and Dibarboure, G.: Fine-scale altimetry and the future SWOT mission, New Frontiers in Operational Oceanography, pp. 191–226, 2018.

Oke, P. R., Larnicol, G., Fujii, Y., Smith, G. C., Lea, D. J., Guinehut, S., Remy, E., Balmaseda, M. A., Rykova, T., Surcel-Colan, D., et al.:
570   Assessing the impact of observations on ocean forecasts and reanalyses: Part 1, Global studies, Journal of Operational Oceanography, 8, s49–s62, 2015.

Ruggiero, G. A., Cosme, E., Brankart, J.-M., Le Sommer, J., and Ubelmann, C.: An Efficient Way to Account for Observation Error Correlations in the Assimilation of Data from the Future SWOT High-Resolution Altimeter Mission, Journal of Atmospheric and Oceanic Technology, 33, 2755–2768, https://doi.org/10.1175/JTECH-D-16-0048.1, 2016.

575   Saulter, A., Bunney, C., Li, J.-G., and Palmer, T.: Process and resolution impacts on UK coastal wave predictions from operational global-regional wave models, in: Proceedings of the 15th International Workshop on Wave Hindcasting and Forecasting and 6th Coastal Hazard Symposium, Liverpool, UK, pp. 10–15, 2017.

Storkey, D., Blockley, E. W., Furner, R., Guiavarc'h, C., Lea, D., Martin, M. J., Barciela, R. M., Hines, A., Hyder, P., and Siddorn, J. R.: Forecasting the ocean state using NEMO:The new FOAM system, Journal of Operational Oceanography, 3, 3–15,
580   https://doi.org/10.1080/1755876X.2010.11020109, 2010.

Tonani, M., Sykes, P., King, R. R., McConnell, N., Péquignet, A.-C., O'Dea, E., Graham, J. A., Polton, J., and Siddorn, J.: The impact of a new high-resolution ocean model on the Met Office North-West European Shelf forecasting system, Ocean Science, 15, 1133–1158, https://doi.org/10.5194/os-15-1133-2019, 2019.

Troccoli, A. and Haines, K.: Use of the temperature–salinity relation in a data assimilation context, Journal of Atmospheric and Oceanic
585   Technology, 16, 2011–2025, 1999.

Uematsu, A., Nakamura, R., Nakajima, Y., and Yajima, Y.: X-band interferometric SAR sensor for the Japanese altimetry mission, COM-PIRA, in: 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS, pp. 2943–2946, IEEE, 2013.

Waters, J., Lea, D. J., Martin, M. J., Mirouze, I., Weaver, A., and While, J.: Implementing a variational data assimilation system in an operational 1/4 degree global ocean model, Quarterly Journal of the Royal Meteorological Society, 141, 333–349,
590   https://doi.org/10.1002/qj.2388, 2015.

Weaver, A. T., Deltel, C., Machu, É., Ricci, S., and Daget, N.: A multivariate balance operator for variational ocean data assimilation, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 131, 3605–3625, 2005.

While, J. and Martin, M. J.: Variational bias correction of satellite sea-surface temperature data incorporating observations of the bias,
595   Quarterly Journal of the Royal Meteorological Society, 145, 2733–2754, 2019.

Xu, Y. and Fu, L.-L.: The Effects of Altimeter Instrument Noise on the Estimation of the Wavenumber Spectrum of Sea Surface Height, Journal of Physical Oceanography, 42, 2229–2233, https://doi.org/10.1175/JPO-D-12-0106.1, 2013.