Technical Memo

ECMWF
European Centre for Medium-Range Weather Forecasts

898

Outcomes from UGROW-IO: Forecast errors in the Eastern Indian Ocean across lead times

Michael Mayer, Magdalena Alonso Balmaseda, Stephanie Johnson, Linus Magnusson, Christopher Roberts, Hao Zuo

22 Jun 2022

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:

http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

Abstract

UGROW is an ECMWF cross-departmental project focused on Understanding systematic error GROWth from hours to seasons ahead. The forecast errors over the Indian Ocean is one of the focus themes (UGROW-IO), which assesses the lead-time-dependent evolution of forecast errors in the Indian Ocean, which happens to be the eastern pole of the Indian Ocean Dipole Index.

Seasonal forecasts rapidly develop an easterly surface wind bias in the Equatorial Eastern Indian Ocean (EEIO) during boreal summer, visible already at week 1. This bias amplifies with time via coupled feedbacks, and it eventually manifests in a cold SST bias, as one of the prominent errors at the seasonal time scales in SEAS5.

The EEIO exhibits two regimes: a warm pool regime or convective regime, characterized by a local negative wind-SST coupled feedback which limits warming, and a cold-tongue regime, characterized by a non-local positive wind-SST-thermocline feedback which enhances and maintains the cold SST. The error in the seasonal forecasts suggest that in the model the cold-tongue regime dominates, while the wind-SST coupling in the warm regime is very weak compared to observations

We have studied the dependency of this error to ocean and atmosphere initial conditions, ocean and atmospheric resolution, and different model cycles. We conclude that there are two fundamental and independent sources of errors that lead to the SST errors in seasonal forecast. The first one is of atmospheric nature and is largely related with too stable easterly circulation present in the equatorial Indian Ocean, further characterized by the lack of response of the local winds to local surface heating in the EEIO. This induces an easterly bias which leaves the model state predominantly in a state with a shallow thermocline and cold SSTs in the EEIO. The second error is of oceanic origin, associated with a too shallow thermocline, which enhances the SST errors arising from errors in the wind. Ocean initial conditions, which depend on both the quality of the assimilation and the ocean model, play an important role in this context. Nevertheless, the version of the ocean model used for the forecast can also play a non-negligible role at the seasonal time scales, by amplifying or damping the subsurface errors in the initial conditions due to the strength of the atmosphere-ocean coupling in this region.

1. Motivation

Investigation of errors in the Indian Ocean was motivated by the fact that the Eastern equatorial Indian Ocean (EEIO) is one of the few regions worldwide where forecast skill in surface parameters has degraded in SEAS5 compared to SEAS4 especially for forecasts started in May. Figure 1 is reprinted from Johnson et al. (2019) and shows the difference in CRPSS between SEAS5 and SEAS4 in JJA, which clearly shows the statistically significant degradation of SEAS5 in the EEIO. The degradation is also associated with a negative SST bias and overly large spread in the EEIO, as will be shown later.

Given the importance of the EEIO for teleconnections and global climate variability, it is essential to better understand these changes. Stockdale et al. (2018) assessed the impact of various modelling choices (coupled vs uncoupled, resolution, stochastic physics etc.) on the SST bias in the EEIO but did not explore the underlying processes. This manuscript aims at characterization of forecast errors in the EEIO at different lead times, to gain a better understanding of mechanisms and provide recommendations for metrics to guide further model development. The rest of the paper is organized as follows. Section 2 shows in more detail the changes

in SST forecasts in the Indian Ocean from SEAS4 to SEAS5. Section 3 investigates errors in atmospheric and oceanic initial conditions used in SEAS5 and compares them to those used for SEAS4. Section 4 discusses evolution of surface and subsurface errors in the EEIO as a function of lead time across the medium, extended and seasonal ranges. Section 5 explores the state dependence of wind and precipitation errors in the Indian Ocean. Section 6 employs diagnostics developed in the earlier sections to evaluate the performance of more recent model cycles. Section 7 concludes with a discussion.

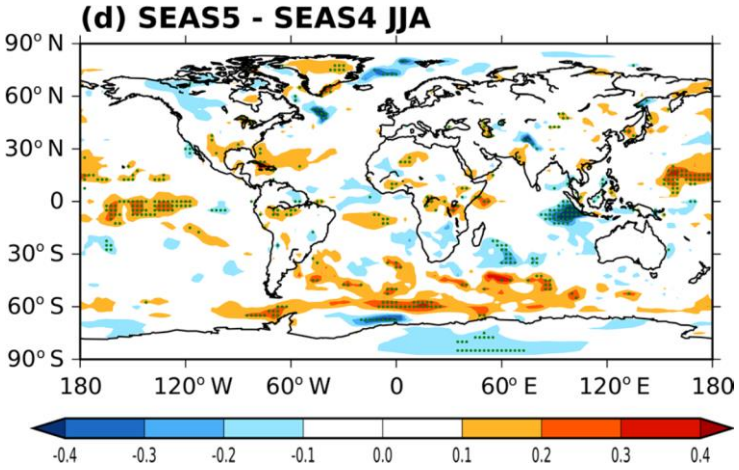

Figure 1 . Difference in CRPSS of t2m forecasts for JJA from SEAS5 and SEAS4. Scores are based on forecasts initialized in May 1981-2016. Reprinted from Johnson et al. (2019).

2. From SEAS4 to SEAS5: Changes in SST forecast scores and indices

Figure 2 is reprinted from Johnson et al. (2019) and is here used to further set the scene. SEAS5 exhibits a positive SST bias in the western equatorial Indian Ocean (WEIO; 10S-10N, 50-70E) with a moderate annual cycle (Figure 2a). Compared to SEAS4, SSTs in SEAS5 have warmed in this region, but the shape of the annual cycle is similar. In the WEIO, other performance metrics than bias are similar in SEAS4 and SEAS5 (Figure 2c, e, g). In the eastern equatorial Indian Ocean (EEIO, 10-0S 90-110E), SEAS5 SST bias has a more pronounced annual cycle (Figure 2b). While it is positive in boreal winter and spring, it is negative in boreal fall, especially in ASON and most pronounced for 4-6 months lead time.

This represents a significant change compared to SEAS4, which has a seasonally more uniform SST bias, which is comparatively small especially for 1-3 months lead time. Also the other metrics shown in Figure 2 show the degradation of SEAS5 SST forecasts in the EEIO, with greatly enhanced amplitude ratio in JASOND (Figure 2d), reduced anomaly correlation especially in boreal fall (Figure 2f), and substantially increased RMS errors in JASON compared to SEAS4. The degraded performance in boreal fall indicates a reduced ability of the system to predict the state of Indian Ocean Dipole (IOD), which usually has its peak during this season.

Figure 3 compares SEAS4 and SEAS5 SST forecasts in the EEIO in ASO, when the negative forecast bias in SEAS5 is most pronounced. SEAS4 shows clear anti-correlation between ensemble mean SSTs and SST spread (Figure 3a). This makes sense physically, as at lower SSTs the SSTs themselves are more sensitive to small changes in the surface winds (see section 5.3). Although SEAS5 forecasts exhibit a similarly negative correlation between ensemble mean SSTs and spread, it is evident that the correlation mainly arises from the few points with SST values >28°C, while most forecasts cluster in the low SST/high spread regime. Figure 3a furthermore suggests that for a given SST, the spread in SEAS5 is higher than in SEAS4, especially for warmer

SSTs. Figure 3b compares SST spread and forecast errors. The spread-error relationship in SEAS5 appears to have two distinct regimes: a more populated cold error/large spread regime, with overestimation of spread in some instances, and a scarcely populated warm error/small spread regime, where the forecasts are clearly under-dispersive. This relationship is not present in SEAS4, where correlation between spread and forecast error is close to 0.

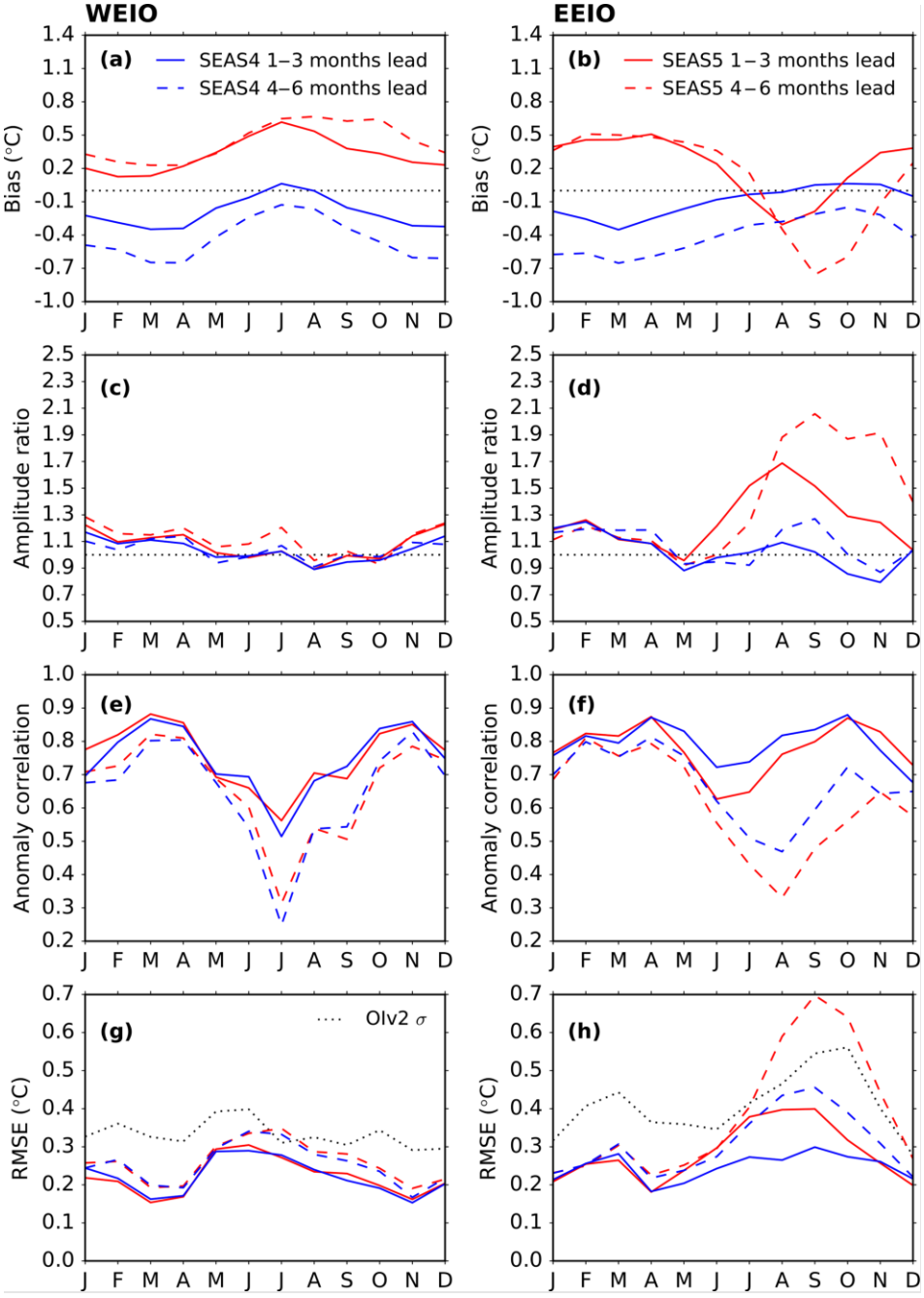

Figure 2. Various SST forecast performance metrics in the Western (WEIO) and Eastern (EEIO) Equatorial Indian Ocean as a function of target month: (a,b) bias, (c,d) amplitude ratio, (e,f) anomaly correlation, (g,h) RMSE. Blue (red) lines show SEAS4 (SEAS5), and solid (dashed) lines show seasonal averages for 1-3 (4-6) months lead time. OIv2 SST data are used as a reference. The observed standard deviation of SST interannual variability in WEIO and EEIO is shown as a function of calendar month in (g) and (h), respectively. Reprinted from Johnson et al (2019).

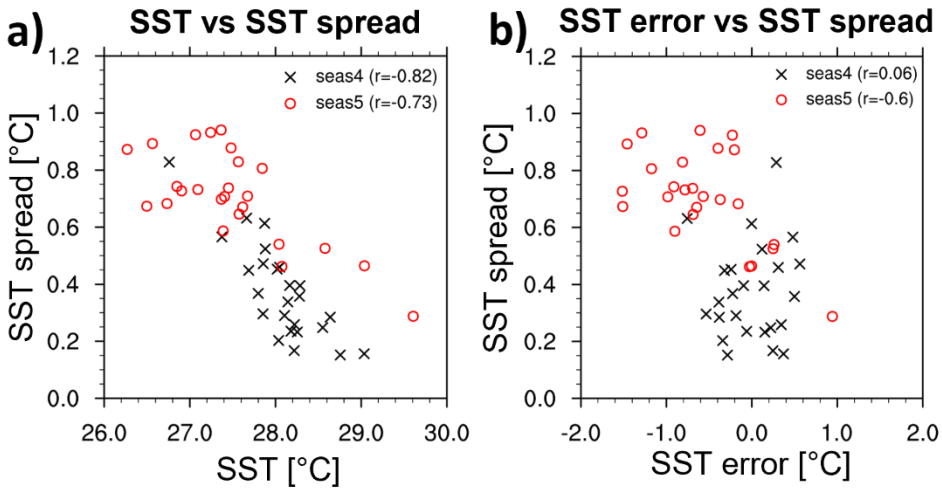

Figure 3. Relationships of different metrics of ASO mean SST forecasts (initialized on 1 May 1993-2017) in the EEIO. a) ensemble mean SSTs versus SST ensemble spread; b) ensemble mean SST forecast error versus SST ensemble spread.

To examine whether these relationships hold in other forecast systems, the characteristics of SEAS4 and SEAS5 were compared to other forecast systems in the C3S multi-system seasonal forecast (https://doi.org/10.24381/cds.68dd14c3). Ten ensemble members from seven forecast systems with reforecasts were included. In Figure 4, the climate mean SST bias in the EEIO is compared to the root-mean-square of the spread and the root-mean-square error of the bias-corrected ensemble reforecasts. A clear relationship between SST bias and RMS error in the EEIO is shown in Figure 4b. When the SST bias is small, the RMS error is also at a minimum. This figure also shows that the largest SST biases and RMS errors are associated with cold SST biases. Figure 4a shows that these cold SST biases are associated with larger spread, as discussed above for SEAS4 and SEAS5 (see also Figure 3). Together, these two relationships suggest that systems with a larger warm SST bias would have low spread-error ratios compared to models with a cold SST bias. In Figure 4c, we show the spread-error ratio of the systems compared to the SST drift. We can see that SEAS4 and SEAS5 are overspread for their RMS errors, despite having very different SST bias and error characteristics. However, the other systems have too little spread for their error. There appears to be little relationship between SST drift and the spread-error ratio, except perhaps a lack of models exhibiting a large warm SST bias (> 0.3 C) and excess spread at the same time. However, this could also be due to the limited number of systems considered here.

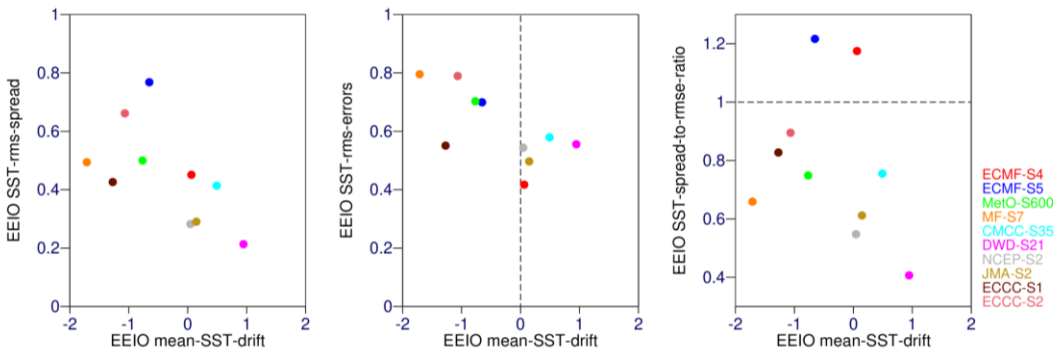

Figure 4. Relationships between ASO mean EEIO SST bias, RMS spread and RMS error in the C3S multi-system seasonal re-forecasts initialized on 1 May from 1993 to 2016. (a) RMS spread versus SST bias (b) RMS error versus SST bias and (c) Spread-error ratio versus SST bias.

Figure 5 explores how these SST characteristics are related to the mean zonal wind bias in the EEIO. In Figure 5a, the mean SST bias and mean zonal wind bias are compared, which shows that an overly strong easterly wind (negative wind bias) always exists with a cold SST bias, and weak easterly wind (positive wind bias) always exists with a warm SST bias. This is consistent with an overly strong easterly bias leading to increased upwelling off the coast of Java and cooling SSTs in the EEIO. However, there is a lot of variation in the relationship between the magnitude of the zonal wind and SST bias – clearly indicating that other characteristics of the model, beyond the zonal wind bias, determine the size of the SST bias. As expected from the relationships with SST bias in Figure 4, Figure 5b and c show that a stronger zonal wind in the EEIO is also associated with increased SST RMS error and increased SST spread, though any relationship with SST spread-error ratio is again less clear (not shown).

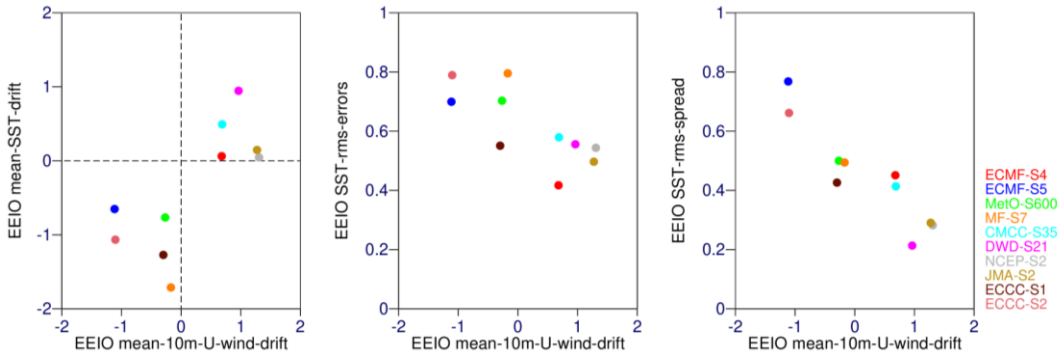

Figure 5. Relationships between ASO mean EEIO 10m zonal wind bias and SST characteristics in the C3S multi-system seasonal re-forecasts initialized on 1 May from 1993 to 2016. (a) SST bias versus zonal 10m wind bias (b) SST RMS error versus zonal 10m wind bias and (c) SST RMS spread versus zonal 10m wind bias.

3. Errors in initial conditions

In this section we assess errors in the EEIO of the initial conditions used in SEAS4 and SEAS5. These are ORAS4 (SEAS4) and ORAS5 (SEAS5) for the ocean and ERA-Interim (both systems) for the atmosphere.

3.1. Atmospheric reanalyses

We validate analysed 10m zonal winds (u10m) from ERA-Interim against in-situ measurements taken from buoys located in the EEIO and are provided through Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction (RAMA; Mcphaden et al. 2009). This parameter is crucial as it is strongly coupled to SSTs in two ways, i.e., driving SST changes and responding to changed SSTs. Figure 6 shows (a) time series and (b) mean annual cycle of u10m bias from ERA-Interim and buoy measurements for a buoy located at 0N, 90E, i.e., at the edge of the EEIO. The mean annual cycles were computed for the months when observations were available, and buoy measurements were adjusted to account for the measurement height of the anemometers (4m) following Bidlot et al. (2002). ERA5 data is included as well for comparison. When considering full time series (left column), reanalysis-based winds are in generally good agreement with the observations, both in terms of interannual variations and the mean annual cycle. Considering the mean annual cycle of the bias compared to observations, ERA-Interim and ERA5 show a weak easterly bias (too weak westerlies) at the equator (0N, 90E) during April-November, which is more pronounced in ERA-Interim than ERA5 (Figure 6b). This is consistent with the analysis by Belmonte Rivas and Stoffelen (2019), who compared ERA-Interim and ERA5 surface winds (taken from short-term forecasts) to ASCAT observations and found relatively small zonal wind errors in the EEIO. Nevertheless, the small easterly bias in the reanalyses despite

observational constraints is already indicative of an easterly bias of the atmospheric model, which becomes much more pronounced in forecast mode (see subsequent sections). We note that climatological near-surface winds at the buoy location have a strong southerly component in boreal summer, but the reanalysis bias of the v-component (too weak southerlies) is smaller than that of the u-component and hence not shown here.

To which degree the easterly bias at 0N,90E during April-November has the potential to induce cooling in the ocean initial conditions is unclear, as it is relatively far off the coast, where one would expect the strongest effect on the upwelling. Moreover, the buoy location is at the northern edge of climatological easterlies in boreal summer. Figure 6b also suggests that ERA5 winds are in better agreement with the observations, with the easterly bias roughly halved. However, as will be shown later (Appendix A1), change from ERA-Interim to ERA5 forcing does not lead to significant changes in the analysed ocean state.

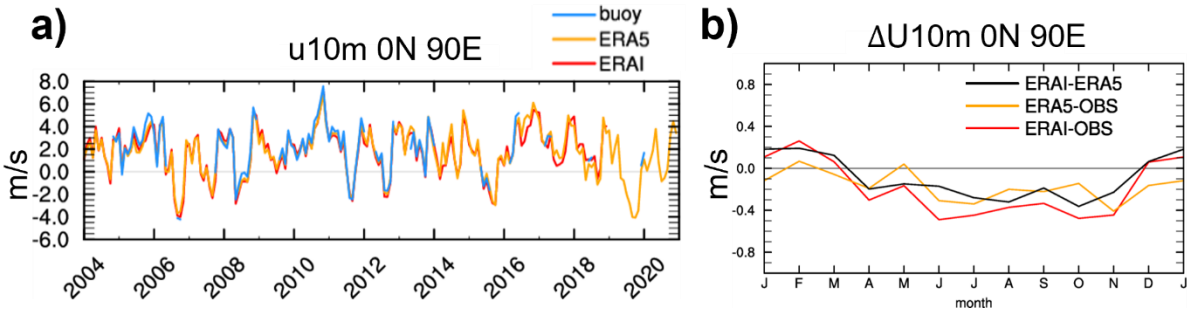

Figure 6. a) time series of u10m for ERA-Interim, ERA5, and the RAMA buoy at 0N, 90E and b) annual cycle of u10m bias from reanalyses (w.r.t. buoy measurements, considering only months when observations are available) and difference between reanalysis winds (considering all months) at 0N, 90E.

We also considered buoy measurements at two locations south of the equator, but there the results are less clear (not shown). ERA-Interim seems to exhibit less reliable winds towards the end of its lifetime, but also the buoys seem to contain spurious values, especially in the early period. Because of these complications, we refrain from including results from the buoys south of the equator here.

3.2. Oceanic reanalyses

In this subsection we focus on the differences in the ocean initial conditions ORAS5 (Zuo et al. 2019) and ORAS4 (Balmaseda et al. 2012). The ocean component differs substantially between the two systems. ORAS5 configuration is ~0.25° in the horizonal, 1m vertical resolution in the upper ocean layers, higher than that in ORAS4 (~ 1° in the horizontal and 10m in the vertical over the upper ocean). There are also changes in the ocean model version and data assimilation. In order to isolate the impact of ocean resolution, we conduct an additional ocean reanalysis similar to ORAS5 but at the lower the horizontal and vertical resolution of ORAS4. We call this ocean reanalysis ORAS5-LR.

We first compare three ocean reanalyses ORAS5, ORAS5-LR, and ORAS4 to in-situ-based analyses Hadley EN4 (Good et al. 2013) in its version 4.2.1. We note that EN4 profiles are assimilated into the ECMWF ocean reanalyses and hence should not be viewed as independent validation data, but systematic differences can indicate biases in the assimilating ocean model. Figure 7 shows the difference between the mean annual cycle (1993-2013) of water temperature in the EEIO from (a) ORAS5, (b) ORAS5-LR, and (c) ORAS4 and EN4, as a function of depth. ORAS5 and ORAS5-LR exhibit a cool bias order ~0.5K around the thermocline almost year-round, with a maximum around 100m depth. Biases in the mixed layer and below 150m are small. The similarity between ORAS5 and ORAS5-LR biases suggests that they are resolution-independent. ORAS4 has

a reduced cool bias in 100m, which only appears in a few calendar months (February and March, August and September). In return, ORAS4 exhibits a warm bias in the mixed layer. To facilitate comparison of the temperature bias with available data of temperature analysis increments of the three reanalyses (discussed next), the second row of Figure 7 repeats the bias plots but for the period 2011-2013.

It is useful to now consider the temperature analysis increments in the different reanalyses, which are plotted in the same manner as the temperature bias in the third row of Figure 7. As averaging period we choose the short period during which all three reanalyses archived the increments (2011-2013). ORAS5 (Figure 7g) and (to slightly lesser degree) ORAS5-LR (Figure 7h) have strongly positive increments between 50 and 150m from May-November (Figure 7d-e) and thus act to reduce the bias during that time of the year (Figure 7d-e) . Increments in boreal winter are generally negative during January-April and limit the temperature bias, which tends to be slightly positive during this time of the year. Thus, in general the data assimilation works effectively to limit biases around the thermocline, as expected. Only in a few instances increments and bias are of the same sign, e.g., in December, which may be related to a temporary reduction of available observations, rejection of observations due to too large deviations, or spatial propagation of temperature increments. Figure 7i shows temperature increments from ORAS4, which are negative around the thermocline almost all year, which is in stark contrast to the increments of the two ORAS5 versions. It is especially surprising that ORAS4 and ORAS5-LR exhibit such differences in the increments, despite the same resolution and atmospheric forcing. This confirms that the differences seen in ORAS4 and ORAS5 do not stem from the resolution. There are other contributing factors, such as differences in the used flux formulation (bulk formulation in ORAS5 versus direct IFS-fluxes in ORAS4), different assimilation options, different observational data sets, and differences in the model versions used in ORAS5 (NEMO3.4) and ORAS4 (NEMO3.0), especially those associated with vertical mixing (inclusion of wave effects or different choices for the tidal mixing).

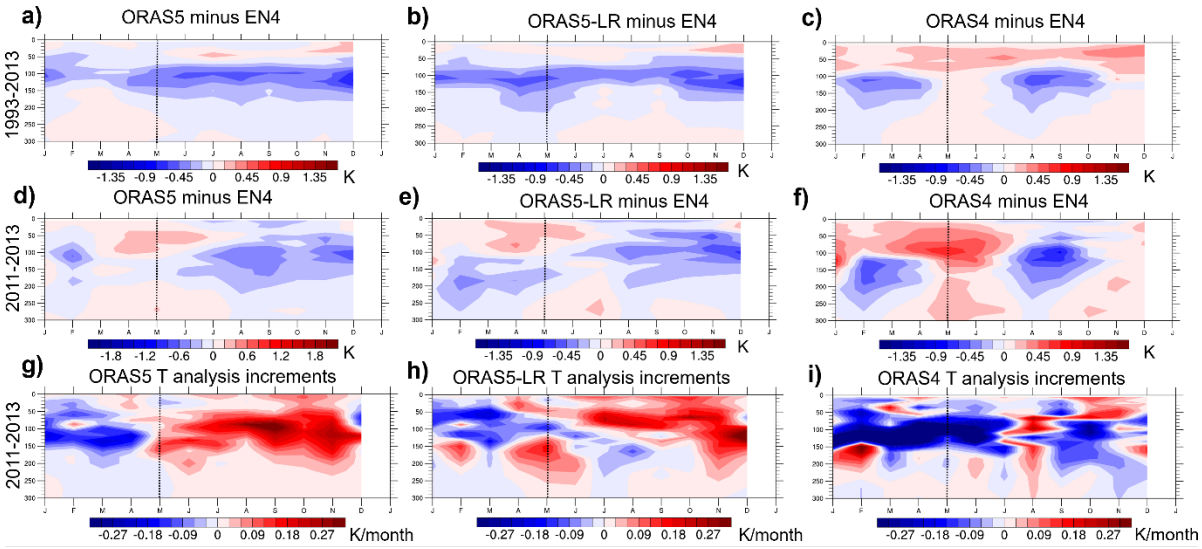

Figure 7. mean annual cycle of subsurface temperature bias in EEIO of a)/d) ORAS5, b)/e) ORAS5-LR, and c)/f) ORAS4 compared to EN4 analyses averaged over (first row) 1993-2015 and (second row) 2011-2013. Temperature increments in EEIO from g) ORAS5, h) ORAS5-LR, and i) ORAS4 averaged over 2011-2013.

We performed the diagnostics shown in Figure 7 also on a prototype of the new ocean analysis system ORAP6.1 (Zuo et al., 2021, see Appendix A1). The main differences between ORAP6 and ORAS5 is related with observational datasets (reprocessed in-situ and altimeter, SST), use of atmospheric forcing from ERA5, and modified subsurface bias correction scheme. The subsurface bias in ORAP6.1 is slightly reduced, but its

vertical structure as well as the seasonal cycle of temperature assimilation increments are very similar compared to ORAS5.

4. Evolution of errors across forecast ranges

After assessing biases in the initial conditions used in the forecasting system, we now turn to forecast biases of zonal wind, SST and subsurface temperature as a function of lead time. As before, we assess atmospheric and sub-surface errors separately. We use results from medium-range forecasts using IFS CY46R1 as well as extended range hindcasts from CY46R1. Table 1 summarizes the used seasonal forecast data.

Name	Initial conditions	Resolution	Model	Exp id
SEAS4	ORAS4, ERA-Interim	T255 L91 ORCA1 Z42	IFS CY36R4	
SEAS5	ORAS5, ERA-Interim/operations	Tco319 L91 ORCA025 Z75	IFS CY43R1	
SEAS5-LR	ORAS5-LR, ERA-Interim	Tco319 L91 ORCA1 Z42	IFS CY43R1	gltf
SEAS5-LR-O4	ORAS4, ERA-Interim/operations	Tco319 L91 ORCA1 Z42	IFS CY43R1	ghzq
SEAS5-obsSST	ERA-Interim/operations	Tco319 L91	IFS CY43R1, uncoupled, forced by observed SSTs	gv90
FC_47R1	ORAS5-LR/ERA5	Tco199 L91 ORCA1 Z75	IFS CY47R1	hcy8
FC_47R2	ORAS5-LR/ERA5	Tco199 L91 ORCA1 Z75	IFS CY47R2	hfnf
FC_47R3	ORAS5-LR/ERA5	Tco199 ORCA1 Z75	IFS CY47R3	hm3f
FC_ORAP6	ORAP6.1, ERA-Interim/operations	Tco199 L137 ORCA025 Z75	IFS CY47R1	hiu0
FC_ORAS5	ORAS5, ERA-Interim/operations	Tco199 L137 ORCA025 Z75	IFS CY47R1	hiqa
FC_N3.4_LR	Low-resolution NEMO3.4 ocean control run/ERA5	Tco199 L137 ORCA1 Z75	IFS CY47R3	hgr8
FC_NX_LR	Low-resolution ocean control run using modified ocean model NEMOX/ERA5	Tco199 L137 ORCA1 Z75	IFS CY47R3	hpa0
FC_N34_LR_PERT	Perturbed NEMO3.4 ocean control run/ERA5	Tco199 L137 ORCA1 Z75	IFS CY47R3	hr1o

Table 1.Seasonal forecast experiments used in this study, including information on oceanic and atmospheric initial conditions and resolution (atmospheric horizontal and vertical resolution, oceanic horizontal and vertical resolution). All forecasts are run in coupled mode unless otherwise stated.

4.1. Errors in the atmospheric forecasts

We begin with the assessment of short-range wind errors in the ECMWF HRES forecasts and put them in relation to other NWP models. Figure 8 shows the bias in forecasts of 10m-u-wind in the IFS HRES at (a) 1 day and (b) 3 days lead time for JJASO 2018. The largest error, especially at day 3, is the too strong easterlies in the Western Tropical Pacific, which is strongest in DJF (not shown) and has been present in ECMWF seasonal forecasts for a long time (see Magnusson et al. 2013). Focusing on the tropical Indian Ocean, we find a strong negative bias along the equator and a positive bias in the northern tropical IO. The wind bias bears the signature of a too strong monsoon circulation. It is a long-standing issue that was extensively discussed in Rodwell et al. (2010). It is clear that the wind bias develops very fast in the Indian Ocean as it is present already during the first days of the forecast.

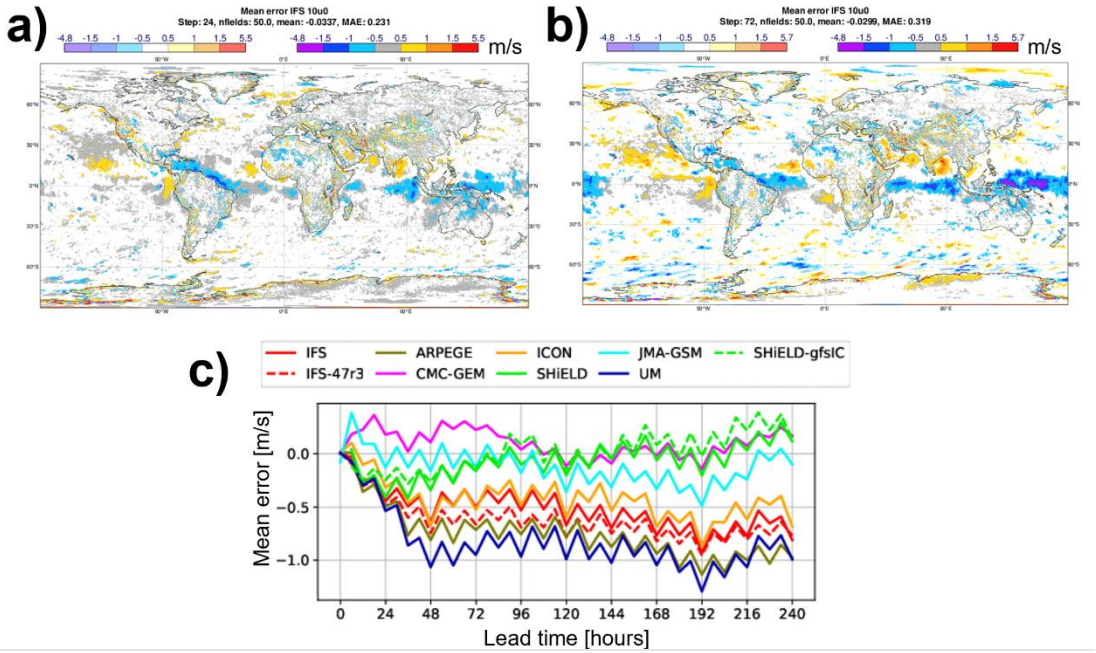

Figure 8. a) 10m u-wind bias of the IFS HRES (a) day-1 and (b) day-3 forecast in JJASO 2018. c) Lead-time evolution for u10m of mean error in the EEIO for different models contributing to DIMOSIC. Winds were verified against own analysis.

To put the ECMWF bias in perspective, Figure 8c shows result from the DIMOSIC project, where different global models were initialised from the same (ECMWF) analysis (Magnusson et al., 2022, submitted to BAMS). Here we make use of forecasts from 6 June 2018 - 1 November 2018 (49 cases). Figure 8c shows the evolution of the bias of u10m in the EEIO from the different models as a function of lead time verified against the ECMWF analysis. Most of the models develop a negative bias over the first 2 days. The two red lines show ECMWF IFS CY47R1 (solid) and IFS-47r3 (dashed). IFS CY47R3 shows a somewhat stronger bias than IFS CY47R1 but gives on the other hand a lower RMSE (not shown). The plot also includes the SHiELD model initialised from GFS initial conditions, and this model starts from a similar mean state and shows a similar model drift, indicating that the issue is not dependent on the initial conditions. A similar wind bias in the EEIO is found in both IFS versions and DWD-ICON, METEOFRANCE-ARPEGE and UKMO-UM, with the strongest bias in UKMO-UM. On the other hand, this bias is not present in GFDL-SHiELD, JMA-GSM and CMC-GEM. However, GFDL-SHiELD and JMA-GSM have strong negative biases in other parts of the tropical Indian Ocean (not shown).

We now turn to 10m zonal wind errors in extended range forecasts, shown for CY46R1 hindcasts 2000-2019, verified against ERA5. Figure 9a and b show the week-1 wind bias for May and August starts, respectively. The moderate (~0.3 m/s) easterly wind bias in the equatorial and southern subtropical Indian Ocean and the stronger easterly wind bias in the equatorial Pacific (most pronounced at order 0.7-0.8 m/s west of the date line) as well as the westerly bias in the northern IO are similar to the bias in short-range forecasts. Week-1 biases for August starts are qualitatively similar, but with a generally higher magnitude, likely because in August the monsoon circulation is fully established, and related model biases appear to be proportional to its strength. Going to weeks 2-4 (Figure 9c and d), we see a strengthening of the easterly bias along the equator, where magnitudes are 1.5 m/s. The easterly wind bias further extends to the subtropical Pacific. While the westerly wind bias persists in the northern subtropical Indian Ocean for May starts (Figure 9c), it diminishes in August starts (Figure 9d).

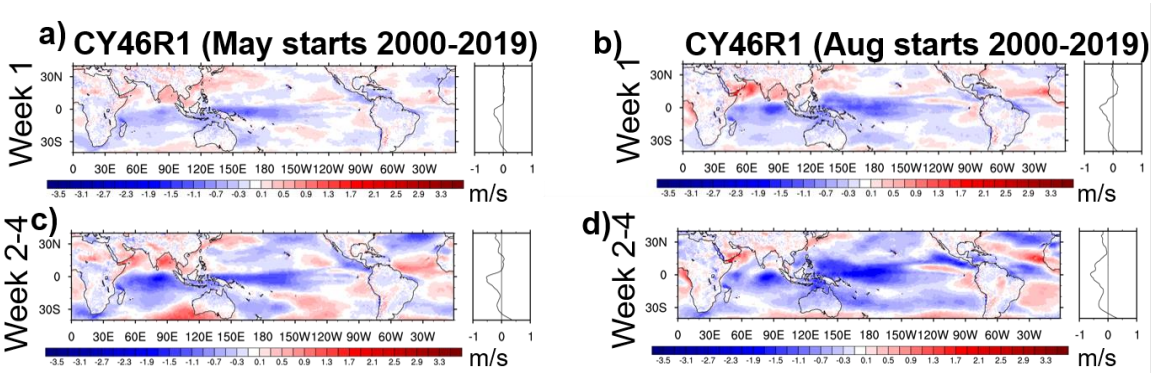

Figure 9. 10m u-wind errors in extended-range forecasts (CY46R1) initialized in (a,c) May and (b,d) August 2000-2019 for week (a,b) 1 and (c,d) 2-4. ERA5 is used as reference.

Figure 10 presents 10m zonal wind biases for coupled and uncoupled seasonal forecasts initialized in May. SEAS5-ObsSST is an atmosphere-only seasonal reforecast experiment, equivalent to SEAS5, where the atmosphere model sees the daily SST prescribed from observations. Figure 10a and b show month-1 (i.e., May) wind biases of SEAS5 and SEAS5-obsSST, respectively. First, they are qualitatively similar to the week 2-4 biases of the extended range hindcasts started in May as shown in Figure 9c. However, the seasonal forecast biases appear generally stronger, but this may also be related to the different periods used for computation. The difference between the wind biases of coupled and uncoupled models (Figure 10c) show that in some regions the coupled model shows more pronounced biases (e.g., western equatorial Pacific), but reduced biases in other regions (e.g., equatorial Indian Ocean).

Progressing to JJA averages (Figure 10d-e), we see further amplification of the bias patterns seen already in May. Especially in the western equatorial Pacific, SEAS5 exhibits a much stronger easterly wind bias than SEAS5-obsSST (Figure 10f), which is likely the result of a positive coupled feedback between the cold SST bias in the eastern equatorial Pacific of SEAS5 (Johnson et al., 2019) and the Equatorial easterlies. Thus, in areas with positive coupled feedback, the wind biases of the coupled model in JJA are mostly stronger compared to those of the uncoupled model. One exception is the north-western subtropical Pacific and northern Indian Ocean, where SEAS5 has slightly smaller biases.

The increased easterly bias of SEAS5 along the equatorial but decreased westerly bias in the northern Indian Ocean is related to regional changes of errors in the coupled model. The former is associated with the enhanced zonal SST gradient along the equatorial Indian Ocean (compare Figure 1a and b), while the latter is related to the improved precipitation climatology of SEAS5 in the Southeastern monsoon region (SEAS5-obsSST exhibits strong overestimation of precipitation in this region; compare Fig. A12 in Stockdale et al. 2018). The

improved monsoons in the coupled model may be related to the warm SST bias present in the northern Indian Ocean (compare Fig. 12d in Stockdale et al., 2018), which acts to reduce the land-sea temperature contrast and consequently could act to reduce the monsoonal inflow to the continent.

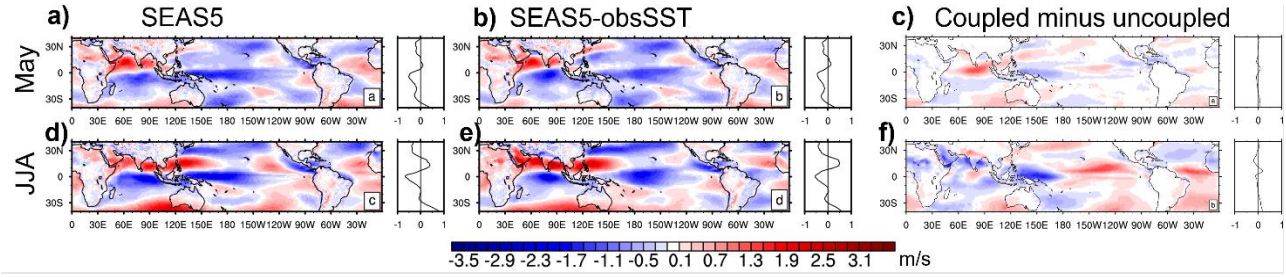

Figure 10. 10m u-wind errors in long-range (a,d) coupled forecasts from SEAS5, (b,e) SEAS5-obsSST, and (c,f) their difference initialized on 1st of May 1993-2018 for (a,b,c) May and (d,e,f) JJA. ERA5 is used as reference.

Figure 11 relates wind and SST biases in the EEIO in extended-range and seasonal forecasts for different lead times. Figure 11a confirms that in the extended range the easterly wind bias develops in week-1 and saturates in weeks 2-4. The evolution is similar for all four months May to August, but the strongest week 2-4 wind bias is found in June. Extended range SST biases (Figure 11a) are positive in week 1 and then diminish, with a similar behaviour for all start months. The initial warming may be related to the shoaling mixed layer which then can be warmed more efficiently by surface heat fluxes. Figure 11b shows the evolution of wind and SST biases in seasonal forecasts started in May. Wind biases of the coupled (SEAS5) and uncoupled forecasts (SEAS5-obsSST) are similar in month 1. They further worsen in the coupled forecast from month 2 onward, which is associated with the development of the cold SST bias from month 2 onward, while the wind bias remains roughly stable in SEAS5-obsSST. The relationship between SST in the EEIO (and WEIO) with the winds in the EEIO will be further investigated in the subsequent sections. SEAS4 shows a different behaviour than SEAS5, despite a weakly positive SST and weak easterly bias in lead month 1, biases reduce to almost zero in JJA. Figure 11b and c additionally include results for other experiments which will be referred to later.

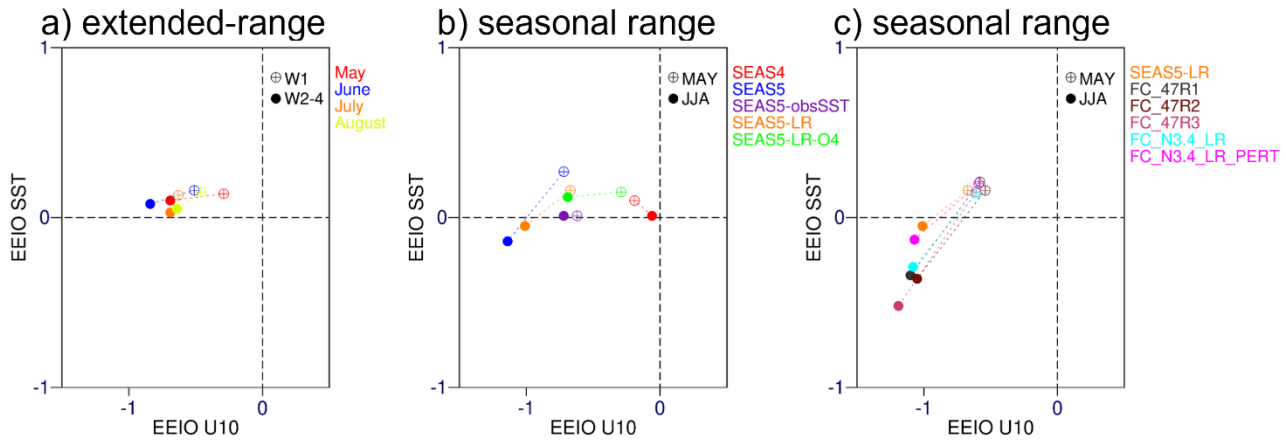

Figure 11: 10m u-wind and SST bias in EEIO at different lead times (a) in extended-range (2000-2019) and (b,c) seasonal forecasts (May starts 1993-2015). ERA5 is used as a reference.

We note that the week 1-4 SST bias (~0.1K) of the extended range forecasts (Figure 11a) is lower than that of SEAS5 in month 1 (~0.3K) bias (Figure 11b). This can partly be attributed to the different time periods considered, as the SEAS5 month-1 SST bias for 2000-2018 is only ~0.2K (not shown). The remaining

difference may be attributable to the different model cycles (CY43R1 for SEAS5, CY46R1 for the extended range results).

4.2. Subsurface errors

Figure 12 shows the lead-time dependent subsurface temperature bias in seasonal forecasts initialized in May 1993-2013. In month-1 SEAS5 (Figure 12a) has a moderate cold bias of ~-0.5K peaking at ~100m depth, which is largely consistent with the temperature biases found for ORAS5 (Figure 7). The negative bias at that depth rapidly amplifies with lead time and becomes lower than -3K from July onward. From July onward, the bias also extends to the surface, consistent with Figure 11b. The forecast bias below ~200m remains small at all lead times. Figure 12b shows the subsurface temperature bias of a low-resolution version of SEAS5 (SEAS5-LR) that is initialized from ORAS5-LR. Consistent with the similar sub-surface biases of ORAS5 and ORAS5-LR, SEAS5-LR exhibits a similar bias in lead month 1. The negative bias around the thermocline subsequently amplifies, but not as pronounced as in SEAS5. Also, the negative bias close to the surface is substantially reduced compared to SEAS5. The reduced surface impact of the cold thermocline bias in SEAS5-LR is consistent with the reduced SST bias and, as a result, a reduced u10m bias compared to SEAS5 (compare Figure 11b). Figure 12c shows the subsurface temperature bias of a low-resolution version of SEAS5 that is initialized from ORAS4, with the same coupled model cycle used in SEAS5 and SEAS5-LR. We call this experiment SEAS5-LR-O4. The negative bias with a maximum around the thermocline is also present for this experiment, but with a magnitude that is further reduced compared to SEAS5-LR. This is most visible in lead month 1, which is consistent with the warmer thermocline temperatures of ORAS4 compared to ORAS5-LR (compare Figure 7b and c). More importantly, in SEAS5-LR-O4 the cold bias does not reach the surface. This is an important difference to SEAS5 and also to SEAS5-LR, which both have the cold bias reaching the surface (at varying degrees). An important consequence is that the subsurface-errors in SEAS5-LR-O4 are fairly uncoupled from the atmosphere in summer, during the monsoon season. Figure 11b confirms the positive impact of ORAS4 initial conditions on surface biases, with the SST bias remaining positive also in JJA, despite the strengthening easterly bias. We conclude that all considered forecast experiments show a strong cold bias developing around the thermocline for May starts, and in all except for the one using ORAS4 initial conditions the cold bias reaches the surface by boreal summer and thus induces a cold SST bias as well. Although the difference in forecast biases between SEAS5-LR and SEAS5-LR-O4 is due only to the ocean initial conditions, since they use the same model setup in forecast mode, the impact of the ocean model version cannot be discarded. Indeed, the different ocean model versions used for the long reanalyses can impact the forecasts through the initial conditions. We also find sensitivity to the resolution, with different bias evolution in SEAS5 compared to SEAS5-LR, a fact that should be kept in mind for testing of newer cycles at low resolution (see also Roberts et al. 2020).

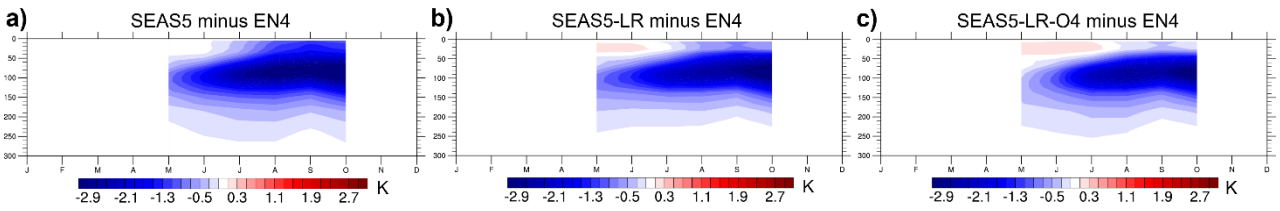

Figure 12. Subsurface ocean temperature bias in EEIO in long-range coupled forecasts from (a) SEAS5, (b) SEAS5-LR, (c) and SEAS5-LR-O4 for lead months 1-6 (May starts 1993-2015). Hadley EN4 analyses are used as reference.

5. State-dependence of errors

In the previous sections we assessed unconditional biases as a function of lead time. To identify conditional biases, here we examine flow-dependent relationships between SST, winds, and precipitation. We focus on JJA, when the SST bias is not yet at its peak (compare Figure 2), to explore relationships and feedbacks during the development phase of the errors.

5.1. Local atmospheric response in the EEIO to SSTs

Figure 13 shows scatter diagrams of SST_{EEIO} versus (first row) zonal wind and (second row) precipitation in the EEIO for forecasts and ERA5. The SST/wind relationship (Figure 13a and b) exhibits a non-linear behavior, with two different regimes for SST above and below a certain threshold T_c (~28.5°C): the sensitivity of wind to SST variations is higher (steeper slope) in the warm regime.

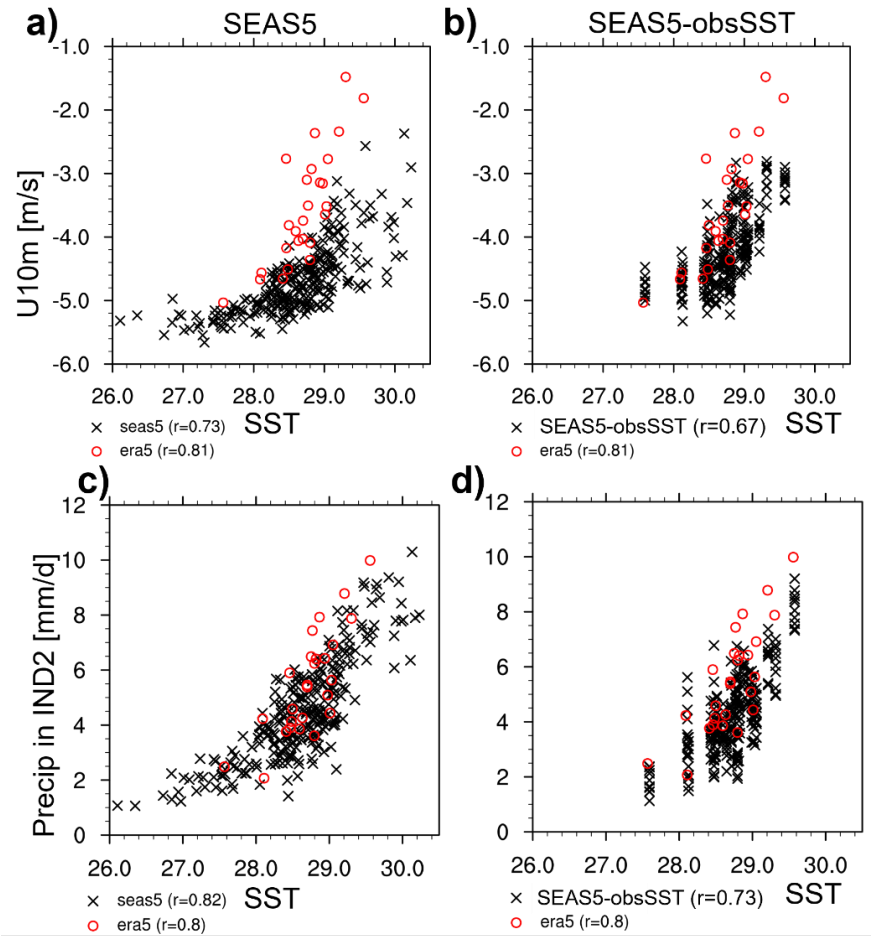

Figure 13. Scatter diagrams of JJA SSTs in EEIO against (a,b) 10m u-wind and (c,d) precipitation in JJA in EEIO. Black crosses show single members (10 per start date) of (a,c) SEAS5 and (b,d) SEAS5-obsSST forecasts initialized on 1 May 1993-2018 and red circles relationships based on ERA5 1993-2018.

Since the winds and SSTs are coupled two-ways, the flattening of the u10m/SST relationship in the cold regime can also be interpreted as an increased SST sensitivity to wind variations. In the observations (Figure 13a), most of the points are in the warm regime, showing a strong reduction of the prevailing easterlies as a response to warmer SSTs. SEAS5 forecasts exhibit three major differences w.r.t. reanalyses. First, for a given SST, the easterlies are too strong. Second, the easterly wind bias increases with SSTs (warm regime slope error). Third, SEAS5 SST_{EEIO} is more populated in the cold regime, with a sizable fraction of forecasts reaching much cooler

values than observed. The SST/wind relationship in SEAS5-obsSST (Figure 13b) is somewhat improved compared to SEAS5, with forecasts lying closer to ERA5 for SSTs < T_c. However, at higher SSTs the warm regime slope error appears similar in coupled and uncoupled forecasts.

The observed and modelled SST/precipitation relationships are closer to observations (compared to the u10m/SST relationship) for both coupled and uncoupled forecasts (Figure 13c, d). Both forecasting systems exhibit a slight underestimation of precipitation at high SSTs, where SEAS5 exhibits weaker underestimation of precipitation at high SSTs (-0.8 mm/d for all cases with SST> T_c and -1.1 mm/d for all cases with SST>T_c and lower the observed maximum SST) than SEAS5-obsSST (-1.5 mm/d for all cases with SST>T_c). For the cool SSTs only present in SEAS5, the SST/precipitation relationship weakens as precipitation goes towards zero.

5.2. EEIO atmospheric response to zonal SST gradients

Next, we look at relationships of u10m and precipitation in the EEIO with the zonal SST gradient across the Indian Ocean, approximated by the difference SST_{EEIO}-SST_{WEIO} (ΔSST). Note that ΔSST differs from the Indian Ocean Dipole Model Index as the latter uses standardized SST anomalies and has the sign reversed. Figure 14 is analogous to Figure 13, but now with ΔSST on the x-axis. It stands out that the observed zonal SST gradient is always positive and ranges between 0 and 2K, i.e., SSTs in the EEIO are always warmer than those in the WEIO. In SEAS5 (left column of Figure 14), the SST gradient can attain large negative values lower than -2K. For the cases where SEAS5 has a positive zonal SST gradient, the ΔSST/u10m relationship in SEAS5 (Figure 14a) agrees reasonably well with ERA5, although the weakening of the winds in response to a positive SST gradient still seems underestimated in terms of slope. Hence, it appears as if the model captures comparatively well the response to the positive zonal SST gradients in spite of the deficient local response over the EEIO. This can be understood by noting the positive SST bias in the WEIO in JJA (compare Figure 2): For a given positive SST_{EEIO} anomaly in SEAS5, ΔSST is smaller compared to observations because of the warm SST_{WEIO} bias. This is confirmed by Figure A2 which shows the relationship between SST_{EEIO} and ΔSST in SEAS5 and ERA5. We conclude that the warm regime slope error of SEAS5 (too weak zonal wind response to SST_{EEIO}; Figure 14a) at least partly results from the warm SST_{WEIO} bias and the fact that the zonal winds are more strongly related to ΔSST rather than SST_{EEIO}. However, from Figure 14a it is also evident that cases with more than moderately positive ΔSST are relatively rare in SEAS5, as the majority of SEAS5 forecasts cluster around neutral SST gradients, which is rarely observed (not to mention the occurrences of negative ΔSST).

The too weak wind response in SEAS5-obsSST for ΔSST>1K may be related to the too weak precipitation response for high SSTs in the EEIO (Figure 13d), which conversely provides too weak diabatic heating and low-level wind convergence. Another aspect is the positive precipitation bias of SEAS5-obsSST over southeast Asia [see Stockdale et al. (2018) or also Lavers et al. (2021)], which is associated with strengthened equatorial easterlies through the enhanced anticyclonic atmospheric circulation over the northern Indian Ocean. Furthermore, the observed anti-correlation between SST_{EEIO} and southeast-Asian precipitation (i.e., higher SST_{EEIO} goes with reduced precipitation over southeast Asia and vice versa in reanalysis) is not reproduced by SEAS5-obsSST (not shown), i.e., the precipitation bias and associated circulation response is especially strong for high SST_{EEIO}, which likely contributes to the conditional easterly wind bias of SEAS5-obsSST in the EEIO. In addition to this, the too weak slope of the ΔSST/wind relationship of SEAS5-obsSST may also indicate a persistent and stable easterly regime which is difficult to destabilize by either local convection or large-scale SST gradients.

SEAS5-obsSST (Figure 14b) exhibits a similar ΔSST/u10m relationship as SEAS5, with reduced wind spread at SST gradients >1K. In the reanalyses, the ΔSST/u10m relationship has a stronger linear slope for positive SST gradients, i.e., the model winds do not weaken sufficiently with stronger positive SST gradients.

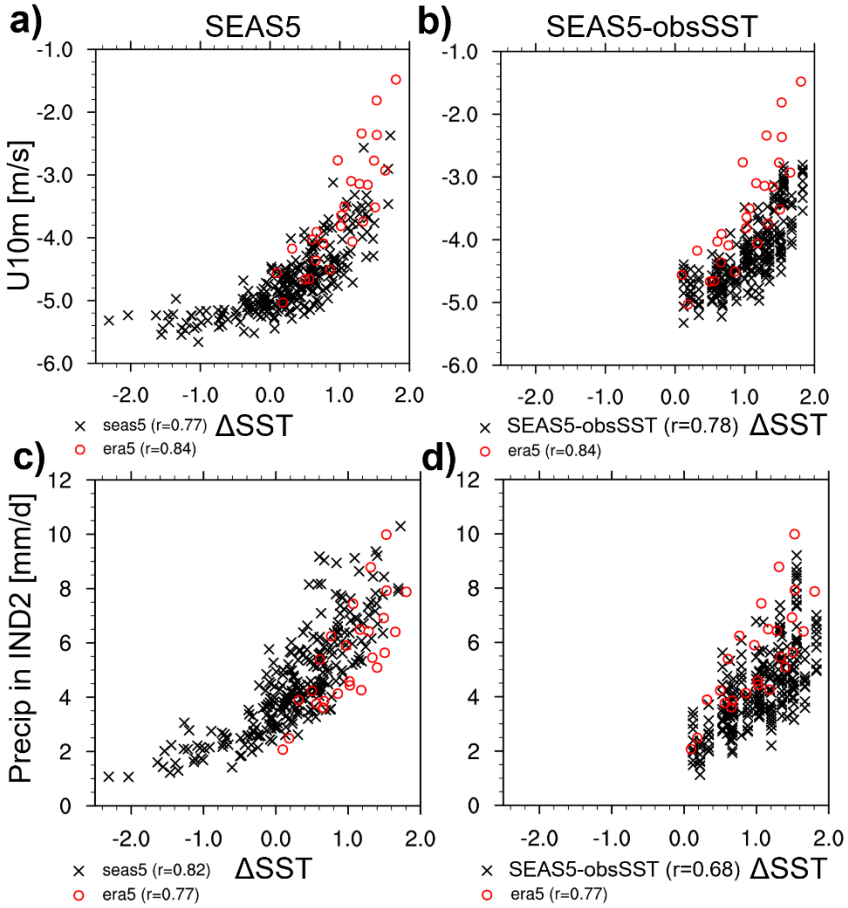

Figure 14. Scatter diagrams of JJA west-east SST gradients in the Indian Ocean against (a,b) 10m u-wind and (c,d) precipitation in JJA in EEIO. Black crosses show single members (10 per start date) of (a,c) SEAS5 and (b,d) SEAS5-obsSST forecasts on 1 May 1993-2018 and red circles relationships based on ERA5 1993-2018.

The ΔSST/precipitation relationship in SEAS5 (Figure 14c) is degraded compared to the local SST/precipitation relationship (compare Figure 13c), with too high precipitation for a given SST gradient. This can be explained by the fact that SEAS5 only produces significant positive zonal SST gradients when SST_{EEIO} attains overly high SSTs (see Figure A2). In these cases, the SST gradient may be corrected by the warm WEIO SST bias, but precipitation responds more strongly to local SSTs and hence is too high for a given SST gradient. The ΔSST/precipitation relationship in SEAS5-obsSST (Figure 14d) shows an underestimation of precipitation for strongly positive SST gradients. This result is in accordance with the found underestimation of precipitation to local SST_{EEIO} (Figure 13d).

We do not show these diagnostics for extended-range forecasts for the sake of brevity, but we note that the main features found for the seasonal range can be seen already in weeks 2-4 of the forecasts.

5.3. Relationships between SSTs and sub-surface temperatures

While the previous sections assessed the relationships between SSTs and atmospheric quantities, we now turn to the relationship between SSTs and sub-surface temperatures in the EEIO. Figure 15 shows scatter diagrams of SSTs versus temperatures in 100m depth (T100) in EEIO. T100 can be interpreted as a proxy for thermocline depth, with high values indicating a deep thermocline and vice versa.

ORAS5 (Figure 15a) expectedly shows a positive correlation between SSTs and T100 in EEIO. Again, the scatter diagram shows the existence of two regimes. The slope is steeper in the warm regime (for SSTs>T_c) with strong T100 variations (~21.5 to 24.5°C) compared to a relatively small range in SSTs (~28.5 to 29.5°C), indicative of a strong vertical temperature gradient around 100m, associated with a deep thermocline in the warm regime. In the cold SST regime (SST < T_c), the slope is relatively flat with a smaller change in subsurface temperatures being associated with a larger change in SSTs. This is indicative of a shallower thermocline (maximum vertical gradient in shallower layers than 100m). We note that in ORAS5 the cold regime is a rare event and represented only by very few cases.

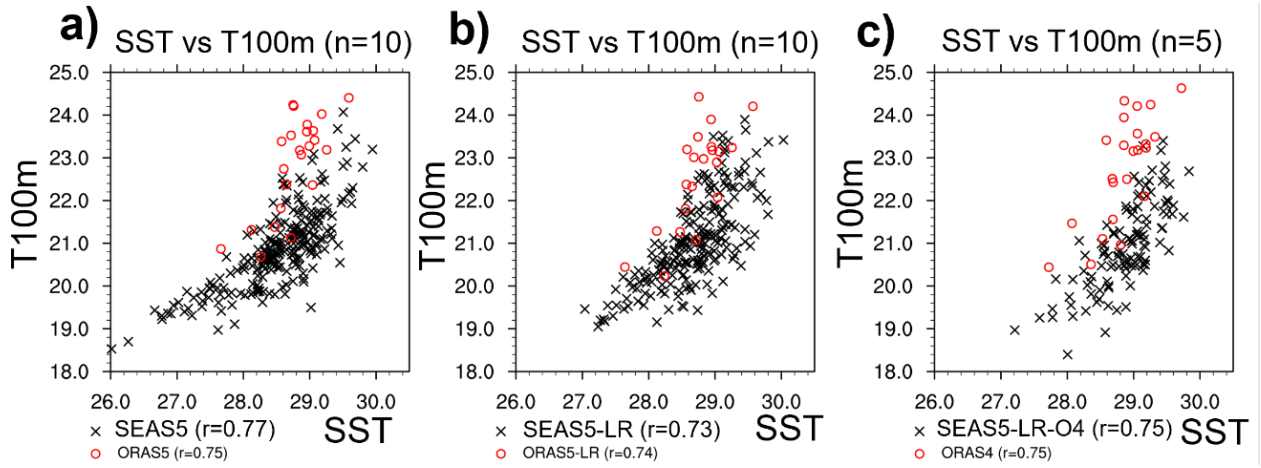

Figure 15. Scatter diagrams of JJA SSTs in EEIO against temperature in 100m depth in JJA in EEIO. Black crosses show single members of (a) SEAS5 (1993-2015) and (b) SEAS5-LR (1993-2013), and (c) SEAS5-LR-O4 (1993-2015) forecasts (May starts). Red circles show relationships based on the respective initial conditions of the forecasts: (a) ORAS5, (b) ORAS5-LR, and (c) ORAS4.

The steep slope in the warm regime can also be interpreted as little SST sensitivity to variations of a relatively deep thermocline, together with the fact that deep atmospheric convection prevents the growth of SST perturbations. Conversely, the flatter slope of the cold regime can be interpreted as a strong sensitivity of the SST to variations of a shallow thermocline, a sensitivity that is not capped due to the weaker local air-sea interaction in the cold regime

The T100/SST distribution based on SEAS5 forecasts (Figure 15a) is shifted vertically compared to ORAS5, i.e., T100 for a given SST is lower in SEAS5 compared to ORAS5. Nevertheless, SEAS5 exhibits a similar distinction of regimes around SST=T_c. However, in contrast to ORAS5, a large fraction of forecasts resides in the cold regime, and very cool T100 values are obtained by some members which have never been observed during the considered period. This is consistent with a shallowing error of the thermocline in the forecasts. The transition from the cold to warm regime appears to happen more gradually, with the T100/SST relationship in the warm regime of SEAS5 being not as steep as for ORAS5.

Figure 15b shows the T100/SST relationship for SEAS5-LR as well as ORAS5-LR (i.e., the corresponding ocean reanalysis used to initialize SEAS5-LR). ORAS5-LR behaves very similar to ORAS5 (Figure 15a), the

notable difference being the slightly cooler lowest two T100 values. SEAS5-LR has a similar cool T100 bias as SEAS5 but has a weaker tendency to go into the cold regime. The long tail for cold SSTs is not present in SEAS5-LR, and the relative frequency of high SSTs is higher compared to SEAS5 (not shown).

A further gradual change in behavior can be seen for SEAS5-LR-O4 (Figure 15c). SEAS5-LR-O4 exhibits a less pronounced distinction between warm and cold regime compared to SEAS5, i.e., the flattening of the T100/SST relationship for SST<T_c is less present, and cool T100 values below 19°C are reached for comparatively warm SSTs. Conversely, in SEAS5-LR-O4, low T100 values do not translate into SSTs as low as in SEAS5. This is consistent with the result that the unconditional cold subsurface bias in SEAS5-LR-O4 does not appear at the surface as prominently as in SEAS5 (compare Figure 12). Another salient difference of SEAS5-LR-O4 with SEAS5 are the increased frequencies of cases in the warm regime (not shown explicitly), but it has to be kept in mind that SEAS5-LR-O4 consists of only 5 ensemble members, which makes robust statements about changes in forecast distributions difficult.

6. Assessment of recent cycles and sensitivity tests

Here we apply the diagnostics developed in previous sections to more recent developments and cycles. The impact of ORAP6 initial conditions on subsurface biases in seasonal forecasts is slightly positive, and results are shown in the Appendix A3. The next two subsections assess the impact of updated moist physics (introduced in IFS CY47R3) and an interim ocean model version different from NEMO3.4, here termed NX.

6.1. Recent IFS cycles

The wind/SST bias has not improved in recent model cycles compared with the performance of CY43R1 (used in SEAS5). In particular, CY47R1 saw a clear degradation of the biases (see FC_47R1 results in Figure 11c). The moist physics changes in CY47R3 have been evaluated with the seasonal reforecast experiment FC_47R3 (also shown in Figure 11c). In this experiment, SSTs in the EEIO are cooler from lead month 1 and u10m in the EEIO is more easterly from lead month 2 (compared to FC_47R1). While the SST cooling means a reduction of the initial warm bias in lead month 1, it increases the cold bias from month 2 onward. Further diagnostics showed that the cooling mainly arises from reduced downwelling shortwave radiation and enhanced evaporation in FC_47R3 compared to FC_47R1 (not shown). The increased easterly bias in FC_47R3 is likely associated with the cooler SSTs in the EEIO.

Inspection of the change in SST/u10m and SST/precipitation distributions from FC_47R1 to FC_47R3 (not shown) suggests that in FC_47R3 there is a decrease in relative frequency of strong easterlies with u<4m/s but an increase of weak easterlies for warm SSTs >T_c. This represents an improvement compared to SEAS5, as Figure 13a clearly shows the over-representation of cases with warm SSTs and strong easterlies. Indeed, the conditional wind bias for SSTs >29°C is reduced by ~0.1m/s in FC_47R3 compared to FC_47R1 (-1.1 vs 1.2 m/s; not shown). However, the effect of the increased frequency of SSTs<T_c in FC_47R3 (which are already overrepresented in SEAS5) clearly dominates, as the unconditional biases in Figure 11c show. The SST/precipitation relationship does not change in FC_47R3 compared to FC_47R1 (not shown). Thus, the changes in the moist physics in FC47R3 do not show a direct impact on the precipitation response to local SSTs in the EEIO. We also assessed a seasonal forecasting experiment using IFS CY47R2 (FC_47R2; biases shown in Figure 11c as well), which behaves more similarly to FC_47R1 than FC_47R3. This confirms the above conclusion that the changes seen in FC_47R3 indeed result from the moist physics changes introduced with that cycle.

6.2. Role of the ocean model

In this section we investigate the relative role of the ocean model in the EEIO errors, trying to separate the impact from the initial conditions and that of the model formulation during the forecasts. The method used to address this question is somewhat opportunistic, taking advantage of existing ocean simulations and reforecast experiments with different ocean model formulation, which turned out to exhibit different behaviour in the forecasts of the EEIO.

The benchmark reforecast experiment is FC_N3.4_LR, which uses NEMO3.4 at low resolution, and it is initialized by an ocean-only simulation (i.e., no ocean data assimilation), where the SST has been constrained via a strong relaxation. FC_N3.4_LR exhibits a similar cold subsurface bias as SEAS5-LR (not shown) and also develops a cold SST and easterly u10m bias in the EEIO (Figure 11c). The reforecast experiment of opportunity is called FC_NX_LR, which is similar to FC_N3.4_LR but uses a different formulation of the ocean model (NX) during the forecast phase and in the production of the ocean initial conditions. FC_NX_LR shows warmer subsurface temperatures (i.e., a reduced subsurface bias) compared to the reference FC_N3.4_LR (Figure 16*a*) as well as a decreased SST bias in EEIO (not shown). Further inspection reveals that the ocean initial conditions produced with version NX show a warmer mixed layer and deeper thermocline in this region compared with those produced by the standard NEMO3.4 (not shown). However, it has to be noted that the ocean-only simulation using NX used a different subsurface climatology for the nudging than the simulation using NEMO3.4 (Sarah Keeley, personal communication), which could partially explain the differences in the initial conditions.

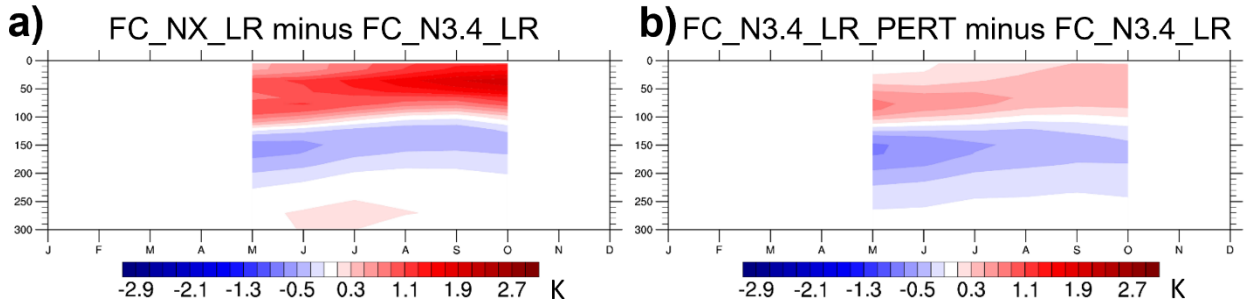

Figure 16. Subsurface temperature differences in EEIO as a function of lead time (May start dates 1993-2015) between a) FC_NX_LR and FC_N3.4_LR and b) FC_N3.4_LR_PERT and FC_N3.4_LR.

*Figure 12*We assess the sensitivity of seasonal forecasts using NEMO3.4 to a warm subsurface perturbation. For this purpose, we compute the temperature difference between the initial conditions produced with NX and NEMO3.4. We applied this difference (in combination with a density-compensating salinity perturbation) over the EEIO in May to the NEMO3.4 initial conditions and carried out seasonal forecasts using these perturbed ICs (initialized 1st May 1993-2015; labelled FC_N3.4_LR_PERT). The difference between subsurface temperature in FC_N3.4_LR_PERT and FC_N3.4_LR is presented in Figure 16b. Differences in lead month 1 (May) are similar to the difference between FC_NX_LR and FC_N3.4_LR (compare Figure 16b with a), explaining about ~87% of the temperature difference between initial conditions of FC_N3.4_LR_PERT and FC_N3.4_LR at the thermocline (80m depth). At longer lead times, the differences decrease, explaining only ~43% of the thermocline initial condition perturbation in 80m by lead month 6. The surface impact of the perturbation in FC_N3.4_LR_PERT is small, with slightly reduced SST bias and hardly changed u10m bias in the EEIO in JJA (Figure 11c). This result suggests that while the memory of subsurface perturbations in the initial conditions lasts into the seasonal forecast, they fail to amplify; in fact, the initial perturbation is damped as the forecast progresses. This implies that the formulation of the ocean model (most likely vertical processes)

plays a role in the growth of the initial condition error. However, it should be kept in mind that the perturbation has been applied only to the relatively small EEIO region, and lateral advection and mixing might play a role in the damping of the initial perturbation as well.

7. Summary and conclusions

This paper summarizes the work carried out within UGROW-IO, which assessed the lead-time-dependent evolution of atmospheric and oceanic forecast errors in the Indian Ocean, most notably the cold SST and easterly wind bias in the EEIO. Analysis reveals that errors are present already in the initial conditions to some extent. Comparison of winds from atmospheric reanalyses to in-situ-observations in the EEIO showed that there is a weak easterly bias at the equator during April-October. The ORAS5 ocean reanalysis exhibits a too cold thermocline in the EEIO, which may be related to the identified easterly wind bias in ERA-Interim. Moreover, the subsurface temperature bias in ORAS5 is increased compared to ORAS4 despite the same atmospheric forcing. Oceanic resolution has been found to play a relatively small role in this context, which suggests that the different versions of the NEMO ocean model (NEMO3.4 compared to 3.0) along with differences in assimilated data and assimilation methods leads to the increased bias in ORAS5.

The easterly wind bias increases rapidly within the first forecast days, suggesting that the atmospheric model establishes strong easterlies in the EEIO (and other tropical regions) in the absence of sufficient observational constraints. A similar behaviour is seen for a range of atmospheric models participating in the DIMOSIC project. At subseasonal-to-seasonal lead times, the easterly wind bias further strengthens and leads to development of the cold SST bias in the EEIO through enhanced upwelling.

Investigation of the state-dependence of atmospheric errors shows that warmer SSTs are associated with overly strong easterlies, even in uncoupled experiments, while cooler SSTs have smaller wind biases. SEAS5 forecasts develop a warm SST bias in the WEIO which likely further enhances the easterly bias in the coupled runs. Local precipitation response to SST anomalies in the EEIO is underestimated when compared to reanalyses, with a stronger underestimation by the uncoupled compared to the coupled forecasts. Considering relationships of winds in the EEIO with zonal SST gradients along the equatorial Indian Ocean yields better agreement of the forecasts with reanalyses, although easterlies are still too strong for high SSTs. For the coupled runs, the cause for the improved relationship is likely the removal of the effect of the warm WEIO bias by considering the zonal SST gradient. A potential cause for the remaining bias is too weak local diabatic heating and associated wind response in the EEIO arising from the underestimation of the local precipitation response. For the uncoupled forecasts, overestimation of the anticyclonic circulation over the northern subtropical associated with the too strong southeast Asian monsoon may play an additional role.

Diagnostics of the relationship between subsurface ocean temperatures and SSTs in the EEIO show that there are two distinct regimes: the warm regime with deep thermocline and weak sensitivity of SSTs to winds and a cold regime with a shallow thermocline and a strong sensitivity of SSTs to winds. Coupled forecasts predominantly reside in the cold regime (opposite to observations), in which SST errors amplify quickly in the presence of wind errors.

The diagnostics developed under UGROW-IO were subsequently applied to ongoing system developments on atmospheric and ocean model versions and ocean data assimilation. The analysis of system developments, assisted by targeted numerical experimentation, has helped us to gain further insight into the nature of the errors in the EEIO. The lessons learnt can be summarized as follows:

- *Recent atmospheric model cycles:* Wind and SST bias in the EEIO have not been improved in recent model cycles when compared to SEAS5 (using CY43R1). In direct comparison to CY47R1, the wind-SST relationship in the convective regime is slightly improved in seasonal forecasts with CY47R3, but overall biases (SST and wind) in the EEIO increased, likely because in CY47R3 reduction of net surface heat flux into the ocean leads to an overall cooling of SST. Although the change in SSTs represents a reduction of the existing positive bias in lead month 1, it represents an increase of the existing negative SST bias from lead month 2 onward.
- *Ocean data assimilation developments in the path towards ORAS6*: We have evaluated the pilot ocean reanalyses (the so-called ORAP6 family), which use surface forcing based on hourly ERA5 and updated observational datasets. Changing ocean initial conditions from ORAS5 to ORAP6.1 improves subsurface biases only slightly. This suggests that the change to ERA5 forcing does not improve the cold thermocline bias in ocean reanalyses significantly despite the improvement of ERA5 10m-u-winds over ERA-Interim in the comparison to winds measured by RAMA buoys.
- *Ocean model developments in the path towards NEMO4:* ocean simulations (control runs) using an interim version of NEMO (NEMOX) showed a deeper and warmer mixed layer in the EEIO and a tighter thermocline. The comparatively warmer mixed layer in the initial conditions persists well into the forecasts conducted with this interim version of the ocean model, thus reducing the cold SST bias in the EEIO. We note though that these results may not be applicable to newer versions of the ocean model proposed for operational implementation.
- *Numerical experimentation on perturbation growth:* A seasonal forecast experiment using NEMO3.4 and modified initial conditions with a warm perturbation added in the EEIO (to mimic the warmer mixed layer in this region from an ocean simulation with modified ocean model version) shows that the initial perturbation is damped out relatively quickly within the first 3 months of the forecasts. This contrasts with the previous results using the modified ocean version in the forecast model, suggesting that, within a given atmospheric model version, the EEIO cold SST forecast bias does not only depend on the ocean initial conditions, but also depends on the formulation of the ocean model, which can damp or amplify the initial state.

We conclude that there are two fundamental and independent sources of errors that lead to the SST errors in seasonal forecast. The first one is of atmospheric nature and is largely related with too stable easterly circulation present over the whole equatorial Indian Ocean, characterized by the lack of response of the local winds to local surface heating in the EEIO. This induces an easterly bias which leaves the model predominantly in a state with a shallow thermocline and cold SSTs in the EEIO. Circulation errors arising from precipitation biases in the southeast Asian monsoon regions play a role as well. The second error is of oceanic origin, associated with a too shallow thermocline, which enhances the SST errors arising from errors in the wind. The markedness of this error varies with ocean initial conditions, which depend on both the quality of the assimilation and the ocean model. The version of the ocean model used for the forecast also plays a non-negligible role at seasonal time scales, by amplifying or damping the subsurface errors in the initial conditions due to the strength of the atmosphere-ocean coupling in this region.

Acknowledgments

The authors thank colleagues involved with ocean model developments (Sarah Keeley, Kristian Mogensen, and Jean Bidlot) for provision of experimental data.

References

Balmaseda, M. A., K. Mogensen, and A. T. Weaver, 2012: Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, **139**, 1132–1161, https://doi.org/10.1002/qj.2063.

Belmonte Rivas, M., and A. Stoffelen, 2019: Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT. *Ocean Science*, **15**, 831–852.

Bidlot, J.-R., D. J. Holmes, P. A. Wittmann, R. Lalbeharry, and H. S. Chen, 2002: Intercomparison of the performance of operational ocean wave forecasting systems with buoy data. *Weather and forecasting*, **17**, 287–310.

Dee, D., et al., 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society* 137.656: 553-597.

Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans (1978–2012)*, **118**, 6704–6716, https://doi.org/10.1002/2013JC009067.

Hersbach, H., and Coauthors, 2020:TheERA5 global reanalysis. Quart. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803.

Johnson, S. J., and Coauthors, 2019: SEAS5: the new ECMWF seasonal forecast system. *Geoscientific Model Development*, **12**, 1087–1117.

Lavers, D. A., S. Harrigan, and C. Prudhomme, 2021: Precipitation biases in the ECMWF integrated forecasting system. *Journal of Hydrometeorology*, **22**, 1187–1198.

Magnusson, L., M. Alonso-Balmaseda, S. Corti, F. Molteni, and T. Stockdale, 2013: Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors. *Climate dynamics*, **41**, 2393–2409.

Mcphaden, M. J., and Coauthors, 2009: RAMA: the research moored array for African–Asian–Australian monsoon analysis and prediction. *Bulletin of the American Meteorological Society*, **90**, 459–480.

Roberts, C., and Coauthors, 2020: *Reduced-resolution ocean configurations for efficient testing with the ECMWF coupled model*. European Centre for Medium-Range Weather Forecasts,.

Rodwell, M. J., and Coauthors, 2010: Developments in diagnostics research. *ECMWF Technical Memorandum*, **637**.

Stockdale, T., and Coauthors, 2018: SEAS5 and the future evolution of the long-range forecast system. *ECMWF Tech Memo*, **835**.

Zuo, H., M. A. Balmaseda, S. Tietsche, K. Mogensen, and M. Mayer, 2019: The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment. *Ocean Sci.*, **15**, 779–808, https://doi.org/10.5194/os-15-779-2019.

Zuo, H., M. A. Balmaseda, E. de Boisseson, S. Tietsche, M. Mayer, and P. de Rosnay, 2021: The ORAP6 ocean and sea-ice reanalysis: description and evaluation, EGU General Assembly 2021, online, 19–30 Apr 2021, EGU21-9997, https://doi.org/10.5194/egusphere-egu21-9997, 2021.

Appendix

A1 Subsurface temperature biases in ORAP6

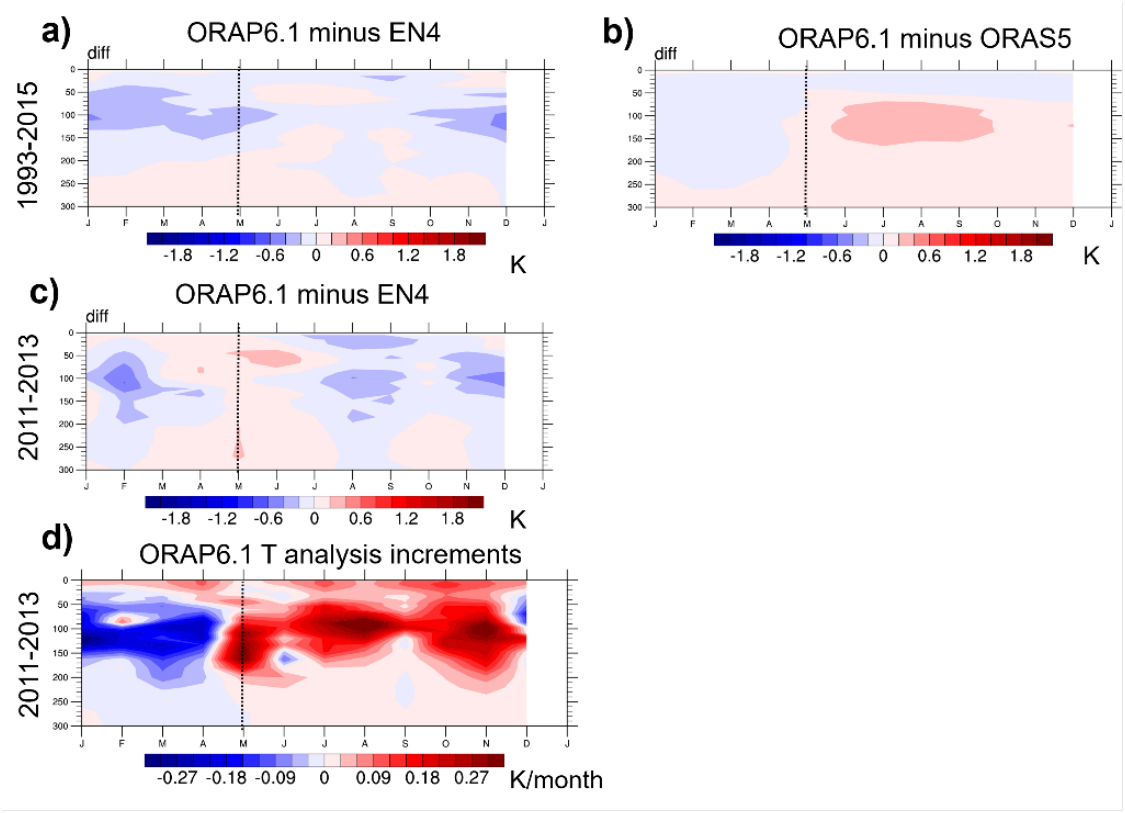

Figure A1. Mean annual cycle of subsurface temperature bias in EEIO of ORAP6.1 during (a) 1993-2015 and c) 2011-2013; b) Mean annual cycle of subsurface temperature differences between ORAP6.1 and ORAS5; d) Temperature increments in EEIO from ORAP6.1 during 2011-2013.

A2 Observed and modelled relationships between SST in EEIO and zonal SST gradient

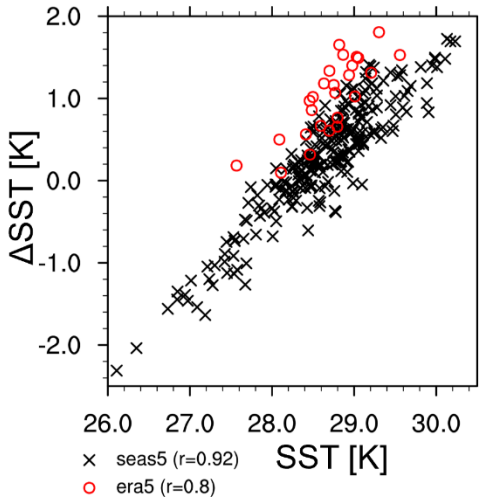

Figure A2. Scatter diagrams of JJA SSTs in the EEIO versus west-east SST gradients in the Indian Ocean. Black crosses show single members (10 per start date) of SEAS5 (May starts) and red circles relationships based on ERA5 1993-2018.

A3 Impact of ORAP6 initial conditions on seasonal forecasts

Figure A3 shows the impact of the updated ocean analysis system ORAP6 on seasonal forecasts by comparing the subsurface temperature forecast bias in seasonal forecasts initialized from ORAP6.1 (FC_ORAP6.1) and ORAS5 (FC_ORAS5), which use the same model setup (Tco199 L137 ORCA025 Z75) and only differ in their initial conditions. The reduced cool subsurface bias in ORAP6.1 is translated to a slight warming of subsurface temperatures in seasonal forecast experiments initialized from ORAP6.1 compared to initialization with ORAS5. This change represents a slight improvement given the strong cold subsurface forecast bias of SEAS5 (compare Figure 12a).

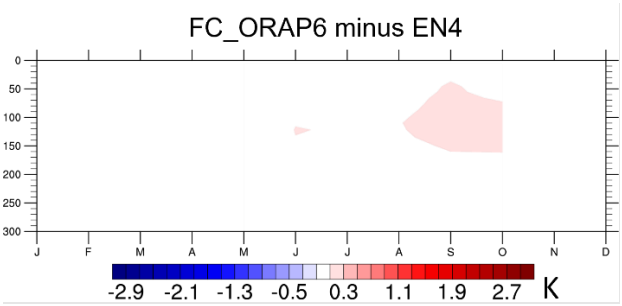

Figure A3. Difference between subsurface temperature bias in EEIO of hindcasts initialized from ORAP6.1 and an identical forecast experiment except for initialization with ORAS5 (May starts 1993-2015).