Coupling of surface air and sea surface temperatures in the CERA-20C reanalysis

RMetS
Royal Meteorological Society

Coupling of surface air and sea surface temperatures in the CERA-20C reanalysis

Xiangbo Feng,[a]* Keith Haines[a] and Eric de Boisséson[b]
[a]*Department of Meteorology, University of Reading, UK*
[b]*European Centre for Medium-Range Weather Forecasts, Reading, UK*

*Correspondence to: X. Feng, Department of Meteorology, University of Reading, Reading RG6 6BB, UK. E-mail: xiangbo.feng@reading.ac.uk

ECMWF has recently produced its first ocean-atmosphere coupled twentieth century reanalysis CERA-20C, with ten ensemble members. CERA-20C is based on a coupled climate model with data assimilation implemented individually for the atmosphere and ocean components. This article focuses on the relationships between 2 m air temperature (T2m) and sea surface temperature (SST) in this new coupled climate reanalysis from the perspective of ensemble statistics. The results show where the atmosphere or the ocean are driving the ensemble variability of temperature at the air–sea interface on varying time-scales.

In CERA-20C, the T2m–SST ensemble relationships are changing on diurnal, seasonal and longer time-scales and also within regions. The T2m ensemble spread is much larger than the SST ensemble spread at high frequencies (3 h), normally with low correlations due to different time-scales of oceanic and atmospheric dynamics, showing the atmospherically driven variability of T2m. On monthly time-scales, the T2m spread is mostly slightly lower than the SST spread, with high correlations and strong coupling, reflecting the SST-forced variability. T2m–SST coupling is strongest where atmospheric and oceanic boundary layers are shallow, e.g. in the summer hemisphere, and weakest where the atmospheric boundary layer becomes deep, e.g. in the ITCZ region. As the twentieth century progresses, ensemble spreads decline, while T2m–SST coupling becomes stronger as both atmosphere and oceans are better constrained to observations. In the Tropics, strong ENSO-related variability is found in T2m–SST coupling as atmospheric convection centres move interannually to break the SST-dominated relationships.

Key Words: CERA-20C; ensemble spread; SST; 2 m air temperature; air–sea interactions; coupled data assimilation; ENSO

Received 21 December 2016; Revised 23 October 2017; Accepted 25 October 2017; Published online in Wiley Online Library

1. Introduction

Large-scale geophysical data assimilation (DA) has mainly been developed within operational weather forecasting centres, with the initialization of numerical weather prediction (NWP) models being the main goal. NWP forecasts have rarely used an active ocean component over short time-scales, e.g. ECMWF have used land–atmosphere–wave models. However, with more attention being paid to medium range, monthly and seasonal forecasts, as well as to reanalyses, an initialized ocean model becomes necessary to allow the ocean surface to evolve more realistically. Seasonal forecasts, for example, usually use coupled atmosphere-ocean (AO) models starting from initial conditions produced with separate DA systems. An advantage of performing the atmospheric and oceanic DA separately is the ability to use different assimilation window lengths, appropriate to the dynamics of the medium, with atmospheric DA restricted to 6–24 h windows (Kalnay *et al.*, 1996; Rawlins *et al.*, 2007; Compo *et al.*, 2011; Dee *et al.*, 2011), while ocean DA typically uses 1 day–1 month windows (Carton and Giese, 2008; Balmaseda *et al.*, 2013; Martin *et al.*, 2015; Zuo *et al.*, 2015).

NWP centres have recently started to develop coupled DA systems to provide more consistent AO analysed states for forecast applications, with the most recent examples being the GFDL's ECDA (Zhang *et al.*, 2007), NCEP's CFSR (Saha *et al.*, 2010), Canadian CanSIPS (Merryfield *et al.*, 2013), UK Met Office's coupled DA system (Lea *et al.*, 2015) and ECMWF's CERA (Laloyaux *et al.*, 2016).* Such systems use a common assimilation window for both ocean and atmosphere, but still with separated

*GFDL = Geophysical Fluid Dynamics Laboratory; ECDA = Ensemble Coupled Data Assimilation; NCEP = National Centers for Environmental Prediction; CFSR = Climate Forecast System Reanalysis; CanSIPS = Canadian Seasonal to Interannual Prediction System; ECMWF = European Centre for Medium-range Weather Forecasts; CERA = Coupled ECMWF ReAnalysis system.

analysis of increments, and are usually called 'weakly' coupled DA as no cross-medium background-error covariances are used (Lu *et al.,* 2015a; Sluka *et al.,* 2016). In contrast, 'strongly' coupled DA would use cross-medium background-error covariances in the DA procedure to allow innovations in the atmosphere and ocean to directly interact. Smith *et al.* (2017) confirm that, in an idealized coupled model, significant cross error correlations mainly exist in the AO boundary layer and are characterized by diurnal and seasonal variations. More information may be gained from each observation near the air–sea interface if observational information can be spread through coupled background-error covariances. Strongly coupled DA has mainly been tested in idealized DA experiments (Zhang *et al.,* 2007; Han *et al.,* 2013; Tardif *et al.,* 2014; Smith *et al.,* 2015, 2017; Lu *et al.,* 2015a), although experiments with more realistic general circulation models have now begun (Lu *et al.,* 2015b; Sluka *et al.,* 2016).

A ten member ensemble AO coupled climate reanalysis over the twentieth century, known as CERA-20C, has recently been produced based on the CERA system (Laloyaux *et al.,* 2016). CERA uses a coupled land–atmosphere–wave–ocean–sea-ice system with DA applied individually for the atmosphere and ocean components, and with sea surface temperature (SST) constrained to the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST2) gridded ensemble product (Titchner and Rayner, 2014). CERA-20C is generated using an Ensemble of Data Assimilations (EDA) scheme in the atmosphere (Isaksen *et al.,* 2010; Bonavita *et al.,* 2016), and then perturbing the ocean observations in the ocean, following the method used in ECMWF's Ocean ReAnalysis System-5 (ORAS5; Zuo *et al.,* 2017). CERA-20C is the first twentieth century AO coupled climate reanalysis ensemble.

Following the definitions above, CERA is a 'weakly' coupled DA system, but the atmosphere and ocean components interact at the outer-loop level of the DA process, e.g. their increments interact through iteration of the coupled model (Laloyaux *et al.,* 2016). This article examines the AO coupling at the interface in this new coupled climate reanalysis by focussing on the relationships between the 2 m air temperature (T2m) and SST ensemble members. Ensemble spreads (standard deviations of the ensemble anomalies) and ensemble correlations between T2m and SST ensemble members, at varying time-scales, are diagnosed and interpreted. The T2m–SST relationships within the ensemble members are established via a two-way control exerted by the coupling processes at the air–sea interface. The study will consider whether the atmosphere or the ocean is dominating the ensemble variability of temperature at the air–sea interface, on different time-scales.

The article is structured as follows. In section 2, CERA and CERA-20C are introduced. In section 3, the ensemble spreads in T2m and SST and their relationships are evaluated on sub-daily to interannual time-scales, where the diurnal and seasonal cycles are also diagnosed. Long-term variations of the ensemble spreads and their relationships are also examined, both globally and regionally, demonstrating sensitivities to the evolution of the assimilated observations throughout the twentieth century. In section 4, the origin of the SST spread is further discussed as the main source for driving the T2m monthly variability. Conclusions and discussions are provided in section 5.

2. CERA system and CERA-20C

The CERA system was initially built under a European Space Agency project (Dee *et al.,* 2014; Laloyaux *et al.,* 2014) and has been further developed under the EU project ERA-CLIM2, based on a land–atmosphere–wave–ocean–sea-ice coupled system. The coupled system includes (i) the Integrated Forecast System (IFS) with the atmosphere model (ECMWF, 2016), the Hydrology Tiled ECMWF Scheme of Surface Exchanges over Land (HTESSEL) land surface model (Balsamo *et al.,* 2009), and the WAve Model (WAM) for waves (Komen *et al.,* 1994), and (ii) the Nucleus for European Modelling of the Ocean (NEMO) model (Madec, 2008) integrated with the Louvain-la-Neuve sea Ice Model (LIM2; Fichefet and Maqueda, 1997; Bouillon *et al.,* 2009). The coupling frequency in CERA is 1 h. The resolution of the atmospheric model is set to T159/L91 (IFS cycle 41R2), which corresponds to a 1.125° (125 km) horizontal grid with 91 vertical levels between the surface and 0.01 hPa. The ocean model (NEMO version 3.4) uses the ORCA1 grid, which has roughly a 1° (110 km) horizontal resolution with meridional refinement at the Equator. The ocean model has 42 vertical levels from the surface down to 5350 m, with a first layer thickness of 10 m.

The atmosphere and ocean components of CERA assimilate observations using incremental variational schemes (4D-Var in the atmosphere and 3D-Var in the ocean), while the land, ocean wave and sea ice models are running freely without constraints from observations. In the atmosphere, an EDA scheme is used to produce flow-dependent estimates of background-error covariance (Isaksen *et al.,* 2010; Bonavita *et al.,* 2016). The EDA scheme in CERA uses the ERA-20C method for atmospheric component of background-error covariance (Poli *et al.,* 2013), based on ten members of a lower-resolution (T95) IFS model run above the coupled SSTs as surface boundary conditions. For the ocean, the background errors are parametrized in proportion to the vertical gradient of the background temperature from each member (Mirouze and Weaver, 2010). The atmosphere observations and model physical tendencies are perturbed to generate the atmosphere ensemble in CERA, while the positions of the *in situ* observations are perturbed to generate the ocean ensemble, following the method used for ECMWF's Ocean ReAnalysis System-5 (ORAS5; Zuo *et al.,* 2017).

In CERA, each member implements 4D-Var and 3D-Var to produce the atmospheric and oceanic increments respectively, using a common 24 h DA window. The increments are added into the coupled states entirely at the start of each DA window. In CERA, the coupling is introduced at the outer-loop level of the variational method, which means that air–sea interactions start being taken into account when the increments have been applied to the initial conditions. In CERA, two iterations (two minimizations and two analysis steps) are performed on each DA cycle, allowing the oceanic (atmospheric) increments to influence the atmospheric (oceanic) components (Laloyaux *et al.,* 2016).

The twentieth century ten-member ensemble reanalysis using the CERA system is known as CERA-20C, and we will use these distinct terms throughout the article. The product is produced from 14 individual streams, each of 10 years (2 year overlap). Each stream is initialized from the uncoupled reanalyses ERA-20C (Poli *et al.,* 2013) and ORA-20C (de Boisséson *et al.,* 2017) for atmosphere and ocean respectively. CERA-20C covers the 1900–2010 period with a temporal resolution of 3 h, and with daily and monthly means also available.

The atmosphere model assimilates only surface pressure and surface marine wind observations from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS; Woodruff *et al.,* 2011) and the International Surface Pressure Databank (ISPD; Cram *et al.,* 2015). The ocean model assimilates sub-surface ocean temperature and salinity profiles from the EN4.02 observational dataset produced by the Met Office Hadley Centre (Good *et al.,* 2013) with depth bias correction of mechanical bathythermograph and expendable bathythermograph profiles (Gouretski and Reseghetti, 2010).

In CERA-20C, the SST represents the temperature of the ocean model top layer (first 10 m). In the analysis process, the ten member ensemble of SSTs are constrained towards the ten realizations of a daily version of the monthly gridded ensemble product HadISST2 (Titchner and Rayner, 2014), in which the ensemble spread is designed to reflect the uncertainties of observational sources and bias adjustments (Kennedy *et al.,* 2011). Please note that where night-time satellite data are used in HadISST products, they have been adjusted to a daily mean using

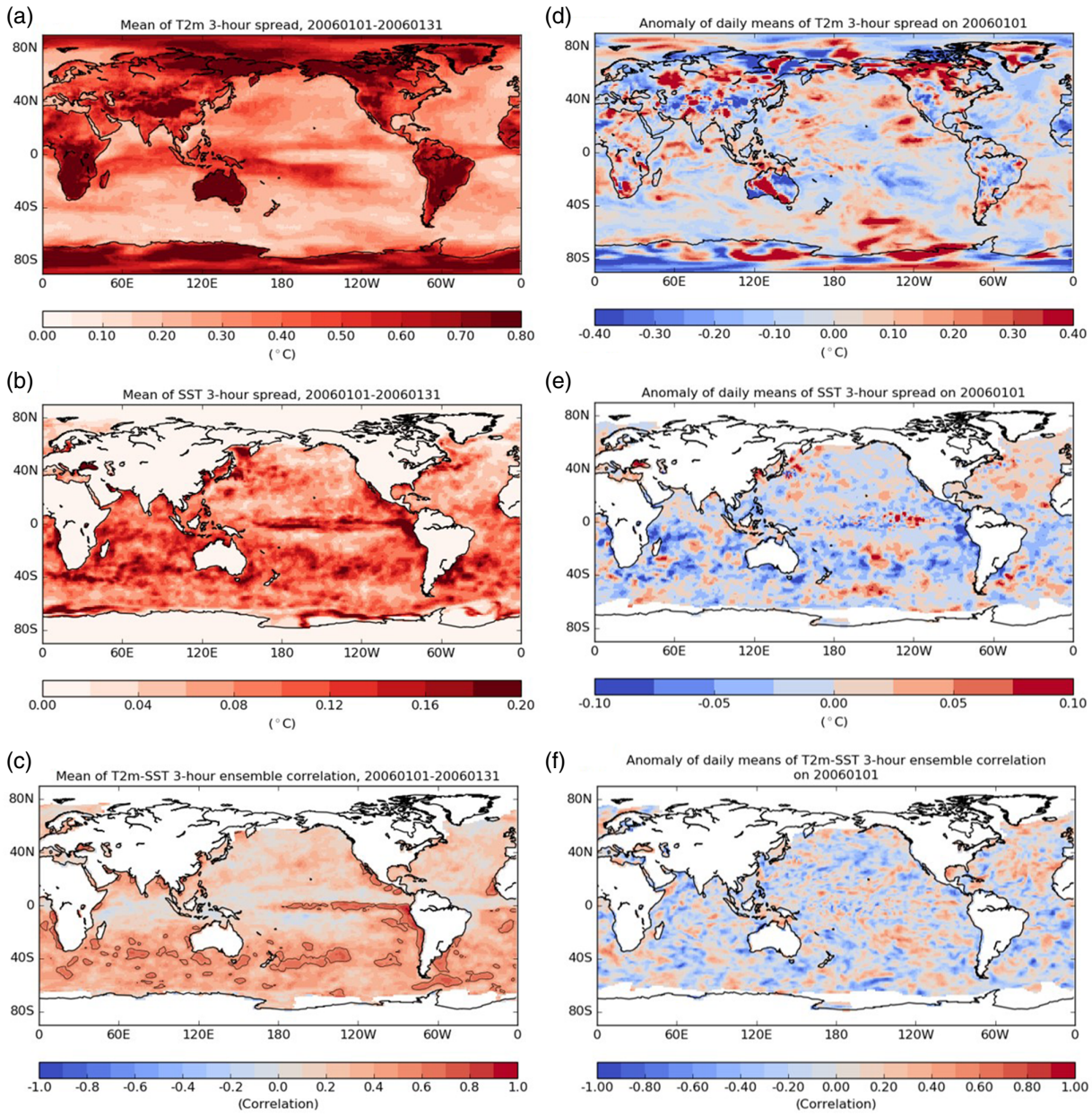

Figure 1. Means of 3 h ensemble spreads for (a) T2m and (b) SST, and (c) means of 3-h ensemble correlations between T2m and SST, for 1–31 January 2006. Daily means of 3-h ensemble spreads for (d) T2m and (e) SST, and daily means of (f) T2m–SSTensemble correlations on 1 January 2006, after removing the monthly means shown in (a)–(c). Note the different scales for T2m and SST values. Black contour in (c) shows values of 0.55 (90% confidence level). Correlation means are calculated using the Fisher transformation. SST values on sea ice-covered areas are excluded. [Colour figure can be viewed at wileyonlinelibrary.com].

a bias-correction based on *in situ* SSTs for all hours (Rayner *et al.*, 2003). A Newtonian relaxation scheme, developed in the ocean reanalysis system ORAS4 (Balmaseda *et al.*, 2013), is applied as a flux correction to the ocean top layer with a relaxation coefficient of $-200\,\mathrm{W\,m^{-2}\,K^{-1}}$, which corresponds to ~2–3 day relaxation time-scale. Because the depth of the ocean top layer (10 m) is not shallow enough to represent the amplitude of the SST diurnal cycle as seen by the atmosphere, a skin SST (SKT) scheme with a warm layer of 3 m depth is applied within the atmospheric component of the coupled system to capture the SST diurnal cycle (Zeng and Beljaars, 2005; Takaya *et al.*, 2010). In the analysis of high-frequency variability (section 3.1), both SST and SKT are assessed.

3. Relationships between T2m and SST ensemble spreads

3.1. *High-frequency variations*

3.1.1. *Three-hour means*

High frequency characteristics of the surface variables are diagnosed by using 3 h fields each day from 0000 to 2100 UTC throughout January 2006. Ensemble spreads and correlations are calculated every 3 h and then averaged over the month. Figures 1(a)–(c) shows the average spreads in T2m and SST, and the average correlations between them. Figures 1(d)–(f) shows daily mean anomalies with respect to the monthly means on 1 January 2006, where a daily mean is used because later we will see that a diurnal cycle also exists.

T2m has much larger spread over land (usually over 0.5 °C) than over oceans (Figure 1(a)). The smallest T2m spread (less than 0.2 °C) is found in the Southern Ocean and the eastern South Pacific and South Atlantic, where it is local summer. The largest T2m spread over the oceans (over 0.3 °C) is along the Intertropical Convergence Zone (ITCZ), the South Pacific Convergence Zone (SPCZ), and over ocean western boundary currents and fronts. Strong atmospheric instability, which can be demonstrated by the large sea–air temperature differences (SST minus T2m ensemble mean) at the surface (Cayan, 1980), is seen in these regions (Figure S1(a) in supplementary material). The T2m spread shows small-scale daily weather-related signatures of up to 0.3 °C, especially in midlatitudes where storms are frequent (Figure 1(d)).

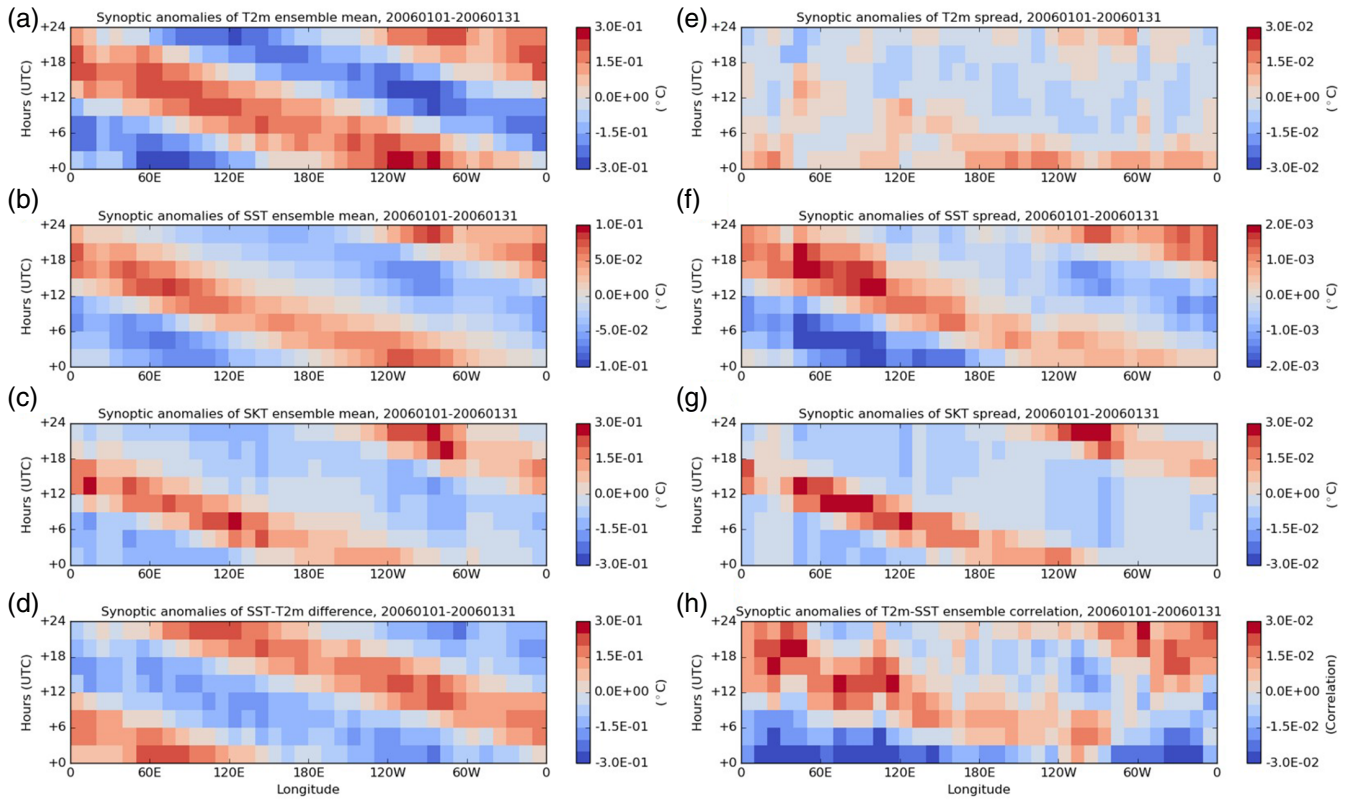

Figure 2. Hovmöller diagrams of synoptic anomalies of ensemble means for (a) T2m (over oceans only), (b) SST, (c) SKT, and (d) synoptic anomalies of SST−T2m difference, against times 0000, 0300 ... 2100 UTC through January 2006. Equivalent anomalies of ensemble spreads for (e) T2m, (f) SST and (g) SKT, and (h) anomalies of T2m−SST ensemble correlations. Values are averaged between 60°N and 60°S within 10° bins of longitude, with land and sea ice-covered areas excluded. Note the different scales in different panels. [Colour figure can be viewed at wileyonlinelibrary.com].

The SST spread is much smaller than the T2m spread over the oceans, and their spatial features are very different (Figure 1(b); note the different colourscales between (a) and (b)), indicating that the 3-h variability in T2m is also atmospherically driven. The largest values of the SST spread (over 0.12 °C) are in midlatitudes of the summer hemisphere and in upwelling regions where a shallow mixed-layer depth (MLD) (Figure S1(b)) allows SST to respond quickly (with large spread) to air−sea flux anomalies. Small-scale daily anomalies are also seen in the SST spread (Figure 1(e)), with largest absolute values where monthly mean values are also large. The SST spread is generated in the coupled ocean model but is also constrained by the SST relaxation scheme. (This will be discussed in section 4.)

Correlations between T2m and SST ensemble members are strongest (over 0.55) in regions of large SST spread (Figure 1(c)), where turbulent air−sea fluxes dominate the T2m variability under a stable atmospheric boundary (Figure S1(a)). In contrast, the weakest correlations are associated with instability in the atmospheric boundary in tropical precipitating regions, such as the ITCZ and SPCZ where the T2m spread is also very large. We show later that precipitation corresponds to cooler near-surface air temperatures, due to the re-evaporation of precipitation at lower levels as implemented in IFS (ECMWF, 2016). This weakens any relationship with the SST variability (this will be discussed in details in section 3.2). Correlation anomalies on daily time-scales (Figure 1(f)) associated with atmospheric weather-scale variations are generally weak. In general, higher coupled ensemble correlations on short time-scales can only be found under calm atmospheric conditions, with shallow atmospheric and oceanic boundary layers, associated with regions of smaller T2m spread and larger SST spread, where the reduced atmospheric variability allows SST and T2m to become coupled.

3.1.2. *Diurnal cycle*

The 3-h frequency in the CERA-20C output allows us to analyse the diurnal cycle of the ensemble mean and spread. The average diurnal cycle of the January 2006 data between 60°N and 60°S is obtained by averaging values at times 0000, 0300, ..., 2100 UTC, and then removing the daily mean. Figure 2 shows a longitudinal Hovmöller diagram of the mean diurnal anomalies. Figures 2(a)−(d) shows anomalies in T2m only over oceans, SST, SKT and SST−T2m temperature difference. Figures 2(e)−(h) shows the same for the ensemble spreads and T2m−SST ensemble correlations.

A westward propagating diurnal cycle is clear in both means and spreads of the temperatures, except in the T2m spread. The diurnal cycle of temperatures is at a maximum in late afternoons and at a minimum around sunrise (Figures 2(a)−(c)), associated with the diurnal solar heating. The diurnal cycle in T2m reflects the impact of the SKT scheme (Zeng and Beljaars, 2005; Takaya *et al.*, 2010) on the atmosphere. As the warm layer depth in the SKT scheme is 3 m and the SST represents the top 10 m, the amplitude of the diurnal cycle of SST (0.05 °C as a global average) is about one third the amplitude of SKT (0.14 °C as a global average). The amplitude of the diurnal cycle of T2m over the oceans (0.2 °C) is slightly larger than the amplitude of SKT. The diurnal cycle of SST and T2m peaks at the same time, corresponding to the minimum SST−T2m difference (Figure 2(d)), while SKT tends to peak slightly earlier (by ∼3 h). The lag of SST and T2m to SKT might be related to a diurnal cycle in wind speed (not shown), which starts building up in later afternoons helping the surface heat to mix downward through the warm layer and increasing the upward surface turbulent fluxes.

Similarly, the diurnal cycle of the SST spread (with amplitude of 0.001 °C as a global average) is much smaller than the diurnal cycles of the SKT spread (with amplitude of 0.01 °C as a global average) (Figures 2(e)−(g)). However, no diurnal cycle of the T2m spread over the oceans is clearly seen. We think this is related to the strong atmospheric variability within 3 h, which masks the diurnal signal of T2m spread, and perhaps additionally the analysis procedure may generate larger T2m spreads just after the increments are added (2100 UTC). The largest diurnal cycle of SST spread (with amplitude of 0.002 °C) is in the Indian Ocean

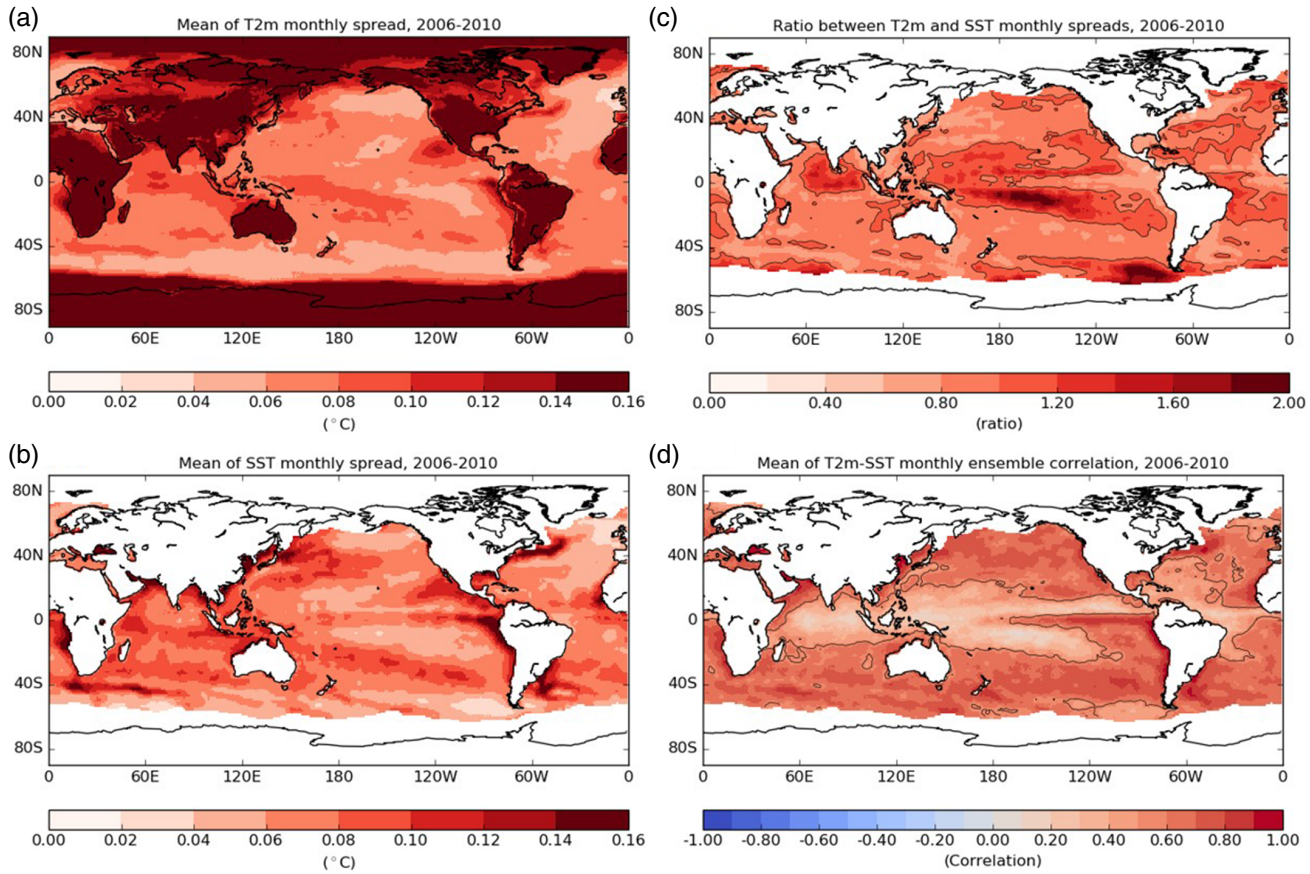

Figure 3. Means of monthly ensemble spreads for (a) T2m and (b) SST, over 2006–2010. (c) shows means of monthly ensemble spread ratio between T2m and SST, and (d) means of monthly ensemble correlations between T2m and SST, over 2006–2010. Contours show values of 1.0 in (c) and 0.55 (90% confidence level) in (d). Correlation means are calculated using the Fisher transformation. SST values on sea ice-covered areas are excluded. [Colour figure can be viewed at wileyonlinelibrary .com].

(30–100°E) where the proportion of subtropical ocean is high. The diurnal cycle of the spread is generally about 1% and 4% of the means of SST and T2m spreads, respectively.

The T2m–SST ensemble correlations also show a diurnal cycle (Figure 2(h)), with an averaged amplitude of 0.03 (8% of the mean of correlations). This indicates the high-frequency variations of the air–sea coupling. T2m and SST ensembles have the highest correlations in late afternoons and the lowest before sunrise, corresponding to the most stable and unstable atmospheric boundary conditions during the day (e.g. air–sea temperature differencesin Figure 2(d)). The diurnal cycle of T2m–SST ensemble correlations can reach up to 0.2 in some areas around the ITCZ (Figure S2), probably due to modulation of the diurnal cycle of precipitation.

3.2. *Monthly time-scale variations*

We expect the air–sea coupling strength to increase over longer time-scales and indeed these are the relationships that would be most useful for initializing longer-range forecasts. To investigate how the coupling strength increases and changes seasonally, and how the coupling strength is phase lagged to the seasonal cycle in the ocean and atmosphere, allowing identification of the forcing patterns, we work with monthly ensemble diagnostics.

Averaging each ensemble member throughout each month before calculating the ensemble mean and spread in T2m and SST yields the lower frequency signals. The seasonal cycle over the last 5 years (2006–2010) is assessed by least squares fitting the monthly properties to the following equation:

$$\eta(t) = \beta_0 + A_a \cos(\,(t - \phi_a)2\pi/12),$$

where $\eta(t)$ is the monthly value at time t (in months), β_0 is the 5-year mean value, and A_a and ϕ_a are the seasonal cycle amplitude and phase. No trend is considered although longer-term trends in earlier periods do exist (section 3.3). The regression is based on bootstrapped confidence intervals running for 1000 times, where 95% is taken as the confidence level.

3.2.1. *Time means*

The T2m monthly spread on land and polar regions is quite high (over 0.2 °C), while the values over the oceans are much lower (less than 0.2 °C) (Figure 3(a)), again indicating the impact of SST on the marine T2m. Over oceans, the T2m and SST monthly spreads (Figures 3(a) and (b)) have the largest values (over 0.1 °C) in regions with strong ocean dynamics. The smallest spreads (less than 0.06 °C) are seen in the northeastern Atlantic and Pacific and some parts of the Southern Ocean where the MLD is deep (over 100 m, Figure 4(a)). Unlike the high-frequency spreads, these T2m and SST monthly spreads have very similar patterns and magnitudes over the oceans, with the SST spread mostly being slightly larger than the T2m spread (Figure 3(c)). Exceptions are again seen in the ITCZ and SPCZ and in some areas close to the Antarctic, where the T2m spread can be twice as large as the SST spread. In the tropical convergence zones, large T2m monthly variability is internal to the atmosphere (e.g. from the Madden–Julian Oscillation), rather than being forced by the SST. Close to the Antarctic, the T2m monthly spread is large (Figure 3(a)) due to the strong temperature gradients over ice, while the SST spread is small (Figure 3(b)) as SST needs to remains near freezing, leading to large T2m–SST spread ratios (Figure 3(c)).

Over most of the oceans, the SST is the main source of T2m monthly variability through upward surface turbulent fluxes (sensible and latent heat fluxes), as the SST is higher than the T2m (Figure 4(b)) and the SST spread is usually larger (Figure 3(c)). In the northern North Pacific and northwest Atlantic in July, temperature gradients are reversed (T2m higher than SST), but latent heat fluxes will dominate here and the SST spread

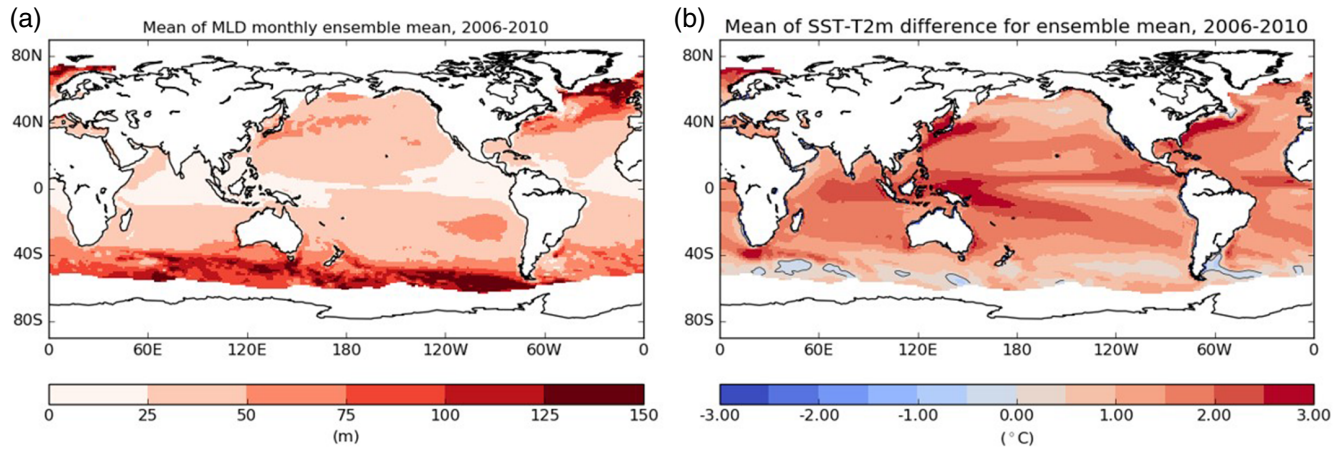

Figure 4. (a) Means of MLD monthly ensemble mean, over 2006–2010. (b) means of SST–T2m monthly ensemble mean difference, over 2006–2010. MLD is defined by using the threshold density criterion of 0.01 kg m^{-3}. The black contour in (b) shows zero. Values on land and sea ice-covered areas are excluded. [Colour figure can be viewed at wileyonlinelibrary.com].

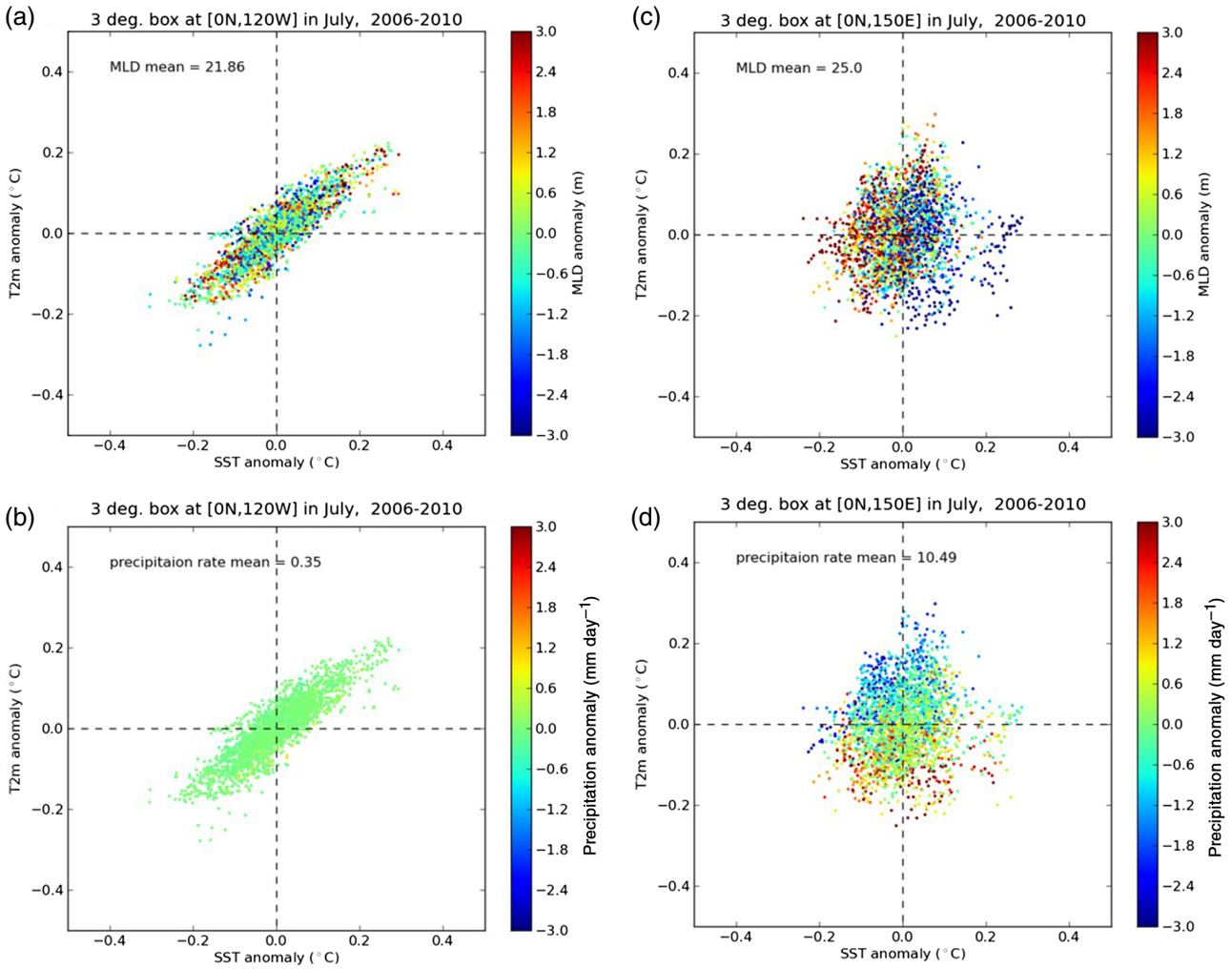

Figure 5. Scatter plots of monthly ensemble anomalies of T2m against SST anomalies in a 3° box centred at [0°N, 120°W] for July data over 2006–2010 with points coloured by (a) MLD and (b) precipitation rate anomalies. (c, d) are as (a, b) but for a 3° box centred at [0°N, 150°E]. Mean MLD (m) and mean precipitation rate (mm day^{-1}) are also shown. [Colour figure can be viewed at wileyonlinelibrary.com].

remains larger than the T2m spread. Low-to-intermediate air–sea temperature differences (SST–T2m ensemble mean difference less than 1.5 °C, Figure 4(b)) allow shallow stable boundary layers to develop where cross-medium fluxes lead to the correlated T2m–SST variability (Figure 3(d)). This is particularly true in summer and in the ocean upwelling regions when the spreads are both largest, corresponding to stronger correlations (over 0.8). This explains some of the largest values of the time means of ensemble correlations (Figure 3(d)).

Over some tropical ocean regions such as the ITCZ and SPCZ, because of the strong atmospheric instability i.e. deep convection (SST–T2m ensemble mean difference over 2 °C, Figure 4(b)), the T2m varies greatly between ensemble members but the SST varies much less (Figure 3(c)). The additional T2m spread that is generated by the internal variability leads to the weakest T2m–SST correlations (less than 0.3). For demonstration, Figures 5(a) and (b) shows the strong relationship between the SST and T2m ensemble anomalies in a 3-degree box centred at 0°N, 120°W in the eastern Equatorial Pacific where the SST anomalies are largely changed by the equatorial currents, using July data. In contrast, at 0°N, 150°E within the ITCZ region, SST and T2m ensemble anomalies show no clear relationship (Figures 5(c) and (d)). At

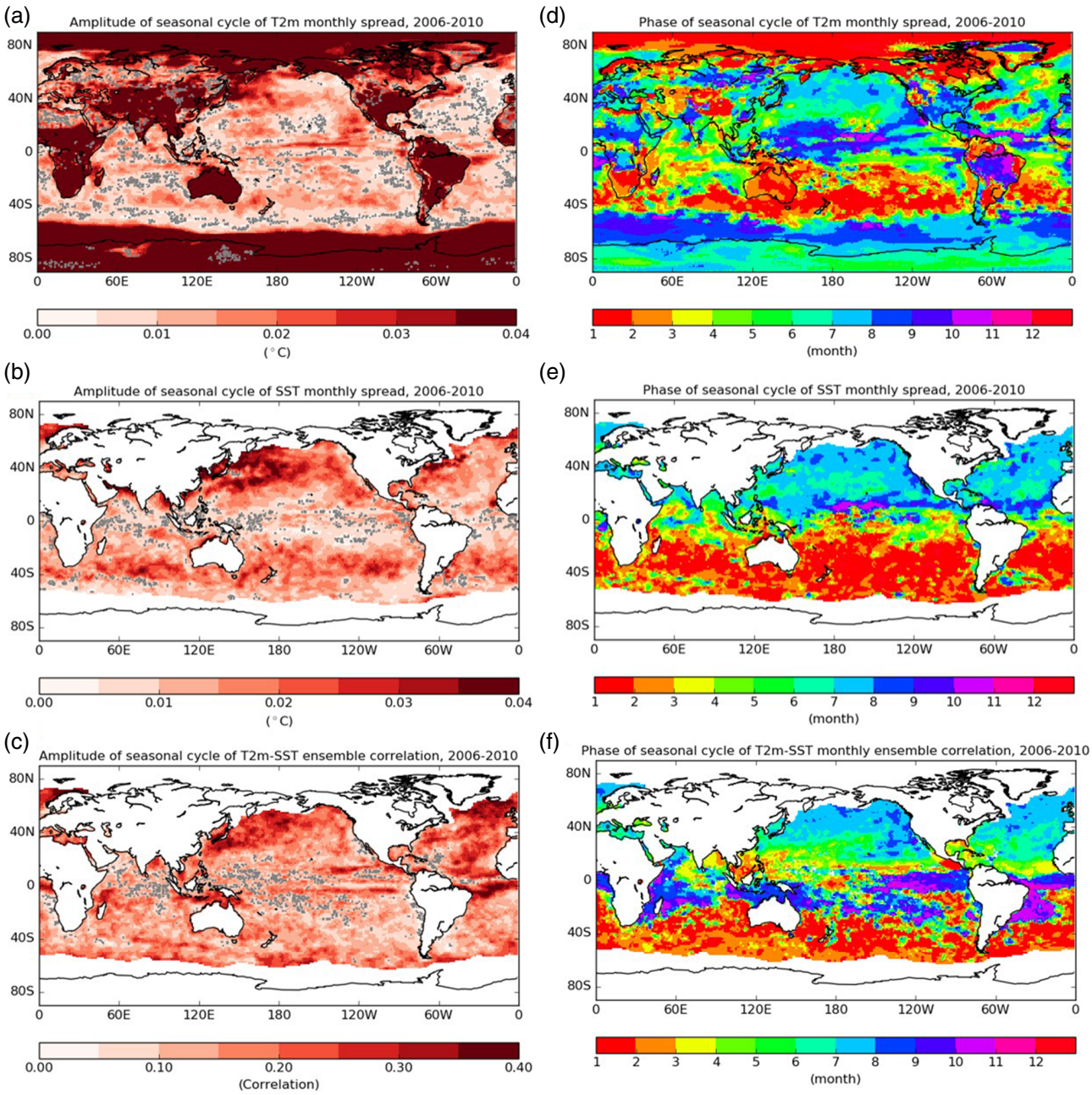

Figure 6. Seasonal cycle amplitudes of monthly ensemble spreads for (a) T2m and (b) SST, and (c) monthly ensemble correlations between T2m and SST, over 2006–2010. (d)–(f) seasonal cycle phases (peaking time) corresponding to the amplitudes in (a)–(c), started at 1st January. Grey areas in (a)–(c) are the locations where the seasonal cycle regression does not pass the significance test at 95% confidence levels. SST values on sea ice-covered areas are excluded. [Colour figure can be viewed at wileyonlinelibrary.com].

this location, ensemble members with higher precipitation rates (coloured red in Figure 5(d)) are associated with negative T2m anomalies, while ensemble members with deeper MLD (red in Figure 5(c)) have negative SST anomalies. We think the link between T2m and precipitation is due to the cooling effect of re-evaporation in the lower atmosphere represented in the IFS atmosphere model (ECMWF, 2016). SST cooling is not associated with increased precipitation (Figure 5(d)), ruling out the solar radiation, which is due to more precipitating clouds, as the cause of cooler T2m. These modulations of the anomalies prevent correlated relationships forming between T2m and SST via the surface turbulent fluxes where deep atmospheric convection is prevalent, and this is also found in the high-frequency (3 h) correlations.

3.2.2. *Seasonal cycle*

The T2m and SST monthly spreads have a significant seasonal cycle over most of the oceans (Figures 6(a) and (b)). The T2m spread has a much bigger seasonal cycle over land and polar regions than over the oceans. The seasonal cycle of the T2m spread is usually less than 0.02 °C over the oceans, about 20% of the time-mean spread, while the seasonal cycle of the SST spread is larger than 0.02 °C in mid-to-high latitudes, about 20–50% of the time-mean spread. The seasonal cycle of the T2m spread is reduced in the areas with small seasonal changes in atmospheric instability (Figure S3), and the seasonal cycle of the SST spread is small in the Tropics (Figure 6(b)) because of the lack of seasonality in MLD (Figure S3).

The SST spread is largest in summer (Figure 6(e)), clearly associated with seasonal variations in MLD (Figure S3). In the North Pacific and Southern Hemisphere (SH) subtropics, the maximum T2m spread is in the summer when the SST spread is also largest (Figures 6(d) and (e)), confirming the T2m spread is driven by the SST. In the Southern Ocean (50–60°S) and along the North Atlantic storm track above the warm ocean currents, the T2m spread peaks in the winter season due to strong instability caused by the large surface temperature gradients. In the tropical North Pacific, the T2m spread peaks in August–September when tropical convection is strongest (Figure S3).

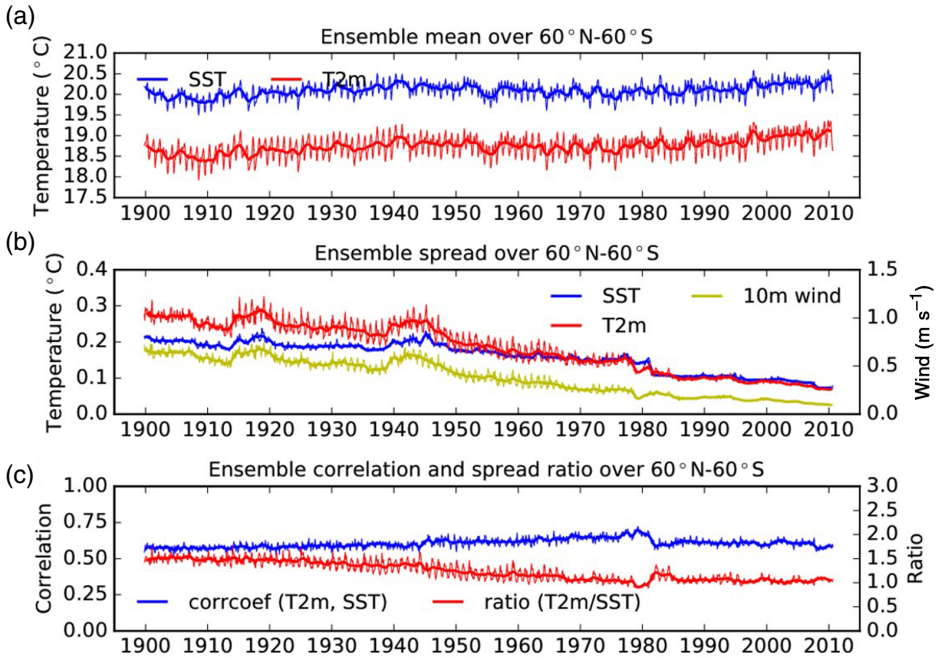

Figure 7. (a) Monthly ensemble means of SST and T2m, (b) monthly ensemble spreads of SST, T2m and 10 m wind speed, and (c) monthly ensemble spread ratio and ensemble correlations between T2m and SST, for a global average (60°N–60°S) with land and sea-ice covered areas excluded, over 1900–2010. Thin lines are for the monthly values, while thick lines are for the 12-month moving average. The spikes in the monthly values indicate seasonal variations. [Colour figure can be viewed at wileyonlinelibrary.com].

The seasonal cycle of T2m–SST ensemble correlations can be clearly seen over most of the oceans (Figures 6(c) and (f)), except in some strong atmospheric convection regions and ocean upwelling regions. The seasonal amplitude of the T2m–SST correlation coefficient is usually larger than 0.1 (more than 15% of the time mean), with largest values (over 0.3) in the Northwestern Pacific and the Northeastern and tropical Atlantic. At mid-to-high latitudes, the correlations peak in local summer time (Figure 6(f)) when the SST spread is also largest (Figure 6(e)), indicating the ocean as the main driver for the T2m–SST relationships in these regions. At low latitudes, the correlations peak in local spring (March–June in the north and September–December in the south) (Figure 6(f)). This is when tropical convection is weakening/enhancing in the northern/southern Tropics (Figure S3), showing the weakening impact of atmospheric instability on the air–sea coupling.

3.3. *Long-term evolution*

3.3.1. *Global average*

In this subsection, we evaluate the long-term changes in the analysis spread over 1900–2010 using monthly mean ensemble products (thin lines in Figure 7). A 12-month moving average excludes the seasonal cycle (thick lines in Figure 7). The global average of the SST is about 1.5 °C higher than the global average of the T2m over the oceans throughout the whole period (Figure 7(a)). SST and T2m show very similar interannual variations in both ensemble mean and ensemble spread (Figure 7(b)), indicating the overall coupling between SST and T2m. The spreads in the early twentieth century are larger with values of ∼0.20 °C for SST and ∼0.27 °C for T2m, and decrease to values less than 0.1 °C by 2010. The long-term variations of SST spread are overall determined by the HadISST2 dataset via the SST relaxation scheme (this will be discussed in section 4). The downward trends reflect the growing availability of observations.

The ensemble spreads suddenly increase in the 1910s and the 1940s (Figure 7(b)) due to data-sparsity during the two World Wars (Rayner *et al.*, 2003). Spreads decrease around 1980 and 2008, coinciding with the introduction of Advanced Very High Resolution Radiometer (AVHRR) data from NOAA, and MetOp satellites, respectively, into the HadISST2 product. Prior to the 1970s the T2m spread is larger than the SST spread, and the two spreads are much closer thereafter with the T2m spread becoming smaller, corresponding to the decreasing ratio and increasing correlations between T2m and SST spreads (Figure 7(c)). We find this is related to the changes in the wind spread (Figure 7(b)). In the early years, when the wind is not well constrained due to the sparse observations, the T2m spread is large among the ensemble members because of the lack of control on atmospheric dynamics. As the wind spread reduces in recent decades, the T2m spread decreases.

The global average for T2m–SST ensemble correlations starts at ∼0.55 in the early twentieth century, and steadily increases to 0.7 by around 1980 at the start of the satellite era when the SST spread suddenly reduces as SST quality improves due to the availability of remote sensing in HadISST2 (Figures 7(b) and (c)). Then, the global average for correlations suddenly decreases to 0.55. The T2m–SST spread ratio is decreasing when the ensemble correlations are increasing. The T2m–SST ensemble correlations initially reduce when new satellite data start to come in around 1980 and 2008, and then recover. This predominantly arises from the changes in the SH and will be discussed in the next subsection.

3.3.2. *Midlatitudes*

In the SH midlatitudes (30–60°S), the ensemble spreads of T2m and SST are both large in the first half of the twentieth century (Figure 8(a)), with values ∼0.4 and 0.2 °C, respectively, reflecting the less constrained atmospheric and oceanic dynamics. As more observations become available, the two spreads converge towards values ∼0.1 °C. In the Northern Hemisphere (NH) midlatitudes, observations increase rapidly in the early twentieth century and the spreads decline correspondingly (Figure 8(d)), while the T2m–SST ensemble correlations increase to values over 0.7 (Figure 8(e)). Note the large seasonal variations of spreads and their relationships (thin lines in Figure 8).

In the SH midlatitudes, the T2m spread is larger than the SST spread in the first half of the century, and the T2m–SST ensemble correlations are weaker during this period (Figure 8(b)), indicating less control of SST on the T2m spread. The spread of surface winds significantly influences the T2m spread in these early years as wind observations are very limited and atmospheric dynamics dominate the T2m spread (Figure 8(a)

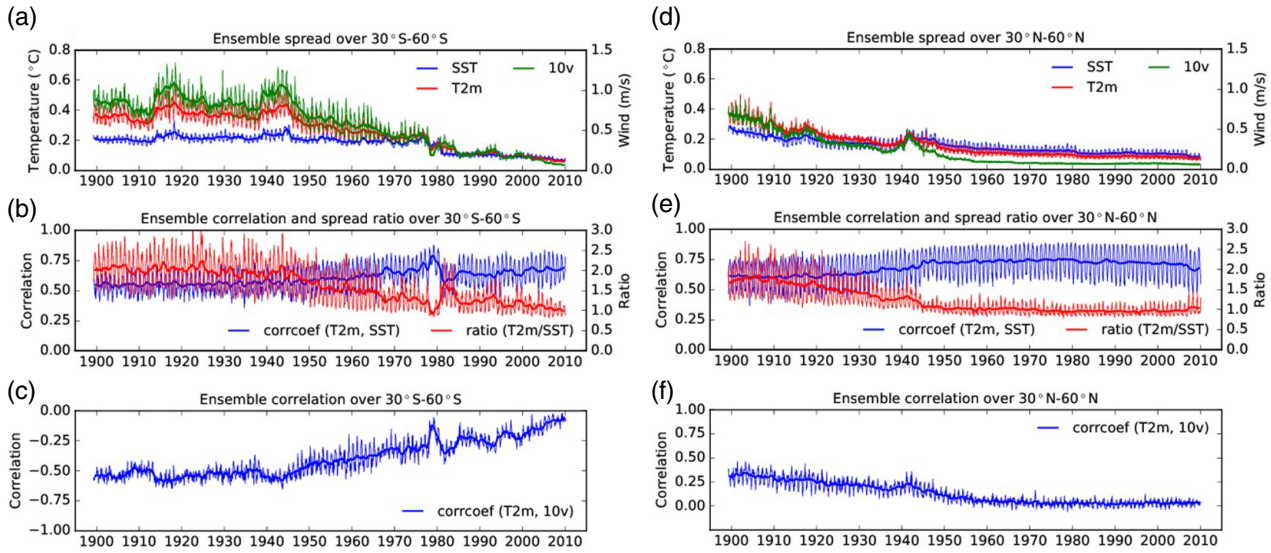

Figure 8. (a) Monthly ensemble spreads of SST, T2m and 10v (meridional 10 m wind speed), (b) monthly ensemble spread ratio and ensemble correlations between T2m and SST, and (c) monthly ensemble correlation between T2m and 10v, for 30–60°S average with land and sea-ice covered areas excluded, over 1900–2010. (d)–(f) are as (a)–(c), but for 30–60°N average. Thin lines are for the monthly values, while thick lines are for the 12 month moving average. The spikes in the monthly values indicate seasonal variations. [Colour figure can be viewed at wileyonlinelibrary.com].

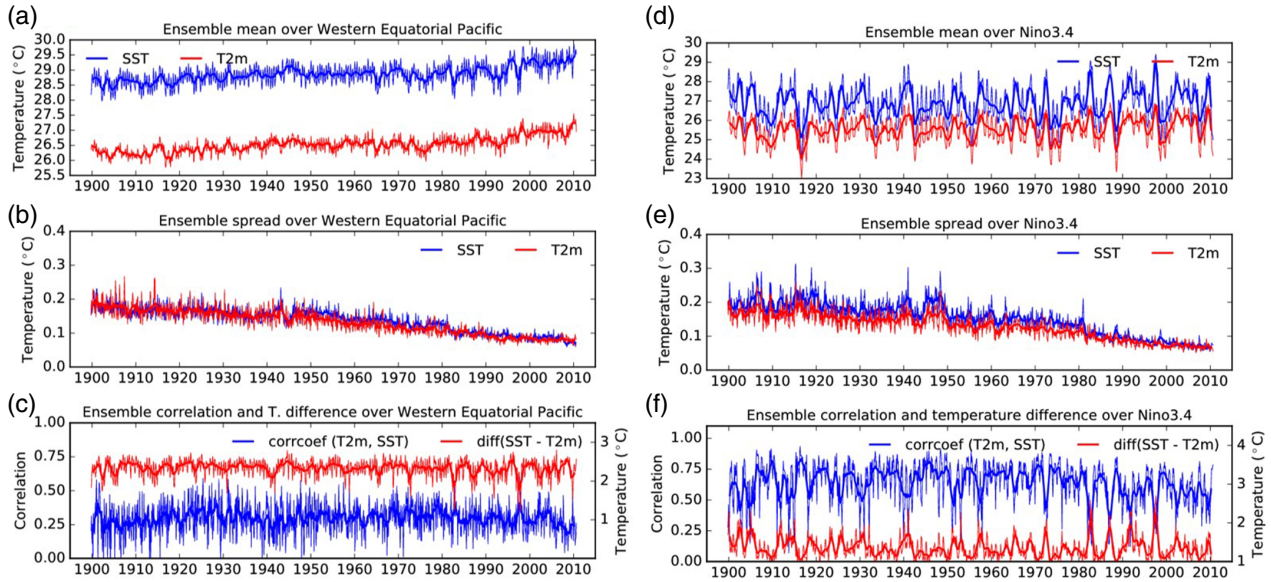

Figure 9. (a) Monthly ensemble means of SST and T2m, (b) monthly ensemble spreads of SST and T2m, and (c) monthly ensemble correlations between T2m and SST, and temperature ensemble mean difference (SST–T2m), for the Western Equatorial Pacific with land areas excluded, over 1900–2010. (d)–(f) are as (a)–(c), but for the Niño3.4 region. Thin lines are for the monthly values, while thick lines are for the 12-month moving average. The spikes in the monthly values indicate the seasonal variations. [Colour figure can be viewed at wileyonlinelibrary.com].

and (c)). The spread of 10 m meridional wind (10v), which regulates the southward/northward movement of warm/cold air in SH midlatitudes, can reach up to 1 m s^{-1} in these early years. In later years when the wind is more constrained, the T2m spread reduces with less association with the wind spread and more coupling with the SST spread.

The larger spreads during the World Wars (in the 1910s and the 1940s) can be seen in both the SH and NH, and the sudden changes in T2m–SST correlations around 1980 can be seen to occur in the SH but not the NH (Figure 8). The SH sees the biggest changes when satellite data become available around 1980, permanently reducing the SST spread dramatically. However, the T2m spread temporarily reduces more than the SST spread, which leads to a brief increase in T2m–SST correlations. The reasons for this are hard to understand, but may be linked to a temporary dip in the surface wind spread, reducing wind impacts on the T2m.

3.3.3. *Tropics*

In the Tropics, the Western Equatorial Pacific and Niño3.4 (120°–170°W; 5°S–5°N) regions are compared in Figure 9. The Western Equatorial Pacific shows a warming trend in both T2m and SST (Figure 9(a)). The monthly spreads gradually decline from 0.2 to less than 0.1 °C over the century, without much interannual variability (Figure 9(b)), and the T2m–SST ensemble correlations are always less than 0.3 (Figure 9(c)), indicating the consistently weak coupling between T2m and SST monthly variability. The weak coupling reflects the deep convection, which is also indicated by the large air–sea temperature differences (SST–T2m ≈ 2.5 °C). When the convection is weaker during El Niño events due to the warm pool moving to the east, for example in November 1997 (Figure 10(a)–(c)), the shallower atmospheric boundary layer allows SST to considerably force T2m in the west, leading to noticeable increase of T2m–SST correlations, although still remaining less than 0.5.

In the Niño3.4 region, both ocean and air surface temperatures have profound interannual variations (Figure 9(d)), related to ENSO. The SST and T2m spreads again both decline with time (Figure 9(e)), and the SST spread is slightly larger than the T2m spread, indicating the normal dominance of SST monthly variability. However, the T2m–SST ensemble correlations exhibit striking interannual variability (Figure 9(f)), ranging from 0.8

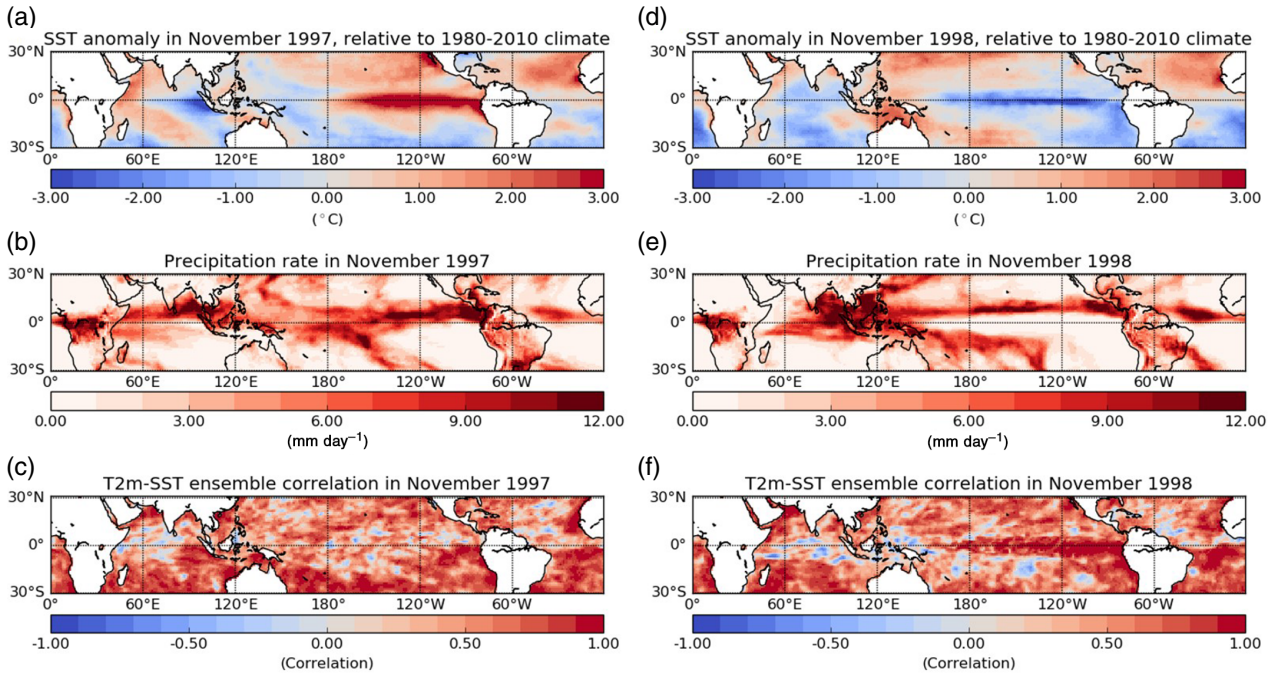

Figure 10. (a) Anomalies of SST monthly ensemble mean between 30°N and 30°S in November 1997 relative to 1980–2010 climate, (b) precipitation ensemble mean, and (c) monthly ensemble correlations between T2m and SST. (d)–(f) are as (a)–(c), but for November 1998. [Colour figure can be viewed at wileyonlinelibrary.com].

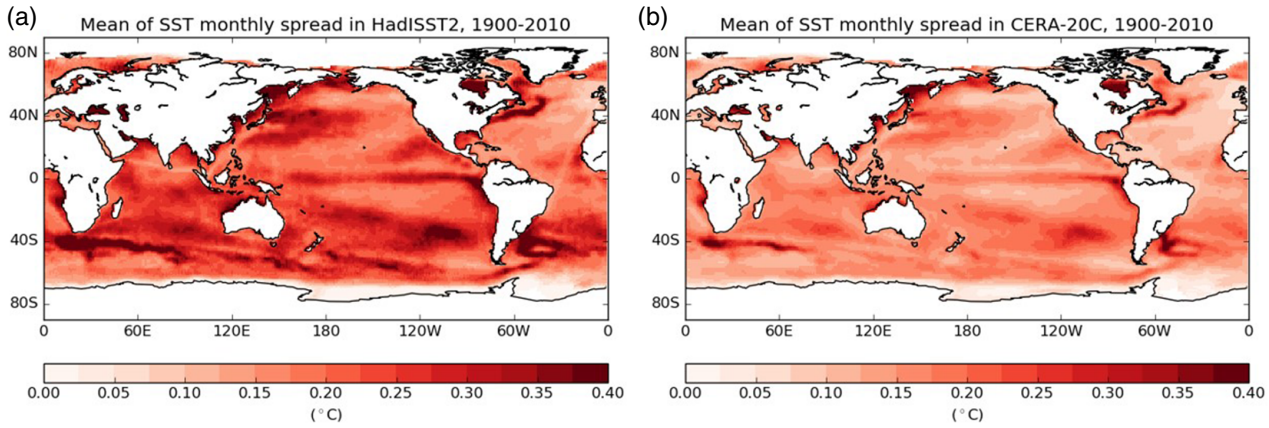

Figure 11. Means of SST monthly ensemble spread in (a) HadISST2 and (b) CERA-20C, averaged over 1900–2010. SST values on sea ice-covered areas are excluded. [Colour figure can be viewed at wileyonlinelibrary.com].

to 0.25. During El Niño events Figure 10(a)–(c), the T2m–SST correlations are weak over Niño3.4 as the Pacific warm pool moves in, building up the deep convection (also corresponding to an increase in air–sea temperature difference; Figure 9(f)), which tends to influence the air temperatures via the strong precipitation. T2m and SST therefore become more independent during El Niño events. Conversely, during La Niña events, for example in November 1998 (Figure 10(d)–(f)), the warm pool moves out the Niño3.4 region, convection and precipitation are weakened, and very stable atmospheric boundary conditions allow the SST to drive the surface air temperature, producing particularly strong T2m–SST correlations.

4. Determining the SST spread

Over most of the oceans, particularly in the recent period with good observation coverage, low-frequency SST spread is larger than the T2m spread and their correlations are high (Figure 3), indicating that the T2m variability is mainly forced by SST, so it is important to understand how the SST ensemble variability is generated. In CERA-20C, SST is constrained towards the ten initializations of the monthly gridded ensemble product HadISST2 (Titchner and Rayner, 2014), via a Newtonian relaxation scheme (Balmaseda *et al.*, 2013), but the analysed SSTs do differ from the HadISST2.

Spatial features of the SST spread in CERA-20C agree well with those in HadISST2 (Figure 11). Long-term variations of the SST spread are also similar between products (Figure 12), reflecting the overall decrease of uncertainties in SST observational sources and bias adjustments (Kennedy *et al.*, 2011). However, CERA-20C SST has a smaller spread than HadISST2 nearly everywhere during the twentieth century (Figures 11 and 12), and the differences are larger in cold seasons when the spread is usually smaller (Figure S4). The global average of the difference is 0.11–0.05 °C (Figure 12), amounting to a reduction ∼25–45% from the HadISST2 spread. Discrepancies are smaller in the satellite era when the HadISST2 spread is smaller.

The global average of CERA-20C SST is about 0.04 °C higher than HadISST2, which remains similar throughout the century (yellow lines in Figure 12). The regions with larger SST mean differences between CERA-20C and HadISST2 are also the regions where the SST ensemble spreads differ more (Figure S4), for example in the western boundary currents and Southern Ocean, implying the ensemble spread and ensemble mean bias may be related. The CERA-20C mean SST bias is also present before the satellite era, so it is likely to be unrelated to uncertainties in satellite data used in HadISST2 or the availability of other observations. We conclude that both the higher mean SSTs than HadISST2, and the smaller SST spread, must be due to background bias in the ocean model of CERA-20C. The ocean model is generally drifting warm

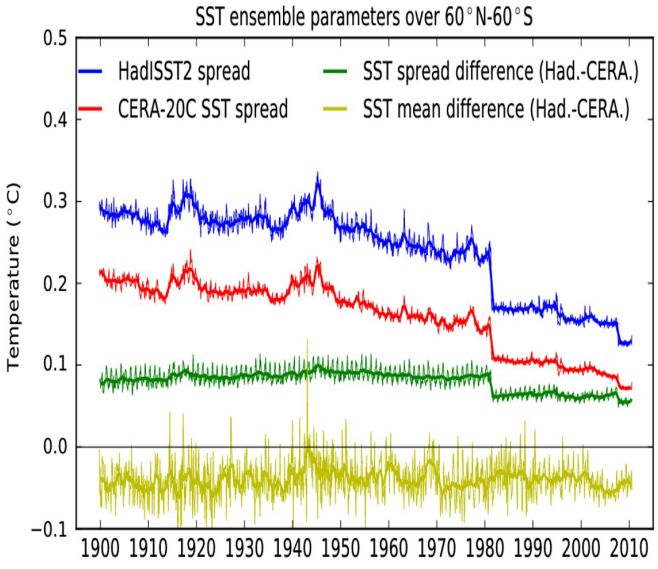

Figure 12. SST monthly ensemble spread in HadISST2 (blue) and CERA-20C (red), and their differences (green), for a global average (60°N–60°S) with sea-ice covered areas excluded, over 1900–2010, together with differences of SST monthly ensemble mean between HadISST2 and CERA-20C(yellow). Thin lines are for the monthly values, while thick lines are for the 12-month moving average. The spikes in the monthly values denote seasonal variations. [Colour figure can be viewed at wileyonlinelibrary.com].

and also has a tendency to collapse the SST spread. This suggests that ensemble inflation would be needed to make the SST spread realistic for using as an estimate of errors (Isaksen *et al.,* 2010).

It is desirable for the ensemble spread to reflect the uncertainties of the ensemble mean for representing the 'truth'. One might expect that since CERA-20C assimilates observations other than just SST, that the CERA-20C mean SST should be more accurate than HadISST2, and that this would be properly represented by a smaller ensemble spread. However, the reduction of CERA-20C SST spread from HadISST2 spread is not a function of the availability of other observations (e.g. wind observations and ocean profile data), and it consistently appears with nearly constant values prior to 1980, even in the early years when very limited observations are assimilated. Therefore, the overall smaller SST spread in CERA-20C than in HadISST2 should not be interpreted as greater confidence in the SST ensemble mean results, and is rather due to model error as discussed above.

5. Conclusions and discussions

ECMWF has produced the first AO coupled reanalysis ensemble for the twentieth century, called CERA-20C, using the CERA coupled DA system. CERA produces DA increments in the ocean and atmosphere individually, with the SST constrained by the HadISST2 product. Because CERA does not use coupled background-error covariances, the system has been labelled as a 'weakly' coupled DA system (Laloyaux *et al.,* 2016). However in CERA the strength of AO coupling in the DA is enhanced by applying multiple iterations in the DA outer loop, allowing the observational increments from ocean and atmosphere to communicate within each 24 h cycle. This paper focuses on the T2m–SST relationships in this coupled climate reanalysis using ensemble statistics. Because the air temperature is not assimilated in CERA-20C, the T2m–SST ensemble relationships are built up through the two-way coupling exerted in the coupled analysis system. This article explains how and where the atmosphere and/or the oceans are driving the ensemble variability of temperature at the interface on different time-scales.

At high frequencies (3-hourly), the T2m ensemble spread is always larger than the SST spread over the oceans, and the two spreads have very different spatial patterns, indicating different sources for the development of the temperature variability. The T2m spread is significantly driven by the underlying SST spread only where the atmospheric boundary conditions are very stable. At high frequencies the rapidly evolving atmosphere usually does not depend on details of the ocean variability which are changing more slowly. Daily-varying features and a large-scale diurnal cycle are both distinguishable in the spreads and their correlations, highlighting the high-frequency dynamics reproduced by CERA-20C.

In contrast, the low-frequency (monthly) T2m ensemble variability is closely coupled with the SST ensemble variability. In recent well-observed period (2006–2010), over most of the oceans the T2m spread is slightly smaller than the SST spread with very similar large-scale spatial distributions, and the two spreads are well correlated. This close relationship on monthly time-scales is established through the SST forcing. However, in some tropical regions e.g. ITCZ and SPCZ, the T2m spread can still be double the SST spread on monthly time-scales, with much weaker correlations due to the strong atmospheric instability impacting on T2m. The T2m–SST relationships exhibit a clear large-scale seasonal cycle, with peak values in summer when both atmospheric and oceanic boundary layers are stable and shallow.

As more observations become available throughout the century, the SST and T2m spreads decline globally and become more correlated. In the early years in SH midlatitudes, the T2m spread is less coupled with the SST spread as the winds are less constrained by observations within the ensemble members. When the atmospheric circulation is better constrained in later decades, the T2m spread is reduced and is more coupled with the SST spread. In the Niño3.4 region, a striking interannual variability of T2m–SST correlations is found related to ENSO, as the convective warm pool moves interannually. We further found that the SST monthly spread in CERA-20C is constrained by HadISST2 in terms of spatial and temporal variability, but with considerable discrepancies, which are changing with time.

In CERA-20C, the air-sea interactions are represented in the two-way coupling between ocean and atmosphere boundary layers. This makes the reanalysis suitable for further analysis of other boundary-layer processes. For example, the SST–precipitation relationship, which is important for tropical variability, should be better represented in this coupled climate data than in uncoupled products. This relationship is being examined in more detail and will be presented in another article. We also note that a higher resolution (1/4°) version of the CERA system (CERA-SAT) is being produced using satellite observations for short periods (Buizza *et al.,* 2017), which will better represent small-scale features of the ocean and atmosphere; the analysis of these may help to explain more detailed aspects of the coupled variability.

Acknowledgements

This work is supported by the ERA-CLIM2 project, which is funded by the European Union Seventh Framework Program (EU FP7). We thank ECMWF and particularly Patrick Laloyaux, for useful discussions and for providing the CERA-20C data. We also thank David P. Mulholland at the University of Reading for useful discussions in the early phases of the work, as well as two reviewers for their constructive comments and suggestions.

Supporting information

The following supporting information is available as part of the online article:

Figure S1. (a) means of temperature ensemble mean differences (SST-T2m) and (b) means of MLD ensemble mean, averaged over daily fields throughout January 2006. Contour line shows values of 2 °C in (a) and 20 m in (b). Values in high latitudes (70–90°N, 70–90°S) and in sea ice-coved areas are excluded.

Figure S2. Amplitude of diurnal cycle of T2m-SST ensemble correlations, in January 2006. Values in high latitudes (70–90°N, 70–90°S) and in sea ice-coved areas are excluded.

Figure S3. Panel (a, b) seasonal cycle amplitudes of temperature monthly ensemble mean difference (SST-T2m) and of MLD monthly ensemble mean, over 2006–2010. Panel (c, d) seasonal cycle phases (peaking time) corresponding to the amplitudes in (a, b), started at 1 January. Grey areas in (a, b) are the locations where the seasonal cycle regression does not pass the significance test at 95% confidence levels. Land and sea ice-covered areas are excluded.

Figure S4. Panel (a, b) SST monthly ensemble mean differences between CERA-20C and HadISST2, averaged in January and July of 1900–2010. Panel (c, d) same as (a, b) but for SST monthly ensemble spread differences. Sea ice-covered areas are excluded.

References

Balmaseda MA, Mogensen K, Weaver AT. 2013. Evaluation of the ECMWF ocean reanalysis system ORAS4. *Q. J. R. Meteorol. Soc.* **139**: 1132–1161. https://doi.org/10.1002/qj.2063.

Balsamo G, Beljaars A, Scipal K, Viterbo P, van den Hurk B, Hirschi M, Betts AK. 2009. A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *J. Hydrometeorol.* **103**: 623–643.

de Boisséson E, Balmaseda MA, Mayer M. 2017. Ocean heat content variability in an ensemble of twentieth-century ocean reanalyses. *Clim. Dyn.* In Press. https://doi.org/10.1007/s00382-017-3845-0.

Bonavita M, Hólm E, Isaksen L, Fisher M. 2016. The evolution of the ECMWF hybrid data assimilation system. *Q. J. R. Meteorol. Soc.* **142**: 287–303. https://doi.org/10.1002/qj.2652.

Bouillon S, Maqueda MAM, Legat V, Fichefet T. 2009. An elastic-viscous-plastic sea ice model formulated on Arakawa B and C grids. *Ocean Model.* **27**: 174–184. https://doi.org/10.1016/j.ocemod.2009.01.004.

Buizza R, Broennimann S, Haimberger L, Laloyaux P, Martin MJ, Fuentes M, Balmaseda MA, Becker A, Blaschek M, Dahlgren P, de Boisseson E, Dee D, Boucher MD, Feng X, John V, Haines K, Jourdain S, Kosaka Y, Lea D, Lemarie F, Mayer M, Messina P, Perruche C, Paylin P, Pullainen J, Rayner N, Rustemeier E, Schepers D, Saunders R, Schulz J, Sterin A, Stichelberger S, Storto A, Testut CE, Valente MA, Vidard A, Vuichard N, Weaver A, While J, Ziese M. 2017. The EU-FP7 ERA-CLIM2 project contribution to advancing science and production of Earth-system climate reanalyses. *Bull. Am. Meteorol. Soc.* accepted.

Carton JA, Giese BS. 2008. A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Mon. Weather Rev.* **136**: 2999–3017.

Cayan DR. 1980. Large-scale relationships between sea surface temperature and surface air temperature. *Mon. Weather Rev.* **108**: 1293–1301.

Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose RS, Rutledge G, Bessemoulin P, Brönnimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY, Jones PD, Kruk MC, Kruger AC, Marshall GJ, Maugeri M, Mok HY, Nordli Ø, Ross TF, Trigo RM, Wang XL, Woodruff SD, Worley SJ. 2011. The twentieth century reanalysis project. *Q. J. R. Meteorol. Soc* **137**: 1–28.

Cram TA, Compo GP, Yin X, Allan RJ, McColl C, Vose RS, Whitaker JS, Matsui N, Ashcroft L, Auchmann R, Bessemoulin P, Brandsma T, Brohan P, Brunet M, Comeaux J, Crouthamel R, Gleason BE, Groisman PY, Hersbach H, Jones PD, Jónsson T, Jourdain S, Kelly G, Knapp KR, Kruger A, Kubota H, Lentini G, Lorrey A, Lott N, Lubker SJ, Luterbacher J, Marshall GJ, Maugeri M, Mock CJ, Mok HY, Nordli Ø, Rodwell MJ, Ross TF, Schuster D, Srnec L, Valente MA, Vizi Z, Wang XL, Westcott N, Woollen JS, Worley SJ. 2015. The international surface pressure databank version 2. *Geosci. Data J.* **2**: 31–46. https://doi.org/10.1002/gdj3.25.

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**: 553–597.

Dee DP, Balmaseda M, Balsamo G, Engelen R, Simmons AJ, Thépaut J-N. 2014. Toward a consistent reanalysis of the climate system. *Bull. Am. Meteorol. Soc.* **95**: 1235–1248.

ECMWF. 2016. *IFS Documentation CY41R2*. Reading, UK: ECMWF.

Fichefet T, Maqueda MMA. 1997. Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. *J. Geophys. Res.* **102**: 12609–12646. https://doi.org/10.1029/97JC00480.

Good SA, Martin MJ, Rayner NA. 2013. EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res.* **118**: 6704–6716. https://doi.org/10.1002/2013JC009067.

Gouretski V, Reseghetti F. 2010. On depth and temperature biases in bathythermograph data: Development of a new correction scheme based on analysis of a global ocean database. *Deep Sea Res. Part I.* **57**: 812–833.

Han G, Wu X, Zhang S, Liu Z, Li W. 2013. Error covariance estimation for coupled data assimilation using a Lorenz atmosphere and a simple pycnocline ocean model. *J. Clim.* **26**: 10218–10231.

Isaksen L, Bonavita M, Buizza R, Fisher M, Haseler J, Leutbecher M, Raynaud L. 2010. *Ensemble of Data Assimilations at ECMWF*. ECMWF: Reading, UK.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds B, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer Meteor. Soc.* **77**: 437–472.

Kennedy JJ, Rayner NA, Smith RO, Parker DE, Saunby M. 2011. Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2. Biases and homogenization. *J. Geophys. Res.* **116**: D14104. https://doi.org/10.1029/2010JD015220.

Komen GJ, Cavaleri L, Donelan M, Hasselmann K, Hasselmann S, Janssen PAEM. 1994. *Dynamics and Modelling of Ocean Waves*. Cambridge University Press: Cambridge, UK.

Laloyaux P, Balmaseda M, Mogensen K, Janssen P, Dee D. 2014. 'The ECMWF coupled assimilation system'. In *Proceedings of the 2014 Conference on Big Data from Space (BiDS'14)*, Frascati, Italy, European Space Agency-European Space Research Institute, 16–19. 10.2788/1823.

Laloyaux P, Balmaseda M, Dee DP, Mogensen K, Janssen PAEM. 2016. A coupled data assimilation system for climate reanalysis. *Q. J. R. Meteorol. Soc.* **142**: 65–78. https://doi.org/10.1002/qj.2629.

Lea D, Mirouze I, Martin M, King R, Hines A, Walters D, Thurlow M. 2015. Assessing a new coupled data assimilation system based on the Met Office coupled atmosphere, land, ocean, sea ice model. *Mon. Weather Rev.* **143**: 4678–4694. https://doi.org/10.1175/MWR-D-15-0174.1.

Lu F, Liu Z, Zhang S, Liu Y. 2015a. Strongly coupled data assimilation using Leading Averaged Coupled Covariance (LACC). Part I: Simple model study. *Mon. Weather Rev.* **143**: 3823–3837.

Lu F, Liu Z, Zhang S, Liu Y, Jacob R. 2015b. Strongly coupled data assimilation using Leading Averaged Coupled Covariance (LACC). Part II: CGCM experiments. *Mon. Weather Rev.* **143**: 4645–4659.

Madec G. 2008. 'NEMO reference manual, ocean dynamics component: NEMO-OPA. Preliminary version', Note du Pole de Modelisation 27. Institut Pierre-Simon Laplace (IPSL), Paris, France.

Martin M, Balmaseda M, Bertino L, Brasseur P, Brassington G, Cummings J, Fujii Y, Lea D, Lellouche JM, Mogensen K, Oke P, Smith G, Testut CE, Waagbø G, Waters J, Weaver A. 2015. Status and future of data assimilation in operational oceanography. *J. Oper. Oceanogr.* **8**: S28–S48.

Merryfield WJ, Lee WS, Wang W, Chen M, Kumar A. 2013. Multi-system seasonal predictions of Arctic sea ice. *Geophys. Res. Lett.* **40**: 1551–1556. https://doi.org/10.1002/grl.50317.

Mirouze I, Weaver AT. 2010. Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Q. J. R. Meteorol. Soc.* **136**: 1421–1443.

Poli P, Hersbach H, Tan D, Dee DP, Thepaut J-N, Simmons A, Peubey C, Laloyaux P, Komori T, Berrisford P, Dragani R. 2013. 'The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C)', ERA Report Series No. 14. ECMWF: Reading, UK.

Rawlins F, Ballard SP, Bovis KJ, Clayton AM, Li D, Inverarity GW, Lorenc AC, Payne TJ. 2007. The Met Office global four-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.* **133**: 347–362.

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A. 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* **108**: 4407. https://doi.org/10.1029/2002JD002670.

Saha S, Moorthi S, Pan H, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu H, Stokes D, Grumbine R, Gayno G, Hou YT, Chuang HY, Juang HMH, Sela J, Iredell M, Treadon R, Kleist D, van Delst P, Keyser D, Derber J, Ek M, Meng J, Wei H, Yang R, Lord S, van den Dool H, Kumar A, Wang W, Long C, Chelliah M, Xue Y, Huang B, Schemm JK, Ebisuzaki W, Lin R, Xie P, Chen M, Zhou S, Higgins W, Zou CZ, Liu Q, Chen Y, Han Y, Cucurull L, Reynolds RW, Rutledge G, Goldberg M. 2010. The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* **91**: 1015–1057.

Sluka TC, Penny SG, Kalnay E, Miyoshi T. 2016. Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophys. Res. Lett.* **43**: 752–759. https://doi.org/10.1002/2015GL067238.

Smith PJ, Fowler AM, Lawless AS. 2015. Exploring strategies for coupled 4D-Var data assimilation using an idealised atmosphere–ocean model. *Tellus A* **67**: 27025. https://doi.org/10.3402/tellusa.v67.27025.

Smith PJ, Lawless AS, Nichols NK. 2017. Estimating forecast error covariances for strongly coupled atmosphere–ocean 4D-Var data assimilation. *Mon. Weather Rev.* **145**: 4011–4035.

Takaya Y, Bidlot JR, Beljaars ACM, Janssen PAEM. 2010. Refinements to a prognostic scheme of skin sea surface temperature. *J. Geophys. Res.* **115**: C06009. https://doi.org/10.1029/2009JC005985.

Tardif R, Hakim GJ, Snyder C. 2014. Coupled atmosphere–ocean data assimilation experiments with a low-order model and CMIP5 model data. *Clim. Dyn.* **45**: 1415–1427. https://doi.org/10.1007/s00382-014-2390-3.

Titchner HA, Rayner NA. 2014. The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations. *J. Geophys. Res.* **119**: 2864–2889. https://doi.org/10.1002/2013JD020316.

Woodruff SD, Worley SJ, Lubker SJ, Ji Z, Eric Freeman J, Berry DI, Brohan P, Kent EC, Reynolds RW, Smith SR, Wilkinson C. 2011. ICOADS Release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.* **31**: 951–967. https://doi.org/10.1002/joc.2103.

Zeng X, Beljaars A. 2005. A prognostic scheme of sea surface skin temperature for modeling and data assimilation. *Geophys. Res. Lett.* **32**: L14605. https://doi.org/10.1029/2005GL023030.

Zhang S, Harrison MJ, Rosati A, Wittenberg A. 2007. System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Weather Rev.* **135**: 3541–3564.

Zuo H, Balmaseda MA, Mogensen K. 2015. The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals. *Clim. Dyn.* **49**: 791–811.

Zuo H, Balmaseda MA, Boisseson E, Hirahara S. 2017. 'A new ensemble generation scheme for ocean analysis', Technical report 795. ECMWF: Reading, UK.