Atmospheric Science Letters

RMetS Royal Meteorological Society

RESEARCH ARTICLE OPEN ACCESS

The Representation of Surface Temperature Trends in C3S Seasonal Forecast Systems

Matthew Patterson[1,2,3] | Daniel J. Befort[4] | Julia F. Lockwood[5] | John Slattery[6] | Antje Weisheimer[7,8]

[1]National Centre for Atmospheric Science (NCAS), University of St Andrews, St Andrews, UK | [2]School of Earth and Environmental Sciences, University of St Andrews, St Andrews, UK | [3]School of Mathematics and Statistics, University of St Andrews, St Andrews, UK | [4]European Centre for Medium-Range Weather Forecasts (ECMWF), Bonn, Germany | [5]Met Office Hadley Centre, Exeter, UK | [6]British Antarctic Survey, Cambridge, UK | [7]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK | [8]National Centre for Atmospheric Science (NCAS), University of Oxford, Oxford, UK

Correspondence: Matthew Patterson (mrp23@st-andrews.ac.uk)

Received: 22 April 2025 | **Revised:** 25 June 2025 | **Accepted:** 14 July 2025

Funding: M.P., D.J.B., and A.W.: EU Horizon 2020, grant agreement 776613; J.S.: NERC and Met Office. J.F.L.: Met Office Hadley Centre Climate Programme funded through the UK Government Departments, BEIS, and Defra.

Keywords: atmospheric and climate dynamics | climate variability, change & impacts | seasonal scale | weather and climate prediction

ABSTRACT

We assess near-surface temperature and sea surface temperature trends in 8 seasonal forecast systems in the Copernicus Climate Change Service archive, over the common hindcast period (1993–2016). All but one of the systems show a faster warming of the global-mean, relative to observations in both boreal summer and winter seasons. On average, systems warm at 0.21 K/decade and 0.22 K/decade for winter and summer, respectively, compared to 0.17 K/decade and 0.19 K/decade for ERA5. In summer, forecast systems tend to show an excessive warming of the tropical Pacific, tropical Atlantic and southern mid-latitudes, which contributes to the difference in global warming rates compared to observations. In contrast, greater warming in the northern mid-latitudes contributes most to trend differences for winter. The faster warming of models over this period has important implications for seasonal forecasts of future global and regional temperature and suggests further work is required to understand this bias.

1 | Introduction

Seasonal forecasts aim to predict seasonally-averaged variables, such as temperature and precipitation, relative to the long-term average, several months in advance (Weisheimer and Palmer 2014). These forecasts provide potentially valuable information to a range of industries including the water, agriculture and energy sectors. They rely on boundary conditions and slowly varying components of the climate system, such as the ocean, which influence the atmosphere in a predictable way (Palmer and Anderson 1994). The pace of anthropogenic climate change is such that increases to greenhouse gas concentrations are also an important predictor for regional near-surface temperature anomalies, even on the seasonal timescale (Doblas-Reyes et al. 2006; Liniger et al. 2007; Patterson et al. 2022, 2024).

In addition to regional temperature variability, efforts to limit global warming levels to 1.5C above pre-industrial levels have generated considerable interest in predictions of shorter term variations in global-mean temperature. For example, the UK Met Office produces a global-mean temperature forecast for the upcoming year using its decadal forecast system, DePreSys[1] and the Copernicus Climate Change Service (C3S) has created an online tool for estimating the year in which temperatures will pass the 1.5C threshold, based on recent temperature trends[2]. Seasonal forecast models have also been used for this purpose (Tippett and Becker 2024).

It is important to assess the ability of seasonal forecast systems to reproduce past anthropogenically-driven trends in order to build confidence in their future predictions. Generally

speaking, ensembles of coupled, free-running simulations, forced with observed greenhouse gas and aerosol concentrations over recent decades, show a substantial spread in global and regional temperature trends due to natural variability of the climate system or depart from observational trends due to biases in model processes. Similar ensembles of atmosphere-only simulations forced with observed sea surface temperatures (SSTs) show a much smaller spread in temperature trends and lie closer to the observed trends due to the strong constraint of the SST forcing, albeit without the physical realism of coupled ocean–atmosphere dynamics. Seasonal forecasts may be expected to lie somewhere between these two extremes as they are coupled simulations, but, through their intialisation each year, are constrained to internal climate variability at the start time of the forecast (Thomas et al. 2025).

On the other hand, Beverley (2024) found that seasonal forecast models in the C3S archive revert to approximately similar temperature and precipitation trends as coupled climate models, in certain regions, in the first few months after initialisation. This is particularly true in the tropical Pacific with both coupled climate models (Coats and Karnauskas 2017; Seager et al. 2019, 2022) and seasonal forecasts (L'Heureux et al. 2022; Beverley et al. 2023) tending towards an El Niño-like trend. In contrast, observed tropical Pacific SSTs have shown a La Niña-like trend in recent decades (Seager et al. 2019). This apparent discrepancy may result from a combination of an incorrect response to forcing in models and internal climate variability (Wills et al. 2022; Rugenstein et al. 2023; Heede and Fedorov 2023). Such trend errors in Pacific SST may also affect skill in other regions via teleconnections (L'Heureux et al. 2022).

The goal of this paper is to assess how well seasonal forecast models within the C3S archive can reproduce observed near-surface temperature and SST trends. We compare global and regional temperature trends in hindcasts from 8 seasonal forecast systems to observed trends over the common hindcast period 1993–2016. This paper is structured as follows: the data and methodology are described in Section 2 with an assessment of global mean trends in Section 3. The spatial structure of the model trends is compared to observations in Section 4 and a discussion of the results and conclusions are provided in Sections 5 and 6, respectively.

2 | Data and Methods

2.1 | Data

This study examines 8 operational or recently operational prediction systems within the Copernicus Climate Change Service (C3S) archive, listed in Table 1. We assess trends during the common C3S hindcast period (1993–2016) in SST and 2 m temperature at 2–4 month lead-time. These runs are initialised at or before the start of November (valid in December–January–February, DJF) and May (valid in June–July–August, JJA), though some systems use lagged initialisation (see table). In this case, initialisation dates in the previous month are also included. We limit our discussion to DJF and JJA as these are the focus of much academic work on seasonal forecasts and represent the two main extremes for each hemisphere in the seasonal cycle.

The forecast systems are initialised with updated concentrations of carbon dioxide each year, though the details of the concentration pathway followed varies between systems. Only two systems (ECMWF and DWD) have any representation of trends in sulfate aerosol concentrations, other systems use a fixed climatology. Table 1 lists the horizontal resolutions in the ocean and atmosphere for each system, ensemble size, as well as the ocean models. Ocean and atmospheric horizontal resolutions range from around 0.25° to 1° and 36 km to over 100 km, respectively. Five of the systems use the NEMO ocean model and are grouped together in the figures to reflect this. We also note that the atmosphere in most systems is initialised with ERA5 or ERA-Interim, with NCEP CFSv2 and JMA/MRI-CPS2 the two exceptions.

We compare temperature trends in these forecast systems to trends in two reanalysis datasets, ERA5 (Hersbach et al. 2020) and JRA-55 (Kobayashi et al. 2015), observational SST datasets COBE-SST (Ishii et al. 2005) and HadISST2 (Titchner and Rayner 2014), and observational near-surface temperature datasets HadCRUT5 (Morice et al. 2021) and GISTEMP (Lenssen et al. 2019). Note that HadCRUT5 comprises near-surface temperatures over land and SSTs over ocean. Notably, ERA5 and JRA-55 use different SST datasets (HadISST and COBE, respectively) which contributes to differences in their trends (Simmons et al. 2021).

TABLE 1 | Table of operational or recently operational seasonal forecast systems used in this study.

Centre	System	*N*	Ocean model	Res. (atm.)	Res. (ocean)	Atm. init.
CMCC	CMCC-SPS3.5	40	NEMO v3.4	0.5°×0.5°/L46	ORCA025	ERA5
DWD	GCFS 2.1	30	MPIOM 1.6.3	T127/L95	TP04 (0.4°)	ERA5
ECCC	GEM5-NEMO	10	NEMO v3.6	110 km/L85	1°×1°	ERA5
ECMWF	SEAS5	51	NEMO v3.4	36 km/L91	ORCA025	ERA-Interim
JMA	**JMA/MRI-CPS2**	10	MRI.COM v3	TL319/L100	0.25°×0.25°	JRA-55
Meteo-France	System 7	25	NEMO v3.6	TL259/L137	OCRA025	ERA5
NCEP	**CFSv2**	30	GFDL MOM4	T128/L64	0.5°×0.5°	CFSR
UKMO	**GloSea5**	28	NEMO v3.4	N216/L85	ORCA025	ERA-Interim

Note: Columns are (left to right) the modelling centre, system name, the number of ensemble members in hindcasts for that system (*N*), the ocean model, horizontal atmospheric resolution, horizontal ocean resolution and the dataset used to initialise the model atmosphere. Bold, underlined model names indicate that the system uses lagged start dates.

2.2 | Methodology

In the presence of the same external forcing, internal variability of the climate system could lead to different temperature trends over a given period. Observed temperature time series are therefore just one possible realization of what could have happened under the same forcing. On the other hand, seasonal forecasts provide an ensemble of realizations from a similar initial state. For each forecasting system, we create ensembles of trends by randomly selecting one member for each year and calculating the trend for this combination of ensemble members. We repeat this process 1000 times to give an ensemble of 1000 trends. If the forecast system accurately represents the variability of the climate system and its response to forcing, we would expect that the observed trend would almost always lie within this ensemble. We take this approach for both trends in the global mean and at individual grid points to create maps of SST and near-surface temperature trends. Trends are calculated by minimizing the squared error of points with respect to the linear trend line.

3 | Global-Mean Temperature Trends

We begin by examining global-mean trends in near-surface temperature. We focus on the average over 60S–60N as observational coverage is low in polar regions and thus polar surface temperatures in reanalyses are relatively uncertain. The results are broadly similar if temperatures poleward of 60N/S are included, but the spread in trend ensembles is much larger (not shown).

Over the 1993–2016 period, observed global-mean temperatures warm at 0.11–0.17 K/decade and 0.13–0.2 K/decade in DJF and JJA, respectively (Figure 1a,b). It can also be seen that there is considerable year-to-year variability in global mean temperature, particularly in DJF (Figure 1a). There is some reanalysis uncertainty in the magnitude of the warming trends over this period, with ERA5 showing larger trends compared to JRA-55 in both DJF and JJA (Figure 1a,b). Observational datasets, HadCRUT5 and GISTEMP, agree better with the higher magnitude trend in ERA5. This difference in trends could be partly the result of the use of different SST datasets. JRA-55 uses COBE-SSTs, whereas ERA5, HadCRUT5, and GISTEMP use HadISST2, HadISST4, and ERSSTv5 (Huang et al. 2017), respectively.

Bearing in mind this uncertainty, we now assess the temperature trends in the seasonal forecast models. In general, almost all of the forecast systems tend to show a faster rate of warming over this period, relative to the observations for both DJF and JJA (Figure 2a–p). The one exception is JMA/MRI-CPS2, which generally warms more slowly than ERA5, GISTEMP, and HadCRUT5 for both seasons (Figure 2o,p), though JRA-55 lies at the cool end of model trend ensemble members in DJF and in the middle for JJA (Figure 2o,p). On average, the systems warm at 0.21 K/decade

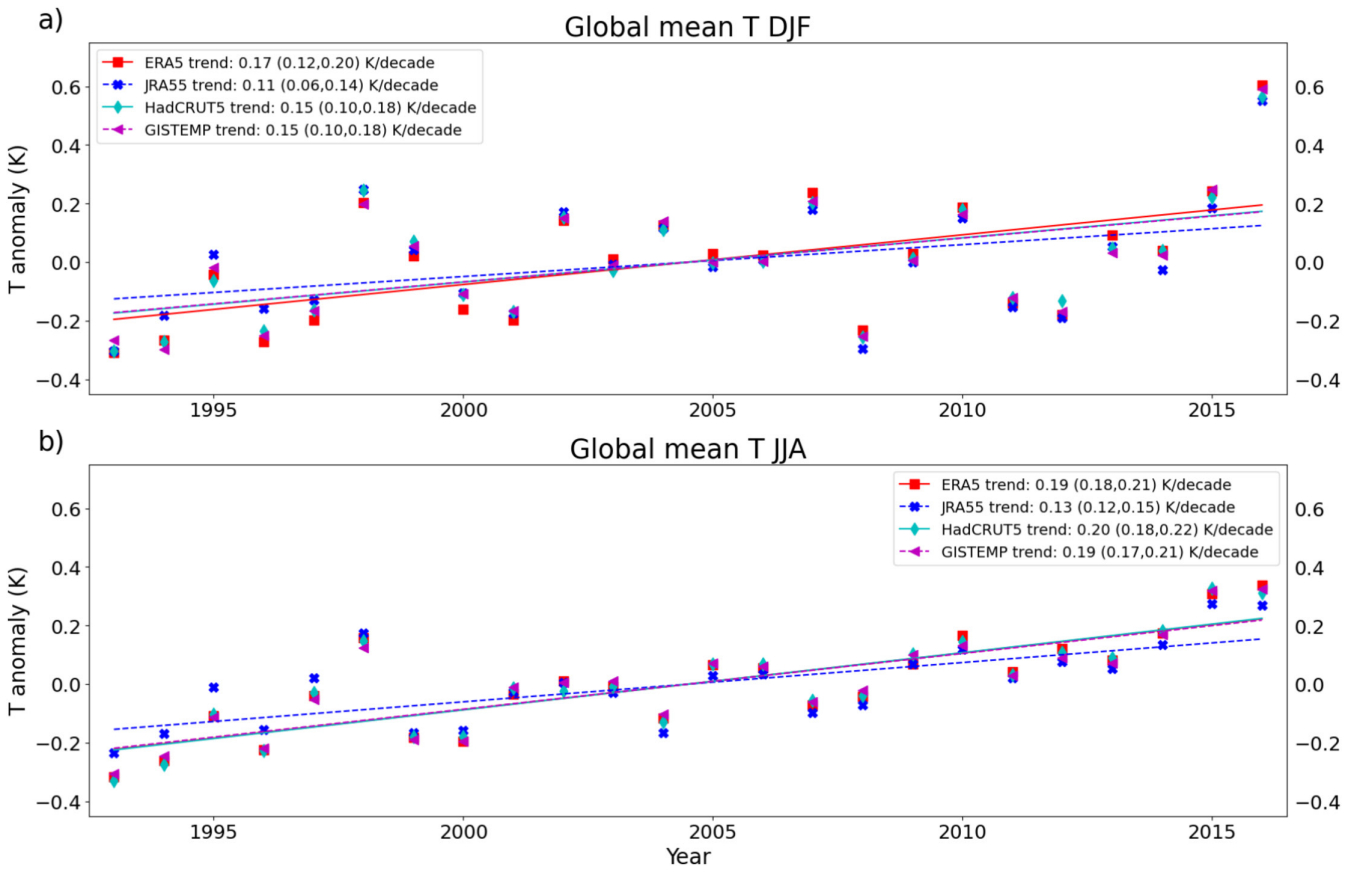

FIGURE 1 | Trends in observation-based datasets of global-mean T2m over the common hindcast period, 1993–2016, for (a) DJF and (b) JJA. T2m anomalies, relative to the 1993–2016 climatology, are plotted for ERA5 (red squares), JRA-55 (blue crosses), HadCRUT5 (cyan diamonds) and GISTEMP (magenta triangles) with linear trend lines also drawn. The values of these trends are shown in the legend with the 95% confidence interval in brackets.

in DJF and 0.22 K/decade in JJA, compared to 0.17 K/decade and 0.19 K/decade in ERA5 for DJF and JJA, respectively.

4 | Spatial Structure of Trends

We now examine trends in area-averaged tropical (20S–20N), northern mid-latitude (20N–60N) and southern mid-latitude (60S–20S) regions and calculate ensembles of trends with the seasonal forecast systems. For DJF, several seasonal forecast models (CMCC SPS3.5, DWD GCFS 2.1, ECCC GEM5-NEMO, ECMWF SEAS5 and Meteo-France system 7) show a similar rate of warming to ERA5, GISTEMP, and HadCRUT5 in the tropics; though others warm at a faster rate (UKMO GloSea5 and NCEP CFSv2). In contrast, almost all show a higher rate of warming in the northern mid-latitudes (Figure 3a). JMA/MRI-CPS2 is the one exception to this, warming at a slower rate than ERA5, GISTEMP, or HadCRUT5 in the tropics and southern mid-latitudes and at a similar rate in the northern mid-latitudes (Figure 3a).

The observed trends in the northern mid-latitudes lie within the spread of most model trend distributions, apart from NCEP CFSv2 (Figure 3a). This is partly because the spread in model trends in the mid-latitudes in DJF is large, which may be due to the influence of internal atmospheric variability affecting temperatures over Northern Hemisphere land. The wider spread of DJF northern mid-latitude temperature trends is consistent with the wide spread of global mean trends in Figure 2a. Overall, the fact that most models (with the exception of JMA/MRI-CPS2) warm faster than observations in DJF appears to be primarily the result of trend differences in the northern mid-latitudes (Figure 3a).

On the other hand, the picture is reversed for JJA. In this case, observed tropical trends are near to or below the lower tails of most model trend distributions, and observed northern mid-latitude trends are above the medians of some models and below the medians of others (Figure 3b). Conversely, JMA/MRI-CPS2 and ECCC GEM5-NEMO warm at roughly the same rate as ERA5 in the tropics, and JMA/MRI-CPS2 warms much more slowly in the northern mid-latitudes than all of the observational datasets. UKMO GloSea5 and NCEP CFSv2 lie at the other end of the spectrum, with the entirety of their trend ensembles warming faster than all four observational datasets, in both seasons and for all regions.

To provide further information about the spatial structure of temperature trends, we show maps of SST trends in HadISST

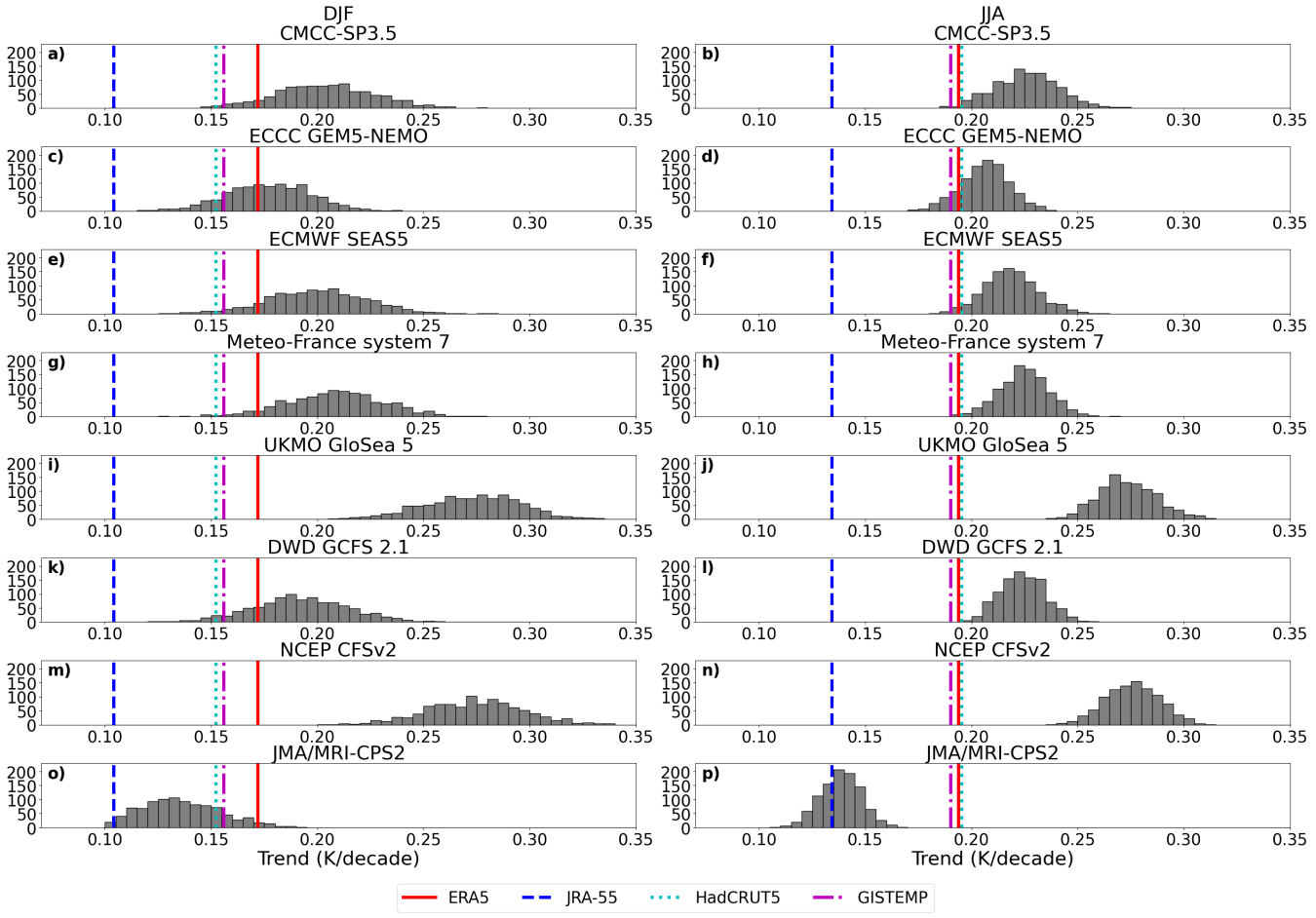

FIGURE 2 | Histograms of ensembles of global-mean trends (60S–60N) for years 1993–2016 for each of the eight seasonal forecast systems in (left) DJF and (right) JJA. Ensembles are constructed by taking a random ensemble member for each year, 1993–2016, and calculating the trend; repeating this process 1000 times. Observed trends calculated from ERA5, JRA-55, GISTEMP, and HadCRUT5 data are also shown by vertical lines, as indicated by the legend.

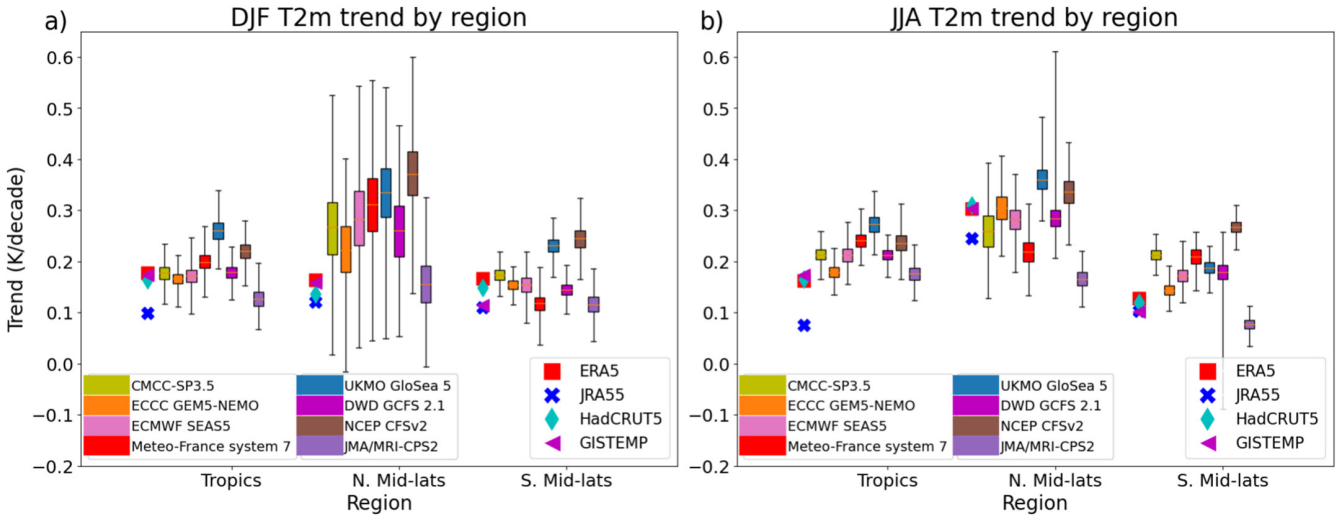

FIGURE 3 | Trends in near-surface temperature averaged within the tropics (20S–20N), northern mid-latitudes (20N–60N) and southern mid-latitudes (60S–20S). Trends are shown by markers for the two reanalysis and two observational datasets, and as boxplots for ensembles of the eight seasonal forecast systems. The boxes indicate the quartiles of the forecast trend distributions, and whiskers are the full range of the trend ensembles.

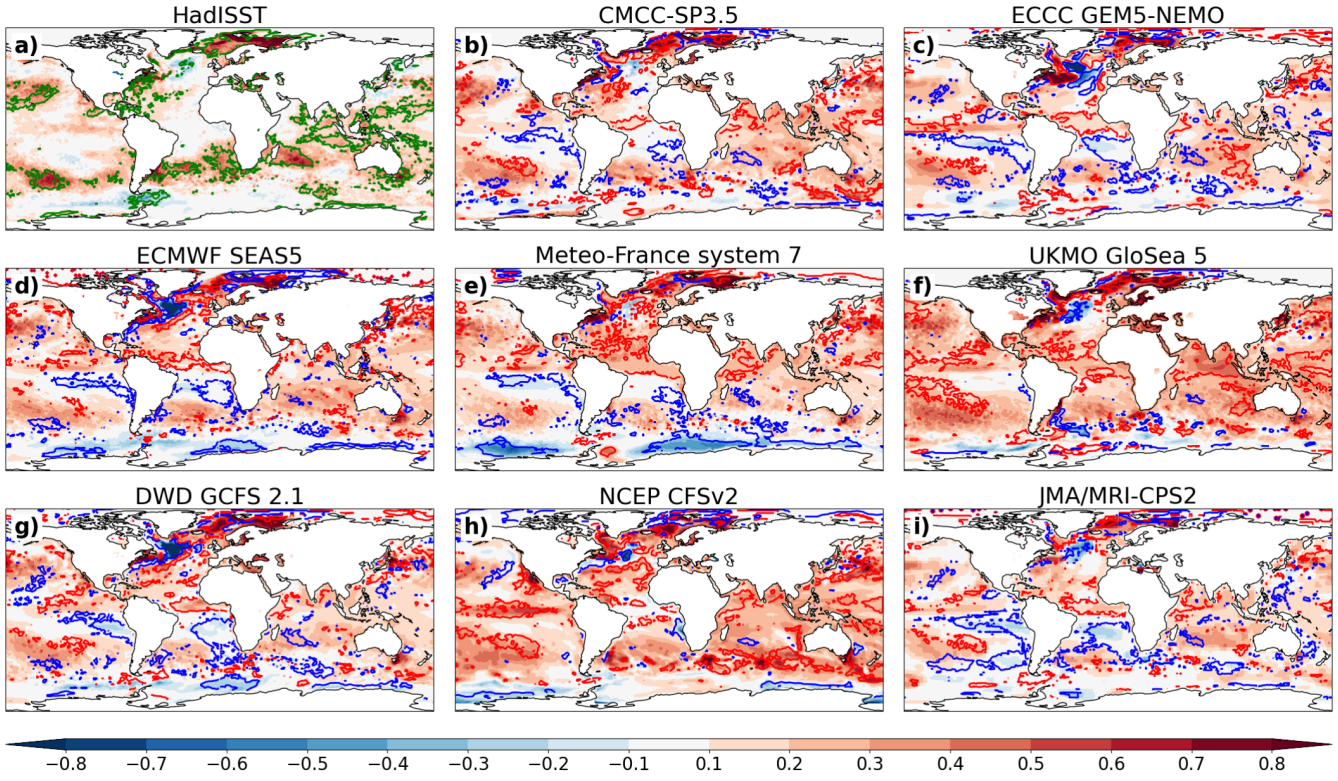

FIGURE 4 | SST trend maps for DJF 1994–2017. Maps are shown of the trend in (a) HadISST and (b–i) ensemble-means of the various seasonal forecast systems. Units are K/decade. The consistency of the model trends with HadISST has been calculated using ensembles of trends. These have been constructed for each model by taking one ensemble member from each year in the study period and calculating the trend, repeating this 1000 times. Unfilled green contours in (a) indicate where trends are statistically significant at the 5% level following a *t*-test. Unfilled red and blue contours in (b–i) show where the ERA5 trend lies below the 2.5th percentile or above the 97.5th percentile of this distribution, respectively.

and the ensemble-mean SST trends in the 8 forecast systems. It can be seen in Figure 4a that the observed DJF SST trends are non-uniform, with the largest warming occurring in the Arctic and parts of the mid-latitudes (in both hemispheres). There are regions of slight cooling around the West Antarctic Peninsula and in the sub-polar North Atlantic (Figure 4a). Other regions, such as the tropical Atlantic and eastern tropical Pacific, show no clear trend. Some of these enhanced warming or weak cooling features may result from natural variability. However, given that the forecasts are initialized with observed data only 2–4 months ahead, we still expect the forecasts to capture these features to some extent, in addition to the forced response to increasing carbon dioxide levels. Note that the pattern of SST trends is similar for the COBE-SST and HadISST datasets (Figure S1).

In contrast to HadISST, almost all models show a warming of the tropical Atlantic (Figure 4b–i) and several show a more substantial warming in the central and western tropical Pacific (Figure 4b,c,e,f,h,i). Conversely, most models show a cooling in the cold upwelling regions, such as subtropical parts of the Pacific and Atlantic Oceans and the Southern Ocean (Figure 4b–e,g,i).

Unfilled blue and red contours in (Figure 4b–i) indicate where the HadISST trend lies below the 2.5th percentile or above the 97.5th percentile of model trend ensembles, respectively. That is, red indicates where the model is warming faster and blue where it is warming more slowly.

In JJA, the largest warming has occurred at high northern latitudes, while the central tropical Pacific and tropical Atlantic show no clear trends in HadISST (Figure 5a). The faster warming of tropical Pacific and Atlantic SSTs in models relative to HadISST is even clearer for JJA than in DJF for this 24-year period (Figure 5b–i). It is notable that in spite of its weaker global warming JMA/MRI-CPS2 also shows a warming in the tropical Pacific and Atlantic (Figure 5i). Two models, UKMO GloSea5 and Meteo-France system 7, also show a much higher rate of warming in the eastern Indian Ocean compared to observations (Figure 5a,e,f).

Examining near-surface temperature in ERA5, the faster warming of the Arctic in DJF is even clearer than for SST (Figure 6a). Variability of temperatures over land tends to be higher than over ocean; hence, there are some large areas of cooling trends, most notably over Siberia, which some have argued is forced by the greater Arctic warming (e.g., Cohen et al. 2012). However, the Siberian cooling trend is not statistically significant over this period (Figure 6a). Note that a similar feature is seen in JRA-55 (Figure S2). The model ensemble mean trends do not tend to show any regions of cooling over land (Figure 6b–i). This is unsurprising given that these are averages over many ensemble members and are less affected by internal variability. However, it is notable that individual members of models show less internal variability than ERA5 over parts of Siberia (Figure S3). Conversely, models show more variability than ERA5 over tropical oceans in JJA (Figure S4).

Overall, there are few regions over land where ERA5 lies outside of the model trend ensembles, likely due to large spread in modeled near-surface temperature trends over land. One exception to this is the larger warming of central Africa in ERA5 (Figure 6a) relative to many of the other model ensembles, which show weak warming or even cooling (Figure 6b–e,h,i).

For boreal summer, ERA5 shows regions of enhanced warming around eastern Europe, Greenland, West Antarctica (in contrast to cooling in DJF, Figure 6a), northern South America, western North America, Japan, and central to northern Africa (Figure 7a). The enhanced eastern European warming is particularly prominent and has been attributed to reductions in sulfate aerosols and low-frequency atmospheric circulation variability (Dong and Sutton 2025). There appear to be more regions over land for which ERA5 shows trends exceeding the bulk of the model ensemble, particularly over central Africa, Greenland, and northern South America (Figure 7b–i).

5 | Discussion on Potential Drivers of Trend Errors

As the signals of climate change emerge from internal variability, it is becoming possible to assess how well climate models

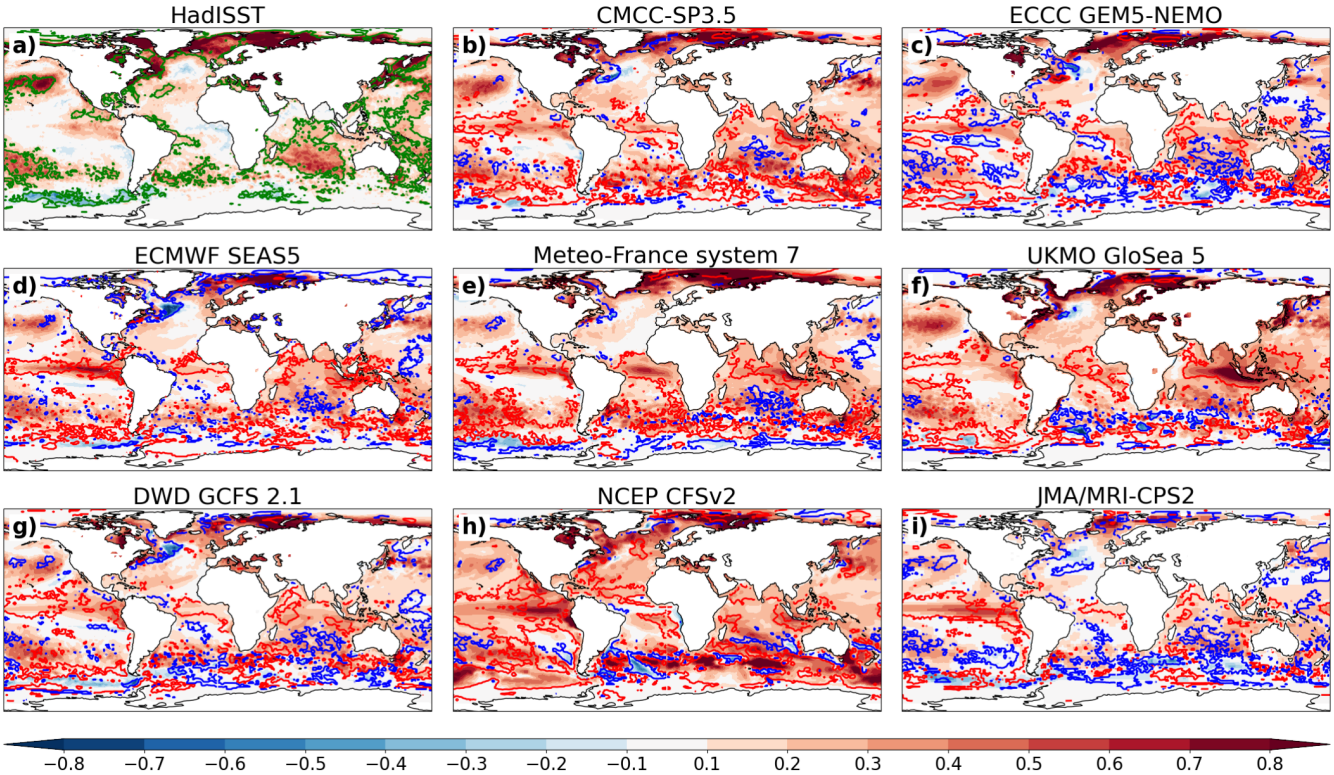

FIGURE 5 | SST trend maps for JJA 1993–2016. As in Figure 4 but for SSTs and for the JJA season.

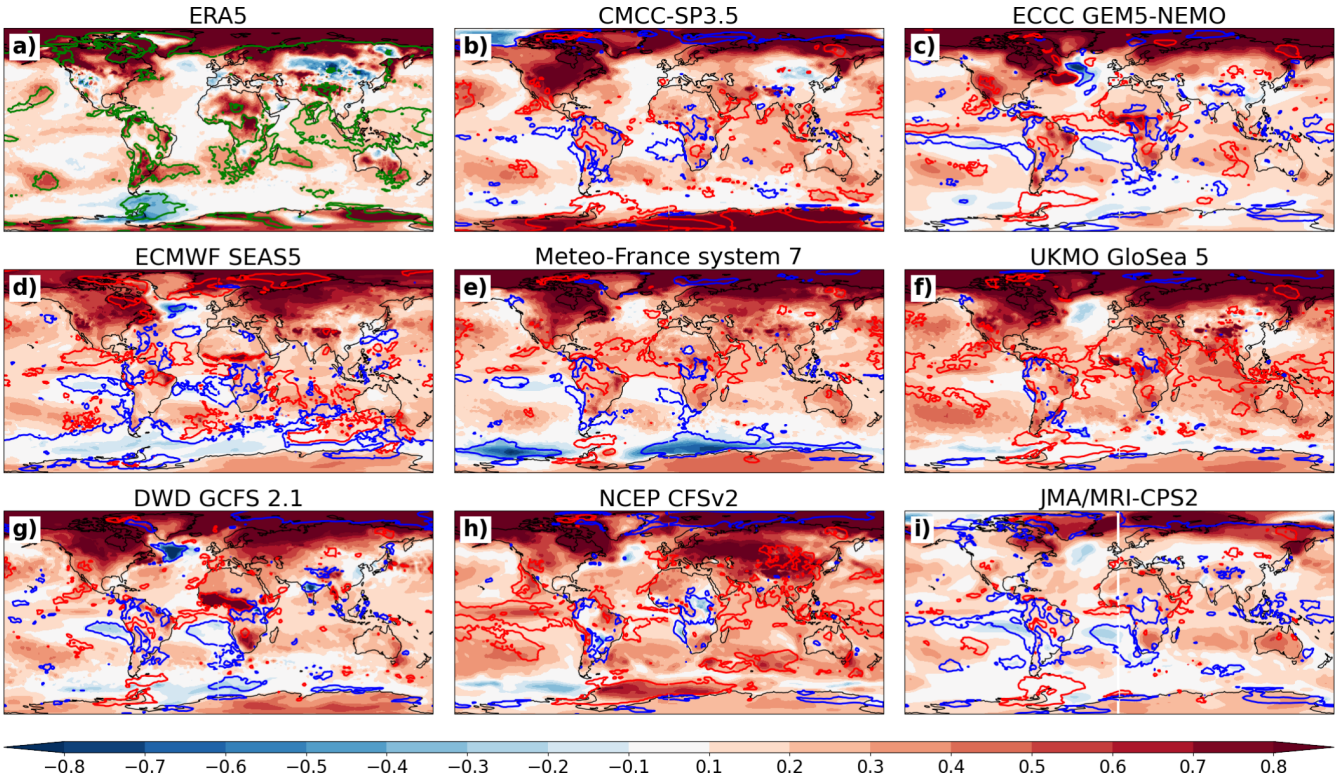

FIGURE 6 | T2m trend maps for DJF 1994–2017. As in Figure 5 but for T2m.

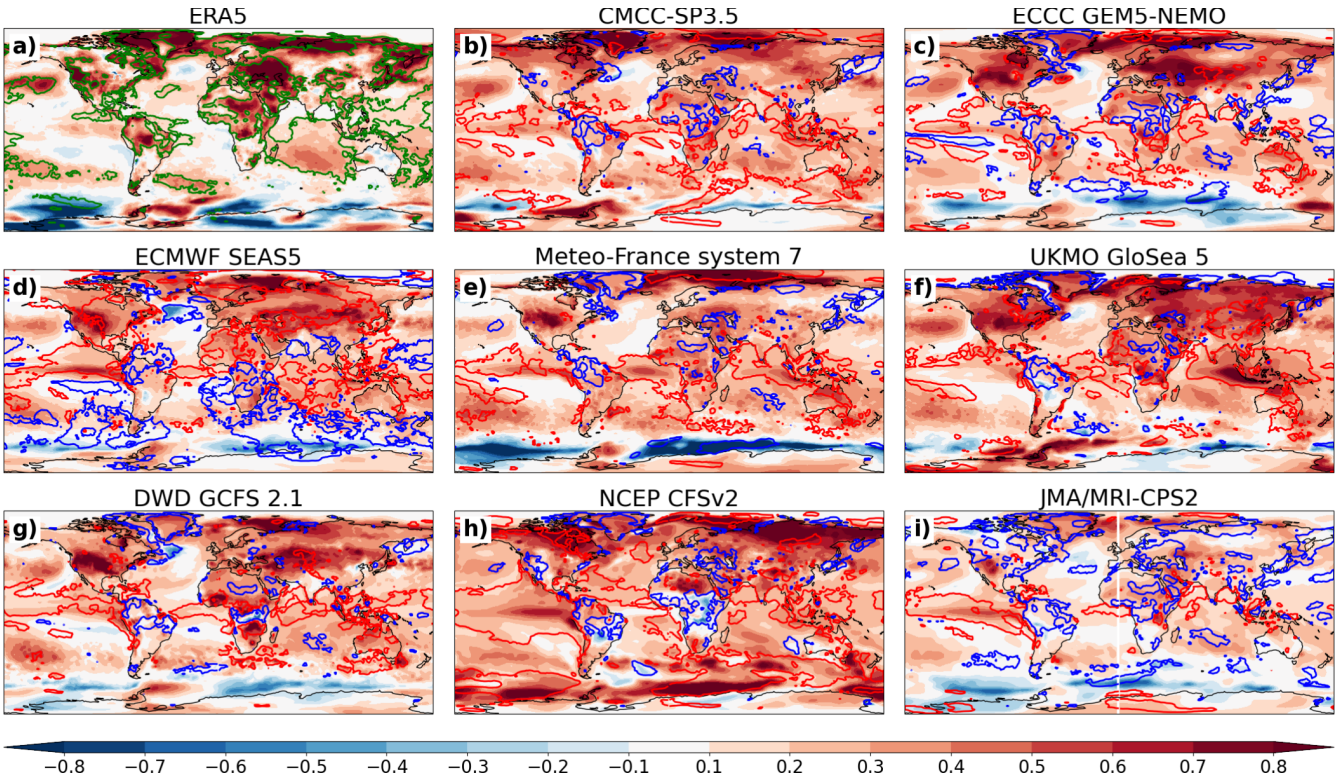

FIGURE 7 | T2m trend maps for JJA 1993–2016. As in Figure 6 but for T2m and JJA.

perform in capturing forced changes (Simpson et al. 2025; Shaw et al. 2024). In general terms, climate models have been found to accurately simulate some changes such as global mean temperature (Hausfather et al. 2020) and total column water vapour (Douville et al. 2022), though some more detailed regional changes like spatial patterns of surface temperature (e.g., Wills et al. 2022; Vautard et al. 2023) and atmospheric circulation (e.g., Blackport and Fyfe 2022; Patterson and O'Reilly 2025) are less well captured.

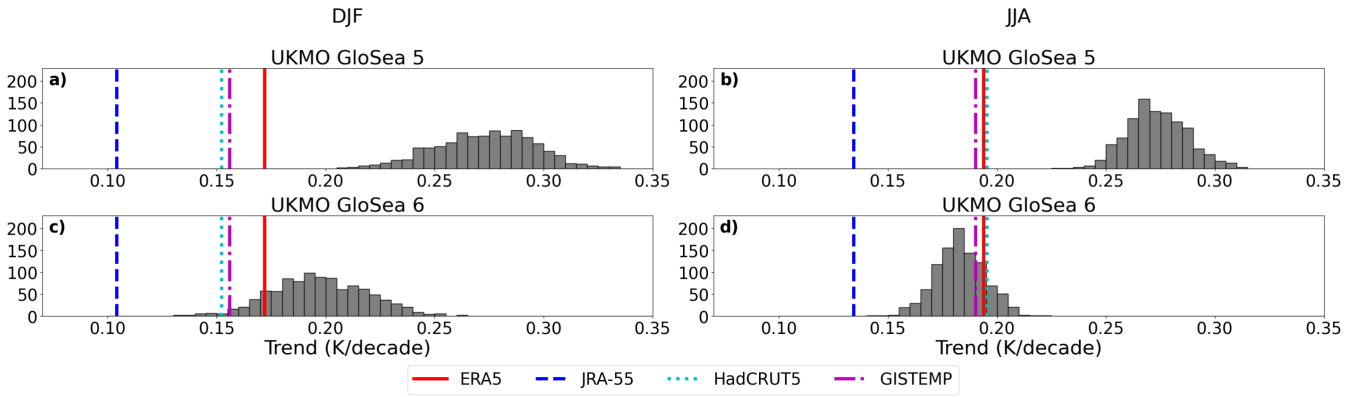

FIGURE 8 | The same as Figure 2 but for UK Met Office systems (a, b) GloSea5 and (c, d) GloSea6.

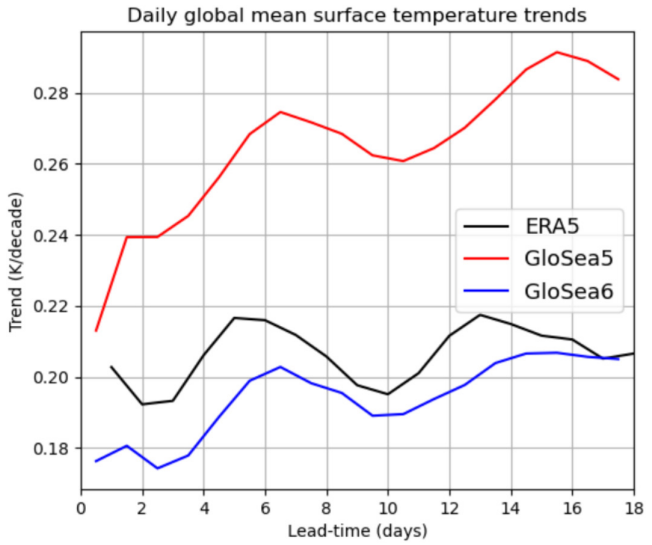

FIGURE 9 | Trends in daily global-mean surface temperature with lead-time in GloSea5 and GloSea6 1993–2016. The daily trend in the ensemble mean, global mean temperature is calculated for initializations in October and November and averaged across these initializations. The daily global mean trend in ERA5 is also shown, averaged over the same days as the forecast systems for the different lead times.

In the case of the seasonal forecast models in this study, there are many factors that could contribute to trend errors, including errors in initialization datasets, the representation of forcing and processes within the models. The fact that all forecast systems but one tend to show a greater warming over the common hindcast period, relative to observations, is suggestive of systematic drivers; though these drivers may differ between seasons (Figure 3). Assessing the relative roles of different sources of trend errors is beyond the scope of this article, but we provide some speculation.

The global mean trend (Figure S5) and regional trend ensembles (Figure S6) for each forecast system are very similar for Month 1 (i.e., the month of initialisation) as for lead-times of 2–4 months. This suggests that either the initialisation may be important or that trend errors develop quickly. To examine potential drivers of trend errors in more depth, we assess trends with lead-time in UKMO GloSea5 and its successor, GloSea6. The motivation for this comparison is that these systems are likely to have many similarities as they largely share the same underlying physical model. However, UKMO GloSea6 warms considerably more slowly than GloSea5 over the common hindcast period in both seasons (Figure 8).

To assess how these differences develop, we plot trends in daily global-mean surface temperature with lead time for boreal winter. Small differences in daily global-mean near-surface temperature trends are present between these systems from the first day of initialization, but daily GloSea5 trends increase quickly with lead time over the first week (Figure 9). The fact that the trend difference takes time to develop suggests that factors other than initialization datasets are important in driving the difference. The rapid growth of this trend difference with time is consistent with Beverley et al. (2023) who found that local trend errors in atmospheric circulation and SSTs developed within the first few weeks after initialization. Understanding why such a difference between GloSea systems exists is complicated by the fact that there are differences in both the underlying models (GloSea5 and GloSea6 are based on HadGEM3-GC2 and HadGEM3-GC3, respectively) and the initialization datasets (the ocean reanalyses differ slightly, while the land surface model, JULES (Walters et al. 2019), is forced with ERA-Interim in GloSea5 and JRA-55 in GloSea6).

For boreal summer, trend differences in global mean temperature between the seasonal prediction systems and observations appear to be more closely related to modeled warming trends in the Tropics and southern mid-latitudes (Figure 3b). Excessive warming in the Tropics may be related to the findings of L'Heureux et al. (2022), who found that seasonal forecast systems consistently show an El Niño-like trend bias, consistent with (Figure 5b–i).

Looking back to the full spread of forecast systems investigated, it is interesting that JMA/MRI-CPS2 is a particular outlier. It is possible that initialisation is important here as the ocean reanalysis, MOVE/MRI.COM-G2, assimilates COBE-SST, which shows noticeably different trends from HadISST over the common hindcast period (Figure S1). It is also possible that the forecast model for JMA/MRI-CPS2 exhibits different characteristics to the other models. For example, the other models may respond more strongly to greenhouse gas forcing.

The faster rate of warming in models relative to observations may also be sensitive to the time period in question. The only

model with a longer hindcast period is ECMWF SEAS5 (1981–2016). Similar to the shorter period, ECMWF SEAS5 shows a faster rate of warming for DJF for this longer period (Figure S7a). However, for JJA, it shows a slower rate of warming than the observations (Figure S7b). This suggests that there may be some sensitivity to the period, at least for JJA.

Notably, the common hindcast period 1993–2016 coincides with the so-called global warming 'hiatus' in which global temperatures warmed more slowly from 1998 to 2012 than in previous decades (Xie and Kosaka 2017). Extensive work on this topic has revealed that the 'hiatus' was largely driven by ocean heat uptake and natural variability of the Pacific Ocean (Kosaka and Xie 2013; Douville et al. 2015; Yan et al. 2016; Medhaug et al. 2017). Nevertheless, it is surprising that models generally overestimated global-mean warming over this period, given their initialization with the observed state of the ocean and atmosphere.

Further analysis would be needed to disentangle the causes of the different trends in models relative to observations. These could include heat budget analyses, ocean–atmosphere heat flux, and assessment of the role of atmospheric circulation (using circulation analogues, e.g., Patterson et al. 2024). Moreover, the role of resolution in the ocean and atmosphere could be investigated as this likely impacts relevant processes such as ocean–atmosphere coupling.

6 | Summary and Conclusions

We have examined near-surface temperature and SST trends in 8 operational or recently operational seasonal forecast systems and compared these to observational trends over the common hindcast period 1993–2016. One important finding is that all but one of the systems tends to warm substantially faster than observations for global-mean temperature, in both DJF and JJA (Figure 2).

Separate analysis of trends in the tropical and northern/southern mid-latitude regions suggests that faster warming in the mid-latitudes in DJF is responsible for the faster global warming in the majority of models, relative to ERA5 (Figure 3). ERA5 shows some regions of cooling or no trend, such as over Siberia in DJF (Figure 6a), likely due to internal atmospheric variability. In JJA, ERA5 shows no warming in the central and eastern tropical Pacific or tropical Atlantic (Figure 5a), whereas most forecast systems do (Figure 5b–i), which may contribute to the elevated global-mean warming in these models.

The faster warming of seasonal forecast models relative to observations has important implications for future predictions of global and regional temperatures. Trend errors may have an effect on model skill and may have a particular influence where the absolute values matter. Moreover, it is often the regional climate extremes that are most impactful for humans and our society. Reliable quantifications of the changing risks of these extremes are thus of foremost importance.

Our analysis shows that there are many open questions as to why the forecast models develop such significant trend errors within the first weeks and months of their forecasts. Further investigations will examine the causes of this overly fast warming to improve future seasonal forecasts. Moreover, running longer hindcast periods will be crucial to assess whether trend errors are robust or specific to this time period.

Author Contributions

Matthew Patterson: methodology, visualization, writing – review and editing, writing – original draft, formal analysis, conceptualization, investigation. **Daniel J. Befort:** writing – review and editing, investigation, conceptualization. **Julia F. Lockwood:** writing – review and editing, investigation, conceptualization. **John Slattery:** investigation, writing – review and editing, methodology, conceptualization. **Antje Weisheimer:** conceptualization, funding acquisition, writing – review and editing, methodology, supervision.

Acknowledgements

We thank the modelling centres and other data centres for making their model and observational temperature datasets freely available. We thank three anonymous reviewers whose thoughtful comments helped to improve this manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

No new data were created as part of this study. The climate data store provides access to seasonal forecast data (https://cds.climate.copernicus.eu/datasets/seasonal-monthly-single-levels?tab=download) and ERA5 (https://doi.org/10.24381/cds.adbb2d47). HadCRUT is available from the Met Office website (https://www.metoffice.gov.uk/hadobs/hadcrut5/), as is HadISST (https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html). GISTEMP is available from the NASA website (https://data.giss.nasa.gov/gistemp/); COBE SST data can be found on the NOAA website (https://psl.noaa.gov/data/gridded/data.cobe.html) and JRA-55 is available from UCAR (https://rda.ucar.edu/datasets/d628001/).

Endnotes

[1] https://www.metoffice.gov.uk/about-us/news-and-media/media-centre/weather-and-climate-news/2024/2025-global-temperature-outlook#:~:text=The%20Met%20Office%20outlook%20for,levels%20for%20the%20first%20time.

[2] https://apps.climate.copernicus.eu/global-temperature-trend-monitor/.

References

Beverley, J. D. 2024. "Climate Model Trend Errors Are Evident in Seasonal Forecasts at Short Leads." *NPJ Climate and Atmospheric Science* 7: 1–13.

Beverley, J. D., M. Newman, and A. Hoell. 2023. "Rapid Development of Systematic ENSO-Related Seasonal Forecast Errors." *Geophysical Research Letters* 50: e2022GL102249. https://doi.org/10.1029/2022GL102249.

Blackport, R., and J. C. Fyfe. 2022. "Climate Models Fail to Capture Strengthening Wintertime North Atlantic Jet and Impacts on Europe." *Science Advances* 8: eabn3112.

Coats, S., and K. B. Karnauskas. 2017. "Are Simulated and Observed Twentieth Century Tropical Pacific Sea Surface Temperature Trends

Significant Relative to Internal Variability?" *Geophysical Research Letters* 44: 9928–9937. https://doi.org/10.1002/2017GL074622.

Cohen, J. L., J. C. Furtado, M. A. Barlow, V. A. Alexeev, and J. E. Cherry. 2012. "Arctic Warming, Increasing Snow Cover and Widespread Boreal Winter Cooling." *Environmental Research Letters* 7: 014007. https://doi.org/10.1088/1748-9326/7/1/014007.

Doblas-Reyes, F. J., R. Hagedorn, T. N. Palmer, and J.-J. Morcrette. 2006. "Impact of Increasing Greenhouse Gas Concentrations in Seasonal Ensemble Forecasts." *Geophysical Research Letters* 33: L07708. https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2005GL025061.

Dong, B., and R. T. Sutton. 2025. "Drivers and Mechanisms Contributing to Excess Warming in Europe During Recent Decades." *NPJ Climate and Atmospheric Science* 8: 1–13.

Douville, H., S. Qasmi, A. Ribes, and O. Bock. 2022. "Global Warming at Near-Constant Tropospheric Relative Humidity Is Supported by Observations." *Communications Earth & Environment* 3: 1–7.

Douville, H., A. Voldoire, and O. Geoffroy. 2015. "The Recent Global Warming Hiatus: What Is the Role of Pacific Variability?" *Geophysical Research Letters* 42: 880–888. https://doi.org/10.1002/2014GL062775.

Hausfather, Z., H. F. Drake, T. Abbott, and G. A. Schmidt. 2020. "Evaluating the Performance of Past Climate Model Projections." *Geophysical Research Letters* 47: e2019GL085378. https://doi.org/10.1029/2019GL085378.

Heede, U. K., and A. V. Fedorov. 2023. "Colder Eastern Equatorial Pacific and Stronger Walker Circulation in the Early 21st Century: Separating the Forced Response to Global Warming From Natural Variability." *Geophysical Research Letters* 50: e2022GL101020. https://doi.org/10.1029/2022GL101020.

Hersbach, H., B. Bell, P. Berrisford, et al. 2020. "The ERA5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146: 1999–2049.

Huang, B., P. W. Thorne, V. F. Banzon, et al. 2017. "Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons." *Journal of Climate* 30: 8179.

Ishii, M., A. Shouji, S. Sugimoto, and T. Matsumoto. 2005. "Objective Analyses of Sea-Surface Temperature and Marine Meteorological Variables for the 20th Century Using ICOADS and the Kobe Collection." *International Journal of Climatology* 25: 865–879. https://doi.org/10.1002/joc.1169.

Kobayashi, S., Y. Ota, Y. Harada, et al. 2015. "The JRA-55 Reanalysis: General Specifications and Basic Characteristics." *Journal of the Meteorological Society of Japan* 93: 5–48.

Kosaka, Y., and S.-P. Xie. 2013. "Recent Global-Warming Hiatus Tied to Equatorial Pacific Surface Cooling." *Nature* 501: 403–407.

Lenssen, N. J. L., G. A. Schmidt, J. E. Hansen, et al. 2019. "Improvements in the GISTEMP Uncertainty Model." *Journal of Geophysical Research: Atmospheres* 124: 6307–6326.

L'Heureux, M. L., M. K. Tippett, and W. Wang. 2022. "Prediction Challenges From Errors in Tropical Pacific Sea Surface Temperature Trends." *Frontiers in Climate* 4: 837483.

Liniger, M. A., H. Mathis, C. Appenzeller, and F. J. Doblas-Reyes. 2007. "Realistic Greenhouse Gas Forcing and Seasonal Forecasts." *Geophysical Research Letters* 34: 2006GL028335. https://doi.org/10.1029/2006GL028335.

Medhaug, I., M. B. Stolpe, E. M. Fischer, and R. Knutti. 2017. "Reconciling Controversies About the 'Global Warming Hiatus'." *Nature* 545: 41–47.

Morice, C. P., J. J. Kennedy, N. A. Rayner, et al. 2021. "An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set." *Journal of Geophysical Research: Atmospheres* 126: e2019JD032361.

Palmer, T. N., and D. L. T. Anderson. 1994. "The Prospects for Seasonal Forecasting—A Review Paper." *Quarterly Journal of the Royal Meteorological Society* 120: 755–793.

Patterson, M., D. J. Befort, C. H. O'Reilly, and A. Weisheimer. 2024. "Drivers of the ECMWF SEAS5 Seasonal Forecast for the Hot and Dry European Summer of 2022." *Quarterly Journal of the Royal Meteorological Society* 150: 4969–4986.

Patterson, M., and C. H. O'Reilly. 2025. "Climate Models Struggle to Simulate Observed North Pacific Jet Trends, Even Accounting for Tropical Pacific Sea Surface Temperature Trends." *Geophysical Research Letters* 52: e2024GL113561.

Patterson, M., A. Weisheimer, D. J. Befort, and C. O'Reilly. 2022. "The Strong Role of External Forcing in Seasonal Forecasts of European Summer Temperature." *Environmental Research Letters* 17, no. 10: 104033.

Rugenstein, M., S. Dhame, D. Olonscheck, R. J. Wills, M. Watanabe, and R. Seager. 2023. "Connecting the SST Pattern Problem and the Hot Model Problem." *Geophysical Research Letters* 50: e2023GL105488. https://doi.org/10.1029/2023GL105488.

Seager, R., M. Cane, N. Henderson, D.-E. Lee, R. Abernathey, and H. Zhang. 2019. "Strengthening Tropical Pacific Zonal Sea Surface Temperature Gradient Consistent With Rising Greenhouse Gases." *Nature Climate Change* 9: 517–522.

Seager, R., N. Henderson, and M. Cane. 2022. "Persistent Discrepancies Between Observed and Modeled Trends in the Tropical Pacific Ocean." *Journal of Climate* 35, no. 14: 4571.

Shaw, T. A., J. M. Arblaster, T. Birner, et al. 2024. "Emerging Climate Change Signals in Atmospheric Circulation." *AGU Advances* 5: e2024AV001297. https://doi.org/10.1029/2024AV001297.

Simmons, A., H. Hersbach, J. Munoz-Sabater, et al. 2021. "Low Frequency Variability and Trends in Surface Air Temperature and Humidity From ERA5 and Other Datasets." European Centre for Medium-Range Weather Forecasts.

Simpson, I. R., T. A. Shaw, P. Ceppi, et al. 2025. "Confronting Earth System Model Trends With Observations." *Science Advances* 11: eadt8035.

Thomas, R., T. Woollings, and N. Dunstone. 2025. "The Role of Internal Variability in Seasonal Hindcast Trend Errors." *Journal of Climate* 1, no. aop. https://doi.org/10.1175/JCLI-D-24-0367.1.

Tippett, M. K., and E. J. Becker. 2024. "Trends, Skill, and Sources of Skill in Initialized Climate Forecasts of Global Mean Temperature." *Geophysical Research Letters* 51: e2024GL110703. https://doi.org/10.1029/2024GL110703.

Titchner, H. A., and N. A. Rayner. 2014. "The Met Office Hadley Centre Sea Ice and Sea Surface Temperature Data Set, Version 2: 1. Sea Ice Concentrations." *Journal of Geophysical Research: Atmospheres* 119: 2864–2889. https://doi.org/10.1002/2013jd020316.

Vautard, R., J. Cattiaux, T. Happé, et al. 2023. "Heat Extremes in Western Europe Increasing Faster Than Simulated due to Atmospheric Circulation Trends." *Nature Communications* 14: 6803.

Walters, D., A. J. Baran, I. Boutle, et al. 2019. "The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 Configurations." *Geoscientific Model Development* 12: 1909–1963.

Weisheimer, A., and T. N. Palmer. 2014. "On the Reliability of Seasonal Climate Forecasts." *Journal of the Royal Society Interface* 11: 20131162.

Wills, R. C. J., Y. Dong, C. Proistosecu, K. C. Armour, and D. S. Battisti. 2022. "Systematic Climate Model Biases in the Large-Scale Patterns of Recent Sea-Surface Temperature and Sea-Level Pressure Change."

Geophysical Research Letters 49: e2022GL100011. https://doi.org/10.1029/2022GL100011.

Xie, S.-P., and Y. Kosaka. 2017. "What Caused the Global Surface Warming Hiatus of 1998–2013?" *Current Climate Change Reports* 3: 128–140. https://doi.org/10.1007/s40641-017-0063-0.

Yan, X.-H., T. Boyer, K. Trenberth, et al. 2016. "The Global Warming Hiatus: Slowdown or Redistribution?" *Earth's Future* 4: 472–482. https://doi.org/10.1002/2016EF000417.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** asl1316-sup-0001-Supinfo.