Summertime circumglobal Rossby waves in climate models: Small biases in upper-level circulation create substantial biases in surface imprint

Fei Luo[1,2], Frank Selten[2], Kathrin Wehrli[3], Kai Kornhuber[4,5], Philippe Le Sager[2], Wilhelm May[6], Thomas Reerink[2], Sonia I. Seneviratne[3], Hideo Shiogama[7], Daisuke Tokuda[8], Hyungjun Kim[8,9,10], and Dim Coumou[1,2]

[1]Institute for Environmental Studies (IVM), VU University Amsterdam, Amsterdam, Netherlands
[2]Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands
[3]Institute for Atmospheric and Climate Science, Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland
[4]Earth Institute, Columbia University, New York, United States
[5]Lamont-Doherty Earth observatory, Columbia University, New York, United States
[6]Centre for Environmental and Climate Science (CEC), Lund University, Lund, Sweden
[7]Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan
[8]Moon Soul Graduate School of Future Strategy, Korea Advanced Institute of Science and Technology, Daejeon, Korea
[9]Department of Civil and Environmental Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea
[10]Institute of Industrial Science, University of Tokyo, Tokyo, Japan

Correspondence to: Fei Luo (fei.luo@vu.nl)

Abstract. In boreal summer, circumglobal Rossby waves can promote stagnating weather systems that favor extreme events like heatwaves or droughts. Recent work highlighted the risks associated with amplified Rossby wavenumber 5 and 7 triggering simultaneous warm anomalies in specific agricultural breadbasket regions in the Northern Hemisphere. These types of wave patterns thus pose a potential threat to human health and ecosystems. The representation of such summertime wave events and their surface imprints in general circulation models (GCMs) has not been systematically analyzed. Here we validate three state-of-the-art global climate models (EC-Earth, CESM, and MIROC), quantify their biases and provide insights into the underlying physical reasons for the biases. To do so, the ExtremeX experiments output data were used, which are (1) historic simulations of a freely running atmosphere with prescribed ocean, and experiments that additionally (2) nudge toward the observed upper-level horizontal winds in the atmosphere, (3) prescribe soil moisture conditions, or (4) both. The experiments are used to trace the sources of the model biases to either the large-scale atmospheric circulation or surface feedback processes. We show that while the wave position and magnitude is generally represented well during high amplitude (> 1.5 s.d.) wave-5 and wave-7 events, the imprint on surface variables is substantially underestimated: typically, by a factor of 1.5 in correlation and normalized standard deviations (n.s.d.) for near-surface temperature and mean sea level pressure. For precipitation, it is also a factor of 1.5 for n.s.d. but 2 for correlation. The correlations and n.s.d. for surface

Deleted: compared to ERA5 reanalysis data

variables do not improve if only the soil moisture is prescribed, but are corrected almost entirely when the upper-level atmospheric circulation is nudged. When applying both soil moisture prescription and nudging of the upper-level atmosphere, both the correlation and n.s.d. values are quite similar to the experiments where only the atmosphere component is nudged. Hence, the near-surface biases can be substantially improved when nudging the upper-level circulation providing evidence that relatively small biases in the models' representation of the upper-level waves can strongly affect associated temperature and precipitation anomalies.

1 Introduction

The past decade has witnessed a series of unprecedented boreal summer weather extreme events around the globe such as the 2010 Russian heatwave, 2012 North American heatwave, and the record breaking heatwaves of 2015, 2018 and 2019 in Europe (Barriopedro et al., 2011; Kornhuber et al., 2019; Krzyżewska and Dyer, 2018; Wang et al., 2014; Huntingford et al., 2019; Xu et al., 2021). Some of these events also happened simultaneously with other types of extremes such as the persistent Russian heatwave and Pakistan flood in 2010 July and August (Lau and Kim, 2012; Martius et al., 2013). These persistent weather extremes can have disastrous impacts on human health and societies such as wide spread crop failure, infrastructure damage and properties loss, especially when they co-occur (Zscheischler et al., 2018). Persistent weather extremes are often induced by quasi-stationary Rossby waves. For instance, Recurring Rossby Waves Packets (RRWPs) can lead to cold spells in winter and hot spells in summer (Röthlisberger et al., 2019). Several other studies have also identified that amplified circumglobal waves favor the occurrence of weather extremes in certain regions (Screen and Simmonds, 2014; Kornhuber et al., 2020). Specifically, in summer wave-5 (Ding and Wang, 2005; Kornhuber et al., 2020) and wave-7 (Kornhuber et al. 2019, 2020) have preferred phase positions and thereby favor simultaneous extremes in major breadbasket regions (Kornhuber et al., 2020).

Several mechanisms can promote quasi-stationary Rossby waves including strong convective forcing from monsoons (Di Capua et al., 2020), extratropical sea surface temperature (SST) anomalies (McKinnon, et al., 2016; Vijverberg, et al., 2020), soil moisture anomalies (Teng and Branstator, 2019), waveguide effect (Hoskins and Ambrizzi, 1993), and wave-resonances (Petoukhov et al., 2013, 2016; Kornhuber et al., 2017; Thomson and Vallis, 2018). Recent work by Di Capua et al. (2020) found that the latent heat release during the Indian Summer Monsoon initiates a circumglobal teleconnection pattern, which reflects a wave-5 type pattern in the northern mid-latitudes. Extratropical SSTs can interact with atmospheric waves creating quasi-stationary atmospheric Rossby waves favorable for e.g. hot days in the eastern United States (McKinnon et al., 2016). Moreover, waves can be excited by reduced soil moisture and then maintained by waveguides in the Northern Hemisphere mid-latitudes (Teng et al., 2019), which could result in high-amplitude wave events occurring more often. Quasi-Resonant Amplification (QRA) theory suggests that synoptic scale Rossby waves can be trapped within the mid-latitude waveguides, where they can get amplified given suitable forcing conditions (Petoukhov et al. 2013). Since the wave's energy is not lost

Deleted: they are defined as compound

Deleted: a

Deleted: effects

Deleted: (Hoskins and Ambrizzi 1993)

via meridional dispersion, waves tend to propagate over long longitudinal distances and can sometimes form circumglobal wave pattern (Hoskins and Ambrizzi, 1993; Branstator, 2002b; Teng and Branstator, 2019).

75

Climate models are important tools for process understanding and assessment of future climate risks. However, most of previous studies that link specific circumglobal Rossby wave patterns to regional extreme events are based on reanalysis/observational data. Although studies such as Garfinkel et al. (2020); Wills et al. (2019) have analyzed waves in models, their focus is not on summer and, also, they have not explored the phase-locking behavior of amplified, quasi-

80 stationary Rossby waves. Furthermore, most studies have not analyzed waves above wave number 6. Studies by Branstator et al. (2002 & 2017) have also looked into models but focus on seasonal means and/or winter. Thus, a multi-model validation study of quasi-stationary Rossby waves in boreal summer is still lacking. Another key issue here is the general underestimation in atmospheric blocking (Davini and D'Andrea 2020) due to the misrepresentation of the atmospheric circulation and processes that maintain blocking; this reduces the reliability of future model projections (Scaife et al. 2010;

85 Shepherd 2014). A recent study by Davini and D'Andrea (2020) analyzed the representation of both winter and summer time blocking frequencies in models from the Coupled Model Intercomparison Project Phase 3 (CMIP3, 2007), CMIP5 (2012), and CMIP6 (2019). Although biases in CMIP6 models were reduced by 50% compared to CMIP3 models in some key regions like Europe, the biases still remain, hence: 1: even CMIP6 models cannot truthfully reproduce wintertime blocking frequencies in Europe, 2: CMIP models are not able to capture the observed strong and significant increase in summertime

90 blocking activities over Greenland (Davini and D'Andrea 2020).

Furthermore, it should also be noted that although it's not examined in detail in this study, extreme events can also occur at other wavenumbers apart from wave-5 and 7. High amplitude slow moving planetary waves are associated with the duration of surface weather conditions. For example, QRA mechanism is able to explain the generation of circumglobal Rossby

95 waves with wavenumbers 6 to 8 in Northern Hemisphere (Coumou et al., 2014). The evidence for QRA in the Southern Hemisphere (SH) is also found to exist in wavenumber 4 and 5 (Kornhuber., 2017). Whether climate models can also reproduce this wide range of wave numbers and their associated surface anomalies needs to be tested.

Thus, to increase confidence in future projections of extreme summer weather, a proper validation of state-of-art climate

100 models in their representation of circumglobal Rossby waves in summer is essential. Both the upper-level dynamical characteristics, in terms of amplitude and phase position, as well as the waves' impact on surface weather is important. Here we systematically validate the representation of summertime Rossby waves in three state-of-the-art climate models, focusing on wave-5 and 7, their phase-locking behavior and surface anomalies. Further, we remove the bias in only one or two components of the model by nudging and/or prescribing to reanalysis, to understand the origin of model biases.

105

This paper aims at addressing the following questions:

1: Can models capture the key characteristics of high-amplitude circumglobal waves in summer?

2: What are the near-surface temperature, precipitation, and mean sea level pressure anomalies from such waves and how do they compare to observations?

3: Do potential model biases originate from the atmospheric circulation or land surface-feedbacks?

2 Data and Methods

2.1 ExtremeX experiment

We use simulation output from three Earth System Models (ESM) that participated in the ExtremeX modeling experiment (Wehrli et al., 2021, in review): European Community Earth System Model version 3.3.1 (EC-Earth 3.3.1; Döscher et al., 2021, in review), Community Earth System Model version 1.2 (CESM1.2; Hurrell et al., 2013), and Model for Interdisciplinary Research on Climate version 5 (MIROC5; Watanabe et al., 2010). The configuration of CESM and MIROC was used for CMIP5 (CMIP5; Taylor et al., 2012), whereas EC-Earth is the latest 3rd generation model used for CMIP6 (Eyring et al., 2016). The ExtremeX modeling experiments were designed to disentangle the influence from atmospheric dynamics vs. soil moisture feedback on extreme events such as heatwaves, droughts, and other extremes. By nudging either the upper-level atmosphere or prescribing the soil moisture state, or both, the individual effects can be compared across different models. Details on the experimental set-up and atmospheric nudging approach are described in a recent study where five individual heatwaves in the period of 2010 – 2016 were examined (Wehrli et al., 2019).

2.2 Model data output

Here we use four out of five sets of simulations from ExtremeX, which are all run in Atmospheric Model Intercomparison Project (AMIP) (Gates et al., 1999) style with prescribed monthly mean SSTs and sea-ice, but differing in whether the other components are free or constrained. The experiments are run with: (1) interactive atmosphere and soil moisture as reference (AISI), (2) nudged atmosphere (mostly above 700 hPa) but interactive soil moisture (AFSI), (3) nudged atmosphere with prescribed soil moisture (AFSF), and (4) interactive atmosphere with prescribed soil moisture (AISF). The experiment period extends from January 1979 to December 2016 for both EC-Earth and CESM and till December 2015 for MIROC. Overall output is provided 6 hourly on different model grids (in number of grid points longitude x latitude): EC-Earth (512 x 256), CESM (288 x 192), and MIROC (256 x 128). There are five ensemble members for the AISI and AISF runs for the

whole period. However, for AFSI and AFSF, only one simulation for each experiment was used. All model and reference data are regridded to the same resolution (256 x 128) for comparisons.

2.3 Atmospheric nudging

To constrain the natural variability in large-scale atmospheric circulation, a grid-point nudging method was implemented in the AFSI and AFSF experiments (Jeuken et al., 2016). This approach forces the atmospheric large-scale circulation by introducing a tendency term in the wind both zonally and meridionally. This added tendency term is calculated between the simulations and reference dataset. Kooperman et al. (2012) demonstrated that when the horizontal wind is nudged towards a reference state, the impact of natural variability is substantially minimized. The strength of the nudging can be modified by a relaxation time scale, which was chosen to be 6h following other studies (Kooperman et al., 2012). All models use 6-hourly wind field data from the ERA-Interim reanalysis as reference data (Dee et al., 2011). The atmospheric nudging vertical profile (See Appendix Figure B1) shows that nudging starts around 700hPa but only with a very weak nudging strength. The nudging strength increases gradually in the upward direction and full nudging is only applied above ca. 400hPa. Thus, the mid-to-upper atmosphere is nudged, and it is important to note that the planetary boundary layer is free to adjust in the nudged experiments.

2.4 ERA5 reanalysis data

For the study period 1979 to 2016, weekly meridional wind data at 250 hPa (v250), near-surface temperature (t2m), and mean sea level pressure (mslp) are taken from the ERA5 reanalysis for the summer months June, July, and August (JJA) (Hersbach et al., 2020). For precipitation (prcp), land-only data is used from bias-adjusted ERA5 (WFDE5_CRU, Cucchi et al., 2020). Also, weekly t2m data is detrended to its climatological mean (1979 – 2016) values of that week.

2.5 Extracting circumglobal waves and phase-locking analysis

High amplitude wave events are selected based on the Fourier transformation analysis of weekly-mean v250 averaged over 35N to 60N, both in ERA5 and models analogous to previous studies (Kornhuber et al. 2019, 2020). Wave events are identified as those weeks with wave amplitudes higher than 1.5 standard deviations (s.d.) of the climatology calculated from 494 weeks (38 years times 13 summer weeks per year) for ERA5. Since the AISI and AISF runs both have 5 ensemble members, the weeks are as follows: EC-Earth 494 x 5, CESM 494 x 5 and MIROC 481 x 5. For AFSI and AFSF, only one member is used for each model. Then, the composite surface imprints of near-surface temperature, precipitation, and mean sea level anomalies were obtained from those wave event periods. The probability density functions for phase positions are obtained for waves 5 to 8 for the weeks with identified high-amplitude events.

2.6 Model bias definition

To isolate the source of the models biases in surface imprint anomalies to either the upper-level atmosphere or to the land surface component, we define the total bias (B_tot), bias from the atmospheric circulation (B_atm), bias from land-atmosphere interactions (B_land), and the remaining residual bias (B_res).

In the AISI experiment, both the atmosphere and land surface component are allowed to interact and evolve freely, and this experiment thus defines the total bias:

B_tot = AISI – ERA5

When prescribing soil moisture in AISF, we assume that the land bias is removed and only the bias from the atmosphere acting upon near-surface variables remains. Thus:

B_atm = AISF – ERA5

In contrast, when nudging the upper-level atmosphere, the upper-level circulation pattern is constrained in the model and thus the bias arises from land-atmosphere interactions.

B_land = AFSI – ERA5

When nudging the upper-level atmosphere and prescribing soil moisture, the model biases are expected to be strongly reduced with only a residual bias remaining:

B_res = AFSF – ERA5

3 Results

3.1 Climatology of summertime Rossby waves

We first assess whether the climate models are able to represent the mean state in terms of wave amplitude and variability for wavenumbers 1-10. Figure 1 compares wave spectra for wavenumbers 1 to 10 from the AISI experiment with those of the ERA5 reanalysis. Overall, the wave amplitudes, regardless of wave numbers and models, are reasonably well reproduced with errors in the model climatology ranging from 5% (wave-10) to 12% (wave-3). This also applies to the variance in wave activity as given by the whisker bars for each model at different wavenumbers. For all models, the wave amplitudes and

200 variabilities follow the same behavior with increasing values from wave number 1 to 5 and decreasing values till wave-10 thereafter. ERA5 shows the peak for both the wave amplitude and variance at wavenumber 6, which might suggest a systematic bias in the models, or alternatively it might be an under-sampling in ERA5.

3.2 Wave phase-locking behaviors

Following Kornhuber et al (2020), we use 1.5 s.d. above the mean wave amplitude as a threshold to define high amplitude
205 wave events to analyze phase-locking behavior of high amplitude waves 4-8. The phase positions of high amplitude events are shown in Fig. 2. It is clear that ERA5 data has inherent phase-locking properties, especially for waves 5 and 7, which remains consistent with the work from Kornhuber et al. (2019) where a different reanalysis data NCEP-NCAR (Kalnay et al., 1996) was used. Also, in the NCEP-NCAR reanalysis data, waves 6 and 8 do not really show a preferred phase position. In our experiments, across all three models, strong phase-locking behavior (a single symmetrical peak in the
210 probability density function) can clearly be observed for wave-7. For wave-5, two models (CESM and MIROC) show phase locking that is comparable to ERA5, however EC-Earth underestimates the peak in the probability density function.

ERA5 shows no phase-locking for wave 6 and wave 8 but only a mild preference for some phase positions. The models capture this, with only MIROC showing fairly pronounced phase-locking behavior for wave 6. Detailed histogram
215 comparisons between the models and ERA5 can be found at Fig. B2.

3.3 High-amplitude wave events and their surface imprint

Let's recall that when summertime amplified circumglobal Rossby waves get locked in their favorite positions, especially for wave-5 and wave-7, they favor a prolongation of simultaneous surface warm anomalies. As we find that the preferred phase position of circumglobal waves 5 and 7 is reasonably well represented in models, we next analyze high-amplitude events
220 (i.e. exceeding 1.5 s.d.) in more detail. The wave event occurrence is calculated as the number of weeks selected as high amplitude wave event divided by the total number of weeks then times one hundred percent. The occurrences of JJA wave-5 and wave-7 events during 1979 to 2016 for ERA5 are 8.1% and 7.1%. The values are quite comparable to the results calculated from the models: 7.7% (EC-Earth), 8.0% (CESM), and 7.9% (MIROC) for wave-5, whereas the occurrences are 8.1% (EC-Earth), 7.8% (CESM), and 8.0% (MIROC) for wave-7. Figure 3 and 4 show the upper-level circulation (v at 250
225 hPa), near-surface temperature anomaly (t2m), precipitation anomaly (prcp), and sea level pressure anomaly (mslp) signals during such high-amplitude events in ERA5 (a) and the 3 climate models (b-d). All the anomalies are calculated as the differences between the selected event period and the full period summertime climatology mean. Here we present the associated surface anomaly imprints during those high amplitude waves-5 and wave-7 events across all models for the AISI experiment with free-running atmosphere and soil moisture modules (Fig. 3 & Fig. 4). Later we extended our analysis for
230 high amplified wave events also for wavenumbers 4, 6, and 8 (See Appendix Fig. B3 to Fig. B5). The results imply that the model is able to reproduce summertime surface anomalies associated to different wavenumber events, thus confirming the

model's ability to cover a wide range of circumglobally Rossby waves in terms of wave numbers 4 to 8. In Fig. 3 and Fig.4 the v250 field is shown in absolute values, together with surface variables in anomalies with respect to climatology mean. The same analysis is also carried out for taking the anomaly of v250 during wave 5 and 7 events (see Appendix Fig. B10 and B11). By comparing Fig.3 and Fig.4 to Fig. B10 and Fig. B11, the observation can be obtained that the spatial patterns of v250 do not differ much.

Furthermore, Fig. 3 and Fig. 4 with significant tests at confident level of 95%, as well as False Discovery Rate (FDR) method (Benjamini & Hochberg, 1995) were applied in Fig. B12 and Fig. B13. Areas with highlighted fuchsia color are the locations passed the significant tests. Considering the spatial coverage of the significant areas, the conclusion can be drawn that, variables such as v250 and t2m during the wave 5 and 7 events are significantly different from the rest period.

Additionally, to quantify the bias of all models and visualize how close the models are to the ERA5 reanalysis data, a Taylor Diagram (Taylor 2001) is constructed for both wave-5 (Fig.5) and wave-7 (Fig.6). A Taylor Diagram presents three key statistics in a single plot: the Pearson correlation between the observed and modeled spatial pattern; the centered Root Mean Square Error (RMSE) of the modeled field; and the normalized spatial standard deviation of the modeled field. Thus, v250 during the wave-5 and wave-7 events, as well as t2m, prcp, and mslp anomalies from the different model experiments with respect to ERA5 reanalysis data, are plotted in the Taylor Diagram.

During the wave-5 events, all models are able to capture the mean upper-level circulation patterns with correlations of the v250 of 0.86 (EC-Earth), 0.95 (CESM), and 0.88 (MIROC) (Fig.5 & Table A1). This is consistent with our findings from Fig. 1 and 2. In terms of the magnitude of the wind speed anomalies, CESM and MIROC have similar strengths compared to ERA5 data, whereas the signal from EC-Earth is weaker. The n.s.d. for EC-Earth is 0.70, for CESM is 1.04. and for MIROC is 1.24. This also holds for surface temperature imprints during wave-5 events, as all models are able to reproduce the patterns found in ERA5, such as the continental-scale patterns of positive and negative signals for central North America (+), western Europe (-), and central Europe (+). But the strength of the patterns is weaker in EC-Earth, especially for eastern Eurasia. However, the correlations of the near-surface temperature anomalies are substantially smaller for EC-Earth (0.55) and MIROC (0.48), but not so much for CESM (0.81). As for precipitation, all model correlations are below 0.50 with MIROC being the lowest (0.18), 0.46 for EC-Earth, and 0.43 for CESM. The correlation values for sea level pressure in models vary from 0.52 (MIROC), 0.58 (EC-Earth) to 0.80 (CESM). As for the multi-model mean (MMM) s.d., there is a decline from v250 (0.99) to the surface variables t2m (0.71), mslp (0.69) and prcp (0.63). Both reanalysis data and models show strong positive anomalies in sea level pressure in the eastern basin of the Atlantic Ocean (west coast of Europe) during wave-5 events (Fig.3)

Deleted: , the centered RMSE

Deleted: s

Deleted: , the correlations reduce greatly

Deleted: the

Deleted: , and mslp (0.69)

As for the wave-7 events, field correlations in Table A1 show that the upper-level circulation patterns compare well to ERA5 data: 0.84 (EC-Earth), 0.84 (CESM), and 0.82 (MIROC). Again, this confirms the previous statement of models' satisfactory performance in producing correct upper-level circulation patterns during high amplitude wave events. The t2m correlations in models are 0.70 (EC-Earth), 0.63 (CESM), and 0.53 (MIROC). The hot anomalies in t2m are quite pronounced in the regions of central North America, western Europe, northern Europe and central Eurasia. All models are able to reproduce the hot t2m anomalies in these regions but all have weaker positive and negative anomalies than ERA5 with n.s.d. values being 0.62 (EC-Earth), 0.61 (CESM), and 0.67 (MIROC) (Table A2). The large-scale precipitation anomaly patterns in EC-Earth relate better to WFDE5_CRU data for North America, whereas in both CESM and MIROC there is more noise. The correlation of precipitation anomalies 0.32 in the MMM. The large-scale patterns of mslp during wave-7 events match relatively well with ERA5 data with a MMM correlation value 0.63 In contrast to wave-5 having strong positive anomalies in mslp at the eastern side of the Atlantic Ocean (west coast of Europe), during wave-7 events strong negative anomalies are found at the same location. Also, positive mean sea level pressure anomalies are found during wave-7 events (Fig.4) at the east coast of North America, whereas the location shows negative anomalies during wave-5 events.

One common finding from both wave-5 and wave-7 events is that the models show relatively minor biases (n.s.d. ≥ 0.75) in upper-level circulation, but substantial biases in t2m, prcp, and mslp anomalies. All models substantially underestimate the magnitude of t2m, prcp, and mslp anomalies associated with wave-5 and wave-7 events, typically by a factor of 1.5 (Table A4).

3.3 Investigating sources of model biases

Next, we systematically assess the origin of the biases in the upper-level wind and surface fields as shown in Figs 3 and 4. As defined in the methods section, the bias maps were computed as the differences between the selected variables' anomalies in the models and in the reanalysis data ERA5 during high amplitude wave-5 and wave-7 events. It has to be noted that the biases that we refer to in surface variables are the biases of the anomalies instead of the absolute bias of the models. Here we present and describe the EC-Earth bias maps only. Equivalent plots for the other 2 models, with qualitatively similar outcomes, can be found in Fig. B6 to Fig. B9.

Here, we also employ the different nudged experiments: AISF (soil moisture prescription), AFSI (upper-level atmosphere nudging), and AFSF (nudging both) (see data section above for details). Overall, when nudging both atmosphere and soil moisture, the residual bias B_res is, as expected, negligible. This is true for both wave-5 and wave-7 events in all models and all analyzed variables (Fig.7 & Fig.8).

By nudging the atmosphere, the bias from the atmospheric part (B_atm) is (of course) almost completely removed for the v250 anomaly across all models (see Fig. 7(a), B_land). More interestingly, Fig. 7(b) shows that most of the EC-Earth t2m

310 anomaly bias is also removed when we nudge the upper-level atmosphere. Thus, the total bias (B_tot) in t2m is almost completely explained by the upper-level atmospheric bias (B_atm), and the land-atm bias (B_land) is negligible (Fig 7(b)).

Similarly, for wave-7 events, Fig. 8 (b) confirms our finding that nudging the upper-level atmosphere alone reduces the bias in surface temperature dramatically. Therefore, the total bias (B_tot) in t2m can be explained almost fully by B_atm, and
315 again the land contribution to the bias is minor (Fig.8(b)). Specifically, with the aid of a Taylor Diagram (Fig.6 (b) & Fig.6 (d)), there is a clear separation and improvement comparing the atmospherically nudged run AFSI to control run AISI and soil moisture nudged run AISF. The t2m bias still remains substantial when nudging the soil moisture. The actual n.s.d and RMSE values for v250 in AFSI are 1.0 and 0.10, compared to that of AISI run 0.8 (n.s.d.) & 0.55 (RMSE). Fig. 6 also exhibits that CESM and MIROC have similar characteristics, with substantial t2m and mslp being removed by upper-level
320 atmospheric nudging. Still in For t2m in EC-Earth , n.s.d. improves from 0.62 to 1.1 and mslp from 0.74 to 1.0. Fig. 6(c) also shows improvement in the spatial pattern correlation for prcp with 0.39 for the free running AISI run compared to 0.80 in the AFSI run. Another interesting observation obtained from comparing wave-5 and wave-7 Taylor Diagrams is that the models are more clustered for all variables for wave-7 compared to wave-5.

325 This shows, that free-running EC-Earth (AISI) has a relatively minor bias in v250 (blue square in Fig 7(a)) with a correlation of ~0.9, RMSE of ~0.5 and n.s.d. ~0.7. In other words, the pattern is very similar with a bit underestimated strength in terms of wind speed. Still the bias in t2m (B_tot, blue square in Fig 5 (b)) is substantially larger with correlation of ~0.6, RMSE ~0.9, and n.s.d ~0.6. Thus, the surface temperature imprint is underestimated with about a factor 1.7 (n.s.d ~0.6). This substantial bias in t2m is almost completely removed when nudging the upper-level wind field, i.e. removing the relatively
330 minor bias in v250. This is given by the blue triangle in Fig 5(b) (AFSI) showing a correlation of 0.94, RMSE of 0.36 and n.s.d. of 1. As can be seen in Fig. 5, the other models behave qualitatively in a similar way, with a substantial bias in near-surface temperature being almost completely removed when the relatively minor bias in upper level wind is removed.

Errors in precipitation anomalies are not fully removed when nudging upper-level circulation. Fig. 5 (c) shows some
335 reduction in the overall magnitude of errors in precipitation where the field correlation improves by almost a factor of 2 from 0.46 to 0.81(Fig.5(c)). However, for the bias in sea surface pressure anomalies, the errors are almost completely removed (Fig.5(b)).

In general, nudging the soil moisture doesn't affect the upper atmospheric flow. It can, however, in AISF runs reduce some
340 errors across models for the t2m and prcp variables. The same conclusion stands for the AISF and AFSI runs for prcp and mslp. The aforementioned observations are location specific as one component within a climate model might, erroneously, be tuned in such a way that it compensates for biases in other components of the climate model. If so, nudging only that

component would not reduce the overall bias. In this case, prescribing only soil moisture part does not guarantee the reduction of overall bias.

4 Discussion and outlook

4.1 Discussion

Large atmospheric circulation patterns, especially amplified wave-5 and wave-7 circumglobal Rossby waves, play an important role in climate variability and can trigger and maintain some extreme events such as heatwaves and heavy precipitation during the summer months. In this study, we demonstrate that amplified circumglobal Rossby waves, with a focus on the characteristics of boreal summer wave-5 (corr. 0.90) and wave-7 (corr. 0.83), wave-7 events are well captured in different climate models in terms of their climatology, variability. The phase-locking behaviour is captured. Both amplitude and week-to-week variability, in terms of standard deviations, are reasonably well reproduced in all models for all relevant wave numbers. The MMM n.s.d. for v250 are 0.99 and 0.91 during wave-5 and wave-7 events, showing that the wave-amplitudes are well-captured. Although the upper-level wind flows are satisfactorily reproduced across all models, their associated surface meteorological imprints (surface temperature and precipitation) during wave-5 and wave-7 high amplitude events is too weak. The MMM n.s.d. are 0.71 & 0.63 for t2m and 0.63 & 0.74 for prcp in wave-5 and wave-7 events respectively. These model biases can be largely corrected by nudging the upper-level atmosphere. For instance, the n.s.d. for the surface temperature field during high amplitude wave-5 (wave-7) events increase from 0.71 (0.63) to 0.99 (1.06), for precipitation from 0.63 (0.74) to 1.05 (1.04), and for sea level pressure from 0.69 (0.72) to 0.97 (1.0). The same scale reduction in errors are not observed when prescribing the soil moisture. This implies that a small bias in the upper atmospheric levels can result in big biases in surface weather conditions and related extreme events. A full analysis of the underlying reasons is outside the scope of this paper, but here we discuss some potential mechanisms. First, nudging zonal (u) and meridional (v) winds in the upper-atmosphere constrains the large-scale vertical wind component (ω) which is a key input for cloud parameterization schemes. In models, large-scale vertical wind is primarily defined by divergence in the horizontal wind fields, ensuring mass conservation, and thus nudging u and v will also effectively nudge ω. Likewise, biases in u and v will propagate in ω and can then have a strong (non-linear) impact on the amount of clouds in models (Satoh et al., 2019; Rio et al., 2019). Regions with anomalously high pressure due to the circumglobal wave, will have pronounced subsidence in ERA5 but this can be disturbed in the models. As a consequence, the models are likely to have more hazy cloud conditions as compared to clear-sky conditions in ERA5. This would impact the surface by reduced short wave radiation and hence less pronounced warm anomalies. Potential limitations in the cloud parametrization schemes could exacerbate this, with models having difficulties reproducing clear sky conditions (Lacagnina and Selten 2014). The resolution for GCMs often do not allow sub-grid scale convective systems and their associated clouds to be resolved. It is also highlighted that, particularly in mid-latitude continents, climatological biases in both clouds and precipitation persist in

major GCMs (Rio et al., 2019). One study specifically focused on clouds and radiative fluxes in EC-Earth and revealed that there are too many clouds in EC-Earth that are optically thick but too little clouds that are optically thin (Lacagnina and Selten 2014). Thus, a small bias in the upper atmospheric wind field propagates, via vertical wind (ω) and cloud scheme, into surface biases in climate models. While previous work has indicated that soil moisture can have pronounced effects on circumglobal waves (Koster et al., 2016; Teng et al., 2019), our analyses show that adjusting for soil moisture biases (by prescription) has little effect on the representation of circumglobal waves nor their surface imprint, or, at least not on the anomaly these events produce. These differences could arise from different time-scales and/or experiment set-ups. Earlier studies focused mainly on monthly to seasonal mean responses while ours analyzed weekly timescale. In addition, in Teng et al. (2019), they prescribed soil moisture as zero at specific locations, whereas in our prescribed soil moisture experiments, the soil moisture is not set to zeros. Instead, soil moisture is set to more realistic values coming from the model's land component froced by atmospheric fields from reanalysis. Our experiment thus represents much smaller forcings than prescribing the soil moisture to zero as done in Teng et al. (2019).

Recent work has shown that the individual extreme events can be examined in climate models, but anomalies or bias-corrections should be applied for during the analysis (Wehrli et al., 2019). Similarly, in our study, we defined the wave-5 and wave-7 events, then the surface variable anomalies against their climatological means and compared them between ERA5 and the models for different experiments.

4.2 Limitations and outlook

One caveat of the Fast Fourier Transform (FFT) method applied in this study is that the results return circumglobal Rossby waves patterns. Since its on the hemispheric scale, some local blocking events that are not extensive in spatial scale, might not be accounted for. Depending on the stationarity of the dataset, the results can also differ substantially with FFT method. The local signal is ignored as the input data for FFT method is the meridional average of wind v250 field. Some small inconsistencies between models and ERA5 reanalysis data can arise from the under-sampling issue in ERA5. Since there are five ensemble members for the AISI experiment, the data used in the analysis are five times bigger for the models than for ERA5. As for the atmospheric nudging experiments, from the nudging vertical profile, it can be observed that in CESM and MIROC, the nudging intensity is identical, whereas in EC-Earth the nudging strength is weaker between 700hPa to 400hPa. This can result some differences in the v250 field.

Our findings have implications for climate model projections of persistent summer weather extremes in the key affected regions. Kornhuber et al. (2020) identified hotspots that are affected by summertime amplified wave-5 patterns (Central North America, Eastern Europe, and Eastern Asia) and wave-7 (Western Central North America, Western Europe and Western Asia). These regions are sensitive to simultaneous heat extremes, and to exacerbate the situation, some identified regions are also considered as global breadbasket regions. In the summer, when wave-5 or wave-7 events persist more than

two weeks, the average reduction in crop production is 4% and even up to 11% on regional level (Kornhuber et al., 2020).

445 In our study, for wave-5 and wave-7 events, all the aforementioned key regions are identified in our three models for near-surface temperature positive anomalies. This allows us to have more confidence in the state-of-the-art climate models, not only for the accurate representation of upper-level atmospheric circulation, but also for the accuracy of the associated hotspots at the surface. Since the strength of the near-surface temperature anomalies is underestimated, the climate models are likely to underestimate heatwaves as well. A potential way to adjust for this is to establish statistical links between upper-

450 level atmosphere slow to near-surface temperature based on observational data, i.e. an emergent constraint. Then use this statistical link to adjust the effect of upper-level atmospheric circulation changes in climate models under future scenarios for heatwave risks. In addition, it will be important to assess how the large-scale circulation pattern and strength change under future greenhouse gas forcing that can lead to changes in the extreme events.

5 Summary and Conclusions

455 Our validation study shows that upper-level wave characteristics are reasonably well reproduced in three GCMs in historic AMIP runs with MMM of n.s.d. for wave-5 and wave-7 events being 0.99 and 0.91. Both the climatology and phase-locking behaviors are captured in models for wave numbers 5 and 7, as the MMM correlation values are 0.90 and 0.83. Surface temperature anomalies are associated with the amplified wave-5 and wave-7 patterns but have weaker anomalies as compared to the ERA5 reanalysis data. The MMM n.s.d. for the surface temperature field during high amplitude wave-5 and

460 wave-7 events increase from 0.71 and 0.63 to 0.99 and 1.06 just by nudging the mid-to-upper level atmosphere.

In summary, for the surface meteorological variables, we find that:

- Overall, v250 is the most accurate and precipitation is poorly estimated among all variables for both wave-5 and wave-7 events.
465 - Prescribing soil moisture does not add much improvement for the anomalies in t2m and prcp. In the case of v250 and mslp, it even made the representation of wave-5 and wave-7 events worse. This is likely due to model tuning issues.
- Nudging the upper-level atmosphere indicates that this is the prime origin of surface anomaly biases. We observe significant improvements from AISI in AISF runs and AFSI runs across all models and all variables.
470

The bias in model surface imprints for amplified wave events mainly originates from smaller biases in the atmospheric circulation component of the model for the mid latitudes. The soil moisture feedback also explains some bias but this is minor compared to the atmospheric part. Our study suggests that climate models can be used to study present and future wave characteristics, but that care should be taken when analyzing the associated surface extremes.

Deleted: 2
Deleted: in summer events and on a regional level
Deleted: correctness
Deleted: 99
Deleted: 63
Deleted: -
Deleted: s
Deleted: s

485

Code and data availability. The code and data can be made available by the authors upon request.

Author contributions. FL, DC and FS designed the analysis with input from KK. FL ran the EC-Earth3 simulations with technical help from FS, WM, PLS, and TR. HS ran the interactive and atmosphere nudged simulations with MIROC5. DT

490   and HK ran the soil 460 moisture nudged simulations with MIROC5. KW ran the CESM1.2 model simulations with technical support by Mathias Hauser. FL analyzed the results from all models. All authors contributed to the discussion of results. FL prepared the manuscript with contributions from all co-authors.

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

495

Acknowledgements. EC-Earth3 simulations were contributed by VU Amsterdam and KNMI. The MIROC5 simulations were contributed by NIES Japan and University of Tokyo. CESM1.2 simulations were contributed by ETH Zurich. FL, DC and FS acknowledge VIDI-award from Netherlands Organization for Scientific Research (NWO) (Persistent Summer Extremes "PERSIST" project: 016.Vidi.171.011). KK was partially supported by the NSF project NSF AGS-1934358.KW and SIS

500   acknowledge funding from the European Research Council (ERC) ("DROUGHT-HEAT" project, Grant 617518). HS was supported by the Integrated Research Program for Advancing Climate Models (JPMXD0717935457). The MIROC5 simulations were performed by using Earth Simulator in JAMSTEC and the NEC SX in NIES. WM is supported through the Swedish strategic research area ModElling the Regional and Global Earth system (MERGE). H.K acknowledges the National Research Foundation of Korea (NRF) grant Funded by the Korea Government (MSIT)(2021H1D3A2A03097768).

505   The authors would like to thank Mathias Hauser and Emanuel Dutra for their help in the discussions for the preparation of soil moisture prescription data and the applications in the models.

510

References

Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., and Garcia-Herrera, R.: "The Hot Summer of 2010 : Map of Europe." *Science* 332 (April): 220–24, doi: 10.1126/science.1201224, 2011.

Benjamini, Y., and Hochberg, Y.:" Controlling the false discovery rate: a practical and powerful approach to multiple testing". *Journal of the Royal Statistical Society Series B*, **57**, 289-300, doi: http://www.jstor .org /stable /23461 01, 1995.

Branstator, G.: "Circumglobal Teleconnections, the Jet Stream Waveguide, and the North Atlantic Oscillation." *Journal of Climate* 15 (14): 1893–1910, doi: 10.1175/1520-0442 (2002) 015<1893:CTTJSW>2.0.CO;2 2002b. 2002a, 2002.

Branstator, G., and Teng, H.: "Tropospheric Waveguide Teleconnections and Their Seasonality." *Journal of the Atmospheric Sciences* 74 (5): 1513–32, doi: 10.1175/JAS-D-16-0305.1, 2017.

Di Capua, G., Kretschmer, M., Donner, R.V., van den Hurk, B., Vellore, R., Krishnan, R., and Coumou, D.: "Tropical and Mid-Latitude Teleconnections Interacting with the Indian Summer Monsoon Rainfall: A Theory-Guided Causal Effect Network Approach." *Earth System Dynamics* 11 (1): 17–34, doi: 10.5194/esd-11-17-2020, 2020.

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Schmied, H. M., Hersbach, H., and Buontempo.C.,: "WFDE5: Bias-Adjusted ERA5 Reanalysis Data for Impact Studies." *Earth System Science Data* 12 (3): 2097–2120, doi: 10.5194/essd-12-2097-2020, 2020.

Davini, P. , and D'Andrea, F.: "From CMIP3 to CMIP6: Northern Hemisphere Atmospheric Blocking Simulation in Present and Future Climate." *Journal of Climate* 33 (23): 10021–38, doi: 10.1175/jcli-d-19-0862.1, 2020.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haim- berger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Ding, Q., and Wang, B.: "Circumglobal Teleconnection in the Northern Hemisphere Summer." *Journal of Climate* 18 (17): 3483–3505, doi: 10.1175/JCLI3473.1, 2005.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., Bergmann, T., et al.: "The EC-Earth3 Earth System Model for the Climate Model Intercomparison Project 6." *Geoscientific Model Development Discussions*, no. February: 1–90, doi: 10.5194/gmd-2020-446, 2021.

Eyring, V., Bony, S., Meehl, G. A., A. Senior, C., Stevens, B., J. Stouffer, R., and Taylor. K. E.: "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization." *Geoscientific Model Development* 9 (5): 1937–58, doi: 10.5194/gmd-9-1937-2016, 2016.

Garfinkel, C. I., White, I., Gerber, E. P., Jucker, M. and Erez, M.: "The Building Blocks of Northern Hemisphere Wintertime Stationary Waves." *Journal of Climate* 33 (13): 5611–33, doi: 10.1175/jcli-d-19-0181.1, 2020.

Gates, W. L., Boyle, J. S., Covey, C., G. Dease, C., Doutriaux, C. M., Drach, R. S., Fiorino, M, et al. : "An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP I)." *Bulletin of the American Meteorological Society*, doi: 10.1175/1520-0477 (1999)080<0029:AOOTRO> 2.0. CO;2, 1999.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., et al.: "The ERA5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146 (730): 1999–2049, doi:10.1002/qj.3803, 2020.

Hoskins, B. J., and Ambrizzi, T.: "Rossby Wave Propagation on a Realistic Longitudinally Varying Flow." *Journal of the Atmospheric Sciences* 50 (12), doi: 10.1175/1520-0469(1993)050<1661:RWPOAR>2.0.CO;2, 1993.

Huntingford, C., Mitchell, D., Kornhuber, k., Coumou, D., Osprey, S., and Allen, M.: "Assessing Changes in Risk of Amplified Planetary Waves in a Warming World." *Atmospheric Science Letters*, no. May 2018: 1–11, doi: 10.1002/asl.929, 2019.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Jennifer E. K., Kushner, P. J., Lamarque, J. F., et al.: "The Community Earth System Model: A Framework for Collaborative Research." *Bulletin of the American Meteorological Society* 94 (9): 1339–60, doi: 10.1175/BAMS-D-12-00121.1, 2013.

Jeuken, A. B. M., Siegmund, P. C., Heijboer, L. C., Feichter, J., and Bengtsson, L.: "On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation", *Journal of Geophysical Research: Atmospheres,* 101, 16 939–16 950, https://doi.org/10.1029/96JD01218, 1996.

Kooperman, G. J., Pritchard, M. S., Ghan, S. J., Wang, M., Somerville, R. C. J., and Russell, L. M.: "Constraining the influence of natural variability to improve estimates of global aerosol indirect effects in a nudged version of the Community Atmosphere Model 5", *Journal of Geophysical Research: Atmospheres,* 117, https://doi.org/10.1029/2012JD018588, d23204, 2012.

Kornhuber, K., Petoukhov. V., Karoly. D., Petri. S., Rahmstorf, S., and D Coumou: "Summertime planetary wave resonance in the Northern and Southern Hemispheres" *Journal of Climate 30* (16), 6133-6150, doi: https://doi.org/10.1175/JCLI-D-16-0703.1, 2017.

Kornhuber, K., Petoukhov, V., Petri, S., Rahmstorf, S., and Coumou, D.: "Evidence for Wave Resonance as a Key Mechanism for Generating High-Amplitude Quasi-Stationary Waves in Boreal Summer." *Climate Dynamics* 49 (5–6): 1961–79, doi: 10.1007/s00382-016-3399-6, 2017.

Kornhuber, K., Coumou, D., Vogel, E., Lesk, C., Donges, J. F., Lehmann, J. and M. Horton, R.: "Amplified Rossby Waves Enhance Risk of Concurrent Heatwaves in Major Breadbasket Regions." *Nature Climate Change* 10 (1): 48–53, doi: 10.1038/s41558-019-0637-z, 2020.

Kornhuber, K., Osprey, S., Coumou, D., Petri, S., Petoukhov, V., Rahmstorf, S., and Gray, L.: "Extreme Weather Events in Early Summer 2018 Connected by a Recurrent Hemispheric Wave-7 Pattern." *Environmental Research Letters* 14 (5): 054002, doi: 10.1088/1748-9326/ab13bf, 2019.

580   Koster, R. D., Chang, Y., Wang, H., and Schubert, S. D.: "Impacts of Local Soil Moisture Anomalies on the Atmospheric Circulation and on Remote Surface Meteorological Fields during Boreal Summer: A Comprehensive Analysis over North America." *Journal of Climate* 29 (20), doi: 10.1175/JCLI-D-16-0192.1, 2016.

Krzyżewska, A., and Dyer. J.: "The August 2015 Mega-Heatwave in Poland in the Context of Past Events." *Weather* 73 (7): 207–14, doi: 10.1002/wea.3244, 2018.

585   Lacagnina, C., and Frank S.: "Evaluation of Clouds and Radiative Fluxes in the EC-Earth General Circulation Model." *Climate Dynamics* 43 (9–10), doi: 10.1007/s00382-014-2093-9, 2014.

Lau, W. K. M., and Kim, K.: "The 2010 Pakistan Flood and Russian Heat Wave: Teleconnection of Hydrometeorological Extremes." *Journal of Hydrometeorology* 13 (1): 392–403, doi: 10.1175/JHM-D-11-016.1, 2012.

Martius, O., Sodemann, H., Joos, H., Pfahl, S., Winschall, A., Croci-Maspoli, M., Graf, M., et al.: "The Role of Upper-Level
590   Dynamics and Surface Processes for the Pakistan Flood of July 2010." *Quarterly Journal of the Royal Meteorological Society* 139 (676): 1780–97, doi: 10.1002/qj.2082, 2013.

McKinnon, K. A., Rhines, A., Tingley, M. P., and Huybers, P.: "Long-Lead Predictions of Eastern United States Hot Days from Pacific Sea Surface Temperatures." *Nature Geoscience* 9 (5): 389–94, doi: 10.1038/ngeo2687, 2016.

Petoukhov, V., Petri, S., Rahmstorf, S., Coumou, D., Kornhuber, K., and Schellnhuber, H. J.: "Role of Quasiresonant
595   Planetary Wave Dynamics in Recent Boreal Spring-to-Autumn Extreme Events." *Proceedings of the National Academy of Sciences* 113 (25): 6862–67, doi: 10.1073/pnas.1606300113, 2016.

Petoukhov, V., Rahmstorf, S., Petri, S., and Schellnhuber, H. J.: "Quasiresonant Amplification of Planetary Waves and Recent Northern Hemisphere Weather Extremes." *Proceedings of the National Academy of Sciences of the United States of America* 110 (14): 5336–41, doi: 10.1073/pnas.1222000110, 2013.

600   Rio, C., Del Genio, A. D., and Hourdin, F.: "Ongoing Breakthroughs in Convective Parameterization." *Current Climate Change Reports*, doi: 10.1007/s40641-019-00127-w, 2019.

Röthlisbergera, M., Frossard, L., Bosart, L. F., Keyser, D., and Martius, O.: "Recurrent Synoptic-Scale Rossby Wave Patterns and Their Effect on the Persistence of Cold and Hot Spells." *Journal of Climate* 32 (11): 3207–26, doi: 10.1175/JCLI-D-18-0664.1, 2019.

605   Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S., Putman, W. M., and Düben, P.: "Global Cloud-Resolving Models," *Current Climate Change Reports*, 172–84, doi:10.1007/s40641-019-00131-0, 2019.

Scaife, A. A., Woollings, T., Knight, J., Martin, G., and Hinton, T.: "Atmospheric Blocking and Mean Biases in Climate Models." *Journal of Climate*, doi: 10.1175/2010JCLI3728.1, 2010.

Screen, J. A., and Simmonds, I.: "Amplified Mid-Latitude Planetary Waves Favour Particular Regional Weather Extremes."
610   *Nature Climate Change* 4 (8): 704–9, doi: 10.1038/nclimate2271, 2014.

Shepherd, T. G. "Atmospheric Circulation as a Source of Uncertainty in Climate Change Projections." *Nature Geoscience* 7 (10): 703–8, doi: 10.1038/NGEO2253, 2014.

Taylor, K. E.: "Summarizing Multiple Aspects of Model Performance in a Single Diagram." *Journal of Geophysical*

Research Atmospheres 106 (D7), doi: 10.1029/2000JD900719, 2001.

615 Taylor, K. E., Stouffer, R. J. and Meehl, G. A.: "An Overview of CMIP5 and the Experiment Design." *Bulletin of the American Meteorological Society*, doi: 10.1175/BAMS-D-11-00094.1, 2012.

Teng, H., and Branstator, G.: "Amplification of Waveguide Teleconnections in the Boreal Summer." *Current Climate Change Reports* 5 (4): 421–32, doi: 10.1007/s40641-019-00150-x, 2019.

Teng, H., Branstator, G., Tawfik, A. B., and Callaghan, P.: "Circumglobal Response to Prescribed Soil Moisture over North
620 America." *Journal of Climate* 32 (14): 4525–45, doi: 10.1175/JCLI-D-18-0823.1, 2019.

Thomson, S. I., and Vallis, G. K.: "Atmospheric Response to SST Anomalies. Part II: Background-State Dependence, Teleconnections, and Local Effects in Summer." *Journal of the Atmospheric Sciences* 75 (12): 4125–38, doi: 10.1175/JAS-D-17-0298.1, 2018.

Vijverberg, S., Schmeits, M., van der Wiel, K., and Coumou. D.: "Subseasonal Statistical Forecasts of Eastern U.S. Hot
625 Temperature Events." *Monthly Weather Review* 148 (12), doi: 10.1175/mwr-d-19-0409.1, 2020.

Wang, H., Schubert, S., Koster, R., Ham, Y. G., and Suarez, M.: "On the Role of SST Forcing in the 2011 and 2012 Extreme U.S. Heat and Drought: A Study in Contrasts." *Journal of Hydrometeorology* 15 (3): 1255–73, doi: 10.1175/JHM-D-13-069.1, 2014.

Watanabe, M., Suzuki, T., O'Ishi, R., Komuro, Y., Watanabe, S., Emori, S., Takemura, T., et al.: "Improved Climate
630 Simulation by MIROC5: Mean States, Variability, and Climate Sensitivity." *Journal of Climate* 23 (23): 6312–35, doi: 10.1175/2010JCLI3679.1, 2010.

Wehrli, K., Guillod, B. P., Hauser, M., Leclair, M., and Seneviratne, S. I.: "Identifying Key Driving Processes of Major Recent Heat Waves." *Journal of Geophysical Research: Atmospheres* 124 (22): 11746–65, doi: 10.1029/2019JD030635, 2019.

635 Wehrli, K., Luo, F., Hauser, M., Shiogama, H., Tokuda, D., Kim, H., Coumou, D., et al.: "The ExtremeX Global Climate Model Experiment : Investigating Thermodynamic and Dynamic Processes Contributing to Weather and Climate Extremes," no. July: 1–31, doi:10.5194/esd-2021-58, 2021.

Wills, R. C. J., White, R. H., and Levine, X. J.: "Northern Hemisphere Stationary Waves in a Changing Climate." *Current Climate Change Reports*. Springer, doi: 10.1007/s40641-019-00147-6, 2019.

640 Xu, P., Wang, L., Huang, P., and Chen, W.: "Disentangling Dynamical and Thermodynamical Contributions to the Record-Breaking Heatwave over Central Europe in June 2019." *Atmospheric Research* 252 (October 2020): 105446, doi: 10.1016/j.atmosres.2020.105446, 2021.

Zscheischler, J., Westra, S., Van Den Hurk, B. J. J. M., Seneviratne, S. I., Ward, P. J., Pitman, A., Aghakouchak, A., et al.: "Future Climate Risk from Compound Events." *Nature Climate Change* 8 (6): 469–77, doi: 10.1038/s41558-018-
645 0156-3, 2018.

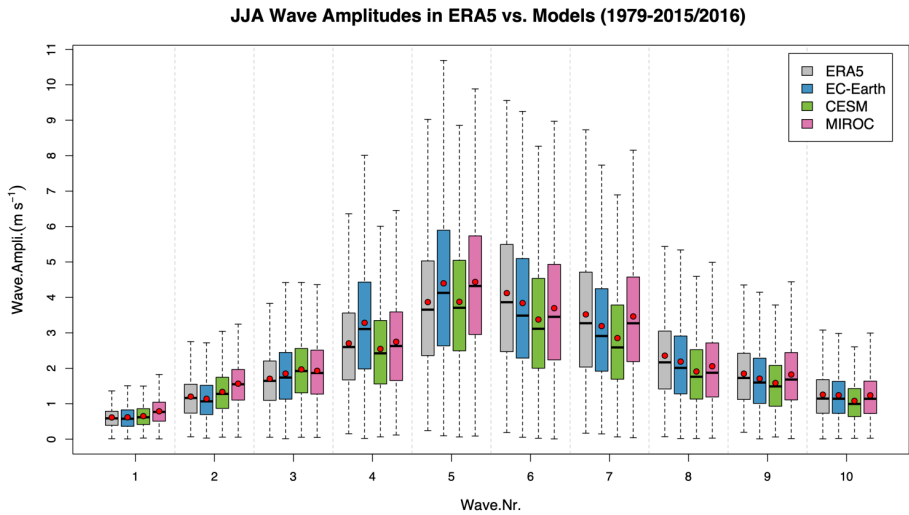
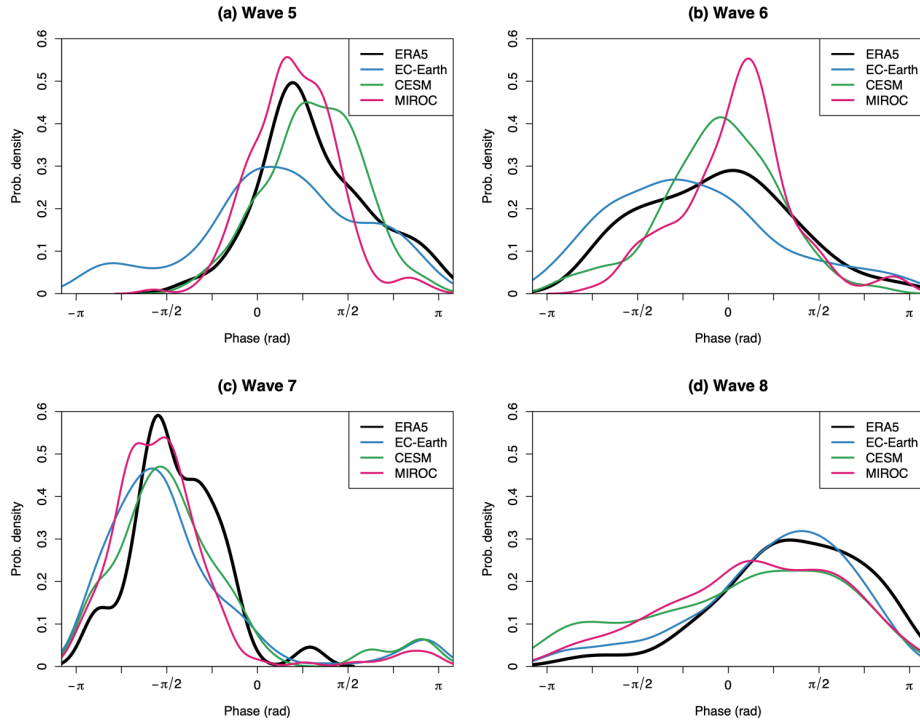



Figure 2: Phase-locking of Rossby waves for JJA ERA5 and model waves 5-8 in control run AISI for high wave amplitude events (> 1.5 s.d.): (a)-(d), Probability density functions of the phase positions of waves 5-8 in ERA5, EC-Earth, CESM, and MIROC during JJA for the period of 1979-2015/2016 (wave 5 (a), wave 6 (b), wave 7 (c), wave 8 (d)). The bandwidth for ERA5 and models are as follows: (a) Wave 5: 0.35(ERA5), 0.40(EC-Earth), 0.25(CESM), 0.22(MIROC), (b) Wave 6: 0.53(ERA5), 0.45(EC-Earth), 0.30(CESM), 0.25(MIROC), (c) Wave 7: 0.25(ERA5), 0.29(EC-Earth), 0.27(CESM), 0.22(MIROC), (d) Wave 8: 0.49(ERA5), 0.39(EC-Earth), 0.52(CESM), 0.46(MIROC).

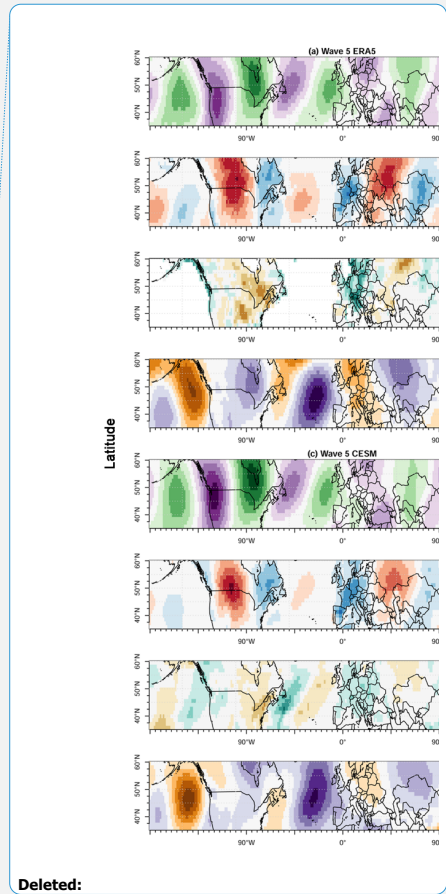
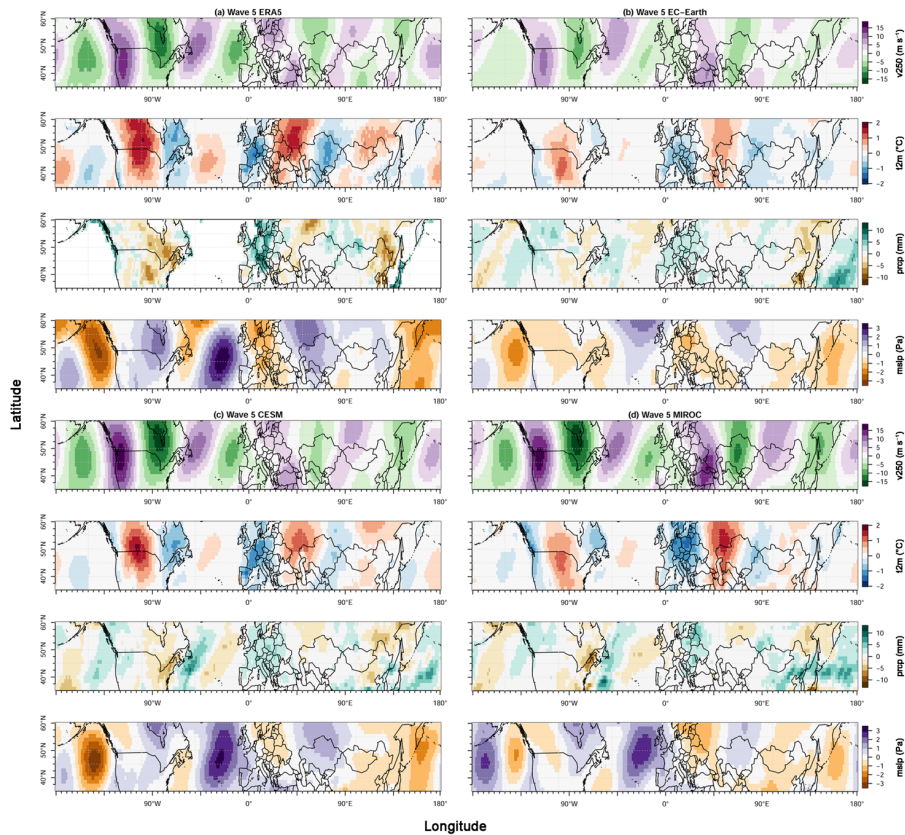



Figure 3: Composite anomaly plots of weeks with high-amplitude waves-5 events for meridional wind velocity at 250hPa (v250, absolute field), near-surface temperature (t2m, anomaly), precipitation (prcp, anomaly), and seal level pressure (mslp, anomaly) in ERA5 (a), EC-Earth (b), CESM (c) and MIROC (d) based on control runs AISI.

675

680

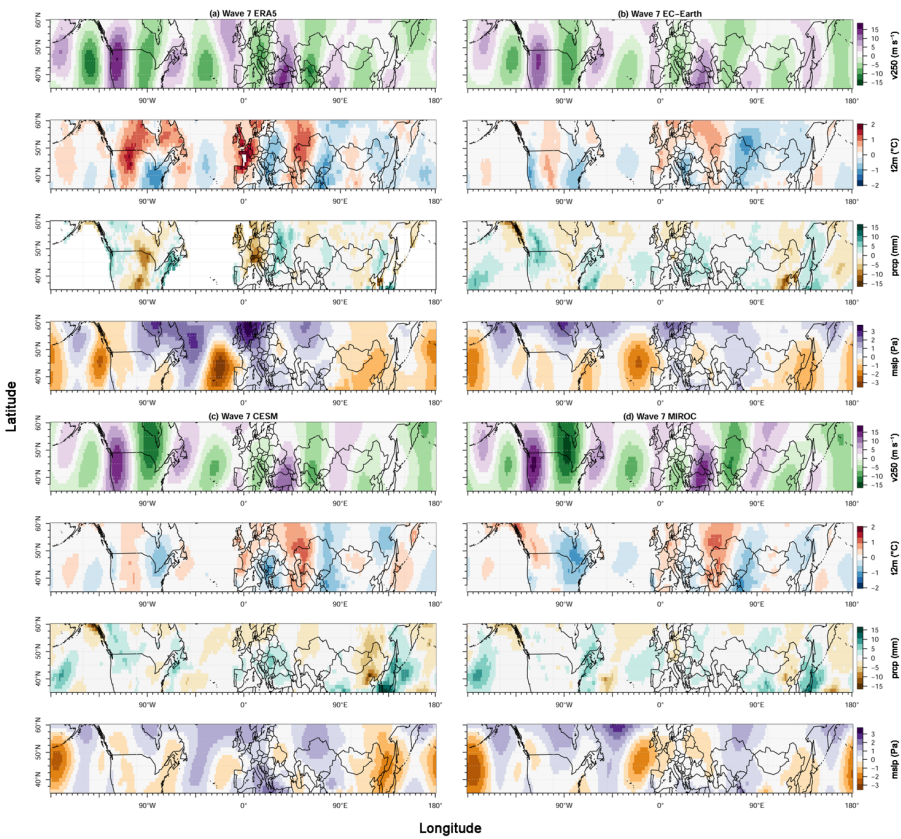



Figure 4: Composite anomaly plots of weeks with high-amplitude waves-7 events for meridional wind velocity at 250hPa (v250, absolute field), near-surface temperature (t2m, anomaly), precipitation (prcp, anomaly), and seal level pressure (mslp, anomaly) in ERA5 (a), EC-Earth (b), CESM (c) and MIROC (d) based on control runs AISI.

685

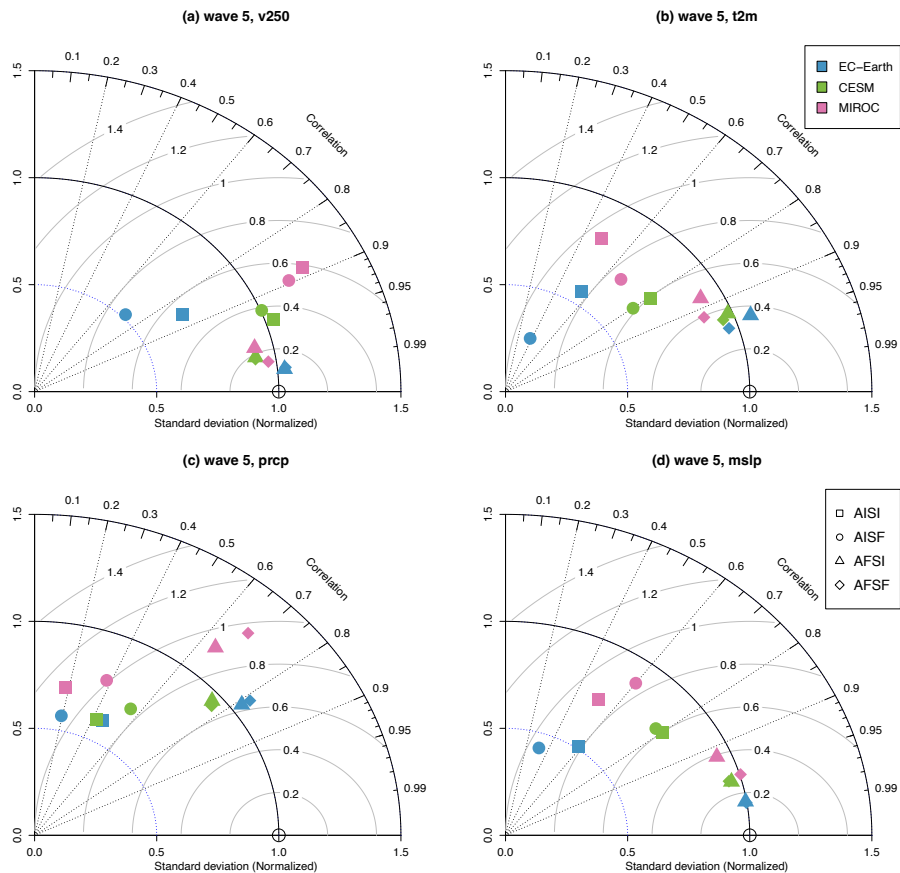



Figure 5: Taylor Diagram for all experiments in models compared to ERA5 for wave-5 events. For a) V250, b) t2m, c) prcp, and d) mslp, the Taylor diagram presents for each model and each experiment, three statistics: the Pearson correlation (dashed lines); the RMS error (grey contours); and the normalized spatial standard deviation (solid black contours).

690

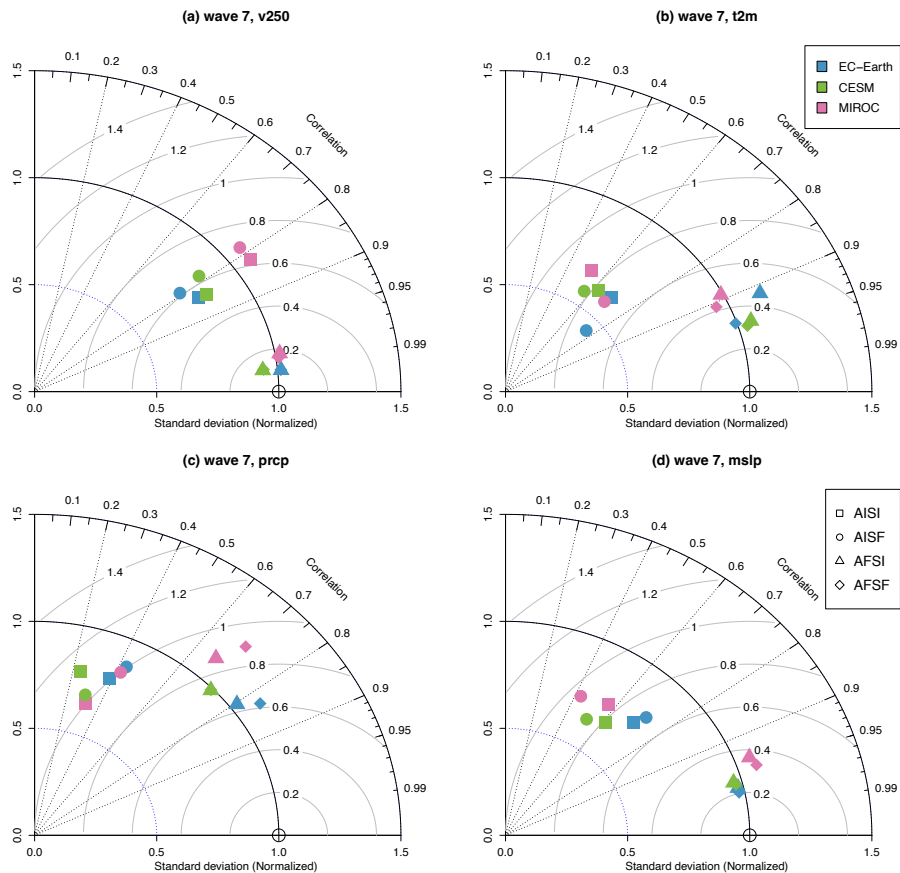



Figure 6: Taylor Diagram for all experiments in models compared to ERA7 for wave-5 event. For a) v250, b) t2m, c) prcp, and d) mslp, the Taylor diagram presents for each model and each experiment, three statistics: the Pearson correlation (dashed lines); the RMS error (grey contours); and the normalized spatial standard deviation (solid black contours).

695

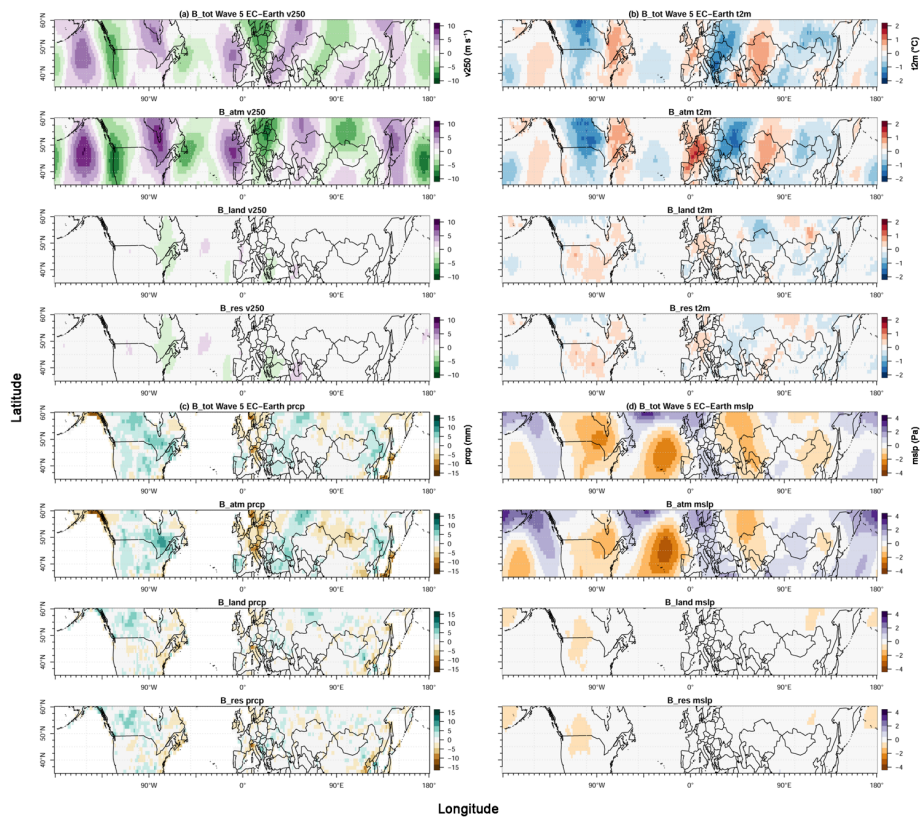



Figure 7: Bias plots for high-amplitude wave-5 events in different experiments for EC-Earth. Total bias(B_tot), atmospheric bias(B_atm), land-atmosphere interaction bias (B_land) and residual bias(B_res) for meridional wind velocity at 250hPa (a), surface temperature (b), precipitation (c), and seal level pressure (d).

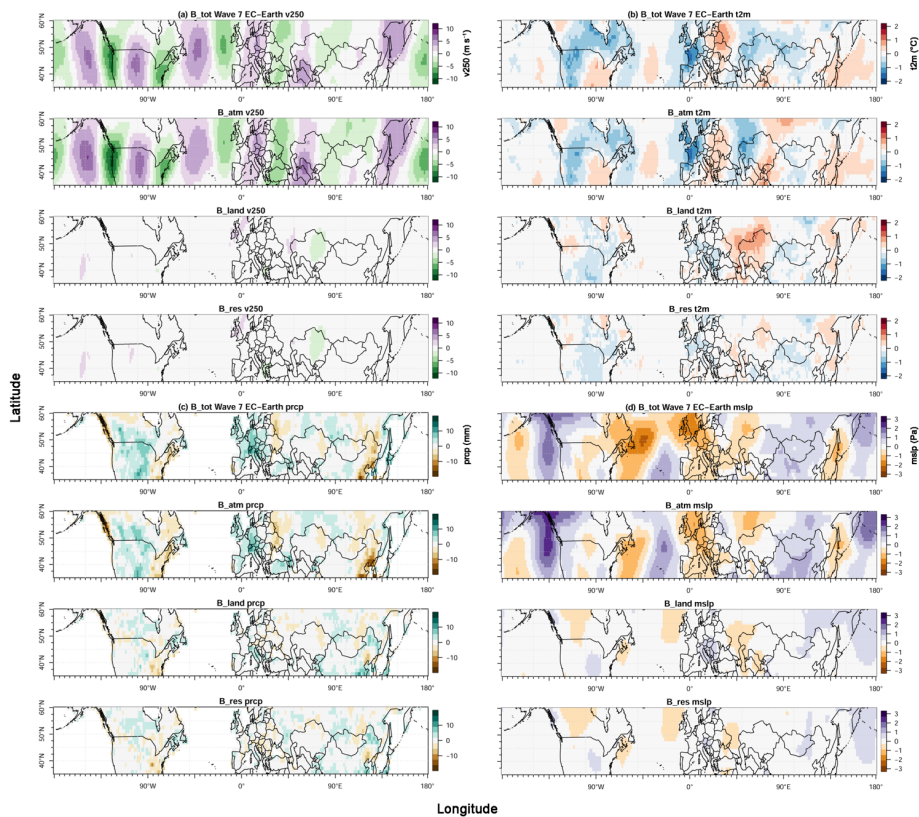



Figure 8: Bias plots for high-amplitude wave-7 events in different experiments for EC-Earth. Total bias(B_tot), Atmospheric bias(B_atm), Land-Atm interaction bias(B_land) and residual bias(B_res) for meridional wind velocity at 250hPa (a), surface temperature (b), precipitation (c), and seal level pressure (d).

Appendix

Wave5	v250				tas				prcp				mslp			
Model	EC-Earth	CESM	MIROC	Model Mean	EC-Earth	CESM	MIROC	Model Mean	EC-Earth	CESM	MIROC	Model Mean	EC-Earth	CESM	MIROC	Model Mean
AISI	0,858	0,945	0,884	0,896	0,552	0,806	0,481	0,613	0,460	0,427	0,179	0,355	0,584	0,801	0,515	0,633
AISF	0,719	0,926	0,895	0,847	0,375	0,801	0,669	0,615	0,192	0,554	0,377	0,374	0,317	0,777	0,600	0,565
AFSI	0,995	0,985	0,975	0,985	0,942	0,928	0,878	0,916	0,812	0,756	0,645	0,738	0,987	0,965	0,921	0,958
AFSF	0,994	0,986	0,989	0,990	0,951	0,936	0,920	0,936	0,814	0,768	0,679	0,754	0,988	0,964	0,959	0,970
Wave7	**v250**				**tas**				**prcp**				**mslp**			
Model	EC-Earth	CESM	MIROC	Model Mean	EC-Earth	CESM	MIROC	Model Mean	EC-Earth	CESM	MIROC	Model Mean	EC-Earth	CESM	MIROC	Model Mean
AISI	0,837	0,839	0,819	0,832	0,703	0,629	0,528	0,620	0,388	0,236	0,323	0,316	0,704	0,613	0,566	0,628
AISF	0,791	0,780	0,780	0,784	0,757	0,565	0,693	0,672	0,431	0,301	0,420	0,384	0,723	0,522	0,429	0,558
AFSI	0,995	0,994	0,984	0,991	0,915	0,950	0,890	0,918	0,804	0,729	0,668	0,734	0,974	0,967	0,939	0,960
AFSF	0,995	0,994	0,988	0,992	0,947	0,955	0,909	0,937	0,832	0,728	0,700	0,753	0,979	0,970	0,952	0,967

Table A1: Summary of Model Taylor Diagram correlation values.

Wave5	v250					tas					pr					psl				
Model	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM
AISI	3,510	5,168	6,203	4,960	0,994	0,292	0,382	0,424	0,366	0,705	1,746	1,729	2,027	1,834	0,634	0,586	0,916	0,847	0,783	0,685
AISF	2,584	5,012	5,805	4,467	0,895	0,139	0,338	0,367	0,281	0,542	1,646	2,054	2,262	1,987	0,687	0,492	0,906	1,016	0,805	0,704
AFSI	5,133	4,584	4,608	4,775	0,957	0,553	0,509	0,472	0,511	0,985	3,024	2,779	3,322	3,042	1,052	1,138	1,096	1,075	1,103	0,965
AFSF	5,158	4,576	4,826	4,853	0,973	0,499	0,494	0,459	0,484	0,933	3,135	2,731	3,723	3,196	1,105	1,141	1,085	1,148	1,125	0,983
Wave7	**v250**					**tas**					**pr**					**psl**				
Model	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM
AISI	4,162	4,344	5,585	4,697	0,907	0,310	0,304	0,334	0,316	0,631	2,526	2,510	2,064	2,366	0,744	0,818	0,734	0,818	0,790	0,717
AISF	3,894	4,467	5,574	4,645	0,897	0,219	0,285	0,292	0,265	0,530	2,775	2,190	2,669	2,545	0,800	0,877	0,700	0,792	0,790	0,717
AFSI	5,250	4,864	5,276	5,130	0,991	0,570	0,530	0,496	0,532	1,063	3,278	3,150	3,537	3,322	1,044	1,076	1,063	1,170	1,103	1,002
AFSF	5,245	4,883	5,240	5,123	0,990	0,498	0,520	0,476	0,498	0,994	3,532	3,154	3,928	3,538	1,112	1,076	1,072	1,188	1,112	1,010

Table A2: Summary of model standard deviation values.

Wave5	v250					tas					prcp					psl				
Model	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM
AISI	2,675	1,695	2,942	2,437	0,488	0,446	0,319	0,498	0,421	0,811	2,302	2,425	2,924	2,550	0,882	0,956	0,687	1,025	0,889	0,778
AISF	3,610	1,927	2,600	2,712	0,544	0,497	0,343	0,397	0,412	0,794	2,671	2,251	2,530	2,484	0,859	1,093	0,724	0,995	0,937	0,820
AFSI	0,542	0,944	1,140	0,875	0,175	0,187	0,209	0,257	0,218	0,420	1,771	1,822	2,490	2,027	0,701	0,189	0,303	0,450	0,314	0,274
AFSF	0,586	0,908	0,740	0,745	0,149	0,160	0,197	0,212	0,190	0,366	1,793	1,767	2,468	2,009	0,695	0,187	0,305	0,330	0,274	0,240
Wave7	**v250**					**tas**					**prcp**					**psl**				
Model	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM	EC-Earth	CESM	MIROC	Model Mean	NORM
AISI	2,838	2,813	3,263	2,971	0,574	0,380	0,397	0,442	0,407	0,812	3,076	3,395	2,988	3,153	0,991	0,785	0,872	0,928	0,862	0,783
AISF	3,174	3,267	3,582	3,341	0,645	0,376	0,420	0,372	0,389	0,777	3,062	3,152	2,878	3,031	0,953	0,785	0,949	1,04764	0,927	0,842
AFSI	0,527	0,638	0,935	0,700	0,135	0,231	0,167	0,237	0,212	0,423	1,988	2,262	2,558	2,269	0,713	0,250	0,306	0,510	0,355	0,323
AFSF	0,536	0,645	0,829	0,670	0,129	0,166	0,155	0,215	0,179	0,356	1,908	2,206	2,560	2,224	0,699	0,226	0,304	0,483	0,338	0,307

Table A3: Summary of model RMSE values.

Wave5	v250			tas			prcp			psl		
Model	corr	std	rmse	corr	std	rmse	corr	std	rmse	corr	std	rmse
AISI	0,896	0,994	0,488	0,613	0,705	0,811	0,355	0,634	0,882	0,633	0,685	0,778
AISF	0,847	0,895	0,544	0,615	0,542	0,794	0,374	0,687	0,859	0,565	0,704	0,820
AFSI	0,985	0,957	0,175	0,916	0,985	0,420	0,738	1,052	0,701	0,958	0,965	0,274
AFSF	0,990	0,973	0,149	0,936	0,933	0,366	0,754	1,105	0,695	0,970	0,983	0,240
Wave7	v250			tas			prcp			psl		
Model	corr	std	rmse	corr	std	rmse	corr	std	rmse	corr	std	rmse
AISI	0,832	0,907	0,574	0,620	0,631	0,812	0,316	0,744	0,991	0,628	0,717	0,783
AISF	0,784	0,897	0,645	0,672	0,530	0,777	0,384	0,800	0,953	0,558	0,717	0,842
AFSI	0,991	0,991	0,135	0,918	1,063	0,423	0,734	1,044	0,713	0,960	1,002	0,323
AFSF	0,992	0,990	0,129	0,937	0,994	0,356	0,753	1,112	0,699	0,967	1,010	0,307

Table A4: Summary of Multi-Model Mean Taylor Diagram values.

ERA5 s.d.	v250	t2m	prcp	mslp
wave 5	4,990	0,519	2,893	1,144
wave 7	5,177	0,501	3,181	1,101

Table A5: Summary of ERA5 standard deviation values.

Deleted:

ERA5 s.d.	v250	t2m	prcp	mslp
wave 5	4,990	0,519	2,893	1,144
wave 7	5,177	0,501	3,181	1,101

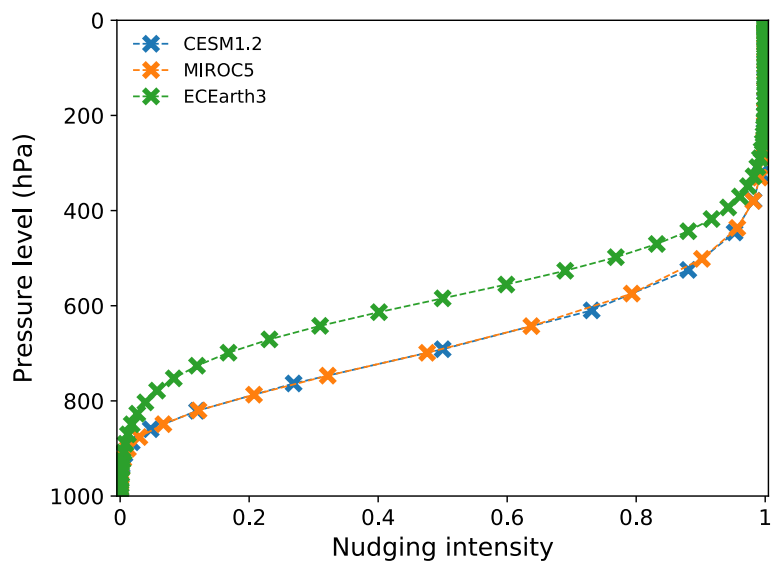



Figure B1. Nudging profile for the three ExtremeX ESMs. The actual pressure levels are marked with an x and joined with lines.
The nudging intensity is given from zero (no nudging) to one (fully nudged) (Wehrli et al., 2021, in review). Taken with permission.

750

755

760

Figure B2. Same with Fig.2 but with histogram added.

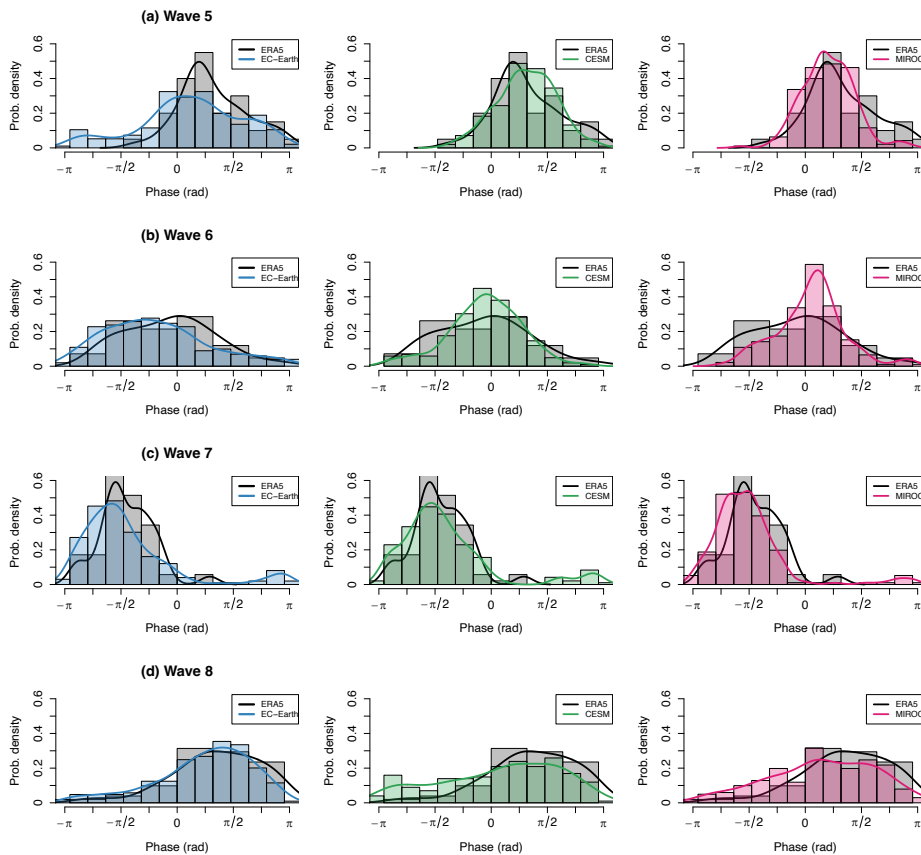
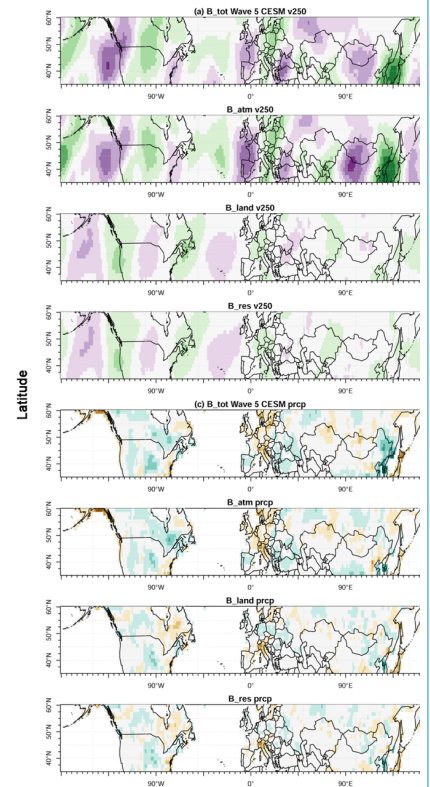



Figure B1: Bias plots for high-amplitude wave-5 events in different experiments for CESM. Total bias(B_tot), atmospheric bias(B_atm), land-atmosphere interaction bias (B_land) and residual bias(B_res) for meridional wind velocity at 250hPa (a), surface temperature (b), precipitation (c), and seal level pressure (d).¶

¶
¶
¶

... [1]

780

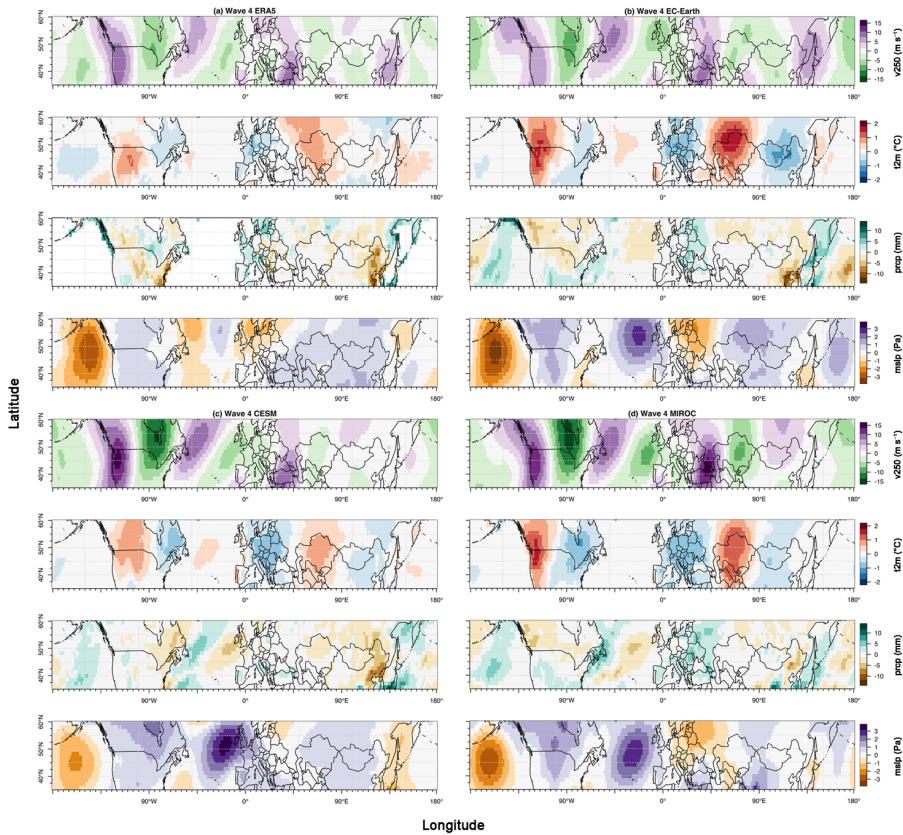



Figure B3: Composite anomaly plots of weeks with high-amplitude waves-4 events for meridional wind velocity at 250hPa (v250, absolute field), near-surface temperature (t2m, anomaly), precipitation (prcp, anomaly), and seal level pressure (mslp, anomaly) in ERA5 (a), EC-Earth (b), CESM (c) and MIROC (d) based on control runs AISI.

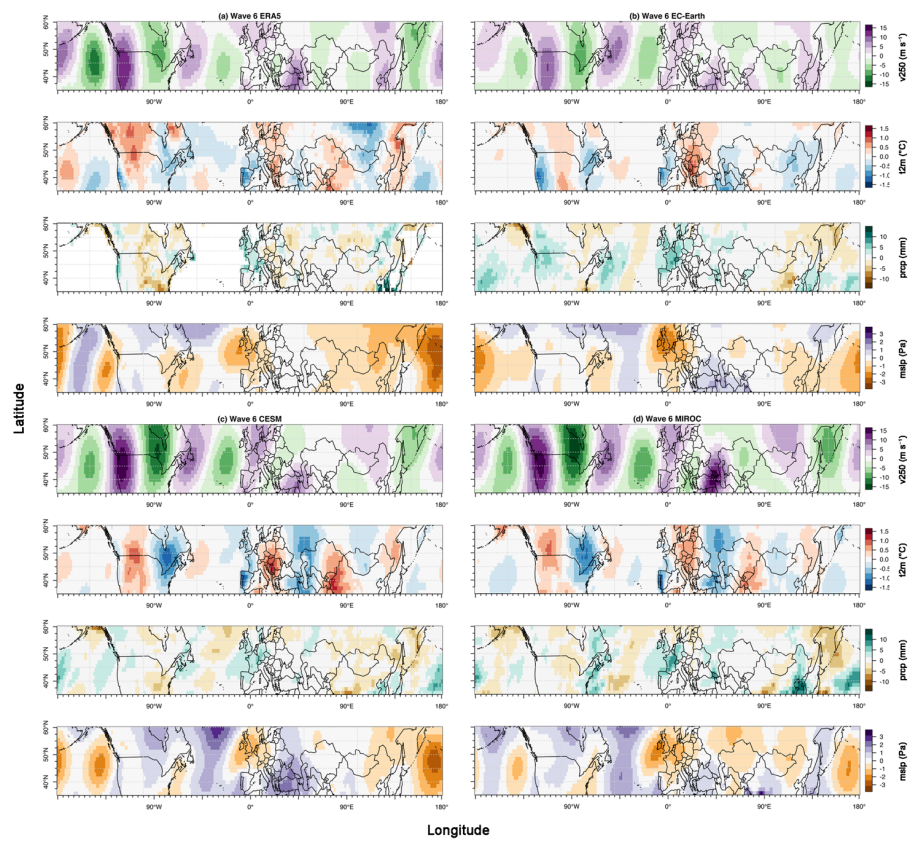



Figure B4: Composite anomaly plots of weeks with high-amplitude waves-6 events for meridional wind velocity at 250hPa (v250, absolute field), near-surface temperature (t2m, anomaly), precipitation (prcp, anomaly), and seal level pressure (mslp, anomaly) in ERA5 (a), EC-Earth (b), CESM (c) and MIROC (d) based on control runs AISI.

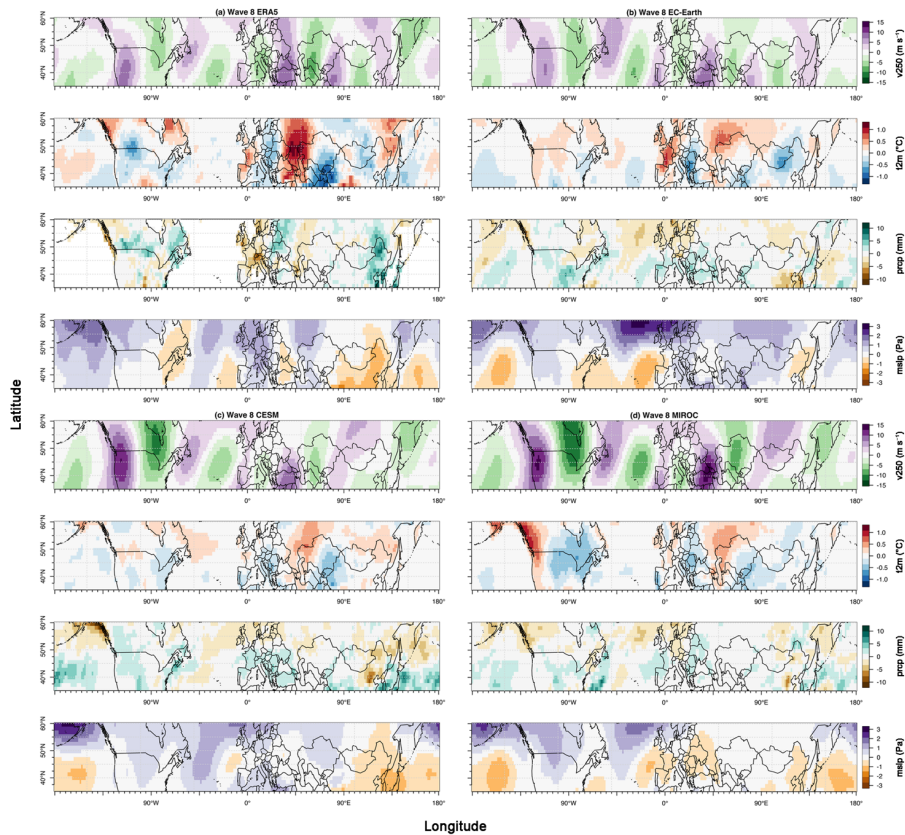



Figure B5: Composite anomaly plots of weeks with high-amplitude waves-8 events for meridional wind velocity at 250hPa (v250, absolute field), near-surface temperature (t2m, anomaly), precipitation (prcp, anomaly), and seal level pressure (mslp, anomaly) in ERA5 (a), EC-Earth (b), CESM (c) and MIROC (d) based on control runs AISI.

790

795

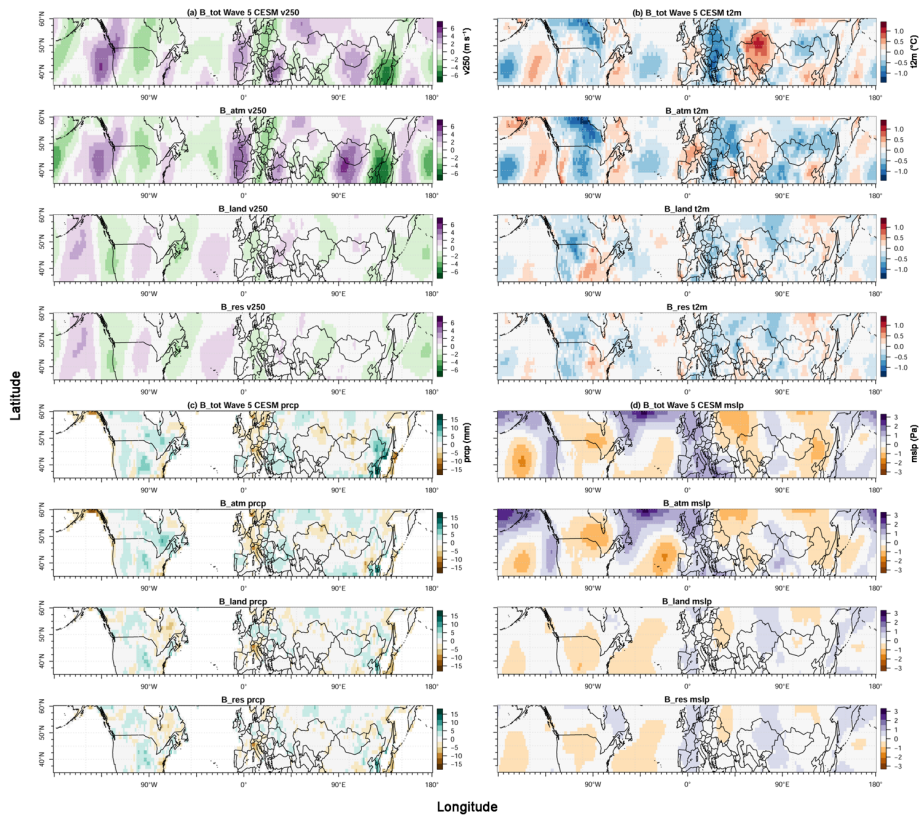



Figure B6: Bias plots for high-amplitude wave-5 events in different experiments for CESM. Total bias(B tot), atmospheric bias(B atm), land-atmosphere interaction bias (B land) and residual bias(B res) for meridional wind velocity at 250hPa (a), surface temperature (b), precipitation (c), and seal level pressure (d).

800

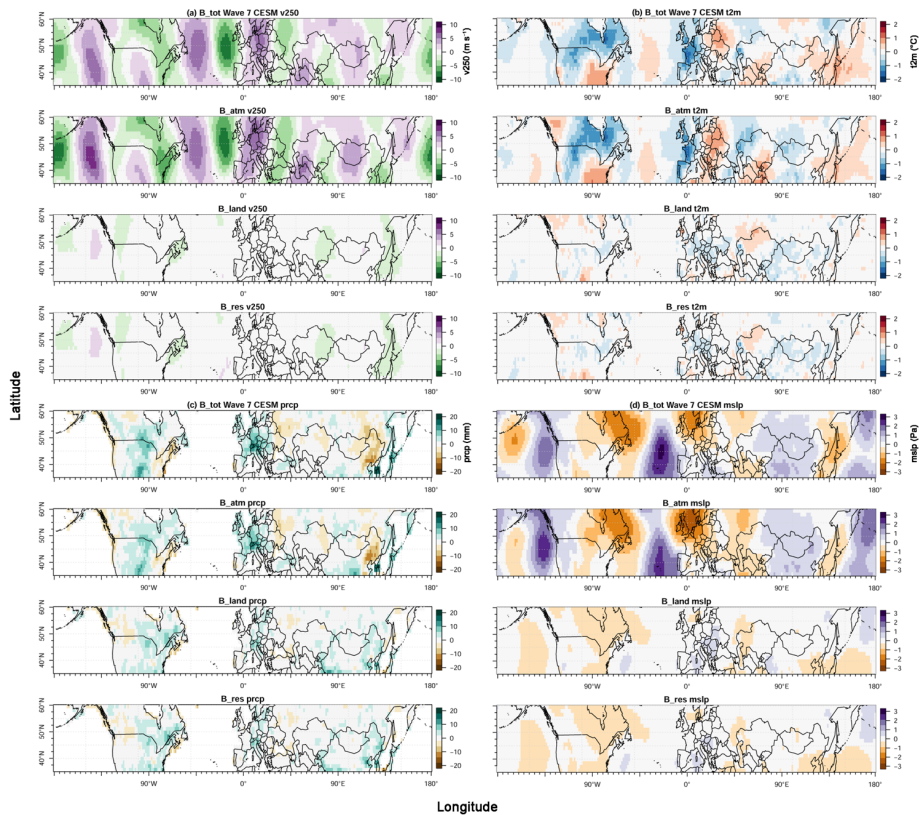



Figure B7: Bias plots for high-amplitude wave-7 events in different experiments for CESM. Total bias(B_tot), Atmospheric bias(B_atm), Land-Atm interaction bias(B_land) and residual bias(B_res) for meridional wind velocity at 250hPa (a), surface temperature (b), precipitation (c), and seal level pressure (d).

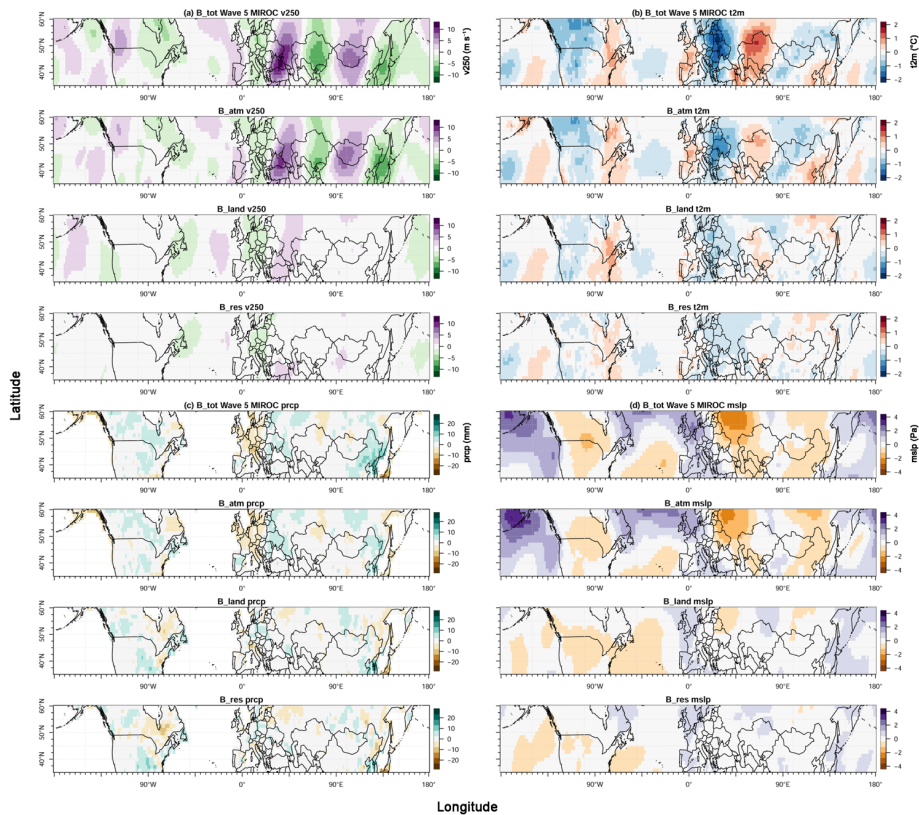
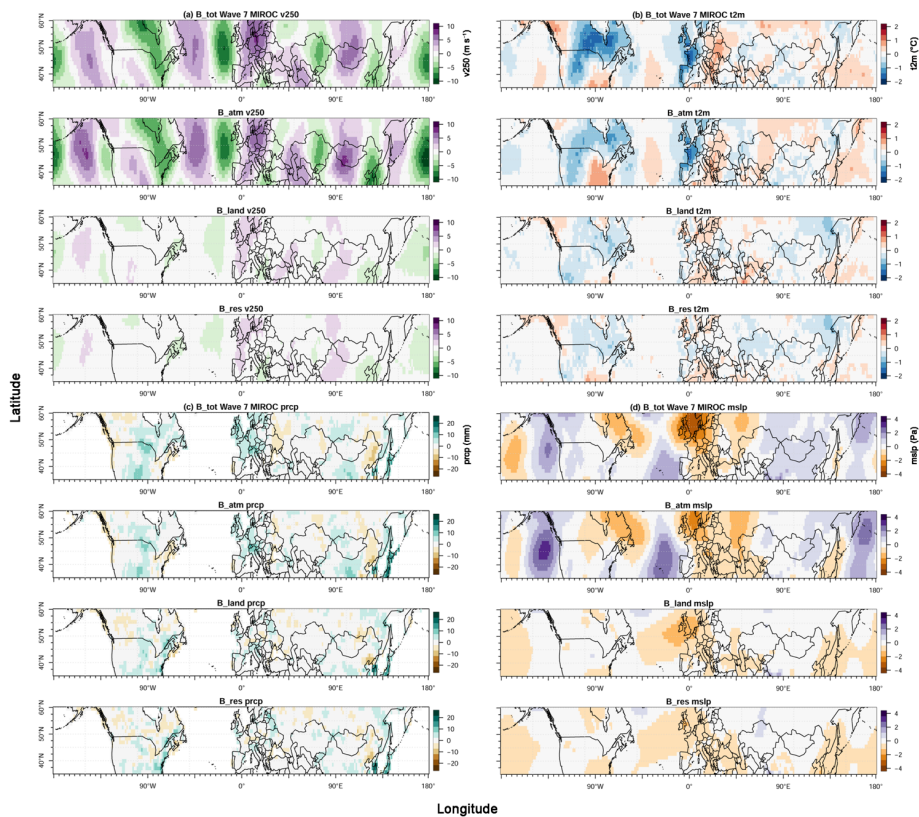
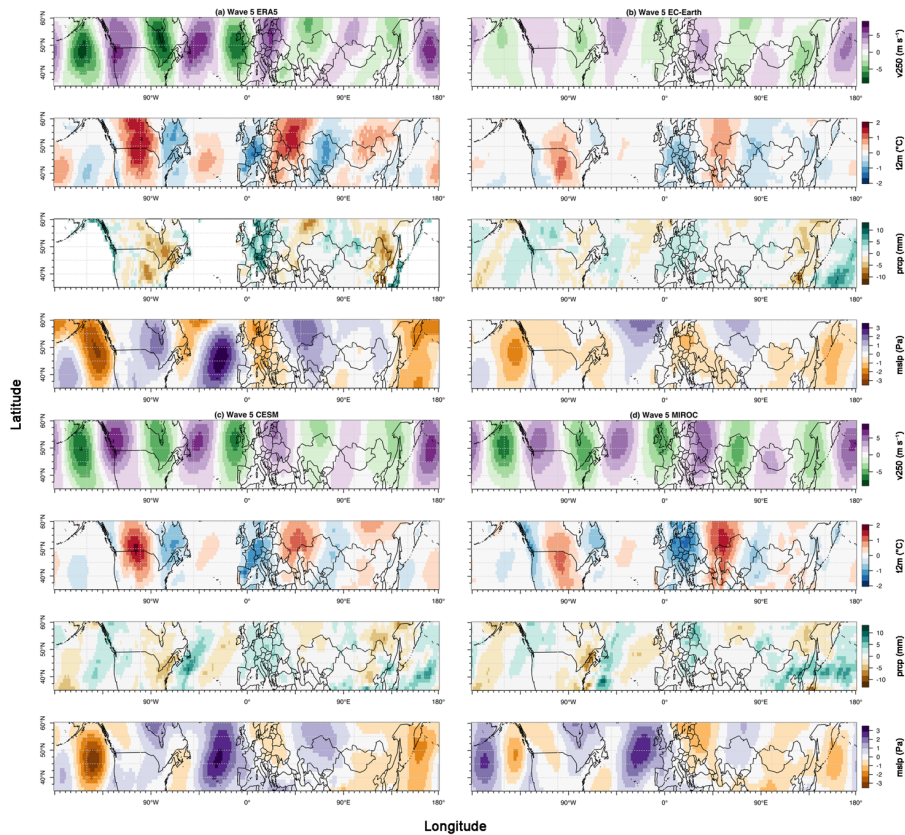
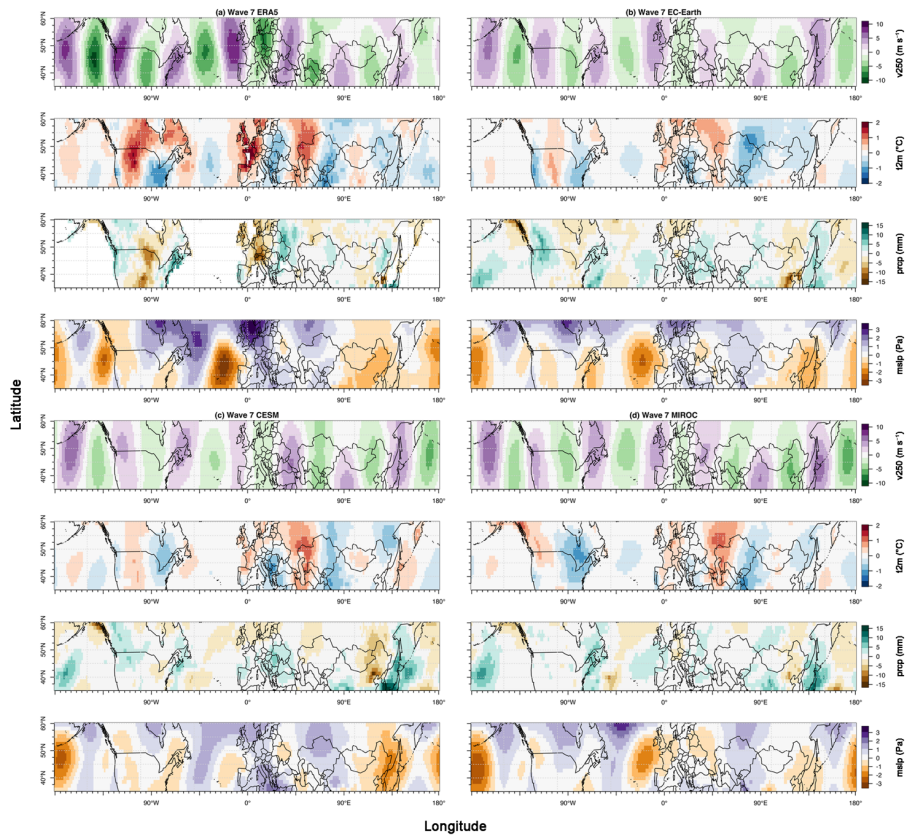



Figure B8: Bias plots for high-amplitude wave-5 events in different experiments for MIROC. Total bias(B_tot), atmospheric bias(B_atm), land-atmosphere interaction bias (B_land) and residual bias(B_res) for meridional wind velocity at 250hPa (a), surface temperature (b), precipitation (c), and seal level pressure (d).

810

815 **Figure B9: Bias plots for high-amplitude wave-7 events in different experiments for MIROC. Total bias(B_tot), Atmospheric bias(B_atm), Land-Atm interaction bias(B_land) and residual bias(B_res) for meridional wind velocity at 250hPa (a), surface temperature (b), precipitation (c), and seal level pressure (d).**

820   Figure B10: Composite anomaly plots of weeks with high-amplitude waves-5 events for meridional wind velocity at 250hPa (v250, anomaly), near-surface temperature (t2m, anomaly), precipitation (prcp, anomaly), and seal level pressure (mslp, anomaly) in ERA5 (a), EC-Earth (b), CESM (c) and MIROC (d) based on control runs AISI.

825 Figure B11: Composite anomaly plots of weeks with high-amplitude waves-7 events for meridional wind velocity at 250hPa (v250, anomaly), near-surface temperature (t2m, anomaly), precipitation (prcp, anomaly), and seal level pressure (mslp, anomaly) in ERA5 (a), EC-Earth (b), CESM (c) and MIROC (d) based on control runs AISI.

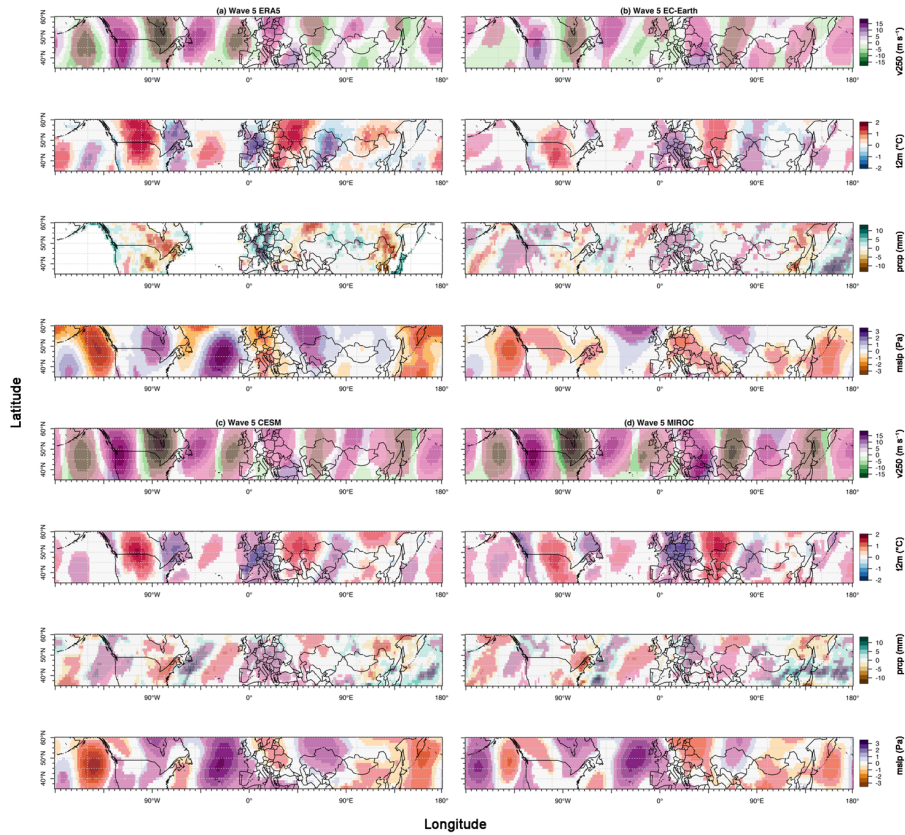



Figure B12: Same as Fig. 3 but with significant applied at 95% confidence level (areas highlighted in shade fuchsia are statistically significant after FDR method accounted for).

830

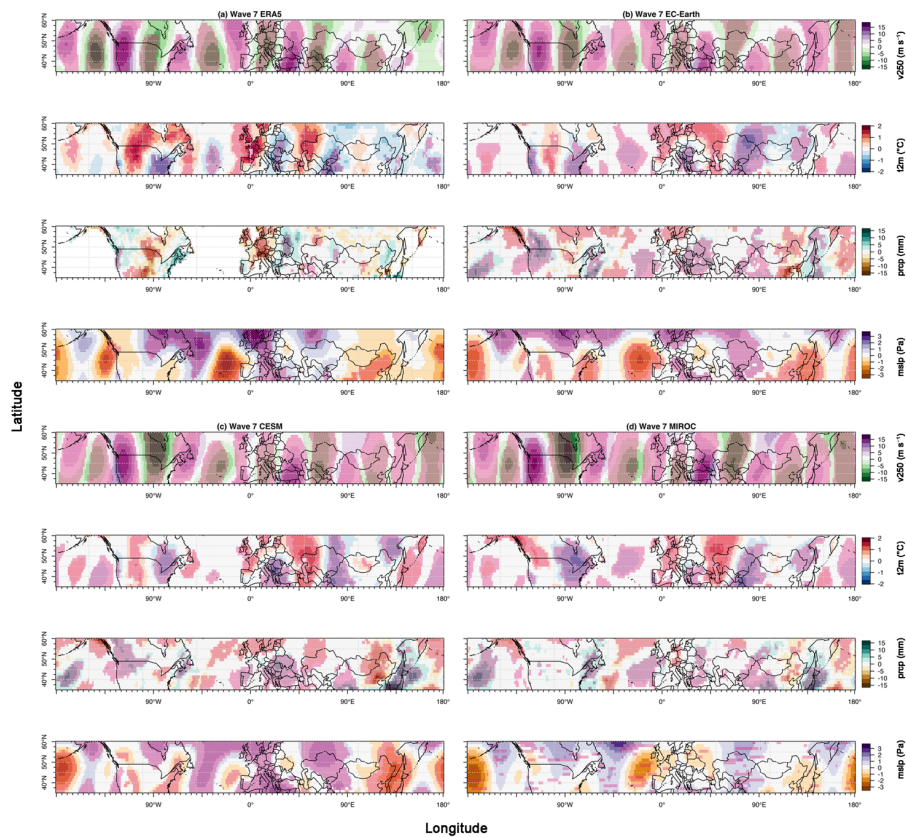



Figure B13: Same as Fig. 4 but with significant applied at 95% confidence level (areas highlighted in shade fuchsia are statistically significant after FDR method accounted for).

835