Validation of boreal summer tropical-extratropical causal links in seasonal forecasts

Giorgia Di Capua[1,2], Dim Coumou[2,3,4], Bart van den Hurk[3,5], Antje Weissheimer[6,7], Andrew G. Turner[8,9]
10 and Reik V. Donner[1,2]

[1]Department of Water, Environment, Construction and Safety, Magdeburg-Stendal University of Applied Sciences, Magdeburg, 39114, Germany
[2]Earth System analysis, Potsdam Institute for Climate Impact Research (PIK) – Member of the Leibniz Association, Potsdam, 14473, Germany
20 [3]Water and Climate Risk Department, Institute for Environmental Studies (IVM), VU University of Amsterdam, Amsterdam, 1081 HV, Netherlands
[4]Royal Netherlands Meteorological Institute (KNMI), De Bilt, 3730 AE, Netherlands
[5]Department of Flood Risk Management, Deltares, Delft, 2629 HV, Netherlands
[6]European Centre for Medium-Range Weather Forecasts, Reading, RG2 9AX, UK
25 [7]National Centre for Atmospheric Science, University of Oxford, Oxford, OX1 4BH, United Kingdom
[8]Department of Meteorology, University of Reading, Reading, RG6 6UR, United Kingdom
[9]National Centre for Atmospheric Science, University of Reading, Reading, RG6 6UR, United Kingdom

Correspondence to: Giorgia Di Capua (dicapua@pik-potsdam.de)

Abstract. Much of the forecast skill in the mid-latitudes on seasonal timescales originates from deep convection in the tropical belt. For boreal summer, such tropical-extratropical teleconnections are less well understood as compared to winter. Here we validate the representation of boreal tropical – extratropical teleconnections in a general circulation model in comparison with observational data. To characterise variability between tropical convective activity and mid-latitude circulation, we identify the South Asian monsoon (SAM) – circumglobal teleconnection (CGT) pattern and the western North Pacific summer monsoon (WNPSM) – North Pacific high (NPH) pairs as the leading modes of tropical-extratropical coupled variability in both reanalysis (ERA5) and seasonal forecast (SEAS5) data. We calculate causal maps, an application of the PCMCI causal discovery algorithm which identifies causal links in a 2D field, to show the causal effect of each of these patterns on circulation and convection in the Northern Hemisphere. The spatial patterns and signs of the causal links in SEAS5 closely resemble those seen in ERA5, independent of the initialization date of SEAS5. However, the strength of causal links in SEAS5 is often too weak (about two thirds of those in ERA5). By performing a subsampling (over time) experiment, we identify those regions for which SEAS5 data well reproduce ERA5 values, e.g. South-eastern US, and highlight those where the bias is more prominent, e.g. North Africa. We demonstrate that different El Niño – Southern Oscillation phases have only a marginal effect on the strength of these links. Finally, we discuss the potential role of model mean-state biases in explaining differences between SEAS5 and ERA5 causal links.

1 Introduction

Seasonal forecasts provide a useful tool to study atmospheric dynamics and predict seasonal variations in wind, rainfall
55 and temperature patterns across tropical and extratropical regions (Bauer et al., 2015; Palmer and Anderson, 1994). To a
certain extent, seasonal forecasts can be used by stakeholders and governments to anticipate and mitigate extreme weather
events, failures in crop yields, water scarcity hazards for infrastructure such as electricity grids (Lazo et al., 2009; Challinor
et al., 2003; Hagger et al., 2018; Meza et al., 2008). Tropical – extratropical interactions are linked to mid-latitude boreal
surface weather conditions and represent a source of predictability at seasonal and subseasonal timescales (Shukla, 1998).
60 Hence, improving the representation of these teleconnections in seasonal forecasts can help to improve our knowledge of
atmospheric dynamics as well as helping to better forecast relevant weather patterns to support early warning.

Obtaining reliable seasonal forecasts is a challenging problem due to the intrinsic nonlinearity of processes governing
atmospheric motions (Holton, 1973). While providing weather forecasts beyond a two-week threshold is a complex problem
due to the chaotic nature of atmospheric processes (Tsonis and Eisner, 1989; Palmer and Anderson, 1994), slowly varying
65 climatic fields such as sea surface temperatures (SST) and soil moisture can provide forecast skill beyond the weekly time
scale (Charney and Shukla, 1981). The representation of the interaction of the atmosphere with other components of the
climate system, e.g., SST, is an important requirement to achieve forecast skill (Roberts et al., 2021; Tietsche et al., 2020).
Historically, both statistical (Gadgil et al., 2005; Kumar, 2012) and dynamical approaches (Jain et al., 2018; Scaife et al.,
2018) have been used to provide seasonal forecasts, often with comparable skill (Seo et al., 2009; Barnston et al., 1999).
70 However, dynamical forecasts, generated by general circulation models (GCMs), provide a more complete representation of
the atmospheric physics that governs weather and climate behaviour (Shukla et al., 2000). Therefore, dynamical seasonal
forecasts are better suited for representing the dynamic and thermodynamic processes and emerging dynamical
teleconnections within the climate system.

To produce accurate seasonal forecasts, GCMs need to represent the physical processes operating at those timescales
75 truthfully in the current climate. Great progress in this field has been made in recent decades, leading to an improved
representation of dynamic and thermodynamic processes and a steady increase in model resolution (Bauer et al., 2015;
Palmer, 2017; Haarsma et al., 2016). Nevertheless, probabilistic reliability of seasonal forecasts remains limited
(Weisheimer and Palmer, 2014). In this context, analysing seasonal forecasts such as the SEAS5 (Johnson et al., 2019)
dataset provided by the European Centre for Medium-range Weather Forecasts (ECMWF), can help to identify, understand
80 and improve biases between observations and model simulations.

Tropical – extratropical interactions in the Northern Hemisphere during boreal summer have been analysed in several
recent studies (O'Reilly et al., 2018; Di Capua et al., 2020a; Ding et al., 2011). Heat generated by tropical convective
activity provides a source of wave activity which can affect weather in the mid-latitudes (O'Reilly et al., 2018; Rodwell and
Hoskins, 1996; Ding and Wang, 2005), while in turn mid-latitude wave activity can modulate rainfall events in the tropical
85 belt (Ding and Wang, 2007; Di Capua et al., 2020b). Here, we focus on the two main modes of covariability between tropical

convection and mid-latitude circulation as defined in Di Capua et al. (2020a) and Ding et al. (2011). The first mode of covariability between tropical convective activity and mid-latitude circulation is represented by the circumglobal teleconnection (CGT) paired with the South Asian monsoon (SAM) convection (Ding et al., 2011; Di Capua et al., 2020a). Circumglobal wave trains such as the CGT are connected to temperature and precipitation anomalies at intraseasonal and

90   interannual timescales in the northern mid-latitudes (Ding and Wang, 2005; Di Capua et al., 2020b). Recent work based on statistically evaluating causal relationships in reanalysis data has shown that the CGT pattern and the SAM circulation system are connected by a two-way causal interaction (Di Capua et al., 2020b). Moreover, the causal effect of each of these patterns on atmospheric circulation and surface conditions can be effectively represented on a 2D map (Di Capua et al., 2020a). The CGT has been studied in seasonal forecasts provided by ECMWF, and the corresponding results show that

95   generally the model can reproduce this pattern (Beverley et al., 2019). However, the CGT pattern in seasonal forecasts is too weak, likely due to a misrepresentation of the SAM convective activity in the tropical belt.

The second mode of covariability between tropical convection and boreal summer circulation is represented by a pair of patterns consisting of the western North Pacific summer monsoon (WNPSM) and the North Pacific high (NPH) (Di Capua et al., 2020a). The NPH is the results of the northward displacement of the North Pacific sub-tropical high due to the onset of

100   the WNPSM activity beginning of July (Di Capua et al., 2020a). In reanalyses, the influence of these two patterns on other atmospheric fields is weak and mostly confined to the Pacific Ocean as compared to the SAM – CGT pair. Nevertheless, the WNPSM and NPH systems can affect typhoon cyclogenesis in the tropical Pacific (Briegel and Frank, 1997) and temperature and circulation patterns in East Asia and North America, respectively, potentially acting as a source of wave activity downstream (Di Capua et al., 2020a; Ding et al., 2011). Therefore, even though the direct area of influence of the

105   WNPSM – NPH pair is found over the ocean, effects of changes in their intraseasonal variability are relevant to remote and highly populated areas (e.g., US west coast or Japan).

Causal discovery algorithms, such as the Peter and Clark Momentary Conditional Independence (PCMCI) method, help overcome issues with commonly used statistical techniques, like correlation measures. When carefully applied, they allow to identify and select causal versus spurious links (Runge, 2018; Runge et al., 2014, 2019). PCMCI has been used to study

110   stratospheric polar vortex variability (Kretschmer et al., 2017, 2016, 2018), the Silk Road pattern interdecadal variability (Stephan et al., 2019), Atlantic hurricane activity (Pfleiderer et al., 2020), and causal interactions between the Indian summer monsoon and mid-latitude wave trains (Di Capua et al., 2020a, b). Moreover, PCMCI has also proven useful in providing early forecasts of Moroccan crops (Lehmann et al., 2020), sub-seasonal statistical forecasts of US surface temperatures (Vijverberg and Coumou, 2022; Vijverberg et al., 2020) and statistical seasonal predictions of Indian summer monsoon

115   rainfall (Di Capua et al., 2019).

Process-based validation can help us to understand and correct biases in seasonal forecasts (Eyring et al., 2019; Horak et al., 2021). Here, we propose to use causal discovery to perform a process-based validation (Nowack et al., 2020) of tropical – extratropical interactions in SEAS5 seasonal forecasts. We compare observed (reanalysis) causal interactions between tropical convective activity and mid-latitude wave trains in the Northern Hemisphere during boreal summer with those

120 provided by seasonal forecasts. The scope of this comparison is three-fold: (i) we validate causal links in a coupled general circulation model (GCM) in forecasting mode against those derived from observations and (ii) we gather information on missing or misrepresented links in the GCM in forecast mode. Finally, (iii) we analyse whether these differences can be attributed to model biases and what impact different phases of the El Niño – Southern Oscillation (ENSO), present in the initial conditions of the forecasts, have on the strength and representation of causal links. Thus, this work represents a

125 preliminary step to improving forecast skill, and the representation of tropical – extratropical teleconnections in GCMs. The remainder of this paper is organized as follows: Section 2 presents the data and methods used. Section 3 describes the results obtained by applying causal maps first to ERA5 reanalysis and then SEAS5 data. Section 4 provides a discussion of the obtained results in the context of the existing literature and the final conclusions.

2 Data and Methods

130 ### 2.1 Data

We analyse intraseasonal (weekly) tropical convective activity and mid-latitude circulation characteristics using gridded data (0.25°x0.25° upscaled to 2°x2°) from the ERA5 reanalysis dataset (Hersbach et al., 2020) and the SEAS5 seasonal retrospective forecast dataset (Johnson et al., 2019), both provided by the European Centre for Medium-range Weather Forecasts (ECMWF). From the ERA5 dataset, we use daily (temporally averaged to obtain weekly samples)

135 geopotential height fields at 200 hPa (Z200), outgoing longwave radiation (OLR), sea surface temperature (SST) and zonal (U200) wind fields for the period 1979-2020 and for the subset 1993-2016 (to be consistent with the available SEAS5 dates). While Z200 is useful for representing the mid-latitude circulation, OLR can be used as proxy of tropical convective activity. Despite SEAS5 providing values for precipitation globally, we prefer OLR to precipitation itself as precipitation is not assimilated in reanalysis (thus less reliable than assimilated fields such as SST), observational precipitation data coverage for

140 tropical regions is sparse and to keep this analysis consistent with Di Capua et al. (2020a). For the general circulation model comparison, we analyse OLR, Z200, SST and U200 fields for the period 1993-2016 of SEAS5. As for ERA5, also SEAS5 data are re-gridded from the original 1°x1° on a 2°x2° grid and daily data are temporally averaged to obtain weekly samples. The interannual variability, seasonal cycle and any long-term trend are removed. While reanalysis data from ERA5 provide one realization per year, SEAS5 provides 25 ensemble members for each year, thus a total of 25x24 (600) model years. A

145 schematic of the SEAS5 ensemble is shown in Fig. 1. Considering only the summer season (June to September) plus the month of May, as required by the causal discovery framework to ensure a proper handling of time lags (see Section 2.3), we have a total of 21 weeks per extended summer (MJJAS) for each year (Fig. 1a). Thus, for the common period 1993-2016, we have a total of 21x24=504 weekly time samples for ERA5 and 21x24x25=12,600 time samples for SEAS5 (Fig. 1b). Note that for both ERA5 and SEAS5 datasets, the first week of the MJJAS period starts on the 7th of May (the first full week

150 common to both datasets when for ERA5 the first week of the year starts on the 1st of January).

Treating the 25 ensemble members per year for 24 years as one unique time series composed of distinct subsequences is possible under the assumption that each member of each ensemble year is independent of the rest of the members. While the ensemble is generated by varying the initial conditions, this uniqueness assumption is not true in general, since each ensemble member for a given year has common lower boundary conditions for the atmosphere, inherited from slowly

155   evolving features in the ocean, e.g. SST anomalies in both tropical and extratropical regions. However, for the purpose of this analysis we are interested in the *relative* effect of a certain (set of) variable(s) on the remaining atmospheric fields *inside* the intraseasonal variability, with a maximum lag of a few weeks (thus on a much shorter time scale than interannual). It must be noted that we do not intend to use SEAS5 data to assess or exploit its forecast skill, but instead assess the ability of a general circulation model in forecasting mode at reproducing observed tropical – extratropical teleconnections. In other

160   words, whether there is shared information between two ensemble members for a certain year, e.g. a certain phenomenon in the analysed climate system is stronger or weaker, or whether a specific year shows better forecast skill than another, does not affect the relative effect of that phenomenon on some selected atmospheric fields. However, we cannot exclude that different initialisation dates (SEAS5 seasonal forecasts are initialized on the first day of each month and run for 7 months) and the vicinity of the target season to the beginning of the simulation may influence the outcome and the resulting causal

165   links. Thus, we choose to analyse SEAS5 forecasts initialized on both the 1st of March and on the 1st of May, with a target season of June-September (see Fig. 1a). This way, the model has up to three (and at least one) months of spin-up to reduce the influence of the initial conditions, and we provide a sensitivity analysis to show which results depend on (or are independent of) the chosen initialization date. In the final step of this work, we will assess whether the effect of ENSO on wind fields and convective activity in the Northern Hemisphere is (i) sufficiently well reproduced in SEAS5 when compared

170   to ERA5, and (ii) influences the identified tropical – extratropical causal links.

2.2 Modes of co-variability

To identify the dominant modes of intraseasonal co-variability between tropical convective activity and mid-latitude circulation in the Northern Hemisphere, we apply maximum covariance analysis (MCA) as described in Di Capua et al. (2020a). The first two MCA modes are calculated by applying MCA to OLR fields in the tropical belt (15°S-30°N, 0°-

175   360°E) paired with Z200 fields in the northern mid-latitudes (25°N-75°N, 0°-360°E), thus using the same geographical borders as in Di Capua et al. (2020a).

By construction, the first MCA mode explains a maximum of *squared covariance between the selected two fields* (Ding et al., 2011; Wiedermann et al., 2017) and MCA modes are ranked according to their explained squared covariance fraction (SCF) (Wilks, 2011). With this method, it is possible to identify pairs of patterns that can explain shared covariance and (to

180   some extent) evolve simultaneously. However, shared covariance, as in general correlation-based techniques, does not imply causality. To check whether these patterns may be causally related (e.g., via dynamical mechanisms) we apply the PCMCI causal discovery algorithm (see Section 2.3).

Each MCA mode provides two coupled (2D) spatial patterns (one for tropical OLR and one for mid-latitude Z200) and two associated time series. These time series are obtained by projecting each (2D) MCA spatial pattern on the corresponding time-varying atmospheric field, and represent the time-dependent MCA scores or pattern amplitudes for both fields. Each time series describes the magnitude (prominence) and phase (sign) of those patterns for each time step of the field's time series. The acronyms associated with the MCA patterns are shown in Table 1.

Here, we apply MCA both to for the full 1979-2020 (ERA-L) and reduced 1993-2016 (ERA-S) periods and to SEAS5 for the period 1993-2016. On one hand, we attempt to be consistent with the available SEAS5 data (ERA-S). On the other hand, we want to provide a direct comparison with Di Capua et al. (2020a), who studied the previous ERA-Interim reanalysis product, by exploiting the full length of the time series (ERA-L). However, the first two MCA modes calculated with SEAS5 data do not provide a close enough representation of the patterns shown in ERA5 (see spatial correlation coefficients shown in Table 2 and Section 3.1). Thus, to provide a meaningful comparison between the two datasets, we define MCA patterns in SEAS5 by projecting the first two ERA5 MCA patterns onto SEAS5 data. This is done in the same way in which the ERA5 MCA scores/time series are calculated, i.e., by calculating the dot product of each ERA5 MCA mode with the corresponding OLR or Z200 field time series for each time step.

2.3 PCMCI and Causal Maps

The PCMCI algorithm is a causal discovery method using partial correlations to iteratively test for causality between two (or more) time series (*actors*) given a certain set of conditioning variables (Runge, 2018; Runge et al., 2014, 2019; Spirtes et al., 2000). The term *causal* builds upon a series of hypotheses, such as causal sufficiency (i.e. all relevant actors are included in the network) and stationarity of the detected causal chain (i.e. the causal links are stationary over time and actors show no trend). Hence, the detected causal links are valid in the set of analysed actors and also depend on the linear or non-linear framework applied as well as a set of parameters, such as the significance threshold α. The causal links identified by PCMCI are represented in a so-called Causal Effect Network (CEN), a graph where each actor is represented by a node and causal links are shown as arrows connecting different nodes (Fig. 2a). The sign and strength of a certain causal link are given by the β coefficient, represented in the CEN by the colour of the arrow. For example, given the causal link $actor2_{\tau=-1} \rightarrow actor1_{\tau=0}$, a β coeff. of 0.25 represent a positive change of 0.25 standard deviation units (s.d.) of actor1 at lag 0 due to a positive change of 1 s.d. in actor2 at lag -1.

Here we apply the concept of *causal maps*, an extension of PCMCI to spatial fields of variables, to analyse the influence of a set of spatial patterns representing tropical – extratropical summer interactions (identified by applying MCA) on a 2D field (Di Capua et al., 2020a). Causal maps make use of the concept of CEN and the PCMCI algorithm. However, instead of showing the typical network-like shape with actors connected by arrows representing the direction, sign and strength of the causal links as shown in Fig. 2a, causal maps provide information in a similar conceptual manner as a classical correlation map (Fig. 2b). In a causal map, however, each grid point represents the sign and strength (given by the β

215   coeff.) of a certain causal link, e.g. between actor1 and actor3, while the direction of the link and the set of actors involved is kept constant throughout each map (more maps are necessary to show different configurations of actors as shown in Fig. 2b).

To provide a meaningful comparison with previous work, in this analysis we apply the same framework as used in Di Capua et al. (2020a), i.e. each causal map is obtained by running at grid point level a CEN analysis with three actors. Of these three actors, two represent the pair of time series obtained for each MCA mode and are kept constant throughout the
220   map (actor1 and actor2 in our example in Fig. 2a). The third time series is the time series for any individual grid point of the considered time-varying field (OLR or Z200) and thereby depends on longitude and latitude (actor3(lat, lon) in Fig. 2a). Thus, for each MCA mode we will have eight causal maps, as we have two target fields (OLR and Z200), two MCA modes and two MCA time series for each MCA mode (one for tropical OLR and one for mid-latitudes Z200). This can be summarized in Eq. (1):

225   $$MCA_i^k \rightarrow field_j | MCA_i^{k \neq l} \qquad\qquad (1)$$

where $i \in \{trop.\ OLR, midlat.\ Z200\}$, $j \in \{Z200, OLR\}$ and $k, l \in \{1,2\}$ being the two MCA modes.

Here, we use the PCMCI version 4.1 from the Python package tigramite (https://github.com/jakobrunge/tigramite). We estimate causal maps with the parameters lag min = -1, lag max = -2 (units of the sampling frequency, i.e., weekly) and pc_alpha = 0.2 (unless otherwise indicated we use the default parameters of PCMCI). Note that pc_alpha is not the final
230   significance threshold adopted to determine the significance of the identified causal links, instead it is a parameter used in the PC step of the PCMCI algorithm, which if taken too strict, e.g. the usual α = 0.05, would prevent the algorithm from retaining any potentially meaningful links (for further details see https://jakobrunge.github.io/tigramite/). The significance threshold adopted for plotting the results is α = 0.05 and we apply a false discovery rate (FDR) correction (Benjamini and Hochberg, 1995) to control for multiple testing among the multiple grid locations in causal maps. Moreover, the robustness
235   of causal map is assessed by calculating causal maps for a range of sub-periods. In 10 trials, we iteratively remove 10% of the record: 60 year out of 600, in the first iteration year 1-to-60 are removed, in the second year 61-to-120 are remover and so on. Finally, we plot β values for those grid points were a significant β value is found for at least 70% of the cases (similarly to what was done in Di Capua et al. 2020a).

3 Results

240   The following results section is organized as follows: first we define and describe MCA patterns both in ERA5 and SEAS5 datasets (Section 3.1). Then we calculate the respective causal maps and compare those obtained in ERA5 with those obtained with SEAS5 (Section 3.2). In Section 3.3 we produce a 1000 member subsampling experiment to determine whether ERA5 β values fall in the range of realisations of SEAS5. Following those studies, in Section 3.4, we analyse the role that model biases may play in explaining ERA5-SEAS5 differences and finally, in Section 3.5, we check whether ENSO
245   may influence these tropical – extratropical teleconnections.

3.1 MCA patterns in SEAS5 and ERA5

We first calculate the two leading maximum covariance analysis (MCA) coupled modes of tropical – extratropical co-variability between mid-latitude (25°-75°N) geopotential height at 200 hPa and tropical (15°S-30°N) outgoing longwave radiation in ERA5 reanalysis data for the period 1993-2016 (ERA-S, Fig. 3). MCA modes calculated for the extended period

250 1979-2020 (ERA-L) are shown in Fig. S1 in the Supplementary Material. The two pairs of patterns identified in this way are the South Asian Monsoon (SAM, Fig. 3d) paired with the circumglobal teleconnection (CGT, Fig. 3a) pattern for MCA mode 1, and the western North Pacific Summer Monsoon (WNPSM, Fig. 3j) paired with the North Pacific High (NPH, Fig. 3g) for MCA mode 2.

The SAM is characterized by a large rainfall band stretching from the Arabian Sea towards the western edge of the South

255 China Sea, with a peak of negative OLR (relatively high rainfall) centred over the Indian peninsula. The CGT pattern shows five centres of positive Z200 anomalies over the Iberian Peninsula, central Asia on the western side of the Caspian Sea, East China, the North Pacific and southeast US. The WNPSM pattern features a region of enhanced convective activity over the tropical western Pacific between 25° and 30°N accompanied by an area of suppressed convective activity on its west side centred over the South China Sea. The main feature of the NPH is a ridge in Z200 located in the western side of the North Pacific, however this pattern

260 also shows a zonally oriented wave train similar to the CGT pattern, but with centres of action shifted in longitude.

The key features described above for the ERA-S MCA modes 1 and 2 (Fig. 3a,d,g,j) are qualitatively very close to those found for the 1979-2020 period (ERA-L, Fig. S1a,d,g,j) or ERA-Interim (Figs. 3 and 4 in Di Capua et al. (2020a)). The spatial correlation between ERA-L and ERA-S MCA modes ranges between 0.8 and 0.9 (Table 2), showing that these patterns are robust across different versions of ERA reanalysis and quite insensitive to the chose period. However, when MCA modes are calculated

265 on SEAS5 seasonal forecasts initialized on the 1st of May and compared to ERA-S, a few key differences arise (Fig. S2). Overall each ERA-S MCA mode (Fig. 3a,d,g,j) shares common features with its corresponding SEAS5 MCA mode (Fig. S2a,d,g,j), the similarity of which can be quantified by calculating the spatial correlation (Table 2, second column), yielding values ranging between 0.4 and 0.6.

However, some features that in ERA-S characterise MCA 1 are also found in SEAS5 MCA2, and vice versa for ERA-S

270 MCA 2. Thus, features that in ERA-S are separated and characterise each of the first two MCA modes, appear mixed in SEAS5 (or vice versa). For example, in ERA-S the CGT-related high east of the Caspian Sea is only visible in MCA 1 (Fig. 3a), while in SEAS5, the same positive signal is found in both MCA 1 and MCA 2 (although stronger in MCA 1). Similarly, the convective activity over the Indian peninsula, which in ERA-S characterizes only MCA 1 (Fig. 3d), is found in both MCA 1 and MCA 2 of SEAS5 (Fig. S2d,j). In contrast, the wave pattern over Eurasia characterising ERA-S MCA 2 showing a high pressure region over

275 Eastern Europe and a low over Central Asia (Fig. 3g) is not found in either of the first two SEAS5 MCA modes (Fig. S2a,g). This can be seen in spatial correlation values calculated between ERA-S MCA 1 (MCA 2) and SEAS5 MCA 2 (MCA 1) which also range between 0.4 and 0.6, although sometimes with reversed sign (opposite phase of the pattern) (Table 2, third column). Therefore, we conclude that despite the general resemblance of the first two SEAS5 MCA modes to those calculated in ERA-S,

working with a mixed signal would hinder the interpretability of parts of the results, since it would make it difficult to effectively
280 separate the causal effect of the SAM – CGT patterns from that of the WNPSM – NPH pair.

To account for the aforementioned problem, SEAS5 MCA modes have been re-calculated by projecting ERA-S MCA patterns onto SEAS5 Z200 and OLR 3D fields, as explained in the Data and Methods section. Then, to visualise the equivalent SEAS5-ERA5 MCA modes, composites of time steps with the MCA time series values higher than 1 standard deviation (s.d.) are calculated (Fig. 4a,d,g,j). As a result, a much closer resemblance between SEAS5-ERA5 and ERA-SMCA patterns is obtained and
285 the spatial correlation coefficients between SEAS-ERA5 and ERA-S MCA patterns reach values of +~0.8-0.9 (Table 2, fourth column).

3.2 Causal maps in SEAS5 and ERA5

Having defined ERA-S and SEAS5-ERA5 MCA modes, we apply the concept of causal maps to detect the causal effect of each of these four patterns (two for each mode) to Z200 and OLR fields, in the range 15°S-75°N.
290 Causal maps calculated for the effect of ERA-S SAM and CGT on Z200 and OLR fields are shown in Fig. 3b,c and 3e,f respectively. The main features are the effect of SAM on Sahel Z200, the wave train in the North Pacific and the negative effect on Z200 in Central Europe (Fig. 3e). A positive causal effect (positive β value) of SAM on Sahel Z200 means that enhanced rainfall over the Indian monsoon region are followed one week later by high Z200 anomalies over the Sahel region and the tropical Pacific. In contrast, increased SAM activity leads to negative Z200 anomalies over Central Europe and, over the North
295 Pacific, to a wave train with positive β values southwest of Alaska and negative β value in the eastern North Pacific (Fig. 3e). A positive causal effect (positive β values) of the CGT pattern on Z200 is mostly concentrated in the subtropical North Atlantic, Southern Europe, Central and East Asia, North Pacific and Southeast US (Fig. 3b). Thus, these regions will experience positive Z200 anomalies one week after an enhanced CGT pattern. In contrast, the North Atlantic (around Iceland), South Asia and the Arabian Sea, the eastern North Pacific, the Philippine Sea and Canada will experience negative Z200 anomalies (negative β
300 values, Fig. 3b). In general, β values for OLR causal maps (Fig. 3c and 3f) mirror those for Z200 field in the mid-latitudes, while in the tropics they add the information on tropical convective activity which is not detected by Z200 anomalies. In Fig. 3e negative β values over Southern India indicate that an enhanced CGT pattern leads to lower OLR values and thus increased convective activity over the region. In Fig. 3f the tilted band of convective activity stretches from the Arabian Sea and central India towards the Maritime continent. This convective activity is related to the Boreal summer intraseasonal oscillation (BSISO),
305 as shown in Di Capua et al. (2020a).

Causal maps calculated for the effect of WNPSM and NPH on Z200 and OLR fields are shown in Fig. 3h,i and 3k,l respectively. Causal maps for the effect of NPH on Z200 and OLR field shows a zonally oriented pattern that encircles the northern mid-latitudes (Fig. 3h,i). Positive β values in Fig. 3h indicate that an enhanced NPH pattern (Fig. 3g) leads to positive Z200 anomalies over the North Atlantic (around Iceland), Southeast China, the NPH region and the north-western US. Negative β
310 values over Central and East Asia and over the eastern North Atlantic indicate that an enhanced NPH pattern is followed one week later by negative Z200 anomalies over these regions. The effect of WNPSM is more localized to East Asia and the North Pacific,

where an arch-shaped pattern stretches from the tropical western Pacific, reaching Alaska and the US west coast (Fig. 3k,l). Suppressed convective activity over the South China Sea and the Philippine Sea and increased convective activity over the WNPSM are followed one week later by negative Z200 anomalies over Southeast Asia and Central Asia and positive Z200 anomalies over Northeast Asia and the US west coast.

In general, a two-way causal link between tropical convection and mid-latitude circulation is shown for both MCA modes: the causal effect of SAM and WNPSM reaches the northern mid-latitudes, while the effect of the mid-latitudes pattern CGT and NPH extends to subtropical latitudes. Consistent results are obtained for ERA-L (see Fig. S1) with more significant causal links (likely due to the increased length of the 1979-2020 time series), improving the clarity of the spatial patterns. These patterns also show great resemblance to those produced using the ERA-Interim reanalysis shown in Di Capua et al. (2020a) (see their Figs. 3 and 4).

Causal maps produced with SEAS5-ERA5 MCA modes and SEAS5 OLR and Z200 fields (Fig. 4) show similar spatial patterns to those obtained with ERA-S (Fig. 3). In general, the sign and the geographical position of the causal links detected in SEAS5 are consistent with those found in ERA5. Thus, the main tropical – extratropical intraseasonal causal relationships in boreal summer in the Northern Hemisphere are at least qualitatively well represented in the SEAS5 system. These causal maps also show that the two-way causal pathway between tropical convective activity and extratropical circulation is captured by the seasonal forecasts. Thus, on the one hand we gain confidence in the interpretation of the earlier ERA-Interim and ERA-S/L causal map analysis, which is reproduced by SEAS5, and on the other hand we show that, to a first approximation, seasonal forecasts can reproduce such causal links.

However, the strengths of the causal links detected in SEAS5 are generally weaker than those in ERA-S (or ERA-L). Note that in Figs. 3 and 4 the same colorbar for β values is used. To visualise the difference in strength and/or sign between SEAS5-ERA5 and ERA-S β values, we provide maps of SEAS5-ERA5 β coefficients (β_{SEAS5}) relative to the strength of ERA-S β coefficients (β_{ERA5}) (Fig. 5 and 6, left column). It should be noted here that a comparison between SEAS5-ERA5 β and ERA-S β is only possible when both causal maps show a significant causal link at a specific geographic location. Maps in Fig. 5 and 6 show values of Δ_β defined following Eq. (2):

$$\Delta_\beta = \frac{\beta_{SEAS5}}{\beta_{ERA5}}. \tag{2}$$

When $0 < \Delta_\beta < 1$, SEAS5-ERA5 β values are smaller in magnitude than those shown by ERA-S but the sign of β is correct; this means for example that increased SAM activity leads to positive Z200 anomalies over the Sahel both in SEAS5-ERA5 and ERA-S, however, the β value in ERA-S in larger than that in SEAS5-ERA5 (Fig. 5c), meaning that the causal effect of SAM on this region is weaker in the model than in the reanalysis product. When $\Delta_\beta > 1$, the sign of SEAS5-ERA5 β is correct but its magnitude is greater than in ERA-S; e.g. over the Tibetan Plateau in Fig. 5c. Finally, when $\Delta_\beta < 0$, the sign of SEAS5-ERA5 β is wrong irrespective of the strength of the link; e.g. over the Philippine Sea in Fig. 5a. The colorbar in these plots has been chosen to visually distinguish between these three cases. Δ_β maps for both MCA mode 1 and MCA mode 2 generally give values of Δ_β between 0 and 1, with mean values averaged over each map ranging between 0.4 for the causal effect of CGT on Z200 (Fig. 5a)

345 up to 0.7 for the causal effect of SAM on OLR (Fig. 5g). ERA-S has a much smaller number of significant β, thus the spatial patterns in Fig. 5 and 6 shrink when compared to those shown in Fig. 4. For MCA mode 1, regions with especially poor β representation are the northern high-latitudes and the western North Pacific in Fig. 5a (CGT → Z200|SAM), the tropical belt and in particular the Sahel region in Fig. 5c (SAM → Z200|CGT), Siberia in Fig. 5e (CGT → OLR|SAM) and the Maritime Continent in Fig. 5g (SAM → OLR|CGT).

350 For MCA mode 2, regions with especially poor β representation are the western US in Fig. 6a (NPH → Z200|WNPSM), the western tropical Pacific in Fig. 6c (WNPSM → Z200|NPH), northern US in Fig. 6e (NPH → OLR|WNPSM) and the West Coast of the US and the western tropical Pacific in Fig. 6g (WNPSM → OLR|NPH). This difference may hint at potential biases or misrepresentations of these teleconnections in the GCM.

Together with these β-difference maps we also show the histograms indicating the percentage of grid points (in the domain 355 15°S-75°N, 0°-360°E) in which the SEAS5 β coefficient has the correct sign/magnitude compared to ERA5 both for MCA mode 1 (Fig. 5, right column) and MCA mode 2 (Fig. 6, right column). In general, these plots underline the results shown in the Δ_β maps: the largest amount (>90%) of SEAS5-ERA5 β have a correct sign with respect to ERA-S β but lower strength, with few exceptions, e.g. for CGT → Z200|SAM (Fig. 5b) where the number of β values with wrong sign is ~13%. Thus, in general the SAM – CGT and WNPSM - NPH patterns have the same causal influence on Z200 and OLR fields both in ERA5 and SEAS5 360 dataset. However, although the sign of β values is generally correct, in almost all cases the majority (more than half) of the Δ_β values are found between 0 and 0.5, indicating that SEAS5 strongly underestimates β values.

The weaker strength of the causal links is not dependent on the use of SEAS5-ERA5 MCA modes instead of SEAS5 MCA modes. When SEAS5 MCA modes, which capture the strongest co-variability signal in SEAS5 seasonal forecasts, are used to produce causal maps (see SI, Fig. S2), a similar strength of the causal links is detected.

365 The dependence of causal maps on the time of initialization of SEAS5 has also been analysed. The same plots produced with SEAS5 seasonal forecasts initialized on the 1st of March (rather than 1st of May) are shown in the Supplementary Material (see Figs. S3-S5). Figures 4-6 are very consistent with Figs. S3-S5, thus the initialization date of SEAS5 does not have a qualitative influence on causal maps calculated over the entire 1993-2016 period. The apparent independence on the initialization date is most likely explained by the independence of the β coefficients on the absolute values of a certain variable in a specific year. As 370 explained in the Methods section, the β coefficients represent the relative change in s.d. of one actor given a certain change in the values of its parents (expressed in s.d.).

3.3 Causal effect spread in the SEAS5 ensemble

We perform a sub-sampling experiment to better understand differences in the strength of causal links between ERA-S and SEAS5-ERA5 and evaluate the spread inside the SEAS5 ensemble. We select 1000 samples of 60 years each (randomly selected 375 out of the total 600 model years), and for each sample we provide the corresponding causal map. In this way, the number of years used in each subsampling experiment (60 years) is of the same order of magnitude as those available from ERA-S (24 years) and ERA-L (42 years) but one order of magnitude smaller than that of the entire SEAS5-ERA5 sample (600 years). Reducing the

length of the time series increases the variability and hence lowers the significance of the obtained β values. However, this should not by itself lower the strength of the β values themselves. Thus, a priori, we might expect fewer regions to show a significant β

380 value in a smaller dataset than in a larger one but not a difference in the strength of the β values. Hence, this 1000-ensemble member subsampling experiment allows us to evaluate the distribution of β values around their mean value and to compare it to the ERA-S (or ERA-L) values of reference.

Due to the complexity of the spatial patterns shown in the causal maps in Fig. 3, and to the smaller number of significant grid-points available in ERA-S compared to SEAS5-ERA5 (as visualized in Figs. 3 and 4), calculating spatial correlations is not

385 the most efficient way to compare the two sets of causal maps. A high (or low) spatial correlation may result from strong or weak agreement in different regions, but it would not be possible to discern from which specific area the signal is originating.

To provide a meaningful comparison, we choose four key regions for each MCA mode, and for those regions we analyse the characteristics of the causal effect in detail. We identify these regions based on (i) the prominence of the signal in Fig. 3 and 4 and (ii) the misrepresentation of β values shown in Fig. 5 and 6. By applying these criteria, the chosen geographical regions (shown in

390 Figs. 7 and 8) include: (a) the Sahel region, (b) southeast US, (c) India and (d) northeast Europe for MCA mode 1 (Fig. 7), and (a) Japan, (b) central western US, (c) the Maritime Continent and (d) central eastern Europe for MCA mode 2 (Fig. 8).

The spatial domains used to define these regions are shown in Table 3. For each region and for each sample in the 1000-ensemble member subsampling experiment, the causal effect is spatially averaged (accounting only for significant values, i.e., zero values are discarded as they are not significant) and the absolute value is taken after averaging. In this way, we obtain a

395 distribution of 1000 β values for each region of interest. Note that we choose to consider spatially averaged *absolute β* values as indices for corresponding teleconnectivity strength to focus on the strength of the causal effect rather than the number of significant grid-points or the sign of the connection. Probability density functions (PDFs) for the subsampling experiments of each β index (one for each of the eight regions in Table 3) are calculated by standardizing each β index by its standard deviation (calculated over the 1000 samples) and centring it around its mean. The reference value calculated in ERA-L and ERA-S are

400 shown as a purple and magenta vertical line in each PDF. Both the ERA-L and ERA-S reference values are standardized by dividing by the standard deviation and mean value of the SEAS5-ERA5 subsample distribution. Despite causal maps for ERA-S and ERA-L showing good agreement, those for ERA-L show a greater number of significant grid-points (due to the length of the time series); showing both standardized values in Figs. 7 and 8 helps to further put into perspective the importance of the choice of years and thus provides a more balanced interpretation of the results.

405 For MCA mode 1, which is characterized by the South Asian monsoon (SAM) – circumglobal teleconnection (CGT) pair, we show that the link from the SAM to Sahel Z200 (SAM → Sahel Z200 | CGT) is the one with the largest bias between SEAS5-ERA5 and ERA-L (or ERA-S), with no subsample in our 1000 subsampling experiment being capable of reproducing the ERA-L (or ERA-S) causal link strength (Fig. 7b). By contrast, the causal effect of SAM on OLR over India (SAM → India OLR | CGT) and of CGT on OLR in northeast Europe (CGT → NW-EU OLR | SAM) is falling in the range of the possibilities of the 1000

410 subsamples: although far above the 90th quantile of the PDF, with only a few subsamples exceeding the observed ERA-L value (Fig. 7f,h). The link with the smallest bias between SEAS5-ERA5 and ERA-L (or ERA-S) MCA mode 1 is the one from the CGT

towards south-eastern US Z200 (CGT → SE-US Z200 | SAM), with both ERA-L and ERA-S β values falling between the 50th and the 90th quantile (Fig. 7d). Thus, in this region the β values show a good, or very good, agreement with reanalysis data.

415   Results for MCA mode 2, analysing the North Pacific high (NPH) paired together with the western North Pacific summer monsoon (WNPSM), are shown in Fig. 8. The links from the NPH towards the north-western US Z200 (NPH → NW-US Z200 | WNPSM) (Fig. 8d) and from the WNPSM towards the Maritime Continent OLR (WNPSM → M.Cont. OLR | NPH) (Fig. 8f) are those with the largest bias between SEAS5-ERA5 and ERA-L (or ERA-S) β values, with ERA-L (or ERA-S) β values falling at the upper edge of the PDF (above the 90th quantile). In contrast, the links from the WNPSM towards Japan Z200 (WNPSM → Japan Z200 | NPH) (Fig. 8b) and from the NPH towards central Europe OLR (NPH → CE-EU OLR | WNPSM) (Fig. 8h) fall well 420  in the centre of the distribution, thus showing the smallest bias among the analysed β values. That is, for those cases in which ERA5 values fall in the middle of the distribution, the particular modelled field is more likely to have a low bias.

  Again, we test for the dependence on the initialization date of the SEAS5 dataset. Figures S6 and S7 in the Supplementary Material show the same figures as Figs. 7 and 8 produced using SEAS5 initialized on the 1st of March. The results obtained are qualitatively and quantitatively very similar, with β values for 7 out of the 8 analysed regions falling above (or below) the 90th 425  quantile consistently between SEAS5 init05 and SEAS5 init03. Hence, the initialisation date does not affect the spread of the subsampling experiment.

3.4 The effect of the ERA5 – SEAS5 bias on tropical – extratropical causal links

  Next, we assess of the effect of biases between ERA5 and SEAS5 in their JJAS climatology for SST, U200 and OLR fields on the calculated causal maps. The climatology and bias (SEAS5 minus ERA-S) for each of these three fields are shown in Fig. 9. 430  The overall spatial patterns and magnitude of the values shown in Figs. 9a,f,k (ERA-S), 9b,g,l (SEAS5 init. 1st of March) and 9d,i,n (SEAS5 init. 1st of May) show a good agreement between ERA5 and SEAS5 data. However, a closer look at the differences between ERA-S and SEAS5 JJAS climatologies (Figs. 9 right column) reveals a warm bias of about 1°C (up to 2°C) in tropical SST temperatures in the SEAS5 model in the western Indian Ocean, at the eastern shores of the Maritime Continent and the west of the South American coast, and east of the African continent (Fig. 9c,e). Moreover, the warm bias in the Pacific SST north of 435  the equator may also likely affect the representation of extratropical teleconnections, particularly in the North Pacific and North America sector (Fig. 9c,e). , Anomalous positive U200 of about ~5 m/s are observed (Fig. 9h,j) in the same region where biases in SST fields are detected. These differences may explain the differences in the magnitude of β values shown in Figs. 5 and 6. The U200 mean-state biases are also large in the mid-latitudes, suggesting a systematic northward shift of the westerly jet all across Eurasia. Since the mid-latitude jet is the main waveguide for both the CGT and the Silk Road patterns, this bias might also play a 440  role in miscommunicating the teleconnection. Notably, while the warm SST bias is larger in SEAS5 init. 1st of March, the U200 bias in the mid-latitude jet over Eurasia is larger in SEAS5 init. 1st of May.

  Biases in convective activity in the tropical Indian Ocean may provide another source for the misrepresentation of causal link in SEAS5 (Fig. 9m,o). Despite the overall good representation of the OLR spatial patterns in SEAS5 when compared to ERA-S, the convective activity in the tropical Indian Ocean shows a positive bias around the Equator and a negative bias over South Asia

445 (Fig. 9m,o). This bias may explain the too weak causal links over North Africa: Previous work by Rodwell and Hoskins (1996) has shown that the heat source provided by the convective activity in the Indian Ocean/Bay of Bengal region generates Rossby waves that reach the Sahara Desert. This phenomenon is also known as the monsoon – desert mechanism. However, the latitudinal position of the heat source is critical: a heat source located in the south (10°N) does not act as a source of Rossby waves capable of reaching the Sahara Desert, while a heat source located around 25°N does.

450 To further test this hypothesis, we create an index representing the difference between convective activity (represented by negative OLR values) shifted northward from the equatorial Indian Ocean towards the Indian peninsula (10°-20°N, 60°-90°E; northward box in Fig. 10 a,b,f,g) and convective activity shifted southward towards the equator (0°-10°N, 60°-90°E; southward box in Fig. 10 a,b,f,g). Next, we select among the 600 years available in SEAS5 only those for which this OLR India index exceeds +1 s.d., corresponding to enhanced convective activity over India (Fig. 10a,f) or falls below -1 s.d., corresponding to

455 enhanced convective activity over the equatorial Indian Ocean (Fig. 10b,g).

To assess whether the latitudinal position of the convective activity in the Indian Ocean affects the causal effect of the SAM on Sahel Z200, we calculate causal maps for the effect of the SAM on Z200 fields for SEAS5 initialized the 1st of March (Fig. 10c,d) and for SEAS5 initialized on the 1st of May (Fig.10h,i). Qualitatively the spatial patterns and strengths of the links shown in the causal maps obtained with years with negative OLR India index are similar to those obtained with years with positive OLR

460 India index, and both are similar to those shown in Fig. 4 for the entire SEAS5 dataset. To analyse changes in the strength of the β values we calculate the difference between causal maps calculated with OLR India negative and OLR India positive both for SEAS5 initialized on the 1st of March and on the 1st of May (shown in Fig. 10c,d and 10h,i respectively). For each grid point, the difference Δ_β between the β value for the OLR India negative (β_{OLRn}) and the β value for the OLR India positive (β_{OLRp}) is calculated and then divided by β_{OLRn}, following Eq. (3):

465 $$\Delta_\beta = \frac{\beta_{OLRn} - \beta_{OLRp}}{\beta_{OLRn}} x 100\% \qquad (3)$$

Hence, Δ_β is expressed as a percentage, where a zero value represents perfect agreement between β_{OLRn} and β_{OLRp}, a positive value of, e.g., 50% means that β_{OLRn} is 50% larger than β_{OLRp} and vice versa. The Δ_β maps in Fig. 10e,j show some agreement in the sign and magnitude of Δ_β over Southeast Asia and the tropical central-eastern Pacific, with larger β values for causal maps obtained by OLR India negative years (~20-60%). Over our region of interest, i.e. the Sahel region, results are mostly

470 inconclusive: causal maps obtained for SEAS5 initialized on the 1st of May show an increase (~20-40%) of β values during OLR India negative years, but those obtained with SEAS5 data initialized on the 1st of March show the opposite tendency. One could argue that in Fig. 9 the seasonal bias between ERA5 and SEAS5 in the tropical Indian Ocean/Indian peninsula and Pacific Ocean is smaller for SEAS5 data initialized on the 1st of May (see Fig. 9c,e). However, this is not a strong enough argument to favour SEAS5 initialized on the 1st of May over those initialized on the 1st of March.

475

3.5 The effect of ENSO on tropical – extratropical causal links

Finally, we analyse the effect of the ENSO state on each causal map shown in Fig. 4 to determine whether SST anomalies in the tropical Pacific can affect the strength of the causal links. Composites for SST, U200 and OLR anomalies during Nino3.4 positive and Nino3.4 negative years are shown in the Supplementary Material in Fig. S8. If a dependence is found, this may also
480 explain the differences between SEAS5 and ERA5 causal maps. Moreover, SEAS5 performance may be better (or worse) for strong (or weak) ENSO conditions. Hence, for each of the 600 model years, we calculate the Niño3.4 index as SST seasonal (JJAS) averaged over the region 5°S-5°N, 190°-240°E. Then, those years that exceed the +0.5°C (-0.5°C) threshold are defined as El Niño (La Niña) years. Next, causal maps for the effect of MCA mode 1 and 2 on Z200 fields are recalculated using only the 102 Niño3.4 positive and 142 Niño3.4 negative years separately. The results for the causal effect of MCA patterns on Z200 fields
485 for both Niño3.4 positive (left column) and Niño3.4 negative years (middle column) separately and for the difference Niño3.4 positive minus Niño3.4 negative (right column) and for different initialization dates (1st of March and 1st of May) are shown in Fig. 11 for MCA mode 1 and Fig. 12 for MCA mode 2.

Finally, we analyse the effect of opposite ENSO states on Z200 causal maps shown in Fig. 11 and 12 is marginal. Despite some differences in the magnitude and shape of the patterns can be noticed, e.g., β values over the Sahel in Fig. 11a (El Niño years) are ~20-40%
490 higher than in Fig. 11b (La Niña years), in general these minor differences are of the same order of magnitude as those shown between different initialization dates. Note that Δ_β values shown in the right column of Figs. 11 and 12 are calculated following Eq. (3) but for $\beta_{Ni\tilde{n}o}$ and $\beta_{Ni\tilde{n}a}$ instead. Moreover, the differences shown in SEAS5 data initialized on the 1st of March do not always match those obtained from SEAS5 data initialized on the 1st of May.

Nevertheless, we can identify a handful of regions for which ENSO impacts the β values consistently between the two
495 different initialization dates. These regions are: (1) western central Africa and the tropical central Pacific for the link from the SAM to the Z200 (compare Fig. 11c and 11f) and (2) Southeast Asia with a negative effect of the CGT pattern on Z200 fields (compare Fig. 11i and 11l) and a positive effect of the WNPSM on Z200 (compare Fig. 12c and 12f). The same causal maps for OLR fields are shown in Fig. S9 and S10 in the Supplementary Material and β values for OLR are consistent with the corresponding regions (1) to (3) identified above. Notably, in the OLR fields, the β values decrease consistently between the two
500 initialization dates over north-eastern India and Pakistan (see Figs. S9c and S9f in the Supplementary Material), showing an increased effect of the SAM over this region during La Niña years. Similarly, decreased OLR activity over north-eastern Europe is also shown to be consistent between the two initialization dates (Figs. S10i and S10l in the Supplementary Material).

This analysis shows that despite different ENSO phases affecting SST anomalies and other atmospheric fields (see previous paragraph and Fig. 9), these differences have a minor effect in altering the β values. However, these results do not imply that
505 ENSO does not affect wind or geopotential heights fields (see Fig. S8 in the Supplementary Material), but only that these changes have little effect on the *relative* influence of each MCA pattern on Z200 and OLR fields. In other words, if the convective activity in the SAM region is weaker due to the ENSO phase, its *total* effect on the Sahel region will be weaker as well but the relative effect represented by the β values will be almost the same.

We also calculate causal maps for sets of positive and negative Niño3.4 years combined together against neutral (all other)
510 years, in order to detect whether the presence of high amplitude tropical forcing alters the presence of causal links to the
extratropics. These results are shown in the Supplementary Material in Figs. S11 and S12. Similarly to what is shown in Figs. 11
and 12, the overall spatial patterns in the causal maps remain qualitatively unchanged. However, some quantitative changes
consistent between the two initialization dates can be identified, namely: for MCA mode 1, the effect of the CGT pattern on (1)
tropical western Pacific (Figs. S11c and S11f in the Supplementary Material), which is stronger during neutral years (likely due to
515 the absence of ENSO teleconnections) and (2) eastern Russia, which is stronger during active ENSO phases (Figs. S11i and S11l
in the Supplementary Material). For the MCA mode 2, the effect of the NPH on Z200 in eastern Russia (Figs. S12c and S12f in
the Supplementary Material), which is stronger during ENSO neutral phases and the effect of the WNPSM on Z200 fields over the
western tropical Pacific, which seems to be stronger during ENSO neutral phases (Figs. S12i and S12l in the Supplementary
Material).

4 Discussion

In this work, we provide a process-based validation, built on causal discovery, of the representation of boreal summer
tropical – extratropical intraseasonal teleconnections in the Northern Hemisphere in SEAS5 seasonal forecasts by ECMWF.
We have analysed the two first modes of covariability identified using maximum covariance analysis (MCA) between
weekly geopotential heights (Z200) and convective activity (OLR) fields in reanalysis data from the ERA5 dataset for the
525 period 1993-2016. The first MCA mode shows the South Asian monsoon (SAM) paired with the circumglobal
teleconnection pattern (CGT). The second MCA mode shows the western North Pacific summer monsoon (WNPSM) paired
with the North Pacific High (NPH). Causal maps showing the causal effect of these four patterns on Northern Hemisphere
Z200 and OLR fields at 1-week lead for periods 1979-2020 (Fig. S1, see Supplementary Material) and 1993-2016 (Fig. 3)
are largely consistent with results obtained with ERA-Interim data for the period 1979-2018 (Di Capua et al., 2020).

530 Here, we focus on assessing the ability of SEAS5 seasonal forecasts in reproducing those results. To achieve this goal
and provide a meaningful comparison we project the first two MCA modes calculated in ERA5 onto SEAS5 data and we
calculate the corresponding causal maps (Fig. 4). In general, causal maps obtained with SEAS5 correctly reproduce the sign
and the spatial patterns of ERA5 causal maps, though with weaker magnitudes (Figs. 5 and 6). Thus, spatial patterns shown
in SEAS5 seasonal forecasts causal maps are validated by those extracted from ERA5: while the SEAS5 forecasting system
535 can reproduce the patterns seen in ERA5 reanalysis, we gain confidence that observed causal maps represent actual physical
mechanisms.

We analyse the negative bias found in SEAS5, i.e., a general underestimation of the causal effect, which may arise for
different reasons: (i) ERA5 causal maps are subject to multidecadal variability or affected by the limited time span
considered or (ii) the SEAS5 forecasting system is missing or misrepresenting a key mechanism for a correct representation
540 of the strength of the analysed causal links. To test the first hypothesis, we perform a subsampling experiment, providing a

thousand causal maps obtained using 60 randomly selected years out of the in total 600 available. Random year selection may overestimate the range for potential β coefficients, thus further emphasizing the underestimation of the strength of these tropical – extratropical teleconnections in SEAS5. We identify a few key regions for which we compared the observed ERA5 causal link strength with the range of SEAS5 values obtained from the subsampling (Figs. 7 and 8). Through this

545   experiment we show that, although ERA5 values are generally higher than the mean of the SEAS5 distribution, approximately half of them fall below the 90th quantile of the SEAS5 distribution (1 out of 4 for MCA 1 and 2 out of 4 for MCA2). Thus, SEAS5 has difficulty to generate intense values of the teleconnection strength, which is generally underestimated. Here, we can identify those regions for which SEAS5 can reasonably represent observed causal links and distinguish from those which are underestimated (when the ERA5 reference values exceed the SEAS5 90th quantile).

550   We calculate the biases between ERA5 and SEAS5 for SST and U200 JJAS climatologies and show that biases are present both in tropical and extratropical SST and in the mid-latitude jet over Eurasia. The U200 mean-state biases in the mid-latitudes suggest a systematic northward shift of the subtropical westerly jet all across Eurasia, possibly affecting the waveguide for the CGT pattern and the Silk Road pattern. Future work may employ nudging experiments to test the sensitivity to this northward shift, despite previous work had shown little effect of this phenomenon on the CGT pattern

555   (Beverley et al., 2019). Changes in the initialization dates of the SEAS5 simulations (here 1st of March and 1st of May) also impact the magnitude of both SST and U200 biases, with reduced SST biases in SEAS5 data initialized on the 1st of May and seemingly increased biases for the Eurasian jet when compared to SEAS5 data initialized on the 1st of March (see Fig. 9).

To explain the negative bias in the first MCA mode, we can refer to the work by Beverley et al. (2019) showing that the CGT pattern is too weak in SEAS4 (ECMWF's previous operational seasonal forecasting system). Climate models struggle

560   to reproduce the climatology of SAM rainfall patterns, both in magnitude and spatial pattern (Menon et al., 2013). In SEAS4, the SAM precipitation is lower than observed and this may explain a too weak CGT pattern in the model (Beverley et al., 2019). If the forcing (SAM) is too weak, the response (CGT) will be too weak even if the causal link strength would be correct. However, our results show that also the causal link strength is too weak, thus even if the forcing in the model would be correct, the response would still be too weak. While in boreal summer the CGT pattern arises even without the heat

565   source provided by SAM (Ding et al., 2011) as it represents a preferred mode of variability of boreal summer circulation that can be ignited by different forcings (Kornhuber et al., 2020; Teng and Branstator, 2019). Recent work has shown that there is a positive causal link from the SAM to the CGT (Di Capua et al., 2020b, a). Hence, the underestimation of the strength of the SAM convective activity and rainfall shown in SEAS5 seasonal forecasts over the Indian peninsula and the Bay of Bengal (see Fig. 9 or Chevuturi et al. (2021)), i.e. a reduced strength of the SAM pattern in SEAS5, may affect the strength of the

570   CGT pattern (Beverley et al., 2019) and explain the negative bias between SEAS5 and ERA5 when the causal effect of these two patterns on other surface variables is analysed.

Despite consistent underestimation of causal link strength in certain regions, these results imply that the ability of the SEAS5 forecast system to reproduce observed causal patterns for boreal summer intraseasonal variability in the Northern Hemisphere is quite satisfying. Although this analysis does not rely on nor imply a skilful forecast, the causal effect of

575   tropical and mid-latitude patterns on circulation and convection in the Northern Hemisphere in SEAS5 seasonal forecasts is qualitatively well comparable with that shown in ERA5 reanalyses. Here we have shown for which regions the agreement between SEAS5 and ERA5 is good (or for which ERA5 values at least fall in the range of values shown in the 1000 subsample experiments) and those for which no subsample of SEAS5 can reproduce values comparable to ERA5. The region with the strongest bias, which cannot be reproduced in SEAS5, is the Sahel region. This may be explained by the southward

580   shift of OLR and rainfall activity towards the equatorial Indian Ocean in SEAS5 (Rodwell and Hoskins, 1996), disrupting the Rossby-forced teleconnection pattern. However, our analysis conducted to determine the importance of the latitudinal position of convective activity in the Indian Ocean basin is inconclusive (Fig. 10).

We further analyse the effect of El Niño – Southern Oscillation (ENSO) on these boreal summer tropical – extratropical links in SEAS5 and we find that, in general, different ENSO phases do not affect the spatial patterns and sign of the causal

585   links substantially (see Figs. 11 and 12). However, depending on the specific region, some effect on the strength of the causal links is detected (see Figs. 11 and 12, right columns). This effect is generally relatively weak (~20-40% of the total strength of the link). These findings are generally in agreement with Di Capua et al. (2020), who showed that the spatial patterns and the sign of the causal links were mainly unaffected by ENSO; however, they noticed a regional dependence of the strength of the causal links on ENSO (see their Fig. 6). Although the effect of ENSO on the strength of the causal links

590   seems to be marginal, we highlight those regions where a change in the strength is identified consistently both for 1st of March and 1st of May initialization dates. These regions are mainly located in the tropical belt, in particular western tropical Pacific, western Africa and Southeast Asia. Similar results are found for the effect of neutral versus ENSO active years (see Figs. S10 and S11 in the Supplementary Material). In general, the spatial pattern and sign of the causal links do not change, however the effect of the identified MCA patterns seems to increase in the tropical Pacific during neutral years, likely due to

595   the absence of a strong ENSO signal. These results align with the finding of Goddard and Dilley (2005) who showed that the active phase of ENSO does not affect the numbers of extremes but their overall predictability. Nevertheless, as discussed above, a change in the forcing (e.g. stronger convective activity induced by ENSO) would still result in a stronger effect even if the (relative) causal effect does not change.

This information becomes even more relevant in the context of climate change. If the SEAS5 forecasting system

600   behaves consistently in EC-Earth (Döscher et al., 2022) (the Earth system model built by the ECMWF which shares the same atmosphere model as SEAS5) as seen in this analysis, we can have some confidence that these tropical – extratropical teleconnections are fairly well represented. Thus, future projections under global warming scenarios may be fairly reliable. The analysed regions (from the South Asian monsoon to the North American and Eurasian continents) are prone to suffering the effect of the ongoing anthropogenic climate change (Pfleiderer et al., 2018; Mann et al., 2018; Coumou et al., 2018,

605   2017; Huntingford et al., 2019; Turner and Annamalai, 2012). Therefore, it is critical to evaluate the model ability at reproducing key seasonal modes of variability and in doing so, identify key targets for model development and motivate the improvement of parameterisation schemes. Ultimately, this could lead to increased reliability of seasonal and sub-seasonal forecasts; helping in improving warning systems and taking sensible early action to protect economy and society in most

vulnerable regions, especially for summer, when the effect of climate change is felt the most (Christidis et al., 2014; Teng
610   and Branstator, 2019; Kornhuber et al., 2020; Coumou et al., 2015, 2017).

Finally, this process-based validation analysis represents a step towards the implementation of hybrid forecasts that unite statistical with dynamical models to increase the skill of seasonal and sub-seasonal weather forecasts (Schepen et al., 2012). By identifying the regions where a certain pattern exerts a significant influence and/or deriving information on which regions have a bias in the model, we provide useful information on how to correct the model representation of these
615   mechanisms and work towards targeted forecasts. Moreover, general circulation models show higher skill at forecasting tropical atmospheric dynamics than in the mid- or high-latitudes (Shukla, 1998; Chen et al., 2015), thus, knowing which regions in the Northern hemisphere are more affected by tropical precipitation (as shown in Fig. 4) provides valuable information to improve seasonal forecast skill.

5 Conclusions

620   In summary, this analysis has shown that ECMWF's seasonal forecasts have good ability at reproducing the causal effect patterns of the two main modes of covariability between tropical convection and mid-latitude circulation in boreal summer on convection and mid-latitude circulation in the Northern Hemisphere. Despite a general underestimation of the causal strength, our subsampling experiment shows that in most of the analysed regions, this negative bias is actually contained in the spread of the SEAS5 seasonal forecasts. Thus, our confidence in both the ability of the PCMCI causal discovery tool to identify meaningful
625   patterns in observation and the ability of SEAS5 forecasting system to correctly represent those causal links is increased. The effect of different ENSO phases (or ENSO versus neutral years) on tropical-extratropical links seems to be marginal, although biases in the SEAS5 model (e.g., reduced convective activity in the SAM region) may explain this discrepancy with observations. Further work is needed to confirm these results and the implications that may be implied. Finally, the causal links represented in these causal maps represent a starting point to produce a new family of hybrid statistical model-based forecasts. In conclusion, this
630   analysis has shown the usefulness of causal discovery algorithms as a tool to provide process-based validation of general circulation model and has increased the knowledge available on the effect of tropical – extratropical teleconnections in boreal summer in the Northern Hemisphere.

Acknowledgments

635   This work has been financially supported by the German Federal Ministry for Education and Research of Germany (BMBF) via the JPI Climate/JPI Oceans project ROADMAP (grant no. 01LP2002B) (G.D.C. and R.V.D.). AGT's contribution to the work was supported by the National Centre for Atmospheric Science.

Data availability. The data used in this article can be accessed by contacting the corresponding author.

640

Author contributions. All authors contributed to the design of the analysis. GDC performed the analysis and wrote the first draft of the paper. All authors contributed to the interpretation of the results and to the writing of the paper.

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

645

650 **References**

Barnston, A. G., Glantz, M. H., and He, Y.: Predictive Skill of Statistical and Dynamical Climate Models in SST Forecasts during the 1997-98 El Nino Episode and the 1998 La Nina Onset, Bull. Am. Meteorol. Soc., 80, 217–244, 1999.

Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, https://doi.org/10.1038/nature14956, 2 September 2015.

655 Benjamini, Y. and Hochberg, Y.: Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing, J. R. Stat. Soc. Ser. B, 57, 289–300, https://doi.org/10.2307/2346101, 1995.

Beverley, J. D., Woolnough, S. J., Baker, L. H., Johnson, S. J., and Weisheimer, A.: The northern hemisphere circumglobal teleconnection in a seasonal forecast model and its relationship to European summer forecast skill, Clim. Dyn., 52, 3759–3771, https://doi.org/10.1007/s00382-018-4371-4, 2019.

660 Briegel, L. M. and Frank, W. M.: Large-scale influences on tropical cyclogenesis in the western North Pacific, Mon. Weather Rev., 125, 1397–1413, https://doi.org/10.1175/1520-0493(1997)125<1397:LSIOTC>2.0.CO;2, 1997.

Di Capua, G., Kretschmer, M., Runge, J., Alessandri, A., Donner, R. V., van den Hurk, B., Vellore, R., Krishnan, R., and Coumou, D.: Long-Lead Statistical Forecasts of the Indian Summer Monsoon Rainfall Based on Causal Precursors, Weather Forecast., 34, 1377–1394, https://doi.org/10.1175/waf-d-19-0002.1, 2019.

665 Di Capua, G., Runge, J., Donner, R. V., van den Hurk, B., Turner, A. G., Vellore, R., Krishnan, R., and Coumou, D.: Dominant patterns of interaction between the tropics and mid-latitudes in boreal summer: causal relationships and the role of timescales, Weather Clim. Dyn., 1, 519–539, https://doi.org/10.5194/wcd-1-519-2020, 2020a.

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 670 approach, Earth Syst. Dyn., 11, 17–34, https://doi.org/10.5194/esd-11-17-2020, 2020b.

Challinor, A. J., Slingo, J. M., Wheeler, T. R., Craufurd, P. Q., and Grimes, D. I. F.: Toward a Combined Seasonal Weather and Crop Productivity Forecasting System: Determination of the Working Spatial Scale, 2003.

Charney, J. G. and Shukla, J.: Predictability of monsoons, in: Monsoon dynamics, 99–109, 1981.

Chen, M., Kumar, A., and Wang, W.: A study of the predictability of sea surface temperature over the tropics, Clim. Dyn., 675 44, 1767–1776, https://doi.org/10.1007/s00382-014-2187-4, 2015.

Chevuturi, A., Turner, A. G., Johnson, S., Weisheimer, A., Shonk, J. K. P., Stockdale, T. N., and Senan, R.: Forecast skill of the Indian monsoon and its onset in the ECMWF seasonal forecasting system 5 (SEAS5), Clim. Dyn., 56, 2941–2957, https://doi.org/10.1007/s00382-020-05624-5, 2021.

Christidis, N., Jones, G. S., and Stott, P. a.: Dramatically increasing chance of extremely hot summers since the 2003 680 European heatwave, Nat. Clim. Chang., 5, 46–50, https://doi.org/10.1038/nclimate2468, 2014.

Coumou, D., Lehmann, J., and Beckmann, J.: The weakening summer circulation in the Northern Hemisphere mid-latitudes, Science (80-.)., 348, 324–327, https://doi.org/10.1126/science.1261768, 2015.

Coumou, D., Kornhuber, K., Lehmann, J., and Petoukhov, V.: Weakened Flow Persistent Circulation, and Prolonged Weather Extremes in Boreal Summer, in: Climate Etremes: Patterns and Mechanisms, Geoph Monogr 226, 61–73, 2017.

685   Coumou, D., Di Capua, G., Vavrus, S., Wang, L., and Wang, S.: The influence of Arctic amplification on mid-latitude summer circulation, Nat. Commun., 9, 2959, https://doi.org/10.1038/s41467-018-05256-8, 2018.

Ding, Q. and Wang, B.: Circumglobal teleconnection in the Northern Hemisphere summer, J. Clim., 18, 3483–3505, https://doi.org/10.1175/JCLI3473.1, 2005.

Ding, Q. and Wang, B.: Intraseasonal teleconnection between the summer Eurasian wave train and the Indian Monsoon, J.

690   Clim., 20, 3751–3767, https://doi.org/10.1175/JCLI4221.1, 2007.

Ding, Q., Wang, B., Wallace, J. M., and Branstator, G.: Tropical-extratropical teleconnections in boreal summer: Observed interannual variability, J. Clim., 24, 1878–1896, https://doi.org/10.1175/2011JCLI3621.1, 2011.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L. P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria,

695   P., Fladrich, U., Fuentes-Franco, R., Gröger, M., Hardenberg, J. V., Hieronymus, J., Karami, M. P., Keskinen, J. P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., Van Noije, T., Nolan, P., O'donnell, D., Ollinaho, P., Van Den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and

700   Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geosci. Model Dev., 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., Gier, B. K., Hall, A. D., Hoffman, F. M., Hurtt, G. C., Jahn, A., Jones, C. D., Klein, S. A., Krasting, J. P., Kwiatkowski, L., Lorenz, R., Maloney, E., Meehl, G. A., Pendergrass, A. G., Pincus, R., Ruane, A. C., Russell, J. L., Sanderson, B. M., Santer, B. D., Sherwood, S. C.,

705   Simpson, I. R., Stouffer, R. J., and Williamson, M. S.: Taking climate model evaluation to the next level, Nat. Clim. Chang., 9, 102–110, https://doi.org/10.1038/s41558-018-0355-y, 2019.

Gadgil, S., Rajeevan, M., and Nanjundiah, R.: Monsoon prediction - Why yet another failure?, Curr. Sci., 88, 1389–1400, 2005.

Goddard, L. and Dilley, M.: El Niño: Catastrophe or opportunity, J. Clim., 18, 651–665, https://doi.org/10.1175/JCLI-

710   3277.1, 2005.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Catherine, A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., Von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J. J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and Von Storch, J. S.: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, Geosci. Model Dev., 9, 4185–4208,

715   https://doi.org/10.5194/gmd-9-4185-2016, 2016.

Hagger, V., Dwyer, J., Shoo, L., and Wilson, K.: Use of seasonal forecasting to manage weather risk in ecological

restoration, 2018.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R.,
Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita,
720 M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A.,
Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G.,
de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Q. J. R. Meteorol.
Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Holton, J. R.: An introduction to dynamic meteorology, Fourth Edition., edited by: Cynar, F., Elsevier Academic Press,
725 1973.

Horak, J., Hofer, M., Gutmann, E., Gohm, A., and Rotach, M. W.: A process-based evaluation of the Intermediate
Complexity Atmospheric Research Model (ICAR) 1.0.1, Geosci. Model Dev., 14, 1657–1680, https://doi.org/10.5194/gmd-
14-1657-2021, 2021.

Huntingford, C., Mitchell, D., Kornhuber, K., Coumou, D., Osprey, S., and Allen, M.: Assessing changes in risk of amplified
730 planetary waves in a warming world, Atmos. Sci. Lett., 20, 1–11, https://doi.org/10.1002/asl.929, 2019.

Jain, S., Scaife, A. A., and Mitra, A. K.: Skill of Indian summer monsoon rainfall prediction in multiple seasonal prediction
systems, Clim. Dyn., https://doi.org/10.1007/s00382-018-4449-z, 2018.

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D.,
Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., and Monge-Sanz, B. M.: SEAS5: The new ECMWF
735 seasonal forecast system, Geosci. Model Dev., 12, 1087–1117, https://doi.org/10.5194/gmd-12-1087-2019, 2019.

Kornhuber, K., Coumou, D., Vogel, E., Lesk, C., Donges, J. F., Lehmann, J., and Horton, R. M.: Amplified Rossby waves
enhance risk of concurrent heatwaves in major breadbasket regions, Nat. Clim. Chang., 20, 48–53,
https://doi.org/10.1038/s41558-019-0637-z, 2020.

Kretschmer, M., Coumou, D., Donges, J. F., and Runge, J.: Using Causal Effect Networks to analyze different Arctic drivers
740 of mid-latitude winter circulation, J. Clim., 29, 4069–4081, https://doi.org/10.1175/JCLI-D-15-0654.1, 2016.

Kretschmer, M., Runge, J., and Coumou, D.: Early prediction of extreme stratospheric polar vortex states based on causal
precursors, Geophys. Res. Lett., 44, 8592–8600, https://doi.org/10.1002/2017GL074696, 2017.

Kretschmer, M., Cohen, J., Matthias, V., Runge, J., and Coumou, D.: The different stratospheric influence on cold-extremes
in Eurasia and North America, npj Clim. Atmos. Sci., 1, 1–10, https://doi.org/10.1038/s41612-018-0054-4, 2018.

745 Kumar, A.: Statistical Models for Long-range Forecasting of Southwest Monsoon Rainfall over India Using Step Wise
Regression and Neural Network, Atmos. Clim. Sci., 02, 322–336, https://doi.org/10.4236/acs.2012.23029, 2012.

Lazo, J. K., Morss, R. E., and Demuth, J. L.: 300 billion served, Bull. Am. Meteorol. Soc., 90, 785–798,
https://doi.org/10.1175/2008BAMS2604.1, 2009.

Lehmann, J., Kretschmer, M., Schauberger, B., and Wechsung, F.: Potential for early forecast of Moroccan wheat yields
750 based on climatic drivers, Geophys. Res. Lett., 1–10, https://doi.org/10.1029/2020gl087516, 2020.

Mann, M. E., Rahmstorf, S., Kornhuber, K., Steinman, B. A., Miller, S. K., Petri, S., and Coumou, D.: Projected changes in persistent extreme summer weather events: The role of quasi-resonant amplification, Sci. Adv., 4, 1–10, https://doi.org/10.1126/sciadv.aat3272, 2018.

Menon, A., Levermann, A., Schewe, J., Lehmann, J., and Frieler, K.: Consistent increase in Indian monsoon rainfall and its variability across CMIP-5 models, Earth Syst. Dyn., 4, 287–300, https://doi.org/10.5194/esd-4-287-2013, 2013.

Meza, F. J., Hansen, J. W., and Osgood, D.: Economic value of seasonal climate forecasts for agriculture: Review of ex-ante assessments and recommendations for future research, https://doi.org/10.1175/2007JAMC1540.1, 2008.

Nowack, P., Runge, J., Eyring, V., and Haigh, J. D.: Causal networks for climate model evaluation and constrained projections, Nat. Commun., 11, 1–11, https://doi.org/10.1038/s41467-020-15195-y, 2020.

O'Reilly, C. H., Woollings, T., Zanna, L., and Weisheimer, A.: The impact of tropical precipitation on summertime euro-Atlantic circulation via a circumglobal wave train, J. Clim., 31, 6481–6504, https://doi.org/10.1175/JCLI-D-17-0451.1, 2018.

Palmer, T.: The primacy of doubt: Evolution of numerical weather prediction from determinism to probability, https://doi.org/10.1002/2017MS000999, 1 June 2017.

Palmer, T. N. and Anderson, D. L. T.: The prospects for seasonal forecasting-A review paper, Q. 1. R. Meteorol. SOC, 755–793 pp., 1994.

Pfleiderer, P., Schleussner, C., and Coumou, D.: Boreal summer weather becomes more persistent in a warmer world, Nat. Clim. Chang., https://doi.org/10.1038/s41558-019-0555-0, 2018.

Pfleiderer, P., Schleussner, C.-F., Geiger, T., and Kretschmer, M.: Robust predictors for seasonal Atlantic hurricane activity identified with causal effect networks, Weather Clim. Dyn., 1, 313–324, https://doi.org/10.5194/wcd-1-313-2020, 2020.

Roberts, C. D., Vitart, F., and Balmaseda, M. A.: Hemispheric Impact of North Atlantic SSTs in Subseasonal Forecasts, https://doi.org/10.1029/2020GL091446, 28 February 2021.

Rodwell, M. J. and Hoskins, B.: Monsoons and the dynamics of deserts, Q. J. R. Meteorol. Soc., 122, 1385–1404, 1996.

Runge, J.: Causal network reconstruction from time series: From theoretical assumptions to practical estimation, 28, 075310, https://doi.org/10.1063/1.5025050, 2018.

Runge, J., Petoukhov, V., and Kurths, J.: Quantifying the strength and delay of climatic interactions: The ambiguities of cross correlation and a novel measure based on graphical models, J. Clim., 27, 720–739, https://doi.org/10.1175/JCLI-D-13-00159.1, 2014.

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D.: Detecting causal associations in large nonlinear time series datasets, Sci. Adv., 5, eaau4996, 2019.

Scaife, A. A., Ferranti, L., Alves, O., Athanasiadis, P., Baehr, J., Dequé, M., Dippe, T., Dunstone, N., Fereday, D., Gudgel, R. G., Greatbatch, R. J., Hermanson, L., Imada, Y., Jain, S., Kumar, A., MacLachlan, C., Merryfield, W., Müller, W. A., Ren, H. L., Smith, D., Takaya, Y., Vecchi, G., and Yang, X.: Tropical rainfall predictions from multiple seasonal forecast systems, Int. J. Climatol., 1–15, https://doi.org/10.1002/joc.5855, 2018.

785   Schepen, A., Wang, Q. J., and Robertson, D. E.: Combining the strengths of statistical and dynamical modeling approaches for forecasting Australian seasonal rainfall, J. Geophys. Res. Atmos., 117, https://doi.org/10.1029/2012JD018011, 2012.

Seo, K. H., Wang, W., Gottschalck, J., Zhang, Q., Schemm, J. K. E., Higgins, W. R., and Kumar, A.: Evaluation of MJO forecast skill from several statistical and dynamical forecast models, J. Clim., 22, 2372–2388, https://doi.org/10.1175/2008JCLI2421.1, 2009.

790   Shukla, J.: Predictability in the Midst of Chaos: A Scientific Basis for Climate Forecasting, Science (80-.)., 282, 728–731, 1998.

Shukla, J., Anderson, J., Baumhefner, / D, Brankovic, # C, Chang, @ Y, Kalnay, E., Marx, L., Palmer, T., Paolino, @ D, Ploshay, J., Schubert, / S, Straus, & D., Suarez, M., and Tribbia, J.: Dynamical Seasonal Prediction, Bull. Am. Meteorol. Soc., 82, 2593–2606, 2000.

795   Spirtes, P., Glymour, C., and Scheines, R.: Causation, prediction, and search, Boston: The MIT Press., 2000.

Stephan, C. C., Klingaman, N. P., and Turner, A. G.: A mechanism for the interdecadal variability of the Silk Road Pattern, J. Clim., 32, 717–736, https://doi.org/10.1175/JCLI-D-18-0405.1, 2019.

Teng, H. and Branstator, G.: Amplification of Waveguide Teleconnections in the Boreal Summer, Curr. Clim. Chang. Reports, 5, 421–432, https://doi.org/10.1007/s40641-019-00150-x, 2019.

800   Tietsche, S., Balmaseda, M., Zuo, H., Roberts, C., Mayer, M., and Ferranti, L.: The importance of North Atlantic Ocean transports for seasonal forecasts, Clim. Dyn., 55, 1995–2011, https://doi.org/10.1007/s00382-020-05364-6, 2020.

Tsonis, A. A. and Eisner, J. B.: Chaos, Strange Attractors, and Weather, 1989.

Turner, A. G. and Annamalai, H.: Climate change and the south Asian summer monsoon, Nat. Clim. Chang., 2, https://doi.org/10.1038/NCLIMATE1495, 2012.

805   Vijverberg, S. and Coumou, D.: The role of the Pacific Decadal Oscillation and ocean-atmosphere interactions in driving US temperature variability, npj Clim. Atmos. Sci., 5, 1–11, https://doi.org/10.1038/s41612-022-00237-7, 2022.

Vijverberg, S., Schmeits, M., van der Wiel, K., and Coumou, D.: Subseasonal Statistical Forecasts of Eastern U.S. Hot Temperature Events, Mon. Weather Rev., 148, 4799–4822, https://doi.org/10.1175/MWR-D-19-0409.1, 2020.

Weisheimer, A. and Palmer, T. N.: On the reliability of seasonal climate forecasts, J. R. Soc. Interface, 11, 20131162,
810   https://doi.org/10.1098/rsif.2013.1162, 2014.

Wiedermann, M., Donges, J. F., Handorf, D., Kurths, J., and Donner, R. V.: Hierarchical structures in Northern Hemispheric extratropical winter ocean–atmosphere interactions, Int. J. Climatol., 37, 3821–3836, https://doi.org/10.1002/joc.4956, 2017.

Wilks, D. S.: Canonical Correlation Analysis (CCA), Int. Geophys., 100, 563–582, https://doi.org/10.1016/B978-0-12-385022-5.00013-0, 2011.

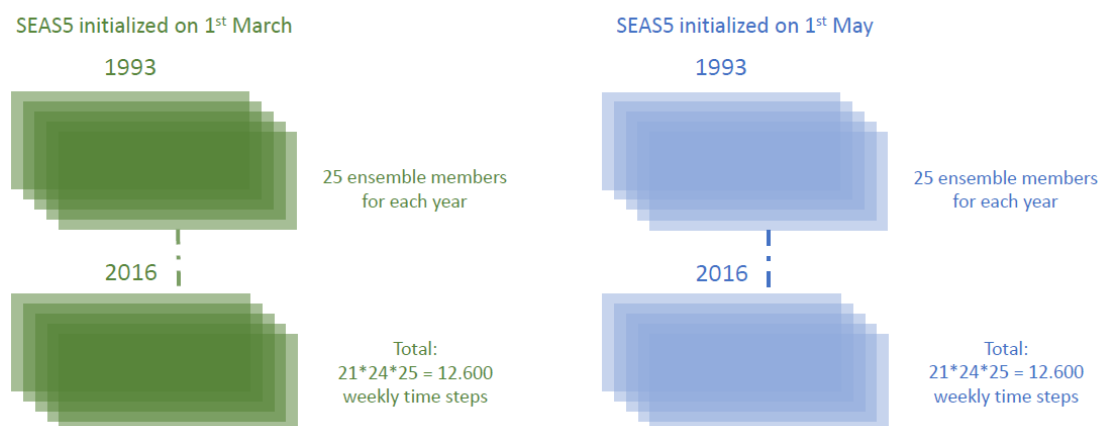
815

Figures

Figure 1. Schematic of the SEAS5 forecasting setup. Panel (a) shows the time line for SEAS5 initialization and target period. Panel (b)
820   shows a schematic of the SEAS5 ensemble members.

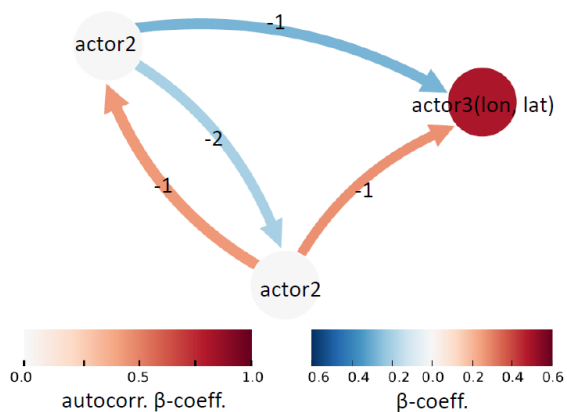
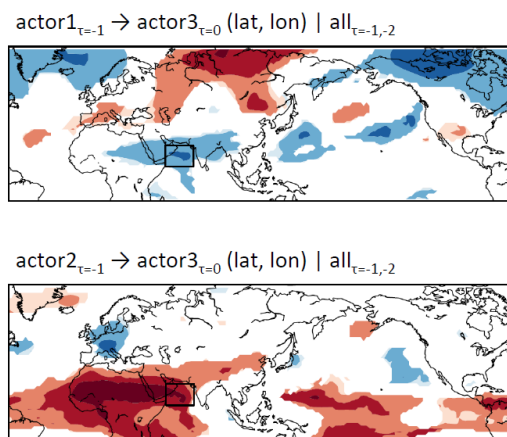



Figure 2. Schematic of CEN and Causal map. Panel (a) an example of a CEN built with three actors and lag$_{max}$ -2. Panel (b) shows an example of a causal map: the β value for the causal effect of actor1 and actor2 on actor3 (a time-varying field) vary with the latitude and longitude in the map.

825

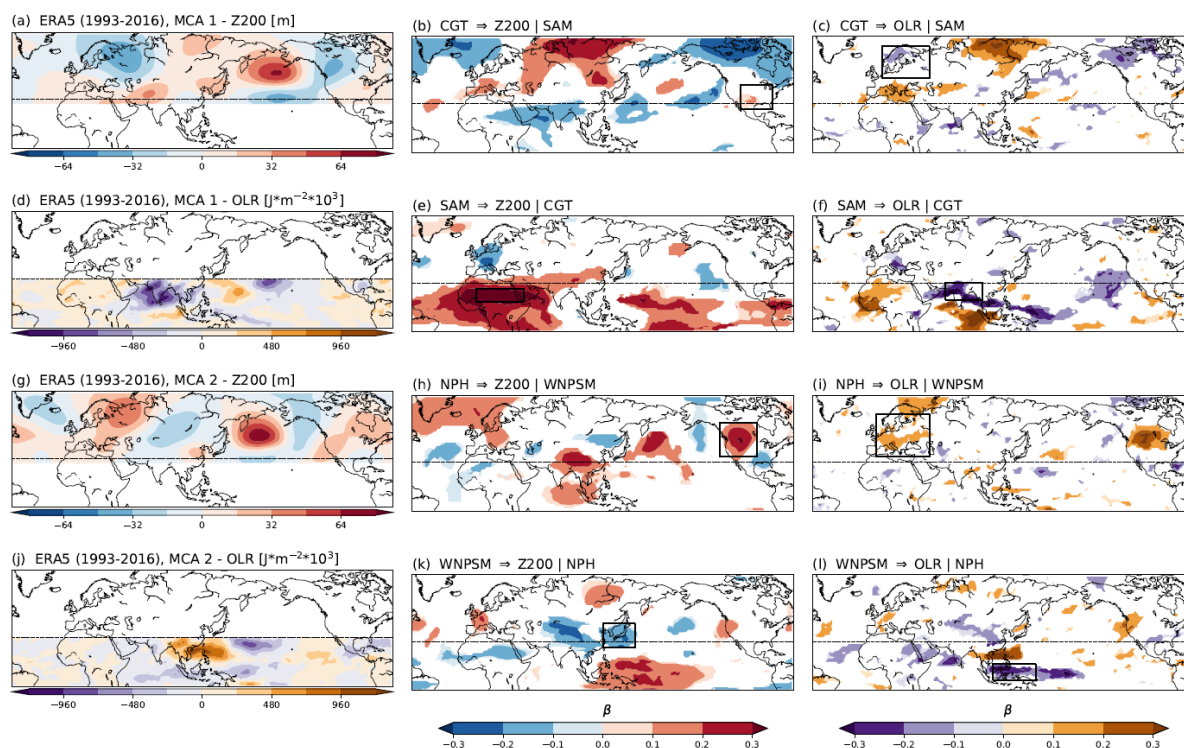



Figure 3. Causal maps for ERA5. Left column: MCA patterns for Z200 and OLR fields for ERA-S. MCA mode 1 Z200 (panel a) shows the CGT pattern. MCA mode 1 OLR shows the SAM (panel d). MCA mode 2 Z200 shows the NPH (panel g). MCA mode 2 OLR shows the WNPSM (panel j). Central column (Panels b,e,h,k): causal maps for the effect of each MCA pattern on Z200 fields. Right column (Panels c,f,i,l): causal maps for the effect of each MCA pattern on OLR fields.

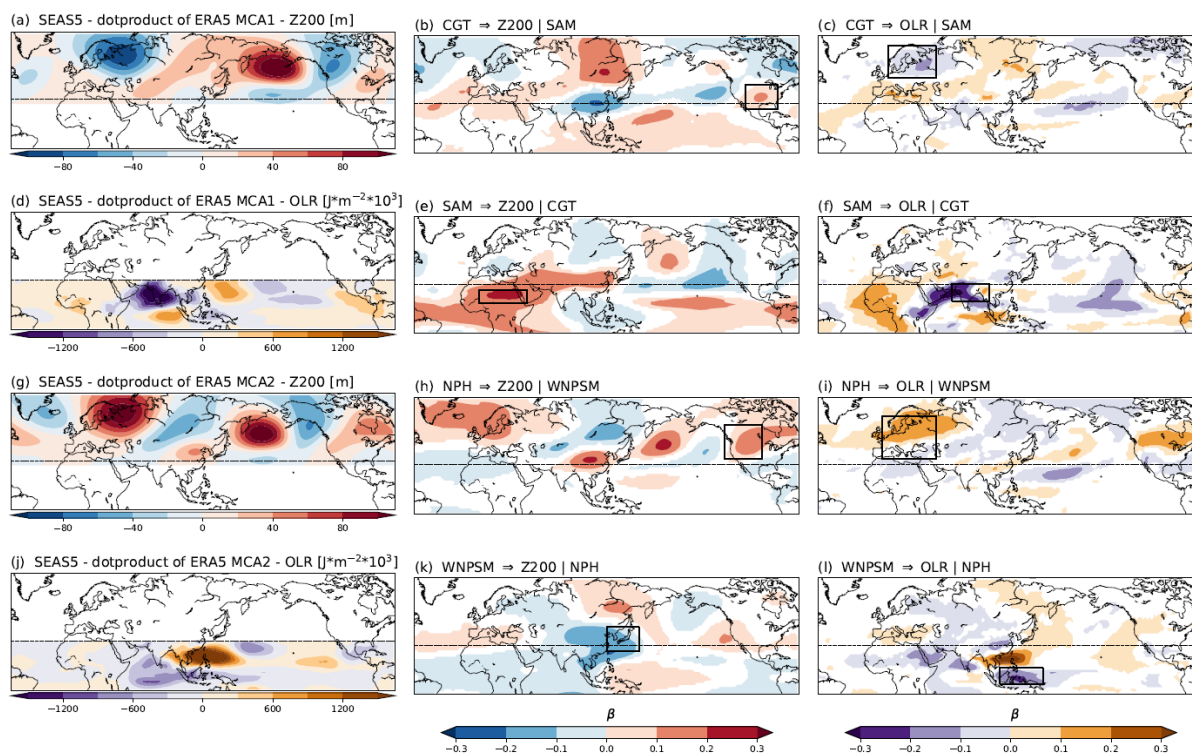



Figure 4. Causal maps for SEAS5. Same as for Fig. 3 but for SEAS5-ERA5 MCA (calculated as the projection of ERA-S on SEAS5 Z200 and OLR fields) and related causal maps.

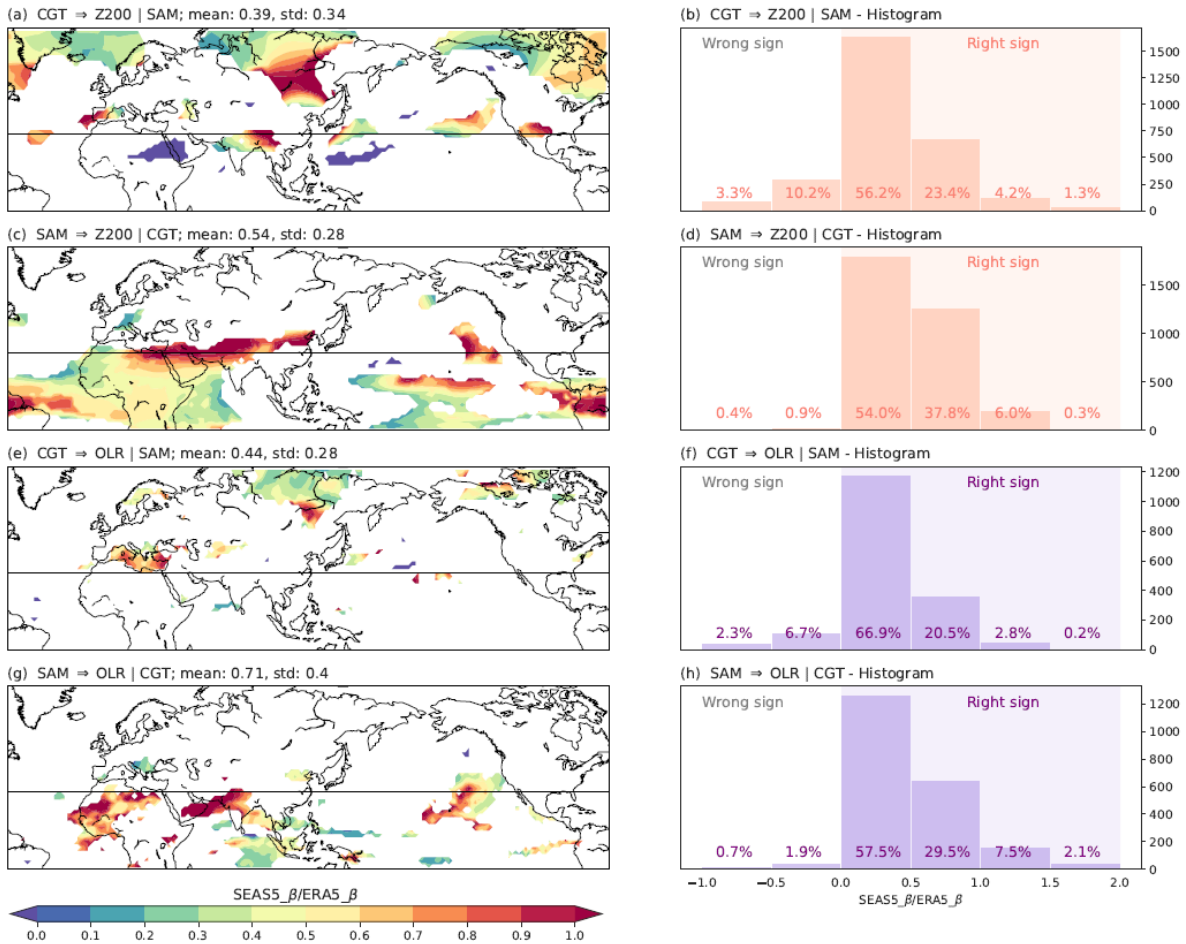



Figure 5. Differences between ERA5 and SEAS5 for MCA mode 1. Left column: Δ_β maps, where $\Delta_\beta = \frac{\beta_{SEAS5}}{\beta_{ERA5}}$ for CGT → Z200|SAM (panel a), SAM → Z200|CGT (panel c), CGT → OLR|SAM (panel e) and SAM → OLR|CGT (panel g). Right column: Histograms for SEAS5-ERA5 Δ_β for CGT → Z200|SAM (panel b), SAM → Z200|CGT (panel d), CGT → OLR|SAM (panel f) and SAM → OLR (panel h). All grid-points in the domain 15°S-75°N, 0°-360°E are used.

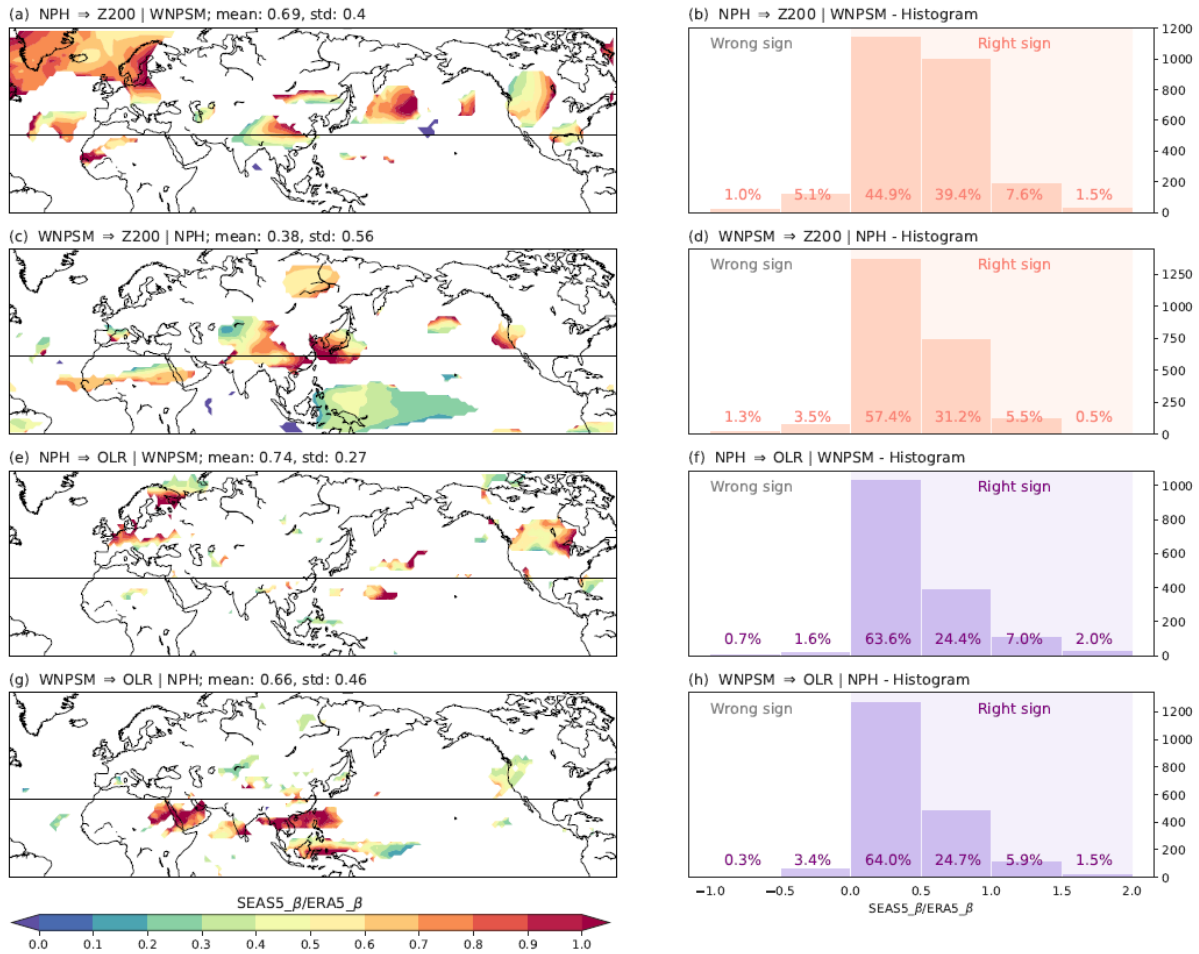



Figure 6. Differences between ERA5 and SEAS5 for MCA mode 2. Left column: Δ_β maps where $\Delta_\beta = \frac{\beta_{SEAS5}}{\beta_{ERA5}}$ for NPH → Z200 (panel a), WNPSM → Z200 (panel c), NPH → OLR|WNPSM (panel e) and WNPSM → OLR|NPH (panel g). Right column: Histogram for SEAS5-ERA5 Δ_β for NPH → Z200|WNPSM (panel b), WNPSM → Z200|NPH (panel d), NPH → OLR (panel f) and WNPSM → OLR|NPH (panel h). All grid-points in the domain 15°S-75°N, 0°-360°E are used.

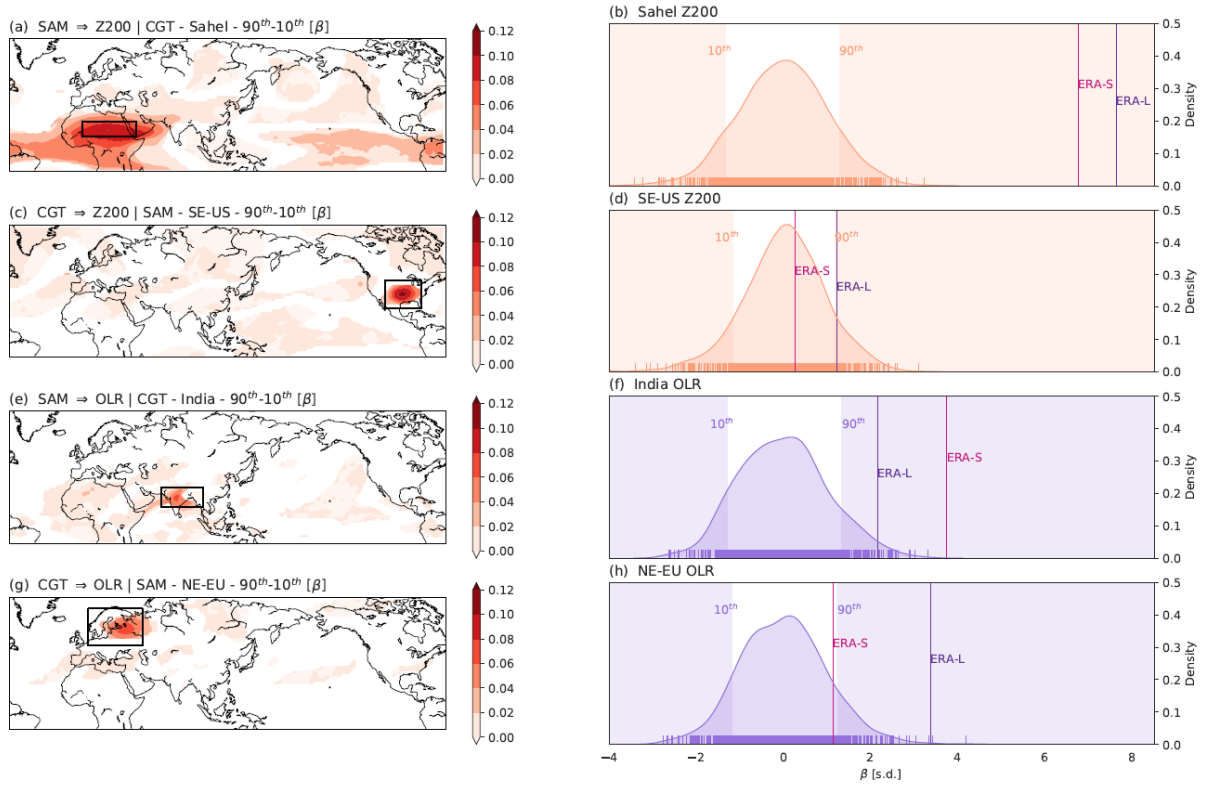



Figure 7. Ensemble spread for MCA mode 1 SEAS5-ERA5. Causal effect averaged over four key regions. Composites of high (>90th quantile) minus low (<10th quantile) causal effect over the Sahel region (Panel a), south-eastern US (Panel c), India (Panel e) and northeaster Europe (Panel g). PDF (standardized) of the distribution with highlighted 10th and 90th quantiles for the Sahel region (Panel b), southeaster US (Panel d), India (Panel f) and north-eastern Europe (Panel h). The value for ERA-L (ERA-S) is shown in the PDFs for comparison by a vertical solid purple (magenta) line.

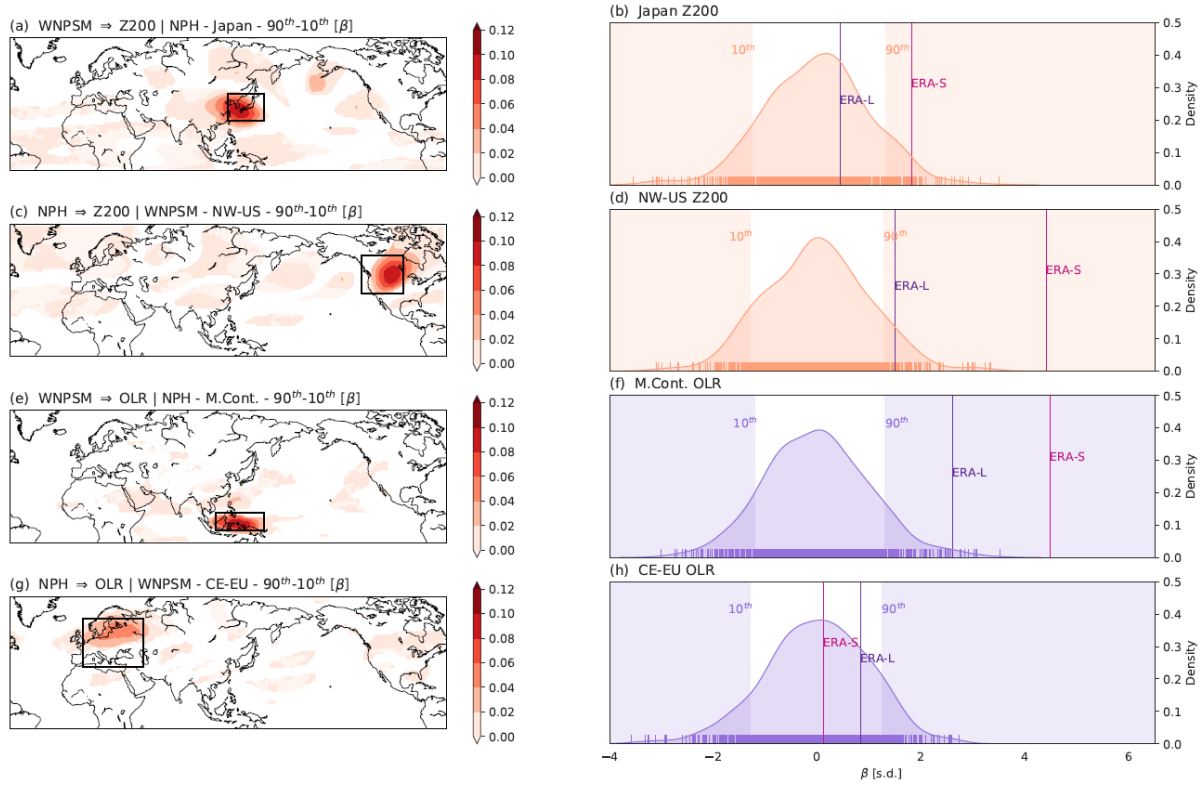



Figure 8. Ensemble spread for MCA mode 2 SEAS5-ERA5. Causal effect averaged over four key regions. Composites of high (>90th quantile) minus low (<10th quantile) causal effect over Japan (Panel a), north-western US (Panel c), Maritime continent (Panel e) and central eastern Europe (Panel g). PDF (standardized) of the distribution with highlighted 10th and 90th quantiles for Japan (Panel b), north-western US (Panel d), Maritime Continent (Panel f) and central eastern Europe (Panel h). The value for ERA-L (ERA-S) is shown in the PDFs for comparison by a vertical solid purple (magenta) line.

860

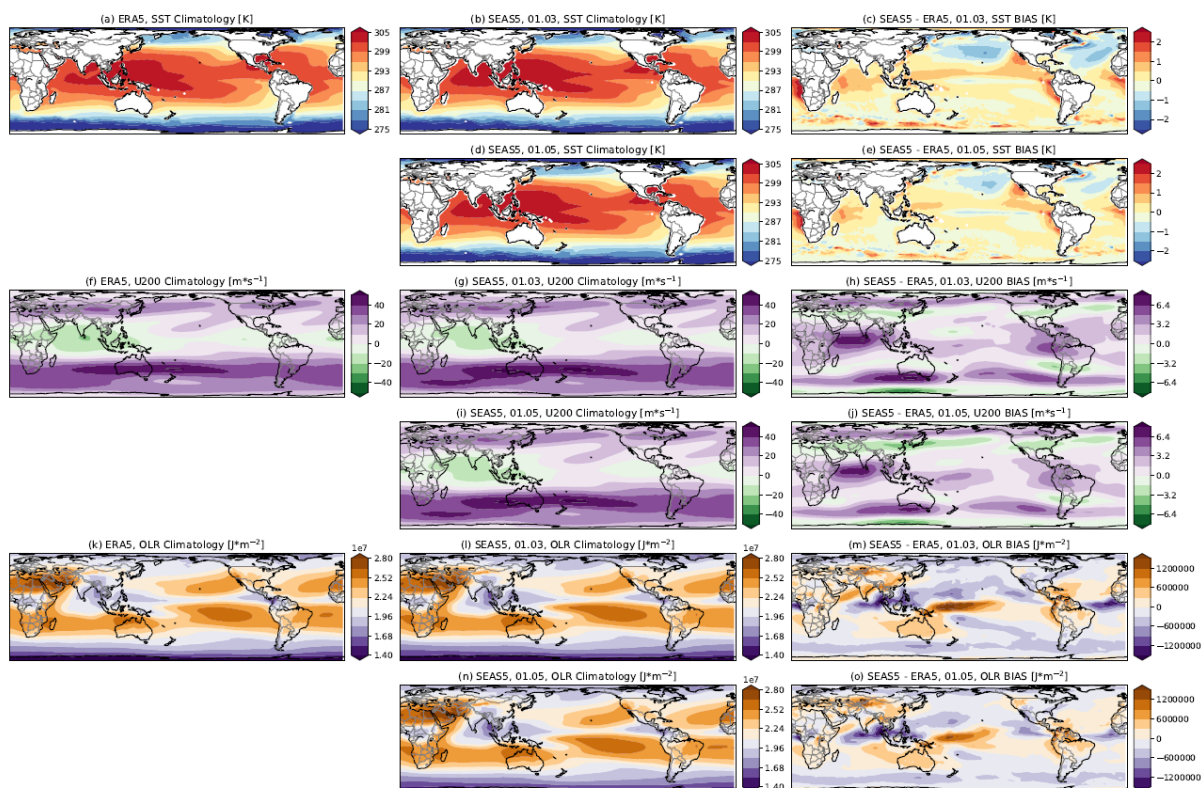



Figure 9. Bias between ERA5 and SEAS5. Panel (a) shows the JJAS climatology of SST for ERA-S. Panels (b) shows the JJAS climatology of SST for SEAS5 initialized on the 1st of March. Panel (c) shows the BIAS between ERA-S and SEAS5 for SST fields calculated by subtracting ERA-S from the SEAS5 climatology. Panels (d) and (e) same as for panels (b) and (c) but for SEAS5 initialized on the 1st of May. Panels (f)-(j) same as for panels (a)-(e) but for U200 fields. Panels (k)-(o) same as for panels (a)-(e) but for OLR fields.

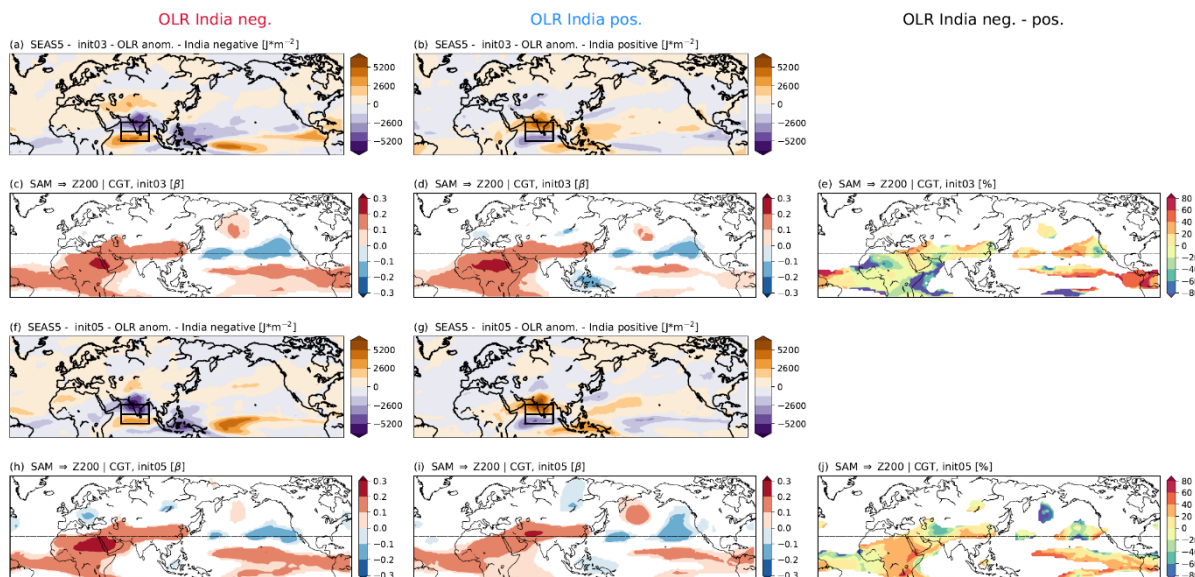



Figure 10. Convective activity bias and the Sahel region. Panel (a) shows the composite of JJAS averaged OLR fields for years with negative OLR anomalies over the Indian peninsula for SEAS5 initialized on the 1st of March. Panel (b) same as for panel (a) but for years with positive OLR anomalies over the Indian peninsula. Panel (c) shows the causal map for the link SAM → Z200|CGT obtained with OLR India negative years. Panel (d) same as for panel (c) but for OLR India positive years. Panel (e) shows the difference Δ_β between β values shown in panels (c) and (d). Panels (f) to (j) same as for panels (a) to (e) but for SEAS5 initialized on the 1st of May. The black boxes show the two areas used to define the OLR India index (see text for further description).

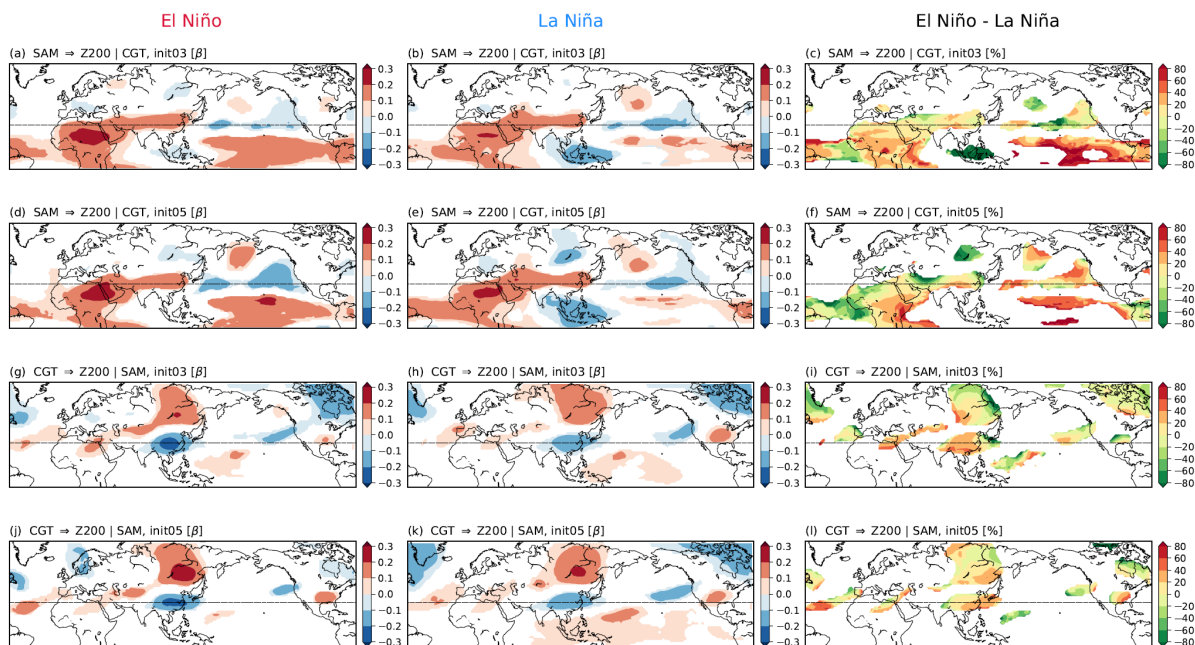



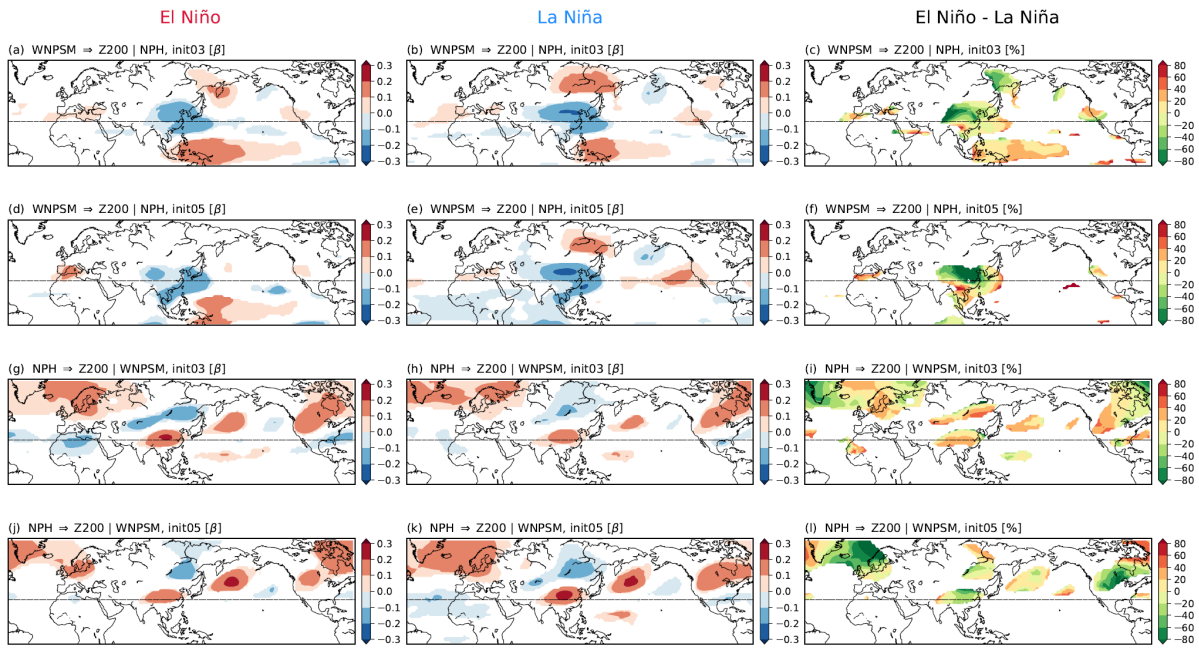
Figure 11. ENSO effect on tropical-extratropical links: MCA mode 1. Panels (a) and (b) show the causal maps for SEAS5 data initialized on the 1st of March for the SAM → Z200|CGT link respectively during Nino 3.4 positive and negative years. Panel (c) shows the differences between Nino 3.4 positive and negative years. Panel (d), (e) and (f): same as for panels (a), (b) and (c) but for SEAS5 data initialized on the 1st of May. Panels (g), (h) and (i): same as for panels (a), (b) and (c) but for the link CGT → Z200|SAM. Panels (j), (k) and (l): same as for panels (g), (h) and (i) but for SEAS5 data initialized on the 1st of May.

885

Figure 12. ENSO effect on tropical-extratropical links: MCA mode 2. Panels (a) and (b) show the causal map for SEAS5 data initialized on the 1ˢᵗ of March for the WNPSM → Z200|NPH link respectively during Nino 3.4 positive and negative years. Panel (c) shows the differences between Nino 3.4 positive and negative years. Panels (d), (e) and (f): same as for panels (a), (b) and (c) but for SEAS5 data initialized on the 1ˢᵗ of May. Panels (g), (h) and (i): same as for panels (a), (b) and (c) but for the link NPH → Z200|WNPSM. Panels (j),

890 (k) and (l): same as for panels (g), (h) and (i) but for SEAS5 data initialized on the 1ˢᵗ of May.

895 **Tables**

ACRONYM	EXTENDED	METHOD – FIELD
SAM	South Asian Monsoon	MCA mode 1, OLR
WNPSM	Western North Pacific summer monsoon	MCA mode 2, OLR
CGT	Circumglobal teleconnection	MCA mode 1, Z200
NPH	North Pacific High	MCA mode 2, Z200
ERA-S	ERA5 short (1993-2016)	
ERA-L	ERA5 long (1979-2020)	
SEAS5-ERA5	SEAS5 (1993-2016), MCA as projection of ERA5-S MCA on SEAS5 data	
SEAS5	SEAS5 (1993-2016), MCA calculated directly on SEAS5	

Table 1: List of the main acronyms used in the manuscript.

	ERA-S – ERA-L	ERA-S – SEAS5		ERA-S – SEAS5-ERA
	$MCA_i – MCA_i\ (i=j)$	$MCA_i – MCA_i\ (i=j)$	$MCA_i – MCA_j\ (i\neq j)$	$MCA_i – MCA_i\ (i=j)$
MCA1 Z200 (CGT)	**0.89**	**0.61**	-0.40	**0.93**
MCA2 Z200 (NPH)	**0.84**	**0.60**	**0.57**	**0.93**
MCA1 OLR (SAM)	**0.85**	**0.58**	-0.47	**0.86**
MCA2 Z200 (WNPSM)	0.77	**0.40**	**0.44**	**0.83**

Table 2: Spatial pattern correlation between MCA modes obtained from ERA5 data over the periods 1979-2020 and 1993-2016,
900 **between ERA5 and SEAS5 data over the common period 1993-2016 and for the same period but as the projection of ERA5 MCA (see**
main text for more details). Numbers highlighted in bold are significant at *p*-value = 0.05.

	REGION	SPATIAL DOMAIN	CAUSAL LINK
MCA MODE 1	Sahel	13°-45°N, 0°-45°E	SAM → Z200|CGT
	Southeast US	25°-45°N, 250°-280°E	CGT → Z200|SAM
	India	15°-30°N, 65°-100°E	SAM → OLR|CGT
	Northeast Europe	50°-70°N, 5°-50°E	CGT → OLR|SAM
MCA MODE 2	Japan	25°-45°N, 120°-150°E	WNPSM → Z200|NPH
	Central western US	35°-60°N, 230°-265°E	NPH → Z200|WNPSM
	Maritime Continent	5°S-10°N, 110°-150°E	WNPSM → OLR|NPH
	Central eastern Europe	35°-60°N, 0°-50°E	NPH → OLR|WNPSM

Table 3: Spatial domains of selected β regions.

905