

# Using causal effect networks to analyze different arctic drivers of midlatitude winter circulation

Article

Accepted Version

Kretschmer, M. ORCID: https://orcid.org/0000-0002-2756-9526, Coumou, D., Donges, J. F. and Runge, J. (2016) Using causal effect networks to analyze different arctic drivers of midlatitude winter circulation. Journal of Climate, 29 (11). pp. 4069-4081. ISSN 1520-0442 doi: 10.1175/JCLI-D-15-0654.1 Available at https://centaur.reading.ac.uk/92431/

It is advisable to refer to the publisher's version if you intend to cite from the work. See <u>Guidance on citing</u>. Published version at: http://dx.doi.org/10.1175/JCLI-D-15-0654.1 To link to this article DOI: http://dx.doi.org/10.1175/JCLI-D-15-0654.1

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the <u>End User Agreement</u>.

www.reading.ac.uk/centaur



# CentAUR

# Central Archive at the University of Reading

Reading's research outputs online

- Using Causal Effect Networks to analyze different Arctic 1 drivers of mid-latitude winter circulation 2 Marlene Kretschmer<sup>\*1,2</sup>, Dim Coumou<sup>1</sup>, Jonathan F. Donges<sup>1,3</sup>, Jakob Runge<sup>1,4</sup> 3 4 <sup>1</sup>Potsdam Institute for Climate Impact Research, Germany 5 <sup>2</sup>Department of Mathematics, University of Potsdam, Germany 6 <sup>3</sup>Stockholm Resilience Centre, Sweden 7 <sup>4</sup>Department of Physics, Humboldt University Berlin, Germany 8 \*Corresponding author: M. Kretschmer, kretschmer@pik-potsdam.de, Telegrafenberg A62, 14473 Potsdam, Germany 9
- 10

## 11 Abstract

In the past years, the northern hemisphere mid-latitudes have suffered from severe winters like 12 13 the extreme 2012/2013 winter in Eastern USA. These cold spells were linked to a meandering 14 upper tropospheric jet stream pattern and a negative Arctic Oscillation Index (AO). However, 15 the nature of the drivers behind these circulation patterns remains controversial. Various studies have proposed different mechanisms related to changes in the Arctic, most of them 16 related to a reduction in sea ice concentrations or increasing Eurasian snow cover. 17 Here, a novel type of time series analysis, called *Causal Effect Networks* (CEN) based on 18 graphical models is introduced to assess causal relationships and their time-delays between 19 20 different processes. The effect of different Arctic actors on winter circulation on weekly to 21 monthly time-scales is studied and robust network patterns are found. Barents and Kara sea ice 22 concentrations are detected to be important external drivers of the mid-latitude circulation, 23 influencing winter AO via tropospheric mechanisms and through processes involving the 24 Stratosphere. Eurasia snow cover is also detected to have a causal effect on sea level pressure in 25 Asia, but its exact role on AO remains unclear. The CEN approach presented in this study overcomes some difficulties in interpreting correlation analyses, complements model 26 27 experiments for testing hypotheses involving teleconnections, and can be used to assess their

validity. Our findings confirm that sea ice concentrations in autumn in the Barents and Kara Seas
are an important driver of winter circulation in the mid-latitudes.

30

# 31 **1. Introduction**

32 The recent cold winters in North America and Eurasia were characterized by a meandering jet stream pattern which allowed cold arctic air to reach lower latitudes [Cohen et al., 2014b]. 33 34 Moreover, these winters were dominated by a negative phase of the Arctic Oscillation Index 35 (AO), which is usually associated with pronounced meridional wind patterns, whereas in a 36 positive AO phase strong zonal flow dominates the wind field. Although a negative AO and meandering flow patterns have been linked to surface extremes [Thompson, 2001; Coumou et 37 al., 2014; Screen and Simmonds, 2014], it is intensively discussed what the mechanisms behind 38 39 AO variability are.

Classical atmosphere dynamic theories relate a meandering jet stream structure to above normal sea surface temperatures in the tropical Pacific [*Palmer and Mansfeld*, 1984; *Palmer and Owen*, 1986; *Trenberth et al.*, 1998]. Warming of the tropical Pacific intensifies evaporation, increasing thunderstorm activity in that region. The associated latent heat release can then trigger large-amplitude planetary waves affecting the mid-latitude flow.

45 In contrast, some recently proposed theories focus on the polar region, claiming that anomalous atmospheric circulations can be linked to low Arctic sea ice concentrations as observed during 46 the last two decades [Petoukhov and Semenov, 2010; Francis and Vavrus, 2012; Jaiser et al., 47 2012; Handorf et al., 2015]. A reduction in sea ice cover in summer leads to the ocean taking up 48 more energy in this season. Since sea ice works as an insulating shield blocking the ocean-49 50 atmosphere interaction, less sea ice in autumn and early winter facilitates larger heat fluxes from the relatively warm ocean into the atmosphere. Kim et al. focus on the Barents and Kara 51 52 Seas in particular and argue that reduction in sea ice concentration preferentially in this area lead to a weakened AO via the stratospheric polar vortex [Kim et al., 2014]. They link the 53 54 additional heat release to the atmosphere caused by sea ice loss in early winter to anomalously

55 high geopotential heights over the Barents and Kara Sea region in addition to lower than normal 56 geopotential heights over Northern Western Europe and Eastern Asia. This observed wave-like 57 structure indicates upward propagation of large scale planetary waves into the Stratosphere, interfering with the predominantly zonal flow in the lower Stratosphere. As a result, the 58 59 stratospheric zonal flow weakens and the geopotential heights and wind anomalies descend to the Troposphere, which is also called a "breakdown" of the polar vortex. As a consequence, cold 60 61 Arctic air reaches lower latitudes thereby forming large meanders. Those pressure anomalies, 62 respectively meandering of the jet stream, are then most often reflected in a negative phase of AO. Kim et al. (2104) base their analysis on theoretical physical considerations and observational 63 data. They validate their results using climate model simulations, which reproduce similar 64 patterns, supporting their proposed theory. 65

66 A similar mechanism was proposed by Cohen et al. who linked increased fall snow cover in 67 Eurasia to changes in surface pressure anomalies, causing a likewise chain of effects [Cohen et al., 2007, 2013, 2014a]. Based on observational data and correlation analysis they hypothesize 68 that an extended Eurasian snow cover in fall, likely resulting from decreasing Arctic sea ice, 69 70 leads to increasing sea level pressures over Central Asia in early winter. As a result a disturbed 71 pressure pattern in the polar region is observed, leading to increased vertical wave activity and poleward heat flux. This is followed by anomalously high geopotential heights in the 72 73 Stratosphere, associated with stratospheric warming and weakening of the polar vortex, and 74 respectively a negative surface AO as described by Baldwin and Dunkerton (1999).

75 In order to study the atmospheric response to changes in the Arctic, different methods have 76 been used. Cross-correlation analysis is widely applied to detect linear relationships and their 77 time delays between different processes [Polvani and Waugh, 2004; Cohen et al., 2014a]. 78 However, correlation can be highly biased by auto-correlation effects, by indirect connections 79 via a third process, or by a common driver leading to non-causal, spurious correlations that 80 limits its interpretability [Runge et al., 2014]. Also, it does not give any answer on the direction 81 of the relationship, such that it is not an adequate tool to study causal effects. Therefore climate 82 models are used, to investigate atmospheric changes due to a controlled perturbation of the system [Deser et al., 2010; Petoukhov and Semenov, 2010; Handorf et al., 2015]. This approach 83

allows interpreting results as causal effects forced by the input data. However, conclusions are
strictly limited to the extent of the physical realism of the climate model used. It remains
questionable whether models capture important processes like ocean-ice feedbacks [*Tremblay et al.*, 2007], land-snow interactions [*Furtado et al.*, 2015], troposphere-stratosphere
interactions [*Manzini et al.*, 2014] and Rossby wave propagation [*Gray et al.*, 2014] accurately.
Thus, both climate model experiments and correlation analysis of observational data are
restricted in their interpretability [*Barnes and Screen*, 2015].

91 Here we analyze observational data with a novel method based on graphical models called 92 *Causal Effect Networks (CEN)*. This method overcomes spurious correlations due to 93 autocorrelation, indirect effects or common drivers (at least among the observed variables included) using a causal discovery algorithm as proposed by Runge et al. (2012a, 2012b, 2014). 94 95 This algorithm is a modified version of the PC-algorithm [Spirtes et al., 2000] (named after its 96 inventors Peter Spirtes and Clark Glymour) which has first been applied to climate research by Ebert-Uphoff and Deng [Ebert-Uphoff and Deng, 2012] to study interactions between major 97 climate modes. Causal discovery approaches have since then been used to study atmospheric 98 99 flows [Deng and Ebert-Uphoff, 2014], causal relationships in the Walker Cell in the Tropics 100 [Runge et al., 2014], the monsoonal dynamics in the Pacific-Indian Ocean [Runge et al., 2015] and decadal ocean circulation in the Atlantic [Schleussner et al., 2014]. 101

The aim of this paper is to explain how to apply this method and show how it can be used for hypothesis testing in the context of teleconnections in climate research. We apply CEN to observational and reanalysis data in order to understand how different mechanisms which might cause a negative AO in winter are causally related with each other. In this study, we limit ourselves to testing a set of proposed Arctic mechanisms. In contrast to tropical mechanisms, those operate on similar sub-seasonal timescales, which facilitates a simultaneous analysis.

The article is structured as follows: In section 2 the data selection is motivated and section 3 gives a detailed description of the two different steps of the CEN-algorithm on the basis of an example. In section 4 the sensitivity of the parameter settings and temporal resolution is analyzed and structure and robustness of the graphs are discussed in the framework of the

tested hypothesis. Finally, in section 5 we conclude and assess the potentials and limitations ofthe presented method.

- 114
- 115

### 116 **2. Data**

Different actors can influence mid-latitudinal winter circulation. The first step of our analysis is hence to come up with a reasonable choice of processes which are expected to be relevant for the analysis. This includes the decision for physical variables which should serve as proxies for the considered processes, the selection of suitable data sources and a reasonable time resolution of the data.

As stated, we limit the analysis to Arctic processes and follow Kim et al. (2014) and Cohen et al. 122 (2014) with respect to data selection. We therefore include Barents and Kara sea ice 123 concentrations (BK-SIC) as well as Eurasia snow cover (EA-snow) in our analysis, as possible 124 125 causal drivers of a negative Arctic Oscillation Index (AO). We further include sea level pressure in the Ural Mountains region (Ural-SLP) as defined in [Cohen et al., 2014] and sea level pressure 126 in the Lake Baikal area as a proxy for Siberian High variability (Sib-SLP). Following Kim et al. and 127 Cohen et al. we include the zonally averaged poleward heat flux v\*T\* at 100 mb (v-flux) to 128 capture the Troposphere-Stratosphere coupling. This is a widely used proxy for vertical wave 129 130 activity, whereby v denotes the meridional wind velocity, T stands for temperature and the asterisk denotes deviations from the zonal mean [Polvani and Waugh, 2004; Dunn-Sigouin and 131 132 Shaw, 2015]. There are many possible ways to describe polar vortex activity (PoV), but for 133 consistency with Kim et al. (2014) and Cohen et al. (2014) we calculate geopotential height anomalies poleward of 65°N, averaged over pressure levels from 10mb to 100mb to define the 134 strength of the stratospheric polar vortex. Eurasia snow data is described in [Robinson et al., 135 1993] and is provided by NOAA<sup>1</sup>. Sea ice concentration data was taken from the Nimbus-7 136 SMMR and DMSP SSM/I-SSMIS passive microwave data set provided by the National Snow & Ice 137

http://gis.ncdc.noaa.gov/all-records/catalog/search/resource/details.page?id=gov.noaa.ncdc:C00756

Data Center<sup>2</sup>. The Arctic Oscillation Index (AO) is provided by NOAA<sup>3</sup> and for the remaining variables we used ERA-Interim reanalysis data<sup>4</sup>.

140 In summary, our analysis contains seven different actors (Tab. 1): Barents and Kara sea ice 141 concentrations (BK-SIC), Eurasia snow cover (EA-snow), the Arctic Oscillation Index (AO), vertical 142 wave activity (v-flux), polar vortex strength (PoV), sea level pressure over the Ural Mountains 143 (Ural-SLP) and Siberian High activity (Sib-SLP). For each variable we consider the time-period 144 01/1979-12/2014, which is most reliable in the reanalysis due to availability of satellite data.

145 We calculate monthly means of daily data for each variable as we are testing mechanisms which are expected to act on monthly time scales. Thereby we perform linear interpolation of the 146 147 snow data and for some years of the sea ice concentration data set, where daily data is not available. To gain additional information on the time-scale of the considered processes we 148 perform additional analysis using half-month means as well as quarter-month means of every 149 variable (Fig. 1). For half-monthly data we take the mean from the 1<sup>st</sup> - 15<sup>th</sup> and from the 16<sup>th</sup>-150 30<sup>th</sup> of each month and for February from 1<sup>st</sup> - 14<sup>th</sup> and 15<sup>th</sup> - 28<sup>th</sup> respectively (thus ignoring the 151 31<sup>th</sup> of all applicable months as well as the 29<sup>th</sup> of February in leap years). To construct quarter-152 monthly time series we calculate the mean from 1<sup>st</sup> - 7<sup>th</sup>, 9<sup>th</sup> - 15<sup>th</sup>, 16<sup>th</sup> - 22<sup>th</sup> and 24<sup>th</sup> - 30<sup>th</sup> 153 (neglecting hence the 8<sup>th</sup>, 23<sup>th</sup> and 31<sup>th</sup> of all applicable months) and for February from 1<sup>st</sup> - 7<sup>th</sup>, 154 8<sup>th</sup> - 14<sup>th</sup>, 15<sup>th</sup> - 21<sup>th</sup> and 22<sup>th</sup> - 28<sup>th</sup> respectively. This approach has the advantage that the 155 156 different time-series are still in sync with each other facilitating the comparison of associated 157 CENs.

For each variable and time-resolution we calculate climatological anomalies (observed value minus the multi-year mean), from which we then compute the area-weighted spatial average over the defined region (see last column in Tab. 1). This way we create single time-series for each time-resolution and each actor (see Fig. 2 for monthly data). Since CEN construction

<sup>3</sup> http://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily\_ao\_index/ao.shtml

<sup>4</sup> http://apps.ecmwf.int/datasets/data/interim-full-daily/

<sup>&</sup>lt;sup>2</sup> http://nsidc.org/data/nsidc-0051

requires stationary time-series, we remove the linear trend if present. For our analysis this is only the case for Barents and Kara sea ice concentrations (BK-SIC). Additionally we change the sign of PoV, such that positive values (negative geopotential height anomalies) indicate a strong polar vortex.

- 166
- 167

### 168 **3. Method**

The Causal Effect Networks approach is based on two steps: (1) Reconstructing the causal 169 170 parents of each actor using a causal discovery algorithm [Runge et al., 2012a, 2012b, 2014], which is a modification of the PC-algorithm [Spirtes et al., 2000] for time series. As explained in 171 172 the following, this step is based on iterative conditional independence tests using partial correlation. (2) In a second step, the strength of causal links is quantified using a linear version 173 of Pearl's causal effect measures [Pearl, 2013]. Thereby the parents are used in a multiple linear 174 regression analysis to test the significance and strength of causal dependencies between all 175 pairs of actors at a range of time-lags. 176

Here we use a linear approach to estimate and interpret causal links, but the two-step 177 procedure of causal reconstruction and quantification can also be embedded in an information-178 179 theoretic framework to study causal information transfer accounting for nonlinear relationships between variables. For a detailed explanation of the method, including a mathematical analysis 180 181 as well as numerical testing, we refer to Runge et al. (2012a, 2012b, 2014). All calculations 182 presented in this study were performed using the Python package TiGraMITe (*Time series graph* based Measures of Information Transfer) which provides the CEN-algorithm and is freely 183 available<sup>5</sup>. 184

185 In the following we explain how to apply CEN to test causality of the hypotheses discussed in 186 the introduction.

https://www.pik-potsdam.de/members/jakrunge

187

#### 188 Step 1: Detecting Causal Effects

The first step of the CEN-algorithm aims to find causal relationships between the different actors and their associated time lags. The scope of this step is to identify past processes which directly influence each actor. We call those processes the parents of an actor and they will be used later to determine the actual strength and the sign of the causal relationships.

193 Cross-correlation can give a first impression of the pairwise linear relationship between two 194 processes X and Y. However, it is not able to identify causal links because the bivariate analysis 195 can be biased by autocorrelation of the two variables, by common drivers or by indirect links via 196 a third process Z (Fig. 3a, b, c). For example cross-correlation of two independent processes X 197 and Y can be high if one of the processes is strongly auto-correlated (Fig. 3a). Also, imagine that 198 Z causes X and Y (Fig. 3c) then cross-correlation analysis would find a strong correlation between X and Y even though there in no direct link between them. In order to detect causal 199 200 links, a multivariate analysis is required which takes all potential actors into account.

Recall that two processes X and Y are conditionally independent given a third process Z if  $P(X \cap Y|Z)=P(X|Z)P(Y|Z)$ , whereby P denotes the probability function. In the linear case this can be tested by removing the linear influence of Z from both X and Y and testing for the correlation between their residuals (partial correlation). In the previous case (Fig. 3c), X and Y would then be conditionally independent given Z. In the example illustrated in Fig. 3b, process X causes Z which in turn influences Y. Process X and Y are thus conditionally independent given Z and a high correlation coefficient between X and Y only occurs due to the indirect link via Z.

208 This section discusses how the CEN-algorithm uses iterative partial correlations to identify non-209 causal correlations as depicted in Fig. 3. The extent to which such a data-based analysis allows 210 to conclude on a physical causal mechanism depends on the included variables, time resolution 211 of the data and assumptions such as stationarity. Two free parameters are involved: the 212 significance level  $\alpha$  for the partial correlation tests and the maximum time delay  $\tau_{max}$ .

#### 213 Calculating the parent processes

As an illustrative example, we start with finding those processes on a monthly time-scale among our actors which have a direct causal effect on the winter (December, January, February) polar vortex (PoV). We look at the monthly time-series for every actor (Fig. 2) having thus a samplelength of 108 time-steps. We define a two-sided significance level  $\alpha$ =0.01 and a maximum timelag of  $\tau_{max}$ =3 months implying that parent processes more than three months ago or those with a significance below 99% will be neglected.

First, for every actor X the cross-correlation function  $\rho(X_{t-\tau}, PoV_t)$  is calculated for time shifts of 220  $\tau$ =1 up to the maximum time-shift  $\tau_{max}$ =3 months. Note that, if we study causal effects on winter 221 PoV, this implies that the monthly time-series PoV<sub>t</sub> only consists of winter data but the lagged 222 223 or driving variable contains data from other seasons (in particular autumn but also summer when  $\tau$ >3). Here the expression "driver" is used in its statistical meaning of being conditional-224 225 dependent and shifted in time. For  $\tau=1$  the expression  $\rho(X_{t-1}, PoV_t)$  denotes the Pearson correlation coefficient of November-December-January data of process X and December-226 227 January-February data of PoV (see Fig. 4) whereas for  $\tau=3$  the linear influence of the three months shifted September-October-November data of actor X on PoV in winter (DJF) is 228 229 measured. For example, for the influence of Eurasia snow cover (X=EA-snow) on the polar 230 vortex with a time delay of  $\tau$ =1 we obtain:

241

$$\rho(EA-snow_{t-1}, PoV_t) = -0.262$$

232 which is significant at the  $\alpha$ =0.01 level. This indicates that there is a negative linear relationship 233 between early winter (NDJ) snow and the winter polar vortex. This seems reasonable since a large snow cover in Eurasia is indicated to induce a weakened polar vortex [Cohen et al., 2014]. 234 The cross-correlation function is now calculated and evaluated for every actor XE{BK-SIC<sub>t-t</sub>, EA-235 snow<sub>t-t</sub>, AO<sub>t-t</sub>, v-flux<sub>t-t</sub>, PoV<sub>t-t</sub>, Sib-SLP<sub>t-t</sub>, Ural-SLP<sub>t-t</sub>} and every time-lag  $\tau \in \{1,2,3\}$ . We find that 236 besides EA-snow (with  $\tau=1$ ), also Ural-SLP (with  $\tau=1$  and  $\tau=2$ ), AO (with  $\tau=1$ ), PoV (with  $\tau=1$ ) and 237 v-flux (with  $\tau$ =1) are significantly correlated with winter PoV. Sorted by the strength of 238 correlation starting with the strongest in absolute value, the set of potential parent-processes of 239 PoV in this zeroth iteration step without any conditioning is: 240

$$\mathbf{P}^{\mathbf{v}} = \{ v - flux_{t-1}, PoV_{t-1}, Ural - SLP_{t-1}, Ural - SLP_{t-2}, AO_{t-1}, EA - snow_{t-1} \}.$$

To test these potential drivers for conditional independence, we next calculate partial correlations:

244

#### ρ(X<sub>t-τ</sub>,Y<sub>t</sub> | **Z**)

245 which measure the linear influence from process X on Y, excluding the influence of some set of 246 variables Z. This thus checks if X and Y are conditionally independent given Z. We choose Z as a subset of  $P^0$  such that Z denotes a set of other processes which potentially influences the 247 bivariate correlation coefficient  $\rho(X_{t-\tau}, Y_t)$ . In each iteration step  $P^1$ ,  $P^2$ , ... we condition on a new 248 Z, whereby the algorithm first takes only one condition and starts with the process which is 249 strongest correlated (in absolute value) with process Y. Then the dimension of the subset 250 251 selected from the remaining parents is increased and different two-dimensional conditions are 252 tested and so on for higher dimensions. If the partial correlation significance test of a pair  $X_{t-\tau}$ and  $Y_t$  is non-significant given **Z**, the process  $X_{t-\tau}$  is removed from the set of potential parents. If, 253 254 however, the partial correlation  $\rho(X_{t-\tau}, Y_t | \mathbf{Z})$  remains significant for all tested **Z**, then actor X is considered to directly influence Y with a time-lag of  $\tau$ . 255

Returning to our example, we first test condition  $\mathbf{Z}=\{v-flux_{t-1}\}$  and find:

257 ρ(EA-snow<sub>t-1</sub>, PoV<sub>t</sub> | v-flux<sub>t-1</sub>) = - 0.147

which is not significantly different from zero at our chosen level and hence we find that EAsnow and PoV are conditionally independent (at a time-delay of one month) if the influence of v-flux from the same time shift is excluded. We thus conclude that there is no direct influence from EA-snow on PoV with a delay of one month and that the significant correlation between them  $\rho(EA-snow_{t-1}, PoV_t) = -0.261$  is due to the influence of v-flux. For example, EA-snow could be linked to PoV indirectly via v-flux (as in Fig. 3b). On the other hand, if we take X=Ural-SLP\_{t-1}  $\in$ **P**<sup>0</sup> and condition on the same **Z**={v-flux<sub>t-1</sub>} we find

265 
$$\rho(\text{Ural-SLP}_{t-1}, \text{PoV}_t | v-flux_{t-1}) = -0.281$$

which is still significantly different from zero. In other words, the linear influence of Ural-SLP<sub>t-1</sub> on PoV<sub>t</sub> cannot exclusively be explained by the linear influence of v-flux. We calculate partial correlations for all the elements from  $P^0$  conditioning on  $Z=\{v-flux_{t-1}\}$  and find that some of them are conditionally independent from PoV given v-flux<sub>t-1</sub>, which can thus be neglected as potential drivers of winter PoV. This way we obtain a much smaller set of potential parent processes of PoV:

272 
$$\mathbf{P}^{1} = \{v-flux_{t-1}, PoV_{t-1}, Ural-SLP_{t-1}\} \subset \mathbf{P}^{0}$$

Now the algorithm proceeds by conditioning on the process which was second strongest correlated with PoV, i.e.  $Z=\{PoV_{t-1}\}$ . We thus check if some of the potential drivers of PoV only occur due to the auto-correlation of PoV. Calculating partial correlations of the elements of  $P^1$ conditioning on  $Z=\{PoV_{t-1}\}$  gives only values significantly different from zero such that  $P^2 = P^1$ . The last possibility of picking only one condition is  $Z=\{Ural-SLP_{t-1}\}$ , where we find again that all the partial correlations remain significantly different from zero such that  $P^3 = P^2 = P^1$  Sorting the elements by the strength of their partial correlation value in the last iteration step we have:

280 
$$P^3 = \{v-flux_{t-1}, Ural-SLP_{t-1}, PoV_{t-1}\}$$

Now we increase the dimension of **Z** and condition on two possible drivers from  $P^3$ . Thus we start with  $Z = \{v-flux_{t-1}, Ural-SLP_{t-1}\} \subset P^3$  and calculate:

283 
$$\rho(PoV_{t-1}, PoV_t | v-flux_{t-1}, Ural-SLP_{t-1}) = 0.268$$

which is still significantly different from zero. When testing for the other possibilities ( $\mathbf{Z}$ ={v-flux<sub>t-</sub> 1, PoV<sub>t-1</sub>} and  $\mathbf{Z}$ ={Ural-SLP<sub>t-1</sub>, PoV<sub>t-1</sub>}), the partial correlations remain significant. Since there are no more combinations for choosing  $\mathbf{Z}$  the algorithm converges and stops.

We have now found the set of direct drivers of winter PoV (relative to the variables taken into account), which we call its parents denoted by:

289 
$$\boldsymbol{\mathcal{P}}_{PoV} = \{v-flux_{t-1}, Ural-SLP_{t-1}, PoV_{t-1}\}$$

290 In other words we found that (given the settings of  $\tau_{max}$ =3 and  $\alpha$ =0.01) winter polar vortex (PoV) 291 is directly driven by itself with a delay of one month and by vertical wave activity (v-flux) and 292 pressure variability in the Ural Mountains region (Ural-SLP) with a delay of one month, but is 293 (linearly) conditionally independent of all other processes. The procedure described for PoV is performed for all actors yielding a set of parents for every actor (see Tab. 2):

 $\boldsymbol{\mathcal{P}} = \{\boldsymbol{\mathcal{P}}_{AO}, \boldsymbol{\mathcal{P}}_{BK-SIC}, \boldsymbol{\mathcal{P}}_{EA-snow}, \boldsymbol{\mathcal{P}}_{v-flux}, \boldsymbol{\mathcal{P}}_{PoV}, \boldsymbol{\mathcal{P}}_{Sib-SLP}, \boldsymbol{\mathcal{P}}_{Ural-SLP}\}.$ 

297 Note that the interpretation of the significance level  $\alpha$  as the probability of false rejections of 298 the hypothesis of a non-causal link is not strictly valid here since we tested every possible link 299 multiple times by conditioning on different processes (see discussion section).

300

#### 301 Step 2: Quantifying Causal Effects

In the second step, we use the sets of parents to determine the strength of causal relationships. The case of  $\tau$ =0, i.e., when there is no time shift between the actors was omitted when calculating the parents. In this step we will nevertheless quantify the significant instantaneous relationships conditional on the parents. As stated above, such contemporaneous links can in general not be interpreted in a causal way. Some might turn out to be causal parents at a higher time resolution, but some might be just due to excluded common drivers. We address this issue later by studying different time lags.

As mentioned, the set of derived parents depends on the significance level  $\alpha$ , which here is, however, not well interpretable due to the multiple testing problem. In order to better assess significance, we therefore test every possible combination of actors and time-lags again (including links from parents) using the causal parents as a conditioning set.

In general, multiple linear regression can be used to measure the influence a system of variables (the independent variables) has on a different (dependent) variable. However, it can often be challenging to define a set of independent variables which can explain the dependent variable. The list of causal parents provides a reasonable choice for those variables with their associated time-lags. We calculate the link strength using standardized multiple linear regression coefficients based on our list of parents for the case of  $\alpha$ =0.01 and up to a maximum lag of  $\tau_{max}$ =3. We found that PoV is influenced from the past by  $\mathcal{P}_{PoV}$ = {v-flux<sub>t-1</sub>, Ural-SLP<sub>t-1</sub>, PoV<sub>t-1</sub>}. To

calculate if process X significantly influences PoV with a time-lag of  $\tau$ ≥0 we formulate the standardized linear regression model

322 
$$POV_{t}^{*} = \beta_{0} + \beta_{1}v - flux_{t-1}^{*} + \beta_{2}Ural - SLP_{t-1}^{*} + \beta_{3}POV_{t-1}^{*} + \beta_{4}X_{t-\tau}^{*} + \epsilon.$$

323 Here the beta coefficients  $\beta_i$  with i $\in$  {0,1,2,3,4} denote the standardized regression coefficients,  $\epsilon$ 324 stands for the error term and the asterisk indicates that the time-series have been normalized and standardized. The regression coefficients express how much the different independent 325 326 variables contribute to variability in PoV in terms of standard deviations. Interpreted causally [Pearl, 2013], this means that if X is increased by one standard-deviation keeping the other 327 variables fixed, then PoV increases by  $\beta_4$  standard-deviations. The  $\beta$ -coefficient of X is tested for 328 329 significance at  $\alpha$ =0.01 with the null-hypothesis  $\beta$ =0, which would mean that variable X does not 330 contribute significantly to the dependent variable PoV.

To test if, for example, X=EA-snow significantly influences winter PoV with a delay of one month  $\tau=1$  we calculate the standardized linear regression model and choose EA-snow t-1 as well as the parents of PoV as independent variables to explain PoV:

334 
$$POV_{t}^{*} = \beta_{0} + \beta_{1}v - flux_{t-1}^{*} + \beta_{2}Ural - SLP_{t-1}^{*} + \beta_{3}POV_{t-1}^{*} + \beta_{4}EA - snow_{t-\tau}^{*} + \epsilon$$

We get  $\beta_4$ = - 0.076 which is not significant at the  $\alpha$ =0.01 level such that the influence from EAsnow on PoV with a delay of one month is considered to be absent. If we, however, calculate the influence of v-flux with  $\tau$ =1 (which is also in  $\mathcal{P}_{Pov}$ ) on winter PoV we obtain a significant beta coefficient  $\beta_1$ = - 0.514. Thus, v-flux is concluded to be causally influencing the winter polar vortex with a delay of one month and with a strength of  $\beta_1$ = -0.514, i.e., a one-standard deviation increase in v-flux leads to a negative change of about half a standard deviation in PoV.

341 We test the influence of every actor XE{BK-SIC<sub>t-τ</sub>, EA-snow<sub>t-τ</sub>, AO<sub>t-τ</sub>, v-flux<sub>t-τ</sub>, PoV<sub>t-τ</sub>, Sib-SLP<sub>t-τ</sub>, 342 Ural-SLP<sub>t-τ</sub>} and every time-lag  $\tau$ E{0,1,2,3} on PoV as well as on every other actor in form of 343 standardized linear regression. The remaining significant links form our *Causal Effect Network*.

Note, that it is possible that in this step significant direct links are identified which had been rejected in the first step. Nevertheless, by testing every potential link again, we can better interpret the statistical meaning of  $\alpha$  as the probability of falsely rejecting the null hypothesis that a lagged variable  $X_{t-\tau}$  is independent of  $Y_t$  given the parents of  $Y_t$  selected with the causal algorithm. However, we will see that our list of parents strongly coincides with the significant strong links identified in the second step.

- 350
- 351

### 352 **4. Results & Discussion**

353 We construct CEN for winter circulation and with different  $\alpha$  and  $\tau_{max}$  settings. Visualization of 354 CEN as a process graph gives an easy to interpret picture of the underlying complex 355 teleconnection pattern. Only the significant links are presented in the graph and the numbers 356 next to the links stand for the associated time lag  $\tau$ . Instantaneous links are represented by 357 dashed links and have no direction or time-shift. The node color (in case the variable influences 358 itself) and the link color represent the standardized regression coefficient (beta values) and 359 hence capture the strength of the causal relationship. If two processes are linked for more than 360 one time-lag then all lags are given (sorted by strength) with the link color based on the strongest connection. The time-lag for auto-driven data is not shown in the graph, but 361 362 predominantly actors are lag-1 auto-correlated.

363 For the settings  $\alpha$ =0.01,  $\tau_{max}$ = 3 and using monthly data we obtain the CEN as in figure 5a. We 364 find evidence that Barents Kara sea ice concentrations (BK-SIC) have a negative effect on sea level pressure over the Ural Mountains region (Ural-SLP) with a time-delay of three months. 365 Thus, low sea ice in autumn can lead to increased surface pressure in winter. We also find a 366 367 positive link from Ural-SLP to v-flux with a delay of one month which means that higher surface pressure can increase the poleward heat flux, respectively the vertical wave activity. This is 368 consistent with the mechanisms proposed by Cohen et al. (2014) and Kim et al. (2014). 369 370 Moreover we can see in figure 5a, that increasing vertical wave activity induces a weakening of 371 the stratospheric polar vortex with a delay of one month. Hence, the CEN depicts the Troposphere-Stratosphere coupling described by Kim et al. (2014) and Cohen et al. (2014). We 372 373 also see a reverse relation from the Stratosphere into the Troposphere, whereby a weak polar

374 vortex (PoV) leads to increasing sea ice in the Barents and Kara Seas and to less vertical wave 375 activity. We find no causal link connecting a weak polar vortex to a negative AO. However, we 376 have a positive instantaneous link between them, which might indicate that this connection is 377 happening on a sub-monthly time-scale. In addition to the mechanisms involving the Stratosphere we also detect a direct positive link from BK-SIC to AO. Thus, we find that Barents 378 and Kara sea ice in fall induces a weakening of AO in winter without any stratospheric 379 380 connection. However, AO is also instantaneously related to sea level pressure in the Ural 381 Mountains region (Ural-SLP) with a negative sign which is in turn strongly positively related with sea level pressure in Siberia (Sib-SLP). Even though the instantaneous links provide no direction, 382 they are in accordance with the expectation that AO is negative when sea level pressure in the 383 Arctic is high. The same is true for the instantaneous link connecting Sib-SLP and Ural-SLP to 384 385 each other resp. to BK-SIC. In addition to the influence of Ural-SLP on PoV via v-flux we also find a weaker direct causal link between them with a delay of one month, suggesting that high sea 386 387 level pressure in Central Asia can induce a weakening of the polar vortex directly, or via 388 processes which are not part of the tested hypothesis. The positive instantaneous link between EA-snow and Sib-SLP is indicating that increasing snow cover in Eurasia is associated to a 389 strengthened Siberian High which is consistent with the hypothesis of Cohen et al. (2014). The 390 391 auto-regressive influence (with a time-lag of one month) is as expected especially high for BK-392 SIC and EA-snow and weaker for PoV and AO. Ural-SLP, Sib-SLP and v-flux are not significantly 393 causally influenced by their values in the months before.

We performed sensitivity analyses of the CEN to the parameter settings used and found the 394 395 detected links to be robust. We limit ourselves to analyzing only links which go back to late summer. Figure 5 shows the winter months CENs associated with different significance levels (a 396 = 0.01, 0.025, 0.05 in the rows) and for maximum time-lags of three and five months (columns). 397 398 Not surprisingly, the number of significant links increases when we increase  $\alpha$ , most of them 399 involving the two actors based on sea level pressure (Fig. 5b, e, c, f). Also links associated with 400 time-lags of more than three months (Fig. 5d-f) appear when increasing the maximum time-lag 401  $\tau_{max}$ , however only for larger  $\alpha$  values. We see that all links in Fig. 5a appear in all other graphs as well. For a significance level  $\alpha$ >0.01 (Fig. 5b, e, c, f), we see that decreasing sea ice 402

403 concentrations in the Barents and Kara Seas (BK-SIC) induce stronger sea level pressure over Siberia (Sib-SLP) with a lag of two months. This is in accordance with the mechanism described 404 405 by Kim et al. (2014). We also see for  $\alpha$ >0.01 that increasing snow cover in Eurasia (EA-snow) is 406 also instantaneously positively linked to surface pressure over the Ural Mountains region (Ural-SLP). For a longer time lag we find that EA-snow is negatively influencing sea level pressure in 407 the Ural Mountains region (Ural-SLP) with a delay of five months (Fig. 5e, f). For  $\alpha$ =0.05 we even 408 409 find some evidence that EA-snow can influence AO directly, and thus it seems again that 410 processes not involving the Stratosphere are present. Overall, the CEN structure as in figure 5a 411 appears for all tested parameters.

412 As explained in the method section, instantaneous links provide no information on the direction. To gain further information on the direction of those links and to further test the 413 414 robustness of our findings, we construct CENs also for half-monthly and quarter-monthly time-415 series (see Fig. 6a, b). Since the data sets are then two-times respectively four-times longer and consist of shorter time-steps we adjust our settings for the CEN-algorithm. In order to make the 416 results comparable with figure 5a we therefore double respectively quadruple  $\tau_{max}$  to refer to 417 the same time-shift. Since for higher time-resolutions more potential links are tested for 418 significance we adjust the  $\alpha$  value accordingly<sup>6</sup>. Comparing figure 5a with CENs based on half-419 monthly (Fig. 6a) and quarter-monthly (Fig. 6b) time series with the same maximum time-shift 420 421 of three months and an adjusted significance level  $\alpha$ =0.005625 for half-monthly and  $\alpha$ =0.003 for quarter-monthly data, we find a robust pattern of the involved causal processes. Especially the 422 Troposphere-Stratosphere connection is clearly visible in all CENs. For the CEN based on half-423 monthly data (Fig. 6a) the connection to vertical wave propagation (v-flux) is via the Siberian 424 region (Sib-SLP) whereby this region is directly influenced by the Ural Mountains (Ural-SLP) area. 425 On a quarter-monthly time-scale both regions directly influence v-flux which in turn influences 426 PoV (Fig. 6b). On the other hand we have a direct link from PoV to AO (Fig. 6b) in the quarter-427

<sup>&</sup>lt;sup>6</sup> If *n* denotes the number of actors, then  $N = n^2(\tau_{max} + 1) - n$  potential links are tested. Thus, N=189 (monthly), N=336 (half-monthly) and N=630 (quarter-monthly). To calculate the adapted  $\alpha$  we use a simple Bonferroni-correction and divide  $\alpha$ =0.01 by the multiplicity of the performed tests.

428 monthly based CEN, which indicates that a breakdown of the polar vortex causes a negative AO 429 on a weekly time-scale. Also, there are direct links connecting Ural-SLP to EA-snow, BK-SIC and 430 Sib-SLP, which shows that the Ural Mountains region has a strong influence on the surrounding 431 regions on sub-monthly time-scales, which is in accordance with the tested hypothesis. However, the strong instantaneous links between tropospheric based actors (AO, Ural-SLP, Sib-432 433 SLP and EA-snow) remain for all time-scales, indicating that those causal processes are occurring 434 on sub-weekly time-scales or are due to common drivers. The darker node-colors show that at 435 sub-monthly time scales auto-regressive processes become larger.

436 In summary, the CEN-algorithm provides robust results, whereby additional links can predominantly be explained by changing parameter settings and by the temporal resolution of 437 the underlying time-series. Barents and Kara sea ice (BK-SIC) is detected to play an important 438 439 role on winter-circulation, especially on the monthly time-scale (Fig. 5) being responsible for 440 changes in the pressure profile over the Ural Mountains region as well as by influencing AO 441 directly. Thus, mechanisms effecting AO not involving the Stratosphere seem to be important, too. We assume that other processes for example as described by Petoukhov and Semenov 442 443 [Petoukhov and Semenov, 2010] not represented by our choice of actors play a role, connecting Arctic sea ice and AO. As stated by Cohen and Kim, we find a connection of surface pressure 444 (Ural-SLP) and upward wave activity (v-flux) into the Stratosphere for all parameter settings and 445 446 time-scales (Fig. 5, 6). On lower time-scales we also have a direct link from Sib-SLP to v-flux (Fig. 6). These findings confirm the hypothesis that higher pressure over Central Asia leads to 447 increasing vertical wave activity into the Stratosphere [Cohen et al., 2014]. The Ural Mountains 448 449 region as a preferred location for atmospheric blocking [Wang et al., 2009] seems to play a central role for winter circulation, being linked to the tropospheric actors AO, BK-SIC, Sib-SLP 450 and EA-snow on all time-scales. Further, the region is responsible for coupling with the 451 Stratosphere (Fig. 5, 6). In this context, we expect that the link connecting Ural-SLP to PoV 452 453 directly, and not via v-flux, can at least partly be explained by hemispheric-wide averaging of the 454 actors v-flux and PoV (in contrast to the regional actor Ural-SLP). Additionally, it is possible that 455 a common driver not included in this analysis is responsible for this direct link. For example, tropical teleconnections like ENSO could influence both the Arctic Stratosphere and sea level 456

457 pressure in Central Asia [Butler et al., 2014]. Additionally, we find that the increased vertical wave activity can induce a weakening of the polar vortex (PoV), whereas PoV is positively 458 connected to surface AO (Fig. 5, 6). Thus, our findings are consistent with the Troposphere-459 460 Stratosphere-Troposphere mechanisms described by Cohen et al. (2014) and Kim et al. (2014). We also find a reverse connection, linking a weak polar vortex (PoV) to increasing Barents and 461 462 Kara sea ice and decreasing vertical wave activity (v-flux). This provides a negative feedback on a 463 time-scale of approximately one to two months. The role of Eurasia snow cover (EA-snow) 464 seems to be more complex. We find no evidence that late autumn snow fall in Eurasia influences the sea level pressure in Central Asia as proposed by Cohen et al. (2014). However, 465 we find that EA-snow is instantaneously linked to Sib-SLP with positive sign and for  $\alpha$ >0.01 also 466 to Ural-SLP (Fig. 5, 6). On a monthly time-scale we also have a direct negative link to Ural-SLP 467 468 (with a lag of five months) and for  $\alpha$ =0.05 also to AO (with a lag of two months). Overall our findings are less robust for EA-snow. 469

470

471

#### 472 **5. Conclusion**

473 In the context of hypothesis testing, we constructed *Causal Effect Networks* (CEN) in order to unravel causal relationships and their time delays between different actors of mid-latitude 474 475 winter circulation. We restricted ourselves to studying Arctic mechanisms, based on those 476 proposed by Kim et al. (2014) and Cohen et al. (2014). For each of the seven actors we 477 constructed one index at different temporal resolutions. CEN-construction was performed by first deriving a set of parents for each actor, consisting of the conditional dependent processes 478 479 (Step 1). Then those parents were used to estimate the strength and statistical significance of links employing linear regression models (Step 2). We only considered effects on winter 480 circulation and applied the method to monthly, half-monthly and quarter-monthly time-series. 481 482 We found that the method provides robust results for different values of the significance level  $\alpha$ 483 and maximum time delay  $\tau_{max}$  as well as for the considered range of temporal resolutions.

484 Figure 7 (respectively Fig. 5a) contains the most robust links on a monthly scale whereby results 485 are presented according to the approximate geographical location of the actors. Overall, our 486 findings are largely consistent with previously proposed hypotheses under consideration, 487 whereby especially Barents and Kara sea ice is detected to be an important external driver for winter circulation. Our CENs confirm the proposed Troposphere-Stratosphere coupling, which is 488 489 evident for all tested parameter settings. However, we also find a robust pattern indicating a 490 direct tropospheric connection of Barents and Kara sea ice and AO, such as for example 491 proposed by Petoukhov and Semenov (2010). The direct link connecting Ural-SLP to PoV might 492 be due to not considered tropical mechanisms influencing both the Stratosphere and sea level pressure in Eurasia as documented by Butler et al. (2014). The role of Eurasia snow cover is less 493 robust but seems to influence sea level pressure in Asia significantly. 494

Since the CEN-algorithm requires the choice of the free parameters  $\tau_{max}$  and  $\alpha$  and depends on 495 496 the temporal resolution of the underlying data, changing settings can produce different graphs. 497 However, by including sensitivity tests for different parameter settings and time scales, we 498 report robust results. Also, it should be noted that the CEN approach assumes stationary time-499 series. Long term trends or changing trends within the studied time period might affect the results [Overland and Wang, 2005] and require a careful treatment of the data. However, here 500 we only found a clear negative linear trend in the sea ice data. The causal interpretation of the 501 502 resulting CENs also depends on the choice of actors such that the inferred parents can still be 503 due to not-yet-included other variables. The challenge of how to choose adequate actors can also be assessed by different methods such as dimension reduction via Principal Component 504 Analysis [Runge et al., 2015]. Nonetheless, the CEN-algorithm is especially useful for testing 505 hypothesis if consistency of the data choice is assured. 506

The scope of this paper was to introduce and explain the CEN-algorithm and how it can be applied to address questions associated with teleconnections in the global climate system. In this context, CENs can overcome ambiguities of correlation analyses and provide a practical supplemental method to model experiments in order to test hypothesis. Moreover, CENs could be used also on model data to assess their validity. Here we limited ourselves to linear measurements, but CENs can also be constructed using non-parametric approaches, e.g., from

information theory [*Runge et al.*, 2012a, 2012b]. Further research should address the question
of how tropical mechanisms contribute to mid-latitude winter circulation [*Palmer*, 2014; *Trenberth et al.*, 2014] and also the different hypotheses related to summer circulation
[*Overland et al.*, 2012; *Coumou et al.*, 2014].

- 517
- 518

# 519 Acknowledgements

The authors thank the editor and three anonymous reviewers for constructive and insightful comments that helped to significantly improve the manuscript. The work was supported by the German Federal Ministry of Education and Research (grant no. 01LN1304A). J.R. received support by the German Federal Ministry of Science and Education (Young Investigators Group CoSy-CC<sup>2</sup>, grant no. 01LN1306A). J.F.D. thanks the Stordalen Foundation (via the Planetary Boundary Research Network PB.net) and the Earth League's EarthDoc program for financial support.

- 527
- 528

### 529 **REFERENCES**

Barnes, E. A., and J. A. Screen (2015), The impact of Arctic warming on the midlatitude jetstream: Can it? Has it? Will it?, *Wiley Interdiscip. Rev. Clim. Chang.*, *6*, 277–286,
doi:10.1002/wcc.337.

- Butler, A. H., L. M. Polvani, and C. Deser (2014), Separating the stratospheric and tropospheric
   pathways of El Niño–Southern Oscillation teleconnections, *Environ. Res. Lett.*, 9(2), 024014,
   doi:10.1088/1748-9326/9/2/024014.
- Cohen, J., M. Barlow, P. J. Kushner, and K. Saito (2007), Stratosphere–Troposphere Coupling and
   Links with Eurasian Land Surface Variability, *J. Clim., 20*(21), 5335–5343,
   doi:10.1175/2007JCLI1725.1.

539 540	Cohen, J., J. Jones, J. C. Furtado, and E. Tziperman (2013), Warm Arctic, Cold Continents, Oceanography, 26(4), 1–12, doi:10.5670/oceanog.2013.70.
541	Cohen, J., J. C. Furtado, J. Jones, M. Barlow, D. Whittleston, and D. Entekhabi (2014a), Linking
542	Siberian snow cover to precursors of stratospheric variability, J. Clim., 27(14), 5422–5432,
543	doi:10.1175/JCLI-D-13-00779.1.
544	Cohen, J. et al. (2014b), Recent Arctic amplification and extreme mid-latitude weather, Nat.
545	<i>Geosci.</i> , 7, 627–637, doi:10.1038/ngeo2234.
546	Coumou, D., V. Petoukhov, S. Rahmstorf, S. Petri, and H. J. Schellnhuber (2014), Quasi-resonant
547	circulation regimes and hemispheric synchronization of extreme weather in boreal
548	summer, Proc. Natl. Acad. Sci., 111(34), 12331–12336, doi:10.1073/pnas.1412797111.
549	Deng, Y., and I. Ebert-Uphoff (2014), Weakening of atmospheric information flow in a warming
550	climate in the Community Climate System Model, <i>Geophys. Res. Lett.</i> , 41(1), 193–200,
551	doi:10.1002/2013GL058646.
552	Deser, C., R. Tomas, M. Alexander, and D. Lawrence (2010), The Seasonal Atmospheric Response
553	to Projected Arctic Sea Ice Loss in the Late Twenty-First Century, J. Clim., 23, 333–351,
554	doi:10.1175/2009JCLI3053.1.
555	Ebert-Uphoff, I., and Y. Deng (2012), Causal discovery for climate research using graphical
556	models, <i>J. Clim.</i> , <i>25</i> (17), 5648–5665, doi:10.1175/JCLI-D-11-00387.1.
557	Francis, J. A., and S. J. Vavrus (2012), Evidence linking Arctic amplification to extreme weather in
558	mid-latitudes, Geophys. Res. Lett., 39(6), doi:10.1029/2012GL051000.
559	Furtado, J. C., J. L. Cohen, A. H. Butler, E. E. Riddle, and A. Kumar (2015), Eurasian snow cover
560	variability and links to winter climate in the CMIP5 models, <i>Clim. Dyn.</i> , 45(9),
561	doi:10.1007/s00382-015-2494-4.
562	Gray, S. L., C. M. Dunning, J. Methven, G. Masato, and J. M. Chagnon (2014), Systematic model
563	forecast error in Rossby wave structure, <i>Geophys. Res. Lett.</i> , 41(8), 2979–2987,
564	doi:10.1002/2014GL059282.
565	Handorf, D., R. Jaiser, K. Dethloff, A. Rinke, and J. Cohen (2015), Impacts of Arctic sea-ice and
566	continental snow-cover changes on atmospheric winter teleconnections, Geophys. Res.
567	<i>Lett., 42</i> (7), doi:10.1002/2015GL063203.

- Jaiser, R., K. Dethloff, D. Handorf, A. Rinke, and J. Cohen (2012), Impact of sea ice cover changes
   on the Northern Hemisphere atmospheric winter circulation, *Tellus A*, *64*,
   doi:10.3402/tellusa.v64i0.11595.
- Kim, B.-M., S.-W. Son, S.-K. Min, J.-H. Jeong, S.-J. Kim, X. Zhang, T. Shim, and J.-H. Yoon (2014),
  Weakening of the stratospheric polar vortex by Arctic sea-ice loss, *Nat. Commun.*, *5*,
  doi:10.1038/ncomms5646.
- Manzini, E. et al. (2014), Northern winter climate change: Assessment of uncertainty in CMIP5
  projections related to stratosphere-troposphere coupling, *J. Geophys. Res. Atmos.*, *119*(13),
  7979–7998, doi:10.1002/2013JD021403.
- 577Overland, J. E., and M. Wang (2005), The Arctic climate paradox: The recent decrease of the578Arctic Oscillation, Geophys. Res. Lett., 32(6), L06701, doi:10.1029/2004GL021752.
- 579 Overland, J. E., J. A. Francis, E. Hanna, and M. Wang (2012), The recent shift in early summer 580 Arctic atmospheric circulation, *Geophys. Res. Lett.*, *39*(19), doi:10.1029/2012GL053268.
- Palmer, T. (2014), Record-breaking winters and global climate change, *Science (80-. ).*,
   344(6186), 803–804, doi:10.1126/science.1255147.
- Palmer, T. N., and D. A. Mansfeld (1984), Response of 2 atmospheric general circulation models
   to sea-surface temperature anomalies in the tropical East and West Pacific, *Nature*, *310*,
   485–485, doi:10.1038/310483a0.
- Palmer, T. N., and J. A. Owen (1986), A possible relationship between some "severe" winters in
  North America and enhanced convective activity over the tropical west Pacific, *Mon. Weather Rev.*, 114(3), 648–651, doi:10.1175/15200493(1986)114<0648:APRBSW>2.0.CO;2.
- 590 Pearl, J. (2013), Linear Models: A Useful "Microscope" for Causal Analysis, *J. Causal Inference*,
   591 1(1), 155–170, doi:10.1515/jci-2013-0003.
- Petoukhov, V., and V. A. Semenov (2010), A link between reduced Barents-Kara sea ice and cold
  winter extremes over northern continents, *J. Geophys. Res. Atmos.*, 115(D21),
  doi:10.1029/2009JD013568.
- Polvani, L. M., and D. W. Waugh (2004), Upward Wave Activity Flux as a Precursor to Extreme
   Stratospheric Events and Subsequent Anomalous Surface Weather Regimes, *J. Clim.*,
   17(18), 3548–3554, doi:10.1175/1520-0442(2004)017<3548:UWAFAA>2.0.CO;2.

- Robinson, D. A., K. F. Dewey, and R. R. Heim (1993), Global Snow Cover Monitoring: An Update,
   *Bull. Am. Meteorol. Soc.*, 74(9), 1689–1696, doi:10.1175/1520 0477(1993)074<1689:GSCMAU>2.0.CO;2.
- Runge, J., J. Heitzig, V. Petoukhov, and J. Kurths (2012a), Escaping the curse of dimensionality in
  estimating multivariate transfer entropy, *Phys. Rev. Lett.*, *108*(25), 258701,
  doi:10.1103/PhysRevLett.108.258701.
- Runge, J., J. Heitzig, N. Marwan, and J. Kurths (2012b), Quantifying causal coupling strength: A
   lag-specific measure for multivariate time series related to transfer entropy, *Phys. Rev. E*,
   *86*(6), 61121, doi:10.1103/PhysRevE.86.061121.
- Runge, J., V. Petoukhov, and J. Kurths (2014), Quantifying the strength and delay of climatic
   interactions: The ambiguities of cross correlation and a novel measure based on graphical
   models, J. Clim., 27(2), 720–739, doi:10.1175/JCLI-D-13-00159.1.
- Runge, J., V. Petoukhov, J. F. Donges, J. Hlinka, N. Jajcay, M. Vejmelka, D. Hartman, N. Marwan,
  and M. Palu<sup>\*</sup> (2015), Identifying causal gateways and mediators in complex spatiotemporal systems, *Nat. Commun.*, *8*, doi:10.1038/ncomms9502.
- Schleussner, C. F., J. Runge, J. Lehmann, and A. Levermann (2014), The role of the North Atlantic
   overturning and deep ocean for multi-decadal global-mean-temperature variability, *Earth Syst. Dyn.*, 5(1), 103–115, doi:10.5194/esd-5-103-2014.
- Screen, J. A., and I. Simmonds (2014), Amplified mid-latitude planetary waves favour particular
   regional weather extremes, *Nat. Clim. Chang.*, doi:10.1038/NCLIMATE2271.
- 618 Spirtes, P., C. Glymour, and R. Scheines (2000), *Causation, Prediction, and Search*.
- Thompson, D. W. J. (2001), Regional Climate Impacts of the Northern Hemisphere Annular
  Mode, *Science (80-. ).*, 293(5527), 85–89, doi:10.1126/science.1058958.
- Tremblay, L. B., M. M. Holland, I. V. Gorodetskaya, and G. A. Schmidt (2007), An Ice-Free Arctic?
  Opportunities for Computational Science, *Comput. Sci. Eng.*, *9*(3), 65–74,
  doi:10.1109/MCSE.2007.45.
- Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski (1998),
  Progress during TOGA in understanding and modeling global teleconnections associated
  with tropical sea surface temperatures, *J. Geophys. Res.*, *103*(C7), 14291,
  doi:10.1029/97JC01444.

- Trenberth, K. E., J. T. Fasullo, G. Branstator, and A. S. Phillips (2014), Seasonal aspects of the
  recent pause in surface warming, *Nat. Clim. Chang.*, 4(10), 911–916,
  doi:10.1038/nclimate2341.
- Wang, L., W. Chen, W. Zhou, J. C. L. Chan, D. Barriopedro, and R. Huang (2009), Effect of the
  climate shift around mid 1970s on the relationship between wintertime Ural blocking
  circulation and East Asian climate, *Int. J. Climatol.*, *30*, 153–158, doi:10.1002/joc.1876.
- 634

# 635 Tables

Abbreviation	Actor	Variable/Unit	Region (Level)				
BK-SIC	Barents Kara sea	Sea ice area fraction	70 °- 80°N, 30°- 105°E				
	ice						
EA-snow	Eurasia snow	snow covered area	40° - 80°N, 30°-180°E				
	cover	fraction					
AO	Arctic Oscillation	Geopotential height in m	20° - 90°N (1000 mb)				
	Index						
v-flux	Vertical wave	Pole-ward eddy heat flux	45° - 75°N (100 mb)				
	propagation	v*T* in K*m/s					
PoV	Polar Vortex	Geopotential height in m	65° - 90°N (10 - 100 mb)				
Sib-SLP	Siberian High	Sea level pressure in mb	40° - 65°N, 85° - 120°E				
Ural-SLP	Ural Mountains	Sea level pressure in mb	45° - 70°N, 40° - 85°E				
	sea level						
	pressure						

636

637 Table 1: Table of variables and regions of every considered actor

Actor	Parents ${m {\cal P}}$				
AO	AO <sub>t-1</sub> , BK-SIC <sub>t-2</sub>				
BK-SIC	BK-SIC <sub>t-1</sub> , PoV <sub>t-2</sub>				
EA-snow	EA-snow <sub>t-1</sub>				
v-flux	PoV <sub>t-1</sub>				
PoV	v-flux <sub>t-1</sub> , Ural-SLP <sub>t-1</sub> , PoV <sub>t-1</sub>				

Sib-SLP	None
Ural-SLP	BK-SIC <sub>t-3</sub>

639

- 640 Table 2: Table of parent processes of each actor for winter (DJF) data and with the settings
- 641  $\alpha$ =0.01 and  $\tau_{max}$ =3. The subscript denotes the time lag in months. The parent processes are then
- 642 used in the second step of the CEN-algorithm in order to quantify the link strength in terms of
- 643 *linear regression coefficients.*

# 644 **FIGURE CAPTIONS**

- Figure 1: Schematic picture of different time scales, whereby each box indicates one time-step.
  Quarter-monthly time series (bottom row) consists of four times respectively two times more
  data points than monthly (top row) and half-monthly (middle row) time series.
- Figure 2: Monthly time-series of all calendar months of climatological anomalies of each actorfrom 01/1979-12/2014.
- Figure 3: Possible scenarios leading to a correlation without a direct causation between process
  X and Y: a) inflated correlation due to autocorrelation b) indirect chain via Z c) common driver Z.
- Figure 4: Schematic picture of time series considered to measure influence of actor X on winter
   polar vortex (PoV) with a time-lag τ=1, whereby the time-series only consist of the dark grey
   boxes.
- 655Figure 5: CENs of actors of winter (DJF) circulation based upon monthly mean data. With a656maximum time-lag of  $\tau_{max}$ = 3 (a, b, c) and  $\tau_{max}$ = 5 (d, e, f) and with significance level  $\alpha$ =0.01 (a,657d),  $\alpha$ =0.025 (b, e) and  $\alpha$ =0.05 (c, f).
- 658Figure 6: CEN of actors of winter (DJF) circulation for a) half-monthly data with  $\tau_{max}$  =6 and659α=0.005625 and b) quarter-monthly data with  $\tau_{max}$  =12 and α=0.003.
- Figure 7: Same as Fig. 5a but the network is embedded in a schematic projection of the earth
   and the atmosphere. The regional actors BK-SIC, Ural-SLP, Sib-SLP and EA-snow are presented
   according to their approximate geographical location and the hemispheric actors AO, v-flux and
   PoV are presented according to their approximate latitude and pressure levels. See Tab. 1 for
   the exact coordinates of all actors.
- 665
- 666

monthly	X		X <sub>t</sub>		X <sub>t+1</sub>			X <sub>t+2</sub>				
half-monthly	X <sub>t</sub>		X <sub>t+1</sub>		X <sub>t+2</sub>		X <sub>t+3</sub>		X <sub>t+4</sub>		X <sub>t+5</sub>	
quater-monthly	X <sub>t</sub>	X <sub>t+1</sub>	X <sub>t+2</sub>	X <sub>t+3</sub>	X <sub>t+4</sub>	X <sub>t+5</sub>	X <sub>t+6</sub>	X <sub>t+7</sub>	X <sub>t+8</sub>	X <sub>t+9</sub>	X <sub>t+10</sub>	<b>X</b> <sub>t+11</sub>

Figure 1: Schematic picture of different time scales, whereby each box indicates one time-step. Quarter-monthly
time series (bottom row) consists of four times respectively two times more data points than monthly (top row) and
half-monthly (middle row) time series.



*Figure 2:* Monthly time-series of all calendar months of climatological anomalies of each actor from 01/1979677 12/2014.



687 with a time-lag  $\tau$ =1, whereby the time-series only consist of the dark grey boxes.









Jral-

SLF



beta-coefficients





- **Figure 6:** CEN of actors of winter (DJF) circulation for a) half-monthly data with  $\tau_{max} = 6$  and  $\alpha = 0.005625$  and b) 697 quarter-monthly data with  $\tau_{max} = 12$  and  $\alpha = 0.003$ .



**Figure 7:** Same as Fig. 5a but the network is embedded in a schematic projection of the earth and the atmosphere.

701 The regional actors BK-SIC, Ural-SLP, Sib-SLP and EA-snow are presented according to their approximate

702 geographical location and the hemispheric actors AO, v-flux and PoV are presented according to their approximate

703 *latitude and pressure levels. See Tab. 1 for the exact coordinates of all actors.*