

## Using Bayesian Networks to Investigate the Influence of Subseasonal Arctic Variability on Midlatitude North Atlantic Circulation

NATHANAEL HARWOOD,<sup>a</sup> RICHARD HALL,<sup>b</sup> GIORGIA DI CAPUA,<sup>c,d</sup> ANDREW RUSSELL,<sup>e,f</sup> AND ALLAN TUCKER<sup>g</sup>

<sup>a</sup> *Institute of Environment, Health and Societies, Brunel University London, Uxbridge, United Kingdom*

<sup>b</sup> *School of Geography, University of Lincoln, Lincoln, United Kingdom*

<sup>c</sup> *Potsdam Institute for Climate Impact Research, Potsdam, Germany*

<sup>d</sup> *Institute for Environmental Studies, Vrije Universiteit Amsterdam, Amsterdam, Netherlands*

<sup>e</sup> *Committee on Climate Change, London, United Kingdom*

<sup>f</sup> *Department of Life Sciences, Brunel University London, Uxbridge, United Kingdom*

<sup>g</sup> *Department of Computer Science, Brunel University London, Uxbridge, United Kingdom*

(Manuscript received 18 May 2020, in final form 11 December 2020)

**ABSTRACT:** Recent enhanced warming and sea ice depletion in the Arctic have been put forward as potential drivers of severe weather in the midlatitudes. Evidence of a link between Arctic warming and midlatitude atmospheric circulation is growing, but the role of Arctic processes relative to other drivers remains unknown. Arctic–midlatitude connections in the North Atlantic region are particularly complex but important due to the frequent occurrence of severe winters in recent decades. Here, dynamic Bayesian networks with hidden variables are introduced to the field to assess their suitability for teleconnection analyses. Climate networks are constructed to analyze North Atlantic circulation variability at 5-day to monthly time scales during the winter months of the years 1981–2018. The inclusion of a number of Arctic, midlatitude, and tropical variables allows for an investigation into the relative role of Arctic influence compared to internal atmospheric variability and other remote drivers. A robust covariability between regions of amplified Arctic warming and two definitions of midlatitude circulation is found to occur entirely within winter at submonthly time scales. Hidden variables incorporated in networks represent two distinct modes of stratospheric polar vortex variability, capturing a periodic shift between average conditions and slower anomalous flow. The influence of the Barents–Kara Seas region on the North Atlantic Oscillation is found to be the strongest link at 5- and 10-day averages, while the stratospheric polar vortex strongly influences jet variability on monthly time scales.


**KEYWORDS:** Arctic; North Atlantic Ocean; Atmospheric circulation; Teleconnections; Algorithms; Machine learning

### 1. Introduction

The Arctic has warmed at more than twice the speed of the global average since the mid-twentieth century, and at more than 6 times the average between 1998 and 2012 (Huang et al. 2017). This process, known as Arctic amplification (AA), is particularly strong in boreal winter. Concurrently, the heavily populated regions of western Europe and the eastern coast of the United States have experienced several cold outbreaks during the winters of recent years. A multitude of studies are supportive of a link between AA and midlatitude circulation (e.g., Samarasinghe et al. 2019; Blackport and Screen 2019), cold-air outbreaks (e.g., Kim et al. 2014; Chen and Luo 2017), and the North Atlantic Oscillation (NAO) (e.g., Pedersen et al. 2016). The strength of the connection between AA and midlatitude circulation remains uncertain; the importance of the linkage relative to other factors such as internal midlatitude variability, tropical forcing, and the stratospheric polar vortex currently represents a striking gap in our knowledge. In addition, the regional and intermittent nature of linkages means that

direct effect attribution studies using Arctic processes like sea ice loss will not provide a way forward for the research area (Overland et al. 2016). Establishing the Arctic's impact on jet stream variability, relative to these factors, is a complex but essential research endeavor as observational analyses often consider these factors in isolation.

Warming of tropical oceans, in particular with above-average sea surface temperatures (SSTs) in the Pacific, is known to impact midlatitude flow through intense convection and latent heat release, which generate planetary-scale Rossby waves (Trenberth et al. 1998). El Niño–Southern Oscillation (ENSO) (Scaife et al. 2017a) and other tropical Rossby wave source regions (Scaife et al. 2017b) provide predictive skill in seasonal midlatitude circulation forecasting, and ENSO has a stronger role in winter. Arctic sea ice concentration has also been put forward as an important driver of meridional jet stream configurations (Francis and Vavrus 2015). Low sea ice concentrations expose more open water which absorbs additional heat, leading to a greater exchange of heat and moisture between the ocean and atmosphere in autumn and thus an anomalously warm Arctic. More recently, studies have emphasized the importance of considering warming over the

 Denotes content that is immediately available upon publication as open access.

Corresponding author: N. Harwood, nathanael.harwood@brunel.ac.uk

DOI: 10.1175/JCLI-D-20-0369.1

© 2021 American Meteorological Society



This article is licensed under a Creative Commons Attribution 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

TABLE 1. Variables used, along with their abbreviation, source, and unit of measurement.

Variable	Abbreviation	Source	Unit
Arctic 850-hPa temperature	Arctic	ERA-Interim	Temperature (°C)
Barents–Kara Seas 850-hPa temperature	$T_{BK}$	ERA-Interim	Temperature (°C)
North Atlantic 850-hPa temperature	$T_{NA}$	ERA-Interim	Temperature (°C)
Stratospheric polar vortex	PoV	ERA-Interim	Geopotential height (m)
Jet latitude	Jet	ERA-Interim	Degrees (°)
Meandering Index	MI	ERA-Interim	—
North Atlantic Oscillation	NAO	NOAA	Sea level pressure
Madden–Julian oscillation	MJO	NOAA	MJO amplitude
El Niño–Southern Oscillation	ENSO	NOAA	Sea surface temperature

Arctic as opposed to sea ice trends (Barnes and Simpson 2017; Cohen et al. 2018a), chiefly because AA is the result of a complex combination of local sensible heat fluxes, evaporation, and the remote transport of heat and moisture from lower latitudes (Cohen et al. 2018a). Anomalous midlatitude circulation drives intrusions of warm, moist air into the Arctic that play an important role in the feedback between Arctic warming and sea ice retreat (Rigor et al. 2002; Zhang et al. 2008; Woods et al. 2013; Liu and Barnes 2015; Woods and Caballero 2016; Olonschek et al. 2019), with transport thought to be particularly pronounced along the North Atlantic pathway due to Atlantic blocking deflecting midlatitude cyclones poleward (Kim et al. 2017; Yang and Magnusdottir 2017). Intrusion events have measurable impacts on sea ice within several days of an event (Kapsch et al. 2016), in turn strengthening AA and any potential midlatitude feedback.

Sea ice loss in the Barents–Kara Seas, a region of pronounced variability, can expand and intensify the Siberian high through the initiation of vertically propagating Rossby waves and the disruption of the stratospheric polar vortex (Kim et al. 2014; Kretschmer et al. 2016). The southward flow of Arctic air that results from this has been associated with intensified cold events over East Asia (Overland et al. 2015). Over North America meanwhile, AA processes increase the likelihood of Alaskan and Greenland blocking events, which can reinforce and prolong cold events (Chen and Luo 2017; Overland and Wang 2018). It is becoming increasingly clear that AA has an impact on the strength and position of the North Atlantic eddy-driven jet in winter (Barnes and Simpson 2017; Blackport and Screen 2019), although the existence of numerous potential drivers of the jet itself (Hall et al. 2015; Smith et al. 2016) means that the importance of the AA contribution remains unresolved.

As robust linkages are difficult to detect, novel statistical analysis of the observational climate record has been put forward as a potential way to move the Arctic–midlatitude field forward (Overland et al. 2016; Kretschmer et al. 2016; Cohen et al. 2018a), acting as a supportive tool for large coordinated modeling projects. Studies that rely on correlation analysis are subject to autocorrelation bias, as well as misleading results due to indirect links or common drivers of correlated variables that were unaccounted for in the analysis (Runge et al. 2014). Such linear relationships are also directionless and so offer less information than graphical models. Atmospheric model studies, while well regarded as tools for identifying causal linkages, are not immune from potential shortcomings: they may not accurately represent ocean–atmosphere coupling in the Arctic

(Cohen et al. 2018a), may respond too weakly to sea ice forcing (Screen et al. 2018; Mori et al. 2019), may underperform in terms of stratosphere–troposphere coupling (Zhang et al. 2018), and focus on the impact of sea ice removal, which may not capture the complex intermittencies thought to define Arctic–midlatitude linkages (Overland and Wang 2018).

The analysis presented here applies dynamic Bayesian networks with hidden variables to the North Atlantic and European midlatitude circulation research area. Structure-learning algorithms are employed to identify regions of AA that might influence winter jet stream variability. A number of Arctic, midlatitude, and tropical variables are included to investigate the relative role of AA as a driver compared to internal atmospheric variability and other remote forcings. Other graphical model approaches focus either entirely on potential Arctic drivers of midlatitude circulation responses (Kretschmer et al. 2016; Barnes and Simpson 2017; Samarasinghe et al. 2019) or on tropical teleconnections like the MJO–NAO link (Barnes et al. 2019). The aim of this study is to establish how effective dynamic Bayesian networks with structure learning algorithms are for investigating this research area, and to measure the impact of hidden variables on model accuracy which is a priority due to the low signal-to-noise ratio of AA linkages and their intermittent nature (Overland et al. 2016). We demonstrate a feedback relationship between North Atlantic midlatitude circulation and two important regions of Arctic warming occurring in winter at submonthly time scales. Finally, the implications for further study are discussed.

## 2. Data

A number of climatological variables that are understood to have an influence on European midlatitude weather during the winter months are included in this analysis (Table 1). Four nonoverlapping time average resolutions consisting of 5-day, 10-day, 15-day, and monthly averages were used to ensure a robust set of conclusions that include variables that may act at a range of time scales. Note that 29 February was removed from each leap year, allowing for a total of 90 days in each year using only the winter months [December to February (DJF)] from the years 1981 to 2018.

To investigate the relative impact of the Arctic on midlatitudes, tropical indices formed part of a network of nonlocal drivers (Fig. 1). Both ENSO (Brönnimann 2007) and the Madden–Julian oscillation (MJO) (Lin et al. 2015) have been

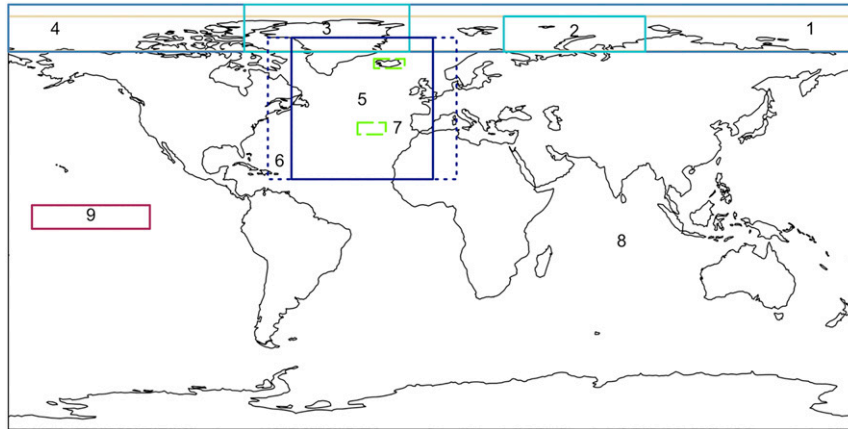


FIG. 1. Map of all variables used in DBN experiments: 1—Arctic (dark blue), 2—Barents–Kara Seas ( $T_{BK}$ ), 3—North Atlantic ( $T_{NA}$ ), 4—stratospheric polar vortex (PoV; brown), 5—jet latitude (Jet), 6—Meandering Index (MI; dotted blue), 7—North Atlantic Oscillation (NAO; dashed green boxes), 8—Madden–Julian oscillation (MJO), and 9—El Niño–Southern Oscillation (ENSO).

found to explain significant amounts of North Atlantic Oscillation variability during wintertime through Rossby wave excitement (Trenberth et al. 1998). A standardized Niño-3.4 index is included, calculated using data from HadISST1 (Rayner et al. 2003), retrieved from the KNMI climate explorer (<https://climexp.knmi.nl/>). The Real-Time Multivariate MJO (RMM) index devised by Wheeler and Hendon (2004, available at <http://www.bom.gov.au/climate/mjo/>) is generated by identifying the principal components (PCs) of two combined empirical orthogonal functions (EOFs) of equatorially averaged 850-hPa zonal wind, 200-hPa zonal wind, and outgoing longwave radiation (OLR) anomalies. The two PC time series (RMM1 and RMM2) are used to define the MJO amplitude, equal to  $\sqrt{(\text{RMM1}^2 + \text{RMM2}^2)}$ . The PC-based NAO index (Hurrell 1995), based on the difference in surface sea level pressure (SLP) between the subtropical high and the subpolar low, provides a secondary indicator of North Atlantic atmospheric variability and was also obtained from KNMI's climate explorer.

To measure the impact of Arctic variability, near-surface 850-hPa temperature from ECMWF interim reanalysis (ERA-Interim) global atmospheric dataset (Dee et al. 2011) was included in the networks, available at <https://apps.ecmwf.int/datasets/data/interim-full-daily/>. The 850-hPa temperature is used in place of sea ice concentration to capture the full effect of AA. Arctic near-surface warming is thought to be a better indicator of AA than sea ice variability because the AA signal is made up of a number of factors that include sea ice retreat, and the impact of heat and moisture transport from lower latitudes is fully accounted for (Cohen et al. 2018a). All data were compiled as  $1^\circ \times 1^\circ$  gridded spatiotemporal data and then processed. The entire Arctic region was prepared as well as important subsections as regional sea ice loss is known to be an important factor in midlatitude circulation responses (Pedersen et al. 2016; Screen 2017). Three regions of Arctic 850-hPa temperature were therefore used: the Arctic ( $70^\circ$ – $90^\circ$ N,

$180^\circ$ W– $180^\circ$ E), the North Atlantic sector ( $T_{NA}$ ) ( $70^\circ$ – $90^\circ$ N,  $10^\circ$ – $80^\circ$ W), and the Barents–Kara Seas region ( $T_{BK}$ ) ( $70^\circ$ – $85^\circ$ N,  $30^\circ$ – $90^\circ$ E). These capture three key areas of sea ice concentration loss: Baffin Bay, the east coast of Greenland, and the Barents–Kara Seas (see Fig. 2 in Overland and Wang 2018). An area-weighted spatial average was taken over these regions, and anomalies were calculated from the resulting univariate time series by subtracting each time step of the multiyear mean (1981–2018) from the value of the matching time step. This was done for each time resolution, with each then detrended to create a time series for the variable.

As a proxy for stratospheric variability, the polar vortex (PoV) was included to examine the impact that stratospheric circulation might have on tropospheric midlatitude circulation. The PoV index uses ERA-Interim geopotential height anomalies from the Arctic region ( $65^\circ$ – $90^\circ$ N,  $180^\circ$ W– $180^\circ$ E) averaged over six pressure levels from 10 to 100 hPa, with the resulting time series used to create anomalies at each time average and detrended. As such, the PoV opens up the potential for linkages through tropospheric and stratospheric pathways, and allows for comparison with previous studies that identify a link between the PoV and midlatitude circulation (Kim et al. 2014; Kretschmer et al. 2016). Barnes et al. (2019) found that the MJO signal via the stratosphere was sensitive to the level at which the PoV was defined. To account for this, a second PoV index was constructed using the 100-hPa level only, and this made no difference to the MJO results in this work.

Two metrics were used to represent North Atlantic midlatitude circulation in the networks. The first, jet latitude, was calculated from ERA-Interim data and defined over the region  $16^\circ$ – $76^\circ$ N,  $0^\circ$ – $60^\circ$ W. Jet latitude was determined using the approach taken by Woollings et al. (2010): zonal winds were height-averaged over 900 to 700 hPa and filtered with a 10-day Lanczos low-pass filter using a 61-day window to ensure that synoptic-scale variability is excluded. The use of lower-level winds isolates the eddy-driven jet as the data are not contaminated

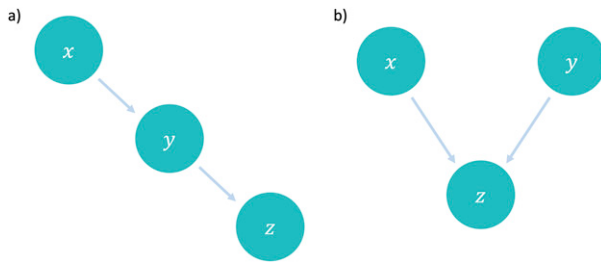


FIG. 2. Example DAGs: (a)  $x$  and  $z$  are conditionally independent given  $y$  therefore no arc between  $x$  and  $z$  is needed or parameterized, and (b) an example of a collider structure, where node  $z$  has two parent nodes,  $x$  and  $y$ .

with the signal of the subtropical jet. Jet latitude output is consistent across pressure levels, and many jet-focused studies have made use of this metric (Hall et al. 2017; Barnes and Simpson 2017; Samarasinghe et al. 2019).

The second metric of jet variability is the Meandering Index (MI), a measure of tropospheric circulation variability that uses geopotential height contours to capture the maximum waviness at each time step, taking into account the full spatial position of each contour (Di Capua and Coumou 2016). The input data used was ERA-Interim 500-hPa geopotential height (gph) for the region 16°–76°N, 10°E–70°W. The MI calculates the length of each isohypse (gph contour) on a 2D grid, which is then normalized to Earth’s circumference at 60°N. The maximum value of this calculation on the vertical profile (e.g., from 4800 to 5600 m) is then taken as the MI, allowing for accurate differentiation between strongly meridional (north–south) deviations and consistently zonal configurations. A full description of the MI can be found in Di Capua and Coumou (2016). The following analysis included both the jet latitude and MI to ensure linkages revealed by the dynamic Bayesian network technique were robust to the use of different metrics. The two metrics also describe different aspects of midlatitude circulation; the jet latitude by definition describes the jet’s latitudinal position on a given day, while the MI focuses on waviness of the middle troposphere, and thus may reveal the waviness of the jet but cannot be used as an indicator of the jet core location.

All data were standardized as a final stage of data preparation, so that each variable had a mean of 0 and a standard deviation of 1. This was done to maximize model accuracy in the parameter learning stage by giving all variables equal means and similar ranges.

### 3. Methods

Graphical models provide an excellent tool for examining relationships between variables of climatological importance. Bayesian networks (BNs) are a form of probabilistic graphical models that provide a useful mechanism for statistically modeling real world relationships, allowing for their visualization in a so-called graph, or network. Graphical model approaches in the climate sciences have included the use of BNs (Ebert-Uphoff and Deng 2012a,b) and causal effect networks (Kretschmer et al. 2016, 2017; Runge 2018; Di Capua et al. 2019).

BNs encode a joint probability distribution, whereby probabilities are assigned for all possible outcomes over a set of random variables taken as input (Friedman et al. 2000). BNs accomplish this by constructing directed acyclic graphs (DAGs) which exploit conditional independence relationships; in effect, a variable that is conditionally independent from another in the graph structure does not need to be parameterized (Fig. 2a), resulting in a significant boost to computational efficiency. Two variables in the graph are conditionally independent if knowing the state of one event does not change the probability of the second. Each variable has a conditional probability distribution (CPD) which encodes the probability of observing values given the values of its “parents” (i.e., the variables on which it is conditionally dependent). A BN factorizes all the CPDs in a joint distribution:

$$p(x_1 \dots x_n) = \prod_{i=1}^n p[x_i | \text{pa}(x_i)], \quad (1)$$

where  $n$  is the total number of variables and  $\text{pa}(x_i)$  denotes the parent set of  $x_i$ .

The DAG graphically represents the BN with nodes (the variables) and conditionally dependent relationships indicated with edges (links between the variables). When a node has two or more parents (a “collider” structure; Fig. 2b), an increase in probability for parent  $A$  that results in a decrease in another  $B$  can be said to “explain away” the likelihood of  $B$  being a driver. As an example, strongly anomalous values for the PoV in a given time step might explain away the need for the  $T_{\text{BK}}$  and  $T_{\text{NA}}$  regions of the Arctic as drivers of jet variability.

A dynamic Bayesian network (DBN), the approach used in this study, is an extension of a BN over time. A DBN makes use of two time slices ( $t - 1, t, \dots$ ) to “unroll” a DBN into  $T$  time slices (Fig. 3), such that model structure and parameters do not change over time and the model stays time invariant (Murphy 2001a). DBNs represent a *Markov process* because the state of a system  $t$  depends only on the preceding time step and state at  $t - 1$  (Mihajlovic and Petkovic 2001). An advantage of BNs is that they can be used to model observed and unobserved data, as hidden nodes can be inferred from the values of the observed nodes.

A hidden variable (HV) can be used to capture the underlying state of a time series or represent a variable of interest to the network that cannot be directly observed (Murphy 2012). Hidden variables may represent something of importance theoretically to the modeled system, or a process or driver that shares interdependencies with the variables but was not explicitly constrained within the model structure for one reason or another (Trifonova et al. 2017). This can occur when no data exist or when the model approach dictates the exclusion of system components; for example, a model where a set of symptoms is observed, but the disease is unknown (Murphy 2012). Here, discrete hidden variables with three possible states are parameterized to identify state switches in observed climate data. Studying the behavior of teleconnections in different atmospheric states is important as many are state-dependent, meaning that the background state of the atmosphere can determine whether a signal (from a driver like the MJO) can propagate to the midlatitudes or whether it acts to negate or diminish it (e.g., Barnes et al. 2019).



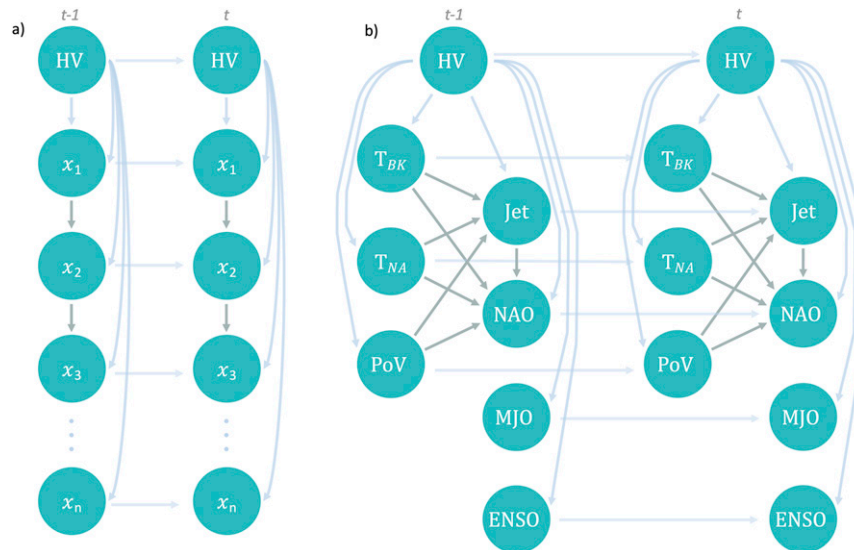


FIG. 3. Example DAGs. (a) A DBN with a single hidden variable unrolled over two time slices with autoregressive links between all variables. Gray arcs in the “intra” (within a time slice) graph represent links between observed variables, with autoregressive links shown in the “inter” (between time slices) arcs. (b) As in (a), but for a summarized graph of the 5-day jet HV DBN used in this study.

Hidden variables encoded within networks can point to any number of observed variables. The hidden variable then reflects the changes in system interactions between the observed nodes it is linked to, as its value is inferred to maximize the fit (log-likelihood) of the model to the data. Graphical models that incorporate hidden variables may result in structures that are significantly more similar to the climate system we are trying to model; simpler models are learned and are less prone to overfitting while being more efficient for inference (Tucker and Liu 2004). Given the challenges associated with identifying AA–midlatitude linkages due to noisy internal dynamics and the time-constrained nature of AA processes, potential improvements in model accuracy make hidden variables worth consideration. Graphical models with hidden variables inferred from observed data have been largely untouched in climate science studies, but their capabilities have been explored and proven in ecological system analyses (Trifonova et al. 2015, 2017, 2019; Uusitalo et al. 2018).

#### a. Structure learning

The PC algorithm, developed by and named after the first initials of Peter Spirtes and Clark Glymour (Spirtes and Glymour 1991), is a simple but effective constraint-based structure-learning algorithm. It connects all nodes in a network initially with undirected edges, and iteratively deleting edges by taking a pair of nodes ( $X, Y$ ) and trying to find a set of nodes  $S$  (exclusive of nodes  $X$  and  $Y$ ) so that  $X$  and  $Y$  are conditionally independent given  $S$  (Ebert-Uphoff and Deng 2012a). If no such set exists, the edge is preserved. Next, “collider” structures (also called  $v$  structures) are identified—where a pair of edges meet a single node, such that the node has two parent nodes that need not be related—and as many directed edges as possible

are added that satisfy the constraints which dictate that loops (cycles) or further addition of collider structures are not allowed.

#### b. Parameter learning

Once the structure is defined, the model is “parameterized” by specifying a CPD and estimating the parameters of a distribution for each variable in the BN and every configuration of its parents. The expectation-maximization (EM) algorithm (Bilmes 1998) was used to parameterize DBNs using the junction tree inference engine. This includes the estimation of parameters for both hidden (HVs) and observed (input data) variables. The EM alternates between finding the expected sufficient statistics using the log-likelihood function, and maximizing the estimated likelihood function until a local maximum is converged upon and the parameter estimates are returned (Dempster et al. 1977). Once all the CPDs are defined, the model can be used to predict the node values in a test dataset to determine the fit.

#### c. Experiments

To test the hypothesis that important climatological relationships exist between North Atlantic atmospheric variability and remote Arctic and tropical drivers, we built a series of networks of increasing complexity. This ensured that any relationships captured by the structure-learning algorithm were consistent across models, and robust in terms of their predictive accuracy. All networks were constructed using the Bayes Net Toolbox (BNT) for MATLAB (Murphy 2001b), with all data preparation and plotting carried out in R.

Autocorrelation functions (ACFs) were used to determine that all variables depended linearly on their values from the previous time step. Thus, autoregressive links were coded into

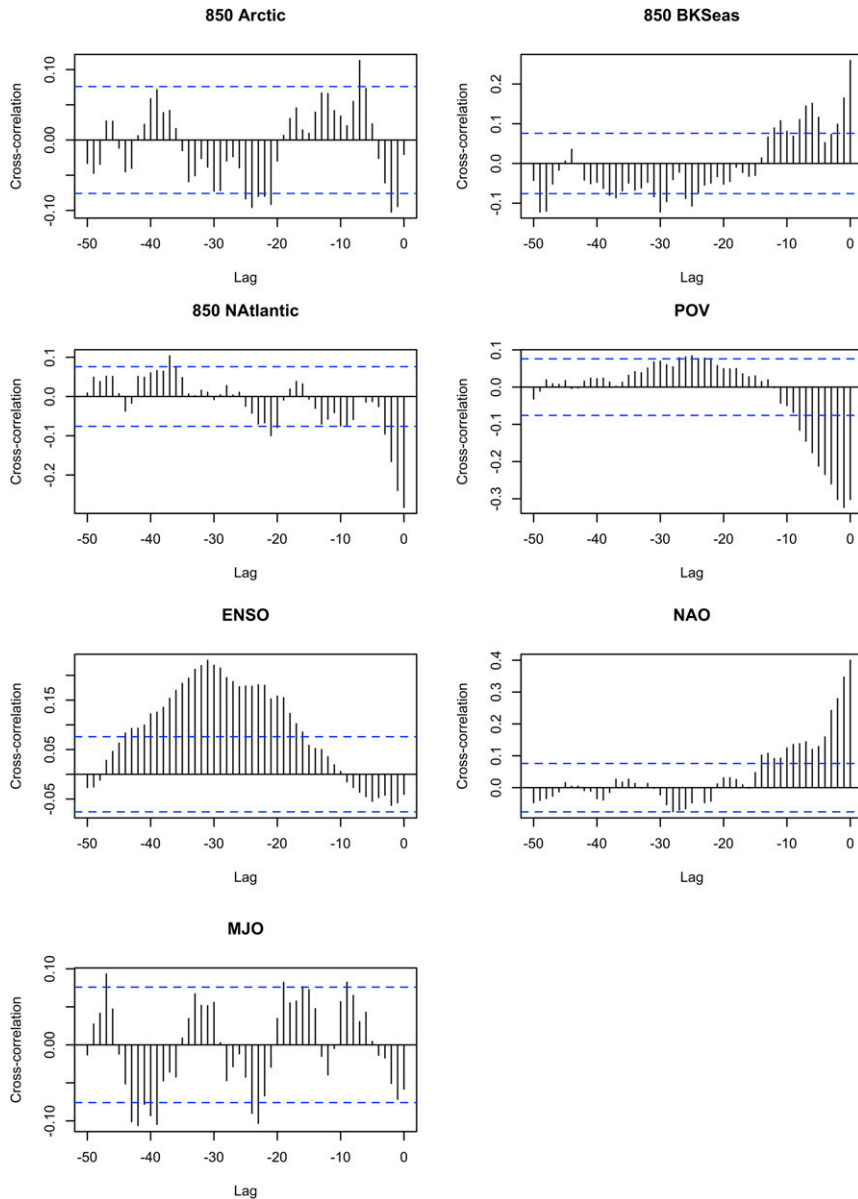


FIG. 4. Cross-correlation functions (CCFs) between all variables and jet latitude, used to determine the lead time, if any, for each dataset. “Lag” represents the number of time steps in 5-day-averaged data, and blue dotted line denotes significance at the 0.05 level.

the “inter” models (i.e., between time slices) for all variables. Cross-correlation functions (CCFs) were then plotted for each variable against jet latitude to determine the number of time steps (if any) at which each variable should be lagged (Fig. 4). The appropriate lead time for each variable was then used to build the datasets used for all analyses; where no clear maximum lag value was returned in the CCFs, the nearest significant lag was selected for the dataset. For the MJO, with a sinusoidal pattern of correlations, a lead time of 45 days was chosen. This lead time is slightly longer than links between the MJO signal to NAO forecasting (Yu and Lin 2016) and variability (Jiang et al. 2017). Further 5-day averaged datasets were

prepared based on previous work with a lead time of 10, 15, and 30 days for the MJO (Henderson et al. 2016; Jiang et al. 2017) and a 14-day lead for ENSO to identify potential influences through stratospheric (Baldwin and Dunkerton 2001) and tropospheric (Scaife et al. 2017a) pathways. Networks were run at each lead time to ensure tropical influences were not missed simply as a result of CCF lead selection. Lead times were run as separate DBNs rather than using different time-lagged variants of the same variable within graphs.

After being prepared with the steps above, the data were loaded into BNT and split into training and testing datasets (80:20) to allow for an unbiased estimate of the generalization

error and prevent overfitting (Shalizi 2013). All DBNs were run first with jet latitude, then with the MI swapped in its place. Three main sets of networks are constructed in the following experiments:

- 1) A DBN with autoregressive links and no hidden variables to act as a control run (i.e., to examine the impact of adding a hidden variable on predictive accuracy).
- 2) A DBN with the full Arctic area (70°–90°N), autoregressive links and one hidden variable linked to all observed nodes (i.e., Barents–Kara Seas and North Atlantic removed).
- 3) A DBN with Arctic subregions (Barents–Kara Seas and North Atlantic), autoregressive links, and one hidden variable linked to all observed nodes (i.e., full Arctic removed).

As a first step, the structure was learned with *PC stable*, which is the more robust, order-independent version of the PC algorithm, using the Fisher- $z$  test for conditional independence and an alpha value of 0.01. Results were robust to the choice of alpha between 0.01 and 0.05. In the case of experiments 2 and 3, a hidden variable with links to all variables and to itself through time (autoregressively to the next time step) is coded into the structure, as in Uusitalo et al. (2018). Next, any bidirectional links (which would lead to cyclical structures) are removed from the learned structure by necessity as DBNs require acyclical graphs; DAGs by definition cannot include loops (Scutari 2010; Scutari and Denis 2014). For example, the structure-learning phase revealed multiple bidirectional links between the jet latitude variable and connected nodes. Where bidirectional links occurred, the direction that preserved collider nodes was chosen because removing arcs from collider nodes with multiple inbound links undermines the individual probability distributions that make up a BN. As the jet was a multiple collider node, incoming conditionally dependent relationships were necessarily kept in order to preserve the “explain away” effect of the parent nodes (as in Ebert-Uphoff and Deng 2012a). This means that while midlatitude-to-Arctic linkages are found to be an important part of jet DBN structures below, they cannot be quantified in terms of strength using this technique.

The result of this process is then taken as the network structure and run through the parameterization and testing steps. The parameters are learned with the EM algorithm and the model is tested on the remainder of the data (the test dataset). BNs perform prediction using inference (Friedman et al. 2000); the  $n$ -step ahead prediction method iterates between entering the observations for all variables at time  $t - 1$  and applying inference to calculate their posterior distributions at time  $t$ , then repeats this step  $n$  times. This makes use of the *junction tree* algorithm for node value prediction in the test dataset (Murphy 1998), which was independent of the input dataset used to create and parameterize the model. The predictive accuracy of networks referred to in this study is an assessment of the model fit: test dataset values are predicted using the structure and parameters learned from the train dataset and then compared against the observed test values. The accuracy of fit is quantified using the sum of squared error (SSE) for each variable over all time resolutions. Network architectures with two hidden variables and multiple

connection orientations formed part of the analysis (not shown), resulting in overly parameterized models with low predictive accuracy.

As a final step, an analysis of variance (ANOVA) was conducted on all networks of experiment 3 to estimate the relative importance of each relationship found. Linear regressions, using all parent nodes as independent variables to predict dependent variable (child node) values, were calculated for all edges in networks across all time resolutions. An incremental sum of squares table was produced using ANOVA for each regression model, and the proportion of variance worked out by dividing the sum of squares for each independent variable by the total sum of squares multiplied by 100 (Montgomery 2012). Strength estimates are limited by the fact that independent variables were determined using the directions set at the DAG stage of network construction, and that ANOVA examines only linear relationships. Here, they are used to assign edge weights (arrow widths) to the DAG plots below.

#### 4. Results

Graphical models provide an easily interpretable interface, allowing for visualization of climate teleconnections as network structures. Nodes representing the variables are colored to indicate their relative geographical position for ease of interpretation. It should be noted that the term “arc” refers specifically to a directed edge between nodes in networks. Faint edges represent relationships hard-coded into the model (i.e., the hidden variable connections) and solid edges show those learned by the structure-learning algorithm in the case of all nonhidden variable links. Two nodes connected by the latter can be said to be conditionally dependent on each other, but BN arcs do not show cause-and-effect relationships unless strict assumptions have been met (Scutari and Denis 2014). Two stages of the network are shown: the partially directed graph (pDAG) returned at the structure-learning stage, which includes the hidden variable and tropical nodes (MJO and ENSO; Fig. 5), and fully directed graphs (DAGs) showing learned relationships used to run the DBNs in the second stage (no hidden or tropical variables; Figs. 6 and 7).

##### a. Network structures: The DAG results

The control run (no HV) DBNs generated the same learned structures as seen in Fig. 5; model accuracy changes with the addition of hidden variables are discussed below. Arctic–midlatitude linkages were only captured by the DBNs when regions of AA were used ( $T_{BK}$  and  $T_{NA}$  in Fig. 1). For experiment 2, DBNs using the entire Arctic revealed no links between the Arctic and the jet or MI, showing only an Arctic–NAO linkage over 5-day averages (not shown). Midlatitude circulation responses to AA are known to be sensitive to the regions selected (Pedersen et al. 2016; Screen 2017). The results for the entire Arctic therefore suggest that Arctic–midlatitude linkages are sensitive to the location of sea ice loss and amplified warming.

In the full networks of experiment 3 (Fig. 5), both the jet and MI DBNs capture a relationship between  $T_{BK}$ ,  $T_{NA}$ , and jet latitude for 5-day averages, pointing to a covariability between

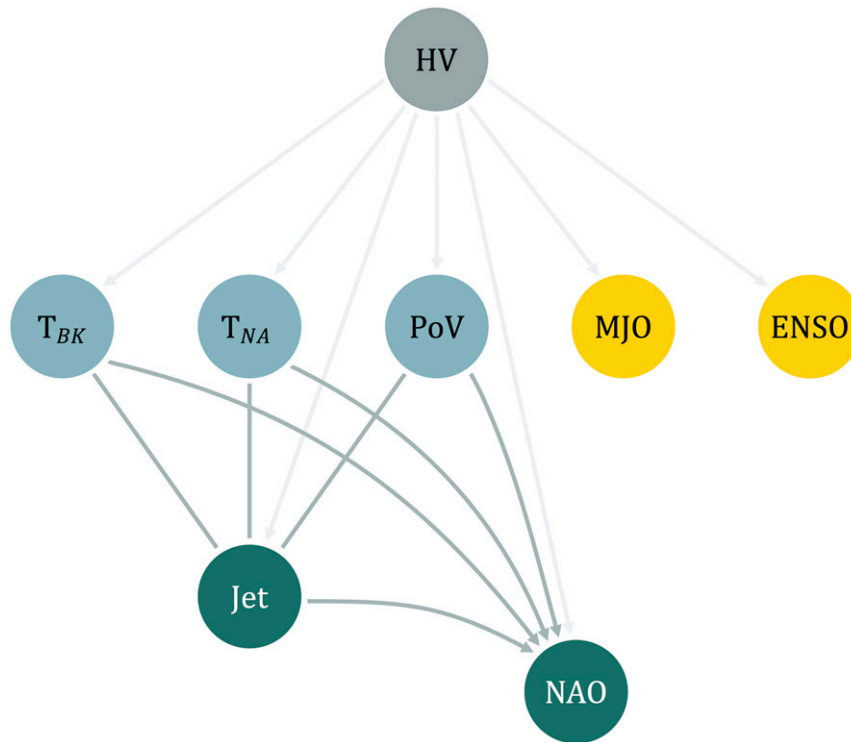


FIG. 5. pDAG returned by the PC algorithm for the 5-day jet HV DBN showing the “intra” (within the time slice) graph. Faint edges indicate edges coded into the model (i.e., HV edges), solid edges represent learned edges, with straight edges (i.e., no arrow) showing bidirectional relationships. Nodes are colored by their relative geographical location: the tropics (yellow), midlatitudes (green), and the Arctic (blue). Only the jet DBN is shown here.

important regions of AA and the jet stream’s latitudinal position. Prior to the removal of cyclical links (the stage of network construction that allows DBNs to function as detailed above) these links were undirected, which may indicate that simplistic cause-and-effect relationships do not explain interactions between AA regions and jet variability (Overland and Wang 2018). This is not the case for the MI DBNs; the meridional component of the jet stream is conditionally dependent on  $T_{NA}$ , a region of rapid warming and sea ice loss that includes Baffin Bay, and seems to impact temperature over the Barents–Kara Seas at 5-day resolutions (Fig. 7a). Note that  $T_{BK}$  is identified as a potential driver of the MI at 10- and 15-day averages, when the direct link between  $T_{NA}$  and the MI is lost.

Evidence of tropospheric–stratospheric coupling with a short lead time of 5–10 days is found for all networks except the 15-day MI DBN (Fig. 7c). Indirect links via a stratospheric pathway are missing however as no links pass through the PoV, which is either a parent or child node depending on the time resolution (Figs. 6 and 7). The PoV becomes the strongest link as a driver of jet variability at monthly intervals, but the vast increase in PoV SSE at 10- and 15-day intervals (Table 2, “Jet HV DBN” section) suggests that the direction of the jet–PoV arc is incorrect, as in monthly MI SSE (Table 2, “MI HV DBN” section). The DAG structures between Arctic ( $T_{BK}$ ,  $T_{NA}$ ), midlatitude (jet, MI, and NAO), and stratospheric (PoV) variables (Fig. 5) were found to be robust to the addition

of extraneous variables with no learned links (i.e., they did not change when unconnected variables were deliberately added).

Neither of the tropical variables are found to be conditionally dependent on any other variable for both the jet and MI DBNs over all time averages. This result was also robust to the choice of alpha value between 0.01 and 0.05, the choice of score or constraint-based algorithms (only results for PC are presented here), and the particular MJO index used (the RMM index is shown here). Figures 6 and 7 thus remove the hidden variable and tropical links for ease of visualization. This finding is consistent across model runs and is discussed in detail below.

#### b. Network performance: Predictive accuracy

The DBN with a HV (Fig. 6) had the most accurate performance for the jet DBNs in terms of SSE (Table 2), indicating that the inference engine was able to accommodate the increase in model complexity. MI DBNs without a HV performed marginally better overall although a few variables were predicted with less accuracy, most notably the MJO and ENSO nodes. The drop in SSE for the MI DBNs seems to be a result of PoV variability dominating the hidden state switches, discussed below in the hidden variable analysis; Fig. 7 demonstrates that the PoV node is less important for MI variability in contrast to jet latitude, where the PoV is connected throughout the time resolutions. The DBN without a HV outperformed it because the



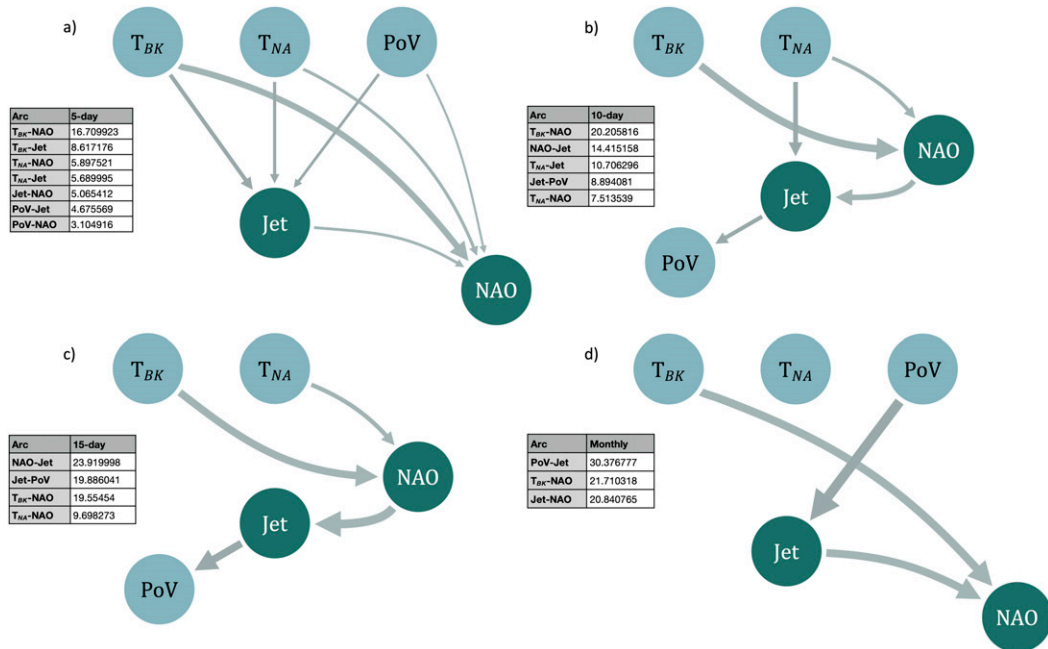


FIG. 6. DAGs for the jet HV DBN at (a) 5-day, (b) 10-day, (c) 15-day, and (d) monthly time averages showing the “intra” (within the time slice) graphs, with the HV and unconnected tropical variables removed for visualization purposes. Arc direction is indicated by arrows and inset tables show relative strength in terms of proportion of variance used to define arrow width.

strongest influence on the hidden state means and switches was the PoV, which is not as central to the learned structure of the MI models as the jet ones. Clearly, there are caveats to using hidden variables as model accuracy is not guaranteed to increase across

all variables; this finding is reflected in [Trifonova et al.’s \(2015\)](#) SSE results for biomass prediction.

Hidden variable “states” are the three possible values the discrete hidden variables can take, inferred from the observed

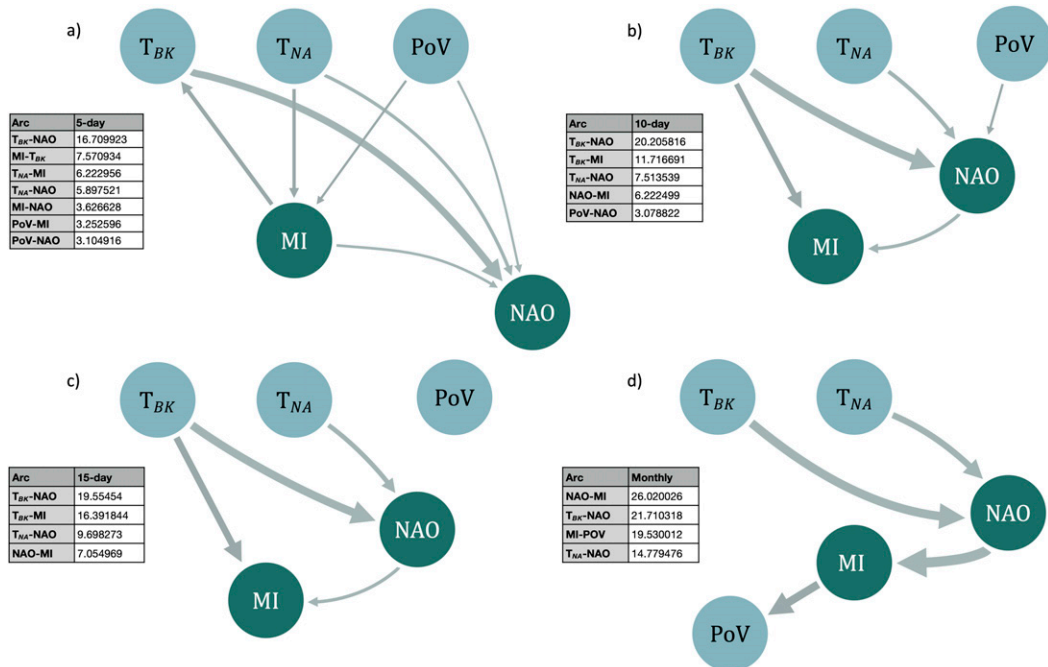


FIG. 7. As in Fig. 5, but for MI HV DBNs at (a) 5-day, (b) 10-day, (c) 15-day, and (d) monthly time averages.

TABLE 2. Unstandardized SSE for each variable at the 5-day, 10-day, 15-day, and monthly time resolutions for jet DBN with HV, jet DBN no HV, MI DBN with HV, and MI DBN no HV.

	5 days	10 days	15 days	Monthly
<b>Jet HV DBN</b>				
$T_{BK}$	25.46	26.67	18.41	13.46
$T_{NA}$	44.52	35.4	43.97	10.95
POV	1.68	20 472.46	34 495.65	4.84
Jet	28.12	258.23	31.21	9.38
NAO	34.49	16.54	35.78	23.66
MJO	7.67	15.77	12.35	10.76
ENSO	0.37	0.1	0.08	0.54
<b>Jet no HV DBN</b>				
$T_{BK}$	34.46	31.79	22.56	9.9
$T_{NA}$	52.29	53.18	40.37	21.15
POV	1.66	166 086.63	608.01	9.4
Jet	31.11	235.46	85.98	9.18
NAO	26.44	31.03	28.03	31.87
MJO	11.05	17.09	17.93	21.43
ENSO	0.56	0.66	0.85	1.68
<b>MI HV DBN</b>				
$T_{BK}$	17 627.71	25.29	25.33	14.87
$T_{NA}$	44.13	46.65	44.34	9.15
POV	0.71	3.65	16.51	259 469.81
MI	96.25	154.68	505.25	371.19
NAO	383.63	31.26	51.61	26.54
MJO	4.56	13.41	9.52	14.53
ENSO	0.11	0.41	2.81	2.95
<b>MI no HV DBN</b>				
$T_{BK}$	1887.33	31.79	22.56	9.9
$T_{NA}$	52.29	53.18	40.37	21.15
POV	1.66	4	5.22	16.68
MI	70.55	140.54	56.51	34.06
NAO	81.94	32.18	28.03	26.31
MJO	11.05	17.09	17.93	21.43
ENSO	0.56	0.66	0.85	1.68

data to maximize model fit. Time series of all variables were split into each hidden variable state, and the three states were then summarized in terms of mean and standard deviation (SD) to investigate what they represent (Table 3). For the jet DBNs, state 3 represents “average” conditions with values close to the mean values of the whole training dataset time series for the  $T_{BK}$ ,  $T_{NA}$ , PoV, and jet nodes. State 1 is associated with higher than average stratospheric polar vortex, lower temperatures over the  $T_{NA}$  region, and a marginally higher average jet latitude. Conversely, state 2 indicates a negative PoV geopotential height anomaly pattern (faster than average polar vortex), higher  $T_{NA}$  temperatures, and average jet conditions. Both states 1 and 2 have values within one standard deviation of the mean, but the hidden variable nonetheless captures three distinct states defined mainly by the PoV,  $T_{NA}$ , and jet variables. The MI HV states are less clear, but both sets of models are characterized by switches between states 1 and 3 (Fig. 8); two obvious differences are that the MI stays in state 1 for longer on average, and that the values for state 1 of the MI HV are closer to each variable’s mean value for the whole dataset than those of the jet HV.

TABLE 3. Hidden state mean values of each variable for jet DBN and MI DBN. Note that the analysis was run using fully prepared data but was reverse standardized here to aid interpretation.

HV state	$T_{BK}$	$T_{NA}$	POV	Jet/MI	NAO	MJO	ENSO
<b>Jet DBN</b>							
1	−0.44	−0.48	440.47	49.06	0.2	1.47	−0.04
2	−0.94	0.98	−490.58	47.83	0.3	1.39	0.09
3	−0.01	0.25	114.27	47.68	0.19	1.36	−0.01
<b>MI DBN</b>							
1	−0.3	−0.31	−41.52	1.67	0.24	1.32	0.01
2	0.15	0.81	−178.6	1.64	0.05	1.37	−0.03
3	−0.14	0.22	289.96	1.73	0.19	1.42	−0.02

### c. MJO phase: Further analysis

The focus on MJO amplitude due to the use of continuous winter (DJF) time series in DBNs gives no consideration to the effect of MJO phase on the existence of linkages. To account for phase, the two PC time series (RMM1 and RMM2) from the MJO index (Wheeler and Hendon 2004) were included in DBNs prepared with the steps detailed above using the CCF method for lead time selection. The MJO findings remained unchanged in 5-, 10-, and 15-day-averaged networks, but an RMM1–NAO link was identified in monthly jet DBNs with no lead time. The addition of this link to the monthly DBN caused the NAO variable to be predicted inaccurately, reflected in large SSE values for the NAO, which was only predicted accurately (as in Table 2, “Jet no HV DBN” section) when the RMM1–NAO link was removed. Clearly, using this specific methodology on monthly-averaged jet DBNs, the RMM1–NAO link identified by structure-learning does not provide any predictive power to the DBN when used for NAO prediction in the test dataset. Further study should focus on incorporating the MJO phase into networks that model Arctic and tropical drivers of midlatitude circulation with accuracy.

## 5. Discussion

The DBNs provide a robust set of results in terms of network structure and model accuracy. The results discussed hereafter generally focus on 5-day averages unless otherwise specified. One of the most consistent linkages picked up by the structure-learning phase points to the relationship between the Arctic regions of the North Atlantic and Barents–Kara Seas and the jet, MI, and NAO nodes. Atmospheric responses in northern Europe to anomalous sea ice loss in the Barents–Kara Seas region were thought to develop through changes in surface turbulent heat fluxes (Petoukhov and Semenov 2010; Liptak and Strong 2014), although recent evidence casts doubt on this (Blackport et al. 2019). The bidirectional edges of the  $T_{BK}$  and  $T_{NA}$  to the jet (Fig. 5) point to a positive feedback of Arctic–midlatitude effects, driven by the thermodynamic impact of temperature and moisture advection events from lower latitudes, which modulate Arctic temperatures (Kapsch et al. 2016; Gong et al. 2017; Cohen et al. 2018a). Figure 5 matches evidence for two-way feedbacks between the Arctic and

midlatitudes found on subseasonal time scales (Messori et al. 2018; Samarasinghe et al. 2019; McGraw and Barnes 2020). This feedback occurs on time scales of <5 days (Kapsch et al. 2016), suggesting that Fig. 5 captures a pathway thought to occur at similar time resolutions in model studies, one that may dictate the nature of the two-way relationship between AA and the midlatitudes. A similar covariability between the Barents–Kara Seas and Eurasian high pressure anomalies has been variably interpreted as warm Arctic anomalies driving high pressure responses observed over Eurasia (e.g., Honda et al. 2009; Kim et al. 2014; Mori et al. 2019) or internal atmospheric variability (McCusker et al. 2016; Sun et al. 2016) and midlatitude circulation as a driver of the observed warm Arctic temperature anomalies (e.g., Messori et al. 2018; Kelleher and Screen 2018; Blackport et al. 2019). Graphical models clearly represent a methodology that is capable of identifying covariability between Arctic warming and midlatitude circulation without the need for a response variable or single-pathway interpretations.

The  $T_{NA}$ –jet linkage, found in both the jet and MI 5-day DBNs and for 10-day averages in the jet, may represent an Arctic–jet connection known to develop intermittently and cause cold outbreaks through the assisted formation and support of blocking patterns (Chen and Luo 2017; Ballinger et al. 2018). Higher geopotential heights in the Greenland and Baffin Bay regions increase the likelihood of Greenland blocking events, contributing to the increased waviness of jet stream patterns (higher MI values) and in turn the persistence of cold events on the eastern coast of the United States (Chen and Luo 2017; Overland and Wang 2018).

The NAO, an index reflecting both jet latitude and speed variability (Woollings and Blackburn 2012), is central to the network structures in both sets of DBNs. While the NAO is likely influenced by a complex set of nonlinear drivers (Smith et al. 2016), near-surface Arctic temperature seems to have a significant impact on the NAO via the troposphere. The NAO is conditionally dependent on the high sea ice variability region of the Barents–Kara Seas, estimated to be the strongest link in 5- and 10-day jet DBNs and all submonthly MI intervals (Figs. 6 and 7). The response of atmospheric circulation indices to AA is constrained entirely within winter in this study as in Blackport and Screen (2019); in contrast, others have found that autumn sea ice conditions provide predictive skill of winter NAO variability in statistical models (Wang et al. 2017; Hall et al. 2017). Figure 7 shows that  $T_{NA}$  and MI are conditionally independent given the NAO node, but that they are connected via the NAO throughout the time averages as the  $T_{BK}$  and  $T_{NA}$  nodes impact NAO variability even to monthly resolutions. This is almost identical for jet latitude up to 15-day intervals.

This conditional independence, added to the strength of the AA–NAO links, implies that NAO phase shifts summarize the AA–midlatitude connection well at lower time resolutions. The negative phase of the NAO is indicative of a southerly displacement of the jet, which has been linked to colder, more severe winters in northern Europe and the eastern United States and warmer conditions over Greenland and the Barents–Kara Seas region (Cohen et al. 2018b). While the overall effect of AA and a warming world is likely to be that

cold outbreaks become less intense (Ayarzagüena and Screen 2016), AA may favor a shift toward meridional circulation patterns, which can promote the increased frequency and persistence of cold-air outbreaks (Cohen et al. 2018a). However, a shift to negative NAO patterns is by itself a potentially misleading trend as the warming trend may offset any dynamical cooling influence (Screen 2017).

#### a. Tropical influence and stratospheric teleconnection pathways

No conditionally dependent relationships were reliably found for the tropical variables; this was robust across all time averages, jet descriptors, and lead times used. This likely reflects on aspects of the data and the model design process, rather than simply pointing to a weak tropical influence on Arctic temperature and jet variability relative to Arctic–midlatitude covariability. Additional analysis using ENSO and MJO lead times based on expert knowledge was conducted to ensure that the lack of tropical input was not simply the result of the CCF analysis providing misleading lead times. Furthermore, the inclusion of the MJO phase in networks revealed an RMM1–NAO link in a monthly averaged jet DBN, although this link significantly inhibited predictive accuracy in the NAO variable. Modifications to the data preparation and model design steps may result in tropical links which enhance accuracy; the use of continuous (DJF) time series for example may be a limiting factor as combined effects and long-duration teleconnections through the stratosphere and upper troposphere may have been effectively masked.

A consistent lack of ENSO influence across networks is significant given that skillful prediction of the wintertime NAO (Scaife et al. 2016, 2017b) and AO (Sun and Ahn 2015) can be achieved with the inclusion of tropical variability in model simulations. With methods similar to those employed here, previous work has identified tropical influences using non-continuous time series analysis to split the data up into weather events; Henderson et al. (2016) for example use only days with an RMM amplitude greater than 1 to assess blocking frequency [as defined by Wheeler and Hendon (2004)], and Barnes et al. (2019) discretize MJO data into active and inactive periods to investigate pathways between the MJO and NAO. If the absence of tropical teleconnections is a result of the continuous time series approach masking intermittent but significant linkages, it is important to note that DBNs consistently pick up the AA–midlatitude covariability. Arctic–midlatitude linkages are thought to be similarly intermittent in nature as they rely on the background jet stream pattern to act as a bridge between thermodynamic forcing and persistent midlatitude extremes (Overland et al. 2016; Overland and Wang 2018; Kolstad and Screen 2019). In this research, this did not prevent their identification at submonthly time scales.

The stratospheric polar vortex has been shown to project onto the NAO and increase the likelihood of Atlantic blocking during vortex weakening events through tropospheric–stratospheric coupling (Kidston et al. 2015). The stratosphere exhibits much more stability than the troposphere below it, requiring Rossby wave activity from below to cause significant disruptions to the flow of the polar vortex. A large body of

studies favor the tropics as the predominant source of Rossby waves necessary to trigger sudden warming events (Liu et al. 2014; Hitchcock and Simpson 2014; Scaife et al. 2017b; Jiang et al. 2017; Hardiman et al. 2019), NAO variability (Scaife et al. 2017a), and Arctic warming itself (Yoo et al. 2012). It is highly probable that uncaptured tropical teleconnections contribute to the strength of the PoV to jet, MI, and NAO linkages; El Niño events can force the Aleutian low to strengthen and move eastward, causing enhanced vertical wave transport and negative NAO conditions through the disruption of PoV flow (Scaife et al. 2017b; Hardiman et al. 2019). Similarly, enhanced tropical convection associated with the MJO may weaken the PoV and disturb the NAO (Liu et al. 2014; Jiang et al. 2017) depending on the background state of the PoV (Barnes et al. 2019), although NAO events do not depend solely on tropical convection. Only the final stages of these teleconnections via the stratosphere are captured in DBN structures.

### b. The implications of hidden state switches

Recall that the hidden variable states are the three discrete values hidden variables can take, inferred from the observed data to maximize model fit, and that a physical interpretation of each state was inferred by splitting and averaging the dataset time series by state (Table 3): a switch between average conditions in state 3 to slower stratospheric polar vortex flow, lower  $T_{NA}$  temperatures, and marginally higher jet latitude values in state 1. The hidden variable state switches did not show a clear delineation between an AA or pre-AA period (Figs. 8a,b), that is, a shift to an amplified warming state signaled in the  $T_{BK}$  and  $T_{NA}$  variables. This is precisely the type of underlying state switch one would expect to be identified in a hidden variable but may have been filtered out by the detrending step of data preparation. The hidden variables also hint that the proportion of variance estimates (Figs. 6 and 7) may not fully represent the importance of stratospheric variability at submonthly time scales as the PoV seems to dictate the hidden state switches by shifting from average (state 3) to slower anomalous PoV flow in state 1 (Fig. 8c). Added to this, Fig. 6 shows that the PoV is connected throughout the time resolutions, and the lack of arcs between the PoV and MI in Figs. 7b and 7c may have caused the drop in model performance (Table 2, “MI HV DBN” section) with the addition of a hidden variable to the MI DBN. Hidden variables maximize model fit by inferring their state from the observed data. The low predictive accuracy of the MI variable in Table 2 (“MI HV DBN” section), particularly toward lower time resolutions, could therefore be a result of the lack of links between the PoV and MI combined with the strong influence of the PoV on state switches.

The importance of stratospheric variability in driving the hidden state switches was further investigated using multinomial logistic regression to predict the probability of the hidden variable switching states from a reference value of state 3 (average conditions) given the observed variables (Menard 2001). The log odds of the states are modeled as a linear combination of the observed variables. The odds ratio, calculated by exponentiating the estimated regression coefficients, describes how much the relative risk of the hidden variable being in state 1 is multiplied by for a unit change in the

observed variable (e.g., the PoV, with all others held constant). Put simply, it models the odds of the hidden variable being in state 1 given that we have high PoV anomaly values. The results suggest that three variables drive the shift from state 3 to state 1: the PoV increases the odds by a factor of 1.212 (or 21.2%), the jet by 1.242 (24.2%), and the MJO by 1.199 (19.9%). This confirms both that the PoV is central to the occurrence of state 1 regimes, as suggested by the state average values (Table 3), and that the MJO forcing of slow stratospheric polar vortex states was captured in hidden variable state shifts but not graphically in network structures.

It is worth noting that recent studies have suggested that the covariability between regions of AA and midlatitude flow characteristics could simply be a result of internal variability, with a forced component originating in the tropics that was uncaptured in the networks. The relationship between  $T_{BK}$  and the NAO was found to be a recent and weak feature in reanalyses and ensemble simulations (Kolstad and Screen 2019). Warner et al. (2020) concluded that internal variability in the Atlantic sector combined with tropical Pacific forcing was more likely to cause  $T_{BK}$ –NAO covariability than sea ice forcing. Nonsimulated contributions from remote forcing could not be ruled out in these studies, but strong anomalous tropical forcing continues to provide skill for forecasting in contrast to extratropical drivers, which remain limited by their low signal-to-noise ratio resulting from natural variability (Trenberth et al. 1998).

## 6. Conclusions

This study represents the first foray into Arctic–midlatitude weather linkages using a graphical model approach paired with hidden variables. A robust covariability between regions of amplified Arctic warming and midlatitude circulation characteristics is found, suggesting that this feedback has a significant influence on variability both in the Arctic and the North Atlantic midlatitudes at submonthly time scales. The two-way nature of the link suggests that unidirectional interpretations need revising in favor of one that takes into account the central importance of poleward heat and moisture fluxes into the Arctic. Of the relationships found, the  $T_{BK}$ –NAO link has the strongest impact on the DBNs at 5- and 10-day intervals, with the PoV–jet link becoming the strongest at monthly resolutions. Midlatitude circulation responses at submonthly time scales are driven by AA processes within winter rather than a lagged response to sea ice losses during autumn. Links to tropical modes of variability were not identified in models shown to have predictive accuracy over a range of lead times, but this is a result of data and model design and in no way implies that their impact on observed jet variability is small. The DBNs found no evidence for a clearly defined AA and pre-AA period through hidden variables included in model architectures.

Clearly, these results only give us part of the picture. Hidden variables revealed two states that essentially represent two different modes of stratospheric polar vortex variability, namely a shift from average conditions to slower anomalous PoV flow. The hidden variables hint that strength estimates may underrepresent the importance of the PoV, with no lagged connections to midlatitude flow identified through the



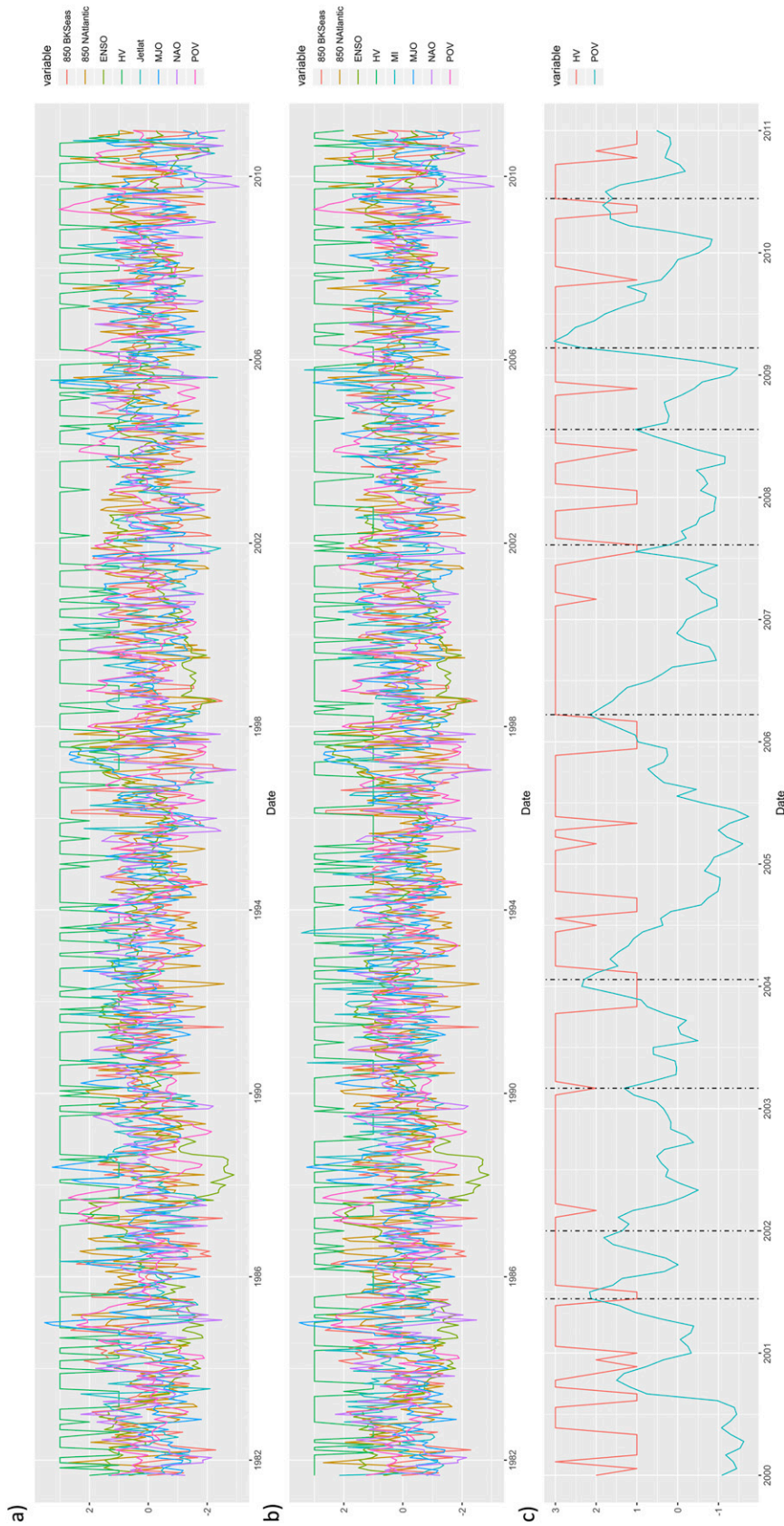


FIG. 8. Winter (DJF) time series of 5-day train dataset data showing observed and hidden variables (green) for (a) jet HV DBN and (b) MJO HV DBN. (c) A subset of (a) showing the final decade (2000–10) of the HV and PoV variables from the training dataset with the onset dates of sudden stratospheric warming events (SSWs) plotted as vertical dashed lines. The central role that strong positive PoV geopotential height anomalies (i.e., a weak polar vortex) have on HV state switches can be observed in (c). SSW onset dates were taken from Cao et al. (2019). Data in all panels are left as standardized for visualization purposes.



stratosphere. The findings focus on submonthly time scales and as such may miss stratospheric links thought to be important for North Atlantic circulation (Scaife et al. 2017a; Hardiman et al. 2019). As a form of time series analysis, the DBNs may not have picked up state-dependent linkages or combined effects like those shown in simulation analyses (e.g., Lee et al. 2015), which could have masked potential tropical influences like the ENSO–PoV linkage. Furthermore, input data were averaged over large areas chosen to maximize the amount of variability captured, which may not represent the true complexity of midlatitude circulation influences. Given that bidirectional edges must be removed for the static BN to be used in a DBN, the strength and accuracy of the midlatitude-to-Arctic relationship found in BNs cannot be quantified due to the requirement of directed graphs in DBNs. The use of multiple lagged copies of a variable in the same network may offer a way to quantify these feedbacks (e.g., Ebert-Uphoff and Deng 2012a; Samarasinghe et al. 2019; Barnes et al. 2019) and would allow for the use of temporal constraints to orient edge direction alongside collider structures.

Graphical models paired with structure learning algorithms prove to be a valuable tool for investigating complex climate teleconnections in a network that includes Arctic, midlatitude, and tropical variables, and may provide a valuable alternative to correlation-based networks and studies (Ebert-Uphoff and Deng 2012a). DBNs with a single hidden variable were found to have predictive accuracy, subject to a number of important caveats; their addition does not guarantee that network accuracy will increase. Network performance was high despite the noisy internal variability of North Atlantic flow and the low signal-to-noise ratio of AA–midlatitude covariability, suggesting that any nonlinearity in linkages did not substantially interfere with accurate node prediction in the test dataset. Further study seeking to apply DBNs should note that they can suffer from overparameterization and will lead to inaccurate variable prediction where graph structure deviates significantly from the climate system we are trying to model, as shown here for jet–PoV links. This study can be seen as part of a growing effort to understand the Arctic contribution to jet variability using machine learning approaches (Kretschmer et al. 2016; Barnes and Simpson 2017; Francis et al. 2018; Samarasinghe et al. 2019). Graphical models would prove themselves more than capable in a supporting role as the results from large coordinated modeling projects are published in the near future.

**Acknowledgments.** The authors wish to thank three anonymous reviewers and the handling editor James Screen, whose comments improved the manuscript. Elisabete Silva provided valuable practical support throughout the work. This work was supported by the Natural Environmental Research Council Grant NE/L002485/1. GDC was supported by the Bundesministerium für Bildung und Forschung Grant 01LP1611A.

**Data availability statement.** All climate index data (NAO, ENSO, and MJO) and reanalysis data ( $T_{BK}$ ,  $T_{NA}$ , and PoV) are derived from public data that is openly available; source websites are given alongside their references in the data section

(section 2). Jet latitude data are derived from ERA-Interim data (700–900 hPa zonal wind), publicly available at <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim>. The Meandering Index (MI) is also derived from ERA-Interim (500-hPa geopotential height) at the same source, but is available in fully processed form at <https://github.com/giorgiadicapua/MeanderingIndex>.

## REFERENCES

- Ayarzagüena, B., and J. A. Screen, 2016: Future Arctic sea ice loss reduces severity of cold air outbreaks in midlatitudes. *Geophys. Res. Lett.*, **43**, 2801–2809, <https://doi.org/10.1002/2016GL068092>.
- Baldwin, M. P., and T. J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science*, **294**, 581–584, <https://doi.org/10.1126/science.1063315>.
- Ballinger, T. J., E. Hanna, R. J. Hall, J. Miller, M. H. Ribergaard, and J. L. Høyer, 2018: Greenland coastal air temperatures linked to Baffin Bay and Greenland Sea ice conditions during autumn through regional blocking patterns. *Climate Dyn.*, **50**, 83–100, <https://doi.org/10.1007/s00382-017-3583-3>.
- Barnes, E. A., and I. R. Simpson, 2017: Seasonal sensitivity of the Northern Hemisphere jet streams to Arctic temperatures on subseasonal time scales. *J. Climate*, **30**, 10 117–10 137, <https://doi.org/10.1175/JCLI-D-17-0299.1>.
- , S. M. Samarasinghe, I. Ebert-Uphoff, and J. C. Furtado, 2019: Tropospheric and stratospheric causal pathways between the MJO and NAO. *J. Geophys. Res. Atmos.*, **124**, 9356–9371, <https://doi.org/10.1029/2019JD031024>.
- Bilmes, J. A., 1998: A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *Int. Comput. Sci. Inst.*, **4**, 126.
- Blackport, R., and J. A. Screen, 2019: Influence of Arctic sea ice loss in autumn compared to that in winter on the atmospheric circulation. *Geophys. Res. Lett.*, **46**, 2213–2221, <https://doi.org/10.1029/2018GL081469>.
- , —, K. van der Wiel, and R. Bintanja, 2019: Minimal influence of reduced Arctic Sea ice on coincident cold winters in mid-latitudes. *Nat. Climate Change*, **9**, 697–704, <https://doi.org/10.1038/s41558-019-0551-4>.
- Brönnimann, S., 2007: Impact of El Niño–Southern Oscillation on European climate. *Rev. Geophys.*, **45**, RG3003, <https://doi.org/10.1029/2006RG000199>.
- Cao, C., Y.-H. Chen, J. Rao, S.-M. Liu, S.-Y. Li, M.-H. Ma, and Y.-B. Wang, 2019: Statistical characteristics of major sudden stratospheric warming events in CESM1-WACCM: A comparison with the JRA55 and NCEP/NCAR reanalyses. *Atmosphere*, **10**, 519, <https://doi.org/10.3390/atmos10090519>.
- Chen, X., and D. Luo, 2017: Arctic sea ice decline and continental cold anomalies: Upstream and downstream effects of Greenland blocking. *Geophys. Res. Lett.*, **44**, 3411–3419, <https://doi.org/10.1002/2016GL072387>.
- Cohen, J., and Coauthors, 2018a: Arctic change and possible influence on mid-latitude climate and weather. U.S. CLIVAR Rep. 2018-1, 41 pp., <https://doi.org/10.5065/D6TH8KGW>.
- , K. Pfeiffer, and J. A. Francis, 2018b: Warm Arctic episodes linked with increased frequency of extreme winter weather in the United States. *Nat. Commun.*, **9**, 869, <https://doi.org/10.1038/s41467-018-02992-9>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.

- Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977: Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc.*, **39B**, 1–22, <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>.
- Di Capua, G., and D. Coumou, 2016: Changes in meandering of the Northern Hemisphere circulation. *Environ. Res. Lett.*, **11**, 094028, <https://doi.org/10.1088/1748-9326/11/9/094028>.
- , and Coauthors, 2019: Long-lead statistical forecasts of the Indian summer monsoon rainfall based on causal precursors. *Wea. Forecasting*, **34**, 1377–1394, <https://doi.org/10.1175/WAF-D-19-0002.1>.
- Ebert-Uphoff, I., and Y. Deng, 2012a: A new type of climate network based on probabilistic graphical models: Results of boreal winter versus summer. *Geophys. Res. Lett.*, **39**, L19701, <https://doi.org/10.1029/2012GL053269>.
- , and —, 2012b: Causal discovery for climate research using graphical models. *J. Climate*, **25**, 5648–5665, <https://doi.org/10.1175/JCLI-D-11-00387.1>.
- Francis, J. A., and S. J. Vavrus, 2015: Evidence for a wavier jet stream in response to rapid Arctic warming. *Environ. Res. Lett.*, **10**, 014005, <https://doi.org/10.1088/1748-9326/10/1/014005>.
- , N. Skific, and S. J. Vavrus, 2018: North American weather regimes are becoming more persistent: Is Arctic amplification a factor? *Geophys. Res. Lett.*, **45**, 11 414–11 422, <https://doi.org/10.1029/2018GL080252>.
- Friedman, N., M. Linial, I. Nachman, and D. Pe'er, 2000: Using Bayesian networks to analyze expression data. *J. Comput. Biol.*, **7**, 601–620, <https://doi.org/10.1089/106652700750050961>.
- Gong, T., S. Feldstein, and S. Lee, 2017: The role of downward infrared radiation in the recent Arctic winter warming trend. *J. Climate*, **30**, 4937–4949, <https://doi.org/10.1175/JCLI-D-16-0180.1>.
- Hall, R. J., R. Erdélyi, E. Hanna, J. M. Jones, and A. A. Scaife, 2015: Drivers of North Atlantic polar front jet stream variability. *Int. J. Climatol.*, **35**, 1697–1720, <https://doi.org/10.1002/joc.4121>.
- , J. M. Jones, E. Hanna, A. A. Scaife, and R. Erdélyi, 2017: Drivers and potential predictability of summer time North Atlantic polar front jet variability. *Climate Dyn.*, **48**, 3869–3887, <https://doi.org/10.1007/s00382-016-3307-0>.
- Hardiman, S. C., N. J. Dunstone, A. A. Scaife, D. M. Smith, S. Ineson, J. Lim, and D. Fereday, 2019: The impact of strong El Niño and La Niña events on the North Atlantic. *Geophys. Res. Lett.*, **46**, 2874–2883, <https://doi.org/10.1029/2018GL081776>.
- Henderson, S. A., E. D. Maloney, and E. A. Barnes, 2016: The influence of the Madden–Julian oscillation on Northern Hemisphere winter blocking. *J. Climate*, **29**, 4597–4616, <https://doi.org/10.1175/JCLI-D-15-0502.1>.
- Hitchcock, P., and I. R. Simpson, 2014: The downward influence of stratospheric sudden warmings. *J. Atmos. Sci.*, **71**, 3856–3876, <https://doi.org/10.1175/JAS-D-14-0012.1>.
- Honda, M., J. Inoue, and S. Yamane, 2009: Influence of low Arctic sea-ice minima on anomalously cold Eurasian winters. *Geophys. Res. Lett.*, **36**, L08707, <https://doi.org/10.1029/2008GL037079>.
- Huang, J., and Coauthors, 2017: Recently amplified Arctic warming has contributed to a continual global warming trend. *Nat. Climate Change*, **7**, 875–879, <https://doi.org/10.1038/s41558-017-0009-5>.
- Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, **269**, 676–679, <https://doi.org/10.1126/science.269.5224.676>.
- Jiang, Z., S. B. Feldstein, and S. Lee, 2017: The relationship between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Quart. J. Roy. Meteor. Soc.*, **143**, 240–250, <https://doi.org/10.1002/qj.2917>.
- Kapsch, M. L., R. G. Graversen, M. Tjernström, and R. Bintanja, 2016: The effect of downwelling longwave and shortwave radiation on Arctic summer sea ice. *J. Climate*, **29**, 1143–1159, <https://doi.org/10.1175/JCLI-D-15-0238.1>.
- Kelleher, M., and J. Screen, 2018: Atmospheric precursors of and response to anomalous Arctic sea ice in CMIP5 models. *Adv. Atmos. Sci.*, **35**, 27–37, <https://doi.org/10.1007/s00376-017-7039-9>.
- Kidston, J., A. A. Scaife, S. C. Hardiman, D. M. Mitchell, N. Butchart, M. P. Baldwin, and L. J. Gray, 2015: Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. *Nat. Geosci.*, **8**, 433–440, <https://doi.org/10.1038/ngeo2424>.
- 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**, 4646, <https://doi.org/10.1038/ncomms5646>.
- , and Coauthors, 2017: Major cause of unprecedented Arctic warming in January 2016: Critical role of an Atlantic winter storm. *Sci. Rep.*, **7**, 40051, <https://doi.org/10.1038/srep40051>.
- Kolstad, E. W., and J. A. Screen, 2019: Nonstationary relationship between autumn Arctic sea ice and the winter North Atlantic Oscillation. *Geophys. Res. Lett.*, **46**, 7583–7591, <https://doi.org/10.1029/2019GL083059>.
- Kretschmer, M., D. Coumou, J. F. Donges, and J. Runge, 2016: Using causal effect networks to analyze different Arctic drivers of midlatitude winter circulation. *J. Climate*, **29**, 4069–4081, <https://doi.org/10.1175/JCLI-D-15-0654.1>.
- , J. Runge, and D. Coumou, 2017: Early prediction of extreme stratospheric polar vortex states based on causal precursors. *Geophys. Res. Lett.*, **44**, 8592–8600, <https://doi.org/10.1002/2017GL074696>.
- Lee, M. Y., C. C. Hong, and H. H. Hsu, 2015: Compounding effects of warm sea surface temperature and reduced sea ice on the extreme circulation over the extratropical North Pacific and North America during the 2013–2014 boreal winter. *Geophys. Res. Lett.*, **42**, 1612–1618, <https://doi.org/10.1002/2014GL062956>.
- Lin, H., G. Brunet, and B. Yu, 2015: Interannual variability of the Madden–Julian Oscillation and its impact on the North Atlantic Oscillation in the boreal winter. *Geophys. Res. Lett.*, **42**, 5571–5576, <https://doi.org/10.1002/2015GL064547>.
- Liptak, J., and C. Strong, 2014: The winter atmospheric response to sea ice anomalies in the Barents Sea. *J. Climate*, **27**, 914–924, <https://doi.org/10.1175/JCLI-D-13-00186.1>.
- Liu, C., and E. A. Barnes, 2015: Extreme moisture transport into the Arctic linked to Rossby wave breaking. *J. Geophys. Res. Atmos.*, **120**, 3774–3788, <https://doi.org/10.1002/2014JD022796>.
- , B. Tian, K. F. Li, G. L. Manney, N. J. Livesey, Y. L. Yung, and D. E. Waliser, 2014: Northern Hemisphere mid-winter vortex-displacement and vortex-split stratospheric sudden warmings: Influence of the Madden–Julian Oscillation and Quasi-Biennial Oscillation. *J. Geophys. Res. Atmos.*, **119**, 12 599–12 620, <https://doi.org/10.1002/2014JD021876>.
- McCusker, K. E., J. C. Fyfe, and M. Sigmond, 2016: Twenty-five winters of unexpected Eurasian cooling unlikely due to Arctic sea-ice loss. *Nat. Geosci.*, **9**, 838–842, <https://doi.org/10.1038/ngeo2820>.
- McGraw, M. C., and E. A. Barnes, 2020: New insights on sub-seasonal Arctic–midlatitude causal connections from a regularized regression model. *J. Climate*, **33**, 213–228, <https://doi.org/10.1175/JCLI-D-19-0142.1>.
- Menard, S., 2001: *Applied Logistic Regression Analysis*. 2nd ed. Sage, 111 pp.
- Messori, G., C. Woods, and R. Caballero, 2018: On the drivers of wintertime temperature extremes in the high Arctic. *J. Climate*, **31**, 1597–1618, <https://doi.org/10.1175/JCLI-D-17-0386.1>.

- Mihajlovic, V., and M. Petkovic, 2001: *Dynamic Bayesian Networks: A State of the Art*. University of Twente, 37 pp.
- Montgomery, D. C., 2012: *Design and Analysis of Experiments*. 8th ed. John Wiley and Sons, 757 pp.
- Mori, M., Y. Kosaka, M. Watanabe, H. Nakamura, and M. Kimoto, 2019: A reconciled estimate of the influence of Arctic sea-ice loss on recent Eurasian cooling. *Nat. Climate Change*, **9**, 123–129, <https://doi.org/10.1038/s41558-018-0379-3>.
- Murphy, K. P., 1998: Inference and learning in hybrid Bayesian networks. Rep. UCB/CSD-98-990, 18 pp., <https://www.cs.ubc.ca/~murphyk/Papers/cg.pdf>.
- , 2001a: An introduction to graphical models, 19 pp., [https://www.cs.ubc.ca/~murphyk/Papers/intro\\_gm.pdf](https://www.cs.ubc.ca/~murphyk/Papers/intro_gm.pdf).
- , 2001b: The Bayes net toolbox for Matlab. *Comput. Sci. Stat.*, **33**, 1024–1034.
- , 2012: *Machine Learning: A Probabilistic Perspective*. Adaptive Computation and Machine Learning Series, MIT Press, 1098 pp.
- Olonscheck, D., T. Mauritsen, and D. Notz, 2019: Arctic sea-ice variability is primarily driven by atmospheric temperature fluctuations. *Nat. Geosci.*, **12**, 430–434, <https://doi.org/10.1038/s41561-019-0363-1>.
- Overland, J. E., and M. Wang, 2018: Resolving future Arctic/midlatitude weather connections. *Earth's Future*, **6**, 1146–1152, <https://doi.org/10.1029/2018EF000901>.
- , J. A. Francis, R. Hall, E. Hanna, S. J. Kim, and T. Vihma, 2015: The melting Arctic and midlatitude weather patterns: Are they connected? *J. Climate*, **28**, 7917–7932, <https://doi.org/10.1175/JCLI-D-14-00822.1>.
- , and Coauthors, 2016: Nonlinear response of mid-latitude weather to the changing Arctic. *Nat. Climate Change*, **6**, 992–999, <https://doi.org/10.1038/nclimate3121>.
- Pedersen, R. A., I. Cvijanovic, P. L. Langen, and B. M. Vinther, 2016: The impact of regional Arctic sea ice loss on atmospheric circulation and the NAO. *J. Climate*, **29**, 889–902, <https://doi.org/10.1175/JCLI-D-15-0315.1>.
- 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.*, **115**, D21111, <https://doi.org/10.1029/2009JD013568>.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, <https://doi.org/10.1029/2002JD002670>.
- Rigor, I. G., J. M. Wallace, and R. L. Colony, 2002: Response of sea ice to the Arctic Oscillation. *J. Climate*, **15**, 2648–2663, [https://doi.org/10.1175/1520-0442\(2002\)015<2648:ROSITT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<2648:ROSITT>2.0.CO;2).
- Runge, J., 2018: Causal network reconstruction from time series: From theoretical assumptions to practical estimation. *Chaos*, **28**, 075310, <https://doi.org/10.1063/1.5025050>.
- , 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. Climate*, **27**, 720–739, <https://doi.org/10.1175/JCLI-D-13-00159.1>.
- Samarasinghe, S. M., M. C. McGraw, E. A. Barnes, and I. Ebert-Uphoff, 2019: A study of links between the Arctic and the midlatitude jet stream using Granger and Pearl causality. *Environmetrics*, **30**, e2540, <https://doi.org/10.1002/env.2540>.
- Scaife, A. A., and Coauthors, 2016: Seasonal winter forecasts and the stratosphere. *Atmos. Sci. Lett.*, **17**, 51–56, <https://doi.org/10.1002/asl.598>.
- , and Coauthors, 2017a: Predictability of European winter 2015/2016. *Atmos. Sci. Lett.*, **18**, 38–44, <https://doi.org/10.1002/asl.721>.
- , and Coauthors, 2017b: Tropical rainfall, Rossby waves and regional winter climate predictions. *Quart. J. Roy. Meteor. Soc.*, **143** (702), 1–11, <https://doi.org/10.1002/qj.2910>.
- Screen, J. A., 2017: Simulated atmospheric response to regional and pan-Arctic sea ice loss. *J. Climate*, **30**, 3945–3962, <https://doi.org/10.1175/JCLI-D-16-0197.1>.
- , and Coauthors, 2018: Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across climate models. *Nat. Geosci.*, **11**, 155–163, <https://doi.org/10.1038/s41561-018-0059-y>.
- Scutari, M., 2010: Learning Bayesian networks with the bnlearn R package. *J. Stat. Software*, **35** (3), 1–22, <https://doi.org/10.18637/jss.v035.i03>.
- , and J.-B. Denis, 2014: *Bayesian Networks with Examples in R*. Taylor & Francis, 225 pp.
- Shalizi, C. R., 2013: *Advanced Data Analysis from an Elementary Point of View*. CRC Press, 801 pp.
- Smith, D. M., A. A. Scaife, R. Eade, and J. R. Knight, 2016: Seasonal to decadal prediction of the winter North Atlantic Oscillation: Emerging capability and future prospects. *Quart. J. Roy. Meteor. Soc.*, **142**, 611–617, <https://doi.org/10.1002/qj.2479>.
- Spirtes, P., and C. Glymour, 1991: An algorithm for fast recovery of sparse causal graphs. *Soc. Sci. Comput. Rev.*, **9**, 62–72, <https://doi.org/10.1177/089443939100900106>.
- Sun, J., and J. B. Ahn, 2015: Dynamical seasonal predictability of the Arctic Oscillation using a CGCM. *Int. J. Climatol.*, **35**, 1342–1353, <https://doi.org/10.1002/joc.4060>.
- Sun, L., J. Perlwitz, and M. Hoerling, 2016: What caused the recent “warm Arctic, cold continents” trend pattern in winter temperatures? *Geophys. Res. Lett.*, **43**, 5345–5352, <https://doi.org/10.1002/2016GL069024>.
- Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling. *J. Geophys. Res.*, **103**, 14 291–14 324, <https://doi.org/10.1029/97JC01444>.
- Trifonova, N., A. Kenny, D. Maxwell, D. Duplisea, J. Fernandes, and A. Tucker, 2015: Spatio-temporal Bayesian network models with latent variables for revealing trophic dynamics and functional networks in fisheries ecology. *Ecol. Inform.*, **30**, 142–158, <https://doi.org/10.1016/j.ecoinf.2015.10.003>.
- , D. Maxwell, J. Pinnegar, A. Kenny, and A. Tucker, 2017: Predicting ecosystem responses to changes in fisheries catch, temperature, and primary productivity with a dynamic Bayesian network model. *ICES J. Mar. Sci.*, **74**, 1334–1343, <https://doi.org/10.1093/icesjms/fsw231>.
- , M. Karnauskas, and C. Kelble, 2019: Predicting ecosystem components in the Gulf of Mexico and their responses to climate variability with a dynamic Bayesian network model. *PLOS ONE*, **14**, e0209257, <https://doi.org/10.1371/journal.pone.0209257>.
- Tucker, A., and X. Liu, 2004: A Bayesian network approach to explaining time series with changing structure. *Intell. Data Anal.*, **8**, 469–480, <https://doi.org/10.3233/IDA-2004-8504>.
- Uusitalo, L., M. T. Tomczak, B. Müller-Karulis, I. Putnis, N. Trifonova, and A. Tucker, 2018: Hidden variables in a dynamic Bayesian network identify ecosystem level change. *Ecol. Inform.*, **45**, 9–15, <https://doi.org/10.1016/j.ecoinf.2018.03.003>.
- Wang, S. Y. S., Y. H. Lin, M. Y. Lee, J. H. Yoon, J. D. D. Meyer, and P. J. Rasch, 2017: Accelerated increase in the Arctic tropospheric warming events surpassing stratospheric warming events during winter. *Geophys. Res. Lett.*, **44**, 3806–3815, <https://doi.org/10.1002/2017GL073012>.
- Warner, J. L., J. A. Screen, and A. A. Scaife, 2020: Links between Barents–Kara sea ice and the extratropical atmospheric circulation explained by internal variability and tropical forcing.

- Geophys. Res. Lett.*, **47**, e2019GL085679, <https://doi.org/10.1029/2019GL085679>.
- Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932, [https://doi.org/10.1175/1520-0493\(2004\)132<1917:AARMMI>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2).
- Woods, C., and R. Caballero, 2016: The role of moist intrusions in winter Arctic warming and sea ice decline. *J. Climate*, **29**, 4473–4485, <https://doi.org/10.1175/JCLI-D-15-0773.1>.
- , —, and G. Svensson, 2013: Large-scale circulation associated with moisture intrusions into the Arctic during winter. *Geophys. Res. Lett.*, **40**, 4717–4721, <https://doi.org/10.1002/grl.50912>.
- Woollings, T., and M. Blackburn, 2012: The North Atlantic jet stream under climate change and its relation to the NAO and EA patterns. *J. Climate*, **25**, 886–902, <https://doi.org/10.1175/JCLI-D-11-00087.1>.
- , A. Hannachi, and B. Hoskins, 2010: Variability of the North Atlantic eddy-driven jet stream. *Quart. J. Roy. Meteor. Soc.*, **136**, 856–868, <https://doi.org/10.1002/qj.625>.
- Yang, W., and G. Magnusdottir, 2017: Springtime extreme moisture transport into the Arctic and its impact on sea ice concentration. *J. Geophys. Res. Atmos.*, **122**, 5316–5329, <https://doi.org/10.1002/2016JD026324>.
- Yoo, C., S. Lee, and S. B. Feldstein, 2012: Mechanisms of Arctic surface air temperature change in response to the Madden–Julian oscillation. *J. Climate*, **25**, 5777–5790, <https://doi.org/10.1175/JCLI-D-11-00566.1>.
- Yu, B., and H. Lin, 2016: Tropical atmospheric forcing of the wintertime North Atlantic Oscillation. *J. Climate*, **29**, 1755–1772, <https://doi.org/10.1175/JCLI-D-15-0583.1>.
- Zhang, J., R. Lindsay, M. Steele, and A. Schweiger, 2008: What drove the dramatic retreat of Arctic sea ice during summer 2007? *Geophys. Res. Lett.*, **35**, L11505, <https://doi.org/10.1029/2008GL034005>.
- Zhang, P., Y. Wu, I. R. Simpson, K. L. Smith, X. Zhang, B. De, and P. Callaghan, 2018: A stratospheric pathway linking a colder Siberia to Barents–Kara Sea sea ice loss. *Sci. Adv.*, **4**, 6025–6050, <https://doi.org/10.1126/sciadv.aat6025>.