Received: 1 October 2020 | Revised: 30 March 2021 | Accepted: 1 April 2021 | Published on: 25 May 2021

DOI: 10.1002/qj.4038

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Predictors and prediction skill for marine cold-air outbreaks over the Barents Sea

Iuliia Polkova*[1] | Hilla Afargan-Gerstman[2] | Daniela I. V. Domeisen[2] | Martin P. King[3] | Paolo Ruggieri[4,5] | Panos Athanasiadis[4] | Mikhail Dobrynin[1,6] | Øivin Aarnes[7] | Marlene Kretschmer[8] | Johanna Baehr[1]

[1]Institute of Oceanography, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, Germany

[2]Institute for Atmospheric and Climate Science, ETH Zürich, zürich, Switzerland

[3]NORCE Climate, and Bjerknes Centre for Climate Research, Bergen, Norway

[4]Euro-Mediterranean Center on Climate Change (CMCC), Bologna, Italy

[5]Department of Physics and Astronomy, University of Bologna, Bologna, Italy

[6]Deutscher Wetterdienst, Hamburg, Germany

[7]DNV-GL, Bergen, Norway

[8]University of Reading, Reading, UK

Correspondence
I. Polkova, Institute of Oceanography, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, Bundesstrasse 53, 20146 Hamburg, Germany.
Email: iuliia.polkova@uni-hamburg.de

Funding information
European Union's Horizon 2020 research and innovation programme Marie Skłodowska-Curie Swiss National Science Foundation Deutsche Forschungsgemeinschaft, Grant/Award Numbers: 727852, 841902, PP00P2_170523, 436413914

Abstract

Marine cold-air outbreaks (MCAOs) create conditions for hazardous maritime mesocyclones (polar lows) posing risks to marine infrastructure. For marine management, skilful predictions of MCAOs would be highly beneficial. For this reason, we investigate (a) the ability of a seasonal prediction system to predict MCAOs and (b) the possibilities to improve predictions through large-scale causal drivers. Our results show that the seasonal ensemble predictions have high prediction skill for MCAOs over the Nordic Seas for about 20 days starting from November initial conditions. To study causal drivers of MCAOs, we utilize a causal effect network approach applied to the atmospheric reanalysis ERA-Interim and identify local sea surface temperature and atmospheric circulation patterns over Scandinavia as valuable predictors. Prediction skill for MCAOs is further improved up to 40 days by including MCAO predictors in the analysis.

KEYWORDS

Arctic climate, causal drivers, marine cold air outbreak, polar low, seasonal prediction

1 | INTRODUCTION

Marine cold-air outbreaks (MCAOs) are surges of cold air escaping from the inner Arctic and becoming exposed to the relatively warm ocean surface (Rasmussen, 1983). MCAOs over the Nordic Seas usually occur within the northerly flow associated with large-scale anomalous anticyclonic circulation over Greenland and anomalous cyclonic circulation over Scandinavia (Kolstad *et al.*, 2009; Mallet *et al.*, 2013; Papritz and Grams, 2018; Afargan-Gerstman *et al.*, 2020). Given the importance of MCAOs for the formation of severe weather events such as polar lows (PLs; Rasmussen, 1983; Kolstad, 2011; Landgren *et al.*, 2019), and for driving the ocean convection in the North Atlantic (Kolstad and Bracegirdle, 2008), we evaluate predictability of MCAOs using a seasonal prediction system and by identifying MCAO precursors. We focus on the North Atlantic and in particular on the Barents Sea MCAOs as this is one of the regions that is exposed to frequent and strong MCAOs (Rasmussen, 1983; Fletcher *et al.*, 2016). Furthermore, this region exhibits a high vulnerability due to the presence of marine infrastructure with a need to improve marine services (Orimolade *et al.*, 2017; Aarnes *et al.*, 2018).

With respect to the role of MCAOs in the genesis of PLs, Ese *et al.* (1988) and Noer *et al.* (2011) describe MCAOs as a necessary yet not a sufficient condition. Sustained dry cold conditions are required for several days before PLs can develop. PLs may occur at the outer edges of MCAOs or as thermal-instability phenomena inside the cold air masses (Rasmussen, 1983; Terpstra *et al.*, 2016). Therefore, previous studies consider the diagnostic index used to detect MCAOs as a large-scale proxy for PLs (Claud *et al.*, 2007; Kolstad *et al.*, 2009; Landgren *et al.*, 2019). Whereas individual PLs can be predicted by nowcasting systems up to two days in advance (Kristiansen *et al.*, 2011), for efficient planning of marine activities, there is an urgent demand for information on PLs and MCAOs on longer lead times. In contrast to PLs, MCAOs are comparably more persistent and large-scale phenomena (PLs about 1–2 days, 200–1,000 km in diameter; MCAOs about 1–2 weeks, >1,000 km). Due to their larger temporal and spatial extent, MCAOs were assumed to be predictable for time horizons beyond a few days as compared to the less predictable PLs (Kolstad, 2017). However, no study has actually quantified prediction skill for MCAOs. Evaluating the prediction skill of the Nordic Sea MCAOs is therefore the major objective of our study.

To further explore potential sources of predictability of MCAOs on the sub-seasonal time-scale, we implement a causal discovery approach. The application of causal discovery algorithms has previously shown high potential for identifying causal Arctic drivers of the midlatitude winter circulation (Kretschmer *et al.*, 2016; Siew *et al.*, 2020), drivers of temperature extremes (Vijverberg *et al.*, 2020) and for understanding Arctic-stratospheric pathways and stratosphere–troposphere coupling (Kretschmer *et al.*, 2018a; Saggioro and Shepherd, 2019). The MCAO development stages identified by Papritz and Pfahl (2015) and Papritz and Spengler (2016) give a first impression of the chronology of a MCAO evolution. It starts with the (a) build-up and (b) advection of cold air masses (Papritz and Pfahl, 2015). In the mature stage, (c) the availability of warm air over the ice-free ocean contributes to MCAO erosion. By contrast, further supply of cold air can sustain MCAOs. Finally, (d) latent heating associated with mesocyclones such as PLs contributes to a complete decay of MCAOs. Pathways of the cold air and the ability of MCAOs to extract energy from the ocean are region-dependent and characterise lifetime and intensity of MCAOs in different regions of their occurrence (Papritz and Spengler, 2016). Using a causal discovery approach, we aim to test potentially causal relationships between the Barents Sea MCAOs and their environmental conditions. These factors include atmospheric circulation patterns enabling cold air flow from the Arctic region (Mallet *et al.*, 2013; Papritz and Grams, 2018; Afargan-Gerstman *et al.*, 2020), sea surface temperature changes as a proxy for ocean-to-air heat loss (Lien *et al.*, 2017) available to erode or fuel MCAOs, as well as sea ice concentration regulating the region of MCAO occurrence (Kolstad and Bracegirdle, 2008). Knowing causal precursors can then help to improve predictions of MCAOs as was shown by Dobrynin *et al.* (2018) and Cai *et al.* (2016) for predictability of the North Atlantic Oscillation (NAO) and cold spells, respectively. For instance, seasonal predictions of the winter NAO can be improved based on the knowledge of the NAO predictors from the previous season (Dobrynin *et al.*, 2018), whereas stratospheric variability can be used to predict low temperature extremes one month ahead (Cai *et al.*, 2016). Our expectation is that establishing causal links between MCAOs and other processes acting in the Arctic region can further improve prediction skill for regional MCAOs.

Thus, our study aims to contribute to knowledge on the predictability of MCAOs in the North Atlantic region and causal precursors of the Barents Sea MCAOs. In Section 2, we describe the MCAO index, prediction skill evaluation and the causal-detection approach. In Section 3, we analyse the prediction skill from the seasonal ensemble predictions performed with the Earth System Model from the Max Planck Institute for Meteorology (MPI-ESM), and provide details on the large-scale conditions that precede the anomalous MCAO variability. Finally, Sections 4 and

5 summarize the study and provide recommendations for future research.

2 | METHODOLOGY

2.1 | Data

We analyze MCAOs and conditions favourable for their development in the atmospheric reanalysis ERA-Interim (Dee *et al.*, 2011). The resolution of the atmospheric reanalysis is T255L60. ERA-Interim is also used as a source of initial conditions for seasonal predictions and a verification dataset for the prediction skill assessment. The prediction skill for MCAOs is evaluated from a 30-member ensemble of seasonal retrospective predictions (hereafter called hindcasts). The hindcast ensembles are carried out with the seasonal prediction system that is based on the Earth System Model from the Max Planck Institute for Meteorology in mixed resolution configuration (MPI-ESM-MR). The ESM consists of the atmospheric component ECHAM6 with a resolution of T63L95 and the oceanic component MPIOM with 0.4° horizontal resolution and 40 vertical levels (Giorgetta *et al.*, 2013). The hindcast ensembles are started every year on 1 November over the period 1980–2016 and provide data over the following six months. Ensemble members are generated by perturbing initial conditions in the ocean using bred-vectors (Baehr and Piontek, 2014) and perturbing the diffusion coefficient in the upper layer of ECHAM6. Full fields of the ERA-Interim atmospheric state and the ORAS4 ocean state (Balmaseda *et al.*, 2013) are nudged into the MPI-ESM-MR. The experimental set-up for seasonal predictions follows that designed by Baehr *et al.* (2015). The prediction experiments were previously used in the study on statistical-dynamical prediction for the NAO (Dobrynin *et al.*, 2018).

2.2 | Marine cold-air outbreak (MCAO) index

The air temperature over the Arctic sea ice can fall far below 0 °C while, for adjacent areas of open water, sea surface temperatures remain close to freezing. This creates a strong temperature gradient across the sea-ice edge. Transports of cold air masses from the sea ice toward the ocean create a large vertical temperature gradient (temperature difference between the ocean surface and atmosphere), which leads to surface sensible and latent heat flux from the ocean into the atmosphere. The MCAO index that we use here is from Kolstad (2017). It measures atmospheric instability conditions which cause convection within the boundary layer over open ocean. The daily MCAO index is then expressed in terms of the sea–air potential temperature difference:

$$MCAO(x,t) = \theta_{\text{sea surface}}(x,t) - \theta_{\text{air}}(x,t), \tag{1}$$

where $\theta_{\text{sea surface}}$ is ocean skin potential temperature (SKT) and θ_{air} is air potential temperature at 850 hPa. Indices x, t denote the space (latitude and longitude) and time dimensions (daily resolution), respectively. A positive $MCAO(x,t)$ index indicates atmospheric instability conditions. We then focus on MCAO events, that is, a positive $MCAO(x,t)$ index, during boreal autumn to spring months (November to April) in the Nordic Seas as this is a season and a region of high MCAO occurrence (Fletcher *et al.*, 2016).

Since the $MCAO(x,t)$ index measures vertical air column stability only over the ocean surface, sea ice grid cells, where SKT is less than −1.9 °C, and land grid cells for time-series analysis are masked. Note that masking or not masking an approximate edge of the sea ice can result in a different value of the long-term trend in the MCAO index as SKT also has values over the sea ice (not shown). Therefore, caution should be taken when comparing the results from different studies. Since the beginning of the analyzed time series in 1980, ERA-Interim shows a drop of sea-ice fraction over the Barents Sea and an upward trend for both SKT and air temperature at 850 hPa, leading to a slight downward trend in the time series for the Barents Sea MCAO index. However, a retreating sea-ice edge may also indicate that MCAOs do not disappear completely but rather occur to the northeast of the Barents Sea (not shown). In order to take into account an approximate sea-ice edge, we performed masking similar to Kolstad (2017). Alternatively, one could apply masking of actual sea-ice fraction above the value of 0.55. Both methods of the sea-ice masking are time varying and thus account for a temporal shift in the sea-ice edge.

In this study, we aim to investigate the merits of using a *seasonal* prediction system to predict MCAOs. Thus, in addition to synoptic characteristics of MCAOs, we analyze an aggregated-over-time quantity such as anomalous frequency of daily MCAO events per 10-day time window:

$$\begin{aligned} MCAO^{\text{freq}} &= \frac{1}{S}\sum_{x=1}^{S}\sum_{t=1}^{N_{\text{D}}} MCAO(x,t)|_{>0}, \\ \Delta MCAO^{\text{freq}} &= MCAO^{\text{freq}} - \frac{1}{N_{\text{Y}}}\sum_{t=1}^{N_{\text{Y}}} MCAO^{\text{freq}}, \end{aligned} \tag{2}$$

where $MCAO(x,t)|_{>0}$ is the positive daily MCAO index value, and x, t are space and time dimensions, respectively. N_{D} stands for number of days in a time-window, N_{Y} for years over the period 1980–2017, and S for number

of grid points in the respective domain. Space averaging is performed for MCAO time-series analysis, otherwise for MCAO maps we do not apply area averaging. The $MCAO^{freq}$ index considers all MCAOs within the time-window regardless of their strength and represents persistent or frequent MCAOs within a time window. Positive (negative) values of the aggregated anomalous index $\Delta MCAO^{freq}$ indicate a higher (lower) number of MCAO events as compared to the climatology of the $MCAO^{freq}$ index for the respective season.

We calculate daily and aggregated MCAO indices for both ERA-Interim and hindcast data. For the hindcasts, the aggregated $MCAO^{freq}$ index is calculated for each ensemble member individually and then the ensemble mean is constructed. For MCAO anomalies, the seasonal cycle is calculated over the period 1980–2017 from the ensemble mean and then removed from each ensemble member. A lead-time-dependent mean-bias correction is applied. The mean bias is calculated as the difference between the MCAO seasonal cycle of the ensemble mean and ERA-Interim, which is then subtracted from each ensemble member. For simplicity of handling the time axis for aggregated MCAO indices during skill assessment and predictor analyses, we use a 360-day calendar.

2.3 | Prediction skill metrics

Prediction skill is assessed by systematically comparing hindcasts and observations for each forecast time, that is, the lead time, to identify the margin (region and lead time), at which the hindcasts start to deviate from observations. Here, we use ERA-Interim instead of observational data. Skill is measured using probabilistic metrics such as relative operating characteristics (ROC) and Brier skill score (Jolliffe and Stephenson, 2012). In addition, the skill is assessed for the ensemble mean using correlation coefficients (COR; Jolliffe and Stephenson, 2012).

The ROC score is used to measure the ability of the hindcast to distinguish an event from a non-event. The MCAO event for the daily MCAO index is determined as $MCAO(x, t)$ being positive. For the aggregated index, the event is determined as a positive $\Delta MCAO^{freq}$ anomaly. An index value in the respective other category is classified as a non-event. The events and non-events are identified in the verification dataset (ERA-Interim). From the hindcasts, we calculate how many ensemble members have predicted or failed to predict an observed event. Based on these results, we can then further estimate the proportion of all events correctly predicted by the hindcasts, termed as the hit rate. The proportion of events that were predicted but not observed represents the false-alarm rate. If the prediction system has no skill in discriminating events and non-events, the hit and false-alarm rates will be identical and the hindcasts are concluded to have skill equivalent to climatology. However, if the prediction system has skill, the hit rate will be larger than the false-alarm rate. The ROC curve is plotted as hit rate against false-alarm rate for different probability categories. The ROC score is calculated as the area under the ROC curve. The hindcast with the skill equivalent to climatology will correspond to a curve that follows the diagonal with the ROC area being 0.5. Hence, for the skilful hindcasts, the ROC score will be higher than 0.5.

The Brier score is a summary verification measure for reliability and resolution of the hindcasts. It has a range from 0 to 1. The value of 0 is for the case that all hindcasts correctly predict the events and non-events with 100% probability. The opposite is true for a score of 1. Thus, low scores indicate better predictions than high scores. The Brier skill score compares the hindcast skill to that from the climatology. We have also estimated the uncertainty in the values of the prediction skill metrics using bootstrapped confidence intervals. For more details on forecast verification, we refer the reader to Mason and Graham (1999), Jolliffe and Stephenson (2012) and Mason (2013). We perform all prediction skill assessments with respect to ERA-Interim over the period November 1980 – April 2017 for daily and aggregated MCAO indices that are introduced in the previous subsection. We provide prediction skill plots for the North Atlantic and the Arctic sectors.

2.4 | Potential MCAO predictors

Here we first review Arctic and North Atlantic climate variability patterns which have been shown in the literature to be associated with MCAOs. In Section 3.3 we will evaluate these hypotheses using Causal Effect Network (CEN). The analysis of MCAO predictors is focused on the Barents Sea region. This is because previous studies showed that MCAO characteristics have regional peculiarities associated with specific local processes and orography (e.g., Papritz and Spengler, 2016). Therefore, causal drivers for MCAOs in other sea basins would need individual predictor assessment. Moreover, we focus on the Barents Sea region due to high marine activity and marine infrastructure which are exposed to frequent MCAOs.

Large-scale atmospheric circulation. The Barents Sea MCAOs usually develop within the northerly flow, which is enabled via anomalous anticyclonic circulation over Greenland and the anomalous cyclonic circulation over Scandinavia (Mallet *et al.*, 2013; Papritz and Grams, 2018). The stratospheric polar vortex variability and associated sudden stratospheric warming (SSW) events (Baldwin

et al., 2021) are a key driver of cold spells over northern Europe (e.g., Kolstad *et al.*, 2010; Kretschmer *et al.*, 2018b; King *et al.*, 2019). While, Papritz and Grams (2018) did not confirm that the strength of the polar vortex systematically affected MCAO frequency in the Nordic Seas, Afargan-Gerstman *et al.* (2020) show that more than half of SSW events in the ERA-Interim atmospheric reanalysis are associated with MCAOs in the Barents Sea. In their study, SSW events with an enhanced MCAO response in the Barents Sea are also linked with a ridge over Greenland and a trough over Scandinavia.

Ocean heat flux. Generation of MCAO events is linked to a warmer ocean surface relative to the air aloft and associated with heat and moisture flux into the atmosphere. Ocean sub-surface temperature changes in the Nordic Seas can be formed locally through changes in air–sea heat flux or advection from the North Atlantic. Whereas Atlantic water inflow into the Barents Sea acts at interannual and longer time-scales (Årthun *et al.*, 2017; Koul *et al.*, 2019), local/regional temperature changes occur on sub-seasonal and seasonal time-scales (Furevik, 2000; Lien *et al.*, 2017). In particular, Lien *et al.* (2017) advocates that an increase in the ocean heat content can have an effect on reduced refreezing on a time-scale of one month in the Barents Sea. They also suggest a link between the ocean heat transport anomaly and the positive NAO index which in turn may favour an increase in the number of low-pressure systems in the Nordic Seas. Kolstad *et al.* (2009) argues that SSTs are important in setting up the seasonal cycle of MCAOs and Claud *et al.* (2007) defines SST as an important condition for the PL development due to thermodynamic fluxes that fuel PLs.

Sea ice. By definition, MCAOs represent transports of cold air from the sea-ice or snow-covered areas toward the ice-free ocean. Thus, sea ice in the Barents Sea might be a necessary condition for generation of a cold-air pool in the first place. The variability of the sea-ice boundary is suggested to largely determine the area over which MCAOs occur (Papritz and Spengler, 2016; Spengler *et al.*, 2017; Landgren *et al.*, 2019). Michel *et al.* (2018) point out that there is no immediate relation between mesoscale cyclogenesis and the Barents Sea sea ice extent, however they indicate a shift in the cyclogenesis as the sea-ice retreats northward.

Based on these prevailing hypotheses of environmental conditions for MCAOs, we will analyze five potentially causal precursors on the sub-seasonal time-scale: changes in atmospheric circulation indices representing the Scandinavian pattern and Greenland blocking, stratospheric polar vortex index (PoV) variability, SST changes as for the upper-ocean state variability and changes in the Arctic sea ice cover (SIC).

2.5 | Causal effect network (CEN)

Previous studies addressing the relationship between MCAOs and atmospheric circulation indices employed composite and cross-correlation analyses (e.g., Kolstad and Bracegirdle, 2008; Kolstad *et al.*, 2009; Papritz and Grams, 2018). As these metrics do not provide information about causality, we here follow a study by Kretschmer *et al.* (2016) in which a causal discovery algorithm is applied to identify the causal links between different mid-latitude winter circulation indices. Thus, we apply the so-called Causal Effect Network (CEN) approach based on the PCMCI algorithm to analyse causal pathways of MCAO predictors. The underlying algorithm was designed and implemented by Runge *et al.* (2017; 2019) and is freely available through https://github.com/jakobrunge/tigramite (accessed 13 April 2021). The CEN algorithm encompasses the following steps (Kretschmer *et al.*, 2016 give a detailed description):

Step 1. Potential predictors of the target variable are identified using pairwise cross-correlation at different lags. Thus, the time-lagged cross-correlation is calculated between the MCAO index and its potential predictors as well as between all pairs of the different considered predictors. Also auto-correlation for each variable is evaluated. We use a maximum time lag of $\tau_{\max} = 9$ which corresponds to 90 days. All statistically significant cross-correlations (for a significance threshold of 0.1) are considered as potentially causal predictors. After this step, a set of MCAO predictors is obtained not only for MCAOs but for all the considered variables in the network.

Step 2. The potential predictors from Step 1 are tested for conditional independence with the target variable by iteratively conditioning on the combinations of the other input variables in the network. The conditional independence test performed by analyzing partial correlations in the iterative algorithm is a version of the Markov discovery algorithm. Those relationships, which remain to be significantly correlated with the target variable even when conditioning on other processes, are interpreted as causal drivers of the target variable. Similar to Step 1, each MCAO predictor is tested in the role of a predictand with other variables acting as its predictors. Step 2 identifies redundant or indirect MCAO drivers providing a refined set of MCAO predictors.

Step 3. Finally, the strength and significance of causal relationships for the set of MCAO predictors from

TABLE 1 Table of potential MCAO predictors (Input for CEN)

Abbreviation	Predictor	Variable (units)	Region (level)	Reference
Scand-Z500	Scandinavian pattern	Z500 (m)	60–70°N, 30–50°E (500 hPa)	Afargan et al. (2020)
Gr-Z500	Greenland pattern	Z500 (m)	60–70°N, 30–50°W (500 hPa)	Afargan et al. (2020)
local-SST	Inflow of ocean temperature	SST (K)	70–78°N, 20–54°E	Lien *et al.* (2017)
Arc-SIC	Arctic sea ice	SIC (%)	40–90°N	—
PoV-T100	Stratospheric polar vortex	T100 (K)	65–90°N (100 hPa)	Domeisen *et al.* (2020)

Note: All the time-series are constructed from the ERA-Interim reanalysis.
Abbreviations: SIC=sea ice concentration; SST=sea surface temperature; T100=air temperature at 100 hPa; Z500=geopotential height at 500 hPa.

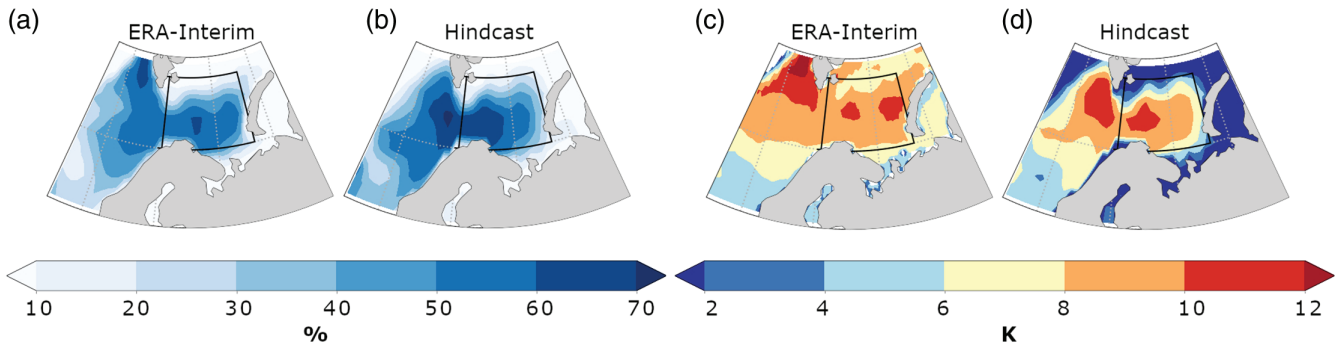

FIGURE 1 (a, b) Frequency of occurrence of MCAO events (in %). The MCAO event is defined as the daily MCAO index exceeding 0 K for longer than 2 days. (c, d) The 90th percentile of MCAO events (in K). Both metrics are plotted from (a, c) the ERA-Interim and (b, d) hindcast data covering November to March 1980–2017. The black outline shows the Barents Sea region, for which we later analyse the MCAO predictors

Step 2 is evaluated in terms of standardized multiple linear regression coefficients. Here, we define a two-tailed significance level $\alpha = 0.1$. This means that potential predictors beyond the maximum time lag and with confidence level below 90% are neglected.

Note that the term 'causal' should be taken with caution, as the identified links can only be interpreted as causal under certain assumptions and only within the set of considered variables (Runge, 2018). Nevertheless, for a reasonable choice of input data, the CEN approach has been shown to successfully remove spurious correlations thereby allowing a more causal interpretation of precursors. Here, the CEN analysis is based on the ERA-Interim data and employs time series of the MCAO index and its potential predictors. The input time series cover the period from January 1980 to December 2017. The target variable is $\Delta MCAO^{\text{freq}}$ for the Barents Sea (20–54°E; 70–78°N) region. Other input variables for the CEN analysis are represented by climatological anomalies, from which an area-weighted spatial average is calculated over the defined region (summarized in Table 1; the exact choice of the region will be explained in the results section). The long-term linear trend is removed as well.

3 | RESULTS

3.1 | MCAOs from ERA-Interim and seasonal predictions

First, to evaluate the model's skill in representing MCAOs, we compare different MCAO characteristics from the seasonal prediction system with those of ERA-Interim. We begin with the frequency of occurrence and the amplitude of extreme daily MCAO values from ERA-Interim and an ensemble member of initialized hindcasts (without bias-correction; Figure 1a,b). The hindcast shows a higher number of MCAO events than ERA-Interim for the period between November and March 1980–2017. The higher frequency of MCAO occurrence reflects the bias of the seasonal prediction system in the Arctic region. For instance, SST bias in winter consists of a warm bias in the western Barents Sea and cold bias in the eastern Barents Sea. Apart from February, hindcasts show a large-scale cold bias for air temperature at 850 hPa. Thus, a warmer ocean surface and colder air aloft might enhance the vertical temperature gradient, especially in the western Barents Sea. Further information on biases in the prediction system are shown in Appendix S1, Figure S1. In

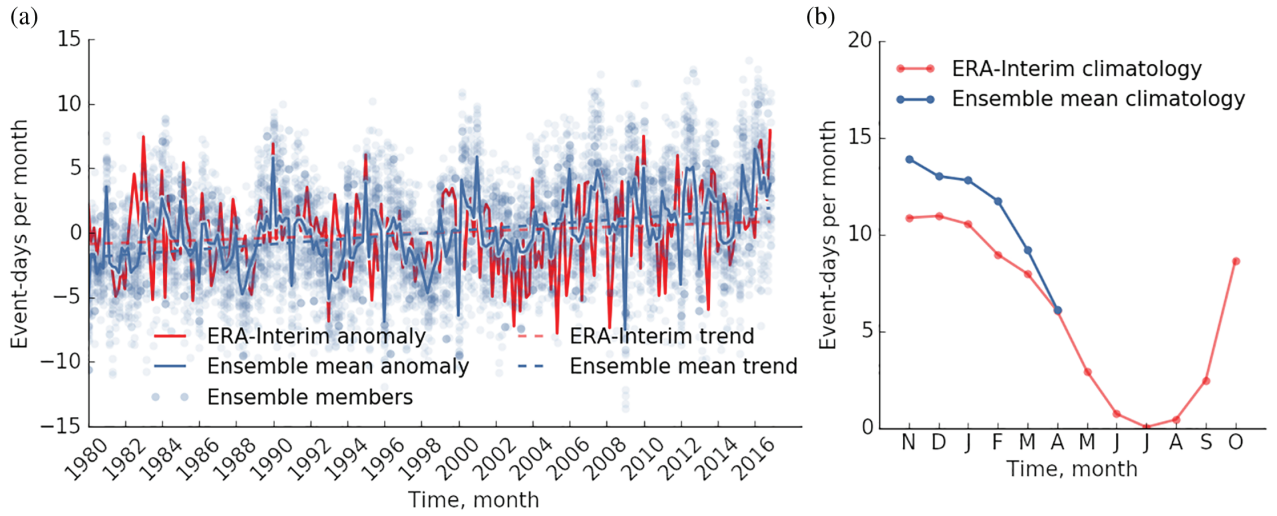

FIGURE 2 Time series of (a) anomalies (November to April) and (b) climatology (for the whole year) of the MCAO index, $\Delta MCAO^{freq}$, aggregated in 30-day time windows. ERA-Interim is shown in red and the hindcasts in blue (ensemble mean in solid and ensemble members in circles). Linear trends from ERA-Interim and the ensemble mean estimated over the 1980–2017 are also shown. The land–ice mask is applied to the daily MCAO index, from which $\Delta MCAO^{freq}$ is calculated

terms of the 90th percentile of the daily MCAO index as shown in Figure 1c,d, in regions of high MCAO occurrence the amplitudes of extreme MCAO values from both ERA-Interim and the hindcast are comparable, indicating that MCAOs from the seasonal prediction system are not necessarily more intense than those from ERA-Interim.

To gain insight into MCAO predictability on sub-seasonal and seasonal time-scales, we analyze the $\Delta MCAO^{freq}$ index aggregated in 10- and 30-day time windows (Section 2.2 gives details). $\Delta MCAO^{freq}$ exhibits low-frequency variability which is more noticeable in the 30-day aggregated index, which we choose to show in Figure 2 for the Barents Sea time series. The climatology of the $\Delta MCAO^{freq}$ index suggests that the seasonal prediction system overestimates the number of MCAO event-days per winter month. Also the time series of the index shows that individual ensemble members go far out of the range of the ERA-Interim variability; these results are consistent with Figure 1, which indicates that the seasonal prediction system tends to overestimate the frequency or the length of MCAO events as compared to ERA-Interim. The trend for $\Delta MCAO^{freq}$ increases, especially after the 2000s. Overall, the variability of the index is the highest during the boreal autumn and winter.

3.2 | Potential MCAO predictors

As noted by previous studies (Section 2.4), there is certain large-scale atmospheric flow and ocean state configurations in which MCAOs typically form over the Nordic Seas: the Scandinavian trough, Greenland ridge and Atlantic ridge patterns at synoptic time-scales (Kolstad *et al.*, 2009; Mallet *et al.*, 2013; Papritz and Grams, 2018), positive ocean temperature anomaly and ice-free ocean (Kolstad *et al.*, 2009). The time lag between the inflow of ocean temperature anomaly through the Barents Sea opening and the associated changes in sea-ice cover downstream was reported to be one month (Lien *et al.*, 2017, and references therein). Additionally, the stratosphere was proposed to drive the Barents Sea MCAOs within one month after SSW events (Afargan-Gerstman *et al.*, 2020). In this section, we first evaluate the relationship between these large-scale conditions and the Barents Sea MCAOs. Further, we test if the variability of any of these large-scale conditions precedes the MCAO variability such that the former can be considered as a precondition of the MCAO variability. For this purpose, we utilize the Barents Sea $\Delta MCAO^{freq}$ on 10-day resolution from the ERA-Interim reanalysis. We cross-correlate $\Delta MCAO^{freq}$ with the variables representing potential MCAO predictors at time lag 0 and −1 (10 days) for different months during high MCAO activity (November to April 1980–2017; Figures 3 and 4).

Consistent with Mallet *et al.* (2013) and Afargan-Gerstman *et al.* (2020), in cross-correlation patterns at time lag 0, we find dipolar anomalies of geopotential height at 500 hPa (Z500) pattern with a ridge over Greenland and a trough over Scandinavia and the Barents Sea. This relation holds from late autumn to early spring (Figure 3). The dipole pattern is also present in air temperature at 850 hPa with a warm anomaly over Greenland and a cold anomaly over the Barents Sea (not shown). For lagged cross-correlations at time lag −1 (i.e., 10 days), the cross-correlation between the Barents Sea MCAOs

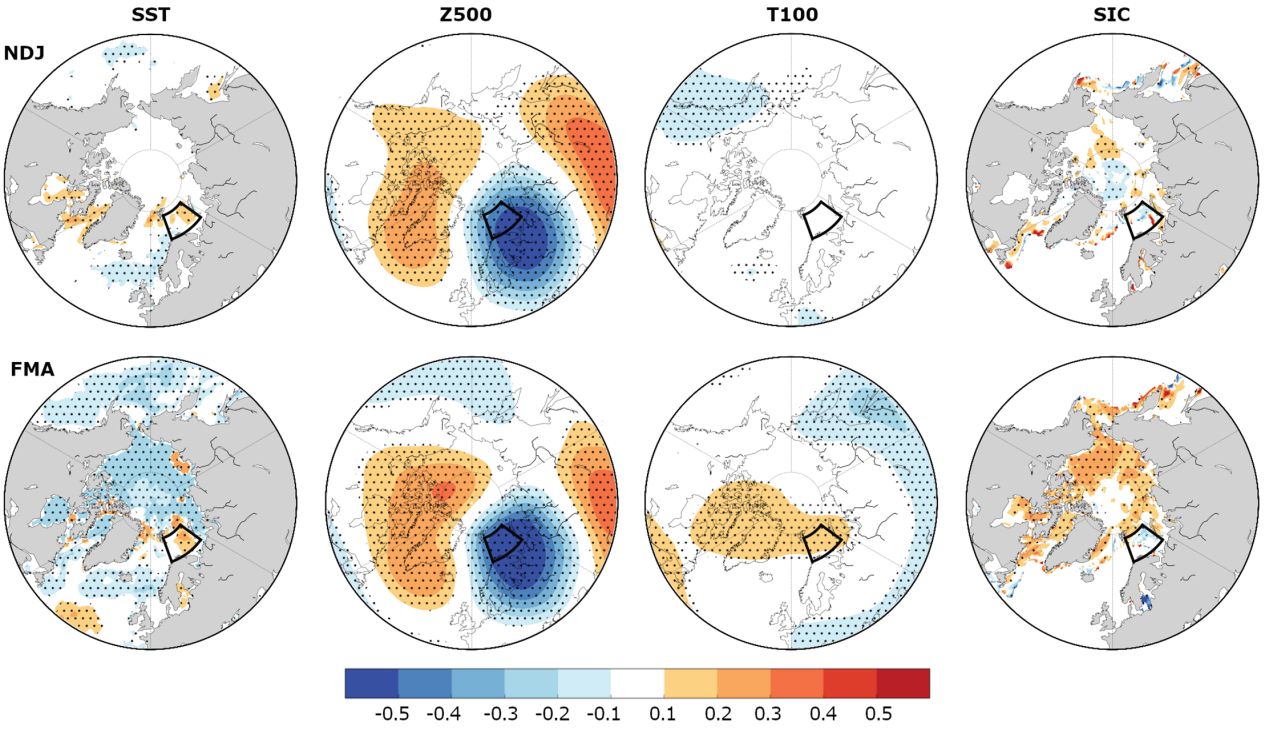

FIGURE 3 Cross-correlation between the time-series of $\Delta MCAO^{freq}$ for the Barents Sea and maps of anomalies for SST, Z500, T100 and SIC (in columns) for the beginning and the end of the MCAO peak season: NDJ (upper panels) and FMA (lower panels). The stippling indicates significant values at the 95% confidence level according to the two-tailed bootstrap test

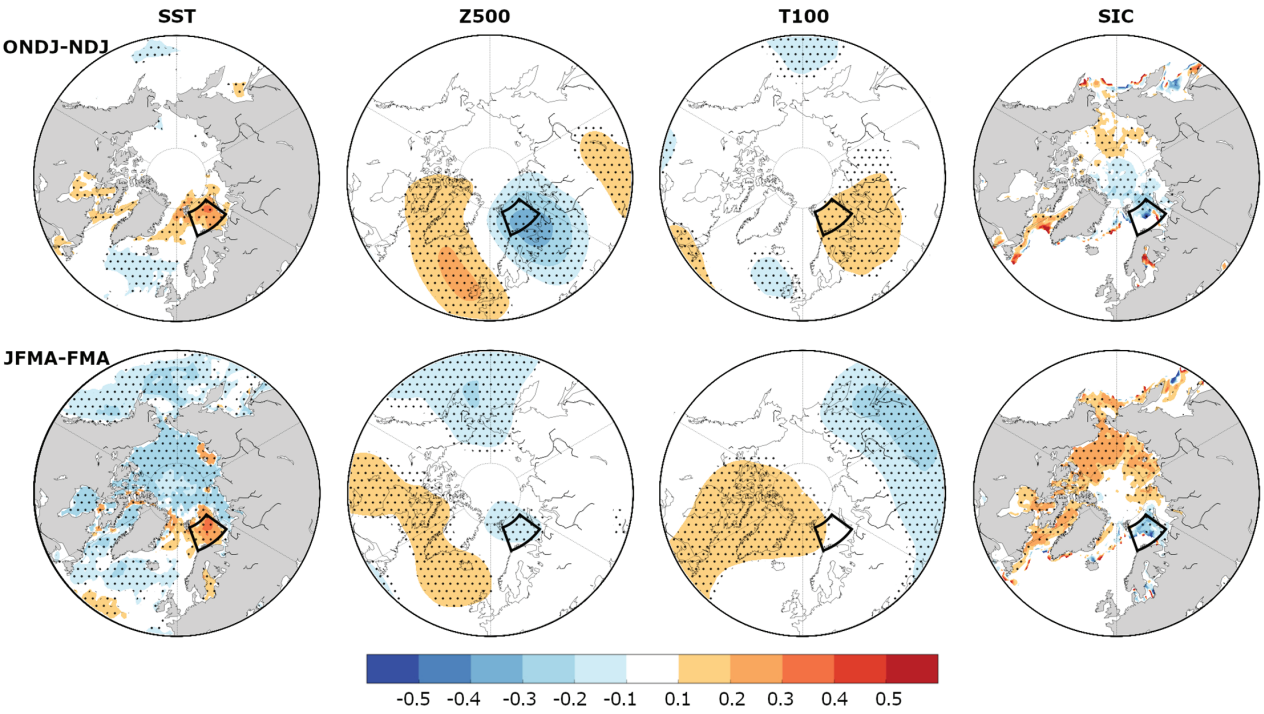

FIGURE 4 As Figure 3, but for SST, Z500, T100 and SIC leading $\Delta MCAO^{freq}$ changes by 10 days (lag −1)

and the dipole Z500 pattern is weaker and is more pronounced in autumn than in spring (Figure 4). The lag 0 cross-correlation with SST shows a positive relationship to the Barents Sea MCAOs at the sea-ice edge throughout the MCAO season (Figure 3). For time lag −1 (Figure 4), MCAOs are associated with warm SST in the whole

Barents Sea basin and, with longer time lag −2 (20 days) and −3 (30 days), the positive SST cross-correlation pattern extends into the Norwegian Sea (not shown). The lag 0 cross-correlation with the Arctic SIC also shows a positive relationship to the Barents Sea MCAOs mostly in spring (Figure 3). The relationship between the Barents Sea MCAOs and the Arctic SIC persists for lags −1 (Figure 4), −2 and −3 (not shown). Interestingly, the low SIC anomaly and warm SST anomaly in the Barents Sea are more correlated with the Barents Sea MCAOs at lag −1 than at lag 0. We use air temperature at 100 hPa (T100) as a representative of the polar vortex variability following Dobrynin *et al.* (2018) and Domeisen *et al.* (2020). In terms of T100, the cross-correlation at time lags 0, −1 and −2 is rather low and does not allow for the identification of a consistent pattern throughout the season. At lag −3 (30 days), the Barents Sea MCAOs positively cross-correlate with T100 over Greenland and Canada (not shown); this relationship is observed in late winter and early spring.

In summary, cross-correlation maps provide a first impression on relevant time-scales and regions. Except for T100, which does not show a consistent relationship to MCAOs, local SSTs and Arctic SIC have an effect on MCAOs in the Barents Sea. The SST–MCAO relationship holds throughout the MCAO peak season, also on longer time-scales. The dipole Z500 pattern (Greenland ridge and Scandinavian trough) relates to MCAOs on short time-scales. In addition, predictors for the Barents Sea MCAOs are not in a stationary relationship, meaning that there are seasonal variations in their influence on MCAOs. This finding is in agreement with previous studies which suggest that, for example, the frequency and intensity of blocking over the Northern Hemisphere is the highest during winter and somewhat lower in autumn and spring (Cheung *et al.*, 2013). A relationship between sea ice and the winter atmospheric circulation seems to be more pronounced in certain multi-year time periods (Kolstad and Screen, 2019). We further investigate conditional dependence of these links using the CEN algorithm (Section 2.5).

3.3 | Causal pathways of the MCAO predictors

Based on the results presented in Figures 3 and 4 and the review of relevant processes in Section 2.4, we analyse causal relationship in the time series of the Barents Sea MCAOs, $\Delta MCAO^{\text{freq}}$, and the following processes (Table 1): Scandinavian (Scand-Z500) and Greenland (Gr-Z500) circulation patterns at 500 hPa, the Barents Sea SST variability (local-SST), Arctic sea-ice variability (Arc-SIC), stratospheric polar vortex variability represented by air temperature anomaly at 100 hPa (PoV-T100). We perform the CEN analysis to identify which of the above processes is causally related to the MCAO changes in the Barents Sea. To address this, we chose a significance level $\alpha = 0.1$, and allow maximum lead/lag of $\tau_{\max} =$ 9 time-steps (90 days). Results from the CEN analysis are visualized in a network (Figure 5), consisting of the Barents Sea MCAOs, $\Delta MCAO^{\text{freq}}$, and their potential predictors stated in Table 1. The links between the variables indicate that a significant conditionally dependent relationship was detected, which is interpreted as the presence of a causal link. The time lag, at which the link is detected, is displayed in the network. The CEN analysis is conducted for the 10-day aggregated data for NDJ, DJF, JFM and FMA (Figure 5). This implies that the time series contain data for three months (e.g., for NDJ, time-steps include 1–10 November to 21–30 January), whereas the lagged variables contain data from the preceding period also represented by three months (e.g, lag −1 (10 days) accounts for time steps 21–30 October to 11–20 January, and lag −9 for 1–10 August to 21–30 October).

The CEN shows three robust links pointing towards the MCAO index: (a) local-SST over the whole season for the time lag of 10 days for NDJ to JFM and 20 days for FMA, (b) Scand-Z500 for months NDJ to JFM, and (c) Arc-SIC with time lag of 60 days for NDJ and DJF. The detected positive direct link from local-SST suggests that positive SST anomalies in the Barents Sea precede an anomalous MCAO frequency. Recall that a positive MCAO index, $\Delta MCAO^{\text{freq}}$, stands for a higher than usual number of MCAO event-days per 10-day window. The link from Scand-Z500 shows that a negative Z500 anomaly leads to an increase in the anomalous MCAO frequency at a lag of 10 days. The CEN further shows that Arc-SIC leads to changes in MCAOs at a lag of 60 days in autumn. Changes in Scand-Z500 appear to precede changes in Gr-Z500 (10–30 days), PoV-T100 (10 days) and Arc-SIC (20 days). PoV-T100 does not show any direct link to the Barents Sea MCAOs at the considered lags.

In summary, the results in Figures 3–5 suggest that the local Barents Sea SST anomalies and the atmospheric circulation index representing variability over Scandinavia can serve as the Barents Sea MCAO predictors almost throughout the whole MCAO season for a lag of 10 days. Thus, the previously reported large-scale atmospheric flow and ocean state configurations in which MCAOs typically form start setting up slightly before the anomalous MCAO activity. In the following, we evaluate the prediction skill for MCAOs and test if the prior knowledge on MCAO predictors can improve prediction skill on the sub-seasonal time-scale.

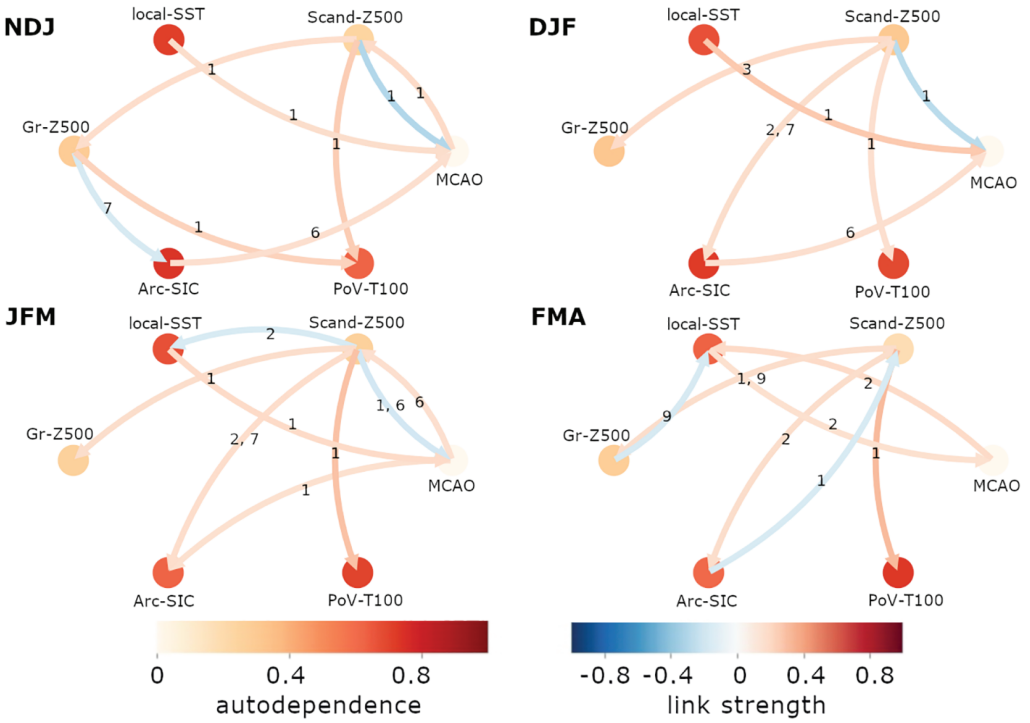

FIGURE 5 CEN of the Barents Sea MCAO drivers from ERA-Interim with a maximum time lag −9 (90 days) and a significance level of 0.1. Nodes represent variables considered (Table 1), while arrows indicate causal links between variables. The colour of nodes and links represent the standardized regression coefficient and indicate the strength of the relationship. The numbers next to arrows indicate time lag τ. Numbers separated by the comma indicate several lags at which the relationship is detected. CEN is constructed based on the time series of predictors for the Barents Sea MCAOs for the 10-day resolution for NDJ, DJF, JFM and FMA. For example, when the MCAO index is considered for JFM, predictors at maximum lag $\tau_{\max} = 9$ represent OND. Only the significant links are shown

3.4 | Prediction skill for seasonal MCAO predictions

The prediction skill for the aggregated index $\Delta MCAO^{\text{freq}}$ is calculated in terms of the ROC skill score and the correlation coefficient (COR) between the ensemble mean and ERA-Interim (Figure 6), and in terms of Brier skill score (Appendix S1, Figure S2). The different skill metrics consistently indicate a good agreement with ERA-Interim in the first 20 days after initialization over the Barents Sea, Labrador Sea and around Kamchatka Peninsula. High skill values are found north of the Barents Sea and the Kara Sea even beyond lead month one; in these months however the Kara Sea is substantially covered with sea ice. Interestingly, a re-emergence of skill appears over the Barents Sea in March in both ROC and COR metrics (shown for correlation skill in Appendix S1, Figure S4). The hindcasts outperform the persistence estimated from an AR(1) model by higher hit rates (Appendix S1, Figure S3). The spread diagnostic shows that the ensemble hindcasts for MCAOs at the beginning have little spread as indicated by the U-shaped Talagrand diagram, whereas the ensemble spread grows and saturates after one month (Appendix S1, Figure S3).

To better understand the performance of the seasonal prediction system, we provide correlation skill for SST, air temperature at 850 and 100 hPa, SLP and the meridional wind component (Figure 7). The correlation skill patterns suggest that the seasonal prediction system is skilful in predicting SST anomalies over the Barents Sea and the adjacent regions on seasonal time-scales. SLP and T100 in the polar region show rather low skill. However, T100 shows a re-emergence of skill in month five (March; not shown), which is potentially related to late-winter sudden stratospheric warmings or early final stratospheric warmings (Butler *et al.*, 2019), and which coincides with the re-emergence of skill for the MCAO index. The air temperature at 850 hPa remains skilful over parts of the Nordic Sea and the Labrador Sea up to lead month four. The meridional wind component is skilfully predicted only in the first lead month. Sea ice is one of the essential conditions for the MCAO development. We could not evaluate the skill of SIC as the model output is not available. However, Bunzel *et al.* (2016) suggest prediction skill for the Arctic sea-ice area for the November-initialized hindcasts up to three months in the MPI-ESM in a low-resolution configuration. The prediction skill of the MPI-ESM in a mixed resolution configuration shown in Figure 7

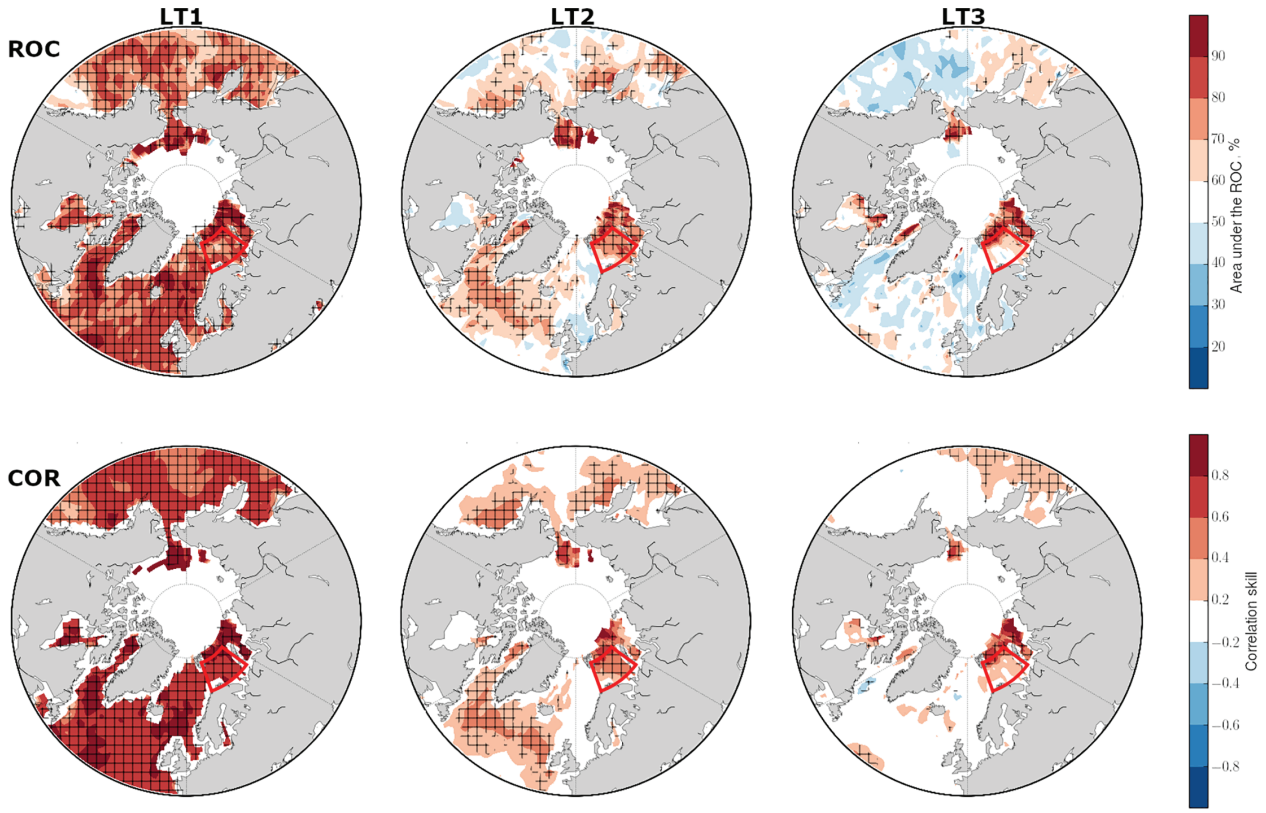

FIGURE 6 Prediction skill for the 10-day $\Delta MCAO^{freq}$ index in terms of the ROC skill score (upper panels) and correlation between the ERA-Interim and hindcast ensemble mean (lower panels) for lead times LT1 (1–10 November), LT2 (11–20 November) and LT3 (21–30 November). The event definition for the aggregated index used for the ROC is a positive value of $\Delta MCAO^{freq}$. The hatching indicates significant values at the 95% confidence level following the bootstrap test

compares to the predictions skill of the state-of-the-art ESMs (Butler *et al.*, 2016; Sanna *et al.*, 2016; Vitart and Balmaseda, 2018). The seasonal prediction system in the mixed-resolution configuration also compares in the level of skill with its siblings in low- and high-resolution configurations (Domeisen *et al.*, 2015; Fröhlich *et al.*, 2021).

In summary, the prediction skill from the seasonal prediction system for MCAOs in the regions where they frequently occur is limited to the first 10–20 lead days. Even though slow-varying components of the climate system (represented by SST and sea ice) are skilfully predicted on seasonal time-scales, MCAOs are only skilfully predicted on sub-seasonal time-scales, which is likely due to the faster changing atmospheric circulation. In the following, we assess if the prior knowledge about MCAO predictors has an impact on the prediction skill of the seasonal MCAO hindcasts. As MCAO predictors we use the causal drivers confirmed by the CEN analysis.

3.5 | Improving MCAO predictions

Dobrynin *et al.* (2018) used NAO predictors in order to improve winter NAO predictions one season in advance. This was achieved by sub-selecting ensemble members of the NAO seasonal predictions that captured the phase of the NAO as indicated by the autumn ocean temperature, stratospheric circulation and Siberian snow cover as predictors. Similarly, Cai *et al.* (2016) used the stratospheric state as a predictor for low temperature extremes resulting in improved prediction skill a month ahead. In this section, we analyse the prediction skill for the sub-sampled $\Delta MCAO^{freq}$ seasonal predictions following the idea by Dobrynin *et al.* (2018). We use the knowledge about the empirical relationship between the Barents Sea MCAOs and their precursors as shown by the CEN. First, we combine the knowledge of multiple predictors in one time series termed the first-guess MCAO prediction. For example, as suggested by the CEN (Figure 5), the warmer SST anomaly (positive link) and the anomalous cyclonic circulation over Scandinavia (negative link) are usually followed by the increased anomalous MCAO frequency at the time lag of 10 days. We combine these multiple predictors into the first guess using the multiple linear regression expressed as

$$Y_t = \sum_i \beta_i X_i + \varepsilon_Y, \tag{3}$$

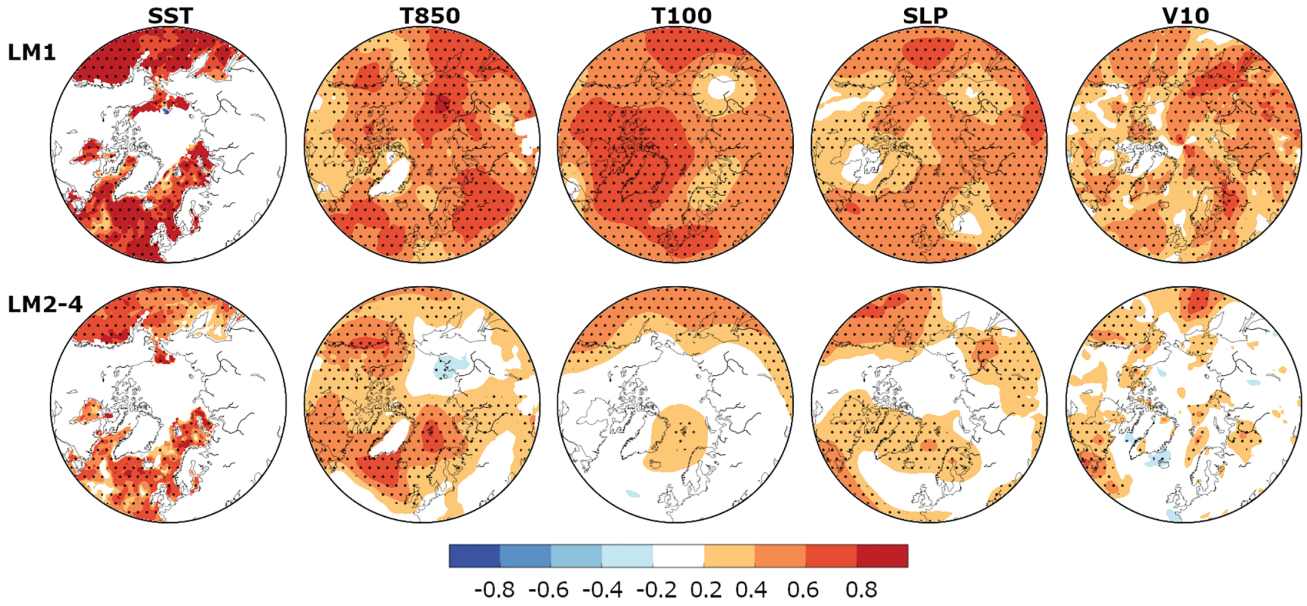

FIGURE 7 Prediction skill in terms of correlation with respect to ERA-Interim for SST, air temperature at 850 and 100 hPa, SLP, and the meridional wind component at 10 m (in columns) for lead months LM1 (November) to LM2-4 (December-January-February). The stippling indicates significant skill at the 95% confidence level estimated with the bootstrap test

where Y_t is the time-series of the first guess for $\Delta MCAO^{\text{freq}}$, X_i is the ith predictor, and β_i is the ith standardized regression coefficient. As we are using standardized time series, coefficients in the regression models can be interpreted as the expected change in the MCAO index (in units of standard deviation; STD) after a change of 1 STD of the predictor. For instance, $\beta = 0.2$ suggests that a change in a causal driver of 1 STD at a considered time lag leads to a change in $\Delta MCAO^{\text{freq}}$ of 0.2 STD. Finally, ε_Y represents the error term.

To construct the first guess for the Barents Sea MCAOs based on the MCAO drivers from the preceding 10 days, we repeat step 3 from the CEN analysis (Section 2.5) with the maximum lag −1. The regression coefficients and the underlying first-guess predictions for NDJ, DJF, JFM and FMA are provided in Table 2 and Figure 8, respectively. Here, changes in local SST, Scandinavian pattern and Arctic SIC are the predictors with the strongest contribution. Over the course of the MCAO season from NDJ to FMA, the contribution from the Scandinavian pattern vanishes with regression coefficients changing from −0.38 for NDJ to −0.21 for JFM. The most robust relationship is with the local SSTs, which is about 0.4 for almost all of the months. Greenland blocking and Arctic SIC contribute to $\Delta MCAO^{\text{freq}}$ changes for some of the months.

The resulting first-guess predictions correlate with the original $\Delta MCAO^{\text{freq}}$ from ERA-Interim by about 0.5 for all considered months (Figure 8). The period of evaluating regression coefficients presented in Table 2 is split into training and testing periods as shown by red and orange colours in Figure 8. The training and testing periods are selected arbitrarily as 80 and 20% of the time series, respectively. The correlation between the first-guess predictions and ERA-Interim decreases in the testing period with correlation coefficients amounting to 0.27–0.45 as compared to the full period with correlation coefficients of 0.45–0.5. Nevertheless, the quality of the first guess is still remarkable considering the high-frequency variability nature of MCAOs. Thus, the first-guess predictions resemble the low-frequency variability of the original $\Delta MCAO^{\text{freq}}$ rather well.

TABLE 2 Standardized regression coefficients for MCAO predictors at time lag −1 from the multiple linear regression model for $\Delta MCAO^{\text{freq}}$ (Figure 8)

Predictors	NDJ	DJF	JFM	FMA
Scand-Z500	−0.38	−0.31	−0.21	—
Gr-Z500	0.13	0.12	—	0.15
local-SST	0.38	0.42	0.45	0.43
Arc-SIC	—	—	0.16	0.30
$\Delta MCAO^{\text{freq}}$	—	—	—	0.11

The obtained first-guess prediction based on multiple linear regression models can be used as a stand-alone statistical MCAO prediction. Following the method by Dobrynin *et al.* (2018), which retains or rejects individual ensemble members based on the causal links, we use the first guess to select ensemble members that are

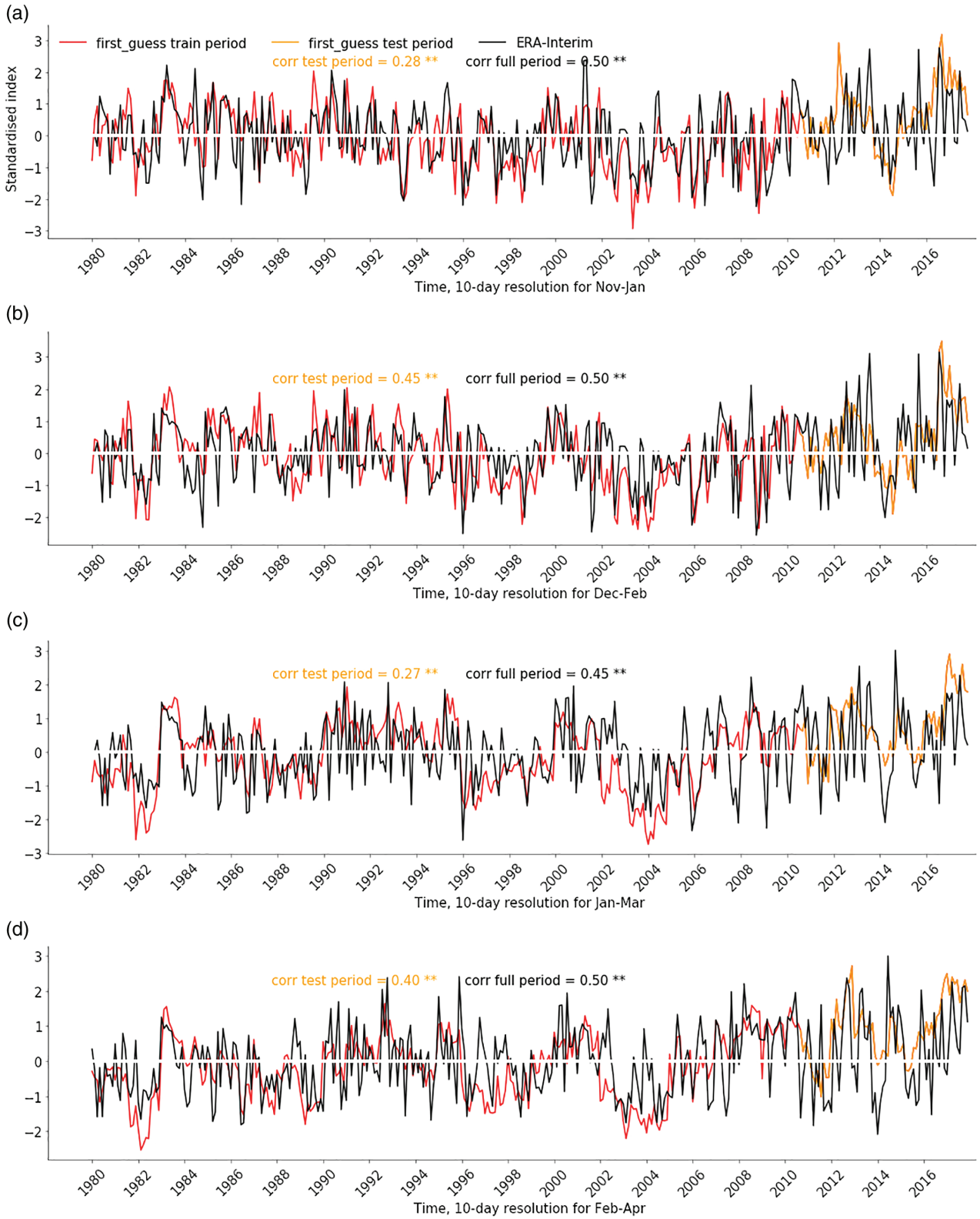

FIGURE 8 Standardized first-guess predictions for $\Delta MCAO^{freq}$ at lag −1 for (a) NDJ and (b) DJF based on Scand-Z500, Gr-Z500 and local-SST, for (c) JFM based on Scand-Z500, local-SST and Arc-SIC and for (d) FMA based on Gr-Z500, local-SST, Arc-SIC and $\Delta MCAO^{freq}$. Table 2 shows predictor contributions in terms of STD. In black is the original $\Delta MCAO^{freq}$ from ERA-Interim. In red (orange) is the first guess from the training (testing) period used to evaluate regression coefficients. The number with asterisks suggests a significant correlation (corr) between the first guess and ERA-Interim estimated with the two-tailed statistical significance at the 90% confidence level based on the *t*-test

nearest to it. Thus, we pick a subset of 15 out of 30 members that are closest to the first guess. The retained ensemble members capture the MCAO evolution as suggested by the MCAO causal predictors. Moreover, the subsampled ensemble has a further advantage over statistical prediction in that one can analyse further variables such as winds, sea level pressure, etc., whereas a statistical prediction only provides information about MCAOs. As our seasonal predictions begin on 1 November, we choose the date from the first guess (Figure 8) which corresponds to the time lag −1 (10 days before initialization). Assuming that a physical mechanism that links the initial state and the predicted variable can be expressed by the first guess, we evaluate skill of the retained ensemble members at different lead times. The results suggest that predictions improve at lead time 2 (11–20 November) as compared to those presented in Figure 6, however only marginally (not shown). This is likely due to low memory in the MCAO index, that is, fast de-correlation time; the ensemble members disperse within few weeks as was shown earlier by the Talagrand diagram (Appendix S1, Figure S3). Thus, even in the subsampled ensemble, spread characteristics become to resemble those of the full ensemble within 1–2 lead times (20 days).

Analysing individual Barents Sea MCAO predictors shows that local-SST correlates with MCAOs up to 2 months (Figure 9), which is likely due to high auto-correlation (persistence) of SST. The MCAO CEN provides links excluding autocorrelation in MCAO predictors. The auto-dependencies are provided separately in the nodes of the CEN. Prediction studies show that persistence is a valuable source of prediction skill. For instance, persistence of soil moisture provides skill for precipitation (Yang and Wang, 2019), and persistence of ocean heat content provides skill for the North Atlantic SST (Polkova *et al.*, 2014). Thus, in the final test, we evaluate the predictor with high auto-correlation, namely local SST, in the role of the first-guess prediction on time lags beyond 10 days (Figure 9). For this, we select ensemble members at each lead time within 2 months according to the local SSTs from 21–30 October. The results show that this exercise improves MCAO skill up to lead time 4 (1–10 December). The subsampled prediction at lead time 4 and the first-guess prediction at lag −4 (local SST from 21–30 October) correlate with the original MCAO index from ERA-Interim at 0.57 and 0.58, respectively; whereas the full ensemble has no significant skill at this lead time (correlation skill is 0.27). The subsampled ensemble according to local SSTs far outperforms the ensembles of randomly subsampled ensemble members (Appendix S1, Figure S5). Thus, ensemble members that follow an expected evolution according to the local SST predictor are able to improve MCAO prediction skill up to 40 days (Figure 10) as compared to the skill of 20 days from the seasonal prediction system (Figure 6).

Our results show that the CEN approach proves to be a useful tool in confirming robust drivers or early-warning changes for MCAOs in terms of the large-scale atmospheric flow and ocean state configurations. The identified precursors can be used for statistical MCAO predictions independently from the seasonal prediction system. Ensemble subsampling technique works if either predictand itself holds sufficient memory from one lead time to another or if long-term variability of a predictor can be associated with predictand as in the example here, where local SST provides memory for the Barents Sea MCAOs up to lead time 4 (40 days).

4 | SUMMARY AND DISCUSSION

Our analysis confirms the relationship between the Barents Sea MCAOs and an anomalous cyclonic circulation pattern over Scandinavia, and an anomalous anticyclonic pattern over Greenland. These findings are in agreement with previous studies (Mallet *et al.*, 2013; Papritz and Grams, 2018; Landgren *et al.*, 2019; Afargan-Gerstman *et al.*, 2020). In terms of conditions preceding changes of the Barents Sea MCAOs, we find a causal link from local SSTs and from the Scandinavian pattern at a lead time of 10 days. The inflow of SST anomaly into the Barents Sea might lead anomalous MCAO frequency on a time-scale of up to 2 months. This is comparable to the time-scales reported by Lien *et al.* (2017) on the response of sea ice cover to inflow of temperature anomaly. Our analysis shows that changes in the Arctic sea-ice concentration also lead the changes in the anomalous frequency of the Barents Sea MCAOs. However, as pointed out by Kolstad and Screen (2019), empirical relationships between the atmospheric circulation and sea-ice variability used for statistical prediction approaches are non-stationary and thus subject to change over time; in particular they questioned the relationship between the fall Barents–Kara sea ice and the winter NAO. We also find that the non-stationary relationships between MCAOs and their predictors present a challenge for deriving the first-guess prediction for MCAOs. Moreover, data quality is critical in establishing empirical relationships. For this reason, Kretschmer *et al.* (2016), whose approach we followed here to identify MCAO precursors, used various observational data for sea ice and snow cover, whereas all other variables were analysed using the ERA-Interim data. In our analysis, we used ERA-Interim for MCAOs and all predictors for consistency. However, it should be noted that ERA-Interim does not analyse sea-ice observations directly; instead, it incorporates external analyses of SIC (until 2009 – the US

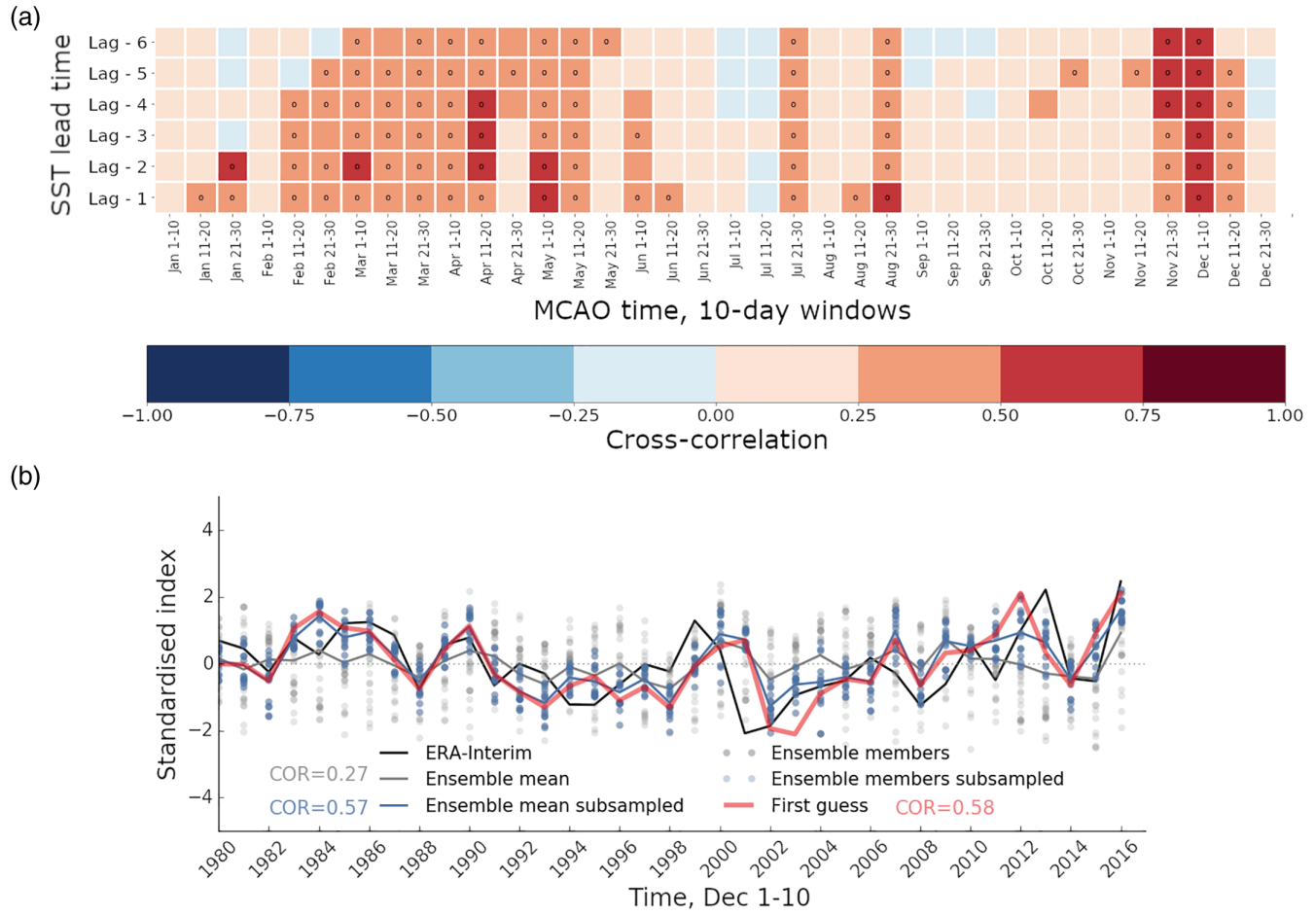

FIGURE 9 (a) Cross-correlation between the Barents Sea $\Delta MCAO^{freq}$ and the local SST at different 10-day lead times, when SST changes lead the MCAO changes. Both variables are from ERA-Interim and are represented in a 10-day resolution, that is, lag −4 corresponds to SST leading changes in MCAOs by 40 days. The stippling suggests significant values estimated with the two-tailed statistical significance at the 90% confidence level based on the *t*-test. (b) Standardised time series for the sub-selection of ensemble members nearest to the first guess based on local SST changes and the corresponding correlation skill (COR) for lead time 4 (1–10 December). $\Delta MCAO^{freq}$ calculated using ERA-Interim is shown in black, first guess based on local SST in red, the full ensemble from the seasonal prediction system in grey (circles for ensemble members and solid curve for the ensemble mean). The subsampled hindcast is in blue

National Centers for Environmental Prediction (NCEP) operational analyses, and from 2009 onwards – the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) product). Thus, we are treating Arctic SIC as a MCAO predictor with caution and suggest further investigation in future studies.

We used a causal effect network (CEN) to analyse causal links between MCAOs and their environment. It is important to note that the interpretation of causal links in the CEN depends on the assumptions of causal sufficiency and causal stationarity of links (Runge 2018 has a detailed discussion). The sufficiency assumption means that all relevant drivers are included in the analysis. Although we have analysed important contributors suggested by the literature, we cannot exclude the possibility that other processes or time-scales not considered here might also play a role. For example, the strong southwesterly winds accompanied by synoptic storms during the positive NAO phase can reach the Norwegian and Barents Seas and thus are possible conditions for MCAOs Claud *et al.* (2007), which we did not investigate here. Causal stationarity represents another challenge: though it was shown in previous studies that Greenland blocking and Scandinavian pattern are linked to MCAOs (Papritz and Grams, 2018; Landgren *et al.*, 2019; Afargan-Gerstman *et al.*, 2020), simple averages over stationary boxes to capture blocking activity has its limitations. Here, a robust connection between the Barents Sea MCAOs and air temperature at 100 hPa representing stratospheric forcing was not found. It is possible that this indicator might not be a good choice to represent the stratopsheric signal, or that the stratosphere drives MCAOs via other processes, namely through the

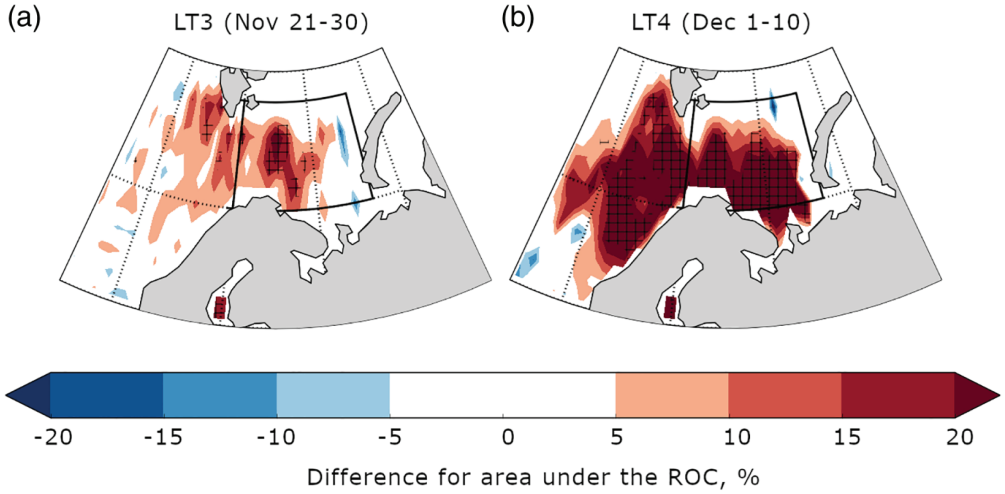

FIGURE 10 ROC skill score difference between the subsampled and full ensemble for (a) lead time 3 (21–30 November) and (b) lead time 4 (1–10 December). The subsampled ensemble follows the first-guess prediction based on local-SST. The hatching indicates significant values at the 95% confidence level following the bootstrap test

geopotential height pattern, which would support findings by Afargan-Gerstman *et al.* (2020).

The seasonal prediction system exhibits a bias which might cause a higher MCAO activity in seasonal hindcasts as compared to ERA-Interim. We used a simple mean bias-correction procedure to de-bias MCAO hindcasts, yet a bias treatment might further complicate the predictor analysis in the prediction system. We find that the cross-correlations between the anomalous MCAO frequency and potential MCAO drivers in the ERA-Interim and the hindcasts agree only in November and December (not shown). We did not investigate if an advanced bias-correction procedure could improve this result. Apart from that, the peak season for MCAOs and the MCAO formation regions in both ERA-Interim and seasonal predictions are in agreement and in line with the results from previous studies (e.g., Kolstad *et al.*, 2009; Fletcher *et al.*, 2016).

The seasonal prediction system shows high skill for MCAOs in the first 20 days, which supports expectations by Kolstad (2017) in MCAOs being more predictable than PLs. The Barents Sea MCAO precursors found by CEN, especially SST and SIC, show skill at longer time-scales than MCAOs. We find that the Barents Sea SST representing upper-ocean heat anomalies with time-scales of variability longer than MCAOs has a potential to serve as predictor for the anomalous MCAO frequency. We can use knowledge about the empirical relationship between local SST changes and the Barents Sea MCAOs to anticipate the future evolution of the MCAO index. Following ensemble members which capture the anticipated evolution of MCAOs based on the local SST predictor, the prediction skill for MCAOs can be extended up to 40 days. For the ensemble subsampling approach, we separated the time series into training and testing periods once, but the most strict way would be to test predictors on the data periods preceding the initialization dates of the seasonal hindcasts for each year.

Our study supports some of the known facts of MCAO large-scale conditions and reveals new aspects such as predictability time-scales from a state-of-the-art seasonal prediction system. We also reveal further important research objectives for future studies. One of them is the evaluation of predictability and predictors throughout the whole year, for example, to understand skills from initialisation in other months than November. A similar analysis of skill assessment for February initialisation showed that MCAOs are less predictable when starting from this date as compared to November initialisation, thus indicating that predictability of MCAOs is sensitive to the initialisation date. A further aspect is related to sufficient resolution for relevant applications. Resolution can also play a role for a more realistic representation of the mechanisms in the region. For instance, using high-resolution simulations Saggiorato *et al.* (2020) showed that convective momentum transport has an effect on the evolution of winds and clouds in a MCAO. Moreover, Papritz and Pfahl (2015) indicated that insufficient model resolution might result in the overestimation of lifetime of MCAOs. We used the MPI-ESM in the mixed resolution (T63L95/TP04L40) configuration, which at the time of analysis showed the best skill for the North Atlantic sector as compared to the high-resolution (T127L95/TP04L40) configuration. However, prediction systems are in constant development, and for future studies one might need to consider the system with higher resolution and reasonable skill in the Euro-Atlantic and Arctic sector not only for MCAOs but also for other large-scale conditions that are relevant for MCAO development. Translating the prediction of MCAOs into statistics on extreme weather such as polar lows is also of practical importance for climate practitioners.

5 | CONCLUSIONS

To summarize, the main conclusions of our study are

1. The seasonal prediction system shows skill for MCAOs over the North Atlantic region for about 10–20 days.
2. Using a causal effect network (CEN) approach, we identify local SST changes and the Scandinavian pattern as robust precursors of the anomalous MCAO frequency in the Barents Sea at a time lag of 10 days. Positive SST anomalies in the Barents Sea and the Norwegian Sea are further found to precede enhanced MCAO activity in the Barents Sea at longer time-scales up to two months.
3. The MCAO prediction skill of the seasonal prediction system can be extended to up to 40 days by sub-selecting ensemble members suggested by the local SST anomalies.

Marine services in the Barents Sea can benefit from using the sub-seasonal MCAO predictions up to 20 days (and up to 40 days when additionally considering the local SST anomalies). Overall, our study contributes to a growing body of literature demonstrating that combining empirical and dynamical predictions can result in more skilful forecasts. This provides new possibilities for decision makers in weather-sensitive sectors. Furthermore, this work shows that understanding the causal drivers of weather and climate events is a crucial step for this task.

ACKNOWLEDGEMENTS

Funding was provided by the Blue-Action project from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727852. Support from the Swiss National Science Foundation through project PP00P2_170523 to D.D. and H.A.-G. is gratefully acknowledged. M.K. has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement [No 841902]. I.P. acknowledges funding from the Deutsche Forschungsgemeinschaft, Project number 436413914. We thank Jacob Runge for the *tigramite* python package (CEN analysis) and David Nielsen for the *polarplots* python package. We also thank Sem Vijverberg and two anonymous reviewers for helpful discussions. Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ORCID

*Iuliia Polkova** https://orcid.org/0000-0003-4940-050X

Mikhail Dobrynin https://orcid.org/0000-0003-3533-3529

REFERENCES

Aarnes, Ø., Kazuyoshi, M. and Mejlaender, L.M. (2018) End-user requirements specification report Blue-Action case study Nr. 3 (D5.11). *Zenodo*. https://doi.org/10.5281/zenodo.1164217.

Afargan-Gerstman, H., Polkova, I., Ruggieri, P., King, M.P., Papritz, L., Athanasiadis, P., Baehr, J. and Domeisen, D.I.V. (2020) Stratospheric influence on marine cold air outbreaks in the Barents Sea. *Weather and Climate Dynamics*, 1, 541–553. https://doi.org/10.5194/wcd-1-541-2020.

Årthun, M., Eldevik, T., Viste, E., Drange, H., Furevik, T., Johnson, H.L. and Keenlyside, N.S. (2017) Skillful prediction of northern climate provided by the ocean. *Nature Communications*, 8. https://doi.org/10.1038/ncomms15875.

Baehr, J. and Piontek, R. (2014) Ensemble initialization of the oceanic component of a coupled model through bred vectors at seasonal-to-interannual timescales. *Geoscientific Model Development*, 7(1), 453–461.

Baehr, J., Fröhlich, K., Botzet, M., Daniela, I.V.D., Kornblueh, L., Notz, D., Piontek, R., Pohlmann, H., Tietsche, S. and Müller, W.A. (2015) The prediction of surface temperature in the new seasonal prediction system based on the MPI-ESM coupled climate model. *Climate Dynamics*, 44(9–10), 2723–2735.

Baldwin, M.P., Ayarzaguena, B., Birner, T., Butchart, N., Butler, A.H., Charlton Perez, A.J., Domeisen, D.I.V., Garfinkel, C.I., Garny, H., Gerber, E.P., Hegglin, M.I., Langematz, U. and Pedatella, N.M. (2021) Sudden stratospheric warmings. *Reviews of Geophysics*, 59(1). https://doi.org/10.1029/2020RG000708.

Balmaseda, M.A., Mogensen, K. and Weaver, A.T. (2013) Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, 139, 1132–1161.

Bunzel, F., Notz, D., Baehr, J., Müller, W.A. and Fröhlich, K. (2016) Seasonal climate forecasts significantly affected by observational uncertainty of Arctic sea ice concentration. *Geophysical Research Letters*, 43(2), 852–859.

Butler, A.H., Arribas, A., Athanassiadou, M., Baehr, J., Calvo, N., Charlton-Perez, A., Déqué, M., Domeisen, D.I.V., Fröhlich, K., Hendon, H., Imada, Y., Ishii, M., Iza, M., Karpechko, A.Y., Kumar, A., MacLachlan, C., Merryfield, W.J., Möller, W.A., O'Neill, A., Scaife, A.A., Scinocca, J., Sigmond, M., Stockdale, T.N. and Yasuda, T. (2016) The climate-system historical forecast project: do stratosphere-resolving models make better seasonal climate predictions in boreal winter?. *Quarterly Journal of the Royal Meteorological Society*, 142, 1413–1427.

Butler, A.H., Perez, A.C., Domeisen, D.I.V., Simpson, I.R. and Sjoberg, J. (2019) Predictability of Northern Hemisphere final stratospheric warmings and their surface impacts. *Geophysical Research Letters*, 46. https://doi.org/10.1029/2019GL083346.

Cai, M., Yu, Y., Deng, Y., van den Dool, H.M., Ren, R., Saha, S., Wu, X. and Huang, J. (2016) Feeling the pulse of the stratosphere: an emerging opportunity for predicting continental-scale cold-air outbreaks one month in advance. *Bulletin of the American Meteorological Society*, 97(8), 1475–1489.

Cheung, H.N., Zhou, W., Mok, H.Y., Wu, M.C. and Shao, Y. (2013) Revisiting the climatology of atmospheric blocking in the Northern Hemisphere. *Advances in Atmospheric Sciences*, 30(2), 397–410.

Claud, C., Duchiron, B. and Terray, P. (2007) Associations between large-scale atmospheric circulation and polar low developments over the North Atlantic during winter. *Journal of Geophysical Research: Atmospheres*, 112(D12). https://doi.org/10.1029/2006JD008251.

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

Dobrynin, M., Daniela, I.V.D., Müller, W.A., Bell, L., Brune, S., Bunzel, F., Düsterhus, A., Fröhlich, K., Pohlmann, H. and Baehr, J. (2018) Improved teleconnection-based dynamical seasonal predictions of boreal winter. *Geophysical Research Letters*, 45(8), 3605–3614.

Domeisen, D.I.V., Butler, A.H., Fröhlich, K., Bittner, M., Müller, W.A. and Baehr, J. (2015) Seasonal predictability over Europe arising from El Niño and stratospheric variability in the MPI-ESM seasonal prediction system. *Journal of Climate*, 28(1), 256–271.

Domeisen, D.I.V., Butler, A.H., Charlton-Perez, A.J., Ayarzagüena, B., Baldwin, M.P., Dunn-Sigouin, E., Furtado, J.C., Garfinkel, C.I., Hitchcock, P., Karpechko, A.Y., Kim, H., Knight, J., Lang, A.L., Lim, E.-P., Marshall, A., Roff, G., Schwartz, C., Simpson, I.R., Son, S.-W. and Taguchi, M. (2020) The role of the stratosphere in subseasonal to seasonal prediction: 2. predictability arising from stratosphere–troposphere coupling. *Journal of Geophysical Research: Atmospheres*, 125(2). https://doi.org/10.1029/2019JD030923.

Ese, T., Kanestrøm, I. and Pedersen, K. (1988) Climatology of polar lows over the Norwegian and Barents Seas. *Tellus A*, 40(3), 248–255. https://doi.org/10.3402/tellusa.v40i3.11798.

Fletcher, J., Mason, S. and Jakob, C. (2016) The climatology, meteorology, and boundary-layer structure of marine cold air outbreaks in both hemispheres. *Journal of Climate*, 29(6), 1999–2014.

Fröhlich, K., Dobrynin, M., Isensee, K., Gessner, C., Paxian, A., Pohlmann, H., Haak, H., Brune, S., Früh, B. and Baehr, J. (2021) The German climate forecast system: GCFS. *Journal of Advances in Modeling Earth Systems*, 13. https://doi.org/10.1029/2020MS002101.

Furevik, T. (2000) On anomalous sea surface temperatures in the Nordic Seas. *Journal of Climate*, 13(5), 1044–1053.

Giorgetta, M.A., Jungclaus, J., Reick, C.H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K.D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J. and Stevens, B. (2013) Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the coupled model intercomparison project phase 5. *Journal of Advances in Modeling Earth Systems*, 5(3), 572–597.

Jolliffe, I.T. and Stephenson, D.B. (2012) *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Chichester: John Wiley & Sons.

King, A.D., Butler, A.H., Jucker, M., Earl, N.O. and Rudeva, I. (2019) Observed relationships between sudden stratospheric warmings and European climate extremes. *Journal of Geophysical Research: Atmospheres*, 124, 13943–13961.

Kolstad, E.W. (2011) A global climatology of favourable conditions for polar lows. *Quarterly Journal of the Royal Meteorological Society*, 137, 1749–1761.

Kolstad, E.W. (2017) Higher ocean wind speeds during marine cold air outbreaks. *Quarterly Journal of the Royal Meteorological Society*, 143, 2084–2092.

Kolstad, E.W. and Bracegirdle, T.J. (2008) Marine cold-air outbreaks in the future: an assessment of IPCC AR4 model results for the Northern Hemisphere. *Climate Dynamics*, 30(7–8), 871–885.

Kolstad, E.W. and Screen, J.A. (2019) Nonstationary relationship between autumn Arctic sea ice and the winter North Atlantic Oscillation. *Geophysical Research Letters*, 46(13), 7583–7591.

Kolstad, E.W., Bracegirdle, T.J. and Seierstad, I.A. (2009) Marine cold-air outbreaks in the North Atlantic: temporal distribution and associations with large-scale atmospheric circulation. *Climate Dynamics*, 33(2–3), 187–197.

Kolstad, E.W., Breiteig, T. and Scaife, A.A. (2010) The association between stratospheric weak polar vortex events and cold air outbreaks in the Northern Hemisphere. *Quarterly Journal of the Royal Meteorological Society*, 136, 886–893.

Koul, V., Schrum, C., Düsterhus, A. and Baehr, J. (2019) Atlantic inflow to the North Sea modulated by the subpolar gyre in a historical simulation with MPI-ESM. *Journal of Geophysical Research: Oceans*, 124(3), 1807–1826.

Kretschmer, M., 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), 4069–4081.

Kretschmer, M., Cohen, J., Matthias, V., Runge, J. and Coumou, D. (2018a) The different stratospheric influence on cold extremes in Eurasia and North America. *npj Climate and Atmospheric Science*, 1(1), 44.

Kretschmer, M., Coumou, D., Agel, L., Barlow, M., Tziperman, E. and Cohen, J. (2018b) More-persistent weak stratospheric polar vortex states linked to cold extremes. *Bulletin of the American Meteorological Society*, 99(1), 49–60.

Kristiansen, J., Sørland, S.L., Iversen, T., Bjørge, D. and Køltzow, M.Ø.D. (2011) High-resolution ensemble prediction of a polar low development. *Tellus A: Dynamic Meteorology and Oceanography*, 63(3), 585–604.

Landgren, O.A., Seierstad, I.A. and Iversen, T. (2019) Projected future changes in marine cold-air outbreaks associated with polar lows in the northern North Atlantic Ocean. *Climate Dynamics*, 1–13.

Lien, V.S., Schlichtholz, P., Skagseth, Ø. and Vikebø, F.B. (2017) Wind-driven Atlantic water flow as a direct mode for reduced Barents Sea ice cover. *Journal of Climate*, 30(2), 803–812.

Mallet, P.-E., Claud, C., Cassou, C., Noer, G. and Kodera, K. (2013) Polar lows over the Nordic and Labrador Seas: synoptic circulation patterns and associations with North Atlantic–Europe wintertime weather regimes. *Journal of Geophysical Research: Atmospheres*, 118(6), 2455–2472.

Mason, S.J. (2013). Guidance on verification of operational seasonal climate forecasts. Commission for Climatology XIV Technical

Report, WMO No. 1220, World Meteorological Organization, Geneva, Switzerland.

Mason, S.J. and Graham, N.E. (1999) Conditional probabilities, relative operating characteristics, and relative operating levels. *Weather and Forecasting*, 14(5), 713–725.

Michel, C., Terpstra, A. and Spengler, T. (2018) Polar mesoscale cyclone climatology for the Nordic Seas based on ERA-Interim. *Journal of Climate*, 31(6), 2511–2532.

Noer, G., Saetra, Ø., Lien, T. and Gusdal, Y. (2011) A climatological study of polar lows in the Nordic Seas. *Quarterly Journal of the Royal Meteorological Society*, 137, 1762–1772.

Orimolade, A.P., Gudmestad, O.T. and Wold, L.E. (2017) Vessel stability in polar low situations. *Ships and Offshore Structures*, 12(Supplement 1), S82–S87.

Papritz, L. and Grams, C.M. (2018) Linking low-frequency large-scale circulation patterns to cold air outbreak formation in the northeastern North Atlantic. *Geophysical Research Letters*, 45(5), 2542–2553.

Papritz, L. and Pfahl, S. (2015) Importance of latent heating in mesocyclones for the decay of cold air outbreaks: a numerical process study from the Pacific sector of the Southern Ocean. *Monthly Weather Review*, 144(1), 315–336.

Papritz, L. and Spengler, T. (2016) A Lagrangian climatology of wintertime cold air outbreaks in the Irminger and Nordic Seas and their role in shaping air–sea heat fluxes. *Journal of Climate*, 30(8), 2717–2737.

Polkova, I., Köhl, A. and Stammer, D. (2014) Impact of initialization procedures on the predictive skill of a coupled ocean–atmosphere model. *Climate Dynamics*, 42(11–12), 3151–3169.

Rasmussen, E. (1983). A review of meso-scale disturbances in cold air masses. pp.,247–283 in Mesoscale Meteorology – Theories, Observations and Models. Dordrecht: Springer.

Runge, J. (2018) Causal network reconstruction from time series: from theoretical assumptions to practical estimation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(7). https://doi.org/10.1063/1.5025050.

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S. and Sejdinovic, D. (2017) Detecting causal associations in large nonlinear time series datasets. arXiv preprint:1702.07007.

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S. and Sejdinovic, D. (2019) Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, 5(11). https://doi.org/10.1126/sciadv.aau4996.

Saggioro, E. and Shepherd, T.G. (2019) Quantifying the timescale and strength of Southern Hemisphere intraseasonal stratosphere–troposphere coupling. *Geophysical Research Letters*, 46(22), 13479–13487.

Saggiorato, B., Louise, N., Siebesma, A.P., de Roode, S., Sandu, I. and Papritz, L. (2020) The influence of convective momentum transport and vertical wind shear on the evolution of a cold air outbreak. *Journal of Advances in Modeling Earth Systems*, 12(6). https://doi.org/10.1029/2019MS001991.

Sanna, A., Borrelli, A., Athanasiadis, P., Materia, S., Storto, A., Navarra, A., Tibaldi, S. and Gualdi, S. (2016). CMCC-SPS3: the CMCC Seasonal Prediction System 3, Research Paper RP0285, Lecce, Italy.

Siew, P.Y.F., Li, C., Sobolowski, S.P. and King, M.P. (2020) Intermittency of Arctic–midlatitude teleconnections: stratospheric pathway between autumn sea ice and the winter North Atlantic Oscillation. *Weather and Climate Dynamics*, 1(1), 261–275.

Spengler, T., Claud, C. and Heinemann, G. (2017) Polar low workshop summary. *Bulletin of the American Meteorological Society*, 98(6), ES139–ES142.

Terpstra, A., Michel, C. and Spengler, T. (2016) Forward and reverse shear environments during polar low genesis over the Northeast Atlantic. *Monthly Weather Review*, 144(4), 1341–1354.

Vijverberg, S., Schmeits, M., van der Wiel, K. and Coumou, D. (2020) Subseasonal statistical forecasts of eastern US hot temperature events. *Monthly Weather Review*, 148(12), 4799–4822.

Vitart, F. and Balmaseda, M. (2018). Impact of sea surface temperature biases on extended-range forecasts, Technical Memorandum 830, Reading, UK.

Yang, K. and Wang, C. (2019) Seasonal persistence of soil moisture anomalies related to freeze–thaw over the Tibetan Plateau and prediction signal of summer precipitation in eastern China. *Climate Dynamics*, 53(3-4), 2411–2424.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Polkova* I, Afargan-Gerstman H, Domeisen DIV, *et al*. Predictors and prediction skill for marine cold-air outbreaks over the Barents Sea. *QJR Meteorol. Soc.* 2021;147:2638–2656. https://doi.org/10.1002/qj.4038