

<https://helda.helsinki.fi>

Stratosphere-troposphere coupling enhances subseasonal predictability of Northern Eurasian cold spells

Statnaia, Irina

2022-07

Statnaia , I , Karpechko , A , Kämäräinen , M & Järvinen , H 2022 , ' Stratosphere-troposphere coupling enhances subseasonal predictability of Northern Eurasian cold spells ' , Quarterly Journal of the Royal Meteorological Society , vol. 148 , no. 747 , pp. 2769-2783 . <https://doi.org/10.1002/qj.4335>

<http://hdl.handle.net/10138/347753>

<https://doi.org/10.1002/qj.4335>

cc_by_nc_nd

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.


This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

RESEARCH ARTICLE

Stratosphere–troposphere coupling enhances subseasonal predictability of Northern Eurasian cold spells

Irina Statnaia^{1,2}  | Alexey Karpechko¹  | Matti Kämäräinen¹  | Heikki Järvinen² ¹Finnish Meteorological Institute, Helsinki, Finland²Institute for Atmospheric and Earth System Research, Faculty of Science, University of Helsinki, Helsinki, Finland**Correspondence**Irina Statnaia, Finnish Meteorological Institute, 00560 Helsinki, Finland.
Email: irina.statnaia@helsinki.fi**Funding information**

Academy of Finland, Grant/Award Number: 317999; Magnus Ehrnroothin Säätiö, Grant/Award Number: 2021

Abstract

Here we explore the stratospheric influence on the predictability of Eurasian cold-spell events using the European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble hindcasts obtained from the Subseasonal-to-Seasonal (S2S) archive. To isolate the stratospheric influence, we subsampled two groups of hindcasts according to the strength of the stratospheric polar vortex preceding the cold spells at the surface. The predicted probability of cold spells, defined as the lowest 10th percentile of weekly mean temperature anomalies over northern Eurasia (10° W–130° E and 50° N–65° N), is systematically higher, by 0.05–0.2, at lags 7–24 days in the forecasts initialized during the weak stratospheric vortex compared to the strong stratospheric vortex group, extending the predictability of cold spells by 3–5 days. Our results suggest that, in the case of the weak polar vortex, stratosphere–troposphere coupling favors the negative Northern Annular Mode (NAM) regime and the cold-air outbreaks in Eurasia. As a consequence, the long stratospheric predictability extends the predictability of the cold spells. On the other hand, when the polar vortex is strong, the stratospheric anomalies do not favor the observed negative NAM regime, which thus results from the internal tropospheric processes only. In this case the predictability of cold-air spells is limited. Furthermore, we show that the extended predictability of cold spells arising from the stratosphere–troposphere coupling is captured by a simple statistical model, suggesting that governing large-scale dynamics behave effectively linearly over some limited periods. Quantified contribution of the stratosphere–troposphere coupling to the enhanced skill of the extended-range cold-spell forecasts documented in our paper may prove useful in the development of forecasting tools.

KEYWORDS

cold spell, NAM, polar vortex, predictability, stratosphere–troposphere coupling, subseasonal predictions

1 | INTRODUCTION

Weather-dependent planning and decision-making can benefit greatly from the Subseasonal-to-Seasonal (S2S) weather predictions made for up to 6 weeks ahead (White *et al.*, 2017; 2021). At this timescale early predictions of severe weather conditions, such as periods of below-normal surface temperatures in winter, become important because they allow more time to prepare. However, numerical weather forecast models still have a rather low skill level on the S2S timescale. Weather predictability in the Northern Hemisphere (NH) extratropics, especially predictability of extreme events, is usually limited to about a week and decreases rapidly with lead time (Vitart *et al.*, 2019). Therefore, it is important to develop a better understanding of the factors that influence extended-range weather forecasts and potential windows of opportunity for the enhanced predictability.

Numerous publications have suggested that remote drivers, such as the state of the stratosphere, can act as a source of enhanced probabilistic predictability of surface winter weather regimes at the S2S timescale (e.g. Baldwin and Dunkerton, 2001; Scaife *et al.*, 2016; Büeler *et al.*, 2020; Domeisen *et al.*, 2020). Stratospheric flow is dominated by planetary waves which are associated with longer timescales, and thus longer predictability, than the synoptic-scale waves in the troposphere (Baldwin *et al.*, 2003).

In winter, the state of the stratospheric circulation can be characterized by the strength of the stratospheric polar vortex. It has been shown that stratospheric extreme events may modulate the large-scale circulation patterns in the troposphere that typically last more than a week (Baldwin and Dunkerton, 2001; Sigmond *et al.*, 2013). Weak stratospheric polar vortex (hereafter, WPV) events, in particular Sudden Stratospheric Warmings (SSW), favor a negative North Atlantic Oscillation (NAO)/Northern Annular Mode (NAM) phase and cold spells over Eurasia. In previous studies addressing the stratosphere–troposphere dynamical coupling it was shown that the below-normal temperatures over northern Eurasia are often preceded by WPVs and SSWs (Thompson *et al.*, 2002; Kidston *et al.*, 2015; Kretschmer *et al.*, 2018; King *et al.*, 2019). Kautz *et al.* (2020) showed an enhanced predictability of the Eurasian cold spell in February–March 2018 at lead times up to 25 days in association with an SSW that occurred in early February. They pointed out that not only did the occurrence of the SSW itself play a crucial role in modulating the tropospheric flow patterns, but also the subsequent evolution of the lower stratosphere and the stratosphere–troposphere coupling. However, it remains unclear whether other Eurasian cold spells that occurred after WPV events are

also better predicted in comparison to cold spells that were not preceded by WPV events.

In this study we aim to quantify the predictability of Eurasian cold spells conditioned on the occurrence of WPV events using the composite analysis of extended-range winter forecasts. While predictability of Eurasian cold spells on S2S timescales and its relation to Madden–Julian Oscillation has been previously investigated (e.g. Ferranti *et al.*, 2018), to the authors' knowledge, ours is the first study that addresses the dependence of the cold-spell forecast skill on stratospheric conditions using a composite analysis. Our goal is to test the hypothesis that WPV events are associated with enhanced weather predictability assuming that the forecast model is capable of properly representing the coupling between the stratosphere and the troposphere.

This paper is organized as follows. The data and methods are described in Section 2. In Section 3 we provide an overview of the composite synoptic situation using the reanalysis data. In Section 4 we assess the predictability of the cold spells in the dynamical model hindcasts and in Section 5 we compare the dynamical model forecasts with those by a simple statistical model. In the final section we conclude the paper with results and discussion.

2 | DATA AND METHODS

We use retrospective forecasts (hindcasts) from the S2S forecast system of European Centre for Medium-Range Weather Forecasts (ECMWF) (Vitart, 2014). The hindcasts are produced to calibrate systematic model biases in the real-time forecasts and are launched on-the-fly at the same calendar date as the real-time-forecasts twice a week (Mondays and Thursdays) for the previous 20 years. Here we use model versions CY41R1, CY41R2, CY43R1, CY43R3, CY45R1 and CY46R1. The resolution of the model is Tco639L91 (about 16 km) for forecast days 0–15 and Tco319L91 (about 32 km) for the rest of the forecast period for all the model versions used in the study but the first one. The resolution in the CY41R1 model version is TL639L91 (about 32 km) up to day 10 and TL319L91 (about 64 km) after day 10. Also, the CY41R1 and CY41R2 model versions have lower ocean resolution (1°) and do not have active sea ice. In the latter versions the ocean resolution is 0.25° and they include active sea ice. Nevertheless, all model versions have 91 vertical levels with the top level at 0.01 hPa, thus fully resolving the stratosphere and representing the stratosphere–troposphere coupling (Kawatani *et al.*, 2019). We use the hindcasts that cover extended winters from November to March (NDJFM) for 25 years (1995–2020). The corresponding operational forecasts cover five winters from 2015/16 to 2019/20. We

analyze hindcasts for 216 operational forecasts, or 4,320 individual hindcasts each consisting of 11 ensemble members that run for 46 days. We pool all hindcasts into a single dataset and assume that model upgrades during this period did not affect our results. The use of different model versions is a potential drawback; however, this allows us to increase the sampling size. Large sampling size is important for the study because we focus on predictability of individual events where case-to-case variability likely overwhelms effects of model improvements. Geopotential height (Z) fields are sampled daily at 0000 UTC, while the 2-m temperature ($T2m$) is daily averaged. Both Z and $T2m$ fields are analyzed at $1^\circ \times 1^\circ$ horizontal resolution. Anomalies in the hindcasts are calculated with respect to the other 19 hindcast years for each re-forecast date and lead time.

We use the ECMWF ERA5 reanalysis (Hersbach *et al.*, 2020) throughout for historical analysis and verification. Anomalies in ERA5 are calculated with respect to ERA5 climatology from 1995 to 2020 to be consistent with the hindcasts.

We identify northern Eurasian cold spells using the ERA5 daily $T2m$ anomaly averaged over northern Eurasia (10°W – 130°E and 50°N – 65°N) as in Karpechko *et al.* (2018) and Kautz *et al.* (2020). Before calculating the anomalies, we detrend temperature datasets linearly over the considered time period. Trends were calculated for each grid point and for each calendar day. Note that detrending did not affect the results. $T2m$ anomalies are averaged weekly and the first week within each season starts on 1 November (e.g. 1–7.11, 8–14.11, etc.). The weekly mean negative $T2m$ anomaly below the 10th percentile of the weekly mean $T2m$ anomaly distribution is considered an extreme event. The threshold of the 10th percentile is chosen to have enough extreme cases for the analysis. The onset date of the extreme event is the first day of a week for the weekly averaged data, irrespective of weekday. Note that our approach does not exclude the possibility that two or more adjacent weeks falling within the 10th percentile are selected for the analysis. In this case the predictability of extreme events might be influenced by persistence. However, we checked that removing the adjacent weeks does not affect our main results (see Section 4).

To investigate the predictability of the selected cold spells in hindcasts and quantify the influence of the stratospheric state on the predictability, we perform a composite analysis by grouping the events according to the state of the polar vortex. We determine the stratospheric polar vortex state using a daily mean zonal mean zonal wind anomaly at 10 hPa 60°N (U10) 2 weeks before the onset of the cold spell. The lag approximately corresponds to the time of the downward anomaly propagation (Baldwin and

Dunkerton, 2001). However, due to the long duration of the stratospheric anomalies, the selection of the events is not expected to be too sensitive to the chosen lag. In the first group the cold-spell events were preceded by a weak polar vortex (WPV) defined using U10 anomalies. We expect that the stratosphere acts as a source of predictability leading to enhanced forecast skill for these events. Note that the WPV group is not always associated with SSWs. The second group is formed by cases preceded by a strong polar vortex (SPV). A strong stratospheric vortex is expected to favor a positive NAO/NAM phase and a warm anomaly over Eurasia (e.g. Tripathi *et al.*, 2015); therefore, the cold spells preceded by the strong stratospheric vortices are expected to result from internal tropospheric dynamics only, without a downward-propagating signal from the stratosphere.

To explore the downward impact of the stratosphere we use the NAM index, defined as an area-weighted polar cap averaged (60° – 90°N) daily geopotential height anomaly normalized by standard deviation and taken with the opposite sign (e.g. Karpechko *et al.*, 2017). The opposite sign is used because NAM phase is conventionally defined as a negative one when polar cap geopotential height anomalies are positive.

Several previous studies used simple linear statistical models to show that enhanced forecast skill can be associated with anomalous states of the lower stratospheric circulation (Baldwin *et al.*, 2003; Christiansen, 2005; Charlton *et al.*, 2003; Karpechko, 2015; Beerli *et al.*, 2017). To test whether the enhanced predictability associated with stratosphere–troposphere coupling can be captured by a linear model, we compare the results we obtained with the ECMWF model to a simple statistical model, implemented at the Finnish Meteorological Institute. The model we use is based on the seasonal model described in Kämäräinen *et al.* (2019). The model is a 50-member ensemble based on random sampling of both the fitting samples (time steps) and the potential predictors. In the fitting process one Least Absolute Shrinkage and Selection Operator regression (LASSO) model was fit to each random sample, thus creating 50 separate members. Here, we apply this model ensemble to predict weekly anomalies at subseasonal scale. The major differences from the previously published model are the use of more Principal Components (PC) per predictor variable and a different set of predictor variables. In Kämäräinen *et al.* (2019), five leading PCs of sea surface temperature anomalies and three PCs of 150-hPa geopotential height anomalies from 20°S – 90°N were used as predictors of the seasonal scale. In the present study the model is run twice with different sets of predictors. In Run 1, for the weekly scale, we use 20 PCs of mean sea level pressure (SLP) anomalies, 20 PCs of 2-m temperature anomalies and 20 PCs of 150-hPa

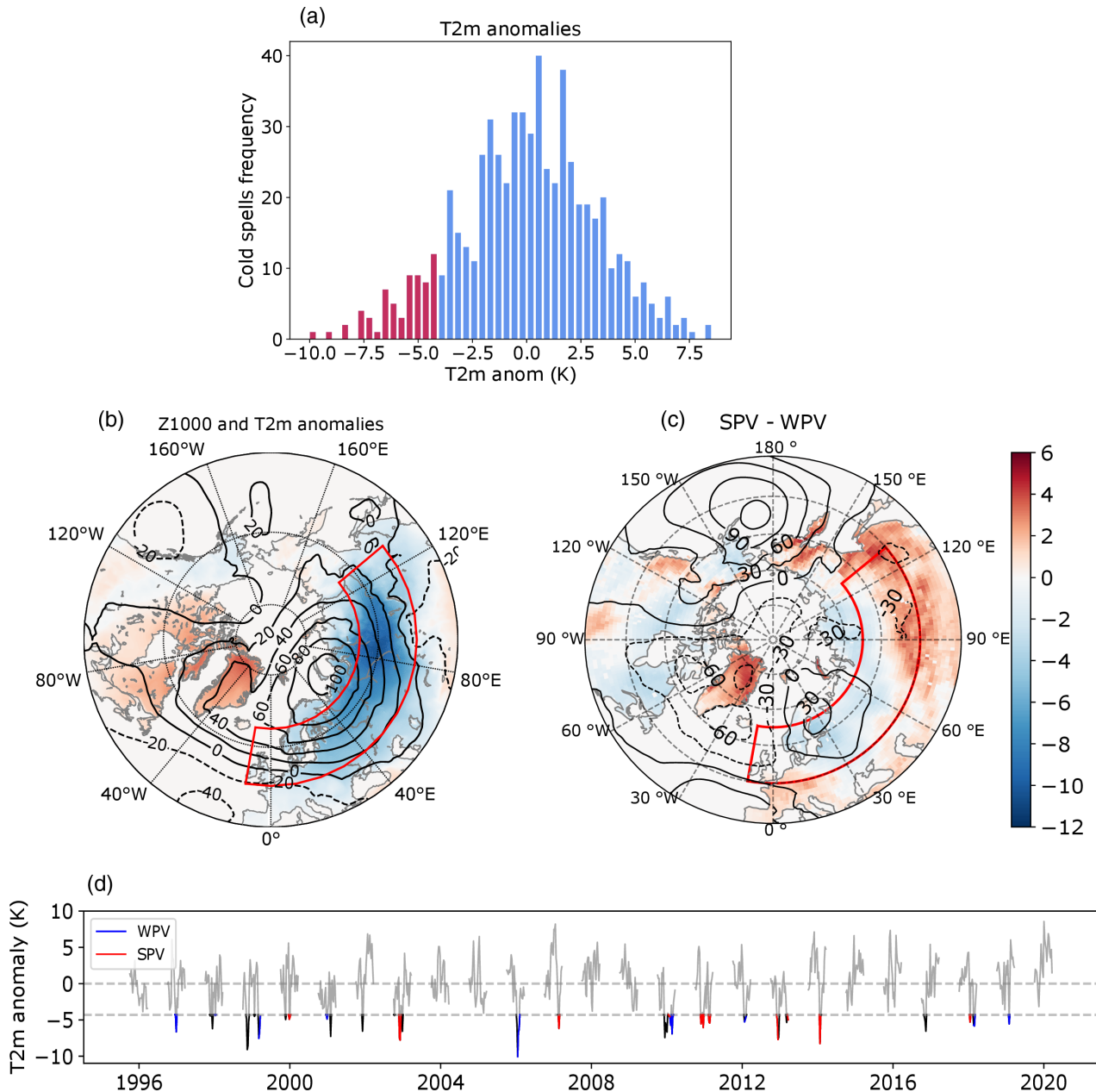


FIGURE 1 (a) Histogram of weekly averaged surface temperature (T2m) anomalies (in K) over Eurasia in ERA5; red-colored bars denote temperatures below the 10th percentile (-4.3 K). (b) Geopotential height anomalies at 1000 hPa (Z1000, in m, contours) and composite of observed 2-m temperature anomalies (T2m, in K, shading) weekly averaged for all detected cold spell events; the red box shows the area in Eurasia used to average the anomalies in panel (a); (c) difference in observed Z1000 and T2m anomalies between the selected strong polar vortex (SPV) and weak polar vortex (WPV) events. (d) Time series of weekly averaged temperature anomalies (in K) over Eurasia during extended winters from 1995 to 2020. All cold spells events below the 10th percentile are shown in black, the WPV group events are shown in blue and the SPV group events in red [Colour figure can be viewed at wileyonlinelibrary.com]

geopotential height anomalies from 20° S to 90° N. To investigate the predictability gain due to the stratospheric predictors, we repeat the experiment without the 150-hPa geopotential height anomalies predictors in Run 2.

A homogenized combination of ERA-20C (Poli *et al.*, 2016) and the ERA5 reanalyses are used in fitting to get enough data for the statistical model to learn from. For homogenization we used quantile mapping adjustment

(e.g. Räisänen and Rätty, 2013) to force the distributions of variables of the ERA-20C reanalysis to match the distributions of ERA5. The overlapping years 1951–2010 were used in the calibration of the correction functions: each variable, each month of year, and each grid cell was adjusted separately. We then created a homogenized dataset for years 1900–2019 by using years 1900–1950 from the adjusted ERA-20C, years 1951–2010 from the mean of

the ERA5 and the adjusted ERA-20C, and years 2011–2019 from ERA5. By doing so we increased the number of fitting years and reduced the biases of the ERA-20C. The period of 1900–1994 is used for model fitting and the prediction period is set to 1995–2020. The model produces forecasts of the mean anomaly for the whole area of study instead of predicting the anomalies at individual grid cells inside it.

While the quality of the Z150 field in ERA-20C in the earlier period is difficult to assess due to the scarcity of upper-level observations, we found in our previous work (Kämäräinen *et al.*, 2019), as well as here, that using the Z150 PCs as predictors improves the statistical model considerably, which increases our confidence in the quality of the Z150 fields.

3 | SYNOPTIC ANALYSIS OF COLD SPELLS

We start by detecting the cold-surface temperature (T2m) events over Eurasia in winter in the reanalysis (Figure 1). Figure 1a shows the distribution of land surface temperature averaged over 10° W–130° E and 50° N–65° N and highlights the extreme events below the 10th percentile.

Overall, 62 cold spells are detected with an average weekly anomaly falling below -4.3 K (Figure 1d). From 25 extended winters analyzed in this study, cold spells occurred only in 17 winters. The composites of T2m and Z1000 of all detected events is shown in Figure 1b. As can be seen, the Eurasian cold spells are associated with a high over the northern Ural, which is similar to the previously studied individual cold events (e.g. Karpechko *et al.*, 2018; Kautz *et al.*, 2020).

To form composites with well-separated stratospheric states, we select 15 events, which represent $\sim 25\%$ of all events (Figure 1d), preceded by the weakest stratospheric zonal winds 2 weeks before the cold-spell onset date, and 15 events preceded by the strongest stratospheric wind anomalies across all events (see Section 2: *Data and Methods*). In this study we mostly analyze these 30 cases, which, as we find, is sufficient to establish statistical significance of the differences in dynamics and predictability between the groups. Note that the T2m and Z1000 composite maps of these 30 events are very similar to those shown in Figure 1b (not shown). Table 1 lists the onset dates of these cases together with mean T2m and stratospheric wind (U10) anomalies. The T2m difference between these two groups (Figure 1c and Table 1) shows that, on average,

TABLE 1 Onset dates of the selected cold spells for the weak polar vortex (WPV, left) and strong polar vortex (SPV, right) groups during 1995–2020. The T2m columns show weekly mean surface temperature over the cold spell area in Eurasia. The U10 column shows zonal mean zonal wind anomalies at 10 hPa 60° N 2 weeks before the onset dates

No.	WPV cold events			SPV cold events		
	Onset date	T2m	U10	Onset date	T2m	U10
1	December 20, 1996	-4.29	-26.13	December 27, 1999	-4.91	9.8
2	December 27, 1996	-6.63	-18.54	January 3, 2000	-4.6	12.45
3	January 10, 1998	-4.37	-19.57	November 29, 2002	-7.57	10.16
4	March 14, 1999	-6.66	-32.2	December 6, 2002	-7.77	7.45
5	March 21, 1999	-4.48	-33.3	February 21, 2007	-6.16	8.28
6	December 20, 2000	-4.29	-17.18	January 24, 2010	-4.46	14.78
7	December 27, 2000	-4.87	-21.65	November 29, 2010	-5.55	8.39
8	January 24, 2006	-6.69	-22.08	December 6, 2010	-4.34	8.67
9	January 31, 2006	-6.28	-28.93	December 20, 2010	-6.05	8.58
10	February 7, 2010	-5.7	-16.7	February 14, 2011	-5.31	8.6
11	February 21, 2010	-6.9	-23.45	February 21, 2011	-5.43	23.9
12	February 28, 2010	-4.56	-19.79	December 6, 2012	-5.03	11.78
13	January 31, 2012	-4.92	-17.18	March 21, 2013	-4.97	10.32
14	February 28, 2018	-5.83	-39.17	January 24, 2014	-8.26	9.96
15	January 31, 2019	-5.55	-30.17	January 17, 2018	-5.29	16.49
Mean	-	-5.47	-24.4	-	-5.71	11.31
SD		0.95	6.69		1.19	4.16

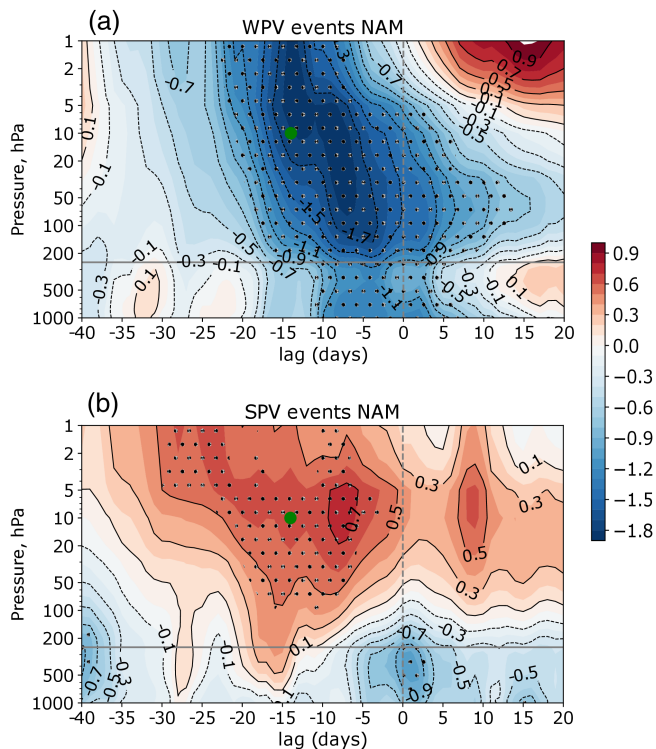


FIGURE 2 Lagged cross-sections of the Northern Annular Mode (NAM) index composite in ERA5 (in standard deviation units) for cold spells preceded by the weak polar vortex (WPV) (a) and by the strong polar vortex (SPV) (b) events. Gray horizontal lines indicate the approximate boundary between the stratosphere and troposphere. Gray dotted vertical lines highlight the onset of events (lag 0). Green dots at a lag of 14 days and the 10-hPa level indicate the point used for the grouping by the polar vortex state. Hatching denotes areas where the NAM index is significantly different from zero ($p < 0.01$) [Colour figure can be viewed at wileyonlinelibrary.com]

the cold-air outbreaks are slightly weaker under the WPV conditions; however, the difference in the strength of the anomaly between the groups is small and not statistically significant.

Extreme states of the polar stratospheric circulation can influence the weather regimes in the extratropical troposphere and shift the probability density function of the NAO/NAM (Baldwin and Dunkerton, 2001; Baldwin *et al.*, 2003). The tropospheric NAO/NAM anomalies can last for up to 2 months, as the troposphere is dynamically coupled with the more slowly evolving stratosphere. To investigate the coupling between the stratospheric polar vortex and the tropospheric anomalies during the cold spells we use lagged vertical cross-sections of the NAM index (Figure 2). As the composites were chosen based on the polar vortex strength 2 weeks before the cold-air outbreaks (onset dates of the cold events correspond to lag 0), the anomalies at the 10 hPa height are the largest at

negative lags from -15 to -7 days for WPV (Figure 2a) and from -8 to -5 days for SPV (Figure 2b). In the WPV composite the tropospheric NAM regime becomes stronger about eight days before the cold-air outbreaks and remains strongly negative for about 2 weeks during the events. The composite cross-section for SPV events NAM (Figure 2b) is different from the WPV group: the positive NAM anomalies are confined to the stratospheric heights, while strongly negative NAM anomalies appear in the troposphere between days -5 and $+10$. In this case the tropospheric NAM regime is decoupled from that in the stratosphere and the extended-range predictability of the cold spells is not expected to be enhanced by the stratospheric signal.

To provide a better overview of the large-scale flow in the lower stratosphere–upper troposphere region and its coupling with the near-surface circulation during the cold-air outbreaks, Figure 3 shows the weekly mean potential vorticity anomalies (PVA) on the 320-K isentropic surface together with Z1000 anomalies for the cold-spell events. The PVA distribution is broadly similar between all cold events (Figure 3a), WPV (Figure 3b) and SPV (Figure 3c) groups because the NAM signal at this level is negative in all cases (Figure 2). The PVA fields in general correspond to the surface geopotential height anomalies, showing positive PVA values in the North Atlantic region for both groups and negative values in the North Pacific region in the SPV group. Over the North Atlantic region, WPV Z1000 composite anomalies resemble a negative NAO phase. In northern latitudes WPV Z1000 exhibits a planetary wave 1 pattern with a low over the North Pacific, which can enhance tropospheric forcing of the stratospheric circulation (e.g. Garfinkel *et al.*, 2010), consistent with a weakened stratospheric vortex. In the SPV composite (Figure 3c) the jet stream, marked by the large PV gradients at the edge of the polar vortex, is more undulating. The NAO/NAM– circulation is weaker but, similarly to the WPV composite, the polar cap is dominated by positive near-surface geopotential height anomalies. Both composites have a near-surface anticyclone in the European Arctic sector which favor cold-air advection into the studied area, consistent with Figure 1. In the WPV composite there are two geopotential height maxima: one over the North Atlantic region resembling a Greenland blocking and an Ural high-like pattern around 80° E. In the SPV composite there is only one maximum near the study area centered near 40° E resembling a Scandinavian blocking. The differences are larger in the Pacific sector where SPV has positive near-surface geopotential height anomaly indicating a weakened Aleutian low, consistent with a reduced wave forcing and a stronger stratospheric vortex (e.g. Orsolini *et al.*, 2009).

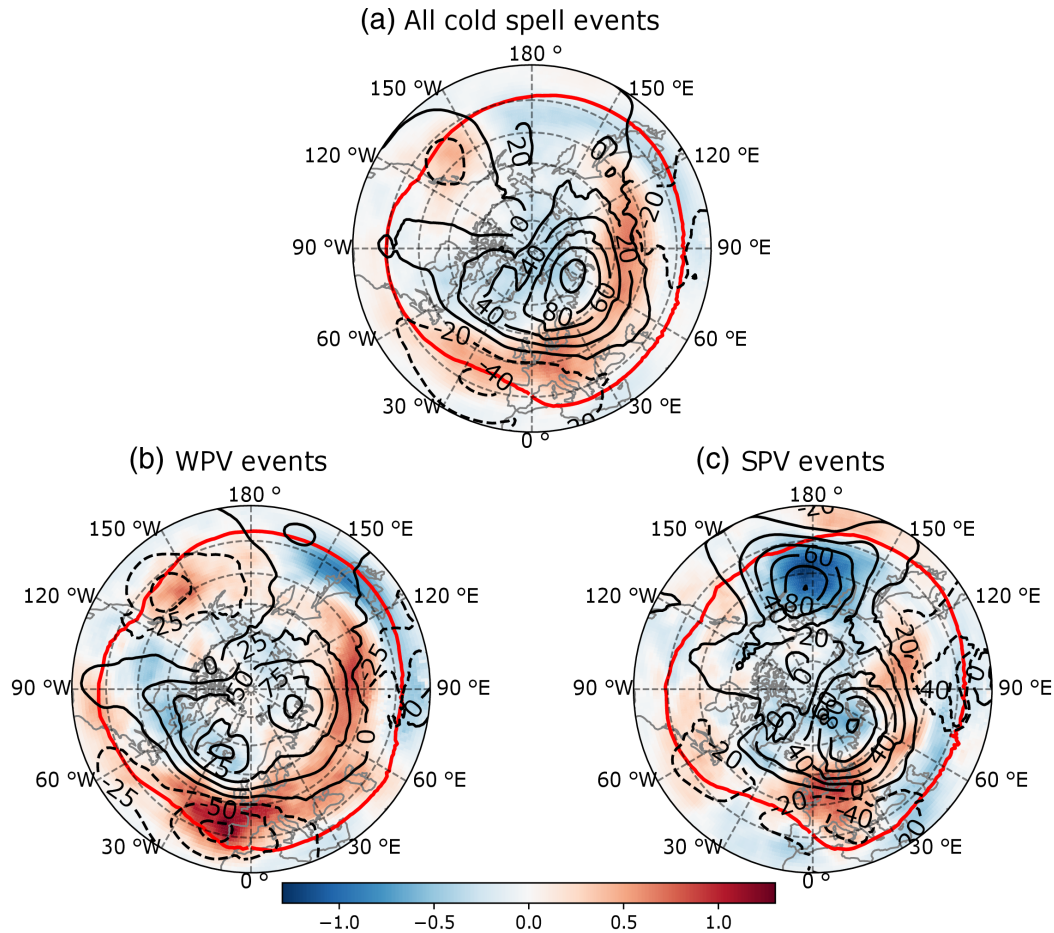


FIGURE 3 Composites of weekly averaged potential vorticity anomalies (PVA) on the 320-K isentropic surface (in potential vorticity unit [PVU], shading) and 1,000-hPa weekly averaged geopotential height anomalies (in m, contours) from ERA5: (a) all detected cold events, (b) 15 weak polar vortex (WPV) events and (c) 15 strong polar vortex (SPV) events. Red contour indicates the PVU = 2 isoline [Colour figure can be viewed at wileyonlinelibrary.com]

4 | PREDICTABILITY OF THE EURASIAN COLD SPELLS

We examine the predictability of extended winter Eurasian cold-spell events in the ECMWF hindcasts during the period of 1995–2020. Figure 4 shows boxplots of the area mean T2m forecasts for two composites of cold-air outbreaks, listed in Table 1, that took place after the weakening (Figure 4a) and the strengthening of the stratospheric polar vortex (Figure 4b) as a function of the initialization lag. For each initialization lag the number of 11-member hindcast ensembles varies between 11 and 29, giving between 121 and 319 individual forecasts for each initialization lag. Since weekly mean anomaly forecasts are considered, lag 0 corresponds to forecast week 1 (forecast days 0–6), lag 7 corresponds to forecast week 2 (forecast days 7–13) and so on.

For the WPV events (Figure 4a) the ensemble mean anomaly starts to consistently decrease after day 20 (week 4). This decrease coincides in time with

strengthening of the stratospheric anomalies (Figure 2) and likely reflects the transition to nearly deterministic forecasts of the WPV conditions that occur approximately six days later (e.g. Karpechko, 2018). The interquartile range of ensemble members is fully in the cold anomaly on day 14 when the WPV conditions are assimilated into the model initialization. However, there is still a considerable spread among the hindcasts until around day 9, when the spread starts to decrease. This is consistent with the previous studies of cold spells (Kautz et al., 2020). For SPV events (Figure 4b) there is very little predictability skill until up to about 12 days in advance, after which the ensemble mean anomaly starts to decrease rapidly. The spread starts to decrease after around day 6 remaining larger than that in WPV composite until day 0. This difference in the time evolution of the forecast skill between WPV and SPV cases at lags beyond 2 weeks suggests the influence of the state of the stratosphere on the predictability of the Eurasian cold spells.

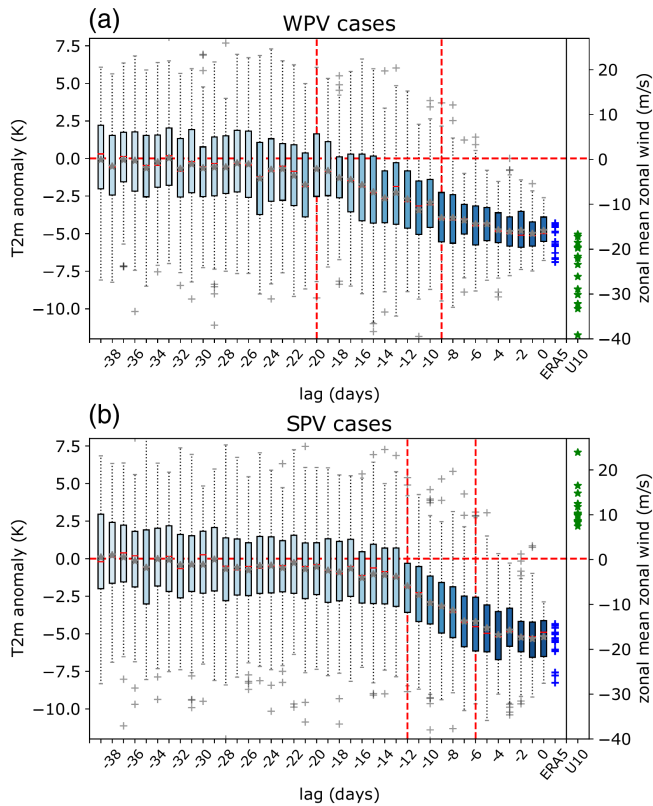


FIGURE 4 Boxplots of the forecast Eurasian 2-m temperature (T2m) anomaly (in K) for different initialization lags of the European Centre for Medium-Range Weather Forecasts (ECMWF) hindcasts for the (a) weak polar vortex (WPV) cases and (b) strong polar vortex (SPV) cases. The boxes show the interquartile range (Q1–Q3), whiskers represent the ranges for the bottom 25% and the top 25% of the data values, excluding outliers, which are shown by gray crosses. The median is the red line inside each box and the mean of each box is shown by a gray star. The intensity of the blue color of boxes is proportional to the mean temperature anomaly. The horizontal dashed red line shows 0 K, the vertical dashed red lines show 9 and 20 days before cold spells for the WPV group and 6 and 12 days for the SPV group to guide the discussion in the text. The blue crosses show the ERA5 T2m anomaly (K) for the cold-spell cases (15 cases in each group), while the green stars show the preceding U10 anomalies (m/s) [Colour figure can be viewed at wileyonlinelibrary.com]

To further illustrate the difference in the signals between WPV and SPV cases, we examine the forecast probability of the cold spells in the hindcasts as a function of initialization lags (Figure 5a). The probability of an extreme event is defined as a fraction of the T2m hindcasts that predicted weekly mean anomalies below the 10th percentile. In most cases, except at the shortest lead times, the 10th percentile threshold is estimated from the hindcast climatology for each 11-member 20-year hindcast set separately (Kautz et al., 2020). At the shortest lead times (0–3 days), the spread of the ensemble forecasts is

negligible and the hindcasts do not allow robust estimation of the 10th percentile threshold. Therefore, at lead times of 0–3 days we define the threshold using ERA5 value of -4.3 K. We average data over 3 days to smooth out fluctuations for display and conciseness reasons; however, a similar result is obtained with daily resolution. Note, that an extreme event probability of 0.1 is expected in the forecasts by definition even if there is no signal. To test whether the forecast probability is significantly different from 0.1 for each lead time we used bootstrapping with replacement and formed the distributions of probabilities by repeating the resampling from all ensemble members of all hindcasts initialized at that lead time 10,000 times. The null hypothesis, that the forecast probability equals 0.1, is rejected if the 2.5th percentile of the distribution exceeds 0.1, corresponding to the 95% confidence interval.

It is clear from Figure 5 that the predicted probability of the extreme event is systematically higher, by 0.05–0.15, at lags 7–24 days in the WPV case. In this composite, the probability starts to be significantly higher than the 0.1 threshold already on days 22–24 before the extreme events. The probability starts to increase rapidly at lags of about 2 weeks when WPV conditions that can provide predictability (Figure 2) are assimilated to the forecast system. On the other hand, the probability in the SPV case starts to be significantly higher than 0.1 later, on day 19–21, and increases slowly until around 10 days in advance. Thus, WPV conditions provide additional skill to extended-range probabilistic forecasts of the Eurasian cold spells at lead times of around days 7–24 (forecast weeks 2–4). For example, the predicted probability of an extreme event starts to be three times higher (0.3) than the climatological probability of 0.1 approximately 3–5 days earlier in the WPV group, as is shown by the gray vertical dotted lines in Figure 5. However, at the lead time less than day 7 (forecast weeks 1–2) the effect of stratospheric conditions is not detectable.

To make sure that persistence did not affect our results we remove the adjacent weeks, keeping only the first week in cases where two consecutive weeks are defined as cold events (Table 1), and repeat our analysis (Figure 5b). In this case the WPV group consists of 10 cold-spell cases and the SPV group – of 11 cases. As can be seen, our key result, that the predictability in the WPV cases is better, is confirmed. Moreover, the difference between the SPV and WPV groups' predictability is even more pronounced now. The predictability in the SPV cases is slightly worse without the effect of persistence: the probability of an extreme event starts to be significantly higher than 0.1 only 16–18 days in advance. On the other hand, removing the adjacent weeks did not worsen the forecast skill in the WPV cases and even improved it at some initialization lags.

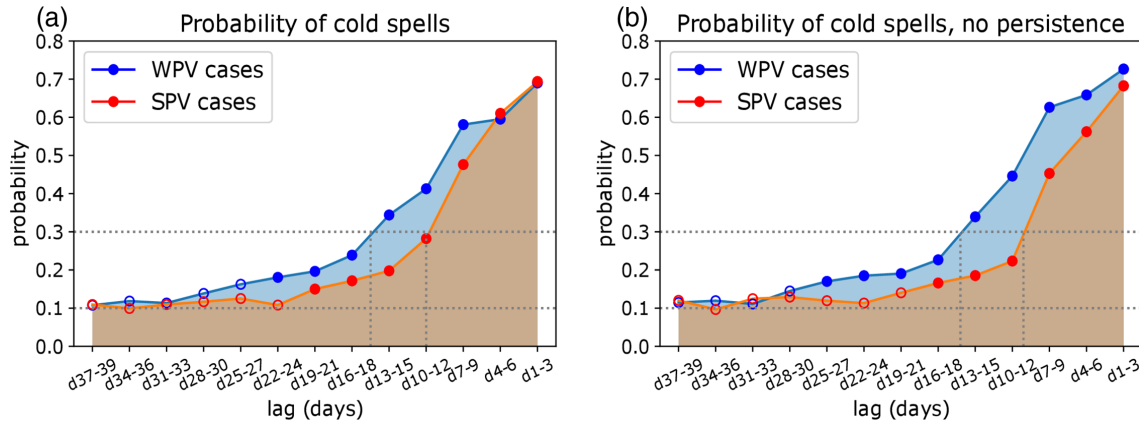


FIGURE 5 (a) Predicted probability of cold spell events as a function of lead time of the European Centre for Medium-Range Weather Forecasts (ECMWF) hindcasts averaged by three days. Weak polar vortex (WPV) events’ probability is shown in blue color and that of the strong polar vortex (SPV) events, in red color. Closed circles denote cases where the lower uncertainty range, defined using bootstrapping with the number of iterations $N = 10,000$, exceeds 0.1. Open circles denote cases with the lower uncertainty range below 0.1. Probabilities of 0.3 and 0.1 are shown by dotted lines. (b) Same as in (a), but for the composites excluding the adjacent weeks [Colour figure can be viewed at wileyonlinelibrary.com]

The comparison of the two groups of cold-spell events with the opposite stratospheric polar vortex conditions allows us to quantify the gain in predictability due to the stratospheric influence, as its contribution to predictability is expected to be suppressed in the SPV group. It is of interest to analyze how the predictability depends on the strength of the stratospheric polar vortex. For this analysis we use all 62 cold-spell events detected during 1995–2020 (Figure 1d) and group them into six groups depending on the corresponding U10 value. We focus on the lead times of 10–24 days because this is where the difference between WPV and SPV is maximal, and group all our cold-spell cases into six groups by 10 cases in each, except for the first and the last group (11 cases). Figure 6 shows that the predictability decreases with the strengthening of the polar vortex; however, only the WPV cases differ significantly from the other events, while the difference between the neutral polar vortex states and the SPV states is not significant. Note that cold spells were not detected after very strong stratospheric polar vortex conditions (U10 anomaly more than ~ 25 m/s, see also Table 1). This is consistent with expectations that SPV favors warm temperature anomalies in the northern Eurasia, rather than cold ones (e.g. Tripathi *et al.*, 2015).

To get insights about the reasons behind the increased predictability in WPV composite we analyze the forecast NAM time–height cross-sections (Figure 7), as well as the forecast Z1000 anomalies and the root mean square errors (RMSE) of the forecasted Z1000 anomalies (Figure 8). We select forecasts initialized from 22 to 24 days before the onset of the cold spells, because at these lags the forecast probability of extreme events becomes significantly different from the climatology in the WPV cases but not in

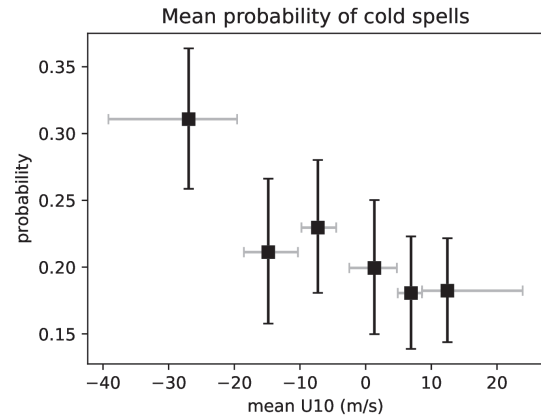


FIGURE 6 Mean predicted probability of cold spell events at lead times of 10–24 days as a function of mean zonal mean zonal wind anomaly at 10 hPa 60° N (U10) two weeks before the onset of the cold spell. Each of the six groups of events consists of 10 members, except for the first and the last group, which consist of 11 members. The confidence intervals shown by black whiskers are defined using bootstrapping with the number of iterations $N = 10,000$. The gray whiskers show the U10 range for each group

SPV cases (Figure 5a). Altogether 57 individual ensemble hindcasts are included into the WPV composite, and 71 ensemble hindcasts are included into the SPV composite. In the WPV group the observed intensification of the stratospheric anomaly and downward propagation of the NAM– conditions are well forecast, although the strongest tropospheric NAM– anomaly is forecast to occur about a week earlier than in observations (cf. Figure 2). Nevertheless, forecast negative NAM anomalies last in the troposphere, as well as in the stratosphere, for several weeks, similarly to what was reported by Karpechko *et al.*,

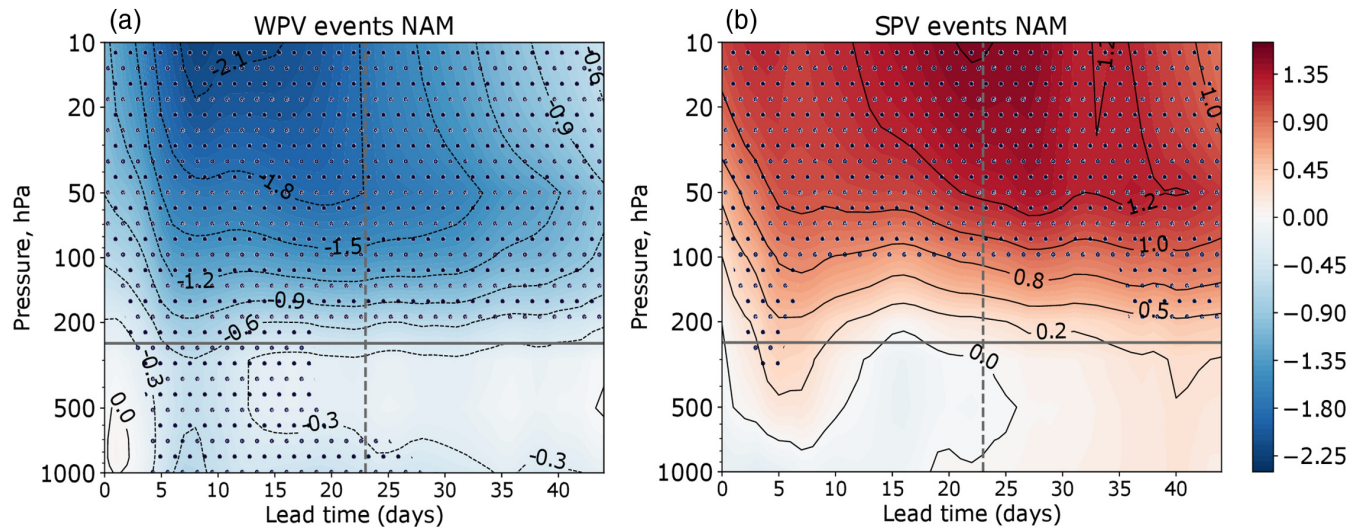


FIGURE 7 Cross-sections of the daily mean Northern Annular Mode (NAM) index composites (in standard deviation units) in hindcasts initialized with the lead time of 22–24 days before the cold air outbreak onset dates for cold spells preceded by the weak polar vortex (WPV) (a) and by the strong polar vortex (SPV) (b) events. Gray horizontal lines indicate the approximate boundary between the stratosphere and the troposphere. Gray vertical dashed lines indicate the onset of the cold spell events. Note that day 23 here corresponds to day 0 in Figure 2. Hatching denotes areas where the NAM index is significantly different from zero ($p < 0.01$) [Colour figure can be viewed at wileyonlinelibrary.com]

(2018) for the 2018 SSW case. This result suggests that the stratosphere–troposphere coupling likely contributes to the negative NAM regime, extending the predictability of the cold-air outbreaks (Figure 7a). On the other hand, in the SPV group the negative NAM phase in the troposphere before and after the onset of the cold spells is weaker than that in the reanalysis (Figure 7b). Moreover, in the hindcasts, insignificant NAM+ signal appears in the troposphere several days after the cold-spell onset dates (after day 25 in Figure 7b), in contrast to the observed NAM– regime (Figure 2) but in agreement with the expected sign of the SPV influence. Note that the stratospheric NAM+ signal in the SPV group is well forecast, including the strengthening of NAM+ before the onset of the cold spells (cf. Figure 2).

Looking at the near-surface forecast, in the WPV events (Figure 8a) the NAO/NAM– circulation pattern is distinct with a strong dipole of high-pressure anomalies over Greenland and the North Atlantic and low-pressure anomalies over the Atlantic Ocean in mid-latitudes. In the SPV composite (Figure 8b) the geopotential height anomalies are much weaker overall. The high-pressure anomalies are centered over northern Europe and, additionally, the second anticyclone is located in the northern Pacific, consistent with observations. Weak low-pressure anomalies are located over the northern part of North America and in the northern Atlantic. The forecast synoptic situation is consistent with our observation analysis in both cases (cf. Figure 3). The RMSE in the SPV cases

is significantly higher than that in the WPV cases starting from approximately day 19 of the forecast (Figure 8c). Figure 8c shows the error saturation level (ESL) defined as the standard deviation of the verifying observations multiplied by the square root of two (Simmons *et al.*, 1995; Bengtsson *et al.*, 2008). The RMSE for the WPV cases remains below ESL at most days, suggesting higher skill compared to the SPV group forecasts.

5 | COMPARISON OF DYNAMICAL AND STATISTICAL FORECASTS

To better understand the mechanisms associated with predictability gain from the stratosphere–troposphere coupling in the ECMWF model, we compare ECMWF hindcasts to forecasts made by the statistical model described in Section 2: *Data and Methods* (Figure 9). The comparison of the predictability as a function of the lead time averaged over seven days shows that the strength of the anomalies predicted by the statistical model is comparable to those by the ECMWF model in both WPV and SPV cases at most lags except at the shortest ones (day 1–7) when the statistical model underestimates the magnitude of the observed anomalies (Figure 9a and b). The predicted probability of an extreme event is shown in Figure 9c. In the statistical model the predicted cold-spell probabilities are comparable to those in the ECMWF hindcasts for both WPV and SPV cases at lead times longer than 14 days. The predicted

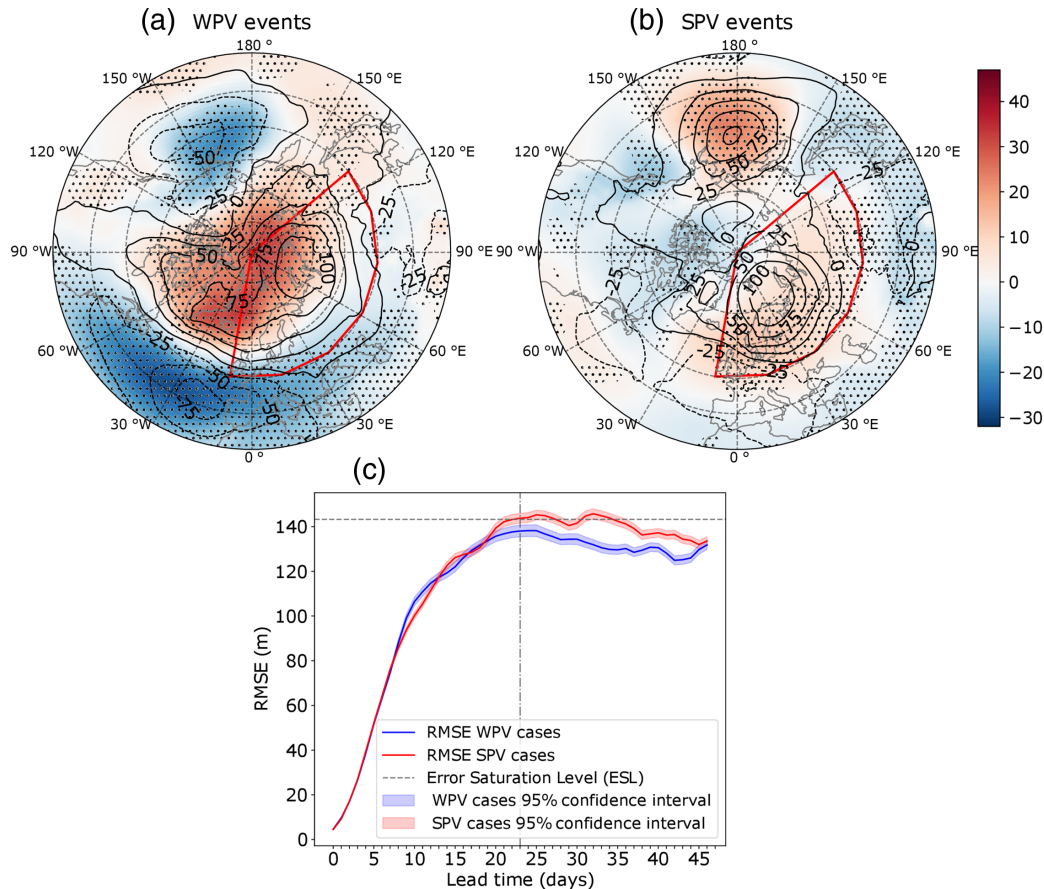


FIGURE 8 Weekly averaged composites of Z1000 anomalies (m) in hindcasts initialized with the lead time of 22–24 days before the cold air outbreak onset dates for (a) weak polar vortex (WPV) events and (b) strong polar vortex (SPV) events. Weekly averaging starts on the cold spell onset date. Contours indicate corresponding Z1000 anomalies in ERA5. Hatching denotes areas where the anomalies are significantly different from zero ($p < 0.01$). The red boxes show the area used to compute the root mean square errors (RMSEs) in panel (c) (10°W – 130°E and 50°N – 90°N). (c) Composites of the RMSE of hindcasts with respect to ERA5 for the WPV events (blue line) and SPV events (red line). The shaded area represents the 95% confidence interval for the RMSE curves. Gray dashed line represents the error saturation level. Gray vertical dash-dotted line indicates the onset of the cold spell events [Colour figure can be viewed at wileyonlinelibrary.com]

probabilities are consistently higher, by 0.1–0.2, for the WPV cases than those for the SPV cases at lead times of 1–21 days, while for longer lead times the probabilities in both groups are close to climatological values.

Removing the stratospheric predictors (PCs of the 150-hPa geopotential height anomalies, see Section 2: *Data and Methods*) from the predictors set of the statistical model provides a neat test for the stratospheric contribution in the predictability. Figure 9d shows that removing the stratospheric predictors results in lower predicted probabilities of the cold spells, by 0.03–0.06, at all lead times for the WPV cases. This result is consistent with findings by Karpechko (2015) for monthly forecasts, and it strongly supports the idea that the stratospheric state provides enhanced predictability of cold spells in Eurasia at S2S timescales. On the other hand, the predicted probabilities in the SPV cases do not differ much between the model runs at all lead times, suggesting that the

stratospheric information does not improve predictability in these cases. Interestingly, the predicted probabilities in the WPV cases remain larger than those in the SPV cases by ~ 0.1 at lead times of 8–21 days. This suggests that the enhanced predictability is also associated with the SLP and T2M states used as predictors in these runs.

6 | DISCUSSION AND CONCLUSIONS

The main purpose of this study is to detect and quantify the enhanced predictability of extreme cold events arising from the downward influence of the stratosphere on the surface weather, extending previous studies (e.g. Kautz *et al.*, 2020). We investigated the predictability of Eurasian cold spells (defined as the lowest 10th percentile across all weekly anomalies) during extended winter by

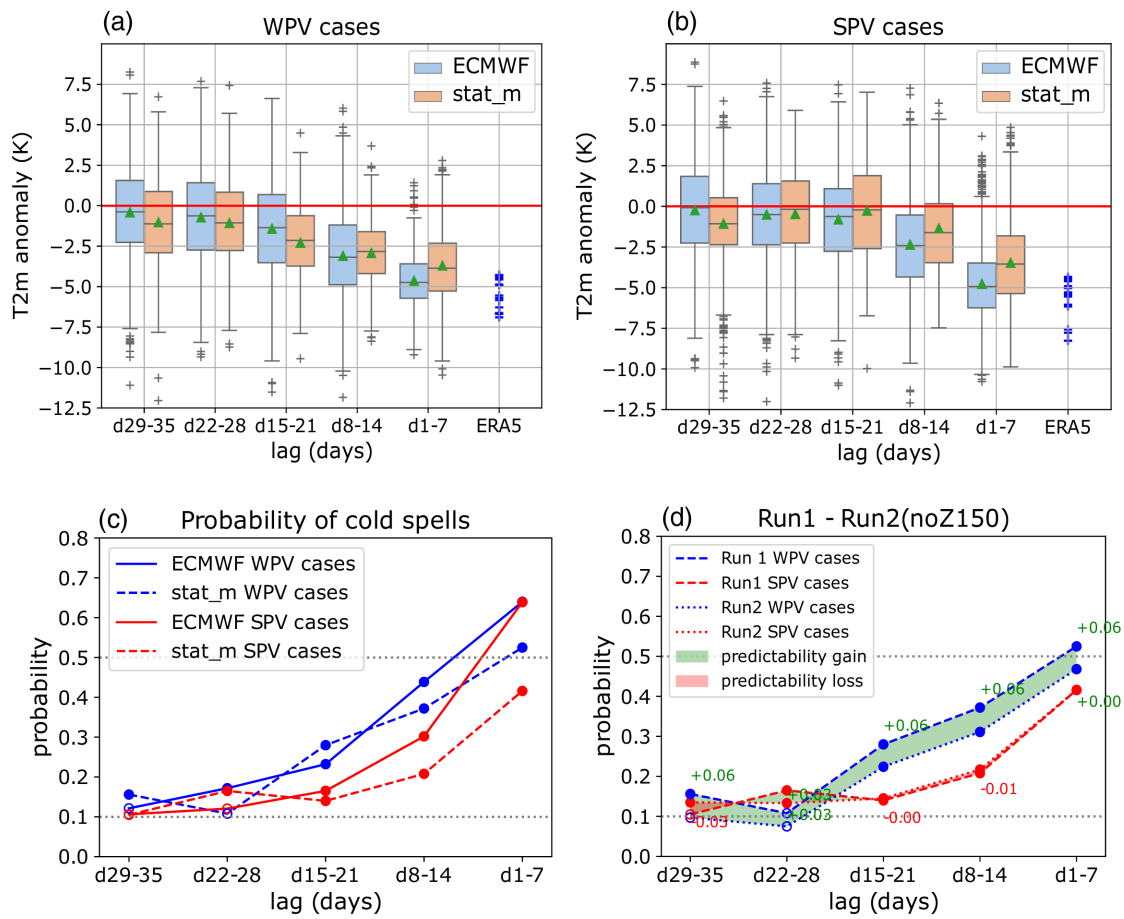


FIGURE 9 Upper panels: Boxplots of the weekly averaged Eurasian 2-m temperature (T2m) anomaly forecasts (in K) by the European Centre for Medium-Range Weather Forecasts (ECMWF) and the statistical models for different initialization lags for the (a) weak polar vortex (WPV) cases and (b) strong polar vortex (SPV) cases. The boxes show the interquartile range (Q1–Q3), whiskers represent the ranges for the bottom 25% and the top 25% of the data values, excluding outliers, which are shown by gray crosses. The median is the black line inside each box and the mean of each box is shown by a green triangle. The horizontal red line shows 0 K. The blue crosses show the ERA5 T2m anomaly (K) for the 15 cold spell cases in each group. Lower panels: (c) predicted probability of the cold-spell events as a function of lead time for both models averaged by seven days. WPV events' probability is shown in blue color and that of the SPV events in red color. Probabilities of 0.5 and 0.1 are shown by dotted lines. Closed circles denote cases where the lower uncertainty range, defined using bootstrapping with the number of iterations $N = 10,000$, exceeds 0.1. (d) Difference in the predicted probability of the cold spells between Run 1 and Run 2 (excluded stratospheric predictor) by the statistical model. Green shading shows a predictability gain in Run 1 with respect to Run 2 and red shading shows a predictability loss. Green and red numbers represent the difference in the predictability between Run1 and Run 2. Closed circles denote cases where the difference between the respective lines is significant at the 0.01 level [Colour figure can be viewed at wileyonlinelibrary.com]

separating cases where predictability can be enhanced by the stratosphere–troposphere coupling from cases where such influence was not expected. We use two groups, each comprising 15 cold-spell events with contrasting stratospheric states preceding the events.

Our results show that, at the S2S range, the predicted probability of extreme cold spells that take place after a weakening of the stratospheric polar vortex is systematically higher at lead times of 2–4 weeks (7–24 days), by 0.05–0.2, than in the cases with a strong stratospheric vortex. This translates into extension of probabilistic predictability by up to 3–5 days for some probability levels. The difference can be explained by the enhanced

predictability of the surface NAO/NAM regime due to the stratosphere–troposphere coupling. Consistent with previous studies we showed that the cold spells in Eurasia take place mainly under the negative surface NAO/NAM conditions. The NAO/NAM– regime is often followed by an increase in frequency of occurrence of tropospheric winter blocking over the North Atlantic region and Greenland and cold-air advection from the polar region. The weakening of the vortex is often associated with the downward propagation of NAM– throughout the troposphere and onset of the surface NAO/NAM– conditions followed by an increase in frequency of occurrence of tropospheric winter blocking over the North Atlantic region

and cold-air outbreaks over northern Eurasia. In this case the group of hindcasts that were initialized with a weak stratospheric vortex properly represents the downward propagation of the signal from the stratosphere. Our results hold for a rather large averaging area and they do not necessarily apply for smaller areas where unpredictable noise dominates (Büeler *et al.*, 2020).

On the other hand, the strong stratospheric vortex is associated with a positive phase of NAM in the stratosphere. In this case, the onset of the negative surface NAO/NAM phase leading to the cold spells results from the internal tropospheric dynamics which is associated with variability at shorter timescales and, therefore, shorter predictability as compared to the stratosphere. Although in the strong vortex cases the ECMWF hindcasts capture the transition of the tropospheric NAM regime into a negative phase, the predicted negative NAM in the troposphere is weaker and shorter-lasting compared to the weak vortex cases, as well as with the reanalysis. While cold spells during the periods of a weak stratospheric polar vortex might be more persistent, we found that the persistence was not the key factor that influenced the different predictability in our groups. Comparison of the forecasts by the ECMWF model with those by a statistical model suggests that the extended predictability originated from the stratosphere–troposphere coupling is an inherent property of the atmosphere which can be captured even by a simple linear model (Beerli *et al.*, 2017; Karpechko *et al.*, 2017). In fact, the superiority of the complex dynamical forecast system is not obvious until the weather timescale (lags less than a week) when the ECMWF model, but not the statistical model, starts to capture the magnitude of the extreme temperatures.

Analysis of the predictability of the Eurasian cold spells after different stratospheric states using the ECMWF ensemble hindcasts offers important insights about the influence of the stratosphere on the tropospheric dynamics. As the WPV can act as a precursor for the onset of the tropospheric NAO/NAM regime that favors the below-average temperatures over Eurasia, the predictability horizon of these events can be extended. There are other potential remote forcings that can influence the predictability of the NAO/NAM regime and cold spells in the troposphere, for example the Madden–Julian Oscillation (MJO) (Lin *et al.*, 2009; Vitart and Robertson, 2018) or the El Niño–Southern Oscillation (ENSO) (Mezzina *et al.*, 2020). In this study we did not analyze the potential influence of these factors, but it would be important to study the relative contributions from remote forcings alongside with the stratospheric influence and develop a better understanding of the teleconnections to improve winter weather predictions.

While the present study focused on the predictability of cold spells during periods of a WPV, it is also known that there is enhanced skill in forecasts initialized during a SPV (e.g. Tripathi *et al.*, 2015). Thus, it is of interest to investigate whether the predictability of anomalously mild periods following SPV events is enhanced. Also, it is of interest to see whether the influence of stratosphere–troposphere coupling on the predictability of Eurasian cold spells is similarly captured by the other S2S models. This is planned as a continuation of the present study.

AUTHOR CONTRIBUTIONS

Irina Statnaia: Conceptualization; data curation; formal analysis; investigation; methodology; visualization; writing – original draft; writing – review and editing. **Alexey Yu. Karpechko:** Conceptualization; methodology; supervision; writing – review and editing. **Matti Kämäräinen:** Data curation; formal analysis; investigation; writing – review and editing. **Heikki Järvinen:** Supervision; writing – review and editing.

ACKNOWLEDGEMENTS


The authors are grateful to Inna Polichtchouk (ECMWF, Reading, UK) and three anonymous reviewers whose comments led to the improvement of this paper. Irina Statnaia is funded by the Magnus Ehrnrooth Foundation and the Finnish Meteorological Institute (FMI). Alexey Karpechko is partly funded by the Academy of Finland (grant 317999). This work is based on (subseasonal-to-seasonal) S2S data. S2S is a joint initiative of the World Weather Research Programme (WWRP) and the World Climate Research Programme (WCRP).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Irina Statnaia  <https://orcid.org/0000-0003-3066-1902>

Alexey Karpechko  <https://orcid.org/0000-0003-0902-0414>

Matti Kämäräinen  <https://orcid.org/0000-0001-9180-6779>

Heikki Järvinen  <https://orcid.org/0000-0003-1879-6804>

REFERENCES

- Baldwin, M.P. and Dunkerton, T.J. (2001) Stratospheric harbingers of anomalous weather regimes. *Science*, 294(5542), 581–584. <https://doi.org/10.1126/science.1063315>.
- Baldwin, M.P., Stephenson, D.B., Thompson, D.W.J., Dunkerton, T.J., Charlton, A.J. and O'Neill, A. (2003) Stratospheric memory and skill of extended-range weather forecasts. *Science*, 301(5633), 636–640. <https://doi.org/10.1126/science.1087143>.

- Beerli, R., Wernli, H. and Grams, C.M. (2017) Does the lower stratosphere provide predictability for month-ahead wind electricity generation in Europe? *Quarterly Journal of the Royal Meteorological Society*, 143, 3025–3036. <https://doi.org/10.1002/qj.3158>.
- Bengtsson, L.K., Magnusson, L. and Källén, E. (2008) Independent estimations of the asymptotic variability in an ensemble forecast system. *Monthly Weather Review*, 136(11), 4105–4112. <https://doi.org/10.1175/2008MWR2526.1>.
- Büeler, D., Beerli, R., Wernli, H. and Grams, C.M. (2020) Stratospheric influence on ECMWF sub-seasonal forecast skill for energy-industry-relevant surface weather in European countries. *Quarterly Journal of the Royal Meteorological Society*, 146(733), 3675–3694. <https://doi.org/10.1002/qj.3866>.
- Charlton, A. J., O'Neill, A., Stephenson, D. B., Lahoz, W. A. and Baldwin, M. P. (2003) Can knowledge of the state of the stratosphere be used to improve statistical forecasts of the troposphere? *Quarterly Journal of the Royal Meteorological Society*, 129(595 PART B), 3205–3224. <https://doi.org/10.1256/qj.02.232>.
- Christiansen, B. (2005) Downward propagation and statistical forecast of the near-surface weather. *Journal of Geophysical Research*, 110, 1–10. <https://doi.org/10.1029/2004JD005431>.
- 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., Marshall, A., Roff, G., Schwartz, C., Simpson, I.R., Son, S. and Taguchi, M. (2020) The role of the stratosphere in subseasonal to seasonal prediction part II: predictability arising from stratosphere - troposphere coupling. *Journal of Geophysical Research: Atmospheres*, 125, 2019JD030923. <https://doi.org/10.1029/2019JD030923>.
- Ferranti, L., Magnusson, L., Vitart, F. and Richardson, D.S. (2018) How far in advance can we predict changes in large-scale flow leading to severe cold conditions over Europe? *Quarterly Journal of the Royal Meteorological Society*, 144(715), 1788–1802. <https://doi.org/10.1002/qj.3341>.
- Garfinkel, C.I., Hartmann, D.L. and Sassi, F. (2010) Tropospheric precursors of anomalous northern hemisphere stratospheric polar vortices. *Journal of Climate*, 23(12), 3282–3299. <https://doi.org/10.1175/2010JCLI3010.1>.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.N. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>.
- Kämäräinen, M., Uotila, P., Karpechko, A.Y., Hyvärinen, O., Lehtonen, I. and Räisänen, J. (2019) Statistical learning methods as a basis for skillful seasonal temperature forecasts in Europe. *Journal of Climate*, 32(17), 5363–5379. <https://doi.org/10.1175/JCLI-D-18-0765.1>.
- Karpechko, A.Y. (2015) Improvements in statistical forecasts of monthly and two-monthly surface air temperatures using a stratospheric predictor. *Quarterly Journal of the Royal Meteorological Society*, 141(691), 2444–2456. <https://doi.org/10.1002/qj.2535>.
- Karpechko, A.Y. (2018) Predictability of sudden stratospheric warmings in the ECMWF extended-range forecast system. *Monthly Weather Review*, 146(4), 1063–1075. <https://doi.org/10.1175/MWR-D-17-0317.1>.
- Karpechko, A. Y., Charlton-Perez, A., Balmaseda, M., Tyrrell, N. and Vitart, F. (2018) Predicting sudden stratospheric warming 2018 and its climate impacts with a multimodel ensemble. *Geophysical Research Letters*, 45(24), 13,538–13,546. <https://doi.org/10.1029/2018GL081091>.
- Karpechko, A.Y., Hitchcock, P., Peters, D.H.W. and Schneider, A. (2017) Predictability of downward propagation of major sudden stratospheric warmings. *Quarterly Journal of the Royal Meteorological Society*, 143(704), 1459–1470. <https://doi.org/10.1002/qj.3017>.
- Kautz, L.A., Polichtchouk, I., Birner, T., Garny, H. and Pinto, J.G. (2020) Enhanced extended-range predictability of the 2018 late-winter Eurasian cold spell due to the stratosphere. *Quarterly Journal of the Royal Meteorological Society*, 146(2019), 1040–1055. <https://doi.org/10.1002/qj.3724>.
- Kawatani, Y., Hamilton, K., Gray, L.J., Osprey, S.M., Watanabe, S. and Yamashita, Y. (2019) The effects of a well-resolved stratosphere on the simulated boreal winter circulation in a climate model. *Journal of the Atmospheric Sciences*, 76(5), 1203–1226. <https://doi.org/10.1175/JAS-D-18-0206.1>.
- Kidston, J., Scaife, A.A., Hardiman, S.C., Mitchell, D.M., Butchart, N., Baldwin, M.P. and Gray, L.J. (2015) Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. *Nature Geoscience*. Nature Publishing Group, 8(6), 433–440. <https://doi.org/10.1038/NGEO2424>.
- 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(24), 13943–13961. <https://doi.org/10.1029/2019JD030480>.
- Kretschmer, M., Cohen, J., Matthias, V., Runge, J. and Coumou, D. (2018) The different stratospheric influence on cold-extremes in Eurasia and North America. *Npj Climate and Atmospheric Science*, 1, 44. <https://doi.org/10.1038/s41612-018-0054-4>.
- Lin, H., Brunet, G. and Derome, J. (2009) An observed connection between the North Atlantic oscillation and the madden-Julian oscillation. *Journal of Climate*, 22(2), 364–380. <https://doi.org/10.1175/2008JCLI2515.1>.
- Mezzina, B., García-Serrano, J., Bladé, I. and Kucharski, F. (2020) Dynamics of the ENSO teleconnection and NAO variability in the North Atlantic-European late winter. *Journal of Climate*, 33(3), 907–923. <https://doi.org/10.1175/JCLI-D-19-0192.1>.
- Orsolini, Y.J., Karpechko, A.Y. and Nikulin, G. (2009) Variability of the northern hemisphere polar stratospheric cloud potential: the role of North Pacific disturbances. *Quarterly Journal of the Royal Meteorological Society*, 135, 1020–1029. <https://doi.org/10.1002/qj.409>.
- Poli, P., Hersbach, H., Dee, D.P., Berrisford, P., Simmons, A.J., Vitart, F., Laloyaux, P., Tan, D.G.H., Peubey, C., Thépaut, J.N., Trémolet, Y., Hólm, E.V., Bonavita, M., Isaksen, L. and Fisher, M. (2016) ERA-20C: an atmospheric reanalysis of the twentieth century. *Journal of Climate*, 29(11), 4083–4097. <https://doi.org/10.1175/JCLI-D-15-0556.1>.
- Räisänen, J. and Rätty, O. (2013) Projections of daily mean temperature variability in the future: cross-validation tests with

- ENSEMBLES regional climate simulations. *Climate Dynamics*, 41(5–6), 1553–1568. <https://doi.org/10.1007/s00382-012-1515-9>.
- Scaife, A.A., Karpechko, A.Y., Baldwin, M.P., Brookshaw, A., Butler, A.H., Eade, R., Gordon, M., MacLachlan, C., Martin, N., Dunstone, N. and Smith, D. (2016) Seasonal winter forecasts and the stratosphere. *Atmospheric Science Letters*, 17(1), 51–56. <https://doi.org/10.1002/asl.598>.
- Sigmond, M., Scinocca, J.F., Kharin, V.V. and Shepherd, T.G. (2013) Enhanced seasonal forecast skill following stratospheric sudden warmings. *Nature Geoscience. Nature Publishing Group*, 6(2), 98–102. <https://doi.org/10.1038/ngeo1698>.
- Simmons, A. J., Mureau, R. and Petroligis, T. (1995) ‘Error growth and estimates of predictability from the ECMWF forecasting system’, *Quarterly Journal of the Royal Meteorological Society*, 121, pp. 1739–1771. doi: <https://doi.org/https://doi.org/10.1002/qj.49712152711>.
- Thompson, D.W.J., Baldwin, M.P. and Wallace, J.M. (2002) Stratospheric connection to northern hemisphere wintertime weather: implications for prediction. *Journal of Climate*, 15(12), 1421–1428. [https://doi.org/10.1175/1520-0442\(2002\)015<1421:SCTNHW>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1421:SCTNHW>2.0.CO;2).
- Tripathi, O.P., Charlton-Perez, A., Sigmond, M. and Vitart, F. (2015) Enhanced long-range forecast skill in boreal winter following stratospheric strong vortex conditions. *Environmental Research Letters*, 10(10), 104007. <https://doi.org/10.1088/1748-9326/10/10/104007>.
- Vitart, F. (2014) Evolution of ECMWF sub-seasonal forecast skill scores. *Quarterly Journal of the Royal Meteorological Society*, 140(683), 1889–1899. <https://doi.org/10.1002/qj.2256>.
- Vitart, F., Cunningham, C., DeFlorio, M., Dutra, E., Ferranti, L., Golding, B., Hudson, D., Jones, C., Lavaysse, C., Robbins, J. and Tippett, M. K. (2019) ‘Sub-seasonal to seasonal prediction of weather extremes’, in Robertson, A. W. and Vitart, F. (eds) *sub-seasonal to seasonal prediction: the gap between weather and climate forecasting*. Elsevier, pp. 365–386. <https://doi.org/10.1016/B978-0-12-811714-9.00017-6>.
- Vitart, F. and Robertson, A. W. (2018) ‘The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events’, *npj climate and atmospheric science*. Springer US, 1(1), pp. 1–7. <https://doi.org/10.1038/s41612-018-0013-0>.
- White, C.J., Carlsen, H., Robertson, A.W., Klein, R.J.T., Lazo, J.K., Kumar, A., Vitart, F., Coughlan de Perez, E., Ray, A.J., Murray, V., Bharwani, S., MacLeod, D., James, R., Fleming, L., Morse, A.P., Eggen, B., Graham, R., Kjellström, E., Becker, E., Pegion, K.V., Holbrook, N.J., McEvoy, D., Depledge, M., Perkins-Kirkpatrick, S., Brown, T.J., Street, R., Jones, L., Remenyi, T.A., Hodgson-Johnston, I., Buontempo, C., Lamb, R., Meinke, H., Arheimer, B. and Zebiak, S.E. (2017) Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorological Applications*, 24(3), 315–325. <https://doi.org/10.1002/met.1654>.
- White, C. J., Domeisen, D. I. V, Acharya, N., Adefisan, E. A., Anderson, M. L., Aura, S., Balogun, A. A., Bertram, D., Bluhm, S., Brayshaw, D. J., Browell, J., Büeler, D., Charlton-perez, A., Christel, I., Coelho, C. A. S., DeFlorio, M. J., Monache, D., Giuseppe, F. Di, García-solórzano, A. M., Gibson, P. B., Goddard, L., Romero, C. G., Graham, R. J., Graham, R. M., Grams, C. M., Halford, A., Huang, W. T. K., Jensen, K., Kilavi, M., Lawal, K. A., Lee, W., Macleod, D., Manrique-suñén, A., Martins, E. S. P. R., Carolyn, J., Merryfield, W. J., Muñoz, Á. G., Olanian, E., Otieno, G., Oyedepo, A., Palma, L., Pechlivanidis, I. G., Pons, D., Ralph, F. M., Dirceu, S., Jr, R., Remenyi, T. A., Risbey, J. S., Robertson, D. J. C., Andrew, W., Smith, S., Soret, A., Sun, T., Todd, M. C., Tozer, C. R., Jr, F. C. Vasconcelos, Vigo, I., Waliser, D. E., Wetterhall, F. and Wilson, G. (2021) ‘Advances in the application and utility of subseasonal-to-seasonal predictions’. *Bulletin of the American Meteorological Society*, 103(6), E1448–E1472. <https://doi.org/10.1175/bams-d-20-0224.1>.

How to cite this article: Statnaia, I., Karpechko, A., Kämäräinen, M. & Järvinen, H. (2022) Stratosphere–troposphere coupling enhances subseasonal predictability of Northern Eurasian cold spells. *Quarterly Journal of the Royal Meteorological Society*, 148(747), 2769–2783. Available from: <https://doi.org/10.1002/qj.4335>