Long-term prediction of Sudden Stratospheric Warmings with Geomagnetic and Solar Activity 2

Article *in* Journal of Geophysical Research Atmospheres · November 2022

4 authors:

Mikhail V. Vokhmyanin
University of Oulu

Timo Asikainen
University of Oulu

Antti Salminen
University of Oulu

Kalevi Mursula
University of Oulu

Long-term prediction of Sudden Stratospheric Warmings with Geomagnetic and Solar Activity

Mikhail Vokhmyanin[1], Timo Asikainen[1], Antti Salminen[1], Kalevi Mursula[1]

[1]Space physics and astronomy research unit, University of Oulu, PO Box 3000, 90014 Oulu, Finland

Key Points:

- We develop models, predicting the wintertime SSW probability in preceding August using QBO, geomagnetic aa index and solar F10.7 index
- The success ratio of the predictions is about 86%
- The study is an important step towards improved long-term forecasting of SSWs

Corresponding author: Mikhail Vokhmyanin, `mikhail.vokhmianin@oulu.fi`

Abstract

The polar vortex is a strong jet of westerly wind which forms each winter around the polar stratosphere. Sometimes, roughly every other winter, the polar vortex in the Northern Hemisphere experiences a dramatic breakdown and associated warming of the polar stratosphere. Such events are called sudden stratospheric warmings (SSW) and they are known to have a significant influence on ground weather in Northern Eurasia and large parts of North America. Typically, these events are thought to occur due to planetary waves propagating to the stratosphere where they may disrupt the vortex. Here, we show that the SSW probability depends significantly on a favorable combination of geomagnetic and solar activity and the phase of the Quasi-Biennial Oscillation (QBO). Using logistic regression models, we find that more SSWs occur when early-winter geomagnetic activity (aa index) is low and QBO winds are easterly and when solar activity (F10.7 index) is high and QBO winds are westerly. We then examine the possibility of using these results to predict the occurrence probability of SSWs with several months lead time and evaluate the optimal lead times for all variables using cross-validation methods. As a result, we find that the SSW probability can be predicted rather well and we can issue a probabilistic SSW prediction for the coming winter season with a success ratio of about 86% already in the preceding August. The results presented here are an important step toward improving the seasonal predictability of wintertime weather using information about solar and geomagnetic activity.

1 Introduction

The wintertime polar stratosphere is characterized by the polar vortex, strong westerly winds circulating the pole. The polar vortex results from cooling of high-latitude air, which starts already in fall and reaches its peak in mid-winter. According to the thermal wind shear balance, the cool air enhances the meridional temperature gradient which, corresponds to enhanced westerly winds of the polar vortex. While the vortex is formed due to radiative cooling of the polar stratosphere, it is also greatly influenced by atmospheric waves, dominantly by planetary-scale Rossby waves, which originate in the troposphere and propagate vertically through relatively weak westerly winds (Charney & Drazin, 1961). When planetary waves converge in the stratosphere, they can deposit easterly momentum on the background flow and therefore decelerate the westerly winds (Matsuno, 1971; Polvani & Waugh, 2004). Such wave activity also enhances the meridional circulation, which adiabatically warms the polar stratosphere (e.g., Salby & Callaghan, 2002). Wave activity is largely responsible for driving a global meridional circulation, so-called Brewer-Dobson circulation (BDC) from lower latitudes toward the winter pole (Butchart, 2014). Often, if planetary wave activity on the vortex is strong enough, the forcing may weaken the vortex so much that it completely breaks down and even reverses or splits into several vortex cells. Such events are called sudden stratospheric warmings (SSWs) due to the accompanied rapid warming of the polar stratosphere (Matsuno, 1971; Dunkerton et al., 1981). There are also so-called final warmings which eventually always occur in the spring and after which the vortex does not recover. While a polar vortex forms in both hemispheres during the local winter season, the SSWs are almost exclusively a Northern Hemisphere phenomenon due to the larger land-sea contrast and to the more intense orographic features (mountains etc.) which cause large planetary wave activity in the Northern Hemisphere (van Loon et al., 1973; Garfinkel et al., 2020). Yet, SSWs have been observed in the Southern Hemisphere as well, e.g., in 2002 (Krüger et al., 2005) and 2019 (Hendon et al., 2019; Rao, Garfinkel, White, & Schwartz, 2020).

Sudden stratospheric warmings are dramatic dynamical events in the wintertime stratosphere and often have long-lasting effects on wintertime ground weather in large parts of the Northern Hemisphere (Baldwin et al., 2021). For example, Northern Europe often experiences cold and dry conditions for weeks after an SSW (Butler et al., 2017; Baldwin et al., 2021), thereby causing significant societal and economic impacts. The surface effect of SSWs is seen in the geopotential anomalies, Northern Annular Mode (NAM) and North Atlantic Oscillation (NAO) (Baldwin & Dunkerton, 1999, 2001), and also in the modulation of mean and extreme climate conditions in Europe (King et al., 2019). The response to SSW events is often described by the negative phase of the NAM/NAO, although it does not occur after every SSW (Charlton-Perez et al., 2018; Domeisen, 2019; White et al., 2019). Palmeiro et al. (2015) found that the stronger and more coherent tropospheric signals related to SSWs are caused by major SSWs when polar vortex reverses, while minor warmings without vortex disruption yield less robust signal. White et al. (2020) also found that the tropospheric response almost linearly depends on the strength of the SSW.

The formation of SSWs has been shown to greatly depend on the intensity of planetary wave propagation into the stratosphere and the state of the polar vortex itself (e.g., de la Cámara et al., 2017; Matsuno, 1971; Scott & Polvani, 2004). Because of this, various internal and external factors influencing planetary wave activity or the state of the polar vortex have been shown to influence the frequency of SSWs. Examples of such internal atmospheric factors are the El Niño–Southern Oscillation (ENSO; Butler & Polvani, 2011; Garfinkel, Butler, et al., 2012; Polvani et al., 2017; Domeisen et al., 2019), the Madden-Julian Oscillation (MJO; Garfinkel & Schwartz, 2017; Schwartz & Garfinkel, 2017), the amount of volcanic aerosols in the stratosphere (van Loon & Labitzke, 1987), and the late-fall snow cover in Eurasia (Cohen et al., 2007; Henderson et al., 2018). Probably the most significant and well known influence on SSW occurrence is exerted by the stratospheric Quasi-Biennial Oscillation (QBO; Holton & Tan, 1980; Anstey & Shepherd, 2014; Garfinkel et al., 2018; Rao, Garfinkel, & White, 2020). QBO is a mode of alternating zonal winds in the tropical stratosphere with an approximate period of 28–34 months, forming a downward propagating pattern of the zonal winds (Baldwin et al., 2001). Holton and Tan (1980) were the first to show that the polar vortex is weaker when QBO at 50 hPa is in the easterly phase and stronger when QBO is in the westerly phase. This result is often referred to as the Holton-Tan effect. The cause of this QBO modulation of the polar vortex is often thought to result from the fact that easterly QBO causes the critical line (location where the zonal wind reverses from westerly to easterly) to be shifted toward the winter hemisphere (Holton & Tan, 1980). Because planetary waves cannot propagate in easterly winds the poleward shift of the critical line guides more planetary waves towards the winter polar stratosphere, which then weakens the polar vortex. Another influence may come from the fact that QBO modulates the meriodional circulation and the easterly QBO enhances the Brewer-Dobson circulation resulting in stronger downwelling and adiabatic warming in the polar stratosphere (Flury et al., 2013). Also other mechanisms to explain the QBO influence have been suggested (e.g., Garfinkel, Shaw, et al., 2012; Silverman et al., 2018; Watson & Gray, 2014; White et al., 2015). Regardless of the exact mechanism the easterly phase of QBO leads to a weaker vortex and because a weaker vortex is more susceptible to planetary wave activity (e.g., Matsuno, 1970) the likelihood of SSWs increases (decreases) during easterly (westerly) QBO phase (Labitzke, 1982).

Solar related factors including solar wind driven energetic particle precipitation into the upper polar atmosphere and the varying solar irradiance have also been found to influence the polar stratosphere and thereby have potential influence on the occurrence of SSWs. Energetic particle precipitation occurs mostly at high latitudes, and comes from several different sources, e.g., electrons from magnetospheric plasma sheet and radiation belts, highly energetic protons related to solar proton events and cosmic rays of galactic origin. Solar proton events are connected to solar eruptions (flares and coronal mass ejections) and therefore are relatively sporadic. However, the energetic electron precipitation (EEP) is driven by solar wind and is more or less continuously present. The particle precipitation and especially EEP in the polar region ionizes neutral atoms and molecules in the lower thermosphere and upper mesosphere forming reactive odd nitrogen (NO_x) and hydrogen (HO_x) oxides. These molecules participate in catalytic reactions resulting in ozone depletion (Crutzen et al., 1975). During winter, in the polar darkness the increased lifetime of NO_x allows them to descend in the downwelling part of the Brewer-Dobson circulation into the stratosphere where they can destroy ozone, leading to the so-called indirect effect of energetic particle precipitation (Randall et al., 2007; Funke et al., 2014).

In the polar mesosphere and upper stratosphere, ozone loss leads to a net radiative heating in mid-winter and to a radiative cooling in late winter and spring due to polar sunrise (Sinnhuber et al., 2018). These thermal changes intensify the polar vortex, which has been confirmed by observations (Lu et al., 2008; Seppälä et al., 2013; Salminen et al., 2019) and by models (Rozanov et al., 2005; Baumgaertner et al., 2011; Arsenovic et al., 2016). Recent studies have also shown that planetary wave activity which is suitably located with respect to the polar vortex is essential in order to allow the EEP effect on the polar vortex to take place (Asikainen et al., 2020; Salminen et al., 2022). One of the consequences of this is that the EEP influence on the polar vortex is predominantly observed during easterly phase of the QBO, when more planetary wave activity is concentrated into the polar region (Salminen et al., 2019). Similar QBO modulation has been found for the EEP effect on the tropospheric NAM indices, where the correlation was stronger in the easterly QBO phase (Palamara & Bryant, 2004; Maliniemi et al., 2013, 2016).

Solar UV irradiance roughly follows the sunspot cycle (L. E. Floyd et al., 2003; Fröhlich, 2006) and varies by up to 6% near 200 nm responsible for ozone production and by up to 4% near 240–320 nm responsible for UV absorption by ozone (Gray et al., 2010). Higher UV irradiance results in a warmer tropical upper stratosphere due to increased ozone production (Soukharev & Hood, 2006; Frame & Gray, 2010), while in the lower stratosphere the UV signal is seen in the circulation (Kodera & Kuroda, 2002; Salby & Callaghan, 2004) although some of the apparent lower stratospheric solar signal has been attributed to aliasing of major volcanic eruptions (Chiodo et al., 2014; Kuchar et al., 2017). Increased UV absorption at low latitudes during winter enhances the meridional temperature gradient and westerly winds in the polar vortex (e.g., Kodera & Kuroda, 2002; Gray et al., 2010). Also the influence of varying solar irradiance on the polar vortex has been found to be modulated by the QBO phase. Labitzke and van Loon (1988) found that during westerly QBO phase the polar lower stratosphere is warmer in solar maxima and cooler during easterly QBO. Camp and Tung (2007) found a positive correlation between late-winter polar stratosphere temperature and sunspot numbers in the westerly QBO phase but no correlation for the easterly QBO. It has also been found that mid-winter SSWs are more frequent when the QBO phase is easterly around solar minimum, while in the westerly QBO phase SSWs occur mostly when solar activity is at maximum (Labitzke, 1987; Gray et al., 2004; Labitzke et al., 2006; Gray et al., 2010). However, as reported by Baldwin et al. (2021) this relationship is modest for data updated to 2019.

The SSWs can be well predicted about two weeks in advance with numerical weather models (Tripathi et al., 2015; Karpechko, 2018; Domeisen et al., 2020) but, considering the above mentioned influences on polar vortex and SSW occurrence, e.g., by solar related factors and QBO, there may be potential for longer lead time predictability. Recently, Salminen et al. (2020) conducted a statistical study on the influence of several different internal and solar related factors on SSW occurrence. They found that the QBO and geomagnetic activity (which is an indirect measure for EEP) were the two most influential drivers affecting the SSW occurrence. More precisely, the latter was greatly enhanced when geomagnetic activity was lower than average and QBO was in the easterly phase. This means that in QBO-E phase enhanced (weakened) particle precipitation makes the polar vortex stronger (weaker) and less (more) vulnerable to SSWs. Motivated by this result, we examine in this paper the long-term predictability of SSW occurrence probability. We develop further the results found by Salminen et al. (2020) and build a model predicting the wintertime SSW occurrence probability before the winter season begins.

We will study how the combined effect of geomagnetic activity (*aa* index) or solar irradiance (F10.7 index) together with the QBO phase modulates SSW occurrence. We investigate the sensitivity of SSW probability on the timing of these explanatory factors in order to find the optimal time lags of these factors for SSW prediction. The paper is organized as follows. The data sets and methods are described in Section 2. In Sections 3 and 4, we investigate the effect of geomagnetic activity and solar irradiance, respectively, and the QBO phase on SSW occurrence rate in order to find the best combinations for prediction. In Section 5, we study what is the optimal length of the time window of geomagnetic/solar activity affecting the SSW occurrence. We also evaluate the performance of the final prediction model in Section 6. The discussion of the results and conclusions are given in Section 7.

2 Data and Methods

2.1 Reanalysis data and identification of SSWs

Over the years, many different definitions for a major SSW event have been suggested (Butler et al., 2015). However, perhaps the most commonly used definition is based on the reversal of the stratospheric zonal-mean zonal wind suggested by Charlton and Polvani (2007), which we also use in this work. According to this definition, the major SSW central date is defined as the day when the daily zonal-mean zonal wind at 10 hPa and 60°N latitude reverses to easterly in any of the northern winter months (November to March). In order to distinguish successive events, zonal wind must have returned to westerly for 20 consecutive days before the next event is identified. To exclude the final warming, the zonal wind has to return to westerly for at least 10 consecutive days before the end of April.

Note that although this definition does not explicitly involve the meridional temperature gradient, the required reversal of the zonal wind implies a reversal or at least a significant weakening of the temperature gradient in accordance with the thermal wind shear balance. An additional criterion for reversed meridional temperature gradient would make only a small difference to the list of SSWs (Charlton & Polvani, 2007). Although SSWs could also be defined based on other criteria (e.g., involving reversal of meridional temperature gradient or considering winds and temperatures at different latitudes and altitudes) the standard definition used here allows for a more direct comparison of our results with other statistical studies using the same definition (Butler et al., 2015) and has been shown to be optimal in terms of the stratospheric changes, wave forcing, and surface impact associated to the event (Butler & Gerber, 2018).

SSW identification requires zonal wind data which can be obtained from atmospheric reanalysis data sets. Reanalysis products are based on numerical weather and climate models, which assimilate a wide variety of atmospheric and other observations providing complete 3D fields of atmospheric variables as a function of time. Therefore, the reanalyses effectively fill the gaps in spatially and temporally irregular observations using numerical models. Because of this approach the reanalysis fields may be more model biased rather than correspond to the actual data within the intervals of sparse observations. Due to the differences in model construction and data quality as well as data assimilation techniques there are also some differences between different reanalysis products. In order to reduce the effects of possible uncertainties in zonal wind fields on the SSW identification, we consider in this work the major SSW events using several different reanalysis data sets: (1) the fifth generation of the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalysis of global climate ERA5 (Hersbach et al., 2020) available in 1950 – 2021, (2) the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) reanalysis (Kalnay et al., 1996) in 1948 – 2021, (3) second generation ECMWF reanalysis ERA40 (Uppala

Table 1. Central dates of the Northern Hemisphere SSWs in reanalysis products. The value after the date indicates the maximum easterly zonal-mean zonal wind at 60° N, 10 hPa during the SSW event. The *** notation indicates that an SSW was not detected in the corresponding reanalysis. The * after the reported date is the time when the zonal wind in the reanalysis almost reached a zero value in the vicinity of the SSW detected by the other reanalyses.

Year[a]	ERA5 1950 - 2021		NCEP/NCAR 1948 - 2021		ERA40 1958 - 2002		ERA-interim 1979 - 2019	
1950	05-Mar	-8.9	***	***				
1951	09-Feb	-7.3	***	***				
1952	19-Feb	-20.0	25-Feb	-6.1				
1953	19-Nov	-1.9	***	***				
1955	12-Jan	-2.7	***	***				
1957	04-Feb	-26.0	08-Feb	-0.4				
1958	01-Feb	-5.2	30-Jan	-13.3	31-Jan	-7.0		
1959	***	***	30-Nov	-5.8	***	***		
1960	17-Jan	-5.3	16-Jan	-2.0	15-Jan	-6.9		
1963	27-Jan	-6.8	12-Feb	0.8	28-Jan	-4.2		
1965	23-Mar*	1.3	23-Mar	-0.4	23-Mar*	1.3		
1966	16-Dec	-5.8	08-Dec	-9.4	16-Dec	-5.5		
	22-Feb	-7.8	24-Feb	-5.0	23-Feb	-7.0		
1968	07-Jan	-5.1	07-Jan*	1.1	07-Jan	-5.2		
1969	28-Nov	-4.8	27-Nov	-6.6	28-Nov	-3.9		
	13-Mar	-0.6	13-Mar	-0.2	13-Mar	-1.0		
1970	02-Jan	-13.6	02-Jan	-9.9	01-Jan	-13.7		
1971	18-Jan	-7.3	17-Jan	-9.4	18-Jan	-11.7		
	20-Mar	-4.0	20-Mar	-4.5	19-Mar	-5.0		
1973	31-Jan	-25.3	02-Feb	-13.8	31-Jan	-28.5		
1977	09-Jan	-2.8	11-Jan	0.0	09-Jan	-4.5		
1979	22-Feb	-13.0	22-Feb	-11.4	22-Feb	-17.1	22-Feb	-15.4
1980	29-Feb	-7.8	29-Feb	-8.7	29-Feb	-5.7	29-Feb	-7.0
1981	04-Mar	-0.7	04-Mar*	0.9	04-Mar	-1.2	04-Mar	-1.0
1982	04-Dec	-2.4	04-Dec	-0.5	04-Dec	-3.6	04-Dec	-2.0
1984	24-Feb	-11.3	24-Feb	-10.7	24-Feb	-10.5	24-Feb	-10.7
1985	01-Jan	-15.4	02-Jan	-11.9	01-Jan	-17.5	01-Jan	-16.3
1987	23-Jan	-20.6	23-Jan	-19.7	23-Jan	-22.5	23-Jan	-22.3
1988	08-Dec	-17.7	08-Dec	-16.1	07-Dec	-17.9	07-Dec	-17.5
	14-Mar	-3.6	14-Mar	-2.9	14-Mar	-4.3	15-Mar	-4.0
1989	21-Feb	-13.8	22-Feb	-12.0	21-Feb	-14.4	21-Feb	-14.5
1995	05-Feb*	0.3	04-Feb*	0.3	05-Feb	+0.3	05-Feb	-0.3
1999	15-Dec	-23.1	15-Dec	-17.4	15-Dec	-23.1	15-Dec	-22.3
	26-Feb	-17.5	25-Feb	-18.0	26-Feb	-18.4	26-Feb	-17.4
2000	20-Mar	-3.1	20-Mar	-3.7	20-Mar	-3.8	20-Mar	-3.3
2001	11-Feb	-12.3	11-Feb	-13.4	11-Feb	-12.3	11-Feb	-12.3
2002	30-Dec	-2.1	02-Jan	-0.8	30-Dec	-1.7	30-Dec	-2.0
	17-Feb	-0.4	17-Feb	1.8	17-Feb	-0.1	17-Feb	0.0
2003	18-Jan	-1.9	18-Jan	-1.5			18-Jan	-2.5
2004	05-Jan	-14.8	07-Jan	-11.1			05-Jan	-15.5
2006	21-Jan	-25.3	21-Jan	-22.3			21-Jan	-25.0
2007	24-Feb	-8.6	24-Feb	-8.8			24-Feb	-8.3
2008	22-Feb	-14.1	22-Feb	-13.4			22-Feb	-15.4
2009	24-Jan	-29.4	24-Jan	-28.6			24-Jan	-31.1
2010	09-Feb	-7.0	09-Feb	-5.7			09-Feb	-6.9
	24-Mar	-2.9	24-Mar	-1.7			24-Mar	-2.5
2013	06-Jan	-12.8	07-Jan	-10.0			06-Jan	-13.3
2017	01-Feb*	0.9	01-Feb*	1.5			01-Feb	-0.3
2018	12-Feb	-24.1	12-Feb	-23.2			11-Feb	-25.1
2019	01-Jan	-10.1	02-Jan	-9.0			01-Jan	-10.5
2021	05-Jan	-9.3	05-Jan	-7.0				

[a] year of winter is defined in January.

et al., 2005) in 1958 – 2002, and third generation ERA-Interim (Dee et al., 2011) in 1979 – 2019.

Table 1 displays the central dates of the major SSW events found in northern winters in 1950 – 2021 using the Charlton and Polvani (2007) definition and different reanalyses. In total, we find 47 SSW events in 40 winters in 72 years for ERA5, 38 events (33 winters) in 74 years for NCEP/NCAR, 29 events (23 winters) in 45 years for ERA40, and 28 events (25 winters) in 41 years for ERA-Interim. Note that the central dates and the maximum daily zonal easterly wind values related to the event often differ somewhat between the reanalyses.

The majority of events after 1960 are found in all reanalyses except for winters of 1965, 1995, and 2017 when the complete reversal of the zonal wind was found only in one reanalysis while in the others , e.g., in ERA5, it did not quite reach zero value. According to the SSW compendium by Butler et al. (2017) available at `https://csl.noaa.gov/groups/csl8/sswcompendium/majorevents.html`, none of these events are seen by JRA-55 (Japanese 55-year Reanalysis) and MERRA-2 (Modern-Era Retrospective analysis for Research and Applications version 2) reanalyses either. Note that the SSW compendium data covers only the years 1958 – 2020 and does not include the ERA5 reanalysis. According to ERA5, there are 6 additional events in 1950 – 1957 and one more in January of 2021 (also seen in NCEP/NCAR).

In Figure 1, we compare daily zonal-mean zonal wind at 10 hPa and 60°N latitude of all the four reanalyses from 1950 to 2021. One can see that until about 1965 the different reanalyses show quite large differences in the wintertime zonal wind. The difference between ERA5 and NCEP/NCAR is noticeable in the 1950s, especially in 1950 and 1951. Although both NCEP/NCAR and ERA5 are able to reproduce the first ever observed SSW event in February 1952 known as “Berlin Phenomenon” (Scherhag, 1952), the events in 1950, 1951, 1953, and 1955 are found only in ERA5, while SSW in the winter of 1959 is found only in NCEP/NCAR (also not seen by JRA-55 according to the SSW compendium).

It is known that the reliability of the reanalysis wind fields at stratospheric levels prior to 1958 gradually decreases toward earlier time due to lack of upper-air observations (Kistler et al., 2001). Butler et al. (2017) also suggests that the evolution of SSW events prior to 1964 should be viewed with caution as radiosonde measurements were very rare at that time. Many more data sources were used in the assimilation of ERA5 compared to older reanalyses, including digitized upper-air observations (Bell et al., 2021). Therefore, in this study, we consider all SSW events seen by ERA5 since 1952. In this way the list of SSWs is based on one reanalysis system during the whole, long time interval, and is expected to provide the most consistent and reliable estimate even in the early period in the 1950s and 1960s.

2.2 QBO data

We also use the ERA5 reanalysis data to calculate the QBO winds in the equatorial stratosphere. The QBO at different pressure levels (10, 20, 30, 50, and 70 hPa) is obtained from the monthly anomalies of the zonal mean zonal wind at the respective height averaged over 10°S—10°N. In order to have a clear separation between the two different QBO phases, we defined the QBO to be in the easterly phase when the equatorial zonal wind anomaly is negative and its magnitude is greater than half of the standard deviation of negative zonal wind anomalies. Similarly, westerly phase is defined by the zonal wind anomaly being larger than half of standard deviation of positive zonal wind anomalies. This approach reduces the uncertainty of defining a winter to the two QBO phases in such cases where the QBO zonal wind anomaly is close to zero. Note that according to Bell et al. (2021), zonal winds below 10 hPa are accurately reproduced by the

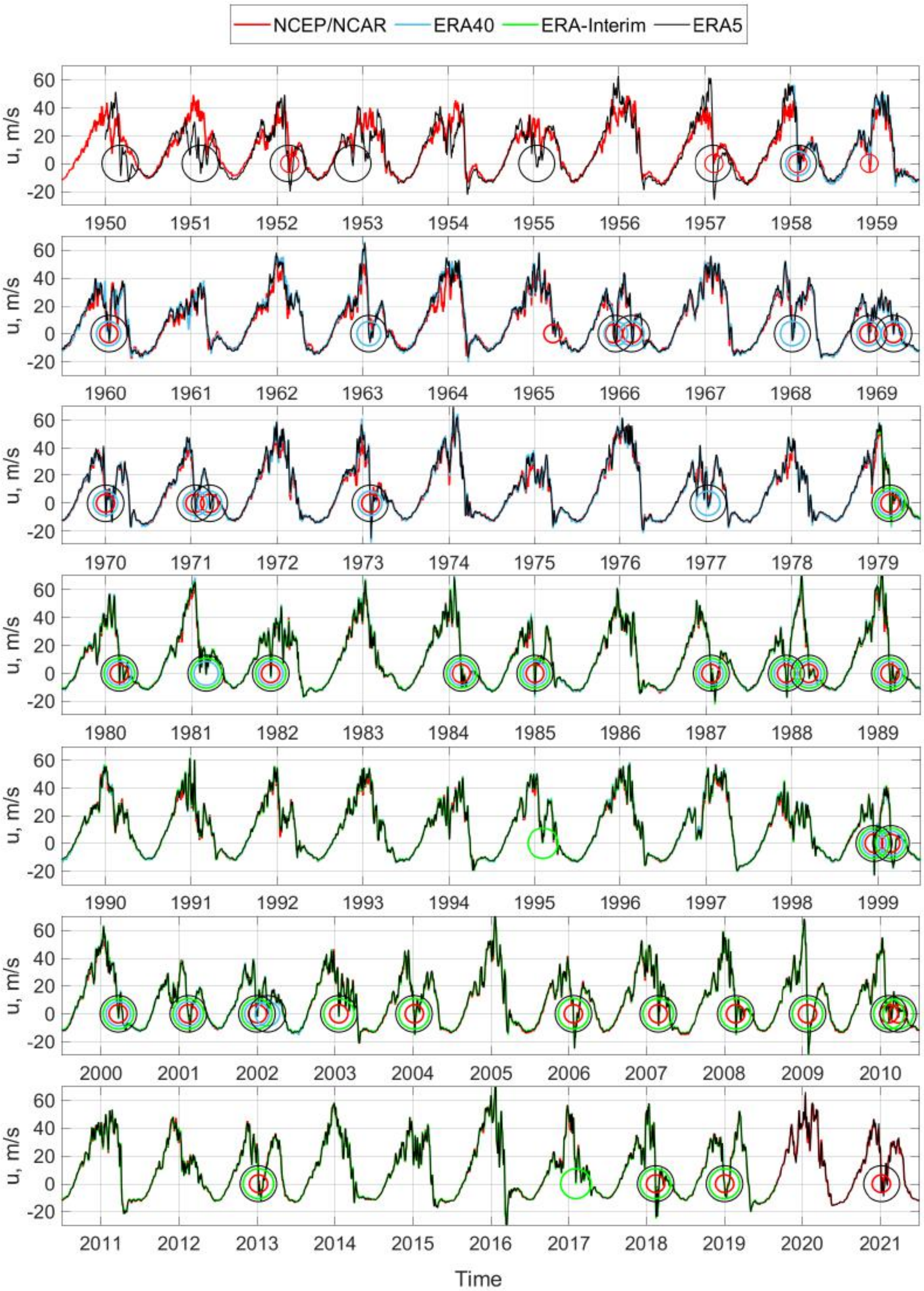

Figure 1. Zonal-mean wind speed at 10 hPa and 60°N according to ERA5 (black), NCEP/NCAR (red), ERA40 (blue), and ERA-Interim (green); circles indicate SSW dates identified using Charlton and Polvani (2007) definition using different reanalyses (same coloring).

ERA5 reanalysis back to 1950 with less than 2 m/s observational error, and are perfectly represented in points co-located with the observations.

2.3 Geomagnetic and solar data

In this work we consider how the occurrence probability of SSWs is influenced by geomagnetic activity and solar activity. We employ the geomagnetic *aa* index which measures global geomagnetic activity and is calculated from the magnetic variations at antipodal observatories in Britain and Australia. The *aa* index has been constructed since 1868 and is the longest continuous record of geomagnetic activity to date. In this study the *aa* index is being used as a proxy measure for energetic electron precipitation into the upper atmosphere.

As the index of solar activity, we use the solar F10.7 radio flux index which correlates with solar UV and total irradiance (L. Floyd et al., 2005; Gray et al., 2010). Solar UV irradiance is strongly absorbed in the stratosphere, thereby affecting the zonal wind distribution, potentially also the polar vortex (Labitzke & van Loon, 1988; Balachandran & Rind, 1995; Gray et al., 2004; Camp & Tung, 2007; Matthes et al., 2013; Mitchell et al., 2015; Garfinkel et al., 2015), see also the reviews by Gray et al. (2010) and Ward et al. (2021).

2.4 Logistic regression

In this study, we develop a model to estimate the probability that an SSW occurs during the winter. Each winter is assigned a binary value: a value of 1 for winters with an SSW event, and a value of 0 for those without an SSW. We model the occurrence probability with the logistic regression method, which is commonly used to model binomial outcomes. The logistic regression model expresses the probability of an event with a dependence on an explanatory variable X (or possibly many variables) according to equation

$$P = \frac{1}{1 + e^{-a-b\cdot X}} \tag{1}$$

This equation describes a curve, which smoothly changes between values of 0 and 1 when X goes from minus to plus infinity. The intercept a defines the location of P=0.5 point and the rate parameter b describes a steepness of the function. The steepness also indicates how strongly X increases ($b > 0$) or decreases ($b < 0$) the occurrence probability. Recently, Salminen et al. (2020) studied the dependence of SSW occurrence rate on several different factors (geomagnetic activity, sunspots, El Niño and QBO). They showed that SSW events occur more often during winters where geomagnetic activity (*ap* index in December) is low and which are preceded by easterly QBO at 30 hPa in September. Motivated by this statistically significant result we first repeated their results using the logistic regression model with geomagnetic *aa* index. We use logarithmic ($\log_{10}$) scale which is found to provide slightly better results. We applied $\log_{10}$ scale to 3-hour *aa* values and calculated normalized daily values from which monthly averages were calculated. We show the results in Figure 2. The logistic regression fit was done using generalized linear regression model built in Matlab programming platform separately for those winters, where the September QBO is easterly (Fig. 2a) and westerly (Fig. 2b), and X in Eq. 1 is $\log_{10}(aa)_{\mathrm{Dec}}$. Due to the threshold of half QBO standard deviation, 62 winters correspond to a certain QBO in September, 31 winters in each phase. The estimated probability is shown by the thick green curves and the 95% confidence interval for the probability is indicated by the dashed green curves. The dashed vertical line indicates the median values for $\log_{10}(aa)_{\mathrm{Dec}}$.

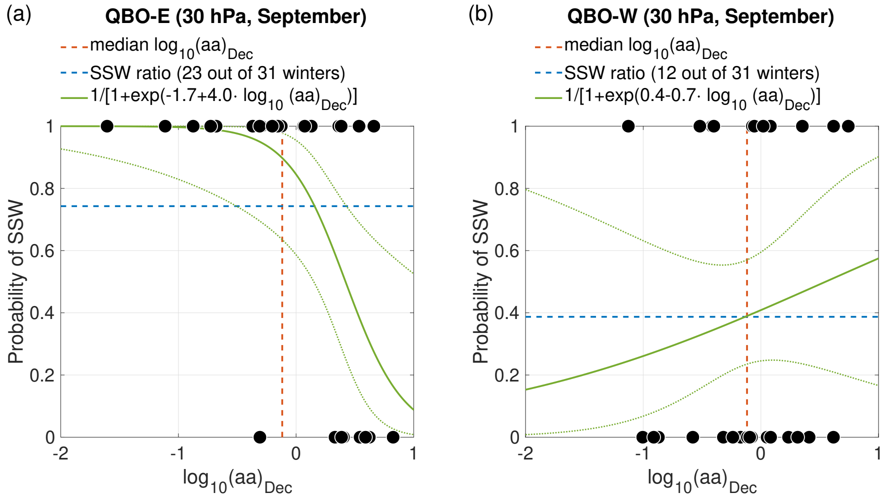

Figure 2. Logistic regression estimate of the SSW probability (thick green curve) using monthly mean $\log_{10}(aa)_{\mathrm{Dec}}$ index in December as explanatory variable. The black dots indicate the binary observations (value of 1 for winters with SSW and value of 0 for winters without SSW), dashed red vertical lines indicate median $\log_{10}(aa)_{\mathrm{Dec}}$ index values in 1951 – 2020, dashed blue horizontal lines are the average occurrence probability of SSWs in the respective QBO phase at 30 hPa in preceding September (a) for QBO-E and (b) for QBO-W (beyond the threshold of half QBO standard deviation). The dotted green curves indicate the 95% confidence interval of the estimated SSW probability.

One can see that best fit probability curves reveal a difference in the distribution of SSW winters between high and low $\log_{10}(aa)$ values. If QBO is easterly (westerly) the probability for an SSW event increases (decreases) with decreasing $\log_{10}(aa)_{\mathrm{Dec}}$ compared to the average occurrence probability of SSWs in the respective QBO phase (shown with horizontal dashed lines). The fitted intercept a is equal to 1.7±0.7 and rate parameter b is -4.0±1.6 in QBO easterly phase. For increasing geomagnetic activity in QBO easterly phase the SSW probability decreases. On the other hand, due to sparse observations, the uncertainty of the model increases significantly. In westerly QBO phase, the influence of aa on SSW probability is opposite to the QBO-E (easterly phase) phase, but much weaker. The corresponding probability does not significantly differ from the average SSW occurrence probability of about 0.4 in QBO-W (westerly phase) for any $\log_{10}(aa)_{\mathrm{Dec}}$ value. This is apparently reflected by the large uncertainty in the fitted parameters with intercept a being -0.4±0.4 and rate parameter b being 0.7±0.8. Although based on a slightly different list of SSW events, these results are in a good agreement with the recent results by Salminen et al. (2020).

2.5 Model performance measures

We use two measures to evaluate the performance of the probabilistic model. The first measure is the so-called Brier score defined as a mean square error of the probability forecast, i.e. the mean squared difference of the continuous-valued probability estimates and binary-valued outcomes (Brier, 1950). Lower values suggest better prediction. A poor model forecasts an event or no-event with probabilities close to 0.5. In this case, the Brier score would be close to 0.25, while in a precise model the Brier score should be lower than 0.25. As the second measure, we use the success ratio, which is defined as the fraction of correct predictions. Because the outcomes are binary valued (0 for winter without SSW and 1 for winter with SSW) the output of the logistic regression model needs to be converted into a binary outcome in order to evaluate the success ratio. This procedure effectively corresponds to a binary classifier, which requires a cutoff value for the probability to indicate an SSW (no SSW) if the predicted probability is larger (smaller) than the cutoff value. A natural choice for the cutoff is often 0.5, but in general the cutoff is an optimizable parameter of the model and can differ from 0.5, e.g., if the data set is significantly imbalanced (different sizes of the two classes).

3 Geomagnetic activity effect on SSW occurrence

As shown above in Fig. 2, the December aa index and September QBO phase at 30 hPa significantly affect the SSW probability. However, these particular choices, which correspond to those in Salminen et al. (2020) are not necessarily optimal for the logistic regression model. Therefore, in order to find the combination of optimal QBO and aa index which yields the strongest influence on the SSW occurrence probability, we fitted the logistic regression model by varying (1) the month and (2) the pressure level used to define the QBO phase, and (3) the timing of the 30-day interval used to calculate the average $\log_{10}(aa)$ index. The QBO month was varied from January in the previous winter to March of the present winter, the QBO pressure level from 10 hPa to 70 hPa, and the $\log_{10}(aa)$ index 30-day interval from September to February in the SSW winter. Running mean $\log_{10}(aa)$ values were calculated with 1-day time step with the same normalization as described in the previous section. In order to preserve a causal connection between the geomagnetic activity (energetic electron precipitation) and the SSWs, winters when SSW event occurs exclusively before the last day of the $\log_{10}(aa)$ time window are considered as winters without SSW. We note that before January, this approach does not significantly affect the analysis as only for one winter in 1982 the SSW event occurred exclusively in December.

The results of this calculation are shown in Figure 3 where the color indicates b values from Equation 1. Higher values signify a steeper change in the SSW probability with increasing $\log_{10} aa$. Negative b values are colored in blue and suggest that the SSW events are more probable during low geomagnetic activity, i.e. similar to Figure 2a. Red colors denote opposite effect, i.e. similar to Figure 2b. The p-value of the b parameter is based on the t-test. We confirmed the validity of the t-test by two types of Monte-Carlo simulations (however, due to computationally expensive calculations only for the optimal model discussed below in Section 5). In the first simulation we introduced random time shifts between QBO, aa, F10.7 and the binary probability time series, refit the model parameters and repeated this for 100000 iterations. In another simulation we used bootstrapping, i.e., resampled each of the time series randomly with replacement blocks of different size (testing from 1 to 10 years) 100000 times. These resampling approaches retain the autocorrelations of the time series but break their mutual relationships. When comparing the original parameter values to the Monte-Carlo distributions of model parameters the p-values of each parameter agreed well with those obtained from the t-test thereby validating the use of t-test.

Here, we show contours for 0.05 (thin dashed), 0.02 (thin solid) and 0.01 p-values (thick). The panels in the left column correspond to winters when QBO in the corresponding month (vertical axis) was in the easterly phase, and in the right column to winters with QBO in the westerly phase. The horizontal axes in Figure 3 indicate the last day of the 30-day interval for averaging $\log_{10} aa$.

Significant negative values of b parameter are clearly visible at all QBO pressure levels in the easterly phase beginning at 10 hPa from previous February and moving to later months with decreasing altitude (increasing pressure level). The downward movement corresponds to the slow downward propagation of the QBO wind shear zones. The strongest response of the SSW probability to $\log_{10}(aa)$ is obtained for winters preceded

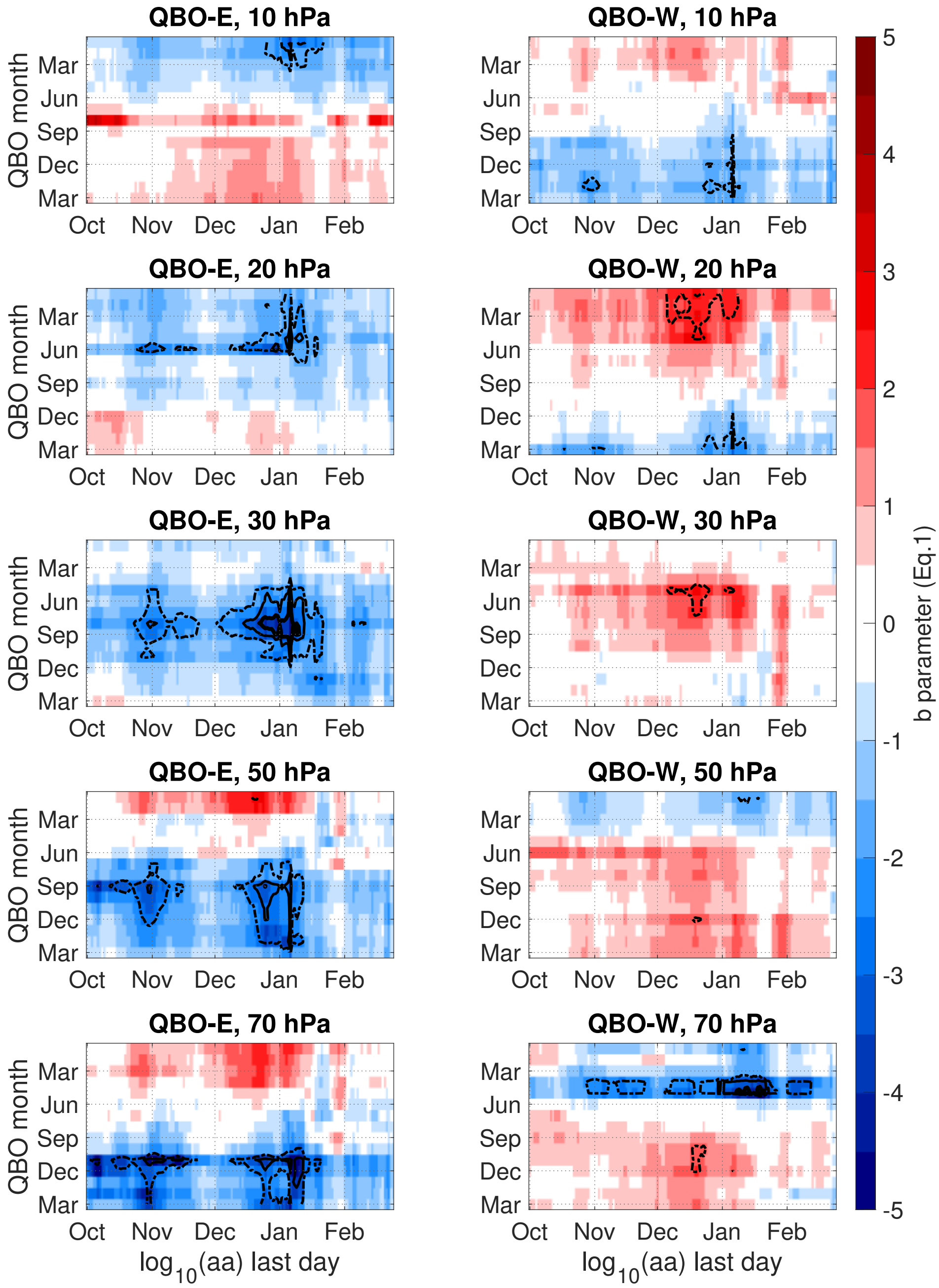

Figure 3. Fitted b parameter in Equation 1 as a function of QBO month (vertical axis) and time of the 30-day $\log_{10}(aa)$ window (horizontal axis). The time of the $\log_{10}(aa)$ window corresponds to the last day of the 30-day interval. The parameter values are calculated separately for the winters when the QBO phase at corresponding month/pressure level was easterly (left column) and westerly (right column). Contours denote statistical significance from a t-test: p=0.05 (thin dashed), p=0.02 (thin solid), and p=0.01 (thick).

by easterly QBO at 30 hPa evaluated in August and seems to begin in November and maximize around beginning of January with 30-day average $\log_{10}(aa)$ taken in December. Since most SSWs happen in January and February, the *aa* related response in Figure 3 mostly disappears after February, when many SSWs are cut out from the analysis. We also note that the *b* parameter has a curious dropout with lower p-values corresponding to the window with last day in December (i.e., November month). We found that this dropout is largely due to three winters; 1957 and 2004, which are SSW winters and 1959 which does not have a SSW. Winters of 1957 and especially 2004 are associated to very large geomagnetic activity (winter of 2003/2004 has the largest November $\log_{10}(aa)$ value in the entire dataset) and therefore they greatly oppose the tendency of most other data points in Figure 2 and lead to a decreased *b*-parameter. In contrast the October or December values of the $\log_{10}(aa)$ for these years are not equally large and do not deviate from the tendency of the other data points, which is why the *b*-parameter for the window with last day in November is again stronger and statistically significant. Note that the winter 2003/2004 has noted to be a strong outlier in other similar studies of EEP influence on the polar vortex (Maliniemi et al., 2013; Salminen et al., 2019). Year 1959, on the other hand, has a rather large *aa* value for the window ending in December, which is why it also greatly diminishes the *b*-parameter there. Removing these influential years from the fit would yield a continuous statistically significant *aa* response from Nov-Jan in Figure 3 (not shown).

In the westerly QBO (right column in Figure 3), weak but significant *b* values are seen mostly for the QBO at 20 and 30 hPa at the beginning of the previous winter and during May, respectively. This is the same result as shown by Figure 2b, where the logistic regression function indicates that less SSWs occur during low geomagnetic activity if the QBO is in the westerly phase. However, this signal is quite weak and appears more intermittently than the response in QBO easterly phase. Overall it therefore seems that the clearest influence of $\log_{10}(aa)$ appears in the QBO-E phase.

The influence of QBO easterly phase together with geomagnetic activity seems to be strongest at 30 hPa level when QBO phase is taken from August, i.e. about 3–4 months before the winter season. Since the QBO phase descends with time a similar, but not quite that strong, *aa*-related response is obtained by taking the QBO phase at 50 hPa in September-October or 70 hPa at November. As shown by Figure 3, the *aa*-related effect is clearly strongest and most significant in August at 30 hPa indicating that the 4–5 month time lag to subsequent winter season is relevant. Some of this significance may be due to the fact that QBO at 30 hPa is more often in the easterly phase (more data points) compared to the lower QBO levels: 30 winters are preceded by easterly QBO at 30 hPa in August, 25 winters are preceded by easterly QBO at 50 hPa in September, and 23 winters are preceded by easterly QBO at 70 hPa in November. This difference may result in higher significance of the fitted logistic regression using QBO phase at 30 hPa.

Often the QBO influence on polar stratosphere is understood via QBO's influence on planetary wave propagation. This is true, e.g., in the so-called Holton-Tan effect (Holton & Tan, 1980), where the easterly QBO in the mid to low stratosphere guides more planetary wave activity into the polar stratosphere thereby making the polar vortex more variable, weaker and even disrupted to the point of SSW formation. Secondly, easterly QBO phase results in the increased ascent rate in the tropical stratosphere, and thus stronger Brewer–Dobson circulation (BDC) (Flury et al., 2013). Therefore, more air including ozone (Salminen et al., 2019) reaches the polar lower stratosphere by winter. Increased adiabatic heating related to the BDC associated downwelling at high latitudes in QBO-E also contributes to weakening of the vortex. Recent studies have shown that the weaker vortex associated with planetary wave forcing favors wave-mean-flow interactions by which the energetic electron precipitation can affect the polar vortex dynamics (Salminen et al., 2019; Asikainen et al., 2020; Salminen et al., 2022). The exact reason why the QBO time lag from August to the winter season seems to be relevant for the EEP effect on the vortex warrants a more detailed study. However, it is likely that the planetary wave forcing and the BDC associated adiabatic heating take some time to build up a significantly weakened vortex by the beginning of the winter, which would favor the EEP effect.

4 Solar activity effect on SSW occurrence

Labitzke (1987), Labitzke and van Loon (1988) and Labitzke et al. (2006) showed that solar activity, specifically F10.7 radio flux index, together with the QBO seem to influence the state of the northern polar vortex. A positive (weaker negative) correlation between F10.7 and polar stratospheric temperature or geopotential height is observed during westerly (easterly) phase of the QBO.

To see if there is evidence for the solar activity influence on SSW probability in our setting we carry out a similar analysis as in the previous section but now using the standardized F10.7 index as an explanatory variable in Equation 1. Results for the easterly and westerly QBO phases at different altitudes are shown in Figure 4 in the same format as in Figure 3. Significant positive (red) values of b parameter are clearly seen mostly in the westerly QBO phase. Positive b indicates that the SSW occurrence probability increases with the F10.7 index. The response is seen with different QBO lags with respect to the winter: 7–11 months before December for the westerly QBO phase at 20 hPa, 4–7 months for the westerly QBO phase at 30 hPa, and no lag for the QBO at 50 hPa. These results are in accordance with the results by Labitzke and van Loon (1988) who showed that during westerly QBO winters the polar stratospheric temperature correlates with solar activity. The decrease of the optimal QBO time lag with decreasing altitude corresponds to the descend of the QBO wind signal in time similarly as in Figure 3. A similar influence of solar activity was also observed by Salminen et al. (2020), who showed that high solar activity together with QBO-W was associated with a higher probability of SSWs than low solar activity in the same QBO phase, although this difference was not statistically very significant. Here, however, the F10.7 effect is significant and does not strongly depend on the timing of the 30-day F10.7 window. This is because F10.7 varies rather slowly with the solar cycle and persists at similar levels over a year. Yet, the solar flux in May – June for the westerly QBO at 30 hPa and in October – November for the westerly QBO at 50 hPa yields a slightly stronger effect on the SSW occurrence as p-values are lower than 0.02 for these combinations (although the differences to other F10.7 timings are not large enough to be significant).

There is also a small but significant region of positive response in QBO-E especially at 50 hPa in June – July and 70 hPa around September – October. Despite a rather slow F10.7 variation on monthly time scale, significant signal is seen only using F10.7 taken around January. The effect emerges when the last day of the window crosses mid-January. Consequently, winters of 1953, 1955, 1985, 2006, and 2019 when SSWs exclusively occur before the end of January become considered as winters without SSWs, i.e. assigned zero P values in Equation 1. The strongest response is seen for winters preceded by QBO-E taken at 70 hPa in October and F10.7 in January. Although, this signal is consistent with earlier studies of the northern hemisphere polar temperature (Gray et al., 2004; Camp & Tung, 2007; Mitchell et al., 2015) it is of little use for predictive purposes.

The mechanism of the solar activity effect on the SSW occurrence is likely partly related to the modulation of the BDC. According to Flury et al. (2013), westerly QBO slows down the ascent rate of the tropical part of the BDC, which is associated with weakened downwelling in the Arctic, making the polar lower stratosphere cooler. Moreover, it has been shown that in westerly phase of the QBO the BDC strengthens (weakens) during high (low) solar activity (Labitzke et al., 2006; Matthes et al., 2010). Gray et al. (2004) found indications that zonal wind anomalies in the equatorial/subtropical upper stratosphere associated with the westerly QBO during high solar activity reinforce each

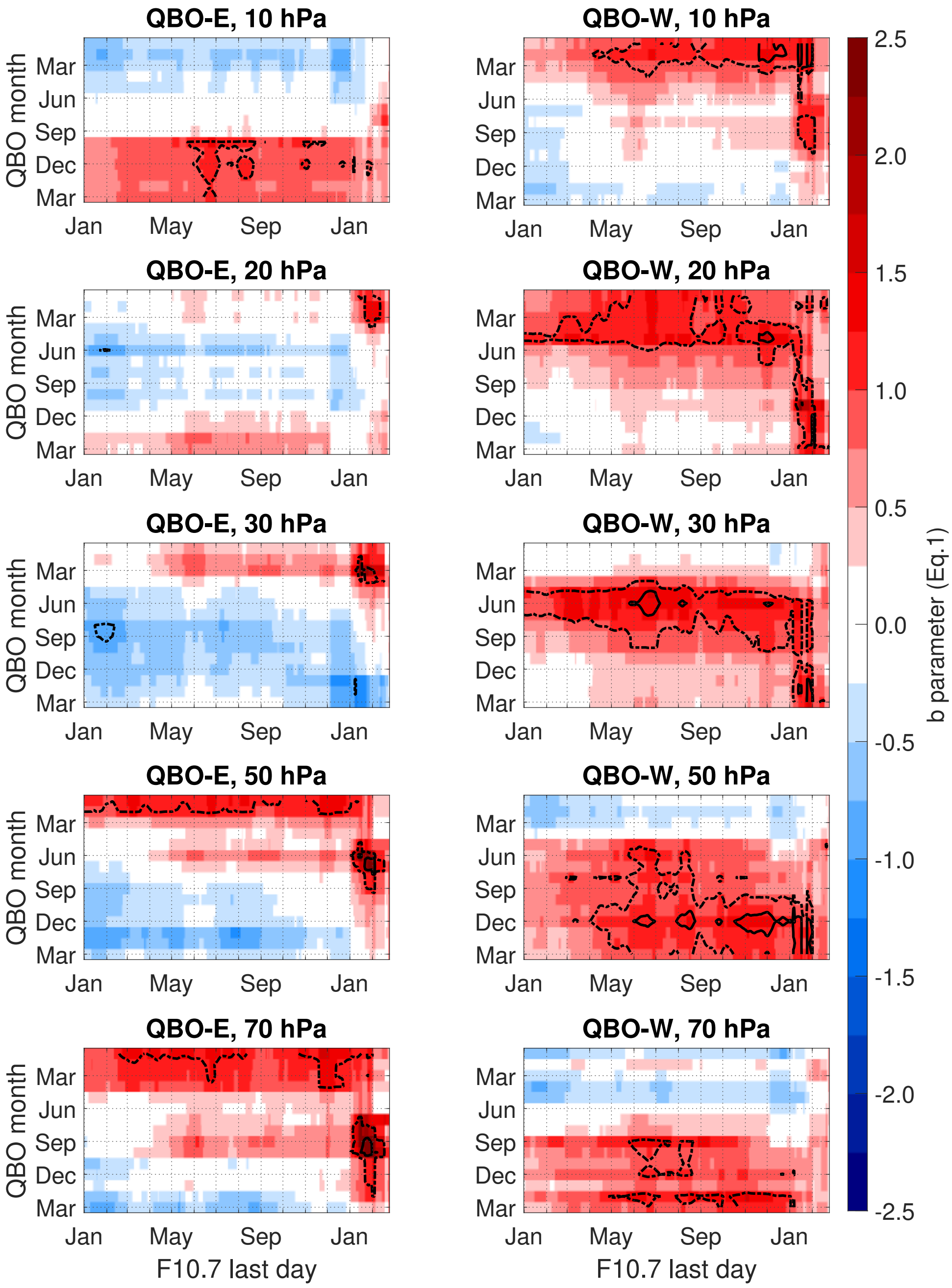

Figure 4. Fitted b parameter in Equation 1 as a function of QBO month (vertical axis) and time of the 30-day F10.7 window (horizontal axis). The time of the F10.7 window corresponds to the last day of the 30-day interval. The parameter values are calculated separately for the winters when the QBO phase at corresponding month/pressure level was easterly (left column) and westerly (right column). Contours denote statistical significance from a t-test: p=0.05 (thin dashed), p=0.02 (thin solid), and p=0.01 (thick).

other in a way which leads to the development of the Aleutian high in the winter stratosphere, which in turn enhances planetary wave formation and propagation into the polar stratosphere. Based on these past findings, in QBO-W the SSW probability is expected to be higher (lower) during high (low) solar activity.

Chiodo et al. (2014) and Kuchar et al. (2017) suggested that solar-related response in the tropical lower stratosphere potentially originates from the aliasing of the solar cycle with the major volcanic eruptions El Chichón in 1982 and Mt. Pinatubo in 1991. However, in this study, we consider only those years when QBO is greater than half of its standard deviation in the corresponding phase. Consequently, winters of 1983 and 1993 are excluded from the analysis, while winter of 1992 corresponds to the easterly QBO phase and, thus, also does not affect the results presented in Figure 4.

For the purposes of building a predictive model for SSWs, we will here use the response to solar F10.7 during westerly QBO at 30 hPa during summer as seen in Figure 4. This allows us to estimate the SSW probability in QBO-W several months before the winter season.

5 Optimal predictive models and their performance

Figures 3 and 4 show the effects of the geomagnetic and solar activity on SSW occurrence probability when the $\log_{10}(aa)$ and F10.7 indices are averaged over a 30-day window. In order to study what is the most optimal length of the time window affecting the SSW occurrence, we use the Brier score and the success ratio to evaluate the performance of the probabilistic model (see Section 2.5). In the following analysis, the performance measures as well as the cutoff probability are obtained using a leave-one-point-out cross validation technique. The idea of this method is to fit and evaluate the model using all the data except one year and then make a prediction for that year. The procedure is then repeated for all years. The model performance and optimal cutoff corresponding to the highest success ratio are then estimated using all the predicted values.

In Figure 5a-f, we varied the length (vertical axis) and timing (the horizontal axis indicates the last day of the window) of the window used to average the explanatory variable (either $\log_{10}(aa)$ or F10.7) in the logistic model and display the Brier score, optimal cutoff probability and the model success ratio as a function of these two factors. The plots of Figure 5a-c correspond to the model for QBO-E phase (QBO taken at 30 hPa from preceding August) with $\log_{10}(aa)$ and plots of Figure 5d-f correspond to QBO-W phase (QBO taken at 30 hPa from preceding June) with solar F10.7 index. For the QBO-E model one can see that the model performs optimally, when the $\log_{10}(aa)$ window ends in the first half of January. The success and Brier score are not very sensitive to the length of the averaging window and the optimal length is somewhere between 25 to 50 days (Figure 5). When the averaging window moves past late-January the model performance decreases considerably. This is because increasingly more SSWs will be dropped out from the fit due to the requirement that the SSW must happen after the last day of the averaging window. The optimal timing of the $\log_{10}(aa)$ window agrees well with the descend of EEP-related NO_y below 0.02 hPa pressure level estimated by Funke et al. (2014). They found that the NO_y amount in December effectively depends on the average geomagnetic activity (*ap* index) in October – December, while the NO_y amount in January and February depends more on *ap* index in December and January. Our results provide similar estimates showing that SSW occurrence is influenced by geomagnetic activity from mid-November to end of January. The optimal cutoff probability around the optimal $\log_{10}(aa)$ window is between 0.5 and 0.75 depending on the exact location of the averaging window. With such cutoffs the binary classification of the logistic regression results yields a rather high success ratio of 87-90%, which is clearly above the average SSW occurrence rate in QBO-E phase: 73% (22 out of 30 winters) for half standard deviation QBO phase threshold and 72% (26 out of 36 winters) with zero QBO phase threshold.

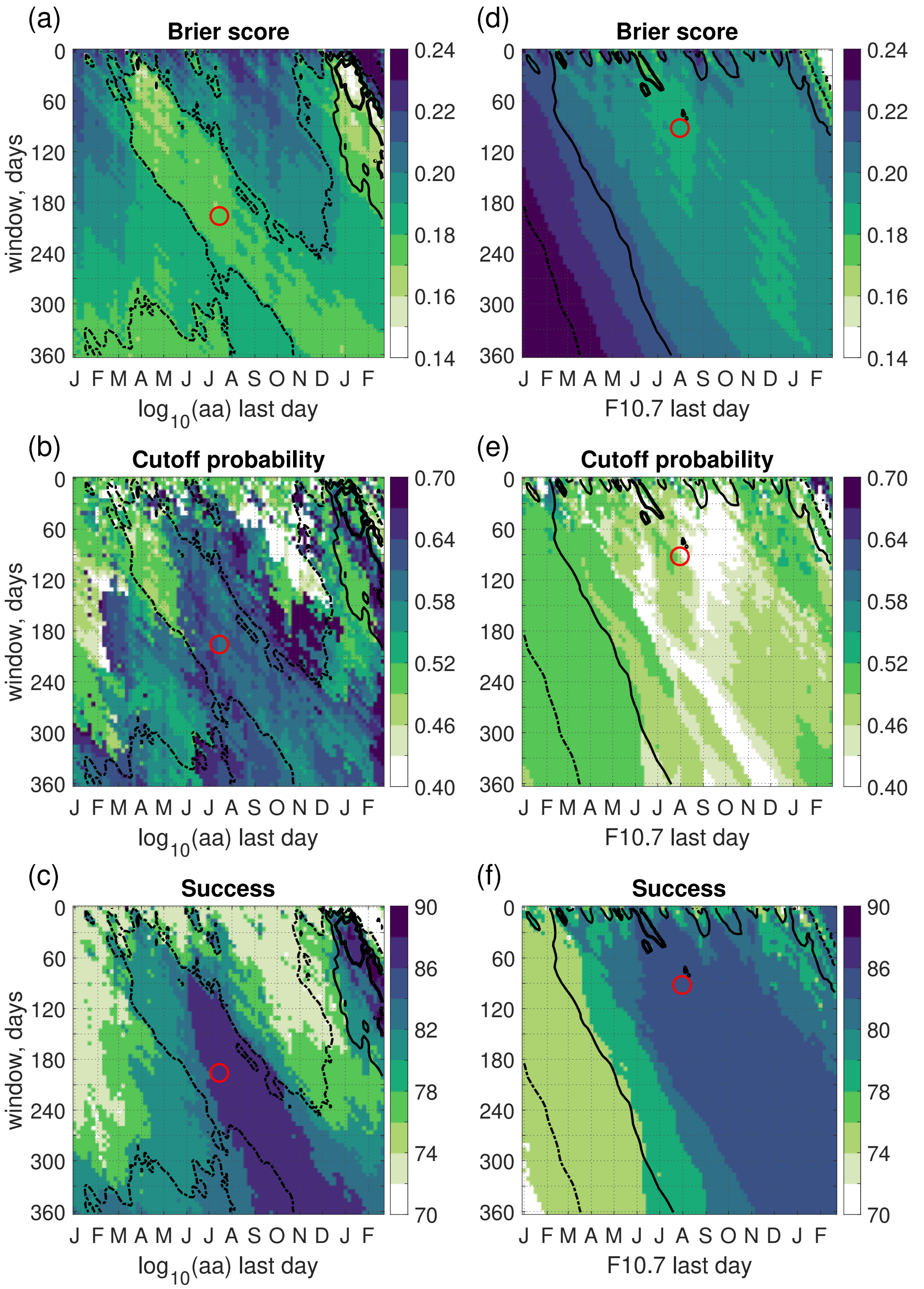

Figure 5. Performance of the SSW prediction models: (a, d) Brier score, (b, e) cutoff probability, and (c, f) success of the prediction. The results have been computed using different lengths and positions of the averaging window of the explaining variable in Equation 1. The models are calculated for the winters with QBO at 30 hPa being in the easterly phase (a, b, c) during August and $\log_{10}(aa)$ as an explaining variable; (d, e, f) for the winters with QBO at 30 hPa being in the westerly phase during June and F10.7 as an explaining variable. Contours denote statistical significance of p=0.05 (dashed), p=0.02 (thin), and p=0.01 (thick). Open red circles indicate optimal combination for the SSW probability model: $\log_{10}(aa)$ in 1 January – 15 July for the easterly QBO phase and 01 May – 31 July for the westerly.

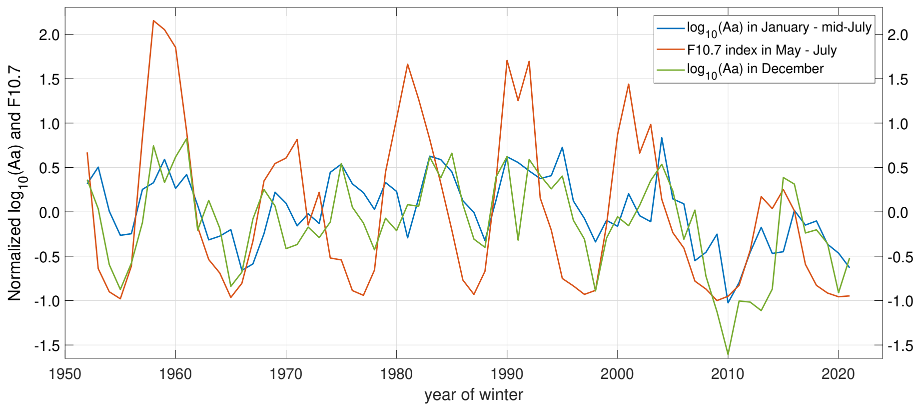

Figure 6. Normalized time series of $\log_{10}(aa)$ averaged in January — mid-July used for the SSW prediction in QBO-E winters (blue). Also shown $\log_{10}(aa)$ in December which was found to be the most significant in modulating the SSW occurrence (green). Red curve corresponds to the F10.7 index averaged over May — July used for the SSW prediction in QBO-W winters.

The overall optimal $\log_{10}(aa)$ window extends roughly from December to January and therefore provides only a rather short lead time, if any at all, for SSW prediction. However, Figures 5a-c indicate that there is another region of $\log_{10}(aa)$, which offers almost as good a model performance as the optimal window. The slanted light-green region of low Brier score in Figure 5a and dark region of high success ratio in Figure 5c extend from June (window length about 90 days) to November-December (window length about a year). For example, for windows extending some 120-180 days backward from mid-July show an average Brier score of about 0.17, which is not much worse than 0.15 of the optimal region. The optimal cutoff in these averaging windows is about 0.6 and with this cutoff the model yields a success ratio of 87%, which is practically as good as in the optimal region.

Taking the average $\log_{10}(aa)$ from the start of the year until mid-July provides the best possibility for predicting the SSW probability of the following winter with a genuine, rather long lead time of about 5 months. The reason why average $\log_{10}(aa)$ evaluated so long before the winter season works here for SSW prediction is possibly due to its strong correlation (cc$\approx$0.8, p$< 10^{-6}$) with the December average of $\log_{10}(aa)$ (blue and green curves in Figure 6, respectively. Figure 7a shows the logistic regression curve for the QBO-E model using the $\log_{10}(aa)$ from start of the year until mid-July along with the overall SSW occurrence rate in the respective QBO-E indicated by the horizontal dashed line. Clearly, for high $\log_{10}(aa)$ values, the model gives the SSW probability which is significantly lower than the overall SSW occurrence rate in QBO-E.

Figures 5d-f show the Brier score, cutoff probability, and model success ratio for the QBO-W phase model, with F10.7 index as the explaining factor. Because the variability of F10.7 index is dominated by slow solar cycle variation the model performance is practically independent of the timing and length of the F10.7 window as long as the window is taken before the winter season. The Brier score of the model is on average about 0.19. The optimal cutoff probability is about 0.4-0.5 and yields a success ratio of about 85%. For the final QBO-W model, we choose the average of F10.7 evaluated over May – July time period. The corresponding values of the solar flux are shown as red curve in Figure 6. The F10.7 evaluated 5 months before the winter period is possible as a predictor because of the strong autocorrelation of F10.7 over several months due to the slow solar cycle variation, which dominates F10.7.

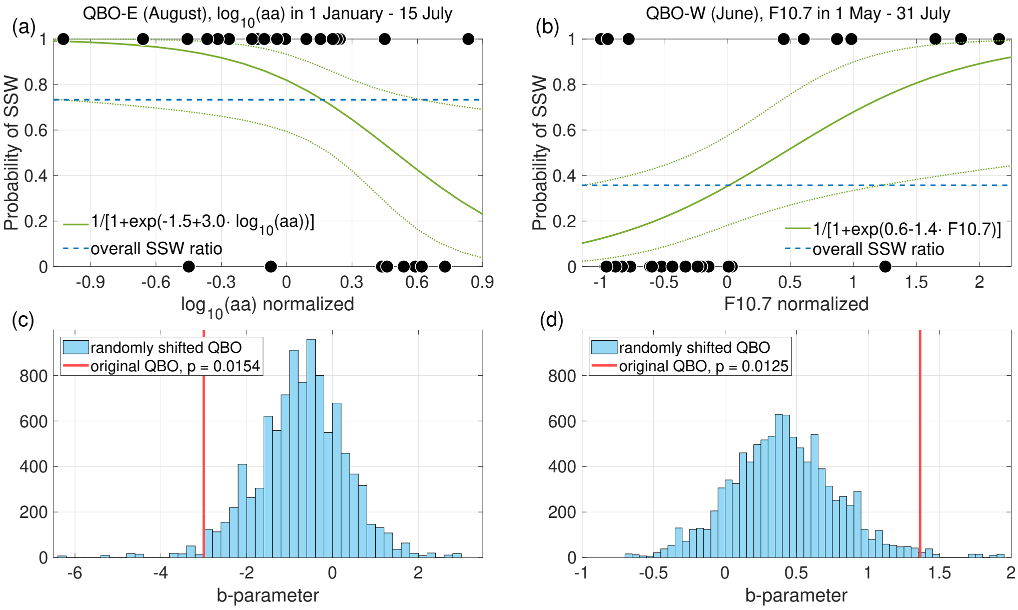

Figure 7. Logistic regression estimate of the SSW probability (thick green curve) using (a) mean $\log_{10}(aa)$ index from 1 January to 15 July, and (b) mean F10.7 index from 1 May to 31 July as explanatory variable; dashed blue horizontal lines are the average occurrence probability of SSWs in the respective QBO phase at 30 hPa (a) in preceding August for QBO-E and (b) in preceding June for QBO-W. The dotted green curves indicate the 95% confidence interval of the estimated SSW probability. Histograms of the b-parameter in QBO-E (c) and QBO-W (d) models (Equation 1) obtained from Monte-Carlo simulation with 10000 iterations that introduces a random time shift to the QBO time series; red vertical lines denote b-parameters found for the original QBO.

Figure 7b shows the logistic regression curve for the QBO-W model using the F10.7 from May to July. One can see that for small F10.7 values the estimated probabilities are quite close to the overall QBO-W SSW occurrence rate, but at high F10.7 values the model gives significantly larger values again indicating that the combination of QBO-W with high F10.7 favors the generation of SSWs.

Salby and Shea (1991) discussed a possibility that when stratifying the data according to the QBO phase, a solar-related signal can be seen because of frequencies higher than half of the QBO frequency can be aliased to low-frequencies. To test the probability of aliasing we performed a Monte Carlo simulation where we shifted the QBO time series randomly and recalculated model b-parameters. This was repeated 10000 times to get a histogram of expected b-parameters under the assumption that the QBO would not influence the b-parameter and that the observed b-parameter was a result of random chance and aliasing due to QBO stratification. Figures 7c-d show the distribution of the corresponding parameters. As indicated by vertical red lines, the fitted b-parameters are significant at 98.5% and 98.7% levels. These results indicate that it would be rather unlikely to obtain the observed b-parameters by random chance due to aliasing introduced by QBO stratification.

As a further check of robustness, the parameters of the optimal models (Figures 7a,b) were also estimated with the leave-k-out cross-validation. In each of 100000 trials k randomly selected points of the time-series are left out of the fitting and then used for the

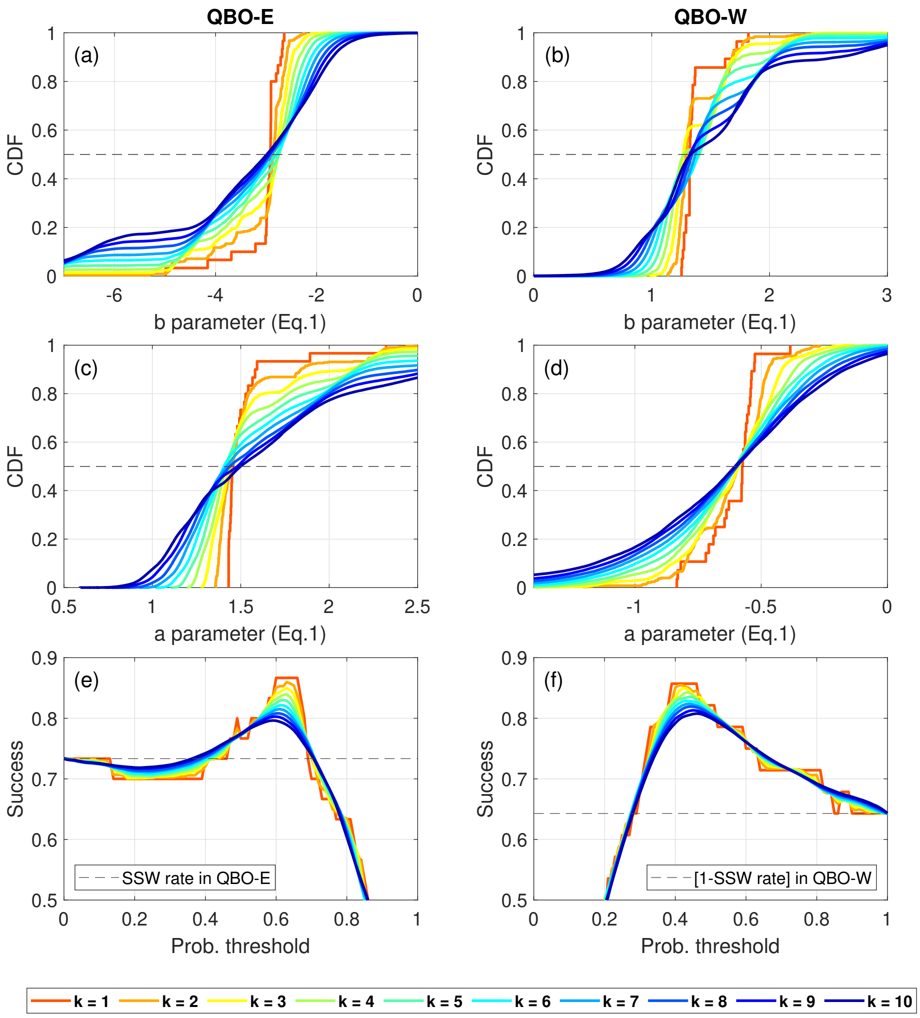

Figure 8. Leave-k-out cross-validation of the models for QBO-E (left column) and QBO-W (right column); k corresponds to the size of the hold out set used for the validation. (a–d) Model parameters; (e–f) success ratios of of the validating points over all trials depending on the probability threshold, while horizontal lines indicate a possible success ratio if no modulation by the geomagnetic or solar activities is considered.

validation of the model. Here, the results are obtained for k ranging from 1 to 10, so the size of the hold-out dataset ranges between 3 and 30% for the QBO-E years and 4 and 36% for the QBO-W years. Figures 8a–d show the cumulative distribution function of the model parameters for different k values. It is clear that leaving more points out results in increasing variance of the a and b parameters. However, for both QBO-E and QBO-W models, the median values do not depend on k which confirms the robustness of the model parameters and performance evaluation metrics based on leave-one-point-out cross validation. We also calculated the success ratio of the prediction for all validation winters which is dependent on the probability threshold for defining the SSW/no-SSW outcome. The performance of the QBO-E (QBO-W) models decreases from 87% (86%) to 80% (81%) with more points left out from the training dataset.

We used the half of a standard deviation threshold to determine the phase of the QBO at 30 hPa. To study the sensitivity of our results on the choice of the QBO threshold, we calculated the cross-validated model success for the QBO-E and QBO-W models using different QBO thresholds. The top row of Figure 9 shows for the QBO-E the model success, the optimal cutoff probability used in calculation of the success, and the number of winters remaining in the analysis as a function of the QBO month and the QBO threshold level. The bottom row of Figure 9 shows the same for the QBO-W model. One can again see the same optimal QBO months of August for QBO-E and June for QBO-W, which produce the best success. These correspond to the optimal months seen earlier in Figures 3 and 4. Overall, Figure 9 shows that the model performance (success and probability cutoff) is not very sensitive to the choice of the QBO threshold in either QBO phase. While the success tends to increase with the QBO threshold the number of retained data points decreases dramatically. The chosen threshold of half of standard deviation seems to produce a good trade-off between having an optimal success ratio but still retaining as many data points as possible. In such a restriction, we lose 3 years when winters are preceded by QBO-E in August, 1 winter preceded by QBO-W in June, and 3 winters preceded by both QBO-E in August and QBO-W in June. In the remaining 58 years the QBO phase is more clearly defined. Note that for another five years, 1964, 1983, 1988, 1993, and 2011, the QBO phase was neither westerly in June, nor easterly in August.

6 SSW forecast

In the previous section, we found that for the QBO-E phase (QBO 30 hPa in August) model the SSW probability can be predicted with average $\log_{10}(aa)$ over January – July and for QBO-W phase (QBO 30 hPa in June) with average F10.7 index over May – July. In both QBO phases, the success ratio of the model exceeds 85%. Using these explanatory variables, we then hindcasted SSW probability for each of those past winters from 1952 to 2021, where the QBO phase could clearly be defined with the criterion discussed in Section 2.2 (for 12 out of 70 years the QBO phase could not be determined with these criteria). The calculation was done using the separate models for QBO-E and QBO-W and the leave-one-point-out cross validation technique that was already used in Figure 5. For each winter, we also evaluated the uncertainty of the SSW probability from the logistic model. The results are presented in Figure 10, which shows the indicator for SSWs for each year with a black dot (value of 1 means SSW and value of 0 means no SSW). The colored background shading indicates the QBO phase (blue for QBO-E and red for QBO-W). The colored dots indicate the predicted SSW probabilities and the 95% confidence limits as well as the upper and lower quartiles of the predicted probabilities. The predicted probabilities were converted to binary outcomes using the optimal cut-off probabilities of 0.6 and 0.45 for QBO-E and QBO-W phases respectively. The green colored dots indicate values, where the binary outcome agrees with the real value (model success) and the red dots indicate values, where the binary outcome disagrees (model

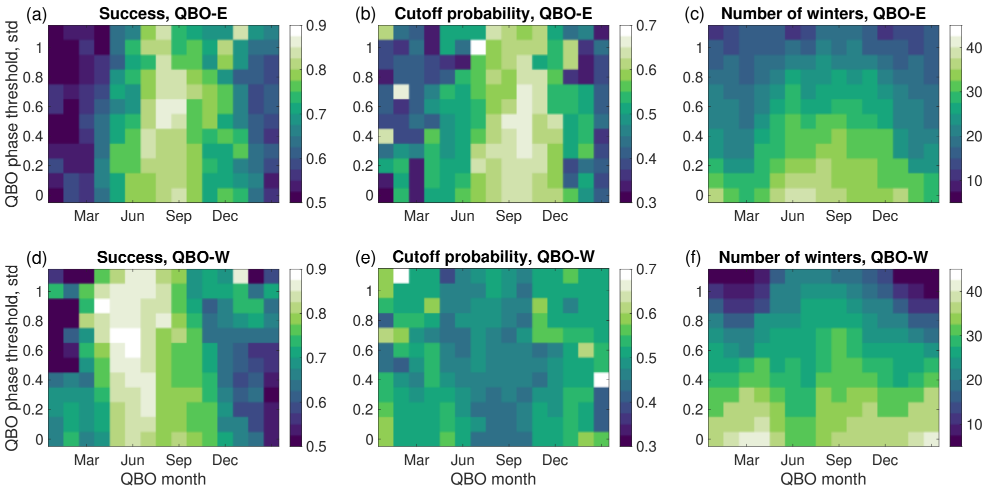

Figure 9. Models success as a function of the QBO month and the QBO threshold level (a, d); the cutoff probability used in calculation of the success (b, e), number of winters remaining in the analysis using different QBO threshold levels (c, f); top row for easterly QBO phase and $\log_{10}(aa)$ in 1 January – 15 July and bottom for westerly QBO and F10.7 in 1 May – 31 July.

failure). The two models together give 50 successful and 8 failed predictions yielding an overall success ratio of 86%.

The overall SSW occurrence rates in QBO-E and QBO-W winters are 73% and 36%, respectively. Therefore, the QBO phase alone could be used as a rough estimate for probability of SSWs so that all QBO-E winters would be predicted to have an SSW while all QBO-W winters would be predicted not to have an SSW. This approach would trivially give success ratios of 73% in QBO-E and 64% in QBO-W and therefore an overall success ratio of 69%. Comparing these numbers to the success ratios obtained by including $\log_{10}(aa)$ and F10.7 into the SSW prediction models shows that the information brought by these parameters greatly improves the accuracy of the SSW/no-SSW prediction. For the QBO-E model the inclusion of $\log_{10}(aa)$ raises the success ratio to 87%, i.e. almost a 20 units of percent of relative increase in success. For the QBO-W phase the inclusion of F10.7 raises the success ratio to 86%, i.e. corresponding to a roughly 34 units of percent of relative increase in success. Overall, the relative increase in the success ratio is about 23 units of percent. Figure 10 indicates with a large black open circle those predicted probabilities, which differ from the prediction based only on the overall SSW occurrence rate in the respective QBO phase. One can see that in QBO-E phase there are 6 winters where the inclusion of $\log_{10}(aa)$ changed the prediction to correct, while only 2 winters where the $\log_{10}(aa)$ information changed the model outcome to incorrect. For the QBO-W phase model there are 7 points where the F10.7 information changed the model outcome to correct and only one point where the model outcome changed to incorrect.

In our QBO-E and QBO-W models above we chose to include only the dominant (either aa or F10.7) effect in order to keep the model structure simpler and reduce the possibility for overfitting. As an additional check to justify this approach we also fitted the prediction model by including both aa and F10.7 as explanatory variables in both QBO phases using the QBO lags and averaging windows with the best response to aa and F10.7 according to the discussion above. We found that for the QBO-E model F10.7 does not have a significant effect and the b-parameter for aa is not influenced by the in-

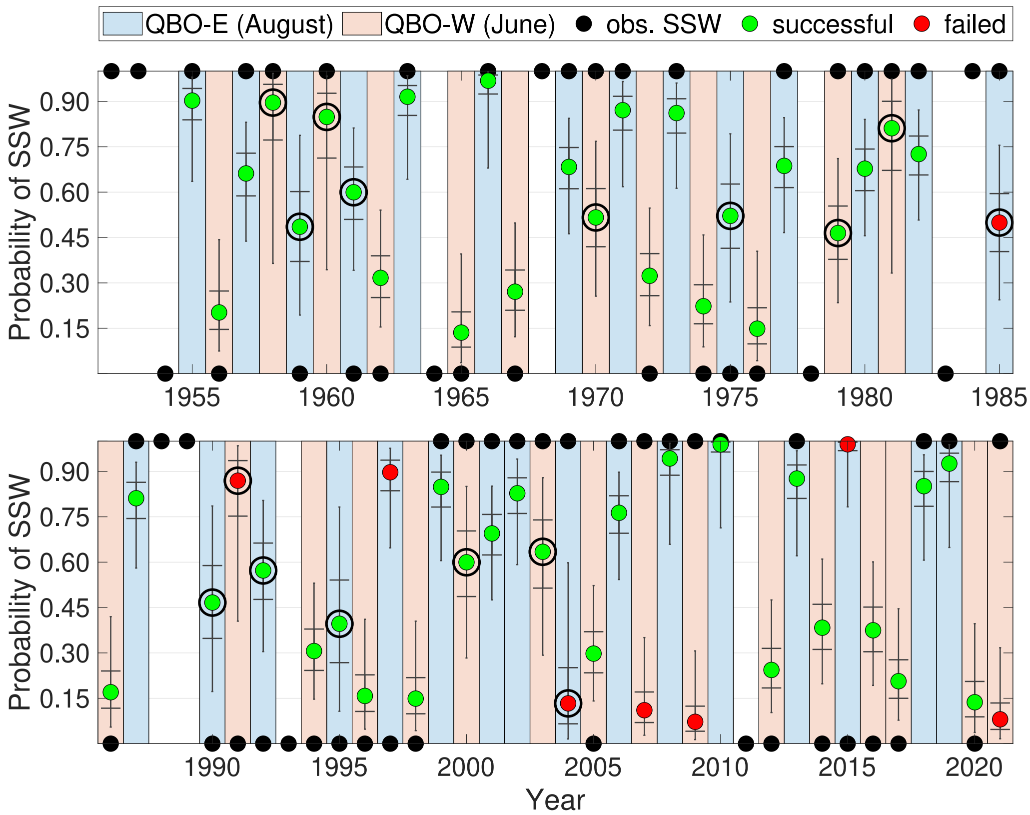

Figure 10. SSW forecast according to the model (Figure 5c) for the winters preceded by the easterly QBO in August (blue background) and model (Figure 5f) for the winters preceded by the westerly QBO in June (purple background). Colored circles indicate median prediction for the corresponding year, green — successful, red — failed relative to cutoff probability thresholds 0.6 for QBO-E and 0.45 for QBO-W. Caps indicate first and third quartiles and vertical lines 95% confidence interval. Black filled circles with SSW probability equal to one (zero) indicate winter with (without) SSW according to the ERA5 reanalysis (Table 1). Large black circles indicate winters when our model and QBO-based forecast disagree.

clusion of F10.7 into the model. The model performance does not significantly improve either.

In the case of QBO-W model inclusion of aa brings a small improvement to the model (e.g., about 8% decrease in Brier score). However, there is some indication that the b-parameter of $\log_{10}(aa)$ changes depending on whether or not F10.7 is included into the model (these changes are not statistically significant though). This tendency might be due to the weak correlation (cc=0.4) between $\log_{10}(aa)$ and F10.7. Despite the the small improvement of the QBO-W model performance by the inclusion of aa it is better to avoid overfitting and potential collinearity (even if estimated to be small) in favor of a simpler and more robust model.

7 Discussion and conclusions

The results presented here show that our SSW prediction models are able to predict fairly successfully the SSW occurrence of the winter season about 4-6 months in advance using information on QBO phase as well as geomagnetic and solar activity. However, one source of uncertainty in the results is the fact that the model performance may actually depend on the criteria used to define the SSW events. Here, we used the procedure by Charlton and Polvani (2007), which is the most common and well recognized method used for the consistency between SSW statistical studies. However, this definition can sometimes miss some events that in slightly different definitions could be classified as major warmings. Another fact, which contributes to the sensitivity of SSW identifications is the uncertainty of the reanalysis products, which are driven by the numerical models assimilating incomplete data with inherent uncertainties and measurement precision. It has been estimated that the uncertainty in the upper-air wind at 10 hPa in ERA5 is close to 3 m/s (Bell et al., 2021). Therefore, the requirement for the wind reversal with a strict 0 m/s threshold may also affect SSW identification in some cases and thereby results of our study. However, apart from the earliest years, we tried to mitigate this problem by verifying the identified SSWs using several reanalyses (see Table 1). Although a complete recalculation and model optimization for all the other re-analyses is out of the scope of this paper, it is worthwhile to roughly estimate how the above results would change if we used the other re-analysis datasets included in Table 1 instead of ERA5. Compared to ERA5 the joint ERA-Interim/ERA40 set has only two winters with different SSW identification (1995 and 2017) and overall has four years (with clearly identified QBO) less than ERA5. These differences in the SSW identifications would likely not lead to significant difference in the model parameters or performance. The NCEP/NCAR re-analysis covers the ERA5 time period, but especially in the 1950s its quality is lower than ERA5's. Apart from the early 1950s, where NCEP/NCAR does not really observe much SSWs, there are only 4 differing SSW years in ERA5 and NCEP/NCAR (in QBO-W 1965 and 1981 and in QBO-E 1959, while the differing year 1968 is not included in the analysis due to QBO being too close to zero). Here 1981 is a solar maximum year (ERA5 has SSW and NCEP/NCAR does not) and 1965 a solar minimum year (NCEP/NCAR has SSW but ERA5 not). Such a change in the binary SSW probability of two points would not significantly influence the model parameters from their ERA5 values as the fit is dominated by a large number of other years (see Figure 7b), where the ERA5 and NCEP/NCAR agree. Year 1959 (NCEP/NCAR has SSW but ERA5 does not) is a QBO-E year with aa value close to its cycle maximum. Figure 7a shows that exchanging this one no-SSW point in the right hand side of the aa-axis to an SSW would not significantly affect the regression fit, because of the dominance of the other no-SSW years at high aa values. Even though these small changes in the number of SSW/no-SSW years from ERA5 and NCEP/NCAR would slightly decrease the number of successful predictions in NCEP/NCAR it would not lead into any significant change in the success ratio if we consider only the slightly shorter and more reliable portion of the NCEP/NCAR dataset after 1959.

In this study, we developed a probabilistic model to estimate the probability for the occurrence of an SSW event in the upcoming winter. We used here the logistic regression method to model the dependence of SSW probability on QBO phase and geomagnetic activity characterized by the *aa* index and solar activity characterized by the F10.7 index. Given the relatively small amount of data we carefully estimated the optimal parameters and performance of the model using cross-validation methods.

We showed that when the QBO phase in preceding August at 30 hPa is easterly the SSW occurrence depends on geomagnetic activity expressed by the $\log_{10}(aa)$ index, which is a proxy for energetic electron precipitation (EEP) into the upper atmosphere. The strongest influence was observed for the geomagnetic activity evaluated in the beginning of the winter from early December to early January. This agrees well with the established influence of energetic electron precipitation on the polar vortex (Salminen et al., 2019). When the geomagnetic activity and level of particle precipitation is lower than average, less ozone is destroyed by the catalytic reactions with EEP-created NO_x. In mid-winter this results in cooler mesosphere and upper stratosphere due to increased infrared radiative cooling (Sinnhuber et al., 2018) and by the following dynamical impact to warmer lower stratosphere (Salminen et al., 2019) and weaker than average polar vortex with more frequent SSWs compared to the average occurrence rate under the easterly QBO. The EEP influence on the polar vortex is known to preferentially occur, when the planetary wave activity to the vortex is suitably enhanced, e.g., during easterly QBO phase (Asikainen et al., 2020; Salminen et al., 2022). On the other hand, the geomagnetic activity influence on SSW probability can also be interpreted so that in QBO-E larger than average geomagnetic activity (EEP) strengthens the polar vortex and makes it less prone to SSWs.

The early winter time window for geomagnetic activity does not allow much long-term predictive capability. However, we found here that due to the autocorrelation of *aa* index we can also use the average $\log_{10}(aa)$ evaluated from the start of the year until mid-July to produce almost an equally successful prediction model for the SSW probability of the subsequent winter season. This model allows us to issue an SSW prediction in August and yields a cross-validated success ratio of about 87%.

We also confirmed here the earlier observations by Labitzke and van Loon (1988); Labitzke et al. (2006); Gray et al. (2010), which indicate that the solar activity modulates the Holton-Tan effect for the westerly QBO phase and consequently leads to a decrease (increase) of the SSW occurrence when solar activity is low (high). We found that this influence could be seen not only for winter solar F10.7 flux and winter westerly QBO at 50 hPa, but also using F10.7 solar flux and QBO phase at 30 hPa during preceding summer. This allowed us to model the SSW probability with May – July average of F10.7 index in QBO-W phase evaluated at 30 hPa pressure level in June. The cross-validated success ratio of this model was about 86%. Some past climate simulation studies, e.g., Matthes et al. (2013) have confirmed these solar UV influences in QBO-W phase in long, over 100-year, simulation runs. However, some others, e.g., Kren et al. (2014), have implied a that the combined solar UV/QBO influence on the polar stratosphere might not be robust feature, but rather observed due to random chance in short climate records of about 40 years. While we cannot completely rule out the possibility of the randomness of the F10.7 response in SSW occurrence frequency during QBO-W we note that given the rather long 69-year observational record used here (nearly twice the length of the 40-year period which Kren et al. (2014) deemed potentially problematic) the results appear statistically significant and the probability of chance occurrence of the results is rather low (Figure 7).

Together the *aa* and F10.7 indices with the QBO phase allow for a rather good prediction of the probability of SSWs in the upcoming winter to be issued already in the preceding August. The overall success ratio of the combined models is about 86%, which is clearly higher than a rough prediction based only on the phase of the QBO (i.e. the

Holton-Tan effect), which yields a success ratio of about 69% (accounting for the both QBO phases).

The current numerical weather forecasting models can successfully predict the occurrence of SSWs about two weeks in advance (Baldwin et al., 2021). While our probabilistic model can only evaluate the probability for a SSW to occur at some point of a winter, it offers a considerably longer lead time of about 4–5 months. Since SSWs are known to have significant impacts on ground weather for several weeks in large regions over the Northern Hemisphere the long lead time prediction offers improved capabilities to mitigate the effects of SSWs on different areas of society dependent on winter time weather conditions, e.g., energy consumption and production. The results obtained here could possibly be further improved, e.g., by including other climate factors known to influence the polar vortex and SSW formation, e.g., El Niño Southern Oscillation, volcanic activity and previous states of the NAO/NAM circulation modes. However, as a first approach to long-term probabilistic prediction of SSWs based on solar related drivers the results obtained here are quite promising.

Acknowledgments

ERA5 reanalysis data set was downloaded from the Copernicus Climate Change Service (C3S) Climate Data Store. ECMWF (`https://www.ecmwf.int/`) is acknowledged for providing ERA40 and ERA-Interim reanalysis products used in this article. NCEP/NCAR reanalysis data set is provided by the National Oceanic and Atmospheric Administration (NOAA)/Earth System Research Laboratory(ESRL) Physical Sciences Division (PSD) from their website at `https://www.esrl.noaa.gov/psd`. The *aa* geomagnetic activity index was downloaded from the International Service of Geomagnetic Indices website at `https://isgi.unistra.fr/`. We are also grateful to the Dominion Radio Astrophysical Observatory (DRAO) in Penticton (Canada) and to NorthWest Research Associates for providing the F10.7 flux data and to OMNIWeb data service (`https://omniweb.gsfc.nasa.gov/`) for access to these data.

We acknowledge the financial support by the Academy of Finland to the PROSPECT (project no. 321440), and by the University of Oulu and The Academy of Finland PROFI4 (Grant 318930). Antti Salminen acknowledges the funding from the Finnish Cultural Foundation (grant no. 00200976).

References

Anstey, J. A., & Shepherd, T. G. (2014). High-latitude influence of the quasi-biennial oscillation. *Quarterly Journal of the Royal Meteorological Society*, *140*(678), 1-21. doi: 10.1002/qj.2132

Arsenovic, P., Rozanov, E., Stenke, A., Funke, B., Wissing, J. M., Mursula, K., … Peter, T. (2016). The influence of Middle Range Energy Electrons on atmospheric chemistry and regional climate. *Journal of Atmospheric and Solar-Terrestrial Physics*, *149*, 180-190. doi: 10.1016/j.jastp.2016.04.008

Asikainen, T., Salminen, A., Maliniemi, V., & Mursula, K. (2020). Influence of Enhanced Planetary Wave Activity on the Polar Vortex Enhancement Related to Energetic Electron Precipitation. *Journal of Geophysical Research (Atmospheres)*, *125*(9), e32137. doi: 10.1029/2019JD032137

Balachandran, N. K., & Rind, D. (1995). Modeling the Effects of UV Variability and the QBO on the Troposphere-Stratosphere System. Part I:. The Middle Atmosphere. *Journal of Climate*, *8*(8), 2058-2079. doi: 10.1175/1520-0442(1995)008⟨2058:MTEOUV⟩2.0.CO;2

Baldwin, M. P., Ayarzagüena, B., Birner, T., Butchart, N., Butler, A. H., Charlton-Perez, A. J., … Pedatella, N. M. (2021). Sudden Stratospheric Warmings.

Reviews of Geophysics, *59*(1), e00708. doi: 10.1029/2020RG000708
Baldwin, M. P., & Dunkerton, T. J. (1999). Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *Journal of Geophysical Research*, *104*(D24), 30,937-30,946. doi: 10.1029/1999JD900445
Baldwin, M. P., & Dunkerton, T. J. (2001). Stratospheric Harbingers of Anomalous Weather Regimes. *Science*, *294*(5542), 581-584. doi: 10.1126/science.1063315
Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., ... Takahashi, M. (2001). The quasi-biennial oscillation. *Reviews of Geophysics*, *39*(2), 179-229. doi: 10.1029/1999RG000073
Baumgaertner, A. J. G., Seppälä, A., Jöckel, P., & Clilverd, M. A. (2011). Geomagnetic activity related NO_x enhancements and polar surface air temperature variability in a chemistry climate model: modulation of the NAM index. *Atmospheric Chemistry & Physics*, *11*(9), 4521-4531. doi: 10.5194/acp-11-4521-2011
Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., ... Thépaut, J.-N. (2021). The ERA5 global reanalysis: Preliminary extension to 1950. *Quarterly Journal of the Royal Meteorological Society*, *147*(741), 4186-4227. doi: 10.1002/qj.4174
Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, *78*(1), 1-3. doi: 10.1175/1520-0493(1950)078⟨0001:VOFEIT⟩2.0.CO;2
Butchart, N. (2014). The Brewer-Dobson circulation. *Rev. Geophys.*, *52*(2), 157-184. doi: 10.1002/2013RG000448
Butler, A. H., & Gerber, E. P. (2018). Optimizing the Definition of a Sudden Stratospheric Warming. *Journal of Climate*, *31*(6), 2337-2344. doi: 10.1175/JCLI-D-17-0648.1
Butler, A. H., & Polvani, L. M. (2011). El Niño, La Niña, and stratospheric sudden warmings: A reevaluation in light of the observational record. *Geophysical Research Letters*, *38*(13), L13807. doi: 10.1029/2011GL048084
Butler, A. H., Seidel, D. J., Hardiman, S. C., Butchart, N., Birner, T., & Match, A. (2015). Defining Sudden Stratospheric Warmings. *Bulletin of the American Meteorological Society*, *96*(11), 1913-1928. doi: 10.1175/BAMS-D-13-00173.1
Butler, A. H., Sjoberg, J. P., Seidel, D. J., & Rosenlof, K. H. (2017). A sudden stratospheric warming compendium. *Earth System Science Data*, *9*(1), 63-76. doi: 10.5194/essd-9-63-2017
Camp, C. D., & Tung, K.-K. (2007). The Influence of the Solar Cycle and QBO on the Late-Winter Stratospheric Polar Vortex. *Journal of Atmospheric Sciences*, *64*(4), 1267. doi: 10.1175/JAS3883.1
Charlton, A. J., & Polvani, L. M. (2007). A New Look at Stratospheric Sudden Warmings. Part I: Climatology and Modeling Benchmarks. *Journal of Climate*, *20*(3), 449. doi: 10.1175/JCLI3996.1
Charlton-Perez, A. J., Ferranti, L., & Lee, R. W. (2018). The influence of the stratospheric state on North Atlantic weather regimes. *Quarterly Journal of the Royal Meteorological Society*, *144*(713), 1140-1151. doi: 10.1002/qj.3280
Charney, J. G., & Drazin, P. G. (1961). Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *Journal of Geophysical Research*, *66*(1), 83-109. doi: 10.1029/JZ066i001p00083
Chiodo, G., Marsh, D. R., Garcia-Herrera, R., Calvo, N., & García, J. A. (2014, June). On the detection of the solar signal in the tropical stratosphere. *Atmospheric Chemistry & Physics*, *14*(11), 5251-5269. doi: 10.5194/acp-14-5251-2014
Cohen, J., Barlow, M., Kushner, P. J., & Saito, K. (2007). Stratosphere Troposphere Coupling and Links with Eurasian Land Surface Variability. *Journal of Climate*, *20*(21), 5335. doi: 10.1175/2007JCLI1725.1
Crutzen, P. J., Isaksen, I. S. A., & Reid, G. C. (1975). Solar Proton Events: Strato-

spheric Sources of Nitric Oxide. *Science*, *189*(4201), 457-459. doi: 10.1126/science.189.4201.457

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., … Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553-597. doi: 10.1002/qj.828

de la Cámara, A., Albers, J. R., Birner, T., Garcia, R. R., Hitchcock, P., Kinnison, D. E., & Smith, A. K. (2017). Sensitivity of sudden stratospheric warmings to previous stratospheric conditions. *J. Atmos. Sci.*, *74*(9), 2857–2877. doi: 10.1175/JAS-D-17-0136.1

Domeisen, D. I. V. (2019). Estimating the Frequency of Sudden Stratospheric Warming Events From Surface Observations of the North Atlantic Oscillation. *Journal of Geophysical Research (Atmospheres)*, *124*(6), 3180-3194. doi: 10.1029/2018JD030077

Domeisen, D. I. V., Butler, A. H., Charlton-Perez, A. J., Ayarzagüena, B., Baldwin, M. P., Dunn-Sigouin, E., … Taguchi, M. (2020). The Role of the Stratosphere in Subseasonal to Seasonal Prediction: 1. Predictability of the Stratosphere. *Journal of Geophysical Research (Atmospheres)*, *125*(2), e30920. doi: 10.1029/2019JD030920

Domeisen, D. I. V., Garfinkel, C. I., & Butler, A. H. (2019). The Teleconnection of El Niño Southern Oscillation to the Stratosphere. *Reviews of Geophysics*, *57*(1), 5-47. doi: 10.1029/2018RG000596

Dunkerton, T., Hsu, C. P. F., & McIntyre, M. E. (1981). Some Eulerian and Lagrangian Diagnostics for a Model Stratospheric Warming. *Journal of Atmospheric Sciences*, *38*(4), 819-844. doi: 10.1175/1520-0469(1981)038⟨0819:SEALDF⟩2.0.CO;2

Floyd, L., Newmark, J., Cook, J., Herring, L., & McMullin, D. (2005, January). Solar EUV and UV spectral irradiances and solar indices. *Journal of Atmospheric and Solar-Terrestrial Physics*, *67*(1-2), 3-15. doi: 10.1016/j.jastp.2004.07.013

Floyd, L. E., Cook, J. W., Herring, L. C., & Crane, P. C. (2003). SUSIM'S 11-year observational record of the solar UV irradiance. *Advances in Space Research*, *31*(9), 2111-2120. doi: 10.1016/S0273-1177(03)00148-0

Flury, T., Wu, D. L., & Read, W. G. (2013). Variability in the speed of the Brewer-Dobson circulation as observed by Aura/MLS. *Atmospheric Chemistry & Physics*, *13*(9), 4563-4575. doi: 10.5194/acp-13-4563-2013

Frame, T. H. A., & Gray, L. J. (2010). The 11-Yr Solar Cycle in ERA-40 Data: An Update to 2008. *Journal of Climate*, *23*(8), 2213-2222. doi: 10.1175/2009JCLI3150.1

Fröhlich, C. (2006). Solar Irradiance Variability Since 1978. Revision of the PMOD Composite during Solar Cycle 21. *Space Science Reviews*, *125*(1-4), 53-65. doi: 10.1007/s11214-006-9046-5

Funke, B., López-Puertas, M., Holt, L., Randall, C. E., Stiller, G. P., & von Clarmann, T. (2014). Hemispheric distributions and interannual variability of NO_y produced by energetic particle precipitation in 2002-2012. *Journal of Geophysical Research (Atmospheres)*, *119*(23), 13,565-13,582. doi: 10.1002/2014JD022423

Garfinkel, C. I., Butler, A. H., Waugh, D. W., Hurwitz, M. M., & Polvani, L. M. (2012). Why might stratospheric sudden warmings occur with similar frequency in El Niño and La Niña winters? *Journal of Geophysical Research (Atmospheres)*, *117*(D19), D19106. doi: 10.1029/2012JD017777

Garfinkel, C. I., & Schwartz, C. (2017). MJO-Related Tropical Convection Anomalies Lead to More Accurate Stratospheric Vortex Variability in Subseasonal Forecast Models. *Geophysical Research Letters*, *44*(19), 10,054-10,062. doi: 10.1002/2017GL074470

Garfinkel, C. I., Schwartz, C., Domeisen, D. I. V., Son, S.-W., Butler, A. H., & White, I. P. (2018). Extratropical Atmospheric Predictability From the Quasi-Biennial Oscillation in Subseasonal Forecast Models. *Journal of Geophysical Research (Atmospheres)*, *123*(15), 7855-7866. doi: 10.1029/2018JD028724

Garfinkel, C. I., Shaw, T. A., Hartmann, D. L., & Waugh, D. W. (2012). Does the Holton-Tan Mechanism Explain How the Quasi-Biennial Oscillation Modulates the Arctic Polar Vortex? *Journal of Atmospheric Sciences*, *69*(5), 1713-1733. doi: 10.1175/JAS-D-11-0209.1

Garfinkel, C. I., Silverman, V., Harnik, N., Haspel, C., & Riz, Y. (2015, August). Stratospheric response to intraseasonal changes in incoming solar radiation. *Journal of Geophysical Research (Atmospheres)*, *120*(15), 7648-7660. doi: 10.1002/2015JD023244

Garfinkel, C. I., White, I., Gerber, E. P., Jucker, M., & Erez, M. (2020, July). The Building Blocks of Northern Hemisphere Wintertime Stationary Waves. *Journal of Climate*, *33*(13), 5611-5633. doi: 10.1175/JCLI-D-19-0181.1

Gray, L. J., Beer, J., Geller, M., Haigh, J. D., Lockwood, M., Matthes, K., ... White, W. (2010). Solar Influences on Climate. *Reviews of Geophysics*, *48*(4), RG4001. doi: 10.1029/2009RG000282

Gray, L. J., Crooks, S., Pascoe, C., Sparrow, S., & Palmer, M. (2004). Solar and QBO Influences on the Timing of Stratospheric Sudden Warmings. *Journal of Atmospheric Sciences*, *61*(23), 2777-2796. doi: 10.1175/JAS-3297.1

Henderson, G. R., Peings, Y., Furtado, J. C., & Kushner, P. J. (2018). Snow-atmosphere coupling in the Northern Hemisphere. *Nature Climate Change*, *8*(11), 954-963. doi: 10.1038/s41558-018-0295-6

Hendon, H. H., Thompson, D. W. J., Lim, E.-P., Butler, A. H., Newman, P. A., Coy, L., ... Nakamura, H. (2019). Rare forecasted climate event under way in the Southern Hemisphere. *Nature*, *573*(7775), 495-495. doi: 10.1038/d41586-019-02858-0

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049. doi: 10.1002/qj.3803

Holton, J. R., & Tan, H.-C. (1980). The Influence of the Equatorial Quasi-Biennial Oscillation on the Global Circulation at 50 mb. *Journal of Atmospheric Sciences*, *37*(10), 2200-2208. doi: 10.1175/1520-0469(1980)037⟨2200:TIOTEQ⟩2.0.CO;2

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., ... Joseph, D. (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, *77*(3), 437-472. doi: 10.1175/1520-0477(1996)077⟨0437:TNYRP⟩2.0.CO;2

Karpechko, A. Y. (2018). Predictability of Sudden Stratospheric Warmings in the ECMWF Extended-Range Forecast System. *Monthly Weather Review*, *146*(4), 1063-1075. doi: 10.1175/MWR-D-17-0317.1

King, A. D., Butler, A. H., Jucker, M., Earl, N. O., & Rudeva, I. (2019). Observed Relationships Between Sudden Stratospheric Warmings and European Climate Extremes. *Journal of Geophysical Research (Atmospheres)*, *124*(24), 13,943-13,961. doi: 10.1029/2019JD030480

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., ... Fiorino, M. (2001). The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. *Bulletin of the American Meteorological Society*, *82*(2), 247-268. doi: 10.1175/1520-0477(2001)082⟨0247:TNNYRM⟩2.3.CO;2

Kodera, K., & Kuroda, Y. (2002). Dynamical response to the solar cycle. *Journal of Geophysical Research (Atmospheres)*, *107*(D24), 4749. doi: 10.1029/2002JD002224

Kren, A. C., Marsh, D. R., Smith, A. K., & Pilewskie, P. (2014, May). Examining the stratospheric response to the solar cycle in a coupled WACCM simulation with an internally generated QBO. *Atmospheric Chemistry & Physics*, *14*(10), 4843-4856. doi: 10.5194/acp-14-4843-2014

Krüger, K., Naujokat, B., & Labitzke, K. (2005). The Unusual Midwinter Warming in the Southern Hemisphere Stratosphere 2002: A Comparison to Northern Hemisphere Phenomena. *Journal of Atmospheric Sciences*, *62*(3), 603-613. doi: 10.1175/JAS-3316.1

Kuchar, A., Ball, W. T., Rozanov, E. V., Stenke, A., Revell, L., Miksovsky, J., ... Peter, T. (2017, September). On the aliasing of the solar cycle in the lower stratospheric tropical temperature. *Journal of Geophysical Research (Atmospheres)*, *122*(17), 9076-9093. doi: 10.1002/2017JD026948

Labitzke, K. (1982). On the Interannual Variability of the Middle Stratosphere during the Northern Winters. *Journal of the Meteorological Society of Japan*, *60*(1), 124-139. doi: 10.2151/jmsj1965.60.1_124

Labitzke, K. (1987). Sunspots, the QBO, and the stratospheric temperature in the north polar region. *Geophysical Research Letters*, *14*(5), 535-537. doi: 10.1029/GL014i005p00535

Labitzke, K., Kunze, M., & Brönnimann, S. (2006). Sunspots, the QBO and the stratosphere in the North Polar Region - 20 years later. *Meteorologische Zeitschrift*, *15*(3), 355-363. doi: 10.1127/0941-2948/2006/0136

Labitzke, K., & van Loon, H. (1988). Associations between the 11-year solar cycle, the QBO (quasi-biennial-oscillation) and the atmosphere. Part I: the troposphere and stratosphere in the northern hemisphere in winter. *Journal of Atmospheric and Terrestrial Physics*, *50*, 197-206. doi: 10.1016/0021-9169(88)90068-2

Lu, H., Clilverd, M. A., Seppälä, A., & Hood, L. L. (2008). Geomagnetic perturbations on stratospheric circulation in late winter and spring. *Journal of Geophysical Research (Atmospheres)*, *113*(D16), D16106. doi: 10.1029/2007JD008915

Maliniemi, V., Asikainen, T., & Mursula, K. (2016). Effect of geomagnetic activity on the northern annular mode: QBO dependence and the Holton-Tan relationship. *Journal of Geophysical Research (Atmospheres)*, *121*(17), 10,043-10,055. doi: 10.1002/2015JD024460

Maliniemi, V., Asikainen, T., Mursula, K., & Seppälä, A. (2013). QBO-dependent relation between electron precipitation and wintertime surface temperature. *Journal of Geophysical Research (Atmospheres)*, *118*(12), 6302-6310. doi: 10.1002/jgrd.50518

Matsuno, T. (1970). Vertical propagation of stationary planetary waves in the winter northern hemisphere. *Journal of the Atmospheric Sciences*, *27*(6), 871-883. doi: 10.1175/1520-0469(1970)027⟨0871:VPOSPW⟩2.0.CO;2

Matsuno, T. (1971). A Dynamical Model of the Stratospheric Sudden Warming. *Journal of Atmospheric Sciences*, *28*(8), 1479-1494. doi: 10.1175/1520-0469(1971)028⟨1479:ADMOTS⟩2.0.CO;2

Matthes, K., Kodera, K., Garcia, R. R., Kuroda, Y., Marsh, D. R., & Labitzke, K. (2013, May). The importance of time-varying forcing for QBO modulation of the atmospheric 11 year solar cycle signal. *Journal of Geophysical Research (Atmospheres)*, *118*(10), 4435-4447. doi: 10.1002/jgrd.50424

Matthes, K., Marsh, D. R., Garcia, R. R., Kinnison, D. E., Sassi, F., & Walters, S. (2010). Role of the QBO in modulating the influence of the 11 year solar cycle on the atmosphere using constant forcings. *Journal of Geophysical Research (Atmospheres)*, *115*(D18), D18110. doi: 10.1029/2009JD013020

Mitchell, D. M., Gray, L. J., Fujiwara, M., Hibino, T., Anstey, J. A., Ebisuzaki, W., ... Tan, D. (2015, July). Signatures of naturally induced variability in the atmosphere using multiple reanalysis datasets. *Quarterly Journal of the Royal Meteorological Society*, *141*(691), 2011-2031. doi: 10.1002/qj.2492

Palamara, D., & Bryant, E. (2004). Geomagnetic activity forcing of the Northern

Annular Mode via the stratosphere. *Annales Geophysicae*, *22*(3), 725-731. doi: 10.5194/angeo-22-725-2004
Palmeiro, F. M., Barriopedro, D., García-Herrera, R., & Calvo, N. (2015). Comparing Sudden Stratospheric Warming Definitions in Reanalysis Data*. *Journal of Climate*, *28*(17), 6823-6840. doi: 10.1175/JCLI-D-15-0004.1
Polvani, L. M., Sun, L., Butler, A. H., Richter, J. H., & Deser, C. (2017). Distinguishing Stratospheric Sudden Warmings from ENSO as Key Drivers of Wintertime Climate Variability over the North Atlantic and Eurasia. *Journal of Climate*, *30*(6), 1959-1969. doi: 10.1175/JCLI-D-16-0277.1
Polvani, L. M., & Waugh, D. W. (2004). Upward Wave Activity Flux as a Precursor to Extreme Stratospheric Events and Subsequent Anomalous Surface Weather Regimes. *Journal of Climate*, *17*(18), 3548-3554. doi: 10.1175/1520-0442(2004)017⟨3548:UWAFAA⟩2.0.CO;2
Randall, C. E., Harvey, V. L., Singleton, C. S., Bailey, S. M., Bernath, P. F., Codrescu, M., ... Russell, J. M. (2007). Energetic particle precipitation effects on the Southern Hemisphere stratosphere in 1992-2005. *Journal of Geophysical Research (Atmospheres)*, *112*(D8), D08308. doi: 10.1029/2006JD007696
Rao, J., Garfinkel, C. I., & White, I. P. (2020). Impact of the Quasi-Biennial Oscillation on the Northern Winter Stratospheric Polar Vortex in CMIP5/6 Models. *Journal of Climate*, *33*(11), 4787-4813. doi: 10.1175/JCLI-D-19-0663.1
Rao, J., Garfinkel, C. I., White, I. P., & Schwartz, C. (2020). The Southern Hemisphere Minor Sudden Stratospheric Warming in September 2019 and its Predictions in S2S Models. *Journal of Geophysical Research (Atmospheres)*, *125*(14), e32723. doi: 10.1029/2020JD032723
Rozanov, E., Callis, L., Schlesinger, M., Yang, F., Andronova, N., & Zubov, V. (2005). Atmospheric response to NO_y source due to energetic electron precipitation. *Geophysical Research Letters*, *32*(14), L14811. doi: 10.1029/2005GL023041
Salby, M., & Callaghan, P. (2004). Evidence of the Solar Cycle in the General Circulation of the Stratosphere. *Journal of Climate*, *17*(1), 34-46. doi: 10.1175/1520-0442(2004)017⟨0034:EOTSCI⟩2.0.CO;2
Salby, M. L., & Callaghan, P. F. (2002). Interannual Changes of the Stratospheric Circulation: Relationship to Ozone and Tropospheric Structure. *J. Clim.*, *15*(24), 3673-3685. doi: 10.1175/1520-0442(2003)015⟨3673:ICOTSC⟩2.0.CO;2
Salby, M. L., & Shea, D. J. (1991, December). Correlations between solar activity and the atmosphere: An unphysical explanation. *Journal of Geophysical Research*, *96*(D12), 22,579-22,595. doi: 10.1029/91JD02530
Salminen, A., Asikainen, T., Maliniemi, V., & Mursula, K. (2019). Effect of Energetic Electron Precipitation on the Northern Polar Vortex: Explaining the QBO Modulation via Control of Meridional Circulation. *Journal of Geophysical Research (Atmospheres)*, *124*(11), 5807-5821. doi: 10.1029/2018JD029296
Salminen, A., Asikainen, T., Maliniemi, V., & Mursula, K. (2020). Dependence of Sudden Stratospheric Warmings on Internal and External Drivers. *Geophysical Research Letters*, *47*(5), e86444. doi: 10.1029/2019GL086444
Salminen, A., Asikainen, T., & Mursula, K. (2022). Planetary Waves Controlling the Effect of Energetic Electron Precipitation on the Northern Polar Vortex. *Geophysical Research Letters*, *49*(6), e97076. doi: 10.1029/2021GL097076
Scherhag, R. (1952). Die explosionsartige Stratosphärenerwärmung des Spätwinters 1951/52. *Berichte des Deutschen Wetterdienstes in der US-Zone*, *6*(38), 51-63.
Schwartz, C., & Garfinkel, C. I. (2017). Relative roles of the MJO and stratospheric variability in North Atlantic and European winter climate. *Journal of Geophysical Research (Atmospheres)*, *122*(8), 4184-4201. doi: 10.1002/2016JD025829
Scott, R. K., & Polvani, L. M. (2004). Stratospheric control of upward wave flux

near the tropopause. *Geophys. Res. Lett.*, *31*(2). doi: 10.1029/2003GL017965

Seppälä, A., Lu, H., Clilverd, M. A., & Rodger, C. J. (2013). Geomagnetic activity signatures in wintertime stratosphere wind, temperature, and wave response. *Journal of Geophysical Research (Atmospheres)*, *118*(5), 2169-2183. doi: 10.1002/jgrd.50236

Silverman, V., Harnik, N., Matthes, K., Lubis, S. W., & Wahl, S. (2018). Radiative effects of ozone waves on the northern hemisphere polar vortex and its modulation by the qbo. *Atmospheric Chemistry and Physics*, *18*(9), 6637–6659. doi: 10.5194/acp-18-6637-2018

Sinnhuber, M., Berger, U., Funke, B., Nieder, H., Reddmann, T., Stiller, G., ... Maik Wissing, J. (2018). NO_y production, ozone loss and changes in net radiative heating due to energetic particle precipitation in 2002-2010. *Atmospheric Chemistry & Physics*, *18*(2), 1115-1147. doi: 10.5194/acp-18-1115-2018

Soukharev, B. E., & Hood, L. L. (2006). Solar cycle variation of stratospheric ozone: Multiple regression analysis of long-term satellite data sets and comparisons with models. *Journal of Geophysical Research (Atmospheres)*, *111*(D20), D20314. doi: 10.1029/2006JD007107

Tripathi, O. P., Baldwin, M., Charlton-Perez, A., Charron, M., Eckermann, S. D., Gerber, E., ... Son, S.-W. (2015). The predictability of the extratropical stratosphere on monthly time-scales and its impact on the skill of tropospheric forecasts. *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 987-1003. doi: 10.1002/qj.2432

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., ... Woollen, J. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, *131*(612), 2961-3012. doi: 10.1256/qj.04.176

van Loon, H., Jenne, R. L., & Labitzke, K. (1973). Zonal harmonic standing waves. *Journal of Geophysical Research*, *78*(21), 4463-4471. doi: 10.1029/JC078i021p04463

van Loon, H., & Labitzke, K. (1987). The Southern Oscillation. Part V: The Anomalies in the Lower Stratosphere of the Northern Hemisphere in Winter and a Comparison with the Quasi-Biennial Oscillation. *Monthly Weather Review*, *115*(2), 357. doi: 10.1175/1520-0493(1987)115⟨0357:TSOPVT⟩2.0.CO;2

Ward, W., Seppälä, A., Yiğit, E., Nakamura, T., Stolle, C., Laštovička, J., ... Pallamraju, D. (2021, December). Role Of the Sun and the Middle atmosphere/thermosphere/ionosphere In Climate (ROSMIC): a retrospective and prospective view. *Progress in Earth and Planetary Science*, *8*(1), 47. doi: 10.1186/s40645-021-00433-8

Watson, P. A. G., & Gray, L. J. (2014). How does the quasi-biennial oscillation affect the stratospheric polar vortex? *Journal of the Atmospheric Sciences*, *71*(1), 391-409. doi: 10.1175/JAS-D-13-096.1

White, I., Garfinkel, C. I., Gerber, E. P., Jucker, M., Aquila, V., & Oman, L. D. (2019). The Downward Influence of Sudden Stratospheric Warmings: Association with Tropospheric Precursors. *Journal of Climate*, *32*(1), 85-108. doi: 10.1175/JCLI-D-18-0053.1

White, I. P., Garfinkel, C. I., Gerber, E. P., Jucker, M., Hitchcock, P., & Rao, J. (2020). The Generic Nature of the Tropospheric Response to Sudden Stratospheric Warmings. *Journal of Climate*, *33*(13), 5589-5610. doi: 10.1175/JCLI-D-19-0697.1

White, I. P., Lu, H., Mitchell, N. J., & Phillips, T. (2015). Dynamical response to the qbo in the northern winter stratosphere: Signatures in wave forcing and eddy fluxes of potential vorticity. *Journal of the Atmospheric Sciences*, *72*(12), 4487 - 4507. doi: 10.1175/JAS-D-14-0358.1