Which Stratospheric Sudden Warming Events are Most Predictable?

Dvir Chwat[1], Chaim I. Garfinkel[1], Wen Chen[2,3], Jian Rao[4]

[1]Fredy and Nadine Herrmann Institute of Earth Sciences, Hebrew University, Jerusalem, Israel
[2]College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, China
[3]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
[4]Key Laboratory of Meteorological Disaster, Ministry of Education (KLME), Joint International Research Laboratory of Climate and Environment Change (ILCEC), Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science and Technology, Nanjing 210044, China

Key Points:

- Predictability of 10 SSWs is compared among 10 S2S models.
- Four factors distinguish SSWs with above average predictability: (1) enhanced convection in the West Pacific associated with the Madden-Julian Oscillation;
- (2) the Quasi-Biennial Oscillation phase with easterlies near 50hPa; (3) a strong SSW; and (4) a strong pulse of wave activity in the week before the event.

Corresponding author: Chaim I. Garfinkel, `chaim.garfinkel@mail.huji.ac.il`

Abstract

The predictability of stratospheric sudden warming (SSW) events are considered in 10 subseasonal to seasonal (S2S) forecast models for 10 SSWs over the period 1999-2009. The 10 SSWs are divided into those with above-average predictability (in one case exceeding 20 days), below-average predictability, and average predictability. The four factors that most succinctly distinguish the composite with above average predictability are an active Madden-Julian Oscillation with enhanced convection in the West Pacific, the Quasi-Biennial Oscillation phase with easterlies in the lower stratosphere, a strong SSW, and a strong pulse of wave activity in the week before the event. Other factors, such as El Nino, stratospheric preconditioning, and the morphology (split vs. displacement) are comparatively less important.

Plain Language Summary

The wintertime stratosphere typically features circumpolar strong westerly winds, but on occasion these strong winds can reverse and temperatures over the pole can rise by tens of degrees. Such an abrupt warming phenomenon, occurred 10 times over the period covered by all models that have contributed re-forecasts to the subseasonal to seasonal (S2S) model archive, and this study considers the factors that distinguish which of these SSWs was most predictable. The most important precursor in the troposphere is the Madden-Julian Oscillation, and it is also shown that the strongest SSWs tend to be more predictable.

1 Introduction

During a major stratospheric sudden warming (SSW), stratospheric westerly winds in the circumpolar region reverse to easterly winds and temperatures rise over the pole by tens of degrees (Schoeberl, 1978; Charlton & Polvani, 2007; Butler et al., 2015; Baldwin et al., 2021). SSWs are typically followed by anomalous cold air outbreaks and precipitation over the midlatitude Northern Hemisphere continents (Thompson et al., 2002; Kolstad et al., 2010; Sigmond et al., 2013; Lehtonen & Karpechko, 2016; C. I. Garfinkel et al., 2017; Kretschmer et al., 2018; Karpechko et al., 2018). As the characteristic time scale of a major stratospheric sudden warming and its surface impact extends for several months, accurately predicting SSWs would open a window of opportunity for more reliable probabilistic predictability of surface weather anomalies on subseasonal time scales (Baldwin et al., 2003; Sigmond et al., 2013; Tripathi et al., 2015).

The factors governing the predictability of SSW events are only partially known. Previous work has found that predictability can range from several days to near a month depending on the model used and the specific SSW focused on (Tripathi et al., 2016; Taguchi, 2014; Noguchi et al., 2016; Karpechko, 2018; Rao, Garfinkel, et al., 2019; Rao et al., 2020). This wide spread may reflect differences in the predictability of different events, in the skill of different forecast systems, and on the method used to quantify successful prediction. For example, raising the model-lid has been shown to lead to an improved predictability of SSW (Marshall & Scaife, 2010), and the high-top S2S models examined by Rao, Garfinkel, et al. (2019), D. I. V. Domeisen et al. (2020), and Rao et al. (2020) typically performed better at capturing SSWs. Some studies have suggested that split SSWs are more difficult to forecast than displacement SSWs (Taguchi, 2016a, 2018; D. I. V. Domeisen et al., 2020; Taguchi, 2020), though because of the limited sample size the statistical significance of this effect is relatively weak. Accurately capturing the anomalous wave flux in both the troposphere and lower stratosphere that usually precedes SSWs has also been pinpointed as important for SSW predictability (Mukougawa et al., 2005; Taguchi, 2016b; Tripathi et al., 2016; Taguchi, 2018; Karpechko et al., 2018). Relatedly, SSWs that were preceded by the phase of the Madden Julian Oscillation with enhanced convection in the west Pacific were more predictable in the two models considered by C. Garfinkel and Schwartz (2017).

The Subseasonal-to-Seasonal (S2S) Prediction project (Vitart et al., 2017) has recently made available a large number of hindcasts covering the past several decades. These simulations are all initialized with observed sea surface temperatures and the atmospheric state, and as they are used operationally, they can be compared directly to observed variability during the duration of their forecast. There are two previous studies which contrasted multi-model predictability of different specific SSWs in the S2S database. Taguchi (2018) considered 4 NH SSWs in the hindcasts of 9 models, while Taguchi (2020) considered 10-11 NH SSWs in the hindcasts of 4 models, and noticed that the predictability of the SSW varies with event types (vortex split or displacement), the model considered, and the ability to represent the anomalous heat flux. Here we revisit the S2S database and considering ten models and ten different SSWs, we attempt to answer the following question: what distinguishes SSWs that were well-forecasted from those that were poorly forecasted?

We demonstrate that the deterministic predictability for SSWs varies from less than ten days to almost twenty days, depending on the SSW in question. This spread in predictability is associated with a range of factors, including two which appear to have been seldom demonstrated before: the strength of the SSW being considered (Rao, Ren, et al., 2019, the lone exception), and the proximity of the anomalous heat flux pulse to the event.

Table 1: S2S Model experiments chosen

model (ensemble members)	years	vertical levels	model top
CMA: BCC-CPS-S2Sv1 (4)	1999-2003	40	0.5hPa
CMA: BCC-CPS-S2Sv2 (4)	2004-2009	56	0.1hPa
NCEP (4)	1999-2009	64	0.02hPa
ECMWF2016 (11)	1999-2009	91	0.01hPa
ECMWF2019 (11)	1999-2009	91	0.01hPa
BoM (33)	1999-2009	17	10hPa
UKMO2015 (3)	1999-2009	85	85km
UKMO2019 (7)	1999-2009	85	85km
KMA (3)	1999-2009	85	85km
Météo France: CNRM-CM 6.1 (10)	1999-2009	91	0.01hPa
CNR-ISAC (5)	1999-2009	54	6.8 hPa
ECCC: GEPS6 (4)	1999-2009	45	0.1 hPa
JMA: GEPS1701 (5)	1999-2009	128	0.01 hPa

Table 1. For the UKMO, we downloaded hindcasts for the operational model in use during 2015 and the winter of 2019/2020, and for the ECMWF, we downloaded data for the model version in use during 2016 and the winter of 2019/2020 (CY41R1/CY41R2 and CY46R1).

2 Data and Methods

We focus on the ten modeling centers that have contributed to the S2S Prediction project (Vitart et al., 2017) with output at 10hPa - the Australian Bureau of Meteorology (BoM), the European Centre for Medium-Range Weather Forecasts (ECMWF), the China Meteorological Administration (CMA), the United Kingdom Met. Office (UKMO), the National Center for Environmental Prediction (NCEP), the Korean Meteorological Agency (KMA), the Japan Meteorological Agency (JMA), the Institute of Atmospheric Sciences and Climate of the National Research Council of Italy (ISAC-CNR), Environment and Climate Change Canada (ECCC) and Meteo France (CNRM). Table 1 summarizes the reforecasts analyzed in this study. We use the high-top version of CMA starting in 2004 when its hindcasts are first available, and the low-top version earlier when the high-top version is unavailable. These various models differ in the quality of their representation of the stratosphere: the stratosphere is less well resolved in BoM and ISAC-CNR as compared to the other models (Table 1). Note that we use the hi-top version of ECCC, and download hindcasts issued both in 2020 and in 2021 to increase temporal resolution. For the UKMO, we downloaded hindcasts for the operational model in use during 2015 and the winter of 2019/2020, and for the ECMWF, we downloaded data for the model version in use during 2016 and the winter of 2019/2020 (CY41R1/CY41R2 and CY46R1).

Table 2: SSWs in common period

	date	typical prediction skill
1	02-26-1999	well
2	03-20-2000	poor
3	02-11-2001	
4	12-30-2001	poor
5	01-18-2003	poor
6	01-05-2004	well
7	01-21-2006	
8	02-24-2007	well
9	02-22-2008	well
10	January 24, 2009	

Table 2. SSWs considered in this paper

We focus on SSWs that occurred in the period common to all models (1999-2009). Ten SSWs occurred in this period, and they are listed in Table 2. For each event, we also consider the ENSO, MJO, and QBO phase immediately before the event. The ENSO state is characterized using the observed Niño3.4 index extracted from monthly mean ERSSTv5 data (Huang et al., 2017) for the calendar month which contains the day of the SSW. The MJO state is defined following Wheeler and Hendon (2004), and specifically we compute the average amplitude and phase of the RMMs from 5 to 15 days before the SSW in order to characterize the MJO state preceding a SSW. If the amplitude is below 1.0, then the MJO is considered to be inactive. The QBO state is characterized using the observed zonal mean zonal wind at 50hPa in monthly mean NCEP CDAS reanalysis data for the calendar month which contains the day of the SSW. The characterization of a SSW as either split or displacement, and also the day of the event, follows Table 1 of Cohen and Jones (2011). Note that the first day of easterly winds can differ among reanalysis products, however for these events all of the modern reanalyses considered by Butler et al. (2017) agree to within one day of each other. An ensemble member is deemed "successful" if it simulates a SSW within ± 3 days of its actual date. This definition of a "success" follows Taguchi (2016a), Taguchi (2020), D. I. V. Domeisen et al. (2020), Rao, Garfinkel, et al. (2019), and Rao et al. (2020).

3 Results

We begin with a hit map for the SSW that occurred on February 22, 2008 in Figure 1, as this SSW turned out to be the most predictable one of the SSWs considered in this study, and the only SSW with predictability comparable to that of the well-predicted January 2, 2019 event (Rao, Garfinkel, et al., 2019). Nearly all models successfully simulated this SSW for initializations up to ten days before the SSW. For several models hit rates exceeding 50% are present fifteen days before the SSW. Half of the CNRM hindcasts initialized 22 days before the SSW still successfully simulate it. The net effect is that the predictability of this event (as for the January 2, 2019 event) for some modeling systems substantially exceeds the predictability limit of around 15 days for SSWs commonly-cited in previous work. In contrast, other SSWs, e.g. the January 18, 2003 event, are poorly predicted (supplementary Figure 1).

Similar hit maps have been created for all 10 SSWs, and Figure 2 summarizes the predictability of each SSW for each forecast system. Specifically, we list the earliest forecast lead in which at least 50% of the ensemble members successfully simulate the SSW. The frequency with which hindcasts are produced and the specific hindcast dates differ among the models, and hence it can be challenging to directly compare forecast skill between models with, say, daily hindcasts to models with hindcasts every 10 days. Nev-

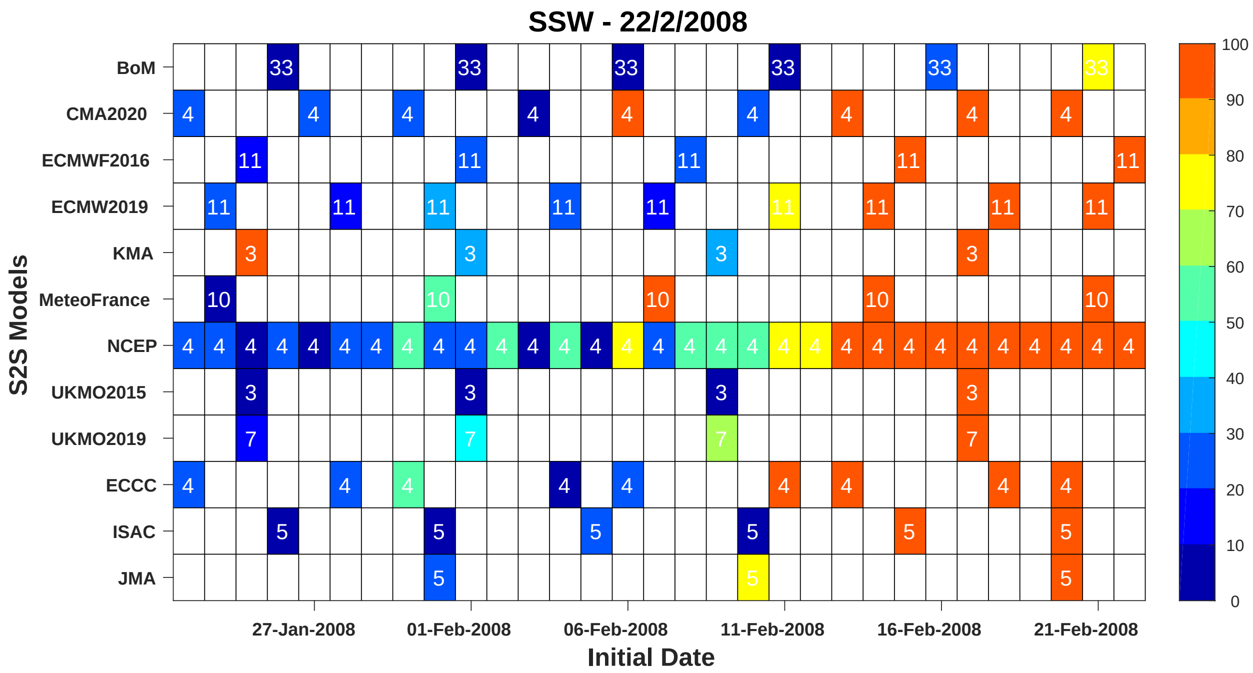

Figure 1. Initialization dates and ensemble sizes of the hindcasts available from each S2S model from 22 January to 22 February 2008. The ensemble size is indicated by the number in each grid cell. The color shading in each grid cell denotes the SSW hit ratio (units: %) of the ensemble members that forecast a reversal of the zonal mean zonal wind at 60N and 10 hPa from 19 February 2008 to 25 February 2008 (i.e., a maximum error of ±3 days is allowed). A blank grid denotes that no real-time predictions were initialized on the specific day for the corresponding model.

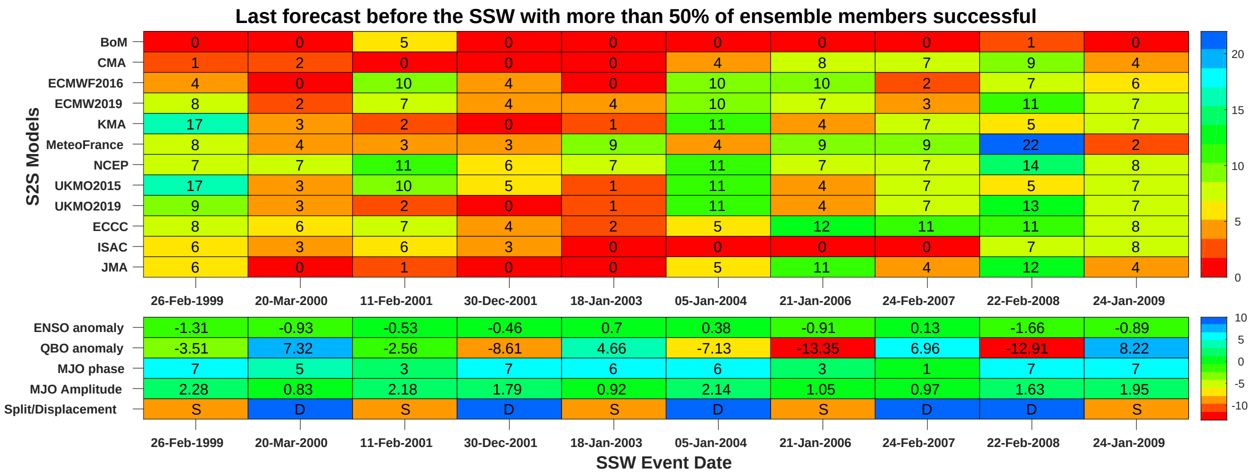

Figure 2. Summary of predictability for all 10 SSWs and 10 models considered in this work. The number of days before the SSW in which at least half of the hindcast ensemble members still simulate a SSW (working backwards from the actual SSW date) is indicated. Also indicated are the ENSO, QBO, and MJO conditions preceding the SSW, as well as the SSW morphology. Note that the low-top version of CMA is used for SSWs in 2003 or earlier, as the hindcasts for the hi-top CMA begin only in 2004.

ertheless, there is a general indication that the low-top BoM and ISAC-CNR struggle as compared to the high-top models, in agreement with previous work. Relatedly, the CMA modeling system is more successful at simulating SSWs starting in 2004 than for earlier SSWs when only the low-top version of CMA is available. The eight high-top models differ in their skill for different events, and given the lack of consistent initialization dates, we refrain from grading them. Rather, our goal going forward is to distinguish between SSWs that are relatively more predictable versus relatively less predictable. Specifically there are four SSWs that were predicted at relatively earlier lead times, and three that were not predicted until much closer to the actual event, as listed in Table 2. Results are similar if instead of using the criterion that 50% of the ensemble members successfully simulate the SSW, we focus on whether the “success” rate is greater than the “false alarm rate”, defined as the number of members that predict an event to occur within vs. outside of ±3 days surrounding the actual event (Supplemental Figure 2).

We begin with the four relatively predictable SSWs. The most predictable SSW was the February 22, 2008 event, which occurred during La Nina, MJO phase 7, and EQBO conditions. Other than La Nina, these are all favorable conditions for enhanced wave driving of the vortex. The next most predictable was the February 26, 1999 event, which occurred during La Nina, MJO phase 7, and neutral QBO. The MJO was the only favorable condition for SSW occurrence. The next most predictable was the January 5, 2004 event, which occurred during neutral ENSO, MJO phase 6, and EQBO. Other than the ENSO phase, these are all favorable conditions for enhanced predictability. The last event with some predictability was the February 24, 2007 event, which occurred during neutral ENSO, inactive MJO, and WQBO. None of these conditions were favorable for SSW predictability, but this event was noticeably more difficult to predict than the previous three.

There were three events with limited or nearly no predictability. The first occurred on March 20, 2000. This event occurred during La Nina, inactive MJO, and WQBO. The lack of a MJO and WQBO led to weak external forcings for a SSW, and hence the mod-

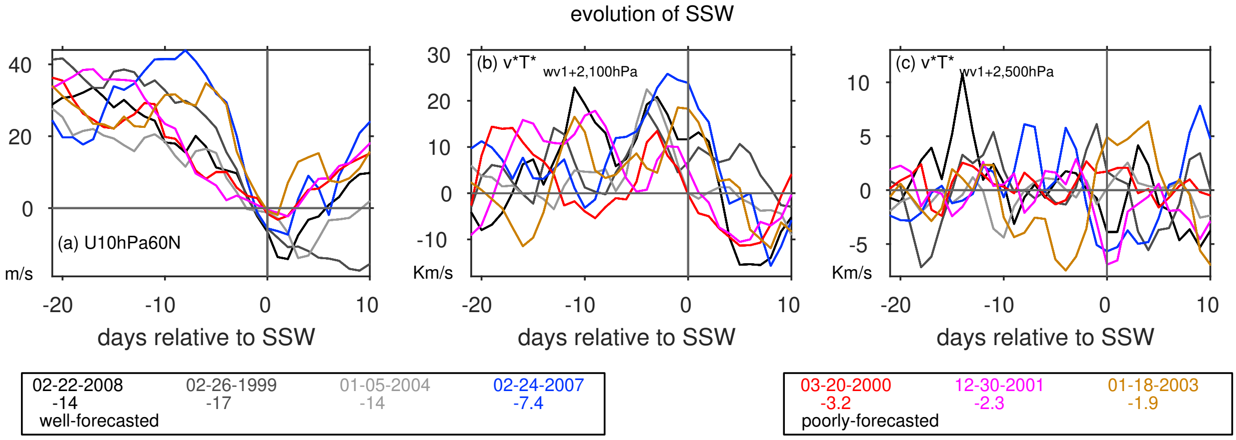

Figure 3. Comparison of well-predicted and poorly-predicted SSWs using ERA-I reanalysis data (Dee et al., 2011). (a) zonal mean zonal wind at 10hPa, 60N; (b) wave-1+wave-2 heat flux anomalies at 100hPa from 40N to 80N; (b) wave-1+wave-2 heat flux anomalies at 500hPa from 40N to 80N. The heat flux anomalies in (b) and (c) are computed relative to a daily climatology.

els struggled. Next, December 30, 2001 was also poorly predicted, and occurred during neutral ENSO, MJO phase 7, and EQBO. Considering that the MJO and ENSO states would seemingly lead to a well-predicted SSW, it is surprising that the models were so unsuccessful, and indeed C. Garfinkel and Schwartz (2017) argued that the MJO helped models predict this event. We will consider this event in more detail later. Finally, the SSW on January 18, 2003 was poorly predicted, and occurred during El Nino, inactive MJO, and WQBO. The lack of MJO and WQBO led to difficulty in simulating this SSW.

These results show that MJO phase 6/7 and to a lesser degree EQBO are more favorable for an improved predictive time among these 10 SSW events, however even the MJO is no guarantee of a successful forecast (Schwartz & Garfinkel, 2017, 2020). In contrast, ENSO seems to be of secondary importance for extended predictability of SSWs. Note that S2S models tend to simulate a stronger seasonal mean vortex for La Nina and a weaker seasonal mean vortex for El Nino, and also a greater probability of easterly winds during El Nino winters, in the full hindcast ensemble(C. I. Garfinkel et al., 2019). However, over the period considered in this study the observed relationship between ENSO and SSW was opposite, with more SSWs during La Nina (D. I. Domeisen et al., 2019), and ENSO also may modulate the impact of the MJO on SSWs (Ma et al., 2020).

Thus far we have demonstrated that long-duration external forcings can contribute to SSW predictability, and now we switch our focus to whether the evolution of the SSW itself helps distinguish well-predicted SSWs from poorly predicted SSWs. Specifically, some SSWs are splits while others are displacement, some are weak with a short period of easterlies or with weak easterlies only, and some events have their peak in wave driving weeks before the central date while others have the peak wave driving immediately before. Is one class of events harder to predict?

We begin by considering vortex morphology. 3 of the 4 well-predicted SSWs were displacements, but so were 2 of the 3 poorly predicted SSWs. (All three SSWs with moderate predictability were splits.) Hence there is little evidence from this analysis for the importance of SSW morphology for predictability, though we acknowledge the caveat that the sample sizes are small.

Next, we consider the importance of the strength of the SSW event. Figure 3a shows the zonal wind evolution for the well-forecasted and poorly forecasted SSWs. The four well-predicted SSWs all featured easterly anomalies of at least -7.4m/s, and most exceeded -14m/s. In contrast the poorly forecasted SSW were all marginal events, which peak easterlies not exceeding -3.2m/s. Hence stronger events were more predictable, though note that the most extreme SSW of the ten we consider, the SSW from January 24, 2009, had near-average predictability.

The evolution of heat flux anomalies before the SSWs also differed between the well-predicted vs. the poorly predicted SSWs (Figure 3b). Specifically, the well-predicted events tended to have anomalously strong heat flux in the week preceding the SSW, while the poorly-predicted SSWs tended to have their anomalously strong heat flux more than 10 days before the SSW, though heat flux anomalies in all cases were positive in the days before the SSW. It is also noteworthy that the well-predicted SSWs all were preceded by robust tropospheric heat flux as well (Figure 3c), though the timing of the heat flux pulse varied among the four events.

Stratospheric preconditioning does not appear to distinguish well-predicted from poorly-predicted events: the 01-05-2004 event was preceded by a weakened vortex for more than 20 days before the event, while the 02-26-1999 was anomalously strong, but both were well-predicted. The stratospheric preconditioning for the poorly predicted SSWs falls within the envelope of these two well-predicted events. Similar results are found if the latitude-height cross section of zonal wind before each of the SSWs is considered (Supplemental Figures 3 and 4). In addition, the rate of deceleration immediately before the SSW also does not seem to be important, as the two events with the most rapid decrease, 01-18-2003 and 02-24-2007, fall in opposite composites. Rather, the only robust distinguishing factor appears to be the minimum wind or the timing of the heat flux pulse.

Finally, we return to the December 30, 2001 event, as this event was poorly-predicted despite favorable tropospheric precursors. Specifically, this event was preceded by a strong MJO phase 6/7 event, and indeed was used by C. Garfinkel and Schwartz (2017) as a case-study of how ensemble members which successfully simulate MJO-related convection tend to better predict the ensuing SSW. We focus on ECMWF hindcasts of this event in Figure 4, with relatively successful ensemble members in blue and other ensemble members in red. On December 26, 2001, zonal winds at 10hPa, 60N in ERA-I weakened to 2.2m/s, though only four days later did they actually reverse. Three of the eleven ECMWF initializations from December 19, 2001 simulated a SSW on December 26, 2001 (indicated in blue), and the ensemble mean vortex strength was weaker than observed (Figure 4a). If we focus on the December 12, 2001 initialization, five of the eleven ensemble members simulated a SSW within three days of December 26, 2001 (indicated in blue), and again the ensemble mean vortex strength was more easterly than observed (Figure 4c). Only the December 5, 2001 initialization can be considered an unambiguous forecast bust: most ensemble members struggle to simulate a weakening of the vortex, and only one member simulates a SSW (Figure 4e). While the December 19, 2001 and December 12, 2001 initializations capture the extremely strong pulse of heat flux in the first week (and over-estimate it for the successful ensemble members initialized on December 12, 2001), the pulse immediately before the SSW is not well represented even in the relatively successful ensemble members, and this late developing pulse appears to have been important for the winds reversing on December 30, 2001. The mid-December pulse of heat flux is under-estimated by the December 5, 2001 initialization, though the three ensemble members that more realistically simulate a weakening of the vortex also do a better job at capturing this wave flux event. The net effect is that the December 30, 2001 SSW was preceded by a strong and long-lasting wave pulse which is generally well represented even in initializations 18 days before the event, however the eventual SSW was comparatively weak and the models struggled to capture its timing due to their failure to capture the late December secondary heat flux pulse. This struggle to capture its tim-

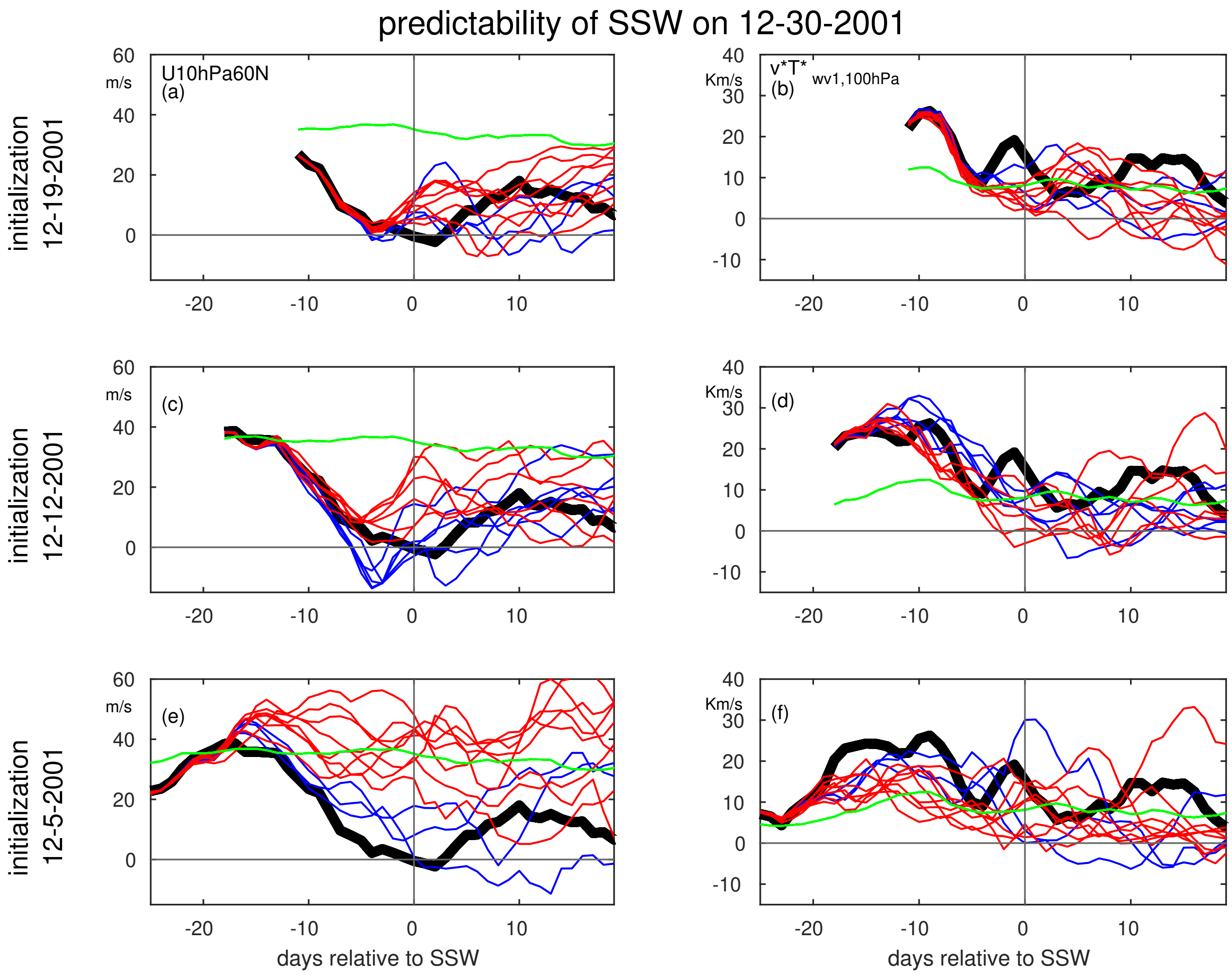

Figure 4. Evolution of the 12-30-2001 SSW in the ECMWF forecast system. ERA-I reanalysis data is shown in thick black, relatively successful ensemble members in blue, poorer ensemble members in red, and the daily climatology in reanalysis in green. (left) zonal mean zonal wind at 10hPa, 60N; (right) wave-1 heat flux at 100hPa from 40N to 80N. We show the wave-1 heat flux only as this was a displacement event.

ing led to a forecast bust if we adopt the ±3 day criteria used by previous work (Taguchi, 2016a, 2020; D. I. V. Domeisen et al., 2020), however alternate definitions may lead to a different conclusion as to the ability of models to forecast this event.

4 Discussion

Stratospheric sudden warmings (SSWs) are associated with a range of surface impacts, and to the extent that SSWs can be predicted on subseasonal timescales, there is hope that the subsequent surface impacts could be predicted at earlier leads. Here we showed that there is wide diversity in the predictability of different SSW events. Some are predictable twenty or more days in advance, while others can only be predicted a week in advance even in the best performing models. Previous work using fewer models and fewer cases has focused on the vortex morphology or the existence of tropospheric precursors as important for SSW predictability, with displacement events and events preceded by the MJO phase 6/7 and by EQBO more predictable. Our results with a relatively larger number of models and SSWs (10 models and 10 SSWs) partially support this previous work: SSWs preceded by MJO phase 6/7 and by EQBO are indeed more predictable (C. Garfinkel & Schwartz, 2017; Rao, Garfinkel, et al., 2019). In addition to these factors, our results also provide evidence for two factors that do not seem to have been noted before. Specifically, stronger SSWs are more predictable, and also SSWs with a peak heat flux immediately before the SSW are also more predictable. (These two factors may be linked, as a late-developing pulse of wave flux can lead to a further rapid deceleration of an already weakening vortex.) Given that stronger SSWs may be expected to have stronger downward impacts (White et al., 2020, 2022), this enhanced predictability for strong events is especially helpful for surface forecasting.

These results stand in contrast to Rao, Ren, et al. (2019) who found that in a single low-top prediction system (40 hybrid sigma/pressure vertical levels, with a lid at 0.5 hPa), more extreme SSWs were harder to predict than more moderate events. It is conceivable that a low-top system will struggle to simulate the dynamics leading up to SSWs, and hence may be less reliable at discerning which SSWs are most predictable. In contrast, the multi-model perspective adopted here indicates that stronger SSWs are more predictable.

A commonly used criteria (and the criteria used in this work) for a successful forecast is that the central date of the simulated SSW falls within ±3 days of the actual event. While this criteria is simple to apply and logical for strong SSWs, it may lead to an underestimate of skill for more marginal events. For example, ECMWF initializations 18 days before the December 30, 2001 event performed remarkably well for the first two weeks of the forecast (Figure 4), yet the ±3 days criteria judges this forecast to be a failure. Future work should consider other possible criterion with which to judge SSW predictability.

5 Open Research

The original S2S database is hosted at ECMWF as an extension of the TIGGE database, and can be downloaded from the ECMWF server `http://apps.ecmwf.int/datasets/data/s2s/levtype=sfc/type=cf/`. The QBO data was downloaded from the NCEP website `https://www.cpc.ncep.noaa.gov/data/indices/qbo.u50.index`. The real time multivariate index of Wheeler and Hendon (2004) was downloaded from the BoM website (`http://www.bom.gov.au/climate/mjo/graphics/rmm.74toRealtime.txt`).

Acknowledgments

DC, CIG, and WC are supported by the ISF-NSFC joint research program (ISF grant No.3259/19 and National Natural Science Foundation of China grant no. 41961144025). CIG and DC were also supported by a European Research Council starting grant under the European Unions Horizon 2020 research and innovation programme (grant agreement No 677756). JR was supported by the National Natural Science Foundation of China (grant no. 42175069). This work is based on S2S data. S2S is a joint initiative of the World Weather Research Programme (WWRP) and the World Climate Research Programme (WCRP).

References

Baldwin, M. P., Ayarzagüena, B., Birner, T., Butchart, N., Butler, A. H., Charlton-Perez, A. J., ... others (2021). Sudden stratospheric warmings. *Reviews of Geophysics*, *59*(1), e2020RG000708.

Baldwin, M. P., Stephenson, D. B., Thompson, D. W. J., Dunkerton, T. J., Charlton, A. J., & O'Neill, A. (2003). Stratospheric memory and skill of extended-range weather forecasts. *Science*, *301*. doi: 10.1126/science.1087143

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.

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.

Charlton, A. J., & Polvani, L. M. (2007). A New Look at Stratospheric Sudden Warmings. Part I: Climatology and Modeling Benchmarks. *Journal of Climate*, *20*, 449-+. doi: 10.1175/JCLI3996.1

Cohen, J., & Jones, J. (2011). Tropospheric Precursors and Stratospheric Warmings. *Journal of Climate*, *24*. doi: 10.1175/2011JCLI4160.1

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., ... others (2011). The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597.

Domeisen, D. I., Garfinkel, C. I., & Butler, A. H. (2019). The teleconnection of el nio southern oscillation to the stratosphere. *Reviews of Geophysics*, *57*. doi: 10.1029/2018RG000596

Domeisen, D. I. V., Butler, A. H., Charlton-Perez, A. J., Ayarzaguena, B., Baldwin, M. P., Dunn-Sigouin, E., ... Taguchi, M. (2020). The role of the stratosphere in sub-seasonal to seasonal prediction. part i: Predictability of the stratosphere. *Journal of Geophysical Research: Atmospheres*, *125*(2), e2019JD030920. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD030920` doi: https://doi.org/10.1029/2019JD030920

Garfinkel, C., & Schwartz, C. (2017). Mjo-related tropical convection anomalies lead to more accurate stratospheric vortex variability in subseasonal forecast models. *Geophysical research letters*, *44*(19).

Garfinkel, C. I., Schwartz, C., Butler, A. H., Domeisen, D. I., Son, S.-W., & White,

I. P. (2019). Weakening of the teleconnection from el niño–southern oscillation to the arctic stratosphere over the past few decades: What can be learned from subseasonal forecast models? *Journal of Geophysical Research: Atmospheres*, *124*(14), 7683–7696.

Garfinkel, C. I., Son, S.-W., Song, K., Aquila, V., & Oman, L. D. (2017). Stratospheric variability contributed to and sustained the recent hiatus in eurasian winter warming. *Geophysical research letters*, *44*(1), 374–382. doi: 10.1002/2016GL072035

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., … Zhang, H.-M. (2017). Extended reconstructed sea surface temperature, version 5 (ersstv5): upgrades, validations, and intercomparisons. *Journal of Climate*, *30*(20), 8179–8205.

Karpechko, A. Y. (2018). Predictability of sudden stratospheric warmings in the ecmwf extended-range forecast system. *Monthly Weather Review*, *146*(4), 1063–1075.

Karpechko, A. Y., Charlton-Perez, A., Balmaseda, M., Tyrrell, N., & Vitart, F. (2018). Predicting sudden stratospheric warming 2018 and its climate impacts with a multimodel ensemble. *Geophysical Research Letters*, *45*(24), 13–538.

Kolstad, E. W., Breiteig, T., & Scaife, A. A. (2010). The association between stratospheric weak polar vortex events and cold air outbreaks in the northern hemisphere. *Quarterly Journal of the Royal Meteorological Society*, *136*(649), 886–893.

Kretschmer, M., Coumou, D., Agel, L., Barlow, M., Tziperman, E., & Cohen, J. (2018). More-persistent weak stratospheric polar vortex states linked to cold extremes. *Bulletin of the American Meteorological Society*, *99*(1), 49-60. doi: 10.1175/BAMS-D-16-0259.1

Lehtonen, I., & Karpechko, A. Y. (2016). Observed and modeled tropospheric cold anomalies associated with sudden stratospheric warmings. *Journal of Geophysical Research: Atmospheres*, *121*(4), 1591–1610. Retrieved from `http://dx.doi.org/10.1002/2015JD023860` (2015JD023860) doi: 10.1002/2015JD023860

Ma, J., Chen, W., Nath, D., & Lan, X. (2020). Modulation by enso of the relationship between stratospheric sudden warming and the madden-julian oscillation. *Geophysical Research Letters*, *47*(15), e2020GL088894.

Marshall, A. G., & Scaife, A. A. (2010). Improved predictability of stratospheric sudden warming events in an atmospheric general circulation model with enhanced stratospheric resolution. *Journal of Geophysical Research: Atmospheres*, *115*(D16). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD012643` doi: https://doi.org/10.1029/2009JD012643

Mukougawa, H., Sakai, H., & Hirooka, T. (2005). High sensitivity to the initial condition for the prediction of stratospheric sudden warming. *Geophysical Research Letters*, *32*(17). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GL022909` doi: https://doi.org/10.1029/2005GL022909

Noguchi, S., Mukougawa, H., Kuroda, Y., Mizuta, R., Yabu, S., & Yoshimura, H. (2016). Predictability of the stratospheric polar vortex breakdown: An ensemble reforecast experiment for the splitting event in january 2009. *Journal of Geophysical Research: Atmospheres*, *121*(7), 3388–3404.

Rao, J., Garfinkel, C. I., Chen, H., & White, I. P. (2019). The 2019 new year stratospheric sudden warming and its real-time predictions in multiple s2s models. *Journal of Geophysical Research: Atmospheres*, *124*. doi: 10.1029/2019JD030826

Rao, J., Garfinkel, C. I., & White, I. P. (2020). Predicting the downward and surface influence of the february 2018 and january 2019 sudden stratospheric warming events in subseasonal to seasonal (s2s) models. *Journal of Geophysical Research: Atmospheres*, *125*(2), e2019JD031919. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD031919` doi: https://doi.org/10.1029/2019JD031919

Rao, J., Ren, R., Chen, H., Liu, X., Yu, Y., Hu, J., & Zhou, Y. (2019). Predictability of stratospheric sudden warmings in the beijing climate center forecast system with statistical error corrections. *Journal of Geophysical Research: Atmospheres*, *124*(15), 8385–8400.

Schoeberl, M. R. (1978). Stratospheric warmings: Observations and theory. *Rev. of Geophys. and Space Physics*.

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*, n/a–n/a. Retrieved from `http://dx.doi.org/10.1002/2016JD025829` (2016JD025829) doi: 10.1002/2016JD025829

Schwartz, C., & Garfinkel, C. I. (2020). Troposphere-stratosphere coupling in subseasonal-to-seasonal models and its importance for a realistic extratropical response to the madden-julian oscillation. *Journal of Geophysical Research: Atmospheres*, *125*(10), e2019JD032043.

Sigmond, M., Scinocca, J., Kharin, V., & Shepherd, T. (2013). Enhanced seasonal forecast skill following stratospheric sudden warmings. *Nature Geoscience*, *6*(2), 98–102. doi: 10.1038/ngeo1698

Taguchi, M. (2014). Stratospheric predictability: Basic characteristics in jma 1-month hindcast experiments for 1979-2009. *Journal of the Atmospheric Sciences*, *71*. doi: 10.1175/JAS-D-13-0295.1

Taguchi, M. (2016a). Connection of predictability of major stratospheric sudden warmings to polar vortex geometry. *Atmospheric Science Letters*, *17*(1), 33-38. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/asl.595` doi: https://doi.org/10.1002/asl.595

Taguchi, M. (2016b). Predictability of major stratospheric sudden warmings: Analysis results from jma operational 1-month ensemble predictions from 2001/02 to 2012/13. *Journal of Atmospheric Sciences*, *73*(2), 789–806.

Taguchi, M. (2018). Comparison of subseasonal-to-seasonal model forecasts for major stratospheric sudden warmings. *Journal of Geophysical Research: Atmospheres*, *123*(18), 10231-10247. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD028755` doi: https://doi.org/10.1029/2018JD028755

Taguchi, M. (2020). Verification of subseasonal-to-seasonal forecasts for major stratospheric sudden warmings in northern winter from 1998/99 to 2012/13. *Advances in Atmospheric Sciences*, *37*(3), 250–258.

Thompson, D. W. J., Baldwin, M. P., & Wallace, J. M. (2002). Stratospheric Connection to Northern Hemisphere Wintertime Weather: Implications for Prediction. *J. Clim.*, *15*, 1421–1428. doi: 10.1175/1520-0442(2002)015

Tripathi, O. P., Baldwin, M., Charlton-Perez, A., Charron, M., Cheung, J. C., Eckermann, S. D., ... others (2016). Examining the predictability of the stratospheric sudden warming of january 2013 using multiple nwp systems. *Monthly Weather Review*, *144*(5), 1935–1960.

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

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., ... others (2017). The sub-seasonal to seasonal prediction (s2s) project database. *Bulletin of the American Meteorological Society*. doi: 10.1175/BAMS-D-16-0017.1

Wheeler, M. C., & Hendon, H. H. (2004). An All-Season Real-Time Multivariate MJO Index: Development of an Index for Monitoring and Prediction. *Monthly Weather Review*, *132*, 1917. doi: 10.1175/1520-0493(2004)132⟨1917:AARMMI⟩2.0.CO;2

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., Garfinkel, C. I., & Hitchcock, P. (2022). On the tropospheric response to stratospheric momentum torques. *Journal of Climate*, *35*. doi: inpress