Predictability of the early winter Arctic oscillation from autumn Eurasian snowcover in subseasonal forecast models

Chaim I. Garfinkel¹ · Chen Schwartz¹ · Ian P. White¹ · Jian Rao¹

Received: 27 November 2019 / Accepted: 19 May 2020 / Published online: 8 June 2020
© The Author(s) 2020

Abstract

The response of the early winter Northern Hemisphere stratospheric polar vortex and tropospheric Arctic Oscillation to anomalous autumn snow cover in Eurasia is evaluated in four operational subseasonal forecasting models. Of these four models, the two with finer stratospheric resolution simulate a weakened vortex for hindcasts initialized with more extensive snow as compared to those with less extensive snow, consistent with the observed effect, though the modeled effect is significantly weaker than that observed. The other two models fail to capture the local Western Eurasian ridge in response to enhanced snow, and hence their failure to simulate a stratospheric response may be due to biases in representing surface–atmosphere coupling rather than their coarser stratospheric resolution per se. There is no evidence of a tropospheric Arctic Oscillation response in early winter in any of these models, which may be related to the weakness of the stratospheric response or (in one model) to too-weak coupling from the stratosphere down to the surface. Overall, the possible contribution of autumn snowcover over Eurasia to improved predictability of the winter Arctic Oscillation in subseasonal forecast models has not yet been realized even in a probabilistic sense.

Keywords October Eurasian snow · Subseasonal forecasting · Stratosphere–troposphere coupling

1 Introduction

Variability of the Northern Hemisphere wintertime stratospheric polar vortex influences tropospheric weather and climate on timescales ranging from the subseasonal to the decadal (Baldwin and Dunkerton 1999; Polvani and Kushner 2002; Limpasuvan et al. 2004; Sigmond et al. 2013; Kidston et al. 2015; Garfinkel et al. 2017). An anomalously strong vortex on average leads to the positive phase of the tropospheric Northern Annular Mode (NAM; also known as the Arctic Oscillation or AO) (Limpasuvan et al. 2005; Tripathi et al. 2015b), while an anomalously weak vortex has largely opposite impacts (Baldwin and Dunkerton 2001; Polvani and Waugh 2004; Limpasuvan et al. 2004; Kidston et al. 2015). Deterministic predictability of the vortex is usually limited to two weeks (Tripathi et al. 2015a, 2016; Karpechko 2018; Domeisen et al. 2020a) though on occasion can extend to almost four weeks (Rao et al. 2019). Probabilistic

predictability, on the other hand, may extend for longer if the planetary wave flux excited in the troposphere is enhanced (Garfinkel and Schwartz 2017; Rao et al. 2019; Domeisen et al. 2020b). A source of predictability for variability of the polar vortex may be autumn Eurasian snow cover.

A thick snowpack tends to both increase surface albedo and decouple the boundary layer from the warmer soil under the snow due to its insulating property (Cohen and Rind 1991). These radiative and thermodynamical processes at the surface can subsequently influence the large-scale atmospheric circulation. Specifically, more extensive snowcover over Eurasia has been shown to precede an increase in upward propagating tropospheric planetary waves which then drives a deceleration of the westerlies of the stratospheric polar night jet (Cohen et al. 2001; Saito et al. 2001; Cohen et al. 2002, 2014). The specific pathway whereby regional snow cover modulates planetary waves is likely via constructive interference with the climatological planetary waves: the climatological planetary waves are characterized by a ridge in Western Eurasia which projects strongly onto both wavenumber 1 and wavenumber 2, and an anomalous ridge in Western Eurasia forced by e.g., enhanced snowcover, constructively interferes with these climatological

✉ Chaim I. Garfinkel
chaim.garfinkel@mail.huji.ac.il

¹ The Fredy and Nadine Herrmann Institute of Earth Sciences,
Hebrew University of Jerusalem, Jerusalem, Israel

planetary waves and leads to stronger planetary waves in the troposphere which can then propagate upwards (Garfinkel et al. 2010; Cohen and Jones 2011; Smith and Kushner 2012; Kretschmer et al. 2017). The net effect is that enhanced October Eurasian snow cover leads to a stronger flux of Rossby waves upwards into the stratosphere in November and a weaker stratospheric polar vortex in December and January, followed by a possible surface negative AO in January and February (Cohen and Entekhabi 1999; Cohen et al. 2007; Fletcher et al. 2009; Allen and Zender 2010) as recently reviewed by Henderson et al. (2018).

Many climate models cannot reproduce the connection between the winter AO and autumn snow cover over Eurasia that appears to exist in reanalysis data (Hardiman et al. 2008; Furtado et al. 2015; Peings et al. 2017). Potential sources of model error include: (i) a boundary layer that does not respond strongly enough to surface forcings, (ii) too-little interannual variability in snow cover, (iii) biased background planetary waves, (iv) coarse stratospheric resolution, (v) interference from other processes such as the Quasi-Biennial Oscillation which also affect the vortex, or (vi) an inability to maintain long-lived ridges or blocking over Western Eurasia after such a ridge is formed by e.g. enhanced snow cover (Peings et al. 2017; Garfinkel et al. 2019; Henderson et al. 2018). Alternately it is possible that the apparent observed effect happened by chance.

Four single-model studies using subseasonal forecast models have analyzed the skill gained by initializing the model with degraded vs. realistic Eurasian snow, most using models that do not resolve the stratosphere (*low-top models*; Jeong et al. 2013; Orsolini et al. 2013, 2016) and one using a model with a lid well above the mesopause (Li et al. 2019). These studies examined differences in forecast skill between pairs of experiments with realistic and alternately degraded initial snow states. The low-top models demonstrate a significant increase in forecast skill in surface air temperature over some regions of the Arctic and Eurasia for up to 2 months, but the positive skill increments generally do not extend far beyond the regions that were snow-covered. For example, Orsolini et al. (2016) used a model with a lid near 5hPa to study the effect of snow cover for the winter of 2009/2010. They found that enhanced snow cover led to surface cooling, stronger wave fluxes leaving the troposphere, a weakening of polar vortex, and a more persistent negative AO, though snow was not the main driver of the AO fluctuations. However, there is strong evidence that stratospheric variability and the stratospheric response to external perturbations are too weak in low-top models (Charlton-Perez et al. 2013; Hurwitz et al. 2014; Garfinkel et al. 2018, 2019; Rao et al. 2019; Domeisen et al. 2020a). Li et al. (2019). formed twin 30-member ensembles of hindcasts, with the snow initialization either degraded or realistic, over the period 1985–2016 using a high-top model. In

addition to regional improvements in surface temperature, the ensemble with realistic snow simulated enhanced wave activity fluxes into the stratosphere after a month, leading to a weakened polar vortex. For the 7 years characterized by a strongly negative phase of the AO, Li et al. (2019) find that snow provides a weak feedback for the negative AO. Overall, studies using individual models have demonstrated a lagged surface response to snow cover variability via the stratosphere. In order to better understand the role of the stratosphere, we here examine four distinct models, two with a lid well above the stratopause and two with a less well-resolved stratosphere.

The Subseasonal-to-Seasonal (S2S) Prediction project (Vitart et al. 2017) has recently organized and collected a multi-model ensemble of hindcasts covering the past several decades. Here we consider whether hindcasts from the 4 S2S models with the requisite data availability successfully simulate an enhanced snow → Western Eurasian high → enhanced planetary wave flux → weak vortex → negative AO chain of connections over the past few decades. We are not aware of any previous multi-model study evaluating the ability of subseasonal forecast models to capture these connections. As these hindcasts are initialized with observed snow-cover and atmospheric conditions (including the QBO, Garfinkel et al. 2018), and as they are used for evaluating the skill of operational forecasts, they can be directly compared to the observed atmospheric state. Furthermore, because the initial model state resembles the observed state, we can track the stages in the development of the response to snow and pinpoint where models fail. Section 2 introduces the data and methods. In Sect. 3, we show that individual high-top S2S models can capture the first few links in this chain of connections, though none can capture the full chain of connections. We also identify the proximate cause of this failure for each model.

2 Data and methods

The lagged connection between October Eurasian snow cover and the winter AO is evaluated in the model output submitted to the S2S Prediction project (Vitart et al. 2017). Because the observed lagged response takes more than a month to develop, we only include forecast centers who submitted data for at least 44 days after initialization and which also archive a product related to land snow cover. Four modeling centers met these criteria at the time we started downloading data—the European Centre for Medium-Range Weather Forecasts (ECMWF), the National Center for Environmental Prediction (NCEP), the Australian Bureau of Meteorology (BoM), and the China Meteorological Administration (CMA). Table 1 summarizes the reforecasts analyzed in this study. For the NCEP model we

Table 1 S2S model experiments chosen

Model (ensemble members)	Years	Reforecasts analyzed	Vertical levels	Model top
NCEP (4)	1999–2010	9 per month	64	0.02 hPa
ECMWF (11)	1996–2013	4 per month	91	0.01 hPa
BoM (33)	1981–2013	6 per month	17	10 hPa
CMA (4)	1999–2014	6 per month	40	0.5 hPa

only downloaded nine reforecasts each month, and for the ECMWF model we downloaded only one reforecast each week, to more closely match the model output available for the other models. The stratosphere is more finely resolved in ECMWF and NCEP relative to the other forecast systems (Table 1), and previous work has found that the quality of their reforecasts and forecasts is indeed higher for processes involving the stratosphere (Garfinkel et al. 2018, 2019; Rao et al. 2019; Domeisen et al. 2020a, b). We consider forecasts initialized in late October (between October 16th and October 30th) and early November (between November 1st and November 15th) and consider the response to snow cover up to two months later, so as to focus on the season when observed snow cover precedes anomalous zonal wind at 10 hPa and 60° N (Cohen et al. 2007). We have also examined early October initializations, and no forecast system shows a significant response to snow cover that mimics the observed response.

The goal of this paper is not to evaluate the fidelity of the snow cover initialization of the respective models. Rather, our focus is on whether the snow cover with which each hindcast was initialized has an impact on the subsequent evolution of the hindcast on large scales. Therefore, the details of how we identify high and low snow cover years differs among the models in order to maintain focus on hindcasts in which the model was actually initialized with anomalous snow coverage. Taking the CMA hindcasts as an example, we first compute the area of the region from 25 N to 60 N and from 0 E to 180 over Eurasia that is covered with snow albedo values larger than 0.805 for all hindcasts and leads. The 0.805 threshold is chosen as the subsequent climatological areal coverage of snow matches that in the NOAA/NCDC snow extent product (<https://nsidc.org/data/NSIDC-0046/versions/4>; updated from Estilow et al. 2015; Brown and Robinson 2011) over this region on the first day of forecasts in late October and early November. For the other models the threshold value is also chosen in order to match the climatological areal coverage of snow in the NOAA/NCDC snow extent product on the first day of forecasts in late October and early November: for the ECMWF model we therefore use a threshold snow albedo value of 0.73. For the BoM and NCEP models we use the snow depth water equivalent product as the snow albedo product is not available, and look for values larger than 2

kg/m² and 5 kg/m² respectively. For consistency, we also show results using the snow depth water equivalent product for CMA and ECMWF, with a threshold of 1 kg/m² and 5 kg/m² respectively. Clearly, one could have a high albedo but a low water equivalent if snow coverage is extensive but the snowpack is thin. Such a situation could mitigate some of the snow thermodynamical feedbacks but radiative effects would be strong. Conversely, relatively low albedo but high snow water equivalent could occur if snow coverage was patchy but locally deep, in which case thermodynamic feedbacks would be stronger than radiative feedbacks. Many land-surface models also include a snow-aging effect whereby snow albedo decreases with time (e.g. ECMWF, Dutra et al. 2010), so one could also have low albedo but high snow water equivalent if there is extensive snow cover but the snow has aged; alternately, fresh snow over a thick snowpack would increase the albedo even as the thermodynamic effect is nearly unchanged. These factors cannot be individually evaluated given the limited output available in the S2S database.

If the modeled snow coverage over the second and third day after initialization is greater than average for that calendar date and that model, then that initialization is composited as high snow, and vice versa for below average snow coverage. All models maintain their anomalies in snow cover in the first 10 days when comparing late October initializations that were initialized with high snow as compared to low snow (Fig. 1), and for ECMWF there are no biases for the duration of reforecast (Fig. 1a). For other models biases in snow cover develop after the first two weeks (the lines corresponding to model and observations diverge), however the biases are less than 20% of observed snow cover. There is only a weak correspondence within a given season between anomalous snow in late October and early November; taking ECMWF as an example, the lag-one week autocorrelation of snow extent anomalies the first day of each reforecast over October and November is 0.28 upon averaging the individual autocorrelations for each autumn and over each ensemble member, while the lag-two week autocorrelation is negative. We exclude hindcasts in which the initial zonal mean wind at 60N, 10hPa is weaker than 10m/s so as to avoid initial stratospheric states without a vortex on which wave propagation may be suppressed (Charney and Drazin 1961).

Fig. 1 Snow cover extent (km^2) as a function of forecast date for the high snow and low snow composites, and also for all years. Weekly-averaged observed snow cover for the dates chosen for each model is also shown. The threshold value of snow water equivalent (SWE) for classifying a gridpoint as snow covered is indicated for each model, and we then sum the area of all gridpoints that are classified as snow-covered

Our next step is to form composites of all hindcasts initialized during high-snow conditions and during low-snow conditions, and compare the weekly averaged response between the two composites for the duration of the hindcasts. We also form corresponding composites based on these same dates but using the observed atmospheric state as estimated by the MERRA (modern-era retrospective analysis for research and applications; Rienecker et al. 2011) reanalysis. We then compare the difference between the high-snow and low-snow composites in the reforecasts and in MERRA in order to assess the fidelity of the response to snow in the S2S forecast systems. Note that each forecast system has contributed hindcasts from different years, and even for a given year the initialization dates differ among the models. We therefore subsample MERRA data separately for each model so as to match the dates simulated by each forecast system, in order to meaningfully compare the modeled and observed stratospheric response to snow cover.

The following metrics of the extratropical circulation are used to quantify the strength of the snow-vortex connection: 70°N and poleward area-weighted averaged geopotential height at 500 hPa, 100 hPa, and 10 hPa (Z_{500} , Z_{100} , Z_{10}); zonal wind at 10hPa and 60°N (U_{10}); heat flux ($\overline{v'T'}$) at 500hPa and 100hPa area-weighted from 40°N to 80°N for zonal wavenumbers 1 and 2. Note that the S2S database only contains data at three stratospheric levels (10 hPa, 50 hPa, and 100 hPa). Polar cap geopotential height and U_{10} are tightly coupled to the Northern Annular Mode (Baldwin and Thompson 2009). The heat flux is proportional to the vertical component of the Eliassen-Palm flux, and hence heat flux at 100hPa serves as a proxy for wave activity in the lower stratosphere (e.g Newman et al. 2001). Anomalous heat flux at 100hPa can alternately reflect altered wave generation in the troposphere or altered conditions for wave propagation within the stratosphere (Cámara et al. 2017; White et al. 2019), and hence we also examine 500 hPa heat flux to better connect snow anomalies to their stratospheric response. We show the weekly average of the aforementioned indices in our figures.

A two-tailed difference of means Student t test is used to assess the statistical significance of the difference between high-snow and low-snow composites, and the null hypothesis is that there is no difference between the two composites. Each hindcast ensemble member is considered to be a separate degree of freedom, as we focus on the response beyond

the first week after internal variability has already led to differences among the ensemble members. For the observed response, each year is treated as one degree of freedom.

The statistical significance of differences between the hindcasts and MERRA in the response to snow is computed via a Monte Carlo resampling technique. As an example, taking the response of Z10 for initializations in late October in ECMWF:

- Step 1: We randomly select late October initializations from the full ECMWF hindcast ensemble to match the number of actual high-snow and low-snow initializations, without allowing the same initialization to be placed in both the high-snow and low-snow fictitious samples.
- Step 2: We then compute the difference in Z10 between the means of these fictitious high-snow and low-snow samples, both for ECMWF and for the corresponding dates in MERRA reanalysis.
- Step 3: We then compute the difference in Z10 between the ECMWF and MERRA responses for these fictitious composites.
- Step 4: Steps 1 through 3 are repeated 3000 times, with different initializations randomly selected for the fictitious high-snow and low-snow composites. We then compute the bottom and top 2.5% quantiles of the difference between the ECMWF and MERRA without making any assumption on the nature of the distribution, to which we compare the actual difference between ECMWF and MERRA.

3 Can the S2S models capture the observed connection between October snow and the vortex?

As discussed in the introduction, Eurasian snow may lead to surface predictability at 2 month leads through an enhanced snow → Western Eurasian high → enhanced planetary wave flux → weak vortex → negative AO chain of associations. We now consider whether the S2S models capture this chain of effects. We begin by considering whether the forecast systems simulate a mid-tropospheric ridge over Western Eurasia.

Figure 2 shows the difference in 500hPa height between forecasts initialized during more extensive snow cover as opposed to forecasts initialized during less extensive snow cover in late October for the ECMWF forecast system using the snow albedo criteria to determine anomalous snow cover. In the second week of the integrations, a ridge appears in Western Eurasia (near the red square) in both the ECMWF forecasts and in reanalysis data. This ridge persists beyond week 3 in the reanalysis data, but is shorter-lived in the ECMWF forecast system.

We now compare across the four forecast systems and focus on 500hPa height over the region 55–65 N and 35–65 E (the red boxed region in Fig. 2) in Fig. 3a, b. The thick black line in Fig. 3a shows the high–low snow difference for the ECMWF model in this region for each week using the albedo criterion, and dots indicate that the composited response to high versus low snow is significantly different using a Student's *t* test. The thick black line in Fig. 3b shows the high–low snow difference in MERRA after subsampling MERRA to match the dates used to compute the thick black line in Fig. 3a, and dots indicate that there is a significant difference between the reanalysis and modeled response to snow upon applying the Monte Carlo test described in Sect. 2. Consistent with Fig. 2, the ECMWF forecast system simulates a ridge over Western Eurasia in response to enhanced snow albedo in week two of similar strength to that observed. The NCEP model (green) also successfully simulates this ridge in week 2 and 3, but the BoM (blue) and CMA (cyan) models do not. The lack of a response in BoM and CMA is consistent with the rapid decays of snow cover extent anomalies in Fig. 1c, d as compared to ECMWF and NCEP (Fig. 1a, b). In the fifth forecast week, there is a statistically significant difference between the Western Eurasian ridge in all forecast systems as compared to that observed. NCEP also simulates a ridge in response to enhanced snow cover for initializations in early November (Figs. 4a,b, 5a, c, e).

As discussed in Garfinkel et al. (2010), a ridge in this region constructively interferes with the climatological stationary wave-1 and wave-2. Figure 3c, d considers whether the sum of wave-1 and wave-2 500hPa heat flux, a proxy for upward propagating tropospheric planetary wave flux, is enhanced in response to enhanced snow cover. The NCEP model and the ECMWF model successfully simulate enhanced planetary heat flux in week 2 (in agreement with the case study of Orsolini et al. 2016), but the enhanced heat flux does not persist for later periods. A similar increase in planetary-wave heat flux is evident at 100 hPa for the NCEP model and the ECMWF model (Fig. 3e, f), though the increase in 100hPa heat flux in NCEP is somewhat smaller than that in ECMWF even though the increase in 500hPa heat flux is larger in NCEP. This tendency for weak upward propagation in NCEP for late October initializations is not evident for early November initializations (Figs. 4c, 7e), and hence may not reflect an actual model bias. In contrast, both the BoM and CMA forecasting systems fail to simulate enhanced heat flux at either 500 hPa or 100 hPa for late October initializations, consistent with the lack of a response in Western Eurasian height. Even in NCEP and ECMWF, the enhanced wave flux is shorter lived than the observed enhanced wave flux, consistent with the relatively fast decay of the Western Eurasia ridge. It is also worth noting that NCEP simulates

Fig. 2 Anomalies (high snow-extent–low snow-extent) in Z500 for initializations between October 16 and October 31 in the ECMWF model in the **a, b** first week, **c, d** second week, **e, f** third week, and **g, h** sixth week after initialization. The right is for MERRA subsampled to match the dates chosen for ECMWF. Difference-of-means between more snow initializations and less snow initializations that are statistically significant at the 95% confidence level as given by a Student-t test are denoted with dots. The zero-line is indicated in black. The 70m contour of the climatological eddy height field is indicated in green. The region 35 E–65 E, 55 N–65 N is indicated with a red box

a subpolar North Pacific low in response to enhanced snow cover for initializations in early November as observed in the first two weeks after initialization (Fig. 5a–d), and such a low has also been linked to enhanced wave driving of the vortex (Garfinkel et al. 2010).

Does the enhanced wave flux in the NCEP and ECMWF forecasting systems affect the stratosphere? Figure 6a, b shows the U10 response to snow for late October initializations, while Fig. 7a, b is similar but for early November

initializations. The only forecasting system which shows a weakened vortex in response to enhanced October snow coverage is ECMWF for the albedo criterion (in thick black), though for later lags the response in ECMWF is too weak. In contrast, only the NCEP model shows a strong response to snow cover for initialization dates in early November (green in Fig. 7a). The CMA and BoM models fail to show any stratospheric response to snow cover, though they are able to persist the slightly weaker vortex for early November

Fig. 3 Difference between high snow-extent and low snow-extent (high snow-extent–low snow-extent) for 4 different S2S models as a function of forecast week for reforecasts initialized between October 16 and October 31. Difference-of-means between high snow-extent and low snow-extent that are statistically significant at the 95% confidence level as given by a Student-t test are denoted with dots on the left column; for the right column dots indicate that there is a significant difference between the modeled and observed response to snow cover as given by the Monte Carlo test described in Sect. 2. The number of reforecast ensemble members for each model and snow composite is indicated on the figure legend. **a, b** 500 hPa geo-

potential height over Western Eurasia, defined here as 35 E–65 E, 55 N–65 N (red box on Fig. 2) in order to capture the region where the observed response to anomalous snow peaks; Area-weighted average of wavenumber 1 and 2 heat flux ($v'T'$) at **c, d** 500 hPa and **e, f** 100 hPa from 40° N to 80° N. Thick black and thick cyan curves are for the snow albedo criterion for CMA and ECMWF, while thin lines are for the snow depth water equivalent criterion for all four forecast systems. Left column is for the S2S models and the right column is for MERRA subsampled to match the specific dates included in each composite for each model

initializations for the duration of their integration (Fig. 7a). Results are similar for Z10 (Fig. 6c, d): the ECMWF and NCEP models indicate that more extensive snow cover leads to relatively higher polar-cap heights and a weaker vortex for initializations in late October and early November respectively, but the low-top models BoM and CMA simulate no such effect (and if anything an opposite signed response for late October initializations).

Previous work has shown that stratospheric anomalies can precede those at the surface (Baldwin and Dunkerton 1999; Polvani and Kushner 2002; Limpasuvan et al. 2004; Sigmond et al. 2013; Kidston et al. 2015), and we now consider whether the mid-stratospheric response in NCEP and ECMWF is followed by a response lower down. While the ECMWF forecasting system captures an effect on Z10, it is

significantly weaker than that observed in weeks 5 and 6, and the modeled effect weakens for Z100 and disappears entirely for Z500 (thick black in Fig. 6e–h). Hence there is no clear evidence of downward propagation or surface impacts from the stratosphere. Consistent with this, Fig. 2g, h shows that while a negative AO was observed following higher snow cover in reanalysis data, no such response is evident in the ECMWF forecast system. The NCEP model captures a strong effect for both Z10 and Z100 (green in Fig. 7c, e), but the impacts in the troposphere are only strong at early lags and disappear at later lags (green in Figs. 7g and 5g, h), though note that the observed response to altered snow in early November also does not resemble an AO pattern.

Both the NCEP and ECMWF forecast systems simulated a weaker lower stratospheric vortex in response to more

Fig. 4 As in Fig. 3 but for initialization between November 1 and November 15

extensive snow cover after ~ 5 weeks, though no tropospheric response followed. Figure 8 considers whether these two models are capable of capturing downward coupling from the lower stratosphere to the surface, neglecting any possible factor that may have led to the weaker vortex to begin with. Specifically Fig. 8 shows the correlation between polar cap height at 100 hPa in week 5 and sea level pressure anomalies in week 6 for initializations in November. In ECMWF the correlations peak at around 0.6, and so more than a third of the variance is shared. Such a correlation is (if anything) larger than that observed (perhaps due to overly persistent vortex anomalies in the parent model; Orsolini et al. 2018), and hence the lack of a tropospheric response in ECMWF is more due to the weakness of the stratospheric response. In contrast, NCEP struggles with downward coupling for November initializations (though for forecasts initialized later in winter, correlations are similar to that observed, Schwartz and Garfinkel 2020), and hence the

lack of a tropospheric response in NCEP to snow may be due to a model inability to represent a tropospheric response to a weakened lower stratospheric vortex (regardless of its source) in early winter.

4 Discussion and summary

October Eurasian snow cover appears to have modulated extratropical surface climate 2–3 months later via its polar stratospheric teleconnection (Cohen et al. 2007, 2014; Henderson et al. 2018), and previous studies have indicated that the snow-vortex connection may provide skill for subseasonal and seasonal forecasting (Orsolini et al. 2016; Li et al. 2019). While previous studies focusing on individual models have reported success in the ability of models to capture the lagged surface response

Fig. 5 As in Fig. 2 but for NCEP initializations between November 1 and November 15

to snow cover variability via the stratosphere, previous multi-model studies using climate models have reached less encouraging conclusions. Here we consider whether models participating in the Subseasonal-to-Seasonal (S2S) Prediction project (Vitart et al. 2017) successfully simulate an enhanced snow → Western Eurasian high → enhanced planetary wave flux → weak vortex → negative AO chain of connections over the past few decades. Because the initial model state resembles the observed state, we can track the stages in the development of the response to snow and

pinpoint where models fail, while in free-running climate models diagnosing the location of model failure is less straightforward.

None of the S2S models analyzed in this work are able to capture the full range of associations between enhanced snow and the negative NAO around 2 months later. The various models fail in capturing the suggested pathway at different points. The BoM and CMA models fail both to maintain the snow cover anomalies beyond the first 2 weeks and to simulate a Western Eurasian tropospheric response to snow

Fig. 6 As in Fig. 3 but for **a, b** zonal mean zonal wind at 60° N, 10 hPa; area-weighted average of geopotential height from 70° N to the pole **c, d** at 10 hPa, **e, f** at 100 hPa, and **g, h** at 500 hPa

cover anomalies, and hence the lack of a response may not be due exclusively to the low model top. The NCEP model was able to capture the Western Eurasian tropospheric response, the subsequent enhanced heat flux, a weaker vortex, and then downward propagation to the lower stratosphere. However the signal did not then propagate downwards into the troposphere, and more generally the troposphere in the NCEP model is not sensitive enough to early winter stratospheric conditions. The strong response for NCEP was evident more for early November initializations and not for late October initializations when the observed response is stronger. For late October initializations, the NCEP model had trouble capturing upward propagation of wave flux from 500 to 100 hPa, and consistent with this the subsequent stratospheric response was weak. Finally, the ECMWF model was able to capture the Western Eurasian tropospheric response, the subsequent enhanced heat flux, and a weaker vortex at 10 hPa, but the subsequent downward propagation was too weak and little tropospheric response was evident. Note, however, that while there is an observed tropospheric NAM response five weeks after the late October initializations, the observed

tropospheric NAM response peaks in January (Cohen et al. 2007) after these hindcasts have ended.

Two models made available multiple snow products (CMA and ECMWF), and while the CMA forecast system fails to simulate a strong response to either snow product, the ECMWF forecast system indicates a strong response to snow in late October only for the albedo product but not for the snow water equivalent product. In other words, ECMWF initializations in late October with extensive snow coverage but a thin snowpack indicate a significant vortex response as compared to initializations with opposite signed anomalies, while ECMWF initializations with a patchy but deep snowpack indicate little vortex response as compared to initializations with opposite signed anomalies. The primacy of albedo suggests that snow thermodynamical feedbacks are relatively unimportant in comparison to radiative effects in late October, at least for the ECMWF forecast system, though future work is needed to confirm this conclusion. In contrast, in early November no response is evident to either snow product in the ECMWF forecast system after the second week. The reduced importance of albedo in early November could be due to the fact that incoming solar radiation is rapidly

Fig. 7 As in Fig. 6 but for initialization between November 1 and November 15

Fig. 8 Correlation between polar cap height at 100 hPa in week 5 and sea level pressure anomalies in week 6 for initializations in November in NCEP and ECMWF. Dots indicate a null hypothesis of no correlation cannot be rejected at the 5% confidence level

decreasing, and thus there is little sensitivity to changes in albedo in early November.

Even though the NCEP and ECMWF models capture a Western Eurasian tropospheric response to snow cover anomalies in the second week after initialization, this ridge is too

short-lived. Garfinkel et al. (2019) also found that these same S2S models fail to maintain ridging over Western Eurasia if initialized with such a ridge. It is possible that model deficiencies in e.g. the representation of Western Eurasia blocking (biases are present in CMIP5 models and also in some

high resolution models, Masato et al. 2013; Dunn-Sigouin and Son 2013; Schiemann et al. 2017) may contribute to the weakened response in this region in the first few weeks of the hindcasts and subsequently may contribute to the weakness of the AO response. On the other hand, NCEP simulates a ridge in response to enhanced snow cover for initializations in early November that is stronger and more persistent than that observed (Figs. 4a, b, 5a–f). A thorough investigation of whether the S2S models are capable of accurately simulating blocking frequency and duration is beyond the scope of this work.

It is worth noting that it is possible that the failure of any of the models to capture the full observed response may not solely reflect model deficiencies, but rather the observed apparent effect may have been enhanced by unrelated variability. The observed effect is non-stationary in time (Peings et al. 2013), and only emerged since the 1970s and peaked over the period 1991–2010 (Peings et al. 2013), before collapsing since 2013 (supplemental figure 1 of Henderson et al. 2018). Hence the composited response in reanalysis data in this paper, which only includes the period in which S2S data are available, may be an overestimate of the “true” magnitude of the response to snow forcing that would be inferred from a longer observational record (Henderson et al. 2018). Comprehensive climate models also cannot reproduce the observed apparent 2-month lead linkage between autumn Eurasian snow cover and the winter mean AO (Hardiman et al. 2008; Furtado et al. 2015), including some studies with controlled forcings (Peings et al. 2017), and this study shows that high-top subseasonal forecast models also suffer from a too-weak response if observations are taken at face-value (in apparent agreement with Li et al. 2019). A small signal-to-noise ratio clearly requires a large ensemble of realizations when exploring the remote extratropical atmospheric response to a boundary forcing, and our ensemble sizes of between 100 and 200 events (for models other than BoM) may simply be too small. Nevertheless, two of the four models analyzed here simulate at least qualitatively several of the large-scale effects found in reanalysis data that are most closely linked to the snow anomalies, and these are the two models that have been shown to perform better in their simulation of other aspects of the stratospheric circulation (Garfinkel et al. 2018, 2019; Rao et al. 2019; Domeisen et al. 2020a, b). Hence the underlying mechanism is likely real.

Acknowledgements This work was supported by a European Research Council starting grant under the European Union’s Horizon 2020 research and innovation programme (Grant agreement no. 677756). 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). 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/>. We thank the two anonymous reviewers and Judah Cohen for their helpful comments. Correspondence should be addressed to C.I.G. (email: chaim.garfinkel@mail.huji.ac.il).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allen RJ, Zender CS (2010) Effects of continental-scale snow albedo anomalies on the wintertime arctic oscillation. *J Geophys Res* 115 (D23105). <https://doi.org/10.1029/2010JD014490>
- Baldwin MP, Dunkerton TJ (1999) Propagation of the Arctic oscillation from the stratosphere to the troposphere. *J Geophys Res* 104(D24):30937–30946
- Baldwin MP, Dunkerton TJ (2001) Stratospheric harbingers of anomalous weather regimes. *Science* 294(5542)
- Baldwin MP, Thompson DWJ (2009) A critical comparison of stratosphere-troposphere coupling indices. *Q J R Meteorol Soc* 135
- Brown RD, Robinson DA (2011) Northern hemisphere spring snow cover variability and change over 1922–2010 including an assessment of uncertainty. *Cryosphere* 5(1):219–229
- Cámara Adl, Albers JR, Birner T, Garcia RR, Hitchcock P, Kinnison DE, Smith AK (2017) Sensitivity of sudden stratospheric warmings to previous stratospheric conditions. *J Atmos Sci* 74(9):2857–2877
- Charlton-Perez AJ, Baldwin MP, Birner T, Black RX, Butler AH, Calvo N, Davis NA, Gerber EP, Gillett N, Hardiman S et al (2013) On the lack of stratospheric dynamical variability in low-top versions of the cmip5 models. *J Geophys Res Atmos* 118(6):2494–2505
- Charney JG, Drazin PG (1961) Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *J Geophys Res* 66:83–109. <https://doi.org/10.1029/JZ066i001p00083>
- Cohen J, Entekhabi D (1999) Eurasian snow cover variability and northern hemisphere climate predictability. *Geophys Res Lett* 26:345–348. <https://doi.org/10.1029/1998GL900321>
- Cohen J, Jones J (2011) Tropospheric precursors and stratospheric warmings. *J Clim* 24. <https://doi.org/10.1175/2011JCLI4160.1>
- Cohen J, Rind D (1991) The effect of snow cover on the climate. *J Clim* 4(7):689–706
- Cohen J, Saito K, Entekhabi D (2001) The role of the Siberian high in Northern Hemisphere climate variability. *Geophys Res Lett* 28:299–302
- Cohen J, Salstein D, Saito K (2002) A dynamical framework to understand and predict the major Northern Hemisphere mode. *Geophys Res Lett* 29(10). <https://doi.org/10.1029/2001GL014117>
- Cohen J, Barlow M, Kushner PJ, Saito K (2007) Stratosphere-troposphere coupling and links with Eurasian land surface variability. *J Clim* 20:5335–5343
- Cohen J, Furtado JC, Jones J, Barlow M, Whittleston D, Entekhabi D (2014) Linking Siberian snow cover to precursors of stratospheric variability. *J Clim* 27(14):5422–5432. <https://doi.org/10.1175/JCLI-D-13-00779.1>
- Domeisen DIV, Butler AH, Charlton-Perez AJ, Ayarzagüena B, Baldwin MP, Dunn-Sigouin E, Furtado JC, Garfinkel CI, Hitchcock P, Karpechko AY, Kim H, Knight J, Lang AL, Lim EP, Marshall A, Rolff G, Schwartz C, Simpson IR, Son SW, Taguchi M (2020a) The role of the stratosphere in sub-seasonal to seasonal prediction.

- Part I: Predictability of the stratosphere. *J Geophys Res Atmos* (in press)
- Domeisen DJV, Butler AH, Charlton-Perez AJ, Ayarzagüena B, Baldwin MP, Dunn-Sigouin E, Furtado JC, Garfinkel CI, Hitchcock P, Karpechko AY, Kim H, Knight J, Lang AL, Lim EP, Marshall A, Rolf G, Schwartz C, Simpson IR, Son SW, Taguchi M (2020b) The role of the stratosphere in sub-seasonal to seasonal prediction. Part II: Predictability arising from stratosphere-troposphere coupling. *J Geophys Res Atmos* (in press)
- Dunn-Sigouin E, Son SW (2013) Northern hemisphere blocking frequency and duration in the cmip5 models. *J Geophys Res Atmos* 118(3):1179–1188. <https://doi.org/10.1002/jgrd.50143>
- Dutra E, Balsamo G, Viterbo P, Miranda PMA, Beljaars A, Schär C, Elder K (2010) An improved snow scheme for the ecmwf land surface model: Description and offline validation. *J Hydrometeorol* 11(4):899–916. <https://doi.org/10.1175/2010JHM1249.1>
- Estilow TW, Young AH, Robinson DA (2015) A long-term northern hemisphere snow cover extent data record for climate studies and monitoring. *Earth Syst Sci Data* 7(1):137–142
- Fletcher CG, Hardiman SC, Kushner PJ, Cohen J (2009) The dynamical response to snow cover perturbations in a large ensemble of atmospheric GCM integrations. *J Clim* 22:1208–1222. <https://doi.org/10.1175/2008JCLI2505.1>
- Furtado JC, Cohen JL, Butler AH, Riddle EE, Kumar A (2015) Eurasian snow cover variability and links to winter climate in the cmip5 models. *Clim Dyn* 45(9–10):2591–2605
- Garfinkel C, Schwartz C (2017) Mjo-related tropical convection anomalies lead to more accurate stratospheric vortex variability in subseasonal forecast models. *Geophys Res Lett* 44(19)
- Garfinkel CI, Hartmann DL, Sassi F (2010) Tropospheric precursors of anomalous northern hemisphere stratospheric polar vortices. *J Clim* 23. <https://doi.org/10.1175/2010JCLI3010.1>
- Garfinkel CI, Son SW, Song K, Aquila V, Oman LD (2017) Stratospheric variability contributed to and sustained the recent hiatus in eurasian winter warming. *Geophys Res Lett* 44(1):374–382. <https://doi.org/10.1002/2016GL072035>
- Garfinkel CI, Schwartz C, Domeisen DI, Son SW, Butler AH, White IP (2018) Extratropical atmospheric predictability from the quasi-biennial oscillation in subseasonal forecast models. *J Geophys Res Atmos* 123(15):7855–7866
- Garfinkel CI, Schwartz C, Butler AH, Domeisen DI, Son SW, White IP (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? *J Geophys Res Atmos* 124(14):7683–7696
- Hardiman SC, Kushner PJ, Cohen J (2008) Investigating the ability of general circulation models to capture the effects of eurasian snow cover on winter climate. *J Geophys Res* 113(D21123). <https://doi.org/10.1029/2008JD010623>
- Henderson GR, Peings Y, Furtado JC, Kushner PJ (2018) Snow-atmosphere coupling in the northern hemisphere. *Nat Clim Change* 8(11):954–963
- Hurwitz MM, Calvo N, Garfinkel CI, Butler AH, Ineson S, Cagnazzo C, Manzini E, Peña-Ortiz C (2014) Extra-tropical atmospheric response to enso in the cmip5 models. *Clim Dyn* 43(12):3367–3376
- Jeong JH, Linderholm HW, Woo SH, Folland C, Kim BM, Kim SJ, Chen D (2013) Impacts of snow initialization on subseasonal forecasts of surface air temperature for the cold season. *J Clim* 26(6):1956–1972
- Karpechko AY (2018) Predictability of sudden stratospheric warmings in the ecmwf extended-range forecast system. *Mon Weather Rev* 146(4):1063–1075
- Kidston J, Scaife AA, Hardiman SC, Mitchell DM, Butchart N, Baldwin MP, Gray LJ (2015) Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. *Nat Geosci* 8(6):433–440
- Kretschmer M, Runge J, Coumou D (2017) Early prediction of extreme stratospheric polar vortex states based on causal precursors. *Geophys Res Lett* 44(16):8592–8600. <https://doi.org/10.1002/2017GL074696>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL074696>
- Li F, Orsolini Y, Keenlyside N, Shen ML, Counillon F, Wang Y (2019) Impact of snow initialization in subseasonal-to-seasonal winter forecasts with the Norwegian climate prediction model. *J Geophys Res Atmos*
- Limpasuvan V, Thompson DWJ, Hartmann DL (2004) The life cycle of the Northern Hemisphere sudden stratospheric warmings. *J Clim* 17:2584–2596
- Limpasuvan V, Hartmann DL, Thompson DWJ, Jeev K, Yung YL (2005) Stratosphere-troposphere evolution during polar vortex intensification. *J Geophys Res* 110. <https://doi.org/10.1029/2005JD006302>
- Masato G, Hoskins BJ, Woollings T (2013) Winter and summer northern hemisphere blocking in cmip5 models. *J Clim* 26(18):7044–7059
- Newman PA, Nash ER, Rosenfield JE (2001) What controls the temperature of the Arctic stratosphere during the spring. *J Geophys Res* 106:19999–20010
- Orsolini Y, Senan R, Balsamo G, Doblas-Reyes F, Vitart F, Weisheimer A, Carrasco A, Benestad R (2013) Impact of snow initialization on sub-seasonal forecasts. *Clim Dyn* 41(7–8):1969–1982
- Orsolini YJ, Senan R, Vitart F, Balsamo G, Weisheimer A, Doblas-Reyes FJ (2016) Influence of the eurasian snow on the negative north atlantic oscillation in subseasonal forecasts of the cold winter 2009/2010. *Clim Dyn* 47:1325–1334. <https://doi.org/10.1007/s00382-015-2903-8>
- Orsolini YJ, Nishii K, Nakamura H (2018) Duration and decay of arctic stratospheric vortex events in the ecmwf seasonal forecast model. *Q J R Meteorol Soc* 144(717):2876–2888
- Peings Y, Brun E, Mauvais V, Douville H (2013) How stationary is the relationship between siberian snow and arctic oscillation over the 20th century? *Geophys Res Lett* 40(1):183–188
- Peings Y, Douville H, Colin J, Martin DS, Magnusdottir G (2017) Snow-(n)ao teleconnection and its modulation by the quasi-biennial oscillation. *J Clim* 30(24):10211–10235
- Polvani LM, Kushner PJ (2002) Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. *Geophys Res Lett* 29. <https://doi.org/10.1029/2001GL014284>
- Polvani LM, Waugh DW (2004) Upward wave activity flux as a precursor to extreme stratospheric events and subsequent anomalous surface weather regimes. *J Clim* 17:3548–3554
- Rao J, Garfinkel CI, Chen H, White IP (2019) The 2019 new year stratospheric sudden warming and its real-time predictions in multiple s2s models. *J Geophys Res Atmos* 124. <https://doi.org/10.1029/2019JD030826>
- Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim GK, Bloom S, Chen J, Collins D, Conaty A, da Silva A, Gu W, Joiner J, Koster RD, Lucchesi R, Molod A, Owens T, Pawson S, Pegion P, Redder CR, Reichle R, Robertson FR, Ruddick AG, Sienkiewicz M, Woolen J (2011) MERRA: NASA's modern-era retrospective analysis for research and applications. *J Clim* 24:3624–3648. <https://doi.org/10.1175/JCLI-D-11-00015.1>
- Saito K, Cohen J, Entekhabi D (2001) Evolution of atmospheric response to early-season eurasian snow cover anomalies. *Mon Weather Rev* 129(11):2746–2760
- Schiemann R, Demory ME, Shaffrey LC, Strachan J, Vidale PL, Mizieliński MS, Roberts MJ, Matsueda M, Wehner MF, Jung T (2017) The resolution sensitivity of northern hemisphere blocking

- in four 25-km atmospheric global circulation models. *J Clim* 30(1):337–358
- Schwartz C, Garfinkel CI (2020) Troposphere-stratosphere coupling in subseasonal-to-seasonal models and its importance for a realistic extratropical response to the Madden-Julian oscillation. *J Geophys Res Atmos* 125:e2019JD032043. <https://doi.org/10.1029/2019JD032043>
- Sigmond M, Scinocca J, Kharin V, Shepherd T (2013) Enhanced seasonal forecast skill following stratospheric sudden warmings. *Nat Geosci* 6(2):98–102. <https://doi.org/10.1038/ngeo1698>
- Smith KL, Kushner PJ (2012) Linear interference and the initiation of extratropical stratosphere-troposphere interactions. *J Geophys Res (Atmospheres)* 117:D13107. <https://doi.org/10.1029/2012JD017587>
- Tripathi OP, Baldwin M, Charlton-Perez A, Charron M, Eckermann SD, Gerber E, Harrison RG, Jackson DR, Kim BM, Kuroda Y, Lang A, Mahmood S, Mizuta R, Roff G, Sigmond M, Son SW (2015a) The predictability of the extra-tropical stratosphere on monthly timescales and its impact on the skill of tropospheric forecasts. *Q J R Meteorol Soc* 141. <https://doi.org/10.1002/qj.2432>
- Tripathi OP, Charlton-Perez A, Sigmond M, Vitart F (2015b) Enhanced long-range forecast skill in boreal winter following stratospheric strong vortex conditions. *Environ Res Lett* 10(10):104007
- Tripathi OP, Baldwin M, Charlton-Perez A, Charron M, Cheung JC, Eckermann SD, Gerber E, Jackson DR, Kuroda Y, Lang A et al (2016) Examining the predictability of the stratospheric sudden warming of January 2013 using multiple nwp systems. *Mon Weather Rev* 144(5):1935–1960
- Vitart F, Ardilouze C, Bonet A, Brookshaw A, Chen M, Codorean C, Déqué M, Ferranti L, Fucile E, Fuentes M et al (2017) The sub-seasonal to seasonal prediction (s2s) project database. *Bull Am Meteorol Soc*. <https://doi.org/10.1175/BAMS-D-16-0017.1>
- White I, Garfinkel CI, Gerber EP, Jucker M, Aquila V, Oman LD (2019) The downward influence of sudden stratospheric warmings: association with tropospheric precursors. *J Clim* 32(1):85–108. <https://doi.org/10.1175/JCLI-D-18-0053.1>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.