Supplemental Information for "Forecast-based attribution of the role of stratospheric variability in temperature and snow extremes"

William J. M. Seviour[1], Justin Finkel[2], Philip Rupp[3], Regan Mudhar[1], Amy H. Butler[4], Chaim I. Garfinkel[5], Peter Hitchcock[6], Blanca Ayarzagüena[7], Dong-Chan Hong[8], Yu-Kyung Hyun[9], Hera Kim[10], Eun-Pa Lim[11], Daniel De Maeseneire[7], Gabriele Messori[12,13,14], Gerbrand Koren[15], Michael Sigmond[16], Isla R. Simpson[17], and Seok-Woo Son[8]

[1]Department of Mathematics and Statistics, University of Exeter, Exeter, UK
[2]Data Science Institute and Department of Geophysical Sciences, University of Chicago, Chicago, IL, USA
[3]Meteorological Institute Munich, Ludwig-Maximilians University, Munich, Germany
[4]NOAA Chemical Sciences Laboratory, Boulder, CO, USA
[5]Fredy and Nadine Herrmann Institute of Earth Sciences, Hebrew University, Jerusalem, Israel
[6]Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY, USA
[7]Department of Earth Physics and Astrophysics, Facultad de CC. Físicas, Universidad Complutense Madrid, Spain
[8]School of Earth and Environmental Sciences, Seoul National University, Seoul, Republic of Korea
[9]National Institute of Meteorological Sciences, Korea Meteorological Administration, Jeju, Republic of Korea
[10]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA
[11]Bureau of Meteorology, Melbourne, Australia
[12]Department of Earth Sciences, Uppsala University, Uppsala, Sweden
[13]Swedish Centre for Impacts of Climate Extremes (climes), Uppsala University, Uppsala, Sweden
[14]Department of Meteorology, Stockholm University, Stockholm, Sweden
[15]Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, Netherlands
[16]Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, British Columbia, Canada
[17]NSF National Center for Atmospheric Research, Boulder, CO, USA

Correspondence: William J. M. Seviour (w.seviour@exeter.ac.uk)

1 Extreme temperatures

In this section we include the figures analogous to Figs. 3, 4 and 5 of the main text, but for SSWJan19 (Figs. S1, S2, and S3) and SSWSep19 (Figs. S4, S5, and S6). Again, we use IFS as the representative model.

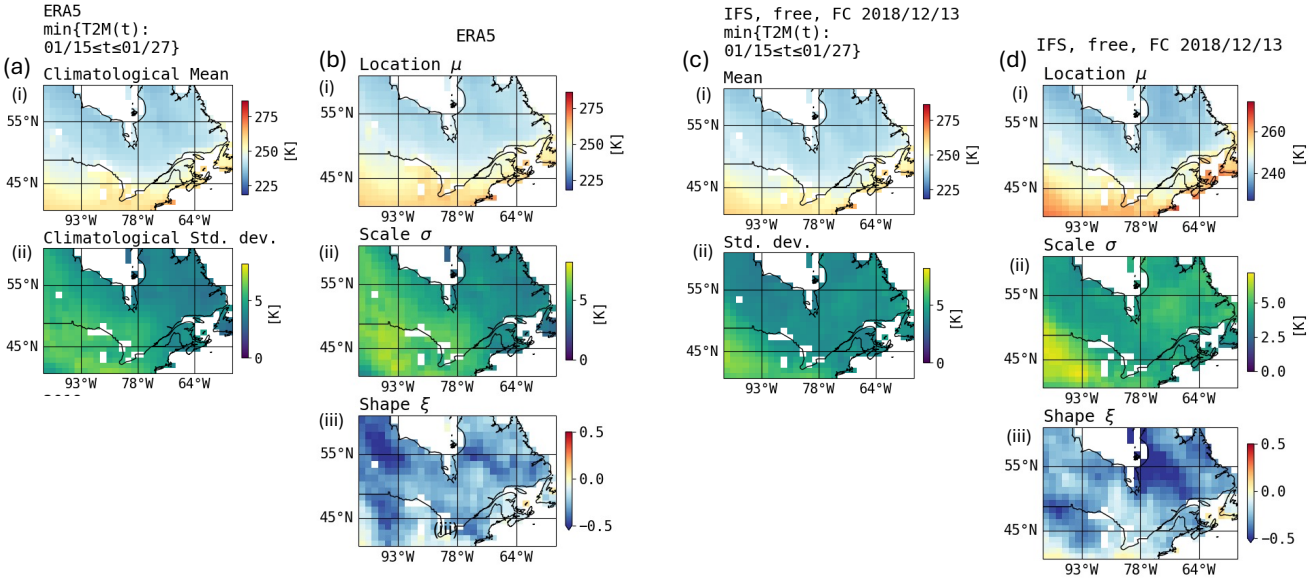

Figure S1. As Fig. 3 of the main text, but for SSWJan19. Statistical summary maps of surface temperature minima in the interval from 15-27 January. (a) Shows 'normal' statistics over the 'ensemble' of years from 1980–2020 in ERA5, (b) shows GEV statistics from ERA5, (c) and (d) show normal and GEV statistics respectively for the early initialization *free* experiment of the IFS model.

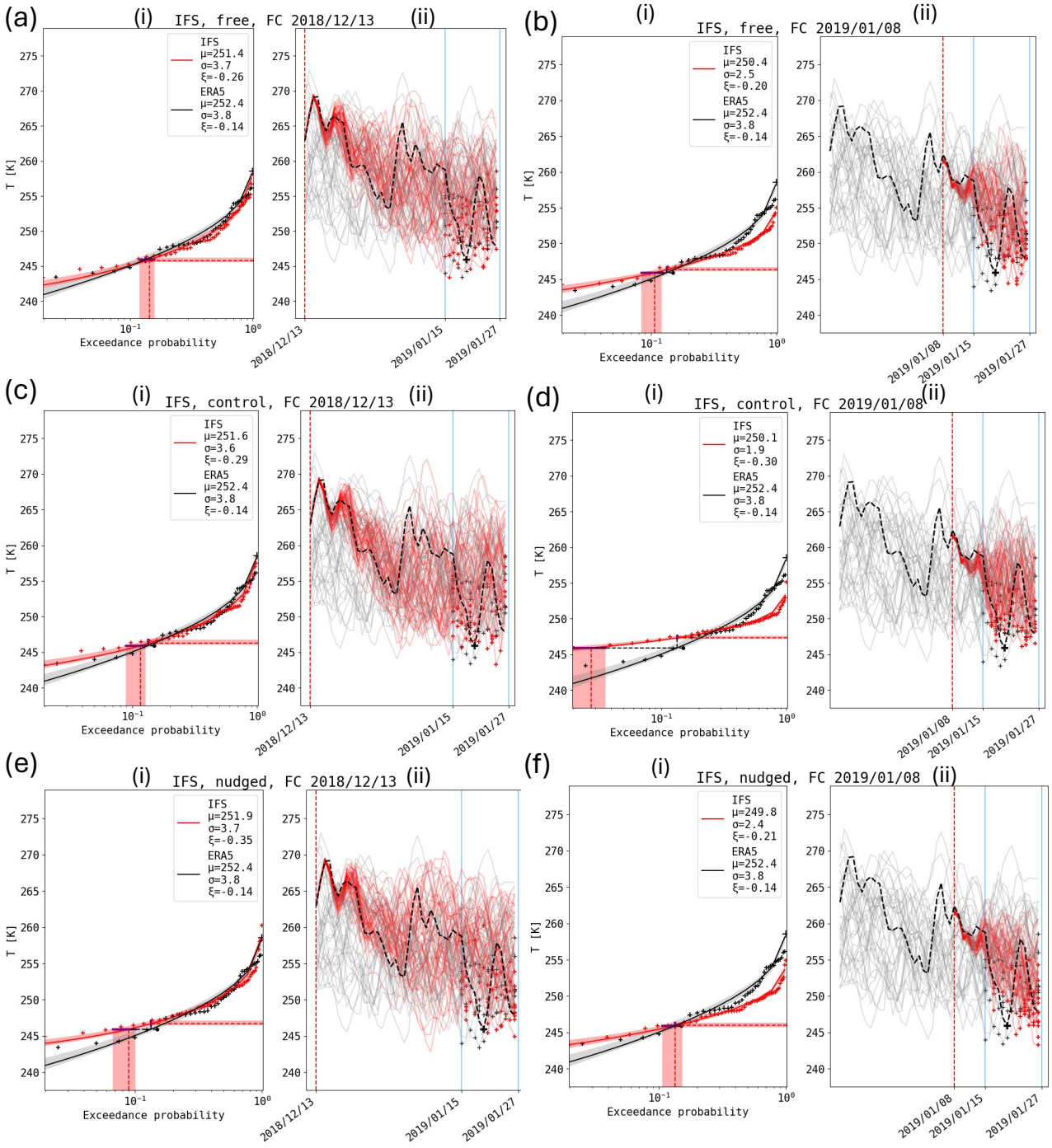

Figure S2. Same as Fig. 4 of the main text, but for SSWJan19. Shown are severity distributions (i) and timeseries (ii) of daily-minimum surface temperature, T, averaged over the full spatial region in Fig. S1. Left (a,c,e) and right (b,d,f) columns show early and late forecast dates. Top (a,b), middle (c,d), and bottom (e,f) rows show *free, nudged*, and *control* experiments.

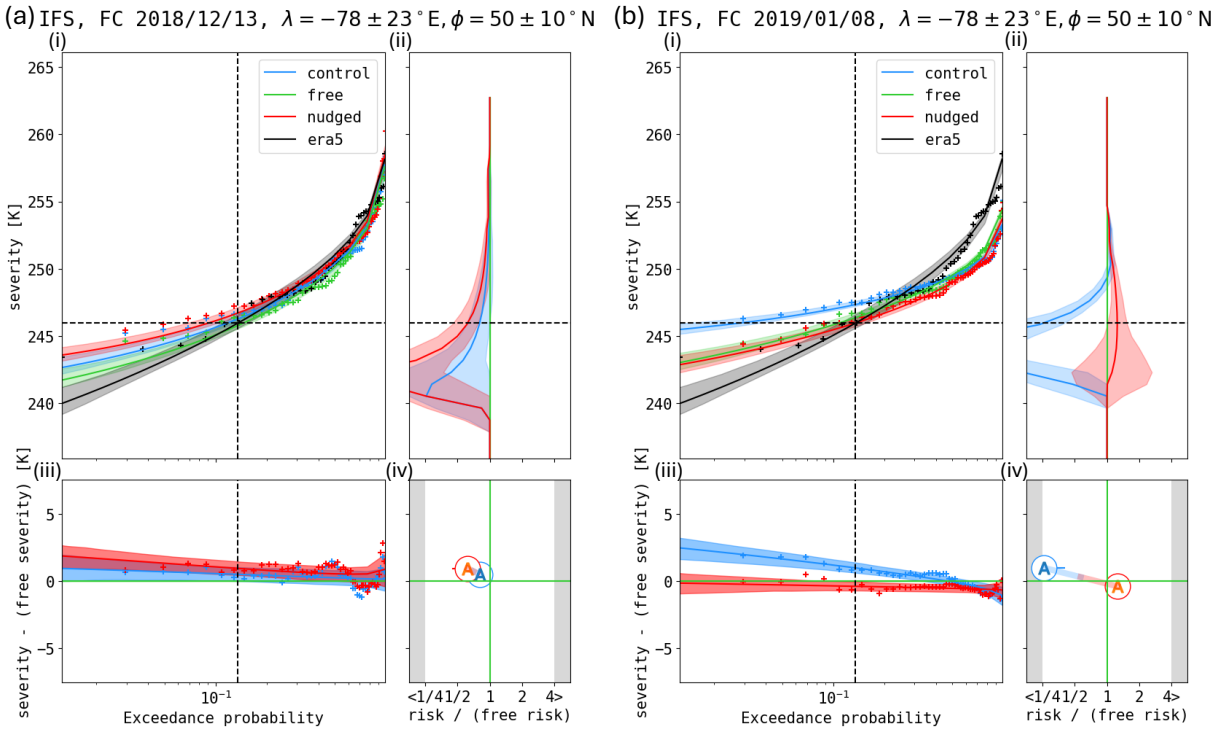

Figure S3. Same as Fig. 5 of the main text, but for SSWJan19. Summary of experimental effects according to IFS for SSWJan19, for both early (a) and late (b) initialization dates. Panels (i) show severities S^* vs. risk, meaning the probability of an event of even greater severity. Crosses represent empirical risks, solid lines represent GEV-fitted risks, and shaded bands represent 50% confidence intervals of GEV fits based on 1000 bootstrap resamplings. Panels (ii) show RR of the *control* and *nudged* distributions relative to *free*, as a function of severity. Panels (iii) show QSs of the *control* and *nudged* distributions relative to the *free*, as a function of risk, as well as empirical differences (available only when the two ensembles are equal in size). In other words, panels (ii,iii) respectively show horizontal and vertical differences between curves in panels (i). In panels (iv), we summarize each response as a single ordered pair (relative risk, quantile shift), where the reference severity is that observed in ERA5 and the reference quantile is that fitted from the ERA5 GEV at its reference severity. Any ratios $< \frac{1}{4}$ or > 4 would be clipped to the gray shading at the margins.

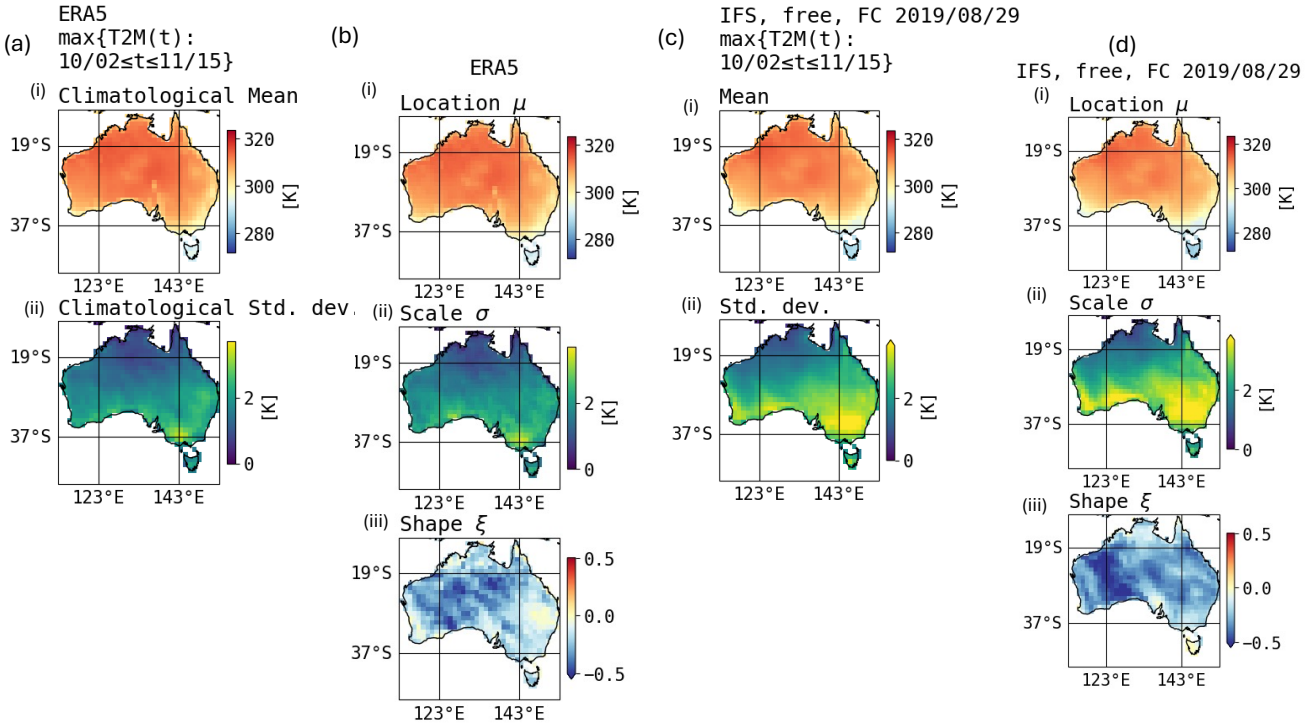

Figure S4. As Fig. S1 and Fig. 3 of the main text, but for SSWSep19, for which temperature maxima are condsidered over the period 2 October-15 November. (a) and (b) show ERA5 'normal' and GEV statistics respectively, (c) and (d) show normal and GEV statistics from the early initialization *free* IFS experiment.

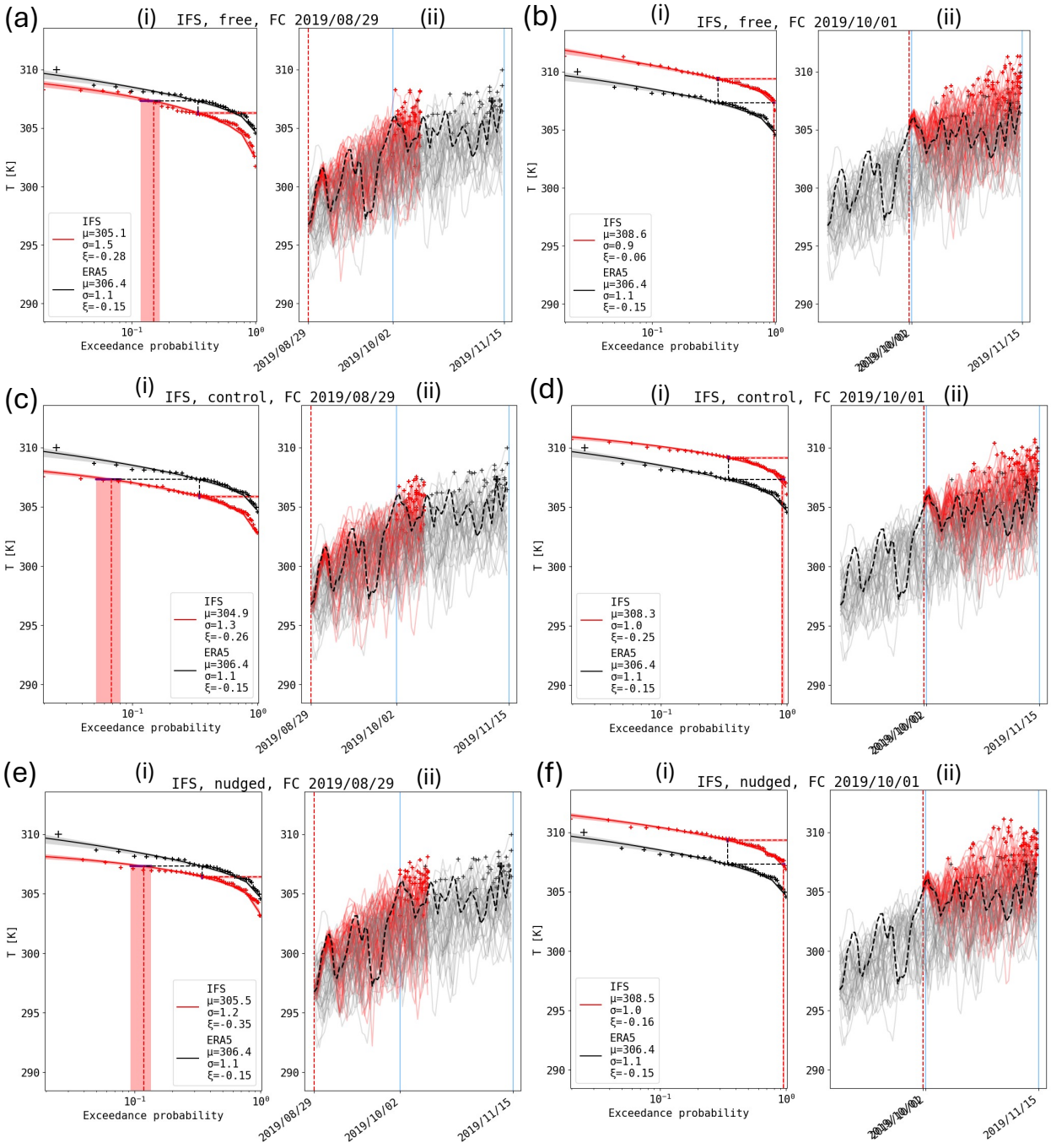

Figure S5. Same as Fig. 4 of the main text, but for SSWSep19. Shown are severity distributions (i) and timeseries (ii) of daily-maximum surface temperature, T, averaged over the full spatial region in Fig. S4. Left (a,c,e) and right (b,d,f) columns show early and late forecast dates. Top (a,b), middle (c,d), and bottom (e,f) rows show *free, nudged*, and *control* experiments.

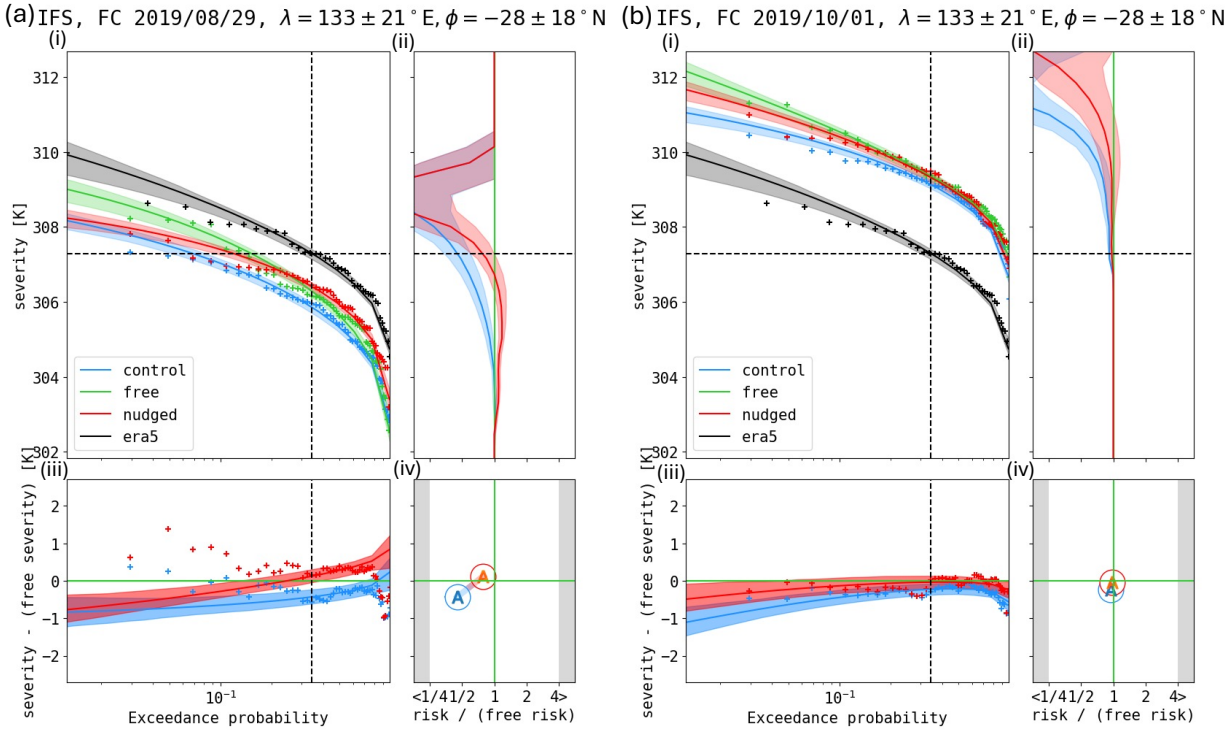

Figure S6. Same as Fig. S3 and Fig. 5 of the main text but for SSWSep19. Note that temperature maxima are considered in this case.

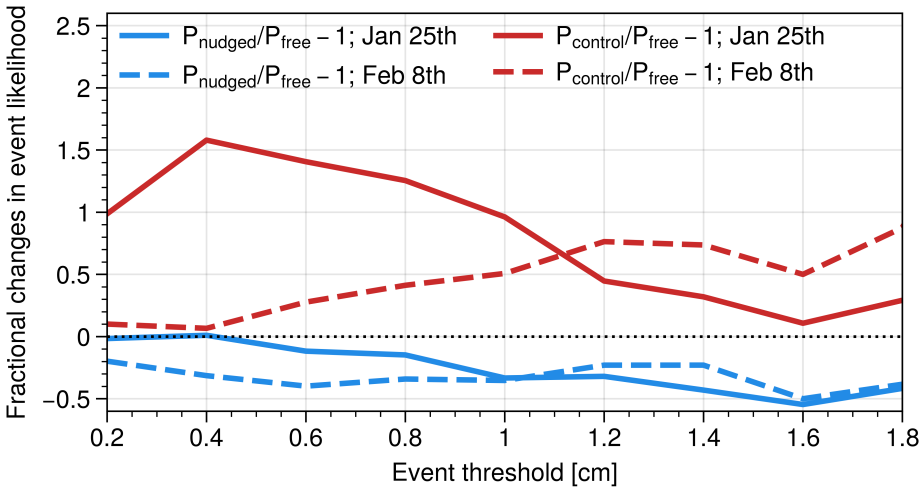

Figure S7. Multi-model mean RR of southern UK and Ireland snow, as Fig. 9 of the main text, but for varying snow depth thresholds. Shown for both early and late initializations of SSWFeb18.

2 Extreme snow

Figure S7 illustrates the sensitivity of the multi-model mean RR for snow accumulation (Fig. 9 of the main text) to the selection of threshold snow depth (0.8 cm shown in the main text.)