

RESEARCH ARTICLE OPEN ACCESS

The Influence of Sudden Stratospheric Warmings on Extreme European Electricity Demand

Regan Mudhar¹ | William J. M. Seviour¹  | Hannah C. Bloomfield²  | Amy H. Butler³ | Chaim I. Garfinkel⁴ | Peter Hitchcock⁵  | James A. Screen¹ | Stephen I. Thomson¹

¹Department of Mathematics and Statistics, University of Exeter, Exeter, UK | ²School of Engineering, University of Newcastle, Newcastle upon Tyne, UK | ³NOAA Chemical Sciences Laboratory, Boulder, Colorado, USA | ⁴Fredy and Nadine Herrmann Institute of Earth Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel | ⁵Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, New York, USA

Correspondence: Regan Mudhar (regan.mudhar@unil.ch)

Received: 23 February 2026 | **Revised:** 26 June 2026 | **Accepted:** 2 September 2026

Keywords: energy sector | extremes | predictability | subseasonal | sudden stratospheric warmings

ABSTRACT

Sudden stratospheric warmings (SSWs) provide a source of surface predictability up to several weeks in advance. They generally precede cold weather across the northern midlatitudes, including extremes. When it is cold, demand for heating typically increases electricity demand. We use reanalysis data and a weather-driven demand model for various European countries to investigate the relationship between SSWs and extreme electricity demand. We find that extreme demand of ~120% of climatology is 1.5–3 times more likely in periods following SSWs across northern and central Europe. We then use subseasonal forecast experiments from the Stratospheric Nudging and Predictable Surface Impacts project to explicitly attribute the impact of the February 2018 SSW, an event that preceded a cold weather outbreak across Europe. These experiments include nudging of the stratospheric state to either its observed or climatological evolution, allowing the causal impact of the SSW to be identified. We find a robust simulated increase in extreme demand risk attributed to the 2018 event. This indicates a window of opportunity for earlier prediction of extreme European electricity demand risks following SSWs.

1 | Introduction

Many sectors of society increasingly seek skilful predictions of weather and its impacts beyond the typical timescales of weather forecasts (White et al. 2017). For a forecast of conditions weeks to months ahead, there are only a few sources of predictability within the climate system. Along with soil moisture and the Madden-Julian Oscillation, a key one is the stratosphere; the layer of the atmosphere ~10–50 km above Earth's surface (Butchart 2022). The stratosphere varies more slowly than the troposphere but can have a lasting influence on surface weather, offering the opportunity to predict tropospheric changes up to weeks in advance (Domeisen et al. 2020a, 2020b; Scaife et al. 2022). A key component of wintertime stratospheric circulation is the polar vortex (Waugh and Polvani 2010). Persistent anomalies in vortex strength are often associated with robust,

and similarly persistent, regional changes in midlatitude surface weather (Domeisen and Butler 2020); though not all stratospheric perturbations necessarily couple to the surface (Sigmond et al. 2013; Karpechko et al. 2017; Oehrlein et al. 2021).

Extreme vortex weakening events, known as sudden stratospheric warmings (SSWs), are dramatic wave-driven breakdowns of the polar vortex. In these events, the otherwise strong and stable westerly vortex winds temporarily reverse, with temperatures inside warming by tens of degrees over just a few days (Baldwin et al. 2021). The downward influence of SSWs on weather and climate can then be felt for up to 30–60 days after onset (Kidston et al. 2015; Domeisen and Butler 2020; Baldwin et al. 2021), potentially providing predictive power on subseasonal-to-seasonal (S2S) timescales of 2 weeks–2 months (Vitart et al. 2017). In the Northern Hemisphere, SSWs occur

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Meteorological Applications* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

on average around every other year and are usually associated with the negative phase of the North Atlantic Oscillation (NAO; Baldwin and Dunkerton 2001). This corresponds to lower than average surface temperatures, including an increase in the likelihood and severity of cold air outbreaks and cold extremes across the northern midlatitudes (Domeisen and Butler 2020; Huang et al. 2021; Spaeth and Birner 2022; Seviour et al. 2026).

Predictions of surface conditions at a range of lead times can provide important information to many different end-users (White et al. 2017; Vitart and Robertson 2018). For example, energy system operators are known to utilise weather forecasts to prepare for potentially impactful events; they can also assess the risks associated with extreme events further in advance by utilising skilful subseasonal forecasts (Soares and Dessai 2016; White et al. 2017; Goutham et al. 2022). Though the energy sector is interested in impacts on the system itself at the short range especially (e.g., damage from extreme weather), the predominant focus for extended range predictions is on renewable generation and demand; this can be leveraged by both the system operators and those in the energy trading industry (S2S4E 2020). Demand is primarily temperature-driven; during European winters, low temperatures drive increases in heating, and thus electricity and gas, demand (Taylor and Buizza 2003; Thornton et al. 2016). Despite its robust electricity network, periods of high electricity demand and low renewable (e.g., wind power and/or solar) generation still present a challenge in Europe. If supply cannot keep up with demand, countries may turn to other sources that are more expensive and/or polluting (van der Wiel et al. 2019)—or even leave millions of customers without electricity in very extreme scenarios, such as the 2021 cold outbreak in Texas (Busby et al. 2021; Doss-Gollin et al. 2021).

There are a number of studies that generally utilise S2S forecast model data for the energy sector. They find forecast skill in energy variables from lead times of up to 2–4 weeks in advance, indicating windows of opportunity for making useful forecasts (e.g., Clark et al. 2017; Soret et al. 2019; Dorrington et al. 2020; Bloomfield et al. 2021; Goutham et al. 2022). There are other studies that use S2S models to study SSWs, their predictability, and the predictability of their impacts (e.g., Rao et al. 2021; Kolstad et al. 2022; Finkel et al. 2023; Ineson et al. 2024; Lee et al. 2025). Furthermore, there is existing work linking SSWs to cold weather-driven societal impacts (such as mortality; Charlton-Perez et al. 2021) and linking weak/strong stratospheric extremes to energy-relevant variables at the subseasonal timescale (Büeler et al. 2020; Beerli et al. 2017). However, there are very few published studies that specifically link SSWs to energy impacts. One recent study examined SSWs' effects on wintertime electricity consumption and wind power generation in Finland, asking whether such effects could be used for seasonal prediction up to 6 months ahead using a geomagnetic activity-based SSW-prediction model (Juntunen, Vokhmyanin, et al. 2025). Our study alternatively generalises across European regions, and focuses on a shorter timescale (sub-seasonal), so is the first to use S2S prediction systems specifically to link SSWs and energy impacts.

As weather-driven impacts on both natural and human systems can vary along with changes in weather and climate, it can

be helpful to understand whether such changes are a result of natural variability or some predictable element of the climate system. This is the realm of 'detection and attribution' studies. These seek to identify (i.e., *detect*) a trend, pattern or change in climate-related variables or systems and then *attribute* it to some specific driver—usually anthropogenic climate change (Hegerl and Zwiers 2011; Otto 2017; Zhai et al. 2018). Attribution studies of specific events, extremes, or impacts typically either take the 'risk-based' approach, which considers classes of events, or the 'storyline' approach, which focuses on a specific event. Storyline studies aim to quantify how observed events could have unfolded in alternative ways by 'conditioning' on specific drivers of extreme events (e.g., Cattiaux et al. 2010; Chan et al. 2022). Though they do not necessarily assess changes in event likelihood as a result, they can be used to investigate how much (less) worse events could have been given changes to their characteristics and/or drivers (Lin et al. 2020; Lloyd and Shepherd 2020). The risk-based approach alternatively tends to calculate changes in likelihood of classes of events, defined by a threshold in a climate variable or an impact, recognising that extreme events can have multiple external drivers (natural or anthropogenic) on top of the system's internal noise, and that human activities can influence the likelihood of such events occurring. Such studies assess extreme events in a representation of the 'factual' world compared to what they might have been under 'counterfactual' climate conditions; this usually relies on models, with the assumption that they can reliably simulate the event (Otto 2017).

Impact attribution can be particularly challenging. It often requires the synthesis of interdisciplinary information that differs in terminology and type, and/or having to interface between models built by research communities with differing priorities. As a result, such studies are not very common (Mudhar et al. 2023). Nevertheless, the risk-based approach has been used to attribute the decreasing trend in domestic energy consumption in the United Kingdom to human-induced climate change (Christidis et al. 2021). However, there is no focus on specific energy impact 'events' so far in the literature.

Furthermore, though attribution of extreme weather events has rapidly developed since the 2010s, there are still concerns over the robustness of the traditional climate model approach. Namely, that such models can struggle to represent key dynamical drivers of extremes (Leach et al. 2021). In recent years, forecast-based attribution has been presented as a framework to synthesise the risk-based and storyline approaches (Leach et al. 2024; Thompson et al. 2025). With a forecast you can specifically condition on the component of an event that was predictable at the time of initialisation; the benefits of using a successful forecast are not only that the model is able to accurately represent the event, but also that the analysis is unquestionably an attribution of that specific event (as opposed to a mixture of many events in some 'class'). Forecast models also typically have greater resolution than climate models and can be run as large ensembles, increasing confidence.

Given the subseasonal predictability afforded by SSWs, we can ask a similar question to that of traditional attribution studies about the attributable impact of such stratospheric events on surface extremes. Unfortunately, fully realising the skill associated with SSWs in operational S2S forecast models is

still challenging. These models must capture both the evolution of the stratospheric polar vortex and the coupling to the troposphere: well-known sources of significant model bias (Lawrence et al. 2022; Garfinkel et al. 2024). Stratospheric Nudging and Predictable Surface Impacts (SNAPSI) is a model intercomparison protocol that uses a set of controlled, subseasonal, ensemble forecasts of three case study vortex weakening events to address a number of science goals. One of these is ‘to attribute specific extreme events to stratospheric variability’ (Hitchcock et al. 2022). To do this, the protocol utilises an artificial nudging in the stratosphere (described further in Section 2.2) to isolate the role of stratospheric variability on predictions from two initialisations per case study. Analysis of the SNAPSI experiments has so far revealed that negative NAO-like regimes systematically follow the simulated Northern Hemisphere SSWs of 2018 and 2019 (Lee et al. 2025), and that they increase the likelihood of observed severe cold spells and extreme snow (Seviour et al. 2026). Here, we aim to take this a step further, linking to an impact model. We investigate two key questions relating to the attributable impact of SSWs on the European energy sector:

1. What is the average historical relationship between SSWs and electricity demand?
2. How much of the risk of a given extreme demand event can be attributed to the occurrence of an SSW?

The first of these aims to understand the links between SSWs and energy impacts, for which reanalysis is used, while the second is more about causation, which can be better understood using the SNAPSI simulations of the 2018 SSW.

We describe the data and analysis methods used in Section 2, before presenting the results answering the questions above in Sections 3 and 4, respectively. Finally, in Section 5 we discuss our findings, including the limitations and implications of the novel analysis presented here.

2 | Data and Methods

In this work, we analyse electricity demand in the periods following SSWs, derived from daily mean 2 m temperature from the reanalysis and SNAPSI datasets. We describe the data, demand model, and extreme event identification and attribution methods used in the following sections.

2.1 | Reanalysis Data

We use ERA5 reanalysis data (Hersbach et al. 2020) covering extended Northern Hemisphere winter periods of October–May from 1979/80 to 2024/25. Specifically, daily-mean zonal-mean zonal wind at 10 hPa, 60° N ($U_{10,60}$) and daily-mean 2 m temperature (T), with the latter regridded to a horizontal resolution of 1° using a distance-weighted average remapping (closer to the typical resolution of SNAPSI-participating models later described in Section 2.2). We use the recent period rather than one extending earlier into the twentieth century, as 1979 marks the introduction of operational satellite sounder data, which led to

improvements in the accuracy of stratospheric representation in reanalysis. Not only is the warming trend in T not as strong than if we extended back to the 1950s, but using the recent period means excluding some of the very cold winters in the early 1960s; it is unclear how likely such winters would be to occur in our current climate, which is a question beyond the scope of the present study.

We use $U_{10,60}$ to identify SSWs following the Charlton and Polvani (2007) approach for ‘major’ warmings: we look for reversals in the sign of $U_{10,60}$ over October–April (inclusive), requiring that winds must return to westerly for at least 20 consecutive days between reversals and that they return to westerly for at least 10 consecutive days before 30 April. This gives us 30 SSWs in the 46 winters of the ERA5 dataset used here, with onset dates spread throughout December–March. The list of onset dates is shown along with the evolution of $U_{10,60}$ in the 10 days before and 45 days after each SSW in Figure 1. With the method explained in Section 2.3, we use these onset dates to analyse the T -derived electricity demand in the 45 days following SSWs (the ‘post-SSW’ period; Mitchell et al. 2013). As the demand model we use was trained using ERA5 T , we also use this reanalysis dataset for consistency. For the use of ERA5 $U_{10,60}$ to identify SSWs, we find onset dates that are broadly consistent with other reanalysis products (as per Butler et al. 2017). Note that we use December–May for the near-surface variables due to the SSW timings and length of the post-SSW period, and that there are no winters in which 2 SSWs occur within 45 days of each other.

2.2 | SNAPSI Protocol

SNAPSI uses S2S forecast models to run retrospective forecasts, or ‘hindcasts’, of 3 case study stratospheric polar vortex weakening events (12 February 2018 and 2 January 2019 in the Northern Hemisphere, and 18 September 2019 in the Southern Hemisphere). For each of the case studies, all of the models run (minimum) 45-day simulations with a (minimum) 50-member ensemble initialised from 2 different start dates for 3 experimental set-ups. A detailed description of the observed dynamical evolution of the case study events and SNAPSI experimental design is given by Hitchcock et al. (2022). Here we present the elements of most relevance for the research questions. First, the key experiments:

- **free**: after initialisation, the ensemble is able to freely evolve—that is, a standard forecast;
- **nudged**: throughout the run, the zonally-symmetric stratospheric state is globally nudged to the ERA5 time-evolution of the stratospheric event;
- **control**: throughout the run, the zonally-symmetric stratospheric state is globally nudged to a time-evolving climatological state calculated from ERA5.

For the nudged and control experiments, SNAPSI leverages a zonally-symmetric nudging technique previously used to identify the influence of the stratosphere on tropospheric circulation (Kautz et al. 2020). The zonal-mean zonal wind and temperature are nudged above 90 hPa, ramping to full strength by 50 hPa

via a cubic profile that is applied equally at all latitudes. For the `control` experiment, the climatological reference state is gradually introduced over the first 5 days to mitigate any spurious initialisation responses. The experimental set-up allows the troposphere to freely evolve in response to the initialisation and nudging, if used; thus, even a perfect simulation of the zonal-mean stratosphere does not mean models will perfectly capture surface events. Nevertheless, we view comparing the surface conditions of the `nudged` to `free` simulations as revealing the impact of an improved stratospheric forecast, while comparing `nudged` to `control` isolates the influence of the anomalous (zonal-mean) stratospheric evolution during SSWs on the surface.

We use daily T from 7 of the SNAPSI-participating models (Table 1) to derive and analyse post-SSW electricity demand, selected based on data availability. Our case study of interest is the Northern Hemisphere SSW on 12 February 2018, which was followed by notable low near-surface temperature anomalies across northern and central Europe, both on average (Figure 2a) and in terms of extreme low temperatures

(Figure 2b), including heavy snow in some regions (Galvin et al. 2018; Lü et al. 2020). Though this event is not considered one of the most extreme energy events across Europe (Bloomfield et al. 2020), there were socioeconomic impacts in certain countries at the time: for example, U.K electricity prices peaked at just over $\pounds 124\text{MWh}^{-1}$ on 3–4 March 2018, almost three times the typical price in the 2010s (Drax Group 2025).

Figure 3a shows the simulated evolution of $U_{10,60}$ for the ensemble mean of all models, and Figure 3b for all ensemble members of the IFS model, compared to that in ERA5, with initialisation dates of 25 January and 8 February 2018, respectively 18 and 4 days before the SSW onset. As expected, for all models and IFS ensemble members, `nudged` closely follows the ERA5 evolution, `control` quickly diverges from it, and there is considerable spread across the `free` experiments: the free-running simulations much more closely follow ERA5 after the later initialisation when the SSW onset is better captured. Note that the minimum 45-day length of the simulations means there is a maximum period of 27 days post-SSW (grey shaded region on Figure 3) that is common to both sets of initialised simulations

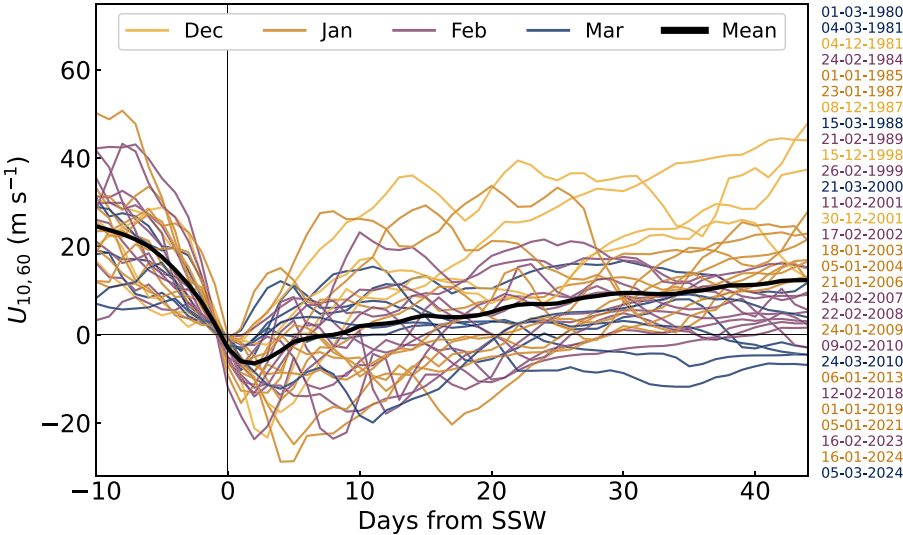


FIGURE 1 | Evolution of the daily-mean zonal-mean zonal wind at 10 hPa, 60°N for each of the 30 SSWs identified from ERA5, with onset dates listed on the right (broadly consistent with Butler et al. 2017). Different colour lines and dates indicate the month of each SSW’s onset date, and the thick black line is the average evolution.

TABLE 1 | The 7 forecast systems analysed. The 2 versions of GloSea6 use the same perturbation strategy but differ in land surface scheme and initialisation.

Model centre	Model name	Spatial resolution	Vertical levels	Model top
CCCma	CanESM5	T63	49	1 hPa
NCAR	CESM2-CAM6	$1.25^{\circ} \times 0.9^{\circ}$	32	2.26 hPa
Météo France	CNRM-CM6-1	TL359	91	0.01 hPa
UKMO	GloSea6-UKMO	N216	85	85 km
KMA	GloSea6-KMA	N216	85	85 km
SNU	GRIMs	T126	60	0.3 hPa
ECMWF	IFS	TCO319	75	0.01 hPa

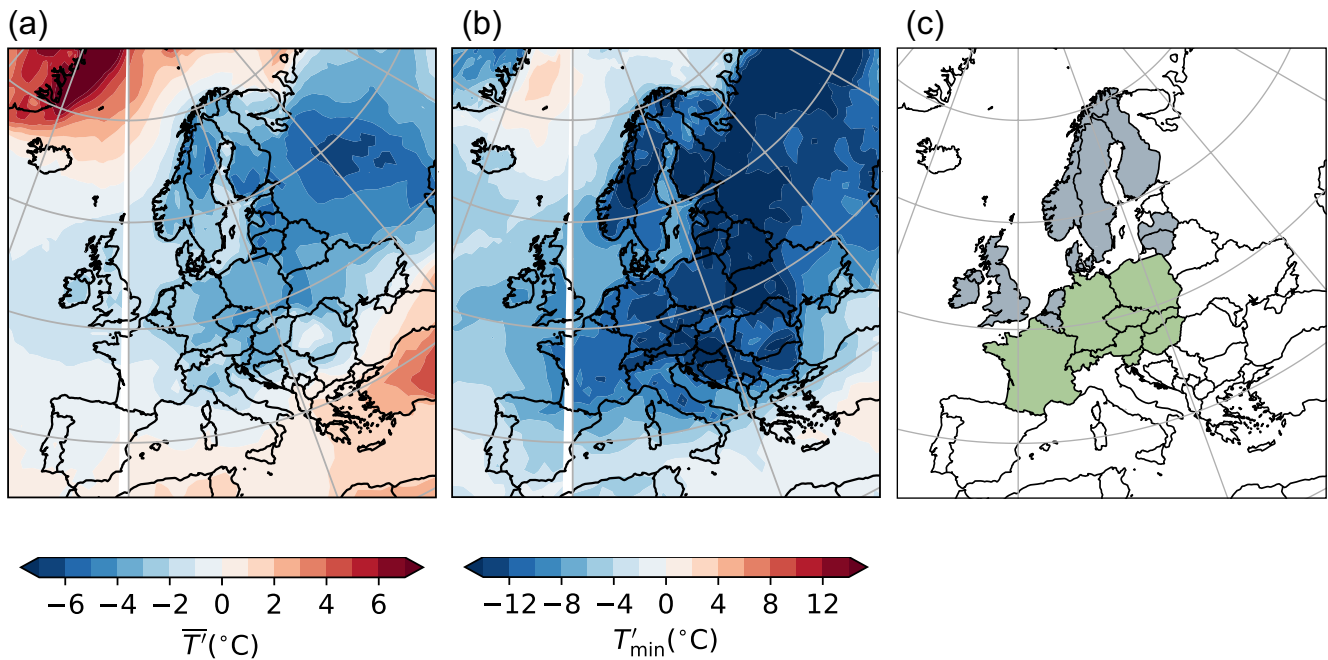


FIGURE 2 | The ERA5-derived (a) mean and (b) minimum temperature anomalies in the 27 days following the 12 February 2018 SSW, calculated for each grid point as described in Section 2.1 (note the different colour scales). Plus, the (c) northern (blue) and central (green) European countries as in Table 2.

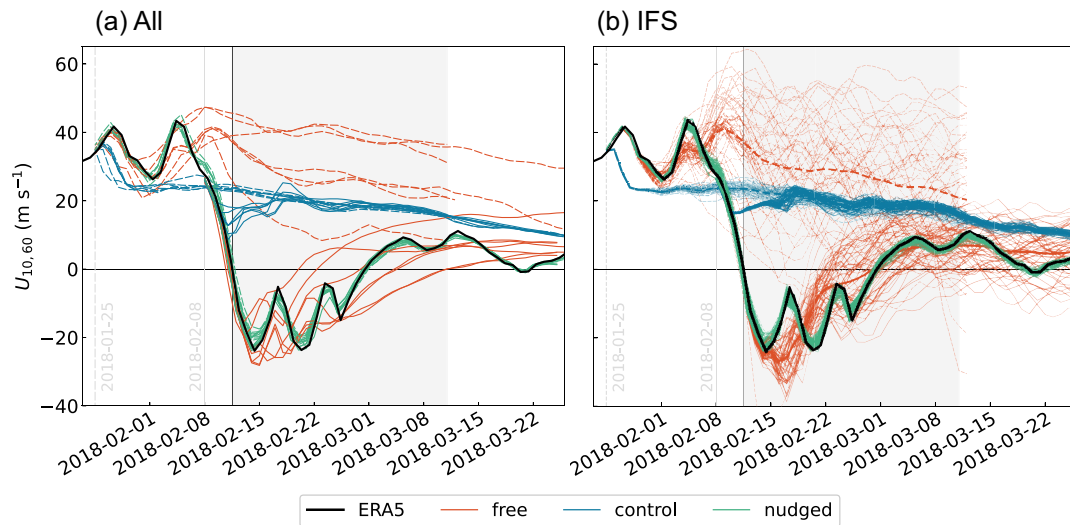


FIGURE 3 | Evolution of the 10hPa, 60°N zonal-mean zonal wind in the *free*, *control* and *nudged* SNAPSI simulations of the 12 February 2018 SSW. Lines are dashed for the first initialisation date and solid for the second, with the vertical grey lines indicating the initialisations, the vertical black line the SSW onset, and the grey shaded period the 27 days post-SSW. (a) each coloured line is the ensemble mean of one of the seven models. (b) each thin line is an ensemble member and the thick line the ensemble mean for the IFS model.

across all the models, which will be the focus of the SNAPSI-related analysis.

2.3 | Demand Model

We use the multiple linear regression model constructed for 28 European countries by Bloomfield et al. (2021) to derive country-aggregated daily electricity demand (E in GW) from daily-mean

2m temperature (T). The training of the multiple linear regression model is described in detail in Bloomfield et al. (2021); the model for each country contains weather-related parameters as well as confounding human behavioural factors, such as ‘day-of-week’ effects (whereby demand is greater on weekdays than weekends) and long-term socioeconomic trends. Here, we choose to treat every day as ‘Monday’ and disregard the long-term changes to enable us to focus on the weather-driven components of demand. Thus, for the winter,

TABLE 2 | European countries per country group in Figure 2c.

Northern Europe	Central Europe
Belgium	Austria
Denmark	Czech Republic (C.R.)
Finland	France
Ireland	Germany
Latvia	Hungary
Lithuania	Luxembourg
Netherlands	Poland
Norway	Slovakia
Sweden	Slovenia
United Kingdom (U.K.)	Switzerland

$$E(t) = \beta_0 + \beta_1 \text{HDD}(t), \quad (1)$$

where t is the daily time step and β_i are the regression coefficients, which are calculated separately for each country: the first term is the constant ‘background’ level of demand (in 2017, as per Bloomfield et al. 2021). Note that a recent retrain of the model on 2021 data modified β_0 slightly, but the peaks in winter demand (as is relevant for this work) are comparable and the skill of the model, of ~2 weeks, was not impacted. The weather-driven electricity demand then relies on variations in ‘heating degree days’ (HDDs); this links E not to some absolute T , but rather the difference between T and a base temperature T_b , below which demand becomes weather-sensitive. Thus, if $T(t) < T_b$ then $\text{HDD}(t) = T_b - T(t)$, else $\text{HDD}(t) = 0$. The temperature difference is key, as the lower the temperature is outside, the more heating, and thus energy, required. Along with the sensitivity of each country’s E to HDDs, T_b typically varies by country depending on their building stock and populations’ resiliency to cold; here, it is set constant at 15.5°C for simplicity, consistent with the European Environment Agency (EEA 2024) and other similar models. The model parameters already estimated from the 2016/17 training period by Bloomfield et al. (2021) can then be used to derive daily demand from T for any period and, in theory, any dataset.

We apply the weather-driven demand model to the ERA5 and SNAPSI datasets, and analyse both mean and extremes of demand following SSWs for each of the relevant countries. For this, we aggregate the daily T timeseries at each grid point within individual countries to compute the national average daily T , from which we calculate HDDs and thus the E timeseries per country. The European Network of Transmission System Operators for Electricity (ENTSO-E) contains a number of regional groups of countries whose own individual national regulated entities operating the high-voltage electric grid can operate synchronously (Foretic et al. 2024). The largest of these is ‘Continental Europe’ (28 countries), with smaller groups of Baltic countries (3), Nordic countries (4), and Great Britain/Ireland (2). As a demand model is not available for all countries (e.g., Estonia is missing from the Baltic group due

to training data availability), we collate the smaller groups to form a ‘northern’ group and separate a subset of the continental countries into a ‘central’ group (Table 2, Figure 2c). Countries in southern Europe are not considered, as the cold-related impacts of SSWs are typically less severe there. Our focus is predominantly on these two aggregated groups of countries, as many of the ENTSO-E countries are interconnected. We thus sum demand over all individual countries within the group, which also has the effect of smoothing out the signal. Figure 4 shows the average daily evolution of ERA5-derived electricity demand in both northern and central Europe over all 46 winters, with 2017/18 highlighted to show the weather-driven demand reaching a new maximum in the weeks following the February SSW.

2.4 | Inferring the Impact of SSWs

We follow a similar analytical framework as in Seviour et al. (2026) to investigate whether there is an attributable signal of SSWs in extreme electricity demand events. This requires a comparison of demand under post-SSW and ‘non-SSW’ conditions in order to quantify the statistical difference between these two states.

Consider some extreme event A , which generally has a probability of occurring $p_0 = p(A)$. The probability of such an event occurring following an SSW, S , is then $p_1 = p(A|S)$. The probability, or risk, of A occurring after an SSW can be framed *relative* to the risk of it occurring generally, using ‘relative risk’ $RR = p_1 / p_0$. RR thus describes the fractional change in risk for a given magnitude and is a measure commonly used in climate change attribution studies (Otto 2017); we use it here to attribute the occurrence of specific extremes to the occurrence of an SSW. As RR frames the attribution result in terms of relative probabilities, it is applicable for both increases and decreases in event likelihood (Wilcox et al. 2018). $RR > 1$ indicates that the occurrence of the SSW makes the extreme more likely to happen, and the greater the RR , the greater the role of the attributed driver; for example, $RR = 0.5$ or 2 would suggest that the extreme is half or twice as likely following the SSW, respectively. While RR is often preferred due to its relative ease of interpretability, it should be noted that it depends on event definition, and that a change in the risk of an event may appear large but only correspond to a marginal change in magnitude, such that the real-life impact is relatively small. Hence, we additionally find the change in magnitude for a given risk, termed the ‘quantile shift’ (QS), to highlight the difference in severity at some exceedance probability—for example, how much more severe does a 1 in 10 event become following an SSW?

2.4.1 | Risk Analysis With Reanalysis

With ERA5, we are interested in both post-SSW mean and extreme anomalies. We remove 30 June from each leap year then compute daily anomalies from the climatology calculated as the multi-year (1979/80–2024/25) mean of each ‘day-of-year’;

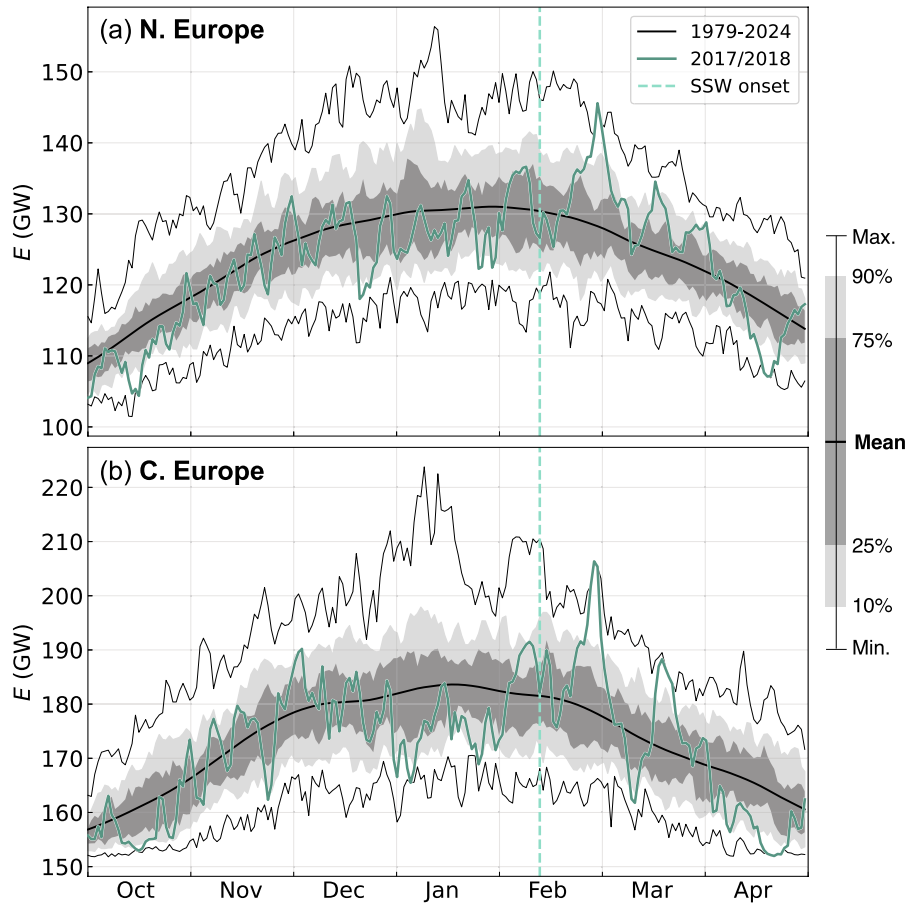


FIGURE 4 | Evolution of ERA5-derived E aggregated over (a) northern and (b) central European countries. Black lines and shaded regions are from all years in 1979/80–2024/25, with a 30-day smoothing applied to the mean, while the green line picks out 2017/18 and the vertical dashed line marks the onset of the 12 February 2018 SSW. Note the difference in the y-axes.

29 February in a leap year is treated as 1 March. Note that we do this conversion of absolute values of temperature and demand into anomalies to better enable comparison between countries and country groups. For mean temperature anomalies, we compute a daily timeseries of anomalies (T') which are then averaged over the post-SSW period to give $\bar{T}'$ (in $^{\circ}\text{C}$). For the extremes, the lowest (minimum) temperature on a single day in the post-SSW period is subtracted from the equivalent day-of-year climatology to calculate the anomaly, $T'_{\min}$. For mean electricity demand anomalies, we calculate the timeseries of anomalies as $E' = (E/E_{\text{clim}}) - 100$ then again average over the post-SSW period to calculate $\bar{E}'$ (in %). For the extremes, the highest (maximum) demand value on a single day in the period is similarly used to calculate the anomaly relative to the equivalent day-of-year climatology as $E'_{\max} = (E_{\max}/E_{\text{clim}}) - 100$.

For the attribution specifically, we focus on $E'_{\max}$ as a way of indicating a ‘most challenging day’ for energy system operators. To analyse extreme event risks, we treat the 45 days after each of the 30 SSWs (onset date inclusive) as our post-SSW periods. We find $E'_{\max}$ after each of the 30 SSWs using the method described above, generating a list of 30 extreme demand values. We then sort these from smallest to largest, in order of highest to lowest

probability of occurring, and fit a generalised extreme value (GEV) distribution (described in detail in Seviour et al. 2026). Each of the values then has an ‘exceedance probability’ associated with it: for example, the most extreme value, or ‘severity’, in the list of 30 has a 1 in 30 probability of occurring, the 10th most extreme a 10 in 30, and so on. Though the use of the GEV may not traditionally be considered appropriate (the variables used are not guaranteed to be ‘large’, ‘independent’ or ‘identically distributed’), it overall captures the distributions here much better than a Gaussian: thus, as in Seviour et al. (2026), we really use the GEV fit here as a regularising tool for the relatively small sample size. When we plot the GEV-fitted risk functions, we also show their 50th percentile-bootstrapped confidence intervals by resampling with replacement 1000 times, re-calculating GEV parameters for each, and shading between the 25th and 75th percentiles.

To generate such distributions for ‘non-SSW’ periods, we sample ‘random’ 45-day periods from the ERA5 dataset using the 30 SSW dates as a basis. Starting from the first SSW in the record on 1 March 1980, we randomise the year. If, for example, the random year is 2015 then the period of interest is the 45 days following the 1 March 2015 (inclusive). Then we find the $E'_{\max}$ within that period, as above. Note that the random

period cannot have the same start date as an actual SSW—that is, there is no duplication in terms of start dates between this and the aforementioned ‘post-SSW’ group, though random periods can overlap with post-SSW 45-day periods. This approach of sampling based on SSW dates enables us to capture the seasonal sampling typical of SSW periods better than if, for example, we picked random 45-day periods over all of December–March. We repeat this randomised-year process for each of the other 29 SSW onset dates (with a new random year each time) to get a list of 30 $E'_{\max}$ values that can be sorted from smallest to largest. We repeat this whole process 1000 times and the result is 1000 lists of 30 extremes, all sorted from smallest to largest (each severity with an associated exceedance probability). The ultimate GEV distribution is fitted to a single list of 30 extremes found by averaging across the 1000 samples, from which the 50% confidence intervals are also computed. In other words, we use an *average* distribution of random extremes. Then, we attempt to detect a signal of SSWs using *RR* and *QS*, with the p_1 derived from the post-SSW and p_0 from the random periods.

2.4.2 | Risk Analysis With SNAPSI

Meanwhile, with SNAPSI, comparing post-SSW to ‘non-SSW’ periods means simply comparing the nudged to the control experiments. In this case, we use absolute values rather than anomalies. Thus, the single highest E on any day in the period following the SSW is identified as the $E_{\max}$ for each ensemble member. This results in a list of 50 $E_{\max}$ values corresponding to the most extreme daily electricity demand in the 27 days after the SSW for each of the 50 ensemble members (of each model, for each initialisation date, for each experiment), to which a GEV distribution is fitted with 50th percentile-bootstrapped confidence intervals, as with ERA5. *RR* is then calculated as the ratio of the severities with respect to the control. Comparing nudged to control is most akin to a traditional attribution, whereby the factual world is one in which the SSW occurred, and the counterfactual is one that follows climatology. Whereas, *free* compared to *control* indicates the stratospheric contribution to forecast risk in standard forecasts.

To compare ERA5 and SNAPSI-derived demand we also construct an ‘ensemble’ from ERA5 using the post-SSW period of 12 February + 26 days (~11 March) in each year from 1980 to 2025 (inclusive). $E_{\max}$ is calculated from the 27-day period of each of these 46 ‘members’, with a GEV distribution fitted to the resultant list of extremes. We use this approach to infer the likelihood of extreme electricity demand within that 27-day period throughout the record; it reveals that the aggregated northern European 2018 $E_{\max}$ derived from ERA5 T is the second most extreme demand on a single day between 12 February and 11 March in the 46-year record, with an associated severity of ~142 GW and exceedance probability of just over 0.06. This is close to a 1 in 16 event—that is, there is a ~6% chance that the most extreme one-day demand in that 27-day period from 12 February in each of the 46 years in the record will be at least ~142 GW. It is this ‘event’ that is used as the threshold for the attribution analysis throughout this study.

3 | The Observed Impact of SSWs

3.1 | Demand Following SSWs

From the T and E timeseries, we can identify the historical influence of SSWs on mean and extreme temperature and electricity demand in northern and central Europe. As described in Section 2.1, these are shown as anomalies, as there is considerable variation in absolute electricity demand values over the winter between countries, as well as the aggregated groups of countries (Figure 4).

Figure 5 shows daily evolution of T' and E' per country in the days leading up to and following SSWs, averaged over all 30 events. We see that, in general, temperatures are lower and demand higher than average across northern Europe in the 45 days after an SSW (Figure 5a,b); there is a weaker effect in the central European countries (Figure 5c,d). Though the evolution varies between countries, and even groups, SSWs are on average followed by anomalously low temperatures for the time of year—notably around days 10–20, extending up to day 40 in some of the Nordic countries especially (Figure 5a). The troughs in T' are also broadly accompanied by peaks in E' , as would be expected, and countries with a higher weather-sensitivity of electricity demand can be identified. For example, though the northern-most European countries generally experience the lowest post-SSW temperatures (and most negative anomalies), it is France that has a particularly strong electricity demand response (Figure 5d). This is indicative not only of France being one of the largest countries in Europe, but also how highly responsive their demand is to, even small, temperature changes, due to their large proportion of electric heating (see Supporting Information: Table 1 in Bloomfield et al. (2021) for details of individual countries’ weather-sensitivity).

Averaging the temperature and demand anomalies across the 45 days following SSWs, $\overline{T'}$ and $\overline{E'}$ are on average lower and greater than climatology, respectively (Figure 6)—but the anomalies are usually small. Conversely, the average $T'_{\min}$ and $E'_{\max}$ values are highly anomalous. This indicates that while SSWs may have a small influence on mean conditions following SSWs, the most extreme days can be far outside climatology, as has previously been found for temperature (King et al. 2019). We thus suggest that the SSW does not equally impact the entirety of the temperature, and thus demand, distribution—that is, the temperature distribution is not simply shifted colder, but colder temperatures (and thus higher demands) are made more extreme. There is a similar story for the individual countries (see Supporting Information), though the demand is very country-dependent, again revealing certain countries’ strong temperature sensitivity.

3.2 | The Risk of Extreme Demand

We now investigate whether there is a greater likelihood of extreme high electricity demand occurring in the days following an SSW. Panel (i) in Figure 7a,b demonstrates the approach described in Section 2.4; the GEV fits are applied to the most extreme one-day electricity demand ($E'_{\max}$) in the 45-day period

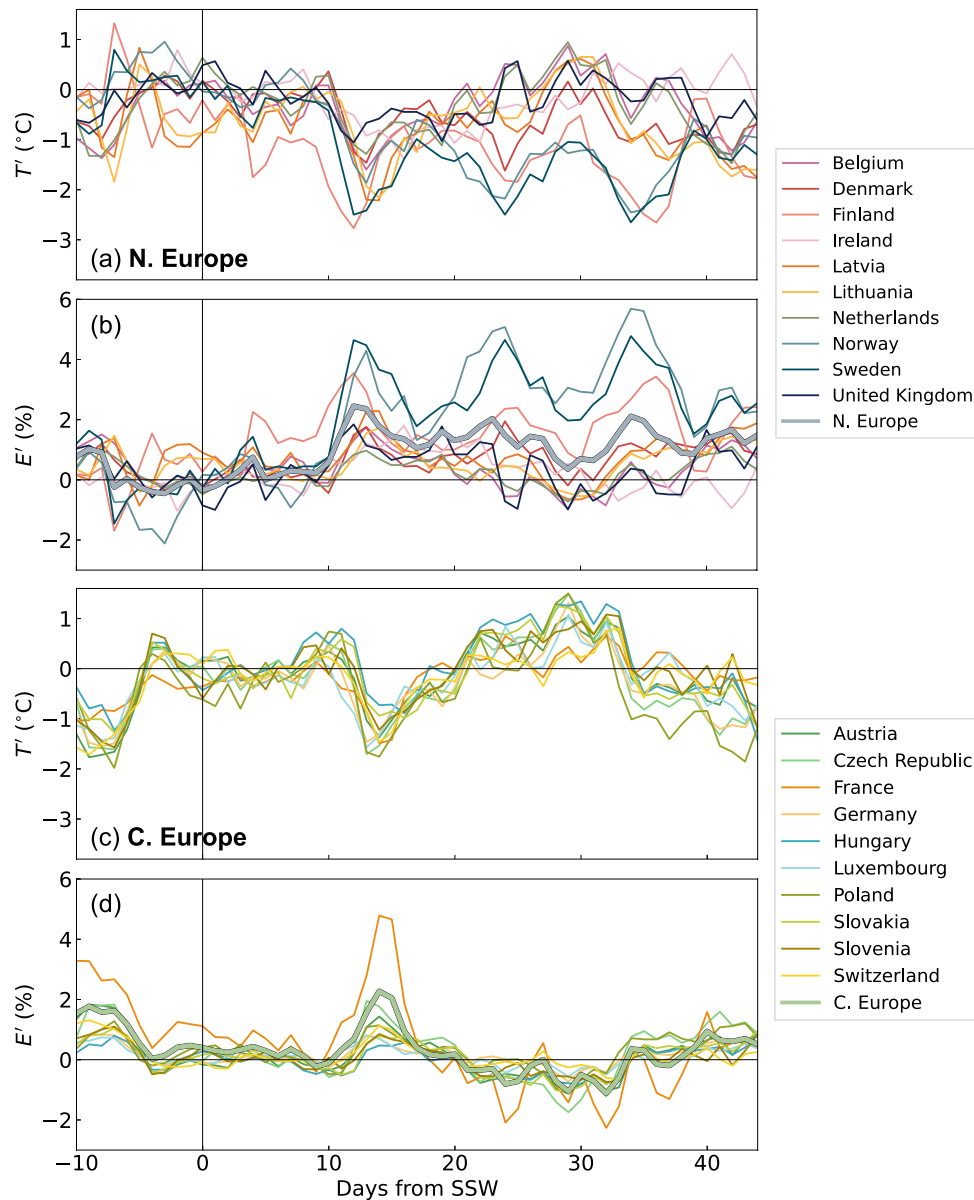


FIGURE 5 | Daily evolution of (a, c) temperature and (b, d) electricity demand anomaly (see Section 2.1) per northern and central European country, averaged over all 30 SSWs, including the E' for the entire country group illustrated with a thicker line.

after each of the 30 SSWs (crosses and solid line) and that from the 'random' periods (dots and dotted line). The RR on panel (ii) is then essentially the ratio of those two lines, and QS on panel (iii) the difference.

Comparing the GEV distributions for the post-SSW to random periods (Figure 7, panel i) indicates that the presence of SSWs generally increases the RR (panel ii) and intensifies the QS (panel iii) for both country groups—and for the most extreme, low probability values particularly. For central Europe (Figure 7b), the risk of lower values of $E'_{\max}$ in the post-SSW periods is fairly indistinguishable from that of the random periods ($RR \sim 1$ and $QS \sim 0$ in panels ii and iii, respectively). However, the most extreme value seems to be three times as likely post-SSW, with a difference in severity of almost 3% of climatological demand (Figure 7b, panels ii and iii). For

northern Europe (Figure 7a), the most extreme value is just over 1.5 times as likely, with a smaller shift in severity of $\sim 1\%$, for the post-SSW periods (panel iii). Though there is wide uncertainty in the RR values for the most extreme $E'_{\max}$ values (panel ii), the 50% confidence intervals are typically still above 1, indicating a significant difference between the SSW and random periods. Furthermore, since $\bar{E}'$ in the post-SSW periods is on average lower than that from random periods (not shown), we suggest that the difference in electricity demand extremes between the two distributions cannot simply be attributed to SSWs shifting the demand distribution; SSWs appear to have greatest influence on the tails of the distribution. Altogether, this analysis gives an indication of the level of extreme electricity demand that could follow any given SSW; or how (un)likely very extreme one-day demand events could be. This could be helpful for decision making at the onset of an SSW.

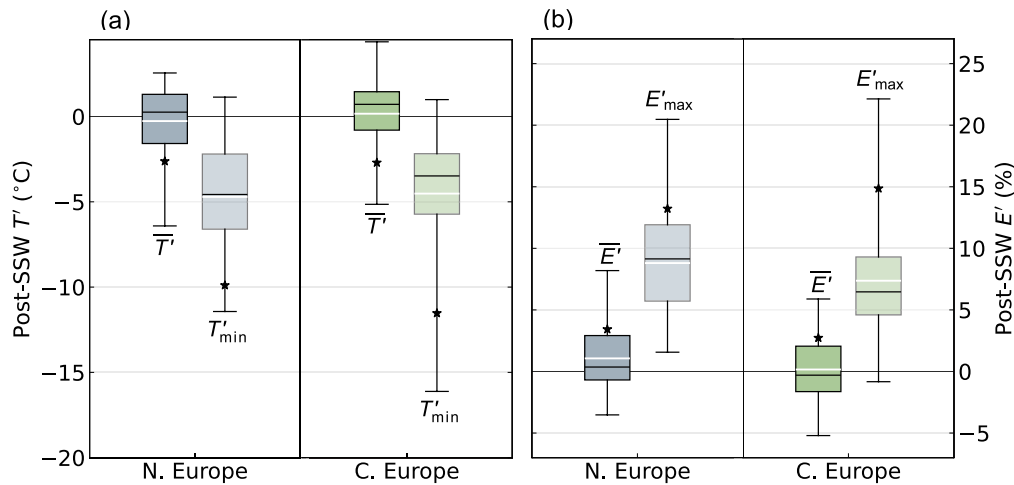


FIGURE 6 | Distribution of mean and extreme anomalous (a) temperatures and (b) electricity demand in the 45 days following the 30 SSWs; the near-central black bar marks the median, upper/lower the extrema, and white the mean. The stars mark values corresponding to the 2018 SSW.

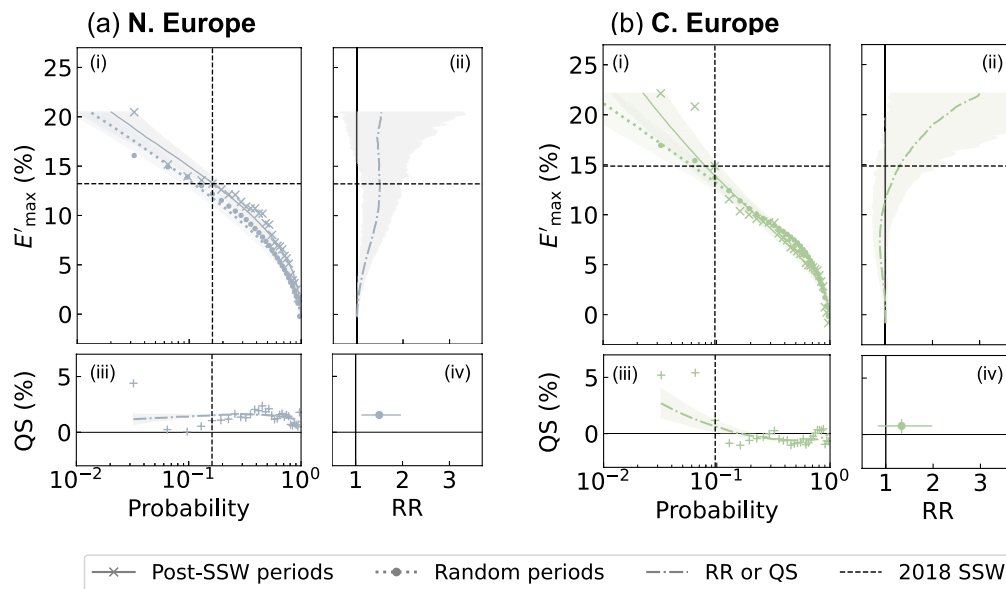


FIGURE 7 | The effects of SSWs in the ERA5 record aggregated over (a) northern and (b) central Europe. (i) shows severities ($E'_{\max}$) versus the risk, or probability, of an event of even greater severity. Symbols (dots and crosses) are used to represent empirical risks and lines the GEV-fitted risks. (ii) shows relative risks of the post-SSW versus random periods as a function of severity, while (iii) shows quantile shifts as a function of risk, as well as empirical differences (crosses). All shaded bands are 50% confidence intervals based on the percentile bootstrap with 1000 resamplings. Then (iv) shows the relative risk and quantile shift corresponding to the 12 February 2018 SSW (the dashed vertical/horizontal black lines on panels i–iii).

As in classic climate change attribution studies, it can be helpful to focus on the change in a particular ‘event’ or threshold. Rather than using a generic class of event, we use a threshold from a real-world example: the 12 February 2018 SSW (marked by the dashed lines on panels i–iii in Figure 7a,b). For northern Europe, the 2018 event $E'_{\max}$ is found to be ~113% of the climatological value, making it the fifth most extreme post-SSW $E'_{\max}$ out of the 30 SSWs in the record. The corresponding exceedance probability (i.e., the probability of being even more extreme) is then $5/30 = 0.16$. This makes it a one in six event—that is, there is an approximately one in six chance of SSWs being followed by an $E'_{\max}$ at least as high as 13%. For

central Europe, the 2018 event is the third most extreme $E'_{\max}$, with a corresponding exceedance probability of 0.10 (i.e., a 1 in 10 event). Comparing to the equivalent extremes in the distribution from the random periods indicates that such an event is 1.51 and 1.26 times as likely following the SSW, for northern and central Europe, respectively. The SSW also causes small shifts in severity of +1.57% and +0.71% of climatology (panel iv in Figure 7a,b, respectively). Additionally, there is high variability at the level of individual countries (see [Supporting Information](#)), but it is worth noting that a high demand event in one country might not ultimately cause a big impact in reality if there is sufficient supply to meet the demand.

4 | The Attributable Impact of the 2018 SSW

Analysis of ERA5 data certainly suggests a *difference* in extreme electricity demand events in post-SSW periods compared to random winter periods: SSWs appear to make extremes both more likely and more severe. However, it does not explicitly show a causal link between the two. We take the extra step to specifically *attribute* anomalous extreme electricity demand to the occurrence of the 2018 Northern Hemisphere SSW using SNAPSI.

For each simulation, T is used to derive E throughout the 45-day forecast, with particular interest in the 27-day period after the SSW (the ‘post-SSW period’, onset date inclusive). Because there are no hindcast ‘climatologies’ for the S2S models used here, SNAPSI-derived electricity demand extremes will be shown as absolute values (in GW) instead of proportions. For this analysis, also note that focus is on the simulated electricity demand for the aggregated northern Europe group only. This is partly because the low temperature response after this SSW was strongest in the north of Europe (Figure 2), but also because most of the seven models struggle to simulate demand extremes comparable to that derived from ERA5 for

central Europe (see [Supporting Information](#)). This is consistent with the nudged experiments simulating warmer temperatures across central Europe, after the later (2 February) initialisation especially, compared to *control* (Figure 8b,d). Figure 8a,b also show that, across the models (in the multi-model mean, MMM), the forecast mean temperatures are lower across northern Europe in the nudged compared to *control* experiments. Though, when compared to Figure 2a, it seems that the anomaly may be much weaker than in reality. Nevertheless, the lower temperatures in the nudged vs. *control* experiments suggest a possible signal of the SSW in temperature-driven electricity demand, too.

4.1 | Demand Following the SSW

Following the early (25 January) initialisation, the MMM electricity demand is typically greater in the nudged compared to *control* and *free* simulations throughout the post-SSW period (Figure 9a). Following the later (2 February) initialisation, the MMM nudged and *free* simulations’ evolution is more similar, with both diverging above the *control* around a week after SSW onset (Figure 9b). This is expected because the

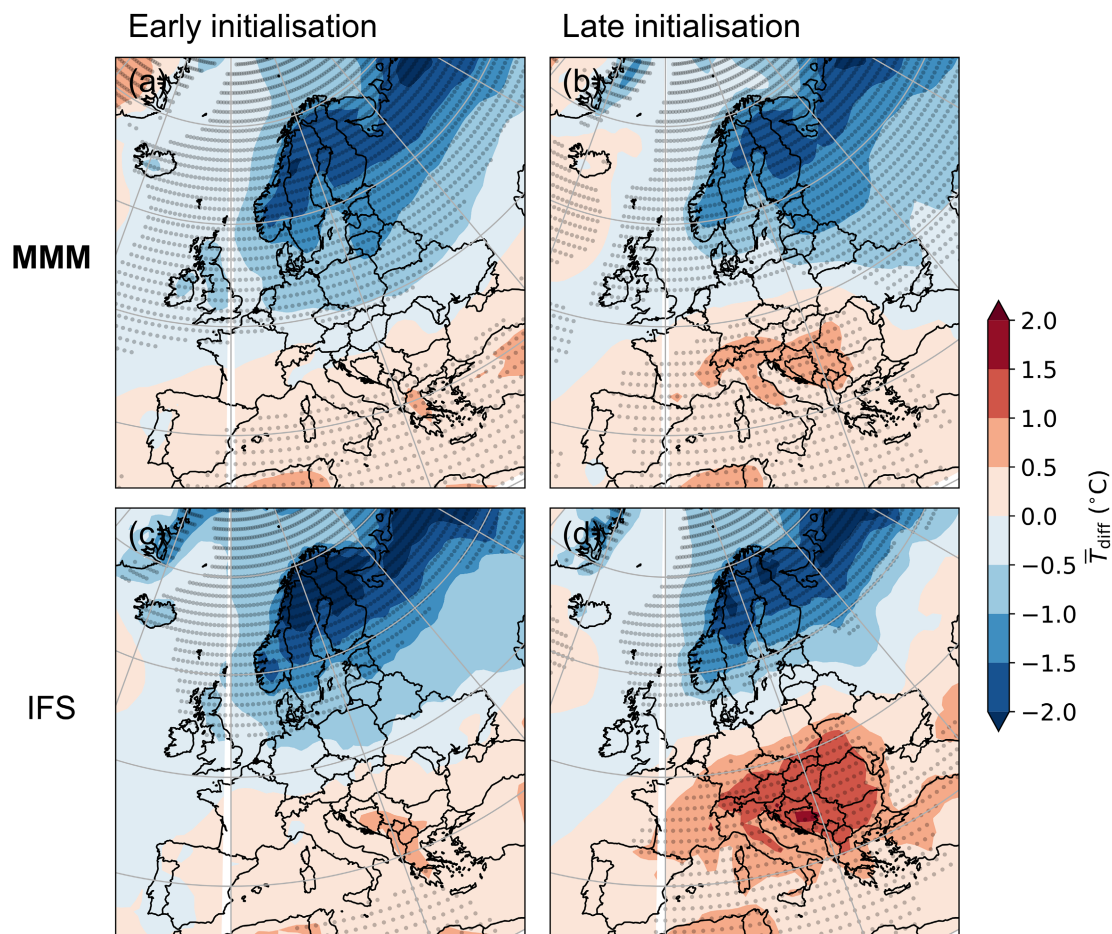


FIGURE 8 | The difference in multi-model (a, b) and IFS ensemble- (c, d) mean 2m temperature in the 27 days after the 12 February 2018 SSW in nudged minus *control* experiments, following the early (a, c) and late (b, d) initialisations. Model data were regridded to $1^\circ \times 1^\circ$ using bilinear interpolation for the computation of the multi-model mean (MMM). Stippling indicates where the difference is significant to the 97.5% confidence level using a bootstrapping approach. The same for all other models, as well as *free* minus *control*, can be found in [Supporting Information](#).

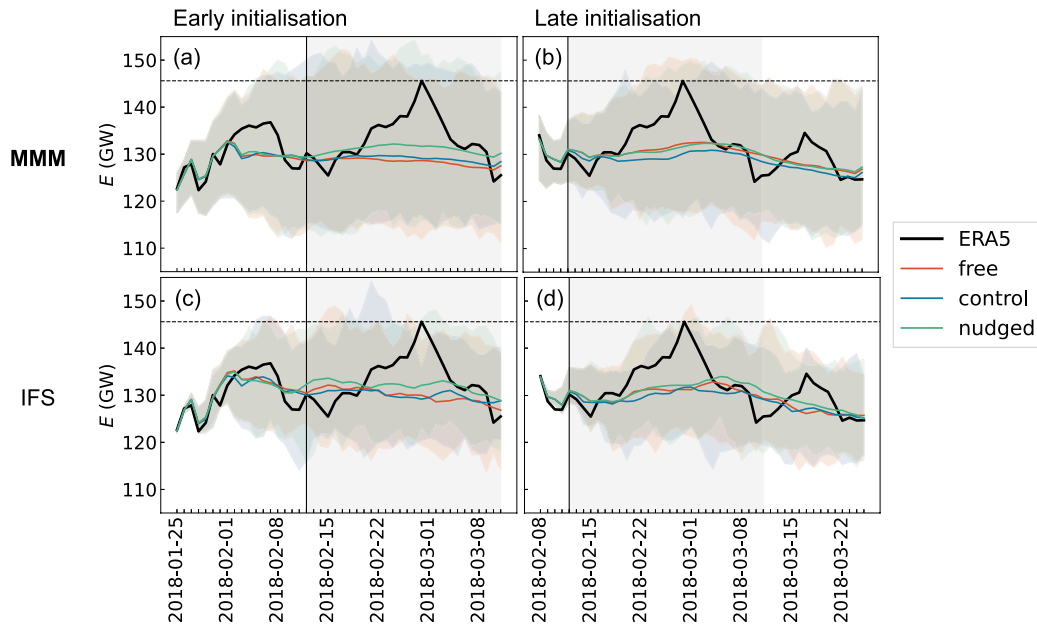


FIGURE 9 | Evolution of derived northern European E following the early (a, c) and late (b, d) initialisations. Coloured lines are the multi-model mean (MMM, a, b) or IFS ensemble mean (c, d) per experiment and the shaded envelope the min./max. daily E across all of the models' or IFS ensemble members, with ERA5 in black. The grey shading marks the 27-day post-SSW period up to ~11 March, the vertical line the SSW onset date, and the dashed horizontal line marks the ERA5 max., used as the event threshold in Section 4.2.

free simulations capture the onset of the SSW at the time of the later initialisation (Figure 3). There is, however, substantial inter-model spread, with some models showing a much greater divergence between the different experiments than others (e.g., CESM2-CAM6 compared to IFS in Figure 9; see [Supporting Information](#) for the other SNAPSI-participating models). The MMMs in Figure 9 show that averaging across models reduces much of the noise (i.e., comparing the single IFS model to the MMM)—in particular from the early initialisation (Figure 9a). The shaded envelopes also show the E extremes across all models' ensemble members, with some members for some experiments simulating demand above even that derived from ERA5. However, this is driven by a few models for each experiment (see [Supporting Information](#) for the additional models); for some, that 'observed' extreme falls well outside of the simulated distributions (e.g., CanESM5 and CNRM-CM6-1, discussed further in the next section). The two GloSea6 models also show some of the highest ensemble mean (and extreme) demand values across the experiments, though particularly in the *nudged* experiment for the early initialisation and the *free* for the late. This, coupled with the offset from ERA5 at initialisation, suggests a possible cold bias compared to some of the other models; CanESM5 shows the opposite, suggesting a warm bias instead.

4.2 | The Risk of Extreme Demand

We now investigate how much of the risk of a given extreme electricity demand 'event' can be specifically attributed to the 2018 SSW. As described in Section 2.4, we use the ERA5-derived 2018 $E_{\max}$ ~142 GW as a threshold with which to conduct an inter-model comparison. We start by using a single model, IFS, to demonstrate the methodology—the simulated evolution of

the stratospheric zonal wind and temperature-derived demand are shown in Figures 3b and 9c,d, respectively.

The GEV distributions fitted to the IFS ensemble's simulation of one-day electricity demand extremes ($E_{\max}$) are comparable to the climatology of extremes calculated from ERA5 (panel i in Figure 10a,b), suggesting that the model is relatively unbiased. Between the two initialisations, the *free* RR becomes more similar to that of the *nudged* when the initialisation date is closer to SSW onset (panel ii). Compared to some of the other models (see [Supporting Information](#)), *nudged* generates extremes comparable to that of *free* where exceedance probabilities are low; the shifts in severity become almost equivalent for the lowest exceedance probabilities (panel iii). The dashed lines on Figure 10 correspond to the ERA5-derived 2018 event of interest. The horizontal dashed line corresponds to the event's severity, which lies well within the IFS model's ensemble spread; the x-axis positions of intersection with the SNAPSI experiment lines indicate absolute risks of exceeding that observed severity. The vertical dashed line alternatively corresponds to the event's absolute risk; the y-axis positions of intersection indicate the severity of an equivalent-risk event in the model ensemble. The ratio of absolute risk values corresponding to the ERA5 2018 threshold gives the associated RR , and the difference in the equivalent-risk severities the QS . For the early initialisation, such an event in IFS is 1.31 and 2.22 times as likely in the *free* and *nudged* vs. *control*, with small shifts in severity, of +1.03 and +2.08 GW (panels iii and iv in Figure 10a,b, respectively). For the late initialisation, these become $RR = 1.90$ and 2.18 times, and $QS = +2.10$ and +2.5 GW. Thus, for IFS, there is a consistent influence of the SSW: *nudged* both increases the RR and intensifies the QS compared to the *control*, though the extent of the effect is lead time-dependent. The difference between the

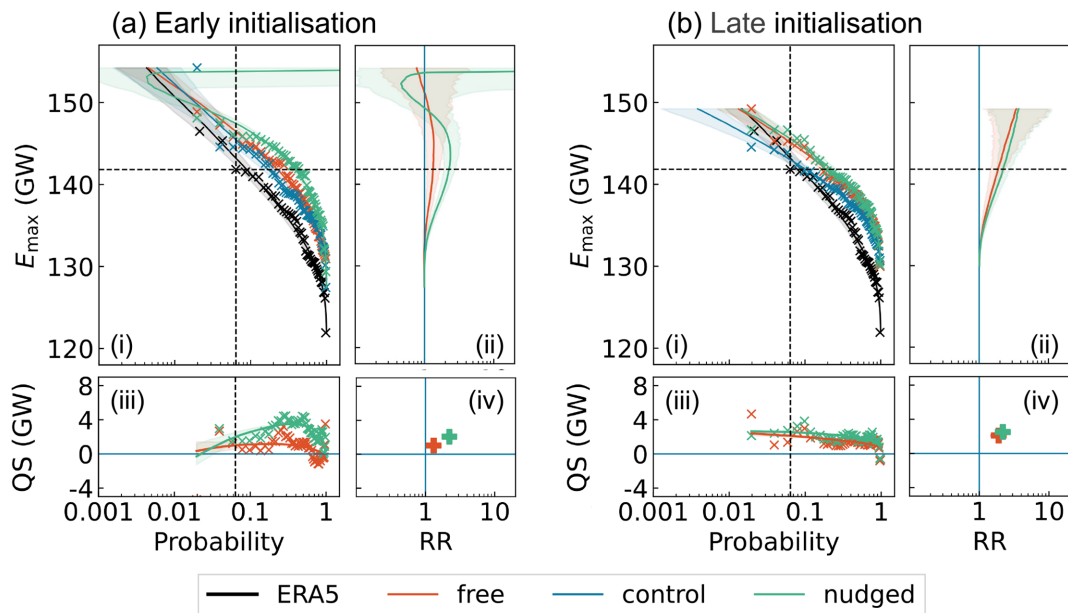


FIGURE 10 | A summary of the experimental effects with the IFS model for the (a) early and (b) late initialisations. (i) shows severities ($E_{\max}$) as a function of exceedance probability of an event of even greater severity; the crosses indicate empirical risks and lines the GEV-fitted risks. (ii) shows relative risks with respect to *control* as a function of severity, while (iii) shows quantile shifts with respect to *control* as a function of exceedance probability, as well as empirical differences (crosses). Then (iv) shows the relative risk and quantile shift corresponding to the ERA5-derived value for the 12 February 2018 SSW (the dashed black lines on panels i–iii). All shaded bands, and error bars in panel (iv), are 50% confidence intervals based on the percentile bootstrap with 1000 resamplings.

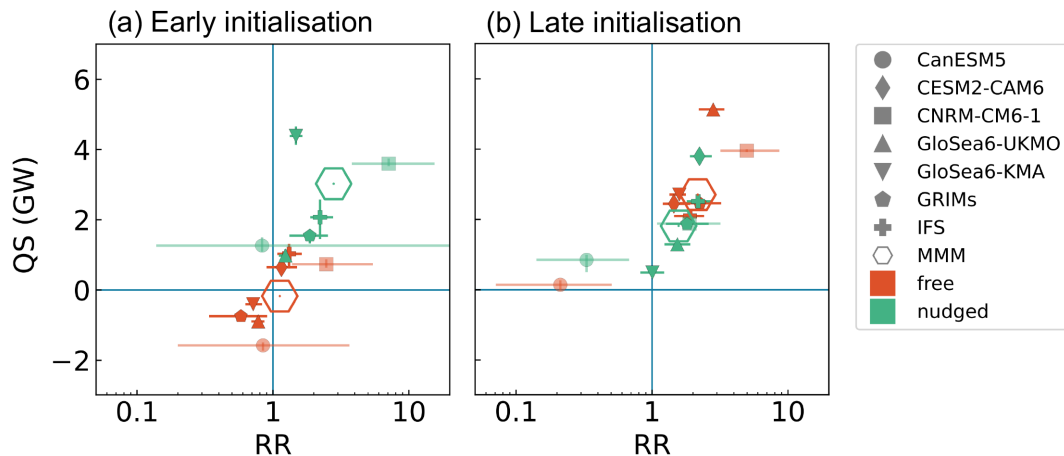


FIGURE 11 | As in panel (iv) of Figure 10, each model's *RR* versus *QS* (w.r.t. *control*), as well as a multi-model mean (MMM). Error bars show the 25th and 75th percentiles and more transparent markers/bars are used for models and experiments where the ERA5 value falls outside of the model ensemble's distribution.

nudged and free forecasts compared to control is small in the IFS—especially after the later initialisation, when the SSW onset is captured by the free (as in Figure 3b). This means that there is little additional information gained by the nudging at that point; more value is gained from nudged for the early initialisation. Indeed, the free simulations show a greater change in *QS* and *RR* between the two initialisations (the former almost doubles), while the nudged *RR* is slightly lower after the late initialisation, despite a greater *QS*.

A summary of the *RR* versus *QS* values corresponding to the 2018 threshold for all of the models is shown in Figure 11. At

the MMM level, shown as the large hexagons, both the free and nudged simulations increase the likelihood of the 2018 extreme compared to the control simulations. The *RR* is greater for the nudged than free following the early initialisation ($RR = 2.79$ compared to 1.12 in Figure 11a), but this reverses for the late initialisation ($RR = 1.57$ compared to 2.17 in Figure 11b). This is indicative of the improvement in the free when the initialisation date is closer to the SSW onset date. In the early initialisation, the MMM *QS* is just over 3 GW for nudged and close to 0 GW for free. Meanwhile for the late initialisation, the MMM *QS* values are more similar between experiments (free and nudged $QS = 2.70$ and 1.81 GW,

respectively), similar to the *RR*. Repeating the ERA5-focused analysis in Section 3.2 using absolute values, rather than anomalies, reveals $QS = 2.40$ GW between the post-2018 SSW period $E_{\max}$ and the equivalent (exceedance probability) event from the random period distribution. This is comparable to that from SNAPSI—indeed, the MMM *nudged* vs. *control* QS averaged across both initialisations is exactly the same value. This suggests that the predictability added by the stratosphere in the *nudged* simulations for this event is consistent with what one might infer from a ‘typical’, observed SSW.

Note that not every model set-up used here generates extremes comparable to ERA5, with some apparently not able to capture an event with a severity like that in 2018. As this impacts on the *RR*, the more transparent markers/error bars on Figure 11 indicate models and experiments where this is the case. The *nudged* and *free* experiments are by-default made more transparent if the *control* value is outside the ERA5 distribution, since *RR* is with respect to *control*. Operational forecasters of demand would typically bias correct models so that an issue such as this could (possibly) be avoided; as will be discussed further in Section 5, this has not been possible here without model climatologies. Altogether, as Seviour et al. (2026) did with temperature extremes, this analysis affirms the original hypothesis of the SNAPSI project: that the 2018 SSW played a causal role in extreme surface impacts—in this case on electricity demand.

5 | Discussion and Conclusions

This analysis builds on previous work utilising stratospheric variability and predictability for the energy sector (e.g., using subseasonal forecasts as in Büeler et al. 2020). We specifically aim to link the occurrence of SSWs in the Northern Hemisphere to extreme electricity demand impacts across Europe. By keeping electricity network and society-relevant parameters in the demand model constant (Section 2.3), we have been able to focus on the impact of post-SSW weather conditions on electricity demand. Using ERA5 near-surface temperature over 46 winters, we find that derived extreme electricity demand in the aggregated northern and central European groups is slightly more likely and severe following an SSW (Figure 7). For central Europe, the most extreme demand anomaly $E'_{\max \sim 20\%}$ is three times more likely to occur in periods following SSWs than in random periods, while for northern Europe the most extreme $E'_{\max \sim 20\%}$ is around 1.5 times more likely. This is consistent with the existing narrative that SSWs precede periods of cold weather, including extremes (Domeisen and Butler 2020; Huang et al. 2021; Spaeth and Birner 2022).

The SNAPSI protocol then enables us to explicitly investigate the attributable impact of the 2018 SSW on extreme northern European demand. We find that improved prediction of the SSW from the early initialisation leads to ~ 3 times greater predicted risk of extreme weather-derived electricity demand in northern Europe at least as large as in 2018 (Figure 11). Even following the late initialisation, the stratospheric influence is important, as the *control* has approximately half the risk of the *free* or *nudged*. However, our risk statements are highly dynamically conditioned, with a strong dependency on initialisation date.

A forecast ensemble initialised just a few days before an event is much more heavily conditioned, with much less ensemble spread, than one initialised weeks before (Figure 3); the interpretation of statistical measures such as *RR* and *QS* is greatly affected by the fact that the probability of an extreme event is highly dependent on the conditions at the point of initialisation (Leach et al. 2021; Hitchcock et al. 2022).

These results are broadly consistent with SNAPSI-focused analyses of weather regimes and temperature extremes conducted so far (Dai et al. 2025; Lee et al. 2025; Seviour et al. 2026). Lee et al. (2025) specifically note that the SNAPSI-participating models they analyse overestimate stratosphere-troposphere coupling. The translation of near-surface variables into energy metrics is likely to be affected by this, on top of the uncertainties associated with the demand model itself. Dai et al. (2025) additionally report that only $\sim 25\%$ of the observed negative NAO anomalies were ‘forced’ by the SSW, as diagnosed from precipitation and mean sea-level pressure. The temperature differences between the *nudged* and *control* experiments (Figure 8) being smaller than the anomalies observed following the event (Figure 2a) suggest a similar situation for temperature, and thus demand. Nevertheless, this nudging framework indicates a way in which to study a variety of vortex events beyond SSWs, such as strong or even stretched vortex events, which can have comparable impacts to SSWs (Cohen et al. 2021). In the context of finding the attributable influence of the vortex in surface events, it is also key to accurately capture the temporal and spatial evolution of the stratospheric polar vortex. Full 3D nudging can facilitate this, and is another part of the SNAPSI protocol that is still being explored.

This analysis utilises just one of the SNAPSI case studies, and with a focus on the northern European countries. It is worth remembering that all SSWs are different; not every SSW will necessarily produce an attributable impact in electricity extremes in Europe, with some instead affecting North America (as in the other Northern Hemisphere SNAPSI case study in January 2019) or Asia, and others producing no surface impact at all (Sigmond et al. 2013; Karpechko et al. 2017; Oehrlein et al. 2021). Indeed, most research in this area, including the ERA5-focused analysis in this study, tends to generalise across SSWs. It would be interesting for future work to explore the impact of different ‘flavours’ of SSWs (Baldwin et al. 2021): wavenumber-1 vs. wavenumber-2; major vs. minor; split vs. displacement; downward-propagating vs. not; those that affect Europe vs. North America vs. Asia. Recent studies focused on Finland also suggest a dependence of post-SSW energy impacts on the phase of the Quasi-Biennial Oscillation (Juntunen and Asikainen 2023; Juntunen, Asikainen, and Salminen 2025; Juntunen, Vokhmyanin, et al. 2025), something not explored at all here. With historical records at least, slicing by these different flavours reduces what is already a fairly small sample even further. However, by generalising, the variation in the weather conditions that can follow SSWs is not properly considered here (Mitchell et al. 2013; Beerli and Grams 2019). Though the general pattern may be a negative NAO, this is comprised of a range of weather sub-patterns which can manifest as a range of responses between countries (Neal et al. 2016; Rouges et al. 2025). At the same time, there are complex sensitivities involved in the conversion of weather to energy variables (Bloomfield et al. 2021), sampling uncertainty,

and also geographical differences that can drive inter-country variations. Most of this analysis has focused on aggregated groups of countries, but there are individual countries whose demand is much more sensitive to temperature changes and which can drive the overall demand (e.g., France in the central Europe group; Figure 6d).

One particularly notable limitation of this work is that the SNAPSI forecasts are neither bias corrected nor calibrated. For the former, there does seem to be some bias in the near-surface temperatures that in turn impacts on the derived electricity demand (Section 4.1). In this case, there is a current lack of climatologies for all of the models in the set-ups used for SNAPSI. The impact of a lack of bias correction is partly reduced, here, because we compare the difference between experiments with the same model, rather than absolute values. However, as the distribution of extremes is non-linear, this does not completely remove bias and a community effort to generate climatologies in order to bias correct the forecasts would undoubtedly be useful going forward (Lawrence et al. 2022; Garfinkel et al. 2024). Forecast calibration could alternatively compare model performance to ERA5 at a range of lead times; for example, ensemble mean demand could be compared to ERA5-derived demand to identify if there is a bias and whether it grows at different lead times. Similar to the climatology issue, this would require the exact set-ups of the SNAPSI-participating models to be run at more than just the two lead times standard in the protocol. Nevertheless, correction of any identified biases could then enable better forecasts of impacts.

Overall, it is clear that the SNAPSI protocol provides an opportunity to explicitly and robustly link specific stratospheric events to tropospheric, including socioeconomic, impacts. Furthermore, the demand model being dependent only on weather variables can help elucidate a clear causal link between weather conditions and fluctuations in electricity demand (and generation). Future research could explore not only variety in SSWs, but other types of extreme stratospheric events and their impacts too, as well as other energy metrics relevant to those in the sector. For example, it could be interesting to go beyond one-day extremes and identify periods reaching the 90th percentile of demand and/or to investigate the net impact of temperature and wind on electricity shortfall (demand minus renewables) events (e.g., Rouges et al. 2025). Then, as the world moves closer to achieving net zero targets amidst a changing climate, better understanding of extreme event drivers can help improve resiliency and enable future planning. It could be useful to compare to real-world electricity records, too, to both better understand biases of the demand model and better quantify the relative influence of weather conditions compared with, for example, population behaviour, electricity infrastructure, or network constraints. To do so would require the re-incorporation of socioeconomic factors (such as long-term trends, day-of-week variability, and country-dependent base temperatures) which we kept constant or did not consider in this analysis in order to focus on weather-driven impacts. It could also be interesting to examine changes to the electricity network over time and into the future, or to combine the nudged forecast with perturbed climate forecast techniques increasingly utilised in climate change attribution (e.g., Leach et al. 2021; Ermis et al. 2024).

Climate projections for the 21st century typically indicate fewer cold days with climate change (Seneviratne et al. 2021; Zheng et al. 2023), with a warming signal in cold periods, and some associated impacts already detected (Lorenz et al. 2019; Christidis et al. 2021). Nevertheless, cold extremes are likely to persist even with the overall warming signal, with the potential risk that society becomes less resilient to them as they become less frequent. Furthermore, though the stratospheric response to climate change is uncertain (Ayarzagüena et al. 2020; Rao and Garfinkel 2021; Karpechko et al. 2024), SSWs are likely to remain important examples of extreme atmospheric events that are both impactful for society and a source of predictive skill for the troposphere on S2S timescales. In the case of the energy sector, improving the skill of weather forecasts beyond the medium range could in turn improve forecasts of electricity demand and renewable generation, and thus prices, at S2S timescales; this would be valuable not only to energy system operators and consumers, but also to energy traders (S2S4E 2020). We believe that this work demonstrates that SSWs can particularly provide helpful predictive power for the energy sector. We would, however, recommend refining the modelling to account for socioeconomic factors, such as the differing base temperatures between countries and variability in demand by weekday, in order to accurately operationalise the forecasting of SSW-driven demand.

Author Contributions

Regan Mudhar: conceptualization, investigation, writing – original draft, methodology, visualization, writing – review and editing, formal analysis. **William J. M. Seviour:** conceptualization, writing – review and editing, methodology, resources, supervision. **Hannah C. Bloomfield:** conceptualization, writing – review and editing, methodology, resources, supervision. **Amy H. Butler:** writing – review and editing, project administration, resources. **Chaim I. Garfinkel:** writing – review and editing, project administration, resources. **Peter Hitchcock:** writing – review and editing, project administration, resources. **James A. Screen:** writing – review and editing, supervision. **Stephen I. Thomson:** writing – review and editing, supervision.

Acknowledgements

R.M. is funded by a NERC GW4+ Doctoral Training Partnership studentship from the Natural Environment Research Council (NE/S007504/1). H.C.B. was supported by Newcastle University Academic Track Fellowship (NU-Act) scheme. The authors thank Justin Finkel for helpful discussion about the statistical methodology. This work used JASMIN, the UK's collaborative data analysis environment (Lawrence et al. 2013, <https://www.jasmin.ac.uk>). The distance-weighted average remapping of the ERA5T data was performed using cdo (Schulzweida 2023). The bilinear interpolation of the SNAPSI data for the MMM-composites in Figure 8 was performed using xESMF (Zhuang et al. 2025).

Funding

This work was supported by a Natural Environment Research Council studentship with grant number NE/S007504/1.

Conflicts of Interest

H.C.B. is an Associate Editor of the journal but was not privy to the peer review process.

Data Availability Statement

This work utilises data generated following the SNAPSI protocol, detailed in Hitchcock et al. (2022), and is publicly available from the Centre for Environmental Data Analysis archive, a UK-based repository part of the Natural Environment Research Council's Environmental Data Service: <https://catalogue.ceda.ac.uk/uuid/0a5a1ce22fb047749e040879efa8e>. The European weather-driven demand model was produced by Bloomfield et al. (2021) and can be accessed at <https://researchdata.reading.ac.uk/272/>. ERA5 reanalysis data are freely available from ECMWF (see Hersbach et al. 2020).

References

- Ayarzagüena, B., A. J. Charlton-Perez, A. H. Butler, et al. 2020. "Uncertainty in the Response of Sudden Stratospheric Warmings and Stratosphere-Troposphere Coupling to Quadrupled CO₂ Concentrations in CMIP6 Models." *Journal of Geophysical Research: Atmospheres* 125: e2019JD032345.
- Baldwin, M. P., B. Ayarzagüena, T. Birner, et al. 2021. "Sudden Stratospheric Warmings." *Reviews of Geophysics* 59: e2020RG000708.
- Baldwin, M. P., and T. J. Dunkerton. 2001. "Stratospheric Harbingers of Anomalous Weather Regimes." *Science* 294: 581–584.
- Beerli, R., and C. M. Grams. 2019. "Stratospheric Modulation of the Large-Scale Circulation in the Atlantic–European Region and Its Implications for Surface Weather Events." *Quarterly Journal of the Royal Meteorological Society* 145: 3732–3750.
- Beerli, R., H. Wernli, and C. M. Grams. 2017. "Does the Lower Stratosphere Provide Predictability for Month-Ahead Wind Electricity Generation in Europe?" *Quarterly Journal of the Royal Meteorological Society* 143: 3025–3036.
- Bloomfield, H. C., D. J. Brayshaw, P. L. Gonzalez, and A. Charlton-Perez. 2021. "Sub-Seasonal Forecasts of Demand and Wind Power and Solar Power Generation for 28 European Countries." *Earth System Science Data* 13: 2259–2274.
- Bloomfield, H. C., C. C. Suitters, and D. R. Drew. 2020. "Meteorological Drivers of European Power System Stress." *Journal of Renewable Energy* 2020: 1–12.
- Büeler, D., R. Beerli, H. Wernli, and C. M. Grams. 2020. "Stratospheric Influence on ECMWF Sub-Seasonal Forecast Skill for Energy-Industry-Relevant Surface Weather in European Countries." *Quarterly Journal of the Royal Meteorological Society* 146: 3675–3694.
- Busby, J. W., K. Baker, M. D. Bazilian, et al. 2021. "Cascading Risks: Understanding the 2021 Winter Blackout in Texas." *Energy Research & Social Science* 77: 102106.
- Butchart, N. 2022. "The Stratosphere: A Review of the Dynamics and Variability." *Weather and Climate Dynamics* 3: 1237–1272.
- Butler, A. H., J. P. Sjöberg, D. J. Seidel, and K. H. Rosenlof. 2017. "A Sudden Stratospheric Warming Compendium." *Earth System Science Data* 9: 63–76.
- Cattiaux, J., R. Vautard, C. Cassou, P. Yiou, V. Masson-Delmotte, and F. Codron. 2010. "Winter 2010 in Europe: A Cold Extreme in a Warming Climate." *Geophysical Research Letters* 37: L20704.
- Chan, W. C., T. G. Shepherd, K. Facer-Childs, G. Darch, and N. W. Arnell. 2022. "Storylines of UK Drought Based on the 2010–2012 Event." *Hydrology and Earth System Sciences* 26: 1755–1777.
- Charlton, A. J., and L. M. Polvani. 2007. "A New Look at Stratospheric Sudden Warmings. Part I: Climatology and Modeling Benchmarks." *Journal of Climate* 20: 449–469.
- Charlton-Perez, A. J., W. T. K. Huang, and S. H. Lee. 2021. "Impact of Sudden Stratospheric Warmings on United Kingdom Mortality." *Atmospheric Science Letters* 22: e1013.
- Christidis, N., M. McCarthy, and P. A. Stott. 2021. "Recent Decreases in Domestic Energy Consumption in the United Kingdom Attributed to Human Influence on the Climate." *Atmospheric Science Letters* 22: e1062.
- Clark, R. T., P. E. Bett, H. E. Thornton, and A. A. Scaife. 2017. "Skilful Seasonal Predictions for the European Energy Industry." *Environmental Research Letters* 12: 024002.
- Cohen, J., L. Agel, M. Barlow, C. I. Garfinkel, and I. White. 2021. "Linking Arctic Variability and Change With Extreme Winter Weather in the United States." *Science* 373: 1116–1121.
- Dai, Y., P. Hitchcock, A. H. Butler, C. I. Garfinkel, and W. J. Seviour. 2025. "Assessing Stratospheric Contributions to Subseasonal Predictions of Precipitation After the 2018 Sudden Stratospheric Warming From the Stratospheric Nudging and Predictable Surface Impacts (SNAPSI) Project." *Weather and Climate Dynamics* 6: 841–862.
- Domeisen, D. I., A. H. Butler, A. J. Charlton-Perez, et al. 2020a. "The Role of the Stratosphere in Subseasonal to Seasonal Prediction: 2. Predictability Arising From Stratosphere-Troposphere Coupling." *Journal of Geophysical Research: Atmospheres* 125: e2019JD030923.
- Domeisen, D. I., A. H. Butler, A. J. Charlton-Perez, et al. 2020b. "The Role of the Stratosphere in Subseasonal to Seasonal Prediction: 1. Predictability of the Stratosphere." *Journal of Geophysical Research: Atmospheres* 125: e2019JD030920.
- Domeisen, D. I. V., and A. H. Butler. 2020. "Stratospheric Drivers of Extreme Events at the Earth's Surface." *Communications Earth & Environment* 1: 59.
- Dorrington, J., I. Finney, T. Palmer, and A. Weisheimer. 2020. "Beyond Skill Scores: Exploring Sub-Seasonal Forecast Value Through a Case-Study of French Month-Ahead Energy Prediction." *Quarterly Journal of the Royal Meteorological Society* 146: 3623–3637.
- Doss-Gollin, J., D. J. Farnham, U. Lall, and V. Modi. 2021. "How Unprecedented Was the February 2021 Texas Cold Snap?" *Environmental Research Letters* 16: 064056.
- Drax Group. 2025. "Drax Energy Insights." <https://electricinsights.co.uk/>.
- EEA. 2024. "Heating Degree Days."
- Ermis, S., N. J. Leach, F. C. Lott, S. N. Sparrow, and A. Weisheimer. 2024. "Event Attribution of a Midlatitude Windstorm Using Ensemble Weather Forecasts." *Environmental Research: Climate* 3: 035001.
- Finkel, J., E. P. Gerber, D. S. Abbot, and J. Weare. 2023. "Revealing the Statistics of Extreme Events Hidden in Short Weather Forecast Data." *AGU Advances* 4: e2023AV000881.
- Foretic, H., I. Asensio Bermejo, V. Kopustinskas, and G. Fulli. 2024. "State of the Art of Regional Coordination Centres (RCCs) and Their Impact on Security of Electricity Supply in Europe." Tech. rep., European Commission, Joint Research Centre, Luxembourg. <https://doi.org/10.2760/01557>.
- Galvin, J., M. Kendon, and M. McCarthy. 2018. "Snow Cover and Low Temperatures in February and March 2018." *Weather* 74: 52.
- Garfinkel, C. I., Z. D. Lawrence, A. H. Butler, et al. 2024. "A Process-Based Evaluation of Biases in Extratropical Stratosphere-Troposphere Coupling in Subseasonal Forecast Systems." *Weather and Climate Dynamics* 6: 171–195.
- Goutham, N., R. Plougonven, H. Omrani, et al. 2022. "How Skillful Are the European Subseasonal Predictions of Wind Speed and Surface Temperature?" *Monthly Weather Review* 150: 1621–1637.
- Hegerl, G., and F. Zwiers. 2011. "Use of Models in Detection and Attribution of Climate Change." *Wiley Interdisciplinary Reviews: Climate Change* 2: 570–591.
- Hersbach, H., B. Bell, P. Berrisford, et al. 2020. "The ERA5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146:

- 1999–2049. <https://cds.climate.copernicus.eu/#/search?text=ERA5&type=dataset>.
- Hitchcock, P., A. Butler, A. Charlton-Perez, et al. 2022. “Stratospheric Nudging and Predictable Surface Impacts (SNAPSI): A Protocol for Investigating the Role of the Stratospheric Polar Vortex in Subseasonal to Seasonal Forecasts.” *Geoscientific Model Development Discussions* 15: 5073–5092.
- Huang, J., P. Hitchcock, A. C. Maycock, C. M. McKenna, and W. Tian. 2021. “Northern Hemisphere Cold Air Outbreaks Are More Likely to Be Severe During Weak Polar Vortex Conditions.” *Communications Earth & Environment* 2: 147.
- Ineson, S., N. J. Dunstone, A. A. Scaife, M. B. Andrews, J. F. Lockwood, and B. Pang. 2024. “Statistics of Sudden Stratospheric Warmings Using a Large Model Ensemble.” *Atmospheric Science Letters* 25: e1202.
- Juntunen, V., and T. Asikainen. 2023. “Electricity Consumption in Finland Influenced by Climate Effects of Energetic Particle Precipitation.” *Scientific Reports* 13: 20546.
- Juntunen, V., T. Asikainen, and A. Salminen. 2025. “Influence of Energetic Electron Precipitation on Wind Power Generation in European Countries Mediated by the Polar Vortex.” *Space Weather* 23: e2024SW004157.
- Juntunen, V., M. Vokhmyanin, A. Salminen, and T. Asikainen. 2025. “Long-Term Prediction of the Influence of Sudden Stratospheric Warmings on the Electricity Consumption and Wind Power Generation in Finland.” *Environmental Research Letters* 20: 104058.
- Karpechko, A. Y., P. Hitchcock, D. H. Peters, and A. Schneidereit. 2017. “Predictability of Downward Propagation of Major Sudden Stratospheric Warmings.” *Quarterly Journal of the Royal Meteorological Society* 143: 1459–1470.
- Karpechko, A. Y., Z. Wu, I. R. Simpson, et al. 2024. “Northern Hemisphere Stratosphere-Troposphere Circulation Change in CMIP6 Models: 2. Mechanisms and Sources of the Spread.” *Journal of Geophysical Research: Atmospheres* 129: e2024JD040823.
- Kautz, L. A., I. Polichtchouk, T. Birner, H. Garny, and J. G. Pinto. 2020. “Enhanced Extended-Range Predictability of the 2018 Late-Winter Eurasian Cold Spell due to the Stratosphere.” *Quarterly Journal of the Royal Meteorological Society* 146: 1040–1055.
- Kidston, J., A. A. Scaife, S. C. Hardiman, et al. 2015. “Stratospheric Influence on Tropospheric Jet Streams, Storm Tracks and Surface Weather.” *Nature Geoscience* 8: 433–440.
- King, A. D., A. H. Butler, M. Jucker, N. O. Earl, and I. Rudeva. 2019. “Observed Relationships Between Sudden Stratospheric Warmings and European Climate Extremes.” *Journal of Geophysical Research: Atmospheres* 124: 13943–13961.
- Kolstad, E. W., S. H. Lee, A. H. Butler, D. I. V. Domeisen, and C. O. Wulff. 2022. “Diverse Surface Signatures of Stratospheric Polar Vortex Anomalies.” *Journal of Geophysical Research: Atmospheres* 127: e2022JD037422.
- Lawrence, B. N., V. L. Bennett, J. Churchill, et al. 2013. “Storing and Manipulating Environmental Big Data With JASMIN.” In *IEEE Big Data*. IEEE.
- Lawrence, Z. D., M. Abalos, B. Ayarzagüena, et al. 2022. “Quantifying Stratospheric Biases and Identifying Their Potential Sources in Subseasonal Forecast Systems.” *Weather and Climate Dynamics* 3: 977–1001.
- Leach, N. J., C. D. Roberts, M. Aengenheyster, et al. 2024. “Heatwave Attribution Based on Reliable Operational Weather Forecasts.” *Nature Communications* 15: 4530.
- Leach, N. J., A. Weisheimer, M. R. Allen, and T. Palmer. 2021. “Forecast-Based Attribution of a Winter Heatwave Within the Limit of Predictability.” *Proceedings of the National Academy of Sciences of the United States of America* 118: e2112087118.
- Lee, R. W., A. J. Charlton-Perez, and S. H. Lee. 2025. “Stratospheric Impacts on Weather Regimes Following the 2018 and 2019 Sudden Stratospheric Warmings.” *Geophysical Research Letters* 52: e2025GL115668.
- Lin, Y. C., S. F. Jenkins, J. R. Chow, S. Biass, G. Woo, and D. Lallemand. 2020. “Modeling Downward Counterfactual Events: Unrealized Disasters and Why They Matter.” *Frontiers in Earth Science* 8: 575048.
- Lloyd, E. A., and T. G. Shepherd. 2020. “Environmental Catastrophes, Climate Change, and Attribution.” *Annals of the New York Academy of Sciences* 1469: 105–124.
- Lorenz, R., Z. Stalhandske, and E. M. Fischer. 2019. “Detection of a Climate Change Signal in Extreme Heat, Heat Stress, and Cold in Europe From Observations.” *Geophysical Research Letters* 46: 8363–8374.
- Lü, Z., F. Li, Y. J. Orsolini, Y. Gao, and S. He. 2020. “Understanding of European Cold Extremes, Sudden Stratospheric Warming, and Siberian Snow Accumulation in the Winter of 2017/18.” *Journal of Climate* 33: 527–545.
- Mitchell, D. M., L. J. Gray, J. Anstey, M. P. Baldwin, and A. J. Charlton-Perez. 2013. “The Influence of Stratospheric Vortex Displacements and Splits on Surface Climate.” *Journal of Climate* 26: 2668–2682.
- Mudhar, R., D. M. Mitchell, P. A. Stott, and R. A. Betts. 2023. “Reviewing Climate Change Attribution in UK Natural Hazards and Their Impacts.” Tech. rep., University of Bristol.
- Neal, R., D. Fereday, R. Crocker, and R. E. Comer. 2016. “A Flexible Approach to Defining Weather Patterns and Their Application in Weather Forecasting Over Europe.” *Meteorological Applications* 23: 389–400.
- Oehrlein, J., L. M. Polvani, L. Sun, and C. Deser. 2021. “How Well Do we Know the Surface Impact of Sudden Stratospheric Warmings?” *Geophysical Research Letters* 48: e2021GL095493.
- Otto, F. E. L. 2017. “Attribution of Weather and Climate Events.” *Annual Review of Environment and Resources* 42: 627–646.
- Rao, J., and C. I. Garfinkel. 2021. “CMIP5/6 Models Project Little Change in the Statistical Characteristics of Sudden Stratospheric Warmings in the 21st Century.” *Environmental Research Letters* 16: 034024.
- Rao, J., C. I. Garfinkel, T. Wu, Y. Lu, Q. Lu, and Z. Liang. 2021. “The January 2021 Sudden Stratospheric Warming and Its Prediction in Subseasonal to Seasonal Models.” *Journal of Geophysical Research: Atmospheres* 126: e2021JD035057.
- Rouges, E., M. Kretschmer, and T. G. Shepherd. 2025. “On the Link Between Weather Regimes and Energy Shortfall During Winter for 28 European Countries.” *Meteorological Applications* 32, no. 4: e70077. <https://doi.org/10.1002/met.70077>.
- S2S4E. 2020. “S2S4E Climate Services for Clean Energy Project (2020): S2S4E White Report: How Subseasonal and Seasonal Forecasts Can Help the Integration of Renewables Into Europe’s Energy Sector.” Tech. rep. <http://www.s2s4e.eu/>.
- Scaife, A. A., M. P. Baldwin, A. H. Butler, et al. 2022. “Long-Range Prediction and the Stratosphere.” *Atmospheric Chemistry and Physics* 22: 2601–2623.
- Schulzweida, U. 2023. “CDO User Guide (2.3.0).” Zenodo. <https://doi.org/10.5281/zenodo.10020800>.
- Seneviratne, S., X. Zhang, M. Adnan, et al. 2021. “Weather and Climate Extreme Events in a Changing Climate.” In *The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*.
- Seviour, W. J. M., J. Finkel, P. Rupp, et al. 2026. “Forecast-Based Attribution of the Role of Stratospheric Variability in Weather Extremes.” *Weather and Climate Dynamics* 7: 1405–1423.

- Sigmond, M., J. F. Scinocca, V. V. Kharin, and T. G. Shepherd. 2013. "Enhanced Seasonal Forecast Skill Following Stratospheric Sudden Warmings." *Nature Geoscience* 6: 98–102.
- Soares, M. B., and S. Dessai. 2016. "Barriers and Enablers to the Use of Seasonal Climate Forecasts Amongst Organisations in Europe." *Climatic Change* 137: 89–103.
- Soret, A., V. Torralba, N. Cortesi, et al. 2019. "Sub-Seasonal to Seasonal Climate Predictions for Wind Energy Forecasting." *Journal of Physics: Conference Series* 1222: 012009.
- Spaeth, J., and T. Birner. 2022. "Stratospheric Modulation of Arctic Oscillation Extremes as Represented by Extended-Range Ensemble Forecasts." *Weather and Climate Dynamics* 3: 883–903.
- Taylor, J. W., and R. Buizza. 2003. "Using Weather Ensemble Predictions in Electricity Demand Forecasting." *International Journal of Forecasting* 19: 57–70.
- Thompson, V., S. Ermis, and M. Athanase. 2025. "The Need for Multi-Method Extreme Event Attribution." *Weather* 81: 40–45.
- Thornton, H. E., B. J. Hoskins, and A. A. Scaife. 2016. "The Role of Temperature in the Variability and Extremes of Electricity and Gas Demand in Great Britain." *Environmental Research Letters* 11: 114015.
- van der Wiel, K., L. P. Stoop, B. R. van Zuijlen, R. Blackport, M. A. van den Broek, and F. M. Selten. 2019. "Meteorological Conditions Leading to Extreme Low Variable Renewable Energy Production and Extreme High Energy Shortfall." *Renewable & Sustainable Energy Reviews* 111: 261–275.
- Vitart, F., C. Ardilouze, A. Bonet, et al. 2017. "The Subseasonal to Seasonal (S2S) Prediction Project Database." *Bulletin of the American Meteorological Society* 98: 163–173.
- Vitart, F., and A. W. Robertson. 2018. "The Sub-Seasonal to Seasonal Prediction Project (S2S) and the Prediction of Extreme Events." *npj Climate and Atmospheric Science* 1: 3.
- Waugh, D. W., and L. M. Polvani. 2010. "Stratospheric Polar Vortices." In *Geophysical Monograph Series*, vol. 190, 43–57. American Geophysical Union.
- White, C. J., H. Carlsen, A. W. Robertson, et al. 2017. "Potential Applications of Subseasonal-to-Seasonal (S2S) Predictions." *Meteorological Applications* 24: 315–325.
- Wilcox, L. J., P. Yiou, M. Hauser, et al. 2018. "Multiple Perspectives on the Attribution of the Extreme European Summer of 2012 to Climate Change." *Climate Dynamics* 50: 3537–3555.
- Zhai, P., B. Zhou, and Y. Chen. 2018. "A Review of Climate Change Attribution Studies." *Journal of Meteorological Research* 32: 671–692.
- Zheng, C., Y. Wu, M. Ting, J. A. Screen, and P. Zhang. 2023. "Diverse Eurasian Temperature Responses to Arctic Sea Ice Loss in Models due to Varying Balance Between Dynamical Cooling and Thermodynamical Warming." *Journal of Climate* 36: 8347–8364.
- Zhuang, J., R. Dussin, P. Bourgauff, et al. 2025. "Pangeo-Data/xESMF: v0.9.2 - Third Time's a Charm (Version v0.9.2) [Computer Software]." Zenodo. <https://doi.org/10.5281/zenodo.17743552>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** met70241-sup-0001-Supinfo.pdf. $\overline{T'}$ (dark) and $T'_{\min}$ (light) for each country, with the near-central black bar marking the median, upper/lower the extrema, and white the mean. Within the country groups, countries are ordered L to R by increasingly negative mean $T'_{\min}$. **Figure S2:** As in Figure S1 but for E , with countries ordered by increasing mean $E'_{\max}$. **Figure S3:** As in Figure 6 of the main article but for individual northern European countries. **Figure S4:** As in Figure 6 of the main article but for individual central European countries. **Figure S5:** Relative risk versus quantile

shift corresponding to the 2018 NH SSW values for each of the individual countries within the (a) northern and (b) central European groups. Figures S3 and S4 show where the 2018 event lies in the distribution of extremes for each country. **Figure S6:** As in Figure 4 of the main article but for the other models (early initialisation). **Figure S7:** As in Figure 4 of the main article but for the other models (late initialisation). **Figure S8:** As in Figure S6 (early initialisation) for free-control for all models, including the multi-model mean (MMM). **Figure S9:** As in Figure S8 but for the late initialisation. **Figure S10:** As in Figure 8 of the main article but for the other models (early initialisation). **Figure S11:** As in Figure 8 of the main article but for the other models (late initialisation). **Figure S12:** As in Figure 9a of the main article but for the other models (early initialisation). **Figure S13:** As in Figure 9b of the main article but for the other models (late initialisation). **Figure S14:** As in Figure S10 (the early initialisation) but for central Europe and all models, including the multi-model mean (MMM). **Figure S15:** As in Figure S11 (the late initialisation) but for central Europe and all models, including the multi-model mean (MMM). **Figure S16:** As in Figure 10 of the main article but for central Europe.