

LETTER • OPEN ACCESS

## Skillful prediction of UK seasonal energy consumption based on surface climate information

To cite this article: Samuel Li *et al* 2023 *Environ. Res. Lett.* **18** 064007

View the [article online](#) for updates and enhancements.

You may also like

- [Stratospheric initial conditions provide seasonal predictability of the North Atlantic and Arctic Oscillations](#)  
Yu Nie, Adam A Scaife, Hong-Li Ren et al.
- [On the predictability of SSTA indices from CMIP5 decadal experiments](#)  
Dipayan Choudhury, Ashish Sharma, Bellie Sivakumar et al.
- [Singular vectors, predictability and ensemble forecasting for weather and climate](#)  
T N Palmer and Laure Zanna

ENVIRONMENTAL RESEARCH  
LETTERS

## LETTER

## OPEN ACCESS

## RECEIVED

30 November 2021

## REVISED

12 April 2023

## ACCEPTED FOR PUBLICATION

26 April 2023

## PUBLISHED

12 May 2023

Original content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.

Skillful prediction of UK seasonal energy consumption based  
on surface climate informationSamuel Li<sup>1,\*</sup>, Ryan Sriver<sup>2</sup> and Douglas E Miller<sup>2</sup><sup>1</sup> Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, United States of America<sup>2</sup> Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, United States of America

\* Author to whom any correspondence should be addressed.

E-mail: [swli2@illinois.edu](mailto:swli2@illinois.edu)**Keywords:** energy demand prediction, climate prediction, statistical modeling

## Abstract

Climate conditions affect winter heating demand in areas that experience harsh winters. Skillful energy demand prediction provides useful information that may be a helpful component in ensuring a reliable energy supply, protecting vulnerable populations from cold weather, and reducing excess energy waste. Here, we develop a statistical model that predicts winter seasonal energy consumption over the United Kingdom using a multiple linear regression technique based on multiple sources of climate information from the previous fall season. We take the autumn conditions of Arctic sea-ice concentration, stratospheric circulation, and sea-surface temperature as predictors, which all influence North Atlantic oscillation (NAO) variability as reported in a previous study. The model predicts winter seasonal gas and electricity consumption two months in advance with a statistically significant correlation between the predicted and observed time series. To extend the analysis beyond the relatively short time scale of gas and electricity data availability, we also analyze predictability of an energy demand proxy, heating degree days (HDDs), for which the model also demonstrates skill. The predictability of energy consumption can be attributed to the predictability of the NAO and the significant correlation of energy consumption with surface air temperature, dew point depression, and wind speed. We further found skillful prediction of these surface climate variables and HDDs over many areas where the NAO is influential, implying the predictability of energy demand in these regions. The simple statistical model demonstrates the usefulness of fall climate observations for predicting winter season energy demand prediction with a wide range of potential applications across energy-related sectors.

## 1. Introduction

Harsh winters and high energy costs are a normal part of life for many people around the world. As a society, we rely on a dependable energy supply to meet our heating demands during the wintertime. However, the winter seasonal mean energy demand varies strongly from year to year, closely related to the variability of weather conditions. For example, the extreme cold weather in the Midwestern United States on 30 January 2019 drove the Midwest natural gas consumption from less than 15 billion cubic feet per day ( $\text{Bcf day}^{-1}$ ) in early January to  $37.9 \text{ Bcf day}^{-1}$ —one of the most significant

single-day natural gas storage withdrawals recorded in the Midwest (U.S. Energy Information Administration Report). Energy security and infrastructure can be affected by extreme cold, as seen by the cold air outbreak in February 2021 that crippled power grids across Texas. The failed infrastructure, unprepared for the extreme cold, caused billions of dollars in damages, left millions of homes without power for up to several days, and created food and water shortages (e.g. Busby *et al* 2021). Skillful predictions of energy demand help ensure the security of energy supply by determining storage necessities and managing available supplies, which protects vulnerable populations from the dangers of extreme cold

weather. The effective matching of energy supply with demand also increases the efficiency of energy allocation, reduces excess energy waste and environmental pollution, and saves money.

Energy consumption is related to surface weather conditions. Although the day-to-day weather is not predictable beyond one or two weeks, some slowly varying components of the climate system (i.e. ocean, land surface conditions, and sea ice) and teleconnection patterns are predictable on the sub-seasonal and even longer time scales (e.g. Miller and Wang 2019a, Merryfield *et al* 2020). One important source of variability is the North Atlantic oscillation (NAO) (Hurrell *et al* 2003). The NAO, a dipole in atmospheric pressure between the Azores high and Icelandic low, strongly influences surface weather conditions over the United States and Europe. The NAO is the dominant mode of variability over the North Atlantic, and it can vary across a wide range of timescales including subseasonal to decadal. Recent studies have shed important light on the predictability of the NAO (Athanasiadis *et al* 2017), particularly for skillful prediction of European and North American wintertime temperature (Folland *et al* 2012, Scaife *et al* 2014). Although eddy feedback, Rossby wave breaking, and blocking are known to influence the NAO on the synoptic time scale (Hurrell *et al* 2003, Woollings *et al* 2010, Miller and Wang 2022), various sources of predictability influence the NAO on subseasonal to longer timescales, including sea surface temperatures (SSTs), stratospheric processes, and Arctic sea ice concentration (SIC). For example, the Atlantic SST tripole can influence the NAO by modulating eddy feedback (Pan 2005, Miller and Wang 2019a). Stratospheric signals can propagate into the troposphere influencing the NAO, specifically during sudden stratospheric warmings (Baldwin and Dunkerton 2001, Butler *et al* 2015, Scaife *et al* 2016). Modulation of SIC anomalies has been shown to influence negative feedback between SIC and the NAO (Deser *et al* 2007, Strong *et al* 2009, Strong and Magnúsdóttir 2011). These three important sources of variability have been used to predict the NAO (Wang *et al* 2017) and Eurasian atmospheric blocking frequency (Miller and Wang 2019b).

The relationship between surface weather conditions and energy usage has been previously utilized for energy prediction. Clark *et al* (2017) demonstrated the skillful prediction of winter energy demand is achievable based on the strong correlation between the observed Great Britain energy consumption and the NAO predicted by the Met Office Global Seasonal forecast system. Thornton *et al* (2019) showed that wintertime gas demand is highly correlated with wintertime temperature and can be predicted with some skill using winter seasonal forecasts of the large-scale atmospheric circulation.

Brayshaw *et al* (2019) constructed a robust long-term weather-related fault-rate climatology, which demonstrated large-scale dependencies and consequences point to the need for new methods to improve skill in the complex area of weather risk management. van der Wiel *et al* (2019) showed that blocked circulation patterns are associated with high energy demand and low renewable power production, thus a higher risk of energy shortfall, suggesting that large-scale weather regimes may be used for seasonal energy prediction. Bloomfield *et al* (2019) investigated the poorly understood winter meteorological drivers of European power systems using a new approach, targeted circulation types (TCTs). By convolving the weather sensitivity of an impacted system with atmospheric circulation, TCTs offer stronger ability in explaining power system variability and extreme events. In a related study, Bloomfield *et al* (2021) investigated high-quality meteorological forecasts in power system terms and presented an extensive power system forecast dataset with daily ensemble reforecasts and brief reviews of forecast skill. Ghalekhondabi *et al* (2017) provided an overview of methods used to predict energy consumption from 2005–2015, including both traditional (e.g. time series modeling) and more sophisticated (e.g. neural networks) methods. Notably, they advocated for simple methods with acceptable accuracies for future studies.

A few aforementioned studies employed sub-seasonal to seasonal (S2S) climate prediction by numerical models (i.e. dynamical prediction). S2S prediction skill has been improved in recent years, mainly when ensemble techniques are employed (e.g. Saha *et al* 2014, Vitart *et al* 2017, Merryfield *et al* 2020). However, seasonal dynamical forecasts are computationally expensive, and ensemble dynamical predictions produce datasets that require substantial amounts of computer resources to analyze and archive.

Previous studies show that statistical models can achieve skillful predictions of seasonal variability comparable to more sophisticated dynamical models (e.g. Barnston *et al* 2012, Huang *et al* 2019). This study presents a simple statistical prediction scheme, providing a skillful approach to predicting winter mean energy demand that is a cheaper alternative to dynamical numerical prediction methods, and can serve as a useful testbed for analyzing sources of seasonal predictability. We seek to predict the wintertime December–February (DJF) seasonal mean energy demand in the United Kingdom (UK), using gas consumption data and heating degree days (HDDs) (an energy demand proxy), combined with global climate indices from the previous September and October. We also explore links between surface air temperature and other surface climate conditions included in

the regression model to understand sources of variability and predictability of seasonal energy demand. The paper demonstrates a traceable example of how real-time observations of climate/weather conditions during the fall season can be utilized to predict wintertime energy demand based on climate proxies and real energy data, in addition to climate prediction, that can be used to inform regional seasonal energy planning.

## 2. Data and methodology

### 2.1. Gas and electricity demand data

Detrended daily UK gas and electricity consumption data from Thornton *et al* (2019) is used to construct a multi-linear regression (MLR) model and assess the model prediction skill. The gas data is available from 1996–2012, and the electricity data is available 1979–2012. We convert daily data into wintertime DJF seasonal averages, which is the predictand in our model. Both residential and large industrial premises, as well as shrinkage (gas leaks and theft), are included in the demand value. Thornton *et al* (2019) demonstrated that low-frequency variability in energy demand in the data period was not driven by weather conditions but likely related to socio-economic changes. Thus, much of this low-frequency variability was removed to accurately examine how weather impacts energy demand and its predictability. This correction was accomplished by replacing the gradual reduction in background energy demand with a climatological-mean annual demand cycle. The resulting detrended energy demand data is used in this study. Readers are referred to Thornton *et al* (2019) for more information.

### 2.2. HDDs

To extend the analysis beyond this time frame of gas and electricity data availability, we introduce an energy consumption proxy, HDD. This proxy fits a sinusoidal wave to daily minimum and maximum temperature and integrates below a threshold ( $20^{\circ}\text{C}$ ). The sinusoidal interpolation employed in this method provides more accurate estimations of energy consumption related to heating demand as compared to simple addition/subtraction functions used in other HDD calculations (D'Agostino and Schlenker 2016). To be consistent with the predictor data source, the HDD proxy is derived from the ERA-Interim reanalysis, which is available from 1979–2018. We found that the HDD proxy over the UK is significantly correlated with the observed UK gas consumption data with  $r = 0.68$  and  $p = 2.7 \times 10^{-03}$ , supporting the validity and potential of the proxy as a cross-check and extension of the energy data analysis.

### 2.3. Predictor selection

Our predictor selection builds from a recent study on the skillful prediction of the NAO (Wang *et al* 2017). (NOAA CPC NAO page: [www.cpc.ncep.noaa.gov/data/teledoc/nao.shtml](http://www.cpc.ncep.noaa.gov/data/teledoc/nao.shtml)). The NAO is the dominant mode of interannual variability over the North Atlantic. It is characterized by a north-south dipole pattern of sea level pressure and is associated with the basin-wide changes in the intensity and structure of the North Atlantic jet stream and storm track (Hurrell 1995). Wang *et al* (2017) used the empirical orthogonal function to extract the dominant modes of interannual variability of SIC and 70 hPa geopotential height (Z70) north of  $20^{\circ}\text{N}$ , and SST north of  $20^{\circ}\text{S}$ . These predictors were selected by Wang *et al* (2017) using the forward stepwise selection method, including the first principal component (PC1) of the October SIC, the PC2 of the October Z70, and the PC3 of the September SST. As discussed by Wang *et al* (2017), these predictors include information on the sea ice conditions in the Barents and Kara Seas, the Atlantic tri-polar SST mode, and the stratospheric circulation anomalies related to the El Niño Southern Oscillation (ENSO) and quasi-biennial oscillation. Using these predictors with an MLR model, Wang *et al* (2017) achieved robust and skillful prediction of the winter seasonal mean NAO index. Similar predictors were used to predict Eurasian winter atmospheric blocking frequency and extreme temperatures (Miller and Wang 2019b).

Given the substantial impacts of the NAO on weather conditions in western Europe, here we apply the same predictors for the NAO prediction as in Wang *et al* (2017) to predict the UK energy consumption. An updated predictor time series processed in the same way and extended to 1979–2018. Readers are referred to Wang *et al* (2017) for more information on the predictors. We focus on the winter seasons (DJF), and three time periods of analysis are taken based on data availability: the prediction of gas demand is explored from 1996–2012, electricity demand from 1979–2012, and HDD from 1979–2018.

### 2.4. Climate conditions over the UK

The 2 m temperature (T2m), 2 m dewpoint depression (TDD; i.e. the difference in degrees between air temperature and dew-point temperature; a lower TDD corresponds to a higher relative humidity for a given temperature), and 10 m wind speed from the ERA-Interim reanalysis are used to examine the impacts of surface climate conditions on UK energy consumption. The ERA-Interim data has spatial resolution  $0.7^{\circ} \times 0.7^{\circ}$  and is available from 1979–2018. The areal means over the UK ( $50^{\circ}\text{N}$ – $59^{\circ}\text{N}$ ,  $8^{\circ}\text{W}$ – $2^{\circ}\text{E}$ ) are calculated for relevant variables.

## 2.5. MLR model and leave-one-out cross-validation

We construct an empirical model for UK wintertime gas consumption prediction using the MLR technique with three selected predictors:

$$Y = a_1 \text{SIC} + a_2 \text{Z70} + a_3 \text{SST} + a_0,$$

where  $Y$  denotes the predictand and  $a_1$ ,  $a_2$ , and  $a_3$  are the regression coefficients for the predictors SIC, Z70 and SST, respectively. The **leave-one-out cross-validation method** is used to evaluate the prediction skill. For example, the observations for 1997–2012 would be used to develop the regression model, which is then used to predict the gas demand for winter 1996. This is repeated for every year from 1996–2012, yielding a time series of the predictand. **The prediction skill was then evaluated against the observed energy consumption using the Pearson correlation coefficient.** The cross validation for HDD is carried out for 1979–2018.

## 3. Results

### 3.1. Predicting the UK winter mean gas consumption

Figure 1 shows the observed and predicted time series of UK winter seasonal mean gas consumption (1996–2012), produced using the leave-one-out cross-validation method. The correlation between the observed and predicted time series is 0.50, which is above the 95% confidence interval based on Student's  $t$ -test. The correlation coefficient suggests that our simple prediction model captures about 25% of the year-to-year variance of energy consumption. While 71% of all space heating is done with gas, and only 8% with electricity (BEIS, [www.gov.uk/government/statistics/energy-consumption-in-the-uk-2020](http://www.gov.uk/government/statistics/energy-consumption-in-the-uk-2020)), we include the observed and predicted time series of UK winter seasonal mean electricity consumption (1979–2012) as well, calculated the same way, due to the high correlation between gas and electricity data ( $r = 0.83$ ) and the longer time span of electricity data availability. There is high agreement between both the observed gas and electricity time series and their corresponding prediction skills. More specifically, the model exhibits prediction skill with the observed energy consumption before 1989 and after 1996 but misses the extreme electricity anomalies in 1990 and 1995. These two years will be further discussed in the next section.

### 3.2. Physical mechanisms

In this section, we explore the physical mechanisms underlying the predictability of UK energy consumption. More specifically, we will address the following two questions: *what surface climate conditions are related to the UK gas and electricity consumption in*

*winter? How well can our predictors predict these surface climate conditions?*

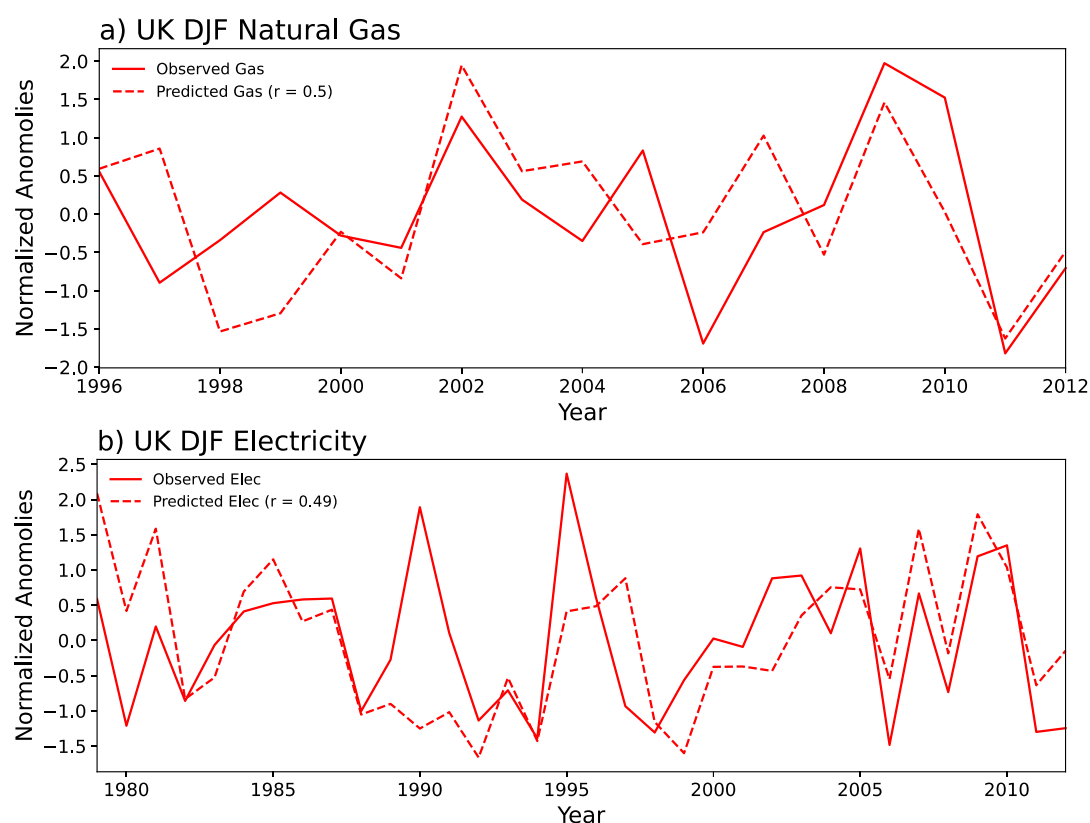
Figure 2 shows the observed and predicted T2m (figure 2(a)) and NAO (figure 2(b)) time series (1979–2012). The predicted T2m and NAO time series are produced using the same leave-one-out cross-validation method and predictors as for gas and electricity consumption prediction. The correlation between the observed and predicted T2m time series is 0.55 and above the 99% confidence interval (with a  $p$ -value less than 0.0001), suggesting that our prediction model captures about 30% of the observed variance of T2m. The correlation between the observed gas consumption and T2m (1996–2012) is  $-0.70$ , and above the 99% confidence interval (see table 1). The correlation suggests that T2m alone can explain about 50% of the variance of energy consumption, which provides the basis for climate condition-based prediction of energy consumption.

The simple MLR model also skillfully predicts the NAO (figure 2(b)), consistent with Wang *et al* (2017). The correlation between the observed and predicted NAO time series is 0.67 (with  $p$ -value less than 0.0001), suggesting that our prediction captures about 45% of the observed variance of the NAO. The correlation between the observed gas consumption and NAO (1996–2012) is  $-0.74$  (see table 1).

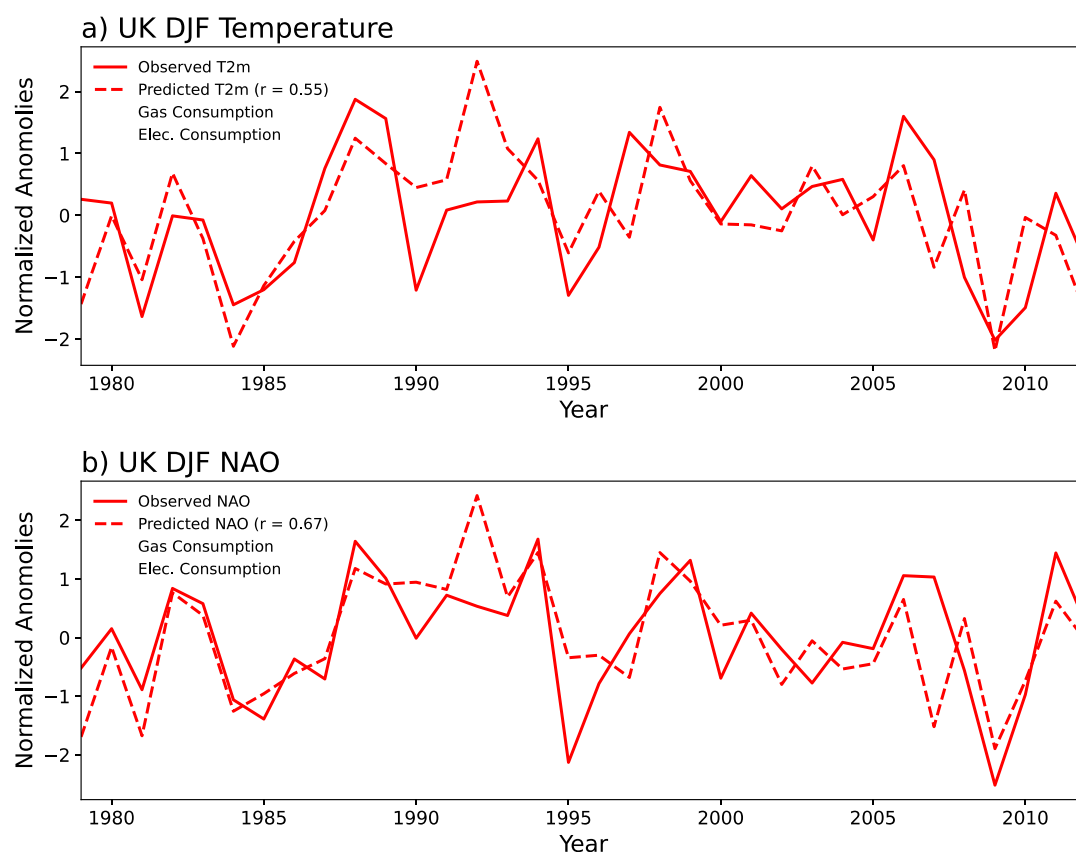
A key finding from figure 2 and table 1 is that both gas and electricity consumption are more strongly correlated with the NAO than with T2m. This implies that the NAO may provide more information than T2m alone in modulating energy consumption. To investigate this possibility, we examined surface climate conditions other than T2m. Both TDD and 10 m windspeed are significantly correlated with gas consumption (Braun *et al* 2014, Vu *et al* 2015; also see table 1). They are also significantly correlated with the NAO, as expected. Since temperature, dewpoint depression, and windspeed are related to the NAO, and all three indices are correlated with gas consumption, it explains why energy consumption is more strongly correlated with the NAO than with any single surface climate condition variable alone. The prediction skill of energy consumption thus arises from the predictability of the NAO and the joint estimation of these surface variables.

To explore the predictability of energy consumption outside the UK, we assessed the prediction skill of T2m, TDD, and wind speed using the same predictors and cross validation method over Eurasia and the eastern U.S. Figure 3 shows that skillful predictions of T2m, TDD, and wind speed are achieved in many regions. In particular, skillful prediction covers Greenland and the Eurasian continent between 50–70° N, and skillful prediction of surface wind speed is obtained over northwestern Europe. Such skillful





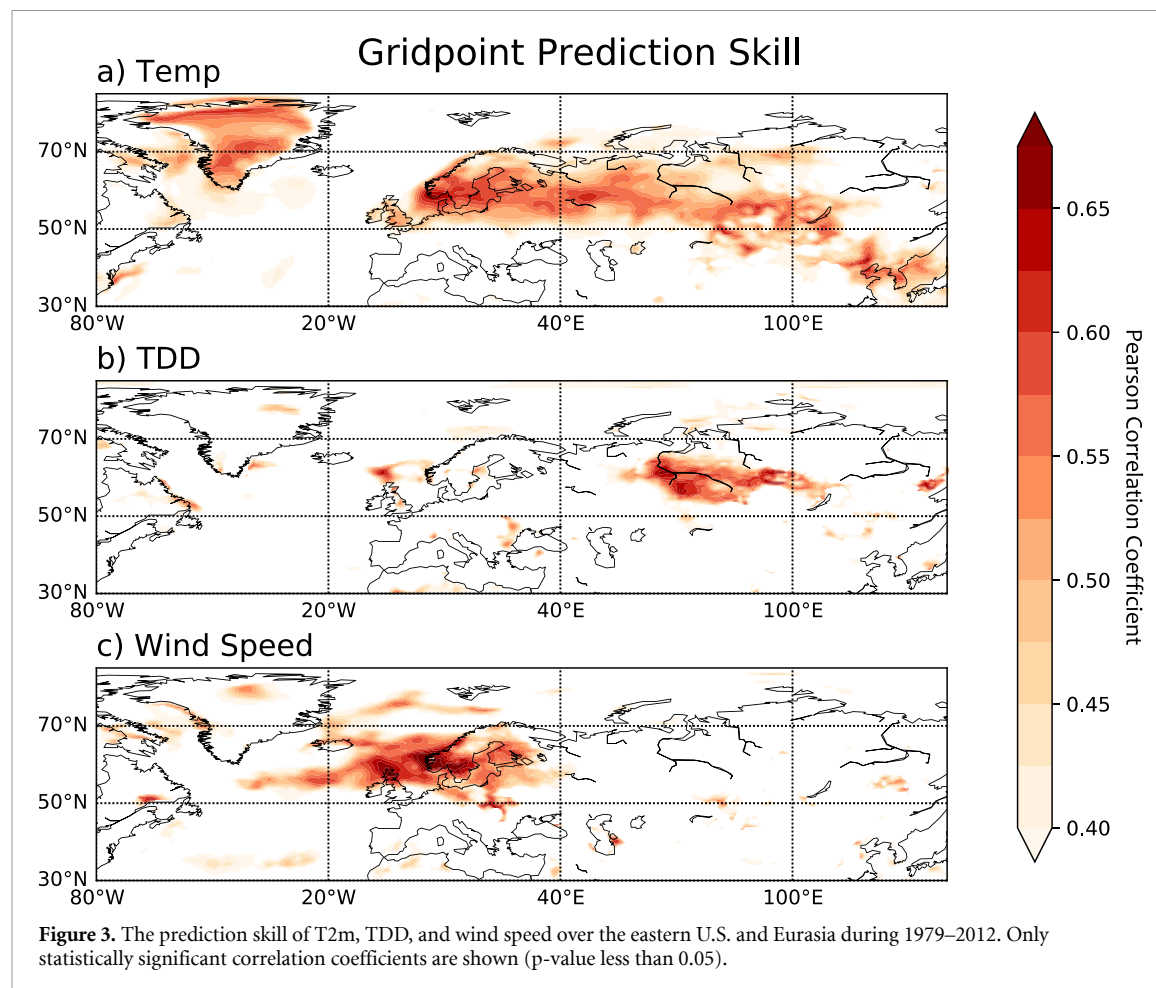
**Figure 1.** The observed (solid) and predicted (dashed) winter seasonal mean (a) natural gas (1996–2012) and (b) electricity (1979–2012) consumption anomalies in the UK. Anomalies are normalized by the corresponding means and standard deviations.



**Figure 2.** The UK observed (red solid) and predicted (red dashed) (a) T2m and (b) NAO along with the observed gas (black) and electricity (blue) consumption. The ACC between the predicted and observed time series is shown in each panel legend.

**Table 1.** Correlation and p-values between observed indices during 1996–2012. Values for electricity (row 3) are calculated over 1979–2012.

Corr. values	NAO	T2m	TDD	Wind
Gas consumption	$-0.74, 7.3 \times 10^{-04}$	$-0.70, 1.63 \times 10^{-03}$	$-0.70, 1.93 \times 10^{-03}$	$-0.65, 4.63 \times 10^{-03}$
NAO	1.0, 0.0	$0.76, 3.93 \times 10^{-04}$	$0.74, 6.23 \times 10^{-04}$	$0.79, 1.53 \times 10^{-04}$
Elec. consumption	$-0.70, 4.53 \times 10^{-06}$	$-0.57, 4.43 \times 10^{-04}$	$-0.37, 3.23 \times 10^{-03}$	$-0.42, 1.73 \times 10^{-02}$



prediction can be primarily attributed to the impacts of the NAO. Given the strong link between surface climate conditions and energy consumption, these regions may point to potential areas of further study.

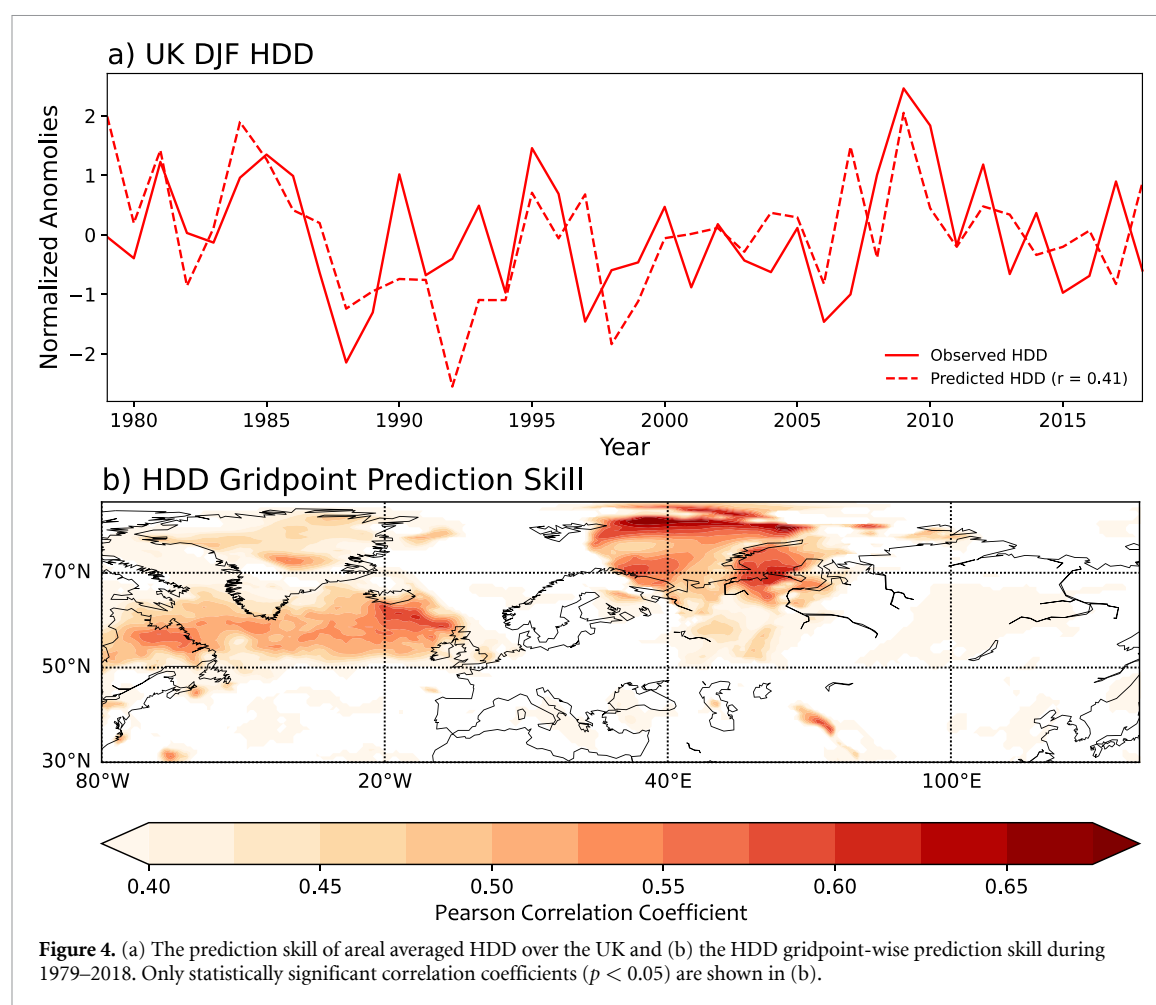
### 3.3. Prediction of HDDs

In the previous section we analyzed the predictability of energy consumption in northern hemispheric regions beyond the UK (figure 3). The investigation of energy demand prediction, however, is often limited by the availability of quality energy consumption data. In this section, we explore energy prediction using HDDs for the extended time range of 1979–2018. As introduced in section 2.2, HDD is an energy demand proxy based on seasonal duration of cold episodes below a temperature threshold. Figure 4(a) shows the time series of observed and predicted UK HDD using the same method as with the gas consumption prediction. The correlation between

the observed and predicted time series is 0.41 with p-value  $8.1 \times 10^{-03}$ , showing statistically significant prediction skill.

The prediction skill of HDD is then assessed using the same NAO predictors and cross validation method over the eastern U.S. and Eurasia (as in figure 3). Figure 4(b) shows that skillful predictions of HDD have large overlap with the regions of predictable surface climate conditions indicated in figure 3, and are achieved in many regions outside the UK, including Greenland and the Eurasian continent between 50–90° N. The skillful predictions support the rigor and usefulness of HDD as a simple energy proxy that can be utilizable in areas with limited energy consumption information.

In addition to ACC, we assessed the ability of the model to predict above the median, above the upper tercile, above the upper quartile, below the lower tercile, and below the lower quartile using the



**Table 2.** The HSS between observed and predicted time series for each energy index.

HSS	Lower quartile	Lower tercile	Median	Upper tercile	Upper quartile
HDD	0.20	0.09	0.10	0.34	0.47
Gas	0.15	0.23	0.29	0.23	0.15
Elec	0.24	0.46	0.65	0.36	0.24

Heidke skill score (HSS) (table 2) (Wilks 2011). HSS varies for different thresholds, which is similar to Thornton *et al* (2019), but it remains positive for all indices, indicating that our prediction is more skillful than random forecasts. In particular, despite slightly lower prediction correlation coefficient, HDD achieves strong HSS in the upper quantiles, both of which are significant at the 5% level. Significance is assessed using a 1000-member bootstrap: the HSS is calculated between the observed time series and randomly sampled (with replacement) hindcast timeseries. An HSS value is significant if it is greater than or equal to the 95th percentile of the corresponding resulting bootstrap distribution. The table shows that the model is particularly skillful in predicting the HDD upper quantiles and the above-median gas and electricity demands, as compared to a random reference forecast. In terms of energy planning, results suggest the models trained from different energy data/proxies may be used in tandem

to determine more confident predictions of (a) the above-median energy demand and (b) potential high energy demand seasons. Note that the gas HSS values are limited by a shorter time record.

#### 4. Summary and discussion

The number of high gas demand days per winter in the UK was demonstrated by Thornton *et al* (2019) to be predictable using UK winter seasonal forecasts. The NAO, which is highly influential in the UK, was demonstrated by Wang *et al* (2017) to be predictable by large-scale climate indices. In this study, we combine and extend findings from both studies to develop a simple statistical prediction model using MLR to analyze the predictability of UK energy consumption directly using large-scale climate indices. We skillfully predict the DJF gas and electricity consumption and the energy demand proxy HDD two months in advance: the correlations are 0.50, 0.49, and 0.4



between the observed and predicted time series of gas consumption (1996–2012), electricity consumption (1979–2012), and HDD (1979–2018), respectively. The predictability of UK energy consumption can be attributed to the predictability of the NAO and its impacts on surface air temperature, dew point depression, and windspeed, all of which affect energy consumption. While the prediction of gas and electricity consumption is focused on the UK, the investigation of the predictability of surface climate conditions and the HDD proxy indicates predictability of energy consumption in many other northern hemispheric regions.

The success of our simple statistical model lies in the choice of predictors. Although surface air temperature, humidity, and windspeed all have strong simultaneous correlations with the winter seasonal mean energy consumption, it is not feasible to use these variables in the preceding season directly as predictors because of their weak persistence. The three selected predictors, SST, SIC, and the large-scale stratospheric circulation, represent three major sources of predictability for the extratropical atmosphere and yield skillful prediction of the NAO and surface climate conditions.

It is worth briefly discussing the poor electricity model performance in winter 1990 and 1995. It is interesting to note that these two winters in the UK are notable for heavy snowfall and extreme cold temperatures. However, our simple statistical prediction model underestimates the negative T2m anomalies in 1995, and it completely fails to predict the negative T2m anomalies in 1990 (figure 2(a)), which explains the poor energy demand prediction in these two winters. The HDD model also mispredicts the peak in winter 1990, and both these winters occur before gas data is available. The cold winter in 1990 occurred with nearly normal NAO conditions (figure 2(b)), suggesting that the cold winter may be attributed to other climate modes. Further analysis indicates that the cold T2m anomalies may have been caused by a Scandinavia blocking high (not shown), one of the prominent weather regimes over the North Atlantic (Yiou and Nogaj 2004). Skillful prediction of weather regimes may help further improve the prediction skill of energy demand, and a longer record of energy consumption data will be useful to sort out the contributing factors for extreme events. Additionally, it is worth pointing out that a caveat of our model is that it does not predict non-weather-related changes in energy consumption.

One of the strengths of our study is the simplicity of our statistical framework. Such practice of using a simple prediction framework and simple proxies is well referenced and useful across many sectors of industry. While more complex models that achieve better energy predication skill exist, these models are often expensive and difficult to train and access. Our proposed model is fast and computationally cheap,

while still providing significant and meaningful prediction skill and information. A natural extension of this project, which is our ongoing work, is incorporating machine learning and exploiting other slowing climate components or teleconnection patterns as potential predictors. Machine learning can be utilized to account for the possible nonlinear relationship between the predictors and predictands. For example, long-short-term-memory models (Huang *et al* 2019) using daily time series of climate mode indices may produce skillful prediction of energy demand on the subseasonal time scale. Other possible methods include ensemble analog forecasting (Comeau *et al* 2017, 2019) and kernel analog forecasting (Wang *et al* 2020).

This project has useful implications for multidisciplinary areas, including applying weather/climate predictions to energy finance. Specific applications include forecasts toward the behavior of the economy and stock markets as influenced by extreme weather/climate events such as heatwaves, cold spells, hurricanes, droughts, and wildfires. Skillful weather/climate prediction can provide valuable and actionable information to stakeholders, government officials, and policymakers in many areas of society.

## Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: [https://github.com/swli2/seasonal\\_energy\\_pred](https://github.com/swli2/seasonal_energy_pred).

## Acknowledgments

SL was supported by the University of Illinois Campus Honors Program. We thank Dr Thornton for making the filtered energy data available and Dr Wang for providing the NAO predictor data. We also thank David Lafferty for providing the HDD calculations and reference. The authors have no conflicts of interest to declare.

## References

- Athanasiadis P J, Bellucci A, Scaife A A, Hermanson L, Materia S, Sanna A, Borrelli A, MacLachlan C and Gualdi S 2017 A multisystem view of wintertime NAO seasonal predictions *J. Clim.* **30** 1461–75
- Baldwin M P and Dunkerton T J 2001 Stratospheric harbingers of anomalous weather regimes *Science* **294** 581–4
- Barnston A G, Tippett M K, L'Heureux M L, Li S and DeWitt D G 2012 Skill of real-time seasonal ENSO model predictions during 2002–11: is our capability increasing? *Bull. Am. Meteorol. Soc.* **93** 631–51
- Bloomfield H C, Brayshaw D J and Charlton-Perez A J 2019 Characterizing the winter meteorological drivers of the European electricity system using targeted circulation types *Meteorol. Appl.* **2020** e1858
- Bloomfield H C, Brayshaw D J, Gonzalez P I M and Charlton-Perez A 2021 Sub-seasonal forecasts of demand and wind power and solar power generation for 28 European countries *Earth Syst. Sci. Data* **13** 2259–74

- Braun M R, Altan H and Beck S B 2014 Using regression analysis to predict the future energy consumption of a supermarket in the UK *Appl. Energy* **130** 305–13
- Brayshaw D J, Halford A, Smith S and Jensen K 2019 Quantifying the potential for improved management of weather risk using sub-seasonal forecasting: the case of UK telecommunications infrastructure *Meteorol. Appl.* **2020** e1849
- Busby J W, Baker K, Bazilian M D, Gilbert A Q, Grubert E, Rai V, Rhodes J D, Shidore S, Smith C A and Webber M E 2021 Cascading risks: understanding the 2021 winter blackout in Texas *Energy Res. Soc. Sci.* **77** 102106
- Butler A H, Seidel D, Hardiman S, Butchart N, Birner T and Match A 2015 Defining sudden stratospheric warmings *Bull. Am. Meteorol. Soc.* **96** 1913–28
- Clark R T, Bett P E, Thornton H E and Scaife A A 2017 Skillful seasonal predictions for the European energy industry *Environ. Res. Lett.* **12** 024002
- Comeau D, Giannakis D, Zhao Z and Majda A J 2019 Predicting regional and pan-Arctic sea ice anomalies with kernel analog forecasting *Clim. Dyn.* **52** 5507–25
- Comeau D, Zhao Z, Giannakis D and Majda A J 2017 Data-driven prediction strategies for low-frequency patterns of North Pacific climate variability *Clim. Dyn.* **48** 1855–72
- D'Agostino A L and Schlenker W 2016 Recent weather fluctuations and agricultural yields: implications for climate change *Agric. Econ.* **47** 159–71
- Deser C, Tomas R A and Peng S 2007 The transient atmospheric circulation response to North Atlantic SST and sea ice anomalies *J. Clim.* **20** 4751–67
- Folland C K, Scaife A A, Lindesay J and Stephenson D B 2012 How potentially predictable is northern European winter climate a season ahead? *Int. J. Clim.* **32** 801–18
- Ghalekhondabi I, Ardjmand E, Weckman G R and Young W A 2017 An overview of energy demand forecasting methods published in 2005–2015 *Energy Syst.* **8** 411–47
- Huang A, Vega-Westhoff B and Srivier R L 2019 Analyzing El Niño–Southern oscillation predictability using long-short-term-memory models *Earth Space Sci.* **6** 212–21
- Hurrell J W 1995 Decadal trends in the North Atlantic oscillation and relationships to regional temperature and precipitation *Science* **269** 676–9
- Hurrell J W, Kushnir Y and Ottersen G 2003 An overview of the North Atlantic oscillation *The North Atlantic Oscillation: Climatic Significance and Environmental Impact (Geophysical Monograph Series)* vol 134 (American Geophysical Union) pp 1–35
- Merryfield W J et al 2020 Current and emerging developments in subseasonal to decadal prediction *Bull. Am. Meteorol. Soc.* **101** E869–96
- Miller D E and Wang Z 2019a Assessing seasonal predictability sources and windows of high predictability in the climate forecast system, version 2 *J. Clim.* **32** 1307–26
- Miller D E and Wang Z 2019b Skillful seasonal prediction of Eurasian winter blocking and extreme temperature frequency *Geophys. Res. Lett.* **46** 11530–8
- Miller D E and Wang Z 2022 Northern hemisphere winter blocking: differing onset mechanisms across regions *J. Atmos. Sci.* **79** 1291–309
- Pan L-L 2005 Observed positive feedback between the NAO and the North Atlantic SSTA tripole *Geophys. Res. Lett.* **32** L06707
- Saha S et al 2014 The NCEP climate forecast system version 2 *J. Clim.* **27** 2185–208
- Scaife A A et al 2014 Skillful long-range prediction of European and North American winters *Geophys. Res. Lett.* **41** 2514–9
- Scaife A A et al 2016 Seasonal winter forecasts and the stratosphere *Atmos. Sci. Lett.* **17** 51–56
- Strong C and Magnusdottir G 2011 Dependence of NAO variability on coupling with sea ice *Clim. Dyn.* **36** 1681–9
- Strong C, Magnusdottir G and Stern H 2009 Observed feedback between winter sea ice and the North Atlantic Oscillation *J. Clim.* **22** 6021–32
- Thornton H E, Scaife A A, Hoskins B J, Brayshaw D J, Smith D M, Dunstone N, Stringer N and Bett P E 2019 Skillful seasonal prediction of winter gas demand *Environ. Res. Lett.* **14** 024009
- van der Wiel K, Bloomfield H C, Lee R W, Stoop I P, Blackport R, Screen J A and Selten F M 2019 The influence of weather regimes on European renewable energy production and demand *Environ. Res. Lett.* **14** 094010
- Vitart F et al 2017 The subseasonal to seasonal (S2S) prediction project database *Bull. Am. Meteorol. Soc.* **98** 163–73
- Vu D H, Muttaqi K M and Agalgaonkar A P 2015 A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables *Appl. Energy* **140** 385–94
- Wang L, Ting M and Kushner P J 2017 A robust empirical seasonal prediction of winter NAO and surface climate *Sci. Rep.* **7** 279
- Wang X, Slawinska J and Giannakis D 2020 Extended-range statistical ENSO prediction through operator-theoretic techniques for nonlinear dynamics *Sci. Rep.* **10** 2636
- Wilks D S 2011 *Statistical Methods in the Atmospheric Sciences* 3rd edn (Cambridge, MA: Academic) ch 8, p 676
- Woollings T, Hannachi A and Hoskins B J 2010 Variability of the North Atlantic eddy-driven jet stream *Q. J. R. Meteorol. Soc.* **136** 856–68
- Yiou P and Nogaj M 2004 Extreme climatic events and weather regimes over the North Atlantic: when and where? *Geophys. Res. Lett.* **31** L07202