Long-range forecasts for the energy market – a case study

Otto Hyvärinen, Antti Mäkelä, Matti Kämäräinen, and Hilppa Gregow

Finnish Meteorological Institute, Erik Palménin aukio 1, P.O. Box 503, 00101 Helsinki, Finland

Correspondence to: Otto Hyvärinen (otto.hyvarinen@fmi.fi)

Abstract. We examined the feasibility of long-range forecasts of temperature for needs of the energy sector in Helsinki, Finland. The work was done jointly by Finnish Meteorological Institute (FMI) and Helen Ltd, the main Helsinki metropolitan area energy provider, and especially provider of district heating and cooling. Because temperatures govern the need of heating and cooling and, therefore, the energy demand, better long-range forecasts of temperature would be highly useful for Helen Ltd. Heating degree day (HDD) is a parameter that indicates the demand of energy to heat a building. We examined the forecasted monthly HDD values for Helsinki using UK Met Office seasonal forecasts with the lead time up to two months. The long-range forecasts of monthly HDD showed some skill in Helsinki in winter 2015–2016, especially if the very cold January is excluded.

1 Introduction

Energy systems are probably in the biggest transition seen in their history. Many nations have agreed to significant reduction of greenhouse gas (GHG) emissions with international commitments, for example most recently with the Paris Agreement (UNFCCC, 2015). However, simultaneously with the agreed emission reductions there is a rapid increase in the energy demand, especially in developing economies. Therefore, there is a need for reliable, secure and more self-sufficient energy supplies and to balance the intermittency of extensive wind and solar power generation that requires different flexible energy systems. These factors combined have caused massive investments worldwide.

Weather and climate play an essential role in the energy systems of the future; they provide the potential (e.g., solar, water, wind, wave energy) but can also disturb the energy production (e.g., storms, heat, snow, freezing). For improved and novel production and usage of energy, improved and novel weather forecasts and applications are needed to ensure that the energy systems work at their best.

The work presented here was done jointly by Finnish Meteorological Institute (FMI) and Helen Ltd, the main Helsinki metropolitan area energy provider. This study is an example of the needs of the Finnish energy sector in the long-range forecasting. Helen Ltd produces and provides electricity and district heating and cooling in Helsinki. The production and consumption is largely governed by the temperatures, which means that the long-range forecasting of temperature would be highly useful for Helen Ltd if the forecast skill is good enough (i.e., better than a value based on climate). Spring and autumn are the most important times of the year for energy providers; decisions have to be made regarding how much energy has to be purchased and sold from/to stock.

Heating degree day (HDD) is a parameter that indicates the demand of energy to heat a building and, thus, it is an useful parameter for the providers of district heating, such as Helen Ltd. In this study, we examine the forecasted monthly HDD values for Helsinki using UK Met Office seasonal forecasts (MacLachlan et al., 2015).

2 Data and methods

2.1 Heating degree day

The HDD is a simple index of the heating demand for a building. It uses only the daily mean 2 m temperature (T_{mean}). The definition used in Finland is:

$$\textbf{if } T_{mean} < \text{Threshold } \textbf{then}$$
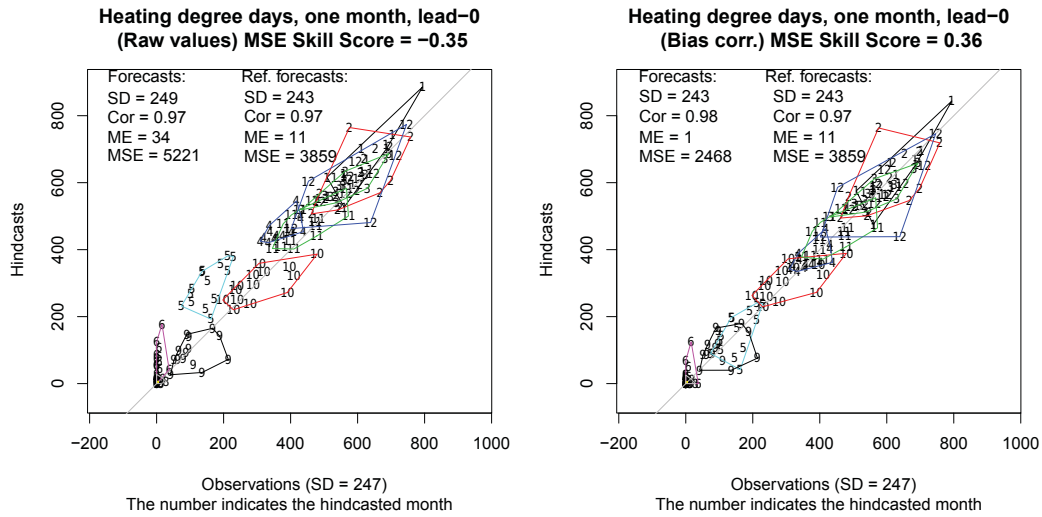$$\text{Index} = 17 - T_{mean}$$

Figure 1. Lead-0 hindcasts (1996–2009) of monthly HDD **(a)** without bias correction and **(b)** with bias correction. MSE Skill score for hindcasts is calculated against the values of reference forecasts. Hindcasts for the same calendar month are numbered and encircled.

```
   else
      Index = 0
   end if
```

where the threshold is $10\,°C$ in spring (from January to June) and $12\,°C$ in autumn (from July to December). These thresholds vary between countries (see, e.g., Valor et al., 2001). Daily index values are then summed for each month. Colder temperatures imply larger HDD values.

2.2　Temperature observations and forecasts

We used daily mean 2 m temperature observations and HDD reference values based on in-situ observations for Helsinki, Kaisaniemi (WMO 02978). The HDD reference values were for the normal period of 1981–2010 and made available by FMI (http://en.ilmatieteenlaitos.fi/heating-degree-days).

We had access to the Met Office seasonal forecasts and hindcasts from the seasonal forecast system GloSea5 (MacLachlan et al., 2015). Its atmosphere and land surface resolution is $0.833° \times 0.556°$. The same model is used both for hindcast and forecast members. Each day four members are run: Two members are run out to 210 days (seasonal forecast members) and two are run out to 60 days (intraseasonal members, not used in this study). The hindcasts, with three members per start date, have the following set of start dates: (1st, 9th, 17th, and 25th of each calendar month).

We chose the nearest grid point to Helsinki and picked daily mean 1.5 m air temperature values for this grid point from the hindcasts data set for years 1996–2009, that is 12 monthly hindcasts for 14 years and 168 hindcasts in total. Our deterministic hindcast was the mean of three hindcasts made for the first day of the month. Hindcasts were used to assess the predictive skill of the model to reproduce the ob-

served HDD values and for the simple bias correction of the operational HDD forecasts.

The operational forecasts were from March 2015 to August 2016 (18 monthly forecasts in total) and our deterministic forecast was the mean of six ensemble members (two seasonal forecast members per day from three days) for the same grid point as hindcasts. Note that this methodology is simpler compared to how the official Met Office seasonal ensembles are constructed (see MacLachlan et al., 2015, for details).

Forecasts (and hindcasts) for the first month (e.g., initialized 1 January and HDD calculated for January) are called lead-0 forecasts, and those for the second and third month are called lead-1 and lead-2, respectively.

3　Results

3.1　Bias correction using hindcasts

In Fig. 1 a, lead-0 HDD hindcasts and observed HDD values are shown. Hindcasts for the same calendar month are numbered and encircled. There is clear bias in hindcasts that is different in each calendar month, so a simple bias correction is performed by subtracting the mean monthly error from hindcasts (Fig. 1b). The bias of HDD was calculated for each calendar month and year separately, without the year to be corrected. In the next section, the mean of these monthly errors is used for bias correction of operational forecasts, also with different lead times. The squared mean of these biases is also the cross-validated mean squared error (MSE) of the monthly HDD.

The MSE of raw hindcasts was larger than the MSE of reference values, meaning that the end user would be better off using the reference values. But after bias correction,
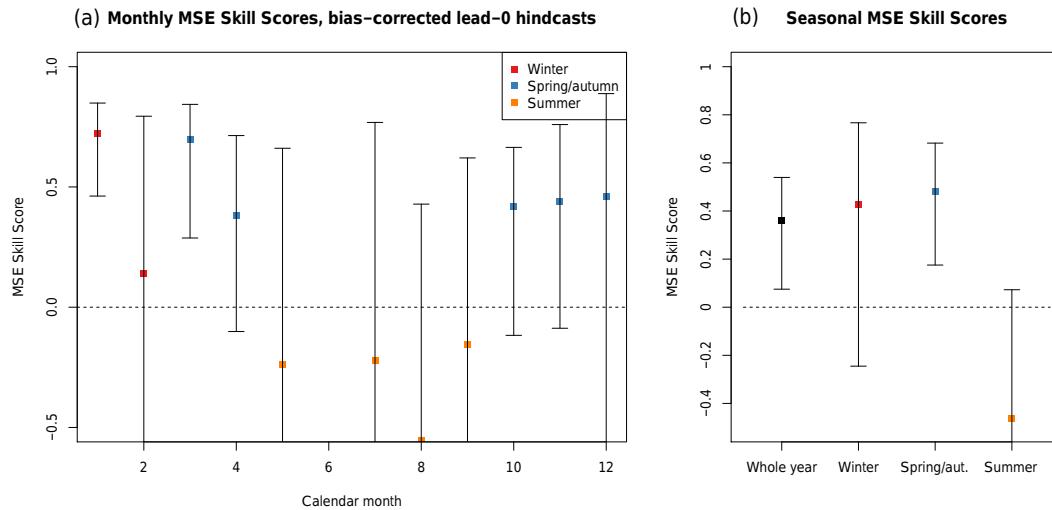
Figure 2. (a) Monthly and **(b)** seasonally aggregated MSE Skill score for bias-corrected lead-0 hindcasts (1996–2009) of monthly HDD. The MSE skill score for hindcasts is calculated against the values of reference forecasts. The 95 % confidence intervals by bootstrap ($R = 25\,000$) are shown. Note that, in **(a)**, the mean June value and its confidence interval are not missing, but outside the scale. The seasonal classification is as suggested by Helen Ltd.

the mean error (ME) drops, correlation (cor) improves and standard deviation (SD) approaches the observed standard deviation, resulting in smaller MSE for bias-corrected hindcasts than the MSE for reference values. In other words, bias-corrected hindcasts show skill, as the MSE Skill Score ($= 1 - \mathrm{MSE_{hindcasts}}/\mathrm{MSE_{reference}}$) is positive (0.36 with correction, -0.35 without correction).

In Fig. 1, the MSE skill score was calculated using the whole data set. In Fig. 2a, the MSE skill score is shown as the function of the calendar month of forecast initialization. The 95 % confidence intervals (CI) by bootstrap ($R = 25\,000$) are also shown. There is variance in the results, because the number of hindcasts per month is still rather small ($n = 14$), and it is unclear how skillful the bias-corrected hindcasts are, especially when the most CIs cover zero, the no-skill threshold. Therefore, months are combined into three groups as suggested by Helen Ltd: winter (January and February), spring/autumn (March, April, October, November, December), and summer (May, June, July, August, September) (Fig. 2b) and their MSE skill scores and CIs were calculated, also for the whole year. Now the CI of hindcasts of the whole year and spring/autumn are positive and do not cover zero, and can be said to be skillful. Bias-corrected hindcasts for summer are likely not to be skillful, because the mean value is very much less than zero. However, the CI of these hindcasts covers zero, and therefore, according to our current data, they are not statistically significantly unskillful. Hindcasts for winter have less data points than others and while the mean value is similar to the whole year and spring/autumn, the results are not statistically significant. The CIs of lead-1 and lead-2 hindcasts for the whole

year are negative and do not cover zero, even if some monthly intervals do (not shown).

3.2 HDD Using Operational Forecasts

In Fig. 3a lead-0, 3b lead-1 and 3c lead-2 operational forecasts of monthly HDD are shown with observations and reference values. As the lead time grows, the number of months forecasted drops. We focus on the winter from November 2015 to April 2016.

The winter was warmer than usual (meaning lower observed HDD values than reference values), except for a very cold spell in January (a rather high observed HDD value). For all lead times, the forecasted HDD is lower than the reference values, so in the forecasts there was a signal for a warmer than usual winter. But only the lead-0 forecasts showed a clear signal for a cold January. Indeed, the cold January degrades the forecasts; without January, forecasts of all lead times were clearly skillful (the MSE skill score is greater than zero), but with January included, only the lead-0 forecast showed skill. The MSE skill score 0.47 of lead-0 is inside the CI of lead-0 hindcasts, so there is not much difference between hindcasts and the forecast. On the other hand, skill scores of lead-1 and lead-2 are above the CI of hindcasts, implying somewhat better forecasts than hindcasts.

4 Discussion

Results of skillful forecasting of HDD are promising; even if only the lead-0 forecasts will turn out to be skillful in further analysis, they would be useful for many users. However, we analysed only one year of operative forecasts in one location.
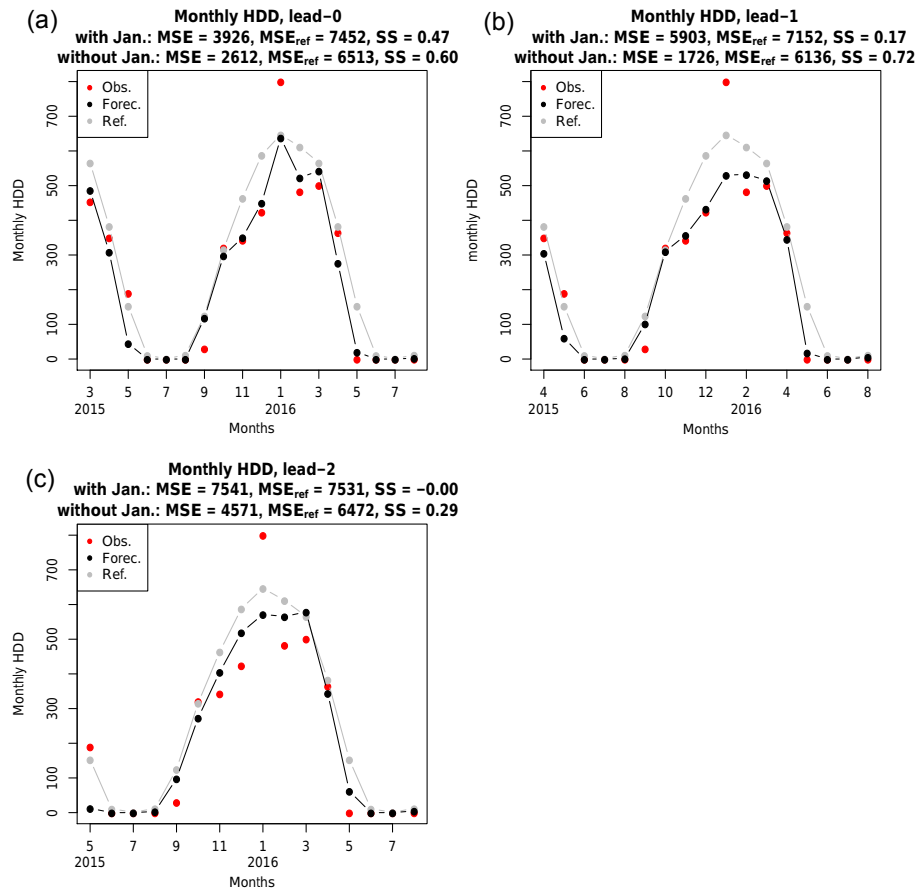
Figure 3. The **(a)** lead-0, **(b)** lead-1 and **(c)** lead-2 operational forecasts of monthly HDD from March 2015 to August 2016 for Helsinki.

So, the number of forecast-observation pairs is rather low, and the results are not statistically significant. The main challenge in validating operative forecasts (in contrast to hindcasts) is that it takes, for example, one year before one year of forecasts are available. Therefore in this study we considered only one year of operational forecasts, because we had no access to more operational forecasts at the time of analysis. Hindcasts can be used to assess the skill of forecasts, but as the number of included members is different, their skills are likely to be different.

Of course, this research should be replicated for other locations and using other forecast systems. In this, the data from the Sub-seasonal to Seasonal Prediction (S2S) project (Vitart et al., 2016) will be useful. In this study the forecasts were deterministic, but in further studies probabilistic methods should also be considered.

It can be argued that the definition of HDD should be updated, as the building stock in Finland has changed while the climate has been changing. In addition to this, the HDD is a rather coarse measure for engineers and most measures of meteorological verification are also not readily useful for them. Therefore direct modelling of the needed heating (and cooling) energy would be more beneficial. The difference in

forecasted and true energy consumption might be more relevant for engineers and more directly connected to the energy market.

The periods covered by hindcasts (1996–2009) and reference values (1981–2010) are different. In this study we did not explore how much effect this had on our results, but this should be done in future studies. Intuitively, using longer reference period should give a better estimate, unless the climate has changed so much that the reference values give too much weight on the old climatological conditions.

5 Conclusions

Skillful seasonal forecasts of temperature would be highly useful for Helen Ltd. The forecasts of the monthly HDD show some skill in Helsinki outside the summer season. This was most apparent in lead-0 hindcasts and in operative lead-0 forecasts. In longer lead times, the operative forecasts can show signal to the correct direction, but even a single month can degenerate results in a way that the numerical measures show no apparent skill. The results are promising, but much work remains to be done. In the context of energy markets, we can tentatively say that at least lead-0 forecasts show skill,

also in spring and autumn periods that are important for He-
len Ltd in their planning.

6 Code availability

The code for the analysis and figures is available on request
from the contact person.

7 Data availability

Reference values based on the normal period 1981–
2010 are available from http://en.ilmatieteenlaitos.fi/
heating-degree-days. Observations are available through
the FMI Open Data system https://en.ilmatieteenlaitos.fi/
open-data. The forecast data from Met Office is not publicly
available, the contact person for data is Adam Scaife. Some
data is available from S2S project, http://s2sprediction.net/.

Author contributions. Otto Hyvärinen analysed data and wrote
most of the text. Antti Mäkelä wrote Introduction and was the con-
tact person to Helen Ltd. Matti Kämäräinen helped with the bias
correction. Hilppa Gregow commented the early draft and was the
liaison to the Met Office.

Edited by: Á. G. Muñoz
Reviewed by: two anonymous referees

References

MacLachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fere-
day, D., Scaife, A. A., Gordon, M., Vellinga, M., Williams, A.,
Comer, R. E., Camp, J., Xavier, P., and Madec, G.: Global Sea-
sonal forecast system version 5 (GloSea5): a high-resolution sea-
sonal forecast system, Q. J. Roy. Meteor. Soc., 141, 1072–1084,
doi:10.1002/qj.2396, 2015.

UNFCCC: Adoption of the Paris Agreement, 21st Conference of
the Parties, Paris: United Nations, 2015.

Valor, E., Meneu, V., Caselles, V., Valor, E., Meneu, V., and
Caselles, V.: Daily Air Temperature and Electricity Load in
Spain, J. Appl. Meteorol., 40, 1413–1421, doi:10.1175/1520-
0450(2001)040<1413:DATAEL>2.0.CO;2, 2001.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M.,
Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M.,
Hendon, H., Hodgson, J., Kang, H., Kumar, A., Lin, H., Liu,
G., Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mas-
trangelo, D., MacLachlan, C., McLean, P., Minami, A., Mladek,
R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A. W.,
Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser, D.,
Woolnough, S., Wu, T., Won, D.-J., Xiao, H., Zaripov, R., Zhang,
L., Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen,
M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes,
M., Hendon, H., Hodgson, J., Kang, H., Kumar, A., Lin, H.,
Liu, G., Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mas-
trangelo, D., MacLachlan, C., McLean, P., Minami, A., Mladek,
R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A. W.,
Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser,
D., Woolnough, S., Wu, T., Won, D.-J., Xiao, H., Zaripov, R.,
and Zhang, L.: The Sub-seasonal to Seasonal Prediction (S2S)
Project Database, B. Am. Meteorol. Soc., doi:10.1175/BAMS-
D-16-0017.1, 2016.