Predicting Natural Gas Demand

Johann-Friedrich Salzmann
j.salzmann@students.hertie-school.org

Danial Riaz
d.riaz@students.hertie-school.org

Finn Krueger
f.krueger@students.hertie-school.org

1. Introduction

The Ukraine war and subsequent suspension of supplies of natural gas by Russia to Europe have resulted in "significant harm to consumers, businesses and entire economies," (Sadamori, 2022). Russia supplied about 40 percent of the European Union's gas consumption by pipeline, and those exports have been cut by 75 percent (Al-jazeera, 2022). In response to this, Europe needs to urgently not only diversify its gas supply but also set minimum storage obligations and implement energy saving measures for the coming winter. Europe's gas prices spiked to record highs in the third quarter of 2022 (IEA, 2022).

The situation has placed a number of sectors in an energy security risk, including transport, services, agriculture, forestry, food and fertilizer production, refrigeration and heating. Subsequently, this will also have an impact on commodity pricing and we are already seeing an 80 percent increase in the price of butter in August, while cheese was up by 43 percent, beef was 27 percent higher, and milk powder was up more than 50 percent (Al-jazeera, 2022). According to the International Energy Agency (IEA), in the European Union, Italian and German families are among the worst hit by surging gas prices. Experts are also seeing "huge levels of fuel poverty" which refers to when people are unable to keep their homes warm. The impact of the cold winters will be particularly harsh towards the poor, elderly and disabled i.e. the most vulnerable members of our community.

2. Motivation

As energy prices are soaring around Europe, policymakers are faced with the tremendous task to deliver support to households and businesses. Whilst price caps on electricity prices including gas can bring immediate financial relief to those in need, it artificially lowers the price of gas translating in higher demand and even higher risks of running out of gas supply. Any response to the energy crisis will have to result in the limiting and rationing of gas demand to make sure that households and business will not run of gas this winter. In our project, we have decided to apply Machine Learning models to forecast daily gas demand within Germany to contribute to a better understanding of how much gas will be demanded.

Machine learning models may have a higher degree of precision when it comes to forecasting in comparison to conventional models. We intend to develop a simple baseline model and to compare it with more sophisticated machine learning models and assess their predictive abilities. Although the gas industry has a long tradition in using classical and machine learning models for their forecasts, the aim of this project is to show that demand can be predicted in a fairly accurate way based only on publicly available, free-of-charge data and with limited computing resources.

Accurate forecasting of gas demand is also essential for market participants including suppliers, industry customers, and regulating authorities to be able to determine when the greatest contingencies need to be planned for. Potential shortages resulting from higher demand than aggregated supply require suppliers to buy additional gas at high prices in order to keep the system balanced.

3. Evaluation

We conducted an initial assessment by training a linear SDG regression to predict national daily aggregate gas demand of the last one year based on gas prices, weather (i.e, nationally aggregated minimum, mean maximum daily temperatures) and GDP in Germany.

The model is based on the assumption that national aggregate gas demand can be explained by a simple decomposition:

$$D = H + I + E$$

, whereas

- D is the total gas demand,
- H is the household (and SME) demand, mostly used for residential heating, we can assume it mostly to be dependent on temperature and not much on gas prices or economics indicators

- I is the industrial demand with regards to gas-intensive production; we assume industry market participants to react more sensitive to the overall economics situation, to short-term gas prices and not much on weather conditions

- E is the gas demand resulting from electricity generation of gas plants; despite this part of the demand has gone down following the scarcity of natural gas, it remains to be a relevant factor when electricity demand is peaking; we assume E to reflect an interaction of short-term gas and electricity prices.We found that, indeed, we are able to predict German aggregated national gas demand fairly well even when applying this straight-forward prototyping approach. The model performed well for both training and test data prediction.

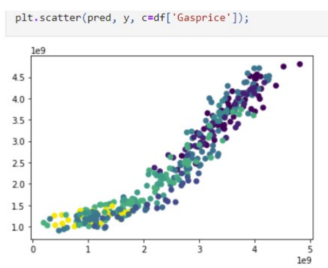

Figure 1. predicted vs. original gas demand, training data

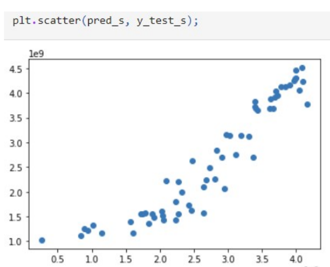

Figure 2. predicted vs. original gas demand, test data

To see the full initial analysis please see here.

This encourages us to tune the model by applying approaches of feature engineering (e.g., different numerical representations and projections of weather dynamics), imputation (accounting for different data resolutions), resampling, and autoregressive techniques.

Time series (TS) models are used to forecast the gas demand based on collected historical demand data. They can be applied for a wide range of forecasting horizons (from annual to hourly) (Zhang, 2020). We will build different TS models.

Autoregressive Integrated Moving Average (ARIMA) models have many uses in many industries. It is widely used in demand forecasting. ARIMA models use differencing to convert a non-stationary time series into a stationary one, and then predict future values from historical data. These models use “auto” correlations and moving averages over residual errors in the data to forecast future values. ARIMA models are generally good for short term prediction and require only prior data of a time series to generalize the forecast. They are however, poor long term forecasting models and not ideal for identifying unprecedented events. ARIMA model was applied to forecast annual or monthly gas demand of Turkey, with the consideration of GDP and price of gas (E. Erdogdu, 2010) .

Seasonal Auto-Regressive Integrated Moving Average with exogenous factors, or SARIMAX, is an extension of the ARIMA class of models. SARIMAX extends on this framework by adding the capability to handle exogenous variables. Unlike the ARIMA, it is able to capture seasonality in the model. This enables it to provide qualified annual forecasting for demand of gas

The project aim is to be able to nowcast and forecast German daily national gas demand. Depending on forecast data availability of our predicting variables, we may either use the forecast data to predict gas demand, or compute forecast approximators ourselves. We expect us to be able to do a decently well performing forecast of next-day and potentially next-week demand. Throughout the model refinement process, we will keep monitor how our models perform vs. what demand is actually realised few days later. Thereby, we will have the opportunity to make use of short feedback and learning cycles.

4. Resources

In order to predict the daily national gas demand in Germany we will utilise weather data including average, minimum and maximum temperatures from various cities within Germany. In addition to that we will base our prediction on economic data including GDP per quartile and daily market spot prices of gas. Our data includes all units on a daily basis for the last couple of years. We include lagged predictors in order to account for inertial systems, industrial processes, and market reaction times. More generally, we will assess the predictive strength of autoregressive models, i.e, by including gas demand as a predictor itself. We use and merge data from the following sources:

1. Daily national gas demand measured in kWh as well as the daily German market spot price in Euros retrieved from the European Energy Exchange (the European Energy Exchange (EEX) is an integration between Powernext and Gaspoint Nordic that publishes market data every business day after the conclusion of its settlement window) and the THE (Trading Hub Europe, joint company of all German regional gas net-

work providers).

2. Daily national electricity spot prices as published by the EEX and the ENTSO-E Transparency platform

3. Average, minimum and maximum daily temperature measured in Celsius taken from Frankfurt, Berlin, Munich and Germany retrieved from Meteostat (Meteostat is one of the largest vendors of open weather and climate data. The platform provides access to long-term time series of thousands of weather stations), allowing to feed the model with sufficient data on different weather conditions within the country. We might experiment with other approaches of weather data aggregation in order to increase models' predictive strength.

4. GDP per quarter retrieved from the German database of the Federal Statistical Office. (The Federal Statistical Office is a federal authority that reports to the Federal Ministry of the Interior and is responsible for collecting, processing, presenting and analysing statistical information concerning the topics economy, society and environment). As this is mostly used as a correction factor, we will assess whether we should increase data resolution, or may test different imputation strategies to account for economic trends.

5. Contributions

1. Data preparation: Danial and Finn will conduct the data preparation

2. Selecting and running training model: Joff and Finn

3. Evaluation and fine-tuning model: Danial

4. Final evaluation: All team

References

1. Al-Jazeera, "How bad could Europe's energy crisis get this winter?", 4th Oct 2022 https://www.aljazeera.com/news/2022/10/4/how-bad-could-europes-energy-crisis-get-this-winter

2. Erdogdu, E. (2010). Natural gas demand in Turkey.Applied Energy,87(1), 211-219.

3. J. D. Hunter, "Matplotlib: A 2D Graphics Environment" (2007), Computing in Science Engineering, vol. 9, no. 3, pp. 90-95

4. Pandas development team, T., 2020. pandas-dev/pandas: Pandas. URL: https://doi.org/10.5281/zenodo.3509134, doi:10.5281/zenodo. 3509134.

5. Pedregosa, F., Varoquaux, Gael, Gramfort, A., Michel, V., Thirion, B., Grisel, O., … others. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825–2830.

6. Sadamori, IEA Press Release, 3rd Oct 2022 https://www.iea.org/news/natural-gas-markets-expected-to-remain-tight-into-2023-as-russia-further-reduces-supplies-to-europe

7. Su, H., Zio, E., Zhang, J., Xu, M., Li, X., Zhang, Z. (2019). A hybrid hourly natural gas demand forecasting method based on the integration of wavelet transform and enhanced Deep-RNN model.Energy,178, 585-597."

8. Vinayak Sharma, Ümit Cali, Bhav Sardana, Murat Kuzlu, Dishant Banga, Manisa Pipattanasomporn (2021). Data-driven short-term natural gas demand forecasting with machine learning techniques. Journal of Petroleum Science and Engineering, (), –. doi:10.1016/j.petrol.2021.108979

9. van der Walt, S., Colbert, S.C., Varoquaux, G., 2011. The numpy array: A structure for efficient numerical computation. Computing in Science Engineering 13, 22–30.