Forecasting the energy consumption for Aarhus.

Manolis Giannitselis, Ólafur Hákon Sigurðarson

Supervisor: Eric Hillebrand

Aarhus University, 15. May 2024

Department of Economics and Business Economics

Abstract

This study focuses on forecasting energy consumption in Aarhus, using Vector Autoregression models with varying data granularities and seasonal dummies. We investigate whether higher data granularity-specifically analyzing energy consumption at the level of municipalities rather than broader regions-improves forecast accuracy, in different lags. Employing Diebold-Mariano and Clark-West tests our findings reveal that NUTS2, which included dummy variables for 7, 12 as well as holidays, provided more consistent and stable results.

Introduction

Forecasting energy consumption is a crucial component overall, which is connected to modern urban planning as well as sustainable development. Accurate predictions of energy demand enable different policy makers and utility companies to make informed decisions regarding resource allocation, energy conservation strategies and infrastructure investments. The city of Aarhus, which is Denmark's second largest city, presents a unique case study for energy consumption. That is due to its diverse economic and demographic characteristics along with the commitment which expresses towards sustainability and green initiatives.

In this project, the aim is to forecast the energy consumption in Aarhus using a Vector Autoregression model (VAR). The VAR is chosen for its robustness in capturing the dynamic relationships among multiple time series variables. In addition, it allows endogenous variables to be explained by their own past values, as well as by the past values of other variables within the system. This characteristic turns the VAR model particularly suitable for analyzing the interconnected and evolving patterns of energy consumption across different regions and periods.

A significant aspect of this research that is worth mentioning is the incorporation of seasonal dummies into the VAR model. Seasonal variants play a crucial role in energy consumption patterns, influenced by factors such as weather changes, economic cycles, and other societal behaviours. This way these periodic fluctuations can be captured, improving the model's predictive accuracy. The seasonal dummies represent different periods within a year, such as quarters or months, capturing the seasonal effects that might impact energy demand.

The primary question guiding this study is: Does data granularity increase forecast accuracy? More specifically, it is investigated whether examining energy consumption at the level of municipalities, as opposed to broader regions, or cities over municipalities etc., enhances the predictability of the VAR model. This inquiry stems from the hypothesis that more granular data provides a clearer resolution of the underlying consumption patterns, potentially even leading to more accurate forecasts.

The motivation behind this question lies in the complexity and localized nature of energy consumption. Different areas within a city or a region can exhibit different consumption behaviors due to numerous socioeconomic factors, industrial activities, and residential characteristics for instance, a residential neighbourhood may have different energy usage patterns compared to a more commercial district. By analyzing the data at a more detailed level, these local variations can be potentially captured more effectively, leading to better forecasts.

To address this question, we employ a VAR model that incorporates both lagged values of energy consumption and the dummy variables as exogenous variables. The endogenous variables represent the energy consumption levels across different regions, including Aarhus (DK751).

The expected outcome is to provide insights into the impact of data granularity on forecast accuracy. If more granular data indeed improves predictability, it could inform the design of more precise and localized energy management strategies. This could also enhance the efficiency and sustainability of energy use in Aarhus and potentially in other urban areas.

Literature Review

Globally, energy consumption is a critical factor. It has a direct bearing on both general well-being and production. The great economists who initiated the field of economics long ago didn't even address energy. Several theories and research projects still utilize the Cobb-Douglass Production Function (also known as the CDPF) as a foundational model. Although it was evident that the model could not account for more than half of the data, its purpose was clear. Energy may be included in K (capital), as many have claimed, but in the end, energy has no bearing on the Cobb-Douglas production function. This led to the proposal of the three-factor CDPF (Solow, R., 1956), which adds energy as a third independent factor.

The importance of energy has increased over time with increasing global energy consumption and demand (Kümmel et al., 2002). Energy's inclusion in the economic output in addition to capital and labor increased. In addition, multivariate cointegration analyses empirically support the important role of energy in macroeconomics and highlight the need for energy in economic forecasting models (Stern., 2000). Furthermore, the theoretical and empirical justifications for the addition of energy to conventional production functions showed how these modifications are consistent with both the dynamics of innovation diffusion and observed economic realities (Ayres and Warr., 2009).

There are specific insights into regional energy consumption trends which become invaluable, while the significance of energy continues to grow globally (Andersen et al. 2019), which in fact offer a critical context that underpins the understanding of Aarhus' specific energy needs and forecasting challenges. Moreover, the evolution of forecasting methods has made significant progress (Aye et al., 2016). The complexity inherent in the VAR model's design, which was used, was mirrored while highlighting the importance of selecting the optimal set of predictors that may impact the accuracy and reliability of the model.

Regional and local factors such as climate, population density, and infrastructure impact energy consumption. Trends, seasonality, and cyclical patterns can be captured in the data by ARIMA models, like in VAR models. Although these models differ in terms of accuracy under certain conditions, they can both shed light on how forecasting can shape policy decisions, particularly in terms of energy management and planning (Angelaccio, 2019).

Similarly, economic and policy frameworks are also impacted by tariff designs (Ansarin et al., 2020) and more of the overall energy consumption patterns can be explained as well. As there can be a deeper dive into energy usage and more specifically, electricity usage, the multifaceted nature of energy consumption drivers are revealed. This shows additionally how the more "the peeling" of the surface, the more of the consumer's behavior is revealed and therefore, the more there is to explain about the energy consumption.

There are additional challenges in estimating and forecasting energy consumption, which can be addressed by seasonal variations. There is a significant impact of seasonal factors on energy demand (Boehme et al., 2018). This aspect is crucial as it underscores the need to integrate and model seasonal variations to improve forecasting accuracy. It has also been shown that the optimization of data flow within networks can inform the data processing techniques used in the VAR model and ensure the efficient handling of extensive data (Fellaou et al., 2020) involved in regional energy forecasting. Of course, the incorporation of a mix of additional models such as ARIMA, MLP (multiplayer perception), and meter data could provide a significant improvement over the benchmark models, which is evidenced by substantial reduction in forecast errors (Karabiber and Xydis, 2020). In addition, the inclusion of exogenous variables like temperature and calendar events appeared to significantly improve the model performance.

Methodology

A fundamental explanation of energy consumption could also be achieved by introducing a VAR model. The data set used, which breaks down Denmark's hourly energy consumption by region, is quite extensive. For simplicity, this has been changed to daily data. After the VAR experiments, it was initially assumed that more exogenous variables would need to be included due to the high chi-square and extremely low p-values. The model had to consider the serial correlation of the residuals. Attempts made included using AIC and BIC to determine the ideal delay or even increasing the delay by one value at a time.

However, the residuals still showed a significant correlation. Holidays and temperature fluctuations also seem to have an influence on energy consumption. They therefore had to be integrated into the model.

The seasonality of the data can be explained in several ways. The effective use of exogenous variables improves the forecasting accuracy (Karabiber and Xydis 2020). To increase the validity of the model, dummies were added to account for seasonal patterns and other cyclical factors, as this action reflects the enhancements achieved in other research models. The dummies represented days, weeks, months and weekends.

Another compelling perspective is the one that applies regime change models in energy forecasting. The potential for incorporating complex interaction effects into energy consumption models was highlighted in a way that employs advanced data analytics to identify and adapt to multiple operational regimes (Kahraman et al., 2021). It is suggested that beyond seasonal and daily variations, additional regimes based on weather conditions or economic cycles could be modeled to refine the forecasting accuracy of the VAR model. However, for simplicity, only the seasonal dummies are retained.

Further, insights have been provided into the modeling of energy consumption under various heating systems, which can further improve the precision of energy forecasts at a municipality level. It has been shown that significant influences on household electricity consumption patterns are exerted by the type of heating system used (Lettau and Ludvigson, 2001). Certain data which can be adapted to similar environmental conditions in Aarhus, were presented, allowing for accurate seasonal adjustments through the incorporation of such differentiated effects in the VAR model, and therefore enhancing the model’s predictability and reliability as well.

When it comes to the assessment of forecasting models, especially when multiple models are to be considered, it becomes crucial to establish a statistical method, which is robust, and use it to compare their predictive accuracy. To this end, the Diebold-Mariano (D.M.) test is employed (Diebold and Mariano, 2002), offering a more comprehensive approach for evaluating whether the differences in the predictive performance in two models are statistically significant. The DM test works by focusing on the forecast error differences between two competing models across multiple periods of time. The null hypothesis of it suggests that the mean difference in forecast errors between those two models is zero, which implies that there is no significant difference in their respective predictive accuracy. However, it is suggested that there is a statistically significant difference in the forecasting performance of the two models under comparison, upon the hypothesis's rejection. The test statistic for this test is calculated by standardizing the mean forecast error (MFE) difference by its SE. In this context, the DM test is applied to compare the predictive accuracy of the VAR model, while being enhanced with seasonal (weekly/ daily) dummies against other potential models. This attempt adheres to rigorous empirical standards, ensuring the model performance being rigorous and statistically validated.

To account for the nested nature of comparing different lags, to further enhance the robustness of the model comparison and address the limitations of traditional forecast accuracy tests, the Clark and West adjusted statistic can be utilized (Clark and West, 2007). This test should be approximately normally distributed for model comparisons of nested forecasts. This is accomplished by comparing the MSFE of both models while accounting for the differences between the models.

Model Specifications

In the context of forecasting the energy consumption in Aarhus, the Vector Autoregression model (VAR) is used, due to its simplicity and robustness in capturing the dynamic relationships in multiple time series variables. In addition, it allows for endogenous variables to be explained by their own past values, as well as by the past values of other variables within it. VAR ensured that the influencing factors are thoroughly accounted for the same way feedback loops amongst different regions do.

To define the model formally, let $\mathbf{Y}_{\mathrm{t}}$ represent the vector of the endogenous variables, which in this case represents the energy consumption levels across *different regions*, including Aarhus (which is denoted as DK751). The vector can be denoted as follows:

$$Y_{\{t\}} = \begin{Bmatrix} Y_{DK751,t} \\ Y_{DK101,t} \\ \vdots \end{Bmatrix}$$

Let $\mathbf{X}_{\mathrm{t}}$ represent the vector of exogenous variables, which includes the dummy variables, that are hypothesized to influence energy consumption. The $\boldsymbol{X}_t$ can be expressed as:

$$\boldsymbol{X}_t = \begin{pmatrix} X_{1,t} \\ \boldsymbol{X}_{2,t} \\ \vdots \end{pmatrix}$$

The coefficients of the lagged endogenous variables are captured in the coefficient matrices $\mathbf{A}_{\mathrm{i}}$ where i ranges from 1 to p, the number of lags. These matrices represent the effect of the past values of the endogenous variables on their current values on their current values and are crucial in understanding the temporal dynamics of the system. The matrix $\mathbf{A}_{\mathrm{i}}$ is of n x n dimensions, where n is the number of exogenous variables. It is expressed as:

$$\mathbf{A}_{\mathrm{i}} = \begin{pmatrix} a_{11,i} & a_{12,i} & \cdots & a_{1n,i} \\ a_{21,i} & a_{22,i} & \cdots & a_{1n,i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1,i} & a_{n1,i} & \cdots & a_{nn,i} \end{pmatrix}$$

Where $a_{jk,i}$ is the coefficient for the ith lag of the kth endogenous variable in the equation for the jth endogenous variable.

The coefficients of the exogenous variables are captured in the matrix $\boldsymbol{B}$, which has dimensions n x k. The impact of exogenous variables on the endogenous ones is captured in the matrix $\boldsymbol{B}$**:**

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{pmatrix}$$

Where b_{jk} represents the coefficient for the kth exogenous variable in the equation for the jth endogenous variable.

Let $\mathbf{C}$ be the vector of intercept terms. This vector represents the constant terms in the VAR equations and is expressed as follows:

$$\mathbf{C} = \begin{pmatrix} \boldsymbol{c_1} \\ \boldsymbol{c_2} \\ \vdots \\ \boldsymbol{c_n} \end{pmatrix}$$

Let $\mathbf{E}_{\mathrm{t}}$ be the vector of the residuals. It captures the error terms for each equation at time t. The unexplained variability in the endogenous variables is explained by this vector. In addition, it is essential for capturing the stochastic nature of the system:

$$\mathbf{E} = \begin{pmatrix} \boldsymbol{e_{1,t}} \\ \boldsymbol{e_{2,t}} \\ \vdots \\ \boldsymbol{e_{n,t}} \end{pmatrix}$$

Where e_j, t represent the residual for the jth endogenous variable at time t.

Reduced-Form VAR Model:

The reduced-form VAR model can be shown as follows:

$$\mathbf{Y}_{\mathrm{t}} = \mathbf{A}_1\mathbf{Y}_{\mathrm{t}-1} + \mathbf{A}_2\mathbf{Y}_{\mathrm{t}-2} + \cdots + \mathbf{A}_{\mathrm{p}}\mathbf{Y}_{\mathrm{t}-\mathrm{p}} + \mathbf{B}\mathbf{X}_{\mathrm{t}} + \mathbf{E}_{\mathrm{t}}$$

The current value of the endogenous variables $\mathbf{Y_t}$ is influenced by their past values, captured by $\mathbf{A_i}$ matrices, the current and past values of the exogenous variables $\mathbf{X}_\mathrm{t}$ and the residuals $\mathbf{E}_\mathrm{t}$.The parameters of the model, including the coefficients $\mathbf{A}_\mathrm{i}$ and $\mathbf{B}$, as well as the intercepts in $\mathbf{C}$, are estimated using historical data on energy consumption and other relevant variables.

A more comprehensive understanding of the factors driving the energy consumption in Aarhus can be achieved by employing the VAR model. It allows the examination of direct and indirect effects on various predictors, easing a tight analysis.

Prediction of specific variable: DK751 (Aarhus)

The task to predict energy consumption for Aarhus, designated as DK751 and involves isolating its specific equation from the overall VAR model. This detailed prediction model accounts for various factors influencing energy consumption over time and across different regions. By focusing on DK751, we aim to understand how past energy usage, external variables, and interactions with other regions impact Aarhus's energy consumption.

To predict the dependent variable for Aarhus, the following components are included in the model:

Past values of DK751 from previous lags, each with their own coefficients.

Contributions from other endogenous variables at past lags that affect DK751.

Current contributions from the exogenous variables through the B matrix.

An error term capturing the residual variability.

The prediction equation for DK751 is formulated as:

$$Y_{DK751,t} = c_{DK751} + \alpha_{DK751} \sum_{i=1}^{p} Y_{DK751,t-i} + \sum_{j \neq DK751} \beta_j \sum_{i=1}^{p} Y_{j,t-i} + \gamma_{DK751} X_t + e_{DK751,t}$$

c_{DK751}: Represents the intercept term for DK751, which also summarizes the baseline level of energy consumption when other factors are held constant.

α_{DK751} and β_j reflect the coefficients for DK751's own lags and other endogenous variables respectively. The vector γ_{DK751} accounts for the coefficients of the exogenous variables, and $e_{DK751,t}$ is the residual error term for DK751.

The Data

The data used come from TSO electric which is part of the Energinet Group that own and operate the overall electricity transmission systems in Denmark. Since Energinet is owned by the Danish Ministry of Energy, Utilities and Climate it is free to copy, change and distribute the data. The base data contains energy consumption in kWh split into three categories: Industry, Public and Private for the years 2021-2023. It is available both in a monthly format and an hourly format.

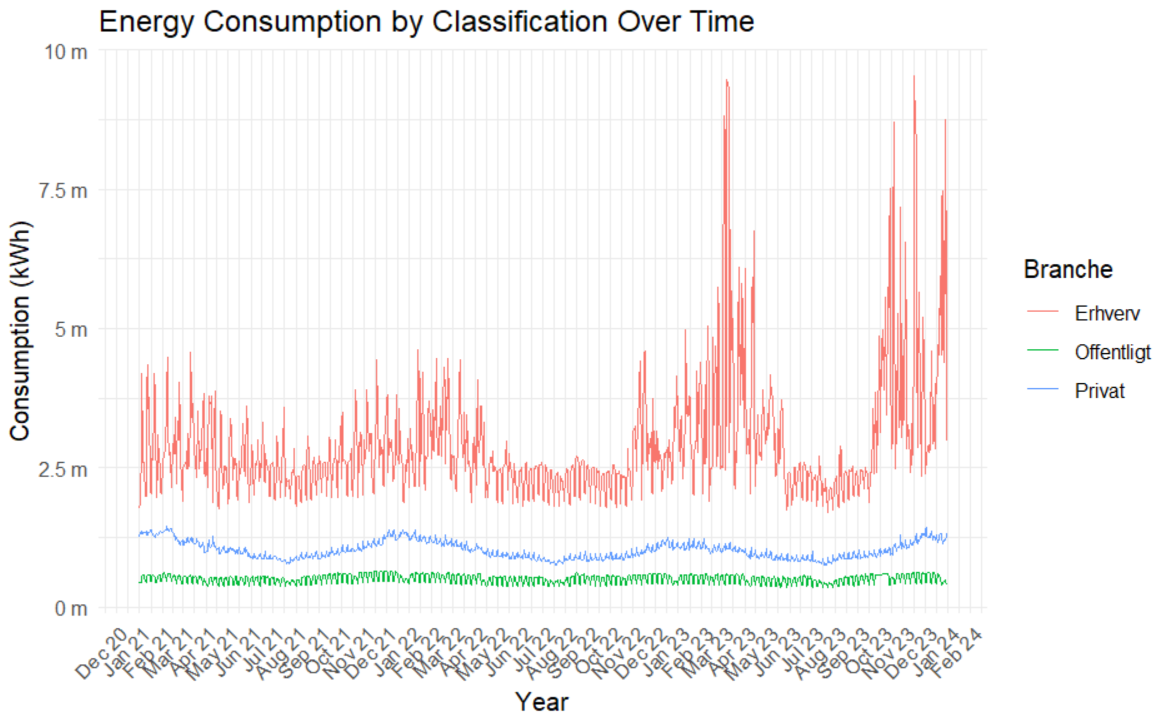

Figure 1: Total Energy Consumption in Aarhus by hour.

Source: Energinet (www.energidataservice.dk)

A very irregular industry (erhverv) consumption can be observed as opposed to public and private consumption. Private consumption looks to be more seasonal than the public sector, which is more stable. The instability of the industry consumption can most likely be accounted for in the description of the data. As reported in the metadata information, "Units connected to a CVR number in the other of Statistics Denmark's categories, including category 11 ('Uoplyst aktivitet') are labeled 'Erhverv' (industry). - Other units (without CVR numbers) are labeled as 'Privat' (private)" (Energidata Service, 2022). The decision to focus on private consumption is mostly driven by this discrepancy between the different energy users.

After inspection and narrowing down the focus to private consumption next up is looking at the differences between other descriptors such as the municipalities. We can also see how the greatest consumers of private electricity of the Danish municipalities the more urban areas of Copenhagen, Aarhus and Aalborg are of course. This is consistent over all three observed years as shown in figure 2 in the appendix.

Granularity

We first took a quick look at how regions are determined and grouped in Denmark according to EU standards. They are split into two: Nomenclature of Territorial Units for Statistics (NUTS) and the local administrative unit (LAU).

Level	Subdivisions	#
NUTS 1	-	1
NUTS 2	Regions (Danish: Regioner)	5
NUTS 3	Provinces (Danish: Landsdele)	11
LAU 1	Municipalities (Danish: Kommuner)	99
LAU 2	Parishes (Danish: Sogne)	2143

Table 1: DK subdivisions

Our data on energy usage is in the granularity of municipalities (LAU1) so we have 99 stochastic processes to work with. The data was then transformed by aggregation into these lower granularities moving from 99 regional variables down to the 11 provinces (NUTS3) and then down to the 5 regions (NUTS2). Since the dependent variable of Aarhus (DK751) is at the subregional level of LAU1 an issue of multicollinearity will arise if not accounted for where the time series of DK751 will be present both in DK751 and its own region DK042 East Jutland. It's possible to either remove the whole NUTS3 or NUTS2 regions in which Aarhus is situated but by subtracting DK751 from its region this can be avoided.

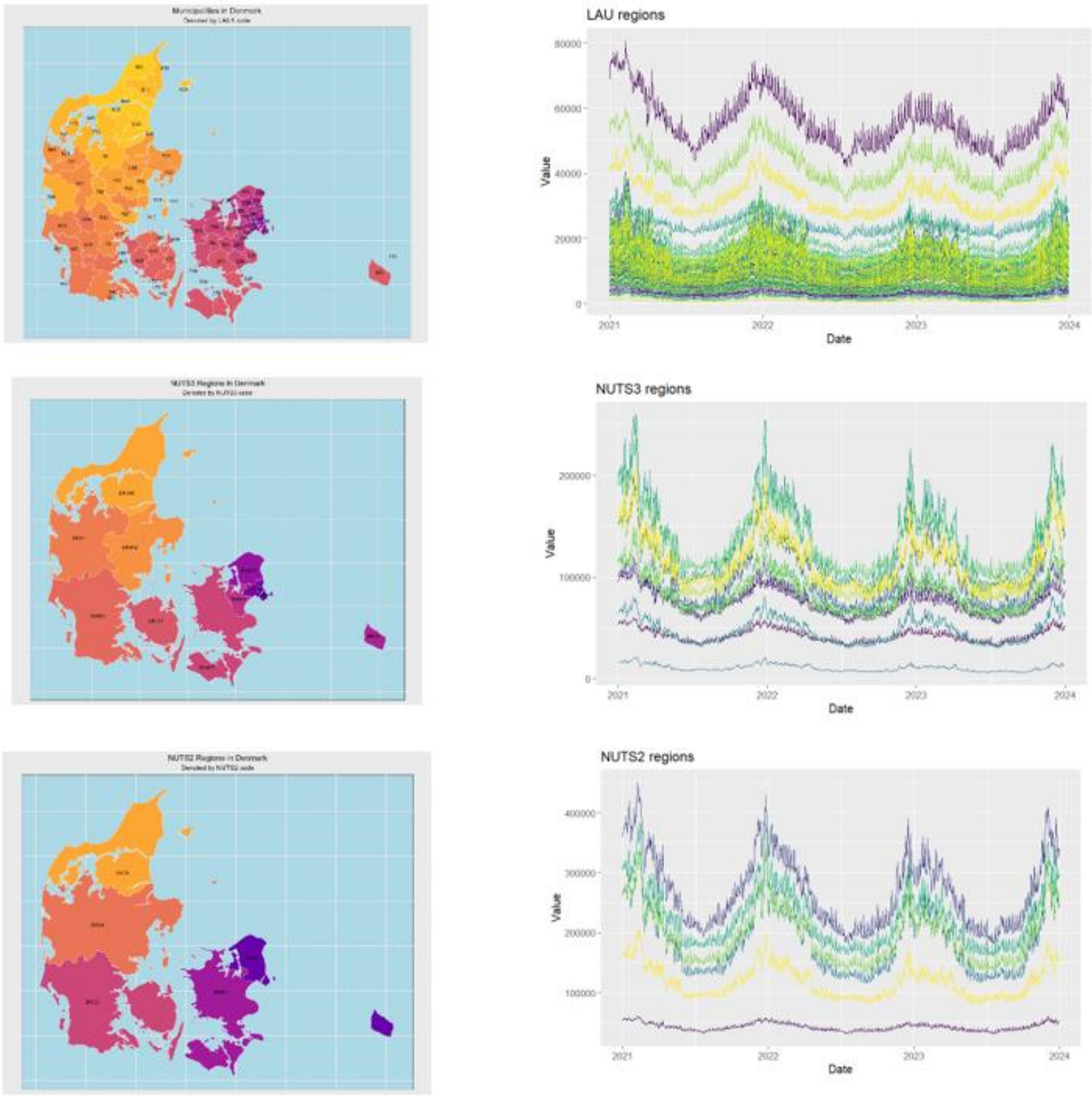

Figure 2:

As we can see the NUTS 2 codes are denoted by a double integer 01-05 whilst the NUTS 3 codes are denoted by a triple integer 011-050 and LAU codes are denoted by a triple integer as well ranging from 101-860. As seen on figure X we can see a general trend of increase as we go East-West in a clockwise manner.

On figures 5-6 we can also see the effect of aggregation and how seasonal the data really is.

Methodology: Static window sample splitting

In the model estimation-and-forecasting, a static window approach is employed, which has three equally sized yearly windows to estimate, forecast, and calculate performance metrics. The data was split into three distinct periods as seen on figure 4:

Presample Period (blue): January 1, 2021 – December 31, 2021

Estimation Period (red): January 1, 2022 – December 31, 2022

Forecast Period (green): January 1, 2023 – December 31, 2023

This choice is grounded in the importance of using Pseudo-Out-of-Sample (Pseudo-OOS) forecasting and practical constraints, aiming to balance model simplicity with robust performance evaluation.

In-sample forecasting makes predictions within the same period used for model estimation, which can lead to overfitting and overly optimistic forecasts. In contrast, Pseudo-OOS forecasting provides a more realistic assessment by mimicking the conditions the model would face in a true out-of-sample forecasting environment. This approach guards against the look-ahead bias inherent in in-sample evaluations.

The generous presample period offers flexibility with lag choice and reduces practical constraints. While the modeling can accommodate a smaller presample period, evenly spacing the periods ensures that any chosen seasonal dummies complete their cycles. This avoids potential distortions in the coefficients of the dummies, which can occur if the estimation period includes partial cycles. The design enhances the robustness of our forecasts by allowing comprehensive model calibration and validation.

Forecasting a full year ahead (n = 365) aligns with macroeconomic standards and provides valuable insights for long-term economic planning and policy-making, as demonstrated by Stern (2000) in his analysis of the causal relationship between GDP and energy use.

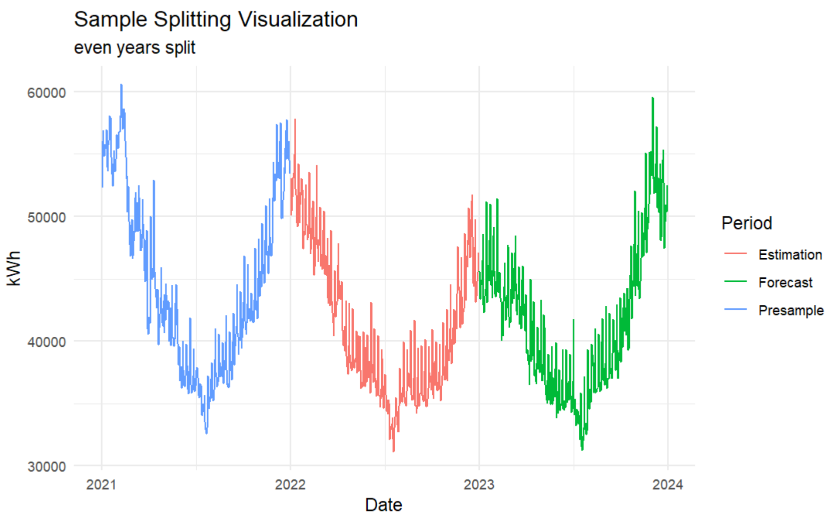

Figure 3

Stationarity

Before starting any sort of modeling stationarity checks had to be performed. ADF-tests were performed before and after taking the first difference for every municipality. The full results table of each region can be found as tables 5 and 6 in the appendix.

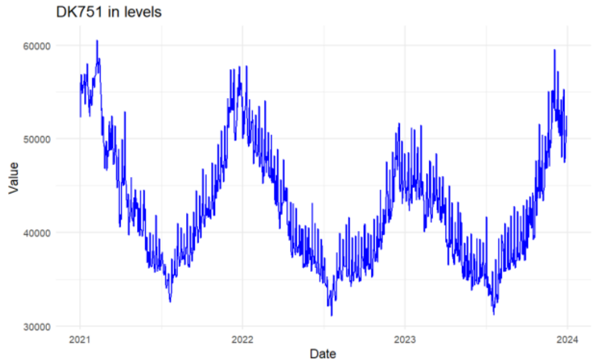

Figure 4

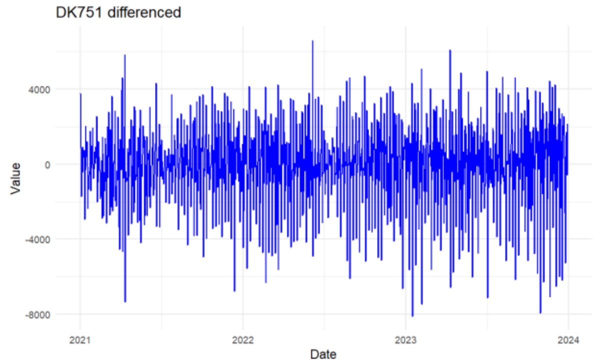

Figure 5

Region	DF_t	P value	Lag used	method	result
DK751	-2.04	0.56	10	ADF-test	non-stationary
ΔDK751	-10.77	0.01	10	ADF-test	stationary

Table 2:

Because of the apparent cyclical nature of the data attempts at weekly and yearly differencing was made, as seen in figures 7 and 8, but the results exhibited higher variance which would negatively affect forecasting.

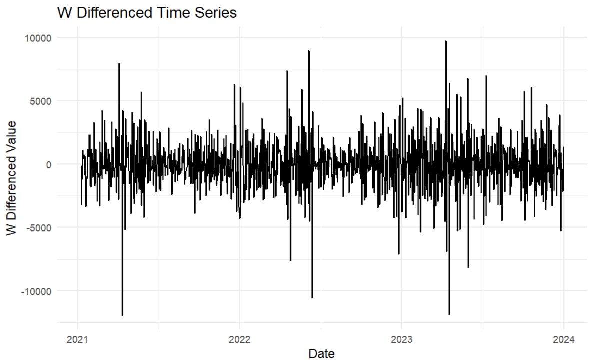

Figure 6

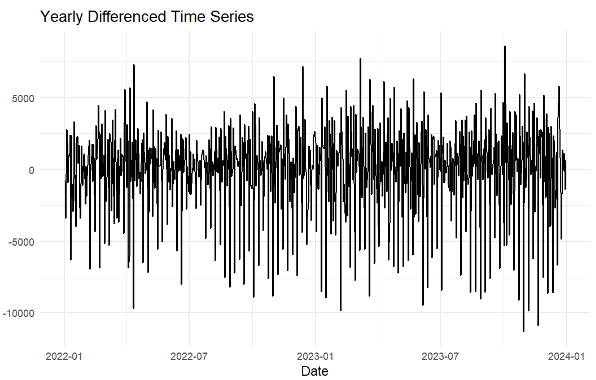

Figure 7

Dummies and spikes

A basic outlier detection was also performed to get a rough idea of if outliers had some obvious grouping which could be accounted for in some way in the model. The outlier detection calculated rolling standard scores (z-scores) with a more lenient threshold of 2. As seen in figure 7 most of the outliers are seen on weekends which on average see a higher consumption when people are more at home.

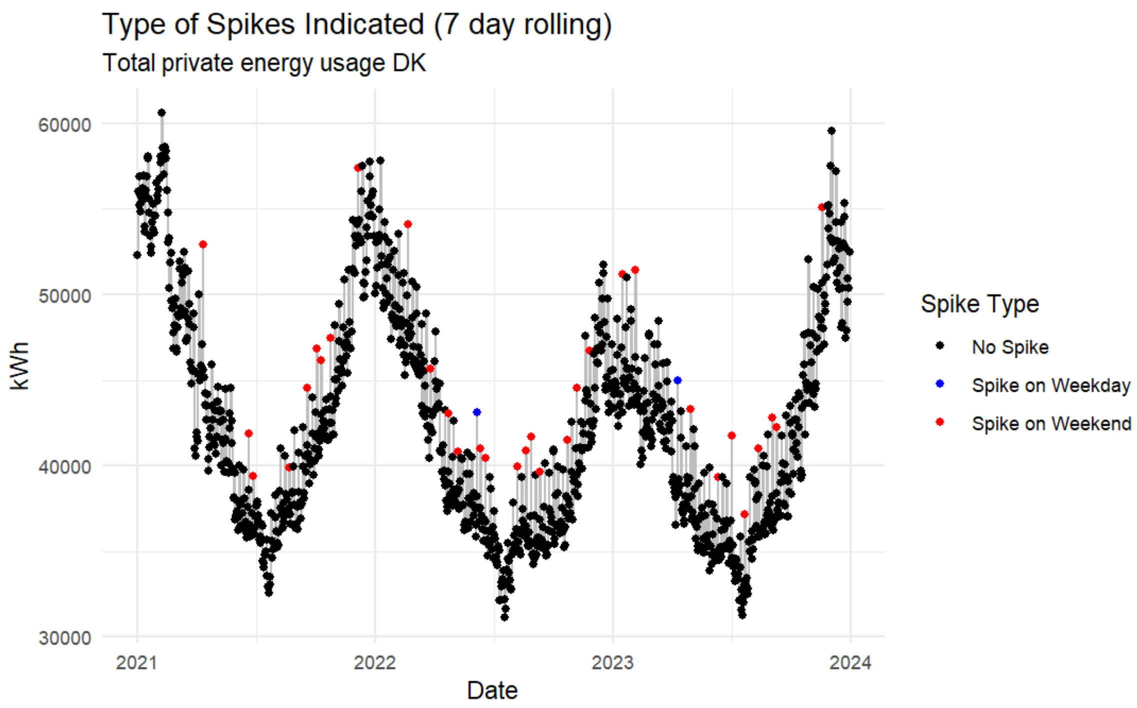

Figure 8

Type	mean (kWh)	median	count
Weekday	41734.18	40849.47	780
Weekend	44302.05	43130.00	312

Table 3:

Which highlights the need for dummies or even more complex methods to address the seasonality of the data. This was further corroborated by an ACF after differencing as seen in figure 10 with significant lags at divisions of 7. After incorporating weekly dummies with Sunday on a non-lagged linear model the ACF was used on the residuals as seen in figure 11. Since no lags were in the model the first spikes signify the need for lagging our dependent variable. Significant spikes around 20-30 show a near monthly pattern and significance at around 50 further shows what we saw in our spike detection in figure 9 where the weekends or holidays have an increased load. We thus added in a monthly dummy and after examining the specific dates of large residuals found that holidays were more likely to have an increase in energy consumption. We then combined the weekday and holiday dummy into one and proceeded to model and compare different lags. Our models thus included the exogenous dummies:

D7, a weekly dummy, **D12,** a monthly dummy, **DH,** a weekend/holiday dummy

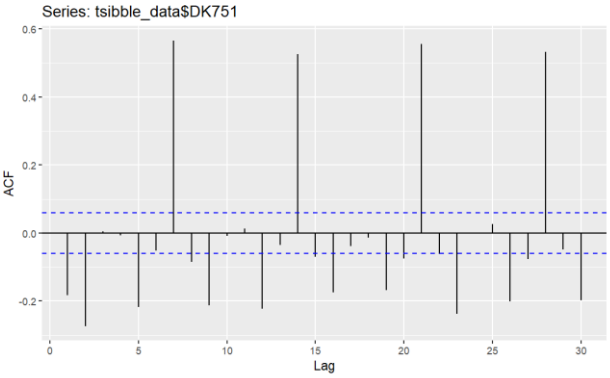

Figure 9

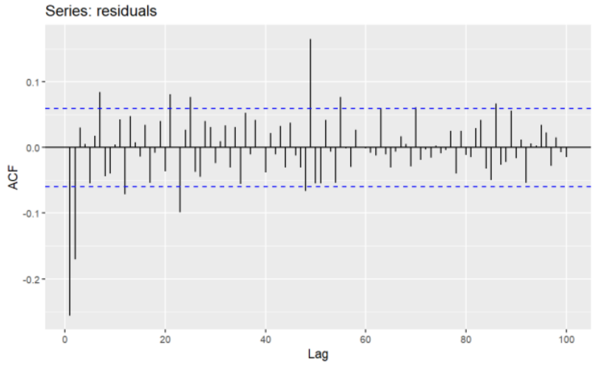

Figure 10

Results: Forecast comparison metrics

Another big choice in forecasting is which measure of accuracy to use. Since the data is differenced, percentage-based measurements such as mean average prediction error (MAPE) will perform worse due to potential issues with zero or near-zero denominators. We opted for mean absolute error (MAE) and root mean squared error (RMSE) as our primary metrics. Both MAE and RMSE are robust to the effects of differencing and provide a clear interpretation of the average errors in prediction. MAE is particularly useful for understanding the average outcome as it provides a linear score that does not overly penalize larger errors. This can be beneficial for macroeconomic data where errors are typically symmetric. RMSE, on the other hand, gives a higher weight to larger errors, which can be beneficial in highlighting worse-case scenarios.

Region	Dummies	Lag.1	Lag.2	Lag.3	Lag.4	Lag.5	Lag.6	Lag.7
NUTS2	7+12	1309.5 (1757.2)	1297.9 (1741.4)	1286.0 (1731.2)	1281.1 (1729.6)	1285.4 (1726.7)	1249.1 (1694.2)	1242.3 (1690.4)
NUTS2	7+12+H	1297.4 (1725.8)	1284.5 (1703.1)	1271.9 (1697.5)	1267.0 (1697.1)	1270.6 (1692.2)	1237.0 (1662.4)	1229.0 (1649.2)
NUTS3	7+12	1276.9 (1724.7)	1272.4 (1717.7)	1269.8 (1714.5)	1267.3 (1713.9)	1270.8 (1716.9)	1245.1 (1694.7)	1247.3 (1696.9)
NUTS3	7+12+H	1267.7 (1701.4)	1263.3 (1693.6)	1266.3 (1698.2)	1264.9 (1697.6)	1269.0 (1694.8)	1251.5 (1673.1)	1259.6 (1675.5)
LAU1	7+12	1217.5 (1669.2)	1230.7 (1690.0)	1370.6 (1859.3)	NA	NA	NA	NA
LAU1	7+12+H	1213.4 (1647.5)	1226.3 (1665.2)	1569.8 (2148.6)	NA	NA	NA	NA

Table 4: Pseudo-OOS results MAE (RMSE)

LAU 1 with 7+12+H dummies seems to be exhibiting the lowest MAE and RMSE at shorter lags. This suggests that it provides more accurate short-term forecasts. On the other hand, NUTS 2 with the 7+12+H dummies offers stable and reliable performance across longer lags, making it a robust choice for longer-term forecasting.

Returning to levels

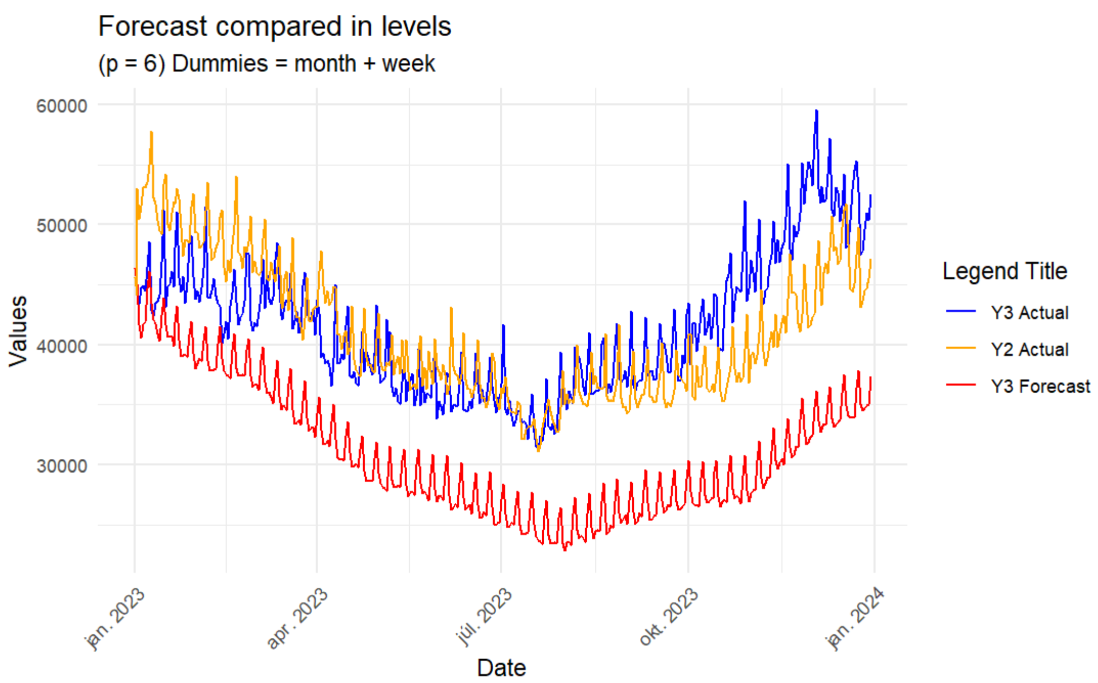

Figure 11: Forecast compared

It is important to mention that the process of transforming differenced data back to levels is a crucial step in similar analysis. While differencing the data is necessary to achieve stationarity and ensure reliable model estimates, it is equally important to present the final forecasts in their original scale- levels. This would allow for a meaningful interpretation as well as for the application of the results.

In this analysis, the differencing approach was initially adopted to handle non-stationarity in the data. Differencing helped stabilize the mean and variance, making the series suitable for VAR modeling. Throughout the forecasting process, all of the computations and model evaluations were conducted using the differenced data. However, due to the substantial complexities involved this was not completed in this analysis.

Results: Multiple forecast comparison

When comparing multiple forecasts there are three main approaches: A pairwise comparison where different models test for equal predictability (one against one), a directional comparison (one against many) where models are tested for dominance and multiple forecast comparison (many against many). For this paper three pairwise comparisons will be tested between three models of varying complexity with one baseline model. The first model, a naïve difference model, predicts no change between periods which implies no change from its last observed value. The second model is a basic VAR(1) model estimated on the regional granularity of NUTS 2. The third model now iterates on lags instead of granularity. The fourth model is then again a basic VAR(1) model but now estimated on NUTS 3 data.

In this way the predictive power of varying granularity and lag structure can easily be interpreted just as the MAE (RMSE) results in table 4. Pseudo-OOS results can thus be seen between the 4 models in figures 13-16.

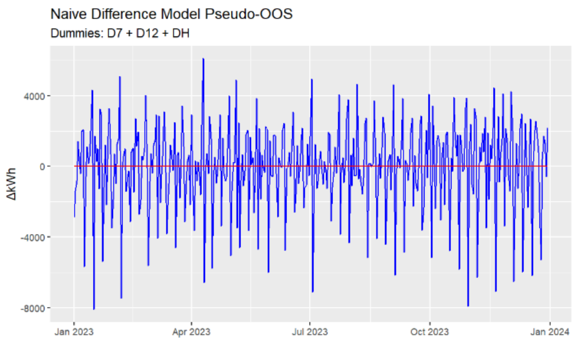

Figure 12

MAE (RMSE): 1871.947 (2497)

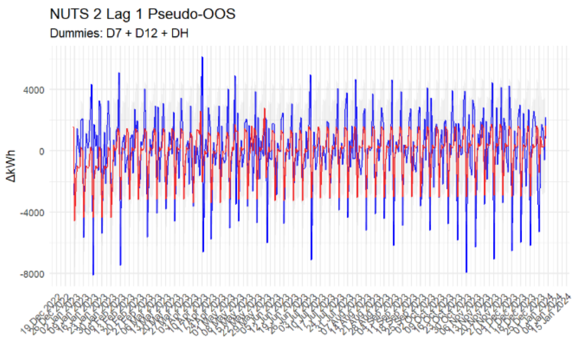

Figure 13

MAE (RMSE): 1297 (1726)

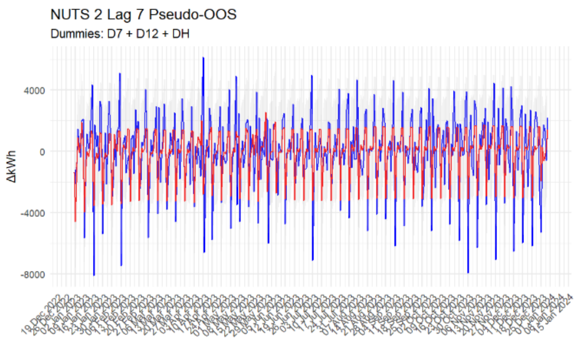

Figure 14

MAE (RMSE): 1229 (1649)

Figure 15

MAE (RMSE): 1202 (1648)

Diebold-Mariano test

The Diebold-Mariano evaluates whether the differences in the predictive performance in two models are statistically significant. The DM test works by focusing on the pairwise forecast error differences. The null hypothesis of it suggests that the mean difference in forecast errors between those two models is zero, which implies that there is no significant difference in their respective predictive accuracy (Diebold and Mariano, 2002).

For an unspecified loss function this difference will be defined as:

$$d_t = L(y_t, \hat{y}_{1,t}) - L(y_t, \hat{y}_{2,t})$$

A Diebold-Mariano test requires covariance stationary residuals of forecasts, so it is essential to first check again to make sure they are stationary, all the models compared reject the null as seen in table 5 in order to move onto the next step.

Granularity	Lag (p)	ADF_Statistic	P_Value	Conclusion
NUTS2	1	−5.88×10^14	0.01	Stationary
NUTS3	1	−6.59×10^16	0.01	Stationary
NUTS2	7	−8.56×10^15	0.01	Stationary

Table 5:

As seen in table 5 when we run a DM-test on residuals between NUTS 2 with lag 1 and NUTS 2 with lag 7 we get a DM-statistic of 2.85 and a p-value of 0.00046. The results suggest that increasing the lags from 1 to 7 does in fact make a statistical significant in forecast accuracy. But we still need to be careful in our interpretation here because model 1 is nested in model 2, all independent variables in model 1 are also present in model 2.

Clark-West test

If we want to account for the nested nature of comparing models with different lags we can utilize the Clark and West (2007) adjusted statistic which should be approximately normally distributed for comparisons of nested forecasts. It accomplishes this by comparing the MSFE of both models while accounting for the differences between the models.

It has the test statistic:

$$CW_i = \left(y_{1,i} - y_{2,i}\right)^2 - \left(y_i - y_{1,i}\right)^2 + \left(y_{2,i} - y_{1,i}\right)^2$$

And the null hypothesis:

$$H_0: MSFE_1 \geq MSFE_2$$

Results table

Regions	Competing models	DM_t	CW_t	P value	Method	Result
NUTS2 NUTS2	Lag 1 vs. Lag 7	2.85	NA	0.0046	Two-sided DM Test	H0: rejected
NUTS2 NUTS2	Lag 1 vs Lag 7	NA	3.68	0.00023	Two-sided CW Test	H0: rejected
NUTS2 NUTS3	NUTS2 vs. NUTS3	2.64	NA	0.0087	Two-sided DM Test	H0: rejected

Table 6:

The results of our model comparison using the Diebold-Mariano and Clark-West tests are summarized in Table 6. The DM test results indicate significant differences between competing models, as evidenced by the rejection of the null hypothesis in all comparisons. Specifically, the DM test results show that Lag 7 outperforms Lag 1 in the NUTS 2 region, with a test statistic of 2.85 and p value of 0.0046. Similarly, the comparison between NUTS2 and NUTS3 regions reveals a significant difference, with a test statistic of 2.64 and a p value of 0.0087 where NUTS 3 outperforms NUTS 2. The CW test further supports these findings even for the instance of nested models with a test statistic of 3.68 and a p value of 0.00023.

Conclusion

Our analysis confirms that higher granularity in data significantly enhances the predictability of VAR energy consumption forecasts for Aarhus, supported by the robust statistical significance observed in the Diebold-Mariano and Clark-West tests.

These findings suggest that utilizing municipal-level data can enhance the accuracy of energy forecasts in Aarhus, potentially aiding in more effective forecasting strategies. However, these results should be viewed as preliminary, encouraging further validation and research.

Building on this paper, future research could explore the use of more sophisticated techniques in removing trend and season, such as Fourier transformations which could more accurately capture and adjust for complex seasonal patterns in energy consumption data. Also integrating more exogenous variables, like temperature data could enhance the model's responsiveness to environmental influences. Expanding the comparison framework to include both more models and evaluating them more rigidly with directional comparisons or multiple forecast comparisons could offer deeper insights into the relative strengths of different forecasting methods, thereby refining the predictive accuracy and reliability of future models.

References

Angelaccio, M., 2019. Forecasting Public Electricity Consumption with ARIMA Model: A Case Study from Italian Municipalities Energy Data. 2019 International Symposium on Advanced Electrical and Communication Technologies (ISAECT), Rome, Italy, pp.1-3. https://doi.org/10.1109/ISAECT47714.2019.9069696.

Andersen, F.M., Henningsen, G., Møller, N.F. and Larsen, H.V., 2019. Long-term projections of the hourly electricity consumption in Danish municipalities. Energy, 186, p.115890. https://doi.org/10.1016/j.energy.2019.115890.

Ansarin, M., Ghiassi-Farrokhfal, Y., Ketter, W. and Collins, J., 2020. Cross-subsidies among residential electricity prosumers from tariff design and metering infrastructure. Energy Policy, 145, p.111736. https://doi.org/10.1016/j.enpol.2020.111736.

Aye, G.C., Miller, S.M., Gupta, R. et al., 2016. Forecasting US real private residential fixed investment using a large number of predictors. Empirical Economics, 51, pp.1557–1580. https://doi.org/10.1007/s00181-015-1059-z.

Ayres, R.U. and Warr, B., 2009. The Economic Growth Engine: How Energy and Work Drive Material Prosperity. Edward Elgar Publishing. Available at: https://books.google.dk/books?hl=el&lr=&id=nLfJKVK9uJsC.

Boehme, P., Berger, M. and Massier, T., 2015. Estimating the building based energy consumption as an anthropogenic contribution to urban heat islands. Sustainable Cities and Society, 19, pp.373-384. https://doi.org/10.1016/j.scs.2015.05.006.

Clark, T.E. and West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics, 138(1), pp.291-311. https://doi.org/10.1016/j.jeconom.2006.05.023.

Diebold, F.X. and Mariano, R.S., 2002. Comparing Predictive Accuracy. Journal of Business & Economic Statistics, 20(1), pp.134–144. https://doi.org/10.1198/073500102753410444.

Kahraman, A., Kantardzic, M., Kahraman, M.M. and Kotan, M., 2021. A Data-Driven Multi-Regime Approach for Predicting Energy Consumption. Energies, 14, p.6763. https://doi.org/10.3390/en14206763.

Karabiber, O.A. and Xydis, G., 2020. Forecasting day-ahead natural gas demand in Denmark. Journal of Natural Gas Science and Engineering, 76, p.103193. https://doi.org/10.1016/j.jngse.2020.103193.

Kümmel, R., Henn, J. and Lindenberger, D., 2002. Capital, labor, energy and creativity: Modeling innovation diffusion. Structural Change and Economic Dynamics, 13(4), pp.415-433.

Lettau, M. and Ludvigson, S., 2001. Consumption, Aggregate Wealth, and Expected Stock Returns. The Journal of Finance, 56, pp.815-849. https://doi.org/10.1111/0022-1082.00347.

Solow, R.M., 1956. A contribution to the theory of economic growth. Quarterly Journal of Economics, 70(1), pp.65-94.

Statistics Denmark. (2024). Units connected to a CVR number in other categories, including category 11 ('Uoplyst aktivitet') [Data set]. Energi Data Service. Available at: https://www.energidataservice.dk/tso-electricity/ConsumptionIndustry

Stern, D.I., 2000. A multivariate cointegration analysis of the role of energy in the US macroeconomy. Energy Economics, 22(2), pp.267-283.

TGARGUIFA, A., BOUNAHMIDI, T. and FELLAOU, S., 2020. Optimal Design of the Distillation Process Using the Artificial Neural Networks Method. 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, pp.1-6. https://doi.org/10.1109/IRASET48871.2020.9092266.

Appendix

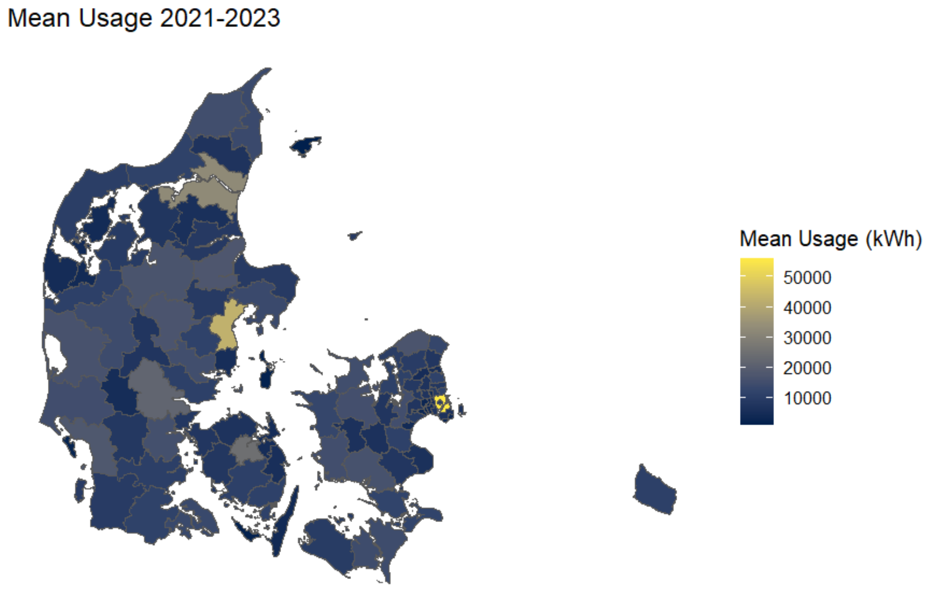

Figure 16

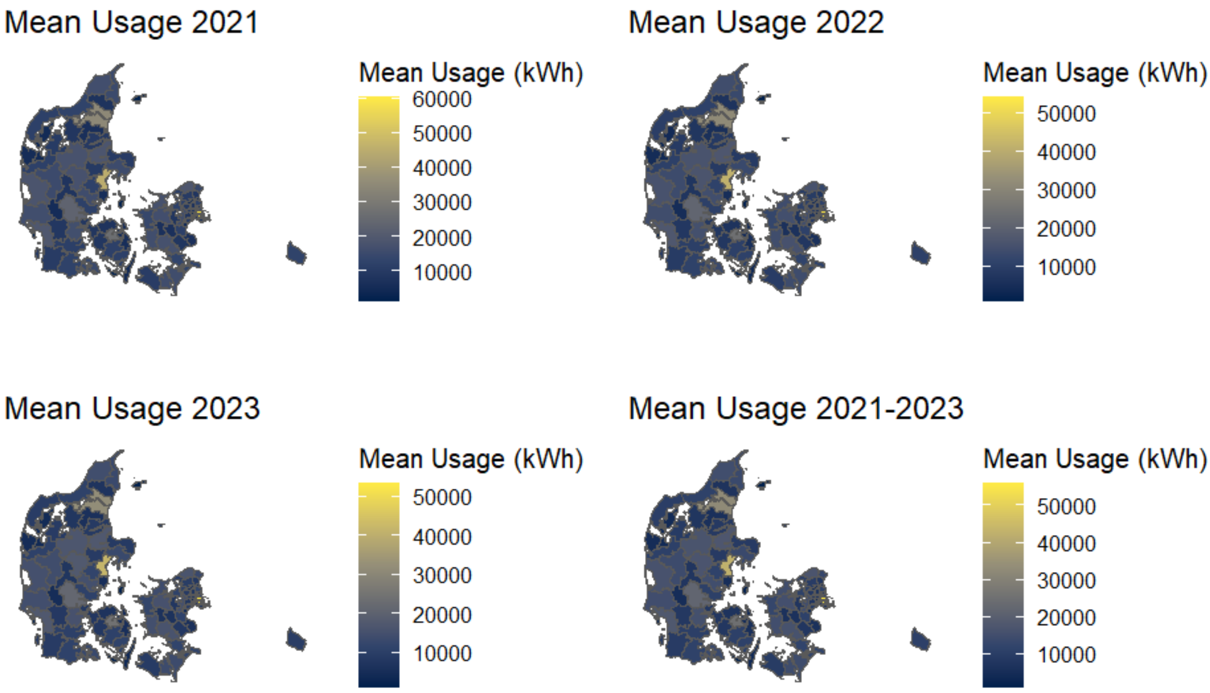

Figure 17

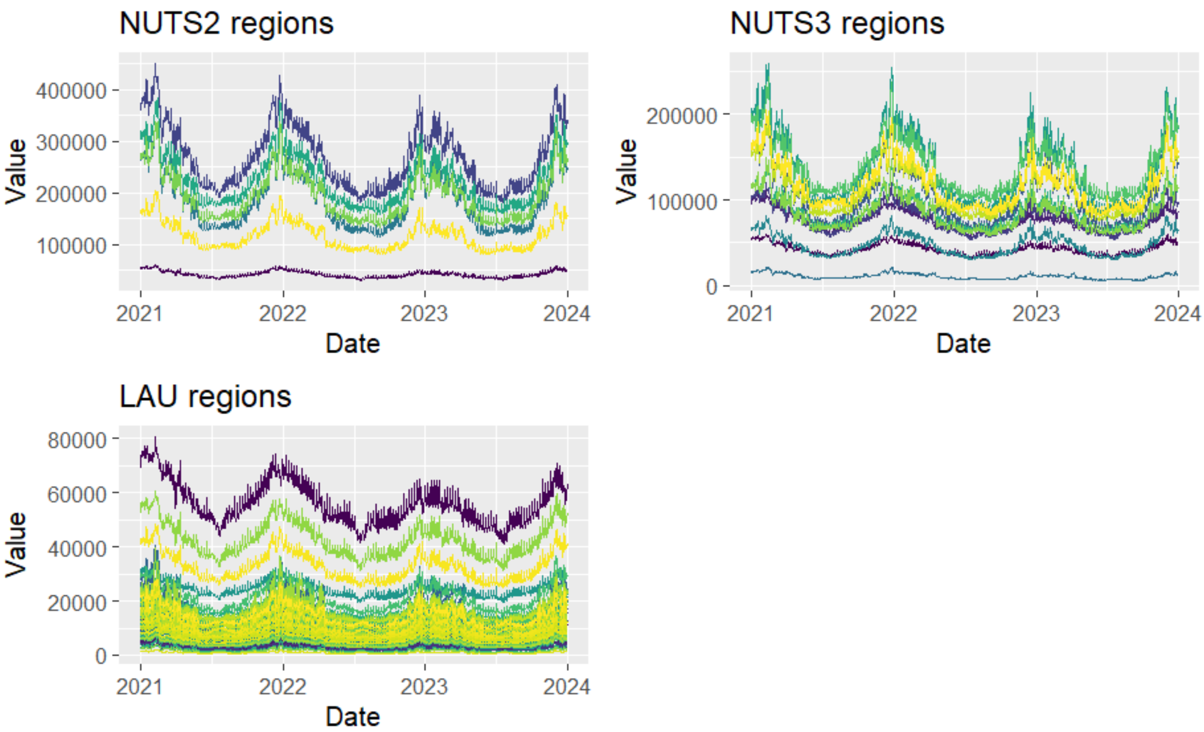

Figure 18

Region	***DF_t***	***p_Value***	***Lag_Used***	***method***	***result***
DK751	*-2.03*	*0.56*	*10*	*ADF-test*	*Non-stationary*
DK011	*-2.10*	*0.54*	*10*	*ADF-test*	*Non-stationary*
DK012	*-1.96*	*0.60*	*10*	*ADF-test*	*Non-stationary*
DK013	*-1.96*	*0.60*	*10*	*ADF-test*	*Non-stationary*
DK014	*-2.26*	*0.47*	*10*	*ADF-test*	*Non-stationary*
DK021	*-1.92*	*0.61*	*10*	*ADF-test*	*Non-stationary*
DK022	*-2.08*	*0.54*	*10*	*ADF-test*	*Non-stationary*
DK031	*-2.17*	*0.51*	*10*	*ADF-test*	*Non-stationary*
DK032	*-2.24*	*0.48*	*10*	*ADF-test*	*Non-stationary*
DK041	*-2.40*	*0.41*	*10*	*ADF-test*	*Non-stationary*
DK042	*-2.09*	*0.54*	*10*	*ADF-test*	*Non-stationary*
DK050	*-2.25*	*0.47*	*10*	*ADF-test*	*Non-stationary*

Table 7: ADF-test on DK751

Region	**DF_t**	**p_Value**	**Lag_Used**	**method**	**result**
DK751	-10.77	0.01	10	ADF-test	stationary
DK011	-10.82	0.01	10	ADF-test	stationary
DK012	-11.67	0.01	10	ADF-test	stationary
DK013	-12.25	0.01	10	ADF-test	stationary
DK014	-12.60	0.01	10	ADF-test	stationary
DK021	-12.15	0.01	10	ADF-test	stationary
DK022	-12.86	0.01	10	ADF-test	stationary
DK031	-12.08	0.01	10	ADF-test	stationary
DK032	-12.46	0.01	10	ADF-test	stationary
DK041	-12.04	0.01	10	ADF-test	stationary
DK042	-12.54	0.01	10	ADF-test	stationary
DK050	-12.36	0.01	10	ADF-test	stationary

Table 8: *ADF-test on ΔDK751*

Region	Lag.1	Lag.2	Lag.3	Lag.4	Lag.5	Lag.6	Lag.7
NUTS2	1309.5 (1757.2)	1297.9 (1741.4)	1286.0 (1731.2)	1281.1 (1729.6)	1285.4 (1726.7)	1249.1 (1694.2)	1242.3 (1690.4)
NUTS3	1276.9 (1724.7)	1272.4 (1717.7)	1269.8 (1714.5)	1267.3 (1713.9)	1270.8 (1716.9)	1245.1 (1694.7)	1247.3 (1696.9)
LAU1	1217.5 (1669.2)	1230.7 (1690.0)	1370.6 (1859.3)	NA	NA	NA	NA

Table 9: MSE(RMSE) metrics

Region	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7
NUTS2	1297.4 (1725.8)	1284.5 (1703.1)	1271.9 (1697.5)	1267.0 (1697.1)	1270.6 (1692.2)	1237.0 (1662.4)	1229.0 (1649.2)
NUTS3	1267.7 (1701.4)	1263.3 (1693.6)	1266.3 (1698.2)	1264.9 (1697.6)	1269.0 (1694.8)	1251.5 (1673.1)	1259.6 (1675.5)
LAU1	1213.4 (1647.5)	1226.3 (1665.2)	1569.8 (2148.6)	NA	NA	NA	NA

Table 10: MAE(RMSE) metrics w DH