Modelling the demand for gas by industry *

Daniel J. Dennis *Newcastle University*

Reducing costs and lowering the environmental footprint are of critical importance across many industries. Natural gas supply and demand is no different, and the gas distribution networks require accurate forecasts for future demand for gas, allowing them to ensure supply is sufficient, in turn lowering their costs and cutting wastage. Demand can be influenced by countless factors, the most prominent for industrial customers being linked to the weather and working patterns. These are incorporated into a four-state hidden Markov model, using a Gibbs sampler to update beliefs about real data provided by Northern Gas Networks. Bayesian inference will be used to test the model through simulations and analyse the results using diagnostic tools. This research focuses on a region of England which has not been discussed and modelled in any previously produced papers, presenting a novel analysis in the domain of gas demand modelling.

Keywords: pandoc, r markdown, knitr

---

1. Data understanding and preparation

1.1 Introduction

As the climate crisis looms, we must not only innovate, but adapt to be more efficient in our energy consumption. Gas is a vital part of the energy sector in the UK, with 57% of industrial and commercial energy needs being met by gas in 2020 (Grid, 2019). Consumption of natural gas for energy use in the UK skyrocketed from 1970 to 2000, increasing from the equivalent of 11.3 million metric tons of oil to 95.9 million metric tons of oil, before falling to the equivalent of 74.3 million metric tons in 2019 (Sönnichsen, 2021). As the UK moves towards its target of net-zero emissions by 2050, satisfying the changing demand for gas will be integral in this fight. As low-carbon energy sees an uptake in consumption, its intermittent and inflexible nature means that, in the future, natural gas is expected to be critical at supplying energy when reserves cannot meet demand (UKOOG, 2019). The National Grid has submitted the Gas Markets Plan (GMaP), in which one of the three focus areas for 2020 was "balancing", describing the process of trying to maintain equal levels of gas brought on and off the network. Herein lies an issue – it is essential to be informed of how much gas will be required in one day's time, one month's time or one year's time.

The National Grid and the companies that distribute gas, such as Northern Gas Networks (NGN), require demand forecasts to enable them to plan for how much gas will be required on the grid at any given time. NGN produces the demand forecasts for the LDZ's it provides to, for both medium-term and long-term horizons. Within each LDZ, consumers are split into three groups; residential (typically represented by a single house), commercial (typically represented by a block of flats or commercial buildings) and industrial (typically represented by small industrial premises). . . . ****What should I include here?****

This paper outlines a model for demand which applied a log transformation to the data, such that the observation equation is:

---

*Replication files are available on the author's Github account (http://github.com/svmiller). **Current version**: August 12, 2021; **Corresponding author**: svmille@clemson.edu.

$$Y_t|S_t = k \sim N(\boldsymbol{x_t}\boldsymbol{\beta}, \sigma_k^2),\ t = 1, ..., N$$

where Y_t is log gas demand on day t, S_t is the state on day t, β is a length $(5+K)$ vector and x_t is a length $(5+K)$ row vector.

This research will focus on a region in the North East of England, known as a local distribution zone (LDZ), which is supplied by Northern Gas Networks (NGN). In total, they deliver gas to 2.7 million homes and businesses covering the North East, Northern Cumbria and parts of Yorkshire. The data used in this paper has been provided by NGN, spanning more than a decade and features daily metered demand from ***large(?)*** industrial users[1], a composite weather variable (CWV) and calendar information (day of the week and date). The composite weather variable accounts for temperature, wind speed, effective temperature and pseudo seasonal normal effective temperature (Grid, 2016). A data set covering the same time period, containing bank holiday information, will be used alongside the main data set. Statistical methods will be applied to the data, namely time series analysis, while model-fitting will be carried out in the Bayesian framework.

1.2 Aim and objectives

The importance of understanding future demands for gas shapes the motivation for this paper; to model and forecast the daily gas demand from industrial consumers in a region of England over a medium-term horizon (up to about 5 years ahead). To achieve this, three statistical goals need to be completed. First, the exploratory data analysis (EDA) and literature will be used to derive an appropriate statistical model for data from the LDZ. This will utilise a number of graphical representations and statistical summaries to provide insight to what is causing variations in gas demand. Second, a Markov chain Monte Carlo algorithm will be constructed to fit the model. This will be built up incrementally and tested throughout, using appropriate materials to gain further knowledge. Third and finally, the model will be fitted to the data and the results interpreted and analysed.

1.3 Literature review

There is a diverse literature regarding modelling the demand for gas, highlighting the sectors need for such research and the many angles the problem can be tackled from. Studies modelling the demand for gas are rarely focused solely on industrial consumers, usually focusing on residential and commercial consumers, or an aggregate of all three, generally providing short-term to medium-term forecasts. These often involve a model based on weather-related and calendar-related factors. The weather-related factors largely influence gas consumption due to users changing their heating demands, Timmer & Lamb (2007), while calendar-related factors generally reference the differences in gas consumption on weekdays, weekends, and public holidays due to changes in working patterns, Franco & Fantozzi (2015). Other factors are noted for their possible impact on demand for gas, such as oil prices, GDP, and population growth. These factors are usually referenced in some scope, with some papers disregarding them after concluding their impact on short and medium-term forecasting horizons is negligible (e.g., Sánchez-Úbeda & Berzosa, 2007), while other papers evidence their importance and integrate such factors into their model (e.g., Zhu et al., 2014).

Early work on the subject by Lyness (1984), describes how the British Gas Corporation uses a Box-Jenkins model to produce short-term and medium-term forecasts, with a view of a top-down or bottom-up approach for long-term forecasting. The bottom-up approach could be to take

[1] Large industrial premises which have their meters read daily, Grid (2016)

assumptions based on the number of gas appliances in the market and their corresponding gas usage, allowing a calculation for total demand. The top-down approach would be looking at the total demand for energy per sector, then breaking this down to the level for gas. Lyness's paper recognised the underlying patterns of demand; where daily consumption follows a "diurnal swing" (the temperature pattern from the daily high to the daily low), a weekly cycle where weekends and weekdays exhibit different consumption patterns, and annual consumption following a roughly sinusoidal curve linked to seasonal temperature variations. For a detailed overview of the topic, both Vitullo et al. (2009), and Soldo (2012), discuss a range of statistical models and the factors these are based upon, with the latter also providing insight on computer science-driven models.

Sánchez-Úbeda & Berzosa (2007), present a novel approach combining quantitative and qualitative forecasts for industrial users in Spain. This uses a model equation which features trend, seasonal, and transitory components, and encodes calendar information categorized by days of the week, weekends, and holidays. Huntington (2007), presents analysis which adopts a general autoregressive distributed lag relationship to model future trends of industrial natural gas consumption in the United States. The variables included in the model are current and lagged values of natural gas consumption and a set of independent explanatory variables, comprising of industrial natural gas price, distillate fuel oil price, structural output, heating degree-days[2] and capacity utilisation. Zhu et al. (2014), similarly include several socioeconomic variables to their model. GDP, total population and urbanization rate were used as the input variables as they constructed a radial basis function neural network quantile regression model. The results of this were combined with those of a Bayesian vector autoregression model, providing a combinational forecast for China's gas consumption. Wang et al. (2017), use a multiverse optimiser algorithm to optimise the parameters of the Nash non-linear grey Bernoulli model, proposing a hybrid of the models to predict natural gas consumption in 30 regions across China. Laib et al. (2018), model demand in Algeria using a Gaussian process regression based on time series data, using lagged consumption observations, three temperature readings and clustering with several combinations of seasonal labels. Heaps et al. (2020), examine the demand for gas in the same region featured in this paper, however their work investigates the effect of public holidays on residential daily demand for gas. A four-state, non-homogeneous hidden Markov model is used, featuring a so-called proximity effect to model the days leading up to and away from public holidays.

The existing body of literature illustrates that modelling and forecasting gas demand is successful through a wide range of techniques. Studies contain many combinations of different variables as the backbone of the model, each providing a compelling case for their inclusion. A good deal of the studies that focus on industrial gas consumption use data from rapidly developing countries, which likely follow a different pattern of demand to the North East of England. A common feature in the existing literature is to group residential, commercial, and industrial demand together, leaving a gap for purely industrial gas demand modelling. This paper is intended to provide unique research by building a model for industrial gas demand in the aforementioned region of England. There has been little work produced in this area, particularly with regards to industrial consumers. A four-state, hidden Markov model is the foundations of the study, incorporating a weather-related variable and several calendar related variables.

Could maybe write more here about novelty.

[2] "Heating degree days are a measure of how much (in degrees), and for how long (in days), outside air temperature was higher than a specific base temperature" - https://www.degreedays.net/

1.4 Exploratory data analysis

To begin, the data was loaded into RStudio from the .csv file provided by Northern Gas Networks. This contained 4031 observations on 33 variables, before being subset to a smaller data frame using the necessary columns - Day, Date, NE_CWV and NE_DM. A row was found with `NA` corresponding to both variables (NE_CWV, NE_DM), and after a check it was clear the entire row was an error in the data, thus it was omitted. No duplicated dates or further errors were present. The entries of the 'Date' column were transformed from *DD-MONTH-YY* format to *YYYY-MM-DD* format using the `lubridate` package, allowing the exploratory data analysis process to begin.

To understand the composite weather variable (CWV), a box plot was created (figure 1) which was subset by the day of the week. This shows that the CWV is constant throughout the week for the region, which is to be expected since the day of the week has no known relationship with the weather.

To help understand how the daily metered demand (DM) is affected by days of the week, another box plot was created (figure 2). This shows that the demand is largely similar on weekdays before decreasing at the weekend. This is likely due to industrial users closing or reducing capacity on weekends, thus requiring less gas for uses such as powering machinery and heating buildings. This suggests that some variable that differentiates between weekdays and weekends may be appropriate for inclusion in the model.

The way in which daily metered demand is changing over time will be vital when it comes to forecasting for the future. Figure 3 shows that the demand on bank holidays is consistently lower than it is on non-bank holidays, so a variable that differentiates between bank holidays and non-bank holidays may be appropriate for inclusion in the model. There may be some seasonality to the data, however it is hard to see due to the general upwards trend over the 10-year period. Figure 3 shows the daily metered demand in the North East following a static path from 2008 to mid 2011, before sharply increasing until the inception of 2012, before again plateauing through to 2019. This is an indication that a model that allows for step changes in the overall level and variability would be desirable. There are four general levels the daily metered demand data fall into: 5000, 13,000, 20,000 and 28,000 (all approximations).

To see how the demand for each year changes, the curves for each yearly cycle can be superimposed on one another, as seen in figure 4. Note that January 1st of each year is represented by 1 on the x-axis, while December 31st will fall on either 365 or 366. Figure 4 shows how the demand varies throughout the year for the entire 11-year period. There is evidence that demand is lower towards the middle of the year, i.e., summer, coinciding with higher temperatures and CWV values. This suggests that weather has some influence on the demand for gas by industrial consumers.

To further investigate the effects of bank holidays on daily metered demand, figure 5 was produced. This box plot clearly shows that bank holidays correspond to lower daily demand for gas than non-bank holidays, as well as covering a smaller range of values across the data. The reason behind the former is that industrial users are more likely pause or reduce operations on bank holidays, thus requiring less gas, while the reason behind the latter is that the number of bank holidays is very small in comparison to non-bank holidays, so you would expect them to cover a smaller range of values.

To explore the idea that the weather plays a part on gas demand, a scatter plot was created showing the CWV against the DM. Figure 6 shows the data is fairly distinctly clustered with respect to the daily metered demand, exhibiting a weak negative correlation for the two clusters featuring the highest daily demand. There does not appear to be any correlation between the CWV and the daily metered demand for the two clusters featuring the lowest daily demand. The data sits on four different levels, as seen in figure 3, explaining the four clusters seen in figure 6.

Add k-means clustering here?

The graph showing the daily metered demand changing through time does not show any clear signs of seasonality, but by plotting the CWV data over the top may give a better insight. From figure 7, there appears to be some inverted relationship between the DM and the CWV. That is, as the CWV peaks, the DM is at a trough, then as the CWV troughs, the DM peaks. This suggests that in the winter, the daily metered demand tends to be at its highest for the year, while in summer it is at its lowest. This is likely due to the industrial users requiring more gas to heat their buildings due to lower temperatures/poorer weather.

There does not appear to be any obvious correlations in the data from figure 5, so using correlation coefficients (Pearson and Spearman) may clarify this, as shown in table 1. Both coefficients indicate a weak negative correlation between the CWV and the DM. The p-values are both very small, thus significant, so the null hypothesis that either correlation coefficient is equal to 0 is rejected, meaning the negative correlations are present, however they are very weak.

Table 1: Table 1. Correlation coefficients for the CWV against the DM

	Pearson	Pearson p-value	Spearman	Spearman p-value
NE_CWV vs NE_DM	-0.0883	1.981e-08	-0.187	$<$2.2e-16

The autocorrelation function of the DM is useful to gain an understanding of the correlation of points separated by various time lags. Figure 8 shows the autocorrelation function of the DM, exhibiting very slow decay. The data is clearly from a non-stationary process due to the varying mean from figure 3, with nearby observations being highly correlated.

To reduce the variance of the daily metered demand data, logs were taken.

2. Modelling

2.1 Deciding modelling technique

The daily metered demand for gas follows the structure of a time series where observations alternate between four well-defined levels. Based on this assessment, a hidden Markov model seems appropriate for the data (see Frühwirth-Schnatter, 2006).

2.2 Hidden Markov model

The parameter Y_t denotes the log gas demand in tenths of a gigawatt-hour for industrial users on day t in the North East LDZ. The log-scale has been utilised to help reduce the variance in the data across the annual patterns and to give the fixed effects in the additive model a multiplicative effect on the original scale. The CWV value for day t is denoted by cwv_t, the weekday/weekend value for day t is denoted by $w_t \in N\{0,1\}$, and the non-bank holiday/bank holiday value for day t is denoted by $b_t \in N\{0,1\}$. It will be assumed that all bank holidays have the same effect on the demand, rather than grouping them by some shared feature. It must also be noted that no bank holidays in England fall on a weekend, so there shall not be a case for t in which $w_t = b_t = 1$.

The hidden Markov model assumes there to be a hidden number of finite states, forming a Markov chain, each corresponding to a time series of observations. A Markov chain is a sequence of states, S_t, with the current observation depending only on the previous observation, with a transition matrix displaying the transition probabilities for each state, denoted by ξ. Each observation, Y_t, depends on the corresponding state, S_t, and its distribution. The structure of a hidden Markov model allows for each state having its own set of parameters so the distributions may vary accordingly. **This needs to be improved**

The distribution for Y_t was assumed to be a normal distribution such that

$$Y_t|S_t = k \sim N(\boldsymbol{x_t\beta}, \sigma_k^2),\ \ t = 1, ..., N$$

where N is the length of the sequence. Here,

$$\boldsymbol{x}_t = (\Theta(S_1), \Theta(S_2), ..., \Theta(S_K), \tilde{\boldsymbol{x}}_t)$$

where

$$\Theta(S_t = k) = \begin{cases} 1 & \text{if } S_t = k, \\ 0 & \text{otherwise} \end{cases}$$

and

$$\tilde{\boldsymbol{x}}_t = (w_t, b_t, cwv_t, sin(\frac{2\pi t}{365.25}), cos(\frac{2\pi t}{365.25}))$$

and

$$\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_K, \tilde{\boldsymbol{\beta}})$$

The observation equation can be rewritten as

$$Y_t|S_t = k \sim N(\beta_k + \tilde{\boldsymbol{x}}_t\tilde{\boldsymbol{\beta}},\ \sigma_k^2)$$

A function to simulate a normal hidden Markov model was derived using the observation equation **- maybe give this a (1.1) or something ? -** , producing a sequence of observations and the corresponding hidden state path. The function requires the following set of inputs to make it run: $\sigma_k^2,\ \xi,\ \beta,\ \tilde{\beta}$, the desired length of sequence, the distribution for the initial state (S_0) and the number of states (K). The results were then plotted to check the function was producing the expected results. This used a number of values for the states as well as varying the respective means, variances and transition matrices. The function was successfully producing the plots resembling hidden Markov models for any number of states and any values for the means, variances and transition matrices.

2.3 Filtering and sampling models

First, a forward-filtering function (appendix . . .) was created to produce the filtered probabilities $Pr(S_t = l|\boldsymbol{y}^t, \boldsymbol{\theta}),\ l = 1, ..., K$ for each $t = 1, ..., T$. The matrix of filtered probabilities essentially shows the probability of being in each state for each value of Y_t. The results of this function were checked by physical calculations a number of times, each time being successful. The so-called 'log-sum-exp trick' was integrated to the function to make the calculations more numerically stable (see chapter 11, Frühwirth-Schnatter, 2006) . Again, another check was made and the function was found to be running correctly. This function was then incorporated to a backward-sampling algorithm, which iterates backwards through the filtered probabilities from $t = T-1, T-2, ..., t_0$. This function creates a vector of the state path based on the filtered probabilities matrix. Together, the functions represent the forward-filtering-backward-smoothing algorithm, returning a vector of smoothed states. The hidden state path from the hidden Markov model function should be identical to this vector of smoothed states. The plots in **appendix ...** show that the forward-filtering-backward-smoothing function was performing correctly. The algorithms 11.1 and 11.2 discussed in Frühwirth-Schnatter (2006) provided the inference for these functions.

2.4 Prior distributions

A prior for each parameter of the observation equation was required, as well as a prior for ξ. It shall be assumed that each parameter follows the same distribution across all states, however the variables of this distribution can change. Firstly, if $\boldsymbol{\beta} = (\beta_1, ..., \beta_K)$ then it shall be that $\beta_k \sim N(m, s^2)$ where$\beta_1 < \beta_2 < ... < \beta_K$ for $k = 1, ..., K$, where m and s^2 are K-length vectors containing means and variances respectively. Second, if $\sigma_k^2 = (\sigma_1^2, , \sigma_K^2)$, then it shall be that $\sigma_k^2 \sim invgamma(a, b)$, where a and b are shape and rate respectively. Third, $\tilde{\boldsymbol{\beta}} \sim N_5(\mathbf{0}, \tilde{s}^2 I_5)$, where $\mathbf{0}$ is a vector of five 0's. Finally, $\xi \sim Dir(c)$ (?).

These priors need variables that will allow the distributions to be updated by the data, so will feature large variances to give wide scope for change.

2.4 Full conditional distributions (FCDs)

The posterior distributions for the data needed to be derived in order to update the prior beliefs. The posterior distribution of the model parameters follows from Bayes' rule where

$$\pi(\boldsymbol{y}|\boldsymbol{s}, \boldsymbol{\beta}, \sigma_k^2, \tilde{\boldsymbol{\beta}}, \xi) \propto p(\boldsymbol{s}|\boldsymbol{y}, \boldsymbol{\beta}, \sigma_k^2, \tilde{\boldsymbol{\beta}}, \xi)\pi(...) \textit{ Is this right}???$$

in which $\boldsymbol{y}$ represents the complete time series. The probability density function for the observation equation is:

$$p(\boldsymbol{y}, \boldsymbol{s}|\beta_1, ..., \beta_K, \sigma_1^2, ..., \sigma_K^2, \xi) = p(\boldsymbol{y}|\boldsymbol{s}, \beta_1, ..., \beta_K, \sigma_1^2, ..., \sigma_K^2)p(\boldsymbol{s}|\xi)$$

for $k = 1, ..., K$ where

$$\begin{aligned}
p(\boldsymbol{y}|\boldsymbol{s}, \beta_1, ..., \beta_K, \sigma_1^2, ..., \sigma_K^2) &= p(y_1|s_1 = s_1, \beta_{s_1}, \sigma_{s_1}^2) \times p(y_2|s_2 = s_2, \beta_{s_2}, \sigma_{s_2}^2) \times ... \times p(y_T|s_T, \beta_{s_T}, \sigma_{s_T}^2) \\
&= \prod_{k=1}^{K} \prod_{t:s_t=k} p(y_t|s_t = k, \beta_k, \sigma_k^2) \\
&= \prod_{k=1}^{K} \prod_{t:s_t=k} \frac{1}{\sqrt{2\pi\sigma_k^2}} exp\left\{-\frac{1}{2\sigma_k^2}(y_t - \beta_k)^2\right\} \\
&\propto \prod_{k=1}^{K} \sigma_k^{2\left(-\frac{N_k}{2}\right)} exp\left\{-\frac{1}{2\sigma_k^2}\sum_{t:s_t=k}(y_t - \beta_k)^2\right\}
\end{aligned}$$

where N_k is the number of observations in state k. The term $p(\boldsymbol{y}|\boldsymbol{s}, \beta_1, ..., \beta_K, \sigma_1^2, ..., \sigma_K^2)$ is the observed data likelihood, and so **(put a 1.12 or something after the final line above)** will be the likelihood used to calculate the parameters' FCDs by multiplying it by the corresponding prior. The full derivations of the FCDs for β, σ_k^2, $\tilde{\boldsymbol{\beta}}$ and ξ can be found in appendices A, B, C and D respectively.

A function was created for each parameter $(\beta, \sigma_k^2, \tilde{\beta}, \xi)$, returning a list (or matrix in ξ's case) of updated parameters based on their respective FCD's.

2.5 Gibbs sampling

Gibbs sampling is an iterative process used here to continually update a distributions parameters. Through a large number of iterations, the parameters of the distribution should tend towards their true values, thus making it an effective method for when the actual parameters are unknown.

A function was created to perform Gibb's sampling on the data. The function called on all previously discussed functions, running a loop over the specified number of iterations, returning a list of matrices containing updated parameters for β, σ_k^2, $\tilde{\beta}$ and ξ. It also provides a matrix containing the proportion of visits to each state per observation for each iteration, allowing for a model to be taken for each observation, returning a smoothed state vector. **That needs changing!!**

Appendices

Figure 1:

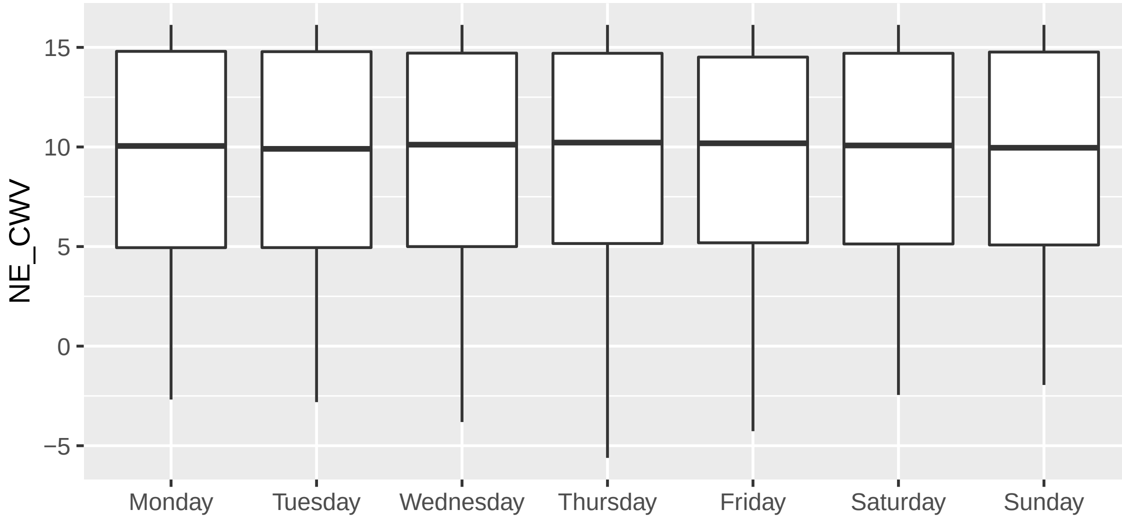


Figure 2:

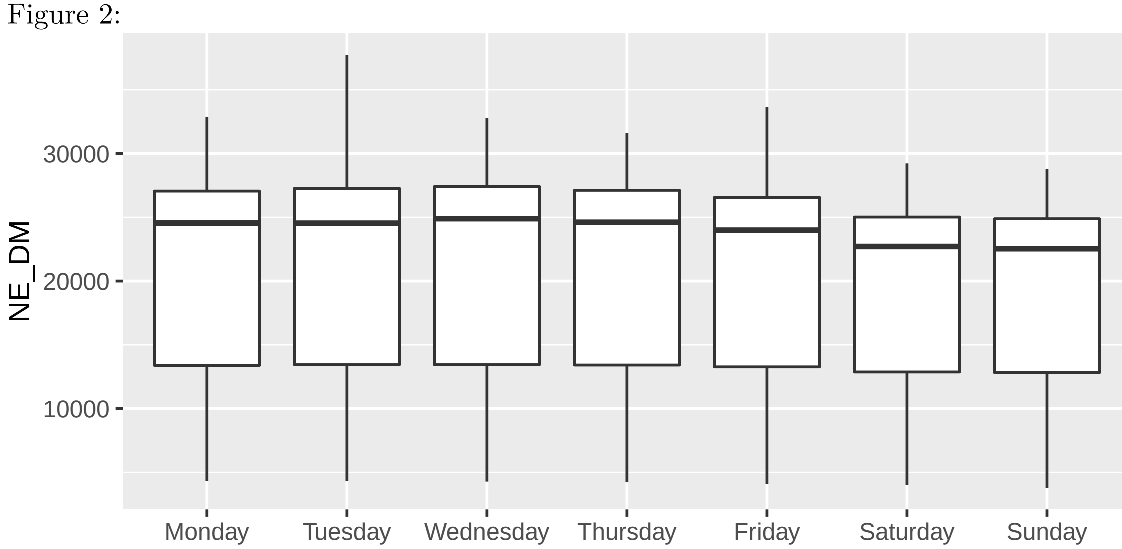


Figure 3:

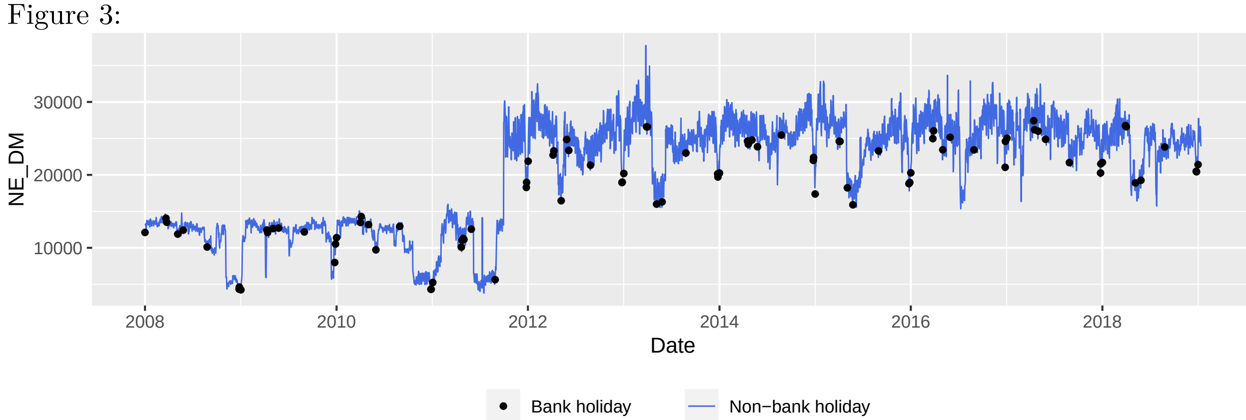


Figure 4:

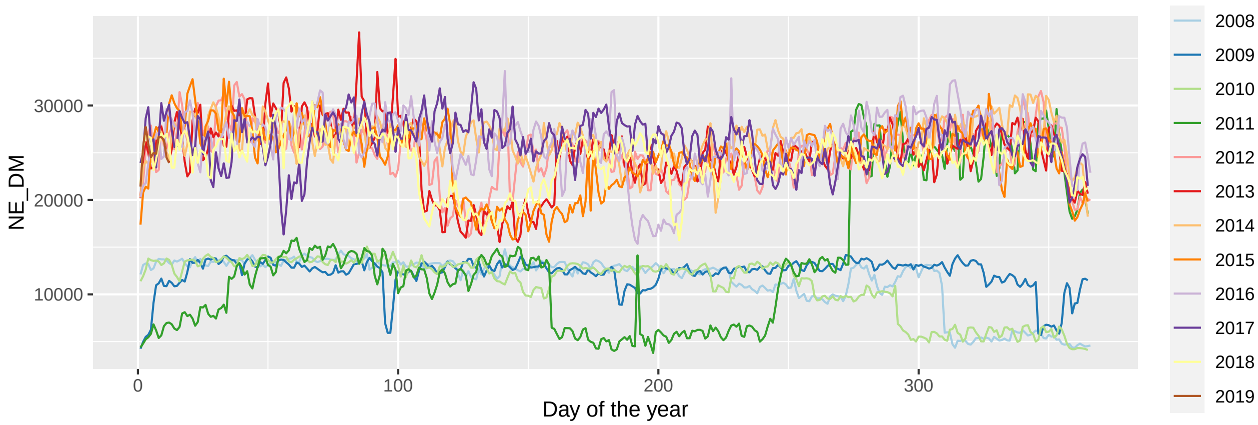


Figure 5:

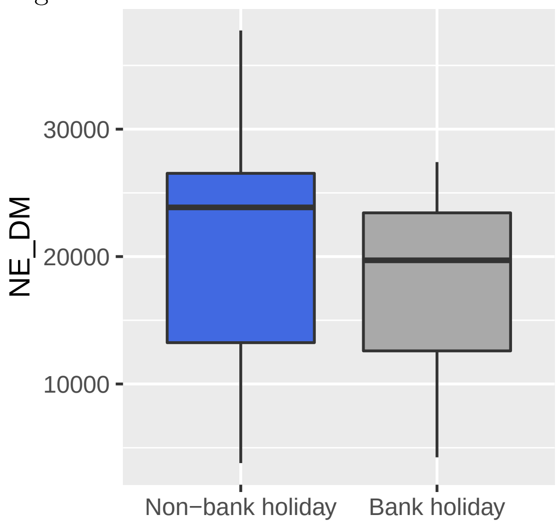


Figure 6:

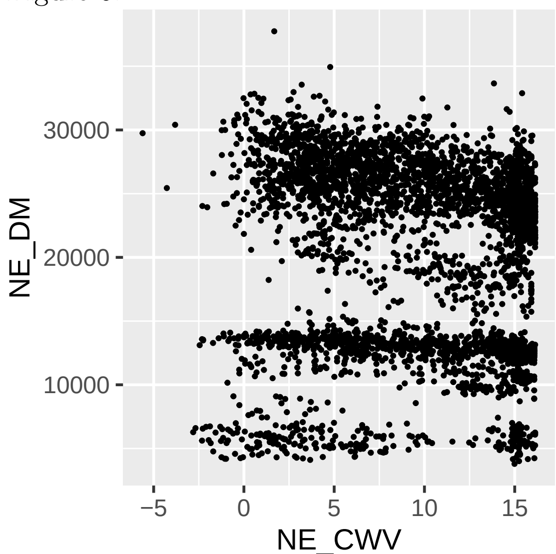


Figure 7:

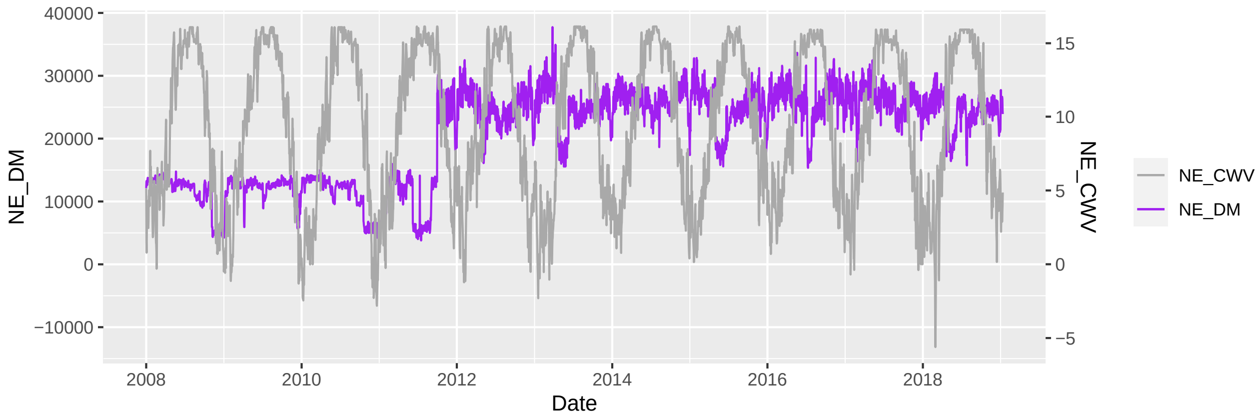
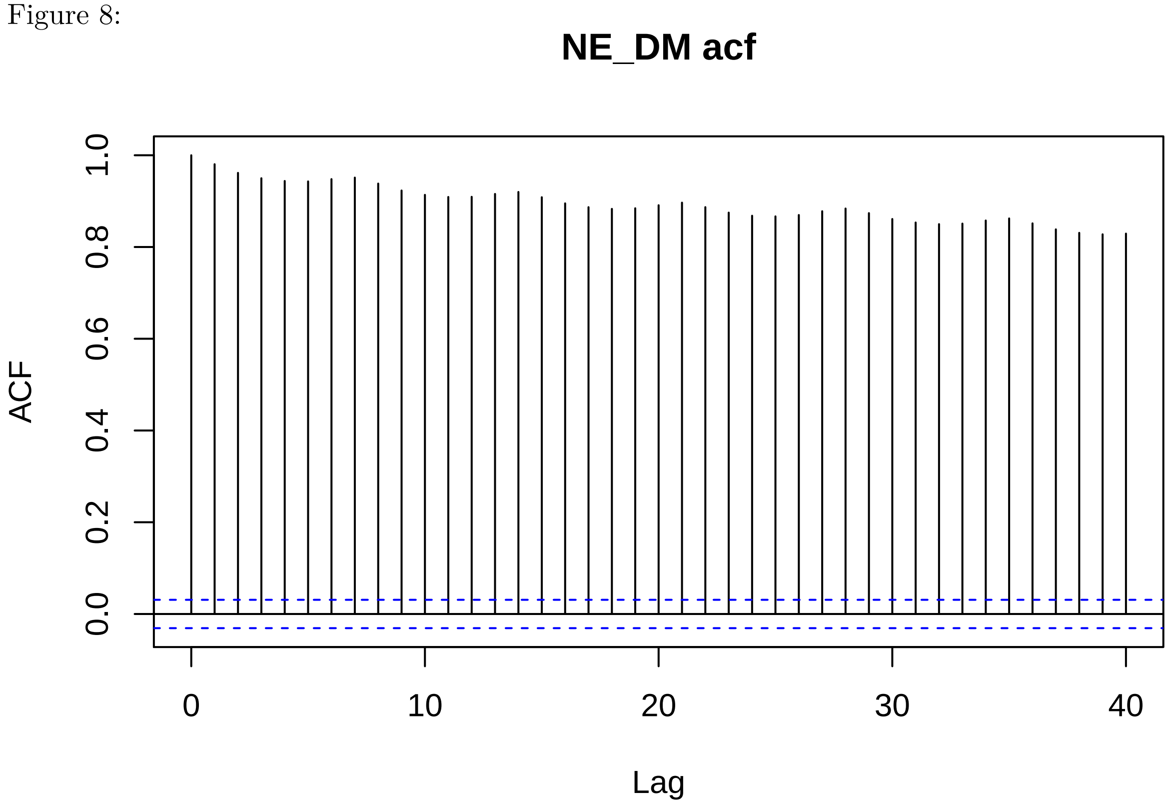


Figure 8:



Appendix A

FCDs for the parameters of a hidden Markov model (?)

The prior for the observation equation is:

$$Y_t | S_t = k \sim N(\boldsymbol{x}_t \boldsymbol{\beta}, \sigma_k^2),\ \ t = 1, ...N$$

Y_t is log gas demand on day t, S_t is the state on day t, $\boldsymbol{\beta}$ is a length $(5+k)$ row vector and $\boldsymbol{x}_t$ with entries

$$\boldsymbol{x}_t = (\Theta(S_t = 1), \Theta(S_t = 2), ..., \Theta(S_t = k), w_t, b_t, cwv_t, sin\left(\frac{2\pi t}{365.25}\right), cos\left(\frac{2\pi t}{365.25}\right))$$

Here,

$$\Theta(S_t = k) = \begin{cases} 1 & \text{if } S_t = k, \\ 0 & \text{otherwise} \end{cases}$$

$$w_t = \begin{cases} 1 & \text{if day } t \text{ is a Saturday/Sunday,} \\ 0 & \text{otherwise} \end{cases}$$

$$b_t = \begin{cases} 1 & \text{if day } t \text{ is a bank holiday,} \\ 0 & \text{otherwise} \end{cases}$$

$cwv_t =$ compositve weather variable on day t

with the final two terms giving a smooth day of the year effect. The vectors $\boldsymbol{\beta}$ and $\boldsymbol{x}_t$ can be written as:

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \\ \beta_k \\ \tilde{\boldsymbol{\beta}} \end{pmatrix} \text{ and } \boldsymbol{x}_t = (\Theta(S_t = 1), ..., \Theta(S_t = k),\ \tilde{\boldsymbol{x}}_t)$$

The observation equation can therefore be written as

$$Y_t | S_t = k \sim N(\beta_k + \tilde{\boldsymbol{x}}_t \tilde{\boldsymbol{\beta}}, \sigma_k^2),\ t = 1, ...N$$

The prior for β_k is

$$\beta_k \sim N(m, s^2) \text{ where } \beta_1 < \beta_2 < ... < \beta_K \text{ for } k = 1, ..., K \text{ and } \tilde{\boldsymbol{\beta}} \sim N_5(\mathbf{0}, \tilde{s}^2 I_5) \text{ for } k = K{+}1, ..., K{+}5$$

where $\mathbf{0}$ is a vector of five 0's.

The probability density function for the observation equation is:

$$p(\boldsymbol{y}, \boldsymbol{s} | \beta_1, ..., \beta_K, \sigma_1^2, ..., \sigma_K^2, \xi) = p(\boldsymbol{y} | \boldsymbol{s}, \beta_1, ..., \beta_K, \sigma_1^2, ..., \sigma_K^2) p(\boldsymbol{s} | \xi)$$

for $k = 1, ..., K$ where

$$\begin{aligned} p(\boldsymbol{y} | \boldsymbol{s}, \beta_1, ..., \beta_K, \sigma_1^2, ..., \sigma_K^2) &= p(y_1 | s_1 = s_1, \beta_{s_1}, \sigma_{s_1}^2) \times p(y_2 | s_2 = s_2, \beta_{s_2}, \sigma_{s_2}^2) \times ... \times p(y_T | s_T, \beta_{s_T}, \sigma_{s_T}^2) \\ &= \prod_{k=1}^{K} \prod_{t: s_t = k} p(y_t | s_t = k, \beta_k, \sigma_k^2) \\ &= \prod_{k=1}^{K} \prod_{t: s_t = k} \frac{1}{\sqrt{2\pi\sigma_k^2}} exp\left\{ -\frac{1}{2\sigma_k^2} (y_t - \beta_k)^2 \right\} \\ &\propto \prod_{k=1}^{K} \sigma_k^{2\left(-\frac{N_k}{2}\right)} exp\left\{ -\frac{1}{2\sigma_k^2} \sum_{t: s_t = k} (y_t - \beta_k)^2 \right\} \end{aligned}$$

where N_k is the number of observations in state k.

The full conditional distribution for $\beta_1, ..., \beta_k$ is therefore:

$$\pi(\beta_1, ..., \beta_K | \boldsymbol{y}, \boldsymbol{s}, \sigma_1^2, ..., \sigma_K^2, \xi) \propto p(\boldsymbol{y}, \boldsymbol{s} | \beta_1, ..., \beta_k, \sigma_1^2, ..., \sigma_K^2, \xi) \pi(\beta_1, ..., \beta_K)$$

$$= p(\boldsymbol{y}|\boldsymbol{s}, \beta_1, ..., \beta_k, \sigma_1^2, ..., \sigma_K^2) p(\boldsymbol{s}|\xi) \pi(\beta_1, ..., \beta_K)$$

$$\propto p(\boldsymbol{y}|\boldsymbol{s}, \beta_1, ..., \beta_k, \sigma_1^2, ..., \sigma_K^2) \pi(\beta_1, ..., \beta_K)$$

$$= p(\boldsymbol{y}|\boldsymbol{s}, \beta_1, ..., \beta_k, \sigma_1^2, ..., \sigma_K^2) \times \prod_{k=1}^{K} \pi(\beta_k)$$

$$\propto \prod_{k=1}^{K} \sigma_k^{2\left(-\frac{N_k}{2}\right)} exp\left\{-\frac{1}{2\sigma_k^2} \sum_{t:s_t=k} (y_t - \beta_k)^2\right\} \times \prod_{k=1}^{K} exp\left\{-\frac{1}{2v}(\beta_k - m)^2\right\}$$

$$\propto \prod_{k=1}^{K} exp\left\{-\frac{1}{2\sigma_k^2}\left(\sum_{t:s_t=k}(y_t^2 - 2y_t\beta_k + \beta_k^2)\right) - \frac{1}{2v}(\beta_k - m)^2\right\}$$

$$= \prod_{k=1}^{K} exp\left\{-\frac{1}{2\sigma_k^2}\left(\sum_{t:s_t=k} y_t^2 - 2\beta_k \sum_{t:s_t=k} y_t + \beta_k^2 \sum_{t:s_t=k} 1\right) - \frac{1}{2v}(\beta_k^2 - 2\beta_k m + m^2)\right\}$$

$$= \prod_{k=1}^{K} exp\left\{-\frac{1}{2\sigma_k^2}\left(\sum_{t:s_t=k} y_t^2 - 2\beta_k N_k \bar{y}_k + N_k \beta_k^2\right) - \frac{1}{2v}(\beta_k^2 - 2\beta_k m + m^2)\right\}$$

$$\propto \prod_{k=1}^{K} exp\left\{-\frac{1}{2\sigma_k^2}(-2\beta_k N_k \bar{y}_k + N_k \beta_k^2) - \frac{1}{2v}(\beta_k^2 - 2\beta_k m)\right\}$$

$$= \prod_{k=1}^{K} exp\left\{-\frac{1}{2\sigma_k^2 v}(v[-2\beta_k N_k \bar{y}_k + N_k \beta_k^2]) + \sigma_k^2(\beta_k^2 - 2\beta_k m)\right\}$$

$$= \prod_{k=1}^{K} exp\left\{-\frac{1}{2\sigma_k^2 v}([vN_k + \sigma_k^2]\beta_k^2 - 2\beta_k N_k \bar{y}_k v - 2\beta_k m \sigma_k^2)\right\}$$

$$= \prod_{k=1}^{K} exp\left\{-\frac{1}{\frac{2\sigma_k^2 v}{vN_k + \sigma_k^2}}\left[\beta_k^2 - 2\left(\frac{vN_k \bar{y}_k + \sigma_k^2 m}{vN_k + \sigma_k^2}\right)\beta_k\right]\right\}$$

$$= \prod_{k=1}^{K} exp\left\{-\frac{1}{\frac{2\sigma_k^2 v}{vN_k + \sigma_k^2}}\left[\left(\beta_k - \frac{vN_k \bar{y}_k + \sigma_k^2 m}{vN_k + \sigma_k^2}\right)^2 - \left(\frac{vN_k \bar{y}_k + \sigma_k^2 m}{vN_k + \sigma_k^2}\right)^2\right]\right\}$$

$$\propto \prod_{k=1}^{K} exp\left\{-\frac{vN_k + \sigma_k^2}{2\sigma_k^2 v}\left[\left(\beta_k - \frac{vN_k \bar{y}_k + \sigma_k^2 m}{vN_k + \sigma_k^2}\right)^2\right]\right\}$$

So, because this factorises as a product of densities for $\beta_1, ..., \beta_K$, the β_k are independent in the full conditional distribution with

$$\beta_k|\boldsymbol{y}, \boldsymbol{s}, \sigma_k^2, \xi \sim N\left(\frac{vN_k \bar{y}_k + \sigma_k^2 m}{vN_k + \sigma_k^2}, \frac{\sigma_k^2 v}{vN_k + \sigma_k^2}\right)$$

where $\bar{y}_k = \frac{1}{N_k}\sum_{t:s_t=k} y_t - \tilde{\boldsymbol{x}}_t \tilde{\boldsymbol{\beta}}$ for $k = 1, ..., K$.

The full conditional distribution for $\sigma_1^2, ..., \sigma_K^2$ is:

$$\pi(\sigma_1^2, ..., \sigma_K^2|\boldsymbol{y}, \boldsymbol{s}, \mu_1, ..., \mu_K, \xi) \propto p(\boldsymbol{y}, \boldsymbol{s}|\mu_1, ..., \mu_k, \sigma_1^2, ..., \sigma_K^2, \xi) \pi(\sigma_1^2, ..., \sigma_K^2)$$

$$
\begin{aligned}
&= p(\boldsymbol{y}|\boldsymbol{s}, \mu_1, ..., \mu_k, \sigma_1^2, ..., \sigma_K^2) p(\boldsymbol{s}|\xi) \pi(\sigma_1^2, ..., \sigma_K^2) \\
&\propto p(\boldsymbol{y}|\boldsymbol{s}, \mu_1, ..., \mu_k, \sigma_1^2, ..., \sigma_K^2) \pi(\sigma_1^2, ..., \sigma_K^2) \\
&= p(\boldsymbol{y}|\boldsymbol{s}, \mu_1, ..., \mu_k, \sigma_1^2, ..., \sigma_K^2) \times \prod_{k=1}^{K} \pi(\sigma_k^2)
\end{aligned}
$$

$$
= \prod_{k=1}^{K} \sigma_k^{2\left(-\frac{N_k}{2}\right)} exp\left\{-\frac{1}{2\sigma_k^2} \sum_{t:s_t=k} (y_t - \mu_k)^2\right\} \times \prod_{k=1}^{K} \frac{\beta}{\Gamma(\alpha)} \sigma_k^{2(-\alpha-1)} exp\left\{-\frac{\beta}{\sigma_k^2}\right\}
$$

$$
\propto \prod_{k=1}^{K} \sigma_k^{2\left(-\alpha-\frac{N_k}{2}-1\right)} exp\left\{-\frac{1}{2\sigma_k^2}\left(\sum_{t:s_t=k} (y_t - \mu_k)^2\right) - \frac{\beta}{\sigma_k^2}\right\}
$$

$$
= \prod_{k=1}^{K} \sigma_k^{2\left(-\alpha-\frac{N_k}{2}-1\right)} exp\left\{-\frac{1}{\sigma_k^2}\left(\beta + \sum_{t:s_t=k} \frac{(y_t - \mu_k)^2}{2}\right)\right\}
$$

So, we can see that σ_k^2 is independent in its full conditional distributions and that

$$
\sigma_k^2|\boldsymbol{y}, \boldsymbol{s}, \mu_k, \xi \sim invgamma\left(\alpha + \frac{N_k}{2}, \beta + \sum_{t:s_t=k} \frac{(y_t - \mu_k)^2}{2}\right)
$$

Franco, A. & Fantozzi, F. (2015) Analysis and clustering of natural gas consumption data for thermal energy use forecasting. *Journal of Physics: Conference Series*. 655.

Frühwirth-Schnatter, S. (2006) *Finite mixture and markov switching models*.

Grid, N. (2019) *Enabling the gas markets plan*.

Grid, N. (2016) *Gas demand forecasting methodology*.

Heaps, S.E., Farrow, M. & Wilson, K.J. (2020) Identifying the effect of public holidays on daily demand for gas. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 183 (2), 471–492.

Huntington, H.G. (2007) Industrial natural gas consumption in the united states: An empirical model for evaluating future trends. *Energy Economics*. 29 (4), 743–759.

Laib, O., Khadir, M.T. & Mihaylova, L. (2018) A gaussian process regression for natural gas consumption prediction based on time series data. *2018 21st International Conference on Information Fusion*. 55–61.

Lyness, F.K. (1984) Gas demand forecasting. *Journal of the Royal Statistical Society: Series D (The Statistician)*. 33 (1), 9–21.

Sánchez-Úbeda, E.Fco. & Berzosa, A. (2007) Modeling and forecasting industrial end-use natural gas consumption. *Energy Economics*. 29 (4), 710–742.

Soldo, B. (2012) Forecasting natural gas consumption. *Applied Energy*. 92 (C), 26–37.

Sönnichsen, N. (2021) *Natural gas consumption for energy use in the united kingdom (UK) 1970-2019*.

Timmer, R. & Lamb, P. (2007) Relations between temperature and residential natural gas consumption in the central and eastern united states. *Journal of Applied Meteorology and Climatology*. 461993–2013.

UKOOG (2019) *Natural gas uses*.

Vitullo, S.R., Brown, R.H., Corliss, G. & Marx, B.M. (2009) Mathematical models for natural gas forecasting. *Can. Appl. Math. Q.* 17 (4), 807–827.

Wang, X., Luo, D., Liu, J., Wang, W. & Jie, G. (2017) Prediction of natural gas consumption in different regions of china using a hybrid MVO-NNGBM model. *Mathematical Problems in Engineering*. 20171–10.

Zhu, Q., Lu, Q., Zhou, X. & Lai, K.K. (2014) A driving force analysis and forecast for gas consumption demand in china. *Mathematical Problems in Engineering*. 2014 (1), 1–11.