

Prediction of Intermitted Gas Flows in Large Gas Networks

Inken Gamrath Milena Petkovic

June 26, 2019

Abstract

About a fifth of the whole energy demand of Germany is supplied by natural gas. Additionally, there is a huge amount of gas just transiting Germany to other countries. However, the network operators, responsible for transporting the gas, have no knowledge of the intentions and planned actions of the gas traders. Thus, the number one priority of the network operators is to ensure security of supply.

We worked with one of Germany's largest network operators to improve the hourly forecast of gas flow at each node during the next 24 hours. Some of the nodes have intermitted gas flows and have therefore the additional challenge to decide whether they have flow at all besides the prediction of the amount of flow.

This article will report a method to predict zeros and one to predict gas flows. Furthermore it will combine both methods to predict intermitted gas flows.

1 Introduction

Since 2005 the transport and trading of natural gas has to be conducted by separate independent companies. The duty of the network operators is to operate the long distance transport network as well as ensuring the security of the network, i.e., that all demands are met. On the contractual level all gas transports have to be balanced, i.e., if gas gets out of the network the same amount has to be induced into the network. Due to the virtual trading point market model, the network nodes where the gas is injected might be very far from the points where it is taken out of the network. Furthermore, the inflow might be actually happen after the outflow. However, since gas flows really slow through the network, the network can act as a buffer. Additionally, consumers like municipal power stations just draw gas out of the network as needed. Nevertheless, they are indentured to also buy the same amount of gas to supply the network.

To ensure security of supply it is important for the network operator to have a precise forecast of the upcoming situation. When a local shortage seems to appear, balancing energy should be bought by the network operator. The situation is going to become worse, once the future gas driven peak load power stations go online, delivering electricity in peak load situations. This means a huge amounts of gas might be drawn from the network on short notice. It is the obligation of the operators to monitor the situation, foresee possible shortages

and react accordingly to ensure safety of supply. Since changes in gas networks happen rather slow it is therefore extremely important to have accurate forecasts on the demands and supply of the network to be able to react on time.

We worked together with one of Germany’s biggest network operators to improve the hourly forecast for demand and supply of the next 24 hours.

Most literature related to forecasting gas demands is focusing on long term issues. There are quite some publications regarding electricity demand forecasts, (see, e.g. [Taylor, 2003], [Mirasgedis et al., 2006],[Alfares and Nazeeruddin, 2002], [Nguyen and Nabney, 2010]) but electricity behaves very different from gas. A survey on works related to natural gas forecast and published between 1949 and 2010 is presented by [Soldo, 2012] who evidences that there are only a few authors that propose models to predict hourly gas flow.

The article is structured as follows: The next section describes the data that was used in this study. Section 3 will give details on the forecast methods we used, and Section 4 will give our reasoning how to evaluate the results as well as our computational experiments. Finally, we try to draw some conclusions.

2 Data

Open Grid Europe GmbH (OGE) supplied us with almost two years of recent hourly measurements of all exit and entry points to the system. Additionally, we were given average daily temperatures as measured at the nodes. And of course we know the day of the week for the nodes.

The nodes in the network can have totally different behavior. They can

- be either entries, exits, or both
- have typically daily or seasonal patterns, or no patterns at all
- have different behavior on workdays, weekend, or holidays
- be temperature dependant
- have a large number of zeros or none at all
- show a high variability

In this study, gas flows are not scaled and, for this reason, nodes with higher flows have higher prediction errors. As can be seen in Figure 1 some nodes have almost continuous flow, while other are active only occasionally.

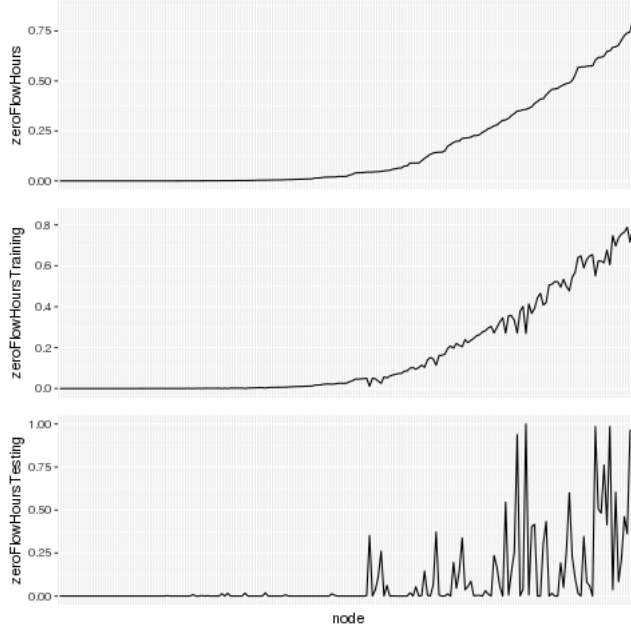
Especially, once the amount of zero flow hours is above 50%, the question whether there is flow at all becomes a question of its own.

For this study, we choose all nodes with 30 - 80 % of zeros for which the behavior does not change too much during the time. The chosen nodes are shown in 2

3 Methods

Let us denote the set of hours by $H = \{1, \dots, 24\}$, the set of days by D , the set of measured values by $m_{d,h} \in M$, and the temperatures by $t_d \in T$.

Figure 1: Percentage of zero flows of the nodes



Our goal is to predict flow values $p_{d,1}$ to $p_{d,24}$ for a given $d \in D$. Of course only data for days before d can be used.

3.1 Zero Detection

To detect zeros we use Nonlinear Autoregressive Networks with exogenous inputs (NARX). These are recurrent dynamic networks with feedback connections enclosing several layers of the network. Dynamic Neural Networks represent a class of Artificial Neural Networks many times proven to be well suited for modeling nonlinear systems, especially for time series due to their capability of accommodating the dynamic and complex nature of real-world time series prediction problems. ([Menezes and Barreto, 2008], [Xie et al., 2009]).

The NARX model is based on the linear ARX model, which is commonly used in time-series modeling.

The equations for the NARX model are:

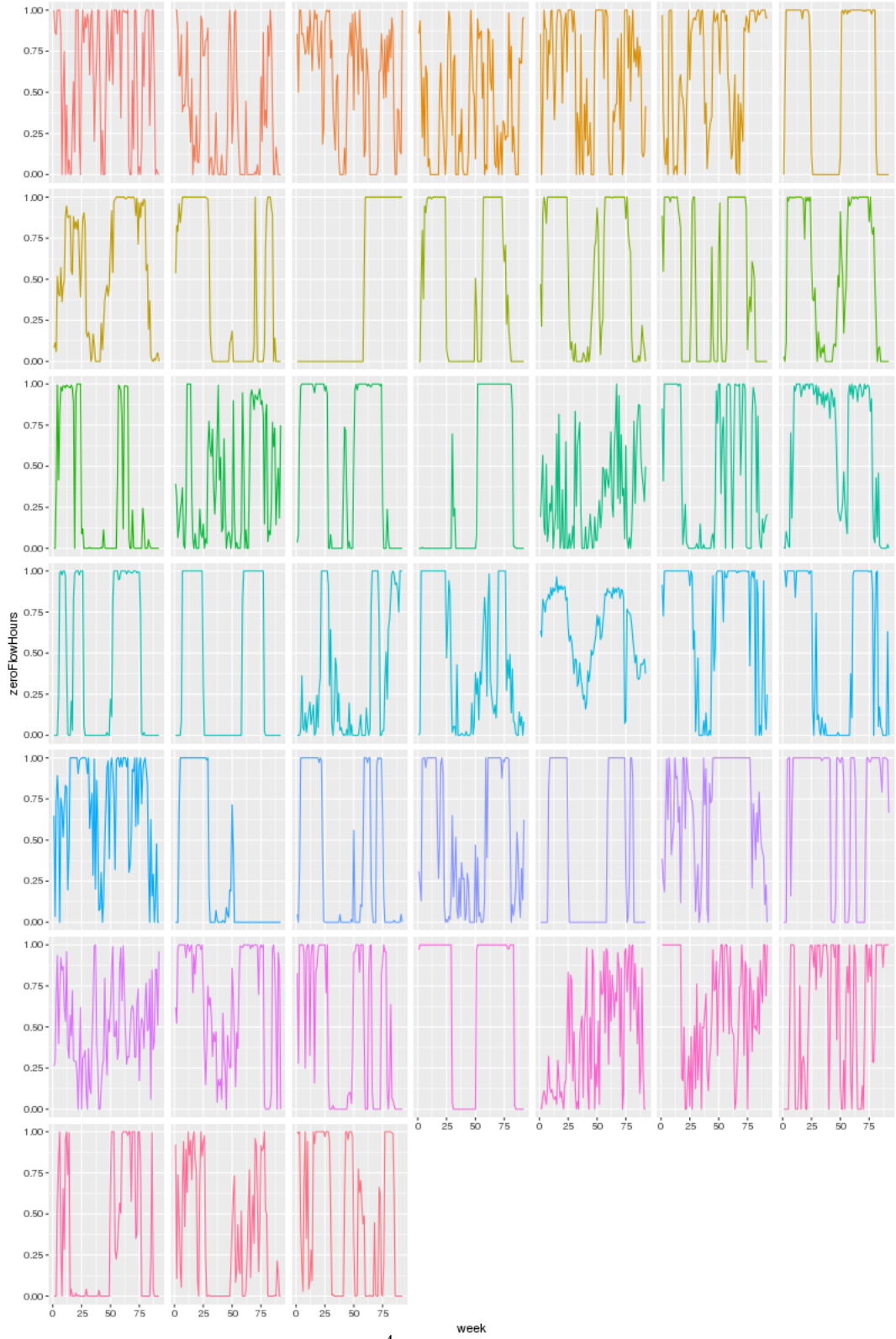
$$p_{d,h} = f_h(M_{d-1}, U_{d-1}, V_d) \quad (1)$$

with

$$\begin{aligned} M_{d-1} &= \{m_{d-1,24}, m_{d-1,23}, \dots, m_{d-i_m, j_m}\}, \\ U_{d-1} &= \{u_{d-1,24}, u_{d-1,23}, \dots, u_{d-i_u, j_u}\}, \\ V_d &= \{v_1, \dots, v_{52}\} \end{aligned}$$

where the next value of the dependent output $p_{d,h}$ is calculated based on the previous values of the output and the previous values of exogenous inputs u and v . For detecting zero flow for nodes with high percentage of zero flow

Figure 2: Percentage of zero flows of the nodes



hours we implemented this equation by using a feed forward neural network to approximate the underlying functions f_h .

The parameters of the NARX model as well as several exogenous inputs were chosen based on the sensitivity analysis. The obtained model has 2 hidden layers both with 100 neurons and 2 lags for output and exogenous inputs. For the transfer function Tansig function is chosen, number of Epoch is 600 and the training function is Scaled Conjugated Gradient Backpropagation (SCG). The final model is the multiple-input multiple-output model with 24 outputs, each one representing forecast of hourly flow in the next 24 hours. Using the multiple output model we avoided the accumulation of errors that will be made by multi-step ahead recursive strategy and at the same time we benefit from modeling the stochastic dependency between predicted values of the time series ([Taieb et al., 2010]). For the exogenous inputs we used the binary feature defined by:

$$u_{d,h} = \begin{cases} 0, & \text{if } m_{d,h} = 0 \\ 1, & \text{otherwise} \end{cases}$$

Also, we used 52 exogeneous features describing the date: Month (12 binary variables), Day of the month (31 binary variables), Day of the week (7 binary variables), Weekend (1 binary variable), Holiday (1 binary variable).

3.2 Predicting Gas Flows

To predict flows—supplies and demands of the gas network—we use a weighted sum of features

$$p_{d,h} = \sum_{i \in F_h} w_{h,i} f_{h,i}(d) \quad (2)$$

where $w_{h,i}$ define the weights, and $f_{h,i}$ are our features. Part of the features $i \in F_h = \{1, \dots, n_h\}$ are defined as historical flow values, the other features are exogenous variables, like mean flow values or temperatures. We use a different set of features for each hour of the day $h \in H = \{1, \dots, 24\}$.

For each day in the test set the forecasted flow values are computed by first computing the weights via a linear program and then using the weighted sum (2) of features for each hour to forecast the flow values. The linear program uses historical data of the last 16 weeks to forecast historical flow values, compares these forecasted values to the real historical flow values and minimizes the absolute error. For the computation of the forecasted flow values, it might be that also forecasted flow values of prior hours are used as input values, if the corresponding hours do not lie in the past.

In the linear program for each $h \in H$ the error $e_{d,h}$ of the prediction is calculated by

$$\sum_{i \in F} f_{h,i}(d) \cdot w_{h,i} - m_{d,h} = e_{d,h}$$

where $w_{h,i} \in [-2, 2]$.

For each hour $h \in H$ the best weights are chosen by taking the ones which minimize the sum of the absolute errors:

$$\min \sum_{d \in D, h \in H} |e_{d,h}|$$

Furthermore, we want to reduce the over- and under-estimation of the flow values:

$$\sum_{d \in D, h \in H} |e_{d,h}| \leq \epsilon$$

The solution of the linear model leads to functions with which we can calculate the flow values for the different hours of the day. It describes a iterative multi-step ahead multiple-input procedure with exogenous inputs.

Training: Feature Selection

In the training procedure of this method, a slightly different model is used which chooses for each hour the features which are important, so that we do not get the problem of over-fitting in the linear program. Therefore, we add additional binary variables $x_{h,i}$ to the problem, which determine whether feature i is chosen for hour h , i.e., whether the weight of feature i and hour h is not equal to zero. Then, we need to link these variables to the weight variables

$$-2 \cdot x_{h,i} \leq w_{h,i} \leq 2 \cdot x_{h,i}$$

and limit the number of chosen features by B

$$\sum_{i \in F} x_{h,i} \leq B$$

for all hours $h \in H$. The solution of the resulting mixed integer program leads for each hour h to one feature set F_h which will be used in the linear program.

4 Testing and Results

We will compute for each predicted node the following 4 values to compare the performance of the predictions.

Information on the zero/non-zero behavior of the node

- Zero-Hit-rate: Percentage of hours where both the measurement and the prediction are both not more than a given threshold away from zero.
- Non-Zero-Hit-rate: Percentage of hours where both the measurement and the prediction are both outside the given threshold around zero in the same direction.

Those two numbers give a good idea regarding the zero-flow behavior.

Here, the Non-Zero-Hit-rate is more important than the Zero-Hit-rate, because all Non-Zero-Hits are forecasted as zero and can't be corrected by another prognosis.

The errors are only computed on hours that are a hit regarding the non-zero-hit-rate for all prognosis methods. These are:

- Mean Absolute Error
 $MAE := 1/n \sum_{h \in H, d \in D} |p_{d,h} - m_{d,h}|$
- Mean Absolute Percentage Error
 $MAPE := 1/n \sum_{h \in H, d \in D} \frac{|p_{d,h} - m_{d,h}|}{|m_{d,h}|} \cdot 100$

This gives us 4 values describing the performance of our prognosis.

4.1 Results

5 Conclusions

6 Acknowledgments

This work was supported by the *Research Campus Modal* funded by the German Federal Ministry of Education and Research (fund number 05M14ZAM).

References

- [Alfares and Nazeeruddin, 2002] Alfares, H. and Nazeeruddin, M. (2002). Electric load forecasting: Literature survey and classification of methods. *International Journal of Systems Science*, 33(1):23–34.
- [Menezes and Barreto, 2008] Menezes, J. M. P. and Barreto, G. A. (2008). Long-term time series prediction with the narx network: An empirical evaluation. *Neurocomputing*, 71(16):3335 – 3343. Advances in Neural Information Processing (ICONIP 2006) / Brazilian Symposium on Neural Networks (SBRN 2006).
- [Mirasgedis et al., 2006] Mirasgedis, S., Sarafidis, Y., Georgopoulou, E., LalasD.P., Moschovits, M., Karagiannis, F., and Papakonstantinou, D. (2006). Models for mid-term electricity demand forecasting incorporating weather influences. *Energy*, 31(2):208 – 227.
- [Nguyen and Nabney, 2010] Nguyen, H. T. and Nabney, I. T. (2010). Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models. *Energy*, 35(9):3674 – 3685.
- [Soldo, 2012] Soldo, B. (2012). Forecasting natural gas consumption. *Applied Energy*, 92:26 – 37.
- [Taieb et al., 2010] Taieb, S. B., Sorjamaa, A., and Bontempi, G. (2010). Multiple-output modeling for multi-step-ahead time series forecasting. *Neurocomputing*, 73(10):1950 – 1957. Subspace Learning / Selected papers from the European Symposium on Time Series Prediction.
- [Taylor, 2003] Taylor, J. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, 54(8):799–805.
- [Xie et al., 2009] Xie, H., Tang, H., and Liao, Y.-H. (2009). Time series prediction based on narx neural networks: An advanced approach. In *2009 International Conference on Machine Learning and Cybernetics*, volume 3, pages 1275–1279.