Paper 3178-2019

SNAM Gas Forecasting with Ensemble Method and Neural Networks

Andrea Magatti, BIP Business Integration Partners; Silvia Lameri, SNAM; Gabriele Oliva, BIP Business Integration Partners; Gabriella Jacoel, BIP Business Integration Partners; Carlo Binda, BIP Business Integration Partners

ABSTRACT

Natural gas is the most important energy source in Italy, fueling domestic heating, industrial facilities, and thermoelectric power plants. Gas demand forecasting is a critical task for energy providers since it ensures safe, reliable and efficient operational planning and impacts gas prices as well as future investment requirements. This paper details a Machine Learning approach to forecasting the daily Italian gas demand, based on an ensemble of artificial neural networks developed utilizing SAS® software. This task results particularly difficult in Italy due to the complicated structure of the gas network as well as volatile weather patterns. The focus will be on the one-day-ahead model, subject to government regulations that impose the quality of the prediction by charging providers with a fee proportional to the daily percentage error.

INTRODUCTION

Natural gas is in the midst of a rapid growth phase. Since 2010, average global gas consumption has grown by 1.8% per year, making it the fastest growing energy source other than renewable power. This growth is a result of the multiple benefits offered by gas as a clean, abundant, flexible, and cost-effective fuel.

In Italy, natural gas is the most common fuel for both power plants and domestic heating. Furthermore, it is used for heating and powering productive processes by various large industrial facilities. In 2018 over 70 billion cubic meters of natural gas were consumed, with an overall increase in demand of 11% compared to 2015. Out of the total gas demand in 2018, about 35% was due to thermoelectric power plants, 20% to industrial facilities and 45% to residential users.

The transport of natural gas in Italy occurs at two main levels. The first, called "primary distribution" (Figure 1), involves transport at a national and regional scale through large pipelines. The second level or "secondary distribution" supplies gas locally for domestic use through a widespread system. Primary distribution is guaranteed by a network of over 32 thousand kilometers of pipelines that spans the whole of Italy, apart from Sardinia. The entity entrusted with transporting energy through primary distribution is defined the transmission system operator (TSO). A TSO receives gas from producers or shippers, transports it via pipeline through an area, and delivers it to second level gas distribution companies or directly to industries and power plants.

Natural gas is introduced into the Italian national network at eight entry points, where the network connects to the import pipelines and at the liquefied natural gas (LNG) regasification terminals. Italy's major supplier is Russia, amounting to more than 45% of importations, followed by Algeria with 32%. Natural gas is also produced in Italy, with it being injected into the network at over 50 entry points.

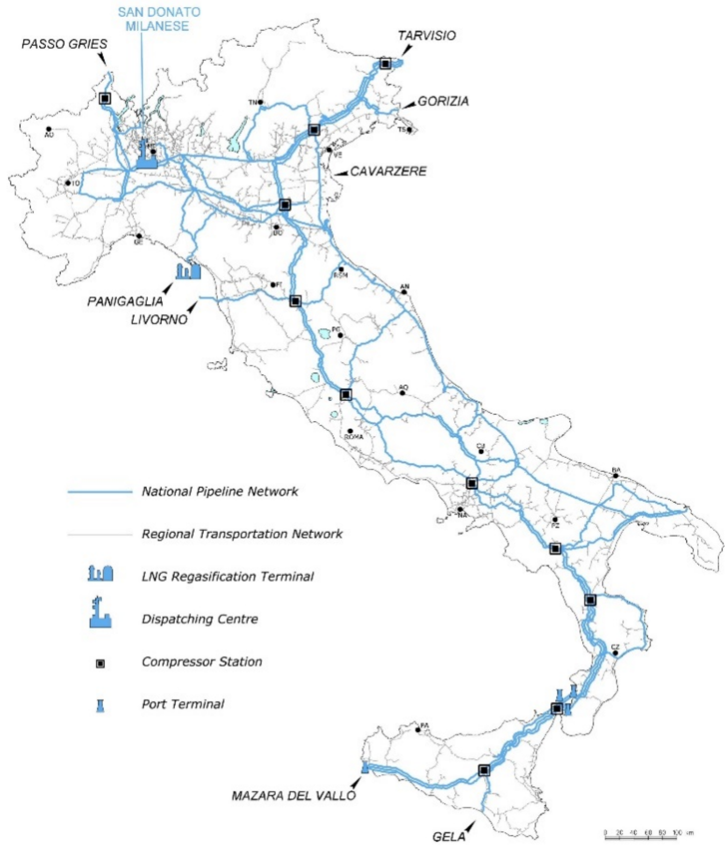

Figure 1 – The Italian primary distribution network

Due to the cost of establishing and maintaining a national transmission infrastructure, a TSO is typically a monopoly and as such is often subjected to regulations. Amongst these procedures, the TSO is required to produce gas demand forecasts for the entire network. Demand forecasting can be classified according to the prediction horizon, ranging from hourly to yearly, and the reference area, which can correspond to a single node of the network up to the total national demand. A variety of approaches can be adopted, from mathematical models based on fluid dynamics to time series models that evaluate demand trends to more complex Machine Learning models that allow a large assortment of input variables. Gas demand forecasting is a critical task for energy providers since these values are used in a wide variety of contexts:

- Operational planning – demand forecasts feed directly into the assessment of the natural gas infrastructure, to ensure its reliability in satisfying gas demand. They provide relevant information to effectively reserve pipe capacity and plan stocks, guaranteeing that sufficient gas is available to support the safe operation of the network;
- Network balancing - energy regulations impose the balance of the network by charging providers with a fee proportional to their unbalanced quantity;
- Pricing - demand forecasts are used in the setting of natural gas futures prices;
- Investment planning - demand forecasts provide insights on network developments, the evolution of gas tariffs and a cross-check on customers' expectations of the evolution of the market.

Our study focuses on the day-ahead forecasting of the daily Italian gas demand. The TSO initially generated forecasts using time series models (see Figure 2): resulting efficient and simple to develop, they satisfied the operational planning needs of energy providers. However, due to the continuous development of the gas network infrastructure and the creation of new protocols associated to network balancing and gas future prices, during the course of 2017 it became apparent that a new and more advanced approach was required. The Italian government in fact promulgated a new regulation that imposed the quality of the day-ahead prediction: the daily forecasts are required to have a percentage error below 5% to avoid a large penalty, with desired average performances of 3.5% during the winter months and 4.5% during the summer months.

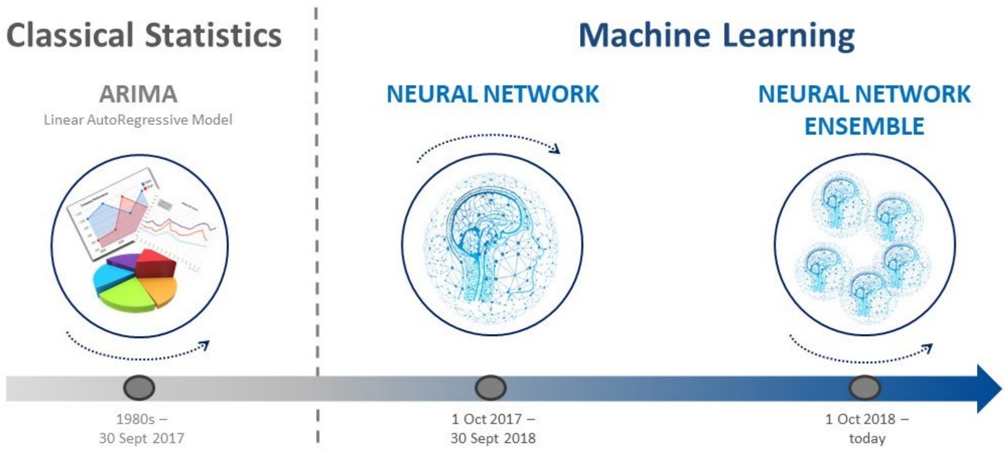

Figure 2 – Evolution of day-ahead forecasting for the daily Italian gas demand

In October 2017 we adopted a Machine Learning approach to exploit a broader set of data and developed a new day-ahead forecasting model based on artificial neural networks. The input data consisted in real-time gas transportation data, demand history, calendar, and basic weather forecasts. The model performed well during the winter months, with an average percentage error below 3%, but in April 2018 volatile weather patterns and unprecedented electricity generation led to bizarre gas demand. The model was unable to capture these unique patterns, leading to a monthly performance above 9% and large government penalties. Even though the model performed well during the remaining summer months, we decided to introduce new data and improve the model. Using SAS® software, we developed a system that provides hourly forecasts utilizing a "Mixture of Experts" ensemble of neural networks called DAFNE (Dynamically Adjusted Forecasting Neural Ensemble). DAFNE is an ensemble of 6 models characterized by different data sources and model architectures. These models then contribute to the final prediction based on dynamically adjusted weights, increasing both the accuracy and the stability.

THE ENSEMBLE APPROACH

Due to the Italian government introducing a regulation that imposes the quality of the day-ahead forecast, requiring that the daily predictions have a percentage error below 5% to avoid a large penalty, in October 2017 we developed a new day-ahead forecasting model. This model was based on artificial neural networks and implemented utilizing SAS/STAT®, SAS/ETS® and Enterprise Miner™ software. The input data consisted in:

- demand history – past values of the daily gas demand. Gas consumption fluctuates according to the season: volumes during the winter are two to three times the volumes during the summer months. This is because domestic heating is typically switched off when the temperature climbs above 18°C. Furthermore, we extracted the seasonality through time series decomposition and discovered that gas demand has a yearly seasonality (Figure 3), due to the temperature and heating requirements, as well as weekly seasonality (Figure 4), with higher volumes during working days compared to weekends and holidays;

Figure 3 – Yearly seasonality of gas demand

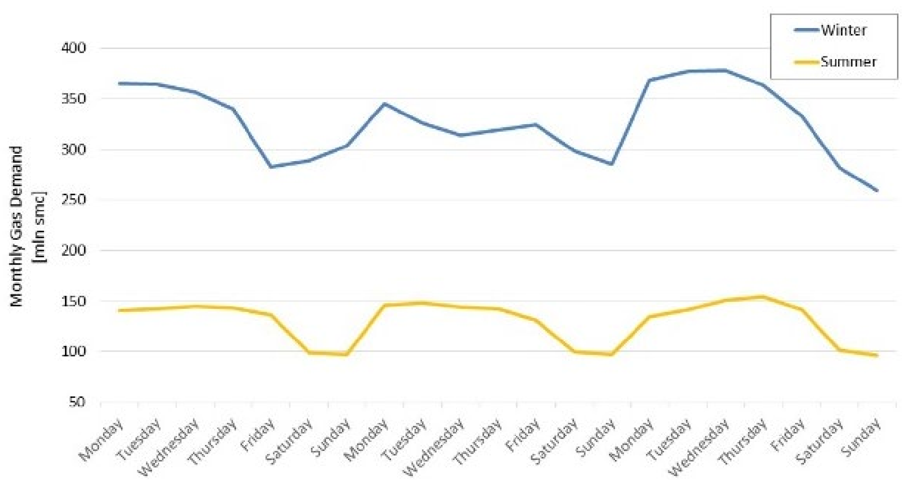

Figure 4 – Weekly seasonality of gas demand

- real-time SCADA gas transportation data – hourly aggregations of real-time SCADA (Supervisory Control and Data Acquisition) measurements of volumes, per entry and exit point of the network as well as storage and pipe capacity;
- historic operational business data – official transported gas balances, containing all the detail about the total intake (gas entering the network) and total offtake (gas being withdrawn from the network). For example, we have the historic values of imports, national production, storage systems, exports and consumption per user;

- calendar – as previously described, gas demand has two levels of seasonality. To capture these phenomena, we introduced a series of binary calendar variables that indicate the period of the year as well as the day type. Some examples are the month and the weekday, encoded as dichotomic features, and extended holidays, a binary feature with value 1 on holidays and weekends;
- weather forecasts – day-ahead weather forecasts of temperature, humidity and precipitation for the main Italian cities as well as sub-regional areas of Italy. They are received from the provider in the morning and remain constant through-out the day.

The model has a rolling operating mode, since it is continuously retrained with the most recent data. We also optimized the hyperparameters and tested the model on a year of data (specifically October 2016 – September 2017), obtaining a yearly mean absolute percentage error or MAPE of 4% with daily error peaks a little below 20%. Once deployed, this model performed well during the winter months, with a MAPE below 3% that respected the desired government performance. However, the model encountered difficulties in April 2018: volatile weather conditions, with temperatures varying greatly over the course of the day, impacted demand associated to domestic heating while unusual electricity generation, with only renewable sources being used, led to extremely low demand from thermoelectric power plants. The model was unable to capture these unique patterns, leading to a monthly MAPE above 9% with various error peaks above 20%. Even though the model performed well during the remaining summer months (see Figure 5) and had a yearly MAPE of 4%, the events in April had convinced us to make some improvements. We set out to develop a new forecasting system, with the goal of incrementing the stability as well as the accuracy of the forecasts. We also wanted to introduce new data, such as electricity demand forecasts and more detailed weather forecasts.

Performance Indicators	1 Oct 2017 – 30 Sep 2018
Average percentage error (MAPE)	4.05%
Variability of error (standard deviation of MAPE)	3.7%
Maximum percentage error (MAPE Max)	23.9%
N° occurances of error below 10%	93%

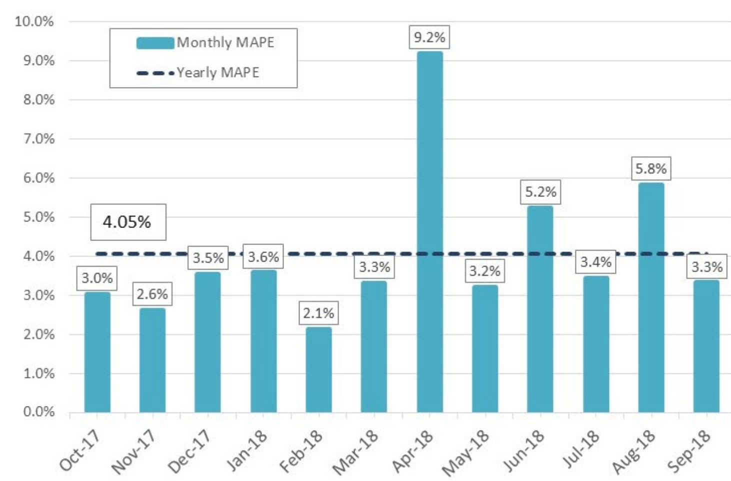

Figure 5 – Gas demand forecasting performances Oct 17 – Sep 18

DYNAMICALLY ADJUSTED FORECASTING NEURAL ENSEMBLE (DAFNE)

On October 1st, 2018 we deployed the new day-ahead gas demand forecasting model called DAFNE or Dynamically Adjusted Forecasting Neural Ensemble. Completely developed using SAS/STAT®, SAS/ETS® and Enterprise Miner™ procedures, the system provides hourly forecasts utilizing a "Mixture of Experts" ensemble of neural networks. Specifically, DAFNE is an ensemble of 6 sub-models characterized by different data sources and model architectures. These sub-models then contribute to the final prediction based on dynamically adjusted weights calculated on the past performances. The input data consists in the previously described sources as well as two new ones:

- electricity demand forecasts – day-ahead forecasts of electricity demand per power station, which is then cross-referenced with the power plants that utilize gas;
- more detailed weather forecasts – extension of the previously described day-ahead predictions to include more detailed sub-regional areas of Italy and updates of all forecasts during the course of the day.

When designing the model, we decided to build a set of models able to forecast gas demand utilizing different points-of-view. To do this we specialized each model on different input data, therefore obtaining six independent experts (see Table 1). Specifically, we developed the following sub-models:

1. Autoregressive model – neural network that reproduces the logics of an autoregressive model by specializing itself in demand history, as well as considering a few other regressors associated to calendar and weather;
2. Company model – neural network corresponding to the first model deployed in October 2017. It is the expert that gives equal weight to operational business data and real-time SCADA gas transportation data;
3. SCADA model - neural network specialized in real-time SCADA gas transportation data, while also considering calendar, weather and demand history;
4. Complete model – model designed to evaluate all data sources, giving similar weights to each input;
5. Similarity model – model based on the notion of similar day, distinguishing between work-days and extended holidays (holidays and weekends). It is actually composed of two neural networks, with complementary training sets, that switch on and off according to the type of the forecast day;
6. Electricity model – neural network that focuses on the fluctuations of electricity demand and therefore gas demand associated to thermoelectric power plants. This expert is particularly important during the summer.

Sub-model	Input data					
	Demand history	SCADA data	Calendar	Business data	Weather	Electricity demand
Autoregressive	✓		✓		✓	
Company	✓	✓	✓		✓	
SCADA	✓	✓	✓	✓	✓	
Complete	✓	✓	✓	✓	✓	✓
Similarity	✓	✓	✓		✓	✓
Electricity	✓		✓		✓	✓

Table 1 – Input data of the DAFNE sub-models

The sub-models are continuously retrained with the most recent data, allowing the system to have a rolling operating mode. We optimized the hyperparameters of each sub-model as well as the logics of the ensemble before testing the system on a year of data (specifically October 2017 – September 2018), obtaining a yearly mean absolute percentage error or MAPE of 2.7% with daily error peaks around 15%.

THE ENSEMBLE ALGHORITM

DAFNE is a "Mixture of Experts" ensemble with the peculiarity that every sub-model contributes to the final prediction based on its past performance, where the performance is calculated considering a dynamic time period. Specifically, the sub-model performances are characterized by evaluating the mean absolute percentage error in five different situations:

- local MAPE: evaluates the short-term performances since it corresponds to the mean absolute percentage error of the last k days of the current month, with k reaching at most a value of 14 days;
- historic MAPE: evaluates the historic performances since it corresponds to the mean absolute percentage error obtained during the current month in past years;
- work-day MAPE: evaluates the performances on work-days, since it corresponds to the mean absolute percentage error calculated on all past work-days;
- extended holiday MAPE: evaluates the performances on extended holidays, since it corresponds to the mean absolute percentage error calculated on past weekends and holidays;
- holiday MAPE: evaluates the performances on holidays, since it corresponds to the mean absolute percentage error calculated on Italian public holidays.

The algorithm used to determine the weight of each sub-model and therefore generate the final prediction can be summarized in two steps:

1. Determine the sub-model weight – for each sub-model, we have that the local MAPE and historic MAPE evaluate the performance during the current month while the remaining three MAPE evaluate the performance based on the day type of the reference date of the prediction. These five indexes are aggregated using a weight system that is a function of the reference date of the prediction (see Figure 6). We can identify three different phases:
    - Phase 1 – for day-ahead predictions referred to the first two days of the month, the local MAPE can't be calculated since the actual gas demand is not yet available: the actual demand of the first day of the month is available on the second day, when we are already generating the day-ahead forecast referred to the third day of the month. For this reason, the monthly performance must be evaluated using exclusively the historic MAPE. In this situation the historic MAPE has a weight of 50% in determining the overall sub-model weight. The remaining 50% is assigned based on the day type performances: on work-days we consider the work-day MAPE, on weekends we consider the extended holiday MAPE, and on holidays we consider the weighted average between the extended holiday MAPE and the holiday MAPE (weights respectively ¼ and ¾);
    - Phase 2 – For day-ahead forecasts referred to the third to sixteenth day of the month, the local MAPE is not at full capacity since less than 14 actuals are available. To calculate the monthly performance, it is necessary to use a weighted average of the local and the historic MAPE (the weights are proportionate to the number of actuals). In this situation the historic MAPE has a weight of 60% in determining the overall sub-model weight. The remaining 40% is assigned based on the day type performances: on work-days we consider the work-day MAPE, on weekends we consider the extended holiday MAPE, and on holidays we consider the weighted average between the extended holiday MAPE and the holiday MAPE (weights respectively ¼ and ¾);

- Phase 3 – For day-ahead forecasts referred to the sixteenth day of the month onwards, we have over 14 actuals available and therefore have a local MAPE at full capacity. In this situation the monthly performance is evaluated using exclusively the local MAPE and has a weight of 60% in determining the overall sub-model weight. The remaining 40% is assigned based on the day type performances, as previously explained.

Day of the Month	1	2	3	4	...	15	16	17	...
Day Type	Work-day	Weekend	Holiday	Work-day		Work-day	Weekend	Holiday	
local MAPE	0%	0%	$\frac{1}{14}$ ***60%**	$\frac{2}{14}$ ***60%**	...	$\frac{13}{14}$ ***60%**	**60%**	**60%**	...
historic MAPE	**50%**	**50%**	$\frac{13}{14}$ ***60%**	$\frac{12}{14}$ ***60%**	...	$\frac{1}{14}$ ***60%**	0%	0%	...
work-day MAPE	**50%**	0%	0%	**40%**	...	**40%**	0%	0%	...
ext. holiday MAPE	0%	**50%**	**8%**	0%	...	0%	**40%**	**8%**	...
holiday MAPE	0%	0%	**32%**	0%	...	0%	0%	**32%**	...

Phase 1 Phase 2 Phase 3

Figure 6 – MAPE aggregation logics to determine the sub-model weight

2. Determine the final prediction – DAFNE's prediction corresponds to the weighted average of the sub-model forecasts, where each sub-model is assigned a weight that is dynamically adjusted based on past performances and the reference data of the prediction as previously described.

$$F_{DAFNE} = \frac{1}{\sum_{i=1}^{6} w_i}\left(w_1 F_{Auto} + w_2 F_{Company} + w_3 F_{Scada} + w_4 F_{Complete} + w_5 F_{Sim} + w_6 F_{Elec}\right)$$

The ensemble algorithm, while dynamic, is rule-based and therefore deterministic. We could not use a stacking approach, that is train a new model which learns from the sub-model predictions, because of the limited amount of training data. The evaluations of the current algorithm are made utilizing forecasts from October 2015 to today, and we could not extend this period further into the past due to restrictions caused by the availability of the input data.

SIMULATED PERFORMANCES OF SUB-MODELS AND DAFNE

Each of the sub-models was optimized on a validation set (Table 2) consisting in the period October 2016 – September 2017 before being tested (Table 3) on the period October 2017 – September 2018. We selected those models that had the lowest MAPE on the validation set and that obtained similar results on the test set. This optimization phase was carried out using a grid search implemented with moving window logics and monthly retraining of the models to reproduce the rolling operating mode.

	Simulated Performances: Validation (Oct 16 – Sep 17)			
	MAPE	**MAPE Max**	**MAPE StdDev**	**N° errors < 10%**
Autoregressive	3.93%	25.6%	3.4%	95.1%
Company	4.01%	17.5%	3.3%	93.7%
SCADA	4.04%	17.9%	3.3%	93.4%
Complete	3.26%	17.9%	2.6%	97.8%
Similarity	3.56%	18.8%	2.8%	97.5%
Electricity	3.54%	19.4%	3.0%	96.4%

Table 2 – Sub-model performances on the validation set (Oct 16 – Sep 17)

	Simulated Performances: Test (Oct 17 – Sep 18)			
	MAPE	**MAPE Max**	**MAPE StdDev**	**N° errors < 10%**
Autoregressive	4.58%	20.1%	4.2%	89.0%
Company	4.05%	21.2%	3.8%	92.1%
SCADA	4.00%	48.2%	4.2%	92.3%
Complete	3.44%	16.3%	2.8%	97.3%
Similarity	3.17%	18.2%	2.8%	96.7%
Electricity	3.48%	13.6%	2.6%	97.3%

Table 3 – Sub-model performances on the test set (Oct 17 – Sep 18)

After selecting the six sub-models, we applied the ensemble algorithm maintaining the same validation and test sets. The validation set was used to tune the logics of the ensemble, while the test set was used to confirm that no overfitting was occurring. The results (Table 4) confirmed that the DAFNE approach boosted both model accuracy and stability: the yearly MAPE was reduced below 3% while the standard deviation of the percentage error is below that of every sub-model.

	Simulated Performances			
	MAPE	**MAPE Max**	**MAPE StdDev**	**N° errors < 10%**
Validation (Oct 16 – Sep 17)	2.87%	13.8%	2.3%	99.2%
Test (Oct 17 – Sep 18)	2.73%	15.6%	2.5%	97.5%

Table 4 – DAFNE simulated performances

DEPLOYED PERFORMANCES OF SUB-MODELS AND DAFNE

On October 1st, 2018 we deployed the new day-ahead forecasting model DAFNE. The model was released into the production environment and began using live data. The models would therefore have to face typical operational issues such as missing values, automatically imputed by the system, or bad quality due to data collection problems of the underlying database.

Considering the period from deployment to February 2019, all sub-models have performed better than anticipated except for the Autoregressive model (Table 5). This may be due to the fact that, even though gas demand has a strong seasonality, the process has become less autoregressive and requires more variables to adjust the forecast.

	Deployed Performances (Oct 18 – Feb 19)			
	MAPE	MAPE Max	MAPE StdDev	N° errors < 10%
Autoregressive	5.34%	20.2%	4.4%	88.7%
Company	3.81%	17.6%	3.1%	94.7%
SCADA	3.15%	19.7%	3.0%	96.0%
Complete	2.66%	9.2%	2.1%	100.0%
Similarity	2.75%	14.4%	2.5%	99.3%
Electricity	3.38%	18.1%	3.2%	93.9%

Table 5 – Sub-model performances after deployment (Oct 18 – Feb 19)

DAFNE has also respected the simulation results (Table 6 and Figure 7), performing slightly better than expected. The system was able to handle the performance issues of the Autoregressive model through the ensemble algorithm, assigning less weight to this model.

	Deployed Performances (Oct 18 – Feb 19)			
	MAPE	MAPE Max	MAPE StdDev	N° errors < 10%
Oct-18	2.06%	7.8%	1.8%	100.0%
Nov-18	2.61%	9.3%	2.2%	100.0%
Dec-18	3.29%	13.0%	2.8%	96.8%
Jan-19	2.08%	11.4%	2.6%	96.8%
Feb-19	2.35%	8.5%	2.0%	100.0%
Overall	**2.48%**	**13.0%**	**2.3%**	**98.7%**

Table 6 – DAFNE performances after deployment (Oct 18 – Feb 19)

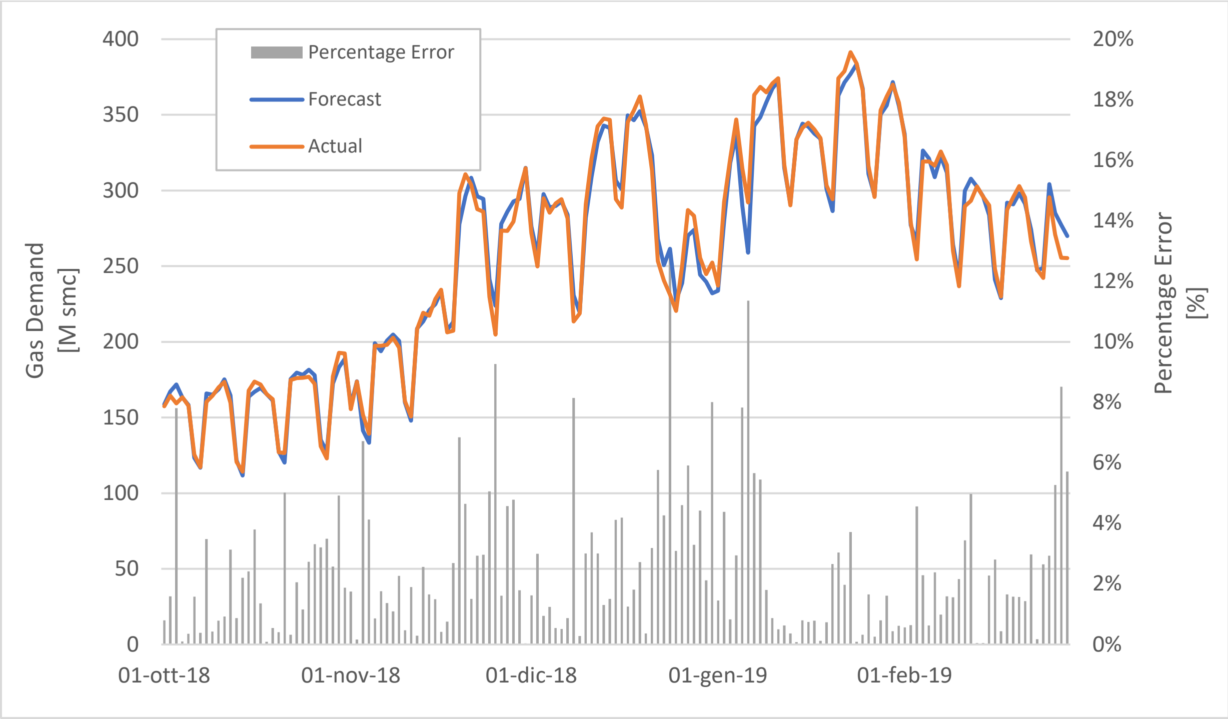

Figure 7 – Daily performances of the DAFNE system

It can be noted that the model had higher error peaks during the Christmas holidays, from roughly December 24th, 2018 to January 6th, 2019. This period, along with Easter and the week of Ferragosto (Italian holiday that occurs on August 15th), is the hardest to predict since there are many bridge holidays, the autoregressive dependency is irrelevant and there is a strong dependency on weather patterns.

CONCLUSION

The DAFNE system provides an effective solution for gas demand forecasting, guaranteeing reliable and efficient operational planning as well as providing dependable insights on gas prices and investment planning. On the one hand, the "mixture of Experts" approach leads to a boost in both accuracy and stability compared to the sub-models. On the other, the ensemble algorithm based on dynamically adjusted weights calculated on the past performances allows the system is to adapt itself to the evolving gas demand patterns.

DAFNE thus far has confirmed the simulation results and respected the government regulations, performing slightly better than expected. It will be interesting to see how the system handles the summer months, where gas demand largely depends on thermoelectric power plants due to the absence of domestic heating.

Finally, the ensemble algorithm, while dynamic, remains rule-based and therefore deterministic. DAFNE could benefit from the use of stacking, with the creation of a model trained directly on the sub-model predictions and performances.

REFERENCES

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.

Hastie, T., Tibshirani R., Friedman J. *The Elements of Statistical Learning*. NY: Springer Series in Statistics.

Hyndman, R., Athanasopoulos, G. 2013. *Forecasting: principles and practice*. OTexts.

Jacoel, G. "Predicting commodity consumption using Google trends and sentiment analysis of news." MA Thesis. Politecnico of Milan. April 2016.

Kasemsumran, S. 2006. "Moving window cross validation". Analyst, 4.

Oliva, G. "Forecasting U.S. Corn Yield Using Machine Learning." MA Thesis. Politecnico of Milan. April 2017.

Zhou, Zhi-Hua (2012). *Ensemble Methods: Foundations and Algorithms*. Chapman & Hall.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Andrea Magatti
Business Integration Partners S.p.a.
andrea.magatti@mail-bip.com

Gabriele Oliva
Business Integration Partners S.p.a.
gabriele.oliva@mail-bip.com

Gabriella Jacoel
Business Integration Partners S.p.a.
gabriella.jacoel@mail-bip.com

Carlo Binda
Business Integration Partners S.p.a.
carlo.binda@mail-bip.com