Improving Gas Load Forecast Accuracy – A Practical Approach

Sophie Jehaes, Fluxys Belgium, Fakher Raza, Energy Solutions UK

ABSTRACT

Fluxys operates as an independent natural gas transportation company at the centre of the European gas market. The Fluxys network and the Zeebrugge Hub are the very heart of cross-border natural gas flows in Western Europe, which has made Fluxys a key player in the gas operations services. The Zeebrugge hub, owned and operated by Fluxys, has an annual throughput of around 1412.6 MMMSCF (40 BSCM) – connecting to 17 neighbouring networks and about 11% of the total demand for Western Europe.

In the wake of gas market deregulation in Europe, Fluxys is committed to safe, reliable and flexible transportation of gas. To effectively anticipate the requirements of gas demands on its network, Fluxys use a Neural Network based forecaster. This strategic tool furnishes Fluxys with the information required to make critical decisions, resulting in greater flexibility, agility, and increasing customer service.

Although traditional methods of forecasting demand, such as Regression Analysis or Neural Networks, have proven to be very useful in the management of gas supply, they are highly dependent on the availability of influencing factors such as weather, calendar, economics and production plan. The quality of data available to the forecasting tool as well the utilization of that data within the tool plays a key role in the level of accuracy of the forecast. As such, if necessary data is unavailable, of poor quality or used inappropriately then the accuracy of the forecast is adversely affected, which has a direct impact on the business.

This paper examines various ways in which the data quality and utilization can be improved to more accurately predict gas demand. Methods to improve forecasting include intelligent selection of the Neural Network model, combination of weather effects, data cleaning (data pre-processing) and within-day forecasting after abrupt changes (data post-processing). These techniques involve calculated manipulation of both the input and the output data with a view to enhance the accuracy of the forecast. Fluxys incorporates this methodology as an integral part of the forecasting system to reduce supply risk and to drive operational efficiencies by significantly reducing forecast unavailability.

Each of these methods is described and the effects of applying each method and subsequent improvement of forecast are given. We also show that the correct use of each of these methods can lead to a consistent forecast accuracy of more than 95%.

INTRODUCTION

Technological advancements have allowed forecasters to choose forecasting methods and techniques suited to their particular application within the business. Accurate Short-Term forecasting not only demands choice of proven methods but also involves direct user intervention to streamline the utilization of the data. Using appropriate and timely data sets enhance the effectiveness of the chosen forecasting mechanism – the result is increased accuracies and productive business planning.

Achieving a reliable daily forecast is rather easy but achieving the same quality of hourly forecast is a real challenge for any gas dispatching team. Balancing supply and demand in today's competitive market has become an important issue both within-day as well as on daily basis. The challenge exist to focus more on tuning the data and external effects than exhausting resources in choosing core forecasting methodology.

Forecasting at Fluxys is performed with a Neural Network based forecasting tool. This tool has shown great benefits to the Fluxys gas business over the years. Its success lies in the way the data processing is managed. This paper demonstrates that reliable and good quality hourly forecasts can be achieved with some straight forward data manipulation techniques being applied to the forecaster.

NEURAL NETWORK FORECAST

Neural networks—or more accurately, Artificial Neural Networks (ANN), are a family of complex, non-linear data fitting methodologies. ANNs can approximate any function and are thus considered universal approximators. ANNs are very appealing in application domains where one has little or incomplete understanding of the relationship between the system parameters and their effect on the process. Where historic behavioural data is available Gas Load Forecasting is an ideal application for ANNs.

The Neural Network model is designed to forecast volume and energy targets at supply points along the pipeline network. The Neural network model finds patterns in complex pipeline operations data resulting from weather fluctuations, economics and calendar information. Based on past activity, the network utilizes a training process, where it develops and adapts parameters used in the algorithms, which are employed when predicting future demand. The model is also periodically "re-trained" to provide the most precise demand forecast possible for every meter on a daily or hourly basis. Changes to weather predictions or other factors generate a new set of forecasts.

Using this technology, Fluxys have accurate, updated projections to use in balancing and planning operations far more quickly than they have been previously able to achieve – and their pipeline staffs are better prepared for handling demand fluctuations.

MODEL OPTIMIZATION

The Gas Load Forecaster at Fluxys has the ability to perform routine evaluation of the available models to determine whether another available model is capable of attaining better training results than the one currently selected. If this is the case then, depending on the selection of associated options the software either advises the user of the situation or automatically selects the best model with a notification to the user.

Higher accuracy is achieved when an appropriate forecast model is chosen for a particular season. See table below:

Group	Model	Accuracy	Status
Group A	Summer	85%	Accepted
Group A	Peak Winter	96%	Accepted
Group A	Mild Winter*	97%	Accepted
Group A	Lag Effect**	98%	Selected

** Current Season – Mild Winter*
*** Mild Winter with load and weather lag effect*

The following optimization sequence proved a significant benefit to the system performance:

- For the most recent historic period use the actual weather data to perform a retrospective load forecast for each available model.
- Compare the actual loads for the forecast period to the retrospective load forecast from each of the available models and determine which one generated the most accurate forecast.
- Set the accuracy threshold level below which no advice on alternate models will be given and no automatic switching of models will take place.
- Set the number of days for which forecast is intended.
- Set the schedule for days of the week and hours of the day to run the optimization process.

Once the meter group priority and schedule is set, the forecast is generated automatically on the defined timings and priority sequence.

Whilst it seems obvious that using a relevant seasonal profile in the neural network model produces more accurate load forecast, the above model selection strategy provides a basis for selecting the appropriate season profile thus overcoming the difficulties encountered during e.g. warm spells in an otherwise severe winter. See Figure 1.

WEATHER EFFECTS COMBINATION

Weather data plays a key role in driving the desired forecast accuracy levels. In particular, accuracy of the weather forecast data needs to be determined to account for any deficiencies.

If the error in the weather forecast in real-time is not known then the ability to plan and manage risk associated with system imbalance can be very high. In the winter, the difference between a forecast uncertainty of 3-5 Deg C can be 10% of total load.

The analysis and magnitude of variations reflected in the forecast weather are shown in Figure 2:

Weather forecast reliability lowers load forecasting errors up to 50% for a selected weather dependant customer.

Besides weather data accuracies there are techniques that may be used to enhance the forecast accuracies which involve combining relevant weather effects.

Method1

In order to associate delayed weather effect with the gas consumption, custom weather effect (TAU) was applied in some neural network models. The formula used is:

TAU (t) = 0.03 * TMP (t) + 0.97 * TAU (t – 1)

Where TMP is the temperature effect at hour t

The above approach produced a positive impact on the accuracy in urban areas where buildings have thermal insulation.

Method2

The forecast with the upper limited product (WXT) of the wind speed (WSP) and the temperature (TMP) was also analyzed. The formula used is:

WXT (t) = Max (0, 15 – TMP (t)) * WSP (t)

The above correlation takes into account the effect of wind speed when the temperature drops below 15°C. See Figure 3 which demonstrates the effect of wind speed being ignored when the temperature is above certain reference level (15°C in this case). This approach produces more accurate forecasts during the periods when temperature variations are high within a day or week.

DATA CLEANING

The management of data sets is an important operation for many forecasters. Data should be looked at critically for errors and consistency. There are two main potential sources of degradation of data quality: the possible existence of statistically significant extreme deviations from the mean and the possibility of erroneous data recorded in the measurement process. Such irregularities in the data often result in predictions with a low accuracy. The use of data that is not reliable and verified is not recommended.

When used for prediction purposes gas load data should be evaluated and cleaned. The possible factors which influence the data quality are:

- Missing Values
- Metering Errors
- Offset Values
- Manual Errors
- Outliers

Fluxys uses Data Cleaning approach which provides the option to analyze the forecasting data beforehand to account for outliers, metering errors and missing values. The implemented data cleaning technique is also equipped with an online data processing module that will enhance the quality of online data as it is stored in the database. The original data remains recoverable for comparison. The outcome of the data cleaning practice enhances the quality of available data in order to achieve higher forecast accuracy. See Table 1 & 2 for detailed information of detection and emulation methods.

An Example:

Data Set – Public Distribution, 1 year period

Detection Method – 5 methods i.e., Absolute Range, Rate of Change, Dynamic range, Relative Range and Update Time

Cleaning Method – Using 2 weeks lead time, 1004 values

Result – Forecast accuracy increased to more than 95%. The results are shown in Figure 5 and 6.

WITHINDAY ADJUSTMENTS

Abrupt changes in gas consumption or non-availability of flow measurements produce a direct impact to the forecast accuracy if the neural network forecast model is trained using data not representative of such unplanned events. The situation is critical when the gas business cycle is running on daily basis and there is a requirement of critical load balance at the end of each gas day. In Belgium, there is a mindset the way dual fuel power plants are operated. The gas fired operation is now carried out for few hours within a day. If it takes place at the beginning of the day, without withinday adjustments, the forecast may be overestimated for the current gas day.

In such a situation a data post processing approach is required that takes into account the most recent flow measurements and modifies the forecast. The methodology involves calculating a moving average of the most recent errors and making adjustment in the forecast accordingly. See Table below.

Actual *a*	**Forecast** *b*	**% Error** *c=(a-b)/b*	**Adjusted Forecast** *d*
81.4	79.1	2.908 %	
79.3	76.8	3.225 %	
78.2	74.4	5.108 %	76.7
	80.9		83.4
	85.3		87.9
	71.9		74.1

Where d = b(1+Moving Average(c-1, c-2))*

The most recent flow measurement in this case may also be the same hour(s) actual load in the previous gas day(s).

The data post processing approach using a withinday procedure as explained in the table above was applied during peak winter season. This resulted in a reduction of up to 25% in forecast errors. See Figure 4.

CONCLUSIONS

By using the above approaches a number of conclusions can be drawn:

- Fluxys has installed a reliable on-line forecasting tool that has run continuously with minimum maintenance since the site acceptance test in April 2004.
- Choosing the best forecasting method for a particular type of data needs to be supported by proven pre- and post-processing methods of both measurement and weather data. This provides a consistent accuracy level.
- The system provides useful information to the operators when the automated routines correct the forecast accuracy level.
- Within the tool there is an option to be able to "optimize" the parameters that affect the forecast accuracy.

The correct use of each of these methods can lead to a consistent forecast accuracy of more than 95%. See Figure 7.

Overall this is an extremely interesting implementation that differs in many ways from the usual forecasting system. The main driver behind its success is the specific automation requirements from the system users which have been directly translated into the core product.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Jon Barley of Energy Solutions International for his valuable input and support which helped to strengthen the ideas.

ABOUT THE AUTHORS

Sophie Jehaes is an IT Project Leader in Fluxys BE. She holds an MSc degree in Civil Engineering from Université Libre de Bruxelles, Belgium (1994) and MBA from United Business Institute, Brussels, Belgium (2002). She started her professional career with BRC in 1994 where she worked as a Road Traffic Project Manager. In 2002, she joined Fluxys, the Belgian Gas Transport System Operator. She is responsible for projects related to Gas Demand Forecast, Shipper Servicing, Public Distribution Servicing and Standardisation of EDI messages.

Fakher Raza is a Senior Project Engineer in Energy Solutions UK. He holds a degree in Mechanical Engineering from the University of Engineering & Technology, Lahore, Pakistan (1994). He started his professional career with Sui Northern in 1994 where he worked as a Pipeline Flow Planning Engineer. After spending 5 years as a Senior Engineer providing technical support for pipeline simulation he joined Energy Solutions International in year 2001.

TABLES

TABLE 1: SUSPECT DATA DETECTION METHODS

Check Name	Accept Condition
Absolute Range	$L_1 \le v \le H_1$
Rate of Change	$L_2 \le v - v_o \le H_2$
Dynamic Range	$-s_\nu * n \le \lvert v - v_\nu \rvert \le n * s_\nu$
Relative Range	$L \le (v - v_r)/v_r \le H$
Update (value) (time tag)	$v <> v_{o,}\ t <> t_o$
Where;	
L	lower limit, (1 Absolute Range 2 Rate of Change)
H	upper limit, (1 Absolute Range 2 Rate of Change)
v	value to be checked
v_o	previous or next value
v_r	reference value
v_ν	average of values from the past time and future interval ν
t	time of reception of v
t_o	time of reception of v_o
s_ν	standard deviation from the past time and future interval ν
n	predefined number (integer)

TABLE 2: DATA EMULATION METHODS

Method Name	Emulated Value
Default value	v_d
Old value	v_o
Average value	v_ν
Forecast Value	v_f
Same hour/day	$v_s = \text{Average}(\ v_1*w_1,\ v_2*w_2, \ldots. v_n*w_n)$
External Value	V_e
Where;	
v_d	default value (constant)
v_o	previous value (i.e. one scan before)
v_ν	time average of measurement from the past time interval ν
v_s	Average value of the same hour or day in history
v_f	Forecast value of the same hour in future
V_e	Value of the same hour coming from another DB source
$v_1, v_2, \ldots. v_n$	Previous values
$w_1, w_2, \ldots. w_n$	Weights of the previous values based on calendar

FIGURES

FIGURE 1 – MODEL OPTIMIZATION

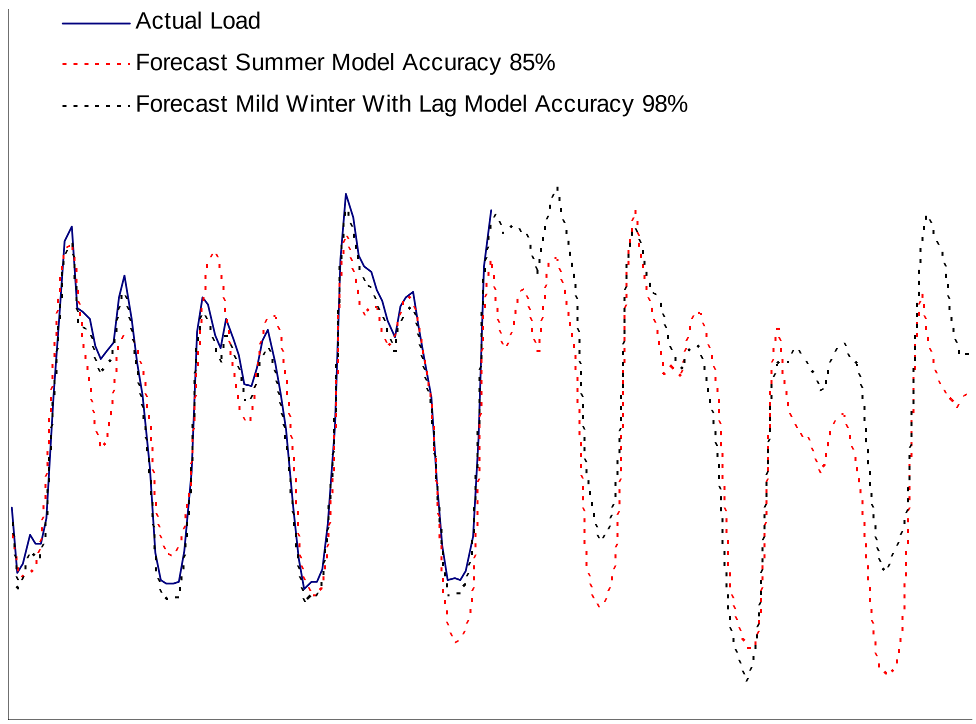

FIGURE 2 – WEATHER FORECAST ACCURACY

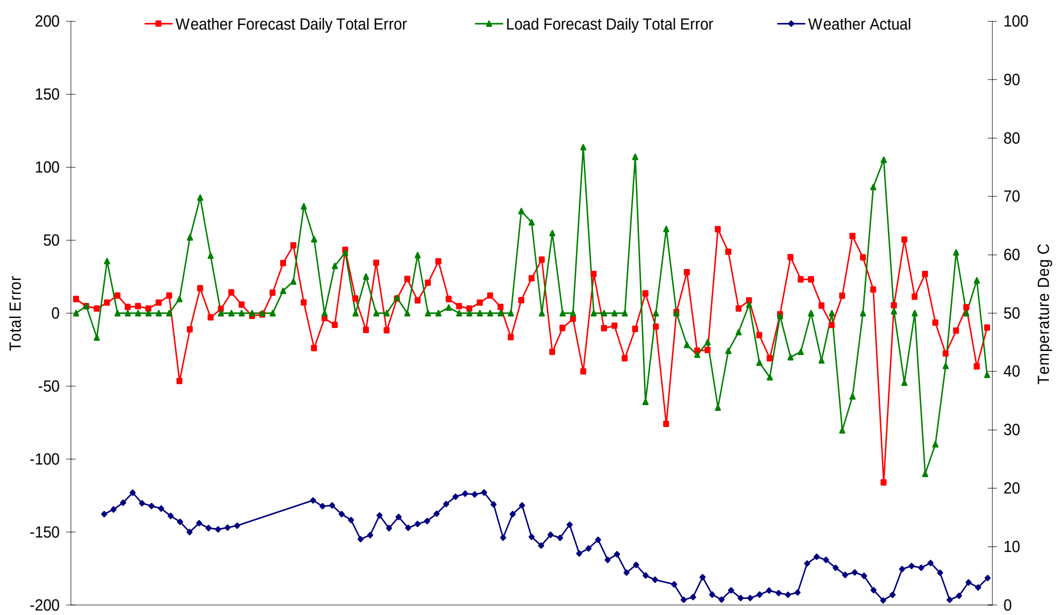

FIGURE 3 - WEATHER EFFECT, TEMPERATURE AND WIND SPEED

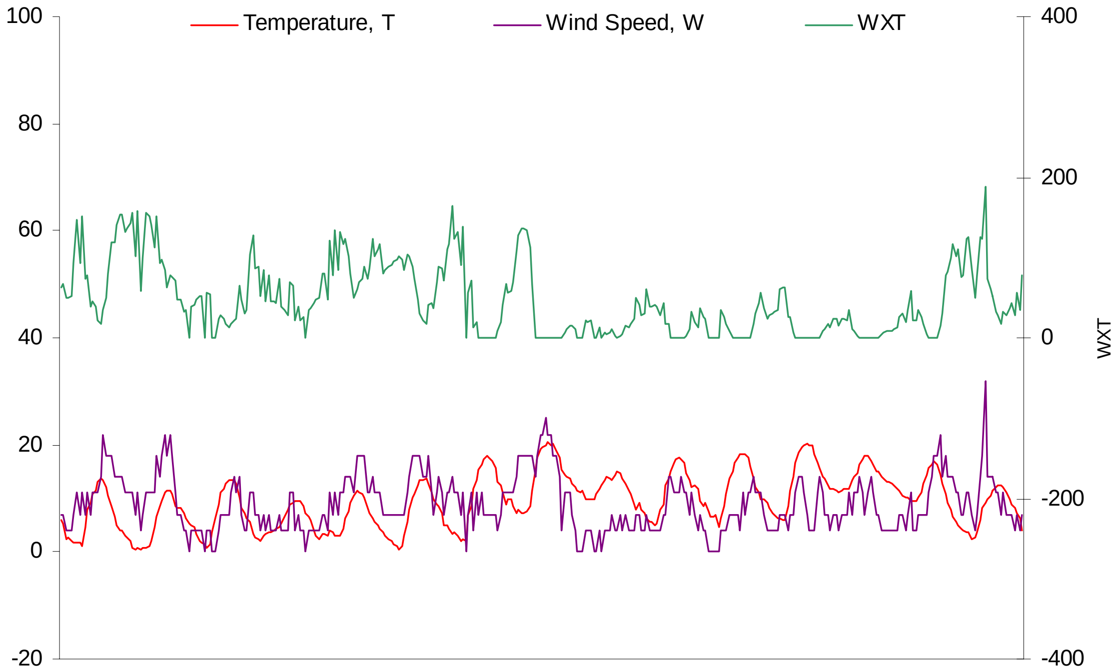

FIGURE 4 - WITHINDAY ADJUSTMENTS

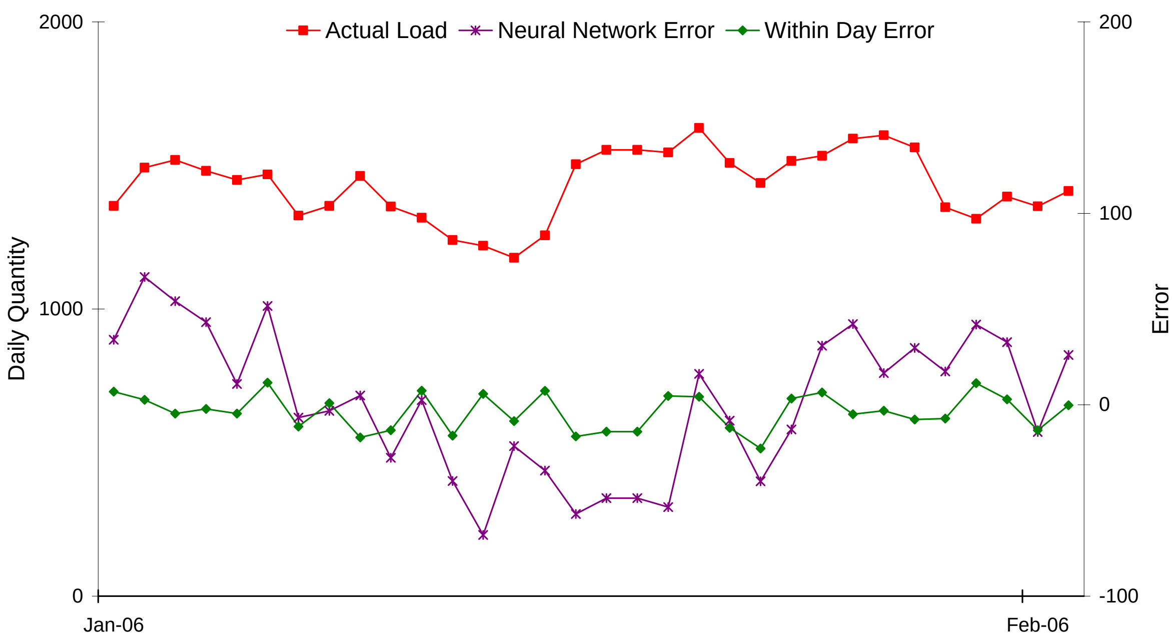

FIGURE 5 – DETECTED DATA

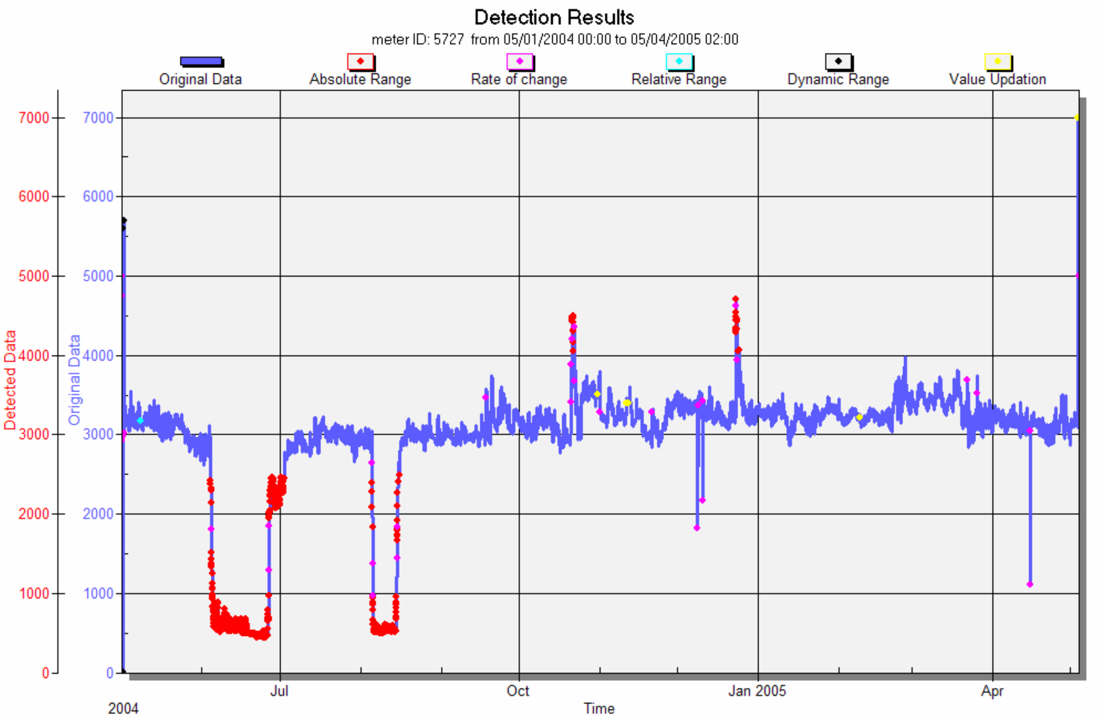

FIGURE 6 – CLEANED DATA

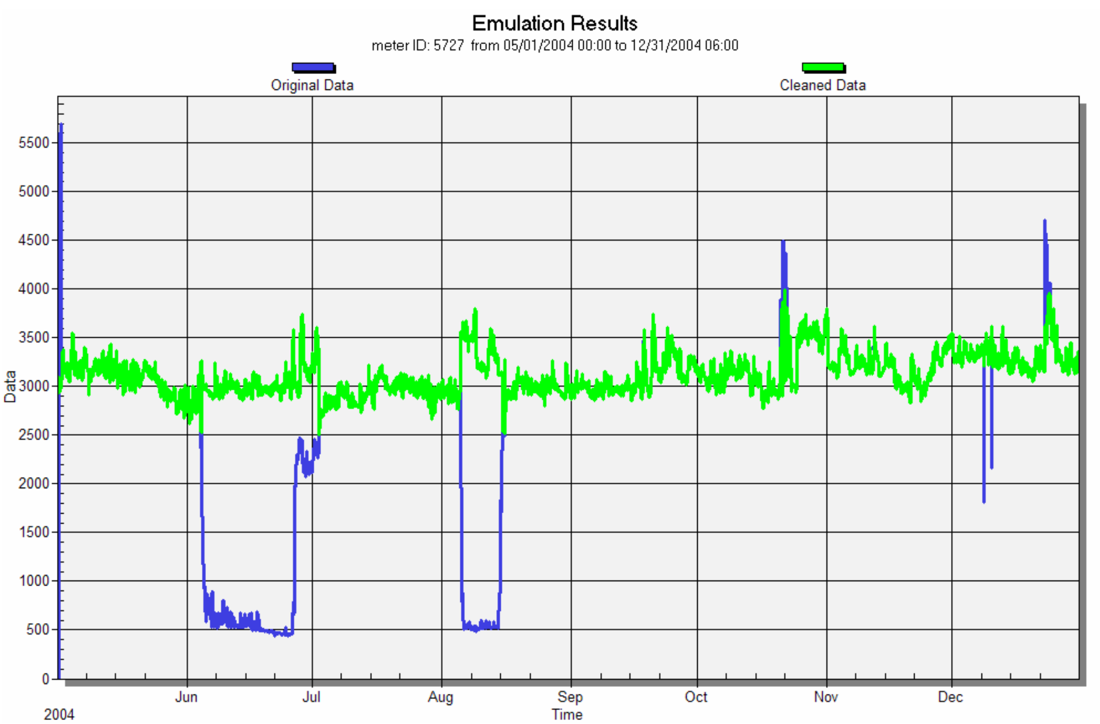

FIGURE 7 – FORECAST ACCURACY