Forecasting Energy Demand

MIKAN
27 rue de Sébastopol, 98800 Nouméa
Nouvelle Calédonie
www.mikan.nc

Summary

In this article we outline different groups of forecast types – short term, medium range, and long range – their characteristics and the potential benefit for companies operating in the energy market. Furthermore we introduce a list of requirements a modern forecast system should fulfil and present a few methods that are used for this purpose. In the last section we presented our forecast and modelling engine DECIDE.

Table of Content

1 Introduction

Over the past years the energy market in many countries has undergone dramatic changes. In countries like New Zealand, Australia, the US, Germany, Austria, and Switzerland (amongst many others) energy and especially electricity has become a tradable commodity. Prices change on a daily basis and companies active in the energy sector are obliged to develop business strategies to cope with a dynamic and volatile market.

The usefulness of energy demand forecasts is obvious in such an environment. Nowadays electricity producers and network operators use forecast systems on a regular basis. In the following we will outline on which time scales predictions are needed, what the benefits are, and which requirements a modern forecast system has to fulfil. In the last section we give a brief description of our forecast and modelling system DECIDE.

2 Time scales and benefits

In the energy sector predictions are needed on different time scales. These time scales can be divided into three groups:

- short term forecasts cover periods from some hours up to 7-10 days
- medium range predictions deal with time lags of a few weeks up to a couple of years
- long term forecasts go beyond one year

In the following sections we will characterise each of those groups, and discuss briefly the possible benefits of forecasts.

2.1 Short term

2.1.1 Type of forecast

Short term forecasts cover periods of a few hours up to about 7-10 days. In the energy sector such forecasts are usually based on meteorological data (temperature, humidity, wind speed, cloud coverage, …), on characteristic diurnal and weekly cycles, public holidays, school holidays, and the current (today/now) energy consumption. In principle, however, many other factors may be relevant. Examples for such factors are socio-economic indices or big cultural and sport events.

Short term forecasts are designed to predict as precisely as possible individual events, i.e. the actual energy demand/consumption on a rather fine time pattern and to forecast unusual energy demands due to special weather patterns or other unusual events.

2.1.2 Benefits

Short term forecasts have become the principal tool for scheduled energy interchange. As an example we give a brief outline of the system established in Germany. To obtain an optimised net schedule every electricity producer, electricity distributor or network operator has to announce its demand of energy. Each day before 10:00 am the needs for the following day have to be announced on a 15 minute time pattern. The network management is based on those announcements. The differences between the actual needs and the announced needs is then assessed later on. If this difference is below a certain threshold – usually 5% - one gets a price reduction of the order of 0.3% to 0.5%.

A second application for short term forecast is the energy exchange (for instance in Leipzig/Germany). This energy exchange is used to trade balance energy on a day to day basis. Again short term forecast systems are the principal tools for every market player to optimised their trading strategies and to minimise the need of purchasing relatively expensive balance energy.

In many countries contracts between electricity producers and distributors or network operators still take into account the peak load (power) measured over one year. Energy distributors usually have to pay a fee which is proportional to the height of the peak. In order to get more favourable contracts, network operators therefore try to cut those peaks. Again, short term predictions are the main instrument for this sort of energy management.

2.2 Medium range

2.2.1 Type of forecast

Medium range forecasts cover periods from a few months up to a few years. For energy predictions relationships between large scale climate variability and/or socio-economic factors are exploited. A quantitative model that relates climate phenomena and socio-economics on one side and energy demand on the other side is needed for that purpose. This model is then used to transform medium range climate forecasts and medium range socio-economic forecasts into energy predictions.

Socio-economic forecasts have been available for a long time. The government or research institutes in basically each country provide such outlooks and their interpretation and use is relatively straight forward.

Medium range climate forecasts have become operational or quasi operational recently. Leading climate research centres and meteorological services offer such climate outlooks. Their use is more complicated - compared to economic outlooks – and requires profound expertise of data handling and preparation.

In contrast to short term forecasts medium range energy forecasts do not aim at predicting individual events. The goal is to forecast changes in the probability distribution of certain events.

An example will clarify this point: An anomalous warm winter in Europe will certainly lower the demand of heating energy. A long range climate forecast for Europe, however, will not predict precisely the temperature for each day between beginning of December and end of February. Instead it may for instance predict a probability of 40% that the average winter temperatures are more than 2 degrees above normal. In order to use this information it is necessary to transform it into a (biased) probability for energy demand. A decision maker in the energy industry, however, is not necessarily interested in climate variability itself. He needs to know how a 40% probability of having a mild winter transforms into a probability distribution for energy demand.

2.2.2 Benefits

As already pointed out, medium range forecasts play on probabilities, not on actual events. Hence, they are mainly used for risk management and risk mitigation. Companies active in the energy market may protect themselves against unusual economic or climatic events on the capital market. For climate related risks weather derivatives are the proper instrument. The first weather contract was signed in 1997 in the US. Since then an enormous growth of the weather derivative market has been observed world wide, which reflects the importance that decision makers nowadays attribute to the possible impact of weather phenomena onto business activities.

In order to use such financial instruments energy producers and distributors need to quantify the impact of weather variability onto their business and they need to predict the possible shift of probability distributions.

2.3 Long term

2.3.1 Type of forecast

Long range climate forecasts typically cover periods of several years to several decades. Those forecasts deal with slowly evolving phenomena on large scales. Such phenomena modify the mean reference state for a certain business. Examples are long term trends, evolution of the population, long term political evolutions and climate change.

2.3.2 Benefits

Long term forecasts are important to adapt and develop long term business strategies. If general conditions change risks and therefore possible losses increase. To secure their own position companies must anticipate long term changes and adapt their strategies accordingly.

3 A comprehensive forecast system

A modern forecasting system for the energy industry has to fulfil a large range of requirements. First of all it has to be easy to use. A forecast system must not require special knowledge in mathematics or statistics from the user. Beside that a forecasting environment should be flexible and produce robust results on relative short training periods. Since data processing is often a difficult and time consuming process a data processing module should be included, and finally meaningful reports should be created automatically.

In the following we will address each of those points.

3.1 Self-learning

Finding a quantitative model to describe the impact of a set of independent so-called predictor variables on the target variables of interest (e.g. energy consumption) is usually a difficult and lengthy process. Nowadays self-learning mathematical methods are available that have largely facilitated this process. Such methods find the proper model automatically from a set of example data. The example data are usually called training data set or training period. The model, derived from the training data is stored and can then be used to compute from a new set of predictor records the new target variables of interest.

As an example we outline the process to find a model for the consumption of electricity each day at 10:00 am as a function of mean temperature, minimum temperature, and maximum temperature of the same day. As training data we use the three temperature values and the corresponding electricity demand for the last 200 days. Those training data allow the self learning method to find the quantitative relationship between the predictors (temperatures) and the target (energy consumption). Once the model training has finished the final model is stored for later use. To predict the electricity consumption for tomorrow the stored model is used to compute the consumption from the forecasted temperature values for the next days. The temperature forecasts can be obtained from a meteorological service.

Real life, however, is more complicated. Usually energy demand has to be predicted on a much finer time grid (typically 15 minutes for electricity and 60 minutes for gas) and the predictor variables do not only include temperatures but also other meteorological data like cloud coverage, wind speed, and humidity, possibly at different locations. Furthermore information about the day of the week, public holidays, and school holidays certainly play an important role. A powerful model must be able to cope with a relatively large number of predictor and target variables at the same time.

Different self learning methods exist with different degrees of complexity. They basically can be divided into two classes, parametric and non-parametric methods. Examples for parametric methods are uni- and multivariate linear regressions and non-linear (parametric) regressions. Examples for non-parametric methods are neural networks, multivariate adaptive regression splines, and support vector machines. Non-parametric methods are in general preferable since they do not assume an a-priori set functional form for the underlying relation between predictors and targets.

3.2 Flexibility

Each forecast system must be adapted to individual needs. Every forecasting problem is at least slightly different, and therefore the optimal choice of parameters to be used or the best data treatment to be applied may differ. To find the setup that yields the best results tests with different input variables and with different data treatments have to be carried out.

To facilitate those tests the forecast system has to be as flexible as possible. It must be easy to add or remove predictor and target data, to use different data treatments, and to check the results of each test quickly. An important issue in this context is the computational time needed for each test. Some methods yield excellent results but require an enormous amount of computational time. If a single test takes several hours it is impossible to run a large number of tests within an acceptable time frame.

3.3 Robustness

Model results should be robust. The structure of the model, and hence the forecasting results must not be oversensitive to the choice of tuning parameters or the period used for the training. Methods that show great sensitivity to such choices may yield excellent results in some cases. To maintain a high level of accuracy over different periods such methods require a large amount of maintenance and a lot of user-expertise.

3.4 Short training periods

Self learning models need example data – the so called training data. To be used in daily business a modern forecast system should be able to cope with relatively small training data sets.

In nowadays volatile energy markets the customer structure is no longer static in time. The customer base changes typically at the end of each month, with bigger changes at the end of the year. If the models need long training periods, it is necessarily trained for a relatively long time interval with a customer base which is no longer up to date. Companies compensate these effects by applying manual corrections either to the forecasts or to the data used to train the model. Those corrections are time consuming, and since they have to be carried out manually they are error prone and expensive.

Methods that can cope with rather short training periods – of the order of 30 to 150 days – can be re-trained easily after major changes of the customer base. Thus the model is adjusted by simply re-running the training cycle with up to date training data. Such a system needs much less data maintenance, reacts quickly and automatically to changes and hence is much easier to use in day to day business.

3.5 Data Processing

Data processing is an important step in setting up a forecast system. Raw data have to be prepared before using them in the actual modelling process. Basic data pre-processing consists of removal of a linear trend, subtraction of mean values and periodic cycles, and normalisation of the data (variance = 1).

More advanced pre-processing methods can be summarised as “filtering”. Filtering is used to damp certain frequencies in the data or to concentrate on phenomena with a pre-selected period range.

Pattern recognition methods can be used to separate unpredictable noise from predictable signals. Another very important application of pattern recognition is data compression. In high dimensional problems very long training periods are needed. This problem can be overcome by reducing the dimension of the problem using data compression methods. This point is especially important for short term energy demand forecasts. For this type of forecasts the number of predictors (i.e. the dimension) easily can reach several hundreds. To obtain a stable and robust model using a short training period it is indispensable to reduce the dimension significantly. Another feature of pattern recognition methods is that they can be used to remove colinearities from the predictor variables, which pose problems to most self-learning algorithms.

3.6 Reports

The final goal of each forecasting and analysis system is to provide easily understandable results. Reports and graphics have to be generated automatically and it must be possible to exchange results easily with colleagues. Graphics must be self-explanatory and reports must use “plain” language, adapted to the user’s background. A strong background in statistics or mathematics must not be anticipated.

It further should be possible to integrate the forecasting system into an existing working environment. Thus the user gets just additional functionality in an environment he is already familiar with.

4 DECIDE: A comprehensive modelling and prediction engine

To fulfil the above requirements we have developed DECIDE, a comprehensive modelling and prediction system. It consists of three main modules: A data pre-processing module, a pattern recognition tool and the actual modelling/forecasting part.

In its current configuration DECIDE can handle up to 2500 different predictor and target variables and a maximum of 2500 cases (a case corresponds either to a time steps or an example). DECIDE uses computationally extremely efficient algorithms and is for many applications more than a hundred times faster than conventional methods (neural networks). DECIDE comprises of a complete data pre-processing module, a powerful pattern recognition tool and a very efficient non-parametric modelling module. In the following a brief description of each module is given.

Currently the DECIDE modelling engine is implemented into mlRAP of metalogic, a software company in Germany. mlRAP is a complete prediction and analysis system that is especially tailored for the needs of energy supplies and distributors.

HAKOM is currently integrating DECIDE into their software package. Due to its speed, flexibility and ease to use they decide to replace their own in-house forecasting module by our system. HAKOM is located in Vienna and is one of the leaders in Austria in the field of Energy Data Management systems.

4.1 Data pre-processing

Included into DECIDE is a complete pre-processing module. The following processing steps can be performed:

- Detection and removal of outliers. Each data point is compared with its neighbours. If the difference is beyond a certain threshold (to be specified by the user) the corresponding data point is rejected. This part acts as a final quality control.
- Linear trend removal: A linear trend can be computed and removed from each data record. A robust regression method is used, that is not sensitive to extreme outliers. Trend removal is required for most statistical modelling methods.
- Subtraction of the mean is needed for pattern recognition
- Subtraction of a period cycle: Often signals of interest are overwhelmed by a simple periodic cycle. A weekly cycle, for instance, may be so strong, that week to week changes - in which the user is actually interested in - are hidden. To investigate and model the week to week signal the weekly cycle has to be removed from the data. The model searches and removes automatically for periodic cycles. The period to look for has to be specified by the user.
- Normalisation: To compare different data types to each other they have to be normalised such that they all have the same order of magnitude. This is usually done by dividing them by their respective variance. This is especially important if pattern recognition is applied to combined data sets.

4.2 Pattern recognition

Pattern recognition is an extremely powerful tool for data compression and noise reduction. It is extensively used in fields like biology, climate research, and psychology, amongst others.

In DECIDE we have implemented a very efficient pattern recognition method. With this module the user can handle high dimensional data sets in combination with short training periods, which otherwise would not be possible. Furthermore, calculations can be sped up in some cases by a factor of 50 to 100 – depending on the size of the problem. To our knowledge DECIDE is the only forecasting system for energy demand that provides a pattern recognition module.

4.3 Modelling/Forecasting

DECIDE has implemented a non-linear, non-parametric modelling algorithm called MARS. As far as forecast precision goes this method yields comparable results with artificial neural networks but at much lower computational costs. MARS can deal with much shorter training periods, it is far more robust, and it is computationally much more efficient. Depending on the application MARS is up to 300 times faster than neural networks. Last but not least MARS modelling results can be easily interpreted, which provides valuable insight into the functioning of the problem of interest.

5 Conclusions and outlook

In this article we have outlined different groups of forecast types – short term, medium range, and long range – their characteristics and the potential benefit for companies operating in the energy market. Furthermore we presented a list of requirements a modern forecast system should fulfil. In the last section we presented our forecast and modelling system DECIDE.

So far we did not address the topics risk management and a monetary assessment of a forecast system. These points are especially important for medium range forecast, where one has to deal with predicted probabilities instead of prediction of events. Future developments will certainly concentrate on these points. Tools to automatically develop optimal business strategies, taking into account the occurrence rate of a certain event and the skill of the actual forecast system will become an integrated part of a prediction system. Other tools may address the pricing of weather derivatives in real time.

MIKAN is actively doing research in this field in collaboration with renowned international research institutes. We are member of the WMO expert team on end-user liaison, which has the mission to provide the often missing link between academic research and industrial applications.