# Predicting short-term wholesale prices on the Irish Single Electricity Market with Artificial Neural Networks

Pengfei Li, Francesco Arci, Jane Reilly
Ark Energy Consulting Limited
Unit 20 Daingean Hall, N4 Axis Centre,
Battery Road, Longford.
{pengfei.li, francesco.arci, jane.reilly}@ark-energy.eu

Kevin Curran, Ammar Belatreche, Yauheniya Shynkevich
Faculty of Computing and Engineering
Ulster University
Northern Ireland
{kj.curran, a.belatreche, Shynkevich-y@ulster.ac.uk

***Abstract:*** *The ability to operate effectively on electricity spot markets relies on the capability to devise appropriate bidding strategies. These in turn can benefit from the inclusion of a reliable forecast of short term system marginal prices (SMPs). In a market with an increasing percentage of renewable generators, reliable forecasts must necessarily consider additional factors such as meteorological forecasts, forecasted demand and constraints imposed by network topology. Traditional time series forecasting algorithms (e.g. based on AutoRegressive Integrated Moving Average models) can perform reasonably well in this context but rely on assumptions being made on behavior over different temporal windows to yield consistent results. Research studies have demonstrated that an adaptive or self-adaptive approach to forecasting would appear more suited to the task of predicting energy demands in territory such as Ireland. Implementing an in-house self-adaptive model should yield good results in the dynamic uncertain Irish energy market. We have identified the features that such a model demands and outline it here.*

## I. INTRODUCTION

Electricity markets are different from other markets as electricity generation cannot be easily stored in large amounts and to avoid blackouts, the generation of electricity must be balanced with customer demand for it on a second-by-second basis. Customers tend to rely on electricity for day-to-day living and cannot replace it easily so when electricity prices increase, customer demand generally does not reduce significantly in the short-term. As electricity generation and customer demand must be matched perfectly second-by-second, and because generation cannot be stored to a large extent, cost bids from generators must be balanced with demand estimates in advance of real-time. The increasing percentage of electricity generated through renewable sources tends to invalidate the assumption of correlation between electricity spot prices and the price of the mix of commodities utilized to supply generators (e.g. gas, coal, oil – depending on the generating asset composition on the specific grid). The variable nature of production of renewable energy sources also increases the volatility of system marginal prices (SMPs) on markets based on a mandatory central pool model. European countries have undertaken substantial investments to boost the amount of energy produced through renewable generation. Ireland in particular is aiming at 40% of its power needs being met by renewable sources by 2020. In this environment, we can expect the wholesale, fine granularity (e.g. half hourly) wholesale price of electricity to become more volatile over time. Previous work (Sharma et al., 2012) have shown that an adaptive or self-adaptive approach to forecasting would appear more suited to the task of predicting energy demands in territory such as Ireland.

The Single Electricity Market (SEM) is the wholesale electricity market for the island of Ireland, regulated jointly by the CER and its counterpart in Belfast, the Utility Regulator. The Commission for Energy Regulation (CER) is the independent body responsible for regulating the natural gas and electricity markets in Ireland. By combining what were two separate jurisdictional electricity markets, the SEM became one of the first of its kind in Europe when it went live on 1st November 2007 (CER, 2011). The SEM is designed to provide for the least cost source of electricity generation to meet customer demand at any one time across the island, while also maximising long-term sustainability and reliability. The SEM is operated by SEMO, the Single Electricity Market Operator, a joint-venture between EirGrid and SONI, the transmission system operators in Ireland and Northern Ireland respectively. SEMO[1] is responsible for administering the market, including paying generators for their electricity generated and invoicing suppliers for the electricity they have bought (CER, 2011). SEM consists of a centralised and mandatory all-island wholesale pool (or spot) market, through which generators and suppliers trade electricity. Generators bid into this pool their own short-run costs for each half hour of the following day, which is

[1] www.sem-o.com.

mostly their fuel-related operating costs. Based on this set of generator costs and customer demand for electricity, the System Marginal Price (SMP) for each half-hour trading period is determined by SEMO, using a stack of the cheapest all-island generator cost bids necessary to meet all-island demand. It is these more efficient generators which are generally run to meet demand in the half hour in what is known as the "Market Schedule". More expensive or inefficient generators are "out of merit" and hence they are not run and are not paid SMP, keeping customers' bills down. The SMP for each half hour is paid to all generators that are needed to meet demand. Suppliers, who sell electricity direct to the final consumer, buy their electricity from the pool at this common price. Overall the SEM facilitates the running of the cheapest possible generators, determined by the stack of generation cost bids, to meet customer demand across the island. This mandatory centralised pool model in SEM, in which all key generators and suppliers must participate, differs from most other European markets in which most trade takes place bilaterally between generators and suppliers. In these bilateral markets only a residual amount of electricity is traded in an exchange, primarily for balancing purposes. In contrast all key players must trade in SEM, so there is more transparency associated with SEM prices and market outcomes (CER, 2011). This paper outlines several forecasting algorithms to predict short-term (72 hours ahead) wholesale prices on the Irish Single Electricity Market so that market participants can make more informed trading decisions. This paper is organised as follows: Section II introduces Artificial Neural Networks & Short-term Load Forecasting, section III presents the short-term forecasting model and section IV provides a conclusion.

## II. ARTIFICIAL NEURAL NETWORKS & SHORT-TERM LOAD FORECASTING

In machine learning and cognitive science, artificial neural networks (ANNs) are a family of models inspired by biological neural networks (the central nervous systems of animals, in particular the brain) and are used to estimate or approximate functions that can depend on a large number of inputs and are generally unknown. (McCulloch & Pitts, 1943). Like other machine learning methods – systems that learn from data – neural networks have been used to solve a wide variety of tasks that are hard to solve using ordinary rule-based programming, including computer vision and speech recognition (Forouzanfar, 2010). For short-term load forecasting, the Back Propagation Network (BP) network is the most widely used one. Due to its ability to approximate any continuous nonlinear function, the BP network has extraordinary mapping (forecasting) abilities. The BP network is a kind of multilayer feed forward network, and the transfer function within the network is usually a nonlinear function such as the Sigmoid function. The typical BP network structure for short-term load forecasting is a three-layer network, with the nonlinear Sigmoid function as the transfer function (Schmidhuber, 2015). Fully connected BP networks need more training time and are not adaptive enough to temperature changes therefore some have moved to using non-fully connected BP models (Graves et al., 2009). Although a fully connected ANN is able to capture the load characteristics, a non-fully connected ANN is more adaptive to respond to temperature changes. Results also show that the forecasting accuracy is significantly improved for abrupt temperature changing days. There is also merit in combining several sub-ANNs together to give better forecasting results such as using recurrent high order neural networks (RHONN). Due to its dynamic nature, the RHONN forecasting model is able to adapt quickly to changing conditions such as important load variations or changes of the daily load pattern (Graves et al., 2009). A back propagation network is a type of array which can realize nonlinear mapping from the inputs to the outputs. Therefore, the selection of input variables of a load forecasting network is very important. In general, there are two selection methods. One is based on experience and the other is based on statistical analysis such as the ARIMA and correlation analysis.

For instance, we can denote the load at hour k as l(k) so a typical selection of inputs based on operation experience will be l(k-1), l(k-24), t(k-1), where t(k) is the temperature corresponding to the load l(k). Unlike those methods which are based on experience, we can apply auto-correlation analysis on the historical load data to determine the input variables. Auto-correlation analysis should show that correlation of peaks occurs at the multiples of 24 hour lags. This indicates that the loads at the same hours have very strong correlation with each other. Therefore, they can be chosen as input variables. In addition to using conventional information such as historical loads and temperature as input variables, wind-speed, sky-cover can also be used. Potential input variables could be historical loads, historical and future temperatures, hour of day index, day of week index, wind-speed, sky-cover, rainfall and wet or dry days. There are no hard fast rules to be followed to determine input variables. This largely depends on engineering judgment and experience. Previous research (Rui & El-Keib, 1995) has found that for a normal climate area, historical loads, historical & future temperatures, hour of day and day of week index are sufficient to give acceptable forecasting results. However, for an extreme weather-conditioned area the other input variables classes were recommended, because of the highly nonlinear relationship between the loads and the weather conditions.

## III. A SHORT TERM ENERGY FORECASTING MODEL

Artificial Neural Networks (ANNs) can only perform what they were trained to do. Therefore to achieve short term load forecasting, the selection of the training data is a crucial one. The criteria for selecting the training set is that the characteristics of all the training pairs in the training set must be similar to those of the day to be forecasted. Choosing as many training pairs as possible is not the correct approach for a number of reasons. On reason is load periodicity. For instance, each day of the week has different patterns. Therefore, using Sundays' load data to train the network which is to be used to forecast Mondays' loads would lead to wrong results. Also, as loads possess different trends in different periods, recent data is more useful than old data. Therefore, a very large training set which includes old data is less useful to track the most recent trends. Therefore to obtain good forecasting results, day type information must be taken into account. We can achieve this by constructing different ANNs for each day type, and feeding each ANN the corresponding day type training sets ( Lee, 1992). Another way is to use only one ANN but contain the day type information in the input variables (Srinivasan, 1994). The two methods have their advantages and disadvantages. The former uses a number of relatively small size networks, while the latter has only one network of a relatively large size. The day type classification is system dependent e.g. the load on Monday may be similar to that on Tuesdays but not always. Therefore one option is to classify historical loads into classes such as Monday, Tuesday-Thursday, Friday, Saturday, and Sunday/Public holiday. The Back Propagation algorithm is widely used in short-term load forecasting and has some good features such as, its ability to easily accommodate weather variables, and its implicit expressions relating inputs and outputs but it is also a time consuming training process and its convergence to local minima (Ciresan et al., 2012). The determination of the optimal number of hidden neurons is a crucial issue. If it is too small, the network cannot possess sufficient information, and therefore yields inaccurate forecasting results. On the other hand, if it is too large, the training process will be very long (Balabin et al., 2009).

Other important factors are to determine how big the prediction window should be. For instance, it could possibly be cold in one month so is this valid 12 months later. The forecast horizon is day + 1 - and for remainder of day. This is for the next available market. The model may also provide predictions for 48/72 hours. This will lead of course to dimensioned results but we associate a corresponding error value. Not all electricity markets follow the same slots so in practice we aim to weather forecast, model network topology and more. Some of the main factors for forecasting are demand forecast, estimated power production capability and available interconnection capacity. Outliers include weather events, solar eclipses so we must also be careful not to factor into our model. The initial stage involves determining the input variables from the demand, power production and price prediction data we download from SEMO[2]. The notation used throughout the paper is provided in Figure 1.

[2] http://www.sem-o.com

| Notation | Meaning |
|---|---|
| **D** | Report date, such as 11/25/16 |
| **D+2** | The delivery date of predicted SMP, such as 11/27/16 (7:00am – 6:30am$^{+1}$) |
| **$SMP_{D+2hh}$** | The output (7:00am – 6:30am$^{+1}$) |
| **$Demand_{D+2hh}$** | The Demand corresponding to the output (7:00am – 6:30am$^{+1}$) |
| **$Power\_Ireland_{D+2hh}$** | The power summation of Solar power and wind power production in the whole Ireland (7:00am – 6:30am$^{+1}$) |
| **$Power\_UK_{D+2hh}$** | The power summation of Solar power and wind power production in the whole UK mainland (7:00am – 6:30am$^{+1}$) |
| **$SMP_{D+1hh}$** | The SMP tomorrow (7:00am – 6:30am$^{+1}$) |
| **$SMP_{D-5hh}$** | The week-ahead SMP of the predicted date |
| **$SMP_{D-12hh}$** | The 2-week ahead SMP of the predicted date |
| **$SMP_{D+1hh-1}$** | The SMP of previous half hour |
| **$SMP_{D+1hh-2}$** | The SMP of previous hour |

Figure 1: Nomenclature used

Feature selection is the process of selecting a subset of relevant features for use in model construction. Feature Selection is placed into two main categories, wrapper methods and filter method. Wrapper methods evaluate multiple features using procedures that add and/or remove predictors to find the optimal combination that maximizes model performance. We use Recursive Feature Elimination with Backwards Selection in our feature selection model and use Random Forecast Method as the forecasting algorithm. An obvious concern is that too few variables are selected or that the selected set of input variables is not sufficiently informative. Half-hourly SMP itself can be divided between the shadow price and uplift price. The SMP follows customer demand, as a more expensive stack of generators is needed to meet demand when it is high, whereas at low demand times demand can be met with cheaper generators. Approximately 80% of the island's electricity generation comes from imported fossil fuels, with most this in the form of gas-fired generation plants, though the amount of renewable generation (especially wind) is increasing. The start date of training date was 20-11-2016 and the last date of training date was 20-1-2017. The preprocessing included normalization, separation of input and output, removal of the column with near zero variance and removal of the column with high correlation. The inputs were ["Delivery_Date", "Delivery_Hour", "Delivery_Interval", "SMP_D_Euro", "SMP_D_Minus_6_Euro", "SMP_D_Minus_13_Euro", "LoadDemand", "Power_Production_Ireland", "Output_SMP_Euro"]. The resampling method is cv (cross validation), the number of divided blocks is 9. The WM method tuning Grid of num.label is 5,7,9,11.

| SMP D Euro | SMP D-1 Euro | SMP D-2 Euro | SMP D-3 Euro | SMP D-4 Euro | SMP D-5 Euro | SMP D-6 Euro | SMP D-13 Euro | SMP HH-1 Euro | SMP HH-2 Euro | Load Demand | Power Prod Ireland | Power Prod UK | Output SMP Euro |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 34.11 | 56.12 | 35.58 | 35.45 | 35.45 | 33.85 | 36.02 | 37.27 | 38.56 | 40.22 | 3332.99 | 2632 | 6432 | 26.82 |
| 34.96 | 53.31 | 34.96 | 36.22 | 35.67 | 33.85 | 39.60 | 37.22 | 34.11 | 38.56 | 3617.22 | 2632 | 6432 | 33.29 |
| 37.35 | 52.05 | 35.93 | 36.22 | 37.65 | 33.93 | 48.55 | 40.12 | 34.96 | 34.11 | 4044.04 | 2622 | 6301 | 33.37 |
| 46.83 | 48.49 | 36.95 | 45.34 | 49.53 | 41.99 | 58.69 | 48.23 | 37.35 | 34.96 | 4598.26 | 2622 | 6301 | 44.12 |
| 53.00 | 45.41 | 39.11 | 58.81 | 54.65 | 50.07 | 49.99 | 52.00 | 46.83 | 37.35 | 4794.32 | 2588 | 6213 | 36.26 |
| 53.00 | 42.83 | 45.16 | 59.50 | 54.65 | 50.07 | 48.91 | 53.24 | 53.00 | 46.83 | 4848.44 | 2588 | 6213 | 36.26 |

Figure 2: Training data set

| Wang and Mendel Fuzzy Inference System | | | Wang and Mendel Fuzzy Rules | | |
|---|---|---|---|---|---|
| Num labels | RMSE | RSquared | Num Labels | RMSE | RSquared |
| 5 | 0.08085976391 | 0.6164148104 | 5 | 0.08243602951 | 0.4944017323 |
| 7 | 0.08348111341 | 0.5985532171 | 7 | 0.08034681858 | 0.5329743329 |
| 9 | 0.08282707433 | 0.6045367775 | 9 | 0.06520352477 | 0.5802984609 |
| 11 | 0.08351732060 | 0.6031938904 | 11 | 0.06158859213 | 0.6117265554 |
| 13 | 0.08297830444 | 0.6087738091 | 13 | 0.06288671368 | 0.5995754154 |
| 15 | 0.08141637713 | 0.6133514129 | 15 | 0.06096258818 | 0.6064996381 |

Figure 3: WM Methods

| Neural Network | | | | Neural Network with Feature Extraction | | | |
|---|---|---|---|---|---|---|---|
| Size | Decay | RMSE | RSquared | Size | Decay | RMSE | RSquared |
| 7 | 0.1 | 0.05270524894 | 0.6929778933 | 7 | 0.1 | 0.05089153350 | 0.7208933843 |
| 7 | 0.2 | 0.05347311140 | 0.6888295076 | 7 | 0.2 | 0.05167995561 | 0.7140019128 |
| 7 | 0.3 | 0.05453963414 | 0.6838997251 | 7 | 0.3 | 0.05250239843 | 0.7054384820 |
| 7 | 0.4 | 0.05573614668 | 0.6818432270 | 7 | 0.4 | 0.05310788152 | 0.6991367997 |
| 7 | 0.5 | 0.05697179184 | 0.6792381810 | 7 | 0.5 | 0.05352101625 | 0.6939291592 |
| 9 | 0.1 | 0.05271452634 | 0.6927232380 | 9 | 0.1 | 0.05119465918 | 0.7175199801 |
| 9 | 0.2 | 0.05344611869 | 0.6889050040 | 9 | 0.2 | 0.05164156293 | 0.7146256683 |
| 9 | 0.3 | 0.05437667388 | 0.6856260338 | 9 | 0.3 | 0.05240655299 | 0.7054188336 |
| 9 | 0.4 | 0.05541222294 | 0.6832885298 | 9 | 0.4 | 0.05300188523 | 0.6998314790 |
| 9 | 0.5 | 0.05666255577 | 0.6808368771 | 9 | 0.5 | 0.05348630999 | 0.6950229190 |
| 11 | 0.1 | 0.05270600467 | 0.6927112655 | 11 | 0.1 | 0.05163404149 | 0.7143136221 |
| 11 | 0.2 | 0.05338348384 | 0.6890494993 | 11 | 0.2 | 0.05159720638 | 0.7150221137 |
| 11 | 0.3 | 0.05428912207 | 0.6860951090 | 11 | 0.3 | 0.05238156971 | 0.7062859365 |
| 11 | 0.4 | 0.05526530966 | 0.6838981978 | 11 | 0.4 | 0.05307634772 | 0.6988441224 |
| 11 | 0.5 | 0.05649060682 | 0.6814349601 | 11 | 0.5 | 0.05339202900 | 0.6959431786 |
| 13 | 0.1 | 0.05267401075 | 0.6930661064 | 13 | 0.1 | 0.05111346554 | 0.7186556348 |
| 13 | 0.2 | 0.05332159021 | 0.6898724679 | 13 | 0.2 | 0.05154455074 | 0.7152574599 |
| 13 | 0.3 | 0.05414921928 | 0.6870251648 | 13 | 0.3 | 0.05250977982 | 0.7055231884 |
| 13 | 0.4 | 0.05518955461 | 0.6841764190 | 13 | 0.4 | 0.05300783963 | 0.6996952093 |
| 13 | 0.5 | 0.05629418211 | 0.6822628708 | 13 | 0.5 | 0.05341948835 | 0.6956999722 |
| 15 | 0.1 | 0.05265961271 | 0.6931825427 | 15 | 0.1 | 0.05098003722 | 0.7217784390 |
| 15 | 0.2 | 0.05330512595 | 0.6899778133 | 15 | 0.2 | 0.05137956282 | 0.7174058104 |
| 15 | 0.3 | 0.05411134212 | 0.6871732710 | 15 | 0.3 | 0.05240654610 | 0.7056927565 |
| 15 | 0.4 | 0.05508926487 | 0.6847492765 | 15 | 0.4 | 0.05293500471 | 0.7004590233 |
| 15 | 0.5 | 0.05619344396 | 0.6827220878 | 15 | 0.5 | 0.05324102449 | 0.6970127299 |

Figure 4: Neural Network Methods

### *A. Rule-Based Models*

The Wang–Mendel (WM) method (Wang & Mendel, 1992) was one of the first methods to design fuzzy systems from data (Brown & Harris, 1994). Others known as “neuro-fuzzy” methods were (Lin & Lee, 1991). The method has been applied to a variety of problems and is one of the benchmark methods in the field (Cox, 1999). In the WM Fuzzy Inference model, RMSE was used to select the optimal model using the smallest value which was 0.08085976391 (5). In the WM Fuzzy Rules model, the final values used for the model were num.labels = 15 and type.mf = GAUSSIAN. In the Subtractive Clustering and Fuzzy c-Means RMSE was used to select the optimal which was r.a = 0.3, eps.high = 0.3 and eps.low = 0.2 as shown in Figure 5.

| r.a, eps.high/low | RMSE | Rsquared |
|---|---|---|
| 0.3,0.3,0.10 | 0.06477678 | 0.5755458 |
| 0.3,0.3,0.15 | 0.06477678 | 0.5755458 |
| 0.3,0.3,0.20 | 0.06477678 | 0.5755458 |
| 0.3,0.5,0.10 | 0.06477678 | 0.5755458 |
| 0.3,0.5,0.15 | 0.06477678 | 0.5755458 |
| 0.3,0.5,0.20 | 0.06477678 | 0.5755458 |
| 0.3,0.7,0.10 | 0.06477678 | 0.5755458 |
| 0.3,0.7,0.15 | 0.06477678 | 0.5755458 |
| 0.3,0.7,0.20 | 0.06477678 | 0.5755458 |
| 0.5,0.3,0.10 | 0.06752837 | 0.5593372 |
| 0.5,0.3,0.15 | 0.06752837 | 0.5593372 |
| 0.5,0.3,0.20 | 0.06752837 | 0.5593372 |
| 0.5,0.5,0.10 | 0.06752837 | 0.5593372 |
| 0.5,0.5,0.15 | 0.06752837 | 0.5593372 |
| 0.5,0.5,0.20 | 0.06752837 | 0.5593372 |
| 0.5,0.7,0.10 | 0.06752837 | 0.5593372 |
| 0.5,0.7,0.15 | 0.06752837 | 0.5593372 |
| 0.5,0.7,0.20 | 0.06752837 | 0.5593372 |
| 0.7,0.3,0.10 | 0.06858001 | 0.5591306 |
| 0.7,0.3,0.15 | 0.06858001 | 0.5591306 |
| 0.7,0.3,0.20 | 0.06858001 | 0.5591306 |
| 0.7,0.5,0.10 | 0.06858001 | 0.5591306 |
| 0.7,0.5,0.15 | 0.06858001 | 0.5591306 |
| 0.7,0.5,0.20 | 0.06858001 | 0.5591306 |
| 0.7,0.7,0.10 | 0.06858001 | 0.5591306 |
| 0.7,0.7,0.15 | 0.06858001 | 0.5591306 |
| 0.7,0.7,0.20 | 0.06858001 | 0.5591306 |
| *Normalised Error* | | |
| Test | 0.05450646 | |
| Training | 0.02600927 | |
| *Actual Error* | | |
| Test | 16.124102 | |
| Training | 7.694062 | |

Figure 5: Subtractive Clustering and Fuzzy c-Means Rules

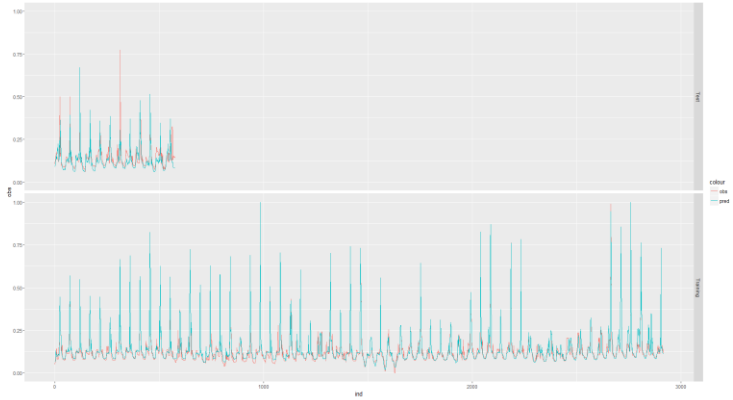


Figure 6: Subtractive Clustering and Fuzzy c-Means Rules

*B. Neural Network Models*

Next we tried Neural Networks with 2916 samples, 13 predictors and no pre-processing. The resampling was Cross-Validated (9 fold) with sample sizes: 2592, 2592, 2592, 2592, 2592, 2592. In the Neural Network model, RMSE was used to select the optimal model using the smallest value which was 15 and decay = 0.1 and in the Neural Network with feature extraction, the final values used for the model were size = 7 and decay = 0.1. The first experiment was the Bayesian Regularization for Feed-Forward Neural Networks model. The input variables are: [SMP_D_Minus_13_Euro, SMP_D_Euro, LoadDemand, SMP_D_Minus_1_Euro]. RMSE was used to select the optimal model using the smallest value which was neurons = 11 as shown in Figure 7.

| Neurons | RMSE | **Rsquared** |
|---|---|---|
| 11 | 0.07401631 | 0.5261056 |
| 13 | 0.09454580 | 0.4266669 |
| 15 | 0.08705493 | 0.4360314 |
| *Normalised Error* | | |
| Test | | 0.05517338 |
| Training | | 0.04931483 |
| *Actual Error* | | |
| Test | | 16.32139 |
| Training | | 14.58831 |

Figure 7: Bayesian Regularization for Feed-Forward Neural Networks

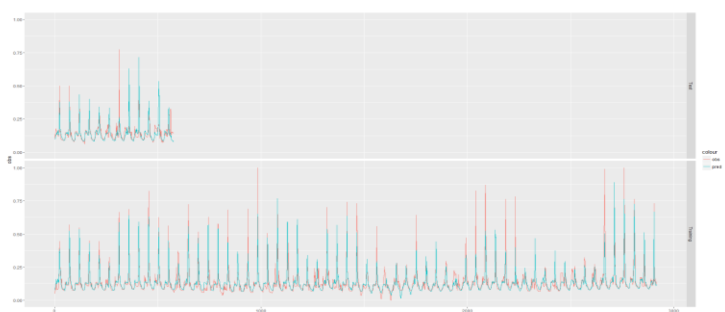


Figure 8: Bayesian Regularization for Feed-Forward Neural Networks

The next experiment was the multi-layer perceptron model. The input variables are: SMP_D_Minus_13_Euro, LoadDemand, SMP_D_Euro, SMP_D_Minus_1_Euro, SMP_D_Minus_6_Euro, SMP_D_Minus_2_Euro, SMP_D_Minus_3_Euro, SMP_D_Minus_5_Euro, SMP_HH_Minus_1_Euro, SMP_D_Minus_4_Euro, SMP_HH_Minus_2_Euro, Power_Prod_IRL]. The best result was neurons = 15 as shown in Figure 9.

| Neurons | RMSE | **Rsquared** |
|---|---|---|
| 11 | 0.05250854 | 0.7048104 |
| 13 | 0.05307737 | 0.7005197 |
| 15 | 0.05249322 | 0.7026434 |
| *Normalised Error* | | |
| Test | | 0.04292271 |
| Training | | 0.04648640 |
| *Actual Error* | | |
| Test | | 12.69740 |
| Training | | 13.75161 |

Figure 9: Multi-layer perceptron

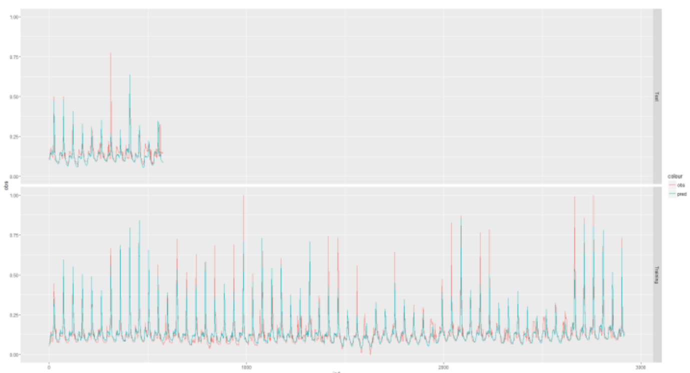


Figure 10: Multi-layer perceptron

In the Neural Networks experiment, the input variables are: [LoadDemand, Power_Production_Ireland, SMP_D_Minus_13_Euro, SMP_HH_Minus_2_Euro, SMP_D_Minus_2_Euro, Power_Production_UK, SMP_D_Minus_5_Euro, SMP_D_Minus_3_Euro, SMP_D_Minus_4_Euro, SMP_D_Euro, SMP_D_Minus_6_Euro]. RMSE was used to select the optimal model using the smallest value which was neurons = 11 and decay = 0.02 as shown in Figure 11.

| **size/decay** | **RMSE** | **Rsquared** |
|---|---|---|
| 11 0.010 | 0.05248126 | 0.7161288 |
| 11 0.015 | 0.05235626 | 0.7148340 |
| 11 0.020 | 0.05235063 | 0.7133702 |
| 13 0.010 | 0.05282604 | 0.7118659 |
| 13 0.015 | 0.05235792 | 0.7148175 |
| 13 0.020 | 0.05235315 | 0.7133737 |
| 15 0.010 | 0.05283664 | 0.7116809 |
| 15 0.015 | 0.05242537 | 0.7139332 |
| 15 0.020 | 0.05236808 | 0.7133572 |
| *Normalised Error* | | |
| Test | 0.04594939 | |
| Training | 0.04620509 | |
| *Actual Error* | | |
| Test | 13.59275 | |
| Training | 13.66839 | |

Figure 11: Neural Network

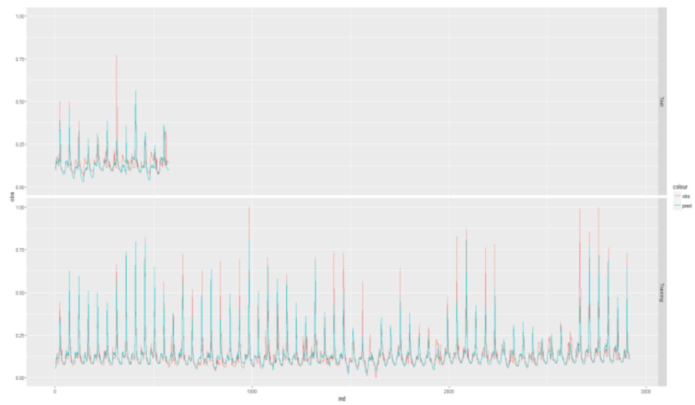


Figure 12: Neural Network

In the Neural Networks with Feature Extraction experiment, the input variables are: [SMP_D_Minus_13_Euro, LoadDemand, SMP_D_Euro, SMP_D_Minus_1_Euro, SMP_D_Minus_2 Euro]. RMSE was used to select the optimal model using the smallest value which was size = 13 and decay = 0.02 as shown in Figure 13.

| **size/decay** | RMSE | Rsquared |
|---|---|---|
| 11 0.010 | 0.06011476 | 0.6211297 |
| 11 0.015 | 0.05982740 | 0.6250083 |
| 11 0.020 | 0.06159925 | 0.6027796 |
| 13 0.010 | 0.06854780 | 0.5366326 |
| 13 0.015 | 0.05902941 | 0.6339155 |
| 13 0.020 | 0.05689748 | 0.6527683 |
| 15 0.010 | 0.06493007 | 0.5633139 |
| 15 0.015 | 0.06272061 | 0.5915866 |
| 15 0.020 | 0.06343059 | 0.5956760 |
| *Normalised Error* | | |
| Test | | 0.06434868 |
| Training | | 0.04389391 |
| *Actual Error* | | |
| Test | | 19.03563 |
| Training | | 12.98470 |
| | | |

Figure 13: Neural Networks with Feature Extraction

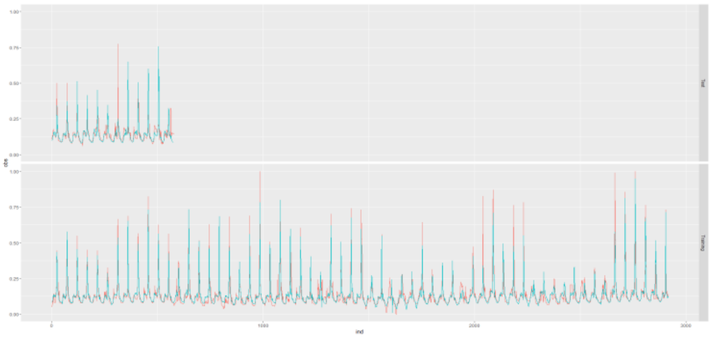


Figure 14: Neural Networks with Feature Extraction

In the Radial Basis Function Network experiment, the input variables are: [SMP_D_Minus_13_Euro, LoadDemand, SMP_D_Euro, SMP_D_Minus_1_Euro, SMP_D_Minus_2_Euro, SMP_D_Minus_6_Euro, SMP_D_Minus_3_Euro, SMP_HH_Minus_1_Euro]. RMSE was used to select the optimal model using the smallest value which was size = 11as shown in Figure 15.

| Size | RMSE | Rsquared |
|---|---|---|
| 11 | 0.06419983 | 0.6427550 |
| 13 | 0.08197198 | 0.6383784 |
| 15 | 0.07501390 | 0.6455758 |
| *Normalised Error* | | |
| Test | 0.04858537 | |
| Training | 0.05931154 | |
| *Actual Error* | | |
| Test | 14.37252 | |
| Training | 17.54554 | |

Figure 15: Radial Basis Function Network

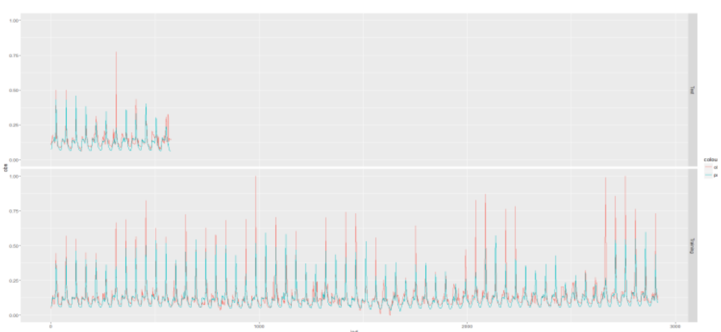


Figure 16: Radial Basis Function Network

In the Multi-Layer Perceptron, with multiple layers experiment, the input variables are: [SMP_D_Minus_13_Euro, LoadDemand, SMP_D_Euro, SMP_D_Minus_1_Euro, SMP_D_Minus_2_Euro, SMP_D_Minus_6_Euro, SMP_D_Minus_3_Euro, SMP_D_Minus_5_Euro, SMP_HH_Minus_1_Euro, SMP_D_Minus_4_Euro, SMP_HH_Minus_2_Euro, Power_Production_Ireland, Power_Production_UK]. RMSE was used to select the optimal model using the smallest value which were layer1 = 13, layer2 = 13 and layer3 = 13 as shown in Figure 17.

| Layer 1,2,3 | RMSE | Rsquared |
|---|---|---|
| 11,11,11 | 0.05234793 | 0.7152431 |
| 11,11,13 | 0.05179760 | 0.7129437 |
| 11,11,15 | 0.05226256 | 0.7113108 |
| 11,13,11 | 0.05290105 | 0.7121994 |
| 11,13,13 | 0.05223151 | 0.7125928 |
| 11,13,15 | 0.05216129 | 0.7157355 |
| 11,15,11 | 0.05263369 | 0.7146328 |
| 11,15,13 | 0.05224950 | 0.7170724 |
| 11,15,15 | 0.05196844 | 0.7155363 |
| 13,11,11 | 0.05212525 | 0.7148120 |
| 13,11,13 | 0.05213235 | 0.7119987 |
| 13,11,15 | 0.05267301 | 0.7136223 |
| 13,13,11 | 0.05222891 | 0.7140849 |
| 13,13,13 | 0.05135011 | 0.7171306 |
| 13,13,15 | 0.05378522 | 0.7171374 |
| 13,15,11 | 0.05153877 | 0.7140732 |
| 13,15,13 | 0.05178785 | 0.7136781 |
| 13,15,15 | 0.05203611 | 0.7165408 |
| 15,11,11 | 0.05686386 | 0.7133883 |
| 15,11,13 | 0.05192465 | 0.7141276 |
| 15,11,15 | 0.05257671 | 0.7138850 |
| 15,13,11 | 0.05195074 | 0.7144070 |
| 15,13,13 | 0.05283617 | 0.7131725 |
| 15,13,15 | 0.05171448 | 0.7176357 |
| 15,15,11 | 0.05223462 | 0.7176678 |
| 15,15,13 | 0.05187319 | 0.7159089 |
| 15,15,15 | 0.05162038 | 0.7169941 |
| *Normalised Error* | | |
| Test | 0.04469686 | |
| Training | 0.04738835 | |
| *Actual Error* | | |
| Test | 13.22223 | |
| Training | 14.01842 | |

Figure 17: Multi-Layer Perceptron, with multiple layers

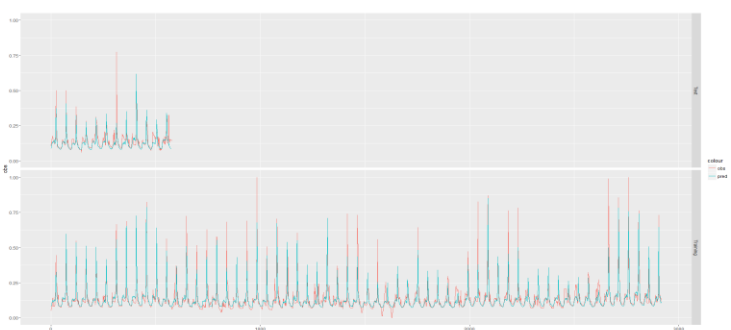
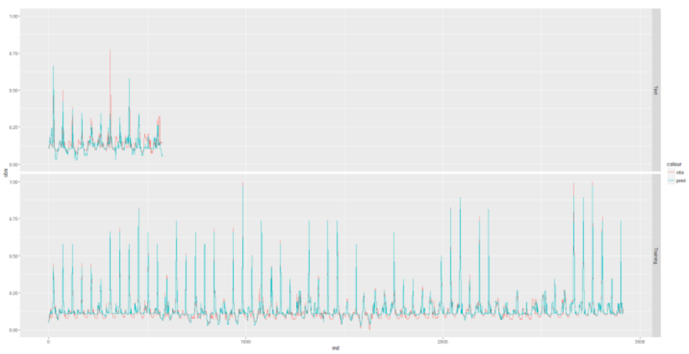


Figure 18: Multi-Layer Perceptron, with multiple layers

In the Wang and Mendel Fuzzy Rules experiment, the input variables are: [SMP_D_Minus_13_Euro, SMP_D_Euro, LoadDemand, SMP_D_Minus_1_Euro]. Tuning parameter 'type.mf' was held constant at a value of GAUSSIAN. RMSE was used to select the optimal model using the smallest value which was num.labels = 15 and type.mf = GAUSSIAN as shown in Figure 19.

| Num labels | RMSE | Rsquared |
|---|---|---|
| 13 | 0.06502759 | 0.5887993 |
| 15 | 0.06490675 | 0.5796969 |
| *Normalised Error* | | |
| | Test | 0.05378375 |
| | Training | 0.02158287 |
| *Actual Error* | | |
| | Test | 15.910308 |
| | Training | 6.384645 |

Figure 19: Wang and Mendel Fuzzy Rules



A comparison between the models is shown in Figure 20.

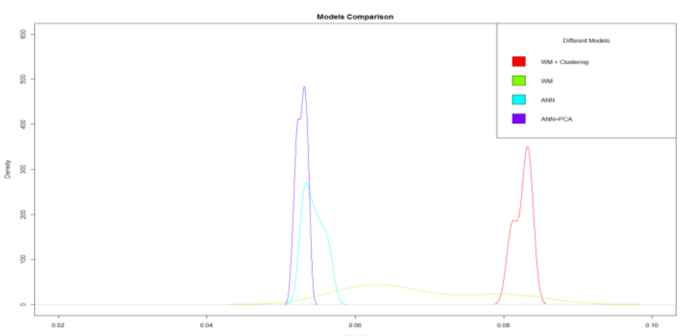


Figure 20: Comparison of Models

## IV. CONCLUSION

Short-term load forecast is an essential part of electric power system planning and operation. This paper presents a series of rules based and neural network based approaches for short-term load forecasting that uses the correlated weather data for training, validating and testing of a neural network. Time series prediction is one of the most important prediction that collect past observations of a variable and analyze it to obtain the underlying relationships between historical observations but time series has properties such as nonlinearity, chaotic, non-stationary and cyclic which cause problems. An adaptive neural network based fuzzy inference system (ANFIS) is where the learning processes are performed by interleaving the optimization of the antecedent and conclusion parts parameters. We believe ANNs permit modelling of complex and nonlinear relationships through training with the use of historical data and can therefore be used in models based on weather information without the need for assumptions for any functional relationship between load and weather variables

## REFERENCES

Balabin, R., Lomakina, E. (2009). Neural network approach to quantum-chemistry data: Accurate prediction of density functional theory energies. *J. Chem. Phys.* Vol. 131, No. 7, pp: 74-104. doi:10.1063/1.3206326. PMID 19708729.

Brown, M. and Harris, C. (1994) Neurofuzzy Adaptive Modeling and Control. Upper Saddle River, NJ: Prentice-Hall, 1994. [4] V. Cherkassky, J. H. Friedman, and H. Wechsler, Eds., From Statistics to Neural Networks. Berlin, Germany: Springer-Verlag, 1994.

CER (2011) CER Factsheet on the Single Electricity Market, Commission for Energy Regulation, April 2011 http://www.cer.ie/docs/000262/cer11075.pdf

Ciresan, D., Meier, U., Schmidhuber, J. (2012) Multi-column Deep Neural Networks for Image Classification. IEEE Conf. on Computer Vision and Pattern Recognition CVPR 2012.

Cox, E. (1999) The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, and Maintaining Fuzzy Systems. San Diego, CA: AP Professionals, 1999.

Graves, A., Liwicki, M., Fernandez, S., Bertolami, R., Bunke, H., Schmidhuber. J. (2009) A Novel Connectionist System for Improved Unconstrained Handwriting Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 5, 2009.

Lin, C. and Lee, C. Neural network-based fuzzy logic control and decision system, *IEEE Trans. Comput.*, vol. 40, pp. 1320–1336, Dec. 1991

Lee, K. (1992) Short-Term Load Forecasting Using an Artificial neural Network, IEEE Trans. *On Power Systems,* Vol. 7, No. 1, pp. 124-131, Feb.1992.

McCulloch, Warren; Walter Pitts (1943). A Logical Calculus of Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics, Vol.* 5, No. 4, pp: 115–133, doi:10.1007/BF02478259.

Rui, Y. & El-Keib, A. (1995) A review of ANN-based Short-term load forecasting models, Proc. Of 27th South Eastern Symposium on System Theory, Starkville, MS, 12-14 March 1995, pp: 78-82

Sharma, L., Chakrawarti, M., Dutta, A., Adhikari, N. (2012) Neural Network Based Approach for Short-term Load Forecasting, International Journal of Science, Environment and Technology, Vol. 1, No 5, 2012, 416 – 424

Srinivasan, D. (1994) A neural network short-term load forecaster, *Electric Power Research*, pp. 227-234, 28 (1994).

Wang, L. and Mendel, J. (1992) Generating fuzzy rules by learning from examples, *IEEE Trans. Syst., Man, Cybern.*, vol. 22, pp. 1414–1427, Dec. 1992.

## ACKNOWLEDGEMENTS

This work was funded by an InterTradeIreland FUSION programme.