❒   4060

Forecasting Short-term Wholesale Prices on the Irish Single Electricity Market

Francesco Arci[1], Jane Reilly[2], Pengfei Li[3], Kevin Curran[4], Ammar Belatreche[5]

[1,2,3]Ark Energy Consulting Limited, Unit 20 Daingean Hall, N4 Axis Centre, Battery Road, Longford, Irlandia
[4]Faculty of Computing, Engineering & Built Environment, Ulster University, Northern Ireland, United Kingdom
[5]Department of Computer Science and Digital Technologies, Northumbria University, United Kingdom

Article Info	ABSTRACT
	Electricity markets are different from other markets as electricity generation cannot be easily stored in substantial 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 considerable extent, cost bids from generators must be balanced with demand estimates in advance of real-time. This paper outlines a a forecasting algorithm built on artificial neural networks to predict short-term wholesale prices on the Irish Single Electricity Market so that market participants can make more informed trading decisions. 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. We have identified the features that such a model demands and outline it here.

Corresponding Author:

Kevin Curran,
Faculty of Computing, Engineering & Built Environment,
Ulster University, Northern Ireland, United Kingdom.
Email: kj.curran@ulster.ac.uk

1. INTRODUCTION

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.

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 take into account additional factors such as meteorological forecasts, forecasted demand and constraints imposed by network topology [1], [2]. 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 [3], [4].

There is a small number of companies providing or working on a robust approach to forecasting both renewable power output and/or marginal prices for electricity. Alba Soluzioni are an independent consultancy providing information, training and bespoke consultancy services in the European gas and power markets [5]. Their primary publication to date is considered a reference on Italian gas & power markets. They currently provide a short-term marginal price forecast service. MKOnline provides online market intelligence services to provide clients with timely and high-resolution forecasts of fundamentals and prices for the short, mid and long term horizon. It also offers a complementary weather service [6].

Meteologica supply forecasts of business variables related to weather through the provision of integrated forecasting solutions, unique to each client. Meteologica specializes in wind and solar power forecasting services all around the world [7]. Their forecasts are utilized by hundreds of plant owners, power traders and grid operators to optimize their business activities. From preliminary conversations as a potential supplier of renewable power output data, Meteologica is currently working on an SMP forecasting service. Meteogroup is a global private weather business with offices around the world [8]. They aim to combine experience and global coverage with local expertise to offer our customers highly accurate and bespoke weather services. Meteogroup has recently launched a portal aimed at presenting meteorological information useful to energy traders.

Used ANN-based load forecasting methods for 24-hour-ahead peak load forecasting by using forecasted temperature [9]. They proposed a one hour-ahead load forecasting method using the most significant weather data. In the proposed forecasting method, weather data is first analyzed to determine the most correlated factors to load changes. The most correlated weather data is then used in training, validating and testing the neural network. Correlation analysis of weather data was used to determine the input parameters of the neural networks and they tested it on actual load data from the Egyptian Unified System.

Outline a neural network approach for forecasting short-term electricity prices using a back-propagation algorithm [10]. The results obtained from their neural network show that the neural network-based approach is more accurate. Present an ANN based short-term load forecasting model for a substation in Kano, Nigeria [11]. The recorded daily load profile with a lead time of 1-24 hours for the year 2005 was obtained from the utility company. The Levenberg-Marquardt optimization technique was used as a back-propagation algorithm for the Multilayer Feed Forward ANN. The forecasted next day 24 hourly peak loads were obtained based on the stationary output of the ANN with a performance Mean Squared Error (MSE) of $5.84e^{-6}$ and compared favorably with the actual Power utility data. The results showed that their technique is robust in forecasting future load demands for the daily operational planning of power system distribution sub-stations in Nigeria.

Short-term load forecast is therefore an essential part of electric power system planning and operation. Forecasted values of system load affect the decisions made for unit commitment and security assessment, which have a direct impact on operational costs and system security. Conventional regression methods are used by most power companies for load forecasting. However, due to the nonlinear relationship between load and factors affecting it, conventional methods are not sufficient enough to provide accurate load forecast or to consider the seasonal variations of load.

We believe artificial neural networks (ANN) based load forecasting methods can deal with 24-hour-ahead load forecasting by using forecasted weather input variables, which can lead to high forecasting errors in case of rapid weather changes [12], [13]. 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. We outline here a novel neural network-based approach for short-term load forecasting that uses the correlated weather data for training, validating and testing of a neural network. Correlation analysis of weather data determines the input parameters of the neural networks. The suitability of the proposed approach is illustrated through an application to the actual load data of the Irish Electricity Market. This paper is organised as follows: Section 2 provides a background to the Single Electricity market in Ireland, section, section 3 introduces Artificial Neural Networks & Short-term Load Forecasting, section 4 presents the short-term forecasting model and section 5 provides a conclusion.

2.   SINGLE ELECTRICITY MARKET

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 [14]. 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 [15] is responsible for administering the market, including paying generators for their electricity generated and invoicing suppliers for the electricity they have bought [14].

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 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 [16]. 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 as shown in Figure 1.
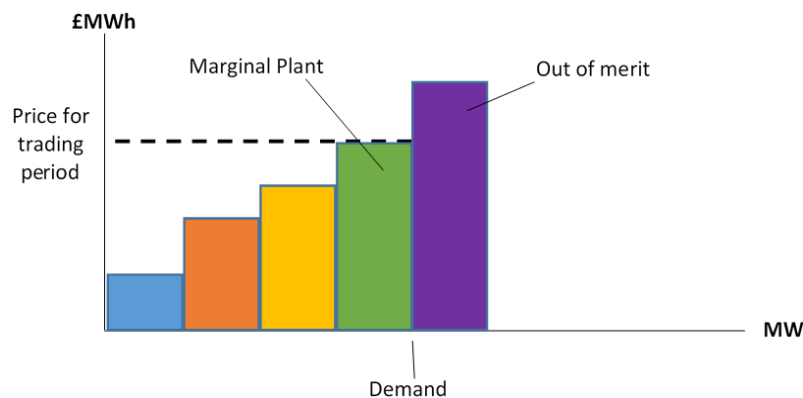

Figure 1. The role of System Marginal Price

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, as illustrated in Figure 2. 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.
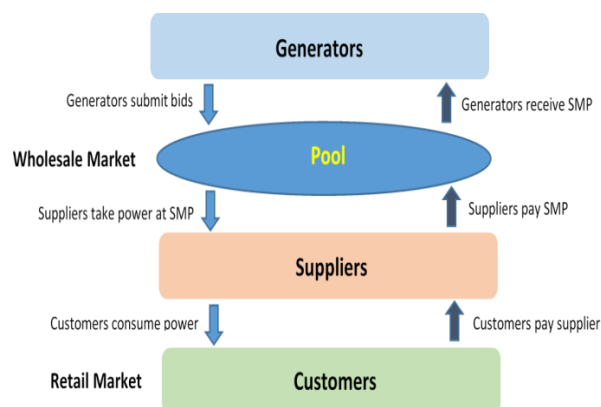

Figure 2. Wholesale and retail market

Generators operating within the SEM also receive separate capacity payments which contribute towards their fixed costs, if they are available to generate. The capacity payment pot of money for generators is set ahead of time by the SEM Committee and is calculated based on the relatively low fixed costs of a peaking plant. As a result, the payments generally cover only a portion of the fixed costs involved in building most plants. Suppliers also pay for these capacity payments and any other system charges, which are typically passed through to customers. To sell electricity into the SEM pool, generators must submit cost bids to SEMO the day before the physical trade/generation takes place, known as D-1. The bids submitted are primarily based on a generator's running or Short Run Marginal Cost (SRMC), i.e. the cost of each extra MW it could produce excluding its fixed costs. The SRMC reflects the opportunity cost of the electricity produced, which is the economic activity that the generator forgoes to produce electricity. For example, in the case of a generator fuelled by gas, the opportunity cost includes the price of gas on the day that it is bidding in, because if the generator was not producing electricity it could sell its gas in the open market. Generator bids also include a generator's start-up costs, which are costs it faces if it needs to be turned on after a period of inactivity, as well as generator no-load costs which are (mostly fuel) costs which are indifferent to output levels.

The generators submit these bids to SEMO up until Gate Closure, currently at 10:00am on D-1. Software is then run by SEMO to determine a Market Schedule which forecasts the SMP for each half hour trading period for the following day. However, no software can predict with complete accuracy what will happen in reality: real-time factors such as a change in wind generation or customer demand, which can affect SMP, must be accounted for. For this reason, SEMO completes two more software runs reflecting the reality of what happened in generator dispatch, one on the day after the trading day (D+1), and another four days after (D+4), to calculate the final SMP for each half hour of the trading day. This D+4 price is the one that is paid to generators and paid by suppliers. The Market Schedule identifies the lowest cost solution at which generation can meet demand for each half hour trading period. It ranks generators with the lowest bids first until the quantity needed for the demand is met - see blue shaded bars in

Figure 3. The marginal generator needed to meet the demand sets the SMP for that trading period. The other generators who have submitted SRMC bids lower than this price are deemed to be "in merit" and will also be scheduled to run. All generators who have submitted bids which are higher than this price (SMP) are deemed to be "out of merit" and will not be scheduled to run - see the green bar in Figure 3. These tend to be old or inefficient plants.
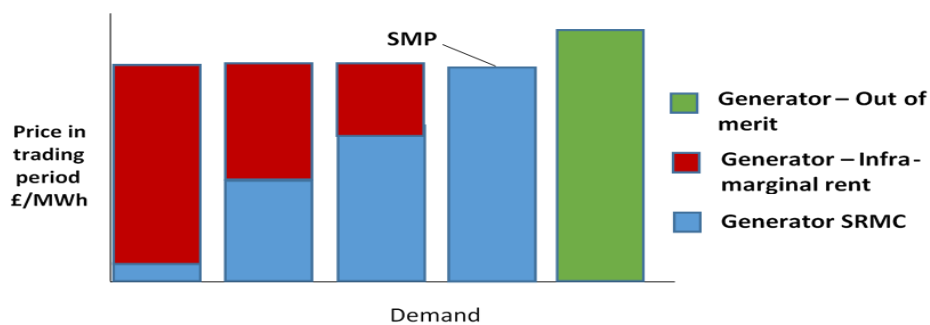

Figure 3. Market schedule

All generators who have submitted a bid which is under the SMP earn a profit, known as "inframarginal rent", on the difference between their SRMC bid offer and the SMP. This is illustrated in red shaded bars in the graph. The plant that sets the marginal price in a half hour, i.e. the one with the highest running costs among those that are scheduled to run, does not receive any infra-marginal rent. However, this is typically a peaking plant which, while it has high short-run costs, has low fixed costs. Hence its costs are covered through the SMP and the capacity payments it receives. Infra-marginal rent is needed for most generators that are run, including efficient modern gas plants and wind farms, because while such plants have relatively low running costs (SRMC), they have much higher fixed costs which the (relatively low) capacity payment does not fully cover. Without infra-marginal rent, it would not be economic to build modern efficient power plants or wind farms, threatening security of electricity supply and driving higher prices in the long-run.

Wind farms are an example of electricity generators that have very low SRMC-the wind is free-and so typically they receive a higher rate of infra-marginal rent than other electricity generators, which

in turn is needed to pay for their much higher fixed costs. If a generator was dispatched more than it was scheduled to in the Market Schedule, for example to compensate for another (cheaper) generator not being brought online due to a network failure or "constraint", it is "constrained on". This means it receives its bid cost to compensate for the extra MW it must produce, though it does not receive infra-marginal rent. Generators who were originally included in the Market Schedule, but not actually run for reasons outside of their control, for example due to a network fault, are said to be "constrained off". They receive the SMP less their bid, i.e. the infra-marginal rent they would have received in the market had they been run. Constraints costs also cover costs associated with "reserve". This is where, to ensure the continued security of the system, for example in the event of a generator tripping, some generators are instructed to run at lower levels than indicated in the Market Schedule. This means there is spare generation capacity available (reserve) which can be quickly brought online if needed. To maintain the demand-supply balance, this reserve means that some generators will be constrained down while others may be constrained on/up, again leading to the actual dispatch deviating from the Market Schedule [14].

3.   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 [17]. Artificial neural networks are generally presented as systems of interconnected "neurons" which exchange messages between each other [18]. The connections have numeric weights that can be tuned based on experience, making neural nets adaptive to inputs and capable of learning. For example, a neural network for handwriting recognition is defined by a set of input neurons which may be activated by the pixels of an input image. After being weighted and transformed by a function (determined by the network's designer), the activations of these neurons are then passed on to other neurons [19]. This process is repeated until finally, an output neuron is activated. This determines which character was read. 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 [20].

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 [21]. 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 [22]. Although a fully connected ANN can 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) [23]. Due to its dynamic nature, the RHONN forecasting model can adapt quickly to changing conditions such as important load variations or changes of the daily load pattern [22]. 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. [24] 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.

4.   A SHORT-TERM ELECTRICITY MARKET 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 like 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.

To obtain good forecasting results, day type information must be considered. We can achieve this by constructing different ANNs for each day type and feeding each ANN the corresponding day type training sets [25], [26]. Another way is to use only one ANN but contain the day type information in the input variables [27]. 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 like 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 [28], [29]. 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 [30].

Other key 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 [15] can see in Table 1.

Table 1. Key data fields

Variables name	The unit of measurement	Example
Trade date	Day of month	1 Feb 2016
Delivery date	Half Hour	1 Feb 2016 06:00
Jurisdiction		ROI/NI
Forecast MW	Megawatts	2551.98
Solarpower	Megawatts	0
Solarpower Utilization	%	0
Windpower	Megawatts	2022
Windpower Utilization	%	81
SMP	Euro	18.9
Shadow Price	Euro	18.809999

We plot a subset of data. Figure 4 shows Solar power production in Northern Ireland. The value of horizontal axis is time domain from 1st Feb 2016 to 9th Feb 2016. The Red line indicates the solar power production (MW) in Northern Ireland and the blue line indicates the solar power utilization rate (%) in Northern Ireland.

In Figure 5, the Red line indicates the wind power production (MW) in Northern Ireland and the blue line indicates the wind power utilization rate (%) in Northern Ireland. In Figure 6, the Red line indicates the wind power production (MW) in the Republic of Ireland and the blue line indicates the wind power utilization rate (%) in the Republic of Ireland. There is no solar power production in the Republic of Ireland. Figure 7 shows the demand prediction (Megawatts) of the Republic of Ireland and Northern Ireland. The Red line indicates the demand (MW) in Republic of Ireland and the blue line indicates the demand (MW) in

Northern Ireland. Figure 8 shows the SMP for North and South. The Red line indicates the SMP (Euro) in the Republic of Ireland and the blue line indicates the shadow price (Euro) in the Republic of Ireland.
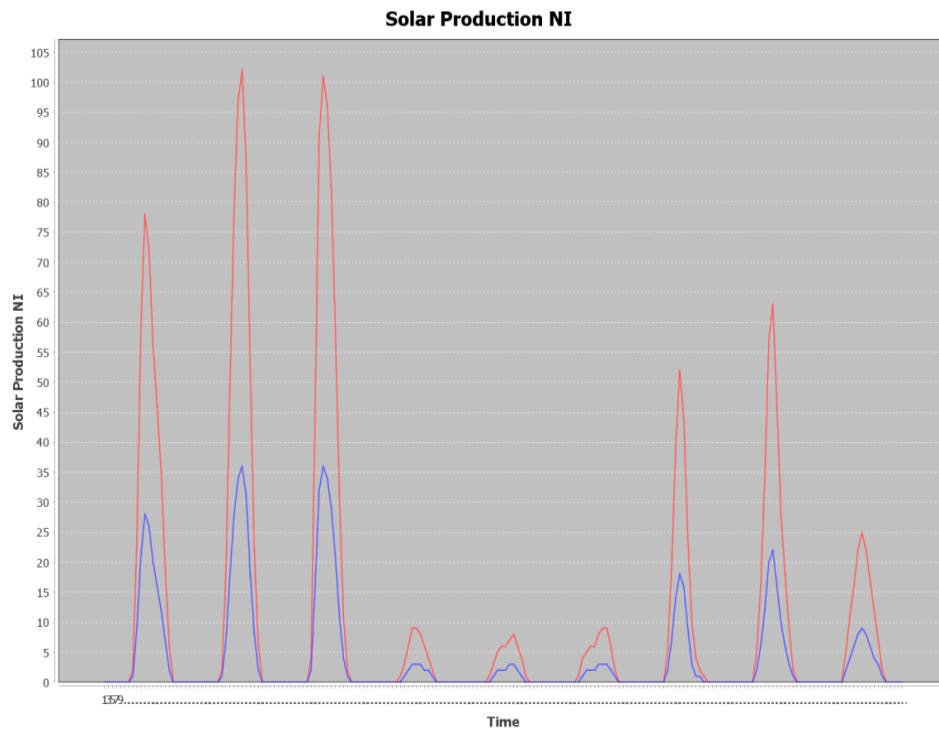

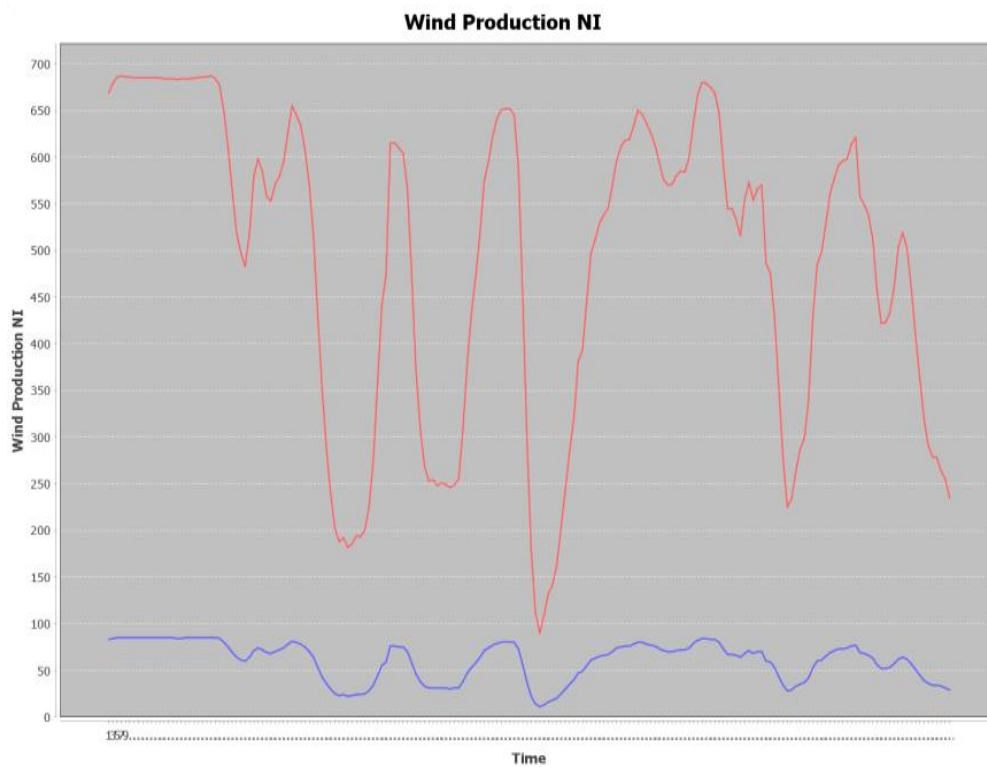Figure 4. Solar power production in Northern Ireland

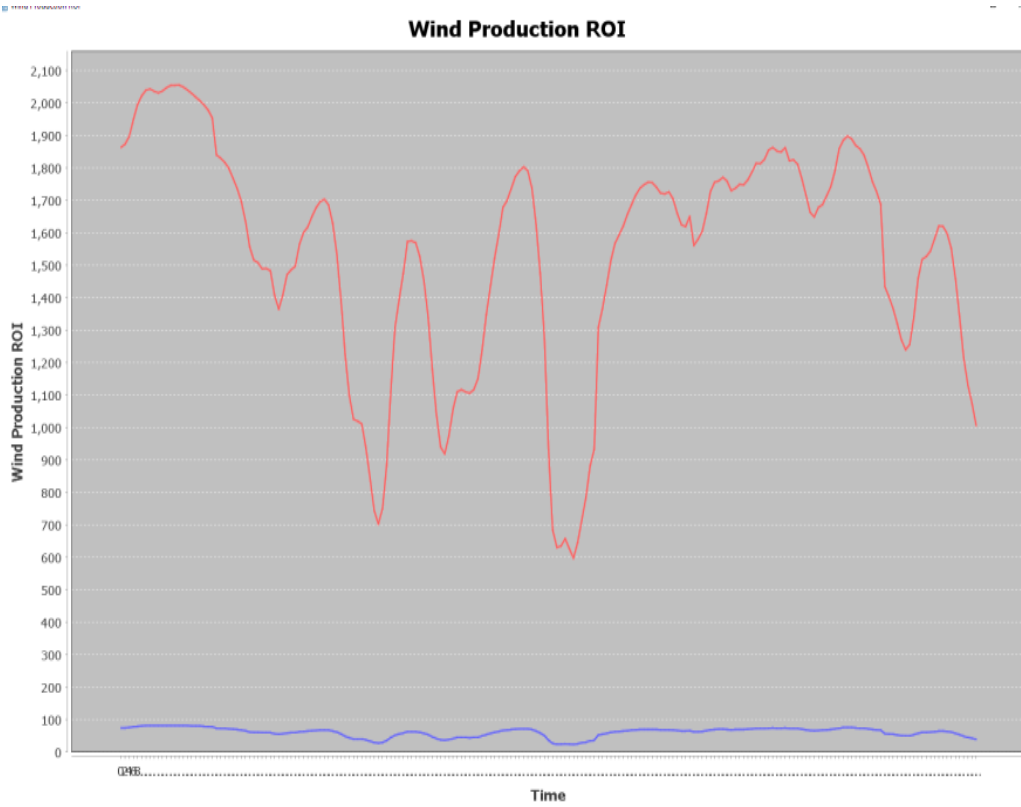Figure 5. Wind Production power in Northern Ireland

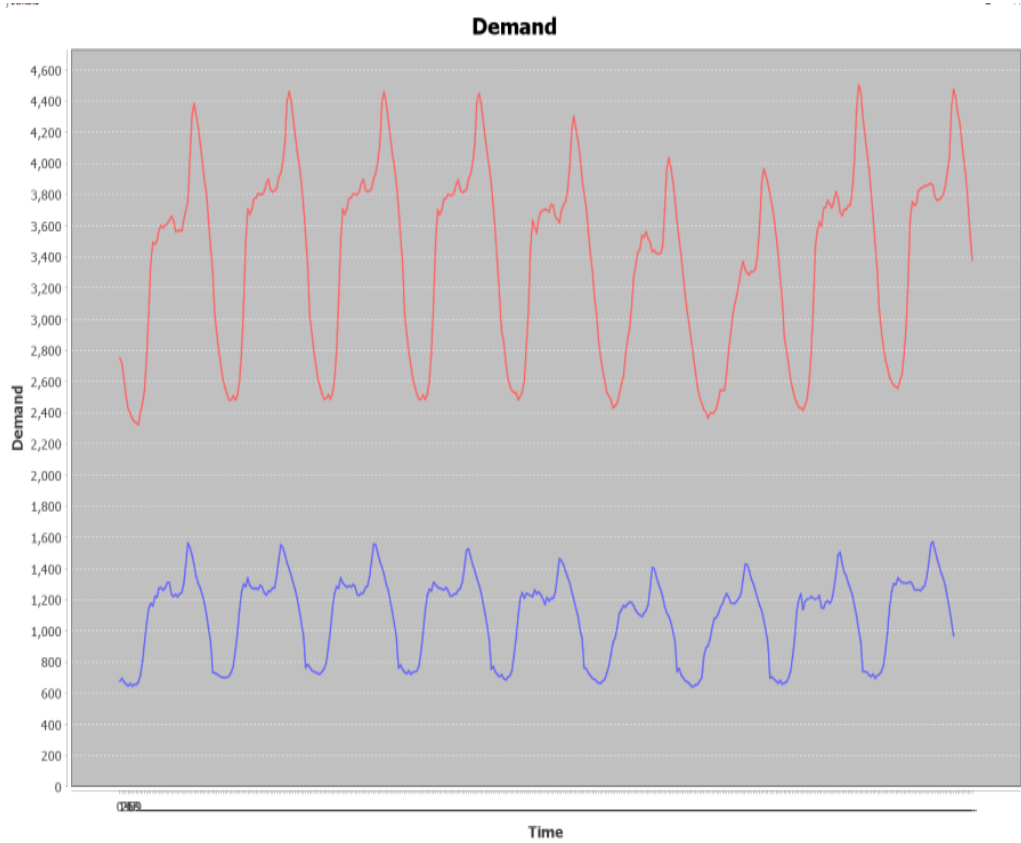Figure 6. Wind Production power in Rep. of Ireland

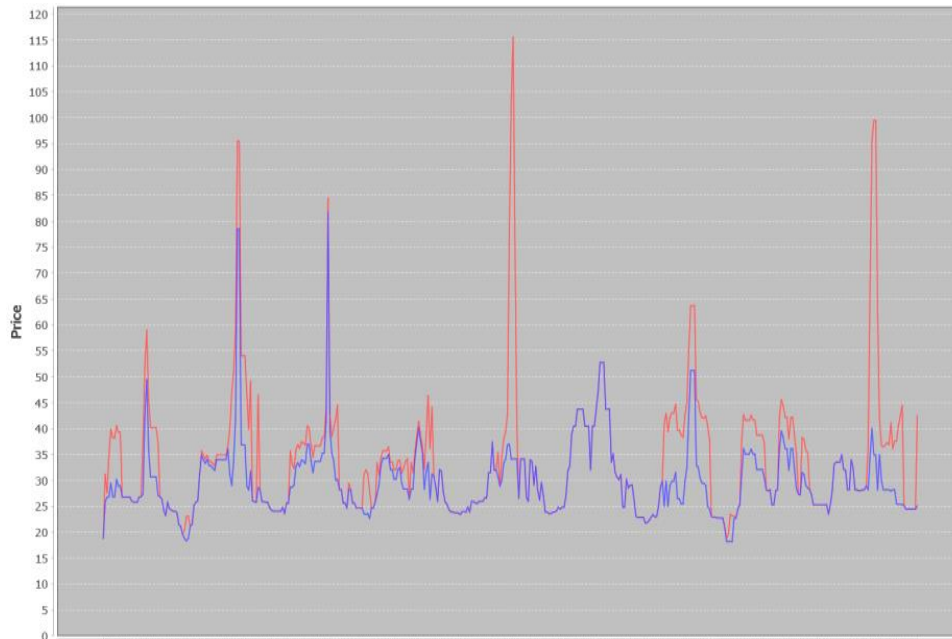Figure 7. Demand prediction in Republic of Ireland

Figure 8. SMP in Northern Ireland and Republic of Ireland

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. The ANFIS model we are using is a Takagi-type Neuro-fuzzy Network which combines neural networks and fuzzy systems. Fuzzy reasoning and network calculation will be available simultaneously.

Before we employ the ANFIS method to forecast the daily electricity SMP data, the raw data needed to be preprocessed to get the proper input and we need to determine the data input variables. One input data sample input consists of Production Forecasting (D-2), Load Forecasting (D-2) and Previous Prices Window (D-9… D-2). The data of production forecasting and load forecasting can be obtained from the Ex-Ante lag-2 file. The data of previous prices window can be obtained from the Ex-Ante files of lag-2, lag-3, …, lag-9. Production forecasting includes 9*2*48 variables, Load Forecasting includes 4*2*48 variables and Previous Prices Window includes 7*2*48 variables. Output (D) includes 48 variables to compare with control data. This output can see in Table 2..

Table 2. Four Day Rolling Load Forecast Sample

	Production Forecasting	Load Forecasting	Previous Prices Window	Output	Control Data (Output)
Data Sample 1	D-2 (9-day Forecasting)	D-2 (4-day Forecasting)	EA(D-9), EA(D-8), …, EA(D-2)	D	EA(D)= (H1, H2 ,…)
Data Sample 2	D-3	D-3	EA(D-10), EA(D-9), …, EA(D-3)	D-1	EA(D-1)

We experimented with other algorithm to determine the parameters of the ANFIS model (Grid Partitioning, subtractive clustering and FCM clustering), training method (SOM algorithm, Levenberg-Marquardt algorithm, Bayesian Regularization and Scaled Conjugate Gradient), AR model, state space model and ARIMAX model, Neural Network and Fuzzy Inference System.

Next, we examine our feature selection methodology. 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. The notation used throughout the paper is provided in Table 3. The training data is shown in Table 4. The WM methods are shown in Table 5 and Table 6 shows the neural networks methods.

Table 3. Nomenclature used

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

Table 4. Training data set

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

Table 5. WM Methods

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

Table 6. Neural Network 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

Table 6. Neural Network Methods

		Neural Network				Neural Network with Feature Extraction	
Size	Decay	RMSE	RSquared	Size	Decay	RMSE	RSquared
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

4.1. Rule-Based Models

The Wang-Mendel (WM) method [31] was one of the first methods to design fuzzy systems from data [32]. Others known as "neuro-fuzzy" methods were [33]. The method has been applied to a variety of problems and is one of the benchmark methods in the field [34]. 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 Table 7.

Table 7. Subtractive Clustering and Fuzzy c-Means Rules

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	

4.2. 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 Figures 9 and 10.
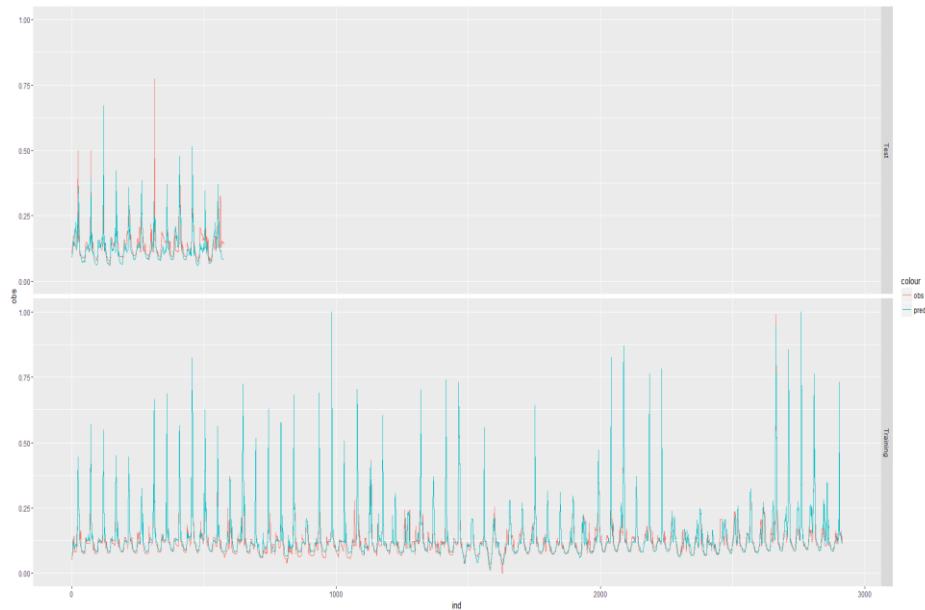

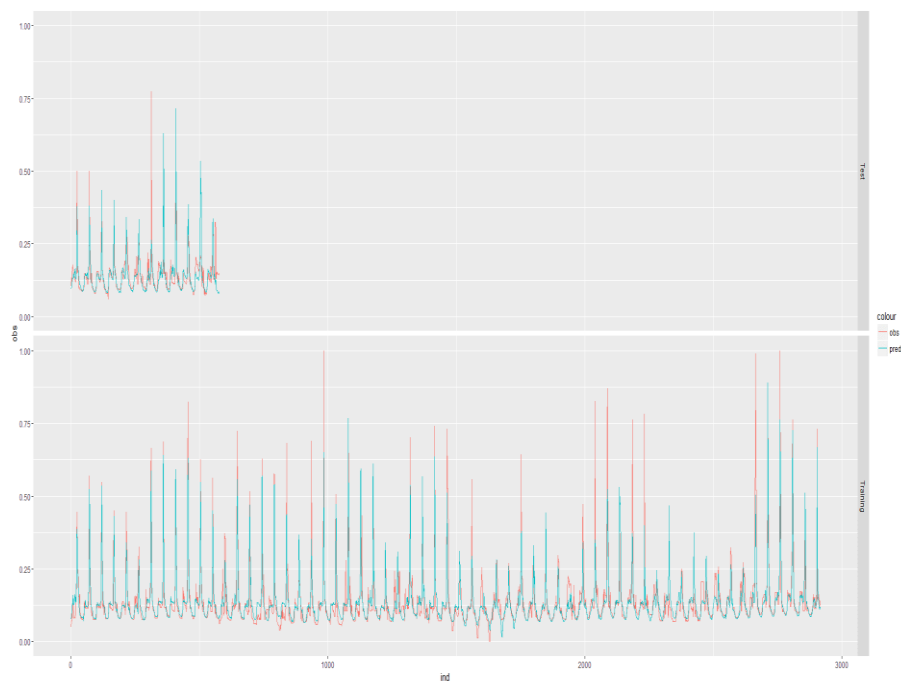Figure 9. Subtractive Clustering and Fuzzy c-Means Rules

Figure 10. 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 Tables 8 and 9.

Table 8. Bayesian Regularization for Feed-Forward Neural Networks

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

Table 9. Multi-layer perceptron

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

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 Figures 11 and 12.
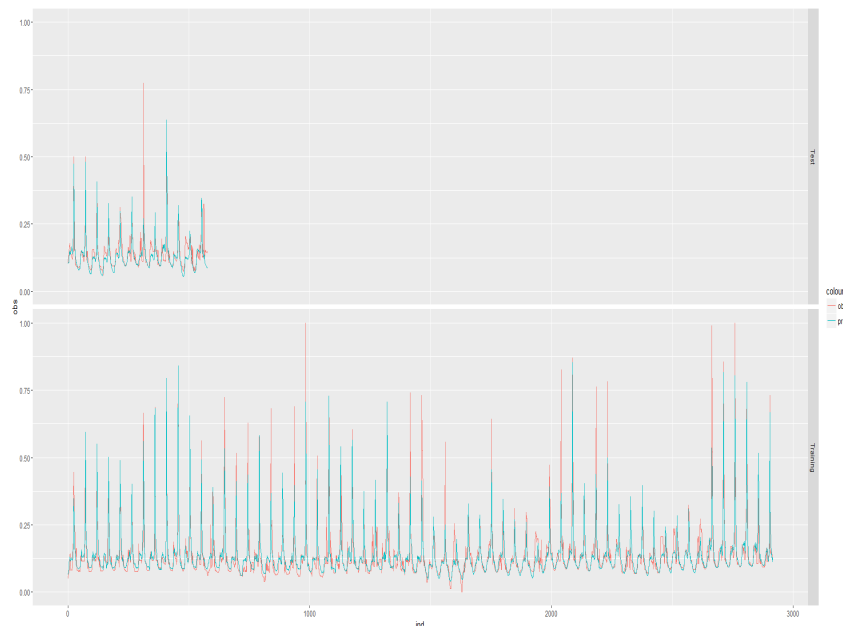

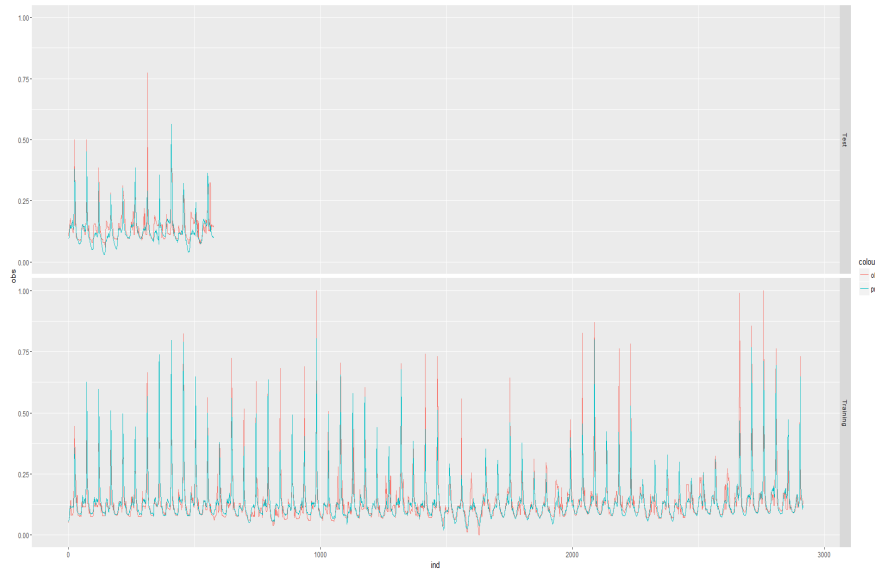Figure 11. Multi-layer perceptron

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 Tables 10 and 11.

Table 4. Neural Network

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

Table 5. Neural Networks with Feature Extraction

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

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 Figures 13 and 14.
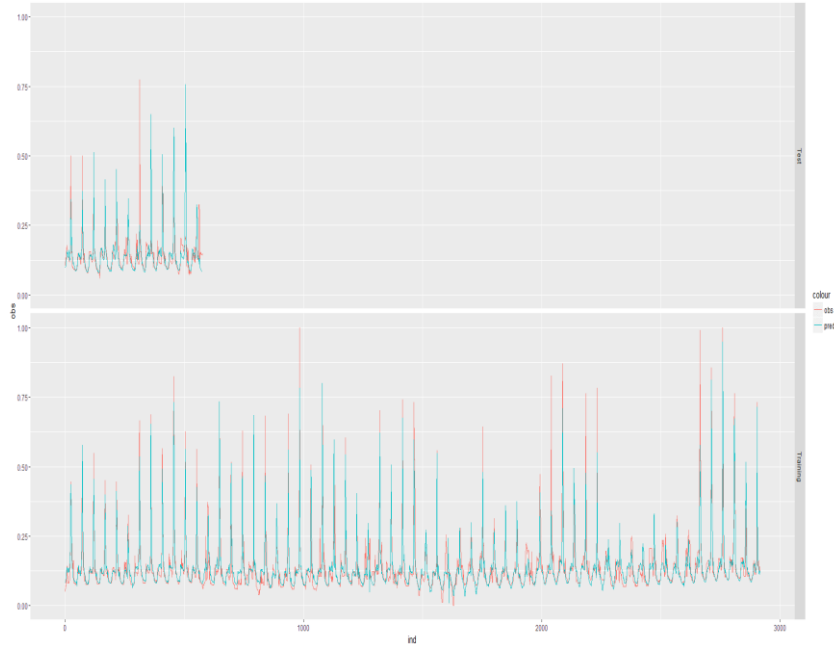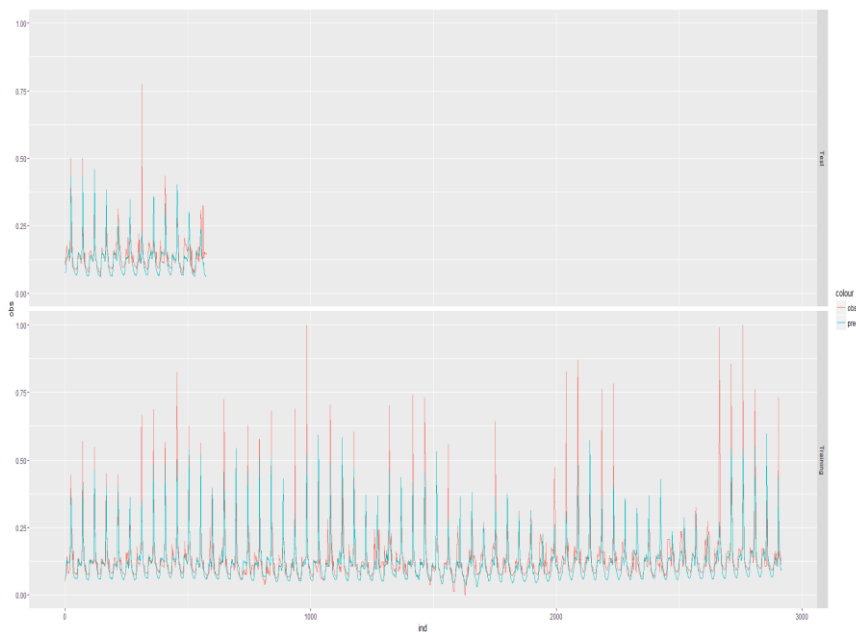

Figure 13. Neural Networks with Feature Extraction

Figure 14. 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 can see in Table 12 and Figure 15.

Table 12. Radial Basis Function Network

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

Table 13. Wang and Mendel Fuzzy Rules

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

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]. The 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 Table 13 and Figure 16. A comparison between the models is shown in Figure 17.
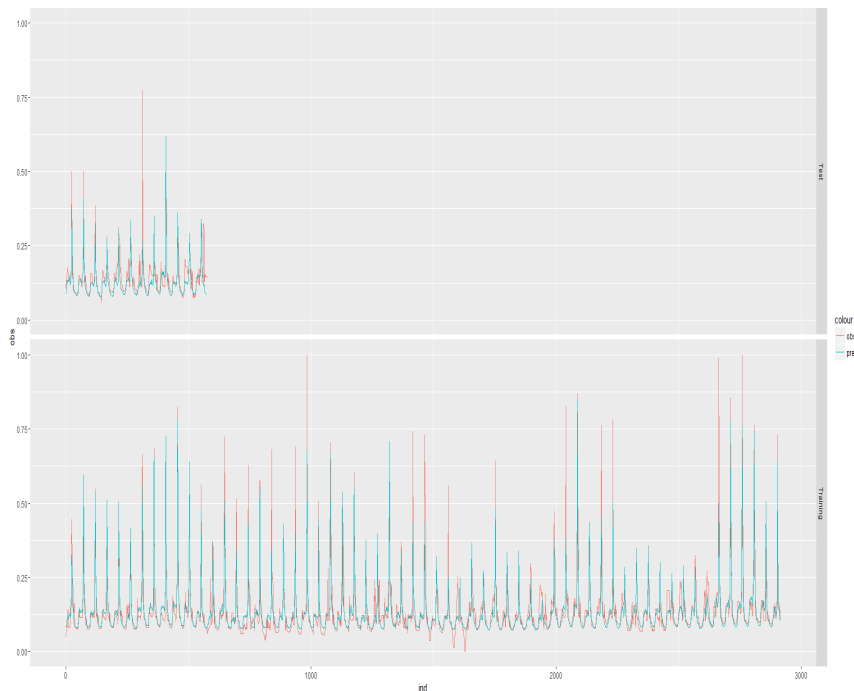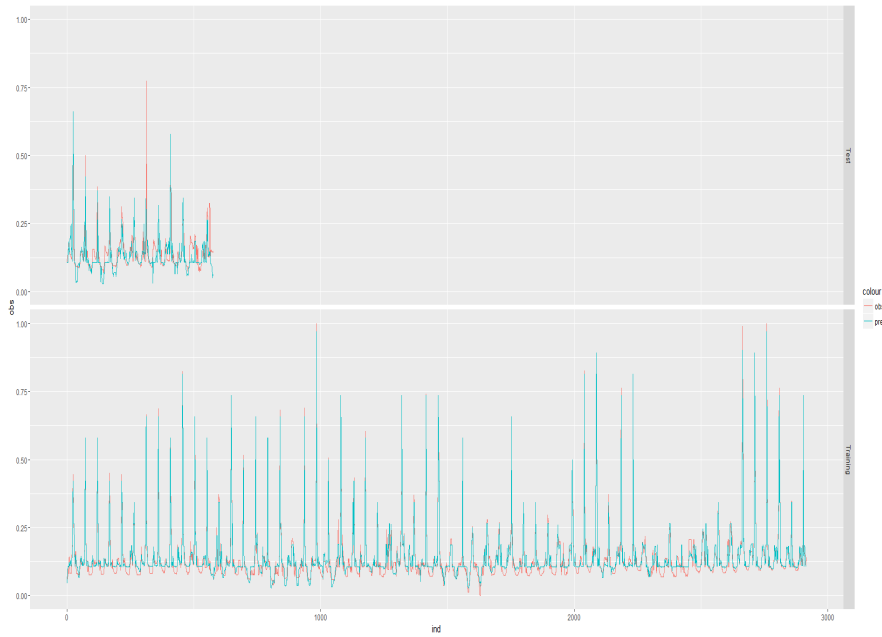

Figure 15. Multi-Layer Perceptron, with multiple layers
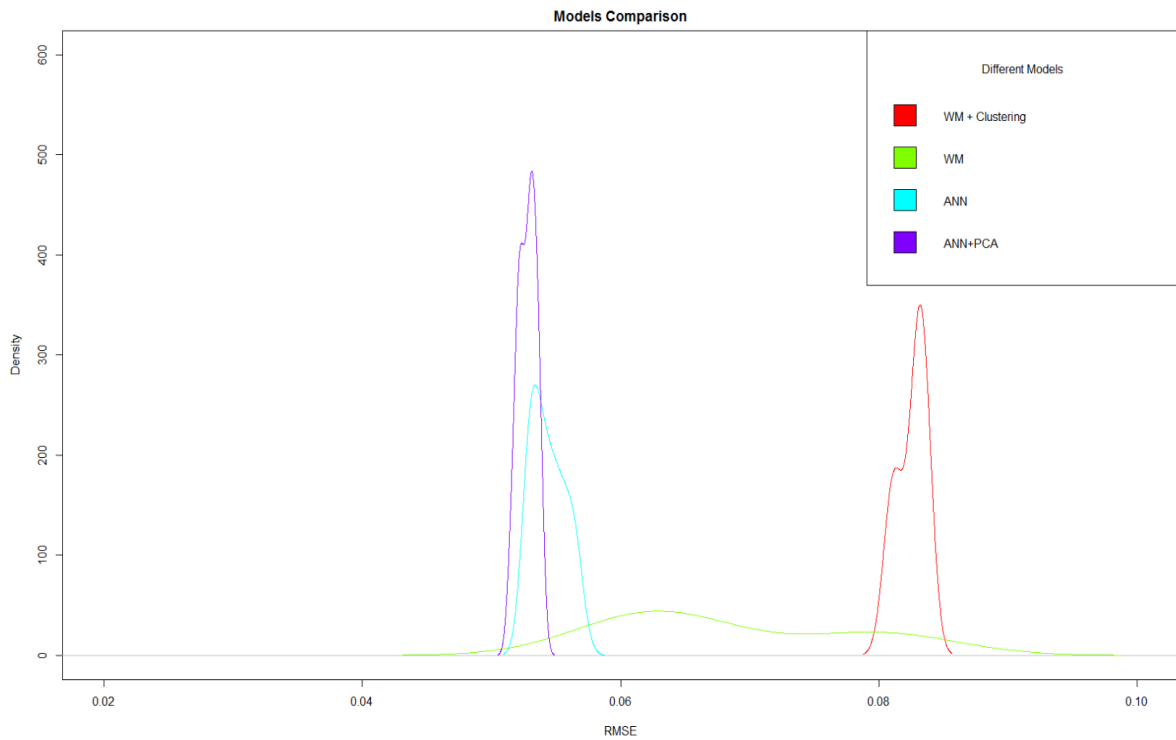
Figure 16. Wang and Mendel Fuzzy Rules

Figure 17. Comparison of Models

5. CONCLUSION

Short-term load forecast is an essential part of electric power system planning and operation. Forecasted values of system load affect the decisions made for unit commitment and security assessment, which have a direct impact on operational costs and system security. Conventional regression methods are used by most power companies for load forecasting. However, due to the nonlinear relationship between load and factors affecting it, conventional methods are not sufficient enough to provide accurate load forecast or

to consider the seasonal variations of load. ANN-based load forecasting methods can deal with 24-hour-ahead load forecasting by using forecasted weather input variables, which can lead to high forecasting errors in case of rapid weather changes. 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.

This paper presents a novel neural network-based approach for short-term load forecasting that uses the correlated weather data for training, validating and testing of a neural network. Correlation analysis of weather data determines the input parameters of the neural networks. The suitability of the proposed approach is illustrated through an application to the actual load data of the Irish Electricity Market. We may also make use of Microsoft Azure Machine Learning and IBM APIs cognos for experimenting with different ML algorithms. Azure ML could be useful as quick and easy to export to web services and bring online with minimum fuss. It has a wide range of plug-ins. A problem however is that it is a black box and but still potentially useful for early stage quick tests.

REFERENCES

[1]   D. Park, "Electric load forecasting using an artificial neural network," *IEEE Trans. on Power Systems,* vol/issue: 6(2), pp. 412-449, 1991.
[2]   P. Li, *et al.*, "Using Artificial Neural Networks to predict short-term wholesale prices on the Irish Single Electricity Market," *ISSC 2016 - The 27th Irish Signals and Systems Conference, Derry, Northern Ireland*, 2016.
[3]   P. J. Werbos, "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences," 1975.
[4]   O. Mohammed, "Practical Experiences with an Adaptive Neural Network short-term load forecasting system," *IEEE/PES 1994 Winter Meeting, Paper # 94 210-5 PWRS,* 1994.
[5]   http://albasoluzioni.com
[6]   http://www.mkonline.com
[7]   Green, R., Vasilakos, N. (2011) the Long-Term Impact of Wind Power on Electricity Prices and Generating Capacity. Department of Economics discussion Paper 11-09. [Online] Available at: https://www.birmingham.ac.uk/Documents/college-social-sciences/business/economics/2010-papers/economics-papers-2011/economics-papers-2011/11-09.pdf
[8]   http://www.meteogroup.com
[9]   Z. Osman, *et al.*, "Neural network-based approach for short-term load forecasting," *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, Seattle, WA, pp. 1-8, 2009.
[10] L. Sharma, *et al.*, "Neural Network Based Approach for Short-term Load Forecasting," *International Journal of Science, Environment and Technology*, vol/issue: 1(5), pp. 416-424, 2012.
[11] M. Buhari and S. S. Adamu, "Short-Term Load Forecasting Using Artificial Neural Network," *Proceedings of the International Multi-conference of Engineers and Computer Scientists*, IMECS 2012, Hong Kong, vol. 1, 2012.
[12] J. Wu, *et al.*, "A Novel Nonparametric Regression Ensemble for Rainfall Forecasting Using Particle Swarm Optimization Technique Coupled with Artificial Neural Network," *6th International Symposium on Neural Networks,* 2009.
[13] A. Baliyan, *et al.*, "A Review of Short Term Load Forecasting using Artificial Neural Network Models," *Procedia Computer Science*, vol. 48, pp. 121-125, 2015.
[14] CER, "CER Factsheet on the Single Electricity Market, Commission for Energy Regulation," 2011. http://www.cer.ie/docs/000262/cer11075.pdf
[15] www.sem-o.com.
[16] W. McCulloch and W. Pitts, "A Logical Calculus of Ideas Immanent in Nervous Activity," *Bulletin of Mathematical Biophysics,* vol/issue: 5(4), pp. 115-133, 1943.
[17] P. Li, *et al.*, "Predicting short-term wholesale prices on the Irish Single Electricity Market with Artificial Neural Networks," *ISSC 2017 - Irish Signals and Systems Conference, Killarney Co. Kerry*, 2017.
[18] https://en.wikipedia.org/wiki/Artificial_neural_network
[19] F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," *Psychological Review*, vol/issue: 65(6), pp. 386-408, 1958.
[20] M. Forouzanfar, *et al.*, "Comparison of Feed-Forward Neural Network Training Algorithms for Oscillometric Blood Pressure Estimation," *4th Int. Workshop Soft Computing Applications*. Arad, Romania: IEEE, 2010.
[21] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85-117, 2015.
[22] A. Graves, *et al.*, "A Novel Connectionist System for Improved Unconstrained Handwriting Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol/issue: 31(5), 2009.

[23] A. Graves and J. Schmidhuber, "Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks," in Y. Bengio, *et al.*, *Advances in Neural Information Processing Systems 22 (NIPS'22), Vancouver, BC*, Neural Information Processing Systems (NIPS) Foundation, pp. 545-552, 2009.

[24] Y. Rui and A. El-Keib, "A review of ANN-based Short-term load forecasting models," *Proc. of 27ᵗʰ South Eastern Symposium on System Theory, Starkville, MS*, pp. 78-82, 1995.

[25] A. Papalexopoulos, "Application of neural network technology to short-term system load forecasting," *Proceedings IEEE Second International Forum on the Applications of Neural Networks to Power Systems,* Yokohoma, Japan, pp. 796-800, 1993.

[26] K. Lee, "Short-Term Load Forecasting Using an Artificial neural Network," *IEEE Trans. On Power Systems,* vol/issue: 7(1), pp. 124-131, 1992.

[27] D. Srinivasan, "A neural network short-term load forecaster," *Electric Power Research*, vol. 28, pp. 227-234, 1994.

[28] Y. Lee, "An Analysis of Premature Saturation index Back Propagation Learning," *Neural Networks,* vol. 6, pp. 719-728, 1993.

[29] D. Ciresan, *et al.*, "Multi-column Deep Neural Networks for Image Classification," *IEEE Conf. on Computer Vision and Pattern Recognition CVPR*, 2012.

[30] R. Balabin and E. Lomakina, "Neural network approach to quantum-chemistry data: Accurate prediction of density functional theory energies," *J. Chem. Phys.* Vol/issue: 131(7), pp. 74-104, 2009.

[31] L. Wang and J. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, pp. 1414-1427, 1992.

[32] M. Brown and C. Harris, "Neurofuzzy Adaptive Modeling and Control," Upper Saddle River, NJ: Prentice-Hall, 1994.

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

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