2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

A Dynamic Forecasting Method for Small Scale Residential Electrical Demand

Andrei Marinescu, Ivana Dusparic, Colin Harris, Vinny Cahill, Siobhán Clarke
Distributed Systems Group, School of Computer Science and Statistics
Trinity College Dublin, Ireland
{marinesa, ivana.dusparic, colin.harris, vinny.cahill, siobhan.clarke}@scss.tcd.ie

***Abstract*—Small scale electrical demand forecasting is an emerging field motivated by the penetration of renewable energy sources and the growth of microgrids and virtual power plants. These advances pose more complex forecasting challenges compared to the already established large scale forecasting approaches.**

Current short term load forecasting methods deal with two types of day, normal and anomalous, which are predicted separately. Anomalous days are classified as such ahead of time, based on key calendar events such as public holidays. However, there are some anomalous days which are not always predictable on a day ahead basis. Due to unforeseen events, a seemingly normal day can progress towards an anomalous case causing high errors in prediction. We propose a new dynamic forecasting mechanism that actively monitors residential electrical demand along a forecasted day, and detects anomalous pattern changes from a previously predicted demand of the day. A self-organising map is employed to detect anomalous days as they progress. Once an anomaly is detected, a neural network based prediction system changes its input neurons according to a previously detected and recorded match found in a database of anomalous days, in order to accommodate the anomalous day prediction.

Results are based on measured power demands recorded in Ireland from domestic smart-meters between 2009-2011, and focus on small scale residential electrical demands of up to 350 kWh. During anomalous days our dynamic prediction approach achieves forecasting results within 3.63% of the real load, down from the 7.37% obtained by the initial prediction algorithm and the 5.41% achieved by standalone re-prediction, without pattern matching.

I. Introduction

Neural networks have been traditionally used in time-series forecasting, and are in particular some of the most commonly found techniques in short term electrical demand estimation [1], [2]. Utilities rely on it for applications such as generator scheduling and electricity market operations. While large scale forecasting is an established field, changes in the structure of the electrical grid pose new challenges for prediction methods. In the attempt to move towards a smarter grid, architectural modifications advance the concept of a distributed electrical network. The system should integrate renewable sources of energy while at the same time supply an increasing demand for electrified appliances such as electric vehicles. As a result, small scale units such as microgrids or virtual power plants (VPPs) come into play as quasi-autonomous entities interconnected in the smart grid. These units have their own sources of energy and can participate in electricity market operations, as well as operate by themselves autonomously, if required by the main grid. For each case, they have to be able to estimate in advance their own demand. Both microgrids and VPPs are aggregates of various sources of electricity generation, storage and consumption. Due to having unsteady supply from renewable sources, these units depend on the electrical demand of their own users for efficient operations.

In recent years, small scale short term load forecasting (STLF) has become of significant interest to microgrids thanks to the previously mentioned penetration of renewable sources of energy. In general, the owners of microgrids try to maximize the use of renewable energy and minimize the overall electricity costs, while maintaining user comfort. We believe that these constraints can be better matched by the use of accurate small scale load forecasting techniques.

One fact worth noting is that small scale demand raises more obstacles in the way of forecasting, mainly through frequent power demand changes and therefore less smooth demand, with more variability in user behaviour when compared to large scale. This is due to the considerable impact of individual users and their irregular behaviour pattern upon the overall power consumption, as observed in recent work investigating the effect of scale on forecasting algorithms [3], [4].

The following parts of this paper are organised as follows: Section II presents the current research focusing on short term load forecasting, with particular interest in those dealing with small scale power systems. Section III describes our own approach to small scale load forecasting with regard to the more general normal days and also with special focus on unanticipated anomalous days, which is an issue that hasn't been addressed before. Section IV shows our results obtained when testing the dynamic forecasting algorithm on both normal and unanticipated anomalous days. Finally, Section V presents our conclusion with regard to the obtained results, and future work based on short term load forecasting on small scale.

II. Background and Related Work

Small scale forecasting has gained more interest in the last decade, through the emergence of microgrids. There are a number of different techniques applied in electrical load forecasting with focus on large scale [1], [5]–[8]. Most of the methods consider weather information and current day of the week as important inputs in the prediction mechanisms. Several techniques of those applied on a large scale have also proven themselves successful in small scale, with artificial

978-1-4799-1484-5/14/$31.00 ©2014 IEEE

neural networks (ANNs), auto-regressive integrated moving average (ARIMA) methods and neuro-fuzzy networks achieving reasonable accuracy in tests. Results have shown prediction accuracy that goes up to 5% mean absolute percentile error (MAPE), according to the state of the art research [4], [9]–[11].

Most successful techniques combine several methods to improve overall results. Among them, one of the most noticeable and effective additions is the classification of daily demands into different sets, generally accomplished with the help of self-organising maps (SOM) [3], [12]–[14]. These classification methods split the forecasted days into separate classes, such as weekdays, weekends, and holidays. Classification is performed based on normal days and anomalous days. Public holidays and sometimes days close to them tend to be seen as anomalous days in most of the research, and are treated separately as such [15]–[17].

Although classification techniques have brought significant increases in forecasting accuracy when compared to more general approaches, anomalous days are still considered based on calendar events. Public holidays are seen as anomalous days to start with, and the otherwise normal days around them are considered anomalous as well in some of the cases (for example when they are the bridge between a weekend and a public holiday). We believe that a purely calendar based approach doesn't deal with real anomalies caused by unforeseen events outside of the calendar and weather range. Some of these events could be electrical grid malfunctions, unexpected climate phenomena or natural disasters. A true smart grid should be able to mitigate such issues, as well as be able to dynamically adapt its previously forecasted demand estimates. Of course, this has to be done with regard to short term load forecasting, where generally a demand estimate is made with a 24 hours ahead basis. The problem is that, most probably, such anomalies will be detected as a forecasted day progresses.

III. Algorithm Design

Our previous work has dealt with short term load forecasting on a small scale, through a hybrid approach that combines neural networks, wavelet smoothing, neuro-fuzzy networks and ARIMA techniques, with accuracy results that surpass the ones of the individual methods involved. The tests were performed over the aggregated measured power demand of 230 residential households from Ireland, recorded during a smart-meter trial which took place between 2009 and 2011. The hybrid method obtained forecasting results of 2.39% normalised root mean square error (NRMSE) during a testing period of 4 consecutive weeks involving only weekdays [18]. However, the forecasting algorithm was evaluated over a time-frame that included only normal days, with anomalous days being considered a separate case.

In this paper we present a new technique, motivated by state of the art research involving load forecasting and SOMs. The technique detects anomalous power usage behaviours on the fly and triggers an appropriate re-prediction mechanism,

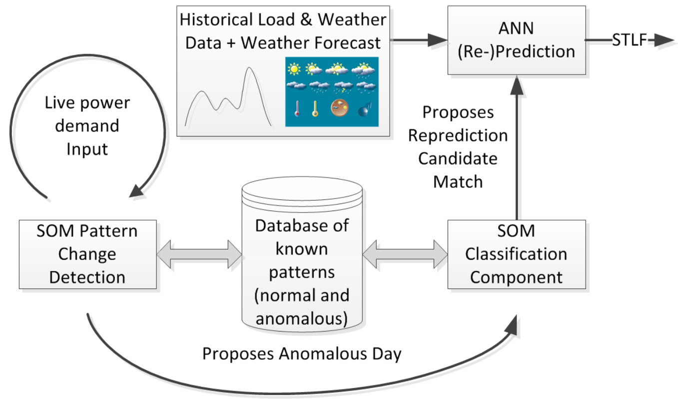

Fig. 1: Anomaly Detection and Re-prediction

as pictured in Fig. 1. Short term load forecasting generally makes estimates ahead between 2 weeks and 24 hours. We believe that some anomalies cannot be anticipated within a day ahead. These can occur as the day progresses, and actions taken at the point of anomaly detection can be critical for the optimal operation of the microgrid. Unlike the calendar-based approaches of the state-of-the-art forecasting methods, we are dealing here with unanticipated anomalous days. Our pattern change detection component continuously monitors power demand during a day to detect if it becomes anomalous, considering that it hasn't been marked as anomalous previously. As a seemingly normal day progresses and anomalous power demands occur, the pattern change detection (PCD) mechanism detects changes from the expected behaviour. Once the type of change is detected, the PCD proposes the reprediction of the demand based on a similarly previously encountered pattern found by the SOM component.

A. *SOM Classification Component*

An anomalous day (24 hour time frame) is normally detected with 100% accuracy only once it has ended, as anomalies can occur even at the end of the day (which would be close to the evening demand peak - a critical point). When detected as such, these anomalous days can only be of use at a further date, which is the case at the moment in state of the art forecasting approaches. Our approach employs self-organising maps for classification and pattern change detection of anomalous days before the day reaches its end, and more importantly before the critical evening peak.

SOMs tend to group similar samples into clusters (also known as classes). Initial analysis involving a large number of classes provided disparate clusters due to the relatively low number of , where grouping of public holidays and anomalous days was scattered across the map. Bringing the number of classes down to 4 led to grouping all Irish bank holidays into a single class, and non-calendar based anomalous days into another class. The remaining two classes comprise only normal days. Therefore, we have further employed a self-organizing map with 4 classes (1a, 1b, 2a, 2b).

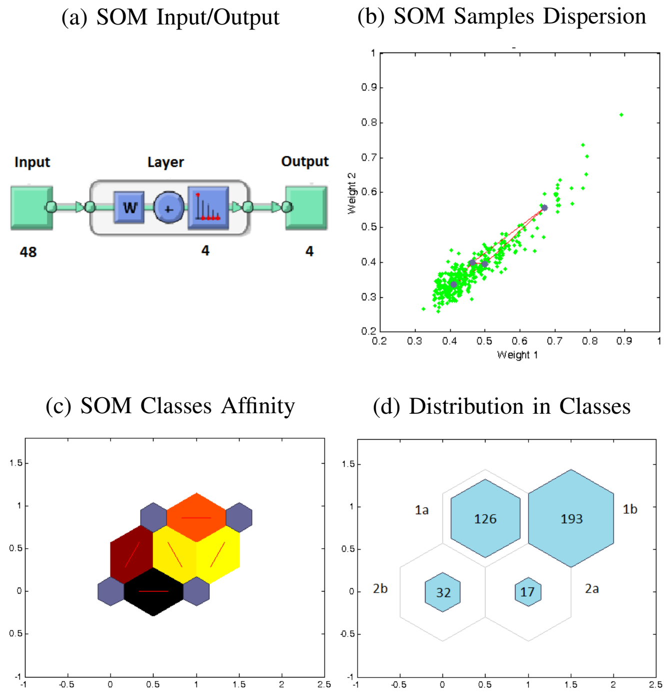

Fig. 2: Self-Organising Map

The input layer of the SOM contains 48 half-hourly measures of power demand, that is one sample of power demand recorded every half an hour over the course of a 24 hour period (one day), as can be observed in Fig. 2a. Once the dispersion of the samples is established, they are allocated into the four different classes based on their similarity Fig. 2b. The four classes share some similarities between themselves. When looking at Figure 2c, we can see the four different classes pictured with blue hexagons. The more similar features between two classes, the lighter the colour representing the connection between them (elongated hexagons). The total number selected for each class is shown in Fig. 2d.

Upon inspection, based on the historical load input, we have observed that the SOM has further clustered samples into: 1a) normal days with higher power demands (cold season); 1b) normal days with lower power demands (warm season); 2a) anomalous days occurring during bank holidays; and 2b) anomalous days outside calendar based events.

The SOM classifies all the occurring 13 bank holidays in the given researched interval (between 01-08-2009 and 31-12-2010) into a single class, together with a few more days around them, particularly in the Christmas/New Years period. The most detached class is the one containing only anomalous days, outside of the holidays range (that cannot be explained by public holidays or proximity to these holidays), which sum up to 32 days. As noted in the result section, while some anomalous days occur in a year (Nov-Dec 2010), they don't occur in the other (Nov-Dec 2009). Part of the anomalous days are around holidays as well, but they have a particular shape, as some people take days off and some don't, resulting in a unique demand pattern. The rest of the days in the given period are split between the other two classes depending on seasonality, as seen by the differences between summer and winter. The lower the daily temperature, the higher the demand, as many household heating units rely on electricity in Ireland. During summertime, as temperature doesn't rise enough to create discomfort among household users, HVAC systems are not employed and as a result we have generally lower power demand patterns.

B. SOM Pattern Change Detection

To deal with anomalous days on the fly, we have developed a pattern change detection system to provide us with valuable information about the state of the day (normal or anomalous). State of the art forecasting approaches that deal with anomalous day prediction employ just a classification component that post-processes the already passed day. Their assumption is that the pattern of the passed day can be used at a further date, possibly same time next year (e.g. for Christmas day).

We considered that the SOM detects anomalous days with 100% accuracy in post-processing mode, when all the days have ended. Based on previous results, we performed experiments to detect when a trade-off would be sufficient so that the anomaly detection algorithm reacts faster, with satisfactory accuracy. The consideration was that at about 50% anomaly detection rate it's more likely that a day is anomalous than not. For this purpose we have initially computed an average demand shape over 24 hours, with samples taken every half hour, represented in Fig. 3 by the blue curve. This average demand is based on all the available historical data from the smart-meter trial. Note that this shape is much smoother compared to real day values, pictured in the same figure by a randomly chosen normal day, highlighted here with red, due to the averaging process.

According to the SOM, the average shape fits in the upper two classes (1a and 1b), the ones of normal days. We have then replaced each value in the 48 element vector representing the average shape with the ones obtained from the day in progress, starting from midnight. For example, if the hour is 06:10, we select 12 values from the actual day from midnight up to and

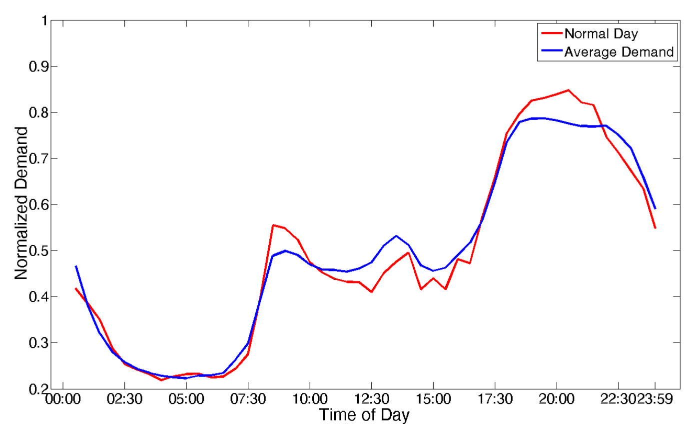

Fig. 3: Average Shape

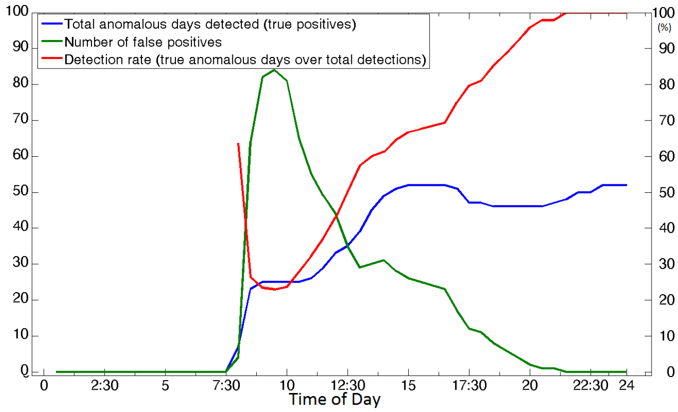

Fig. 4: Detection rates.
Left axis represents total samples, right axis accuracy rates

including 06:00, followed by 36 samples of the average day representing values between 6:30 up to and including 23:30.

As results show, possible anomalies start to be detected from 8:00AM on. If we look at Fig. 4, we can observe the false positive anomalies reported by the SOM, pictured here with a green line. The blue line marks the actual anomalies (true positives) out of the total possible proposed, as determined by the SOM at the end of the day. Both the green and blue lines total number of days are represented on the Y axis on the left. The red line shows the total accuracy rate, that is the ratio between the possible anomalies detected which turn out to be real anomalies over the total number of proposed anomalies, including false positives. Its values are represented on the Y axis pictured on the right side of Fig. 4. If we follow the red curve, we can notice that in the early morning period (midnight-8AM) a lot of anomalous days (about 10 out of 50) are detected quite accurately (40% accuracy), due to their particularly different demand patterns over the beginning of the day. As the time progresses a lot more false positives appear, due to the seasonality factor, and the overall accuracy decreases because of the inclusion of these false detections.

The total number of false positives detected starts to significantly decrease at 11:00, and at 12:30 we can see that its representation in the figure, the green line, intersects the blue line (true positives), meaning that they both share the same value. At this point we have reached 50% accuracy in detection, our target. Note that only four and a half hours were required to reach this level of accuracy, as during night time the demand tends to be insignificant, thus no anomalies occurred in our search interval. While this SOM method sacrifices accuracy for the morning peak, it prepares the forecasting mechanism for the most critical part of the day, the evening peak, where the highest demand occurs.

Under specific conditions re-prediction can be accomplished just by employing a SOM anomaly detection algorithm that overlaps values on top of the average shape from midnight up to 12:30, in order to trigger the re-prediction in case it detects anomalies in demand. Even though at this stage an anomaly is detected with 50% accuracy, the SOM requires another 2 hours of demand to properly match the type of anomaly, after looking into previously encountered anomalous patterns. Another observation is that, in our samples, the afternoon/evening period (14:30-23:59) accounts for more than half of the total energy consumption, more precisely 55% based on the same results which we have employed for the average demand shape in Fig. 3. Given that, at 14:30 we already have 65% detection accuracy, with 98% of the true positives detected, along with a few more false positives. However, as seen later in the results section, even if we re-predict the false positives, this will not effect the forecasting accuracy much.

At 14:30 we look for a similarly encountered demand in our database of anomalous days with regard to the demand. Once the closest match is found, the demand obtained so far and the rest of the demand belonging to the closest match are fed into the re-prediction system.

If the possibility of an anomaly is detected after 5 hours of monitoring, re-prediction occurs. Once the type of anomaly is detected (based on closest pattern match and requiring another 2 hours for increased accuracy), the reprediction mechanism proposes the new demand estimate.

Our pattern change detection mechanism requires only 5 hours from the beginning of the day in order to be able to detect anomalous days with a 50% accuracy rate, which we regard as important enough to consider the re-evaluation of prediction for the day in question. Also, by detecting an anomalous day in the morning we are able to prepare the microgrid in advance with the change in expected demand for the mentioned critical evening peak, by far the period of highest power demand during the day.

C. ANN Prediction and Re-prediction

In our previous work we evaluated several techniques and combined them in an adaptive hybrid method for increased accuracy [18], [19]. The mentioned approach is more computationally intensive than it's subcomponents and requires several consecutive normal days to accurately forecast a following normal day. For the purpose of this work, where dealing with anomalous days, we have selected for forecasting only one of these techniques, an improved version of the neural network component presented in [19]. ANNs don't require additional learning once trained, thus forecasting is instantaneous when input data is provided, a useful feature in critical applications. The component involves a multilayer perceptron artificial neural network (ANN) trained through resilient backpropagation. The implementation was accomplished by using the Fast Neural Network Toolbox [20], which is an open source software tailored to our own needs.

We have noticed patterns of overfitting in our previous work when it came to some particular days, therefore we decided to reduce the total number of neurons in the ANN. While previously we used 55 neurons for input, we decreased this to 43 to compensate the overfitting issue, as described by Fig. 5. They are further divided as follows:

- 24 neurons are used for previous load input (one for each hour)
- 5 neurons (down from 7) are used for the day code input, since we deal only with weekdays
- 14 neurons (down from 24) are used for weather forecast input along the day, 8 for temperature and 6 for humidity; more inputs were chosen for temperature as it is more relevant than humidity according to our correlation tests

The neurons in the input layer considered for load, temperature, and humidity use extrapolated values from samples taken during a whole day, depending on their correlation level with the load. The neurons corresponding to the former load each use two consecutive half-hourly values averaged to represent the demand over one hour; the neurons corresponding to the temperature average 3 consecutive hourly values and use the result as input for each neuron; and the neurons corresponding to humidity average 4 consecutive hourly values and use the result as input for each, since they are less relevant than the load and temperature neurons.

Another change from the initial ANN described in the previous work occurs in the training process. Due to the small number of anomalous days in our sample set, we have artificially increased the training set three-fold by adding small random variations to the demand for each real day recorded. The validation and testing set were not affected by this measure, and comprise only actual recorded demands.

The output layer totals a number of 24 neurons, which represent the short term load forecast. Each of the output neurons provides a demand estimate corresponding to the equivalent hour of the day.

Each prediction is based on the weather forecast for the day it attempts to predict, together with the historical recorded demand occurring over the same day of the previous week. The reason behind choosing the load that belongs to the same day of the week for prediction (e.g. previous Tuesday for predicting the following Tuesday) is that each day of the week tends to have a particular demand shape in the samples collected by us from the aggregate of residential users.

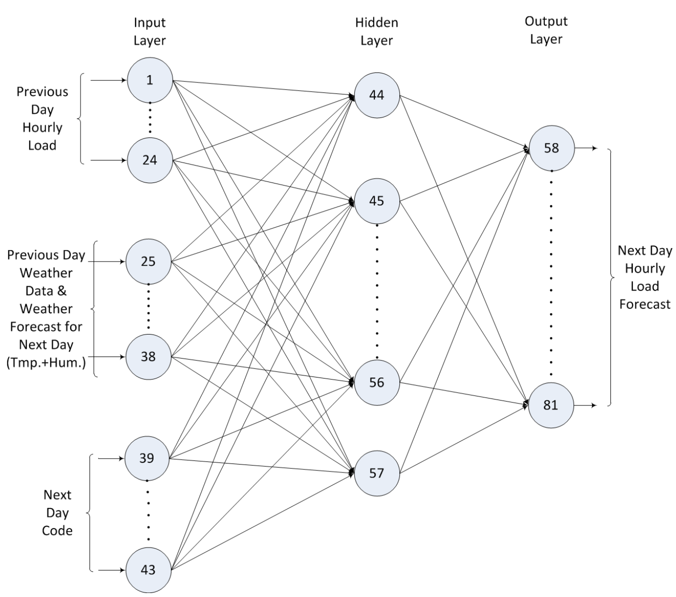

Fig. 5: Neural Network Structure

For the re-prediction mechanism component, the input part with regard to the historical load changes to accommodate the shape provided by the pattern change detection and matching mechanism. Since the network doesn't require any more training at this stage, re-prediction is instantaneous once the network is provided with the appropriate anomaly matching input from the SOM. In the case of re-prediction, the first 14 neurons are substituted with the values obtained from the day in progress up to hour 14, while the last 10 are based on the closest fit found by the pattern matching mechanism.

IV. Results

In order to test the dynamic forecasting algorithm we have employed historical electrical demand recorded during a smart meter trial, which was organized by the Commission of Energy Regulation in Ireland [21]. The trial comprises anonymised smart-meter data from both residential and commercial users. For our microgrid/VPP scenario we have employed only demand from residential users, aggregating the demand from 230 random households in order to roughly approximate the demand that goes through a 630 KVA transformer in an urban area. This is based on a power factor of 0.85 and an over-sizing factor of 0.5-0.6 for the transformer, considering that the aggregate demand of the 230 houses peaks at about 350 kW.

The trial recorded half hourly demand over a period of 17 months, during 2009-2010. In order to adjust the prediction algorithm to seasonality, historical Irish weather information was also involved, as provided by OGIMET [22]. Our evaluation only considers weekdays, since weekends are a special case with different daily behaviour from weekdays, and lower peak demands. Weekdays also have a larger sample dataset available compared to the weekends, and their power demand is also greater than the one of weekends when it comes to the critical evening peaks. The technique would be applicable on weekends too, but would require a separately trained ANN and possibly more samples.

In total we have used 370 weekdays, out of a sample of 517 days (weekends and weekdays). Furthermore, the training and validation sets involve data recorded over one year (260 weekdays), to account for seasonality. Tests were performed on the remaining 5 months (110 weekdays), due to the time limitations of the recorded data available. The ANN and SOM use a normalised version of the actual demand and weather information.

Forecasting accuracy rates are calculated in normalized root mean square error (NRMSE), RMSE and MAPE, and are computed based on the formulas in Eq. 1.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n}}$$

$$NRMSE = \frac{RMSE}{x_{max} - x_{min}} \quad (1)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i - \hat{x}_i|}{x_i}$$

,where x_i the actual value, $\hat{x}_i$ the forecasted value, x_{max}, x_{min} the maximum and minimum values from the tested set.

The tests were split into two different categories, mainly testing the accuracy of prediction during normal days and the special case with regard to prediction of anomalous days.

Testing the ANN prediction algorithm over normal days provided us with an accuracy that was quite close to the one of our hybrid approach, although not surpassing it. Over the same testing period of 20 consecutive weekdays (months of August and September), without any anomalous days involved, we have obtained an accuracy of 2.69% Normalised Root Mean Square Error (NRMSE), or 4.83% in Mean Absolute Percentage Error (MAPE), as shown in Table I. For more details, predictions over 4 consecutive days are pictured in Fig. 6. The purpose of this work though is to employ the ANN approach only for the anomalous days, when triggering time critical re-prediction. This avoids the more computationally expensive hybrid approach calculations when given such short notice, such as a few seconds/minutes in the middle of the day.

While in other work results are presented in the non-normalised RMSE (e.g. [3], [4], [23]), we believe NRMSE is a better way to compare, as the power demand scale doesn't have to be exactly the same as in the other test cases when comparing accuracy rates. Therefore even offset power demands (generally represented on Y axis in graphs) would be comparable with other similarly evaluated demands, when normalised. Additionally, NRMSE poses a more realistic estimation of error when compared to MAPE at very small scale, the latter being very strict when it comes to deflections of the forecast from the true load at low periods of demand (low morning peaks) while not as strict at periods of high demand (evening peak). This is evidenced by the high discrepancy in MAPE (7.59% to 12.04%) from similar values of NRMSE

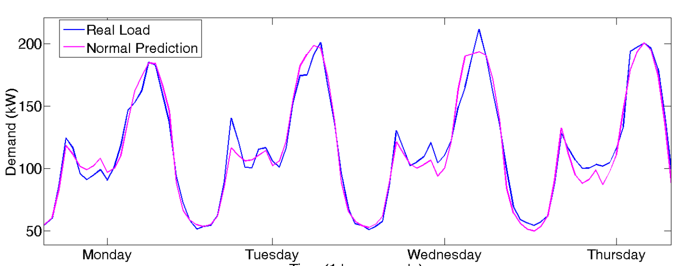

Fig. 6: Prediction over 4 consecutive days (Aug 2010)

TABLE I: Prediction Accuracy over 20 Consecutive Normal Days (Aug/Sept 2010)

Method	**RMSE (kW)**	**NRMSE (%)**	**MAPE (%)**
ANN	7.81	2.69	4.83
Hybrid	6.94	2.39	4.55

(7.31% to 7.37%) when comparing whole days from Table II versus the evening periods in Table III. As a result, we believe that NRMSE is a more evenly distributed way of evaluating forecasting accuracy in very small scale and therefore we use it in the rest of our evaluation section.

Some detailed results are presented in Table II, where we can see prediction accuracies during normal days versus anomalous days. These occur in the same time of the year. Fig. 7 visualises the forecasting accuracy obtained for each day, between the 10th of November and 29th of December 2010. Note the two periods of anomalies, one in the beginning of December due to possibly a very cold North Atlantic front. The second one is due to the Christmas holidays and New Year's eve. In the first anomalous period we can observe that generally the first five days have produced high errors in forecasting of up to 11% NRMSE in the standalone prediction. Standalone re-prediction (ANN+REP) tends to reduce this accuracy error after the anomaly detection, while the SOM enhanced re-prediction (ANN+SOM+REP) minimizes it to similar levels as the ones of normal days. After five days the normal algorithm adjusts itself and starts using the input of an anomalous day for the prediction of the 6th day. This can be noticed on the standalone prediction of Monday (the 6th of December) and Tuesday (the 7th), in Fig. 7.

The anomalous days evaluation was performed over a whole week (5 consecutive weekdays) occurring at the end of November and beginning of December. The pure ANN based prediction algorithm faced significant decreases in accuracy over that time, with values of 7.37% NRMSE, 12.04% MAPE, 22.18 kW RMSE, which is still an improvement over our closest related work from the state of the art [4]. While this might be considered somewhat satisfactory, we believe that the error is large enough compared to the average obtained by our normal prediction mechanism to consider the involvement of re-prediction through pattern matching techniques.

TABLE II: Prediction Error Normal vs. Anomalous Days (Nov/Dec 2010)

Method	**Normal Days**		**Anomalous Days**	
	NRMSE (%)	MAPE (%)	NRMSE (%)	MAPE (%)
ANN	3.03	5.03	7.37	12.04
ANN+REP	2.83	4.50	5.41	6.49
ANN+SOM+REP	2.81	4.36	3.63	4.71

TABLE III: Prediction Error Anomalous Days: 14:30-23:59 Interval (Nov/Dec 2010)

Method	**Anomalous Days**	
	NRMSE (%)	MAPE (%)
ANN	8.02	9.68
ANN+REP	7.31	7.59
ANN+SOM+REP	4.84	4.99

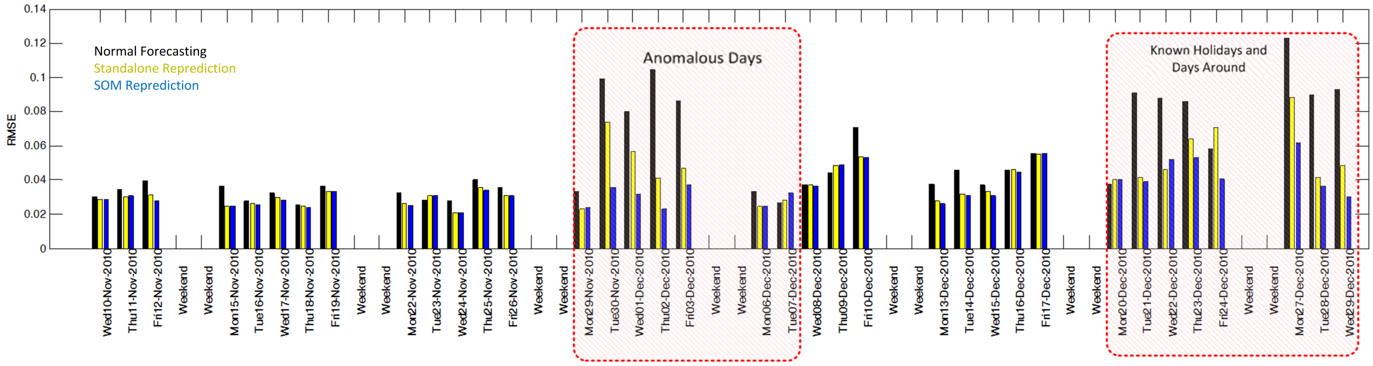

Fig. 7: Anomalous Month

As seen in Fig. 8, the real demand is somewhat higher than expected. This is actually not the only anomalous behaviour for the two days presented. Typical features of our aggregated residential demand include the two peaks that occur in the morning and early afternoon, before the very high evening peak. The morning peak is higher in the normal days, but here we can see a higher demand in the early afternoon peak, particularly in the second day (Friday).

The pattern change detection and matching mechanism triggered re-prediction over the five anomalous days in question, bringing down the RMSE to 10.89 kW, from the previous 22.18. (NRMSE 3.63%, MAPE 4.71%). More details can be seen in Table II, where we also present the results obtained by ANN+REP, prediction which is accomplished without the involvement of pattern matching techniques.

The more relevant testing period is occurring after the 14:30 interval though, which is the moment where the fitting re-prediction is triggered, and therefore it's the more appropriate one for comparison. The results obtained are shown in Table III. During this specific interval (14:30-23:59), the pure ANN prediction algorithm provides an accuracy of 9.68% MAPE (8.28% NRMSE) over the anomalous days. ANN+REP (re-prediction with pattern change detection enabled but without the pattern matching enhancements) achieved 7.59% MAPE (7.55% NRMSE) in the same given period, while the SOM enhanced re-prediction, ANN+SOM+REP, reached 4.99% MAPE (5.00% NRMSE). Some of the anomalous days together with the predictions are presented in Fig. 8.

It is worth noting that, while ANN+SOM+REP has the better accuracy out of the three evaluated forecasting methods (ANN, ANN+REP, ANN+SOM+REP), the errors that occur in its forecasting attempts also tend to be in general overestimates of the evening peak when compared to the other methods, which underestimate it. We believe that this is more important, as it is more relevant towards possible demands that could reach the transformer's capacity limits.

The power demand used for forecasting in our evaluation ranges between 40 kW and 340 kW (depending on the season), close both in demand patterns and power usage to the one presented in [4] at distribution substation level. As already mentioned, it is difficult to compare with the previous work because of the slight differences in scale. However, for illustration purposes, we have observed that they have at distribution substation level a demand which is between 100 and 300 kW [4]. This is according to the evaluation period presented in the graphs. This results in approximately 10.72% NRMSE, with their best considered result of 21.43 kW RMSE obtained through the auto-regressive (AR) model.

For our previously evaluated months, between August and September, the demand is actually between 40 kW and 220 kW. Our neural network forecasting algorithm provides a 7.81 kW RMSE over a period of 4 consecutive weeks (20 weekdays), in comparison to their best result of 21.43 kW RMSE (obtained with AR).

Another interesting test with regard to power demand level comparisons were forecasting evaluations performed during 2 consecutive weeks (10 weekdays) in November 2010, because it's a highly variable period due to the proximity of holidays. Here the power demands range between 50 kW and 340 kW. These demand values are close to the ones in the previously mentioned work [4]. The ANN forecasting algorithm developed by us provided a RMSE of **9.11** kW. This accuracy was reached despite the slightly wider range (subject to a higher RMSE) and the high variability in the given period. We motivate the variability of demand by the fact that exactly after this period of 10 consecutive weekdays we have several

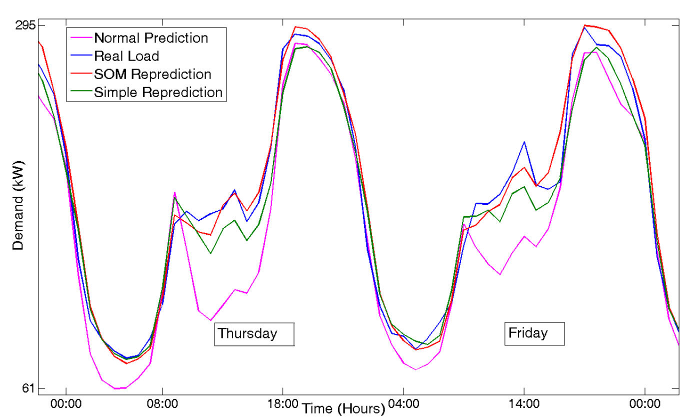

Fig. 8

days marked as anomalous by the SOM, with the weekdays in question being close or on the borderline of anomalous. In other systems of measurements, the results were of 5.03% MAPE, an insignificant decrease in accuracy when compared to the results obtained over the more settled summer period.

V. Conclusion and Future Steps

We have presented a new method for adaptive anomaly detection and re-prediction based on pattern matching techniques. We evaluated this method in the area of small scale residential electrical demand forecasting, a field of high interest due to the emergence of important smart grid actors such as microgrids and VPPs.

The results obtained by our method are better than state of the art approaches in small scale, as shown in the results section. Even more, as far as we know, this is the only forecasting approach in small scale that deals with normal days as well as anomalous days without classification on a predetermined basis, thus enabling on-the-fly anomaly detection, pattern matching, and re-prediction techniques in case of unanticipated anomalous days occurring. We believe that the electrical demand forecasting results achieved are very good at residential transformer level, which in our case is considered to be of up to 350 kW.

We plan to improve our classification techniques in order for them to be based on seasonality and day of the week patterns. For such improvements to occur, we need a larger dataset, one which spans several years, unlike our case which was limited to 17 months and therefore didn't allow too much tinkering in terms of SOM classes. Ultimately, we will connect our prediction techniques with demand response algorithms. For this purpose our future work will test demand response multi-agent systems at household and community level based on accurate power demand predictions, in order to suit the limits of the transformer providing power and also optimize the use of available renewable sources. Furthermore, intelligent learning techniques involving collaboration should help in fulfilling primary or critical objectives when combined to appropriate forecasting techniques. Some of our preliminary tests already point out the benefits of such multi-agent systems in demand shifting [24]. Overall, these developments should all contribute to increased stability of the power system and a lower carbon footprint through efficient use of renewables, reduced user costs, and optimal operation of critical systems.

Acknowledgment

This work was supported, in part, by Science Foundation Ireland grant 10/CE/I1855 to Lero - the Irish Software Engineering Research Centre (www.lero.ie)

References

[1] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: Literature survey and classification of methods," *International Journal of Systems Science*, vol. 33, no. 1, pp. 23–34, Jan. 2002.

[2] H. Hippert, C. Pedreira, and R. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Transactions on Power Systems*, no. 1, pp. 44–55.

[3] J. Llanos, S. Member, D. Sáez, S. Member, and A. Núñez, "Load Profile Generator and Load Forecasting for a Renewable Based Microgrid Using Self Organizing Maps and Neural Networks," *International Joint Conference on Neural Networks*, pp. 10–15, 2012.

[4] A. Tidemann, B. A. Høverstad, H. Langseth, and P. Öztürk, "Effects of scale on load prediction algorithms," *22nd International Conference on Electricity Distribution*, June 2013.

[5] M. Beccali, M. Cellura, V. Lo Brano, and a. Marvuglia, "Forecasting daily urban electric load profiles using artificial neural networks," *Energy Conversion and Management*, no. 18-19, pp. 2879–2900, Nov.

[6] J. W. Taylor and P. E. McSharry, "Short-term load forecasting methods: An evaluation based on european data," *Power Systems, IEEE Transactions on*, vol. 22, no. 4, pp. 2213–2219, 2007.

[7] A. Motamedi, H. Zareipour, S. Member, and W. D. Rosehart, "Electricity Price and Demand Forecasting in Smart Grids," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 664–674, 2012.

[8] N. Amjady, "Short-Term Bus Load Forecasting of Power Systems by a New Hybrid Method," *IEEE Transactions on Power Systems*, no. 1, pp. 333–341, Feb.

[9] M. F. I. Khamis, Z. Baharudin, N. H. Hamid, M. F. Abdullah, and S. Solahuddin, "Electricity Forecasting For Small Scale Power System Using Artificial Neural Network," no. June, pp. 6–7, 2011.

[10] M. Fatimie, Z. Baharudin, N. Hisham, M. Faris, and S. S. Yunus, "Electricity forecasting for small scale power system using fuzzy logic," *2010 Conference Proceedings IPEC*, pp. 1040–1045, Oct.

[11] N. Amjady and F. Keynia, "Short-Term Load Forecast of Microgrids by a New Bilevel Prediction Strategy," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 286–294, 2010.

[12] P. P. K. Chan, W.-C. Chen, W. W. Y. Ng, and D. S. Yeung, "Multiple classifier system for short term load forecast of Microgrid," *2011 International Conference on Machine Learning and Cybernetics*, pp. 1268–1273, Jul.

[13] C. Baladrón, J. M. Aguiar, B. Carro, A. Sánchez-Esguevillas, J. Lloret, J. J. Gomez-Sanz, and D. Cook, "A multi-agent system architecture for smart grid management and forecasting of energy demand in virtual power plants," *IEEE Communications Magazine*, p. 107, 2013.

[14] I. Koprinska, M. Rana, A. Troncoso, and F. Martínez-álvarez, "Combining pattern sequence similarity with neural networks for forecasting electricity demand time series," pp. 940–947, 2013.

[15] R. Lamedica, A. Prudenzi, M. Sforna, M. Caciotta, and V. O. Cencellli, "A neural network based technique for short-term forecasting of anomalous load periods," *Power Systems, IEEE Transactions on*, vol. 11, no. 4, pp. 1749–1756, 1996.

[16] G. Chicco, R. Napoli, and F. Piglione, "Load pattern clustering for short-term load forecasting of anomalous days," in *Power Tech Proceedings, 2001 IEEE Porto*, vol. 2. IEEE, 2001, pp. 6–pp.

[17] K.-H. Kim, H.-S. Youn, and Y.-C. Kang, "Short-term load forecasting for special days in anomalous load conditions using neural networks and fuzzy inference method," *Power Systems, IEEE Transactions on*, vol. 15, no. 2, pp. 559–565, 2000.

[18] A. Marinescu, I. Dusparic, C. Harris, S. Clarke, and V. Cahill, "A hybrid approach to very small scale electrical demand forecasting," *Innovative Smart Grid Technologies (ISGT)*, p. to appear, 2014.

[19] A. Marinescu, C. Harris, I. Dusparic, S. Clarke, and V. Cahill, "Residential electrical demand forecasting in very small scale: An evaluation of forecasting methods," *2nd International Workshop on Software Engineering Challenges for the Smart Grid*, pp. 25–32, May 2013.

[20] S. Nissen, "Implementation of a fast artificial neural network library (fann)," Department of Computer Science University of Copenhagen, Tech. Rep., 2003.

[21] Comission for Energy Regulation, Ireland. (2011) Smart meter trial data. [Online]. Available: http://www.ucd.ie/issda/data/commissionforenergyregulationcer/

[22] (2013, Dec.) Professional information about meteorological conditions in the world. [Online]. Available: http://www.ogimet.com

[23] M. Espinoza, C. Joye, R. Belmans, and B. D. Moor, "Short-Term Load Forecasting , Profile Identification , and Customer Segmentation : A Methodology Based on Periodic Time Series," vol. 20, no. 3, pp. 1622–1630, 2005.

[24] I. Dusparic, C. Harris, A. Marinescu, V. Cahill, and S. Clarke, "Multi-agent residential demand response based on load forecasting," in *Technologies for Sustainability (SusTech), 2013 1st IEEE Conference on*. IEEE, 2013, pp. 90–96.