Evaluation of Forecasting Methods for Very Small-Scale Networks

Jean Cavallo, Andrei Marinescu, Ivana Dusparic, and Siobhán Clarke

Distributed Systems Group, School of Computer Science and Statistics,
Trinity College Dublin, Ireland
jean.cavallo@centrale-marseille.fr
{marinesa,ivana.dusparic,siobhan.clarke}@scss.tcd.ie
http://www.dsg.cs.tcd.ie/

Abstract. Increased levels of electrification of home appliances, heating and transportation are bringing new challenges for the smart grid, as energy supply sources need to be managed more efficiently. In order to minimize production costs, reduce the impact on the environment, and optimize electricity pricing, producers need to be able to accurately estimate their customers' demand. As a result, forecasting electricity usage plays an important role in smart grids since it enables matching supply with demand, and thus minimize energy waste. Forecasting is becoming increasingly important in very small-scale power networks, also known as microgrids, as these systems should be able to operate autonomously, in islanded mode. The aim of this paper is to evaluate the efficiency of several forecasting methods in such very small networks. We evaluate artificial neural networks (ANN), wavelet neural networks (WNN), autoregressive moving-average (ARMA), multi-regression (MR) and autoregressive multi-regression (ARMR) on an aggregate of 30 houses, which emulates the demand of a rural isolated microgrid. Finally, we empirically show that for this problem ANN is the most efficient technique for predicting the following day's demand.

Keywords: Auto-Regressive Moving-Average (ARMA), Auto-Regressive Multi-Regression (ARMR), Artificial Neural Networks (ANN), Comparison of Statistical Methods, Forecasting, Microgrid, Multi-Regression (MR), Smart-Grid, Wavelet Neural Networks (WNN).

1 Introduction

As traditional power systems are transitioning towards smart grids, resilience and robustness requirements have created a demand for the decentralisation of the grid, where subcomponents of the network have to be able to operate autonomously, in islanded mode, whenever the grid requires. These subcomponents are known as microgrids, and comprise both power generation (e.g., combined-heat and power units, solar panels, wind turbines, fuel cells, diesel generators) and consumption units (e.g., factories, university campuses, households). When a microgrid is operating autonomously, it needs to be able to match its supply

and demand, therefore requiring an estimate of future energy usage. Moreover, by combining production of electricity with heat (e.g., combined heat and power units), producers would be able to optimize their process according to both demands. Microgrids pose new challenges for electrical energy forecasting due to their very small scale. Recent research in microgrids investigate forecasting in a group of 19 customers [25], or over groups of 90/230 households [16]. Even though current microgrids are mainly implemented in campus of universities [9], they can be expanded to small towns. In order to make towns self-sufficient in terms of energy, new specific infrastructure is being built and custom forecasting tools are being developed for this kind of networks. For example, main expectations in Ireland based on 2020 European energy targets are: 40% of electricity to be provided by renewable sources, 10% of transport energy to be fuelled through renewable sources, 10% penetration rate for electric vehicles and energy efficiency to be increased by 20% [20]. However, to optimize supply and maximize integration of renewable sources, the electric energy consumption for each city has to be predicted with high reliability and accuracy. Even though some solutions for energy storage exist [2], they still do not have expected and sufficient yield. The purpose of this paper is to provide an evaluation of forecasting methods in very small scale networks, where electric energy consumption forecasting is accomplished over 50 synthetic groups of 30 houses. The demand of these groups are based on smart-meter data recorded from a set of 900 Irish houses through a trial operated by the Irish Commission of Energy Regulation in 2011. The forecasting is done on a day-ahead basis, with different statistical methods being considered.

1.1 World electrical consumption repartition

The world's electric energy consumption has almost tripled[1] from 7,323 billion of KWh in 1980 to 19,710.36 in 2012. Moreover, as shown in Fig. 14, this increase is mainly due to the energy demand from developing countries in Asia. For BRIC's [2] demand, the evolution is country dependent. On the one hand, Russia's consumption stagnates thanks to their gas production. On the other hand, China has multiplied its energy usage by a factor of 17 due to strong industrial development. For developed countries, the energy use throughout the years tends to be more consistent. For example, Germany's demand has barely increased, whereas France's has doubled, and Ireland's has multiplied by a factor of three. As shown in the table 1 on the next page[3], the biggest recorded increases are often occurring in small countries. These are also sometimes underdeveloped countries or countries which have experienced widespread political instability. Hence, building directly efficient systems would limit the gap between them and developed countries.

[1] **See Fig. 14 on page 17**: `http://www.eia.gov`
[2] **BRIC**: Brazil, Russia, India and China
[3] **Source**: `http://www.eia.gov` and `http://data.worldbank.org/`

$N°$	Country	EC (b. KWh)	Relation 2012/1980	GDP in 2012 (b. US$)
1	Bhutan	1.642	82.59	1.823
2	Maldives	0.2671	71.80	2.113
3	Turks and Caicos Islands	0.1674	36.00	nonOECD
4	Cambodia	3.004	33.65	14.054
5	Vietnam	108.27	32.96	155.820
6	Oman	20.36	25.51	77.497
7	Bangladesh	41.52	22.64	133.355
8	Cape Verde	0.29	20.47	1.757
9	Macau	4.22	18.50	42.981
10	Jordan	13.9	17.33	31.015
11	United Arab Emirates	93.28	17.10	372.313
12	China	4,467.92	17.09	9,240.270

Fig. 1: Top 12 of electrical consumption increases between 1980 and 2012

1.2 Indicators impacting on electrical energy consumption

Electric energy usage is dependent on several factors. Along this study, only the weather aspect is treated, whereas the social aspects of each house are outside the scope of this paper.

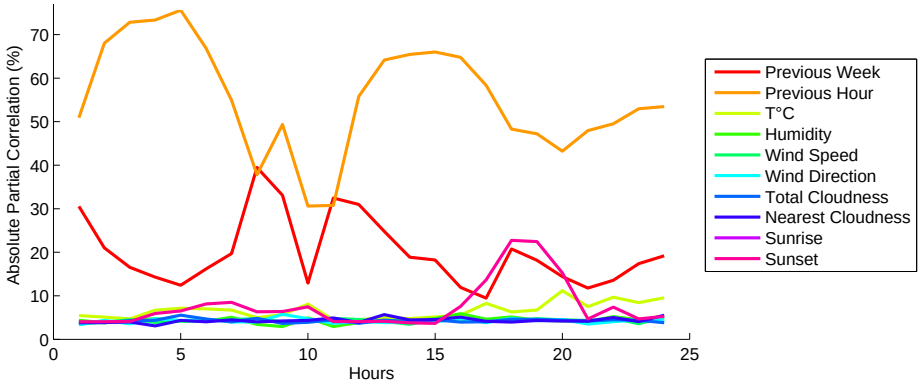

Fig. 2: Absolute partial auto-correlation function for a group of 30 houses

For each house of the forecast group, social elements impact differently on the energy demand (e.g, the social class, the size of the family and of the house, the number of rooms and the power and tariff established). However, energy demand depends also on weather conditions, and thus relies on generally available information, such as the forecast hour, the temperature along the day, the humidity, the cloudiness and the sunrise/sunset hours. Furthermore, the previous hour's consumption is impacting on future hours' demand, as shown in Fig. 2. As a side note, one of the most significant influencing factor for the overall

consumption of a country is the political stability, which has a real effect on the economical stability of the country. The long term industrial and household consumptions are significantly affected by political and social situations, whereas for short term they depend on factors such as energy consumption over the previous hours/days and weather conditions.

1.3 Networks scale and windows time

Most research on forecasting electrical energy usage has been done on large networks. At this scale, energy usage exhibits regular patterns. Below a certain number of houses, energy usage patterns tend to become more stochastic [12]. Such irregular patterns can be addressed through the use of artificial neural netwokrs (ANNs). ANNs are effective when used to solve non-linear problems. This is because ANNs minimize the impact of randomness on the forecasting accuracy. This is the main reason why ANNs are used in microgrid forecasting [8]. In this case, the day of the week and the month have each one two inputs. These inputs are calculated by using trigonometric functions. As shown in [8], the separation of patterns is essential when mitigating errors. As such, at least one input needs to be dedicated for the day of the week to be forecasted. The use of inputs for the month in [8] is likely due to a weak variation of temperature in Bilbao (Spain), the city where the experiments were performed. ANNs with multiple hidden layers have also been shown to be effective [7]. Nevertheless, there are some differences with the models implemented in this case: only the wind speed and the forecasted average humidity of the forecast day are considered along the input. It can be noted that the data set is not representing a microgrid but a wide area in Iran. Further work shows more focus on classical times-series forecast methods such as ARIMA, ARMA, Multi-Regression, Double Exponential Smoothing and Random Forest, where these algorithms are applied on 6 microgrids [23]. In these scenarios ARIMA is shown to be the most accurate method. There is further research where forecasting was implemented through ANNs [25], Auto-Regressive Integrated Moving-Average (ARIMA) [24, 27], multi-regression (MR) [17], wavelent neural networks (WNNs) [28]. These models have been applied on different scales: large, medium, small and very small [15, 25]. Finally, most of these methods forecast electricity usage on a day-ahead basis, which is classified as a short-term forecasting.

2 Background

2.1 Artificial Neural Network

A neural network comprises three type of layers: input layer, hidden layer(s) and output layer. One input corresponds to one feature and (in fully connected networks) is linked to each neuron belonging to the first hidden layer. The link weights the value of the neuron through a training process. In a neural network, there is a minimum of one hidden layer, and generally no more than three hidden

layers. Each layer's neurons are linked to the previous layer's neurons. The neural network goes through a training stage, where the input sets and desired output sets are provided, and where the weights for each link auto-configure throughout the training through a process known as backpropagation. Once the network is properly trained, it is able to provide by itself outputs that closely match desired outputs. The inputs of the neural network are correlated to the outputs, as these need to give an indication of the final results of the forecast.

Remark 1. Even if there is no mathematical demonstration, the rule of thumb is for the number of neuron in a hidden layer to be between the number of inputs and outputs. Other methodology has been proposed in the literature [11, 22]. According to this, when working with several hidden layers, results tend to be better when the layers have the same number of neurons.

For the forward propagation, each value of the previous layer has to be weighted and add up. Thus, Eq. (1a) allows to get the value of each hidden neurons, whereas Eq. (1b) gives the output values thanks to the same reasoning.

$$H_p = \sum_{n=1}^{N} \theta_{n,k}.I_n \tag{1a}$$

$$O_q = \sum_{p=1}^{P} \theta_{p,q}.H_p \tag{1b}$$

where θ_{ij} is the weight from $neuron_j$ of the $layer_J$ to $neuron_i$ of the $layer_I$

Eq. (1a) and Eq. (1b) characterize the forward propagation. However, as every statistical method, there are also equations to correct the model during the fitting. The most used algorithm - especially because of its calculation speed - is the Resilient backPROPagation (RPROP) [10, 21] . Backpropagation is very often used in supervised learning [21]. This is because it allows neural networks to adjust the weights of each neuron. This idea is characterized by the following equation:

$$\frac{\partial E}{\partial \theta_{ij}} = \frac{\partial E}{\partial s_i}.\frac{\partial s_i}{\partial net_i}.\frac{\partial net_i}{\partial \theta_{ij}} \tag{2}$$

After computing the partial derivative of an arbitrary error function E, each weight is corrected through an advanced optimization method, such as Broyden-Fletcher-Goldfarb-Shannon (BFGS) algorithm [4], or by a gradient descent:

$$\theta_{ij}(t+1) = \theta_{ij}(t) - \epsilon.\frac{\partial E}{\partial \theta_{ij}}(t) \tag{3}$$

2.2 Wavelet Neural Network

A Wavelet Neural Network (WNN) [1] is composed of two main parts [3]: the wavelet decomposition which splits the signal into several sub-signals according to high-pass and low-pass filters, and neural networks which are different

for each sub-signals. Two main decomposition algorithms derive from here: Discrete Wavelet Transform (DWT) [5] and Continuous Wavelet Transform (CWT). Many algorithms can be used for DWT: Haar Wavelet, Daubechies Discrete Wavelets, Gaussian Derivative or Battle-Lemarie [26]. In this paper, only the Daubechies Discrete Wavelets (DDW) technique is presented. The common equations which have to fit function ϕ to be a wavelet [26] are: zero-mean (4a), normalization (4b) and admissibility condition (4c).

$$\int_{-\infty}^{+\infty} \phi(x)\mathrm{d}x = 0 \tag{4a}$$

$$\int_{-\infty}^{+\infty} \phi^2(x)\mathrm{d}x = 1 \tag{4b}$$

$$0 < C_\phi < +\infty, \quad C_\phi = \int_0^{+\infty} \frac{|\Phi(u)|^2}{u}\mathrm{d}u \tag{4c}$$

Each wavelet is forecasting through a separate neural network, which is calibrated during the training period. Afterwards all the results are added.

2.3 ARMA

ARMA comprises of two distinct parts. The Auto-Regressive (AR) part is a linear combination of lags of the studied time-series. The Moving Average (MA) part is a linear combination of lags of the summarized series. An ARMA model can be summarized by the process in Eq. 5. This process takes into account seasonal aspects.

$$ARMA(p,q) : y_t = \mu + \sum_{i=1}^{p} \phi_i . y_{t-i} - \sum_{j=0}^{q} \theta_j . e_{t-j} \tag{5}$$

2.4 Multi-Regression

Multi-Regression attempts firstly to study the impact of external parameters which influence the signal to be predicted. If with these parameters an expert can predict the output, then the number of features is correct, otherwise, other features have to be added. After this step, the maximum polynomial degree for each feature and also the combination between each of them have to be considered. To avoid overfitting or underfitting, as described in [18], some parameters need to be computed, such as cost the function of the polynomial degree or the regularization parameter.

3 Methodology

3.1 Introduction

The data employed in the experiments is from the same database which was used in [15, 16]. 900 households are grouped into 30 households sets. These 900

households comprise recorded smart-meter data from mid-July 2009 to the end 2010. 50 groups are generated in order to validate the techniques over several very small-scale networks and be able to generalize the results of the experiment. The first particularity is the common origin of all the groups. The selection of best models are done with a unique set, which represents the latest 15 weeks of the initial sampling. The final evaluation between the most efficient version of each model is done along the year of 2010 with the same groups of houses. The data used is only for the week days, as these represent a wider sample and present highest demand patterns throughout the week [16]. As the used data sets do not involve computationally demanding operations, the time difference between C++ and Matlab programs was considered irrelevant. As such, Matlab was chosen to implement all models.

3.2 Artificial Neural Network

For the implementation of this method, two solutions are possible. To build the ANN, Matlab's Neural Network Toolbox can be employed. The free open source Fast Artificial Neural Network (FANN) library [19] allows to implement ANNs in several languages such as C, PHP, Python, Octave and C++. To see how it works, a Neural Network was built without GUI. Also, a free open source GUI called FANNTool[4] is readily available for users interested to work with Neural Networks. For this forecasting method, there are two different main types of models characterized by inputs, with particular focus on the first 24 inputs used. The first model, abbreviated as ANN(W,.) is quite similar to the one described in [16]. The 31 inputs are: the previous day's hourly consumptions (24 inputs), minimum, maximum and mean of future temperature and humidity (6 inputs) and the type of the day (1 input). The only difference from the other model is due to the allocation of only one input for the type of the day for this model instead of 7. There are 24 outputs, one for each hour of the next day's elctrical energy consumption, which is the forecasted information. The second variation of the model, abbreviated as ANN(D,.), uses the demand of the same day of the previous week.

3.3 Wavelet Neural Network

In this part, there are two different challenges. The first one is to find out the right Daubechies number. The second one is to use the same process described in 3.2. Based on a process inspired by methods presented in [1], the time-series represented by electrical demand is split into a subset of time-series. To ascertain the right Daubechies number, the process presented in detail in Fig. 15 was employed. This is further summarized in 3. Inputs are the same as described in [15, 16] (i.e., 24 hours consumption of same day of the previous week, temperature and humidity, and day type. Outputs are the 24 hours forecast of the following day).

[4] **Source**: https://code.google.com/p/fanntool/

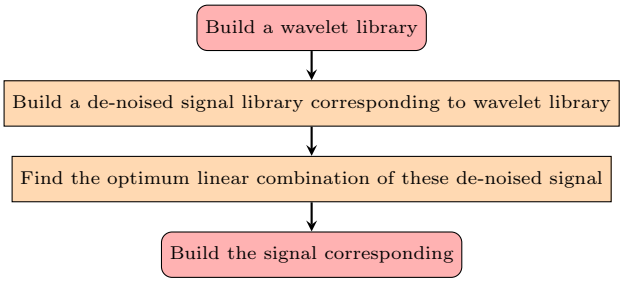

Fig. 3: Summary of process to build an optimal WNN

As described in literature [6], Daubechies number cannot be too large. The same observation is done with regard to the previous experiment. Often the right interval is between 4 and 8. That is why, the different configuration for Daubechies number are 4,5,6,7 and 8. Also, for each sub-signals, such as the approximation and detail signals, the combination of wavelets characterized by their Daubechies number can be different. To simplify the reasoning, the combination has been made with final signals. The decomposition-level is 3 for all WNN models.

3.4 ARMA

Because the studied data set is already stationary, only ARMA models are set up. ARMA is built by following two different types of fitting. The first one attempts to fit the whole training set. The second type is the same as a Time-Varying Auto-Regressive Moving-Average (TVARMA), because the model's parameters are constantly updated. It works as a sliding window concept, with a window following the time frame used in the training set. Several experiments need to be carried out to estimate the optimal window-time size. The selected models for in ARMA used forecasting the energy demand comprise windows size of 4,5,6,7 and 8 weeks before the testing step.

3.5 Multi-Regression

For Multi-Regression, the additional parameters taken into account besides historical information are temperature and the humidity, because of their considerable impacts on electrical energy consumption. This is shown in Fig. 2.

The equation characterizing Multi-Regression is:

$$\widehat{Y}_t = \sum_{i=0}^{n} \Theta_i . f\left(X_i^{(training)}\right) \tag{6}$$

The different sub-methods tested consider f as polynomial function, where the degree is ranging between a minimum of 1 to a maximum 5. Further tests

were done also with Gaussian and sigmoid functions. The model MR(f,nW) corresponds to the model with Multi-Regression method, taking into account the function f in Eq. 6 and nW weeks in its training set.

3.6 ARMA and Multi-Regression combined

Both models are combined to try to exploit each of their individual advantages. This method returns the average between the (ARMA(1,0,0),nW) model, where Weeks is varying according to the various models evaluated, and the different MR(f,nW) models. This is further referred to as ARMR(f,nW).

3.7 Surrounding forecasts using a frequency approach

In this part, the aim is to evaluate the risk taken by an user by evaluating the upper and lower boundaries. For that, the training errors are assimilated to a Gaussian distribution. Then, μ and σ are calculated and both sub-signals (upper and lower), built based on $\mu + 2.\sigma$ and $\mu - 2.\sigma$, are added to the forecast signal.

4 Experiments

4.1 Introduction

For the selection of the best models for each method, one half-season has been chosen because of the high variance of consumption. Indeed, this leads to an increased level of difficulty when attempting to accurately forecast the demand. The choice of only one period for this kind of selection is due to the small size of the available data set.

4.2 Dataset description

The initial dataset used comprises recordings from approximately 900 houses, starting from the 14th of July 2009 to the 31st of December 2010. For the purpose of evaluation, the total consumption of the 50 groups of 30 houses were considered. Thanks to the OGIMET website (www.ogimet.com), weather indicators for aviation were obtained under SYNOP code. A short function developed with Matlab has allowed to extract requested weather indicators. References for the weather station where data was measured are 03969 (WMO Index) and EIDW (ICAO Index).

Remark 5. To obtain humidity from dew-point temperature, the Heinrich Gustav Magnus-Tetens' formula was used. For some data - when temperature is under $10C$ -, $T_R > T$ - in the OGIMET file - that is implies $RH > 100\%$. To solve this kind of problem, another simple filter is set to ensure that the maximum bound is at 100%.

4.3 ANN Model Selection

Empirical tests show that the number of neurons in the hidden layer does not have a significant impact on the precision of ANNs methods with the same day in the previous week. However, both ANN(W,25) and ANN(W,26) seem to be the most accurate and reliable for this kind of models.

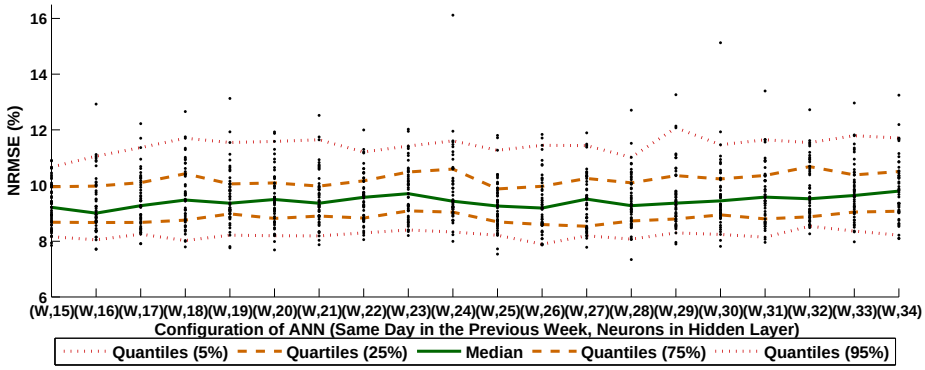

Fig. 4: Repartition of total NRMSE for each configuration - ANN - previous week

Fig. 4 and Fig. 5 emphasize the efficiency of retaining electric energy consumption of the previous day instead of electric energy consumption of the same day in the previous week. Indeed, it provides on average a 5% better median. Thus, the 5 models should be chosen among the second type of models. The contrast between accuracy and reliability for some models should be also noticed. For instance, the model ANN(D,23) is more accurate than ANN(D,20) according to the median, but it remains less reliable according to the dispersion: the 95% quantile illustrates a difference around 1.25% of total NRMSE.

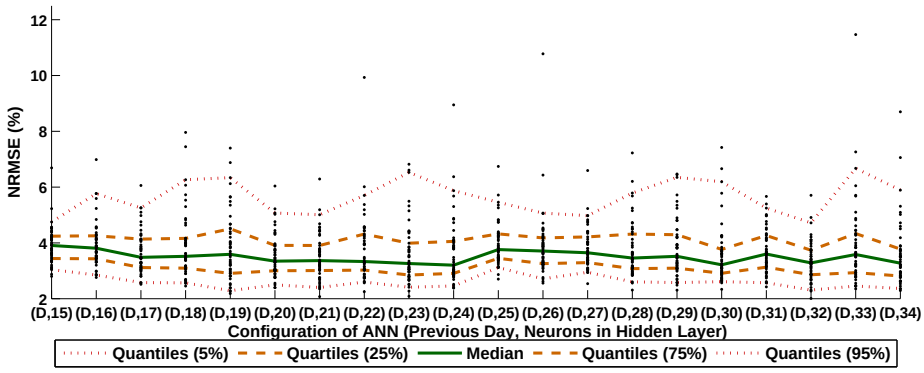

Fig. 5: Repartition of total NRMSE for each configuration - ANN - previous day

The consequences that can be drawn from evaluation are contrary to those found in [15, 16]] and this is in all likelihood due to size of the network studied. Actually, due to a high presence of residuals from forecasting errors, the more recent consumption is closer related to the network's future demand.

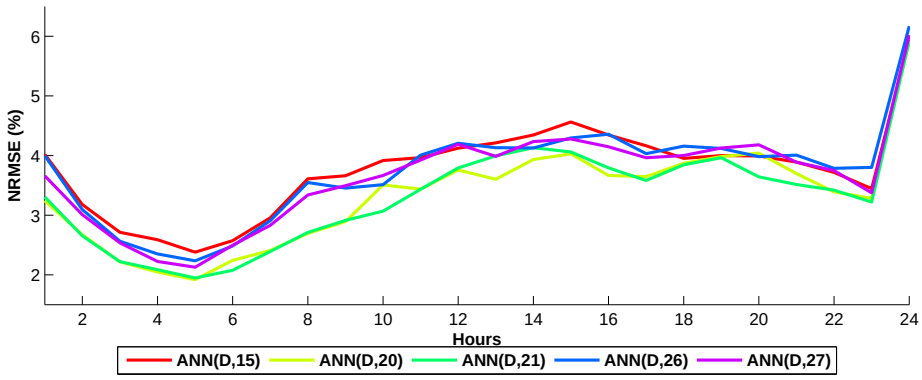

Fig. 6: Hourly NRMSE for 5 configurations - ANN

In order to refine the selection of the final method, Fig. 6 is plotted. Two methods, ANN(D,20) and ANN(D,21), have obtained best results. Finally, ANN(D,21) was selected because of its better reliability around the evening peak, which is a critical time when additional generating units can be commissioned to provide electricity. It is also because of this residual that small-scale networks forecasting is difficult.

4.4 WNN Model Selection

Fig. 17 shows that for Wavelet Neural Networks, the optimal number of neurons in the hidden layer is not necessary between the number of inputs and number of outputs, as suggested in [11, 22]. The Daubechies numbers do not have a high impact on the average NRMSE. For the previous day, as it can be seen in Fig. 16, the results are different: the Daubechies numbers affect the forecasting accuracy more than the number of neurons in the hidden layer. After a similar study to the one in Section 4.3, the model selected for WNNs is WNN(D,7,15). This model comprises 15 neurons in the hidden layer, a Daubechies number of 7 and the electric energy consumption of the previous day.

4.5 ARMA Model Selection

Due to the seasonality of the electric energy consumption, the model ARMA-NWM[5] has a very poor accuracy compared to other ARMA models using Windows-Time Moving, illustrated in Fig. 7.

[5] **ARMA-NWM:** ARMA No Window-Time Moving

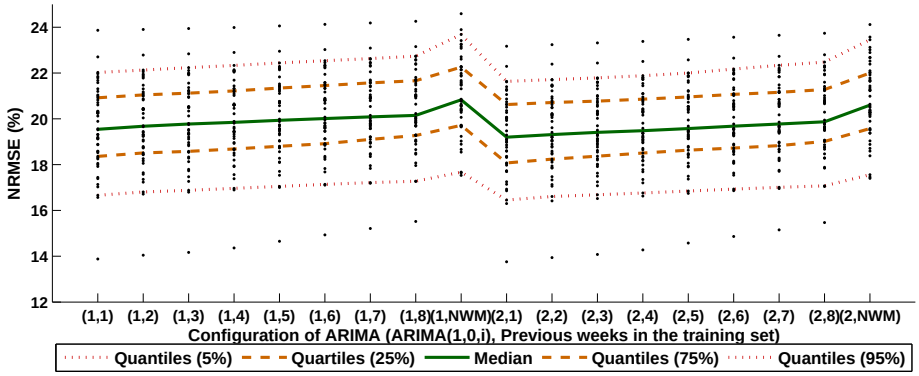

Fig. 7: Repartition of total NRMSE for each configuration - ARMA

Moreover, ARMA-WM models do not show satisfactory results, as the total NRMSE median is approximately 20%. In this case, the 5 models which have to be selected are the first 5 models with error coefficients equal to 2. Finally, following a similar procedure to the one in Section 4.3, the model selected for ARMA method is (ARIMA(1,0,2),1). It can be re-written as an ARMA(1,2) model with a training set formed by only the previous week.

4.6 MR Model Selection

For the MR models, the results show that a model built without using a Window-Time Moving for the training set creates a decrease in accuracy. This decrease reaches more than 4% of total NRMSE for the median. The second result is linked to the evolution of quartiles which gives approximately the same confidence interval, independently of the degree of the polynomial function used. It does not lead to overfitting thanks to the regularized cost function. However, it does not imply better forecasting.

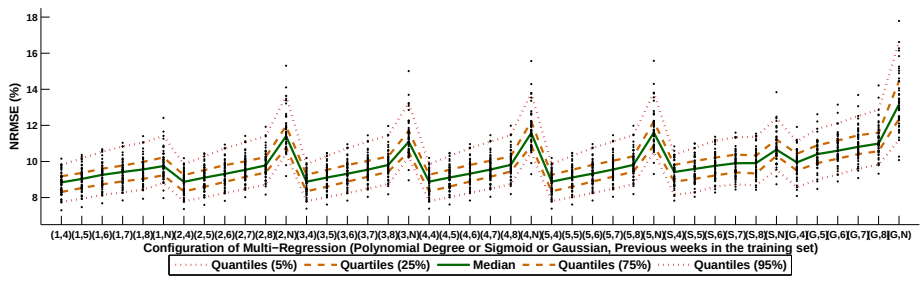

Fig. 8: Repartition of total NRMSE for each configuration - MR

Sigmoid or Gaussian functions give worst results for the corresponding models in terms of training set used. Finally, the 5 models selected as this stage are MR(2,4), MR(3,4), MR(4,4), MR(5,4) and MR(S,4). The best number of previous weeks to load as training set was concluded to be 4. Multi-Regression, even when there are 24 sub-models created (one for each hour of the day), has very low accuracy. The main conclusion is that the consumption of the previous hour has a significant impact on the model. Another approach is to build one model and include as input the hour of the forecast. The selected model is MR(5,4) which uses a polynomial function of degree 5 and takes into account the 4 previous weeks as training set.

4.7 ARMR Model Selection

The conclusions that can be drawn from Fig. 9 are similar to those from Section 4.6, which are shown in Fig. 8. The compensation expected between ARMA and MR models is not efficient, and leads to an overall accuracy which is between the accuracy of both models. Finally, the model selected is ARMR(5,4).

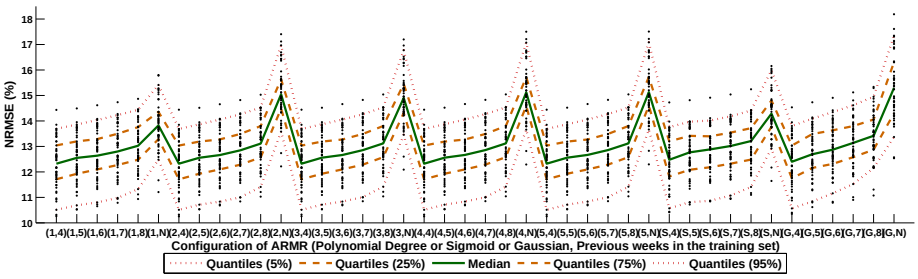

Fig. 9: Repartition of total NRMSE for each configuration - ARMR

4.8 Evaluation of Selected Models

In this part, the 5 different models are evaluated along one year of consumption. The ranking regarding models accuracy fits with the one illustrated in the first experiments. In fact, the ARMA model is the least performing one, with a NRMSE average values 12.86% higher compared to the ARMR model, as can be seen from Table 12. Moreover, the gap between ARMA and ANN is between 19.45% - the 253th day - and 8.38% - the 122nd day. Thus, even during the summer, which is on both sides of the middle of Fig. 10, when the electric energy consumption is not fluctuating too much, ARMA remains inefficient. Moreover, it can be observed in Fig. 10 that the shape of daily NRMSE of the ARMR model and MR model are similar.

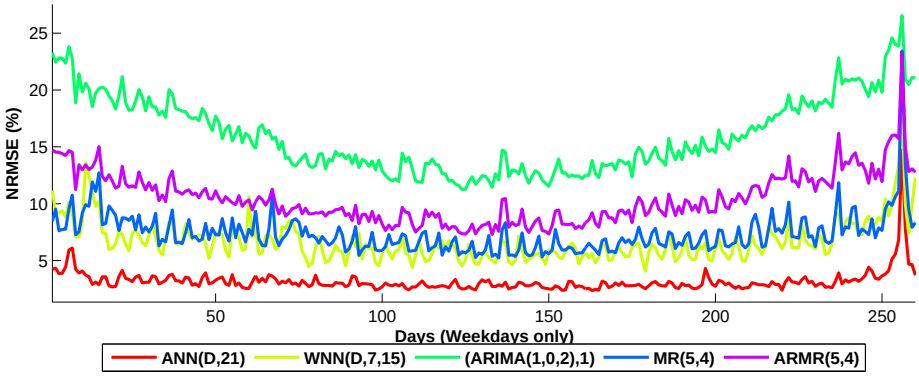

Fig. 10: Comparison of the 5 selected methods with daily NRMSE

The peak of daily NRMSE for the 256th day involves the limit of each method. Indeed, no procedure of anomaly detection of energy demand has been implemented. Thus, this day, which corresponds to Monday, 27th December 2010, is directly following Friday, 25th December 2010. It is because of only using week days in the experiments, due to their particular difference in demand patterns from weekend days. The consumption during this particular bank holiday is anomalous, and leads to overestimated inputs for the following day.

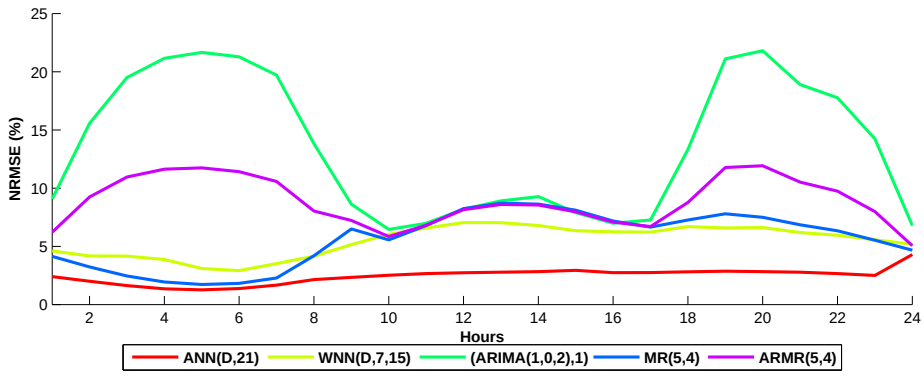

Fig. 11: Comparison of the 5 selected methods along the day with average of NRMSE

As shown in Fig. 11, from 10am to 5pm, due to the regularity of the working days, most of the models achieve the highest forecasting accuracy. Only ANN(D,21) is more robust throughout the day. The ARMA model is limited by the fluctuations of the demand during peak hours, specifically between 5-11pm, even achieving an average of 21.8% NRMSE at 8pm.

Method	1	2	3	4	5	6	7	8	9	10	11	12
ANN	2.40	2.01	1.63	1.35	1.26	1.38	1.67	2.15	2.34	2.52	2.66	2.73
WNN	4.61	4.18	4.16	3.87	3.10	2.92	3.52	4.16	5.15	6.03	6.57	7.05
ARMA	9.07	15.57	19.51	21.16	21.66	21.28	19.71	13.82	8.62	6.45	6.98	8.20
MR	4.14	3.23	2.46	1.94	1.72	1.82	2.28	4.19	6.49	5.56	6.78	8.24
ARMR	6.21	9.24	10.96	11.63	11.74	11.42	10.58	8.04	7.23	5.85	6.81	8.16

Fig. 12: The 5 selected methods along the day with average of NRMSE - part 1

Method	13	14	15	16	17	18	19	20	21	22	23	24
ANN	2.78	2.83	2.94	2.74	2.75	2.81	2.87	2.83	2.78	2.66	2.50	4.29
WNN	7.03	6.79	6.35	6.25	6.23	6.69	6.58	6.62	6.19	5.96	5.57	5.18
ARMA	8.91	9.27	7.90	6.99	7.26	13.32	21.10	21.80	18.90	17.77	14.24	6.79
MR	8.71	8.63	8.11	7.18	6.65	7.27	7.80	7.49	6.85	6.34	5.53	4.67
ARMR	8.60	8.55	7.97	7.08	6.70	8.77	11.78	11.93	10.52	9.74	7.99	5.06

Fig. 13: The 5 selected methods along the day with average of NRMSE - part 2

As for WNN model, the difference of accuracy when compared to the ANN model stems from the addition of forecasting errors on the 4 predicted sub-signals. ANN(D,21) obtains 1.26% NRMSE at 5am on all the configurations and days tested, and has a maximum NRMSE of 4.29% at midnight. Taking into account the unavoidable residual errors when forecasting electric energy consumption in such a small-scale network, the result is considered to be sufficiently accurate for generator scheduling for this model.

5 Conclusion

Frequentist inference approaches allow to set the 5% and 95% error boundaries to ensure that the user can take a decision according to risk evaluation, by taking into account the previous errors occurring in the forecast model. ARMA models are not very efficient in such scales and seem to be more compatible with larger groups of houses. The main conclusion which can be extracted from the models evaluated in this set of experiments is that ANNs are the most powerful models that can be used to forecast electric energy consumption of very small-scale networks. Furthermore, a potential problem is that the training set is not as comprehensive as other sets used in the literature. Another method which could improve provide improvements is the use of Bayesian networks to fit the ANN, such those described in [13, 14].

Acknowledgments. This work was supported by Trinity College Dublin.

References

1. A. K. Alexandridis and A. D. Zapranis, *Wavelet neural networks: A practical guide*, in Neural Networks, vol. 42, June 2013, pp. 1–27.
2. J. Auer and J. Keil, *State-of-the-art electricity storage systems: Indispensable elements of the energy revolution*, March 2012.
3. D. Clancy and U. Ozguner, *Wavelet neural networks: a design perspective*, in Intelligent Control, 1994., Proceedings of the 1994 IEEE International Symposium on, Aug 1994, pp. 376–381.
4. Y.-H. Dai, *A perfect example for the bfgs method*, Mathematical Programming, 138 (2013), pp. 501–530.
5. T. Edwards, *Discrete wavelet transforms: Theory and implementation*, 1992.
6. C. Guan, P. Luh, L. Michel, Y. Wang, and P. Friedland, *Very short-term load forecasting: Wavelet neural networks with data pre-filtering*, Power Systems, IEEE Transactions on, 28 (2013), pp. 30–41.
7. M. Hayati and Y. Shirvany, *Artificial neural network approach for short term load forecasting for illam region*, World Academy of Science, Engineering and Technology, 28 (2007), pp. 280–284.
8. L. Hernandez, C. Baladrón, J. M. Aguiar, B. Carro, A. J. Sanchez-Esguevillas, and J. Lloret, *Short-term load forecasting for microgrids based on artificial neural networks*, Energies, 6 (2013), pp. 1385–1408.
9. E. Hossain, E. Kabalci, R. Bayindir, and R. Perez, *Microgrid testbeds around the world: State of art*, Energy Conversion and Management, 86 (2014), pp. 132 – 153.
10. C. Igel and M. Hsken, *Improving the rprop learning algorithm*, 2000.
11. S. Karsoliya, *Approximating number of hidden layer neurons in multiple hidden layer bpnn architecture*, in International Journal of Engineering Trends and Technology, vol. 3, pp. 714–717.
12. T. Kivipõld and J. Valtin, *Regression analysis of time series for forecasting the electricity consumption of small consumers in case of an hourly pricing system.*
13. O. Kocadal and B. Akgil, *Nonlinear time series forecasting with bayesian neural networks*, Expert Systems with Applications, 41 (2014), pp. 6596 – 6610.
14. F. Liang, *Bayesian neural networks for nonlinear time series forecasting*, Statistics and Computing, 15 (2005), pp. 13–29.
15. A. Marinescu, I. Dusparic, C. Harris, V. Cahill, and S. Clarke, *A dynamic forecasting method for small scale residential electrical demand*, in Neural Networks (IJCNN), 2014 International Joint Conference on, July 2014, pp. 3767–3774.
16. A. Marinescu, C. Harris, I. Dusparic, S. Clarke, and V. Cahill, *Residential electrical demand forecasting in very small scale: An evaluation of forecasting methods*, in Software Engineering Challenges for the Smart Grid (SE4SG), 2013 2nd International Workshop on, May 2013, pp. 25–32.
17. M. Nezzar, N. Farah, and T. Khadir, *Mid-long term algerian electric load forecasting using regression approach*, in Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE), 2013 International Conference on, May 2013, pp. 121–126.
18. A. Ng, *Machine Learning (MOOC)*, Stanford University, Coursera.
19. S. Nissen, *Implementation of a fast artificial neural network library (fann)*, (2003).
20. S. S. E. A. of Ireland), *Ireland - your smart grid opportunity*, September 2010.
21. M. Riedmiller and H. Braun, *A direct adaptive method for faster backpropagation learning: the rprop algorithm*, in Neural Networks, 1993., IEEE International Conference on, 1993, pp. 586–591 vol.1.

22. K. G. Sheela and S. N. Deepa, *Review on methods to fix number of hidden neurons in neural networks*, Mathematical Problems in Engineering, (2013), p. 11.

23. S. Subbayya, J. Jetcheva, and W.-P. Chen, *Model selection criteria for short-term microgrid-scale electricity load forecasts*, in Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES, Feb 2013, pp. 1–6.

24. Suhartono, I. Puspitasari, M. Akbar, and M. Lee, *Two-level seasonal model based on hybrid arima-anfis for forecasting short-term electricity load in indonesia*, in Statistics in Science, Business, and Engineering (ICSSBE), 2012 International Conference on, Sept 2012, pp. 1–5.

25. M. Tasre, V. Ghate, and P. Bedekar, *Hourly load forecasting using artificial neural network for a small area*, in Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on, March 2012, pp. 379–385.

26. D. Veitch, *Wavelet neural networks and their application in the study of dynamical systems*, August 2005.

27. L. Wei and Z. zhen gang, *Based on time sequence of arima model in the application of short-term electricity load forecasting*, in Research Challenges in Computer Science, 2009. ICRCCS '09. International Conference on, Dec 2009, pp. 11–14.

28. Z. yang Dong, B.-L. Zhang, and Q. Huang, *Adaptive neural network short term load forecasting with wavelet decompositions*, in Power Tech Proceedings, 2001 IEEE Porto, vol. 2, 2001, pp. 6 pp. vol.2–.

Appendix

5.1

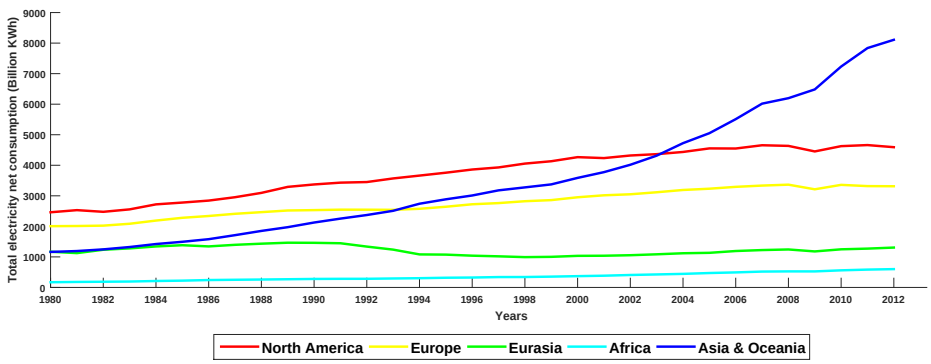

Fig. 14: Evolution of world electrical consumption by continent from 1980 to 2012

5.2

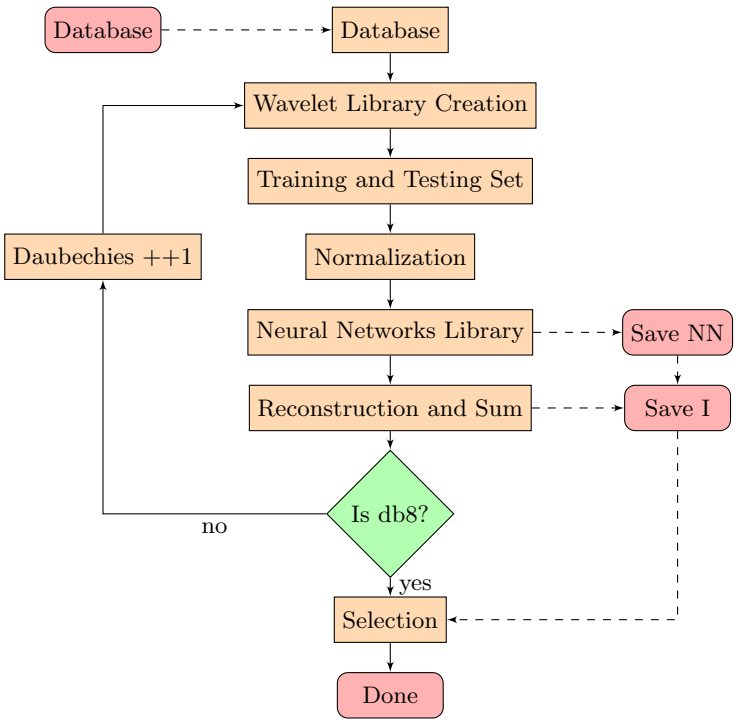

Fig. 15: Used process to build a WNN

5.3

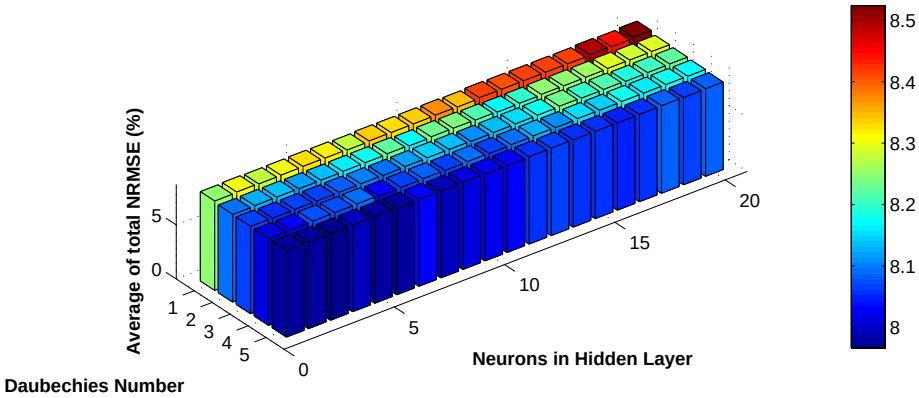

Fig. 16: Average of total NRMSE for WNN - previous day

5.4

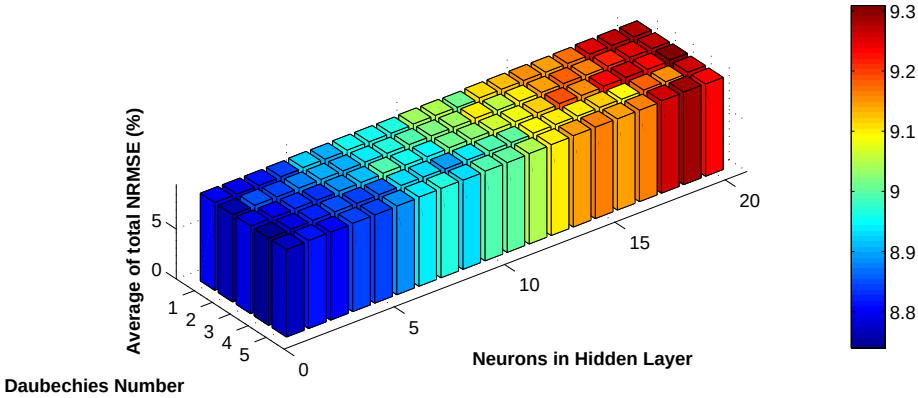

Fig. 17: Average of total NRMSE for WNN - previous week

5.5

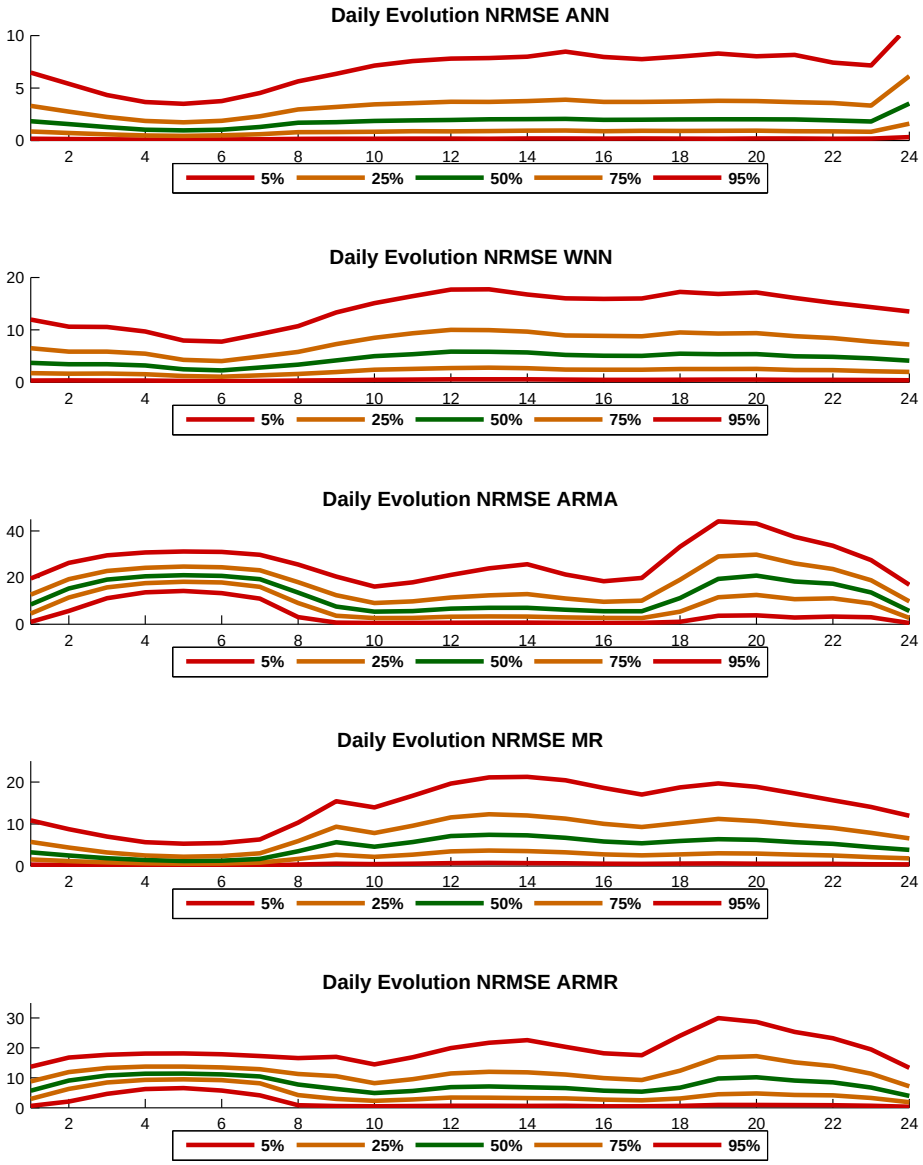

Fig. 18: Comparison of Hourly NRMSE along one year