New Forecasting Method for the Residual Demand Curves using Time Series (ARIMA) Models

Agustín Martín Calmarza, José Ignacio de la Fuente

Instituto de Investigación Tecnológica
Escuela Técnica Superior de Ingeniería (ICAI)
Universidad Pontificia Comillas
28015 Madrid
agustin.martin@iit.upco.es, nacho@iit.upco.es

Abstract—**In this paper a new methodology to forecast the day ahead electricity market behaviour is presented. This behaviour can be easily modelled by means of the so called residual demand curves (RDC's) . The pattern of these curves (as the spot market is an hourly market there is one RDC for each hour) changes greatly according with the type of the day (labour-non labour) and the hour (peak ,valley, plateau,...) so this fact must be taken into account. Firstly, a classical ARIMA analysis without explanatory variables is carried out. Afterwards, adequate explanatory variables are searched in order to build a more accurate Transfer Function Model. Next a new procedure called weighted estimation is developed and the differences between these two methods are pointed out. Finally, a case study is presented in order to check the validity of the weighted estimation model.**

Keywords--**Spot Market, Residual Demand Curves (RDC's), ARIMA models, TF models, weighted estimation, explanatory variables, Market behaviour estimation.**

I. NOMENCLATURE

RDC's: Residual Demand Curves.
ACF/PACF: Autocorrelation/Partial Autocorrelation Function.
∇_x: Differencing Operator of order x.
a_t: Random shock (white noise).
B^x: Backshift Operator. It delays x lags when applied to a given variable.
SO/MO: System/Market Operator.
WE: Weighted Estimation Method.
MA: Moving average term
AR: Autoregressive term
RSE: Residual Standard Error.
TF: Transfer Function Model.
LTF: Linear Transfer Function. Method used to adjust TF models.

II. INTRODUCTION

Under the new Spanish competitive environment where generator companies should develop their activities, it is important to envisage techniques to forecast the day ahead electricity market behaviour. One of the simplest ways to model market conditions is by means of the so called Residual Demand Curves (RDC´s). These curves relate the amount of power dispatched to a given agent in a certain hour with respect to the resulting marginal price. The process to build up RDC´s curves is shown in figure 1.

The Spanish electricity system has recently changed its rules, so every utility has access, with only one day of delay, to the aggregated demand and supply energy day-ahead bid curves, which is published by the market operator. Then, it results very easy for all the utilities the calculation of their corresponding residual demand curves. They must just subtract their own supply bid curve from the aggregated supply curve in order to obtain the **aggregated supply bid curve of the rest of agents** shown in figure 1. Under this procedure, an accurate forecasting of the day-ahead residual demand curve can be carried out using *time series analysis.*

Therefore, RDC's are important inputs in order to carry out the dispatch optimization, which every utility has to solve in order to maximize its expected profits. A wrong prediction of this curves may provide misleading results in the dispatch optimization process.

The present article is focused on the forecasting and estimation of the **hourly** RDC's. For generation agents of a moderate size, the shape of the RDC´s can usually be approximated using a linear interpolation. Thus, the problem is reduced to the analysis of the time evolution of two variables, the slope and the intercept of the linearized RDC´s (see figure 1). Both of them can fluctuate sharply depending on the hour of the day and also the day of the week. Thus, hours with a strong demand of electricity tend to have greater values of slope and intercept than the rest, due to the higher level of demand served and the more expensive generation technologies dispatched. As the prediction's horizon is very short (day-ahead market or spot market) tools like ARIMA time series are useful to achieve a good RDC´s estimation. In particular this approach will allow the following analysis:

a) Determining whether a unique time series model for each one of the variables is enough or more detailed models (i.e, distinction between labor days and weekends or between peak/valley demand hours) lead to more accurate results. In the latter case a weighted estimation is developed which allows to obtain optimal parameters for each of the time zones considered. The study of possible data's seasonalities is considered as well.
b) Evaluating the influence of some explanatory variables in the shape of the residual demand curves, like the expected system demand, the expected level of run of river power or the expected amount of nuclear power in the system. This fact leads to the use of transfer function (TF) models.
c) Classification/filtering of outliers which can come up in the data.
d) Analysis of calendar variations (i.e, holidays like Easter or Christmas) and known disturbances (like a sudden increase in the run of river power's level due to heavy rainfalls), which can distort the performance of the models used.

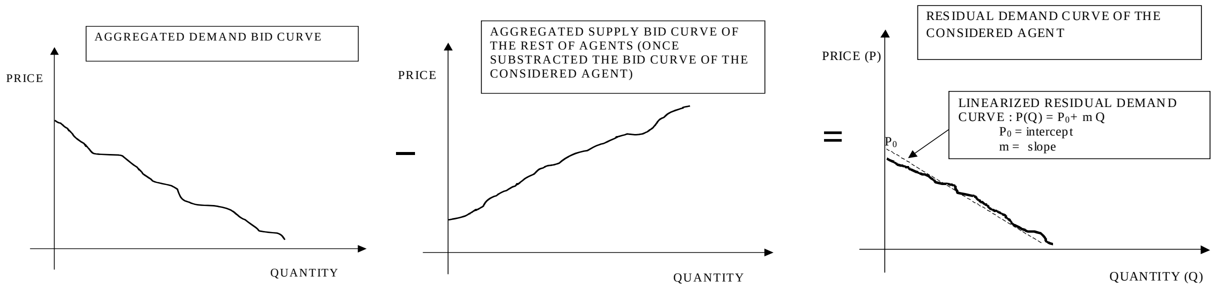

Figure 1: Construction process of a residual demand curve.

The paper is structured as follows: in section III a classical ARIMA analysis of the two variables (i.e, intercept and slope) involved in the process corresponding to the linearized RDC's is explained and a ARIMA model for each one is developed. In Section IV a useful combination of explanatory variables is looked at in order to improve the performance of the original ARIMA model. Besides this, a weighted estimation method is presented which will allow the achievement of optimal models for each of the different time intervals that can be considered depending on the type of day (labour, weekend) and the hour (peak, valley, plateau). Section V shows a case study where each of the above models (ARIMA, TF and weighted estimation) performance is compared. Finally, Section VI contains a summary of conclusions obtained by means of a detailed analysis of the models' performance.

III. ANALYSIS USING AN ARIMA MODEL

Firstly, it must be determined whether a transformation of the original series is required in order to deal with the conditions required to perform an ARIMA analysis. Standard ARIMA analysis relies on the assumption that the time series is stationary. This means that the mean and the variance of the data series are constant through time. The first of the problems (stationary mean) is solved by differencing the series and it is clearly visible in the pattern of the ACF.

Possible modifications to induce a stationary variance must be applied (when needed) before any further analysis of the data. It is fairly common to find data series where their variance is proportional to its level. In this case a constant variance is induced by transforming the data, being the Box-Cox transformations the most widely used [1],[6].

An analysis of a two months sample (Fig.2) of the linearized RDC's intercept (a similar graph is obtained using the variable slope) shows that the variance tends to be more or less constant, regardless of the level reached. So a transformation in this case is not necessary.

Once these previous steps have been accomplished, it is time to analyse the ACF and PACF of the model, in order to identify the autoregressive (AR) and moving average (MA) parameters of the model. During the identification procedure the principle of parsimony will be applied. According to this principle, models with as few coefficients as needed to adequately explain the behaviour of the data will be selected.

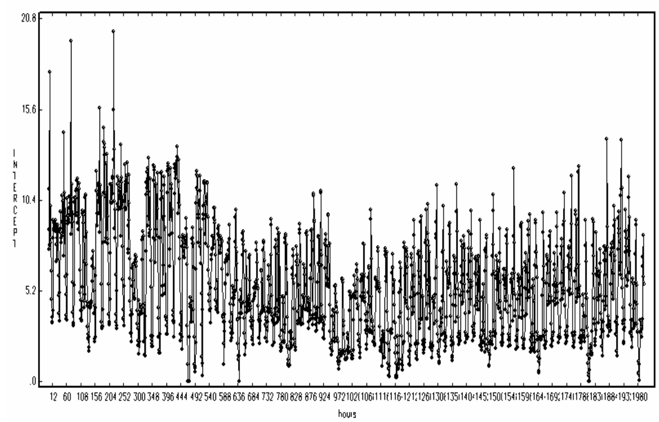

Figure 2. Variable intercept, hourly evolution during two months.

When a tentative model is proposed, then it must be estimated and checked to verify if it complies with the stationarity (mean, variance and ACF are constant through time) and invertibility (which ensures that more recent data have more weight than the distant past) conditions and if the value of the obtained coefficients seem reasonable.

The following ACF/PACF graphs, correspond to the variable intercept. The graphs obtained with the variable slope present the same kind of pattern and behavior.

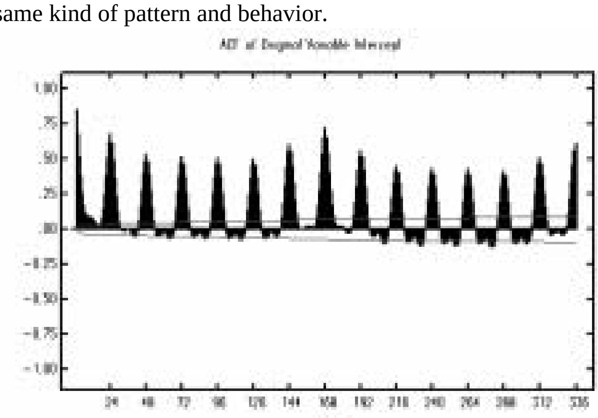

Figure 3. ACF for the variable intercept.

The outstanding feature of the ACF shown in figure 3 is a slow decay at the first seasonal lags (24,48 ,72) mixed with a slow growth during the following seasonal lags (96,120,144) until it reaches a local maximum at lag 168.

Then the procedure repeats itself from lag 168 to lag 336 but with slightly lower values.
So it seems that there is a double periodicity, one with a daily scope (this is the reason of the peaks of the ACF function at lags *t=24*k*, with *k=1,2,3,...*) and the other with a weekly horizon (ACF peaks at lags *t=168*k, k=1,2,3,....*). Due to the ACF pattern at these seasonal lags, a double seasonal differencing is needed. Regular differencing (i.e, between adjacent values) is not necessary because the ACF decays fairly quickly at the first lags.

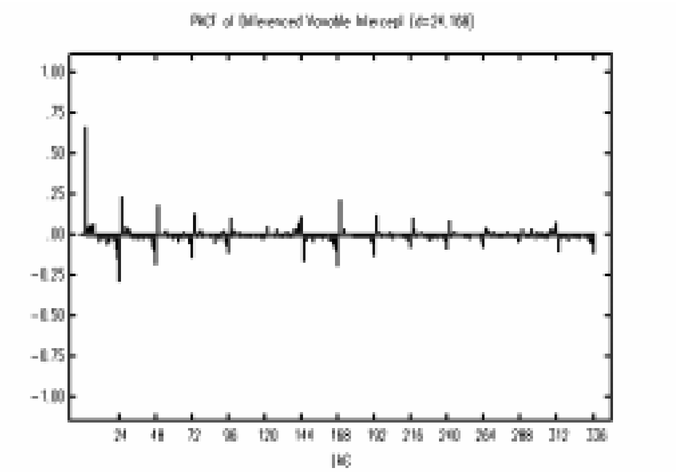

Figure 4. PACF for the variable intercept with double seasonal differencing.

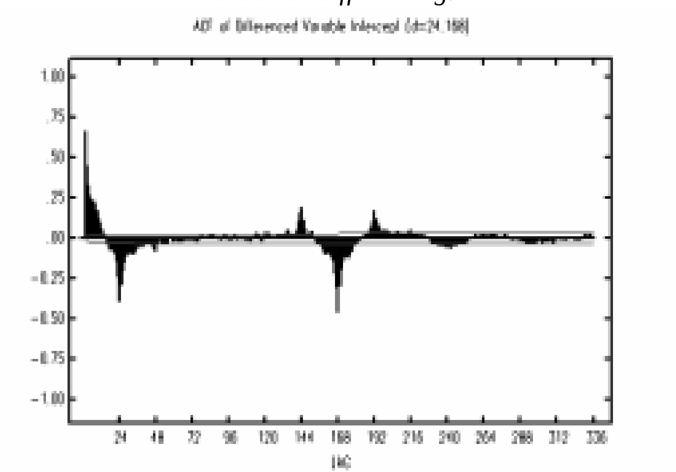

Figure 5. ACF for the variable intercept with double seasonal differencing.

Figure 4 shows a slow decay in the PACF pattern at the seasonal lags *t1=24*k* and *t2=168*k, k=1,2,3,...* .Figure 5 shows high spikes at lags *t=24* and *t=168* and a lower one at lag *t=48*. This means that a moving average behaviour is present at these lags. In contrast, the first lags show a different pattern with a spike at lag *t=1* in the PACF (Figure 4) and a decay in the ACF (Figure 5), so a first order autoregressive parameter looks adequate. Many of the other large autocorrelations (e.g., lags 144 and 192) observed in figure 5 are probably a reflection of the weekly seasonality, thus it is not necessary to consider more differencings. It is also very important to avoid overdifferenced models, which always drive to pure moving average models (as the classical example of the number of passengers of an airline [2],[5]). Thus, a tentative model which complies with all the above characteristics is proposed.

$$\nabla_{24}\nabla_{168}Y = k + \frac{(1-\theta_1\ B^{24}-\theta_2\ B^{48})\,(1-\theta_3\ B^{168})}{(1-\phi_1 B)}a_t \qquad (1)$$

where Y corresponds to the model output variable (either intercept or slope) of the linearized RDC's, and 'k' is a constant. The rest of nomenclature has been defined in Section I.

SCA© (the statistical software program used) provides the following particular values for the model given in (1):

Coefficients.	Type	Order	Value	t-Value
ϑ1	MA	24	0.6379	67.81
ϑ2	MA	48	0.1551	16.27
ϑ3	MA	168	0.7708	118.26
Φ1	AR	1	0.7368	114.98
k		0	-0.0006	-0.41
RSE			0.82	

Table 1.SCA results for the variable intercept. ARIMA model.

Coefficients.	Type	Order	Value	t-Value
ϑ1	MA	24	0.6406	68.04
ϑ2	MA	48	0.1455	15.24
ϑ3	MA	168	0.8006	126.63
Φ1	AR	1	0.7208	109.75
k		0	-0.0963	-0.4
RSE			151.65	

Table 2.SCA results for the variable slope. ARIMA model.

The 't-value' indicates whether a parameter of the models is meaningful ('t-value' greater than two) or not by means of a t-Student significance test. As it was pointed out, both variables present the same characteristics (same periodicities and differencings) so the results of the estimated parameters are very closed. It is straightforward to check that both determined models deal successfully with the stationarity and invertibility conditions required for an ARIMA model (All the roots of the AR and MA polynomials lie outside the unit circle [1],[2],[3]). Once a tentative model is specified, it is mandatory to check if the residuals obtained are normally distributed. There are a great variety of tests to accomplish this task (e.g, the histogram of the residuals, the normal probability plot or the Box-Ljung test [1],[2],[3]).

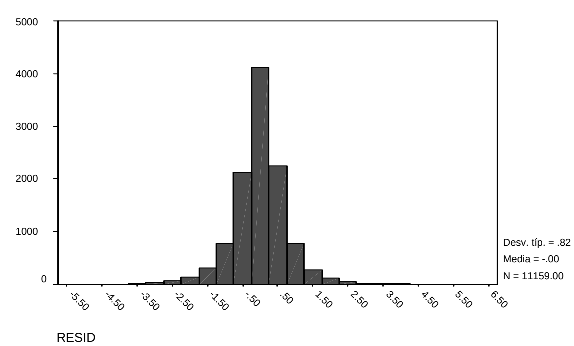

Figure 6 .Histogram of the residuals for the variable intercept.

The histogram of the intercept variable residuals shown in figure 6 (same results obtained for the residuals of the variable slope) is quite symmetrical, suggesting that the residuals are normally distributed.

The normal probability plot (not shown here) [1],[5], does not have large deviations from a straight line except from very large residuals' values, again suggesting that the residuals are normal distributed and also that an outlier test should be done to analyse the mentioned large residuals values (this task will be briefly commented on in Section IV.D).

Also the ACF of the residuals series is totally clean (i.e, each residual autocorrelation falls well short of its two standard error limit) concluding that the model adequately captures the autocorrelation patterns in the data.

The ARIMA models built in this section will be used as a comparison base regarding the performance of an improved model which will be proposed in the following section.

IV. ANALYSIS USING TF MODELS AND WEIGHTED ESTIMATION

In order to improve the performance of the above ARIMA model, adequate explanatory variables should be searched. A useful explanatory variable should be easily available and, if possible, known with little uncertainty. Three meaningful variables have been determined in order to explain the evolution (level and slope) of the linearized RDC´s (results will only be shown for the variable intercept, but the interpretation with the variable slope is straightforward). For the sake of clarity **results are shown in weekly average values although real data are hourly values**. Thus, these explanatory variables are (in order of importance):

a)**The national hourly electricity demand**. It is obvious that the higher the demand, the higher and more abrupt should be the RDC´s. At present, existing demand forecasting models are very accurate, so the demand will be treated as a non stochastic variable.

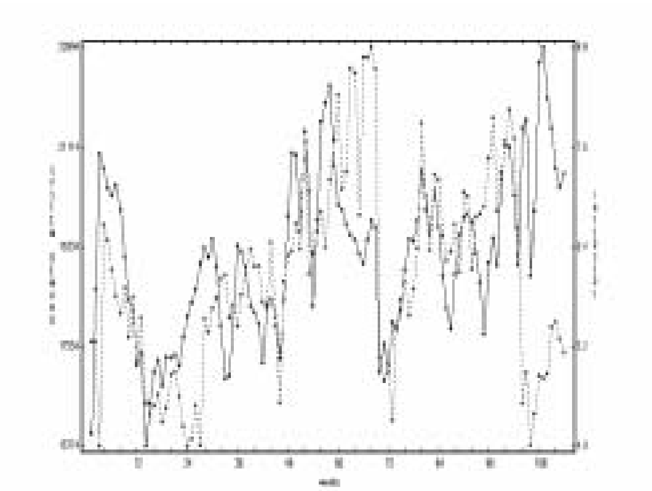

Figure 7.Evolution of the variable system demand (continuous line) versus variable intercept (dashed line).(Weekly average values).

Although it can be seen in figure 7 that there is a meaningful correlation between both variables, there are periods where the demand was high and the intercept low, and vice versa. So there are more factors involved in the RDC's time evolution.

b) **Run of river power.**

There is always a fraction of the whole generation power available which is bided at the lowest prices. This portion of power (known as base generation) corresponds to the amount offered by both nuclear power plants and run of river units (hydro plants without regulating capacity, because they do not have dams). This base hydro generation remains rather constant through a day and it is easy to forecast, knowing the dams hydro conditions and the weather forecast. Figure 8 shows the correlation between the variable intercept and the variable run of river power with a weekly average value basis.

It can be observed that when there has been a sudden increase in the level of the run of river power (due to heavy rainfalls), the values of the variable intercept has fallen dramatically. Also in this situation the RDC´s tend to be flatter so the values of the variable slope are lower than in normal conditions.

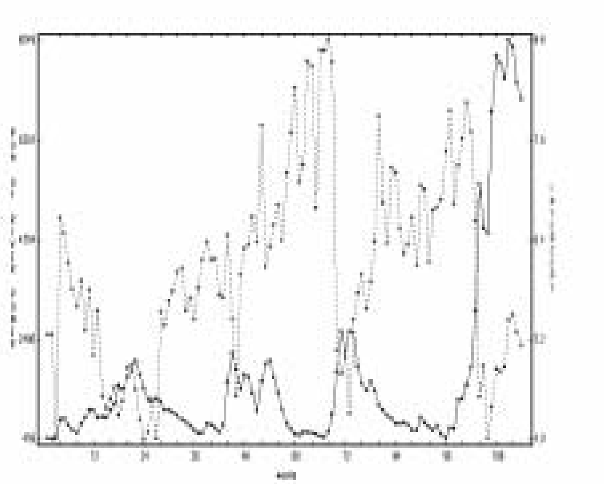

Figure 8.Evolution of the variable run of river power (continuous line) versus variable intercept (dashed line).(Weekly average values).

c) **The level of nuclear power.**

As it was said previously, nuclear power is a base generation component. The scheduled maintenance of the nuclear power plants use to be known in advance, so this variable becomes a useful explanatory variable. The effect of the level of this variable is quite similar to the one shown with the variable run of the river power.

B. AGGREGATION IN A UNIQUE EXPLANATOTY VARIABLE

If a TF model is tried using separately the above three explanatory variables, all the variables except from the electricity demand will be considered as non-meaningful. The reason is clear: On one hand the two output variables (intercept and slope) change greatly from hour to hour, but the variables run of river and nuclear power are rather constant during the whole day, so the estimated correlation between the output variables and these two input variables will be very small. On the other hand, if these variables are compared in a weekly basis (as in figure 8), it is obvious that exists a relationship between the output variables and the level of run river power (or nuclear power).

Therefore a explanatory variable with an hourly scope and that takes into account all the above characteristics needs to be defined. If the sum of run of river power and nuclear power (nearly constants for the whole day) is subtracted from the hourly system demand, a useful explanatory variable is obtained. It can be explained as the portion of demand which is not covered by the base generation or, in brief terms, the non-base demand. Formally, the proposed explanatory variable is defined as:

Expected Non-Base Demand(h) = Expected System Demand(h) - Expected Run of river power(h) - Expected Nuclear power(h) *(2)*

Where h=1,…,24 (hours of the day).

This new explanatory variable covers all the possible cases. It will take high values if the system demand is high and/or there is a lack of base generation (due to low values of run of river power and/or nuclear power) and lower values if the system demand is low and/or there is a high level of base generation, so it presents an adequate hourly variation. For the sake of clarity, figure 9 shows in weekly values, the relationship between the new explanatory variable and the variable intercept.

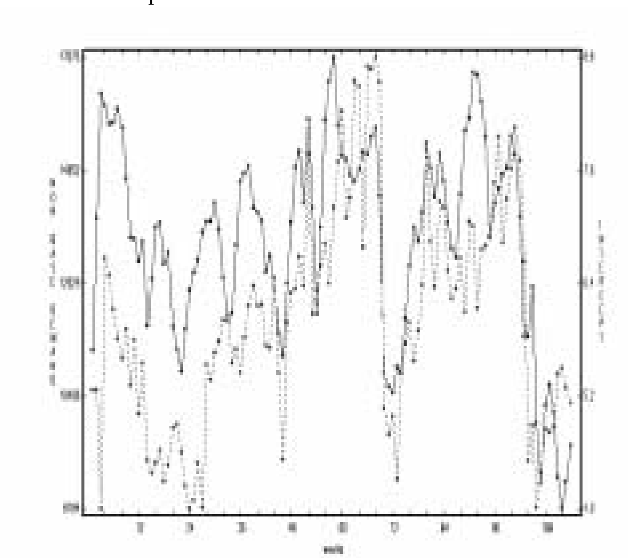

Figure 9.Evolution of the variable non-base demand (continuous line) us variable intercept (dashed line).(Weekly average values).

It is observed in figure 9 that the peaks and valleys of the new explanatory variable evolution bear strong resemblances with those of the variable intercept, and their correlation is much better than when only the variable demand (figure 7) is considered. For example, focusing on the last weeks of figures 7 and 9, which corresponds with a period of both high levels of system demand and run of river power, the new explanatory variable is able to capture the sudden drop of the variable intercept, while if only the system demand is used the results will be poor, because higher prices are expected. The same fact occurs with the variable slope of the linearized RDC's, which has low values (i.e, the RDC's are flatter) if the explanatory variable non-base demand has also low values and vice versa.

C . TF MODELS AND A WEIGHTED ESTIMATION APPROACH

Firstly, a non-weighted TF model with non-base demand as explanatory variable will be proposed. The LTF approach ([5],[6]) has been applied to adjust the model stated in (3), because it presents some advantages with respect to the more classical **prewithening** method ([2],[3]). Using this method and skipping the non meaningful lags of the explanatory variable, the following structure results:

$$\nabla_{24}\nabla_{168} Y = v_0 \nabla_{24}\nabla_{168}\left[non-base\ demand\right] + \frac{(1-\theta_1 B^{24} - \theta_2 B^{48})(1-\theta_3 B^{168})}{(1-\phi_1 B)} a_t \qquad (3)$$

The TF model shown in (3) does not dramatically reduce the RSE in comparison with the ARIMA model proposed in Section III, but when an outlier analysis is carried out, a great number of the outliers caused by changes in the input variable are explained, so its total amount is strongly reduced using the TF model.

Regarding the model considered in (3), it has some serious **drawbacks**.

a) If the terms containing differencings are fully developed, the model shows a contemporaneous relation between the values of the output and input variables at current time and the values which took place twenty four (the day before) and one hundred sixty eight (the week before) hours ago. In fact there are some more relations than these, but they are much less important.

The fact stated above can give problems when one is focused on days like Mondays or Saturdays, when there are great differences between their RDC´s pattern and those of the previous day (Sunday and Friday respectively). So it seems reasonable that Saturdays, Sundays and Mondays will follow a different model than that stated in (3), therefore a daily discrimination must be considered.

b)The TF model shows a linear and unique relation between the output and the input variable (coefficient *v0* in (3)), regardless of the type of hour considered. This is not totally true, because the ratio between the values of the output and input variables changes depending on the hour and the day of the week. Thus both daily and hourly discriminations should be taken into account.

The **weighted estimation method(WE)** provides a good and easy way to deal with problems like non-linear relations or saturations between the variables [6]. In order to use the weighted estimation approach, a prior classification of the available data series divided in time ranges must be done. As it was addressed, two discriminations have been done.

a)**Daily discrimination:** One independent model for each of the following groups: 1) Saturday, 2) Sunday, 3) Monday and 4) Tuesday through Friday.

b)**Hourly discrimination:** One independent model for each of the following groups, according to the pattern of the system demand curve: 1) From 2 a.m to 7 a.m., 2) 8 a.m and 9 a.m., 3) from 10 a.m to 14 p.m, 4) from 15 p.m to 18 p.m., 5) from 19 p.m to 22 p.m, and 6) from 23 p.m to 1 a.m.

The weighted estimation method searches for the optimal values of the parameters which minimize the RSE at the time range considered [6]. It is important to note that, despite the weights applied, the whole series is still considered, so the dynamic pattern of the data is not lost. The objective is to adequately represent the non-linear relationship between output and input variables as a group of linear models focused in different time intervals, taking advantage of the simplicity and clarity of linearity. The following model equations and tables are focused on the time period 8 a.m-9 a.m and results are displayed for different days and for the variable intercept.

Firstly, the adjusted model and its estimated parameters focusing on Saturdays is shown.

$$Y = v_0 \left[non-base\ demand\right] + \frac{(1-\theta_1 B)}{(1-\phi_1 B)} a_t \qquad (4)$$

Coefficients.	**Type**	**Order**	**Value**	**t-Value**
v0		0	0.0002	14.62
ϑ1	MA	1	0.5746	10.21
Φ1	AR	1	0.8962	32.16
Differencings	None			
RSE(Time zone)			0.51	

Table 3 .SCA results for time interval Saturday, 8 a.m-9 a.m. Weighted estimation. Model equation (4).

Next, the optimal model for the same hourly range (8 a.m-9 a.m) focusing on Tuesday through Friday is developed.

$$\nabla_{24} Y = v_0 \ \nabla_{24} \left[non-base\ demand\right] + \frac{(1-\theta_1\ B\ )\ (1-\theta_2\ B^{24})}{(1-\phi_1 B)\ (1-\phi_2 B^{24})\ (1-\phi_3 B^{168})}\, a_t \qquad (5)$$

Coefficients.	Type	Order	Value	t-Value
v0		0	0.0008	13.07
ϑ1	MA	1	0.2802	3.04
ϑ2	MA	24	0.7904	13.86
Φ1	AR	1	0.7426	11.01
Φ2	AR	24	0.4392	6.01
Differencings	Yes	24		
Φ3	AR	168	0.1562	3.6
RSE(Time zone)			0.9	

Table 4.SCA results for time interval Tuesday through Friday, 8 a.m-9 a.m. Weighted estimation. Model equation (5).

Finally, the optimal model focused on Sundays with the previous hourly range (8 a.m-9 a.m) is expressed:

$$Y = v_0 \ \left[non-base\ demand\right] + \frac{(1-\theta_1\ B)}{(1-\phi_1 B)(1-\phi_{31} B^{168})}\, a_t \qquad (6)$$

Coefficients.	Type	Order	Value	t-Value
v0		0	0.0001	4.41
ϑ1	MA	1	0.2285	2
Φ1	AR	1	0.848	14.43
Φ3	AR	168	0.1727	2
Differencings	None			
RSE(Time zone)			0.59	

Table 5. SCA results for time interval Sunday, 8 a.m-9 a.m. Model equation (6).

The tables above show a great variety in the structure of the models depending on the time interval considered. They should be interpreted in the following way: For instance, focusing on the row corresponding to parameter ‘ϕ2’ in table 4, it means that a seasonal daily (order is 24) autoregressive (AR) parameter labelled as ‘ϕ2’, which takes a value of *0.4392* is considered in the model, because it is a meaningful parameter (its *t-value* is greater than two).
Although the base model used has been always the one stated in (3), the weighted estimation method focused on the time zone selected skips in each case the non meaningful parameters and differencings or add new parameters like ‘ϕ2’ and ‘ϕ3’ in table 4, until an optimal parsimonious model which deals with the stationarity and invertibility conditions has been reached ([1],[2],[3]).
It can be noted that tables 3 and 5, which correspond to weekend days, do not include any differencing, because due to the strong variation in the RDC´s pattern from Friday to Saturday and from Saturday to Sunday, there does not exist any relation between these days and its previous days.
In order to avoid confusion, it is necessary to remark that the ARIMA model developed in Section III, which includes two seasonal differencings, does not discriminate the data series in time intervals, so a model which minimizes the **overall** RSE is obtained.

However, using the **WE** method, it is possible to see that depending on the time interval analysed, some of them do not need any differencing at all (like the ones detailed in tables 3 and 5), while others may need one (table 4) or more differencings.
Each time zone model also includes its RSE (residual standard error), whose magnitude will usually have a strong resemblance with the range of variation of the values of each time interval (i.e, the lower the output variable dispersion, the lower the value of the RSE).

This fact can be used to generate multiple RDC´s scenarios in order to build up the 24 optimal bid curves to be submitted to the MO.

D. A BRIEF COMMENT ON OUTLIERS

SCA provides the capability to detect the most common type of outliers (for example a level shift or a temporary change in the data series) and its deviation from the expected value (i.e, obtained in normal conditions) in order to evaluate their impact. The effects of outliers should always be considered, because they can bias the estimation of the parameters of the model. The problem is that the outlier detection procedure takes a lot of time if the data series has a big size, as is usually common with hourly data.
A possible approach to this topic would be to build auxiliary daily and weekly TF models, in which both input and output variables are the daily/weekly average values, so the size of the series is shortened. The daily model would detect those days where there has been an abnormal behaviour (usually non labor days different from the weekends) and the weekly model would be suitable for more longer events like Easter or Christmas.
Finally, as the RDC´s must be hourly generated, the average daily or weekly deviations should be converted into hourly deviations taking into account the different characteristics of each period of time.

V. CASE STUDY

In this case study the advantages of an accurate prediction of the variables slope and intercept corresponding to linearized RDC’s are going to be shown. Firstly, a week where there has been a step increase of run of river power with respect to the previous week has been chosen. It is expected that the RDC´s forecasts of the ARIMA model without explanatory variables described in Section III will deviate greatly from the real RDC´s curves, while the more complex weighted estimation model which uses the explanatory variables described in Section IV (denoted as WE in the next figures) should capture the change in the RDC pattern caused by the new hydro conditions. Figures 10, 11 and 12 display the different performance of the forecasting models in three different type of hours of a Wednesday (one valley hour, one plateau hour and one peak hour).
As it was expected, ARIMA model forecasts present big departures from the real RDC curves. In this case, forecasted RDC’s tend to be higher and more abrupt, whereas weighted estimation model forecasts match better to the real pattern.

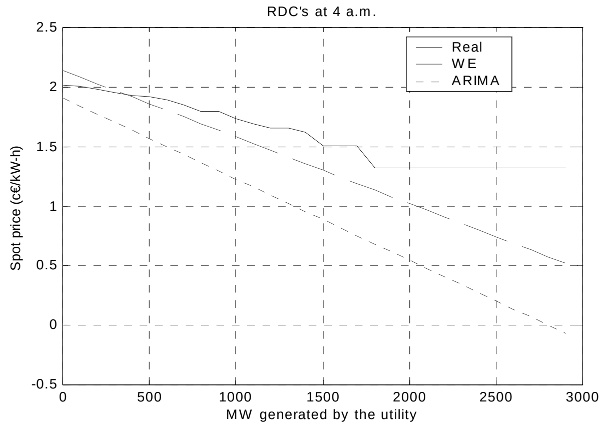

Figure 10.Real RDC (continuous line) versus WE forecast (dashed line) and ARIMA forecast (dotted line). Valley hour.

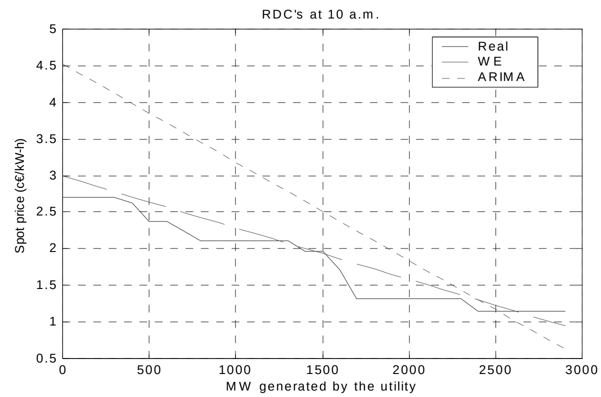

Figure 11.Real RDC (continuous line) versus WE forecast (dashed line) and ARIMA forecast (dotted line). Plateau hour.

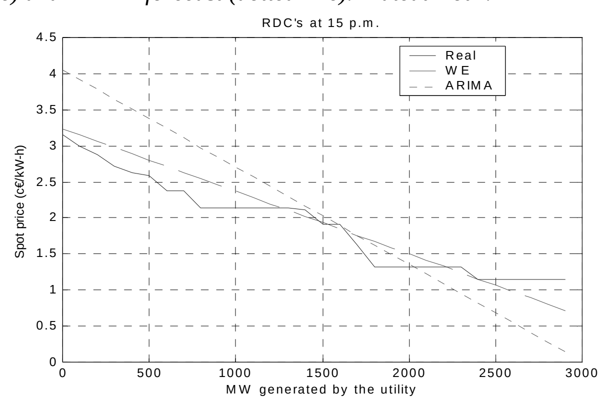

Figure 12.Real RDC (continuous line) versus WE forecast (dashed line) and ARIMA forecast (dotted line). Peak hour.

VI CONCLUSIONS

In this paper the advantages of an accurate prediction of the residual demand curves of a given utility have been shown. An adequate hourly explanatory variable has been defined which takes into account different situations of the national system. A weighted estimation method has been proposed in order to deal with the non linear relationship between the explanatory and the input variables, depending on the time period (day and hour) considered. Finally, a case study has been carried out, showing the advantages of the weighted estimation using the proposed explanatory variable.

VII ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of the staff at *Centro de Gestión de la Energía* of *Unión Fenosa Generación, S.A.* for their work and collaboration in this research. Also we appreciate the helping hand of SCA company members William Lattyak and Prof.Lon Mu Liu, who have provided great support and brilliant ideas.

VIII REFERENCES

[1] A. Pankratz,, " Forecasting with Dynamic Regression models", *John Wiley&Sons,1991.*

[2] George E.P.Box, Gwilym M.Jenkins " Time Series Analysis", *Holden-Day, 1976.*

[3] George E.P.Box, Gwilym M.Jenkins,Gregory C.Reinsel, " Time Series Analysis", *Prentice Hall,1994.*

[4] S.Makridakis et al, " Forecasting Methods and Applications", *John Wiley&Sons, third edition,1998.*

[5] Lon Mu Liu et al, " SCA Statistical System",*1997, Volumes 1 & 2.*

[6] Lon Mu Liu , " Effective Forecasting and Time Series Data Mining", *Course Material, Madrid, December 13-15,2001.*

[7] Lon Mu Liu, " Multivariate Time Series Analysis using VARMA models", *SCA publications,1996.*

[8] Binroth W. et al, " A Comparison of Commodity Price Forecasting by Box-Jenkins and Regression-based Techniques", *Technological Forecasting and Social Change 14,1979.*

[9] Espasa A. et al, "Automatic Modelling of Daily Series of Economic Activity", *Compstat 12th Proceedings in Computational Statistics, Barcelona ,1996.*

[10] Granger C. , " Forecasting stock market prices: Lesson for forecasters " ,*International Journal of Forecasting, vol 8, 1998* .

[11] Gelow M.E., " Economic evaluation of commodity price forecasting models ", *International Journal of Forecasting, vol 9, 1993..*

[12] Taylor S.J, " Forecasting market prices", *International Journal of Forecasting, vol 4, 1988..*

[13] Ip W.H., " Integration of Simulation and Expert System through Intervention Modeling", *Computer Integrated Manufacturing,3rd International Conference, Singapore.*

IX BIOGRAPHIES

Agustín Martín Calmarza was born in Las Palmas de Gran Canaria, Spain, in 1977 and has just obtained his Bachelor Degree in electrical Engineering from Universidad Pontificia Comillas (UPCo), Madrid, 2001. His interest areas include power system control, time series analysis and ancillary services topics.

José Ignacio de la Fuente was born in Soria, Spain, in 1967. He obtained both his B.S. and Ph.D. degrees in Electrical Engineering from Universidad Pontificia Comillas (UPCo), Madrid, in 1990 and 1997 respectively. Currently, he is a researcher at the Instituto de Investigación Tecnológica (IIT-UPCo). He also teaches Power Systems at the Engineering School ICAI, in the same university. His interest areas include power systems control and optimization, ancillary services procurement and remuneration, energy bidding strategies and power quality issues.