UNIVERSIDAD PONTIFICIA COMILLAS

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA (ICAI)

OFFICIAL MASTER'S DEGREE IN THE ELECTRIC POWER INDUSTRY

Master's Thesis

ANALYSIS OF ENERGY PRICES SEASONALITY

Author: **Mikel Vega Andrés**

Supervisor **Juan Ignacio Gil Gómez**

Madrid, July 2014

Master's Thesis Presentation Authorization

THE STUDENT:

Mikel Vega Andrés

……………………………………………….

THE SUPERVISOR

Juan Ignacio Gil Gómez

Signed:  Date: ……/ ……/ ……

THE CO-SUPERVISOR

Signed: …………………… Date: ……/ ……/ ……

Authorization of the Master's Thesis Coordinator

Dr. Javier García González

Signed: …………………… Date: ……/ ……/ ……

To my grandfather

For all the moments we spent drawing

INDEX

TABLE INDEX:

GRAPH INDEX:

1 ABSTRACT

The main problem the time series present, is that the only thing that can be analyzed at a glance is the evolution of the price as function of time. The purpose of this Master Thesis is to develop different econometric models in Python in order to be able to analyze the seasonality of seasonal time series.

The econometric models that have been coded are multiplicative moving averages, additive moving averages and trigonometric model. Once they were coded, we use an optimization algorithm to find the set of parameters that maximize the likelihood and, afterwards, the likelihoods are used in order to know which model fits better for each one of the underlying assets that were analyzed. The main difference between the moving averages and the trigonometric is that, while the former is a discrete model, the latter gives a continuous result. Different tests were also developed in order to check the goodness of the developed models.

The obtained results show the model that best fits to most of the underlying assets is the trigonometric one. It is the model that obtains the lowest likelihood for those underlying assets.

When the seasonal coefficients are extracted, they are employed to complement Monte Carlo simulations for options valuation. They can be used also to analyze futures curves and analyze whether the prices are really varying or is only a matter of the season of the year.

El principal problema que presentan las series temporales, es que a simple vista lo único que se puede analizar es la evolución del precio en función del tiempo. El objetivo de esta Tesis de Máster es desarrollar diferentes modelos econométricos en Python para ser capaces de analizar la estacionalidad de las series temporales.

Los modelos econométricos que han sido desarrollados son trigonométrico, medias móviles multiplicativo y medias móviles aditivas. Una vez que han sido desarrollados, se ha verificado la máxima verosimilitud para ajustar los modelos y, después, ha servido para saber qué modelo se ajusta mejor a cada activo subyacente que ha sido analizado. La principal diferencia entre los modelos de medias móviles y el trigonométrico reside en que, mientras que los primeros son modelos discretos, el segundo aporta un resultado continuo. También se han desarrollado diferentes pruebas para comprobar la bondad de los modelos desarrollados.

Los resultados obtenidos muestran que el modelo que mejor se ajusta a la mayoría de activos subyacentes es el trigonométrico. Es el modelo que obtiene la menor verosimilitud para los activos subyacentes que han sido analizados.

Cuando se obtienen los coeficientes estacionales, son empleados para complementar las simulaciones de Monte Carlo para valorar opciones. Además, también se pueden emplear para analizar curvas de futuros y de este modo saber si el precio está variando o simplemente es una consecuencia de la época del año en la que se está.

2 APPRECIATION

I would like to thank Ignacio for the great opportunity he gave me.

The challenge you proposed was not an easy task but with your help, finally, we did it. For me was a pleasure working with you and I am pretty sure we will come across again sooner or later.

I would like to thank Ana, the director of the trading department of Gas Natural Fenosa, because having the opportunity to work with such a great team is something that anybody would be proud of.

You all know that you can count on me for whatever.

3 INTRODUCTION

3.1 Motivation

3.1.1 Technical motivation

Through this Master Thesis, the main objective is to help the trading department by creating different tools in order to be able to analyze the seasonality of different assets.

Once these models have been developed, we have calibrated them looking for the parameters that maximize the likelihood, a Student T-Test is going to be applied to all of them in order to be able to decide for a moving averages model , whether any underlying asset has significant seasonality or not. When the probability given by the T-Test is below 0.05 does not mean that the seasonality is not significant, and whether the probability is above 0.05 means that the analyzed time series has enough seasonality and cannot be ignored.

Afterwards, the maximum likelihood is going to be computed in order to decide which one of models adjust better for each one of the underlying assets that are going to be analyzed.

After deciding which model fits better for each underlying asset, the obtained results will help within the trading department in the following issues:

- Improving a Monte Carlo model for pricing options
- Analyzing the futures curves.

Regarding the former issue, the way of improving this methodology is that once the seasonal coefficients are obtained for certain underlying asset, when the Monte Carlo simulation is launched now is possible to introduce the seasonality that was computed in the estimation.

Talking about the latter one, when you have to analyze the futures curve of the electricity or gas, you cannot know if the prices are going up or down just because of the effect of the season of the year. So a good way to analyze this problem is deseasonalizing the futures curve using the coefficients that were computed earlier and this way obtaining the real evolution of the price without being affect by the season.

The models that have been developed in Python are:

- Trigonometric model
- Multiplicative moving averages
- Additive moving averages

These models are going to be explained later in the document.

3.1.2 Business motivation

Gas Natural Fenosa is very well positioned in gas and power markets, what makes the company get additional margins in the markets through a proper trading activity, taking place throughout all the value chain of all the commodities related to its business.

The company owns a trading department which operates physical and financial products internationally, with some key targets:

- Managing the risk price of all the commodities of the businesses of Gas Natural Fenosa (gas, oil and its derivatives, power, coal and emission allowances).
- Contributes creating trading opportunities linked to the availability/flexibility of gas, power and coal assets of the company.
- Offer a high quality service to the clients of the company and third parties, providing opinion, products and expert management in the markets.

The main characteristics of these models are that give the possibility to analyze the seasonality of the electricity and gas prices.

3.2 Liberalization of the market

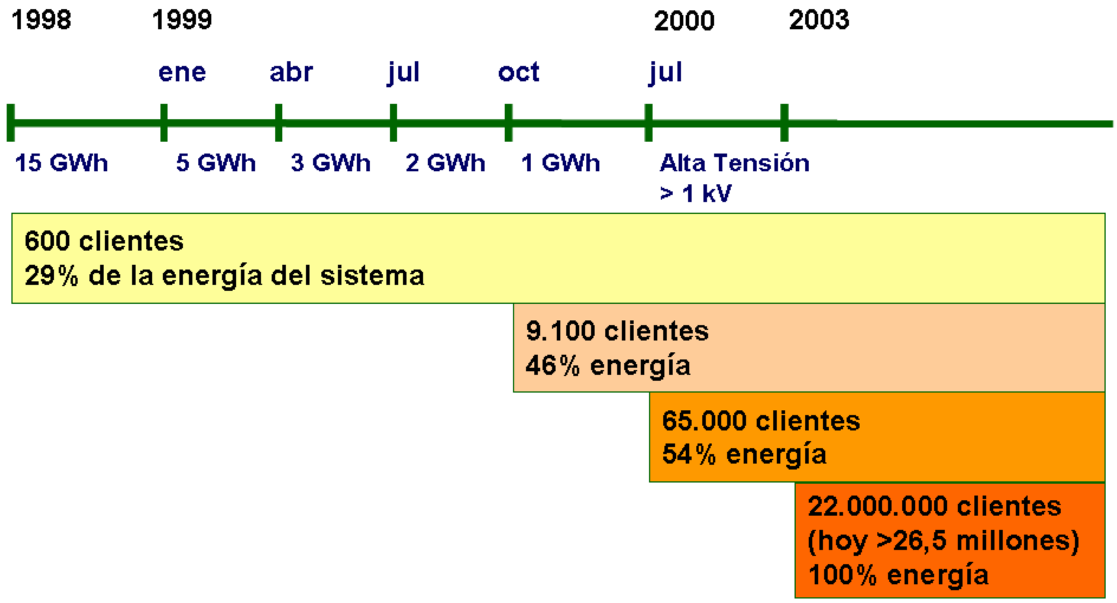
From 1998, when the Spanish electricity sector was liberalized, it has suffered a huge transformation. Up to that year, the companies were vertically integrated and were exercising monopoly in the different regions of Spain.

The unbundling of activities between the regulated activities and the non-regulated activities entered into force with the Law 54/1997 of the Electric Sector. The regulated activities from then on are the transmission and distribution of electricity, while the liberalized ones are the generation and retailing (activities that are carried out under free market conditions).

Now, the main activities in the power business are presented:

- Generation: the production of power using different technologies, such as coal power plants, combined cycle gas turbines, hydro, solar thermal…
- Transmission: carry the electricity from the production points to the consumption nodes.
- Distribution: carry the electricity from the final nodes to each consumption point.
- Retailing: is the activity of buying and selling electricity. They represent their customers in the wholesale market and acquire the electricity for them.

In the next picture we can see the process of the liberalization of the Spanish electricity market:



3.3 Type of markets

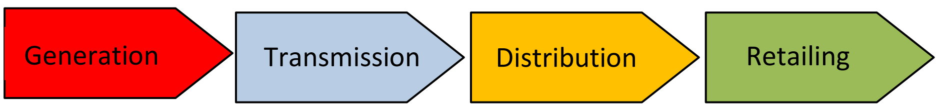
The liberalization of the Spanish electricity market starts with the Law 54/1997 of the power sector.

The Law enforces the European Union member states to adapt to European Union and Council Directive 96/92/CE. This one establishes the principles in order to create a healthy European electricity markets. That directive was modified by 2003/54/CE, which in the Spanish law is shown as Law 17/2007 in order to amend the 1997 Law and adapt to the current structure of the market.

The main goal of these rules is to achieve a market that reflects objectivity, transparency and free competition, creating a field for having free business initiative.

The fact of being such strategic sector provokes that the regulation of the sector must be very strong, for having a sector both economically and technically coordinated. It is also very important to make sure that having the liberalized market does not affect other pillars of the sector: energy efficiency, consumption reduction and the protection of the environment.

The main areas involved are these:



Now the main differences are going to be explained:

Generation: the goal is to produce electricity, and there are two different regimes of power generation plants

- Ordinary regime: power plants that use tradition technology for the production of electricity, such as: coal, hydraulic power, nuclear power…
- Special regime: photovoltaic, solar power, wind energy, biomass, cogeneration (up to 50 MW)…

Transmission: transport the energy through the high voltage lines. It is a regulated activity and only can be carried out by the Transmission System Operator (TSO). Is possible to find two different types:

- Primary: consists of lines, parks, transformers and other elements with nominal voltages equal to or greater than 380 kV or international, insular and extra peninsular interconnection.
- Secondary: consists of lines, parks, transformers and other elements with nominal voltages equal to or greater than 220 kV or lower voltage installations with transportation functions.

Distribution: the main function is to transport the electricity to the end-users. They are in charge of building, maintaining and operating the distribution gris and also the metering. They must ensure the following services:

- Guaranty the power supply.
- Guaranty the quality of supply.
- Make this at the lowest possible cost.

Retailing: buy the energy in the wholesale market and sell to the end-users. Is possible to distinguish two different agents (in the table below is possible to see the consumers that can contract power with the last resort suppliers):

- Free market retailer: the customer negotiates with him a free price that will include the cost of the energy, the cost of the regulated activities and a little margin for the retailer.
- Last resort supplier: only customers with special characteristics can contract the power with these retailers

Type of customer	Means of signing up for power supply
Low voltage customer (<1000 V) with contracted power of <=10 kW (almost all domestic customers).	Supplier of last resort. Free market supplier.
High voltage and low voltage customers (<1000 V) with contracted power of >10 kW.	Free market supplier

Table 1 Supplier and consumer types

OMIE manages the wholesale electricity market (referred to as cash or “spot”) on the Iberian Peninsula. Like any other, the electricity market caters for the trading of electricity between agents (producers, consumers, retailers, etc.) at a price that is known, transparent and accessible.

3.3.1 Day ahead market

The electricity prices in Europe are established at 12:00 for the 24 hours of the following day, and this is what is known as day-ahead market. The price and volume to supply is set at the point where the supply and demand curves match.

This is a marginal market, where the price is set as the cost of supplying one extra MW, and the same methodology is applied for all Europe. The matching algorithm that is used in the process is called EUPHEMIA. The algorithm is already working in Spain and Portugal, France, Germany, Austria, Belgium, The Netherlands, Luxemburg, Finland, Sweden, Denmark, Norway and the United Kingdom. In the short run is going to be implemented also in Italy, the Czech Republic, Hungary and Rumania.

The traders, independently of where they are, either in Spain or Portugal, must bid in the market, and the bids will be accepted according to the merit order until the interconnection capacity between Spain and Portugal becomes congested. Whether the capacity of the line is not congested, the price will be the same both for Spain and Portugal (situation that happened the 89% of the time in 2013). And whether the capacity of the interconnection is congested in a certain hour, the EUPHEMIA algorithm is run for the two countries separately, so they will have different prices (situation that happened only the 11% of the times in 2013).

When the two countries are separated is called market splitting.

When the economic efficiency is taken into account, the results obtained coming from the free trade among agents is the optimal one. Due to the particularities of the electricity, apart from the economic point of view, it is also important to have a feasible schedule for the power generation according the technical constraints. Once the feasible schedule is obtained (regarding the economic point of view), the Transmission System Operator must approve the technical part (process known as management of the system’s technical limitations) and makes sure that the market results can be accommodated considering the technical part of the networks. This means that the results that were obtained in the market the previous day can be modified in order to solve the technical constraints of the system. These variations affect to around 5% of the traded volume. As new constraints are introduced in the system, these situations lead to higher prices than the ones obtained in the day-ahead market.

3.3.2 Intraday market

Once the day-ahead market has been cleared and the production scheduled according to the technical constraints, the market operators can buy and sell energy in the intraday market for correcting deviations in the schedule.

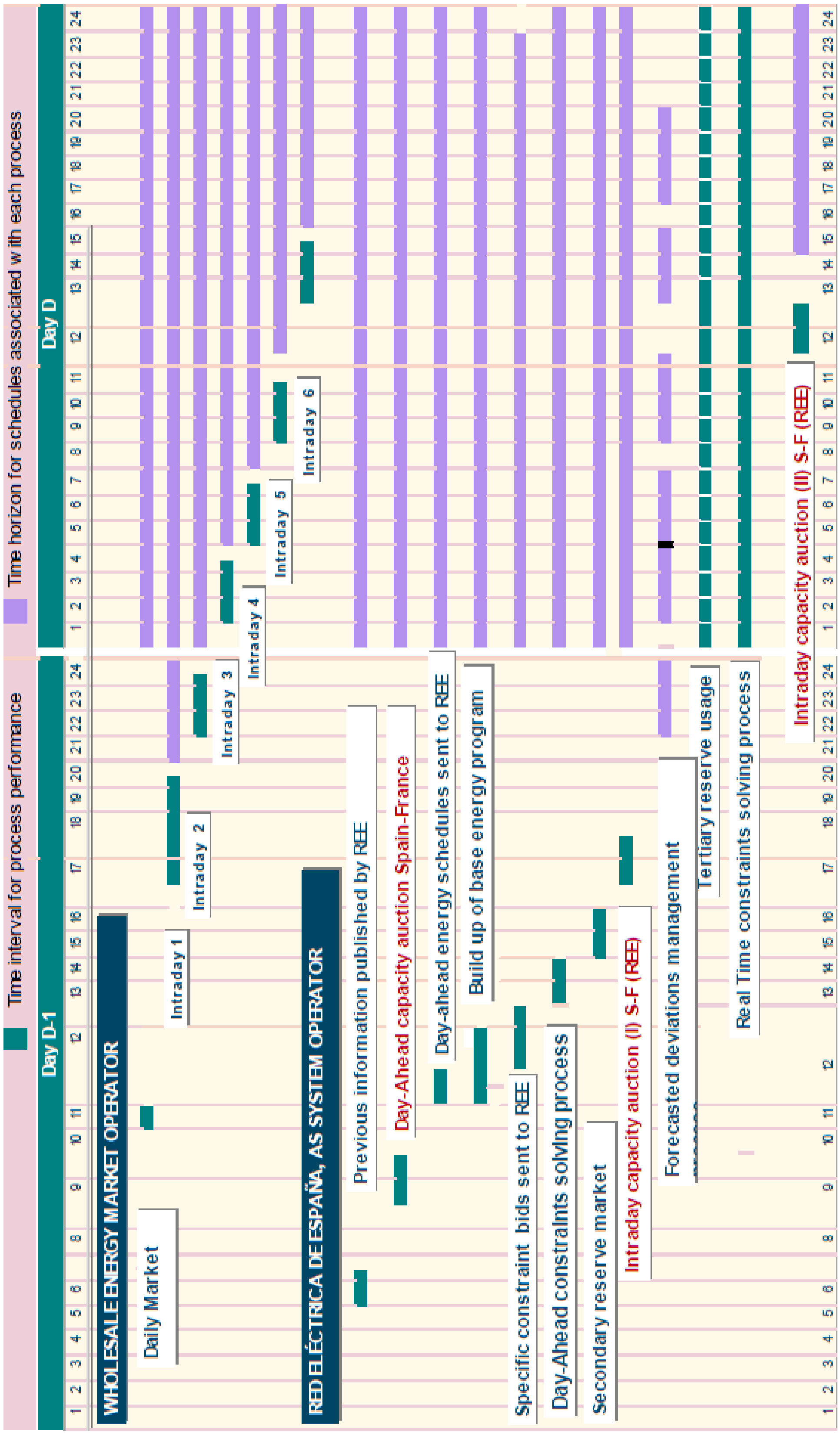
In this market they have six trading sessions with a scheme equal to the day-ahead market, and the resulting price is cleared where the supply and demand curves are met.

The MIBEL’s intraday market makes it the most liquid within European Union, so this makes it easier to adjust the positions for the market operators. Other characteristic features is that the prices obtained are similar to the ones obtained in the day-ahead, making the MIBEL intraday market one of the most competitive within Europe. You can trade in the intraday market electricity 4 hours in advance of the real time operation.

	SESSION 1	SESSION 2	SESSION 3	SESSION 4	SESSION 5	SESSION 6
Session opening	17:00	21:00	01:00	04:00	08:00	12:00
Session closing	18:45	21:45	01:45	04:45	08:45	12:45
Matching results	19:30	22:30	02:30	05:30	09:30	13:30
Reception of breakdowns	19:50	22:50	02:50	05:50	09:50	13:50
Publication PHF	20:45	23:45	03:45	06:45	10:45	14:45
Schedule Horizon (Hourly periods)	27 h (22-24)	24h (1-24)	20h (5-24)	17h (8-24)	13h (12-24)	9h (16-24)

Table 2Intraday market sessions

In the following table is shown the overall view of the day-ahead and intraday markets:



Graph 1 “Ancillary services in Spain: Dealing with high penetration of RES” de la Fuente, Ignacio

3.3.3 Ancillary services

Are very frequent experiencing short unbalances in the operation of the system, and in order to correct them.

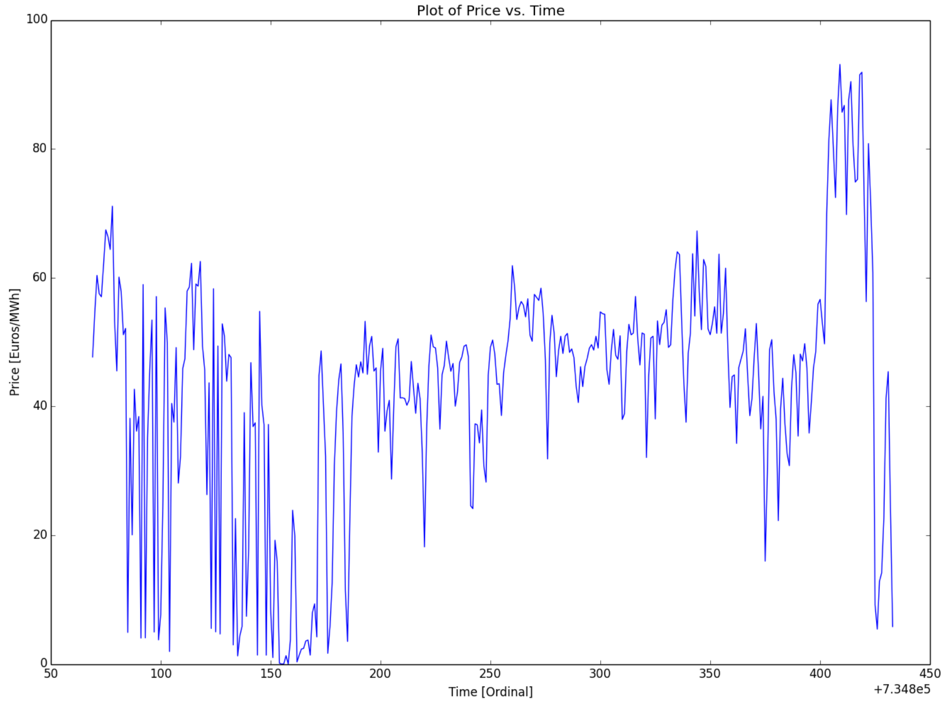
Here are going to be presented the different types of balancing services:

- Primary regulation: action of speed regulators from generator units responding to changes in system frequency (less than 30 seconds)
- Secondary regulation: automatic and hierarchical control that faces changes in system frequency and power deviations with respect to France-Spain exchange program (less than 100 seconds).
- Tertiary regulation: manual power variation with respect to a previous program in less than 15 minutes
- Slow reserve: running reserves of connected thermal units (39 minutes to 4-5 hours)

4 GRAPHICAL REPRESENTATION AND COMPONENTS OF A SERIE

The graphic representation of a time series is done using the Cartesian axis. On the abscissa is represented the time (t) and on the ordinate the values of the observed magnitude (y_t), with which is obtained a series of points, and when these points are linked is possible to obtain the first conclusions of the historical evolution of the series.

In the following graph is shown the evolution of the daily spot electricity prices from Jan 1 2013 to December 31 2013.



Graph 2 Spanish spot prices from January 1st 2013 to December 31st 2013

From the classical point of view, a series is made up by four components:

- Trend (T): this is a component that shows the evolution of the series in the long run. It can be of a seasonal component (it would be represented by a straight line parallel to abscissa axis), of lineal nature (increasing or decreasing), of parabolic nature, exponential nature and so on and so forth.

- Cycle (C): is a component of the series that shows the periodic oscillations with higher amplitude than a year, and these are due to the alternation process between economic growth and contraction. The higher the period of a cycle, the higher the number of observations of the variable for being able to recognize it. In the Master Thesis we are not taking into account this component.[1]

[1] The reason for not taking into account this component is that the electricity markets are very young and there are not enough number of years in order to have a reliable model of these cycles. The

- Seasonal (S): it shows the oscillations happening in periods lower or equal than a year regularly repeated throughout the years. Whether is considered a year as repetition framework, is possible to see fluctuations throughout the months, quarters, … The origin of the seasonal variations comes from the climate or cultural factors, such as Christmas holidays, summer holidays… For example, the climate affects to the demand, such as in Texas, where in summer the consumption increases around 30 GW due to the installed air conditioning. And from the production side the same happens with the RES affecting the production mix depending on the seasonal climate characteristics.

- Residue (R): this component gathers the erratic fluctuations of the time series due to not only unforeseen phenomena, but also the activity of the markets and the clearing of the demand and offer. Both are influenced by such a big amount of unpredictable factors that they are presented as a random variable. For example, talking about the electricity, climate catastrophes.

Once the different components of a time series have been defined, we can now explain the different ways in which they can be combined in order to get the classical two schemes of the time series:

- Additive scheme: The observed values of any time series are the result of adding its four components (although we are not taking into account C_t in our models)

$$(1) \qquad y_t = T_t + C_t + S_t + R_t$$

- Multiplicative scheme: The observed values of any time series are the result of multiplying its four components

$$(2) \qquad y_t = T_t C_t S_t R_t$$

In the multiplicative scheme, it is possible to have a variation, better known as mix model, which is used when it is assumed that the residual component is independent from the others. In the mix model, each one of the observed values is obtained by adding the residual component and the product of the trend, cyclic and seasonal component (and this is the way we decided to develop it):

$$(3) \qquad y_t = T_t C_t S_t + R_t$$

evolution of the markets, switch in technology, transnational interconnection… make these changes in the long run more evolutional than cyclical.

5 WHAT OTHERS DID

As Lucia & Schwartz (2006) pointed out, "*the number of papers addressing the specific valuation problems of electricity contracts is still scarce. […]. Others, on the contrary, have addressed the importance of the periodic seasonal behavior of electricity prices. […]. Nevertheless, the much needed empirical work is on its early stages. Only very limited and tentative work has been published to date…*"

In their research, they say that the systematic behavior of electricity prices through time can be explained by changes in demand following the business activity, and the periodic behavior of consumption arising from the seasonal evolution of temperatures.

For their analysis they used a one-factor model (with a deterministic part and a stochastic part) and a two-factor model. With the two-factor model they solved the problem presented when using the one-factor model: that they are assuming the spot price and the futures price are perfectly correlated.[2]

They found that all the years, in the late April, as the snow season is over, the uncertainty about reservoir levels is solved and they saw a violent downward jump.

Cartea & Figueroa (2005) presented a paper in which they captured the main characteristics of electricity spot prices such as seasonality. What they did for stripping the seasonality was following a common approach which is to subtract the mean of every day across the series according to

$$R_t = r_t - r_d \tag{4}$$

Where the deseasonalized return at time t is R_t, r_t is the return at time t and r_d is the corresponding mean (throughout the series) of the particular day r_t represents.

Within the research that Geman and Roncorini (2006) did, they pointed out three main characteristics for the electricity:

- Mean reversion in electricity prices toward a level that represents marginal cost and may be constant.
- The existence of small random moves around the average trend, which could be due to temporary supply and demand imbalances in the system. This is unpredictable and could be represented as white-noise.
- The spikes of the electricity prices are an intrinsic characteristic.

[2] With the one-factor, two-factor… they mean the number of random variables used in order to describe an asset.

6 COMPUTATIONAL TOOLS

In order to develop these models, the programming language that has been used is Python. Basically, this was the used language because is very useful for the mathematical calculus and is very used in the industry.

Python is an interpreted, interactive, object-oriented programming language. One of the strengths of Python is that program development using Python is said to be 5-10 times faster than with C/C++ and 3-5 times faster than using Java. A nice mathematical library is “SciPy”. It gives a decent computer algebra system within python. It also offers the same syntax as Matlab plus the power of the active “NumPy” community. There is also a widely used in the industry open source toolkit called “Pandas” very useful for financial time series.

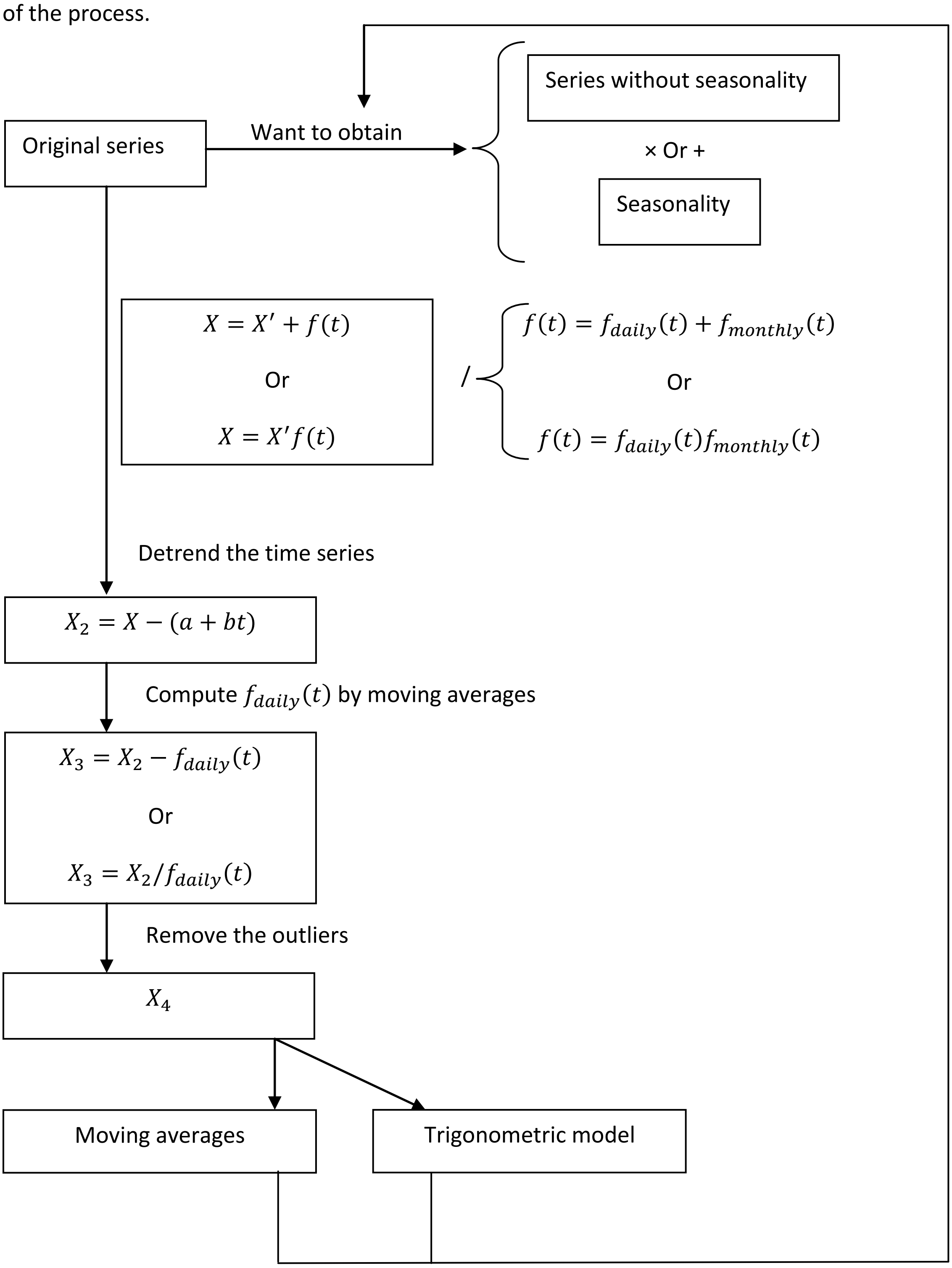
Gas Natural Fenosa has a business intelligence system which contains an Oracle database where all the market data are stored. It was necessary to develop from scratch a tool to access this database and execute SQL commands from Python functions, to feed the seasonality models with the market prices time series

Apart from all the previous advantages that Python presents, it is also quite straightforward to interact with Excel sheets for saving the results.

It is also easy to interact with Matlab, and this way the results obtained by these models can be fed into the models that already are developed in Matlab.

7 DESCRIPTION OF THE FOLLOWED PROCEDURE

In this section the steps that have been applied will be explained in order to have an overall view of the process.



8 DESCRIPTION OF THE DIFFERENT METHODOLOGIES IMPLEMENTED

One of the main reasons for using these methods and not others is that, for example, the Autocorrelation Function is not the best method due to the high volatility of the time series related to the electricity.

8.1 Initial procedure

In this section are going to be explained the first treatments that have to be applied to the original time series in order to be ready to feed the models. This way, it is going to be possible to start the computations.

The first step that must be applied is for removing the trend. For doing so, we compute a linear regression and we subtract it to the data, like this: $X_2 = X - (a + bt)$

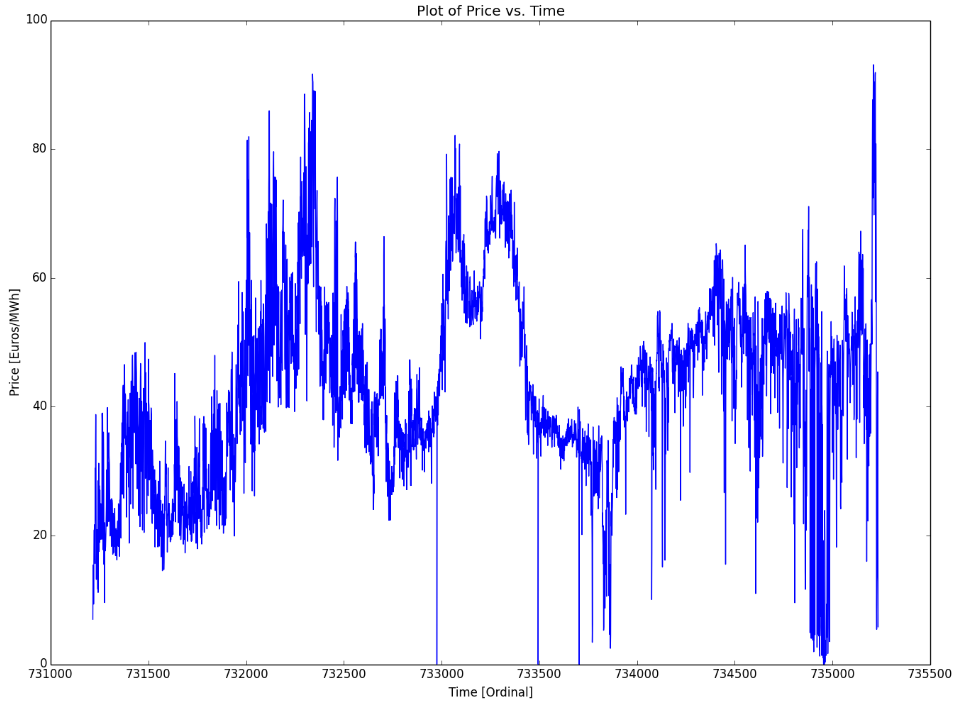
Secondly $f_{daily}(t)$ is computed and finally the spikes are located and removed. The steps follow this order because removing the spikes before computing $f_{daily}(t)$ could distort the daily coefficients. The distortion could happen due to the way we chose for removing the spikes (substituting the value of one day by its average).

8.1.1 Removing spikes

One of the first things that must be done is the following: decide a threshold for removing the spikes from the original time series, following Saz-Micheler-Görlich. They comment that the first thing they do is setting a threshold in order to remove the spikes, because they distort too much the series. We believe that depending on how you do it, affect too much to the final price you will get, so we started analyzing the different results obtained using different thresholds.

For removing them, we go one by one, computing the corresponding average for each day taking into account the previous three days and the following three. In the paper mentioned above, they do it taking into account only the previous and the following day, and the threshold they set is chosen arbitrarily.

In the next graph, the daily spot prices are shown for the period Jan 1 2003-Dec 31 2013.



Graph 3 Spanish spot prices from January 1st 2003 to December 31st 2013

These spikes (such as the ones that can be seen in the above graph from the period January 1, 2003 until December 31, 2013) happen because of the problems that Cuaresma J. C. et al (2004) say in their paper. Common originators of these spikes are used to being unplanned outages of large power plants or high increases from the demand side, either due to heat waves in summer time or cold waves in winter. Another problem that the market agents must face is the possibility of having network constraints that can cause spikes, because they are moving away from perfect competition to oligopolystic market in which they can exercise market power.

The different limits that were used are the following ones.

- Without removing any spike, that is to say, using the original time series (in this case the threshold is set in 180, because is the price cap set by the market operator).
- Setting the limit in 30.
- Setting the limit in 15.
- Setting the limit in 10.
- Setting the limit in 5.

What does "Setting the limit in…" mean? This means that whether the real price differs more than 5 €/MWh (depending on each limit) from the mean for that day, the real price is going to be substituted by its mean for that day, leading to the elimination of the spikes.

Just as an example, in the following graph is possible to see how this procedure works. In cyan is shown the original time series, and as it is not out of the limit set in 30€, is kept its original value. But the day 2448, the price goes down up to the value zero, and as it moves further than 30, is substituted by its average for that day.

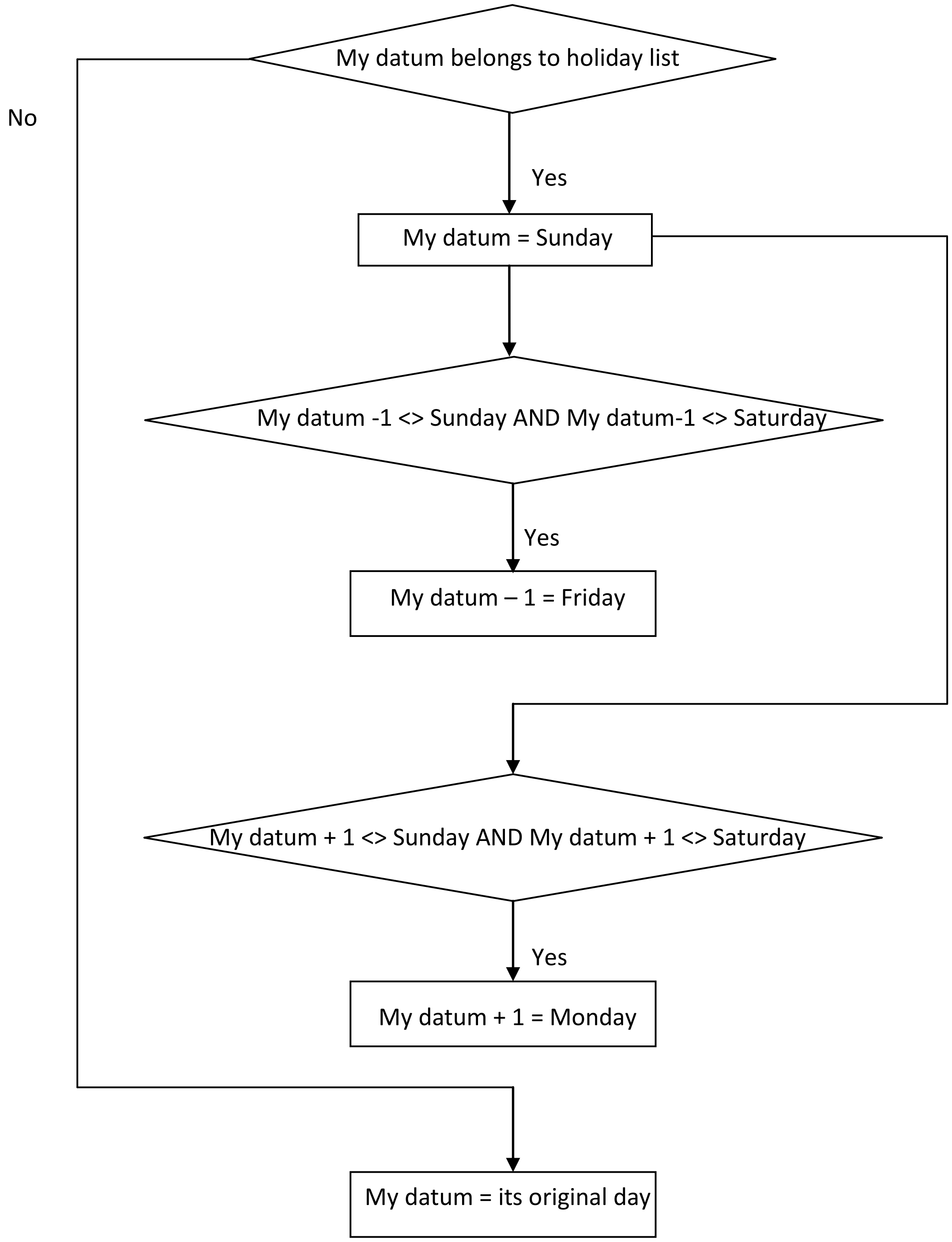
In order to carry out this research project, the threshold has been set at 30€/MWh.

Graph 4 Removing the spikes

8.1.2 Reassigning the days of the week

For carrying out this step, the first thing that was done was getting a calendar from the data base with the holidays. Once this calendar was obtained, the next step is to try one by one whether each one of your datum is in the list of holidays. The objective of doing this is to treat the bank holidays as Sundays, the days before as Fridays and the following days as Mondays.

The flowchart of the procedure is as follows:



Graph 5 Flowchart for reassigning the days of the week

8.2 Dummy Variables representation

The deterministic seasonality describes behavior in which the unconditional mean of the process varies with the season of the year.

First of all a brief introduction will be done on the dummy variable representation, and after this it is going to be explained the trigonometric model in depth.

The conventional dummy variable representation of seasonality, as very frequently used in applied econometric studies, can be written as:

$$(5) \qquad y_t = \sum_{s=1}^{S} \gamma_s \delta_{st} + z_t \ \forall \ \ \text{t=1... Tt}$$

Where z_t is a weakly stationary stochastic process with mean zero and δ_{st} (s=1…S) are seasonal dummy variables. Thus, for season s of year t,

$$(6) \qquad \mathrm{E}(y_{st}) = \gamma_s$$

The property of (18) is the one of primary interest for deterministic seasonality, as it implies that the process has a seasonally shifting mean. Because of this time-varying mean, the process y_t is not stationary. This nonstationarity is often overlooked because it is simple to remove by subtraction of the mean for each season. Thus, the deviations $y_t - E(y_t) = z_t$ are weakly stationary, so that the second order properties of these deviations are invariant to the season s in addition to the year t.

One disadvantage of the simple dummy variable representation of (5), however, is that it mixes seasonality and the overall mean when the latter is nonzero. Since a proportion 1/S of observations relates to each of the S seasons, then the overall mean of y_t is

$$(7) \qquad E(y_t) = \mu = \frac{1}{S}\sum_{s=1}^{S} \gamma_s$$

When this overall mean is extracted, the deterministic seasonal effect for season s is $m_s = \gamma_s - \mu$. Obviously, this definition imposes the restriction $\sum_{s=1}^{S} m_s = 0$, with the natural interpretation that there is no deterministic seasonality when observations are summed over a year. When the level of the series (represented here by μ) is separated from the seasonal component, then

$$(8) \qquad y_t = \mu + \sum_{s=1}^{S} m_s \delta_{st} + z_t$$

It might be noted that the discussion of deterministic seasonality implicitly assumed μ=0.

When μ is replaced by $\mu_0 + \mu_1 t$, the last equation can be generalized to include a trend component that is constant over the seasons. A further generalization would allow a separate trend for each season, so that

$$(9) \qquad y_t = \mu_0 + \mu_1 t + \sum_{s=1}^{S} (m_{0s} + m_{1s} t)\delta_{st} + z_t$$

Once again, the seasonal coefficients are restricted so that $\sum_{s=1}^{S} m_{0s} = \sum_{s=1}^{S} m_{1s} = 0$ to ensure that the seasonality sums to zero over the year. Such trending deterministic seasonality may, however, be unrealistic in practice for many economic series, since it implies that observations for the seasons of the year diverge over time. Therefore, any trend included is typically taken to be non-seasonal.

Whether in the process (8) or (9), one implication of stationarity for $z_t = y_t - E(y_t) = y_{st}E(y_{st})$ is that each observation deviates from its respective seasonal mean with a variance that is constant over both s and t. In this sense, the deterministic seasonal process ensures that the observations cannot “wander” too far from their underlying mean.

8.3 Moving averages

In the moving averages models what is being done is the following procedure.

Being S’ the original time series, S the seasonally adjusted time series and f(t) the seasonality, for the multiplicative model is as follows: $S = S'f(t)$ Where $f(t) = f_{daily}(t)f_{monthly}(t)$

And for the additive model, $S = S' + f(t)$ where $f(t) = f_{daily}(t) + f_{monthly}(t)$

8.3.1 Multiplicative

In this section how the moving averages method has been developed is going to be explained. The steps that have been followed are the following:

1) Obtain the original time series y_t
2) Compute the arithmetic averages for each y_t
    - If p is odd → the averages are centered in time

$$y_{\frac{p+1}{2}} = \frac{y_1+y_2+\cdots+y_p}{p} \quad (10)$$

$$y_{\frac{p+3}{2}} = \frac{y_2+y_3+\cdots+y_{p+1}}{p} \quad (11)$$

$$y_{\frac{p+5}{2}} = \frac{y_3+y_4+\cdots+y_{p+2}}{p} \quad (12)$$

    - If p is pair → the averages are not centered in time

(13) $$y_{\frac{p+2}{2}} = \frac{y_{\frac{p+1}{2}} + y_{\frac{p+3}{2}}}{2}$$

(14) $$y_{\frac{p+4}{2}} = \frac{y_{\frac{p+3}{2}} + y_{\frac{p+5}{2}}}{2}$$

(15) $$y_{\frac{p+6}{2}} = \frac{y_{\frac{p+5}{2}} + y_{\frac{p+7}{2}}}{2}$$

3) Determine the trend by the moving averages method

(16) $$Trend_t = \frac{y_t + y_{t-1}}{2}$$

4) Remove the trend component from the previously computed arithmetic averages

(17) $$\frac{y_t}{T*C} = \frac{TCER}{TC} = ER$$

5) Remove the residue by computing the arithmetic average for each season

6) Compute the yearly average

7) Obtain the seasonal component.

(18) $$I_1 = \frac{M_1}{MA}100$$

(19) $$I_2 = \frac{M_2}{MA}100$$

8) For finally obtaining the seasonally adjusted time series each value must be divided by its own coefficient

Now, a simple example is going to be used in order to explain more clearly the multiplicative methodology. During the process quarterly data of three years is going to be used, although for our models monthly data was used. In this example quarterly data is used basically because the process is shorter.

The first step is to compute the average of all the periods of the first year, and afterwards you must move one step forward as it is represented in the picture of the right, just as it is shown in the next pictures.

Quarter	Serie	Off-centered	Trend	E x R
Q1 14	150			
Q2 14	165	152,5		
Q3 14	125	153,75	153,125	0,81632653
Q4 14	170	155	154,375	1,10121457
Q1 15	155	157,5	156,25	0,992
Q2 15	170	156,25	156,875	1,08366534
Q3 15	135	157,5	156,875	0,86055777
Q4 15	165	160	158,75	1,03937008
Q1 16	160	161,25	160,625	0,99610895
Q2 16	180	165	163,125	1,10344828
Q3 16	140			
Q4 16	180			

Quarter	Serie	Off-centered	Trend	E x R
Q1 14	150			
Q2 14	165	152,5		
Q3 14	125	153,75	153,125	0,81632653
Q4 14	170	155	154,375	1,10121457
Q1 15	155	157,5	156,25	0,992
Q2 15	170	156,25	156,875	1,08366534
Q3 15	135	157,5	156,875	0,86055777
Q4 15	165	160	158,75	1,03937008
Q1 16	160	161,25	160,625	0,99610895
Q2 16	180	165	163,125	1,10344828
Q3 16	140			
Q4 16	180			

Table 3 Computing off-centered series

Once the averages have been computed, is possible to find the trend of the time series, and for so the average of two consecutive values of the off-centered series are used. When the first one is computed, you must move forward as it is shown in the picture of the right. You have to repeat the process up to the last value.

Quarter	Serie	Off-centered	Trend	E x R
Q1 14	150			
Q2 14	165	152,5		
Q3 14	125	153,75	153,125	0,81632653
Q4 14	170	155	154,375	1,10121457
Q1 15	155	157,5	156,25	0,992
Q2 15	170	156,25	156,875	1,08366534
Q3 15	135	157,5	156,875	0,86055777
Q4 15	165	160	158,75	1,03937008
Q1 16	160	161,25	160,625	0,99610895
Q2 16	180	165	163,125	1,10344828
Q3 16	140			
Q4 16	180			

Quarter	Serie	Off-centered	Trend	E x R
Q1 14	150			
Q2 14	165	152,5		
Q3 14	125	153,75	153,125	0,81632653
Q4 14	170	155	154,375	1,10121457
Q1 15	155	157,5	156,25	0,992
Q2 15	170	156,25	156,875	1,08366534
Q3 15	135	157,5	156,875	0,86055777
Q4 15	165	160	158,75	1,03937008
Q1 16	160	161,25	160,625	0,99610895
Q2 16	180	165	163,125	1,10344828
Q3 16	140			
Q4 16	180			

Table 4Computing the trend

The following step is to compute the seasonal coefficient for each one of the values that are been analyzed. For so, you must divide the original value of the time series (the red rectangle on the left) by the trend (the right rectangle on the right), and you get as result ExR. In the next pictures the process is shown for the first two values, but you must apply this to all of them.

Quarter	Serie	Off-centered	Trend	E x R
Q1 14	150			
Q2 14	165	152,5		
Q3 14	125	153,75	153,125	0,81632653
Q4 14	170	155	154,375	1,10121457
Q1 15	155	157,5	156,25	0,992
Q2 15	170	156,25	156,875	1,08366534
Q3 15	135	157,5	156,875	0,86055777
Q4 15	165	160	158,75	1,03937008
Q1 16	160	161,25	160,625	0,99610895
Q2 16	180	165	163,125	1,10344828
Q3 16	140			
Q4 16	180			

Quarter	Serie	Off-centered	Trend	E x R
Q1 14	150			
Q2 14	165	152,5		
Q3 14	125	153,75	153,125	0,81632653
Q4 14	170	155	154,375	1,10121457
Q1 15	155	157,5	156,25	0,992
Q2 15	170	156,25	156,875	1,08366534
Q3 15	135	157,5	156,875	0,86055777
Q4 15	165	160	158,75	1,03937008
Q1 16	160	161,25	160,625	0,99610895
Q2 16	180	165	163,125	1,10344828
Q3 16	140			
Q4 16	180			

Table 5 Computing seasonality

After computing ExR, the average of each one of the seasons must be cleared. In order to do so, you must compute the average of all the Q1, Q2, Q3 and Q4. Below are indicated for the Q3 and Q4.

E x R
0.81632653
1.10121457
0.992
1.08366534
0.86055777
1.03937008
0.99610895
1.10344828

M Q3	0.83844215
M Q4	1.07029233
M Q1	0.99405447
M Q2	1.09355681

E x R
0.81632653
1.10121457
0.992
1.08366534
0.86055777
1.03937008
0.99610895
1.10344828

M Q3	0.83844215
M Q4	1.07029233
M Q1	0.99405447
M Q2	1.09355681

Table 6 Computing averages of each period

Before finishing, the average of the whole year is computed.

M Q3	0.83844215
M Q4	1.07029233
M Q1	0.99405447
M Q2	1.09355681

MA =	0.99908644

Table 7 Computing the average of the year

The last step is to divide the average of the quarter by the average of the year and the seasonal coefficients are obtained.

M Q3	0.83844215
M Q4	1.07029233
M Q1	0.99405447
M Q2	1.09355681
MA =	0.99908644
I 3 =	0.83920882
I 4 =	1.071271
I 1 =	0.99496343
I2 =	1.09455675

Table 8 Computing the final coefficients

When the coefficients have been obtained, you already can seasonally adjust the original time series dividing the original values by the coefficients:

Quarter	Original time series	Deseasonalized series
Q1 14	150	150,759309
Q2 14	165	150,745953
Q3 14	125	148,949817
Q4 14	170	158,690005
Q1 15	155	155,78462
Q2 15	170	155,314012
Q3 15	135	160,865803
Q4 15	165	154,022652
Q1 16	160	160,80993
Q2 16	180	164,45013
Q3 16	140	166,823795
Q4 16	180	168,024711

Table 9 The seasonally adjusted series

In the case of the research that has been carried out, the data that has been used are daily prices, so after applying the methodology that has been explained above, the days were gathered according to their month and afterwards the monthly coefficients were computed.

So finally, for adjusting seasonally the original time series with this model, you have to divide the original datum by the coefficient of the corresponding month, and afterwards divide it by the coefficient of the corresponding month.

8.3.2 Additive

The first three steps are the same as for the multiplicative method, and after applying them, the next one must be applied

1) Remove the trend and the cyclical component, subtracting to the original time series the trend component (remember that in the research the cyclic component has not been taken into account, but it is going to be analyzed in the future)

$$y_t - (T + C) = T + C + E + R - (T + C) = E + R \quad (20)$$

2) Remove the residue by computing the arithmetic average for each season. Whether the time series is composed by monthly data, they will have twelve averages (M1, M2…).

3) The yearly average must be computed (MA).

4) The seasonal component is computed subtracting to each average of each season the yearly average

$$E1=M1\text{-}MA \;;\; E2=M2\text{-}MA\ldots \quad (21)$$

5) The original time series is seasonally adjusted

Now a brief example is going to be developed with the previous time series. The process up to the point of computing the trend is the same as for the multiplicative method. The first difference arises when computing ExR. In this case you have to subtract to the original time series the value of the trend.

Quarter	Time series	Off-centered	Trend	E x R	Quarter	Time series	Off-centered	Trend	E x R
Q1 14	150				Q1 14	150			
Q2 14	165	152,5			Q2 14	165	152,5		
Q3 14	125	153,75	153,125	-28,125	Q3 14	125	153,75	153,125	-28,125
Q4 14	170	155	154,375	15,625	Q4 14	170	155	154,375	15,625
Q1 15	155	157,5	156,25	-1,25	Q1 15	155	157,5	156,25	-1,25
Q2 15	170	156,25	156,875	13,125	Q2 15	170	156,25	156,875	13,125
Q3 15	135	157,5	156,875	-21,875	Q3 15	135	157,5	156,875	-21,875
Q4 15	165	160	158,75	6,25	Q4 15	165	160	158,75	6,25
Q1 16	160	161,25	160,625	-0,625	Q1 16	160	161,25	160,625	-0,625
Q2 16	180	165	163,125	16,875	Q2 16	180	165	163,125	16,875
Q3 16	140				Q3 16	140			
Q4 16	180				Q4 16	180			

Table 10 Computing seasonality

After computing all the ExR values, you have to obtain the average for each season, as shown in the next pictures for M Q3 and MQ4:

E x R
-28,125
15,625
-1,25
13,125
-21,875
6,25
-0,625
16,875

M Q3	-25
M Q4	10,9375
M Q1	-0,9375
M Q2	15

E x R
-28,125
15,625
-1,25
13,125
-21,875
6,25
-0,625
16,875

M Q3	-25
M Q4	10,9375
M Q1	-0,9375
M Q2	15

Table 11 Computing averages of each period

When the coefficients are computed, it is time to seasonally adjust the original time series subtracting to it the seasonal coefficients:

Quarter	Original time series	Deseasonalized series
Q1 14	150	150,93750
Q2 14	165	150,00000
Q3 14	125	150,00000
Q4 14	170	159,06250
Q1 15	155	155,93750
Q2 15	170	155,00000
Q3 15	135	160,00000
Q4 15	165	154,06250
Q1 16	160	160,93750
Q2 16	180	165,00000
Q3 16	140	165,00000
Q4 16	180	169,06250

Table 12 The seasonally adjusted series

8.4 Trigonometric model

Although perhaps less intuitive, it is may be useful to adopt the trigonometric representation of deterministic seasonality. The deterministic seasonal process can equivalently be written in terms of trigonometric functions[3]:

$$(22) \qquad y_t = \mu + \sum_{k=1}^{S/2} [\alpha_k \cos\left(\frac{2\pi kt}{365.25}\right) + \beta_k \sin\left(\frac{2\pi kt}{365.25}\right)] + z_t$$

Which also explicitly recognizes the overall mean μ. Notice that is needed to consider $\alpha_k \ \forall \ k = 1, \dots, \frac{S}{2}$ but $\beta_k \ \forall \ k = 1, \dots, \frac{S}{2} - 1$, since $\beta_{S/2}$ multiplies a sine term that is always zero. This representation is the basis of spectral analysis of seasonality and seasonal adjustment; an early reference to this is Hannon, Terrell and Tuckwell (1970).

For the case of quarterly data, (22) implies that the seasonal dummy variable coefficients of (5) are related to the deterministic components of the trigonometric representation by

$$(23) \qquad \gamma_1 = \mu + \beta_1 - \alpha_2$$

$$(24) \qquad \gamma_2 = \mu - \alpha_1 + \alpha_2$$

$$(25) \qquad \gamma_3 = \mu - \beta_1 - \alpha_2$$

$$(26) \qquad \gamma_4 = \mu + \alpha_1 + \alpha_2$$

Notice that, in this quarterly case, α_1 and β_1 each give rise to a component of y_t that has half-cycle every two periods and a full cycle every four periods. That is, these components cycle each year. The coefficients, α_1 and β_1 are associated with the spectral frequency $\pi/2$, since these multiply $\cos(\frac{t\pi}{2})$ and $\sin(\frac{t\pi}{2})$ respectively, for t=1,2,… (note that it is a property of these trigonometric functions that they cycle every four periods through the values 1, 0, -1, 0). Further, it might be remarked that although α_1 is associated with the second and fourth quarters in (23, 24, 25 and 26), while β_1 is associated with the first and third quarters, these

[3] In equation (22), we decided to use 365.25 instead of 365 or 366 because it is more comfortable and the error it introduces is minimum

two cycles cannot be distinguished in the frequency domain because they both correspond to four-period cycles.

Continuing the logic, the coefficient, α_2 is associated with the spectral frequency π, since it is the coefficient of terms of the form $\cos(t\pi)$ for t=1,2,… in (22). As these terms alternate between -1 and 1, α_2 contributes a full cycle every two periods. With quarterly data, the spectral frequencies π/2 and π are referred to as the seasonal frequencies, because any deterministic seasonal pattern over the four quarters of the year can be expressed as a linear function of terms at these two frequencies, namely as $\alpha_1 \cos(\frac{t\pi}{2})$ + $\beta_1 \sin(\frac{t\pi}{2})$+$\alpha_2 \cos(t\pi)$. By construction (through these trigonometric functions), the seasonal pattern necessarily sums to zero over any four consecutive values of t.

Note that (23, 24, 25, 26) can be written more compactly as

$$(27) \qquad \text{T=RB}$$

Where T= $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$', B= $(\mu, \alpha_1, \beta_1, \alpha_2)$' and

$$(28) \qquad R = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

This 4x4 nonsingular matrix R captures the one-to-one relationship between the trigonometric representation (22) and the usual dummy variable representation in (5) for the quarterly case. The general matrix equation, (27), also, of course, applies for data sampled at other frequencies. In particular, the elements of R and B for monthly data can be defined from (22) with S=12. In that case the seasonal frequencies are$\frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}$ and π, with two cycles occurring at each of these frequencies (one through each of the cosine and sine terms) except for the frequency π, where only the term $\cos(t\pi)$ applies. Any deterministic seasonal pattern in monthly data can be written by using the trigonometric cosine and sine functions at these seasonal spectral frequencies. Note that, once again by construction, the representation separates the overall mean μ form the deterministic seasonal component, with the latter necessarily summing to zero over the twelve consecutive values of t.

Returning to the general case of S seasons, the matrix R in (27) has some useful properties. With the overall mean included in the vector B, this is a square matrix and, since there is a one-to-one relationship between the seasonal dummy variable representation considered earlier in this section and the trigonometric representation, this matrix must be nonsingular. Indeed, the columns of this matrix are orthogonal to each other. That is, when the i^{th} column is denoted as the vector R_i, so that R=(R_1, …, R_1), then R'_iR_j= 0, i≠j. This property ensures that R'R = D is a diagonal matrix. Thus, whether the i^{th} diagonal element of D is d_i, then

$$(29) \qquad R^{-1} = \begin{bmatrix} \frac{1}{d_1} R'_1 \\ \frac{1}{d_2} R'_2 \\ \cdot \\ \cdot \\ \cdot \\ \frac{1}{d_s} R'_s \end{bmatrix}$$

So that the inverse of R (after appropriate scaling row by row) is the transpose of itself. For example, in the quarterly case it can be verified that

$$(30) \qquad R^{-1} = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0 & -0.5 & 0 & 0.5 \\ 0.5 & 0 & -0.5 & 0 \\ -0.25 & 0.25 & -0.25 & 0.25 \end{bmatrix}$$

Note also that the first column of R is a vector of ones. In this case, $R_1'R_1 = d_1 = S$ so that each element of the first row of R^{-1} is $\frac{1}{S}$. Consequently, the inverse yields the definitional relationship $\mu = \frac{1}{S}\sum_{s=1}^{S} \gamma_s$.

It is sometimes useful to associate the overall mean with the zero spectral frequency. This association follows as μ can be written in terms of trigonometric function as $\alpha_0 \cos\left(\frac{2\pi kt}{S}\right)$ with k=0, and hence (22) is equivalent to:

$$(31) \qquad y_t = \sum_{k=0}^{S/2} [\alpha_k \cos\left(\frac{2\pi kt}{S}\right) + \beta_k \sin\left(\frac{2\pi kt}{S}\right)] + z_t$$

Since sin(0)=0. However, having recognized that the overall mean has a spectral interpretation, we will continue to write it as μ and hence typically use the representation (22) in preference to (31).

It should be clear that the coefficients of the trend in the seasonally varying trend model of (9) can also be written by using a trigonometric representation. With appropriate definition of elements of T, R and B in (27), this equation applies also to the relationships between the trend coefficients in the dummy variable and trigonometric representations.

Each term is considered as a disturbance of the previous one, and it has been computed this way by an iterative process. The minimization process that has been used is an annealing process. The optimization is done for computing the alphas and betas and this way minimizing the sum of the squares of the residues.

Minimize a function using simulated annealing. Uses simulated annealing, a random algorithm that uses no derivative information from the function being optimized. Other names for this family of approaches include: “Monte Carlo”, “Metropolis”, “Metropolis-Hastings”, *etc*. They all involve:

- Evaluating the objective function on a random set of points

- Keeping those that pass their randomized evaluation criteria,
- Cooling (*i.e.*, tightening) the evaluation criteria
- Repeating until their termination criteria are met. In practice they have been used mainly in discrete rather than in continuous optimization.

8.4.1 Maximum likelihood

The objective is to obtain the parameters that maximize the likelihood, in order to calibrate the model. Under the usual assumption of the residues following a normal distribution with zero mean and constant volatility, maximizing the likelihood is equivalent to minimizing the sum of the squares of the residues:

$$S = f(t) + \varepsilon \rightarrow$$

$$\rightarrow N(0,1) \rightarrow$$

$$\rightarrow MAX\ [P] = MAX[\prod \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-\varepsilon^2}{2\sigma^2}}] \rightarrow$$

$$\rightarrow MAX[\log(P)] = MAX[-n\sqrt{2\pi}\sigma - \sum_{i=1}^{n} \frac{\varepsilon_i^2}{2\sigma^2}] \rightarrow$$

$$\rightarrow MIN[-\log(P)] = MIN[n\sqrt{2\pi}\sigma + \sum_{i=1}^{n} \frac{\varepsilon_i^2}{2\sigma^2}] \rightarrow$$

$$\rightarrow MIN[\sum \varepsilon_i^2]$$

What has been done is compute this function for the geometric Brownian function. Applying this method to a time series with a known seasonality allow us to check the validity of the model.

In the case of the trigonometric model, we chose the next array because of its simplicity, but any other could have been chosen:

[1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

Once we have generated a time series with this known seasonality, we use our models to extract the seasonality and compare it with the real value.

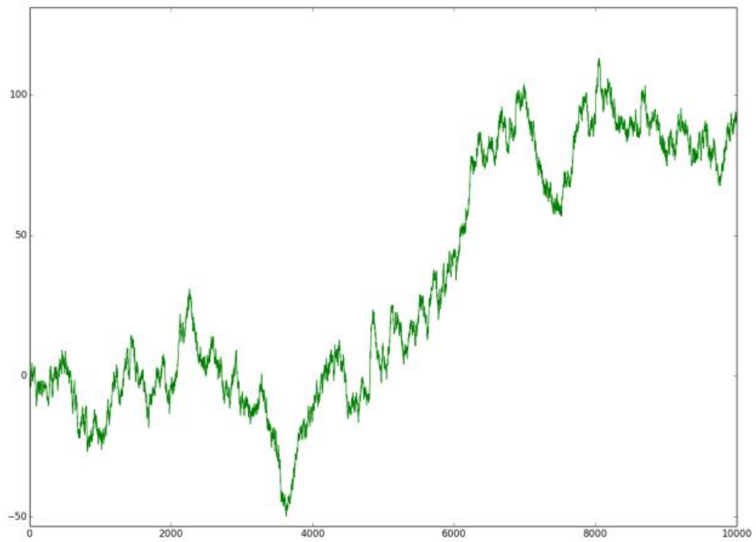
The last step is choosing the one with the lowest value and closer to the value computed with the right result.

8.4.2 Geometric Brownian motion

The Geometric Brownian Motion is also known as exponential Brownian motion, and is a continuous-time stochastic-process in which the logarithm of the process follows a Brownian motion. A Brownian motion is the random motion of particles that are suspended in a fluid that results from the collision with the quick molecules in the fluid.

The phenomenon receives this name from the botanist Robert Brown. In 1827, while Brown was observing through a microscope pollen grains floating on the water, he realized that they were moving but he was not able to determine what was the mechanism that caused that motion. The molecules and atoms have been theorized, but Albert Einstein was in 1905 the one that explained in detail that the motion Brown observed was the result of the water molecules moving the pollen. The explanation that Einstein gave was the definite confirmation that molecules and atoms exist. This was later confirmed by Jean Perrin in 1908, and he was awarded the “Nobel Prize in Physics in 1926 for his research on the discontinuous structure of matter.

An example of GBM is shown in the next graph:



Graph 6 GBM graphic

8.4.2.1 Trigonometric model

What we do in this section is this:

- Generate a time series with a known seasonality and a white noise component like the Geometric Brownian motion
- Use the model in order to compute the seasonality of the generated time series
- Make sure that the obtained result is similar to the real one, which was already known

We use a set of known parameters in order to generate a time series with a known seasonality and a given noise generated by a Brownian motion

Afterwards the model was fed with the values that the model computes and the maximum likelihood is again computed.

Once the two values have been obtained, the perfect situation is in which the two of them are equal or very similar.

The computation was done for the geometric Brownian motion with data 27 years, and the obtained values are these:

- Using the seasonality known in advance, the value is: 102.575551749
- With the solution seasonally adjusted by the model: 101.236106701

8.4.2.2 Multiplicative moving averages

Computing the value for the likelihood with this method, the result is: 100.697005184

8.4.2.3 Additive moving averages

Computing the value for the likelihood with this method, the result is: 100.120033804

9 HYPOTHESIS TESTING OF THE SIGNIFICANCE OF THE MOVING AVERAGES

This test has been carried out in order to test the accuracy of the moving averages model. What we are expecting to obtain through this test is whether we can say that the seasonality is significant or not. So whether the T-Test provides us with a result lower than 0.05 we can discard $\mu = 1$, that is to say, we can discard that there is no seasonality.

What it is done in this process, is to create a statistical as follows.

$$t = \frac{x - \mu_0}{\frac{s}{\sqrt{n}}}$$

Where x is the average of the time series, μ_0 is the null hypothesis (in this case equals 1), s is the variance and n the number of elements of the time series.

9.1.1 Geometric Brownian motion (multiplicative)

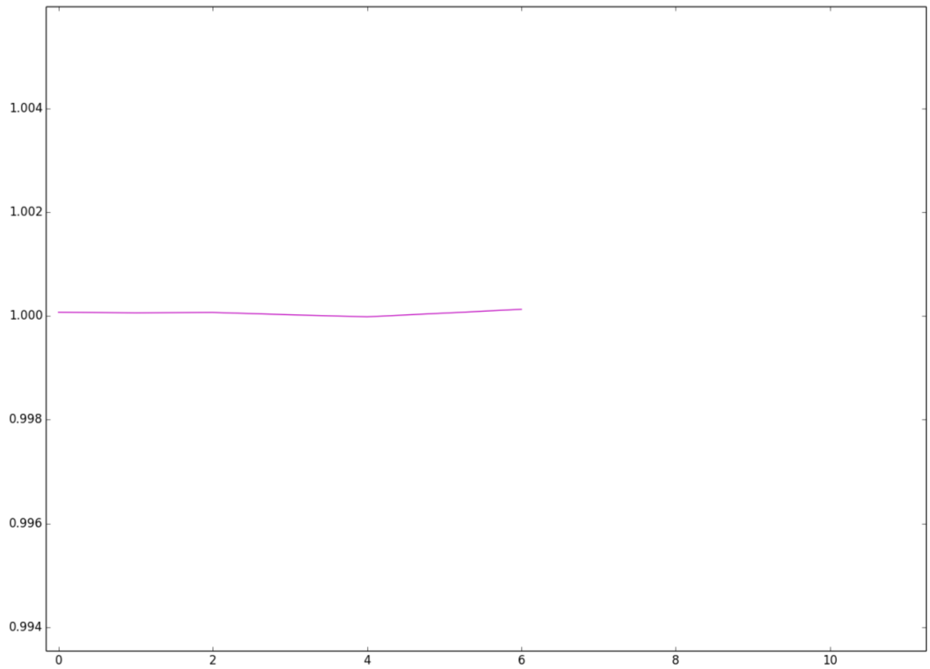
So now a T-Test is going to be applied to the GBM. The aim of applying this test is checking whether is possible to determine with a certain confidence interval that the average of a series is going to be different from 1.

In this test, the confidence interval that has been established is the 95%. The result we are expecting to obtain for the GBM is a high probability of being one the seasonality (just the results that are shown in Table 13, with higher probability than 0.05), because it is a time series without seasonality.

The results say that in no case we can discard that $\mu = 1$ (not having seasonality)

Day of the week	Coefficient	Probability T Test
Monday	1,00007475	0,86422659
Tuesday	1,00006378	0,88437109
Wednesday	1,00007092	0,87163633
Thursday	1,00002568	0,9532879
Friday	0,99998671	0,97569114
Saturday	1,00005773	0,89446788
Sunday	1,00013128	0,76355052

Table 13 T-Test for the days of the week of a GBM

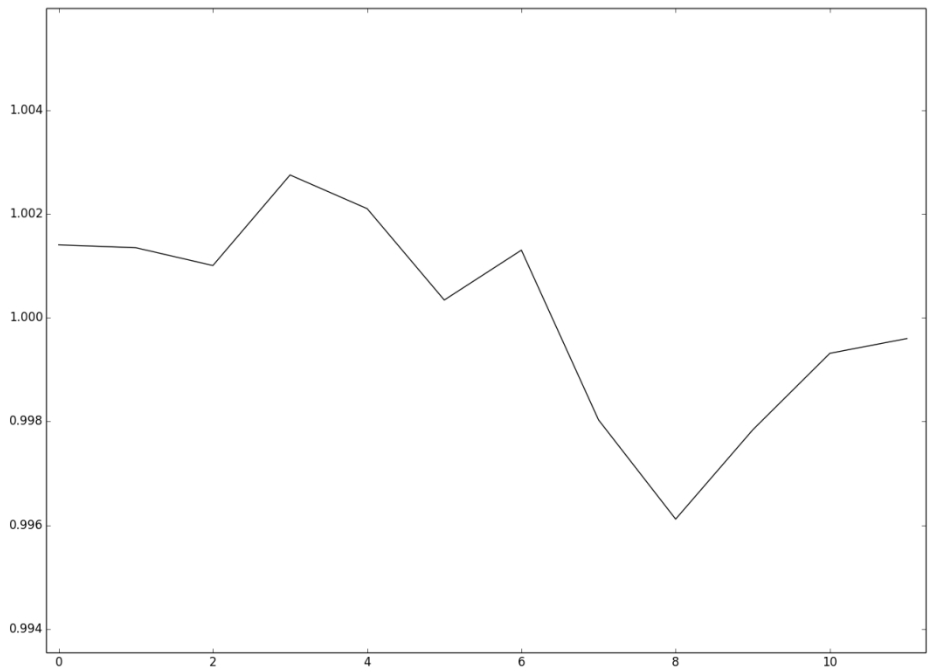


Graph 7 Seasonality of the days of the week for the GBM

In the Table are shown the coefficients and results of the T-Test for the monthly seasonality:

Month	Coefficient	Probability T Test
January	1,00140841	0,01169787
February	1,00135346	0,03169124
March	1,00100959	0,09423151
April	1,00275662	5,6866E-08
May	1,00210715	3,18E-07
June	1,00034401	0,42174541
July	1,00130638	0,00375929
Agoust	0,99803163	0,00024387
September	0,99611712	1,6593E-09
October	0,99783958	0,00062901
November	0,99931659	0,27707219
December	0,99960067	0,52542145

Table 14 T-Test for the months of a GBM



Graph 8 Seasonality of the months for GBM

The conclusions that are extracted from the analysis of these two tables are that:

- For the seasonal coefficients of the days of the week is not possible to discard $\mu = 1$. This means that the seasonality they present is very low.
- For march, june, November and December we cannot discard $\mu = 1$, that is to say, could happen not having significant seasonality.

In this graph the results are shown, and is possible to see the seasonality for the days of the week and for the months, feeding the moving averages model with a geometric Brownian motion. It was expected not to have seasonality, and for the days of the week is possible to see that is almost one (as it is shown in the Table 14).

The explanation having so much variation in the monthly seasonality is due to pure chance. However, whether the average of the monthly seasonality is computed, is shown that it is still higher than 0.05, so we can say that we can ignore its seasonality because it is not significant.

Here we are computing the average for the probability, that are what they tell us whether we can discard or not the hypothesis. This is a rough way for deciding that, in general, there is no seasonality.

9.1.2 EURUSD

The EURUSD is part of our testing procedures, because although we do not know its seasonality, we believe it must be very low.

In the following graph is shown the result for the seasonality of the EURUSD index depending on the days of the week. Watching the probability of the T Test, is possible to say that the seasonality it presents is very low and can be discarded with a confidence interval of at least 95%.

In this case the coefficients are computed only from Monday to Friday because is an index that only is traded these days.

Day of the week	Coefficient	Probability T Test
Monday	0,99999579	0,94571403
Tuesday	1,00000974	0,87306133
Wednesday	0,99999173	0,89168434
Thursday	0,99998859	0,85186276
Friday	0,99998008	0,74828807

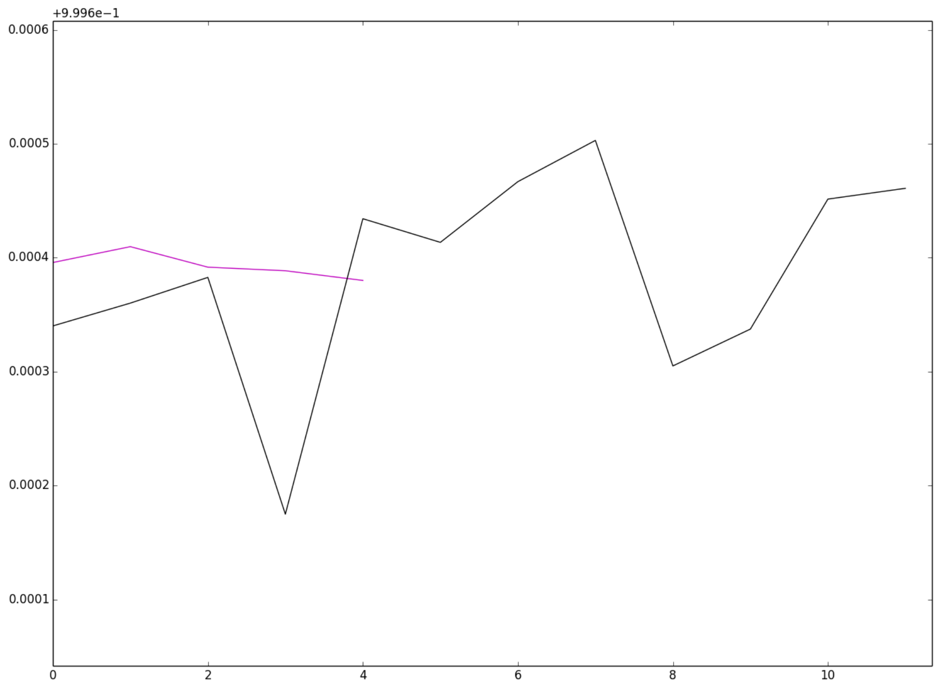
Table 15 T-Test for the days of the week of the EURUSD

In the Table 16 the high values simply tell us that there are not arguments opposed to $\mu = 1$ hypothesis (that implies no seasonality).

Date	Coefficient	Probability T Test
January	0,99994014	0,53638983
February	0,99996017	0,69326164
March	0,99998276	0,85088567
April	0,9997749	0,01341821
May	1,00003426	0,64460491
June	1,0000135	0,88537419
July	1,00006679	0,46182333
Agoust	1,00010301	0,17248448
September	0,99990502	0,1925926
October	0,99993738	0,4748821
November	1,00005148	0,58030438
December	1,00006095	0,50110747

Table 16 T-Test for the months of the EURUSD

Here we see the see the seasonality (in pink for the days of the week and in black for the months). For the days of the week you must analyze from the fifth decimal on and you will be able to see some seasonal pattern.



Graph 9 Seasonality for the days of the week and months

For the EURUSD we were expecting to have very low seasonality because it is almost a pure financial index, and the result we obtained confirmed our hypothesis.

Our test finishes with a positive result, because we have seen that our models work quite good and they have fulfilled our expectations.

10 ANALYSIS OF THE OBTAINED RESULTS

In this section, we are going to analyze the maximum likelihood of the different models for each one of the underlying assets, as well as some assets we have chosen in order to study their seasonality.

10.1 Maximum likelihood

In this section are going to be analyzed the maximum likelihood for each one of the underlying assets and each method.

The lower the value of the residues, the better the method for that underlying asset (as it was explained in section 9.1).

10.1.1 EURUSD

Multiplicative moving averages: 0.199269933047

Additive moving averages: 0.201042087566

Trigonometric: 0.200226783345

When the different values of the maximum likelihood are analyzed for the EURUSD index, the main conclusion extracted is that best method for analyzing this index is the multiplicative moving averages, because is the one with the lowest value, although the three of them are very similar and close to zero. The fact of being so similar to the obtained values is because this index has almost no seasonality or volatility, so it is quite stable.

10.1.2 Spanish electricity spot prices

Multiplicative moving averages: 228991.281376

Additive moving averages: 231840.967527

Trigonometric: 201580.822015

Regarding the Spanish daily spot prices, the method that best fits with this particular underlying asset is the trigonometric model, while both multiplicative moving averages and additive moving averages are very similar between them.

For the Spanish electricity, as it has very strong seasonality and volatility, the difference among models becomes much bigger than for the EURUSD asset.

10.1.3 Spanish base load

Multiplicative moving averages: 106285.924891

Additive moving averages: 108542.677719

Trigonometric: 87285.4818715

Talking about the Spanish electricity base load, the best model for analyzing this asset is the trigonometric one. The difference among models for the Spanish electricity spot prices and the base load is slightly smaller in this case, and this could be due to the lower seasonality the base load presents.

10.1.4 French spot electricity

Multiplicative moving averages: 2203822.2832

Additive moving averages: 1690562.17438

Trigonometric: 500664.087767

In the case of the French electricity prices, there is a very big difference among models, being the best one by far the trigonometric, the second one the additive moving averages and finally the multiplicative moving averages.

This situation could be due to the bigger difference in the seasonality during the weekdays in the French spot than in the Spanish spot. These seasonal coefficients of the Spanish and French spot are going to be shown forward in the document, in the sections 10.2 and 10.5 of the document.

10.1.5 French Gas prices

Multiplicative moving averages: 6128.58111496

Additive moving averages: 6263.56514695

Trigonometric: 6155.41503121

For the analysis of the French gas prices the tree methods are very similar, although is shown that the best one (not for a big difference) is the multiplicative moving averages, the second one the trigonometric and in the last position the additive moving averages.

In order to have an outlook of what model fits each underlying asset, here are presented for each model which of them adjusts better. The numbers in the table represent the level of adjustment, using the number 1 for the one with the lowest likelihood and 3 for the model with highest likelihood for that underlying asset.

	Spanish electricity spot (OMELRTRB)	Spanish base load (OMELRTRPTB)	French electricity spot (PNXBASE)	French gas spot (ARG_PEG NORD_D1)
Trigonometric	1	1	1	2
Multiplicative Moving Averages	2	2	3	1
Additive Moving Averages	3	2	2	3

Table 17 Comparison of the three models depending on the underlying asset

The best model is the trigonometric one, because is the best in three of the underlying assets, and is the second best for the French gas spot (very close to the multiplicative moving averages, that is the first one; but the three of them provide with very similar results).

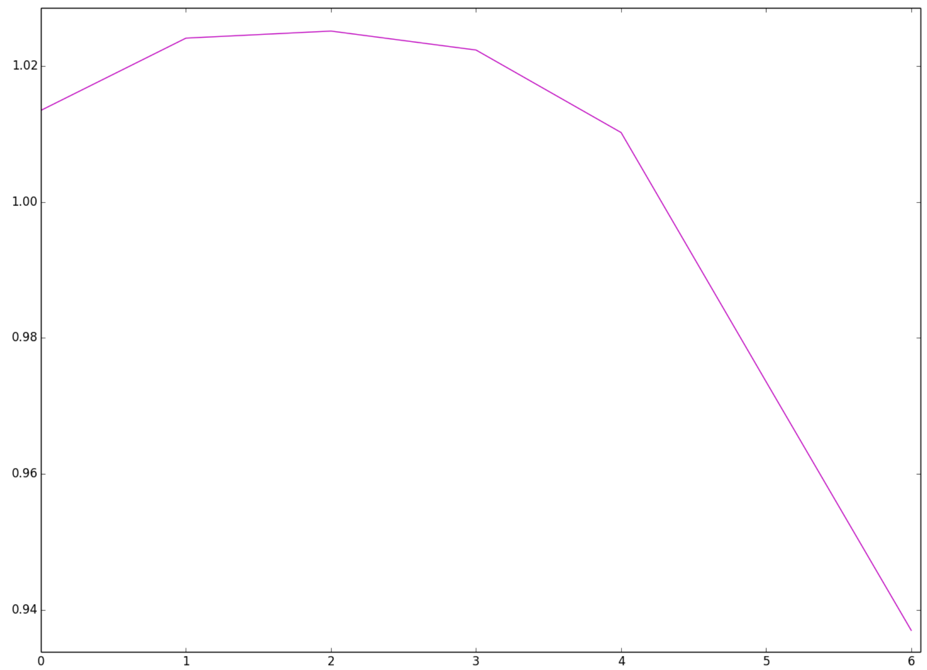
10.2 Spanish daily electricity spot prices

10.2.1 Moving averages multiplicative

In the next table are shown the seasonal coefficients for the days of the week.

Day of the week	Coefficient
Monday	1,01425285
Tuesday	1,02431159
Wednesday	1,02559701
Thursday	1,02127763
Friday	1,01003123
Saturday	0,97358621
Sunday	0,93568041

Table 18 Coefficients of the Spanish electricity spot price for the days of the week (multiplicative)



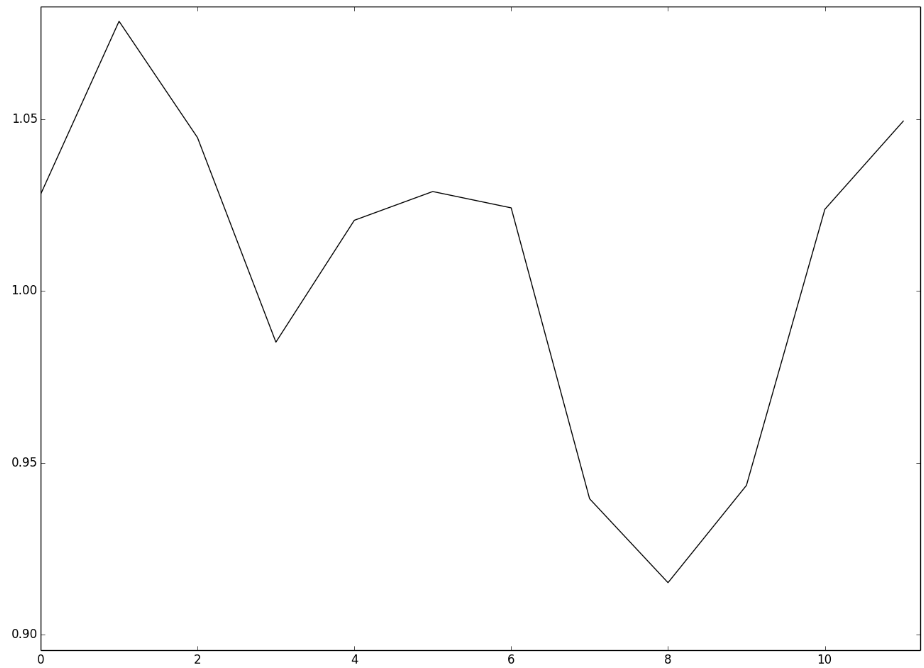
Graph 10 Seasonality of the days of the week of the Spanish electricity spot prices (multiplicative moving averages)

Analyzing the results, the Spanish electricity spot prices show a clear seasonality for the days of the week. As it is normal, the Tuesdays, Wednesdays and Thursdays present the highest seasonality, being the Mondays and Fridays a bit lower.

As the energy intensive industries are stopped for the weekend, the consumption goes down as the seasonal coefficients reflect.

Month	Coefficient
January	1,02941355
February	1,07758805
March	1,04408898
April	0,98267288
May	1,01961933
June	1,02856915
July	1,02366346
Agoust	0,93867988
September	0,91539844
October	0,94293947
November	1,02335631
December	1,04845991

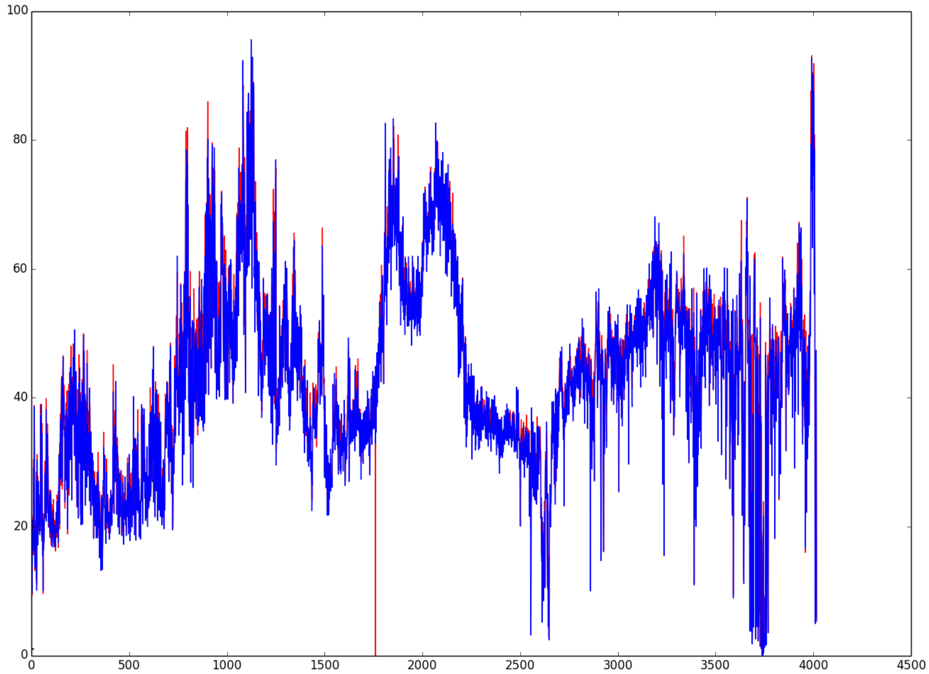
Table 19 Monthly coefficients of the Spanish electricity spot price (multiplicative)



Graph 11 Seasonality of the months of the Spanish electricity spot prices (multiplicative moving averages)

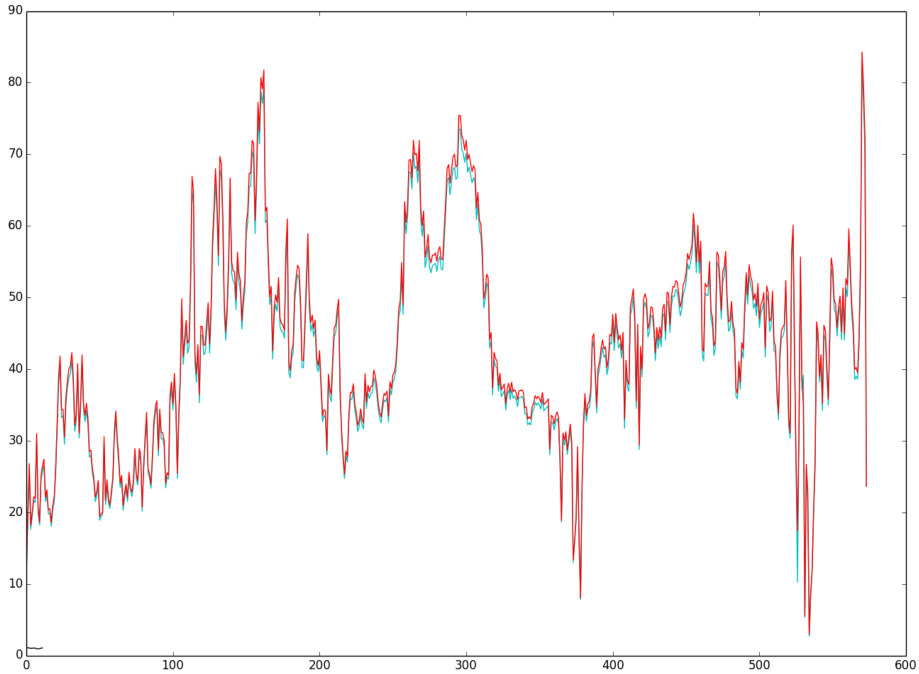
The surprising conclusion regarding the monthly seasonality is that the fact of being in summer time makes the electricity price be lower than it should be. In April (number 3 in the horizontal axes) is seen that the seasonality goes down, and this is due to the Easter holidays, that every year are around those dates. As many people go on holidays, the power consumption is lower than during the year. It could be similar to what happens the weekends and bank holidays.

In the following graph the prices without seasonality are shown. In red are plot the original prices, and in blue the seasonally adjusted prices:



Graph 12 Original and seasonally adjusted time-series of the Spanish spot prices (multiplicative moving averages)

For being able to analyze with less volatility the results, the two time series (the original prices and the seasonally adjusted one) have been plotted computing the weekly averages. In red is painted the weekly average of the original prices and in cyan the weekly averages of the seasonally adjusted time series by the method of the multiplicative moving averages.

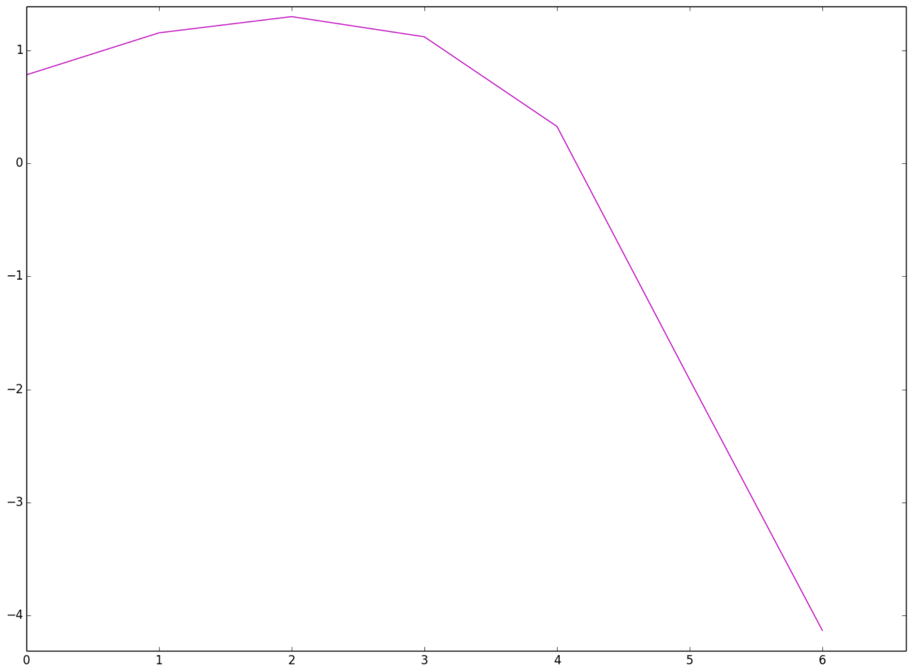


Graph 13 Flattened original and seasonally adjusted time-series of the Spanish spot prices (multiplicative moving averages)

10.2.2 Moving averages additive

Day of the week	Coefficient
Monday	0,78188813
Tuesday	1,15345709
Wednesday	1,29668115
Thursday	1,11830312
Friday	0,32520782
Saturday	-1,91381576
Sunday	-4,13049112

Table 20 Coefficients of the Spanish electricity spot price for the days of the week (additive)



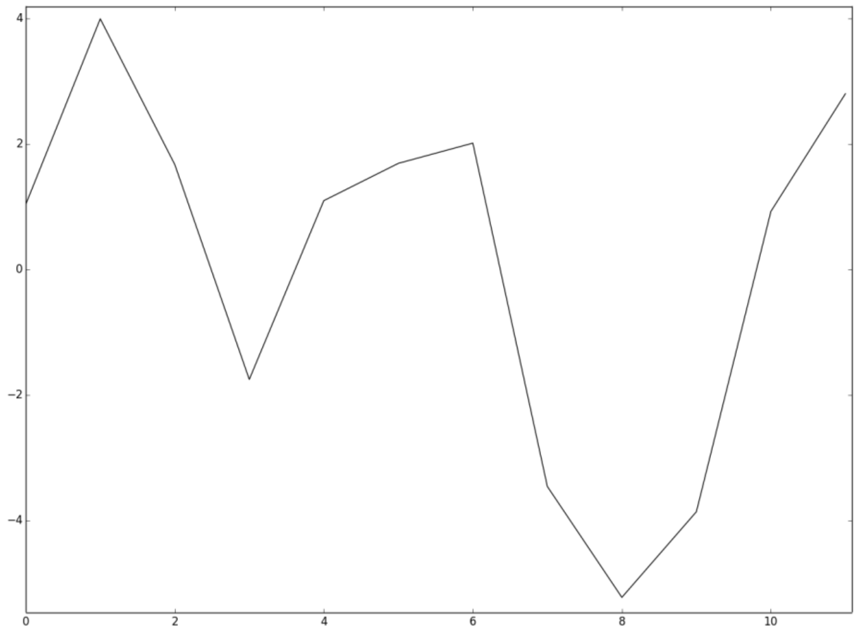
Graph 14 Seasonality of the days of the week of the Spanish electricity spot prices (additive moving averages)

The seasonality that is shown by the additive moving averages has the same shape as the multiplicative moving averages, but the difference is that in the multiplicative for representing no seasonality you use the coefficient one, and in the additive you use the coefficient zero. The same differences are presented here for the weekdays and for the weekends.

Following, the coefficients for the monthly seasonality are shown:

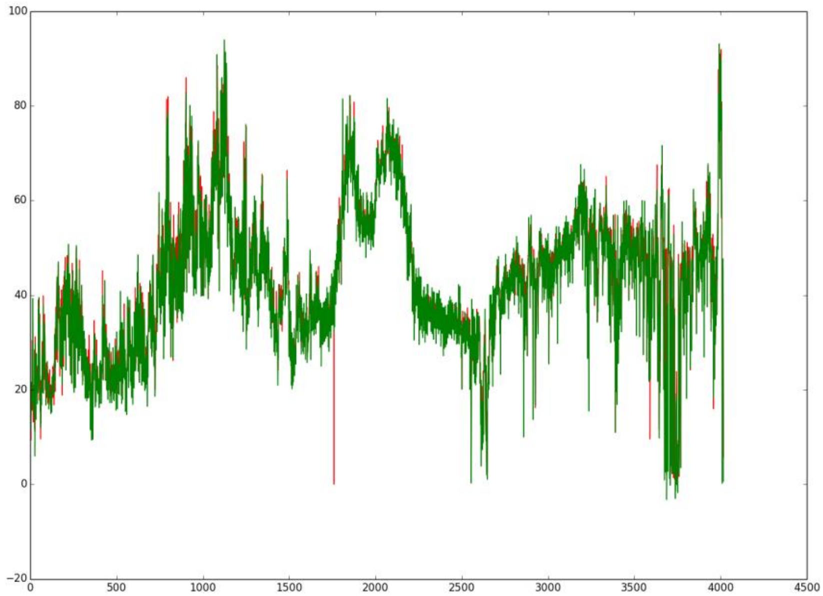
Month	Coefficient
January	1,04043788
February	3,99937301
March	1,6772709
April	-1,74774433
May	1,1014734
June	1,69584534
July	2,01829872
Agoust	-3,45442425
September	-5,22169942
October	-3,85668316
November	0,92979706
December	2,80567016

Table 21 Monthly coefficients of the Spanish electricity spot price (additive)



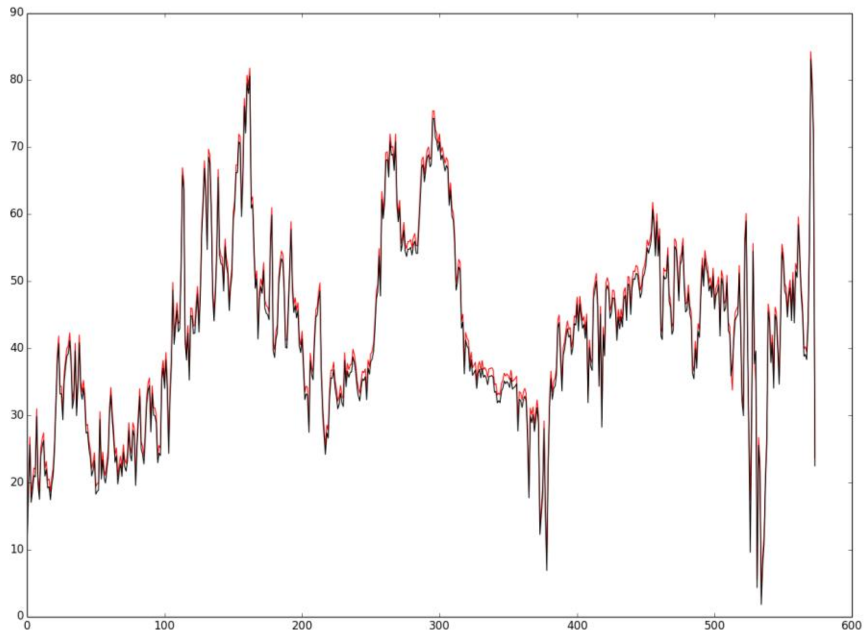
Graph 15 Seasonality of the months of the Spanish electricity spot prices (additive moving averages)

Once the seasonal coefficients for the days of the week and the month are computed, the next graph is painted with the original prices and the same prices seasonally adjusted (in green the seasonally adjusted time series of the Spanish spot prices and in red the original prices):



Graph 16 Original and seasonally adjusted time-series of the Spanish spot prices (additive moving averages)

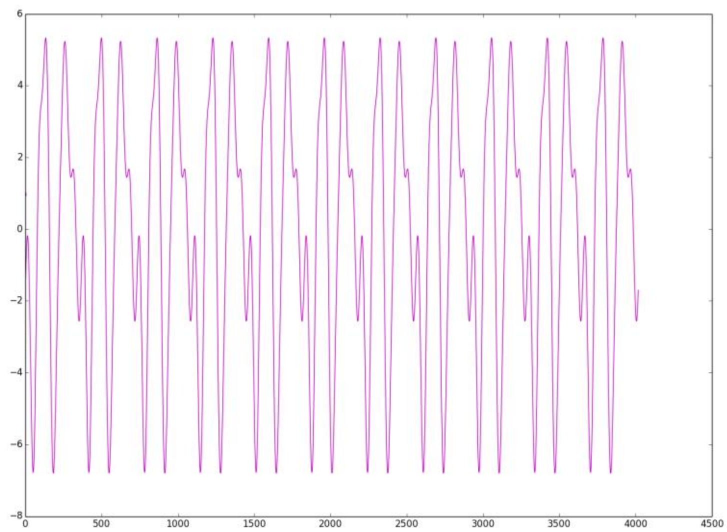
In the next graph, in red can be seen the weekly averages for the original prices and in black, the weekly averages for the seasonally adjusted prices by the additive moving averages method.



Graph 17 Flattened original and seasonally adjusted time-series of the Spanish spot prices (additive moving averages)

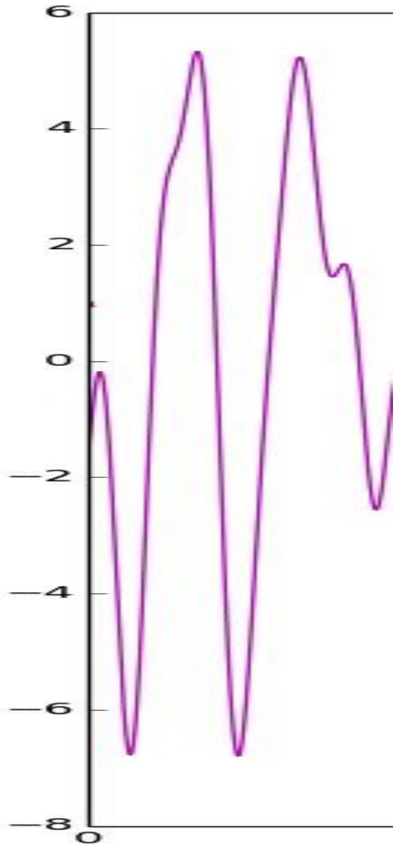
10.2.3 Trigonometric model

Now, the seasonality is going to be represented using the trigonometric model.

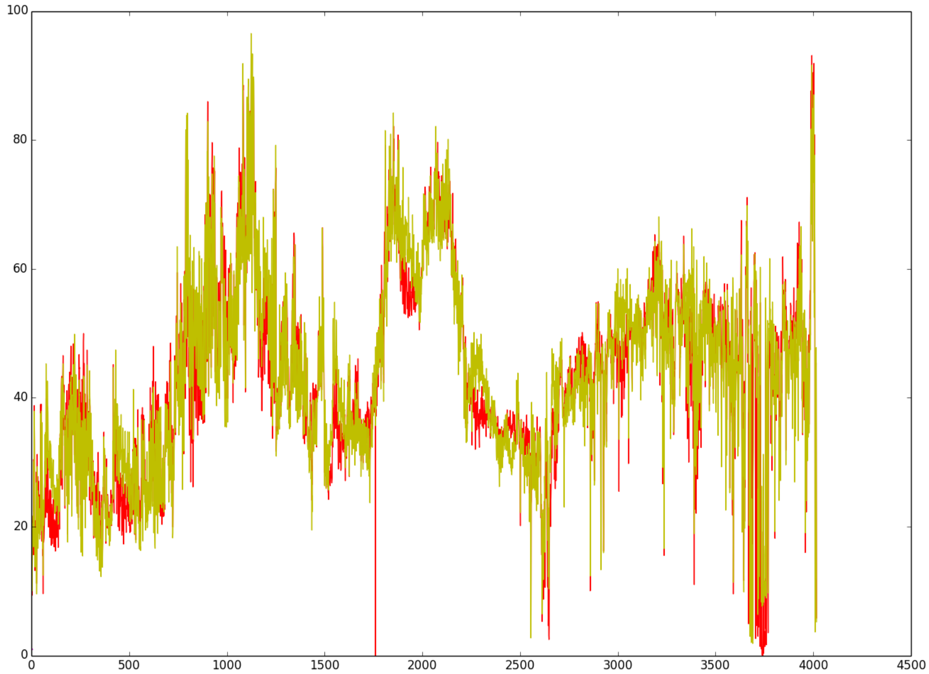


Graph 18 Seasonal coefficients for the Spanish electricity spot prices (trigonometric)

And in the next one is painted only a year, in order to be able to see better the differences among months:

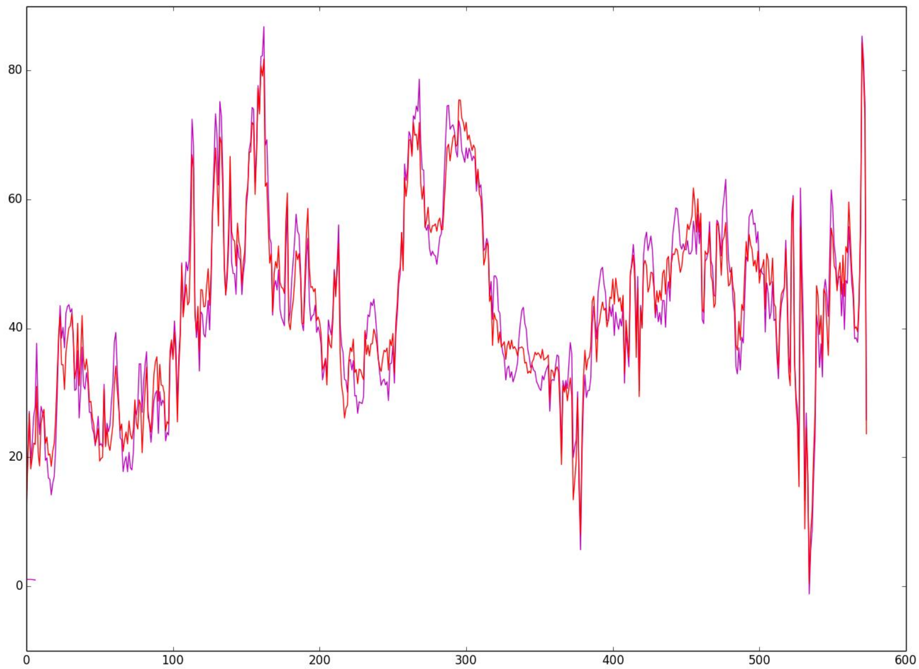


In the Graph 19, the two time-series are painted. In red the original one, and in yellow the seasonally adjusted time series.



Graph 19 Original and seasonally adjusted time-series of the Spanish spot prices (trigonometric)

Now we present the two time series, the original one in red and the seasonally adjusted in magenta, once the weekly averages have been computed for being able to remove the high volatility (the original time series is represented in red and the seasonally adjusted in magenta):



Graph 20 Flattened original and seasonally adjusted time-series of the Spanish spot prices (trigonometric)

The differences between the original price and the seasonally adjusted one are much bigger than in the case of the additive and multiplicative moving averages. This could happen because the trigonometric model adjusts better the seasonality and manages to go further than with the previous methods.

10.3 Spanish electricity base load

In this section is analyzed the seasonality of the Spanish base-load electricity

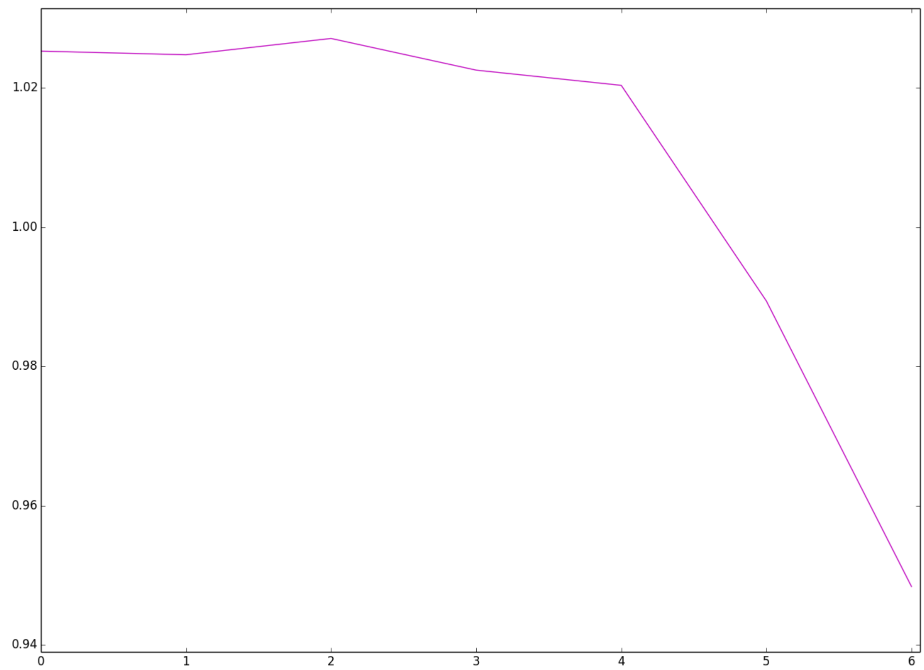
10.3.1 Multiplicative moving averages

The most remarkable characteristic of the base-load is that the seasonality of the weekdays is quite flat comparing to the daily spot prices. This could happen basically because the base-load is contracted in the long-run, only for the consumption that the industrial companies know the will have in the future. For the remaining part they bid in the spot market and this way adjust to their actual needs.

Obviously, they also know that during the weekends their consumption is lower than during the weekdays, so it is reflected in the coefficients for those days as it is shown here:

Date	Coefficient
Monday	1,0252446
Tuesday	1,02472447
Wednesday	1,02706412
Thursday	1,02251832
Friday	1,0203343
Saturday	0,98938241
Sunday	0,9484064

Table 22 Coefficients of the days of the week for the Spanish electricity base-load (multiplicative)

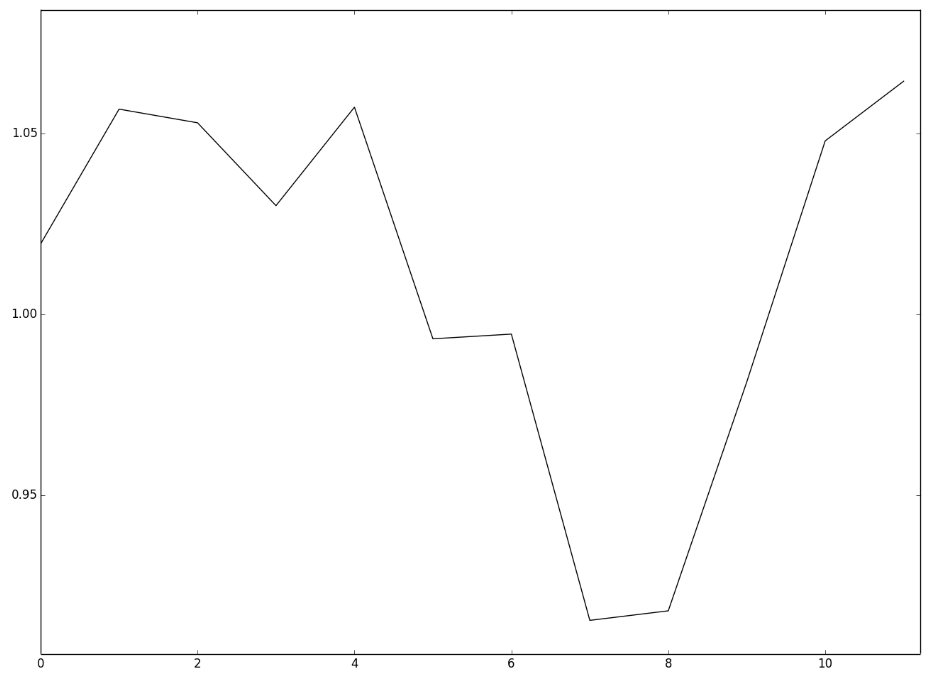


Graph 21 Seasonality of the days of the week for the Spanish base-load (multiplicative moving averages)

Regarding the monthly coefficients, the fact of being in hot season makes the prices be below one, so the prices are lower than what they should be.

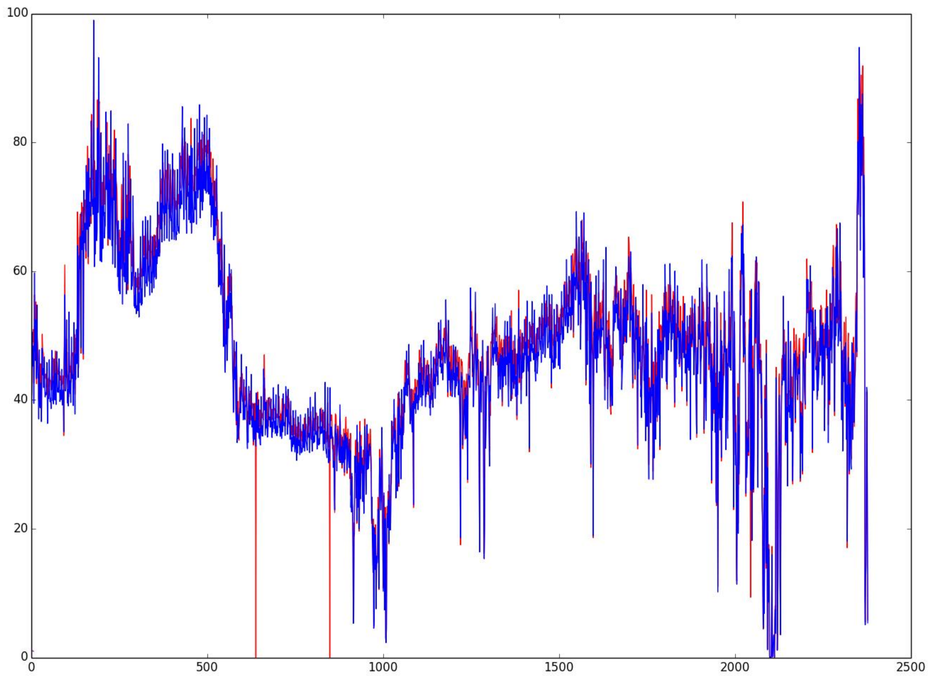
Date	Coefficient
January	1,01944235
February	1,05670134
March	1,05292499
April	1,03002288
May	1,05724132
June	0,99321907
July	0,99451795
Agoust	0,91539313
September	0,91802575
October	0,98127179
November	1,04792837
December	1,06442154

Table 23 Monthly coefficients for the Spanish base-load (multiplicative)



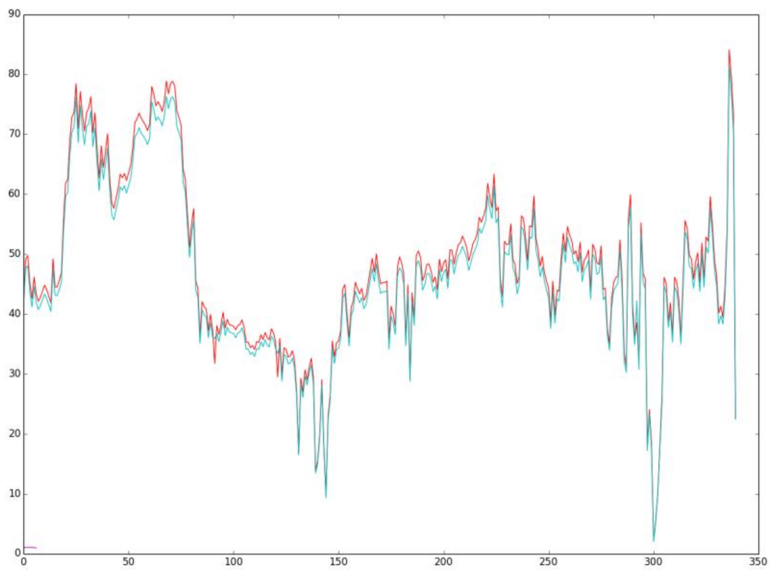
Graph 22 Seasonality of the months of the Spanish base-load (multiplicative moving averages)

Once the coefficients are computed, is possible to plot both the real (in red) and seasonally adjusted (in blue) time series.



Graph 23 Original and seasonally adjusted time series of the Spanish base-load (multiplicative moving averages)

In order to have a clearer view of the two series and be able to see how they are affected, the two of them were flattened computing weekly averages.



Graph 24 Flattened original and seasonally adjusted time series of the Spanish base-load (multiplicative moving averages)

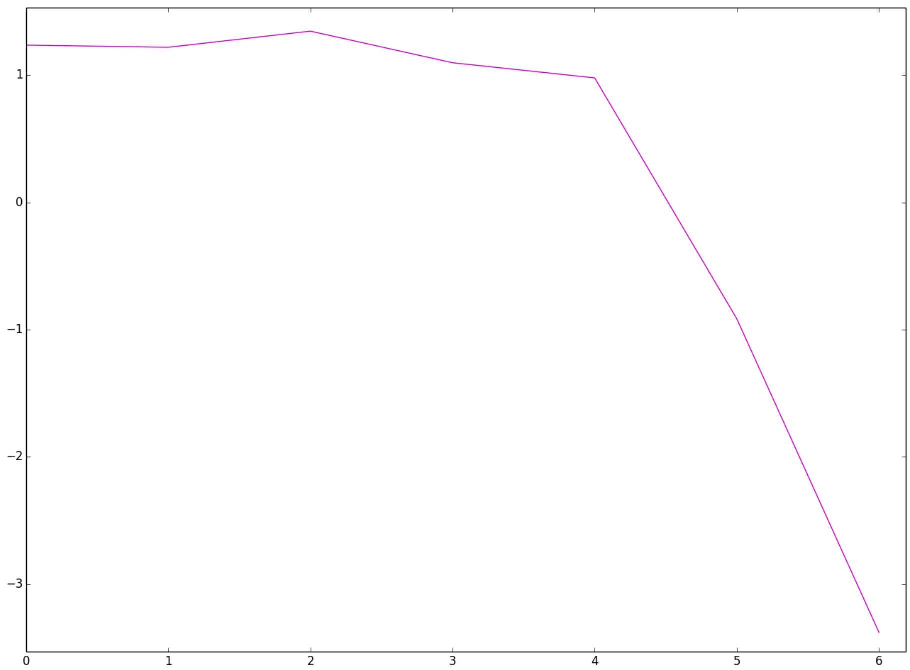
In the previous graphs for the Spanish electricity base-load, the volatility has much bigger effect than the seasonality; this is the main reason for not having clear differences between the original and the seasonally adjusted time series.

10.3.2 Additive moving averages

Like in the case of the Spanish spot prices, here the shape is the same both for the multiplicative and additive moving averages.

Date	Coefficient
1	1,23557924
2	1,21870092
3	1,34542136
4	1,09714593
5	0,97819165
6	-0,91462694
7	-3,37791568

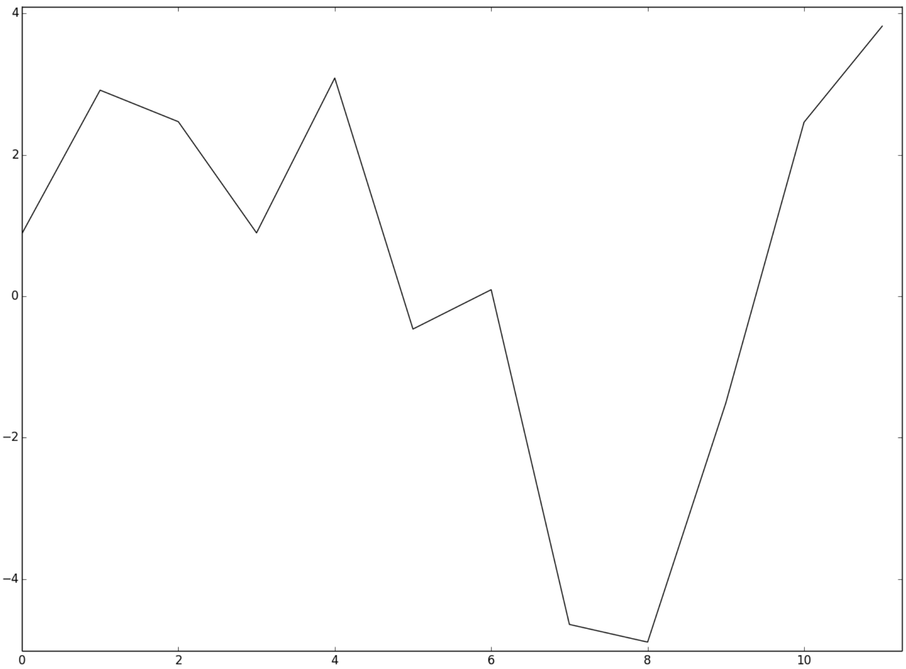
Table 24 Coefficients of the days of the week for the Spanish base-load (additive)



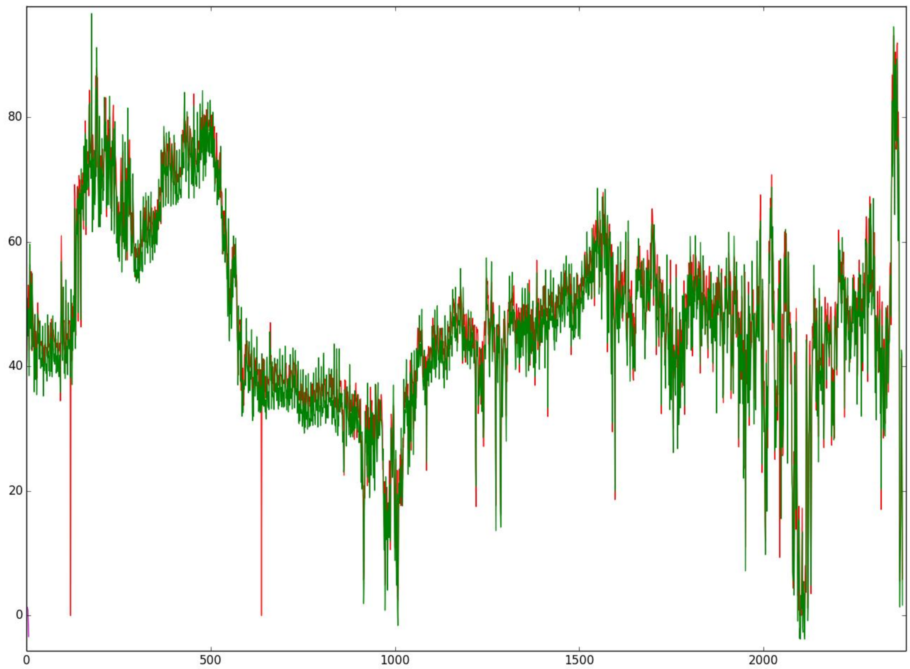
Graph 25 Seasonality of the days of the week for the Spanish base-load (additive moving averages)

Date	Coefficient
1	0,88276688
2	2,91361747
3	2,46788659
4	0,89505039
5	3,08535386
6	-0,46479358
7	0,09173203
8	-4,64191061
9	-4,89196068
10	-1,51039621
11	2,46022858
12	3,81966733

Table 25 Monthly coefficients for the Spanish base-load (additive)

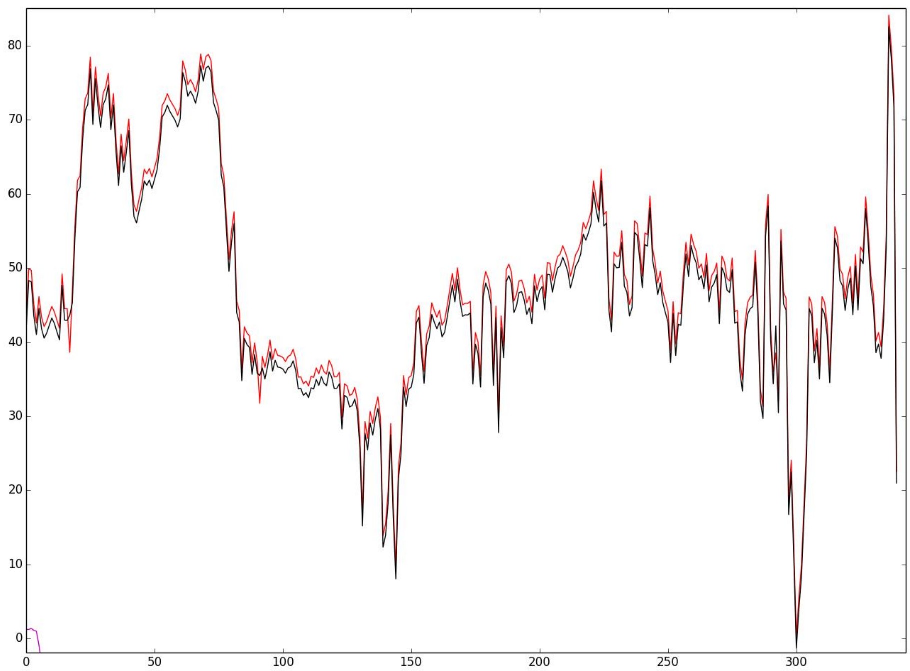


Graph 26 Seasonality of the months for the Spanish base-load (additive moving averages)



Graph 27 Original and seasonally adjusted time series of the Spanish base-load (additive moving averages)

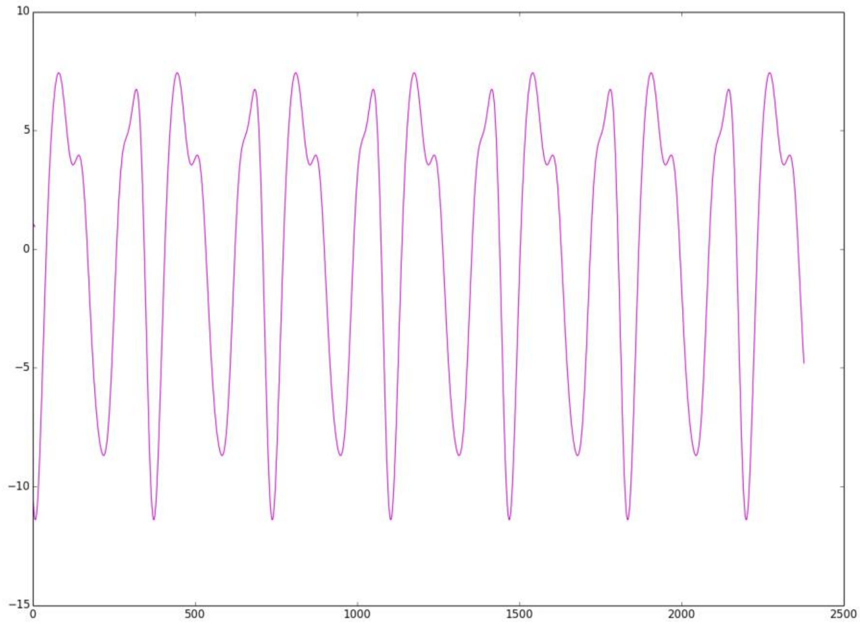
The main difference comes from the comparison of both the spot prices and base-load (Spanish). While in the case of the base-load the seasonality is quite remarkable, in the spot prices it is much lower. In the case of the base-load, almost all the time the seasonally adjusted prices (in black) are below the original ones (in red), as is shown in the following graph:



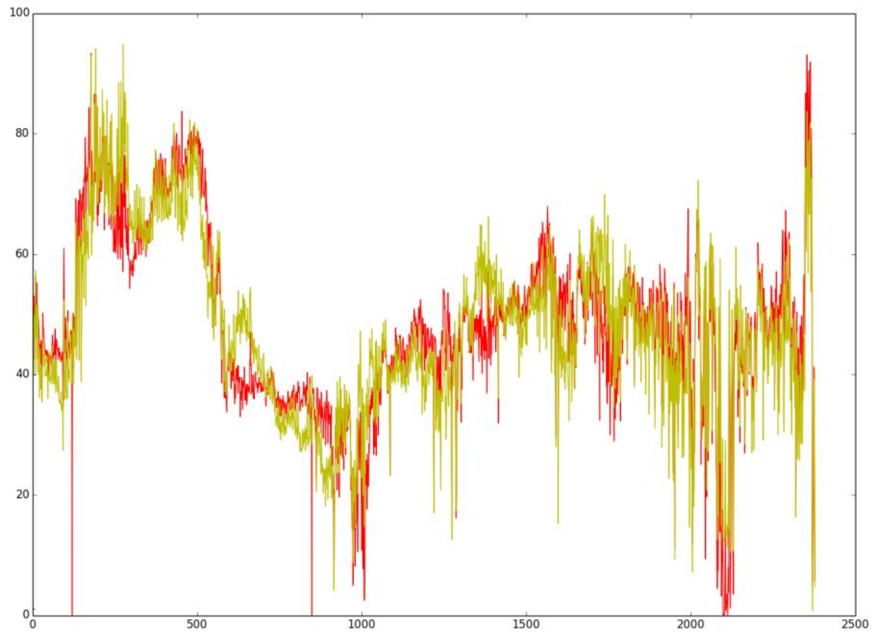
Graph 28 Flattened original and seasonally adjusted time series of the Spanish base-load (additive moving averages)

In this method, as for the multiplicative moving averages, the seasonally adjusted time series is most of the time below the real one, so this might mean that the fact of being base-load, could imply having prices higher than what really should be.

10.3.3 Trigonometric method

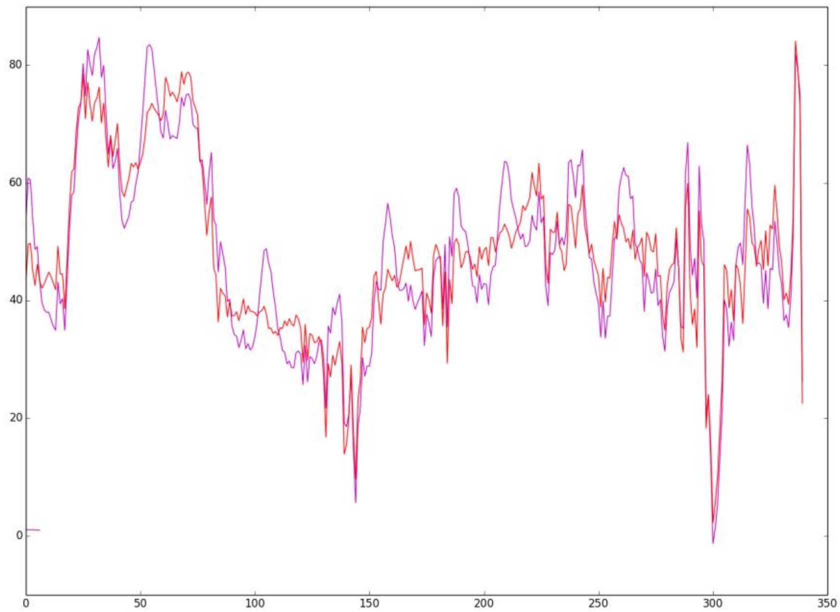


Graph 29 Seasonal coefficients for the Spanish base-load (trigonometric)



Graph 30 Original and seasonally adjusted time-series of the Spanish base-load (trigonometric)

When the seasonally adjusted (in magenta) and the original time series (in red) are plotted using the trigonometric model, much more spikes are plotted than for the two previous models (the multiplicative and additive moving averages).



Graph 31 Flattened original and seasonally adjusted time-series of the Spanish spot prices (trigonometric)

10.4 French daily electricity spot prices

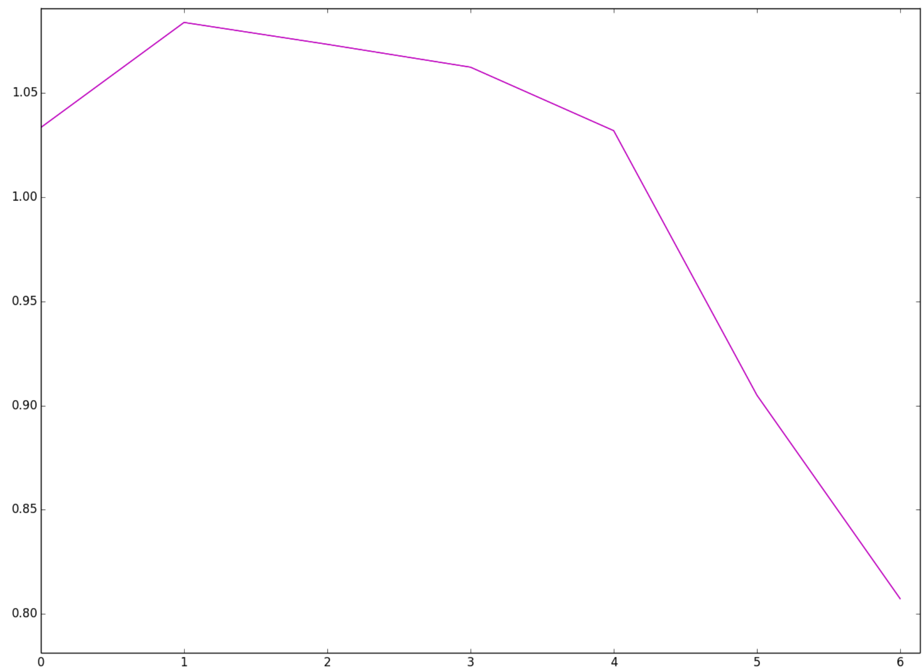
10.4.1 Multiplicative moving averages

When talking about the French electricity spot prices, is worthy to point out that also the seasonality of the days of the week is very strong, with remarkable differences among the weekdays and also between the weekdays and the weekends.

When the weekdays are analyzed, the seasonal coefficients for Mondays and Fridays are very similar, while for Tuesdays, Wednesdays and Thursdays are higher. Talking about the weekends, it makes the days have a lower price.

Date	Coefficient
Monday	1,03327185
Tuesday	1,08364609
Wednesday	1,07315476
Thursday	1,06219811
Friday	1,03175888
Saturday	0,90495308
Sunday	0,80737598

Table 26 Coefficients of the days of the week for the French electricity spot prices (multiplicative)

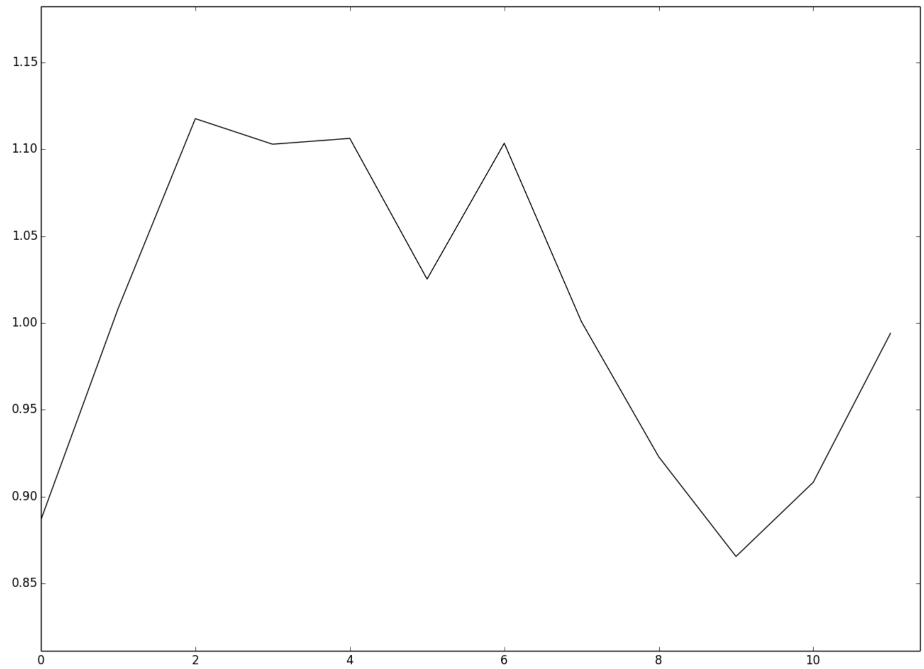


Graph 32 Seasonality of the days of the week of the French electricity spot prices (multiplicative moving averages)

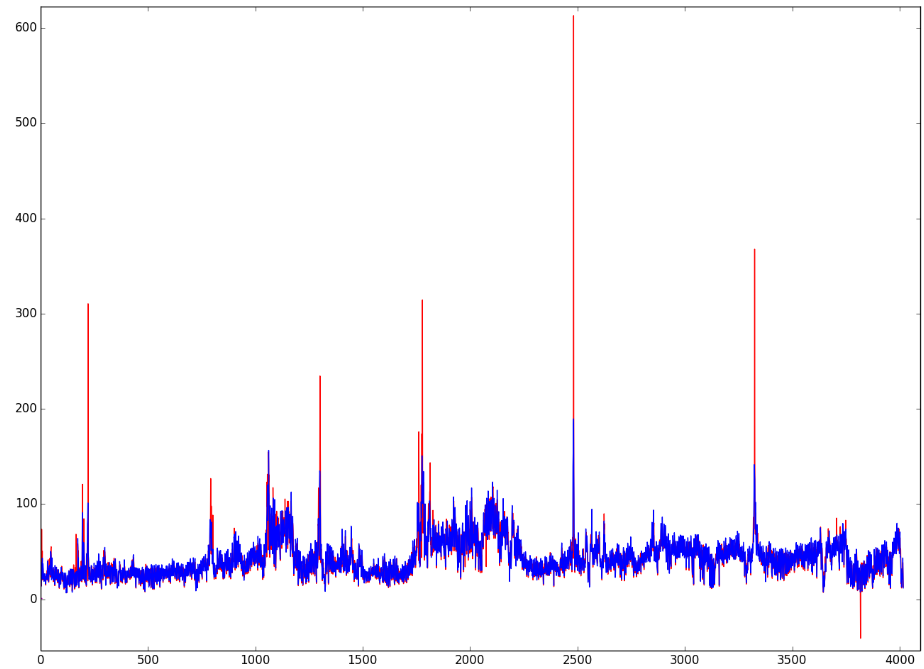
Studying the results of the monthly coefficients, the spring and summer make the prices be higher, and on the other hand, the autumn and winter make them be lower.

Date	Coefficient
January	0,88653992
February	1,00825302
March	1,11761827
April	1,10290809
May	1,10625068
June	1,02524064
July	1,10349652
Agoust	1,00064004
September	0,9229062
October	0,86553196
November	0,90824165
December	0,99403019

Table 27 Monthly coefficients for the French electricity spot prices (multiplicative)

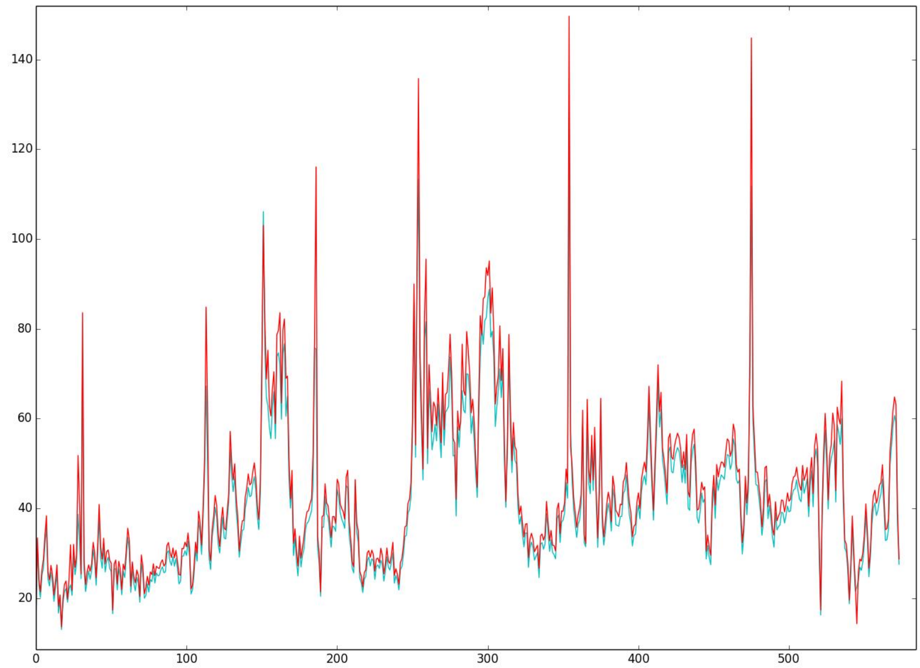


Graph 33 Seasonality of the months of the French electricity spot prices (multiplicative moving averages)



Graph 34 Original and seasonally adjusted time series of the French spot prices (multiplicative moving averages)

In red the original prices are shown and in cyan the seasonally adjusted ones.



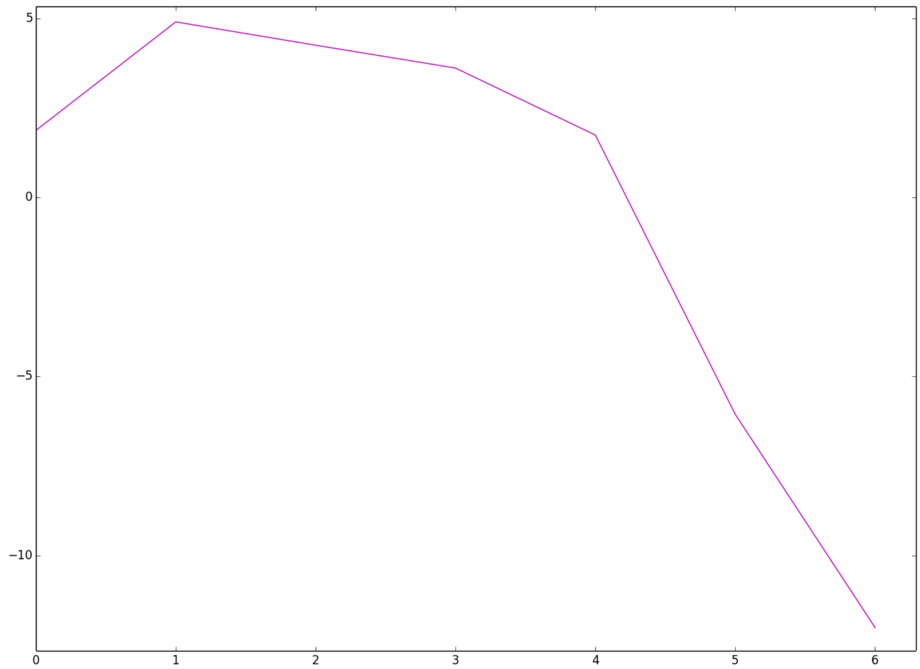
Graph 35 Flattened original and seasonally adjusted time series of the French spot prices (multiplicative moving averages)

10.4.2 Additive moving averages

In the case of the additive model for the French electricity prices, the shape also is the same as in the multiplicative model.

Date	Coefficient
1	1,86772119
2	4,89595033
3	4,24533203
4	3,60775913
5	1,73660124
6	-6,05733982
7	-12,0120319

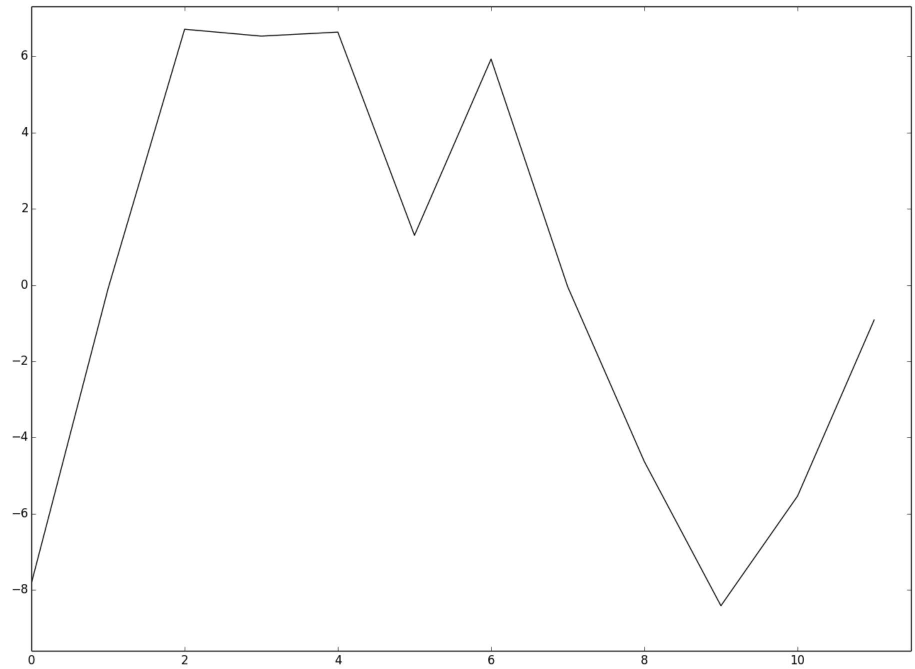
Table 28 Coefficients of the days of the week for the French electricity spot prices (additive)



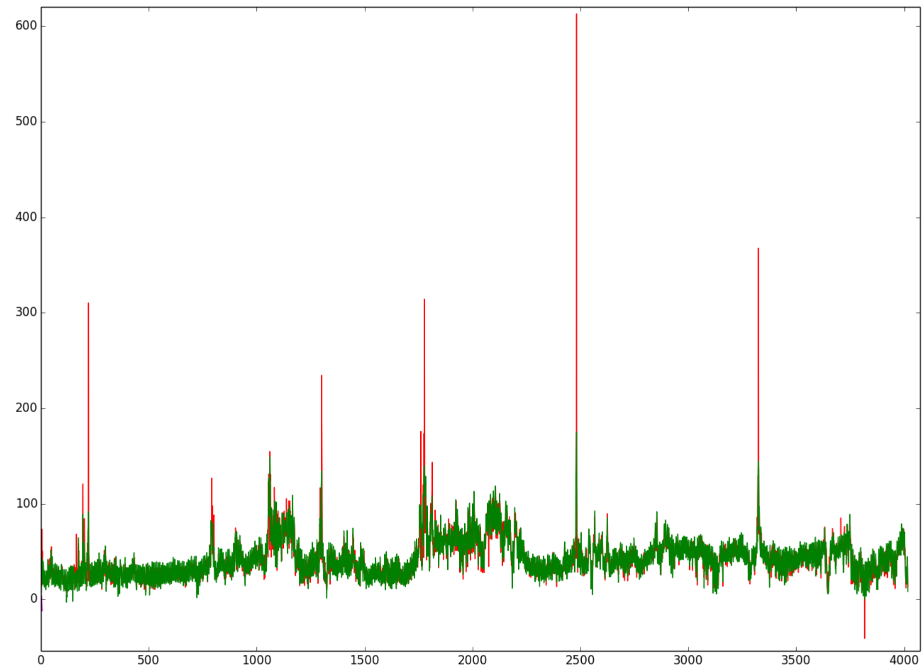
Graph 36 Seasonality of the days of the week of the French electricity spot prices (additive moving averages)

Date	Coefficient
1	-7,82662382
2	-0,10204956
3	6,71373192
4	6,53327275
5	6,63841844
6	1,30550533
7	5,93101178
8	-0,04323538
9	-4,63018949
10	-8,41560314
11	-5,53747282
12	-0,91618159

Table 29 Monthly coefficients for the French electricity spot prices (additive)

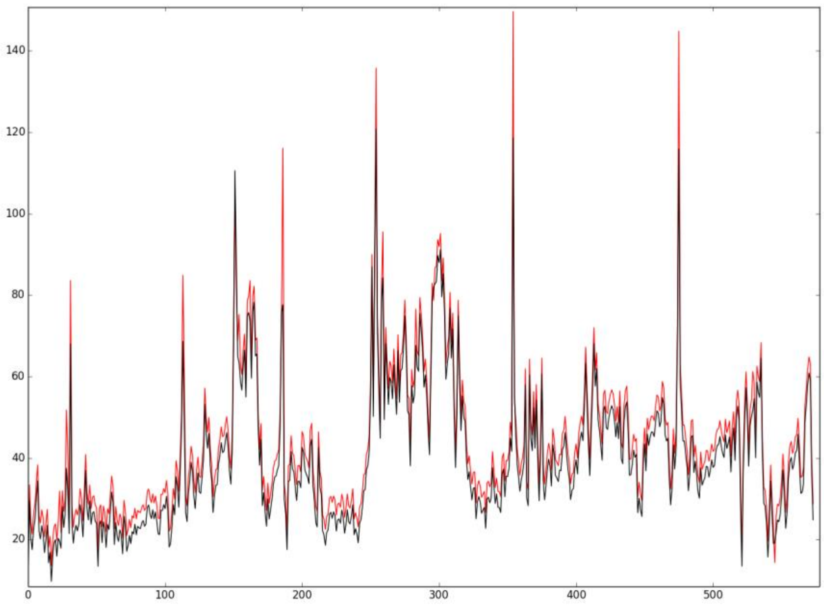


Graph 37 Seasonality of the months of the French electricity spot prices (additive moving averages)



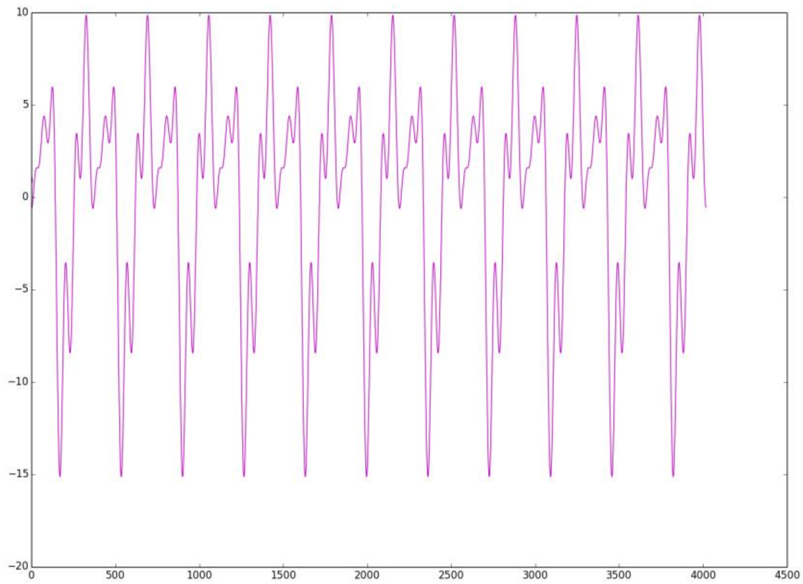
Graph 38 Original and seasonally adjusted time series of the French spot prices (additive moving averages)

When the two curves (the original prices in red and the seasonally adjusted ones in black) are studied, is possible to see the high similarities between the two model for the moving averages.

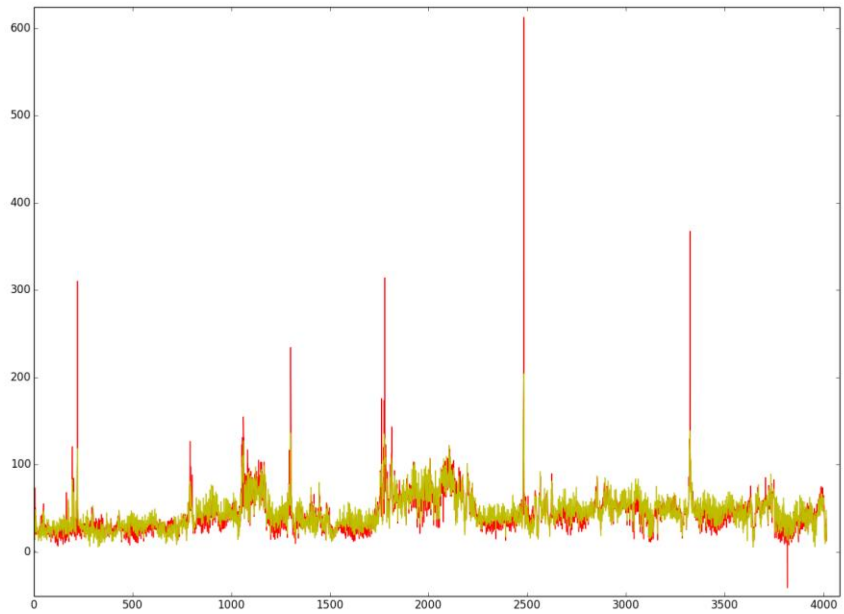


Graph 39 Flattened original and seasonally adjusted time series of the French spot prices (additive moving averages)

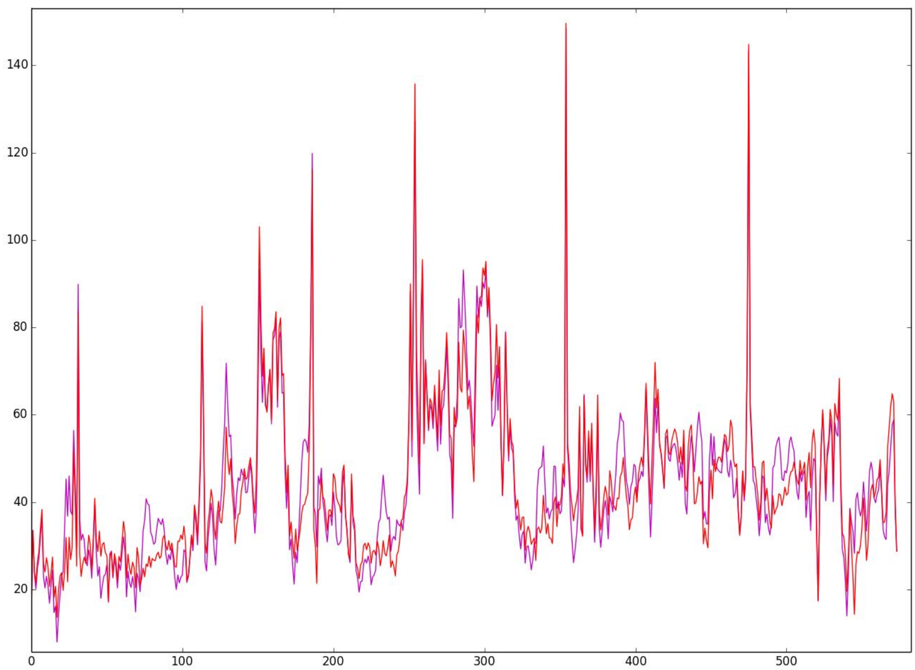
10.4.3 Trigonometric model



Graph 40 Seasonal coefficients for the French spot prices (trigonometric)



Graph 41 Original and seasonally adjusted time-series of the French spot prices (trigonometric)



Graph 42 Flattened original and seasonally adjusted time-series of the French spot prices (trigonometric)

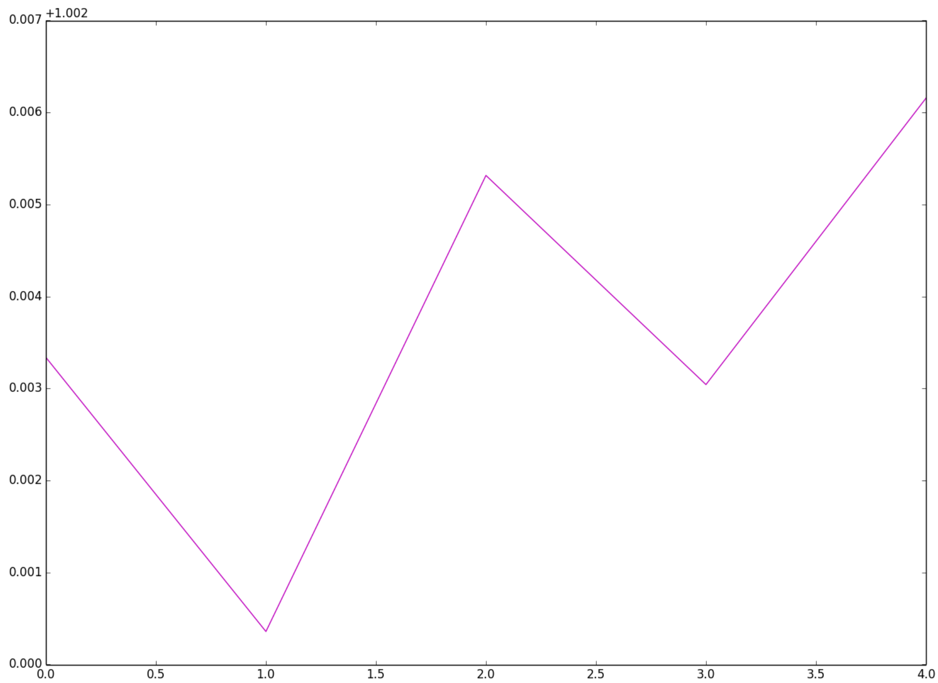
10.5 French gas PEG Nord

10.5.1 Multiplicative moving averages

Analyzing the PEG Nord, is possible to see that the Fridays have the strongest seasonality.

Date	Coefficient
Monday	1,00533803
Tuesday	1,00235834
Wednesday	1,00731596
Thursday	1,00504257
Friday	1,00815588

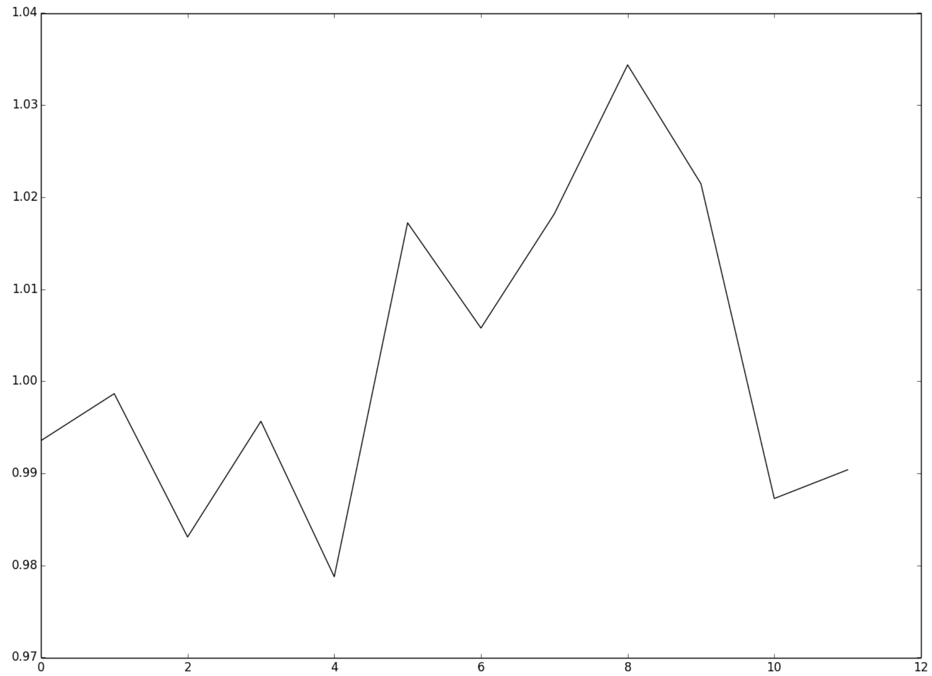
Table 30 Coefficients of the days of the week for the Peg Nord (multiplicative)



Graph 43 Seasonality of the days of the week of the Peg Nord (multiplicative moving averages)

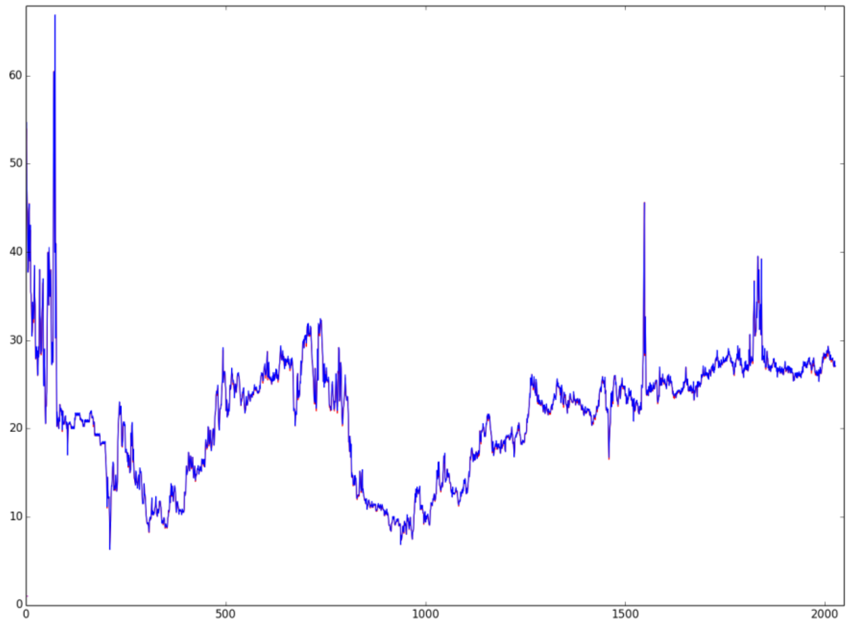
Date	Coefficient
January	0,9935501
February	0,99865415
March	0,98308439
April	0,99565993
May	0,9787722
June	1,01721231
July	1,00577838
Agoust	1,01817828
September	1,03436405
October	1,02144784
November	0,98726731
December	0,99038586

Table 31 Monthly coefficients for the Peg Nord (multiplicative)



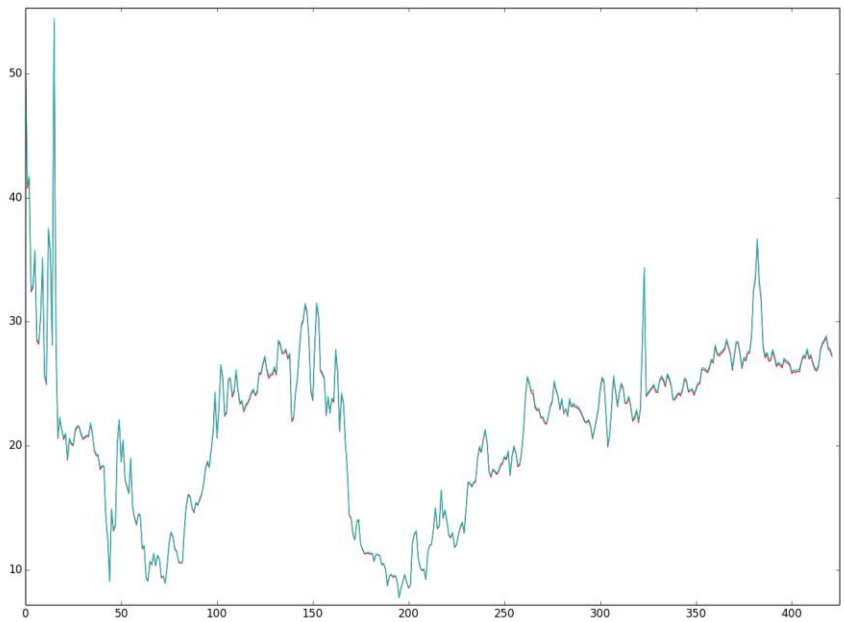
Graph 44 Seasonality of the months of the Peg Nord (multiplicative moving averages)

In the following graph we see the seasonally adjusted (in blue) and the original (in red) time series.



Graph 45 Original and seasonally adjusted time series of the Peg Nord (multiplicative moving averages)

In the next graph we have flattened both the original prices (in red) and the seasonally adjusted ones (in cyan).

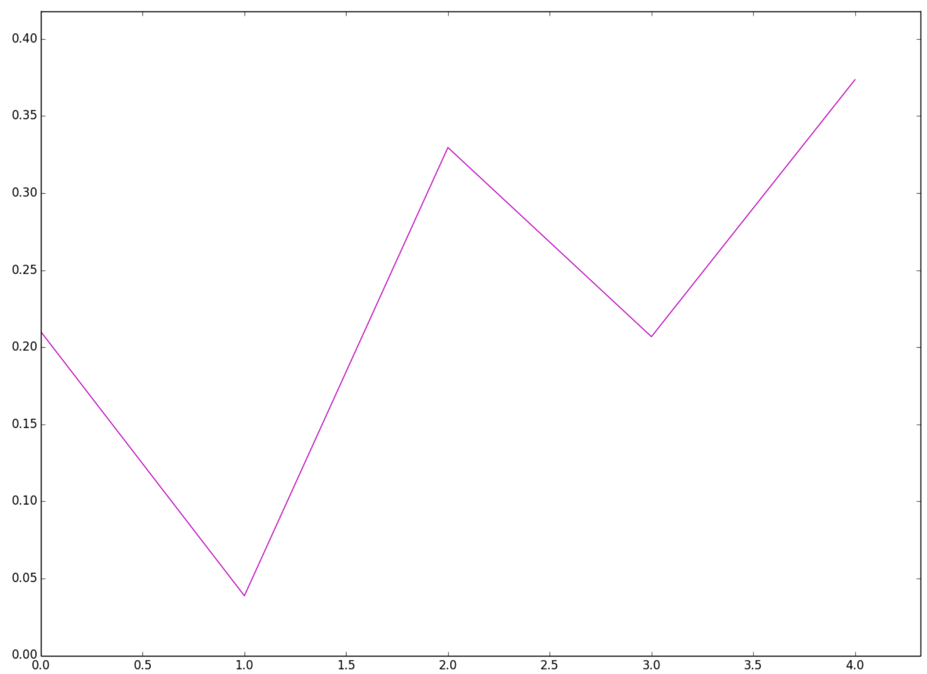


Graph 46 Flattened original and seasonally adjusted time series of the Peg Nord (multiplicative moving averages)

10.5.2 Additive moving averages

Date	Coefficient
1	0,21002299
2	0,03874783
3	0,32955779
4	0,20681084
5	0,3735466

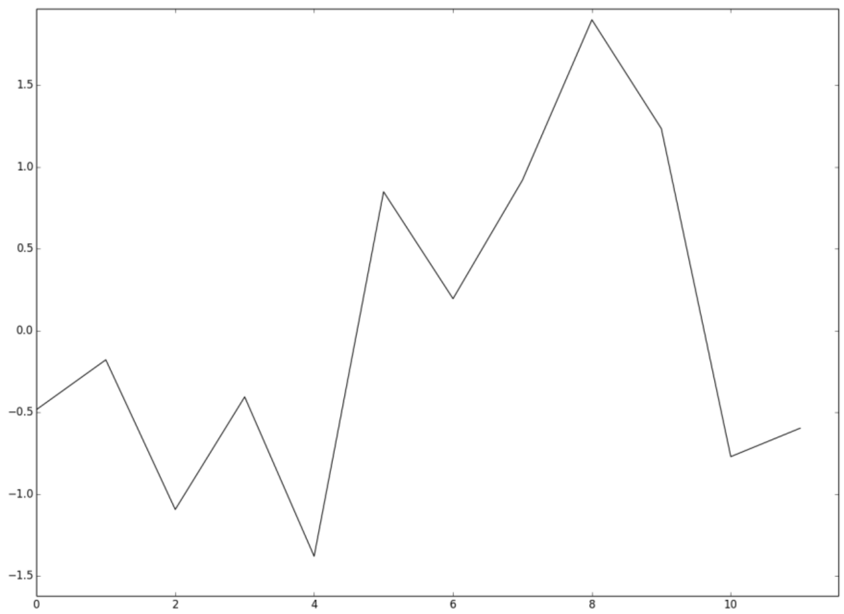
Table 32 Coefficients of the days of the week for the Peg Nord (additive)



Graph 47 Seasonality of the days of the week of the Peg Nord (additive moving averages)

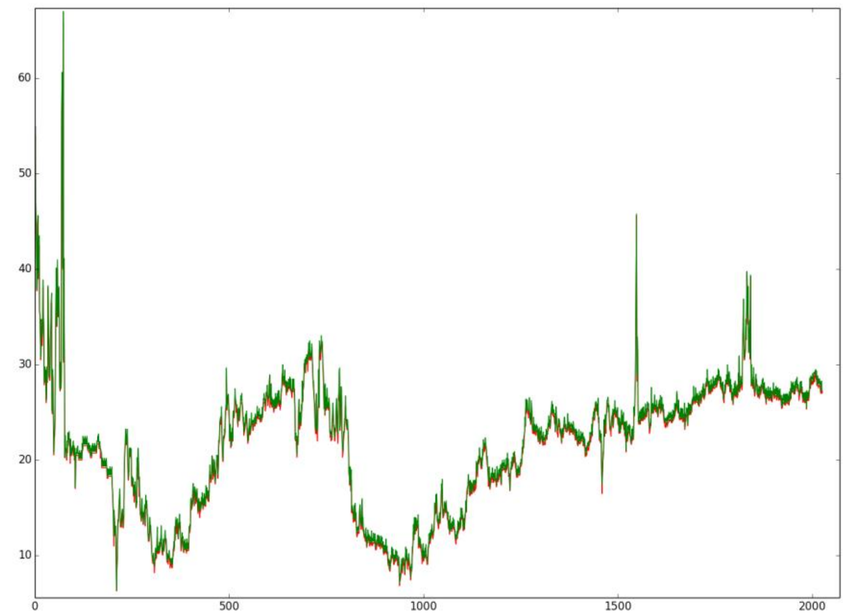
Date	Coefficient
1	-0,48248372
2	-0,17777537
3	-1,09271755
4	-0,40440539
5	-1,37872389
6	0,84936421
7	0,19573036
8	0,92050375
9	1,90069987
10	1,23552581
11	-0,76972007
12	-0,59564187

Table 33 Monthly coefficients for the Peg Nord (additive)



Graph 48 Seasonality of the months of the Peg Nord (additive moving averages)

Original price and seasonally adjusted



Graph 49 Original and seasonally adjusted time series of the Peg Nord (additive moving averages)

Original and seasonally adjusted flattened

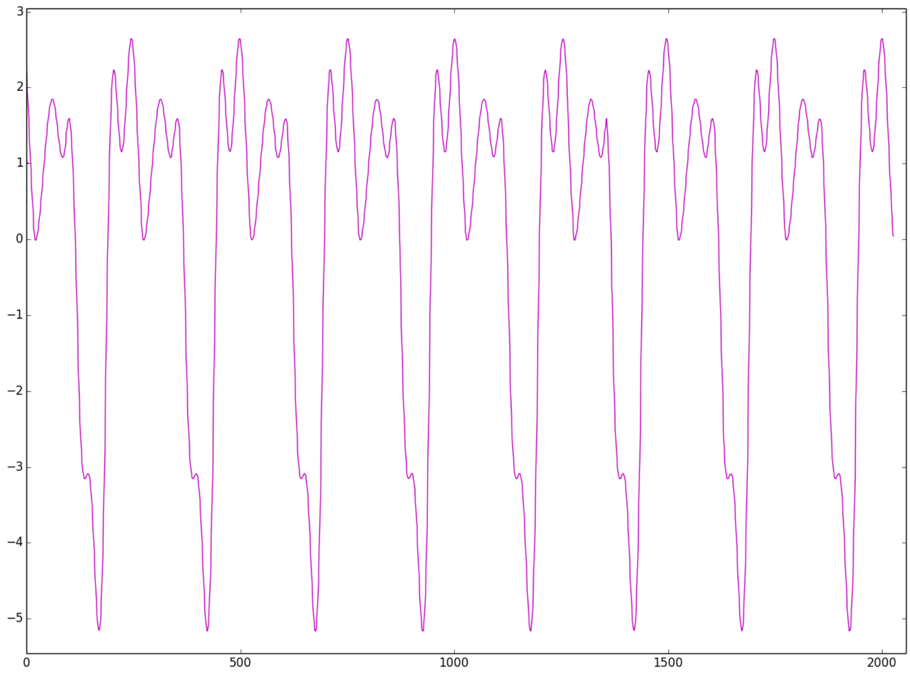
Graph 50 Flattened original and seasonally adjusted time series of the Peg Nord (additive moving averages)

10.5.3 Trigonometric model

In the Graph 51 is painted the seasonality that the trigonometric model provides for the French gas prices.

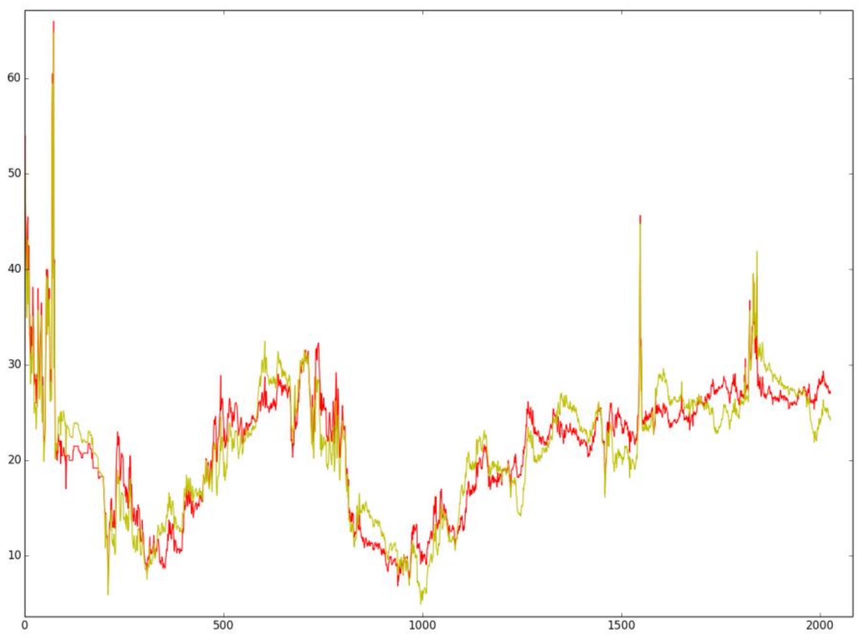
This underlying asset has much more spikes than in the case of the electricity, as it is commented in the conclusions.

Is shown that the fact of being summer makes the gas prices go down, and in cold months they go up.



Graph 51 Seasonal coefficients for the French spot prices (trigonometric)

In the graph 52 are shown the original prices and the seasonally adjusted ones. As with other underlying assets, in this model is seen that the seasonally adjusted prices go further than with the moving averages models.



Graph 52 Original and seasonally adjusted time-series of the Peg Nord (trigonometric)

Graph 53 Flattened original and seasonally adjusted time-series of the Peg Nord (trigonometric)

11 CONCLUSIONS

Now, the main conclusions of the Master Thesis are going to be presented, divided in two sections: the first one regarding a comparison between Spanish and French electricity, and the second one more general.

11.1 Spanish and French comparison

- When comparing both Spanish and French electricity spot prices, the two markets show a correlation with the monthly seasonality. In the two cases, the fact of being summer makes them to have lower prices than what they should have. While in Spain the lowest seasonality is achieved in September, in the French case it is obtained in October (it is like one month delayed).

- Other difference is that while in the French electricity March, April and May are quite flat; in Spain these three months are very different among them, having the lowest seasonality in April. For the Spanish case, this could be due to the Easter holidays.

- In spring, in France the lowest seasonality is for May, and the highest for March-April-May (they are very similar). In Spain, however, the highest seasonality is for May.

- In January, France has got a seasonal coefficient below one, and Spain has got higher than one, so it affects in the opposite way. While in France the fact of being January makes the prices go down, in Spain happens the other way around.

- The highest coefficients obtained for the French electricity are slightly above 1.1, in March, April, May and July.

- The highest coefficient obtained for the Spanish electricity is located between 1.05 and 1.1, and is only for February.

11.2 In general

- The surprising thing is the difference in the seasonality for the Spanish electricity depending on whether you are talking about spot prices or base-load. The main conclusion is that the base-load for the consumer is much more predictable, and that makes a clearer differentiation between week days and weekends. The remaining part (the one that was not contracted as base-load) is contracted in the spot market, and this is the reason why the difference is much bigger. As result, in the base-load the seasonality is much flatter during the week days and has bigger differences for the spot prices.

- As shown in the table 17, the best model for representing the seasonality is the trigonometric one. Its main advantage is that any curve can be represented by sines and cosines and can be very accurate.

- One of the possible drawbacks of the trigonometric model is that could happen that apart from representing the seasonality, could also be overfitting it. So developing a test for checking this possible problem would be suitable.

- When the French electricity prices must be analyzed, the best model for representing the seasonality is the trigonometric one, because when testing the maximum likelihood the result is much lower than for the moving averages models.

- One of the main differences between the trigonometric model and the two moving averages, is that while for the moving averages is possible to differentiate among months (is a discrete model), in the trigonometric a continuous seasonality is obtained, as shown in the plotted results.

- One of the main issues to improve is related to the methodology for removing the spikes. Earlier in this document was proposed the method explained by Cuaresma J. C. et al (2004), and apart from this reference, Lapuerta and Moselle (2001) proposed another one. Whether an accurate model is developed, the differences in prices are not going to be so high and the uncertainty is going to be lower.

- When analyzing the French electricity and gas trigonometric coefficients, is shown the high correlation between them. In summer they go down at the same time and before summer they present two spikes, as well as after the summer. In the particular case of the electricity, the spikes are much higher than in the case of the gas.

- In the future, would be a good idea to develop an autocorrelation function in order to check the behavior for electricity time series. We believe it is not a good model for the electricity because it could lead us to amplify the distortions due to the spikes. But we will have to wait until the model is developed and compare the results.

- In a future analysis, the different cycles (periodic oscillations with higher amplitude than a year) that affect to the electricity should be taken into account.

- A model should be developed for differencing the pure stochastic intrinsic error from the error that is generated due to our models.

12 REFERENCES

Bierbrauer, M., Menn, C., Rachev, S. T., & Trück, S. (2007). Spot and derivative pricing in the EEX power market. *Journal of Banking & Finance, 31*(11), 3462-3485.

Burger, M., Graeber, B., & Schindlmayr, G. (2008). *Managing energy risk: An integrated view on power and other energy markets* John Wiley & Sons.

Cartea, A., & Figueroa, M. G. (2005). Pricing in electricity markets: A mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance, 12*(4), 313-335.

Clewlow, L., & Strickland, C. (2000). *Energy derivatives: Pricing and risk management* Lacima Publ.

Crespo Cuaresma, J., Hlouskova, J., Kossmeier, S., & Obersteiner, M. (2004). Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy, 77*(1), 87-106.

Espeja, E. (2003). *Valoración de contratos" forward" de la electricidad* Centro de Estudios Monetarios y Financieros.

Geman, H., & Roncoroni, A. (2006). Understanding the fine structure of electricity prices*. *The Journal of Business, 79*(3), 1225-1261.

Gustavo Niemeyer. Retrieved from http://labix.org/python-dateutil#head-de6843b3f85a4d89902482be0a545d424674bc6f

Hamilton, J. D. (1994). *Time series analysis* Princeton university press Princeton.

Hannan, E. J., Terrell, R., & Tuckwell, N. (1970). The seasonal adjustment of economic time series. *International Economic Review,* , 24-52.

Joshi, M. S. (2003). *The concepts and practice of mathematical finance* Cambridge University Press.

Knoll, D. A., & Keyes, D. E. (2004). Jacobian-free Newton–Krylov methods: A survey of approaches and applications. *Journal of Computational Physics, 193*(2), 357-397.

Lindstädt, H.Energy trading and its relevance for european energy companies.

Lucia, J. J., & Schwartz, E. S. (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research, 5*(1), 5-50.

Marzal, A., & Gracia, I. (2003). Introducción a la programación con python. *Universitat Jaume I,*

Oracle.Instant client downloads for microsoft windows. Retrieved from http://www.oracle.com/technetwork/topics/winx64soft-089540.html

Osborn, D. R. (2001). *The econometric analysis of seasonal time series* Cambridge University Press.

Perktold, J., Seabold, S. & Taylor, J.StatsModels: Statistics in python. Retrieved from http://statsmodels.sourceforge.net/devel/index.html

Picca, C.Python desde cero: Bases de datos. Retrieved from http://codehero.co/python-desde-cero-bases-de-datos/

Python Software Foundation.The mission of the python software foundation is to promote, protect, and advance the python programming language, and to support and facilitate the growth of a diverse and international community of python programmers. Retrieved from https://www.python.org/

Sintes Marco, B.Introducción a la programación con python. Retrieved from http://www.mclibre.org/consultar/python/index.html

The Scipy community.Optimization (scipy.optimize). Retrieved from http://docs.scipy.org/doc/scipy/reference/tutorial/optimize.html

van Werven, M., & Scheepers, M. (2005). The changing role of distribution system operators in liberalised and decentralising electricity markets. *Future Power Systems, 2005 International Conference on,* 6 pp.-6.

13 ANNEXES

13.1 Regarding the method for removing spikes with monthly seasonality

Regarding the first procedure of removing the spikes, is important to consider the results of this system. Although is out of the scope of this project, is very important to implement an accurate system to remove the spikes, because setting different thresholds, in some months of the year is possible to see differences in prices that could leave you out of the market.

In this table is shown the seasonality using the moving averages method setting different thresholds. Limit in 5 means that whether the real price differs more than 5 €/MWh from the mean for that day, the spike is substituted by its mean. Is possible to see that the month with biggest differences is July, and the one with the smallest differences is September.

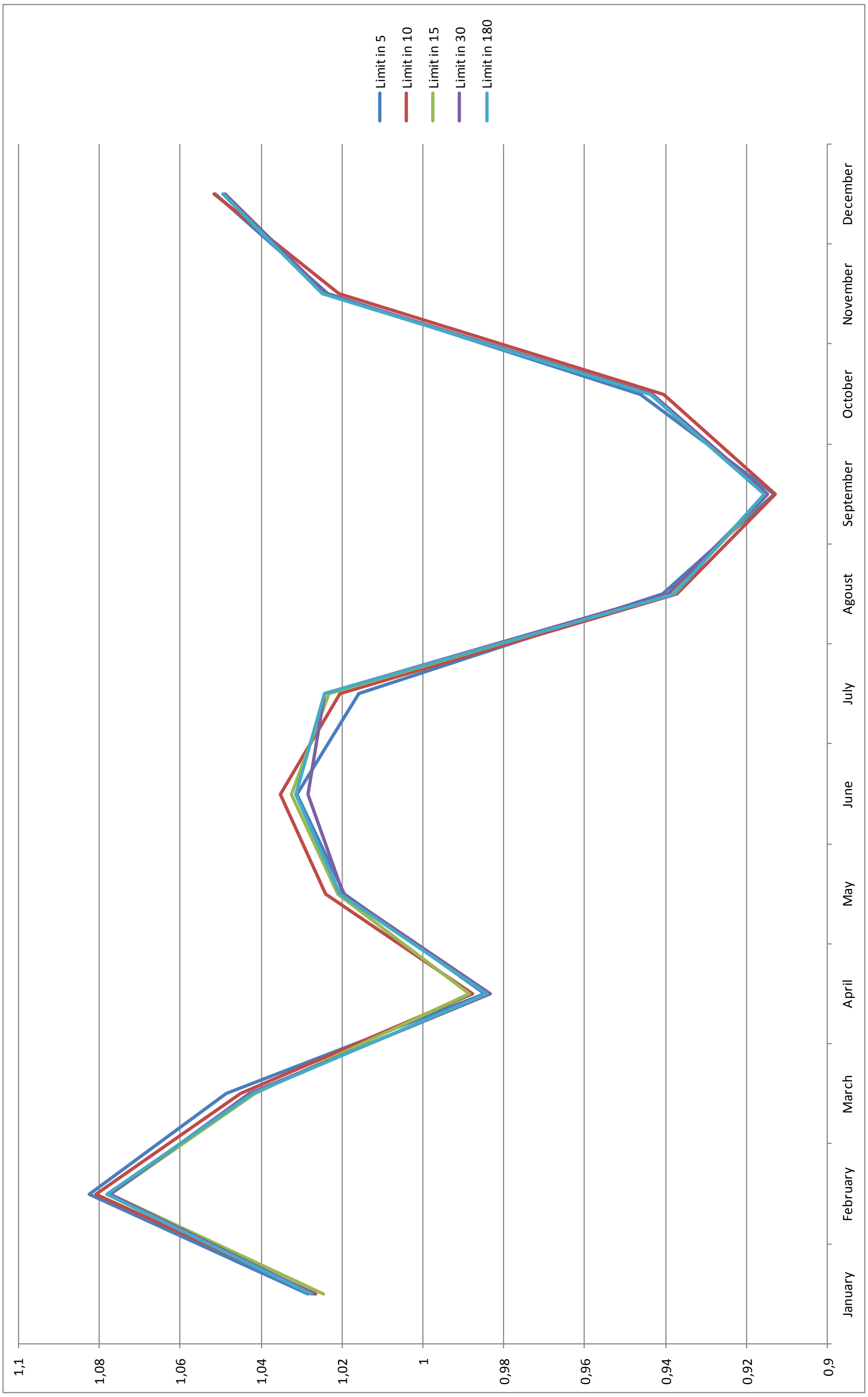
These results show that whether in July are trading electricity in the range of 50€/MWh, the difference in price is going to be around 0.43€/MWh (50*0.00860198), and September it is going to be around 0.14€/MWh (50*0.00269493).

This has been computed with the Spanish spot prices.

	Limit in 5	Limit in 10	Limit in 15	Limit in 30	Limit in 180	MAX value for each month	MIN value for each	Difference
January	1,02852455	1,026458147	1,02455614	1,02695228	1,02770372	1,028524545	1,02455614	0,00396841
February	1,0824777	1,080839587	1,077462591	1,077455726	1,078172094	1,082477698	1,07745573	0,00502197
March	1,04871188	1,045020957	1,041530716	1,042868674	1,042170098	1,048711881	1,04153072	0,00718117
April	0,98484429	0,987646281	0,988613285	0,983284421	0,984041795	0,988613285	0,98328442	0,00532886
May	1,01948554	1,024093115	1,020962134	1,01976941	1,020568811	1,024093115	1,01948554	0,00460758
June	1,03125496	1,035245516	1,032441197	1,028553642	1,031414424	1,035245516	1,02855364	0,00669187
July	1,01582729	1,020549086	1,02317952	1,024200166	1,024429276	1,024429276	1,01582729	0,00860198
Agoust	0,94076811	0,937085721	0,938397	0,939490486	0,937690213	0,940768107	0,93708572	0,00368239
September	0,91327495	0,912833381	0,914865155	0,914872161	0,91552831	0,91552831	0,91283338	0,00269493
October	0,94612559	0,940478099	0,943539073	0,943178161	0,943842655	0,946125594	0,9404781	0,00564749
November	1,02349935	1,020884909	1,02460896	1,024020601	1,024761058	1,024761058	1,02088491	0,00387615
December	1,05147841	1,051597221	1,049470674	1,048808214	1,049548956	1,051597221	1,04880821	0,00278901

Table 34 Differences in monthly seasonal coefficients depending on where is set the threshold

In the next graph is represented the seasonality of the different months based on the results shown in the previous table.



Graph 54 Monthly seasonal coefficients

13.2 Regarding the method for removing spikes with daily seasonality

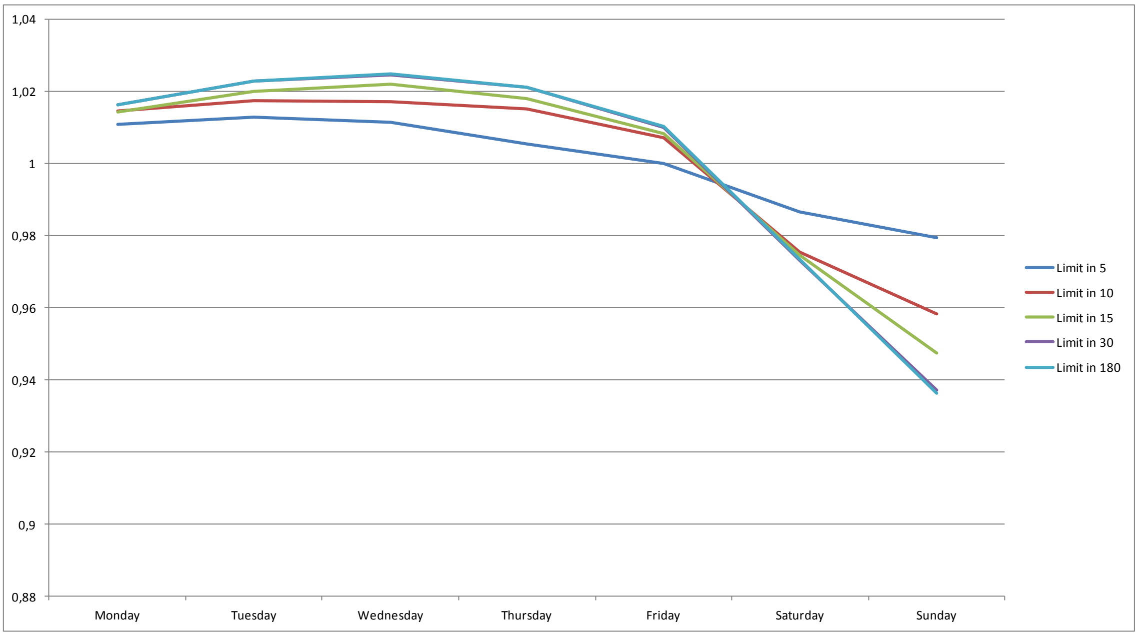
In this case, is possible to see that the differences in prices are much higher than in the case of the monthly seasonality.

In the next table are shown the different coefficients depending on the day of the week. The highest difference is for the Sundays (in red) and the lowest is for the Mondays (in green). This means that if it is assumed that the electricity is being traded at a price of 50 €/MWh, on Mondays it is going to have a price gap of 0.28€/MWh and on Sundays the difference is going to be around 2.16 €/MWh.

	Limit in 5	Limit in 10	Limit in 15	Limit in 30	Limit in 180	MAX value for each day	MIN value for each day	Difference
Monday	1,01086192	1,014473534	1,014300871	1,016173161	1,016389329	1,016389329	1,010861917	0,00552741
Tuesday	1,012931	1,017309365	1,020075122	1,022782481	1,022998832	1,022998832	1,012931001	0,01006783
Wednesday	1,01149854	1,01710852	1,022110009	1,024574379	1,024795584	1,024795584	1,01149854	0,01329704
Thursday	1,00555985	1,015153322	1,017993798	1,021067746	1,021288628	1,021288628	1,005559847	0,01572878
Friday	1,00001828	1,007110818	1,008189591	1,009958481	1,010186447	1,010186447	1,000018282	0,01016816
Saturday	0,98654716	0,975417466	0,974606323	0,973164817	0,9733905	0,986547159	0,973164817	0,01338234
Sunday	0,97956181	0,958224365	0,947414634	0,937039608	0,936373421	0,979561813	0,936373421	0,04318839

Table 35 Differences in daily seasonal coefficients depending on where is set the threshold

In the graph of the following page are shown the results in a picture. It is deduced that the seasonal coefficients with the threshold set in 30 and without a threshold are practically the same. When the differences among the thresholds set in 15 and 30 are compared, we get that the difference in prices is 0.07€ for the Saturdays (the smallest one) and 0.52€ for the Sundays (always assuming that the spot price traded for these days is 50€/MWh).

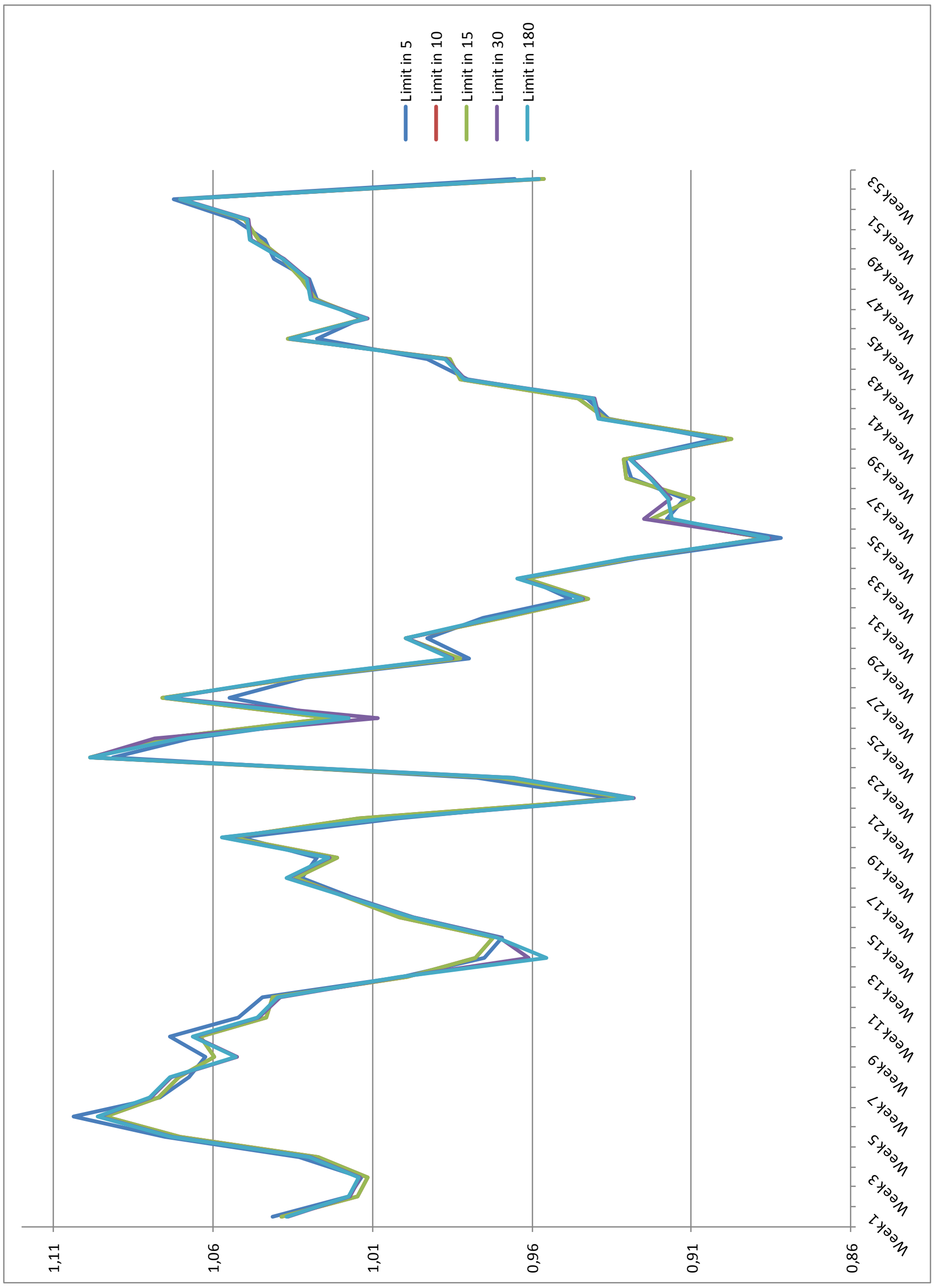


Graph 55 Daily seasonal coefficients

13.3 Regarding the method for removing spikes with weekly seasonality

In the case of the weekly seasonality, the fact of removing the spikes by the method that was explained earlier does not affect as much as for the daily seasonality, but still having some influence where you set the threshold.

The highest difference in prices comes from the week 14, and multiplying that spread time the spot price (assuming 50€/MWh) it is obtained 1.12€, and taking into account the smallest difference in the week 17 the gap is around 0.083€.



Graph 56 Weekly seasonal coefficients

	Limit in 5	Limit in 10	Limit in 15	Limit in 30	Limit in 180	MAX value for each week	MIN value for each week	Difference
Week 1	1,04119403	1,03870033	1,03870033	1,03655725	1,03697159	1,04119403	1,03655725	0,00463678
Week 2	1,01733955	1,0147081	1,0147081	1,01702264	1,01743736	1,01743736	1,0147081	0,00272927
Week 3	1,01344808	1,01168339	1,01168339	1,01401817	1,01442626	1,01442626	1,01168339	0,00274287
Week 4	1,03290262	1,02693023	1,02693023	1,02929518	1,02970189	1,03290262	1,02693023	0,00597239
Week 5	1,0753736	1,07029036	1,07029036	1,07264771	1,07304759	1,0753736	1,07029036	0,00508323
Week 6	1,10374863	1,09375536	1,09375536	1,09603365	1,0964341	1,10374863	1,09375536	0,00999326
Week 7	1,07695024	1,07714873	1,07714873	1,07946633	1,07987769	1,07987769	1,07695024	0,00292745
Week 8	1,06769518	1,0706942	1,0706942	1,07301313	1,07343309	1,07343309	1,06769518	0,00573792
Week 9	1,06246286	1,05981638	1,05981638	1,05247998	1,05287081	1,06246286	1,05247998	0,00998288
Week 10	1,07379229	1,06387932	1,06387932	1,06616677	1,06658945	1,07379229	1,06387932	0,00991296
Week 11	1,05202116	1,04351252	1,04351252	1,04578835	1,04620873	1,05202116	1,04351252	0,00850864
Week 12	1,04474939	1,0412315	1,0412315	1,0391915	1,03960204	1,04474939	1,0391915	0,00555789
Week 13	1,0006868	0,99927233	0,99927233	1,00137113	1,00179225	1,00179225	0,99927233	0,00251992
Week 14	0,97513179	0,97784445	0,97784445	0,96123138	0,95535565	0,97784445	0,95535565	0,0224888
Week 15	0,96953308	0,9722191	0,9722191	0,97039904	0,97085321	0,9722191	0,96953308	0,00268602
Week 16	0,99723821	1,00158517	1,00158517	0,99716853	0,99762389	1,00158517	0,99716853	0,00441664
Week 17	1,01667326	1,01833855	1,01833855	1,01753395	1,01803734	1,01833855	1,01667326	0,0016653
Week 18	1,03216178	1,03409113	1,03409113	1,03648941	1,03699443	1,03699443	1,03216178	0,00483265
Week 19	1,02754905	1,02099028	1,02099028	1,02351715	1,0240155	1,02754905	1,02099028	0,00655877
Week 20	1,05144299	1,05433284	1,05433284	1,05680565	1,05734383	1,05734383	1,05144299	0,00590084
Week 21	1,00184604	1,01422435	1,01422435	1,00466138	1,00518294	1,01422435	1,00184604	0,01237832
Week 22	0,93482831	0,93001151	0,93001151	0,92805017	0,92855277	0,93482831	0,92805017	0,00677814
Week 23	0,97721421	0,97061093	0,97061093	0,96587127	0,96640427	0,97721421	0,96587127	0,01134294
Week 24	1,09142991	1,0987297	1,0987297	1,09828564	1,09883252	1,09883252	1,09142991	0,00740261
Week 25	1,06721554	1,07586515	1,07586515	1,07855673	1,07032843	1,07855673	1,06721554	0,01134119
Week 26	1,02045038	1,02294675	1,02294675	1,00822728	1,01742117	1,02294675	1,00822728	0,01471947
Week 27	1,05481563	1,07601874	1,07601874	1,07453933	1,07493412	1,07601874	1,05481563	0,02120311
Week 28	1,03089277	1,03205272	1,03205272	1,03459483	1,03490744	1,03490744	1,03089277	0,00401467
Week 29	0,97988514	0,98235899	0,98235899	0,98484581	0,98525916	0,98525916	0,97988514	0,00537401
Week 30	0,99274518	0,99945388	0,99945388	0,99924316	0,9996138	0,9996138	0,99274518	0,00686862
Week 31	0,97552512	0,96940193	0,96940193	0,97150478	0,97181679	0,97552512	0,96940193	0,00612319
Week 32	0,94782544	0,942456	0,942456	0,94405711	0,94448788	0,94782544	0,942456	0,00536945
Week 33	0,96203506	0,96165924	0,96165924	0,96404261	0,96450159	0,96450159	0,96165924	0,00284236
Week 34	0,92653287	0,92808333	0,92808333	0,92959561	0,93006147	0,93006147	0,92653287	0,0035286
Week 35	0,88188679	0,88670735	0,88670735	0,8853025	0,88579448	0,88670735	0,88188679	0,00482056
Week 36	0,91775152	0,92242083	0,92242083	0,92496432	0,9159274	0,92496432	0,9159274	0,00903692
Week 37	0,91201207	0,90942845	0,90942845	0,91665349	0,91713036	0,91713036	0,90942845	0,00770192
Week 38	0,92873488	0,93050276	0,93050276	0,92257081	0,92302244	0,93050276	0,92257081	0,00793196
Week 39	0,93083858	0,93133242	0,93133242	0,92873609	0,92923532	0,93133242	0,92873609	0,00259633
Week 40	0,90222949	0,89733683	0,89733683	0,89946792	0,89983717	0,90222949	0,89733683	0,00489266
Week 41	0,93588483	0,93667251	0,93667251	0,93886265	0,93924748	0,93924748	0,93588483	0,00336265
Week 42	0,94293231	0,94542018	0,94542018	0,94028464	0,9406701	0,94542018	0,94028464	0,00513553
Week 43	0,98003748	0,98270514	0,98270514	0,9807698	0,98114259	0,98270514	0,98003748	0,00266766
Week 44	0,99304007	0,9857961	0,9857961	0,98683018	0,98721276	0,99304007	0,9857961	0,00724398
Week 45	1,02727699	1,03659995	1,03659995	1,0356193	1,03600811	1,03659995	1,02727699	0,00932296
Week 46	1,01348473	1,0122492	1,0122492	1,01164002	1,01203568	1,01348473	1,01164002	0,00184472
Week 47	1,02727116	1,02778088	1,02778088	1,028996	1,02939886	1,02939886	1,02727116	0,0021277
Week 48	1,02994853	1,03212357	1,03212357	1,03035222	1,03075009	1,03212357	1,02994853	0,00217504
Week 49	1,04081219	1,03838342	1,03838342	1,0376852	1,03808329	1,04081219	1,0376852	0,00312699
Week 50	1,0437705	1,04560515	1,04560515	1,04797431	1,04837895	1,04837895	1,0437705	0,00460845
Week 51	1,05329657	1,05000046	1,05000046	1,0488366	1,04924031	1,05329657	1,0488366	0,00445997
Week 52	1,07243432	1,0705739	1,0705739	1,06992032	1,07032627	1,07243432	1,06992032	0,00251399
Week 53	0,96547969	0,95644176	0,95644176	0,9578276	0,95779952	0,96547969	0,95644176	0,00903793

Table 36 Differences in weekly seasonal coefficients depending on where is set the threshold