Acta logistica - International Scientific Journal about Logistics
Volume: 9 2022 Issue: 2 Pages: 151-169 ISSN 1339-5629

doi:10.22306/al.v9i2.290 *Received: 14 Jan. 2022; Revised: 18 Mar. 2022; Accepted: 02 Apr. 2022*

DEVELOPMENT OF A NEW NUMERICAL MODEL OF DYNAMIC HARMONIC REGRESSION FOR THE FORECAST OF SELLING FUEL PRICE IN THE MOROCCAN PETROLEUM SECTOR

Younes Fakhradine El Bahi
Mechanical and Integrated Engineering, Industrial department, 15290 ENSAM, Meknes 50500, ENSAM School, Marocco, elbahiyounesfakhradine@gmail.com
Latifa Ezzine
Mechanical and Integrated Engineering, Industrial department, 15290 ENSAM, Meknes 50500, ENSAM School, Marocco, latifae@yahoo.com
Zineb Aman
Mechanical and Integrated Engineering, Industrial department, 15290 ENSAM, Meknes 50500, ENSAM School, Marocco, zineb.aman@gmail.com (corresponding author)
Haj El Moussami
Mechanical and Integrated Engineering, Mechanical department, 15290 ENSAM, Meknes 50500, ENSAM School, Marocco, hajelmoussami@yahoo.com

Keywords: liberalization, forecast, time series model, dynamic harmonic regression, decision making.
Abstract: The liberalization of the petroleum sector in Morocco has a significant effect for petroleum product distributors. Since the beginning of December 2015, fuel prices are freely determined. This event presents many constraints affecting the balance of the sector plus the competition among its economic players. As all fuel products are imported, we will be interested in the evolution by making forecasts of the price of fuels in the Moroccan market. In this context, our paper aims mainly to study the time series of diesel and gasoline in order to provide precise forecasts to the company and to respect the permissible error margin of 3%. To this end, the harmonic dynamic regression model through the proposed process approach yielded excellent forecasting results for the first quarter of 2017 with an average error margin of 1.617%. Compared to ARIMA model, the harmonic dynamic regression proves its strength manifested in the low rate of error. In addition, the assumption that the residuals are a Gaussian white noise has always been verified. The forecasts obtained are very crucial for managers to take good decisions at the strategic level.

1 Introduction

Oil is an important sector given the increasingly competitive nature of the industry today. Large companies and oil suppliers want the most economical and reliable forecasting mechanism to evaluate the market. For this reason fuel price forecasting is one of the most important problems beyond all strategic and planning decisions in any company [1], but difficult in the research areas of the analysis and prediction of because of the uncertain interactive factors that determine the oil market.On the one hand, oil is directly associated with various uncertain market factors, such as supply and demand, competition among suppliers, substitution for other forms of energy, economic development, population growth, and technical development [2]. On the other hand, as crude oil is a dominant resource for energy security, the oil market is highly sensitive to various uncertain external factors, such as political instability, war and conflict [3,4]. Since 01 December 2015, fuel prices are freely fixed. This event presents a lot of constraints impacting the balance of the sector and the competition among its economic actors. The lack of support measures by the state makes it vulnerable. With the cessation of the activity of the only Moroccan refinery, distributors must, from their part, build up large stocks. As all fuel products are imported, we will be interested in the evolution by making forecasts of the price of fuels on the Moroccan market. In order to remain competitive on the market, distributors must be able to source fuel at the optimal price while maintaining a significant profit margin. In order to achieve their objectives, oil companies must rely on very specific forecast mechanisms.The overall mechanism by which the price of oil affects most countries or national economic factors are generally well understood, and thus the forecast of the price of oil has been perceived as an important research topic.One of the most commonly used approaches for oil price prediction is the statistical time series method [5,6], characterizing the price of oil as a time trend, a seasonal factor, a cyclical element and a term of error. Many techniques are available to break a series of oil prices in these components.They include the Akarca and Andrianacos model medium autoregressive movement (ARIMA) [7], Lanza et al.'S (ECM) error correction model [8], and the vectorial regression of Mirmirani and Li (VAR) model [9].Other types of approaches assume that stochastic quantification, the relationship between oil prices and the latent economy are factors that can provide a more relevant prediction than attempting to discover the underlying structure of the series itself.

Our article studies the price forecast of the Moroccan oil market, which is an important area for the determination of future value, from the point of view of oil distribution companies, considering global variables. We will develop a new numerical model of dynamic harmonic regression that aims to determine the price of fuel in the forecasting process.

2 Literature review

Several articles have been published in the literature, aiming at comparing and improving the prediction capabilities of dynamic harmonic regression and ARIMA models more precisely. Church and Curram (1996) [10] have attempted to predict the decline in the growth rate of Consumer spending in the late 1980s using a network of artificial neurons and a model ARIMA. They suggest that neural network models describe decline as well as, but not better, econometric specifications. Finally, they suggest that the selection of the methodology used to make predictions should be based solely on the variables available. Prybutok et al. (2000) [11] develop a neural network model for predicting daily maximum ozone levels and compare with two conventional statistical models, regression and ARIMA Box-Jenkins models. The results show that the neural network is superior to the other two models tested. Gutierrez-Estrada et al. (2004) [12] use linear multiple regression, univariate time series models (ARIMA models), and neural network computations to predict the average daily ammonia concentration in water-recirculating ponds. The results show that the non-linear ANN model approach provides better prediction of ammonia concentration than the multiple linear regression and univariate time series analysis, when the correlation between the data series is weak and when models are forced to predict in a situation for which they have not been specifically calibrated. Ho et al. (2002) [13] present a comparative study of integrated self-regulating movement Box-Jenkins models of medium and two artificial neural networks with a different architecture (multilayer power supply and recurrent network) in the prediction of failures in computer networks. The study showed that the recurrent network architecture significantly outperforms the results produced by the other two methodologies. In many scientific or technical applications, the data is generated in the form of a time series. As a result, time series analysis is one of the main tools for research and development (Cryer 1986, Chatfield 1991) [14-15]. Modelling of univariate and multivariate time series and structural time series methods have been useful for describing and forecasting (Prybutok et al., 2000, Tsitsika et al., 2007) [11,16].

3 Elaboration of the new methodology

In our study, we will develop a new numerical model of dynamic harmonic regression that aims to ameliorate the forecast of the price of fuel in the forecasting process.

In our previous work, El Bahi et al. (2018) [17] demonstrates the utility of ARIMA model in the forecasting field, we found that the ARIMA model (1,1,1) gave accurate forecasts to the price of gasoline near the margin to be met for the first quarter of the current year with an average error margin of 2,855% respecting thus the margin permissible given by the company. However, in this paper, we will present a new model of dynamic harmonic regression through a proposed new process approach.

3.1 Dynamic harmonic regression

We have developed the basics of the new harmonic dynamic regression model, while performing various tests to validate the proposed model. Our study has focused on the price of diesel fuel in Morocco.

The classic model of additive time series with hidden components is shown in (1).

$$\boldsymbol{Y_t = T_t + C_t + S_t + f(u_t) + N_t + e_t\,;\;\; e_t \sim N(0, \sigma^2)} \quad (1)$$

- $\boldsymbol{Y_t}$: the observed time series
- $\boldsymbol{T_t}$: the trend
- C_t : the cyclical component
- $\boldsymbol{S_t}$: the seasonal component
- $\boldsymbol{f}$: the influence of an exogenous vector with variables $\boldsymbol{u_t}$
- $\boldsymbol{N_t}$:stochastic disturbance
- e_t : Gaussian white noise: $\boldsymbol{e_t \sim N(0, \sigma^2)}$[18].

Nevertheless, we worked only with the following model:

$$\boldsymbol{S_t = \sum_{i=1}^{n} A_i cos(\omega_i t + \varphi_i)} \quad (2)$$

$$\boldsymbol{C_t = \sum_{j=1}^{m} B_i cos(\omega_j t + \varphi_j)} \quad (3)$$

The ω_i andω_j are the Fourier frequencies determined from the discrete Fourier transform through a spectral analysis. This study in the frequency and non-temporal domain allows a consequent gain in parameters [19].

It is interesting to separate the cyclical and seasonal components because even if they are written in the same way, the cycles have relatively large periods compared to the observations and are therefore hardly detectable. In our work, we consider the cyclical component equal to zero because the observations relate only to one year (24 observations / fortnight in all). The parameters n and m depend on the number of observations p. We will see later that:

$$\boldsymbol{n = \frac{p}{3} - 1} \quad (4)$$

For the trend, we will see that the regression line can be a good approximation. The most important thing is that we will detail later that everything is calculated at the same time [20]!

In this study, we will work on the prices of diesel fuel for the development of the model. The points cloud and regression line are presented in figure 1.

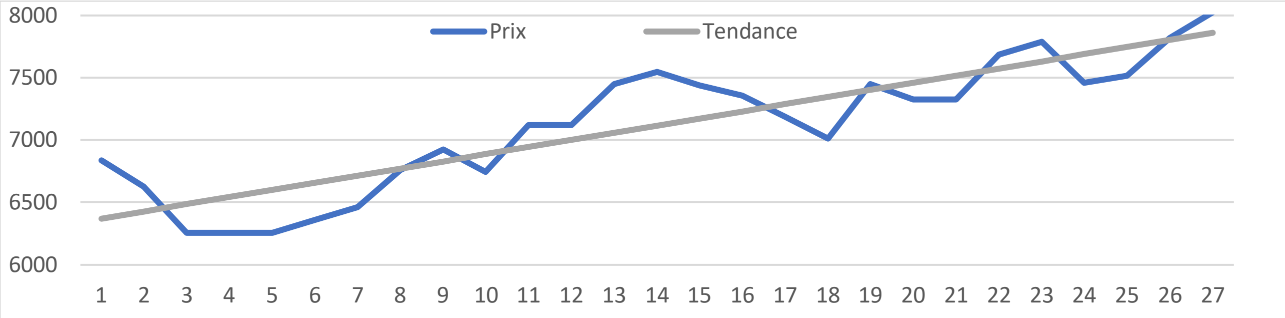

Figure 1 Gas price excluding taxes and trend

3.2 Determination of Fourier frequencies

3.2.1 Methodology

To determine the frequencies of Fourier, we will present two methods: the direct method of the periodogram and the indirect method of Blackman-Tuckey.

Under the assumptions of stationary and ergodic signals, the autocorrelation function and the power spectral density are defined by:

$$r_{xx}(m) = E\{x^*(n)x(n+m)\} = r_{xx}^*(-m) \quad (5)$$

$$S_x(f) = \sum_{m=-\infty}^{+\infty} r_{xx}(m)e^{-2j\pi mf} = \lim_{N\to\infty} E\left\{\frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)e^{-2j\pi nf}\right|^2\right\} \quad (6)$$

It is assumed to have N samples $\{x(n)\}_{n=0}^{N-1}$ of the signal to be analyzed and we, therefore, try to estimate the power spectral density from these data. There are two major classes of non-parametric spectrum estimation, each of which is related to one of the equalities in the spectral density equation.

Periodogram method:

It is a method that uses the signal directly. The spectrum is estimated by (7).

$$\hat{S}_{PER}(f) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)e^{-2j\pi nf}\right|^2 \quad (7)$$

Because of the truncation of the signal, the periodogram is in fact the convolution of the spectrum by a cardinal sinus window. Truncating the signal induces two main phenomena:

- Widening of the main lobe which leads to a decrease of the resolution.
- Appearance of secondary lobes.

The resolution of the periodogram is $1/N$.

To remedy truncation problems, windows are generally used to either reduce the main lobe or attenuate the sidelobes. Also, the periodogram is a biased estimator.

$$E\{\hat{S}_{PER}(f)\} = \int_{-1/2}^{1/2} W_b(f-u)S_x(u)du \quad (8)$$

Where $W_b(f) = \frac{1}{N}\left[\frac{sin(\pi Nf)}{sin(\pi f)}\right]$ is the Fourier transform of the triangular window.

The periodogram is therefore on average the convolution of the true spectrum with the Fourier transform of the triangular window. Nevertheless, when $N \to \infty$, the bias becomes zero. It can also be shown that the variance is virtually independent of N and proportional to the spectrum.

$$var\{\hat{S}_{PER}(f)\} \cong S_x(f)^2 \quad (9)$$

Thus, the periodogram is therefore not a consistent estimator of the Spectral Power Density. In order to reduce the variance of this estimator, an averaged periodogram can be used.

This consists of separating the signal into K slices (of length N / K), calculating the periodogram on each slice and averaging. Due to the average K, the variance is almost divided by K: nevertheless, with the slices being shorter, the resolution decreases.

Indirect method: Blackman-Tuckey:

Another approach is to use the definition of the spectrum from the correlation function. The spectrum is then estimated as in (10).

$$\hat{S}_{BT} = \sum_{m=-M}^{M} \hat{r}_{xx}(m)e^{-2j\pi mf} \quad (10)$$

Where $\hat{r}_{xx}$ is an estimator of the correlation function, given by (11).

$$\hat{r}_{xx} = \frac{1}{N}\sum_{k=0}^{N-m-1} x^*(k)x(k+m) \quad (11)$$

It is a biased estimator. The following estimator is unbiased:

$$\hat{r}_{xx} = \frac{1}{N-m}\sum_{k=0}^{N-m-1} x^*(k)x(k+m) \qquad (12)$$

Blackman-Tuckey suggested taking M of the order of 10% N. The application of a window is possible to reduce the variance on the estimate of the correlation function. Indeed, we have:

$$\hat{S}_{PER} = \sum_{m=-(N-1)}^{N-1} \hat{r}_{xx}(m)e^{-2j\pi mf} \qquad (13)$$

However, the variance over $\hat{r}_{xx}(m)$ increases as m approaches N, this explains why only M correlation points are used. Nevertheless, when M decreases, the bias increases.

=> bias-variance dilemma.

Disadvantages:

- Resolution in 1 / N which implies a certain difficulty in finding two very close lines for short duration signals.
- Difficulty in finding weak signals compared to strong signals.
- Secondary lobes and negative spectra (BT) due to windowing.
- Inconsistent estimate of the PSD.

Advantages

- Very fast and inexpensive computing algorithms (FFT).
- Estimated spectrum proportional to the power.
- Robust behavior on a wide range of signals.

The choice of a window results essentially from the compromise between the width of the main lobe and the height of the secondary lobes. In our study, we will work with Hannig's window because it allows a good compromise.

3.2.2 Spectral analysis

This analysis will allow us to bring out the most significant frequencies (which allow a maximum information gain: signal processing theory).
We will classify the frequencies of the most important (spectral density of the largest to the smallest). Here is the result on Table 1.

Using the SPSS software, we will calculate the other parameters by a non-linear regression. The first idea that comes to mind is to take all Fourier frequencies and build the forecast model for these frequencies. Figure 2 shows the result obtained. The results seem satisfactory, but the strong oscillations push us to study them more closely.

3.3 Oscillation study

To be able to do this study, we have to map the model exactly. But we only have 27 observations. Shannon's theorem states that the sampling frequency (observations) must be greater than or equal to twice the maximum frequency. All calculations done, the theoretical sampling frequency must be greater than 50. We propose to complete our samples to make a polynomial interpolation between each two consecutive points so as to keep a line that passes through two consecutive points and the middle point created. The result obtained is shown in Figure 3. We notice a very strong oscillation among the values which pushes us to consider the application of a filter.

3.4 Static filter proposed

The idea is to build a band pass filter that filters high and low frequencies as shown in figure 4.

New constraints are then added to our model. Recall that the model exclaims as in (14).

Table 1 Fourier frequencies

Frequency	Period	Periodogram	Spectral density
0.2327	27.0000	3705812.5981	294899.1966
0.4654	13.5000	2004167.0706	159486.5480
0.9308	6.7500	307353.1302	24458.3850
1.3963	4.5000	171525.5708	13649.5712
0.6981	9.0000	162633.9992	12942.0024
1.6290	3.8571	159402.0988	12684.8160
1.8617	3.3750	85748.8769	6823.6788
2.3271	2.7000	72878.7657	5799.5079
3.0252	2.0769	62642.4908	4984.9310
1.1636	5.4000	57555.0920	4580.0887
2.7925	2.2500	47251.8233	3760.1806
2.0944	3.0000	26826.4541	2134.7814
2.5598	2.4545	20257.9969	1612.0802
0.0000		0.0000	0.0000

Volume: 9 2022 Issue: 2 Pages: 151-169 ISSN 1339-5629

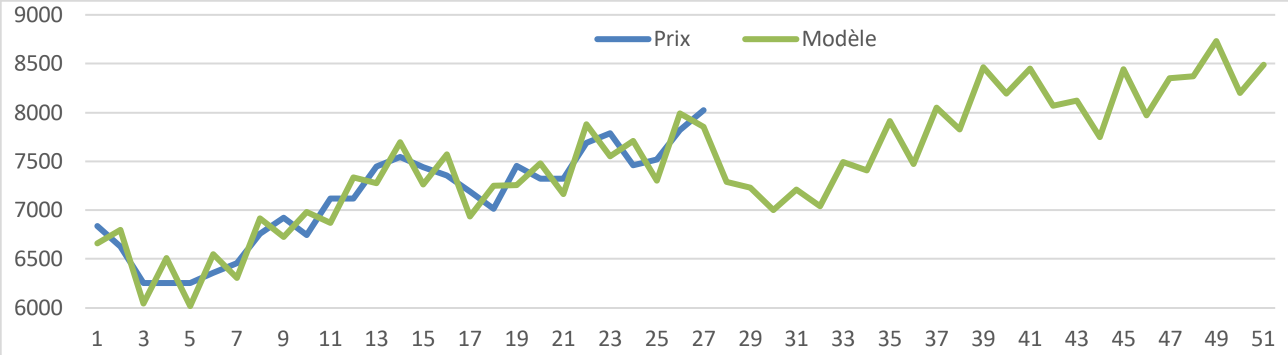

Figure 2 Model of all frequencies

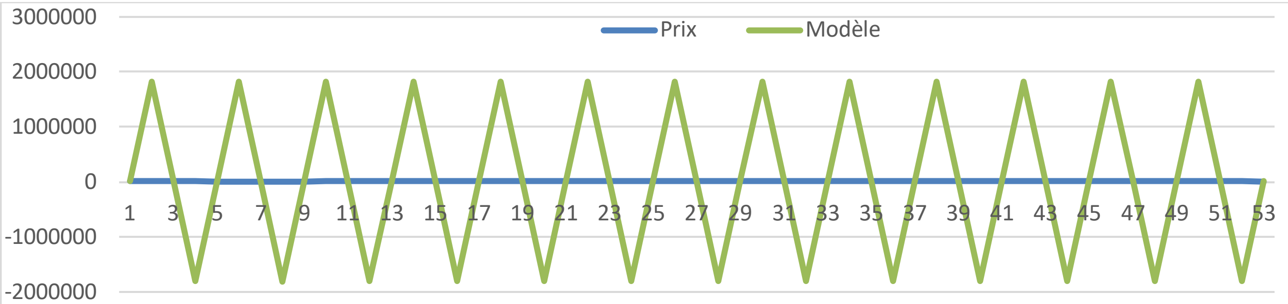

Figure 3 Oscillatory study of the model of all the frequencies

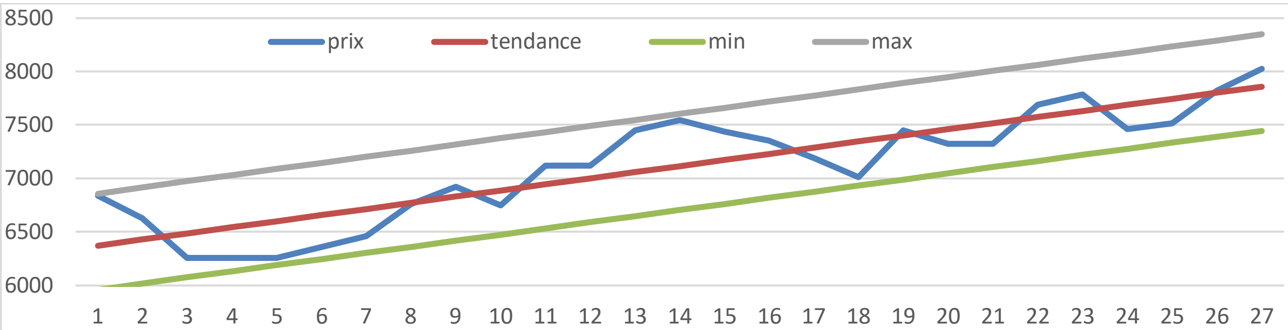

Figure 4 Low band filter

$$Y_t = a + bt + \sum_{i=1}^{p-2} A_i \cos(\omega_i t + \varphi_i) + e_t e_t \sim N(0, \sigma^2) \quad (14)$$

To achieve the filter, we took the bandwidth so that all the elements of the series are between the minimum and the maximum of this one. So each A_i must be less than or equal to a constant (450 in our case). So we will have:

$$S_t = \sum_{i=1}^{\frac{p}{2}-1} A_i cos(\omega_i t + \varphi_i) \quad (15)$$

Similarly, we will do an oscillatory study of the signal or the time series. The result is in figure 5.

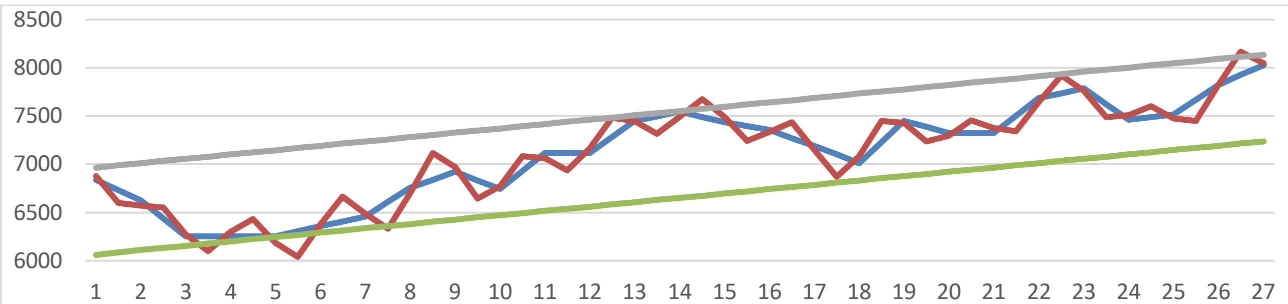

Figure 5 Oscillation study of the static filter

The observed overruns are justified by the fact that we have made a naive assumption that the amplitudes must not exceed one constant sequentially. However, the damping effect of the other amplitudes must be considered. We must therefore consider the sum. Despite this, this filter already demonstrates excellent forecasting qualities. First, let's check the hypothesis that$\boldsymbol{e_t \sim N(0,\sigma^2)}$.

We perform tests on XLSTAT. For normality, we performed two tests, that are Kolmogorov-Smirnov and Khi². Results are grouped in Table 2 [21]:
H_0: The sample follows a Normal law.
H_1: The sample does not follow a Normal law.

Table 2 Normality tests of the static filter

Test		**Kolmogorov-Smirnov**		**Khi² Degree of Freedom =6**	
Parameter	Value	D	0.1178	Observed value	12.3801
µ	-0.0762	P-value	0.8239	Critical value	12.5916
Sigma	42.1218	Alpha	0.05	P-value	0.0540

Interpretations:

- *Kolmogorov-Smirnov:*

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected. The risk of rejecting the null hypothesis H_0 when it is true is 82.39%.

- *Khi²:*

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected. The risk of rejecting the null hypothesis H_0 when it is true is 5.40%.

The last test to be performed is the heteroskedasticity test [22]:

- *White's test:*

Table 3 White's Test

LM (Observed value)	0.6505
LM (Critical value)	5.9915
Degree of Freedom	2
P-value (bilateral)	0.7224
Alpha	0.05

Also, two hypotheses are formulated:
H_0: The residues are homoscedastic
H_1: The residues are heteroscedastic

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected. The risk of rejecting the null hypothesis H_0 when it is true is 72.24%.

Although this static filter seems to provide very good results, it prompts us to think of a new filter that takes into account the aspect of amplitude attenuation: the birth of the dynamic filter.

3.5 Dynamic filter proposed

The idea of the filter is simple: to be able to get as close as possible to the time series with the minimum of oscillations. The figure 6 give us explanation of all this.

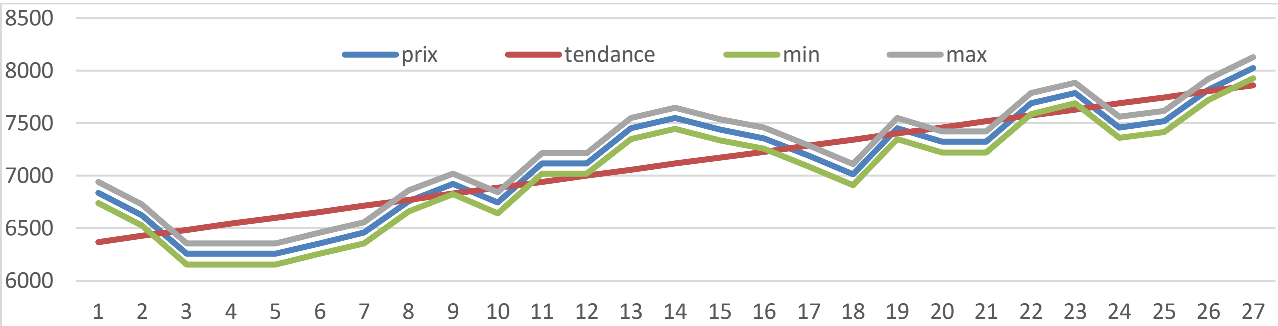

Figure 6 Dynamic filter construction

We will consider a parameter α so that the minimum and maximum are closest to the time series. The new seasonal component becomes:

$$\boldsymbol{S_t = \sum_{i=1}^{\frac{p}{3}-1} A_i cos(\beta_i \omega_i t + \varphi_i)} \tag{16}$$

Because we will have new constraints: First: $-1 < \beta_i \leq 1$ (for all i ranging from 1 to n)
This parameter is inspired by the wavelet theory, notably the Morlet wavelet.

The exponential form is set aside expressly so that there is no exponential envelope (reverse funnel effect) because the series is stationary around the trend.
Two other constraints are added to attenuate the signal or the time series.

$$\boldsymbol{\sum_{i=1}^{n} A_i + \alpha \leq \max(y_t)} \tag{17}$$

$$\boldsymbol{\sum_{i=1}^{n} A_i - \alpha \geq \min(y_t)} \tag{18}$$

According to these new constraints the number of allowed sinusoids becomes: $\mathbf{n = \frac{p}{3} - 1.}$

The first remark is that this filter is greedy and it should only be used in case of significant observations! Secondly, this filter creates a compromise between the loss of information and the attenuation of oscillations. So we have to make sure that $\boldsymbol{e_t \sim N(0, \sigma^2)}$. Results after applying this filter are presented in Figure 7. The test results are grouped in Table 4.

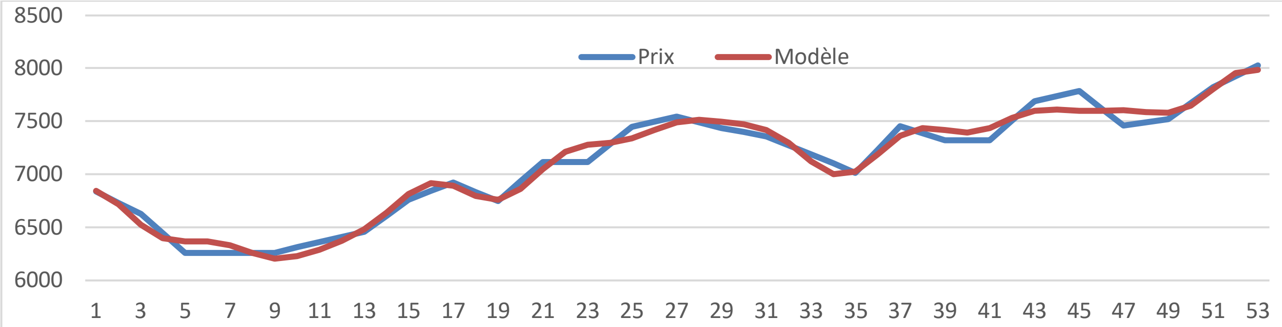

Figure 7 Dynamic filtering

Table 4 Normality tests of the dynamic filter

Estimated parameters		Kolmogorov-Smirnov		khi^2 Degree of Freedom=2	
Parameter	Value	D	0.0994	Observed value	0.5926
μ	0.2795	P-value	0.9410	Critical value	3.8415
Sigma	86.3681	Alpha	0.05	P-value	0.4414

Interpretations:

H_0: The sample follows a Normal law.
H_1: The sample does not follow a Normal law.

- *Kolmogorov-Smirnov:*

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected. The risk of rejecting the null hypothesis H_0 when it is true is 94.04%.

- *Khi²:*

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected. The risk of rejecting the null hypothesis H_0 when it is true is 44.14%.

- *White's test (Table 5)*

Table 5 White's Test

LM (Observed value)	3.3172
LM (Critical value)	5.9915
Degree of Freedom	2
P-value (bilateral)	0.1904
Alpha	0.05

H_0: The residues are homoscedastic
H_1: The residues are heteroscedastic

Given that the calculated p-value is greater than the threshold level of significance alpha = 0.05, we cannot reject the null hypothesis H_0. The risk of rejecting the null hypothesis H_0 when it is true is 19.04%.

We note that the dynamic filter makes it possible to normalize the residues with a very good power (94% for the Kolmogorov-Smirnov test), but there is a risk that these residues are heteroscedastic opposite to the static filter.

3.6 Comparison between the static filter and the dynamic filter

3.6.1 Oscillation study

Similarly, through Shannon's theorem and interpolation we obtain the result presented in figure 8. It is clear that the oscillations are much better attenuated by the dynamic filter which makes the curve much smoother.

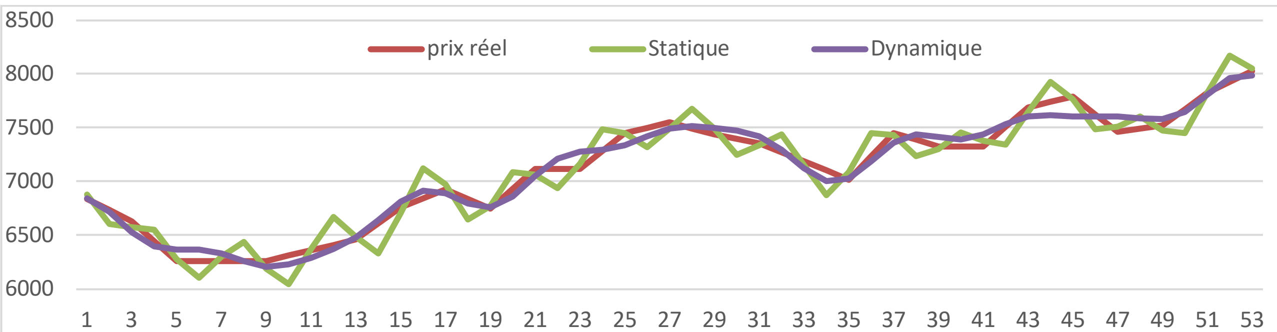

Figure 8 Oscillatory study of static and dynamic filtering

3.6.2 Forecast Horizon

Since both models are stationary around the trend, we can broaden the forecast horizon until a new cycle appears. The figure 9 shows the result of the two filters. The two filters give asymptotically equal series for the first two years then there is a shift! The new harmonic dynamic regression model is under the form of an optimization problem. The proposed new model is in (19).

$$\mathbf{Min}\ Y_t = a + bt + \sum_{i=1}^{\frac{p}{3}-1} A_i \cos(\beta_i \omega_i t + \varphi_i) + e_t e_t \sim N(0, \sigma^2 \quad (19)$$

$$(\textit{Under Constraints}) \begin{cases} \sum_{i=1}^{\frac{p}{3}-1} A_i \cos(\beta_i \omega_i t + \varphi_i) + \alpha \le \max(y_t) \\ \sum_{i=1}^{\frac{p}{3}-1} A_i \cos(\beta_i \omega_i t + \varphi_i) - \alpha \ge \min(y_t) \\ -1 \le \beta_i \le 1 \end{cases}$$

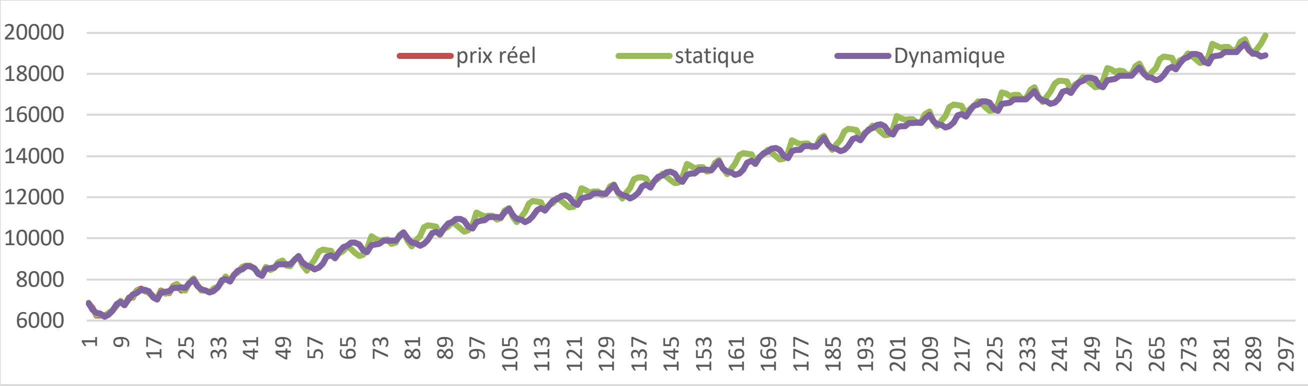

Figure 9 Forecast horizon for both filters

Since it is difficult to implement this model, we propose to work using the numerical model in (20).

$$\mathbf{Min}\ Y_t = a + b * t + \sum_{i=1}^{\frac{p}{3}-1} A_i * \cos(\beta_i * \omega_i * t + \varphi_i) + e_t e_t \sim N(0, \sigma^2) \quad (20)$$

$$(\textit{Under Constraints}) \begin{cases} \sum_{i=1}^{\frac{p}{3}-1} A_i + \alpha \le \max(y_t) \\ \sum_{i=1}^{\frac{p}{3}-1} A_i - \alpha \ge \min(y_t) \\ -1 \le \beta_i \le 1 \end{cases}$$

In this section, we have developed the basics of the new harmonic dynamic regression model, while performing various tests to validate the proposed model. This study has focused on the price of diesel fuel duty free and the conclusive results push us to test it on the price of SSP " Super Sans Plomb" all taxes included.

4 Case study

In this section, we will focus on in-depth SSP prices to determine the best possible DHR-based forecasting model. We will be constrained by the total permissible margin of the order of 3% which corresponds to our tolerance interval.

4.1 General model

In this first part, we will model, through the HDR, the prices of the last four years. The result found is presented in figure 10.

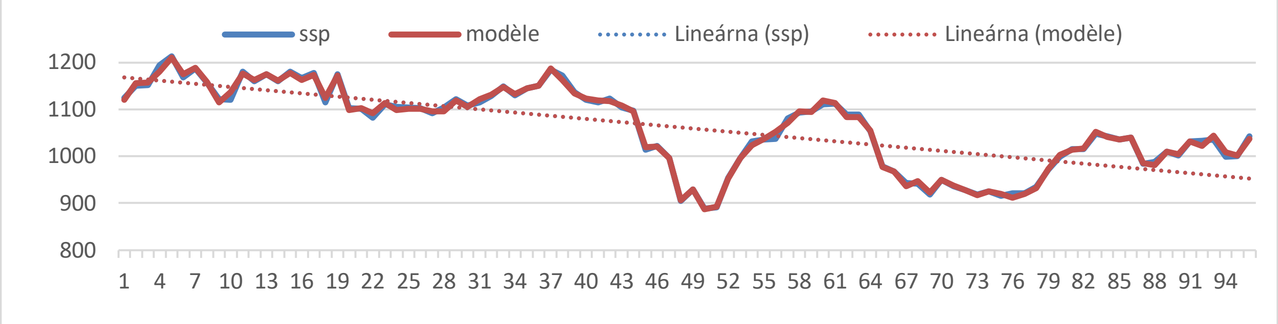

Figure 10 Model of the last 4 years

We will test the hypothesis of normality of errors through Khi^{2}and Kolmogorov-Smirnov tests:
H_0: The sample follows a Normal law.
H_1: The sample does not follow a Normal law.

Table 6 Normality test

Test		Kolmogorov-Smirnov		Khi^2 Degree of Freedom =2	
Parameter	Value	D	0.1104	Observed value	0.7108
μ	0.0012	P-value	0.7009	Critical value	5.9915
Sigma	0.4782	Alpha	0.05	P-value	0.1805

Table 7 White's Test

LM (Observed value)	2.2118
LM (Critical value)	5.9915
Degree of Freedom	2
P-value (bilateral)	0.3309
Alpha	0.05

Interpretations

- Kolmogorov-Smirnov test:

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected.
The risk of rejecting the null hypothesis H_0 when it is true is 70.09%.

- Khi^2 test:

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected.
The risk of rejecting the null hypothesis H_0 when it is true is 18.05%.

- Heteroskedasticity test:

H_0: The residues are homoscedastic
H_1: The residues are heteroscedastic

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected. The risk of rejecting the null hypothesis H_0 when it is true is 33.09%.
The next step will be to make price projections for the SSP based on this general model.

4.2 *Forecasts from the general model*

One-time forecasts

From the data of the last four years, we have been able to calculate the following RDH model shown in Figure 11.

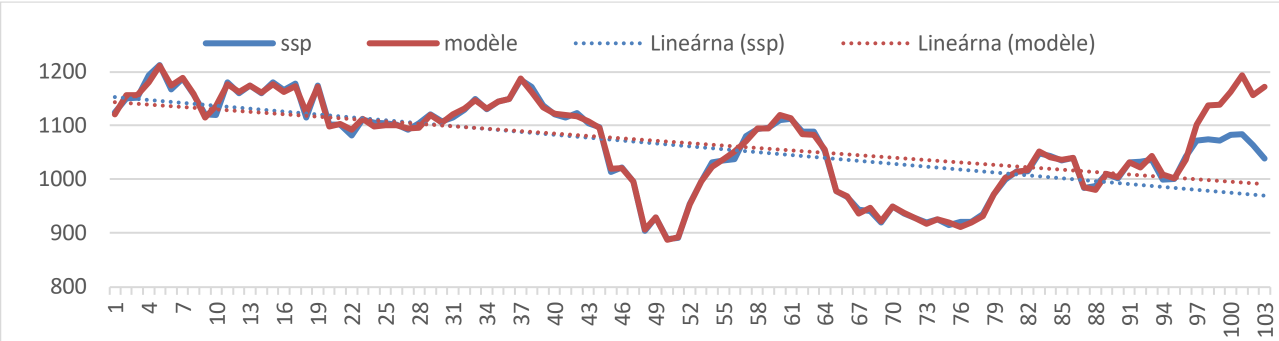

Figure 11 Predictions from the general model

Here are the results of the first quarter forecast summarized in Table 8.

Table 8 Forecast Results

Fortnight	Real Price	Model	% Error
1Q January	1072	1101.89	2.788246
2Q January	1074	1137.86	5.945996
1Q February	1072	1138.83	6.234142
2Q February	1082	1162.02	7.395564
1Q March	1084	1193.62	10.11255
2Q March	1064	1156.71	8.713346

The results obtained were not satisfactory compared to the tolerated margin of at least 3%. This is mainly due to data from the last three years before 2016 where the price was set by the state. Also, the compensation fund played a major role in covering the excessive increases in the price of a barrel on an international scale.

We therefore have only one year's data to make the predictions, i.e. 24 observations.

Forecast by confidence interval

Current forecasts are built for a 95% confidence interval.

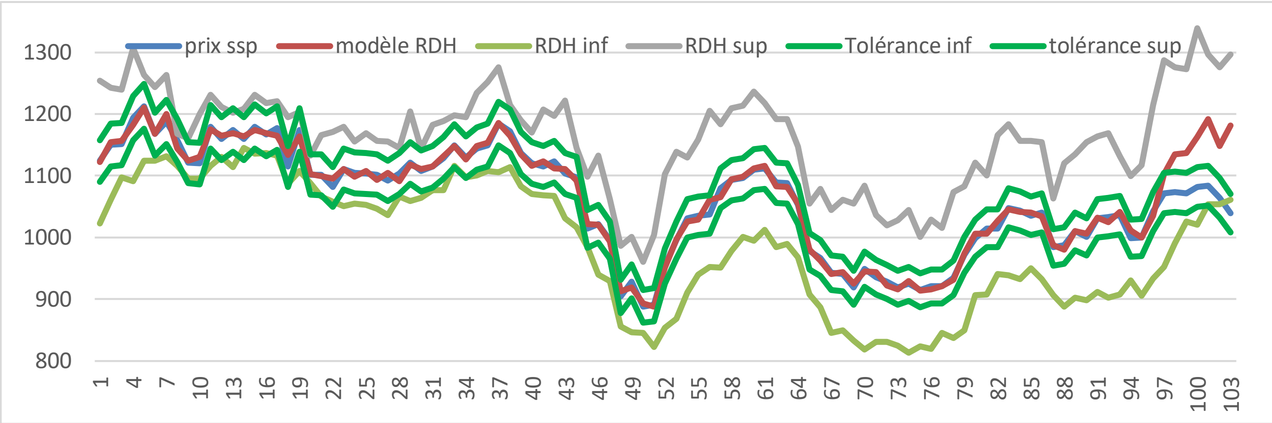

Figure 12 Confidence interval forecasts of the general model

4.3 Model reduced

To be able to work with data from a single year, we must choose between two approaches:

• Static filtering

• Dynamic filtering

We decided to compare between the two approaches.

4.3.1 Static filtering

Figure 13 reflects the results found.

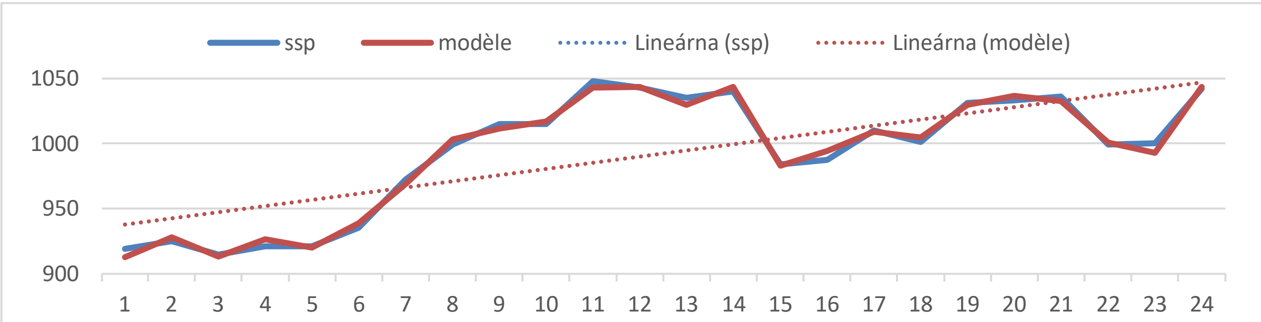

Figure 13 Static model

H_0: The sample follows a Normal law
H_1: The sample does not follow a Normal law

Table 9 Normality test results

Test		Kolmogorov-Smirnov		Khi^2 Degree of Freedom =2	
Parameter	Value	D	0.1251	Observed value	2.0978
μ	-0.0004	P-value	0.3503	Critical value	5.9915
Sigma	0.3962	Alpha	0.05	P-value	0.825

- Kolmogorov-Smirnov test:

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H0 cannot be rejected.
The risk of rejecting the null hypothesis H_0 when it is true is 82.5%.

- Khi^2 test:

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected.
The risk of rejecting the null hypothesis H_0 when it is true is 35.03%.

- Heteroskedasticity test:

Table 10 White's Test

LM (Observed value)	1.1608
LM (Critical value)	5.9915
Degree ofFreedom	2
P-value (bilateral)	0.5597
Alpha	0.05

H_0: The residues are homoscedastic
H_1: The residues are heteroscedastic

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected. The risk of rejecting the null hypothesis H_0 when it is true is 55.97%.

Predictions from the static model

• One-time forecasts

The forecast results for the first quarter are summarized in Table 11.

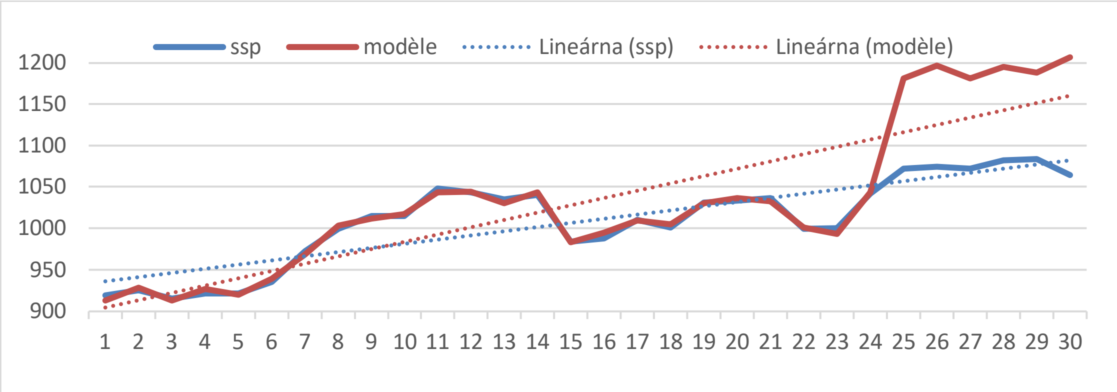

Figure 14 Static filter predictions

Table 11 Static filter prediction results

Fortnight	Real Price	Model	% Error
1Q January	1072	1180.96	10.16433323
2Q January	1074	1196.05	11.36442522
1Q February	1072	1180.91	10.16710017
2Q February	1082	1194.66	10.41225348
1Q March	1084	1187.98	9.592779328
2Q March	1064	1206.78	13.41920589

These results are not at all satisfactory and the percentage of error large enough is explained by the fact that the variables are overestimated and therefore there is a fairly large oscillation between two observations.

• Forecast by confidence interval

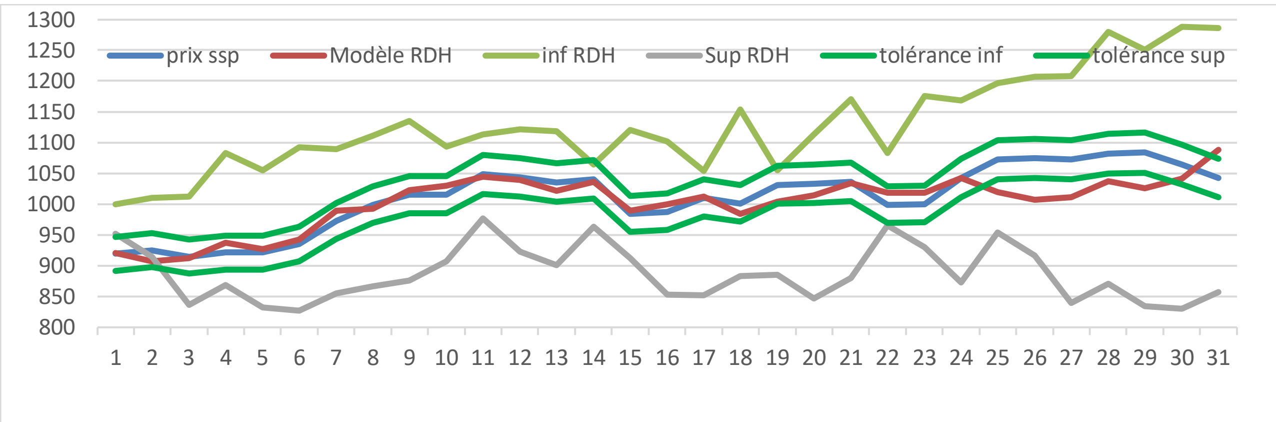

Figure 15 Forecast by confidence interval of the static model

4.3.2 Dynamic filtering

The found results are shown in Figure 16.

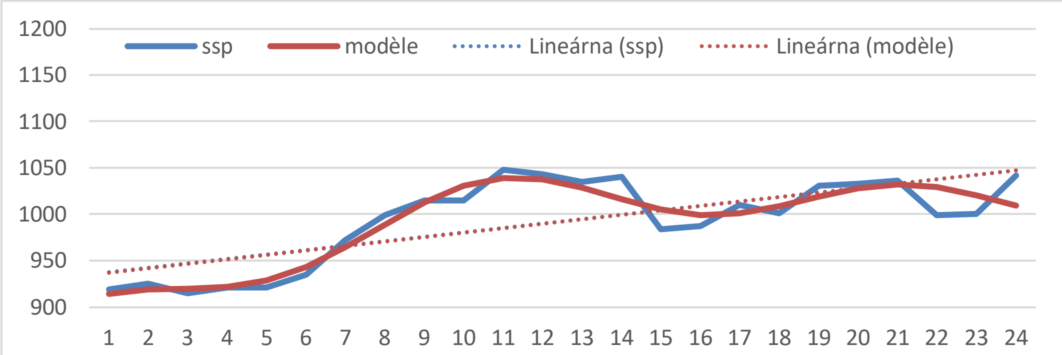

Figure 16 Dynamic Filter Results

H_0: The sample follows a Normal law
H_1: The sample does not follow a Normal law

Table 12 Normality test results

Test		Kolmogorov-Smirnov		Khi² Degree of Freedom =2	
Parameter	Value	D	0.1458	Observed value	2.7145
μ	-0.0189	P-value	0.2574	Critical value	5.9915
Sigma	1.4038	Alpha	0.05	P-value	0.6526

- Kolmogorov-Smirnov test:

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H0 cannot be rejected.
The risk of rejecting the null hypothesis H_0 when it is true is 65.26%.

- Khi² test:

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected.
The risk of rejecting the null hypothesis H_0 when it is true is 25.74%.

- Heteroskedasticity test:

Table 13 White's Test

LM (Observed value)	9.6522
LM (Critical value)	5.9915
Degree ofFreedom	2
P-value (bilateral)	0.0080
Alpha	0.05

H_0: The residues are homoscedastic
H_1: The residues are heteroscedastic
Interpretations:

Since the calculated p-value is less than the significance level alpha = 0.05, the null hypothesis H_0 must be rejected, and the alternative hypothesis H_1 must be accepted. The risk of rejecting the hypothesis H_1 when it is true is less than 0.80%.

Predictions from the dynamic model

• One-time forecasts

The results of the first quarter forecast are summarized in Table 14.

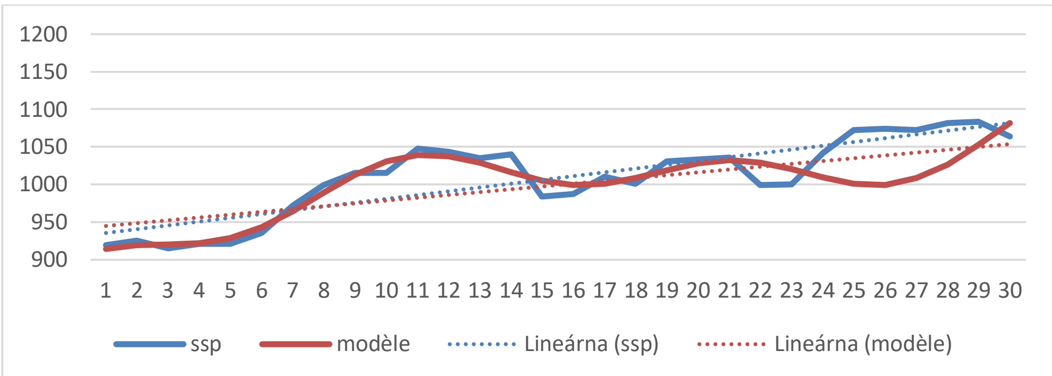

Figure 17 Dynamic Model Forecasts

Table 14 Dynamic filter prediction results

Fortnight	Real Price	Model	% Error
1Q January	1072	1000.899217	-6.632535701
2Q January	1074	999.5125295	-6.935518669
1Q February	1072	1008.11424	-5.959492561
2Q February	1082	1026.784783	-5.103069996
1Q March	1084	1052.821886	-2.876209805
2Q March	1064	1081.646185	1.658476047

These results are still far from the margin of 3% fixed. This leads us to review our method to make changes.

• Forecast by confidence interval

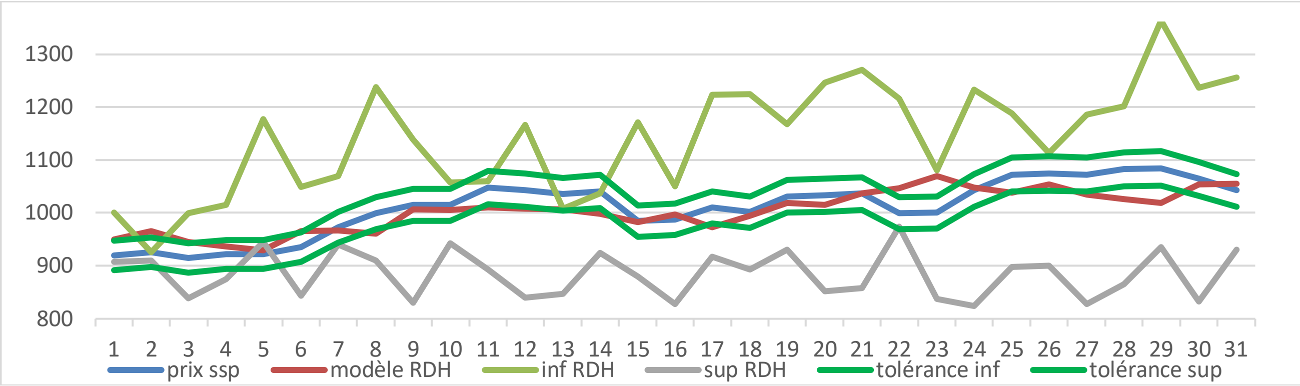

Figure 18 Confidence interval forecasts for the dynamic model

4.4 *Process approach*

We have seen in the previous section the importance of a good data analysis. In addition, we found that the static and dynamic models do not give the same results. Also, we notice that the assumed linear trend of the model no longer follows the real regression line of the data. On these three points in addition to the fixed margin of 3%, we will try to build a process approach for DHR.

4.4.1 Data analysis

In the previous section, we assumed that the company adapted to the withdrawal of the compensation fund from the outset, but it took a gradual adaptation time of about 6 months. So we can only work with the data of the last 6 months or 12 observations. We will work with the static model.

4.4.2 Local Dynamic Harmonic Regression Model

Figure 19 shows the results found.

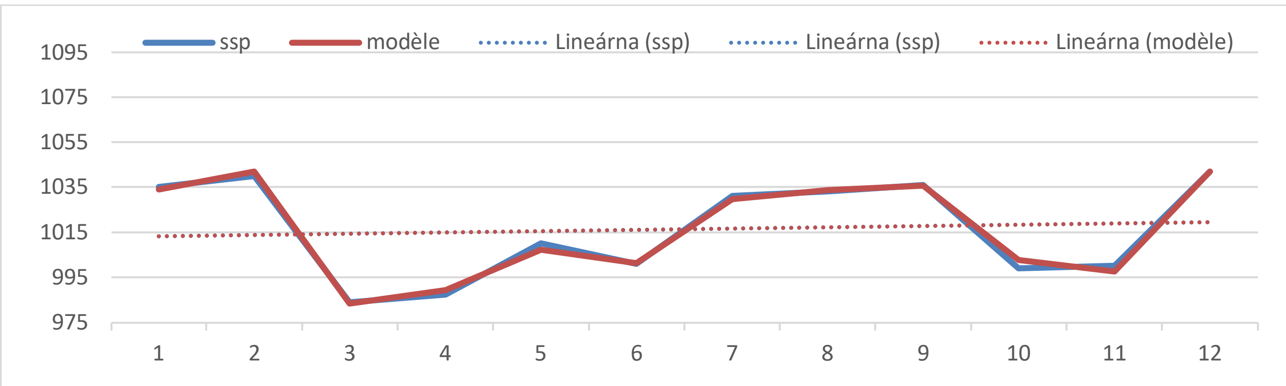

Figure 19 Local RDH Model

• Kolmogorov-Smirnov test:

H_0: The sample follows a Normal law

H_1: The sample does not follow a Normal law

Table 15 Normality test results

Test		Kolmogorov-Smirnov	
Parameter	Value	D	0.1288
µ	-0.0005	P-value	0.9818
Sigma	0.1812	Alpha	0.05

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H0 cannot be rejected.

The risk of rejecting the null hypothesis H_0 when it is true is 98.18%.

• Heteroskedasticity: White test:
H_0: The residues are homoscedastic
H_1: The residues are heteroscedastic

Table 16 White's Test

LM (Observed value)	0.6213
LM (Critical value)	5.9915
Degree of Freedom	2
P-value (bilateral)	0.7330
Alpha	0.05

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected.

The risk of rejecting the null hypothesis H_0 when it is true is 73.30%. We can move to the forecasts now.

Table 17 Local Model Forecast Results

Fortnight	Real Price	Model	% Error
1Q January	1072	1000.899217	0.717024748
2Q January	1074	999.5125295	1.251208536
1Q February	1072	1008.11424	-4.004232553
2Q February	1082	1026.784783	-4.363472259
1Q March	1084	1052.821886	-2.873636521
2Q March	1064	1081.646185	-1.617299405

The new process approach is to adjust the model as forecasts are made by adjusting the parameters a and b. Indeed, we will readjust the trend of the model to each forecast so that it is as close as possible to the regression line of the quantitative variable. At each step, we must ensure that the predictions remain within the pre-defined tolerance interval (plus at least 3% in our case); otherwise we will have to recalculate the updated RDH model. If no update has been made after one-third of the observations (two months in our case), update the data.

Here is an explanatory diagram for this new approach shown in Figure 20.

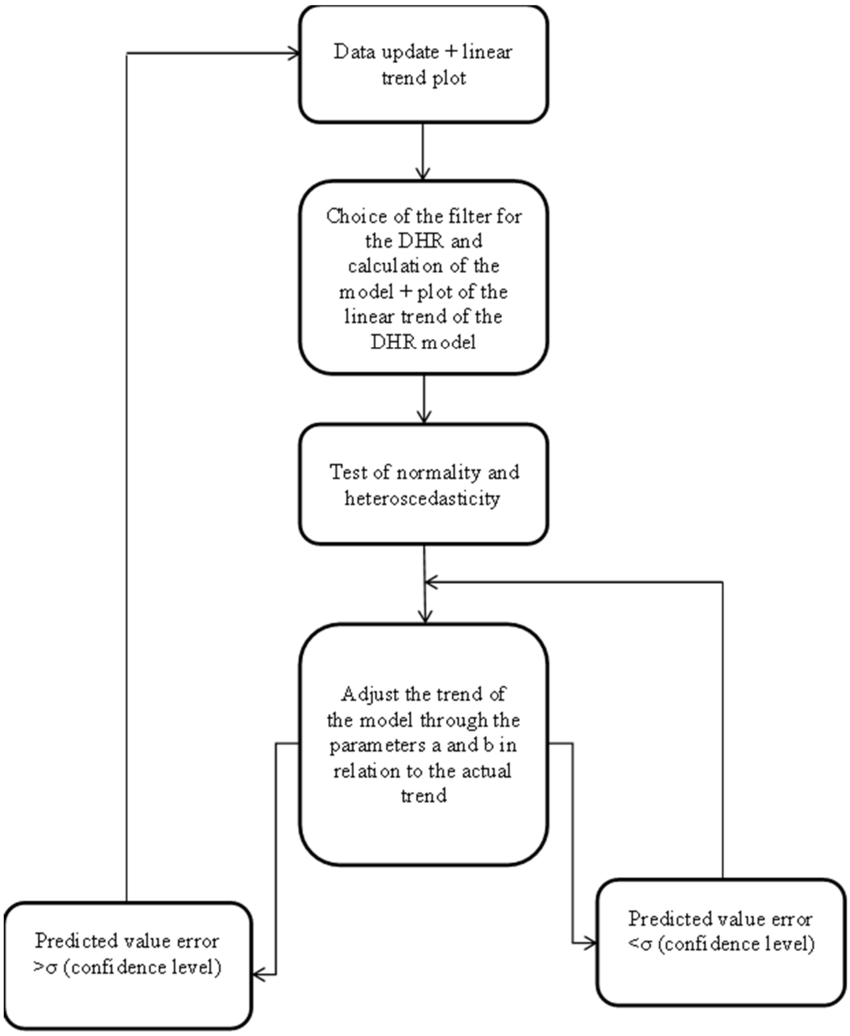

Figure 20 Process approach

4.5 Results of the process approach

In Table 18, we will list the values of the parameters a and b after each forecast. At t = 0, we modeled our data and did not make any predictions until then.

Table 18 First evolution of parameters a and b

t	**0**	**1**	**2**
a	991.813	993	995
b	3.802	3.55	3.2

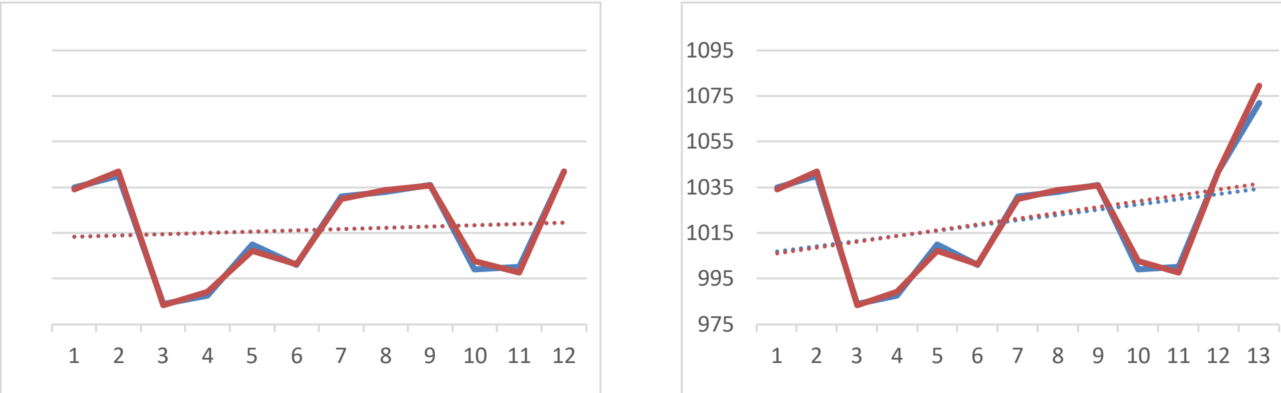

Figure 21 Prediction Using the Initial Model

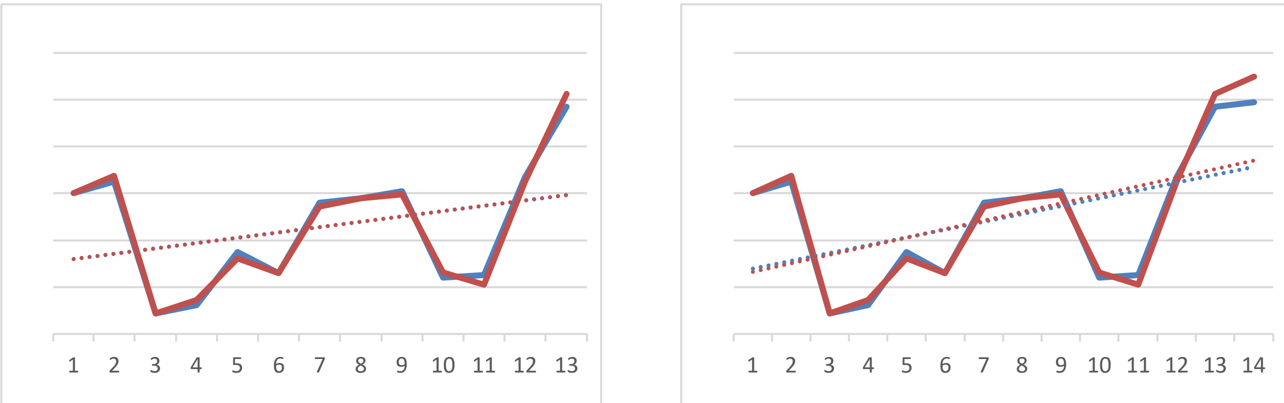

Figure 22 Forecast Using the Adjusted Model

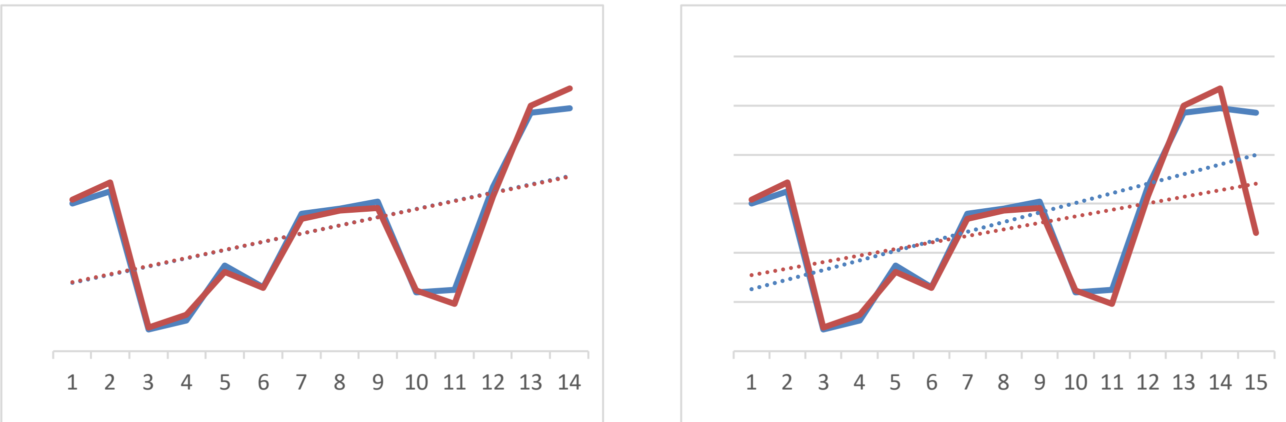

Figure 23 Forecast Using the Adjusted Model 2

Table 19 Forecast results

Fortnight	Real Price	Model	% Error
1Q January	1072	1079.686505	0.717024748
2Q January	1074	1085.09698	1.033238331
1Q February	1072	1023.231627	-4.549288523

It is now necessary to recalculate the model before making the forecasts of the next fortnights. To do this, we will integrate the last three measured values and therefore we will have 15 observations.

Here is the new DHR model shown in Figure 24.

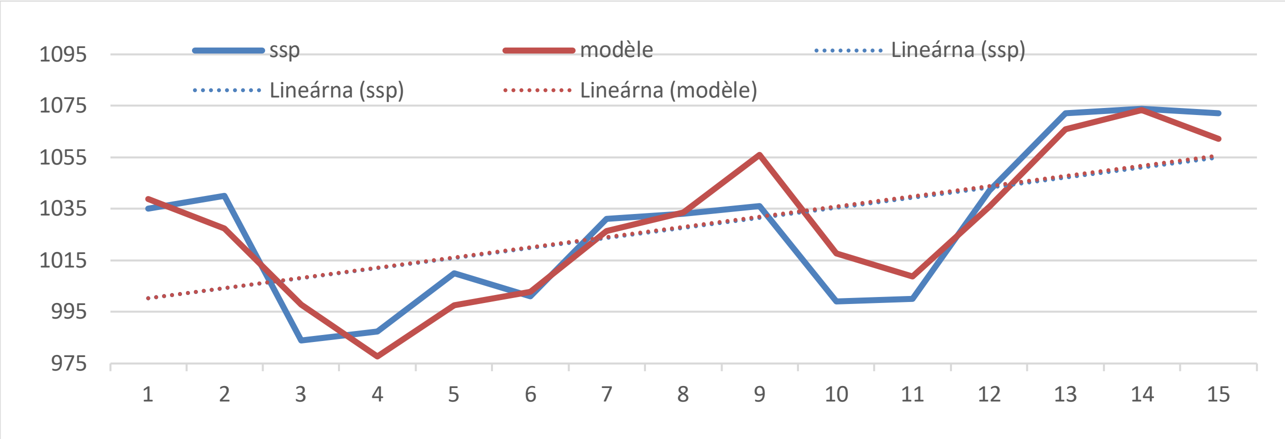

Figure 24 Updated DHR model

• Kolmogorov-Smirnov test:
H_0: The sample follows a Normal law
H_1: The sample does not follow a Normal law

Table 20 Normality test results

Test		Kolmogorov-Smirnov	
Parameter	Value	D	0.1420
μ	0.0428	P-value	0.9002
Sigma	1.0609	Alpha	0.05

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H0 cannot be rejected.

The risk of rejecting the null hypothesis H_0 when it is true is 90.02%.

• Heteroskedasticity: White test:
H_0: The residues are homoscedastic
H_1: The residues are heteroscedastic

Table 21 White's Test

LM (Observed value)	1.1290
LM (Critical value)	5.9915
Degree of Freedom	2
P-value (bilateral)	0.5686
Alpha	0.05

Since the calculated p-value is greater than the alpha threshold significance level = 0.05, the null hypothesis H_0 cannot be rejected. The risk of rejecting the null hypothesis H_0 when it is true is 56.86%.

The table 22 summarizes the evolution of the parameters a and b.

Table 22: Second evolution of parameters a and b

t	**0**	**1**	**2**
a	1005.622	1005.622	1004
b	2.809	2.809	3.08

Forecast results

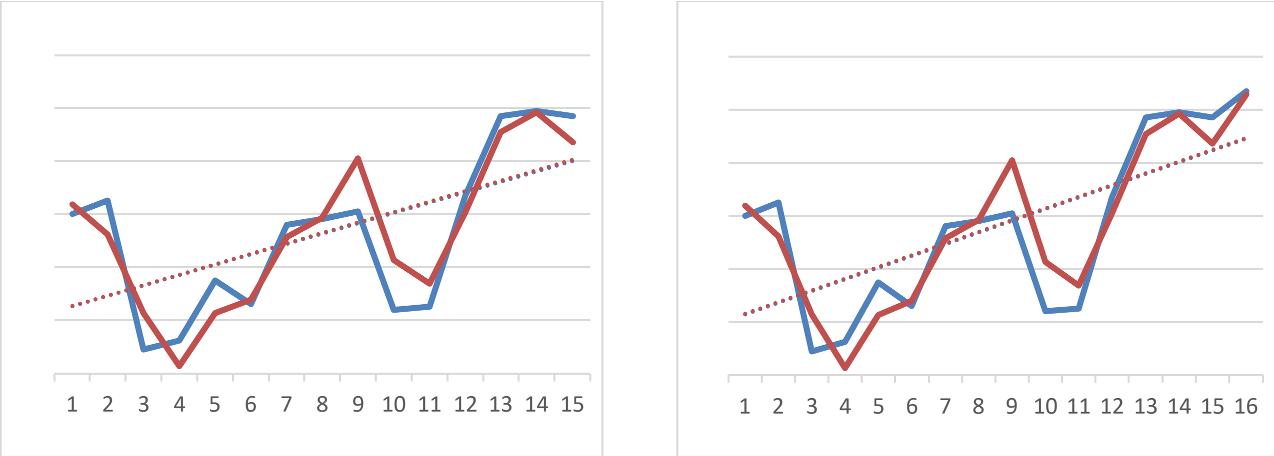

Figure 25 Prediction Using the Initial Updated Model

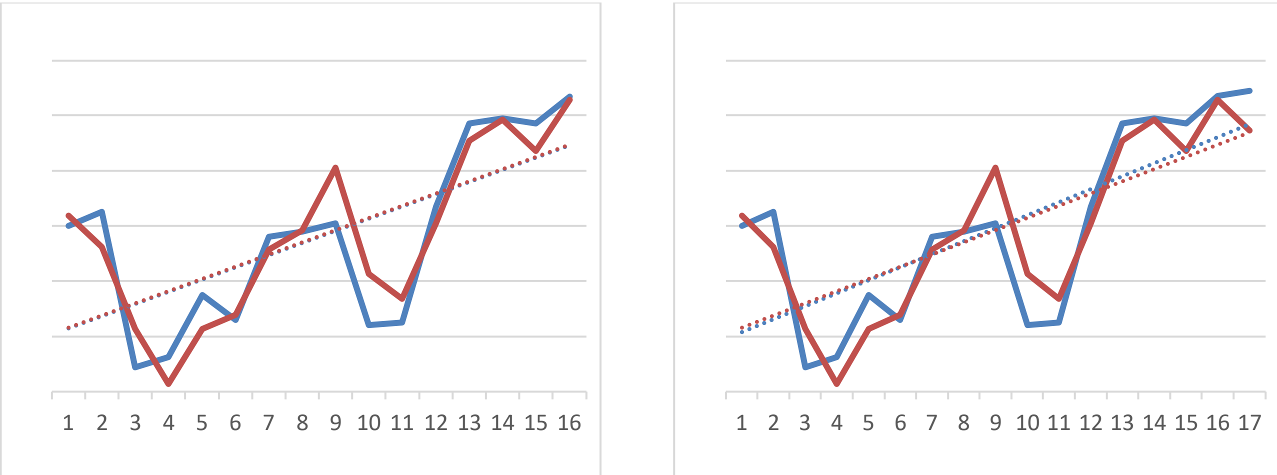

Figure 26 Forecast Using the Updated Updated Model 1

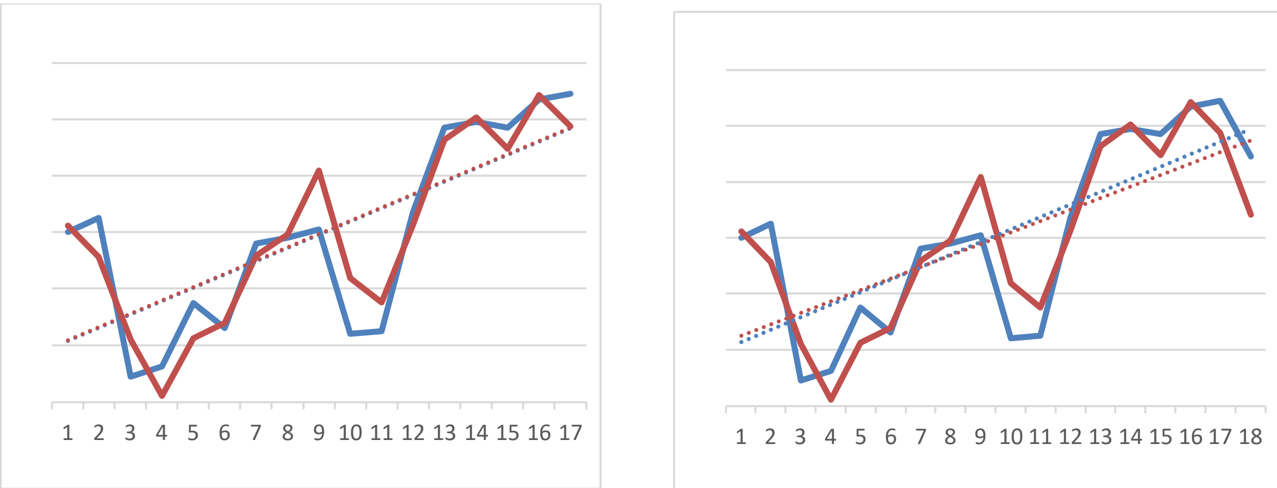

Figure 27 Forecast Using the Updated Updated Model

Table 23 Forecast results

Fortnight	Real Price	Model	% Error
2Q February	1072	1080.850305	-0.106256505
1Q March	1074	1072.436995	-1.342066846
2Q March	1072	1043.173202	-1.957405788

We have seen in this section that the DHR model does not only require the choice between the dynamic filter and the static filter, it must also make a rigorous analysis of the data and ensure that the objectives are achieved. This allowed us to put in place a process approach that proved to be very relevant for price analysis and forecasting of the SSP and thus, enable managers to take strategic decisions based on these accurate forecasts.

5 Conclusions

In this context, our work mainly aimed at studying the time series of diesel and SSP fuel in order to provide accurate forecasts and to respect the permissible margin of error of 3% setted by the company. For this purpose, we developed a new numerical method of harmonic dynamic regression. This new harmonic dynamic regression model through the proposed process approach yielded excellent forecast results for the first quarter of 2017 with an average margin error of 1.617%. We can therefore retain our model to make future forecasts, but each time we must feed our database to improve the results obtained. In this, the predicted performance can be evaluated, the best methodology and approach can be selected, and projections can be made. The increased predictions thus made will allow the managers to manage the business well in order to increase the income.

References

[1] ZHAO, X., XIE, J., LAU, R.S.M.: Improving the supply chain performance: use of forecasting models versus early order commitments, *International Journal of Production Research*, Vol. 39, No. 17, pp. 3923-3939, 2001. https://doi.org/10.1080/00207540110072236

[2] YU, L., ZHAO, Y., TANG, L.: A compressed sensing-based AI learning paradigm for crude oil price forecasting, *Energy Economics*, Vol. 46, pp. 236-245, 2014.

[3] YU, L., XU, H., TANG, L.: LSSVR ensemble learning with uncertain parameters for crude oil price forecasting, *Applied Soft Computing*, Vol. 56, pp. 692-701, 2017.

[4] NASER, H.: Estimating and forecasting the real prices of crude oil: A data rich model using a dynamic model averaging (DMA) approach, *Energy Economics*, Vol. 56, pp. 75-87, 2016.

[5] KNETSCH, T.A.: Forecasting the price of crude oil via convenience yield predictions, *Journal of Forecasting*, Vol. 26, pp. 527-549, 2007.

[6] GHOSH, S.: Import demand of crude oil and economic growth: evidence from India, *Energy Policy*, Vol. 37, No. 2, pp. 699-702, 2009.

[7] AKARCA, A.T., ANDRIANACOS D.: Detecting break in oil price series using the Box–Tiao method, *International Advances in Economic Research*, Vol. 3, No. 2, pp. 217-224, 1997.

[8] LANZA, A., MANERA, M., GIOVANNINI, M.: Modeling and forecasting cointegrated relationships among heavy oil and product prices, *Energy Economics*, Vol. 27, No. 6, pp. 831-848. 2005.

[9] MIRMIRANI, S., LI H.C.: A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil, *Applications of Artificial Intelligence in Finance and Economics*, Vol. 19, pp. 203-223, 2004.

[10] CHURCH, K.B., CURRAM, S.P.: Forecasting Consumers' Expenditure: A Comparison between Econometric and Neural Network Models, *International Journal of Forecasting*, Vol. 12, No. 2, pp. 255-267, 1996.

[11] PRYBUTOK, V.R., YI, J., MITCHELL, D.: Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations, *European Journal of Operational Research*, Vol. 122, No. 1, pp. 31-40, 2000.

[12] GUTIÉRREZ-ESTRADA, J.C., DE PEDRO-SANZ, E., LÓPEZ-LUQUE, R., PULIDO-CALVO, I.: Comparison between traditional methods and artificial neural networks for ammonia concentration forecasting in an eel (Anguilla Anguilla L.) intensive rearing system, *Aquacultural Engineering*, Vol. 31, No. 3-4, pp. 183-203, 2004.

[13] HO, D.C., AU, K.F., NEWTON, E.: Empirical research on supply chain management: a critical review and recommendations, *International Journal of Production Research*, Vol. 40, No. 17, pp. 4415-4430, 2002.

[14] CRYER, J.D.: *Analyse des séries chronologiques*, Boston, Duxbury Press, 1986.

[15] CHATFIELD, C.: Avoiding statistical pitfalls, *Statistical Science*, Vol. 6, No. 3, pp. 240-252, 1991.

[16] TSITSIKA, A., CRITSELIS, E., KORMAS, G., FILIPPOPOULOU, A., TOUNISSIDOU, D., FRESKOU, A., SPILIOPOULOU, T., LOUIZOU, A., KONSTANTOULAKI, E., KAFETZIS, D.: Internet use and misuse: a multivariate regression analysis of the predictive factors of internet use among Greek adolescents, *European Journal of Pediatrics*, Vol. 168, No. 6, pp. 655-665, 2009.

[17] EL BAHI, Y.F., EZZINE, L., AMAN, Z., EL MOUSSAMI, H.: *Modeling and Forecasting of Fuel Selling Price Using Time Series Approach: Case Study*, Proceedings of the International Conference on Control, Decision and Information Technologies, Thessaloniki, Greece, pp. 283-288, 2018. https://doi.org/10.1109/CoDIT.2018.8394835

[18] YOUNG, P., PEDREGAL, D., TYCH, W.: Dynamic harmonic regression, *Journal of Forecasting*, Vol. 18, pp. 369-394, 1999.

[19] TRAPERO, J.R., PEDREGAL, D.J.: Frequency domain methods applied to forecasting electricity markets, *Energy Economics*, Vol. 31, No. 5, pp. 727-735, 2009.

[20] PEDREGAL, D.J., TRAPERO, J.R.: Electricity prices forecasting by automatic dynamic harmonic regression models, *Energy Conversion and Management*, Vol. 48, pp. 1710-1719, 2007.

[21] SAPORTA, G.: *Probabilités, analyse des données et statistique*, Paris, Editions Technip, 2006.

[22] BREUSCH, T.S., PAGAN, A.R.: A simple test for heteroscedasticity and random coefficient variation, *Econometrica*, Vol. 47, No. 5, pp. 1287-1294, 1979.

Review process

Single-blind peer review process.