Lancaster University Management School
Working Paper
2009/022

Non-Linear Identification of Judgmental Forecasts Effects at SKU-Level

Juan Trapero Arenas, Robert Fildes and A Davydenko

The Department of Management Science
Lancaster University Management School
Lancaster LA1 4YX
UK

Non-Linear Identification of Judgmental Forecasts Effects at SKU-Level

Juan R. Trapero *, R. Fildes, A. Davydenko

Lancaster University
Department of Management Science, LA1 4YX, UK

Abstract

Prediction of demand is a key component within supply chain management. Improved accuracy in forecasts affects directly all levels of the supply chain, reducing stock costs and increasing customer satisfaction. In many application areas, demand prediction relies on statistical software which provides an initial forecast subsequently modified by the expert's judgment. This paper outlines a new methodology based on State Dependent Parameter (SDP) estimation techniques to identify the non-linear behaviour of such managerial adjustments. This non-parametric SDP estimate is used as a guideline to propose a non-linear model that corrects the bias introduced by the managerial adjustments. One-step-ahead forecasts of SKU sales sampled monthly from a manufacturing company are utilized to test the proposed methodology. The results indicate that adjustments introduce a non-linear pattern undermining accuracy. This understanding can be used to enhance the design of the Forecasting Support System in order to help forecasters towards more efficient judgmental adjustments.

Key words: Forecast adjustment, Supply chain, Non-linear system identification

1 Introduction

Companies working within supply chains use forecasts of demand to drive purchasing and supply chain management. Accurate forecasts can affect positively

* Corresponding author.
Email addresses: j.traperoarenas@lancaster.ac.uk (Juan R. Trapero), R.Fildes@lancaster.ac.uk (R. Fildes), a.davydenko@lancaster.ac.uk (A. Davydenko).

to the operational management of companies leading to significant monetary savings, greater competitiveness, enhanced channel relationships and customer satisfaction, lower inventory investment, reduced product obsolescence, improve distribution operations, schedule more efficient production and distribution, and more profitable financial decisions (Moon et al., 2003). For most of these companies, a particular type of a Decision Support System, known as a Forecasting Support System (FSS) is employed to prepare the forecasts (Fildes et al., 2006). These FSS integrate a statistical forecasting approach with managerial judgment from forecasters in the organization.

The manager's judgment is an important element within the forecasting process. For instance, judgment influences a wide range of decisions which range from the selection of a more appropriate statistical method to direct modification of the quantity being forecast. In fact, managers by adjusting the statistical system forecast often add information to the final forecast which may be difficult to include in a statistical model such a future product promotion (Fildes et al., 2006).

Despite the importance of judgment, for example around 89.5% of model forecasts were adjusted in the case study presented by Franses and Legerstee (2009), the literature devoted to study its effect with regards to company forecasts of the demand for SKUs is scarce with only three data sets covering 6 companies Mathews and Diamantopoulos (1990), Fildes et al. (2009), Syntetos et al. (2009), Franses and Legerstee (2009).

The recent literature suggests the existence of a bias towards making overly positive adjustments (Fildes et al., 2009) or as a consequence of a non-symmetric loss function of the managers (Franses and Legerstee, 2009). Mello (2009) analyzes the biases introduced by means of forecast game playing, defined as the intentional manipulation of forecasting processes to gain personal, group, or corporate advantage. Eroglu and Croxton (2009) explore the effects of particular individual differences and suggest that a forecaster's personality and motivational orientation influence significantly the forecasting biases. Since the companies in the supply chain are interdependent, the bias introduced into sales forecasts by one company affects the rest of companies along the chain. Therefore, the reduction of biases in sales forecasts is of paramount importance.

In order to correct the presence of the bias several works have modeled the appropriate weight that statistical forecasting and judgmental forecasting should have. For instance, Blattberg and Hoch (1990) took the mean of each approach which proved effective. Fildes et al. (2009) propose an Optimal Adjust model based on Linear Regression classifying the data depending on the adjustment sign. In contrast to Blattberg and Hoch (1990), it was found that negative adjustments were more precise than positive ones. This discontinuity between positive and negative adjustments may indicate the desirability of adopting non linear models to describe the judgmental process, knowing the form of the adjustment process has the potential to influence the design of the forecasting support system in order to mitigate the worst effects of such biases.

The present work reports the non-linear effect of adjustments on the final forecast accuracy on the basis of a manufacturing company database containing one-step-ahead forecasts and the actual sales. Assuming the expert adjustment is predictable (Franses and Legerstee, 2009) or fixed this non-linear identification is employed to propose a model which can correct the aforementioned bias and improves overall forecasting accuracy.

A State Dependent Estimation (SDP) approach is used to study the non-linearities involved in the manager's adjustment. SDP nonlinear estimation belongs to a family of methods within the data-based mechanistic modelling (DBM) developed by Young and co-workers, see Young et al. (2001), Young (2006), Young and Garnier (2006) and references therein among others. The SDP technique uses recursive methods like Fixed Interval Smoothing (FIS) combined with special data re-ordering and "backfitting" procedures which shows in a non-parametric way, i.e. through a graph, the state dependency between the parameter under study and an associated state variable (Young et al., 2001).

The outline of the paper is the following: Section 2 describes the problem formulation; Section 3 explains the SDP approach; Section 4 analyzes a case study to verify the model proposed and finally, Section 5 reports the main conclusions and their implications for achieving improvements in practice.

2 Problem formulation

The Optimal Adjust model, proposed by Fildes et al. (2009) aimed at optimally combining two of the sources of information available to the forecaster, the system forecast and the forecaster's subjective adjustment in order to deliver a more accurate forecast. It is given by:

$$y_{i,t} = \alpha_1 SF_{i,t} + \alpha_2 Adj_{i,t} + \nu_{i,t} \tag{1}$$

where $y_{i,t}$ is the actual value for the i^{th} product of the analyzed company at time t. The regressors are $SF_{i,t}$ and $Adj_{i,t}$ which stand for the System Forecast (Statistical Forecast) and the Adjustment Forecast one step ahead at time t, respectively. The Adjustment Forecast variable is computed as:

$$Adj_{i,t} = FF_{i,t} - SF_{i,t} \tag{2}$$

where $FF_{i,t}$ is the Final Forecast employed by the FSS. The error term is $\nu_{i,t}$.

In order to assess the influence of the judgmental adjustment on the accuracy of forecasts we propose a more flexible version of the Optimal Adjust model. Fildes et al. (2009) provided statistical tests which indicated that coefficients α_1 and α_2 are different depending on the adjustment sign. Also they claimed that, according to the data extracted from the companies analyzed, negative adjustments tended to improve the forecast accuracy and the size of the adjustment affected the accuracy. In turn, positive adjustments tended to decrease the forecast accuracy. In order to correct the aforementioned bias a non-linear model is proposed:

$$y_{i,t} = \alpha_1(v_1(i,t))SF_{i,t} + \alpha_2(v_2(i,t))Adj_{i,t} + \nu_{i,t} \tag{3}$$

The aim is to determine the potential states $v_1(i,t)$ and $v_2(i,t)$, as well as, to estimate the unknown SDP $\alpha_1(v_1(i,t))$ and $\alpha_2(v_2(i,t))$ which may offer a better explanation of the non-linear process described by $y_{i,t}$. By understanding how the company's forecasters misweight the information available it may be possible to develop a FSS that overcomes some of the worst excesses (Fildes et al., 2009)

3 A State Dependent Parameter Estimation approach

Following the works of Fildes et al. (2009) and Syntetos et al. (2009) we have grouped all the observations as cross-sectional data, dealing with each observation as an individual case. Since the variance of each SKU can be different, data normalization is also required. For instance, it is possible to normalize with respect to the standard deviation of each SKU as proposed by Fildes et al. (2009). Furthermore, since the parameters are expected to vary depending on the adjustment size, the data is sorted with respect to the adjustments. Accordingly, the data can be reindexed by $k = 1, \ldots, N$, where N is the sample size. In this sense equation (3) is rewritten as:

$$y_k = \alpha_1(v_1(k))SF_k + \alpha_2(v_2(k))Adj_k + \nu_k \tag{4}$$

The State Dependent Parameters are expressed by $\alpha_1(v_1(k))$ and $\alpha_2(v_2(k))$, where $v_i(k),\ i = 1,2$ is the variable which drives the behaviour of the aforementioned SDP. The random noise ν_k is assumed Gaussian with zero mean and variance σ^2.

In order to determine $\alpha_1(v_1(k))$ and $\alpha_2(v_2(k))$ described in (4) several assumptions have been made to capture the adjustment process. Firstly, it is assumed that $v_1(k)$ remains constant. In other words, as a company usually needs to predict the demand of a vast number of products, an "automatic" forecasting technique[1] implemented in a Forecasting Support System is used for this purpose, providing the first regressor (SF_k) in (1). Therefore, it is expected the weight of SF_k is approximately the same for the wide range of products yielding a constant α_1. In the second place, since the effectiveness of the adjustments may differ depending on the adjustment sign, it is assumed the parameter α_2 does not remain constant. Indeed, we assume that α_2 is function of the adjustments represented by Adj_k.

The SDP modeling procedure allow us to incorporate this form of non-linearity. Therefore, taking into account the previous assumptions the SDP-Optimal Adjust model is proposed as an extension of the Optimal Adjust model described in (1), such as:

$$y_k = \alpha_1 SF_k + \alpha_2(Adj_k)Adj_k + \nu_k \tag{5}$$

Note that it is possible to formulate more complicated models based on the SDP procedure, for example, we can assume α_1 is also state dependent. Nevertheless, we prefer expression (5) because even when it is a non-linear model each term can be easily interpreted following the DBM philosophy (Young, 2006).

In order to let the parameter α_2 vary with adjustment, a first approach would be to define stochastically the parameter α_2 as a two-dimensional stochastic state vector, whose stochastic properties are defined by a Generalized Random Walk (Jakeman and Young, 1984). There is a wide range of options (Pedregal and Young, 2002). Generally, the stochastic state vector is also called Time Varying Parameter (TVP) because the data is ordered in a temporal fashion associated to time series problems. However, we are interested in looking for the variations of α_2 with respect to the adjustments instead of time. Since we are not working on-line, the data can be sorted with respect to the adjustments and run the TVP procedure. After that, the data is unsorted to its original time order.

Fortunately, the SDP technique allow us to sort and unsort the data and run the TVP procedure. Additionally, the algorithm includes back-fitting procedures employing recursive FIS algorithms to achieve estimations of any state dependent parameter. The outcome of the SDP algorithm is a non-parametric estimate displayed as a graph of the state dependent parameter ($\alpha_2(Adj_k)$) against the variable which affects it in a non-linear fashion (Adj_k).

[1] For instance, an exponential smoothing method

The stochastic behaviour of α_2 chosen in this application is an *Integrated Random Walk* which consists of:

$$\begin{pmatrix} \alpha_2(k+1) \\ \alpha_2^*(k+1) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_2(k) \\ \alpha_2^*(k) \end{pmatrix} + \begin{pmatrix} 0 \\ w^*(k) \end{pmatrix} \tag{6}$$

where α_2 and α_2^* are associated to the changing level and slope of the state dependent parameter; the flexibility in the model is introduced by the random Gaussian noise $w^*(k)$ with mean zero and variance σ_α^2.

The full model is formulated as a State Space (SS) system by assembling the observation equation in (5) and the State equations in (6). The SS formulation is well-suited for optimal recursive estimation accomplished by well-known recursive algorithms as the Kalman Filter (KF) in Kalman (1960) and the Fixed Interval Smoothing (FIS) in Bryson and Ho (1969). However, in order to use these algorithms all the system matrices must be assumed known. In this case the unknown parameters (often called hyper-parameters to distinguish them from the main parameters or states in (6)) are the noise variances of the Observation equation (σ^2) and the State equations (σ_α^2).

Usually, the variances are normalized by the innovations variance (σ^2) reducing to one the number of unknown parameters. In this sense, the Noise Variance Ratio (NVR) is defined as σ_α^2/σ^2. The optimization of the NVR can be done by Maximum Likelihood (ML) in the time domain obtained via "prediction error decomposition", see Harvey (1989).

A complete description of the technique with numerous examples can be found in Young et al. (2001), and some applications to environmental systems are shown in Young and Garnier (2006) and Young (2006) among others. Additionally, SDP algorithms are available within the CAPTAIN[2] toolbox (Taylor et al., 2007), developed for using with Matlab/Simulink ^{TM} software.

4 Case study

Data from a manufacturing company specialized in household products has been collected. The data has been split in three series which represent: i) one-step-ahead systems forecasts; ii) one-step-ahead final forecasts; and iii) corresponding actual outcomes for 413 Stock Keeping Units (SKU). The data comprises 7544 completed triplets that have been sampled monthly between

[2] see http://www.es.lancs.ac.uk/cres/captain/

2004 and 2007. It is an extended data set from that analysed as company A in Fildes et al. (2009).

Basically, the final forecast produced by the company is the result of two sources of information (Fildes et al., 2009). On the one hand, there is available computer software which provides the statistical system forecasts. On the other hand, various meetings, which involve the forecasters meeting with personnel in sales, marketing, and production occur to share pieces of information that cannot be included in a statistical model. Thus, the previous system forecast is adjusted accordingly with the meeting group decisions obtaining an agreed forecast (final forecast). For instance, approximately 65% of the 7544 complete triples was modified by the demand forecaster's adjustments.

4.1 Data selection

Since the characteristics of the data are heterogeneous a pre-treatment step is required in order to generate a homogeneous sample. We agree with Syntetos et al. (2009) about the fact that data selection has been overlooked in the past. Basically, this step eliminates those time series that are not useful for the experimental analysis (depending the goals of the research). For instance, Fildes et al. (2009) removes SKUs without the required continuous forecast history or those with low volume SKUs because these are the result from special circumstances such as the particular items have been withdrawn from the market. In contrast, Syntetos et al. (2009) was focused on intermittent demand.

Our aim here is to develop a model of the adjustment process for established SKUs that captures any non-linear effects, so the pre-treatment stage in this work removes the time series which fulfils any of these conditions:

- Time series with less than 12 months history available.
- Time series with any Actual observation equal to zero.

After this pretreatment the number of SKU's is reduced to 91 with 2882 triplets. An example of the time series considered per each SKU can be seen in Figure 1, where actual values are in solid line, System Forecasts are in dotted line and Final Forecasts are in dashed line. Recall that $FF_k = SF_k + Adj_k$.

4.2 Exploratory Data Analysis

In order to take advantage of the judgemental adjustments we have to check that adjustments improve the forecasting accuracy provided by the Statistical

Forecast. Basically, if there is no evidence that the Final Forecasts based on adjustments beat the Statistical Forecast accuracy it would be quite unlikely to propose a model, which uses as a regressor those adjustments, capable of beating the Statistical Forecast. In this sense, Table 1 assesses the forecasting performance provided by the company under study, where cases with lower error are shown in bold. In this table the Mean Absolute Percentage Error (MAPE) and the Median Absolute Percentage Error (MdAPE) were chosen as accuracy measures, such as:

$$\begin{aligned} MAPE &= mean(|p_t|) \\ MdAPE &= median(|p_t|) \end{aligned} \tag{7}$$

where p_t is the percentage error given by $p_t = 100|Y_t - F_t|/Y_t$, $t = 1, \ldots, N$. In this expression Y_t stands for the actual value at time t and F_t is the forecast at that time. All forecasts considered are one-step-ahead and N is the sample size.

In particular, these error measures have been computed across time for each SKU, and then the mean of these values were calculated across SKUs. The last row in Table 1 shows the total Mean of the MAPE and MdAPE, where the values in bold highlight the best performance method. In general terms, the FF is more accurate than the SF. Additionally, we can breakdown the errors according to the adjustment sign obtaining three rows which analyze the forecasting performance of the positive and negative adjustments, as well as, when there is no adjustment. The row denoted by Overall adjusted comprises positive and negative adjustments excluding those observations that were not modified by judgmental adjustments. In addition, the second column also shows the sample size of each kind of adjustment, where it is possible to verify that positive adjustments are more frequent.

In relation to the negative adjustments, we can see from Table 1 that the FF is more accurate than the SF. Nevertheless, this same conclusion cannot be extrapolated to the positive adjustments case. In fact, there is no clear conclusion about whether the FF is more accurate than the SF because according to the Mean(MAPE) the SF outperforms the FF but, conversely, assessing the Mean(MdAPE) the FF beats the SF. A possible explanation to this discrepancy is that the Absolute Percentage Error puts a heavier penalty on positive errors than on negative ones (Hyndman and Koehler, 2006) and whilst the MdAPE is robust to this penalty, the MAPE is quite sensitive to it. In order to solve this discrepancy we can normalize the data and then compute the Mean Absolute Error (MAE=$mean(|Y_t - F_t|)$) which is not a percentage error measure. In the next section such data normalization is carried out, where it will be shown that not only is it valuable for error comparison purposes but also it is necessary to identify a non-linear pattern in the process of adjustments.

4.2.1 Data Normalization

One of the main objectives of this work is to find out if there is a bias in the adjustments accomplished by the forecasters. If so this information can be used to propose a model which improves the forecasting accuracy. Nonetheless, we are mixing different SKUs with different statistical properties. Thus, it is convenient to provide a framework where it is possible to compare them. This can be done by means of a data normalization. In particular, each product can be normalized with respect to its standard deviation (Fildes et al., 2009).

A statistical description of the normalized data can be found in Table 2. It is interesting to note that central measures like the mean and median corresponding to the Actual column are higher than their System Forecast counterpart. This difference between the real and SF has been detected and managerial adjustments were imposed to compensate this SF bias as can be seen in the statistics for the FF. However, this compensation were too optimistic achieving a FF mean and median higher than the actual ones. In the last two rows of this Table we can also find two dispersion measures, the standard deviation (Std.) and the Median Absolute Deviation (MAD). These measures show a similar dispersion between the SF, FF and Actual values.

Figure 2 depicts the boxplot of the actual values, as well as the System and Final Forecast provided by the company. Note that there are a higher number of extreme values in the Final Forecast which do not correspond to any actual value, see Figure 2. This implies that some large positive adjustments have been made incorrectly.

Figure 3 shows the histogram of the normalized adjustments, where it is possible to see that the adjustments are positively biased. In this sense, Table 3 analyzes the overoptimism in adjustments. For instance, the first row shows that positive adjustments tend to overestimate. In particular, 62% of them are positively biased because they were too large or they were in the wrong direction. In contrast, negative adjustments are less biased since only 51.7% of them were overoptimistic.

Previously, it was mentioned that the Mean(MAPE) and the Mean(MdAPE) might not be good error measures to compare the accuracy of the System and Final Forecast for positive adjustments because of the penalization of positive errors associated to those error measures. Therefore, we can take advantage of the suggested normalization to solve this problem by computing the MAE on this normalized dataset. In this way, Table 4 shows the MAE accomplished by the SF and FF. Unlike Table 1, Table 4 shows that the FF is more accurate than SF even for positive adjustments. The explanation of the difference between the results for the normalized versus the percentage error measures lies in the different weight given to longer errors associated with volatile SKUs.

Another advantage of the normalization is that we can analyze the evolution of the MAE with respect to the size of the adjustment. Indeed, Figures 4 and 5 depict the aforementioned evolution with regards to positive and negative adjustments, respectively. We can observe that: i) FF is more accurate than SF for larger adjustments; ii) there is no big differences between SF and FF methods for small adjustments; iii) the improvement achieved by the FF in negative adjustments is larger than for the positive counterparts. It should be pointed out that these findings agree with those suggested in Fildes et al. (2009).

4.3 Non-parametric SDP estimation

Under the assumption that some parameters may be state dependent, SDP algorithms have been employed to obtain non-parametric estimates. The model estimated assumed that the weight of the adjustments in (5) is potentially dependent of the Adjustments as a state. In order to perform this estimation an Integrated Random Walk described in (6) is used to model variations in the SDP, where the observations are ordered by adjustments size. As a result of this stage, a graph is provided giving us an indication of the possible non-linearity shape. This stage has been accomplished with the $MATLAB^{TM}$ toolbox called $CAPTAIN$ (Taylor et al., 2007).

Figure 6 depicts the estimation of $\alpha_2(Adj_k)$ as a solid line, and the standard errors of its estimation in dashed lines. According to this figure, there is a variation of the parameter α_2 depending on the adjustment sign. It is interesting to note that the weight of the adjustments represented by α_2^* is greater for negative adjustments. This means that positive adjustments tend to be optimistic and the SDP is tuned to damp this optimism. In other words, negative adjustments are more accurate than positive adjustments. Additionally, confidence intervals show that there is a big uncertainty for adjustments values close to zero. One explanation is that forecasters may make small adjustments when they mistake noise for patterns in the signal. Furthermore, confidence intervals are tighter for positive small adjustments, indicating that a large quantity of data is concentrated in this range, see Figure 3.

4.4 Identification and estimation of non-linearities regarding SDP

With the non-parametric estimate computed in the previous step, we now examine the graph obtained to propose a non-linear model capable of capturing the source of such non-linearities. Typically, this task can be done via non-linear parametric models which may range, for instance, from a radial basis function to a sigmoidal law, see (Young, 2006) and (Young et al., 2001). Then, the parametric model is efficiently estimated by nonlinear least squares, Prediction Error minimization or Maximum Likelihood optimization.

Considering Figure 6 the following non-linear model is proposed:

$$y_k = \beta_1 \cdot SF_k + \beta_2 \cdot Adj_k + (\beta_3 + \beta_4 \cdot e^{-\beta_5 \cdot Adj_k}) Adj_k \cdot X_d + \nu_k \tag{8}$$

where X_d is a dummy variable such as:

$$X_d = \begin{cases} 0 & \text{if} \quad Adj_k < 0 \\ 1 & \text{if} \quad Adj_k > 0 \end{cases} \tag{9}$$

The estimates of the model parameters are given below (the respective estimated standard deviation is between brackets) :

$$\begin{array}{ll} \hat{\beta}_1 = 0.94919 \quad (7.36 \cdot 10^{-5}) & \hat{\beta}_2 = 0.824 \quad (2.74 \cdot 10^{-3}) \\ \hat{\beta}_3 = -0.94 \quad (1.36 \cdot 10^{-2}) & \hat{\beta}_4 = 1.323 \quad (6.70 \cdot 10^{-3}) \\ \hat{\beta}_5 = 0.247 \quad (2.52 \cdot 10^{-3}) & \end{array}$$

Hereafter the non-linear expression in (8) is referenced as (NL).

4.5 Comparison with previous methodologies

Once we have estimated the SDP parameter in (5) and the non-linear function in (8), we will compare those results with two approaches. Firstly, we will use the Blattberg-Hoch (B-H) “50% model, 50 % manager” as a benchmark (Blattberg and Hoch, 1990), where:

$$\begin{aligned} y_k &= 0.5 \cdot SF_k + 0.5 \cdot (SF_k + Adj_k) + \nu_k \\ &= SF_k + 0.5 \cdot Adj_k + \nu_k \end{aligned} \tag{10}$$

We also analyze the Optimal Adjust (OA) model proposed by Fildes et al. (2009) that can be expressed using the dummy variable X_d defined in (9) such as:

$$y_k = \gamma_1 \cdot SF_k + \gamma_2 \cdot Adj_k + \gamma_3 \cdot SF_k \cdot X_d + \gamma_4 \cdot Adj_k \cdot X_d + \nu_k \tag{11}$$

The estimates of the model described in (11) are given below:

$$\hat{\gamma_1} = 0.96 \quad (0.014) \qquad\qquad \hat{\gamma_2} = 0.81 \quad (0.073)$$
$$\hat{\gamma_3} = 0.07 \quad (0.018) \qquad\qquad \hat{\gamma_4} = -0.42 \quad (0.076)$$

In order to compare models (8), (10) and (11) we related them to the general equation in (5). Assuming that the System Forecast weight of the aforementioned models is approximately 1, it is possible to plot in the same graph the Adjustments weight computed by the different models. For instance, α_1 and α_2 are defined as the SF and Adjustments weights, respectively in (5). The equivalent of α_1 in the non-linear model (8) is β_1 and the equivalent of $\alpha_2(Adj_k)$ is given by $\alpha_2(Adj_k) = \beta_2 + (\beta_3 + \beta_4 \cdot e^{-\beta_5 \cdot Adj_k})X_d$. Furthermore, we can see that $\hat{\beta}_1 = 0.949 \approx 1$.

Figure 7 depicts the estimation of the Adjustment weight accomplished by (8) in solid line; the non-parametric SDP estimation is depicted by a dashed line and the Optimal Adjust model in (11) is in dotted line. The dashdot line shows the Blattberg and Hoch model described in (10).

According to the non-parametric SDP estimation shown in Figure 7, the explanatory weight of the managerial adjustments depends on its sign. Basically, without the SDP guidance one might consider as a starting point the adjustments average (Blattberg and Hoch, 1990) to ponder the influence of adjustments on the forecasting accuracy. Nevertheless, the analysis carried out by Fildes et al. (2009) over different companies shows that adjustments accuracy was asymmetric with respect to its sign, i.e., negative adjustments was shown to be more precise than the positive ones. Effectively, the Optimal Adjust Model proposed in that reference is a better approximation to the non-linear nature of the adjustment process. Nonetheless, the Optimal Adjust Model only allows the variation of α_2 between constant values. This restriction is valid for negative adjustments (see Figure 7) but it is apparently not the best method to describe the positive adjustments in relation to the non-parametric SDP. In order to resolve this limitation, the non-linear function given by (8) is proposed, which models the negative adjustments with a constant (as the Optimal Adjust Model) but it uses an exponential function to describe the positive adjustments. Note this non-linear function is inspired by the non-parametric SDP estimate.

4.6 Model validation

In this section predictive validation is used to compare models, where we expect that if a better description of the adjustment process is offered by the non-linear model(s), these models should contribute to reduce the forecasting error compared to the simpler linear models. For this purpose, 20% of the data (582 triplets) constituted by the last months of each SKU, which were not used for the parameter estimation of the models, were employed as the hold-out sample to compare the performance of the proposed models. This hold-out sample design results in a more demanding experiment than selecting 20% of the data randomly (Fildes et al., 2009).

Table 5 shows the Mean(MAPE) and Mean(MdAPE) on the validation dataset. In the lower part of the table, we can see the overall performance of the methods analyzed. In order to get a deeper insight into the adjustment sign influence, the results have been separated according to the adjustment sign. Essentially, the Final Forecast beats the System Forecast (quite substantially for negative adjustments). The NL method outperforms the non-parametric "state dependent model" and in particular the linear models. The number of observations taken into account are shown in the second column. Again, the NL method delivers very promising results except for positive adjustments.

However, the Mean(MAPE) and Mean(MdAPE) are not well-suited to measure the positive adjustments performance for the reasons given in section 4.2. As previously, the normalized data was employed to compute the Mean Absolute Error (MAE) on the validation dataset, shown in Table 6. From this Table we can corroborate the good performance of the NL model for positive adjustments as well. Additionally, Table 7 shows the standard deviation of the Absolute Error in order to analyze the forecasting error dispersion. Assessing Tables 6 and 7 we can conclude that the non-linear model(s) proposed achieve a lower forecasting error and also reduces the variance of such errors compared to the Final Forecast.

Figures 8 and 9 depict the MAE against the size of positive and negative adjustments for the validation dataset, respectively. From these figures we can observe that the FF is more accurate than the SF. Furthermore, the margin of improvement is more visible for larger adjustments. Note that these findings are consistent with those reached in the exploratory data analysis assessing the SF and FF, in section 4.2

Regarding the proposed models for positive adjustments in Figure 8, the NL approach outperforms the rest of the models. In addition, the OA model works slightly worse than the FF and the B-H models. Since neither the OA model nor the B-H model are flexible enough to describe the non-linear process associated with positive adjustments, see Figure 7, the differences found between them may not be systematic. This means that this result may not be consistent for another sample.

In relation to the negative adjustments shown in Figure 9, the analyzed methods, except for the B-H technique, achieve a similar performance, where the

NL method outperforms them slightly. The B-H method performs rather worse than the rest of methods. This poor performance can be explained by analyzing negative adjustments in Figure 7. From this Figure we can see that the weight adjustment $\alpha_2(Adj_k)$ suggested by the B-H for negative adjustments is lower than the one computed by the other methods based on parameter estimation from the data. Therefore, this discrepancy results in a bigger forecasting error for the B-H method.

Considering the case data where no managerial adjustment is made the NL method also improves forecasting accuracy. This indicates that there is room to improve the SF design.

Finally, it is interesting to note that NL beats SDP but only slightly since NL is estimated in a more efficient way than the SDP.

5 Conclusions

The use of State Dependent Parameter estimation was exploited in a new application in order to understand the non-linear complexity involved in judgmental adjustments. These adjustments are of paramount importance in numerous companies since they have a direct and substantial influence on the forecasting accuracy of supply chain demand. Actual data sampled monthly were collected from a manufacturer company to verify the approach. In fact, an SDP estimate was the baseline to formulate a non-linear model which was employed to reduce the forecasting errors on the basis of a better description of the nonlinearity observed in managerial adjustments. In order to compare the performance of the methods considered, several well-known error measures were considered including the MAPE and MdAPE. Nonetheless, it was shown that normalization of the data can be very helpful to get a better understanding of the influence of the adjustment size. Basically, this normalization allow us to use another error measure (MAE) that avoid the heavy penalization which the percentage errors apply to positive errors. Therefore, this MAE gave a better description of the methods' relative performance when positive adjustments are considered.

Putting these together, several conclusions can be drawn: i) there were not big differences between the methods analyzed for small adjustments; ii) FF forecasts outperform the SF when adjustments are larger; iii) the NL model proposed on the basis of a non-parametric SDP estimation was shown to provide a description of the non-linear behaviour involved in the adjustment process of the company analyzed by means of an efficient estimation. This ability was translated into a reduction of the forecasting error on the hold-out sample data. The non-linear weights derived from the model(s) we have proposed have implications for the design of Forecasting Support Systems. If such systems are to be effective in supporting judgmental interventions, they need to help users distinguish between the various sources of information, guiding them in weighting reliable and major pieces of information much more effectively.

Since there is considerable potential in State Dependent Parameters models in this field, further research is needed to analyze a wider range of datasets from more companies with different features, as Fildes et al. (2009) have already shown that companies differ in their responses to information when making adjustments.

Acknowledgment

J. R. Trapero is partly supported by La Consejería de Educación y Ciencia de la Junta de Comunidades de Castilla-La Mancha, and the European Social Fund.

References

Blattberg, R. C., Hoch, S. J., 1990. Database models and managerial intuition: 50% model + 50% manager. Management Science 36, 887–899.

Bryson, A., Ho, Y., 1969. Applied Optimal Control, Optimization, Estimation and Control. Blaisdell Publishing Company.

Eroglu, C., Croxton, K. L., 2009. Biases in judgmental adjustments of statistical forecasts: The role of individual differences. doi:10.1016/j.ijforecast.2009.02.005.

Fildes, R., Goodwin, P., Lawrence, M., 2006. The design features of forecasting support systems and their effectiveness. Decision Support Systems 42, 351–361.

Fildes, R., Goodwin, P., Lawrence, M., Nikolopoulos, K., 2009. Effective forecasting and jugdmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. International Journal of Forecasting 25, 3–23.

Franses, P. H., Legerstee, R., 2009. Properties of expert adjustments on model-based SKU-level forecasts. International Journal of Forecasting 25, 35–47.

Harvey, A., 1989. Forecasting Structural Time Series Models and the Kalman Filter. Cambridge University Press.

Hyndman, R. J., Koehler, A. B., 2006. Another look at measures of forecast accuracy. International Journal of Forecasting 22, 679–688.

Jakeman, A. J., Young, P. C., 1984. Recursive filtering and the inversion of ill-posed causal problems. Utilitas Mathematica 35, 351–376.

Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. ASME Transactions, Journal Basic Engineering 83-D, 95–108.

Mathews, B., Diamantopoulos, A., 1990. Judgmental revision of sales forecasts-effectiveness of forecast selection. Journal of Forecasting 9, 407–415.

Mello, J., 2009. The impact of sales forecast game playing on supply chains. Foresight: The International Journal of Applied Forecasting 13, 13–22.

Moon, M. A., Mentzer, J. T., Smith, C. D., 2003. Conducting a sales forecasting audit. International Journal of Forecasting 19, 5–25.

Pedregal, D. J., Young, P. C., 2002. Statistical approaches to modelling and forecasting time series. Companion to Economic Forecasting Blackwell Publishers, Ch. 4, pp. 69–104.

Syntetos, A. A., Nikolopoulos, K., Boylan, J. E., Fildes, R., Goodwin, P., 2009. The effects of integrating management judgement into intermittent demand forecasts. International Journal Production Economics 118, 72–81.

Taylor, C. J., Pedregal, D. J., Young, P. C., Tych, W., 2007. Enviromental time series analysis and forecasting with the captain toolbox. Enviromental Modelling & Software 22 (6), 797–814.

Young, P. C., 2006. The data-based mechanistic approach to the modelling, forecasting and control of environmental systems. Annual Reviews in Control 30, 169–182.

Young, P. C., Garnier, H., 2006. Identification and estimation of continuous-time, data-based mechanistic (DBM) models for environmental systems. Enviromental Modelling & Software 21, 1055–1072.

Young, P. C., McKenna, P., Bruun, J., 2001. Identification of non-linear stochastic systems by state dependent parameter estimation. International Journal of Control 74, 1837–1857.

FIGURES

List of Figures

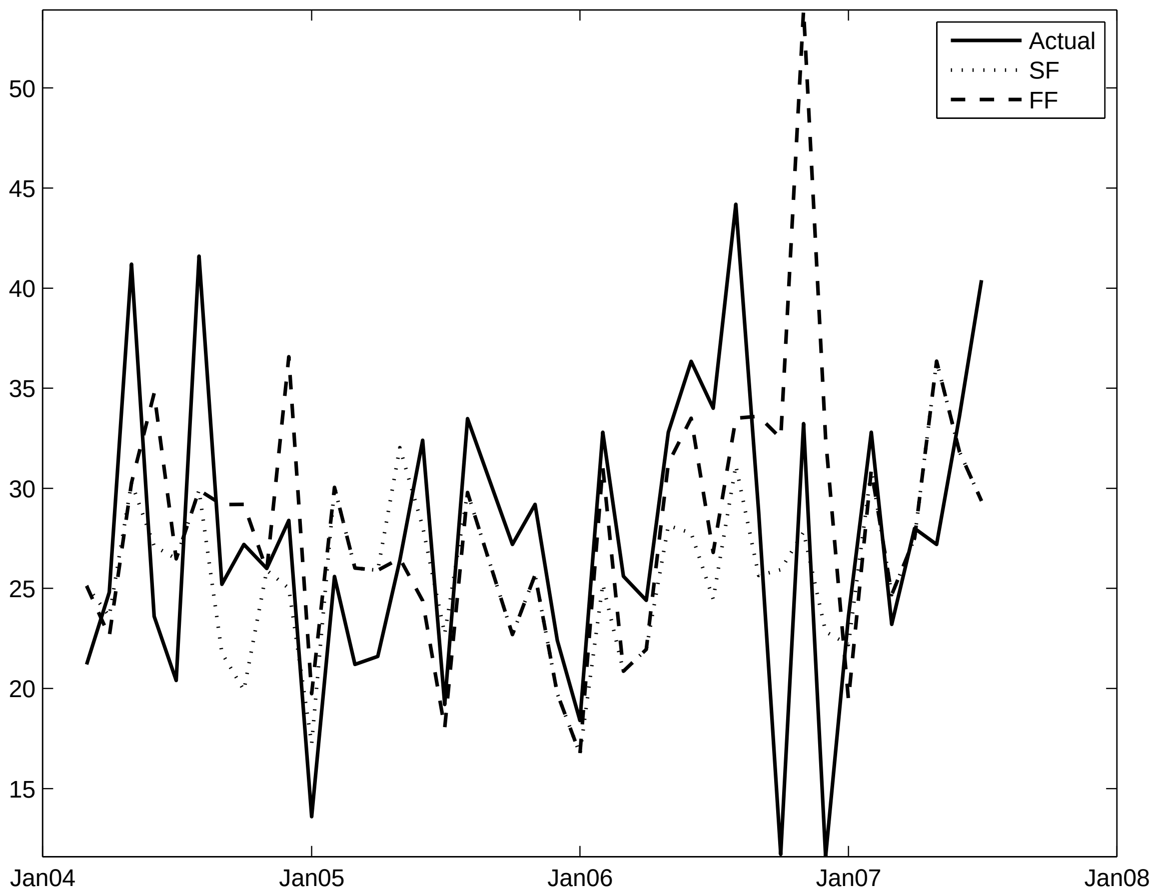

Fig. 1. Example of time series considered per each SKU. Actual values are in a solid line (-), System Forecasts in a dotted line (:) and Final Forecasts in a dashed line (- -)

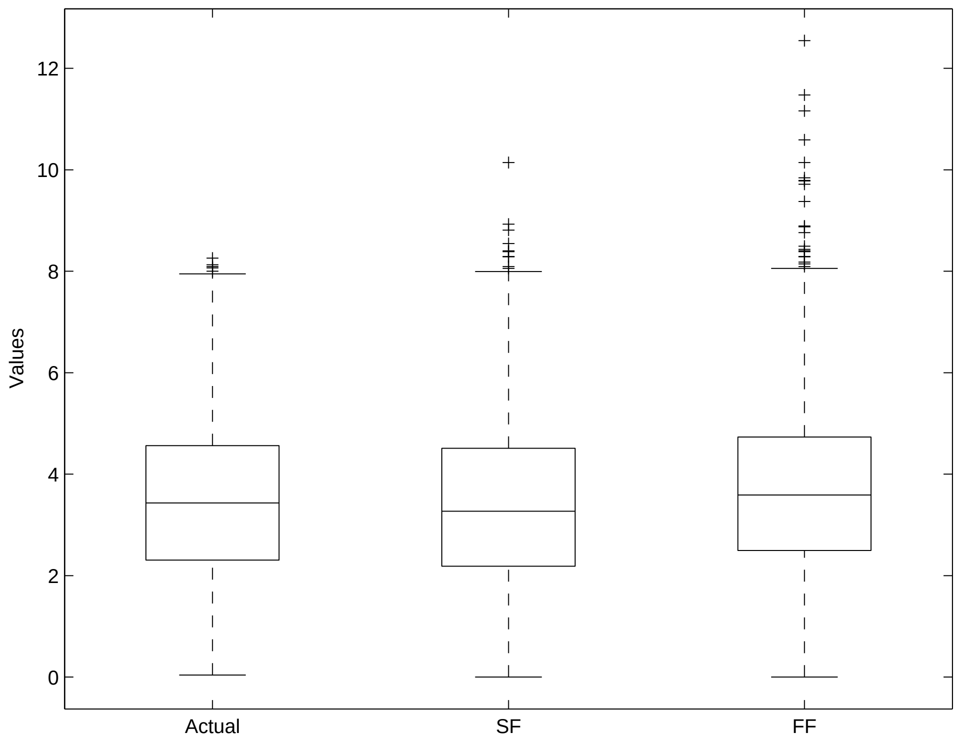

Fig. 2. Boxplots of the normalized Actual values, System and Final forecasts

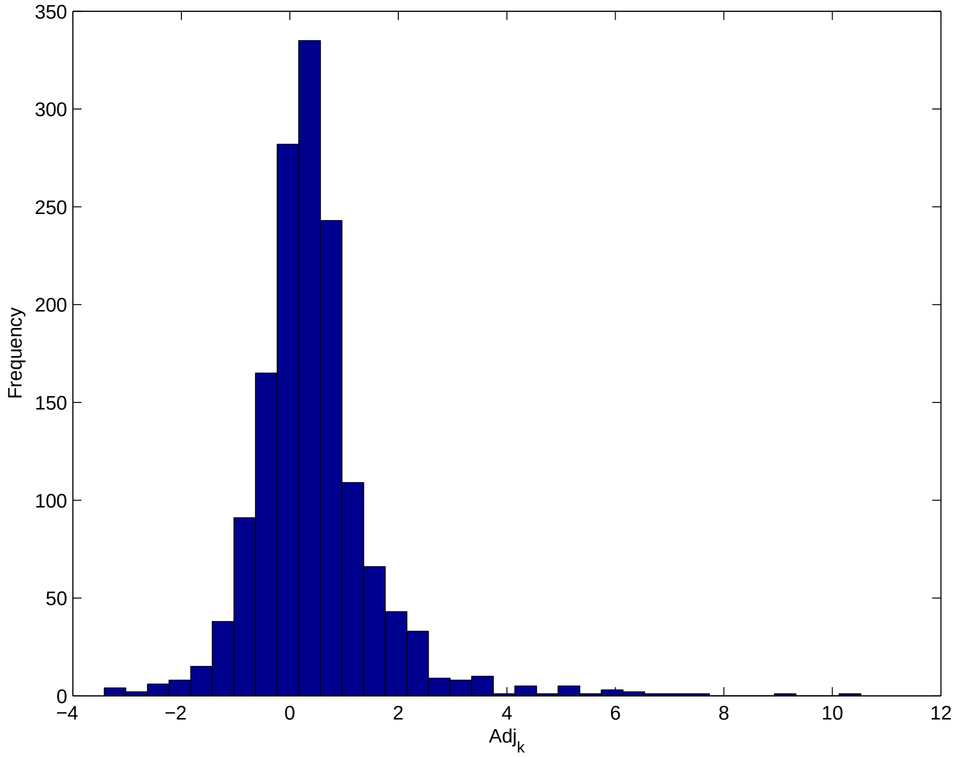

Fig. 3. Histogram of normalized adjustments

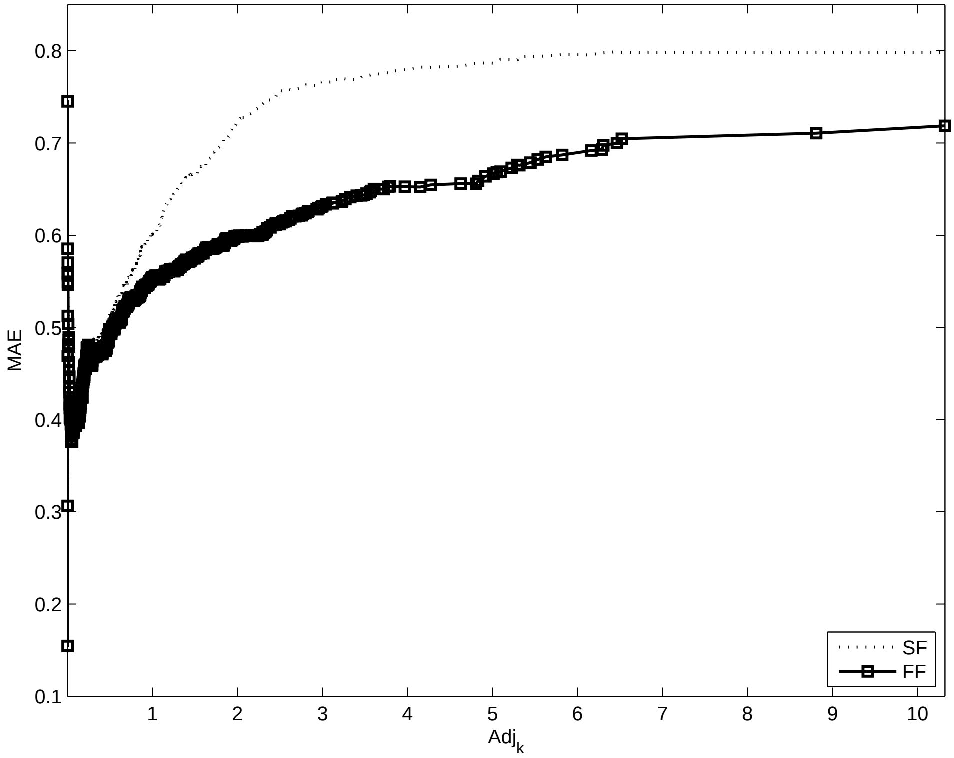

Fig. 4. MAE for normalized positive adjustments

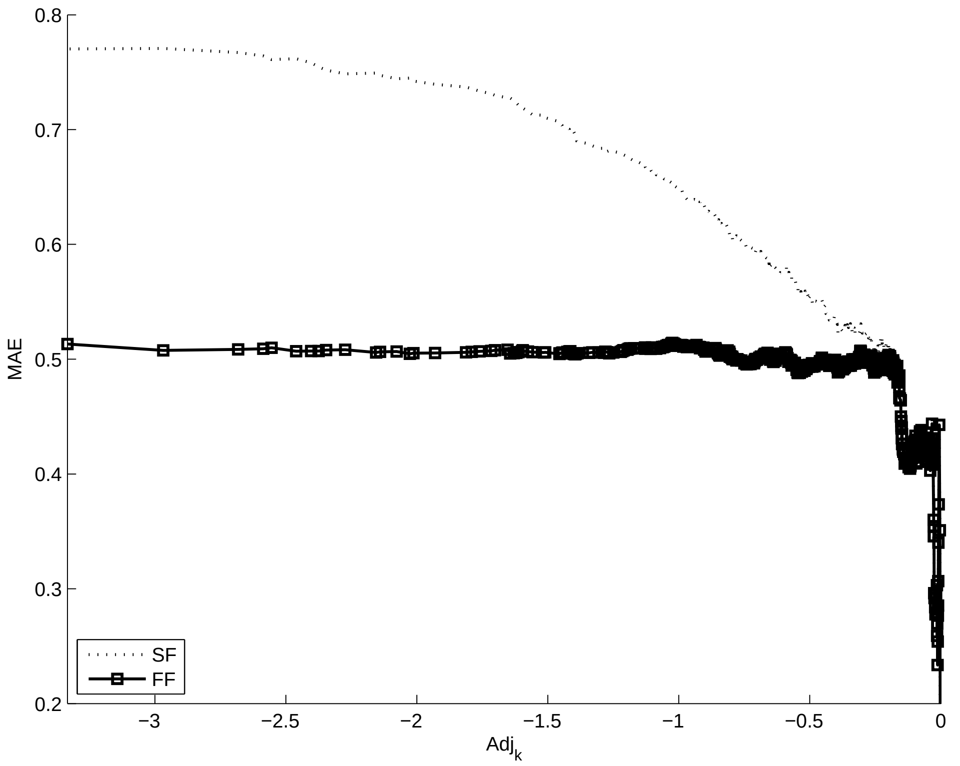

Fig. 5. MAE for normalized negative adjustments

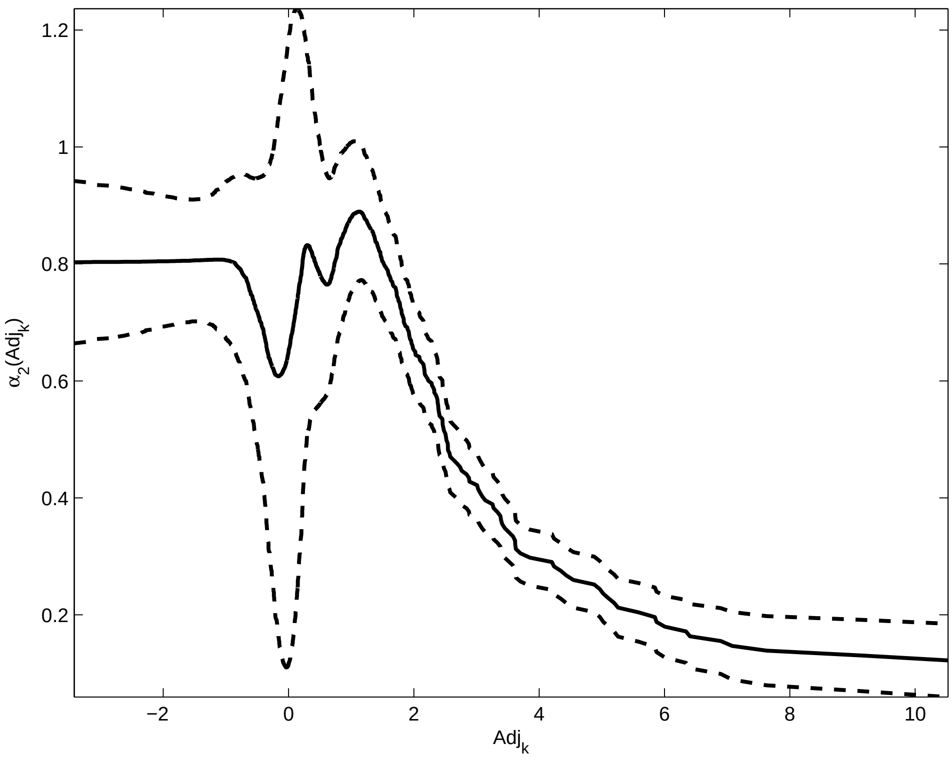

Fig. 6. Non-parametric estimation of $\alpha_2(Adj_k)$ against Adj_k

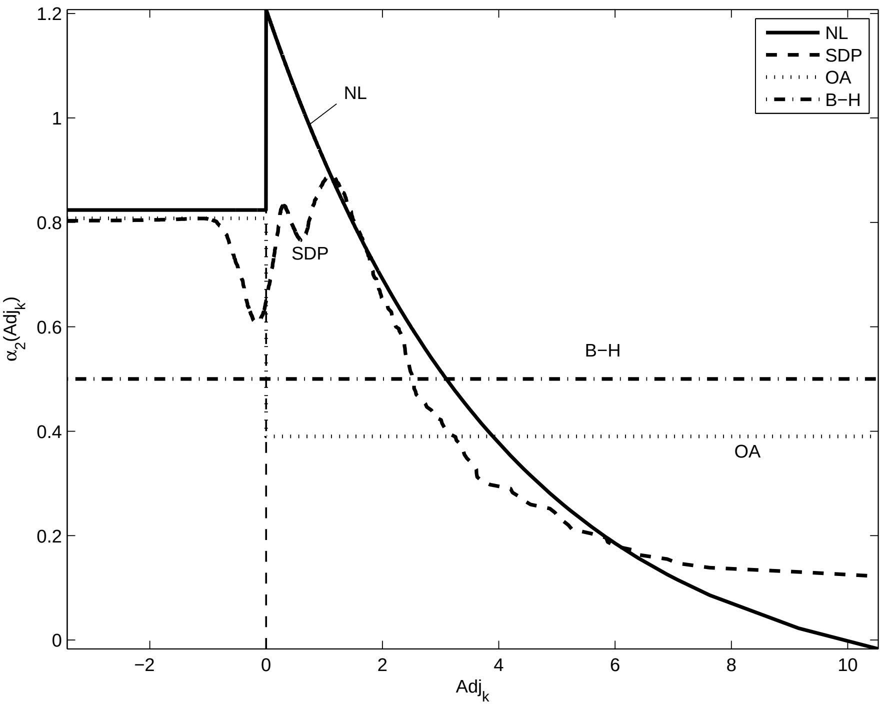

Fig. 7. Adjustment weight estimation accomplished by the non-linear (NL) model (8), the State Dependent Parameter (SDP) model, the Optimal Adjust (OA) model (11) and the Blattberg and Hoch (B-H) model on normalized data (10)

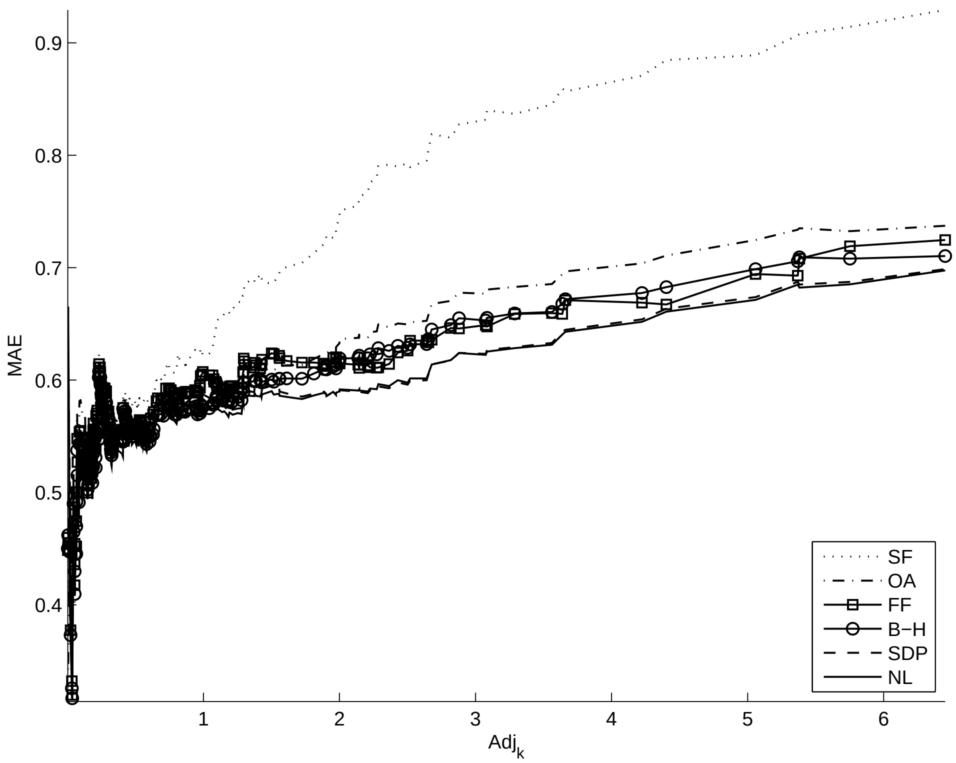

Fig. 8. MAE for normalized positive adjustments on the validation dataset

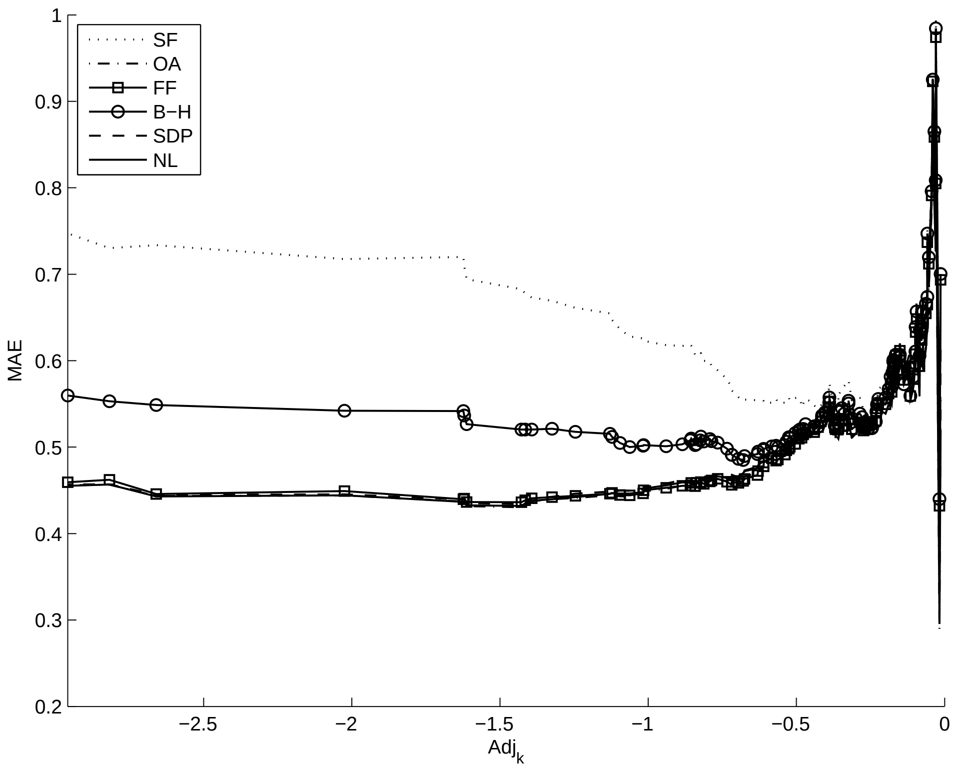

Fig. 9. MAE for normalized negative adjustments on the validation dataset

TABLES

List of Tables

Table 1
Mean of the MAPE and MdAPE for the SF and FF.

Adjustment	No. of observations	Mean(MAPE)		Mean(MdAPE)	
		SF	FF	SF	FF
Positive	1249	**26.99**	32.72	23.37	**18.48**
Negative	601	71.42	**38.24**	34.13	**23.82**
None	1032	30.34	30.34	21.32	21.32
Overall adjusted	1850	40.39	**33.38**	24.04	**17.47**
Total	2882	38.05	**32.71**	21.49	**17.13**

Table 2
Exploratory normalized data analysis

	Actual	System Forecast	Final Forecast
Mean	3.5	3.4	3.7
25th. percentile	2.3	2.2	2.5
Median	3.4	3.3	3.6
75th. percentile	4.6	4.5	4.7
Std. deviation	1.6	1.6	1.6
MAD	1.1	1.1	1.1

Table 3
Evidence of optimism bias in adjustments

Adjustments	% of times adjustment is too large	% of times adjusment is in wrong direction	Total % of adjustments that are overoptimistic
Positive	34.3	27.7	62.0
Negative	25.3	23.0	51.7

Table 4
MAE on the basis of normalized data

Adjustment	System Forecast	Final Forecast
Positive	0.798	**0.719**
Negative	0.779	**0.513**
None	0.552	0.552
Overall adjusted	0.789	**0.652**
Total	0.704	**0.616**

Table 5
Mean of MAPE and MdAPE for the validation dataset

Adjustment	No. of observations	Error (Mean)	System Forecast	Final Forecast	SDP	NL	Optimal Adjust	Blattberg-Hoch
Positive	235	MAPE	**27.04**	33.97	28.13	27.79	27.91	27.71
		MdAPE	25.09	26.78	24.70	24.22	23.71	**23.16**
Negative	125	MAPE	70.72	36.58	39.90	**36.14**	37.10	50.69
		MdAPE	68.06	35.49	39.30	**35.15**	36.07	49.15
None	222	MAPE	35.28	35.28	35.92	**33.06**	33.44	35.28
		MdAPE	24.34	24.34	24.77	**22.82**	23.07	24.34
Overall	360	MAPE	39.00	30.66	28.75	**27.46**	28.06	31.95
Adjusted		MdAPE	26.38	21.30	20.40	**20.15**	20.21	21.34
Total	582	MAPE	38.52	31.08	30.12	**28.53**	29.34	32.61
		MdAPE	23.73	19.09	18.26	**18.23**	18.78	19.46

Table 6
MAE for the normalized validation dataset

Adjustment	System Forecast	Final Forecast	SDP	NL	Optimal Adjust	Blattberg-Hoch
Positive	0.929	0.725	0.699	**0.698**	0.737	0.710
Negative	0.747	0.459	0.457	**0.455**	0.456	0.560
None	0.579	0.579	**0.563**	0.564	0.564	0.579
Overall Adjusted	0.866	0.633	0.615	**0.613**	0.640	0.658
Total	0.757	0.612	0.595	**0.594**	0.611	0.628

Table 7
Standard deviation of the Absolute Error for the normalized validation dataset

Adjustment	System Forecast	Final Forecast	SDP	NL	Optimal Adjust	Blattberg-Hoch
Positive	0.912	0.823	0.693	0.686	0.672	**0.651**
Negative	0.611	0.446	0.433	0.428	**0.426**	0.461
None	0.561	0.561	**0.54**	0.549	0.548	0.561
Overall Adjusted	0.825	0.726	0.626	0.620	0.613	**0.597**
Total	0.749	0.669	0.595	0.594	0.590	**0.585**