

## The advanced Forecasting Information System PYTHIA: an application in time series of mortgage loans in U.K.

**Elli Pagourtzi<sup>a,\*</sup>, Vassilios Assimakopoulos<sup>a,b</sup>, Akrivi Litsa<sup>a</sup>, Fotios Petropoulos<sup>a</sup>**

<sup>a</sup> Forecasting Systems Unit, School of Electrical and Computer Engineering, National Technical University of Athens, 9<sup>th</sup> Iroon Polytechniou Str., 15773 Zografou Athens, Greece.

<sup>b</sup> Secretary for the Information Society, Ministry of Economics, 5-7<sup>th</sup> Nikis Str., 10180, Athens, Greece.

\* Corresponding author and delegate in the Conference. Tel.: +30 210 7723741, Fax +30 210 7723740. E-mail address: [ellipag@epu.ntua.gr](mailto:ellipag@epu.ntua.gr)

### **Corresponding data**

Tel.: +30 210 7723741  
Fax +30 210 7723740

E-mail addresses:

Dr. Elli Pagourtzi: [ellipag@epu.ntua.gr](mailto:ellipag@epu.ntua.gr) (Corresponding author)

Prof. Vassilios Assimakopoulos: [vassim@epu.ntua.gr](mailto:vassim@epu.ntua.gr)

Akrivi Litsa: [alitsa@tpd.gr](mailto:alitsa@tpd.gr)

Fotios Petropoulos: [fotpetr@gmail.com](mailto:fotpetr@gmail.com)

# The advanced Forecasting Information System PYTHIA: an application in time series of mortgage loans in U.K.

## **Abstract**

*The paper presents PYTHIA, a business forecasting platform designed and developed at the Forecasting Systems Unit of National Technical University of Athens. PYTHIA has many functionalities that can be useful not only to individuals that want to perform forecasts but also to managers from companies and organisations that need to forecast the progress of their product or services make the appropriate adjustments and organize their activities. As a case study we used mortgage data in UK that we found from the Council of Mortgage Lenders. Specifically, we use and compare the Theta method [2], [3] and other forecasting methods, available in PYTHIA, to forecast the time series mortgage loan data. We also point out how a bank could use this data and improve its loan policy. The mortgage loan time series used for the forecasts cover the categories of all buyers, first-time buyers and home-movers. Data is organized in quarters, from the first quarter of 1979 up to the first quarter of 2007.*

**Keywords:** advanced forecasting information systems, loan, modelling, monitoring, mortgage, PYTHIA, real estate time series, theta.

## **1. Introduction**

The main scope of the paper is to demonstrate the capabilities of PYTHIA forecasting platform and to compare time series forecasting techniques that were used to forecast mortgage loans in UK. One of the methods used is the Theta method, introduced by Vassilis Assimakopoulos and Kostas Nikolopoulos [2], [3] and presented thoroughly in previous papers [8], [9] as part of the Theta Forecaster application [1].

The paper also analyzes the recent trends concerning the mortgage market in U.K., shows how the loan forecasts provided from PYTHIA can be useful to a Bank and compares the two forecasting applications developed at our Unit, PYTHIA that we present here and Theta Forecaster presented in previous papers [1], [8], [9].

The tool used to provide the forecasts, PYTHIA, is a forecasting platform designed and developed at the Forecasting Systems Unit of National Technical University of Athens. PYTHIA aims mainly at managers, as it provides managerial forecasting that includes data analysis and adjustments, budgets, long term forecasts, monitoring, reporting etc. It includes the Theta Method and several others, well-established forecasting methods. The main feature of the Theta Method is that it applies different techniques to deal with short-term and long-term forecasts and allows giving different weights in the short and long-term components [2], [3]. PYTHIA allows the combination of two or more techniques, which in many cases produces better forecasts than using a stand-alone method [4].

The loans time series used for the forecasts cover the categories of all buyers, first-time buyers and home-movers. Data is organized in quarters, from the first quarter of 1979 up to the first quarter of 2007. The mortgage loan data was found from the Council of Mortgage Lenders, the trade association for the mortgage lending industry, the members of which account for around 98% of UK residential mortgage lending [10].

The aim of the Council of Mortgage Lenders is to help to foster a favorable operating environment in the UK housing and mortgage markets. It is the representative voice for the residential mortgage lending industry, and the central provider of economic, statistical, legal, research and other market information. Its members are banks, building societies and other mortgage lenders. The Council of Mortgage Lenders also has associates, drawn from a variety of related businesses, including lawyers, conveyancers, search companies and management consultants.

## **2. PYTHIA – Business Forecasting application**

PYTHIA is basically aimed at practicing managers (at the level of financial directors, product managers, production/inventory managers and planners/analysts) and it is designed and developed with a single purpose in mind, that of making the task of managerial forecasting as straightforward, user-friendly and practical as possible.

### **2.1 Objectives**

PYTHIA is designed and built to satisfy four objectives:

- First, any manager, without a technical/statistical background, should be able to use it easily and intuitively.
- Second, since practically all data contains “systematic noise” from external sources, the user should be able to identify and adjust his/her data in a simple, straightforward manner.
- Third, information about the data and the forecasts should be readily available to the user who should, in addition, be able to input his/her knowledge and experience into such forecasts by incorporating judgmental knowledge and/or budgetary and other objectives in order to improve the accuracy of the statistical methods.
- Finally, human forecasters are often overoptimistic while also including their own biases into their forecasts. Forecasters are, therefore, given comparative information about the accuracy of their own forecasts, in comparison to the statistical ones, in order to identify systematic biases and correct them in the future when forecasting again.

## 2.2 Modules

PYTHIA is made up of seven modules each focusing on a particular aspect. PYTHIA also utilizes a wizard which advises the user about ways of adjusting his/her data and the best module to utilize. Those modules and the wizard are briefly presented in following:

- Data Analysis and Adjustments: All data is influenced by external events and/or actions (strikes, extreme weather conditions, promotions/advertising, big price changes etc. and even recording or other mistakes) that “corrupt” past patterns and established relationships. For this reason they need to be adjusted. The first module of PYTHIA recommends to the user ways of adjusting his/her data before forecasting. In addition, to the usual adjustments (trading/working days, missing values, outliers) PYTHIA also provides the option for adjusting for Special Events/Actions (SEA) that usually exert a significant influence on sales/revenues and greatly affect the accuracy and reliability of forecasting that is based on such data. When data adjustments are made they are kept and are subsequently utilized by all modules.
- Budgets: PYTHIA makes available forecasts for the entire company, each division, country/geographical area, major product category, product and item. It also allows forecasting using a bottom-up approach, or basing its predictions at any specified level of aggregation. Consequently, the person in charge can modify the statistical forecasts to incorporate judgmental information and/or budget objectives. Once the adjusted forecasts are finalized they can be allocated proportionally to all categories below, or in some precise manner specified by the user.
- Production/Inventory: Many organizations require many thousands of forecasts on a monthly, weekly or, in some cases, even daily basis. These forecasts need to be produced automatically, efficiently and as accurately as possible. PYTHIA supplies such forecasts in a mechanical manner when and where they are needed. It also provides an option to resolve the difference between bottom-up and aggregate forecasts, including judgmental and/or target adjustments, and allocates possible differences proportionally or according to the preferences of the user.
- Long Term: Predicting long term trends and estimating growth rates requires different methods and considerably more assumptions as environmental, competitive and technological changes affect long term trends. This module provides information and advice about making long term predictions while presenting information about their accuracy and reliability.
- Estimating Relationships: The above modules deal with time series forecasting while the present module allows for the estimation of relationships based on multiple regression. In addition to such estimation it also provides pertinent information that can be used for forecasting purposes and for getting a better understanding of the factor driving changes in the future.
- Monitoring: Past patterns and established relationships can and do change invalidating the forecasts that inevitably are based on the extrapolation of such patterns and the continuation of relationships. The Monitoring module continuously checks for the difference between the actual and predicted values and provides a warning signal when such differences (i.e., the forecast errors) cease to be random. In such a case the user can re-forecast and/or take appropriate actions, if needed, to correct the situation.
- Reporting: PYTHIA supplies detailed reporting information on all aspects of data analysis and forecasting while also showing detail information about the accuracy of different types of predictions (original data without adjustments, adjusted data, statistical forecasts, judgmental overrides, budget objectives, and final forecasts). The reports can be customized and are exported in an Excel sheet for further usage.
- Wizard: PYTHIA uses a wizard which advises the user about ways of adjusting his/her data and the best module, from those listed above, to utilize. Once such a module has been selected, PYTHIA analyzes the available data, informs the user on their characteristics and selects the most appropriate method for such data with the objective of improve forecasting accuracy and efficiency. Expert users can override the suggested method and/or modify its utilization according to his/her preferences.

### **2.3 Audience**

PYTHIA's intended audience is practicing managers working at various functions in the organization. More specifically each module is geared to the following type of managers:

- Budget: Financial Directors at the level of the company, division or department concerned with the preparation and/or approval of budgets and Product Managers at the level of the product/service, channel or geographical area who are in charge of selling/providing services and who have to achieve certain revenue or profit targets.
- Production/Inventory: Production/Inventory managers who have to ensure that there are satisfactory levels of inventories of raw and other material to be able to produce enough products to meet market demand.
- Long Term: Heads of Divisions/departments as well as Business Planners who need to assure satisfactory growth objectives and profitability for the medium and long term.
- Data Analysis: Practicing managers can be assisted by business analysts in the preparation stages of inputting and adjusting the data and of obtaining the forecasts. Managers can then concentrate their efforts in incorporating their knowledge and experience into adjusting the statistical forecasts and setting the budget and other objectives to be achieved in each category.

### **2.4 Some applications**

PYTHIA can be used for example by financial directors at headquarters to set their budgets based on growth objectives as well as to forecast established trends by major product category, by country (in the case of multinational companies), by major customer or any other desired clarification. Country managers can prepare their own budgets (also based on the targets set by headquarters and local trends) and submit them to headquarters for approval. Product managers can prepare estimates for the sales of their own products and/or territories while manufacturing/inventory managers can predict how much to produce, or stock, at the level of each specific item (SKU, Stock Keeping Unit). PYTHIA can become part of an ERP, such as Navision, while it can also provide the forecasts required by practically all business situations requiring accurate predictions.

### **2.5 Examples of screens**

Some of the main screens of PYTHIA are indicatively presented in following. The example used and shown in the following screens is forecasting sales of Microsoft products.

After starting the application the user is asked to provide an appropriate username and password to log in into the application (Figure 5). After providing this information the user can enter one of the modules of the application. When loading data from Excel files there is no need to follow this procedure (no username and password needed). The application is being terminated when the user presses the "EXIT" button. When leaving a module this screen reappears.

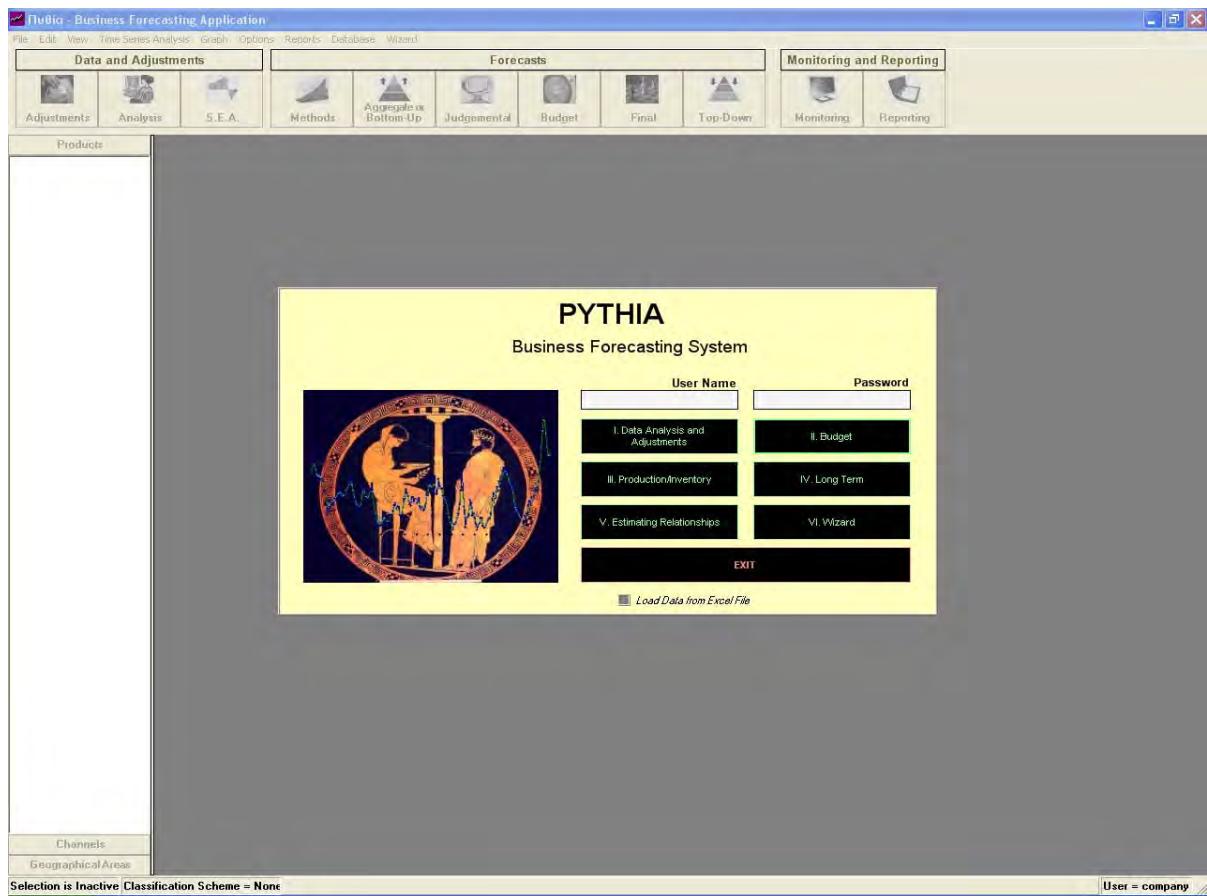


Figure 5: PYTHIA Logging in

Figure 6 shows the adjustment of missing values. Missing Values are values which are not available for some reason. In order to produce accurate forecasts, "PYTHIA" requires all data. So missing values can be 'reproduced' by three methods. User can select one of them just by clicking the appropriate radio box. Those three methods are:

- Set Missing Value at the average of previous and next value.
- Set Missing Value at the average of respective time periods. This refers to seasonal time series. For instance, if monthly data is available and there is a missing value for one July, "PYTHIA" set this missing value at the average of all other Julys.
- Set Missing Value at a certain level. This certain level is default zero (0) but user can define a different one if wants to.

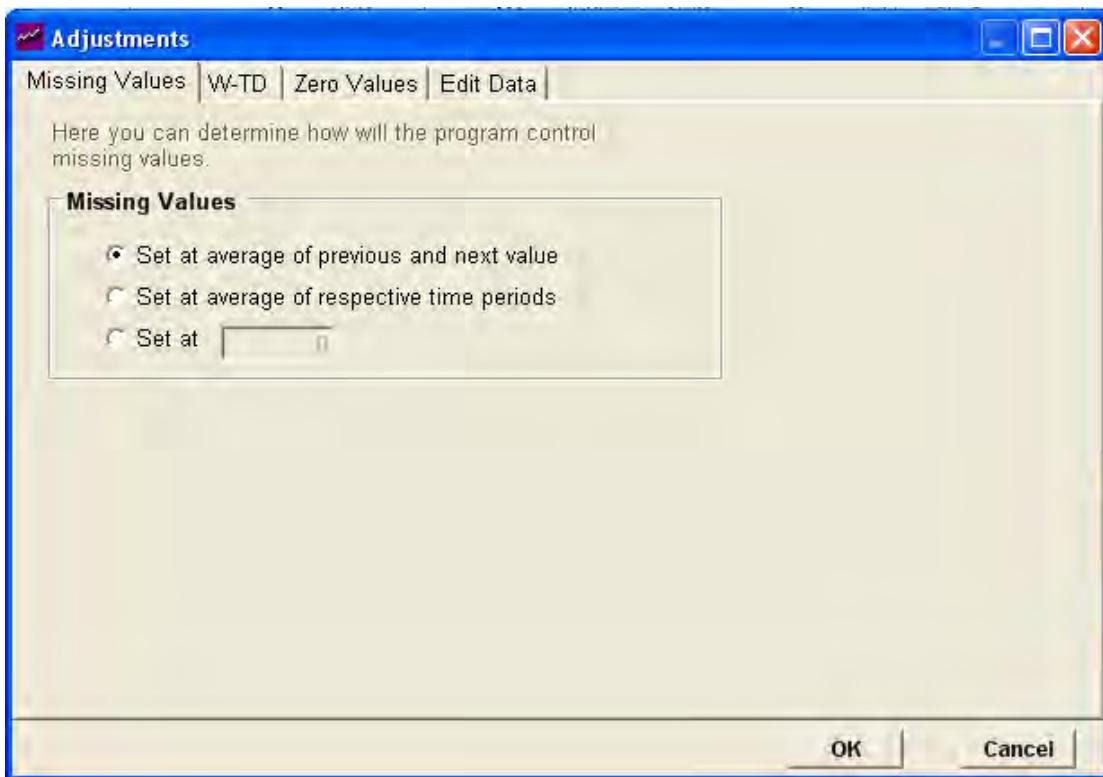


Figure 6: Adjustments - Missing Values

Time series containing large number of zero values can cause false forecasts and undesired results. It is considered essential that zero values should be adjusted if user wishes. The options of this adjustment are presented in Figure 7 and the user can select one of them by clicking the appropriate radio box:

- No change of zero values. Usually this is selected if zero values are few.
- Treat zero values as missing values. If this option is selected, then zero values will be treated as missing values as described above.

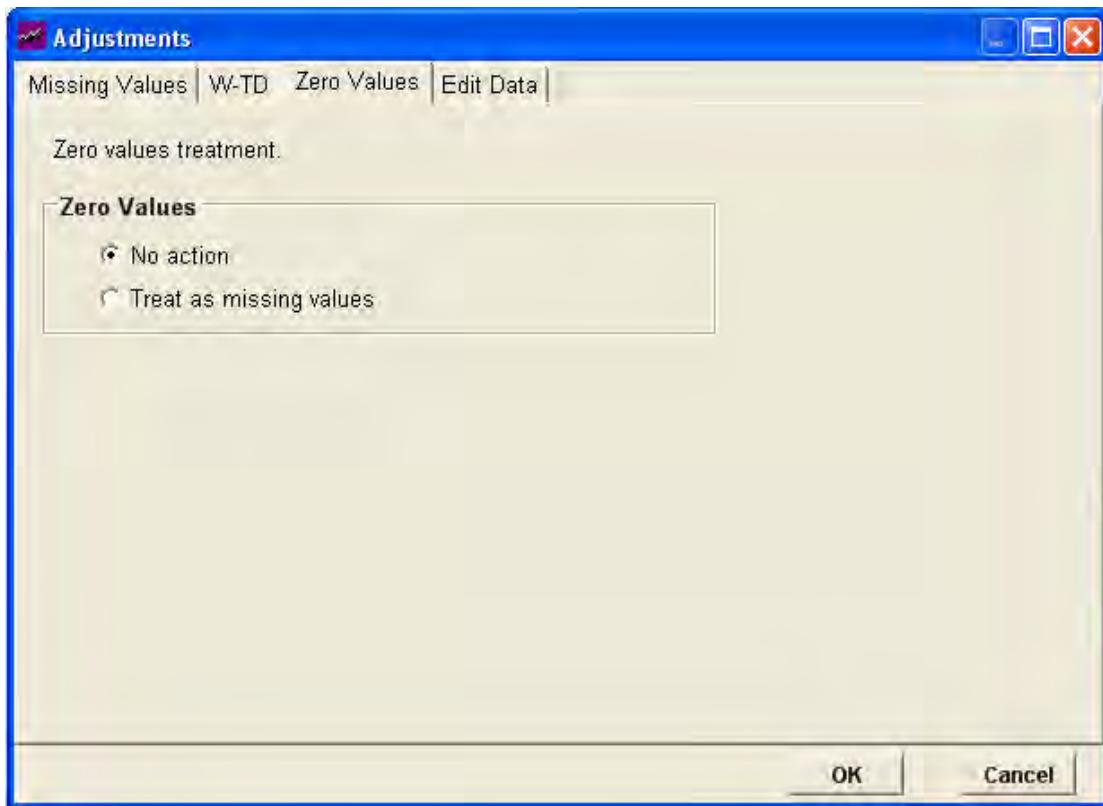


Figure 7: Adjustments – Zero Values

Working Days of the company such as Bank Holidays can cause undesired abnormalities in the time series. In the tab shown in Figure 8, user can smooth the time series by selecting “Make Date Adjustments” radio box, so the date adjustments are enabled. The date adjustments are of two kinds. Firstly, user can define the country in which the company is based, so the bank holidays can be calculated. This can be enabled by checking “Calculate Bank Holidays” checkbox and selecting the country. Secondly, user can select the working and trading days of the company easily by checking the respective checkboxes in “Days” group box.

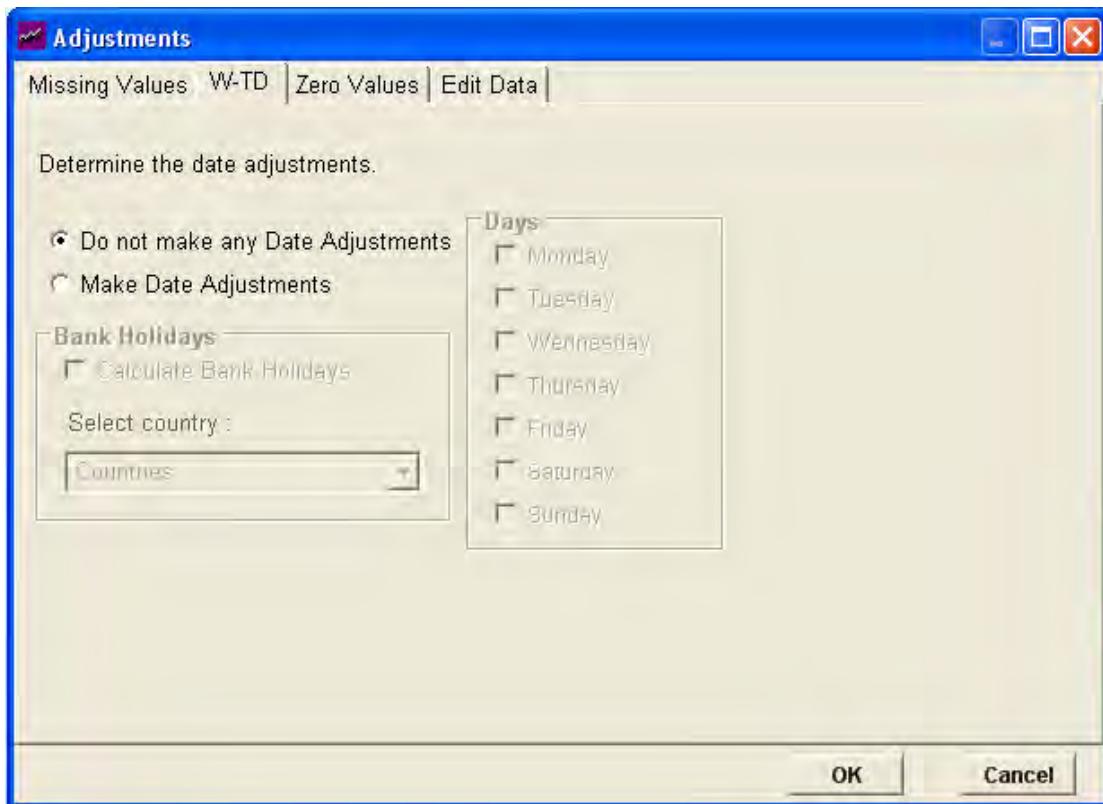


Figure 8: Adjustments – Working & Trading Days (W-TD)

"PYTHIA can calculate easily and accurately a large number of statistical values. The study of those values can provide interesting conclusions about the time series and expert users use them in order to select the appropriate forecast method which minimizes the forecasts errors. The statistic measures calculated by "Pythia" are shown in Figure 9 that follows:

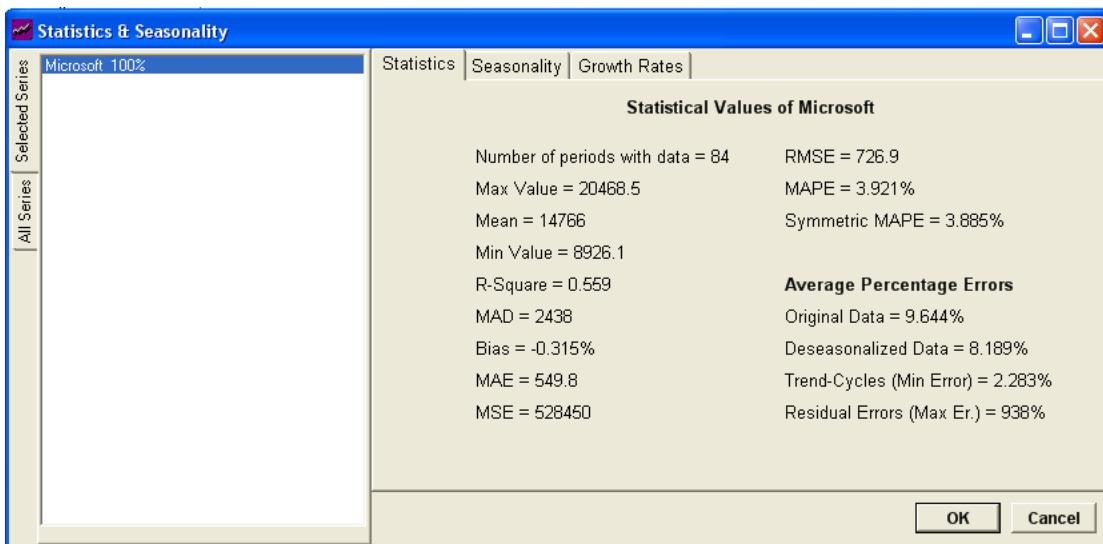


Figure 9: Statistic analysis

A seasonal pattern exists when a series is affected by seasonal factors. Sales of products such as ice creams have this type of pattern. This pattern is calculated by several methods. Once calculated, time series is adjusted. "PYTHIA" uses the Classical Decomposition Method to calculate seasonal patterns. In the tab shown in Figure 10, "PYTHIA" informs the user upon the

seasonal indices of the time series. There are two possible views of the indices. "Initial Seasonality" and "Final Seasonality" can be viewed by clicking on one of them. The difference between those is that the initial seasonal indices are calculated in the original data, without any Date or SEA (Special Events and Actions) adjustment where the final seasonal indices contains those adjustments.

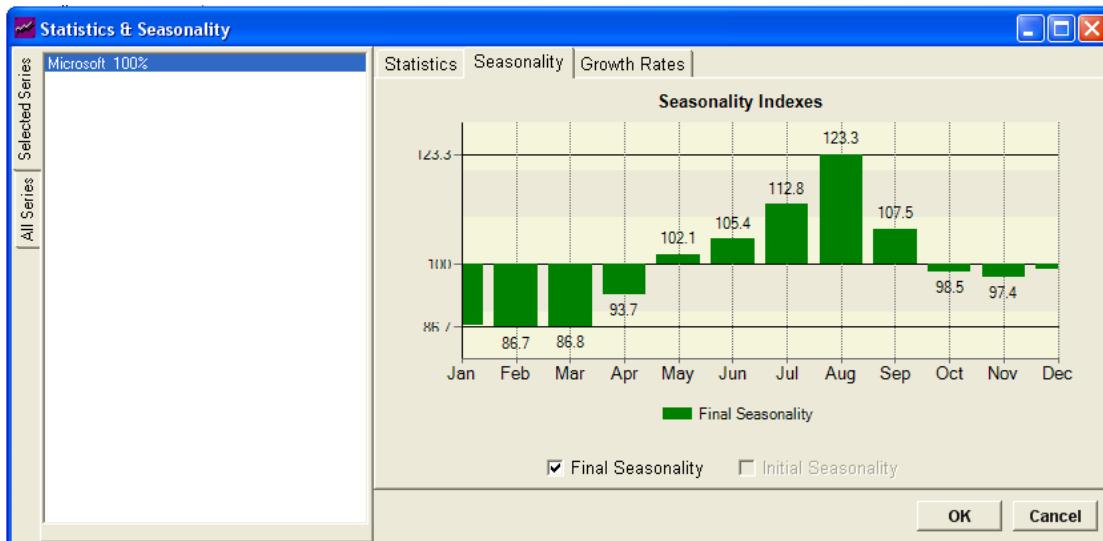


Figure 10: Seasonality

Special Events or Special Actions (SEA) are the below cases:

- Promotions (Amount/Type)
- Advertising (Amount/Media)
- Price Increases
- Price Decreases
- Substitute Products: Price Increases
- Substitute Products: Price Decreases
- New Products (Performance/Price Ratio)
- Insufficient Inventories
- Insufficient Inventories of Competitors
- Competitive Promotions (Amount/Type)
- Competitive Advertising (Amount/Media)
- Competitive Price Increases
- Competitive Price Decreases
- Competitive Substitute Products: Price Increases
- Competitive Substitute Products: Price Decreases
- New Products Introduced by Competitors (Performance/Price Ratio)
- Unusual Weather Conditions
- Strikes
- Strikes in Competitors

Once a SEA is entered in this list, user can modify / edit its aspects or delete it (see Figure 11). This can happen by clicking on it and press the appropriate button "Edit" or "Delete". If "Edit" button is pressed, the "Edit SEA" form is appeared (Figure 12). Here user can define the characteristics of the selected SEA. Firstly, user must select the type of the SEA. Types of SEA are presented above. Secondly, user can make changes at the date where SEA begins or ends. This can change via functional calendar tool or via up-down buttons. Thirdly, user can define the budget of the SEA (if the SEA has one such as a promotion). Fourthly, Comments on the SEA can be saved by entering text in the appropriate textbox. Lastly, the impact of SEA in the selected period can be changed. These impacts are presented on the graph. The grey line presents the

automatic calculated impacts and the blue line-bars presents the impacts that will be considered as true. By default the impacts are zero. If user presses button "Set custom impacts as automatic" then the automatic impacts are presented. Users can also change the impacts manually easily by dragging the desired bar (period) on the graph towards the desired direction. This action is followed by a confirmation dialog box in which user can make his manually change more accurate. Changes are saved by clicking "OK" button or discarded by clicking "Cancel" button.

Once the changes are saved (by pressing the "OK" button) the programs returns to the "Special Events & Special Actions" form. If impact is changed, an additional graph is now visible, just under the main graph. This graph presents the total impacts of all the identified SEA. Its visibility can change simply by clicking on the "View Impact Chart" checkbox. Additionally, there is the option of "View Adjusted Data". By clicking on this checkbox, user can see on the main graph how the data would be if there was no SEA. This function is presented by a green line. If impact is set as zero, then there will be no difference from the original data (yellow line). If the program has not identified a SEA, user can insert it manually by clicking the "Add" button. "Edit SEA" form will appear and user continues as before.

It is mentioned that PYTHIA uses four different methods in order to identify SEA. Only one method can be applied at a time. User can change the method of identification simply by selecting the appropriate one from the "Ratio" combo box which is included in the "Methods" group box. User can also define the levels of sensitivity of the methods by changing their thresholds. Note that if a threshold is increased then the method becomes more sensitive and so more SEA are identified. Each method has one threshold except the first method which has two thresholds. Once all SEA are identified and modified, user can press "OK" button to save the changes or else "Cancel" button.

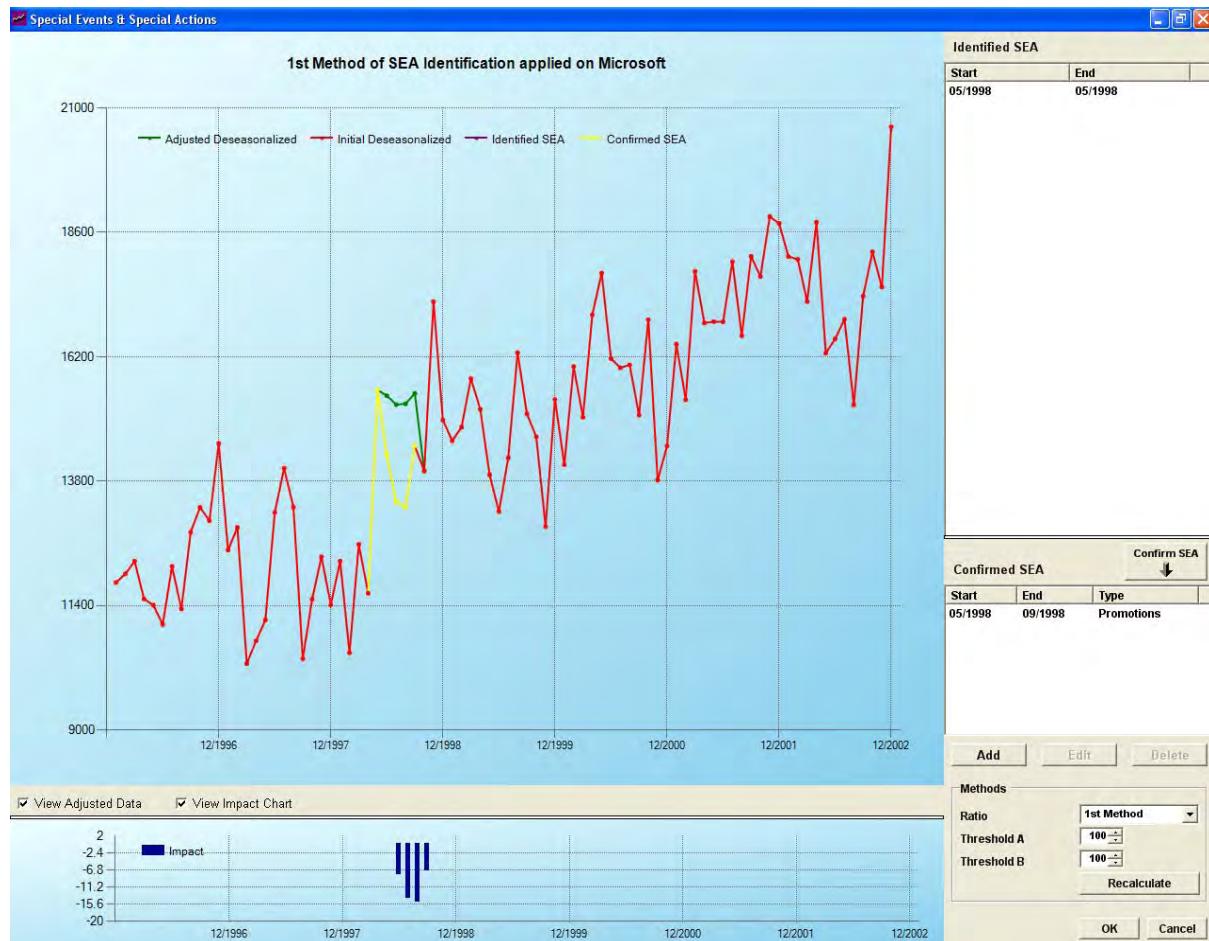


Figure 11: Special Events & Actions (SEA)

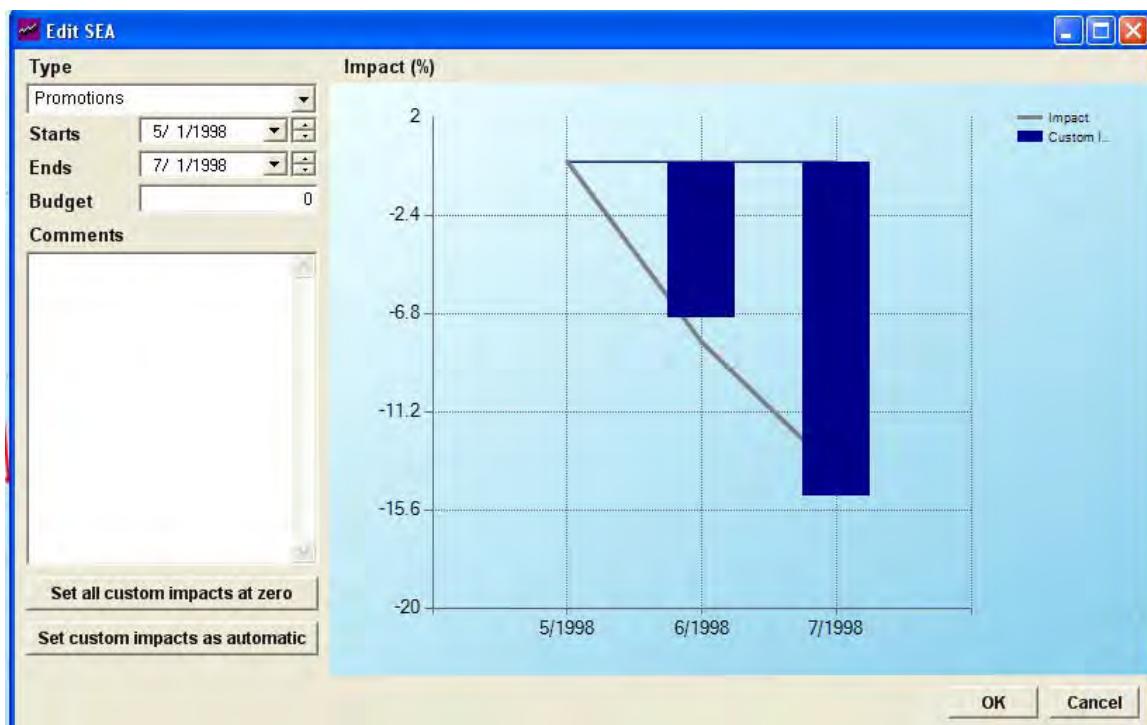


Figure 12: Edit Special Event Action

By pressing the “Methods” button in the initial log-in screen, the user has the opportunity to choose one of the following methods to produce the statistical forecasts, as shown in Figure 13:

- Single Exponential Smoothing (SES)
- Holt
- Winters
- Damped
- Linear Trend
- Exponential Trend
- Theta
- Averaging
- Eclectic

The Eclectic Method, which is the default method used by the application, chooses automatically a combination of methods to produce forecasts. The Averaging Option allows producing forecasts using a specific combination of two or three different methods. Custom weights for each of the contributing methods can be determined.

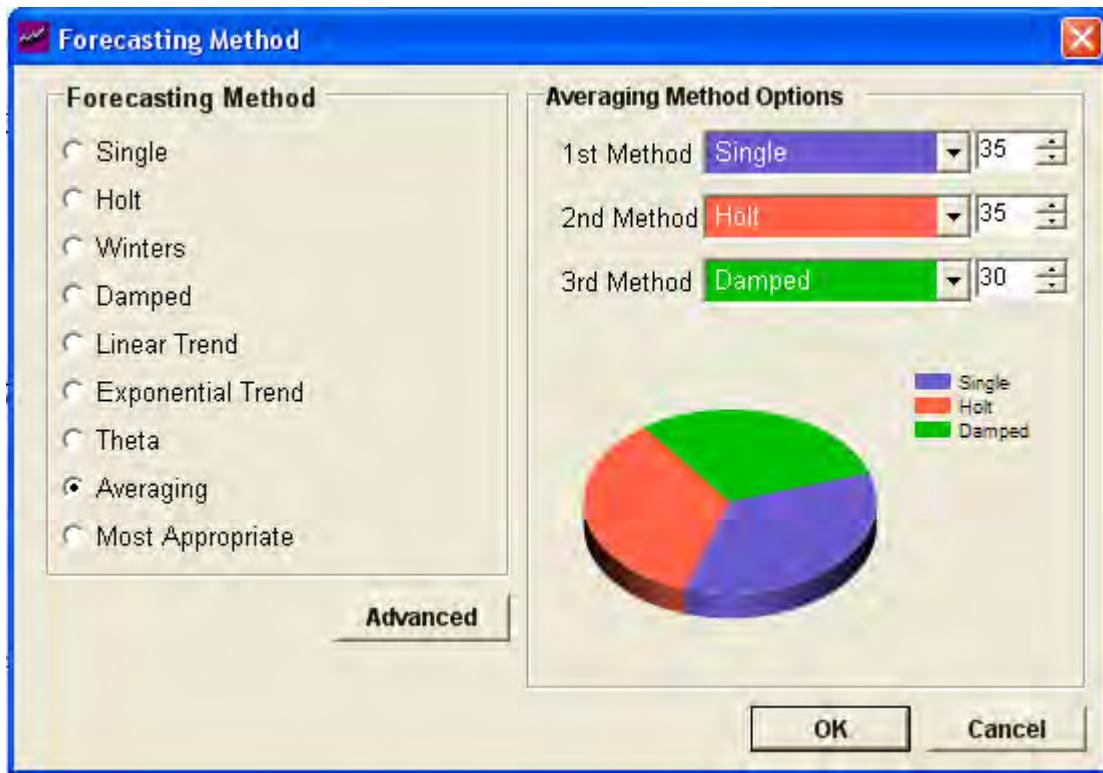


Figure 13: Forecasting Methods

After choosing the forecasting method the user can choose the “Aggregate or Bottom-up” button in the initial screen. Thus he is able to compare the produced forecasts (aggregate statistical forecasts) with the bottom-up forecasts (which are the sum of forecasts that have been produced for the subcategories of the item that the user examines). The differences of these two forecasts have to be reconciled. There are three options, to use the aggregate forecasts, to use the bottom up forecasts or to use the average of the two as statistical forecasts (see Figure 14):

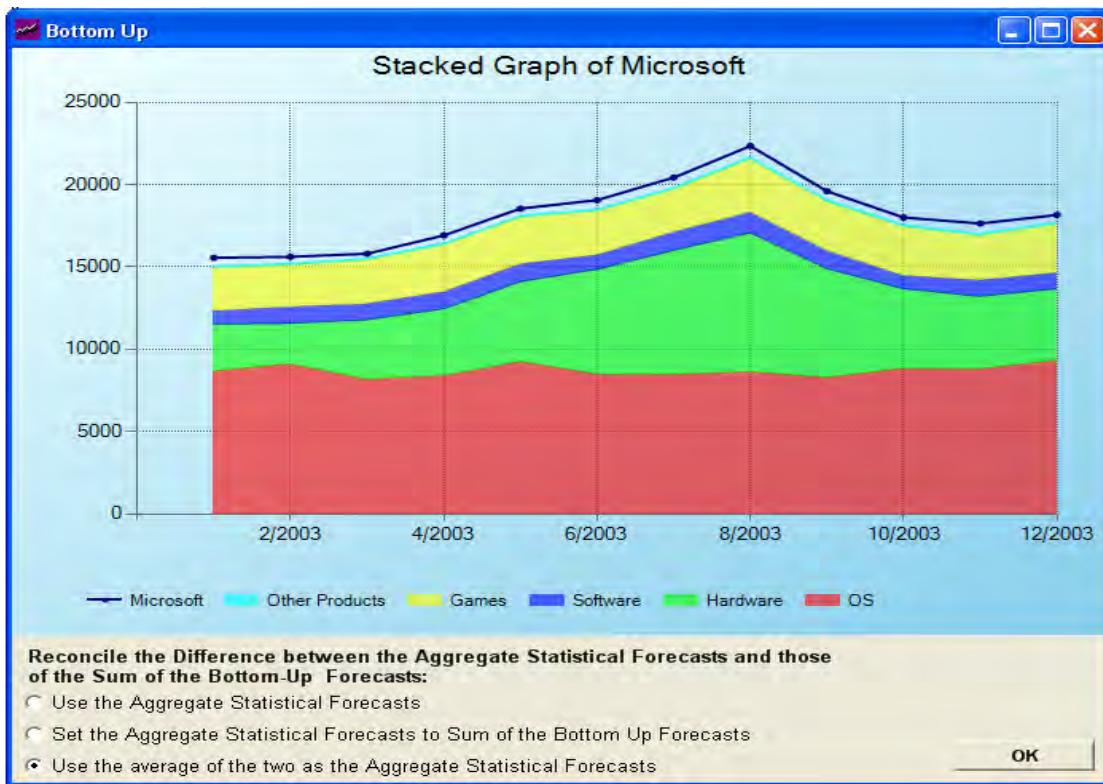


Figure 14: Aggregate or Bottom Up Forecasts

By pressing the “Judgmental” button, the user is able to define forecasts based on his/hers judgment. These forecasts can reflect the optimism or pessimism of the user based on knowledge that a forecasting method can not have. For example, if a great promotion is going to take place in the near future, the judgmental forecasts could include this fact and show increased values in comparison to the statistical forecasts. There is also an opportunity to see the forecast without seasonality by pressing the related button and undo or redo changes by pressing the undo or redo button. Judgmental forecasts are shown in Figure 15:

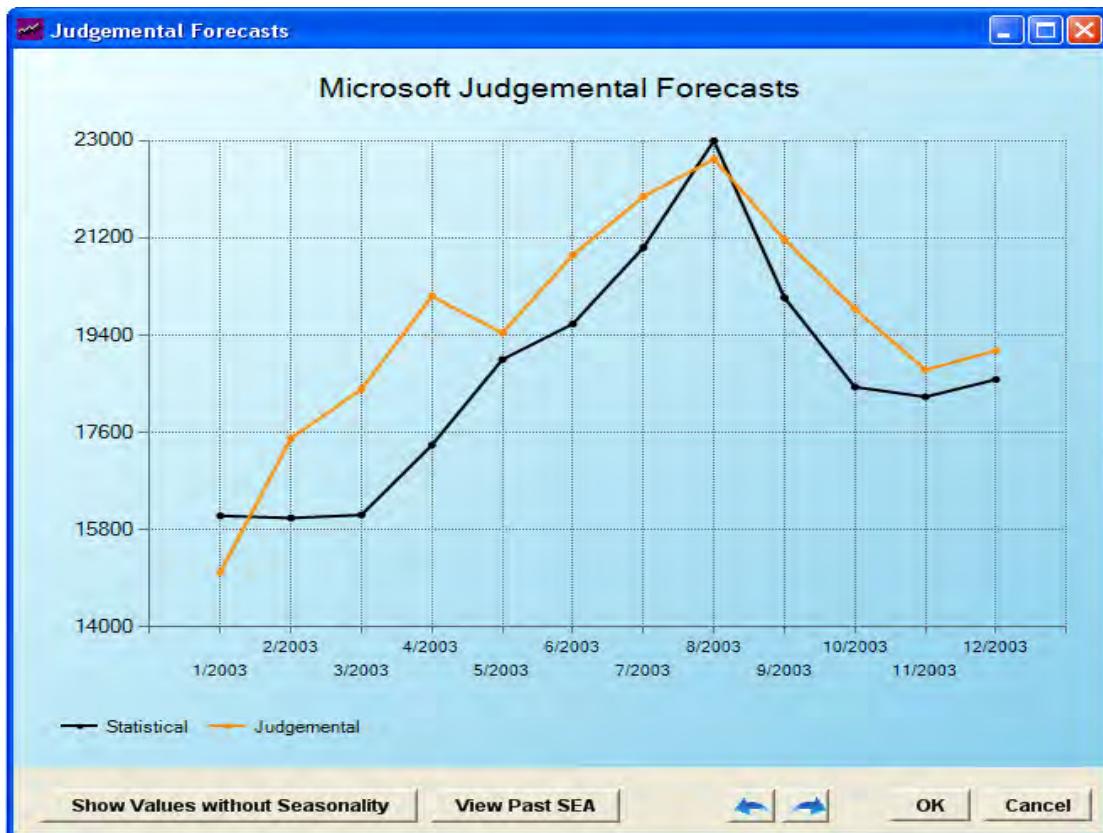


Figure 15: Judgmental forecasts

By pressing the “Budget” button the user can define the additional growth of the forecasts due to the increase or decrease of the budget. This additional growth can be applied on the statistical or the judgmental forecasts. It can be allocated proportional to all periods or show its effects on the late periods. We can see an example in Figure 16:

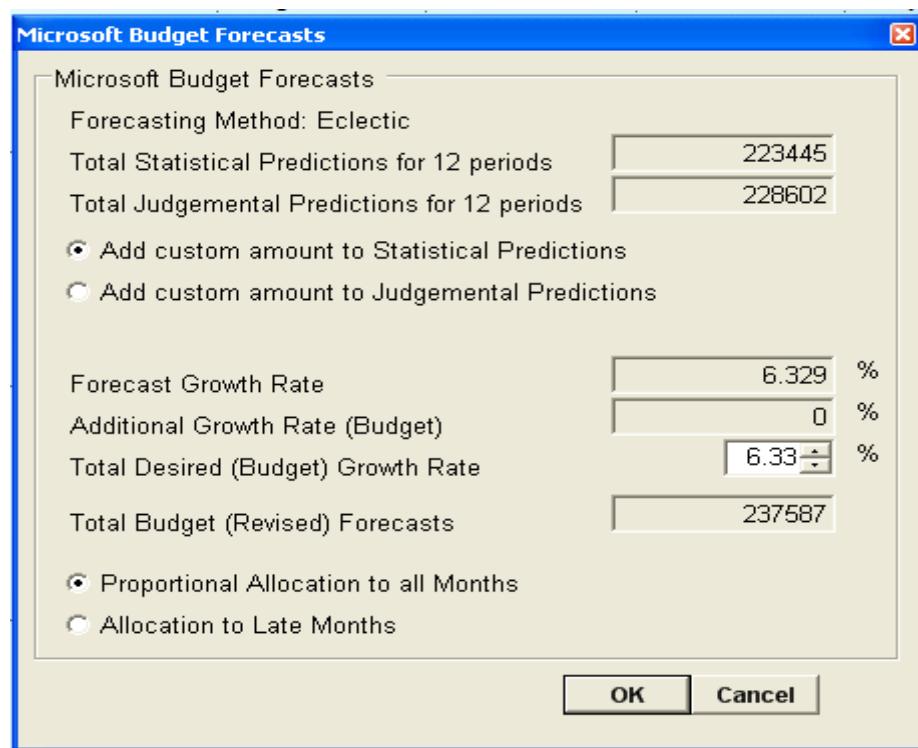


Figure 16: Budget Forecasts

The final forecasts can be defined by pressing the “Final” button and the screen that appears is shown in Figure 17. Here the user has the ability to view all the forecasts that have been produced (statistical, judgmental and budget) and choose which of them or which combination of them will be the final forecast. There is also an opportunity to see the forecast without seasonality by pressing the related button and undo or redo changes by pressing the undo or redo button.

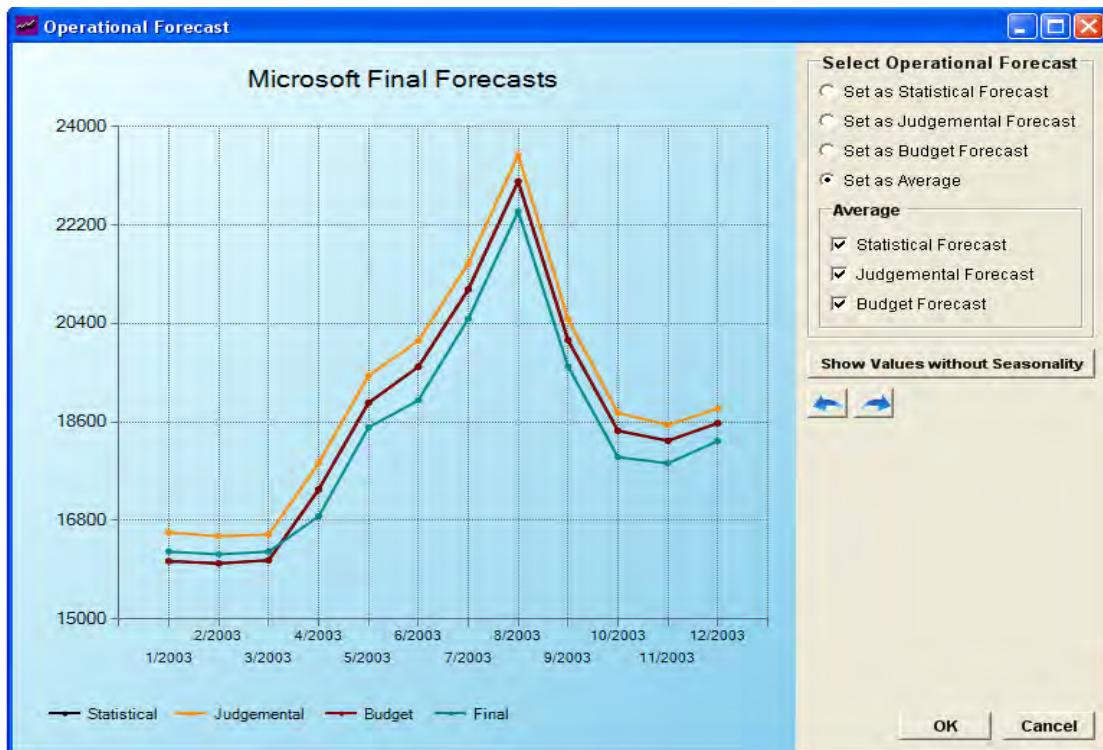


Figure 17: Final / operational Forecasts

After having specified the final forecasts, the user has to determine the forecasts for the underlying categories. By pressing the “Top Down” button the user is able to allocate the final forecasts to each subcategory. This can happen by reconciling the differences between the final forecast and the sum of the final forecasts of the subcategories proportional to each period or according to the grand total. This produces new forecasts for each subcategory which are called Top Down Forecasts (see Figure 18):

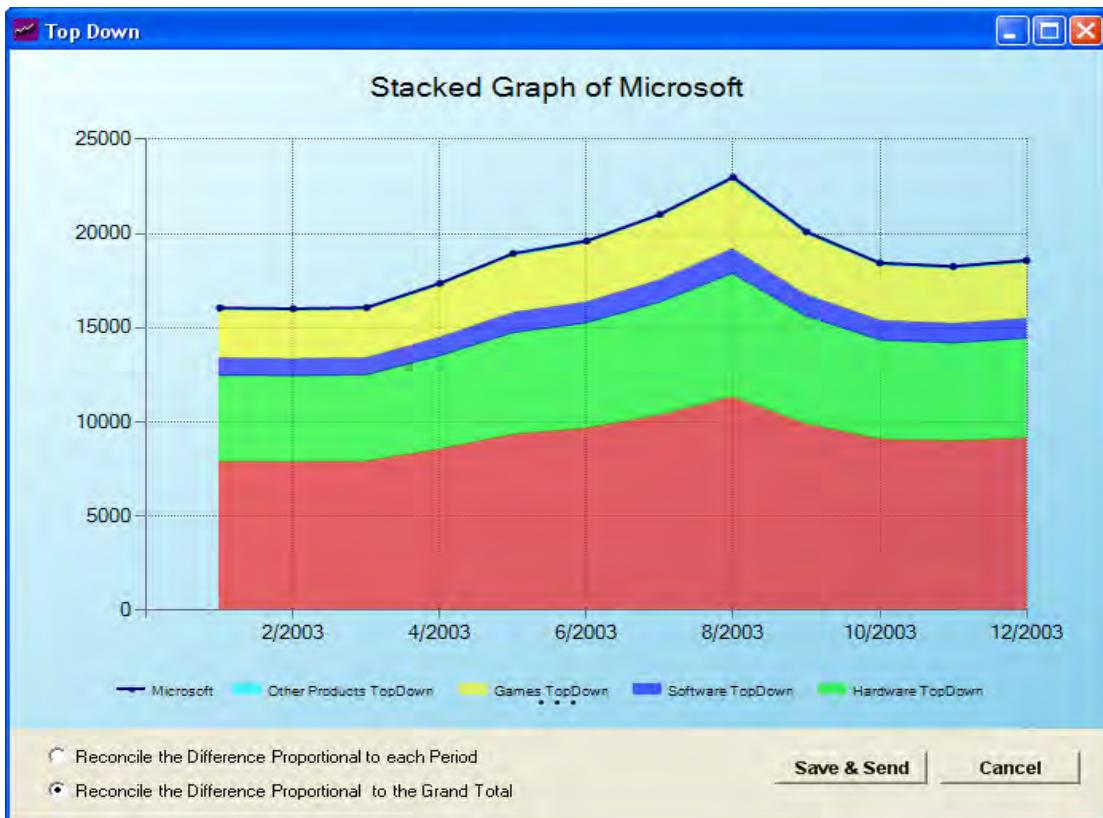


Figure 18: Top down forecasts

Once the new data is available, the user can monitor the error of the forecasts via the option 'Monitoring' and the screen shown in Figure 19 appears. Firstly the user selects the type of forecast to monitor among the available ones which are in the listbox (Statistical, Judgmental, Budget, Final, Top-Down-if available). Once a forecast is selected, "PYTHIA" calculates the statistics for this forecast. Those are several types of errors (Bias, MAE, MSE, RMSE, MAPE and Symmetric MAPE).

At the same time the graph is updated so as to present the randomness of the errors for the selected forecasts. The randomness is calculated by a method called "Trigg's Method". If the errors are into the green area ( $e_i \leq |0.5|$ ) then errors considered as random. Else, errors are not random, due to the pessimistic or optimistic attitude of the user and so the errors are continuously positive or negative. If this happens, program suggests the user to reforecast the data so that this attitude will be avoided in the future.



Figure 19: Monitoring

The option “Reporting” can generate complete and accurate reports in MS Excel Spreadsheet. This can be very useful for presentation of the results or further analysis and adjusting. The report form is separated in two parts, as can be seen in Figure 20.

On the left part, time series are shown up in a pyramid view. Here the user can select more than one of them. The produced MS Excel will provide information for all the selected time series. On the right part, user can select which of the statistics and information will be shown and saved in the spreadsheet. The available ones are Original Data, Deseasonalized Data, Trend-Cycles, Forecast Model, Error, Seasonal Indices, Statistical Forecasts, Judgemental Forecasts, Budget Forecasts, Top-Down Forecasts, Final Forecasts, Confidence Intervals, Chart, Method Parameters, Basic Statistical Values, ACF. The default selection includes Original Data, Deseasonalized Data, Statistical Forecasts, Chart and Basic Statistical Values. Using the combo box, user can select all the checkboxes at once, deselect them all or enable again the default selections.

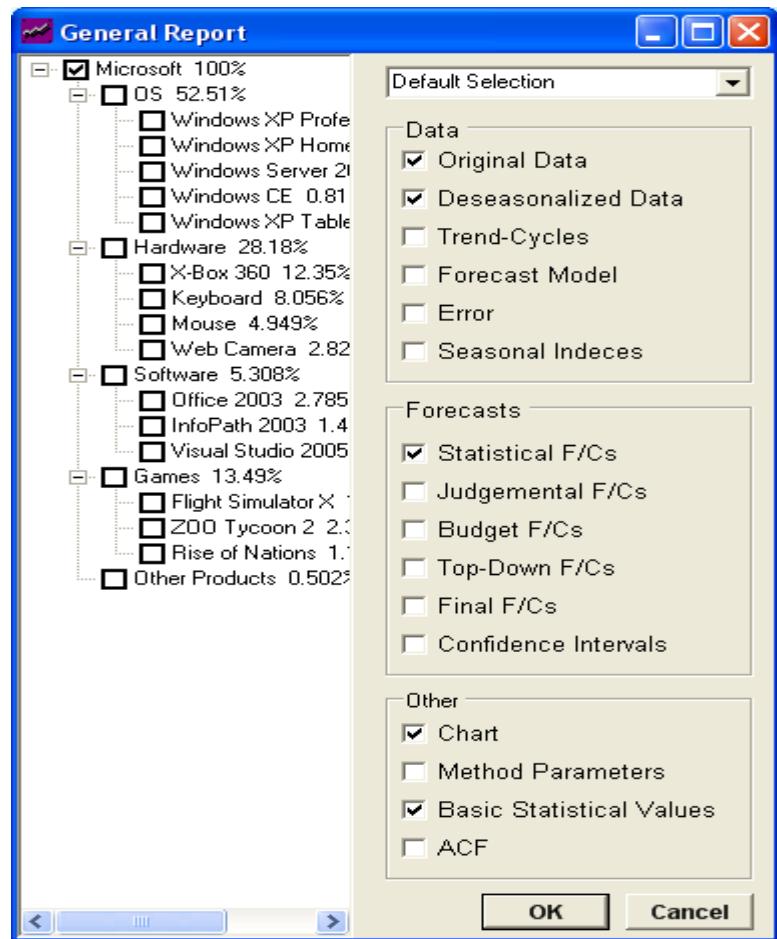


Figure 20: General Report

The module “Long Term” provides long-term forecasts. Long term forecasts have to take into account technological, environmental changes as well as changes in the markets and have to indicate mega-trends. The Main Screen of the module (Figure 21) shows the pyramid of the products (items) and the graphs of the selected categories – products (items). The system indicates automatically the most appropriate method for the long-term forecast.

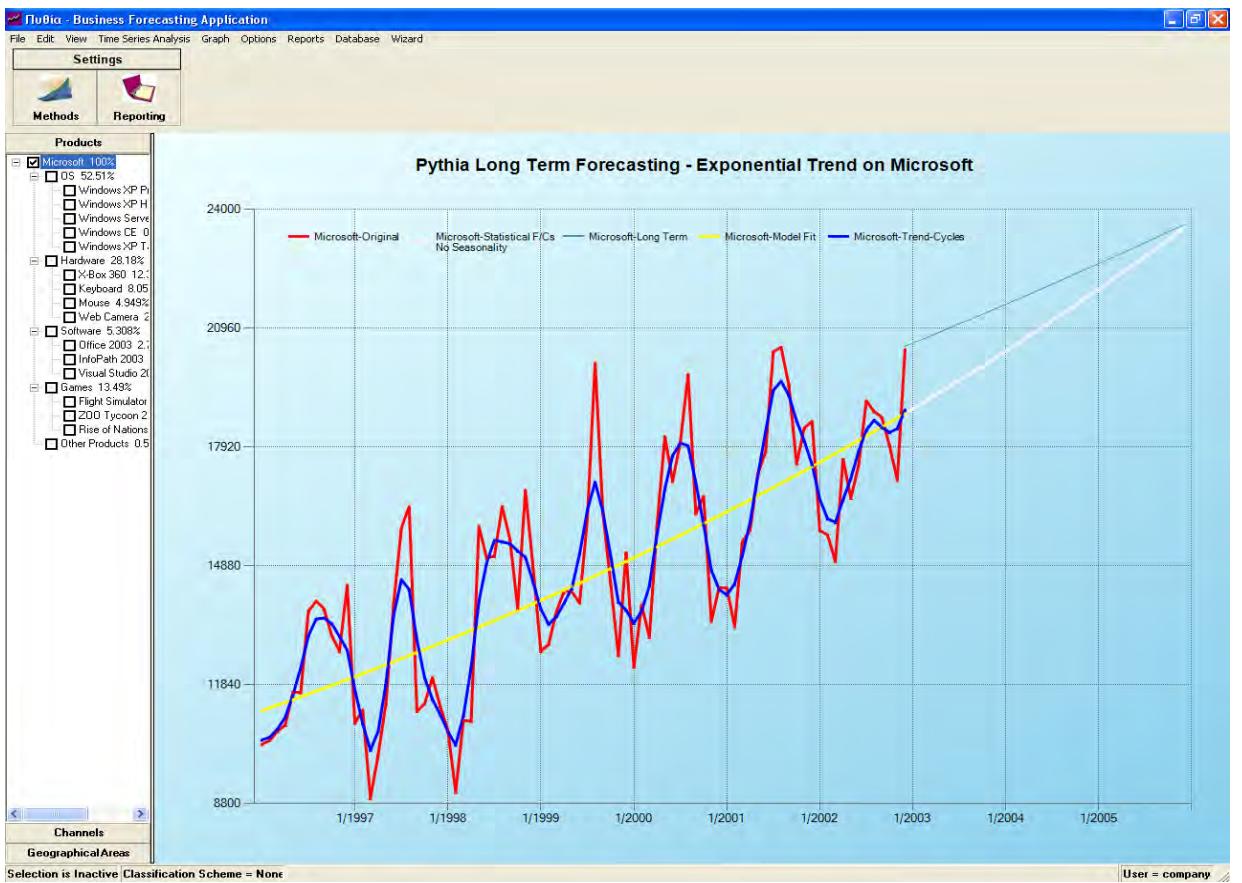


Figure 21: Long term forecasts

In the “Regression” section the user defines which variable will be the dependent, which variables will be independent and which of them will be excluded. There is a correlation matrix which shows the correlations between the variables that have been selected to contribute to the linear model. The model is being defined when the user presses the button “Calculate Regression Coefficients”. After the model is defined, the values of the coefficients in the regression model of each variable appear. The respective values of the standard error, the t-test and if the variable is significant or not (at Significance level 0.05) and the values of the  $R^2$  and the F-test value of the model are also displayed. The significant variables of the model according to the t-test are highlighted in red color in contrast to the insignificant which are highlighted in green color. All the above are shown in Figure 22:

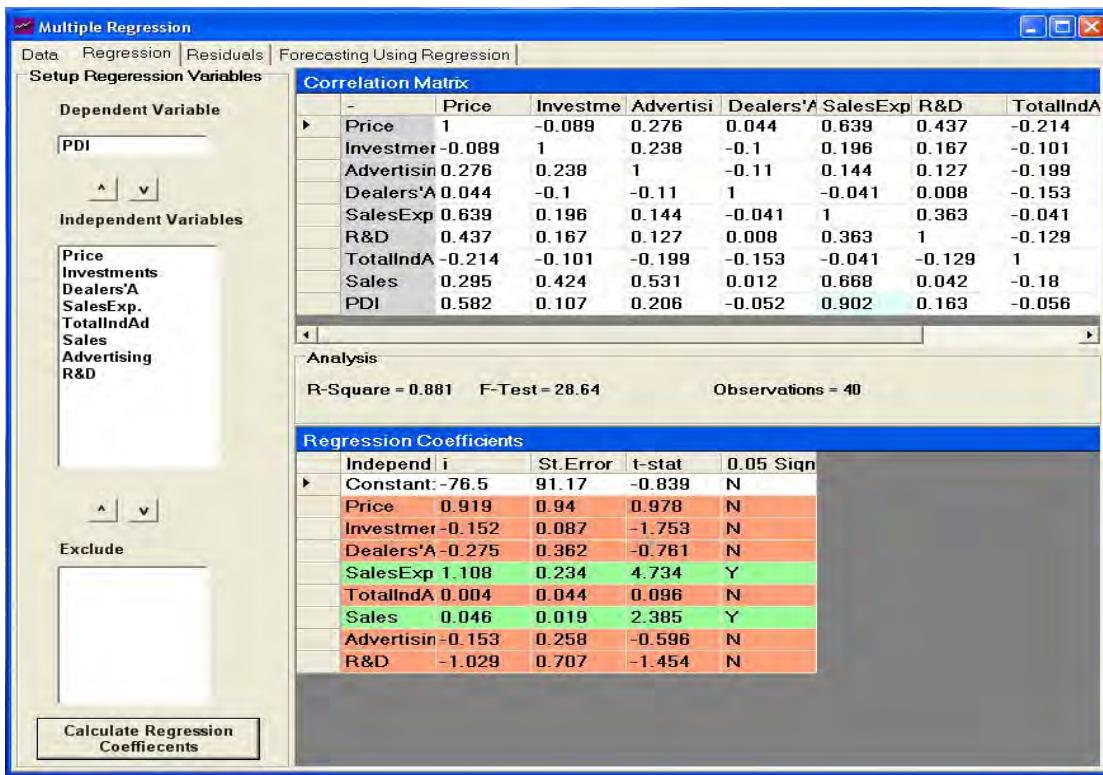


Figure 22: Regression

### 3. Case study

#### 3.1 Using PYTHIA at a Bank

As we said previously, PYTHIA can provide the forecasts required by practically all business situations demanding accurate predictions. Since the data we use in this paper comes from mortgages in UK, it would be very useful to a bank. A bank would be interested to forecast time series concerning mortgages such as the average value of loans, the number of loans, the value of loans in total, the average advance of loans etc. By taking into account this data, the bank could design its future products concerning loans, and be also able to make appropriate adjustments.

Within the bank, PYTHIA can be also used by financial directors at headquarters to set their budgets based on growth objectives as well as forecasting established trends by major product category, by country (in the case of multinational banks), by major customer or any other desired clarification. Country managers can prepare their own budgets (also based on the targets set by headquarters and local trends) and submit them to headquarters for approval, while product managers can prepare estimates for the sales of their own products.

#### 3.2 Forecasting Methods

The methods we used to provide the forecasts of the time series data are presented in following. These methods that are provided by PYTHIA, are:

- Theta model
- Linear Trend Exponential Smoothing (HOLT),
- Damped Trend Exponential Smoothing (DAMPED),
- Simple Exponential Smoothing with Multiplicative Seasonality (WINTERS)

### 3.2.1 The Theta Model

Theta model is a new univariate forecasting method introduced by Assimakopoulos and Nikolopoulos [2], [3]. The model is based on the concept of modifying the local curvatures of the time series. This change is obtained from a coefficient, called  $\theta$ -coefficient, which is applied directly to the second derivatives of the time series:

$$X_{\text{new}}^{''} = \theta \cdot X_{\text{data}}^{''} \quad (1)$$

If the local curvatures are gradually reduced then the time series is deflated. The smaller the value of the  $\theta$ -coefficient, the larger is the degree of deflation. In the extreme case where  $\theta=0$  the time series is transformed to a linear regression line. The progressive decrease of the fluctuations diminishes the absolute differences between successive momentary trends and is related, in qualitative terms, to the emerging of the data's long-term trends. To the opposite direction, if the local curvatures are increased ( $\theta>1$ ), then the time series is inflated. The larger the degree of inflation, the larger is the magnification of the short-term behavior.

Following both directions, a set of new time series, the so-called  $\theta$ -lines, are constructed. Their primary qualitative characteristic is the improvement of the approximation of the long-term behaviour of the data or the augmentation of the short-term features, depending on the value of the Theta coefficient. The placement of the lines in relation to the original data can be done in many different ways. If the fitting is an OLS estimation procedure then the mean and the slope of the  $\theta$ -lines remain the same compared to those of the original data.

The general formulation of the method becomes as follows: The initial time series is disintegrated into two or more  $\theta$ -lines. Each of the  $\theta$ -lines is extrapolated separately and the forecasts are simply combined. Any forecasting method can be used for the extrapolation of a  $\theta$ -line according to existing experience. A different combination of  $\theta$ -lines can be employed for each forecasting horizon.

The essence of the method is described in the following statement by Ord [7]:

In 'The Theta Model: A Decomposition approach to Forecasting' Vassilis Assimakopoulos and Kostas Nikolopoulos first remove any seasonal pattern by a classical decomposition method and then decompose the series into two parts. In essence, these two components may be viewed as a trend line to describe longer-term behaviour and a second difference for short-term movements. These components are then extrapolated separately and the elements are recombined to produce the final forecasts. [6]

Theta Model is one of the top performers of the M3 competition (particularly for monthly series and for microeconomic data), the bigger forecasting competition ever conducted [5], [7].

### 3.2.2 Exponential Smoothing Methods

Exponential smoothing is a forecasting technique that projects objects of the pattern of history data, as trends and seasonal cycles, to the future. The forecasts are computed after smoothing the data, in order to isolate the real patterns from the clearly random variations.

The main principle of the exponential smoothing methods is that the more recent the data is, the more information it contains. For that reason, greater significance is given to the most recent data, which declines exponentially.

#### HOLT

The Linear Trend Exponential Smoothing Model (HOLT) is an expansion of the regression model. However, in this model the initial point and the slope fluctuate in each time period through two parameters, since gradually more weight is given to the most recent data.

#### DAMPED

The Linear Trend Exponential Smoothing Model can be appropriately transformed in order to be also adjusted for non-linear trends (DAMPED). This is achieved with the use of one more, third parameter, which controls the rate that the forecast values increase. This parameter is called trend correction parameter.

### WINTERS

WINTERS is the combination of fixed level exponential smoothing model (simple) with Multiplicative Seasonality. Simple assumes the absence of trend from the data. That is to say that the time series is supposed to have a relatively stable average and the forecast results from the projection of a horizontal straight line. If the data has a seasonal pattern, then a seasonal index is used for each period of the year. The multiplicative seasonality means that in order to take the deseasonalized data we have to divide the actual data with the seasonal indexes.

### **3.3 Time-series data**

In following we point out some matters concerning the origin of the data provided by the Council of Mortgage Lenders (CML). Totals are estimates grossed up from the sample of lenders reporting to reflect total market size. All figures from April 2005 onwards are based on Product Sales Data reported to CML. Figures pre-April 2005 are taken from the Survey of Mortgage Lenders. Prior to 1992-Quarter2 figures are taken from the Building Societies 5% sample of mortgage completions. There are material differences in both the reporting methodologies and the sample of contributing lenders for the different surveys. Figures after April 2005 are not strictly comparable with those up to that point.

The data we found from CML included among others, time-series with number of loans and value of loans in total. In our case study, it would be rather more useful to a Bank to forecast average values of loans, so we divided the respective elements of the two time-series to produce average values. Specifically, we use three time-series to forecast average mortgage loan prices that are grouped in:

- Model 1: All buyers - average loan prices in UK,
- Model 2: First-time buyers - average loan prices in UK,
- Model 3: Home-movers - average loan prices in UK

Here we present in detail the case of 'Model 1: All buyers - average loan prices in UK', which includes the graphs and the statistics produced by each of the four forecast methods we mentioned above. The process and results are similar for the other Models too, but for the sake of space are not presented here. However, we include a table with forecasting errors of each method applied on each time-series, in order to evaluate the performance of the methods. All three time-series with historical data [11] are included in the Appendix of the paper.

In the following forecasts, PYTHIA automatically hides a number of observations in order to test the performance of the methods by comparing forecasts with historical data. The forecast horizon is again automatically chosen by PYTHIA and is 11 points in the future, but we could have also chosen another number of forecasts. Each forecast point corresponds to a quarter and so the actual forecasts we obtain are for the second, third and fourth quarter of 2007 and for all quarters in the years 2008 and 2009. In the graphs that follow, the red line represents the original time series, the blue line the trend-cycle series, the black line the statistical forecasts and the grey one the confidence intervals.

### **3.4 Model 1: All buyers – average loan prices in UK**

In following we present the results of each forecasting method applied in Model 1:

### 3.4.1 Theta

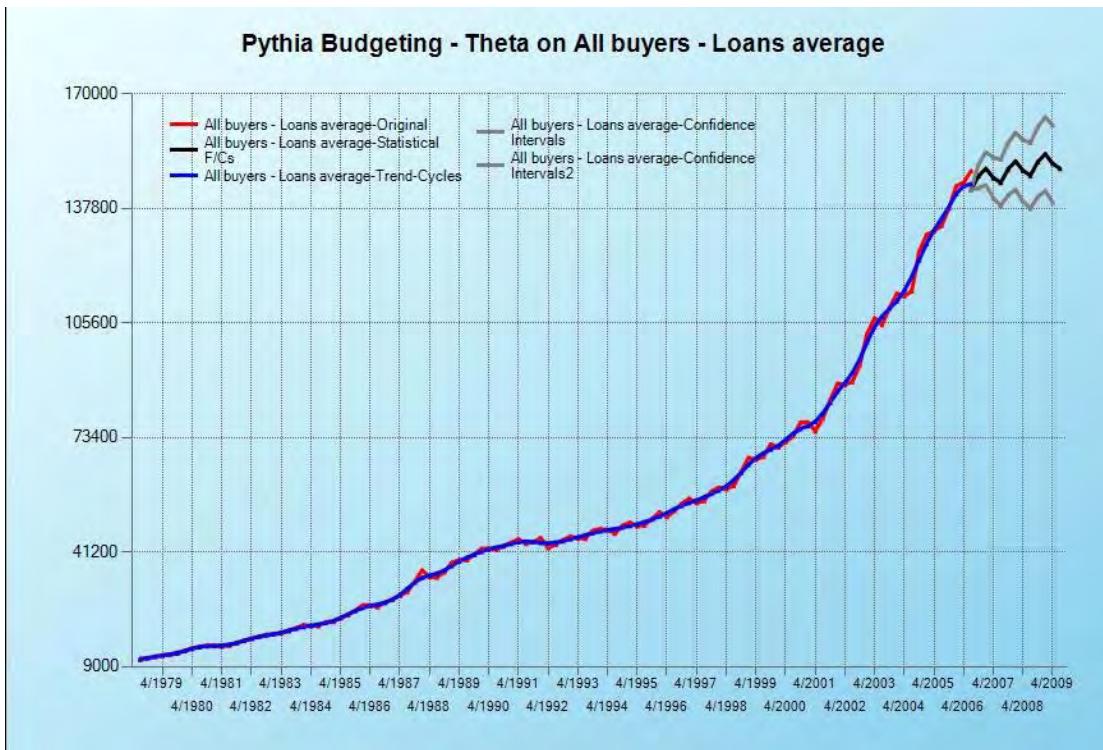


Figure 23: Graph of historical data and forecasts data using Theta Method

MEAN ABSOLUTE ERROR (MAE)	MEAN SQUARE ERROR (MSE)	MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)
1235	3957513	2,107%

Table 1: Statistical Errors using Theta Method

### 3.4.2 HOLT

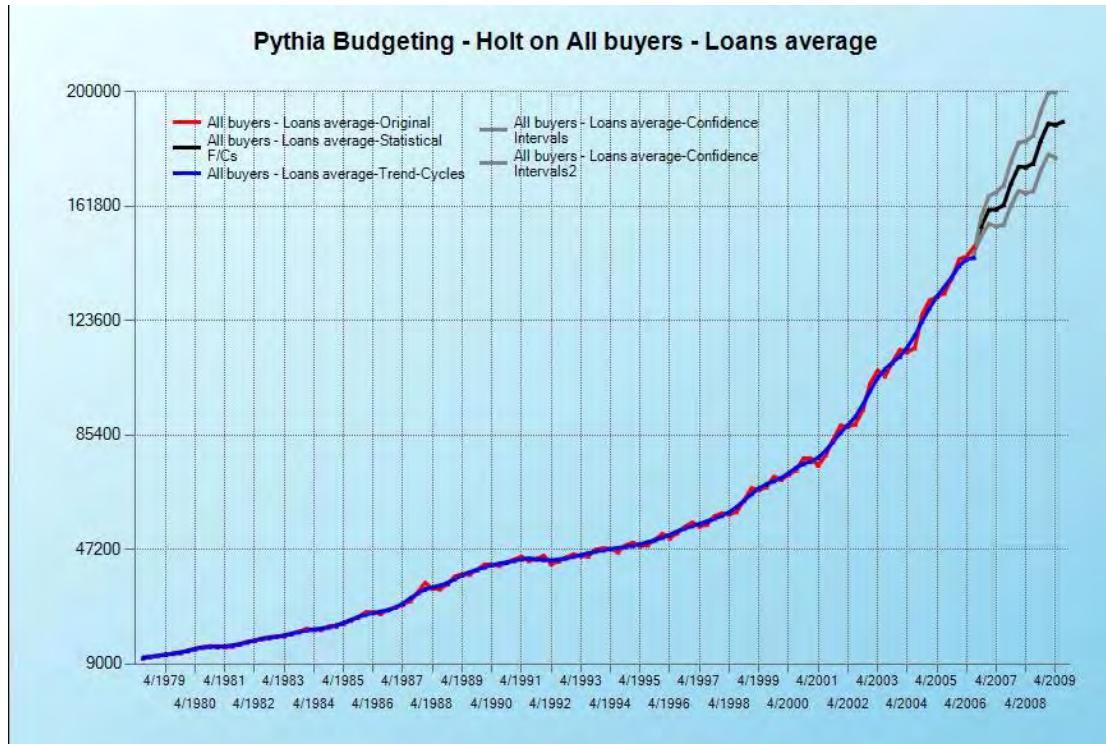


Figure 24: Graph of historical data and forecasts data using Holt Method

MEAN ABSOLUTE ERROR (MAE)	MEAN SQUARE ERROR (MSE)	MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)
1106	3970614	3,819%

Table 2: Statistical Errors using Holt Method

### 3.4.3 DAMPED

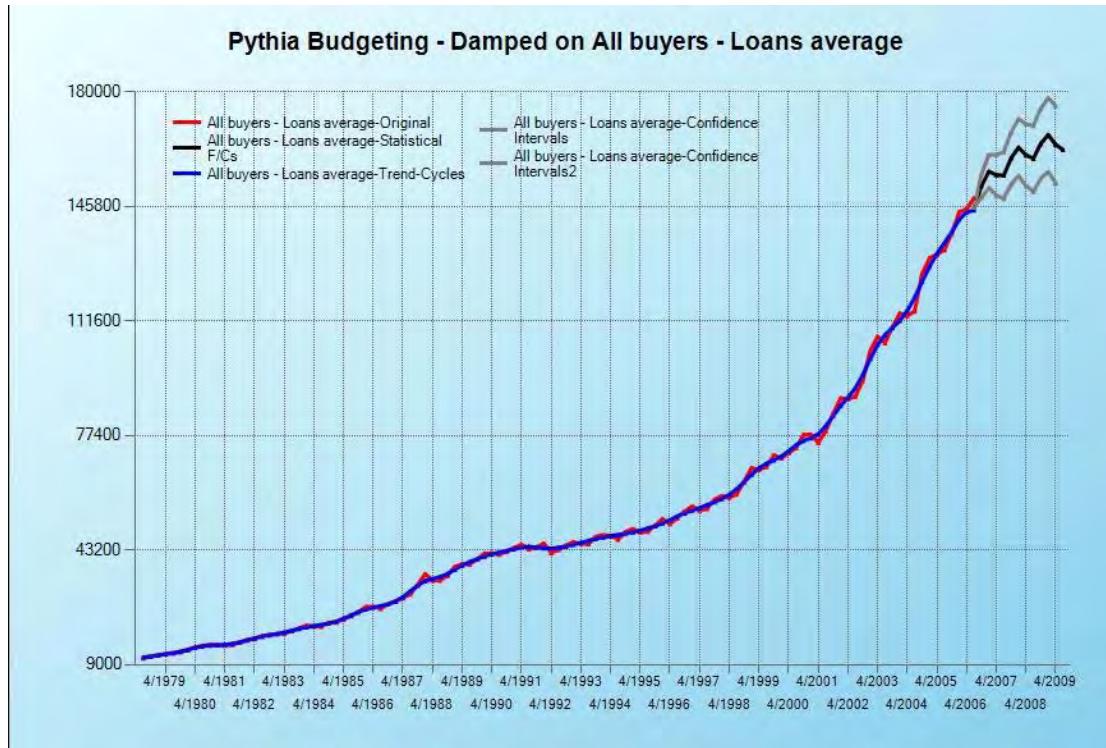


Figure 25: Graph of historical data and forecasts data using Damped Method

MEAN ABSOLUTE ERROR (MAE)	MEAN SQUARE ERROR (MSE)	MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)
1203	4471608	3,832%

Table 3: Statistical Errors using Damped Method

### 3.4.4 WINTERS

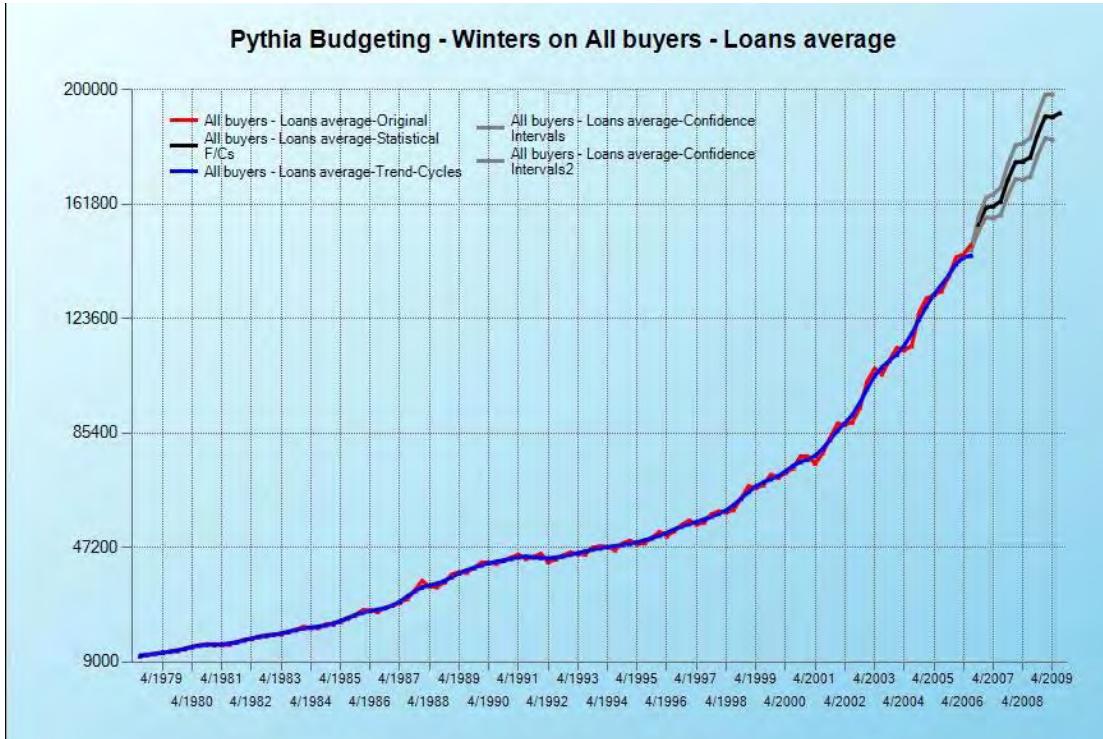


Figure 26: Graph of historical data and forecasts data using Winters Method

MEAN ABSOLUTE ERROR (MAE)	MEAN SQUARE ERROR (MSE)	MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)
890,5	1933250	1,769%

Table 4: Statistical Errors using Winters Method

### 3.5 Results

Other statistics that PYTHIA produces for Model 1 are shown in Table 5. PYTHIA also computes four seasonal indices, one for each quarter of the time-series that are presented in Table 6:

MEAN	MIN VALUE	MAX VALUE	R-SQUARE	MAD
52310	10999,15	148332,32	0,871	27204

Table 5: Time-series statistics of Model 1

SEASONAL INDEX 1	SEASONAL INDEX 2	SEASONAL INDEX 3	SEASONAL INDEX 4
98,3 %	100,5 %	101,7 %	99,5 %

Table 6: Seasonal indices of Model 1

The same procedure was followed for the rest time series too. The next table presents the mean absolute percentage errors (MAPE) of each forecasting method when applied to each Model of time series. The forecasting errors are major indicators of a good forecast and except from MAPE we could use MSE, MAE, Symmetric MAPE etc in order to compare the forecasting methods:

MEAN AVERAGE PERCENTAGE ERRORS (MAPE)			
	Model 1 All buyers - average loan prices in UK	Model 2 First-time buyers - average loan prices in UK	Model 3 Home-movers - average loan prices in UK
<b>Theta</b>	2,107 %	2,358 %	2,252 %
<b>Holt</b>	3,819 %	3,93 %	4,003 %
<b>Damped</b>	3,832 %	3,915 %	3,817%
<b>Winters</b>	1,769 %	2,148 %	1,982 %

Table 7: Competition of forecasting methods

From the table we can see that the best forecasts according to MAPE are produced by WINTERS and this is maybe due to the fact that there is seasonality in the data. The Theta Method comes next in the row and generally produces good forecasts with small MAPE. In order to tell with greater certainty which method produces the most accurate forecasts we could compare the rest error statistics provided by PYTHIA too.

### 3.6 Adjustments useful to the Bank

After the production of the statistical forecasts of the average mortgage loan values, the manager at the Bank can do adjustments if he thinks that the future values may differ from the forecasts. So, he is able to interfere and change the forecasts manually on the graph, thus producing judgmental forecasts. He can also alter the forecasts to a positive or negative percent, which provides budget forecasts. Then he can choose which of the above methods he wants to combine in order to produce the final forecasts.

Another useful functionality for the Bank is the bottom-up and top-down analysis that PYTHIA provides. PYTHIA lets the manager to make bottom-up predictions by adding the forecasts up to any desired level of aggregation. A top manager can then allocate the forecasts to each subcategory or division. In the case that the manager's decision is affected by the desired growth rate of the Bank or division, he can see through PYTHIA various growth rates related to the forecasts and change the forecasts accordingly. As the two forecasts (bottom-up and top-down) may differ, the top manager can reconcile the difference and allocate it appropriately to the various sub-aggregation levels.

PYTHIA can also identify Special Events and Actions (SEA) when an unexpected change in the level of data is observed in the time-series. A SEA for our case study could be a raise or a decline of interest rates, which affects loans and buyers. The manager can confirm an identified SEA, change the impact of a SEA in the selected period or change the period that happened. In another case, if the Bank wanted to forecast number of loans or sales of other products and services, it would probably find useful to adjust the working days or calculate the Bank Holidays in UK, which is also provided by PYTHIA.

## 4. Comparing PYTHIA with Theta Forecaster

The PYTHIA forecasting platform offers many more functionalities to the user than Theta Forecaster [1], [8], [9]. The main differences between the two forecasting tools are presented in following:

- New modules are incorporated in PYTHIA, which are those of 'Budget', 'Production / Inventory', 'Long Term', 'Estimating Relationships', 'Monitoring' and 'Wizard'.
- In addition, the modules of 'Data Analysis and Adjustments' and 'Reporting' that we also find in Theta Forecaster, have more features and better functionality in PYTHIA.
- PYTHIA gives the capability to the user to deal with missing values, zero values, working days and special events or actions.
- PYTHIA provides the user with more statistical information about the time series, such as many types of error statistics.
- In PYTHIA the user can choose to see more statistical lines in the main chart than in Theta Forecaster, e.g. the trend / cycle line.

- Theta Forecaster allows the combination of more forecasting methods than PYTHIA does, but PYTHIA allows the user to adjust the proportion that each method has in the averaging.
- PYTHIA allows the user to adjust the smoothing parameters of the forecasting methods. Moreover, PYTHIA automatically applies the appropriate forecasting method to the time-series inserted by the user.
- Theta Forecaster produces only statistical forecasts, whereas PYTHIA can also produce judgmental and budget forecasts. The user can then combine them to produce the final forecasts.
- PYTHIA lets the user to make bottom-up predictions by adding the forecasts up to any desired level of aggregation.
- Additionally, in PYTHIA after the person in charge produces the final forecasts he is able to allocate them to each category or division below. This can happen by reconciling the difference and allocating it appropriately to the various sub-aggregation levels.
- While Theta Forecaster can be used only for time-series forecasting, PYTHIA provides multiple regression too, through the module 'Estimating Relationships'.
- Finally, PYTHIA aims rather more to professionals than simple users and is more suitable for use in a business environment than Theta Forecaster.

## 5. Perspectives

For the years 2007 and 2008, mortgage and housing markets will probably remain resilient. With the wider economy continuing to provide a sound backdrop, the level of property sales in 2007 is expected to be broadly similar to 2006. But affordability will remain stretched. Asking prices for homes continue to be challenging for many first-time buyers, and movers will find it increasingly difficult to trade up as price differences between properties widen.

With interest rates rising in 2006, payment difficulties are likely to increase over the next year. Prospects for employment and income are positive, but higher interest rates look set to result in a temporary rise in short-term arrears later in 2007. Most borrowers will be able to resolve these problems by speaking with their lender early if they are in difficulty. The number of possessions is expected to be broadly stable over the years 2007 and 2008.

## 6. Conclusions

In this paper we show how the PYTHIA forecasting platform can be used by the managers of a Bank to forecast mortgage loan values. In fact, PYTHIA can provide the forecasts required by practically all business situations demanding accurate predictions. It is designed and developed with the purpose of making the task of managerial forecasting straightforward, user-friendly and practical. It incorporates a lot of knowledge and experience in the field of forecasting while fully utilizing new capabilities of computers and software. PYTHIA gives more capabilities to the user than Theta Forecaster and can be used by professionals that need a flexible tool to forecast, analyze and organize their data.

## 7. References

- [1] Assimakopoulos, V. and Nikolopoulos, K. (2001) Theta Intelligent Forecasting Information System, *International Symposium on Forecasting ISF 2001*, June 17-20, Callaway Gardens, Atlanta Georgia U.S.A.
- [2] Assimakopoulos, V., Nikolopoulos K. (2000) The theta model: a decomposition approach to forecasting, *International Journal of Forecasting*, Vol. 16, pp. 521 –530
- [3] Assimakopoulos, V. and Nikolopoulos, K. (1999) The Theta Model, *Decision Sciences Institute, 5th International Conference*, Proceedings, pp. 584-586, July 4-7, Athens, Greece.
- [4] Clemen, R. (1989) Combining forecasts: a review and annotated bibliography with discussion, *International Journal of Forecasting* 5, 559–584.

- [5] Makridakis, S. and Hibon, M. (2000) The M3-Competition: Results, conclusions and implications, *International Journal of Forecasting*, Vol. 16, pp. 451-476.
- [6] Nikolopoulos, K., Maris, K., Pantou, G., Metaxiotis, K. and Assimakopoulos, V. (2003) Forecasting Volatility with the Theta Model, *Journal of Empirical Economics Letters* 2 (6): 216-227.
- [7] Ord, K., Hibon, M., Makridakis, S. (2000). The M3-Competition, *International Journal of Forecasting*, Volume 16, Issue 4, October-December 2000, Pages 433-436
- [8] Pagourtzi, E., Assimakopoulos V. and Litsa A., (2007) Theta Model forecasts quarterly & monthly Dwelling Prices in UK, *Briefings in Real Estate Finance*, 2007; 1-30.
- [9] Pagourtzi, E., Assimakopoulos V. and Litsa A. (2006) Theta Model forecasts Real Estate values, *Bauhaus Universität Weimar, 13th Annual European Real Estate Society Conference*, Proceedings, June 7-10, Weimar, Germany.
- [10] Council of Mortgage Lenders (CML), “About us”,  
<http://www.cml.org.uk/cml/about/>
- [11] Council of Mortgage Lenders (CML), “Statistics”,  
<http://www.cml.org.uk/cml/statistics>

## Appendix

The table that follows contains the time series we used to produce the forecasts. In PYTHIA, each time series was inserted separately in Excel format, but here for convenience we present them as a whole. All prices are given in English pounds (£):

Name	Average loan prices			
Periods Per Year	4			
Seasonality	4			
Data				
Quarter	Year	All buyers	First-time buyers	Home-movers
1	1979	10.999	10.566	11.385
2	1979	11.504	10.995	11.894
3	1979	11.951	11.776	12.094
4	1979	12.256	11.975	12.492
1	1980	12.473	12.221	12.683
2	1980	12.789	12.534	12.996
3	1980	13.472	13.013	13.777
4	1980	14.298	13.843	14.649
1	1981	14.650	14.196	14.993
2	1981	15.088	14.475	15.530
3	1981	14.958	14.328	15.449
4	1981	14.787	14.016	15.543
1	1982	14.987	14.121	16.018
2	1982	15.731	14.708	16.863
3	1982	16.461	15.349	17.552
4	1982	16.906	15.440	18.255
1	1983	17.535	15.722	19.272
2	1983	18.035	16.412	19.637
3	1983	18.241	16.469	19.913
4	1983	18.419	16.950	19.837
1	1984	19.074	17.518	20.762
2	1984	19.878	18.419	21.238
3	1984	20.794	19.278	22.094
4	1984	20.489	19.045	21.906
1	1985	20.481	18.946	22.332
2	1985	21.634	20.123	23.263
3	1985	21.711	20.300	23.202
4	1985	22.653	21.268	24.119
1	1986	23.583	21.933	25.281
2	1986	24.769	23.147	26.342
3	1986	26.345	24.326	28.113
4	1986	26.364	24.465	28.197
1	1987	25.815	23.843	27.917
2	1987	27.097	25.045	29.026
3	1987	27.931	25.994	29.596
4	1987	28.971	26.886	30.900
1	1988	30.109	27.641	32.527
2	1988	32.948	30.015	35.477
3	1988	36.165	33.020	38.440
4	1988	34.223	30.613	37.592
1	1989	34.181	29.556	39.429
2	1989	35.649	32.143	39.340
3	1989	38.286	34.036	42.529
4	1989	39.084	35.865	42.620
1	1990	39.009	34.580	44.305
2	1990	40.483	36.697	44.638
3	1990	42.212	39.019	45.632
4	1990	42.325	38.882	46.047
1	1991	41.900	37.924	45.883
2	1991	42.905	38.726	46.910

3	1991	43.966	39.381	47.566
4	1991	44.888	39.352	49.493
1	1992	43.587	39.198	48.099
2	1992	44.047	39.218	48.317
3	1992	45.253	40.226	50.116
4	1992	42.447	36.936	48.917
1	1993	43.322	37.976	49.530
2	1993	44.658	39.837	50.239
3	1993	45.675	39.618	52.213
4	1993	45.139	39.114	51.547
1	1994	45.063	38.325	53.211
2	1994	47.271	40.493	55.059
3	1994	47.762	40.432	56.195
4	1994	47.568	40.903	55.363
1	1995	46.460	40.557	54.077
2	1995	48.668	42.072	56.270
3	1995	49.564	41.333	57.726
4	1995	48.515	41.602	55.535
1	1996	48.796	41.069	57.541
2	1996	50.583	43.900	56.411
3	1996	52.539	44.795	59.243
4	1996	51.149	43.866	58.437
1	1997	52.840	43.870	60.396
2	1997	54.841	46.024	61.862
3	1997	56.292	47.862	63.589
4	1997	55.076	47.283	62.057
1	1998	55.580	47.754	62.899
2	1998	58.318	51.123	65.459
3	1998	59.406	53.007	65.330
4	1998	58.968	51.271	65.920
1	1999	59.911	52.699	66.711
2	1999	63.736	57.695	68.987
3	1999	67.751	59.070	74.966
4	1999	67.246	58.956	74.314
1	2000	68.079	59.104	76.088
2	2000	71.565	61.516	79.444
3	2000	70.669	60.084	78.789
4	2000	72.193	60.854	80.418
1	2001	73.783	63.157	81.821
2	2001	77.667	66.091	85.911
3	2001	77.760	68.175	86.650
4	2001	75.294	69.711	82.021
1	2002	78.803	74.168	83.633
2	2002	83.889	79.054	88.149
3	2002	88.574	84.482	92.469
4	2002	88.336	85.075	91.237
1	2003	89.136	83.913	90.624
2	2003	93.856	85.390	96.079
3	2003	102.676	91.050	105.836
4	2003	106.974	95.205	110.090
1	2004	105.089	93.467	109.355
2	2004	109.840	101.217	113.115
3	2004	113.897	105.365	117.541
4	2004	113.164	105.039	117.438
1	2005	114.533	106.559	118.879
2	2005	125.690	108.053	136.588
3	2005	130.399	111.016	141.890
4	2005	131.514	113.096	142.689
1	2006	132.854	114.321	144.325
2	2006	137.855	119.789	148.008
3	2006	144.108	123.736	155.226
4	2006	145.033	125.792	155.268

1	2007	148.332	127.395	160.035
---	------	---------	---------	---------