Dirección y Organización 65 (2018) 20-35

www.revistadyo.com

Modelling performance management measures through statistics and system dynamics-based simulation

Hanzel Grillo*, Francisco Campuzano-Bolarin**, Josefa Mula*

Recibido: 25 de Enero de 2018 / Aceptado: 18 de Febrero de 2018

Abstract

The objective of this paper is to establish a methodology that combines performance measurement, a statistical record of measures to identify any relations among them, and system dynamics-based simulation modeling with the aim of supporting operations decision systems. This methodology intends to provide the comprehensive analysis of performance in such a way that it also analyzes the sensitivity and optimization of certain metrics according to requirements in each case. In the literature, this appears as a poorly developed research area. Some relevant studies have been identified which have attempted this combination, but have not completely established it.

Keywords

Performance management measures; simulation; system dynamics; statistics.

1 Introduction

Measuring performance is a relevant matter of widespread use among researchers and practitioners. Otto and Kotzab (2003) present a review of suitable metrics to measure the effectiveness of supply chain management from six perspectives: system dynamics, operations research/information technology, logistics, marketing, organization and strategy. Tung et al. (2011) analyze the suitability of the multidimensional control and classification of these measures. Other works, like Burgess (1998), Kleijnen and Smits (2003), Barnabè (2011) and Bianchi (2012), have gone one step forward and designed measures defined by system dynamics, which is the modeling approach used in this paper. Cai et al. (2009), among others, propose a framework that uses a systematic approach to improve the accomplishment of key performance indicators (KPIs) in a supply chain context.

This proposal is about handling an initial set of measures whose record suffices to ensure statistical representativeness, with which it is possible to identify correlative performances. In the final set of measures, those measures that are particularly interesting for managers must be identified, which are classified as such. Strategic measures must be explained according to the remaining set of measures classified as basic or input measures. A multiple linear regression analysis is used to interpret the relations between input measures and their effect on each strategic measure. To the strategic measures, it is also possible to add a third group of measures, known as derivative measures, which are not obtained through linear regression equations, but as formulae predefined by the decision maker (DM), and can involve both basic and other strategic measures. Lastly, a fourth set of metrics, known as ratios, is established to measure the ratios of change among the strategic measures in terms of time or any input measure. Modeling these groups of measures can be done by system dynamics, and in such a way that a sensitivity analysis of the strategic measures can be done with the changes made in the input measures, or by predictive studies on the optimum values of the input measures to accomplish the levels expected in the strategic measure. To illustrate the proposed methodology, a generic analysis of the operations in a supermarket of retail product sales was done using the ratios among the performance measures that resulted from this operation.

The remainder of the paper is organized as follows: Section 2 presents a literature review where the proposed methodology is compared to alternative approaches. Section 3 develops the proposal of a methodology based on statistics and system dynamics for modeling KPIs. Section 4 illustrates this methodology in a supermarket application. Section 5 presents the conclusions drawn and further research.

2 Literature review

Nudurupati et al. (2011) thoroughly reviewed the progress made in measuring performance from the management information system concepts viewpoint and development toward performance measurement systems (PMS). Gunasekaran et al. (2004) did a literature review of the different

Hanzel Grillo*
hangries@upvnet.upv.es

Francisco Campuzano-Bolarin**
francisco.campuzano@upct.es

Josefa Mula*
fmula@cigip.upv.es

*CIGIP (Research Centre on Production Management and Engineering), Universitat Politècnica de València, Plaza Ferrándiz y Carbonell, 2, 03801, Alcoy, Alicante, Spain
**Business Management Department, Universidad Politécnica de Cartagena, Campus Muralla del Mar, s/n, 30202, Cartagena, Murcia, Spain

metrics that must be used to measure performance in supply chain settings but does not carry out statistical or simulation analyses. Sousa (2004) reviewed various performance measurement techniques in export companies in detail, and analyzed the empirical research conducted. Despite the variety of the statistical analysis employed, it mentioned no study that has combined system dynamics and simulation. Kleijnen and Smits (2003) provided details of applying simulation and system dynamics in performance measurement, and evidenced how system dynamics, and simulation in general, can be used as a means to evaluate performance.

Alfaro-Saiz et al. (2007) developed a PMS for company chains based on three levels: network enterprise, supply chain and individual. Akkermans and van Oorschot (2004) conducted studies on balanced scorecards using system dynamics to evaluate scenarios with conflicting objectives. Angerhofer and Angelides (2006) considered a performance measurement model for supply chains in collaborative settings by analyzing the relations among the metrics with non linear differential equations and ratios of change and developed their model through system dynamics. This study was the most similar one to that considered herein, but the difference lies in our analysis being based on correlative interactions modeled by multiple regression and not being specific for collaborative supply chain environments. Verdecho et al. (2012) also analyzed performance measurement in collaborative settings based on the analytical network process (ANP), where performance measurement was analyzed from the collaborative supply chains perspective in which inter-company relations had to be taken into account as factors to be measured. Boj et al. (2014) also resorted to the ANP methodology, but they used it in this particular case to measure intangible assets and intellectual capital to relate them with the strategic objectives defined in the performance measurement system, which is normally the balanced scorecard.

Cai et al. (2009) presented a study which analyzed interdependences between performance measurement measures and analyzed the iterative cost of achieving them. Mora-Monge et al. (2006) did a multiple regression analysis of performance measures in advanced manufacturing technology based on the sampling results taken from a survey that evaluated North and Central American companies. Finally, Santos et al. (2002) did a performance measurement study based on system dynamics and a multi-criteria analysis. This study is similar to our proposal in that it applied a multi-criteria approach to evaluate tendencies among measures that could come into conflict with each other, and it was possible to perform an optimization analysis. The difference lies in the fact that these authors did not analyze the relations among the defined measures, based on statistical correlations and linear regression analyses.

Of all the works that we reviewed which came closer to the proposed methodology, we found that by Jusoh et al. (2008), who conducted a study about measuring performance. This work was based on statistically analyzing the correlation of a set of multiple measures. Although this study is similar to our proposal, the main difference lies in our work using linear regression, based on a record of the measure itself. System dynamics-based simulation was also done with possible sensitivity and optimization analysis approaches. Rodriguez et al. (2009) conducted a study that took measures of a previously established MIS, which were statistically analyzed to identify the relations among them. The methodology consisted in identifying the relations among the measures and then to make forecasts of them. Although this study used a similar methodology to ours, it neither performed a direct linear regression analysis to define relations among measures nor made any type of simulation or establish system dynamics as part of the methodology. Therefore, it did not establish optimum levels or a sensitivity analysis. Finally, Rodriguez et al. (2010) extended the work of Rodríguez et al (2009), based on a PCA to identify the relations between the measures in an already existing PMS. The difference with this study is that it included a scenario analysis, which mainly aimed to make future forecasts of the levels of the measures that could be influenced by previously identified external factors. This methodology approaches our sensitivity analysis proposal, but the difference is that it used a PCA. No measure was seen as an optimization objective, nor was system dynamics used to simulate these measures.

3 Methodology for modeling performance management measures

This methodology attempts to determine the impact that each measure's variations may have on the set of measures by describing their statistical performance and, based on this, establishing a system dynamic analysis that allows organizations to suitably re-adjust decision making in operations management. The first step involved is to delimit the problem or the scope of the simulation. Correct prioritization will enable more rigorous controls to be made of measures that mainly influence overall performance.

With the analysis done of the KPIs, the intention was to establish a system dynamics-based simulation model that interprets the generic performance. Here the specific objectives to pursue with such a model were to: Determine the effects that imply variations in the results of the different measures; that is, statistically describe the relation between several monitored measurements; and provide managers with a rapid swift consultation tool for strategic decision making which allows a sensitivity analysis of the model to be done to set the best levels for the values of the variables that ensure good results in cost terms, and without sacrificing a good level of customer satisfaction.

Of all the so-called strategic measures, those which based on the DM's experience can be classified as dependent variables; that is, those that give the level of important in control terms which the DM wishes to confer them can be calculated as the result of the interaction with other measures, which will be the independent variables, known as the basic measures in the proposed methodology. Next a multiple linear regression analysis is done (for example, see Draper and Smith (1998) and Cohen et al (2013)) for each one in accordance with the stepwise method in the IBM SPSS Statistics 21 software (Field, 2013). This approach aims to select from the independent variables those that mainly describe the variance of the dependent variable with no redundancy or collinearity among them (Chong and Jun, 2005). After checking the appropriateness and validity of the calculation model for each defined dependent variable, it is also necessary to check a series of necessary premises for the study to be robust in statistical terms. Among the minimum requirements evaluated in this study we find:

- Explaining the variance of the dependent variable. The R2 value explains the percentage of variance of the considered measure. As a general criterion, R2 must come as close to 1 as possible because this scenario will mean that variance will be totally explained. A value of 0.6 or above is considered acceptable.

- Statistical independence of observations. The set of historical data of each measure must show statistical independence. For this purpose, the value by Durbin and Watson (1950) test must be over 0.5 for all the calculation models of the dependent variables.

- Level of significance. It measures the effect of the total set of variables selected for the stepwise method on each dependent variable. This test is done in SPSS with an ANOVA. Records must take a value below 0.05 to be considered acceptable.

- Existence of influential observations. The distance test by Cook (1977) is used and the value of influence is centered (Hazewinkel, 2001). Both test statistics need to take a value below 1.

- Normality in residuals. For the model to be acceptable, these residuals must show a tendency that fits normal tendency. This verification is done by the Kolmogorov-Smirnof test (Hazewinkel, 2001) to fit normal probability distribution. The statistical K-S value must be above 0.05 to accept normality in residuals.

- Homoscedasticity of residuals. It indicates that the error variance of residuals remains constant in measurements and is checked graphically in this proposal. When we graph the predicted dependent variable against its respective residual value in a dispersion graph, the values must be uniformly distributed with no marked tendencies.

Also, we considered the possibility of including variables that can be calculated using the values of the initial measures dubbed as derivative measures. In this way, a derivative measure can be obtained from the calculation made by the formula predefined by the DM. Finally, using a fourth group of metrics, called ratios of change, is considered, whose function consists in relating the ratio of change between the previously defined measures.

The methodology used for the modeling and analyses in this research work was system dynamics (Forrester, 1961; Sterman, 2000). For further information about supply chain system dynamics-based simulation, we refer readers to Campuzano and Mula (2011) and Mula et al. (2013). The sensitivity analysis test aims to explore the performance of the model when submitted to extreme and atypical situations or conditions. The proposed simulation software is Vensim DSS®. For the evalution of improvements test, the optimization analysis of this simulation software was used. The final part of the developed methodology, preparing a scorecard interface based on the systems dynamics model is proposed. Figure 1 summarizes the methodology proposed in this section.

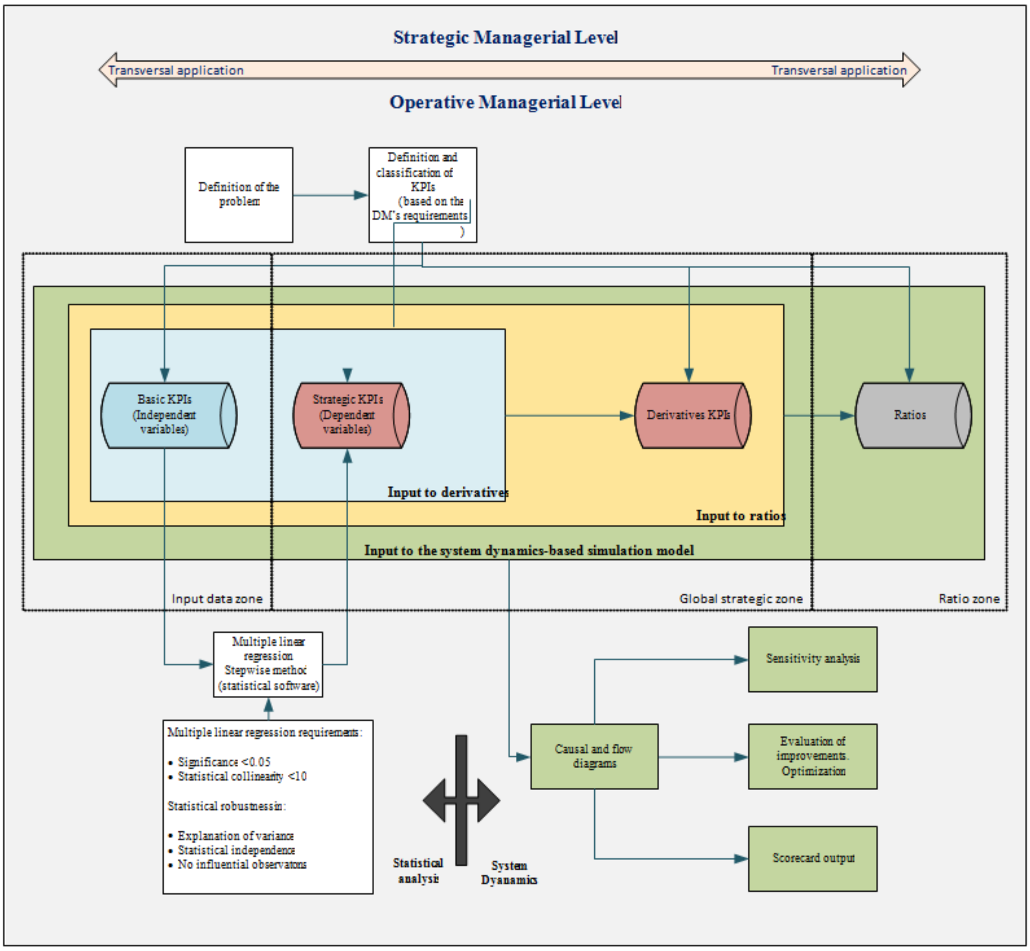


Figure 1 Methodology for modeling KPIs.

4 The supermarket application

This section provides an example of applying the proposed methodology using the input data of the operations in a supermarket. In order to delimite the problem, the internal operation is studied because it deals with third parties, e.g. suppliers and transport organizations, which normally adapt quite well to the organization's requirements if we assume that, given the level of demand involved, the supplier is willing to adapt to requirements and conditions. This study centers on the operational plan: a single store. Additionally, some very useful strategical measures exist for top management decision making, which can and must be measured at the operational level to then make an aggregate strategic decision. For instance, the number of staff members that manage a single supermarket is extremely important when it comes to performing aggregation for a supermarket chain, or for forecasting either new supermarkets or a group of supermarkets in the future. For this reason, the quantity of human resources is a purely strategic measure in the proposed methodology. Here everything measured at the operational level can be measured at the strategic level, which is why the name strategic has been given to the high value measures by the DM.

Selecting KPIs implies taking a set of variables and following a priority criterion, and choosing those with a higher interaction between them, and those with the strongest impact on the finally expected store performance. Table 1 provides a list of the measures that are usually controlled in stores of the retail trade. Here we find that the set of measures is divided into the groups that the proposed methodology required. The first group includes the so-called basic measures, which sporadically provide specific information of each activity and/or measure referred to, are static and are used as a basis to calculate other strategic measures. The second group contains the strategic measures defined in this category in accordance with their importance for the DM, along with the strategic and operational decisions that may be made based on them. A derivative measure is also included, whose formula is predefined by the DM. Finally, the ratios to be contemplated in the study are identified. To conduct the present study, we used a set of real store operation data that respond to the basic and strategic measures found in Table 1, measured monthly, and a record with 60 periods. This record is the initial KPI data set.

Table 1 . List of KPIs. Definition of influence.

Type	Code	Measure/indicator	Description	Order	Variable Type
Basic	IN1	Average charging rate (items/min)	Speed of the cashier charging customers	2	Auxiliary
	IN2	Average customer count	Total average of customers served	2	Auxiliary
	IN3	Building area (m2)	Total building area	2	Auxiliary
	IN4	Customer traffic	Measuring customers concentration per trade area	2	Auxiliary
	IN5	Inventory turnover	Monthly rotation of inventory	2	Auxiliary
	IN6	Staff turnover	Measure of hiring-firing activity	2	Auxiliary
	IN7	Total items	Total items registered in cash registers	2	Auxiliary
	IN8	Total sales (€)	Total store sales	2	Auxiliary
	IN9	Trade area (m2)	Area where exhibition and sale take place	2	Auxiliary
Strategic	IN10	Cash registers (c, r)	Cash registers installed in the store	1	Auxiliary
	IN11	Customer satisfaction	Customer satisfaction measured with a poll	1	Auxiliary
	IN12	Store presentation	Store presentation measured with a poll	1	Auxiliary
	IN13	Total staff	Total people working on the store	1	Auxiliary
	IN14	Total staff hours	Total staff paid hours	1	Auxiliary
Derivatives	IN15	Cashier staff	Total people working specifically at chash registers	1	Auxiliary
Ratio	IN16	Items per customer (Items/customer)	Average items per customer	2	Flow
	IN17	Items per staff hour (Item/hour)	Items sold per paid hour of staff	2	Flow
	IN18	Sales per cash register (€/c.r)	Sales per cash register	2	Flow
	IN19	Sales per staff hour (€/hour)	Sales per paid working hour	2	Flow
	IN20	Sales per trade area (€/m2)	Sales per trade area	2	Flow

Regarding to the statistical analysis of the measures, the strategic measures in Table 1 (from IN10 to IN14) are classified as the dependent variables for the multiple linear regression analysis. In this way, the following dependent variables are defined: Cash registers (c.r.), Store presentation, Total staff, Customer satisfaction and Total staff hours. Table 2 summarizes the results of the multiple linear regression analysis with the stepwise method for all the dependent variables. Column B offers the final coefficients of all the required variables for the equation to calculate the dependent variables. Eq. (1) to (5) show the result of combining the coefficients in Column B. The Significance column in Table 2 provides the significance values of each variable for each dependent variable. Significance must take a value below 0.05 to assume that the considered variable has a significant effect on explaining the variance of the dependent variable. We can see that all the significance values for each considered variable fulfill this condition.

Cash registers (c.r)

$$IN10=3.204+0.092\cdot IN13+0.005\cdot IN9-5.217\cdot 10^{-7}\cdot IN8 \quad (1)$$

Customer satisfaction

$$IN11=1.757-0.093\cdot IN10-0.013\cdot IN1-1.45\cdot 10^{-7}\cdot IN7+6.23\cdot 10^{-7}\cdot IN8-1.23\cdot 10^{-2}\cdot IN2+5.13\cdot 10^{-2}\cdot IN14 \quad (2)$$

Store presentation

$$IN12=3.012-0.016\cdot IN1-0.002\cdot IN3-0.011\cdot IN4+7.78\cdot 10^{-7}\cdot IN8+7.66\cdot 10^{-5}\cdot IN14 \quad (3)$$

Total staff

$$IN13=15.813-1.47\cdot 10^{-5}\cdot IN8+3.94\cdot 10^{-4}\cdot IN2-33.28\cdot IN6-0.962\cdot IN5+2.7\cdot 10^{-6}\cdot IN7 \quad (4)$$

Total staff hours

$$IN14=349.41+191.11\cdot IN13 \quad (5)$$

Finally, the Statistical collinearity-VIF (variance inflation factor) value column presents the inflated variance factor, a statistic that measures collinearity among independent variables which, according to Kutner et al. (2004), must take a value below 10. As we can see, all the independent variables fulfill the condition. Statistical evidence was obtained for the equation's representativeness to calculate each dependent variable, and for lack of collinearities that confers redundant information or can distort the end value. It is noteworthy that eq. (1) to (5) evidence the interaction of the 9 basic measures in Table 1 with all the strategic measures. We can even observe how a dependent variable can play the role of an independent variable in the equation of another variable, provided the collinearity principle is respected. The measures that only appear as independent variables are the data which will have to be supplied as input variables in the system dynamics model in order to obtain the dependent variables values by eq. (1) to (5). In this way, it is possible to run a sensitivity analysis to first evaluate the model's performance when faced with possible brusque changes in any of the basic measures, then to optimize any strategic measure to obtain a given level or value in the objective, and to observe the values required in the basic measures or in the other strategic measures to establish them as performance targets. It will be necessary to follow the same procedure if the measures do not represent a single store, but the performance of a group of them. In this case, the only aspect that might vary are eq. (1) to (5). For this reason, the methodology can be applied cross-sectionally, as seen in Figure 1, at the strategic, managerial and operational levels.

Table 2 .Results for the multiple regression stepwise method (IBM SPSS Statistics 21).

Dependent Variable-DV	Independent variables-IV (Calculation model of DV)	B	Significance (Criteria < 0.05)	Statistical collinearity – VIF value (Criteria < 10)
Cash registers (c.r)	Constant	3.204	0	
	Total staff	0.092	0	1.77
	Trade area	0.005	3.00E-06	1.02
	Total sales	-5.217E-07	8.02E-03	1.76
Customer satisfaction	Constant	1.757	0	
	Cash registers (c.r)	-0.093	7.00E-03	4.24
	Average charging rate	-0.013	3.00E-06	1.64
	Total Items	-1.45E-07	3.48E-03	1.83
	Total sales	6.23E-07	1.00E-06	4.43
	Average customer count	-1.23E-02	1.30E-05	3.64
	Total staff hours	5.13E-02	9.92E-04	3.15
Store presentation	Constant	3.012	1.92E-16	
	Average charging rate	-0.016	6.52E-06	1.53
	Building area	-0.002	5.21E-10	1.43
	Customer traffic	-0.011	4.26E-10	2.33
	Total sales	7.78E-07	2.48E-08	2.75
	Total staff hours	7.66E-05	1.67E-04	2.92
Total staff	Constant	15.813	3.17E-18	
	Total sales	-1.47E-05	1.22E-08	1.84
	Average customer count	3.94E-04	6.95E-10	1.73
	Staff turnover	-33.28	5.96E-03	1.86
	Inventory turnover	-0.962	4.83E-03	1.41
	Total Items	2.70E-06	4.86E-02	1.64
Total staff hours	Constant	349.41	2.83E-01	
	Total staff	191.11	0.00E+00	1.00

Table 3 shows the results of the statistical tests that explain dependent variable variance, the statistical independence of the observations, level of significance, existence of any influential observations, normality in residuals and homoscedasticity in residuals for each calculation model that results from the multiple linear regression done of the dependent variables. As observed, the results indicate that all the models comply with the statistical assumptions required for the multiple linear regression analysis to be suitable.

Table 3 Statistical test results.

Test	Objective	Criteria	Cash registers (c.r)	Customer satisfaction	Store presentation	Total staff	Total staff hours
R2	Explanation of the model's variance	Near to 1	0.897	0.895	0.883	0.825	0.694
Durbin Watson	Statistical independence of the observations	> 0.5	1.061	0.842	1.310	1.037	1.972
Significance level	Influence level of the independent variables over the dependent one	< 0.05	0.000	0.000	0.000	0.000	0.000
Cook distance	Existence of influential observations	< 1	0.292	0.217	0.388	0.143	0.102
Value of influence centered	Existence of influential observations	< 1	0.149	0.294	0.282	0.248	0.051
Z de Kolmogorov-Smirnov	Normal adjust of the residual values	> 0.05	0.477	0.573	1.000	0.971	0.869
Dispersion plot of the residuals	Homocedasticity of the residual values	Dispersed plot	ok	ok	ok	ok	ok

In this case, a single derivative measure is included (see Table 1), called Cashier staff, Eq. (6), which represents the staff members who operate all the Cash registers (c.r) in a supermarket. In this equation, additional measures IN21, IN22 and IN23 are fixed data that the DM must previously define according to the organization's policy about each one. As a general rule, measures IN21 to IN23 can be added to the group of basic measures since they are set input data and are involved in the derivative measure calculation.

$$Cashier\ staff\ (IN15) = \frac{\left[\frac{IN7}{60 * IN1} + IN2 * IN21\right]}{IN22 * IN23} \tag{6}$$

Regarding to the system dynamics-based simulation model, Table 1 offers the classification made of the level, flow and auxiliary variables. Figure 2 depicts the causal diagram of the proposed model. To validate the simulation model, the following tests were considered (Sterman, 2000): test to reproduce known performances, sensitivity analysis and evaluation of improvements. The test to reproduce known performances was done by simulating the 60 periods and by taking the 60 real data records we had available as the input data for Total sales (€), Average customers count and Total items. Table 4 shows the summary of the input data used to validate the model. With these data, the validation process consisted in running the model in Vensim DSS® and extracting the values that were forecast during the 60 monthly periods for the five variables calculated by linear regression. The difference between the forecast data and the record value defines the general error with which the model operates. The general error average was expected to be around 0%, and this error progressively lowered as we advanced in all 60 forecasts. Note that for the DM, an estimation with an error below 5% is quite good for this measurement type in this specific business kind (supermarkets). Table 5 shows the validation results for all five strategic variables calculated by linear regression. We can see how the general tendency of the errors throughout the consecutive executions during each period diminished and gave values close to 0%. We can also see how the general error average came close to 0%.

Table 6 provides the original values of each variable, along with the minimum and maximum values used for the sensitivity analysis. As we can see, a sufficiently wide range is considered for each measure for the atypical and extreme scenarios to be included.

For the sensitivity analysis five hundred simulations are made. In each simulation, the measures found in Table 6 may vary between the minimum value and the maximum value, with a uniform aleatory probability distribution. The results are measured using the dependent and derivative variables, and the obtained results are shown in Figure 3. We can observe how the remaining data constitute the extreme situations that the model can adopt if parameters vary considerably. For example, the average charging rate is a measure of the speed at which cashiers pass articles over scanners at cash registers. This value depends on each store and the mean at which staff can charge customers. Hence if the original value shown in Table 4 is 20.51 and an 81% customer satisfaction level can be achieved with this value (Table 5), the optimization analysis can obtain an optimum average charging rate value, which can lead to customer satisfaction levels of over 90%, if desired. We now go on to provide an example of the optimization analysis for the case study . Of the input data shown in Table 4, only Average charging rate, Inventory turnover, Staff turnover, Average service time per customer and % Cashier utilization were considered to allow some range of variation in order to set a goal set by the DM. So these five variables were used for the optimization analysis (Table 7). The objective was to achieve a maximum value for Customer service and Store presentation (values as close to 1 as possible). These results evidence how minor variations can be made to the model's input data to accomplish higher Customer satisfaction and Store presentation levels. Although it is true that improvements in optimization did not obtain a very good result for Customer satisfaction, where improvement was only 0.8%, it did improve Store presentation to a better extent (2.1%). Moreover, the scorecard output proposal is depicted in Figure 4, where we can see that all the measures that the model needs to be run are customizably and graphically grouped in the input data zone. Only the measures Average customer count, Total sales and Total items are not directly customizable in this panel because their values must be supplied in a spreadsheet. The optimization analysis results are shown in Tables 8 and 9.

Table 6 Variation ranges for the sensitivity analysis.

Code	Measure/indicator	Original value	Minimum	Maximum
IN1	Average charging rate (items/min)	20.51	10	30
IN3	Building area (m2)	929	0	1500
IN5	Inventory turnover	2.41	0	5
IN6	Staff turnover	0.034	0	3
IN9	Trade area (m2)	540	0	1000
IN21	Average service time per customer	0.0464	0	0.1
IN23	%Cashier utilization	0.5066	0	1

Table 7 Optimization analysis of the input ranges.

Code	Measure/indicator	Minumum	Maximum
IN1	Average charging rate (items/min)	20	23
IN5	Inventory turnover	2	3
IN6	Staff turnover	0.02	0.05
IN21	Average service time per customer	0.017	0.083
IN23	%Cashier utilization	0.4	1

Table 8 Optimization values for the input data.

Code	Measure/indicator	Initia value	Optimal value	% Var
IN1	Average charging rate (items/min)	20.51	20	-2.5%
IN5	Inventory turnover	2.41	2	-17.0%
IN6	Staff turnover	0.034	0.02	-41.2%
IN21	Average service time per customer	0.0464	0.051	9.9%
IN23	%Cashier utilization	0.5066	0.5066	0.0%

Table 9 Optimization results for the objectives.

	Customer satisfaction	Store presentation
Initial value	81.3%	82.8%
Optimal value	82.1%	84.9%
% dif	0.8%	2.1%

5. Conclusions

The main objective of the methodology proposed herein was to identify the relation between the various performance measures monitored in operations decision support systems. Each performance measure is normally monitored separately without being included in the analysis of their effects on other measures, and they can define overall business performance as a whole set. In this study, both types were analyzed: interactions between measures, and their simulation and optimization to establish objectives and goals. This methodology was not based on traditional performance measurement categories for balanced scorecards, rather we sought a comprehensive cross-sectional analysis of the various metrics considered. This methodology is sufficiently generic for it to be applied to any organization or even supply chain, but is completely flexible for it to be adapted to the specific conditions in each case; i.e. strategic or input data can be varied as required.

In this specific case, by way of example we used the operation of a supermarket of retail product sales. This example evidenced how this methodology is able to take the initial set of measures and transform them into an all-round analysis tool for key variables in decision making, such as customer satisfaction, store presentation, total staff and cash registers. These variables are often crucial in such businesses when making decisions about an already open store, such as re-adjusting staff, investing in its infrastructure, extending or reducing customer movements, exhibiting products, and customer charging processes. Eventually, these decisions match the main cost measures in such businesses, and evidently match end consumer satisfaction. This methodology is also useful for evaluating explansion plans, which is another relevant purpose of supermarket supply chains, which can be used for other organization types. Finally, further research is oriented to apply the proposed methodology in other real world cases. A forthcoming work is about the proposal of a simulation model and scorecard output, according to the proposed methodology, to support the analysis stage of new facility implementations and global supplier network developments in the context of the internationalisation of operations decision making.

References

1. Akkermans HA, van Oorschot KE. 2004. Relevance assumed: a case study of balanced scorecard development using system dynamics. Journal of Operational Research Society 56: 931–941.

2. Alfaro-Saiz J, Ortiz-Bas A, Rodríguez-Rodríguez R. 2007. Performance measurement system for enterprise networks. International Journal of Production and Performance Management, 56: 305–334.

3. Angerhofer BJ, Angelides MC. 2006. A model and a performance measurement system for collaborative supply chains. Decision Support Systems 42: 283–301.

4. Barnabè F. 2011. A "system dynamics based Balanced Scorecard" to support strategic decision making: Insights from a case study. International Journal of Production and Performance Management 60: 446–473.

5. Bianchi C. 2012. Enhancing performance management and sustainable organizational growth through system-dynamics modelling. In Grösser S.N., Zeier R. (eds.) Systemic Management for intelligent organizations. Springer Berlin Heidelberg, 143-161.

6. Boj JJ, Rodriguez-Rodriguez R, Alfaro-Saiz JJ. 2014. An ANP-multi-criteria-based methodology to link intangible assets and organizational performance in a Balanced Scorecard context. Decision Support Systems 68: 98–110.

7. Burgess TF. 1998. Modelling the impact of reengineering with system dynamics. International Journal of Operations and Production Management 18: 950–963.

8. Cai J, Liu X, Xiao Z, Liu J. 2009. Improving supply chain performance management: A systematic approach to analyzing iterative KPI accomplishment. Decision Support Systems 46: 512–521.

9. Campuzano F, Mula J. 2011. Supply Chain Simulation: A System Dynamics Approach for Improving Performance. Springer, London.

10. Chong IG, Jun CH. 2005. Performance of some variable selection methods when multicollinearity is presented. Chemometrics and Intelligent Laboratory Systems 78: 103–112.

11. Cohen J, Cohen P, West SG, Aiken LS. 2013. Applied Multiple Regression/Correlation Analysis for Behavioral Sciences. Routledge, New Jersey.

12. Cook RD. 1977. Detection of influential observation in linear regression. Technometrics 19: 15–18.

13. Draper NR, Smith H. 1998. Applied Regression Analysis. 3rd edition. Wiley, New York.

14. Durbin J, Watson GS. 1950. Testing for serial correlation in least squares regression: I. Biometrika 37: 409–428.

15. Field A. 2013. Discovering Statistics Using IBM SPSS Statistics. Sage, London.

16. Forrester JW. 1961. Industrial Dynamics. The MIT Press. Massachusetts Institute of Technology, Cambridge, Massachusetts.

17. Gunasekaran A, Patel C, McGaughey RE. 2004. A framework for supply chain performance measurement. International Journal of Production Economics 87: 333–347.

18. Hazewinkel M. 2001. Encyclopaedia of Mathematics, Supplement III. Springer, Netherlands.

19. Jusoh R, Nasir Ibrahim D, Zainuddin Y. 2008. The performance consequence of multiple performance measures usage: Evidence from the Malaysian manufacturers. International Journal of Production and Performance Management 57: 119-136.

20. Kleijnen JPC, Smits MT. 2003. Performance metrics in supply chain management. Journal of the Operational Research Society 54: 507–514.

21. Kutner MH, Nachtsheim C, Neter J. 2004. Applied Linear Regression Models. McGraw-Hill/Irwin.

22. Mora-Monge CA, Subba-Rao S, Gonzalez ME, Sohal AS. 2006. Performance measurement of AMT: a cross regional study. Benchmarking International Journal 13: 135–146.

23. Mula J, Campuzano-Bolarin F, Díaz-Madroñero M, Carpio KM. 2013. A system dynamics model for the supply chain procurement transport problem: comparing spreadsheets, fuzzy programming and simulation approaches. International Journal of Production Research 51: 4087–4104.

24. Nudurupati SS, Bititci US, Kumar V, Chan FTS. 2011. State of the art literature review on performance measurement. Computers and Industrial Engineering 60: 279–290.

25. Otto A, Kotzab H. 2003. Does supply chain management really pay? Six perspectives to measure the performance of managing a supply chain. European Journal of Operational Research 144: 306-320.

26. Rodriguez RR, Saiz JJA, Bas A. 2009. Quantitative relationships between key performance indicators for supporting decision-making processes. Computers and Industry 60: 104–113.

27. Rodriguez-Rodriguez R, Alfaro Saiz JJ, Ortiz Bas A, Carot JM, Jabaloyes JM. 2010. Building internal business scenarios based on real data from a performance measurement system. Technological Forecasting and Social Change 77: 50–62.

28. Santos SP, Belton V, Howick S. 2002. Adding value to performance measurement by using system dynamics and multicriteria analysis. International Journal of Operations and Production Management 22: 1246–1272.

29. Sousa CM. 2004. Export performance measurement: an evaluation of the empirical research in the literature. Journal of the Academic of Marketing Science 9: 1–23.

30. Sterman JD. 2000. Business Dynamics: Systems Thinking and Modeling for a Complex World. Irwin/McGraw-Hill, Boston.

31. Tung A, Baird K, Schoch HP. 2011. Factors influencing the effectiveness of performance measurement systems. International Journal of Operation and Production Management 31: 1287–1310.

32. Verdecho MJ, Alfaro-Saiz JJ, Rodriguez-Rodriguez R, Ortiz-Bas A. 2012. A multi-criteria approach for managing inter-enterprise collaborative relationships. Omega 40: 249–263.

Annex

Figure 2 Causal loop diagram.

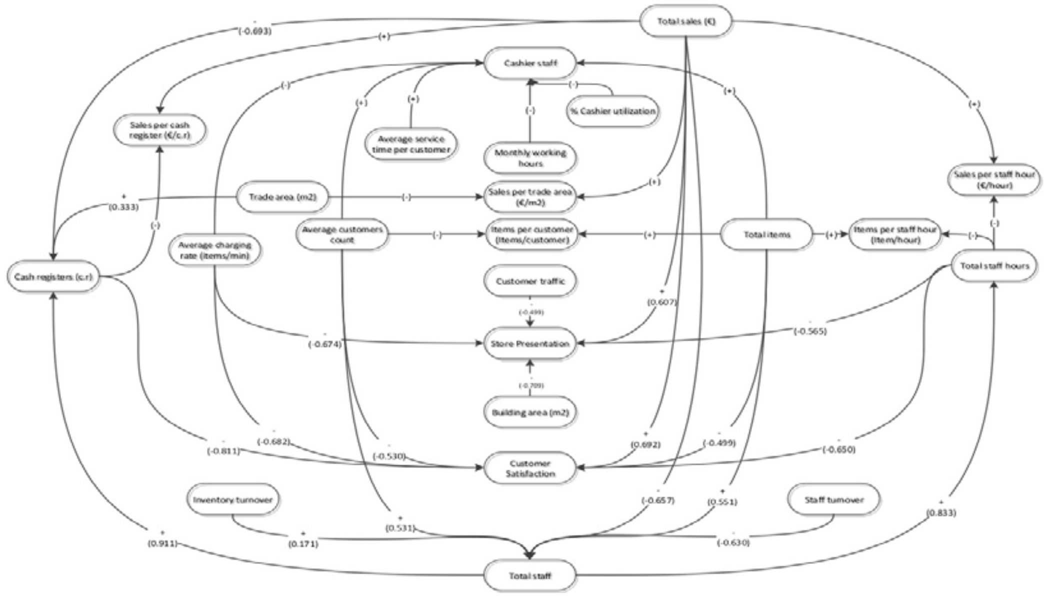
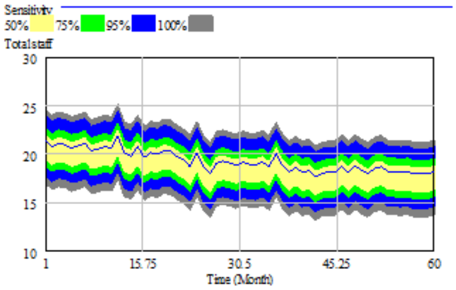
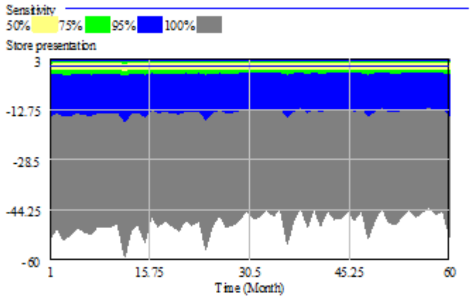


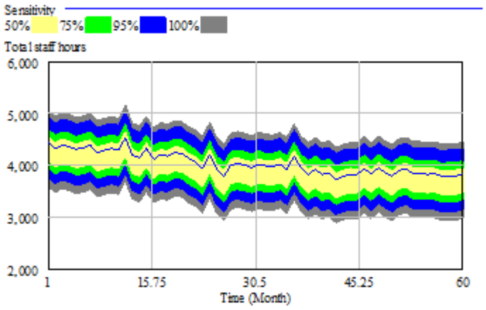
Figure 3 Causal loop diagram.



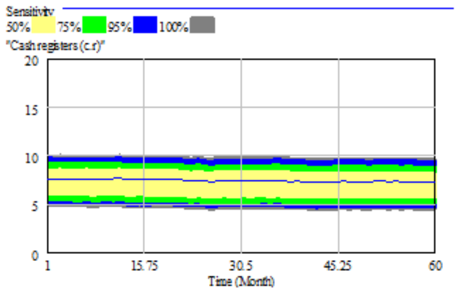
Sensitivity analysis result. Total Staff



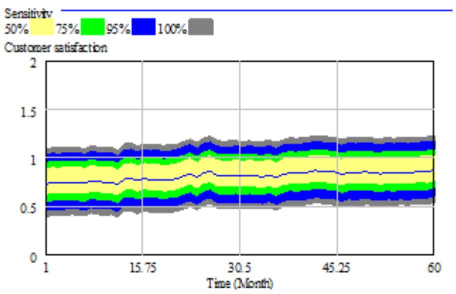
Sensitivity analysis result. Store presentation



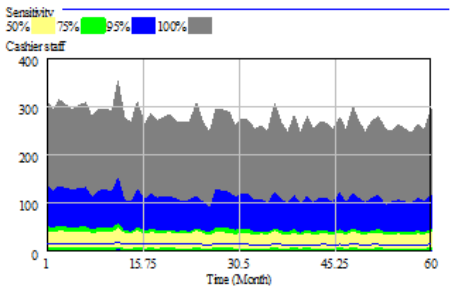
Sensitivity analysis result. Total Staff hours



Sensitivity analysis result. Cash registers



Sensitivity analysis result. Customer satisfaction



Sensitivity analysis result. Cashier staff

Figure 4 Scorecard output

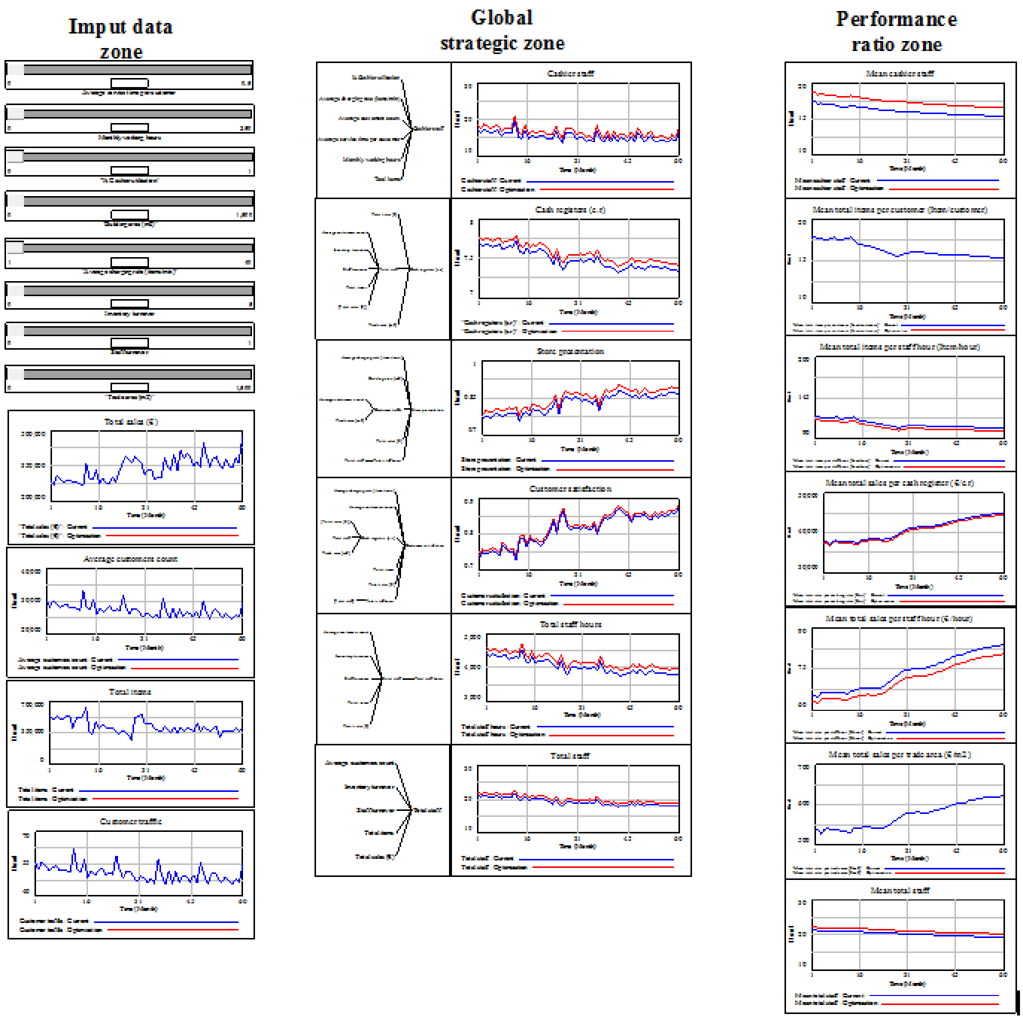


Table 4 Input data

Period	Average customers count	Total sales (€)	Total Items
1	29963	290368.6	535396
2	28191	272769.8	487921
3	30295	312431.2	531490
4	29409	294149.3	497898
5	28152	282502.9	500128
6	28812	292993.4	535712
7	29347	291312.7	538969
8	27969	279294.6	400192
9	27997	285291.2	511976
10	28018	274105.6	519854
11	27399	270529.2	528086
12	33289	359012.5	626895
13	28472	299810.4	346572
14	27543	292532.3	338579
15	30569	332535.1	472351
16	26150	273879.7	393661
17	27991	301016.6	437954
18	26911	277413.5	406026
19	27847	275676.1	401306
20	28726	303663.1	392907
21	26823	286389.7	376539
22	27639	325923.8	341386
23	26980	359426.7	369730
24	31868	389542.1	360635
25	27967	379126.3	331724
26	26043	364436.4	267830
27	28013	388867.8	517942
28	27754	371872.8	529835
29	26789	363265.7	551395
30	25118	314516.0	449960
31	26577	334180.9	455216
32	26235	337326.5	455084
33	24971	294074.9	393103
34	26067	302759.5	378152
35	24807	300402.8	353115
36	30899	384075.6	424138
37	26846	347810.9	384012
38	24630	332472.9	355971
39	27989	398220.4	415942
40	24773	333154.7	345918
41	27820	416238.2	415348
42	25219	380902.4	363285
43	26778	407135.2	403976
44	26447	386273.1	390972
45	25237	348914.9	361430
46	26873	370260.0	461150
47	25105	346713.8	359727
48	30048	445847.5	426542
49	26876	385752.6	389852
50	24686	347892.6	360021
51	27097	377510.2	394819
52	27700	389190.2	418179
53	26174	358093.1	302654
54	24987	338220.1	369830
55	26361	380066.5	374306
56	25442	349026.4	363065
57	24593	344488.4	359228
58	25898	384774.2	407147
59	25296	361835.5	366164
60	29796	477205.7	410579

Code	Indicator	Validation value
IN1	Average charging rate (items/min)	20.51
IN3	Building area (m2)	929
IN5	Inventory turnover	2.41
IN6	Staff turnover	0.034
IN9	Trade area (m2)	540
IN21	Average service time per customer	0.0464
IN22	Monthly working time	207
IN23	%Cashier utilization	0.5066

Table 5 The model validation results

Period	Customer satisfaction			Store presentation			Cash registers (c.r)			Total staff			Total staff hours		
	Real	Model	% error	Real	Model	% error	Real	Model	% error	Real	Model	% error	Real	Model	% error
1.0	0.73	0.73	-0.1%	0.73	0.8	4.5%	7.75	7.7	-0.6%	21.56	21.4	-0.9%	4692.0	4431.9	-5.5%
2.0	0.75	0.75	0.0%	0.74	0.8	5.0%	7.71	7.7	-0.7%	21.60	20.8	-3.7%	4376.8	4323.2	-1.2%
3.0	0.75	0.74	-0.6%	0.75	0.8	3.4%	7.69	7.7	-0.2%	21.21	21.2	-0.2%	4398.4	4393.1	-0.1%
4.0	0.74	0.75	1.0%	0.75	0.8	3.2%	7.69	7.7	-0.3%	21.22	21.0	-1.1%	4692.1	4360.2	-7.1%
5.0	0.75	0.75	0.8%	0.73	0.8	8.2%	7.69	7.6	-0.6%	21.34	20.7	-3.2%	4470.5	4299.3	-3.8%
6.0	0.73	0.75	2.7%	0.73	0.8	6.7%	7.69	7.7	-0.4%	21.33	20.9	-2.2%	4335.3	4338.1	0.1%
7.0	0.72	0.74	2.6%	0.72	0.8	7.0%	7.68	7.7	0.0%	21.33	21.1	-1.0%	4409.7	4384.7	-0.6%
8.0	0.72	0.77	7.0%	0.72	0.8	8.7%	7.60	7.6	0.2%	21.14	20.4	-3.6%	4549.3	4242.9	-6.7%
9.0	0.72	0.76	5.4%	0.72	0.8	9.1%	7.60	7.6	0.4%	19.79	20.6	4.1%	4174.5	4286.0	2.7%
10.0	0.72	0.75	3.7%	0.72	0.8	9.1%	7.60	7.7	0.7%	20.85	20.8	-0.3%	4171.2	4323.0	3.6%
11.0	0.72	0.75	4.9%	0.72	0.8	10.1%	7.85	7.6	-2.7%	21.58	20.6	-4.4%	4209.7	4290.6	1.9%
12.0	0.72	0.73	0.5%	0.72	0.8	4.3%	7.64	7.7	1.0%	21.08	21.9	4.0%	4257.8	4537.4	6.6%
13.0	0.73	0.78	6.7%	0.73	0.8	7.4%	7.64	7.6	-0.8%	20.02	20.1	0.5%	4122.9	4195.6	1.8%
14.0	0.75	0.79	5.6%	0.74	0.8	7.0%	7.64	7.6	-1.1%	20.21	19.8	-1.8%	3925.1	4141.9	5.5%
15.0	0.75	0.76	1.8%	0.75	0.8	4.0%	7.63	7.6	-0.1%	20.03	20.8	3.9%	4354.9	4326.9	-0.6%
16.0	0.74	0.79	6.4%	0.75	0.8	7.8%	7.63	7.6	-1.0%	19.79	19.7	-0.4%	4292.1	4117.7	-4.1%
17.0	0.75	0.78	3.8%	0.73	0.8	9.6%	7.63	7.6	-0.6%	19.61	20.2	2.9%	4219.0	4203.2	-0.4%
18.0	0.76	0.78	2.3%	0.74	0.8	7.3%	7.63	7.6	-0.6%	19.57	20.0	2.2%	4048.4	4171.5	3.0%
19.0	0.77	0.77	0.3%	0.80	0.8	-2.0%	7.63	7.6	-0.1%	19.93	20.4	2.3%	4153.1	4244.4	2.2%
20.0	0.76	0.78	1.7%	0.79	0.8	-0.1%	7.62	7.6	-0.3%	19.87	20.3	2.1%	4430.7	4227.9	-4.6%
21.0	0.78	0.79	1.7%	0.78	0.8	2.6%	7.61	7.6	-0.7%	20.00	19.8	-1.2%	4143.5	4124.5	-0.5%
22.0	0.78	0.81	3.7%	0.78	0.8	3.9%	7.58	7.5	-1.0%	20.10	19.4	-3.5%	4115.2	4057.1	-1.4%
23.0	0.79	0.84	5.8%	0.80	0.8	5.6%	7.58	7.4	-2.1%	20.26	18.7	-7.6%	4067.0	3928.3	-3.4%
24.0	0.79	0.80	1.5%	0.68	0.8	15.9%	7.53	7.5	0.1%	19.99	20.2	1.0%	3967.0	4207.4	6.1%
25.0	0.86	0.84	-2.2%	0.91	0.84	-7.7%	7.53	7.4	-1.6%	19.81	18.7	-5.5%	4081.8	3927.9	-3.8%
26.0	0.86	0.87	0.3%	0.90	0.86	-5.3%	7.53	7.4	-2.4%	19.61	18.0	-8.2%	3764.0	3791.1	0.7%
27.0	0.85	0.82	-3.1%	0.90	0.85	-5.6%	7.43	7.4	0.2%	19.83	19.1	-3.7%	4165.7	4000.3	-4.0%
28.0	0.84	0.81	-3.4%	0.90	0.84	-6.0%	7.43	7.5	0.5%	19.55	19.3	-1.3%	4371.1	4034.5	-7.7%
29.0	0.85	0.82	-3.6%	0.89	0.85	-3.8%	7.25	7.5	2.8%	19.26	19.1	-0.9%	3998.8	3997.1	0.0%
30.0	0.85	0.82	-3.7%	0.88	0.85	-4.0%	7.46	7.5	-0.1%	19.26	18.9	-2.1%	3840.8	3955.3	3.0%
31.0	0.85	0.81	-3.8%	0.89	0.84	-5.8%	7.46	7.5	0.1%	19.31	19.2	-0.7%	4033.4	4012.8	-0.5%
32.0	0.84	0.82	-2.4%	0.90	0.84	-6.1%	7.44	7.5	0.2%	19.94	19.0	-4.8%	4385.2	3978.2	-9.3%
33.0	0.85	0.81	-3.7%	0.90	0.84	-7.1%	7.44	7.5	0.5%	19.36	19.0	-2.1%	4177.0	3972.1	-4.9%
34.0	0.86	0.81	-5.4%	0.90	0.82	-8.7%	7.44	7.5	0.7%	19.35	19.2	-0.7%	4358.5	4022.6	-7.7%
35.0	0.87	0.83	-4.5%	0.90	0.84	-6.9%	7.44	7.4	0.1%	19.13	18.7	-2.3%	3847.0	3921.4	1.9%
36.0	0.87	0.80	-7.7%	0.91	0.80	-11.9%	7.44	7.5	1.2%	19.24	20.1	4.3%	4216.2	4182.5	-0.8%
37.0	0.87	0.83	-4.3%	0.90	0.84	-7.3%	7.44	7.4	0.0%	18.95	18.9	-0.4%	4082.7	3958.1	-3.0%
38.0	0.87	0.85	-2.2%	0.90	0.86	-4.3%	7.44	7.4	-0.8%	18.59	18.2	-2.3%	3536.3	3819.7	8.0%
39.0	0.86	0.84	-2.1%	0.90	0.85	-5.5%	7.46	7.4	-0.9%	18.47	18.7	1.2%	3852.4	3919.6	1.7%
40.0	0.86	0.85	-1.1%	0.86	0.86	-0.3%	7.46	7.4	-1.1%	18.27	18.2	-0.5%	4016.4	3823.4	-4.8%
41.0	0.86	0.86	0.1%	0.90	0.86	-4.1%	7.46	7.4	-1.4%	18.36	18.3	-0.1%	3863.6	3856.1	-0.2%
42.0	0.87	0.87	0.2%	0.89	0.88	-1.2%	7.45	7.3	-1.7%	18.35	17.7	-3.5%	3654.6	3732.3	2.1%
43.0	0.88	0.87	-1.9%	0.89	0.87	-1.7%	7.43	7.3	-1.3%	18.27	18.0	-1.3%	3830.3	3797.2	-0.9%
44.0	0.89	0.86	-3.1%	0.88	0.87	-1.1%	7.43	7.4	-0.9%	18.31	18.2	-0.7%	4045.0	3824.0	-5.5%
45.0	0.88	0.85	-2.7%	0.87	0.86	-1.4%	7.43	7.4	-0.7%	18.69	18.2	-2.8%	3764.3	3822.2	1.5%
46.0	0.87	0.83	-4.1%	0.87	0.85	-1.6%	7.38	7.4	0.6%	17.96	18.8	4.5%	3828.6	3937.2	2.8%
47.0	0.87	0.85	-2.2%	0.87	0.86	-0.7%	7.32	7.4	0.7%	18.03	18.1	0.7%	3784.8	3817.6	0.9%
48.0	0.87	0.85	-2.5%	0.86	0.85	-1.8%	7.27	7.4	1.6%	18.18	18.8	3.5%	3902.0	3946.7	1.1%
49.0	0.87	0.85	-2.2%	0.86	0.86	0.3%	7.27	7.4	1.4%	17.96	18.4	2.2%	3757.4	3857.2	2.7%
50.0	0.86	0.86	-0.3%	0.85	0.87	1.9%	7.27	7.4	1.2%	17.92	18.0	0.3%	3369.2	3782.9	12.3%
51.0	0.86	0.84	-1.3%	0.87	0.85	-2.1%	7.27	7.4	1.8%	17.84	18.6	4.1%	3824.3	3899.5	2.0%
52.0	0.84	0.84	-0.2%	0.88	0.85	-3.4%	7.27	7.4	1.9%	17.66	18.7	5.9%	3800.1	3924.2	3.3%
53.0	0.86	0.86	-0.5%	0.88	0.85	-3.0%	7.27	7.4	1.5%	17.55	18.2	4.0%	3850.5	3836.7	-0.4%
54.0	0.86	0.85	-1.5%	0.88	0.86	-2.0%	7.27	7.4	1.7%	17.53	18.3	4.1%	3691.2	3837.7	4.0%
55.0	0.85	0.86	1.1%	0.88	0.86	-1.7%	7.27	7.4	1.3%	17.28	18.2	5.3%	3495.3	3826.3	9.5%
56.0	0.84	0.85	0.6%	0.87	0.86	-1.6%	7.27	7.4	1.6%	17.33	18.3	5.3%	3978.9	3838.2	-3.5%
57.0	0.83	0.86	3.2%	0.87	0.87	-0.1%	7.27	7.4	1.3%	17.01	18.0	5.7%	3524.1	3785.0	7.4%
58.0	0.83	0.86	3.7%	0.88	0.87	-0.1%	7.27	7.3	1.0%	17.02	18.0	5.9%	3527.2	3795.2	7.6%
59.0	0.83	0.86	3.2%	0.87	0.87	0.1%	7.27	7.4	1.2%	16.96	18.0	6.2%	3487.6	3792.9	8.8%
60.0	0.83	0.88	5.1%	0.87	0.87	-0.3%	7.27	7.3	0.6%	17.13	18.2	6.3%	3736.6	3831.7	2.5%
Max	0.89	0.88	7.0%	0.91	0.88	15.9%	7.85	7.71	2.8%	21.60	21.91	6.3%	4692.08	4537.44	12.3%
Min	0.72	0.73	-7.7%	0.68	0.76	-11.9%	7.25	7.31	-2.7%	16.96	17.70	-8.2%	3369.24	3732.27	-9.3%
Average	**0.81**	**0.81**	**0.2%**	**0.83**	**0.83**	**0.4%**	**7.48**	**7.48**	**0.0%**	**19.28**	**19.29**	**0.2%**	**4033.65**	**4036.91**	**0.3%**