Knowledge-Based Systems 97 (2016) 1–10

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

An expert system for financial performance assessment of health care structures based on fuzzy sets and KPIs

Cinzia Muriana[a,*], Tommaso Piazza[a], Giovanni Vizzini[b]

[a] University of Pittsburgh Medical Center (UPMC) Italy, Via Discesa dei Giudici 4, 90133 Palermo, Italy
[b] IRCCS – ISMETT (Istituto Mediterraneo per i Trapianti e Terapie ad alta specializzazione), Via Tricomi 5, 90133 Palermo, Italy

ARTICLE INFO

Article history:
Received 11 September 2015
Revised 20 January 2016
Accepted 20 January 2016
Available online 29 January 2016

Keywords:
Performance assessment of health care
Key performance indicators
Adaptive models
Expert system
Fuzzy sets

ABSTRACT

Interest in the field of performance assessment of health care structures has grown in recent decades. In fact, the possibility of determining overall performances of health care structures plays a key role in the optimization of resource allocation and investment planning, as it contributes to reducing the uncertainty of future performance. In this context, key performance indicator (KPI) tools have been developed to assess the performance of health care structures from process, organizational, cost, financial, and output points of view. In practice, they are periodically calculated, and the effect of several KPIs on the overall performance of health care structures is determined by management through human judgment or software that provides synthetic dashboards. Given their non-stationary nature, performance assessment and forecasting are generally tackled by employing adaptive models, but these approaches cannot reflect the holistic nature of performance itself, nor take into account the impact of KPIs on the overall performances. In order to overcome these shortcomings, this study presents an expert system whose engine relies on fuzzy sets, in which the input–output relations and correlations have been modeled through inference rules based on time-series trends. The focus is on the financial performance assessment of a health care structure, such as a hospital. The approach is of an interdisciplinary kind, as several indicators were taken as inputs that relate to output, process, and cost KPIs, and their impact on the output measure, which is of a financial kind (namely the total reimbursement). The output measure calculated by the expert system was then compared with that predicted using only adaptive forecasting models, and the error with respect to the actual value was determined. Results showed that measures determined by fuzzy inference, able to effectively model actual input–output relations, outperform those of adaptive models.

1. Introduction

Today, the possibility of assessing and forecasting health care performance is of fundamental importance in properly planning investments and allocating financial resources. In accomplishing these tasks, managers usually rely on Key Performance Indicators (KPIs) able to support them in the decision-making process, providing process, organizational, output, cost, and financial indicators. Managers analyze KPIs and seek to determine the overall effect of such variables on health care performance, usually relying on human judgment or software that provides synthetic dashboards. In order to forecast future health care performance, the non-stationary behavior of KPIs is usually modeled through adaptive forecasting models, but these are unable to capture the effect of several variables simultaneously. However, in order to make decisions, the effect of all the variables affecting the overall performance has to be taken into consideration. Thus, traditional approaches, aimed at providing KPI dashboards are not sufficient; in fact, they do not allow us to infer the performance in relation to multiple factors interacting simultaneously, as they do not account for the holistic behavior of health care performances, assessment of which has to be characterized by multidisciplinary approaches. In other words, traditional tools do not allow us to synthesize the overall impact of several input variables on the global performance. For this reason, it is necessary to model the simultaneous role of different KPIs in determining the final score. In this regard, there is little in the literature, with only a few cases of frameworks aimed at determining the global health care performance based on a set of input factors. This paper presents an expert system for the assessment of financial performance of health care structures that takes into consideration the simultaneous impact of the process, cost, and output KPIs, relying on a fuzzy-based inference engine,

* Corresponding author. Tel.: +39 3290128446.
E-mail addresses: cmuriana@ismett.edu (C. Muriana), tpiazza14@gmail.com (T. Piazza), gvizzini@ismett.edu (G. Vizzini).

http://dx.doi.org/10.1016/j.knosys.2016.01.026
0950-7051/

whose knowledge base is represented by a data warehousing tool developed within the Smart Health 2.0 project (PON04a2_C).

2. Review of the literature and goal of the study

2.1. KPIs and health care performance assessment

Health care performance assessment has attracted the interest of researchers in recent decades, as the possibility of monitoring performances through a set of KPIs is seen as a suitable tool for investigating the actual state of health care structures from organizational, processing, and clinical standpoints.

Based on classification reported by Kalinichenko et al., [20], health care KPIs can be divided into the structure, process, outcome, and output measures. "Structure" involves organizational characteristics of the caregivers, including human, physical, and financial resources. In particular, financial resources can refer to reimbursements recognized by the regional/national government, and to parameters used to determine the entity of reimbursements, such as diagnosis-related group (DRG) weights. "Process" indicates those activities involving health care practitioners and patients, such as length of stay in hospitals, procedures, and other treatment practices, and use of prescribed medicines. "Outcome" refers to the impact of these activities on a patient's current and future health status. Finally, "Output" indicates the quantity of health services provided, without taking into account effects of these activities on patients' health, e.g., the number of visits or patient volume.

It is worth mentioning that based on the use one can make of KPIs, they can be divided into internal, namely that used to monitor and improve the outcomes of care processes, and external, used by governments, patient organizations, and payers to assess the quality of a health care provider, and to compare it with the performance of other health caregivers [4].

There are several studies in the literature that deal with the topic of KPIs for health care structure assessment. In particular, the attention of researchers has been focused on identifying the most suitable panel of KPIs in relation to the characteristic of the specific health care structure. As an example, Berg et al., [4] described the development and implementation of the first national, public, and obligatory set of hospital performance indicators in Holland. They focused on effectiveness and safety KPIs, and developed a set of indicators with the aim of monitoring the quality of the care delivered by providers, enhancing the transparency of the hospital sector, and prompting individual hospitals to improve their scores. Burge et al., [8] developed a set of quality indicators (QIs) for primary care practice, for the primary prevention and chronic disease management of ischemic heart disease, hypertension, hyperlipidemia, and heart failure using a four-stage modified Delphi approach. Bradley et al., [6] determined a set of 33 indicators to assess the quality of a childhood cancer system. Finally, Boulkedid et al., [7] proposed a panel of indicators for maternity units, while Cruppè et al. (2015) investigated the feasibility of 48 quality indicators in ambulatory care in Germany through a cross-sectional observational study. However, the authors mentioned limited their contribution to the determination of set of KPIs for health care, and did not consider the simultaneous impact of different KPIs on the overall health care structure performance.

KPIs can commonly be characterized by a non-stationary behavior in time, as they are affected by several exogenous variables that cannot be controlled by health care managers. For example, the diffusion of specific bacteria can promote some diseases that involve a greater number of patients to be hospitalized, with a consequent increase in health care costs. Moreover, hot summers can increase the need for care of respiratory problems and heart attacks, while cold winters can increase the need for pulmonary care. Financial and legislative decisions, such as decisions about the amount of reimbursements for health care services provided, also affect health care performances.

The possibility of forecasting future behavior of health care performance is a topic frequently addressed in the literature, as reported for example in Jones and Spiegelhalter [19]. Approaches usually used are based on forecasting tools that rely on regressive models (see, for example, [16]) that are simple to apply but, undoubtedly, not very flexible. In fact, given that health care policies and governance are affected by several external conditions (exogenous variables), e.g., legislative decisions, seasonality affecting some kind of diseases, the performance assessment is in turn a holistic issue. Thus, holistic methodologies, such as adaptive forecasting models should be preferred for modeling such complex systems. Adaptive forecasting usually allows us to model the health care structure behavior through a series of additive components (level, trend, seasonality) that characterize the structures themselves.

In this setting, the possibility of employing adaptive models and generalized exponential smoothing methods, which are holistic approaches, can be of great help in modeling and forecasting such phenomena. The classical Bayesian linear regression models are unable to reproduce some of the features frequently observed in non-stationary processes, while, on the contrary, in such cases time-series methods are extremely effective. Linear regression models allow us to model only phenomena in which the future behavior depends on that of previous periods. In particular, by definition, they are used when a phenomenon has a linear behavior in time (see, for example, [13]). In this context, non-stationary phenomena, such as those considered in this work, cannot be interpreted by models that are not flexible and able to capture the variation of data within short time periods. Conversely, adaptive forecasting methods can easily model phenomena that usually characterize non-stationary processes where the trend of data changes in the short term, and can be found, for example, in Simple Exponential Smoothing (SES), and Holt's and Winter's models [11]. Such models allow us to interpret level, trend, and seasonality of data by taking into account short-term variations, modeled through the use of constants, able to represent the impact that past data can have on future trends. In order to highlight these differences, in this study a fuzzy-based expert system is presented, and adaptive models are employed to test the effectiveness of results given by the designed fuzzy system, comparing data forecasted through adaptive models with those that arise from the fuzzy system itself.

2.2. KPI-based frameworks for performance assessment

Drivers for assessing health care performances through suitable frameworks arise from the need of measuring and raising the productivity of health care systems themselves [21]. For this reason, setting up a panel of KPIs for health care performance assessment can be a useful approach to improving knowledge with respect to specific aspects of health care performances, but is unsuitable when global and cross-dimensional knowledge is required. As observed by Toplicianu et al., [38], "*the atypical nature of health care services market and the specificity of the activity in hospitals, determines that performance analysis is a complex, multifactorial process.*" In this respect, methodologies in the literature neglect the importance of a comprehensive assessment that allows us to know how several different input variables simultaneously impact global performance. In order to satisfy this requirement, frameworks that allow us to determine overall performance should be applied. In other words, a holistic approach is needed to model the dependence of global health care performance on multiple factors that simultaneously interact. Examples of such approaches can be found in Neumann et al., [29], who investigated the suitability of boosted decision trees for the case-mix adjustment involved in comparing the performance of various health care entities, and found that boosting decision trees are a powerful tool for case-mix adjustment in health care performance measurement. Kruk and Freedman [24] proposed a framework for the assessment of health care systems, considering three major dimensions of performance: effectiveness, equity, and efficiency. Inputs taken into account were policies, funding, and organization performance. Moreover, they presented a systematic review of the literature on performance indicators of health care structures, with a focus on developing countries. Gauld et al., [12] developed a national scorecard for assessing health structure performance derived from routine data. They concluded that such a framework was a useful method for combining a range of data to provide an overall view of health care performances. The literature mentioned typically aimed at determining health care KPIs that can be employed in health care assessment, but none of them provide a multidisciplinary approach that takes into account the simultaneous impact of KPIs arising from different fields on the overall performance. In this paper, output, cost, and process KPIs are employed with regard to the setting of an expert system (ES), for determining the financial performance. The general characteristics of the system are detailed in Section 2.3.

2.3. Expert systems as a framework for KPI-based performance assessment

Expert systems (ESs) are interactive computer-based decision tools that use both facts and heuristics to solve difficult decision-making problems, based on knowledge acquired from an expert. They model the problem-solving behavior of an expert in a narrow domain. ESs can implement both judgmental knowledge and formal knowledge of established theories. One of the characteristics that make ESs easily implementable is flexibility, i.e., the faculty of integrating new knowledge incrementally into its existing store of knowledge. ESs consists of a set of components, such as a knowledge base, working storage, inference engine, user interface, and individuals who interact with the system [15]. The operating principle of an ES consists of a dialogue conducted by the user interface between the user and the system. The user provides information on the problem to be solved, and the system then attempts to provide insights derived (or inferred) from the knowledge base [1]. These insights are provided by the inference engine after examining the knowledge base. The knowledge base consists of some encoding of the domain of expertise for the system. This can be in the form of semantic nets, frames, or production rules [30]. The inference engine is a control mechanism that allows us to manipulate the knowledge, and deduce results in an organized manner. It applies the knowledge present in the knowledge base to infer conclusions. In this paper, the inference engine of the ES is realized using fuzzy sets. This topic is addressed in Section 2.4.

2.4. Fuzzy sets for performance assessment

Fuzzy logic is a set theory, introduced by [39] in 1965, as an extension of classical set theory. Historically, this was closely related to the concept of fuzzy measure, proposed just after by Sugeno [37]. Fuzzy set theory is an effective methodology for tackling complex problems that are characterized by a high number of input variables impacting one or more output variables. For detailed discussion about fuzzy set theory see, for example, Klir and Yuan [23], Passino and Yurkovich [31], Zimmermann [40], Baczynski and Jayaram [3].

Within the scope of a health care performance assessment, a review of the literature found no case study that used fuzzy theory.

The main applications of fuzzy sets in the health care domain are related to health status assessment, as can be found in Abdullah, [2], who studied health-related quality of life (HRQL), seen as a means for assessing health conditions among patients who suffer from specific diseases or illnesses. The study aimed to model the relationship between HRQL variables using an integrated model of a fuzzy inference system, and linear regression. The methodology allowed the authors to determine the strength of the relationship between multiple variables of HRQL and health status. The approach of Abdullah [2] is much closer to our approach because it is of a multidisciplinary kind. However, it does not consider the holistic nature of health care performances. Another example related to fuzzy-logic-based health care was reported by Patil and Mohsin [32], who employed wireless sensor network systems to continuously monitor the pulse and temperature of patients at a remote place or in a hospital through a wearable wireless sensor system that gathers data with established frequency. Data stored in the database is passed to a fuzzy logic controller to improve accuracy and amount of data and information to be sent to the remote patient. Bingchuan and Herber [5] proposed a fuzzy-logic-based context model and a related context-aware reasoning middleware that provides a personalized, flexible, and extendible reasoning framework for Context Aware Real-time Assistant (CARA). It provides context-aware data fusion and representation as well as inference mechanisms that support remote patient monitoring and caregiver notification. Finally Medjahed et al., [28] developed a health care monitoring system based on a fuzzy inference engine capable of learning and recognizing human activities in daily living.

Our review of the literature found no application to financial performance assessment of fuzzy systems based on a multidisciplinary approach that takes into account the non-stationary and holistic behavior of the health care field. For this reason, this paper can be considered original in itself. The goals of the study are detailed in Section 2.5.

2.5. Goals of the study

As can be seen from the review of the literature mentioned in previous sections, no significant approaches aimed at taking into account the holistic behavior of performances of health care structures have been identified in the literature, beyond those here reported. Thus, in order to bridge this gap, our study presents a general framework for a financial health care performance assessment based on KPIs, and relying on a fuzzy approach. The goal is to forecast health care performance based on KPI assessment, and determine the impact on the overall performance. In this context, fuzzy systems can be effectively applied for determining variables (input of the model) responsible for the system performance (output of the model), and the way they affect the performance itself. These rely on inference models to manage input–output relations. Despite the obvious utility that such a methodology can provide, there are no examples of its application in the field of health performance assessment. Unlike present models that do not apply the holistic and multidisciplinary nature of the health care field, our principal aim was to design an ES based on a fuzzy approach, employing KPIs that can determine the overall performance of a health care structure. In particular, the aim is to show that the accuracy of performance determined with a fuzzy system is higher than that arising from adaptive forecasting models. As regards the fuzzy systems, its knowledge base is represented by a data warehousing tool, designed and developed within the Smart Health 2.0 project (PON04a2_C). The paper focuses on the possibility of modeling the impact of multiple input factors on the overall performance by using a fuzzy system, providing management with a decision-making tool that can monitor health care performance

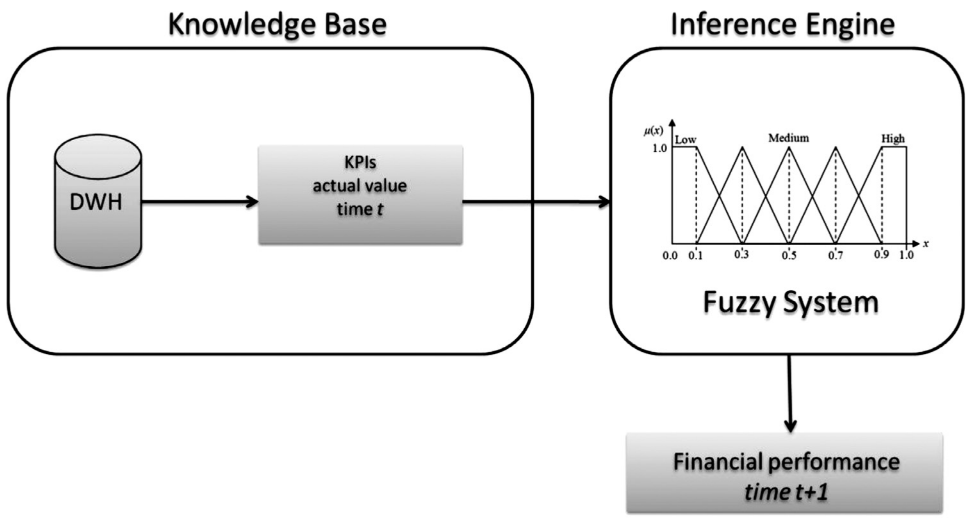

Fig. 1. Expert system based on a fuzzy system.

for given input variables. The focus of the study is not only on the possibility of determining health care performances, but also on the comparison between results arising from the fuzzy system, in terms of performance assessment, and that of adaptive forecasting methods, in order to show how fuzzy systems can outperform the accuracy of measure assessment with respect to adaptive models. The study first presents a general framework and addresses methodological issues; then a case study is proposed to test the feasibility of the designed ES, using data extracted by the data warehousing tool. The goal is to determine the financial health care performance in terms of total DRG fee (DRGF), taking into account input variables such as Patient Volume (PV), Length of Stay (LoS), and Diagnosis-Related Group Weights (DRGW) that affect such performance. Input and output variables of the model were determined by managers of the UPMC Italy Management Control Department of the IRCCS-ISMETT health care structure, in Sicily, Italy.

3. Materials and methods

3.1. Fuzzy-based expert system for performance assessment

The process of health care performance assessment is usually carried out by the Management Control Department, which based on data related to *t−1* time determines the expected reimbursement that is usually received at time *t*. The designed ES uses output, process, and cost KPIs as input in order to evaluate financial performance. In particular, the knowledge base for input data is represented by a data warehouse (DWH), from which data are extracted monthly. These data are employed to define the inference engine of the ES, realized through a fuzzy system. Because the input data are updated monthly, the quality of data managed by the fuzzy system will improve over time, thus providing a feedback loop for the expert system. This allows us to enrich inference rules and gradually reduce the standard deviation of measures, improving in turn the quality of the output assessed. A qualitative scheme of the proposed system is reported in Fig. 1.

Fig. 1 shows the components of the ES, which are the DWH, the knowledge base from which KPIs are extracted, and the fuzzy system, which allows us to combine the knowledge base by means of inference rules. From Fig. 1 it can be seen that at time *t* input data are gathered from the DWH and employed in the fuzzy system, and the output variable for the reimbursement expected at time *t+1* is calculated. As the actual value of the output variable is made available at time *t+1*, the inference engine data will be updated, the correctness of inference rules and the goodness of distribution fitting will be checked, and new rules will be inserted, if necessary.

4. Case study: The fuzzy-based expert system

4.1. Data setting

In order to apply the fuzzy method, input and output variables have to be defined. This task was requested of the UPMC Italy Management Control Department, as field experts. The variable chosen as output was the monthly DRGF (financial KPI) paid to health care structures by the national health system. This measure was preferred for its significance in determining the global financial performance of health care structures, as it is usually reported in official documents for health care assessment. Thus, input variables were selected by applying statistical analysis to monthly time series of a panel of KPIs. In particular, variables responsible for 99% of the variance were selected through principal component analysis. Among them, those having a higher correlation with the output variable, and lower or no correlation among them, were chosen. The variables were PV (output KPI), average LoS (process KPI), and average DRGW (cost KPI), and refer to monthly periods according to the output variable. In particular, the PV (output KPI) is responsible for 52.21% of the total variance, the average LoS (process KPI) for 38.65%, and the average DRGW (cost KPI) for 8.45%. For the purpose of building the fuzzy engine, the time series considered was that corresponding to years 2011–2014 at the IRCCS-ISMETT health care structure.

In order to properly employ data, some typical analyses were carried out. First, mean and variance of time series were calculated based on the following equations:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

where n is the sample size.

$$s^2(n) = \sum \frac{(x_i - \bar{x})^2}{n-1} \tag{2}$$

Thus, confidence intervals were determined for each of the input–output variables, as:

$$\bar{x}(n) \pm t_{1-\frac{\alpha}{2}} \sqrt{\frac{S^2(n)}{n}} \tag{3}$$

Table 1
Confidence interval and precision of measures.

	Average	Standard deviation	Upper limit	Bound limit	Half width	precision
PV (units)	225.1250	22.3249	219.7519	230.4981	5.373095	0.952266
LoS (days)	9.692	1.0779	9.432586	9.951429	0.259422	0.946467
DRGW	3.072	0.2848	3.002967	3.140037	0.068535	0.955373
DRGF (€)	2548461.6329	299648.2266	2,476,343	2,620,580	72118.51	0.943402

Table 2
Autocorrelation analysis.

Autocorrelation (Lag)	PV (units)	LoS (days)	DRGW	DRGF (€)
1	0.444577	0.213969	−0.07681	0.06276
2	0.209612	0.026285	0.032714	0.18566
3	0.224732	−0.16828	−0.0196	0.038409
4	0.130707	−0.04639	−0.04697	−0.03681
5	0.139871	−0.01901	0.007912	0.141736
6	−0.01953	−0.11204	0.067404	−0.14662

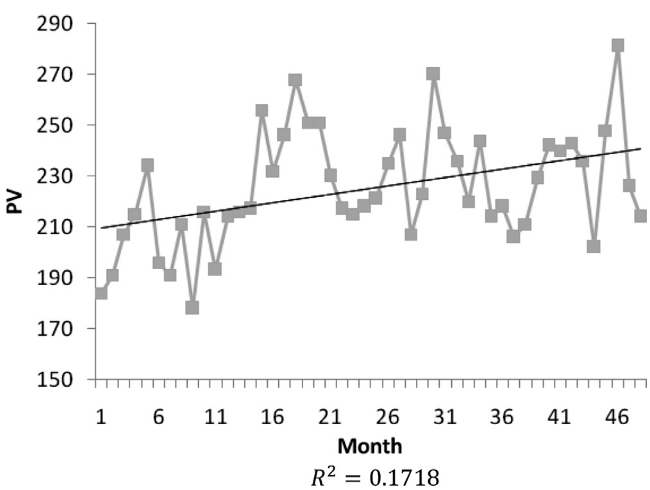

Fig. 2. Regression analysis of PV with time.

where $t_{1-\frac{\alpha}{2}}$ is the $(1-\alpha)$ percentile of the t-student distribution with $n-1$ degrees of freedom. The quantity $t_{1-\frac{\alpha}{2}}\sqrt{\frac{S^2(n)}{n}}$ represents the half width of the confidence interval indicated by h.

Finally, the accuracy of measures was calculated with the following equation:

$$P = 1 - \frac{2h}{\bar{x}} \tag{4}$$

For further details about such statistical aspects see Kleijnen, [22].

Results reported in Table 1 shows that the precision of measures is in all cases about 95%. This means that the sample size is sufficient to ensure a high accuracy of measures calculated in this study.

After measuring accuracy, data were detected in order to highlight possible trends and seasonality. Results reported in Table 2 and Figs. 2–5 show that data are not affected by seasonality (autocorrelation analysis), and are generally not trended (regression analysis and determination coefficient).

In particular, Table 2 shows that the maximum autocorrelation period is 1 for PV and LoS, while for DRGW it is 5, but the autocorrelation index is very low. Finally, the autocorrelation period of the DRGF is 5, but is very low as well.

Figs. 2–5 show that the data are not trended, as the determination coefficient is very low. The determination coefficient relates to the possibility of explaining the dependency of the output variable in relation to the input variable. In our case, there is no correlation between input and output variables, meaning that output data, such as PV, LoS, DRGW and DRGF are stable with time and do not depend on the progression of time.

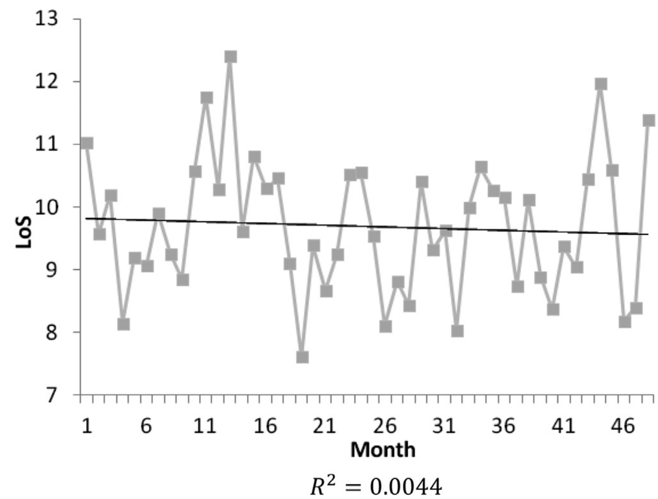

Fig. 3. Regression analysis of LoS with time.

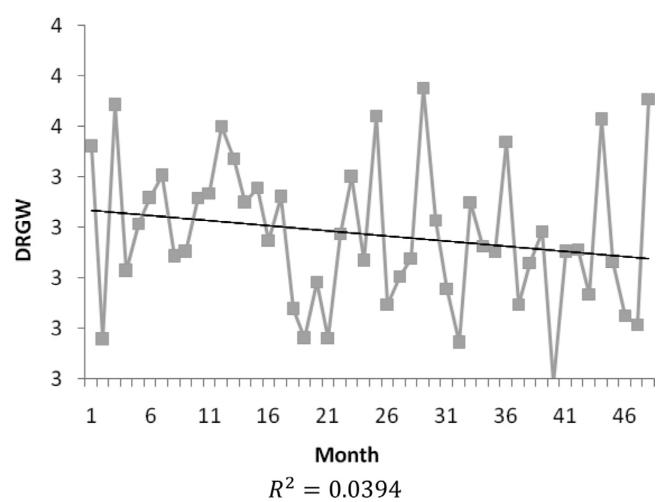

Fig. 4. Regression analysis of DRGW with time.

4.2. Fuzzy system

Having properly analyzed the data in order to set up the fuzzy system, the correlation between variables was investigated. First, the correlation between each of the input variables and the output variable has to be taken into consideration.

The correlation between the DRGW and the DRGF is positive, as is the correlation between the LoS and the DRGF. In fact, the DRGF recognized for the ith speciality is:

$$DRG\ Fee_i = F(DRGW,\ LoS) \tag{5}$$

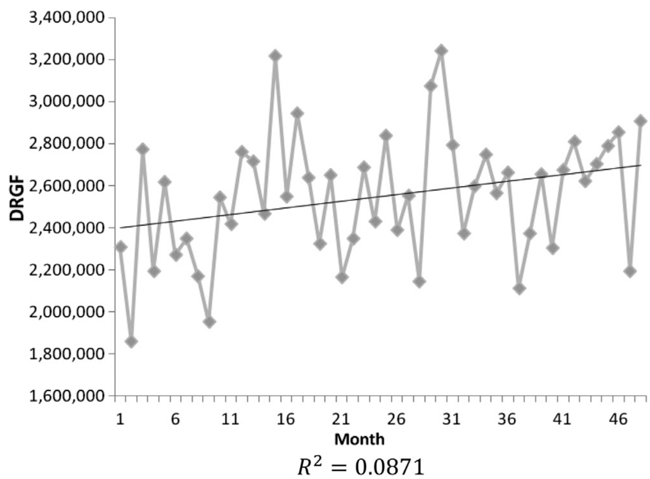

Fig. 5. Regression analysis of DRGF with time.

Table 3
Fitted distributions.

Variable	Distribution fitted	P Value	Average	Standard deviation
PV (LOW)	Normal	0.9224	186.500	6.8557
PV (MEDIUM)	Normal	0.2582	213.238	6.5643
PV (HIGH)	Normal	0.8793	239.778	7.7805
PV (VERY HIGH)	Normal	0.9676	268.750	10.2429
LoS (LOW)	Normal	0.7758	8.119	0.2643
LoS (MEDIUM)	Normal	0.9945	9.140	0.3428
LoS (HIGH)	Normal	0.9913	10.322	0.3622
LoS (VERY HIGH)	Normal	0.9993	11.879	0.4284
DRGW (LOW)	Normal	0.7706	2.692	0.0952
DRGW (MEDIUM)	Normal	0.5903	2.973	0.0663
DRGW (HIGH)	Normal	0.6244	3.208	0.0550
DRGW (VERY HIGH)	Normal	0.9951	3.533	0.0808
DRGF (LOW)	Normal	0.4151	2099133.854	124671.0287
DRGF (MEDIUM)	Normal	0.9607	2336784.806	38804.3565
DRGF (HIGH)	Normal	0.9999	2676061.339	139505.7230
DRGF (VERY HIGH)	Normal	0.7254	3178008.457	90424.8729

As can be seen from Eq. (5), the DRGF is a function of the DRGW and the LoS. The DRGW is computed in relation to age, gender, state at discharge, primary and secondary diagnosis, and procedures/surgical interventions. This information can be derived from the discharge form, and employed in specific software, named Grouper, to determine the DRGW. The national rate associates with each DRGW a fixed threshold, taking into account the average LoS in relation to procedures. Once this threshold has been surpassed, e.g., due to complications, the health care system recognizes an extra fee that is proportional to the number of days exceeding the threshold. This means that the DRGF for the *i*th speciality is computed as a sum of a base tariff depending on the DRGW and the average LoS plus an extra fee that is a function of days exceeding the average length of stay.

Moreover, analyzing data reported in the national tariff (see Decree Oct. 18, 2012, “Remunerazione prestazioni di assistenza ospedaliera per acuti, assistenza ospedaliera di riabilitazione e di lungodegenza post acuzie e di assistenza specialistica ambulatoriale” [33]), a correlation between LoS and the DRGF of 66.5% was found, while the correlation between the DRGW and the DRGF recognized was 98%.

From Eq. (5) it can be seen that the DRGF depends on the DRGW and the LoS, but a linear dependence cannot necessarily be enforced, as the extra fee is generally less than the base fee.

The correlation between the PV and the DRGF is positive, as the DRGF increases the PV increases as well, based on the following equation:

$$Total\ DRGF = \sum_{i=1}^{n} DRGF_i V_i \tag{6}$$

where V_i is the volume of patients related to specialty *i*. Eq. (6) shows that the total DRGF is a weighted sum of the DRGF recognized for each specialty. Thus, the DRGF is proportional to the PV.

Because the linear dependence of the output variable on each of the input variables cannot be enforced, the possibility of employing a multiple linear regression was excluded.

As regards the correlation between input variables, it does not make sense to determine the correlation between variables PV and LoS, or PV and DRGW, as there is no conceptual relation between them. Regarding the correlation between the DRGW and LoS, it can be seen from Eq. (5) that the two variables are only partially correlated. In particular, analyzing data reported in the aforementioned national tariff, it was possible to see that there is a moderate correlation of 65% between such variables. This is due to the fact that the DRGW is determined based on a number of factors, among which the LoS. Thus, in order to apply fuzzy sets, one of the two variables should be removed. On the other hand, doing so would mean neglecting some important components of the effect of such variables on the output. First, removing the DRGW or the LoS would lead to a loss of the effect of correlation between such input variables and the output. Moreover, removing the DRGW would involve neglecting the effect of other variables based on which the DRGW is determined by Grouper. Finally, removing the LoS would lead to losing the effect of extra days on the DRGF. As a result, the two variables, even if partially correlated, were included in the model.

The second step of the process involved the definition of membership functions for input and output variables. Thus, a frequency analysis of monthly time series was done, enabling the identification of the number of levels for each variable, and the most feasible distribution function. The frequency analysis allowed us to find four levels or classes for each variable, which were denominated “LOW”, “MEDIUM”, “HIGH,” and “VERY HIGH.” This was the only step that required human intervention in the definition of how many levels were to be set based on the frequency analysis. Then, in order to establish the function that best fit the data, a fitting analysis was done with the Minitab 17 tool. The results are reported in Table 3, where ranges for each variable are also included. As can be seen, some distributions partially overlap.

The third step consisted in defining inference rules, such as IF-THEN rules, allowing us to model the inference engine behavior. For that reason, time series data were analyzed once again, determining for each variable the level to which each data in the sample belongs, based on the above-mentioned classification. So, for example, if the single observation of the PV has a value of “268,“ the LoS “9.097 days, “ the DRGW “2.774,” and the DRGF “€2,639,002,“ they belong respectively to levels VERY HIGH, MEDIUM, MEDIUM, and HIGH of the respective input–output variables, based on ranges reported in Table 3, and the IF-THEN rule that arises is “IF PV is VERY HIGH AND Average LoS is MEDIUM AND Average DRGW is HIGH, THEN the DRG Fee is HIGH.” In cases in which, due to overlapping between fuzzy sets, data could be attributed to more than one fuzzy set, they were attributed to that set with respect to data that had the greater membership degree. In this way, the first set of IF-THEN rules was determined. Thus, inference rules were submitted to the Management Control Department for assessment. They pruned inference rules from inconsistencies (two rules in which same antecedents correspond to different consequents), and checked for redundancies (two rules with identical consequents and different but not contradictory antecedents),

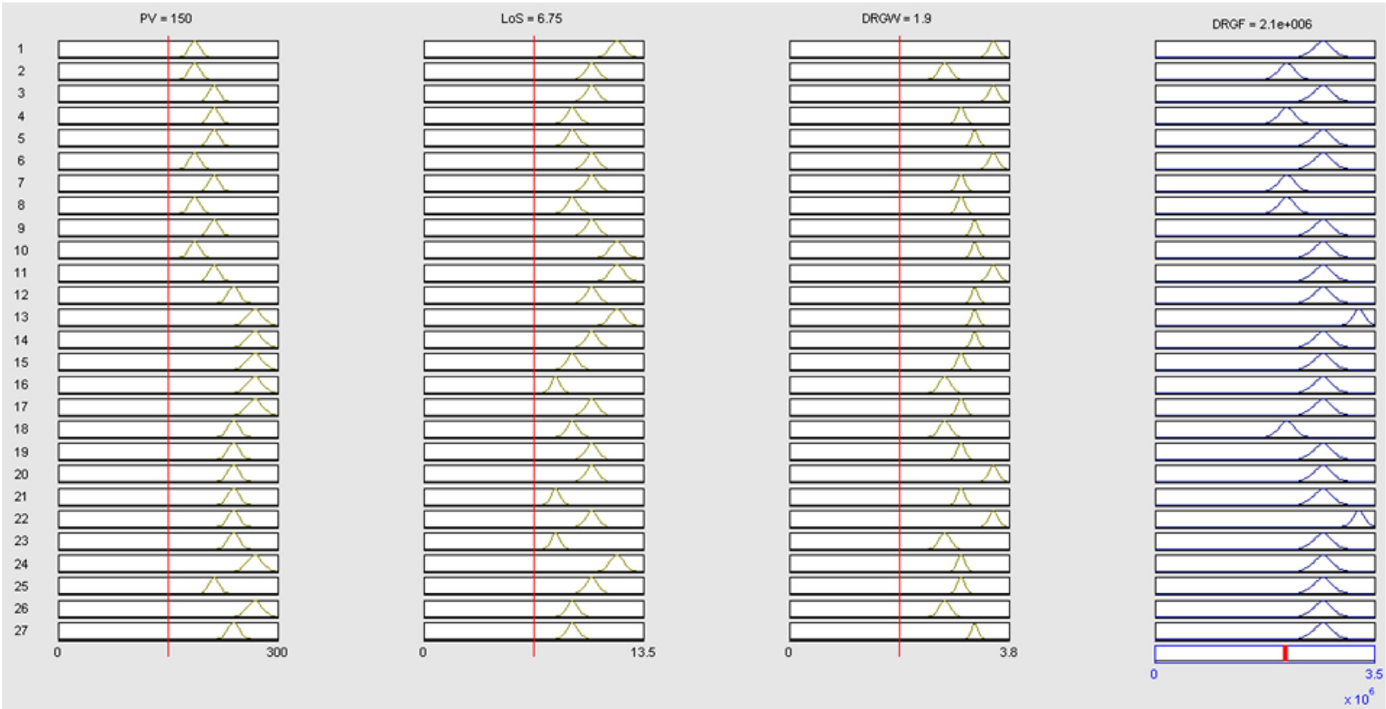

Fig. 6. Fuzzy inference engine.

Table 4
Inference rules.

Rule	PV	LoS	DRGW	DRGF
1	LOW	VERY HIGH	VERY HIGH	HIGH
2	LOW	HIGH	LOW	LOW
3	MEDIUM	HIGH	VERY HIGH	HIGH
4	MEDIUM	MEDIUM	MEDIUM	LOW
5	HIGH	MEDIUM	HIGH	HIGH
6	MEDIUM	MEDIUM	HIGH	HIGH
7	LOW	HIGH	VERY HIGH	HIGH
8	MEDIUM	HIGH	MEDIUM	LOW
9	LOW	MEDIUM	MEDIUM	LOW
10	MEDIUM	HIGH	HIGH	HIGH
11	LOW	VERY HIGH	HIGH	HIGH
12	MEDIUM	VERY HIGH	VERY HIGH	HIGH
13	HIGH	HIGH	HIGH	HIGH
14	VERY HIGH	VERY HIGH	HIGH	VERY HIGH
15	VERY HIGH	HIGH	HIGH	HIGH
16	VERY HIGH	MEDIUM	MEDIUM	HIGH
17	VERY HIGH	LOW	LOW	HIGH
18	VERY HIGH	HIGH	MEDIUM	HIGH
19	HIGH	MEDIUM	LOW	LOW
20	HIGH	HIGH	MEDIUM	HIGH
21	HIGH	HIGH	VERY HIGH	HIGH
22	HIGH	LOW	MEDIUM	HIGH
23	HIGH	HIGH	VERY HIGH	VERY HIGH
24	HIGH	LOW	LOW	HIGH
25	VERY HIGH	VERY HIGH	MEDIUM	HIGH
26	MEDIUM	HIGH	MEDIUM	HIGH
27	VERY HIGH	MEDIUM	LOW	HIGH

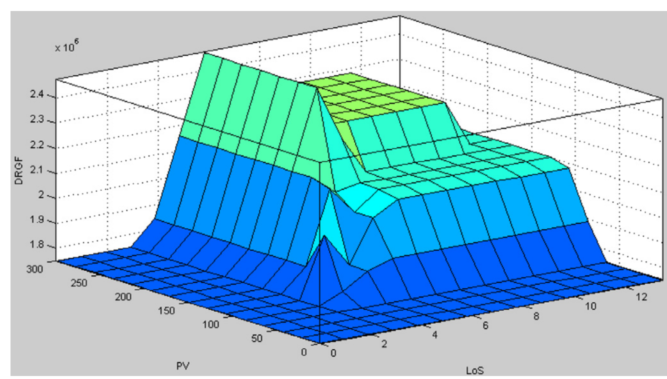

Fig. 7. Surface that shows the relation between the input variables PV and LoS and the output variable.

in order to identify unnecessary rules, thus guaranteeing the consistency and completeness of the rules themselves. For details on consistency of fuzzy rules see, for example, Leung and So, [26], Chiu [10], Gonzalez and Perez, [14], Jin et al., [18], Roychowdhury and Wang, [35], Reusch, [34], Linkens and Chen, [27], Sindelár and Babuska, [36], Lee et al., [25]. IF-THEN rules are reported in Table 4.

Finally, the fuzzy system was built with the Matlab 7.1 tool, using the Mamdani method. The defuzzification method chosen was that of the centroid. Fig. 6 reports the rules, while Figs. 7–9 report results in terms of surfaces representing the relations of input–output variables.

Fig. 6 shows, for each of the rules given in Table 4, the validity range of each input and output variable, based on the distributions fitted.

Fig. 7 shows that analyzing the impact of the PV and the LoS on the DRGF, for fixed LoS the DRGF initially is constant as the PV increases, and then decreases as the PV further increases. This means that there is an optimal PV that allows us to reach a maximum DRGF. For fixed PV, the DRGF initially increases, then reaches a maximum, and finally decreases. This means that there is an optimal LoS that maximizes the DRGF.

Fig. 8 shows that analyzing the impact of the DRGW and the LoS on the DRGF, for fixed DRGW the DRGF initially is constant as the LoS increases, and then decreases as the LoS further increases. This means that there is an optimal LoS that allows us to reach a maximum DRGF. For fixed LoS the DRGF always increases, as the DRGW increases as well.

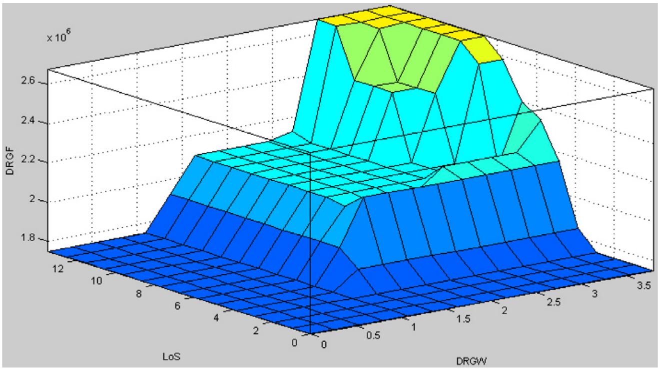

Fig. 8. Surface that shows the relation between the input variables DRGW and LoS and the output variable.

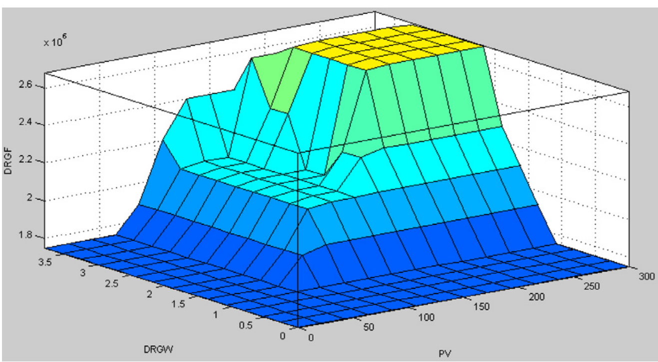

Fig. 9. Surface that shows the relation between the input variables DRGW and PV and the output variable.

Fig. 9 shows that analyzing the impact of the DRGW and the PV on the DRGF, for fixed DRGW the DRGF always increases, as the PV increases as well. For fixed PV, the DRGF initially increases as the DRGW increases as well, and then decreases as the DRGW further increases. This means that there is an optimal DRGW that allows us to reach a maximum DRGF.

These results allow us to use the fuzzy set not only to infer the output behavior depending on the simultaneous variation of the input variables, but also to find the combination of the input variables that maximize the DRGF. Once the fuzzy system was set up, it was tested by carrying out a set of case studies, and assigning a random number for the input variables, and determining the output variable. The validation of results was assessed by the Management Control Department, which confirmed the goodness of the results.

The goodness of the calculated output measure will improve as the number of data available increases with time. This means that as new performance measures for input and output variables are made available, the input–output variable membership functions will be updated and, consequently, the inference engine will be checked, allowing the fuzzy system to improve its overall goodness of fit with respect to the real system.

4.3. Forecasting

In the field of health care performance assessment, the usefulness of multiple adaptive short-term time series, such as auto regressive (AR), moving average (MA), and smoothed methods is well recognized, as documented, for example, by Jones and Spiegelhalter [19]. They affirmed the importance of knowing future health care performances for the proper planning of service provision. In addition, they pointed out that as the next set of time series data

Table 5
MAD of forecasted measures with respect to the actual DRGF value.

	MAD	MAPE
DRGF Fuzzy system	168686.01	6625%
DRGF SES	243359.63	9749%
DRGF MA (N = 3)	248060.31	9768%
DRGF MA (N = 4)	258153.18	10,140%
DRGF MA (N = 6)	204744.43	9974%
DRGF MA (N = 8)	225280.17	10,135%
DRGF MA (N = 12)	232969.24	9501%

becomes available, short-term predictive distributions can be used to identify units that have experienced recent changes. Moreover, they observed that assessments of predictive performance can be used to compare models. The work carried out in this paper employed adaptive models and MA for the comparison of measures determined with the fuzzy system in assessing health care performances. As pointed out in Section 4.1, time series seem to be neither trended nor affected by seasonality. For this reason, forecasting methods such as MA and SES can be effectively employed to model data with time.

For MA the following formula was employed to forecast data:

$$P_{t+1}(t) = \frac{\sum_{j=0}^{N-1} D_{t-j}}{N} \tag{7}$$

where Eq. (7) means that the forecasted value $P_{t+1}(t)$ is the average of the actual value of the last N periods.

For SES, future KPI values were predicted based on the following equations:

$$\hat{y}_2 = \alpha y_1 + (1-\alpha)L_0 \tag{8}$$

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_{t-1} \tag{9}$$

where y_t is the observation at time t, $\hat{y}_{t-1}$ is the forecasted value at time $t-1$, and $0 \leq \alpha \leq 1$ is the smoothing constant for the level. The initial level L_0 is determined as the average of available observations. For references about MA models, see Holt [17].

Eq. (8) results in the evaluation of the KPIs at a future time, and $t+1$ as a weighted average of the (adjusted) previous estimate and the most recent information acquired at time t. Concerning the establishment of the smoothing constants α, practical issues are discussed in detail in [9]); however, a common approach is to determine the values of α that minimize the mean or median absolute error, or a similar measure. As the number of observed KPIs increases, the accuracy of the forecast will eventually improve.

Time series representing input variables were forecasted based on Eqs. (8) and (9), and the optimal value of α parameter determined by minimizing the mean absolute deviation (MAD) using the Excel Solver tool. The optimal α was found to be 0.071.

Output measures forecasted through SES and MA were compared with those determined by the fuzzy system in order to test the effectiveness of the proposed framework. MA forecasts were determined for N = 3,4,5,6,7. Results are reported in Fig. 10, while measures of performance are reported in Table 5, where MAD and mean absolute percent error (MAPE) are shown.

Fig. 10 shows the comparison of the actual values of the DRGF with the forecasted value of the DRGF in the case in which the MA is adopted (with N = 3, 4, 6, 8, 12), the forecasted value in the case in which the SES is adopted, and the forecasted value of the fuzzy system. It is worth noting that the values forecasted by the fuzzy system are closer to the actual values with respect to values forecasted with other methodologies. This can be further supported by the statistics reported in Table 5.

Table 5 shows that the MAD of the output forecasted with SES and MA is greater than that calculated with the fuzzy system. This

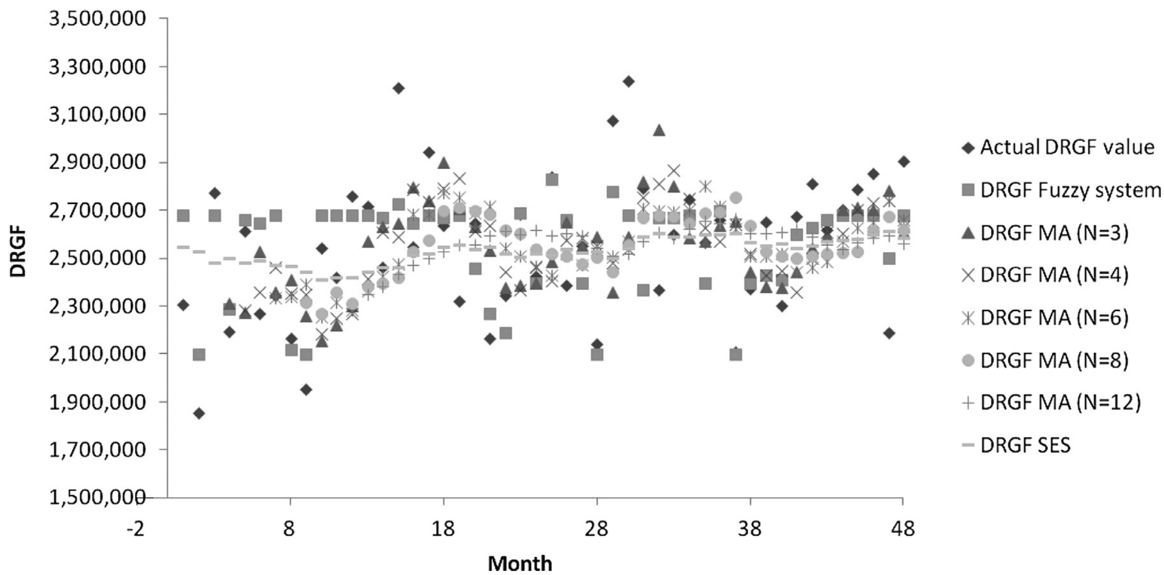

Fig. 10. Comparison of results.

means that the fuzzy system, taking into consideration the correlation of input–output variables and, thus, the simultaneous impact of all input variables on the overall performance, outperforms the results achievable by simply forecasting the output variable. In fact, because the reimbursement system is affected by several factors, such as those considered as input variables, it can be deemed a holistic system, in which neglecting the role of input variables leads to wrong results. In this context, applying forecasting methods that do not reflect the interdisciplinary nature of the phenomenon is not sufficient to ensure the correctness of the conclusions drawn from data behavior. The framework proposed can be effectively seen as an expert system whose inference engine is represented by fuzzy sets. In fact, as the amount of data gathered increase with time, the deviation measures will decrease, meaning that the forecasted values tend to converge toward the true value.

5. Conclusions

KPI-based health care performance assessment is a topic of relevant interest, as it involves considering non-stationary KPI behavior and the holistic nature of health care performances. In this study, an expert system for financial performance assessment was designed, relying on fuzzy sets as an inference engine. Unlike traditional decision-making systems, which are used to determine a KPI panel synthesized by human judgments, the fuzzy approach allows us to determine the simultaneous impact that several KPIs can have on health care performances, taking into account the correlation between input variables and the output variable. In order to test the effectiveness of the proposed methodology, a comparison between fuzzy inference engine forecasting and traditional adaptive models was made. The results showed that a comprehensive approach, such as that proposed, outperforms adaptive forecasting methods in minimizing MAD and MAPE of measures, as it takes into account the correlation between variables and their effect on the output variable. The methodology proposed can be considered sufficiently general, as new input–output variables can be added in order to better reflect the real system characteristics.

However, the study focused on the relation between some cost, process, and output KPIs, taken as inputs, with a financial KPI, taken as output, though the possibility of determining a unique output score, synthesizing the performance from the process, output, financial, and cost standpoints, should also be investigated. This could allow us to improve information on the overall health care performance. In this field, multi-criteria decision-making (MCDM) methods may be usefully employed to determine such output scores, taking the input variables as criteria, and their monthly score as alternatives. Such an approach would allow us to take into account the direction of preference (criteria that have to be maximized/ minimized) and the impact of each input variable on the performance assessment through the definition of criteria weights.

Acknowledgments

This research was supported by the Smart Health project (PON04a2_C), carried out within the Smart Cities program 2007–2013.

References

[1] M. Abdelhadi, T.G. Krishna, G. Kanann, New developed model for arabic information retrieval system based on knowledge base system, Int. J. Emerg. Res. Manag.Technol. 2 (11) (2013) 1–7.
[2] L. Abdullah, Modeling of health related quality of life using an integrated fuzzy inference system and linear regression. International Conference on Robot PRIDE 2013-2014 - Medical and Rehabilitation Robotics and Instrumentation, ConfPRIDE 2013-2014, Procedia Computer Science 42 (2014) 99–105.
[3] M. Baczynski, B. Jayaram, Fuzzy Implications, Springer Verlag, Heidelberg, 2008.
[4] M. Berg, Y. Meijerink, M. Gras, et al., Feasibility first: developing public performance indicators on patient safety and clinical effectiveness for Dutch hospitals, Health Policy 75 (1) (2005) 59–73.
[5] Y. Bingchuan, J. Herbert, Fuzzy CARA - A fuzzy-based context reasoning system for pervasive healthcare, Procedia Comput. Sci. 10 (2012) 357–365.
[6] N. Bradley, P.D. Robinson, M.L. Greenberg, et al., Measuring the quality of a childhood cancer care delivery system: assessing stakeholder agreement, Value Health 16 (4) (2013) 639–646.
[7] R. Boulkedid, C. Alberti, O. Sibony, Quality indicator development and implementation in maternity units, Best Pract. Res. Clin. Obstet. Gynaecol. 27 (2013) 609–619.
[8] F. Burge, K. Bower, W. Putnam, J. Cox, Quality indicators for cardiovascular primary care, Can. J. Cardiol. 23 (5) (2007) 383–388.
[9] C. Chatfield, M. Yar, 'Holt-winters forecasting: some practical issues', The Statistician 37 (1988) 129–140.
[10] S. Chiu, Selecting input variables for fuzzy models, J. Intell. Fuzzy Syst. 4 (4) (1996) 243–256.
[11] Chopra, S., Meindl, P., 2015. Demand Forecasting in a Supply Chain. Chapter 7. In "Supply Chain Management: Strategy, Planning, and Operation". Pearson, 6th.
[12] R. Gauld, S. Al-wahaibi, J. Chisholm, et al., Scorecards for health system performance assessment: the New Zealand example, Health Policy 103 (2011) 200–208.

[13] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin, Bayesian Data Analysis, Chapter 8, CRC Press, Taylor & Francis Group, NY, 2003.

[14] A. Gonzalez, R. Perez, Completeness and consistency conditions for learning fuzzy rules, Fuzzy Sets Syst. 96 (1998) 37–51.

[15] Griffin, N.L., Lewis, F.D., 1989. A Rule-based Inference Engine which is Optimal and VLSI Implementable. Technical report, Edition 149, Part 89, University of Kentucky Dept. of Computer Science.

[16] Health Performance Framework, 2012. Report of the Australian Government, Department of Health and Ageing.

[17] C. Holt, Forecasting seasonals and trends by exponentially weighted moving averages, Int. J. Forecast. 20 (2004) 5–10.

[18] Y. Jin, W. Seelen von, B. Sendhoff, On generating fc fuzzy rule systems from data using evolution strategies, IEEE Trans. Syst. Man Cybern.—Part b: Cybern. 29 (6) (1999) 829–845.

[19] H. Jones, D.J. Spiegelhalter, Improved probabilistic prediction of healthcare performance indicators using bidirectional smoothing models, J. R. Stat. Soc. Assoc. 175 (3) (2012) 729–747 Part.

[20] Kalinichenko, O., Amado, CAF., Santos, SP., 2013. Performance Assessment in Primary Health Care: A Systematic Literature Review. CEFAGE-UE Working Paper. Available at: http://econpapers.repec.org/paper/cfewpcefa/2013_5f03.htm. Accessed: 03 July 2015.

[21] Kelley, E., Hurst, J., 2006. Health Care Quality Indicators Project Conceptual Framework Paper, OECD HEALTH WORKING PAPERS nÂ°23. Available at: http://www.oecd.org/els/health-systems/36262363.pdf. Accessed: 03/07/2015.

[22] J. Kleijnen, An overview of the design and analysis of simulation experiments for sensitivity analysis, Eur. J. Oper. Res. 164 (2) (2005) 287–300.

[23] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic : Theory and Applications, Prentice Hall, NJ, 1995.

[24] E.M. Kruk, L.P. Freedman, Assessing health system performance in developing countries: a review of the literature, Health Policy 85 (2008) 263–276.

[25] H.M. Lee, J.M. Chen, C.L. Liu, Inconsistency resolution and rule insertion for fuzzy rule-based systems, J. Inf. Sci. Eng. 18 (2015) 187–210.

[26] K.S. Leung, Y.T. So, Consistency checking for fuzzy expert systems, Int. J. Approx. Reas. 9 (1993) 263–282.

[27] D.A. Linkens, M.Y. Chen, Input selection and partition validation for fuzzy modeling using neural network, Fuzzy Sets Syst. 107 (1999) 299–308.

[28] Medjahed, H., Istrate, D., Boudy, J., Baldinger, J.L., Bougueroua, L., Dhouib, M.A., Dorizzi, B., 2012. A Fuzzy Logic Approach for Remote Healthcare Monitoring by Learning and Recognizing Human Activities of Daily Living. In “Fuzzy Logic - Emerging Technologies and Applications”, Edited by Prof. Elmer Dadios, INTECH.

[29] A. Neumann, j. Holstein, J.R. Le Gall, E. Lepage, Measuring performance in health care: case-mix adjustment by boosted decision trees, Artif. Intell. Med. 32 (2) (2004) 97–113.

[30] S.A.O. Ogirima, S.O. Olabiyisi, E.O. Omidiora, A. Oluwafunke, Web-based decision support system for prescription in herbal medicine, Trans. J. Sci. Technol. 2 (11) (2012) 17–33.

[31] K.M. Passino, S. Yurkovich, Fuzzy Control, Addison Wesley Longman, 1998.

[32] P. Patil, S. Mohsin, Fuzzy logic based health care system using wireless body area network, Int. J. Comput. Appl. 80 (12) (2013) 46–51.

[33] “Remunerazione prestazioni di assistenza ospedaliera per acuti, assistenza ospedaliera di riabilitazione e di lungodegenza post acuzie e di assistenza specialistica ambulatoriale”, Decreto Oct 18, 2012. Available at: http://www.gazzettaufficiale.it/atto/stampa/serie_generale/originario Accessed: 06 July 2015.

[34] B. Reusch, Computational Intelligence, in: Theory and application Proceedings of the International conference, 64th Fuzzy Days, Dortmund, Germany, 1999 May 1999.

[35] S. Roychowdhury, Bh. Wang, Measuring Inconsistency in Fuzzy Rules, in: Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE International Conference on Fuzzy system, 2, Anchorage, AK, 1998, pp. 1020–1025.

[36] R. Sindelár, R. Babuska, Input selection for nonlinear regression models, IEEE Trans. Fuzzy Syst. 12 (5) (2004).

[37] Sugeno, M., 1974. Theory of fuzzy integrals and its applications. Doct. Thesis, Tokyo.

[38] SI. Toplicianu, V. Toplicianu, E. Brezeanua, Performance in the public health system —is only a purely medical issue? Procedia Econ. Financ. 20 (2015) 301–307.

[39] L.A. Zadeh, Fuzzy sets as a basis for theory of possibility, Fuzzy Set Syst. 3-28 (1978) 9–34.

[40] H.J. Zimmermann, Fuzzy Set Theory and its Applications, Springer Science+Business Media, New York, 2001 ISBN 978-94-010-3870-6.