Developing a hybrid comparative optimization model for short-term forecasting: an 'idle time interval' roadmap for operational units' strategic planning

Filippou, Miltiades and Zervopoulos, Panagiotis

10 August 2011

Developing a hybrid comparative optimization model for short-term forecasting: An 'idle time interval' roadmap for operational units' strategic planning

Miltiades Filippou
Institut EURECOM
2229 route des Crêtes, BP 193 Sophia-Antipolis, France
Email: filippou@eurecom.fr
Panagiotis D. Zervopoulos
China Center for Health Development Studies, Peking University,
38 Xueyuan Rd, Beijing, China
Department of Business Administration of Food and Agricultural Enterprises
University of Ioannina, 2 Georgiou Seferi St, Agrinio, Greece
Email: pzervopoulos@uoi.gr

Abstract

Data drain and data uncertainties for rival units affect the reliability and effectiveness of strategic plans for individual operational units. This study introduces a stochastic, multi-stage, optimization technique for short-term forecasting that intends to assist policy makers in developing 'flawless' plans for their organizations during the idle time interval in which official data and balance-sheet reports of the competitors are unavailable. The developed technique, called SDEANN, draws on the 'deterministic' data envelopment analysis (DEA) method, 'regression-type' artificial neural networks (ANNs), and the contamination of the outputs of the DEA analysis with statistical noise. Statistical noise represents the bias of a 'deterministic' sample optimum production frontier when generalization or the uncertainty of the data used becomes the issue. The SDEANN model respects the monotonicity assumption that prevails in microeconomic theory, uses the DEA definition of efficiency, and addresses the dimensionality issues of ANNs with minimum sample size requirements.

Keywords: forecasting, optimization, efficiency, data envelopment analysis (DEA), artificial neural networks (ANNs), statistical noise

1. Introduction

Optimum relative efficiency is a primary driver for strengthening profitability for private companies (Banker et al., 1984). It is also one of the main goals of public organizations that embrace the New Public Management directions summarized in the "3Es" acronym that stands for Efficiency, Effectiveness and Economy (Ferlie et al., 2007; Pollitt and Bouckaert, 2004; Worthington and Dollery, 2000). Although the measurement of relative efficiency is important in the strategic planning of an operational unit, the lack of data for the peer units operating in the same market (e.g. level and cost of the inputs used, level of outputs produced, and revenues obtained) is regarded as a source of uncertainty in the analyses of decision makers. During the short-term data drain for the rival units associated with the announcement of official financial and production data (e.g. release of balance-sheet reports), only ceteris paribus analyses can be conducted, in which either the peers are deemed to be inactive or the last optimum production frontier available is the benchmark that ignores temporal changes.

In addition, strategic plans can lead to flaws when sample data are used and the population production function is unknown or is arbitrarily selected. Consequently, the bias increases when data drain occurs and limited information is available on population attributes.

We tackle both issues at stake by developing a stochastic efficiency optimization tool for bridging the gap between the period *t* and the future period *t+1* in which data uncertainties exist for peer operational units. This tool, which we call the "stochastic data envelopment analysis artificial neural network" (SDEANN), yields generalized optimum input and output values for every operational unit by taking into account short-term perturbations of the production frontier and also any possible inconsistencies of the frontier due to a lack of population data.

The SDEANN model applies a hybrid analysis based on the data envelopment analysis (DEA) method and the artificial neural networks (ANNs). The optimization forecasting model is generalized in order to be applicable to the existing sample units and to operational units that are not included in the dataset under evaluation. Generalization is achieved through the transformation of the 'deterministic' empirical production frontier, specified by DEA, into a stochastic production frontier by incorporating statistical noise in the dataset. Depending on the orientation of the analysis, different levels of noise added to the target inputs or outputs that are located in the production frontier help us to identify the appropriate level of noise, which is the level that does not distort the attributes of the original dataset (i.e. does not yield unacceptable input and output values).

The novelty of the SDEANN model is that it introduces the generalized short-term prediction of the optimum outputs or inputs for every sample operational unit towards efficiency attainment. The developed model is deemed to be dynamic because it anticipates all possible future actions of the sample and of the unknown population of the peer operational units by constructing a stochastic optimum production frontier which is tolerant to the perturbations of short-term and missing data. At the same time, the SDEANN model respects the monotonicity assumption, the Pareto efficiency principle, and the dimensionality issues associated with ANN theory with minimum sample size and data requirements.

The rest of the paper is organized as follows. In Section 2, we discuss the results of literature review of DEA (input-oriented variable returns to scale – VRS), ANNs (feed-forward neural networks) and hybrid DEA-ANN models for forecasting. In Section 3, we analyze the SDEANN model, and in Section 4, we apply SDEANN to real data. In Section 5 we elaborate on the managerial implications of the SDEANN model and present our conclusions and recommendations for future research.

2. Literature review

The two methods the SDEANN model draws on (i.e. DEA and ANNs), the properties of short-term forecasting methods, and the optimization methods that rely on DEA and ANNs are discussed below. This review provides brief analyses of variable returns to scale (VRS or BCC) DEA and feed-forward (multi-layer perceptrons - MLPs) ANNs in order to provide insight into their functional underpinning and their potential in efficiency optimization forecasting.

2.1 Data envelopment analysis (DEA)

The DEA, developed by Charnes et al. (1978), is a non-parametric, comparative efficiency measurement technique. Based on linear programming, DEA assesses the relative efficiency status of the sample operational units, or decision making units (DMUs), in order to distinguish the efficient DMUs (efficiency score = 1.000) from the inefficient units (efficiency score < 1.000). The efficient units form the sample optimum production frontier that is the reference set for all the inefficient DMUs. In other words, by applying DEA, each inefficient unit is projected to the reference set in order to specify a customized road map towards efficiency attainment.

The sample used consists of homogeneous peer DMUs that use common inputs to produce common outputs by performing common operations. The homogeneity is not associated with the size of the DMUs, the level of inputs used and the level of outputs produced.

The orientation of the DEA analysis is mainly input-oriented or output-oriented. In the first case, the goal is the determination of the optimum input levels, which are usually lower than the original inputs, while keeping the outputs fixed. In the second case (output-oriented analysis), the goal is output optimization (commonly maximization) with constant inputs. The decision on the orientation of the analysis depends on the features of the market in which the DMUs under evaluation operate, the priorities of the policy makers, the availability of resources, and the controllability of the operational units over the resources.

The simplicity of the application of DEA to real world data and the assumption-free, empirically-identified production function made this method popular among academics and practitioners. Some of the application areas in which DEA has been used are banking, healthcare, and education (Emrouznejad et al., 2008).

The input-oriented BCC DEA model, introduced by Banker et al. (1984), is used in this study. This model assumes that variable returns to scale (VRS) technology underlies the input-to-output transformation process. As a result, the production frontier specified by the BCC DEA model provides a better fit of the sample data than the CCR DEA model, which assumes that constant returns to scale dominate the production technology (Charnes et al., 1978). In addition, the former model is more suitable for use with samples that consist of various size DMUs (Cooper et al., 2007).

The linear programming model developed for the input-oriented BCC DEA model is shown below:

$$
\begin{aligned}
& \theta^* = \min \theta \\
\text{subject to} \quad & \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta x_{io} \quad i = 1, \ldots, m \\
& \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro} \quad r = 1, \ldots, s \\
& \sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0 \qquad j = 1, \ldots, n
\end{aligned} \tag{1}
$$

where DMU_o expresses the unit under assessment of the sample; x_{io} and y_{ro} stand for the *i*th input and the *r*th output of DMU_o, respectively; and *lambdas* (λ_j) denote the non-negative weights of the input and output.

2.2 Artificial neural networks (ANNs)

An artificial neural network (ANN) is a "global" optimization system that draws on real input and output data (Siegelman, 1999; Maqsood and Abraham, 2007). By applying an ANN, an adaptive, "regression-type", best-practice frontier with stochastic underpinnings can be identified (Wang, 2003). This stochasticity is the cornerstone of generalization for the optimum results obtained by the ANNs.

A functional form that connects the inputs and the outputs of the system underlies the best-practice frontier. This implicit functional form is the outcome of a training process of the system, and it is expected to be highly significant to the real-world optimization function. The optimization property of this generalized, "mechanical-learning", functional form is associated with the training specifications determined for the system by the analyst. In economic studies, this optimization functional form expresses the production function.

The ANNs are broken down into three stages or layers (i.e. input, hidden, and output layers) and they consist of multiple nodes that are associated with the input and output data and with the topology specified for the system according to the attributes of the data.

A brief review of the feed-forward multilayer perceptrons (MLPs) that follows is sufficient for the requirements of this study.

2.2.1 Multilayer perceptrons (MLPs)

MLPs are the most popular feed-forward ANNs applied for clustering, ordering, forecasting, and pattern recognition (Mostafa, 2009; Yan et al., 2006; Liang and Wu, 2005). MLPs are "universal approximators" (Hornik et al., 1990), and they identify the most adaptive functional form for linking inputs with outputs of the provided dataset when non-linearity prevails, while simultaneously incorporating a degree of stochasticity. The "universal approximation" property of the MLPs is the key to the generalization of the estimated functional form and also of the inputs and outputs obtained.

The input, hidden, and output layers involved in the optimization process of the MLPs are fully connected in one direction only (Şeker et al., 2003). The relationship between the nodes of the input and hidden layers, and of the nodes between the hidden and output layers depends on the weighting factors computed iteratively during the training phase of the system.

In this context, let $\mathbf{w1}$ and $\mathbf{w2}$ be the weight vectors of the input-hidden layer and of the hidden-output layer respectively (we assume the existence of a single hidden layer), and let $\mathbf{x}$, $\mathbf{h}$ and $\mathbf{y}$, be the input-layer, hidden-layer and output-layer vectors. These vectors consist of the following elements:

$$w1_{ij} \in \mathbf{w1} \;\forall\; i = 1,\ldots,n \;\text{ and }\; j = 1,\ldots,p$$
$$w2_{jr} \in \mathbf{w2} \;\forall\; j = 1,\ldots,p \;\text{ and }\; r = 1,\ldots,s \tag{2}$$

and
$$x_i^{(\psi)} \in \mathbf{x} \;\;\forall\; i = 1,\ldots,n$$
$$h_j^{(\psi)} \in \mathbf{h} \;\;\forall\; j = 1,\ldots,p \tag{3}$$
$$y_r^{(\psi)} \in \mathbf{y} \;\;\forall\; r = 1,\ldots,s$$

where (ψ) denotes the ψ th iteration of the system, n, p, and s are the number of inputs, hidden-layer nodes and outputs of the system, respectively.

The elements of the output layer of the MLP are calculated by applying equation (4):

$$y_r^{(\psi)} = f\left(\sum_{j=1}^{p} w2_{jr} h_j^{(\psi)} + b_r^{(\psi)}\right)$$

or
$$y_r^{(\psi)} = f\left[\sum_{j=1}^{p} w2_{jr}\, f\left(\sum_{i=1}^{n} w1_{ij} x_i^{(\psi)} + b_j^{(\psi)}\right) + b_r^{(\psi)}\right] \tag{4}$$

where $f(\cdot)$ expresses the activation function, and b_r and b_j stand for the biases of the output layer and the hidden layer respectively.

The input and output vectors (i.e. $\mathbf{x}$, $\mathbf{y}$) and also the input and output elements (i.e. x_i, y_r) associated with the ANNs, and particularly with the MLPs, are not in any case the same as the inputs and outputs applied to DEA. To be more precise, in ANN theory, the term "output" is associated with the results obtained by the system and not the output yielded by a production process. Therefore, for ANNs, if the outputs of a production process are provided to the system, then the optimum input levels will be calculated and the latter expresses the outputs of the system. This processing path of the ANNs is similar to the output-oriented analysis of the DEA in which the outputs are given and the minimum levels of inputs are requested. However, in case of the input-oriented analysis, the terms "inputs" and "outputs" that are used both by ANNs and DEA are interpreted evenly.

The selected network that is described by equation (4) complies with the criterion of minimum mean square error (MMSE) which is expressed by equation (5):

$$MSE(w) = \frac{1}{2}\sum_{k=1}^{\varphi}\sum_{r=1}^{s}\left[Y_r^{(\psi)} - y_r^{(\psi)}\right]^2 \tag{5}$$

where $Y_r^{(\psi)}$ is the optimum output at the ψ th iteration.

The MMSE criterion secures the best fit of the MLP functional form to the provided dataset and its generalization potentials.

2.3 Literature related to the synergy between DEA and ANNs

Both DEA and ANNs are commonly regarded as non-parametric techniques. However, they perform well even when statistical errors occur in the dataset provided proving their stochastic underpinnings (Banker, 1993; Banker et al., 1993; Wang, 2003). As a result, the outputs of the analyses of both methods have a limited degree of tolerance to data uncertainties that is not adequate for the purposes of this study, i.e., for forecasting short-term changes in inputs and outputs, and for “broad generalization” purposes. The deviation of the target dataset from the original input and output levels could be significant enough for forecasting the performance of population data based on sample data or for simulating their performance.

Short-term, comparative optimization forecasting of input and output data of peer operational units, and “broad generalization” go beyond the application area of the synergy between DEA and ANN demonstrated in the literature. The studies of joint DEA-ANNs handle the estimation and prediction of efficiency (Yaghoobi et al., 2010; Wu et al., 2006; Wu et al., 2004; Pendhakar and Rodger, 2003; Wang, 2003), and they also provide comparisons of the accuracy of DEA and ANNs in estimating relative efficiency scores (Athanassopoulos and Curram, 1996). In the first case, DEA is applied in first-stage analysis for preprocessing and filtering the actual inputs and outputs and for calculating efficiency scores. ANNs are used in the second-stage analysis for short-term forecasting.

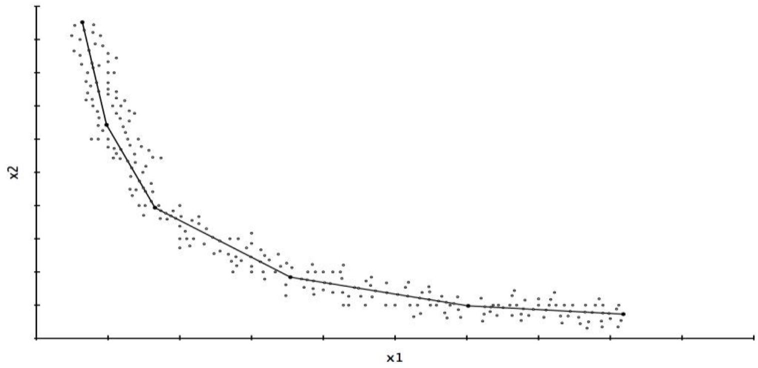
The common DEA-ANN methods for short-term forecasting, which are the methods the SDEANN model is based on, do not respect the Pareto efficiency principle for satisfying the monotonicity assumption. To be more precise, for training the ANN, these methods select both efficient and inefficient units that are arbitrarily regarded as efficient. Consequently, the results obtained by the system are biased and inconsistent with the “regression-type” optimum production frontier. Such a distortion has been suggested for overcoming the dimensionality issues associated with ANNs (Trout et al., 1996) when the available sample size is not adequate. This arbitrariness raises incompatibility issues between DEA and ANNs.

Another limitation of the existing DEA-ANN models for short-term optimization forecasting is that they neglect the estimation of input and output levels for future periods, whereas they put emphasis on forecasting efficiency. However, the ability to forecast inputs and outputs is a prerequisite for successful strategic planning.

3. SDEANN methodology

The SDEANN model relies on the synergy of DEA and ANN for short-term optimization forecasting. The scope of this model goes beyond local maxima based on the sample DMUs and identifies the global maxima providing, in this case, generalized solutions that serve short-term optimization forecasting needs. Our intention is to estimate a stochastic production frontier by modifying the deterministic DEA frontier using statistical noise. By incorporating statistical noise into the DEA production frontier, we test the degree of robustness of our model and form a “tolerant” reference set to errors that has greater adaptability to population data than to sample data (generalization) and is consistent with short-term data variations (Figure 1).

Figure 1. Error-contaminated production frontier



To be more precise, by applying input-oriented or output-oriented VRS DEA, we identify the efficient and inefficient DMUs of the sample (efficiency score = 1.000). Additionally, we specify the target inputs and outputs that lead the inefficient sample units to the reference set. In other words, by first-stage analysis, we uncover the efficient and the "conditionally" efficient DMUs, and we also construct a non-parametric production frontier. In order to estimate a "global" reference set from the available "local" frontier we introduce statistical noise (i.e. additive white Gaussian noise – *awgn*). The impact of the additive noise depends on its power, which is measured either on a linear scale or, as in the present case, in a logarithmic scale (dB). The statistical noise represents the random errors of the production frontier that are associated with its generalized expression. The errors (ε_i) are independently and identically distributed random variables [$\varepsilon_i \sim i.i.d.\, N(\mu, \sigma^2)$] that perturb the deterministic frontier either positively or negatively. Relying on the stochastic frontier, we apply a "regression-type" technique (i.e. ANNs) that allows for short-term forecasting of the optimum input-output mix, even for DMUs that are not part of the evaluation sample. A critical point for the application of the last stage of the SDEANN model is the selection of the appropriate type of ANN (e.g. feed-forward, recurrent), architecture, topology, and computational strategies (e.g. number of hidden layers, number of nodes per hidden layer, training algorithm) in order to specify the most adaptive functional form to the input and output data. We use the least mean square error (MSE) as a performance metric for fitting the specified functional form.

Thus, our purpose is to achieve the highest degree of generalization possible, with a high probability of yielding acceptable results (i.e. $x_i \geq 0 \;\forall\; i = 1,\ldots,n$ or $y_j \geq 0 \;\forall\; j = 1,\ldots,s$).

The SDEANN model is applied either for input-oriented analysis or for output-oriented analysis that respects the priorities of the policymakers of the operational units, the restrictions of the market in which the units operate, and the market trends. In other words, by selecting input-oriented analysis, the policymaker must provide the ANN with outputs for the operational unit in order to predict the optimum input levels associated with the "global" stochastic, best-practice frontier that is applicable for the period from *t* to *t+1*. In case of output-oriented analysis, input values are requested for projecting the optimum output levels for a unit, taking into account every possible movement of its counterpart operational units within the same period (i.e. the period from *t* to *t+1*).

By taking into consideration the stages of analysis incorporated into the SDEANN model, its algorithm runs as follows:

Step1: Apply DEA (BCC) in order to compute the efficiency scores of the sample DMUs and target input and output values for every inefficient sample DMU.

Step 2: Contaminate the DEA-filtered inputs or outputs with statistical noise. The level of perturbation varies according to the magnitude of the data that are subject to noise contamination.

In case of the input-oriented analysis we suggest the contamination of the outputs of the best-practice frontier and vice versa for the output-oriented analysis.

Thus, we estimate a stochastic – noisy production frontier from a deterministic (low-consistent to errors) production frontier.

Step 3: Train the appropriate ANN model with the input-output data of the stochastic production frontier and validate the accuracy of its results using the minimum mean square error (MMSE) criterion.

Step 4: Introduce inputs (output-oriented analysis) or outputs (input-oriented analysis) to the functional form of the SDEANN model in order to predict the "global" optimum short-term outputs or inputs, respectively, for attaining efficiency.

The stochastic production frontier formed by the SDEANN model is highly "tolerant" of short-term data perturbations to the generalization issue of "local" best practices and to uncertainties related to the dataset used. Although the two methods the SDEANN model draws on, i.e. DEA and ANNs, are non-statistical (DEA literature: Cooper et al., 2007; Simar, 2007; Coelli et al., 2005; ANN literature: Liang and Wu, 2005; Hornik et al., 1990), Banker (1993) and Banker et al. (1993) proved that the former technique yields consistent DEA estimators in cases with low measurement errors and when the production function is monotone increasing and concave and when the probability of errors embedded in the transformation process of the inputs to outputs is strictly positive. Additionally, the latter method (i.e. ANNs) is a 'regression-type' technique with stochastic underpinnings (Wang, 2003).

In order to enhance the adaptability of the SDEANN model to noisy settings in which a noise-corrupted production function is possible with unknown magnitude of the error to the best-practice frontier, we import additive white Gaussian noise (*awgn*) either to the optimum output or the optimum inputs, for input-oriented or output-oriented analysis, respectively, located in the 'deterministic' production frontier. The level of *awgn* injected depends on the sensitivity of the sample data. For instance, when applying error-contaminated input data or output data for predicting their output or input pairs, respectively, the predicted values should respect the properties of the original values (e.g. non-negativity). To be more precise, let's assume a deterministic production function:

$$y_j = g(x_i) \tag{6}$$

where y_j expresses the *j*-number outputs ($j = 1, \ldots, s$) and x_i denotes the *i*-number inputs ($i = 1, \ldots, n$). The y_js and the x_is are realizations of the *s*-number and *n*-number population outputs (Y) and inputs (X) that are independently and identically distributed (*i.i.d.*) random variables.

In order to estimate a stochastic production frontier from the 'deterministic' DEA frontier we perturb the original data that compose the reference set by adding random two-sided error $\varepsilon \in \mathbb{R}$ with a normal $f_{\varepsilon}(\cdot)$. As a result, the stochastic production function is written as:

$$\tilde{y}_j^{(l)} = h(x_i) \tag{7}$$

where $$\tilde{y}_j^{(l)} = y_j + \varepsilon_j$$

and l denotes the level of noise added to the data.

Similarly, the input-oriented expression of equation (7) is:

$$\tilde{y}_j^{(l)*} = h(x_i) \tag{8}$$

where $$\tilde{y}_j^{(l)*} = y_j^* + \varepsilon_j$$

Error contamination is applied either to the DEA target outputs or the DEA target inputs in order to form a production function that consists of the maximum number of the available DMUs. This is particularly useful for the third-stage analysis of the SDEANN model where the appropriate ANN model is incorporated. The selected ANN should be trained solely with the input-output bundles of DEA-efficient DMUs to ensure that the SDEANN model can meet a major economic assumption (i.e. monotonicity) (Pendhakar and Rodger, 2003). However, most of the studies associated with DEA and ANN applications use for the training stage of the ANN a mix of efficient DMUs and DMUs that are arbitrarily regarded as efficient DMUs (i.e. DMUs with efficiency scores less than unity) by violating the DEA definition for efficiency (e.g. Yaghoobi et al., 2010; Hu et al., 2008; Wu et al., 2006; Wu et al., 2004; Wang, 2003; Costa and Markellos, 1997). This decision is made in order to overcome the dimensionality weakness of the ANNs, which leads to flaws when the number of DMUs used for the training of the ANN is not greater than ten times the number of the inputs incorporated (Troutt et al., 1996).

In this study, by using DEA target input and output values, particularly, noisy target-input ($\tilde{x}_i$) and target-output bundles, or noisy target-output ($\tilde{y}_j$) and target-input bundles, we address the dimensionality "curse" of the ANNs when the number of the "pure" efficient sample DMUs in conjunction with the number of input variables is not adequate, or when the sample size is small enough to apply ANNs. As a result, we manage to extend the synergy of the DEA and ANNs.

4. Numerical example

4.1 Data description

The SDEANN model is applied to real data collected from the Citizen Service Centers (CSCs) in Greece. CSCs are governmental one-stop shops dedicated to the provision of administrative services to citizens and companies that operate locally in most of the municipalities of the country. The CSCs are homogeneous and independent operational units, the Ministry of the Interiors assesses them annually to determine their efficiency and effectiveness.

The sample of CSCs consists of 100 of the 1020 units that operate in four prefectures of Greece. The sample units serve about 70% of the citizens who apply to CSCs for administrative issues. The number of inputs and outputs used to specify empirically the best-practice frontier are five and three, respectively. The five inputs are the number of full-time employees, weekly hours of work, number of PCs, number of fax machines, and number of printers). The three outputs are the number of electronic protocol registered services provided, the number of manual services provided, and the number of citizens served.

4.2 SDEANN application

By applying the input-oriented BCC DEA to the original dataset, we calculate the efficiency scores of the sample operational units and the target input (DEA-filtered inputs). The input-oriented approach is selected in this study, although, the SDEANN algorithm also is applicable to the output-orientated approach.

According to the second step of the SDEANN algorithm, the DEA-filtered dataset, which consists of the original input and output values of the efficient CSCs as well as the target inputs and outputs of the "conditionally" efficient units, is contaminated with two levels of noise, i.e., with 40dB and 60dB of *awgn*. Two levels of noise are applied to the data in order to test the performance and adaptability of the data, their capacity to yield acceptable results, and consequently, the robustness of our hybrid method. In this study, respecting the input-orientation of the analysis, the statistical noise is added to the outputs that are the "constant" variables of the model. The error contamination of both the inputs and the outputs would distort the original production frontier which finally will refer to an extrinsic dataset other than the given dataset.

The perturbation of the outputs after the addition of the two levels of noise varies. The deviation between the noise-free DEA-filtered outputs and the outputs contaminated by statistical noise is mainly negative for up to 84% of the outputs, indicating that an upward bias is associated with the DEA results (Tables 1 and 2).

Table 1. Deviation of the noise-free DEA-filtered outputs and the DEA-filtered outputs contaminated with additive white Gaussian noise (*awgn* = 40dB)

DMUs	Deviation			DMUs	Deviation		
	eProtocol services	Manual services	Served citizens		eProtocol services	Manual services	Served citizens
1	-0.0700	-0.0948	-0.0719	51	0.0080	-0.0559	-0.0996
2	-0.4218	-0.5037	-0.4211	52	-0.0475	0.0040	0.0407
3	-0.0347	-0.0627	-0.0623	53	0.0040	-0.0045	0.0254
4	-0.5007	-0.5572	-0.5618	54	-0.0655	-0.0696	-0.0642
5	-0.3660	-0.7077	-0.6529	55	0.0043	0.0004	0.0058
6	-0.1627	-0.2300	-0.2167	56	0.0012	-0.0009	-0.0030
7	-0.2116	-0.2547	-0.2114	57	0.0103	-0.0390	0.0156

8	-0.5654	-0.6053	-0.5565	58	-0.0225	-0.0012	0.0149
9	-0.4348	-0.4556	-0.3951	59	0.0756	-0.0207	-0.0218
10	-0.3838	-0.3763	-0.3715	60	-0.6211	-0.7200	-0.6913
11	-0.6992	-0.7648	-0.7580	61	-0.0058	-0.0023	0.0023
12	-0.2085	-0.5103	-0.4666	62	-0.4961	-0.6211	-0.5933
13	-0.5180	-0.7037	-0.5350	63	-0.5079	-0.5079	-0.5119
14	-0.0454	0.1648	-0.0452	64	-0.0001	0.0018	-0.0078
15	-0.1795	-0.2222	-0.1933	65	-0.0075	0.0389	0.0193
16	-0.7656	-0.7788	-0.7967	66	-0.5011	-0.5028	-0.4926
17	0.0051	0.0013	-0.0028	67	-0.0063	0.0115	-0.0017
18	-0.4772	-0.6163	-0.5877	68	-0.0035	-0.0066	-0.0067
19	0.0139	-0.0014	0.0060	69	-0.5824	-0.6021	-0.5870
20	-0.0016	0.0016	-0.0023	70	0.0003	-0.0027	-0.0109
21	0.2807	0.2047	-0.0627	71	-0.3946	-0.3894	-0.4312
22	-0.1855	-0.5057	-0.3039	72	-0.0042	0.0024	0.0014
23	-0.0014	0.0031	-0.0005	73	-0.5595	-0.5695	-0.5429
24	-0.0054	0.0006	0.0016	74	-0.5983	-0.5950	-0.6815
25	-0.4755	-0.4766	-0.5099	75	0.0277	-0.0110	-0.0078
26	-0.0011	0.0065	0.0019	76	-0.5551	-0.5878	-0.5650
27	-0.0019	0.0037	0.0007	77	-0.0028	-0.0066	0.0030
28	-0.0016	0.0004	0.0048	78	-0.0125	0.0239	-0.1045
29	0.0090	0.0037	-0.0021	79	-0.0057	-0.0009	-0.0007
30	-0.0019	0.0023	-0.0008	80	-0.5713	-0.6459	-0.6222
31	0.0097	-0.0135	-0.0134	81	-0.5328	-0.5389	-0.5129
32	-0.0013	-0.0001	-0.0008	82	0.0001	-0.0031	0.0012
33	-0.7734	-0.7754	-0.8335	83	-0.0419	0.1152	-0.0430
34	-0.6347	-0.6673	-0.6646	84	-0.2694	-0.4518	-0.3484
35	0.0045	0.0145	0.0150	85	-0.4335	-0.4290	-0.4178
36	-0.5835	-0.8072	-0.7759	86	-0.0061	0.0030	-0.0018
37	-0.0111	0.0011	0.0407	87	-0.1889	-0.1952	-0.1923
38	-0.0009	0.0001	0.0001	88	-0.6237	-0.6546	-0.6551
39	-0.5414	-0.5679	-0.5605	89	-0.0064	0.1433	-0.0339
40	-0.0580	-0.0614	-0.0699	90	0.0581	0.0476	-0.0028
41	-0.0026	-0.0003	-0.0010	91	-0.6261	-0.6911	-0.6821
42	-0.4912	-0.6774	-0.6452	92	-0.0189	0.0907	0.0215
43	-0.4540	-0.4006	-0.4277	93	0.0473	-0.0186	-0.0207
44	-0.6368	-0.7942	-0.7625	94	0.0313	-0.0257	-0.0526
45	-0.4121	-0.4732	-0.4491	95	-0.0435	0.0254	0.0064
46	-0.2587	-0.3127	-0.3780	96	-0.0952	-0.0468	-0.0145
47	-0.5177	-0.6640	-0.6069	97	-0.4689	-0.4742	-0.4321
48	-0.4214	-0.4023	-0.3590	98	0.0211	0.0299	-0.0456
49	0.0028	-0.1700	-0.0277	99	-0.3502	-0.3556	-0.3505
50	-0.6216	-0.6880	-0.6812	100	0.0261	0.0162	-0.1208

Table 2. Deviation of the noise-free DEA-filtered outputs and the DEA-filtered outputs contaminated with additive white Gaussian noise (*awgn* = 60dB)

DMUs	Deviation			DMUs	Deviation		
	eProtocol services	Manual services	Served citizens		eProtocol services	Manual services	Served citizens
1	-0.0096	-0.1172	-0.0329	51	0.7472	-0.1436	-0.3033
2	-0.4182	-0.4942	-0.4236	52	0.7870	-0.2581	0.9977
3	0.0106	-0.0689	-0.0402	53	0.0284	-0.0057	-0.0102
4	-0.5156	-0.5574	-0.5759	54	-0.1165	-0.1357	-0.0425
5	-0.3694	-0.7131	-0.6647	55	0.0718	-0.0162	-0.0645
6	-0.1495	-0.2327	-0.2322	56	-0.0226	-0.0135	-0.0215
7	-0.1978	-0.2539	-0.2047	57	-0.1170	-0.2397	-0.4404
8	-0.6011	-0.6096	-0.5637	58	0.3460	0.1818	0.1678
9	-0.4809	-0.4315	-0.4040	59	-0.0848	-0.0806	0.1045
10	-0.4846	-0.3046	-0.3381	60	-0.6944	-0.7155	-0.7017
11	-0.8049	-0.7443	-0.8295	61	0.0320	0.0520	0.0312
12	-0.1964	-0.5261	-0.4947	62	-0.4786	-0.6210	-0.6037
13	-0.5260	-0.7161	-0.5107	63	-0.5065	-0.5034	-0.5071
14	-0.0635	0.1308	-0.0253	64	0.0317	0.0304	0.0771
15	-0.2603	-0.1979	-0.2579	65	0.0682	0.0362	0.0204
16	-0.7388	-0.7247	-0.8586	66	-0.4571	-0.4853	-0.5339
17	0.0250	0.0095	-0.0213	67	-0.0519	-0.0720	-0.0863
18	-0.4790	-0.6276	-0.5818	68	0.0426	0.0310	-0.0095
19	0.1379	0.0150	0.1026	69	-0.5788	-0.6234	-0.5923
20	-0.0410	-0.0326	0.0038	70	-0.0133	-0.0374	-0.0293
21	0.2952	0.1659	-0.0954	71	-0.4108	-0.3780	-0.4429
22	-0.1693	-0.4848	-0.2882	72	0.0220	-0.0222	-0.0279
23	0.0222	-0.0297	-0.0349	73	-0.5640	-0.5644	-0.6201
24	-0.0454	0.0082	-0.0374	74	-0.5575	-0.6129	-0.6553
25	-0.4554	-0.4393	-0.5329	75	-0.0049	-0.0025	0.0010
26	-0.0205	0.0127	-0.0755	76	-0.6154	-0.6015	-0.5711
27	-0.0395	0.0377	-0.0236	77	-0.0691	0.0271	-0.0730
28	-0.0486	-0.0262	-0.0329	78	0.2443	0.2729	0.7644
29	0.0947	0.0377	-0.0288	79	-0.0064	0.0427	0.0118
30	-0.0535	0.1177	-0.0269	80	-0.5780	-0.6505	-0.6320
31	-0.0147	0.0557	-0.0467	81	-0.3444	-0.6605	-0.5144
32	-0.0286	-0.0048	-0.0028	82	0.0097	-0.0168	-0.0074
33	-0.7894	-0.7806	-0.7962	83	-0.0374	0.1533	-0.1110
34	-0.7839	-0.6816	-0.6390	84	-0.2198	-0.4721	-0.3555
35	-0.2784	0.0955	0.0175	85	-0.4584	-0.4144	-0.4376
36	-0.6250	-0.8034	-0.7933	86	0.0259	0.0046	-0.0298
37	-0.2728	0.0699	-0.3467	87	-0.1655	-0.1677	-0.2280
38	-0.0042	-0.0006	-0.0097	88	-0.6511	-0.7426	-0.6784
39	-0.5809	-0.5892	-0.3581	89	0.0059	0.1385	-0.1542
40	0.0258	-0.0774	-0.0287	90	-0.0055	0.0034	0.0459
41	-0.0338	0.0136	0.0159	91	-0.7323	-0.5342	-0.6375
42	-0.5301	-0.7390	-0.7464	92	0.0335	0.2684	-0.0090

43	-0.5232	-0.4060	-0.4800	93	0.1069	0.0578	-0.0229
44	-0.6761	-0.7858	-0.7263	94	0.9120	-0.3513	-0.4477
45	-0.4870	-0.4878	-0.3734	95	-0.2095	-0.1421	0.4014
46	-0.2472	-0.3027	-0.4270	96	0.0911	-0.1131	-0.0720
47	-0.4939	-0.6635	-0.6155	97	-0.4773	-0.4063	-0.4758
48	-0.4126	-0.4231	-0.3621	98	**-1.0179**	-0.1156	0.8663
49	-0.4124	0.1590	**-1.3432**	99	-0.3714	-0.3384	-0.3584
50	-0.7941	-0.7147	-0.5838	100	-0.9431	**-1.0538**	-0.3855

When higher-level statistical noise (i.e. *awgn* = 60dB) is added to the original (DEA-filtered) dataset for estimating the stochastic production frontier from the empirical deterministic frontier, the deviation between the noise-free and the noise-contaminated outputs in three cases is negative, and its absolute value is greater than unity (See the bold values in Table 2). These usually high deviations, compared with the other deviations in Table 2, appear to be due to the negative outputs produced by the addition of random *awgn*. In this case, it is obvious that an *awgn* value of 60dB is too high for the original production function because it leads to unaccepted outputs. Thus, the addition of *awgn* = 40dB is a fair level of noise for this particular case in that it respects the attributes of the data without being regarded as a low-significance intervention.

The training phase of the ANN model is crucial for the accuracy of the projected results of the short-term, comparative optimization forecasting model. The stochastic optimum inputs and outputs of the whole sample that are used for the training of the appropriate ANN provide the best fit of the SDEANN model to the generalized production frontier. Additionally, they enhance the forecasting properties of this model to short-term data changes while keeping the sample size requirements short.

For the training phase of the ANN, 70 CSCs were selected randomly from the sample. This phase included the training (using 80% of the selected units), the validation (using 10% of the selected units) and the testing of the network (using 10% of the selected units).

After experimentation with various ANN models, architectures, topologies, and training functions, we concluded that the most appropriate network for our dataset and for the input-orientation of our analysis was the *Levenberg-Marquardt* network (Table 3). Relying on this ANN model, we identified the most statistically significant functional form for both noise-embedded production frontiers (*awgn* = 40dB and *awgn* = 60dB) with the minimum MSE.

Table 3. ANN properties of the input-oriented SDEANN model

Properties	**Level of additive statistical noise**	
	40dB	60dB
Inputs	3	3
Hidden Layer(s)	1	1
Neurons	6	6
Outputs	5	5
Training Function	Levenberg-Marquardt	Levenberg-Marquardt
Mean Square Error	10^1	4.3491
R		

Training	0.9766	0.9861
Validation	0.9794	0.9880

Figure 2. Selected ANN's fitting to the stochastic production frontier - training phase (*awgn* = 40dB)

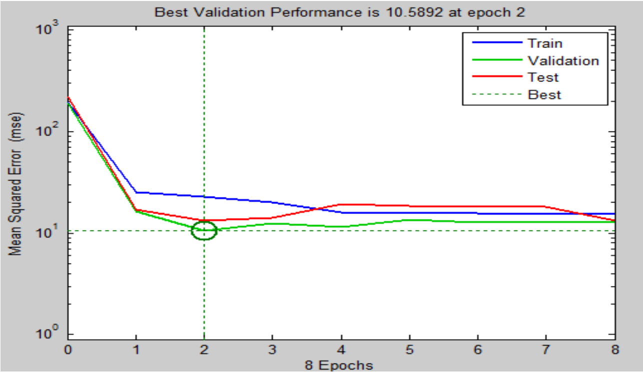
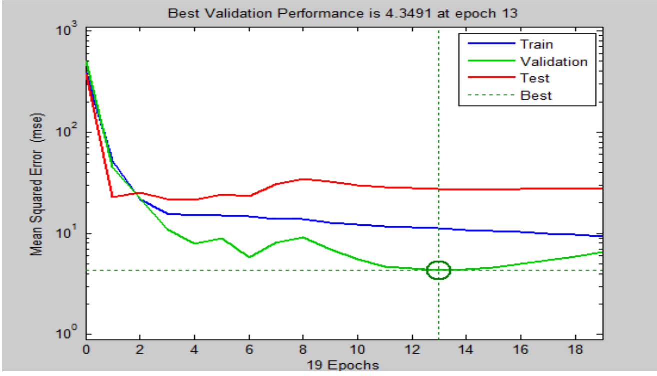


Figure 3. Selected ANN's fitting to the stochastic production frontier – training phase (*awgn* = 60dB)



Following the training process of the ANN, we introduced the noisy outputs of the 30 remaining operational units to the SDEANN model in order to test the projected inputs against their DEA-filtered counterparts. The majority of the inputs projected by the SDEANN model have moved upward, comparing to the DEA target inputs, for cases of noise contamination (i.e., for 40dB, 56.7% of the inputs were moved upward, whereas, for 60dB, 62.7% of the inputs were moved upward) (Tables 4 and 5).

Table 4. Deviation of the DEA-filtered and the SDEANN (*awgn*=40dB) input values (input-oriented analysis)

DMUs	Deviation					DMUs	Deviation				
	Full-time employees	Working hours	PC	Fax	Printers		Full-time employees	Working hours	PC	Fax	Printers
71	-0.1795	-0.0317	0.0168	0.0632	-0.0046	86	-0.2984	-0.2095	0.4648	-0.5965	-0.7931
72	0.0128	-0.0345	0.6425	-0.8436	-0.7132	87	-0.1054	-0.1079	0.0934	0.0203	-0.0483
73	0.1581	-0.0814	0.0183	0.9633	0.8613	88	0.2261	-0.0207	0.1028	1.1574	0.9396
74	0.0217	0.1280	-0.0580	1.0641	0.1794	89	-0.1924	-0.0245	-0.0608	1155.7527	-0.2044
75	-0.2460	-0.0627	0.1310	-0.0551	0.1483	90	-0.1803	-0.0902	0.0447	1128.7619	0.5358
76	-0.0071	0.0567	-0.0977	1.0343	0.4034	91	0.4530	-0.1524	-0.1181	1.0897	0.6166
77	-0.3064	-0.1882	0.5029	1048.3629	-0.3166	92	0.5511	-0.1055	-0.3864	0.9811	1.6491
78	0.5760	-0.3910	-0.2615	1742.3385	1.7373	93	-0.3194	0.0484	-0.1467	0.8929	0.8136
79	-0.0395	-0.1698	0.2783	742.0480	-0.3059	94	0.5516	-0.2022	-0.1886	0.7008	1.6535
80	0.1683	-0.0960	-0.3070	0.5193	0.6998	95	0.5541	-0.1085	-0.3913	1.8714	1.0505
81	0.3845	-0.2066	0.0525	1.0029	1.5354	96	0.5844	-0.0365	-0.4623	1755.4599	1.7633
82	0.2272	-0.1595	0.2400	-0.4489	-0.7516	97	-0.0549	0.2034	-0.4123	3.0526	0.3079
83	-0.1565	-0.1375	-0.0193	898.9825	-0.4426	98	0.5790	-0.1915	0.0953	0.7481	1.7467
84	-0.2711	0.0611	0.0692	0.2632	-0.3445	99	-0.2749	0.0570	0.1815	0.3989	0.2163
85	-0.2001	0.0037	0.0516	0.5824	0.0674	100	0.5785	-0.0909	0.0934	0.7479	1.7458

Table 5. Deviation of the DEA-filtered and the SDEANN (*awgn*=60dB) input values (input-oriented analysis)

DMUs	Deviation					DMUs	Deviation				
	Full-time employees	Working hours	PC	Fax	Printers		Full-time employees	Working hours	PC	Fax	Printers
71	0.0543	-0.0019	0.1401	-0.1196	0.4148	86	-0.0673	-0.1326	0.5831	-0.6851	-0.5197
72	-0.1203	0.1125	0.3596	-0.6135	1.1233	87	0.2313	0.0015	0.2062	-0.0850	0.8785
73	0.1554	0.0190	0.2066	0.0112	0.3725	88	0.0640	0.1094	0.4146	-0.3870	0.2095
74	0.0645	0.2454	0.0972	0.2113	-0.0903	89	0.0039	0.0396	-0.0624	1179.5034	-0.0077
75	0.0283	-0.0077	0.2099	-0.1317	1.0493	90	0.0416	-0.1422	0.1082	917.5272	0.8426
76	0.0073	-0.0229	-0.0204	0.1709	0.1606	91	0.0748	-0.0628	0.2436	-0.7474	-0.0972
77	-0.0997	-0.2079	0.4567	1075.5911	-0.0415	92	0.1943	0.1445	-0.0359	-0.7608	0.4308
78	0.1064	-0.2501	0.2541	109.6907	0.3899	93	-0.1969	0.0849	-0.1081	0.6095	0.8076
79	0.1938	-0.0807	0.3113	725.3382	1.0651	94	0.4081	0.1938	0.3471	-0.5442	0.5216
80	0.1930	-0.0385	-0.2117	-0.1441	0.3214	95	0.0548	-0.0189	-0.0652	-0.8449	0.0545
81	0.1251	-0.0473	0.6274	-0.5119	0.4355	96	0.1141	0.2310	-0.0676	86.2006	0.3883
82	0.2528	-0.1037	0.1534	-0.3522	1.0182	97	-0.1189	0.2981	-0.2677	0.7280	-0.1048
83	0.1190	-0.0072	0.0464	946.3375	0.1814	98	-0.1306	-0.1984	0.7206	**-1.1721**	0.2761
84	0.0494	0.1780	0.1613	0.2145	0.1491	99	-0.1932	0.0122	0.3446	-0.0264	0.2122
85	-0.0229	-0.0143	0.1936	0.1874	0.2545	100	-0.0491	0.0291	0.7387	**-1.1466**	0.3077

Significant increases appear at the fax machines for the units 77, 78, 79, 83, 89, 90, and 96, when *awgn* of 40dB was applied, because of the doubling of the values projected. Namely, in all these cases, the DEA-filtered number of fax machines was null, whereas the SDEANN model yielded unity as the optimum generalized number for the same variable. Similar changes are occurred for the fax machines when an *awgn* of 60dB was added.

The consistency of the results obtained by the SDEANN model when an *awgn* of 40dB was incorporated was validated in Table 4 while all the projected inputs respected the attributes of the original data. Unlike the 40dB SDEANN model, the 60dB expression yielded distorted results for DMUs 98 and 100 for the fax machines.

The two deviations (Table 5) express the negative input values calculated by the 60dB SDEANN model. As a result, statistical noise at this level was not accepted for the available dataset.

By introducing any output (input-oriented approach) or input values (output-oriented approach) to the appropriately trained SDEANN model, taking into consideration the attributes of the case applied, optimum inputs or outputs, respectively, were obtained that were valid for the period starting from time t and ending at the future time of $t+1$.

5. Concluding remarks and further research

The goal of this research was to develop a short-term, stochastic optimization forecasting model. The purpose of the model is to assist the decision makers in operational units in forecasting optimum production equilibria, preparing a strategic plan, and restructuring their units when there is a lack of information, or when only, imprecise information is available about their rivals. The SDEANN model anticipates possible deviations of the sample production frontier due to missing information for the population or due to uncertainties associated with the dataset in order to yield the generalized optimum input or output levels, depending on the orientation selected.

Our model draws on a stochastic production frontier for training the appropriate ANN to the case set each time without any assumptions concerning the production function, which is consistent with stochastic methods (e.g. Stochastic Frontier Analysis). The stochastic production frontier is a 'deterministic' empirical production frontier specified by DEA that is contaminated with statistical noise. Therefore, the training material for the ANN is real data that express the attributes of the case.

The SDEANN model respects the monotonicity assumption that prevails in microeconomic theory and DEA theory for efficiency by using solely efficient operational units for the estimation of the generalized production frontier. Additionally, it addresses the dimensionality issues associated with ANNs with minimum sample size so that its applicability can be extended to sectors with small number of operational units. Associated with the minimum sample size requirements for applying the SDEANN model is the significant reduction of computational complexity that is particularly important in real-world problems.

Further research is needed to define a pattern for adjusting the statistical noise injected to the production frontier according to the attributes of the dataset used in order to prevent distortion of the inputs or outputs and to preclude the production of faulty SDEANN results. Another extension of the SDEANN methodology could be the "global" long-term comparative optimization forecasting.

References

Athanassopoulos, A., Curram, S., 1996. A comparison of Data Envelopment Analysis and artificial neural networks as tools for assessing the efficiency of decision making units. Journal of the Operational Research Society 47, 1000-1016.

Banker, R.D., 1993. Maximum likelihood, consistency and Data Envelopment Analysis: A statistical foundation. Management Science 39, 1265-1273.

Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. Management Science 30, 1078-1092.

Banker, R.D., Gadh, V.M., Gorr, W.L., 1993. A Monte Carlo comparison of two production frontier estimation methods: Corrected ordinary least squares and Data Envelopment Analysis. European Journal of Operational Research 67, 332-343.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European Journal of Operational Research 2, 429-444.

Coelli, T.J., Rao, P., O'Donnell, C.J., Battese, G., 2005. An Introduction to Efficiency and Productivity Analysis (2nd ed.). Springer, New York.

Cooper, W.W., Seiford, L.M., Tone, K., 2007. Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-solver Software (2nd ed.). Springer, New York.

Costa, A., Markellos, R.N., 1997. Evaluating public transport efficiency with neural network models. Transportation Research Part C: Emerging Technologies 5, 301-312.

Emrouznejad, A., Parker, B., Tavares, G., 2008. Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. Journal of Socio-Economics Planning Science 42(3), 151-157.

Ferlie, E., Lynn, L., Pollitt C., 2007. The Oxford Handbook of Public Management. Oxford University Press, New York.

Hornik, K., Stinchcombe, M., White, H., 1990. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Networks 3, 551-560.

Hu, S.C., Chung, Y.K., Chen, Y.S., 2008. Using Hopfield neural networks to solve DEA problems. IEEE International Conference on Cybernetics and Intelligence Systems, 606-611.

Liang, L., Wu, D., 2005. An application of pattern recognition on scoring Chinese corporations financial conditions based on backpropagation neural network. Computers and Operations Research 32, 1115-1129.

Maqsood, I., Abraham, A., 2007. Weather analysis using ensemble of connectionist learning paradigms. Applied Soft Computing 7, 995-1004.

Mostafa, M.M., 2009. Modeling the efficiency of top Arab banks: A DEA-neural network approach. Expert Systems with Applications 36, 309-320.

Pendhakar, P., Rodger, J., 2003. Technical efficiency-based selection of learning cases to improve forecasting accuracy of neural networks under monotonicity assumption. Decision Support Systems 36, 117-136.

Pollitt, C., Bouckaert, G., 2004. Public Management Reform: A Comparative Analysis (2nd ed.). Oxford University Press, New York.

Şeker, S., Ayaz, E., Türkcan, E., 2003. Elman's recurrent neural network applications to condition monitoring in nuclear power plant and rotating machinery. Engineering Application of Artificial Intelligence 16, 647-656.

Siegelmann, H.T., 1999. Neural Networks and Analog Computation: Beyond the Turing Limit. Birkhäuser, Boston.

Simar, L., 2007. How to improve the performances of DEA/FDH estimators in the presence of noise?. Journal of Productivity Analysis 28, 183-201.

Troutt, M.D., Rai, A., Zhang, A., 1996. The potential use of DEA for credit applicant acceptance systems. Computers & Operations Research 4, 405–408.

Wang, S., 2003. Adaptive non-parametric efficiency frontier analysis: A neural-network-based model. Computers & Operations Research 30, 279-295.

Worthington, A., Dollery, E., 2000. Measuring efficiency in local governments planning and regulatory function. Public Productivity & Management Review 23, 469-485.

Wu, C., Chen, X., Yang, Y., 2004. Decision-making modeling method based on artificial neural network and data envelopment analysis. IEEE International Conference on Geoscience and Remote Sensing Symposium.

Wu, D., Yang, Z., Liang, L., 2006. Using DEA-neural network approach to evaluate branch efficiency of a large Canadian bank. Expert Systems with Applications 31, 108-115.

Yaghoobi, R., Aryanezhad, M.B., Hosseinzadeh Lotfi, F., 2010. Application of multi-layer recurrent neural network and fuzzy time series in input/output prediction of DEA models: Real Case Study of a Commercial Bank. IEEE International Conference on Computers and Industrial Engineering, 1-6.

Yan, H., Jiang, Y., Zheng, J., Peng, C., Li, Q., 2006. A multilayer perceptron-based medical decision support system for heart disease diagnosis. Expert Systems with Applications 30, 272-281.