Int. J. Production Economics 183 (2017) 159–170

Contents lists available at ScienceDirect

Int. J. Production Economics

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

Exploring the predictive potential of artificial neural networks in conjunction with DEA in railroad performance modeling

He-Boong Kwon

Hasan School of Business, Colorado State University-Pueblo, 2200 Bonforte Blvd., Pueblo, CO 81001, USA

A R T I C L E I N F O

Keywords:
Artificial neural network
Benchmarking
Data envelopment analysis
Class I railroads

A B S T R A C T

This paper is an investigation into the feasibility of using artificial neural networks (ANN) in conjunction with data envelopment analysis (DEA) for performance measurement and prediction modeling of Class I railroads in the United States. For this exploratory study, DEA-ANN are combined into a two-stage modeling approach. While it is frequently used as a benchmarking tool, DEA lacks predictive capabilities. However, ANN has strong nonlinear mapping and adaptive prediction functionality. In this study, the advantages of combining these complementary methods into an integrated performance measurement and prediction model are explored. For this combined approach, a Charnes, Cooper and Rhodes (CCR) DEA model is used to evaluate the efficiency of each decision making unit (DMU) and to capture the efficiency trend of each railroad. Based upon those DEA results, the follow-on backpropagation neural network (BPNN) model predicts an efficiency score and target output for each DMU. This is a new attempt to extend the BPNN model for purposes of best performance prediction. The resulting framework is an effective benchmarking and decision support system which adds adaptive prediction capabilities to current benchmarking practices.

1. Introduction

Freight railroads in the U.S are growing and regaining momentum after suffering a decline in 2009 as a result of the economic recession. Overall transportation volume is expected to grow in the upcoming years, and the industry anticipates a 50% increase of rail freight by 2040. This optimistic forecast is motivating railroads to reassess their long term plans to prepare for growth in future demand, and has inspired the commitment of additional capital to their existing infrastructure to better meet current and foreseeable service needs. The record high level of capital investments in 2013 reflects these hopeful growth prospects in the rail industry (Global Railroads Industry Profile, 2013; Morris, 2013). Freight railroads are one of the most capital intense industries in the U.S and invest vast amounts of capital to retain necessary capacity. In fact, more than 40% of their revenue dollars are reinvested for maintaining and improving rail capacity which exceeds the average manufacturing rate by a factor of five times (AAR, 2012, 2013a, 2013b).

As of 2013, seven Class I railroads dominated the U.S. freight transportation industry by representing 69% of total rail mileage, 90% of employees, and 94% of revenue (AAR, 2013c). These railroads privately own most of the rail infrastructure (i.e., tracks, locomotives, freight wagons, and etc.) and these assets are operated by their own labor forces. In this vertically integrated, capital intense environment, the efficient utilization of infrastructure resources has been a major managerial concern in rail operations. Furthermore, when considering the substantial commitment of additional resources for capacity expansion, sustaining a high level of efficiency in current operations is a precondition for the healthy growth of a railroad. From this perspective, efficiency and productivity have not only been of keen interest to railroad managers but also an important research topic in railroad studies in the U.S. and other countries as well (AAR, 2012; Bitzan and Keeler, 2003; Himola, 2007; Lim and Lovell, 2008; Shi et al., 2011). In their productivity analysis of Class I railroads during 1996–2003 period, Lim and Lovell (2008) pointed out the underlying inefficiencies in rail operations and addressed the necessity of utilizing a benchmarking approach to improve rail efficiency. Subsequent analysis of railroad operations between 2002 and 2007 inclusive still shows varying efficiency levels among Class I railroads, which implies the potential for significant improvements for most railroads (Shi et al., 2011). The logical question following this discussion will be on '*how much improvement should an inefficient railroad make to become superior to their peer entities?*' The managers of inefficient firms will have subsequent questions including '*Can I set target performance in terms of direct output beyond efficiency*?' and '*Is it feasible to monitor progress by testing 'what-if' scenarios during implementation stages?*' These questions pose the need for a model that possesses predictive capabilities. In railroad performance analysis studies, despite success-

E-mail addresses: heboong.kwon@csupueblo.edu, palmbay95@hotmail.com.

http://dx.doi.org/10.1016/j.ijpe.2016.10.022
Received 19 September 2013; Received in revised form 2 May 2016; Accepted 28 October 2016
Available online 04 November 2016

ful modeling to measure the comparative performance of competing entities, efforts have seldom been made to incorporate prediction capability into the performance measurement framework. In reality, managers of inefficient firms are eager to know how much improvement needs to be made to achieve an efficient level of operation. Therefore, the capability to establish specific and actionable performance targets is a practical necessity for managers pursuing improvement during their planning stages. Besides, using an adaptive prediction capability to test hypothetical scenarios would be a valuable addition that would facilitate the monitoring process during the implementation stage. In this sense, an integrated performance management system is becoming a business imperative in railroads and other industries as well.

In an effort to implement a feasible solution for this challenging need, this study aims at developing a performance measurement and prediction model for Class I railroads by utilizing time series data covering the period of 2005–2012 inclusive. The combined Data Envelopment Analysis (DEA)-Artificial Neural Network (ANN) model is presented as an integrated benchmarking tool. DEA, as a linear programming based benchmarking method, has been widely used to assess the efficiency of competing entities, known as Decision Making Units (DMUs). The resulting DEA efficiency has been commonly accepted as an indirect performance measure in benchmarking studies. Despite its wide range of successful applications, DEA as an extreme point method lacks predictive capability. In comparison, ANN as an artificial intelligence information processing system is capable of learning nonlinear functional relationships and underlying patterns between variables. Consequently, its strength is in its predictive capability. Accordingly, exploration of the advantages of combining these two complementary methods is the focus of this attempt to develop an integrated performance benchmarking system for railroads. Most previous studies have used DEA as a standalone technique to measure efficiency scores, and ANN applications in railroad performance are a rarity. This study, to the best of the author's knowledge, is the first attempt to integrate DEA and ANN for railroad performance analysis.

This paper proposes a new approach to predict best performance output beyond the surrogate measure of the efficiency index. This paper is distinguished from previous DEA studies in the following ways: 1) DEA is used to evaluate the industry and to generate a data set for ANN analysis as a front end preprocessor. 2) ANN, as a post learning and prediction processor, is used to predict DEA efficiency. 3) A new attempt is made to predict best performance through the approximation of efficient frontiers. In so doing, this study extends the use of ANNs to benchmarking studies by proposing the combined method as a promising performance modeling tool for railroads, as well as other industries.

This paper is organized as follows. Section 2 reviews related studies. In Section 3, research methodology, including DEA and ANN, is introduced. Section 4 discusses the empirical analysis results by describing the data sets, variables, proposed framework, and analytic results of DEA and ANN. Concluding remarks and future research directions are presented in Section 5.

2. Related studies

2.1. Railroad efficiency measurement using DEA

DEA has been widely accepted as a benchmarking tool and its application covers a broad range of industry sectors, and a variety of functional areas of management, including railroad performance measurement. Recent research by Liu et al., (2013a, 2013b) comprehensively summarizes the DEA literature and reports the evolution of its application by indicating that transportation and logistics is one of the consistently growing areas utilizing DEA. In this capital intensive and cost conscious industry, the efficient utilization of existing infrastructure and cost elements has been a key managerial concern. These variables have commonly been used as key inputs to measure DEA efficiency in generating railroad output.

In an analysis to identify best practices for Class I railroads, Shi et al. (2011) used aggregated measures of infrastructure (e.g., locomotives, freight cars, way and structure, and number of employees) and fuel as DEA inputs by utilizing revenue per ton-mile (RTM) as a measure of single quantity output. In European rail applications, Himola (2007) utilized infrastructure inputs (e.g., freight wagons, total route km, total locomotives, and staff) by using two separate outputs, freight ton-km (FTK) and freight tons (FTONS). Infrastructure has consistently been a popular DEA input for railroad benchmarking studies (Abate et al., 2013; Coelli and Perelman, 2000, 1999; Himola, 2007; Salerian and Chan, 2005). Another stream of railroad studies has focused on cost efficiency measurement, utilizing cost and expense variables such as labor, fuel, material, and maintenance (Andrikopoulos and Loizides, 1998; Merkert et al., 2010; Parisio, 1999). The combined utilization of infrastructure and cost variables also enables rich contextual analysis by capturing industry specific characteristics (Chapin and Schmidt, 1999; George and Rangaraj, 2008; Lim et al., 2009; Shi et al., 2011).

In this study, infrastructure resources are selected as a critical input that impacts rail performance. The excess capacity resulting from underutilized resources is wasteful, and retards healthy growth for capital intensive railroads. Therefore, capitalizing on current resources and achieving operational efficiency is a primary concern for rail managers. At the same time, railroads need to preserve an appropriate scale of infrastructure so as to accommodate their foreseeable growth in demand, which requires a level of capital commitment that carries the potential cost of undermining efficiency. From this perspective, managerial focus should be centered on balancing the three key requirements of enhancing internal efficiency, increasing external demand, and investing capital, accordingly (Lim and Lovell, 2009). From this perspective, the prediction of the level of targeted performance needed to become efficient and superior to one's rivals is a crucial step toward setting improvement goals in operations, marketing, finances, and other strategically critical areas. For railway outputs, two-dimensional quantity-distance measures such as RTM (or FTK in Europe) has been the preferred metric and, often times, a single dimensional output, such as FTONS, has been used as a complementary or alternative quantity measure (Coelli and Perelman, 2000; Himola, 2007; O'Mahony and Oulton, 2000; Shi et al., 2011). This study utilizes RTM and FTONS as a single rail output over four infrastructure inputs which include staff, locomotives, freight cars, and miles of track. Table 1 displays selected railroad studies using DEA and summarizes the variables used for each study.

2.2. A combined DEA and ANN approach

ANN, as a complementary method to DEA, excels at estimation because of its pattern learning and mapping capabilities. A review of the literature shows the recent growth of ANN applications in various forecasting areas such as bankruptcy (Kim, 2011), customer demand (Kourentzes, 2013), exchange rates (Sermpinis et al., 2013), manufacturing costs (Cavalieria et al., 2004; Chou, 2010), product development (Parry et al., 2011; Thieme, 2000), project length (Li and Liu, 2012), sales (Kumar and Mittal, 2012), and quality (Carlucci et al., 2013; Kuo et al., 2013). Despite its proven capability and its potential for application, the literature shows limited use of ANN for performance benchmarking, and a very limited number of papers report combined DEA and ANN applications. Athanassopoulos and Curram (1996) first attempted to compare DEA and ANN as an alternative tool for measuring efficiency, and a subsequent stream of research hinted at the potential benefits of combining these two methods (Liu et al., 2013; Liao et al., 2007; Santin et al., 2004; Santin, 2008; Wang, 2003). These explorative studies stimulated complementary usage of both methods

Table 1
Selected DEA applications to railroad efficiency measurement.

Authors	Input variables (categories or elements)														Output variables							Data
	Staff	Loco motives	Freight cars	Passenger cars	Rolling stock	Track	Land	Horse power	Labor	Fuel	Energy	Material	Operating Expenses	Repair	Track	Revenue Ton Mile (RTM)	Price per RTM	Freight Ton Kilometer	Freight Tons	Freight miles	Passenger Kilometers	
Abate et al. (2013)	•	•		•		•															•	7 Euro-railways 1997-2007
Shi et. al (2011)	•	•	•			•				•		•				•						US Class I railroads 2002-2007
Merkert et al. (2010)	•											•						•			•	43Euro-railways 2006-2007
Lim and Lovell (2008, 2009)	•					•			•	•		•				•	•					US Class I railroads 1996-2003
George and Rangaraj (2008)	•		•	•		•		•					•							•	•	Indian railways 2004-2005
Himola (2007)	•	•	•			•												•	•			31 Euro-railways 1980-2003
Salerian and Chan (2005)	•	•	•	•		•												•			•	20 national railways 1990-1998
Coelli and Perelman (2000)	•				•	•												•			•	17 Euro-railways 1988-1993
Chapin and Schmidt (1999)	•		•			•		•		•				•	•						•	US Class I railroads 1980-1993
Coelli and Perelman (1999)	•				•	•												•			•	17 Euro-railways 1988-1993
Cowie (1999)	•	•		•			•														•	57 railways in Swiss 1995
Parisio (1999)									•		•	•						•			•	8 Euro-railways 1973-1989
Andrikopoulos and Loizides (1998)									•		•		•							•	•	10 Euro-railways 1969-1993

and a handful of research outcomes reported advantages of combining DEA and ANN in conducting efficiency analysis (Azadeh et al., 2011; Kwon, 2014; Mostafa, 2009a; Ülengin et al., 2011; Wu et al., 2006). Wu et al. (2006) and Mostafa (2009a), in their combined modeling for banking systems, used probabilistic neural networks (PNNs) to categorize DMUs according to predicted efficiency levels. Wu (2009) successfully applied a DEA-Backpropagation neural network (BPNN) model in predicting efficiency scores of suppliers and their relevant efficiency classes. In this combined approach for efficiency prediction, BPNN was the most preferred neural network, with recent applications to technology oriented industries (Liu et al., 2013; Kwon, 2014). Kwon (2014), in his application to the smartphone industry, demonstrated the predictive power of BPNNs in estimating efficiency determined by varying returns to scale assumptions and proposed the possibility of extending the model for the optimization analysis. These previous studies added meaningful value to the existing literature as summarized in Table 2. However, a practical implementation scheme of the combined model in promoting performance improvement has not been proposed, especially within a benchmarking context.

As summarized in the table, another stream of papers introduces utilization of DEA efficiency to screen training data for neural network prediction modules and consequent improvement of prediction outcomes (Emrouznejad and Shale, 2009; Pendharkar, 2005; Pendharkar and Roger, 2003). Pendharkar and Roger (2003), in their healthcare application, used a subset of 50 ‘efficient’ DMUs selected from 100 DMUs and reported notable improvement in predicting the number of employees needed for sound operation. Emrouznejad and Shale (2009) further demonstrated effectiveness of the prescreening approach by employing a large scale data set of 10,000 DMUs. This particular application revealed the enhanced prediction performance of neural networks based upon the use of monotone increasing data sets while exploiting better performance patterns. However, prescreened subsets thus obtained can be considered ‘approximately monotonic’ due to the inclusion of both DEA_efficient (score of 1) and ‘efficient (score less than 1)’ DMUs. This implies the potential for further improvement of the combined model through the exploration of the efficient frontier. In addition, screening of a subsample for training can limit the practical utility of the proposed model in empirical analysis when using small data sets. Indeed, Table 2 shows frequent utilization of small data such as 23 suppliers (Wu, 2009) and 43 banks (Mostafa, 2009a) despite these applications being limited to the efficiency analysis. Aside from the aforementioned DEA-ANN modeling, the table presents another prospective use of ANNs for data preprocessing, thus forming a ANN-DEA sequence (Celebi and Bayraktar, 2008; Kheirkhah et al., 2013; Liao and Li, 2008). In summary, in spite of encouraging research outcomes and promising potential for combined modeling approaches, there is still a lack of theoretical advancement and empirical support, especially in optimization and benchmarking studies (Azadeh et al., 2010; Kwon, 2014). Furthermore, research outcomes on railroads have not been reported to date.

3. Research methodology

3.1. DEA

DEA's capability to deal with multiple variables without a priori assumptions about their distributions has led to DEA being widely accepted as a methodology for assessing the relative efficiency of firms. DEA's rate of use is growing and gaining momentum in various industry sectors (Barros, 2013; Banaszewska et al., 2012; Liu et al., 2013a, 2013b). DEA, as a linear programming based nonparametric method, has proven to be effective in railroad performance analysis as evidenced through successful research outcomes (Abate et al., 2013; Couto and Graham, 2009; Merkert et al., 2010).

DEA (Charnes et al., 1978) is a frontier technology with its theoretical basis in Farrell's work (1957). By analyzing DMUs with multiple inputs and outputs, the DEA model envelopes extreme data points to represent efficient frontiers. DEA then measures the distance of each DMU from the frontier, which represents its level of inefficiency. Therefore, the DMUs that lie on the efficient frontier are considered efficient and hold efficiency scores of 1 while the remaining DMUs below it are given nonnegative fractional scores. Assuming n-

Table 2
Literature on combined DEA and ANN models.

Category	ANN usage	Authors	Model sequence	Data (DMUs)
Comparative studies	Efficiency prediction	Athanassopoulos and Curram (1996)	DEA vs. BPNN	16 Simulated data, 250 Bank branches
		Liu et al. (2013)	DEA vs. BPNN	29 Semi-conductor firms (6 years)
		Liao et al. (2007)	DEA vs. BPNN	7 Countries (35 years)
		Santin (2008)	DEA vs. BPNN	100 Simulated data
		Santin et al. (2004)	DEA vs. BPNN	50–300 Simulated data
		Wang (2003)	DEA vs. BPNN	49 Education levels in a PFT program
Complementary studies	Efficiency prediction	Kwon (2014)	DEA-BPNN	8 Companies (8 years)
		Azadeh et al. (2010)	DEA-BPNN	19 Power plants (8 years)
		Azadeh et al. (2011)	DEA-BPNN	102 Bank branches
		Sreekumar and Mahapatra (2011)	DEA-BPNN	49 Business schools
		Ülengin et al. (2011)	DEA-BPNN	45 Nations
		Vaninsky (2004)	DEA-BPNN	50 Companies (four day stock)
		Wu (2009)	DEA-BPNN	23 Suppliers
	Efficiency classification	Mostafa (2009a)	DEA-PNN[a]	43 Banks
		Mostafa (2009b)	DEA-PNN	85 Banks
		Mostafa (2009c)	DEA-PNN	62 Companies
		Wu (2009)	DEA-BPNN	23 Suppliers
		Wu et al. (2006)	DEA-PNN	142 Bank branches
	Screening of training data	Emrouznejad and Shale (2009)	DEA-BPNN	10,000 DMUs
		Pendharkar and Roger (2003)	DEA-BPNN	100 Hospitals
		Pendharkar (2005)	DEA-BPNN	275 Hospitals
		Pendharkar (2005)	DEA-BPNN	175 Hospitals
	Data processing	Celebi and Bayraktar (2008)	BPNN-DEA	20 Suppliers
		Kheirkhah et al. (2013)	BPNN-DEA	130 Monthly electricity data
		Kuo et al. (2010)	BPNN-DEA	12 Suppliers
		Liao and Li (2008)	SOFM[b]-DEA	43 Simulated data
		Pendharkar (2011)	RBFN[c]-DEA	30 Simulated data
		Samoilenko and Osei-Bryson (2013)	BPNN-DEA	18 Nations (10 years)

(continued on next page)

Table 2 *(continued)*

Note:
[a] PNN- Probabilistic neural network.
[b] SOFM: Self organizing feature map.
[c] RBFN: Radial basis function neural network.

DMUs with r-input and s-output vectors, the CCR efficiency of kth DMU can be found by using the following formula:

$$\text{Maximize } h_k = \frac{\sum_{j=1}^{s} o_j y_{jk}}{\sum_{i=1}^{r} q_i x_{ik}} \tag{1}$$

The problem can be solved by transforming Eq. (1) into a linear programming format as in Eqs. (2)–(4).

$$\text{Maximize } h_k = \sum_{j=1}^{s} o_j y_{jk} \tag{2}$$

s.t.

$$\sum_{i=1}^{r} q_i x_{ik} = 1 \tag{3}$$

$$\sum_{j=1}^{s} o_j y_{jp} - \sum_{i=1}^{r} q_i x_{ip} \le 0 \quad \text{p} = 1, \ldots, \text{n} \tag{4}$$

$o_j{\cdot}q_i \ge \rho \quad \forall_{j,i} \quad \rho$: a positive infinitesimal value

where

y_{jp}: quantity of *j*th output of DMU_p.
x_{ip}: quantity of *i*th input of DMU_p.
o_j: weight assigned to *j*th output.
q_i: weight assigned to *i*th input.

Note that the maximization of h_k indicates the output orientation of efficiency measurement. In this research, the output oriented model is used by focusing on output maximization rather than resource minimization. The choice of orientation depends on the characteristics of the problem to be solved, the controllability of variables, and the discretionary power of variables (Abate et al., 2013). However, the orientation itself is not considered an important factor when railroads are compared to other industries (Coelli and Perelium, 2000). In general, output orientation is considered an appropriate choice for planning and strategy formulation purposes while input orientation is better suited for the application of operational level improvement (Fernandes and Pacheco, 2002; Wang et al., 2003). More importantly, this research aims at predicting desired output levels attainable for given levels of infrastructure inputs when setting improvement goals. Under these premises, the output-oriented version of the CCR (CCR-O) model is used for this study. The CCR DEA model running on constant returns to scale (CRS) assumptions has assigned theoretical foundations to the variations of DEA models and is still one of the most popular models used for efficiency analysis (Chandra, et al., 1998; Couto and Graham, 2009; Himola, 2007; Tsai et al., 2006). Therefore, in this exploratory study of performance prediction, the CCR model is considered a suitable choice.

3.2. ANNs

ANNs are a data driven nonparametric intelligent modeling tool which learns nonlinear functional relationships and pattern associations between pairs of input and output variables. ANNs are character-

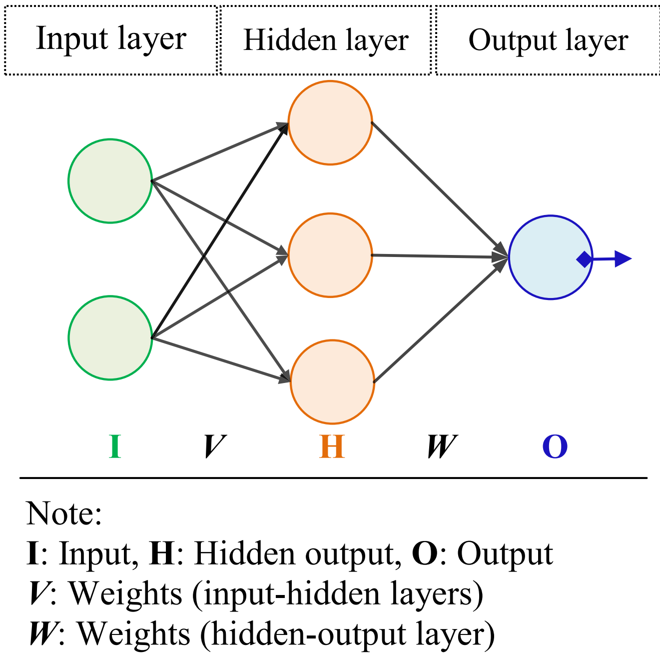

Fig. 1. A simple prediction neural network (2-3-1 structure).

ized by network structures represented by a topology of neural arrays and learning paradigms. ANN learning, which is constructed to mimic the human brain, enables generalization and adaptive pattern matching capability without the need for *a priori* assumptions or imposed constraints (Fausett, 1994). Given the different available neural network models, the BPNN model is used in this study to capitalize on its proven predictive capability.

Fig. 1 shows a simple BPNN model structured for a prediction task. As shown from the figure, a typical BPNN has a multilayered structure which includes an input layer, a hidden layer, and an output layer. Neurons in adjacent layers are linked together via highly interconnected weights for information sharing and processing. The network is trained to determine the near optimal weights for capturing the underlying characteristics of the training data. In general, one hidden layer is deemed sufficient in most nonlinear mapping problems. The number of hidden neuron layers is dependent on the complexity of the problem to be solved (Ciampi and Gordini, 2013; Ülengin et al., 2011).

The back-propagation algorithm adopts a gradient search method of network training by minimizing the error term, represented by the Euclidean distance E, between the target and actual outputs of all training pairs (Fausett, 1994). The error can be represented by:

$$E = \frac{1}{2}\sum_{k} [T_k - Y_k]^2 \tag{5}$$

where, T is the target (desired) output and Y the activated (actual) output of neuron k. Back-propagation learning starts with the calculation of network outputs through a feed forward process, as shown in the following formula. In this process, pairs of input and output patterns are presented to the network.

$$Y_k = f(y_{netK}) = f(\sum_{j} H_j w_{jk}) \tag{6}$$

f () is a nonlinear activation function applied to net outputs of neuron k, y_{netK}. H_j represents the inputs from the hidden neuron j to output neuron k and w_{jK} is the weight between neurons j and k. Followed by feed forward input presentation and error calculations, the back-propagation of errors occur from the output layer to the hidden layer, then from the hidden layer to the input layer, to ensure that weight adjustments between adjacent neurons are done in such a way that they minimize E. Let $\Delta v_{i,j}(t)$and $\Delta w_{j,k}(t)$represent weight changes between input (i) and hidden (j) neurons and between hidden (j) and output (k) neurons as expressed below.

$$\Delta v_{i,j}(t) = -\frac{\partial E}{\partial v_{i,j}} \tag{7}$$

$$\Delta w_{j,k}(t) = -\frac{\partial E}{\partial w_{j,k}} \tag{8}$$

Then, the new weights between associated neurons at epoch t using learning rate ρ can be obtained by using the following rules.

$$v_{i,j}(t+1) = v_{i,j}(t) + \rho\Delta v_{i,j}(t) \tag{9}$$

$$w_{j,k}(t+1) = w_{j,k}(t) + \rho\Delta w_{j,k}(t) \tag{10}$$

To summarize, the back propagation algorithm adjusts preset weights to minimize errors using training data and this process continues until convergence conditions are met, in other words, until the network is generalized. More detailed description on learning rules and variations are available from the extensive literature (Fausett, 1994; Kheirkhah et al., 2013).

4. Empirical analysis and results

4.1. Data and variables

The U.S. Class I railroads were selected for this empirical analysis, which include seven companies: BNSF Railway Company (BNSF), CSX Transportation Inc (CSX), Grand Trunk Corporation (GTC), Kansas City Southern Railway Company (KCS), Norfolk Southern Combined Subsidiaries (NS), Soo Line Railroad Company (SOO), and the Union Pacific Railroad (UP). The data set used for this study was collected from the annual R-1 reports of Class I railroads, which are available from the Surface Transportation Board (STB). The time series data captured railroad specific operations for an eight year period, 2005–2012 inclusive. Each railroad's yearly observation is treated as an independent DMU. Therefore, the collected data sets constitute 56 DMUs. In this combined DEA-ANN modeling approach, four input variables were selected: locomotive units, freight cars, track mileage operated, and number of employees. These input variables were then paired with two single output variables RTM and FTON to form RTM and FTON models, respectively. The combination of these variables has proven to be meaningful in previous analyses of European railroads (Himola, 2007). Table 3 shows the summary statistics of each variable and its usage in the DEA and ANN modules.

4.2. The proposed framework

The purpose of this empirical study is to explore the predictive potential of ANN when used in conjunction with DEA. Performance prediction, especially in a competitive and capital intensive environment such as the railroad industry, is a critical benchmarking step toward achieving superior performance over competitors. In addition, the capability to predict actual output requirements beyond existing relative efficiency levels, under various assumptions and changing scenarios, is a practical necessity for the advancement of performance benchmarking. As an exploratory study, the focus of this research is centered on 'best practice' modeling in the sense that the designed model predicts the desired level of efficient output needed for a given DMU to become a superior performer. This pilot study can be the foundation for further extension in support of 'better practice' benchmarking, given that the model can predict the appropriate performance output level needed to support an incremental improvement process. As shown in Fig. 2, this empirical experiment is conducted in three stages. First, DEA efficiency is measured by using an output oriented CCR (CCR-O) model. Second, BPNN models are designed to test the

Table 3
Descriptive statistics of variables.

	Locomotive units	Freight cars	Track mileage	Employees	RTM (M)	FTON (K)
Max	8721	106,743	52,667	53,309	664,384	658,867
Min	359	10,526	4361	2428	20,361	47,850
Average	3486	56,794	28,125	22,991	244,303	336,260
SD	2937	35,820	18,973	17,741	231,783	210,224
DEA	Input	Input	Input	Input	Output[a]	Output[b]
ANN[c]	Input	Input	Input	Input	Input[a]	Input[b]
ANN[d]	Input	Input	Input	Input	Output[a]	Output[b]

Note:
[a] RTM model.
[b] FTON model.
[c] ANN for efficiency prediction (DEA efficiency is used as output).
[d] ANN for best performance prediction.

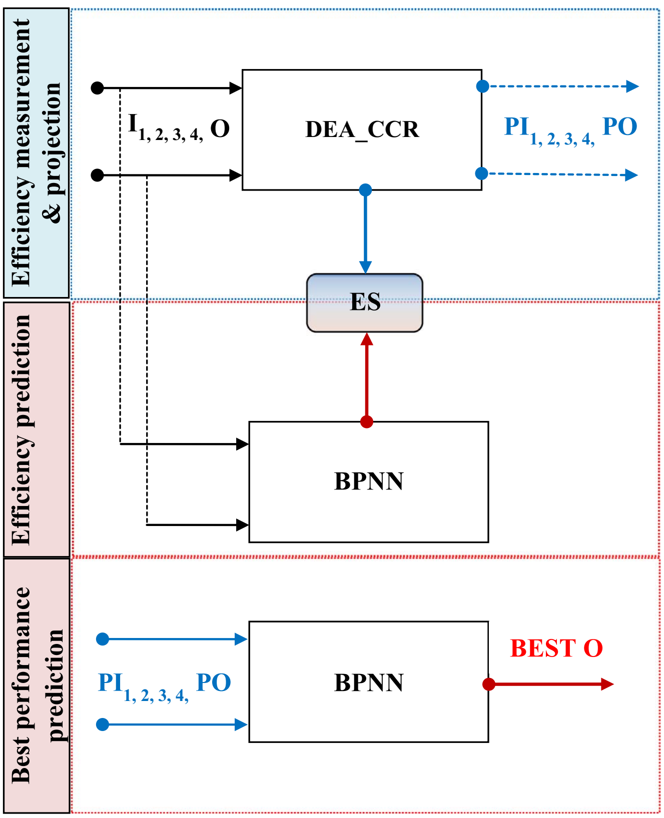

Fig. 2. A proposed DEA-ANN framework.

feasibility of predicting efficiency scores produced by the CCR model. In the last stage, the BPNN is trained utilizing DEA projected data sets. In this innovative approach, the BPNN approximates the data envelopment surface by learning the optimal performance pattern of DMUs lying on the efficient frontier. In so doing, ANN can predict superior performance for seen or unseen hypothetical DMUs.

As discussed, DEA is used in the first stage to evaluate the performance of each railroad and the overall efficiency trends of Class I railroads. The DEA module receives the input and output vectors ($I_{1,2,3,4}$, O), calculates the efficiency score (ES) for each DMU, and generates projections ($PI_{1,2,3,4}$, PO) of the performance needed for each DMU to become efficient. In the second stage of the experiment, the BPNN module is first trained to learn the efficiency patterns needed to predict the efficiency score of each DMU. The extended experiment is then conducted to explore the potential of the BPNN model to predict best performance outputs by utilizing DEA projected optimal data sets. The BPNN learns the input-output patterns of efficient DMUs and the predicted output indicates the desired performance level needed for inefficient DMUs to become superior performers.

4.3. DEA modeling and analysis

Table 4 shows the results of the DEA experiment utilizing infrastructure variables as input and two quantity variables, RTM and FTON, as a separate output.

As shown in the table, the CCR DEA model produces relative efficiency score, rank and reference set for each DMU. The results show varying efficiency scores among competitors, thus implying potential performance improvement that can be made in railroad as hinted by earlier studies (Lim and Lovell, 2009; Shi et al., 2011). Noteworthy are the differing efficiency levels depending on selected output variable. For example, BNSF appears to be an industry leader throughout the observation period in RTM model, but not in the FTONS model. Instead, GTX demonstrates superior performance in the FTON model and becomes a benchmark for other companies. As discussed earlier, RTM and FTON complement each other as a meaningful quantity measure for the integral analysis of railroads (Himola, 2007; O'Mahony and Oulton, 2000).

Figs. 3 and 4 visualize the efficiency trends of each railroad for the RTM and FTON models, respectively.

As discussed, BNSF in the RTM model consistently demonstrates superior efficiency along with GTX in 2012. These two railroads are followed by UP in keeping their distance from less efficient rails including CSX, KCS, SOO, and NS. Interestingly, each railroad reveals a rapid downturn in 2009 due to the impact of the recession, and shows recent recovery effects in both models. In contrast, in the FTON model, BNSF shows consistently inferior performance to GTX and KCS in terms of efficiency. GTX, which shows steady improvement in the RTM model, has been a leader in the FTON model throughout the observation period except for 2009, making it a benchmark for others. On the contrary, NS exhibits consistently lower performance than peer railroads in the RTM model and marks the lowest in recent operations in both models, therefore, managerial priority should be given to improvement initiatives. This initial stage of DEA analysis clearly exposes the competitive edge of each railroad and hints at the potential improvements that can be made. Note that detailed analysis of individual rail performance is beyond the scope of this research. Rather, DEA efficiency and projected data obtained thus far are further processed for the prediction task conducted by neural network modules.

4.4. ANN modeling for efficiency prediction

The next stage of the experiment is aimed at exploring the nonlinear mapping capabilities of ANNs by using both DEA inputs and outputs as neural network inputs and by assigning CCR efficiency scores as ANN outputs. For the implementation of neural network

Table 4
DEA experiment results.

DMU	RTM model					FTON model			
	Score	Rank	Reference			Score	Rank	Reference	
B12	1.0000	1	B12			0.6712	18	G12	
B11	0.9888	7	B12	B10	B08	0.6457	26	G12	G10
B10	1.0000	1	B10			0.6583	24	G12	
B09	0.9212	10	B10	B08		0.5968	40	G12	G10
B08	1.0000	1	B08			0.6706	20	G12	G10
B07	1.0000	1	B07			0.4994	55	G12	G10
B06	0.9877	8	B07	G12		0.7124	17	G12	G10
B05	0.9925	6	B08	B07	G12	0.6662	21	G12	G10
C12	0.5198	45	B07	G12		0.5705	49	G10	
C11	0.5336	44	B08	G12		0.6038	38	G10	
C10	0.5422	43	B08	G12		0.5924	42	G10	
C09	0.4947	50	B08	G12		0.5456	54	G10	
C08	0.5742	39	B07	G12		0.6425	28	G10	
C07	0.5893	37	B07	G12		0.6456	27	G10	
C06	0.6235	27	B07	G12		0.6625	23	G10	
C05	0.6173	28	B07	G12		0.6505	25	G10	
K12	0.6020	33	B10			0.8728	14	G10	
K11	0.6162	30	B10			0.9411	10	G10	
K10	0.6275	26	B10			0.9215	12	G10	
K09	0.5984	34	B10			0.8906	13	G10	
K08	0.5653	42	B10			0.9540	9	G10	
K07	0.5831	38	B10			0.9362	11	G10	
K06	0.6164	29	B10			0.9551	8	G10	
K05	0.5167	47	B10			0.8668	15	G10	
N12	0.4420	55	B07	G12		0.5632	51	G10	
N11	0.4600	53	B07	G12		0.5879	44	G10	
N10	0.4466	54	B07	G12		0.5779	46	G10	
N09	0.3857	56	B08	B07	G12	0.4982	56	G10	
N08	0.4746	52	B07	G12		0.6135	35	G10	
N07	0.4787	51	B07	G12		0.6164	33	G10	
N06	0.5007	49	B07	G12		0.6274	32	G10	
N05	0.5122	48	B07	G12		0.6152	34	G10	
S12	0.6037	32	B08	G12		0.5800	45	G12	G08
S11	0.5902	36	B08	G12		0.5737	48	G08	G07
S10	0.5656	41	B08	G12		0.6014	39	G07	G06
S09	0.5193	46	B08	G12		0.5698	50	G05	
S08	0.5691	40	B10	B08		0.5621	52	G10	G05
S07	0.6678	25	B08	G12		0.5751	47	G10	G05
S06	0.5978	35	B08	G12		0.5946	41	G10	G05
S05	0.6053	31	B08	G12		0.5883	43	G05	
U12	0.8236	17	B12			0.6630	22	G12	
U11	0.8400	14	B12			0.6709	19	G12	
U10	0.7856	19	B12	B08		0.6096	36	G12	
U09	0.7092	22	B12	B08		0.5499	53	G12	G10
U08	0.8329	15	B08			0.6372	29	G12	G10
U07	0.8295	16	B08			0.6334	30	G10	
U06	0.8233	18	B08			0.6325	31	G10	
U05	0.7830	20	B08			0.6056	37	G10	
G12	1.0000	1	G12			1.0000	1	G12	
G11	0.9266	9	B08	G12		1.0000	1	G11	
G10	0.8622	13	B07	G12		1.0000	1	G10	
G09	0.7063	23	B08	G12		0.8559	16	G12	G08
G08	0.9146	11	B08	G12		1.0000	1	G08	
G07	0.8683	12	B08	G12		1.0000	1	G07	
G06	0.7741	21	B08	G12		1.0000	1	G06	
G05	0.6861	24	B08	G12		1.0000	1	G05	

models, NeuralWorks Predict 3.1 simulator software was used by exploiting its built-in design capabilities. The neural simulator partitions data into training and test sets according to a 7:3 ratio. Accordingly, 39 randomly selected DMUs were used for training and the remaining 17 DMUs were used for testing throughout this experiment. For each of the RTM and FTON models, two BPNN models were built by using two different sets of training data selected by round robin and random sequences. The BPNN model formed 5-7-1 and 5-8-1 network structures for the RTM model and 5-8-1 and 5-7-1 structures for the FTON model by generating the hidden neurons necessary for nonlinear pattern mapping. Neural network performance is dependent on the model building parameters such as training sequence, number of hidden neurons and learning mechanisms, however, network optimization still remains a challenging task (Emrouznejad and Shale, 2009; Hu et al., 1999; Kheirkhah et al., 2013; Li et al., 2012).

In this particular application, the trained neural network demonstrated a generalized pattern learning capability by showing high correlations and low prediction errors between pairs of actual CCR efficiencies and predicted neural outputs for both training and test data sets. Table 5 summarizes performance of BPNN for both the RTM and FTON models. The output shows a good model fit with slightly higher values on training data with the exception of the RTM output on a random training sample. Overall, the results demonstrate the robustness of neural network models with regard to sample variations by showing high performance on all training subsets without sacrificing network performance on test data.

For further analysis, better performance models in terms of high correlation and low errors were selected, thus yielding RTM and FTON models trained by round robin and random sequence, respectively.

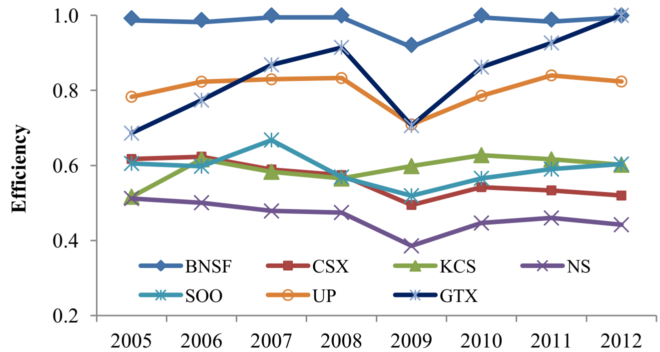

Fig. 3. Efficiency trends (RTM model).

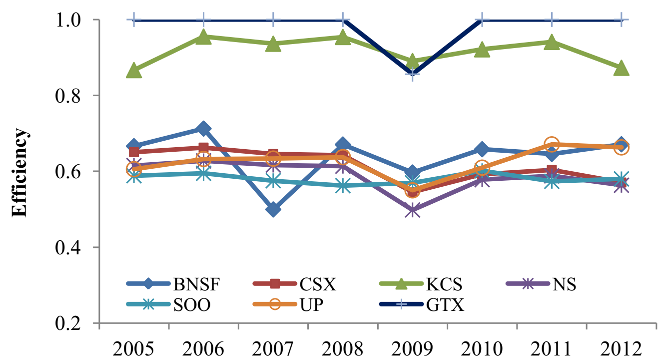

Fig. 4. Efficiency trends (FTON model).

Table 5
BPNN results on efficiency prediction.

Model (neurons)	Neurons	Data	R	AAE[a]	MAE[b]	MAPE[c]	Data selection
RTM	5-7-1	Train	0.979	0.034	0.130	5.601%	Round
		Test	0.987	0.029	0.082	4.634%	robin
RTM	5-11-1	Train	0.971	0.032	0.103	4.885%	Random
		Test	0.964	0.040	0.117	6.077%	
FTON	5-8-1	Train	0.945	0.043	0.137	6.816%	Round
		Test	0.942	0.057	0.143	9.215%	robin
FTON	5-7-1	Train	0.991	0.017	0.054	2.558%	Random
		Test	0.958	0.034	0.080	5.408%	

Note:
[a] Average absolute error.
[b] Max. absolute error.
[c] MAPE: Mean absolute percentage error.

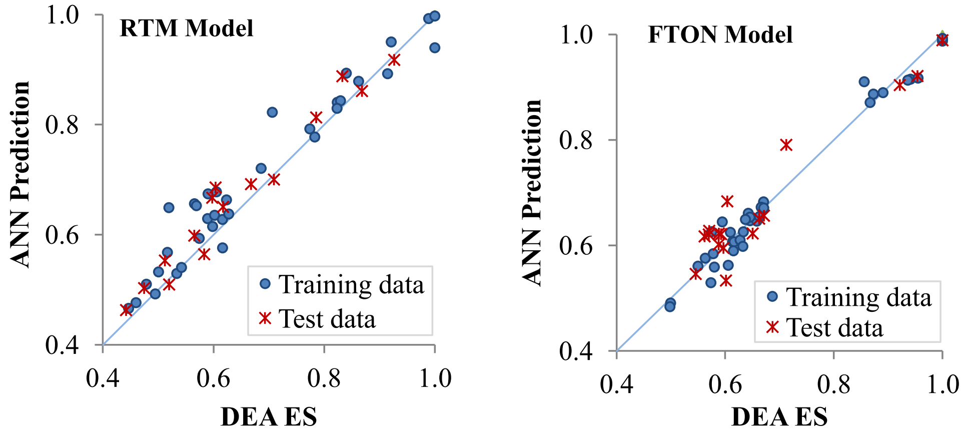

Fig. 5. BPNN prediction of DEA efficiency.

Fig. 5 displays a two dimensional plot of the actual and predicted efficiency scores (ES) of the DMUs used for training and testing in both models. The high correlation between the actual and the predicted values demonstrates the generalized learning property of BPNN and exhibits a regression type processing of BPNN to learn central tendency of presented data as well.

The capability of the combined model to predict accurate efficiency scores of DMUs is encouraging in that it can be the basis for further extension of the combined model toward 'better practice' benchmarking and incremental improvement. Up to this date, the main focus of the combined model has been limited to predicting DEA efficiency scores with rare exploration into the models potential for being used in promoting efficiency improvement. Furthermore, despite DEA solutions being mathematically sound, DEA targets are often hard to realize, thus limiting practical utility of the method (Mostafa, 2009a, 2009b, 2009c; Wang et al., 2013). In practice, pursuit of stepwise improvement might be a feasible option when considering environmental constraints of firms.

To further elaborate this practically demanding but less exploited managerial issue, additional experimentation is conducted by taking NS12 as a test case, which marked the lowest efficiency among other peer DMUs in 2012 as shown in Table 4. DEA efficiency of NS12 was determined to be 0.4420 and 0.5632 for the RTM and FTON model, respectively. For this inefficient DMU, DEA suggested a 0% (30%, 0%, 7%) reduction of locomotives (freight cars, miles, employees) and a 126% increase in RTM and a 58% (23%, 0%, 23%) reduction of resources and 78% increase in FTON to boost its efficiency level to 1. These figures might be impractical and impossible to be sought in the short term. Consequently, managers need to deal with this dilemma in pursuit of better performance. Tables 6 and 7 present prediction outcomes on possible improvement scenarios and show increasing efficiency patterns according to input reductions and output increases. As demonstrated, the proposed model has the potential to support managers in setting stepwise improvement goals through what if scenarios. The ANN's predictive power and the promising results obtained thus far prompted this study to take this pioneering exploration of ANN's capability to predict best performance outputs beyond efficiency scores.

4.5. ANN implementation for best output prediction

The last stage of this experiment is designed to predict direct output quantity in terms of RTM and FTON rather than the indirect performance measure of an efficiency index. Differing from the efficiency prediction experiment of the previous stage, the BPNN prediction module utilizes DEA projected inputs/outputs and learns the efficient frontiers of the RTM and FTON models. In other words, the neural network is trained to approximate the data envelopment surface formed by a virtual set of efficient DMUs that are transformed from real DMUs. The trained neural network module then predicts the best performance output for each DMU, accordingly. The motivation for this innovative approach is to explore the efficient frontiers determined by DEA, which preserve monotonicity properties. Monotonicity, for a simple univairate input x_1 and x_2, satisfies the following functional relationships:

$$[x_1 > x_2] \leftrightarrow [f(x_1) > f(x_2)]$$

DMUs lying on the efficient frontier are assumed to hold monotonicity, and the BPNN model trained by these data sets also preserves monotonicity assumptions and minimizes the unpredictability of the model (Archer and Wang, 1993; Pendharkar, 2005; Pendharkar and Roger, 2003). Advanced from past studies of using efficient subsets, this paper presents utilization of DEA projected optimal data which constitute virtual DMUs lying on the efficient frontier. To exhibit the effectiveness of the proposed approach, the BPNN outputs using optimal data sets were compared to the BPNN outputs obtained by using the original nonmonotonic data sets. As shown in Fig. 6, monotonic data sets result in notably small error scales implying stable learning in both RTM and FTON models. In contrast, large scale errors from the original data sets reflect unpredictability of BPNN caused by nonmonotonic data sets. This observation confirms previous research findings by Pendharkar (2005) and affirms prospective advantages of using DEA-BPNN in predicting best outputs.

Note that the BPNN output obtained by using optimal data sets represents the performance target needed for each DMU to become a superior performer and the discrepancy between the predicted and the real outputs of each DMU indicate performance gaps to be filled. This adaptive prediction capability can support managers in predicting optimal performance goals in terms of direct measure of outputs, RTM and FTON. In addition, the trained model can also predict optimal output for hypothetical inputs, thus providing managers alternative options in their paths toward best practices. As previously shown in Fig. 2, five dimensional vectors projected by the DEA preprocessor were used for this experiment and two neural networks were implemented using different training and test subsets. The training results in Table 8 demonstrate high correlation between actual and predicted outputs with low scales of prediction errors across all models and data sets regardless of sample variations. As discussed, monotonic data sets contributed to the construction of the robust models and consequent comparable performance on both training and test data showing no signs of overtraining. The BPNN models thus trained preserve adaptive and generalized learning properties.

Figs. 7 and 8 graphically display the prediction capability of the BPNN models by contrasting the optimal performance projections of the DEA module and subsequent prediction outcomes. As shown from the figures, the BPNN outputs follow the patterns of the efficient DMUs

Table 6
Incremental prediction results on RTM model.

DMU	Locomotive	F. cars	Miles	Employees	RTM (M)	Efficiency
N12	4029	86,037	36,160	30,459	185,639	0.442
Projection	4029	60,666	36,160	28,198	419,982	
Slacks	**(0%)**	**(−30%)**	**(0%)**	**(−7%)**	**(+126%)**	
Prediction scenarios	4029	77,433	36,160	28,198	204,202	0.482
	(0%)	(−10%)	(0%)	(−7%)	(+10%)	
	4029	8830	36,160	8198	222,766	0.527
	(0%)	(−20%)	(0%)	(−7%)	(+20%)	
	4029	60,666	36,160	28,198	241,330	0.597
	(0%)	(−30%)	(0%)	(−7%)	(+30%)	
	4029	60,666	36,160	28,198	259,894	0.690
	(0%)	(−30%)	(0%)	(−7%)	(+40%)	
	4029	60,666	36,160	28,198	278,458	0.788
	(0%)	(−30%)	(0%)	(−7%)	(+50%)	
	4029	60,666	36,160	28,198	297,022	0.875
	(0%)	(−30%)	(0%)	(−7%)	(+60%)	
	4029	60,666	36,160	28,198	315,585	0.942
	(0%)	(−30%)	(0%)	(−7%)	(+70%)	
	4029	60,666	36,160	28,198	334,149	0.988
	(0%)	(−30%)	(0%)	(−7%)	(+80%)	
	4029	60,666	36,160	28,198	352,713	1.019
	(0%)	(−30%)	(0%)	(−7%)	(+90%)	
	4029	60,666	36,160	28,198	371,277	1.039
	(0%)	(−30%)	(0%)	(−7%)	(+100%)	
	4029	60,666	36,160	28,198	389,841	1.052
	(0%)	(−30%)	(0%)	(−7%)	(+110%)	
	4029	60,666	36,160	28,198	408,405	1.060
	(0%)	(−30%)	(0%)	(−7%)	(+120%)	

in both models. The innovativeness of this modeling approach is in the exploration of the neural networks to approximate data envelopment surfaces and construction of the monotonicity preserving neural network models in support of best performance prediction. In practical applications, different from the aforementioned efficiency prediction model, this approach can support mangers to predict optimal output (e.g. RTM or FTON) for varying input levels. In its core, RTM (or FTON) is a monotonic function of 4 input variables that can be expressed as RTM (or FTON)=*f* (locomotives, freight cars, miles, employees). In essence, the adaptive prediction capability provides flexible managerial options for managers in setting target goals and promoting stepwise improvement in pursuit of best practice.

The experimental results obtained so far highlight the promising performance of the ANN-based benchmarking and performance modeling approach. This pilot study has established a foundation for further deployment of ANN models in the railroad industry and beyond. The combined ANN and DEA model can enhance the current benchmarking process by adding prediction capability to the performance measurement framework. The generalization and intelligent learning capacity of ANNs can be a valuable addition, which enables adaptive prediction over previously unseen data sets. In this capacity, managers can generate possible scenarios and test the feasibility of various alternatives and options during the planning and implementation of their improvement initiatives. In addition, timely prediction even under conditions of insufficient information will be a potential benefit derived from using ANNs. After training, the ANN model preserves abstract information in its weight sets and decodes incoming data into a relevant prediction output.

As observed from the experiment results, most of the Class I railroads are revealed to have inefficient operations. Accordingly, managers are engaged in benchmarking processes which typically include planning, analysis, and implementation stages. The proposed combined model can support managers in identifying the best performers to learn from, predict levels of "best performance" needed to gain superiority, determine the performance gap to be overcome to initiate improvement, and monitor progress and take appropriate actions. The

Table 7
Incremental prediction results on FTON model.

DMU	Locomotive	F. cars	Miles	Employees	FTON (K)	Efficiency
N12	4029	86,037	36,160	30,459	387,858	0.555
Projection	1688	65,893	36,160	22,017	688,664	1.000
Slacks	**(−58%)**	**(−23%)**	**(0%)**	**(−28%)**	**(78%)**	
Prediction scenarios	3626	77,433	36,160	27,413	426,644	0.674
	(−10%)	(−10%)	(0%)	(−10%)	(+10%)	
	3223	68,830	36,160	24,367	465,430	0.817
	(−20%)	(−20%)	(0%)	(−20%)	(+20)	
	2820	65,893	36,160	22,017	504,215	0.929
	(−30%)	(−23%)	(0%)	(−28%)	(+30%)	
	2417	65,893	36,160	22,017	543,001	0.978
	(−40%)	(−23%)	(0%)	(−28%)	(+40%)	
	2015	65,893	36,160	22,017	581,787	0.991
	(−50%)	(−23%)	(0%)	(−28%)	(+50%)	
	1612	65,893	36,160	22,017	620,573	0.995
	(−58%)	(−23%)	(0%)	(−28%)	(+60%)	
	1612	65,893	36,160	22,017	659,359	0.996
	(−58%)	(−23%)	(0%)	(−28%)	(+70%)	

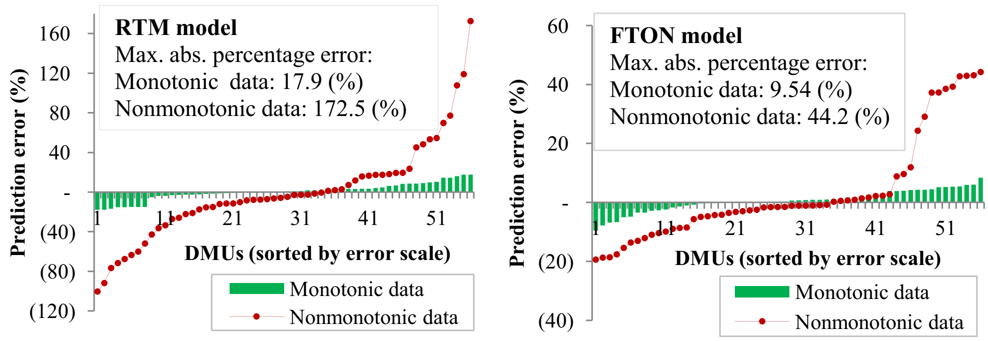

Fig. 6. Prediction performance over monotonic vs. nonmontonic data.

Table 8
BPNN results on best output prediction.

Model (neurons)	Neurons	Data	R	AAE[a]	MAE[b]	MAPE[c]	Data selection
RTM	4-3-1	Train	0.9993	6661	37,482	4.853%	Round
		Test	0.9994	6580	22,810	4.589%	robin
RTM	4-2-1	Train	0.9990	9190	36,778	6.804%	Random
		Test	0.9989	9368	17,562	5.010%	
FTON	4-4-1	Train	0.9996	7827	24,550	2.824%	Round
		Test	0.9997	6531	23,824	2.555%	robin
FTON	4-5-1	Train	0.9995	8355	33,010	4.386%	Random
		Test	0.9994	10,223	29,887	4.881%	

Note:
[a] Average absolute error.
[b] Max. absolute error.
[c] MAPE: Mean absolute percentage error.

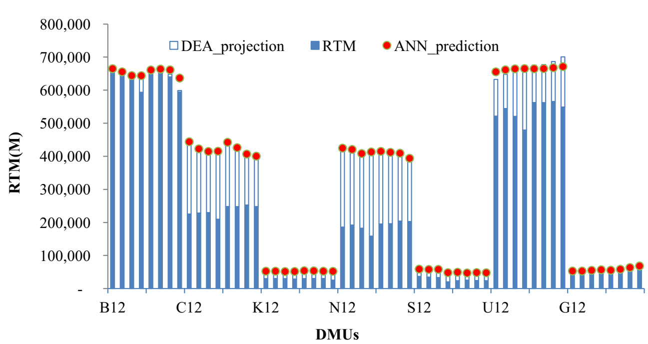

Fig. 7. BPNN prediction on target RTM.

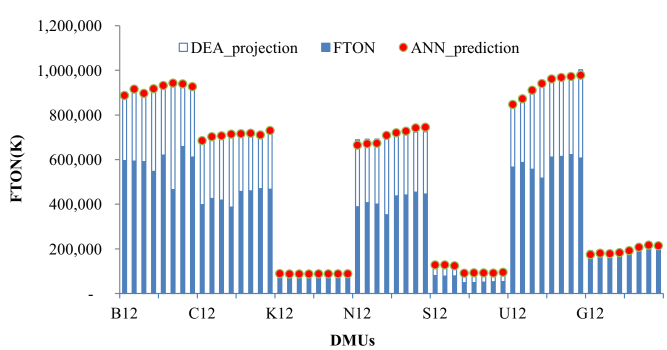

Fig. 8. BPNN prediction on FTON.

approach presented in this study can eventually provide managers with the quantitative capability needed to initiate performance benchmarking and improvement initiatives.

5. Concluding remarks

The primary purpose of this paper is to explore the potential capabilities of ANN in railroad performance measurement and prediction, and for this task, a combined modeling approach was presented with encouraging research results. DEA, despite its powerful optimization capability and breadth of applications, has limitations when faced with handling unseen or new data sets. This restricts its use for estimation and prediction tasks. ANN, on the other hand, has its strength in its predictive power stemming from its adaptive learning and generalization capabilities, even when utilizing a small subset of data without *a priori* knowledge regarding the underlying distribution. Taking advantage of the complementary nature of these two methods, this paper introduces an uncommon approach of combining DEA-ANN methods. In spite of a growing necessity, prediction capability has rarely been built into the measurement framework. However, with the successful implementation of integrated measurement and prediction models, this study advances the current benchmarking paradigm. In addition, the encouraging empirical results highlight promising advantages of utilizing the combined model and potential extension of ANNs in benchmarking studies.

The contribution of this paper can be summarized as follows. First, the presented model successfully evaluated the current comparative performance of Class I railroads and efficiency trends over an eight year period utilizing industry specific infrastructure and quantity output variables. Second, a combined DEA-ANN model was successfully designed to predict efficiency scores for each DMU. Adaptive prediction capability can assist managers in assessing hypothetical efficiencies utilizing 'what-if' scenarios. Third, an innovative attempt has been made to predict best performance output beyond the indirect measure of efficiency scores. In this approach, the BPNN can predict the desired level of best performance of a DMU through approximation of the efficient frontier. This research advances the current paradigm of best practice benchmarking with empirical support from the rail industry, thus enabling incremental improvement through integration of measurement and prediction modules. In practice, capability to predict achievable performance levels can be a significant decision aid to prudent managers who are seeking stepwise improvement, especially under contingent operational environments and resource constraints. Therefore, it will be valuable to further investigate the feasibility of the proposed model to predict '*better performance*' and '*above average performance*' in addition to '*best performance*.' In this predictive analysis scheme, target performance can be leveraged by desired efficiency as a reference to predict relevant performance output. Moreover, exploration of varying performance patterns differentiated by efficiency categories can provide meaningful insight on the strategic impact of key variables on different levels of firm performance, thus leading to an explanatory prediction model. In this extended experiment, it will be worthwhile to collect more data from additional years and increase the sample size by filling performance gaps. The utilization of cost variables will also produce interesting research outcomes. Through its distinctive methodological advancement, the proposed model can provide a sound foundation for modeling consecutive multi-stage production systems (e.g., bank operation, technology innovation, and etc.) as a promising prediction model. As discussed in Seiford and Zhu (1999), bank operation reflects a typical two-stage production process which includes three layers of input (e.g., employees and assets), intermediate output (e.g., revenues and profits), and final output (market value and earnings per share). Technology innovation processes can be viewed as two consecutive subprocesses, knowledge creation and subsequent commercialization, which utilize input (e.g., assets and R & D resources), intermediate output (e.g., patents and citations), and final output (e.g., profit and market value) in a layered structure (Kaihua and Mingting, 2014; Wang et al., 2013). Recent literature shows emerging interest in two-stage DEA models for efficiency analysis of multi-stage production systems, however, the scope is still limited to the efficiency measurement (Barros and Wanke, 2014; Kao and Huang, 2008). Therefore, as a fruitful future research agenda, the proposed combined model can be further explored to augment two-stage measurement framework toward predictive analysis for the prediction of intermediate output and subsequent final output as a direct performance measure beyond efficiency for each DMU.

References

Abate, M., Lijesen, M., Pels, E., Roelevelt, A., 2013. The impact of reliability on the productivity of railroad companies. Transp. Res.: Part E, 5141–5149.
Andrikopoulos, A.A., Loizides, J., 1998. Cost structure and productivity growth in European railway systems. Appl. Econ. 30 (12), 1625–1639.
Archer, N., Wang, S., 1993. Application of the back propagation neural network algorithm with monotonicity constraints for two-group classification problems. Decis. Sci. 24 (1), 60–75.
Association of American Railroads, 2012. America's Freight Railroads: Global Leaders. Available at ⟨https://www.aar.org/keyissues/Documents/Background-Papers/Americas-RRs-Global-Leaders-October-2012.pdf⟩ (20. 05.13).
Association of American Railroads, 2013a. Railroad infrastructure investment. Available at ⟨https://www.aar.org/keyissues/Pages/Infrastructure-Investment.aspx#.UYveUbWTjwB⟩ (20. 05.13).
Association of American Railroads, 2013b. We invest so America moves: how railroads spend their money. Available at ⟨https://www.aar.org/keyissues/Pages/Infrastructure-Investment.aspx#.UYveUbWTjwB⟩ (20. 05.13).
Association of American Railroads, 2013c. Overview of America's Freight Railroads. Available at ⟨https://www.aar.org/keyissues/Documents/Background-Papers/Overview-US-Freight-RRs.pdf⟩ (20. 05.13).
Athanassopoulos, A.D., Curram, S.P., 1996. A comparison of data envelopment analysis and artificial neural networks as tools for assessing. J. Oper. Res. Soc. 47 (8), 1000–1016.
Azadeh, A., Saberi, M., Anvari, M., 2010. An integrated artificial neural network algorithm for performance assessment and optimization of decision making units. Expert Syst. Appl. 36 (8), 5688–5697.
Azadeh, A., Saberi, M., Moghaddam, R.T., Javanmardi, L., 2011. An integrated data envelopment analysis-artificial neural network-rough set algorithm for assessment of personnel efficiency. Expert Syst. Appl. 38 (3), 1364–1373.
Banaszewska, A., Cruijssen, F., Dullaert, W., Gerdessen, J.C., 2012. A framework for measuring efficiency levels – the case of express depots. Int. J. Prod. Econ. 139 (2), 484–495.
Barros, C.P., Wanke, P., 2014. Insurance companies in Mozambique: a two-stage DEA and neural networks on efficiency and capacity slacks. Appl. Econ. 46 (29), 3591–3600.
Barros, C.P., Bin Liang, Q., Peypoch, N., 2013. The efficiency of French regional airports: an inverse -convex analysis. Int. J. Prod. Econ. 141 (2), 668–674.
Bitzan, J.D., Keeler, T.E., 2003. Productivity growth and some of its determinants in the deregulated U.S. railroad industry. South. Econ. J. 70 (2), 232–253.
Carlucci, D., Renna, P., Schiuma, G., 2013. Evaluating service quality dimensions as antecedents to outpatient satisfaction using back propagation neural network. Health Care Manag. Sci. 16 (1), 37–44.
Cavalieria, S., Maccarrone, P., Pinto, R., 2004. Parametric vs. neural network models for the estimation of production costs: a case study in the automotive industry. Int. J. Prod. Econ. 91 (2), 165–177.
Çelebi, D., Bayraktar, D., 2008. An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information. Expert Syst. Appl. 35 (4), 1698–1710.
Chandra, P., Cooper, W., Shanling, L., Rahman, A., 1998. Using DEA to evaluate 29 Canadian textile companies—considering returns to scale. Int. J. Prod. Econ. 54 (2), 129–141.
Chapin, A., Schmidt, S., 1999. Do mergers improve efficiency? J. Transp. Econ. Policy 33 (2), 147–162.
Charnes, A., Cooper, W., Wand, E., 1978. Measuring the efficiency of decision making units. European. J. Oper. Res. 2 (6), 429–444.
Chou, J.S., Tai, Y., Chang, L.J., 2010. Predicting the development cost of TFT-LCD manufacturing equipment with artificial intelligence models. Int. J. Prod. Econ. 128 (1), 339–350.
Ciampi, F., Gordini, N., 2013. Small enterprise default prediction modeling through artificial neural networks: an empirical analysis of Italian small enterprises. J. Small Bus. Manag. 51 (1), 23–45.
Coelli, T., Perelman, S., 1999. A comparison of parametric and non-parametric distance functions: with application to European railways. Eur. J. Oper. Res. 117 (2), 326–339.
Coelli, T., Perelman, S., 2000. Technical efficiency of European railways: a distance function approach. Appl. Econ. 32 (15), 1967–1976.
Couto, A., Graham, D.J., 2009. The determinants of efficiency and productivity in European railways. Appl. Econ. 41 (22), 2827–2851.
Emrouznejad, A., Shale, E., 2009. A combined neural network and DEA for measuring efficiency of large scale data sets. Comput. Ind. Eng. 56 (1), 249–254.
Farrell, M.J., 1957. The measurement of productive efficiency. J. R. Stat. Soc., Ser. A (III), 253–290.
Fausett, L., 1994. Fundamentals of Neural Networks: Architectures, Algorithms, and Applications. Prentice Hall, Englewood Cliffs, NJ.
Fernandes, E., Pacheco, R.R., 2002. Efficient use of airport capacity. Transp. Res.: Part A: Policy Pract. 36 (3), 225–238.
Global Railroads Industry Profile, 2013. Railroads Industry Profile: Global. pp. 1–32.
Himola, O.P., 2007. European railway freight transportation and adaptation to demand decline. Int. J. Product. Perform. Manag. 56 (3), 205–225.
Hu, M., Zhang, G., Jiang, C., Patuwo, E., 1999. A cross validation analysis of neural network out-of-sample performance in exchange rate forecasting. Decis. Sci. 30 (1), 197–216.
Kaihua, C., Mingting, K., 2014. Staged efficiency and its determinants of regional innovation systems: a two-step analytical procedure. Ann. Reg. Sci. 52 (2), 627–657.
Kao, C., Hwang, S., 2008. Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. Eur. J. Oper. Res. 185 (3), 418–429.
Kheirkhah, A., Azadeh, A., Saberi, M., Azaron, A., Shakouri, H., 2013. Improved estimation of electricity demand function by using of artificial neural network, principal component analysis and data envelopment analysis. Comput. Ind. Eng. 64 (1), 425–441.
Kim, S.Y., 2011. Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis. Serv. Ind. J. 31 (3), 441–468.
Kourentzes, N., 2013. Intermittent demand forecasts with neural networks. Int. J. Prod. Econ. 143 (1), 198–206.
Kumar, C., Mittal, M.L., 2012. Application of artificial neural networks for sales forecasting in an Indian automobile manufacturing company. Adv. Manag. 5 (9), 20–24.
Kuo, C.F., Hsu, C.T., Fang, C.H., Chao, S.M., Lin, Y.D., 2013. Automatic defect inspection system of colour filters using Taguchi-based neural network. Int. J. Prod. Res. 51 (5), 1464–1476.
Kuo, R.J., Wang, Y.C., Tien, F.C., 2010. Integration of artificial neural network and MADA methods for green supplier selection. J. Clean. Prod. 18 (12), 1161–1170.
Kwon, H., 2014. Performance modeling of mobile phone providers: a DEA-ANN combined approach. Benchmark.: Int. J. 21 (6), 1120–1144.
Li, Y., Liu, L., 2012. Hybrid artificial neural network and statistical model for forecasting project total duration in earned value management. Int. J. Netw. Virtual Organ. 10 (3/4), 402–413.
Liao, H., Li, Z., 2008. Multiobjective design of equivalent accelerated life testing plans. Int. J. Reliab. Qual. Saf. Eng. 15 (6), 515–538.
Liao, H., Wang, B., Weyman-Jones, T., 2007. Neural network based models for efficiency frontier analysis: an application to East Asian economies' growth decomposition. Glob. Econ. Rev. 36 (4), 361–384.
Lim, S.H., Lovell, C., 2008. Short-run total cost change and productivity of US Class I railroads. J. Transp. Econ. Policy 42 (1), 155–188.
Lim, S.H., Lovell, C., 2009. Profit and productivity of US Class I railroads. Manag. Decis. Econ. 30 (7), 423–442.
Liu, H.H., Chen, T.Y., Chiu, Y.H., Kuo, F.H., 2013. A comparison of three-stage DEA and artificial neural network on the operational efficiency of semi-conductor firms in Taiwan. Mod. Econ. 4 (1), 20–31.
Liu, J.S., Lu, L.Y., Lu, W., Lin, B.Y., 2013a. Data Envel. Anal. 1978–2010: A Cit. Based Lit. Surv. Omega 41 (1), 3–15.
Liu, J.S., Lu, L.Y., Lu, W., Lin, B.Y., 2013b. A survey of DEA applications. Omega 41 (5), 893–902.
Merkert, R., Smith, A.J., Nash, C.A., 2010. Benchmarking of train operating firms - a transaction cost efficiency analysis. Transp. Plan. Technol. 33 (1), 35–53.
Morris, B., 2013. Boom times on the tracks: rail capacity, spending soar. Wall Str. J..
Mostafa, M., 2009a. A probabilistic neural network approach for modeling and classifying efficiency of GCC banks. Int. J. Bus. Perform. Manag. 11 (3), 236–258.
Mostafa, M., 2009b. Modeling the efficiency of top Arab banks: a DEA-neural network approach. Expert Syst. Appl. 36 (1), 309–320.
Mostafa, M., 2009c. Modeling the competitive market efficiency of Egyptian companies: a probabilistic neural network. Expert Syst. Appl. 36 (5), 8839–8848.
O'Mahony, M., Oulton, N., 2000. International comparisons of labor productivity in transport and communications: the US, the UK and Germany. J. Product. Anal. 14 (1), 7–30.
Parisio, L., 1999. A comparative analysis of European railroads efficiency: a cost frontier approach. Appl. Econ. 31 (7), 815–823.

Parry, M.E., Cao, Q., Song, M., 2011. Forecasting new product adoption with probabilistic neural networks. J. Prod. Innov. Manag. s1, 78–88.
Pendharkar, P., 2005. A data envelopment analysis-based approach for data preprocessing. IEEE Trans. Knowl. Data Eng. 17 b (10), 1379–1388.
Pendharkar, P., 2011. A hybrid radial basis function and data envelopment analysis neural network for classification. Comput. Oper. Res. 38 (1), 256–266.
Pendharkar, P.C., Rodger, J.A., 2003. Technical efficiency-based selection of learning cases to improve forecasting accuracy of neural networks under monotonicity assumption. Decis. Support Syst. 36 (1), 117–136.
Salerian, J., Chan, C., 2005. Restricting multi-output multi-input DEA models by disaggregating the output-input vector. J. Product. Anal. 24, 5–29.
Samoilenko, S., Osei-Bryson, K., 2013. Using Data Envelopment Analysis (DEA) for monitoring efficiency-based performance of productivity-driven organizations: Design and implementation of a decision support system. Omega 41 (1), 131–142.
Santin, D., 2008. On the approximation of production functions: a comparison of artificial neural networks frontiers and efficiency techniques. Appl. Econ. Lett. 15 (8), 597–600.
Santín, D., Delgado, F.J., Valiño, A., 2004. The measurement of technical efficiency: a neural network approach. Appl. Econ. 36 (6), 627–635.
Seiford, L.M., Zhu, J., 1999. Profitability and marketability of the top 55 U.S. commercial banks. Manag. Sci. 45 (9), 1270–1288.
Sermpinis, G., Theofilatos, K., Karathanasopoulos, A., Georgopoulos, E.F., Dunis, C., 2013. Forecasting foreign exchange rates with adaptive neural networks using radial-basis functions and Particle Swarm Optimization. Eur. J. Oper. Res. 225 (3), 528–540.
Shi, F.X., Lim, S.H., Junwook, C., 2011. Railroad productivity analysis: case of the American Class I railroads. Int. J. Product. Perform. Manag. 60 (4), 372–386.
Sreekumar, S., Mahapatra, S.S., 2011. Performance modeling of Indian business schools: a DEA-neural network approach. Benchmark.: Int. J. 18 (2), 221–239.
Thieme, R., Song, M., Calantone, R.J., 2000. Artificial neural network decision support systems for new product development project selection. J. Mark. Res. (JMR) 37 (4), 499–507.
Tsai, H.C., Chen, C.M., Tzeng, G.H., 2006. The comparative productivity efficiency for global telecoms. Int. J. Prod. Econ. 103 (2), 509–526.
Ülengin, F., Kabak, Ö., Önsel, S., Aktas, E., Parker, B.R., 2011. The competitiveness of nations and implications for human development. Socio-Econ. Plan. Sci. 45 (1), 16–27.
Vaninsky, A., 2004. Combining data envelopment analysis with neural networks: application to analysis of stock prices. J. Inf. Optim. Sci. 25 (3), 589–611.
Wang, C.H., Lu, Y.H., Huang, C.W., Lee, J.Y., 2013. R & D, productivity, and market value: an empirical study from high-technology firms. Omega 41 (1), 143–155.
Wang, S., 2003. Adaptive non-parametric efficiency frontier analysis: a neural-network-based model. Comput. Oper. Res. 30 (2), 279–296.
Wu, D., 2009. Supplier selection: a hybrid model using DEA, decision tree and neural network. Expert Syst. Appl. 36 (5), 9105–9112.
Wu, D., Yang, Z., Liang, L., 2006. Using DEA-neural network approach to evaluate branch efficiency of a large Canadian bank. Expert Syst. Appl. 31 (1), 108–115.

Dr. He-Boong Kwon is an Associate Professor of Management at Colorado State University-Pueblo, USA. Dr. Kwon holds a doctoral and Master's degree in Electrical Engineering from Florida Institute of Technology. He was enrolled in an MBA program with emphasis on Technology and Innovation Management from Pacific Lutheran University, and finished AACSB AQ Bridge to Business Program from University of Toledo. He has 22 years of field experience in operations and quality management, program management, and technology management throughout his tenure in Korean Air Force and Samsung-Thales Company in Korea. His articles have been published in journals including *European Journal of Innovation Management, Expert Systems with Applications, International Journal of Procurement Management, International Journal of Logistics Management, Benchmarking: International Journal, and International Journal of Productivity and Quality Management*. His research interest includes Operations and Supply Chain Management and Technology and Innovation Management.