Int. J. Engng Ed. Vol. 25, No. 2, pp. 391–402, 2009
Printed in Great Britain.

0949-149X/91 $3.00+0.00

Two-Step Data Envelopment Analysis Approach for Efficient Engineering Enrollment Management*

ELIF KONGAR
Departments of Mechanical Engineering and Technology Management, University of Bridgeport, 221 University Avenue, 141 Technology Building, Bridgeport, CT 06604, USA. E-mail: kongar@bridgeport.edu

TAREK M. SOBH
University of Bridgeport, 221 University Avenue, Bridgeport, CT 06604, USA

MAHESH BARAL
University of Bridgeport, 221 University Avenue, Bridgeport, CT 06604, USA

In this paper, we present two consecutive Data Envelopment Analysis (DEA) models to measure the relative efficiency of applicants to graduate programs in engineering and to compare these efficiencies with the success of these students in the program. The proposed performance criteria are determined depending on the current evaluation criteria in the School of Engineering at the University of Bridgeport. The steps and implementation of the proposed methodology are explained with the help of a numerical example for the Fall 2004 semester.

Keywords: graduate enrollment; engineering; decision making; engineering education; Data Envelopment Analysis.

INTRODUCTION

EVALUATING CANDIDATES for graduate degree programs has always been a concern for both academic and administrative personnel at universities. The difficulty of this task has increased over time owing to the growing complexity and size of the pool of applicants as educational programs extend to the global arena. Many universities are facing a significant increase in the number of international student applications to graduate degree programs.

With this being the motivation, this study aims at determining the key criteria for applicants to the graduate programs at the University of Bridgeport, School of Engineering. In this regard, a two-step approach is developed. In the first step, an output oriented Data Envelopment Analysis (DEA) model has been used to evaluate and rank the accepted applicants depending on various criteria, for example, GRE and TOEFL scores, GPA, number of below-B grades in the Bachelor of Science transcripts, and other parameters. Following this, an additional ranking algorithm is implemented and run to determine the degree of success among the same set of accepted students, following their progress in the program until they graduate.

The results of the two ranking algorithms are then compared to validate the appropriateness of the selection criteria. A case study is included to demonstrate the steps and applicability of the proposed DEA approach.

Data envelopment analysis (DEA) is a widely applied linear programming-based technique first developed by Charnes et al. [1] in 1978 to evaluate the efficiency of a set of decision-making units. Since then, DEA has been recognized as an excellent methodology for modeling operational processes, and its empirical orientation and absence of a priori assumptions has resulted in its use in a number of studies involving efficient frontier estimation in the nonprofit sector, in the regulated sector, and in the private sector [2].

The paper is organized as follows: A brief list of previous studies is summarized in the following section. A summary of the Data Envelopment Analysis approach is then provided in the next section. This is followed by a section focusing on the problem description and a case study. The conclusions and thoughts for future research are then provided in the last section.

LITERATURE REVIEW

Even though there has been a large body of literature aiming at increasing the overall student achievement via alternative course designs, such as interactive learning, teamwork, etc., there is a lack of literature monitoring a student's progress starting from the admission phase through to graduation to detect the significant performance

* Accepted 29 November 2008.

indicators. Therefore, admission decisions are currently made by depending on varying criteria for each university and department, in a way that is not necessarily analytically justified.

One of the few studies aiming at detecting student performance indicators was by Deniz and Ersan [3]. In this study, stating the need of universities to have extensive analysis capabilities of student achievement levels in order to make appropriate academic decisions, the authors developed a software package entitled 'the Performance-based Academic Decision-Support System (PADSS)', which is based on various performance parameters that are visible at the student department, school and university levels.

Furthermore, in order to ensure the desired outcome achievement for students, Besterfield-Sacre et al. [4] developed, evaluated, and validated a representative model for the engineering education system at the University of Pittsburgh. The authors' model provides insight into success factors and educational processes that influence outcome achievement.

With a similar motivation, this study aims at determining the key criteria for applicants to graduate programs using a two-step Data Envelopment Analysis approach.

Usually modeled as a linear programming (LP) model, Data Envelopment Analysis (DEA) provides relative efficiency score for each decision making unit under consideration. The approach has the ability to accommodate multiple inputs and multiple outputs, allowing these variables to be included in the model with different units of measurement. Owing to these advantages and its ease of use, the approach has been employed extensively in various areas, such as health care, education, banking, manufacturing, and management.

One of the most relevant studies was by Johnson and Zhu [5]. In their work, the authors employed DEA to select the most promising candidates to fill an open faculty position. In this regard, the authors proposed a DEA aided recruiting process that: (1) determines the performance levels of the 'best' candidates relative to other applicants; (2) evaluates the degree of excellence of 'best' candidates' performance; (3) forms consistent tradeoff information on multiple recruiting criteria among search committee members, and then (4) clusters the applicants.

DEA also found a large variety of applications in the environmental arena. To this extend, Sarkis [6] proposed a two-stage methodology to integrate managerial preferences and environmentally conscious manufacturing (ECM) programs. Consequently, Sarkis and Cordeiro [7] investigated the relationship between environmental and financial performance at the firm level.

Furthermore, Talluri et al. [8] applied DEA and Goal Programming methods to a Value Chain Network (VCN) considering the cross efficiency evaluations of Decision Making Units (DMUs).

Methods other than DEA have also been used to study the efficiency of application and admission processes. Moore [9] built an operational two-stage expert system to examine the admission decision process for applicants to an MBA program, and predict the degree completion potential for those actually admitted. A similar study is also published by Nilsson [10] to investigate any differences in the predictive relationships between the scores of the Graduate Record Examination (GRE), the graduate grade point average, the scores of the Graduate Management Admission Test (GMAT), and the graduate grade point average. Furthermore, Landrim et al. [11] constructed a value tree diagram for fifty-five graduate institutions offering the Ph.D. degree in psychology. The authors made use of this diagram to indicate the relative weight of admission factors used in the decision making process.

This study is a follow up on Kongar and Sobh's [12] previously published work where the authors proposed a DEA approach to measure the relative efficiency of applicants to the graduate programs in engineering. The proposed performance criteria in the study were determined depending on the current evaluation criteria in the School of Engineering at the University of Bridgeport.

INTRODUCTION TO THE DATA ENVELOPMENT ANALYSIS APPROACH

Data Envelopment Analysis (DEA) is a non-parametric approach that compares similar entities, i.e. decision making units (DMUs), against the 'best virtual decision making unit'. Usually modeled as a linear programming (LP) model, the method provides a relative efficiency score for each decision making unit under consideration.

The most appealing advantage of DEA is that, unlike parametric approaches such as regression analysis (RA), DEA optimizes on each individual observation and does not require a single function that suits all observations best (Charnes et al. [13]). Comparisons of DEA and RA have been well recorded in the literature. The majority of the published work accepts that DEA is more advantageous in comparing decision making units, even though there are some studies emphasizing the advantages of both (i.e. see Thanassoulis [14]).

One of the above mentioned comparative studies is by Banker et al. [15], comparing estimates of technical efficiencies of individual hospitals obtained from the econometric modeling of the translog cost function, and the application of DEA. In their study, the authors reported that DEA estimates were highly related to capacity utilization, whereas translog estimates failed to provide such a relationship.

In addition, Bowlin et al. [16] compared DEA with RA using 15 hypothetical hospitals and concluded that DEA outperformed RA with its ability to identify the sources of inefficiencies by underlining the resources that are used in excess in inefficient hospitals. Furthermore, the authors stated that DEA performed better in estimating and returning scale characterizations. In addition, Sarkis [17] compared DEA with conventional multiple criteria decision making (MCDM) tools in terms of efficiency and concluded that DEA appeared to perform well as a discrete alternative MCDM tool.

DEA algorithms can be classified into two categories: *input-* and *output-oriented* DEA models, according to the 'orientation' of the model. Input-oriented DEA models concentrate on reducing the amount of input by keeping the output constant. Output-oriented DEA models, on the other hand, focus on maximizing the amount of output with the identical amount of input. In DEA modeling, inputs are considered as the items that are subject to minimization (i.e., less is better), whereas, outputs are the items that are subject to maximization (i.e., more is better).

Further classification of DEA models can be given depending on the 'optimality scale' criterion. Here, DEA models can work under the assumption of Constant Returns to Scale (CRS), or non-constant returns to scale, i.e. 'Increasing Returns to Scale (IRS)', 'Decreasing Returns to Scale (DRS)', and 'Variable Returns to Scale (VRS)', implying that not all DMUs are functioning at an optimality scale. Here, CRS assumes changes in output values subsequent to a proportional change in the input values. VRS was initially introduced by Banker et al. [18] as an extension of the CRS DEA model. In this paper, we employ an output oriented CRS DEA model. Further explanation regarding the CRS model follows.

As also mentioned above, a basic DEA model allows the introduction of multiple inputs and multiple outputs and obtains an 'efficiency score' of each DMU with the conventional output/input ratio analysis. Defining basic efficiency as the *ratio of weighted sum of outputs to the weighted sum of inputs*, the relative efficiency score of a test DMU p can be obtained by solving the following DEA ratio model (CCR) proposed by Charnes et al. [1]:

$$\begin{aligned} \max \quad & \frac{\sum_{k=1}^{s} v_k y_{kp}}{\sum_{j=1}^{m} u_j x_{jp}} \\ \text{s.t.} \quad & \frac{\sum_{k=1}^{s} v_k y_{ki}}{\sum_{j=1}^{m} u_j x_{ji}} \leq 1 \quad \forall \text{ DMUs } i \\ & v_k, u_j \geq 0 \qquad \forall\ k,j. \end{aligned} \tag{1}$$

where

$k = 1$ to s,
$j = 1$ to m,
$i = 1$ to n,
y_{ki} = amount of output k produced by DMU i,
x_{ji} = amount of input j produced by DMU i,
v_k = weight assigned to output k,
u_j = weight assigned to input j.

Equation (1) can easily be converted into a linear program as in Equation (2). We refer the reader to the study by Charnes et al. [13] for further explanation of the model.

$$\begin{aligned} \max \quad & \sum_{k=1}^{s} v_k y_{kp} \\ \text{s.t.} \quad & \sum_{j=1}^{m} u_j x_{jp} = 1 \\ & \sum_{k=1}^{s} v_k y_{ki} - \sum_{j=1}^{m} u_j x_{ji} \leq 0 \quad \forall \text{ DMUs } i \\ & v_k, u_j \geq 0 \qquad \forall\ k,j, \end{aligned} \tag{2}$$

where the

$$\sum_{j=1}^{m} u_j x_{jp} = 1$$

constraint sets an upper bound of 1 for the relative efficiency score.

In the CCR model provided in Equation (2), evaluating the efficiency of n DMUs correspond to a set of n LP problems. Using duality, the dual of the CRS model can be represented as in Equation (3):

$$\begin{aligned} \min \quad & \theta \\ \text{s.t.} \quad & \sum_{i=1}^{n} \lambda_i x_{ji} - \theta x_{jp} \leq 0 \quad \forall \text{ Inputs } j \\ & \sum_{i=1}^{n} \lambda_i y_{ki} - y_{kp} \geq 0 \quad \forall \text{ Outputs } k \\ & \lambda_i \geq 0 \qquad \forall \text{ DMUs } i. \end{aligned} \tag{3}$$

Equation (3) corresponds to the dual of the basic input-oriented CCR model assuming constant returns to scale for all the inputs and outputs. Using Talluri's [19] notation, the dual of a basic output-oriented CRS model can be written as follows:

$$\begin{aligned} \max \quad & \phi \\ \text{s.t.} \quad & x_{jp} - \sum_i \lambda_i x_{ji} \geq 0 \quad \forall \text{ Inputs } j \\ & -\phi y_{kp} + \sum_i \lambda_i y_{ki} \geq 0 \quad \forall \text{ Outputs } k \\ & \lambda_i \geq 0 \qquad \forall \text{ DMUs } i. \end{aligned} \tag{4}$$

In the case where the assumption is that not all DMUs are functioning at an optimality scale, Equation (4) could be converted into a VRS model by including the constraint

$$\sum_i \lambda_i \geq 0$$

to the set of technological constraints.

The result of the model, ϕ is the relative efficiency score of each DMU. Inverse of the variable

ϕ ($1/\phi$) provides the technical efficiency value (TE) for each DMU. Here, given that the technical efficiency value is equal to one ($TE = 1$), DMU p is considered 'efficient' for its selected weights. In this case, DMU p lies on the optimal frontier and is not dominated by any other DMU. Using similar reasoning, if the technical efficiency value is less than one ($TE < 1$), then it can be claimed that DMU p is not on the optimal frontier and there exists at least one efficient DMU in the population.

The following demonstrates the application of the CRS DEA model to the evaluation process of the applicants for graduate engineering programs.

APPLYING DATA ENVELOPMENT ANALYSIS TO THE APPLICATION REVIEW PROCESS

The proposed DEA model in this study aims at (1) accepting students, (2) comparing the accepted students with the DEA model results, and (3) preparing a base to observe the students' future success to evaluate the performance criteria fed into the model.

To achieve these objectives, the data for all 37 candidates ($n = 37$) for the Masters of Science (M.S.) in the Computer Science program in the School of Engineering for Fall 2004 semester was collected.

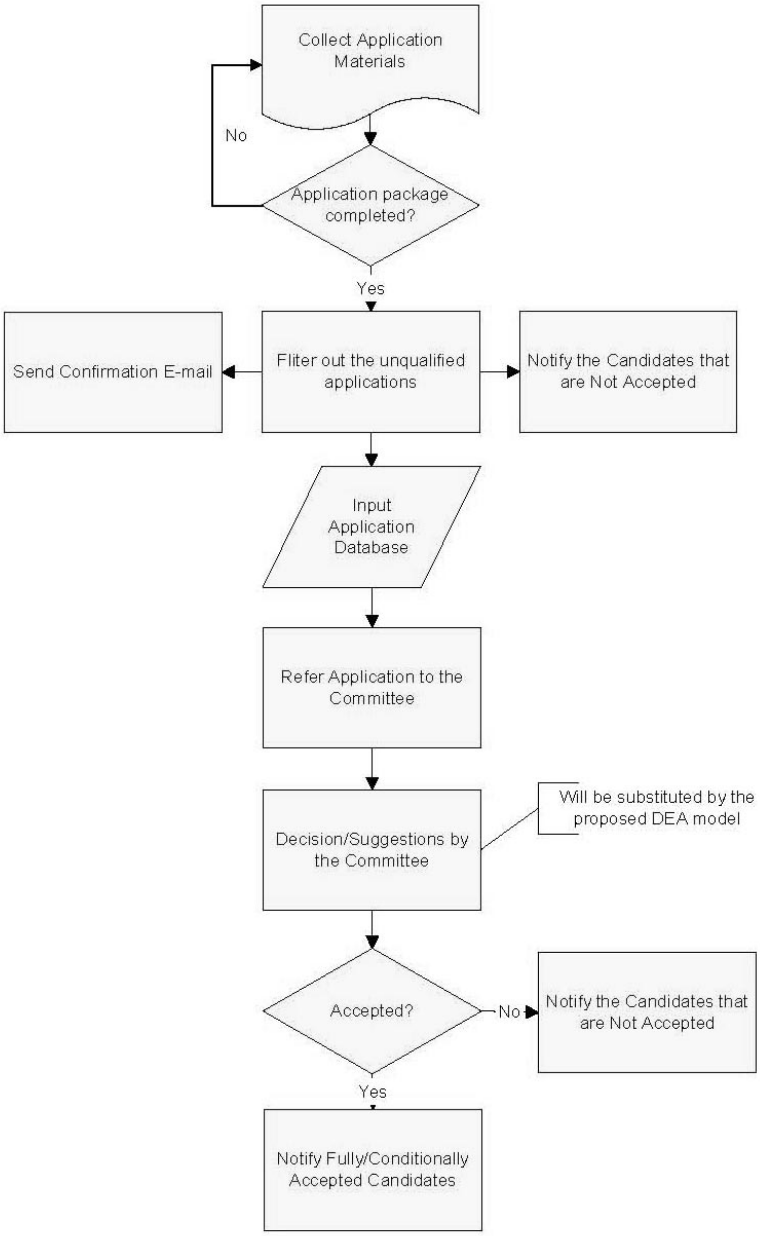

Fig. 1. Simplified schematic diagram of the application evaluation and decision making process.

After reading in the relevant data, a DEA model was employed to evaluate the relative efficiency of each candidate using six performance criteria, viz., the Bachelors of Science (B.S.) GPA (B.S. GPA), TOEFL and GRE Quantitative (GRE-Q) scores, number of years of work experience, number of undergraduate semesters until B.S. degree completion, and the number of below-B grades in math-related and technical courses in the B.S. degree transcript.

DEA model for the evaluation process

Following the retrieval of the complete application materials, related data is entered into the applications database. The office of admissions then sends each applicant a confirmation e-mail with an assigned University of Bridgeport (UB) identification number confirming that the application has been received.

Subsequently, the applications are filtered by the office of admissions depending on basic application criteria, filtering out unqualified applicants. These applicants are then notified regarding the result of their applications. Remaining applications that meet the basic requirements are then sent to the relevant Faculty for decision making (Fig. 1).

The information provided by this study enables users to identify the best candidates for the graduate engineering program. In the following sections, we illustrate how the evaluation process can be enhanced using the DEA approach introduced earlier.

DEA model I to evaluate the efficiency of candidates for graduate study

In the proposed model, the applications to the graduate program correspond to decision-making units in DEA, while application data correspond to criteria in DEA, dependent on the definition of the indicators (inputs or outputs in the DEA model) [12].

In total, the model embodies 108 decision-making units and six criteria. These criteria include two inputs and four outputs. Input criteria consist of e_1, and e_2, whereas output criteria include e_3, e_4, $e_{5,}$ and, e_6, where:

$e_1 =$ number of below-B grades in math-related/technical courses in the B.S. transcript of the applicant,

$e_2 =$ number of semesters that the applicant spent to complete the B.S. degree,

$e_3 =$ B.S. GPA of the applicant,

Table 1. Initial data for the DEA model I

DMU #	e_1	e_2	e_3	e_4	e_5	e_6	DMU #	e_1	e_2	e_3	e_4	e_5	e_6	DMU #	e_1	e_2	e_3	e_4	e_5	e_6
1	13	8	2.87	597	720	0	37	18	8	2.75	637	700	1	73	11	8	3.20	507	770	0
2	26	8	2.77	563	620	0	38	13	10	2.82	593	780	2	74	0	8	2.37	574	693	0
3	19	8	3.00	597	780	0	39	16	8	3.14	473	690	0	75	5	6	3.14	490	750	0
4	9	6	2.90	560	640	4	40	23	10	2.94	473	530	0	76	0	8	3.98	553	800	0
5	32	12	2.34	613	650	0	41	5	8	3.35	620	720	0	77	18	8	2.92	677	790	1
6	39	8	1.71	563	630	0	42	15	8	2.82	637	660	0	78	20	10	2.97	633	780	0
7	20	8	3.09	567	590	0	43	15	10	2.85	610	770	0	79	8	8	3.10	563	660	2
8	22	8	2.95	473	650	0	44	10	8	3.07	637	780	0	80	2	8	3.56	593	800	0
9	16	8	3.07	627	570	0	45	19	8	2.61	620	720	0	81	23	8	2.98	523	660	2
10	6	8	3.50	560	710	0	46	0	6	2.10	473	690	0	82	15	8	3.24	563	700	0
11	26	10	2.19	610	620	0	47	18	8	3.13	603	720	0	83	0	6	3.77	597	600	0
12	20	8	2.98	567	520	0	48	16	8	3.04	573	720	0	84	6	8	3.41	593	660	0
13	23	8	2.94	610	750	0	49	13	8	3.24	473	630	0	85	1	8	3.85	600	770	0
14	24	10	2.63	537	740	0	50	20	8	2.70	670	710	0	86	11	8	3.33	550	570	0
15	21	8	2.81	587	750	0	51	8	8	3.33	567	750	0	87	1	8	3.68	480	693	2.5
16	15	8	2.68	543	690	1	52	20	8	2.30	567	590	0	88	0	6	4.00	603	660	0
17	15	8	3.20	550	690	0	53	23	8	2.79	547	690	0	89	1	8	3.92	643	800	0
18	11	8	2.95	650	770	0	54	20	8	2.44	473	690	0	90	9	8	3.37	627	710	0
19	20	8	2.60	637	690	0	55	17	8	2.74	593	710	0	91	17	8	3.11	560	610	0
20	34	10	2.52	593	680	0	56	33	8	1.70	647	710	0	92	12	8	3.32	610	730	0
21	21	8	2.69	620	620	0	57	17	8	2.78	500	720	0	93	6	6	3.68	574	693	2
22	18	8	2.90	560	710	0	58	20	8	2.93	530	770	1	94	0	6	3.40	574	693	5
23	24	8	2.87	560	690	0	59	16	8	3.13	560	650	0	95	12	8	3.24	577	730	0
24	4	6	2.84	473	690	0	60	8	8	3.40	587	690	1	96	9	8	3.04	583	580	0
25	24	8	2.98	527	440	0	61	17	8	3.12	633	550	0	97	0	8	2.97	560	760	0
26	19	8	3.08	650	720	0	62	36	8	2.18	627	750	0	98	14	8	3.03	550	730	0
27	29	8	2.40	483	340	0	63	18	8	2.97	587	760	0	99	7	8	3.34	560	640	0
28	26	10	2.70	567	680	0	64	3	6	3.00	587	570	2	100	9	8	3.34	550	620	0
29	9	8	3.20	530	730	0	65	12	8	2.94	677	750	0	101	11	8	3.07	647	630	0
30	11	8	3.43	550	140	0	66	23	8	2.84	537	380	0	102	7	8	3.52	563	670	0
31	12	10	2.90	637	770	0	67	11	8	3.04	587	670	0	103	1	6	3.38	653	760	7
32	16	8	3.24	577	590	0	68	13	8	2.37	577	680	0	104	3	8	3.67	560	610	0
33	17	8	3.17	560	650	0	69	21	8	2.78	537	550	0	105	2	6	3.50	574	693	8
34	22	8	3.03	620	710	0	70	19	8	3.10	597	740	0	106	0	8	3.44	587	770	0
35	34	10	2.50	563	760	0	71	13	8	3.04	620	730	0	107	10	8	3.00	567	540	0
36	14	10	2.90	553	640	0	72	8	8	3.22	477	640	0	108	18	8	2.57	547	670	0
														Ave.	14	8	3.01	575	675	0.4

e_4 = TOEFL score of the applicant,
e_5 = GRE-Q score of the applicant,
e_6 = number of years of work experience of the applicant.

The first input introduced to the model is the number of below-B grades in math-related/technical courses in the B.S. transcript (e_1). Following the notation of the first DEA model, the first input formulation for each DMU i (x_{1i}) can be written as follows:

$$x_{1i} = e_{1i} \qquad \forall \text{ DMUs } i. \tag{5}$$

The second input introduced to the model is the number of semesters spent to complete the B.S. degree, (e_2). Hence, the second input formulation for each DMU i (x_{2i}) can be written as follows:

$$x_{2i} = e_{2i} \qquad \forall \text{ DMUs } i. \tag{6}$$

The output variables in the proposed DEA model are selected as: the B.S. GPA of the applicant (e_3); the TOEFL score of the applicant (e_4); the GRE-Q score of the applicant (e_5); and the number of years of previous work experience (e_6) of the applicant.

Therefore, with similar reasoning, Equations (7), (8), (9), and (10) can be expressed mathematically as follows:

$$x_{1i} = e_{3i} \qquad \forall \text{ DMUs } i. \tag{7}$$

$$x_{2i} = e_{4i} \qquad \forall \text{ DMUs } i. \tag{8}$$

$$x_{3i} = e_{5i} \qquad \forall \text{ DMUs } i. \tag{9}$$

$$x_{4i} = e_{6i} \qquad \forall \text{ DMUs } i. \tag{10}$$

This completes the formulation of the DEA model. Selected application data for a total of 108 candidates are provided in Table 1.

Using this data set, the output-oriented DEA model is run for each applicant in the sample using DEA-Solver-PRO 5.0. DEA-Solver-PRO is a DEA software designed on the basis of the textbook by Cooper et al. [20] to solve and analyze DEA models. The results of the model are presented in Table 2 in descending order of *TE* I values.

According to the DEA results depicted in Table 2, Candidates 105, 103, 94, and 88 are efficient in terms of their pre-application academic performances with technical efficiency (*TE* I) values equal to 1. All other applicants have a potential to increase the relative efficiency of academic performances by 1 minus the *T*E value. For instance, the efficiency of candidate 9 could be increased by 28.0%. The two lowest technical

Table 2. Relative efficiency score (TE I) and rank of each candidate

Rank	DMU	Score	Rank	DMU	Score	Rank	DMU	Score
1	105	1.000	37	19	0.732	73	96	0.671
2	103	1.000	38	37	0.732	74	68	0.671
3	94	1.000	39	42	0.732	75	30	0.670
4	88	1.000	40	70	0.730	76	33	0.667
5	46	0.996	41	90	0.728	77	79	0.667
6	83	0.990	42	92	0.727	78	72	0.666
7	75	0.987	43	61	0.727	79	100	0.666
8	93	0.986	44	41	0.724	80	59	0.664
9	24	0.908	45	29	0.720	81	86	0.663
10	64	0.899	46	98	0.720	82	49	0.662
11	76	0.868	47	95	0.720	83	7	0.661
12	4	0.858	48	71	0.720	84	108	0.661
13	106	0.833	49	9	0.720	85	91	0.657
14	89	0.823	50	45	0.712	86	107	0.655
15	97	0.823	51	21	0.712	87	81	0.655
16	85	0.799	52	34	0.712	88	12	0.654
17	80	0.790	53	60	0.712	89	52	0.651
18	77	0.780	54	102	0.711	90	8	0.647
19	65	0.778	55	1	0.711	91	2	0.647
20	3	0.770	56	47	0.711	92	6	0.647
21	44	0.770	57	48	0.711	93	25	0.622
22	50	0.770	58	57	0.711	94	66	0.621
23	18	0.760	59	84	0.703	95	69	0.617
24	73	0.760	60	82	0.703	96	78	0.616
25	58	0.760	61	55	0.701	97	38	0.616
26	74	0.750	62	22	0.701	98	31	0.608
27	63	0.750	63	104	0.694	99	43	0.608
28	26	0.747	64	17	0.693	100	35	0.600
29	101	0.743	65	39	0.688	101	14	0.584
30	56	0.743	66	54	0.681	102	11	0.560
31	13	0.740	67	16	0.681	103	27	0.555
32	15	0.740	68	53	0.681	104	20	0.545
33	62	0.740	69	23	0.681	105	28	0.537
34	51	0.740	70	32	0.679	106	36	0.510
35	87	0.739	71	99	0.677	107	5	0.469
36	10	0.732	72	67	0.674	108	40	0.462
							Ave.	0.719

efficiency values are calculated for Candidates 5 and 40 with 46.9%, and 46.2%, respectively.

These low values are most probably driven by e_1, the high numbers of below-B grades in math-related/technical courses in the B.S. transcript (32 and 23, respectively) and e_3, the low GPAs of the applicants (2.34 and 2.94, respectively). In addition, the TOEFL score of the applicant 40 is at a very low level, i.e., $e_4 = 473$, placing the applicant to the lowest rank level.

The average efficiency for the sample is 71.9%. Figure 2 represents the average efficiency and the *TE* I values for the 108 candidates in the population. As illustrated by Fig. 2, 59 candidates fall below the average efficiency value (approx. 55% of the candidates).

As we analyze the results further, we can easily observe that all of the efficient candidates have completed their B.S. degrees in an identical number of semesters (6). In addition, the efficient candidates are characterized by either significantly high GPAs, GRE-Q scores, years of work experience, significantly low numbers of below-B grades in math-related/technical courses, or a combination of these criteria.

With this in mind, depending on the importance of each criterion, the input data can be normalized and weighed according to the decision maker preferences, so that the more important criterion would provide competitive advantage to the candidate.

In the following, a subsequent DEA model (DEA model II) is proposed to measure the relative efficiency of the future success of M.S. candidates.

DEA model II to evaluate the efficiency of candidates for graduate study

In this section, an output-oriented DEA model (DEA model II) is constructed to seek a relationship between the relative efficiency measures of the graduate students and their success in the graduate program. In this regard, 37 DMUs are selected representing the applicants that are accepted to the Computer Science graduate degree program at the University of Bridgeport. These DMUs are provided at the bottom 37 of Table 2, i.e., DMUs 72 through 108.

The model embodies four criteria, including three inputs and one output. The input criteria include t_1, whereas the output criteria include, t_2, t_3, and, t_4, where:

t_1 = number of below-C grades in the M.S. transcript of the M.S. candidate,
t_2 = GPA of the M.S. candidate,
t_3 = application status for the Curricular Practical Training (CPT) or Optional Practical Training (OPT) programs for the M.S. candidate; indicating whether they have applied to an industry internship during the program or a full-time position immediately following graduation,
t_4 = graduation status of the M.S. candidate.

The first input introduced to the DEA model II is the number of below-C grades in the M.S. transcript (t_1). Following the notation of the first DEA model, the first input formulation for each DMU i (x_{1i}) can be written as follows:

$$x_{1i} = t_{1i} \qquad \forall \text{ DMUs } i. \tag{11}$$

The output variables in the proposed DEA model are selected as, the GPA of the M.S. candidate (t_2), the application status for CPT or OPT of the M.S. candidate (t_3), and the graduation status for of the M.S. candidate.

Therefore, with similar reasoning, Equations

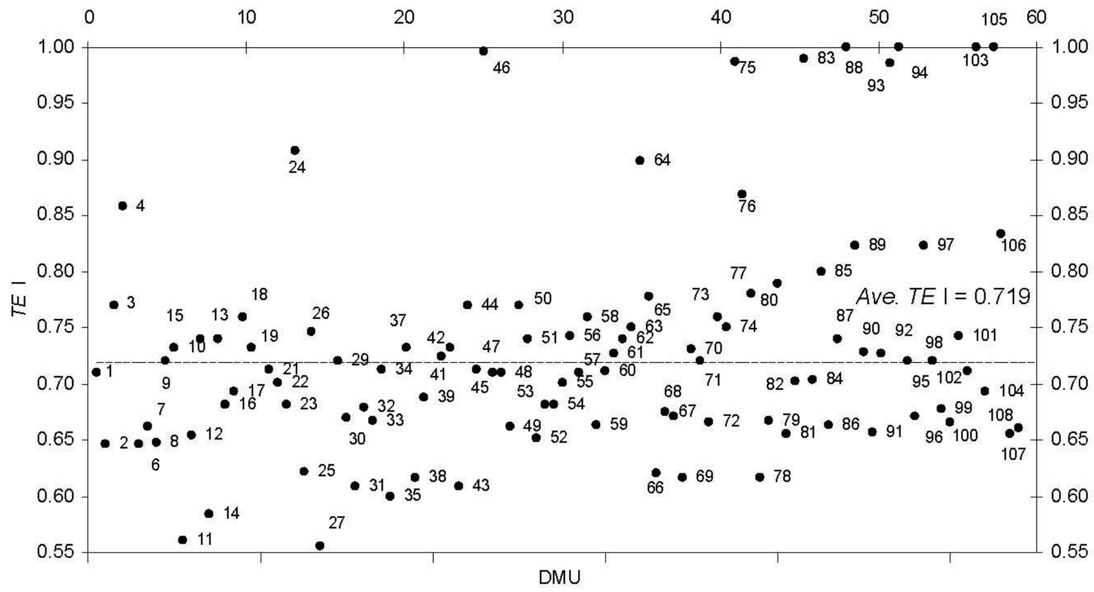

Fig. 2. Performance efficiencies of 108 candidates according to the DEA I model results. Ave. *TE* I = 0.719.

Table 3. Initial data for the DEA model II

DNU #	t_1	t_2	t_3	t_4	DNU #	t_1	t_2	t_3	t_4
72	1	3.12	2	2	91	0	2.34	1	1
73	0	3.21	2	2	92	0	3.42	2	2
74	0	0	1	1	93	0	3.38	2	2
75	0	3.03	2	1	94	3	2.07	1	1
76	0	4	1	1	95	0	2.67	1	1
77	0	3.58	2	2	96	0	3.58	2	2
78	0	3.49	2	2	97	0	3.24	2	2
79	0	3.56	2	1	98	0	2	1	1
80	0	3.46	1	2	99	2	0	1	1
81	2	2.4	1	1	100	0	3.14	2	2
82	0	3.18	2	2	101	0	3.43	1	2
83	0	3.27	2	2	102	0	2.45	1	1
84	0	3.3	2	2	103	0	3.72	1	1
85	0	3.45	2	2	104	0	2.89	1	1
86	0	3.11	2	2	105	0	3.37	2	2
87	0	3.21	2	2	106	0	3.7	2	2
88	0	3.58	2	2	107	0	3.15	1	2
89	0	3	1	1	108	0	3.58	2	2
90	0	3.43	2	2	*Ave.*	0.2	3.01	1.6	1.6

(12), (13), and (14) can be expressed mathematically as follows:

$$y_{1i} = t_{2i} \quad \forall \text{ DMUs } i. \tag{12}$$

$$y_{2i} = t_{3i} \quad \forall \text{ DMUs } i. \tag{13}$$

$$y_{3i} = t_{4i} \quad \forall \text{ DMUs } i. \tag{14}$$

Here, for the application status for CPT or OPT of the M.S. candidate (y_{2i}), a positive integer value, '2', is assigned if the M.S. candidate has applied for either CPT or OPT, where as the remaining variables are assigned the value of '1'.

Similar logic has been applied to the graduation status of the M.S. candidate (y_{3i}) and a positive integer value, '2', is assigned if the M.S. candidate has graduated from the graduate degree program, where as '1' is assigned if the student has transferred out or if she/he is currently enrolled, but has not yet graduated.

The application data for a total of 37 candidates are given in Table 3.

The results of the model are presented in Table 4 in descending order of TE II values.

According to the DEA results shown in Table 4, twenty five candidates are efficient in terms of their post-application academic performances, with technical efficiency (TE II) values equal to 1. All other applicants have a potential to increase the relative efficiency of academic performances by 1 minus the TE II value.

These low values are most probably driven by the lack of OPT or CPT applications and failure to graduate.

Figure 3 represents the average efficiency, 82.2%, and the TE II values for the 37 candidates in the population. As illustrated by Fig. 3, 11 candidates fall below the average efficiency value.

Furthermore, it is difficult to establish a straight-forward or an obvious emerging pattern

Table 4. Relative efficiency score (TE II) and rank of each candidate

Rank	DMU #	*TE* II	Rank	DMU #	*TE* II
1	108	1	1	90	1
1	107	1	1	100	1
1	73	1	1	92	1
1	106	1	1	93	1
1	75	1	1	97	1
1	76	1	1	96	1
1	77	1	26	103	0.936
1	78	1	27	89	0.768
1	79	1	28	104	0.742
1	80	1	29	95	0.69
1	105	1	30	102	0.639
1	82	1	31	91	0.613
1	83	1	32	98	0.535
1	84	1	33	74	0.5
1	85	1	34	72	1.0×10^{-5}
1	86	1	35	81	3.1×10^{-6}
1	87	1	36	99	2.5×10^{-6}
1	88	1	37	94	1.8×10^{-6}
1	101	1	*Ave.*		0.822

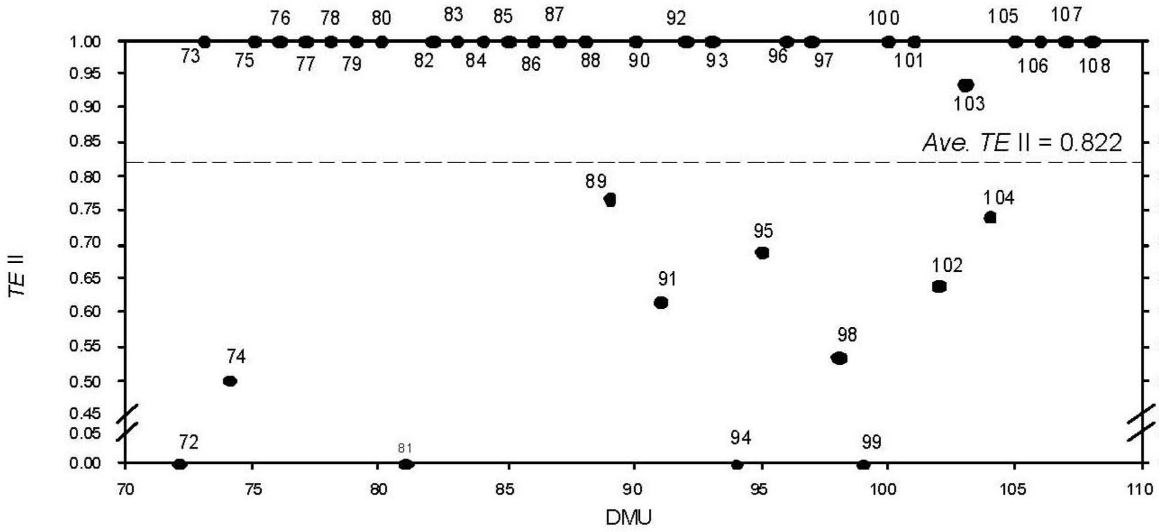

Fig. 3. Performance efficiencies of 37 candidates according to the DEA II model results. Ave. TE II = 0.822.

between the pre- and post-application relative efficiencies. However, as given in Fig. 4, we can observe that the proposed two-step DEA approach is more successful in determining the success of the applicants and is opposed to failures. In addition, it is also interesting to note that, 59 candidates selected by the DEA I model algorithm as inefficient and only 9 of these (app. 15%) are ranked as efficient, depending on their post-application performances by the DEA II model. This can be interpreted as both DEA models providing reasonable results.

In order to perform further analysis, one can compare the current admission process with the DEA model results. As explained above, candidates 72 through 108 are the applicants actually selected by the School of Engineering. According to the proposed DEA algorithms the top 37 candidates are provided in Table 5 with an average efficiency value of 82.7%.

Out of these, 24% of the applicants are selected as efficient units as per the DEA algorithm whereas the current selection process rejected the students. Since we lack the post-admission data for these candidates, it is not possible to perform a performance comparison, even though the proposed models can increase the commonality between the successful DEA model I and II applicants.

In addition, DEA I model assumes that the efficiency of each graduate degree program applicant is a function of simply five variables, viz., the number of below-B grades in math-related/technical courses in the B.S. transcript of the applicant (e_1), the number of semesters that the applicant took to complete the B.S. degree (e_2), the B.S. GPA of the applicant (e_3), the TOEFL score of the applicant (e_4), and the GRE-Q score of the applicant (e_5).

Furthermore, DEA II model structure is built

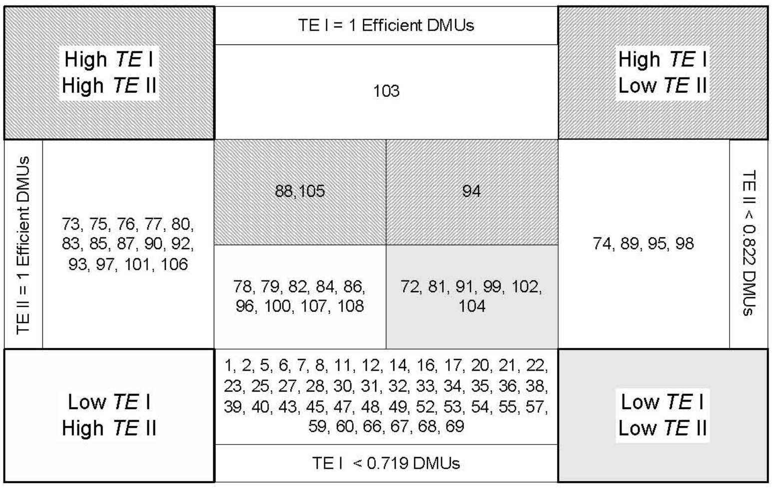

Fig. 4. Results summary of the DEA models I and II.

Table 5. Top 37 candidates

Rank	DMU	Score	Rank	DMU	Score
1	105	1	20	3	0.770
2	103	1	21	44	0.770
3	94	1	22	50	0.770
4	88	1	23	18	0.760
5	46	0.996	24	73	0.760
6	83	0.990	25	58	0.760
7	75	0.987	26	74	0.750
8	93	0.986	27	63	0.750
9	24	0.908	28	26	0.747
10	64	0.899	29	101	0.743
11	76	0.868	30	56	0.743
12	4	0.858	31	13	0.740
13	106	0.833	32	15	0.740
14	89	0.823	33	62	0.740
15	97	0.823	34	51	0.740
16	85	0.799	35	87	0.739
17	80	0.790	36	10	0.732
18	77	0.780	37	19	0.732
19	65	0.778		*Ave.*	0.827

under the assumption that the efficiency of candidates for graduate study solely depends on (1) the number of below-C grades in the M.S. transcript of the M.S. candidate (t_1), the GPA of the M.S. candidate (t_2), the application status for the Curricular Practical Training (CPT) or Optional Practical Training (OPT) programs for the M.S. candidate (t_3), and the graduation status of the M.S. candidate.

Therefore, DEA I and DEA II are far from reflecting real details of the efficiency measures of graduate degree applicants and graduate students even though both models cover the most influential factors for each corresponding performance measure.

CONCLUSIONS AND FUTURE RESEARCH

In this study, implementations of two output-oriented DEA models are considered and applied to a sample of 108 M.S. candidates to the Computer Science graduate degree program at the University of Bridgeport to determine the relative efficiency score of applicants based on their credentials. The model provides a basis to conduct a fast and reliable automated application evaluation process.

There is a significant difference between the manually accepted candidates and the candidates ranked according to the DEA model results. This was most likely caused by (1) the inconsistency of the manual evaluation process and/or (2) the presence of factors that are not included in the model; for example: the ranking of the university providing the B.S. degree, the B.S. major, the strength of the recommendation letters, etc. [21].

This study also looked at the accepted candidates and analyzed their future performance to seek a correlation between the students' performance in the graduate program after admission and to compare the existing evaluation results, towards the eventual implementation of an automated graduate application admission system.

Both DEA steps proposed in the paper utilize the data for students who are both accepted and enrolled in the graduate engineering program. However, a considerable portion of accepted students, approximately 70–75%, do not enroll in the degree program even though they are accepted. This is due to visa acquisition problems and/or personal preference in attending a different university. Furthermore, data for rejected students are either unavailable or unreliable due to the recording and privacy laws limitations. Hence, recording the applications to the school and tracking each application so that the data set will include every student who applied to the program would certainly provide much more reliable results.

In addition to the criteria used in the second DEA model, the duration of study, number of total credits and courses completed, and the numbers of grades less then C were also available. However, these data points were omitted, as they were considered to be not correlated with the graduate GPA and graduation status, which we considered, for the purpose of this study, to be the main indicators of success within the graduate course of study.

Furthermore, applications to OPT or CPT does not necessarily in all cases imply that the M.S. candidate has been employed by an organization. It only shows the intention of the M.S. candidate to seek employment in the U.S. after graduation. The employment data cannot be obtained in a reliable manner since keeping track of the employment status of graduate students is often difficult to accomplish in a timely manner.

In summary, the quality of the data greatly affects the outcome of the proposed models. In the future, we are planning to collect the data solely for this purpose and track students from the application stage and follow their progress until they graduate. We plan to perform more correlation studies between the admission and graduate performance models and vary/change the number of the parameters for both models in order to fine-tune our system; towards the eventual goal of successfully implementing a fully-automated graduate admission system.

Acknowledgements—The authors would like to acknowledge the significant contributions, effort and support provided by Bryan Gross and Isabella Varga, of the Office of Admissions, and Yumin Wang, Director, International Affairs Office, at the University of Bridgeport, and Mert Ozan Bahtiyar, MD at Yale University.

REFERENCES

1. A. Charnes, W. Cooper and E. Rhodes, *Measuring the Efficiency of Decision-Making Units, European Journal of Operational Research*, **2**(6), 1978, pp. 429–444.
2. L. M. Seiford and R. M. Thrall, Recent developments in Dea—the mathematical-programming approach to frontier analysis, *Journal of Econometrics*, **46**(1–2), 1990, pp. 7–38.
3. D. Z. Deniz and I. Ersan, *An* Academic decision–support system based on academic performance evaluation for student and program assessment, *International Journal of Engineering Education*, **18**(2), 2002, pp. 236–244.
4. M. Besterfield-Sacre, L. J. Shuman and H. Wolfe, Modeling undergraduate engineering outcomes, *International Journal of Engineering Education*, 2002, **18**(2), pp. 128–139.
5. S. A. Johnson and J. Zhu, Identifying 'best' applicants in recruiting using data envelopment analysis, *Socio-Economic Planning Sciences*, **37**(2), 2003, pp. 125–139.
6. J. Sarkis, A methodological framework for evaluating environmentally conscious manufacturing programs, *Computers & Industrial Engineering*, **36**(4), 1999, pp. 793–810.
7. J. Sarkis and J. J.Cordeiro, An empirical evaluation of environmental efficiencies and firm performance: pollution prevention versus end-of-pipe practice, *European Journal of Operational Research*, **135**(1), 2001, pp. 102–113.
8. S. Talluri, R. C. Baker and J. Sarkis, Framework for designing efficient value chain networks, *International Journal of Production Economics*, **62**(1–2), 1999, pp. 133–144.
9. J. S. Moore, An expert system approach to graduate school admission decisions and academic performance prediction, *Omega*, **26**(5), 1998, pp. 659–670.
10. J. E. Nilsson, The GRE and the GMAT: A comparison of their correlations to GGPA, *Educational and Psychological Measurement*, **55**, 1995, pp. 637–40.
11. R. E. Landrim, E. B. Jeglum and J. R. Cashin, The decision-making process of graduate admissions committees in psychology, *Journal of Social Behavior and Personality*, **9**(2), 1994, pp. 239–248.
12. E. Kongar and T.Sobh, Evaluating the efficiency of candidates for graduate study via data envelopment analysis, in *2007 ASEE Annual Conference & Exposition*, 2007, Honolulu, Hawaii.
13. A. Charnes, W. W. Cooper, A. Y. Lewin and L. M. Seiford, *Data Envelopment Analysis: Theory, Methodology, and Applications*, Kluwer, Boston, (1994).
14. E. Thanassoulis, A comparison of regression analysis and data envelopment analysis as alternative methods for performance assessments, *The Journal of the Operational Research Society*, **44**(11), 1993, pp. 1129–1144.
15. R. D. Banker, R. F. Conrad and R. P. Strauss, A comparative application of data envelopment analysis and translog methods: An illustrative study of hospital production, *Management Science*, **32**(1), 1986, pp. 30–44.
16. W. F. Bowlin, A. Charnes, W. W. Cooper and H. D. Sherman, Data envelopment analysis and regression approaches to efficiency estimation and evaluation,*Annals of Operations Research*, **V2**(1), 1984, pp. 113–138.
17. J. Sarkis, *Ecoefficiency:* How data envelopment analysis can be used by managers and researchers, in *Environmentally Conscious Manufacturing*, Society of Photo-Optical Instrumentation Engineers, Boston, MA, (2000).
18. R. D. Banker, A. Charnes and W. W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science*, **30**(9), 1984, pp. 1078–1092.
19. S. Talluri, Data envelopment analysis: Models and extensions. *Decision Line*, **31**(3), 2000, pp. 8–11.
20. W. W. Cooper, L. M. Seiford and K. Tone, *Data Envelopment Analysis—A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Springer, (2000).
21. E. Kongar and T. Sobh, Are we accepting the right students to graduate engineering programs: easuring the success of accepted students via data envelopment analysis, in *2008 ASEE Annual Conference & Exposition*, Pittsburgh, PA, (2008).

Elif Kongar received her BS and MS degrees from the Industrial Engineering Department of Yildiz Technical University, Istanbul, Turkey. She was then awarded full scholarship for graduate studies in the USA. She obtained her Ph.D. degree in Industrial Engineering from the Department of Mechanical, Industrial and Manufacturing Engineering at Northeastern University, Boston, Massachusetts, USA. Dr. Kongar is currently an Assistant Professor at Bridgeport University and a Part-Time Researcher in the Center for Industrial Ecology at Yale University. Her research interests include the areas of supply chain management, logistics, environmentally conscious manufacturing, product recovery, disassembly systems, production planning and scheduling and multiple criteria decision making. She has co-authored several technical papers presented at various national and international conferences and published in their respective proceedings. She is a member of the Scientific Research Society, Sigma Xi, the Industrial Engineering Honor Society, Alpha Pi Mu, the Phi Beta Delta Honor Society and the Phi Kappa Phi Honor Society.

Tarek Sobh is the Vice President for Graduate Studies and Research and Dean of the School of Engineering at the University of Bridgeport. He is the Founding Director of the Interdisciplinary Robotics, Intelligent Sensing, and Control (RISC) laboratory and a Professor of Computer Engineering, Computer Science, Electrical Engineering and Mechanical Engineering. Professor Sobh's background is in the fields of computer science and engineering, control theory, robotics, automation, manufacturing, AI, computer vision and signal processing. He has been awarded many research grants to pursue his work in robotics, automation, manufacturing, and sensing. He is also an internationally known expert in the areas of online engineering, engineering education and e-learning. He has supervised over 120 graduate students working on different projects within his research areas. Dr. Sobh is active in consulting and serves on the board of several organizations and companies. He has provided services for industry and organizations in the U.S., Switzerland, India, Malaysia and Egypt. He has published over 200 journal and conference papers and book chapters, in addition to ten books in the areas of robotics, automation, systems engineering, simulation, computer science, engineering education, industrial electronics, instructional technology, assessment, and e-learning. Dr. Sobh has actively participated in, organized and chaired many international conferences, workshops and symposia, and has made many presentations, keynote speeches, invited talks, invited lectures and colloquia, seminars, and panel participations, at research meetings, University departments, research centers, and companies. Dr. Sobh's current research interests include self-reproducing autonomous electromechanical systems and mechatronics frameworks for developing and creating self-sustainable environments.

Mahesh Baral received his BS degree in Computer Engineering from Kathmandu University, Nepal in 2005. In August 2006, he started his MS degree in Technology Management (Advanced Database) at the School of Engineering, University of Bridgeport, Bridgeport, USA. He has been awarded a Graduate Assistantship by the same department since January 2007. He worked as a software engineer at Personal Broadband Nepal Pvt Ltd, Nepal for one year in 2005/2006, where he designed and developed database software projects. His area of expertise and interests include .NET-MsSQL and PHP-MySQL. Mahesh Baral has co-authored and presented various technical papers at several international conferences. He received the 'Honorable Mention' award for his project entitled 'Variation analysis and quality control chart development for 'determination of specific heat of a metal' experiment' at the ASEE Conference, University of Rhode Island, Kingston, RI, in April 2007.