Performance Evaluation of Academic Research Activity in a Greek University: A DEA Approach

Gregory Koronakos, Dimitris Sotiros, Dimitris K. Despotis and Dimitris Apostolou

Abstract In this paper we present a methodology developed for assessing the research performance of the academic staff, as part of the internal evaluation of the School of Information and Communication Technologies of a Greek university. We incorporate a qualitative aspect in the assessment by categorizing the research outcomes according to their quality. Also, value judgments over the evaluation criteria are elicited from the supporting committee and incorporated in a Data Envelopment Analysis assessment framework in the form of assurance region constraints.

Keywords Higher education · Academic research assessment · Quality of academic research · Data envelopment analysis (DEA) · Assurance region

1 Introduction

Academic research is considered as one of the most important activities of academic staff in higher education. The extent and quality of academic research are determinants for the academics' appointment and advancement. However, the quality is a controversial topic because of the existence of a large volume of publications in journals of low quality. As the research activity in a university is strictly designated

G. Koronakos (✉) · D. Sotiros · D.K. Despotis · D. Apostolou
Department of Informatics, University of Piraeus, 80 Karaoli and Dimitriou, 18534 Piraeus, Greece
e-mail: gkoron@unipi.gr

D. Sotiros
e-mail: dsotiros@unipi.gr

D.K. Despotis
e-mail: despotis@unipi.gr

D. Apostolou
e-mail: dapost@unipi.gr

R. Neves-Silva et al. (eds.), *Intelligent Decision Technologies*,
Smart Innovation, Systems and Technologies 39,
DOI 10.1007/978-3-319-19857-6_33

by the research activities of its staff members, the outcomes leverage its recognition and affect its position in international academic rankings (competitiveness). Moreover, there are countries where quality and performance issues play a crucial role in determining the funding that they receive from the government (e.g. in the UK and Australia). Therefore, the policy makers as well as the public draw significant attention to the results of the assessment of Higher Education Institutions (HEIs) and of their departments or faculties. Governments in many countries have already delivered policies with the aim to handle issues of accountability, cost control and enhancements of the quality of HEIs. In line with the above policies, in many countries periodical exercises are carried out by assessment bodies (committees). In the UK, for instance, the primary objective of Research Excellence Framework (REF) is the evaluation of the quality of research in publicly funded UK HEIs. It replaced the previous assessment system, last conducted in 2008, and named Research Assessment Exercise (RAE). In Australia, the Excellence in Research for Australia (ERA) initiative evaluates the quality of the research in Australian universities in order to provide advice to the Government on research matters and assist the National Competitive Grants Program (NCGP).

Beyond the aforementioned initiatives, complementary policies, such as internal assessments are often adopted in many institutions. For instance, the research development group at Helsinki School of Economics established a two-person team in order to assess the research performance and assist the administration to the allocation of the resources [1]. Recently, Greek higher education institutions have started developing diagnostic tools in order to better understand their comparative strengths and weaknesses. The goals are to provide incentives to the academics to improve the quality of their research and to enable the policy makers to allocate effectively the available resources. In [2], 20 Greek universities were evaluated by applying Data Envelopment Analysis (DEA) and econometric models. The inputs considered were: the number of academic and non-academic staff, the number of active registered students and the operating expenses, whereas the outputs considered were the number of graduates and the research income. A similar evaluation scheme was used in [3] in a department level of a Greek University. Further applications of DEA in education can be found in [4–11] among others.

In this paper we present a systematic framework developed for assessing the research performance of the academic staff, as part of the internal assessment of the ICT School in a Greek university. In this context, we focus on the efficiency assessment of 40 academic staff members undertaken by a committee of experts. Value judgments of the experts on the quality of the research outcome were incorporated in the assessment on the basis of the ERA 2010 journal classification system.

The rest of the paper unfolds as follows. Section 2 discusses the proposed methodology. Section 3 examines the research activity of the 40 academics. Finally conclusions are drawn.

2 Methods and Tools

The assessment exercise focuses on measuring the academics' research performance in the ICT School since their appointment. Therefore, the committee considered as inputs the years since the academic joined the institution and the total academic salary that s/he has received. As outputs were considered the publications in journals and the citations per year. The data of the academics are drawn from Scopus and their CVs.

Inputs

- Number of years in post.
- Total academic salary: total income received since appointment (euros).

Outputs

- Publications: number of single-author equivalent (SAE) journal papers published after the appointment.
- Citations per year: the index derives by dividing the total number of citations of journal papers by the number of years since the first publication. Self-citations are not taken into account.

In case of multiple co-authors (say N), each author is credited with the 1/N of the publication. Summing up over the publications of a particular academic, the single-author equivalent (SAE) is calculated as a measure of the quantity of his research outcome. Note here that the contribution of all the co-authors in the paper is assumed isomeric.

As long as quality matters, we made our assessments by taking into account the quality of the research records as well. First, we limited the research publications of each academic by taking into account only journal papers indexed in Scopus. Then we used the quality of the journal to classify the papers in five categories in accordance to the ERA2010 journal classification system, which ranks the journals in four tiers of quality rating (A+, A, B and C). Scopus papers not indexed by ERA2010 (unranked papers) formed a fifth category D. We note that the choice of the ERA journal classification system was made by the evaluators, which is an assumption in our assessment framework. However, other journal classification systems could be used instead. In our case though, because of the wide range of scope covered by the publications of the 40 academics of the ICT School under evaluation, a classification system that includes a wide list of journals was needed. Such a classification system is ERA2010, which comprises 20712 of a wide spectrum of scientific fields. To the best of our knowledge, there is not any other classification system so valid and thorough. For instance, the UK's Association of Business Schools (ABS) journal ranking includes a short list of journals relative to business and management science; as a result it does not meet our needs.

In many studies, the aspect of quality is incorporated in the assessment by using different weighting schemes to aggregate the different quality categories of publications. For instance in [6], the research quantity and quality were assessed by using weighted indexes of research publications. In a similar manner, in [1] weighted averages of outputs are obtained in order to shorten the outputs consisted of many factors and to account for the quality. In terms of this assessment though, the outputs are treated separately in order to obtain better insight of the quality of each academic's research activity.

2.1 Incorporating Preferences in Data Envelopment Analysis

As already mentioned, data envelopment analysis (DEA) has been widely used as an assessment framework in higher education. Indeed, a few years after the DEA was introduced [12], the technique was straightforwardly applied to the higher education sector. For instance, six DEA models used in [13] for the assessment of 20 English accounting departments. The remarks made in [11] about the attributes of higher education (absence of input and output prices, non-profit character and production of multiple outputs from multiple inputs) render it an attractive domain for DEA. In addition, DEA enables the decision makers to identify the best practice and provides methods for improvements.

For the assessment of the academic research we use the DEA model introduced in [14], under variable returns-to-scale (VRS) assumption. An output orientation is selected so as to determine the improvements for the inefficient academics by increasing the level of their outputs, given the levels of income and years in post. Assume n units, each using m inputs to produce s outputs. We denote by y_{rj} the level of the output r $(r = 1,\ldots,s)$ produced by unit j $(j = 1,\ldots,n)$ and by x_{ij} the level of the input i $(i = 1,\ldots,m)$ consumed by the unit j. The multiplier form of the output-oriented VRS DEA model for evaluating the relative efficiency of the unit j_0 is as follows:

$$
\begin{aligned}
&\min h_{j_0} = \sum_{i=1}^{m} v_i x_{ij_0} - w_0 \\
&s.t. \\
&\sum_{r=1}^{s} u_r y_{rj_0} = 1 \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + w_0 \leq 0, \quad j = 1,\ldots,n \\
&u_r \geq 0 \quad (r = 1,\ldots,s) \\
&v_i \geq 0 \quad (i = 1,\ldots,m) \\
&w_0 \in \Re \\
&u_r \in \Omega
\end{aligned}
\tag{1}
$$

Notice that in model (1), the free in sign variable w_o corresponds to the convexity constraint of the dual (envelopment) DEA model. Also, Ω generally denotes the set of restrictions imposed on the weights that limit the freedom of the evaluated unit in selecting its optimal weights [15, 16]. In our case, these restrictions reflect the aggregated opinions of the members of the evaluation committee with respect to the intensity of differentiation in quality among the five categories of journal publications. Pairwise comparisons matrices are used to assess the priorities of the five quality categories by each individual member of the evaluation committee. The comparison scheme and scale is as in the analytic hierarchy process (AHP) c.f. [17]. Then the individual priorities are used to obtain assurance region constraints that restrain the weights given to the papers of different categories. A modified version of method proposed in [18] is used to translate the individual priorities obtained by each evaluator into assurance region constraints that define a weight space reflecting the committee as a whole. For every pair of factors (O_i, O_j), the ratio of their weights (u_i/u_j) is bounded as follows:

$$L_{ij} \leq u_i/u_j \leq U_{ij} \tag{2}$$

where the bounds are defined by the priorities w_{ki} and w_{kj} that express the priority given by the k^{th} evaluator to the i^{th} and the j^{th} factor as follows:

$$L_{ij} = \min_k \frac{w_{ki}}{w_{kj}},\ U_{ij} = \max_k \frac{w_{ki}}{w_{kj}} \tag{3}$$

It is clear from (3) that the lower and upper bounds L_{ij} and U_{ij} are defined by the extreme judgments (priorities). Thus, the preferences' central tendency is not captured adequately. We treat this issue by excluding, for each pair of factors, the ratios of priorities, which are deemed as outliers. We provide the method in detail in the next section.

3 Assessment of Academic Research Performance

In this section, we provide the results obtained from the internal assessment of the faculty members of the ICT School. Descriptive statistics for the data set (40 faculty members) are presented in the Table 1:

Table 1 Descriptive statistics of the data set

	Inputs		Outputs					
	Y*	TS*	A+*	A*	B*	C*	D*	Citations*
Max	28.50	105.68	1.33	6.42	7.42	9.20	4.26	71.75
Min	3.00	6.42	0	0	0	0	0	0.41
Average	11.73	35.55	0.14	1.07	1.57	1.73	0.79	14.94
St. Dev.	6.90	25.81	0.31	1.72	1.69	2.04	1.06	13.63
Variance	47.61	666.38	0.10	2.94	2.85	4.18	1.12	185.80

* Y: Years in post; TS: Total salary since appointment (10,000 euros); A+, A, B, C: Number of single author equivalent (SAE) papers in class A+, A, B, C journals (ERA2010); D: Number of single author equivalent (SAE) papers in unranked journals; Citations: Citations per year

The five evaluators, members of the evaluation committee, provided independently their opinion with respect to the five quality categories of journals, indicating how much one category excels or is exceled by another in terms of quality. The evaluation scheme employed was the pairwise comparisons matrix used in the context of AHP [17]. Judgments were made on the [1–9] scale proposed in [17] whereas the eigenvector method was used to calculate the priorities. The priorities are the principal right eigenvector i.e. the eigenvector corresponding to the largest eigenvalues of the reciprocal response matrix. Tables 2 and 3 exhibit the pairwise comparisons of two of the evaluators (E4 and E5).

Table 2 Pairwise comparisons matrix of evaluator E4

Categories	A+	A	B	C	D
A+	1	2	4	9	9
A	0.5	1	4	9	9
B	0.25	0.25	1	4	4
C	0.11	0.11	0.25	1	1
D	0.11	0.11	0.25	1	1

The evaluator E4 identifies three main categories (A+, A), B and (C, D) with a slight superiority of class A+ over A, whereas the evaluator E5 is more strict as he identifies only two main categories (A+, A) and (B, C, D) with a slight difference between class A+ and class A.

Table 3 Pairwise comparisons matrix of evaluator E5

Categories	A+	A	B	C	D
A+	1	3	9	9	9
A	0.333	1	7	7	7
B	0.111	0.143	1	1	1
C	0.111	0.143	1	1	1
D	0.111	0.143	1	1	1

The priorities (weights) obtained for the five categories from the pairwise comparison matrices of the evaluators E1–E5 are summarized row-wise in Table 4 and are exhibited in Fig. 1. The last column of Table 4 portrays the Consistency Ratio (CR) of each evaluator's comparison matrix. As all CRs are below 10 % the priorities are acceptable.

Table 4 Evaluators' priorities for the five journal categories

Evaluator	A+	A	B	C	D	CR
E1	0.434	0.306	0.162	0.060	0.037	0.030
E2	0.503	0.299	0.107	0.056	0.035	0.021
E3	0.570	0.250	0.076	0.076	0.027	0.1
E4	0.456	0.344	0.123	0.038	0.038	0.025
E5	0.550	0.301	0.050	0.050	0.050	0.020

As shown in Table 4 and Fig. 1, for the priorities of all the evaluators holds that $w^{A+} > w^{A} > w^{B} \geq w^{C} \geq w^{D}$. However, the intense of preference among the different journal categories is not the same. The evaluators E1 and E2 rank order the five categories with no ties (strict order), whereas E3, E4 and E5 differentiate by assuming ties for some categories.

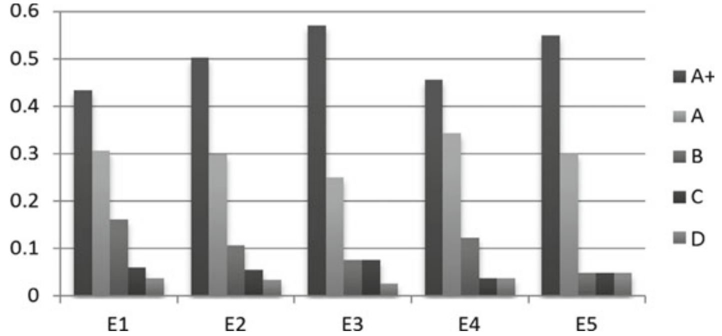

Fig. 1 Weighting preferences of five evaluators

Next we eliminate the outlier ratios of priorities across the evaluators to avoid the effect of extreme judgments. As shown in the boxplot of Fig. 2, the ratio For example, the ratio w^A/w^C of the priorities given from the evaluator E4 for class A over class C is detected as outlier and, thus, omitted before calculating the upper bound as in (3).

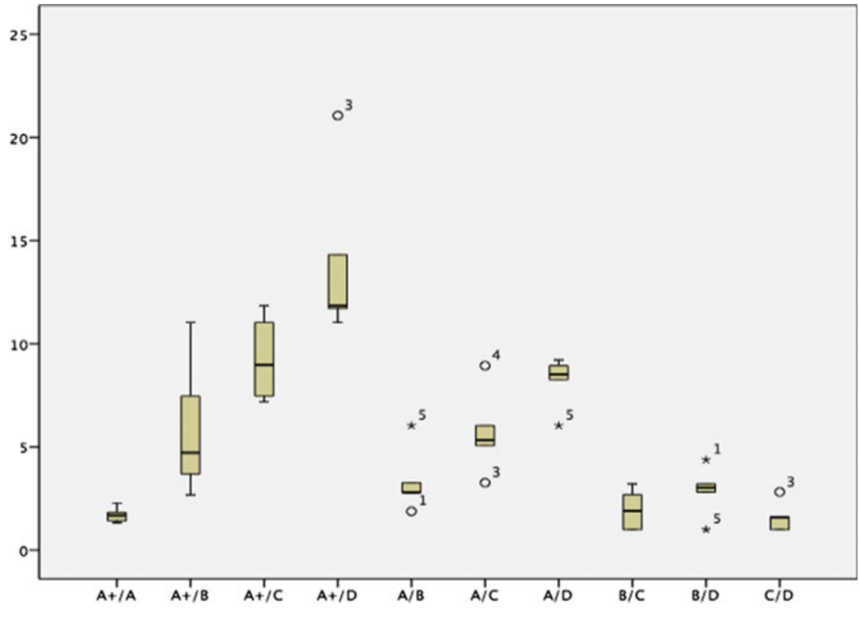

Fig. 2 Outlier detection

Table 5 contains the calculated lower and upper bounds according to Eqs. (2) and (3) after removing the upper outliers.

Table 5 Lower and upper bounds

	$\frac{u^{A+}}{u^{\mathrm{A}}}$	$\frac{u^{A+}}{u^{\mathrm{B}}}$	$\frac{u^{A+}}{u^{C}}$	$\frac{u^{A+}}{u^{Unr.}}$	$\frac{u^{A}}{u^{B}}$	$\frac{u^{A}}{u^{C}}$	$\frac{u^{A}}{u^{D}}$	$\frac{u^{B}}{u^{C}}$	$\frac{u^{B}}{u^{D}}$	$\frac{u^{C}}{u^{D}}$
LB	1.325	2.674	7.188	11.034	1.885	3.268	6.036	1	1	1
UB	2.284	11.034	11.847	14.315	3.268	6.036	9.220	3.209	3.209	1.629

To highlight how the incorporation of the quality dimension into the assessments affects the performance results, we run two models (scenarios). In the first one we consider the total SAE papers, regardless the journal class they are published, as one aggregate output. In the second scenario we consider five distinct outputs for the publications, one for each category of journals and we introduce assurance region constraints in the assessment model by restricting the ratios of the output weights between the lower and the upper bounds given in Table 5.

In the first scenario, 6 academics (namely A4, A18, A21, A22, A24 and A39) were deemed efficient. In the second one, only 4 of them (namely A4, A22, A24 and A39) maintained their 100 % efficiency score. For the 75 % of the academics the efficiency scores in the second scenario are lower than in the first one. This is attributed to the high number of publications in low and medium quality journals. For instance, A21 is efficient in the first scenario as he has the highest total number of publications. However, when the quality of the journals is taken into account in the second scenario the same academic becomes inefficient. The low efficiency score of A21 in the scenario 2 is justified by the low number of publications in journals ranked as A+ and A (cf. Table 6), which are of the greatest importance according to the committee. On the contrary, A3 has most of his papers published in class A journals but his total number of publications is relatively low (cf. Table 6). That is why A3 appears more efficient in the second scenario (160 %) than in the first one (211 %). Notice here that, due to the output orientation assumed in the assessments, the lower is the figure the more efficient is the academic. Similarly, comparing the research profiles of the academics A2 and A3 (cf. Table 6) justifies the differentiation of their efficiency scores in the two scenarios, although they have the same years in post.

Table 6 Data and efficiency scores of three academics

Ac.	Eff. 1	Eff. 2	Y	TS	A+	A	B	C	D	Citations	Total Publ.
A2	139 %	381 %	24	84.91	0	0	4	7.08	4	2.11	15.08
A3	211 %	160 %	24	76.73	0	6.25	0.83	1.67	0.5	15.4	9.25
A21	100 %	132 %	12.5	37.54	0	1.98	5.47	9.2	4.26	24.59	20.9

Figure 3 depicts the top 3 layers of efficient academics. The first layer comprises the efficient academics discussed above. The second layer is obtained by rerunning the scenario 2 after excluding the academics of the first layer. The third layer is obtained similarly by excluding the academics of the first two layers.

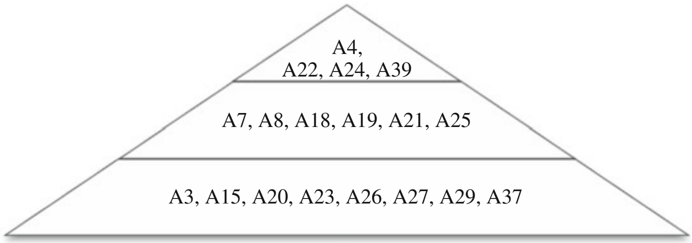

Fig. 3 The top 3 layers of efficient academics

The dual solution of the model (1) provides the necessary information to project the inefficient units on the efficient frontier. The profiles of the efficient academics in the second scenario are presented in the Table 7.

Table 7 Profile of the efficient academics in scenario 2

Ac.	Y	TS	A+	A	B	C	D	Citations
A4	16	54.749	1	4.417	3.450	7.617	1.417	41.348
A22	3	6.422	0	0.333	0	0	0	7.6
A24	7	15.607	0	0	1.450	0.500	0	71.75
A39	3.5	14.419	0	0.750	0.250	0	0	15.938

Table 8 exhibits the profile of three inefficient academics and their potential improvements so as to be rendered efficient. Specifically, it provides the levels of improvement that the academics should exhibit on each output so as to be more competitive.

Table 8 Profile and targets of three inefficient academics

		Original values						Target values					
	Eff.	A+	A	B	C	D	Cit.	A+	A	B	C	D	Cit.
A6	313 %	0.2	0.75	0.95	0.7	0.25	11	0.63	2.35	2.98	2.19	0.78	34.47
A20	199 %	0.5	0	0.8	0.33	1.67	15.8	0.99	0	1.59	0.66	3.32	31.43
A27	218 %	0	0.25	0.33	2.08	0.33	19.08	0	0.54	0.73	4.53	0.73	41.53

4 Conclusions

We evaluated the performance of research activity of 40 academics in the context of the internal assessment of a Greek ICT School. We also discussed about the crucial role of categorizing the publications according to their quality. Our approach utilizes an AHP-like method to elicit the experts' preferences with respect to the relative importance of evaluation criteria and integrates it with DEA. We also illustrated that the incorporation of expert judgments as weight restrictions in DEA yields efficiency scores that more accurately reflect the common academic practice of giving credit to research outcome of higher quality.

Acknowledgements This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program 'Education and Lifelong Learning' of the National Strategic Reference Framework (NSRF) – Research Funding Program: THALES. Investing in knowledge society through the European Social Fund.

References

1. Korhonen, P., Tainio, R., Wallenius, J.: Value efficiency analysis of academic research. Eur. J. Oper. Res. **130**, 121–132 (2001)
2. Katharaki, M., Katharakis, M.G.: A comparative assessment of Greek Universities' efficiency using quantitative analysis. Int. J. Ed. R. **49**, 115–128 (2010)
3. Kounetas, K., Anastasiou, A., Mitropoulos, P., Mitropoulos, I.: Departmental efficiency differences within a Greek University: an application of a DEA and Tobit analysis. Int. T. Oper. Res. **18**, 545–559 (2011)
4. Beasley, J.E.: Comparing university departments. Omega **18**(2), 171–183 (1990)
5. Beasley, J.E.: Determining teaching and research efficiencies. J. Oper. Res. Soc. **46**(4), 441–452 (1995)
6. Athanassopoulos, A.D., Shale, E.: Assessing the comparative efficiency of higher education institutions in the UK by the means of data envelopment analysis. Educ. Econ. **5**(2), 117–134 (1997)
7. Johnes, J., Johnes, G.: Research funding and performance in U.K. University departments of economics: a frontier analysis. Econ. Educ. Rev. **14**(3), 301–314 (1995)
8. Avkiran, N.K.: Investigating technical and scale efficiencies of Australian Universities through data envelopment analysis. Socio. Econ. Plan. Sci. **35**(1), 57–80 (2001)
9. Abbott, M., Doucouliagos, C.: The efficiency of Australian Universities: a data envelopment analysis. Econ. Educ. Rev. **22**(1), 89–97 (2003)
10. Kao, C., Hung, H.-T.: Efficiency analysis of university departments: an empirical study. Omega **36**, 653–664 (2008)
11. Johnes, J.: Data envelopment analysis and its application to the measurement of efficiency in higher education. Econ. Educ. Rev. **25**(3), 273–288 (2006)
12. Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. Eur. J. Oper. Res. **2**, 429–444 (1978)
13. Tomkins, C., Green, R.: An experiment in the use of data envelopment analysis for evaluating the efficiency of UK University departments of accounting. Financ. Acc. Man. **4**(2), 147–164 (1988)
14. Banker, R.D., Charnes, A., Cooper, W.W.: Some models for estimating technical and scale inefficiencies in DEA. Manage. Sci. **32**, 1078–1092 (1984)

15. Charnes, A., Cooper, W.W., Huang, Z.M., Sun, D.B.: Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. J. Econometrics **46**(1–2), 73–91 (1990)
16. Thompson, R.G., Langemeier, L.N., Lee, C., Lee, E., Thrall, R.M.: The role of multiplier bounds in efficiency analysis with application to Kansas firms. J. Econometrics **46**(1–2), 93–108 (1990)
17. Saaty, T.L.: A scaling method for priorities in hierarchical structures. J. Math. Psychol. **15**(3), 234–281 (1977)
18. Takamura, Y., Tone, K.: A comparative site evaluation study for relocating Japanese government agencies out of Tokyo. Socio. Econ. Plan. Sci. **37**(2), 85–102 (2003)