

Evaluation of research activity in higher education: A Data Envelopment Analysis approach

Yannis G. Smirlis*
Department of
Informatics, University
of Piraeus
80, Karaoli and
Dimitriou, 18534
Piraeus, Greece

Dimitris Sotiros
Department of
Informatics, University
of Piraeus
80, Karaoli and
Dimitriou, 18534
Piraeus, Greece

Dimitris K. Despotis
Department of
Informatics, University
of Piraeus
80, Karaoli and
Dimitriou, 18534
Piraeus, Greece

Gregory Koronakos
Department of Informatics,
University of Piraeus
80, Karaoli and
Dimitriou, 18534
Piraeus, Greece

Keywords: Data Envelopment Analysis (DEA); piece-wise linear Data Envelopment Analysis (PL-DEA); research activity; higher education

Abstract

Academic research is considered as one of the most important activities of academic staff in Higher Education. The extent and the quality of the research records are determinants of the academic profiles and play, or they should play, a key role in the advancement of the academic staff members. As the research activity of a University department is strictly determined by the research activity of its members, quality individual research outcomes improve the recognition of the University department and affect its position in international academic rankings.

In this paper we develop a Data Envelopment Analysis (DEA) assessment framework to evaluate the research activity of individuals in comparable University departments. The selected factors (inputs and outputs) have a meaningful interpretation in the analysis and, moreover, provide us the ability to perform the assessments by taking into account both the extent as well as the quality of the research records. We take as inputs the duration of the research activity, the funds received (in monetary values) and the teaching load (adjusted average). We take as outputs the number of publications in journals ranked as A⁺ or A, the number of publications in journals ranked as B or C, the publications in unranked journals, publications in conferences, the number of research programs in which the researcher participates and the number of citations (excluding self-citations). We draw the journal rankings from the Excellence in Research for Australia (ERA) 2010 journal classification system. The data are drawn from Scopus, Google Scholar, university personal records and CVs. To facilitate the incorporation of a quality aspect in our assessments, we use a piece-wise linear variant of the DEA model with assurance region constraints. For example, assuming convex value functions for the publications in highly ranked journals and concave value functions for the publications in unranked journals, we reward the quality research records while diminishing the contribution of extensive publications in non-quality journals in the overall research performance. We illustrate our assessment approach with anonymous data and we provide a meaningful interpretation of the results.

Evaluation of research activity in higher education: A Data Envelopment Analysis approach

Yiannis G. Smirlis, Dimitris Sotiros, Dimitris K. Despotis, Gregory Koronakos

University of Piraeus, Department of Informatics
80, Karaoli and Dimitriou, 18534
Piraeus, Greece

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.

Layout of presentation

- Introduction
- Data Envelopment Analysis (DEA)
- Application
- Conclusions

Introduction

- Research is an important activity of the members of academic staff in higher education.
- The extent and the quality of their research records
 - determine their profile and play, or they should play, a key role in their advancement
 - improves the recognition of the University departments and affects their position in international academic rankings



Introduction

- We propose a new assessment framework that
 - encounters additional factors : duration of the research, amount of received funds, quality of publications, conference participation etc.
 - uses standard DEA methodology for the comparative assessment
 - facilitates the incorporation of a quality aspect in our assessments by using a piece-wise linear variant of the DEA model

Data Envelopment Analysis

- Data envelopment analysis (DEA) is the leading technique for assessing the efficiency of decision making units (DMU), when multiple inputs and outputs are to be taken into account.
- It is based on Linear Programming
- Units under evaluation are set free to estimate the weights to their best

The DEA model*

* Variable Returns to Scale (VRS) -output oriented

x_{ij} : value of researcher j on input i

y_{rj} : value of researcher j on output r

The weights

$v_1, \dots, v_i, \dots, v_m \quad u_1, \dots, u_r, \dots, u_s$

are the variables under estimation

The model is executed n times, one for each researcher.

$h_{j_0}^*$ is the optimal value of the objective function obtained by the assessment of researcher j_0

Efficient are the researchers that achieve score 1

Non-efficient researchers get score > 1

$$\min h_{j_0} = \sum_{i=1}^m v_i x_{ij_0} - w_0$$

s.t.

$$\sum_{r=1}^s u_r y_{rj_0} = 1$$

$$\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s u_r y_{rj} - w_0 \geq 0, \quad j = 1, \dots, n$$

$$u_r, v_i \geq \varepsilon \quad \forall r, i$$

w_0 free in sign

$$h_{j_0}^* = 1$$

$$h_{j_0}^* > 1$$

Linearity in virtual inputs / outputs

$$\min h_{j_0} = \sum_{i=1}^m v_i x_{ij_0} - w_0$$

s.t.

$$\sum_{r=1}^s u_r y_{rj_0} = 1$$

$$\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s u_r y_{rj} - w_0 \geq 0, \quad j = 1, \dots, n$$

$$u_r, v_i \geq \varepsilon \quad \forall r, i$$

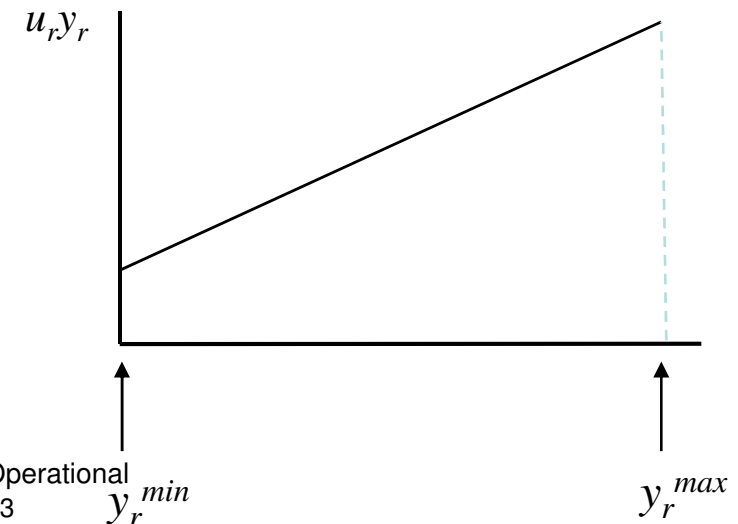
w_0 free in sign

The summations

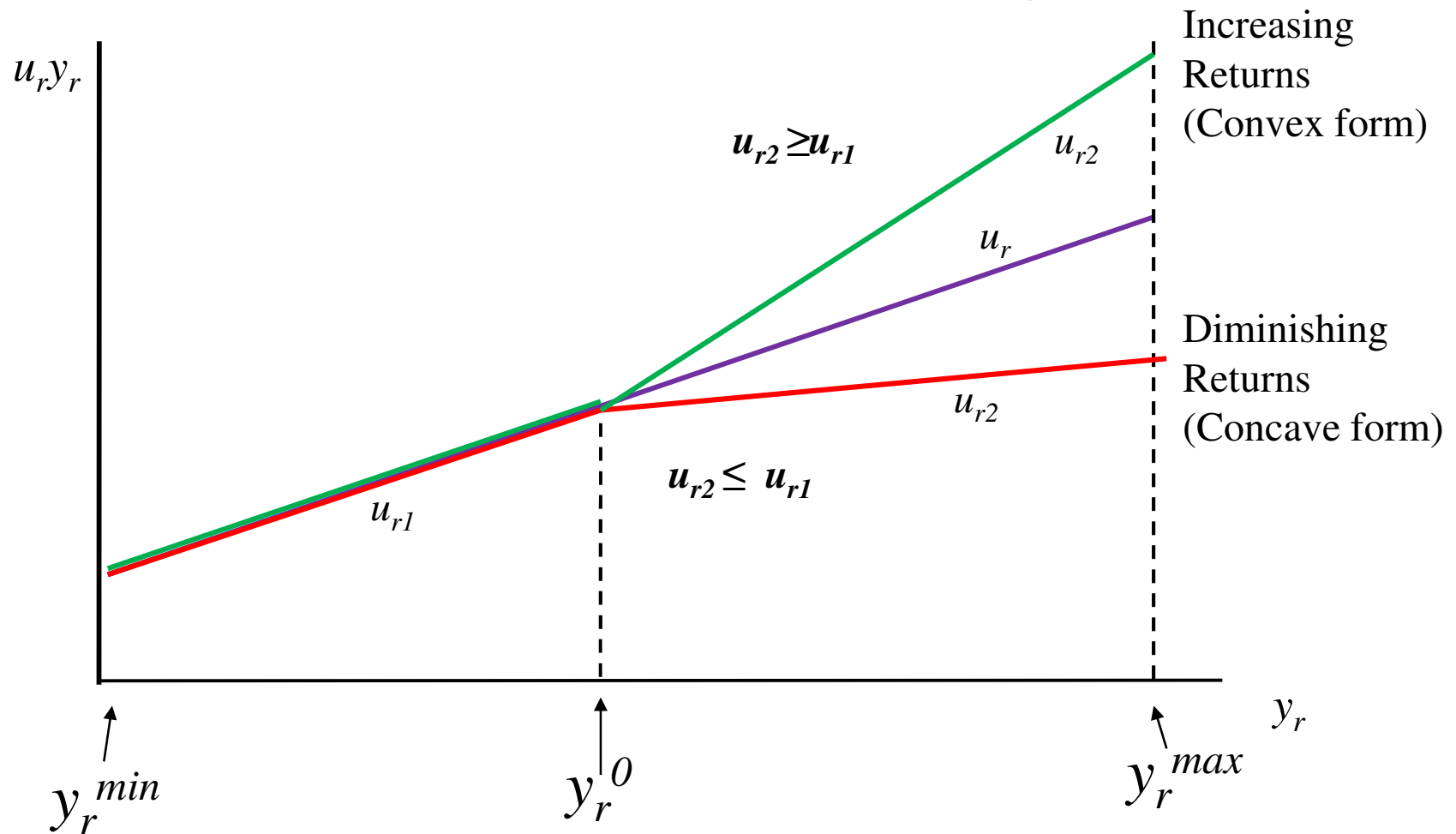
$$\sum_{r=1}^s u_r y_{rj}, \quad \sum_{i=1}^m v_i x_{ij}$$

represent the total virtual output and input respectively for unit j

and are **value functions of the weights and have linear increasing shape**

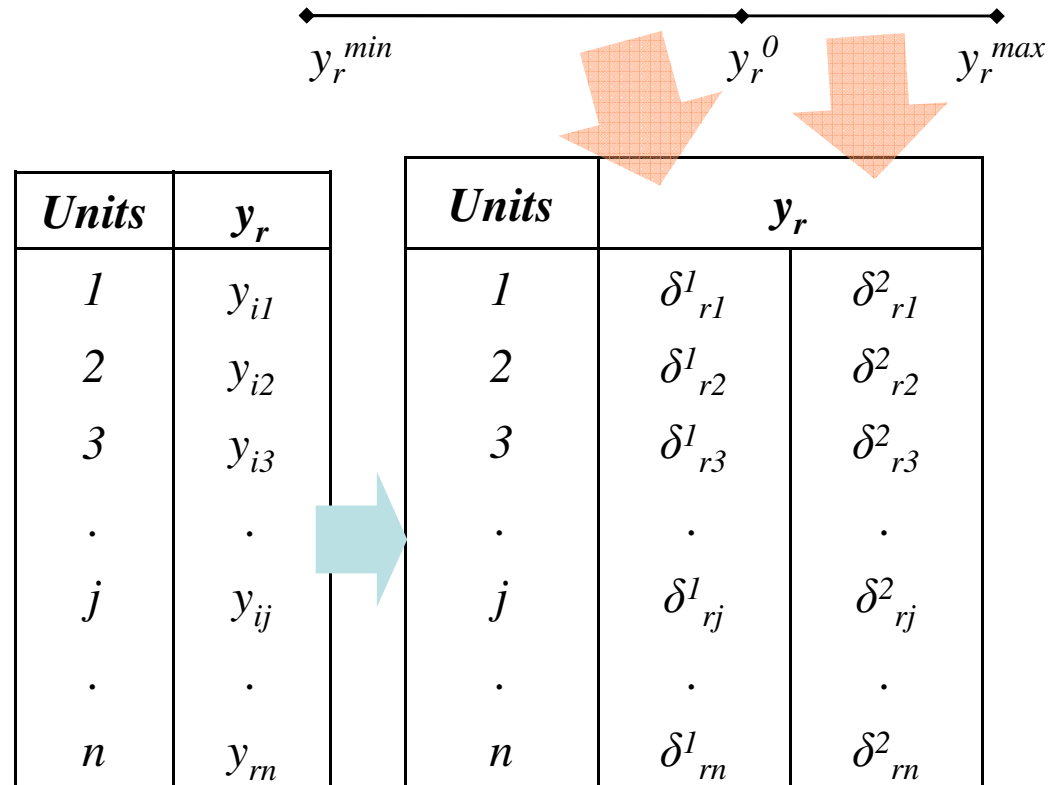


Piecewise Linear virtual outputs



Threshold value to distinguish
the non-desirable / highly
desirable output

Decomposition of factors into parts (outputs)

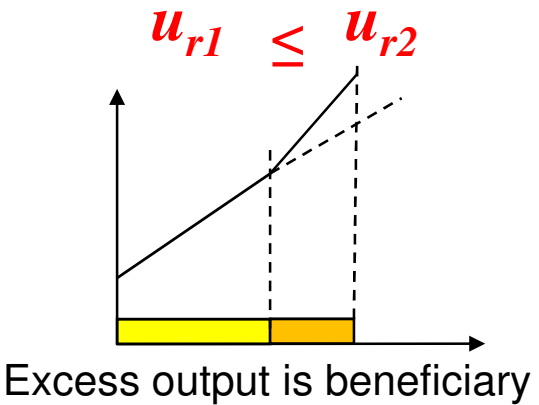
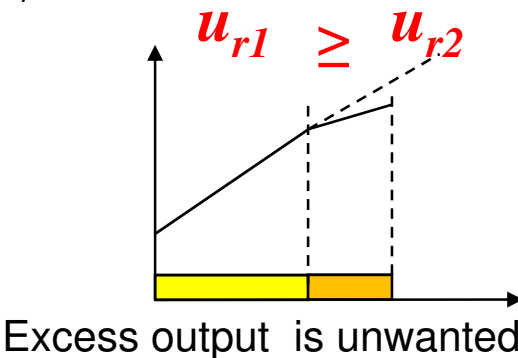
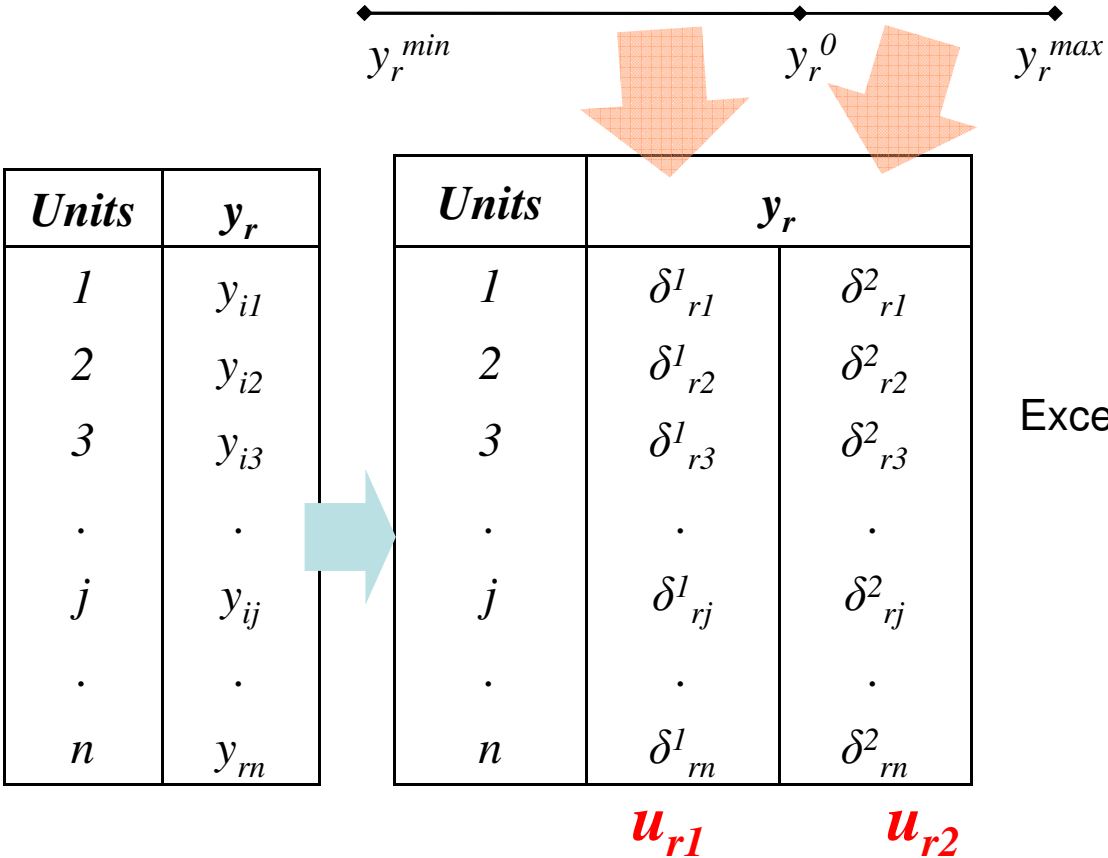


The r -th output value $y_{ri} \in [y_r^{\min}, y_r^{\max}]$ of any unit j is decomposed in two parts

$$y_{rj} = \delta_{rj}^1 + \delta_{rj}^2$$

$$\delta_{rj}^1 = \begin{cases} y_{rj} & \text{if } y_{rj} \in [y_r^{\min}, y_r^o] \\ y_r^o & \text{if } y_{rj} \in (y_r^o, y_r^{\max}] \end{cases}, \quad \delta_{rj}^2 = \begin{cases} 0 & \text{if } y_{rj} \in [y_r^{\min}, y_r^o] \\ y_{rj} - y_r^o & \text{if } y_{rj} \in (y_r^o, y_r^{\max}] \end{cases}$$

Decomposition of factors into parts (outputs)



The piecewise linear model with one breakpoint for inputs and outputs

$$\min h_{j_0} = U(Y_{j_0}) - w_0$$

s.t.

$$U(X_{j_0}) = 1$$

$$U(X_j) - U(Y_j) - w_0 \geq 0 \quad j = 1, \dots, n$$

$$U(Y_j) = \sum_{r=1}^d u_r y_{rj} + \sum_{r=d+1}^s (u_{r1} \delta_{rj}^1 + u_{r2} \delta_{rj}^2)$$

$$U(X_j) = \sum_{i=1}^t v_i x_{ij} + \sum_{i=t+1}^m (v_{i1} \gamma_{ij}^1 + v_{i2} \gamma_{ij}^2)$$

$$u_r, v_i \geq \varepsilon \quad r = 1, \dots, d, i = 1, \dots, t$$

$$u_{r1}, u_{r2}, v_{i1}, v_{i2} \geq \varepsilon \quad r = d+1, \dots, s, i = t+1, \dots, m$$

$$u_{r1}, u_{r2}, v_{i1}, v_{i2} \in \Omega$$

Ω = set of weight restrictions
to express increasing /
diminishing returns

Application

- Use of an anonymous data set to evaluate research activity for 112 faculty members in Business Administration and Economics Departments of Greek Universities
- The data are collected from Scopus, Google Scholar, university staff records and CVs.

The factors for the assessment

Inputs	Outputs
x1. Duration of the research activity (years past since first publication)	y1. Number of publications in high ranked journals (A+ or A) *
x2. Funds received from research projects (in €)	y2. Number of publications in medium ranked journals (B or C) *
x3. Available time for research (estimated working hours excluding teaching load and administrative work)	y3. Number of publications in unranked journals*
	y4. Number of publications in conferences
	y5. Number of research programs
	y6. Number of citations (excluding self-citations)

* *The journal rankings are drawn from the Excellence in Research for Australia (ERA) 2010 journal classification system –www.arc.gov.au*

Descriptives

Variable	Mean	StDev	Minimum	Median	Maximum
<u>INPUTS</u>					
Duration of research act	17,777	6,723	5	17	30
Funds (in €)	122215	109901	1000	85421	410334
Avail. Time	5045	1467	1920	4992	7680
<u>OUTPUTS</u>					
A+, A	6,063	4,527	0	5	19
B, C	12	7,2	0	11	29
UNRANKED	15,161	10,434	0	14	47
CONFERENCES	35,81	20,43	4	33,5	104
RESEARCH PROGRAMS	6,321	4,515	1	5	15
CITATIONS	52,81	72,66	2	23	350

Intra-variable weight restrictions

Outputs

$$U(A^+, A) \geq 2 * U(B, C)$$

$$U(A^+, A) \geq U(\text{Citations})$$

$$U(B, C) \geq 3 * U(\text{Unranked})$$

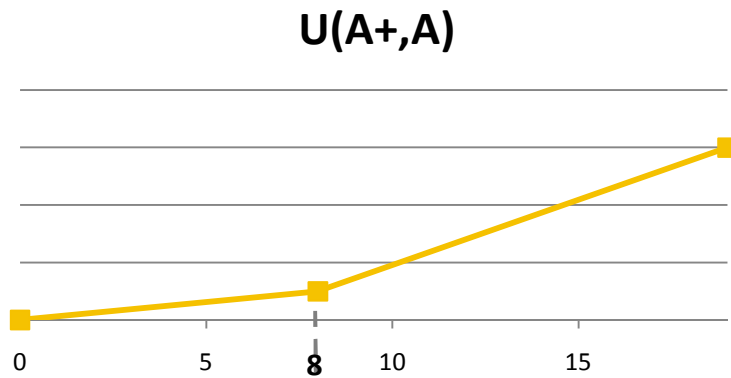
$$U(B, C) \geq 3 * U(\text{Conferences})$$

Inputs

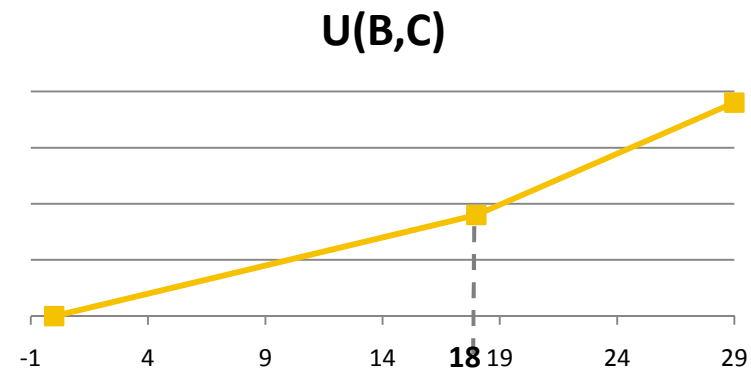
$$V(\text{Duration}) \geq 3 * V(\text{Available Time})$$

$$V(\text{Duration}) \geq 3 * V(\text{Funds})$$

Inter-variable weight restrictions (1)



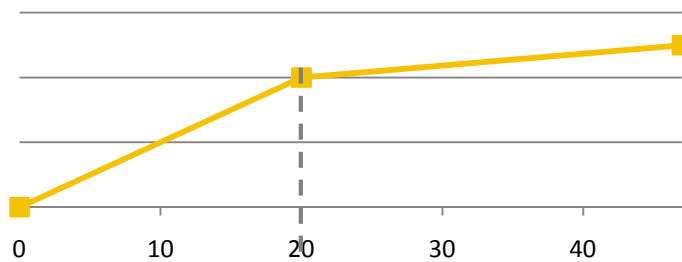
$$u_2^{A+,A} \geq 2 * u_1^{A+,A}$$



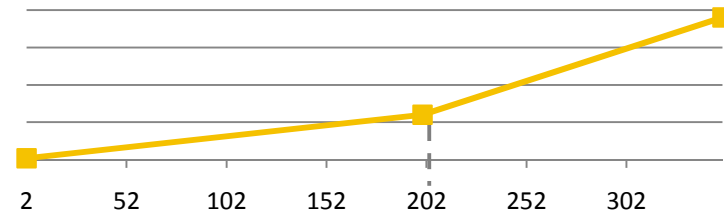
$$u_2^{B,C} \geq u_1^{B,C}$$

Inter-variable weight restrictions (2)

U(unranked)



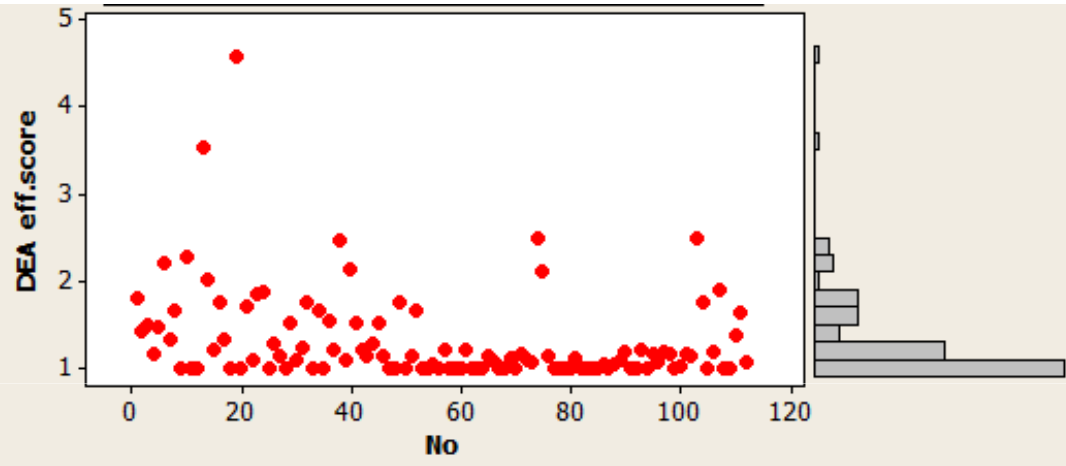
U(citations)



$$u_2^{Unranked} \leq \frac{u_1^{Unranked}}{2}$$

$$u_2^{Citations} \geq 2 * u_1^{Citations}$$

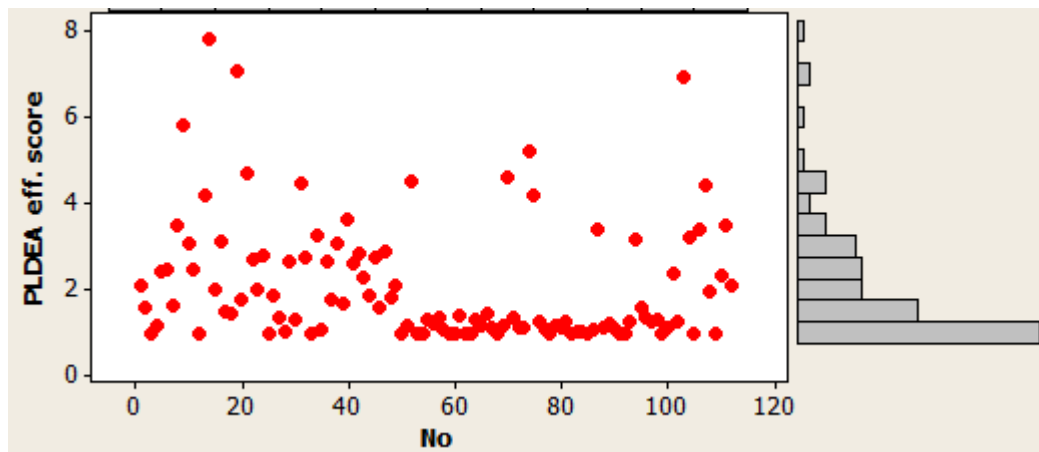
Results



DEA

Number of efficient researchers = 29

Average score = 1,323



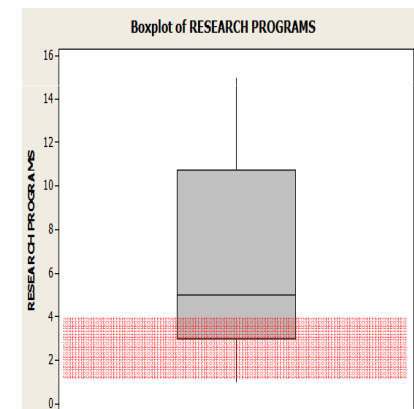
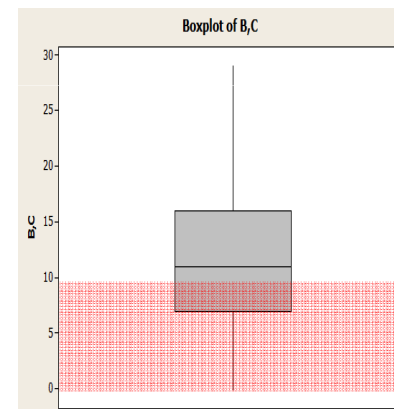
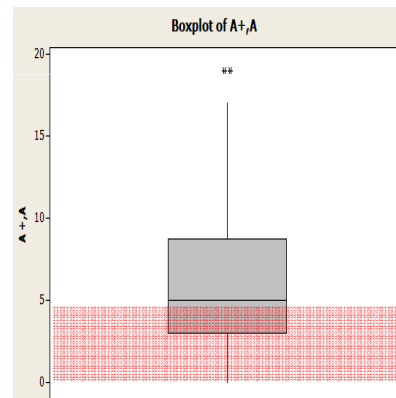
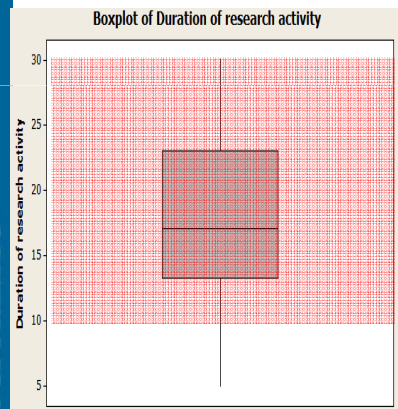
Piecewise linear DEA

Number of efficient researchers = 19

Average score = 2,091

The effect on low performance researchers

- We focus on a group of 25 researchers with relative low performance :
 - Duration of research activity 10+ years
 - Number of A+, A journals [0... 5]
 - Number of B, C journals [0... 9]
 - Number of research programs [0.. 4]



DEA

Number of efficient researchers = 3
Average score = 1,759

Piecewise linear DEA

Number of efficient researchers = 0
Average score = 3,656

Impact on selected cases

Evaluation factors	Case #1	Case #2
Duration of research	27 years	30 years
Funds	2000€	155976€
Publications in A+, A ranked journals	0	17
Publications in B, C ranked journals	2	23
Unranked journals	47	29
Conference presentations	34	104
Research programs	1	14
Citations	5	350
RESULT	DEA score=1 PLDEA score=5,802	DEA score=1 PLDEA score=1

Conclusion

- Our approach facilitates the incorporation of a quality aspect in the assessment :
 - it penalizes the extensive publications in non-quality journals and the time available for research
 - It rewards the quality research records giving higher importance to publications in highly ranked journals

Thank you for your attention...