Building a Global Performance Indicator to Evaluate Academic Activity Using Fuzzy Measures

Marta Cardin[1], Marco Corazza[1,3], Stefania Funari[2], and Silvio Giove[1]

[1] Department of Economics, Ca' Foscari University of Venice
[2] Department of Management, Ca' Foscari University of Venice
[3] Advanced School of Economics of Venice
Cannaregio 873, 30121 Venice, Italy
{mcardin,corazza,funari,sgiove}@unive.it

Abstract. The aim of this note is to provide a global performance index that allows to evaluate the performance of each faculty member and which is able to consider the multidimensional nature of the academic activity in terms of research, teaching and other activities that academics should/might exercise. In order to model also the case in which there could be synergic and redundant connections among the different areas of the academic activity, we propose to use fuzzy measures and the Choquet integral as an aggregator of the different components.

Keywords: Fuzzy measures, Choquet integral, performance evaluation.

≪*Work. Finish. Publish.*≫
Michael Faraday, 22 September 1791 – 25 August 1867

1 Introduction

The evaluation of the performance of the academic activity has gained a growing interest in recent years and many Universities undertake periodic assessments of their faculty members which act as a first step for a wider performance evaluation process involving, at different levels, departments, faculties and Universities.

Regarding in particular the research activity, a vast literature is interested in the measurement of the research output of scholars. Several approaches have indeed been proposed in order to assess the scientific production.

We remind the use of bibliometric indicators, the well-known h-index proposed by Hirsh [6] and other indices that complement and extend it, such as the g-index, the m-index, the $h^{(2)}$-index, the h_α-index, the A-index, the R-index, the normalized h-index, the dynamic h-type index and many others indicators related to the h-index (see [1] for a review). This great variety of research output measures, even if with some differences, intends to capture both the quantitative and the qualitative dimension of the research, summarized in the number of publications and the number of citations that each publication receives, by focusing in many cases on *the most productive core* of a researcher ([3]).

B. Apolloni et al. (Eds.): *Neural Nets and Surroundings*, SIST 19, pp. 217–225.
DOI: 10.1007/978-3-642-35467-0_23

The bibliometric indices allow to obtain a ranking of the authors; however, having each indicator some advantages but also some drawbacks, none of them can be considered the best measure of the research outputs. Without emphasizing a particular advantage or drawback, Marchant [8] proposed a general theory of bibliometric ranking, which explores the properties of the scoring rules that permit to rank the authors by taking into account all their publications, although weighted by some partial scores.

There is therefore a vast debate on the performance measurement of the research production; on the other hand, turning towards the teaching area, we may note that the discussion on the assessment of the teaching activity is as much active.

At present, many university departments have activated some teacher evaluation systems, but also in this case there is no a uniform system of indicators fully recognized. A reasonable solution is to consider not only the classroom teaching and thus the student ratings as a measure of teaching effectiveness, rather to take account of all educational activities of a faculty member, including the development of new courses, the restyle of the old ones, the use of innovative teaching strategies, the offer of seminars, workshops, and so on. Therefore, experts in the field of teacher evaluation tend to prefer the combined use of multiple approaches to evaluate faculty teaching (see for instance [7], [5]).

Although both research and teaching are important components of the university production, the literature on the university performance mainly focused separately either on the evaluation of the research dimension or on the teaching component, obtaining a partial view of the goodness of the academic activity. In order to provide an overall assessment of the performance of each faculty member, so that to capture the multidimensional nature of the evaluation problem, it should be more desirable to simultaneously consider the different faces of the academic activity, shared mainly among research, teaching and service.

Based on this thought, Bana e Costa and Oliveira ([2]) have recently proposed a faculty evaluation model that takes into consideration the whole range of academic activities pursued by each faculty member. They used a multicriteria decision analysis based on value functions, in order to associate to each faculty member an overall score, according to which each member of the faculty can be finally assigned to one of the four rating categories (inadequate, sufficient, relevant and excellent). The authors used value functions which transform individual performance into value for the university and used a hierarchical aggregation procedure which involves first an aggregation of the criteria (in terms of value) within each area of activity and then an aggregation of values across areas, leading finally to an overall score for each faculty member.

However, when different evaluation criteria within each area are simultaneously considered, some synergic effects may arise. For example, referring to the research area, Tucci, Fontani and Ferrini ([12]) noted that the joined activity of the researcher in terms of publishing and of accomplishment of other activities (as for example organizing meetings and conferences, attending to seminars, coordinating research groups, being editor and being referee), should be awarded in the computation of the R-Factor index that measures the research output. The idea is that the involvement in several activities of the researcher requires an additional effort that should be considered in the researcher's performance evaluation.

A similar synergic behaviour can also occur among the different areas of the academic activity. For example, why not reward (in terms of performance score) an academic that performs well both in research and teaching or even in all the areas considered? In this case a sort of premium on performance could be considered in addition to the partial performance evaluation scores, which concurs to increase the final overall performance score. Otherwise, academics working on management activity but who do not directly contribute to publishing activity, are penalized.

In decision making problems, a natural approach that allows to model the interaction among criteria is the use of the Choquet integral with respect to a fuzzy measure, which is proved to be an adequate aggregator operator that extends the classical weighted arithmetic mean when the criteria adopted interact ([9]).

In this contribution, we propose to use fuzzy measures and the Choquet integral in order to compute a global performance index which allows to evaluate the performance of faculty members by taking into account multiple areas of the academic activity, multiple evaluation criteria within each area and also their possible connections. In particular, we adopt a two-step Choquet integral which permits to model a hierarchy in the aggregation process.

The rest of the paper is structured as follows. Section 2 reminds the concepts of Choquet integral and fuzzy measures and that of multi-step Choquet integral. Section 3 first shows how the Choquet integral can be used in the performance evaluation process of the academic activity and then presents a numerical example. Section 4 reports some final remarks.

2 Non Additive Measures for Aggregation

Non additive measures are up to now a commonly used method to to represent interactions between the elements of a set. In cooperation with integral aggregation functions, they are a well-founded framework able to aggregate information from several sources. Usually the monotonicity property is required in most of practical applications.

In particular in this paper we consider non-additive monotonic measures, also called fuzzy measures and Choquet integral for a n-dimensional vector as in [9] for example. Let N be a finite index set, $N = \{1, \dots, n\}$.

Definition 1. *A set function* $v : 2^N \to [0,1]$*, with* $v(\emptyset) = 0$*,* $v(N) = 1$ *and* $A \subseteq B \subseteq N$ *implies that* $v(A) \leq v(B)$*, is called a non-additive measure on* N*.*

We note that if $S \subseteq N$, $v(S)$ can be viewed as the importance of the set of elements S. Note also that non-additive measures encompass probability measures, belief functions and capacities.
Let now introduce the discrete Choquet integral on N viewed as an aggregation function that generalizes the weighted arithmetic means.

Definition 2. *Let* v *a non-additive measure* $v : 2^N \to [0,1]$ *and* $x \in \mathbb{R}^n$*. Let* σ *be a permutation of* N *such that* $\{x_{\sigma(n)}\}$ *forms a non decreasing sequence and we define* $x_{\sigma(0)} = 0$ *. Then the Choquet integral of* x *is:*

$$C_v(x) = \int x dv = \sum_{i=1}^{n} (x_i - x_{i-1}) v(\{\sigma(i), \ldots\ldots, \sigma(n)\}) \ . \tag{1}$$

This formula can be interpreted as an expectation operator with respect to a generalized measure. Notice that in the case in which the measure $v(\{\cdot, \ldots, \cdot\})$ is additive it is immediate to prove that (1) if the well known arithmetic weighted mean.

Now we give some basic definitions and on multi-step Choquet integral. The two-step Choquet integral has been investigated mainly in [10] and [11]. Let us now give a formal definition of a multi-step Choquet integral.

Definition 3. *Let $\Gamma \subseteq \mathbb{R}^n$. For any $i \in N$, the projection* $\mathbf{x} \mapsto x_i$ *is a* 0-step Choquet integral. *Let us consider $F_i : \Gamma \to \mathbb{R}$, $i \in M := \{1, \ldots, m\}$, being k_i-step Choquet integrals, and a non-additive measure v on M. Then*

$$F(\mathbf{x}) := C_v(F_1(\mathbf{x}), \ldots, F_m(\mathbf{x}))$$

is a k-step Choquet integral, *with $k := \max\{k_1, \ldots, k_m\} + 1$. A* multi-step Choquet integral *is a k-step Choquet integral for some integer $k > 1$.*

3 Performance Evaluation of Academic Activity: A Numerical Example

The aim of this section is twofold:

- to illustrate, in a synthetic way, the main academic activities, the related criteria, their hierarchical organization and their interactions that, in our opinion, have to take into account in order to evaluate an academic subject (a scholar, a department, a faculty, ...);
- to provide an exemplification about the "mechanics" of our two-step evaluation procedure.

As far as the first point is concerned, we propose an extension of the Research Factor index (on following: R), recently introduced in [12] for the evaluation of the research quality of an academic subject. In short, the R is obtained by a two-level process. In the first level this index is determined by a (simple) aggregation of two macro-criteria and their interaction; in the second level each of these macro-criteria is achieved by (simple) aggregation of various criteria and their interactions.

Given these bases, in order to globally evaluate the activities of an academic subject, we propose a Global index (on following indicated by C_G) obtained by a two-step Choquet integral based aggregation procedure. In particular, going backward, in the second step the C_G is achieved by the aggregation of three indexes and their interactions. These indexes are respectively related to the research (from which the x_R index), to the teaching (from which the x_T index) and to services[1] (from which the x_S index). We recall

[1] By "services" we mean, for instance, consultancy activities, membership on non-academic committees and so on.

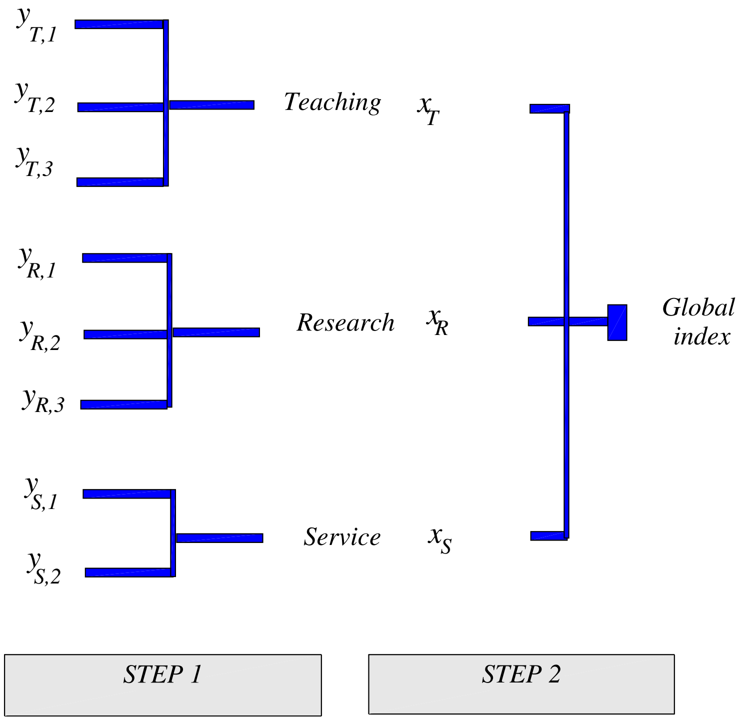

Fig. 1. Two-steps evaluation process

that we are interested in the evaluation of the overall quality of an academic subject. This is the main reason for which we consider all the various typologies of academic activities.

Formally, we can represent the output of this second step as follows:

$$C_G = C_v\left(x_R, x_T, x_S\right),$$

where v indicates the non-additive measure associated to the three indexes.

In the first step each of these indexes is determined by the aggregation of various criteria and the interactions among them.

Figure 1 schematizes the two-steps evaluation process.

Of course, also each of these criteria could be obtained by the aggregations of suitable sub-criteria, so leading to a three-step Choquet integral based aggregation process (the generalization to an n-step evaluation procedure is obvious).

About the choice of the criteria for each of the considered indexes, there exists a wide enough literature, and our personal expertises, from which drawing inspiration (see, for instance, [2], [5], [6], [8] and [12]). As the specification of such criteria does not constitute a goal of this note, the list of them we take into account has to be simply considered as exemplificative and not prescriptive. In any case, as general recommendation, we suggest that the number of criteria chosen for each index is not particularly large given the "curse of dimensionality" implied in the specification of the non-additive measures respectively related to the various indexes themselves (see, for details, definition 2).

Now we can present the list of these criteria. Notice that they are generally relating to a prefixed time period (for instance, the last five academic years). With reference to the research activities, we consider the following $N_R = 3$ criteria:

- $y_{R,1}$: an indicator of the quantity and the quality of the publications;
- $y_{R,2}$: an indicator of the numbers of the self-citations and the others-citations received by those publications;
- $y_{R,3}$: an indicator of the quantity and the quality of the coauthors of those publications;

(see, for more details on these criteria, [4]). With reference to the teaching activities, we consider the following $N_T = 3$ criteria:

- $y_{T,1}$: an indicator of the quantity and the quality of the degree courses;
- $y_{T,2}$: an indicator of the quantity and the quality of the master and the Ph.D. courses;
- $y_{T,3}$: an indicator of the student evaluations.

Finally, with reference to the service activities, we consider the following $N_S = 2$ criteria:

- $y_{S,1}$: an indicator of the quantity and the quality of the consultancy activities;
- $y_{S,2}$: an indicator of the numbers of membership on non-academic committees.

Formally, we can represent the outputs of this first step as follows:

$$x_j = C_{v_j}\left(y_{j,1}, \ldots, y_{j,N_j}\right), \text{ for all } j \in \{R, T, S\},$$

where v_js are the non-additive measures respectively associated to the various subset of multicriteria.

At this point we can pass to deal with the second of the two starting points, that is to give a numerical example of our approach. Since now we premise that in this application we do not use real data. However, the considered data are generated in such a manner to embody our views about the interactions among the three indexes (second step) and among the criteria respectively relating to each of these indexes (first step).

Now we first need to specify the transformations, defined in section 2, of the non-additive measures associated to the three indexes (see Table 1) and to the various subset of multicriteria (see Table 2).

We emphasize that our main views incorporated in $\pi(\cdot)$, $\pi_R(\cdot)$. $\pi_T(\cdot)$ and $\pi_S(\cdot)$ are:

- "*to do more is better than to do less*", in fact all the considered measures are strictly increasing with respect to their respective arguments
(for instance, $\pi_T(y_{T,2}) + \pi_T(y_{T,3}) < \pi_T(y_{T,2}, y_{T,3})$);
- "*academic activities are better than non-academic activities*", in fact, among all the considered measures, the only case of subadditivity is related to the one associated to the three indexes with respect to the service inputs
(for instance, $\max\{\pi(x_T), \pi(x_S)\} < \pi(x_T, x_S) < \pi(x_T) + \pi(x_S)$).

Table 1. Transformations of the measures associated to the three indexes

$\pi(\cdot)$
$\pi(\emptyset) = 0.00$
$\pi(x_R) = 0.35$
$\pi(x_T) = 0.30$
$\pi(x_S) = 0.15$
$\pi(x_R, x_T) = 0.75$
$\pi(x_R, x_S) = 0.45$
$\pi(x_T, x_S) = 0.40$
$\pi(Universe) = 1.00$

Table 2. Transformations of the measures associated to the various subset of multicriteria

$\pi_R(\cdot)$	$\pi_{TFi}(\cdot)$	$\pi_S(\cdot)$
$\pi_R(\emptyset) = 0.00$	$\pi_T(\emptyset) = 0.00$	$\pi_S(\emptyset) = 0.00$
$\pi_R(y_{R,1}) = 0.50$	$\pi_T(y_{T,1}) = 0.45$	$\pi_S(y_{S,1}) = 0.55$
$\pi_R(y_{R,2}) = 0.35$	$\pi_T(y_{T,2}) = 0.30$	$\pi_S(y_{S,2}) = 0.45$
$\pi_R(y_{R,3}) = 0.20$	$\pi_T(y_{T,3}) = 0.25$	$\pi_S(Universe) = 1.00$
$\pi_R(y_{R,1}, y_{R,2}) = 0.95$	$\pi_T(y_{T,1}, y_{T,2}) = 0.85$	–
$\pi_R(y_{R,1}, y_{R,3}) = 0.80$	$\pi_T(y_{T,1}, y_{T,3}) = 0.70$	–
$\pi_R(y_{R,2}, y_{R,3}) = 0.65$	$\pi_T(y_{T,2}, y_{T,3}) = 0.65$	–
$\pi_R(Universe) = 1.00$	$\pi_T(Universe) = 1.00$	–

Then, in order to make directly comparable the various criteria among them, we need to suitably normalize the criteria themselves in a prefixed interval, let us say $[0, 1]$. Also this aspect does not constitute a goal of this note. Anyway, a suggestion for a possible normalization function can be a piecewise linear function of the following type:

$$y = \begin{cases} 0 & \text{if } x \leq x_{min} \\ a + bx & \text{if } x_{min} < x \leq x_{max} \\ 1 & \text{if } x > x_{max} \end{cases}, \text{ with } x_{min} < x_{max},$$

where y indicates the normalized criterion, x indicates the non-normalized criterion, $a = 1 / (x_{max} - x_{min}) > 0$ and $b = -x_{min} / (x_{max} - x_{min}) < 0$.

Notice that both the specification of the non-additive measures and the definition of the normalized functions used in our approach should generally be performed by teams of experts or by focus groups.

Now, we can provide the numerical exemplification. Let us consider two academic subjects, respectively characterized by the following values of the criteria:

- $y^1_{R,1} = 0.6$, $y^1_{R,2} = 0.4$, $y^1_{R,3} = 0.2$, $y^1_{T,1} = 0.5$, $y^1_{T,2} = 0.1$, $y^1_{T,3} = 0.3$, $y^1_{S,1} = 0.2$, $y^1_{S,2} = 0.4$;
- $y^2_{R,1} = 0.5$, $y^2_{R,2} = 0.1$, $y^2_{R,3} = 0.3$, $y^2_{T,1} = 0.6$, $y^2_{T,2} = 0.4$, $y^2_{T,3} = 0.2$, $y^2_{S,1} = 0.2$, $y^2_{S,2} = 0.4$.

where the superscript identifies the academic subject.

Notice that the difference between the two academic subjects consists in the "exchange" of the values associated to the research criteria ($y_{R,1}$, $y_{R,2}$ and $y_{R,3}$) with the values associated to the teaching criteria ($y_{T,1}$, $y_{T,2}$ and $y_{T,3}$). In the following Table 3 we report in details for both the academic subjects the calculations related to our approach.

Table 3. Calculations related to our two-step Choquet integral based aggregation procedure

Step	Choquet integral
First	$x_R^1 = (0.20 - 0.00)1.00 + (0.40 - 0.20)0.95 + (0.60 - 0.40)0.50 = 0.4900$
	$x_T^1 = (0.10 - 0.00)1.00 + (0.30 - 0.10)0.70 + (0.50 - 0.30)0.45 = 0.3300$
	$x_S^1 = (0.20 - 0.00)1.00 + (0.40 - 0.20)0.45 = 0.2900$
	$x_R^2 = (0.10 - 0.00)1.00 + (0.30 - 0.10)0.80 + (0.50 - 0.30)0.50 = 0.3600$
	$x_T^2 = (0.20 - 0.00)1.00 + (0.40 - 0.20)0.85 + (0.60 - 0.40)0.45 = 0.4600$
	$x_S^2 = (0.20 - 0.00)1.00 + (0.40 - 0.20)0.45 = 0.2900$
Second	$C_G^1 = (0.29 - 0.00)1.00 + (0.33 - 0.29)0.75 + (0.49 - 0.33)0.35 = 0.3760$
	$C_G^2 = (0.29 - 0.00)1.00 + (0.36 - 0.29)0.75 + (0.46 - 0.36)0.30 = 0.3725$

As $C_G^1 > C_G^2$, academic subject 1 is "better" than academic subject 2.

4 Final Remarks

In this paper we propose a Choquet integral based method to evaluate academic performance. With respect to the previous literature one of the main contribution of our proposal concerns the computation of a numerical index which takes into account not only the scientific productivity but also all the other activities linked to academic functions. Again the non linear computation method permits to consider interactions among the criteria in a formal way, and, at the same time, to realize an easy approach to understand, implement and apply. We fill that this Choquet based linear aggregation algorithm will constitute an innovative way to approach this kind of evaluation method in contrast with most of the methods actual in use.

References

1. Alonso, S., Cabrerizo, F.J., Herrera-Viedma, E., Herrera, F.: h-Index: A review focused in its variants, computation and standardization for different scientific fields. Journal of Informetrics 3, 273–289 (2009)
2. Bana e Costa, C.A., Oliveira, M.D.: A multicriteria decision analysis model for faculty evaluation. Omega 40, 424–436 (2012)
3. Burrell, Q.: On the h index, the size of the Hirsh core and Jin's A index. Journal of Informetrics 1, 170–177 (2007)
4. Cardin, M., Corazza, M., Funari, S., Giove, S.: A fuzzy-based scoring rule for author ranking. An alternative of h index. In: Apolloni, B., Bassis, S., Esposito, A., Morabito, C.F. (eds.) Neural Nets WIRN 2011. Proceedings of the 21st Italian Workshop on Neural Nets. Frontiers in Artificial Intelligence and Applications, vol. 234, pp. 36–45. IOS Press (2011)

5. Felder, R.M., Brent, R.: How to evaluate teaching. Chemical Engineering Education 38, 200–202 (2004)
6. Hirsch, J.: An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences 102, 16569–16572 (2005)
7. Hoyt, D.P., Pallett, W.H.: Appraising Teaching Effectiveness: Beyond Student Ratings. IDEA PAPER 36 (1999)
8. Marchant, T.: Score-based bibliometric rankings of authors. Journal of the American Society for Information Science and Technology 60, 1132–1137 (2009)
9. Marichal, J.L.: Aggregation of interacting criteria by means of the discrete Choquet integral. In: Calvo, T., Mayor, G., Mesiar, R. (eds.) Aggregation Operators: New Trends and Applications. STUDFUZZ, vol. 97, pp. 224–244. Physica-Verlag (2002)
10. Mesiar, R., Vivona, D.: Two-step integral with respect to fuzzy measure. Tatra Mountains Mathematical Pubblications 16, 359–368 (2006)
11. Narukawa, Y., Murofushi, T.: The n-step Choquet integral on finite spaces. In: Proceedings of the 9th International Conference on Information Processes and Management of Uncertainty in Knowledge-Based Systems, Annecy, France (2002)
12. Tucci, M.P., Fontani, S., Ferrini, S.: L'R-Factor: un nuovo modo di valutare l'attività di ricerca. Studi e Note di Economia 1, 103–140 (2010)