Talanta 132 (2015) 285–293

Contents lists available at ScienceDirect

Talanta

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

Evaluation of analytical performance based on partial order methodology

Lars Carlsen [a,b,*], Rainer Bruggemann [c], Olga Kenessova [d], Erkin Erzhigitov [d]

[a] Awareness Center, Linkøpingvej 35, Trekroner, DK-4000 Roskilde, Denmark
[b] Kazakh British Technical University, Department of Chemical Engineering, Almaty, Kazakhstan
[c] Leibniz – Institute of Freshwater Ecology and Inland Fisheries, Department Ecohydrology, Müggelseedamm 310, D-12587 Berlin, Germany
[d] RSE Kazakh Institute of Metrology, Almaty, Kazakhstan

ARTICLE INFO

Article history:
Received 27 June 2014
Received in revised form
27 August 2014
Accepted 5 September 2014
Available online 19 September 2014

Keywords:
Analytical performance
Partial order
Posets
Hasse diagram
Multi-indicator systems

ABSTRACT

Classical measurements of performances are typically based on linear scales. However, in analytical chemistry a simple scale may be not sufficient to analyze the analytical performance appropriately. Here partial order methodology can be helpful. Within the context described here, partial order analysis can be seen as an ordinal analysis of data matrices, especially to simplify the relative comparisons of objects due to their data profile (the ordered set of values an object have). Hence, partial order methodology offers a unique possibility to evaluate analytical performance. In the present data as, e.g., provided by the laboratories through interlaboratory comparisons or proficiency testings is used as an illustrative example. However, the presented scheme is likewise applicable for comparison of analytical methods or simply as a tool for optimization of an analytical method. The methodology can be applied without presumptions or pretreatment of the analytical data provided in order to evaluate the analytical performance taking into account all indicators simultaneously and thus elucidating a "distance" from the true value. In the present illustrative example it is assumed that the laboratories analyze a given sample several times and subsequently report the mean value, the standard deviation and the skewness, which simultaneously are used for the evaluation of the analytical performance. The analyses lead to information concerning (1) a partial ordering of the laboratories, subsequently, (2) a "distance" to the Reference laboratory and (3) a classification due to the concept of "peculiar points".

1. Introduction – the need for several indicators

Evaluating analytical performance may be assessed, e.g., through interlaboratory comparisons and/or proficiency testings or comparisons of analytical methods or through the optimization process of an analytical method. In the following we will apply the analytical performance of a series of laboratories as an illustrative example how analytical performance advantageously can be evaluated beyond the restrictions imposed by a single-quantity-approach such as the evaluation by checking only the accuracy (see below). Participation in interlaboratory comparisons or proficiency testings and a satisfying result from these are typically a prerequisite for obtaining accreditation to perform a given analysis.

Result of the single laboratory maybe expressed of the *z*-score (see below) only, which measure of how "far away" the single analytical result is from the accepted value, i.e., the *z*-score [1,2]. The *z*-score of the reference laboratory by definition being 0 (Eq. (1)). Obviously, the *z*-score does not as such tell the full story and other measures should be considered simultaneously. These (in the present case three) measures will a priori form a grid in the three-dimensional space and laboratories can be represented as a point in the space spanned by the three coordinates. The projection to each of these three coordinates bears information of its own right and it seems to be careless to aggregate the three quantities in order to obtain an easy to handle scale for ranking purposes. Various approaches for aggregating the indicators have been reported [3,4]. However, such aggregation of the indicators may unequivocally lead to compensation effects where high values in one indicator may compensate low values in another [5].

Taking all three indicators into account simultaneously is a bit more complicated than the use of a single scalar as we are facing a multi-indicator system (MIS). Partial order theory applied in multi-indicator systems is also known as Hasse diagram technique (HDT) [6,7] and appears in this connection as a highly advantageous tool. The visual representation of a partial order can be performed by a so-called Hasse diagram. In the Hasse diagram comparable objects are connected by a sequence of lines [6] and the diagram is characterized by its structure that comprises levels, chains and antichains (see

* Corresponding author at: Awareness Center, Linkøpingvej 35, Trekroner, DK-4000 Roskilde, Denmark.
E-mail address: LC@AwarenessCenter.dk (L. Carlsen).

http://dx.doi.org/10.1016/j.talanta.2014.09.009
0039-9140/

below). An often heard criticism of partial order ranking approaches is that the method is lacking of a unique or absolute ranking, i.e., the presence of incomparabilities. In a Hasse diagram incomparabilities are causing that the Hasse diagram (cf. Fig. 2) has not only a vertical (ranking) but also a horizontal geometrical configuration. Hence, applying partial ordering implies comparison among different objects. The consequent application of this concept leads to figures, which are known in algebra as (directed) graphs.

2. Theory and methods

2.1. Indicators

The z-score denotes the standard score [2] or accuracy [1,8] of the measurements, which enables a comparison between two scores from different normal distributions.

$$z = (m - \mu_a)/s_a \quad (1)$$

where m is the reported value from the laboratory and μ_a and s_a are the accepted value (mean) and standard deviation, respectively, as generated by the Reference laboratory, i.e., the laboratory assigned to perform the interlaboratory comparisons and proficiency testings [9]. In other words: a Reference laboratory is by definition considered as the "best". Also the standard deviation of the laboratory measurements is important to judge the laboratory performance, i.e., the precision of the measurements [1]. Finally a further possible factor to bring into play to elucidate the laboratory performance is the skewness of the distribution of the laboratory measurements [10]. This factor (indicator) tells about the shape of the distribution and thus to what extent the measurements are in analytical or statistical control [11]. The three parameters are all standard output from statistical treatment of the results from a series of replicate measurements (vide infra)

Obviously these factors – or indicators – for evaluating laboratory performance on an individual basis can be easily compared and thus each of them then may constitute the basis for ranking of the laboratories, i.e. the performance of the laboratory relative to each other as well as to the Reference laboratory. As such this may constitute an evaluation tool for selecting laboratories for specifics task. However, only part of the story is told and the information basis is obviously limited.

In the present study the absolute values of the z-score, |z|, and the skewness, |skew|, together with the standard deviation, s, are applied as indicators. The [0, 1] normalized indicators are called r_i (i=1, 2, 3) and any possible triple of values of a certain laboratory is consequently called r. Thus, each laboratory is characterized by a data profile, i.e., an ordered set of indicator values (|z|,s,|skew|).

Generally indicators r_i can be normalized to a [0, 1]-scale (called rn_i) by

$$rn_i(x) = (r_i(x) - r_i\ \mathrm{min})/(r_i\ \mathrm{max} - r_i\ \mathrm{min}) \quad (2)$$

where $r_i$max and $r_i$min being the maximum, minimum value with respect to the objects. Here the indicators r_i are assumed to be already normalized and we consider r as the triple (r_1, r_2, r_3)=(normalized z-score, normalized standard deviation, normalized skewness).

2.2. Partial order

2.2.1. Overview

In this paper the numerical aggregation for example by calculating a composite indicator CI is avoided.

$$\mathrm{CI} = \sum g_1 \cdot rn_i \quad g_i : weights, \quad i = 1, \ldots, 3 \quad (3)$$

The alternative is based on partial order methodology. The subject "Partial Order" is of increasing interest in the mathematical field as well as in many other applied fields. Here a pretty simple variant of partial order is applied to study MIS, according to a recent paper by Bruggemann and Carlsen [12]. This methodology can be applied, whenever a MIS and a data matrix are available and a ranking is of interest. For example partial order was applied in environmental sciences, in biology, in economy and social sciences. Good overviews can be found in Refs. [7,13]. Partial ordering can be considered as a non-parametric method as, in contrast to standard multidimensional analyses no assumptions about linearity or distribution of the indicators are made and are necessary.

2.2.2. The basical equation of Hasse diagram technique

In Partial Ordering the only mathematical relation among the objects is "$\leq$" [6,7,14]. The "$\leq$"-relation is the basis for a comparison of laboratories and constitutes a graph. A laboratory x is connected with laboratory y if and only if the relation $x \leq y$ holds. The crucial question is, under which conditions it can be claimed that $x \leq y$. Here, a system is described by a series of indicators r_j, the MIS. A laboratory, x, characterized by the a set of indicators $r_j(x)$, j=1, …, m can be compared to another laboratory y, characterized by the indicators $r_j(y)$, when

$$\mathrm{r}_i(x) \leq r_i(y) \quad for \quad all \quad i = 1, \ldots, m \quad (4)$$

Eq. (4) is a very hard and strict requirement for establishing a comparison. It demands that all indicators of laboratory x should be better (or at least equal) than those of laboratory y through comparison of the single indicators, respectively. To be still a little bit more technically: Let X be the group of laboratories studied, i.e., X={REF, lab1, lab2, …., lab20}, the laboratory lab$x$ will be ranked higher than laboratory lab$y$, i.e., lab$x$ > lab$y$ or still more condensed $x > y$, if at least one of the indicator values for lab$x$ is higher than the corresponding indicator value for lab$y$ and no indicator for lab$x$ is lower than the corresponding indicator value for lab$y$. However, if r_j(lab$x$) > r_j(lab$y$) for some indicator j and r_i(lab$x$) < r_i(lab$y$) for some other indicator i, lab$x$ and lab$y$ will be called incomparable (notation: lab$x$||lab$y$) expressing the mathematical contradiction due to conflicting indicator values. A set of mutual incomparable objects is called an antichain. When all indicator values for lab$x$ are equal to the corresponding indicator values for lab$y$, i.e., r_j(lab$x$)=r_j(lab$y$) for all j, the two objects will have identical rank and will be considered as equivalent, i.e., lab$x$~lab$y$.

Eq. (4) is the basic for the Hasse diagram technique (HDT) [6,7,14], which is a special (statistically oriented part of partial order theory). Hasse diagrams are visual representation of the partial order. In the Hasse diagram comparable objects are connected by a sequence of lines [14–17]. For construction of the Hasse diagram, a uniform orientation of the indicators should be secured, i.e., high indicator values correspond to "bad" objects and low values to "good" objects. In Fig. 2 (Section 3) the Hasse diagram based on the data given in Table 1 is shown. Thus, in the present case the higher a laboratory is placed in the Hasse diagram the farther away from the performance of the Reference laboratory it is.

2.2.3. Structure of a Hasse diagram

A Hasse diagram is characterized with it structure that comprises levels, chains and antichains (cf. Fig. 2).

Level: The horizontal arrangement of objects within a Hasse diagram. The level structure gives a first approximation to a weak order (i.e. tied ranks are not excluded: thus, for example in the ordering of four objects $a < b = d < e$ there is a tie, i.e., the equivalence of b and d with respect to the rank.) of the objects from "good" (bottom) to "bad" (top). Unfortunately, this will often give rise to many tied ranks as all objects in a level automatically will be assigned identical ranks. Typically, it is desirable that the degree of tied ranks should be as low as possible, i.e., there is a need for a linear