Ecological Indicators 80 (2017) 12–22

Contents lists available at ScienceDirect

Ecological Indicators

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

Weights and importance in composite indicators: Closing the gap

William Becker, Michaela Saisana, Paolo Paruolo, Ine Vandecasteele

European Commission, Joint Research Centre, Via Enrico Fermi, 2749, 21027 Ispra VA, Italy

ARTICLE INFO

Keywords:
Composite indicators
Sensitivity analysis
Correlation
Splines
Gaussian process
Smoothing
Optimisation

ABSTRACT

Composite indicators are very popular tools for assessing and ranking countries and institutions in terms of environmental performance, sustainability, and other complex concepts that are not directly measurable. Because of the stakes that come with the media attention of these tools, a word of caution is warranted. One common misconception relates to the effect of the weights assigned to indicators during the aggregation process. This work presents a novel series of tools that allow developers and users of composite indicators to explore effects of these weights. First, the importance of each indicator to the composite is measured by the nonlinear Pearson correlation ratio, estimated by Bayesian Gaussian processes. Second, the effect of each indicator is isolated from that of other indicators using regression analysis, and examined in detail. Finally, an optimisation procedure is proposed which allows weights to be fitted to agree with pre-specified values of importance. These three tools together give developers considerable insight into the effects of weights and suggest possibilities for refining and simplifying the aggregation. The added value of these tools are shown on three case studies: the Resource Governance Index, the Good Country Index, and the Water Retention Index.

1. Introduction

Composite indicators (also known as synthetic indices or performance indices) are popular tools for assessing the performance of countries/entities on human development, sustainability, perceived corruption, innovation, competitiveness, or other complex phenomena that are not directly measurable and not uniquely defined. Examples include the Human Development Index (Jahan, 2015), the Sustainable Society Index (Van de Kerk and Manuel, 2008), the Financial Secrecy Index (Cobham et al., 2013) and the Environmental Performance Index (Hsu, 2016). Composite indicators are employed for many purposes, including policy monitoring, communication to the public, and generating rankings.

The popularity of rankings owes to two main reasons. First, their simplicity: they provide a summary picture of the multiple facets or dimensions of complex, multidimensional phenomena in a way that facilitates evaluation and comparison. Second, rankings force institutions and governments to question their standards; rankings are drivers of behaviour and of change (Kelley and Simmons, 2015). Hence, it comes at no surprise that over the past two decades there has been a turbulent growth of performance indices. Bandura (2011) provides a comprehensive inventory of over 400 country-level indexes monitoring complex phenomena from economic progress to educational quality. Similarly, a more recent inventory by the United Nations (Yang, 2014) details 101 composite measures of human well-being and progress, covering a broad range of themes from happiness-adjusted income to environmentally-adjusted income, from child development to information and communication technology development.

Even though considerable attention is given to the rankings of composite indicators, many subjective choices are made in their construction: this has motivated studies which perform uncertainty and sensitivity analysis on composite indicator assumptions (Saisana et al., 2005). One important step is the aggregation of indicators, where typically the variables are combined in a weighted average to give the resulting value of the composite indicator. Apart from the decision of which kind of weighted average to use (e.g. arithmetic, geometric), the developer must select values of weights to apply to each variable. The values of these weights can have a large impact on the subsequent rankings, which often goes unnoticed. Understanding the impact of weights on the variation of the composite indicator scores is therefore important.

A possible misconception is that the weight assigned to a variable can be directly interpreted as a measure of importance of the variable to the resulting value of the composite indicator. Indeed, in common approaches to composite indicator weighting such as budget allocation (expert input) and equal weighting, this appears to be the supporting logic. However this is rarely the case: different variances and correlations among variables, for instance, prevent the weights from corresponding to the variables' importance. Two questions immediately arise: first, given a set of weights and a sample, what is the influence of

E-mail addresses: william.becker@ec.europa.eu (W. Becker), michaela.saisana@ec.europa.eu (M. Saisana), paolo.paruolo@ec.europa.eu (P. Paruolo), ine.vandecasteele@ec.europa.eu (I. Vandecasteele).

http://dx.doi.org/10.1016/j.ecolind.2017.03.056
Received 23 February 2017; Received in revised form 30 March 2017; Accepted 30 March 2017
Available online 02 May 2017
1470-160X/ © 2017 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/BY/4.0/).

each variable on the output? Second, how can weights be assigned to reflect the desired importance?

This article puts together, for the first time, tools that allow developers to examine in detail the effect of weights and to subsequently refine the composite indicator. First, building on an earlier proposal of some of the present authors, see Paruolo et al. (2013), the relative importance of each indicator is measured with the *Pearson correlation ratio*, which is a variance-based measure that accounts for (possibly nonlinear) dependence between input variables and the composite indicator (see Section 2). Going beyond Paruolo et al. (2013), who use local-linear regression to estimate the correlation ratio, the main tool used here is Bayesian Gaussian processes, which have the advantage of providing confidence intervals on the Pearson correlation ratio (see Section 3). This overall approach echoes the work of Da Veiga et al. (2009) in the sensitivity analysis literature, but with the additional advantage that it does not require the use of an independent sample from the marginal distribution of each indicator to estimate the variance of the conditional expectation.

As a further step, to better understand the influence of variables on the composite indicator, an approach based on the correlated sensitivity analysis work of Xu and Gertner (2008a,b) is proposed in Section 4, which uses additional regressions to decompose the influence of each variable into influence caused by correlation, and influence caused by the composite indicator structure (aggregation and weights).

Finally, the issue of optimisation of weights is considered in Section 5. Although a linear solution to the optimisation problem is available in Paruolo et al. (2013) using linear regression, the proposal here is to extend it to nonlinear regression, to account for nonlinear main effects. This is done by numerically minimising the difference between correlation ratios and their desired values. Given that this involves a large number of regression fits, penalised splines are preferred over Gaussian processes because of their low computational cost.

The various strands of methodology in this article are demonstrated on three case studies in Section 6, which are chosen as examples related to sustainability and quality of life: the Resource Governance Index, the Good Country Index, and the Water Retention index. The first two indices were chosen because of their potentially high policy impact owing to a considerable media presence. The latter index is chosen for academic purposes, in order to evidence the advantage of using penalised splines in studies with large number of entities (thousands of drainage basins in this case).

The tools to perform these analyses in Matlab are available for free download on the author's web page (Becker, 2017).

2. Measures of importance

Consider the case of a composite indicator y (output) calculated by aggregating over d normalised input variables (indicators) $\{x_i\}_{i=1}^{d}$. The most common aggregation scheme is the weighted arithmetic average, i.e.

$$y_j = \sum_{i=1}^{d} w_i x_{ji}, \quad j = 1, 2, \cdots, n \tag{1}$$

where x_{ji} is the normalised score of individual j (e.g., country) based on the value X_{ji} of the ith raw variable $i = 1, \ldots, d$ and w_i is the nominal weight assigned to the i-th variable, such that $\sum_{i=1}^{d} w_i = 1$ and $w_i > 0$.

Now consider an importance measure I, which captures the influence of each x_i on y, and which is also normalised to sum to one over all d variables. The fundamental underlying principle of this paper is that $I_i \neq w_i$, nor is I necessarily linearly related to w, although this fact is sometimes overlooked by developers. In fact, the importance of x_i is also strongly dependent on its (possibly nonlinear) correlations with other variables, which are in turn correlated with each other. Therefore determining and isolating the effect of x_i on y is by no means trivial.

For any given composite indicator, one can define measures of importance of each of the input variables x_i with respect to the output y of the composite indicator. One approach is to measure the dependence of y on x_i. Consider the decomposition,

$$y_j = f_i(x_{ji}) + \varepsilon_j, \tag{2}$$

where $f_i(x_{ji})$ is an appropriate function, possibly nonlinear, that models the conditional mean of y given a sample point x_{ji}—and ε_j is an error term which accounts for variation due to indicators other than x_i. A well-known way to measure the *linear* dependence of y on x_i is to use the coefficient of determination R_i^2: in sample, this can be computed as:

$$\mathrm{SS}_{\mathrm{reg},i}/\mathrm{SS}_{\mathrm{tot}}, \tag{3}$$

where (in the case of R^2) $\mathrm{SS}_{\mathrm{reg},i} = \sum_{j=1}^{n} (\widehat{f_i}(x_{ji}) - \bar{y})^2$ is the sum of squares explained by the *linear* regression, $\bar{y} := n^{-1}\sum_{j=1}^{n} y_j$ is the sample average, $\widehat{f_i}(x_{ji}) = \widehat{\beta}_0 + \widehat{\beta}_1 x_{ji}$ is the *linear* fit for observation y_j and $\mathrm{SS}_{\mathrm{tot}} = \sum_j (y_j - \bar{y})^2$ is the total sum of squares. R_i^2 can hence be seen as the ratio of the sum of squares explained by the linear regression of y on x_i, and the total sum of squares of y. Since this measure is based on linear regression, it does not account for any nonlinearities between y and x_i.

In composite indicators, although the aggregation formula is often linear, a nonlinearity in the relationship between y on x_i can be introduced by the correlation between variables. In such cases, R_i^2 may underestimate the degree of dependence. The nonlinear measure adopted here is the *correlation ratio*, S_i, $i = 1, 2, .., d$ also widely known as the *first order sensitivity index*, or *main effect index* (see e.g. Becker and Saltelli, 2015). This measure is meant to measure the (possibly nonlinear) influence of each variable on the composite indicator, and is a nonlinear generalisation of R_i^2, such that R_i^2 equals S_i when $f_i(x_{ji})$ is linear. It can be interpreted as the expected variance reduction of the composite indicator scores, if a given variable were fixed. It is defined as:

$$S_i \equiv \eta_i^2 := \frac{\mathrm{V}_{x_i}(\mathrm{E}_{\mathbf{x}_{\sim i}}(y|x_i))}{\mathrm{V}(y)}, \tag{4}$$

where $\mathbf{x}_{\sim i}$ is the vector containing all the variables $(x_1, \ldots, x_d)$ except variable x_i. The term $E_{\mathbf{x}_{\sim i}}(y|x_i)$ is explicitly stated here with its subscript $\mathbf{x}_{\sim i}$ to emphasise that it is the expected value of y (the composite indicator), at a given value of x_i, with the expectation taken over $\mathbf{x}_{\sim i}$ (hereafter the subscript $x_{\sim i}$ is omitted to avoid cluttering the notation, such that $E_{\mathbf{x}_{\sim i}}(y|x_i) \equiv E(y|x_i)$). In other words, it is *conditional on* x_i, e.g. with x_i fixed at one value in its interval of variation. However, $E(y|x_i)$ is specified so that the value that y is conditioned on is not specified. Therefore $E(y|x_i)$ is a *function* of x_i. It is also known as the *main effect* of x_i, and is equivalent to the $f_i(x_{ji})$ discussed previously; therefore it is nothing more than a nonlinear regression fit on a scatter plot of y against x_i.

Now let $m_{ji} := \widehat{f_i}(x_{ji})$, corresponding to the fitted regression value (of y on x_i) corresponding to the jth sample point y_j. The correlation ratio S_i can then be estimated in sample as

$$\hat{S}_i = \sum_{j=1}^{n} (m_{ji} - \overline{m_i})^2 / \sum_{j=1}^{n} (y_j - \bar{y})^2 \tag{5}$$

where $\overline{m_i}$ is the mean of the m_{ji}, i.e. $\overline{m_i} := n^{-1}\sum_{j=1}^{n} m_{ji}$, $m_j := \hat{m}(x_{ji})$. Eq. (5) mimics (3).

In summary, this means that the correlation ratio, which is a nonlinear measure of dependence of y on x_i, can be estimated simply by fitting a nonlinear regression to a scatter plot of y against x_i, taking the variance of the resulting curve, and standardising by the unconditional variance of y.

3. Estimating the main effect

As discussed, the main effect $\mathrm{E}(y|x_i)$ is simply the nonlinear regression fit of y against x_i: therefore it can be estimated by various

Download English Version:

https://daneshyari.com/en/article/5741363

Download Persian Version:

https://daneshyari.com/article/5741363

Daneshyari.com