



Effect of sub-indicator weighting schemes on the spatial dependence of multidimensional phenomena

Matheus Pereira Libório¹ · João Francisco de Abreu¹ ·
Petr Iakovlevitch Ekel^{1,2} · Alexei Manso Correa Machado^{1,2}

Received: 8 September 2021 / Accepted: 22 November 2022 / Published online: 29 December 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

The weighting of sub-indicators is widely debated in the composite indicator literature. However, these weighting schemes' effects on the composite indicator's spatial dependence property are still little known. This research reveals a direct relationship between the weighting scheme of sub-indicators and the spatial autocorrelation of the composite indicator. The Global Moran's Index (I) of composite indicators built using Data-driven (Moran's $I=0.636$) and Hybrid (Moran's $I=0.597$) weighting schemes is, on average, eleven percent higher than in the Equal-weights (Moran's $I=0.549$) and Expert opinion (Moran's $I=0.560$) weighting schemes. The average score of the composite indicator is higher when they are built by weighting schemes that better describe the spatial dependence. The spatial dependence of sub-indicators and composite indicators are not related. All fifteen sub-indicators show lower spatial autocorrelation than the composite indicators built by Expert opinion, Hybrid, and Data-driven weighting schemes. The spatial weighting matrix influences the spatial autocorrelation but does not change the robustness and quality parameters of the composite indicator. The research develops a Data-driven weighting scheme that allows individually or simultaneously considering the opinion of experts and parameters of quality and robustness of the composite indicator. It also offers the means to reduce judgment errors and evaluation biases in Expert opinion sub-indicator weighting schemes.

Keywords Composite indicators · Spatial dependence · Sub-indicators weighting · Moran's Index

JEL Classification C02 · C43 · C44 · D6 · I3

✉ Matheus Pereira Libório
m4th32s@gmail.com

Extended author information available on the last page of the article

1 Introduction

Composite indicators are one-dimensional measures that facilitate the interpretation and analysis of multidimensional phenomena in different areas of knowledge (Mazziotta and Pareto 2017; Kuc-Czarnecka et al. 2020; Bernardes et al. 2021; Ekel et al. 2022). In particular, composite indicators are especially useful for geographic analysis because they represent the several sub-indicators adjacent to the multidimensional phenomenon on a single map (Carley et al. 2018; Marzi et al. 2019). However, building composite indicators involves significant shortcomings during normalization, weighting, and aggregation of sub-indicators (Saltelli 2007; Dialga and Le Giang 2017; Cinelli et al. 2021).

The weights of the sub-indicators can be assigned using the Equal-weight, Data-driven, Expert opinion, or Hybrid weighting schemes (Becker et al. 2017). There is no single best weighting scheme, as none of them are exempt from criticism, and each offers a different perspective of analysis (Greco et al. 2019). From a geographical point of view, the weighting scheme may or may not consider spatial heterogeneity.

Spatial heterogeneity is considered in methods that assign different weights to the same sub-indicator in different spatial units. In the Benefit-of-the-Doubt method, higher weights are assigned to higher performance sub-indicators of spatial units, offering a composite indicator that highlights the positive characteristics of each spatial unit (Fusco et al. 2018; Libório et al. 2022c). In the Geographic Weighted Principal Component Analysis method, weights maximize the variance extracted from the original data in the first component in each spatial unit, providing a composite indicator that captures locally as many sub-indicators of the multidimensional phenomenon as possible (Harris et al. 2015; Kallio et al. 2018; Cartone and Postiglione 2021). In the Ordered Weighted Averaging operator, the assignment of weights makes it possible to regulate the compensation levels between sub-indicators of poor and above-average performance, offering a composite indicator that highlights the positive or negative aspects of each of them the spatial units (Badea et al. 2011).

In weighting schemes that disregard spatial heterogeneity, the weights of a sub-indicator do not vary between spatial units, offering better comparability of composite indicator scores. The Equal-weights is an objective weighting scheme, easy to implement, and appropriate when there is a lack of a theoretical framework that justifies the differentiation of weights, when the experts disagree on the weights, or when the statistical and empirical knowledge is inadequate (Greco et al. 2019). Expert opinion weighting schemes are based on the concept of the multidimensional phenomenon.¹ Therefore, greater weights are assigned to sub-indicators with greater conceptual importance, offering a composite indicator closer to the conceptual structure of the multidimensional phenomenon (Nardo et al. 2005). Data-driven weighting schemes use the data to assign weights to sub-indicators. In the principal

¹ The most popular Expert opinion weighting methods are the Analytic hierarchy process, Conjoint analysis, Public opinion and Budget allocation process (Saisana & Tarantola 2002; Nardo et al. 2005; Otoo et al. 2021).

component analysis, the weights of the sub-indicators are defined to maximize the variance extracted from the data in the first component (Libório et al. 2020a). This approach offers a composite indicator that explains most of the information contained in the original sub-indicators (Saisana and Tarantola 2002). Enhanced Scatter Search assigns weights to sub-indicators to reduce rank reversals in robustness analyses, offering a more stable composite indicator (Maricic et al. 2019). The weights can also be assigned to minimize the effects of correlations between sub-indicators on the weighting, offering a composite indicator that maximizes the transfer of information in the sub-indicators (Lindén et al. 2021).

Considering spatial heterogeneity or not, research on the weighting of sub-indicators does not address a relevant topic for the geographic area: the spatial dependence of the multidimensional phenomenon. Spatial dependence is a property in which observations from different geographic areas can be correlated. Thus, observations from neighboring areas tend to be more similar than observations from distant areas (Goodchild 1991). The property of spatial dependence is associated with the First Law of Geography: "everything is related to everything else, but near things are more related to each other" (Tobler 1970).

The Moran's (1950) Index and Geary's (1954) Coefficient are two popular spatial autocorrelation measures for measuring the overall spatial dependence of a set of observations. The Moran's Index is based on the covariance between associated locations, while Geary's Coefficient considers numerical differences between the associated locations (Getis 1999). These indices are calculated based on a spatial weighting matrix built by proximity, contiguity, or geographic distance criteria.

Several studies employ the Moran's Index and the Geary Coefficient to analyze the spatial dependence of multidimensional phenomena. These researches bring examples of spatial autocorrelation of multidimensional phenomena in urban and rural areas and at different geographic scales: district, municipal, state, and regional (Jha and Gundimeda 2019; Katumba et al. 2019; Aljoufie and Tiwari 2020; Chauhan et al. 2020; Adeleke and Alabede 2021; Mavhura et al. 2021). However, these studies ignore that sub-indicator weighting schemes directly correlate with the composite indicator's results (Greco et al. 2019) and, consequently, influence its spatial autocorrelation.

This research aims to examine the effect of sub-indicator weighting schemes on the spatial dependence of a multidimensional phenomenon. This analysis allows answering questions such as: How does spatial autocorrelation relate to the robustness and quality parameters of the composite indicator? What are the effects of the weighting scheme in representing the multidimensional phenomenon? What is the relationship between the spatial autocorrelation of the sub-indicators and the composite indicator? What is the effect of the spatial weighting matrix on spatial autocorrelation? How is this effect reflected in the robustness and quality parameters of the composite indicator?

The first innovation of this research is associated with the introduction of procedures that avoid judgment errors and evaluation biases in the Expert opinion weighting scheme. The first procedure allows the processing of different assessment formats, offering flexibility and psychological comfort that reduce judgment errors. The second procedure allows for measuring the extent to which individual and collective

assessments differ by sub-indicator, signaling possible biased assessments. The second innovation of this research is associated with developing a flexible Data-driven sub-indicator weighting scheme that allows for emphasizing the property of spatial dependence of the composite indicator, taking into account, individually or simultaneously, parameters of robustness and quality.

2 Materials and methods

The details of the materials and methods used in the research are divided into six subsections. Section 2.1 presents the steps for building the distance and spatial weighting matrices and calculating Moran's Index. In Sects. 2.2 and 2.3, the stages of the building of composite indicators are presented. The building method of composite indicators is presented in Sect. 2.2, and the Equal-weights, Expert opinion, Data-driven, and Hybrid weighting schemes are presented in Sect. 2.3. The robustness and quality of the composite indicator and the consistency of the results are verified in Sects. 2.4, 2.5, and 2.6.

2.1 Spatial model of the study area: weighting matrix and spatial autocorrelation

The study area of this research is an urban conurbation composed of three cities in the state of Paraná, Brazil. The cities of Maringá-Sarandi-Paiçandu have an estimated population of 577,611 inhabitants² and are divided into 661 census tracts (IBGE, 2010). Figure 1 shows these census tracts and the location map of the study area.

A spatial weighting matrix was built considering the 661 census tracts of the study area. This matrix reflects the arrangement of spatial interactions through spatial weights that represent the degree of connection between the census tracts. The degree of connection between the census tracts represented by the spatial weights was obtained using the non-binary geographic distance proximity criterion. This criterion has two advantages. First, it avoids the unbalanced connectivity that occurs due to census tracts with many neighbors and census tracts with few neighbors. Second, it eliminates subjectivity in defining the influence of distance from neighboring areas. The spatial weighting matrix was built as follows:

First, the Euclidean distances were calculated using the latitude and longitude of the centroids of the census tracts $u, v \in U$ where U is the total number of census tracts:

$$d_{uv} = \sqrt{(u_i - v_i)^2 + (u_j - v_j)^2}, \quad u, v \in U. \quad (1)$$

² Population Estimates carried out by the Brazilian Institute of Geography and Statistics in July 2021. Accessible at <https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao>.

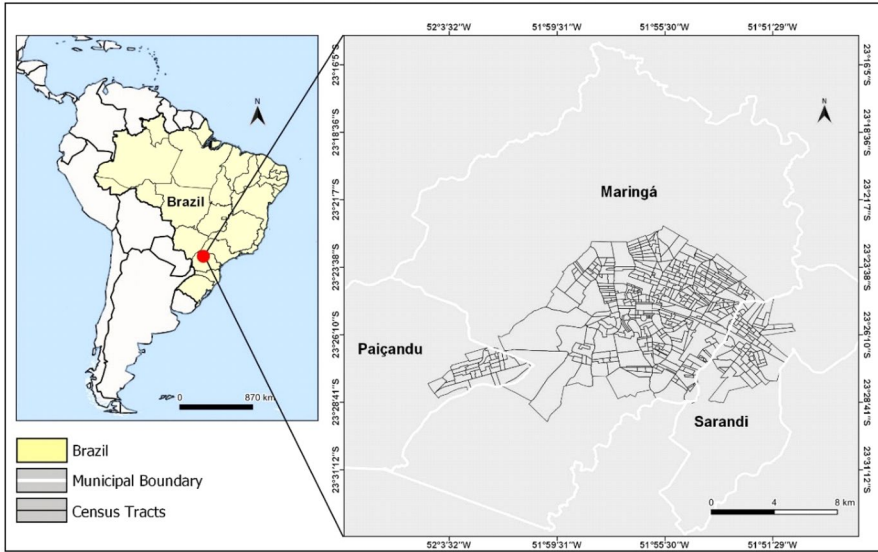


Fig. 1 The study area location map: Maringá-Sarandi-Paiçandu, Paraná, Brazil

Then, the spatial weight matrix can be elaborated based on the distance matrix d_{uv} and the distance sensitivity parameters b :

$$w_{uv} = \exp(-bd_{uv}), \quad u, v \in U \tag{2}$$

where the distance sensitivity parameter b was set to 1.

Finally, the spatial weighting matrix was normalized as

$$w_{uv}^* = \frac{w_{uv}}{\sum_{v \in U} w_{uv}}, \text{ so that } \sum_{v \in U} w_{uv}^* = 1, \tag{3}$$

where w_{uv}^* is the normalized matrix w_{uv} . $u = \sum_{v \in U} w_{uv}$.

Based on w_{uv}^* , the Moran's Index global spatial autocorrelation coefficient was calculated as

$$\text{Moran's } I = \frac{N}{\sum_u \sum_v w_{uv}^*} \frac{\sum_u \sum_v w_{uv}^* (X_u - \bar{X})(X_v - \bar{X})}{\sum_u (X_u - \bar{X})^2}, \tag{4}$$

where N is the number of census tracts indexed by u and v , X corresponds to the Composite Indicator (CI) score, and the \bar{X} corresponds to the average of the CI scores. CI is to be further explained in Sect. 2.2

After calculating the Moran's Index, the statistical significance test of spatial effects is performed. This test measures the model's pseudo significance (Anselin 1996).

2.2 Composite indicator of social vulnerability

Social vulnerability is a multidimensional social phenomenon that reflects the condition of a group of individuals on the margins of society. The condition of social vulnerability is not limited to the problem of the economic dimension. It includes dimensions that portray household conditions, quality of education, access to clean water and sanitation, and the environment (Davino et al. 2021). Multidimensional social phenomena such as social vulnerability are also characterized by the differentiation of weights of sub-indicators (Mazziotta and Pareto 2017; Libório et al. 2020a, b) and by the property of spatial dependence (Cutter and Finch 2008) that can vary according to or sub-indicator (Rufat et al. 2019).

The composite indicator of social vulnerability was built following these characteristics and aggregating the fifteen objective-quantitative³ sub-indicators of the five dimensions listed below:

- *Demographic* Dm-1. Residents up to one year old, Dm-2. Residents per household, Dm-3. Dependents per head of household, and Dm-4. Heads of households between 10 and 19 years old.
- *Economic* Ec-1. Heads of households without income, Ec-2. Heads of households with income of up to two minimum wages, and Ec-3. Heads of households with income above 20 minimum wages.
- *Educational* Ed-1. Heads of households without literacy and Ed-2. Illiterate residents aged between 10 and 14 years.
- *Environmental* En-1. Vegetation Coverage Index (Normalized Difference Vegetation Index).
- *Households* Hd-1. Households rented or leased, Hd-2. Households connected to the water network, Hd-3. Households connected to the sewage network, Hd-4. Households without a bathroom, and Hd-5. Households with more than three bathrooms.

The data from the fifteen sub-indicators refer to the last census carried out in Brazil in 2010 (IBGE, 2010). The data were collected and then normalized as

$$I_{lu} = \frac{x_{lu} - \min(x_l)}{\max(x_l) - \min(x_l)}, \quad (5)$$

³ The nature of a sub-indicator is objective-quantitative when the variable is directly measurable, objective-qualitative when the variables are objectively verifiable by the presence or absence of something even if not directly measurable, and qualitative-subjective when the variables are obtained from Experts opinion (Libório et al. 2022a).

where x_{lu} is the value of sub-indicator $l \in L$ of census tract $u \in U$.

Simple Additive Weight (SAW) offers a more flexible and straightforward way of implementing different weighting schemes than other methods applied to build composite indicators (El Gibari et al. 2019). The aggregation of the sub-indicators in the SAW is performed in a compensatory way. Compensatory aggregation allows sub-indicators of poor performance to be offset by sub-indicators of above-average performance (Ekel et al. 2020). This compensation between poor and above-average performance sub-indicators increases the composite indicator scores, reducing its capability to represent multidimensional social phenomena (Libório et al. 2022a). The formulas used in non-compensatory aggregation schemes involve normalization adjustments for the calculation of geometric or harmonic means (Cinelli et al. 2021), data standardization with the inclusion of penalties (Mazziotta and Pareto 2018), or the ordered aggregation of sub-indicators (Libório et al. 2021). In turn, the building of the composite indicator by the SAW method is quite simple, facilitating the implementation of the weighting scheme proposed in this research through the expression

$$CI = \sum_{l \in L} w_l I_{lu}, \quad (6)$$

where w_l is the weight of I_l so that:

$$\sum_{l \in L} w_l = 1 \quad (7)$$

and

$$0 \leq w_l \leq 1, l \in L. \quad (8)$$

2.3 Weighting of sub-indicators

Assigning weights by Data-driven weighting schemes avoids judgment errors and evaluation biases in the Expert opinion weighting scheme. In turn, assigning weights based on expert opinion prevents weights from disagreeing with the importance of sub-indicators in the concept of the multidimensional phenomenon. Finally, assigning Equal-weights avoids judgment errors and evaluation biases while ignoring the differentiation of the relative importance of indicators (Greco et al. 2019). This research analyzes how the advantages and disadvantages of these four weighting schemes are reflected in the spatial dependence property of the composite indicator. For this, the following methods were used:

- *Equal-weights* is an objective weighting scheme that assigns an equal w_i value to all x_i sub-indicators from the ratio $1/N$ where N is the number of sub-indicators.
- *Data-driven* is an objective weighting scheme that assigns weights that maximize the value of the Moran's Index, subject to constraints (7) and (8), through the Generalized Reduced Gradient algorithm (GRG; Abadie 1969). The GRG

algorithm is appropriate when the objective functions and constraints have exponential or trigonometric powers and calculations (Lasdon et al. 1974). The weighting of sub-indicators by the GRG was implemented in Microsoft Excel (Powell and Batt 2008).

- *Expert opinion* is a subjective weighting scheme that assigns weights considering the conceptual importance of the sub-indicator in the phenomenon, taking as a reference the opinion of a group of experts. The assignment of weights to the sub-indicators was performed using the Fuzzy extension of Saaty's (1988) Analytic Hierarchy Process, the so-called Fuzzy-AHP (Van Laarhoven and Pedrycz 1983). The Fuzzy-AHP allows the processing of evaluations in different formats, offering psychological comfort to specialists when evaluating alternatives (sub-indicators) and avoiding judgment errors (Ekel et al. 2020). Assessments were performed by four experts using four different assessment formats (see "Appendix" for details). The degree of consensus between individual and collective opinions was calculated and served as a basis to verify possible evaluation biases (Pedrycz et al. 2011).
- *Hybrid* is an objective-subjective weighting scheme that assigns weights by combining elements from the GRG and Fuzzy-AHP. This weighting scheme seeks an intermediate solution between the Data-driven and Expert opinion schemes. It assigns weights from the data without disregarding the sub-indicators' relative importance in the multidimensional phenomenon concept. The GRG algorithm was configured to maximize the value of the Moran's Index, taking as constraints the lowest and highest weight of the sub-indicator according to the group of experts (see Table 10 in "Appendix").

2.4 Robustness check

The robustness check consists of analyzing how the different input factors propagate through the composite indicator structure and impact the positions of the spatial units (Nardo et al. 2005). The uncertainty analysis makes it possible to verify the robustness of the composite indicator through the absolute variation of the difference in the positions of the census tracts N with each change in the input variables (Becker et al. 2017). The uncertainty analysis performed in this research measures the changes in the positions of the spatial units with each change in the weighting scheme through the following expression

$$\bar{R}_S = \frac{1}{N} \sum_{u \in U} \left| \text{Rank}_{\text{ref}}(\text{CI}_u) - \text{Rank}(\text{CI}_u) \right| \quad (9)$$

where $\text{Rank}_{\text{ref}}(\text{CI}_u)$ is the position occupied by the census sector score u in the composite indicator of reference CI, $\text{Rank}(\text{CI}_u)$ is the position occupied by the census sector score u in the composite indicator.

The uncertainty analysis reveals whether the classifications of spatial units are stable, providing a measure of the composite indicator's robustness and

capacity to represent the multidimensional phenomenon regardless of the variations in the weights of the sub-indicators (Saisana et al. 2005).

2.5 Quality check

Two quality checks were performed. The first quality check measures how well the composite indicator captures the main concept of the multidimensional phenomenon. This verification is performed through Pearson's r correlation coefficient between the composite indicator and the main individual indicator in the concept of the multidimensional phenomenon.

The individual indicator Average Income was used in the quality check for two reasons. First, because the average income is pointed out by many researchers as the main individual indicator of the concept of multidimensional social phenomena, such as social vulnerability, poverty, inequality, and social exclusion (Libório et al. 2022a). Second, because the individual indicator Average Income is not aggregated in the composite indicator of social vulnerability.

The second quality check measures how compatible the weights and relative importance of the sub-indicators in the concept of the multidimensional phenomenon are. This verification is carried out through the degree of consensus between the weights obtained by the different weighting schemes and the opinion of the experts, applying the following expression

$$C^p = 1 - \sum_{l=1}^m \frac{|O^G(x_k) - O^{E_i}(x_k)|}{m}, \quad p \in P \quad (10)$$

where $O^G(z_k)$ is the position of the k th alternative of the group's opinions and $O^{E_i}(z_k)$ is the k th position of the alternative of the expert E_i results, m is the number of experts, and n is the number of sub-indicators.

The decision-maker stipulates the threshold value of the degree of consensus (Ekel et al. 2020; Libório et al. 2022a, b, c). The weight of the sub-indicator is compatible with its relative importance in the concept of the multidimensional phenomenon when the consensus threshold is reached.

2.6 Consistency of results

The consistency analysis verifies how much the spatial weighting matrix influences the weights of the sub-indicators, the spatial dependence, and the robustness and quality parameters of the composite indicator. The analysis is performed using alternative spatial models to the base model. In particular, alternative models were created by changing the sensitivity parameter of distance b and recalculating the spatial weighting matrix

$$w_{uv} = d_{uv}^{-b}. \quad (11)$$

Table 1 The composite indicator's spatial dependence, robustness, and quality by weighting scheme

Parameter	Equal-weights	Expert opinion	Hybrid	Data-driven
Moran's index ^a	0.549	0.560	0.597	0.636
Uncertainty	29	27	19	23
Correlation with the average income	0.654	0.654	0.657	0.630
Degree of consensus	0.581	0.810	0.745	0.710

^aStatistical significance of spatial effects p value < 0.05

Model 2 was built by applying Eq. (2) with $b = 2$. Models 3 and 4 were built by applying Eq. (11). Model 3 uses $b = 1$, and Model 4 uses $b = 2$.

3 Effect of the sub-indicator weighting scheme on the spatial dependence property, robustness, and quality of the composite indicator

Table 1 shows the sub-indicator weighting scheme's effect on the composite indicator's spatial dependence property. The intensity of this effect is greater in the Data-driven and Hybrid weighting schemes. On average, the Moran's Index in these weighting schemes is thirteen and nine percent higher than in the Equal-weights and Expert opinion weighting schemes. In particular, weighting the sub-indicators by the Data-driven scheme allows obtaining a Moran's Index fourteen percent higher than the Equal-weights scheme.

The results in Table 1 show that the composite indicator of social vulnerability has a positive and significant spatial dependence, regardless of the weighting scheme. However, the spatial dependence of the composite indicator of social vulnerability is more prominent when the sub-indicators are weighted using the Hybrid and Data-driven schemes.

The composite indicators obtained by the four weighting schemes have an uncertainty of lower than 30 positions. This level of uncertainty means that changing the weights of the sub-indicators implies an unimportant variation in the positions of the spatial units. The composite indicator with the highest level of uncertainty was obtained using the Equal-weights weighting scheme. The uncertainty associated with the positions of the spatial units of the composite indicator built with Equal-weights is, on average, twenty-nine positions. The uncertainty associated with the weighting scheme is lower when the composite indicators are built using the Hybrid and Data-driven schemes. The uncertainty of the composite indicators obtained in the Hybrid and Data-driven schemes is twenty-five percent lower than in the composite indicators obtained in the Equal-weights and Expert opinion schemes. These results suggest that composite indicators built in Hybrid and Data-driven weighting schemes offer a robust and representative measure of the multidimensional phenomenon (Saisana et al. 2005).

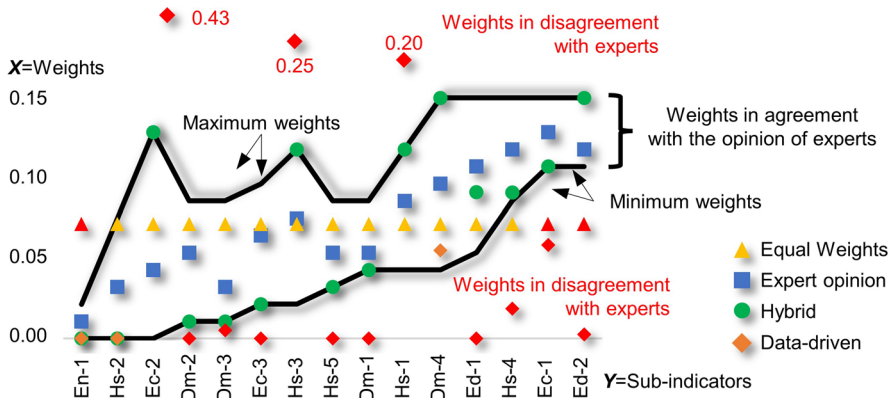


Fig. 2 Agreement between the weights of the sub-indicators according to the opinion of the experts and the weighting schemes (for sub-indicators descriptions see Table 2)

The correlation coefficients between the composite indicators and the individual average income indicator exceeded 0.60 in the four weighting schemes. These results indicate that the main individual indicator of the multidimensional phenomenon concept is captured independently of the weighting scheme.

The Expert opinion, Hybrid, and Data-driven weighting schemes generated composite indicators compatible with the expert group’s opinion. The composite indicators built through these weighting schemes converge with the expert group’s opinion above the threshold of 0.70. This high degree of consensus indicates that these composite indicators are compatible with the concept of the multidimensional phenomenon. These results reinforce the argument that the Equal-weights weighting scheme should be considered when there is no theory to justify weight differentiation and no consensus among experts (Greco et al. 2019).

However, several sub-indicators weighted through Equal-weights, and Data-driven weighting schemes have weights that exceed the upper and lower limits defined by expert opinion. Figure 2 reveals that the Equal-weights scheme has three sub-indicators with weights that exceed upper and lower limits defined by specialists. The number of sub-indicators that exceed the upper and lower limits rises to 12 in the Data-driven weighting scheme.

The Hybrid weighting scheme assigns weights within the maximum and minimum limits to two sub-indicators and weights equal to the minimum and maximum limits to five and eight sub-indicators, respectively.

The joint analysis of the Moran’s Index, uncertainty, correlation with average income, and degree of consensus suggests that Hybrid and Data-driven weighting allow obtaining a better composite indicator than composite indicators built by Equal-weights and, Expert opinion weighting schemes. Table 2 shows that four sub-indicators were decisive for the best performance of the Hybrid and Data-driven weighting in the analyzed parameters. These sub-indicators are identified based on the variance criterion (σ^2).

Table 2 Weights of sub-indicators by weighting scheme

Code	Sub-indicator	Equal weights	Expert opinion	Hybrid	Data-driven	σ^2
Dm-1	Residents per household	0.067	0.050	0.040	0.000	0.001
Dm-2	Heads of households between 10 and 19 years old	0.067	0.050	0.010	0.000	0.001
Dm-3	Dependents per head of household	0.067	0.030	0.010	0.005	0.001
Dm-4	Residents up to 1 year old	0.067	0.090	0.140	0.051	0.001
Ec-1	Heads of households without income	0.067	0.120	0.100	0.055	0.001
Ec-2	Heads of households with income above 20 min. wages	0.067	0.040	0.120	0.426	0.024
Ec-3	Heads of households with income up to 2 min wages	0.067	0.060	0.020	0.000	0.001
Ed-1	Illiterate residents aged between 10 and 14 years	0.067	0.100	0.085	0.000	0.001
Ed-2	Heads of households without literacy	0.067	0.110	0.140	0.002	0.003
Hs-5	Household without bathroom	0.067	0.080	0.110	0.195	0.003
Hs-5	Households with more than 3 bathrooms	0.067	0.030	0.000	0.000	0.001
Hs-5	Households connected to the sewage network	0.067	0.070	0.110	0.249	0.005
Hs-5	Households connected to water network	0.067	0.110	0.085	0.017	0.001
Hs-5	Households rented or leased	0.067	0.050	0.030	0.000	0.001
En-1	Vegetation coverage index	0.067	0.010	0.000	0.000	0.001

σ^2 is the variance of the sub-indicators weights in the four weighting schemes.

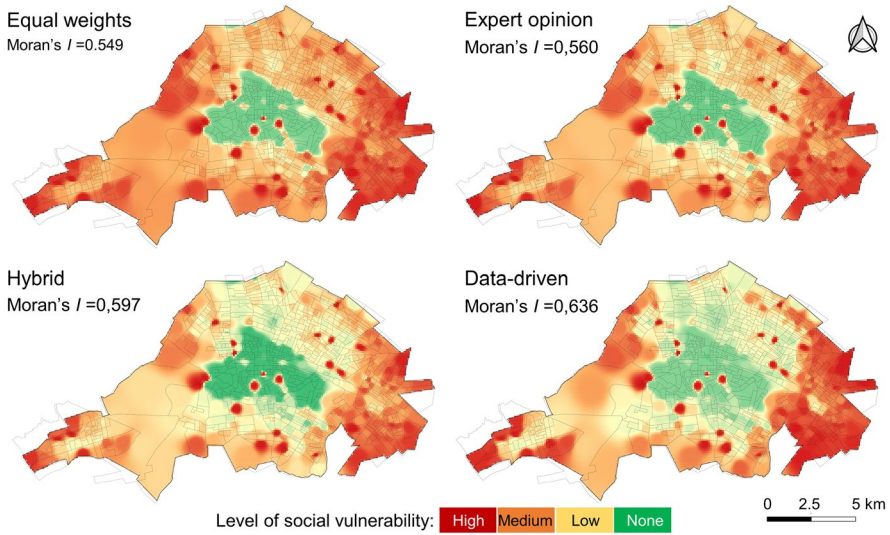


Fig. 3 Effect of sub-indicator weighting schemes on Social Vulnerability in the study area

The results show that the sub-indicators “Heads of households with income above 20 minimum wages” and “Households connected to the sewage network” increase the spatial autocorrelation of social vulnerability in the study area. The weight of the sub-indicator "Heads of households with income above 20 minimum wages" is 0.37 greater in the Data-driven weighting scheme than in the Equal-weights and Expert opinion weighting schemes. In the sub-indicator "Households connected to the sewage network," the weight obtained in the Data-driven weighting scheme is 0.17 greater than in the other weighting schemes. These results suggest that experts cannot consider the spatial dependence of sub-indicators in the weighting process. The weighting by Equal-weights ignores that the spatial dependence property is not equal between the sub-indicators.

In the Data-driven weighting scheme, the composite indicator with the highest spatial autocorrelation reduced the weight of the sub-indicators "Heads of households without literacy" and "Household without bathroom" by 0.10 and 0.11, respectively.

Table 3 Social vulnerability of census tracts by weighting scheme

Level of social vulnerability	None	Low	Medium	High	Average score
Equal-weights	126	8	268	259	0.38
Expert opinion	127	22	311	201	0.42
Hybrid	133	151	250	127	0.49
Data-driven	138	205	144	174	0.51

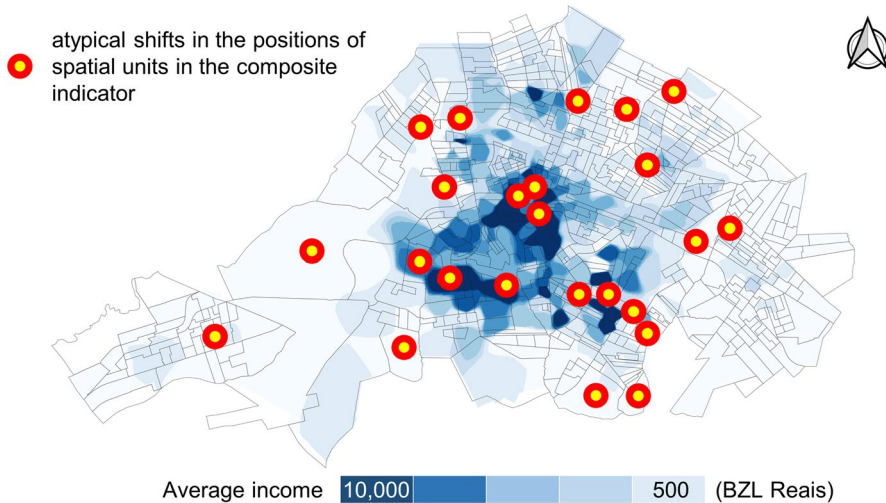


Fig. 4 Atypical shifts in the census tracts' position in the composite indicator ranking. *Note* The three z-score rule was applied to identify the atypical shifts corresponding to 27.6 positions

4 Influence of weighting schemes in the representation of social vulnerability

The influence of weighting schemes on the representation of social vulnerability in the study area can be seen in Fig. 3. There is a pattern in the distribution of social vulnerability regardless of the weighting scheme. The high social vulnerability is concentrated in the peripheral census tracts, especially in the eastern and western peripheries. Census tracts classified as having no social vulnerability are concentrated in the central region of the urban conurbation. However, it is possible to observe in the maps that the intensity of social vulnerability varies with the weighting scheme. The map shows that changes in the intensity of social vulnerability are accompanied by changes in the Moran's Index (Moran's I).

The Equal-weights and Expert opinion weighting schemes have a lower Moran's Index and a higher number of census tracts classified as high and medium social vulnerability. Table 3 shows how the Moran's Index relates to social vulnerability in the study area. First, it shows a positive relationship between the Moran's Index and the number of census tracts classified as having None social vulnerability. Second, it shows a negative relationship between the Moran's Index and the sum of census tracts classified as High and Medium social vulnerability. Third, it reveals that the Moran's Index is positively related to the average score of the composite indicator scores.

These results show that disregarding the opinion of experts through Data-driven weighting increases the spatial autocorrelation, the composite indicator scores, and the number of census tracts classified as Medium and High social vulnerability. At the same time, it reduces the number of census tracts classified as None and Low social vulnerability. The number of census tracts of None and Low social

vulnerability is 2.30 times and 1.21 times greater in the Data-driven weighting scheme than in the Expert opinion and Hybrid weighting schemes, respectively. The Data-driven weighting scheme presents 1.61 times and 1.19 times fewer census tracts classified as Medium and High social vulnerability concerning the Expert opinion and Hybrid weighting schemes.

These results indicate that positive spatial autocorrelation increases the composite indicator's average score due to the propagation of scores between census tracts. Therefore, disregarding the opinion of experts through Data-driven weighting can make it challenging to evaluate and formulate public policies to reduce social vulnerability, as this weighting scheme of sub-indicators tends to hide areas of Medium and High social vulnerability.

Furthermore, the composite indicator built by the Data-driven weighting scheme presents 26 census tracts with atypical positions. These census tracts show variations above 27.6 positions concerning the composite indicators built by the Hybrid, Equal-weights, and Expert opinion weighting schemes. The map in Fig. 4 shows that most of these atypical shifts occurred in census tracts located in the central area of the study area that presents a higher average income.

The census tracts highlighted with dots on the map shifted an average of 94.7 positions in the ranking of the composite indicator. The average income in these census tracts is 8% higher than the average for the study area. This result is associated with the weighting schemes' significant variance in the weights of the sub-indicator Heads of households with income above 20 minimum wages ($\sigma^2=0.024$). The weight of this sub-indicator in the Data-driven and Hybrid weighting schemes is, on average, 5.1 times greater than in the Equal-weights and Expert opinion weighting schemes. The sub-indicator Heads of households with income above 20 minimum wages have a greater weight in the weighting schemes that have a greater impact on the Moran's Index, suggesting a relationship between these two elements.

5 Relationship between the spatial autocorrelation of the sub-indicators and the composite indicator

The direct relationship between the Moran's Index and the average score of the composite indicator suggests that the Data-driven weighting scheme assigns greater weights to sub-indicators with greater spatial dependence. This weighting logic suggests that the spatial autocorrelation of the sub-indicator determines its weight. Therefore, conceptually important sub-indicators and low spatial autocorrelation may be underrepresented. Table 4 demystifies this assumption by showing a non-significant correlation between the weights obtained by the Data-driven weighting scheme and the Moran's Index of the sub-indicator.

These results allow us to reach a peculiar conclusion. The average Moran Index of the sub-indicators is 0.280 and is far below the Moran Index of the composite indicators. The Moran's Index of the composite indicator built by Data-driven weights is 1.96 times higher than the average Moran's Index of the sub-indicators. None of the sub-indicators presents greater spatial autocorrelation than the composite indicator built by Expert opinion, Hybrid, and Data-driven weighting. Furthermore, the

Table 4 Weights and Moran's Index of sub-indicators

Code	Sub-indicator	Data-driven	Weight order	Moran's Index	Index order
Dm-1	Residents per household	0.000	9	0.481	3
Dm-2	Heads of households between 10 and 19 years old	0.000	9	0.025	14
Dm-3	Dependents per head of household	0.005	7	0.409	6
Dm-4	Residents up to 1 year old	0.051	5	0.295	8
Ec-1	Heads of households without income	0.055	4	0.132	10
Ec-2	Heads of households with income above 20 min wages	0.426	1	0.552	2
Ec-3	Heads of households with income up to 2 min wages	0.000	9	0.466	5
Ed-1	Illiterate residents aged between 10 and 14 years	0.000	9	0.002	15
Ed-2	Heads of households without literacy	0.002	8	0.331	7
Hs-5	Household without bathroom	0.195	3	0.055	12
Hs-5	Households with more than three bathrooms	0.000	9	0.475	4
Hs-5	Households connected to the sewage network	0.249	2	0.555	1
Hs-5	Households connected to water network	0.017	6	0.028	13
Hs-5	Households rented or leased	0.000	9	0.259	9
En-1	Vegetation coverage index	0.000	9	0.131	11

$r = 0.36$ and p value = 0.06

Table 5 Relationship between the weighting scheme and Moran's Index in composite social indicators

Weighting scheme	Method	Scale	Index	Moran's <i>I</i>	Source
Data-driven	Principal component analysis	District	Vulnerability	0.437	Jha and Gundimeda (2019)
Equal weights	Simple additive weighting	Municipal	Poverty	0.296 and 0.363	Katumba et al. (2019)
Participatory	Analytical hierarchical process	Rural	Social vulnerability	0.207	(Chauhan et al. (2020)
Equal weights	Simple additive weighting	State	Financial inclusion	0.563	Adeleke and Alabede (2021)
Data-driven	Principal component analysis	District	Community resilience	0.437	Mavhura et al. (2021)
Data-driven	Principal component analysis	Country	Social vulnerability	0.490 and 0.510	Cutter and Finch (2008)

results show that emphasizing spatial dependence does not distort the representation of the multidimensional phenomenon.

The discrepancy of the Moran Index between social indicators is not new and has been reported in previous research (Rufat et al. 2019). However, this research results indicate that the composite indicator's spatial dependence property is not related to the Moran Index of its sub-indicators. The aggregation of sub-indicators has a positive and strong impact on the Moran Index of the composite indicator, regardless of the weighting scheme. However, the strength of this impact is higher in the Data-driven weighting scheme and lower in the Equal-weights and Expert opinion weighting schemes.

Table 5 provides evidence that reinforces this conclusion. Moran Indices of composite social indicators are normally higher when built by the Data-driven weighting scheme than when built by the Equal-weights and Participatory weighting schemes (see Adeleke and Alabede (2021) for exception).

The Hybrid weighting scheme developed in this research offers a solution that counterbalances the property of spatial dependence, the compatibility of the composite indicator with the multidimensional phenomenon, and the agreement of the weights of the sub-indicators with the opinion of experts. This solution makes it possible to know the spatial association of the multidimensional social phenomenon between neighboring areas without overestimating the conditions of these areas or weighing sub-indicators with weights that disagree with the opinion of experts.

6 Effect of the spatial weighting matrix on the spatial autocorrelation, robustness, and quality of the composite indicator

Composite indicators built by Hybrid and Data-driven weighting schemes, which emphasize spatial dependence, are more robust ($\bar{R}_S=19$ and $\bar{R}_S=23$). Quality checks indicate that these composite indicators provide good representations of social vulnerability in the study area. The correlation of these composite indicators with the individual average income indicator was $r=0.630$ in the Hybrid weighting and $r=0.657$ in the Data-driven weighting. The degree of consensus was $C^G=0.710$ in the Hybrid weighting and $C^G=0.745$ in the Data-driven weighting, surpassing the acceptance threshold of $C^G > 0.60$.

Tables 6, 7, 8, 9 shows that changes in the spatial weighting matrices do not change the relationships between the weighting schemes, spatial dependence and robustness, and quality parameters, indicating that the results are reliable. For example, the Moran's Index and uncertainty remain higher in composite indicators built with Hybrid and Data-driven weights.

These results show that the spatial weighting matrix does not change the relationship of the weighting scheme with the Moran's Index but shows that the effect of the weighting scheme on spatial autocorrelation is greater in spatial weighting matrices that are less sensitive to distance (Models 2 and 4).

Tables 7, 8, 9 shows that the spatial weighting matrix does not change the uncertainty and correlation of the composite indicator with the most important individual

Table 6 Moran's Index of the composite indicator per model

	Equal-weights	Expert opinion	Hybrid	Data-driven
<i>a</i>				
Model 2	0.73	0.73	0.75	0.77
Model 3	0.26	0.26	0.28	0.30
Model 4	0.62	0.62	0.64	0.66

Model 2 applies Eq. (11) with $b = 1$, Model 3 applies Eq. (2) with $b = 2$, and Model 4 applies Eq. (11) with $b = 2$

Table 7 Uncertainty of the composite indicator per model

	Equal-weights	Expert opinion	Hybrid	Data-driven
<i>b</i>				
Model 2	29	27	18	23
Model 3	29	27	19	24
Model 4	29	0.62	18	25

Model 2 applies Eq. (11) with $b = 1$, Model 3 applies Eq. (2) with $b = 2$, and Model 4 applies Eq. (11) with $b = 2$

Table 8 Degree of consensus of the composite indicator per model

	Equal-weights	Expert opinion	Hybrid	Data-driven
<i>c</i>				
Model 2	0.58	0.81	0.72	0.68
Model 3	0.58	0.81	0.75	0.70
Model 4	0.58	0.81	0.71	0.69

Model 2 applies Eq. (11) with $b = 1$, Model 3 applies Eq. (2) with $b = 2$, and Model 4 applies Eq. (11) with $b = 2$

indicator in the multidimensional phenomenon. Finally, Table 8 shows that the weighting matrix changes the degree of consensus regarding the weights of the sub-indicators in the Hybrid and Data-driven schemes. The degree of consensus is lower in models based on spatial weighting matrices that are more sensitive to distance. In models 2 and 4, the degree of consensus between the Data-driven and Expert group weights did not exceed the limit of $C^G = 0.70$.

On the one hand, the spatial weighting matrix influences the spatial autocorrelation of the composite indicator. On the other hand, the spatial weighting matrix does not change the effects of the weighting scheme on spatial autocorrelation, nor does it influence the robustness and quality parameters of the composite indicator.

Table 9 Correlation with the average income of the composite indicator per model

	Equal-weights	Expert opinion	Hybrid	Data-driven
<i>d</i>				
Model 2	0.654	0.654	0.657	0.623
Model 3	0.654	0.654	0.653	0.624
Model 4	0.654	0.654	0.652	0.627

Model 2 applies Eq. (11) with $b = 1$, Model 3 applies Eq. (2) with $b = 2$, and Model 4 applies Eq. (11) with $b = 2$

7 Conclusions

The results gathered in this research reveal that the property of spatial dependence of the composite indicator is influenced by the weighting scheme of the sub-indicators and the spatial weighting matrix. It is possible to build composite indicators that emphasize the spatial aspects of the neighborhood by applying Hybrid and Data-driven weighting schemes. Robustness and quality checks indicate that emphasizing the property of spatial dependence does not diminish the ability of composite indicators built through these weighting schemes to represent the multidimensional phenomenon.

The results offered in this research are limited to Moran's Global Index. Therefore, there are still numerous investigations to be carried out on the effects of sub-indicator weighting schemes on Moran's Local Index. In addition, the Data-driven weighting scheme developed in the research has the flexibility to deal with other spatial dimension problems. It is possible to configure the GRG algorithm for at least three other purposes. First, to maximize the spatial autocorrelation between two composite indicators. Second, maximize the correlation coefficient of a spatial regression model. Third, maximize the spatial autocorrelation based on Geary's (1954) Coefficient. These models can be configured to individually or simultaneously incorporate robustness and quality parameters. The flexibility of this weighting scheme also allows finding weights that maximize the spatial dependence of the composite indicator and, at the same time, ensure that the correlation of the composite indicator with the main individual indicator of the concept of the multidimensional phenomenon exceeds a certain threshold. For example, it is possible to emphasize the spatial dependence of the composite indicator and guarantee an $r=0.65$. The composite indicator obtained from the weighted sub-indicators with this GRG algorithm configuration allows reaching a Moran's Index equal to 0.635 and $r=0.650$. Finally, it is possible to adapt the Data-driven weighting scheme introduced to build composite indicators based on other methods. Future research may adapt the GRG algorithm to implement the Mazziotto-Pareto Index to consider no compensation between sub-indicators, rather than the Ordered Weighted Average operator that simultaneously considers no compensation between sub-indicators, spatial dependency, and spatial heterogeneity.

Appendix: weighting of sub-indicators by expert opinion

Selection of experts

Three criteria were used to select the experts. First, have a doctorate in human and social sciences, such as sociology and geography. Second, have a publication in scientific journals on social phenomena such as inequality, social exclusion, poverty, and social vulnerability. Third, know the social aspects of the study area.

The selection of experts from this set of criteria has advantages and disadvantages. On the one hand, this set of criteria limits the number of experts qualified to carry out the assessments. On the other hand, this set of criteria favors the quality and homogeneity of the evaluations.

Four experts were selected based on the virtual curriculum system created and maintained by the Brazilian National Council for Scientific and Technological Development. Different evaluation formats and the degree of consensus were adopted to avoid these experts' judgment errors and evaluation biases.

Assessment of alternatives

The assignment of weights by Expert opinion is commonly associated with the problem of judgment errors (Greco et al. 2019). One way to reduce judgment errors is to offer experts the possibility to choose the format for assessing alternatives (sub-indicators) that best suits them. Alternative assessment formats are not without limitations, but the flexibility of choosing the assessment format gives the expert psychological comfort that reduces judgment errors (Ekel et al. 2020).

- *Ordering of Alternatives (OA)* an array $O = \{o(x_1), o(x_2), \dots, o(x_k), \dots, o(x_n)\}$, where $o(x_k)$ is a permutation function that defines the position of the alternative x_k among the integer values $\{1, 2, \dots, k, \dots, n\}$.
- *Multiplicative Preference Relations (MR)* reflect the preference intensity ratio between the alternatives x_k and x_l , being understood as x_k is $m(x_k, x_l)$ times as good as x_l .
- *Utility Values (UV)* the preferences in X are given as a set of n UV: $U = \{u(x_1), u(x_2), \dots, u(x_k), \dots, u(x_n)\}$, where $u(x_k) \in [0, 1]$ represents the UV assigned to the alternative x_k .
- *Fuzzy Estimates (FE)* a fuzzy number that can be specified directly or through a linguistic variable in which the elements of X are directly assessed by experts using a set of estimates $L = \{l(x_1), l(x_2), \dots, l(x_k), \dots, l(x_n)\}$, where $l(x_k)$ is the FE associated with the alternative x_k .
- *Non-Reciprocal Fuzzy Preference Relations (FR)* indicate the degree to which the alternative x_k is, at least, as good as the alternative x_l , employing its membership function $0 \leq \mu(x_k, x_l) \leq 1$.

Table 10 Weights by experts, highest and lowest weights, and weights according to the group

Code	Sub-indicator	Exp1	Exp2	Exp3	Exp4	Group	Min	Max
Dm-1	Residents per household	0.08	0.06	0.04	0.04	0.05	0.04	0.08
Dm-2	Heads of households between 10 and 19 years old	0.08	0.07	0.01	0.04	0.05	0.01	0.08
Dm-3	Dependents per head of household	0.08	0.01	0.02	0.02	0.03	0.01	0.08
Dm-4	Residents up to 1 year old	0.07	0.04	0.11	0.14	0.09	0.04	0.14
Ec-1	Heads of households without income	0.10	0.11	0.13	0.14	0.12	0.10	0.14
Ec-2	Heads of households with income above 20 min. wages	0.00	0.01	0.12	0.04	0.04	0.00	0.12
Ec-3	Heads of households with income up to 2 min. wages	0.08	0.09	0.06	0.02	0.06	0.02	0.09
Ed-1	Illiterate residents aged between 10 and 14 years	0.10	0.10	0.05	0.14	0.10	0.05	0.14
Ed-2	Heads of households without literacy	0.10	0.11	0.10	0.14	0.11	0.10	0.14
Hs-5	Household without bathroom	0.10	0.11	0.08	0.04	0.08	0.04	0.11
Hs-5	Households with more than 3 bathrooms	0.00	0.00	0.07	0.04	0.03	0.00	0.07
Hs-5	Households connected to the sewage network	0.08	0.11	0.09	0.02	0.07	0.02	0.11
Hs-5	Households connected to water network	0.08	0.11	0.10	0.14	0.11	0.08	0.14
Hs-5	Households rented or leased	0.08	0.05	0.03	0.04	0.05	0.03	0.08
En-1	Vegetation coverage index	0.01	0.02	0.00	0.00	0.01	0.00	0.02

The experts assessed the weights of the sub-indicators through the Utility Values (two experts), Ordering of Alternatives, and Multiplicative Preference Relations evaluation formats. The assessments were standardized in the Non-Reciprocal Fuzzy Preference Relations format using the preference format transformation functions:

$$FR(x_k, x_l) = \begin{cases} \frac{1}{2} + \frac{OA_l - OA_k}{2(n-1)} & \text{if } OA_k > OA_l \\ 1 & \text{if } A_k \leq OA_l \end{cases} \tag{13}$$

for the transformation $OA \rightarrow FR$;

$$FR(x_k, x_l) = \begin{cases} 1 + \frac{1}{2} \log_m \frac{MR(x_k, x_l)}{MR(x_l, x_k)} & \text{if } \log_m M(x_k, x_l) < 0 \\ 1 & \text{if } \log_m M(x_k, x_l) \geq 0 \end{cases} \tag{14}$$

for the transformation $MR \rightarrow FR$;

$$FR(x_k, x_l) = \begin{cases} \frac{UV(x_k)}{UV(x_l)} & \text{if } UV(x_k) < UV(x_l) \\ 1 & \text{if } UV(x_k) \geq UV(x_l) \end{cases} \tag{15}$$

for the transformation $UV \rightarrow FR$.

Table 11 Expert and group degree of consensus

Code	Sub-indicator	Exp1	Exp2	Exp3	Exp4	Group
Dm-1	Residents per household	0.71	1.00	0.86	0.79	0.84
Dm-2	Heads of households between 10 and 19 years old	0.71	0.93	0.64	0.79	0.77
Dm-3	Dependents per head of household	0.43	1.00	1.00	0.93	0.84
Dm-4	Residents up to 1 year old	0.50	0.57	0.86	0.71	0.66
Ec-1	Heads of households without income	1.00	1.00	1.00	1.00	1.00
Ec-2	Heads of households with income above 20 min wages	0.86	0.93	0.29	0.57	0.66
Ec-3	Heads of households with income up to 2 min wages	0.79	0.93	0.93	0.71	0.84
Ed-1	Illiterate residents aged between 10 and 14 years	0.79	0.86	0.57	0.79	0.75
Ed-2	Heads of households without literacy	0.93	0.93	0.86	0.93	0.91
Hs-5	Household without bathroom	0.64	0.64	0.93	1.00	0.80
Hs-5	Households with more than 3 bathrooms	0.93	0.86	0.64	0.50	0.73
Hs-5	Households connected to the sewage network	0.86	0.57	0.93	0.64	0.75
Hs-5	Households connected to water network	0.79	0.93	0.86	0.93	0.88
Hs-5	Households rented or leased	0.71	0.93	0.79	0.79	0.80
En-1	Vegetation coverage index	0.86	0.79	1.00	1.00	0.91
	Expert and group degree of consensus	0.76	0.86	0.80	0.79	0.81

Applying the transformation functions is to obtain the evaluations in a single format. Then, it is possible to calculate the weights of the sub-indicators according to the group of experts, applying the following expression:

$$F_p^c(x_k) = \sum_{y=1}^v w_y F_p^y(x_k) \tag{16}$$

where $x_k \in X$, $0 \leq w_y \leq 1$, for $y = 1, 2, \dots, v$, taking into account the condition $\sum_{y=1}^v w_y = 1$.

Table 10 shows the weights assigned by each expert, the highest and lowest weights per sub-indicator, and the weights according to the group of experts.

Individual and group degree of consensus

Another common problem of weighting by Expert opinion is associated with evaluation bias. This problem is especially relevant when the number of experts is small because biased assessments more strongly influence the results (Musa et al. 2019).

The degree of individual consensus measures how much an expert opinion diverges from the collective opinion, signaling possible evaluation biases. The individual degree of consensus is obtained applying the expression:

$$C^I = \sum_{l \in L}^m \left(1 - \frac{|O^G(z_k) - O^{E_l}(z_k)|}{n-1} \right) \quad (17)$$

Evaluation biases can be identified by considering the degree of consensus between the expert's assessments and the expert group's assessment. The acceptance threshold of the degree of consensus is a subjective measure defined by the decision-maker (Ekel et al. 2020). When the acceptance threshold of the degree of consensus is not reached, the decision-maker may request the reassessment of the alternatives or disregard the expert's assessments (Pedrycz et al. 2011). After identifying and eliminating evaluation biases, it is possible to calculate the degree of consensus of the group by the following expression:

$$C^G = \frac{\sum_{k=1}^n C^I}{n} \quad (18)$$

The research adopts the threshold of 0.70 as the individual and group degree of consensus. Table 11 shows that the four experts reached the adopted degree of consensus threshold.

Funding This work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel—Brazil (CAPES)—Financing Code 0001 and the National Council for Scientific and Technological Development of Brazil (CNPq)—Productivity Grant, Grant 311922/2021-0 and 151518/2022-0.

Data availability Martinuci, Oseias da Silva; Machado, Alexei Manso Correa; Libório, Matheus Pereira (2021), “Data for: Time-in-space analysis of multidimensional phenomena”, Mendeley Data, V4, <https://doi.org/10.17632/m3y4jncvch.4>.

Declaration

Human and/or animal rights No human participants and/or animals are involved in this research.

References

- Abadie J (1969) Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints. *Optimization*, 37–47
- Adeleke R, Alabede O (2021) Understanding the patterns and correlates of financial inclusion in Nigeria. *GeoJournal* 87:1–18
- Aljoufie M, Tiwari A (2020) Exploring housing and transportation affordability in Jeddah. *Housing Policy Debate*, 1–27
- Anselin L (1996) The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In: Fischer M, Scholten HJ, Unwin D (eds) *Spatial analytical perspectives on GIS in environmental and socioeconomic sciences*. Taylor and Francis, London, pp 111–125
- Badea AC, Tarantola S, Bolado R (2011) Composite indicators for security of energy supply using ordered weighted averaging. *Reliab Eng Syst Saf* 96(6):651–662
- Becker W, Paruolo P, Saisana M, Saltelli A (2017) Weights and importance in composite indicators: mind the gap. *Handbook of uncertainty quantification*, pp 1187–1216

- Bernardes P, Ekel PI, Rezende SF, Pereira Júnior JG, dos Santos AC, da Costa MA, Carvalhais RL, Libório MP (2021) Cost of doing business index in Latin America. *Qual Quant* 56:1–20
- Carley S, Evans TP, Graff M, Konisky DM (2018) A framework for evaluating geographic disparities in energy transition vulnerability. *Nat Energy* 3(8):621–627
- Cartone A, Postiglione P (2021) Principal component analysis for geographical data: the role of spatial effects in the definition of composite indicators. *Spat Econ Anal* 16(2):126–147
- Chauhan N, Shukla R, Joshi PK (2020) Assessing impact of varied social and ecological conditions on inherent vulnerability of Himalayan agriculture communities. *Hum Ecol Risk Assess Int J* 26(10):2628–2645
- Cinelli M, Spada M, Kim W, Zhang Y, Burgherr P (2021) MCDA Index Tool: An Interactive Software To Develop Indices And Rankings. *Environ Syst Decis* 41(1):82–109
- Cutter SL, Finch C (2008) Temporal and spatial changes in social vulnerability to natural hazards. *Proc Natl Acad Sci* 105(7):2301–2306
- Davino C, Gherghi M, Sorana S, Vistocco D (2021) Measuring social vulnerability in an urban space through multivariate methods and models. *Soc Indic Res* 157(3):1179–1201
- Dialga I, Le Giang TH (2017) Highlighting methodological limitations in the steps of composite indicators construction. *Soc Indic Res* 131(2):441–465
- Ekel P, Pedrycz W, Pereira J Jr (2020) Multicriteria decision-making under conditions of uncertainty: a fuzzy set perspective. Wiley, Chichester
- Ekel P, Bernardes P, Vale GMV, Libório MP (2022) South American business environment cost index: reforms for Brazil. *Int J Bus Environ* 13(2):212–233
- El Gibari S, Gómez T, Ruiz F (2019) Building composite indicators using multicriteria methods: a review. *J Bus Econ* 89(1):1–24
- Fusco E, Vidoli F, Sahoo BK (2018) Spatial heterogeneity in composite indicator: a methodological proposal. *Omega* 77:1–14
- Geary RC (1954) The contiguity ratio and statistical mapping. *Inc Stat* 5(3):115–146
- Getis A (1999) Spatial statistics. *Geogr Inf Syst* 1:239–251
- Goodchild MF (1991) Geographic information systems. *Prog Hum Geogr* 15(2):194–200
- Greco S, Ishizaka A, Tasiou M, Torrisi G (2019) On the methodological framework of composite indices: a review of the issues of weighting, aggregation, and robustness. *Soc Indic Res* 141(1):61–94
- Harris P, Clarke A, Juggins S, Brunson D, Charlton M (2015) Enhancements to a geographically weighted principal component analysis in the context of an application to an environmental data set. *Geogr Anal* 47(2):146–172
- IBGE (2010) Censo demográfico 2010. <https://censo2010.ibge.gov.br/>. Accessed 23 Dec 2022
- Jha RK, Gundimeda H (2019) An integrated assessment of vulnerability to floods using composite index—a district level analysis for Bihar, India. *Int J Disaster Risk Reduct* 35:101074
- Kallio M, Guillaume JH, Kumm M, Virrantaus K (2018) Spatial variation in seasonal water poverty index for Laos: an application of geographically weighted principal component analysis. *Soc Indic Res* 140(3):1131–1157
- Katumba S, Cheruiyot K, Mushongera D (2019) Spatial change in the concentration of multidimensional poverty in Gauteng, South Africa: evidence from quality of life survey data. *Soc Indic Res* 145(1):95–115
- Kuc-Czarnecka M, Lo Piano S, Saltelli A (2020) Quantitative storytelling in the making of a composite indicator. *Soc Indic Res* 149(3):775–802
- Lasdon LS, Fox RL, Ratner MW (1974) Nonlinear optimization using the generalized reduced gradient method. *Revue française d'automatique, informatique, recherche opérationnelle. Recherche Opérationnelle* 8(V3):73–103
- Libório MP, Martinuci ODS, Ekel PI, Hadad RM, Lyrio RDM, Bernardes P (2021) Measuring inequality through a non-compensatory approach. *GeoJournal* 87:1–18
- Libório MP, Laudares S, Abreu JFD, Ekel PY, Bernardes P (2020a) Property tax: dealing spatially with economic, social, and political challenges. *Urbe. Revista Brasileira de Gestão Urbana*, p 12
- Libório MP, Martinuci ODS, Laudares S, Lyrio RDM, Machado AMC, Bernardes P, Ekel P (2020b) Measuring intra-urban inequality with structural equation modeling: a theory-grounded indicator. *Sustainability* 12(20):8610
- Libório M, Abreu JF, Martinuci ODS, Ekel PI, Lyrio RDM, Camacho VAL, Melazzo ES (2022a) Uncertainty analysis applied to the representation of multidimensional social phenomena. *Papers in Applied Geography*, 1–24

- Libório MP, Ekel PY, Martinuci ODS, Figueiredo LR, Hadad RM, Lyrio RDM, Bernardes P (2022b) Fuzzy set based intra-urban inequality indicator. *Qual Quant* 56(2):667–687
- Libório MP, Martinuci ODS, Machado AMC, Ekel PI, Abreu JFD, Laudares S (2022c) Representing multidimensional phenomena of geographic interest: benefit of the doubt or principal component analysis?. *The Professional Geographer*, pp 1–14
- Lindén D, Cinelli M, Spada M, Becker W, Gasser P, Burgherr P (2021) A framework based on statistical analysis and stakeholders' preferences to inform weighting in composite indicators. *Environ Model Softw* 145:105208
- Maricic M, Egea JA, Jeremic V (2019) A hybrid enhanced Scatter Search—Composite I-Distance Indicator (eSS-CIDI) optimization approach for determining weights within composite indicators. *Soc Indic Res* 144(2):497–537
- Marzi S, Mysiak J, Essenfelder AH, Amadio M, Giove S, Fekete A (2019) Constructing a comprehensive disaster resilience index: the case of Italy. *PLoS ONE* 14(9):e0221585
- Mavhura E, Manyangadze T, Aryal KR (2021) A composite inherent resilience index for Zimbabwe: an adaptation of the disaster resilience of place model. *Int J Disaster Risk Reduct* 57:102152
- Mazziotta M, Pareto A (2017) Synthesis of indicators: the composite indicators approach. In: *Complexity in society: from indicators construction to their synthesis*. Springer, Cham, pp 159–191
- Mazziotta M, Pareto A (2018) Measuring well-being over time: the adjusted Mazziotta-Pareto index versus other non-compensatory indices. *Soc Indic Res* 136(3):967–976
- Moran PA (1950) Notes on continuous stochastic phenomena. *Biometrika* 37(1/2):17–23
- Musa HD, Yacob MR, Abdullah AM (2019) Delphi exploration of subjective well-being indicators for strategic urban planning towards sustainable development in Malaysia. *J Urban Manag* 8(1):28–41
- Nardo M, Saisana M, Saltelli A, Tarantola S (2005) Tools for composite indicators building. *Eur Com Ispra* 15(1):19–20
- Otoiú A, Pareto A, Grimaccia E, Mazziotta M, Terzi S (2021) Open issues in composite indicators. A starting point and a reference on some state-of-the-art issues, vol 3. Roma TrE-Press, Rome
- Pedrycz W, Ekel P, Parreiras R (2011) *Fuzzy multicriteria decision-making: models, methods and applications*. Wiley, Chichester
- Powell SG, Batt RJ (2008) *Modeling for insight*. Wiley, Hoboken
- Rufat S, Tate E, Emrich CT, Antolini F (2019) How valid are social vulnerability models? *Ann Am Assoc Geogr* 109(4):1131–1153
- Saaty TL (1988) What is the analytic hierarchy process?. In: *Mathematical models for decision support*. Springer, Berlin, Heidelberg, pp 109–121
- Saisana M, Tarantola S (2002) State-of-the-art report on current methodologies and practices for composite indicator development, vol 214. Ispra: European Commission, Joint Research Centre, Institute for the Protection and the Security of the Citizen, Technological and Economic Risk Management Unit
- Saisana M, Saltelli A, Tarantola S (2005) Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *J R Stat Soc A Stat Soc* 168(2):307–323
- Saltelli A (2007) Composite indicators between analysis and advocacy. *Soc Indic Res* 81(1):65–77
- Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 46(sup1):234–240
- Van Laarhoven PJ, Pedrycz W (1983) A fuzzy extension of Saaty's priority theory. *Fuzzy Sets Syst* 11(1–3):229–241

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Matheus Pereira Libório¹  · **João Francisco de Abreu¹**  ·
Petr Iakovlevitch Ekel^{1,2}  · **Alexei Manso Correa Machado^{1,2}** 

João Francisco de Abreu
jofabreu@hotmail.com

Petr Iakovlevitch Ekel
petr.ekel2709@gmail.com

Alexei Manso Correa Machado
alexeimcmachado@gmail.com

¹ Pontifical Catholic University of Minas Gerais, Belo Horizonte 30535-012, Brazil

² Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil