

# Article **Problems of Aggregation of Sustainable Development Indicators at the Regional Level**

Ludmila Petkovová <sup>1,3</sup>\*<sup>(D)</sup>, David Hartman <sup>2,3,4</sup><sup>(D)</sup> and Tomáš Pavelka <sup>5</sup>

- <sup>1</sup> Department of Economics and Management, Unicorn University, 130 00 Prague, Czech Republic
- <sup>2</sup> Department of Information Technology, Unicorn University, 130 00 Prague, Czech Republic; hartman@cs.cas.cz
- <sup>3</sup> TREXIMA, spol. s r.o., 763 02 Zlín, Czech Republic
- <sup>4</sup> Institute of Computer Science of the Czech Academy of Sciences, 182 07 Prague, Czech Republic
- <sup>5</sup> Department of Managerial Economics, University of Economics in Prague, 130 67 Prague, Czech Republic; tomas.pavelka@vse.cz
- \* Correspondence: ludmila.petkovova@unicornuniversity.net

Received: 31 July 2020; Accepted: 30 August 2020; Published: 2 September 2020



**Abstract:** The aim of the paper is to evaluate the possibilities of evaluating sustainable development in regions based on the 2030 Agenda, and in particular to identify issues that need more attention. Our interest is focused on issues with compiling CIs for a small number of regions with limited available data. The article offers a critical discussion of various methods of aggregation, which are derived from the respective level of compensability, robustness, and sensitivity of the results.

Keywords: sustainable development; Agenda 2030; NUTS 3 regions; composite indicators

JEL Classification: Q01; Q56; R11; C10; C43

# 1. Introduction

Sustainable development has the worldwide attention of researchers and of the general public. An important milestone was the United Nations Sustainable Development Summit 2015 in New York. The UN adopted 17 Sustainable Development Goals, which followed on from the so-called Millennium Development Goals, which focused on the problems of developing countries [1]. These goals cover many areas of human activity, including fiscal systems, governmental processes, infrastructure as well as agricultural or ecological aspects that together build a complicated puzzle of sustainable development. There are various methods to evaluate sustainable development. Building a composite indicator (CI) is probably the most common and popular method of assessing sustainability, see e.g., [2–4]. The goals of the Agenda 2030 are regularly evaluated based on 17 composite indicators, which are evaluated for most of the UN countries. Simply put, a composite indicator is a simple variable evaluating the quality of a particular area of sustainable development via aggregating several components, often being composite indicators themselves., This approach facilitates the final evaluation because only a limited number of indicators are taken into account in the final assessment. On the other hand, we need to perform the relatively sensitive step of determining weights and aggregation rules [5]. The problem of the in-comparability of various composite indicators arises not only because of the potentially improper selection of the underlying components, but also because of the methodology used for their aggregation and further analysis. Despite the disadvantages caused by a potentially insufficient CI design or an inaccurate determination of subjective weights [6], the advantages of CI use usually lead to the adoption of this method. One of the most important benefits is the ability to summarize different perspectives of different criteria into a single variable or its ability to reduce the number of indicators [6,7].



Our goal is to assess the possibilities of assessing sustainable development at the regional level in a country with a small number of regions on the basis of Agenda 2030 and to identify the issues that require more attention. Such an assessment should provide analysts with a guidance on how to address specific issues that may arise in the process when applying this specific analytical approach. To assess these possibilities, we study a potential pipeline to determine an effective evaluation of sustainable development by using the components and a decomposition of goals from Agenda 2030 while also making use of various combinations of aggregation and corresponding processing methods.

The major goal of this work, therefore, is to identify potential issues in the whole process and to assess the applicability of sustainable development evaluations using parameters from Agenda 2030 for the regional level for a country with a small number of regions. This goal can be approached by providing answers to the following interlinked questions. 1. Can the goals of Agenda 2030 be assessed at the regional level using composite indicators? This question can be rephrased and disassembled into questions touching upon the efficiency of the decomposition in Agenda 2030, e.g., of goals into indicators. An inappropriate dependency structure between pairs of indicators can cause a non-robust character of composite indicators for the corresponding goal. The same problem may apply on the level of goals when combining them into key areas. 2. What problems in the data need to be considered when evaluating a small number of regions? Is it possible to solve these problems, and, if so, by what methods? Let us also note that despite us being convinced about the correct decomposition of the composite indicators into corresponding components, we are well aware of the various issues that arise around the data, e.g., the variability or existence of outliers.

The article is divided into five main parts. The first part presented in the section Theoretical background concerns the theoretical background of a sustainable development assessment. The second part represented by the section Materials and methods describes the principles and issues of the corresponding methodology. The third part in the section Results presents the results when applying various combinations of methods to evaluate sustainable development for a chosen country with a small number of regions. This section also presents details of the corresponding dataset. The fourth part in the section Discussion contains a discussion on the obtained results and suggestions for further research, while the last part in the section Conclusions presents the conclusion.

#### 2. Theoretical Background

The first references to the definition of sustainable development are linked to the works of Brundtland [8] or the team of Donella and Denis Meadows and Randers [9]. Some authors, however, put the origin of sustainable development back in the 19th century [10], when there was first observed an interest in the mutual interplay of the limitations of natural resources and their best use. One of the major goals in this area is to evaluate the quality of sustainable development to assess the development in a particular region or to compare developments in different regions. This task is complicated by the fact that sustainable development is undoubtedly a multidimensional phenomenon not to mention the potentially different characters of each dimension. Due to its complicated nature, a common definition does not exist, see [11,12]. Unfortunately, this complicates the comparability of results given by different authors. The most commonly accepted definition is certainly the definition introduced in [8]. The work says that: "Sustainable development is a development that meets the needs of the present without compromising the ability of future generations to meet their own needs", but many other definitions can be found, see [13,14]. In 2008, the Commission for Measuring Economic Performance and Social Progress (OECD), known as the Stiglitz Commission, was set up. This commission has run a large study concerning sustainability. In the final report, the authors of this study provide a comprehensive overview of alternatives in measuring and evaluating sustainable development. The report points out that future generations may be sensitive to the relative lack of natural goods, which are of little importance today. This requires that more attention be paid to these goods immediately. However, the question is to what extent today's society can determine what the goods are [15]. In 2015, 17 sustainable development goals were adopted: No poverty,

3 of 20

Zero hunger, Good health and well-being, Quality education, Gender equality, Clean water and sanitation, Affordable and clean energy, Decent work and economic growth, Industry, innovation and infrastructure, Reduced inequalities, Sustainable cities and communities, Responsible consumption and production, Climate action, Life below water, Life on land, Peace, justice and strong institutions and Partnership for the Goals [1].

## Evaluating Sustainable Development

Sustainable development is commonly described as an effort towards creating a balance between the economic, social and environmental pillars. When searching for currently existing CIs that cover at least a part of sustainable development issues, there are many examples all around the world. The most well-known are probably the Index of Sustainable Economic Welfare (ISEW), see [16,17], the Sustainable Dashboard published by Joint Research Centre, both of which cover all three pillars, the Environmental Performance Index (EPI) published by Yale University, the Human Development Index (HDI), the Genuine Saving Index (GSI), the Ecological Footprint (published by World Wild Fund) and many others, see their description in [4]. Some of such CIs tend to focus more on one individual pillar of sustainability, e.g., the Ecological Footprint on the environmental pillar, or HDI on the social pillar, as explained in [10]. CIs at the regional level are less common, see e.g., [18].

In recent years, composite indicators have been responding to the emergence of Agenda 2030. Such work is, for example, SDG composite indicators from the Sustainable Development Solutions Network [19]. It is a system of CIs pursuing all the objectives of the 2030 Agenda [1,19,20]. In 2019, the Joint Research Center (JRC) audited the SDGs [21]. As part of this audit, minor changes were recommended concerning the structure, and especially the methods of aggregation. SDGs are aggregated using only an arithmetic mean, which is a completely compensatory technique. The audit, as well as other work, recommends the use of non-compensatory aggregation techniques.

#### 3. Materials and Methods

#### 3.1. Data

For our analysis, we selected data from the Czech Republic, which is a small country with a low number of regions at the NUTS3 level. Nevertheless, there are data available for this level.

We use the following mechanism for CI design operating on two levels. On the global level, we have several key areas that are further subdivided into goals. For each goal, we need to identify a set of corresponding indicators that contribute to its state. To define CI for a key area we need to perform two consecutive aggregations: (1) an aggregation of indicators into a goal which combines the contributions of selected indicators, and (2) an aggregation of goals into key areas.

We use the Strategic Framework Czech Republic 2030 [22] to compile the composite indicator framework. In this strategy, there are 6 *key areas* defined within the framework of sustainable development: People and Society, Economy, Resilient Ecosystems, Municipalities and Regions, Global Development, and Good Governance. Each key area is further divided into goals. Each such *goal* represents a particular qualitative phenomena in a corresponding key area. These goals were designed based on goals defined in Agenda 2030 [1].

We have selected only the areas of People and Society, the Economy, Resilient Ecosystems, and Good Governance for the regional level analysis. However, the last area (Good Governance) coincides with the goals of the People and Society area at the level of regional evaluations. This area has therefore been incorporated as one of the goals into the area of People and Society. This preprocessing results in three areas that approach the classical concept of sustainable development through three pillars: economic, social, and environmental. For this reason, we further label them ECO, SOC, and ENV for clarity.

The search for suitable indicators was based on the recommendations of the Strategic Framework Czech Republic 2030, previous works of the Czech Statistical Office [23] and the recommendations

in [24–26]. In the case of some goals, it was not possible at the NUTS3 regional level to fill the indicators with data that would describe the given goal well. Therefore, within the People and Society area, two goals were completely removed, namely "social climate is universally favorable to families; barriers and social pressures have been minimized; family, parenthood and marriage are covered by special legal protection and are highly valued in society" and "greater public investment supports key cultural functions and equal access to culture and creativity". Apart from a simple removal, there were situations in which goals were merged to avoid a lack of data. In particular, within the area of Resilient Ecosystems we merged two goals "The Landscape of the Czech Republic is conceived as a complex ecosystem and ecosystem services provide a suitable framework for the development of human society", and "the Czech landscape is diverse and biological diversity is being restored".

Further preliminary analysis was done based on the evaluation of the mutual correlation between indicators. An example of such modification is represented by the goal "Natural resources are used as efficiently and economically as possible to minimize the external costs caused by their consumption" in the Economy area. This goal has a negative correlation with other goals in this area which may cause problems in a second level aggregation. At the same time, this goal has positive correlations with goals in the Resilient Ecosystems area. For these reasons, we move this goal into this area. This process results in 55 indicators pursuing 12 goals in three areas. The list of indicators, goals and areas is in Tables A1–A3.

#### 3.2. Normalization

We assume that we have several *indicators* stored in the variable  $X_{qr}$  representing the value of the indicator q for the region r, where q = 1, ..., Q and r = 1, ..., R. A common situation is that sustainable development indicators are in different units of measurement. Therefore, data normalization is required prior to the aggregation step. Data normalization aims at adjusting different measurement units and different ranges of variation. There are several ways to perform such operations. One popular way of handling this that is also strong enough is the *min-max method*. This method takes the original data represented by variable  $X_{qr}$  and produces the corresponding *normalized indicator* variable  $I_{qr}$ . For indicators with a positive direction (higher values mean better performance) the following formula is used:

$$I_{qr} = \frac{X_{qr} - \min_r(x_q)}{\max_r(x_{qr}) - \min_r(x_{qr})},\tag{1}$$

on the other hand, indicators with a negative direction (higher values mean worse performance) are normalized according to:

$$I_{qr} = \frac{\max_{r}(x_{qr}) - X_{qr}}{\max_{r}(x_{qr}) - \min_{r}(x_{qr})},$$
(2)

This method is very popular and has been used for the construction of many CIs. The most known is the Human Development Index (HDI) by the United Nations [27] and SDGs. The min-max method rescales the data into various intervals based on minimum and maximum values. The output is dimensionless but the relative distances remain. The normalized data range from 0 to 1 and have the same positive direction (a higher value means better performance and vice versa). Each indicator reaches a value between 0 and 1 even though there can be an extreme value. Since this method depends on the minimum and the maximum value, outliers—if they appear—produce a strong impact on the final output. However, compared to, for example, the z-score, this impact is weaker.

#### 3.3. Weights in an Aggregation Step

During the aggregation step, indicators as well as goals can be weighted using different methods, but the simplest approach is to assign an equal weight for each indicator:

$$w_q = \frac{1}{Q'} \tag{3}$$

where  $w_q$  is the weight for *q*th indicator (q = 1, ..., Q). This means that weights of indicators in one goal are the same for all indicators as well as regions. The main strength of the method is simplicity. Nevertheless, there is a risk that indicators of a goal with a low number of indicators will have a higher impact on the final CI. An equal weighting may be justified when there is no clear interpretability of unequal weighting. In general, there are two types of methods for setting weights. The first class of methods is represented by participatory approaches such as expert-based weights, public opinion, conjoint analysis, etc. Weights in these approaches find their base in subjective value judgments. The second type is represented by statistical methods—e.g., the principal component analysis/factor analysis, the benefit of doubt, the unobserved components regression. Weights in these cases are derived directly from data (for discussion about weights derivation see in [6]). However, using any kind of weights may reduce the interpretability and makes the results less understandable for the

observations (regions). The possibilities of using multidimensional methods are therefore very limited. Another method of the weights determination and aggregation is the Benefit of Doubt (BoD). The determination of weights using BoD is based on Data Envelopment Analysis models (DEA) [6]. The DEA method is a widely used tool to assess technical efficiency and performance of decision making. The basic component of such a model is represented by decision making units (DMUs) that constitute a decision hierarchy that is further evaluated to optimize a chosen criterion. In our case these DMUs represent individual NUTS 3 regions. For more detailed information see [6].

public. The problem in our application is the size of the dataset, which has a very low number of

Using DEA, each region gains its own weights that maximize (or minimize) the impact of the criteria where the region performs relatively well (or poorly) compared to the others. The DEA analysis also provides information about the necessary amount of the reduction of an inefficient region's inputs to become efficient, but this is not the subject of this paper, and regarding the nature of the used data, an interpretation would be very difficult.

BoD model itself leads to maximization problems of linear programming

Maximize 
$$I_r = \sum_{q=1}^{Q} w_{qr} I_{qr}$$
  
Under conditions  $\sum_{q=1}^{Q} w_{qr} I_{qr} \le 1$   
 $w_{qr} \ge 0$  (4)

where  $w_{qr}$  is the weight and  $I_{qr}$  is the value of the indicator q (q = 1, ..., Q) for the region r (r = 1, ..., R). The weights obtained by this method are potentially different for each region to maximize the influence of strengths and, conversely, to minimize the influence of weaknesses. The method is suitable when we would like to avoid assigning larger weights to certain indicators due to the improvement of the position of certain regions. At the same time, the method is given a motivating approach for lagging regions, because in contrast to other methods of determining weights, it points to their strengths.

The disadvantage of the method is the frequent assignment of the highest weights to only a limited number of indicators, which can be adjusted by adding conditions related to the weight of individual indicators and thus setting the interval in which the weights can move [28]. More methods can be found to add constraints. Here we focus only on the most well-known method, where we extend the basic method by adding conditions bounding the ratio of the weighted value to the total score of the region, i.e., for all  $q \in \{1, ..., Q\}$  and for all  $r \in \{1, ..., R\}$  we define a lower bound  $\alpha_q$  and an upper bound  $\beta_q$  such that

$$\alpha_q \le \frac{w_{qr} x_{qr}}{\sum_{q=1}^{Q} w_{qr} x_{qr}} \le \beta_q.$$
(5)

To enable this method, we need to determine the values of these bounds. The common method is, however, to determine the midpoint between these bounds and to determine the width of the

considered interval. The common approach to find this midpoint is to use methods based on subjective evaluation of indicators and determination of weights [29]. Due to the low number and diversity of regions, we did not include the calculation of AHP or any other method depending on subjective evaluation. We therefore calculated the midpoint based on equal weights. The calculated share for each indicator was further adjusted by 20% (in the case of the lower bound by a decrease and in the case of the upper bound by an increase). The results are also taken as a basis for robust and sensitivity analyses.

Due to the difficulty of interpretation, we consider only equal weights for the second level of aggregation.

#### 3.4. Methods of Aggregation

In general, a particular combination of CI's components and a chosen aggregation method usually determines the level of compensability, i.e., the level to which a potentially weaker value of one component can be compensated by others [6]. This work considers a selected family of methods that are commonly used for aggregation, namely linear aggregation using arithmetic mean, geometric mean, Borda scoring rule, Copeland's approach and data wrapper analysis. Let us also note that a sole choice of a particular method is not a final solution. There are several other processing parameters and choices that can influence the whole process such as computing weights of aggregated components. This task can be solved by various approaches including determination by an expert or automatically e.g., via factor analysis.

Let us remember that the major steps in the ranking are two levels of aggregations. The first is the aggregation of indicators into goals and the second is the aggregation of goals into key areas. For the construction of CIs, Papadimitriou [21] suggests using a linear approach to combine indicators within a goal and a geometric approach to merge goals into one single CI. This choice implies that a trade-off between indicators within one goal is allowed but the goals should not be fully compensable. In this work we recommend a further extension based on the use of linear aggregation within each goal and consequent non-fully compensable aggregation of goals. SDGs also only use linear aggregation [20]. Even though this approach has been audited [21] we decided to include linear aggregation for the second level for reasons of comparison. Hence the arithmetic average, the Borda rule or BoD was applied for merging indicators within the goal and the geometric average, and the Copeland rule and the arithmetic average was employed for merging resulting goals. Altogether 9 alternative CIs defined by particular choices of methods were calculated (see Table 1).

	Weighting Method of Indicators	Aggregation within a Goal	Aggregation of Goals
CI 1	Equal weighting	Linear aggregation	Copeland approach
CI 2	BoD with constrained	Linear aggregation	Copeland approach
CI 3	Equal weighting	Borda scoring rule	Copeland approach
CI 4	Equal weighting	Linear aggregation	Geometric aggregation
CI 5	BoD with constrained	Linear aggregation	Geometric aggregation
CI 6	Equal weighting	Borda scoring rule	Geometric aggregation
CI 7	Equal weighting	Linear aggregation	Linear aggregation
CI 8	BoD with constrained	Linear aggregation	Linear aggregation
CI 9	Equal weighting	Borda scoring rule	Linear aggregation

Table 1. Overview of constructed composite indicators.

#### 3.5. Methods in Detail

Linear aggregation allows for full compensability, i.e., poor performance in one indicator can be compensated by sufficiently high values of other indicators. In practice, linear aggregation is the most widespread. The simplest method is weighted average:

$$CI_r = \sum_{q=1}^{Q} I_{qr} \cdot w_q, \tag{6}$$

subject to

$$\sum_{q=1} w_q = 1$$
 and  $0 \le w_q \le 1$ ,

where  $w_q$  is the weight for the indicator q (remember that we use equal weights for the arithmetic mean Borda rule). Because linear aggregation implies full compensability, this aggregation is appropriate for merging within one goal of sustainable development. The indicators within one goal can be considered compensable because they measure a similar phenomenon.

Another example of compensable aggregation is the Borda rule which is known from the multi-criteria decision theory and the theory of social choice. For the fixed indicator, the Borda rule is a scoring rule which assigns no point if a region ranks last and one point if it ranks last but one. Given a total of R regions, the process continues like this up to R—1 points awarded to the region ranking first. Finally, the Borda rule sums up these scores across indicators. The Borda method is based on ordinal information and therefore the interval level information is lost. Evidently the impact of outliers is eliminated.

Compensability between indicators is only desirable if the indicators in question are considered to be substitutes [30]. Even if full compensability is weakened by the weighting scheme, different aggregation rules can completely suppress that. Geometric aggregation is only partially compensable which is illustrated by the following formula:

$$CI_r = \prod_{q=1}^{Q} I_{qr} \cdot w_q, \tag{7}$$

where  $I_{qr}$  is a normalized indicator q(q = 1, ..., Q) for region r(r = 1, ..., R) and  $w_q$  weight for indicator q(q = 1, ..., Q). Geometric aggregation rewards regions with higher scores in stronger intensity because marginal utility of an increase in a low score is much larger than in a high score. It means that a region which wants to rank better should focus on dimensions/indicators where it performs poorly.

If it is not desired to compensate the shortage in one indicator by a surplus in another one, a non-compensatory approach based on multi-criteria approach is to be considered. The Condorcet theorem is non-compensable and became a base for other aggregation rules [31].

The Condorcet approach is based on the determination of an outranking matrix *E* indexed by regions and defined for a fixed region *q*. Considering the application of min-max normalization, each element  $E_{ij}$  contains either 1 if  $I_{qr_i} > I_{qr_j}$ ,  $\frac{1}{2}$  if  $I_{qr_i} = I_{qr_j}$  or 0 otherwise. Furthermore, the Condorcet approach proceeds with a column sum. The *i*-th element of this sum expresses the number of wins for the region  $r_i$  over all other regions (considering cases where  $I_{qr_i} = I_{qr_j}$  as half-wins). Finally, the Condorcet method filters out those sums that are less than or equal to  $\frac{R-1}{2}$  leaving their corresponding values as zeros. In this way a final ranking is obtained, see details in [31]. Since this approach is based on pairwise comparisons within the given group of regions, computational issues can arise. There are several rules which have been developed to overcome the computational problems [6,32]. Apart from compensability, the properties of this method are the same as those of the Borda aggregation, i.e., interval level information is lost, only ordinal information is used and thus it is independent of outliers.

We have decided to use a slight modification of the Condorcet approach called the Copeland method [33]. This method modifies Condorcet's approach by reducing the number of wins in pairwise comparisons by the number of comparisons that result in the unit losing. In practice, this means that instead of matrix *E* we use a modified matrix *E* that can be produced out of *E* via the following substitution. All units in the matrix *E* greater than 0.5 are replaced by 1, elements less than 0.5 by -1, and finally elements equal to 0.5 by 0. The rest of the algorithm remains the same.

# 3.6. Robustness and Sensitivity Analysis

The simplest measure of the uncertainties of a constructed composite indicator is the range of all selected variants and combinations for the design. This is a very simple characteristic, which is computationally unpretentious, but it must be supplemented by further analyses. Another option is the average absolute difference of the rank. The measure is constructed using the absolute differences of the rank of the regions and the reference rank for each calculated composite indicator (c = 1, ..., C)

$$\bar{R}_{Sc} = \frac{1}{R} \sum_{r=1}^{R} |rank_{ref}(CI_r) - rank(CI_{cr})|, \qquad (8)$$

where  $rank_{ref}(CI_r)$  is the order of the region r(r = 1, ..., R) of the median of CIs and  $rank(CI_{cr})$  is the order of the region r(r = 1, ..., R) based on the observed composite indicator c(c = 1, ..., C) [7]. The reference order can be a combination of methods and thus a certain composite indicator or median (or average) of the computed variants of composite indicators. It is possible to proceed similarly to create a measure for individual regions with monitoring of the differences in the order of the given region over all variants of the composite indicator and the reference region

$$\bar{R}_{Sr} = \frac{1}{C} \sum_{c=1}^{C} |rank_{ref}(CI_r) - rank(CI_{cr})|, \qquad (9)$$

The sensitivity analysis based on the decomposition of variability is based on the method published by Sobol [34] and subsequently modified by Saltelli [35]. The calculation is based on the decomposition of the variability of the output variable Y—in our case this variable is the rank of the region, i.e.,  $rank(CI_{cr})$ . To determine this decomposition we need to introduce factors  $X_i$  that represents major steps in our analytical pipeline that are sources of uncertainty. In our case these factors are (1) determining weights represented by factor  $X_1$ , (2) first-level aggregation represented by factor  $X_2$ , and (3) second level aggregation represented by factor  $X_3$ . The values of these factors are usually just a coding by number of used methods, e.g.,  $X_1 = 1$  for equal weights and  $X_1 = 2$  for others or similarly for first-level aggregation  $X_2 = 1$  for arithmetic average,  $X_2 = 2$  for Borda scoring rule and  $X_2 = 3$  for BoD. For the *k*-th factor  $X_k$  we can decompose the variability V(Y) into the so-called main effect  $V_k$  and the residual effect  $V_{r_k}$  as

$$V(Y) = V_{X_k}(E_{X_{-k}}(Y|X_k)) + E_{X_k}(V_{X_{-k}}(Y|X_k)) = V_k + V_{r_k},$$
(10)

where  $E_{-k}(Y|X_k)$  is the mean value of Y conditioned on  $X_k$  computed over all factors but  $X_k$  and  $V_{X_k}(\cdot)$  is the variance of the corresponding variable computed over values of  $X_k$ . The proportion of the main effects and the total variability indicates the sensitivity coefficient of the first order  $S_k$ 

$$S_k = \frac{V_{X_k}(E_{X_{-k}}(Y|X_k))}{V(Y)} = \frac{V_k}{V(Y)}.$$
(11)

Similarly, the method can be used for a combination of several factors. The total effect of the factor *k* in the model with interactions can then be expressed using the sensitivity coefficient of the total effect  $S_{T_k}$ 

$$S_{T_k} = \frac{V(Y) - V_{X_{-k}}(E_{X_k}(Y|X_{-k}))}{V(Y)} = \frac{E_{X_{-k}}(V_k(Y|X_{-k}))}{V(Y)}.$$
(12)

In the case of an additive model without interactions, it applies

$$V(Y) = \sum_{k} V_k.$$
(13)

The significance of the difference between the total effect of  $S_{T_k}$  and the first-order effect of  $S_k$  for a certain factor of  $X_k$  indicates the importance of interactions for Y. The analysis of both coefficients and their differences is a tool that significantly helps to understand the model.

The data used for the identified problems in the evaluation of the low number of small regions were taken from the Czech Republic on the basis of Strategic Framework Czech Republic 2030 [22], which focuses on sustainable development and quality of life.

## 4. Results

To assess the suitability of the dataset for composite indicators design, the first step we undertook was the analysis of the initial data. Let us summarize the problems observed during the initial analysis of the dataset. The major feature that has been studied is the correlation structure between individual indicators that exhibited four fundamental problems:

- negative correlations occurring between individual indicators within the goals,
- the existence of an overly strong correlation between indicators in a particular goal,
- negative values of correlation for pairs of indicators from different goals within one area and
- high positive values of pairwise correlation coefficients for indicators from different goals in one area.

The resulting correlation structure calculated using the pairwise correlation coefficients of the normalized indicators within the individual areas is shown in Figure 1.



Figure 1. Pearson's rank correlation coefficient between indicators.

For the first-level of aggregation, we use methods with compensability within the goals, i.e., arithmetic mean, the BoD and the Borda's approach. The weights were determined to be equal or determined by the BoD method. For all indicators, pair correlation coefficients were calculated with the respective goal (first level of aggregation); see Figure 2. For the second level aggregation, we use linear aggregation, geometric aggregation and Copeland's rules; see Figure 3. Table 2 shows the correlations between the individual goals for the arithmetic mean and the Borda approach used in the first level of aggregation and the correlations of the individual goals with the respective area.



Figure 2. Pearson's rank correlation coefficient between indicators and goals.



**Figure 3.** Comparison of regions ranks determined by various composite indicators (CI1 to CI9) in three main areas (ECO, SOC, and ENV).

Pairwise correlation coefficients between individual indicators and respective targets are influenced both by the potential change in weights in the case of the BoD method as well as by the ordinalization in the case of Borda's rule. When using Borda's rule, the values of all pairwise correlation coefficients are higher than 0.5, and in most cases, also lower than 0.9. Much higher values occur mainly within linear methods, with the arithmetic mean being less affected. When using BoD weights, the variability of correlation between coefficient values within individual targets is higher than when using equal weights.

In Table 2 we can see a comparison of pairwise correlations between goals when using either the arithmetic mean or Borda's rule. We can see a significant difference between the correlations of goals depending on the method used in the economic area.

We can continue further and evaluate the process with the second level of aggregation, thus considering various combinations of methods producing composite indicators CI1 to CI9 as given in Table 1. Already Table 2 shows correlations of these composite indicators with goals in each area. We can see that if we use a combination of the arithmetic average or the geometric average in the second level primarily with linear methods, i.e., CI4, CI5, CI7, and CI8, in the first-level aggregation, the results in the economic area are quite dependent on goals E1 and E3.

	E1	E2	E3	E4		<b>S</b> 1	S2	<b>S</b> 3	<b>S4</b>		N1	N2
E1		0.73	0.90	0.59	<b>S1</b>		0.74	0.84	0.67	N1		0.51
E2	0.37		0.49	0.32	<b>S2</b>	0.75		0.83	0.52	N2	0.47	
E3	0.68	0.09		0.50	<b>S</b> 3	0.75	0.63		0.58			
E4	0.03	0.19	-0.26		<b>S4</b>	0.71	0.57	0.57				
CI1	0.77	0.57	0.68	0.46	CI1	0.85	0.73	0.84	0.85	CI1	0.66	0.89
CI2	0.72	0.66	0.66	0.35	CI2	0.82	0.72	0.87	0.68	CI2	0.54	0.89
CI3	0.89	0.54	0.55	0.18	CI3	0.91	0.78	0.81	0.81	CI3	0.82	0.86
CI4	0.94	0.62	0.90	0.65	CI4	0.92	0.88	0.91	0.79	CI4	0.81	0.91
CI5	0.96	0.89	0.95	0.80	CI5	0.89	0.83	0.91	0.68	CI5	0.79	0.77
CI6	0.81	0.49	0.71	0.38	CI6	0.93	0.84	0.86	0.82	CI6	0.85	0.85
CI7	0.94	0.62	0.90	0.65	CI7	0.92	0.88	0.91	0.79	CI7	0.81	0.91
CI8	0.96	0.89	0.95	0.80	CI8	0.89	0.83	0.91	0.68	CI8	0.79	0.77
CI9	0.81	0.49	0.71	0.38	CI9	0.93	0.84	0.86	0.82	CI9	0.85	0.85

**Table 2.** Correlation between goals and between goals and CIs (for the arithmetic mean and the Borda's method in the first level of aggregation).

Further results concerning composite indicators are shown in Figure 3. These results show high variability of the order in the case of the economic area. There is a significant effect of using linear methods on both levels of aggregation, i.e., CI7 and CI8. In the second level of aggregation, the full compensability of the arithmetic mean is manifested. However, observed outliers in ranks mainly belong to composite indicators with the arithmetic mean and BoD in the first level of aggregation. The variability in the data is reflected in the final results also when using the geometric mean, which allows only partial compensability and is sensitive to remote observations.

For further robustness and sensitivity analyses, we therefore used only methods without full compensability on the second level of aggregation, i.e., CI1 to CI6.

To assess the robustness of the results for individual composite indicators as well as regions, we used the average absolute differences of the order  $R_s$  for regions as well as individual combinations of methods—composite indicators CI1 to CI6, see the upper part of Figure 4. Similar results for the first level of aggregation for particular goals can be observed in the lower part of Figure 4.



**Figure 4.** *R*<sup>s</sup> for computed over regions and over combinations of methods, i.e., composite indicators CI1 to CI6, and for goals.

In these results, we can see the problems mentioned above. We can see an apparent sensitivity to changes in the weighting system in combination with the geometric mean in the second degree of aggregation, i.e., the values for CI4 and CI5. The sensitivity to the weighting system is also indicated in the lower part of Figure 4. One can observe a significant difference in the results of the BoD approach and the arithmetic mean in the economic area.

For a closer analysis of the influence of individual construction steps, a sensitivity analysis was performed using the decomposition of variability into the main effect and the residual effect and the associated first-order sensitivity coefficients. Main effects (left) and sensitivity coefficients (right) are shown in Figure 5.



Figure 5. Decomposition of region variance (CI1 to CI6).

The analysis confirms the results described above. The economic area is characterized by a high variability. A nonlinear and thus non-additive component causes the major effect of variability. Apart from this component, the strongest first-order, main effect is at the second level of aggregation, see regions of strongest variance in the economic area in Figure 5. An interesting result is that the effect, as mentioned above, of the second level of aggregation, which can be seen within an economic as well as environmental area, is not present in the social area, where the main effect is represented by the weighting step, at least for the regions with the strongest variance.

## 5. Discussion

## 5.1. Composite Indicators Framework

We analyzed issues that can potentially occur during the compilation of a framework for composite indicators for sustainable development when using datasets of a smaller, regional scale. Using such a domain commonly brings many problems, including the small dimensions of the dataset as well as high variability. Small dimensions are usually caused by the limited availability of suitable data at the regional level. Solving the suitability of data means resolving the interplay between the particular needs of the composite indicator framework, such as the clear definition of goals, statistical properties, i.e., a good correlation structure, and the real-world suitability of the source data. For these reasons, we suggest a particular process of constructing a simplified framework suitable for further data analysis. Particularly for the case of regional sustainable development, we make use of a particular strategic approach that gives rise to a three-level indicator construction use of key areas that are further subdivided into goals. Thus, at the beginning of our work, it was necessary to make changes in the structure of key areas and the goals based on the Strategy 2030 and in the proposed individual indicators. In this step, the analyses of pairwise correlations between indicators

correlation structure of indicators.

and goals were involved. Each region is unique—some regions are mountainous, others lowland, and others completely urban. Moreover, regions differ in their industrial focus—some regions focus more on industry and others on agriculture or services. The ideal solution for the analysis would be to cluster the datasets along with these properties. Due to the low number of regions, corresponding partitions into classes would produce quite small datasets (for example, the Capital City of Prague is the only completely urban region). In this step, we identified four fundamental problems based on the

The first issue that emerged was represented by negative correlations occurring between individual indicators within the goals. This problem was identified in our case, mainly in the areas of SOC and ENV. In the case of indicators describing the goals of the social area, it was found that some pairs of indicators have high positive correlations. However, there are other pairs of indicators in the goal connected by an indirect linear dependence. Although from a socio-economic point of view, the inclusion of the given indicators was assessed as beneficial, it was necessary to exclude them or add them to another goal. The reasons for the necessity of such a step are hidden in a problematic interplay between the structure of the framework of Agenda 2030 and the capabilities of the methods used as well as the specifics of the dataset represented by the Czech Republic's regions. Please note that small regions have a higher tendency to have narrow specialisations.

The next problem is represented by the existence of a strong correlation between indicators in a particular goal. Examples include indicators of goals in the ENV and ECO areas. This might cause issues representing by the existence of a major effect artificially neglecting the others caused by the adoption of a particular weighting system, note that all indicators with pairwise correlation high represent one effect. In the ENV area, this is due to the nature of the observed phenomena expressed by indicators. In the case of the economic area, however, the issue is deeper. This phenomenon is additionally supported by various properties of the dataset as well as the steps used in the processing pipeline. The monitored indicators have great variability, and at the same time, the structures of their outliers are similar; see the already mentioned situation in the region of the Prague Capital, which acts as a remote observation for some indicators. Moreover, this is further strengthened using the normalization step. However, in the case of a small number of regions, the common approach of the use of more robust methods (windsorization, etc.) is problematic.

Another problem is represented by the negative values of correlation for pairs of indicators from different goals within one area. Even though these indicators belong to different goals, this structure causes problems when applying standard methods as has been noted above. In the area of sustainable development, this is a common phenomenon. For this reason, one of the objectives in our dataset was transferred from the economic area to the environment area even before the start of the analysis; see also the corresponding discussion about N1 above. Even more complicated seems to be the social area, particularly the "Structural inequalities in society are low" goal. This goal has a different internal correlation structure compared to other goals in this area, so the values of the pairwise correlation coefficients were small. At the same time, some of the indicators in this goal possessed (in some cases also significant) negative values of pairwise correlation coefficients with indicators from other goals in this area. For these reasons, this goal has been dropped from the model. It may be beneficial to rethink the design of social objectives as well as to suggest new potentially valuable indicators, whose data can be collected to enable a wide understanding of this area. In our situation, we had to rectify this problem, mainly caused by the lack of supporting data, by the above-given modification.

The last of the major problems resulting from the correlation structure of the indicators is represented by the high positive values of pairwise correlation coefficients for indicators from different goals of one area. The reason is the proximity of the phenomena that characterize the goals (the case of the environmental area), but also the background influences, affecting the phenomena characterizing multiple goals (e.g., S2 and S3). In the environmental area, this phenomenon occurs in the "The landscape of the Czech Republic is conceived as a complex ecosystem and ecosystem services provided an appropriate framework for the development of human society" and "The Czech landscape

is diverse, and biodiversity is being restored" goals. These two goals are generally problematic in terms of data availability. At the regional level, very few indicators that track these goals are monitored and published. Moreover, the observed phenomena are highly interconnected, which causes statistically significant values of correlation coefficients ( $\alpha = 0.05$ ). For these reasons, the goals were merged into one goal—N2.

## 5.2. Problems in the Data and the Use of Different Aggregation Methods

In our work, pairwise correlation coefficients were also used to analyze the correlation between individual indices and relevant goals, between individual goals and between goals and relevant key areas.

When we analyzed the pairwise correlations between the individual indicators, and the respective goals, much higher values were found, mainly in linear methods, with the arithmetic mean being less affected. The reasons for this behavior are both the existence of multiple outliers in the data, as well as the sensitivity to the particular choice of weights (BoD approach). When using BoD weights, the variability of correlation coefficient values within individual goals is higher than when using equal weights.

Significant differences were identified in the results in pairwise correlation coefficients between goals using the arithmetic mean and Borda's rule in the economic area. The reason for this difference is hidden in the different structures of values of indicators from two different goals. It is manifested in the low correlation of these indicators, e.g., between indicators of the E2 and E3 goals. The difference in structure is indicated by Borda's rule but is often levelled by the arithmetic mean. The correlation structure between indicators is more significant for the social area, and consequently, the values of pairwise correlations of goals are much higher as well. Contrary to that, the goal E4 represents an extreme case of the above-discussed problem that we can also see in the analysis of the pairwise correlations between goals and areas. For that reason, if we were like to choose Borda's rule, it could be quite beneficial to redesign the area to avoid its application in an actual form (e.g., additional indicators with a better correlation structure).

Overlapping goals and high correlations between goals cause, when using arithmetic and geometric aggregation at the second level of aggregation in combination with linear methods in the first-level of aggregation, results dependent on only a few goals. However, this phenomenon cannot be eliminated when evaluating the goals of Agenda 2030. The solution would be the restructuring of key areas and goals, which would, however, also worsen the evaluation of the goals of Agenda 2030.

The analysis of sensitivity confirms that the economic area is characterized by high variability. A nonlinear and thus non-additive component causes the major effect of variability. For this reason, we recommend further exploration of this component in future studies. Apart from this component, the strongest first-order sensitivity coefficient is the main effect at the second level of aggregation. The reason for this is the use of the partially compensatory, as well as the non-compensatory methods. Hence, the values of CIs with the geometric aggregation in the second level of aggregation shows higher sensitivity to changes in the weighting system as one can observe a great difference in the results of the BoD approach and the arithmetic mean in the economic area. That is due to the excellent results in a few indicators at some regions, which are otherwise below average. In this case, BoD determines high weights for these indicators, thus clouding the overall results. For these reasons, the choice of indicators should be made very carefully regarding these impacts. Such an approach can easily run into the obstacle of a lack of data for particular goals on the regional level. For these reasons, Borda's rule seems to be more satisfactory.

One of the potential approaches to resolving some of the issues in this part is represented by the community detection algorithms [36]. These algorithms are based on the analysis of weighted graphs, where nodes represent particular indicators and weights represents the dependency between these nodes. These weights can be represented by correlations since, as has been shown in many systems, network analysis with correlation is often satisfactory, see examples for analysis of the human

brain [37], Earth's climate [38] or stock markets [39]. The goal of a community detection algorithm is to find a decomposition of the graph into subsets called communities, where inside communities we expect a large number of highly weighted edges while in-between communities we expect a low number of low weighted edges. Many approaches are analyzing either weighted of a filtered unweighted graph, such as [40–42]. Nonetheless, its satisfactory assessment is beyond the scope of this work. This blind method can provide some new viewpoints into the structure of the composite indicators. However, the obvious problem is also that for such an analysis, more abundant data availability is needed.

#### 5.3. Evaluation and Selection of Methods

In line with the suggestion of a particular structure of an analytical pipeline, we also propose a list of methods to be used in particular levels of this pipeline. For the normalization step, we used the min-max method, which is partially able to resolve problems caused by outliers. The handling of this phenomenon is even more important because the negative effects of remote observations are strengthened by the low availability of data commonly present at the regional level.

On the other hand, eliminating remote observations before an analysis can cause a significant loss of information contained in the data, and thus violates the final assessment of some units. We suggest resolving this issue by using methods involving ranking the original values. We have also shown some positive effects of this step. The complete assessment of this approach is, however, the question for future research. This would, however, require a more detailed simulation study to assess the contributions and negatives of such a step.

When using BoD, the problems of the data matrix manifest themselves even more deeply. To simply put it, although the idea of using this method for very diverse regions sounds promising, it has a large number of pitfalls. Although the use of the same weights for all units can be considered "unfair", the use of a different weighting system provided by BoD requires a very careful selection of input indicators. In the case of an improper selection, BoD weights can make the situation even worse. This represents the main problem of using BoD at the regional level, where data availability is problematic.

Based on the performed analyses, Borda's rule seems to be the most suitable of the first-level aggregation methods. For further work, we propose an analysis and comparison of methods from the theory of multi-criteria decision making, such as the TOPSIS, VIKOR methods, etc. TOPSIS is based on the concept that the chosen alternative should have the shortest geometric distance from the positive ideal solution and the longest geometric distance from the negative ideal solution. VIKOR ranks alternatives and determines the solution named compromise that is the closest to the ideal. For more detail about these methods see [43].

Any problem of aggregation at the first level is usually made even worse at the second level. Commonly used or recommended methods for this level are the arithmetic and geometric mean. The clear advantages of these methods are their simplicity and general awareness of their calculation. For the arithmetic mean, we see a problem in allowing full compensability for goal aggregation. Geometric aggregation is known to have problems when working with remote observations. These problems are even more apparent when the geometric approach is used in combination with linear aggregation methods at the first stage. In addition, the question is whether it is not appropriate to use non-compensatory techniques that are "stricter" in terms of the possibility of compensation for the regions, as sustainable development is a coherent system and none of its objectives should be neglected.

#### 6. Conclusions

For the construction of composite indicators at the regional level, we propose to maintain the usual two-level design adopted from Agenda 2030. With correspondence with this program, this approach entails aggregating individual indices into goals at the first level and then aggregating them into key areas at the second level of aggregation. Key areas need to be adapted to the regional level. The goals

pursued must be adjusted so that they do not overlap, but complement appropriately for it to describe key areas sufficiently along with keeping the availability of corresponding suitable quantitative data for the construction of the composite indicators. Even though the original framework in our example on the Czech Republic was relatively rich, with six key areas altogether, the final simplified framework ended up with only three key areas. These three areas correspond to standard pillars of sustainable development. It suggests that richer models might tend to be reduced into standard type models using statistical, phenomenological and data availability arguments. All results indicate the impossibility of directly evaluating the objectives of the 2030 Agenda using composite indicators in the case of a regional level with a small number of regions.

In this work, we list problems that can potentially occur during the compilation of a framework for composite indicators for sustainable development when using datasets of a smaller, regional scale. Two major problems in the data represented the main issues that were identified.

The first problem identified is the availability of quality data. It is the largest problem in the regional concept of sustainable development assessment. The low availability of data is problematic both in terms of the explanatory power of composite indicators, which are very demanding for the selected individual indicators and their quality and in terms of statistics. In the above example, it was clear that for this reason, it is necessary to significantly change the framework of composite indicators and adapt methods for their construction, such as modification of the goals structure according to unsatisfactory statistical properties or avoid methods working without data ordinalization.

The second problem is high data variability and sensitivity to short-term fluctuations. One of the reasons causing this is hidden in the existence of various types of regions in the dataset. Unfortunately, the number of regions of a particular type is usually quite small. For this reason, it makes it impossible to divide regions into groups that would make the data more homogeneous. The results given by the commonly used methods are, therefore, non-robust and sensitive to fluctuations. The solution to this problem in all levels of aggregation is the ordinalization of the data or the use of aggregation and Copeland rule in the second level). The reason for these data issues is also hidden in the sensitivity to changes in the weighting system, which is eliminated during the initial ranking of the data or by the use of the Borda's method.

Altogether, we can claim that assessment of sustainable development under presented conditions, i.e., stated in a simpler form as regional level with the low number of regions, cannot be directly applied. There must be either methodological or data availability enhancements.

**Author Contributions:** Conceptualization, L.P. and D.H.; methodology, L.P. and D.H.; writing—original draft preparation, L.P., D.H. and T.P.; writing—review and editing, L.P., D.H. and T.P.; resources, L.P., D.H. and T.P.; formal analysis, L.P., D.H. and T.P.; project administration, L.P. and T.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Technology Agency of the Czech Republic We support research and innovation grant number TL02000206.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

# **Appendix A. Indicators**

Table A1. The list of selected indicators in the Economy area. Bold lines indicating goals and consecutive rows represent *indicators* for a corresponding goal.

#### E1: The economy grows in long-term and the domestic sector is strong [44,45]

E1C1: Labor Productivity

E1C2: Gross Value Added in Services per inhabitant (Share of the Tertiary Sector in Gross Value Added in %) E1C3: Net Disposable Income of Households per inhabitant

#### E2: A working and stable infrastructure promotes economic activities [45,46]

E2C1: Internet access

E2C2: Water supply connection

E2C3: Sewer system connection

E2C4: Railway Lines Density (in km per 100 km<sup>2</sup>)

E3: The Czech Republic has well-functioning and stable institutions to support applied research and development and to identify opportunities in this area [45]

E3C1: Expenditure on research and development

E3C2: Share of persons working in research and development (FTE) in overall employment in the region

E3C3: Patents and new research activities

E3C4: License money

E3C5: Rate of companies applying innovations

E4: The fiscal system as a prerequisite for a successful economy is stable [45]

E4C1: Public budget deficit / surplus

E4C2: Non-investment transfers to non-profit, contributory and similar non-business organizations per capita

# Table A2. The list of selected indicators in the People and Society area. Bold lines indicating goals and consecutive rows represent indicators for a corresponding goal.

## S1: Technological and social development extend the approach to dignified work [45,47–49]

S1C1: General unemployment rate

S1C2: The share of 90/50 percentile of monthly wage

- S1C3: Share of specialists in science, technology and ICT in total employees in the region (in %)
- S1C4: Median monthly salary

S1C5: Households with a net income below the subsistence level (at-risk-of-poverty rate or social exclusion) S1C6: Long-term unemployment rate

## S2: Education develops individuals' unique potential and ability to manage and influence change, and promotes a cohesive, sustainable society oriented towards sustainable development [47,50–52]

S2C1: Highest educational level attained by economic active population

S2C2: Participation of the adult population in further education in the last 4 weeks

S2C3: Reading literacy

S2C4: Average ESCS index

## S3: The health of all population groups is improving [44,45]

S3C1: Mortality rate

S3C2: Life expectancy

S3C3: Average percentage of temporary incapacity for work

S4: Good governance [45,53,54]

S4C1: Average length of court proceedings

S4C2: Civil society—political participation S4C3: Civil society—civic participation

**Table A3.** The list of selected **indicators in the Resilient Ecosystems area**. Bold lines indicating *goals* and consecutive rows represent *indicators* for a corresponding goal.

N1: Natural resources are used as efficiently and economically as possible to minimize the external costs caused by their consumption [55]

N1C1: Nitrogen oxide emissions per capita N1C2: Carbon monoxide emissions per capita

N1C3: Production of business waste per capita

N2: The landscape of the Czech Republic is diverse and conceived as a complex ecosystem and ecosystem services provide an appropriate framework for the development of human society [55–57]

N2C1: Ecological agriculture

- N2C2: Number of days with usable water supply in the profile of medium-heavy soil below 30% of usable water capacity in at least 10% of the territory
- N2C3: Coefficient of ecological stability
- N2C4: Rate of permanent grasslands

N2C5: Coverage of landscape by woods

# References

- 1. United Nations. *Transforming Our World: The 2030 Agenda for Sustainable Development;* United Nations: New York, NY, USA, 2015.
- 2. Hassan, O.A.B. Assessing the sustainability of a region in the light of composite indicators. *J. Environ. Assess. Policy Manag.* **2008**, *10*, 51–65. [CrossRef]
- 3. Munda, G.; Saisana, M. Methodological Considerations on Regional Sustainability Assessment Based on Multicriteria and Sensitivity Analysis. *Reg. Stud.* **2011**, *45*, 261–276. [CrossRef]
- 4. Parris, T.M.; Kates, R.W. Characterizing and measuring sustainable development. *Annu. Rev. Environ. Resour.* **2003**, *28*, 559–586. [CrossRef]
- 5. Freudenberg, M. Composite indicators of country performance: A critical assessment. In *OECD Science, Technology and Industry Working Paper;* OECD Publishing: Paris, France, 2003; Volume 16.
- 6. Nardo, M.; Saisana, M.; Saltelli, A.; Tarantola, S.; Hoffman, A.; Giovannini, E. *Handbook on Constructing Composite Indicators: Methodology and User Guide*; OECD: Paris, France, 2008.
- 7. Nardo, M.; Saisana, M.; Saltelli, A.; Tarantola, S. *Tools for Composite Indicators Building, Report EUR 21682 EN;* European Commission Joint Research Centre: Ispra, Italy, 2005.
- 8. World Commission on Environment and Development; Brundland, G.H. *Our Common Future*; Oxford University Press: Oxford, UK, 1987.
- 9. Meadows, D.H.; Meadows, D.L.R.; Ers, J. *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*; Universe Books: New York, NY, USA, 1972.
- 10. Maier, K.; Čtyřoký, J.; Vorel, J.; Franke, D. *Územní plánování a udržitelný rozvoj;* ABF-Arch: Praha, France, 2008.
- 11. Byrch, C.; Kearins, K.; Milne, M.J.; Morgan, R.K. Sustainable Development: What does it really mean? *Univ. Auckl. Bus. Rev.* **2009**, *11*, 1–7.
- 12. Marsden, G.; Kimble, M.; Nellthorp, J.; Kelly, C. Sustainability Assessment: The Definition Deficit. *Int. J. Sustain. Transp.* **2010**, *4*, 189–211. [CrossRef]
- 13. Ciegis, R.; Ramanauskiene, J.; Martinkus, B. The Concept of Sustainable Development and its Use for Sustainability Scenarios. *Eng. Econ.* **2009**, *62*, 28–37.
- 14. Rassafi, A.A.; Poorzahedy, H.; Vaziri, M. An alternative definition of sustainable development using stability and chaos theories. *Sustain. Dev.* **2006**, *14*, 62–71. [CrossRef]
- 15. Stiglitz, J.E.; Sen, A.; Fitoussi, J.-P. *Report by the Commission on the Measurement of Economic Performance and Social Progress;* Commission on the Measurement of Economic Performance and Social Progress: Paris, France, 2009.
- 16. Custance, J.; Hillier, H. Statistical issues in developing indicators of sustainable development. J. R. Stat. Soc. Ser. A Stat. Soc. 1998, 161, 281–290. [CrossRef]
- Ščasný, M.; Kopecký, O.; Cudlínová, E. Alternativy k Ukazateli HDP—Zhodnocení Předpokladů a Využití Indikátoru Trvale Udržitelného Ekonomického Blahobytu (ISEW) Pro Českou Republiku; Centrum University Karlovy pro OtáZky žIvotníHo ProstřEdí: Praha, Czech Republic, 2000.

- 18. Scipioni, A.; Mazzi, A.; Mason, M.; Manzardo, A. The Dashboard of Sustainability to measure the local urban sustainable development: The case study of Padua Municipality. *Ecol. Indic.* **2009**, *9*, 364–380. [CrossRef]
- 19. Diaz-Sarachega, J.M.; Jato-Espino, D.; Castro-Fresno, D. Is the Sustainable Development Goals (SDG) index an adequate framework to measure the progress of the 2030 Agenda? *Sustain. Dev.* **2018**, *6*, 663–671. [CrossRef]
- Lafortune, G.; Fuller, G.; Moreno, J.; Schmidt-Traub, G.; Kroll, C.H. SDG Index and Dashboards Detailed Methodological Paper; Sustainable Development Solutions Network: New York, NY, USA, 2018. Available online: https://raw.githubusercontent.com/sdsna/2018GlobalIndex/master/2018GlobalIndexMethodology. pdf (accessed on 10 April 2020)
- 21. Papadimitriou, E.; Neves, A.R.; Becker, W. JRC Statistical Audit of the Sustainable Development Goals Index and Dashboards; European Commission—JRC: Brussels, Belgium, 2019.
- 22. Office of the Government of the Czech Republic. *Strategic Framework Czech Republic* 2030; Office of the Government of the Czech Republic: Prague, Czech Republic, 2017.
- 23. Czech Statistical Office. *Vybrané Oblasti Udržitelného Rozvoje v Krajích České Republiky;* Český Statistický Úřad: Praha, Czech Republic, 2010.
- 24. Fischer, J.; Helman, K.; Kramulová, J.; Petkovová, L.; Zeman, J. Sustainable development indicators at the regional level in the Czech Republic. *Stat. Econ. J.* **2013**, *93*, 5–18.
- 25. Hák, T.; Kovanda, J.; Weinzettel, J. A method to assess the relevance of sustainability indicators: Application to the indicator set of the Czech Republic's Sustainable Development Strategy. *Ecol. Indic.* **2012**, *17*, 46–57. [CrossRef]
- Hák, T.; Janoušková, S.; Moldan, B. Sustainable Development Goals: A need for relevant indicators. *Ecol. Indic.* 2016, 60, 565–573. [CrossRef]
- 27. Klugman, J. *Human Development Report 2011. Sustainability and Equity: A Better Future for All;* United Nations Development Programme; United Nations: New York, NY, USA, 2011.
- 28. Cherchye, L.; Moesen, W.; Rogge, N.; Van Puyenbroeck, T.; Saisana, M.; Saltteli, A.; Tarantola, S. Creating composite indicators with DEA and robustness analysis: The case of the Technology Achievement Index. *J. Oper. Res. Soc.* **2008**, *2*, 239–251. [CrossRef]
- 29. Pakkar, M.S. Using data envelopment analysis and analytic hierarchy process to construct composite indicators. *J. Aslied Oper. Res.* **2014**, *3*, 174–187.
- 30. Hudrlíková, L.; Petkovová, L. Regional Sustainable Development Czech Region Ranking Using Multi-criteria Decision Analysis. *Eur. J. Sustain. Dev.* **2013**, *2*, 253–262. [CrossRef]
- 31. Munda, G. Social Multi-Criteria Evaluation for a Sustainable Economy; Springer: Heidelberg, Germany, 2008.
- 32. Munda, G.; Nardo, M. Noncompensatory/nonlinear composite indicators for ranking countries: A defensible setting. *Appl. Econ.* **2009**, *41*, 1513–1523. [CrossRef]
- 33. Saisana, MATarantola, S. *State-of-the-Art Report on Current Methodologies and Practices for Composite Indicator* Development, Report EUR 20408 EN; European Commission Joint Research Centre: Ispra, Italy, 2002.
- 34. Sobol, I.M. Sensitivity Estimates for Non-linear Mathematical Models. *Mat. Model.* **1990**, *1*, 112–118. (In Russian). English translation in: *Math. Model. Comput. Exp.* **1993**, *4*, 407–414.
- Saltelli, A. Making best use of model valuations to compute sensitivity indices. *Comput. Phys. Commun.* 2002, 145, 280–297. [CrossRef]
- 36. Fortunato, S. Community detection in graphs. Phys. Rep. 2010, 486, 75–174. [CrossRef]
- 37. Hartman, D.; Hlinka, J.; Paluš, M.; Mantini, D.; Corbetta, M. The role of nonlinearity in computing graph-theoretical properties of resting-state functional magnetic resonance imaging brain networks. *Chaos* **2011**, *21*, 013119. [CrossRef]
- 38. Hlinka, J.; Hartman, D.; Vejmelka, M.; Novotná, D.; Paluš, M. Non-linear dependence and teleconnections in climate data: Sources, relevance, nonstationarity. *Clim. Dyn.* **2014**, *42*, 1873–1886. [CrossRef]
- 39. Hartman, D.; Hlinka, J. Nonlinearity in stock networks. Chaos 2018, 28, 083127. [CrossRef]
- 40. Blondel, V.D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, *10*, P10008. [CrossRef]
- 41. Newman, M.E.J.; Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **2004**, *69*, 026113. [CrossRef]
- 42. Pimentel, S.D. Choosing a Clustering: An A Posteriori Method for Social Networks. *J. Soc. Struct.* **2014**, *15*, 1–21. [CrossRef]

- Piwowarski, M.; Miłaszewicz, D.; Łatuszyńska, M.; Mariusz, B.; Nermend, K. Application of the Vector Measure Construction Method and Technique for Order Preference by Similarity Ideal Solution for the Analysis of the Dynamics of Changes in the Poverty Levels in the European Union Countries. *Sustainability* 2018, 10, 2858. [CrossRef]
- 44. Czech Statistical Office. Vývoj Obyvatelstva v Krajích České Republiky—2018. Available online: https://www.czso.cz/csu/czso/cri/vyvoj-obyvatelstva-v-krajich-ceske-republiky-2018 (accessed on 4 April 2020).
- 45. Czech Statistical Office. Regions of the Czech Republic—2018. Available online: https://www.czso.cz/csu/ czso/regions-of-the-czech-republic-2018 (accessed on 8 May 2020).
- 46. Ministry of Transport. Transport Yearbook Czech Republic. Available online: https://www.sydos.cz/cs/ rocenky.htm (accessed on 16 May 2020).
- 47. Czech Statistical Office. Labor Market 2018 in Regional Comparison. Available online: https://www.czso. cz/documents/10180/91564589/25011919.pdf (accessed on 4 April 2020).
- 48. Czech Statistical Office. Household Income and Living Conditions—2018. Available online: https://www.czso.cz/csu/czso/prijmy-a-zivotni-podminky-domacnosti-kf03f95ff5 (accessed on 8 March 2020).
- 49. Trexima, s.r.o. Average Earnings Information System 2018. Available online: https://www.ispv.cz/ (accessed on 2 May 2020).
- 50. Czech School Inspectorate. PISA 2018—Národní Zpráva. Available online: https://www.csicr.cz (accessed on 10 May 2020).
- 51. Czech Statistical Office. Teachers in Regional Education and Their Salaries—In Year 2018. Available online: https://www.czso.cz/documents/10180/134650497/230066-20.pdf (accessed on 4 May 2020).
- 52. OECD. PISA 2018. Available online: https://www.oecd.org/pisa/data/ (accessed on 10 May 2020).
- 53. Department of Justice. Czech Judiciary 2018: Annual Statistical Report. Available online: https://www.justice.cz/web/msp/statisticke-udaje-z-oblasti-justice (accessed on 9 May 2020).
- 54. Supreme audit office. Analýza—Nestátní Neziskoví Organizace v ČR. Available online: https://www.nku. cz/cz/kontrola/analysy/analysa--nestatni-neziskove-organizace-v-cr-id10354/ (accessed on 6 May 2020).
- 55. Czech Environmental Information Agency. Statistical Yearbook of the Environment of the Czech Republic 2018. Available online: https://issar.cenia.cz/publikace-o-stavu-zivotniho-prostredi/statisticka-rocenka-zivotniho-prostredi-cr/ (accessed on 20 April 2020).
- 56. Czech Environmental Information Agency. Report on the Environment in the Regions of the Czech Republic 2018. Available online: https://issar.cenia.cz/zivotni-prostredi-v-krajich/ (accessed on 20 April 2020).
- 57. Czech Environmental Information Agency. Vulnerability Indicators. Available online: https://issar.cenia.cz/ indikatory-zranitelnosti/ (accessed on 4 May 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).