Istanbul Management Journal
IMJ 96, 2024, 59–73
DOI: 10.26650/imj.2024.96.1403313

RESEARCH ARTICLE

Review of the Environmental Performance Index (EPI): Methods, Constraints and Recommendations

Uğur Arcagök[1]

[1] Assist. Prof., Muş Alparslan University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Muş, Türkiye

ABSTRACT

During the formation of composite index constructs, some constraints are encountered with respect to considerations such as the availability or accessibility of indicators, weights of such indicators, sub-indicators or variables to be used for reaching the indicators, data pertaining to such variables, and the determination of the most convenient method to be used for arranging the data set. In addition, it is required for the methods to be used to comply with the aim of the index to be formed. Researchers who want to develop a composite environmental index are required to consider all these considerations. This study aims to indicate and discuss the composite environmental index formation process through the examination of the EPI, which is a global composite index. One might argue that it contributes to the explanation of the composite index construction process, which has been examined, albeit limitedly, in the literature. Eventually, the methods that may be used in the index formation process by the researchers who want to form a composite index were discussed, and suggestions that may improve the methodological strength of the indices to be developed by them were presented. Our findings indicate the absence of an established theoretical methodology for composite sustainability indices. The creation of these indices has depended entirely on the expertise of the involved researchers.

Keywords: Composite Index, Weighting Methods, Indicators, Sustainability, Environmental Performance Index
JEL Code: C30, C38, C82

1. Introduction

We have entered a new era in recent years in terms of data-driven approaches concerning environmental sustainability. The sustainable development goals were determined, especially in the Millennium Development Goals declared by the United Nations in 2000 and at the Paris Conference organized in 2015, and the states were asked to explain their attainability of such goals using quantitative criteria. Thus, it has been ensured that the states adopt a more data-driven and empirical approach to the determined goals. In the global sense, governments have come together to determine the problems, monitor the tendencies, and measure the success or failure of the determined policies. In addition, the way has been paved for the development of composite indices that would ensure the measurement of states' performances on economic, social, or environmental issues. This process has pioneered the emergence of numerous indices such as the Human Development Index, Sustainable Development Goals Index, Global Green Economy Index, Ocean Health, European Innovation Scoreboard, Social Progress Index, and Environmental Quality Index, as well as the Environmental Performance Index (EPI).

The Environmental Performance Index (EPI), which is the first among these indices and which was published with the name Environmental Sustainability Index (ESI) before 2006, is published biennially and updated by Yale University and Columbia University in cooperation with the World Economic Forum and the EU's Joint Research Centre. The EPI presents a strong policy instrument supporting the efforts to attain the UN's Sustainable Development Goals and carry society to a sustainable future (Wolf et al., 2022). The EPI was published under the name Environmental Sustainability Index (ESI) before the 2006 ESI was aimed at decision-makers, the public, and analysts wishing to compare the nations' long-term environmental orbits. ESI tried to determine the nations' performances in terms of environmental sustainability. However, in 2006, the ESI underwent extensive modifications and began to be published under the name EPI. The EPI intends to measure sustainability at the global level, determine the problems, define the goals, follow up on the trends, understand the consequences, and determine effective policy methods (WEF, 2002). In other words, the EPI addresses the environmental dimension of sustainable development more extensively.

The goals and political categories of EPI have been determined to adhere to international policy agreements (Srebotnjak, 2007).

Corresponding Author: Uğur Arcagök **E-mail:** u.arcagok@alparslan.edu.tr

Submitted: 11.12.2023 • **Revision Requested:** 07.02.2024 • **Last Revision Received:** 08.02.2024 • **Accepted:** 28.02.2024

It can be specified that enabling environmental sustainability (MDG7), among the Millennium Development Goals (MDGs), caused the formation of EPI. Hence, in the EPI report of 2012, it was specified that EPI was the complementary factor of the Millennium Development Goals (J.W. Emerson et al., 2012). EPI had two main goals until the year 2020: goals of environmental health and ecosystem vitality, which covered the political categories concerning long-term public health. In the report published in 2022, the number of main goals was increased to three, with the addition of the climate change goal. During the formation of the index, performance goals were determined for each indicator based on international agreements, scientific literature, and expert consultancy.

This study aims to indicate and discuss the composite environmental index structure process through the examination of the EPI, which is a global composite index. The reason for the preference of the EPI as the index is that it is the first study providing a measure at an international scale regarding how close the states are to the determined goals and political categories (Hsu et al., 2013). The present study, it was based on EPI reports, which were published from 2006 to 2022, and the JRC Technical Reports, published by the Joint Research Centre (JRC) affiliated with the European Commission. The statistical methods that may be used in the global composite index formation process were defined by the methods used in the EPI. Moreover, it referred to the strengths and weaknesses of the methods used and the global indices in which such methods were used. Ultimately, recommendations were made concerning the methods that researchers who want to form a composite index may resort to. The present study may be a load-star for researchers who form a new composite index or who perform studies concerning these indices. When the literature on the indices is examined, there are limited sources in which the methodological development of a global index has been indicated. The present study will also contribute to the literature in this sense.

2. Literature Review

A composite index is a mathematical composition consisting of numerous indicators and representing more than one dimension of a concept (Saisana and Tarantola, 2002). In another definition, a composite index has been defined as the synthetic index of numerous independent indicators (Freudenberg, 2003). Today, the indices are used more extensively to facilitate communication among policy-makers, the public, and scientists (Reisi et al., 2014). Under composite indices, tendencies regarding themes such as poverty, food safety, humanitarian development, and biological diversity are followed. In other words, composite indices enable us to observe the multi-dimensional and complex constructs around us.

Environmental indices inform policy-makers and the public regarding the development process of environmental themes (Dobbie and Dail, 2001). Thus, they encourage the accountability of the states both to each other and to the public regarding the determined goals. In addition, such indices contribute to the formation of political and media awareness regarding environmental themes (Fischer et al., 2022). However when such indices are formed insufficiently or poorly, they may hinder the environmental efforts of policy-makers and the public, and may direct policy messages and decisions incorrectly (Alberti and Parker, 1991).

When the composite environmental performance indices are examined, some categories and sub-groups enable the identification of the indices (Mendola and Volo, 2017). For instance, the EPI consists of three goals: environmental health, ecosystem vitality, and climate change. The referred goals consist of eleven sub-groups (air quality, biodiversity and habitat, alleviation of climate change etc.). The ocean health index (OHI) consists of two goals: current status and possible future status. The referred goals consist of four dimensions: status, trend, pressures, and resilience. Such dimensions consist of 10 sub-groups within themselves (Halpern et al., 2012). These sub-groups formed in composite environmental indices try to establish a clear relationship between the measured structure and the structure of the index. Thus, it becomes easier for the users of the index to understand the index.

In general, three institutions lead the formation of environmental indices with their reports that they have shared with the public for many years. Since 1990, the OECD has followed more than 50 indicators from 30 member and 17 non-member states (Lankoski and Lankoski, 2023). The referred indicators addressed by the OECD have focused on individual aspects of environmental performance instead of a general evaluation of environmental sustainability. The World Bank, which is another institution that has shared reports with the public for many years, annually publishes its report, which examines the quality of life in 127 world economies under seven themes. Under the seven themes referred to in the report, data is shared regarding 18 dimensions reflecting how the use of natural resources and activities interfere with nature and environmental growth (García-Sánchez et al., 2015). The Joint Research Centre (JRC), affiliated with the European Commission, which is another institution, shares the data regarding EPI in cooperation with Yale University and Columbia University biennially. The primary characteristic of this index is that it is an integrated model considering the economic, social, demographic, and environmental dimensions. Moreover, another significant characteristic of this index is that it is also being used by the United Nations. Member states of the UN regularly inform the data of EPI to the UN Sustainable Development Commission (García-Sánchez et al., 2015).

Another significant characteristic of the EPI is that it is an index that is both use-based and scientific-based. Eyles and Furgal (2002) divide the index criteria into two groups: scientific-based and use-based. Scientific criteria cover scientific quality themes such as data usability and compatibility, indicator validity, indicator representation, reliability, and decomposition ability. These criteria are accepted in many studies (Edwards et al., 1999; Eylenbosch and Noah, 1988; Eyles et al., 1996; Von Schirnding Y.E.R., 1997). Considering the use-based criteria, they depend on the goals of the indicators. These goals are applicability, manageability, balanceability, manipulability, and the ability to serve as a catalyst. When the literature is examined, different variations of these criteria are found. For instance, such as indicator sensitivity, understandability by policy-makers, cost-effectiveness, minimum environmental effect to be gained, audience interpretability, and applicability to the population (Barber, 1994; Cairns and Mc-Cormick, P. V. Niederlehner, 1993; Edwards et al., 1999). When the EPI reports are examined, it is observed that they meet both the use-based and scientific-based criteria. In this sense, it can be said that the EPI has the quality of being for scientists, policy-makers, and the public.

3. Method

3.1. Research Design

In this section, the methods that may be used during the formation of the composite index were defined. The first step is the determination of the goal. These goals are divided in two: use-based and scientific-based (Eyles and Furgal, 2002). Then, concrete variables regarding the policies for realizing the determined goals are determined. Afterwards, forming and arranging the data set for such variables to be used in statistical analyses is needed. This stage consists of the steps of missing value imputation to variables containing missing values and normalization of the data set. Then, the type of correlation to be used for observing the strength and direction of the relationship among the variables is decided. Afterwards, the type of mean to be used in the index is determined. The most extensively used mean types are arithmetic mean and geometric mean. Finally, the weights of variables for the calculation of the index score is decided.

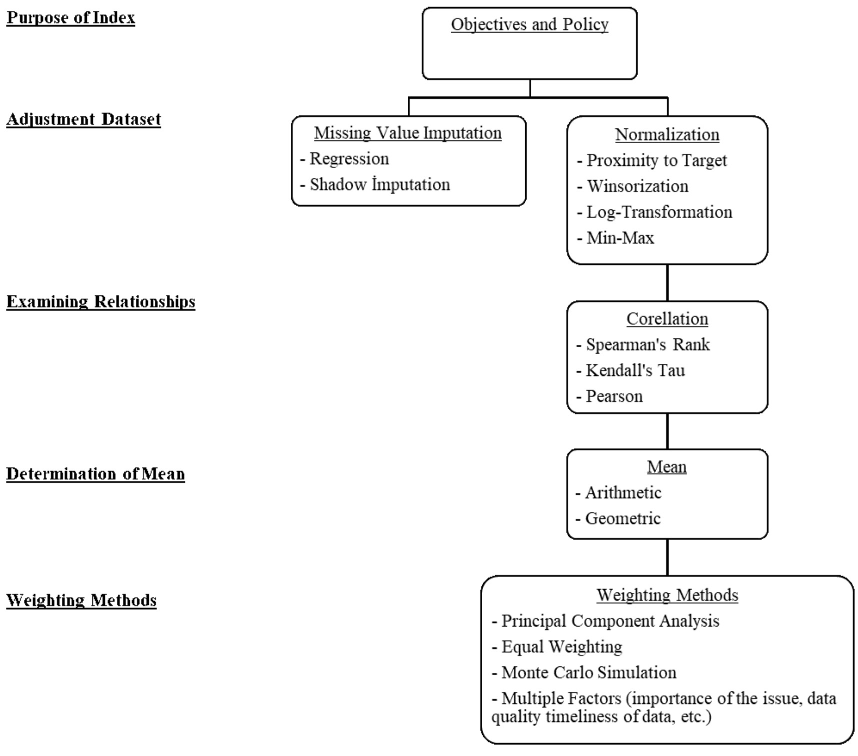

Figure 1. Research Design

3.2. Goals and Political Categories

The number and categories of EPI's indicators differ in each publication period. These indicators are selected through policy agreements and expert appraisal emerging from the Millennium Development Goals determined by the UN. The first index published in 2006 consisted of two goals (environmental health and ecosystem vitality), six political categories, and sixteen indicators. In the report published in 2022, there were three goals (environmental health, ecosystem vitality, and climate change), eleven political categories, and forty indicators (all the variables are shown in detail in Appendix Table 1).

In the first EPI report, while the goal of environmental health measured the protection of human health from environmental damages, the goal of ecosystem vitality measured the protection of ecosystems and resource management. The six political categories under these goals covered sixteen indicators in total, with at least two and at most five. In the index published in 2008, the number of indicators under six political categories for the same two goals was increased to 25 from 16 (Esty et al., 2008). However, the policy category of "sustainable energy" included in 2006 was replaced with the category of "climate change" in 2008. In 2010, the number of indicators in the index remained the same as in the previous index, but the number of political indicators had changed. The political indicator of "productive natural resources" under the goal of ecosystem vitality was redefined as three separate policies: "forestry", "fisheries", and "agriculture" (J. Emerson et al., 2010). In 2012, the political indicator of "environmental burden of disease" was replaced with the political indicator of "environmental health". Moreover, two indicators were removed from both of the political indicators of "air pollution" and "water", one indicator was removed from "agriculture", and one indicator was added to both of the political indicators of "forestry" and "climate change" (J.W. Emerson et al., 2012). In 2014, even if the number of indicators was specified as twenty, the index was calculated at nineteen indicators as the states' electricity score couldn't be calculated. The political indicator of "water", included under the goals of environmental health and ecosystem vitality, was removed from both goal and instead, it was represented in the goal of environmental health with the name of "water and sanitation". The change was made in three political goals and included under the goal of environmental health. Instead, the political indicators of "health impacts", "water and sanitation", and "air quality" were added. The number of indicators included under the goal of ecosystem vitality decreased to 14 from 17 (Hsu et al., 2014).

"Climate change", which was initially included in 2008 as a political indicator under the goal of ecosystem vitality, was named "climate change and energy" in the following years, and in 2022, "climate change" was determined as a main goal. Thus, the number of goals in the EPI was increased to three for the first time since 2006 (Wolf et al., 2022). Another first in 2022 was the use of this many indicators (40) in the calculation of the EPI score.

3.3. Adjustment of the Data Set

During the adjustment of the data set, the raw data were standardized to make the data comparable (as per population, acreage, gross domestic product, or other measures) among the states (Hsu, A., and L.A. Johnson, 2013). Additionally, the distance-to-target method was used. In the EPI, the distance-to-target technique was normalized using all the time series containing data and the states (Papadimitriou et al., 2020). This method is also used as a weighing method to evaluate the distance from the current status to the desired status (target) (Castellani et al., 2016). In 2012, the scores of proximities to the annual target were used in a simple linear regression model to determine the rate of increase or decrease for each indicator. In addition, the time series analysis was used for a few indicators in 2010. However, in 2012, the availability of time series data for nearly all the indicators enabled an increase in data quality by allowing the evaluation of trends and disparities. Moreover, for the 2012 EPI and Pilot Trend EPI, all the time series data was used to determine both the low and high-performance criteria.

Even if the distance-to-target method was used to normalize the data set in EPI, this method is a weighing method extensively used in the measurement of environmental policies (Bjørn and Hauschild, 2015; Frischknecht and Büsser Knöpfel, 2014; Lin et al., 2005; Wang et al., 2011). The distance-to-target is measured as the ratio of the current environmental load to the future environmental target value (Li et al., 2015). his method focuses on to what extent a society or a state is unsuccessful in attaining environmental standards. This approach has strengths as well as weaknesses. The progress of states towards predetermined targets, the states' exhibition of their strengths and weaknesses, the performance of standardization using a common scale, and guidance constitute the strengths. In case the determined targets and policies are subjective, however, the possibility of the calculation of distance-to-target causes biases or uncertainties, and in case the data obtained from the states are limited or inconsistent, the possibility of the data affecting the accuracy of measurement may be indicated as the weaknesses of this method (Valipour et al., 2015). Moreover, as is known, environmental systems and policies are complex and include dependency. In such cases, measuring environmental systems by reducing them to a single measure may be deceptive.

In the second step of the adjustment of the data set in the EPI, the skewness of distributions was corrected using a logarithmic transformation for the raw data of each indicator. As the transformed data was generally substantially skewed, a logarithmic transformation is required in specific data sets (namely, data sets with left or right skewness). This method has two benefits. The first is that if there are a large number of states that are very close to the target in an indicator, a logarithmic scale makes a clearer distinction among the ones exhibiting the best environmental performance. The second is that a logarithmic transformation facilitates the interpretation of differences among the sub-nation units at the opposite ends of the scale. While the log scale more accurately reflects the nature of differences in all the performance ranges, it doesn't exhibit the differences among the states in the best and worst statuses in a sufficiently distinct manner.

The use of logarithmic transformations can be beneficial in statistical modelling or regression analysis. In cases of a change in the data's variance along with the level of the independent variable, the data may exhibit heteroscedasticity. The application of a logarithmic transformation can make the variable more constant at different levels by fixing the variance (J.H. Curtiss, 1943).

Moreover, logarithmic transformation is also able to assist in the normalization of skew data distributions (Feng et al., 2014). This method may reduce the effect of extreme values and converge the data distribution to a normal distribution. Moreover, it may also be beneficial while applying the statistical techniques assuming normality or while comparing the variables with different scales.

In the third step, during the adjustment of the data set, the Winsorization method was used to prevent the states with extreme values from forming skewness. Outlier observations were adjusted as per the percentiles of 5%, 95%, and 97%. This method was used until 2018. In the 2020 and 2022 reports, the regression method was used to prevent and determine the outlier observations. While the Winsorization method provides advantages such as the providing resilience against outliers, preserving the original order and sequencing of data, and being a simple method (Jennifer Anne Haley, 2001). Disadvantages include the possibility of the results being affected due to the subjectivity of changes in extreme values, changes in the distribution and form of data, and the determination of a threshold (Barnett and Lewis, 1994).

In the report published in 2022, numerous changes were made. One of these is the adjustment of the data set. During the adjustment of the data set, the data set was normalized using the min.-max. approach as the first step (Smallenbroek et al., 2023). The purpose here is to ensure the contribution of all the indicators at equal rates. In other words, it was ensured that all the indicators would contribute to the total score between 0 and 100. The use of the min.-max. approach, among the normalization methods, ensured the data was both cleared from their units and drawn to a specific range. In other words, it facilitates the interpretation of results as well as the ease of measurement.

According to the method to be used in the index, it is required to normalize the data set because the variables forming the data set generally constitute different units. Normalization is applied to draw these variables to a common scale. The most extensively used normalization methods are standardization, min.-max., categorical scaling, and reference distance (Ruiz et al., 2020). Standardization is used to transform the data set into a scale with a mean of 0 and a standard deviation of 1. The min.-max. method is used for the data set to be in a specific range. On the categorical scale, value imputation is performed. The reference distance method is used to normalize a specific indicator as per its relative position to a reference level.

To overcome the problem of missing values, the regression method and the shadow imputation method were used. The shadow imputation method was not included in the previous reports. In the preferred method, imputation is made to the point of the missing value by ignoring the missing value and averaging out the rows or columns. This method was used for some variables. For instance, it was used in environmental variables that are valid for states that don't have access to large water bodies. Missing value imputation was generally performed by estimating with the regression method. Moreover, in the 2020 report, the median value was used for missing value imputation. In 2022, a penalty was also applied while performing missing value estimation for states with missing values. In the report, the missing value, the final step of the data set adjustment, was also referred to. When the variables had an absolute skewness higher than 2.0 and an absolute kurtosis higher than 3.5 at the same time, they were assessed as outlier observations (Groeneveld and Meeden, 1984).

Hair (2009) classified the data sets containing missing values into three groups. He specified that modelling is not required for the ones in which the ratio of missing values to the data set is below 10%, or that value imputation may be performed with the mean. If the data set contains missing values between 10% and 20%, then the hot deck imputation method should be applied for MCAR values, and the model-based missing value process should be applied for MAR cases. If the data set contains a missing value above 20% and if imputation is desired, then regression should be applied for MCAR cases and model-based imputation methods should be applied for MAR cases (Hair, 2009).

When the EPI reports are examined, the meticulous performance of the process of preparation of the data set for the analyses draws attention. Relying on these many changes in the methods used emphasizes the fact that the method to be used in research is required to comply with the data set and the purpose of the index. However, the lack transparency regarding the cause of changes made in the methods used in the reports compared to the previous period indicates the imperfection of the reports.

3.4. Correlation

In the first index published in 2006, the relationship among the variables and the relationship between the variables and the EPI were not considered. In the interpretation part of the results, the relationship between the GDP and EPI scores of the states was considered by Pearson's correlation coefficient. The reports after 2008 examined the relationships among the EPI score, goals, and political categories. In 2008, the strength of the relationship among the EPI score, two goals, and six political categories was examined by Spearman's rank correlation coefficient. In 2010, the relationship among the EPI score, two goals, and ten political categories was examined by Kendall's Tau correlation coefficient. In the reports published between 2012 and 2018, the relationship among the EPI score, goals, and political categories was examined by Pearson's correlation. In the reports published in 2020 and 2022, Spearman's rank correlation was used.

In the EPI, the results of the correlation analysis were considered at many points. If a variable contains missing values and if it has a high correlation with the other variables, then missing value imputation was performed for that variable. Otherwise, the value zero, or no data, was entered. A correlation analysis was performed to examine the relationship between the EPI score and the indicator scores, policy scores, and finally, goal scores. As each of the political categories represents different aspects of environmental performance, it was observed that they had a high correlation with the EPI while they had a low correlation among themselves. A low correlation among political categories is a status desired in the development of an index. Moreover, to determine whether environmental success is sacrificed for economic competitiveness or not, the relationship between the EPI scores and GDP, the human development index, the global competitiveness index, the voice and accountability index, and the government effectiveness index was examined by way of correlation.

In different periods of the EPI reports, three different correlation coefficients, namely Pearson's, Spearman's, and Kendall's Tau correlation coefficients were used. As is known, correlation defines the strength and direction of the relationship among the variables in the widest sense. Pearson's correlation coefficient is used to measure the strength and direction of the relationship for variables having a linear relationship, having a normal distribution, and being in the interval and ratio scale. Spearman's rank correlation coefficient and Kendall's Tau correlation coefficient are used to measure the strength and direction of the relationship for the variables without a linear relationship, without a normal distribution, ranked, and in the interval and ratio scale (Schober et al., 2018). While Spearman's rank correlation coefficient is based on ranking differences, Kendall's Tau correlation coefficient is based on the number of compatible and incompatible pairs (Kowalski and Tu, 2008). It is important to decide the type of correlation to be used as per the type of relationship (linear or non-linear), the type of scale (categorical, ordinal, interval, and ratio), and the type of distribution (normal or non-normal). Moreover, Spearman's rank correlation considers the strength of the monotonous (increase or decrease at the same time) relationship among the variables. In the EPI, as the relationship among ranks (Spearman's) is more significant than the relationship among variables (Pearson's), the use of Spearman's rank correlation in the recent two index reports (2020 and 2022) was the right decision.

3.5. Mean

The decision regarding the use of the arithmetic mean or geometric mean is made as per the distribution of the data. For instance, if the data set exhibits a normal distribution and does not contain outlier observations, and if it is a data set with an interval or ratio scale, then an arithmetic mean is suggested (Gaddis and Gaddis, 1990). If the data set exhibits a log-normal distribution and contains multiplicative relationships or outlier observations, then a geometric mean is suggested (Olsen et al., 2003). When the EPI reports were examined, it was determined that the index had been formed using the arithmetic mean method. During the adjustment of the data set, missing value imputation was performed, the problem of outlier observations was solved, showing that the use of the arithmetic mean after observation of the normal distribution of the data set was the right decision.

3.6. Weighting

In the three EPI reports published prior to 2012, two goals were equally weighted (50% - 50%). In 2012, the use of this weighting method was discontinued. In 2012, it was observed that the EPI score was affected to a high degree by the goal of environmental health in cases of equal weighting. It was determined that the referred inequality was arising from the variance differences in the scores of the environmental health and ecosystem vitality goals. The equal weighting caused the formation of a much higher correlation between the scores of the general EPI and environmental health compared to the score of ecosystem vitality. In other words, the states exhibiting high performance in the environmental health goal would generally exhibit better performance in the EPI, independent of the score of the ecosystem vitality goal. To eliminate this statistical imbalance, the environmental health goal was weighted at 30 percent and the ecosystem vitality goal at 70 percent during the 2012 determination of the total EPI score. This weighting implies the prioritization of the indicators of ecosystem vitality compared to environmental health. The purpose is to ensure a balance among the contributions of these policy goals to the general EPI. In the EPI published in 2014, equal weights were assigned to the goals. In the political categories, if any indicator under a political category is less reliable or less related compared to other indicators under the same category, then it is weighted with a lower score. In 2016, the goals were again subjected to equal weighting. In addition, the political categories forming these goals were also subjected to equal weighting. In 2018, this equilibrium was disrupted again, and during the determination of the total EPI score, the goal of environmental health was weighted at 40 percent and the goal of ecosystem vitality at 60 percent, and the statistical strength of these weights was tested by way of Monte Carlo simulation (Wendling et al., 2018). However, the weights of the political categories were not as equal as in the previous year. The same weighting method was also used in 2020 (Wendling, Z.A. et al., 2020). In 2022, the weights of the political goals were amended again. The weights of political goals were determined using statistical analyses for balancing the significance of the theme, quality of data, relevance of data, and distribution of scores. Accordingly, ecosystem vitality was determined at 42%, climate change at 38%, and environmental health at 20%.

In the EPI, the states' scores were scaled in a manner that gave them a value between 0 and 100. A composite index was formed with a three-step cumulative model. To combine the indicators into a single composite performance score, the scores of individual components were combined into a general score after the assignment of numerical weights. This combination was performed with a linear combination (combination of weighted normalized indicator scores). When the weights are equal, it is similar to calculating the simple arithmetic mean. Uncertainty and sensitivity analyses were performed to ensure the validity of the results of policies obtained from EPI and measure the sensitivity of the index to alternative methodological assumptions. In other words, the results were examined under different scenarios to determine how the ranks and scores changed when the weights of the index were differentiated. Moreover, in the reports published in some of the years, clustering analysis was used in the interpretation of the states' EPI scores, and interpretations were made on a regional basis.

It can be specified that the weighting process is both a political and a scientific process. To form the EPI score, weights were assigned to indicators, policy categories, and goals forming the EPI. The EPI was formed considering the experts' recommendations on weights, perceived data quality, the significance of indicators and categories in terms of policy-making, and the indicators' degree of enabling the direct measurement of environmental performance, etc. Moreover, another significant issue considered while determining the weights is the basic distribution of the indicator's policy category and purpose scores, or the number of variations in the data. Principal component analysis (PCA) was used for the determination of the loads of these political categories forming the goals, the formation of suitable groupings, and the determination of their weights. As a result of PCA, PCA factor loads were used as weights for these indicators. The indicators without clear references in the PCA results were grouped as per the literature review and the experts' opinions.

As the determination of weights has an effect on the results during the formation of the composite index, it is a subject that must be considered. In many composite indices, equal weights are assigned to all the variables. In cases where equal weights are not assigned, the methods of PCA, conjoint analysis, and data envelopment analysis are generally used for the determination of weights (Ruiz et al., 2020). Moreover, the analytic hierarchy process (Saaty, 1977, 1988) or MACBETH (Benito and Romera, 2011) techniques, among multiple-criteria decision-making techniques, are also used. All these methods can affect the strengths and weaknesses depending on the purposes and goals of the index desired to be formed.

4. Discussion

4.1. Goals and Political Categories

The goals of an effective environmental index should be based on both scientific and practical criteria. The indices should be quantitative, sensitive to change, sensitive to analyses, and traceable in terms of the determined policies. When the example of the EPI is examined, it can be said that the goals and policies were formed as per scientific criteria. The EPI consists of three inclusive environmental goals and a political category for attaining such goals. These goals cover the policies prioritized by the global environment authorities, the environmental dimension of the Millennium Development Goals, and the net zero greenhouse gas emission goals of the Glasgow Climate Pack. These goals are based as much as possible on international agreements and contracts (Moldan et al., 2012). The goals not included in this category were formulated assertively by the ones forming the index, and they were formed in a manner that will allow improvement for all the states.

4.2. Adjustment of the Data Set

During the formation of the composite index, steps such as identification of the indicators, evaluation of the missing data, normalization, mean, weighting, summation, uncertainty, and sensitivity analysis should be performed carefully. An examination of the literature on sustainability reveals a discussion on the normalization method to be used and whether its use is required or not (Diaz-Balteiro et al., 2018). It is not required to perform normalization before some analysis methods (DEA or multiple benefit theory) used during the formation of the composite index, because these methods contain normalization. The normalization technique to be used can uncover different results (Pollesch and Dale, 2016). Which technique would be optimal should depend on the characteristics of the specific problem analyzed. While the rate of not using normalization in the sustainability indices was determined to be 70% (Ibáñez-Forés et al., 2014) by the study performed in 2014, it was determined that the relevant rate decreased to 30% in the study performed in 2017 (Diaz-Balteiro et al., 2017). Moreover, in the context of sustainability, if it is required for the final solutions to reflect the analyzed truth with the minimum error, it is required for the variables used to be normalized.

The missing value imputation is another significant issue that must be considered. An examination of the literature shows that the most extensively used missing value imputations have been case deletion, single imputation, or multiple imputations. In single imputation, mean, median, mode, hot deck and cold deck imputation, unconditional mean imputation, regression imputation, and expectation-maximization imputation are used. In multiple imputations, the Monte Carlo algorithm is used (Kondyli, 2010). The benefit of this imputation is its assistance in estimating the missing values in the present data. When the EPI example is examined, it is not clearly specified how the problem of missing values was solved, in other words, which method was used in the reports before 2020. The lack of discussion on missing value imputation indicates the deficiency of the reports published until 2020. In

2020, the referred deficiency was filled, and the median method, among the single imputation methods, was used in missing value imputation. In 2022, linear regression was used to perform missing value imputation on the data set.

4.3. Mean

Deciding on the type of mean to be used while forming the composite index is a significant step. In general, there are two mean types in composite indices: arithmetic mean and geometric mean. Which one of these would be used should be determined as per the distribution of data. For instance, if the data set exhibits a normal distribution and does not contain outliers, and if it is a data set with an interval or ratio scale, then an arithmetic mean is suggested (Gaddis and Gaddis, 1990). If the data set exhibits a log-normal distribution and contains multiplicative relationships or outliers, then a geometric mean is suggested (Olsen et al., 2003). Hence, the question of whether the geometric mean or the arithmetic mean should be used in the calculation of EPI was discussed in the reports. The mean to be used was determined based on to what extent it would change the states' EPI scores. It was observed that the geometric mean had a moderate level of effect on the EPI ranking. In other words, when the geometric mean was used instead of the arithmetic mean at the level of policy, a remarkable skewness arose in the ranking of states. The use of a geometric mean caused skewness in the median in the ranking of one-tenth of the states. Consequently, the general environmental performance index was calculated based on the arithmetic mean of the target scores. Moreover, the arithmetic mean was also used in the missing value imputation in 2010.

To indicate the states' performance in the EPI, three policy goals were combined under a single score using a weighted arithmetic mean. The use of an arithmetic mean instead of a geometric mean to compare the two summing approaches and emphasize the states with varying profiles was a correct decision, because as is known, the geometric mean tends to penalize the presence of a very low value in the data set. Easy interpretability of the arithmetic mean can ensure a balance (a high score in one goal may completely balance the low scores in another goal) among the policy goals in the EPI.

4.4. Determination of Weights

During the formation of the composite index, the correct weighting of the goals and policies is as important as the meticulous determination of such goals and policies by the relevant experts. The weighting methods that may be used by the researchers are shown in Figure 2. When the globally used composite indices are examined, the most extensively used weighting methods are equal weighting (Human Development Index, Sustainable Development Goals Index, Global Green Economy Index, Ocean Health Index, European Innovation Scoreboard) and principal component analysis (Social Progress Index, Environment Quality Index). The greatest advantage of the equal weighting method is its simplicity. However, this method has disadvantages such as the consideration of the significance levels of the indicators as equal, double weighting, and the inability to recommend concrete policies for policy-makers (Hermans et al., 2008). In weighting with PCA, there is the advantage of making individual interpretations for each factor as the indicators are grouped. However, in the preferred method, there is the risk of having the weights differing in truth as they are based on correlation. An examination of the EPI reports shows that the most extensively used weighting methods are those with equal weighting at the level of policies, equal weighting at the level of indicators, and weightings derived from factor analysis and principal component analysis. These methods were also discussed in different scenarios.

4.5. Uncertainty and Sensitivity

During the formation of the EPI, it is important to have uncertainty and sensitivity analyses performed because uncertainty and sensitivity analyses increase transparency and determine the strength of the index (Munda and Saisana, 2011; Quadus and Siddique, 2001; Saisana et al., 2005). Sensitivity analysis is used to calculate the share of uncertainty that the indicators cause in the composite index (Freudenberg, 2003). In the composite index, it is important to determine the propagation of uncertainty to input values. The formation of a composite index is performed by including or not including the uncertainty indicators and using alternative normalization, weighting, and addition schemes (Kwatra et al., 2020). Uncertainty and sensitivity analyses are very important for checking the strength of global composite indices such as the EPI, by which states' sustainability is measured and their performances can be compared.

5. Conclusion

In the present study, the EPI reports published from its initial year of publication in 2006 to its last year of publication in 2022 were methodologically examined. Before moving on to methodological discussion, the change in years of goals and political categories forming the index was also shown in detail (Table 1). The purpose here was to show that the discussed index's goal or policy level can easily adapt to changing conditions. For a composite index to be formed at the global level and gain recognition, it is required to be able to adapt to changes.

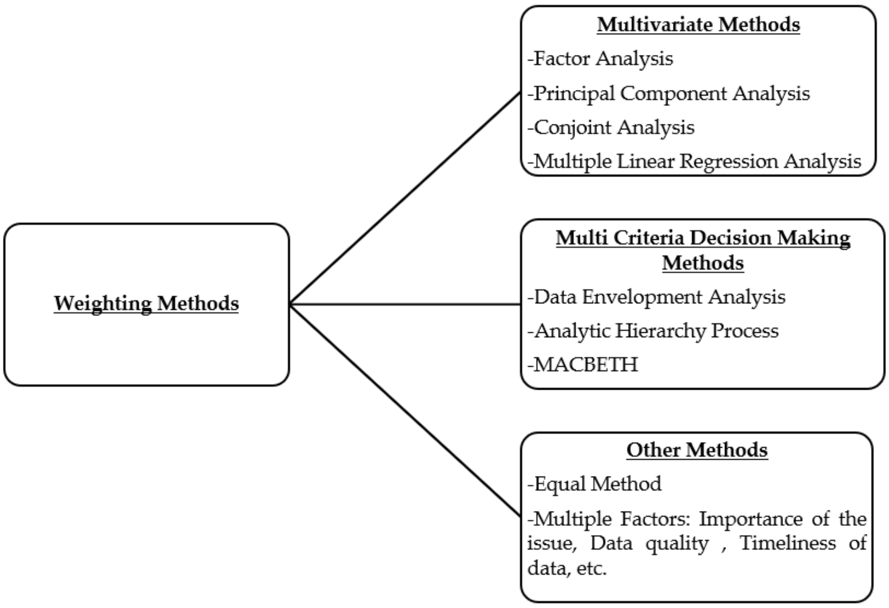

Figure 2. Weighting Methods

The composite indices have a complex construct. Therefore, it is important to consider that the statistical methods to be used do not play a dominant role in the index: in other words, they do not directly affect the scores or rankings that will be formed as a result of the index. For that reason, in the present study, it is very important to decide the methods, correlation types, mean types, and weighting methods that may be used in the adjustment of the data set. In the present study, the strengths and weaknesses of the methods and the circumstances in which they have to be used were identified. However, the causes of differentiation as per the years of the statistical methods used in the EPI reports or the justifications of preference for the methods used were not reported. This circumstance reveals the weakness of EPI reports. Resorting to these many changes in the statistical sense in the calculation of EPI may eliminate the ability to compare the index with the score of the previous index. In other words, it may cause the index to be inconsistent. The composite indices should be consistent and comparable within themselves. However, the differentiation of the determined goals or political indicators may be explained by the periodicity of such goals and policies. For instance, while climate change is a political indicator, it has become one of the main goals in 2022. Likewise, the effect of periodicity is observed in the determination of the goals' weights.

The composite index may give inconsistent messages in the case of its incorrect construction or interpretation or in the case of ignoring the significant dimensions (Singh et al., 2009). To avoid this circumstance or to strengthen the index, it is required to test the results with uncertainty analysis and sensitivity analysis. Moreover, deciding only by considering the index's score may cause information loss. In other words, the decision-making of the decision-makers only by considering the index's score may mislead them. During the formation of the index, the determined goals and the policies formed for attaining such goals or the performances in sub-indicators should also be considered. Only in this manner can the decision-makers develop the correct policies for their strengths and weaknesses.

Global composite indices, which will enable the measurement of countries' performances, are expected to increase gradually over the years. The creation and availability of consistent and comparable indicators regarding countries' structures and performances are important for the analysis of various environmental, economic, and social policy areas. Researchers who will construct a composite index need to extend the scope of indicators to broader and more flexibly defined conceptual areas. Otherwise, the failure to calculate the relative contributions of indicators could lead to the formation of misleading index scores. In other words, if the composite index created is theoretically weak, the index may rely on spurious quantitative degrees, thereby resulting in erroneous comparisons."

In conclusion, the composite indices formed for measuring sustainability generally enable the decision-makers to assess the decisions they will make in the future or allow them to raise the awareness of society. As these indices contain quantitative and temporal goals, they ensure the measurement of the states' results. Thus, this enables the ranking and comparison of the states. In addition, it may ensure transparency and accountability and facilitate decision-making on complex issues. However, such indices do not indicate the sustainable status of a state or a state's position on a sustainable route. In the sense of attaining the determined common goals, they give an idea regarding the states' or regions' degree of performance. Under the composite index, the general direction of progress towards the determined goals is determined.

Peer Review: Externally peer-reviewed.
Conflict of Interest: Author declared no conflict of interest.
Financial Disclosure: Author declared no financial support.

ORCID:

Uğur Arcagök 0000-0002-4469-9525

REFERENCES

Alberti, M., & Parker, J. D. (1991). Indices of environmental quality. Environmental Impact Assessment Review, 11(2), 95–101. https://doi.org/10.1016/0195-9255(91)90026-G

Barber, M. C. (1994). Environmental Monitoring and Assessment Program: Indicator development strategy. In US Environmental Protection Agency (Issue January 2014).

Barnett, V., & Lewıs, T. (1994). Outliers in statistical data (Vol. 3, No). Wiley.

Benito, M., & Romera, R. (2011). Improving quality assessment of composite indicators in university rankings: a case study of French and German universities of excellence. Scientometrics, 89(1), 153–176. https://doi.org/10.1007/s11192-011-0419-5

Bjørn, A., & Hauschild, M. Z. (2015). Introducing carrying capacity-based normalisation in LCA: framework and development of references at midpoint level. The International Journal of Life Cycle Assessment, 20(7), 1005–1018. https://doi.org/10.1007/s11367-015-0899-2

Cairns, J., & McCormick, P. V. Niederlehner, B. R. (1993). A proposed framework for developing indicators of ecosystem health. Hydrobiologia, 263, 1–44.

Castellani, V., Benini, L., Sala, S., & Pant, R. (2016). A distance-to-target weighting method for Europe 2020. International Journal of Life Cycle Assessment, 21(8), 1159–1169. https://doi.org/10.1007/s11367-016-1079-8

Diaz-Balteiro, L., Belavenutti, P., Ezquerro, M., González-Pachón, J., Ribeiro Nobre, S., & Romero, C. (2018). Measuring the sustainability of a natural system by using multi-criteria distance function methods: Some critical issues. Journal of Environmental Management, 214, 197–203. https://doi.org/10.1016/j.jenvman.2018.03.005

Diaz-Balteiro, L., González-Pachón, J., & Romero, C. (2017). Measuring systems sustainability with multi-criteria methods: A critical review. European Journal of Operational Research, 258(2), 607–616. https://doi.org/10.1016/j.ejor.2016.08.075

Dobbie, M. J., & Dail, D. (2001). Encyclopedia of Environmetrics. In A. H. El-Shaarawi & W. W. Piegorsch (Eds.), Encyclopedia of Environmetrics. Wiley. https://doi.org/10.1002/9780470057339

Edwards, R., Hancock, T., & Labonte, R. (1999). Indicators that count! Measuring population health at the community level. Canadian Journal of Public Health, 90(1), 22–26.

Emerson, J., Esty, D. C., Levy, M. A., Kim, C. H., de Sherbinin, A., Srebotnjak, T., Mara, V., & Jaiteh, M. (2010). 2010 Environmental Performance Index. 87.

Emerson, J. W., Hsu, A., Levy, M. A., Sherbinin, A. de, V. Mara, D. C. E., & Jaiteh., M. (2012). 2012 Environmental Performance Index and Pilot Trend Environmental Performance Index. 99.

Esty, D. C., Kim, C., Srebotnjak, T., Levy, M. A., Sherbinin, A. de, & Mara, V. (2008). Environmental performance. 2008 Environmental Performance Index, 82–89. https://doi.org/10.1787/862890ed-en

Eylenbosch, W., & Noah, N. (1988). Surveillance in Health and Disease. Oxford University Press.

Eyles, J., Cole, D., & Gibson, B. (1996). Human Health in Ecosystem Health: Issues of Meaning and Measurement. International Joint Commission (IJC) Digital Archive, 1–62. http://scholar.uwindsor.ca/ijcarchive%0Ahttp://scholar.uwindsor.ca/ijcarchive

Eyles, J., & Furgal, C. (2002). Indicators in environmental health: Identifying and selecting common sets. Canadian Journal of Public Health, 93(SUPPL. 1). https://doi.org/10.1007/bf03405121

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., & Tu, X. M. (2014). Log-transformation and its implications for data analysis. Shanghai Archives of Psychiatry, 26(2), 105–109. https://doi.org/10.3969/j.issn.1002-0829.2014.02

Fischer, S. M., Joy, M. K., Abrahamse, W., Milfont, T. L., & Petherick, L. M. (2022). The Use and Misuse of Composite Environmental Indices. BioRxiv, 2022.03.15.484501. https://doi.org/10.1101/2022.03.15.484501

Freudenberg, M. (2003). Composite indicators of country performance: a critical assessment. OECD Science, Technology and Industry Working Papers, 16(September), 35. https://doi.org/10.1787/405566708255

Frischknecht, R., & Büsser Knöpfel, S. (2014). Ecological scarcity 2013—new features and its application in industry and administration—54th LCA forum, Ittigen/Berne, Switzerland, December 5, 2013. The International Journal of Life Cycle Assessment, 19(6), 1361–1366. https://doi.org/10.1007/s11367-014-0744-z

Gaddis, M. L., & Gaddis, G. M. (1990). Introduction to biostatistics: Part 1, basic concepts. Annals of Emergency Medicine, 19(1), 86–89. https://doi.org/10.1016/S0196-0644(05)82149-3

García-Sánchez, I.-M., Almeida, T. A. D. N., & Camara, R. P. D. B. (2015). A proposal for a Composite Index of Environmental Performance (CIEP) for countries. Ecological Indicators, 48, 171–188. https://doi.org/10.1016/j.ecolind.2014.08.004

Groeneveld, R. A., & Meeden, G. (1984). Measuring Skewness and Kurtosis. The Statistician, 33(4), 391. https://doi.org/10.2307/2987742

Hair, J. (2009). Multivariate data analysis. Pearson.

Halpern, B. S., Longo, C., Hardy, D., McLeod, K. L., Samhouri, J. F., Katona, S. K., Kleisner, K., Lester, S. E., O'Leary, J., Ranelletti, M., Rosenberg, A. A., Scarborough, C., Selig, E. R., Best, B. D., Brumbaugh, D. R., Chapin, F. S., Crowder, L. B., Daly, K. L., Doney, S. C., . . . Zeller, D. (2012). An index to assess the health and benefits of the global ocean. Nature, 488(7413), 615–620. https://doi.org/10.1038/nature11397

Hermans, E., Van den Bossche, F., & Wets, G. (2008). Combining road safety information in a performance index. Accident Analysis & Prevention, 40(4), 1337–1344. https://doi.org/10.1016/j.aap.2008.02.004

Hsu, A., L.A. Johnson, and A. L. (2013). Measuring Progress: A Practical Guide From the Developers of the Environmental Performance Index (EPI). 80. http://epi.yale.edu/content/measuring-progress-practical-guide-developers-environmental-performance-index-epi

Hsu, A., Emerson, J., Levy, M., Sherbinin, A. de, Johnson, L., Malik, O., Schwartz, J., & Jaiteh., M. (2014). 2014 Environmental Performance Index. Wiley StatsRef: Statistics Reference Online. https://doi.org/10.1002/9781118445112.stat03789

Hsu, A., Lloyd, A., & Emerson, J. W. (2013). What progress have we made since Rio? Results from the 2012 Environmental Performance Index (EPI) and Pilot Trend EPI. Environmental Science and Policy, 33, 171–185. https://doi.org/10.1016/j.envsci.2013.05.011

Ibáñez-Forés, V., Bovea, M. D., & Pérez-Belis, V. (2014). A holistic review of applied methodologies for assessing and selecting the optimal technological alternative from a sustainability perspective. Journal of Cleaner Production, 70, 259–281. https://doi.org/10.1016/j.jclepro.2014.01.082

J. H. Curtiss. (1943). On Transformations Used in the Analysis of Variance. 14(2), 107–122. https://www.jstor.org/stable/2235813

Jennifer Anne Haley. (2001). Information To Users. Dissertation, 274.

Kondyli, J. (2010). Measurement and evaluation of sustainable development. Environmental Impact Assessment Review, 30(6), 347–356. https://doi.org/10.1016/j.eiar.2009.08.006

Kowalski, J., & Tu, X. M. (2008). Modern applied U-statistics. John Wiley & Sons.

Kwatra, S., Kumar, A., & Sharma, P. (2020). A critical review of studies related to construction and computation of Sustainable Development Indices. Ecological Indicators, 112(January), 106061. https://doi.org/10.1016/j.ecolind.2019.106061

Lankoski, J., & Lankoski, L. (2023). Environmental sustainability in agriculture: Identification of bottlenecks. Ecological Economics, 204(PA), 107656. https://doi.org/10.1016/j.ecolecon.2022.107656

Li, X., Su, S., Shi, J., & Zhang, Z. (2015). An environmental impact assessment framework and index system for the pre-use phase of buildings based on distance-to-target approach. Building and Environment, 85, 173–181. https://doi.org/10.1016/j.buildenv.2014.11.035

Lin, M., Zhang, S., & Chen, Y. (2005). Distance-to-Target Weighting in Life Cycle Impact Assessment Based on Chinese Environmental Policy for the Period 1995-2005 (6 pp). The International Journal of Life Cycle Assessment, 10(6), 393–398. https://doi.org/10.1065/lca2004.10.185

Mendola, D., & Volo, S. (2017). Building composite indicators in tourism studies: Measurements and applications in tourism destination competitiveness. Tourism Management, 59, 541–553. https://doi.org/10.1016/j.tourman.2016.08.011

Moldan, B., Janoušková, S., & Hák, T. (2012). How to understand and measure environmental sustainability: Indicators and targets. Ecological Indicators, 17, 4–13. https://doi.org/10.1016/j.ecolind.2011.04.033

Munda, G., & Saisana, M. (2011). Methodological Considerations on Regional Sustainability Assessment Based on Multicriteria and Sensitivity Analysis. Regional Studies, 45(2), 261–276. https://doi.org/10.1080/00343401003713316

Olsen, G. W., Church, T. R., Miller, J. P., Burris, J. M., Hansen, K. J., Lundberg, J. K., Armitage, J. B., Herron, R. M., Medhdizadehkashi, Z., Nobiletti, J. B., O'Neil, E. M., Mandel, J. H., & Zobel, L. R. (2003). Perfluorooctanesulfonate and other fluorochemicals in the serum of American Red Cross adult blood donors. Environmental Health Perspectives, 111(16), 1892–1901. https://doi.org/10.1289/ehp.6316

Papadimitriou, E., Neves, A. R., & Saisana, M. (2020). JRC statistical audit of the 2020 environmental performance index. https://doi.org/10.2760/849820

Pollesch, N. L., & Dale, V. H. (2016). Normalization in sustainability assessment: Methods and implications. Ecological Economics, 130, 195–208. https://doi.org/10.1016/j.ecolecon.2016.06.018

Quaddus, M. A., & Siddique, M. A. B. (2001). Modelling sustainable development planning: A multicriteria decision conferencing approach. Environment International, 27(2–3), 89–95. https://doi.org/10.1016/S0160-4120(01)00066-6

Reisi, M., Aye, L., Rajabifard, A., & Ngo, T. (2014). Transport sustainability index: Melbourne case study. Ecological Indicators, 43, 288–296. https://doi.org/10.1016/j.ecolind.2014.03.004

Ruiz, F., El Gibari, S., Cabello, J. M., & Gómez, T. (2020). MRP-WSCI: Multiple reference point based weak and strong composite indicators. Omega (United Kingdom), 95. https://doi.org/10.1016/j.omega.2019.04.003

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology, 15(3), 234–281. https://doi.org/10.1016/0022-2496(77)90033-5

Saaty, T. L. (1988). What is the Analytic Hierarchy Process? In Mathematical Models for Decision Support (pp. 109–121). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-83555-1_5

Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and Sensitivity Analysis Techniques as Tools for the Quality Assessment of Composite Indicators. Journal of the Royal Statistical Society Series A: Statistics in Society, 168(2), 307–323. https://doi.org/10.1111/j.1467-985X.2005.00350.x

Saisana, M., & Tarantola, S. (2002). State-of-the-art report on current methodologies and practices for composite indicator development Institute for the Protection and Security of the Citizen State-of-the-art Report on Current Methodologies and Practices for Composite Indicator Development. Eu, July, 1–72. https://doi.org/10.13140/RG.2.1.1505.1762

Schober, P., Schwarte, L. A., & Boer, C. (2018). Correlation coefficients: Appropriate use and interpretation. Anesthesia and Analgesia, 126(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Singh, R. K., Murty, H. R., Gupta, S. K., & Dikshit, A. K. (2009). An overview of sustainability assessment methodologies. Ecological Indicators, 9(2), 189–212. https://doi.org/10.1016/j.ecolind.2008.05.011

Smallenbroek, O., Caperna, O., & Papadimitriou, G. (2023). JRC audit of the Environmental Performance Index 2022. https://doi.org/10.2760/381247

Srebotnjak, T. (2007). The role of environmental statisticians in environmental policy: the case of performance measurement. Environmental Science and Policy, 10(5), 405–418. https://doi.org/10.1016/j.envsci.2007.02.002

Valipour, M., Ziatabar Ahmadi, M., Raeini-Sarjaz, M., Gholami Sefidkouhi, M. A., Shahnazari, A., Fazlola, R., & Darzi-Naftchali, A. (2015). Agricultural water management in the world during past half century. Archives of Agronomy and Soil Science, 61(5), 657–678. https://doi.org/10.1080/03650340.2014.944903

Von Schirnding Y.E.R. (1997). Indicators for Policy and Decision-making in Environmental Health. Draft report prepared for the Office of Global and Integrated Environmental Health. World Health Organization.

Wang, H., Hou, P., Zhang, H., & Weng, D. (2011). Towards Life Cycle Sustainability Management. In M. Finkbeiner (Ed.), Towards Life Cycle Sustainability Management (Issue July 2011). Springer Netherlands. https://doi.org/10.1007/978-94-007-1899-9

WEF. (2002). Pilot Environment Performance Index. World Economic Forum: Annual Meeting, An Initiative of the Global Leaders of Tomorrow Environment Task Force.

Wendling, Z. A., Emerson, J. W., de Sherbinin, A., & Esty, D. C. (2020). Environmental Performance Index 2020: Country Profile Egypt. 2–4.

Wendling, Z. A., Emerson, J. W., Esty, D. C., Levy, M. A., & de Sherbinin, A. (2018). 2018 Environmental Performance Index.

Wolf, M. J., Emerson, J. W., Esty, D. C., Sherbinin, A., & Wendling, Z. A. (2022). Environmental Performance Index 2022: Ranking country performance on sustainability issues Summary for Policymakers. In Report: Summary for Policymakers. epi.yale.edu

How to cite this article

Arcagök, U. (2024). Review of the Environmental Performance Index (EPI): methods, constraints and recommendations. *Istanbul Management Journal*, 96, 59-73. http://doi.org/10.26650/imj.2024.96.1403313

6. Appendix

Table 1. Methodological change of EPI as per years

Year	Objectives and Policy Categories	Indicators	Method	Difference from the Previous Index
2006	Environmental Health (4) Ecosystem Vitality (12): Air Quality (2), Water Resources (1), Biodiversity and Habitat (3), Productive Natural Resources (3), Sustainable Energy (3)	16	Proximity to Target Winsorization PCA Sensitivity Analyses Pearson’s Correlation Coefficient (GDP with EPI)	-
2008	Environmental Health (6): Environmental Burden of Disease (1) Water (2), Air Pollution (3), Ecosystem Vitality (19): Air Pollution (2), Water (2), Biodiversity and Habitat (4), Productive Natural Resources (8), Climate Change (3)	25	Proximity to Target Winsorization Pearson’s Correlation Coefficient PCA K-Means Method Sensitivity Analyses	The number of variables was increased to 25 from 16. Environmental health was divided into three political categories and covered six indicators in total. The category of sustainable energy was removed. The category of climate change was added.
2010	Environmental Health (5): Environmental Burden of Disease (1) Water (2), Air Pollution (2) Ecosystem Vitality (20): Air Pollution (4), Water (3), Biodiversity and Habitat (3), Forestry (2), Fisheries (2), Agriculture (3), Climate Change (3)	25	Proximity to Target Winsorization PCA Time Series Kendall's Tau Correlation Coefficient Logarithmic Transformation	The number of variables remained the same, but the political categories and the number of variables under such categories had changed. A variable (local ozone) under the political category of “air pollution”, included under the environmental health goal, was removed. When ecosystem vitality was examined, two variables (nitrogen oxide emission and volatile organic compound emission) were added under the political category of air quality. A variable (the water shortage index) was added under the political category of water. Two variables (protection risk index and effective protection) were removed and one variable (protection of living beings) was added to the political category of biodiversity and habitat. The political category of productive natural resources was removed, and instead, the sub-indicators forming the referred political category became political categories. The number of variables in forestry was increased by one (growing stock), and fisheries remained the same. In agriculture, only the variable of agricultural support remained among the variables of the previous year, and all the other variables were removed. Two variables

Table 1. Continued

				(agricultural pesticide adjustment and agricultural water density) were added instead of them. No change had occurred in the political category of climate change.
2012	Environmental Health (5): Environmental Health (1) Water (2), Air (2), Ecosystem Vitality (17): Air (2), Water (1), Biodiversity and Habitat (3), Forestry (3), Fisheries (2), Agriculture (2), Climate Change & Energy (4)	22	Proximity to Target Winsorization Logarithmic Transformation Pearson's Correlation Coefficient Linear Regression Sensitivity Analyses	The political indicator of environmental disease burden was removed, and the political indicator of environmental health was added. Two indicators were removed from both of the political indicators of air pollution and water, and one indicator was removed from agriculture. One indicator was added to both of the political indicators of forestry and climate change.
2014	Environmental Health (6): Health Impacts (1) Water ve Sanitation (2), Air Quality (3) Ecosystem Vitality (14): Water Resources (1), Biodiversity and Habitat (4), Forestry (1), Fisheries (2), Agriculture (2), Climate Change & Energy (4)	20	Proximity to Target Winsorization Logarithmic Transformation Pearson's Correlation Coefficient Linear Regression Sensitivity Analyses	Even if the number of indicators was specified as 20, the index was calculated with 19 indicators. The states' electricity score couldn't be calculated. The political indicator of water, including the goals of environmental health and ecosystem vitality, was removed from both of these goals and instead, the political indicator of water quality was added under the goal of environmental health. A change was made in the three political goals included under the goal of environmental health, and instead, the political indicators of health impacts, water quality, and air quality were added. The number of indicators included in the goal of ecosystem vitality decreased to 14 from 17.
2016	Environmental Health (7): Health Impacts (1) Water and Sanitation (2), Air Quality (4) Ecosystem Vitality (14): Water Resources (1), Biodiversity and Habitat (5), Forestry (1), Fisheries (1), Agriculture (2), Climate Change & Energy (2)	20	Proximity to Target Winsorization Logarithmic Transformation Pearson's Correlation Coefficient Linear Regression	Even if the number of indicators was specified as 20 in the year 2014, the index was calculated with 19 indicators. The states' electricity score couldn't be calculated. The number of indicators under the political category of air quality, included under the goal of environmental health, was increased by 1. While there was no change in the total number of indicators under the political categories included under the goal of ecosystem vitality, the political category of water quality was replaced with the political category of water resources.

Table 1. Continued

2018	Environmental Health (6): Heavy Metals (1) Water and Sanitation (2), Air Quality (3) Ecosystem Vitality (18): Air Pollution (2), Water Resources (1), Biodiversity and Habitat (6), Forestry (1), Fisheries (2), Agriculture (1), Climate Change & Energy (5)	24	Proximity to Target Winsorization Logarithmic Transformation Pearson's Correlation Coefficient Linear Regression	The political category of health effects, included under the goal of environmental health, was removed, and instead, the political category of heavy metals was added. The number of political categories under the goal of ecosystem vitality was increased by one. The political category of air pollution was added.
2020	Environmental Health (7): Heavy Metals (1) Sanitation & Drinking Water (2), Air Quality (3), Waste Management (1) Ecosystem Vitality (25): Pollution Emissions (2), Water Resources (1), Biodiversity and Habitat (7), Fisheries (3), Agriculture (1), Climate Change (8), Ecosystem Services (3)	32	Proximity to Target Logarithmic Transformation Spearman's Correlation Coefficient Linear Regression	The political indicator of waste management was added under the goal of environmental health. The name of the political indicator of air pollution, included under the goal of ecosystem vitality, was amended to pollution emission. The name of the political indicator of forestry was amended, and it was named ecosystem services. Moreover, the name of the political indicator of climate change and energy was amended to climate change.
2022	Environmental Health (13): Heavy Metals (1) Sanitation & Drinking Water (2), Air Quality (7), Waste Management (3) Ecosystem Vitality (18): Biodiversity and Habitat (7), Ecosystem Services (3), Fisheries (3), Acid Rain (2), Agriculture (2), Water Resources (1) Climate Change (9): Climate Change Mitigation (9)	40	Proximity to Target Logarithmic Transformation Spearman's Correlation Coefficient Linear Regression	The political indicator of climate change, included in the index in previous years under the goal of ecosystem vitality, had become a goal. The number of goals was increased to three for the first time, and the EPI was formed with these many indicators for the first time.