Ecological Indicators 132 (2021) 108281

Contents lists available at ScienceDirect

Ecological Indicators

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

Original Articles

Monitoring the environmental sustainability of countries through the strong environmental sustainability index

Arkaitz Usubiaga-Liaño *, Paul Ekins

Institute for Sustainable Resources, University College London, UK

ARTICLE INFO

Keywords:
ESGAP
Strong Environmental Sustainability Index
Environmental sustainability
Sustainability gap
Environmental indicators

ABSTRACT

Countries still lack adequate metrics to monitor environmental sustainability across a range of relevant environmental and resource issues. The Strong Environmental Sustainability Index (SESI), which is based on the Environmental Sustainability Gap (ESGAP) framework, is intended to fill this gap. SESI is the result of aggregating 21 indicators across different dimensions. Each of the underlying indicators is related to the functions of natural capital and normalised using science-based targets. SESI uses the geometric mean to aggregate in order to reflect the limited substitutability between the functions of natural capital.

The results of the index, which is computed for 28 European countries, show that several functions of natural capital are impaired in Europe. Countries tend to perform worse in indicators related to pollution and ecosystem health, compared to indicators that describe the provision of natural resources, and human health and welfare. Because the results are sensitive to assumptions in the normalisation, weighting and aggregation processes, the relevant choices have been aligned with the theoretical underpinnings of the ESGAP framework. SESI responds to the demands of the 'Beyond GDP' community on the need for a single environmental sustainability metric that can complement GDP in its (mis-)use as a headline indicator for development.

1. Introduction

Metrics are a key part of environmental governance. Among their key uses are provision of information on the state of the environment, identification of key factors behind environmental problems, comparison of different countries' performance over time, monitoring the effects of policies and progress towards their objectives, and raising awareness on environmental issues (EEA, 1999). Over time, the phrase "we cannot manage what we cannot measure" has become part of the vocabulary of those using quantitative tools to produce policy-relevant information. Although the statement cannot be taken as an absolute truth, it highlights the relevance of having a clear and scientifically-sound information base around which decisions can be made.

The recent Global Environmental Outlook report (UN Environment, 2019) concluded that current paths of economic development will lead to unprecedented hardship for billions of people, as the most basic systems that support human life on Earth start to unravel. From this outlook, it is clear that the current development model is far from being environmentally sustainable. And yet, despite the existence of hundreds of environmental metrics, countries still lack resonant and robust metrics to monitor their environmental sustainability performance across a range of relevant environmental and resource issues (Usubiaga-Liaño and Ekins, submitted). Instead, countries tend to measure their performance against policy targets and best performing countries, as opposed to science-based sustainability reference values that represent the level at which the functioning of natural capital is not impaired. Thus, unless policy targets and best performers represent environmental sustainability conditions (as currently frequently they do not), a measurement gap remains.

The environmental sustainability gap (ESGAP) framework (Usubiaga-Liaño and Ekins, submitted) is an extension and update of the original sustainability gap framework proposed by Ekins (Ekins and Simon, 1999; Ekins et al., 2003; Ekins, 2011) two decades ago to address this metric gap. The framework uses the concepts of strong sustainability, critical natural capital, environmental functions and science-based reference values to provide the theoretical basis to develop policy relevant indices of environmental sustainability. To date the ESGAP framework has only been implemented in a case study 20 years ago (Ekins and Simon, 2001). This paper presents the first comprehensive effort to implement the framework since then.

---

* Corresponding author.
E-mail address: arkaitz.usubiaga.15@ucl.ac.uk (A. Usubiaga-Liaño).

https://doi.org/10.1016/j.ecolind.2021.108281
Received 30 June 2021; Received in revised form 20 September 2021; Accepted 8 October 2021

Table 1
Functions of natural capital and environmental sustainability principles.

Function	Objective	Principle
Source	Maintain the capacity to supply resources	Renew renewable resources Use non-renewables prudently
Sink	Maintain the capacity to neutralise wastes, without incurring ecosystem change or damage	Prevent global warming and ozone depletion Respect critical levels and critical loads for ecosystems
Life support	Maintain the capacity to sustain ecosystem health and function	Maintain biodiversity and ecosystem health
Human health and welfare	Maintain the capacity to maintain human health and generate human welfare in other ways	Respect standards for human health Conserve landscape and amenity

Source: Adapted from Ekins and Simon (1999); Ekins et al. (2003).

Here we describe and compute the Strong Environmental Sustainability Index (SESI). SESI is an index that normalises, weights and aggregates indicators of environmental sustainability that use science-based sustainability reference values to measure absolute country performance across different environmental and resource issues related to the functions of natural capital. Thus, it provides a static perspective on the environmental sustainability of nations. SESI has been built following the guidance provided by the most comprehensive manual on composite indicators (OECD and JRC, 2008; JRC, 2019), which was not available for the original effort two decades ago. The main goal of the index is to provide an easy-to-understand message about the environmental sustainability of countries to decision makers and the general public. In this paper, SESI is calculated for the 27 European Member States and the UK (hereinafter Europe for readability purposes). A different indicator, the Strong Environmental Sustainability Progress Index, which is currently under development, provides a dynamic – instead of a static – perspective by measuring progress over time.

Against this background, the paper is structured as follows. Section 2 summarises the main features of the ESGAP framework. Section 3 summarises the indicator selection process, while section 4 describes the methodology followed to compute the index. Sections 5 and 6 present and discuss the results respectively. Section 7 concludes.

2. The ESGAP framework

Natural capital is a key contributor to human well-being. Its environmental functions that make this contribution can in many cases not be replaced by those provided by other types of capital such as manufactured, social and human capital. The elements of natural capital that fulfil irreplaceable functions are termed ‘critical natural capital’. The limited substitution capacity between different types of capital and between the different elements of natural capital is at the core of the strong sustainability proposition (Costanza and Daly, 1992; Neumayer, 2003). Depending on the substitution capacity assumed, other positions exist, e.g. weak sustainability through to very strong sustainability (Turner, 1993).

There are different typologies of environmental functions. The ESGAP framework groups them in four main categories (Ekins et al., 2003):

- *Source functions* represent the capacity of natural capital to sustain the supply of biotic and abiotic natural resources.
- *Sink functions* represent the capacity of natural capital to neutralise wastes without incurring ecosystem change or damage.
- *Life support functions* refer to the capacity of natural capital to maintain ecosystem health and function.
- *Human health and welfare functions* represent the capacity of natural capital to provide other services to humans, very often of a non-economic kind, which maintain health and contribute to human well-being in other ways.

In this context, environmental sustainability requires maintaining important environmental functions and consequently maintaining the capacity of the capital stock to provide those functions (Ekins et al., 2003). Despite progress in linking environmental functions to human well-being, it is still difficult to isolate the specific functions that need to be maintained in different social and geographical settings. Hence, the framework proposes a set of sustainability principles that can help with this task (Table 1).

From these principles a set of reference values intended to reflect environmental sustainability conditions can be derived. These reference values are termed ‘environmental standards’. Environmental standards differ from environmental policy targets in that they are primarily science-based and therefore are less normative. After all, environmental policy targets are the result of a process in which, beyond scientific considerations, other factors such as economic cost, technological feasibility or social acceptance are weighted.

All these concepts are further elaborated by Usubiaga-Liaño and Ekins (submitted). When brought together, they provide the basis for construction of SESI and SESPI.

3. Indicator selection and structure

3.1. Criteria for selection

The selection of indicators is a critical step in the construction of an index. Different criteria can be used to select metrics to populate indicator systems or the structure of indices (e.g. Srebotnjak et al. (2009); UNSD (2015); Eurostat (2020)). Here we adapt the criteria used by Eurostat for their 2020 SDG indicator set, which considers the relevance, the statistical and methodological soundness, and the data quality of indicators in the selection process. These criteria are interpreted as follows.

Relevance considers the alignment of the indicators to the theoretical framework and their capacity to reflect environmental sustainability, which is defined as the “maintenance of important environmental functions and therefore, the maintenance of the capacity of the capital stock to provide those functions” (Ekins et al., 2003). Thus, for an indicator to be relevant it has to meet three sub-criteria. First, the indicator needs to be linked to the environmental functions of natural capital: source, sink, life support, and human health and welfare. Second, an appropriate reference value is required against which performance can be measured. That reference value should be defined through science-based environmental standards that ultimately represent the conditions under which the functioning of natural capital is not altered in a way that it threatens its capacity to provide ecosystem services in the long-term. Third, the indicator must be relevant at the national level, for this is the geographical scale at which SESI is produced.

Statistical and methodological *soundness* reflects the readiness and sustainability of statistical production, the methodological soundness, accessibility and transparency and compliance with existing methodological standards. In practice, this can be assumed to be true when the indicators are sourced from official statistical offices and well-established international institutions.

Data *quality* considers aspects related to the frequency of dissemination, timeliness, time and geographical coverage and data comparability. These aspects are checked through a score-based system as explained in the supplementary material.

3.2. Selection of indicators

The selection of indicators is documented in detail in the

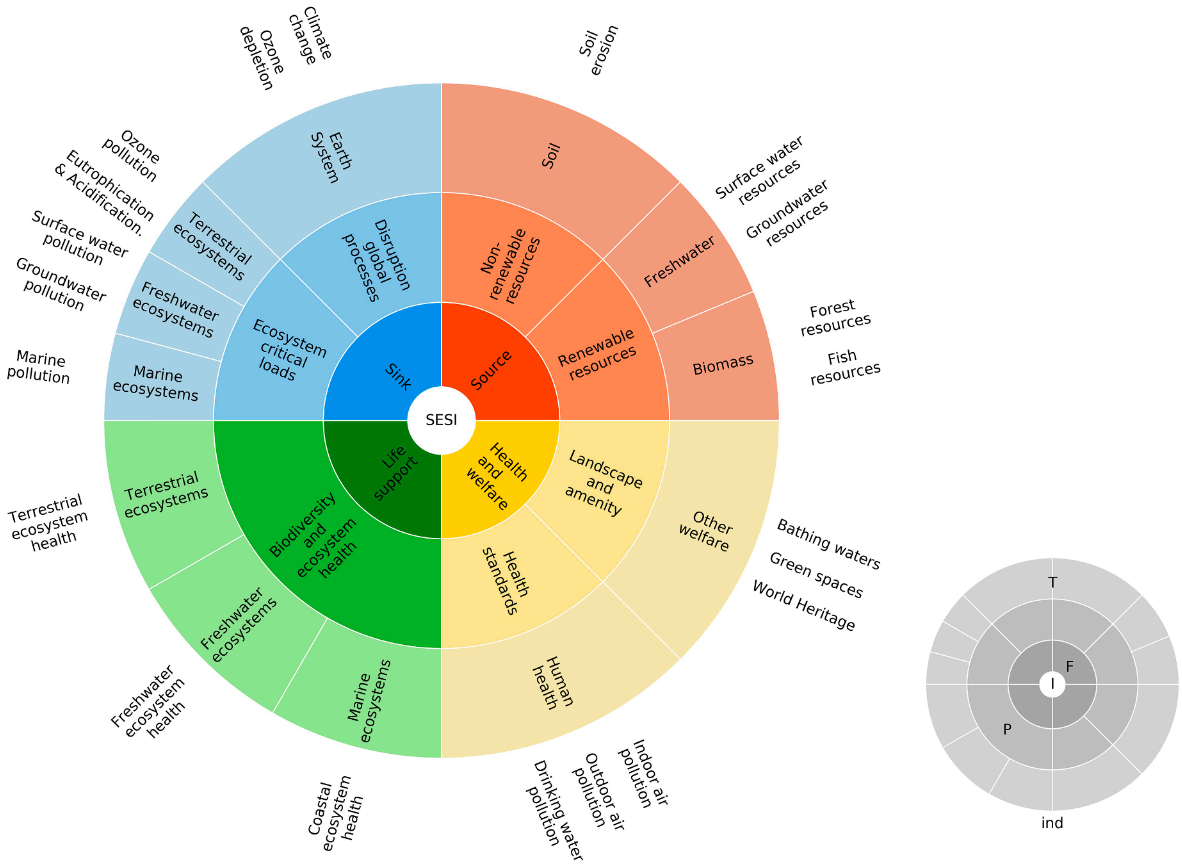

Fig. 1. Structure of the Strong Environmental Sustainability Index.
Note: I (index), F (function), P (principle), T (topic), ind (indicator).

Table 2
List of indicators in the Strong Environmental Sustainability Index.

Function	Indicator
Source	Forest utilization rate
	Fish stocks within safe biological limits
	Freshwater bodies not under water stress
	Groundwater bodies in good quantitative status
	Area with tolerable soil erosion
Sink	CO_2 emissions
	ODS consumption
	Cropland and forest area exposed to safe ozone levels
	Ecosystems not exceeding the critical loads of eutrophication and acidification
	Surface water bodies in good chemical status
	Groundwater bodies in good chemical status
	Coastal water bodies in good chemical status
Life support	Terrestrial habitats in favourable conservation status
	Surface water bodies in good ecological status
	Coastal water bodies in good ecological status
Human health and welfare	Population exposed to safe levels of outdoor air pollutants
	Population using clean fuels and technologies for cooking
	Samples that meet the drinking water criteria
	Recreational water bodies in excellent status
	Population with nearby green areas
	Natural and mixed world heritage sites in good conservation outlook

supplementary material. It uses a stepwise process in which an initial list of indicators is assessed against the relevance, the statistical and methodological soundness, and the data quality criteria above.

The initial list comprises 30 indicators linked to functions of natural capital. Rather than evaluating the relevance of hundreds of environmental indicators, this initial selection was informed by a literature review on environmental standards and complemented by feedback received in different meetings. After checking for relevance, the initial list was reduced to 23 indicators. In the statistical and methodological soundness check, three indicators were replaced by proxies because the original indicators were not reported by relevant institutions. After the data quality check, 21 indicators remained.

Each of the 21 indicators shows whether a specific element of natural capital is managed sustainably in that its functioning is not altered in a way that threatens its capacity to provide ecosystem services in the long-term. In order to do so, each indicator is matched to an environmental standard that represents a sustainable reference value. These standards are taken from the scientific literature when possible, thereby avoiding having to make additional judgements. When appropriate, the standards are based on relevant international environmental agreements and EU-level environmental legislation if these are based on scientific considerations. The indicators have been structured around the environmental functions and sustainability principles described in the theoretical framework and shown in Fig. 1. Table 2 below contains the full name of the indicators, while Table 3 in the appendix provides information on the indicators, their data sources and the environmental standards that have been used in each case. Each indicator and the environmental standards are further described in fiches in the supplementary material.

Indicators for the source function cover renewable and non-renewable resources. Renewable resources include forest, fish, groundwater and freshwater resources. The environmental standards for these indicators tend to describe exploitation rates (i.e. extraction vs annual availability) that are deemed environmentally sustainable (Raskin et al., 1997; EC, 2009; EEA, 2017), except in the case of fish resources, which represents an exploitation status that uses criteria on fishing mortality and spawning stock biomass to define overexploitation (EC, 2010).

Indicators of non-renewable resources are restricted to soil resources, in this case represented through soil erosion. The environmental standard is defined as the tolerable soil erosion rate (Verheijen et al., 2009). Other factors such as the content of organic matter, salinization and sealing are also linked to the functioning of soils, but these lack a credible environmental standard (Loveland and Webb, 2003; Huber et al., 2008). Indicators on the extraction of raw materials are also missing from the final selection because of the lack of environmental standards related to scarcity.

Indicators for the sink function cover those related to global processes and those related to regional or local processes. In the former, we have translated global standards related to climate change and stratospheric ozone depletion into emission levels relevant at the national scale. Regarding the latter, we consider indicators related to the exceedance of critical levels and critical loads of pollutants in terrestrial, freshwater and coastal ecosystems. In the absence of data on marine ecosystems, coastal ecosystems have been chosen instead.

Indicators of life support function take the form of composite indicators of ecosystem condition that consider different parameters to determine whether an ecosystem is in good condition or not. Biodiversity indicators would also feature in this category, but appropriate metrics and/or environmental standards at broad scales are commonly missing (Mace et al., 2014).

Indicators for human health and other welfare functions cover human exposure to environmental factors as well as indicators associated with amenity and landscape value. The first group covers outdoor and indoor exposure to $PM_{2.5}$ and exposure to water pollution. There are many more substances not covered in the selected indicators that can lead to harmful effects on human health (e.g. persistent organic pollutants, pesticides, etc.), although air pollution and drinking water quality are among the most relevant environmental factors behind health issues (Landrigan et al., 2018). The functions related to amenity and landscape value are represented by standards on the quality of bathing water bodies, the population with nearby green areas next to dwellings and the conservation outlook of relevant World Heritage sites. It should be noted that the indicators selected fall short from covering all non-use values of natural capital, which are not only difficult to capture through indicators, but in many cases also lack science-based environmental standards.

All in all, there are five indicators for the source function, seven for sink, three for life support and six for human health and welfare. Although at first sight, the difference in the number of indicators assigned to each function might seem striking, it should be noted that some of the indicators in the sink and life support functions are composite metrics of ecosystem condition, each of which consider dozens of parameters. That is the case for those indicators related to ecosystem health and pollution (e.g. conservation status of terrestrial ecosystems, and the chemical and ecological status of water bodies). The exception would be the chemical status of terrestrial ecosystems. Since it was not possible to generate a single composite metric for this one, two separate indicators have been used: one for (tropospheric) ozone pollution and one for eutrophication and acidification. The latter is the result of spatially aggregating the critical load exceedance maps for eutrophication and acidification, and considering as sustainable only the area that is not affected by any one the two pollution types.

4. Methodology

Several steps need to be undertaken to convert indicators into a single index. As mentioned earlier, we used the OECD manual on composite indicators (OECD and JRC, 2008) to guide the construction of SESI. There are four steps between the selection of indicators and the generation of the results. Data treatment covers filling data gaps and dealing with outliers. The normalisation process requires all the indicators to be transformed into a comparable scale. After that, weights are assigned to the indicators and the dimensions to which they are allocated, and lastly the normalised scores are aggregated across dimensions into a single unitless score. In addition to the results, two additional steps are recommended in the manual. The first one is an analysis of the statistical and conceptual coherence of the index. In short, this analysis intends to shed light on whether the choices made during the selection of indicators and the construction of the index are aligned with the theoretical framework on the one hand, and on the extent to which the information contained in the individual indicators is transferred across layers into the final index, or if, on the contrary, some information is lost in the process. The second step recommends users to undertake an uncertainty analysis to understand how the choices made while constructing the index affect the final results.

Because of space limitations, in the main text we focus on the normalisation, weighting and aggregation steps, as well as on the uncertainty analysis. The supplementary material also includes a description of the data treatment approach used to fill the few data gaps in the dataset and of the statistical and conceptual coherence analysis undertaken.

4.1. Normalisation

Most of the indicators in an index usually have different units, which makes them incomparable unless transformed into a common unitless scale. This is the goal of the normalisation process. There are multiple normalisation methods, so the selection of a method is not trivial. The relevance of environmental standards in the conceptual framework of SESI demands the goalpost method to be used in the normalisation process. In this method, user-defined values are used as upper and lower bounds to transform indicators into a scale between 0 and 100. For the normalisation process to be aligned with the strong sustainability narrative, these upper and lower bounds need to be consistent with environmental standards.

The normalised scores are calculated as shown in equation below, where the normalised value of an indicator (NI) depends on the value of the indicator (I), and the minimum and maximum values assigned as goalposts (gp_{min} and gp_{max}). Scores lower than 0 and higher than 100 are assigned 0 and 100 values.

$$NI = 100\frac{I - gp_{min}}{gp_{max} - gp_{min}} \tag{1}$$

4.2. Weighting

The weights assigned to the indicators are a reflection of their importance, yet this does not necessarily represent how much they impact the final index score (Becker et al., 2017). When assigning weights to environmental functions, life support functions should take preference over source, sink and human health and welfare functions because without life support functions, the functions for humans would not be able to be sustained in the long-term. At lower levels, prioritising sustainability principles becomes more difficult. For instance, in the source functions, the relevance of renewable and non-renewable resources depends on the domestic endowments. In sink functions, prioritising global vs regional pollution neutralisation processes is not straightforward. In the case of human health and welfare functions, human health should come before the functions related to other aspects of human welfare. At the level of indicators, it becomes almost impossible to assign weights based on relevance, since different natural capital endowments and the uneven contribution of pollutants to overall environmental and health impacts differ considerably between countries.

In absence of a method, or a consensus criterion that would allow translating these arguments into weights, equal weights are assigned to all the indicators and dimensions from top to bottom.

4.3. Aggregation

The concepts of ‘strong sustainability’ and ‘critical natural capital’ are at the core of the ESGAP framework. In combination, both concepts address the substitution capacity between natural capital and other types of capital, as well as between the different functions of natural capital. The limited substitutability between the different types of capital is reflected in that SESI comprises different indicators that address the environmental dimension of sustainable development. In order to incorporate the limited substitutability between the functions of natural capital into SESI, we have selected a weighted geometric mean for aggregating the normalised scores across layers as shown in Equation (2), where i represents one of n indicators in the dimension, and α the weight assign to it. The sum of the weights in each dimension equals one.

$$\prod_{i=1}^{n} NI_i^{\alpha_i} \tag{2}$$

A geometric mean penalises low performances as opposed to the weighted arithmetic mean, where a poor performance in one dimension is linearly compensated for by high achievement in another dimension and therefore it implicitly assumes that the functions provided by natural capital are interchangeable. Nonetheless, it is important to bear in mind that the use of the geometric mean in some contexts also has its drawbacks. In this case, the main limitation of the geometric mean is that it collapses to zero when any indicator has a value of zero. This is commonly dealt with by replacing zeros and small values by a user-defined value. In this case, we have chosen to replace all the values lower than five, by five. The rationale and argumentation are described in the supplementary material.

4.4. Uncertainty analysis

The construction of an index requires making assumptions on the different steps of the process. From the inception to the generation of the results, different methods and strategies exist to select indicators and to transform the raw data into a single score. In this context, understanding the effects of the choices made is critical to properly interpret the results. For this reason, in the uncertainty analysis we have opted for testing different approaches in key steps such as the normalisation, weighting and aggregation, since the assumptions made in these steps translate key parts of the conceptual framework from theory to practice.

In the normalisation process, we tested the min–max method, which measures performance relative to the best and worst performers. Thus, min–max represents relative performance, as opposed to the goalpost method, which measures absolute performance against environmental standards. The min–max method has been implemented through Equation (3). I_{min} and I_{max} values are calculated as the 2.5th and 97.5th percentiles of the values of each indicator across the 28 European country sample.

$$NI = 100\frac{I - I_{min}}{I_{max} - I_{min}} \tag{3}$$

In the weighting process, we increased the weights of the life support functions, arguably the most important functions. Thus, we undertook two tests with the following weights:

- Life support: 0.4, others 0.2
- Life support: 0.7, others 0.1

Last, in the aggregation we tested the arithmetic mean (as opposed to the geometric mean) and the minimum normalised score across the individual indicators, which represent weak and very strong sustainability propositions as defined by Turner (1993) (i.e. lack of substitution capacity).

The different normalisation, weighting and aggregation options described above were tested separately for the index scores at European level.

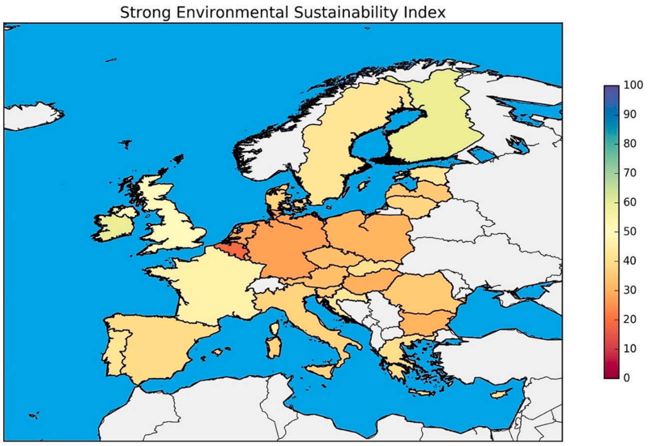

Fig. 2. Strong Environmental Sustainability Index score for 28 European countries.
SESI scores countries from 0 to 100 in terms of their environmental sustainability performance. A score of 100 indicates the compliance of all the indicators across the four environmental functions with their corresponding environmental standard. A score of 0 indicates the opposite.

5. Results

5.1. Index scores

Fig. 2 shows the index score of the 28 European countries according to their most recent data point. The Anglo-Celtic isles and the Scandinavian countries seem to perform better than Mediterranean, and central and eastern European countries. Nonetheless, the absolute scores are low in most cases, suggesting that one or more environmental functions are currently jeopardised in many countries. Only three countries score more than 50 points and the maximum score is 60, which is obtained by Finland. After the frontrunners, 18 countries obtain scores between 30 and 45, while six countries score lower than 30, with Belgium being at the bottom with 19 points. When considered as a block, Europe gets a score of 47. Of course, at the index level, the score is influenced by the use of geometric mean in the aggregation, since this penalises low performances in individual indicators. Thus, countries that perform poorly in several indicators will see their aggregate score reduced, thereby reflecting the limited substation capacity between the environmental functions represented by the indicators.

As with any index, the total score can hide disparities in the performance at lower levels of aggregation. In this context, Fig. 3 shows country scores for the four broad environmental functions and the seven sustainability principles used to characterise environmental sustainability. Countries perform very differently in source, and human health and welfare functions, with countries in the first positions scoring relatively high in those two functions. In the source function, which covers the provision of forest and fish biomass, surface and groundwater, and soil, former Soviet Union and Scandinavian countries hold the first five positions with scores over 70. Most countries obtain scores between 40 and 65. The European block sits at the upper side of the range with a score of 62. Former Soviet Union and Scandinavian countries, as well as the Anglo-Celtic isles are the frontrunners in the human health and welfare function. Countries such as Finland, Sweden and Ireland score over 90. This means that these countries almost comply with the science-based standards used for (indoor and outdoor) air pollution, drinking water, bathing waters, access to green spaces and the conservation of relevant World Heritage sites. Europe obtains a score of 64 in this category.

The sink and life support functions describe a different picture. Scores are more homogeneous with almost every country performing poorly. In the case of sink functions, none of the countries reaches 50

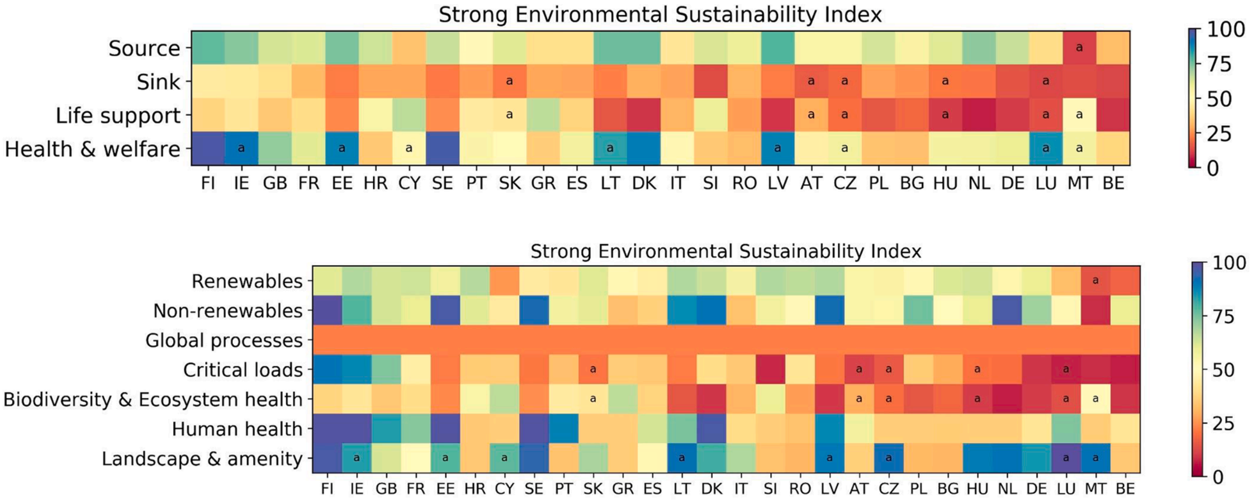

Fig. 3. Heatmap of the country index scores by (a) environmental function, (b) sustainability principle.
The top figure shows the scores of each country for the four environmental functions. The bottom figure shows the scores of each country for seven sustainability principles. Dark red indicates low scores, while dark blue indicates high scores. Countries are sorted by the total index score from higher to lower. The labels in the y axis in the bottom figure are equivalent to the following principles. Renewables: renew renewable resources; Non-renewables: use non-renewables prudently; Global processes: prevent global warming; Critical loads: respect critical loads for ecosystems; Biodiversity & Ecosystem health: maintain biodiversity and ecosystem health; Human health: respect standards for human health; Landscape & amenity: conserve landscape and amenity. The label [a] in the heatmap indicates that one of the indicators assigned to the function or principle is blank because it does not apply to the country (e.g. coastal areas in landlocked countries). Country codes (in alphabetical order): AT: Austria, BE: Belgium, BG: Bulgaria, CY: Cyprus, CZ: Czech Republic, DE: Germany, DK: Denmark, EE: Estonia, ES: Spain, FI: Finland, FR: France, GB: United Kingdom, GR: Greece, HR: Croatia, HU: Hungary, IE: Ireland, IT: Italy, LT: Lithuania, LU: Luxembourg, LV: Latvia, MT: Malta, NL: Netherlands, PL: Poland, PT: Portugal, RO: Romania, SE: Sweden, SI: Slovenia, SK: Slovakia.

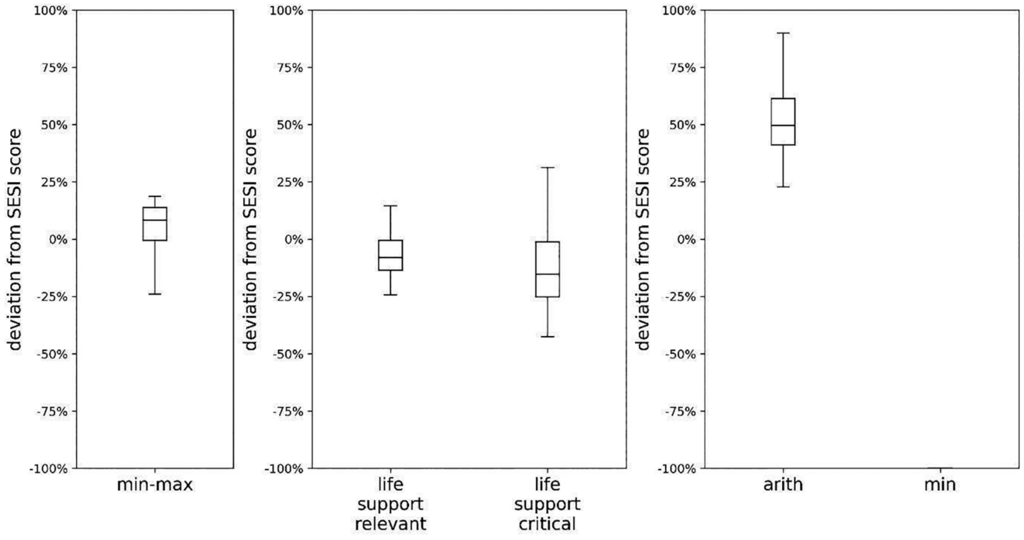

Fig. 4. Distribution of differences between the default method and alternative methods associated with the (a) normalisation, (b) weighting and (c) aggregation processes.
The upper and lower edges of the rectangle in the boxplot represent the 75th and 25th percentiles, while the top and bottom markers represent the maximum and minimum values. In the middle figure, 'life support relevant' assigns 0.20–0.20–0.40–0.20, while 'life support critical' assigns 0.10–0.10–0.70–0.10. In the last figure, 'arith' means arithmetic mean; and 'min' means minimum value.

(Europe scores 33). 23 countries have scores below 30. In the countries with the highest scores, the main explanatory factor is the poor performance in CO_2 emissions, where none of the countries are in the sustainable emission range (0.5–2.5 t CO_2 per capita). In the remaining countries, poor performance in CO_2 emissions is combined with poor performance in terrestrial, freshwater or coastal ecosystems' pollution. Scores in life support functions are also generally low with 23 countries getting scores under 50. Five Mediterranean countries (Cyprus, Greece, Slovenia, Croatia and Malta) are the top performers with scores up to 66. This is driven by high scores in the ecological condition of their coastal ecosystems. It should be noted that this indicator is used as a proxy for marine ecosystems, which currently lacks data. Given that Mediterranean fish stocks are largely overexploited, it seems unlikely that countries will report such good conditions in marine ecosystems.

Because of the format of SESI, the information provided at the higher levels of its structure is a summary of the information provided at indicator level. For space issues, the results at the level of indicators are reported in the supplementary material.

5.2. *Uncertainty analysis*

The results of the uncertainty analysis are shown in Fig. 4.

Environmental sustainability demands measuring absolute performance, which can be interpreted as calculating normalised country scores in relation to science-based environmental standards. SESI uses the goalpost method for such a task. The left side figure shows the effects of measuring relative performance instead of absolute performance. Relative performance is represented through the min–max method where country scores are a function of the best and worst performers. As shown in the figure, relevant deviations are seen in both directions when using the min–max method to normalise the underlying indicators (median = 8.34%).

Different sets of weights have been tested by assigning more weight to the life support function. As shown in the figure in the centre, changing the sets of weights leads to lower index values for most countries. The more weight is assigned to the life support function (0.4 under 'life support relevant', and 0.7 under 'life support critical'), the lower the overall index score. This is the result of assigning more weight to a dimension in which countries tend to perform worse than in others.

The aggregation method used also affects the results of SESI at country level as shown in the right-hand side figure. At country level using the arithmetic mean, the median increase is 49.68%. Using the minimum indicator value, which assumes no substitution capacity at all between the functions of natural capital, completely disrupts the results, since all the countries have a normalised value of zero for the indicator on CO_2 emissions. As a result, the index value will be zero for all the countries.

6. Discussion

6.1. A novel index of environmental sustainability

SESI is a single metric that represents the extent to which countries comply with environmental standards intended to represent the conditions under which the functioning of natural capital is not threatened in the long term. The index comprises indicators that are aggregated across different layers: from indicators to topics, from topics to sustainability principles, from principles to the main function of natural capital (source, sink, life support, and human health and welfare), and from the latter to a single index. Each of these indicators measures absolute performance against a science-based environmental standard, thereby showing whether specific functions of natural capital are potentially compromised.

The selection of indicators has been done based on the relevance of the indicator, its statistical and methodological soundness, and data quality. This allowed shortening the initial list of 30 candidates into the 21 indicators that form SESI. The relevance criterion required the indicator to (a) be related to the functions of natural capital, (b) to have a science-based reference value against which performance could be measured, and (c) to be defined at the national level. Relevance assesses consistency with the definition of environmental sustainability adopted in the ESGAP framework. Statistical and methodological soundness, as well as data quality considered more generic criteria that can be applied to any other index. While the 21 indicators cover quite a lot of ground, there are some topics such as extraction of abiotic raw materials, organic soil matter, marine systems or several aspects of human welfare that are not covered in this version of the index. This is the result of a lack of environmental standards or data for the relevant indicators. As knowledge on environmental standards and data availability improves, the list of indicators should be revised.

SESI covers a space in sustainability science that none of the most widely known environmental (sustainability) indices covers. The main novelty of SESI is the use of science-based environmental standards to measure the absolute environmental sustainability performance of countries. Other indices such as the Environmental Performance Index or the environmental dimension of the SDG indicators either measure absolute performance against policy or international targets or relative performance against frontrunners (Usubiaga-Liaño and Ekins, submitted). Although some of the targets used are aligned with environmental standards, this is not the general rule. The information contained in these alternative indices is useful in many ways, but these metrics do not allow the assessment of the environmental sustainability of nations from a strong sustainability perspective.

Furthermore, it is worth noting that the index has been built following the guidance provided by the most comprehensive manual on composite indicators (OECD and JRC, 2008; JRC, 2019), and therefore contains additional complementary analyses related to its conceptual and statistical soundness (see supplementary information), which is already a distinctive feature compared to other metrics that tend to focus on the main results (Kwatra et al., 2020).

6.2. Environmental sustainability in Europe

Our results suggest that the functioning of different elements of natural capital is impaired as a result of excessive environmental degradation in Europe. Most European countries obtain index scores below 50, including the European block, which scores 47 points. In this context, it is important to bear in mind that only a score of 100 reflects compliance with the environmental standards of each of the 21 indicators selected to represent relevant environmental functions of natural capital. Even in the case of the highest scoring country, Finland, the gap between the current and sustainable conditions is of 40 points.

Performance across environmental functions is quite uneven, with those related to environmental integrity being the most affected. In the sink function, countries perform very poorly with regard to CO_2 emissions and the chemical pollution of ecosystems, especially terrestrial ecosystems. Scores are also very low in the life support function, arguably the most important function of all, which covers biodiversity and ecosystem health. Functions associated with the provision of resources seem to be in better shape than those associated with the neutralisation of waste and life support. One can only hypothesise if the fact that biotic and abiotic resources have a market value can partially explain this pattern, which does not hold in every country. An exception in the source function are fish stocks, which are consistently overexploited across countries.

Countries tend to obtain relatively high scores when health standards are on the line as in the case of drinking water and indoor air pollution. Outdoor air pollution is an exception, arguably because the policy targets set are more permissive than the guideline values proposed by the World Health Organisation. When it comes to the amenity function, countries tend to have high scores in relation to bathing sites and access to green spaces, while with World Heritage sites, performance is very uneven, with many countries not having any natural site within their territory.

The interpretation of the results needs qualifications on several grounds. First, the index provides a snapshot of whether countries meet science-based environmental standards from a territorial perspective across a variety of environmental and resource issues. While doing so, the indicators that form the index represent whether or not the environmental standards have been transgressed, but do not capture the severity of the transgression. For instance, the outdoor air pollution indicator represents the percentage of the population that is exposed to $PM_{2.5}$ concentrations higher than the guideline values proposed by the World Health Organisation. In theory, it would be possible for two countries to have the same normalised score (e.g. 75), while in the first country a quarter of the population is exposed to air pollution levels slightly above the environmental standards, while in the second a quarter of the population is exposed to air pollution levels that are several times higher than the environmental standard.

Second, the indicators that form the index adopt a territorial perspective, as opposed to the consumption perspective that is characteristic in environmental footprint indicators. SESI seeks foremost to be useful for policy making, and therefore is restricted to the elements of natural capital that can be most easily influenced by policy makers. Nonetheless, consumption-based indicators can provide a complementary perspective to the results.

Third, the environmental standards used to characterise environmental sustainability have either been taken from the scientific literature or from relevant environmental legislation informed by expert input. Nonetheless, the standards do not have a homogeneous meaning in that they can refer to acceptable health risks, acceptable environmental impacts, precautionary expert guesses or safe distance from tipping points. Thus, the level of consensus around the standards chosen differs considerably. In all cases though, their transgression flags a potential problem that requires further policy attention. As the knowledge base improves, existing environmental standards might change, or new ones might be formulated. Likewise, it is important to bear in mind that potential trade-offs might arise when trying to meet environmental standards. For instance, the reduction of CO_2 emissions through bioenergy and carbon capture and storage would have negative impact on terrestrial habits (Heck et al., 2018). Thus, interventions intended to address the environmental and resource issue covered in SESI should consider the potential consequences they might have in other areas.

Lastly, SESI provides a static perspective on countries' environmental sustainability, and therefore fails to reflect whether progress towards the standard is being made over time. The Strong Environmental Sustainability Progress Index (currently under development) is intended to fulfil this role.

6.3. Choices in the construction of the index matter

Indices have the potential to help make sense of complex systems through numbers. Nonetheless, the big picture they intend to show can be unintendedly distorted or even manipulated if the choices made during the construction of the index are not clear or properly justified (Greco et al., 2019). SESI is not exempt of such risk and, therefore, its construction has been guided by the most comprehensive manual on composite indicators (OECD and JRC, 2008). The computation of SESI required several methodological choices to be made in relation to data treatment, normalisation, weighting, and aggregation. When possible, key choices related to the indicator selection, normalisation or the aggregation have been aligned with the key features of the ESGAP framework, which reflects a more accurate and restrictive vision on the concept of strong sustainability as opposed to other indices such as the Sustainable Development Goals Index or the Environmental Performance Index.

The goal of the uncertainty analysis undertaken was to understand how choices in the construction of the index affect the results and the narrative developed based on them. Thus, different choices related to the normalisation, weighting, treatment of zeros and small values, and the aggregation were tested separately. The results show that the index and function scores are affected by these choices. Measuring the absolute performance of countries, which depends on the environmental standards, as opposed to measuring the relative performance of countries, which depends on the frontrunners and laggards leads to lower index scores, although functions and indicators are affected differently.

With regard to the substitution capacity between the functions of natural capital, SESI uses the geometric mean with treatment of zeros and small values to represent a limited capacity in line with the strong sustainability discourse. Assuming full substitutability through aggregating with arithmetic means or no substitutability with the adoption of the minimum normalised score of any indicator as final index score significantly impacts the results. The use of the arithmetic mean leads to higher scores, especially in the functions in which countries perform worse. This makes it more challenging to identify which functions of natural capital are threatened if the low scores in the underlying indicators are linearly compensated by high scores. On the opposite end, when assuming no substitution capacity between functions, only the information on the worst performance is aggregated, which ultimately limits the usefulness of the index because it omits the information contained in all the other topics covered by the indicators.

The weighting method remains the most controversial choice in the construction of SESI. Equal weights have been assigned to all the indicators and (sub)dimensions, including functions. Indicator weights could be set based on the natural endowments of each country, but this would hinder the comparability of the results. At the level of function, the life support function has been identified as being more relevant than source, sink, and human health and welfare functions, but because of the lack of a credible method to weight each function, equal weights have been used as well. The uncertainty analysis has tested different sets of weights at function level and the results show that their effect is by no means negligible. As in the previous case, showing the function scores alongside the index scores minimises this effect. In any case, the issue of weighting remains unresolved in this version of the index and should be revisited in the future.

All in all, the uncertainty analysis has shown that the choices made during the construction of SESI are not trivial and therefore need to be aligned with the theoretical framework. After all, measuring absolute or relative performance, or assumptions on the substitution capacity between the functions of natural capital do not only have an impact on the results, but also on the narrative built from them.

7. Conclusion

It is remarkable that countries still lack meaningful metrics that allow them to measure their environmental sustainability performance from a strong sustainability perspective. SESI is based on the ESGAP framework, which builds on key concepts such as strong sustainability, critical natural capital, environmental functions and science-based reference values. The limited substitution capacity between different types of capital and between the different functions provided by natural capital, and the notion that some elements of natural capital provide irreplaceable functions are much closer aligned with the biophysical reality that governs the natural system and the socioeconomic systems embedded within, than the concept of weak sustainability, which assumes that the loss of nature can be fully compensated by increases in manufactured, human or social capital. For these reasons, metrics of weak sustainability can be misleading and lead to poor decision making.

Although this first version of SESI can only be considered a proof of concept, it can provide policy-relevant information by helping countries navigate the environmental sustainability agenda beyond single issues and providing scores that allow comparisons, trends analysis and benchmarking across countries. In this context, SESI provides a snapshot of the absolute performance of countries against environmental standards intended to represent whether the capacity of natural capital to provide ecosystem services is compromised. As a result, SESI provides a different perspective on the environmental sustainability of nations compared to most environmental indices and indicator systems that tend to measure the performance of countries against their peers or against policy targets, rather than science-based reference values.

SESI and the sub-indices for environmental functions (source, sink, life support, and health and human welfare) could be used as headline indicators when assessing progress towards sustainable development at country level, thereby complementing the narratives around social and economic welfare. A single metric such as SESI shows the absolute performance of countries with regard to environmental sustainability and responds to the demands made from the 'Beyond GDP' community on the need for a single environmental sustainability metric that can complement GDP in its (mis-)use as a headline indicator for development.

In the future, feedback provided by different stakeholders as well as an increased availability of relevant data or scientific evidence that supports changes in existing environmental standards or the inclusion of different ones will require the structure and indicator selection of SESI to be revisited. Hopefully, the robustness of the framework and the potential applications of SESI will create the momentum for such review of the evidence and for relevant data to be generated. In the meantime, SESI is being computed in New Caledonia, Vietnam, Kenya, Japan and China considering their own data availability. In contrast to the work presented here, which intends to make the results as comparable as possible across countries, these case studies are adapting the ESGAP framework to their own national context in order to maximise its policy

Table 3
Final indicator set for SESI.

Function	Principle	Topic	Indicator [Unit]	Data	Standard	References
Source	Renew renewable resources	Biomass	Forest utilization rate [%]	Forest Europe (2020)	Fellings / Net Annual Increment	EEA (2017)
			Fish stocks within safe biological limits [%]	EEA (2019b)	Fishing mortality consistent with Maximum Sustainable Yield Spawning stock biomass consistent with Maximum Sustainable Yield	EC (2010)
		Freshwater	Freshwater bodies not under water stress [%]	EEA (2018b)	Blue water consumption / Mean quarterly flows	Raskin et al. (1997)
			Groundwater bodies in good quantitative status [%]	EEA (2018a)	Good quantitative status as defined in European legislation	EC (2009)
	Use non-renewables prudently	Soil	Area with tolerable soil erosion [%]	Panagos et al. (2020)	Tolerable soil erosion rate	Jones et al. (2004); Huber et al. (2008); Verheijen et al. (2009)
Sink	Prevent global warming, ozone depletion	Earth system	CO_2 emissions [tonnes per capita]	Eurostat (2019)	Long-term CO_2 emissions consistent with a 1.5–2 °C increase in global mean temperature compared to pre-industrial levels.	IPCC (2018)
			ODS consumption [tonnes per capita]	Ozone Secretariat United Nations Environment Programme (2019)	ODS consumption consistent with reducing the ozone hole	UN (1987); EEA (2019a)
	Respect critical levels and loads for ecosystems	Terrestrial ecosystems	Cropland and forest area exposed to safe ozone levels [%]	Horálek et al. (2020)	Critical levels of tropospheric ozone	Karlsson et al. (2003); Karlsson et al. (2007); Mills et al. (2007)
			Ecosystems not exceeding the critical loads of eutrophication and acidification [%]	Tsyro et al. (2020)	Critical load of eutrophication and acidification	CLRTAP (2017)
		Freshwater ecosystems	Surface water bodies in good chemical status [%]	EEA (2018a)	Good chemical status as defined in European legislation	European Parliament and European Council (2008)
			Groundwater bodies in good chemical status [%]	EEA (2018a)	Good chemical status as defined in European legislation	EC (2009)
		Marine ecosystems	Coastal water bodies in good chemical status [%]	EEA (2018a)	Pollution-related elements of good environmental status as defined in European legislation	EC (2017)
Life support	Maintain biodiversity and ecosystem health	Terrestrial ecosystems	Terrestrial habitats in favourable conservation status [%]	EEA (2020)	Favourable conservation status based on range, area, structure and function.	Röschel et al. (2020)
		Freshwater ecosystems	Surface water bodies in good ecological status [%]	EEA (2018a)	Good ecological status as defined in European legislation based on biological, physicochemical and hydromorphological parameters	EC (2003)
		Marine ecosystems	Coastal water bodies in good ecological status [%]	EEA (2018a)	Good environmental status as defined in European legislation based on biological, physicochemical and hydromorphological parameters	EC (2017)
Human health and welfare	Respect standards for human health	Human health	Population exposed to safe levels of outdoor air pollutants [%]	Horálek et al. (2020)	Critical levels of $PM_{2.5}$	WHO (2005)
			Population using clean fuels and technologies for cooking [%]	WHO (2020)	Critical levels of $PM_{2.5}$	WHO (2005)
			Samples that meet the drinking water criteria [%]	EC (2016)	Safe drinking water criteria as defined in European legislation based on microbiological, chemical and other parameters	European Council (1998)
	Conserve landscape and amenity	Other welfare	Recreational water bodies in excellent status [%]	EEA (2019c)	'Excellent' quality criteria as defined in European legislation based on the concentration of Intestinal Enterococci and Escherichia Coli in recreational waters	EC (2002)
			Population with nearby green areas [%]	Poelman (2018)	Green areas that can be reached within 10 min' walking.	Poelman (2018)
			Natural and mixed world heritage sites in good conservation outlook [%]	Osipova et al. (2020)	Good conservation outlook based on three elements: the current state and trend of values, the threats affecting those values, and the effectiveness of protection and management	Osipova et al. (2014)

impact and to more adequately reflect the national natural capital endowment. In the future, we expect a combination of internationally comparable case studies and case studies that are more reflective of individual nation's context.

CRediT authorship contribution statement

Arkaitz Usubiaga-Liaño: Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Paul Ekins:** Writing – original draft, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to thank Thomas Scheusnar (Umweltbundesamt), Panos Panagos (European Commission) and Nihat Zal (European Environment Agency) for sending additional data that was not available online at the time. The authors would also like to thank the staff from the JRC's Competence Centre on Composite Indicators and Scoreboards for providing feedback on earlier versions of the index through the COIN open days.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2021.108281.

References

Becker, W., Saisana, M., Paruolo, P., Vandecasteele, I., 2017. Weights and importance in composite indicators: Closing the gap. Ecol. Ind. 80, 12–22.

CLRTAP, 2017. Mapping critical loads for ecosystems, Chapter V of Manual on methodologies and criteria for modelling and mapping critical loads and levels and air pollution effects, risks and trends. UNECE Convention on Long-range Transboundary Air Pollution.

Costanza, Robert, Daly, Herman E., 1992. Natural Capital and Sustainable Development. Conserv. Biol. 6 (1), 37–46.

EC, 2002. Proposal for a Directive of the European Parliament and of the Council concerning the quality of bathing water. COM(2002) 581 final. Brussels: European Commission.

EC, 2003. Overall approach to the classification of ecological status and ecological potential. No. 13. Brussels: European Commission. Common Implementation Strategy for the Water Framework Directive (2000/60/EC) Guidance Report.

EC, 2009. Guidance Document No. 18. Guidance on Grounwater Status and Trend Assessment. Technical Report - 2009 - 026. Luxembourg: European Commission.

EC, 2010. Commission Decision of 1 September 2010 on criteria and methodological standards on good environmental status of marine waters. Brussels.

EC, 2016. Synthesis Report on the Quality of Drinking Water in the Union examining Member States' reports for the 2011-2013 period, foreseen under Article 13(5) of Directive 98/83/EC. COM(2016) 666 final. Brussels.

EC, 2017. Commission Decision (EU) 2017/848 of 17 May 2017 laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardised methods for monitoring and assessment, and repealing Decision 2010/477/EU.

EEA, 1999. Environmental indicators: Typology and overview. No 25. European Environment Agency. Technical report, Copenhagen.

EEA, 2017. Forest: growing stock, increment and fellings. European Environment Agency, Copenhagen.

EEA, 2018a. European waters. Assessment of status and pressures 2018. European Environment Agency. EEA Report, No 7/2018. Copenhagen.

EEA, 2018b. Use of freshwater resources. European Environment Agency, Copenhagen.

EEA, 2019a. Production and consumption of ozone-depleting substances in Europe. European Environment Agency, Copenhagen.

EEA, 2019b. Status of the assessed European commercial fish and shellfish stocks in relation to Good Environmental Status (GES) per EU marine region in 2015–2017. European Environment Agency, Copenhagen.

EEA, 2019c. Country reports 2018 bathing season. European Environment Agency, Copenhagen.

EEA, 2020. National summary dashboards - Habitats Directive – Art.17. Copenhagen: European Environment Agency.

Ekins, P., 2011. Environmental sustainability: From environmental valuation to the sustainability gap. Prog. Phys. Geogr. 35 (5), 629–651.

Ekins, P., Simon, S., 1999. The sustainability gap: a practical indicator of sustainability in the framework of the national accounts. Int. J. Sustain. Dev. 2 (1), 32–58.

Ekins, P., Simon, S., 2001. Estimating sustainability gaps: methods and preliminary applications for the UK and the Netherlands. Ecol. Econ. 37 (1), 5–22.

Ekins, P., Simon, S., Deutsch, L., Folke, C., De Groot, R., 2003. A framework for the practical application of the concepts of critical natural capital and strong sustainability. Ecol. Econ. 44 (2–3), 165–185.

European Council, 1998. Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption.

European Parliament and European Council, 2008. Directive 2008/105/EC on environmental quality standards in the field of water policy, amending and subsequently repealing Council Directives 82/176/EEC, 83/513/EEC, 84/156/EEC, 84/491/EEC, 86/280/EEC and amending Directive 2000/60/EC of the European Parliament and of the Council.

Eurostat, 2019. Greenhouse gas emissions by source sector. Eurostat, Luxembourg.

Eurostat, 2020. EU SDG Indicator set 2020. Result of the review in preparation of the 2020edition of the EU SDG monitoring report. Eurostat, Luxembourg.

Forest Europe, 2020. State of Europe's Forests 2020. Bratislava: Ministerial Conference on the Protection of Forests in Europe.

Greco, S., Ishizaka, A., Tasiou, M., Torrisi, G., 2019. On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness. Soc. Indic. Res. 141 (1), 61–94.

Heck, V., Gerten, D., Lucht, W., Popp, A., 2018. Biomass-based negative emissions difficult to reconcile with planetary boundaries. Nat. Clim. Change 8 (2), 151–155.

Horálek, J., Schreiberová, M., Schneider, P., Kurfürst, P., Schovánková, J., Ďoubalová, J., 2020. European air quality maps for 2017: PM10, PM2.5, Ozone, NO2 and NOx spatial estimates and their uncertainties. ETC/ATNI 2019/9. Kjeller: t. European Topic Centre on Air pollution, noise and industrial pollution.

Huber, S., Prokop, G., Arrouays, D., Banko, G., Bispo, A., Jones, R.J.A., Kibblewhite, M. G., Lexer, W., Möller, A., Rickson, R.J., Shishkov, T., Stephens, M., Toth, G., Van den Akker, J.J.H., Varallyay, G., Verheijen, F.G.A., Jones, A.R., 2008. Environmental Assessment of Soil for Monitoring. Volume I: Indicators & Criteria. Luxembourg: Joint Research Centre - Institute for Environment and Sustainability.

IPCC, 2018. Global Warming of 1.5°C.An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Edited by V. Masson-Delmotte, et al.: Intergovernmental Panel on Climate Change,.

Jones, R.J.A., Y.L. Bissonnais, P. Bazzoffi, J. Sánchez, O.D. Díaz, G. Loj, L. Øygarden, V. Prasuhn, B. Rydell, P. Strauss, J.B. Üveges, L. Vandekerckhove, Y. Yordanov, 2004. Soil Erosion. In: Reports of the Technical Working Groups Established under the Thematic Strategy for Soil Protection. Volume II: Erosion, edited by L. Van-Camp, et al. Luxembourg: Joint Research Centre - Institute for Environment and Sustainability.

JRC, 2019. Your 10-Step Pocket Guideto Composite Indicators & Scoreboards. Joint Research Centre, Ispra.

Karlsson, P.E., Uddling, J., Braun, S., Broadmeadow, M., Elvira, S., Gimeno, B.S., Le Thiec, D., Oksanen, E., Vandermeiren, K., Wilkinson, M., Emberson, L., 2003. New Critical Levels for Ozone Impact on Trees Based on AOT40 and Leaf Cumulated Ozone Uptake. In: Karlsson, P.-E. (Ed.), Establishing Ozone Critical Levels II, UNECE Workshop Report. Gothenburg.

Karlsson, P.E., Braun, S., Broadmeadow, M., Elvira, S., Emberson, L., Gimeno, B.S., Le Thiec, D., Novak, K., Oksanen, E., Schaub, M., Uddling, J., Wilkinson, M., 2007. Risk assessments for forest trees: The performance of the ozone flux versus the AOT concepts. Environ. Pollut. 146 (3), 608–616.

Kwatra, S., Kumar, A., Sharma, P., 2020. A critical review of studies related to construction and computation of Sustainable Development Indices. Ecol. Ind. 112, 106061.

Landrigan, P.J., Fuller, R., Acosta, N.J.R., Adeyi, O., Arnold, R., Basu, N., Baldé, A.B., Bertollini, R., Bose-O'Reilly, S., Boufford, J.I., Breysse, P.N., Chiles, T., Mahidol, C., Coll-Seck, A.M., Cropper, M.L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., Hanrahan, D., Hunter, D., Khare, M., Krupnick, A., Lanphear, B., Lohani, B., Martin, K., Mathiasen, K.V., McTeer, M.A., Murray, C.J.L., Ndahimananjara, J.D., Perera, F., Potočnik, J., Preker, A.S., Ramesh, J., Rockström, J., Salinas, C., Samson, L.D., Sandilya, K., Sly, P.D., Smith, K.R., Steiner, A., Stewart, R.B., Suk, W. A., van Schayck, O.C.P., Yadama, G.N., Yumkella, K., Zhong, M., 2018. The Lancet Commission on pollution and health. Lancet 391 (10119), 462–512.

Loveland, P., Webb, J., 2003. Is there a critical level of organic matter in the agricultural soils of temperate regions: a review. Soil Tillage Res. 70 (1), 1–18.

Mace, G.M., Reyers, B., Alkemade, R., Biggs, R., Chapin, F.S., Cornell, S.E., Díaz, S., Jennings, S., Leadley, P., Mumby, P.J., Purvis, A., Scholes, R.J., Seddon, A.W.R., Solan, M., Steffen, W., Woodward, G., 2014. Approaches to defining a planetary boundary for biodiversity. Global Environ. Change 28, 289–297.

Mills, G., Buse, A., Gimeno, B., Bermejo, V., Holland, M., Emberson, L., Pleijel, H., 2007. A synthesis of AOT40-based response functions and critical levels of ozone for agricultural and horticultural crops. Atmos. Environ. 41 (12), 2630–2643.

Neumayer, E., 2003. Weak versus Strong Sustainability: Exploring the Limits of Two Opposing Paradigms. Edward Elgar, Cheltenham.

OECD and JRC, 2008. Handbook on Constructing Composite Indicators: Mehodology and User Guide. Organisation for Economic Co-operation and Development, Paris.

Osipova, E., Shi, Y., Kormos, C., Shadie, P., C, Z., Badman, T., 2014. IUCN World Heritage Outlook 2014: A conservation assessment of all natural World Heritage sites. International Union for the Conservation of Nature, Gland.

Osipova, E., Emslie-Smith, M., Osti, M., Murai, M., Åberg, U., Shadie, P., 2020. IUCN World Heritage Outlook 3: A conservation assessment of all natural World Heritage sites. International Union for the Conservation of Nature, Gland.

Ozone Secretariat United Nations Environment Programme, 2019. Data centre. Secretariat for The Vienna Convention for the Protection of the Ozone Layer & The Montreal Protocol on Substances that Deplete the Ozone Layer.

Panagos, P., Ballabio, C., Poesen, J., Lugato, E., Scarpa, S., Montanarella, L., Borrelli, P., 2020. A Soil Erosion Indicator for Supporting Agricultural. Environ. Clim. Policies European Union 12 (9), 1365.

Poelman, H., 2018. A walk to the park? Assessing access to green areas in Europe's cities. WP 01/2018. European Commission.

Raskin, P.D., Gleick, P., Kirshen, P., Pontius, G., Strzepek, K., 1997. Comprehensive assessment of the water resources of the world. Stockholm Environment Institute, Stockholm.

Röschel, L., Noebel, R., Stein, U., Naumann, S., Romão, C., Tryfon, E., Gaudillat, Z., Roscher, S., Moser, D., Ellmauer, T., Löhnertz, M., Halada, L., Staneva, A., Rutherford, C., 2020. State of Nature in the EU - Methodological paper Methodologies under the Nature Directives reporting 2013–2018 and analysis for the State of Nature 2000. European Topic Centre on Biological Diversity, Paris.

Srebotnjak, T., H. Gerdes, A. Best, and S. Cavalieri, 2009. Evaluation of Indicators for EU Policy Objectives. D2.1. IN-STREAM project.

Tsyro, S., Aas, W., Solberg, S., Benedictow, A., Fagerli, H., Scheuschner, T., 2020. Status of transboundary air pollution in 2018. In: Fagerli, H. (Ed.), Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components. Norwegian Meteorological Institute, Oslo.

Turner, R.K., 1993. Sustainability: Principles and Practice. In: Turner, R.K. (Ed.), Sustainable Environmental Economics and Management: Principles and Practice. Belhaven Press, New York/London.

UN, 1987. Montreal Protocol on Substances that Deplete the Ozone Layer. Vol. 1522, 1-26369. Montreal: United Nations.

UN Environment, 2019. Global Environment Outlook – GEO-6: Summary for Policymakers. Nairobi.

UNSD, 2015. Discussion paper on Principles of Using Quantification to Operationalize the SDGs and Criteria for Indicator Selection. United Nations Statistics Division, New York.

Usubiaga-Liaño, A., P. Ekins, submitted. Where do we stand in measuring the environmental sustainability of countries? An assessment of existing metrics and the ESGAP framework. Front. Environ. Sci.

Verheijen, F.G.A., Jones, R.J.A., Rickson, R.J., Smith, C.J., 2009. Tolerable versus actual soil erosion rates in Europe. Earth Sci. Rev. 94 (1–4), 23–38.

WHO, 2005. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment. Geneva: World Health Organisation.

WHO, 2020. Percentage of the population using clean and polluting fuels and technologies for cooking. World Health Organisation, Geneva.