Building a composite indicator for biodiversity through supervised learning and linked indicator sets

Kilian Rouge and C. Adam Schlosser
Building a composite indicator for biodiversity through supervised learning and linked indicator sets

Kilian Rouge\textsuperscript{1,2}, C. Adam Schlosser\textsuperscript{1}

Abstract: Understanding and predicting fate of global biodiversity amidst an increasingly complex and changing world is a major challenge facing the Earth-system science community. Among the core research objectives within this challenge lies the ability to construct a comprehensive metric that not only faithfully quantifies the current and observed state of biodiversity, but also captures future trends that are driven by a variety of stressors across environmental, social, and economic systems. In order to give a better overview of our impact on biodiversity despite the obvious complexity inherent to the multi-sectoral nature of the problem, we have chosen to group together the indicators currently assessed and used internationally in a linked indicator set categorized according to the “Pressure-State-Response” framework. This approach stems from a desire to highlight and quantify the links between these different indicators in a logical and objective manner and allows us to construct a systematic synthesis of the key drivers of biodiversity. We develop a new methodology using predictive supervised learning to propose a statistical weighting of the linked indicator metric.

1. INTRODUCTION
2. CREATION OF A COMPOSITE INDICATOR
   2.1 GENERAL CONSIDERATIONS
   2.2 WEIGHING INDICATORS TO COMPUTE COMPOSITE INDICATOR
   2.3 USING A LINKED INDICATOR SET TO ASSESS BIODIVERSITY
   2.4 USING FEATURE IMPORTANCE TO DETERMINE THE RELATIVE WEIGHT OF INDICATORS IN THE LINKED SET
3. RESULTS
   3.1 CHOICE OF INDICATORS COMPOSING THE LINKED SET
   3.2 EXPERIMENTAL PROTOCOL
   3.3 COMPUTATION OF THE RELATIVE IMPORTANCE
   3.4 NATIONAL COMPUTATION OF THE RELATIVE IMPORTANCE OF THE INDICATORS
4. COMPUTATION OF THE WEIGHTED INDEX
5. CLOSING REMARKS
6. REFERENCES
APPENDIX

\textsuperscript{1} MIT Joint Program on the Science and Policy of Global Change, Cambridge (United States)
\textsuperscript{2} École Polytechnique, Palaiseau (France)
1. Introduction

Among the planetary boundaries proposed by Rockström et al. (2009), that of change in the integrity of the biosphere is one of the six considered exceeded in 2022. Two aspects are assessed to reach this conclusion: genetic diversity reflecting the resilience of the biosphere and the diversity of functional traits reflecting ecosystem health. This choice of measure to qualify this limit has always been highly criticized both because of the limitations associated with the calculation of relevant indicators due to missing data and because of the uncertain links between the measurements made and the true state of biodiversity. The Convention on Biological Diversity (CBD), signed in the presence of 196 parties at the Earth Summit in Rio de Janeiro in 1992, initiated important work on indicators for global monitoring of biodiversity as well as national monitoring of the National Biodiversity Strategies and Action Plans (NBSAPs) that the Convention requires each party to implement. In order to monitor these indicators, the Biodiversity Indicator Partnership (BIP) was created in 2007 to select, evaluate and monitor indicators of the state of biodiversity, the threats it faces, and the measures implemented. The sixty indicators assessed, although sometimes with a lack of data, made it possible to ensure with certainty that the targets set for 2010 had not been reached. The 10th Conference of the Parties held in Nagoya, Japan in 2010 set new targets for 2020: the Aichi Targets (United Nations, 2010). These set a broader and stricter conservation target. In addition, the targets are subdivided into 5 strategic goals. BIP had to rework its indicators, as only 13 of the 20 targets were covered by at least one indicator. Today, BIP differentiates between a set of primary and secondary indicators for each target, and monitoring is facilitated by the fact that, at least at the global and national levels, many indicators are available and calculated annually. A new monitoring framework has been proposed more recently with the development of the Kunming-Montreal Global Biodiversity Framework (United Nations, 2022). The number of indicators is therefore greater than ever, making it particularly difficult to raise awareness among the public and decision makers.

Biodiversity is intrinsically multi-sectoral. The deterioration of biodiversity is caused by multitude anthropogenic pressures: the impact of overfishing or deforestation on natural resources, the increase in temperature or eutrophication of the oceans due to climate change, the impact of pollution and fine particles, and the artificialization and pollution of soils due to industrial, agricultural or urban activities are only some of them. The public policies implemented by the different countries or the awareness of the populations, just as important as the assessment of the state of biodiversity, are still not sufficiently accessible to decision makers.

The growing complexities among the links between climate, natural and societal systems require scientists and policy makers to explore these interdependencies in the study of risks and instabilities. These complex interactions lead to the appearance of tipping points, the analysis of which requires the use of increasingly complex and often specific models. For the general public as well as for decision-makers, it is becoming more difficult to provide a meaningful diagnosis that provides a summary of the underlying phenomena that drives risks on a macro scale. It has also become increasingly difficult to track global or relative risks that result from stressors related to social and environmental phenomena. Therefore, the scientific community is challenged with a task to construct a more comprehensive, time-evolving metric of biodiversity that logically and objectively combines data relating both to climatic and natural risks and societal factors about vulnerability of populations and resilience of communities such as: access to health care, standards of living, or education.

This paper aims to provide a new methodology to assess weights used to compute a composite indicator using predictive neural networks and feature importance algorithms. Our methodology is, furthermore, based on the use of linked indicator sets to logically categorize sub-indicators. We apply this methodology to a multisectoral set of biodiversity indicators categorized according to the “Pressure-State-Response” framework and compute an aggregated index as an example of use case.

2. Creation of a composite indicator

2.1 General considerations

To give a systematic overview of the absolute risk level in the case of biodiversity indicators, composite indicators have become more common practice to international organizations and governments. They allow the monitoring of countries over time while considering a single indicator instead of several separate indicators while keeping the underlying information.

The choice of the normalization and aggregation methods is crucial when considering indicators with very different measurement units such as US Dollars, Percentages, and Biodiversity-specific units for trophic levels. Jacobs et al. (2004) and Freudenberg (2003) lists several normalization methods. The proposed linked indicator set must, however, be standardized or be run through a MinMax scaler as the minimum and maximum don’t always represent the worst and best value. The other option is to use categorical scales consisting in assigning a score to each indicator. This can be based on a ranking of the different countries and categorical values representing the position of a given country within the histogram of values or the position with regards to a given threshold. However, the very existence...
of thresholds and tipping points for biodiversity is in fact questioned (e.g., Montoya et al., 2018).

### 2.2 Weighing indicators to compute composite indicator

To calculate a composite indicator from an indicator set one can compute an arithmetic mean aggregation:

\[
l = \sum_{k=1}^{n} w_k i_k
\]

where \(w_k\) is the relative weight of indicator \(i_k\) in the sum.

In the score evaluation presented above, the score of each indicator can be taken into account in the same way in the final score evaluation by considering. The addition of superfluous indicators is sufficient to convince oneself that the score of each individual indicator should, however, be weighted in the construction of the total score. The question of the relative importance of the indicators in relation to each other and therefore of the choice of weights is a delicate one, as it is important to consider the entirety of the physical and socio-economic phenomena and therefore not to eliminate indicators that may be relevant to one specific aspect of biodiversity.

Two main paths exist when considering weighting of indicators (see Sharpe et al., 2012). The first one is explicit weighting which consists in manually analyzing the meta-data and general analysis resulting from the first step of the methodology described above and making subjective decisions about the weights to give to each indicator in the composite indicator mix. Several methods have been used by governmental and non-governmental organizations such as expert weighting where the weights are decided by professionals (as described for example by Gómez-Limón et al., 2020) of the relevant field or survey weighting where one considering that accurate weight should reflect valuation of the society. Several methodologies for theses participatory methods have been described in the literature, the best-known being Budget Allocation Process (BAP) based on allocation of a given number of points to the indicators included in a set and the Analytic Hierarchy Process based on pairwise comparisons. Finally, the analyst can decide to give the final user of the composite indicator the choice of weights he wants to use (e.g., Schlosser et al., 2022).

The other path is to use “data-driven” weighting methods (Decancq and Lugo, 2013) where a numerical analysis of the data enables to compute the weights for the composite indicator. These methods can therefore be considered as objective methods as the analyst is not involved in the decision-making regarding the choice of weights. However, this approach results in a tradeoff as both the analyst and the final user lose the transparency associated to the explicit understanding of the process of choosing the weights. Furthermore, criticism arises about these methods both on the fact that statistical relationships don’t always accurately represent the relationship between indicators and on method-specific issues (see Greco et al., 2019). Most of the data-driven methods are based on algorithms trying to maximize the variance in the indicator set with as few components as possible. Principal Component Analysis (PCA) for example consists of finding the eigenvalues of the covariance matrix of the indicator set corresponding to the variance of the principal components. The goal is to determine optimal weights in order to maximize the variation explained by only the first principal component used as the composite indicator.

Recently, other statistical approaches based on machine learning have been developed. In 2020, Paulvannan Kanmani et al. (2020) proposed a method to better understand indicator sets by using clustering and unsupervised learning. The method developed in the study aims to overcome the limits introduced by rankings of countries as in traditional sustainability indicators. By tracking the evolution of each country within a self-organized map, grouping countries by similarities in the evolution of their sustainability, one can compare the trajectory of a country with some having or that have had similar characteristics. This method is more of an alternative than an ultimate solution to weighing and aggregation of composite indicators. More recently Jiménez-Fernández et al. (2022) and Jiménez-Fernández et al. (2022) used Multivariate Adaptive Regression Splines instead of traditional distance minimization to best approximate the indicator set with the composite indicator on one side and P2 minimization specifically applied to the benchmark of units of the different variables.

### 2.3 Using a linked indicator set to assess biodiversity

To provide an overview of the state of biodiversity and the complex interactions that govern it, many indicators are needed. The only description of the state of biodiversity (richness or abundance for example) can be done through measures of specific ecosystem, genetic or functional diversity, the latter considering the interactions between society and nature (i.e., resources, landscapes, and cultural wealth). This multitude of indicators, although scientifically necessary for the comprehension of the state of biodiversity, can also be detrimental to the understanding of the underlying problems by an uninformed public. Our desire to synthesize the information stems from this problem. The main issues surrounding the development of a relevant set of indicators are proposed in Levrel et al. (2007). In order to achieve a systematic vision of biodiversity while maintaining the underlying complex dynamics, we seek to achieve an educated, categorized, hierarchized and contextualized selection of indicators.
Given these considerations, an important milestone in this construction was the scientific consensus that emerged in 2010 on a Pressure-State-Response framework to organize biodiversity indicators (e.g., Sparks et al., 2011). This conception recognized that biodiversity is affected by human activities, the state of the environment, as well as the resource status and the actions of economic and environmental agents. The strategic goals adopted at the COP 2010 are in the spirit of this framework. Strategic Goal B (Reduce the direct pressures on biodiversity and promote sustainable use) encompasses the pressures, Strategic Goal C (To improve the status of biodiversity by safeguarding ecosystems, species and genetic diversity) the state and Strategic Goal A (Address the underlying causes of biodiversity loss by mainstreaming biodiversity across government and society), D (Enhance the benefits to all from biodiversity and ecosystem services) and E (Enhance implementation through participatory planning, knowledge management and capacity building) the responses. The idea of a linked framework of indicators based on the relationship between the drivers of change, the state of the diversity of flora and fauna and the corrective actions taken is not new and is often applied in the monitoring of forests and marine environments. Linking biodiversity metrics in an organized framework can support indicator development, enable stronger predictions of biodiversity change, and provide policy-relevant advice (e.g., Sparks et al., 2011). Many linked sets have been developed to assess the status of biodiversity at a national scale (e.g., Sparks et al., 2011; Han et al., 2014; Marques et al. (2014), Hill et al., 2016) following this model. It is important to emphasize, however, that causes, state and responses are not perfectly correlated.

### 2.4 Using feature importance to determine the relative weight of indicators in the linked set

Within statistical weighing methods, a main criticism is the difficulty of interpretations of the weights given to each of the different sub-indicators. We believe that the use of predictive neural networks to compute weights can make these interpretations easier by highlighting quantitative relationships between indicators. Feature importance allows to understand the relative importance of the input parameters and to make accurate interpretations on the most important input variables for the prediction of an output. In the case of indicator sets, training a model to predict the value of one indicator based on the value of others teaches it temporal and logical relationship between indicators. Feature importance therefore allows to highlight indicators whose values can be deduced from one another. Weights given to each of the indicators in a linked set can be deduced from those relationships by making the assumption that in the Pressure-State-Response framework, an indicator in one category may be considered crucial if its value has an impact on many indicators in another category. In other words, this means that a variation at a given moment has an impact on the value of the indicators of the other categories shortly thereafter, indicating a response to a common phenomenon related to a risk that increases tenfold with the number of domains affected. For this hypothesis to be valid, the linked indicator set has to be a partition of the space of indicators of the studied question. In other words, the indicators in the different categories have to cover all the underlying phenomena and the categories have to be of null intersection so an indicator can be assigned to only one category.

The Pressure-State-Response framework is particularly useful in that matter because the relationships between categories are logical. It is expected of a very information-rich indicator in the “Pressure” category to impact a lot of indicators in the “State” and “Response” categories. This approach, furthermore, limits the problem of double counting encountered with other statistical methods. By disregarding relationships between indicators of the same category, we allow ourselves to consider indicators that might partly describe the same phenomenon and therefore be easily inter-predictable.

In order to determine the weights given to each indicator in the linked set, we successively train neural networks to predict the value of an indicator from those of the other categories. After successful training, we evaluate the relative importance of each input indicator, that is determined by its predictive strength. The relative importance (RI) of each indicator within the set of k indicators is then given by:

$$ R_{Ij} = \frac{1}{\Sigma_{I \in \text{cat} j} \Sigma_I c_{ij} \frac{\Sigma_k c_{kj}}{\Sigma_k c_{ki}}} $$

(2)

In Equation 2 the weights given to each indicator is the average importance of that indicator in the prediction of indicators across other categories weighted by the importance those indicators have. Thus, adding the last weighting term allows to give a low weight to the relative importance of the considered indicator in the prediction of an indicator that would be irrelevant. The relative importance of each indicator in the prediction of noise would for example not be considered in our indicator weighting method.

### 3. Results

#### 3.1 Choice of indicators composing the linked set

We have chosen our different clusters of indicators to cover the entirety of the Aichi Targets while making some adjustments and grouping some of them together to consider the Pressure-State-Response framework. The indicators finally selected in each cluster were chosen according to the following criteria:
1. Data availability and quality
2. Ability to aggregate to a national scale
3. Availability over a time scale greater than 5 years

The resulting linked indicator set is presented in Table 1. In many cases, an indicator used for the national monitoring of the Aichi Targets could be obtained directly or by calculation from existing BIP data. Otherwise, we turned to other sources. Some of the indicators used in the pollution, land use, and climate change clusters are derived from the indicators or data used in the calculation of the composite indicator Environmental Performance Index (see Wolf et al., 2022). For this indicator, it is worth mentioning the normalization between a minimum and a maximum set at given quantiles to carry out the aggregation and an empirical weighting based on “the importance of the issue, data quality, timeliness of data, and statistical analyses to balance the spread of score” (see Wolf et al., 2022). Finally, in some cases we have identified relevant indicators that are unfortunately not yet available or substitutable today. This is the case for the Ecosystem Area Index, Ecosystem Health Index and Red List Index of Ecosystems (see Rowland et al., 2020) as well as the Biodiversity Engagement Indicator (see Cooper et al., 2019).

Joint Research Centre-European Commission and others (2008) describes a methodology of ten steps for the development of composite indicators. Even though this handbook can be viewed as a sequence towards the creation of a composite indicator, one must make several decisions in the process to have a heavy impact on the outcome. It is therefore important not to forget that composite indicators might oversimplify the issues at stake and mislead policy construction.

The first few steps of developing a theoretical framework, selecting variables, imputation of missing data and multivariate analysis are briefly described in the previous sections of this report. The use of a linked indicator set allows to create a nested structure to describe a wide-ranging phenomenon such as biodiversity under the scope of its state, drivers and consequences. Furthermore, the summary tables and in particular the consideration of the limitation of the individual indicators and the multivariate analysis allow for a better understanding of the weaknesses of the resulting composite indicator and interpretation of the weights that will be computed in the next steps of the methodology.

### 3.2 Experimental protocol

The raw data first undergoes different steps of pre-processing. To add more data points to the set, a spline is fit to the different indicators per country and 100 points are added in between two years. A quick study of Figure 1 leads us to eliminate the period to ensure that enough indicators of each category are non-null. In order to facilitate the learning of temporal relationships, each entry is composed of the last 10 values for each indicator and not only of the value corresponding to the date of the entry. The ISO country codes are one hot encoded and the data normalized. Due to the aggregation of many indicators spanning different

---

**Table 1.** Summary table of the identified indicators used in a “Pressure-State-Response” framework. For a more comprehensive description of the indicators, please refer to the appendix Tables A1-A3.

<table>
<thead>
<tr>
<th>Pressure</th>
<th>GHG</th>
<th>SNM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pollution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land Use</td>
<td>TCL</td>
<td>WLL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GLL</td>
</tr>
<tr>
<td>Climate Change</td>
<td>CO₂</td>
<td>DIS</td>
</tr>
<tr>
<td>Resource Extraction</td>
<td>MTI</td>
<td></td>
</tr>
<tr>
<td>Invasive Species</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>State</th>
<th>RLI</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Species Richness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species Abundance</td>
<td>BII</td>
<td></td>
</tr>
<tr>
<td>Ecosystem Health and Services</td>
<td>NaN</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response</th>
<th>PAKA</th>
<th>SPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protected Ecosystems</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pollution Policy</td>
<td>REC</td>
<td></td>
</tr>
<tr>
<td>Finance</td>
<td>CRS</td>
<td></td>
</tr>
<tr>
<td>Public Awareness</td>
<td>NaN</td>
<td></td>
</tr>
</tbody>
</table>
time ranges, the data has many missing values (See Figure 1). These are replaced by the value “-8888”.

The model used is a 3-layer Long Short-Term Memory (LSTM) network. This special type of recurrent network is capable of learning long-term dependencies and is therefore particularly suited in our case. We train it separately for the determination of each indicator. The inputs used for an indicator are then all indicators not belonging to the same category as . Thus the obvious correlation between the Red List Index and the Biodiversity Index, both in the State category and both concerning specific hazard issues, is not taken into account in the calculation of weights. We also eliminate for each indicator the input data with fewer than two non-zero values. Once the model is properly trained and tested on a random set, we compute the importance of the input features using the Deep Learning algorithm Important Features proposed by Shrikumar et al. (2017). Only the values corresponding to indicators are considered in the final results, thus excluding the years and country. The relative importance of the different lags is aggregated for each indicator and the results are presented in Table 3. The relative importance given to each indicator is then calculated by considering the results of feature importance of the set of models each predicting an indicator from the others of a different category according to Equation 2.

In the following discussion, we use a socio-geographic regional decomposition to analyse the results more finely. The macro-regions chosen are the following: Southeast Asia, Pacific Asia, Global West, Former Soviet countries, Sub-Saharan Africa, Middle East, Latin America, Eastern Europe also presented on Figure 2.

![Figure 1. Years of existence of at least one value per indicator considered](image)

![Figure 2. Socio-geographical macroregion chosen for the analysis of the results](image)
### 3.3 Computation of the relative importance

To check the results given by the Feature importance and to identify possible aberrations, we perform the Feature Importance calculation on a specific input. The results are presented in Table 2 and highlight the following:

- The country considered is of primary importance as expected. On the contrary, all the other countries are of zero importance. One would have expected to see more complex links between the values of a country and those of its neighbouring countries.
- The indicators considered with no value are not considered by the model. On the contrary, all other indicators have a strictly positive importance.
- Contrary to what one might have expected, the year considered is not that important. The main part of the prediction is therefore based on the value of indicators of different categories at the same point in time and on learning patterns in their evolution. It should be noted that the importance of the year fluctuates a lot and becomes particularly important in the absence of value for the most relevant indicators.

The results are presented in Table 3 and in a more visual form in Figure 3. Finally, we calculate the relative importance of each indicator according to Equation 2 which is summarized in the graph Figure 4.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Importance</th>
<th>Input</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISOUSA</td>
<td>47.03</td>
<td>1</td>
</tr>
<tr>
<td>EF</td>
<td>11.58</td>
<td>6.22</td>
</tr>
<tr>
<td>MTI</td>
<td>10.32</td>
<td>3.79</td>
</tr>
<tr>
<td>SPI</td>
<td>7.46</td>
<td>35.06</td>
</tr>
<tr>
<td>WLL</td>
<td>7.04</td>
<td>0.05</td>
</tr>
<tr>
<td>GLL</td>
<td>6.01</td>
<td>0.00</td>
</tr>
<tr>
<td>GHG</td>
<td>3.53</td>
<td>0</td>
</tr>
<tr>
<td>PAKA</td>
<td>3.00</td>
<td>34.72</td>
</tr>
<tr>
<td>REC</td>
<td>2.54</td>
<td>0.86</td>
</tr>
<tr>
<td>Year</td>
<td>1.02</td>
<td>2001.54</td>
</tr>
<tr>
<td>CO2</td>
<td>0.33</td>
<td>0.01</td>
</tr>
<tr>
<td>SNM</td>
<td>0.13</td>
<td>0.39</td>
</tr>
<tr>
<td>TCL</td>
<td>0</td>
<td>-8888</td>
</tr>
<tr>
<td>DIS</td>
<td>0</td>
<td>-8888</td>
</tr>
<tr>
<td>CRS</td>
<td>0</td>
<td>-8888</td>
</tr>
<tr>
<td>ISOXXX</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. Feature Importance results for an entry of the BII-prediction model

Table 3. Feature Importance computed according to the DeepLift Algorithm of each indicator (rows). Each column represents the entry features considered. The Relative importance of each indicator is computed according to Equation 2 and presented in the last column. The results in other columns are expressed as a percentage of use for the prediction of each indicator.
Figure 3. Graphical representation of matrix Table 2. Arrows represent the use of one indicator for the prediction of another. The thickness of the arrows is proportional to the coefficients in matrix Table 2. Arrows corresponding to coefficients of less than 2% are not represented.

Figure 4. Relative Importance computed with Feature Importance according to the DeepLift Algorithm of each indicator.
Note first that Table 3 is a block matrix of diagonal null blocks. This is due to the fact that by design, the predictions and calculations of relative importance of an indicator of a category only take into account indicators that are not part of this category. For example, for the prediction of the Red List Index (RLI) of the “State” category, only the indicators of the “Pressure” and “Response” categories are used. In the case of this example, the prediction made by the neural network is based 50% on the species protection index (SPI) and 26% on the biodiversity protection index (PAKA). This is not surprising, as the three indicators explain the risks of extinction incurred by the species and the means put in place to fight against it. Note, however, that the opposite is not necessarily true and that the land use indicators (WLL, TCL and GLL) are preferred to the red list index (RLI) for predicting SPI and PAKA. The Species Protection Index (SPI) is used for the prediction of many pressure and response indicators. It is therefore strongly weighted in our study. The population impact of natural disasters (DIS) is a weak predictor of indicators in other categories. It characterizes the impact of extreme events on populations, which is, as one might expect, a poor indicator of the state of biodiversity. Its prediction from the indicators of the “State” and “Response” categories is, however, achieved and it is important to note that Equation 2 considers its uselessness in the predictions of other indicators to decrease the weighting of the coefficients brought by DIS in the calculation of the relative importance of the other indicators.

Finally, we note that the use of some indicators in the prediction of others is unexpected. This is the case, for example, of the use of wetland cover changes (WLL) in the prediction of the recycling rates of the different countries (REC). The heavy reliance on WLL in the prediction of REC can only be explained by the fact that the algorithm could not find a better indicator than WLL to perform the prediction of REC. The addition of many other indicators would improve the predictions and reveal relationships more in line with our expectations.

### 3.4 National computation of the relative importance of the indicators

We also perform the relative importance calculations at a national level. To do this, the methodology used to obtain Table 3 is repeated for each country. Starting from the models trained for the prediction of each of the indicators, the Feature Importance calculations are made from the data for each of the countries and a relative importance of each indicator is calculated for each country. The results are presented in Figure 5. One can see that the variance of the different relative importance depends on the indicators. Some of them, such as the impact of natural disasters (DIS) on populations, have a relative importance that varies enormously from one country to another, while others, such as the ecological footprint (EF) or the rational use of nitrogen in agriculture (SNM), have very limited variance. It should also be noted that the geographical distribution chosen does not seem to give any particular meaning to the distribution of the relative importance of the different indicators by country. Indeed, there does not seem to be an obvious correlation between the chosen geographical categorisation and, by extension, the...
development or not of a country and the relative importance of the different indicators. Finally, it should be noted that the calculation of the relative importance of each indicator at the national level is not related to the global level. It is therefore not surprising that an indicator such as the impact of natural disasters (DIS) has a much greater relative importance in the case of the individual calculation for each country than in the global case.

4. Computation of the weighted index

To give an example of the use of the weighted indicator set, we propose an aggregated index based on the indicators presented above. As this paper aims to focus on the development of a weighting method, we will use the following simple methodology for the aggregation: (i) Normalization of the different indicators with a StandardScaler (ii) Capping of values between 0 and 1 by setting 0 and 1 respectively at the 5th and 95th quintile (iii) Inversion of the indicators if necessary, so that 1 is systematically the most favorable value. For indicators such as the Marine Trophic Index for which there is no favorable value as its value represents the mean trophic level of fish catches, we apply the same methodology to the 5-year rolling average change rate.

We can now use the weights obtained in the calculation of a biodiversity index as presented in the first part of this chapter. The map Figure 7 presents the map of our biodiversity index calculated using the weights from the use of our predictive model compared to an unweighted index. Our weighting highlights as expected countries for which a low biodiversity index was expected, such as China, India and Indonesia, where individual metrics are initially low. Two countries that initially had a similar unweighted index, such as Argentina and Brazil for 2018, may have a very different weighted index (see Table 4). Indeed, the higher
values of the Species Protection Index (SPI) and the Red List Index (RLI), which are heavily weighted, tend to make the difference between the two countries at the expense of other differences in the values of some indicators for these two countries that would have favoured Argentina more, but which are less weighted.

In putting this aggregate metric into a global context, recently efforts have led to the development and continual updating to an "environmental performance index" (EPI, Wolf et al., 2022). While designed to look more broadly at sustainability issues, the EPI metric carries factors that reflect a nation’s ability to protect environmental health and enhance ecosystem vitality. Analysis of the EPI (Wolf et al., 2022) has shown a positively correlated relationship between a nation’s EPI score to GDP per capita, with higher environmental performance generally associated with wealthier countries. In view of this, we construct a similar analysis with our aggregated metric to assess the extent to which such a relationship exists (Figure 8), whereby countries’ aggregate metric values are mapped as a function of the log GDP per capita. We find evidence of a consistent re-

### Table 4. Comparison of the value of each indicator and the aggregated indexes for Argentina and Brazil in 2018. Missing values are non-existent for those countries in 2018

<table>
<thead>
<tr>
<th>ISO</th>
<th>Year</th>
<th>GHG</th>
<th>SNM</th>
<th>TCL</th>
<th>WLL</th>
<th>GLL</th>
<th>EF</th>
<th>MTI</th>
<th>CO2</th>
<th>DIS</th>
<th>RLI</th>
<th>PAKA</th>
<th>SPI</th>
<th>REC</th>
<th>CRS</th>
<th>Unweighted Index</th>
<th>Weighted Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARG</td>
<td>2018</td>
<td>0.33</td>
<td>0.57</td>
<td>0.95</td>
<td>0.98</td>
<td>1</td>
<td>0.74</td>
<td>0.57</td>
<td>0.38</td>
<td>0.47</td>
<td>0</td>
<td>0.60</td>
<td>0.61</td>
<td></td>
<td></td>
<td>0.60</td>
<td>0.61</td>
</tr>
<tr>
<td>BRA</td>
<td>2018</td>
<td>0.19</td>
<td>0.44</td>
<td>0.94</td>
<td>0.93</td>
<td>0.78</td>
<td>0.57</td>
<td>0.75</td>
<td>0.50</td>
<td>0.89</td>
<td>0</td>
<td>0.60</td>
<td>0.69</td>
<td></td>
<td></td>
<td>0.60</td>
<td>0.69</td>
</tr>
</tbody>
</table>

### Table 5. Comparison of the country rankings between the aggregate index developed in this study and the Environmental Performance Index (EPI) for 2018.

<table>
<thead>
<tr>
<th>Rank</th>
<th>EPI</th>
<th>Aggregate Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Denmark</td>
<td>Luxembourg</td>
</tr>
<tr>
<td>2</td>
<td>United Kingdom</td>
<td>Estonia</td>
</tr>
<tr>
<td>3</td>
<td>Finland</td>
<td>Latvia</td>
</tr>
<tr>
<td>4</td>
<td>Malta</td>
<td>Denmark</td>
</tr>
<tr>
<td>5</td>
<td>Sweden</td>
<td>Finland</td>
</tr>
<tr>
<td>6</td>
<td>Luxembourg</td>
<td>Namibia</td>
</tr>
<tr>
<td>7</td>
<td>Slovenia</td>
<td>Czechia</td>
</tr>
<tr>
<td>8</td>
<td>Austria</td>
<td>Slovakia</td>
</tr>
<tr>
<td>9</td>
<td>Switzerland</td>
<td>Venezuela, Bolivarian Republic of</td>
</tr>
<tr>
<td>10</td>
<td>Iceland</td>
<td>Lithuania</td>
</tr>
<tr>
<td>11</td>
<td>Netherlands</td>
<td>Croatia</td>
</tr>
</tbody>
</table>

![Figure 8. Aggregated Index as a function of the decimal logarithm of GDP per capita. The solid line indicates the least-squares linear fit to the data and the shaded region indicates the error estimate.](image)
rEPOrT 365 MIT JOINT PrOGraM ON ThE ScIENcE aND POLIcY OF GLOBaL chaNGE


## Appendix

**Table A1. Summary table of pressure indicators**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Abbr.</th>
<th>What it shows</th>
<th>Methodology</th>
<th>Data Source</th>
<th>Time Frame</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>GHG Emission Trends</td>
<td>GHG</td>
<td>Emission intensity of Greenhouse gases</td>
<td>Potsdam Institute for Climate Impact Research</td>
<td>1999-2019</td>
<td></td>
<td>International transport. Land use change and deforestation emissions are not taken into account</td>
</tr>
<tr>
<td>Sustainable Nitrogen Management Index</td>
<td>SNM</td>
<td>Nitrogen use efficiency in agriculture</td>
<td>Nitrogen use efficiency is divided by crop yield and compared to a reference value</td>
<td>Zhang and Davidson (2016)</td>
<td>1961-2015</td>
<td>Based on an arbitrary base value</td>
</tr>
<tr>
<td>Tree Cover Loss</td>
<td>TCL</td>
<td>Loss of canopy surface</td>
<td>Satellite imagery</td>
<td>Global Forest Watch</td>
<td>2006-2020</td>
<td></td>
</tr>
<tr>
<td>Wetland Loss</td>
<td>WLL</td>
<td>Loss in wetland cover</td>
<td>Satellite imagery</td>
<td>Copernicus</td>
<td>1997-2020</td>
<td></td>
</tr>
<tr>
<td>Grassland Loss</td>
<td>GLL</td>
<td>Loss in grassland cover</td>
<td>Satellite Imagery</td>
<td>Copernicus</td>
<td>1997-2020</td>
<td></td>
</tr>
<tr>
<td>Ecological Footprint</td>
<td>EF</td>
<td>Human impact on natural resources</td>
<td>Biocapacity is compared to ecological assets needed to support human demand</td>
<td>Global Footprint Network</td>
<td>1961-2014</td>
<td>Impact of fossil fuels only partly taken into account Part of anthropic emissions not taken into account Regenerative capacities of the biosphere and ecosystems not taken into account</td>
</tr>
<tr>
<td>Marine Trophic Index</td>
<td>MTI</td>
<td>A measure of whether fish stocks overexploitation</td>
<td>Catch observations</td>
<td>SeaAroundUs</td>
<td>1950-2019</td>
<td>Only available for countries with marine border Quality and consistency of the measures Use of abundance reports to assess integrity of ecosystems</td>
</tr>
<tr>
<td>CO(_2) Emission Trends</td>
<td>CO(_2)</td>
<td>Emission intensity of CO(_2)</td>
<td>Potsdam Institute for Climate Impact Research</td>
<td>1999-2019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tree Cover Loss</td>
<td>TCL</td>
<td>Climate change impacts and human resilience</td>
<td>Data is collected from governmental organisations</td>
<td>United Nations</td>
<td>2005-2021</td>
<td>Impacts of climate change and higher risks due to low resilience of societies are difficult to disjoin</td>
</tr>
</tbody>
</table>
### Table A2. Summary table of state indicators

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Abbr.</th>
<th>What it shows</th>
<th>Methodology</th>
<th>Data Source</th>
<th>Time Frame</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red List Index</td>
<td>RLI</td>
<td>Species extinction risk</td>
<td>The Red List Index is computed from species threat levels and geographical trends</td>
<td>IUCN</td>
<td>2000-2022</td>
<td>Extrapolation is used because of different assessment years Regional extinction has a strong impact on national values Not very sensitive due to classification</td>
</tr>
<tr>
<td>Biodiversity Intactness Index</td>
<td>BII</td>
<td>Human impacts on terrestrial biodiversity</td>
<td>Using two models. the variation of abundance is modelled as a function of human pressures (land use and population distribution).</td>
<td>PREDICTS Project. National Zoological Society London</td>
<td>1970-2014</td>
<td>Systematic statistical error when categorising land use Each area is considered equally</td>
</tr>
</tbody>
</table>

### Table A3. Summary table of response indicators

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Abbr.</th>
<th>What it shows</th>
<th>Methodology</th>
<th>Data Source</th>
<th>Time Frame</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protected Area Coverage in Key Biodiversity Areas</td>
<td>PAKA</td>
<td>Protecting measures implemented by governments</td>
<td>Overlay of protected areas and key areas for biodiversity disaggregated per country</td>
<td>IUCN. BirdLife. United Nations</td>
<td>2000-2021</td>
<td>Effectiveness of the management of the protected area not taken into account</td>
</tr>
<tr>
<td>Species Protection Index</td>
<td>SPI</td>
<td>Protecting measures implemented by governments</td>
<td>Comparison of optimal and effectively protected habitat surfaces</td>
<td>Map of Life</td>
<td>1980-2021</td>
<td>Biodiversity abundance not taken into account Aggregation at national level creates bias Countries with less biodiversity have higher scores Arbitrary and species independent optimal species protection habitat surface</td>
</tr>
<tr>
<td>Recycling Rates</td>
<td>REC</td>
<td>Amount of recyclable waste effectively recycled</td>
<td></td>
<td>World Bank</td>
<td>1990-2020</td>
<td>Regional differences in data collection</td>
</tr>
<tr>
<td>Official development assistance for biodiversity</td>
<td>CRS</td>
<td>International aid targeting environmental protection received by each country</td>
<td>CRS Logs</td>
<td>OECD</td>
<td>2008-2020</td>
<td>A large number of countries are not considered The source and nature of funding is not taken into account</td>
</tr>
</tbody>
</table>
365. Building a composite indicator for biodiversity through supervised learning and linked indicator sets. Rouge and Schlosser, Mar 2023


363. A Large Ensemble Global Dataset for Climate Impact Assessments. Gao et al., Feb 2023


360. The MIT EPPA7: A Multisectoral Dynamic Model for Energy, Economic, and Climate Scenario Analysis. Chen et al., Jun 2022

359. A Tool for Air Pollution Scenarios (TAPS v1.0) to Facilitate Global, Long-term, and Flexible Study of Climate and Air Quality Policies. Atkinson et al., Jun 2022


357. The Changing Nature of Climate-Related Risks in Global Wind Power Resources. Schlosser et al., Feb 2022

356. Transition Scenarios for Analyzing Climate-Related Financial Risk. Chen et al., Jan 2022


353. Predictability of U.S. Regional Extreme Precipitation Occurrence Based on Large-Scale Meteorological Patterns (LSMPs). Gao & Mathur, Jun 2021


351. Meeting Potential New U.S. Climate Goals. Yuan et al., Apr 2021


347. Representing Socio-Economic Uncertainty in Human System Models. Morris et al., Feb 2021

346. Renewable energy transition in the Turkish power sector: A techno-economic analysis with a high-resolution power expansion model, TR-Power. Kat, Feb 2021

345. The economics of bioenergy with carbon capture and storage (BECCS) deployment in a 1.5°C or 2°C world. Fajardy et al., Nov 2020

344. Future energy: In search of a scenario reflecting current and future pressures and trends. Morris et al., Nov 2020


341. Emulation of Community Land Model Version 5 (CLM5) to Quantify Sensitivity of Soil Moisture to Uncertain Parameters. Gao et al., Feb 2020

340. Can a growing world be fed when the climate is changing? Dietz and Lanz, Feb 2020

339. MIT Scenarios for Assessing Climate-Related Financial Risk. Landry et al., Dec 2019


336. Did the shale gas boom reduce US CO2 emissions? Chen et al., Apr 2019


333. Statistical Emulators of Irrigated Crop Yields and Irrigation Water Requirements. Blanc, Aug 2018