Working Paper **18 June 2021**

Global Commons Stewardship Index:

A Statistical Review of the Pilot Methodology

Zachary A. Wendling[1], Reed Miller[2], Salma Dahir[1], Guillaume Lafortune[1], Daniel C. Esty[2], Guido Schmidt-Traub[3], Naoko Ishii[4], and Akiyuki Kawasaki[4]

[1] Sustainable Development Solutions Network (SDSN)

[2] Yale Center for Environmental Law & Policy, Yale University

[3] SYSTEMIQ

[4] Center for Global Commons, University of Tokyo

Abstract

This Working Paper is a companion to the Pilot Global Commons Stewardship (GCS) Index (SDSN et al., 2020), released at the Tokyo Forum in December 2020 by the Center for Global Commons (CGC) at the University of Tokyo. The Pilot Index Report contains details about its data and construction, especially in its appendices containing country profiles and methodology, but here, we present more detailed statistical analyses of the metrics, choices, and assumptions. This Working Paper provides additional transparency about the Pilot Index and discrete guidelines for improving further iterations. In particular, we address the categorization of the metrics, describe imputations and outliers, recommend the use of geometric means for aggregation, note areas needing greater data availability, and suggest new ways of interpreting this work in the narrative of the report. Increased scrutiny allows for deeper insights into the findings and conclusions of the Pilot Index, including understanding the robustness of the Pilot GCS Index to various alternative techniques.

About the SDSN

The UN Sustainable Development Solutions Network (SDSN) mobilizes scientific and technical expertise from academia, civil society, and the private sector to support practical problem solving for sustainable development at local, national, and global scales. The SDSN has been operating since 2012 under the auspices of the UN Secretary-General. The SDSN is building national and regional networks of knowledge institutions, solution-focused thematic networks, and the SDG Academy, an online university for sustainable development.

Table of Contents

Introduction

This Working Paper is a companion to the Pilot Global Commons Stewardship Index (SDSN et al., 2020), released at the Tokyo Forum in December 2020 by the Center for Global Commons at the University of Tokyo. Here, we present more detailed statistical analyses and methods behind the Pilot Index. These analyses aim to provide more information about the pilot results, sensitivity tests, and also aim to inform refinement in the methodology and indicators selected. Further details are provided in the Pilot Index report, especially its appendices containing country profiles and a detailed methodology.

Five major principles guided the design of the Pilot GCS Index. First, it uses a framework that integrates multiple dimensions of the Global Commons into a comprehensive assessment of impacts. Second, the Pilot GCS Index tracks impacts within territorial borders and transboundary impacts or spillovers through trade and physical flows. Spillovers are attributed to the country of final consumption. Third, we estimated the distance to target for all metrics in the Pilot GCS Index to quantify and compare priorities within and across countries. Fourth, the index focuses on outcome-based measures of environmental impacts at the country level. And fifth, the Index relies on data that are timely and can be updated regularly.

The design of the initial Pilot GCS Index follows established methods for building sound composite indicators. The construction of the Pilot GCS Index follows the various steps identified in the OECD-JRC Handbook on constructing composite indicators (Nardo & Saisana, 2008). We organized the index into two pillars covering domestic environmental impacts and spillovers, respectively (see Figure 1). Each pillar is divided into six sub-pillars: aerosols, biodiversity, climate change, land, oceans, and water. Scores are normalized for each indicator and first aggregated by sub-pillar, then by pillar, and finally for the entire index. We presented ratings based on per capita impacts and also identify the countries with the greatest absolute impacts on the Global Commons. In the initial edition, we were able to include 34 indicators using data from official and non-official sources. All metrics are globally relevant, statistically valid and reliable, up to date, collected according to internationally approved methods, and available for a large range of countries. The Pilot GCS Index reports results for 50 countries that have some of the greatest impacts on the Global Commons.

The Pilot GCS Index generated four initial findings:

1. Most countries generate large negative impacts on the Global Commons, but variations across countries are substantial. This variation generates opportunities for poor performers to learn from countries that generate lower per capita impacts on the Global Commons. Many developing countries have better ratings, but no country achieves the best or second-best rating (AAA and AA) on the index or within any of its sub-pillars. Small, rich countries with high trade-intensities rate worst on the Pilot GCS Index.

2. International spillovers account for a large share of countries' impacts on the Global Commons, particularly in relation to greenhouse gas emissions, biodiversity threats, and water scarcity embodied in imports.

3. In absolute terms, the greatest negative impacts arise from the world's largest economies: China, the United States, India, Japan, the EU, and Russia.

4. There are major gaps in availability and coverage of data for the Global Commons, particularly in relation to biodiversity loss at the genetic and population levels; disruptions to the phosphorus cycle; land degradation, especially from agriculture; hazardous waste; and water quality and scarcity.

This Working Paper presents a thorough analysis of the data and methodological choices used to construct the Pilot GCS Index and identifies key next steps.

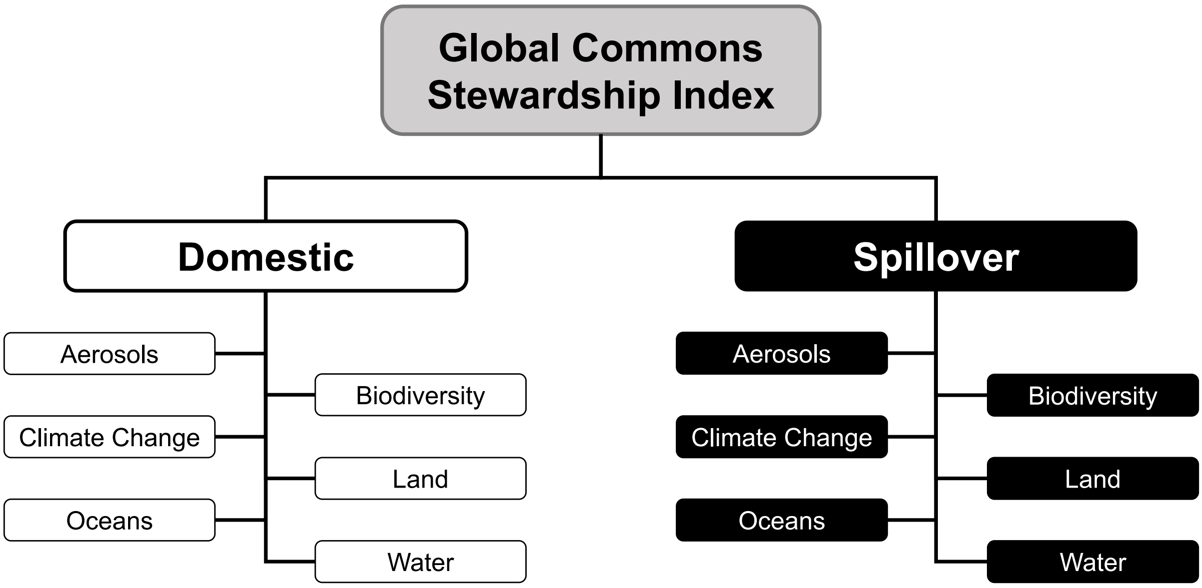
Abbreviations

CCPI	Climate Change Performance Index
CO_2	Carbon dioxide
EPI	Environmental Performance Index
EU	European Union
EVI	Environmental Vulnerability Index
GCI	Good Country Index
GCS	Global Commons Stewardship
GDP	Gross domestic product
GGI	Green Growth Index
GHG	Greenhouse gas
HPI	Happy Planet Index
JRC	Joint Research Centre (EU)
MRIO	Multi-regional Input-Output
NH_3	Ammonia
NO_X	Nitrogen oxides
OECD	Organisation for Economic Co-operation and Development
PCA	Principal Component Analysis
$PM_{2.5}$	Particulate Matter (≤ 2.5 µm in diameter)
SDG	Sustainable Development Goals
SDSN	Sustainable Development Solutions Network
SNMI	Sustainable Nitrogen Management Index
SO_2	Sulfur dioxide
UN	United Nations
UNEP	United Nations Environment Programme

1. Pilot Global Commons Stewardship Index

Based on the emerging definitions of the Global Commons and available data, we covered six sub-pillars in the Pilot GCS Index: aerosols, biodiversity, climate change, land, oceans, and water (Figure 1). For each sub-pillar we track domestic or territorial impacts and international spillovers or transboundary impacts.

Figure 1. Conceptual framework of categories within the Pilot Global Commons Stewardship Index



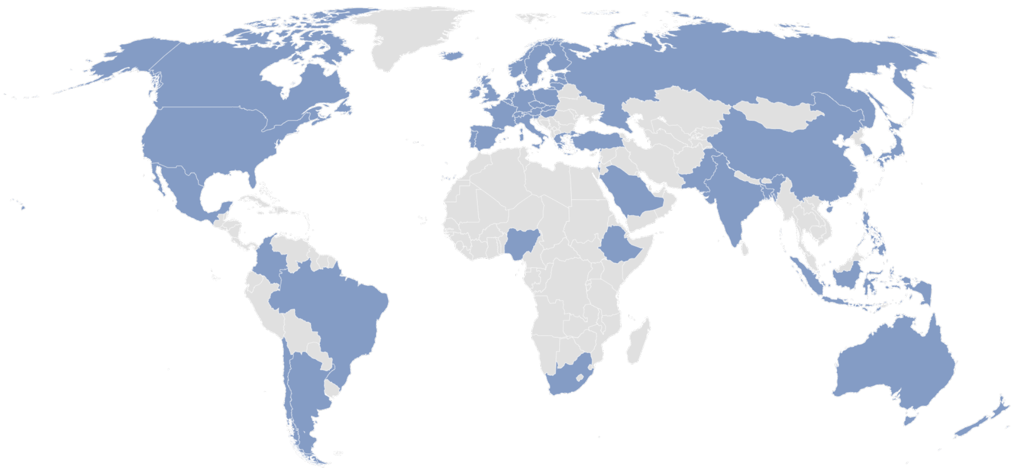
For the initial Pilot GCS Index, we identified a total of 34 indicators from a variety of sources – 23 domestic indicators and 11 spillover indicators (Table 1). As described in Appendix B of the report, the indicators are globally relevant, valid and reliable, up to date, collected according to internationally approved methods, and available for a large range of countries.

Table 1. Indicators included in the Pilot Global Commons Stewardship Index.

Pillar	Sub-Pillar	Indicator
Domestic	*Aerosols*	Domestic NOx emissions
		Domestic SO_2 emissions
		Annual mean concentration of $PM_{2.5}$
Domestic	*Biodiversity*	Terrestrial biodiversity threats embodied in domestic production
		Freshwater biodiversity threats embodied in domestic production
		Marine biodiversity threats embodied in domestic production
		Red List Index of species survival
		Mean area that is not protected in terrestrial sites important to biodiversity
		Mean area that is not protected in freshwater sites important to biodiversity
		Mean area that is not protected in marine sites important to biodiversity
Domestic	*Climate*	Domestic greenhouse gas (GHG) emissions
		Domestic black carbon emissions
Domestic	*Land*	Domestic ammonia emissions
		Sustainable Nitrogen Management Index
		Non-Recycled Municipal Solid Waste
		Permanent deforestation (5-year average)
		Human Trophic Level
Domestic	*Oceans*	Fish caught from overexploited or collapsed fish stocks
		Fish caught by trawling
Domestic	*Water*	Anthropogenic wastewater that does not receive treatment
		Nitrogen exportable to water bodies
		Domestic scarce water consumption
		Freshwater withdrawal
Spillover	*Aerosols*	NOx emissions embodied in imports
		SO_2 emissions embodied in imports
Spillover	*Biodiversity*	Terrestrial biodiversity threats embodied in imports
		Freshwater biodiversity threats embodied in imports
		Marine biodiversity threats embodied in imports
Spillover	*Climate*	Greenhouse gas emissions embodied in imports
		CO_2 emissions embodied in fossil fuel exports
Spillover	*Land*	Ammonia emissions embodied in imports
Spillover	*Oceans*	Ocean Health Index: Clean Waters
Spillover	*Water*	Nitrogen exportable to water bodies embodied in imports
		Scarce water consumption embodied in imports

While other global indices seek to include as many countries as possible, the Pilot GCS Index began with a smaller set. First, the current state of data availability means indicators can only meaningfully be constructed for certain countries. Second, a subset of countries will have economies and populations large enough to meaningfully impact the Global Commons. Our sample of countries included all members of the OECD and the G20. To this list, we also added the five next most-populous countries: Pakistan, Nigeria, Bangladesh, Ethiopia, Philippines. In total, the Pilot GCS Index included indicators for 50 countries:

Argentina	Czech Rep.	India	Netherlands	Slovak Rep.
Australia	Denmark	Indonesia	New Zealand	Slovenia
Austria	Estonia	Ireland	Nigeria	South Africa
Bangladesh	Ethiopia	Israel	Norway	South Korea
Belgium	Finland	Italy	Pakistan	Spain
Brazil	France	Japan	Philippines	Sweden
Canada	Germany	Latvia	Poland	Switzerland
Chile	Greece	Lithuania	Portugal	Turkey
China	Hungary	Luxembourg	Russia	United Kingdom
Colombia	Iceland	Mexico	Saudi Arabia	United States



Map 1. The 50 countries in the Pilot Global Commons Stewardship Index.

In the Pilot Index, we reported ratings instead of rankings and scores. The goal of the pilot was to propose a methodology and indicator set while acknowledging that further refinements would impact results. Full rankings would have provided a false impression of the relative performance of each country. Future versions of the GCS Index, following further refinements and the addition of new indicators, may well provide a sufficient basis for ranks and scores. We wished to avoid the appearance of a conflict between the results provided by more mature versions of the GCS Index and those results shown in the pilot version. Rating countries provides an ordinal scale on which performance can be measured without the danger of false precision. We use the following scale:

Table 2. Ratings categories as defined by score ranges in the Pilot GCS Index.

Rating	Score Range
AAA	100
AA	90–99
A	80–89
BBB	70–79
BB	60–69
B	50–59
CCC	0–49
no data	n/a

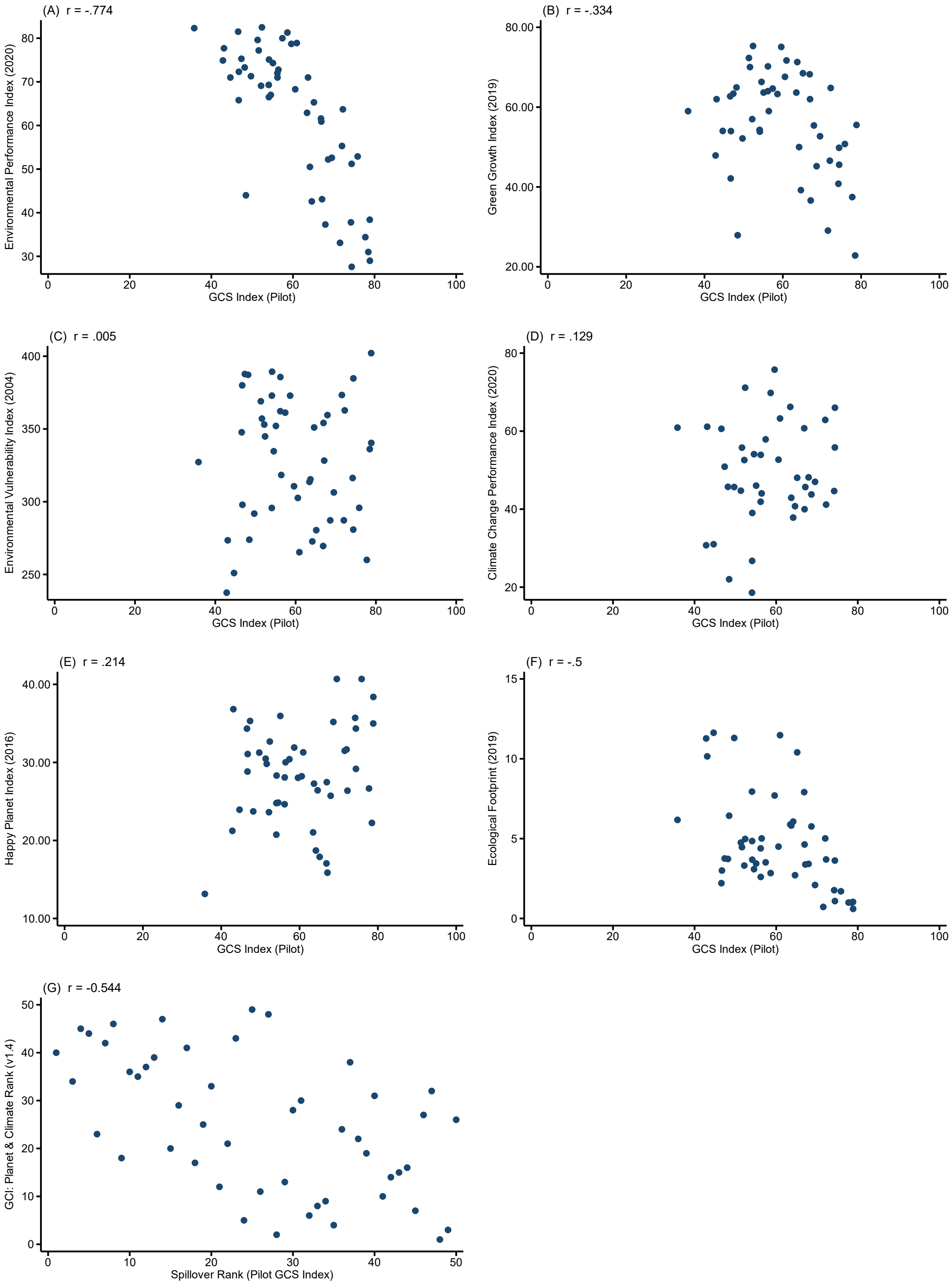
2. Unique contribution

As described in the pilot report, our goal with the GCS Index is to provide a novel perspective on threats to the Global Commons that transcend current approaches that are heavily rooted in territorial metrics or flawed conceptions of transboundary issues. Evidence of the distinction of the GCS Index, and the empirical support for our higher-level construct, lies in comparing our results with similar composite indices.

To gauge the added value of our project, we compared the results from our Pilot Index with the country scores on seven other composite indices. We use the most recent country scores on the Environmental Performance Index (EPI) (Wendling et al., 2020), the Green Growth Index (GGI) (Acosta et al., 2019), the Environmental Vulnerability Index (EVI) (SOPAC & UNEP, 2005), the Climate Change Performance Index (CCPI) (Burck et al., 2019), the Happy Planet Index (New Economics Foundation, 2016), the Ecological Footprint (Wackernagel et al., 2019), and the "Planet & Climate" category from the Good Country Index (GCI) (Anholt, 2021). We use our total GCS Index scores for these correlations, except for the correlation the Good Country Index where we correlate only the Spillover pillar since the Good Country Index focuses specifically on measuring countries' positive and negative contributions vis-à-vis the rest of the world (and not domestic performance).

Figure 2 shows the results of our correlation analysis between the Pilot GCS Index and similar composite indices. Overall, these provide robust evidence of an original and distinct contribution to this field. In four of the seven comparisons, the correlation coëfficient fails to be statistically significant, whereas in three, the indices are negatively correlated, *viz*., the EPI (Figure 2, Panel A, r = -0.774), the Ecological Footprint (Panel F, r = -0.500) and the GCI: Planet & Climate category (Panel G, r = -0.544). Besides differences in methods and data, the contrary findings in the Pilot Index might be attributed to the unique focus of our project. As our index expands beyond territorial metrics, as found in the EPI, and includes impacts along international supply chains, we capture a broader suite of threats to the Global Commons, revealing poor performance that is otherwise masked by traditional approaches.

Figure 2. Comparisons of Pilot GCS Index scores and ranks with other composite indices of environmental impacts.



3. Descriptive statistics

Table 3 shows descriptive statistics for each of the 34 indicators used in the Pilot Index. Rather than the raw values, cross-indicator comparisons are best made using the scores. We present a summary of the scores without the imputation of missing values.

3.1 Data availability

While data coverage is generally good for the 50 countries used in the Pilot, seven indicators have missing values. For four of these – unprotected marine sites, fish caught from overexploited or collapsed stocks, fishing by trawling, and the Clean Waters sub-score of the Ocean Health Index – missing data can be mostly explained by seven countries in our sample that are landlocked: Austria, the Czech Republic, Hungary, Luxembourg, Slovakia, Slovenia, and Switzerland. The indicator with the most missing values, non-recycled solid waste, is only available for OECD countries, highlighting a major gap in global coverage of this indicator. The other two indicators with missing values may be due to the inapplicability of the subject matter, as with Saudi Arabia and Iceland and deforestation (neither country has substantial forests) or Bangladesh and greenhouse gas (GHG) emissions embedded in exported fossil fuels (which Bangladesh no longer exports). The Pilot Index imputes missing values by using the regional average, though alternative methods for dealing with missing values could influence ultimate scores and ratings.

3.2 Skewness

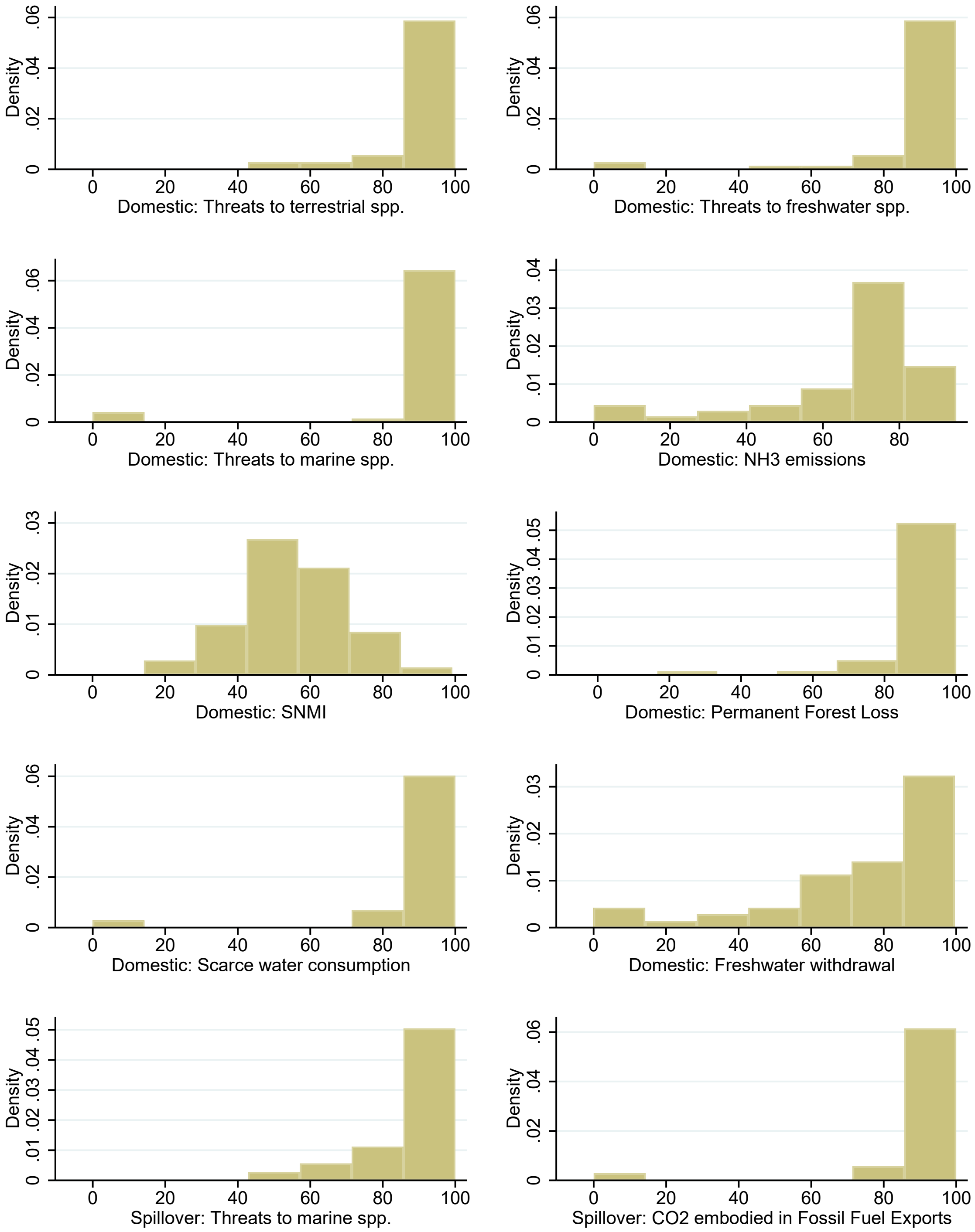
Outliers and skewed data can bias the scores at higher levels of aggregation. Winsorization during the scoring process limits the influence of extreme values. Distributions with long tails, however, may still need to be addressed through normalization techniques. In general, indicators with absolute skewness above 2.0 and kurtosis above 3.5 signify problematic distributions. Ten indicators, shaded in Table 3 and depicted in Figure 3, fulfill one or both of these criteria and so may be candidates for further treatment.

Table 3. Summary statistics for the 34 indicators used in the Pilot Index.

Indicator	Obs	Mean	Min	Median	Max	Skew	Kurtosis
Domestic NOx emissions	50	59.7	0.0	60.8	97.8	-0.74	3.38
Domestic SO_2 emissions	50	68.3	0.0	79.8	99.3	-1.09	3.06
Annual mean [PM2.5]	50	65.5	0.0	77.0	100.0	-1.07	2.94
Terrestrial biodiversity threats	50	92.0	43.6	97.5	100.0	-2.27	7.77
Freshwater biodiversity threats	50	90.3	0.0	97.5	99.9	-3.34	13.90
Marine biodiversity threats	50	89.9	0.0	98.5	100.0	-3.38	13.19
Red List Index	50	61.4	0.0	63.5	97.7	-0.48	2.09
Unprotected terrestrial sites	50	40.1	0.0	36.8	96.9	0.14	1.39
Unprotected freshwater sites	50	39.6	0.0	31.4	97.8	0.27	1.40
Unprotected marine sites	43	39.4	0.0	41.8	96.9	0.22	1.58
Domestic GHG emissions	50	52.6	0.0	53.0	94.9	-0.56	2.72
Domestic black carbon emissions	50	50.8	0.0	57.9	88.2	-0.63	2.26
Domestic NH_3 emissions	50	67.3	0.0	75.2	94.9	-1.61	5.11
Sustainable Nitrogen Mgmt. Index	50	55.4	23.9	55.6	99.2	0.28	4.06
Non-Recycled Municipal Solid Waste	36	45.5	4.4	46.0	79.2	-0.40	2.70
Permanent deforestation	48	94.0	29.9	99.3	100.0	-3.09	13.47
Human Trophic Level	50	62.8	41.7	59.9	95.3	0.74	3.29
Overexploited or collapsed fish stocks	39	35.7	0.0	29.5	97.2	0.42	1.81
Fish caught by trawling	43	58.0	12.2	61.8	100.0	-0.12	1.97
Untreated wastewater	50	56.6	0.0	62.6	100.0	-0.39	1.72
Nitrogen in water bodies	50	57.0	0.0	66.0	93.5	-0.88	2.73
Domestic scarce water consumption	50	92.5	0.0	98.6	100.0	-3.95	18.47
Freshwater withdrawal	50	74.7	0.0	83.0	99.6	-1.53	4.71
NOx emissions embodied in imports	50	57.1	0.0	56.7	99.6	-0.30	2.20
SO_2 emissions embodied in imports	50	51.4	0.0	48.8	99.6	0.02	1.74
Imported terrestrial biodiv. Threats	50	62.7	0.0	70.2	99.8	-0.66	2.15
Imported freshwater biodiv. Threats	50	74.9	14.9	83.5	99.9	-0.79	2.44
Imported marine biodiv. Threats	50	89.4	49.4	96.2	100.0	-1.39	4.06
GHG embodied in imports	50	49.1	0.0	44.5	99.6	0.15	1.72
CO_2 emissions in fossil fuel exports	49	93.8	0.0	99.9	100.0	-4.22	19.83
NH_3 emissions embodied in imports	50	52.1	0.0	50.6	99.7	-0.01	1.73
Ocean Health Index: Clean Waters	43	43.0	4.5	44.3	92.1	0.21	2.67
Imported Nitrogen in water bodies	50	59.6	0.0	58.2	99.8	-0.25	2.14
Imported scarce water consumption	50	58.4	0.0	61.5	99.7	-0.30	2.00

Notes: Cells shaded in the second column show indicators with missing values, *i.e.*, the number of observations is less than the number of countries, 50. Cells shaded in the columns for skewness and kurtosis show indicators with absolute values greater than or equal to 2.0 and 3.5, respectively.

Figure 3. Histograms of indicators with skewed distributions.



Note: *SNMI* = Sustainable Nitrogen Management Index.

4. Correlation analysis

Composite indices provide insights when they are able to summarize disparate facets of an overall concept. In theory, the various indicators should not be so coherent (collinear) that some metrics provide no additional information and lead to double counting and implicit weighting. Yet within a higher-level construct, the indicators should not contradict each other since each indicator should measure different dimensions of the same phenomenon. While we selected the indicators in this Pilot Index and arranged them in sub-pillars according to theoretically sound criteria and expert judgement, we measure our categorization through correlation analysis. The results provide further insights on our indicator selection and categorization.

4.1 Indicators

We first show the correlations between the 34 indicators and their respective sub-pillar, pillar, and overall score. If the indicators are capturing information about impacts to the Global Commons, we expect all correlations to be positive and statistically significant. But if each indicator has a unique contribution to the index, we would not expect the correlation coëfficients to approach the value of 1. Very high correlations are problematic because these indicators are effectively double-counted in the weighting within the index – as well as complicating the model without adding any value to the analysis.

Table 4 shows the results of the correlation analysis of the indicators. The Domestic pillar is more heterogeneous, with many weak and even negative correlations across indicators. Domestic indicators come from a variety of data sources and aim to capture a more diverse set of elements under each sub-pillar. By contrast, the Spillover pillar indicators are much more homogeneous. This correlation reflects the common underlying methods used to track spillover effects through extended multi-regional input-output (MRIO) tables. Biodiversity stands apart as a sub-pillar with high levels of correlation. One possible remedy would be to move the indicator on marine threats to the Ocean sub-pillar to reduce the threat of collinearity; another would be to merge terrestrial and freshwater biodiversity threats into a single indicator, though this metric would be less informative about impacts to the Global Commons. In the next version of the GCS Index, the indicator on CO_2 emissions in fossil fuel exports will be moved to the Domestic pillar, since it measures domestic emissions for exporting fossil fuels.

Table 4. Correlations between indicators, their respective sub-pillars and pillars, and the overall score in the Pilot Index.

Indicator	Sub-Pillar	Pillar	Overall
Domestic NOx emissions	0.624	0.579	0.689
Domestic SO_2 emissions	0.686	0.380	0.501
Annual mean [PM2.5]	0.215	0.323	-0.027
Terrestrial biodiversity threats	0.359	0.325	0.052
Freshwater biodiversity threats	0.266	0.340	0.218
Marine biodiversity threats	0.547	0.594	0.274
Red List Index	0.762	0.208	-0.338
Unprotected terrestrial sites	0.879	0.421	-0.213
Unprotected freshwater sites	0.905	0.543	-0.124
Unprotected marine sites	0.816	0.378	-0.289
Domestic GHG emissions	0.999	0.697	0.725
Domestic black carbon emissions	0.179	0.103	-0.067
Domestic NH_3 emissions	0.679	0.348	0.178
Sustainable Nitrogen Mgmt. Index	-0.026	-0.010	-0.118
Non-Recycled Municipal Solid Waste	0.766	0.430	0.230
Permanent deforestation	0.105	-0.090	-0.514
Human Trophic Level	0.512	0.090	0.595
Overexploited or collapsed fish stocks	0.823	0.401	0.314
Fish caught by trawling	0.708	0.395	0.361
Untreated wastewater	0.709	0.073	-0.658
Nitrogen in water bodies	0.083	0.291	0.093
Domestic scarce water consumption	0.736	0.386	0.049
Freshwater withdrawal	0.651	0.285	-0.012
NOx emissions embodied in imports	0.988	0.863	0.784
SO_2 emissions embodied in imports	0.986	0.863	0.787
Imported terrestrial biodiv. Threats	0.978	0.846	0.772
Imported freshwater biodiv. Threats	0.964	0.807	0.725
Imported marine biodiv. Threats	0.814	0.579	0.661
GHG embodied in imports	0.879	0.934	0.870
CO_2 emissions in fossil fuel exports	0.624	0.175	0.271
NH_3 emissions embodied in imports	n.a.	-0.055	-0.233
Ocean Health Index: Clean Waters	n.a.	0.978	0.898
Imported Nitrogen in water bodies	0.967	0.974	0.896
Imported scarce water consumption	0.972	0.883	0.857

Notes: Cells shaded in gray show correlation coëfficients whose absolute value falls below the 1% significance level (r_{crit} = 0.363). Cells shaded in red show correlation coëfficients that are both statistically significant and negative. Cells shaded in blue show correlations coëfficients that are very high ($r \geq 0.92$). *n.a.* = not applicable, for sub-pillars with only one indicator.

4.2 Sub-Pillars

We next examine the correlations between the sub-pillars and pillars and their relationship to the overall score, shown in Table 5. This analysis confirms the higher homogeneity of the Spillover pillar compared with the Domestic pillar. The Oceans sub-pillar is measured through the Ocean Health Index and not extended MRIO (unlike all the other indicators included under the Spillover sub-pillar). Currently our spillover indicator set does not capture spillovers embodied into physical flows (air and water), which may not be correlated with trade-related spillovers.

Table 5. Correlations between sub-pillars, pillars, and the overall score in the Pilot Index.

	Overall	Pillar	Aerosols	Biodiv.	Climate	Land	Oceans
Domestic	0.460						
Aerosols	0.313	0.570					
Biodiversity	-0.146	0.572	0.130				
Climate Change	0.716	0.697	0.559	0.054			
Land	0.319	0.369	-0.025	0.165	0.256		
Oceans	0.432	0.492	-0.075	0.048	0.200	0.261	
Water	-0.345	0.432	0.204	0.474	-0.073	-0.118	0.002
Spillover	0.909						
Aerosols	0.796	0.875					
Biodiversity	0.780	0.831	0.559				
Climate Change	0.825	0.831	0.844	0.598			
Land	0.898	0.978	0.853	0.814	0.807		
Oceans	-0.247	-0.097	-0.325	-0.103	-0.415	-0.185	
Water	0.903	0.955	0.802	0.817	0.761	0.956	-0.189

Notes: Cells shaded in gray show correlation coëfficients whose absolute value falls below the 1% significance level (r_{crit} = 0.363). Cells shaded in red show correlation coëfficients that are both statistically significant and negative. Cells shaded in blue show correlations coëfficients that are very high ($r \geq 0.92$).

5. Principal component analysis

We also conduct a principal component analysis (PCA) of the scores in the Pilot Index to examine how coherently our indicators fit within the category framework. First, we run PCAs of the two sets of indicators, domestic and spillover. Second, we run a second PCA of the sub-pillar scores to see how these contribute to the pillar scores. Our results provide an assessment of both the empirical strength of our categorization and the contributions of indicators to the overall analysis.

5.1 Indicators

Table 6 shows results from the PCAs on the 23 domestic and 11 spillover indicators. For the domestic indicators, the PCA shows seven or eight components with eigenvalues greater than or close to 1, which explain 79–83% of the total variance (Table 6). This is more than the six sub-pillars in our pilot categorization scheme, and further visualization of the results in Figure 4, Panel A shows that the components do not map thematically onto those sub-pillars, either. For example, the Biodiversity indicators load onto separate components.

Table 6. Results of a Principal Component Analysis of the indicators.

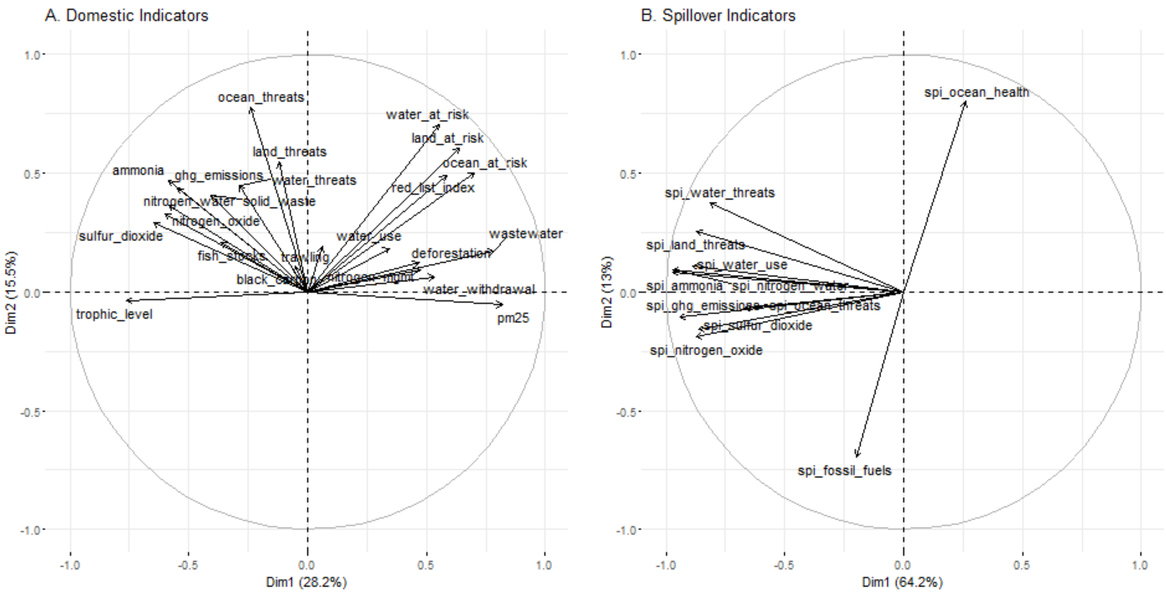
Domestic Indicators

Component	Eigenvalue	Prop.	Cum.
Comp1	6.49	0.28	0.28
Comp2	3.56	0.16	0.44
Comp3	2.34	0.10	0.54
Comp4	1.88	0.08	0.62
Comp5	1.51	0.07	0.69
Comp6	1.31	0.06	0.74
Comp7	1.02	0.04	0.79
Comp8	0.91	0.04	0.83
Comp9	0.74	0.03	0.86
Comp10	0.61	0.03	0.89
Comp11	0.52	0.02	0.91
Comp12	0.49	0.02	0.93
Comp13	0.38	0.02	0.95
Comp14	0.36	0.02	0.96
Comp15	0.21	0.01	0.97
Comp16	0.17	0.01	0.98
Comp17	0.14	0.01	0.98
Comp18	0.11	0.00	0.99
Comp19	0.10	0.00	0.99
Comp20	0.06	0.00	1.00
Comp21	0.04	0.00	1.00
Comp22	0.03	0.00	1.00
Comp23	0.02	0.00	1.00

Spillover Indicators

Component	Eigenvalue	Prop.	Cum.
Comp1	7.06	0.64	0.64
Comp2	1.43	0.13	0.77
Comp3	1.18	0.11	0.88
Comp4	0.62	0.06	0.94
Comp5	0.27	0.02	0.96
Comp6	0.19	0.02	0.98
Comp7	0.12	0.01	0.99
Comp8	0.06	0.01	0.99
Comp9	0.04	0.00	1.00
Comp10	0.02	0.00	1.00
Comp11	0.01	0.00	1.00

Figure 4. Factor maps of the indicators used in the Pilot Index.



For the spillover indicators, only three components explain 88% of the total variance (Table 6) – half of the number of sub-pillars into which we classify these impacts. As suggested by the results in Table 4, Figure 4, Panel B shows that two indicators, the Clean Waters component of the Ocean Health Index and GHG embodied in exported fossil fuels, load onto two distinct components, while all other indicators, generated from MRIO tables, load onto a single factor that explains nearly two-thirds of the total variance.

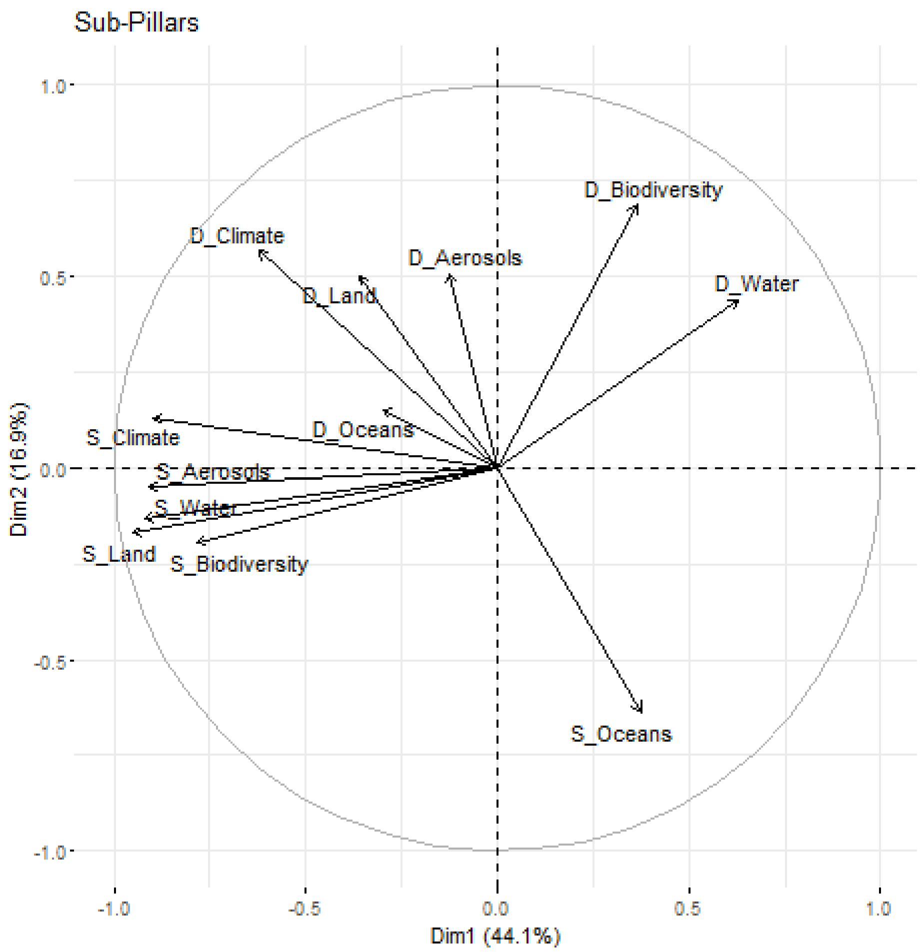
5.2 Sub-Pillars

We similarly analyze all twelve sub-pillar scores, with the results of the PCAs shown in Table 7. This analysis shows that the scores load onto four components, explaining 82% of the total variance. Mapping these factors visually, in Figure 5, reveals the patterns of association. The first pattern that emerges is a close correspondence among the Spillover sub-pillars. The only outlier is the Spillover sub-pillar for Oceans; as noted above, this is likely due to the unique conceptual and methodological basis for its lone indicator. Otherwise, the Domestic sub-pillars are largely dispersed, reflecting the lower overall coherence within this pillar.

Table 7. Results of a Principal Component Analysis of the sub-pillars.

Component	Eigenvalue	Prop.	Cum.	Component	Eigenvalue	Prop.	Cum.
Comp1	5.30	0.44	0.44	Comp7	0.39	0.03	0.94
Comp2	2.03	0.17	0.61	Comp8	0.27	0.02	0.96
Comp3	1.35	0.11	0.72	Comp9	0.17	0.01	0.98
Comp4	1.14	0.10	0.82	Comp10	0.14	0.01	0.99
Comp5	0.64	0.05	0.87	Comp11	0.08	0.01	0.99
Comp6	0.46	0.04	0.91	Comp12	0.03	0.00	1.00

Figure 5. Factor maps of the Sub-pillars used in the Pilot Index.

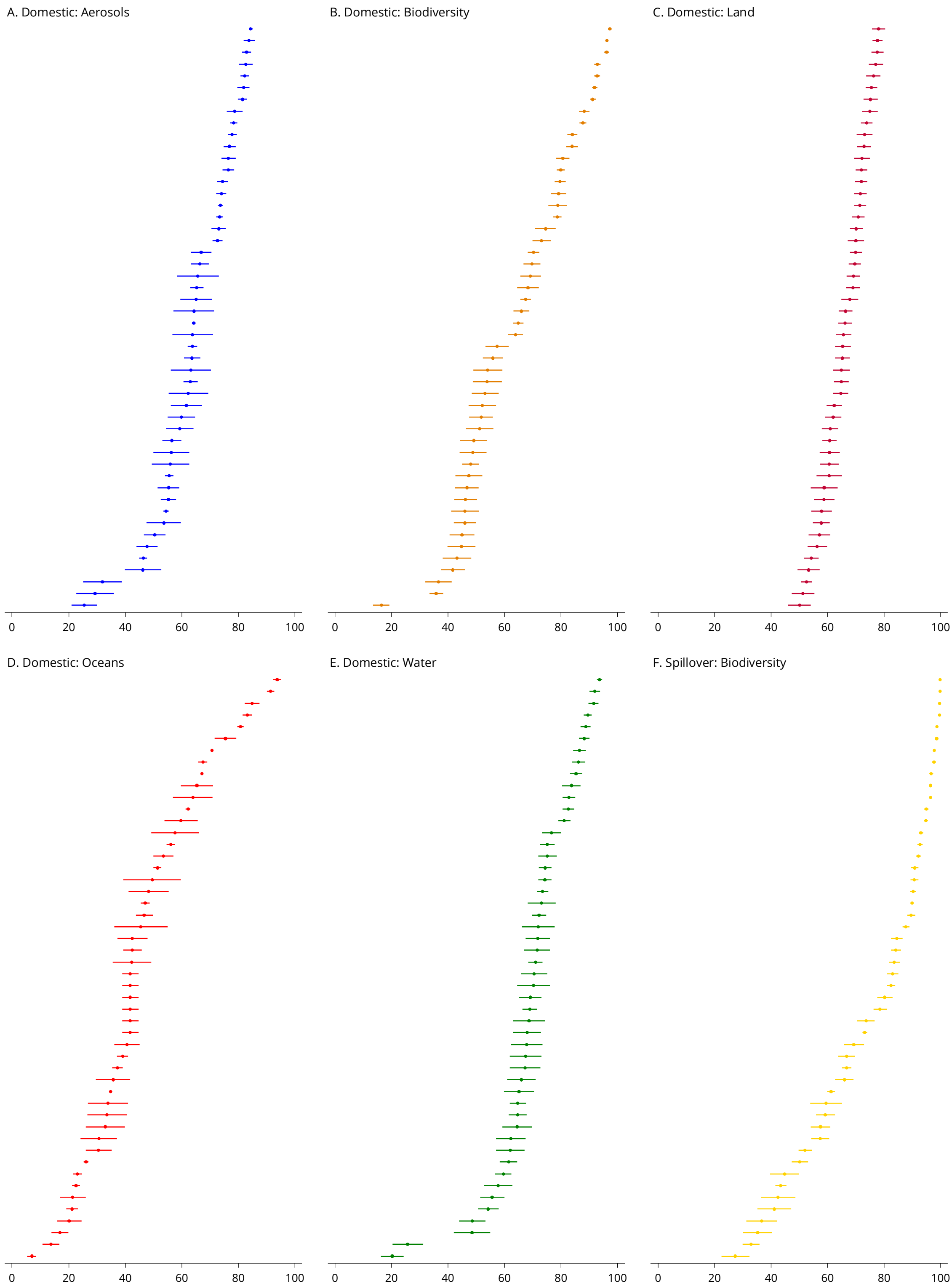


6. Score sensitivity to weighting scheme

There are four main ways to generate weights for building composite indicators (Lafortune et al., 2018): equal weights, expert weights, mathematical weights, and subjective or user-driven weights. For the sake of simplicity, the Pilot GCS Index uses equal weights, with the exception of the Domestic sub-pillar for Climate Change, which is based on expert judgement. The mathematical technique would use the results of PCA to derive weights, but as shown in the previous section, we cannot rely on PCA because our theoretical framework and indicator categorization does not fully align with the components suggested by this analysis. Subjective weights are most feasible as a *post hoc* tool that allows users to emphasize aspects of the index most important to them, and we can foresee subsequent versions of the GCS Index to include such an interface online. Ultimately, the GCS Index needs a default weighting scheme, and equal weighting is the most transparent and intuitive approach.

As a transparency test, we provide a sensitivity analysis in which we allow the original weights to vary ±25%. For example, in our Domestic Land sub-pillar, the five indicators have a weight of 0.20 each. Under our uncertainty analysis, we generate new weights that take random values in the interval 0.15–0.25 (or, 0.20 ± 0.05) and still sum to 1. To survey the impacts of this variation, we conduct a Monte Carlo simulation, repeating this process 1000 times, and compare the aggregated scores. Based upon the analyses in the previous sections, we are particularly interested in gauging the impact of the choice of weights within six of the twelve sub-pillars, shown in Figure 6, and within both pillars, shown in Figure 7.

Figure 6. Monte Carlo simulation of sub-pillar scores using 1000 draws of weights ±25% of the original scheme. Each panel shows the original score for the 50 countries with a 90% confidence interval of the aggregated scores.



6.1 Sub-pillars

Within the six sub-pillars analyzed for alternative weights, the results show low to moderate sensitivity to the choice of weights. Figure 6 shows country scores for the sub-pillars with 90% confidence intervals. In general, top-scoring countries are relatively insensitive to weighting schemes, with mid- to lower-scoring countries having wider confidence intervals. This pattern fits our expectations, as high aggregate scores come from generally high scores among all indicators, whereas there is more variation in performance further down. Changes in relative rankings of countries capture the practical effects of alternative weights, which are summarized in Table 8. Two sub-pillars are relatively insensitive to the choice of weights: Biodiversity impacts in both the Domestic and Spillover pillars, with low mean and maximum change in ranks, and few or no countries with substantial 90% confidence intervals. The choice of weights is more influential within the Domestic sub-pillars for Aerosols, Land, and Water, with at least half of all countries varying in rank by five places or more. This variation in ranks indicates uneven scores across the indicators within these sub-pillars for many countries.

Table 8. Descriptive statistics for the 90% confidence intervals in country ranking from 1000 alternative weighting schemes for six sub-pillar scores.

				Number of countries with ranking range …	
Sub-Pillar	Mean	Min	Max	> 5	> 10
Domestic					
Aerosols	7.9	1.3	18.4	26	16
Biodiversity	2.4	0	8.2	4	0
Land	6.0	1.0	16.7	26	6
Oceans	4.6	0	20.7	11	6
Water	6.9	0	22.2	25	14
Spillover					
Biodiversity	1.3	0	4.1	0	0

6.2 Pillars

A Monte Carlo simulation using alternative weighting schemes shows moderate changes in the aggregated pillar scores. Figure 7 shows all 50 countries with a 90% confidence interval centered on the original pillar scores. The width of the confidence intervals is roughly comparable between pillars and across the spectrum of countries. The rankings of the countries under each weighting scheme show the practical effects of deviating from equal weighting, summarized in Table 9. Here, the Domestic pillar is much more sensitive to the choice of weights, with a higher mean spread of rankings and 30 countries showing a 90% confidence interval of five places or more. As shown in previous sections, there are high levels of correlation among Spillover sub-pillar scores and low correlations among Domestic sub-pillars scores, which corresponds to the findings here.

Figure 7. Monte Carlo simulation of pillar scores using 1000 draws of weights ±25% of the original scheme. Each panel shows the original score for the 50 countries with a 90% confidence interval of the aggregated scores.

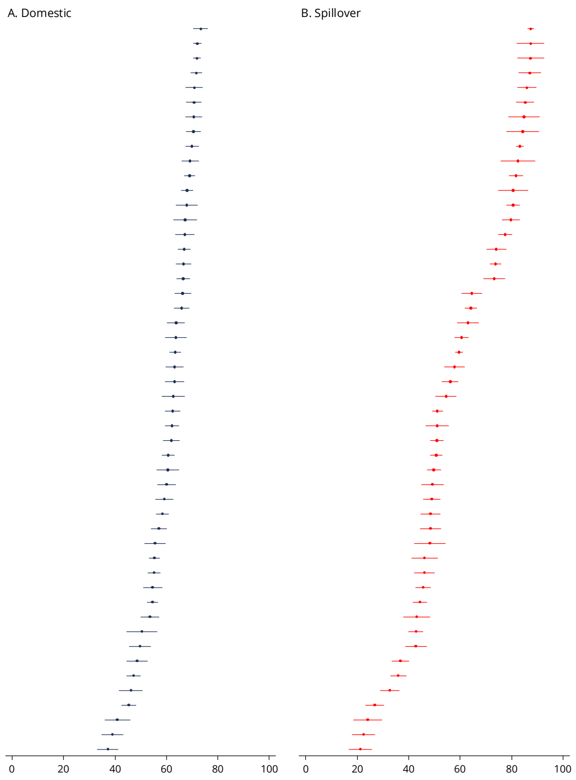


Table 9. Descriptive statistics for the 90% confidence intervals in country ranking from 1000 alternative weighting schemes for the pillar scores.

				Number of countries with ranking range …	
Pillar	Mean	Min	Max	> 5	> 10
Domestic	5.8	0.6	13.4	30	4
Spillover	3.3	0	10.1	10	1

7. Score sensitivity to aggregation technique

We aggregate the scores in the Pilot Index using the arithmetic mean. While simple, this approach has the disadvantage of implying that indicators are fully *compensatory*, that is, harmful impacts in one issue may be compensated for by good performance in another. For example, in the sub-pillar of Domestic Land impacts, a country might harm the Global Commons through permanent deforestation yet make up for this harm through high levels of recycling of municipal solid waste. An arithmetic mean offsets the low score on the former with the high score on the latter. When aggregation combines disparate kinds of impacts, as often is the case in this Pilot Index, the assumption that scores are compensatory may not be warranted. To return to our example, the losses in ecosystem services from permanent deforestation are hardly mitigated by high levels of waste recycling. In order to test the results of this assumption empirically, we recalculate the aggregated scores in the Pilot Index using the geometric mean, a non-compensatory technique.

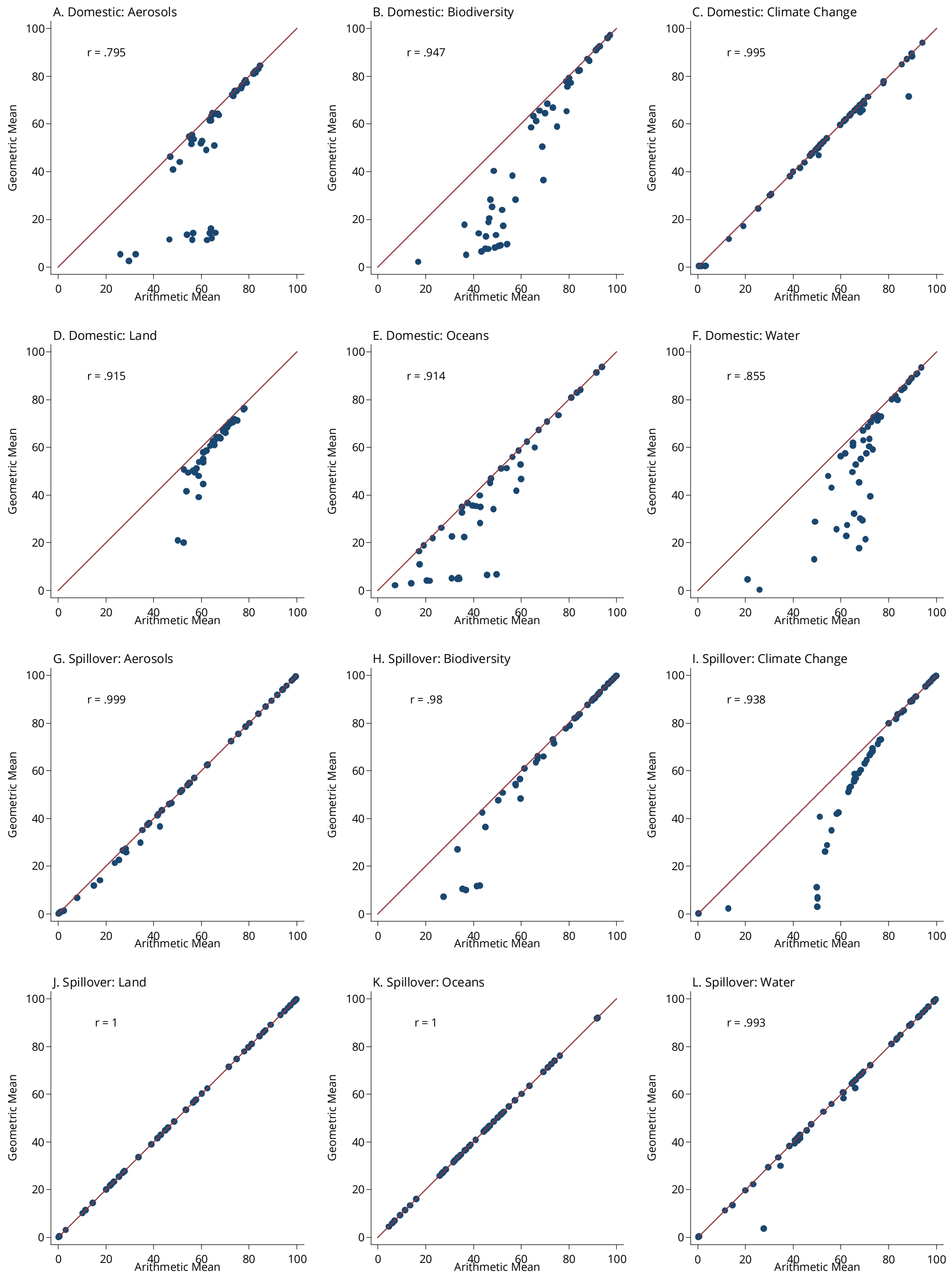
7.1 Sub-pillars

Figure 8 compares the scores for all twelve sub-pillars using the arithmetic and geometric means. These results show that the scores are robust to the choice of aggregation technique only in a few sub-pillars that have one or two indicators, shown in Panels C, G, J, K, and L. Otherwise, the changes in scores can be substantial, described in Table 10. Figure 8 further illustrates that the largest changes in sub-pillar scores happen among the countries with low indicator scores, who otherwise benefit from the compensatory effect of arithmetic aggregation. In some sub-pillars, the practical effects on country rankings are modest, as in Panels D, H, and I, whereas in other sub-pillars, the rankings under the geometric aggregation are substantially different, as in Panels A, B, and E.

Table 10. Descriptive statistics for changes in sub-pillar scores between arithmetic and geometric aggregation techniques.

	Domestic				*Spillover*			
		Difference in scores				Difference in scores		
Sub-pillar	r	Mean	Min	Max	r	Mean	Min	Max
Aerosols	0.795	11.7	0	52.3	0.999	0.6	0	6.0
Biodiversity	0.947	16.6	0	44.5	0.980	3.9	0	30.8
Climate Change	0.995	1.0	0	16.8	0.938	9.0	0	47.0
Land	0.915	5.1	1.4	32.4	1	–	–	–
Oceans	0.914	8.0	0	42.9	1	–	–	–
Water	0.885	12.7	0.2	49.7	0.993	0.9	0	24.0

Figure 8. Comparison of sub-pillar scores using arithmetic and geometric means and equal weights.



7.2 Pillars and overall scores

We also test the effects of geometric means in the aggregation of sub-pillars into our two pillars, Domestic and Spillover, and of the two pillars into the overall scores. Figure 9 illustrates the results from the geometric and arithmetic aggregations, and Table 11 summarizes the changes in pillar and overall scores. Again, the differences in pillar scores are most pronounced for countries at the lower end of impacts on the Global Commons, though the overall changes in scores and rankings are relatively modest. In aggregating the overall score, the results are very robust to the choice of aggregation technique, with little change in the score or rankings.

Figure 9. Comparison of pillar and overall scores using arithmetic and geometric means and equal weights.

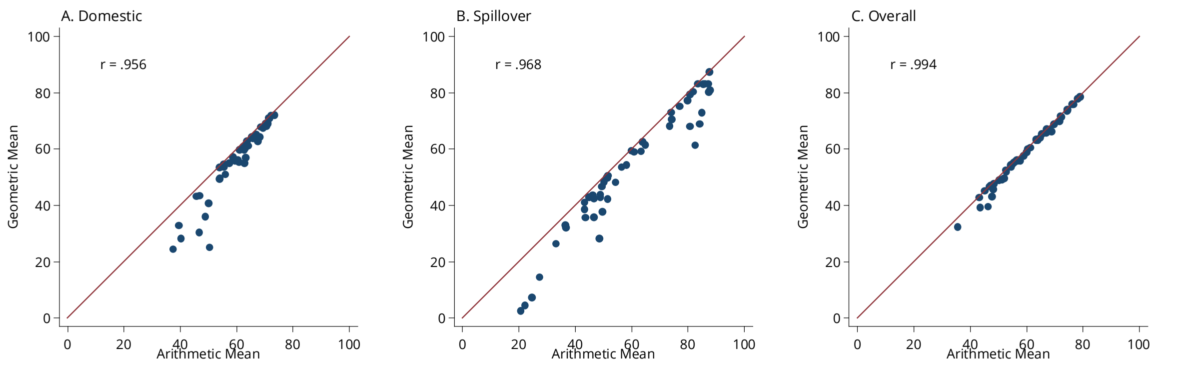


Table 11. Descriptive statistics for changes in pillar and overall scores between arithmetic and geometric aggregation techniques.

		Difference in scores		
	r	Mean	Min	Max
Domestic	0.956	4.1	0.4	25.3
Spillover	0.968	6.0	0.2	21.1
Overall	0.994	1.1	0	6.8

8. Conclusions and future directions

Calculating a composite index requires making many choices and assumptions. This Working Paper, including multivariate analyses and various sensitivity tests, provides further insights on how to refine the methodology, indicator selection and categorization. The sensitivity analyses conducted in this Working Paper also provide additional transparency about the construction and conclusions of the report and offer insights for understanding the robustness of the Pilot GCS Index to various weighting and aggregation schemes.

Based on the findings reviewed in this Working Paper, we underline six key findings which will inform our future work and next versions of the GCS Index.

1. Adjust our indicator categorization.

The multivariate analyses presented in this Working Paper – but also expert judgement – suggests some adjustments to our indicator selection and categorization. These include

- Moving the indicator on CO_2 emissions embodied into fossil fuel exports to the Domestic pillar, and
- Moving the indicators on marine biodiversity threats to the Oceans sub-pillar within both the Domestic and Spillover pillars.

2. Refine our imputation method and method for dealing with outliers (especially for highly skewed data).

We might revisit our approach for imputing values in cases of missing data, for instance on oceans indicators for landlocked countries. The skewness and kurtosis analysis presented in this Working Paper calls for increased attention on how outlier values are treated for 10 of the 34 indicators.

3. Use the geometric mean instead of the arithmetic mean as our main aggregation method at the pillar and sub pillar level.

Correlation and Principal Component Analysis underline the greater heterogeneity of the domestic pillar. Arithmetic means allow countries to balance out low scores in some indicators with high scores in others. The implication that these indicators are compensatory is not warranted on theoretical grounds. Therefore, this analysis indicates that the next iteration of the GCS Index could use geometric means as a non-compensatory aggregation method.

4. Work with partners to increase data availability, especially to track spillovers embodied into physical flows (*e.g.*, air and water).

The Pilot GCS Index provides an overview of countries' domestic and spillover impacts on the Global Commons. Yet, data gaps impede our ability to provide a comprehensive and adequate picture. One important data gap is related to the attribution of spillovers generated through physical flows (air, water) to the countries of origin. The homogeneity of the spillover pillar reflects the fact that we currently measure only one type of spillover – those embodied into trade.

5. Call for prudence in interpreting small differences in countries' ratings and scores, especially on the Domestic pillar.

Our Monte Carlo simulation reveals that scores, especially for the Domestic pillar, are sensitive to changes in the weighting schemes. The Pilot GCS Index published in December 2020 presented "ratings" instead of "scores and ranks" to reflect the preliminary and on-going nature of the work. As we consider presenting scores and ranks in the next version of the GCS Index, this sensitivity suggests that we should emphasize that small difference in scores and ranks should not be over-interpreted. Scores and ranks on the Spillover pillar are, however, very robust.

6. Dig deeper into countries' results, including supply chains and sectors responsible for trade-related spillover impacts.

This Working Paper and Pilot GCS Index report focused extensively on the overall country findings. Yet, the policy response and stakeholders involved vary across industries, sectors, and commodities. Looking ahead, further insights could be generated by looking at the sectors, supply chains, time series, and geographic impacts of international spillovers to support better policies to address spillovers based on robust science and data.

9. References

Acosta, L. A., Maharjan, P., Peyriere, H., Galotto, L., Mamiit, R. J., Ho, C., Flores, B. H., & Anastasia, O. (2019). *Green Growth Index: Concepts, Methods and Applications* (No. 5 GGGI Technical Report). Green Growth Performance Measurement (GGPM) Program, Global Green Growth Institute. http://greengrowthindex.gggi.org/

Anholt, S. (2021). *Good Country Index*. https://index.goodcountry.org/

Burck, J., Hagen, U., Höhne, N., Nascimento, L., Bals, C., & Germanwatch. (2019). *Climate Change Performance Index Results 2020*.

Lafortune, G., Fuller, G., Moreno, J., Schmidt-Traub, G., & Kroll, C. (2018). *SDG Index and Dashboards. Detailed Methodological paper*. Bertelsmann Stiftung and Sustainable Development Solutions Network. http://sdgindex.org/assets/files/2018/Methodological%20Paper_v1_gst_jmm_Aug2018_FINAL.pdf

Nardo, M., & Saisana, M. (2008). *OECD/JRC Handbook on Constructing Composite Indicators, Putting Theory into Practice*.

New Economics Foundation. (2016). *Happy Planet Index*. http://happyplanetindex.org/

SDSN, Yale Center for Environmental Law & Policy, & Center for Global Commons at the University of Tokyo. (2020). *Pilot Global Commons Stewardship Index*.

South Pacific Applied Geoscience Commission & United Nations Environment Programme. (2005). *Environmental Vulnerability Index*. http://www.vulnerabilityindex.net/

Wackernagel, M., Beyers, B., & Rout, K. (2019). *Ecological footprint: Managing our biocapacity budget*. https://www.footprintnetwork.org/

Wendling, Z. A., Emerson, J. W., de Sherbinin, A., Esty, D. C., & et al. (2020). *2020 Environmental Performance Index*. Yale Center for Environmental Law & Policy. epi.yale.edu