Proceedings of the EUROPEAN ACADEMY of Sciences & Arts PEASA

Proceedings of the European Academy of Sciences & Arts, doi: 10.4081/peasa.12

Open Access. Online

doi: 10.4081/peasa.13

Topics: ORIGINAL RESEARCH

Category: TECHNICAL & ENVIRONMENTAL SCIENCES

ON A MATHEMATICAL THEORY OF SUSTAINABILITY ASSESSMENT

Yannis A. Phillis*[1], Vassilis S. Kouikoglou[1], Evangelos Grigoroudis[1], Fotis D. Kanellos[2]

[1]School of Production Engineering and Management, Technical University of Crete, Chania 73100, Greece
[2]School of Electrical and Computer Engineering, Technical University of Crete, Chania 73100, Greece

Received: 23/07/2022
Accepted: 12/10/2022

**Corresponding author: Yannis A. Phillis (phillis@dpem.tuc.gr)*

ABSTRACT

We compile a number of postulates to be satisfied by a function that mathematically assesses sustainability of some type: national, urban, energy, etc. These postulates lay the mathematical foundations of a sustainability theory and lead to a simple model based on shifted geometric means combining basic sustainability indicators into an overall index. The model has a number of desirable properties and generalizes the weighted arithmetic and weighted geometric means, which are commonly used aggregation functions in sustainability assessments. Numerical results demonstrate the closeness of the model to other established techniques of sustainability.

KEYWORDS: Sustainability assessment, mathematics of sustainability, sustainability indicators, data aggregation

1. INTRODUCTION

Sustainability assessments of an entity, be it a nation, a city, an energy or a transportation system etc., rely on data to evaluate human welfare and environmental integrity. Each of these two fundamental components is a combination of more specialized indicators organized into various hierarchical levels. For example, the human dimension of sustainability encompasses socio-economic, technological and political aspects which are further elaborated using different groups of indicators. Most assessment models normalize indicators from their physical domains into a common interval representing a range from the lowest to the highest levels of sustainability. Normalized indicators of the same group are then combined into more composite indicators through aggregation functions which capture the relationships among and the relative importance of indicators belonging to the same dimension. Composite indicators are further aggregated by exploiting the hierarchy from bottom to top to eventually arrive at a single numerical value of overall sustainability.

Existing approaches to measure sustainability differ in scope, suite of indicators, and normalization and aggregation procedures. We review some commonly used definitions and assessment frameworks of national sustainability and sustainable development.

The Human Development Index (HDI) uses four indicators: life expectancy at birth, expected years of schooling for children, mean years of schooling for adults aged 25 years and older, and the logarithm of gross national income per capita. The indicators are first transformed into a scale from 0 to 1 by a linear interpolation between thresholds of unsustainable and sustainable values and then combined into an overall index using arithmetic and geometric means. HDI was first released in 1990 and is updated annually by the United Nations Development Programme (2022).

The Environmental Performance Index (EPI) measures the closeness of a country's performance to established environmental policy targets. Its latest version (Wolf et al., 2022) uses 40 indicators relevant to three policy objectives: climate change, environmental integrity, and ecosystem vitality. The indicators are first converted to dimensionless numbers from 0 to 100 and then successively aggregated using weighted arithmetic means into more composite indicators and the overall index. Missing data are imputed either via predictive models involving past or correlated variables or by averaging data of neighboring countries. For some indicators the imputed values include a penalty for failing to report information. EPI is successor to the Environmental Sustainability Index (ESI), which contained additional indicators to assess social and political aspects of sustainability (Esty et al., 2005).

Sustainability Assessment by Fuzzy Evaluation (SAFE) is a hierarchical fuzzy system whose first version appeared in 2001. Its most recent release (Grigoroudis et al., 2021) uses 69 time series of indicators grouped in various components of ecological and human sustainability. The Pressure-State-Response classification of OECD (1991) is used to describe each component. Pressure indicators assess the negative impacts on the corresponding component, state describes the prevailing conditions, and response indicators reflect the actions taken to improve the state. Each indicator time series is transformed into a single value which captures the latest trends, it is then normalized in [0, 1], and finally it is combined with other normalized indicators through a sequential fuzzy reasoning process to obtain the SAFE index.

A model of national environmental sustainability proposed by Liu (2007) uses the indicators of ESI. The Analytic Hierarchy Process (Saaty, 1980) is used to assign weights to and aggregate well-defined indicators such as water quality, while a fuzzy reasoning scheme similar to SAFE is used to aggregate composite indicators using subjective criteria and qualitative information.

The Sustainable Society Index (SSI) introduced by van de Kerk and Manuel (2008) is a combination of separate indices corresponding to three dimensions: Human, Environmental and Economic Wellbeing. Each dimension is assessed using five to nine indicators which are normalized and aggregated using geometric means. All variables are assigned a numerical value from 1 (weakest sustainability) to 10 (strongest sustainability). SSI is updated every second year since 2006.

All the above models provide country rankings. About one hundred other models of sustainability, sustainable development, and human well-being have been surveyed by Yang (2014). Recent comprehensive reviews on weighting and aggregation methods for constructing composite indicators can be found in Gan et al. (2017) and Greco et al. (2019).

The simplest aggregation functions are the arithmetic and geometric means and their weighted counterparts, also known as additive and multiplicative functions. These are the most commonly used models for sustainability assessments. However they have certain drawbacks that limit their applicability. For example, an additive function cannot be used to describe an indicator that is critically important for the sustainability of the whole entity since a decline of such an indicator below a certain unsustainability level can, by the additive property, be compensated for by an increase in sustainability levels of other indicators. Multiplicative functions have the opposite property, i.e., they treat all indicators as being critically important. Fuzzy logic and other universal function approximation tools overcome these limitations, sometimes at the price of increased computational complexity and need for sophisticated parameter tuning.

In the present paper we propose an extension to the geometric mean which avoids the drawbacks of the previous models. The new function satisfies a number of sustainability postulates and provides sustainability assessments and country rankings which have strongly positive correlations with some of the models reviewed above and variants thereof.

2. AGGREGATION OF SUSTAINABILITY INDICATORS

2.1. *Symmetry and homogeneity*

We start with n indicators relevant to an entity whose sustainability is assessed. We assume that these indicators are normalized into dimensionless numbers x_i on [0, 1], $i = 1, \ldots, n$, where 0 corresponds to a range of indicator values deemed totally unsustainable and 1 corresponds to absolutely sustainable ones. The objective is to aggregate these indicators into an overall measure of sustainability $S(x_1, \ldots, x_n)$ whose range is [0, 1] as well.

It is reasonable to require that S satisfy the following postulates:

(P.1) S is a continuous and increasing function.

(P.2a) S is *internal*, i.e., $\min x_i \leq S(x_1, \ldots, x_n) \leq \max x_i$, $\forall (x_1, \ldots, x_n) \in [0, 1]^n$.

Continuity is assumed for analytical simplicity. Monotonicity ensures that S does not decrease whenever an indicator is improved. "Internness" is a measuring convention that assigns to each combination of $(x_1, \ldots, x_n)$ an intermediate aggregate value. It has been shown (Beliakov et al., 2007) that (P.1) and (P.2a) together are equivalent to (P.1) and (P.2) below

(P.2) S is *idempotent*, i.e., $S(x, \ldots, x) = x$, $\forall x \in [0, 1]$.

Several functions S satisfy (P.1) and (P.2). The arithmetic mean

$$\mathrm{AM} = (x_1+\ldots+x_n)/n,$$

the geometric mean,

$$\mathrm{GM} = (x_1\ldots x_n)^{1/n},$$

and the harmonic mean,

$$\mathrm{HM} = n/(x_1^{-1}+\ldots+x_n^{-1}),$$

are examples of such aggregation functions, with AM and GM commonly used in sustainability assessments.

Idempotence implies that

$$S(1, \ldots, 1) = 1 \quad (1)$$

$$S(0, \ldots, 0) = 0 \quad (2)$$

two fundamental properties of any sustainability function. These two conditions along with monotonicity reminisce of the joint distribution function of *n random variables* defined on the unit interval [0, 1]. However, distributions in general do not satisfy (P.2).

In 1930 Kolmogorov defined *regular means* as functions which satisfy (P.1), (P.2), and two additional conditions:

(P.3) S is a symmetric function, i.e., . $S(\ldots, x_i, \ldots, x_j, \ldots) = S(\ldots, x_j, \ldots, x_i, \ldots)$ $\forall$ i, j.

(P.4) S is *decomposable* in the sense that if $S(x_1, \ldots, x_n) = y$ then $S(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+m}) = S(y, \ldots, y, x_{n+1}, \ldots, x_{n+m})$, for all natural numbers n and m.

According to (P.4), any subset of values $x_1, \ldots, x_n$ can be replaced by their aggregate value without affecting the overall aggregate value. Symmetry is necessary for the validity of the next result which will be later relaxed.

Theorem 1 (Kolmogorov, 1930). If (P.1)–(P.4) hold then S is a *quasi-arithmetic mean* of $x_1, \ldots, x_n$, i.e., it has the form

$$S(x_1, \ldots, x_n) = \varphi^{-1}\left(\frac{\varphi(x_1)+\ldots+\varphi(x_n)}{n}\right)$$

where φ and its inverse, φ^{-1}, are continuous and strictly increasing univariate functions having the same range.

A similar theorem was proved by Nagumo (1930), who also examined regular means that are homogeneous of degree 1, i.e.,

(P.5) $S(tx_1, \dots, tx_n) = t\, S(x_1, \dots, x_n)\ \forall\ x_i \in [0, 1]$ and t: $tx_i \in [0, 1]$.

Theorem 2 (Nagumo, 1930). If (P.1)–(P.5) hold, then S is a quasi-arithmetic mean generated by $\varphi(x) = x^p$ for $p \in (-\infty, \infty)$ or, equivalently, the *power mean* (also known as the Hölder mean)

$$S(x_1, \dots, x_n) = \mathrm{PM}_p = \left(\frac{x_1{}^p + \dots + x_n{}^p}{n}\right)^{1/p}.$$

Note that PM_1 = AM, PM_{-1} = HM, $\lim_{p\to-\infty} \mathrm{PM}_p = \min x_i$, $\lim_{p\to\infty} \mathrm{PM}_p = \max x_i$, and, by L' Hôpital's rule on $\ln(\mathrm{PM}_p)$, PM_0 = GM.

2.2. *The shifted geometric mean*

In this section we extend the power and geometric means to relax the symmetry and homogeneity assumptions. The expression we adopt in equation (5) below is a novel analytical model for the assessment of sustainability. It satisfies postulates P.1-P.4 and, by consequence, properties (1) and (2), and generalizes the arithmetic and geometric mean approaches providing them with a solid theoretical rationale.

If indicator 1 is deemed twice as important as the others, then this is equivalent to having *two indicators with the same value* x_1. The quasi-arithmetic mean is

$$S(x_1, \dots, x_n) = \varphi^{-1}\left(\frac{2\varphi(x_1) + \dots + \varphi(x_n)}{n+1}\right).$$

This approach can be extended to any combination of integer weights $m_i \geq 1$; thus

$$S(x_1, \dots, x_n) = \varphi^{-1}\left(\frac{m_1\varphi(x_1) + \dots + m_n\varphi(x_n)}{m_1 + \dots + m_n}\right).$$

If we define the rational weights $w_i = m_i/(m_1 + \dots + m_n)$, the above is written as

$$S(x_1, \dots, x_n) = \varphi^{-1}[w_1\varphi(x_1) + \dots + w_n\varphi(x_n)]. \quad (3)$$

The same expression holds true for irrational weights as well because, by Theorem 1, φ^{-1} is continuous. Eq. (3) defines a *weighted* quasi-arithmetic mean and relaxes the symmetry assumption.

Idempotence is an intuitively rooted convention for sustainability assessments. It can be derived from the homogeneity requirement and the boundary condition $S(1, \dots, 1) = 1$. Indeed, under those two conditions $S(x, \dots, x) = S(1x, \dots, 1x) = x\, S(1, \dots, 1) = x$. Therefore, homogeneity is justified as well, albeit only for equal x_i. Also, by Theorem 2, homogeneity gives rise to explicit generating functions, $\varphi(x) = x^p$ and power means.

To relax the homogeneity requirement we introduce a *shifted* version of the weighted power mean, generated by $\varphi(x) = (C + x)^p$ for $p \neq 0$:

$$S(x_1, \dots, x_n) = \mathrm{PM}_{C,p} = [w_1(C+x_1)^p + \dots + w_n(C+x_n)^p]^{1/p} - C \quad (4)$$

where the parameter C is nonnegative so that $C + x \geq 0$ for all $x \in [0, 1]$, and $w_1 + \dots + w_n = 1$ with $w_i > 0$. A single-parameter family of aggregation functions is the *shifted geometric mean*

$$S(x_1, \dots, x_n) = \mathrm{GM}_C = (C + x_1)^{w_1} \dots (C + x_n)^{w_n} - C \quad (5)$$

which is a special case of (4) for $p \to 0$. Despite their simple form, shifted geometric means have important properties, in addition to (P.1)–(P.4), which provide a continuum of choices between the weighted geometric and arithmetic means defined by

$$\mathrm{GM} = x_1{}^{w_1} \dots x_n{}^{w_n} \text{ and } \mathrm{AM} = w_1x_1 + \dots + w_nx_n.$$

We discuss these properties in the next section.

2.3. *Further properties of means*

Let $C \geq 0$, $w_i > 0$, and $w_1 + \dots + w_n = 1$. Then the following are true:

(a) GM_C is increasing in both, C and $x_i\ \forall i$

(b) GM_0 = GM

(c) $\lim_{C\to\infty} \mathrm{GM}_C$ = AM.

To prove (a) we take the derivative of GM_C,

$$\frac{d\mathrm{GM}_C}{dC} = \sum_{i=1}^{n}\left(w_i(C + x_i)^{w_1 - 1}\prod_{k\neq i}(C + x_k)^{w_k}\right) - 1 = \left(\sum_{i=1}^{n}\frac{w_i}{C + x_i}\right)\left(\prod_{k=1}^{n}(C + x_k)^{w_k}\right) - 1;$$

the property follows from the inequality of weighted arithmetic and geometric means (Hardy et al. 1934, Theorem 9),

$$\sum_{i=1}^{n} w_i y_i \geq \prod_{i=1}^{n} y_i^{w_i}$$

with $y_i = 1/(C+x_i)$. The monotonicity in x_i can be verified by taking the partial derivative of GM_C. (b) is obvious. Finally

$$\begin{aligned}
\lim_{C\to\infty}\left[\prod_{i=1}^{n}(C+x_i)^{w_i} - C\right] &= \lim_{C\to\infty} C\left[\prod_{i=1}^{n}\left(1+\frac{x_i}{C}\right)^{w_i} - 1\right] \\
&\underset{B \triangleq 1/C}{=} \lim_{B\to 0}\frac{\prod_{i=1}^{n}(1+Bx_i)^{w_i} - 1}{B} \\
&= \lim_{B\to 0}\frac{\sum_{i=1}^{n}\left[w_i(1+Bx_i)^{w_i-1} x_i \prod_{k\neq i}(1+Bx_k)^{w_k}\right]}{1} \quad \text{(by L' Hôpital's rule)} \\
&= \sum_{i=1}^{n} w_i x_i
\end{aligned}$$

which is (c).

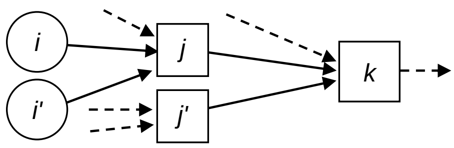

Figure 1. Hierarchical assessment by successive aggregations in stages *j*, *j'* and *k*.

Next, consider a multi-stage assessment procedure which successively combines indicators into more composite variables. As shown in Figure 1, the indicators *i*, *i'*, … on the left are aggregated into variable *j*, which is next combined with other variables *j'* of the same hierarchical level to form a yet more composite variable *k*, and so on. The input variables of each stage *j* or *k* are assigned *relative weights* and the output variables, x_j and x_k, are given by

$x_j = S(x_i, x_{i'}, \ldots)$ with weights w_{ij} such that $\sum_i w_{ij} = 1$

$x_k = S(x_j, x_{j'}, \ldots)$ with weights w_{jk} such that $\sum_j w_{jk} = 1$.

Using the shifted geometric mean with the same C in both stages gives

$$x_j = \prod\nolimits_{i\in V_j}(C+x_i)^{w_{ij}} - C$$

$$x_k = \prod\nolimits_{j\in V_k}(C+x_j)^{w_{jk}} - C$$

where V_j and V_k are the subsets of variables which are combined to generate variables *j* and *k* respectively.

Substituting the first equation above into the second we get

$$\begin{aligned}
x_k &= \prod\nolimits_{j\in V_k}\left[C + \prod\nolimits_{i\in V_j}(C+x_i)^{w_{ij}} - C\right]^{w_{jk}} - C \\
&= \prod\nolimits_{j\in V_k}\left[\prod\nolimits_{i\in V_j}(C+x_i)^{w_{ij}}\right]^{w_{jk}} - C \\
&= \prod\nolimits_{j\in V_k}\prod\nolimits_{i\in V_j}(C+x_i)^{w_{ij}w_{jk}} - C
\end{aligned}$$

The last expression is identical to the shifted geometric mean of a *single aggregation stage k* (without the intermediate variables *j*, *j'*, …) whose inputs are the same as the primary inputs *i*, *i'*, … of the multistage system of Figure 1 with weights $w_i \triangleq w_{ij}w_{jk}$ such that

$$\sum_j \sum_i w_{ij}w_{jk} = 1 = \sum_i w_i \qquad (6)$$

Conversely, any shifted geometric mean can in many different ways be represented as a hierarchical (nested) composition of simpler GM_C functions all with the same parameter C. We refer to this property as

(d) All GM_C functions with $n > 2$ inputs (and $C \geq 0$) are *hierarchically decomposable*.

We close this section by stating a number of other elementary properties of shifted means.

(e) GM_C is concave in C and concave in $(x_1, \ldots, x_n)$. These properties are consequences of the inequality $d^2GM_C/dC^2 \leq 0$ and the negative definiteness of the Hessian, $H = [\partial^2 GM_C/\partial x_i \partial x_j]$, which in turn can be derived by applying the Cauchy inequality, $(\sum a_i b_i)^2 \leq \sum a_i^2 \sum b_i^2$ with $a_i = w_i^{0.5}$ so that $\sum a_i^2 = 1$ and suitable b_i. Concavity in $(x_1, \ldots, x_n)$ implies that GM_C exhibits diminishing marginal returns to each i, that is, if all indicators but i are fixed, then any increase in x_i leads to a smaller increase in the overall sustainability as x_i approaches 1 (see Figure 2).

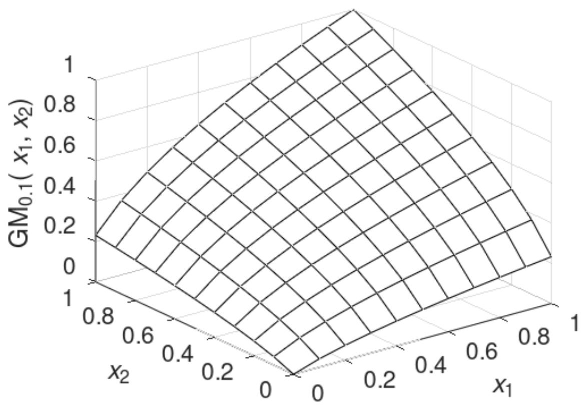

Figure 2. Function $(0.1 + x_1)^{0.5}(0.1 + x_2)_{0.5} - 0.1$ exhibiting diminishing marginal returns.

(f) Shifted power means, (2), are also hierarchically decomposable and increasing in C and $(x_1, \ldots, x_n)$. It can be shown that these functions are convex in C and in $(x_1, \ldots, x_n)$ for $p > 1$ and concave for $p < 1$, so that they can model either increasing or decreasing marginal returns.

(e) Finally we consider the case where a subset I of indicators are *critically important* in the sense that $S = 0$ whenever $x_i = 0$ for *at least one* $i \in I$. The nine planetary boundaries identified by Rockström et al. (2009) could be viewed as such critical points of sustainability. These boundaries are defined by a number of indicator threshold values which, if exceeded for too long, will have harmful or even catastrophic consequences for the planet. For example, humanity might not be able to avert collapse if atmospheric CO_2 concentrations become too high despite improvements of, say, GDP. To take into account both critical and noncritical indicators we propose the following extension to GM_C:

$$GM_{B,C} = \left(\prod_{i \in I} x_i^{w_i}\right)^{\omega} \left(\prod_{j \notin I} (C + Bx_j)^{w_j} + 1 - C - B\right)^{1-\omega}, \tag{7}$$

where i and j index critical and noncritical indicators, respectively, with relative weights satisfying

$$\sum_{i \in I} w_i = 1 \text{ and } \sum_{j \notin I} w_j = 1,$$

and C, B, and ω are model parameters satisfying $C > 0$ and $\omega, B \in [0, 1]$. These three parameters characterize sustainable operating spaces where $x_i > 0$ $\forall i \in I$. For example, suppose that all critical indicators are $x_i > 0$ and all noncritical indicators are $x_j = 0$. From (7) we get

$$GM_{B,C} = \left(\prod_{i \in I} x_i^{w_i}\right)^{\omega} (C + 1 - C - B)^{1-\omega} = \left(\prod_{i \in I} x_i^{w_i}\right)^{\omega} (1 - B)^{1-\omega}.$$

For $B = 1$, the expression given above is zero as well, which implies that the set of noncritical indicators is critical if considered *as a whole*. Yet choosing $B < 1$ renders this set noncritical as a whole since $GM_{B,C} > 0$ even when $x_j = 0$ for all $j \notin I$.

3. NUMERICAL RESULTS

In this section we report on the application of the shifted geometric mean using the same indicator data and weights as those of the most recent versions of several established models: HDI, EPI, a TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) model fitted to EPI data, and SAFE. These models use different aggregation approaches:

- HDI is the geometric mean of three variables, one of which is the average of two specialized indicators.
- EPI is a hierarchical additive model that generates a weighted arithmetic mean.

- TOPSIS is a multi-criteria decision-making model.
- SAFE is a hierarchical fuzzy inference system.

Against all the above we tested the model

$$\mathrm{GM}_1 = (1+x_1)^{w_1} \dots (1+x_n)^{w_n} - 1 .$$

From numerical experiments not reported here we have observed that this function gives similar assessments as the functions GM_C and $\mathrm{PM}_{C,p}$ with weights w_i and parameters C and p chosen so as to maximize the overall sustainability correlation coefficient or minimize the mean absolute ranking deviations from the original models.

Let $q = 1, \dots, Q$ be entities whose sustainability is assessed and $t = 1, 2$ the estimation method (e.g., 1 for the shifted geometric mean and 2 for any other model). S_{qt} denotes the overall sustainability estimate of entity q by method t. We use Pearson's r correlation coefficient to measure the linear relationship between the values S_{q1} and S_{q2} for all q. We also use Kendall's τ rank correlation coefficient to compare the rankings generated by the two methods from most to least sustainable countries. Both coefficients range from −1 (complete negative association) to +1 (complete positive association). The relevant formulas are

$$r = \frac{\sum_{q=1}^{Q}(S_{q1}-\bar{S}_1)(X_{q2}-\bar{S}_2)}{\sqrt{\sum_{q=1}^{Q}(S_{q1}-\bar{S}_1)^2}\sqrt{\sum_{q=1}^{Q}(S_{q2}-\bar{S}_2)^2}}\,,\quad \bar{S}_t = \frac{1}{Q}\sum_{q=1}^{Q} S_{qt}\,,\ t = 1, 2.$$

Rank correlation is based on the numbers of entity pairs having concordant, antithetical, and tied rankings in the two models. Consider all distinct entity pairs, i.e., $\{q, p\}$ such that $q = 1, \dots, Q - 1$ and $p = q + 1, \dots, Q$. There are $N = Q(Q - 1)/2$ such pairs; thus $\{q, p\} = \{1, 2\}, \dots, \{1, Q\}, \{2, 3), \dots \{Q - 1, Q\}$. We define the following auxiliary quantities:

N_c = number of concordant entity pairs $\{q, p\}$, i.e., $(S_{q1} - S_{p1})(S_{q2} - S_{p2}) > 0$

N_d = number of discordant entity pairs $\{q, p\}$, i.e., $(S_{q1} - S_{p1})(S_{q2} - S_{p2}) < 0$

N_t = number of ties $\{q, p\}$ in model t, i.e., $S_{qt} = S_{pt}$ (regardless of the other model), where t=1, 2.

The Kendall correlation is defined by

$$\tau = \frac{N_c - N_d}{\sqrt{(N-N_1)(N-N_2)}}$$

with maximum value 1 assumed when $N_d = 0$ (in which case we have that $N = N_c + N_1$ and $N_2 = N_1$).

HDI uses four indicators:

LE = life expectancy at birth

GNI = gross national income per capita in 2017 PPP $

ES = expected years of schooling if prevailing patterns of age-specific enrolment rates persist

MS = current mean years of schooling of adults aged above 25

Auxiliary indices are computed by interpolation between thresholds of values deemed unsustainable and sustainable

$$y_1 = (\mathrm{LE}–20)/(85–20),\ y_2 = [\ln(\mathrm{GNI})–\ln(100)]/[\ln(75000)–\ln(100)],\ y_3 = \mathrm{ES}/18,\ y_4 = \mathrm{MS}/15,$$

which are then restricted on [0, 1] as follows:

$$x_i = \min[1, \max(0, y_i)],\ i = 1, \dots, 4.$$

The overall index is given by

$$\mathrm{HDI} = x_1^{1/3} x_2^{1/3} \left(\frac{x_3 + x_4}{2}\right)^{1/3}.$$

We apply GM_1 with four inputs and weights $w_1 = w_2 = 1/3$ and $w_3 = w_4 = 1/6$. Table 1 shows the ten highest and ten lowest ranking countries according to HDI.

Table 1. Country rankings and indices according to HDI and GM1 (2019 data).

Country	HDI rank	GM_1 rank	HDI index	GM_1 index
Norway	1	1	0.9570	0.9567
Switzerland	2	2	0.9555	0.9559
Ireland	3	3	0.9553	0.9549
Hong Kong, China (SAR)	4	4	0.9490	0.9493
Iceland	5	5	0.9489	0.9485
Germany	6	6	0.9469	0.9469
Sweden	7	7	0.9452	0.9447
Australia	8	8	0.9441	0.9437
Netherlands	9	9	0.9439	0.9434
Denmark	10	10	0.9399	0.9395
Eritrea	180	177	0.4586	0.4837
Mozambique	181	181	0.4559	0.4610
Burkina Faso	182	182	0.4521	0.4595
Sierra Leone	183	183	0.4517	0.4508
Mali	184	184	0.4335	0.4456
Burundi	185	186	0.4334	0.4427
South Sudan	186	185	0.4326	0.4430
Chad	187	188	0.3978	0.4044
Central African Republic	188	189	0.3972	0.4013
Niger	189	187	0.3937	0.4142

The correlation coefficient of HDI and GM_1 is $r > 0.999$ and the Kendall rank correlation is $\tau > 0.99$. The two indices give mostly identical rankings; this can be justified by the fact that they are both geometric means since HDI is GM_0 in x_1, x_2 and $(x_3+x_4)/2$.

EPI ranks 180 countries using 40 variables which assess the national efforts to protect environmental health, enhance ecosystem vitality, and mitigate climate change. These variables are transformed to points on the 0–100 scale and all the missing data are imputed. Let A be the set of all variables available for a given country, with values z_i on [0, 100] and weights w_i. When applying the shifted geometric mean we normalize the original data by dividing by 100 and also *disregard the missing variables* (5.2% of the total data) *by renormalizing* w_i. The resulting approximation is given by EPI ≈ 100 GM_1 where

$$\mathrm{GM}_1 = \sum_{i\in A}\left(1+\frac{z_i}{100}\right)^{v_i} - 1\,, \quad v_i = \frac{w_i}{\sum_{k\in A} w_k}\,. \tag{8}$$

The rankings of EPI with imputed data and $100GM_1$ with missing data are quite similar, as shown in Table 2. Despite the fact that the EPI score is higher than $100GM_1$ for most countries, very strong positive correlations are observed between sustainability indices ($r > 0.99$) and rankings ($\tau = 0.93$).

Table 2. Country rankings and indices according to EPI and GM1 (2019 data).

Country	EPI rank	GM1 rank	EPI index	100GM1 index
Denmark	1	1	77.9	74.78
United Kingdom	2	2	77.7	74.65
Finland	3	3	76.5	74.24
Malta	4	4	75.2	72.99
Sweden	5	6	72.7	70.04
Luxembourg	6	5	72.3	70.40
Slovenia	7	7	67.3	64.67
Austria	8	9	66.5	62.67
Switzerland	9	8	65.9	64.65
Iceland	10	10	62.8	60.25
Sudan	171	163	27.6	25.72
Turkey	172	171	26.3	24.13
Haiti	173	169	26.1	24.52
Liberia	174	174	24.9	23.07
Papua New Guinea	175	175	24.8	22.91
Pakistan	176	176	24.6	22.33
Bangladesh	177	177	23.1	21.50
Viet Nam	178	178	20.1	18.73
Myanmar	179	179	19.4	18.06
India	180	180	18.9	17.15

HDI and EPI are strongly correlated with GM_1. To a great extent, this is due to two facts: (i) the main functional evaluations in HDI and EPI involve GM_C, with C equal to 0 and ∞, respectively; and (ii) when $x_i \in [0, 1]$ GM_C appears to be robust to variations in C. To further explore the approximation capability of GM_1 we examine models involving several computational steps rather than a single functional evaluation.

TOPSIS (Hwang and Yoon, 1981) calculates the overall assessment for a given country using distances of the country indicators from the so-called negative and positive ideal points, comprising the worst and the best indicator values, respectively, over all countries. We apply TOPSIS and a variant thereof of the EPI data and compare the results with GM_1.

The environmental performance estimates obtained by the standard TOPSIS method have a strong correlation with the original EPI indices with coefficients $r = 0.87$ and $\tau = 0.67$ and a slightly stronger correlation with GM_1 with $r = 0.88$ and $\tau = 0.69$.

SAFE uses 69 inputs, each mapped into one or more fuzzy sets, and employs fuzzy inference engines and rule bases to aggregate inputs into composite indicators and, finally, the overall sustainability index. HDI and SAFE have a strong correlation ($r = 0.80$ and $\tau = 0.62$) and common indicators LE, GNI and ES, but SAFE uses many more inputs to allow for a global assessment and more sophisticated methods to impute missing data and assign weights to indicators and intermediate components of sustainability.

We applied GM_1 to the SAFE data with and without imputation. To compute overall indicator weights, we repeatedly used (6) and the relative weights shown in Figure 1 of Grigoroudis et al. (2021). The results are shown in Table 3.

Despite their different features the two models have very strongly correlated index values and rankings ($r = 0.98$, $\tau = 0.90$). We also tested GM_1 with missing data (about 7% of total) and obtained similar results. The rather significant deviations in the rankings of the two models for the Netherlands and the Democratic Republic of Congo is attributed to model differences and the number of fuzzy sets used in SAFE. For 9 fuzzy sets SAFE ranks the Netherlands 8th among countries. However, a finer reasoning scheme with 21 sets placed the same country 17th. A numerical investigation showed that the rankings of most countries end up quite close to those of GM_1 when a larger number of fuzzy sets are used in SAFE. One could thus fine-tune SAFE by just enhancing the number of its fuzzy sets.

Table 3. Country rankings and indices according to SAFE and GM_1 (2016 data; imputed).

Country	SAFE rank	GM_1 rank	SAFE index	GM_1 index
Denmark	1	3	0.8691	0.8574
Sweden	2	2	0.8618	0.8592
Norway	3	1	0.8578	0.8643
Switzerland	4	5	0.8387	0.8375
Austria	5	4	0.8284	0.8487
Finland	6	6	0.8189	0.8312
Slovenia	7	7	0.8067	0.8248
Netherlands	8	22	0.8044	0.7746
Slovakia	9	13	0.8043	0.8066
United Kingdom	10	10	0.8041	0.8161
Dem. Rep. of the Congo	155	148	0.4039	0.5118
Iraq	156	158	0.3957	0.4740
Guinea-Bissau	157	155	0.3934	0.4858
Yemen	158	162	0.3862	0.4620
Central African Rep.	159	157	0.3830	0.4763
Eritrea	160	163	0.3747	0.4522
Sudan	161	161	0.3664	0.4622
Mauritania	162	160	0.3643	0.4637
Haiti	163	164	0.3640	0.4355
Afghanistan	164	159	0.3621	0.4683

4. LIMITATIONS

The benchmark assessments of the previous section indicate that the shifted geometric function gives remarkably similar estimates as those of other common sustainability assessments. This function is based on a set of intuitively appealing postulates and unifies existing additive and multiplicative sustainability approaches.

A shortcoming of the shifted geometric function is that it can only assign constant weights w_i, independent of the indicator values x_i. Value-dependent weights are necessary to model more elaborate sustainability concepts. For example, a group of important indicators may exhibit a compensatory behavior, that is, a "bad" value of one indicator is compensated by the "good" value of another, thus cancelling individual effects on sustainability. SAFE on the other hand avoids such an effect by its involved reasoning layout. In a sense the two models are complementary. The simplicity of the shifted geometric model comes with the already mentioned compensatory price and SAFE with its computationally involved structure.

5. CONCLUSIONS

Sustainability assessment models, by their nature are subjective. In this paper we departed from traditional ad hoc approaches by proposing a mathematical model based on a number of natural postulates, adopting a bottom up approach. The resulting model is novel, has interesting properties, and correlates very well with other established approaches. Our intention is to provide a solid mathematical basis for the measurement of sustainability and derive insights which are not available in existing models.

Besides the model's interesting properties, a major advantage is its simplicity. It would be important to derive conditions for a uniqueness result but presently this seems to be elusive. More work is needed to apply our approach to systems other than nations, such as energy, transportation, urban systems, etc. Also, sensitivity analyses are needed to provide guidance regarding the most important indicators that would improve sustainability as fast as possible.

One could contend that the postulates we used are subjective and the resulting model is also subjective. However, every mathematical theory is founded on axioms or postulates and its information content is basically that of its axioms. We hope our approach will start new discussions about sustainability assessment and provide a more rigorous framework for this important problem.

Author Contributions: Conceptualization, Y.A.P; methodology and investigation, Y.A.P., V.S.K, E.G, and F.D.K.; validation and data curation, V.S.K.; writing Y.A.P. and V.S.K. All authors have read and agreed to the published version of the manuscript.
Funding: This research was carried out under the support of the Technical University of Crete.
Institutional Review Board Statement: Not applicable.
Data Availability Statement: Research data and a spreadsheet for reproducing the calculations of this article are available at https://doi.org/10.17632/85h8njxvz6.1
Institutional ethical review information: Not applicable; this research did not involve human participants or animal experimentation.

ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their constructive comments.

REFERENCES

Beliakov, G., Pradera, A., and Calvo, T. (2007). *Aggregation Functions: A Guide for Practitioners.* Berlin, Springer-Verlag.

Esty, D. C., Levy, M., Srebotnjak, T., and de Sherbinin, A. (2005). *Environmental Sustainability Index: Benchmarking National Environmental Stewardship*. New Haven, CT, Yale Center for Environmental Law and Policy. http://sedac.ciesin.columbia.edu/es/esi/ESI2005.pdf (accessed 22 July 2022).

Gan, X., Fernandez, I. C., Guo, J., Wilson, M., Zhao, Y., Zhou, B., and Wu, J. (2017). When to use what: Methods for weighting and aggregating sustainability indicators. *Ecological Indicators*, Vol. 81, pp. 491-502.

Greco, S., Ishizaka, A., Tasiou, M., and Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*, Vol. 141, No. 1, pp. 61-94.

Grigoroudis, E., Kouikoglou, V. S., and Phillis, Y. A. (2021). SAFE 2019: Updates and new sustainability findings worldwide. *Ecological Indicators,* Vol. 121, 107072.

Hardy, G. H., Littlewood, J. E., and Polya, G. (1934). *Inequalities*. London, Cambridge University Press.

Hwang, C. L., and Yoon, K. (1981). Multiple Attribute Decision Making. Berlin, Springer.

van de Kerk, G., and Manuel, A. R. (2008). A comprehensive index for a sustainable society: The SSI—the Sustainable Society Index. Ecological Economics, Vol. 66, No. 2-3, pp. 228-242.

Kolmogorov, A. (1930). Sur la notion de la moyenne. Atti dell'Accademia Nazionale dei Lincei, Vol. 12, No.9, pp. 388-391. Also in: Kolmogorov, A. (1991). Selected Works I: Mathematics and Mechanics. Dordrecht, Kluwer Academic Publishers, pp. 144-146.

Liu, K. F. (2007). Evaluating environmental sustainability: an integration of multiple-criteria decision-making and fuzzy logic. *Environmental Management*, Vol. 39, No. 5, pp. 721-736.

Nagumo, M. (1930). Über eine klasse der mittelwerte. *Japanese Journal of Mathematics: Transactions and Abstracts*, Vol. 7, pp. 71-79.

OECD (1991). *Environmental Indicators: A Preliminary Set*. Paris, OECD.

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin III, F. S., Lambin, E., et al. (2009). Planetary boundaries: exploring the safe operating space for humanity. *Ecology and Society*, Vol. 14, No. 2.

Saaty, T.L., 1980. *The Analytic Hierarchy Process*. New York, McGraw Hill.

United Nations Development Programme (2022). *Human Development Index (HDI)*. https://hdr.undp.org/data-center/human-development-index#/indicies/HDI (accessed 2 July 2022).

Wolf, M. J., Emerson, J. W., Esty, D. C., de Sherbinin, A., Wendling, Z. A., et al. (2022). 2022 *Environmental Performance Index*. New Haven, CT, Yale Center for Environmental Law and Policy. https://epi.yale.edu (accessed 11 July 2022).

Yang, L., (2014). *An Inventory of Composite Measures of Human Progress*. Technical report, United Nations Development Programme Human Development Report Office.