Ecological Economics 130 (2016) 195–208

Contents lists available at ScienceDirect

Ecological Economics

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

Analysis

Normalization in sustainability assessment: Methods and implications

N.L. Pollesch[a,b,*], V.H. Dale[b]

[a]Department of Mathematics, The University of Tennessee, 1403 Circle Drive, Knoxville, TN 37996-1320, United States
[b]Center for BioEnergy Sustainability, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6036, United States

ARTICLE INFO

Article history:
Received 29 November 2015
Received in revised form 24 June 2016
Accepted 25 June 2016
Available online 8 August 2016

Keywords:
Aggregation
Bioenergy sustainability
Composite indicator
Multi-criteria assessment
Normalization
Sustainability assessment
Standardization

ABSTRACT

One approach to assessing progress towards sustainability makes use of multiple indicators spanning the environmental, social, and economic dimensions of the system being studied. Diverse indicators have different units of measurement, and normalization is the procedure employed to transform differing indicator measures onto similar scales or to unit-free measures. Given the inherent complexity entailed in interpreting information related to multiple indicators, normalization and aggregation of sustainability indicators are common steps after indicator measures are quantified. However, it is often difficult for stakeholders to make clear connections between specific indicator measurements and resulting aggregate scores of sustainability. Motivated by challenges and examples in sustainability assessment, this paper explores various normalization schemes including ratio normalization, target normalization, Z-score normalization, and unit equivalence normalization. Methods for analyzing the impacts of normalization choice on aggregate scores are presented. Techniques are derived for general application in studying composite indicators, and advantages and drawbacks associated with different normalization schemes are discussed within the context of sustainability assessment. Theoretical results are clarified through a case study using data from indicators of progress towards bioenergy sustainability.

1. Introduction

Sustainability is an inherently complex topic with different meanings pertaining to different contexts. Indicators of progress toward sustainability are measures that characterize conditions under which resource uses are more sustainable and are often tracked over time or compared for alternative practices. A variety of indicators of progress toward sustainability have been identified and both the breadth as well the number of indicators used for each assessment varies by application (Dale et al., 2013; Mori and Christodoulou, 2012; McBride et al., 2011; Singh et al., 2009; Mayer, 2008). Indicators typically involve social, economic and environmental measures in order to capture the three major aspects of sustainability.

Sustainability assessments often rely on a variety of indicators. Different indicators are measured and reported in units pertinent to the particular metric. Having a common unit of measure is useful for comparison and synthesis of indicators. The synthesis of indicators can be done analytically, statistically, or graphically. Combining of measurements of multiple indicators to produce sustainability scores, composite indices, or aggregates is done to reduce dimensionality and can provide a single holistic value. Industry reports and national inventories are typically based on these highly aggregated data (Heijungs et al., 2007; Bare et al., 2006).

Normalization is the process of transforming units of measurement from the original units to common measurement units or to measurements that are unit less. This process is also referred to as unit scaling or standardization, with terminology varying based on the functions utilized in the process and by discipline. For clarity, this paper uses the term *normalization* to refer to all such processes transforming diverse units to common or unit-less quantities. When indicator units vary, normalization is seen as a necessary step prior to aggregation (Nardo et al., 2005). Mathematical research into the structure of sustainability assessments has focused on the aggregation step (Pollesch and Dale, 2015; Langhans et al., 2014; Roberts, 2014; Zhou et al., 2006; Ebert and Welsch, 2004). Freudenberg (2003) gives a comparison of two different normalization procedures on a composite assessment outcome. Although researchers are often aware of the effect that a given choice of normalization scheme has on assessment outcome, no formal analysis of the implications of the normalization procedure on assessment outcome has emerged in the sustainability assessment literature.

In this paper the consequences of using different normalization functions within an aggregate score of sustainability are

* Corresponding author.
E-mail addresses: pollesch@math.utk.edu (N. Pollesch), dalevh@ornl.gov (V. Dale).

http://dx.doi.org/10.1016/j.ecolecon.2016.06.018
0921-8009/

explored. The normalization and subsequent aggregation process used to derive composite sustainability scores vary greatly (Mori and Christodoulou, 2012; Singh et al., 2009; Mayer, 2008). Moldan et al. (2012) observe that for sustainability assessment “the selection and definition of pertinent indicators, to a large extent, defines the whole issue.” Indicator selection may define the “whole issue,” yet how measurements of those indicators are interpreted depends critically on the assessment structure. Even a well-defined set of indicators and accompanying high quality data can lead to completely different assessments of a system depending on the normalization and aggregation procedure employed. A recent example illustrating how something as arbitrary as the order in which indicators appear on a spider diagram can have large effect on sustainability ratings as calculated through surface area is given in Dias and Domingues (2014).

After terminology is established, this paper provides examples of normalization methods in sustainability assessment to show variety and form. Next, we move to an in-depth look at four common normalization procedures in combination with different aggregation functions to elucidate exact dependencies between aggregate sustainability scores and the normalization schemes utilized in their calculation. Through a case study of composite scores of bioenergy sustainability, the question of how changes in non-normalized indicator measures impact aggregate scores of sustainability is addressed. The paper concludes with the discussion of advantages and drawbacks for normalization schemes, reinforced by results derived in the case study provided.

2. Normalization Methods

The normalization process is used throughout scientific research and is motivated by a variety of circumstances. In sustainability assessment, the major motivation for normalization is to transform measurements of indicators, typically obtained in different units, to a common unit of measurement to compare them or to prepare them for inclusion in an aggregate score of sustainability.

2.1. Terminology and Notation

There are a wide variety of functions that can be applied to data in order to normalize. To aid in the discussion and analysis of these different normalization procedures and functions, the underlying terminology is established below.

- *Indicator bearing:* Sustainability indicators can differ as to whether smaller or larger values of the indicators are interpreted as being ideal[1] or if there is some ideal value from which the measure should not differ in magnitude too much. In this paper we use the term *indicator bearing* to describe this attribute of indicators. Indicator bearings are referenced through a variety of terminology in the literature, for example “direct correlation with utility” and “inverse correlation with utility” (Maxim, 2014). Krajnc and Glavič (2005) use “positive impact” or “negative impact,” and Dias and Domingues (2014) use “criteria is to maximize” or “criteria is to minimize.” In this paper, the terms *larger-the-better* (LTB), *smaller-the-better* (STB), and *distance-to-ideal* (DTI) are used given their straightforward meaning. It should be noted that these are not the only types of indicator bearings for normalization and that some normalization strategies, such as *unit equivalence normalization* (Table 1), do not discriminate indicators in these regards.
- *Normalization schemes:* Since each assessment may include indicators that are of many bearing types, families of normalization functions are used to take these differences into account. For example, if *target normalization* (Table 1) is employed, the form of the function applied differs by indicator bearing types. Families, or groups, of normalization functions are referred to here as *normalization schemes*. In the case where the normalization procedure does not discriminate among indicator bearing types, the scheme may consist of just a single function; Z-score normalization (Table 1) is an example.
- *Internal normalization:* Another differentiating factor between normalization functions occurs if they use the entire data set for a given indicator to normalize a single measurement of that indicator. These normalization functions are referred to as *internal*[2]. Examples of this type of normalization function are ratio normalization functions for STB and LTB type indicators (Table 1). Normalization functions that are not internal depend on predefined, exogenous values, such as target and baseline levels or unit conversion factors.
- Notation for normalized and non-normalized indicator measurements also varies. In this paper, non-normalized measures are denoted by a superscript ‘*’. Subscripts are used to convey a variety of information, such as which indicator is being considered and/or which measurement of that indicator is being referenced. For example, if measurement j of indicator i is normalized, the notation for the normalized measure would be x_{ij} and the non-normalized measure would be x^*_{ij}. For consistency and clarity, examples referenced in this paper are translated into this notation when possible.

2.2. Examples of Normalization in Sustainability Assessment

A plethora of normalization functions are utilized in sustainability assessment. Examples given next provide a glimpse into the variety and form.

Krajnc and Glavič (2005) propose two different normalization schemes, the second of which is employed in the *Sustainable Development Index*. The first scheme normalizes measurements relative to the average of the indicator measures (in this case a total of T measures have been taken over time) so that

$$x_i = \frac{x^*_i}{\frac{1}{T}\sum_{j=1}^{T} x^*_j}$$

The second scheme is given by

$$x_i = \frac{x^*_i - \min_j\left\{x^*_j\right\}}{\max_j\left\{x^*_j\right\} - \min_j\left\{x^*_j\right\}} \quad \text{and} \quad x_i = 1 - \frac{x^*_i - \min_j\left\{x^*_j\right\}}{\max_j\left\{x^*_j\right\} - \min_j\left\{x^*_j\right\}}$$

for larger-the-better and smaller-the-better type indicators, respectively. All the normalization functions from Krajnc and Glavič (2005) given above are internal.

In the *Holistic Sustainability Assessment Tool for Bioenergy* of Hayashi et al. (2014), the indicator measures are normalized by

$$x_i = \begin{cases} \frac{(x^*_i - T_i)}{(x_{max_i} - T_i)} & x^*_i > T_i \\ \frac{(x^*_i - T_i)}{(T_i)} & x^*_i \le T_i \end{cases} \quad \text{and} \quad x_i = \begin{cases} \frac{(x^*_i - T_i)}{(x_{max_i} - T_i)} & x^*_i > T_i \\ \frac{(x^*_i - T_i)}{(T_i)} & x^*_i \le T_i \end{cases}$$

[1] The term *ideal* can be interpreted in a variety of ways. Frequently, analysis goals and assessment context guide the interpretation as to what is considered an ideal and, correspondingly, a non-ideal or baseline measurement value.

[2] The term *internal*, with respect to a normalization function, is used to identify those functions that utilize the entire data set for an indicator to normalize any given measurement value from the set. This term is not to be confused with the *internality* of an aggregation function, which describes the aggregation function's compensatory behavior (see Pollesch and Dale (2015) or Grabisch et al. (2009) for a formal definition).

Table 1
Common normalization function definitions and notations: internal normalization functions, those for which the normalized value of x_j depends on the entire data set $\boldsymbol{x}^*$, and the normalization functions that create dimensionless quantities are identified.

Scheme, notation, and definition	Indicator bearing	Internal	Dimensionless
Ratio normalization			
$R_{L,j}(\boldsymbol{x}^*) = \frac{x_j^*}{\max\{\boldsymbol{x}^*\}}$	LTB	✓	✓
$R_{S,j}(\boldsymbol{x}^*) = \frac{\min\{\boldsymbol{x}^*\}}{x_j^*}$	STB	✓	✓
$R_{D,j}\left(x_j^*, T\right) = \frac{\min\left\{x_j^*, T\right\}}{\max\left\{x_j^*, T\right\}}$	DTI		✓
Z-score normalization			
$Z_j(\boldsymbol{x}^*) = \frac{x_j^* - \bar{x}^*}{S_N(\boldsymbol{x}^*)}$ where $\bar{x}^* = \frac{1}{n}\sum_{j=1}^n x_j^*$, $S_N = \left(\frac{1}{n}\sum_{j=1}^n \left(x_j^* - \bar{x}^*\right)^2\right)^{1/2}$	n/a	✓	✓
Unit equivalence			
$C_j\left(x_j^*, c_f\right) = x_j^* c_f$ where c_f is a *conversion factor* from x_j^*'s to desired units	n/a		
Target normalization to interval $[0,1]$			
$T_{L,j}\left(x_j^*, T, B\right) = \begin{cases} 0, x_j^* & \leq B \\ 1 - \frac{T - x_j^*}{T - B}, B < x_j^* & < T \\ 1, x_j^* & \geq T \end{cases}$	LTB		✓
$T_{S,j}\left(x_j^*, T, B\right) = \begin{cases} 1, x_j^* & \leq T \\ 1 - \frac{x_j^* - T}{B - T}, T < x_j^* & < B \\ 0, x_j^* & \geq B \end{cases}$	STB		✓
$T_{D,j}\left(x_j^*, T, B_l, B_u\right) = \begin{cases} 1 - \frac{T - x_j^*}{T - B_l}, B_l < x_j^* < T \\ 1, \quad x_j^* = T \\ 1 - \frac{x_j^* - T}{B_u - T}, T < x_j^* < B_u \\ 0, \quad \text{else} \end{cases}$	DTI		✓

Note: LTB: larger-the-better, STB: smaller-the-better, DTI: distance-to-ideal, $\boldsymbol{x}^* = \{x_1^*, x_2^*, \ldots, x_n^*\}$, T is a target or ideal value for a given indicator, B is a baseline or non-ideal value for a given indicator (B_l and B_u used when an upper and lower baseline are required), $\bar{x}^*$ is the sample mean, S_N is the sample standard deviation, and C_f is a conversion factor to change units of $\boldsymbol{x}^*$ to alternate units (ex. dollars or greenhouse gas equivalents).

for larger-the-better and smaller-the-better indicators, respectively. The authors state that x_{max_i} is determined either by historical data or legislation and T_i is a threshold value. These functions score indicator measures from −1 to 1, with −1 being the least sustainable and 1 being the most sustainable. The score is 0 when the indicator is the same measure as the threshold value, T_i. This scheme can be internal or not depending on how x_{max_i} is defined.

Maxim (2014) uses the following functions for the sustainability assessment of electricity generation technology:

$$x_i = \frac{x_i^* - \min_j\{x_j^*\}}{\max_j\{x_j^*\} - \min_j\{x_j^*\}} \quad \text{and} \quad x_i = \frac{\min_j\{x_j^*\} - x_i^*}{\max_j\{x_j^*\} - \min_j\{x_j^*\}}$$

for larger-the-better and smaller-the-better indicators, respectively. Thus, normalized values fall into the interval $[0,1]$. This is another example of an internal normalization scheme.

Castoldi and Bechini (2010) use a set of normalization functions in their construction of an integrated sustainability assessment of cropping systems. Normalization is carried out by use of *continuous simple functions* such that $x_i = 1$ if x_i^* is within some range of sustainability optimality thresholds; x_i takes on values $(0,1)$ for x_i^* measures between optimal and anti-ideal thresholds and takes on the value of 0 outside of the anti-ideal thresholds. This may be seen as a generalization on the *distance to ideal* normalization function within the target normalization scheme. This normalization scheme is not internal and depends on predefined optimal thresholds for indicators.

Sadamichi et al. (2012) convert all measures to greenhouse gas equivalents in their sustainability assessment of biomass utilization for energy in east Asian countries. Transformation of indicator measures to a different and common unit of measurement for comparison can also take place by transforming to monetary units, such as dollars, or embodied energy units, such as emJoules, see Odum et al. (2000) for example, and falls broadly under the normalization scheme that is referred to in this paper as *unit-equivalence normalization*.

The normalization method utilized in Pinar et al. (2014) for the *FEEM (Fondazione Eni Enrico Mattei) Sustainability Index* is termed *benchmarking*. A benchmarking function is defined that assigns a normalized value to each indicator based on its level of sustainability, determined by "reliable and authoritative literature and international legislation sources." Specifically, Pinar et al. (2014) use the function given in Table 2. The benchmarking normalization function is not an internal normalization function, as it depends on indicator values each being mapped to some value based on a qualitative valuation of their level of sustainability.

Table 2
Benchmarking normalization function from Pinar et al. (2014). Indicators are normalized to values between 0 and 1 based on expert judgments of their sustainability level.

Normalized value	Sustainability level
0	Extremely unsustainable
0.25	Still not sustainable but not as severely as in the previous case
0.50	Discrete level of sustainability, but still far from target
0.75	Satisfactory level of sustainability, yet not on target
1	Fully sustainable

A variety of normalization procedures are employed in sustainability assessment, each of which has its own properties and unique impact on any aggregate measure of sustainability derived from the normalized measures. Although it is not within the scope of this paper to fully analyze the normalization functions provided above, the analysis and case study included in this paper sheds light on some of the behavior of common normalization procedures encountered, namely ratio normalization and target normalization to the interval [0,1]. The examples in the case study provide background for describing the ways in which normalization functions can be analyzed to understand important properties of their behavior.

2.2.1. Ratio and Target Normalization Schemes

Given that ratio normalization and target normalization to the interval [0,1] are used for analysis in this paper, a brief discussion of these two schemes is useful. Ratio normalization is named as such because measures are transformed by taking the ratio of individual measurements to extremal measurements (minimum or maximum) of the data set. When indicators are of smaller-the-better type, the minimum value from the data set is used to transform all other measurements, and hence the minimum value normalizes to a value of 1. For larger-the-better type indicators, it is the maximum value that is used to transform all other measurements, and the maximum value normalizes to a value of 1. The smaller-the-better and larger-the-better normalization functions are internal, so normalized values do not have meaning relative to exogenous system-defined targets or baselines. Also, for indicators with these bearings, the normalized value's significance comes from their relation only to the extremal elements from the data set in which they belong.

Target normalization compares individual measurements to predefined baseline and target values. These values can be system specific and tied to the environmental or socioeconomic sensitivities of the system being studied. They can also be uniform values, such as those provided by government regulations in the case of baselines, that may apply to multiple systems included in the study. Arguments for linking sustainability assessment outcomes to target or ideal levels, which is what target normalization accomplishes, can be found in the work of Moldan et al. (2012), Stiglitz et al. (2009) and Mayer (2008). Moldan et al. (2012) argues that "The benefit of specific, quantitative, time bound targets is then straightforward: The indicators can be linked to them and interpreted clearly on a distance-to-target basis." Mayer (2008) states that " indicators are more helpful if they give information on the state of the system with respect to policy targets or biophysical limits." Unlike ratio normalization, extremal elements in a data set do not influence normalized values when target normalization is used.

Beyond advocating for the inclusion of targets and baselines in sustainability, there is some discussion about how to determine and assign specific target and baseline values. Moldan et al. (2012) discusses ways in which target levels can be defined for sustainability indicators and provides example resources for their definition. Specifically, that study cites EEAiS: Star Portal Smeets et al. (1999), Millennium Development Goals, Eurostat, and Organisation for Economic and Co-Operation and Development (OECD) as potential resources for target references (Smeets et al., 1999; Eurostat, 2009; Nations, 2010; OECD, 2003). Further discussion and comparison of these normalization methods are provided throughout this paper.

3. Analyzing Normalization Functions

Studying the mathematical structure of the normalization functions provides insights into the implications that a given choice of normalization scheme may have on sustainability assessment outcomes. The four normalization schemes that are considered in this paper are *ratio normalization, Z-score normalization, unit equivalence normalization, and target normalization to the interval [0,1]* (Table 1). This paper analyzes how changes in the original, non-normalized data for a given indicator can cascade to alter composite sustainability scores. This investigation has implications for, not only the sensitivity of aggregate outcomes based on the normalization scheme chosen, but also the implicit weight or impact that a given normalization scheme has on particular measures of indicators. The comparability of assessment results based on the normalization procedure employed is also discussed.

3.1. Internal Normalization

Whether a normalization function is internal or not can have a large impact on how changes in non-normalized values can affect the total aggregate outcome. This effect occurs because each normalized indicator measurement depends on the full data set for that indicator. Internal normalization functions from literature were identified in the previous section. Of the four normalization schemes defined in Table 1, Z-score normalization is an internal normalization function; and unit equivalence normalization and target normalization to interval [0,1] are not. Within the ratio normalization scheme, the larger-the-better and smaller-the-better normalization functions are internal, while the distance-to-ideal function is not internal; hence internality is not necessarily a property of the normalization scheme but rather of individual normalization functions. In the case of ratio normalization functions, the use of $\min\{\boldsymbol{x}^*\}$ and $\max\{\boldsymbol{x}^*\}$ cause them to be internal, while Z-score normalization is internal from both the explicit use of the mean value, $\bar{x}^*$, and the calculation of the standard deviation, S_N. The case study provided in Section 5 motivates the importance of knowing if a normalization function is internal or not.

Table 3
Normalization function derivatives: using functions defined in Table 1, change in normalized value with respect to a change in the data point, x_j^*, is presented.

Change in normalized value with respect to change in x_j^*
Ratio normalization
$\frac{\partial}{\partial x_j^*}(R_{L,j}(\boldsymbol{x}^*)) = \begin{cases} \frac{1}{\max\{\boldsymbol{x}^*\}}, & x_j^* < x_k^* \forall k \neq j \\ 0, & \text{else} \end{cases}$
$\frac{\partial}{\partial x_j^*}(R_{S,j}(\boldsymbol{x}^*)) = \begin{cases} \frac{-\min\{\boldsymbol{x}^*\}}{\left(x_j^*\right)^2}, & x_j^* < x_k^* \forall k \neq j \\ 0, & \text{else} \end{cases}$
$\frac{\partial}{\partial x_j^*}(R_{D,j}(x_j^*, T)) = \begin{cases} \frac{1}{T}, & x_j^* < T \\ 0, & x_j^* = T \\ \frac{-T}{\left(x_j^*\right)^2}, & x_j^* > T \end{cases}$
Z-score normalization
$\frac{\partial}{\partial x_j^*}(Z_j(\boldsymbol{x}^*)) = \frac{\left(\left(\frac{1}{n}\sum_{j=1}^n \left(x_j^* - \bar{x}^*\right)^2\right)^{\frac{1}{2}}\left(1-\frac{1}{n}\right)\right) - \left(\frac{n^{-1/2}\left(x_j^* - \bar{x}^*\right)^2}{\left(\sum_{j=1}^n \left(x_j^* - \bar{x}^*\right)^2\right)^{1/2}}\right)\left(\frac{\sum_{k \neq j}\left(x_k^* - \bar{x}^*\right)}{-n\left(x_j^* - \bar{x}^*\right)} + \left(1-\frac{1}{n}\right)\right)}{\frac{1}{n}\sum_{j=1}^n (x_j^* - \bar{x}^*)^2}$
Unit equivalence normalization
$\frac{\partial}{\partial x_j^*}(C(x_j^*, c_f)) = c_f$
Target normalization to interval [0,1]
$\frac{\partial}{\partial x_j^*}(T_{L,j}(x_j^*, T, B)) = \begin{cases} \frac{1}{T-B}, & B < x_j^* < T \\ 0, & \text{else} \end{cases}$
$\frac{\partial}{\partial x_j^*}(T_{S,j}(x_j^*, T, B)) = \begin{cases} \frac{-1}{B-T}, & T < x_j^* < B \\ 0, & \text{else} \end{cases}$
$\frac{\partial}{\partial x_j^*}(T_{D,j}(x_j^*, T, B_l, B_u)) = \begin{cases} \frac{1}{T-B_l}, & B_l < x_j^* < T \\ \frac{-1}{B_u-T}, & T < x_j^* < B_u \\ 0, & \text{else} \end{cases}$

Note: $\boldsymbol{x}^* = \{x_1^*, x_2^*, \ldots, x_n^*\}$, T is a target or ideal value for a given indicator, B is a baseline or non-ideal value for a given indicator (B_l and B_u used when an upper and lower baseline are required), $S_N = \left[\frac{1}{n}\sum_{j=1}^n (x_j - \bar{x})^2\right]^{1/2}$ is the sample standard deviation, $\bar{x}^* = \frac{1}{n}\sum_{j=1}^n x_j^*$ is the sample mean, and c_f is a conversion factor to change units of $\boldsymbol{x}^*$ to alternate units, such as dollars or greenhouse gas equivalents.

3.2. Derivatives of Normalization Functions

The goal of investigating changes in the output of some normalization function naturally leads to the calculation and investigation of the derivatives of the normalization functions. Table 3 presents derivatives of the functions included in the four normalization schemes from Table 1 with respect to an arbitrary j^{th} non-normalized measurement, x_j^*. Two important properties considered here are piecewise differentiability of the normalization functions and how the derivative function depends on the variable being differentiated.

Many of the normalization functions are piecewise-defined and thus are differentiated piecewise. For ratio normalization functions, the presence of the $\min\{\boldsymbol{x}\}$ and $\max\{\boldsymbol{x}\}$ have a particular influence on the calculation of the derivative. Specifically, for $R_{L,j}(\boldsymbol{x}^*)$, as long as the non-normalized value being changed is not the maximum of the data set, $\max\{\boldsymbol{x}\}$, and does not become the maximum of the data set, the derivative is a constant value $\frac{1}{\max\{\boldsymbol{x}\}}$. However, in the case where the value changing is the maximum (or becomes the maximum), the behavior is quite different. Thus a complete characterization of the derivative must take these different possibilities into account (see Section 6.1). Similarly, target normalization behavior changes as measurements near the target and baselines values are varied and surpass these thresholds. How normalized values change near these threshold values and the impact of this behavior on aggregate sustainability scores are shown in further detail in the example included in Section 6.

In both internal and non-internal cases, how the variable of differentiation x_j^* appears in the derivative function is important (see Section 5.4). For example, the derivative of $R_{S,j}(\boldsymbol{x}^*)$ has fundamentally different behavior from nearly all other normalization functions considered due to the appearance of $(x_j^*)^2$ in the derivative (see Table 3). This difference leads to an impact on an aggregate score of sustainability that varies depending both on the value that is changing and the magnitude of the change, whereas the impact of normalization functions whose derivatives do not contain an x_j^* term is proportional to the change alone. An example of this effect is shown in the case study and discussed further in Section 6 below.

3.3. Comparability and Normalization

Assessments are often created with the goal of comparing alternative scenarios, different systems, or the same system at different points in time. The normalization scheme chosen has an affect on the comparability of results. Internal normalization schemes transform indicator measures based on the values present only in a particular data set. For a very simple example, consider two systems that are to be assessed and compared through measurements of a single, smaller-the-better type indicator. Let the first system have values (2, 5, 6, 2, 10) and the second system have values (20, 50, 60, 20, 100) for the indicator measured. If ratio normalization is used, these two very different data sets would normalize to equivalent the measures (1/5, 1/3, 2/5, 1, 1). Z-score standardization in this case behaves identically, since both data sets normalize to measures of $(0, -\frac{3}{\sqrt{11}}, -\frac{3}{\sqrt{11}}, \frac{1}{\sqrt{11}}, \frac{5}{\sqrt{11}})$. However, if these measures were normalized using unit equivalence normalization, the order of magnitude difference would be maintained. Depending on context and the indicator being measured, the order of magnitude difference showing up in normalized values may or may not be a necessary or desirable trait. For target normalization, the transformed values for an indicator depend not only on the individual measurement but also on targets and/or baseline(s) defined. If the same targets and baselines were used for both systems, the results would be distinguishable. Dependence, in both cases, of normalized values on targets and baselines leads to questions of comparability.

4. Normalization and Aggregate Measures of Sustainability

Thus far we have defined terminology, presented examples of common normalization functions, and shown a sample of the variety of these functions that can be found in the sustainability assessment literature. We have also provided derivatives of the functions included in four normalization schemes and defined relevant properties that can be used to classify types and behaviors of these normalization functions. Consideration now moves to how a change in a non-normalized value impacts an aggregate score of sustainability given a choice of normalization scheme and aggregation function(s). To carry out this analysis, its helps to place the normalization process into a context relevant to sustainability assessment.

Interpretation of composite sustainability scores is predicated by an understanding of the differential impacts of indicators. Langhans et al. (2014) present trade-off diagrams to show, for given aggregation functions and two indicators (each indicator represented by a single measurement), how much an increase in one indicator needs to be accompanied by an increase in the other indicator to have the same impact on the composite score. Pinar et al. (2014) also provide examples of how relative importance of indicators and interaction among indicators can be computed within their FEEM Sustainability Index.

In this paper, differential impacts, or sensitivities, of the aggregate sustainability score to changes in non-normalized indicator measurements are investigated by the use of derivative functions. Let $\mathcal{S}$ denote an aggregate sustainability score, where $\mathcal{S} = A(x_1, x_2, \ldots, x_n)$ for some aggregation function A. The change of the aggregate output, $\mathcal{S}$, with respect to a change in an indicator, x_i, is investigated by computing the partial derivative

$$\frac{\partial}{\partial x_i}(A(x_1, x_2, \ldots, x_n)) \tag{1}$$

However, the partial derivative in Eq. (1) assumes that there is one representative value for each indicator, x_i. In practice, the x_i values are often the aggregate of multiple measurements, j, for a given indicator, i. Adding in this detail, let $x_i = \alpha_i(x_{ij})$, for some aggregation function α_i that combines the various measurements for indicator i. Now we can ask how $\mathcal{S}$ is impacted by changes of individual indicator measurements, x_{ij}. Calculating the change in $\mathcal{S}$ as an indicator measurement, x_{ij}, changes leads to computing the partial derivative by use of the chain rule

$$\begin{aligned}&\frac{\partial}{\partial x_{ij}}(A(\alpha_1(x_{1j}), \alpha_2(x_{2j}), \ldots, \alpha_n(x_{nj})) \\ &\quad = A'(\alpha_1(x_{1j}), \alpha_2(x_{2j}), \ldots, \alpha_n(x_{nj}))\alpha_i'(x_{ij})\end{aligned} \tag{2}$$

where the prime notation ' ′ ' represents the partial derivative with respect to x_{ij}. Again, this is equation is often not representing the full picture because normalization is performed on individual indicator measurements before aggregation. Let $x_{ij} = f_i(x_{ij}^*)$ be the output of some normalization function, f_i, for indicator i that is acting on the raw indicator data, x_{ij}^*, where all raw data for a given indicator are normalized using the same normalization function, f_i. Thus, for a full treatment of how $\mathcal{S}$ is impacted by changes in non-normalized indicator data, we need to add this final detail to our derivatives. This leads to the following:

$$\begin{aligned}&\frac{\partial}{\partial x_{ij}^*}(A(\alpha_1(f_1(x_{1j}^*)), \alpha_2(f_2(x_{2j}^*)), \ldots, \alpha_n(f_n(x_{nj}^*))) \\ &\quad = A'(\alpha_1(f_1(x_{1j}^*)), \alpha_2(f_2(x_{2j}^*)), \ldots, \alpha_n(f_n(x_{nj}^*)))\alpha_i'(f_i(x_{ij}^*))f_i'(x_{ij}^*)\end{aligned} \tag{3}$$

In this case, the prime notation ' ′ ' represents the partial derivative with respect to x_{ij}^*. Eq. (3) is calculated to determine the impact of a

Table 4
Arithmetic and geometric mean definitions and derivatives: for the arithmetic mean (*AM*), geometric mean (*GM*), weighted arithmetic mean (*WAM*), and weighted geometric mean (*WGM*) a change in aggregate value with respect to a change in an input component, x_i, is presented.

$A(\boldsymbol{x})$: aggregation function	$\frac{\partial}{\partial x_i}(A(\boldsymbol{x}))$
Arithmetic means	
$AM(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i$	$\frac{\partial}{\partial x_i}(AM(\boldsymbol{x})) = \frac{1}{n}$
$WAM(\boldsymbol{x},\boldsymbol{w}) = \sum_{i=1}^{n} w_i x_i$	$\frac{\partial}{\partial x_i}(WAM(\boldsymbol{x},\boldsymbol{w})) = w_i$
Geometric means	
$GM(\boldsymbol{x}) = \prod_{i=1}^{n}(x_i)^{1/n}$	$\frac{\partial}{\partial x_i}(GM(\boldsymbol{x})) = \frac{1}{n}(\prod_{j\neq i} x_j)(\prod_{i=1}^{n} x_i)^{1/n-1}$
$WGM(\boldsymbol{x},\boldsymbol{w}) = \prod_{i=1}^{n}(x_i^{w_i})$	$\frac{\partial}{\partial x_i}(WGM(\boldsymbol{x},\boldsymbol{w})) = w_i x_i^{w_i-1}\prod_{j\neq i}(x_j^{w_j})$

Note: $\boldsymbol{x} = \{x_1, x_2 \ldots, x_n\}$ and $\boldsymbol{w} = \{w_1, w_2 \ldots, w_n\}$ where $0 \le w_i < 1$ and $\sum_{i=1}^{n} w_i = 1$.

variable, in this case a raw data measurement x^*_{ij}, that has been acted on by a normalization and aggregation function, f_i and α_i, respectively, before being acted on by A to determine the final output for $\mathcal{S}$. Although Eq. (3) is beginning to look like a bit of a monster, if the aggregation functions A and α_i are those of common employ, such as the weighted or unweighted arithmetic or geometric mean, the derivatives are quite straightforward (see Table 4). The derivative in Eq. (3) serves as the road map for the analysis that takes place in the case study that follows.

A summary of how raw data are transformed is as follows: First raw data, x^*_{ij}, for indicators are normalized using f_i, then normalized data, x_{ij}, are aggregated for each indicator using α_i, and finally multiple indicator aggregates, x_i, are combined using some aggregation function A to derive a sustainability score $\mathcal{S}$. Fig. 1 summarizes this procedure and presents a flowchart that describes the process. Fig. 1 also identifies examples of relevant properties and parameters for consideration at each step as measurements move from raw data, x^*_{ij}, to a sustainability score, $\mathcal{S}$.

5. Case Study

The following case study is presented to link the varying properties of normalization functions discussed in Section 3 to behaviors in example aggregate scores of sustainability. The sustainability assessment structure outlined in Fig. 1 is followed. This application uses Eq. (3) to understand how changes in raw data impact a sustainability score, $\mathcal{S}$, under eight different scenarios. In this case each scenario is a choice of a normalization scheme, which determine functions f_i, and the choice of an aggregation function A, that is used to compute $\mathcal{S}$ from the combined normalized indicator measures. The scenarios included in the case study are the combinations of two different normalization schemes, ratio and target normalization to [0,1], with four different aggregation functions: the weighted and non-weighted arithmetic and geometric means.

5.1. Background Information on Assessing Progress Towards Bioenergy Sustainability

This case study builds from work to identify a limited set of indicators of progress toward sustainability for bioenergy systems and data collected for those indicators. Researchers at the Center for BioEnergy Sustainability at Oak Ridge National Laboratory have identified 35 indicators covering environmental, social, and economic aspects of sustainability of bioenergy systems (Dale et al., 2013, McBride et al., 2011). Under research of the Southeastern Partnership for Integrated Biomass Supply Systems (IBSS), data were collected for a number of these indicators for switchgrass (*Panicum virgatum*) to examine actual yields and production costs under a wide range of

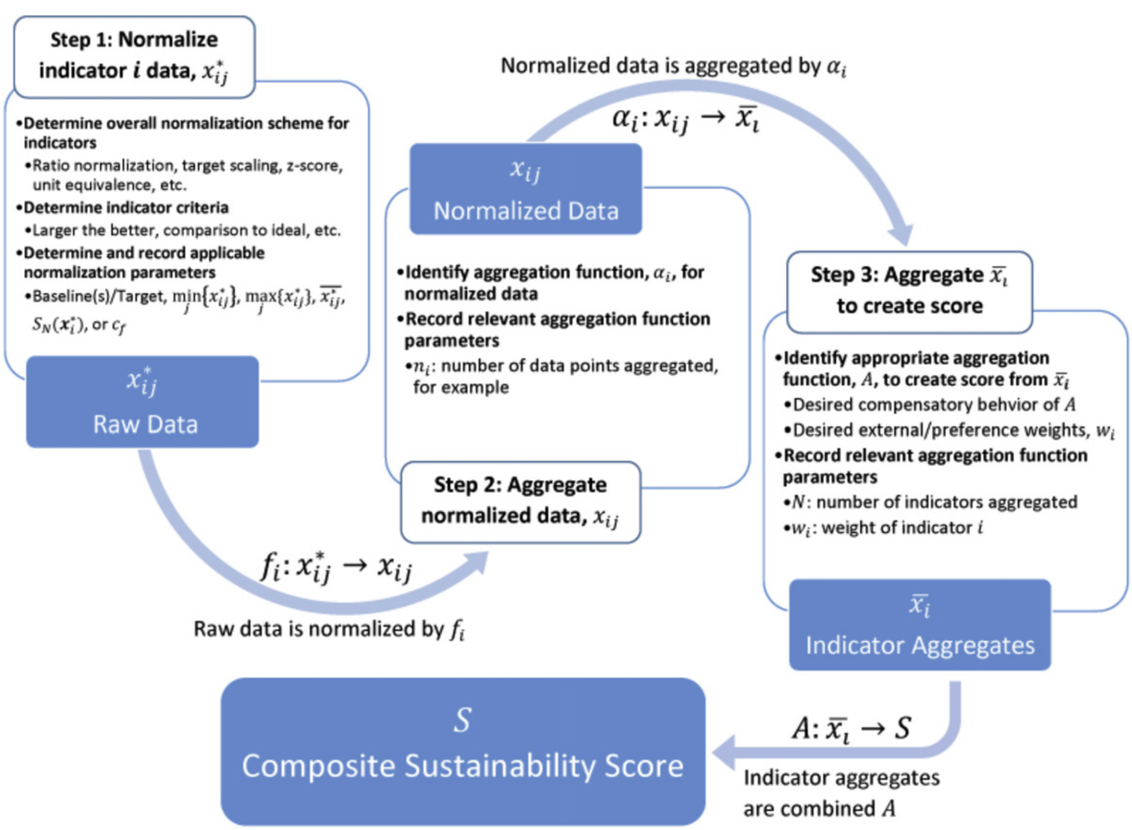


Fig. 1. Flowchart of a normalization and aggregation procedure utilized in multi-criteria sustainability assessment. Beginning with raw data for an indicator, data is transformed and aggregated multiple times before its eventual inclusion in a sustainability index or score.

physical settings and realistic farm management conditions in east Tennessee (Parish et al., 2016). Switchgrass is native to the southeastern United States and was planted within an eleven-county area to support a demonstration-scale ethanol production biorefinery in Vonore, Tennessee, that was operated by DuPont Cellulosic Ethanol (Tiller, 2011). To illustrate how normalization can affect aggregate scores of sustainability, this case study focuses on aggregation of three indicators: phosphorus levels in an adjacent water catchment (mg/L), yield of switchgrass (tons/acre), and percent organic matter (%OM) of soil in fields within the study using two normalization scenarios. The combination of these indicators has been chosen due to availability of high quality data sets. A comprehensive assessment of progress towards bioenergy sustainability would utilize many more of the 35 indicators identified in McBride et al. (2011) and Dale et al. (2013). Given that the focus of this paper is on the analysis methodology presented, using this small set of indicators is done to aid in clarity and tractability. The approach developed in this paper is general and can be applied to a variety of assessments scenarios when aggregate scores are derived through normalized indicator measurements where the number of indicators, normalization scheme, and aggregation functions can vary.

5.2. Ratio Normalization Scenarios

In the ratio normalization scenarios, the derivation of the composite sustainability score $\mathcal{S}$ follows the procedure outlined below. Results are provided in Table 5.

1. Individual measures, x^*_{ij}, of each indicator are normalized to values, x_{ij}, under the function $x_{ij} = f_i(x^*_{ij})$. In this example, $f_1 = R_{S,j}$ (smaller-the-better, ratio normalization), $f_2 = R_{L,j}$ (larger-the-better, ratio normalization), and $f_3 = R_{L,j}$ (larger-the-better, ratio normalization). See Table 1 for definitions.
2. The arithmetic mean is employed to aggregate the normalized measurements, such that $\alpha_i(f_i(x^*_{ij})) = \frac{1}{n_i}\sum_{j=1}^{n_i} f_i(x^*_{ij})$, where n_i is the number of measurements for indicator i, to give an aggregate value for the normalized measures, which is denoted as $\bar{x}_i = \alpha_i(f_i(x^*_{ij}))$.
3. The aggregate normalized measures for each indicator, $\bar{x}_i$, are used to to calculate $\mathcal{S}$ under the aggregation function $A(\bar{x}_1, \bar{x}_2, \bar{x}_3)$. In this case A is taken to be each of the arithmetic mean (AM), geometric mean (GM), weighted arithmetic mean (WAM), and weighted geometric mean (WGM).

5.3. Target Normalization Scenarios

For the target normalization scenarios, the derivation of the composite sustainability score $\mathcal{S}$ follows the procedure outlined below. Notation has been changed from x to y in the target normalization scheme to distinguish between the two different normalization procedures; the underlying data sets for the indicators are the same in both cases. Results, along with baseline and target values used for the target normalization scenarios, are provided in Table 6.

1. Individual measures, y^*_{ij}, of each indicator are normalized to values, x_{ij}, under the function $y_{ij} = f_i(y^*_{ij})$. In this example, $f_1 = T_{S,j}$ (smaller-the-better, target normalization), $f_2 = T_{L,j}$ (larger-the-better, target normalization), and $f_3 = T_{L,j}$ (larger-the-better, target normalization). See Table 1 for definitions.
2. The arithmetic mean is employed to aggregate the normalized measurements, such that $\alpha_i(f_i(y^*_{ij})) = \frac{1}{n_i}\sum_{j=1}^{n_i} f_i(y^*_{ij})$, where n_i is the number of measurements for indicator i, to give an aggregate value for the normalized measures, which will be denoted as $\bar{y}_i = \alpha_i(f_i(y^*_{ij}))$.
3. The aggregate normalized measures for each indicator, $\bar{y}_i$, are used to to calculate $\mathcal{S}$ under the aggregation function $A(\bar{y}_1, \bar{y}_2, \bar{y}_3)$. In this case A is taken to be each of the arithmetic mean (AM), geometric mean (GM), weighted arithmetic mean (WAM), and weighted geometric mean (WGM).

Table 5
Ratio normalization of bioenergy sustainability indicators: Indicators, notation, and appropriate parameters for the indicator data sets are presented.

Ratio normalization				
x_i : Indicator (*units*)	n_i	$\min_j\{x^*_{ij}\}$	$\max_j\{x^*_{ij}\}$	$\bar{x}_i$
x_i : Phosphorus (*mg/L*)	113	0.003	0.491	0.036
x_2 : Yield (*tons/acre*)	10	0	6.58	0.422
x_3 : % Organic matter (%)	120	0	7.24	0.410
$\mathcal{S}$: Composite score derived from ratio normalized indicators				
$\mathcal{S} = AM(\bar{x}_1, \bar{x}_2, \bar{x}_3) = 0.290$				
$\mathcal{S} = GM(\bar{x}_1, \bar{x}_2, \bar{x}_3) = 0.185$				

Note: n_i is the number of measurements for indicator i, $\bar{x}_i$ is the arithmetic mean of the ratio normalized measurements of indicator i.

Table 6
Target normalization of bioenergy sustainability indicators: indicators, notation, and appropriate parameters for the indicator data sets are presented.

Target normalization				
y_i : Indicator (*units*)	n_i	Baseline	Target	$\bar{y}_i$
y_1 : Phosphorus (*mg/L*)	113	$B = 0.1$	$T = 0$	0.123
y_2 : Yield (*tons/acre*)	10	$B = 0$	$T = 8$	0.348
y_2 : % Organic matter (%)	120	$B = 1.4$	$T = 4$	0.545
$\mathcal{S}$: Composite score derived from target normalized indicators				
$\mathcal{S} = AM(\bar{y}_1, \bar{y}_2, \bar{y}_3) = 0.339$				
$\mathcal{S} = GM(\bar{y}_1, \bar{y}_2, \bar{y}_3) = 0.286$				

Note: The notation y_i is used to distinguish between the two different normalization procedures, the underlying data set for each indicator is the same in both cases. n_i is the number of measurements for indicator i, $\bar{y}_i$ is the arithmetic mean of the target normalized measurements of indicator i.

Baseline and target levels for yield (*tons/acre*) are derived from expert opinion based on extensive data for the case study as described in Parish et al. (2016). Phosphorus concentration (*mg/L*) target and baseline levels have been set to 0 (*mg/L*) and 0.1 (*mg/L*) based on what is considered a critical concentration (Walker, 2000). The values for % organic matter baseline and target levels are based on Brady et al. (1996), who provide 1.5%–4.0% as a range of values for %OM in Ultisol, the dominant soil type in the bioenergy cropping region for this case study.

5.4. Quantifying Impacts of Indicator Measurements

Aggregate values for each indicator and for the composite sustainability score $\mathcal{S}$ are different when internally normalized by ratio normalization and when tied to external target and baseline values in the target normalization process. These differences are to be expected. Comparing and applying meaning to the different aggregate results derived from ratio and target normalization is not recommended, given how different the two normalization approaches are. However, what can be contrasted is how the composite score of sustainability, $\mathcal{S}$, is impacted, as non-normalized data measures change in each normalization scenario explored.

The impact, or weight, that individual indicator measurements carry into the score $\mathcal{S}$ can become unclear in composite scores of sustainability. For example, in this case study, one may ask if the measurements of phosphorus are having more influence on $\mathcal{S}$ than the measurements of yield? One might also ask, what role the

Table 7
Change in composite scores of ratio normalized bioenergy sustainability indicators as a function of non-normalized indicator measurements. The weighted arithmetic and geometric mean derivatives have been left in a general form to show influence of the weights, w_i, without need of a particular specification.

$\frac{\partial}{\partial x^*_{ij}}(\mathcal{S})$: Change in composite score with respect to change in x^*_{ij}
Arithmetic mean: $\mathcal{S} = AM$
$\frac{\partial}{\partial x^*_{1j}}(AM(\bar{x_1},\bar{x_2},\bar{x_3})) = (\frac{1}{3})(\frac{1}{n_1})\left(\frac{-\min_j\{x^*_{1j}\}}{(x^*_{1j})^2}\right) = (-8.85\times 10^{-6})(x^*_{1j})^{-2}$
$\frac{\partial}{\partial x^*_{2j}}(AM(\bar{x_1},\bar{x_2},\bar{x_3})) = (\frac{1}{3})(\frac{1}{n_2})\left(\frac{1}{\max_j\{x^*_{2j}\}}\right) = 0.0051$
$\frac{\partial}{\partial x^*_{3j}}(AM(\bar{x_1},\bar{x_2},\bar{x_3})) = (\frac{1}{3})(\frac{1}{n_3})\left(\frac{1}{\max_j\{x^*_{3j}\}}\right) = 0.0004$
Geometric mean: $\mathcal{S} = GM$
$\frac{\partial}{\partial x^*_{1j}}(GM(\bar{x_1},\bar{x_2},\bar{x_3})) = (\frac{1}{3})(\bar{x_2}\bar{x_3})^{1/3}(\bar{x_1})^{-2/3}(\frac{1}{n_2})\left(\frac{-\min_j\{x^*_{1j}\}}{x^*_{1j}}\right) = -(4.33\times 10^{-6})(x^*_{1j})^{-2}(\bar{x_1})^{-2/3}$
$\frac{\partial}{\partial x^*_{2j}}(GM(\bar{x_1},\bar{x_2},\bar{x_3})) = (\frac{1}{3})(\bar{x_1}\bar{x_3})^{1/3}(\bar{x_2})^{-2/3}(\frac{1}{n_2})\left(\frac{1}{\max_j\{x^*_{2j}\}}\right) = 0.0012(\bar{x_2})^{-2/3}$
$\frac{\partial}{\partial x^*_{3j}}(GM(\bar{x_1},\bar{x_2},\bar{x_3})) = (\frac{1}{3})(\bar{x_1}\bar{x_2})^{1/3}(\bar{x_3})^{-2/3}(\frac{1}{n_3})\left(\frac{1}{\max_j\{x^*_{3j}\}}\right) = 0.0001(\bar{x_3})^{-2/3}$
Weighted arithmetic mean: $\mathcal{S} = WAM$
$\frac{\partial}{\partial x^*_{1j}}(WAM(\bar{x_1},\bar{x_2},\bar{x_3})) = (w_1)(\frac{1}{n_1})\left(\frac{-\min_j\{x^*_{1j}\}}{x^*_{1j}}\right)$
$\frac{\partial}{\partial x^*_{2j}}(WAM(\bar{x_1},\bar{x_2},\bar{x_3})) = (w_2)(\frac{1}{n_2})\left(\frac{1}{\max_j\{x^*_{2j}\}}\right)$
$\frac{\partial}{\partial x^*_{3j}}(WAM(\bar{x_1},\bar{x_2},\bar{x_3})) = (w_3)(\frac{1}{n_3})\left(\frac{1}{\max_j\{x^*_{3j}\}}\right)$
Weighted geometric mean: $\mathcal{S} = WGM$
$\frac{\partial}{\partial x^*_{1j}}(GM(\bar{x_1},\bar{x_2},\bar{x_3})) = (w_1)(\bar{x_2}^{w_2}\bar{x_3}^{w_3})(\bar{x_1})^{1-w_1}(\frac{1}{n_2})\left(\frac{-\min_j\{x^*_{1j}\}}{x^*_{1j}}\right)$
$\frac{\partial}{\partial x^*_{2j}}(GM(\bar{x_1},\bar{x_2},\bar{x_3})) = (w_2)(\bar{x_1}^{w_2}\bar{x_3}^{w_3})(\bar{x_2})^{1-w_2}(\frac{1}{n_2})\left(\frac{1}{\max_j\{x^*_{2j}\}}\right)$
$\frac{\partial}{\partial x^*_{3j}}(GM(\bar{x_1},\bar{x_2},\bar{x_3})) = (w_3)(\bar{x_1}^{w_1}\bar{x_2}^{w_2})(\bar{x_3})^{1-w_3}(\frac{1}{n_3})\left(\frac{1}{\max_j\{x^*_{3j}\}}\right)$

Note: The derivatives above hold for the case when, for indicator 1, $x^*_{1j} > x^*_{1k} \forall k \neq j$, otherwise $x^*_{1j} = \min_j x^*_{1j}$ and thus takes on the constant normalized value of 1, and thus the derivative is 0. For indicators 2 and 3, the case is similar, but the derivatives hold when $x^*_{2j} < x^*_{2k}$ and $x^*_{3j} < x^*_{3k} \forall k \neq j$. The weights must satisfy $0 \leq w_i < 1$ and $\sum_{i=1}^{n} w_i = 1$.

normalization and aggregation functions chosen have on determining any differential impacts on $\mathcal{S}$ for specific indicators? Tables 5 and 6 give values for $\mathcal{S}$ as calculated through the arithmetic and geometric mean for the ratio and target normalization schemes. Tables 7 and 8 give the partial derivatives of those scores with respect to changes in non-normalized indicator measures for each of the three indicators as calculated using Eq. (3); these derivatives serve as the starting point in elucidating differing impact on changes in $\mathcal{S}$ from different normalization schemes. The derivatives in Tables 7 and 8 give exact formulas for analysis; the plots given in Figs. 2 and 3 provide another way to visualize differences in impact of changes in the non-normalized indicator measures given the different normalization functions applied. All derivative functions can be calculated with respect to an arbitrary non-normalized measurement x^*_{ij}; however, in order to create the visualizations in Figs. 2 and 3, a specific indicator measurement in the data set must be chosen to vary. In this case the median value was chosen and varied, and the corresponding value of $\mathcal{S}$ was calculated and plotted. Together, the derivatives and the visualizations provide two tools that can be used to study how normalization affects this composite score of sustainability.

With the results presented in Tables 5 through 8, one can see not only how the different normalization functions affect $\mathcal{S}$ but also how the aggregation functions for individual indicators, α_i, and the aggregation function A affect the value of $\mathcal{S}$ as non-normalized measures are changed. As discussed in Section 3.2, given the piecewise definition of many of the normalization functions, a similar piecewise definition of the derivatives is needed. For clarity, the derivatives presented in Tables 7 and 8 represent the behavior of $\mathcal{S}$ for x^*_{ij} values changing away from the minimum and maximum values, for the ratio normalization scheme, and in between the target and baseline values, for the target normalization scheme. Further discussion of how $\mathcal{S}$ changes as non-normalized values take on the minimum and maximum values is given in detail in Section 6.1.

6. Discussion

In order to understand the different influences that a normalization function can have on a composite sustainability score, properties of normalization functions have been discussed, and an analysis using partial derivatives of the aggregation and normalization functions has been presented. Even though calculation of the derivatives shown in Table 7 and Table 8 adds a step to the assessment process, the application and interpretation of the quantities derived improves overall understanding of the composite score for the sustainability indicators.

The first useful information added by calculating the derivatives is nearly by definition; derivatives indicate the per unit change in the composite score, $\mathcal{S}$, due to a per unit change in a non-normalized measurement, x^*_{ij}. As such, differences in the derivatives quantify differential sensitivities in the sustainability score by indicator. For example, Table 7 shows that the composite score $\mathcal{S}$ is nearly 12 times as sensitive to a per unit change in a yield measurement as it is to a unit change in a %OM measure when ratio normalization and the arithmetic mean are used.

Beyond just the different sensitivities of the aggregate score, $\mathcal{S}$, with respect to changes in indicator measurements, Figs. 2 and 3 show that different normalization procedures lead to fundamentally different ways in which indicator measurements impact the composite score. This difference is apparent not only across the

Table 8
Change in composite scores of target normalized bioenergy sustainability indicators as a function of non-normalized indicator measurements. The weighted arithmetic and geometric mean derivatives have been left in a general form to show influence of the weights, w_i, without need of a particular specification.

$\frac{\partial}{\partial y^*_{ij}}(\mathcal{S})$: Change in composite scores with respect to change in y^*_{ij}
Arithmetic mean: $\mathcal{S} = AM$
$\frac{\partial}{\partial y^*_{1j}}(AM(\bar{y_1},\bar{y_2},\bar{y_3})) = (\frac{1}{3})(\frac{1}{n_1})(\frac{-1}{B_1-T_1}) = -0.0295$
$\frac{\partial}{\partial y^*_{2j}}(AM(\bar{y_1},\bar{y_2},\bar{y_3})) = (\frac{1}{3})(\frac{1}{n_2})(\frac{1}{T_2-B_2}) = 0.0042$
$\frac{\partial}{\partial y^*_{3j}}(AM(\bar{y_1},\bar{y_2},\bar{y_3})) = (\frac{1}{3})(\frac{1}{n_3})(\frac{1}{T_3-B_3}) = 0.0011$
Geometric mean: $\mathcal{S} = GM$
$\frac{\partial}{\partial y^*_{1j}}(GM(\bar{y_1},\bar{y_2},\bar{y_3})) = (\frac{1}{3})(\bar{y_2}\bar{y_3})^{1/3}(\bar{y_1})^{-2/3}(\frac{1}{n_1})\left(\frac{-1}{B_1-T_1}\right) = -0.0169(\bar{y_1})^{-2/3}$
$\frac{\partial}{\partial y^*_{2j}}(GM(\bar{y_1},\bar{y_2},\bar{y_3})) = (\frac{1}{3})(\bar{y_1}\bar{y_3})^{1/3}(\bar{y_2})^{-2/3}(\frac{1}{n_2})\left(\frac{1}{T_2-B_2}\right) = 0.0017(\bar{y_2})^{-2/3}$
$\frac{\partial}{\partial y^*_{3j}}(GM(\bar{y_1},\bar{y_2},\bar{y_3})) = (\frac{1}{3})(\bar{y_1}\bar{y_2})^{1/3}(\bar{y_3})^{-2/3}(\frac{1}{n_3})\left(\frac{1}{T_3-B_3}\right) = 0.0004(\bar{y_3})^{-2/3}$
Weighted arithmetic mean: $\mathcal{S} = WAM$
$\frac{\partial}{\partial y^*_{1j}}(AM(\bar{y_1},\bar{y_2},\bar{y_3})) = (w_1)(\frac{1}{n_1})(\frac{-1}{B_1-T_1})$
$\frac{\partial}{\partial y^*_{2j}}(AM(\bar{y_1},\bar{y_2},\bar{y_3})) = (w_2)(\frac{1}{n_2})(\frac{1}{T_2-B_2})$
$\frac{\partial}{\partial y^*_{3j}}(AM(\bar{y_1},\bar{y_2},\bar{y_3})) = (w_3)(\frac{1}{n_3})(\frac{1}{T_3-B_3})$
Weighted geometric mean: $\mathcal{S} = WGM$
$\frac{\partial}{\partial y^*_{1j}}(GM(\bar{y_1},\bar{y_2},\bar{y_3})) = (w_1)(\bar{y_2}^{w_2}\bar{y_3}^{w_3})(\bar{y_1})^{1-w_1}(\frac{1}{n_1})\left(\frac{-1}{B_1-T_1}\right)$
$\frac{\partial}{\partial y^*_{2j}}(GM(\bar{y_1},\bar{y_2},\bar{y_3})) = (w_2)(\bar{y_1}^{w_1}\bar{y_3}^{w_3})(\bar{y_2})^{1-w_2}(\frac{1}{n_2})\left(\frac{1}{T_2-B_2}\right)$
$\frac{\partial}{\partial y^*_{3j}}(GM(\bar{y_1},\bar{y_2},\bar{y_3})) = (w_3)(\bar{y_1}^{w_1}\bar{y_2}^{w_2})(\bar{y_3})^{1-w_3}(\frac{1}{n_3})\left(\frac{1}{T_3-B_3}\right)$

Note: These derivatives hold for the case when y^*_{ij} falls within the interval created by the targets (T_i) and baselines (β_i) for the respective indicators. The weights must satisfy $0 \leq w_i < 1$ and $\sum_{i=1}^{n} w_i = 1$.

normalization schemes presented, but fundamental differences can emerge within the same scheme. For example, setting aside behavior changes near the extremal values, within the ratio normalization scheme if an indicator is of the type smaller-the-better, then a change in a measurement of that indicator has an impact on the composite score that depends on the value of the measurement changing. This is due to the presence of x^*_{1j} term in the derivative (see Figs. 2a and 2b). However, if the indicator is of the type larger-the-better, then the impact of a change in a measurement of that indicator on $\mathcal{S}$ is independent of the value that is changing (see Figs. 2c, 2e, 2d, and 2f). With respect to the extremal values, it can also be seen that as the minimum (maximum) value changes in smaller-the-better (larger-the-better) type indicators, there is a dramatic change in the score of $\mathcal{S}$ (see Fig. 2). This change is due to the internal normalization property of the ratio normalization functions, and, when the minimum or maximum changes for a data set, all of the other measurements in the data set also change causing $\mathcal{S}$ to be very sensitive to changes in the extremal values of the data set.

In the target normalization scheme, Fig. 3 shows that in both smaller-the-better and larger-the-better indicator bearing, behavior does not display the fundamental differences that can be be seen between these two bearings in the ratio normalization scheme. In the case of the final aggregate score calculated through the arithmetic mean, changes in $\mathcal{S}$ are constant when non-normalized indicator measures change between the baseline and target values. For the geometric mean, although the change between baseline and target values appears to be constant, it is not, as the derivatives in Table 8 show. In both cases changes in $\mathcal{S}$ are 0 when non-normalized measures change beyond the baseline and target measures defined. Target normalization is free from the dramatic changes in $\mathcal{S}$ that appear in ratio normalization as extremal values change in the data set. In target normalization, if a non-normalized measure changes to a value beyond the baseline or target, it ends up having no impact on $\mathcal{S}$ because the normalized value for that measurement is constant at either 0 or 1 beyond those threshold values.

6.1. On the Piecewise Nature of Derivatives Encountered

All functions of the ratio normalization and target normalization schemes are piecewise differentiable; this fact leads to complicated behavior of the derivatives. Up to this point of the paper, there has been limited discussion on how to analyze the behavior of $\mathcal{S}$ when non-normalized measures change to become the minimum or maximum values in the ratio normalization scheme and when non-normalized measures move beyond the baseline and target values in target normalization.

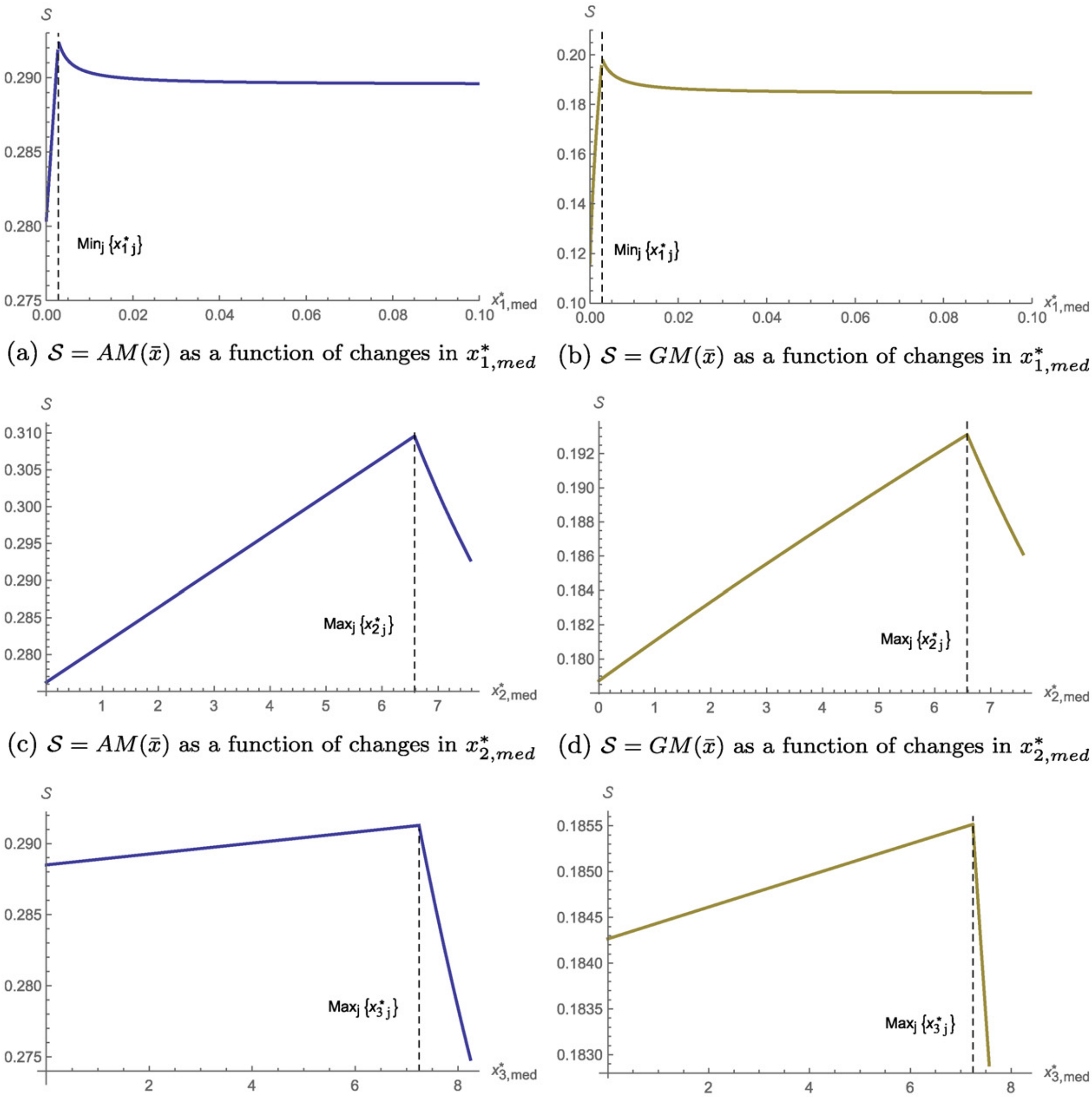
Fig. 2 addresses the behavior of $\mathcal{S}$ in the ratio normalization scenarios. Notice that in all six plots that the derivatives given in Table 7 hold until the indicator value becomes the maximum or minimum value for the data set, at which point one needs to consider how a change not just in x^*_{ij} impacts $\mathcal{S}$, given by $\frac{\partial \mathcal{S}}{\partial x^*_{ij}}$ but also how changing $\max_j\{x^*_{ij}\}$ impacts $\mathcal{S}$, given by $\frac{\partial \mathcal{S}}{\partial \max_j\{x^*_{ij}\}}$. In practice for ratio normalization, there are four different cases one needs to consider to capture all the behaviors that may occur as indicator values change:

1. x^*_{ij} *is not the maximum* (or minimum, respectively), and changes *do not cause* it to become so. In this case, one can analyze the impact of changes in x^*_{ij} without need to consider $\frac{\partial \mathcal{S}}{\partial \max_j\{x^*_{ij}\}}$ (these are the derivatives shown in Table 7).
2. x^*_{ij} *is not the maximum* (or minimum, respectively), and changes *cause* it to become so. In this case, one must consider both $\frac{\partial \mathcal{S}}{\partial x^*_{ij}}$ and $\frac{\partial \mathcal{S}}{\partial \max_j\{x^*_{ij}\}}$.
3. x^*_{ij} *is the maximum* (or minimum, respectively), and changes *do not cause* it to become otherwise. In this case one need only consider $\frac{\partial \mathcal{S}}{\partial \max_j\{x^*_{ij}\}}$.
4. x^*_{ij} *is the maximum* (or minimum, respectively), and changes *cause* it to become otherwise. In this case one again needs to consider both $\frac{\partial \mathcal{S}}{\partial x^*_{ij}}$ and $\frac{\partial \mathcal{S}}{\partial \max_j\{x^*_{ij}\}}$.

In the target normalization scheme, the changes in $\mathcal{S}$ as a function of the changes in indicator values are more easily captured, even though they are also defined piecewise. This behavior is due to the fact that target normalization is not an internal normalization process. The cases for target normalization have to do with the measurement value changing between the baseline and target values defined and the changing beyond those values. As it is shown in Fig. 3, there is no dramatic change as indicator values move outside the interval defined by the baselines and targets. In fact, once an indicator measurement moves beyond the baseline or target values, the change in $\mathcal{S}$ becomes exactly 0.

7. Opportunities for Further Research

Further research may seek to investigate additional normalization functions not included in this paper. The techniques developed in this paper are general in their application, with regard to the type of normalization functions and aggregation functions that can be analyzed and used. In addition to expanding analysis to other normalization schemes, determining the sensitivity in the calculation of a composite sustainability score, $\mathcal{S}$, to other normalization scheme parameters, such as the targets and baseline(s) defined in target normalization, would also be valuable. Additional topics of interest include the quantification of implicit weights and studying normalization in the context of meaningful aggregation. These two topics are discussed next and examples are included.



(a) $\mathcal{S} = AM(\bar{x})$ as a function of changes in $x^*_{1,med}$ (b) $\mathcal{S} = GM(\bar{x})$ as a function of changes in $x^*_{1,med}$

(c) $\mathcal{S} = AM(\bar{x})$ as a function of changes in $x^*_{2,med}$ (d) $\mathcal{S} = GM(\bar{x})$ as a function of changes in $x^*_{2,med}$

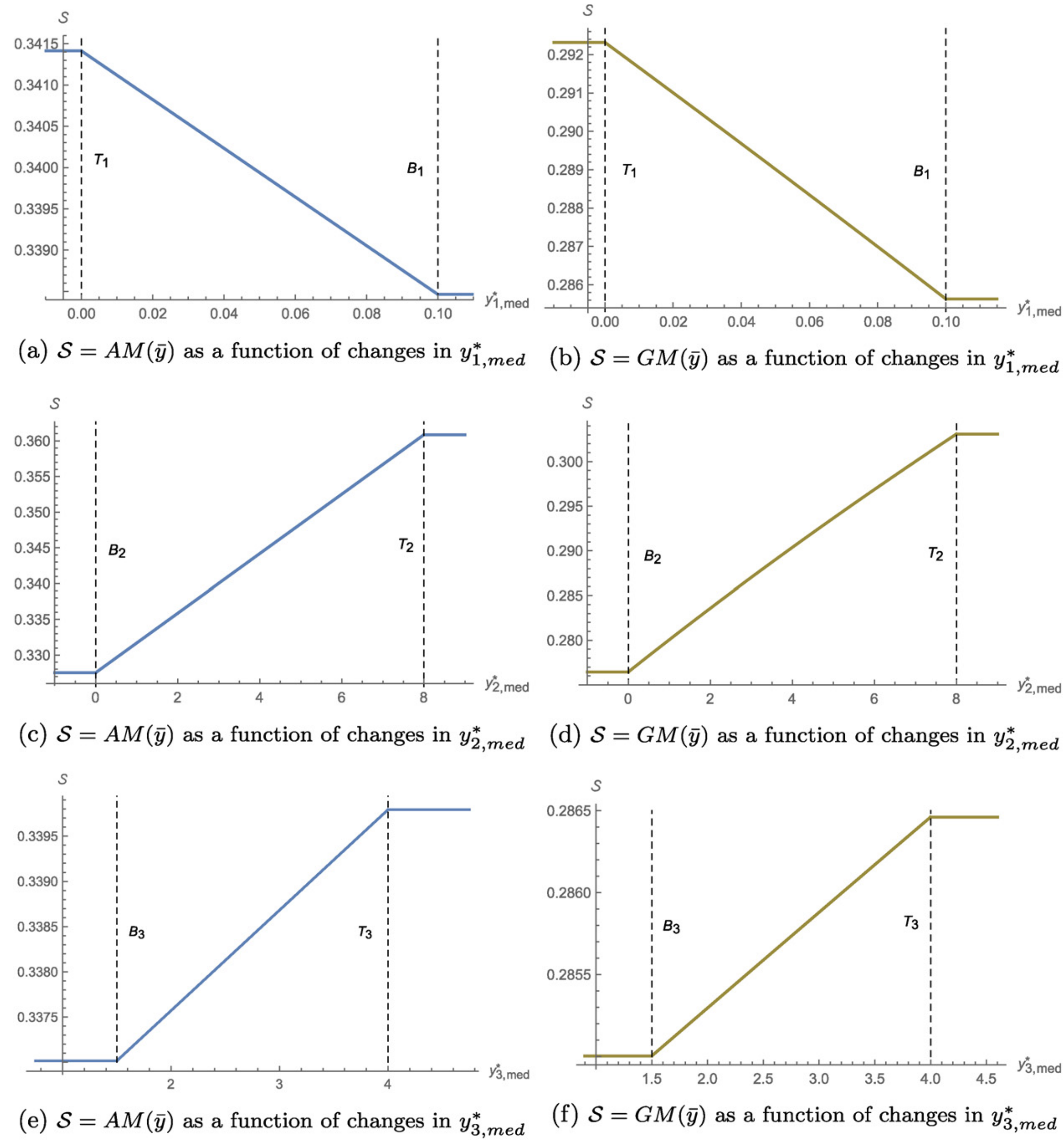
(e) $\mathcal{S} = AM(\bar{x})$ as a function of changes in $x^*_{3,med}$ (f) $\mathcal{S} = GM(\bar{x})$ as a function of changes in $x^*_{3,med}$

Fig. 2. Changes in $\mathcal{S}$ in response to changes in median values of from the data sets for phosphorus ($x^*_{1,med}$), yield ($x^*_{2,med}$), and %OM ($x^*_{3,med}$) under ratio normalization scheme. Dashed lines show $\min_j x^*_{ij}$ and $\max_j x^*_{ij}$ values from Table 5. In all cases, the median value from each data set was varied in order to show the effect that changing non-normalized indicator measures has on $\mathcal{S}$. Notice the behavior of $\mathcal{S}$ when the median value becomes the $\min_j x^*_{ij}$ in (a,b) and the $\max_j x^*_{ij}$ in (c,d,e,f). This dramatic change is due to the fact that ratio normalization is an internal normalization process, and the dependence of all normalized values in the data set on the minimum or maximum value of that data set. All functions depicted correspond to functions presented in Table 7.

7.1. Quantification of Implicit Weights

In some specific cases, the differential impacts on $\mathcal{S}$ can be written as a set of weights associated with each indicator given the normalization function, f_i, indicator aggregation function, α_i, and final aggregation function, A that are chosen. Further development of methods to quantify implicit weights could prove useful and would provide stakeholders a quick way to determine the relative importance placed on each indicator resulting from the mathematical structure of the sustainability score. The simplest case of when these implicit weights can be calculated occurs when the the partial derivatives, found using Eq. (3), are constant. For example, consider the derivatives in Table 8 when $\mathcal{S} = AM$, the arithmetic mean and values are not changing beyond baseline or target measures. We have the derivatives for $\frac{\partial \mathcal{S}}{\partial y^*_{1,j}} = -0.0295$, $\frac{\partial \mathcal{S}}{\partial y^*_{2,j}} = 0.0042$, and $\frac{\partial \mathcal{S}}{\partial y^*_{3,j}} = 0.0011$. The implicit weights, call them w_i, for each indicator measurement are

$$w_i = \frac{\left|\frac{\partial \mathcal{S}}{\partial y^*_{i,j}}\right|}{\sum_{k=1}^{3}\left|\frac{\partial \mathcal{S}}{\partial y^*_{k,j}}\right|},$$



(a) $\mathcal{S} = AM(\bar{y})$ as a function of changes in $y^*_{1,med}$ (b) $\mathcal{S} = GM(\bar{y})$ as a function of changes in $y^*_{1,med}$

(c) $\mathcal{S} = AM(\bar{y})$ as a function of changes in $y^*_{2,med}$ (d) $\mathcal{S} = GM(\bar{y})$ as a function of changes in $y^*_{2,med}$

(e) $\mathcal{S} = AM(\bar{y})$ as a function of changes in $y^*_{3,med}$ (f) $\mathcal{S} = GM(\bar{y})$ as a function of changes in $y^*_{3,med}$

Fig. 3. Changes in $\mathcal{S}$ in response to changes in median values of from the data sets for Phosphorus ($y^*_{1,med}$), Yield ($y^*_{2,med}$), and %OM ($y^*_{3,med}$) under target normalization scheme. In all cases, the median value from each data set was varied in order to show the effect that changing non-normalized indicator measures has on $\mathcal{S}$. Dashed lines show normalization parameters from Table 6. When the median value is changed to values outside the baseline and target intervals, there is no response in $\mathcal{S}$ to further changes due to the normalized value becoming a constant 0 or 1. All functions depicted correspond to functions presented in Table 8.

specifically, this scenario produces weights of $w_1 = 0.85$, $w_2 = 0.12$, $w_3 = 0.03$ for phosphorus, yield, and %OM matter indicators, respectively. However, it should be pointed out that these are changes in $\mathcal{S}$ per a unit change in the indicators within the target and baseline range. For indicators, such as the water quality indicator of Phosphorus, a unit change (say from 0 to 1 mg/L) is very large and would in fact move any measurement within the target and baseline values to a value outside of that range. Once again, the challenge of working with multiple indicators on various scales shows up. In an instance such as this, which is likely to be very common in sustainability assessment, the question becomes, how can one use the information contained in the derivatives to quantify implicit weights adjusted to the scales of the indicators?

Using the derivatives, $\frac{\partial \mathcal{S}}{\partial y^*_{1,j}} = -0.0295$, $\frac{\partial \mathcal{S}}{\partial y^*_{2,j}} = 0.0042$, and $\frac{\partial \mathcal{S}}{\partial y^*_{3,j}} = 0.0011$, we can capture a more relevant quantity related to changing an indicator measure by, instead of considering a single unit change, using the baseline and target values to provide a range for changes that are relevant to the indicator. Specifically, one can multiply the derivative value by the difference in the baseline and targets,

- $(-0.0295)|T_1 - B_1| = (-0.0295)|0 - 0.1| = 0.00295$
- $(0.0042)|T_2 - B_2| = (0.0042)|8 - 0| = 0.0336$
- $(0.0011)|T_3 - B_3| = (0.0011)|4 - 1.5| = 0.00275,$

for Phosphorus, Yield, and %OM, respectively. This use of the derivatives, baselines, and targets has immediately provided something useful; this calculation is the analytical analog to derive the numerical quantities that one can gather by taking the difference of the maximum and minimum values of the plots provided in Figs. 3a, 3c, 3e, respectively. With these scale adjusted responses of $\mathcal{S}$ to changes in indicator measurements, we can now revisit our quantification of implicit weights and calculate *scale-adjusted implicit weights*, $\hat{w}_i$, in the following way,

$$\hat{w}_i = \frac{|T_i - B_i| \left| \frac{\partial \mathcal{S}}{\partial y^*_{i,j}} \right|}{\sum_{k=1}^{3} |T_k - B_k| \left| \frac{\partial \mathcal{S}}{\partial y^*_{k,j}} \right|}. \tag{4}$$

Using this formulation, the scale adjusted weights are $\hat{w}_1 = 0.075$, $\hat{w}_2 = 0.855$, $\hat{w}_3 = 0.070$ for phosphorus, yield, and %OM matter indicators, respectively. These scale-adjusted weights now represent the relative impact each indicator measurement has on the aggregate output $\mathcal{S}$ as the measurement varies between the baseline and target values.

7.2. Normalization and Meaningful Aggregation

How normalization functions transform *measurability scales* of data can also be investigated in order to utilize results from previous research into *meaningful* statements made with aggregate values, sometimes just referenced as *meaningful aggregation*. The topic of meaningful aggregation arises in sustainability and environmental assessment (Pollesch and Dale, 2015;Roberts, 2014;Böhringer and Jochem, 2007;Zhou et al., 2006;Ebert and Welsch, 2004) and uses the measurability scale of indicator data to provide a method for selecting aggregation functions. Examples of *measurability scales* include ratio, interval, ordinal, and nominal (Stevens, 1946). Knowledge of these scales, along with an application of Luce's principle (Luce, 1959), is used to ensure that, when data are transformed, they are transformed in such a way that the information contained in the data is maintained; such transformations are referred to as *meaningful transformations*. If one determines how or if the normalization function changes the scale of measurement of the data being considered, it is then possible to utilize results of previous research in meaningful aggregation and to create an aggregate score of sustainability that adheres to the principles therein. For example, ratio-scale measurable indicators occur frequently in sustainability assessment. These indicators are identified by differences between data points having meaning, ratios of data points having meaning, and the existence of a non-arbitrary zero point for the data being measured. Pollesch and Dale (2015) showed that of the 19 environmental indicators for bioenergy sustainability identified in McBride et al. (2011), all but one indicator is ratio-scale measurable. For an example of how normalization functions affect scales of measurement, consider an indicator that is ratio-scale measurable.

- Applying any of the functions in the ratio normalization scheme to a ratio-scale measurable indicator results in a unit less ratio-scale measurable indicator. The non-arbitrary zero value stays the same, and the normalized value now defines a new ratio scale.
- Z-score standardization of ratio-scale measurable data assigns a value of zero to the mean value of the data set, and the unit less quantity represented by a Z-score is also ratio-scale measurable. The normalized value represents the number of standard deviations the original value is away from the mean, thus Z-score standardization transforms ratio-scale measurable data to a new ratio scale.
- Unit equivalence normalization is scalar multiplication, and thus for any non-zero conversion factor c_f, the normalized measurability scale of ratio-scale measurable data is a new ratio scale.

As an opportunity for future research, further investigation as to how the normalization process changes measurability scale can be carried out for different combinations of measurability scale types and normalization functions. This analysis would allow identification of meaningful aggregation functions for indicators included in a sustainability assessment.

8. Conclusions

This paper investigates properties of normalization functions and explores the implications that different choices of normalization schemes can have when normalized values are included in aggregate measures of sustainability. Ratio normalization, Z-score normalization, unit equivalence normalization, and target normalization schemes are analyzed for their behavior in terms of internal normalization, the structure of their derivatives, and comparability of normalized values. We introduce the term *bearing* to unify the variety of terminology present in literature that is used to discuss this property of indicators. The case study motivates the theoretical analysis of normalization schemes by demonstrating how the properties of normalization functions manifest in the simple three-indicator bioenergy sustainability assessment provided.

Quantification of sustainability is approached using a variety of metrics, many of which utilize indicators as stand-alone measures or within aggregate values. Indicator approaches for assessing progress towards sustainability include, at a minimum, information about the economic, social, and environmental aspects of the system being studied. Given the large number of indicators that can be used within an assessment, there is often stakeholder interest and a benefit in combining sustainability indicators. Although clarity is seen as a benefit when combining indicators, as indicator measurements are combined, this benefit comes at the cost of lost information and data resolution. This is inherent in any aggregation procedure. Gasparatos and Scolobig (2012) provide a good discussion on tradeoffs arising in sustainability assessment. Normalization of indicators, although almost always prerequisite for aggregation of indicators, elicits tradeoffs within the analysis as well.

The case study shows differences of behavior between ratio and target normalization schemes. The consequence of ratio normalization functions being internal is especially evident in Fig. 2 where the aggregate score of sustainability $\mathcal{S}$ is impacted greatly as measures change the extremal values of the data set. The ratio normalization scheme has fundamental differences in the behavior of smaller-the-better and larger-the-better normalization functions. Specifically, for smaller-the-better bearing indicators, the impact of changes on $\mathcal{S}$ in non-normalized measures differs depending on where those measures are in relation to the minimum value for that data set (see Figs. 2a and 2b), whereas larger-the-better type indicators do not have this dependence. This discrepancy between indicator bearing type does not occur when target normalization is used. Both normalization schemes have behaviors that change as non-normalized measures are varied near threshold values. These are the minimum or maximum values, in the case of ratio normalization, and baseline and target values in target normalization. This set of behaviors differs near threshold values in complexity and influence on predictability of how aggregate outcomes are impacted by changes in non-normalized values across these two normalization schemes.

This research highlights some of the advantages and disadvantages associated with normalization schemes used in sustainability assessment and the calculation of a composite score of sustainability. The internal normalization procedures, Z-score and ratio normalization are easier to implement on a data set given that they do not require externally defined targets and baselines encountered in target normalization or the multitude of conversion factors required for unit equivalence normalization. However, the internal normalization procedures have disadvantages when it comes to the dependence exhibited in the normalized values on extremal values of the data set and how that dependence manifests in aggregate sustainability scores derived from those normalized values. With respect to ratio normalization, the difference between how changes in smaller-the-better and larger-the-better type indicators can impact normalized values and aggregate values derived thereof is of concern. The different cases that one might encounter due to the piecewise differentiability of ratio and target normalization functions are not present in Z-score and unit equivalence normalization. The change in normalized value with respect to changes in non-normalized measures presented in Table 3 show that unit-equivalence normalization has the simplest partial derivative expression of the four schemes, while Z-score normalization produces quite a complicated expression for the partial derivative, even without needing to consider the different scenarios of the piecewise defined derivatives for ratio and target normalization.

Of the four normalization schemes explored in-depth in this paper, target normalization stands out as a candidate for use within sustainability assessment. In sustainability assessment, context is extremely important, and a strength of target normalization is that it allows for the inclusion of contextually relevant normalization parameters in the forms of baseline and target values. This context specificity also aids in the interpretation of normalized values. Additionally, as discussed previously, functional forms across bearing type within target normalization are more consistent than those used within the ratio normalization scheme. Although target normalization is a stand-out when it comes to sustainability assessment for the reasons just provided, it is recommended that advantages and disadvantages of normalization schemes be considered before inclusion in any assessment application; this paper will aid researchers in this regard.

This paper will also help researchers and stakeholders by providing methods to clarify connections between normalization scheme and the accompanying impact that normalization functions choice can have on the aggregation of indicators measuring progress towards sustainability. The derivatives based approach shown in this paper was chosen to elucidate how general properties of normalization functions, such as internality, manifest to create specific dependencies in aggregate assessment outcomes. The derivatives based approach also provides a foundation upon which other analysis can be developed. Specifically, the scale adjusted implicit weights formulation (Eq.(4)) shows promise, with further development, to become a standard method for reporting indicator specific sensitivities that can accompany aggregate scores of sustainability.

Acknowledgments

This research was supported by the U.S. Department of Energy (DOE) under the Bioenergy Technologies Office. Oak Ridge National Laboratory is managed by the UT-Battelle, LLC, for DOE under contract DE-AC05-00OR22725. Discussions with Fred Roberts, Esther Parish, Harlan Stech, and Bruce Peckham have been helpful. Comments by Lou Gross on an earlier draft are also appreciated.

The case study in the paper builds from research of the Southeastern Partnership for Integrated Biomass Supply Systems (IBSS), which is funded through Agriculture and Food Research Initiative Competitive grant no. 2011-68005-30410 from the USDA National Institute of Food and Agriculture.

References

Bare, J., Gloria, T., Norris, G., 2006. Development of the method and US normalization database for life cycle impact assessment and sustainability metrics. Environ. Sci. Technol. 40 (16), 5108–5115.

Böhringer, C., Jochem, P.E., 2007. Measuring the immeasurable — A survey of sustainability indices. Ecol. Econ. 63 (1), 1–8.

Brady, N.C., Weil, R.R., et al. 1996. The nature and properties of soils. 11th, Prentice-Hall Inc.

Castoldi, N., Bechini, L., 2010. Integrated sustainability assessment of cropping systems with agro-ecological and economic indicators in northern italy. Eur. J. Agron. 32 (1), 59–72.

Dale, V.H., Efroymson, R.A., Kline, K.L., Langholtz, M.H., Leiby, P.N., Oladosu, G.A., Davis, M.R., Downing, M.E., Hilliard, M.R., 2013. Indicators for assessing socioeconomic sustainability of bioenergy systems: a short list of practical measures. Ecol. Indic. 26, 87–102.

Dias, L.C., Domingues, A.R., 2014. On multi-criteria sustainability assessment: spider–gram surface and dependence biases. Appl. Energy 113, 159–163.

Ebert, U., Welsch, H., 2004. Meaningful environmental indices: a social choice approach. J. Environ. Econ. Manage. 47 (2), 270–283.

Eurostat, 2009. Measuring progress towards a more sustainable europe. In 2009 monitoring report of the EU sustainable development strategy. Office for Official Publications of the European Communities. Luxembourg.

Freudenberg, M., 2003. Composite indicators of country performance. OECD Science. Technology and Industry Working Papers.

Gasparatos, A., Scolobig, A., 2012. Choosing the most appropriate sustainability assessment tool. Ecol. Econ. 80, 1–7.

Grabisch, M., Marichal, J.-L., Mesiar, R., Pap, E., 2009. Aggregation Functions. first, Cambridge University Press, New York.

Hayashi, T., van Ierland, E.C., Zhu, X., 2014. A holistic sustainability assessment tool for bioenergy using the Global Bioenergy Partnership (GBEP) sustainability indicators. Biomass and Bioenergy 66, 70–80.

Heijungs, R., Guinée, J., Kleijn, R., Rovers, V., 2007. Bias in normalization: causes, consequences, detection and remedies. Int. J. Life Cycle Assess. 12 (4), 211–216.

Krajnc, D., Glavič, P., 2005. A model for integrated assessment of sustainable development. Resour. Conserv. Recycl. 43 (2), 189–208.

Langhans, S.D., Reichert, P., Schuwirth, N., 2014. The method matters: a guide for indicator aggregation in ecological assessments. Ecol. Indic. 45, 494–507.

Luce, R.D., 1959. On the possible psychophysical laws. Psychol.Rev. 66 (2), 81.

Maxim, A., 2014. Sustainability assessment of electricity generation technologies using weighted multi-criteria decision analysis. Energy Policy 65, 284–297.

Mayer, A.L., 2008. Strengths and weaknesses of common sustainability indices for multidimensional systems. Environ. Int. 34 (2), 277–291.

McBride, A.C., Dale, V.H., Baskaran, L.M., Downing, M.E., Eaton, L.M., Efroymson, R.A., Garten, C.T., Kline, K.L., Jager, H.I., Mulholland, P.J., Parish, E.S., Schweizer, P.E., Storey, J.M., 2011. Indicators to support environmental sustainability of bioenergy systems. Ecol. Indic. 11 (5), 1277–1289.

Moldan, B., Janoušková, S., Hák, T., 2012. How to understand and measure environmental sustainability: indicators and targets. Ecol. Indic. 17, 4–13.

Mori, K., Christodoulou, A., 2012. Review of sustainability indices and indicators: towards a new City Sustainability Index (CSI). Environ. Impact Assess. Rev. 32 (1), 94–106.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., Giovannini, E., 2005. Handbook on Constructing Composite Indicators. OECD publishing.

Nations, U., 2010. The Millennium Development Goals Report 2010. United Nations Publications, New York.

Odum, H.T., Doherty, S.J., Scatena, F.N., Kharecha, P.A., 2000. Emergy evaluation of reforestation alternatives in Puerto Rico. For. Sci. 46 (4), 521–530.

OECD, O., 2003. Environmental Indicators-development, Measurement and Use. OCDE, Paris.

Parish, E.S., Dale, V.H., English, B.C., Jackson, S.W., Tyler, D.D., 2016. Assessing multimetric aspects of sustainability: application to a bioenergy crop production system in East Tennessee. Ecosphere 7 (2).

Pinar, M., Cruciani, C., Giove, S., Sostero, M., 2014. Constructing the feem sustainability index: a choquet integral application. Ecol. Indic. 39, 189–202.

Pollesch, N., Dale, V., 2015. Applications of aggregation theory to sustainability assessment. Ecol. Econ. 114, 117–127.

Roberts, F.S., 2014. Meaningful and meaningless statements in landscape ecology and environmental sustainability. In Clusters, Orders, and Trees: Methods and Applications, Springer., pp. 297–312.

Sadamichi, Y., Kudoh, Y., Sagisaka, M., Chen, S.S., Elauria, J.C., Gheewala, S.H., Hasanudin, U., Romero, J., Shi, X., Sharma, V.K., 2012. Sustainability assessment methodology of biomass utilization for energy in East Asian countries. J. Jpn. Inst. Energy 91 (10), 960–968.

Singh, R.K., Murty, H., Gupta, S., Dikshit, A., 2009. An overview of sustainability assessment methodologies. Ecol. Indic. 9 (2), 189–212.

Smeets, E., Weterings, R., voor Toegepast-Natuurwetenschappelijk, N.C.O., 1999. Environmental indicators: Typology and overview. European Environment Agency Copenhagen.

Stevens, S., 1946. On the theory of scales of measurement. Science 103 (2684), 677–680.
Stiglitz, J., Sen, A., Fitoussi, J.-P., 2009. The measurement of economic performance and social progress revisited. Reflections and Overview. Commission on the Measurement of Economic Performance and Social Progress, Paris.
Tiller, K., 2011. Tennessee case study in full-scale development and integration of switchgrass feedstocks with biorefinery operations: Genera Energy and University of Tennessee biofuels initiative. Ind. Biotechnol. 7 (5), 357–364.
Walker, F., 2000. Best management practices for phosphorus in the environment. Agricultural Extension Service, University of Tennessee,PB, 1645,
Zhou, P., Ang, B., Poh, K., 2006. Comparing aggregating methods for constructing the composite environmental index: an objective measure. Ecol. Econ. 59 (3), 305–311.