

## Environmental Research Letters



## LETTER

## OPEN ACCESS

RECEIVED  
20 May 2015

REVISED  
8 February 2016

ACCEPTED FOR PUBLICATION  
12 February 2016

PUBLISHED  
29 February 2016

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



# Making environmental assessments of biomass production systems comparable worldwide

Markus A Meyer<sup>1,2</sup>, Ralf Seppelt<sup>1,3,4</sup>, Felix Witing<sup>5</sup> and Joerg A Priess<sup>1</sup>

<sup>1</sup> UFZ—Helmholtz Centre for Environmental Research, Department Computational Landscape Ecology, Permoserstr. 15, 04318 Leipzig, Germany

<sup>2</sup> University of Munich (LMU), Department of Geography, Luisenstr. 37, 80333 Munich, Germany

<sup>3</sup> iDiv—German Centre for Integrative Biodiversity Research, Deutscher Platz 5E, 04103 Leipzig, Germany

<sup>4</sup> Martin-Luther-University Halle-Wittenberg, Institute of Geoscience & Geography, 06099 Halle (Saale), Germany

<sup>5</sup> UFZ—Helmholtz-Centre for Environmental Research, Department of Soil Physics, Theodor-Lieser-Str. 4, 06120 Halle (Saale), Germany

E-mail: [markus.meyer@lmu.de](mailto:markus.meyer@lmu.de) and [markus.a.meyer@googlemail.com](mailto:markus.a.meyer@googlemail.com)

**Keywords:** environmental heterogeneity, threshold, potential natural vegetation, stratification, bioenergy, environmental impact, ecosystem service

Supplementary material for this article is available [online](#)

## Abstract

Global demand for agricultural and forestry products fundamentally affects regional land-use change associated with environmental impacts (EIs) such as erosion. In contrast to aggregated global metrics such as greenhouse gas (GHG) balances, local/regional EIs of different agricultural and forestry production regions need methods which enable worldwide EI comparisons. The key aspect is to control environmental heterogeneity to reveal man-made differences of EIs between production regions. Environmental heterogeneity is the variation in biotic and abiotic environmental conditions. In the present study, we used three approaches to control environmental heterogeneity: (i) environmental stratification, (ii) potential natural vegetation (PNV), and (iii) regional environmental thresholds to compare EIs of solid biomass production. We compared production regions of managed forests and plantation forests in subtropical (Satilla watershed, Southeastern US), tropical (Rufiji basin, Tanzania), and temperate (Mulde watershed, Central Germany) climates. All approaches supported the comparison of the EIs of different land-use classes between and within production regions. They also standardized the different EIs for a comparison between the EI categories. The EIs for different land-use classes within a production region decreased with increasing degree of naturalness (forest, plantation forestry, and cropland). PNV was the most reliable approach, but lacked feasibility and relevance. The PNV approach explicitly included most of the factors that drive environmental heterogeneity in contrast to the stratification and threshold approaches. The stratification approach allows consistent global application due to available data. Regional environmental thresholds only included arbitrarily selected aspects of environmental heterogeneity; they are only available for few EIs. Especially, the PNV and stratification approaches are options to compare regional EIs of biomass or crop production such as erosion, biodiversity, or water quality impacts worldwide and thereby complement existing metrics assessing global EIs such as GHG emissions.

## 1. Introduction

The EU Renewable Energy Directive (EU RED), developed to mitigate climate change, to promote energy security, and to create jobs [1–5], has led to an increase in global biomass trade for bioenergy. Previous policies such as EU forest protection policies

displaced the demand for and production of forestry and agricultural products to other countries with weaker governance structures [6]. Increasing biomass trade results in remote land-use change and environmental impacts (EIs) (desirable or undesirable), i.e., telecouplings [7]. Increasing pressures on resources from global trade of forest and agricultural

commodities require comparative approaches to assess the EIs [8] in different production systems and world regions.

Biomass for bioenergy originates from various sources (e.g., agriculture or forestry) and production systems (e.g., managed forests or plantations) from different parts of the world. Lamers *et al* [9] expect an increase in EU imports of solid biomass from 2010 until 2020 by approximately 300% to 236 PJ. EU RED sustainability requirements implemented through certification schemes [10] should reduce unwanted EIs and social impacts of biomass production from global trade; but these requirements are weakly specified. Global EIs such as greenhouse gas (GHG) emissions are assessed on the basis of harmonized and standardized life-cycle assessments [11–13]. Such global EI assessments rarely represent major local/regional socio-economic (e.g., local societies' preferences) and environmental processes (e.g., water and matter fluxes). Therefore, Creutzig *et al* [14] request spatially explicit regional environmental assessments of bioenergy production. Specifically, Dale *et al* [15] indicate that soil and water quality impacts of biomass production are not suitably assessable by standardized and non-place-based indicators as the ones typically used in life-cycle assessments. Others address the need for the comparison of biodiversity [16, 17], and food and income security [18] in regional impact assessments. One of the key problems for interregional comparisons are missing broadly accepted methods taking the heterogeneity of environmental conditions into account.

Environmental heterogeneity is the spatial variation in biotic (vegetation and land cover) and abiotic conditions (climate, topography, and soil) [19]. We need to control the environmental heterogeneity of different regions to enable comparisons of regional environmental assessments. Biodiversity studies have confirmed that environmental heterogeneity affects for instance species richness [20]. Therefore, using approaches explicitly taking environmental heterogeneity into account would enable us to address the question whether species richness is affected by differences in land use or environmental conditions [21]. Kienast *et al* [22] indicate that environmental heterogeneity similarly affects the occurrence of beneficial and harmful EIs and that it requires spatially explicit quantification.

Options to account for the variation in EIs due to environmental heterogeneity are the use of (i) baseline conditions or (ii) thresholds [23, 24]. Two options exist to obtain (i) baseline conditions: a first approach is to divide a region into classes of comparable environmental conditions and to relate each EI value to the best and worst value of its class (environmental stratification, e.g., Metzger *et al* [25]). A second baseline approach is to relate EIs simulated for current land use/land cover (LU/LC) to potential natural vegetation (PNV) as a benchmark [26]. For example, West

*et al* [27] compare carbon (C) storage between current LU/LC and PNV at the global scale. Stratification and PNV approaches have not been used to compare EIs between regional-scale case studies, which is fundamental for ecosystem services (ESS) as positive EIs mostly assessed at the regional scale [28]. Instead, several studies have defined arbitrary thresholds to distinguish desirable and undesirable locations with respect to ESS. For example, Qiu and Turner [29] define the highest 20% of C storage values as desirable and the lowest 20% as undesirable locations.

(ii) Thresholds are typically set by policy makers or regional stakeholders such as farmers or environmental agency representatives. Stakeholders may set critical pollutant loads taking into consideration the regional ecosystem and its desired state [23, 30].

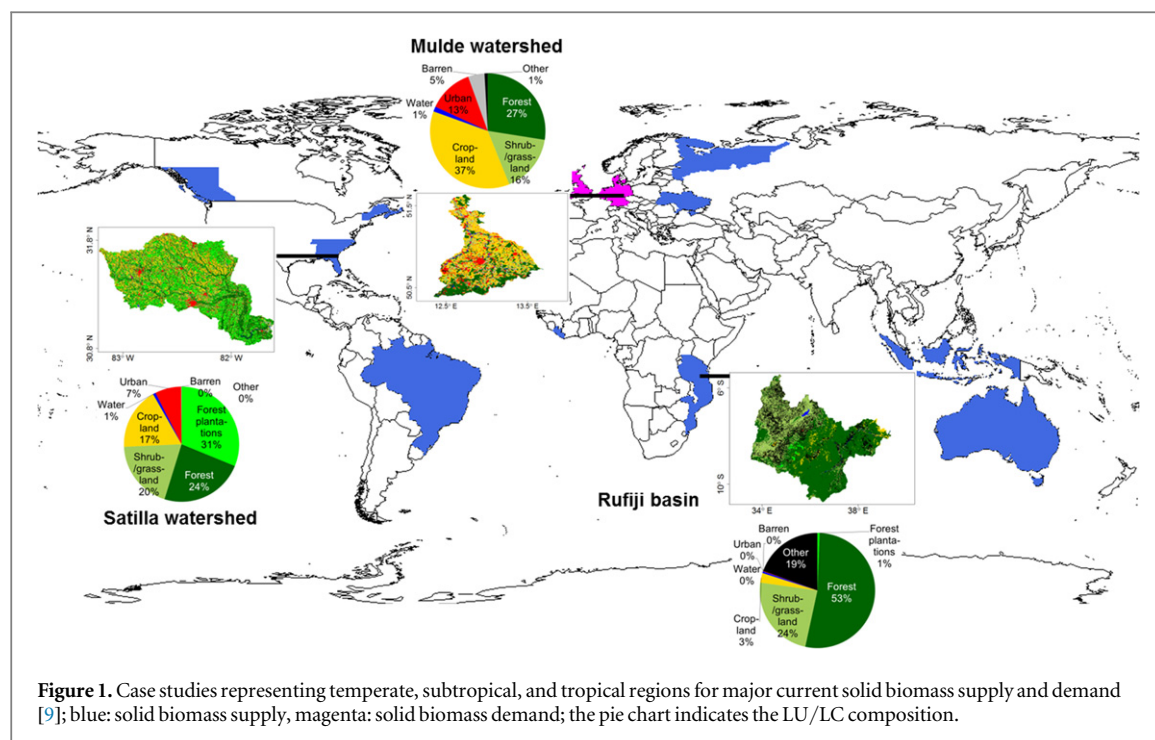
In the present study, first, we hypothesized that EIs are higher in regions with more intensive biomass production. Second, we hypothesized different EI outcomes for the different approaches that control environmental heterogeneity. We tested and compared methods to account for environmental heterogeneity with the objective to compare EIs of land use between different world regions. To achieve this objective, we stratified EIs, used PNV as a benchmark, and applied environmental thresholds of biomass production in three world regions. Third, we hypothesized that the different approaches differ with respect to reliability, feasibility, and relevance. We analyzed the methods for a broader use beyond bioenergy production in terms of reliability, feasibility, and relevance.

## 2. Material and methods

### 2.1. Production regions and assessed environmental impacts

The three selected production regions (figure 1) represent two major global solid biomass supply regions (Southeastern US and Tanzania) and one major demand and supply region (Central Germany) [9]. These regions cover the following socio-economic and regulatory conditions:

1. Southeastern US (Satilla watershed): developed country with legislation, industry standards, and voluntary schemes to ensure the sustainability of bioenergy [31].
2. Tanzania (Rufiji basin): developing country with existing, but weakly enforced legislation, and without broadly applied standards and schemes for bioenergy sustainability [32].
3. Central Germany (Mulde watershed): developed country with enforced legislation for the sustainability of bioenergy that is binding for biofuels and bioliquids including imports; enforced via certification schemes; future application of



**Figure 1.** Case studies representing temperate, subtropical, and tropical regions for major current solid biomass supply and demand [9]; blue: solid biomass supply, magenta: solid biomass demand; the pie chart indicates the LU/LC composition.

regulations to solid bioenergy production is expected [33].

These regions also represent different climate zones: humid continental (Mulde watershed: 5791 km<sup>2</sup>), humid subtropical (Satilla watershed: 8760 km<sup>2</sup>), and tropical wet and dry savanna (Rufiji basin: 176 301 km<sup>2</sup>).

C storage, sediment export and retention, and P export and retention were modeled with the InVEST package (Integrated Valuation of Ecosystem Services and Tradeoffs) [34, 35], while biodiversity was simulated with the GLOBIO model [36]. Further details are provided in the supporting information. We selected the local/regional EI based on scientific literature [31, 37, 38] and followed the selection of a global stakeholder panel for sustainability assessment of bioenergy production [39].

## 2.2. Congruence of sustainability assessment approaches

We applied three approaches to control environmental heterogeneity in the biomass production regions presented in section 2.1:

1. bioclimatic environmental stratification according to [40],
2. PNV [41–43] as a benchmark, and
3. regional environmental thresholds or target values [44–46].

For environmental stratification as best-in-class approach, we divided a region into classes or strata of comparable environmental conditions. We

standardized each EI value to the range of minimum (0) and maximum (1) of its class. In the PNV approach, we modeled the EIs for the current LU/LC and PNV. We compared the EI values for current LU/LC to the EI values for PNV as a benchmark. In the threshold approach, we modeled EI values for current LU/LC and compared them with regional threshold values.

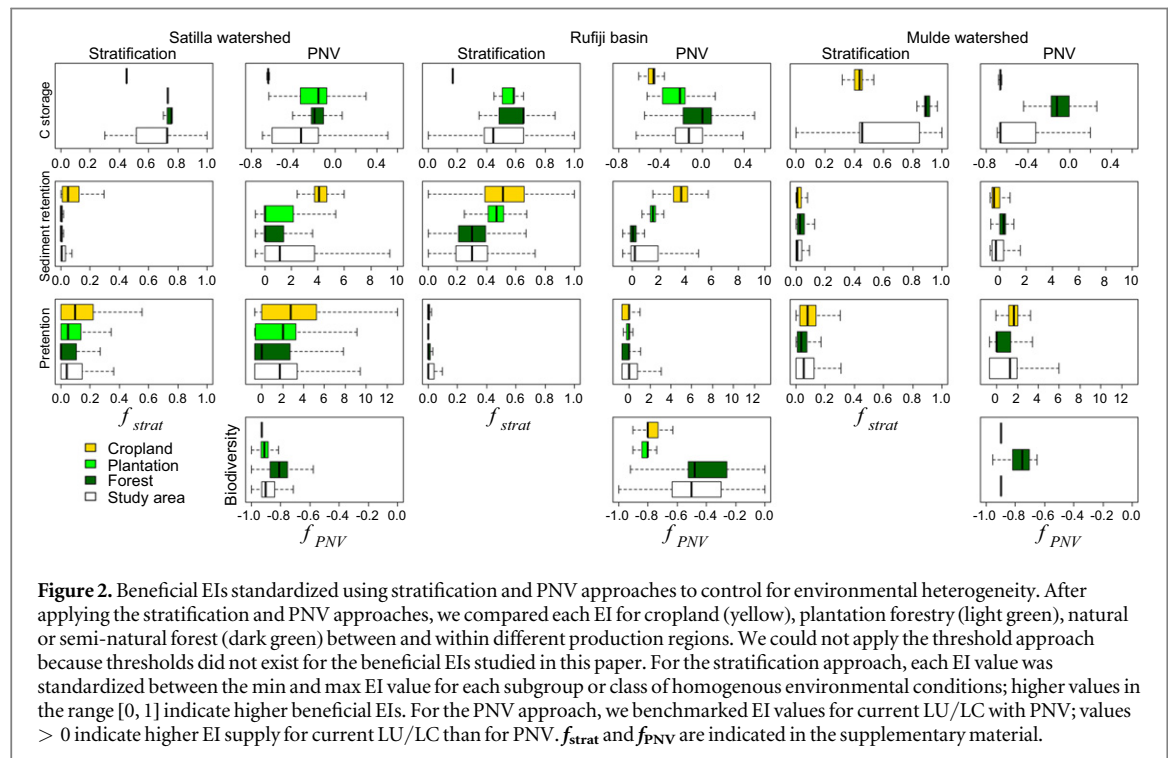
To test the congruence of results, i.e., the agreement between the different approaches, we calculated the cross-predictive capacity with diagnostic test statistics [47], using the R statistics package [48]. We obtained a score between 0 and 1 for the agreement of two approaches correctly and incorrectly identifying desirable and undesirable locations of EI. We aimed at identifying cases in which PNV, stratification and thresholds could be used interchangeably, i.e., to identify the degree of congruence of the three approaches. To characterize the congruent locations, we compared (1) the spatial extent and (2) the dominating LU/LC for the stratification and the PNV approaches. Further details can be found in the second and third sections of the supplementary material available at [stacks.iop.org/ERL/11/034005/mmedia](http://stacks.iop.org/ERL/11/034005/mmedia).

## 2.3. Reliability, feasibility, and relevance of the approaches

We evaluated the three approaches explained in the previous section with three major quality criteria from indicator evaluation and environmental management: reliability, feasibility, and relevance for the end user [2 49–53]. From existing literature, we collected the subcriteria for reliability, feasibility, and relevance listed in table 1.

**Table 1.** Evaluation subcriteria for the categories reliability, feasibility, and relevance.

	Subcriteria	Reference
Reliability	Worldwide consistent methodology and datasets	[54]
	Regional stakeholders and experts involved	[24]
	Environmental heterogeneity within a region	[55]
	Range of environmental factors considered	[20]
	Global data coverage	[55]
Feasibility	Easily applicable to various environmental impacts	[56]
	Development effort if no data available	[30]
Relevance	Classification in desirable and undesirable locations	[30]
	Relative comparison of globally distributed production regions	[55]



**Figure 2.** Beneficial EIs standardized using stratification and PNV approaches to control for environmental heterogeneity. After applying the stratification and PNV approaches, we compared each EI for cropland (yellow), plantation forestry (light green), natural or semi-natural forest (dark green) between and within different production regions. We could not apply the threshold approach because thresholds did not exist for the beneficial EIs studied in this paper. For the stratification approach, each EI value was standardized between the min and max EI value for each subgroup or class of homogenous environmental conditions; higher values in the range [0, 1] indicate higher beneficial EIs. For the PNV approach, we benchmarked EI values for current LU/LC with PNV; values  $> 0$  indicate higher EI supply for current LU/LC than for PNV.  $f_{strat}$  and  $f_{PNV}$  are indicated in the supplementary material.

### 3. Results

#### 3.1. Comparison of environmental assessments of biomass production systems across three world regions

Biomass production systems were located in the Southeastern USA (Satilla watershed), in Tanzania (Rufiji basin), and in Central Germany (Mulde watershed). For the targeted solid biomass, the PNV and stratification approaches had the following consistent rankings: C storage and biodiversity were highest for forestry in the Mulde watershed (figure 2). Sediment retention was highest for the plantation forestry in the Rufiji basin, whereas P retention was highest for the plantation forestry in the Satilla watershed.

The stratification and PNV approaches provided a consistent ranking of individual LU/LC classes between the production regions for three cases of beneficial EIs: sediment retention for cropland and

sediment and P retention for plantation forestry (figure 2 and table S3). For harmful EIs (figure S1 and table S3), the stratification and PNV approaches had a consistent ranking for P export. The PNV and threshold approaches had consistent rankings for sediment export for cropland and forestry and P export for forestry.

C storage and biodiversity for current LU/LC declined the least in the Rufiji basin compared to PNV as a benchmark. P and sediment retention increased most in the Satilla watershed compared to PNV. Sediment export increased most in the Satilla watershed and P export in the Mulde watershed compared to PNV. Therefore, the change from PNV to current land use in the Satilla watershed more strongly affected the supply of P retention than in the Mulde watershed. In the PNV approach, P retention compared with other EIs increased most for the study regions' mean, for forestry, and for plantation forestry. Sediment retention increased and C storage decreased less compared

**Table 2.** Congruence of approaches to identify desirable and undesirable locations of EIs; the percentages indicate the quantiles for the stratification approach which are most congruent with the PNV or threshold approaches. The EI values are split into two groups with 10%-quantiles to obtain desirable and undesirable locations. We indicated the best fitting value for all EIs (total) and for the individual EIs per production region; sensitivity (congruence of pixels between two approaches classified as (un)desirable locations) and specificity (congruence of pixels between two approaches not classified (un)desirable locations) scores are indicated in brackets; score values of 0 indicate no congruence and 1 indicates complete congruence between two approaches; see tables S4–14 for the entire results.

Beneficial EIs	Total	C storage	Sediment retention	P retention
Stratification/PNV				
<i>Satilla</i>	50% (0.5/0.9)	90% (0.2/1.0)	60% (0.5/0.9)	40% (0.8/0.9)
<i>Rufiji</i>	70% (0.5/0.8)	70%–90% (0.8/0.8)	mean (0.5/0.7)	70% (0.4/0.8)
<i>Mulde</i>	50% (0.6/0.9)	90% (0.7/1.0)	60% (0.7/0.8)	40% (0.9/0.9)
Harmful EIs			Sediment export	P export
Stratification/PNV				
<i>Satilla</i>			mean (1.0/0.2)	50% (0.7/1.0)
<i>Mulde</i>			60% (1.0/0.4)	40% (0.6/0.7)
Stratification/threshold				
<i>Satilla</i>			mean (1.0/0.0)	NA
<i>Mulde</i>			NA	mean (0.0/0.4)
PNV/threshold				
<i>Satilla</i>			(0.5/1.0)	NA
<i>Mulde</i>			(0.9/0.0)	(0.8/0.0)

to PNV. Cropland compared with other LU/LC class deviated most compared with PNV.

The ranking within the production regions was mostly consistent between the stratification and PNV approaches. Beneficial EIs were higher and harmful EIs were lower for plantation forestry than for cropland. As an exception, sediment and P retention were higher for cropland than for plantation forestry than for cropland in the Satilla watershed and thereby followed the ranking of P and sediment export. The PNV approach ranked biodiversity for plantation forestry slightly better than for cropland in the Rufiji basin.

### 3.2. Congruence of sustainability assessment approaches

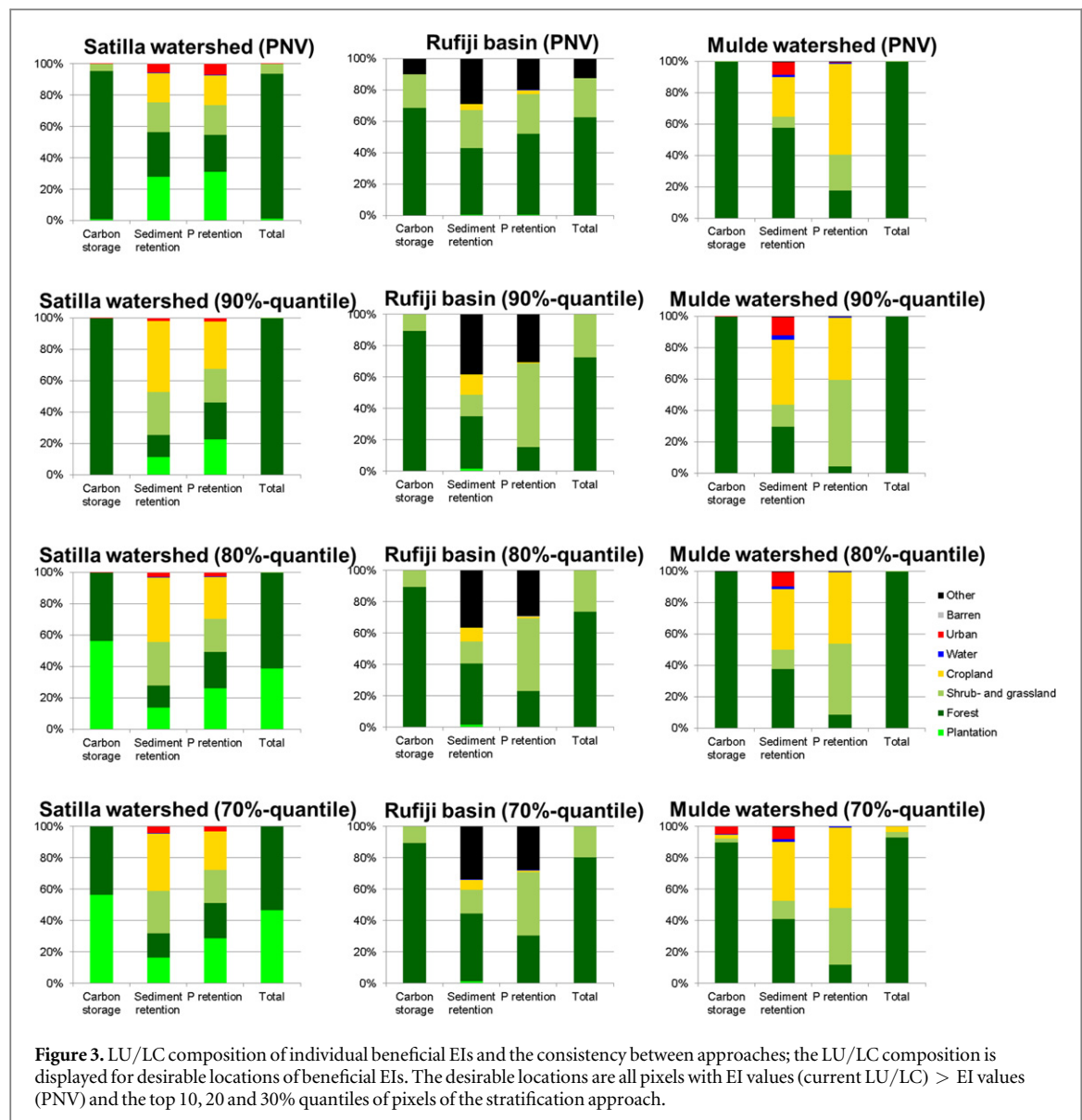
We analyzed if the PNV and stratification approaches detected congruent desirable locations of beneficial EIs, see table 2. The PNV and stratification approaches in the Satilla and Mulde watersheds detected desirable locations at similar quantiles. In contrast, the PNV and stratification approaches in the Rufiji basin detected desirable locations of EIs at different quantiles. This difference showed that the identification of desirable locations of the stratification approach depended on the overall land-use intensity and the resulting range of worst- and best-in-class values. If stratification with quantiles was used instead of PNV, the chance that locations of strongly desirable EIs would not be detected is high; the chance that locations of strongly desirable EIs would be included would be low. This showed the lower sensitivity than the specificity (see table 2).

For harmful EIs, the PNV and stratification approaches selected undesirable locations of harmful EIs in the Satilla and Mulde watersheds at different quantiles. Stratification compared with PNV overestimated the size and number of undesirable locations of sediment export and underestimated those of P export. The NA-values and low specificity when comparing stratification with the threshold approach, and PNV with the threshold approach, respectively, reflected the low P and sediment export rates above locally set thresholds. Thresholds only selected extremely undesirable locations of harmful EIs.

### 3.3. LU/LC differences between desirable locations (PNV versus stratification approaches)

In the Satilla watershed, 1.1% of the area was desirable locations of EIs in the PNV and 0.8% in the stratification approaches (80%-quantile). Plantation forestry accounted for 1% (PNV approach) and 39% (80%-quantile in the stratification approach) of total desirable locations (figure 3). Beneficial EI values were high for plantation forestry compared with other LU/LC classes, but were mostly lower than the natural state, i.e., PNV. In the Rufiji basin, 13.8% of the area was classified as desirable locations in PNV and only 0.7% in the stratification approaches (80%-quantile). The larger desirable locations of the PNV approach in the Rufiji basin showed the lower decline of beneficial EIs due to the lower land-use intensity compared with the Satilla and Mulde watersheds. The stratification approach did not reveal this lower land-use intensity comparing the production regions. In the Mulde watershed, 3.1% of the area was classified as desirable





locations in the PNV and 0.5% in the stratification approach (80%-quantile). In both cases, all desirable locations were forests. A comparison of the 90%- and 70%-quantiles with the 80%-quantile showed that the stratification approach increasingly included LU/LC with higher land-use intensity (cropland, plantation forestry) for lower quantiles.

### 3.4. Reliability, feasibility, and relevance of the approaches

The reliability of the comparative approaches in the present study depends on *consistent methodology* and the data required to analyze EIs in different production systems. The environmental stratification approach applied with a single global dataset promises more comparability than using regional PNV approaches (see table 3). The PNV approach is less comparable because various ecological concepts are used inconsistently. For example, natural disturbance through

fire is only considered in the Satilla watershed. The involvement of *regional stakeholders and experts* creates inconsistencies because production regions with different regional strategies are compared. The available environmental thresholds are defined for an entire area and do not consider *environmental heterogeneity within a region* in contrast to the stratification and PNV approaches. The *range of environmental factors* considered to control environmental heterogeneity varies between the approaches. The PNV approach includes more abiotic (topography and soil) and biotic factors (vegetation and land cover) that contribute to environmental heterogeneity. The threshold approach implicitly considers abiotic and biotic factors as in the ‘critical load’ concept [57] applied in the Satilla watershed or the water framework directive [58] applied in the Mulde watershed.

Environmental stratification is suitable for all EIs in the current study (*global data coverage* and *easily applicable*). Data on PNV and thresholds does not

**Table 3.** Analyzing the strengths and weaknesses of the environmental stratification, PNV, and threshold approaches concerning their reliability, feasibility and relevance; the subcriteria have been collected from previous studies [19, 24, 30, 54–56]. Whether the subcriteria are fulfilled or not is indicated as following: green: +, yellow: +/–, red: –.

		Stratification	PNV	Thresholds
<b>Reliability</b>	Worldwide <b>consistent methodology</b> and datasets	• single global approach	• regional approaches • homogenous concept • heterogeneously applied	• regional approaches • heterogeneous concepts
	<b>Regional stakeholders and experts</b> involved	• no involvement due to global data use	• typically set up with regional experts	• partly set up with regional experts
	<b>Environmental heterogeneity within a region</b>	• spatially explicit	• spatially explicit	• not spatially explicit
	<b>Range of environmental factors</b> considered	• some abiotic factors (climate and topography)	• biotic and abiotic factors	• variable and implicit consideration (expert knowledge)
	<b>Global data coverage</b>	• single global dataset	• partial coverage of the world	• partial coverage of the world
<b>Feasibility</b>	<b>Easily applicable</b> to various environmental impacts	• all spatially explicit environmental impacts	• environmental impact modeling with spatial LU/LC datasets	• only available for few environmental impacts
	<b>Development effort</b> if no data available	• not applicable due to available global dataset	• regional data and experts needed	• regional experts and stakeholders needed
	Classification in <b>desirable and undesirable locations</b>	• not given	• classification given • hardly attainable environmental limits	• classification based on stakeholder consultation
<b>Relevance</b>	<b>Relative comparison</b> of globally distributed production regions	• possible	• possible	• possible

have a global coverage. The PNV approach allows comparing multiple EIs if a spatially explicit modeling approach based on LU/LC datasets exists. The threshold approach was limited due to the few EIs having regional thresholds (nutrient or sediment export rates and concentrations). Due to the broader range of input data or the need to establish expert panels, both PNV and threshold approaches require more local knowledge and effort than the stratification approach; this effect multiplies when comparing a (large) number of production regions worldwide (*development effort*).

In contrast to the threshold and PNV approaches, the stratification approach does not distinguish between *desirable and undesirable locations* of EIs. The stratification approach requires a reference case for comparison, e.g., another LU/LC class or production region. All approaches allow *relative comparison* of EIs of different LU/LC classes and between different land-use and production systems or by relating the EIs of current LU/LC with PNV or threshold values for EIs.

## 4. Discussion

### 4.1. Different approaches—consistent impact/sustainability assessments?

The stratification and PNV approaches identified cropland with the most harmful EIs, plantation forestry with more beneficial EIs, and forestry with

the most beneficial EIs. With decreasing land-use intensity, beneficial EIs (e.g., C storage) increased and harmful EIs (e.g., P export) decreased. This effect is consistent with the results of Brockerhoff *et al* [59]. Both approaches allowed the comparison of EIs of biomass production between and within production regions and with each other. Forestry in the Mulde watershed had the most beneficial EIs as similarly identified by PNV and stratification approaches. Practical consequences would be (i) to increase biomass sourcing from such desirable production locations. (ii) An analysis of factors distinguishing the production regions could generate knowledge how to raise beneficial EIs at undesirable locations of EIs. Exemplary factors distinguishing production regions are land management, e.g., differences in forestry practices, or governance instruments, e.g., certification schemes.

### 4.2. Congruence of approaches

We compared the congruence of the stratification, PNV, and threshold approaches with the desirable and undesirable locations (hot- and cold-spots) concept from ESS research [29, 60]. The stratification and PNV approaches did not select the same desirable and undesirable locations for EIs. This difference may result from (i) different land-use intensities or the degrees of modification of the natural environment and (ii) the set of EIs assessed. The stratification approach did not reliably reveal strong or weak human

modifications of the environment due to a missing benchmark (a natural or desired state of the environment). For example, the share of plantation forestry in the Satilla watershed at desirable EI locations was significantly larger for the stratification (80%-quantile) (39%) than for the PNV approach (1%) for a comparable area (1.1 versus 0.8% of each watershed). Stratification overestimated the beneficial EIs of plantation forestry compared with PNV. In the Rufiji basin, the PNV approach classified 13.8% and the stratification approach only 0.7% of the production region as desirable locations. In total, the stratification and PNV approaches will be more comparable if the quantiles to determine desirable and undesirable locations are set for individual EIs. The highest congruence is at different quantiles for the individual EIs (stratification approach). Equally, the LU/LC composition more strongly agrees between the PNV and stratification approaches if looking at individual EIs (figure 3). Therefore, maximizing a set of beneficial EIs or ESS using the hot- and cold-spots concept, e.g., Qiu and Turner [29], is unsuitable to identify desirable locations for multiple EIs simultaneously. The concept rather suits to assess whether EIs or ESS are balanced [26] in an ecosystem or watershed.

Thresholds are less useful. They may indicate (i) weak sustainability requirements or (ii) low EI/sustainable land-use activities. Thus, thresholds mainly tend to identify strongly undesirable locations of harmful EIs.

#### 4.3. Reliability, feasibility and relevance of the approaches

The PNV approach is most reliable, but less feasible and only partly relevant (table 3), e.g., due to the large range of considered abiotic and biotic factors of environmental heterogeneity. The stratification approach is more feasible for application, e.g., due to a globally available dataset. The threshold approach is more relevant due to a longer history of application by stakeholders and authorities (e.g., the 'critical load' concept in the US [56]).

We propose to improve the PNV approach by (i) assessing a consistent standard set of biotic and abiotic factors, and (ii) providing more transparency to reveal remaining inconsistencies between regional applications. A standardized protocol, listing the biotic and abiotic factors included and describing the modeling approach, would make the PNV approach much more transparent and reproducible.

The available global dataset is the major advantage of the stratification approach, which overcomes the heterogeneity of the expert-based and regionally specific PNV approach. The major advantage of PNV over stratification is the neutral benchmark or baseline independent from regional minima and maxima of EI values. For example, if an entire production region is managed at a high land-use intensity, the stratification

as best-in-class approach likely does not reveal sites, where EIs strongly differ from the natural state. Information about land-use intensity enhanced the reliability of the stratification approach. The PNV and stratification approaches would more likely identify similar (un)desirable locations for production regions with comparable land-use intensities (e.g., Satilla and Mulde watersheds). To evaluate the congruence between the approaches, we recommend to compare regions with either similar environmental conditions or land management, i.e., the two main parameters that affect the heterogeneity in EIs.

Environmental thresholds for beneficial EIs or ESS are hardly available, except applications based on the 'critical load' concept for environmental pollutants [56]. Major reasons may be that thresholds require (i) more effort to consider regional environmental conditions and (ii) different methodologies to develop them for individual EIs or ESS. Existing studies therefore use ESS capacities (potential ESS supply) with flexible thresholds that vary with regional environmental conditions, e.g., [61], but lack a universal approach to control environmental heterogeneity. ESS capacities are based on individual methodologies for each ESS and therefore are less suitable for an increasing number of studies assessing sets of EIs or ESS and their desirable and undesirable locations and interactions, e.g., Mouchet *et al* [62]. For example, critical sediment loads are set depending on soil type and topography or thresholds for nutrient inputs depending on the vulnerability of ecosystems (e.g., peatlands). Differing stakeholder preferences and regional regulations additionally lower the comparability of threshold approaches. Thus, thresholds are considered ill-suited as universal sustainability standard, but can reflect regional sustainability expectations/regulations of natural resource use. In a regional context, the threshold approach classifies land use in desirable and undesirable locations of EIs based on stakeholders' preferences and governmental regulations. As an approach with stakeholder consultation, it can provide clear and socially accepted information for regional environmental management.

There is a need to include stakeholders' preferences regarding the quantity and type of harmful EIs and beneficial EIs or ESS in general [63], including the context of biomass for bioenergy [50]. For example, certification schemes for bioenergy and for agricultural and forestry products partly prescribe environmental thresholds to ensure low levels of harmful EIs [50]. Although stratification and PNV approaches overcome some reliability and feasibility deficiencies of thresholds, governments or authorities rather use environmental thresholds. A natural state of the environment as a reference as given in the PNV approach may be difficult or impossible to obtain after long histories of land-use in many parts of the world, even if they were stopped [64]. Governments and authorities may reject this approach arguing



with lacking realism for current land-use and land management. However, PNV as benchmark reveals the degree of modification of the natural system through land use.

The comparable quantification of regional EIs of agricultural and forestry production between different countries (i) standardizes the magnitude of regional EIs and (ii) allows to discuss improvements or recompensation for severe environmental degradation in exporting regions (polluter-pays principle). The PNV and stratification approaches meet the need to evaluate regional EIs of past, current, and future policies, e.g., bioenergy or forest policies [6, 7], with a global impact on trade of agricultural or forestry production. Standardized EIs provide a basis to discuss stakeholders' preferences on environmental sustainability in exporting and importing regions worldwide by assessing the regional EIs in different land-use systems. Globally applicable and comparative environmental assessments provide the basis to discuss politically how to distribute the regional environmental burden of globally traded products. The approaches in this study complement tools such as the carbon footprint for global EIs and thereby considerably broaden environmental sustainability assessments in global discussions on the polluter-pays principle as requested by Laurent *et al* [65].

## 5. Conclusions

The stratification, PNV, and threshold approaches used in this study facilitate the comparison of EIs of biomass production systems between different world regions despite large environmental heterogeneity. However, comparative EI assessments are required for much broader sets of food, feed and bioenergy production systems and for a wider range of environmental conditions. We recommend to combine major environmental conditions and socio-economic factor (e.g., applied in [66]) to determine the desirable and undesirable levels of EIs for a larger set of land systems. These studies should also further investigate the consistency between the PNV and stratification approaches as more reliable and feasible approaches, respectively. Further insights on (in)consistencies between approaches would enable us to determine the conditions under which it is sufficient to apply the stratification approach with less environmental parameters and when it is necessary to use the more complex PNV approach.

Both PNV and stratification approaches may address and contribute to two major issues in research and governance:

### i. Globally comparable sustainability assessments

The PNV and stratification approaches enable comparisons of the regional EIs of alternative

production locations of a product by taking environmental heterogeneity into account. Existing approaches focus on the comparison of global EIs, such as GHG emissions and lack spatially explicit components necessary for water or soil quality impacts [7]. Both approaches support the link of global trade flows and remote EIs of agricultural and forestry products [67] through standardizing regional EIs. The comparability of regional and spatially explicit studies facilitates the identification of production locations with the lowest levels of harmful EIs based on combinations of production systems and environmental conditions for current (e.g., coffee [68]) or future (e.g., rice [69]) globally traded agricultural products.

### ii. Transferability of regional case study results

It is unfeasible to implement case studies of regional EIs or ESS for all (major) agricultural and forestry production systems and world regions. Thus, we propose (i) to assess EIs or ESS for specific commodities, e.g., major traded agricultural crops such as coffee [68] or forestry products, or land-system archetypes and (ii) to transfer the results to other world regions. Land-system archetypes are representative combinations of land-use intensity, environmental conditions, and socio-economic factors [66]. Both PNV and stratification approaches support transferability through the control of environmental heterogeneity. Beyond environmental heterogeneity, we propose to consider social heterogeneity (i.e., how stakeholders' preferences between regions (e.g., developed or developing countries) and societal groups differ [30, 70]) to enable a comprehensive sustainability assessment of global trade flows and their remote regional impacts.

## Acknowledgments

This work was partly funded by the EU FP7 project SECTOR (Production of Solid Sustainable Energy Carriers by Means of Torrefaction), grant agreement number 282826, and the Helmholtz Programme 'Terrestrial Environmental Research'. The authors would like to thank Christian Hoyer, Christian Schweitzer, and Charlotte Main for data preparation, Matthew McCartney (IWMI) for providing data on the hydrology of the Rufiji River, and Christian Hoyer, Carola Pätzold, and three reviewers for helpful comments on earlier versions of the manuscript.

## References

- [1] Demirbas A 2009 Political, economic and environmental impacts of biofuels: a review *Appl. Energy* **86** Supplement 1 S108–17
- [2] Lewandowski I and Faaij A P C 2006 Steps towards the development of a certification system for sustainable bio-energy trade *Biomass and Bioenergy* **30** 83–104
- [3] Ragauskas A J *et al* 2006 The path forward for biofuels and biomaterials *Science* **311** 484–9

- [4] WG III 2012 *IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation* (Cambridge: Cambridge University Press)
- [5] Ragauskas A J *et al* 2006 The path forward for biofuels and biomaterials *Science* **311** 484–9
- [6] Meyfroidt P, Lambin E F, Erb K-H and Hertel T W 2013 Globalization of land use: distant drivers of land change and geographic displacement of land use *Current Opinion in Environmental Sustainability* **5** 438–44
- [7] Liu J *et al* 2015 Systems integration for global sustainability *Science* **347**
- [8] Lambin E F and Meyfroidt P 2011 Global land use change, economic globalization, and the looming land scarcity *Proc. Natl. Acad. Sci. USA* **108** 3465–72
- [9] Lamers P, Hoefnagels R, Junginger M, Hamelinck C and Faaij A 2015 Global solid biomass trade for energy by 2020: an assessment of potential import streams and supply costs to North-West Europe under different sustainability constraints *GCB Bioenergy* **7** 618–34
- [10] European Commission 2009 On the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC 2009.04.23 (Directive 2009/28/EC) Off J EU pp 16–62
- [11] Cherubini F 2010 GHG balances of bioenergy systems—overview of key steps in the production chain and methodological concerns *Renew. Energy* **35** 1565–73
- [12] Wang M, Han J, Dunn J B, Cai H and Elgowainy A 2012 Well-to-wheels energy use and greenhouse gas emissions of ethanol from corn, sugarcane and cellulosic biomass for US use *Environ. Res. Lett.* **7** 045905
- [13] Polasky S, Tallis H and Reyers B 2015 Setting the bar: standards for ecosystem services *Proc. Natl. Acad. Sci. USA* **112** 7356–61
- [14] Creutzig F *et al* 2015 Bioenergy and climate change mitigation: an assessment *GCB Bioenergy* **7** 916–44
- [15] Dale V H, Lowrance R, Mulholland P and Robertson G P 2010 Bioenergy sustainability at the regional scale *Ecology and Society* **15** 23
- [16] Bustamante M *et al* 2014 Co-benefits, trade-offs, barriers and policies for greenhouse gas mitigation in the agriculture, forestry and other land use (AFOLU) sector *Glob. Change Biol.* **20** 3270–90
- [17] Meller L, van Vuuren D and Cabeza M 2015 Quantifying biodiversity impacts of climate change and bioenergy: the role of integrated global scenarios *Regional Environ. Change* **15** 961–71
- [18] Creutzig F, Corbera E, Bolwig S and Hunsberger C 2013 Integrating place-specific livelihood and equity outcomes into global assessments of bioenergy deployment *Environ. Res. Lett.* **8** 35047
- [19] Stein A and Kreft H 2014 Terminology and quantification of environmental heterogeneity in species-richness research *Biol. Rev.*
- [20] Stein A, Gerstner K and Kreft H 2014 Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales *Ecol. Lett.* **17** 866–80
- [21] Metzger M, Schröter D, Leemans R and Cramer W 2008 A spatially explicit and quantitative vulnerability assessment of ecosystem service change in Europe *Regional Environ. Change* **8** 91–107
- [22] Kienast F *et al* 2009 Assessing landscape functions with broad-scale environmental data: insights gained from a prototype development for Europe *Environmental Management* **44** 1099–120
- [23] Moldan B, Janoušková S and Hák T 2012 How to understand and measure environmental sustainability: indicators and targets *Ecological Indicators* **17** 4–13
- [24] Willemen L, Veldkamp A, Verburg P H, Hein L and Leemans R 2012 A multi-scale modelling approach for analysing landscape service dynamics *J. Environ. Manag.* **100** 86–95
- [25] Metzger M J, Rounsevell M D A, Acosta-Michlik L, Leemans R and Schröter D 2006 The vulnerability of ecosystem services to land use change *Agriculture, Ecosystems & Environment* **114** 69–85
- [26] Foley J A *et al* 2005 Global consequences of land use *Science* **309** 570–4
- [27] West P C *et al* 2010 Trading carbon for food: global comparison of carbon stocks versus crop yields on agricultural land *Proc. Natl. Acad. Sci. USA* **107** 19645–8
- [28] Crossman N D *et al* 2013 A blueprint for mapping and modelling ecosystem services *Ecosystem Services* **4** 4–14
- [29] Qiu J and Turner M G 2013 Spatial interactions among ecosystem services in an urbanizing agricultural watershed *Proc. Natl. Acad. Sci. USA* **110** 12149–54
- [30] Bishop K *et al* 2009 Nature as the ‘natural’ Goal for water management: a conversation *AMBIO* **38** 209–14
- [31] van Dam J, Junginger M and Faaij A P C 2010 From the global efforts on certification of bioenergy towards an integrated approach based on sustainable land use planning *Renewable and Sustainable Energy Rev.* **14** 2445–72
- [32] Maltsoğlu I and Khwaja Y 2010 *Bioenergy and Food Security—The BEFS Analysis for Tanzania* (Rome: FAO) 2071–0992 2071–0992
- [33] EC 2010 Report from the Commission to the Council and the European Parliament on sustainability requirements for the use of solid and gaseous biomass sources in electricity, heating and cooling (Com 2010/11) Publications Office
- [34] Kareiva P, Tallis H, Ricketts T H, Daily G C and Polasky S 2011 *Natural Capital: Theory and Practice of Mapping Ecosystem Services* (Oxford: Oxford University Press)
- [35] Nelson E *et al* 2009 Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales *Frontiers in Ecology and the Environment* **7** 4–11
- [36] Alkemade R, van Oorschot M, Miles L, Nellemann C, Bakkenes M and ten Brink B 2009 GLOBIO3: a framework to investigate options for reducing global terrestrial biodiversity loss *Ecosystems* **12** 374–90
- [37] Efromson R A *et al* 2012 Environmental indicators of biofuel sustainability: what about context? *Environ. Manag.* **51** 291–306
- [38] McBride A C *et al* 2011 Indicators to support environmental sustainability of bioenergy systems *Ecological Indicators* **11** 1277–89
- [39] GBEP Task Force 2011 *The Global Bioenergy Partnership Sustainability Indicators for Bioenergy* (Rome: GBEP Secretariat —FAO)
- [40] Metzger M J, Bunce R G H, Jongman R H G, Sayre R, Trabucco A and Zomer R 2013 A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring *Global Ecol. Biogeogr.* **22** 630–8
- [41] Landfire 2010 *Landfire 1.2.0—Environmental Site Potential Layer* (Washington, DC: Geological UDOTI, Survey)
- [42] LfULG 2011 Potentielle natürliche Vegetation (pnV) in Sachsen *Saxon State Office for Environment/Agriculture/Geology* (Freiberg)
- [43] van Breugel P *et al* 2012 *Potential Natural Vegetation Map of Eastern Africa: An interactive vegetation map for Ethiopia, Kenya, Malawi, Rwanda, Tanzania, Uganda and Zambia* version 1.1 (Copenhagen/Nairobi: Forest and Landscape (Denmark)/ICRAF)
- [44] EPD 2001 *Satilla River Basin Dissolved Oxygen TMDLs* (Atlanta: Georgia Department of Natural Resources)
- [45] EPD 2005 *Total Maximum Daily Load Evaluation for Headstall Creek in the Savannah River Basin for Sediment* (Atlanta: Georgia Department of Natural Resources)
- [46] SMUL, PLA, POH 2007 Bericht über die Entwicklung der Wasserbeschaffenheit in ausgewählten deutsch-tschechischen Grenzwasserläufen—Auswertung der abgestimmten physikalischen, chemischen und biologischen Kenngrößen von 1993 bis 2006. Sächsische Landesamt für Umwelt und Geologie, Povodí Labe, Povodí Ohře
- [47] Altman D G and Bland J M 1994 Diagnostic tests 1: sensitivity and specificity *BMJ* **308** 1552

- [48] R Development Core Team 2009 R: a language and environment for statistical computing (<http://r-project.org/>) (accessed 5 November 2013)
- [49] Kurtz J C, Jackson L E and Fisher W S 2001 Strategies for evaluating indicators based on guidelines from the environmental protection agency's office of research and development *Ecological Indicators* **1** 49–60
- [50] Meyer M A and Priess J A 2014 Indicators of bioenergy-related certification schemes—An analysis of the quality and comprehensiveness for assessing local/regional environmental impacts *Biomass and Bioenergy* **65** 151–69
- [51] Bockstaller C, Guichard L, Keichinger O, Girardin P, Galan M-B and Gaillard G 2009 Comparison of methods to assess the sustainability of agricultural systems: a review ed E Lichtfouse, M Navarrete, P Debaeke, S Véronique and C Alberola *Sustainable Agriculture* (Netherlands: Springer) pp 769–84
- [52] Buchholz T, Luzadis V A and Volk T A 2009 Sustainability criteria for bioenergy systems: results from an expert survey *J. Cleaner Production* **17** Supplement 1 S86–98
- [53] Niemeijer D and de Groot R S 2008 A conceptual framework for selecting environmental indicator sets *Ecological Indicators* **8** 14–25
- [54] Chiarucci A, Araújo M B, Decocq G, Beierkuhnlein C and Fernández-Palacios J M 2010 The concept of potential natural vegetation: an epitaph? *J. Vegetation Sci.* **21** 1172–8
- [55] Metzger M J *et al* 2013 Environmental stratifications as the basis for national, European and global ecological monitoring *Ecological Indicators* **33** 26–35
- [56] Groffman P *et al* 2006 Ecological thresholds: the key to successful environmental management or an important concept with no practical application? *Ecosystems* **9** 1–13
- [57] Pardo L H *et al* 2011 Effects of nitrogen deposition and empirical nitrogen critical loads for ecoregions of the United States *Ecological Applications* **21** 3049–82
- [58] Birk S *et al* 2012 Three hundred ways to assess Europe's surface waters: an almost complete overview of biological methods to implement the Water Framework Directive *Ecological Indicators* **18** 31–41
- [59] Brockerhoff E, Jactel H, Parrotta J, Quine C and Sayer J 2008 Plantation forests and biodiversity: oxymoron or opportunity? *Biodiversity and Conservation* **17** 925–51
- [60] Eigenbrod F *et al* 2010 The impact of proxy-based methods on mapping the distribution of ecosystem services *J. App. Ecol.* **47** 377–85
- [61] Schröter M, Barton D N, Remme R P and Hein L 2014 Accounting for capacity and flow of ecosystem services: a conceptual model and a case study for Telemark, Norway *Ecological Indicators* **36** 539–51
- [62] Mouchet M A *et al* 2014 An interdisciplinary methodological guide for quantifying associations between ecosystem services *Glob. Environ. Change* **28** 298–308
- [63] Wong C P, Jiang B, Kinzig A P, Lee K N and Ouyang Z 2015 Linking ecosystem characteristics to final ecosystem services for public policy *Ecol. Lett.* **18** 108–18
- [64] Somodi I, Molnár Z and Ewald J 2012 Towards a more transparent use of the potential natural vegetation concept—an answer to Chiarucci *et al* *J. Vegetation Sci.* **23** 590–5
- [65] Laurent A, Olsen S I and Hauschild M Z 2012 Limitations of carbon footprint as indicator of environmental sustainability *Environmental Science & Technology* **46** 4100–8
- [66] Václavík T, Lautenbach S, Kuemmerle T and Seppelt R 2013 Mapping global land system archetypes *Glob. Environ. Change* **23** 1637–47
- [67] Liu J *et al* 2013 Framing sustainability in a telecoupled World *Ecology and Society* **18**
- [68] De Beenhouwer M, Aerts R and Honnay O 2013 A global meta-analysis of the biodiversity and ecosystem service benefits of coffee and cacao agroforestry *Agriculture, Ecosystems & Environment* **175** 1–7
- [69] Lakkakula P, Dixon B L, Thomsen M R, Wailes E J and Danforth D M 2015 Global rice trade competitiveness: a shift-share analysis *Agricultural Economics* **46** 667–76
- [70] Daw T M *et al* 2015 Evaluating taboo trade-offs in ecosystems services and human well-being *Proc. Natl Acad. Sci. USA* **112** 6949–54