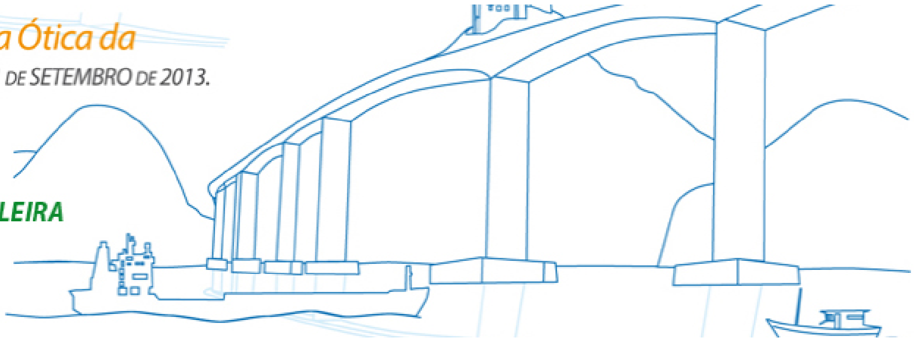
Inovação e Sustentabilidade sob a Ótica da Economia Ecológica. VITÓRIA/ES, 17 A 21 DE SETEMBRO DE 2013.
Hotel Vitória Grand Hall

X ENCONTRO DA SOCIEDADE BRASILEIRA DE ECONOMIA ECOLÓGICA



X ENCONTRO DA ECOECO
Setembro de 2013
Vitória - ES - Brasil

DRIVERS OF ENVIRONMENTAL IMPACT: A PROPOSAL FOR GLOBAL SCENARIO DESIGNING

Ely José de Mattos (PUCRS) - ely.mattos@pucrs.br
Professor do Programa de Pós-Graduação em Economia da PUCRS

Eduardo Ernesto Filippi (UFRGS) - edu_292000@yahoo.com.br
Professor do Programas de Pós-Graduação em Estudos Estratégicos Internacionais [PPGEEI] e em Desenvolvimento Rural [PGDR]

Drivers of environmental impact:
a proposal for global scenario designing

Eixo temático:
F) Crescimento e meio ambiente
- Valoração macroeconômica e quantificação de modelos sistêmicos;

Abstract: Drivers of environmental impact are commonly studied in the related literature through the IPAT and STIPAT models. The first is an accounting model that calculates the environmental impact caused by population, affluence, and technology. The second model is a stochastic approach that enables both statistics tests of significance of the drivers and the consideration of a larger set of drivers. These methodologies however, do not deal with scenarios of environmental impact because they are unable to take account of the level of the drivers in a nonlinear structure, i.e., different impacts according to the level of the variable. This paper presents a global Ordered Logistic Model that estimates the probability of four ordinal categories of environmental impact (four defined categories of Ecological Footprint). The results further the analysis of environmental impact because they offer an additional understanding of what to expect in terms of pressure on the ecosystem when the current level of a particular driver is about to increase or decrease.

Key-words: IPAT, STIRPAT, Ordered Logistic Model, environmental impact.

1. Introduction

The complex relationship between economic growth and the environment has been dealt with in specialized literature based on the Environmental Kuznets Curve (EKC) since the 1990s (Grossman and Krueger, 1991, Dinda, 2004). Nevertheless, significant efforts to understand the impacts of human activities on the environment date back to the 1970's (Ehrlich and Holdren, 1971, Commoner et al. 1972). A classic model of analysis, credited to Paul Ehrlich and John Holdren, is known in the literature as "IPAT". This model proposes that environmental impact results from the multiplicative relation between population, affluence and technology ($Impact = Population \cdot Afluence \cdot Technology$).

The IPAT model offers an accounting feature: from these three variables it is possible to directly estimate the environmental impact (providing there are adequate units of measure). However, this model is not able to address the statistical significance tests of the variables, nor can it take into consideration a larger set of variables. A model that could be considered more advanced in such aspects is the Stochastic Impacts by Regression on Population, Affluence, and

Technology (STIRPAT) (Dietz and Rosa, 1994, 1997). Although the model is theoretically based on the IPAT approach, it allows testing the statistical significance of the variables since it is a regression model. The STIRPAT estimates the elasticity of each variable in relation to the environmental impact and it is also able to incorporate other variables in addition to the three of the IPAT.

The study proposed by York et al. (2003), for instance, presents a STIRPAT model estimated for countries using the Ecological Footprint as a variable of environmental impact. The model was estimated based on more than 10 socioeconomic variables as drivers of environmental impact – among them GDP, population, and institutional aspects. The elasticity estimated by the authors, points towards the confirmation of GDP as an important driver of environmental impact as well as urbanization and area per capita.

The elasticity offered by the regular STIRPAT model, however, is constant through the entire range of the variables. This means that whatever the level of the driver, the environmental impact is proportionally the same. For instance, if we are observing the impact of population growth on the environment, it makes no difference if the original population size is small or large – the proportional impact would be the same. This is clearly a caveat for scenario designing, because it is reasonable to consider that the impact caused by a driver strongly depends on its level (Stern, 2004; Dinda, 2004; Cavlovic et al. 2000; Cole et al. 1997). Using this same example, it would be expected that the growth of a small population implies proportionally different impacts when compared to the growth of a large population.

It is important to note that the STIRPAT models have not been proposed to explore scenarios in the way described above. To be able to perform a scenario investigation it is necessary to draw on a model that is able to take into account levels of drivers in a nonlinear fashion. A potential alternative for this is a nonlinear probability model. This sort of model can measure the impact of independent variables on the probability of a specific outcome, whether it is binary, categorical or ordered. Although these models do not estimate the environmental impact itself as IPAT and STIRPAT do, they are able to present interesting scenario analyses in terms of probabilities.

In this context, **the main contribution of this paper is to propose a nonlinear probability model that can design scenarios based on the theoretical scope of IPAT and STIRPAT models**. Secondly, it aims to present some specific scenarios based on the model proposed in order to demonstrate the feasibility and characteristics of the methodology. To accomplish these objectives the paper is composed of three sections besides the introduction and conclusion. Section 2 deals with theoretical aspects regarding the IPAT and STIRPAT models. Section 3 presents the methodology, and Section 4 demonstrates the results and some illustrative scenarios.

2. The IPAT and STIRPAT models

Barry Commoner, Paul Ehrlich and John Holdren can be seen as precursors of a quantitative and systematic framework for studying the impacts of human activity on the environment (Ehrlich and Holdren, 1971; Commoner, 1972)[1]. Their thoughts have led to the proposal of the IPAT as an analytical model. The central idea of this model is an accounting identity which sustains the following: environmental impact (I) equals the product of population size (P), affluence (A), measured in terms of production per capita, and technology (T), measured as the impact per monetary unit of production. Algebraically we have:

$$I = P \cdot A \cdot T \tag{1}$$

Although the model is seminal and widely replicated[2], the IPAT presents some significant limitations. One of them – which had already been pointed out by Ehrlich and Holdren during debates with Commoner – is discussed in the well-known paper by Thomas Dietz and Eugene Rosa "Rethinking the environmental impacts of population, affluence and technology" (Dietz and Rosa, 1994). According to the authors, typical applications of the IPAT models are grounded on data for population and affluence. Technology, however, is indirectly obtained

---

[1] Chertow (2001) presents a historical perspective of IPAT.

[2] It is also replicated with some adjustments, for instance: Waggoner and Ausubel (2002), Schuzle (2002).

from the other variables: T = I/(P·A). So, since one has the three variables available (impact, population and affluence), the fourth one is automatically determined.

From an empirical point of view, this characteristic of the model can eventually underestimate the impacts of population and affluence – because technology is defined endogenously and might be incorporating factors other than just the technological aspect. Dietz and Rosa (1994) call attention to the observations of Ehrlich and Holdren on this matter indicating that "… calculations underestimate the effect of population on the environment by attributing to the T term changes that could more properly be allocated to P or A" (p. 10). So, on one hand, IPAT is useful as a model with an accounting characteristic, which can generate conclusions on the intensity of the environmental impact from population size and from the environmental efficiency of production. On the other hand, the model does not prove to be suitable for relative analysis where the motivation is to test the hypothesis of the significance of the human drivers of environmental impact, for instance.

It is exactly this limitation that forms the basis of the work of Dietz and Rosa (1994). The authors suggest that the IPAT should be reconsidered to establish a wider debate about the role played by population, economic growth and technology in terms of the environment. Two points are especially important. The first is that the model should be stochastic instead of an accounting exercise, to make it possible to test the hypothesis on the significance of the drivers. The second point is the explicit need to incorporate a larger number of variables to be tested and studied.

In this context, an important step forward in the formulation of models of environmental impact is the proposition of a model named STIRPAT (Stochastic Impacts by Regression on Population, Affluence, and Technology) (Dietz and Rosa, 1994, 1997). The model is formulated as follows:

$$I_i = aP_i^b A_i^c T_i^d e_i \tag{2}$$

It can firstly be observed that in this formula, the index "*i*" appears by the variables. These indexes indicate that the quantities vary across the observations[3]. The coefficients (*a*, *b*, *c* e *d*) are the terms that have to be estimated from the set of observations considered (countries, for example). For the sake of recognizing that this model theoretically derives from IPAT it has to be clear that the IPAT classic is obtained from this very formulation when we have the special case where $a=b=c=d=1$.

The STIRPAT's theoretical equation should be estimated in natural logarithms. In this case the model is presented as follows:

$$\ln(I) = a + b[\ln(P)] + c[\ln(A)] + \sum_{i=1}^{n} \beta_i[\ln(X_i)] + e \tag{3}$$

where a and e are logarithms of those same terms as in the multiplicative formulation of the model. The technology component (T) is incorporated into the error term in the same way it is executed when dealing with the traditional IPAT model. Considering the logarithm formulation the results are basically presented and discussed as elasticity.

As previously mentioned, an important characteristic of this model is the possibility of expansion of the drivers in the formulation. The technology is not frequently considered because of the absence of an adequate variable that can work properly as proxy[4]. However, variables that represent other dimensions such as institutions, culture, and geography, for instance, can be added to the model since they are conceptually consistent with the original multiplicative formulation. This is done through the X_i in the formula above. They represent all variables that the researcher would like to evaluate as a significant driver of environmental impact.

Studies carried out by researchers such as Sztukowski (2010), Shi (2003), York, et al. (2003), Knight and Rosa (2012), Wei et al. (2011), York and Rosa (2012) and others, have applied the STIRPAT model in different contexts and with different propositions. Sztukowski (2010), for example, analyzes the impact

[3] For the classic IPAT there is no need for the index because the accounting is supposed just for one observation or point in the time series.
[4] The paper of Wei (2011) presents a discussion on this matter.

of population, income per capita and climate oscillation on the emissions of CO_2 in several sectors for municipalities in USA. Knight and Rosa (2012), for their part, applied the STIRPAT model to investigate the impact of household dynamics on the consumption of fuelwood.

These applications, however, are especially dedicated to the identification of main drivers of environmental impact and relative importance when compared with each other. There are few references in these studies about prognostics or scenarios though. What would happen if the population grew by 10%? What would happen if the population grew by 10% when the population is small? What about a very large population? The answers to these questions are not statistically explored in these studies because they are not meant to design scenarios.

Returning to the STIRPAT formulation, it is worth noting that the estimated coefficients *b*, *c* and *d* are obviously constants, i.e., they represent fixed effects. This means that regardless of the level or frontiers of the driver, the proportional impact caused by them is constant through the entire range of the variable. However, there is plenty of evidence in the literature emphasizing that the effects of the drivers of environmental impact are level-dependent – for example, the recent “Planetary boundaries: exploring the safe operating space for humanity” by Rockström, et al. (2009) empirically demonstrates some irreversible turning points of sustainability.

Dietz and Rosa (1997) argue that the stochastic model would handle the nonproportional effects if the coefficients were replaced by more complex functions. However, since the coefficients are constant it represents a caveat for scenario studies. The STIRPAT models do represent an important advance in terms of the studies of impact of human activities on ecosystem, but they are not intended to offer scenario analysis where the nonlinearities of drives could be scrutinized.

We believe, though, that the theoretical background of the IPAT and STIRPAT models can work as a suitable starting point for an expansion of objectives in environmental impact analysis. Our proposal is therefore based on these studies but with an alternative target, that of scenario designing. The model applied is presented in the next section.

3. A nonlinear model for scenario designing: methodology, variables and the sample

For scenario designing of environmental impact we are assuming that three characteristics must be present: i) adequate theoretical background; ii) identification of a set of statistically significant drivers of environmental impact; and, iii) analysis of potential environmental impacts that consider the level of the drivers in nonlinear fashion. The first element is met by the IPAT/STIRPAT literature and also the studies that deal with the different determinants of pressure on nature. The second and third elements are fully attended to by the application of a nonlinear probability model named Ordered Logit Model (OLM). The following sections will present the statistical model, the variables of environmental impact, the potential drivers, and the sample.

3.1. The Ordered Logistic Model: a brief statistical presentation

An important aspect to be emphasized from the beginning of this methodology section is that the proposed model is not supposed to estimate directly the environmental impact as do both IPAT and STIRPAT. The model will estimate probabilities of impact – taking into consideration that we are proposing it to study scenarios, the estimated probabilities will meet this requirement. The OLM is a model generally used to estimate probability of outcomes that have more than two categories (where the classic logistic model could be enough). To be specific, the OLM is applied to situations where the ordinal rather the cardinal aspect matters. For example, a Likert scale in five categories applied to the evaluation of agreement on a subject: (1) strongly disagree, (2) disagree, (3) neutral, (4) agree, and (5) strongly agree. In this example, it is the order that really matters and the numbers themselves do not. In this case, the OLM would be able to estimate the probability of response for each category based on the explanatory variables considered. Most importantly: it does so using a logistic function (nonlinear).

The OLM is based on the same scope of the classic logistic regression. The logistic function presents the characteristic that is important for the scenario analysis proposed in this paper: marginal effects of drivers are less representative on both extremes of the distribution and more representative on intermediary portions. For instance, consider the hypothetical impact of population on the probability of a collapse on the food supply. It would be reasonable to consider that the marginal impact of population growth on this probability is minor when there is a small population or an already established large population, than when compared to an intermediary population. This is the reasoning of the logistic model and also to some extent, an underlying assumption of it.

The same applies to the OLM, although there are ordinal categories rather than a dichotomous outcome variable. The model has the same coefficients of impact for the explanatory variables but it presents different constants, which are cut points (or threshold parameters) to differentiate the categories. So the odds are proportional throughout the categories although at different levels. According to Wooldridge (2002) the OLM can be derived from a latent variable model. Assuming a latent variable is determined as follows:

$$y* = \mathbf{x}\boldsymbol{\beta} + e \tag{4}$$

where $\boldsymbol{\beta}$ is a vector of coefficients of the explanatory variables $\mathbf{x}$. There is no constant in this specification because there are different cut points which will play this role – the number of cut points is (J-1), J being the number of categories. Define the still unknown cut points (α_i) as follows:

$$\begin{aligned} &y = 0 \quad \text{if} \quad y* \leq \alpha_1 \\ &y = 1 \quad \text{if} \quad \alpha_1 < y* \leq \alpha_2 \\ &\vdots \\ &y = J \quad \text{if} \quad y* > \alpha_J \end{aligned} \tag{5}$$

Under the standard normal assumption for e[5], the conditional distributions of y given $\mathbf{x}$ can be computed for each response:

$$\begin{aligned}
&\mathrm{P}(y=0|\mathbf{x}) = \mathrm{P}(y* \leq \alpha_1|\mathbf{x}) = P(\mathbf{x}\boldsymbol{\beta} + e \leq \alpha_1|\mathbf{x}) = \varphi(\alpha_1 - \mathbf{x}\boldsymbol{\beta})\\
&\mathrm{P}(y=1|\mathbf{x}) = \mathrm{P}(\alpha_1 < y* \leq \alpha_2|\mathbf{x}) = \varphi(\alpha_2 - \mathbf{x}\boldsymbol{\beta}) - \varphi(\alpha_1 - \mathbf{x}\boldsymbol{\beta})\\
&\vdots\\
&\mathrm{P}(y=J-1|\mathbf{x}) = \mathrm{P}\left(\alpha_{J-1} < y* \leq \alpha_J\middle|\mathbf{x}\right) = \varphi\left(\alpha_J - \mathbf{x}\boldsymbol{\beta}\right) - \varphi(\alpha_{J-1} - \mathbf{x}\boldsymbol{\beta})\\
&\mathrm{P}(y=J|\mathbf{x}) = \mathrm{P}\left(y* > \alpha_J\middle|\mathbf{x}\right) = 1 - \varphi\left(\alpha_J - \mathbf{x}\boldsymbol{\beta}\right)
\end{aligned} \tag{6}$$

It is possible to identify that these probabilities sum to 1,0 and also that when there exists just one category (J=1) it turns out to be the regular binary model. The parameters α and β are estimated by MLE. The main specification test to be performed must guarantee the parallel regression assumption, i.e., the assumption that the angular coefficients are statistically the same for all categories – this test is proposed by Brant (1990). In summary, the OLM will estimate the probabilities of different categories of impact according to the drivers defined. Obviously the environmental impact has to be expressed as categories of impact rather than a continuous scale. The next section discusses this dependent variable as well as the use of it for the modeling process.

3.2. Ecological Footprint as a variable of environmental impact

Several studies focusing on environmental impact are based on the measurement of the inputs of human activity into the ecosystem: basically pollution of water and air. As already indicated in this paper, the majority of the studies analyzing drivers of environmental impact are concentrated on gas emissions. These measures however, do not effectively capture the transformation of nature caused by the human incursion (York et al. 2003). According to the authors, to accomplish a wider understanding of environmental impact it is advisable to count on more comprehensive or aggregate indicators.

---

[5] Wooldridge (2002) presents a Probit model based on a normal distribution function. The logistic model, however, works in the same way.

A fair candidate to such an index is the Ecological Footprint (EF) by Wackernagel and Rees (1996). This measure relies on the concept of carrying capacity. According to the "Calculation Methodology for the National Footprint Accounts" (Ewing et al. 2010), this measure aims to quantify the demand and supply of ecosystem goods and services in a static scope. The demand is the EF itself as the supply is measured by the productive biocapacity of the ecosystem. The balance between the two tells if the society is exploiting the natural resources more than it is able to offer or not.

The EF is in short, a measure of the pressure of human activity and resulting demands on the environment. It can be easily understood as the measure of the biologically productive area that is needed to sustain a certain society. The Global Hectare (Gha) is the standard unit measure considering different lands and its yield and equivalent factors: cropland, grazing, fishing, forest, built-in land, and forest for carbon sequester. Although the calculation is complex and data demanding, the final result is quite intuitive and also informative: the higher the EF, the greater the pressure of society on the environment.

As York et al. (2003), Wei et al. (2011) and other studies have developed STIRPAT models using the EF as the variable of environmental impact, we too propose to base our methodology on the same measure. However, the EF is a continuous variable (Gha) and the OLM is a model applied to ordered outcomes, as previously stated. In such a case it is mandatory to convert the EF into an ordered categorical variable. This process shall be executed considering that the ordered variable obtained has to present a clear and objective ordinal aspect. This means that the categories of EF could not be delimited by equidistant cuts on the scale – if processed this way, one obtains just a categorical variable with a sense of cardinality, which is not appropriate to the model proposed here.

To handle this question we adopted a categorization based on the quartiles of the distribution of the EF per capita of the countries composing the database[6]. In doing this we obtained a variable composed of four categories with different amplitudes of EF per capita, so the order is the element that defines the

[6] As a continuous variable the EF per capita of our sample presents a mean of 2,90 Gha and a standard-deviation of 1,88 Gha (coefficient of variation of 64,8%).

new variable rather than just the size of the EF determined *a priori*, for example[7]. The categories received an intuitive nomenclature according to their characteristics in terms of environmental impact: Low_EF (EF per capita ≤ 1,42), Medium-low_EF (1,42 < EF per capita ≤ 2,27), Medium-high_EF (2,27 < EF per capita ≤ 4,39), and High_EF (EF per capita > 4,39) – Appendix 1 contains the list of countries in each category. Through these categories it was possible to assign four ordered levels of environmental impact based on the EF per capita, which will be the dependent variable for the model. The data is by the Global Footprint Network and is for 2007 covering 128 counties[8].

3.3. The potential drivers of environmental impact

The EF portrays the level of pressure of societies on the environment, and this pressure comes basically from the consumption pattern. As can be seen in Appendix 1, the countries with a higher level of development (in the sense of the material/income approach) are in general also responsible for higher levels of EF. We also know that the consumption is strongly associated with a series of characteristics of the society. These characteristics can be considered the drivers of environmental impact and they come from different dimensions present in the theoretical scope.

As already discussed in this text, the IPAT/STIRPAT models are theoretical references for the study of the drivers of environmental impact. Considering that our study is particularly focused on the methodology of scenario designing, we opted for following the indications of the potential drivers that have already been identified in works such as York et al. (2003) and Dietz et al. (2007). We suggest the drivers are divided into two different categories: population and economics. For the *population* set of drivers, we proposed the initial following variables: total population (*pop_total*); percentage of total population that lives in urban areas (*pop_urb*); percentage of nondependent population, i.e., population

---

[7] The test for parallel regression can asseverate that the variable is adequate. As we will show later in this paper it is well constructed.

[8] This was the latest data available at the time of the estimations. We have not made use of imputation data techniques – only countries where all data was available were considered. This explains why we refer to 128 countries instead of the roughly 200 available at Global Footprint Network.

between 15 e 64 years old (*pop_15_64*); and, demographic density in terms of people per Km (*pop_denskm*). These variables aim to take account of the population pressure on the ecosystem based on size, interaction with the environment and geographical distribution.

For the *economic* drivers the initial variables were selected: GDP per capita (*gdp_pc*); percentage of GPD which does not come from services (*gdp_noserv*); and agricultural area (*land_agr*). As opposed to the economic pressure translated directly by the GDP per capita, the GDP of industry and agriculture (not services) and the agricultural area attempt to incorporate different aspects of the economic activity as drivers of environmental impact.

All the explanatory variables selected are theoretically suitable for the model, as mentioned by York et al. (2003). Technology in turn, is not present as it is not generally a component of the STIRPAT models. However, it is interesting to observe that the EF, in some sense, does incorporate technology itself: the changes in productivity of lands, for example, are sources of improvements of the measure of footprints – remembering that yield factors are used to compute the EF. In such a case, the population and economic drivers are in a certain way linked to technology.

All the variables for the 128 countries of the database come from the World Bank repository of statistics and refer to the year 2007. Table 1 shows the descriptive statistics of each driver. Except for the independent population (*pop_15_64*), the other variables present a considerable dispersion with higher coefficients of variation (CV=standard-deviation/mean).

Table 1 – Descriptive statistics of the explanatory variables

	Mean	Standar-	Minimum	Maximum	CV
pop_total	49.191.426	155.263.741	1.133.007	1.320.000.000	316%
pop_urb_per	55,0	21,3	12,6	97,3	39%
pop_15_64	62,8	6,4	48,8	72,3	10%
pop_denskm	109,3	143,5	1,7	1.105,9	131%
gdp_pc	10.543,9	16.119,1	164,2	82.294,2	153%
gdp_noserv	46,6	13,9	22,8	80,8	30%
land_agr	44,3	20,6	2,3	85,6	47%

Fonte: Elaborated by the authors. Data from World Bank repository.

The correlation between the drivers is also an important element for investigation before using them in a regression exercise. Box 1 shows the correlation matrix between all variables. There are just two correlations higher than 0,50. None of the estimated correlations invalidate or offer preliminary statistical evidence about multicolinearity at this point. The next section presents the estimated model, the required statistical tests, and also examples of scenario designing.

Box 1 – Correlation matrix between explanatory variables*

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
pop_total **(1)**	1,0000						
pop_urb_per **(2)**	-0,0803	1,0000					
pop_15_64 **(3)**	0,1189	0,5083	1,0000				
pop_denskm **(4)**	0,1766	-0,0932	0,1598	1,0000			
gdp_pc **(5)**	-0,0407	0,5251	0,4095	0,0838	1,0000		
gdp_noserv **(6)**	0,0352	-0,4240	-0,4330	-0,2399	-0,5066	1,0000	
land_agr **(7)**	0,0655	-0,0744	-0,0666	0,1842	-0,1477	-0,1171	1,0000

Source: elaborated by the authors.
* Pearson correlation.

4. Results: the model and the scenarios

The model has been estimated using adequate computational methods which follow all the statistical assumptions indicated in the previous sections. We have not applied any stepwise method because it is important to evaluate the significance of each driver, case by case. Having the EF per capita categorized in four levels of environmental impact as the dependent variable, the first estimated Ordered Logistic Model has counted on all the seven drivers previously selected for this exercise (Model I in Table 3). Two variables had coefficients not statistically significant: total population (*pop_total*) and percentage of GDP not from services (*gdp_noserv*)[9]. The former could be justified by the fact that the dependent variable is per capita which means that total population is no longer effective as an explanatory variable itself – the studies which present this variable as highly significant have aggregate dependent variables instead of per capita

[9] The lack of significance for these two drivers has also been demonstrated by including them as centered quadratic conversion and in combination with other variables. None of these options has worked properly.

ones. The latter, however, is probably due to iteration with the other economic variables – York et al. (2003) found the same lack of significance for this variable.

Table 3 – Estimated Ordered Logistic Model
Dependent variable: ordinal EF per capita (4 categories)

	Model I			**Model II**		
	Coefficients	SE	p-value	Coefficients	SE	p-value
pop_total	-0,0000	0,0000	0,384	-	-	-
pop_urb_per	0,2302	0,0127	0,071	0,0255	0,0124	0,041
pop_15_64	0,9997	0,0439	0,023	0,0906	0,0420	0,031
pop_denskm	-0,0044	0,0021	0,037	-0,0043	0,0020	0,037
gdp_pc($ 1.000)	0,0003	0,0000	0,000	0,0003	0,0000	0,000
gdp_noserv	-0,0121	0,0171	0,477	-	-	-
land_agr	0,0240	0,0104	0,021	0,0244	0,0103	0,018
Cut1	7,0926	2,8628		7,4262	2,4511	
Cut2	9,3804	2,9710		9,6761	2,5743	
Cut3	12,8019	3,0514		13,0938	2,6730	
N	128			128		
LR (Chi-sqr)	167,72 (p=0,000)			163,22 (p=0,000)		
Pseudo-R^2	0,4642			0,4600		
Mean VIF/Highest VIF*	1,44/1,77			1,38/1,71		

Source: estimated by the authors.
* Simulating linear regression.

The model without those two drivers remained well adjusted. Model II in Table 3 shows all the dependent variables significant at $p < 0.05$. The estimated cut points are displayed as "*Cut1*", "*Cut2*", and "*Cut3*". The table also shows that the Likelihood Ratio test is highly significant and there is no evident reason for concern over multicolinearity (VIF measures are well within acceptable limits).

Regarding the assessment of the model's specification it is vital to check the parallel regression assumption through the Brant test (Brant, 1990). The results of this test are presented in Table 4. As can be seen the test asseverates that the estimated OLM does not violate this important assumption.

Table 4 – Brant test (Model II)

	Chi^2	$p>Chi^2$*	Degrees of freedom
All	6,03	0,81	10
pop_urb_per	1,13	0,50	2
pop_15_64	2,10	0,35	2
pop_denskm	0,99	0,61	2
gdp_pc	2,29	0,32	2
land_agr	0,19	0,91	2

* The absence of significance guarantees the parallel regression assumption.
Source: results estimated by the authors.

The model presented in Table 3, therefore, is well adjusted and holds the statistical characteristics and properties required by the OLM. These results fulfill the second requirement we identified for scenario designing: the identification of statistically significant drivers of environmental impact. Starting from this point we are able to explore the results with the design of scenarios in mind. As a first step it is interesting to focus on the estimated coefficients for a brief comparison with the findings of other studies – though there are just a few works in the relevant literature about drivers of environmental impact that use the EF as the dependent variable.

The positive coefficients for urbanization (*pop_urb_per*) and nondependent population (*pop_15_64*) are corroborated by studies like York et al. (2003) and Wei (2011)[10]. It is worth pointing out that the impact of nondependent population on the environment is bigger than the urbanization *per se* – as was also found by York et al. (2003). York and Rosa (2012), on the other hand, found the urbanization and dependency ratio to be not significant when the dependent variables are pollutants like SO_2 and CO. The recent work of Wei (2011) suggests that differences in the estimations' results are due mainly to the differences in the specification of STIRPAT models.

The demographic density (*pop_denskm*) presented a negative impact – meaning the higher the population concentration, the smaller the potential impact. York et al. (2003) found the same sign for this variable[11]. This is rather a controversial result in terms of the literature. Ehrlich and Holdren (1971) pointed

[10] A study by Liddle and Lung (2010) presents a model for emissions of carbon dioxide and also energy consumption. The authors suggest that different age groups offer different directions of impact, and that older age groups exert negative impacts.
[11] Using the inverse of density the authors have found a positive sign.

out that it is a usual mistake "… to assume that population density (people per square mile) is the critical measure of overpopulation or underpopulation" (p. 1214). An important aspect that should be considered is the distribution of that population in relation to natural resources. Griffith (1981), for instance, gives us an idea as to how complicated this sort of variable is for planning cities – and consequently for the environment.

Finally, GDP per capita and the percentage of the area for agriculture are both consistently significant. GDP is without the doubt the variable that is almost universally important as a driver of environmental impact. The majority of the literature on Environmental Kuznets Curve as well as the studies of IPAT and STIRPAT models recognizes that the production/income per capita is directly connected to environmental pressures. Given that EF is a measure that strongly incorporates agriculture production, it is quite reasonable to understand that correlated variables (such as agricultural area) represent a specific pressure of human activity on the environment – for example, Cropper and Griffiths (1994) delineate some analysis on the deforestation resulting from population growth and its consequent increment in demand.

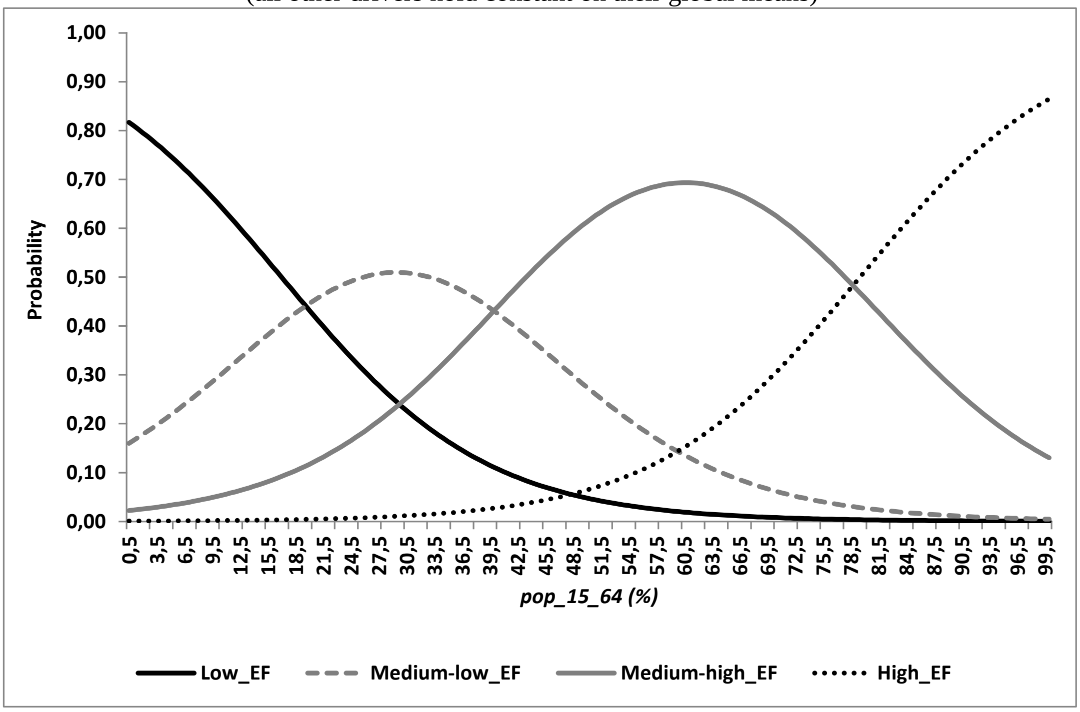
Since we have identified that the model is valid and that the basic results – significance and sign of the drivers – are in line with current literature it is possible to move toward the innovative aspect of this paper, that of scenario designing based on the OLM. The next section will construct some scenarios to demonstrate the methodology proposed and explore its potential.

4.1. Nonlinear scenarios of environmental impact

Through the coefficients displayed in Table 3 it is possible to estimate the probability of occurrence of each category of environmental impact given the values of the drivers (see equation 6). In other words, we are able to estimate probability for any of the four categories of environmental impact given the levels of drivers. Holding all other drivers constant on their global means it is feasible to analyze the trajectory of the probability of environmental impact when only one of them varies. This is the core of the scenario analysis proposed in this study: assuming a ceteris paribus condition we can vary the drivers of interest and check the probability of the environmental impact throughout the range of the variable. To demonstrate this procedure we present in detail the analysis for nondependent population (*pop_15_64*) and GDP per capita (*gdp_pc*).

In Figure 1 four curves can be identified, each of them representing the probability of occurrence of a specific level of environmental impact (category of EF) – it is important to bear in mind that at any point of the graph the four curves sum up to unit (100% of probability). So, we can see that a probability of low environmental impact (Low_EF) is higher than the other categories up to a level of approximately 20,0% of nondependent population. On the other extreme, the probability of high environmental impact (High_EF) detaches from the other categories when the nondependent population reaches a level around 80,0%.

Figure 1 – Scenario of probability of environmental impact
Target driver: nondependent population
(all other drivers hold constant on their global means)

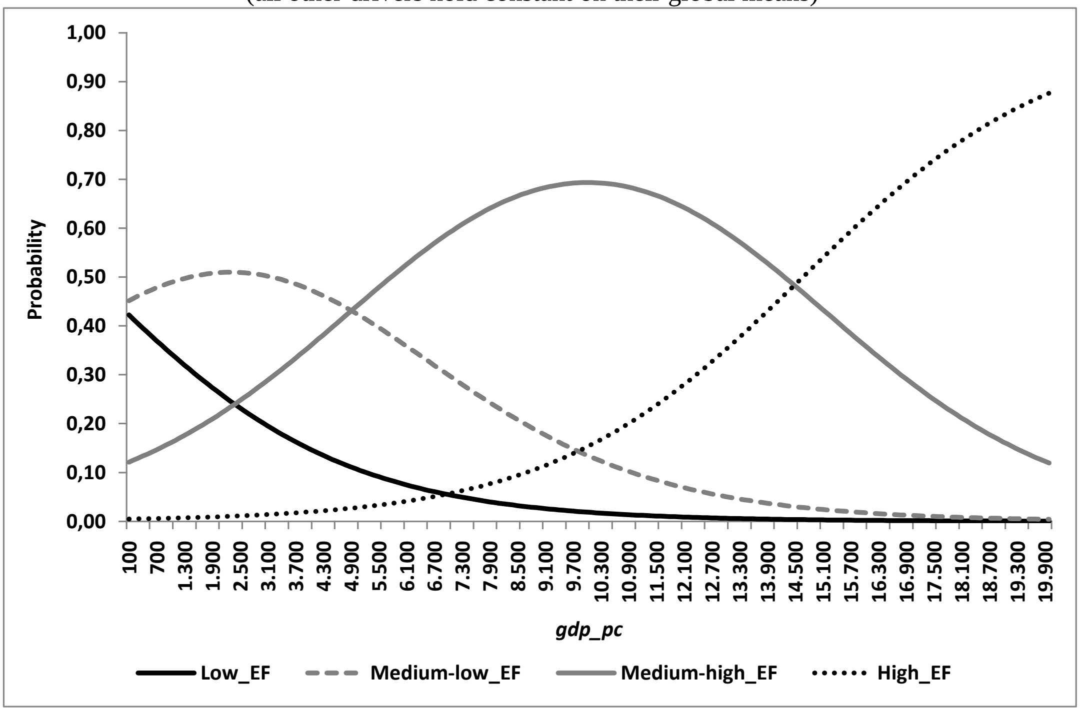


Source: elaborated by the authors.

We can observe that at the global mean level of nondependent population (62,8%), the probability of environmental impact is turning: the probability of a medium impact starts to decrease and the probability of high impact starts to grow faster. This can therefore be considered in some sense, a critical point for this variable. Ceteris paribus, if we foresee new increments in this variable this may result in more environmental pressure in the future. However, according to the estimations of the United Nations Department of Economic and Social Affairs, the population of this age group is supposed to stabilize (or at least to grow more slowly) by 2050.

If on one hand we have the prognostics of a decrease in the speed of growth of the nondependent population, on the other hand we must consider that there is not much optimism about the future of consumption (and consequently production) levels (see for instance the alerts contained in the Living Planet Report (WWF, 2010)). The specific impact of economic growth can be visualized through the same exercise performed above. Figure 2 depicts the estimated probabilities for the GDP per capita holding all the other drivers constant on their global means.

Figure 2 – Scenario of probability of environmental impact
Target driver: GDP per capita
(all other drivers hold constant on their global means)



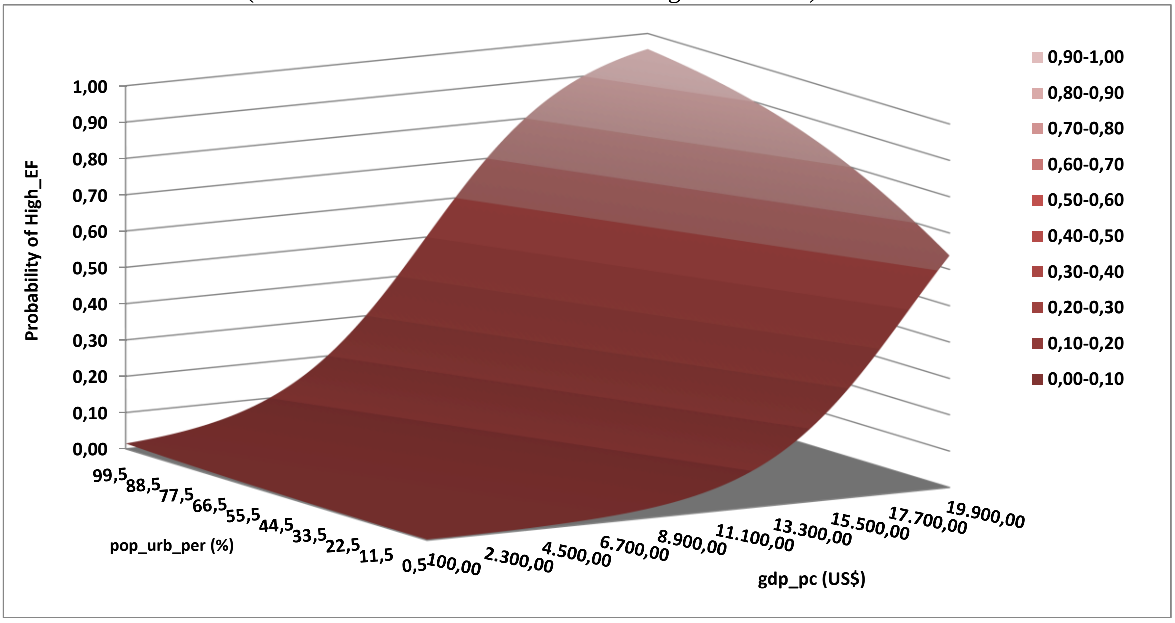
Source: elaborated by the authors.

The highest category of environmental impact (High_EF) is estimated to be dominant when the GDP per capita breaks the level of US$14.500,00 – our sample of 128 countries counts 26 that have exceeded this limit. The results inform that a constant growth of GDP per capita implies a direct increase on environmental pressure. Based on this result it is possible to check that the argument for the Environmental Kuznets Curve is not confirmed – as in the work of York et al. (2003). It would be necessary to eventually verify an inversion in the estimated probability of high environmental impacts to confirm this hypothesis. The lack of evidence of an Environmental Kuznets Curve may be associated with the essence of the environmental variable that is absolutely different from emissions of greenhouse gases, for instance.

It is also important to consider that higher levels of GDP per capita are present in countries that are, for the most part, developed – Sweden, Japan, Italy, USA, Germany, France, among others. All these nations register a high ecological footprint on earth basically because of their consumption pattern. In such cases it is feasible to comprehend that since other countries struggle for higher levels of economic performance the impact on the ecosystems could be worsened – in our model this conclusion is illustrated by the increasing probability of high EF.

Because of lack of space in this paper for a detailed analysis of the three other drivers, we have placed the graphs for these scenarios in Appendix 2 – all of them preserve the same logic applied to the previous one. An interesting additional resource that might emerge from the methodology developed in this study is the possibility of dealing with a scenario where two drivers can be visualized at the same time. Figure 3 displays a graph that illustrates a scenario where the percentage of population in urban areas (*pop_urb_per*) and the GDP per capita (*gdp_pc*) are varying at the same time – all other drivers hold constant on their global means as usual.

Figure 3 – Scenario of probability of environmental impact
Target drivers: GDP per capita and percentage of population in urban areas
(all other drivers hold constant on their global means)



Source: elaborated by the authors.

As we are plotting a 3D graph it would not be appropriate to have one surface for each category. So we only plotted a surface for the probability of the highest level of environmental impact (High_EF). Since both drivers are positive, what can be seen is a surface that demonstrates an increasing probability of high environmental impact as both variables increase. The interesting part of this scenario, however, is an observation of the shape of the surface that illustrates the combined impact of the drivers on the probability of High_EF: if urbanization is maintained at a low level (around 10,0%), even considering a higher GDP per capita the probability does not exceed 60,0%. But when urbanization is paired with economic growth, the probability of the highest category of environmental impact crosses the 90,0% mark. The same analytical scheme can be performed for the combination of all other drivers.

5. Conclusions

In this paper we have presented a methodology for scenarios of environmental impact. We achieved this task following the line proposed by the STIRPAT models. These models, primarily developed in the 1990s, are quite important in the process of understanding the impact of human activity on ecosystems. The main focus of such literature, however, is the static measures of elasticity of impact. This means that prognostics are not accounted for in this sort of tool because the estimated elasticity is independent on the level of the drivers.

The main focus of our methodology is a nonlinear probability model that tackles this limitation by considering the level of the explanatory variables. The Ordered Logistic Model estimates the probability of categories of environmental impact. Although the results represent probability of impact rather than impacts themselves, they do provide some insight into what to expect in terms of pressure on nature if the driver is increasing or decreasing. We have illustrated this feature presenting some scenarios for a couple of variables. Analysis enabled by the OLM is wider in the sense that it enables the comprehension of a new dimension for environmental impact: prognostics. However, the absence of an estimation of the direct impact may require the combination of this methodology with the previous one (IPAT/STIRPAT).

References

BRANT, R. Assessing proportionality in the proportional odds model for ordered logistic regression. **Biometrics**, Arlington, v. 46, n. 4, p. 1171-78, dec. 1990.

CAVLOVIC, T. A. et al. A meta-analysis of environmental Kuznets curve studies. **Agricultural and Resource Economics Review,** v. 29, n. 1, p. 32-42, 2000.

CHERTOW, M. R. The IPAT equation and its variants: changing views of technology and environmental impact. **Journal of Industrial Ecology**, v. 4, n. 4, p. 13-29, 2001.

COLE, T. A.; RAYNER, A. J.; BATES, J. M. The environmental Kuznets curve: an empirical analysis. **Environment and Development Economics**, v. 2, p. 401-16, 1997.

COMMONER, B. The environmental cost of economic growth. In: RIDKER, R. G. (Ed.). **Population, Resources and the Environment.** Washington, DC: U.S. Government Printing Office, 1972. p. 339-63.

COOPER, M.; GRIFFITHS, C. The interaction of population growth and environmental quality. The American Economic Review, v. 84. N. 2, Papers and Proceedings of the Hundred and Sixth Annual Meeting of the American Economic Association, 1994.

DIETZ, T.; ROSA, E. A. Effects of population and affluence on CO2 emissions. **Proceedings of the National Academy of Sciences of the USA**, v. 94, n. 1, p. 175-9, 7 jan. 1997.

______. A. Rethinking the environmental impacts of population, affluence and technology. **Human Ecology Review**, v. 1, p. 277-300, 1994.

DIETZ, T.; ROSA, E. A.; YORK, R. Driving the human ecological footprint. **Frontiers in Ecology and the Environment**, Washington DC, v. 5, n. 1, p. 13-8, 2007.
DINDA, S. Environmental Kuznets curve hypothesis: a survey. **Ecological Economics**, v.49, p. 431-44, 2004.
EHRLICH, P. R.; HOLDREN, J. P. Impact of population growth. **Science**, v. 171, n. 3977, p. 1212-17, 26 mar. 1971.
EWING, B.; REED, A.; GALLI, A.; KITZES, T.; WACKERNAGEL, M. **Calculation methodology for the national footprint accounts, 2010 edition**. Oakland: Global Footprint Network, 2010.
GROSSMAN, G. M.; KRUEGER, A. B. Economic growth and the environment. **The Quarterly Journal of Economics,** Cambridge, v. 110, n. 2, p. 353-77, may. 1995.
GRIFFITH, D. A. Modelling urban population density in a multi-centered city. Journal of Urban Economics, 9, p. 298-300, 1981.
KNIGHT, K. W.; ROSA, E. A. Household dynamics and fuelwood consumption in developing countries: a cross-national analysis. **Population and Environment**, n. 33, p. 365-378, 2012.
LIDDLE, B.; LUNG, S. Age-structure, urbanization, and climate change in developed countries: revisiting STIRPAT for disaggregated population and consumption-related environmental impacts. **Population and Environment**, n. 31, p. 317-343, 2010.
ROCKSTRÖM, J. et al. Planetary boundaries: exploring the safe operating space for humanity. **Ecology and Society**, v. 14, n. 2, art. 32, 2009.
SCHULZE, P.C. ‘I=PBAT’. **Ecological Economics,** 40, p.149–150, 2002.
SHI, A. The impact of population pressure on global carbon dioxide emissions, 1975-1996: evidence from pooled cross-country data. **Ecological Economics**, v. 44, p. 29-42, 2003.
STERN, D. I. The rise and fall of the environmental Kuznets curve. **World Development**, v. 32, n. 8, p. 1419-39, 2004.
SZUTUKOWSKI, J. **A STIRPAT model of sectoral CO2 emissions at the county scale**. Master Thesis, Colorado State University, Colorado, 2010.
WACKERNAGEL, M.; REES, W. **Our ecological footprint**. Gabriola Island, BC and Stony Creek, CT: New Society Publishers, 1996.
WAGGONER, P. E.; AUSUBEL J. H. A framework for sustainability science: a renovated IPAT identity. Proc National Academy of Science, v. 99 n. 12, 2002.
WEI, T. What STIRPAT tells about effects of population and affluence on environment? **Ecological Economics**, v. 72, p. 70-4, 2011.
WEI, T.; ZHONG, X.; LIU, SHAOQUAN. Analysis of major driving forces of ecological footprint based on STIRPAT model and RR method: a case of Sichuan Province, Southwest China. Journal of Mountain Science, v. 8, n. 4, p. 611-618. 2011.
WOOLDRIDGE, J. M. **Econometric analysis of cross section and panel data**. Cambridge: The MIT Press, 2002.
World Wild Foundation International (WWF). **Living Planet Report 2010**. Gland, Switzerland, 2010.
YORK, R.; ROSA, E. A. Choking on modernity: a human ecology of air pollution. Social Problems, v. 59, n. 2, p. 282-300, 2012.

______. Footprints on the Earth: the environmental consequences of modernity. **American Sociological Review**, Washington DC, v. 68, n. 2, p. 279-300, abr. 2003.

Appendix 1

Countries by category of Ecological Footprint per capita

Low_EF *(PE ≤ 1,42)*	**Medium-low_EF** *(1,42 < PE ≤ 2,27)*	**Medium-high_EF** *(2,27 < PE ≤ 4,39)*	**High_EF** *(PE > 4,39)*
Afghanistan	Albania	Argentina	Australia
Angola	Algeria	Belarus	Austria
Bangladesh	Armenia	Bolivia	Belgium
Cambodia	Azerbaijan	Bosnia and Herzegovina	Canada
Cameroon	Chad	Botswana	Czech Republic
Central African Rep.	China	Brazil	Denmark
Congo	Colombia	Bulgaria	Estonia
Congo, Dem. Rep.	Dominican Rep.	Chile	Finland
Côte d'Ivoire	Ecuador	Costa Rica	France
Eritrea	Egypt	Croatia	Germany
Ethiopia	El Salvador	Gambia	Ireland
Gabon	Georgia	Hungary	Italy
India	Ghana	Iran, Islamic Republic of	Japan
Indonesia	Guatemala	Lebanon	Kazakhstan
Kenya	Guinea	Libyan Arab Jamahiriya	Korea, Rep. of
Kyrgyzstan	Honduras	Mauritania	Latvia
Lao People's Dem. Rep.	Jamaica	Mauritius	Lithuania
Lesotho	Jordan	Mexico	Macedonia TFYR
Liberia	Madagascar	Nepal	Malaysia
Malawi	Mali	Panama	Mongolia
Moldova	Namibia	Paraguay	Netherlands
Morocco	Nicaragua	Poland	Norway
Mozambique	Nigeria	Romania	Portugal
Pakistan	Papua New Guinea	Serbia	Russian Federation
Philippines	Peru	Slovakia	Saudi Arabia
Rwanda	Sudan	South Africa	Slovenia
Senegal	Swaziland	Thailand	Spain
Sierra Leone	Syrian Arab Republic	Trinidad and Tobago	Sweden
Sri Lanka	Tunisia	Turkey	Switzerland
Tajikistan	Uganda	Turkmenistan	United Kingdom
Tanzania, United Rep.	Uzbekistan	Ukraine	USA
Viet Nam		Venezuela, Boliv. Rep.	Uruguay
Zambia			

Source: elaborated by the authors based on data of the Ecological Footprint..