

ORIGINAL ARTICLE

Open Access

Biofuels, cropland expansion, and the extensive margin

Farzad Taheripour^{1*}, Qianlai Zhuang², Wallace E Tyner¹ and Xiaoliang Lu²

Abstract

Background: Recently, several papers have assessed land use consequences of biofuel expansion. In the absence of empirical evidence, these papers assigned subjective values to extensive margin (productivity of new croplands over productivity of existing croplands).

Methods: This paper fills the gap in this area and provides a new data set which estimates land productivity at $0.5^\circ \times 0.5^\circ$ (longitude \times latitude) grid-cell level using a process-based biogeochemistry model, the terrestrial ecosystem model (TEM) calibrated for a C4 crop.

Results: The results obtained from the TEM can be used in connection with economic models which are designed to assess land use changes induced by economic factors. To show a real application, a set of regional extensive margins are calculated based on the new data set. The calculated regional extensive margins are then introduced in a computable general equilibrium (CGE) economic model which has been frequently used to assess the land use implications of ethanol production. Finally, land use changes due to US ethanol production are examined using the augmented CGE model with the new extensive margins.

Conclusions: The approach developed here provides estimates of extensive margins disaggregated by the country and agroecological zone, replacing the earlier assumption of a globally uniform value. Using these new parameter values, the estimation of land required for ethanol production is 25% lower than earlier published results.

Keywords: Extensive margin, Cropland productivity, Indirect land use change, Biofuels, Computable general equilibrium, GTAP-BIO, Terrestrial ecosystem model

Background

The increased demand for bioenergy in the USA, European Union, and other regions across the world has raised debates on land use consequences of biofuel production. In response to these debates, several attempts have been made to evaluate land use impacts of biofuels and their consequent land use emissions [1-9]. These evaluations are obtained from partial and general equilibrium models representing different modeling structures and assumptions. Despite the differences, models used in this area, explicitly or implicitly, make assumptions on the elasticities of cropland yield and land supply with respect to crop prices. These two elasticities are often referred to as *intensive* and *extensive* margins, respectively.

Several studies have estimated intensive yield elasticities. Kenney and Hertel [10] have reviewed many of these papers. Huang and Khanna [11] and Berry [12] represent the most recent studies in this area. However, there is almost no reliable work on extensive margin. In the absence of empirical evidence, economic modelers often rely on their personal conjecture and assign subjective values to the extensive margins in their analyses. For example, Searchinger et al. [5]^a and Hertel et al. [6]^b, whose articles are well known and frequently cited in this area, made simple assumptions in their land use assessments. The former article assumed that intensive margin due to ethanol production will wash out its extensive margin. The latter article simply assumed that on average, productivity of new lands converted to crop production is about two-thirds of the average productivity of existing croplands in each region.

* Correspondence: tfarzad@purdue.edu

¹Department of Agricultural Economics, Purdue University, 403 West State St, West Lafayette, IN 47907-2056, USA

Full list of author information is available at the end of the article

This paper fills a gap in this area and provides a data set which estimates land productivity at $0.5^\circ \times 0.5^\circ$ (longitude \times latitude) grid-cell level. To generate this data set, a process-based biogeochemistry model, the terrestrial ecosystem model (TEM) [13-15] along with spatially referenced information on climate, elevation, soils, and vegetation land use data are used. This data set can be used in connection with economic partial equilibrium and computable general equilibrium models which are used to assess land use changes induced by economic factors. To show a real application, the obtained data set is used to generate a set of parameters which estimates extensive margins for 18 agroecological zones (AEZ) at a regional level (Additional file 1 provides the list of regions and their members). Finally, the Global Trade Analysis Project with biofuels (GTAP-BIO) model used by Hertel et al. [6] is augmented with the new extensive margins and used to reassess land use implications due to the US ethanol program estimated in the study of Hertel et al. [6]. Results indicate that new extensive margins reduce the estimated land requirement for ethanol production by 25%.

The intensive and extensive margins play important roles in land use assessments of biofuel production. This paper concentrates on the latter concept. Crop producers may expand their croplands to produce more crops in response to a higher demand for crops, for example, due to an increase in biofuel production. The size of expansion in cropland depends on the productivity of new cropland. A simple example can be used to show the role of extensive margin in cropland expansion. Consider a closed economy which plans to produce 10 billion liters of ethanol. To meet this target, given the existing technology, about 25 million metric tons of corn are needed. Assume that this economy can reduce its current corn consumption by 12.5 million metric tons through reductions in corn consumption in food industries, replacement of corn used in livestock industry with ethanol by-product produced in conjunction with ethanol production, and all other factors which affect corn consumption. Hence, this economy needs to produce 12.5 million metric tons of additional corn. Assume that in this economy, on average, corn yield is about 10 metric tons/ha on the existing cropland and that about 40% of the additional demand for corn can be produced on the existing cropland due to conversion of cropland from other crops to corn production. In addition, assume no yield improvement.

Under these assumptions, this economy needs to expand its cropland to produce about 7.5 million metric tons of the remaining corn demand. The magnitude of cropland expansion depends on the productivity of new lands. If the average productivity of new lands is the same as the existing corn yield (i.e., 10 metric tons/ha), then this economy needs to convert about 0.75 million ha of its natural land (forest and pasture land) to

cropland. If the average productivity of new lands is smaller than the existing yield, say 7.5 metric tons/ha, then the expansion in cropland will be 1 million ha. Finally, if the average productivity of new lands is higher than the existing yield, say 12.5 metric tons/ha, then the expansion in cropland will be 0.6 million ha.

Consider now four terms. The first term is *marginal land*. This term refers to newly converted lands from its current use (forest or pasture land) to crop production. The second term is *average productivity of marginal land*, Y_m , which refers to the average productivity of newly converted lands to cropland and measured in terms of crops produced on new lands. The third term is *average productivity of existing croplands*, Y_c , measured in terms of crops produced on the existing croplands. The last term is the *extensive margin*, ETA, which measures the ratio of Y_m/Y_c . The example presented above assumed three cases for Y_m , and they are 7.5, 10, and 12.5 metric tons/ha. From these figures, three possible values of 0.75, 1.0, and 1.25 can be assigned to ETA. In general, Y_c can be obtained from actual observations, but the true values for Y_m and ETA are unknown in each region. Given an estimate for ETA, one can estimate the average productivity of marginal land for a given Y_c .

The GTAP-BIO model used by Hertel et al. divides the global managed land cover into three categories of managed forest, pasture, and cropland; traces land conversions among these land types due to changes in economic factors (for example, higher demand for biofuels) at a country/regional level by AEZ; and assigns value added to these land categories. The GTAP-BIO model excludes unmanaged forest from economic activity and does not take into account possible conversion of unmanaged forest to other alternative uses. As mentioned above, this model assumes a value of $ETA = 2/3$ regardless of the location of land conversion. This means the model assumes that on average, productivity of marginal land is two-thirds of the existing cropland all across the world. This value is identical to the global extensive margin calculated by Field, Campbell, and Lobell [16] for abandoned croplands. These authors evaluated average yields for different types of lands at a global scale in terms of net primary product (NPP). Their results indicate that the average yield on global available *abandoned* lands (cultivated and used prior to 2000 but were not used this year) was two-thirds of the existing cropland. This definition for land abandonment is comparable to the definition of land category 'cropland pasture' in the USA. However, as mentioned earlier, in GTAP, new lands are indeed converted from their current uses (managed forest and pasture) to crop production, and they are not necessarily low-quality abandoned land. Hence, we need an alternative estimate of

the productivity of marginal lands. This paper fills the gap in this area and estimates extensive margins based on estimated NPP values using the TEM.

In general, there is a positive and strong relationship between NPP and crop yields in an ecosystem, other factors being constant. Based on this relationship, several papers have evaluated NPP using harvested yield data, and some articles have estimated crop yields or potential biomass yield using NPP estimates [16-22]. This paper follows the literature in this area and defines a method to assess land productivity in terms of NPP of a C4 crop. It then uses the C4 NPP yields and calculates extensive margins at a country/AEZ level. While this paper concentrates on the estimated extensive margins using a C4 crop, the same practice is also repeated using a C3 crop. Since the resulting ETA values are ratios of productivities of uncultivated land versus cultivated land, we find that the estimated extensive margins for both C4 and C3 crops are close to each other.

Methods

Extensive margin and productivity measurement

NPP, the difference of gross primary production via plant photosynthesis and autotrophic respiration in a given time period (e.g., 1 year), indicates the rate of storage of biomass in a given time period. The magnitude of NPP per unit of land varies with the soil quality, climate condition, type of land cover, and by species. For a given soil quality and climate condition, the magnitude of the NPP changes when the land cover changes from one type to another. Our goal in this paper is to estimate extensive margins by comparing the productivities of existing croplands and their natural land counterparts regardless of their current vegetation covers.

Measuring productivities of the existing cropland and marginal land in terms of NPP of their current land cover is misleading. For example, consider a hectare of forest land with an annual NPP of 6 million grams of carbon per hectare per year ($\text{Mg C ha}^{-1} \text{ year}^{-1}$) and a hectare of cropland with NPP of 3 $\text{Mg C ha}^{-1} \text{ year}^{-1}$. In this case, the NPP of forest land is twice the NPP of cropland. However, it does not mean that if we convert the forest to cropland, then we can get productivity twice that of the existing cropland. The issue is that the existing land covers of these two land parcels are different. Thus, the productivities of croplands and their natural land counterparts should be evaluated regardless of their current vegetation covers. To be neutral with respect to the current land covers, as explained later on in this paper, we defined a C4 crop. The parameters of the TEM are then calibrated according to the characteristics of the C4 crop. Finally, the TEM is used to assess the NPP for the C4 crop (henceforth, NPP^{C4}) for all grid cells all across the global land surface. Given the calculated NPP^{C4}

for all grid cells, the following approach is used to estimate the extensive margin productivity measured as the ratio of average productivity of the natural land to the average productivity of cropland at a country/AEZ level. At this point, it is important to note that the GTAP land use module considers average productivity of cropland at the AEZ level and does not distinguish variations in productivity within an AEZ. Also, this model uses a constant elasticity transformation function to handle the supply side of market for land. Hence, it ignores important factors such as accessibility of land, existing infrastructure, and land conversion cost.

Consider now an AEZ in region r with n types of original land cover classifications and assume a portion of each type of land classification is already converted to cropland. In this AEZ, each parcel of land is recognized by its original and current land covers. For example, a parcel of land with index $m1$ indicates that this land parcel belongs to the land cover category 1 and currently is covered by its original land cover, or a parcel of land with index of $c1$ indicates that this parcel of land is originally in the land cover category 1 but currently is under crop production. The productivity of each parcel of land is measured in terms of NPP^{C4} regardless of the original and current covers. For example, $\text{NPP}_{c1}^{\text{C4}}$ represents the calculated NPP^{C4} for a parcel of land with original land cover type 1, which is currently under crop production, or $\text{NPP}_{m1}^{\text{C4}}$ represents NPP^{C4} for a parcel of land with original land type 1 which is country covered by its original vegetation. Given the estimated NPP^{C4} for all parcels of a grid cell, the extensive margin for that grid can be calculated using the following relationship:

$$ETA = \frac{A_{m1}\text{NPP}_{m1}^{\text{C4}} + A_{m2}\text{NPP}_{m2}^{\text{C4}} + \dots + A_{mn}\text{NPP}_{mn}^{\text{C4}}}{A_{m1} + A_{m2} + \dots + A_{mn}} \quad (1)$$

$$\frac{A_{c1}\text{NPP}_{c1}^{\text{C4}} + A_{c2}\text{NPP}_{c2}^{\text{C4}} + \dots + A_{cn}\text{NPP}_{cn}^{\text{C4}}}{A_{c1} + A_{c2} + \dots + A_{cn}}$$

In this equation, A_{ci} represents the area of land types i which is already converted to crop production and A_{mi} represents the area of land types i which is still in its original classification. $A_i = A_{ci} + A_{mi}$ represents the total area of land type 1 and $\sum_i^n A_i$ measures the total area of AEZ. In this equation, the numerator of the right-hand side is the weighted average of the productivity of existing natural lands and the denominator measures the weighted average productivity of the existing cropland in the AEZ both measured in terms of NPP^{C4} .

To make sure that the estimated ETA ratios do not vary with the type of simulated crop, we repeated the whole process explained above for a C3 crop. As presented later in this paper, we find that the calculated NPP^{C3} values and NPP^{C4} values are very different, but the estimated ETA ratios obtained for the C3 crop are close to the results obtained for the C4 crop. This shows

that the estimated ETA ratios are independent from the choice of crop and their productivities.

The TEM

We use a process-based biogeochemistry model, the TEM [13], to estimate NPP for each $0.5^\circ \times 0.5^\circ$ (longitude \times latitude) of the global terrestrial ecosystems. TEM uses spatially referenced information on climate, elevation, soils, and vegetation to make monthly estimates of C and N fluxes and pool sizes of the terrestrial biosphere. In TEM, the net ecosystem exchange of CO_2 between the land ecosystems and atmosphere is calculated as the difference between the uptake of atmospheric CO_2 associated with photosynthesis (i.e., gross primary production (GPP)) and the release of CO_2 through autotrophic respiration (R_A) and heterotrophic respiration (R_H) associated with decomposition of organic matter. The fluxes GPP, R_A , and R_H are influenced by changes in atmospheric CO_2 , climate variability and change, and the freeze-thaw status of the soil. Figure 1 represents this model and its major components.

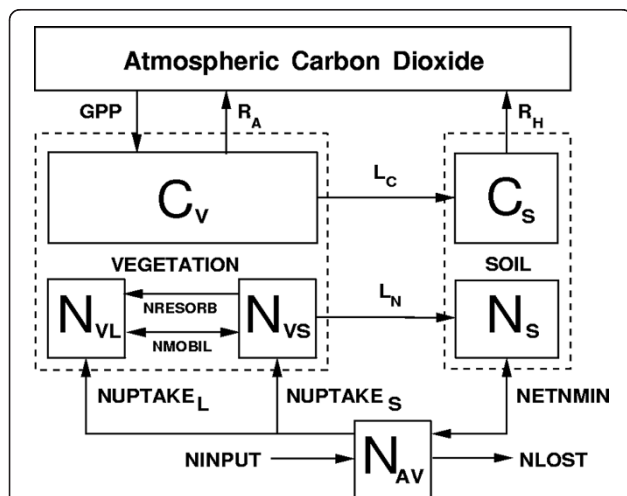


Figure 1 The terrestrial ecosystem model. In this figure GPP, R_A , L_C , and R_H represent gross primary production, autotrophic respiration, litterfall, and heterotrophic respiration, respectively. These carbon fluxes are in grams of carbon per square meter per month ($\text{g C m}^{-2} \text{ month}^{-1}$). C_V , C_S , N_{VL} , N_{VS} , N_S , and N_{AV} represent mean annual values of vegetation carbon, soil carbon, labile nitrogen within vegetation, structural nitrogen within vegetation, soil nitrogen, and inorganic nitrogen in soils, respectively. The mean annual carbon values are in grams of carbon per square meter (g C m^{-2}), while nitrogen values are in grams of nitrogen per square meter (g N m^{-2}). NRESORB, NMOBIL, L_N , NUPTAKE_L, NUPTAKE_S, NETNMIN, NINPUT, and NLOST represent the amounts of resorbed nitrogen as labile nitrogen from structural nitrogen, nitrogen mobilized for production via recycling, nitrogen in litter production, nitrogen uptake by vegetation for labile nitrogen, nitrogen uptake by vegetation for structural nitrogen, net nitrogen mineralization, nitrogen input to the ecosystem, and nitrogen lost from the ecosystem, respectively. These nitrogen fluxes are in grams of nitrogen per square meter per month ($\text{g N m}^{-2} \text{ month}^{-1}$).

The model has been extensively used to evaluate C dynamics [13-15,23-26]. Its structure, algorithm, parameterization, calibration, and performance have been documented [13,26,27].

Parameters in TEM may be specific to different vegetation types, specific to different soil textures, or constant for all vegetation types and soil textures. Most of the parameters in TEM are assigned values derived from the literature, but some parameters are calibrated to the carbon and nitrogen pools and fluxes of intensively studied sites (see [25] and [27] for details). In this paper the model is calibrated for a C4 crop based on literature review [21,28]. The pools and fluxes of ecosystem carbon and nitrogen of the calibrated crop ecosystems are shown in Table 1.

Input data sets

To apply TEM to make spatially and temporally explicit estimates of ecosystem carbon storage and net primary production in this study, we use the same input data sets as were used by Zhuang et al. [13]. These input data sets are important for directly affecting processes in the model (e.g., the effects of soil temperature on heterotrophic respiration) and for defining the parameters that are specific to vegetation types and soil textures. We use a potential vegetation data set similar to that described by Melillo et al. [25]. Soil texture and elevation do not vary in time but vary in space in our simulations. Transient historical atmospheric CO_2 concentrations are used. Data sets describing historical changes in monthly air temperature and precipitation are gridded at $0.5^\circ \times 0.5^\circ$ spatial resolution for our simulations [13].

Table 1 Carbon and nitrogen pools and fluxes used for a C4 crop parameterization

Variable	Values ^a for C4	Source and comments
C_V	649	Evrendilek [28]
N_V	9.9	Evrendilek [28]
C_S	3071.5	Evrendilek [28]
N_S	307.1	Evrendilek [28]
N_{AV}	2.64	Based on 0.86%, the mean N_{AV}/N_S ratio
GPP	649	Evrendilek [28]
NPP	296.6	Evrendilek [28]
NPPSAT	296.6	Evrendilek [28]
NUPTAKE	3.98	Calculated from NPP_n , $75\% NPP_n = NUPTAKE$

^aUnits for annual gross primary production (GPP), net primary production (NPP), and NPPSAT are in grams of carbon per square meter per year ($\text{g C m}^{-2} \text{ year}^{-1}$). Units for vegetation C (C_V) and soil C (C_S) are in grams of carbon per square meter (g C m^{-2}). Units for vegetation N (N_V), soil N (N_S), and inorganic N (N_{AV}) are grams of nitrogen per square meter (g N m^{-2}). Units for annual N uptake by vegetation (NUPTAKE) are in grams of nitrogen per square meter per year ($\text{g N m}^{-2} \text{ year}^{-1}$).

Global simulations

We run TEM at $0.5^\circ \times 0.5^\circ$ (latitude \times longitude) resolution from 1900 to 2000. For the simulations of C4 crops, we assume that each grid cell was replaced with the C4 crop and keep the information of soils, elevation, and climate as the same as the simulation for natural ecosystems. For each grid cell, we first run TEM to equilibrium for an undisturbed ecosystem using the long-term averaged monthly climate and CO₂ concentrations from 1900 to 1949 to account for the influence of inter-annual climate variability on the initial conditions of the undisturbed ecosystem. We then run the model with transient monthly climate data from 1900 to 2000. The simulated NPP for C4 crop simulations of the year 2000 are used for our study and are available at the regional level by AEZ in Additional file 2.

Results and discussion

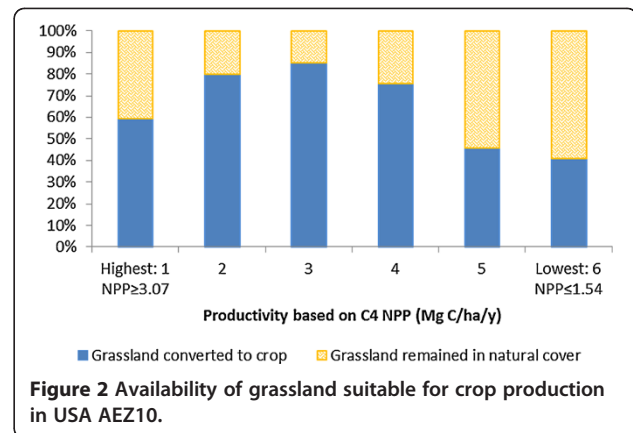
Using NPP data to obtain ETA

We use NPP^{C4} data as a proxy for yield to calculate the regional land conversion factors by AEZ using the relationship defined in Equation 1. In this process, we first matched the results from TEM with GTAP land use databases to assign AEZs to all grid cells across the world. We then imposed several restrictions to drop lands which are not good for crop production. In particular, we dropped the grid cells with the following types of land cover^c:

- Alpine tundra and polar desert,
- Forested boreal wetlands,
- Non-forested boreal wetlands,
- Temperate forested wetlands,
- Xeric shrublands,
- Tropical forested wetlands,
- Deserts,
- Tropical non-forested wetlands,
- Tropical non-forested floodplains,
- Temperate non-forested wetlands,
- Temperate forested floodplains,
- Temperate non-forested floodplains.

In addition, we dropped all grid cells with cells with median of terrain slopes greater than or equal to 16%^d. We dropped these because they are not appropriate for crop production. We then used the cleaned database to derive the land conversion factors.

To explain the derivation process, we first analyze our data for two sample regions: USA AEZ10 and Brazil AEZ4. Figures 2 and 3 represent the shares of available and converted natural grasslands in these two sample areas. In each graph, we classified the land into six groups of productivities measured in terms of NPP^{C4}. Figure 2 indicates that in USA AEZ10, a big portion of



the natural grassland is already converted to crop production. A small amount of grassland is available to be converted to crop production in this AEZ. However, the available land is distributed across all productivity groups. Note that the AEZ10 of the USA covers a large area with relatively different land qualities, weather conditions, and length of growing periods between 180 to 240 days. While the estimated NPPs at the grid-cell level can be used to examine heterogeneity in yield in sub-AEZs, we developed our analyses at the AEZ level because the GTAP-BIO model works at the AEZ level and ignores variations within an AEZ.

Now, consider Figure 3 which indicates that in the Brazil AEZ4, a lot of grassland remains in natural cover, and only a small portion of grassland in this AEZ has been converted to crop production. In this AEZ, available land is distributed across all productivity groups as well.

Now, consider another aspect of the NPP data in these two AEZs. Figure 4 compares the average productivity of grassland converted to crop production in the past with the productivities of all grassland parcels that remained in natural cover in the USA AEZ10. In this figure, grid cells are sorted according to their productivity. Thus, when we move from the left side to the right side of the

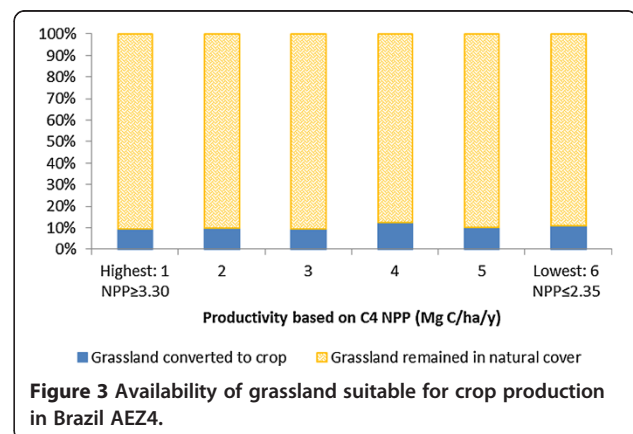
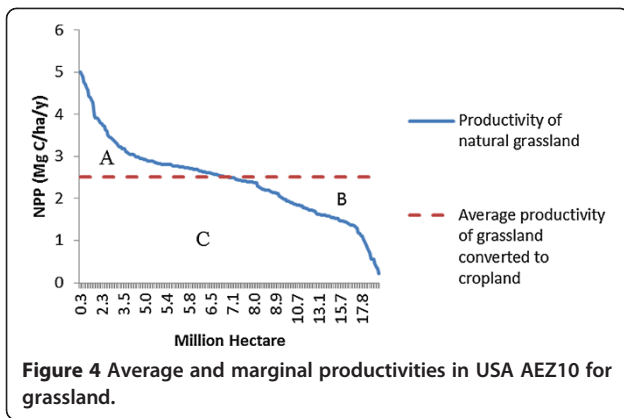
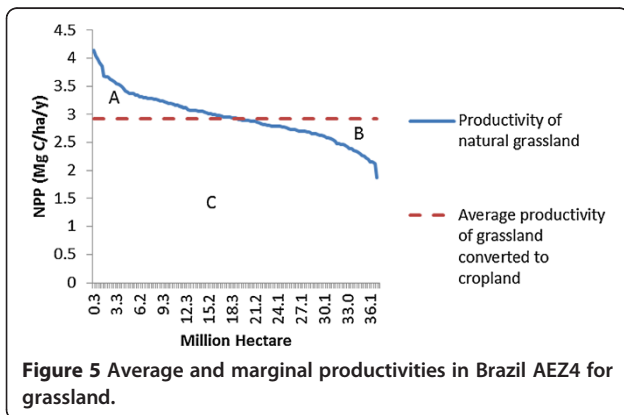


Figure 3 Availability of grassland suitable for crop production in Brazil AEZ4.



horizontal axis, we move from grid cells with higher productivities to the grid cells with lower productivities. In this graph, the ratio of the area A + B (area below the productivity of grassland curve) over the area B + C (area below the red line) provides us a land conversion factor for this type of land in this AEZ. All of the land pixels in area A represent pixels with productivity (for C4) higher than the average productivity of existing cropland (the straight line). All of the pixels in area B have a productivity that is less than the average cropland. Thus, area A + C over area B + C shows the average productivity of new land versus the average productivity of existing cropland. The assumption then is that the marginal unit of land has this productivity. Figure 5 provides the same information for Brazil AEZ4. The calculated ETA obtained for the USA AEZ10 and Brazil AEZ4 are 0.93 and 1, respectively. These figures represent ETA obtained from the mix of data on grassland and forest in these two AEZs.

While we are able to derive the conversion factors for all types of land cover, we pooled all land types in each AEZ in each region and defined the geographical land conversion factors at the AEZ level. It is important to point out that the model does not take irrigation into account. However, in the real world, in some areas, lands are under crop production with irrigation. For this reason, we reduced the productivity of all natural land by



10%^e and assumed that there is no land conversion factor greater than 1^f. The results of these calculations are shown in Table 2. In this table, a value of 0 means no land is available and a value of 1 shows that the average productivity of new land is equal to the average productivity of the existing cropland. Table 2 indicates that the US land conversion factors range from 0.51 to 1, depending on the AEZ. The GTAP-BIO value for the land conversion factor (i.e., $ETA = 0.66$) falls within this range. However, Table 2 shows that the Brazil land conversion factors range from 0.89 to 1 and that most of them are around 0.9. This means that the estimated ETA values for Brazil are higher than the simple assumption made by Hertel et al.

To check the sensitivity of results with respect to the type of simulated crop, another set of new ETA parameters are obtained using C3 NPP for all regions by AEZ. In this practice, the TEM is tuned for a C3 crop, and then the whole process described above is repeated for the C3 crop as well. The results indicate that there is no major difference between the ETA parameters obtained from the C4 NPP and C3 NPP. For example, consider the results presented in Table 3 for USA by AEZ. This table shows that the calculated NPP^{C4} values are significantly larger than the calculated NPP^{C3} values for the same categories of land in each AEZ. For example, the weighted averages of C4 NPP and C3 NPP for cropland in USA AEZ7 are 1.5458 and 0.4229 metric tons $ha^{-1} year^{-1}$, respectively. However, the ETA ratios obtained from the NPP^{C4} and NPP^{C3} values are very similar. This means that the estimated ETA parameters are robust to the representative C3 and C4 crops simulated by the TEM. However, we have not examined the results for all types of crops. The results may change for a particular type of crop. The next section applies the extensive margins reported in Table 2 and reevaluates the results obtained by Hertel et al.

Land use impacts of US ethanol production

In this section, we assess the land use impacts of the 2015 US ethanol mandate using the GTAP-BIO model under two set of assumptions: (1) assume that $ETA = 0.66$ everywhere across the world and (2) assume that ETA varies across the world according to the values presented in Table 2. In both simulations, we follow Hertel et al. to define our simulation process. The results indicate that the new extensive margins estimated using the TEM at the region/AEZ level result in a reduced estimated land requirement for ethanol production by about 25% compared with earlier estimates based on the global *ad hoc* value of $ETA = 0.66$. Geographical distributions of induced cropland expansions are almost identical for both simulations (with the *ad hoc* uniform and new regional extensive margins). The results show that the

Table 2 Regional land conversion factors obtained from NPP data for a C4 crop

AEZ ^a	Region ^b																		
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19
1	0.00	0.00	0.91	0.00	0.00	0.00	0.93	1.00	0.95	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.68	0.61	1.00
2	0.00	0.00	0.92	0.00	0.00	0.00	0.89	1.00	0.81	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.59	1.00	1.00
3	0.00	0.00	0.93	0.00	0.00	0.00	0.86	1.00	0.90	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.89	0.74
4	0.00	1.00	0.89	0.00	0.00	1.00	0.93	1.00	0.88	0.00	0.88	0.89	1.00	0.00	0.00	0.00	0.86	0.92	0.92
5	0.00	0.00	0.93	0.00	0.00	0.90	0.98	0.88	0.90	0.00	0.90	0.91	0.98	0.00	0.00	0.00	0.00	1.00	0.96
6	0.00	0.00	0.91	0.00	0.00	0.88	0.98	0.97	0.85	0.00	0.88	0.95	0.78	0.00	0.00	0.00	0.00	1.00	0.88
7	0.73	0.00	0.00	0.89	0.00	0.80	0.90	0.59	1.00	1.00	0.00	0.00	0.43	1.00	0.98	0.00	0.46	0.80	0.65
8	0.71	0.90	0.00	0.91	0.00	1.00	0.71	0.72	0.90	1.00	0.00	0.00	0.60	0.84	0.84	0.00	0.71	0.79	0.86
9	1.00	1.00	0.00	0.85	1.00	0.98	0.88	1.00	0.91	1.00	0.00	0.00	1.00	0.94	0.82	0.00	0.77	0.84	0.93
10	0.93	0.96	0.88	0.88	0.96	0.84	1.00	0.89	1.00	0.93	0.00	1.00	0.92	0.89	0.89	0.87	0.98	0.88	0.92
11	0.96	0.83	1.00	1.00	0.94	0.95	0.90	1.00	0.87	0.84	0.00	1.00	0.79	0.89	1.00	0.00	0.00	0.77	0.96
12	0.89	0.86	0.91	0.00	0.95	0.92	0.90	1.00	0.84	0.00	0.00	1.00	1.00	0.00	0.89	0.00	0.00	1.00	0.98
13	0.92	1.00	0.00	0.55	0.00	1.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.63	0.97	0.00	0.00	0.00	0.00
14	0.51	0.89	0.00	0.80	0.00	0.92	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.90	1.00	0.95	0.00	0.00	0.00
15	0.71	0.90	0.00	0.83	1.00	1.00	1.00	0.00	0.64	1.00	0.00	1.00	1.00	0.90	1.00	0.87	0.00	0.00	1.00
16	1.00	0.89	0.00	1.00	0.00	1.00	1.00	0.00	0.92	0.00	0.00	1.00	1.00	0.85	1.00	1.00	0.00	0.00	1.00
17	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

In this table, a value of 0 means no land is available. In general, when $ETA = 1$, it shows that the average productivity of new land is equal to the average productivity of existing cropland. ETA values greater than 1 are also truncated to 1. ^aRows are AEZs from AEZ1 to AEZ18. ^bColumns R1 to R19 represent the following regions, respectively: USA, EU27, Brazil, Canada, Japan, China, India, C-America, S-America, E-Asia, Mala-Indo, R-SE-Asia, R-S-Asia, Russia, E-Europe-RFSU, Other Europe, M-East-N-Africa, Sub-Saharan Africa, Oceania. These regions and their members are listed in Additional file 1.

model with new extensive margins predicts a smaller share for forest in the total land conversion (23.3% for uniform ETA versus 14.4% for regional/AEZ $ETAs$). The lower share for forest in induced land conversion leads to a lower estimate for induced land use emissions due to ethanol production. Thus, having a more robust and region/AEZ specific calculation of ETA is very important in estimating the land use changes induced by biofuels.

Uncertainties and limitations

The estimated land margins provided in this paper, as with any other estimate, are subject to restrictions and uncertainties. One restriction is related to the GTAP aggregation level. The GTAP model operates at the AEZ level. At this aggregation level, the model uses the most productive lands first in response to the need for cropland expansion in each region. This means that if more

Table 3 C4 and C3 ETA parameters for USA by AEZ

AEZ	C4 NPP for cropland	C4 NPP for natural land	C4 ETA	C3 NPP for cropland	C3 NPP for natural land	C3 ETA
7	1.5458	1.1320	0.73	0.4292	0.3101	0.72
8	1.8578	1.3194	0.71	0.5203	0.3687	0.71
9	2.1272	2.2656	1.07	0.6314	0.6884	1.09
10	2.7180	2.5288	0.93	0.8522	0.8004	0.94
11	3.5494	3.3902	0.96	1.1363	1.1019	0.97
12	2.7455	2.4367	0.89	0.8750	0.7762	0.89
13	1.5184	1.4000	0.92	0.4053	0.3698	0.91
14	3.3390	1.7192	0.51	0.9340	0.4577	0.49
15	3.3484	2.3937	0.71	1.0813	0.7410	0.69

NPPs are measured in metric tons per hectare per year ($metric\ tons\ ha^{-1}\ year^{-1}$). In this table, AEZs with no or limited land are dropped, and the upper bound of ETA is not enforced.

cropland is needed, it will be taken from the most productive AEZ where land is available and profitable to be converted to crop production. In this process, the model compares average productivities at the AEZ level to determine in what AEZs the cropland should be expanded.

Prior to this research, the GTAP model was using a global land margin (0.66) everywhere all across the world. To remove this deficiency, we developed a set of regional land margins at the AEZ level. In developing these land margins, we ignored heterogeneity within AEZs and aggregated our gridded productivity estimates to AEZ level to obtain average productivities at this level to match the GTAP aggregation level. However, our data indicate that productivity can vary significantly within an AEZ. A typical AEZ covers millions of hectares of cropland, pasture, and forest and includes many landowners/farmers who have different land qualities and are also faced with different constraints and economic factors. In the real world, given the existing constraints and relative returns to alternative land types, landowners/farmers convert lands with different productivities to cropland to satisfy the need for more crops. GTAP simplifies these details and works with the average productivities at the AEZ level which are obtained from actual observations. Therefore, while in the real world, landowners/farmers convert different land qualities to cropland; the GTAP model simplifies the real world and simulates land conversion process using averages at the AEZ level. This simplification and aggregation scheme could introduce uncertainties in the estimation process of induced land use changes due to ethanol production. For example, one can imagine a situation where farmers convert most productive lands to crop production in response to a higher demand for ethanol. In this case, the model which operates based on average productivity overestimates the land requirement for ethanol production. On the other hand, it is quite possible that farmers convert low-quality land due to the existing constraints. In this case, the model underestimates the land requirement. A more precise land use module which operates at a finer grid-cell resolution and takes into account heterogeneity within each AEZ can reduce uncertainties due to the aggregation scheme at the AEZ level.

The TEM assumes no irrigation in its calculations, while in the real world, a portion of cropland is irrigated. In this paper, we reduced productivity of uncultivated land by 10% to take into account the fact that a portion of crops are irrigated. A sensitivity test indicates that the ETA values increase by about 5% to 6% if a uniform irrigation adjustment rate of 5% is used, and they drop by 4% to 6% if the adjustment rate is 15%. This sensitivity test indicates that the ETA values are not very sensitive with respect to changes in the uniform irrigation adjustment factor. However, incorporating regional irrigation

adjustment factors may change this picture for the regions where productivities of irrigated and rainfed lands are very different, and water resources are available to expand irrigated crops. Of course, using a simple irrigation adjustment factor introduces some uncertainties in the results. However, without comprehensive data analyses, it is almost impossible to evaluate these uncertainties. To evaluate uncertainties related to the simple irrigation adjustment factor, data on harvested area and crop produced by irrigation and rainfed categories at the grid cell are needed. This is a major research activity which can be addressed in the future.

The GTAP model only considers land transition among accessible managed lands. However, the data set behind the TEM which represents historical changes in land cover does not distinguish between accessible managed land and inaccessible land. For this reason, we dropped lands which are not suitable for crop production or their conversion to cropland is prohibited across the world for environmental protection reasons. We also truncated ETA values >1 to 1. As mentioned earlier in this paper, ETA values >1 usually represent regions with large areas of inaccessible land with small areas of cropland. If farmers convert a significant amount of inaccessible land to cropland and the productivities of inaccessible lands are significantly different from their accessible counterparts, then our ETA values could be biased.

In this paper, the TEM is used to evaluate NNPs for cultivated and uncultivated lands. The estimation process is tested for a C4 and a C3 crop. Results indicate that the NPP values vary when we switch from C4 to C3. This means that the estimated NNPs are sensitive to the calibration process and the type of simulated crop and hence are subject to uncertainties. However, we showed that the NPP ratios developed and used in this paper are not sensitive to the calibration process.

Finally, as mentioned earlier, the original GTAP model was using a global value of $ETA = 0.66$ everywhere across the world. We also mentioned that based on the work of Field et al. [16], this figure is equal to the ratio of global average productivity of abandoned cropland over the global average productivity of active cropland both estimated in terms of NPP. Our method indeed extends the approach of Field et al. [16] to estimate ETA values at a regional scale by AEZ for converting uncultivated land to cropland. In the estimation process, we used detailed historical information at a grid-cell level. The data items used in this process are mainly representing actual observations at a grid-cell level resolution. In a few cases where proper regional data were not available, global values were used. There is no doubt that the new method provides a better representation of reality than the prior method. While the new method provides better estimates compared to the prior work, we cannot

accurately evaluate the accuracy of the estimated land margins with respect to their true values because we just do not know the true values. If we knew the true values, there was no need to estimate them.

Conclusions

The increased demand for bioenergy has raised debates on land use consequences of biofuel production. Several attempts have been made to evaluate land use impacts of biofuels and their consequent land use emissions. These evaluations are obtained from partial and general equilibrium models which relied on *ad hoc* assumptions on the productivity of new croplands. To establish a method of estimation for this parameter, this paper provides a new data set which can be used to estimate land productivity at $0.5^\circ \times 0.5^\circ$ (longitude \times latitude) grid-cell level regardless of existing economic conditions. However, we used this data set to estimate land productivity at the region/AEZ level to match with the GTAP aggregation level. Basically, instead of assuming that all countries and AEZs are the same, our approach takes advantage of information on productivity differences by country and AEZ and makes use of that information. In general, these data can be used in connection with economic models to assess regional and global land use implications of expansion/contraction in agricultural markets. As an example, this data set is used to estimate a set of regional extensive margins by AEZ according to the regional classifications used by Hertel et al. In general, the regional/AEZ extensive margins obtained from this data set fall in the range between 0.42 and 1. However, in 43 spots, estimated margins are between 1 and 1.05. In eight spots, estimated values are between 1.05 and 1.15, and in 13 spots, estimated margins are between 1.15 and 2. The cells with $ETA > 1$ mainly belong to the regions and AEZs with large natural land areas with very small areas of cropland. The regional/AEZ extensive margins obtained from this data set fall in the range between 0.42 and 1. The uniform global extensive margin of 0.66 used by Hertel et al. falls in this range. Finally, the economic model used by these authors is augmented with the new extensive margins and used to reassess land use implications due to the US ethanol program. Results indicate that new extensive margins reduce the estimated land requirement for ethanol production by 25%. This result shows that having a region/AEZ-specific calculation of extensive margins is very important in estimating the land use changes induced by biofuels.

The regional/AEZ extensive margins obtained in this paper are subject to some limitations and uncertainties. Ignoring heterogeneity in land productivity within AEZs, random selection of new land, using a global factor to adjust yield for irrigation, and ignoring land accessibility factors are the main limitations and sources of uncertainty which deserve more attention in future research.

Endnotes

^aThese authors used the partial equilibrium model developed at the Center for Agricultural and Rural Development at Iowa State University. This model is a revised version of the model originally developed by the Food and Agricultural Policy Research Institute at Iowa State University and the University of Missouri.

^bThese authors used the computable general equilibrium model known as GTAP-BIO which was developed at the Center for Global Trade Analysis Project at Purdue University. This model is a variety of the GTAP model developed originally by Hertel [29].

^cThe wetland is dropped because conversion of this type of land to cropland is prohibited in many countries. Of course, in the real world, when there is no credible regulation, people may convert wetlands to cropland as well. However, we do not know to what extent these illegal land conversion activities may happen in future. Introducing this important issue requires further research in this area.

^dIn this research (following Fischer et al. [30]), terrain slopes are divided into seven classes of 0% to 2%, 2% to 5%, 5% to 8%, 8% to 16%, 16% to 30%, 30% to 45%, and >45%.

^eThe global average yield of coarse grains (including irrigated and rainfed areas) is about 9.8% higher than the global average yield for rainfed coarse grain. We acquired this figure from the data set provided by Portmann et al. [31]. Based on this figure, the productivity of uncultivated land is reduced by 10% to take into account for the difference between average productivities. The sensitivity of the results with respect to changes in this parameter is tested with values of 5% and 15% instead of 10%. This test shows that the ETA values increase by about 5% to 6% if the irrigation adjustment rate of 5% is used, and they drop by 4% to 6% if the adjustment rate is 15%.

^fAn $ETA > 1$ means that the uncultivated land is more productive than the existing land. When $ETA > 1$, one can ask this question: why is less productive land used first? Several reasons such as high costs of land conversion, distance to market, lack of infrastructure, restrictions due to conservation programs, and many other factors could explain this result. Hiedere et al. [32] developed a method to identify marginal land subject to these types of constraints. When the most productive land is not used due to economic or other constraints, then one can observe an $ETA > 1$. We limit the upper bound of ETA to 1 because GTAP does not take into account factors such as those mentioned here and only consider land transition among accessible lands. ETA values > 1 usually represent regions with large areas of inaccessible land with small areas of cropland.

Additional files

Additional file 1: Appendix A. Regions and their members.

Additional file 2: Appendix B. Average NPP-C4 for cropland and for natural land (metric tons of carbon per hectare per year).

Abbreviations

AEZ: agroecological zone; GPP: gross primary production; NPP: net primary product; R_A : autotrophic respiration; R_H : heterotrophic respiration; TEM: terrestrial ecosystem model.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

FT, QZ, and WET jointly developed the theoretical foundation and carried out main developments. QZ and XL developed simulations to assess NNPC⁴ using TEM. FT and WT jointly incorporated new extensive margins into the GTAP-BIO model. All authors read and approved the final manuscript.

Acknowledgments

The research underlying this paper was partially funded by DOE and Argonne National Laboratory. We are deeply indebted to Dr. Michael Wang for his many contributions to this research. We are also indebted to two anonymous referees of the journal of *Energy, Sustainability, and Society* for their constructive and detailed comments.

Author details

¹Department of Agricultural Economics, Purdue University, 403 West State St, West Lafayette, IN 47907-2056, USA. ²Department of Earth, Atmospheric and Planetary Sciences, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN 47907-2051, USA.

Received: 12 January 2012 Accepted: 29 November 2012

Published: 28 December 2012

References

- Fritsche UR, Sims REH, Monti A (2010) Direct and indirect land-use competition issues for energy crops and their sustainable production - an overview. *Biofuels, Bioprod Biorefin* 4(6):692–704
- Timilsina GR, Beghin JC, van der Mensbrugghe D, Mevel S (2012) The impacts of biofuels targets on land-use change and food supply: a global CGE assessment. *Agric Econ* 43(3):315–332
- Dumortier J, Hayes DJ, Carriquiry M, Dong F, Du X, Elobeid A, Fabiosa JF, Tokgoz S (2011) Sensitivity of carbon emission estimates from indirect land-use change. *AEPP* 33(3):428–448
- Arima EY, Richards P, Walker R, Caldas MM (2011) Statistical confirmation of indirect land use change in the Brazilian Amazon. *Environ Res Lett* 6(2):024010
- Searchinger T, Heimlich R, Houghton RA, Dong F, Elobeid A, Fabiosa J, Tokgoz S, Hayes D, Yu T-H (2008) Use of US croplands for biofuels increases greenhouse gases through emissions from land use change. *Science* 319:1238–1240
- Hertel TW, Golub A, Jones AD, O'Hare M, Plevin RJ, Kammen DM (2010) Global land use and greenhouse gas emissions impacts of US maize ethanol: estimating market-mediated responses. *Bioscience* 60:223–231
- Taheripour F, Hertel TW, Tyner WE, Beckman JF, Birur DK (2010) Biofuels and their by-products: global economic and environmental implications. *Biomass Bioenergy* 34:278–289
- Taheripour F, Hertel TW, Tyner WE (2011) Implications of the biofuels boom for the global livestock industry: a computable general equilibrium analysis. *Agric Econ* 42:325–342
- Brits W, Hertel TW (2011) Impacts of EU biofuels directives on global markets and EU environmental quality: an integrated PE, global GCE analysis. *Agric Ecosyst Environ* 142(1–2):102–109
- Kenney R, Hertel TW (2009) Indirect land use impacts of US biofuels policies: the importance of acreage, yield and bilateral trade responses. *Am J Agric Econ* 91:895–909
- Huang H, Khanna M (2010) An econometric analysis of US crop yield and cropland acreage. Discussion paper, University of Illinois, Urbana-Champaign
- Berry S (2011) Biofuels policy and the empirical inputs to GTAP models. <http://www.arb.ca.gov/fuels/lcfs/workgroups/ewg/010511-berry-rpt.pdf> Accessed 12 Nov 2012
- Zhuang Q, McGuire AD, Melillo JM, Clein JS, Dargaville RJ, Kicklighter DW, Myneni RB, Dong J, Romanovsky VE, Harden J, Hobbie JE (2003) Carbon cycling in extratropical terrestrial ecosystems of the Northern Hemisphere during the 20th century: a modeling analysis of the influences of soil thermal dynamics. *Tellus B* 55(3):751–776
- Qin Z, Zhuang Q, Zhu X, Cai X, Zhang X (2011) Carbon consequences and agricultural implications of growing biofuel crops on marginal agricultural lands in China. *Environ Sci Technol*. doi:10.1021/es2024934
- Qin Z, Zhuang Q, Chen M (2012) Impacts of land use change due to biofuel crops on carbon balance, bioenergy production, and agricultural yield, in the conterminous United States. *Glob Change Biol Bioenergy* 4(3):277–288
- Field CB, Campbell JE, Lobell DB (2007) Biomass energy: the scale of the potential resource. *Trends Ecol Evol* 23(2):65–72
- Prince SD, Haskett J, Steinger M, Strand H, Wright R (2001) Net primary production of US Midwest croplands from agricultural harvest yield data. *Ecol Appl* 11(4):1194–1205
- Hicke JA, Lobell DB, Asner GP (2004) Cropland area and net primary production computed from 30 years of USDA agricultural harvest data. *Earth Interact* 8:1–20
- Li X, Wang H, Ge Y, Long H, Zhang C (2007) Estimation of winter wheat yield in Hebei plain of China by improved CASA model. In: (ed) Geoscience and Remote Sensing Symposium IGARSS 2007, IEEE, Barcelona, 23–28 July 2007
- Campbell JE, Lobell DB, Genova RC, Field CB (2008) The global potential of bioenergy on abandoned agriculture lands. *Environ Sci Technol* 42(15):5791–5794. doi:10.1021/es800052w
- Lu X, Zhuang Q (2010) Evaluating climate impacts on carbon balance of the terrestrial ecosystems in the Midwest of the United States with a process-based ecosystem model. *Mitig Adapt Strateg Glob Change* 15:467–487
- Tum M, Günther KP (2011) Validating modeled NPP using statistical yield data. *Biomass Bioenergy* 35(11):4665–4674
- Euskirchen ES, McGuire AD, Kicklighter DW, Zhuang Q, Clein JS, Dargaville RJ, Dye DG, Kimball JS, McDonald KC, Melillo JM, Romanovsky VE, Smith NV (2006) Importance of recent shifts in soil thermal dynamics on growing season length, productivity, and carbon sequestration in terrestrial high-latitude ecosystem. *Glob Chang Biol* 12:731–750
- Balshi MS, McGuire AD, Zhuang Q, Melillo JM, Kicklighter DW, Kasischke E, Wirth C, Flannigan M, Harden J, Clein JS, Burnside TJ, McAllister J, Kurz WA, Apps M, Shvidenko A (2007) The role of historical fire disturbance in the carbon dynamics of the pan-boreal region: A process-based analysis. *J Geophys Res* 112:G02029
- Melillo JM, McGuire AD, Kicklighter DW, Moore III B, Vorosmarty CJ, Schloss AL (1993) Global climate change and terrestrial net primary production. *Nature* 363:234–240
- McGuire AD, Melillo JM, Joyce LA, Kicklighter DW, Grace AL, Moore III B, Vorosmarty CJ (1992) Interactions between carbon and nitrogen dynamics in estimating net primary productivity for potential vegetation in North America. *Global Biogeochem Cycles* 6:101–124
- Raich JW, Rastetter EB, Melillo JM, Kicklighter DW, Steudler PA, Peterson BJ, Grace AL, Moore III B, Vorosmarty CJ (1991) Potential net primary productivity in South America: application of a global model. *Ecol Appl* 1:399–429
- Evrendilek F, Wali MK (2004) Changing global climate: historical carbon and nitrogen budgets and projected responses of Ohio's cropland ecosystems. *Ecosystems* 7(4):381–392
- Hertel TW (1997) *Global trade analysis, modeling and applications*. Cambridge University Press, Cambridge
- Fischer G, Velthuisen H, Nacgtergaele F, Medow S (2000) *Global Agro-Ecological Zones (Global-AEZ): Food and Agricultural Organization of the United Nation and International Institute for Applied System Analysis [CD-ROM]*. FAO/IIASA, Rome/Laxenburg

31. Portmann T, Siebert S, Döll P (2010) MIRCA2000-global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochem Cycles* 24:GB1011
32. Hiederer R, Ramos F, Capitani C, Koeble R, Blujdea V, Gomez O, Mulligan D, Marelli L (2010) Biofuels: a new methodology to estimate GHG emissions from global land use change - a methodology involving spatial allocation of agricultural land demand and estimation of CO₂ and N₂O emissions. EC Joint Research Centre, Ispra

doi:10.1186/2192-0567-2-25

Cite this article as: Taheripour *et al.*: Biofuels, cropland expansion, and the extensive margin. *Energy, Sustainability and Society* 2012 2:25.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
