Applied Geography 62 (2015) 366–376

Contents lists available at ScienceDirect

Applied Geography

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

Predicting 21st century global agricultural land use with a spatially and temporally explicit regression-based model

Nicholas Haney, Sagy Cohen*

Department of Geography, University of Alabama, Box 870322, Tuscaloosa, AL 35487, USA

ARTICLE INFO

Article history:
Received 19 February 2015
Received in revised form
24 May 2015
Accepted 24 May 2015
Available online xxx

Keywords:
Global
Land use
Model
GLUDM
Cropland
Agriculture

ABSTRACT

The extensive alteration of the earth's land cover during the anthropocene had widespread, and in some cases unknown, effects on terrestrial and atmospheric conditions and processes. Predicting future changes to the earth system therefore mandate a future-predicting framework of land use dynamics. However while future-predicting earth surface and atmospheric models tend to explicitly incorporate projected climatic conditions they all but ignore or overly simplify land use dynamics. As most surface and atmosphere dynamics models use gridded input datasets, and land use is a highly spatially-dynamic phenomena, a need clearly arise for spatially explicit representation of future land use dynamics. While a number of such datasets exists at regional and country scales, no fully gridded future-predicting global land use model and database has been reported to date. Here we present the Global Land Use Dynamics Model (GLUDM), a gridded and temporally explicit agricultural land use predictor. GLUDM calculates the relative area of a land use category (e.g. cropland) in each grid-cell by generating unique regression coefficients in each grid-cell based on local historic trends and global population dynamics. Spatial expansions or abandonment of agricultural land is simulated by propagating excesses or deficiencies in agricultural areas between neighboring grid-cells. This spatial connectivity is restricted by topographic, latitudinal and urban characteristics. A validation analysis shows that GLUDM corresponds well to observed land use distribution. GLUDM-predicted global cropland area dynamics between 2005 and 2100 are described herein. Globally, 18% increase in cropland area is predicted between 2005 and 2050 which corresponds very well to previous estimations. Following 2050, a general decrease in cropland area is predicted. The results reveal new insights about global cropland dynamics, demonstrating, for example, that changes in its spatial distribution will be highly heterogeneous, at both micro and macro scales, in some locations worldwide.

1. Introduction

Over the last 300 years humans have greatly altered the natural environment to meet demands for food, fiber, and settlement. The pre-existing ecosystems have been continually relegated to ever shrinking marginal undeveloped and managed areas. As a result the world's natural land cover has been substantially modified. It has been estimated that as much as 50 percent of the Earth's land surface has its biological production completely dominated by humans (Vitousek, Mooney, Lubchenco, & Melillo, 1997). Similarly, Ellis, Goldewijk, Siebert, Lightman, and Ramankutty (2010) found that 39 percent of the earth's ice-free land area had either been converted to agriculture or to urban areas. This modification of natural systems has disrupted a number of important biogeochemical cycles such as the carbon and nitrogen cycles. This has led to increased levels of greenhouse gases, a decline in the health of aquatic ecosystems, and has altered precipitation.

The primary driver of this expansion is the expanding human population (Doos & Shaw, 1999). From 1900 to 2000 the population of the earth experienced a 400 percent increase. While the growth in human biomass is itself a factor, the resulting increase in natural resource consumption to feed, cloth, and house a population of this multitude has had a far greater impact on the environment. While impoundments, mining operations, and forestry make significant changes to the landscape, nothing has altered the natural landscape as much as conversion to agriculture. Over the last 300 years agricultural expansion has resulted in a global net loss of between 8 and 11 million km^2 of forestland (Foley, DeFries, Asner, Barford, &

* Corresponding author.
E-mail address: sagy.cohen@ua.edu (S. Cohen).

http://dx.doi.org/10.1016/j.apgeog.2015.05.010
0143-6228/

Bonan, 2005). Conversion to agriculture has lead to increased runoff, soil erosion, denitrification, desertification, the extinction and endangerment of many species, and an altered atmospheric composition (Tilman, 1998; Foley et al., 2005).

Scientists have long understood the consequences of conversion to agricultural and sought accurate estimates of the global amount of land under agricultural production. Until the 1960's this was impossible as many nations were unable to inventory the amount of agricultural land. The Food and Agriculture Organization (FAO) of the United Nations began keeping detailed records of the amount of agricultural land in each of its member nations in the 1960's (FAO, 2013). By the 1990's the global coverage of IR satellite imagery and the greater availability of agricultural data increased the accuracy and ease of making these estimates. Ramankutty and Foley (1999) developed a comprehensive map of the extent of modern agriculture by combining remotely sensed data with cropland inventories where available. Using recent trends in agricultural development they were able to use a simple land use allocation model and run the model in reverse to the 1700's using available land use data as a model constraint. The HYDE database, using a similar approach to Ramankutty and Foley (1999), was developed to test the IMAGE 2 climate change model and was able to model land use back to 10,000 B.C. (Goldewijk, Beusen, van Drecht, & de Vos, 2010).

Global and continental scale numerical models are increasingly being developed and used for predicting current and future atmospheric, biospheric, hydrospheric and lithospheric conditions and fluxes. As most of these models use gridded representation and land use is often an important parameter, predicting future land use dynamics at global scale on a gridded surface is an important and timely undertaking. For example, the WBMsed model (Cohen, Kettner, Syvitski, & Fekete, 2013) is a gridded model that predicts daily water, sediment and nutrients flux in global rivers (Cohen, Kettner, & Syvitski, 2014). The model can be used to predict 21st century fluxes using future-predicted climatic datasets form a suite of GCMs (General Circulation Models) outputs. As land use is a key parameter in water, sediment and nutrient input to river systems, developing a spatially and temporally explicit land use input dataset would be instrumental for reliably predicting these fluxes into the future.

Based on two independent review papers (Heistermann, Muller, & Ronneberger, 2006; Schmitz et al., 2014) we conclude that, to date, no future-predicting, global, fully gridded and temporally explicit land use dynamics predictions have been published. CLUE (Veldkamp & Fresco, 1996) is a process based modeling framework that allows the user to develop spatially explicit future land use dataset based on multiple scenarios. However this model works only at the regional scale and requires numerous variables for which global data is not available. If the data was available this approach is still impracticable for global scale modeling as data would be collected within political units and the grid cells would overlap international boundaries. Thus economic policies, political decisions, and other variables would not be applicable. Other factors such as a global set of detailed soil types, types of crops grown, available water supply, and agricultural practices are difficult or impossible to obtain even at regional scales. This necessitates significant abstraction if agricultural land use is to be modeled at the global scale using a spatially explicit gridded model.

Despite these difficulties modeling land use, using a gridded structure can be achieved by focusing on the one variable that has the most influence on determining agricultural land use. This variable should be easy to obtain and one that is universally understand to influence the amount of agricultural land necessary. This variable is global population. While this may not be the single most important variable at a sub global scale it is appropriate to use at a global scale given our globalized agricultural system.

In this paper we describe the theoretical and algorithmic framework of the Global Land Use Dynamics Model (GLUDM), present validation results and discuss future agricultural land use dynamics, focusing on 21st century changes in cropland.

2. Methods

2.1. Theoretical framework

Historically, the most significant controlling factor on global agricultural extent has been human population (Doos & Shaw, 1999). While in pre-industrial times the population of each country or region controlled the local extent of agriculture, in the current industrialized economy global population seems to be the main control on the amount of global agricultural land (Trostle, 2008). Thus as global population increases the total amount of agricultural production must also increase.

Another major factor is the technology used in agricultural production. Advancements in agricultural technologies mean that increasing global population will require a relatively smaller increase in agricultural lands (Heistermann et al., 2006). The global standard of living is another important factor. Wealthy societies have a higher caloric intake than poorer societies, requiring a greater agricultural area to sustain them. Economic factors, water availability and human decision making controls which types of crops is planted with less productive crops requiring a greater amount of agricultural land to sustain the same number of people (Heistermann et al., 2006). An absolute constraint on the spatial extent of agricultural land is the environmental variables acting at a given point in space and time. These includes factor such as latitude, altitude, and climate. However environmental constraints can change overtime in response to human impacts or advances in technology.

The question now becomes which of these variables are easily available and are applicable at the global scale. Estimates of population are readily available from the aforementioned HYDE dataset. The environmental variables latitude and longitude are also easy to implement. However the other variables, described above, are either unavailable or inapplicable at the global level. A way to account for these variables implicitly is through the development of regression equations based upon the principle driver of agricultural development, population.

Creating a global regression equation is illogical and will convey incorrect information. Therefore each location on continental earth (i.e. grid cell) would require a unique regression equation relating the changes in agricultural land use for that area in the past. This method can account for recent changes in the fertility of agricultural land and technology. The basic principle is to read in values from a gridded input from a number of years and use the relationship between the population of the world at that year to calculate regression coefficients for that grid cell. Then the model can calculate the extent of agriculture at a given point in time by inserting the population at that year into the regression equation.

2.2. Modeling algorithm

The HYDE 3.1 dataset (Goldewijk et al., 2010) of gridded cropland and grazing land from the years 1960, 1970, 1980, 1990, 2000, and 2005 were chosen to serve as the dependent variables when calculating the regression equations while the total global population served as the independent variables (Fig. 1). The input files are scale independent as the GLUDM model can adjust the internal variables according to the desired output scale. In this paper we use 5 arc-minute spatial resolution to readily align our results to the HYDE 3.1 maps. For each grid cell GLUDM reads in the values from