Global Change Biology (2012) **18**, 3125–3148, doi: 10.1111/j.1365-2486.2012.02759.x

A Land System representation for global assessments and land-use modeling

SANNEKE VAN ASSELEN and PETER H. VERBURG

Institute for Environmental Studies, VU University Amsterdam, De Boelelaan 1087, 1081 HV, Amsterdam, The Netherlands

Abstract

Current global scale land-change models used for integrated assessments and climate modeling are based on classifications of land cover. However, land-use management intensity and livestock keeping are also important aspects of land use, and are an integrated part of land systems. This article aims to classify, map, and to characterize Land Systems (LS) at a global scale and analyze the spatial determinants of these systems. Besides proposing such a classification, the article tests if global assessments can be based on globally uniform allocation rules. Land cover, livestock, and agricultural intensity data are used to map LS using a hierarchical classification method. Logistic regressions are used to analyze variation in spatial determinants of LS. The analysis of the spatial determinants of LS indicates strong associations between LS and a range of socioeconomic and biophysical indicators of human-environment interactions. The set of identified spatial determinants of a LS differs among regions and scales, especially for (mosaic) cropland systems, grassland systems with livestock, and settlements. (Semi-)Natural LS have more similar spatial determinants across regions and scales. Using LS in global models is expected to result in a more accurate representation of land use capturing important aspects of land systems and land architecture: the variation in land cover and the link between land-use intensity and landscape composition. Because the set of most important spatial determinants of LS varies among regions and scales, land-change models that include the human drivers of land change are best parameterized at sub-global level, where similar biophysical, socioeconomic and cultural conditions prevail in the specific regions.

Keywords: global, human-environment interactions, land cover, land system, land-change models, spatial determinants

Received 8 December 2011 and accepted 21 May 2012

Introduction

Humans have affected the natural global environment for millennia, predominantly by converting natural land into settlements, cropland and grazing land (Pongratz *et al.*, 2008; Ellis *et al.*, 2010; Klein Goldewijk *et al.*, 2010). During the last three centuries rapid land changes occurred due to the increasing demand for agricultural products (Klein Goldewijk, 2001; Turner *et al.*, 2007). These changes have significant impacts on ecosystems, for example because they affect climate, nutrient- and hydrological cycles, and biodiversity (Foley *et al.*, 2005; Turner *et al.*, 2007; Hibbard *et al.*, 2010; Klein Goldewijk *et al.*, 2010; Verburg *et al.*, 2011b). Most global scale studies focus on changes in land cover, such as expansion of agricultural land or deforestation, based on large-scale monitoring through remote sensing (DeFries, 2008; Hansen *et al.*, 2010). Such studies provide insight in the patterns of the main land cover but ignore more subtle changes in the land system that may have important environmental and socioeconomic impacts. For example, extensively managed croplands have often less impact on the natural environment compared to intensively managed croplands. The importance of intensification and other changes in land management as a determinant of sustainability has called for a more integrated whole-landscape approach (DeFries & Rosenzweig, 2010).

Integrated Assessment Models (IAM's) are used to assess environmental consequences of interactions between different systems, such as economic, social, and biophysical systems. Land use is a fundamental component in IAM's, most importantly because it is both a cause and consequence of many socio-environmental processes. In these global scale models land management is represented mostly in a simplified and aggregated manner, e.g., by a single, uniform, management factor per world region (Bouwman *et al.*, 2006; Bondeau *et al.*, 2007). Land cover is often represented by the dominant land cover type of large pixels of 0.5 × 0.5 degree (Bouwman *et al.*, 2006; Lotze-Campen *et al.*, 2008; Havlík *et al.*, 2011). An advancement is made in the global land-use change model LandSHIFT, which uses a higher resolution of 5 arcminutes, representing land use by the dominant land cover in a

Correspondence: Sanneke van Asselen, tel. + 31 20 598 5404, fax + 31 20 59 895 53, e-mail: sanneke.van.asselen@vu.nl

cell (Schaldach *et al.*, 2011). Recent efforts to integrate historical and scenario-based land cover data represent fractional land cover (crop, pasture, urban) at a resolution of 0.5 × 0.5 degree as a common base for climate modeling (Hurtt *et al.*, 2011). The Global Land Cover 2000 dataset (2003) and Globcover Land Cover (2008) are frequently used to derive such aggregated representations. The architecture of land use and land management can in such a representation not be represented adequately (Turner, 2010). Relatively small land cover types (including urban area) are not represented as result of the aggregation procedures (Schmit *et al.*, 2006; Nol *et al.*, 2008), and mosaic landscapes that deliver a range of ecosystem services simultaneously as result of their composition and spatial structure, are erroneously characterized by a single homogenous land cover type (Verburg *et al.*, 2009).

Land-use intensity is a critical characteristic of agricultural land systems (Rudel *et al.*, 2009; Licker *et al.*, 2010; Neumann *et al.*, 2010), and is a major cause of environmental damage (Foley *et al.*, 2005). In agricultural systems, drivers of intensification are manifold (Lambin *et al.*, 2001; Keys & McConnell, 2005; Neumann *et al.*, 2010). Also the increasing growth and transformation of the livestock sector has significant social and environmental consequences (FAO, 2006). To examine and predict the negative effects on the environment of increasingly intensified agricultural and livestock systems, both land-use intensity and livestock information are important variables to use besides the land cover composition in global land classifications for use in global assessments and land-change models.

Ellis & Ramankutty (2008) were the first to prepare a novel representation of human-environment interactions at a global scale, providing a new classification of the earth's biomes. They classified anthropogenic biomes (anthromes), which are defined as globally significant ecological patterns created by sustained interactions between humans and ecosystems (Ellis & Ramankutty, 2008). Besides information on land cover and irrigation, population density is used as the main factor for representing the intensity of human-environment interactions, following the theory of Boserup (1965) that relates population density to land-use intensity. The Land Degradation Assessment in Drylands (LADA) project developed a global land-use characterization focusing on land degradation caused by human interventions on the land, from an ecosystem perspective (FAO, 2011). This classification is based on dominant land cover, irrigation, and livestock density data. Letourneau *et al.* (2012) developed a classification of Land Use Systems (LUS), extending the anthromes classification, based on land cover, irrigation data, population density, livestock type and density, and market accessibility.

This article aims at improving the representation of land use at the global scale by using a relatively high spatial resolution (5 arcminute) and by explicitly addressing land management as a component of land systems. The method aims to improve the representation of future land change in IAM's, and to help the evaluation of adaptation and mitigation options with such models through increasing the ability to address more subtle changes in land system architecture than are possible, based on current representations. In our study, we based ourselves on the before-mentioned efforts but explicitly aimed to classify land systems based on land cover composition, agricultural intensity and livestock density. These are variables that characterize land-use system properties directly. We therefore provide an alternative to the above-mentioned efforts by including the intensity of agricultural land management more directly as well as livestock density and the fractional composition of the different land cover types within the unit of analysis. In contrast to the anthromes map of Ellis & Ramankutty (2008) we have not used population as a classification criterion for the land systems. Population does not necessarily characterize a land use system directly and is often seen as a driver of land-use change (Boserup, 1965; Keys & McConnell, 2005; Geist *et al.*, 2006).

In this article, this new integrated approach to represent land use at a global scale, using land cover, agricultural intensity and livestock data, is presented. We evaluate the possibilities to use this land system classification in IAM's by analyzing the extent to which the spatial patterns of the classified land systems can be explained by variables that are frequently used as location factors in land change models or land change modules of IAM's. Preferably globally uniform allocation rules based on these spatial determinants are used to allocate future changes of land use. We analyze the spatial determinants of the classified land systems at both global and regional scale to investigate if globally uniform allocation rules are valid across multiple regions.

Materials and methods

Classification method

The representation of land use in this article is based on Land Systems (LS). LS do not necessarily represent typical management systems at a particular level of organization in which biophysical, economic and human elements are interdependent, as is the case for farming systems (Dixon *et al.*, 2001). The LS classification rather aims to classify combinations of land cover composition, livestock system, and land-use intensity in a series of LS that can be used as entities for land change modeling. The integration of land management and land cover aspects allows us to synchronously address multiple trajectories of land system change upon changes in driving factors. Upon commodity demand the same LS may change into a more intensively managed system or toward a LS with a higher fraction of agricultural land, depending on the conditions at the location. Like farming systems, the LS classification system focuses on agricultural systems, but in contrast to farming systems, LS are not restricted to agricultural land (van de Steeg *et al.*, 2010). A LS may contain several farms belonging to the same or different farming systems or no farm at all. Furthermore, farming systems are mostly studied at the household level. Van de Steeg *et al.* (2010) used a logit model based on location factors and household characteristics to create a regional level farming system map for the Kenyan Highlands. Messerli *et al.* (2009) created a landscape mosaic map at the meso-level (district to national level) by first delineating land cover maps, which were subsequently interpreted in terms of human-environment interactions. Global level maps of farming systems are available for livestock systems in developing countries (Dixon *et al.*, 2001; Thornton *et al.*, 2002; Kruska *et al.*, 2003). LS have full global coverage and focus on agricultural land, but also represent natural vegetation mosaics.

In this study, LS are classified at a spatial resolution of 5 arcminutes (~9.25 × 9.25 km). Although this is more detailed than the current representation of land use in most global scale IAM's, individual grid cells at this resolution still represent a mixture of different land cover types. Hence, LS represent mosaic land cover/use patterns that result from variation in both natural processes and human influences.

The LS classification is based on three main classification factors: (1) land cover, (2) livestock, and (3) agricultural intensity (Table 1). Land cover represents the composition of the landscape, while livestock and agricultural intensity data represent important characteristics of land management and farming systems. Variables used to represent land cover are tree cover (%), bare soil cover (%), cropland cover (%), and built-up area (%). These variables characterize important components of land cover patterns but do not necessarily add up to 100%. For tree and bare soil cover the MODIS Vegetation Continuous Field dataset was used (Hansen *et al.*, 2003). The built-up area map is based on MODIS 500 m satellite data (Schneider *et al.*, 2009). Cropland cover is derived from the global dataset of croplands in the year 2000 developed by Ramankutty *et al.* (2008). They combined sub-national agricultural inventory data and satellite-derived land cover data to create a cropland map at 5 arcminute resolution. This cropland dataset is consistent with statistical (sub-) national data, which is advantageous for application in IAM's that often link to macro-economic models that use statistical data as main data source (Eickhout *et al.*, 2007).

Livestock data are obtained from the Gridded Livestock of the World information system (FAO, 2007). In the LS classification two groups of livestock types are distinguished: (1) pigs and poultry, and (2) bovines, goats, and sheep. Although these two groups may also occur simultaneous, in many cases the distinction represents different livestock systems because pigs and poultry (*pp*) are monogastric species, which are less directly dependent on local land resources as compared to bovines, goats, and sheep. Moreover, this distinction may also be caused by cultural differences, for example bovines in India and pigs/poultry in China.

As an indicator for the intensity of land use global maps of the efficiency of agricultural production are used (Neumann *et al.*, 2010). These maps are constructed based on stochastic frontier production functions, which represent the maximum yield given environmental conditions. Deviations from the frontier function are caused by inefficiency of production and statistical noise. Neumann *et al.* (2010) calculated frontier yields and efficiencies of wheat, maize, and rice at a global scale, using:

$$\ln(q_i) = \beta_0 + \beta_1 ln(temp_i) + \beta_2 ln(precip_i) + \beta_3 ln(par_i) + \beta_4 ln(soil\_constr_i) + v_i - u_i \quad (1)$$

where q_i is the actual grain yield (from Monfreda *et al.*, 2008), $temp_i$ is the deviation from the optimal monthly mean temperature, $precip_i$ is the precipitation, par_i is the Photosynthetically Active Radiation, $soil\_constr_i$ are soil fertility constraints, v_i is a random error (statistical noise), and u_i represents inefficiency effects of production. Here, inefficiency is a function of irrigation, slope, agricultural population (proxy for labor availability), market accessibility, and market influence. The main advantage of using this approach is that the efficiency is a proxy for the intensity of land management independent of variations in biophysical conditions, and multiple biophysical and land management-related variables, for which global datasets were available, are used for the calculation. In principle, the yield may be relatively high with relatively little human influence (low efficiency) in areas where local conditions are favorable for crop growth. In areas with unfavorable conditions for crop growth the yield may be relatively low, despite relatively efficient land management practices (high

Table 1 Resolution, unit, and literature reference of the classification factors

Main group	Classification factor	Resolution	Unit	Reference
Land cover	Tree & bare cover	500 m	%	Hansen *et al.* (2003)
	Cropland cover	5 arcminute	%	Ramankutty *et al.* (2008)
	Built-up area	500 m	%	Schneider *et al.* (2009)
Livestock	Livestock density	3 arcminute	nr km^{-2}	FAO (2007)
Agricultural intensity	Efficiency of agricultural production	Interpolated from point data	ratio (0–1)	Neumann *et al.* (2010)

efficiency). The efficiency map is calculated using the efficiency of the dominant crop type represented in the data (maize, rice or wheat). An Inverse Distance Weighted interpolation was used to create a map covering all cropland areas in the world. Maize, rice, and wheat make up about half of the total global harvested areas. The extrapolation of the management intensities to other crops may induce some bias and the efficiency can therefore only be regarded as a proxy for the management intensity. Licker *et al.* (2010) used a similar method to calculate the yield gap in different climate zones, based on information on the growing degree days and a crop soil moisture index.

All the input maps that are input to the classification are resampled to a resolution of 9.25 × 9.25 km in equal area projection (Eckert IV) which resembles a resolution of 5 arc-minutes at the equator in geographic projection.

A hierarchical classification procedure is used to identify LS (Fig. 1; Appendix S1 in Supporting Information). Both hierarchical classification procedures based on expert-rules (Dixon *et al.*, 2001; Kruska *et al.*, 2003; Van der Steeg *et al.*, 2010; FAO, 2011) as well as statistical clustering techniques (e.g., Ellis & Ramankutty, 2008; Letourneau *et al.*, 2012) have been used to delineate land cover/use and farming systems. We have chosen a hierarchical classification procedure, similar to for example the LADA project (FAO, 2011) and Van der Steeg *et al.* (2010). Classifications based on cluster analysis are sensitive to the distance metric selected and the criterion for determining the order of clustering. Moreover, clusters identified may be (statistically) optimal for the current distribution of the land system parameters but no longer be optimal for future conditions. In our approach, the classification thresholds are arbitrarily defined by identifying natural breaks in the variable distributions while ensuring that resulting patterns correspond to common understanding of the spatial distribution of land systems. Increasing or decreasing the used thresholds by 20% did not result in different spatial patterns on a global scale, but only shifted the class boundaries. First, two very distinctive systems are delineated: (1) settlement systems, which are characterized by a relatively high percentage of built-up area, and (2) bare systems (including snow and ice), which are characterized by a high percentage of bare cover. The remaining unclassified land surface is classified into (mosaic) cropland, grassland, and forest land systems. These LS are distinguished based on the percentages of cropland, tree and bare surface cover, and are further subdivided based on the type and density of livestock. (Mosaic) Cropland systems are subdivided based on the efficiency of agricultural production. To prevent delineation of too many classes, which would make the classification more complex, posing possible problems for modeling purposes (mosaic) cropland systems were only further subdivided if the parent class still had fair world coverage (>1%).

Investigating spatial determinants of land systems

To explore how a series of biophysical and socio-economic factors can explain the spatial distribution of LS, binominal logistic regressions were performed. The fitted logistic regression models can be used to predict, for each location on the map, the probability of finding a specific LS based on the location factors that are used as independent variables in the regression approach. Logistic regressions are frequently used as input to spatial land change allocation models (e.g., Geoghegan *et al.*, 2001; Serneels & Lambin, 2001; Braimoh & Onishi, 2007; Verburg & Overmars, 2009; Letourneau *et al.*, 2012). In such models, the probability maps derived from the regression equations are used to indicate the suitability of a certain

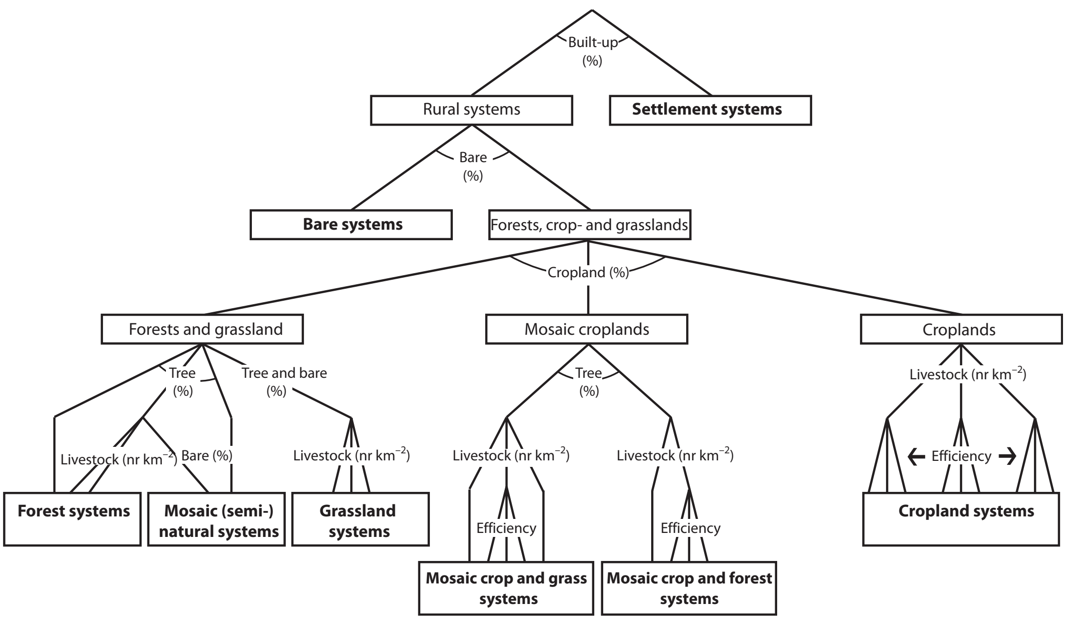

Fig. 1 Diagram showing the hierarical classification procedure. Main land systems in bold, classification variables in *italic*. Threshold values of the classification and the complete classification are given in Appendix S1.

location (grid cell) for the occurrence of a land system or land cover type, possibly accounting for temporal changes in the location factors. Based on these suitability maps, demands for various goods (e.g., livestock, crop production, urban land) are allocated spatially (using alternative algorithms), and therewith, the spatial distribution of land changes is predicted. An alternative to binominal regressions would be to use multinominal regressions (Chomitz & Gray, 1996), but this requires definition of a reference class to which the probabilities for the other land systems will relate. For the Land Systems it is difficult to specify one particular class as a reference (for example, different LS can be the reference system for intensively managed cropland). Therefore, we chose to use binominal logistic regressions, where the probability of a LS is calculated as compared to all other systems.

The variability of spatial determinants among different regions was tested by conducting the regressions separately for four regions that partly have similar LS (the Great Plains, part of Europe, north India and part of China) in addition to an analysis at the global extent. Fifteen variables that were hypothesized as potential determinants of the LS and for which global datasets with a sufficiently high resolution were available, were selected for the regression analyses (Table 2). Many of these are biophysical variables, which determine the local suitability for specific LS. Two climatic variables were selected: temperature and precipitation. The type and growth of vegetation is highly dependent on these factors. Six soil characteristics were selected: sand-, silt- and clay content, organic-matter content, Cation Exchange Capacity (CEC), and pH. In general, soils with a relatively high CEC, and intermediate levels of pH, organic-matter content, and mixed texture, are especially suitable for growing crops (FAO, 1999). Hydrological conditions are expressed by a drainage class, ranging from very poorly drained soils to excessively drained soils. Soils with intermediate drainage classes are favorable for growing crops, since these soils are well aerated whereas water (and nutrients) available for plants are not easily leached out, nor does the soil become waterlogged quickly (FAO, 1999). Two terrain variables are used: altitude and slope, which influence the landscape mosaic and the management constraints. Agricultural systems are expected to mostly occur on relatively gentle slopes and low altitudes. The last environmental variable is the global distribution of biomes (potential natural vegetation), which is expected to explain the type of vegetation in LS through capturing the natural vegetation characteristics.

Three socioeconomic variables were selected. First, a market influence index in US$/person is used (Verburg *et al.*, 2011a). This is a measure for the capital available to invest in expansion or intensification of agricultural land. Intensively managed cropland systems are expected to occur in regions with a relatively high market influence. Secondly, accessibility to national and international markets is an important factor that influences agricultural activities through providing options for marketing LS products (Verburg *et al.*, 2011a). Intensively managed agricultural systems are expected to occur close to markets given transport costs and time and the availability of inputs, while (semi-)natural systems likely occur far from markets following the classical Von Thünen model (Von Thünen, 1966). Third, population density is an important factor influencing the degree of human impact on the environment (Boserup, 1965).

The regression analysis was based on a sample of grid cells that represent a balanced sample. A balanced sample is obtained by randomly omitting observations of the over-represented class. Also, a minimum distance of one cell between the

Table 2 Units and sources of spatial determinants used for the regression analyses

Main category	Spatial determinant	Unit	Source
Climatic	Temperature (mean of monthly mean)	°C	http://www.worldclim.org/
	Precipitation (total of monthly mean)	mm	http://www.worldclim.org/
Soil characteristics	Sand content	mass%	http://www.isric.org/data/data-download
	Silt content	mass%	http://www.isric.org/data/data-download
	Clay content	mass%	http://www.isric.org/data/data-download
	Organic content	mass%	http://www.isric.org/data/data-download
	Cation Exchange Capacity	cmol/kg	http://www.isric.org/data/data-download
	pH	$-\log(H^+)$	http://www.isric.org/data/data-download
	Drainage	class	http://www.isric.org/data/data-download
Terrain	Altitude	m	http://www.worldclim.org/
	Slope	degree	derived from Altitude
Vegetation	Potential Natural Vegetation	–	Ramankutty and Foley (1999); if dominated by land use based on potential vegetation (Haxeltine and Prentice, 1996), else based on currently observed vegetation from a satellite (DISCover dataset).
Socio-economic	Market influence	USD/person	Verburg *et al.* (2011a)
	Market accessibility	index (0-1)	Verburg *et al.* (2011a)
	Population density	people km^{-2}	http://sedac.ciesin.columbia.edu/gpw/

observations was chosen to minimize spatial autocorrelation. We tested the possible influence of remaining spatial autocorrelation by fitting an alternative model based on a regular sample with at least four cells between the sampling points. At this distance we assume spatial autocorrelation to be negligible. The resulting regression models were almost exactly similar. Therefore, we conclude that our models were not affected by bias originating from remaining spatial autocorrelation in our data. Different methods of entering the independent variables in the model are explored, forward conditional and backward conditional, to test the sensitivity toward the stepwise variable selection methodology. Regression results were interpreted and compared to hypotheses, to identify causality in the estimated relations. If the input variables showed a high correlation (Pearson Correlation >0.7) only one of them was used in the regression analysis to avoid multicolinearity. The ROC (Receiver Operating Characteristic) was used as a measure of the goodness of fit of the regressions (Swets, 1988). A value of 0.5 indicates the regression model is as good as random; a value of 1 indicates a perfect fit.

Results

Land System characteristics

The resulting LS map for the world and the four regions are shown in Figs. 2 and 3, respectively. Characteristics of LS, analyzed at a global scale, are summarized in Table 3 and Appendix S2.

Cropland systems

Cropland systems cover about 8% of the world land surface and are characterized by an average cropland cover of about 70% (Table 3). Although these systems are dominated by croplands they still contain fair amounts of other land cover such as trees and barren land. Also, they can contain a significant number of livestock. Cropland systems show similar average yields of wheat, rice, and maize combined compared to other LS, but the harvested area of these crops in cropland systems is significantly larger (Fig. 4; Appendix S2), and hence, total crop production will be higher (per land-use intensity level). The highest yields are attained in most intensively managed cropland systems; most often high management efficiencies coincide with areas that have the highest potential productivity. Despite of the relatively low world coverage, 28% of the world population lives in cropland systems, where mixed crop-livestock systems are on average more densely populated compared to cropland systems with few livestock (Fig. 5; Appendix S2).

The nine cropland systems are distinguished based on livestock type and density, and agricultural intensity (Fig. 1). Most extensive cropland systems occur especially in Africa and India, whereas intensive cropland systems occur especially in central-eastern United States, Europe, south-western Russia, north-eastern China, and north India.

Mosaic cropland and grassland systems

Mosaic cropland and grassland systems cover about 5% of the world land surface. These systems have a cropland cover of on average 30–35% and a low tree cover (~5%). Similar to cropland systems, the highest yields are attained in intensively managed systems, and the mixed cropland-livestock systems are most densely populated. In total, 10% of the world population lives in mosaic cropland and grassland systems.

Mosaic cropland and grassland systems are subdivided based on livestock type and density (Fig. 1). Cropland and grassland systems with few livestock are further subdivided based on agricultural intensity. Extensive mosaic cropland and grassland systems occur especially in Africa, while more intensively managed systems occur in the United States, Argentina, Europe, and on the border between Russia and Kazakhstan. Mosaic crop- and grassland systems with livestock, occur especially in China, India, and Africa.

Mosaic cropland and forest systems

Mosaic cropland and forest systems cover 4% of the world land surface. The average cropland cover percentages are 30–35%, with an average tree cover of ~35%. 9% of the world population lives in mosaic cropland and forest systems.

Mosaic cropland and forest systems are subdivided based on agricultural intensity. A separate class cropland and forest with pigs and poultry exists, which occurs especially in South-East Asia. Cropland and forest systems with few livestock occur all over the world, with extensive systems predominantly found in Africa and Central America, and more intensive systems found in Europe, North America, South America, and Southeast Asia.

Forest systems

Forest systems cover 21% of the world land surface. These systems have an average tree cover of about 55% for open forest systems and about 80% for dense forest systems. Despite the large world coverage, only 8% of the world population lives in forest systems (Fig. 5). The highest mean population densities occur in open forest with pigs and poultry (~119 nr km^{-2}). Open forest systems originate both from natural processes (climate, soil), at the edge of dense forests, and from human intervention such as shifting cultivation. Open

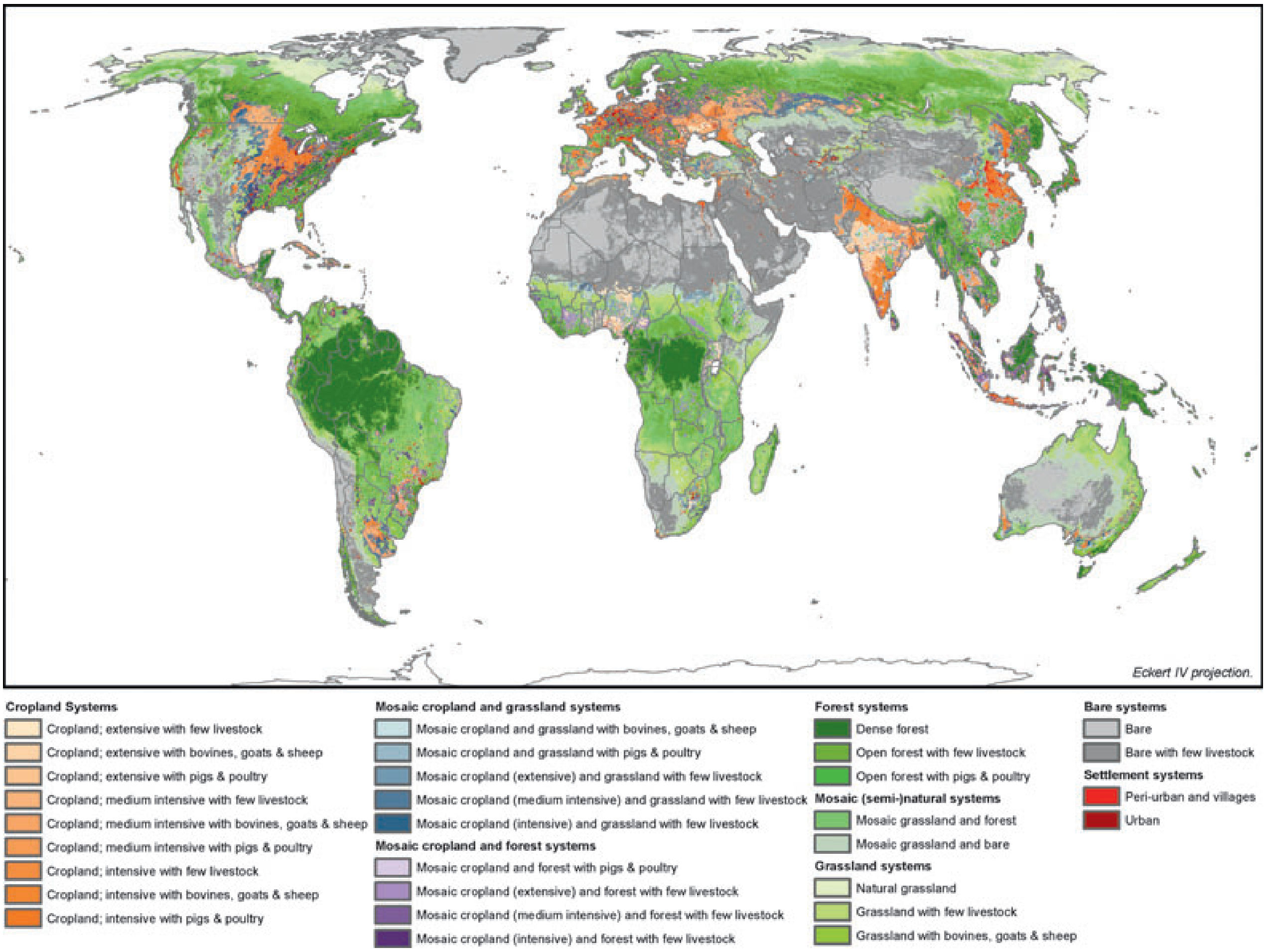

Fig. 2 Global Land System classification map (5 arcminute resolution).

forest systems with pigs and poultry have on average 8% cropland coverage, which may concern both shifting cultivation and permanent cultivation in forests with low population pressure. Open forest systems with few livestock and dense forest have on average <3% cropland cover.

Dense forest systems mostly concern tropical forests, and to a lesser extent, temperate forests at higher latitudes. Open forest systems with pigs and poultry especially occur in China, Japan, and Southeast USA. Open forest systems with few livestock not only occur especially in subarctic regions, where the open character of the landscape is mainly a result of the natural variation in soil and climate conditions, but are also found in other parts of the world (South-Russia, Central Africa, South America, and Southeast Asia).

Grassland systems

Grassland systems cover about 12% of the world land surface, which is mainly attributed to grassland with few livestock (8%). This relatively large surface area explains that still 4.6% of the world population lives in this system, in spite of the low population density (Fig. 5). The two other grassland systems are natural grassland and grassland with bovines, goats, and sheep. Grassland systems have a low average tree and cropland cover (<10%). The yields of wheat, rice and maize combined are similar compared to yields in other LS, but because cropland areas are very small, the total crop production is low (Fig. 4).

Natural grassland systems occur in arctic regions (tundra). Grasslands with few livestock occur all over the world. Grassland with bovines, goats and sheep occur especially in Uruguay, southern Brazil, and Central-East Africa.

Mosaic (semi-)natural systems

The two mosaic (semi-)natural systems are very important and widely spread LS, together covering about 24% of the world land surface. The grassland and forest

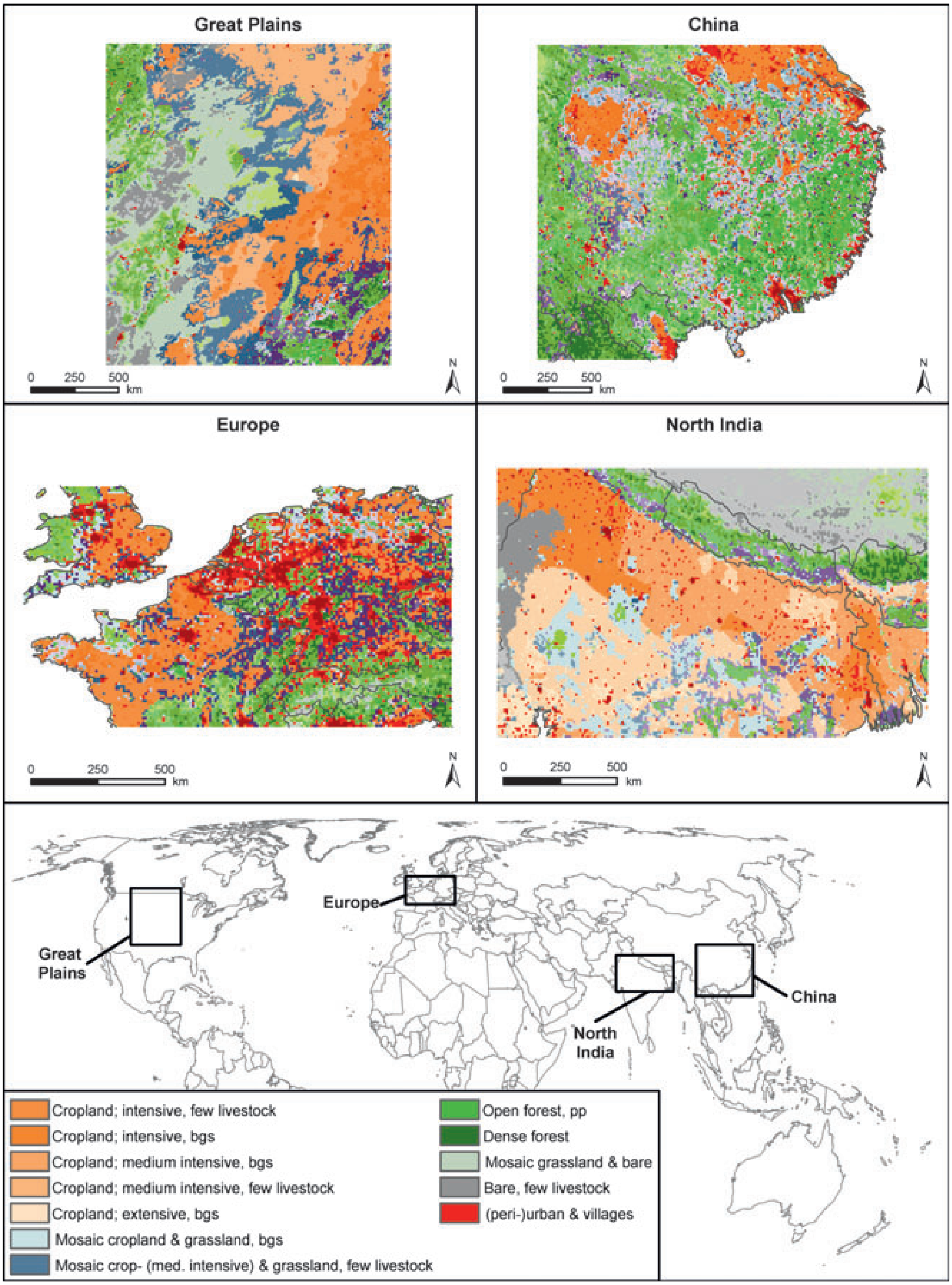

Fig. 3 Land System classification and location of the four regions. For the complete legend see Fig. 2.

system has an average tree cover of ~27%, and low livestock densities. Despite the relatively low average population density, still 8% of the world population lives in this LS (Fig. 5). The grassland and bare system is characterized by an average bare soil coverage of ~35%, and also by low livestock densities. Only 1.5% of the world population lives in this LS.

Mosaic grassland and forest systems especially occur in Canada and Russia (boreal woods), and in South America, Central Africa, and China (often savannas).

Table 3 Average values of land cover and land-use variables used in the classification per Land System, standard deviation between brackets. bgs = bovines, goats, and sheep, pp = pigs and poultry

Land System: Main system	Land System: Subdivisions	World coverage (%)	Cropland (%)	Tree (%)	Bare (%)	Built-up (%)	Livestock *pp* (nr km^{-2})	Livestock *bgs* (nr km^{-2})	Efficiency (–)
Cropland systems	Cropland; extensive, few livestock	0.8	69.8 (*14.1*)	15.8 (*15.5*)	10.6 (*14.0*)	0.3 (*0.8*)	85 (*96*)	33 (*29*)	0.3 (*0.1*)
	Cropland; extensive, bgs	0.5	72.9 (*14.8*)	7.5 (*8.6*)	14.5 (*11.9*)	0.4 (*0.9*)	210 (*430*)	198 (*163*)	0.3(*0.1*)
	Cropland; extensive, pp	0.1	67.2 (*14.0*)	19.3 (*18.1*)	7.2 (*11.0*)	0.3 (*0.9*)	979 (*1882*)	42 (*28*)	0.3 (*0.1*)
	Cropland; medium intensive, few livestock	2.4	67.7 (*12.0*)	10.9 (*14.1*)	11.4 (*10.9*)	0.2 (*0.7*)	84 (*100*)	22 (*21*)	0.6 (*0.1*)
	Cropland; medium intensive, bgs	0.6	74.9 (*14.8*)	6.6 (*8.3*)	11.1 (*11.1*)	0.4 (*0.9*)	386 (*825*)	259 (*184*)	0.5 (*0.1*)
	Cropland; medium intensive, pp	0.4	66.2 (*12.6*)	15.5 (*16.4*)	7.4 (*9.7*)	0.5 (*1.0*)	1076 (*1073*)	37 (*25*)	0.6 (*0.1*)
	Cropland; intensive, few livestock	1.6	69.3 (*12.6*)	10.8 (*13.7*)	6.6 (*9.4*)	0.4 (*0.9*)	87 (*105*)	24 (*23*)	0.8 (*0.1*)
	Cropland; intensive, bgs	0.7	72.3 (*13.0*)	5.1 (*7.0*)	9.0 (*9.9*)	0.8 (*1.3*)	1437 (*3093*)	247 (*184*)	0.8 (*0.1*)
	Cropland; intensive, pp	0.6	69.0 (*12.9*)	11.3 (*14.7*)	3.6 (*5.3*)	0.6 (*1.1*)	1824 (*2431*)	39 (*26*)	0.8 (*0.1*)
Mosaic cropland & grassland systems	Cropland & grassland, bgs	0.8	34.4 (*8.9*)	5.1 (*4.4*)	14.5 (*13.4*)	0.3 (*0.8*)	449 (*1578*)	187 (*158*)	0.5 (*0.2*)
	Cropland & grassland, pp	0.6	35.6 (*8.7*)	6.8 (*4.6*)	7.2 (*9.9*)	0.5 (*1.1*)	1395 (*1830*)	42 (*25*)	0.7 (*0.2*)
	Crop- (extensive) & grassland, few livestock	0.8	31.1 (*8.3*)	5.7 (*4.9*)	17.7 (*14.2*)	0.2 (*0.6*)	55 (*76*)	40 (*30*)	0.3 (*0.1*)
	Crop- (med. intensive) & grassland, few livestock	1.6	34.4 (*8.8*)	5.2 (*4.3*)	16.8 (*12.8*)	0.2 (*0.7*)	59 (*88*)	26 (*26*)	0.6 (*0.1*)
	Crop- (intensive) & grassland, few livestock	0.8	34.6 (*8.8*)	6.0 (*4.6*)	11.0 (*13.2*)	0.4 (*1.0*)	61 (*92*)	32 (*27*)	0.8 (*0.1*)
Mosaic cropland & forest systems	Cropland & forest, pp	1.0	33.0 (*8.6*)	37.9 (*17.1*)	0.9 (*2.6*)	0.4 (*0.9*)	1448 (*2736*)	48 (*82*)	0.6 (*0.2*)
	Cropland (extensive) & forest, few livestock	1.1	30.9 (*8.4*)	33.9 (*16.0*)	1.1 (*2.8*)	0.1 (*0.5*)	71 (*90*)	32 (*71*)	0.2 (*0.1*)
	Cropland (med. intensive) & forest, few livestock	1.4	32.2 (*8.4*)	38.5 (*18.1*)	2.0 (*4.1*)	0.2 (*0.7*)	97 (*110*)	34 (*75*)	0.6 (*0.1*)
	Cropland (intensive) & forest, few livestock	0.9	32.0 (*8.4*)	38.6 (17.2)	1.2 (*2.9*)	0.3 (*0.9*)	81 (*105*)	33 (*68*)	0.8 (*0.1*)
Forest systems	Dense forest	8.1	1.7 (*3.7*)	78.4 (*5.1*)	0.2 (*1.4*)	0.0 (*0.3*)	61 (*404*)	4 (*30*)	–
	Open forest, few livestock	11.8	2.6 (*4.5*)	54.5 (*8.6*)	1.2 (*3.0*)	0.1 (*0.4*)	31 (*65*)	10 (*48*)	–

Table 3 (continued)

Land System: Main system	Land System: Subdivisions	World coverage (%)	Cropland (%)	Tree (%)	Bare (%)	Built-up (%)	Livestock *pp* (nr km^{-2})	Livestock *bgs* (nr km^{-2})	Efficiency (–)
	Open forest, pp	0.8	8.0 (*5.6*)	55.0 (*8.3*)	0.5 (*1.9*)	0.2 (*0.7*)	1207 (*2128*)	35 (*87*)	–
Grassland systems	Natural grassland	3.2	0.0 (*0.0*)	5.9 (*4.3*)	13.1 (*6.2*)	0.0 (*0.0*)	0.0 (*0.0*)	0.0 (*0.0*)	–
	Grassland, few livestock	8.0	4.9 (*5.7*)	7.9 (*4.6*)	10.1 (*7.5*)	0.1 (*0.5*)	70 (*403*)	21 (*25*)	–
	Grassland, bgs	0.8	7.9 (*6.0*)	8.1 (*4.6*)	5.9 (*6.8*)	0.1 (*0.6*)	182 (*726*)	204 (*197*)	–
Mosaic (semi-) natural systems	Grassland and forest	15.2	3.8 (*5.3*)	27.1 (*7.2*)	3.3 (*4.5*)	0.1 (*0.4*)	83 (*540*)	21 (*59*)	–
	Grassland and bare	8.3	2.6 (*4.6*)	2.5 (*4.4*)	35.9 (*7.0*)	0.0 (*0.3*)	22 (*250*)	17 (*62*)	–
Bare systems	Bare	11.6	0.3 (*3.0*)	0.2 (*2.0*)	90.7 (*15.4*)	0.0 (*0.2*)	0 (*0.2*)	0 (*0.2*)	–
	Bare, few livestock	13.6	3.2 (*10.9*)	0.1 (*1.3*)	87.3 (*16.5*)	0.0 (*0.3*)	53 (*721*)	22 (*108*)	–
Settlement systems	Peri-urban & villages	1.6	33.8 (*27.1*)	17.1 (*19.3*)	11.4 (*20.5*)	10.8 (*5.1*)	762 (*2301*)	60 (*141*)	0.7 (*0.2*)
	Urban	0.4	20.6 (*22.6*)	14.0 (*16.2*)	12.8 (*18.4*)	48.2 (*20.6*)	563 (*2403*)	33 (*129*)	0.7 (*0.2*)

Mosaic grassland and bare systems occur all over the world, especially surrounding large deserts.

Bare systems

Bare systems cover 25% of the world land surface. But, only 5% of the world population lives in these systems. The average bare cover is about 90%. Due to irrigation, yields of wheat, maize, and rice can be relatively high in these systems, but only a very low proportion of the area is used for cultivation (Fig. 4).

There are two bare systems; bare systems with and without livestock. Bare systems with livestock have relatively low livestock densities compared to other LS. Bare systems without livestock occur especially in the Sahara, western China, and Australia. Bare systems with few livestock occur in the same areas, as well as in the Middle East region, Mongolia, Kazakhstan, South Argentina, and Western United States.

Settlement systems

Settlement systems cover only 2% of the world land surface, although 25% of the world population lives in this LS (Fig. 5). Settlement systems are subdivided into peri-urban and village (on average ~11% built-up area and ~360 people km^{-2}) and urban systems (on average ~48% built-up area and ~1329 people km^{-2}). The livestock density is relatively high in both classes, slightly higher in the class peri-urban and villages (Table 2). The average cropland cover in the class urban is 21%, whereas in the class peri-urban and villages the average cropland cover is 34%. Hence, there is a lot of agricultural activity in such systems. Crop yields are relatively high compared to most other LS (Fig. 4). Both classes occur all over the world, but are especially found in India, Southeast Asia, Europe, and Eastern United States.

Spatial determinants of Land Systems

The results of the binominal logistic regressions performed at a global scale are summarized in Table 4 and Appendix S3. Regression results for the 10 most important LS in Europe, the Great Plains, China, and north India are summarized in Tables 5, 6, 7 and 8 respectively.

Spatial determinants at a global scale

Cropland systems

Markets are very important for farmers because here they can sell products and buy agricultural inputs (Keys & McConnell, 2005; Lambin & Meyfroidt, 2010;

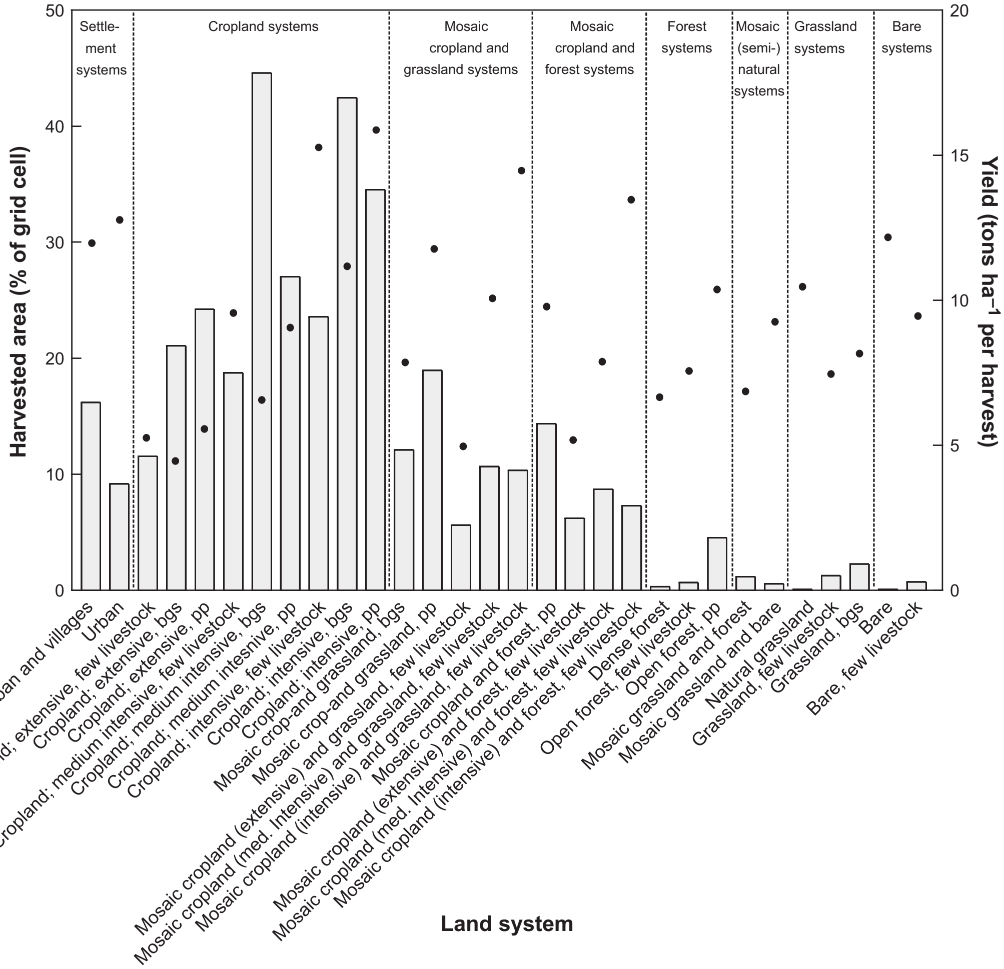

Fig. 4 Sum of the average harvested area (bars) and average yield (black dots) of wheat, rice and maize (bars) for all Land Systems.

Verburg *et al.*, 2011a). Hence, compared to other LS (mosaic) cropland systems are relatively easily accessible (high market accessibility; Table 4).

Especially the more intensively managed (mosaic) cropland systems are located in regions with soil characteristics that are beneficial for growing crops; i.e., soils with relatively high clay and/or silt content, and a relatively high pH and/or CEC. Soil conditions of extensively managed (mosaic) cropland systems are less favorable, for example, extensive cropland with bovines, goats, and sheep is negatively related to organic-matter content. Generally, (mosaic) cropland systems occur on relatively flat and low-lying areas, which are easy to manage and access.

Intensively managed cropland systems require high input costs, for example fertilizers, pesticides, and machinery. These systems are mainly found in regions with a high market influence (positive relation; Table 4). The causality of this relation may work in both directions as high production also provides returns that may be invested in the land.

The most obvious difference between (mosaic) cropland systems with and without livestock is that systems with livestock are positively related with population density.

Forest systems

The distribution of forest systems is especially determined by the potential natural vegetation classification as it captures the envelop of natural systems. Forests often grow in wet areas on acidic soils. These relations reflects the occurrence of large tropical forest regions where precipitation levels are high and soils are old and deeply leached, having poor suitability for agricultural use. Also in many temperate and boreal forests precipitation levels are still relatively high at the global scale, and soils are relatively acid because of leaching and, in case of boreal forest soils, because of cold and often waterlogged conditions that slow down decomposition processes. All forest systems remain especially in sparely populated regions. Still, open forest systems

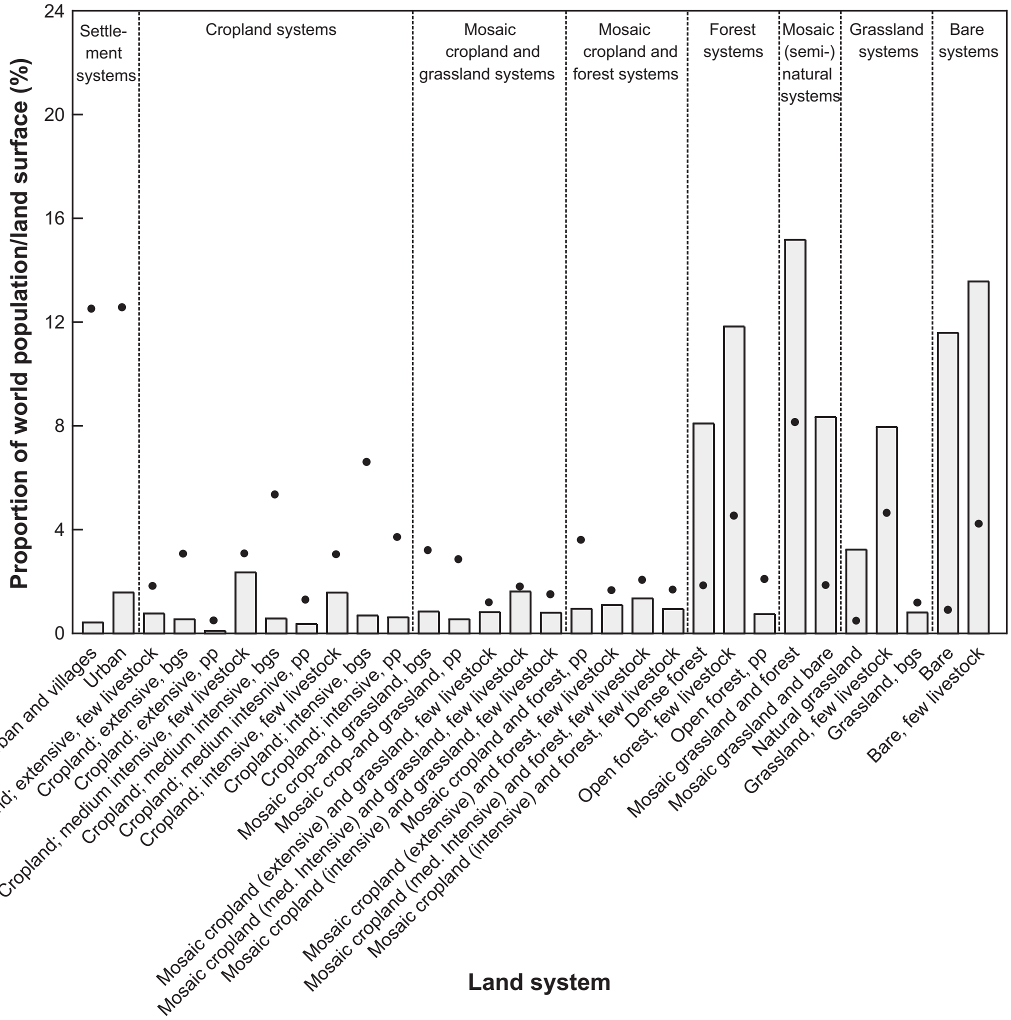

Fig. 5 World coverage (bars) and proportion of the total world population (black dots) for all Land Systems.

with livestock are relatively easy to access on a global scale.

Grassland systems

Natural grassland systems, without cropland and livestock, occur in (sub)arctic regions, which are inaccessible cold regions where the natural vegetation is tundra and mixed woodland. Mosaic grassland and forest systems comprise both boreal and mixed woodlands, and savannas, and have a relatively low population density. These systems receive relatively much precipitation for trees to grow. In contrast, mosaic grasslands and bare systems occur in dry regions where the potential natural vegetation is tundra, open shrubland or grassland and steppe. Grassland systems with few livestock and grassland with bovines, goats, and sheep mainly occur in savanna and steppe regions, where last-mentioned system generally occurs in relatively warm and accessible regions. However, although these areas dominate globally extensive grassland areas are also found in the much colder central Asia region.

Bare systems

Bare systems obviously occur in dry regions (negative relation with precipitation), where the potential natural vegetation is tundra, open shrub or barren land. Bare systems with no livestock occur in inaccessible regions.

Settlement systems

As expected, the distribution of (peri-)urban and village systems is explained by the population density and market accessibility (positive relation). Also, these systems are found in relatively flat areas. Further variation in these systems is most likely the result of other factors not included in our analysis.

Table 4 Spatial determinants and ROC of all Land Systems at a global scale. bgs = bovines, goats, and sheep, pp = pigs and poultry, '+' (light gray) and '-' (dark gray) indicate the sign of the regression coefficient

		Soil parameters						Terrain		Climate		Socio-economic factors			Other		
Land System	Area (% of total area)	Clay (%)	Silt (%)	Sand (%)	CEC (–)	pH (–)	Organic content (%)	Altitude (m)	Slope (°)	Temperature (°C)	Precipitation (mm)	Market accessibility (0–1)	Population density (nr km^{-2})	Market influence ($/person)	Drainage (a–g)*	Potential natural vegetation (1–12)**	ROC
Cropland extensive, few livestock	0.8					+			–			+				+ (1,2,7,8)	0.82
Cropland extensive, bgs	0.5				+		–			+		+	+			+ (7,8,10)	0.94
Cropland extensive, pp	0.1		+							+		+				+ (1,2)	0.87
Cropland med. intensive, few livestock	2.4	+	+			+		–	–			+	–			+ (1,4,7,8)	0.84
Cropland med. intensive, bgs	0.6		+		+					+		+	+			+(2,9)	0.95
Cropland med. intensive, pp	0.4	+	+			+		–			+	+	+				0.83
Cropland intensive, few livestock	1.6		+			+		–	–		+			+		+(4,7,8)	0.94
Cropland intensive, bgs	0.7		+			+			–	+			+	+			0.96
Cropland intensive, pp	0.6		+			+		–	–		+		+	+			0.93
Mosaic cropland & grassland, bgs	0.8	+				+				+		+	+			+ (2,4,7)	0.83
Mosaic cropland & grassland, pp	0.6		+					–		+	+	+	+				0.86
Mosaic cropland (extensive) & grassland, few livestock	0.8	+	–				–					+				+ (7,8)	0.81
Mosaic cropland (med. intensive) & grassland, few livestock	1.6		+			+			–			+				+ (7,8,9)	0.83
	0.8					+			–					+		+ (4,8)	0.89

Table 4 (continued)

		Soil parameters						Terrain		Climate		Socio-economic factors			Other		
Land System	Area (% of total area)	Clay (%)	Silt (%)	Sand (%)	CEC (–)	pH (–)	Organic content (%)	Altitude (m)	Slope (°)	Temperature (°C)	Precipitation (mm)	Market accessibility (0–1)	Population density (nr km^{-2})	Market influence ($/person)	Drainage (a–g)*	Potential natural vegetation (1–12)**	ROC
Mosaic cropland (intensive) & grassland, few livestock																	
Mosaic cropland & forest, pp	1.0		+			–	+	–	+	+	+	+				+ (3,4)	0.88
Mosaic cropland (extensive) & open forest, few livestock	1.1					–	+		+	+	+	+			+(d,e)		0.81
Mosaic cropland (med. intensive) & forest, few livestock	1.4		+								+	+				+ (3,4,6)	0.78
Mosaic cropland (intensive) & forest, few livestock	0.9		+			–					+		–	+		+ (3,4)	0.91
Dense forest	8.1					–					+	–	–			+ (1,3,4,6)	0.93
Open forest, few livestock	11.8					–					+	–	–			+ (2,3,4, 5,6)	0.84
Open forest, pp	0.8		+			–				+	+	+				+ (3,4,6)	0.87
Mosaic grassland & forest	15.2					–	+				+	–	–			+ (5,6,7)	0.75
Mosaic grassland & bare	8.3						–				–	–	–			+ (8,10,11)	0.80
Natural grassland	3.2									–	+	–				+ (6,11)	0.91
Grassland, few livestock	8.0					+	–					–				+ (6,7,8)	0.69
Grassland, bgs	0.8					+				+		+				+ (4,7,8)	0.76
Bare	11.6										–	–				+ (11,12)	0.89
Bare, few livestock	13.6						–				–					+ (10,12)	0.91
	1.6								–			+	+				0.90

Table 4 (continued)

Land System	Area (% of total area)	Soil parameters Clay (%)	Silt (%)	Sand (%)	CEC (–)	pH (–)	Organic content (%)	Terrain Altitude (m)	Slope (°)	Climate Temperature (°C)	Precipitation (mm)	Socio-economic factors Market accessibility (0–1)	Population density (nr km^{-2})	Market influence ($/person)	Other Drainage (a–g)*	Potential natural vegetation (1–12)**	ROC
Peri-urban & villages																	
Urban	0.4								–			+	+				0.97

*Drainage classes: a = very poorly drained, b = poorly drained, c = imperfectly drained, d = moderately drained, e = well drained, f = somewhat excessively drained, g = excessively drained.
**Potential natural vegetation: 1 = tropical evergreen woodland, 2 = tropical deciduous woodland, 3 = temperate evergreen woodland, 4 = temperate deciduous woodland, 5 = boreal woodland, 6 = mixed woodland, 7 = savanna, 8 = grassland and steppe, 9 = dense shrubland, 10 = open shrubland, 11 = tundra, 12 = desert and barren.

Spatial determinants at a regional scale

The regressions analysis shows that the set of variables determining the spatial distribution of LS differs among regions and scale (Tables 5–8). Even a number of contradicting relations between some spatial determinants and LS are found. Contradicting relations are usually caused by different ranges of values per region or scale; a high value in one region may be relatively low in another region.

Cropland systems

Both globally and in the different regions, terrain is an important spatial determinant of cropland systems. Especially intensively managed cropland systems are found in flat low-lying areas. On a global scale (mosaic) cropland systems are generally accessible areas. On a regional scale, accessibility is a less important spatial determinant of (mosaic) cropland systems. Some of these systems are even negatively related to market accessibility, like in India and in the Great Plains.

At a global scale, many mixed (mosaic) cropland-livestock systems occur in densely populated areas. In the region analyzed in Europe, however, population density is relatively high in general, and (mosaic) cropland systems (with and without livestock) occur in areas with a relatively low population density.

Many cropland systems are positively related with precipitation at a global scale. In Europe and at the Great Plains, however, many cropland systems occur in relatively dry areas. In these regions, extremely high precipitation levels occur in mountainous areas (the Alps, Rocky Mountains), along the west coast of the British Isles and Southeast of the Great Plains. Compared to these extremely wet areas, most cropland systems occur in relatively dry parts of these regions, although precipitation levels may still be relatively high on a global scale.

Forest systems

In general, forest systems have similar spatial determinants in the different regions and at a global scale. Dense forests and forests with few livestock usually occur in wet areas with woodland where the population density is low. Forest soils often have a low pH. In Europe and China, forest mainly occurs on higher altitudes and steep slopes (altitude and slope are correlated), where precipitation levels are relatively high. In India and the Great Plains, the distribution of steep slopes is an important spatial determinant for the occurrence of open forest with few livestock.

Table 5 Spatial determinants and ROC for the 10 most important Land Systems in Europe

		Soil parameters						Terrain		Climate		Socio-economic factors			Other		
Land System	Area (% of total area)	Clay (%)	Silt (%)	Sand (%)	CEC (–)	pH (–)	Organic content (%)	Altitude (m)	Slope (°)	Temperature (°C)	Precipitation (mm)	Market accessibility (0–1)	Population density (nr km^{-2})	Market influence ($/person)	Drainage (a–g)*	Potential natural vegetation (1–12)**	ROC
Cropland; intensive, few livestock	17		+			+				+	–		–				0.80
Cropland; intensive, pp	5		+			+		–					–				0.83
Mosaic cropland (intensive) & grassland, few livestock	5	+							–		–		–			+ (4)	0.66
Mosaic cropland (intensive) & forest, few livestock	9					+		+			–	+	–			+ (4)	0.66
Mosaic cropland & forest, pp	4									+			–		+ (d,e)	+ (4,6)	0.64
Grassland, few livestock	4							+				–	–				0.68
Mosaic grassland & forest	9								+		+		–				0.77
Open forest, few livestock	9					–		+			+		–			+ (4,6)	0.80
Peri-urban & villages	16								–			+	+				0.75
Urban	5								–			+	+				0.89

*Drainage classes: a = very poorly drained, b = poorly drained, c = imperfectly drained, d = moderately drained, e = well drained, f = somewhat excessively drained, g = excessively drained.
**Potential natural vegetation: 1 = tropical evergreen woodland, 2 = tropical deciduous woodland, 3 = temperate evergreen woodland, 4 = temperate deciduous woodland, 5 = boreal woodland, 6 = mixed woodland, 7 = savanna, 8 = grassland and steppe, 9 = dense shrubland, 10 = open shrubland, 11 = tundra, 12 = desert and barren.

Table 6 Spatial determinants and ROC for the 10 most important Land Systems in the Great Plains, United States

		Soil parameters						Terrain		Climate		Socio-economic factors			Other		
Land System	Area (% of total area)	Clay (%)	Silt (%)	Sand (%)	CEC (–)	pH (–)	Organic content (%)	Altitude (m)	Slope (°)	Temperature (°C)	Precipitation (mm)	Market accessibility (0–1)	Population density (nr km^{-2})	Market influence ($/person)	Drainage (a–g)*	Potential natural vegetation (1–12)**	ROC
Cropland; med. intensive, few livestock	14		+						–		–	–				+ (7,8)	0.89
Mosaic cropland (med. intenisve) & grassland, few livestock	12				+	+			–		–	–				+ (8)	0.79
Cropland; intensive, few livestock	15						+		–				–	+		+ (7,8)	0.90
Cropland; intensive, pp	4							–	–					+		+ (7,8)	0.86
Mosaic cropland (intensive) & grassland, few livestock	6								–	+				+		+ (8)	0.74
Mosaic grassland & bare	16						–	+			–			–		+ (8,10)	0.87
Grassland, few livestock	6				–			+						–		+ (8)	0.75
Mosaic grassland & forest	6							+	+		+					+ (3)	0.90
Open forest, few livestock	6								+		+				+ (d)	+ (3,4,5)	0.95
Bare, few livestock	5							+			–					+ (10)	0.95

*Drainage classes: a = very poorly drained, b = poorly drained, c = imperfectly drained, d = moderately drained, e = well drained, f = somewhat excessively drained, g = excessively drained.
**Potential natural vegetation: 1 = tropical evergreen woodland, 2 = tropical deciduous woodland, 3 = temperate evergreen woodland, 4 = temperate deciduous woodland, 5 = boreal woodland, 6 = mixed woodland, 7 = savanna, 8 = grassland and steppe, 9 = dense shrubland, 10 = open shrubland, 11 = tundra, 12 = desert and barren.

Table 7 Spatial determinants and ROC for the 10 most important Land Systems in China

		Soil parameters						Terrain		Climate		Socio-economic factors			Other		
Land System	Area (% of total area)	Clay (%)	Silt (%)	Sand (%)	CEC (–)	pH (–)	Organic content (%)	Altitude (m)	Slope (°)	Temperature (°C)	Precipitation (mm)	Market accessibility (0–1)	Population density (nr km^{-2})	Market influence ($/person)	Drainage (a–g)*	Potential natural vegetation (1–12)**	ROC
Cropland; intensive bgs	3								–		–						0.97
Cropland; intensive, pp	6		+						–				+	+			0.89
Mosaic cropland & grassland, pp	8								–			+				+ (3,4)	0.75
Mosaic cropland & forest, pp	10										+			+	+ (e)	+ (3,4)	0.70
Grassland, few livestock	6										–		–			+ (3,4,8)	0.63
Mosaic grassland & forest	21											–	–		+ (e,f)	+ (4.9)	0.69
Open forest, few livestock	12					–			+					–			0.81
Open forest, pp	14					–	+		+				–		+ (e,f)	+ (3)	0.70
Dense forest	4			+				+		+	+	–					0.89
Peri-urban & villages	5								–			+	+				0.83

*Drainage classes: a = very poorly drained, b = poorly drained, c = imperfectly drained, d = moderately drained, e = well drained, f = somewhat excessively drained, g = excessively drained.
**Potential natural vegetation: 1 = tropical evergreen woodland, 2 = tropical deciduous woodland, 3 = temperate evergreen woodland, 4 = temperate deciduous woodland, 5 = boreal woodland, 6 = mixed woodland, 7 = savanna, 8 = grassland and steppe, 9 = dense shrubland, 10 = open shrubland, 11 = tundra, 12 = desert and barren.

Table 8 Spatial determinants and ROC for the 10 most important Land Systems in north India

		Soil parameters						Terrain		Climate		Socio-economic factors			Other		ROC
Land System	Area (% of total area)	Clay (%)	Silt (%)	Sand (%)	CEC (–)	pH (–)	Organic content (%)	Altitude (m)	Slope (°)	Temperature (°C)	Precipitation (mm)	Market accessibility (0–1)	Population density (nr km^{-2})	Market influence ($/person)	Drainage (a–g)*	Potential natural vegetation (1–12)**	
Cropland; extensive, bgs	16					+				+		–			+ (d,e)	+ (2,7,10)	0.81
Cropland; med. intensive, bgs	16							–				+	+		+ (e)	+ (1,2)	0.87
Cropland; intensive, bgs	9					+			–					+		+ (9,10)	0.82
Mosaic cropland & grassland, bgs	9									+		–	–			+ (2,7)	0.76
Mosaic grassland & forest	4								+		+		–			+ (6,7)	0.90
Open forest, few livestock	4								+		+		–			+ (6)	0.95
Mosaic grassland & bare	5											–				+ (8,11,12)	0.95
Bare, few livestock	5										–	–				+ (10, 11,12)	0.94
Bare	9											–	–			+ (11,12)	0.95
Peri-urban & villages	3								–			+	+				0.77

*Drainage classes: a = very poorly drained, b = poorly drained, c = imperfectly drained, d = moderately drained, e = well drained, f = somewhat excessively drained, g = excessively drained.

**Potential natural vegetation: 1 = tropical evergreen woodland, 2 = tropical deciduous woodland, 3 = temperate evergreen woodland, 4 = temperate deciduous woodland, 5 = boreal woodland, 6 = mixed woodland, 7 = savanna, 8 = grassland and steppe, 9 = dense shrubland, 10 = open shrubland, 11 = tundra, 12 = desert and barren.

Grassland systems

At a global scale, mosaic grassland and forest systems occur in relatively wet areas with a low population density, where the potential natural vegetation is often savanna, or mixed- or boreal woodland. This LS often occurs on soils with a low pH. At a regional scale, other variables are important spatial determinants as well. For example, in the Great Plains and in north India this LS is mainly found on relatively steep slopes. A difference between these areas is that in India it mostly concerns savanna-like systems, whereas in the Great Plains it mostly occurs in temperate and boreal woodland systems. In China, this LS is especially found in inaccessible areas with well drained soils where the potential natural vegetation is temperate deciduous woodland or dense shrub.

Mosaic grassland and bare systems occur in dry areas with a low population density, where the potential natural vegetation is tundra, open shrubland or grassland and steppe. In India, this LS mainly occurs on the dry and cold tundra's on the Tibetan high plateau, but at the Great Plains this LS mainly occurs on dry (and warm) open shrublands or steppes, also at high altitudes.

Globally, extensive grassland systems occur in areas where the potential natural vegetation is mixed woodland, grassland and steppe, or savanna. Market accessibility and influence are usually low. In Europe this LS occurs predominantly on high altitudes in areas with a low population density. In the Great Plains this LS also occurs on high altitudes (and relatively steep slopes) in areas with a low market influence, on soils with a low CEC.

Bare systems

At the global scale, the distribution of bare systems is mainly determined by the occurrence of tundra, open shrubland and, desert and barren land. These are dry regions, where bare land without livestock occurs in inaccessible areas. At the regional-scale analysis, bare systems with or without livestock occur in north India and the Great Plains. In north India, bare systems without livestock mainly occur on the cold and dry on the Tibetan high plateau, where market accessibility and influence are low. At the Great Plains, bare systems with livestock occur in dry open shrubland areas on relatively high altitudes.

Settlement systems

Both regionally and globally, peri-urban and village and urban systems occur in relatively flat regions that are easily accessible (close to markets) and have a high population density.

Discussion

Land system classification

In many existing global land cover datasets one grid cell represents one land cover type, whereas in reality, cells often represent mosaics of different land cover/use types. Ignoring the heterogeneity of land cover may lead to an under- or overestimation of the actual coverage of specific land covers, which may have serious implications for example climate change assessments (Verburg *et al.*, 2011b). Global scale analyses, for which usually medium to coarse scale data are used, are especially susceptible to this. Some recent land cover datasets have acknowledged the heterogeneous character of landscapes by including mosaic classes (Bartholomé & Belward, 2005; Bontemps *et al.*, 2011). On a micro- to meso-scale, farming system classifications mostly represent (crop-)livestock systems, but these are classifications of management entities confined to agricultural systems and no global-scale spatial explicit farming system maps are currently available. The LS classification presented in this article incorporates farming system elements, while at the same time delineating (semi-)natural land systems globally, including mosaics of agricultural systems and (semi-)natural vegetation. The new LS classification method incorporates sub-pixel information on land cover, type and density of livestock, and agricultural intensity to delineate land systems. Although attaining a hugely improved representation of land use, care must be taken since the scale of the input data used for the LS classification may still be too coarse to detect small landscape features that are, however, important for the dynamics of for example hydrological or climatic systems (Nol *et al.*, 2008; Ellis *et al.*, 2009). The LS classification is designed for land-change modeling and assessment purposes, and hence, important variables that are often used as driving factors in land-change models, such as population density, are not used as a classification factor but rather as a spatial determinant of the LS instead to ensure independence between the dependent variable (LS) and the independent location factors and drivers of change in LS.

Population density is useful as an indicator for the intensity of human-environment interactions and high population densities are expected to correlate with intensively managed land systems (Boserup, 1965). However, the results of our study show that intensively managed (mosaic) cropland systems do not necessarily have a high population density. The results rather show that, on average, especially (mosaic) cropland systems with livestock have a high population density (Fig. 6). Most land systems have a wide range of population

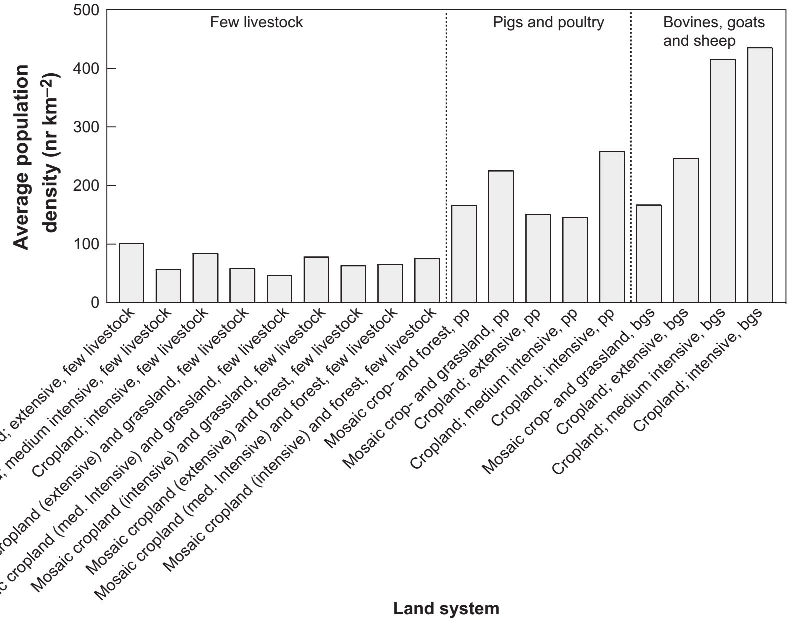

Fig. 6 Average population density of (mosaic) cropland land systems. pp = pigs & poultry, bgs = bovines, goats, and sheep.

density values. Thus, an extensive cropland system with livestock may still be densely populated. Mixed cropland-livestock LS with a high population density occur for example extensively in India (cropland with bovines, goats, and sheep) and China (cropland with pigs and poultry). Cropland systems with few livestock are often located outside densely populated areas but supply a large proportion of the worlds agricultural production (e.g., at the Great Plains, United States). These findings support the notion that population density cannot straightforwardly be used to classify agricultural intensity; population is only one of the multiple determinants of land systems.

Uncertainties in the LS classification especially arise from the quality of input data, which is influenced by (1) the techniques used to process and interpret remote sensing data (Fritz & See, 2008), (2) the original spatial resolution of the data sets, and (3) the quality of census data, which partly depends on different reporting methods of census data (Verburg *et al.*, 2011b). To minimize the influence of such uncertainties, we used the land cover variables with a relatively low level of uncertainty. For example, pasture data were not used as a classification criterion. The main reason for this is that the definition of pasture differs between countries. It is not always clear whether or not forest- and semi-arid land used for grazing is included in pasture inventory datasets. For example, as described by Ramankutty *et al.* (2008), the FAO reports 1.7 million km^2 of pastures in the year 2000 in the mostly arid Saudi Arabia, while sub-national census data only report 486 km^2 of pasture. In other regions grazing does not always occur on pastures but also along roads, in dry river beds and in croplands after harvest (Verburg & Keulen, 1999). Also in India mixed crop-livestock systems are common (Devendra & Thomas, 2001). For these reasons, it was chosen not to rely on current pasture datasets, but only to use land cover (including cropland cover), livestock and agricultural intensity data for the classification. A relatively small number of input variables were used to keep the classification relatively simple. Despite this relative simplicity much information can be extracted from the classification and provide a stratification for global scale environmental (change) assessments. It should be realized that some datasets are not true 5 arcminute datasets, because they are downscaled from (sub)national census data. Examples of these datasets are maps representing crop yield (used for calculating efficiency), market influence and livestock density.

The direct interpretation of the LS classification is obviously influenced by the used classification thresholds (see Appendix S1). Although this does not affect the quality of the results, care should be taken when comparing the LS classification to other land cover/use data sets.

Spatial determinants of Land Systems

Although the selected variables as potential spatial determinants of LS are often listed as important driving factors of land-use change, in many regions land change is the result of the interplay of many more factors acting at different temporal and spatial scales (Lambin *et al*., 2001; Geist & Lambin, 2002; Rudel *et al*., 2005). For example, the influence of institutions (political, legal, economic, and traditional) can be an important determinant of LS (Geist *et al*., 2006). Especially government policy may induce land-use change through subsidies and control of access to land, labor, capital, technology, and information (Lambin & Geist, 2006). Furthermore, cultural factors influence choices made by land managers, and therewith, land-use (change). For example, in India cows are held as sacred animals, which largely explains the high number of bovines in mixed cropland-livestock systems in this country. Limited spatial data are available to characterize such conditions at a global scale. However, including such conditions in explaining global LS patterns is a key priority (Neumann *et al*., 2011).

The results show that the set of factors that determines the spatial distribution of specific LS often differs between regions and scales. This is especially true for LS with a relatively high human impact on the natural environment, like (mosaic) cropland systems, and grassland systems with livestock. Although (semi-)natural land systems often follow globally uniform (biophysical) processes, human interactions with the environment strongly vary by region, depending on the available resources, traditions, and governance. Also the value and sign of the regression coefficients often differ between regions and scales. This may be not only caused by different ranges of a variable in different regions, but also by the possible influence of variables (e.g., governance, cultural) that are not included in the regression analysis, as mentioned above. Also, it can be questioned to what extent the associations between location factors and LS are robust in time; societal change and technology may affect these relations. Hence, the exact value of the coefficients in regression equations in different regions cannot be compared directly. Still, the regional regression analysis does indicate the relative importance of the explanatory factors used in this study when comparing different regions, which is not shown by the regression analysis at the global scale.

Application in land-use change models

Land-change models are used to explore (future) land-use dynamics. Although numerous local to regional scale land-use models exist (e.g., Lambin & Geist, 2006; Matthews *et al*., 2007; Verburg & Overmars, 2009), only few global land-use change models have been developed (Heistermann *et al*., 2006). There is an urgent need to further develop those existing and new global land-use change models, taking stock of the development at regional scale (Rounsevell & Arneth, 2011). Land-change models are often an important component of global IAM's (Verburg *et al*., 2011b). Such models have been increasingly used during the past few decades to assess climate change, biodiversity and energy issues, and need to best represent our understanding of global land change.

The new Land System map provides a new and more integrated classification of land use, and may serve as input for global land change models and other applications for earth system modeling. Advantages of using LS in such models mainly concern the relatively high resolution and representation of (the intensity of) human-environment interactions in mosaic landscapes at a global scale, providing more accurate representation of the interactions in the socio-ecological system. The regression analyses presented in this article demonstrate that the set of factors determining the occurrence of specific LS may differ per region and scale. Most current models use globally uniform allocation algorithms. Our results imply that changes in these biophysical and socioeconomic factors are likely to have different impacts in different regions. Therefore, it is important not to apply uniform drivers of land change but rather apply a region specific parameterization accounting for specific regional determinants of land change. For biophysical processes uniform drivers may hold, but for LS with significant human influence regional parameterization is recommended. This finding is supported by studies of Geist & Lambin (2001, 2002) on tropical deforestation, which showed that drivers of deforestation interact differently per region and that a thorough understanding of these interactions at the regional scale is necessary to generate realistic projections of land-cover change based on simulation models.

The results also indicate that the value of regression coefficients partly depends on the range of the variable in a region; therefore, care should be taken in using these values for predicting future land changes. For each spatial determinant it should be evaluated if and how fast the range may change over time. Depending on this, regression equations should be adapted when predicting land change over long timescales, accounting for the causality in the identified associations. Moreover, as mentioned before, predicting the suitability for land systems at certain locations would probably become more accurate if more spatial determinants are included, such as institutional, governance and cultural data. But, at present, including a wider range of socioeconomic and institutional data in global assessments is a major challenge, mainly as a result of limited data availability.

Acknowledgements

This article is based on research funded by the Netherlands Organization for Scientific Research (NWO; project IGLO). The work presented in this article contributes to the Global Land Project (http:\\www.globallandproject.org). The authors thank Elke Stehfest, Kees Klein Goldewijk, and Tom Kram for discussions on the classification system and Erle Ellis for comments on an earlier version of the paper. We also thank the GCB editor Ivan Janssens and two anonymous reviewers for their comments on an earlier version of this article. The global land system data can be downloaded from http://www.ivm.vu.nl/landsystems.

References

Bartholomé E, Belward AS (2005) GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, **26**, 1959–1977.

Bondeau A, Smith PC, Zaehle S *et al.* (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology*, **13**, 679–706.

Bontemps S, Defourny P, Van Bogaert E, Arino O, Kalogirou V, Ramos Perez J (2011) *Globcover 2009. Products Description and Validation report*. Available at: Ionia ESA website http://ionia1.esrin.esa.int/ (accessed 21 December 2010), 53 pp.

Boserup E (1965) *The Conditions of Agricultural Growth: the Economics of Agrarian Change Under Populaton Pressure*. 108 p. Aldine, Chicago.

Bouwman AF, Kram T, Klein Goldewijk K (2006) *Integrated Modelling of Global Environmental Change. An Overview of IMAGE 2.4*. 228 pp. Netherlands Environmental Assessment Agency (MNP), Bilthoven.

Braimoh AK, Onishi T (2007) Spatial determinants of urban land use change in Lagos, Nigeria. *Land Use Policy*, **24**, 502–515.

Chomitz KM, Gray DA (1996) Roads, land use, and deforestation: a spatial model applied to Belize. *The World Bank Economic Review*, **10**, 487–512.

DeFries R (2008) Terrestrial vegetation in the coupled human-earth system: contributions of remote sensing. *Annual Review of Environment and Resources*, **33**, 369–390.

DeFries R, Rosenzweig C (2010) Toward a whole-landscape approach for sustainable land use in the tropics. *Proceedings of the National Academy of Sciences*, **107**, 19627–19632.

Devendra C, Thomas D (2001) Crop-animal systems in Asia: importance of livestock and characterisation of agro-ecological zones. *Agricultural Systems*, **71**, 5–15.

Dixon J, Gulliver A, Gibbon D, Hall M (2001) *Farming Systems and Poverty: Improving Farmers' Livelihoods in a Changing World*. 412 pp. FAO and World Bank, Rome and Washington D.C.

Eickhout B, van Meijl H, Tabeau A, van Rheenen T (2007) Economic and ecological consequences of four European land use scenarios. *Land Use Policy*, **24**, 562–575.

Ellis EC, Ramankutty N (2008) Putting people in the map: anthropogenic biomes of the world. *Frontiers in Ecology and the Environment*, **6**, 439–447.

Ellis E, Neerchal N, Peng K *et al.* (2009) Estimating long-term changes in China's village landscapes. *Ecosystems*, **12**, 279–297.

Ellis EC, Klein Goldewijk K, Siebert S, Lightman D, Ramankutty N (2010) Anthropogenic transformation of the biomes, 1700 to 2000. *Global Ecology and Biogeography*, **19**, 589–606.

FAO (1999) *Soil Physical Constraints to Plan Growth and Crop Production* (eds Gardner CMK, Laryea KB, Unger PW), 106 pp, FAO, Rome.

FAO (2006) *Livestock's Long Shadow: Environmental Issues and Options* (eds Steinfeld H, Gerber P, Wassenaar T, Catel V, Rosales M, De Haan C), 390 pp, FAO, Rome.

FAO (2007) *Gridded Livestock of the World* (eds Wint GRW, Robinson TP), 141 pp, FAO, Rome.

FAO (2011) *Land Degradation Assessment in Drylands. Mapping Land Use Systems at Global and Regional Scales for Land Degradation Assessment Analysis. Version 1.1* 84 pp, FAO, Rome.

Foley JA, DeFries R, Asner GP *et al.* (2005) Global consequences of land use. *Science*, **309**, 570–574.

Fritz S, See L (2008) Identifying and quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. *Global Change Biology*, **14**, 1057–1075.

Geist HJ, Lambin EF (2001) *What Drives Tropical Deforestation? A Meta-Analysis of Proximate and Underlying Causes of Deforestation Based on Subnational Case Study Evidence*. A LUCC Report Series No. 4., 136 pp, CIACO, Louvain-la-Neuve.

Geist HJ, Lambin EF (2002) Proximate causes and underlying driving forces of tropical deforestation. *BioScience*, **52**, 143–150.

Geist HJ, McConnell WJ, Lambin EF, Moran E, Alves D, Rudel TK (2006) Causes and trajectories of land-use/cover change. In: *Land-use and Land-Cover Change. Local Processes and Global Impacts* (eds Lambin EF, Geist HJ), pp. 41–70, Springer-Verlag, Berlin Heidelberg.

Geoghegan J, Cortina Villar S, Klepeis P (2001) Modeling tropical deforestation in the southern Yucatán peninsular region: comparing survey and satellite data. *Agriculture, Ecosystems and Environment*, **85**, 25–46.

Global Land Cover 2000 database. European Commission, Joint Research Centre, 2003. Available at: http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php (accessed 26 January 2004).

Globcover Land Cover v2. ESA, 2008. Available at: http://ionia1.esrin.esa.int/index.asp (accessed 3 October 2008).

Hansen M, DeFries R, Townshend JR, Carroll M, Dimiceli C, Sohlberg R (2003) *Vegetation Continuous Fields MOD44B*. University of Maryland, College Park.

Hansen MC, Stehman SV, Potapov PV (2010) Reply to Wernick *et al.*: global scale quantification of forest change. *Proceedings of the National Academy of Sciences*, **107**, E148.

Havlík P, Schneider UA, Schmid E *et al.* (2011) Global land-use implications of first and second generation biofuel targets. *Energy Policy*, **39**, 5690–5702.

Heistermann M, Müller C, Ronneberger K (2006) Land in sight?: Achievements, deficits and potentials of continental to global scale land-use modeling. *Agriculture, Ecosystems & Environment*, **114**, 141–158.

Hibbard K, Janetos A, van Vuuren DP *et al.* (2010) Research priorities in land use and land-cover change for the Earth system and integrated assessment modelling. *International Journal of Climatology*, **30**, 2118–2128.

Hurtt GC, Chini LP, Frolking S *et al.* (2011) Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climate Change*, **109**, 117–161.

Keys E, McConnell WJ (2005) Global change and the intensification of agriculture in the tropics. *Global Environmental Change Part A*, **15**, 320–337.

Klein Goldewijk K (2001) Estimating global land use change over the past 300 years: the HYDE database. *Global Biogeochemical Cycles*, **15**, 417–434.

Klein Goldewijk K, Beusen A, Janssen P (2010) Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1. *The Holocene*, **20**, 565–573.

Kruska RL, Reid RS, Thornton PK, Henninger N, Kristjanson PM (2003) Mapping livestock-oriented Agricultural production systems for the developing world. *Agricultural Systems*, **77**, 39–63.

Lambin EF, Geist HJ (2006) *Land-Use and Land-Cover Change; Local Processes and Global Impacts*, 222 pp, Springer, Heidelberg.

Lambin EF, Meyfroidt P (2010) Land use transitions: socio-ecological feedback versus socio-economic change. *Land Use Policy*, **27**, 108–118.

Lambin EF, Turner BL II, Geist HJ *et al.* (2001) The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change*, **11**, 261–269.

Letourneau A, Verburg PH, Stehfest E (2012) A land-use systems approach to represent land-use dynamics at continental and global scales. *Environmental Modelling & Software*, **33**, 61–79.

Licker R, Johnston M, Foley JA, Barford C, Kucharik CJ, Monfreda C, Ramankutty N (2010) Mind the gap: how do climate and agricultural management explain the yield gap of croplands around the world? *Global Ecology and Biogeography*, **19**, 769–782.

Lotze-Campen H, Müller C, Bondeau A, Rost S, Popp A, Lucht W (2008) Global food demand, productivity growth, and the scarcity of land and water resources: a spatially explicit mathematical programming approach. *Agricultural Economics*, **39**, 325–338.

Matthews R, Gilbert N, Roach A, Polhill J, Gotts N (2007) Agent-based land-use models: a review of applications. *Landscape Ecology*, **22**, 1447–1459.

Messerli P, Heinimann A, Epprecht M (2009) Finding homogeneity in heterogeneity – a new approach to quantifying landscape mosaics developed for the lao pdr. *Human Ecology*, **37**, 291–304.

Monfreda C, Ramankutty N, Foley JA (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, **22**, GB1022, doi:10.1029/2007GB002947.

Neumann K, Verburg PH, Stehfest E, Müller C (2010) The yield gap of global grain production: a spatial analysis. *Agricultural Systems*, **103**, 316–326.

Neumann K, Stehfest E, Verburg PH, Siebert S, Müller C, Veldkamp A (2011) Exploring Global Irrigation Patterns: a multilevel modeling approach. *Agricultural Systems*, **104**, 703–713.

Nol L, Verburg PH, Heuvelink GBM, Molenaar K (2008) Effect of Land Cover Data on Nitrous Oxide Inventory in Fen Meadows. *Journal of Environmental Quality*, **37**, 1209–1219.

Pongratz J, Reick C, Raddatz T, Claussen M (2008) A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, **22**, GB3018.

Ramankutty N, Evan AT, Monfreda C, Foley JA (2008) Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, **22**, GB1003, doi: 10.1029/2007GB002952.

Rounsevell MDA, Arneth A (2011) Representing human behaviour and decisional processes in land system models as an integral component of the earth system. *Global Environmental Change*, **21**, 840–843.

Rudel TK, Coomes OT, Moran E, Achard F, Angelsen A, Jianchu X, Lambin EF (2005) Forest transitions: towards a global understanding of land use change. *Global Environmental Change*, **15**, 23–31.

Rudel TK, Schneider L, Uriart M *et al*. (2009) Agricultural intensification and changes in cultivated areas, 1970–2005. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 20675–20680.

Schaldach R, A Priess Jr, Alcamo J (2011) Simulating the impact of biofuel development on country-wide land-use change in India. *Biomass and Bioenergy*, **35**, 2401–2410.

Schmit C, Rounsevell MDA, La Jeunesse I (2006) The limitations of spatial land use data in environmental analysis. *Environmental Science & Policy*, **9**, 174–188.

Schneider A, Friedl MA, Potere D (2009) A new map of global urban extent from MODIS satellite data. *Environmental Research Letters*, **4**, 044003, doi: 10.1088/1748-9326/4/4/044003.

Serneels S, Lambin EF (2001) Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model. *Agriculture, Ecosystems and Environment*, **85**, 61–81.

van de Steeg JA, Verburg PH, Baltenweck I, Staal SJ (2010) Characterization of the spatial distribution of farming systems in the Kenyan Highlands. *Applied Geography*, **30**, 239–253.

Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285–1293.

Thornton PK, Kruska RL, Henninger N, Kristjanson PM, Reid RS, Atieno F, Odero AN, Ndegwa T (2002) *Mapping Poverty and Livestock in the Developing World*, 118 pp, ILRI, Nairobi, Kenya.

Turner BL II (2010) Sustainability and forest transitions in the southern Yucatán: the land architecture approach. *Land Use Policy*, **27**, 170–179.

Turner BL II, Lambin EF, Reenberg A (2007) The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences*, **104**, 20666–20671.

Verburg PH, Keulen HV (1999) Exploring changes in the spatial distribution of livestock in China. *Agricultural Systems*, **62**, 51–67.

Verburg P, Overmars K (2009) Combining top-down and bottom-up dynamics in land use modeling: exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model. *Landscape Ecology*, **24**, 1167–1181.

Verburg PH, van de Steeg J, Veldkamp A, Willemen L (2009) From land cover change to land function dynamics: a major challenge to improve land characterization. *Journal of Environmental Management*, **90**, 1327–1335.

Verburg PH, Ellis EC, Letourneau A (2011a) A global assessment of accessibility and market influence for global environmental change studies. *Global Environmental Change*, **6**, 034019, doi: 10.1088/1748-9326/6/3/034019.

Verburg PH, Neumann K, Nol L (2011b) Challenges in using land use and land cover data for global change studies. *Global Change Biology*, **17**, 974–989.

Von Thünen JH (1966) Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie. In: *Von Thünen's isolated state* (ed Hall P). Pergamion, Oxford.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Hierarchical classification procedure to delineate Land Systems. Classification thresholds are given in the rectangles, pp = pigs & poultry (nr km^{-2}), bgs = bovines, goats & sheeps (nr km^{-2}), eff = efficiency of agricultural production (int1 = extensive system, int2 = medium intensive system, int3 = intensive system), built-up, bare, crop and tree cover in percentages.

Appendix S2. Average yield and harvested area of wheat, rice and maize, and population data per Land System. Standard deviations are given in italic in between brackets. pp = pigs & poultry, bgs = bovines, goats & sheep, ls = livestock.

Appendix S3. Regression coefficients. A binominal logistic regression is used: $\log(P_i/(1-P_i)) = \alpha_0 + \alpha_1 X_{1,I} + \alpha_2 X_{2,i} \ldots + \alpha_n X_{n,i}$, where P_i is the probability of a grid cell for the occurrence of the considered LUS type on location i, X_1 to X_n are the independent variables (driving factors) and α_0 to α_n are coefficients estimated through logistic regression. pp = pigs & poultry, bgs = bovines, goats & sheep, ext = extensive, m. int = medium intensive, int = intensive, ls = livestock. *Drainage classes: a = very poorly drained, b = poorly drained, c = imperfectly drained, d = moderately drained, e = well drained, f = somewhat excessively drained, g = excessively drained. **Biome: 1 = tropical evergreen woodland, 2 = tropical deciduous woodland, 3 = temperate evergreen woodland, 4 = temperate deciduous woodland, 5 = boreal woodland, 6 = mixed woodland, 7 = savanna, 8 = grassland and steppe, 9 = dense shrubland, 10 = open shrubland, 11 = tundra, 12 = desert and barren.