Exploring global irrigation patterns: A multilevel modelling approach

Kathleen Neumann, Elke Stehfest, Peter H. Verburg, Stefan Siebert, Christoph Müller, Tom Veldkamp

Kathleen Neumann (corresponding author)
PBL Netherlands Environmental Assessment Agency
PO Box 303, 3720 AH, Bilthoven, The Netherlands
Email: kathleen.neumann@pbl.nl
Telephone: +31 (30) 274 8694

Elke Stehfest
PBL Netherlands Environmental Assessment Agency
PO Box 303, 3720 AH, Bilthoven, The Netherlands
Email: elke.stehfest@pbl.nl

Peter H. Verburg
Institute for Environmental Studies, VU University Amsterdam,
De Boelelaan 1087, 1081 HV, Amsterdam, The Netherlands
Email: peter.verburg@ivm.vu.nl

Stefan Siebert
Institute of Crop Science and Resource Conservation (INRES), University of Bonn
Katzenburgweg 5, 53115 Bonn, Germany
Email: s.siebert@uni-bonn.de

Christoph Müller
Potsdam Institute for Climate Impact Research (PIK), Telegrafenberg,
PO Box 601203, 14412 Potsdam, Germany
Email: christoph.mueller@pik-potsdam.de

Tom Veldkamp
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente
PO Box 6, 7500 AA Enschede, The Netherlands
Email: veldkamp@itc.nl

Abstract

Globally, areas that are equipped for irrigation have almost doubled in size over the past 50 years and further expansions are expected for the future, to meet a growing food demand. For developing countries, the Food and Agriculture Organization of the United Nations (FAO) expects these areas to be expanded by 40 million ha, by 2030. Knowledge about the constraints to irrigation and spatially explicit information about the potential for irrigation expansion, however, are lacking on a global scale. The objective of our study was to explain the global pattern of irrigated croplands and to identify cropping regions where irrigation is likely to be expanded. We accounted for biophysical determinants, such as humidity and slope, mainly at grid-cell level. Socio-economic and governance determinants, for example, Gross Domestic Product (GDP) and control of corruption, were primarily considered on a country level, given the limitations in availability of sub-national data and the role of national level governance in irrigation decisions. To identify the variability of the determinants within these two spatial levels, we conducted a multilevel analysis. This is a method employing regression models that explicitly account for hierarchically structured data. Results show significant variability in terms of irrigation. While 56% of the global variance in irrigation occurs between countries, 44% occurs within countries. Our results suggest that it is necessary to consider biophysical, socio-economic and governance information for identifying cropland areas that are likely to be under irrigation. Under current conditions, conversion from rainfed to irrigated cropland is most likely in eastern China, northern Africa, and parts of the Mediterranean region.

Keywords: Irrigation, global scale, countries, biophysics, governance, multilevel analysis

1 Introduction

Globally, the area size with irrigation facilities in place has almost doubled, over the past half century. At present, 301 million hectares of land are equipped for irrigation (Siebert *et al.*, 2010). Twenty-four per cent of the total harvested cropland is under irrigation (Portmann *et al.*, 2010) and more than 40% of the global cereal production takes place on irrigated land (Siebert and Döll, 2010). The immense importance of irrigated cropland to the global food supply has been stressed repeatedly (Kendall and Pimentel, 1994; Matson *et al.*, 1997; Rosegrant *et al.*, 2009; Sauer *et al.*, 2010). The global irrigation pattern, however, is very diverse, with more than two thirds of the world's irrigated land being located in Asia, with hardly any irrigation in Sub-Saharan Africa (Molden, 2007). Latitudinal distribution of irrigated area peaks between 20 and 45 degrees north (Wisser *et al.*, 2008). The highest level of irrigation required for achieving optimal crop growth has been recorded for India, Pakistan (Indus basin), Uzbekistan, Iraq, Turkey, and Egypt, with levels of over 500 mm/yr. When inefficient use of irrigation water is also taken into account the required amounts of water even triple in some of these regions (Doell and Siebert, 2002). Irrigation requirements are likely to increase in some cropping regions where precipitation levels are expected to decline in the future. If current crops such as dates, rice, cotton, citrus and sugar cane were not irrigated, their production would decrease by up to 60% (Siebert and Döll, 2010). Moreover, many countries with expected serious water scarcity in the future (e.g., in the semi-arid regions of Asia, the Middle East, Sub-Saharan Africa), will see their food production from irrigated land threatened (Seckler *et al.*, 1999; Zeitoun *et al.*, 2010). Given their importance to global food supply, irrigated areas are expected to expand in the future. Economic growth and urbanisation in many developing countries may lead to dietary shifts from staple food towards more green vegetables and fruit; for example, in China, which requires more irrigation (Molden, 2007). The area equipped for irrigation is expected to expand in developing countries by 40 million ha, up to 2030 (FAO, 2003). This expansion represents an annual growth rate of 0.6%, which is lower than the 1.9% achieved over the past four decades. The lower growth rate reflects a decrease in suitable areas and water resource availability in some countries, as well as increased investment costs for irrigation. The largest expansion in irrigated areas is expected in the land-scarce regions in South Asia, East Asia, the Near East and northern Africa (FAO, 2003).

Although a few attempts have been made to map irrigated areas at a continental or global scale (Siebert *et al.*, 2005; Siebert *et al.*, 2006; Portmann *et al.*, 2008; Thenkabail *et al.*, 2009; Wriedt *et al.*, 2009; You *et al.*, 2009), projected expansions in irrigated areas have not been mapped. Moreover, spatially explicit, global information about irrigation potential is lacking, as well as an understanding of its drivers and constraints. This may be traced back to limited data availability, at the global scale. An earlier study has indicated large regional differences in irrigation across Europe, namely as a consequence of spatial and temporal variability in climate and water availability (Hartmann, 1979). However, political changes also play a role in the development or maintenance of irrigation facilities, as O'Hara and Hannan (1999) and Fraser and Stringer (2009) have shown for Turkmenistan and Romania. Such studies, however, are restricted to regional and national scales. One attempt to identify suitable areas for irrigation was made by Heistermann *et al.* (2006). The authors simulated an expansion in irrigated areas in

Africa, based on the results from a multi-criteria analysis. Barreteau and Bousquet (2000) developed a regional multi-agent model to assess the viability of irrigated agriculture in the Senegal River Valley, based on access to credit for crop management, water distribution among famers, cropping periods, and farm management.

All these studies illustrate the importance of irrigation for agriculture and highlight regional differences. A globally consistent analysis of irrigation and its determining factors, however, is lacking. The objective of our study is to explain global patterns of irrigated cropland. Furthermore, we aim to identify cropping regions where a conversion from rainfed to irrigated cropland is likely. The latter play a central role in exploring the potential for meeting the future food demand of a growing global population. We took location factors into account at two different spatial levels: countries and grid cells (5 arc minutes, which are approximately 9.2 x 9.2 km on the equator). To identify the variability of the constraints, we conducted a multilevel analysis. Multilevel analysis is a statistical method, in which regression models are applied that explicitly account for data at different (spatial) levels. We considered biophysical aspects mainly at grid-cell level and socio-economic and governance aspects primarily at country level. Based on the results, we identified areas where irrigated cropland is likely to be expanded.

2 Data and methods

2.1 Data

Grid-cell data

We obtained global, spatially explicit information on harvested areas that had been irrigated and those that were rainfed, from the MIRCA2000 data set (Portmann *et al.*, 2010). The MIRCA2000 data set indicates, for each 5 arc-minute grid cell, the areas of irrigated and rainfed cropland, referring to the year 2000. MIRCA2000 was developed by combining spatially explicit information on cropland (Ramankutty *et al.*, 2008), harvested cropland (Monfreda *et al.*, 2008), and areas equipped for irrigation (Siebert *et al.*, 2006), with census-based statistics on irrigated and rainfed crop areas and cropping calendars, collected from many different sources. The main improvement of MIRCA2000 compared to earlier products (Siebert *et al.*, 2005; Siebert *et al.*, 2006) is that it presents areas actually under irrigation, instead of areas equipped for irrigation – globally, 84% of the equipped area is actually irrigated (Siebert *et al.*, 2010). Another improvement is the explicit consideration of multi-cropping practices. We used the MIRCA2000 map to calculate the fraction of irrigated cropland as the ratio of harvested irrigated cropland and harvested total cropland (Figure 1, upper part). For our analysis, we excluded marginal cropping areas that covered less than 0.1% of a grid-cell area. To reduce errors of commission, all areas with at least 10% irrigated cropland were classified as irrigated (irrigation = 1) and all areas with less than 10% irrigated cropland were considered as rainfed (irrigation = 0). Figure 2 shows the distribution of grid cells containing irrigated cropland and how they are classified in this study. By applying this 10% threshold, 24% of all cropland cells included in our analysis represents cropland that is irrigated, which corresponds with the total global irrigated cropland area (Portmann *et al.*, 2010). The resulting map of irrigated areas is shown in the lower map of Figure 1. From this data set we drew a random, balanced sample of 5%, in which irrigated and rainfed grid cells are represented in equal numbers. This random sampling reduced the potential bias in the estimation procedure, as a result of spatial autocorrelation.

Insert Figure 1 approximately here

Insert Figure 2 approximately here

A set of gridded data available at a 5-arc-minute spatial resolution was obtained to explain the spatial variation in the occurrence of irrigation in cropland. We considered population density, market accessibility, slope, humidity, river discharge, evaporation, transpiration, potential evapotranspiration (PET), and run-off. Subsequently, Pearson correlation coefficients were calculated to assess relations between all grid-cell variables. Where variables were highly correlated ($\geq$0.7), selected ones were excluded from further statistical analysis, to avoid multicollinearity. This section and Table 1 describe all grid-cell data that were included in the analysis. We applied climate data from the Climate Research Unit (CRU) to calculate river discharge using the dynamic vegetation model LPJmL (Mitchell and Jones, 2005; Bondeau *et al.*, 2007). River discharge, defined as channelled access water that could not infiltrate the soil or percolate to near-surface groundwater via a pre-defined river network, was simulated according to Rost *et al.* (2008), taking the given land-use pattern into account. In the simulation, daily evaporation of water from bare soil occurred from the fraction of the grid cell not covered by vegetation. The size of this fraction depends on the plant-covered area and the daily status of the PFT(plant functional type)-specific leaf phenology. Evaporation is assumed to decline linearly as soil dries, a simple approach that is suitable for large scales (Huang *et al.*, 1996; Gerten *et al.*, 2004). Daily equilibrium evapotranspiration rate depends primarily on net radiation and temperature. Net radiation (net short-wave flux minus net long-wave flux) was calculated from latitude, day of the year, sunshine hours and air temperature (see Prentice et al., 1993; Gerten et al., 2004). PET was determined by multiplying daily equilibrium evapotranspiration with the Priestley–Taylor coefficient, a=1.32 (Priestley and Taylor, 1972). Humidity was calculated by normalising precipitation with PET, making irrigation potentially comparable across the globe. For the hydrological and climate variables, the average of their monthly mean values for the 1990-2005 period were used. This 15-year period was chosen to account for the variability in the reference years of input data used for generating the global irrigation maps. This period furthermore allowed for obtaining representative hydrological information. We also included information on population density and market accessibility at the 5-arc-minute grid-cell level (Table 1).

Insert Table 1 approximately here

Country data

Given the lack of sufficient data at grid-cell level and the role of national governance in irrigation decisions, we considered additional socio-economic and political parameters at country level. For all countries, we obtained governance information from the Worldbank (2009) and the Polity IV Project (Marshall and Jaggers, 2009) for the year 2000, or, if that was not available, for a year close to 2000. Indicators considered were control of corruption, government effectiveness, political stability, GDP per capita, as well as level of democracy and autocracy (Table 1). These indicators, however, showed strong correlations, potentially causing multicollinearity in the multilevel analysis. To reduce redundancies in the data and to obtain independent variables, we conducted a factor analysis with principal component extraction and varimax rotation. Factor analysis is a frequently used tool in land-use studies to analyse correlations among a large number of land use-related variables, and to reduce the information contained in these variables to a smaller set of factors with a minimal loss of information (e.g., Veldkamp and Fresco, 1997; Lise, 2000; Dunjó *et al*., 2003; Herzog *et al*., 2006). The factor analysis resulted in two factors which explain most of the variance. Both factors together explain 87.7% of the total variance, with one factor contributing 69.1% and the other 18.7%. The factor communalities (Table 2) shows that between 78 and 95% of the variance occurring in each variable is accounted for. The rotated component matrix in Table 2 presents the factor loadings for each of the variables. The variables 'control of corruption', 'government effectiveness, 'GDP', and 'political stability' were highly correlated with the first factor. 'Autocracy' and 'democracy' were highly correlated with the second factor. Therefore, the first factor could be interpreted as a factor mainly comprising aspects of government performance and was labelled 'government_performance'. The second factor essentially expressed the level of autocracy and democracy and, therefore, was called ‘government_type’. These two new country-specific factors were then used in the multilevel analysis.

We also included the total natural renewable water resources as an independent variable at country level (AQUASTAT, 2010). An overview of all independent variables and their mean values per country is provided in Table 1 of the online supplementary material.

Insert Table 2 approximately here

2.2 Methodology

The aim of our study is to explain the presence of global irrigated cropland using biophysical and socio-economic information. While global biophysical information is widely available at grid-cell level, the availability of socio-economic information is largely limited to the national level. Hence, multi-scale approaches need to be applied to account for these structural data differences, as emphasised by several authors (Kok and Veldkamp, 2001; Veldkamp *et al.*, 2001; Evans and Kelley, 2004; Verburg *et al.*, 2008; Neumann *et al.*, 2011). In this study, we applied different logistic models as described in the remainder of the methodology section. First, we conducted a multilevel analysis with all grid-cell level variables to explore how available grid-cell data (which is largely restricted to biophysical information) explain irrigation patterns. Second, we conducted a multilevel analysis using all grid-cell and country-level variables to also account for national socio-economic and political factors. Finally, results from the multilevel analysis were compared with results from a logistic regression to assess the added value of the multilevel analysis.

Multilevel analysis

Hierarchically structured data, such as the presented grid-cell and country-level data, are preferably analysed by accounting for statistical within-group and between-group relationships within a single analysis. Multilevel analysis (Snijders and Bosker, 1999) is a method particularly suited for this. A 'group' is an aggregated unit (e.g., country) and is referred to as a level-two (or higher) unit. Units at the lowest hierarchical level are called 'individuals' (e.g., grid cells) and are being referred to as level-one units. In multilevel models, each of these levels are represented by their own sub-models, which represent relationships between variables of the respective levels but also specify how variables at this level influence relationships occurring at other levels (Raudenbush and Bryk, 2002). Although multilevel models mostly have been used in disciplines such as educational research, they have recently become more popular in analyses of environmental issues, such as land use (Polsky and Easterling, 2001; Pan and Bilsborrow, 2005; Overmars and Verburg, 2006; Vance and Iovanna, 2006; Gray *et al.*, 2008; Reidsma *et al.*, 2009).

Multilevel analyses are explained in detail by Snijders and Bosker (1999) and Raudenbush and Bryk (2002). In this study we have applied an important tool for multilevel analysis: the hierarchical linear model, which is an advanced type of regression model (Snijders and Bosker, 1999). As with all regression models, its objective is to explain one dependent variable Y by a set of independent variables. As the hierarchical linear model explains a process or situation at the most detailed level (e.g. grid cells), the dependent variable must be a variable at that level. Therefore, the essence of multilevel modelling is that the dependent variable Y has both a group and an individual element. The same applies to the other level-one variable X which may also contain a group element (Snijders and Bosker, 1999).

Specifying the multilevel models

In this study we distinguish between two levels, the grid-cell level (level one) and the country level (level two). Variables at the grid-cell level, such as slope and humidity, may explain part of the grid-cell variability in irrigated cropland. Grid-cell variables, in addition, may explain part of the country level variability in irrigated cropland, if the respective variables (e.g., humidity) in some countries are significantly higher or lower than their mean. Variables at country level, that is, government type and performance, may explain differences in irrigation between countries. Equation 1 shows a simple two-level model containing one independent variable at each level.

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right)=\beta_{0j}+\beta_{1j}x_{ij}+\beta_2 z_j$$

Equation 1

The probability of presence of irrigation is p_{ij}, The independent variable x refers to the grid-cell level and the independent variable z refers to the country level. Grid-cell-level effects are marked with the index $_i$ while all country-level effects have the index $_j$. The regression coefficients to be estimated are β_n. Equation 1 implies that the grid-cell-dependent residuals are assumed to be mutually independent with a mean of zero. To account for between-country variability, country-dependent intercepts are included in the model as indicated by β_{0j}. Hence, the probability of the presence of irrigation, in this case, not only depends on the grid-cell variabilities, but also on the countries (Snijders and Bosker, 1999).

The country-dependent intercept β_{0j} can be split into an average intercept γ_{00} and the country-dependent deviation U_{0j}.

$$\beta_{0j}=\gamma_{00}+U_{0j}$$

Equation 2

U_{0j} are the unexplained group effects, so-called group residuals, which control for the effects of variable x. U_{0j} are mutually independent and with zero mean. In our study, U_{0j}

represents the country-specific deviation of the probability of finding irrigation (assuming all other conditions equal). If we exchange the relevant part of Equation 1 with Equation 2, and exclude all explanatory variables *x* and *z*, we obtain the simplest case of the hierarchical model, the so-called unconditional or empty model (Equation 3).

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right)=\gamma_{00}+U_{0j}$$

Equation 3

The unconditional model (multilevel model 0) only contains random effects. As it analyses which part of the total variance results from between and within level components, it is a good starting point for multilevel analysis. With this model, we can estimate whether the variance at country level is significant. The intra-class correlation coefficient (ρ_I) is used for separating the variance on country level from the total variance (Snijders and Bosker, 1999). In our study ρ_I represents the proportional variance between countries.

Multilevel model 1 includes all independent biophysical grid-cell variables: slope, discharge, humidity, evaporation and evapotranspiration (Table 1), to explain the variance at grid-cell level (Equation 4). Because linear hierarchical models require normality, the *slope* and *discharge* were log transformed. All other variables show normally distributed values.

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right)=\gamma_{00}+\beta_1\ln(slope_{ij})+\beta_2\ln(discharge_{ij})+\beta_3(humidity_{ij})+\beta_4(evap_{ij})+\beta_5(ET_{ij})+U_{0j}$$

Equation 4

In addition to these variables, multilevel model 2 (Equation 5) also includes the socio-economic grid-cell variables (*access* and *population*), as well as all country-level variables (*water*, *government_performance* and *government_type*; Table 1). A log transformation was necessary for *access*, *population* and *water*, to achieve normally distributed variables.

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right)=\gamma_{00}+\beta_1\ln(slope_{ij})+\beta_2\ln(discharge_{ij})+\beta_3(humidity_{ij})+\beta_4(evap_{ij})+\beta_5(ET_{ij})+$$
$$\beta_6\ln(access_{ij})+\beta_7\ln(population_{ij})+\beta_8\ln(water_j)+\beta_9(government\_performance_j)+$$
$$\beta_{10}(government\_type_j)+U_{0j}$$

Equation 5

All independent variables were centred prior to the multilevel analysis, to control for correlations amongst random components and to stabilise the model (Kreft and De Leeuw, 2002; Paccagnella, 2006; Bickel, 2007). We centred the grid-cell-level variables around their respective country means. Hence, the scale of the variables was shifted by subtracting the country mean from the data points. The country-level variables were grand-mean centred, that is, they were centred around the respective global mean (Raudenbush and Bryk, 2002). We used the hierarchical linear and nonlinear modelling

(HLM6) software for estimating the different models (http://www.ssicentral.com). All parameters were estimated using the restricted maximum likelihood (RML) method and the penalised quasi-likelihood (PQL) approach (Raudenbush and Bryk, 2002). The calculated regression coefficients can be interpreted in the same way as regression coefficients obtained from ordinary logistic regression.

To evaluate the goodness-of-fit, we conducted a relative operating characteristic (ROC) analysis. The ROC measure indicates how well the estimated model explains the dependent variable. A ROC value of 0.5 indicates a pure random model, whereas a ROC value equal to 1 indicates a perfect fit (Swets, 1988).

To assess whether multilevel models perform better than ordinary least squares analyses, we applied a logistic regression to all biophysical and socio-economic variables at grid-cell level. All grid cells of the global 5% sample were included in the regression analysis (Equation 6).

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 \ln(slope_i) + \beta_2 \ln(discharge_i) + \beta_3 (humidity_i) + \beta_4 (evap_i) + \beta_5 (ET_i) + \beta_6 \ln(access_i) + \beta_7 \ln(population_i)$$

Equation 6

The probability of the presence of irrigated cropland is p_i. The regression coefficients to be estimated are β_n, the intercept is β_0. Similar to the multilevel analysis, a ROC analysis was conducted to assess the goodness-of-fit.

Based on the estimated regression models, the probabilities predicted for the presence of irrigated cropland also is calculated to indicate areas where conversion from rainfed to irrigated cropland is likely if all considered constraints remain constant. For this purpose, we mapped the estimated probabilities and compared these to the location of currently irrigated areas.

3 Results

3.1 Currently irrigated areas

Results from the multilevel analysis show that the unconditional model (multilevel model 0) has a highly significant variance component ($P< 0.001$). This means there is a significant variability between countries in terms of irrigation. The intra-class correlation coefficient ρ_I is 0.56, indicating that 56% of the variance occurs between countries while 44% of the variance occurs within countries. Table 3 presents the results from the grid-cell-level model, multilevel model 1, and multilevel model 2. Multilevel model 1 results show that humidity and river discharge significantly explain the variability at grid-cell level ($P<0.001$ for both). Hence, the probability of cropland being irrigated is higher under arid conditions and if river water resources are available. Slope, evaporation and evapotranspiration have no significant influence on the probability of irrigated cropland ($P>0.05$ for all). The intra-class correlation coefficient ρ_I is comparable to the unconditional model, indicating that these variables cannot explain the variance at country level, as identified by the unconditional model. The results from multilevel model 2 show that irrigation is significantly correlated with population density ($P<0.01$) and market accessibility ($P<0.001$). The variables *government_performance* and *government_type* also make a significant contribution to the model at national level ($P<0.05$ and $P<0.01$, respectively). The likelihood of cropland being irrigated increases in economically strong and politically stable countries (i.e., countries with a higher government performance). Availability of renewable water resources per country cannot explain variability in irrigated cropland at country level. Results from the logistic regression analysis indicate that all grid-cell-level variables contribute to the explanation of the variation in irrigation in cropland areas. The Wald test (Table 3) indicates that population density explains most of the global irrigation pattern. Population density alone explains 63% of the variation in irrigation, while all independent variables together explain only slightly more (66% of the total variation). High population densities and short travel times to markets often result in intensive agriculture, including irrigation practices. The ROC value increases from multilevel model 0 (ROC = 0.786) and multilevel model 1 (ROC=0.801) to multilevel model 2 (ROC = 0.812). This indicates that, in spite of the significant contribution of the variables, they only have a limited effect on the overall explained variance. The model fit of the logistic regression analysis is 0.724, which is below those of the three multilevel models.

Insert Table 3 approximately here

Figures 3 and 4 show the likelihood of current croplands being irrigated, based on multilevel model 1 and multilevel model 2 calculations. The likelihood of cropland being irrigated is high in most arid areas, such as northern Africa, Arabian Peninsula, India, Central Asia, Sub-Saharan Africa, South Africa and the western parts of North and South America. Croplands in many humid countries, for example, those in northern Europe and Southeast Asia, have a low probability for being irrigated. Accounting for socio-economic and governance constraints increases the likelihood of croplands being irrigated in countries with a generally better *government_performance* (e.g. France, Germany, the United States, Japan, parts of China, Australia) (Figure 4). The opposite trend is observed for some countries which are characterised by lower political and economic stability, with a decrease in the likelihood of irrigation (e.g. in Bolivia, Central America, Nigeria, Zambia, Angola, the Ukraine). Hence, unfavourable socio-economic and political conditions are a barrier to irrigated agriculture in certain countries. Differences between multilevel model 1 and multilevel model 2 results are not only remarkable between countries but also within them. Figure 4 shows that probabilities of irrigation in the multilevel model 2 calculation are higher than in multilevel model 1, for many metropolitan areas with high population densities and better market accessibility, representing the capital investment needed to establish irrigation (e.g. in California, eastern United States, around Lake Victoria, Java, eastern China). The likelihood of irrigation in the United States and eastern China, according to the multilevel 1 calculation, is dominated by the occurrence of river discharges. In multilevel model 2 calculations, the contribution of socio-economic and political factors significantly changes the regional pattern of irrigation likelihood.

Insert Figure 3 approximately here

Insert Figure 4 approximately here

3.2 Potential irrigation expansion areas

The predicted probabilities for the presence of irrigated cropland may also indicate areas where irrigation is likely to be expanded. Therefore, a map of these areas currently not irrigated was made by applying multilevel model 2 (Figure 5). This map shows cropping regions where irrigation is likely to be applied if demand for irrigation increases given the considered biophysical, socio-economic and governance conditions. Under these circumstances, expansion is most likely near already irrigated areas, for example, in eastern China, northern Africa, and parts of the Mediterranean region. Expansion, furthermore, is most probable close to urban areas, for example, in the eastern United States.

Insert Figure 5 approximately here

4 Discussion

This study has shown how different logistic models can be applied in the spatial analysis of land-use management, such as irrigation, accounting for processes operating on different spatial scales. This section first discusses the methodological issues related to this analysis, followed by a discussion on the findings.

4.1 Cropland currently irrigated

Figures 3 and 4 show the remarkable difference in irrigation probabilities between multilevel model 1 and multilevel model 2 calculations. While the former model accounts for biophysical information only, the latter accounts for biophysical, socio-economic and governance information. The socio-economic variables that were included at national level, however, explain only some of the differences in irrigated cropland between countries. We consider there to be two main reasons for this. To some degree it may be explained by recent political changes. Irrigation systems were often established decades ago when economic and political conditions may have been different from those of today. Hence, the current global irrigation pattern is inherited from earlier circumstances, which makes it difficult to explain on the basis of current political and economic information. For example, heavy investments in irrigation systems in Central Asia, made by the former Soviet government as an important part of its agricultural development policy, resulted in an expansion of irrigated areas during the Soviet period (Saiko and Zonn, 2000). Irrigation is still largely applied in the Republic of Uzbekistan, formerly part of the Soviet Union, irrespective of drastically changed political and economic conditions. The different circumstances at the time of the construction of the irrigation systems and the independent variables prevent a complete explanation of the current situation. The same applies to the constraints on grid-cell level, such as market accessibility and population density. However, many of the variables on grid-cell level are assumed to be less dynamic over time.

A second explanation for the limited influence of variables on country level is the large size of a number of countries. Using country average socio-economic data for large countries, such as the United States, Brazil or China, may be less appropriate. For example, Akita (2003) has shown that there are significant differences in GDP per capita between different regions within China. To capture such regional disparities, intermediate levels would be necessary, such as the province level. However, globally consistent and complete socio-economic data at such levels is not available.

Results from the logistic regression analysis show that all independent grid-cell variables contribute to the explanation of the presence of irrigated cropland. This contrasts the results from the multilevel analyses, from which is concluded that slope, evaporation and evapotranspiration do not significantly contribute to the explanation of the irrigated cropland pattern. Moreover, individual contributions of the significant independent variables differ between the two applied methodologies. Multilevel analysis variables at grid-cell level can explain variability both within and between countries. Certain grid-cell variables may largely vary across the globe. However, they may vary less within certain countries where biophysical effects play a role, due to these countries' geographic location. This may result in non-significant variables in the multilevel analysis. Furthermore, it may also explain the small increases in ROC values over the three multilevel models.

4.2 Probabilities of irrigation expansion

The regions identified as very likely areas for irrigation expansion largely coincide with potential areas for irrigation as identified by the FAO (2003). An exception is Southeast Asia (Indonesia, Papua New Guinea, Malaysia, the Philippines), for which we have identified low probabilities of irrigation expansion, given the high humidity in this region. The estimates of the FAO considered the importance of irrigated paddy fields in this region, which was not explicitly taken into account in this study. The calculated probabilities not only indicate grid cells containing rainfed areas where irrigation may be introduced. They also indicate grid cells containing already irrigated areas where the percentage of irrigation may be increased. Only 12% of all grid cells classified as irrigated cropland were fully irrigated, while the majority of grid cells contained at least some rainfed areas. Hence, potential for irrigation expansion may also exist within already irrigated regions. However, it has to be noted that large parts of the variation could not be explained (Table 3). Therefore, high probabilities of irrigation in rainfed areas either indicate that this potential has not been used, or that other constraining factors play a role.

The presented probabilities of irrigation expansion for croplands reflect current biophysical, socio-economic and governance constraints that are included in the statistical analysis. The results do not necessarily imply that these factors alone will trigger an expansion of irrigated croplands. Public investments in market and irrigation infrastructure, technological progress and water rights can be important stimuli, as well. Biophysical, socio-economical and governance constraints and the expected temporal changes are to be accounted for when assessing potential future irrigation expansion. However, such analysis was beyond the scope of our study, given the limited availability of spatially explicit projections of socio-economic parameters.

Increasing demand for agricultural products on local, regional and global levels may trigger irrigation. However, the effect of a growing global food demand on local irrigation is complex and difficult to assess, as it is influenced by agricultural trade policies, market orientation and competitiveness between regions and countries. Chapagain et al. (2006) estimate that more than half of the water that is used for irrigation in Pakistan, China, Uzbekistan and India, is used in the production of exported cotton products. The amount of water used not only depends on local availability. Import and consumption of water-intensive agricultural commodities (e.g., coffee, tea, rice, cotton) largely influences the amount of water used in irrigation in the producing regions and may trigger further investments in irrigation infrastructure (Kumar and Singh, 2005; Hoekstra and Chapagain, 2007). Information on all these issues is required for a thorough assessment of the potential for expansion of irrigated croplands. Given the lack of detailed and consistent global data sets, their exploration was beyond the scope of this paper (see also Section 4.3).

To make an assessment of the potential for future irrigated areas, climate change also is an important parameter. It influences precipitation, evapotranspiration, snowpacks, groundwater recharge, and other factors of water resources and water demand. This, in turn, will affect both the demand for irrigation and the availability of water (Döll, 2002; Izaurralde *et al.*, 2003; Rosenzweig *et al.*, 2004). Hence, to further investigate potential new areas for irrigation, scenarios on future climate change are recommended to be used for the projection of changes in irrigation water requirements. Rosegrant *et al.* (2009) argue that, up to 2050, irrigation will remain the world's largest user of fresh water, although its share is projected to decline compared to that used domestically and in industry. Our results indicate possible locations for irrigation expansion under current constraints. However, when modelling future expansions of irrigated areas, possible future water scarcities should be taken into account. The results from our study also indicate that political developments may influence investments in irrigation systems. For example, the recent land reforms in Zimbabwe have led to agricultural failure and a loss of nearly all irrigation systems throughout the country (Moss and Patrick, 2006). Furthermore, the collapse of communism in former Eastern Europe has caused a drastic decline in the number of irrigated areas (Fraser and Stringer, 2009; Theesfeld, 2004). In contrast, in the Republic of Uzbekistan, this collapse did not lead to a similar decline, as the production on irrigated cotton and rice fields has a strong tradition, and continues to be supported by the Uzbek government.

Besides the benefits of irrigation for food production it has also led to adverse environmental impacts in many places across the globe. Globally, water used for irrigation has caused drastic modification of the water flows of many rivers. In some cases, rivers no longer reach the sea during dry periods, and water tables are falling due to unsustainable extraction of groundwater (Foley *et al.*, 2005). In addition, salinisation of irrigated land is causing an annual loss in arable land of approximately 1.5 million hectares (Wood *et al.*, 2000). To evaluate the potential for expansion of irrigation areas and for the implementation of irrigation programmes, it is essential to assess potential environmental impacts.

4.3 Data issues

Water availability and climate are important drivers of irrigation. Therefore, in our study, we considered river discharge, evaporation, evapotranspiration and humidity at grid-cell level. These variables, however, may vary within and between years. For example, seasonal variations in precipitation levels within a particular year influence farmers' decisions to irrigate, a factor difficult to capture in averaged monthly data. Furthermore, irrigation facilities are installed in many cropping regions in anticipation of periods of uncertain precipitation and to minimise the risk of crop losses in exceptionally dry periods. Often these irrigation facilities are used infrequently, and their presence cannot be explained by long-term average climate data as used in this study. This potentially causes statistical noise.

Another hydrological aspect to be considered when explaining the global distribution of irrigated areas is access to groundwater, as this is the main source of irrigation water for certain countries. Groundwater has been extracted for irrigation purposes for many years, in many countries, including the Unites States, India, Pakistan, China, Mexico and Yemen (Shah et al., 2000; Shah et al., 2006). Access to groundwater, therefore, is not only important for explaining the current pattern of irrigated areas, but also for identifying potential areas for irrigation. Global inventories of groundwater use for irrigation have only recently become available, but the spatial detail is not sufficient for conducting spatially explicit analyses at grid cell level. Furthermore, information about access to groundwater resources and aquifer properties cannot be obtained (Siebert et al., 2010). Therefore, aquifers properties were not explicitly included as an explanatory factor for irrigated cropland.

The most spatially detailed units of analysis are 5 arc-minute grid cells. This resolution was chosen because irrigation data, as well as a number of socio-economic variables, were available at this resolution. The hydrological data, however, were only available at an 0.5 degree resolution. We directly applied the 0.5 degree data to the 5 arc-minute resolution, which may have introduced some bias. Including the 0.5 degree resolution as an extra level of analysis would not have been an adequate solution, as it is rather a spatial scale for representing the data and not a distinct spatial level. In addition, resampling the finer scale data to 0.5 degrees would have resulted in the loss of too many spatial details. Next to data resolution, there may be significant uncertainty in the data sets we used. The irrigation data are based on a large variation of sources from different years, which are likely to have introduced uncertainties. Second, independent variables were collected from different sources and since these were not validated, their accuracies may differ. Thus, the data used in our study are inherently uncertain. They do, however, comprise a selection of the most appropriate data available.

Land-management decisions such as irrigation are often made at household level. At this level, information on income sources, ethnicity, land ownership, water-sharing rules and access to funding, may be taken into accounted. Several authors have considered the household level in multilevel analyses in land-use studies (Overmars and Verburg, 2006; Pan and Bilsborrow, 2005; Vance and Iovanna, 2006). Considering the global context of this study, such aspects could not be included, as consistent, global data were unavailable.

The data and methodology used in this study indicate probabilities of existing cropland being irrigated, but do not allow any estimates on the total extent of irrigated land. Under more humid and temperate climate conditions the probability of cropland being irrigated is low, because these lands are predominantly rainfed, but the total extent of irrigated land may be large because of the high cropland density. In contrast, almost all cropland in very arid regions is irrigated, resulting in high probabilities in our study (e.g., in Egypt or Central Asia), while the total extent of irrigated land may be low because of the low total cropland density. This difference needs to be taken into account when comparing estimates of the extension of irrigated land in absolute terms (e.g. FAO, 2003) to results of this study indicating probabilities that cropland is irrigated.

5 Conclusions

The results from this study have shown that local biophysical factors alone may explain global irrigation patterns only to a limited degree. Taking socio-economic and governance information into account does improve the explanation. However, much socio-economic data and governance information is only available at country level. Given this lack of data at finer spatial scales and the role of national governance, multi-level analysis was used as a methodology to explain the global irrigation pattern explicitly accounting for the hierarchical structure of the data. Compared to the ordinary logistic regression, multilevel analysis allows for a better prediction of where irrigation is likely to expand.

Information on areas with high probabilities of irrigation expansion is crucial for any analysis of the spatio-temporal dynamics of future food production. The current study is a first step to address irrigated land more explicitly in global analysis and models of land-use change and food production, based on empirical analysis. A possible land-use modelling exercise would combine the spatially explicit information on cropland suitability for irrigation with quantitative projections on cropland expansion per country as, for example, those from the FAO (FAO, 2003). Hence, the identification of factors explaining global irrigation patterns and that of areas with high probability for expansion of irrigation are steps forward in exploring spatial-temporal dynamics of irrigated areas. Yet, the results are not only valuable for visualising possible dynamics of irrigated areas on a global scale. They may further support the description and definition of international objectives for irrigation expansion. Respective policies, however, are often instrumented and implemented on a national or regional scale, considering the specific bio-physical, economic and/or political conditions. Hence, the suitability of such global data sets for underpinning national or regional policies on irrigation expansion is recommended to be carefully explored before data application.

Acknowledgement

This research is a contribution to the Global Land Project (GLP). The authors acknowledge the BSIK RvK IC2 project ‘Integrated analysis of emission reduction over regions, sectors, sources and greenhouse gases’, for funding the research that led to the present publication. We especially thank Hester Biemans for explaining several hydrological modelling aspects. The spatial data sets of irrigation likelihoods as presented in this study are freely available from PBL Netherlands Environmental Assessment Agency (please contact Kathleen.neumann@pbl.nl).

References

Akita, T., 2003. Decomposing regional income inequality in China and Indonesia using two-stage nested Theil decomposition method. The Annals of Regional Science 37, 55-77.

FAO, 2010. AQUASTAT, http://www.fao.org/nr/water/aquastat/main/index.stm, last access: xx/xx/2010..

Barreteau, O., Bousquet, F., 2000. SHADOC: a multi-agent model to tackle viability of irrigated systems. Annals of Operations Research 94, 139-162.

Bickel, R., 2007. Multilevel analysis for applied research: it's just regression! The Guilford Press, New York.

Bondeau, A., Smith, P.C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., Mueller, C., Reichstein, M., Smith, B., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Global Change Biology 13, 679–706.

Chapagain, A.K., Hoekstra, A.Y., Savenije, H.H.G., Gautam, R., 2006. The water footprint of cotton consumption: An assessment of the impact of worldwide consumption of cotton products on the water resources in the cotton producing countries. Ecological Economics 60, 186-203.

Doell, P., Siebert, S., 2002. Global modeling of irrigation water requirements. Water Resour. Res. 38.

Dunjó, G., Pardini, G., Gispert, M., 2003. Land use change effects on abandoned terraced soils in a Mediterranean catchment, NE Spain. CATENA 52, 23-37.

Evans, T.P., Kelley, H., 2004. Multi-scale analysis of a household level agent-based model of landcover change. Journal of Environmental Management 72, 57-72.

FAO, 2003. World Agriculture: Towards 2015/2030. An FAO Perspective. In: Bruinsma, J. (Ed.). Food and Agriculture Organization (FAO), Rome.

Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N., Snyder, P.K., 2005. Global Consequences of Land Use. Science 309, 570-574.

Fraser, E.D.G., Stringer, L.C., 2009. Explaining agricultural collapse: Macro-forces, micro-crises and the emergence of land use vulnerability in southern Romania. Global Environmental Change 19, 45-53.

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., Sitch, S., 2004. Terrestrial vegetation and water balance--hydrological evaluation of a dynamic global vegetation model. Journal of Hydrology 286, 249-270.

Gray, C.L., Bilsborrow, R.E., Bremner, J.L., Lu, F., 2008. Indigenous Land Use in the Ecuadorian Amazon: A Cross-cultural and Multilevel Analysis. Human Ecology 36, 97–109.

Hartmann, H.D., 1979. The Situation of Irrigation and Water Requirements in Western Europe. Acta Horticulturae (ISHS) 89, 19-28.

Heistermann, M., 2006. Modelling the Global Dynamics of Rain-fed and Irrigated Croplands. Center for Environmental Systems Research. Kassel University, Kassel, p. 161.

Herzog, F., Steiner, B., Bailey, D., Baudry, J., Billeter, R., Bukacek, R., Blust, G.d., Cock, R.d., Dirksen, J., Dormann, C.F., Filippi, R.d., Frossard, E., Liira, J., Schmidt, T., Stoeckli, R., Thenail, C., Wingerden, W.v., Bugter, R., 2006. Assessing the intensity of temperate European agriculture at the landscape scale. European Journal of Agronomy 24, 165-181.

Huang, J., van den Dool, H.M., Georgarakos, K.P., 1996. Analysis of Model-Calculated Soil Moisture over the United States (1931–1993) and Applications to Long-Range Temperature Forecasts. Journal of Climate 9, 1350-1362.

Kendall, H.W., Pimentel, D., 1994. Constraints on the Expansion of the Global Food Supply. Ambio 23, 198-205.

Kok, K., Veldkamp, A., 2001. Evaluating impact of spatial scales on land use pattern analysis in Central America. Agriculture, Ecosystems & Environment 85, 205-221.

Kreft, I., de Leeuw, J., 2002. Introducing Multilevel Modeling. Sage Publications, London.

Lise, W., 2000. Factors influencing people's participation in forest management in India. Ecological Economics 34, 379-392.

Marshall, M.G., Jaggers, K., 2009. Polity IV Project. Political Regime Characteristics and Transitions, 1800-2007. Center for Systemic Peace.

Matson, P.A., Parton, W.J., Power, A.G., Swift, M.J., 1997. Agricultural Intensification and Ecosystem Properties. Science 277, 504-509.

Mitchell, T.D., Jones, P.D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology 25, 693-712.

Molden, D., 2007. Water for Food, Water for Life. A Comprehensive Assessment of Water Management in Agriculture. Earthscan, London and International Water Management Colombo.

Monfreda, C., Ramankutty, N., Foley, J.A., 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochem. Cycles 22.

Moss, T., Patrick, S., 2006. After Mugabe: Applying Post-Conflict Recovery Lessons to Zimbabwe. Africa Policy Journal 1.

Neumann, K., Verburg, P.H., Elbersen, B., Stehfest, E., Woltjer, G.B., 2011. Multi-scale scenarios of spatial-temporal dynamics in the European livestock sector. Agriculture, Ecosystems & Environment 140, 88-101.

O'Hara, S.L., Hannan, T., 1999. Irrigation and Water Management in Turkmenistan: Past Systems, Present Problems and Future Scenarios. Europe-Asia Studies 51, 21-41.

Overmars, K.P., Verburg, P.H., 2006. Multilevel modelling of land use from field to village level in the Philippines. Agricultural Systems 89, 435-456.

Paccagnella, O., 2006. Centering or Not Centering in Multilevel Models? The Role of the Group Mean and the Assessment of Group Effects. Eval Rev 30, 66-85.

Pan, W.K.Y., Bilsborrow, R.E., 2005. The use of a multilevel statistical model to analyze factors influencing land use: a study of the Ecuadorian Amazon. Global and Planetary Change 47, 232-252.

Polsky, C., Easterling, W.E., 2001. Adaptation to climate variability and change in the US Great Plains:: A multi-scale analysis of Ricardian climate sensitivities. Agriculture, Ecosystems & Environment 85, 133–144.

Portmann, F., Siebert, S., Bauer, C., Döll, P., 2008. Global data set of monthly growing areas of 26 irrigated crops. Frankfurt Hydrology Paper 06. Institute of Physical Geography, University of Frankfurt, Frankfurt am Main, Germany.

Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Global Biogeochem. Cycles 24, GB1011.

Prentice, I.C., Sykes, M.T., Cramer, W., 1993. A simulation model for the transient effects of climate change on forest landscapes. Ecological Modelling 65, 51-70.

Priestley, C.H.B., Taylor, R.J., 1972. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. Monthly Weather Review 100, 81-92.

Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Global Biogeochem. Cycles 22.

Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical Linear Models: Applications and Data Analysis Methods. Sage Publications, Thousand Oaks.

Reidsma, P., Ewert, F., Oude Lansink, A., Leemans, R., 2009. Vulnerability and adaptation of European farmers: a multi-level analysis of yield and income responses to climate variability. Regional Environmental Change 9, 25-40.

Rosegrant, M.W., Ringler, C., Zhu, T., 2009. Water for Agriculture: Maintaining Food Security under Growing Scarcity. Annual Review of Environment and Resources 34, 205-222.

Rost, S., Gerten, D., Bondeau, A., Lucht, W., Rohwer, J., Schaphoff, S., 2008. Agricultural green and blue water consumption and its influence on the global water system. Water Resour. Res. 44.

Saiko, T.A., Zonn, I.S., 2000. Irrigation expansion and dynamics of desertification in the Circum-Aral region of Central Asia. Applied Geography 20, 349-367.

Sauer, T., Havlík, P., Schneider, U.A., Schmid, E., Kindermann, G., Obersteiner, M., 2010. Agriculture and resource availability in a changing world – the role of irrigation. Water Resources Research.

Seckler, D., Barker, R., Amarasinghe, U., 1999. Water Scarcity in the Twenty-first Century. International Journal of Water Resources Development 15, 29-42.

Shah, T., Molden, D., Sakthivadivel, R., Seckler, D., 2000. The global groundwater situation: Overview of opportunities and challenges. International Water Management Institute, Colombo, Sri Lanka.

Shah, T., Singh, O., Mukherji, A., 2006. Some aspects of South Asia's groundwater irrigation economy: analyses from a survey in India, Pakistan, Nepal Terai and Bangladesh. Hydrogeology Journal 14, 286-309.

Siebert, S., Burke, J., Faures, J.M., Frenken, K., Hoogeveen, J., Döll, P., Portmann, F.T., 2010. Groundwater use for irrigation – a global inventory. Hydrol. Earth Syst. Sci. 14, 1863-1880.

Siebert, S., Döll, P., 2010. Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. Journal of Hydrology 384, 198-217.

Siebert, S., Döll, P., Hoogeveen, J., Faures, J.M., Frenken, K., Feick, S., 2005. Development and validation of the global map of irrigation areas. Hydrol. Earth Syst. Sci. 9, 535-547.

Siebert, S., Hoogeveen, J., Frenken, K., 2006. Irrigation in Africa, Europe and LatinAmerica. Update of the Digital Global Map of Irrigation Areas to Version 4. In: Institute of Physical Geography, U.o.F., Frankfurt am Main, Germany, (Ed.), Frankfurt Hydrology Paper p. 134.

Snijders, T., Bosker, R., 1999. Multilevel Analysis: An Introduction to basic and advanced multilevel modeling. Sage Publications, London.

Swets, J., 1988. Measuring the accuracy of diagnostic systems. Science 240, 1285-1293.

Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M., Gangalakunta, O.R.P., Turral, H., Cai, X., Vithanage, J., Schull, M.A., Dutta, R., 2009. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. International Journal of Remote Sensing 30, 3679 - 3733.

Vance, C., Iovanna, R., 2006. Analyzing spatial hierarchies in remotely sensed data: Insights from a multilevel model of tropical deforestation. Land Use Policy 23, 226-236.

Veldkamp, A., Fresco, L.O., 1997. Reconstructing Land Use Drivers and their Spatial Scale Dependence for Costa Rica (1973 and 1984). Agricultural Svstems 55, Pp. 19-43.

Veldkamp, A., Verburg, P.H., Kok, K., Koning, G.H.J.d., Priess, J., BergsmaA.R., 2001. The need for scale sensitive approaches in spatially explicit land use change modeling. Environmental Modeling and Assessment 6, 111–121.

Verburg, P., Eickhout, B., van Meijl, H., 2008. A multi-scale, multi-model approach for analyzing the future dynamics of European land use. The Annals of Regional Science 42, 57-77.

Wisser, D., Frolking, S., Douglas, E.M., Fekete, B.M., Vörösmarty, C.J., Schumann, A.H., 2008. Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets. Geophys. Res. Lett. 35, L24408.

Wood, S., Sebastian, K., Scherr, S.J., 2000. Pilot analysis of global ecosystems: agroecosystems. International Food Policy Research Institute and World Resources Institute, Washington, DC.

Worldbank, 2009. Governance Matters 2009. Worldwide Governance Indicators, 1996-2008.

Wriedt, G., van der Velde, M., Aloe, A., Bouraoui, F., 2009. A European irrigation map for spatially distributed agricultural modelling. Agricultural Water Management 96, 771-789.

You, L., Wood, S., Wood-Sichra, U., 2009. Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. Agricultural Systems 99, 126-140.

Zeitoun, M., Allan, J.A., Mohieldeen, Y., 2010. Virtual water `flows' of the Nile Basin, 1998-2004: A first approximation and implications for water security. Global Environmental Change 20, 229-242.

Figure captions

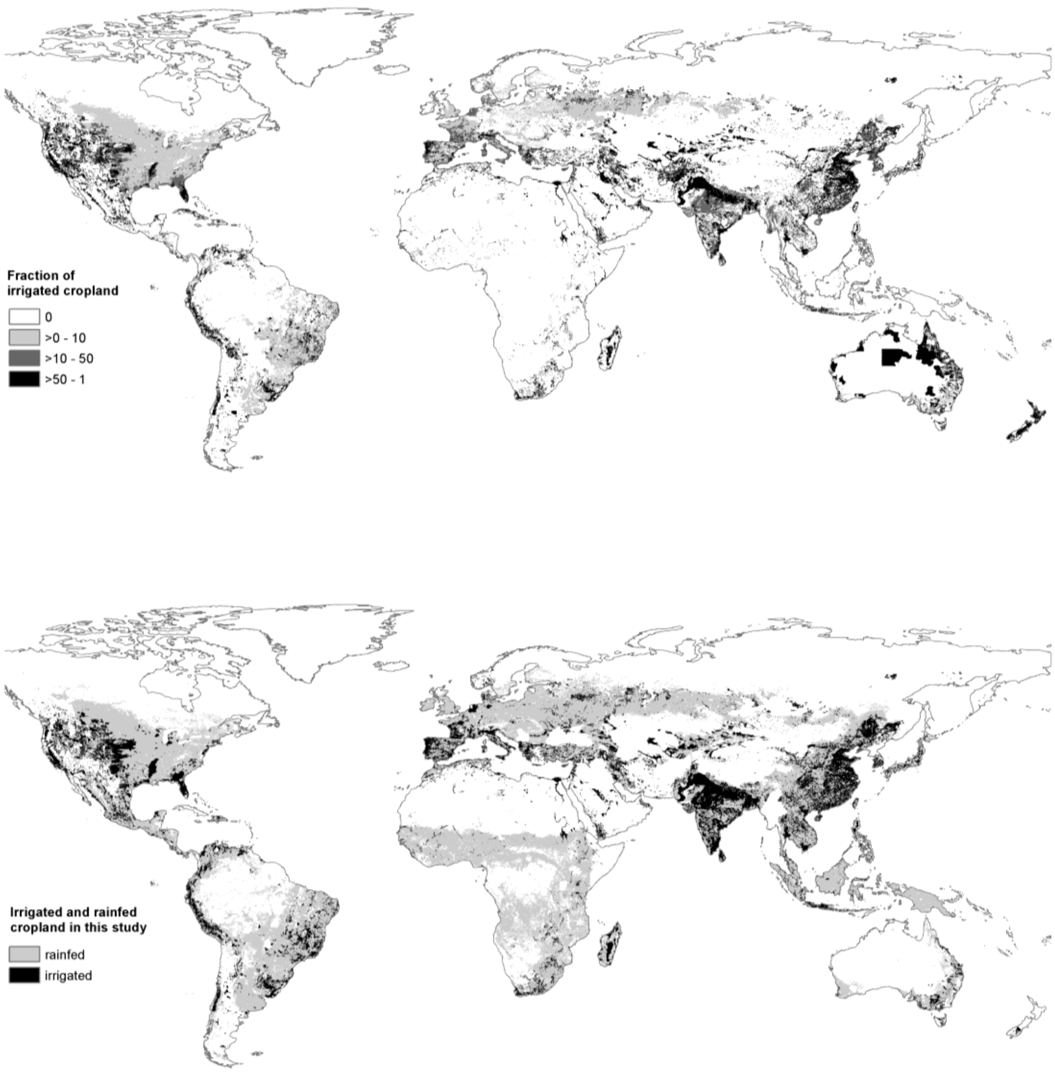

Figure 1: Fraction of irrigated cropland according to MIRCA2000 (upper map) and irrigated cropland and rainfed cropland as considered in this study (lower map).

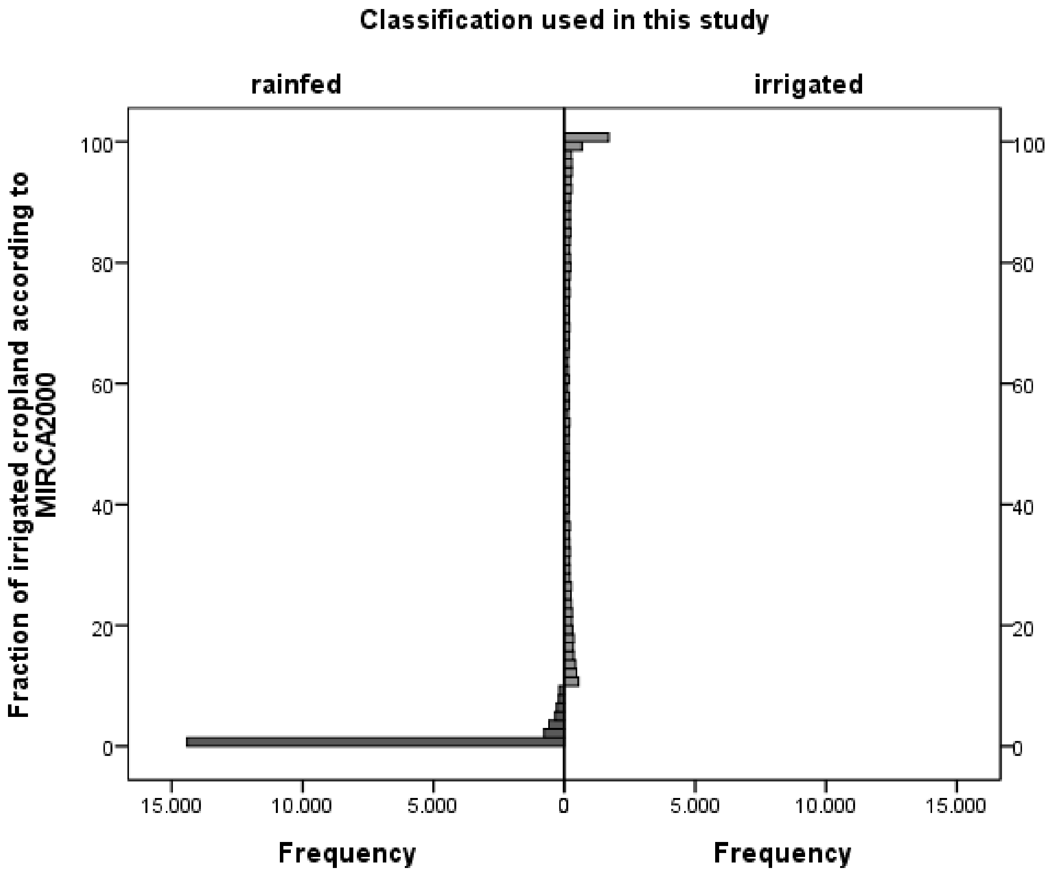

Figure 2: Distribution of grid cells classified as irrigated cropland and their definition in this study.

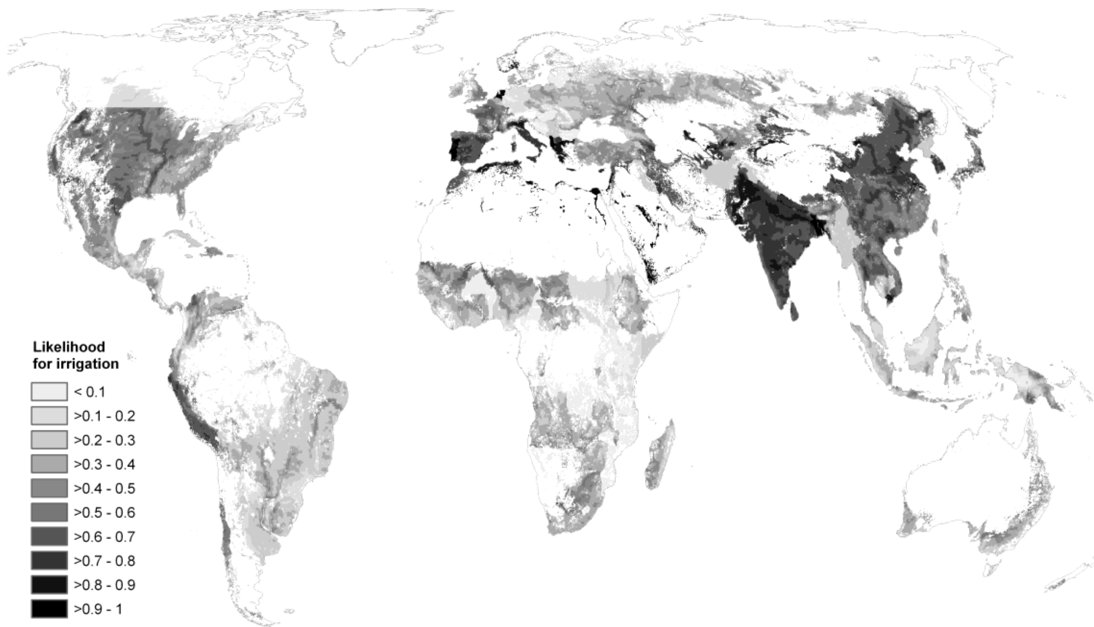

Figure 3: Likelihood of current cropland being irrigated, considering biophysical information only (model 1).

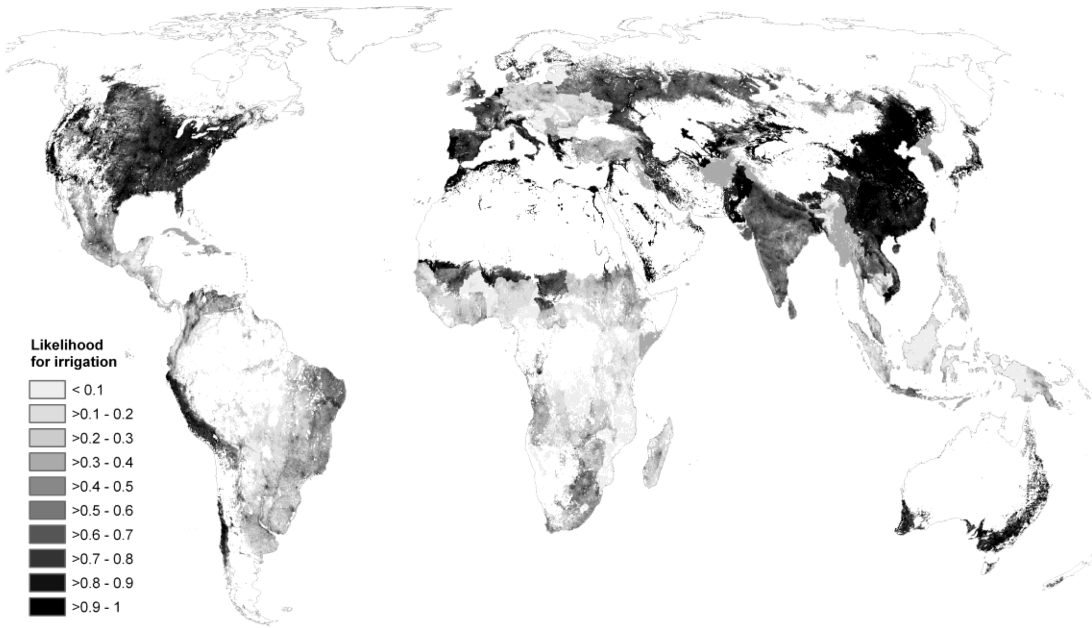

Figure 4: Likelihood of current cropland being irrigated, considering biophysical, socio-economic and governance information (model 2).

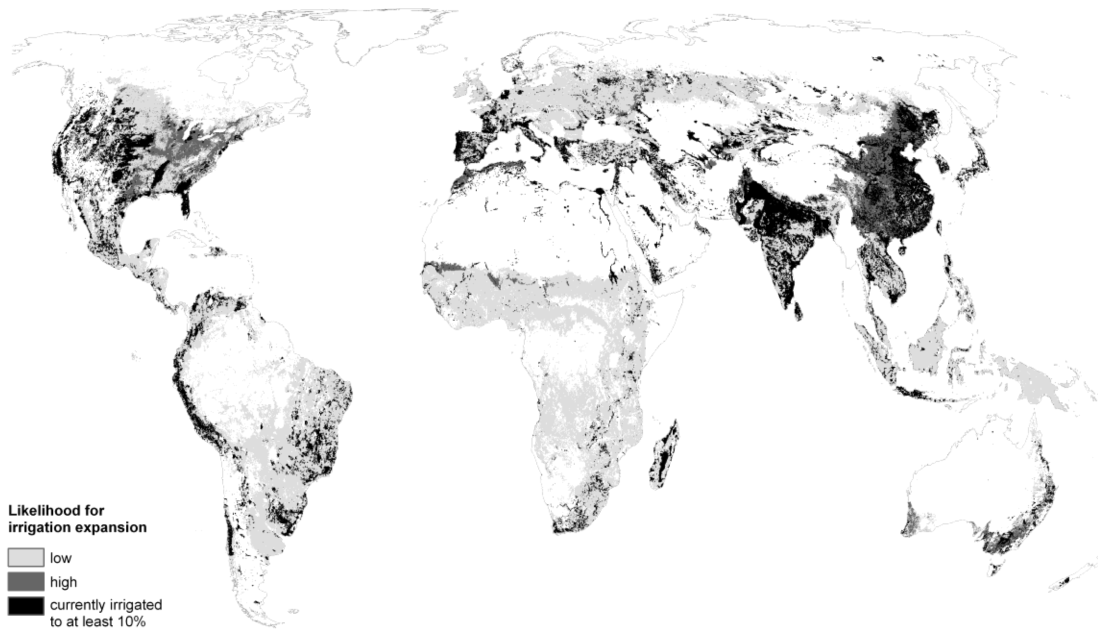

Figure 5: Potential areas for irrigation expansion.