Natural Environment Valuation Online Tool

Technical Documentation

Version 1.0, June 2019

Chapter 1: Agriculture Model

Land, Environment, Economics and Policy (LEEP) Institute

University of Exeter

Corresponding authors

Carlo Fezzi C.Fezzi@exeter.ac.uk

Ian Bateman

David Hadley

Amii Harwood

Authors

Brett Day

Amy Binner

Nathan Owen

Ian Bateman

Greg Smith

Patrick Collings

Louis Haddrell

Lorena Luizzo

Collaborators

Forestry Commission / Forest Research

UCL

JNCC

University of Aberdeen

1. The agricultural module: Land division, crops and livestock

1.1. Agriculture in NEVO

The agricultural model used in NEVO is a structural model of agricultural land use and production for Great Britain estimated using Farm Business Survey (2005 – 2011) and June Agricultural Census data. The model predicts the decisions made by profit maximising farmers about how to allocate their agricultural land between arable and grassland and, within these uses, what crops and livestock to produce. To achieve this the module links a top level land allocation model and two production models (arable and livestock).

The top level land allocation model predicts the division of agricultural land into arable and grassland production taking inputs on the environmental characteristics of the land, such as the soil type and quality, whether the farm is in an environmentally sensitive area or located on the greenbelt, and climate data, including data on temperature and rainfall.

The production models take the outputs of the top level model along with information on output prices, fertiliser prices, distance to market, environmental characteristics and climate to predict the share of arable land devoted to different crops and the heads of livestock produced. The arable production model predicts the production of 8 different crop types (wheat, winter barley, spring barley, oil seed rape, potatoes, sugar beet, protein crops (peas and beans) and other). Land use shares devoted to each crop are scaled by average yields and then multiplied by time dependent output prices (SOURCE) to calculate the value of food production from crops. The livestock production model predicts the production of 3 livestock types (Beef, Dairy and Sheep) in terms of the number of livestock heads. Total farm profits are calculated by adding profits from crops and livestock together, which themselves are a function of crop and livestock prices.

The agricultural model is able to predict outputs on an annual basis. In NEVO future climate and output price projections from UKCP09 are used to predict agricultural decisions, the background model calculates these for each year in the future, these background outputs are then converted into predicted average annual farm profits in each decade, in pounds, which are displayed in the details panel in the NEVO tool. Profits are discounted into the future using a default discount rate of 3.5%.

Farm prices are taken from the John Nix Pocketbook.

Limitations

- Agricultural woodland is assumed to be exogenous and is not included in the top level model which allocates agricultural land to arable or grassland uses.
- Constant returns to scale allows easy up and down scaling to different spatial scales – models were initially estimated at farm level using average 10km input data on environmental and climate characteristics.
- The agricultural model assumes crop yields are constant across space and are fixed across the four decades. Food production from livestock is currently not included.

1.2. Introduction

The agricultural land use component in NEVO builds upon the approach developed by Fezzi and Bateman (2011) at the heart of the UK National Ecosystem Assessment (NEA, Bateman et al., 2013) and UK NEA Follow On studies (Bateaman et al., 2014). It extends the original approach by using Farm Business Survey (FBS) data in order to estimate agricultural input and output price effects on farm land use and livestock production. On the other hand, it still maintain the spatial resolution (2x2 km grid squares) necessary to fully capture the effect of fundamental drivers of land use such as soil, terrain and climate. To the best of our knowledge, this is the first spatially explicit, econometric analysis using farm level data rather than agricultural census data for the United Kingdom. This model merges a structural economic framework based on a multi-output profit function (see Fezzi and Bateman, 2011, for a detailed illustration) with sounds statistical analysis combining the rich FBS panel (more than 2000 farms for the years 2005-2011), agricultural census data (50,000 2x2 km cells every year covering the entirety of Great Britain from 1976 onwards) and GIS information on the environmental determinants of land use including soil quality, climate and land gradient. The model can be used to evaluate changes in land use and farm profitability generated by a number of different drivers, including climate, prices and policy. The user can change any of these drivers to simulate complex scenarios.

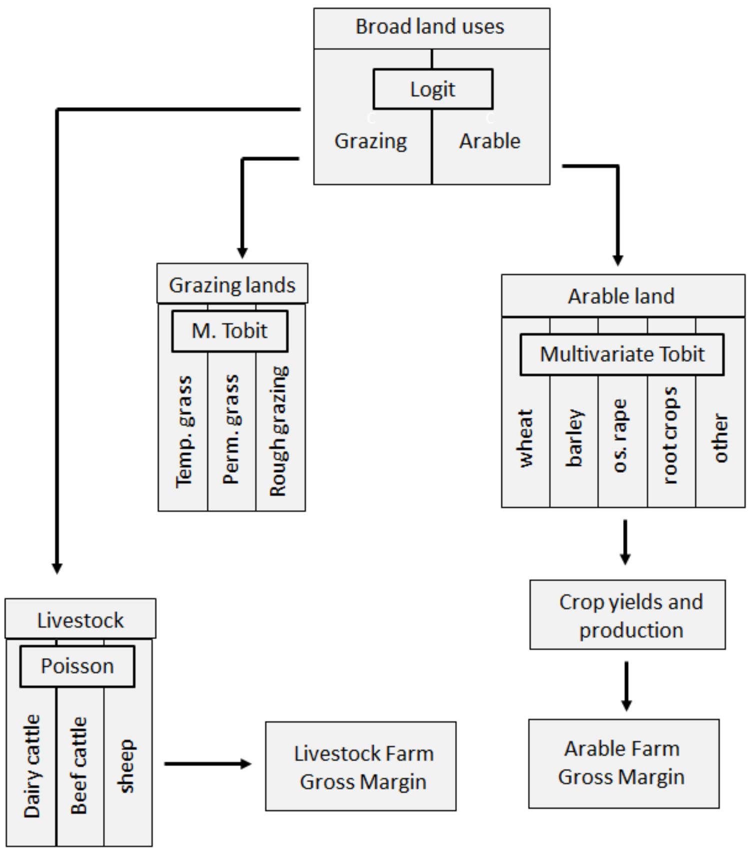
This modelling framework follows a stepwise approach. The first-step separates all agricultural land (with the exception of farmed woodland, which is simply assumed fixed) between grassland and arable. In the second step, arable land use is allocated among wheat, barley, oilseed rape, root crops and other land (including other arable, fallow, uncropped etc.). Grassland land use, on the other hand, is allocated among temporary grassland, permanent grassland and rough grazing. Livestock numbers and density for dairy cattle, beef cattle and sheep are also calculated for the grassland areas. Finally, using historical data on yield and prices we compute farm gross margins for both arable and livestock farms.

The drivers of this land use model are: climate, other environmental determinants (e.g. soil types, slope, etc.), input and output prices and policy measures (e.g. environmentally sensitive areas, nitrate vulnerable zones etc.). Model predictions are at the 2x2km grid square level. Details on model estimation and predictions are provided in the technical report in the following pages.

2. Overview

Figure 1 provides a diagram representing the overall structure of our modelling framework. While some of these models are estimated on June Agricultural Census (JAC) data and some other on Farm Business Survey (FBS) data, all the models are used for predictions at the 2x2 km level, matching the resolution of the other NEVO modules. This framework can be used to evaluate changes in land use and farm profitability generated by a number of different drivers, including climate, prices and policy. The user can change any of these drivers to simulate complex scenarios.

Figure 1: Modelling approach



At the **first step**, the model separates all the available agricultural land (with the exception of farmed woodland, which is not modelled here) between grassland and arable. This model is estimated as a quasi-maximum likelihood (QML) Logit model on 2x2 km grid JAC data from 1972 to 2010 (11 unevenly spaced years) coupled with climate, environmental, price and policy information (see the data Section 10 for more details).

In the **second step**, arable land use is allocated among wheat, barley, oilseed rape, root crops and other land (including other arable, fallow, uncropped etc.). Grassland land use, on the other hand, is allocated among temporary grassland, permanent grassland and rough grazing. Both models use the multivariate Tobit specification developed by Fezzi and Bateman (2011). However, the arable model is estimated on 2005-2011 FBS data, which contains rich information on prices, while the livestock model is based on AC data. The reason is that livestock data in the FBS is too volatile and not necessarily correlated with prices and livestock heads. For example, in one year the revenues from cattle can be high and the number of cattle low because the farmer sold many livestock at the beginning of the year without replacing them, and vice versa, revenues can be even negative if the farmer has bought young cattle to be raised and sold in the following years for profit. In addition, there is no information on the weight of the animals, which strongly influence prices. For this reason, it is very hard to calculate meaningful farm-level price indexes with this data. JAC data coupled with regional price indexes smooth away this important source of variability allowing the estimation of our livestock and grassland models.

In the **third step**, we calculate gross margin. For arable farms, we use information on yield and prices to calculate revenues. We then simply assume that gross margin corresponds to 45% of revenues, a proportion that has been surprisingly stable in the past 15 years (more information in the following sections). Regarding livestock, we also assume gross margins to be related to output prices and historical information.

The next sections illustrate in detail each modelling component. We then show how this approach performs in predicting the spatial patterns of land use in GB. We also present a few climate change scenarios to further illustrate both the capabilities and the caveats of our framework.

3. The broad land use model

The broad land use model (BLU) model allocates agricultural land between arable and grassland. The only type of agricultural land use not included is woodland, which is determined outside this NEVO module. BLU is estimated on JAC data from 1976 to 2010. We include 11 unevenly spaced years, since this data is not available to the public for every year. The spatial extent of this data is the entirety of GB, which generates more than 55,000 cells each year. In order to address potential spatial autocorrelation, we follow the approach by Fezzi and Bateman (2011) and sample a cell every four along both the orizontal and vertical axis, in effect using only about 6% of our cells. This still leaves more than 25,000 observations for our analysis.

Since the dependent variable is the share of agricultural land devoted to arable (and not a 0/1 dummy variable) the correct specification is a QML Logit (Wooldridge and Papke, 1996). The explanatory variables are rainfall and temperature during the growing season (April to September), soil types, slope and other environmental determinants, policy (e.g. share of the agricultural land in the 2x2 square defined as nitrate vulnerable zone) and the ratio of one-year lagged wheat (the major GB arable crop) price and fertilizer price. Lagged price is used because at the time when cropping decisions are taken the future price is not known. To capture potential non-linear effects (e.g. Fezzi et al., 2015) we model rainfall and temperature as piecewise linear functions. We determine the location of the knots (i.e. the points that connect the various linear "pieces") of such function by fitting a generalized additive model with smoothing cubic regression splines prior to the estimation of BLU (Wood, 2006). Finally, we do not model directly the effect of extreme heat (i.e. > 29/30oC, see for example Schlenker and Roberts, 2009) since days with high temperature are very rare in GB. For example, the highest monthly average maximum temperature in our climate data is barely 23oC. Conveniently for our analysis, accounting for irrigation is not an issue in GB, since even in the driest areas most of the farmland is rainfed.

Table 1 reports the estimates. This model has a satisfactory fit, as shown by the extremely high pseudo-R2.[1] As we would expect, the parameter of price is strongly significant and positive. In other words, the share of arable land increases when the prices of arable crops (which are all strongly correlated to wheat) increase compared to the price of fertilizer. The effect of the environmental variables are also in line with our expectations, i.e. flatter land and better soils increase the share of arable. Likewise, policy determinants that impose constraints on agricultural activities (e.g. environmentally sensitive areas, national parks etc.)

[1] Technically, the pseudo-R^2 and other likelihood-based test are not strictly valid in a QML framework. However, we report this measure anyway since it gives an indication of the fit of the model.

reduce the share of arable. Most of the piecewise linear coefficients of rainfall and precipitation are estimated to be highly significant. We plot the relations in Figure 1.

Table 1: BLU model parameter estimates

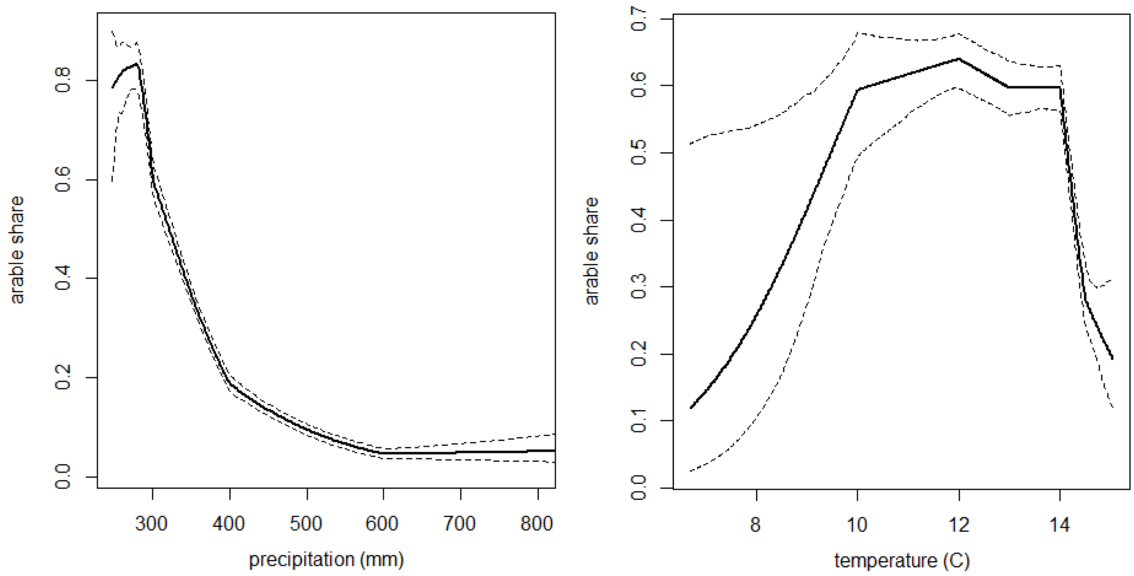
	Estimate	Std.error	Z-test	P-value	
(Intercept)	-15.279	11.523	-1.326	0.185	
rain	0.032	0.043	0.747	0.455	
rain >= 260	-0.037	0.053	-0.705	0.481	
rain >= 280	-0.059	0.022	-2.723	0.006	**
rain >= 300	0.042	0.009	4.662	<2e-16	***
rain >= 400	0.011	0.001	7.931	<2e-16	***
rain >= 600	0.009	0.001	8.848	<2e-16	***
temp	0.636	0.325	1.957	0.050	*
temp >= 10	-0.631	0.313	-2.014	0.044	*
temp >= 12	-0.269	0.127	-2.113	0.035	*
temp >= 13	0.146	0.127	1.148	0.251	
temp >= 14	-2.320	0.243	-9.555	<2e-16	***
elev	-0.003	0.000	-8.087	<2e-16	***
slope	-0.050	0.010	-4.935	<2e-16	***
ph	0.247	0.034	7.264	<2e-16	***
npark	-0.004	0.001	-3.567	<2e-16	***
esa	-0.003	0.001	-4.049	<2e-16	***
greenbelt	-0.001	0.001	-1.925	0.054	.
dist300	-0.001	0.000	-3.030	0.002	**
s_peat	-0.199	0.173	-1.151	0.250	
s_gravel	-0.599	0.123	-4.887	<2e-16	***
s_stoney	0.140	0.077	1.814	0.070	.
s_fragipan	-1.224	0.123	-9.916	<2e-16	***
s_coarse	0.282	0.062	4.565	<2e-16	***
s_fine	-0.315	0.053	-5.938	<2e-16	***
trend	0.013	0.000	28.443	<2e-16	***
lag(wheat_p)/fert_p	0.888	0.019	47.813	<2e-16	***
rain*temp	0.000	0.000	1.134	0.257	
pseudo R-sq.(adj) = 0.77					
Deviance explained = 76%					

Notes: . / * / ** / *** indicate significance levels from 10% to 0.1%. Model estimated via QML. N = 25444. Variable definitions reported in Section 10.

As shown in Figure 1, the effects of both precipitation and temperature on arable land share are highly non linear. The right-hand plot shows that precipitation initially increases arable shares, until values of about 300mm. After this threshold, the effect is reversed, becoming strongly negative. The level of 300mm may appear to be very close to the observed minimum level of rainfall in our data (240mm) but the current GB climate is characterized by relatively abundant precipitation. In fact, irrigation is extremely uncommon in the British islands. On the other hand, an excess of precipitation can create soil waterlogging, delay ploughing and sowing operations, and diminish plant growth.

The effect of temperature, displayed in the right-hand side plot, has an inverted U-shape. The positive effect of warm temperatures on plant growth appears to increase strongly the share of arable crops until an average temperature during the growing season of about 14oC, while we observe a significant negative effect afterwards. After 14.5 oC the confidence intervals become very wide, since summer and spring are very mild in GB and our sample does not contain many observations above this last threshold.

Figure 2: The estimated impact of temperature and precipitation on the share of arable land



Notes: dashed line indicate 95% confidence intervals

4. The arable model

The arable models allocate arable land among five different land uses: wheat, barley, oilseed rape, root crops and "other land uses", which includes all other arable crops, vegetables, uncropped and fallow land. For estimation, we use FBS data from 2005 to 2011. Since year 2005, the UK government removed arable area subsidies and replaced them by the single farm payment, allowing us to model directly the impact of prices on crop allocation. In line with the FBS classification, we define arable farms as enterprises that have at least 70% of their revenues derived from crops. In addition, we restrict our analysis to farms that cover at least 5ha. This selection leaves a bit more than 3000 farms-years records for our analysis. Considering these selected farms, wheat is the major land use, covering almost 50% of the arable area, and oilseed rape is the second, with almost 16%. The miscellaneous group of land uses included in the "other" category, taken together sum to less than 15% of the total arable land.

Unlike the JAC, FBS includes only farms located in England. Since we know the approximate location of each farms on a 10x10km grid square grid, we are able to link FBS accounts to spatial information on physical environment conditions (soil, slope, etc.), climate, policies etc. These variables are practically the same ones we use in the BLU model.

Regarding prices, we calculate them as follows. For each farms and crop, price is given by the ratio between revenues and production in the FBS. If a crop is not produced in that farm-year, we assign the average price in the county in that year. If it is not produced in the county, we assign the average price in the region, then in the nation. Prices can be highly variable with a few outliers. Since this could be caused by particular crop varieties, we removed from the analysis the farms for which the price is 50% higher or lower than the average price in the same county in the same year. Since land use share decisions happen before knowing crop prices, it is necessary to model expectations. As common in the literature, we use the price in the previous year, exactly as in the BLU model. Fertilizer is chosen as the numeraire of our normalized quadratic profit function. This means that the crop prices are defined as ratios with the fertilizer price.

The theoretical development of the model is illustrated in detail by Fezzi and Bateman (2011). Briefly, the starting point is a dual multi-output profit function which is used to derive land use allocation equations following the intuition that, in equilibrium, the shadow values of all the different land uses needs to be equal (Chambers and Just, 1989). Assuming a normalized quadratic profit function, we derive linear land use allocation equations. Since not all farms cultivate all crops, censoring is a significant issue for the estimation the parameters of such equation, because the standard linear regression approach would generate biased results. For this reason, again following Fezzi and Bateman (2011), we use a QML heteroskedastic Tobit specification. In this framework, in order to ensure sum to unity, one of the shares needs to be calculated as a difference between one and the sum of the other land uses. In our context, "other land" arises naturally as such share. Therefore, we will estimate equations for wheat, barley, oilseed rape and root crops only.2

It is important to make sure that the price parameters in our data are not affected by omitted variable bias. For instance, one can imagine that farms (or counties) with better soils receive

[2] An alternative approach is to to use a multivariate version of the QLM logit specified for the BLU model. However, this specification would require the development of an ad-hoc theoretical model. While this is a potentially interesting extension, in this work we remain with the established framework by Fezzi and Bateman (2011).

a price premium and, also, have a higher share of premium crop varieties. This would bias the price coefficients upwards. While these differences in land quality should be encompassed by our set of environmental and climate determinants, differences in farmers' ability and/or risk aversion, which could have a similar implications, may not be removed so easily. An approach to remove the variation in price potentially correlated with these omitted factors is to use panel fixed effects. Unfortunately, a fixed-effect Tobit model does not yet exist in the literature. However, we use a series of linear regressions to examine if these potential omitted variables are an issue in our context.

Since wheat is the major land use in our dataset we present detailed results for this crop (results for the other crops are analogous). Table 2 present the estimates of the effect of wheat price (normalized by the fertilizer price) estimated via a) ordinary least squares (OLS) , b) county fixed-effects, c) farm fixed-effects and d) the QML Tobit model (our final specification). The first three approaches are downward biased because they ignore the censoring, while the last one is biased only if omitted variables are a significant concern.

The first row reports the OLS parameter, which is about 15. Considering that the average winter wheat price in the sample is 111$/tonne, and that the average price of fertilizer is 240$/tonne, this mean that if there is an increase of wheat price of 50$, the effect the average farmer will allocate an additional 3.1% of his total arable area to wheat.

Table 2: Wheat equation own-price effect

MODEL	Coef.	Std. Err.	x/t-stat	p-value
OSL	14.92	4.55	3.29	0.001 **
COUNTY FE	9.54	4.05	2.35	0.019 *
FARM FE	9.57	2.94	3.25	0.001 **
QML TOBIT	10.49	4.11	2.55	0.011 *

Notes: . / * / ** / *** indicate significance levels from 10% to 0.1%

The second row reports the panel county level fixed effects results (note that while not reported in the table, these fixed effects capture 50% of the variability of wheat share). The price effect is reduced to about 10, i.e. 2/3 of the OLS coefficient. Therefore, it appears that a significant part of the variability of price is indeed correlated with the quality , i.e. counties with better land have both an higher share of wheat and enjoy higher prices (for example from premium varieties).

The third row reports farm fixed-effect estimates (this fixed effect captures more than 70% of the variability of price). The coefficients is practically the same as the one obtained using county fixed effect. This means that the farm variability within county is not significantly correlated with the variability in prices. In other words, it appears that differences in abilities and attitudes across farmers should not significantly bias our estimates. In addition, our environmental and climatic variables should be able to capture the relevant farm heterogeneity, because they are defined at a level (10x10km cells) that is finer than the county fixed effects and include all the main drivers of agricultural land use. This intuition is confirmed by the coefficient estimated by our QML Tobit reported in the fourth row: the parameter is not significantly different from the estimate obtained using farm fixed effects. Altogether, these results strongly support the ability of our QML Tobit approach to estimate both price and climatic effects within one consistent framework.

Table 3: Arable model estimates

Variable	**Wheat**		**Barley**		**Oilseed rape**		**Root crops**	
Wheat price	12.44	**	-7.26		-8.66	**	-11.20	.
Barley price			0.00					
Osr price	-4.42	*			6.45	***		
Potatoes price							5.33	.
elev	0.04	***	-0.03		0.08	***	-0.20	***
% slope > 6^o	-0.01		0.00		-0.30		-1.57	**
depth to root	0.14		-0.25	***	0.09	***	-0.29	***
s_carbon5	-0.02		0.31	***	0.04		-0.28	***
s_carbon6	0.00		-0.08	**	-0.11	***	0.15	**
Ph	161.30	**	-290.23	**	20.44	***	12.56	
Ph^2	-26.70	**	55.58	***	-1.57	***	-11.35	
Ph^3	1.47	**	-3.46	***	0.00		1.07	
rain	-1.60	**	-1.43		-2.67	***	0.04	
$rain^2$	0.00	***	0.00		0.00		0.00	
temp	-90.48	*	-1.57		-183.37	***	207.97	**
$temp^2$	2.90	*	-1.63		4.70	***	-7.10	**
Rain*temp	0.04		0.12	***	0.17	***	-0.02	
s_stony	0.06		-0.24	***	0.04		0.51	***
s_gravelly	-0.06	.	-0.13	**	0.06	**	0.14	.
s_notex	0.41	***	-0.32	***	-0.28		-0.66	***
npark	-0.12	.	0.04		0.05		-0.42	**
esa	-0.20	***	0.30	***	0.03		0.08	**
greenbelt	-0.07	***	0.10	***	-0.03		-0.01	
Dist_300	-0.12	***	0.16	***	-0.07	***	0.10	***
Dist_sb	-0.01		0.03	***	-0.01		-0.73	***
Dist_sb >= 20							0.64	**
Dist_sb >= 40							-0.45	***

Dist_sb >= 80							0.71	***
Dist_sb >= 120							-0.37	***
island	4.53		-22.90	**	-0.59		-17.52	
trend	-0.09		0.20		0.53	**	0.01	
intercept	617.79		815.12		1623.24	***	-1313.49	
N	2371		2392		2381		2376	
pseudo-loglik	-9661.93		-7414.58		-6504.80		-4519.99	

Notes: Heteroskedastic Tobit models estimated via QLM (the parameters of the variance equation are not reported but include regional dummies, total arable area and trend). Symbols . / * / ** / *** indicate significance levels from 10% to 0.1%. Variable definitions reported in Section 10.

The estimates of the four crop equations are reported in Table 3 (“other” is calculated as a difference). In each equation, we include own-price effect, an alternative-crop price effect (in the wheat equation oilseed rape price, and in all other equations wheat prices, the main GB crop) and fertilizer price as the numerarie. Multicollinearity is an issue if we include more than three prices, since agricultural output prices are highly correlated. Therefore, this specification allows us to capture own price, cross-price and input price effect in a parsimonious way. Results are in line with our expectations, with an increase in wheat price having a negative effect on the shares of other crops that, on the other hand, respond positively to an increase of their own prices. Since this is a non-linear model, the magnitude of such effects is not constant. However, we can calcualate average effects. For example, considering that the average wheat price in the sample is 111$/tonne, and that the average price of fertilizer is 240$/tonne (we consider a 20.10.10 mix blend), our estimate means that an increase of wheat price of 50$/tonne translates into a wheat share increase of about 3% of the arable area (not the total farm area).

Furthermore, root crops (potatoes and sugar beet) are the ones that generate the highest revenues per hectare and, in fact, are more abundant on flatter land, better soils and more advantageous climatic conditions. In addition, sugar beet cultivations are strongly affected by the nearby presence of processing farms, given the considerable transportation costs. For this reason in the root crop equation we included a piecewise linear function of distance to the closest sugar beet factory, which proved to be highly significant. Finally, the number of observations across model differs because of the farms eliminated for having prices too low or too high compared to the average.

5. The grassland model

The grassland model separates total grassland into temporary grassland, permanent grassland and rough grazing. Temporary grassland is grass being sown every 3 to 5 years, it receives significant input of fertilizer and it usually used for dairy or premium beef cattle. Permanent grassland is maintained perpetually without reseeding but can also receive some fertilizer input. Rough grazing is the least intensive land use, it does not receive any fertilizer application and it is mainly used for raising sheep. Even though these grassland types do not have a direct effect on farm gross margins (which, in our framework, are calculated as a function of crop production and livestock heads only) they still generate different environmental impact and are used as inputs to the other modules of NEVO. The grassland model follows the same approach as the arable one, i.e. it is based on Fezzi and Bateman (2011). However, it is estimated on JAC data (the same used to estimate the BLU model) rather than FBS data, for the reasons illustrated in Section 1. An additional motivation is that FBS only refers to England while both Scotland and Wales contains very significant grassland areas and almost the entire share of rough grazing of GB. As in the arable models, fertilizer price is used as the numerarie, i.e. all prices are calculated as ratios, with fertilizer price at the denominator.

While in the arable model, the "other land" category raised naturally as the share to be calculated as a difference, with grassland there is not such clear-cut distinction. After testing the three different possible specifications, we opted to explicitly model temporary and permanent grassland, and calculate rough grazing as a difference, since this solution was the one generating the smallest prediction errors.

Table 4: Grassland model estimates

Variable	**Temporary grassland**		**Permanent grassland**	
Milk price	15.40	***	-28.16	***
Beef price	1.57	***	6.20	***
Sheep price	-1.10	**	4.23	***
Elev	-0.02	***	-0.04	***
Elev >= 200	-0.04	***	-0.11	***
Slope	-0.22	***	-0.16	
S_peat	-15.59	***	-11.14	***
Temp	-1.08		-5.49	
Temp >= 9	2.18		6.30	
Temp >= 10	11.22	***	4.64	
Temp >= 11	-6.47	***	14.48	***
Temp >= 12	-4.01	***	-20.77	***
Temp >= 13	-2.65	***	-1.02	
Temp >= 14	0.00		7.76	***
Rain	-0.02		0.24	***
Rain >= 300	0.11	**	-0.12	**
Rain >= 350	-0.16	***	0.04	.
Rain >= 400	0.04	***	-0.05	***
Rain >= 500	0.00		-0.15	***

Rain >= 600	0.04	***	0.05	***
Rain * temp	0.00		0.00	.
s_medium	12.52	***	-7.57	***
s_fine	21.45	***	-36.35	***
s_stoney	1.76	***	3.02	***
s_gravel	2.47	***	2.98	***
s_saline	-5.71	***	-0.71	
s_fragipan	-3.83	***	6.08	***
s_silt	-23.28	***	-6.94	
s_clay	-45.17	***	114.05	***
npark	-0.02	***	-0.05	***
esa	-0.01	**	0.00	
nvz	0.00		-0.01	**
greenbelt	-0.01	**	-0.02	**
dist300	0.00	***	-0.04	***
trend	-0.26	***	-0.06	**
wales	-1.30	***	7.29	***
scotland	10.53	***	-19.42	***
constant	23.56		15.69	
Pseudo log-lik	-85145.85		-98219.68	

Notes: Heteroskedastic Tobit models estimated via QLM (the parameters of the variance equation are not reported but include regional dummies, total arable area and trend). Symbols . / * / ** / *** indicate significance levels from 10% to 0.1%. Variable definitions reported in Section 10. N = 25444.

Estimates are reported in Table 4. An increase in milk price (which also takes into account the cost of purchasing milk quota) increases the share of temporary grassland, which is typically used for feeding dairy cows, while beef and sheep price have a positive effect on permanent grassland, which is mainly used for these two types of livestock. Elevation and slope have negative signs, since in upland areas they are replaced by rough grazing. Finally, as in Fezzi et al. (2015) we find strong and non-linear effects of rainfall and temperature.

6. The livestock model

The livestock model estimates the livestock density for dairy cattle, beef cattle and sheep. Here we do not distinguish among the different types of grassland but simply calculate the density as the total number of livestock heads for each type divided by the total grassland (i.e. the sum of temporary grassland, permanent grassland and rough grazing). For estimation, we use the same 1976-2010 JAC data that we have already used to estimate both the BLU and the grassland model. Since the JAC data are derived from parish information, in some rare cases they concentrate in some cells the entire livestock of the neighbouring area, thereby generating implausibly high livestock densities (see Fezzi and Bateman, 2011, for more details). For this reason we classify as outlier and remove from the analysis all cells in which dairy cattle density is higher than 4 heads/ha, beef cattle density is higher than 7 heads/ha and sheep density is higher than 30 heads/ha.

Analogous to the land use equations, livestock numbers can be derived from a multi-output profit function. Rather than equalizing land shadow prices, livestock density can be calculated by simply applying Hotelling's lemma (see Fezzi and Bateman, 2011). A normalized quadratic profit function generates linear livestock density equations. However, rather than applying the usual QML Tobit estimator, in this context we use a QML Poisson model. We compared both approaches and the Poisson provides a better fit and much lower prediction errors when used to project our estimates for the entire GB. Again, fertilizer price is the numeararie.

Table 5 reports the parameter estimates. Again, because of high multicollinarity we include only own-price effects that, as we would expect, are all positive. Climate and environmental variables are all strongly significant. Typically, the best conditions for grass growth increase the number of dairy and premium beef cows, while upland areas are characterized by mainly sheep farms. Again, in line with our expectations, policies aimed at reducing land use intensity (nitrate vulnerable zones, environmental sensitive areas, national parks) significantly reduce the density of all three livestock types. Finally, in the two cattle equations we include a trend and a dummy variable to capture the effect that the Bovine Spongiform Encephalopathy (BSE) outbreak had on the beef and dairy industries. Both variables are strongly significant and negative, generating a noticeable reduction in stocking rates.

Table 5: Livestock model estimates

	Dairy cattle		**Beef cattle**		**Sheep**	
(Intercept)	-66.630	***	-13.920	***	-15.780	***
milk price	5.316	***				
beef price			0.165	***		
sheep price					0.403	***
Rain	0.066	***	-0.011	***	-0.001	.
Temp	7.604	***	2.679	***	3.012	***
rain^2	0.000	***	0.000	**	0.000	***
temp^2	-0.242	***	-0.116	***	-0.131	***
Alt	-0.001	.	-0.002	***	-0.001	***
alt200	-0.009	***	0.001	***	0.000	***
Slope	0.031	.	-0.010	.	0.076	***
slope^2	-0.007	***	-0.001	**	-0.003	***
s_clay	1.921	**	1.214	***	0.724	***
s_silt	2.534	***	0.117		-0.154	
s_saline	-0.473	***	0.115	***	0.171	***
s_fragipan	0.198	*	-0.384	***	0.200	***
s_gravel	0.341	***	-0.043		0.217	***
s_stoney	-0.083	.	0.196	***	0.091	***

s_fine	-1.291	***	-0.510	***	-0.024	
s_medium	-0.769	***	-0.075		0.080	
s_peat	-0.819	***	-0.727	***	-0.311	***
Nvz	-0.001	.	0.000		-0.002	***
Esa	-0.002	**	-0.001	***	0.001	***
greenbelt	-0.001		-0.001	***	-0.005	***
dist300	-0.001	**	0.000	***	-0.001	***
npark	-0.001	*	0.000		0.000	
Wales	-0.251	***	-0.081	***	0.439	***
Scotland	-0.541	***	0.340	***	-0.514	***
Trend	0.004		0.004	*	0.002	***
Bse	-0.292	*	-0.581	***		
Trend_bse	-0.010	*	-0.029	***		
rain*temp	-0.003	***	0.001	***	0.000	.
N	26527		26530		26529	

Notes: QLM Poisson models. Symbols . / * / ** / *** indicate significance levels from 10% to 0.1%. Variable definitions reported in Section 10.

7. Farm gross margin

This section illustrates how we calculate farm gross margins (FGMs) from the hectares of agricultural land and livestock heads predicted by our model. FGMs are defined as the difference between revenues and variable costs. Due to the difficulty of allocating fixed costs (such as labour, machinery etc.) to specific activities within the farm, FGM is the standard measure of economic performance used in agricultural economics. Overall, this is a good indicator of farmers' self-sufficiency, albeit it does not take into account fixed cost. Note that our definition of FGMs does not include decoupled subsidies such as the single farm payment (now called basic payment scheme).

For arable land, we derive FGMs margin by first calculating revenues. To do so, we assume that the yield of each crop is constant across GB, and equal to the average yield in the FBS data. This is supported by the literature, which finds crop yield to be somewhat similar across GB regions (e.g. Lord, Anthony and Goodlass, 2002) and by the FBS data itself, which shows that year-to-year yield variation is much larger than difference across farms.

Therefore, to calculate production we can simply multiply predicted crop hectares by their average yields. We assume a yield of 8.19t/ha for wheat, 6.74t/ha for winter barley, 5.57t/ha for spring barley, 37.06 t/ha for potatoes, 60.66 t/ha for sugar beet and 3.53 t/ha for oilseed rape. Note that our arable model predicts total barley and root crops. Therefore, to calculate production we split barley between winter and spring varieties and root crops between potatoes and sugar beet by using the average historical shares of each crop within each region. Regarding "other" land uses, since this is a mixture of completely different crops, calculating production is not meaningful. For simplicity, we will assume its revenue per hectare equal to the one of wheat (data supporting this assumption are in Fezzi et al., 2010).

We calculate revenues by simply multiplying crop outputs and prices. Perhaps surprisingly, historical data (Defra, 2018) show that for arable farms the ratio of FGMs to crop revenues is remarkably stable over time. For example, during the past 10 years it always remained between 42% and 45%. Here for simplicity we obtain FGMs by multiplying revenues by 0.45.

Regarding livestock, we use historical data (Fezzi et al., 2010; Nix, 2009, and various years) to estimate a relationship between FGMs per head of livestock (dairy cattle, beef cattle, sheep) and price. For dairy, we estimate the following relationship: FGM/head = 183 + 20 * milk price (in pence/litre). This means that if the price is 20p/l (the lowest price during the years 2012-

2018) the FGM/dairy = 583£/head, while if price is equal to 35p/l (the very peak during the same period) FGM/dairy increases to more than 800£/head. We did not find a very strong relationship between price and FGM/head for beef and sheep, so we assume mean values in the Nix data, i.e. FGM/beef = 70£/head and FGM/sheep = 9£/head.

8. Validation

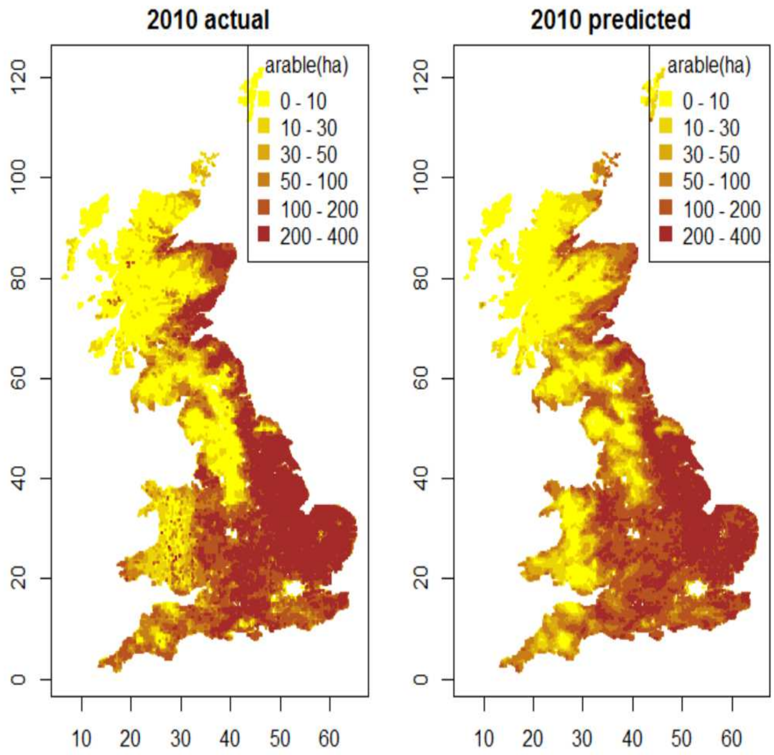
In this section, we use our modelling framework to predict the spatial patterns characterizing current land use in GB. We compare the model predictions with the 2010 agricultural census data (the most recent year for which 2x2 km cell information is available). Considering that to remove potential spatial autocorrelation we used only 6% of the cells in every year, this is primarily out of sample prediction. This test is particularly challenging for the arable models, since they are estimated on English FBS data and they are used to predict JAC data for the entire Great Britain. Not only the datasets are different, but also the spatial extent of the prediction area is much larger and diverse than the one used for estimation. Despite these difficulties, this section shows that the arable model generates very satisfactory predictions.

For prediction, our modelling framework follows the process we have already presented in Figure 1. In other words, for each 2x2km cell we first use the BLU model (section 2) to calculate the share of arable and the share of grassland. After that, we use the arable model (section 3) to divide arable land among wheat, barely, oilseed rape, root crops and other crops, and the grassland model (section 4) to separate grassland among temporary, permanent and rough grazing. Regarding grassland, by using the livestock model (section 5) we also calculate livestock heads for dairy cattle, beef cattle and sheep. We then calculate both arable and livestock gross margin (section 6).

Figure 3 compares 2010 JAC data with predictions from the BLU model. The model captures the overall spatial patterns of arable land use in GB, with the upland areas in the Peak and Lake District, in Scotland and Wales clearly visible in yellow. In these areas, particularly in Wales, the JAC data presents a few cells with high values for arable, which are clearly visible in the left-hand map as dark-brown points. These strange values are caused by the AC data manipulation approach (which may be in place also to preserve farmers' anonymity) rather than being areas with intense arable cultivations in the real world. In other words, these are not land use features that are not captured by our model, but rather artefacts in the JAC data. Even including these outliers, the overall mean absolute error (MAE) of the BLU model is 29.3

ha, which we believe to be a good result given that the standard deviation of arable land is 84.2 ha.

Figure 3: AC data and BLU model predictions



We now use the BLU model prediction as in input for our arable, grassland and livestock models and compare predictions with the JAC data. Table 6 reports measures of goodness of fit, while Figure 4 presents maps of selected variables.

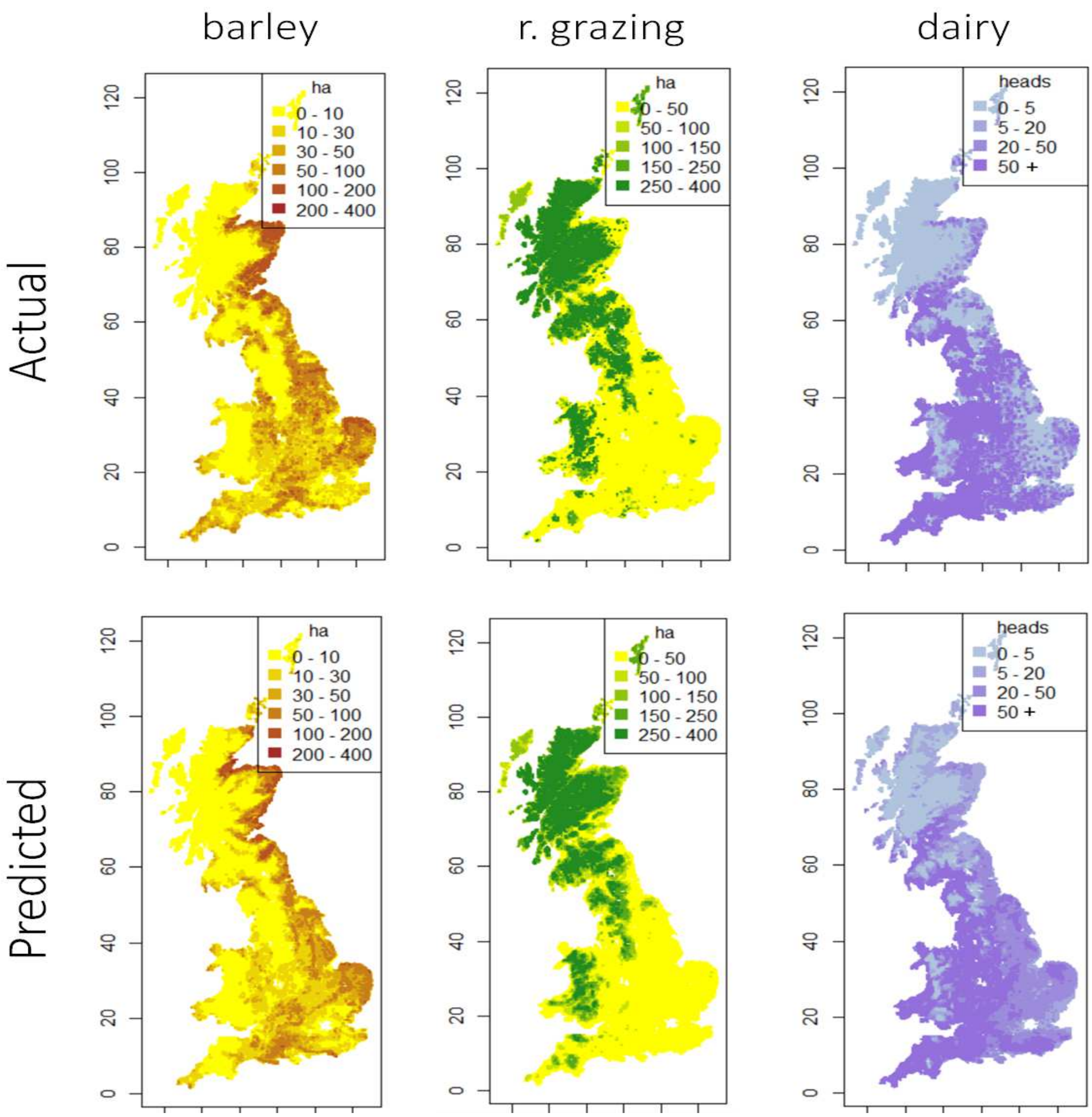
Table 6: prediction performance

Variable (units)	Standard Deviation	MAE
Wheat (ha)	46.52	14.56
Barley (ha)	26.87	11.57
Oilseed rape (ha)	13.79	6.32
Root crops (ha)	14.07	4.34
Other land (ha)	31.45	17.41
Temporary grassland (ha)	20.33	10.82
Permanent grassland (ha)	82.11	34.43

Rough grazing (ha)	119.46	27.16
Dairy cattle (heads)	82.97	35.40
Beef cattle (heads)	124.93	66.10
Sheep (heads)	771.34	315.48

Notes: statistics calculated using 2010 AC data (53811 2x2km cells).

Figure 3: AC data and model predictions



The results in Table 4 indicate a very good forecasting performance. The MAEs for all models are lower than half the standard deviations of the predicted variables (the only exception is the MAE for “other land” which is slightly higher than the half) and often much smaller. For example, rough grazing has a MAE of 27.2ha while its standard deviation is 119.5ha. With the exception of beef cattle, all MAEs are smaller than the ones of the corresponding equations in the Fezzi and Bateman (2011), signalling an improved fit. The arable model is particularly noteworthy: while estimated on FBS data, it still maintain a very good performance when used to predict JAC data. This finding is highlighted by the leftmost comparison in Figure 3. The barley model is able to predict very well the spatial distribution of this crop not only in England, but also in Scotland, even though the model is estimated only on English farms data. The predicted rough grazing maps, in the middle figure, is also very similar to the actual data above. Finally, dairy cattle, in the rightmost comparison, appears also to have a good fit, with the model capturing the East-West gradient, with the largest herds concentrated in the wetter part of the country. However, the JAC appears to be blockier and presents abrupt changes, while the model predict a smoother incline.

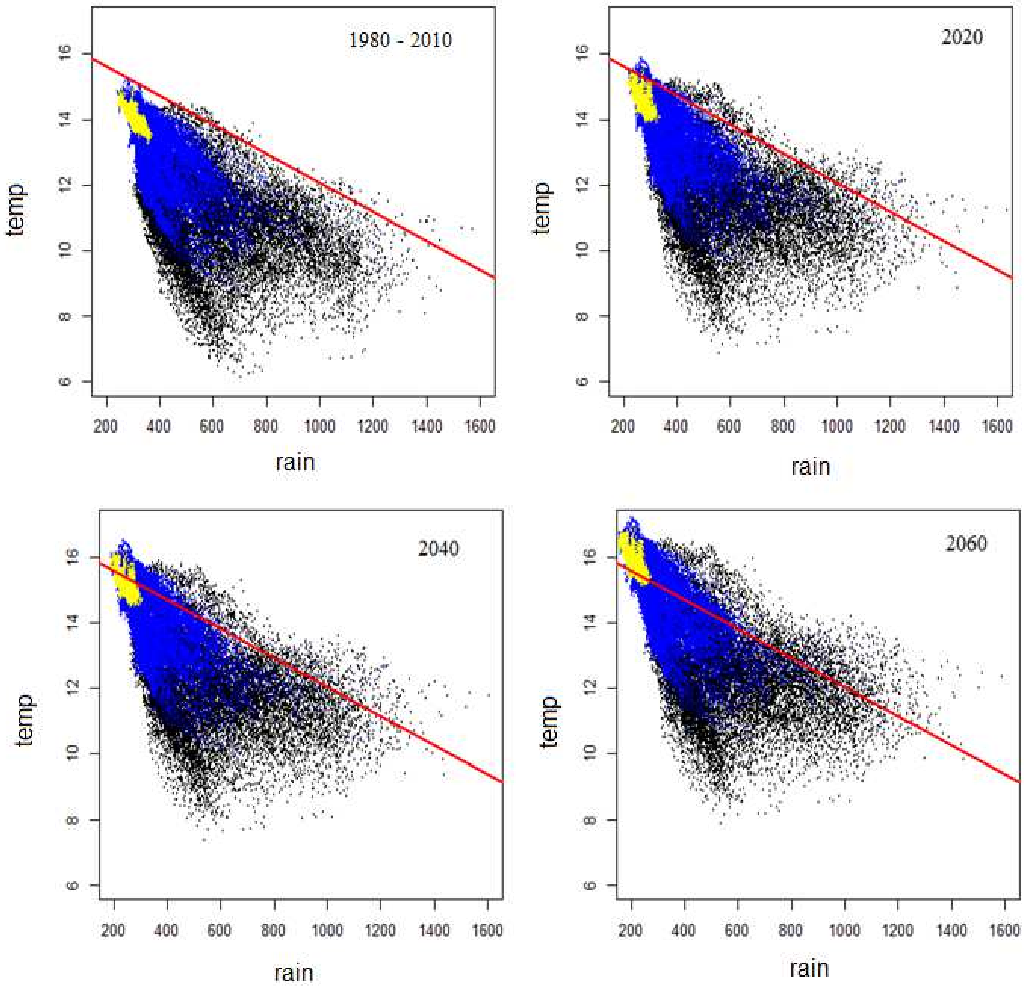
9. Climate change scenarios

The framework illustrate in the previous sections can be used to predict how agricultural land use would respond to changes in climate, prices and policy. In this section, we simulate a simple climate change scenario to a) illustrate both the capabilities and the weaknesses of our modelling approach and b) clarify the assumption behind our predictions. We focus on the BLU model, but these considerations are valid also for the other modelling components.

First, our scenarios are not projections of the future, but rather illustrate, ceteris paribus, the impact of climate change. These scenarios all technological responses (e.g. the introduction of new crop types), prices and policy constant at their values in the baseline, the year 2010. The only exception is irrigation, which, as illustrated below, is modelled rather crudely. In addition, we also leave unchanged all non agricultural land allocation and farm woodland, which is mainly driven by area-specific governmental policies. Finally, we do not allow for possible expansion of urban areas.

We use UK Climate Impacts Programme (UKCIP) (www.ukcip.org.uk), 25x25km grid climate change projections, the most up-to-date projections of future changes for the UK climate. As an illustration we use the projections for the years 2020s (averages from 2005 to 2035), 2040s (averages from 2025 to 2055) and 2060s (averages from 2045 to 2075) corresponding to the medium emission (SRES A1B) scenarios in the IPCC Special Report on Emissions Scenarios (Nakicenovic and Swart, 2000). We applied the projected changes to the baseline 5x5km data for years 1960-1990. Our models include average temperature and total rainfall in the growing season (the months from April to September).

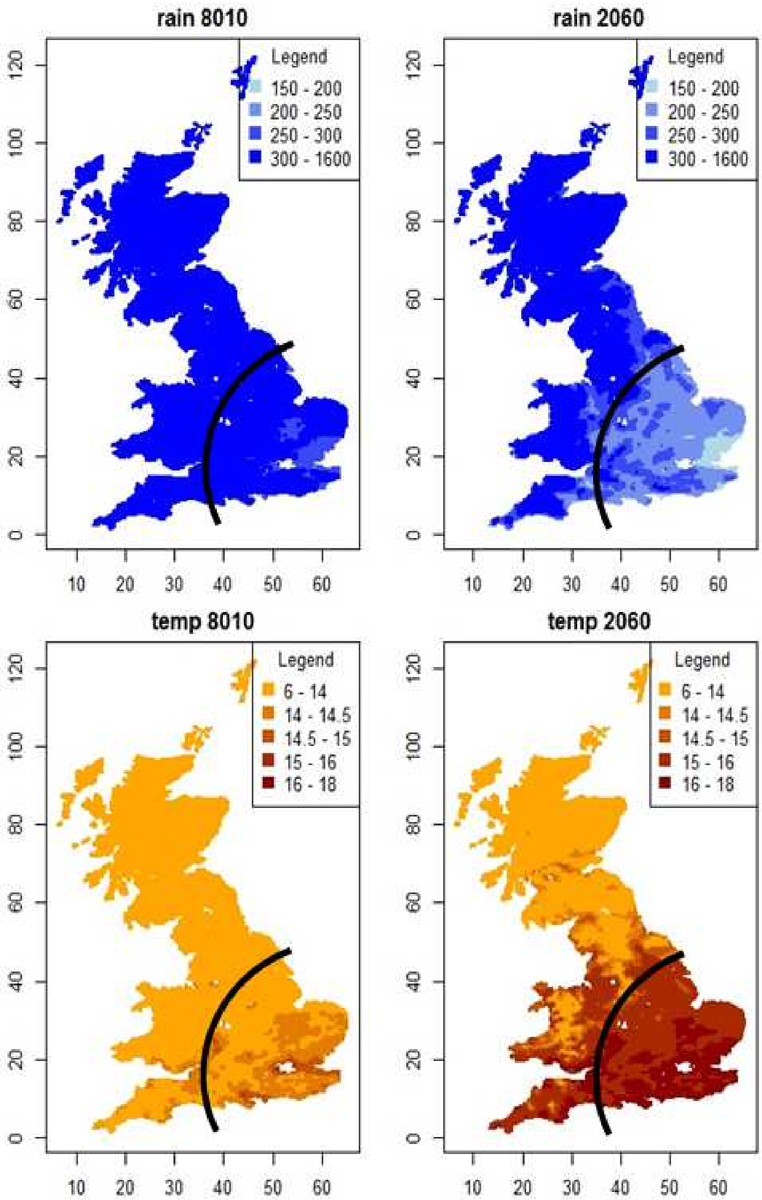
Figure 4: Current climate and climate change



Notes: temperature (oC) and rainfall (mm) during the growing season (April – September).

Figure 4 presents the scatter plot of rainfall and temperature in the growing season in the data used to estimate the BLU model and in the three climate change scenarios. The red line delimits the range of observed conditions in the estimation data. Yellow dots identifies the East of England (currently the driest and warmest region of the country), the blue dots the remaining of England and black dots are Wales and Scotland (the wettest and coldest areas of GB). Comparing the four plots, it is evident how climate change projections envisage GB becoming progressively warmer and drier. On average, by 2060 we see an increase of about 2.5C and a drop in rainfall of about 80mm. While these numbers do not look very impressive, the scatter plots show that by 2060 the entirety of East Anglia and other parts in England will experience conditions outside the range of data used for estimation, i.e. warmer and drier than anything we are observing right now. Interestingly, even some parts of Wales and even Scotland are projected to get warmer than the current South of England. Figure 5 provides a spatial illustration by mapping temperature and rainfall in the current climate (1980-2010) and in the 2060 scenario.

Figure 5: Current climate and 2060 climate

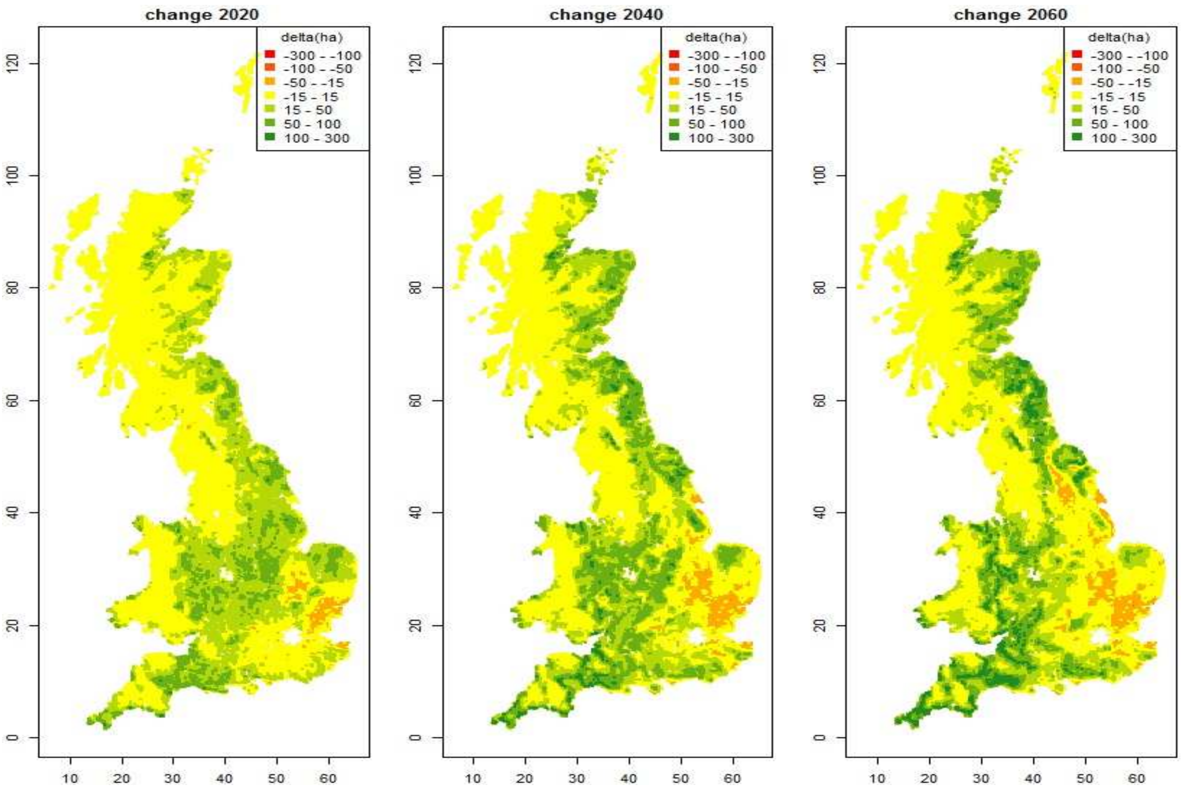


Notes: temperature (oC) and rainfall (mm) during the growing season (April – September).

These plots illustrate that, by 2060, the entire area delimited by the curved line is going to be outside the range used to estimate the model, considering both temperature and precipitation. For this reason, in this climate change scenario our projections for this part of England have to be taken cautiously. On the other hand, the projections for the West of England, Wales and Scotland should be more precise. Obviously, less extreme climate change scenarios will reduce the uncertainty. For example, if we use 2020s (2005-2035 averages) climate, we can provide accurate projections for the entirety of GB.

It is not advisable to extrapolate the non-linear relationships between land use and climate (which we previously reported in Figure 1) outside the range of the data used for estimation. Therefore, as standard in the climate change literature (e.g. Schlenker et al., 2005; Fezzi and Bateman, 2015; Albouy et al., 2016), we assume that the effect of climate flattens out outside such range (this approach is sometimes referred to as "restricting the climate" to the range used for estimation). In our specific case this is also supported by the wide (and widening) confidence intervals for the effect of temperature > 14.5oC and rain < 250mm, signifying that our data is not informative on which shape the relationships are after these thresholds.

Figure 6: projected changes in arable land

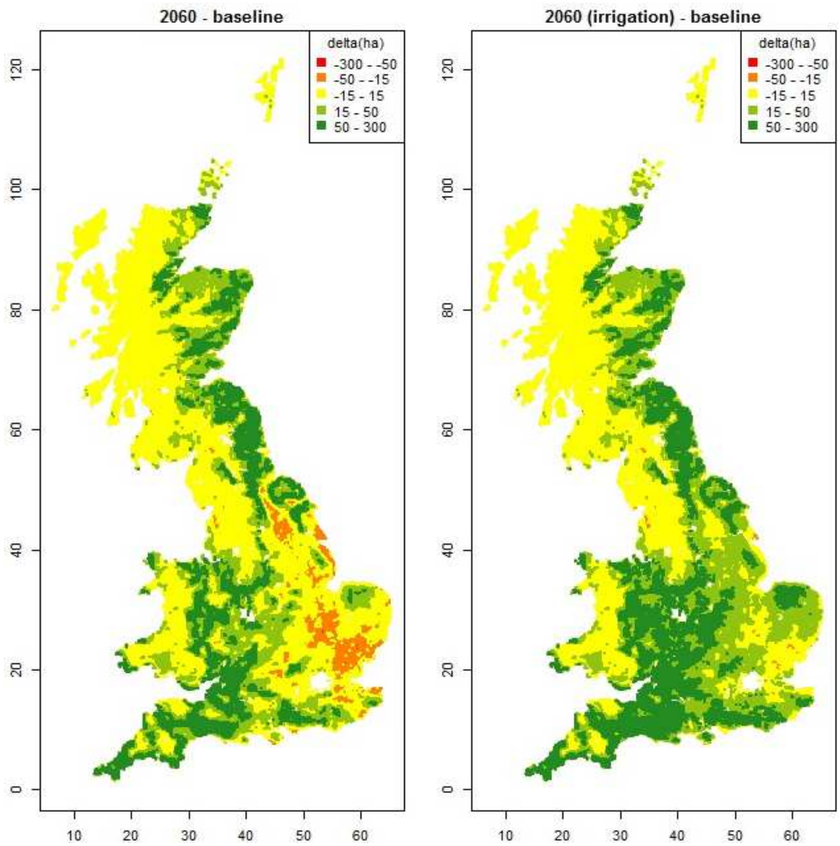


Notes: Changes are calculated as differences with the predictions for the 1980-2010 climate (BLU model).

Bearing in mind these assumptions, Figure 6 presents changes in arable hectares from the baseline climate (1980-2010) to 2020s, 2040s and 2060s. We expect a gradual increase in arable area in the flat regions of Scotland, North of England and South-West of England, where the warmer temperatures will boost crop growth. We observe a similar evolution in the Midlands, but the effect is non-linear: the highest gains are expected in the 2040s, while in the 2060s we start to observe the decreasing curve of an inversed U-shaped relationship. In the East of England, this negative effect is more pronounced, and we expect a small but significant decrease in arable hectares in this area. Such negative effect is mainly caused by the reduction in precipitation, which will increase the risk of drought in a region that is already on the edge of non-irrigated agriculture.

What would happen if farmers introduced irrigation? Figure 7 tries to answer this question. To do so, we assume that irrigation would be able to deliver the optimal amount of water (in our context this is the amount which maximises arable share in the BLU model) to arable land. In other words, we assign precipitation to 280mm for all those cells for which precipitation is lower than 280mm. We do not explicitly model the cost of building an irrigation system, but simply the effect that installing such system would generate on land use. The impact is highly significant. In the areas characterized by a moderate decrease in arable, we now observe a significant increase. This means that by installing irrigation systems the flat areas in the South-East of England could take advantage of the increasing temperatures without suffering drought risk, and become even more productive. Projections for Scotland, North of England and Cornwall are unchanged: in these areas, rainfall will be still far above the irrigation threshold even under climate change.

Figure 7: projected changes in arable land: no irrigation vs. irrigation



Notes: Changes are calculated as differences with the predictions for the 1980-2010 climate (BLU model).

10. Conclusions and caveats

A few caveats need to be taken into account when using this model for projections. First, it may be useful to remind the reader that model scenarios are not predictions of the future. For example, in our climate change scenarios we assumed market prices and government involvement (subsidies, levies, milk quota, etc.) to stay constant. Any other assumption would undermine the clarity of these results and confound the estimate of the impact of climate change adaptation. Of course, since prices and (some) policies are explicitly modelled our framework, more complex scenarios with more than one driver of change can also be derived.

Aside from irrigation, technology also have massive impacts on land use and needs to be taken into account within any robust model of farm decision making. Between 1920 and 1980, for example, cereals yield in the UK have grown more than threefold, mainly driven by technological improvement. Decision makers need to consider all factors when considering the formation of future policy. However, the present paper does not model explicitly such technological advances but simply capture them via a trend. For this reason, for predictions technology should be assumed as constant at baseline levels, i.e. year 2010. In our simulation in Section 7, this allows us to observe the pure, unalloyed effects of climate change.

A further aspect of the approach to potential technical change is that we deliberately do not consider the introduction of possible new crops. This also includes diversions into crops types (such as vineyards) used in warmer countries but not represented in our historical data. For this reason, our climate change projections for the warmest areas of the country (e.g. the South of England) are subject to the highest degree of uncertainty. Since this uncertainty inflates with the extent of climate change, the results for the extreme scenarios (e.g. 2060) for these areas should be interpreted cautiously. Conversely, the results for the North of England, Wales, and Scotland, for which climate change lies within the range of our historic data, are certainly more robust.

Considering our measure of financial impacts, FGM is the most diffused measure of overall farm productivity used in agricultural economics. However, two limitations need to be acknowledged. First, since FGM is defined as the difference between revenues and variable costs, all farm fixed costs (e.g. machineries, buildings, rent, etc.) are not included in the analysis. Secondly, conversion costs are also excluded. In other words, all changes in land use and FGM refer to equilibrium conditions and do not take into account possible costs encountered in order to reach these new equilibriums.

In addition, our climate change simulation focuses on the impact of changes in temperature and precipitation, and does not include other things that might be affected by climate change. For example, increased CO2 fertilization may improve crop yields, albeit recent research indicates that this improvement may not be very significant. Another ignored pressure is climate change permitting the transmigration of new crop pests and diseases. A major difficulty facing the incorporation of these and further effects is the incomplete science base available for such analyses.

Finally, on the technical use of the model, note that to calcualate measures of uncertainty of predictions (e.g. confidence intervals) one needs to take into account the uncertainty in all the model components. A possible approach to address this issue is using Monte Carlo simulation.

11. Variable definitions and data sources

The land use and livestock data are derived from the **June Agricultural Census** (JAC) panel (source, EDINA, www.edina.ac.uk) which, collected on a 2km grid square (400ha) basis, covers the entirety of GB for ten unevenly spaced years from 1972 to 2010. This yields roughly 55,000 grid-square records per year, amounting to more than 500,000 grid-square records for the overall analysis. The only significant agricultural land use category excluded from the farm model is rural woodland, whose expansion and contractions are mainly driven by governmental subsidies which we assume remain constant across our climate change scenarios. As described on the source data website, grid square land use estimates can sometimes overestimate or underestimate the amount of agricultural land within an area, since their collection is based on the location of the main farm house. As in Fezzi and Bateman (2011), we correct this feature by rescaling the sum of the different agricultural land use areas assigned to each grid square to match with the total agricultural land derived using satellite land cover data and ancillary spatial data (Ordnance Survey (OS, www.ordnancesurvey.co.uk/) Meridian Developed Land Use Areas, OS roads, OS railways; the National Inventory for Woodland and Trees) to locate areas that are used for agricultural production, urban activities, etc.

Regarding **Farm Business Survey** (FBS) data, we use years 2005 to 2011. The FBS data includes about 2500 farms, it is updated annually (about 500 farms are rotated each year) and it is published by the Department of Environment, Food and Rural Affairs (Defra). It provides information on the financial characteristics and physical and economic performance of farm businesses throughout England. For each sampled farm, the data available include costs, output, revenues, and margins of farm enterprises; machinery, labor use and other assets; revenues of non-agricultural activities; approximate location (on a 10x10 km grid) and other farm characteristics. Since we focus on arable farms, our dataset includes about 550 farms per year.

Potential environmental drivers of agricultural land use decisions were also calculated for a) a complete 2x2km grid square coverage of GB and b) a 10x10km grid square coverage for England to be linked to the FBS data.

Climate related variables for the growing season (April-September) include average temperature (temp) and accumulated rainfall (rain). These were initially obtained as 5km grid square values, calculated as the average climate between years 1981 and 2010 (http://www.metoffice.gov.uk/ climatechange/science/monitoring/ukcp09/). Sometime we include them in polynomial function and sometimes as linear piecewise function, with for example temp>=10 indicates a variable equal to 0 when temp < 10 and equal to temp – 10 for temp >= 10.

Environmental and **topographic** variables which may influence farmers' decisions include soil characteristics (shares of peat, s_peat, represented also as share of carbon content belonging to the two highest classes, s_carbon4 and s_carbon5, of stoney soil, s_stoney, of gravel s_gravelly, of fragipan soil, s_fragipan, etc.) texture (share of fine s_fine, medium s_medium, coarse s_coarse), ph (ph and its polynomials) We used data from the Harmonised World Soil Database (HWSD): a 30 arc-second (approximately 1km resolution) raster (regular gridded) database with over 16,000 different soil mapping units10. Finally, we include mean altitude (elev) and slope (represented as mean slope, slope, or as % of the cell with slope higher then 6, % slope > 6o, depending on the equation), both derived from the 50m resolution Integrated Hydrological Digital Terrain Model (IHDTM) licensed from the Centre for Ecology and Hydrology (CEH)11.

While for the FBS data, **prices** are derived directly as described in Section 3, for the JAC analysis we included regional-level output prices (e.g. wheat, barley, oilseed rape, etc.) and national level fertilizer price by merging information from information from Defra (2006), the Ministry of Agriculture, Fisheries and Food (MAFF,1986), and the Office of National Statistics.

Regarding the **policy determinants** of land use decisions, we include the share of each grid square designated as National Park, npark, Environmentally Sensitive Area (esa) and Greenbelts (greenbelt). ESAs, introduced in 1987 and extended in subsequent years, were launched to conserve and enhance areas of particular landscape and wildlife significance. Digital boundary data were downloaded from data from Natural England and the Scottish Government. Spatial data for English greenbelt were licensed by Defra from the Ordnance Survey. Presently, there is no national digital spatial boundary dataset for Scottish greenbelt. Each council provided information and PDF maps or ESRI shapefiles. For Wales, there is currently only one area of greenbelt (Newport and Cardiff), and its boundaries were derived from local development plans.

We also control for **transportation costs** by including the distance to the closest major market, defined as an urban centre with more the 300 thousand inhabitants (according to the 2011 Census data): dist300. In the root crop equation, we also include the distance to the closest sugar beet processing factory (sb_dist and piecewise functions). Travel times are calculated from the centroid of a 2km to the nearest urban border. Travel time using motor vehicle is computed via the GB road network (Motorway, A road, B road and minor roads). Urban extent is defined from OS Meridian Developed Land Use Areas. Travel time calculations were performed using a cost distance operation taking into account different driving speed in Geographical Information System (GIS).

References

Albouy D., Graf W., Kellogg R., Wolff H. (2016) Climate amenities, climate change, and American quality of life, Journal of the Association of Environmental and Resource Economists, vol. 3, pp. 205-246.

Bateman et al. (2013) Bringing ecosystem services into economic decision-making: land use in the United Kingdom, Science, vol. 341 (6141), pp. 45-50.

Bateman et al. (2014) Economic value of ecosystem services, Final Report from WP3 to the UK National Ecosystem Assessment – Follow-On programme, UNEP-WCMC, London.

Chambers R.G., Just R.E. (1989) Estimating multioutput technologies, American Journal of Agricultural Economics, vol. 71, pp. 980-995.

Fezzi C. and Bateman I.J. (2011) Structural agricultural land use modeling for spatial agro-environmental policy analysis, American Journal of Agricultural Economics, vol. 93, pp. 1168-1188.

Fezzi C., Bateman I.J. (2015) The impact of climate change on agriculture: nonlinear effects and aggregation bias in Ricardian models of farmland value, Journal of the Association of Environmental and Resource Economists, vol. 2, pp. 57-92.

Fezzi C., Harwood A., Lovett A., Bateman I.J. (2015) The environmental impact of climate change adaptation on land use and water quality, Nature Climate Change, vol. 5, pp. 255–260.

Fezzi C., Hutchins M., Rigby D., Bateman I.J., Posen P., Hadley D. (2010) Integrated assessment of Water Framework Directive nitrate reduction measures, Agricultural Economics, vol. 41, pp. 123-134.

Defra (2006) Agriculture in the United Kingdom 2005 (and following years). Joint report by Defra, the Scottish Executive Environment and Rural Affairs Department, the Department of Agriculture and Rural Development (Northern Ireland), and theWelsh Assembly Government.

Defra (2018) Total income from farming – dataset. Available at: https://www.gov.uk/government/statistics/total-income-from-farming-in-the-uk (last access 09/2018).

Lord, Anthony, Goodlass (2002) Agricultural nitrogen balance and water quality in the UK, Soil Use and Management, vol. 18, pp. 363–369.

Ministry of Agriculture, Fisheries and Food (MAFF, 1986) Agricultural Statistics in the United Kingdom, HMSO, London.

Nakicenovic N., Swart R. (2000) Special Report on Emissions Scenarios, Cambridge University Press, Cambridge.

Nix J. (2009) The farm management pocketbook 2007: 37th edition, The Anderson Centre.

Papke L.E, Wooldridge J.M. (1996) Econometric methods for fractional response variables with an application to 401(k) plan participation rates, Journal of Applied Econometrics, vol. 11, pp. 619-632.

Schlenker W., Roberts, M.J. (2009) Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change, Proceedings of the National Academy of Science, pp. 15594–15598.

Wood S. (2006) Generalized addictive splines: an introduction with R. Chapman & Hall.