Understanding spatial heterogeneity in GB agricultural land-use for improved policy targeting

Patrick Wongsa-art[1,2]   Namhyun Kim[3]   Ian Bateman[4]

Abstract

Today, one of the biggest challenges facing the UK is the new target set when the nation became first major economy to pass net zero emissions law, which requires the country to bring all greenhouse gas emissions to net zero by 2050. On the one hand, there are already a few ideas about how we should farm and use land in order to deliver such a target. On the other hand, the government has a new strategy which is to pay farmers for providing public goods, especially for climate change mitigation through the reduction and storage of greenhouse gas emissions. The most critical task is to find a solution to such a question as *"How should public spending on farm public goods be allocated ?"* In this paper, we argue that formulating an effective subsidy scheme cannot focus on the public need alone, but should also take into consideration what farmers must endure and the opportunities they must forgo. This requires a good understanding about the generating process behind the spatial heterogeneity of agricultural land-use at a fine spatial scale. We aim to provide government and its agents with decision support for policy making post-Brexit in two directions. Firstly, we employ detailed spatial resolution data and establish a new statistical tool that can help: (i) to effectively capture the the spatial heterogeneity of agricultural land-use, (ii) to disentangle the contributions of terrain formulations, environmental characteristics, climatic conditions, policies, and other legacy and agglomeration effects in the generating process of the land-use patterns, and (iii) accurately gauge their relative importance across different regions of GB for more targeted subsidies schemes. Secondly, we employ our new method and provide policy advice and evaluation.

[1]Cardiff Business School, Cardiff University, United Kingdom
[2]Corresponding Author: (Email) wongsa-artp@cardiff.ac.uk
[3]Department of Economics, University of Exeter, United Kingdom
[4]Land, Environment, Economics & Policy Institute, University of Exeter, United Kingdom

1. Introduction

The EU has influenced the UK's environmental policy agenda in a number of ways. Whilst leaving the EU may potentially uncouple the nation from the plethora of flaws surrounding the EU's environmental policies, the government and its agents still faces various challenges. Currently government agencies and statutory bodies, especially Climate Change Committee (CCC), are under enormous pressure to formulate land-use policies that will help to deliver the UK Government's Net Zero greenhouse gas emissions target by 2050. Furthermore, a massive shift in policy away from the EU Common Agricultural Policy's direct payments is approaching in the post-Brexit world. The government's new agriculture bill intends to pay farmers for providing public goods, which may involve water quality improvement, flood risk reduction, air quality improvements, et cetera. It is also important to note that the new system of environmental land management contracts may be voluntary in many cases. Although there are a number of nuances which any subsidy scheme would have to generate to be successful, we feel that Department for Environment, Food and Rural Affairs (Defra) schemes must (i) be able to incentivise farmers to participate, and (ii) be targeted to those locations which deliver the best value for money in terms of the public benefits delivered. In the other words, these schemes must take into consideration both *"what farmers must endure (or opportunities they have to forgo)"* and *"what the public needs"*.

Regarding the latter, Defra have already invested in developing a number of decision support tools (e.g. the ORVal model; see H.M. Government, 2018) and more are in development (e.g. Rose et al., 2016; NEVO, 2018). Nonetheless, formulating such schemes to satisfy the former requires a good understanding about the generating process behind the spatial heterogeneity of agricultural land-use at a fine spatial scale. Such spatial heterogeneity can be linked to the existence of spatial correlation and hence deserves to be interpreted as the "spatial patterns" or more specifically the "land-use patterns". To understand these land-use patterns, we can begin by learning that they are likely generated based on spatial heterogeneity of terrain formulations (e.g. slope and altitude), environmental characteristics (e.g. soil characteristics and textures), climatic conditions, or that which is due to policies, and other legacy and agglomeration effects. Much more information is required however to be successful in formulating effective subsidies schemes and policies in general. This paper aims to provide Defra, government agencies and statutory bodies, e.g. CCC, with decision support for policy making post-Brexit. We do so in two directions, namely methodological development, and policy advice and evaluation.

Regarding the former, we employ detailed spatial resolution data, i.e. a large panel of 2-km^2 grid records, and establish in Section 2 a new statistical tool that is capable of performing the following tasks: (1) Effective capturing the land-use patterns; (2) Disentangling the contributions of terrain formulations, environmental characteristics, climatic conditions, policies, and other legacy and agglomeration effects in the generating process of the land-use patterns; (3) Accurate gauging of their relative importance across different regions of GB for more targeted subsidies schemes. Fezzi et al. (2011), Fezzi et al. (2015) and Bateman et al. (2020) stress an important advantage of using such data, which resides in the ability to capture the heterogeneity in topographic, soil, and climatic characteristics, alongside the corresponding variation in agricultural practices. Nonetheless, their methods involve estimating an empirical model that includes a large number of potential drivers (e.g. soil-textures, soil-characteristics and slope) in order to capture the effects of physical environment alone. On the contrary, our model non-parametrically summarises the effects of spatial heterogeneity of terrain formulations, environmental characteristics, and other legacy and agglomeration effects entirely based on land physical locations. We show that our model is a semi-parametric generalisation of the spatial lag dependence (SLD) model, which is one of the most widely applied models in the spatial econometric literature. Furthermore, by using both the semi- and parametric methods to study selected agricultural land-use shares, which cover up to 88% of the total agricultural land in the UK, in Section 4 we show that the latter are overly restrictive and prone to modeling mis-specification, and has relatively poor model fit.

Moreover, policy analysis is presented in Section 3, while some important findings can be summarised as follows: (1) The majority of land in England's National Parks is farmed and associated positively with the land-use shares studied. These suggest that English National Park Authorities have been quite successful in building up strong relationships with their local farming and commoning communities. (2) Land physical locations (and therefore the underlying effects they represent) have stronger explanatory power on the generating process of land-use patterns than climatic and policy effects. Such a finding has important policy implications. For example, it suggests that the CCC's plan to restore lowland peatland, where it has been artificially drained (e.g. in East Anglia), will be problematic and is likely to meet a great deal of resistance. Moreover, government needs to develop an effective strategy and provide resources to address the historical productivity gap in UK agriculture, which includes skills training, knowledge exchange and delivering R&D at farm level. This way farmers would be less reliant on the natural elements (especially terrain formulations and environmental characteristics) but more flexible to adapt alternative environmentally friendly practices. (3) In GB, the way land is used did not evolve significantly between 1976 and 1981. However, introduction of nitrate vulnerable zones (NVZs) and environmentally sensitive areas (ESAs) leads to substantial changes in the generating process of the land-use patterns. (4) NVZs and ESAs are good examples of schemes that do not try to unnaturally change the way land is used. Hence, they are congruent with the Climate Change Committee's goal for its pragmatic ways of achieving a Net Zero UK, which is to minimise the impact on farmers and their businesses (Climate Change Committee (2020)). Instead, the scheme embraces farmers' ways of using the land and has over time become an important determinant of the data generating process of land-use. (5) Importantly we have also found evidence to suggest that, via the schemes, public subsidies have been orientated towards improvements in public goods (e.g. water quality improvement for the NVZ and conservation and enhancement of biodiversity for ESA) rather than increasing the private production.

2. The proposed model and method

Below we present the proposed model and estimation procedure, then finally discuss the asymptotics.

2.1. The proposed Tobit-SPL model

The model is a combination of two well-known concepts, namely Two-limit Tobit model and semi-parametric partially-linear (SPL) specification, and is therefore referred to hereafter as the Tobit-SPL model. The goal of the former is to take into account possible censoring problem in the land-use shares. The SPL specification allows the most general method of capturing spatial correlation/patterns, while enables parsimonious modelling of the importance of other determinants.

To discuss these in more detail, let $(X_{k,i}, S_i, Y_{k,i})$ denote observations collected from the ith parcel of land and kth crop, where $1 \leq i \leq N$ and $1 \leq k \leq K$. Also, let $X_{k,i} = (X_{k,1,i}, \ldots, X_{k,q,i})^\top \in \mathbb{R}^q$ denote the q-dimensional vector of regressors, $S_i = (S_{1,i}, S_{2,i})^\top \in D$ denote a fixed location by which $D \subset \mathbb{R}^2$ is a spatial domain and

$$Y_{k,i} = \begin{cases} 0 & \text{if } Y^*_{k,i} \leq 0 \\ Y^*_{k,i} & \text{if } 0 < Y^*_{k,i} < 1 \\ 1 & \text{if } Y^*_{k,i} \geq 1 \end{cases} \tag{1}$$

signifies the land-use share, which is censored from below at 0 and above at 1. Otherwise, the land-use share is determined by the partially linear semi-parametric regression model of the form

$$Y^*_{k,i} = X_{k,i}\beta_k + \vartheta_k(S_i) + e_{k,i}, \tag{2}$$

where β_k is a vector of the unknown coefficients, $\vartheta_k(S_i)$ is a smooth function and $e_{k,i}$ denotes the disturbance.

The latent variable specification in (2) is well founded. It can be viewed as a reduced form model of a well-known structural profit-maximisation introduced in Chambers and Just (1989) and extended to the context of the agricultural land-use by e.g. Fezzi and Bateman (2011) and Bateman et al (2020). In this regard, the regressors in the model may include known, well-measured or easily parameterised land-use determinants such as climate conditions and environmental policies. Furthermore, previous studies in the semi-parametric literature have shown that $\vartheta_k(S_i)$ can purely captures the importance of land locations in the d.g.p. of the land-use patterns by excluding their effects on the regressors. Particularly, in the light of Robinson (1988) we can write

$$\vartheta_k(S_i) = \vartheta_{kY}(S_i) - \vartheta_{kX}(S_i)\beta_k, \tag{3}$$

where $\vartheta_{kY}(S_i)$ and $\vartheta_{kX}(S_i)$ are smooth functions that capture the explanatory power of land locations on the land-use shares and regressors, respectively (see e.g. Robinson (1988), Härdle et al. (2000), Li and Racine (2007) and Goa et al. (2012)). Hence, $\vartheta_k(S_i)$ summaries the importance of the spatial heterogeneity of terrain formulations, environmental characteristics, and other legacy and agglomeration effects, which can not be conveniently measured, represented and parameterised in the model.

A useful advantage of the semiparametric model in (2) resides in the fact that it can be considered a semi-parametric generalisation of the SLD model, which is well-known in spatial econometrics and applied extensively to different disciplines (see e.g. Anselin (1988), Anselin and Bera (1998), LeSage and Pace (2009) and Elhorst (2014)). This helps to further confirm the close connection between the two spatial concepts, namely heterogeneity and dependence. To elaborate, let us first introduce some matrix notations. Let $Y^*_{k,N} = (Y^*_{k,1}, \dots, Y^*_{k,N})^\top$, $X_{k,N} = (X^\top_{k,1}, \dots, X^\top_{k,N})^\top$, $e_{k,N} = (e_{k,1}, \dots, e_{k,N})^\top$ and $\vartheta_{k,N} = (\vartheta_k(S_1), \dots, \vartheta_k(S_N))^\top$. In this regard, the semiparametric model in (2) can be expressed as

$$Y^*_{k,N} = X_{k,N}\beta_k + \vartheta_{k,N} + e_{k,N}. \tag{4}$$

Similarly to the construction of (3), here a previous knowledge in the semiparametric regression suggests that we write

$$\vartheta_{k,N} = \vartheta_{kY,N} - \vartheta_{kX,N}\beta_k, \tag{5}$$

where $\vartheta_{kY,N} = (\vartheta_{kY}(S_1), \dots, \vartheta_{kY}(S_N))^\top$ and $\vartheta_{kX,N} = (\vartheta^\top_{kX}(S_1), \dots, \vartheta^\top_{kX}(S_N))^\top$. Then, the Tobit-SDL model follows the following parameterisation

$$\vartheta_{kY,N} = \rho_k W_N Y^*_{k,N} \quad \text{and} \quad \vartheta_{kX,N} = W_N X_{k,N}, \tag{6}$$

where W_N is an $N \times N$ weighting matrix of known constants and ρ_k determines how much land-use shares in the neighboring lands affect that of the i-th land given their distances.

In the other words,

$$Y^*_{k,N} = \rho_k W_N Y^*_{k,N} + \{X_{k,N} - W_N X_N\}\beta_k + e_{k,N}. \tag{7}$$

The close inter-relation between the spatial heterogeneity and spatial dependence is stressed by defining the required weights as a function of the distance between the land grids. Within the context of our discussion, such a distance can be computed based on

$$d^p_{i,j} = \{|S_{1i} - S_{1j}|^p + |S_{2i} - S_{2j}|^p\}^{1/p}. \tag{8}$$

The so-called Euclidean metric and Manhattan metric are obtained by setting $p = 1$ and $p = 2$, respectively. Accordingly, the (i, j)-element of the weighting matrix, W_N, can be computed as

$$\omega_{ij} = \begin{cases} 1/d^p_{ij} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}. \tag{9}$$

2.2. Estimation procedure

Hereafter, let $W_{Nj}(s) = W_{Nj}(s : S_1, \ldots, S_N)$ denote a probability weight function, whose exact definitions will be given later. Then, expression (3) suggests that we compute the unknown function $\vartheta_k(S_i)$ for every given β_k based on

$$\widetilde{\vartheta}_k(s; \beta_k) = \sum_{j=1}^{N} W_{Nj}(s)Y^*_{k,j} - \sum_{j=1}^{N} W_{Nj}(s)X_{k,j}\beta_k. \tag{10}$$

This enables the following approximation of the model in (2)

$$\dot{Y}^*_{k,i} = \dot{X}_{k,i}\beta_k + e_{k,i}, \tag{11}$$
$$\dot{Y}^*_{k,i} = Y^*_{k,i} - a_k(S_i) \quad \text{and} \quad \dot{X}_{k,i} = X_{k,i} - b_k(S_i),$$

where $a_k(S_i) = \sum_{j=1}^{N} W_{Nj}(S_i)Y^*_{k,j}$ and $b_k(S_i) = \sum_{j=1}^{N} W_{Nj}(S_i)X_{k,j}$. In this regard the model in (11) is also a latent variable model in the sense that

$$\dot{Y}_{k,i} = Y_{k,i} - a_k(S_i) = \begin{cases} 0 - a_k(S_i) & \text{if } \ Y^*_{k,i} \leq 0 \\ \dot{Y}^*_{k,i} & \text{if } \ 0 < Y^*_{k,i} < 1, \\ 1 - a_k(S_i) & \text{if } \ Y^*_{k,i} \geq 1 \end{cases} \tag{12}$$

where $0 - a_k(S_i) \leq 0$, $-\, a_k(S_i) < \widetilde{Y}^*_{k,i} < 1 - a_k(S_i)$, and $1 - a_k(S_i) \geq 0$.

In the light of these materials, we suggest estimating the model, which comprises equations (1) and (2), in three steps as follows:

Step A1: Firstly, it is to construct $\dot{X}_{k,i}$ and $\dot{Y}_{k,i}$. This clearly requires computing $a_k(S_i)$ and $b_k(S_i)$. To take into consideration the peculiar feature of the land-use data, we shall follow the suggestion made by Fan and Gijbels (1992) and Cai (2003) and attempt on bias-reduction by defining the probability weight function using the local-linear method as follows

$$W_{Nj}(s) = w_i(s) \Big/ \sum_{i=1}^{N} w_i(s) \tag{13}$$
$$w_i(s) = K_h(S_j - s)\left[1 - \mathscr{S}_1^{\top}\mathscr{S}_2^{-1}(S_j - s)\right],$$

where $\mathscr{S}_1(s) = \sum_{i=1}^{N} K_h(S_i - s)(S_i - s)$, $\mathscr{S}_2(s) = \sum_{i=1}^{N} K_h(S_i - s)(S_i - s)(S_i - s)^{\top}$, $K_h(S_i - s)$ is the product kernel

$$K_h(S_i - s) = \left\{\frac{1}{h_1}k\left(\frac{S_{1,i} - s_1}{h_1}\right)\right\}\left\{\frac{1}{h_2}k\left(\frac{S_{2,i} - s_2}{h_2}\right)\right\},$$

$k(u)$ is a probability kernel function, and h_1 and h_2 are bandwidth parameters. In multiple regressor kernel regression, cross-validation remains a recommended method for bandwidth selection and is therefore used here in association with the leave-one-out residuals.

Step A2: Secondly, it is to estimate the unknown parameters β_k (for $k = 1, \ldots, K$) based on the latent variable model comprising equations (11) and (12). This can be performed based on the standard two-limit Tobit maximum likelihood estimation (see e.g. Fezzi and Bateman (2011), and Bateman et al. (2020) for related land-use applications). Let $\widehat{\beta}_k$ denote the resulting estimates.

Step A3: The third step involves estimating $\vartheta_k(S_i)$ and constructing the the semi-parametric estimates of the land-use patterns based on

$$\widehat{\vartheta}_k(S_i; \widehat{\beta}_k) = a_k(S_i) - b_k(S_i)\widehat{\beta}_k \quad \text{and} \quad X_{k,i}\widehat{\beta}_k + \widehat{\vartheta}_k(S_i; \widehat{\beta}_k), \tag{14}$$

respectively.

2.3. Asymptotics

Assumption 1. *We also assume that the parcel of land i, where $i = 1, \ldots, N$, faces the same relative input/output prices (see e.g. Blundell et al. (2007)).*

Assumption 2. $E[e_{k,i}] = 0$ *and* $0 < var[e_{k,i}] = \sigma^2 < \infty$.

Assumption 3. *(a) Conditions on the kernel functions and more discussion on the asymptotic based on Robinson (1988).*

Asymptotic theory, i.e. convergence rate and the asymptotic distribution, of $\widehat{\vartheta}_k(S_i; \widehat{\beta}_k)$ has been considered extensively in the literature for the case of non-censoring data. In theory, these results are also applicable to our case since $\widehat{\beta}_k$ in Step 2 obtained via the standard two-limit Tobit maximum likelihood estimation also achieves the parametric $\sqrt{N}$-rate, which is faster than that of the nonparametric local linear estimator.

3. Empirical analysis based on the Tobit-PL model

In this section we analyse the agricultural land-use in GB using the above-introduced Tobit-SPL model. To do so, we employ a set of data from a unique database created and maintained by Land, Environment, Economics and Policy (LEEP) Institute, University of Exeter. These include: *Agricultural land-use data.* We focus on four categories of land-use, namely (1) permanent grassland (grassland maintained perpetually without reseeding), (2) temporary grassland (grass typically being part of an arable crop rotation), (3) rough grazing (uncultivated land used for grazing livestock), and (4) arable (cereals (wheat, barley and oats), oilseed rape, root crops (potatoes and sugar beet), other crops and horticulture). These are the main types of land-use in the UK agricultural sector, which usually cover up to 88% of the total agricultural land (Committee on Climate Change (2018)). The data on agricultural land-use are derived from the June Agricultural Census (JAC) on a 2-km^2 (i.e. 400 hectares) grid, which are available from the Edinburgh University Data Library. These data cover seventeen unevenly spaced years between 1976 and 2010. However, there are only four of these years, namely 1976, 1981, 2003 and 2004, in which the data cover the entire GB, i.e. England, Scotland and Wales. Our analysis concentrates on 1976, 1981 and 2004. *Climatic data.* Climate related variables are the average temperature and accumulated rainfall associated with a growing season, namely April to September. These data were initially obtained on a 5-km^2 grid from the Met Office, which are calculated as the average climate between years 1981 and 2010, and then interpolated to the 2-km^2 grid. *Policy determinants of land use decisions.* These include shares of each grid square designated as national park (NPK), greenbelt (GBT), environmentally sensitive areas (ESAs) and nitrate vulnerable zone (NVZs). *Transportation Cost.* Transportation cost is represented by the the distance to the closest major market (defined as an urban centre with more the 300 thousands inhabitants according to the 2011 Census data).

While the full dataset consists of data from up to 55,000 parcels of land each year, the JAC based land-use data can be quite noisy. For instance, since grid-square land-use estimates are derived from parish summaries, when a farm's agricultural land belongs to more than one parish all the land-use is assigned to the one in which the main farm house is registered. Furthermore, Fezzi et al. (2011) observe that the actual JAC data appear rather "blocky" with very high and very low values appearing immediately adjacent to each other. To alleviate this problem, we select a subset of the data by randomly extracting one grid square and then sampling every fourth grid cell along both latitude and longitude axes (see also Carrión-Flores and Irwin (2004), Nelson and Hellerstein (1997) for a similar use of the spatial sampling method). By doing so, there remains a large number of observations in each year and such a data selection should not affect the models' asymptotic results. In addition, to ensure that such a selection does not distort our estimation results we have examined empirical distribution functions of the resulting data and found that they can capture all important features of the original observations (see, for example, Figure 15 in the supplemental document). Also, an important question remains whether the censored regression specification is necessary. To answer this question, we examine two types of the Kaplan-Meier estimates of the cumulative distribution function (CDF) of the four selected land-use shares for 1976, 1981 and 2004, which take into consideration the data censoring problem, and non-censoring estimates. In this regard, we have found that the non-censoring estimates are distorted since they are almost always statistically significantly different from the censoring counterparts.

Finally, we finish this section by discussing a set of descriptive statistics summarised by boxplots of the land-use data in Figure 1. These are particularly useful since the median should be a more accurate measure of the central tendency than the mean when there exists the data censoring problem. The boxplots show that medians of the shares of temporary grassland decline significantly between 1976 and 2004, while those for arable and permanent grassland increase slightly during the period. While the former is due to declining reseeding rates and the reclassification of grass over 5 years to permanent, the latter is brought about by improved commodity prices. On the other hand, shares of rough grazing do not evolve very much and express some distributional stability.

3.1. Estimation results and policy implications

The objective of our work is threefold: (1) To empirically capture the land-use patterns. (2) To disentangle the contributions of terrain formulations, environmental characteristics, climatic conditions, policies, and other legacy and agglomeration effects in the d.g.p. of the land-use patterns. (3) To gauge their relative importance across different regions of GB for more targeted subsidies schemes. Below we satisfy these objectives in turn.

3.1.1. Objectives 1 and 2:

Let us concentrate first on the estimation results in Figures 3 to 14. The top-left panels of these figures display the actual observed land-use shares for each of the selected categories for 2004, 1981 and 1976, while the top-middle panels present the semi-parametric estimates of land-use patterns (see Step 3 in Section 2.2 for details). Permanent grassland dominates large areas of Wales and the mountainous regions NW of England. In these regions, slope, wetness, and often a heterogeneous pattern of conditions render even occasional cultivation unsuitable. Meanwhile, in the Scottish northwest highlands, steep, poorly drained, acid or shallow soils combined with wet, cool or cold climates zones render land capable of

Table 1: Regressors included in Models 2, 1 and 0†

	Abbreviations	Definitions
Model 2	*npk*	Share of each grid square designated as NPK
	esa	Share of each grid square designated as ESA
	gbt	Share of each grid square designated as GBT
	nvz	Share of each grid square designated as NVZ
Model 1	*rain*	Accumulated rainfall for growing season
	temp	Average temperature for growing season
	ratemp	$rain \times temp$, i.e., an interaction term
	dist	Distance to the closet major market
Model 0	*intc*	Intercept term

† *sud*, *nor* and *mid* are summed to one, so *mid* is omitted in the estimation. Similarly, *scoarse*, *fine* and *smedium* are summed to one and so *smedium* is omitted.

supporting only rough grazing. On the contrary, arable dominates prime agricultural land in the SE regions of England. Furthermore, the figures show that the estimated land-use patterns are able to capture all of the important features in the row data. The residuals plots, which are displayed in the top-left panels of the figures, show some observations with large magnitude. These should not be cause for concern, however, since they are most likely caused by a limitation of the raw JAC data, in which very high and very low values appearing immediately adjacent to each other due to the allocation of parish-level records to grid squares.

The bottom-left and bottom-middle panels present accumulated climatic and policy effects, respectively. Not surprisingly, wet, cool and cold climates zones are associated positively with shares of rough grazing, while the opposite is true for other categories of land-use. Since the majority of land in England's National Parks is farmed, it is also not surprising to see that NPK and GBT are associated positively with some categories of the land-use shares, i.e. arable and rough grazing. The English National Park Authorities have been quite successful in building up strong relationships with their local farming and commoning communities. Furthermore, the resulting positive (negative) association between NVZ with temporary grassland and arable (rough grazing) is reasonable given the substantial increase in fertilizer price during 2002 to 2006 (Defra (2006)). On the other hand, ESA payments seem to have encouraged significant switching from arable land to

Table 2: LR Statistics 2004

2004	Permanent G	Temporary G	Rough G	Arable
LR_{10}	175.048 (0.000)	57.543 (0.000)	61.378 (0.000)	1549.382 (0.000)
LR_{21}	1594.688 (0.000)	1209.462 (0.000)	1275.897 (0.000)	2362.658 (0.000)
LR_{32}	2526.835 (0.000)	2498.674 (0.000)	5193.822 (0.000)	3384.680 (0.000)
Log-likelihood	-10331.598	-7472.921	-7890.289	-10199.087

rough grazing. Their negative association with permanent and temporary grasslands is consistent with livestock intensity prediction reported in e.g. Fezzi et al. (2011).

The bottom-right panel displays the non-parametric estimates of $\vartheta_k(S_i)$. Based on (3), we have shown that these estimates encapsulate the importance of land locations in the generating process of the land-use patterns. In the other word, they summarise the importance of terrain formulations, environmental characteristics, and other legacy and agglomeration effects, which can not be conveniently measured and parameterised in the model. From the figures, it is evident that geography and the quality of the farmland does have a significant impact. This suggests that plans to switch land-use (e.g. restoring peatland in SE England or increasing woodland in Wales) will be costly, problematic and are likely to meet a great deal of resistance. Nonetheless, it is intriguing to see that climatic effects seem to play much more critical role in the d.g.p. of the rough grazing shares. Hence, Scotland's ambition to restore 250,000 hectares of degraded peat by 2030 is quite possible and can coexist with sheep farming. Moreover, the results in these figures show that the land-use patterns do not alter substantially between 1976 to 2004. We shall examine some of these claims more thoroughly below.

3.1.2. Objective 3:

The above-mentioned censoring issue renders formal examinations of the above claims difficult. Nonetheless, a method that can be empirically adapted to our queries is the well-known Likelihood Ratio (LR) test procedure. To apply such a procedure requires estimating the following Tobit models:

Model 3 $\quad Y^*_{k,i} = X_{k,i}\beta_k + \alpha_{k,1}\{\vartheta_{kY}(S_i)\} + \alpha_{k,2}\{-\vartheta_{kX}(S_i)\beta_k\} + e_{k,i}$

Model 2 $\quad Y^*_{k,i} = X_{k,i}\beta_k + e_{k,i}$

Model 1 $\quad Y^*_{k,i} = X^{(P)}_{k,i}\beta^{(P)}_k + e_{k,i},$ where $X^{(P)}_{k,i}$ is a subset concerning policies

Model 0 $\quad Y^*_{k,i} = \beta_{k,0} + e_{k,i}$

Model 3 can be viewed as an unrestricted model, which is brought about by substituting expression (3) into (2). Below we shall also check that the resulting estimates of $\alpha_{k,1}$

Table 3: LR Statistics 1981

1981	Permanent G	Temporary G	Rough G	Arable
LR_{10}	42.954 (0.000)	53.312 (0.000)	59.539 (0.000)	167.895 (0.000)
LR_{21}	1548.651 (0.000)	1018.210 (0.000)	1321.858 (0.000)	1917.196 (0.000)
LR_{32}	2800.450 (0.000)	2304.857 (0.000)	4919.573 (0.000)	4657.913 (0.000)
Log-likelihood	-10818.266	-8463.431	-9769.720	11375.714

and $\alpha_{k,2}$ are not statistically different from one. Model 2 is obtained by restricting $\alpha_{k,1} = \alpha_{k,2} = 0$ and contains a full set of regressors listed in Table 1, while Model 1 imposes further restrictions that coefficients associated with the climatic determinants are all equal to zero. Finally, Model 0 is the most restrictive since it only involves the intercept term.

To evaluate the difference between these nested models, we compute a set of LR statistics, namely $LR_{32} = 2(loglik(M_3) - loglik(M_2))$, $LR_{21} = 2(loglik(M_2) - loglik(M_1))$ and $LR_{10} = 2(loglik(M_1) - loglik(M_0))$, where $loglik(M_j)$ are natural log of the models final likelihood (i.e. the log likelihood) for $j = 1, 2, 3$. In this regard, LR_{32} expresses the explanatory power of the spatial heterogeneity of terrain formulations, environmental characteristics, and other legacy and agglomeration effects, while LR_{21} and LR_{10} quantifies those of climatic conditions and environmental policies, respectively.

The results are presented in Tables 2 to 4. Although not reported here to keep the tables simple, it should be noted that the resulting estimates for $\alpha_{k,1}$ and $\alpha_{k,2}$ are found to be close to and not statistically different from one in all cases. In addition, the associated p-values in the parentheses are virtually zero, which suggest that differences in the log likelihoods of the pair of models are statistically significant. More importantly, LR_{32} are found to be larger than LR_{21} and LR_{01} in all cases. These suggest that land physical locations (and therefore the underlying effects they represent) have stronger explanatory power on the d.g.p. of land-use patterns than climatic and policy effects. This finding has a number of important policy implications. For example, it suggests that the Committee on Climate Change (2020) pragmatic plan to restore peatland in the uplands can continue to coexist with sheep farming, which is already heavily reliant on subsidies. Nonetheless, restoring lowland peatland, for example in East Anglia, where it has been artificially drained, will be problematic and is likely to meet a great deal of resistance as it is on prime productive croplands. Moreover, government needs to develop an effective strategy and provide resources to address the historical productivity gap in UK agriculture including: skills, training and knowledge exchange; and delivering R&D at farm level. This way farmers would be less reliant on natural elements (e.g. terrain formulations and

Table 4: LR Statistics 1976

1976	Permanent G	Temporary G	Rough G	Arable
LR_{10}	37.723 (0.000)	58.427 (0.000)	61.014 (0.000)	182.614 (0.000)
LR_{21}	1648.632 (0.000)	1132.100 (0.000)	1966.041 (0.000)	1797.593 (0.000)
LR_{32}	3016.721 (0.000)	2371.397 (0.000)	5390.241 (0.000)	4948.571 (0.000)
Log-likelihood	-11187.773	-8447.931	-11005.132	-11692.999

environmental characteristics) but more flexible to attempt more environmentally friendly practices.

From the table, we can also confirm that agricultural land-use in GB did not evolve very much between 1976 and 1981. Nonetheless, the introduction of NVZ and ESA leads to considerable changes to the d.g.p. of the land-use patterns. The most significant change occurs to that of the arable shares. In this regard, even though spatial heterogeneity of terrain formulations, environmental characteristics, and other legacy and agglomeration effects still have the highest explanatory power, environmental policies do play a much bigger role compared to in 1976 and 1981. The same can also be said about grasslands, but with much smaller extent. An important feature of these policies resides in the fact that they do not attempt to directly interfere with (i) the levels of the land-use shares and (ii) land-use patterns. To see this let us recall two important points made earlier, namely (a) boxplots of at least three land-use categories in Figures 1 look relatively stable between 1976 to 2004 (the change to rough grazing shares was due to declining reseeding rates and the reclassification of grass over 5 years to permanent), and (b) the land-use patterns presented in the top-middle panels of Figures 3 to 14 do not evolve very much between 1976 and 2004. In the other words, NVZ and ESA are good examples of schemes that do not try to unnaturally change the way land is used (which could be costly and therefore lead to opting out). Instead, the scheme embraces farmers' ways of using the land and has over time become an important determinant of the d.g.p. of land-use. Importantly this evidence also suggests that, via the schemes, public subsidies have been orientated towards improvements in public goods (e.g. water quality improvement for the NVZ and conservation and enhancement of biodiversity for ESA) rather than increasing the private production.

Table 5: LR Statistics for Tobit-TL based on 2004

Tobit-TL	Permanent G	Temporary G	Rough G	Arable
LR_{10}	175.048 (0.000)	57.543 (0.000)	61.378 (0.000)	1549.382 (0.000)
LR_{21}	1594.688 (0.000)	1209.462 (0.000)	1275.897 (0.000)	2362.658 (0.000)
LR_{32} (Tobit-TL)	876.876 (0.000)	879.190 (0.000)	961.1344 (0.000)	1210.915 (0.000)
Log-likelihood	-11469.592	-8491.655	-10006.633	-11445.148
LR_{32} (Tobit-SLD)	2357.738 (0.000)	1763.989 (0.000)	1508.925 (0.000)	4644.482 (0.000)
Log-likelihood	-10729.168	-8049.256	-9728.908	-10943.744

4. Model comparison

This section aims to examine whether the Tobit-SPL model is able to provide improvement in term of model fit beyond the parametric models. For the sake of comparison, we shall study two alternative parametric methods, namely (i) Fezzi et al. (2015) and Bateman et al. (2020)) approach, whose two-limit Tobit (Tobit-TL hereafter) models contains a large number of regressors to help to parametrically capture the heterogeneity in topographic, soil, and climatic characteristics, and (ii) the Tobit-SLD model. Our treatment for each of the models follows closely the procedure introduced in Section 3.1.2. In particular, we begin by retaining the above-discussed Models 0, 1 and 2, formulate the corresponding unrestricted versions of the models, and compute a set of LR Statistics in a similar fashion to those in Table 2. To this end, model (7) can be viewed as the unrestricted latent variable model for the Tobit-SLD. To specify the unrestricted Tobit-TL model, we employ the latent-variable model in Model 2 as a starting point, then introduce additional regressors in order to take into consideration possible non-linearity, terrain formulations and other environmental characteristics, i.e. in the spirit of Fezzi et al. (2015). Table 6 presents a list of these additional regressors and their definitions.

The estimation results are presented in Table 5. It is not surprising to see that the first two rows are exactly the same as those in Table 2.

- Regarding the Tobit-TL model, the results in the third row suggest a much smaller explanatory power incurred by introducing the additional regressors listed in Table 6 into the model compared to LR_{32} in Table 2. We argue that this might have been caused either by (i) modeling mis-specification due to a possible exclusion of relevant regressors, or (ii) overly restrictive parametric model specification. Such an argument is supported by the log-likelihood scores reported in the forth row, which are much smaller than those reported in in Table 2. In the other words, the

Table 6: Additional regressors for Model 3†

Abbreviations	Definitions
$alt0$	$d_{eb200} \times elev$, where $d_{eb200} = 1$ if $elev < 200$ and 0 otherwise
$alt200$	$d_{ea200} \times elev$, where $d_{ea200} = 1$ if $elev > 200$ and 0 otherwise
$alt200d$	$alt200d = 1$ if $elev > 200$ and 0 otherwise
$slope6$	Share of each grid square with slope higher than 6 degrees
$speat$	Proportion of soil characteristic "Peat"
$sgravel$	Proportion of soil characteristic "Gravel"
$sstoney$	Proportion of soil characteristic "Stone"
$sfragipan$	Proportion of soil characteristic "Fragipan Soil"
$scoarse$	Proportion of soil texture "Coarse"
$sfine$	Proportion of soil texture "Fine"
$smedium$	Proportion of soil texture "Medium"
sud	$sud = 1$, if the grid square is located in the Sorthern England
nor	$nor = 1$, if the grid square is located in the Northern England
mid	$mid = 1$, if the grid square is located in the Midlands
$rain_\ell$	$rain_\ell = (rain - \ell)d_{r\ell}$ for $\ell = 300, 350, 400, 450, 500, 600^\ddagger$
$temp_\ell$	$temp_\ell = (temp - \ell)d_{t\ell}$ for $\ell = 9, 10, 11, 12, 13, 14$

‡ $d_{r\ell} = 1$ if $rain < \ell$ and 0 otherwise.

Tobit-SPL model has a much better model fit compared to the Tobit-TL model.

- With respect to the Tobit-SLD model, the test statistics reported in the fifth row can be used for testing the null hypotheses that $\rho_k = 0$, which are clearly rejected in all cases since the reported p-values are virtually zero. In spite of the fact that LR_{32} (Tobit-SLD) in some cases are quite close to those reported in Table 2, such a parametric model might still be overly restrictive. The log-likelihood scores reported in the sixth row are much smaller than those reported in in Table 2. In the other words, the Tobit-SPL model also has a much better model fit compared to the Tobit-SLD model.

- Finally, the results in the table suggest that the Tobit-SLD model has a much better model fit compared to the Tobit-TL model.

5. Conclusions

This paper aimed to provide Defra, government agencies and other statutory bodies with decision support for policy making post-Brexit. Regarding the government plan to pay farmers for public goods they provide, we have identified two important aspects, namely "what farmers must endure (or opportunities they have to forgo)" and "what the public needs". While a large amount of resources have concentrated on the latter, the former seems to be somewhat neglected. To address such a shortfall, this paper developed a new statistical tool that can help to better understand the generating process behind the spatial heterogeneity of agricultural land-use at a fine spatial scale. We analysed agricultural land-use in GB for 1976, 1981 and 2004 and found that land physical locations (and therefore the underlying effects they represent) have stronger explanatory power on the generating process of land-use patterns than climatic and policy effects. Hence, government needs to develop an effective strategy and provide resources to address the historical productivity gap in UK agriculture so that farmers would be less reliant on the natural elements (especially terrain formulations and environmental characteristics) but more flexible to adapt alternative environmentally friendly practices. Moreover, introduction of nitrate vulnerable zones (NVZs) and environmentally sensitive areas (ESAs) leads to substantial changes in the generating process of the land-use patterns. NVZs and ESAs are good examples of schemes that do not try to unnaturally change the way land is used, which are congruent with the CCC's goal for its pragmatic ways of achieving a Net Zero UK. Importantly we have also found evidence to suggest that, via the schemes, public subsidies have been orientated towards improvements in public goods (e.g. water quality improvement for the NVZ and conservation and enhancement of biodiversity for ESA) rather than increasing the private production.

6. References

Amemiya, T. (1973). Regression analysis when the dependent variable is truncated normal. Econometrica: Journal of the Econometric Society, 997-1016.

Anselin, L. (1988). Spatial econometrics: Methods and models. Kluwer Academic: Boston, MA.

Anselin, L. & Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics: Regression models. In Handbook of applied economic statistics (pp. 257-259). CRC Press: Boca Raton, FL.

Ay, J. S., Chakir, R., & Le Gallo, J. (2017). Aggregated versus individual land-use models: Modeling spatial autocorrelation to increase predictive accuracy. Environmental Modeling & Assessment, 22(2), 129-145.

Baltagi, B. H. (1980). On seemingly unrelated regressions with error components. Econometrica: Journal of the Econometric Society, 1547-1551.

Baltagi, B. H. (2008). Econometric analysis of panel data. John Wiley & Sons.

Baltagi, B. H., Bresson, G., & Pirotte, A. (2012). Forecasting with spatial panel data. Computational Statistics & Data Analysis. 56(11):3381-97.

Baltagi, B. H., & Li, D. (2004). Prediction in the panel data model with spatial correlation. In Advances in spatial econometrics (pp. 283-295). Springer: Berlin, Heidelberg.

Baltagi, B. H., & Li, D. (2006). Prediction in the panel data model with spatial correlation: the case of liquor. Spatial Economic Analysis, 1(2), 175-185.

Baltagi, B. H., & Pirotte, A. (2011). Seemingly unrelated regressions with spatial error components. Empirical Economics, 40(1), 5-49.

Bhattacharjee, A., & Jensen-Butler, C. (2006). Estimation of spatial weights matrix, with an application to diffusion in housing demand. Centre for Research into Industry, Enterprise, Finance, and the Firm Discussion Paper, 519.

Breusch, T. S., & Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. The Review of Economic Studies, 47(1), 239-253.

Cameron, A. C., & Trivedi, P. K. (2005). Microeconometrics: methods and applications. Cambridge University Press.

Carrion-Flores, C., & Irwin, E. G. (2004). Determinants of residential land-use conversion and sprawl at the rural-urban fringe. American Journal of Agricultural Economics, 86(4), 889-904.

Chambers, R. G., & Just, R. E. (1989). Estimating multioutput technologies. American Journal of Agricultural Economics, 71(4), 980-995.

Chakir, R., & Gallo, J.L. (2013). Predicting land use allocation in France: A spatial panel data analysis. Ecological Economics, 92(0), 114-125.

Cliff, A. D. (1973). Spatial autocorrelation (No. 04; QA278. 2, C5.).

Cliff, A. D., & Ord, J. K. (1981). Spatial processes: models & applications. Taylor & Francis.

Committee on Climate Change (2018). Land use: Reducing emissions and preparing for climate change. Retrieved from https://www.theccc.org.uk/publication/land-use-reducing-emissions-and-preparing-for-climate-change.

Committee on Climate Change (2020). Land use: Policies for a Net Zero UK.
Retrieved from https://www.theccc.org.uk/publication/land-use-policies-for-a-net-zero-uk/.

Dong, D., Gould, B. W., & Kaiser, H. M. (2004). Food demand in Mexico: an application of the Amemiya-Tobin approach to the estimation of a censored food system. American Journal of Agricultural Economics, 86(4), 1094-1107.

Fezzi, C., & Bateman, I. J. (2011). Structural agricultural land use modeling for spatial agro-environmental policy analysis. American Journal of Agricultural Economics, 93(4), 1168-1188.

Greene, W. H. (2008). The econometric approach to efficiency analysis. The measurement of productive efficiency and productivity growth, 1(1), 92-250.

Kapoor, M., Kelejian, H. H., & Prucha, I. R. (2007). Panel data models with spatially correlated error components. Journal of Econometrics, 140(1), 97-130.

Li, M., JunJie, W., & Deng, X. (2013). Identifying drivers of land use change in China: A spatial multinomial logit model analysis. Land Economics, 89(4), 632-654.

Liu, S. F., & Yang, Z. (2015). Asymptotic distribution and finite sample bias correction of QML estimators for spatial error dependence model. Econometrics, 3(2), 376-411.

Kelejian, H. H., & Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review, 40(2), 509-533.

Marcos-Martinez, R., Bryan, B. A., Connor, J. D., & King, D. (2017). Agricultural land-use dynamics: Assessing the relative importance of socioeconomic and biophysical drivers for more targeted policy. Land Use Policy, 63, 53-66.

Mattison, E. H., & Norris, K. (2005). Bridging the gaps between agricultural policy, land-use and biodiversity. Trends in Ecology & Evolution. 20(11):610-6.

Meyerhoefer, C. D., Ranney, C. K. & Sahn, D. E. (2005). Consistent estimation of censored demand systems using panel data. American Journal of Agricultural Economics. 87(3), pp.660-672.

Moscone, F., Knapp, M., & Tosetti E. (2007). Mental health expenditure in England: a spatial panel approach. Journal of Health Economics. 26(4):842-64.

Nelson, G. C., & Hellerstein, D. (1997). Do roads cause deforestation? Using satellite images in econometric analysis of land use. American Journal of Agricultural Economics, 79(1), 80-88.

Lacroix, A., & Thomas, A. (2011). Estimating the environmental impact of land and production decisions with multivariate selection rules and panel data. American Journal of Agricultural Economics, 93(3), 784-802.

Reidsma, P., Tekelenburg, T., Van den Berg M., & Alkemade R. (2006) Impacts of land-use change on biodiversity: An assessment of agricultural biodiversity in the European Union. Agriculture, ecosystems & environment. 114(1):86-102.

Sterling, S. M., Ducharne, A., & Polcher, J. (2013). The impact of global land-cover change on the terrestrial water cycle. Nature Climate Change, 3(4), 385-390.

Tsay, R. S. (2005). Analysis of financial time series (Vol. 543). John Wiley & Sons.

Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. Journal of Econometrics, 68(1), 115-132.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.

Yang, Z. (2013). Quasi-maximum likelihood estimation for spatial panel data regressions.

Yen, S. T., Lin, B. H., & Smallwood, D. M. (2003). Quasi-and simulated-likelihood approaches to censored demand systems: food consumption by food stamp recipients in the United States. American Journal of Agricultural Economics, 85(2), 458-478.

7. Supplemental document

Figures 16 to 18 present the Kaplan-Meier estimates of the cumulative distribution function (CDF) of the four selected land-use shares for 1976, 1981 and 2004, respectively. For the sake of comparison, we present two types of these estimates. The first type is referred to as censoring estimates, which take into consideration the data censoring problem and are presented in association with the 95% confidence bounds. The second type comprises the non-censoring estimates. In this regard, the non-censoring estimates in the green colour are clearly distorted since they are almost always outside of the 95% confidence bounds.

List of Figures

Figure 1: Boxplots

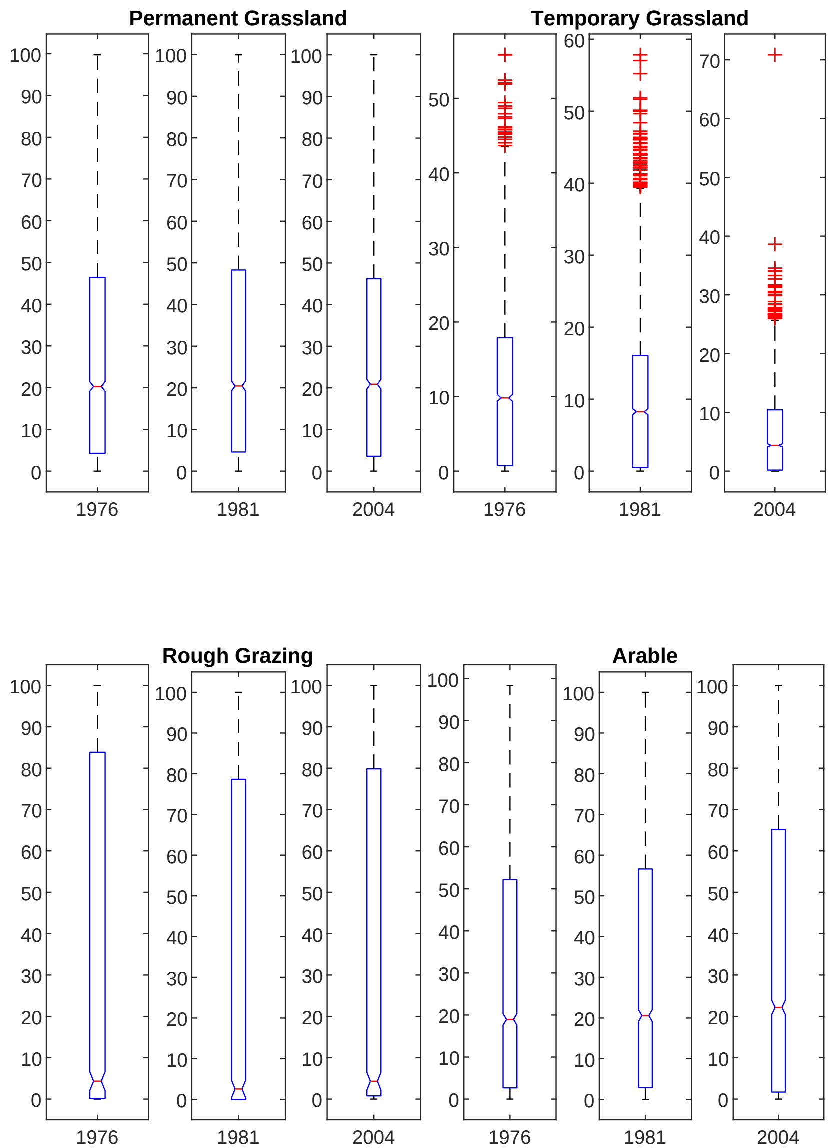

Figure 2: Policies 2004

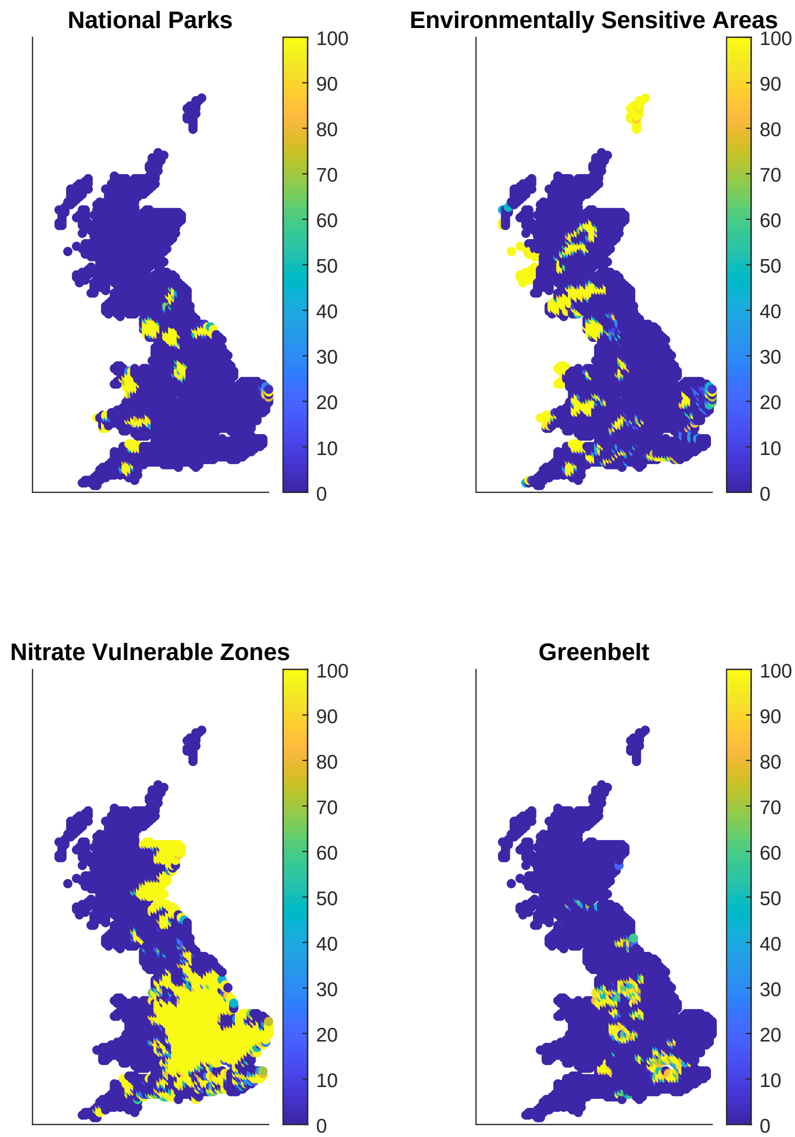

Figure 3: Permanent Grassland 2004

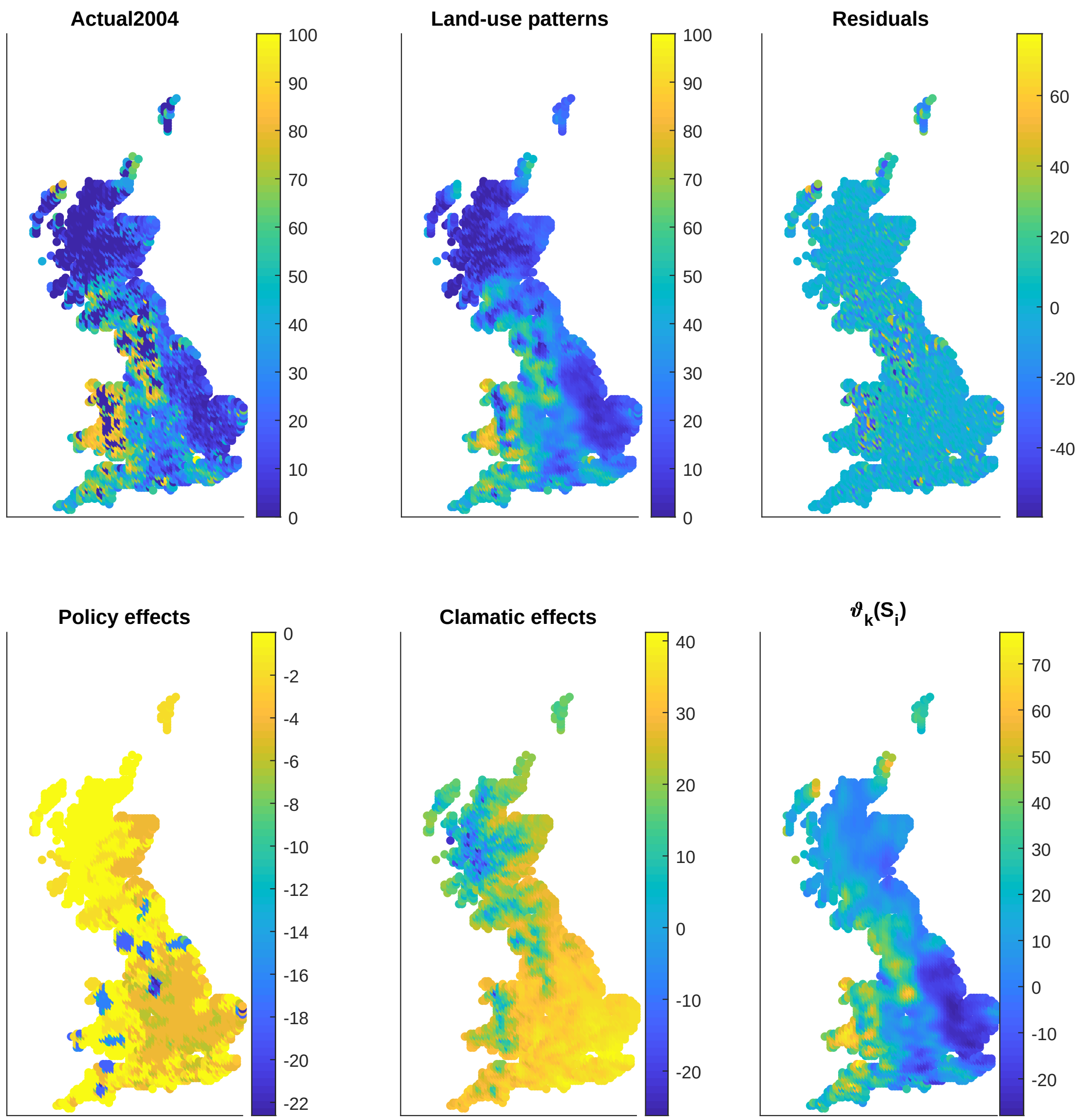

Figure 4: Permanent Grassland 1981

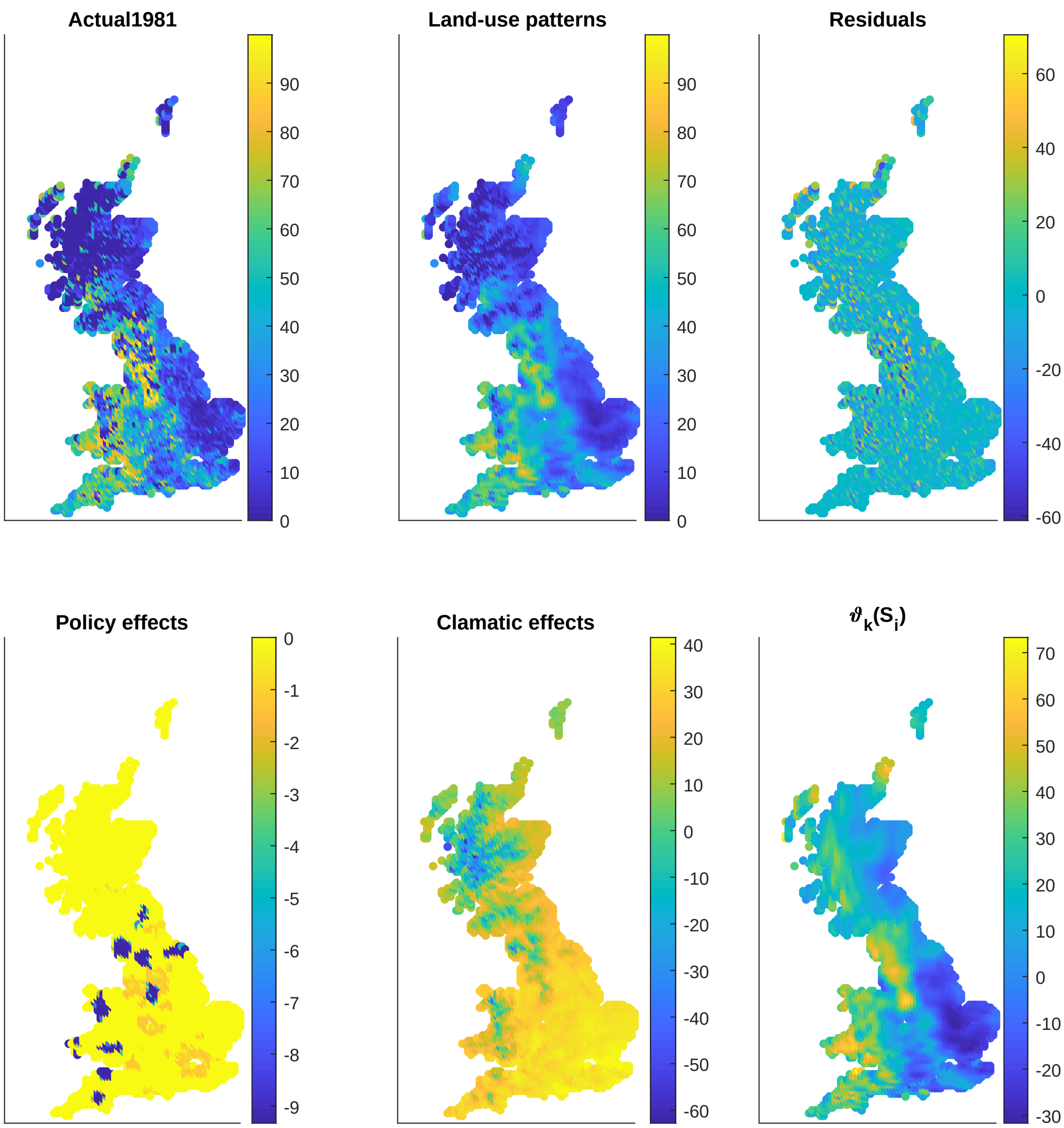

Figure 5: Permanent Grassland 1976

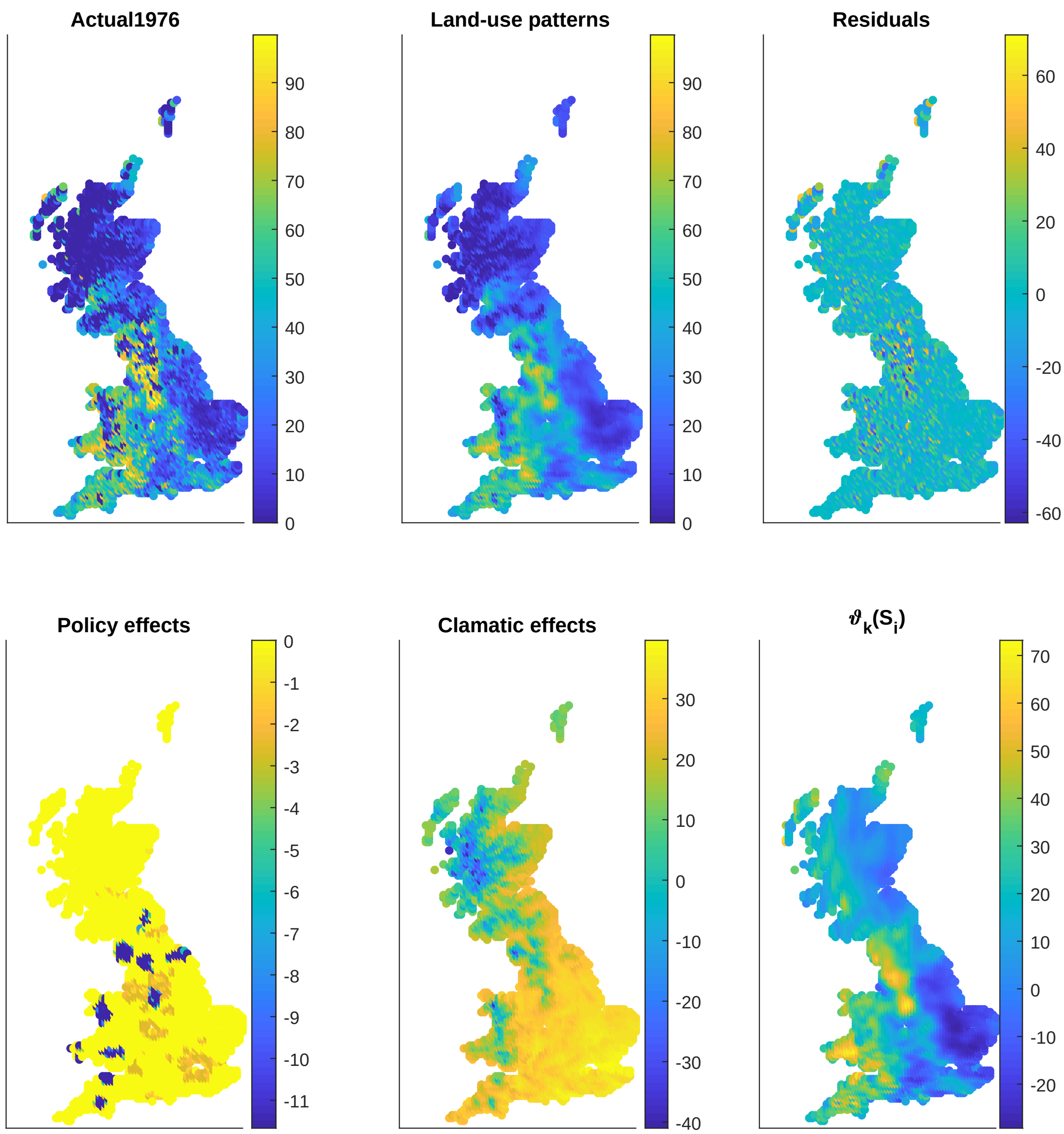

Figure 6: Temporary Grassland 2004

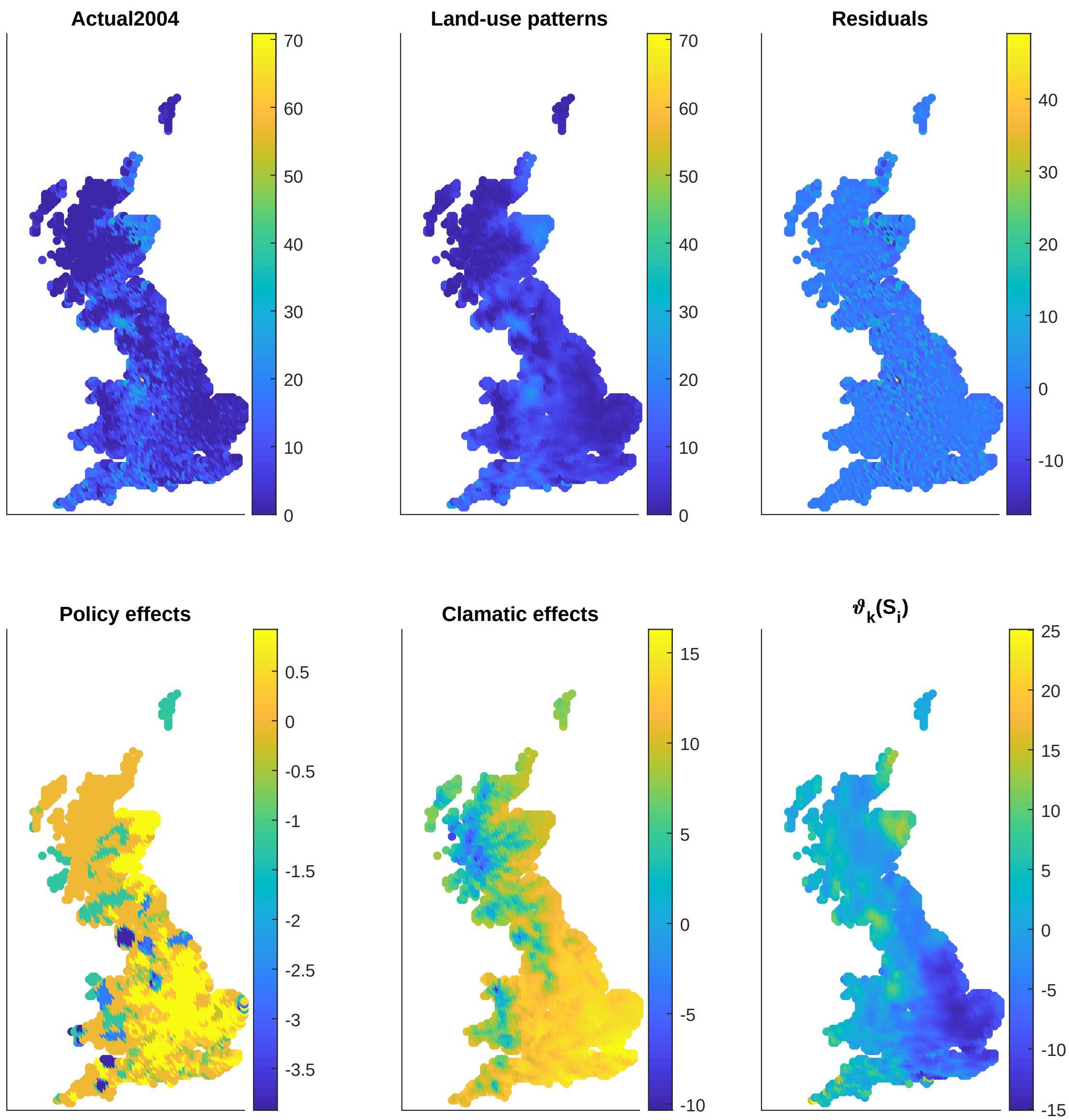

Figure 7: Temporary Grassland 1981

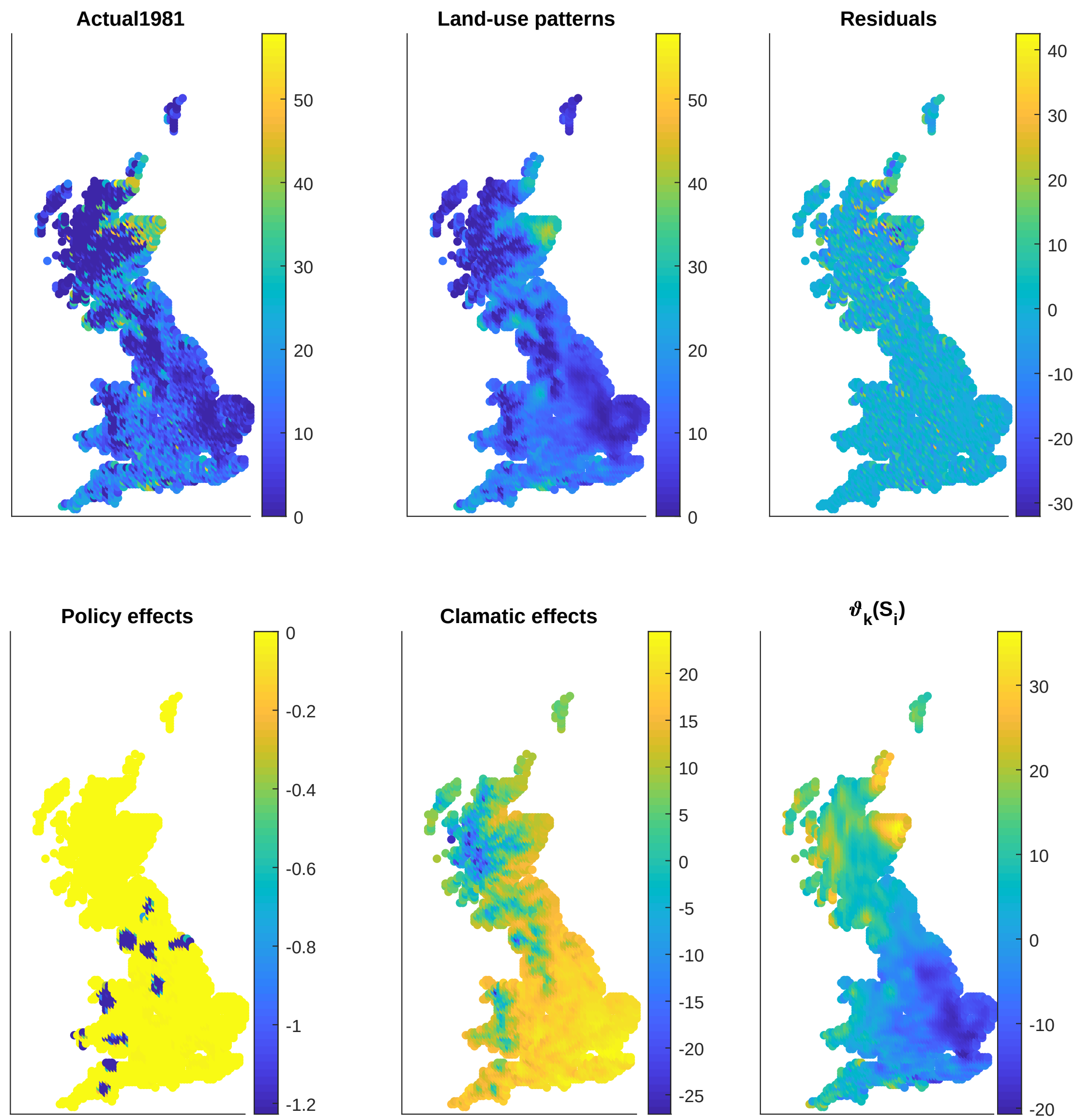

Figure 8: Temporary Grassland 1976

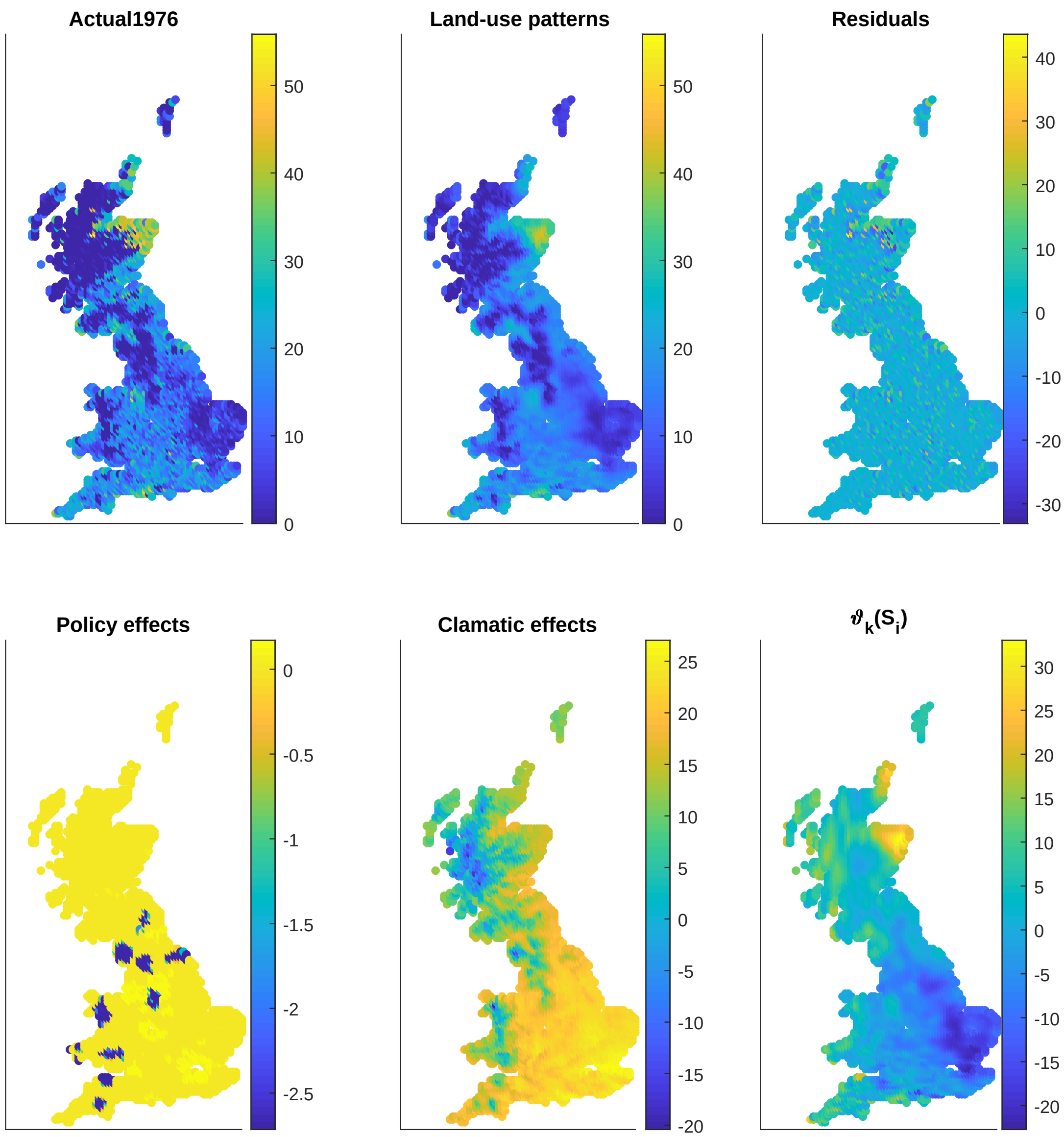

Figure 9: Rough Grazing 2004

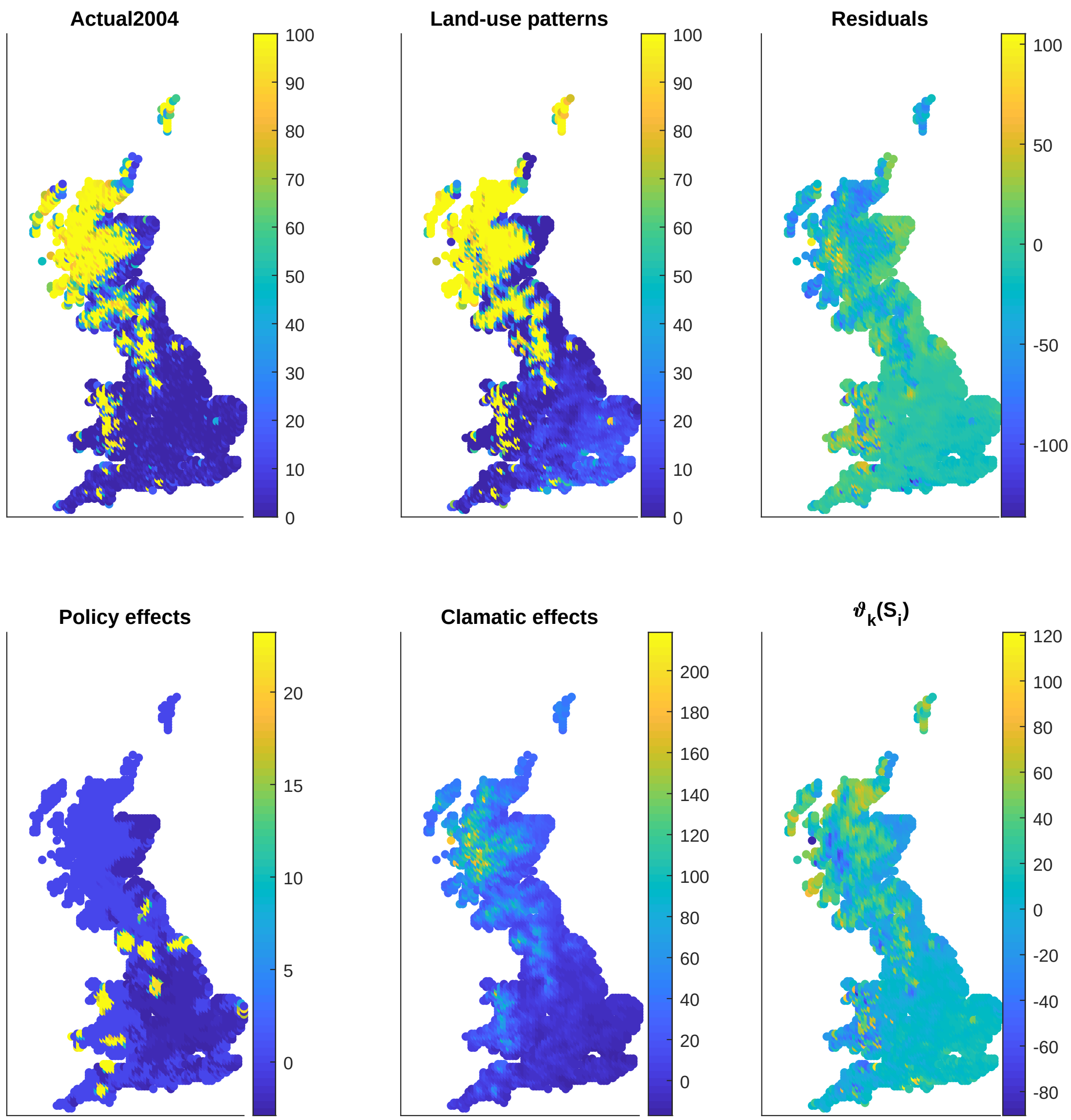

Figure 10: Rough Grazing 1981

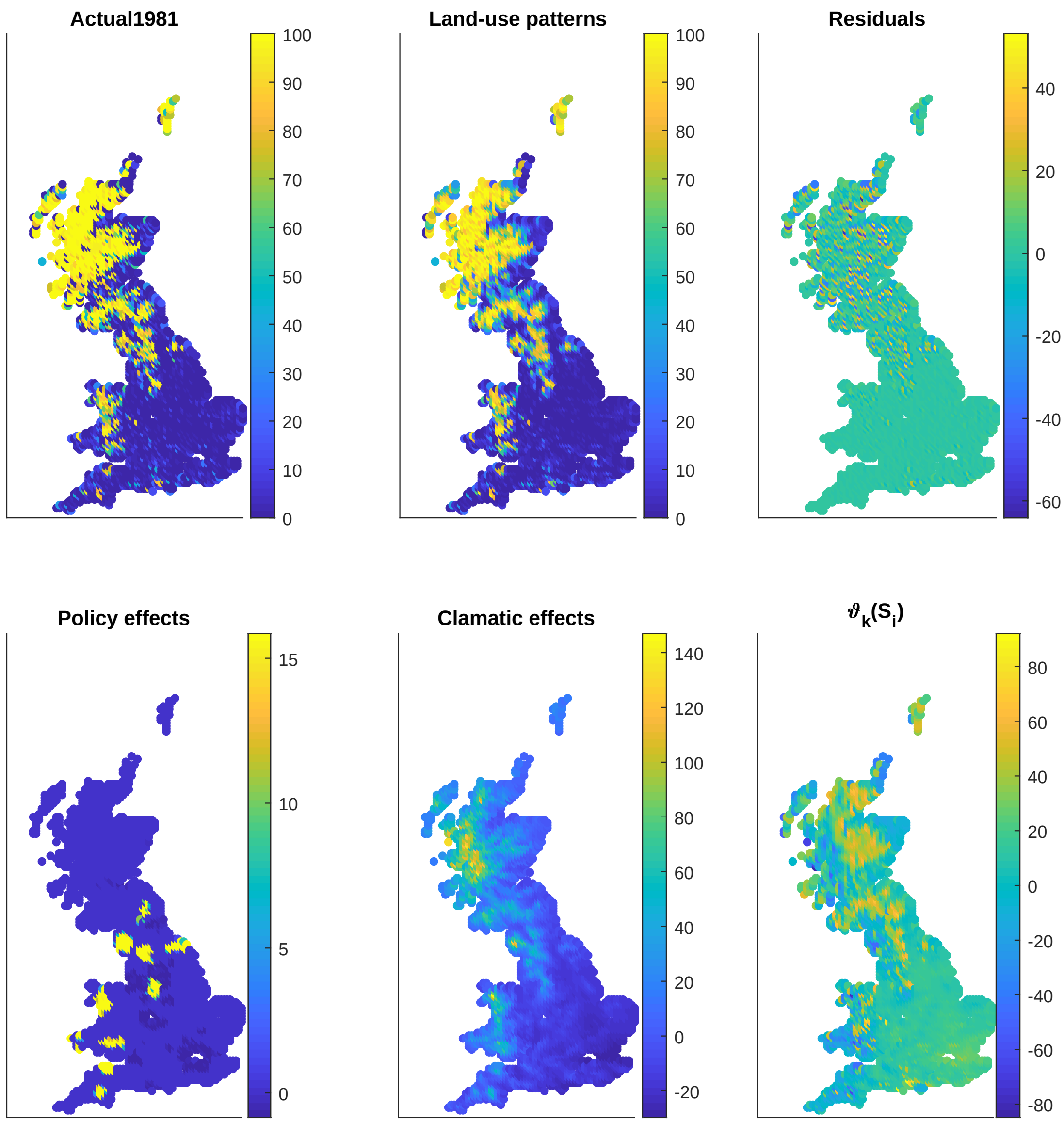

Figure 11: Rough Grazing 1976

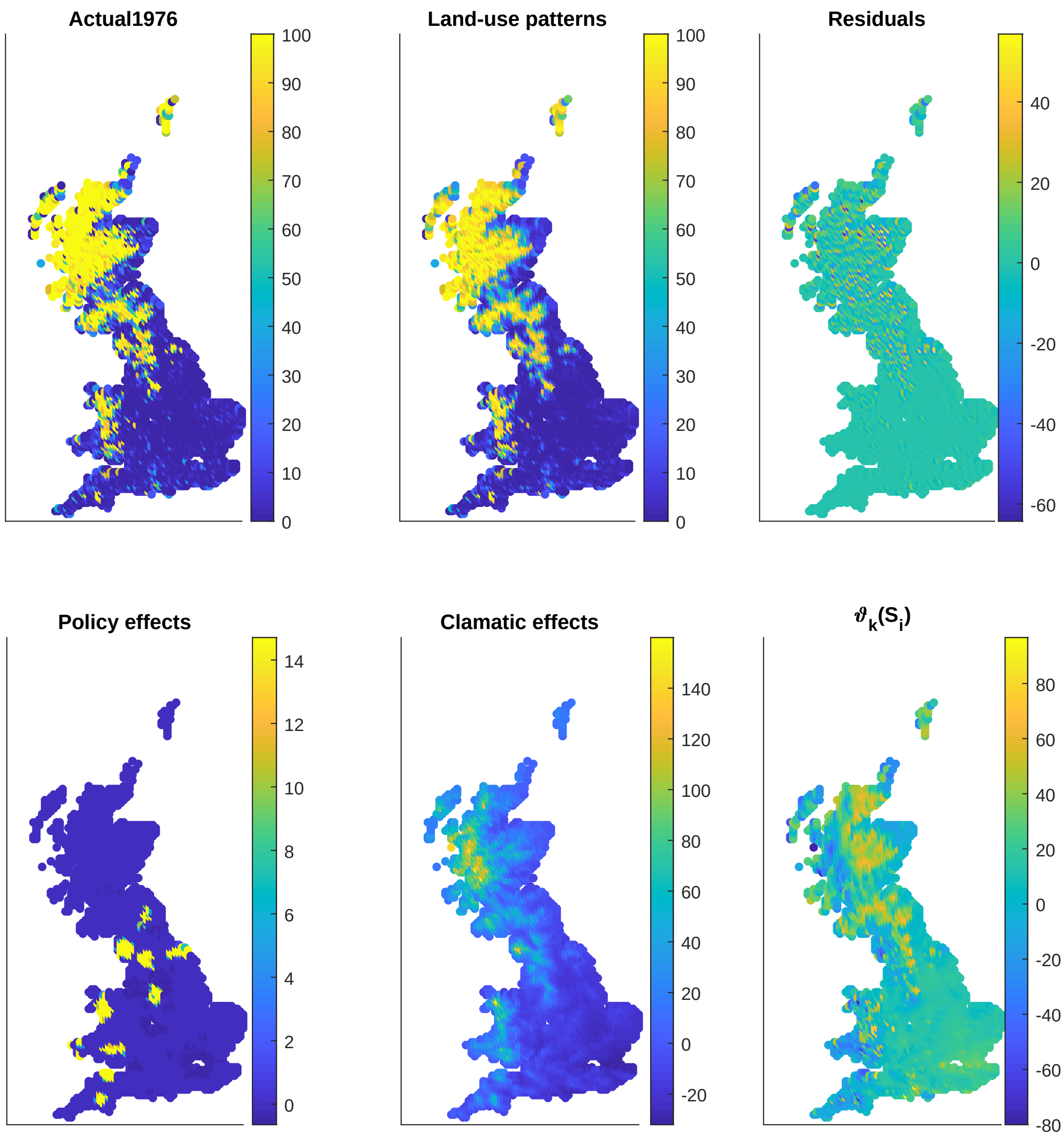

Figure 12: Arable 2004

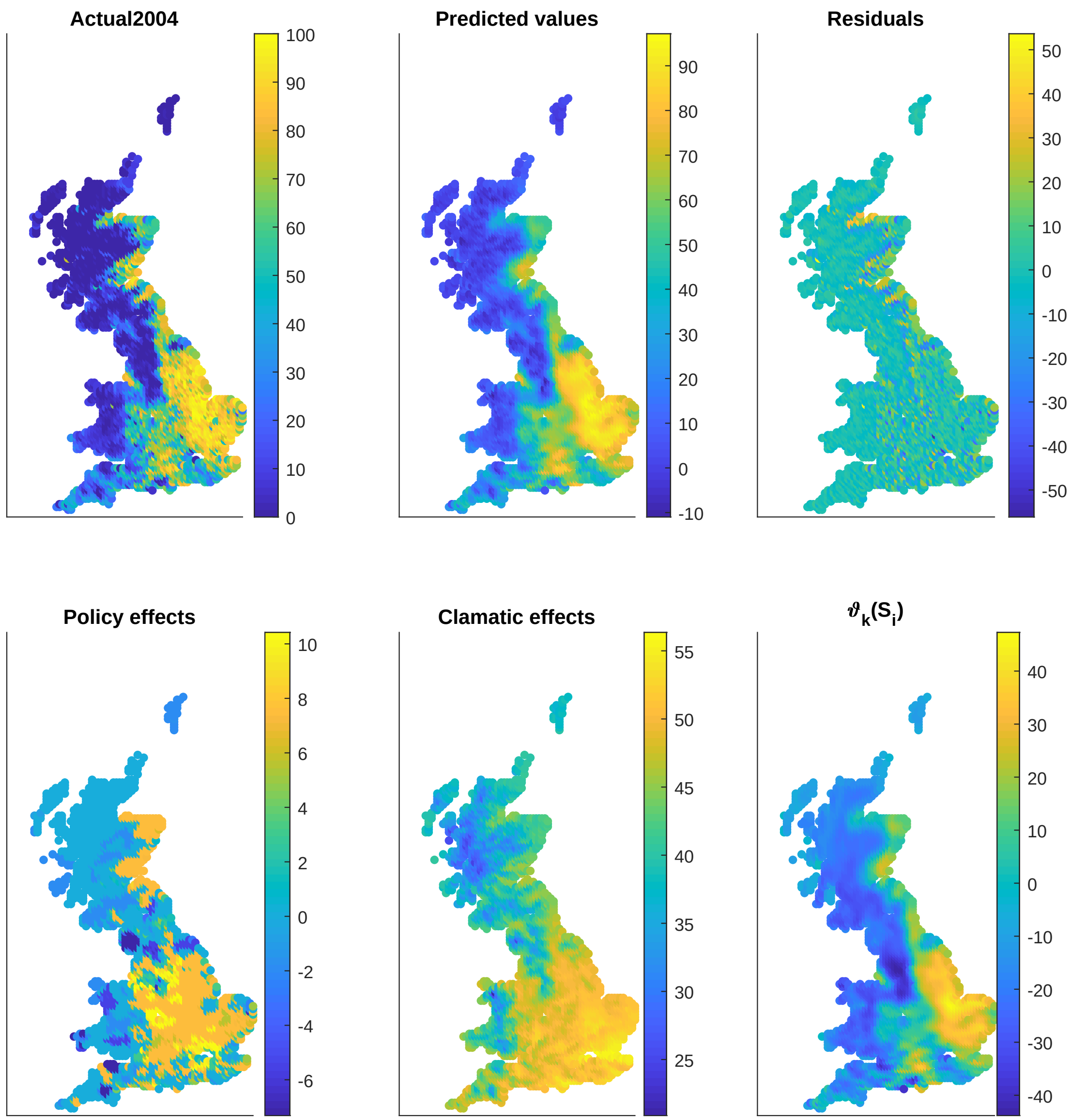

Figure 13: Arable 1981

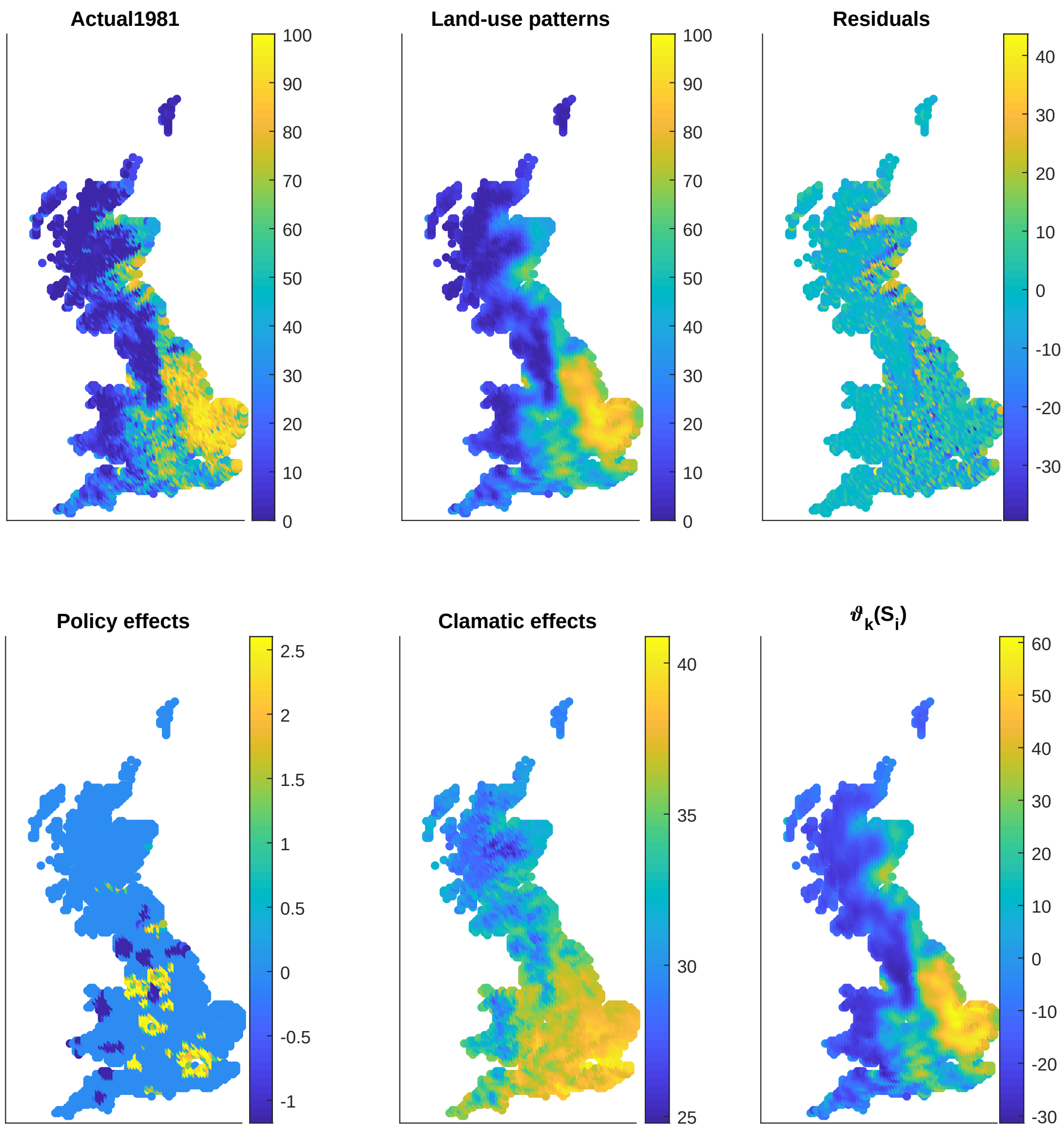

Figure 14: Arable 1976

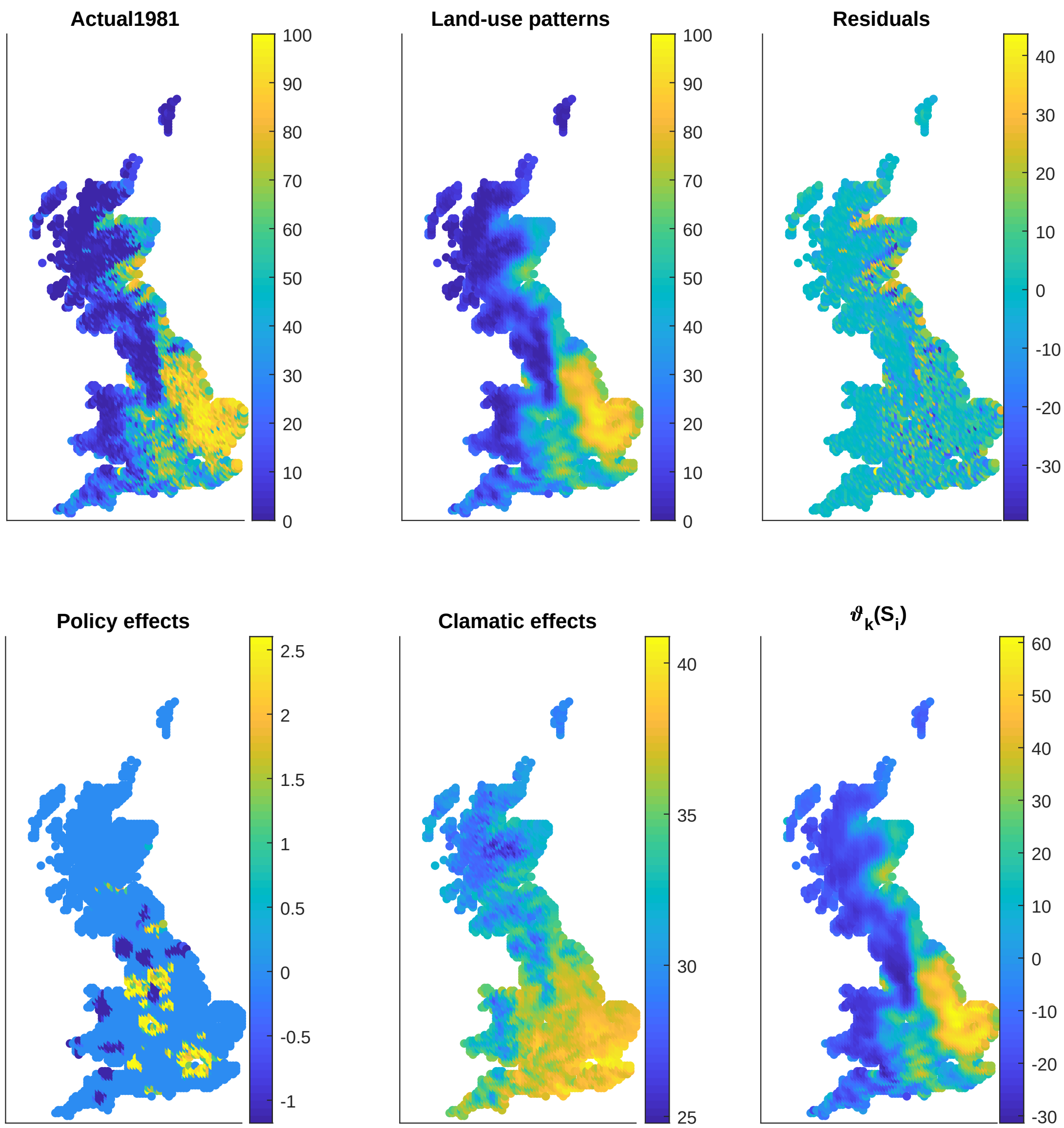

Figure 15: HistSelectedData

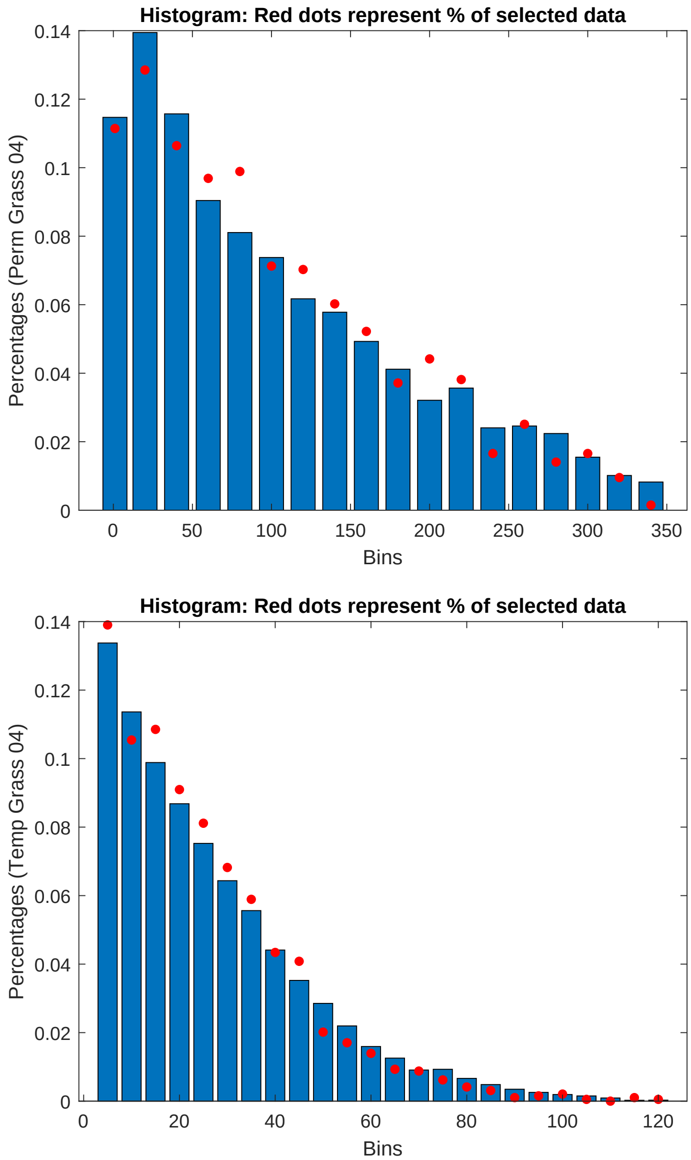

Figure 16: Kaplan-Meier estimates for 1976 with 95% confidence bounds. Estimates ignoring the censoring problem are presented in green colour.

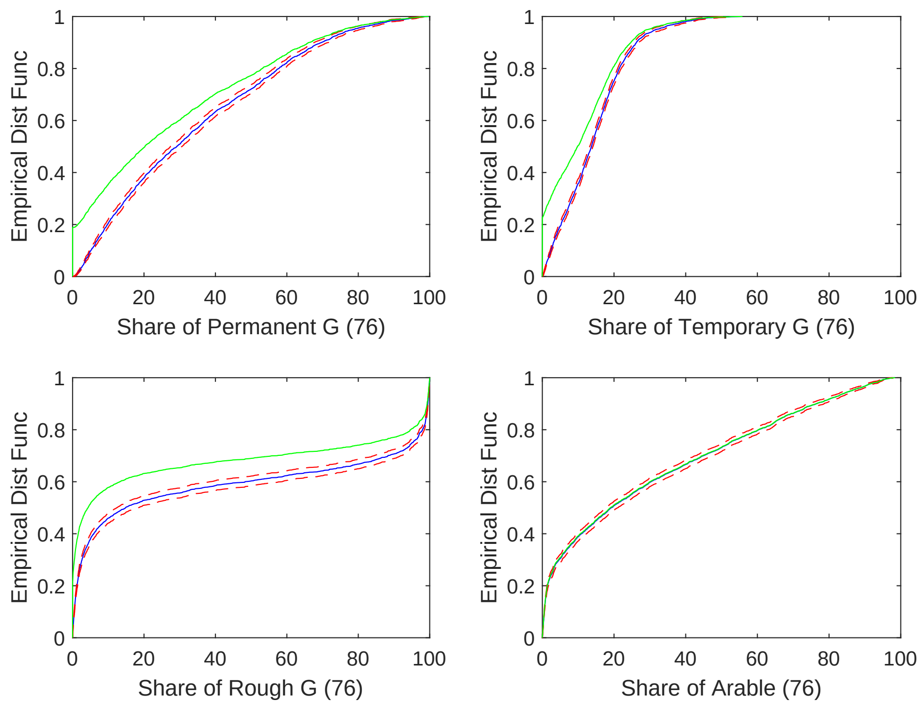

Figure 17: Kaplan-Meier estimates for 1981 with 95% confidence bounds. Estimates ignoring the censoring problem are presented in green colour.

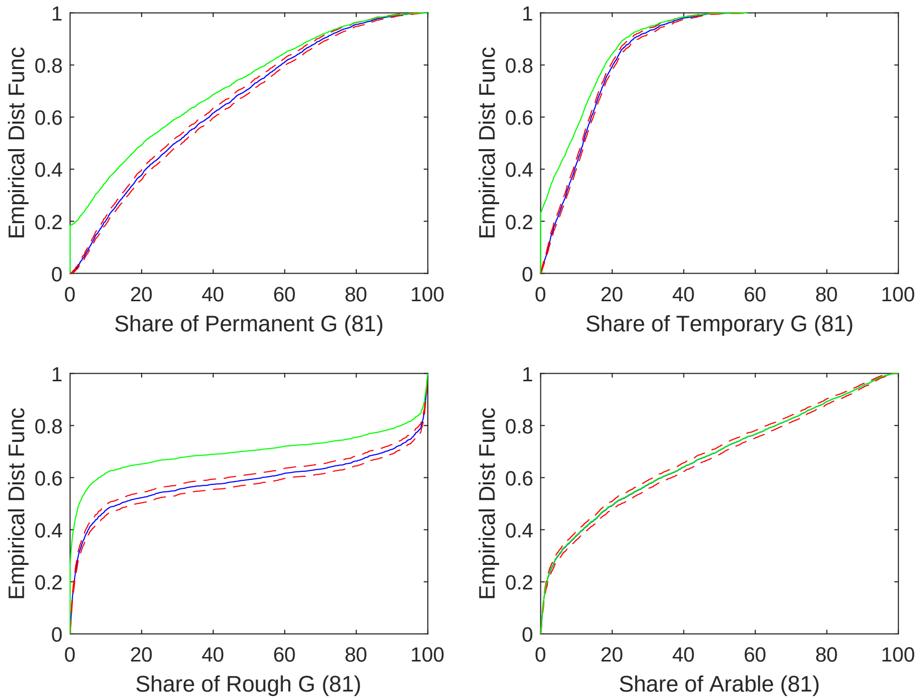

Figure 18: Kaplan-Meier estimates for 2004 with 95% confidence bounds. Estimates ignoring the censoring problem are presented in green colour.