



UNIVERSITY OF WARSAW

Faculty of Economic Sciences

WORKING PAPERS

No. 17/2016(208)

WIKTOR BUDZIŃSKI
DANNY CAMPBELL
MIKOŁAJ CZAJKOWSKI
URŠKA DEMŠAR
NICK HANLEY

USING GEOGRAPHICALLY WEIGHTED CHOICE MODELS TO ACCOUNT FOR SPATIAL HETEROGENEITY OF PREFERENCES

WARSAW 2016



Using geographically weighted choice models to account for spatial heterogeneity of preferences

WIKTOR BUDZIŃSKI

Faculty of Economic Sciences

University of Warsaw

e-mail: wbudzinski@wne.uw.edu.pl

MIKOŁAJ CZAJKOWSKI

Faculty of Economic Sciences

University of Warsaw

e-mail: miq@wne.uw.edu.pl

DANNY CAMPBELL

University of Stirling

Stirling Management School

e-mail: danny.campbell@stir.ac.uk

URŠKA DEMŠAR

University of St Andrews,

School of Geography and Geosciences Development

e-mail: urska.demsar@st-andrews.ac.uk

NICK HANLEY

School of Geography and Geosciences Development

e-mail: ndh3@st-andrews.ac.uk

Abstract

In this paper we investigate the prospects of using geographically weighted choice models for modelling of spatially clustered preferences. The data used in this study comes from a discrete choice experiment survey regarding public preferences for the implementation of a new country-wide forest management and protection program in Poland. We combine it with high-resolution geographical information system data related to local forest characteristics. Using locally estimated discrete choice models we obtain location-specific estimates of willingness to pay (WTP). Variation in these estimates is explained by the socio-demographic characteristics of respondents and characteristics of the forests in their place of residence. The results are compared with those obtained from a more typical, two stage procedure which uses Bayesian posterior means of the mixed logit model random parameters to calculate individual-specific estimates of WTP. The latter approach, although easier to implement and more common in the literature, does not explicitly assume any spatial relationship between individuals. In contrast, the geographically weighted approach differs in this aspect and can provide additional insight on spatial patterns of individuals' preferences. Our study shows that although the geographically weighted discrete choice models have some advantages, it is not without drawbacks, such as the difficulty and subjectivity in choosing an appropriate bandwidth. We find a number of notable differences in WTP estimates and their spatial distributions. At the current level of development of the two techniques, we find mixed evidence on which approach gives the better results.

Keywords:

discrete choice experiment, contingent valuation, willingness to pay, spatial heterogeneity of preferences, forest management, passive protection, litter, tourist infrastructure, mixed logit, geographically weighted model, weighted maximum likelihood, local maximum likelihood

JEL:

Q23, Q28, I38, Q51, Q57, Q58

Acknowledgements:

This study was carried out as a part of the POLFOREX project ("Forest as a public good. Evaluation of social and environmental benefits of forests in Poland to improve management efficiency"; PL0257) funded by EEA Financial Mechanism, Norwegian Financial Mechanism and Polish Ministry of Science and Higher Education. Funding support is gratefully acknowledged. MC gratefully acknowledges the support of the Polish Ministry of Science and Higher Education and the Foundation for Polish Science.

Working Papers contain preliminary research results.

Please consider this when citing the paper.

Please contact the authors to give comments or to obtain revised version.

Any mistakes and the views expressed herein are solely those of the authors.

1. Introduction

Preferences for environmental goods are likely to display spatial patterns such as clustering. One reason for this is that since there are differences in the spatial configuration of these goods, peoples' preferences adapt to their local environments ([Nielsen, Olsen and Lundhede, 2007](#)) and to the availability of substitutes ([Munro and Hanley, 1999](#)). So, for instance, living in an area with high levels of forest cover may induce people to value the conservation of forests more highly than those who live in less-forested landscapes. Another line of reasoning concerns residential sorting: people's preferences for environmental goods partly determine where they choose to live, so that measures of preferences tend to be correlated with measures of environmental quality or with distance to environmental amenities ([Timmins and Murdock, 2007](#); [Timmins and Schlenker, 2009](#); [Baerenklau et al., 2010](#)). These two views on the reasons why spatial patterns are present in values and preferences for the environment can be thought of as different perspectives in the *directional drivers* which link values with environmental features. Recent developments in Geographical Information Science (GIScience) allow the researcher to obtain rich datasets containing detailed information about the spatial configuration of environmental goods and indeed in the socio-economic characteristics of households, and to investigate spatial patterns in stated and revealed preferences for such environmental goods.

In this paper, we employ a method which, while widely used in other areas of social science, remains relatively unknown in environmental economics – geographically weighted regression, as proposed by [Fotheringham, Charlton and Brunsdon \(1998\)](#). Specifically, we propose geographically weighted choice models to investigate the spatial relationship of willingness to pay for landscape characteristics. The rationale of this statistical approach is that, if spatial clusters of preferences do exist, a locally-weighted maximum likelihood method can be used to derive location-specific estimates. Estimation of such models can provide us with insight regarding nonlinear spatial patterns of preferences and welfare measures. This is a non-parametric approach in that no *a priori* assumptions about the spatial distribution of preferences are made.

An important contribution of our study comes from the comparison of the results obtained using geographically weighted discrete choice (GWDC) models with a more standard “two stage”

analysis, which involves estimating a Mixed Logit (MXL) model to derive individual-specific WTP values, and using these individual-specific WTP estimates as dependent variables in a GIS-based spatial regression, as previously done by [Czajkowski *et al.* \(Forthcoming\)](#). This comparison reveals that although there are similarities in the spatial distributions of preferences identified using the two methods, the results differed in several aspects between the two modelling approaches, the most important of which was the difference in the estimated median WTP, and reversed influence of some of the landscape variables.

Overall, we find that GWDC models allow for new insights regarding the spatial distribution of preferences. It is difficult to conclude whether this approach is superior to using the Bayesian posterior means from an MXL model, as both methods have several shortcomings, which are discussed. Two areas of future research which could improve the GWDC approach are extending the model to incorporate unobserved preference heterogeneity, and developing robust methods to avoid some of the arbitrariness in their setup and estimation.

2. Geographically weighted models in the literature

Geographically weighted models belong to the general class of “locally estimated” models. These recognize that the relationship between analyzed variables may be highly nonlinear and, therefore, is difficult to determine it parametrically. Early examples of such models consist of use of spline functions ([Wahba, 1990](#)), LOWESS regression ([Cleveland, 1979](#)) and kernel regressions ([Fan and Gijbels, 1996](#)). The geographically weighted approach differs, because it recognizes nonlinear relationships with respect to spatial dimensions.

Early applications of geographically weighted models were based solely on linear local models. They were used for analysis of morbidity ([Fotheringham, Charlton and Brunsdon, 1998](#)), house price data ([Brunsdon, Fotheringham and Charlton, 1999](#)), economic growth ([LeSage, 1999](#)), school performance ([Fotheringham, Charlton and Brunsdon, 2001](#)) and urban temperatures ([Páez, Uchida and Miyamoto, 2002](#)). In the context of non-market valuation, this approach has been used with

hedonic price models of house prices in [Cho, Poudyal and Roberts \(2008\)](#) or [Saphores and Li \(2012\)](#).

Local likelihood models were introduced by [Fan, Heckman and Wand \(1995\)](#) and [Fan, Farmen and Gijbels \(1998\)](#). These use weighted maximum likelihood estimators for inference. Applications of these techniques in discrete choice models are very limited, and were undertaken mostly in context of transportation. Locally estimated models are used either to recover WTP distribution non parametrically such as in [Fosgerau \(2007\)](#), [Börjesson, Fosgerau and Algers \(2012\)](#) and [Koster and Koster \(2015\)](#) or to analyze behavioral tendencies such as the implications of prospect theory ([Hjorth and Fosgerau, 2012](#)) and preference dynamics ([Dekker, Koster and Brouwer, 2014](#)). None of these approaches used local discrete choice models to analyze spatial heterogeneity. Geographically weighted models for discrete response variables were used in the past, but in different context such as in [Luo and Kanala \(2008\)](#) or [Wang, Kockelman and Wang \(2011\)](#). Our paper aims at filling this gap and exploring the advantages and limitations of this approach in the context of understanding the spatial heterogeneity of environmental values.

3. Methodology

3.1. Geographically weighted multinomial logit

The geographically-weighted multinomial logit (GWMNL) model is defined as follows. A respondent n 's utility from choosing alternative i in the j -th choice task is given by:

$$U_{ijn} = V_{ijn} + \varepsilon_{ijn} = \beta_l' \mathbf{X}_{ijn} + \varepsilon_{ijn} , \quad (1)$$

where the error term ε_{ijn} is assumed to be i.i.d with a Gumbel distribution. β_l is a set of parameters for location l . The assumption which allows for the estimation of such a model is that individuals located close to each other are assumed to have more similar preference parameters than individuals located far away from each other, which is consistent with either of the directional drivers in section 1. As a result, the parameters become spatially correlated. For convenience and ease of comparison between this approach and the [Czajkowski et al. \(Forthcoming\)](#) approach

(described in detail in 3.2 below) we estimated the GWMNL in WTP-space ([Train and Weeks, 2005](#)).

This means that equation (1) was reformulated as:

$$\begin{aligned} U_{ijn} &= \beta_l^{\text{non-cost}} \mathbf{X}_{ijn}^{\text{non-cost}} - \beta_l^{\text{cost}} X_{ijn}^{\text{cost}} + \varepsilon_{ijn} = \\ &= \beta_l^{\text{cost}} \left(\frac{\beta_l^{\text{non-cost}}}{\beta_l^{\text{cost}}} \mathbf{X}_{ijn}^{\text{non-cost}} - X_{ijn}^{\text{cost}} \right) + \varepsilon_{ijn} = \beta_l^{\text{cost}} \left(\mathbf{a}_l' \mathbf{X}_{ijn}^{\text{non-cost}} - X_{ijn}^{\text{cost}} \right) + \varepsilon_{ijn} \end{aligned} \quad (2)$$

where now $\mathbf{a}_l, \beta_l^{\text{cost}}$ are parameters to be estimated. Note that because every local model has an MNL form, this reformulation does not influence the estimates and may only change the standard errors.

Estimation of the GWMNL model is conducted by estimating L ‘local’ models, where L is a number of distinct locations. In the case of our study, there were 253 distinct locations of respondents (unique postal codes) and therefore this is the number of the local models. Each local model is estimated via the weighted maximum likelihood method. The likelihood of individual n making a choice in a j -th choice task in the l -th local model is given by a standard multinomial logit formula:

$$L_{j,n}^l = \sum_{i=1}^I y_{ijn} \frac{\exp(\beta_l' \mathbf{X}_{ijn})}{\sum_k \exp(\beta_l' \mathbf{X}_{kijn})}. \quad (3)$$

In order to take the panel nature of the data into account, we have calculated robust standard errors that are clustered at the individual level. The weighted log-likelihood for l -th model is defined as follows:

$$WL^l = \sum_{n=1}^N \sum_{j=1}^J \lambda(Lat_n, Long_n, b, l) \log(L_{j,n}^l), \quad (4)$$

where $\lambda(Lat_n, Long_n, b, l)$ is a geographical weight (kernel), which depends on latitude and longitude of individual n 's location, b which is called the ‘bandwidth parameter’ and the location l for which the local model is estimated. Note that geographically weighted models normally use projected data, with the location given as metric coordinates X and Y (easting and northing), and indeed this was the same in our case. However, for clarity and to avoid potential confusion with

independent variables $\mathbf{X}_{ijn}$ we refer to the two geographic coordinates here as latitude and longitude. There are a few functional forms of $\lambda(\cdot)$ proposed in the literature. In what follows we use the Gaussian kernel defined as:

$$\lambda(Lat_n, Long_n, b, l) = \exp\left(-0.5 \frac{(Lat_n - Lat_l)^2 + (Long_n - Long_l)^2}{b^2}\right) \quad (5)$$

This is simply an exponential function of minus half of squared Euclidean distance of individual n 's location from location l divided by the square of the bandwidth parameter. If a respondent lives exactly in location l – this weight is equal to 1.¹ The use of this weight implies the clustering of similar values because observations near to location l have a larger bearing on the local model's log-likelihood compared to observations that are further away.

It is worth noting, however, that the choice of bandwidth may have a greater impact on the results than the choice of a specific weighting scheme ([Fosgerau, 2007](#)). There are several methods for choosing the bandwidth parameter available in the literature, with no apparent dominant approach. We tested three approaches, namely: using the corrected Akaike Information Criterion ([AIC, Dekker, Koster and Brouwer, 2014](#)), taking the lowest bandwidth for which all local models converge ([Dekker, Koster and Brouwer, 2014](#)), and using leave-one-individual-out cross-validation criterion ([Fotheringham, Brunson and Charlton, 2003](#)). To evaluate them, we used simulated data which utilized the designs utilized in our study. The results indicated that the available methods led to either under or over-smoothing and were unsatisfactory – a conclusion earlier voiced by [Koster and Koster \(2015\)](#). We therefore decided to use the 'eye-balling' approach they propose. In this approach, a researcher chooses the lowest bandwidth for which the model estimates satisfy a set of *a priori* specified conditions (e.g., achieving identification of all the models or avoiding

¹ For robustness check, we also tried different weighting functions, such as the spatially varying kernel ([Fotheringham, Brunson and Charlton, 2003](#)): $\lambda(Lat_n, Long_n, b, l) = \exp\left(-\frac{R_{n,l}}{b}\right)$, where $R_{n,l}$ is the rank of the n -th location from l -th location in terms of the distance n is from l . The results were not much different from the Gaussian kernel.

extreme estimates). [Pagan and Ullah \(1999\)](#) recommend using this approach when the number of bandwidth parameters is not greater than 2, which is the case in our analysis. We decided that all WTP estimates should lie in interval [-100, 100] EUR, for results to be reliable. Bandwidths that result in WTP estimates outside of this range can be considered as inappropriate. Robustness tests of our results with respect to the bandwidth parameter are presented in the Appendix A.

3.2. Individual and location specific mixed logit

The baseline for the comparison of the performance of GWMNL approach is provided by the traditional, individual-specific conditional distributions retrieved from the mixed logit model ([Czajkowski et al., Forthcoming](#)) which can be estimated in WTP-space ([Train and Weeks, 2005](#)), and the same model with location specific conditional distributions. In the former, respondent n 's utility from choosing alternative i in the j -th choice task is given by:

$$U_{ijn} = \beta_n^{\text{cost}} (\mathbf{a}_n' \mathbf{X}_{ijn}^{\text{non-cost}} - X_{ijn}^{\text{cost}}) + \varepsilon_{ijn} . \quad (6)$$

As before, it is assumed that every individual has a separate, independent set of parameters. Individual parameters are not directly observed, but it is possible to estimate their values implied by each respondents' choices conditional on the population-level estimates of parameter distributions (Bayesian posterior means) using the Bayes theorem:

$$E(\mathbf{a}_n | \mathbf{y}_n, \mathbf{X}_n, \theta) = \int \mathbf{a}_n \frac{p(\mathbf{y}_n | \mathbf{X}_n, \theta, \mathbf{a}_n, \beta_n^{\text{cost}}) f(\mathbf{a}_n, \beta_n^{\text{cost}} | \theta)}{p(\mathbf{y}_n | \mathbf{X}_n, \theta)} d(\mathbf{a}_n, \beta_n^{\text{cost}}), \quad (7)$$

where $p(\mathbf{y}_n | \mathbf{X}_n, \theta, \mathbf{a}_n, \beta_n^{\text{cost}})$ is the likelihood of individual n making the observed choices conditional on the values of random parameters, $p(\mathbf{y}_n | \mathbf{X}_n, \theta)$ is the same likelihood but unconditional (so it is likelihood function for MXL) and $f(\mathbf{a}_n, \beta_n^{\text{cost}} | \theta)$ is the assumed pdf function of random parameters (normal distribution for all attributes except for the cost, which was assumed to be log-normally distributed). For more details about this approach see [Czajkowski et al. \(Forthcoming\)](#).

When comparing the two approaches it is necessary to bear in mind that individual-specific MXL results can include different sources of unobserved preference heterogeneity, while the GWMNL accounts for the spatial heterogeneity only.

Also note that contrary to the MXL model, in the GWMNL there is no need to specify a distribution from which the parameters are drawn. However, while the GWMNL directly accommodates spatial correlation, a drawback of this approach is that individuals who are located in the exactly same location are assumed to have the same set of preference parameters. In order to make these approaches comparable we estimated the MXL model with location-specific parameters – we assumed that individuals from the same location have the same set of preference parameters² and calculated individual conditional expected values of random parameters in the same way as described with equation (7). It is important to note that in this specification the MXL model parameters associated with different locations are independent. Spatial dependence is, therefore, accommodated indirectly as it arises from the calculation of conditional expected values of random parameters.

4. Data

4.1. Discrete Choice Experiment

The dataset used in this study was investigated in [Czajkowski *et al.* \(2014\)](#) and [Czajkowski, Giergiczny and Greene \(2014\)](#). The original survey was conducted in 2010 on a representative sample³ of 1001 Polish adults. The main objective of the survey was estimate public preferences

² Formally, this requires that parameters $\alpha_n, \beta_n^{\text{cost}}$ in equation (6) need to be indexed over l (locations) instead of n (individuals).

³ We hired a professional polling agency that collected the questionnaires using high-quality, face-to-face computer-assisted surveying techniques. A multi-stage sampling strategy was employed, in which communities were randomly selected to represent different community types, and then within each of the selected communities. A starting point address was randomly selected and then a set of addresses was chosen using the random route method. Finally, a

over management options for national forests. The attributes used to describe these management options were (1) passive protection of the most ecologically valuable forests,⁴ (2) reducing the amount of litter (garbage, rubbish) in forests through tougher law enforcement and by increasing forest cleaning services and (3) increasing the level of recreational infrastructure, such as improved signposting of forest trails.

The first attribute concerns the 3% of all 90,000 km² of Polish forests which are perceived as the most ecologically valuable. Currently only about half of these forests are properly protected, so that levels for this attribute were set as:



Status quo

Passive protection of 50% of the most ecologically valuable forests
(1.5% of all forests area)



Partial improvement

Passive protection of 75% of the most ecologically valuable forests
(2.25% of all forests area)



Substantial improvement

Passive protection of 100% of the most ecologically valuable forests
(3% of all forests area)

For the second attribute, the amount of litter in forests was used. Littering decreases the recreational value of forests and potentially also reduces non-use values. The following levels were chosen for this attribute:

random selection of an adult household member was employed.

⁴ By passive (as opposed to active) protection of the forest, we mean leaving the forest ecosystem without any human intervention, even if this results in (natural) changes in ecosystems. It was highlighted that passive protection does not preclude recreational use.

**Status quo**

No changes in amount of litter

**Partial improvement**

Decrease amount of litter in forests of about 50%

**Substantial improvement**

Decrease amount of litter in forests of about 90%

Finally, recreational facilities in the forest were described as taking 3 possible levels:

**Status quo**

No changes in infrastructure

**Partial improvement**

Appropriate tourist infrastructure in an additional 50% of forests

**Substantial improvement**

Appropriate tourist infrastructure available in twice as many (100 % more) forests

An annual cost of making changes to national forest management was also included, defined in terms of increases in annual income tax payments of 0, 10, 25, 50 or 100 PLN.⁵

Every respondent completed 26 choice tasks⁶ with 4 alternatives. In every choice task a Status Quo alternative with no changes in each forest attribute and a zero additional tax cost was included.

⁵ The bid levels and their range were selected based on qualitative pretesting and pilot study results. They were chosen so that they were credible for respondents, used round numbers and to assure good efficiency of the design (given the prior expectations).

⁶ The preference stability and learning or fatigue effects in the course of the 26 choice tasks are analyzed in [Czajkowski](#).

The experimental design was optimized for Bayesian (median) d-efficiency of the MXL model ([Sándor and Wedel, 2001](#); [Ferrini and Scarpa, 2007](#); [Bliemer, Rose and Hess, 2008](#); [Scarpa and Rose, 2008](#)). To obtain initial estimates (priors) and verify the qualitative properties of the questionnaire itself, we conducted a pilot study on a sample of 50 respondents. An example choice card is provided as Figure 1. For more details about design and survey see [Czajkowski et al. \(2014\)](#).

Figure 1. An example choice card used in the DCE study (translation)

	Alternative 1	Alternative 2	Alternative 3	Alternative 4
Protection of ecologically valuable forests	 Status quo Passive protection of 50% of the most ecologically valuable forests (1.5% of all forests)	 Status quo Passive protection of 50% of the most ecologically valuable forests (1.5% of all forests)	 Status quo Passive protection of 50% of the most ecologically valuable forests (1.5% of all forests)	 Substantial improvement Passive protection of 100% of the most ecologically valuable forests (3% of all forests, 100% increase)
Litter in forests	 Status quo No change in the amount of litter in the forests	 Partial improvement Decrease the amount of litter in the forests by half (50% reduction)	 Status quo No change in the amount of litter in the forests	 Partial improvement Decrease the amount of litter in the forests by half (50% reduction)
Infrastructure	 Status quo No change in tourist infrastructure	 Status quo No change in tourist infrastructure	 Partial improvement Appropriate tourist infrastructure in an additional 50% of the forests (50% increase)	 Substantial improvement Appropriate tourist infrastructure available in twice as many forests (100% increase)
Cost	0 PLN	10 PLN	25 PLN	100 PLN
Your choice	☐	☐	☐	☐

[Giergiczny and Greene \(2014\)](#).

4.2. GIS data

Information on forest characteristics used in this study was obtained from two different sources. Firstly, the CORINE Land Cover (CLC) dataset was used. This project is coordinated by the European Environment Agency with the objective of collecting high resolution data for the whole continent.⁷ CLC databases contain area data for objects with a minimum area of 5 ha and a width of more than 100 meters. The second source of information used was the Polish Information System of State Forests. This contains very precise data about the characteristics of forests in Poland. The data from these sources was aggregated to 10x10 km squares.⁸ In total, 3,307 such squares cover the area of Poland. Figure 2 presents a map with a distribution of DCE study respondents. The GIS data were associated with particular respondents using their ZIP-codes identifying their place of residence. For every respondent, the explanatory variables were calculated as weighted averages of forest characteristics in the 10x10 km area common with respondents' ZIP area code. The GIS variables used in this study are described in Table 1.

⁷ See <http://www.eea.europa.eu/publications/COR0-landcover> for further information on the CORINE program.

⁸ We also tested aggregating the 50x50 km resolution which provided equivalent results, although the models were worse fitted.

Figure 2. Respondents and forest area spatial distribution⁹

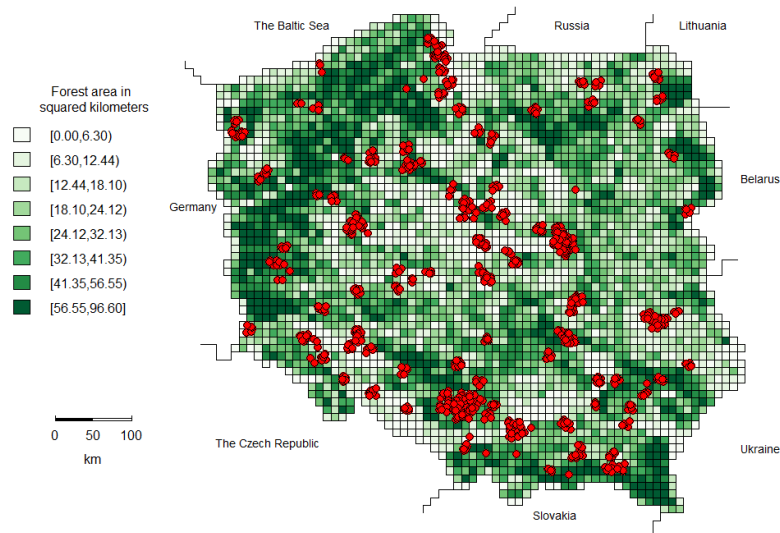


Table 1. GIS variables used to characterize the locations in which respondents' lived

Variable name	Description	Source
Area of coniferous forests	Sum of areas of all coniferous forests [km ²]	Corine Land Cover
Area of deciduous forests	Sum of areas of all deciduous forests [km ²]	Corine Land Cover
Area of mixed forests	Sum of areas of all mixed forests [km ²]	Corine Land Cover
Average Euclidean distance to forest	It is average distance from any point in 10x10 km square to the nearest forest	Corine Land Cover
Area of forests with age > 120	Sum of areas of all forests older than 120 years [km ²]	Information System of State Forests
Area of forests with no. of species > 6	Sum of areas of all forests with no. of tree species greater than 6 [km ²]	Information System of State Forests
Built-up area	Built-up area [km ²]	Corine Land Cover

⁹ As some respondents reported the same ZIP-codes we jittered all of them with uniform random variable on [-5,5] x [-5,5] km square. This also allow us to compute spatial weights matrix.

Maps presenting the distribution of the environmental characteristics of forests used in this paper are presented in Appendix B. Basic characteristics of variables considered in our model are presented in Table 2.

The input for our geographically weighted models was the spatial data set of the respondents, where the location was given as coordinates of ZIP-codes, and the locations were linked to responses and environmental variables. Prior to running the GWR models, the original WGS1984 coordinates were projected using the ETRS_1989_Poland_CS92 coordinate system and the projected coordinates were normalized.

Table 2. Characteristics of variables used in regressions on estimated WTP.

GIS variables			Socio-demographic variables		
Name	Mean	St. Dev.	Name	Mean	St. Dev.
Area of coniferous forests	11.3202	13.3060	Age	44.2957	16.0257
Area of deciduous forests	4.2290	3.9805	Higher Education	0.2288	0.4203
Area of mixed forests	6.5767	6.1084	Income ¹⁰	3.2777	0.9984
Area of forests with age >120	0.9586	1.3336	No. of forests visited in last 12 months	2.4076	4.5873
Average Euclidean distance to the forest	1.3075	0.8921	Number of trips to the forests in last 12 months	49.4276	68.5458
Built-up area	19.5532	19.3520	Sex	0.4216	0.4941
Area of forests with no. of species > 6	5.9285	7.1911	Household size	2.9501	1.3811

¹⁰ Since many respondents did not report their income in the survey the variable used as an income proxy was the question “How would you rate the financial situation of your household?” measured on 5 point Likert scale with 1=“Bad” and 5 = “Good”. We made sure the results are robust with respect to using other measures of respondents’ income (mean income in the area, unemployment rate, percentage population employed, population density, build-up area).

5. Results

Following the approach outlined in section 3.1 we estimated the GWMNL models for each of the 253 distinct locations in which our 1,001 respondents were located.¹¹ Table 3 presents the summary statistics of the estimated parameters for this model.¹² In all cases, the means of the estimated parameters are positive and significant, while the standard deviations are significantly different from zero, which implies the presence of spatially driven preference heterogeneity.

Table 3. Characteristics of the distributions of estimated parameters from the GWMNL model. (standard errors in [] brackets, coefficients in WTP-space, in EUR per year¹³).

	Mean	Std. Dev.
NAT₁	15.7107***	6.8786***
(passive protection of most valuable forests – substantial improvement)	[0.1258]	[0.1701]
NAT₂	23.0767***	10.0173***
(passive protection of most valuable forests – partial improvement)	[0.1826]	[0.2575]
TRA₁	28.3018***	11.0293***
(the amount of litter in forests – partial improvement)	[0.2001]	[0.2316]
TRA₂	37.8590***	14.7594***
(the amount of litter in forests – substantial improvement)	[0.2721]	[0.3624]
INF₁	12.7121***	5.1219***
(tourist infrastructure – partial improvement)	[0.0915]	[0.1003]
INF₂	20.6085***	8.2931***
(tourist infrastructure – substantial improvement)	[0.1388]	[0.1543]
SQ	39.3765***	26.2943***
(alternative specific constant for the no-choice alternative)	[0.3866]	[0.4566]
COST (preference-space equivalent)	0.0575***	0.0240***
(annual cost – tax increase)	[0.0004]	[0.0005]

¹¹ In the estimation, we used the bandwidth parameter of 0.475 which was the lowest value to satisfy our a priori (albeit arbitrarily) specified condition that all models converge and in no location the estimated WTP is larger than 100 EUR. See section 3.1 for discussion and Appendix A for the robustness analysis of this assumption.

¹² Confidence intervals were calculated using Monte-Carlo simulation with 10,000 repetitions, in which parameters of every locally estimated model were assumed to follow multivariate normal distribution.

¹³ At 1 PLN $\approx$ 0.23 EUR $\approx$ 0.25 USD.

As our model was estimated in WTP-space parameters for all attributes can be interpreted directly as willingness to pay. Qualitative results mimic those found in [Czajkowski et al. \(Forthcoming\)](#), namely that on average the individuals are willing to pay the most for reducing of amount of litter and the least for improvements in infrastructure.

For comparison, table 4 presents the results of the location- and individual-specific MXL which mimic the results reported by [Czajkowski et al. \(Forthcoming\)](#). When comparing the two it can be seen that (as expected) the latter fits the data much better. When comparing the WTP estimates, location-specific MXL has higher means and lower standard deviations in all cases aside from the standard deviation of the status quo (SQ) parameter. The comparison of WTP characteristics from location-specific MXL and GWMNL reveals that, in general, in the latter case the mean WTP values are higher. Interestingly, the alternative specific constant parameter (SQ) appears to have a reversed sign – an indication that GWMNL approach can lead to significant changes of the results.¹⁴ Finally, we note that the standard deviations are of similar magnitude in both approaches (except for SQ which has a higher standard deviation in the location-specific MXL model).

¹⁴ This is probably caused by the fact that in GWMNL we assume that every local model has a form of multinomial logit, and therefore assume independence of error terms, which is likely to not be true for non-SQ alternatives. The SQ standard deviation is also relatively large (the coefficient of variation is over 13), which implies a share of respondents with negative and positive marginal utility for the SQ alternative.

Table 4. A comparison of location- and individual- specific mixed logit model results (coefficients in WTP-space, in EUR per year)

Variable	MNL model	Location specific MXL model		Individual specific MXL model	
		Mean	Std. Dev.	Mean	Std. Dev.
	coef. (st. err.)	coef. (st. err.)	coef. (st. err.)	coef. (st. err.)	coef. (st. err.)
NAT₁ (passive protection of most valuable forests – partial improvement)	14.8307*** (0.5673)	11.7961*** (0.4288)	7.6114*** (0.4559)	9.8917*** (0.3436)	11.8622*** (0.5881)
NAT₂ (passive protection of most valuable forests – substantial improvement)	21.8207*** (0.7248)	16.8630*** (0.6129)	12.1756*** (0.6531)	13.5450*** (0.4791)	17.3510*** (0.8286)
TRA₁ (the amount of litter in forests – partial improvement)	26.6697*** (0.8298)	17.4477*** (0.6066)	8.3347*** (0.4433)	11.5526*** (0.3746)	12.8895*** (0.6352)
TRA₂ (the amount of litter in forests – substantial improvement)	35.6782*** (1.0664)	25.2329*** (0.9080)	14.1371*** (0.7508)	17.6876*** (0.5818)	21.4890*** (0.9262)
INF₁ (tourist infrastructure – partial improvement)	12.1400*** (0.5269)	8.2641*** (0.3997)	4.5826*** (0.3251)	6.2377*** (0.2740)	6.1410*** (0.3710)
INF₂ (tourist infrastructure – substantial improvement)	19.5598*** (0.6553)	12.1145*** (0.5044)	6.5138*** (0.3723)	8.6357*** (0.3161)	8.6104*** (0.4837)
SQ (alternative specific constant for the no-choice alternative)	37.2443*** (1.4032)	-3.2461*** (0.7457)	43.2764*** (2.1355)	-13.7474*** (0.9304)	30.9090*** (1.7497)
COST¹⁵ (annual cost – tax increase)	0.0538*** (0.0014)	-2.2472*** (0.0400)	0.7066*** (0.0245)	-1.5776*** (0.0338)	1.0971*** (0.0400)
Model characteristics					
Log-likelihood (constant only)	-36,045.3765	-36,045.3765		-36,045.38	
Log-likelihood	-29,708.2771	-22,632.3017		-17,169.76	
Ben-Akiva Lerman's pseudo-R ²	0.3282	0.4626		0.5266	
McFadden's pseudo-R ²	0.3308	0.3721		0.5236	
AIC/n	2.2836	1.7426		1.3228	
n (observations)	26,026	26,026		26,026	
k (parameters)	8	44		44	

*** p-value < 1%, ** p-value in [1%,5%), * p-value in [5%, 10)

There are, at least, three possible reasons for why we observe significant changes in the mean WTP values. Firstly, one can expect that not allowing for spatial correlation in the specification of the MXL model may lead to biased estimates, and therefore the GWMNL model could be seen as

¹⁵ Underlying normal random variable parameters of the preference-space equivalent of *COST* are reported.

superior. Secondly, as we already noted, the assumption of the MNL model form of local models in GWMNL may not be justified, because the error terms could in reality be correlated. Lastly, and crucially, it may be driven by the distributional assumptions of the MXL model. While the MXL model assumes that the cost*scale parameter is log-normally distributed and that the marginal WTP distributions are all normally distributed, the GWMNL model is a non-parametric approach and thus makes no such assumptions.

In order to compare the relative fit to the data provided by each of the four models (GWMNL, MNL, location- and individual-specific MXL models) we propose to use the Ben-Akiva-Lerman's pseudo- R^2 ([Ben-Akiva and Lerman, 1985](#)). This is a measure of predicted probabilities of choosing the alternatives which were actually chosen by respondents – an intuitive way of illustrating how well a model predicts the observed choices. We adapt this measure to the panel character of our data – because each respondent or each location was associated with n choice tasks, the joint probability of the observed series of n choices is normalized by taking its n -th root. The mean and 95% interquartile ranges are provided in Table 5, while their spatial distribution is illustrated in Figure 3. The pattern that emerges is clear – although the GWMNL approach provides a better fit than the MNL model, it is worse than the location- and individual-specific MXL models. Apparently the ability to generically account for the unobserved preference heterogeneity offers more of an improvement in fit than explicitly accounting for spatial correlations in the MNL model. However, we note that the predicted probabilities are highly correlated – the regions in which respondents' choices are relatively better or worse predicted are unchanged across the four models. This observation is further illustrated with the results provided in Table 6 – the correlation coefficients between location specific choice probabilities predicted by different models.

Table 5. Descriptive statistics of the model-specific predicted probabilities of the observed choices (Ben-Akiva-Lerman's pseudo- R^2)

	GWMNL	MNL	Location specific MXL model	Individual specific MXL model
Mean	0.3550	0.3282	0.4626	0.5266
2.5'th percentile	0.1874	0.2066	0.2682	0.2727
97.5'th percentile	0.5294	0.4396	0.9097	0.9370

Figure 3. Model-specific predicted probabilities of the observed choices (Ben-Akiva-Lerman's pseudo- R^2)

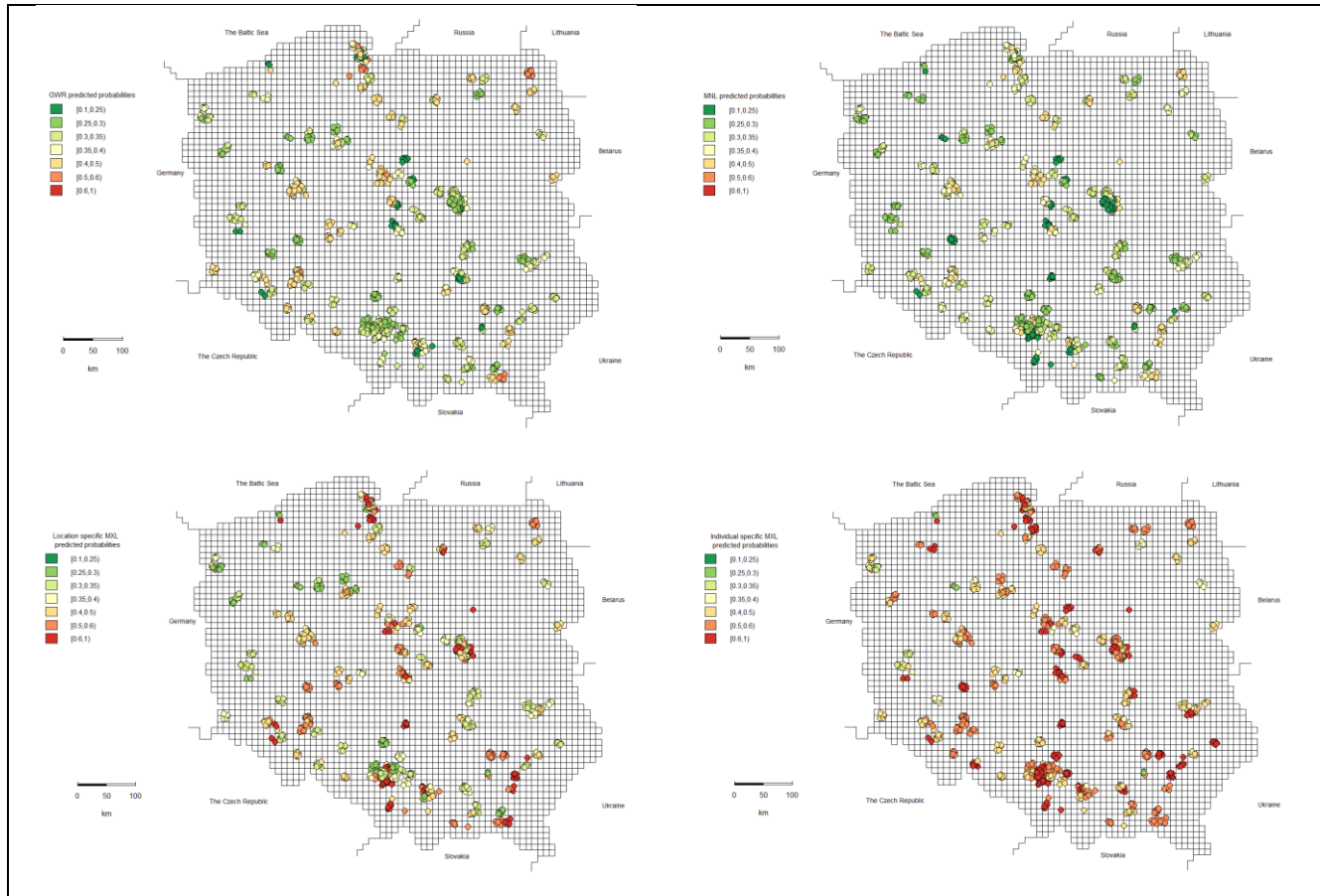


Table 6. Correlation coefficients of model-specific predicted probabilities of the observed choices (Ben-Akiva-Lerman's pseudo- R^2)

	GWR	MNL	Location specific MXL model	Individual specific MXL model
GWR	1.0000	0.7782	0.4110	0.2380
MNL	0.7782	1.0000	0.2313	0.0922
Location specific MXL model	0.4110	0.2313	1.0000	0.7780
Individual specific MXL model	0.2380	0.0922	0.7780	1.0000

In order to analyze the discrepancies between the geographically weighted and the traditional two-step approach further we can compare the differences between WTP estimates for every location presented on map of Poland. This is done in the 7 panels of Figure 4. In these panels positive values depict locations where GWMNL estimates are higher (red) and negative values depict locations where conditional expected values from location-specific MXL are higher (green). Distributions of these differences are not symmetric with respect to 0 (as every interval consists of 10% of the sample). For all attributes, more than 80% of observations have positive values (higher values of WTP from GWMNL).

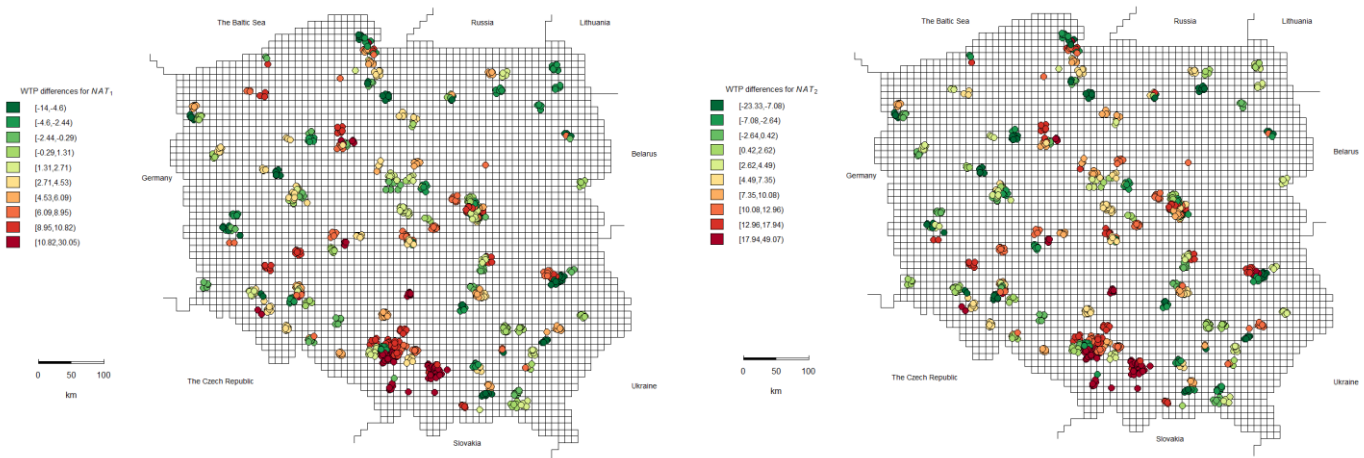
Graphical analysis reveals several spatial patterns of between-estimate differences which are consistent across all attributes. First of all, the largest differences can be observed in the central-south part of Poland near Cracow and Katowice cities, where GWMNL approach leads to much higher estimates of WTP. Secondly, in the west and north-east parts of Poland the differences seem to be much lower, sometimes negative. In the other parts of the country there is no clear spatial pattern, although it seems that also in the central part the differences are rather low (but positive). The fact, that some spatial patterns can be observed is an indication that these the two approaches recover different spatial dependencies of preferences. As GWMNL is designed to recover such spatial dependencies we can expect it to work better in this regard.

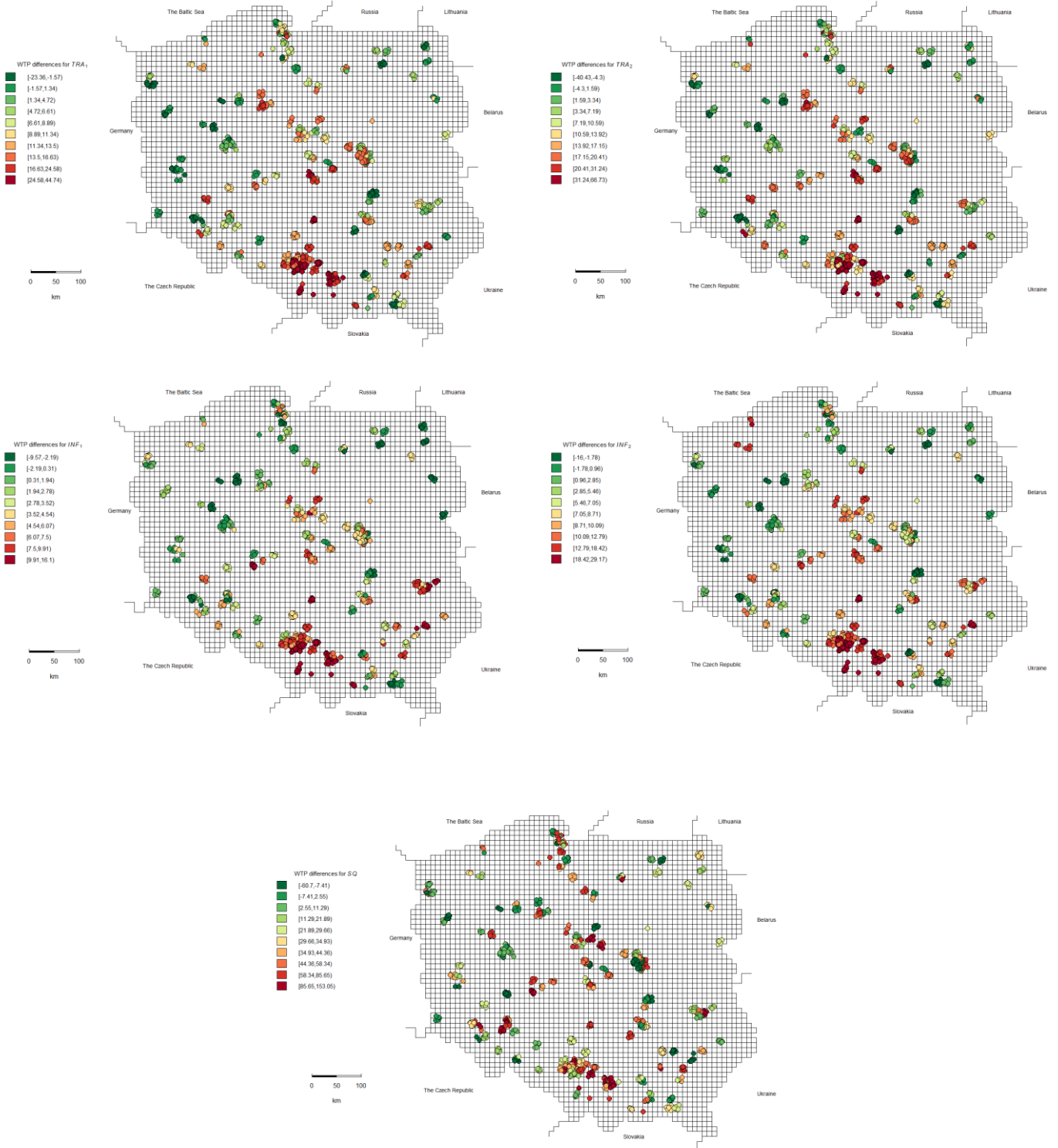
The correlation between the WTP estimates resulting from the GWMNL model and the posterior means from the individual-specific and location-specific MXL models is investigated further using Pearson and Spearman's correlation coefficients. The results are provided in Table 7. As expected, the correlation coefficients are positive, although they are lower than one could have hoped. The individual-specific MXL model gives much lower correlation, which is understandable because it results in different estimates even for the same location. While there is a connection in the WTP distributions, the relatively low correlations signal a disconnect between the models. We note, however, that there is no clear indication whether it is the GWMNL or the traditional two-step approach that is closer to the true data generating process. The ability to explicitly consider spatial correlations and accounting for the unobserved preference heterogeneity could be providing different benefits to a model, and both offer ways to provide a better fit to data, and offer different insights into the data generating process.

Table 7. Correlation coefficients of model-specific predicted probabilities of the observed choices (Ben-Akiva-Lerman's pseudo- R^2)

	MXL location-specific			MXL individual-specific		
	Pearson product-moment correlation coefficient	Spearman's rank correlation coefficient	Mann-Whitney U test statistic	Pearson product-moment correlation coefficient	Spearman's rank correlation coefficient	Mann-Whitney U test statistic
NAT_1	0.3374***	0.3065***	850540***	0.1395***	0.1355***	849407***
NAT_2	0.3699***	0.3421***	829556***	0.1736***	0.1660***	835692***
TRA_1	0.2701***	0.3058***	702578***	0.0254	0.0520	643904***
TRA_2	0.3045***	0.3369***	754979***	0.0851***	0.1004***	726605***
INF_1	0.3469***	0.3520***	761014***	0.0382	0.0327	694666***
INF_2	0.3145***	0.3654***	718047***	0.0237	0.0434	638309***
SQ	0.2784***	0.3089***	728741***	0.1566***	0.1208***	557149***

Figure 4. Spatial distribution of differences between WTP estimates from GWMNL and conditional expected values from location-specific MXL.





Lastly, it is possible to perform a decomposition of the estimated WTP using GIS variables, similar [Czajkowski et al. \(Forthcoming\)](#). To this end, 7 regressions were estimated in which WTP were

explained by the same GIS variables as in [Czajkowski et al. \(Forthcoming\)](#). The results from the spatial lag model on conditional expected values of random parameters from location-specific MXL and individual-specific MXL are presented in Table 8. The results of the current analysis are presented in Table 9.¹⁶ In all cases we decided to use only GIS variables and omit socio-demographic variables, since these were insignificant in both the GWMNL and location-specific MXL approaches, as expected. Spatial lag models for individual-specific MXL were re-estimated without socio-demographic variables to be more comparable.

Most of GIS variables are highly significant. We can compare these estimates with results in Table 9 in order to investigate structural differences in spatial heterogeneity in WTP estimates between the GWMNL and Bayesian posterior means from location-specific MXL. Greyed out cells of Table 8 indicate coefficients which have a different sign than analogous coefficients in Table 9. This issue is the most prominent in regression for the *SQ*, where 3 variables have different signs, although they are all insignificant in this case. This problem also occurs, for almost every attribute, with “*Area of forests with age > 120*” variable. In case of location-specific MXL, this variable has a positive (although insignificant) coefficient for most attributes, whilst for the individual-specific MXL this variable has positive, highly significant coefficient for all attributes. What is also interesting is that in the current analysis variables “*Built-up area*” and “*Area of forests with no. of species > 6*” were significant in almost all cases. This differs to what we discovered when using the conditional expected values of random parameters from the MXL models (both location-specific and individual-specific) where these variables were insignificant in all cases.

¹⁶ Note that in current analysis we used Ordinary Least Square instead of spatial lag model. It was not possible to estimate spatial lag model on WTP estimated from GWMNL as from definition these are highly spatially autocorrelated and therefore coefficient ρ (which is coefficient for spatial lag term) was almost equal to 1 and no other variable would then occur significant.

Table 8. Results of regressions in which WTP estimates from GWMNL are explained by GIS variables

	SQ (alternative specific constant for the no-choice alternative)	NAT₁ (passive protection of most valuable forests – partial improvement)	NAT₂ (passive protection of most valuable forests – substantial improvement)	TRA₁ (the amount of litter in forests – partial improvement)	TRA₂ (the amount of litter in forests – substantial improvement)	INF₁ (tourist infrastructure – partial improvement)	INF₂ (tourist infrastructure – substantial improvement)
Constant	34.7688*** (5.7861)	17.6132*** (1.5842)	27.0083*** (2.4027)	30.8844*** (2.4394)	44.8497*** (3.4203)	11.8696*** (1.0303)	22.0946*** (1.8038)
Area of coniferous forests	-0.1544 (0.1362)	-0.0727* (0.0373)	-0.1088* (0.0566)	-0.0732 (0.0574)	-0.1519* (0.0805)	0.0239 (0.0243)	-0.0449 (0.0425)
Area of deciduous forests	1.0730** (0.4567)	-0.1207 (0.1251)	-0.3464* (0.1897)	-0.0479 (0.1926)	-0.3411 (0.2700)	0.2376*** (0.0813)	0.2164 (0.1424)
Area of mixed forests	-0.1680 (0.3356)	-0.2419*** (0.0919)	-0.3022** (0.1393)	-0.2930** (0.1415)	-0.4331** (0.1984)	-0.0364 (0.0598)	-0.2270** (0.1046)
Area of forests with age >120	-7.0385*** (1.4597)	-0.4937 (0.3996)	-0.5343 (0.6061)	-1.3759** (0.6154)	-1.0958 (0.8628)	-1.4039*** (0.2599)	-2.0912*** (0.4550)
Average Euclidean distance to a forest	-1.7794 (2.3522)	-1.1623* (0.6440)	-1.7118* (0.9768)	-2.3111** (0.9917)	-3.8154*** (1.3904)	0.0347 (0.4189)	-1.1040 (0.7333)
Built-up area	0.2859*** (0.0793)	0.0813*** (0.0217)	0.1123*** (0.0329)	0.1454*** (0.0334)	0.1703*** (0.0469)	0.0129 (0.0141)	0.0405 (0.0247)
Area of forests with no. of species > 6	1.0447*** (0.3331)	0.2774*** (0.0912)	0.3147** (0.1383)	0.3567** (0.1405)	0.4133** (0.1969)	0.1161* (0.0593)	0.3829*** (0.1039)
Model characteristics							
R ²	21.95%	12.15%	9.80%	12.93%	9.25%	15.56%	16.48%
n (observations)	253	253	253	253	253	253	253
k (parameters)	8	8	8	8	8	8	8

*** p-value < 1%, ** p-value in [1%,5%), * p-value in [5%, 10%)

Table 9. Results of spatial lag models in which Bayesian posterior means from MXL models are explained by GIS variables.

	SQ (alternative specific constant for the no-choice alternative)	NAT₁ (passive protection of most valuable forests – partial improvement)	NAT₂ (passive protection of most valuable forests – substantial improvement)	TRA₁ (the amount of litter in forests – partial improvement)	TRA₂ (the amount of litter in forests – substantial improvement)	INF₁ (tourist infrastructure – partial improvement)	INF₂ (tourist infrastructure – substantial improvement)
Location-specific MXL							
Constant	-35.0845*** [9.0012]	15.2120*** [1.9167]	23.2420*** [2.9933]	16.2849*** [2.1964]	26.6299*** [3.5287]	5.9459*** [1.0342]	8.9798*** [1.4828]
Area of coniferous forests	0.4807** [0.2108]	-0.0828** [0.0371]	-0.1442** [0.0595]	-0.0674* [0.0402]	-0.1537** [0.0668]	0.0054 [0.0202]	-0.0129 [0.0284]
Area of deciduous forests	2.5479*** [0.6502]	-0.3785*** [0.1139]	-0.6469*** [0.1829]	-0.1931 [0.1226]	-0.5320*** [0.2043]	0.0471 [0.0616]	0.0096 [0.0867]
Area of mixed forests	0.7688* [0.4489]	-0.1880** [0.0791]	-0.3133** [0.1270]	-0.2060** [0.0859]	-0.3404** [0.1425]	-0.0634 [0.0430]	-0.1022* [0.0607]
Area of forests with age >120	-0.6381 [2.2125]	0.6013 [0.3907]	0.9906 [0.6268]	0.4653 [0.4225]	0.9354 [0.7025]	-0.0776 [0.2117]	0.0971 [0.2983]
Average Euclidean distance to a forest	10.8442*** [3.3104]	-1.6825*** [0.5811]	-2.8926*** [0.9328]	-1.6013** [0.6309]	-3.2686*** [1.0499]	-0.2967 [0.3153]	-0.6871 [0.4454]
ρ	0.2555*** [0.0723]	0.1810** [0.0760]	0.1824** [0.0758]	0.3397*** [0.0681]	0.3325*** [0.0682]	0.3700*** [0.0664]	0.4090*** [0.0637]
Model characteristics							
Log Likelihood	-1255.8570	-814.7967	-934.2825	-838.5166	-966.3454	-665.4564	-753.4436
AIC/n	10.0246	6.5194	7.4643	6.7622	7.7718	5.4062	6.1279
n (observations)	253	253	253	253	253	253	253
k (parameters)	8	8	8	8	8	8	8
Individual-specific MXL							
Constant	-16.8574*** [2.5533]	10.8869*** [1.1637]	15.2921*** [1.6918]	4.0306*** [0.4830]	5.5010*** [0.6690]	8.7263*** [1.0712]	14.6495*** [1.8260]
Area of coniferous forests	0.1481*** [0.0572]	-0.0688*** [0.0253]	-0.1005*** [0.0370]	-0.0153 [0.0105]	-0.0239 [0.0146]	-0.0537** [0.0234]	-0.0965** [0.0401]
Area of deciduous forests	0.5595*** [0.1877]	-0.3243*** [0.0832]	-0.4731*** [0.1215]	-0.0526 [0.0344]	-0.0837* [0.0477]	-0.1848** [0.0767]	-0.3405*** [0.1313]
Area of mixed forests	0.4011*** [0.1382]	-0.1808*** [0.0611]	-0.2643*** [0.0892]	-0.0540** [0.0253]	-0.0796** [0.0352]	-0.1560*** [0.0566]	-0.2749*** [0.0968]
Area of forests with age >120	-1.2737** [0.6278]	0.7988*** [0.2781]	1.1563*** [0.4061]	0.1274 [0.1152]	0.1932 [0.1597]	0.4748* [0.2571]	0.9019** [0.4400]
Average Euclidean distance to a forest	3.2966*** [0.9756]	-1.5019*** [0.4310]	-2.2111*** [0.6295]	-0.3825** [0.1785]	-0.5901** [0.2478]	-1.0983*** [0.3987]	-1.9057*** [0.6820]
ρ	0.5214*** [0.0204]	0.3879*** [0.0243]	0.3942*** [0.0242]	0.5395*** [0.0198]	0.5647*** [0.0189]	0.5482*** [0.0195]	0.5249*** [0.0203]
Model characteristics							
Log Likelihood	-4453.7110	-3597.8680	-3978.4130	-2764.8760	-3101.9990	-3570.4460	-4099.0780
AIC/n	36.8664	29.2900	32.3286	23.6074	26.4754	30.0571	34.0545
n (observations)	1001	1001	1001	1001	1001	1001	1001
k (parameters)	8	8	8	8	8	8	8

*** p-value < 1%, ** p-value in [1%,5%), * p-value in [5%, 10%)

Overall, the results indicate that there are significant discrepancies with regard to spatial patterns recovered with the two methods. Differences in the signs and significance of coefficients for GIS variables demonstrate that the WTP distributions differ in structure between these two approaches. As the GWMNL model explicitly deals with spatial heterogeneity (rather than trying to recover it indirectly post estimation) it could be considered more reliable. It is also important to note that R^2 in all models is very low, which suggests that the GIS variables we used explain only a small fraction of the observed variance. Therefore, most of the preference heterogeneity is caused by some other factors, which were not accounted for. This may partly be due to the spatial distribution of forest characteristics in Poland. As can be seen on maps in Appendix B, in some cases the differences between forests which lie next to each other are large, and therefore significant variance in their values may occur even on a local level. It is possible that the GWMNL approach would perform better in the case of preferences for environmental goods which change more gradually.

6. Discussion and conclusions

In this paper we investigate alternative methods for investigating spatial patterns in willingness to pay for changes to an environmental good. The case study used to generate the data concerned the management of forests in Poland. We introduced a novel method of geographically weighted discrete choice models to account for spatial heterogeneity of model parameters, and compared this to a “two-step” approach using the MXL model and posterior Bayesian means of random parameters ([Czajkowski *et al.*, Forthcoming](#)). The comparison focused on willingness to pay, as this is often a focus in environmental economics.

Our analysis revealed significant differences in the estimates of WTP between the two methods. Specifically, for all forest attributes, mean WTP was much higher when obtained using GWMNL. We also found some structural differences – several GIS variables appeared to have a reversed effect on WTP when the models are compared.

We note that both methods have some shortcomings. MXL assumes that individuals' parameters are drawn from spatially independent distributions and relies heavily on distributional assumptions. Any spatial correlations that are observed are obtained from posterior Bayesian means, and they are, obviously, conditional on the MXL assumptions and the distributions used. On the other hand, GWMNL ignores other (non-spatial) sources of preference heterogeneity.

This shortcoming of the GWMNL model could be addressed by using a more complicated weighting function, which accounts for socio-demographic characteristics – this may also lead to the necessity to use multiple bandwidth parameters. We tried to incorporate heterogeneity with respect to socio-demographic characteristics by using a second bandwidth parameter, but the results were not satisfactory, particularly when using the eye-balling technique to determine the optimal bandwidth value. First of all, in such a setting there was no unique 'lower' bandwidth for which our criterion was fulfilled, e.g. in case of having criterion of all WTP being below 100 EUR, there could exist several pairs of bandwidths for which all WTP values are indeed below this chosen level, such that an increase in any of bandwidths would lead to exceeding this level. Researchers would therefore need some additional criterion to choose the preferable model. The other problem was that the number of models that one needs to estimate increases very quickly with the number of bandwidth parameters, e.g. if for a single bandwidth researcher wants to evaluate 20 different models, then for two bandwidth parameters about $20 \times 20 = 400$ models should be evaluated.

The other possibility is to include unobserved heterogeneity in local models via estimation of latent class models instead of simple multinomial logits. Such approach was used by [Koster and Koster \(2015\)](#) (although not in the spatial context) and may be a preferable approach to GWMNL. It also introduces additional computational burdens, however.

In light of the above, it is difficult to conclude that one of the methods is superior. Further research is needed. Firstly, additional analysis of the reliability of Bayesian posterior means and their vulnerability to MXL assumptions is needed. To some extent, [Hess \(2010\)](#) approach this issue, but with no focus on welfare measures. Secondly, the methods of choosing an appropriate bandwidth parameter in GWDC models are currently underdeveloped. Many of the methods proposed in

literature are unsatisfactory and lead to poor results. This is especially important for more advanced kernels, with multiple bandwidth parameters, which would allow for spatial sources of heterogeneity.

References

- Baerenklau, K. A., González-Cabán, A., Paez, C., and Chavez, E., 2010. Spatial allocation of forest recreation value. *Journal of Forest Economics*, 16(2):113-126.
- Ben-Akiva, M., and Lerman, S. R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, MA.
- Bliemer, M. C. J., Rose, J. M., and Hess, S., 2008. Approximation of Bayesian Efficiency in Experimental Choice Designs. *Journal of Choice Modelling*, 1(1):98-127.
- Börjesson, M., Fosgerau, M., and Algers, S., 2012. Catching the tail: Empirical identification of the distribution of the value of travel time. *Transportation Research Part A: Policy and Practice*, 46(2):378-391.
- Brunsdon, C., Fotheringham, A. S., and Charlton, M., 1999. Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*, 39(3):497-524.
- Cho, S.-H., Poudyal, N. C., and Roberts, R. K., 2008. Spatial analysis of the amenity value of green open space. *Ecological Economics*, 66(2):403-416.
- Cleveland, W. S., 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829-836.
- Czajkowski, M., Bartczak, A., Giergiczny, M., Navrud, S., and Żylicz, T., 2014. Providing Preference-Based Support for Forest Ecosystem Service Management. *Forest Policy and Economics*, 39:1-12.
- Czajkowski, M., Budziński, W., Campbell, D., Giergiczny, M., and Hanley, N., Forthcoming. Spatial heterogeneity of willingness to pay for forest management.
- Czajkowski, M., Giergiczny, M., and Greene, W. H., 2014. Learning and fatigue effects revisited. Investigating the effects of accounting for unobservable preference and scale heterogeneity. *Land Economics*, 90(2):323-350.
- Dekker, T., Koster, P., and Brouwer, R., 2014. Changing with the Tide: Semiparametric Estimation of Preference Dynamics. *Land Economics*, 90(4):717-745.
- Fan, J., Farman, M., and Gijbels, I., 1998. Local maximum likelihood estimation and inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(3):591-608.
- Fan, J., and Gijbels, I., 1996. Local polynomial modelling and its applications: monographs on statistics and applied probability 66. CRC Press.
- Fan, J., Heckman, N. E., and Wand, M. P., 1995. Local polynomial kernel regression for generalized linear models and quasi-likelihood functions. *Journal of the American Statistical Association*, 90(429):141-150.
- Ferrini, S., and Scarpa, R., 2007. Designs with a priori information for nonmarket valuation with choice experiments: A Monte Carlo study. *Journal of Environmental Economics and Management*, 53(3):342-363.
- Fosgerau, M., 2007. Using nonparametrics to specify a model to measure the value of travel time. *Transportation Research Part A: Policy and Practice*, 41(9):842-856.
- Fotheringham, A. S., Brunsdon, C., and Charlton, M., 2003. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons.
- Fotheringham, A. S., Charlton, M. E., and Brunsdon, C., 2001. Spatial variations in school performance: a local analysis using geographically weighted regression. *Geographical and Environmental Modelling*, 5(1):43-66.
- Fotheringham, S., Charlton, M., and Brunsdon, C., 1998. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and planning A*, 30(11):1905-1927.

- Hess, S., 2010. Conditional parameter estimates from Mixed Logit models: distributional assumptions and a free software tool. *Journal of Choice Modelling*, 3(2):134-152.
- Hjorth, K., and Fosgerau, M., 2012. Using prospect theory to investigate the low marginal value of travel time for small time changes. *Transportation Research Part B: Methodological*, 46(8):917-932.
- Koster, P. R., and Koster, H. R. A., 2015. Commuters' preferences for fast and reliable travel: A semi-parametric estimation approach. *Transportation Research Part B: Methodological*, 81, Part 1:289-301.
- LeSage, J. P., 1999. A spatial econometric examination of China's economic growth. *Geographic Information Sciences*, 5(2):143-153.
- Luo, J., and Kanala, N. K., Year. Modeling urban growth with geographically weighted multinomial logistic regression. Geoinformatics 2008 and Joint conference on GIS and Built Environment: The Built Environment and its Dynamics, International Society for Optics and Photonics, 71440M-71440M-71411.
- Munro, A., and Hanley, N., 1999. Information, Uncertainty and Contingent Valuation. In: *Contingent Valuation of Environmental Preferences: Assessing Theory and Practice in the USA, Europe, and Developing Countries*, I. J. Bateman and K. G. Willis, eds., Oxford University Press.
- Nielsen, A. B., Olsen, S. B., and Lundhede, T., 2007. An Economic Valuation of the Recreational Benefits Associated with Nature-Based Forest Management Practices. *Landscape and Urban Planning*, 80:63-71.
- Páez, A., Uchida, T., and Miyamoto, K., 2002. A general framework for estimation and inference of geographically weighted regression models: 1. Location-specific kernel bandwidths and a test for locational heterogeneity. *Environment and planning A*, 34(4):733-754.
- Pagan, A., and Ullah, A., 1999. Nonparametric econometrics. Cambridge university press.
- Sándor, Z., and Wedel, M., 2001. Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, 38(4):430-444.
- Saphores, J.-D., and Li, W., 2012. Estimating the value of urban green areas: A hedonic pricing analysis of the single family housing market in Los Angeles, CA. *Landscape and Urban Planning*, 104(3):373-387.
- Scarpa, R., and Rose, J. M., 2008. Design Efficiency for Non-Market Valuation with Choice Modelling: How to Measure it, What to Report and Why. *Australian Journal of Agricultural and Resource Economics*, 52(3):253-282.
- Timmins, C., and Murdock, J., 2007. A revealed preference approach to the measurement of congestion in travel cost models. *Journal of Environmental Economics and Management*, 53(2):230-249.
- Timmins, C., and Schlenker, W., 2009. Reduced-Form Versus Structural Modeling in Environmental and Resource Economics. *Annual Review of Resource Economics*, 1(1):351-380.
- Train, K., and Weeks, M., 2005. Discrete Choice Models in Preference Space and Willingness-to-Pay Space. In: *Applications of Simulation Methods in Environmental and Resource Economics*, R. Scarpa and A. Alberini, eds., Springer Netherlands, 1-16.
- Wahba, G., 1990. Spline models for observational data. Siam.
- Wang, Y., Kockelman, K., and Wang, X., 2011. Anticipation of land use change through use of geographically weighted regression models for discrete response. *Transportation Research Record: Journal of the Transportation Research Board*, (2245):111-123.

Appendix A

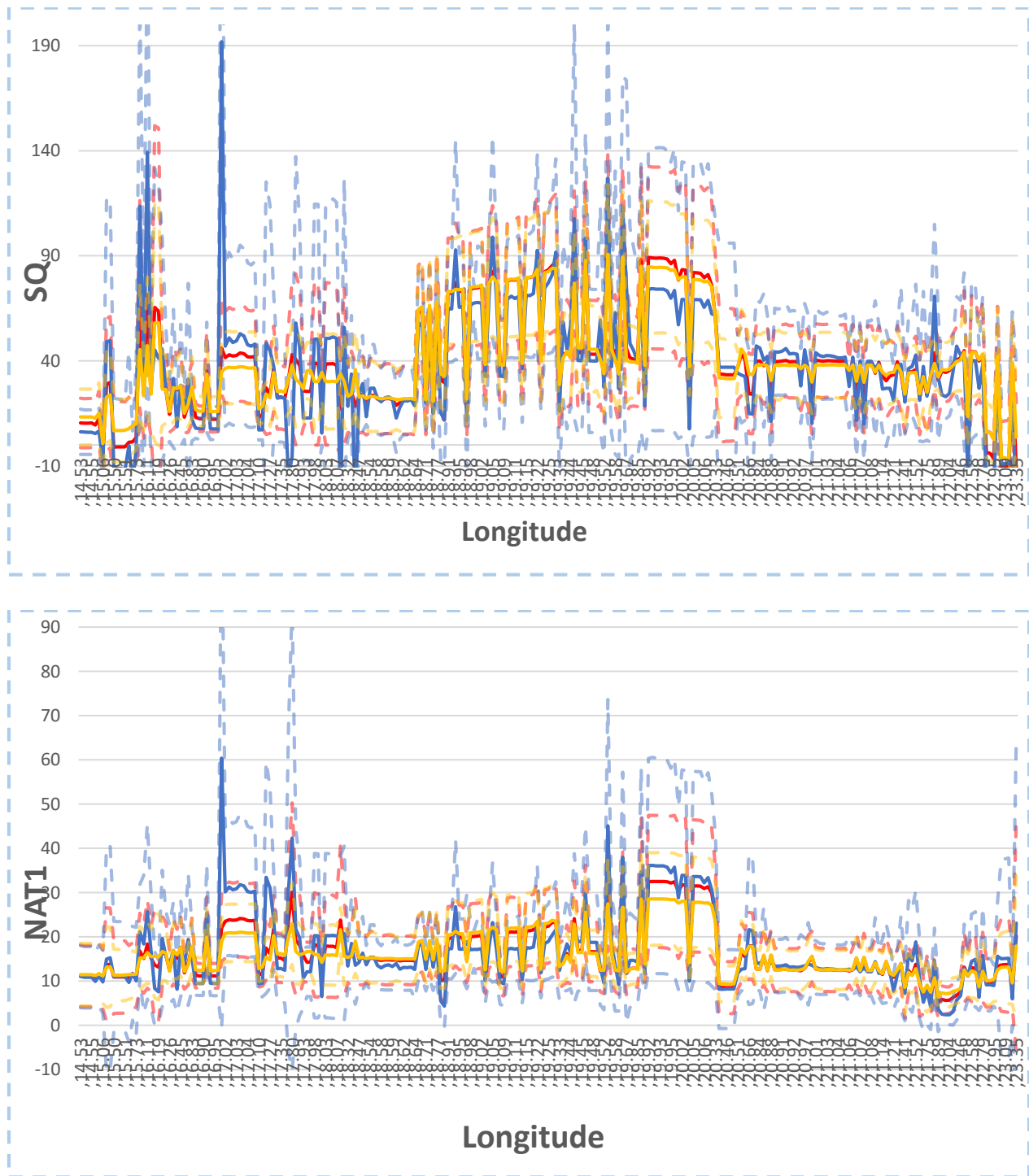
In this section we check robustness of our results by analyzing differences in estimates for 3 different values of bandwidth parameter: 0.275, 0.475 and 0.6. 0.475 is a value used throughout the paper, 0.275 is the lowest bandwidth value for which all WTP except SQ are below 100 EUR and 0.6 is the highest value we considered (as it is generally recommended to use lowest bandwidth value which fulfill chosen criteria). In Table A1. we provide characteristics of WTP for these values of bandwidth analogously as it was done in Table 4. We note that mean WTP are quite robust with respect to choice of a bandwidth. Although mean estimates for bandwidth values of 0.275 and 0.6 lie more than two standard errors from estimates for bandwidth = 0.475, they are fairly close to each other. Dispersion of WTP change significantly between different bandwidth values. To investigate how change in dispersion influence our results we conducted a graphical analysis presented in Figure A1. In every panel we plotted parameters estimates for chosen bandwidth values: 0.275 (blue lines), 0.475 (red lines) and 0.6 (yellow lines) with their 95% confidence intervals (dotted lines) against longitude. We see that, indeed increased dispersion for bandwidth of 0.275 leads to several 'peaks', with very wide confidence intervals, for every parameter. Because of that for lower bandwidth value many WTP estimates in local models occurred insignificant. Estimates for 0.6 bandwidth usually lies within confidence interval for estimates for bandwidth of 0.475. We therefore conclude that choosing bandwidth of 0.475 is justified as it allows to avoid implausible 'peaks' in WTP (especially for SQ) and insignificant estimates. We also do not see any reason to choose bandwidth above 0.475 as the plots does not reveal any particular under-smoothing when comparing lines for bandwidth of 0.475 with lines for bandwidth of 0.6.

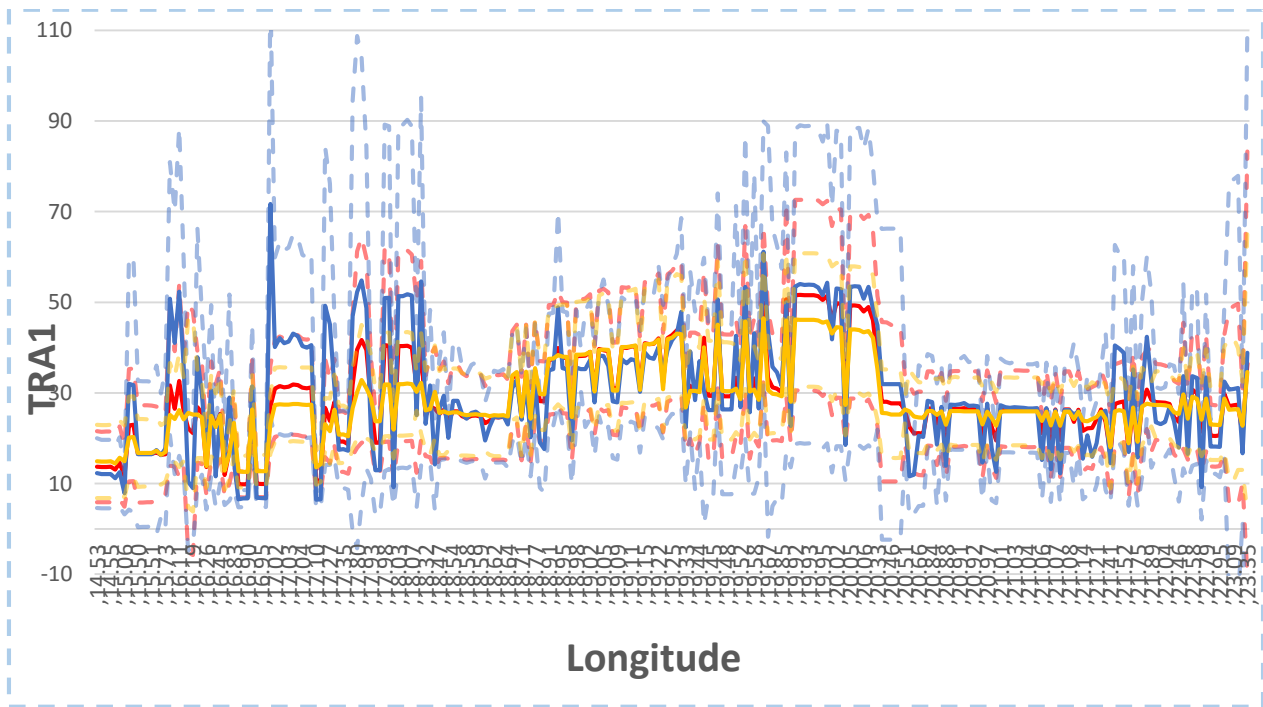
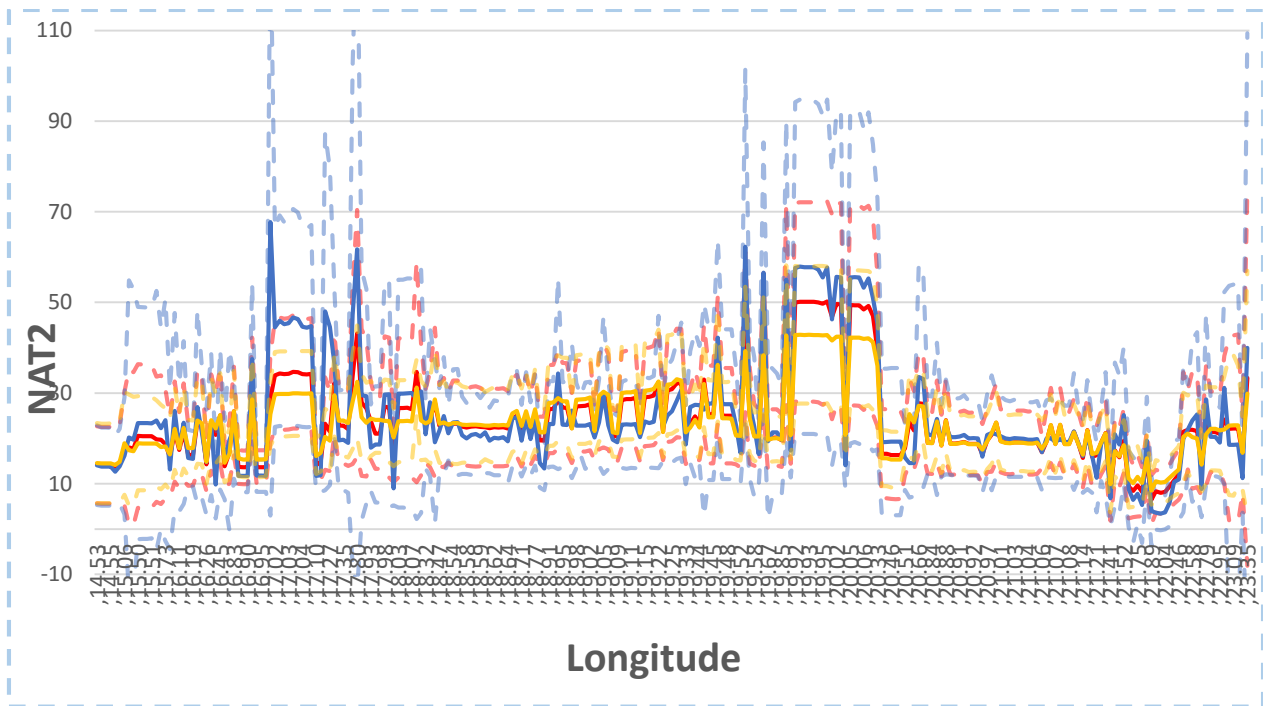
Table A1. Characteristics of distributions of estimated WTP for different bandwidth value (standard errors in [] brackets).

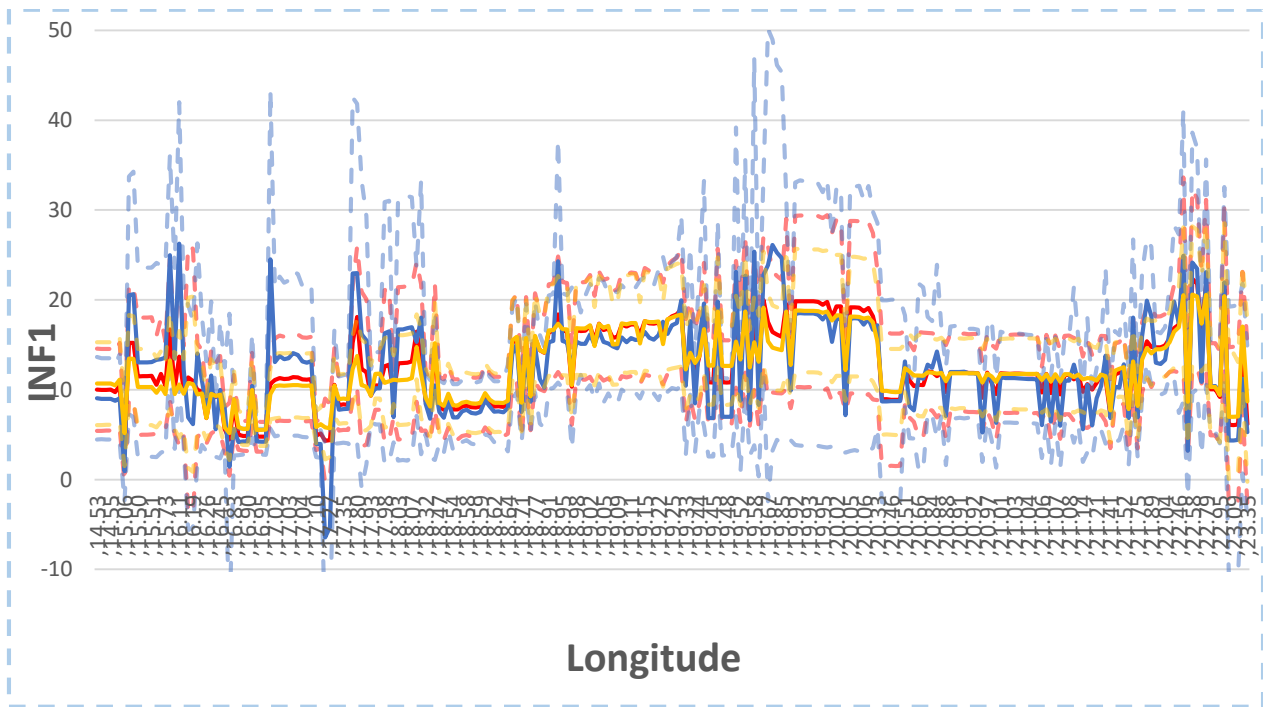
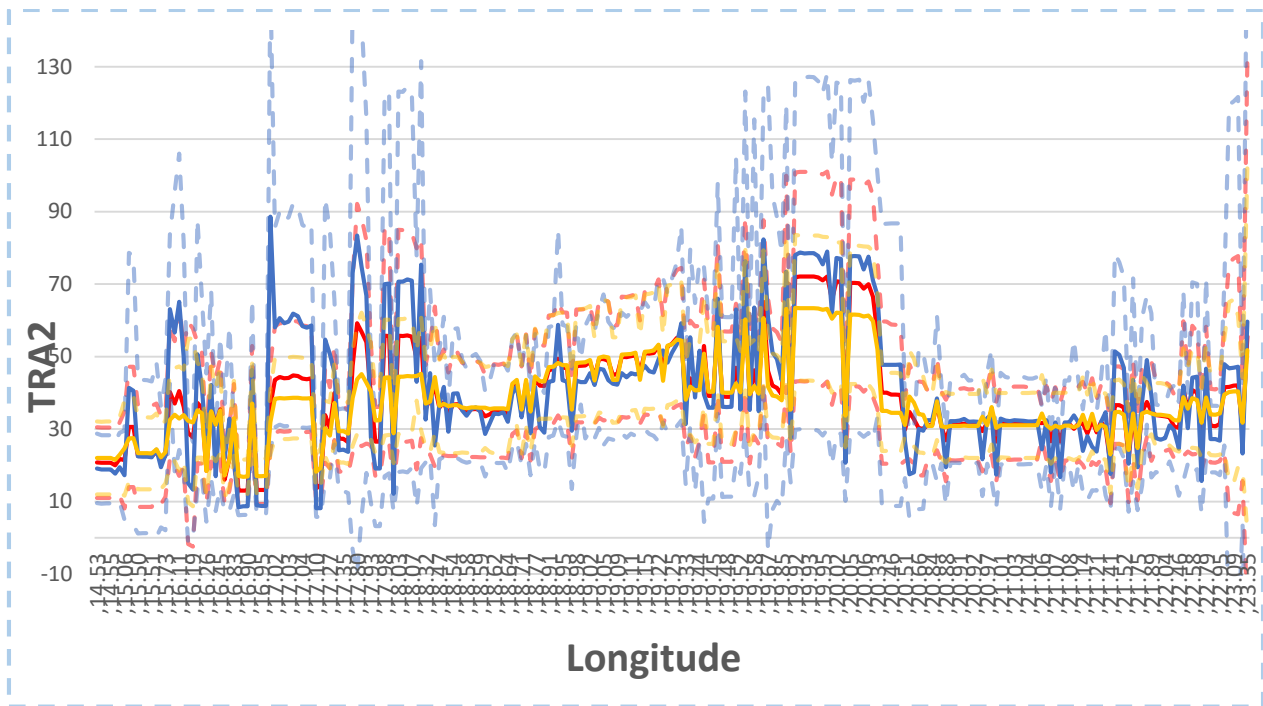
	Bandwidth = 0.275		Bandwidth = 0.475		Bandwidth = 0.6	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
NAT₁ (passive protection of most valuable forests – substantial improvement)	16.4541*** [0.2225]	10.9360*** [0.5033]	15.7107*** [0.1258]	6.8786*** [0.1701]	15.5389*** [0.0968]	5.7223*** [0.1138]
NAT₂ (passive protection of most valuable forests – partial improvement)	24.0868*** [0.3136]	15.4706*** [0.5758]	23.0767*** [0.1826]	10.0173*** [0.2575]	22.7602*** [0.1390]	8.1253*** [0.1682]
TRA₁ (the amount of litter in forests – partial improvement)	29.6563*** [0.3318]	16.7661*** [0.4383]	28.3018*** [0.2001]	11.0293*** [0.2316]	27.8463*** [0.1530]	9.0786*** [0.1704]
TRA₂ (the amount of litter in forests – substantial improvement)	39.8062*** [0.4591]	22.7725*** [0.6970]	37.8590*** [0.2721]	14.7594*** [0.3624]	37.1348*** [0.2041]	11.8575*** [0.2446]
INF₁ (tourist infrastructure – partial improvement)	12.9375*** [0.1469]	7.5310*** [0.1862]	12.7121*** [0.0915]	5.1219*** [0.1003]	12.6410*** [0.0736]	4.3811*** [0.0793]
INF₂ (tourist infrastructure – substantial improvement)	21.1835*** [0.2223]	11.5530*** [0.2611]	20.6085*** [0.1388]	8.2931*** [0.1543]	20.4056*** [0.1064]	7.1172*** [0.1127]
SQ (alternative specific constant for the no-choice alternative)	38.7161*** [0.6308]	37.5901*** [1.1447]	39.3765*** [0.3866]	26.2943*** [0.4566]	39.5860*** [0.3046]	23.3231*** [0.3690]
COST (annual cost – tax increase)	0.0626*** [0.0005]	0.0403*** [0.0008]	0.0575*** [0.0004]	0.0240*** [0.0005]	0.0561*** [0.0003]	0.0182*** [0.0003]

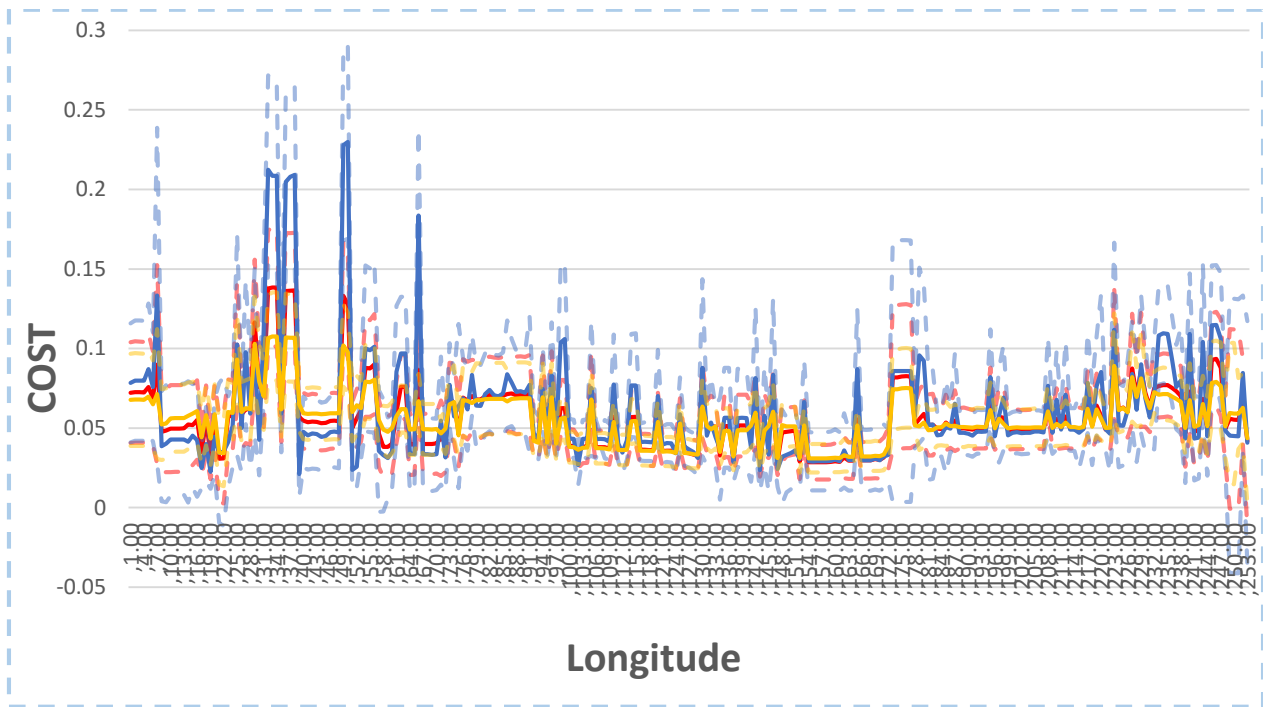
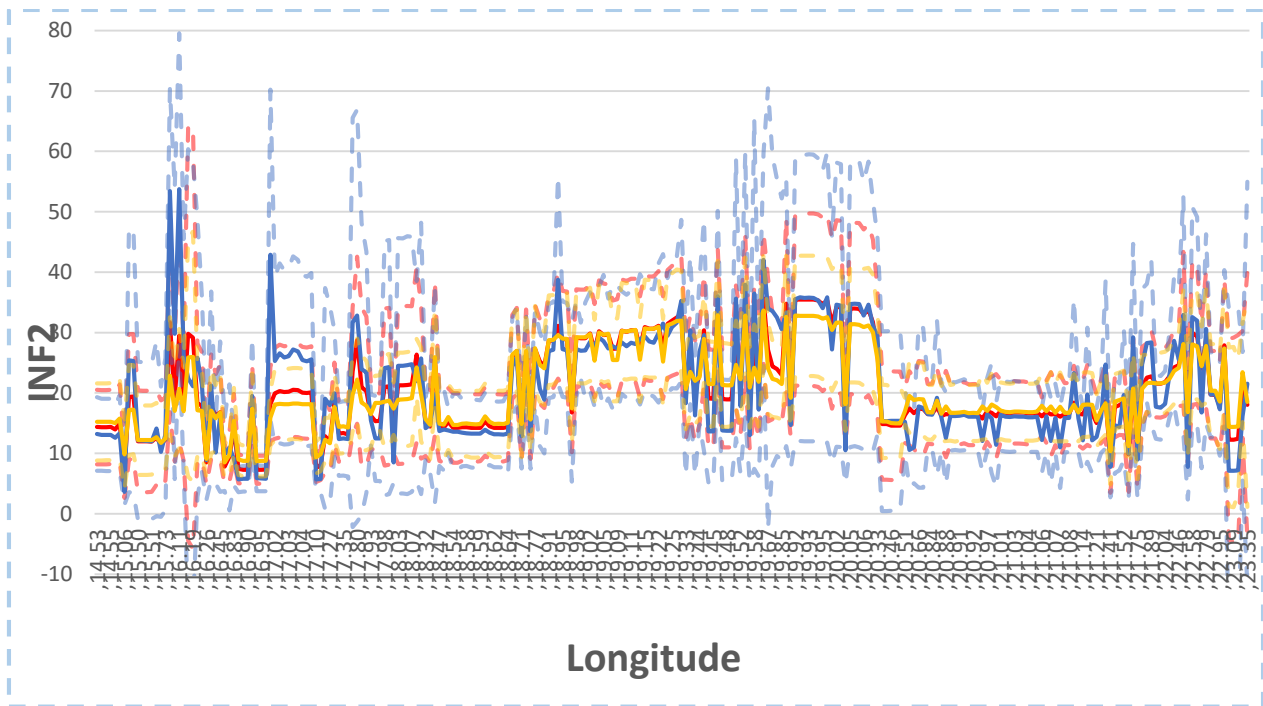
*** p-value < 1%, ** p-value in [1%,5%), * p-value in [5%, 10%)

Figure A1. Parameters values with different bandwidths values plotted against longitude

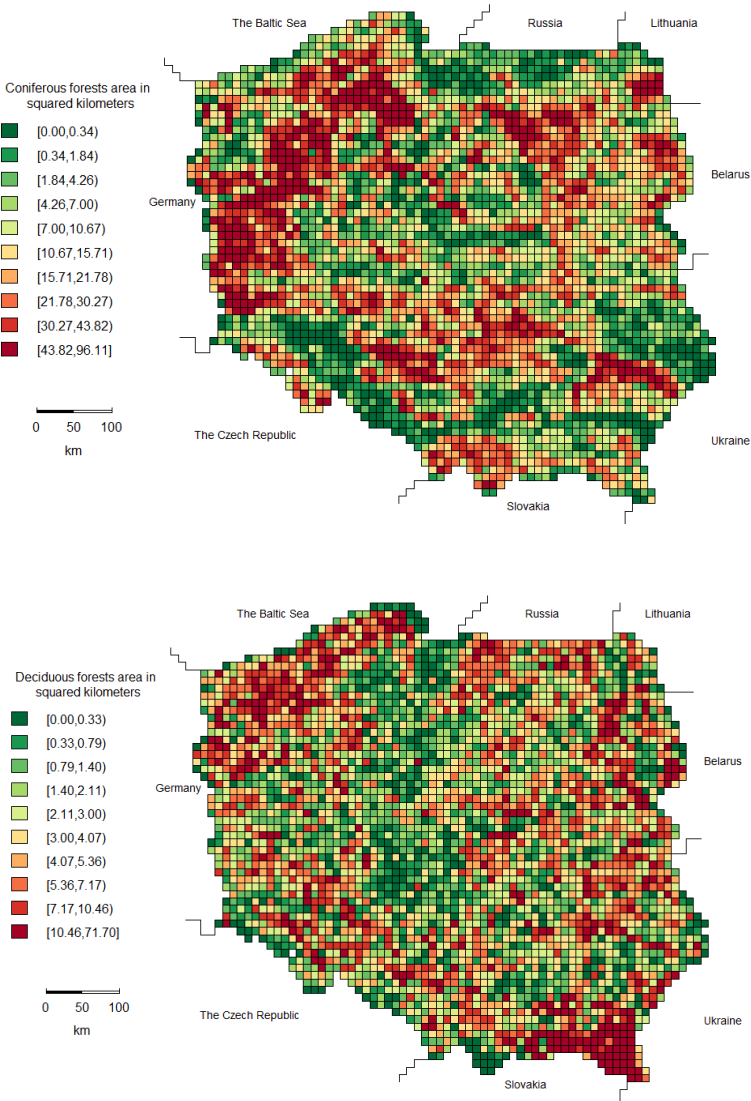


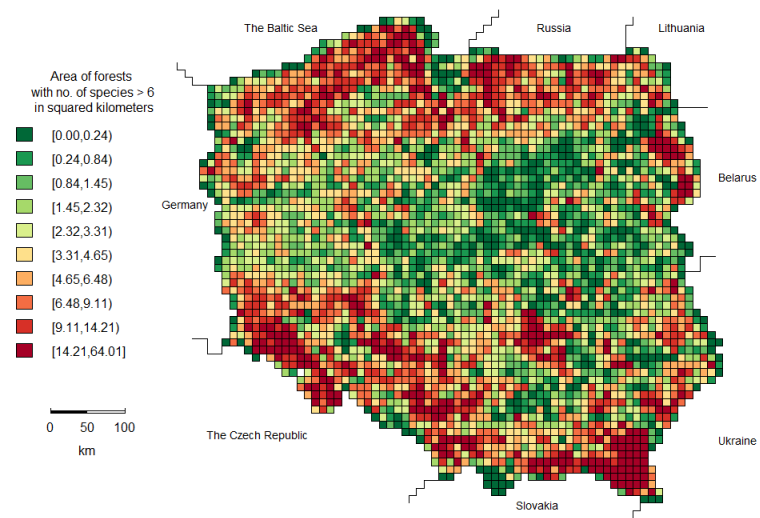
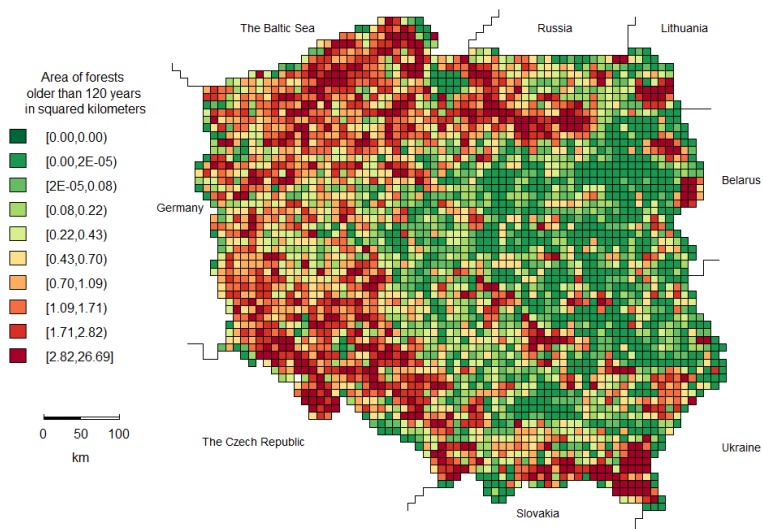
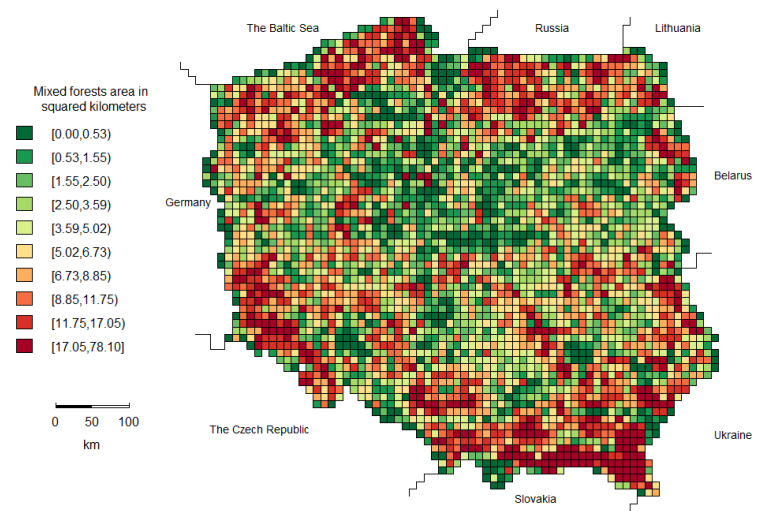






Appendix B. Distribution of the environmental characteristics of forests in Poland







FACULTY OF ECONOMIC SCIENCES
UNIVERSITY OF WARSAW
44/50 DŁUGA ST.
00-241 WARSAW
WWW.WNE.UW.EDU.PL