Ecological Indicators 121 (2021) 107020

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Evaluating Google Street View for tracking invasive alien plants along roads

Dorota Kotowska [a,b,*], Tomas Pärt [b], Michał Żmihorski [c]

[a] Institute of Nature Conservation, Polish Academy of Sciences, Mickiewicza 33, 31-120 Kraków, Poland
[b] Department of Ecology, Swedish University of Agricultural Sciences, Box 7044, SE 750 07 Uppsala, Sweden
[c] Mammal Research Institute, Polish Academy of Sciences, Stoczek 1, 17-230 Białowieża, Poland

ARTICLE INFO

Keywords:
Agricultural abandonment
Farmland
Goldenrods
Invasive alien plants
Monitoring
Remote data collection
Road ecology
Roadside vegetation

ABSTRACT

Invasive alien plants are considered a major driver of global biodiversity loss. Therefore, there is a huge demand of spatial and temporal data on their distribution for investigating possible drivers of species invasions and for predictions of future distributions. We use Google Street View imagery (GSV) as a new source of spatial and temporal data. GSV provides millions of panoramic views along road networks worldwide allowing for the identification of many plant species, including invasive ones. Thus, GSV has a great potential to support ecological research in documenting species distribution, but reliable validation of its precision and accuracy is lacking. Here, we describe and evaluate an approach using GSV to visually track the spread of invasive alien plants, the North American goldenrods (*Solidago canadensis* and *S. gigantea*) occurring abundantly along road network in Poland (Central Europe). We determined presence/absence of the species along 160 randomly selected transects of a length of 500 m by visual inspection of GSV images and compared it with field surveys at the same transects. We show that the occurrence of goldenrods in GSV is a reliable predictor of their occurrence in the wild. Sampling parameters, like road width, season when GSV pictures were taken and number of months elapsed since taking the GSV pictures, did not change the correlation between outputs of the two methods (GSV and field sampling). Furthermore, both the occurrence of goldenrods observed in the field and their occurrence in GSV have similar relations to habitat characteristics investigated (the same direction of relationship and similar effect size). We suggest Google Street View images may be an additional tool to be used in the detection and tracking of the spread of invasive alien plants along roadsides. The approach may be useful in assessing temporal changes in roadside vegetation and managing problematic plant species across large spatial scales and may contribute to the further development of more efficient sampling methods in ecological studies.

1. Introduction

Global civilization changes taking place over the last centuries have brought an intensive development of international trade, transport and tourism (Mascie-Taylor and Krzyżanowska, 2017). These changes have not only resulted in an increased human mobility, but also enabled unintentional or intentional introductions of many plant species outside their natural ranges (Lenda et al., 2014; Lockwood et al., 2005). Once established, some of these alien plants have become invasive posing serious ecological problems to the native fauna and flora (Vitousek et al., 1997). Invasions of many species of alien plants have been identified as a major and growing driver of global biodiversity loss. They may inflict significant damages to native ecosystems through excessive use of resources, disruption of ecological processes and habitat modification (Richardson et al., 2000) thus, negatively affecting richness, diversity and composition of native communities (Hejda et al., 2009), leading to the extinction of vulnerable indigenous species and homogenization of plant communities (Schwartz et al., 2006; Wilcove et al., 1998). Moreover, some species of invasive alien plants also may have an adverse impact on economy (e.g. by substantial production losses in agriculture or forestry) and human health (e.g. by causing allergies, including dermatitis, or accumulation and transferring toxins to human food; Neill and Arim, 2011; Pyšek and Richardson, 2010). It has been estimated that annual economic damages caused by invasive alien species in European Union are as high as 12 billion EUR and these costs are expected to rise (Shine et al., 2010). Therefore, a considerable legislative effort is currently being implemented to minimize the spread and negative impact of invasive plants on economy and environment (e.g. EU Regulation on Invasive Alien Species). Furthermore, numerous conservation initiatives are being taken to tackle threats from invasive alien plants to

* Corresponding author at: Institute of Nature Conservation, Polish Academy of Sciences, Mickiewicza 33, 31-120 Kraków, Poland.
E-mail address: kotowska@iop.krakow.pl (D. Kotowska).

https://doi.org/10.1016/j.ecolind.2020.107020
Received 26 May 2020; Received in revised form 29 August 2020; Accepted 28 September 2020
Available online 10 October 2020

biodiversity (e.g. projects co-funded by LIFE – the EU financial instrument supporting nature conservation).

Given the growing problem of invasive alien plants, there is an increasing demand of spatial data on invasive species for the purpose of monitoring and identifying drivers of species invasions. Such detailed spatial information are usually constrained to small scale inventories due to laborious and time-consuming surveys (but see atlas projects such as the Atlas Florae Europaeae; Jalas and Suominen, 1972–1996). Several studies used aerial photography and satellite remote sensing techniques to track alien plants (e.g. Müllerová et al., 2013) but usefulness of these methods is often limited because of the considerable financial costs of gathering high resolution images. However, recently, with the introduction of Google Street View (GSV) imagery, featured in Google Maps and Google Earth, a new source of georeferenced, open-access data has become available (Anguelov et al., 2010). GSV technology provides millions of vertical, panoramic views (i.e. photographic pictures) along the road network worldwide covering both urban and rural areas. It includes high quality images of the surroundings which allows for the identification of many plant species and habitats (see Hardion et al., 2016), as well as other structural or social features of neighborhood, such as buildings, sidewalks, road signs, aesthetics/disorder (e.g. graffiti) and pedestrian activity. Therefore, using GSV for virtual streetscape audits is an increasingly popular method of characterizing environment in the vicinity of roads for the purpose of urban planning or human health research (e.g. Li et al., 2015; Rundle et al., 2011; Steinmetz-Wood et al., 2019). GSV has also the potential to greatly support ecological studies in documenting species distribution. However, the possibility of its application in this field have been poorly evaluated so far. To date only few studies worldwide have used this source of data for identifying vulture habitats (Olea and Mateo-Tomás, 2013), monitoring the prevalence of the pine moth (Rousselet et al., 2013), and determining plant distribution: Russian olive (Collette and Pither, 2015), giant cane (Hardion et al., 2016), Persian hogweed (Meier et al., 2017), pampas grass (Pardo-Primoy and Fagúndez, 2019) and eucalypt (Queirós et al., 2020). Rousselet et al. (2013) and Deus et al. (2016) tested GSV as an alternative method to car surveys. However, reliable evaluation of precision and accuracy of GSV data needs to be investigated by comparisons to data collected in the field, but at present such validations are lacking.

Several studies have shown that one of the main corridors allowing invasive species to spread throughout different regions and environments are roadsides (e.g. Christen and Matlack, 2009; Pauchard and Alaback, 2004). Roadsides are linear habitats that dissect landscape interior and generate disturbances related with traffic (light, noise, pollution with oil and salt) and management activities (e.g. regular mowing of roadside vegetation), and thus can mediate biological processes including dispersal through vehicle tires or air flow (Forman, 2003; Rew et al., 2018; Speziale et al., 2018). Presence of these dispersal vectors and linear character of the road network that connects isolated populations make roadsides crucial objects facilitating expansion of invasive alien plants (Gelbard and Belnap, 2003; Ibisch et al., 2016). Thus, GSV imagery is a promising source of data covering dispersal pathways of many plant species and as such should cover highly relevant data to detect expanding populations of invasive species.

In this study we describe and evaluate a novel approach using GSV images to perform a large-scale inventory of two problematic plant species considered invasive in Eurasia: Canadian goldenrod (*Solidago canadensis*) and giant goldenrod (*S. gigantea*). These are highly competitive perennial herbs originated from North America which have spread across Europe and Asia as a result of intentional introduction for ornamental purposes. They have become one of the most successful invasive species in this region (Weber, 2001) due to their capacity for vigorous growth, rapid propagation by rhizomes, producing large number of small seeds spread by wind for long distances and because of an exertion of allelopathic effects on other plants. Consequently, the two goldenrod species may form dense stands outcompeting the native plants (Lenda et al., 2019) and may have a negative impact on native pollinators (Fenesi et al., 2015; Moroń et al., 2009), ants (Kajzer-Bonk et al., 2016; Lenda et al., 2013) and birds (Skórka et al., 2010). In their alien range invasive goldenrods are especially abundant in disturbed ruderal environments such as roadsides, riverbanks as well as in agricultural fields, mostly abandoned fields or meadows (Kabuce and Priede, 2010; Weber, 2017).

Here, we aim to validate the use of GSV by comparing data collected by virtual transect sampling using visual inspection of the vegetation on GSV images with corresponding transect data sampled in the field. We predict that these two datasets are positively correlated, and thus hypothesize that a GSV-based method properly identifies occurrences of the studied species along roads. As we expect that the reliability of GSV approach may be dependent on spatial scale, we used data on presence/absence of invasive goldenrods collected along: (1) c. 500 m-long transects and (2) c. 20 m-long sections of these transects. Moreover, we collected data on sampling parameters, such as time elapsed between the field survey and taking the GSV picture, width of the road and presence of road verge mowing, as we hypothesize that the degree of similarity between the two methods may be dependent on these variables (e.g. we expect that the larger the time lag between field survey and date of taking GSV picture, the higher dissimilarity and worse prediction of the goldenrod occurrence). Finally, we tested the usefulness of GSV data in addressing ecological questions. For this purpose, we compared whether GSV data and field survey data produced similar relationships to relevant environmental variables, in this case proportion of uncultivated areas (mainly abandoned fields and grasslands). We chose this variable because previous research found that uncultivated land is a main habitat of the goldenrods and may be their invasion pool (Lenda et al., 2019; Skórka et al., 2007). Thus, we expect positive association between the goldenrod occurrence and cover of uncultivated land in the vicinity of transects, and that this association (the effect size) is similar between the two methods.

2. Materials and methods

2.1. Study area and transect selection

The study was conducted in agricultural areas of Polish lowlands where the two goldenrod species are widespread and still expand (Tokarska-Guzik et al., 2012). In this area we randomly selected 40 districts (average size: 1,030 km^2) and chose all the 160 non-urban communes located within their boundaries. In each commune we randomly selected a point placed along road network (Fig. 1) using GIS tools and Open Street Map vector data in ArcGIS 10.4 software. The selected point was used to locate the beginning of a transect of about 500-m length. Each transect was subsequently divided into sections (see the following sub-chapter for details). If the transect was not covered by Google Street View imagery available in Google Maps web mapping service (as found on 10.05.2017) or intersected patches of forests, water bodies or urban areas (delineated basing on CORINE Land Cover database), it was rejected and the next randomly selected transect was used instead. The transect was also replaced by another one when it run along an unpaved road or was fenced with acoustic barriers. In total we selected 160 transects (Fig. 1) located along roads of different types in variable agricultural landscapes (both heterogeneous with a mosaic of small extensively managed fields, semi-natural and natural open habitats, forest patches and wastelands, and homogenous ones, i.e. with large fields intensively managed for crop production or large intensively grazed pastures). The spatial data processing was made with the use of ArcGIS 10.4 software.

2.2. Goldenrod survey using Google Street View

The selected 160 transects were remotely surveyed in Google Maps application with the use of GSV images taken between 2011 and 2014.

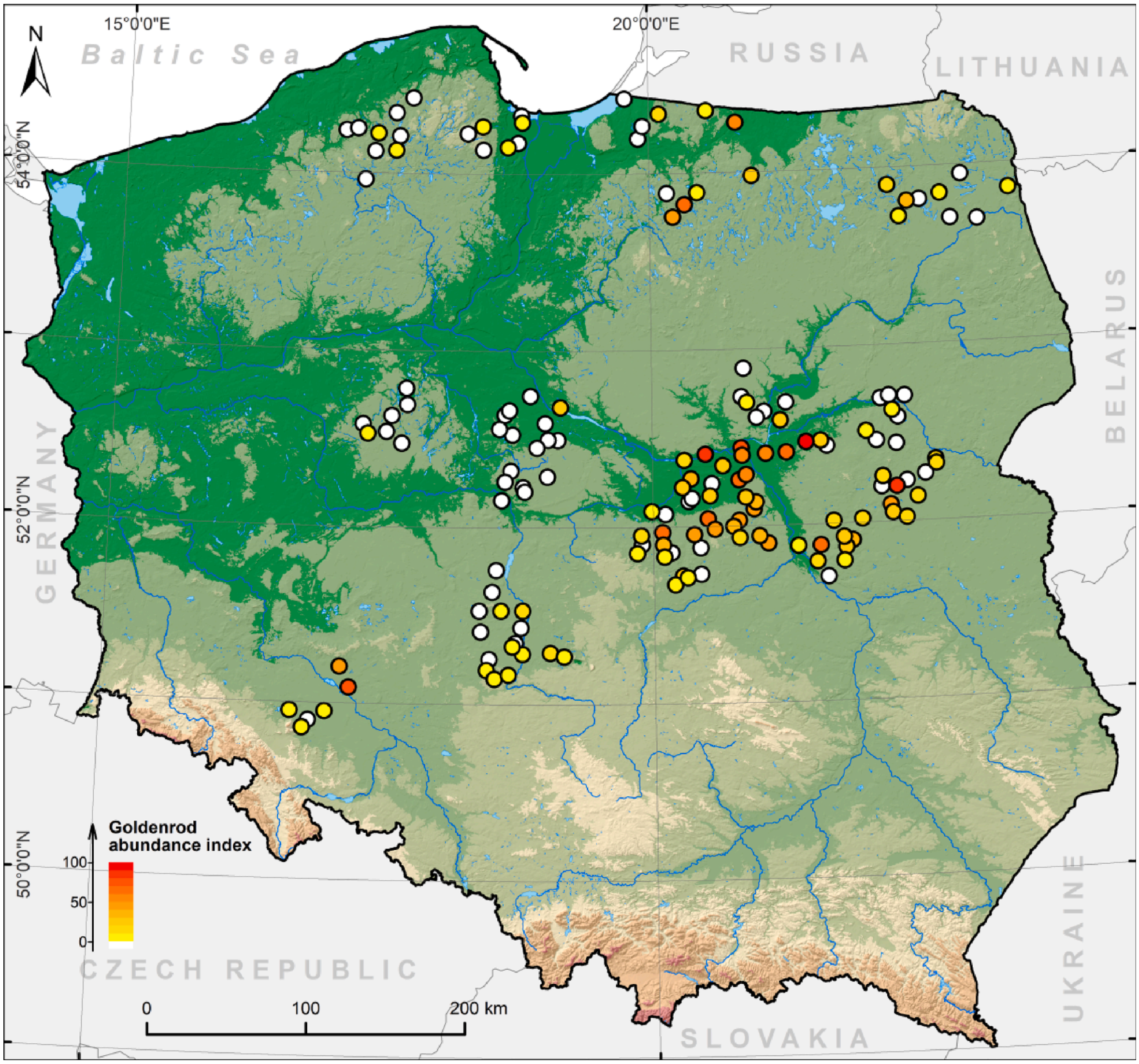

Fig. 1. Distribution of 160 transects surveyed for goldenrods in Poland and the goldenrod abundance index, i.e. share of transect sections (each about 20 m, c. 25 sections/transect) invaded by the goldenrods.

The vegetation was visually analyzed by virtual driving along the transects. As the GSV dataset consists of 360-degree images distributed unevenly along the Google vehicle's route, to keep the sampling effort equal among transects we considered two subsequent pictures located at a given transect line as a single transect section. The length of transect sections ranged between 10 and 61 m (at different roads images were taken at different distance intervals) with an average of 22 m and the mean number of sections per transect of 500 m length was 23 (range: 19–31). Using visual identification, for each transect section we manually determined the occurrence (i.e. presence/absence) of the two invasive goldenrods within 30 m (distance visually estimated) on each side of the road. Then, we used these data to obtain the occurrence of goldenrods at the transect level. As the surveyed species are superficially similar and often co-occur, we considered them together in the study (hereafter termed "goldenrods"). They were easily distinguished from other species by their characteristic shape, size and contrasting color of flowers (i.e. up to 2 m-tall aboveground shoots with numerous alternate single leaves on the stem, yellow inflorescences forming pyramidal panicles, often growing in clumps or dense stands). We also noted the date at which GSV pictures were taken and the coordinates of the picture using the 'googleway' (Cooley, 2018) and 'httr' (Wickham, 2018) packages in R (R Core Team, 2018). Since road managers in Poland are obligated to maintain the roadside vegetation in a way that improves safety of road users, the road verges are mown ensuring proper visibility along the roadway. The vegetation is often cut twice a season (in June and August). However, the mowing frequency and time may vary depending on local conditions, type of road and its localization (urban vs. rural areas). Therefore, during the virtual data collection we also assessed the presence of road verge mowing (yes or no) within 30 m of each transect section on both sides of the road. The 30 m zone of the transect section line for a given side of the road was considered mowed if the visual inspection of the GSV picture indicated that the vegetation covering more than 50% of this area had been previously cut and was not yet fully re-grown in height.

2.3. Goldenrod survey in the field

To validate the GSV method we performed field sampling along the same set of 160 transects (separately for each transect section) previously used for remote data collection (GPS receiver was used to localize GSV sections in the field so potential mismatch should not exceed 5 m). The fieldwork was conducted during the vegetation season of 2017. Each transect was visited once, between July and September (i.e. the peak period of flowering). The observer walked along the road at a constant pace of about 2 km per hour and noted the presence/absence of the goldenrods and whether mowing was applied within 30 m on both sides of the road for each transect section. The vegetation was classified as mowed using the same approach as in the GSV method. Analogous to the GSV method, the data on goldenrod occurrence along transect sections were subsequently used to determine the species occurrence at the scale of transects.

2.4. Transect characteristics

For each section of the transect we noted three characteristics. First, we calculated the share of uncultivated open area (i.e. abandoned arable land and grasslands) within 30 m on each side of the transect section using GIS tools based on analysis of freely available historical satellite imagery obtained from Google Earth. It was calculated both for the time when the GSV pictures were taken and for the time when the fieldwork was conducted. Second, based on the satellite images, we determined the width of the road as an average of three measurements taken in the start, middle and end point of the transect line (as it was not constant, however in all cases the difference between the three measurements was not larger than 1 m). Third, we calculated the length of the transect sections and the number of transect sections per transects using geographic coordinates of the GSV images. All the calculations were computed in ArcGIS 10.4 software.

2.5. Statistical analyses

To evaluate an approach using GSV for detection of invasive alien plants along roads we used generalized additive models (GAM), generalized linear mixed models (GLMM) and general linear models (GLM) implemented in 'mgcv' and 'lme4′ packages (Bates et al., 2015; Wood, 2017) in R (R Core Team, 2018). We performed three types of analyses.

First, we tested whether GSV data predicts presence/absence of goldenrods in the field. For this purpose we performed GLMs with binomial error distribution and logit link in which the occurrence of goldenrods based on the field survey was a response variable (1/0) while the occurrence of goldenrods observed in GSV (termed "GSVSol", see Table 1 for list of variables) was an explanatory variable (1/0). As the accuracy of GSV method may depend on the spatial resolution of sampling units we fitted two models: using transect sections ($GLM1_{ALL}$) and transects ($GLM2_{ALL}$) as single data records. In all cases each side of the transect or transect section was treated independently, thus 160 transects resulted in 320 data records for the transect scale and 7426 for the transect section scale analyses. Moreover, as grass cutting on road verges may weaken the correlation between results of field survey and GSV, we repeated the two models only for transect sections without evidence of mowing both in GSV and field data ($GLM1_{UNMOWED}$, $GLM2_{UNMOWED}$, respectively). Additionally, in $GLM2_{ALL}$ and $GLM2_{UNMOWED}$ the number of transect sections per transect was used as a covariate (NSections; continuous variable).

We evaluated the performance of our four above GLM models by using "leave-one-out" cross validation approach (LOOCV). For each dataset we first selected all sampling units where the goldenrods were recorded in the field and the same number of random sampling units where they were absent, to keep presence to absence ratio equal (this is necessary to keep expected classification error as 50%). Among the selected data subsets a single observation *n* was excluded and used for validation, while remaining observations were used for model fit. Basing on this model, a prediction was made for the excluded observation *n*. The procedure was repeated for all data records in a given data subset. The ratio of number of correct predictions to the total number of predictions is an approximately unbiased estimate for the model classification accuracy (James et al., 2013). Moreover, we calculated two other model performance measures: sensitivity (proportion of sites correctly classified by the model as occupied by goldenrods) and specificity (proportion of sites correctly classified by the model as unoccupied by goldenrods). The described process (starting from the random selection of goldenrod-free sampling units) was replicated 10 times to include different sets of random sampling units in the validation and results of 10 replications were averaged.

Second, we aimed to determine factors that influence the correct classification of a transect or transect section as occupied or unoccupied by goldenrods based on GSV pictures (i.e. drivers of similarities and dissimilarities between outputs of GSV method and field survey). As mowing will affect these results for obvious reasons, we included in the analysis only the sections that were not mowed both in GSV and field data. To account for spatial autocorrelation in observations we fitted two binomial GAMs: for the transect section scale (GAM1) and the scale of transect (GAM2; note that transect length now varied among transects because of the removal of mowed sections). In both models the agreement between data on goldenrod occurrence obtained using the two methods was used as a response variable (1 – agreement, 0 – disagreement). The outputs of GSV method and field survey within a given sampling unit were considered as agreement if both methods detected the presence or both reviled the absence of goldenrods. The other cases were referred to as disagreement. In GAM1 one fixed categorical variable (GSVSeason) and three continuous variables (Length, WdthRoad, MonthSinceGSV) were explanatory variables. To account for spatial autocorrelation among adjacent transect sections, the transect section number (TransectSectionNo) was fitted with a spline with number of degrees of freedom set to 4. Moreover, the transect identity (TransectID) was introduced as a random factor. An analogous model using the same set of predictors was performed for the scale of transects (GAM2; with NSections instead of Length; see Table 1 for description).

Third, we compared estimates of the effect of uncultivated land for the goldenrod occurrence using datasets obtained with GSV and field survey (in both cases mowed sections were included) to test the usefulness of GSV method for predicting occurrence by environmental variables in the surrounding landscape. We computed two generalized linear mixed models ($GLMM1_{GSV}$, $GLMM1_{FIELD}$) for the scale of transect sections and two generalized linear models ($GLM3_{GSV}$, $GLM3_{FIELD}$) for the scale of transects. In all four models we used binomial error distribution with logit link function, goldenrod occurrence as a response

Table 1
Description of explanatory variables used in the models.

#	Variable	Description	Model
1	GSVSol	Categorical. Presence (yes/no) of goldenrods detected by GSV method.	$GLM1_{ALL}$, $GLM1_{UNMOWED}$, $GLM2_{ALL}$, $GLM2_{UNMOWED}$
2	NSections	Continuous. Number of transect sections established along a given transect	$GLM2_{ALL}$, $GLM2_{UNMOWED}$, GAM2
3	GSVSeason	Categorical. Season of the year when the GSV pictures were taken: spring (May-June), summer (July-August), fall (September-October).	GAM1, GAM2
4	Length	Continuous. Length of the transect section in meters.	GAM1
5	WdthRoad	Continuous. Average width of the road in meters, computed using three measurements taken in the starting, middle and ending point of the transect line based on the satellite images.	GAM1, GAM2
6	MonthSinceGSV	Continuous. Number of months elapsing the GSV pictures and the field survey.	GAM1, GAM2
7	Uncultivated	Continuous. Share of uncultivated open area (abandoned arable land and grasslands) in the area of a given sampling unit.	$GLMM1_{GSV}$, $GLMM1_{FIELD}$, $GLM3_{GSV}$, $GLM3_{FIELD}$

variable (1 – present, 0 – absent) and share of uncultivated land in the area of a given sampling unit (Uncultivated; continuous) as an explanatory variable. Moreover, in $GLMM1_{GSV}$ and $GLMM1_{FIELD}$ the transect ID was included as a random factor. As a result, we were able to compare parameter estimates of the uncultivated land effect between models using different sources of goldenrod occurrence.

3. Results

3.1. Goldenrod occurrence based on Google Street View and field data

At the transect sections scale, the field survey revealed goldenrods in 1081 out of 7426 sections (i.e. 14.5%, Table 2) and within 738 out of 3486 sections without mowing (21.2%; i.e. including only the sections that were not mowed both in the field and GSV). Figures based on Google Street View were lower and equaled to 8.6% sections occupied by goldenrods (12.3% without mowing). At the transect level, we observed goldenrods in 47.5% of the transects surveyed in the field (17.8% without mowing, i.e. including only the unmowed sections both in GSV and field data) while the corresponding figures for GSV data were 35% and 11.3% respectively (Table 2).

For each of the four datasets considered (i.e. transect sections and transects, with and without mowing), the occurrence of goldenrods based on field survey was significantly positively correlated with the GSV occurrence (Table 3). The models using goldenrod occurrence based on GSV method correctly classified 72–85% of sites surveyed in the field (as found by cross-validation, Table 3). All the models better predicted actual presences (ca. 94–97% of observed presences were classified correctly) than absences (ca. 64–78% of observed absences were classified correctly; Table 3).

3.2. Factors explaining similarities between Google Street View and field data

Similarity of information on goldenrod occurrence derived from the two compared methods (GSV and field survey) was hardly explained by the sampling characteristics. The season of the year when the GSV pictures were captured, width of the road and number of months elapsed since taking the GSV pictures did not affect the similarity between outputs of two sampling methods, both at the scale of transect section and transect (Table 4). Of considered explanatory variables, only the length of a transect section was negatively correlated with the probability of correct classification using GSV: the longer the transect section, the less similar are results of GSV method and field survey (Table 4). However, we did not find such association for the number of sections (i.e. reflecting average section length) at the scale of transects (Table 4).

3.3. Comparison of models using Google Street View and field data

The share of uncultivated land in the vicinity of transect section was a significant positive predictor of goldenrod occurrence as found in the field ($GLMM1_{FIELD}$, Uncultivated effect: estimate = 0.70 (SE = 0.05), $p < 0.001$) and GSV ($GLMM1_{GSV}$, Uncultivated effect: estimate = 0.70 (SE = 0.04), $p < 0.001$). Similarly, the share of uncultivated areas positively predicted goldenrod occurrence in the transect scale for both field-based data and GSV data ($GLM3_{FIELD}$, Uncultivated effect: estimate = 0.96 (SE = 0.16), $p < 0.001$, $GLM3_{GSV}$, Uncultivated effect: estimate = 1.18 (SE = 0.21), $p < 0.001$, respectively; Fig. 2).

Table 2
Contingency table showing number of sampling units with and without goldenrod presence records based on GSV and field data in two spatial scales (transect sections and transects) and for two datasets (including all sections and only the sections that were not mowed both in the field and GSV).

	All sections		Unmowed sections	
	Field presence	Field absence	Field presence	Field absence
	Scale: transect section			
GSV presence	497	144	381	70
GSV absence	584	6201	357	2678
	Scale: transect			
GSV presence	110	4	21	5
GSV absence	42	164	20	184

Table 3
Summary of GLMs (parameter estimates followed by SE in parentheses) explaining the occurrence of goldenrods in the field survey in relation to goldenrod occurrence based on GSV images (GSVSol) along roads. Separate models were fitted for transect sections and transects as well as for full dataset and a subset of data without mowing. Significant effects are marked in bold. Significance levels (p-values) are indicated by asterisks, and are explained below the table. Performance of the models based on cross-validation (LOOCV) is given at the bottom.

Scale: Predictors	**transect section** $GLM1_{ALL}$ n = 7426	$GLM1_{UNMOWED}$ n = 3486	**transect** $GLM2_{ALL}$ n = 320	$GLM2_{UNMOWED}$ n = 230
Intercept	**−2.36 (0.04)*** **	**−2.02 (0.06)*** **	−1.59 (1.40)	**−2.10 (0.34)*** **
GSVSol: yes	**3.60 (0.10)*** **	**3.71 (0.14)*** **	**4.68 (0.54)*** **	**3.70 (0.56)*** **
NSections	not included	not included	0.01 (0.06)	−0.02 (0.04)
$LOOCV_{ACCURACY}$	71.8%	74.4%	85.1%	74.0%
$LOOCV_{SENSITIVITY}$	95.1%	94.8%	96.9%	94.4%
$LOOCV_{SPECIFICITY}$	64.4%	66.8%	78.0%	66.5%

Statistical significance: *** < 0.001.

Table 4
Summary of GAMs (parameter estimates followed by SE in parentheses) explaining similarity between outputs of GSV method and field sampling at the scale of transect sections and transects in relation to sampling parameters. Significant effects are marked in bold. Significance levels (p-values) are indicated by asterisks, and are explained below the table.

Scale:	**transect section**	**transect**
Predictors	GAM1 n = 3486	GAM2 n = 230
Intercept	−0.10 (3.24)	7.09 (5.63)
GSVSeason: fall	0.22 (0.84)	−1.28 (1.27)
GSVSeason: spring	−1.20 (1.19)	−2.10 (1.46)
Length	**−0.31 (0.08)*** **	not included
NSections	not included	−0.01 (0.07)
WdthRoad	−0.32 (0.36)	−0.33 (0.49)
MonthSinceGSV	0.06 (0.06)	−0.04 (0.10)

Statistical significance: *** < 0.001.

4. Discussion

We show that Google Street View can be used as an additional tool for surveying plant species at road verges and their immediate surroundings. As many invasive alien plants are occurring and dispersing along roadsides, GSV can be an important, effective tool for the future tracking of the spread of these species. More specifically, we showed that occurrence of goldenrods detected using GSV predicted their occurrence as observed in the field 3–5 years after GSV images were taken and this was true for both spatial scales considered (i.e. transects and transect sections). The GSV method performed especially well in predicting actual goldenrod presences. Sampling parameters, like presence of road verge mowing, road width, season when GSV pictures were taken and number of months elapsed since taking the GSV pictures, did not change the correlation between the two methods (except for transect section length suggesting a negative effect, most likely due to the difference in sampling effort between the two methods at longer sections). Finally, models based on GSV or field survey data produced very similar

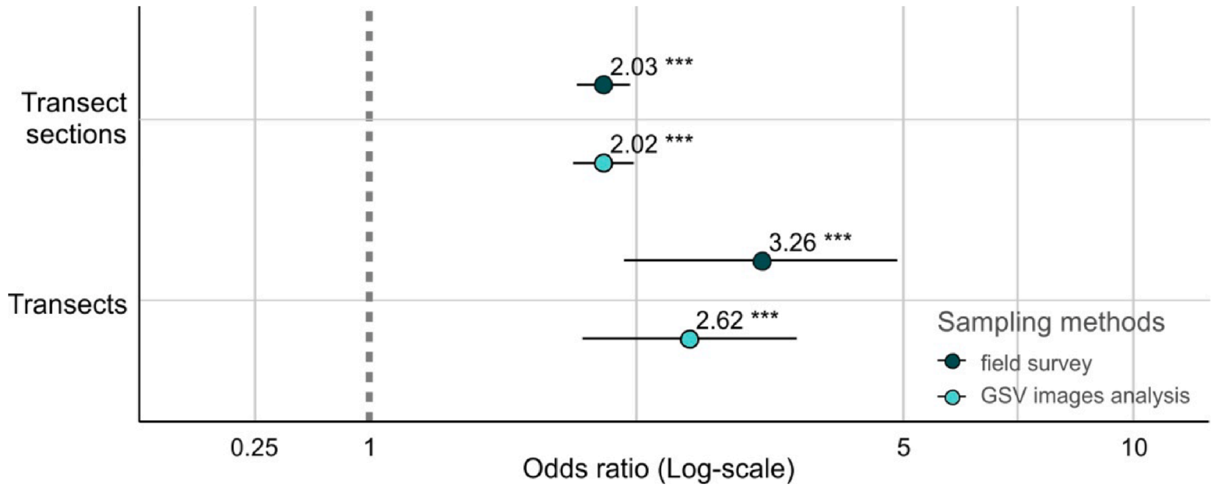

Fig. 2. Effect of share of uncultivated land on the occurrence of goldenrods (log odds with 95% CIs) from models using field survey and GSV image analysis in the scale of transect sections and transects.

estimates in the importance of the uncultivated land for the occurrence of goldenrods. Although these results are suggesting that GSV data appears to perform well in detecting and studying roadside vegetation, it may also have some limitations as discussed below.

GSV seems to be an effective tool for detecting species occurring in large numbers, being tall, having distinct shape or contrasting color. That is, species that are easily visually detected and easy to recognize at distance. For example, while browsing GSV for goldenrods, we noticed also the presence of several other invasive plant species: hogweeds (*Heracleum mantegazzianum* and *H. sosnowskyi*), knotweeds (*Reynoutria japonica*, *R. sachalinensis*, *R.* × *bohemica*), wild cucumber *Echinocystis lobata*, garden lupin *Lupinus polyphyllus*, box elder *Acer negundo*, black locust *Robinia pseudoacacia*, staghorn sumac *Rhus typhina* and black cherry *Prunus serotina.* However, as GSV was not intentionally designed for collecting plant data, some geometric distortions in the images could potentially affect both species detectability and mapping accuracy. Thus, smaller invasive species or species difficult to distinguish, like small-flower touch-me-not *Impatiens parviflora* or common beggar-ticks *Bidens frondosa*, commonly occurring along roadsides (Tokarska-Guzik et al., 2012), may be overlooked in GSV images.

Several factors may potentially weaken the observed high correlation between GSV and field data. First, some challenges may be linked to the frequency of GSV updates. The images used in our study were taken at different times (from morning hours until evening), months (from May to October), and years (2011 to 2014). Nevertheless, we did not find any effect of season on the probability of correct presence/absence detection of goldenrods, most likely because the studied species are relatively easily distinguishable all year round. However, in case of other species seasonal changes in vegetation growth will be more important. Furthermore, the differences between GSV and field data can also be driven by colonization and extinction events in the period between capturing GSV images and field survey, especially if the period extends to several years. Comparison of data from GSV and field survey (conducted 3–5 years after capturing GSV images) suggests an increase in goldenrod occurrence (see Table 2). However, with such an assumption one should expect a negative relationship between time elapsed since the GSV picture was taken and similarity between methods, which was not confirmed in our models (see Table 4).

Second, the probability of visual plant detection may be reduced by road verge mowing, and thus may substantially limit usefulness of GSV as a source of data. However, given there is some variation in the practice of road verge mowing, when using GSV one may also identify where mowing occurs and thus where the risk of establishment of invasive species is lower. Furthermore, mowing also reduces detection of plants during field surveys, so this limitation is not exclusive for GSV method. Fortunately, it is relatively easy to detect recent mowing and account observed species occurrences for the mowing effect.

Finally, as highlighted in previous research (Rousselet et al., 2013), the image availability may be a crucial limitation for GSV-based sampling. GSV was primarily introduced in 2007 and covered major cities of North America. Since then it has been developed to include urban, suburban and rural areas from all around the world. Until now it has collected 170 billion Street View images captured along more than 16 million kilometers of roads across more than 220 countries and territories (Google, 2020). The GSV coverage is being successively enhanced, however there are still many places where GSV data is unavailable or its availability is limited (e.g. Africa, Central America, Middle East). Also, small gravel roads are excluded although they may be important habitats for many invasive plant species. Thus, the GSV method should be applied with caution as it may be biased by omission of some important areas.

Nevertheless, the great advantage of GSV method is that it appears to be more time- and cost effective and has much lower carbon footprint than collecting data in a traditional manner. Field sampling of vegetation usually is highly laborious (Hill et al., 2005). Given the large spatial scale of our survey, it required significant amounts of travel time. Hence, we managed to visit in the field on average six transects per day, while using GSV we were able to virtually sample the same area within an hour. We estimated that during the fieldwork an observer travelled by car a distance of about 7 700 km generating costs equaling 2 000 EUR and releasing 1.29 tons of carbon dioxide to the atmosphere. Sampling with GSV costed about 120 EUR and the estimated emission of carbon dioxide during virtual driving along transects was about 0.14 tons (see Appendix A for detailed calculations). Another future benefit of GSV data is that new GSV pictures are planned to be taken every several years, thus opening up for investigations of distributional shifts in plant species associated with climate and environmental change.

5. Conclusions

Being aware of the limitations discussed above, we conclude that Google Street View imagery, publicly available for substantial proportion of roads worldwide, is a valuable source of data on species distribution patterns. As the GSV-based method allows for considerable sampling effort reduction, it provides an opportunity to investigate some ecological phenomena (e.g. plant invasions) across large spatial scales with relatively low costs. Moreover, since the library of GSV images is permanently being updated, the tool has the potential to be used for assessing temporal changes in roadside vegetation. We emphasize that utilization of GSV data for studying roadside environments (e.g. determining species distribution) should be further developed to include machine learning techniques for a fast identification of species and their occurrences. This would enable automatic detection of some objects (e.g. plant species) to open up for large scale analyses on the spread of invasive plant species across whole continents in order to identify new ways of how to manage these species in the future.

CRediT authorship contribution statement

Dorota Kotowska: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Visualization, Project administration, Funding acquisition. **Tomas Pärt:** Conceptualization, Writing - review & editing, Supervision. **Michał Żmihorski:** Conceptualization, Methodology, Software, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Piotr Skórka for helpful comments on the manuscript. The study was supported by the Polish National Science Centre grant: 2016/21/N/NZ8/01286 and the Polish National Science Centre doctoral scholarship: 2019/32/T/NZ8/00343.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2020.107020.

References

Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., Weaver, J., 2010. Google street view: Capturing the world at street level. Computer (Long. Beach. Calif) 43, 32–38. https://doi.org/10.1109/MC.2010.170.

Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67.

Christen, D.C., Matlack, G.R., 2009. The habitat and conduit functions of roads in the spread of three invasive plant species. Biol. Invasions 11, 453–465. https://doi.org/10.1007/s10530-008-9262-x.

Collette, L.K.D., Pither, J., 2015. Modeling the potential North American distribution of Russian olive, an invader of riparian ecosystems. Plant Ecol. 216, 1371–1383. https://doi.org/10.1007/s11258-015-0514-4.

Cooley, D., 2018. googleway: Accesses Google Maps APIs to Retrieve Data and Plot Maps. R package version 2 (7), 1. https://CRAN.R-project.org/package=googleway.

Deus, E., Silva, J.S., Catry, F.X., Rocha, M., Moreira, F., 2016. Google Street View as an alternative method to car surveys in large-scale vegetation assessments. Environ. Monit. Assess. 188 https://doi.org/10.1007/s10661-016-5555-1.

Fenesi, A., Vágási, C.I., Beldean, M., Földesi, R., Kolcsár, L.-P., Shapiro, J.T., Török, E., Kovács-Hostyánszki, A., 2015. Solidago canadensis impacts on native plant and pollinator communities in different-aged old fields. Basic Appl. Ecol. 16, 335–346. https://doi.org/10.1016/J.BAAE.2015.03.003.

Forman, R.T.T., 2003. Road ecology : science and solutions. Island Press.

Gelbard, J.L., Belnap, J., 2003. Roads as Conduits for Exotic Plant Invasions in a Semiarid Landscape. Conserv. Biol. 17, 420–432. https://doi.org/10.1046/j.1523-1739.2003.01408.x.

Google, 2020. Google Maps Blog. https://www.blog.google/products/maps/ (accessed 1.20.20).

Hardion, L., Leriche, A., Schwoertzig, E., Millon, A., 2016. Species Distribution 2.0: An Accurate Time- and Cost-Effective Method of Prospection Using Street View Imagery. PLoS One 11, e0146899. https://doi.org/10.1371/journal.pone.0146899.

Hejda, M., Pyšek, P., Jarošík, V., 2009. Impact of invasive plants on the species richness, diversity and composition of invaded communities. J. Ecol. 97, 393–403. https://doi.org/10.1111/j.1365-2745.2009.01480.x.

Hill, D., Fasham, M., Tucker, G., Shewry, M., Shaw, P. (Eds.), 2005. Handbook of Biodiversity Methods: Survey, Evaluation and Monitoring, Handbook of Biodiversity Methods. Cambridge University Press. https://doi.org/10.1017/cbo9780511542084.

Ibisch, P.L., Hoffmann, M.T., Kreft, S., Pe'Er, G., Kati, V., Biber-Freudenberger, L., DellaSala, D.A., Vale, M.M., Hobson, P.R., Selva, N., 2016. A global map of roadless areas and their conservation status. Science. 354, 1423–1427. https://doi.org/10.1126/science.aaf7166.

Jalas, J., Suominen, J. (Eds.), 1972-1996. Atlas Florae Europaeae. The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning with applications in R. Springer, New York. https://doi.org/https://doi.org/10.1007/978-1-4614-7138-7.

Kabuce, N., Priede, N., 2010. NOBANIS – Invasive Alien Species Fact Sheet – Solidago canadensis. – From: Online Database of the European Network on Invasive Alien Species - NOBANIS. https://www.nobanis.org (accessed 1.22.20).

Kajzer-Bonk, J., Szpiłyk, D., Woyciechowski, M., 2016. Invasive goldenrods affect abundance and diversity of grassland ant communities (Hymenoptera: Formicidae). J. Insect Conserv. 20, 99–105. https://doi.org/10.1007/s10841-016-9843-4.

Lenda, M., Skórka, P., Knops, J., Żmihorski, M., Gaj, R., Moroń, D., Woyciechowski, M., Tryjanowski, P., 2019. Multispecies invasion reduces the negative impact of single alien plant species on native flora. Divers. Distrib. 25, 951–962. https://doi.org/10.1111/ddi.12902.

Lenda, M., Skórka, P., Knops, J.M.H., Moroń, D., Sutherland, W.J., Kuszewska, K., Woyciechowski, M., 2014. Effect of the Internet Commerce on Dispersal Modes of Invasive Alien Species. PLoS One 9, e99786. https://doi.org/10.1371/journal.pone.0099786.

Lenda, M., Witek, M., Skórka, P., Moroń, D., Woyciechowski, M., 2013. Invasive alien plants affect grassland ant communities, colony size and foraging behaviour. Biol. Invasions 15, 2403–2414. https://doi.org/10.1007/s10530-013-0461-8.

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban For. Urban Green. 14, 675–685. https://doi.org/10.1016/j.ufug.2015.06.006.

Lockwood, J.L., Cassey, P., Blackburn, T., 2005. The role of propagule pressure in explaining species invasions. Trends Ecol. Evol. 20, 223–228. https://doi.org/10.1016/j.tree.2005.02.004.

Mascie-Taylor, C.G.N., Krzyżanowska, M., 2017. Biological aspects of human migration and mobility. Ann. Hum. Biol. 44, 427–440. https://doi.org/10.1080/03014460.2017.131344.

Meier, S., Taff, G.N., Aune, J.B., Eiter, S., 2017. Regulation of the Invasive Plant *Heracleum persicum* by Private Landowners in Tromsø. Norway. Invasive Plant Sci. Manag. 10, 166–179. https://doi.org/10.1017/inp.2017.11.

Moroń, D., Lenda, M., Skórka, P., Szentgyörgyi, H., Settele, J., Woyciechowski, M., 2009. Wild pollinator communities are negatively affected by invasion of alien goldenrods in grassland landscapes. Biol. Conserv. 142, 1322–1332. https://doi.org/10.1016/J.BIOCON.2008.12.036.

Müllerová, J., Pergl, J., Pyšek, P., 2013. Remote sensing as a tool for monitoring plant invasions: Testing the effects of data resolution and image classification approach on the detection of a model plant species Heracleum mantegazzianum (giant hogweed). Int. J. Appl. Earth Obs. Geoinf. 25, 55–65. https://doi.org/10.1016/J.JAG.2013.03.004.

Neill, P.E., Arim, M., 2011. Human Health Link to Invasive Species, in: Encyclopedia of Environmental Health. Elsevier, pp. 116–123. https://doi.org/10.1016/B978-0-444-52272-6.00528-6.

Olea, P.P., Mateo-Tomás, P., 2013. Assessing Species Habitat Using Google Street View: A Case Study of Cliff-Nesting Vultures. PLoS One 8, e54582. https://doi.org/10.1371/journal.pone.0054582.

Pardo-Primoy, D., Fagúndez, J., 2019. Assessment of the distribution and recent spread of the invasive grass Cortaderia selloana in Industrial Sites in Galicia, NW Spain. Flora Morphol. Distrib. Funct. Ecol. Plants. https://doi.org/10.1016/j.flora.2019.151465.

Pauchard, A., Alaback, P.B., 2004. Influence of Elevation, Land Use, and Landscape Context on Patterns of Alien Plant Invasions along Roadsides in Protected Areas of South-Central Chile. Conserv. Biol. 18, 238–248. https://doi.org/10.1111/j.1523-1739.2004.00300.x.

Pyšek, P., Richardson, D.M., 2010. Invasive Species, Environmental Change and Management, and Health. Annu. Rev. Environ. Resour. 35, 25–55. https://doi.org/10.1146/annurev-environ-033009-095548.

Queirós, L., Deus, E., Silva, J.S., Vicente, J., Ortiz, L., Fernandes, P.M., Castro-Díez, P., 2020. Assessing the drivers and the recruitment potential of Eucalyptus globulus in the Iberian Peninsula. For. Ecol. Manage. 466, 118147 https://doi.org/10.1016/j.foreco.2020.118147.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rew, L.J., Brummer, T.J., Pollnac, F.W., Larson, C.D., Taylor, K.T., Taper, M.L., Fleming, J.D., Balbach, H.E., 2018. Hitching a ride: Seed accrual rates on different types of vehicles. J. Environ. Manage. 206, 547–555. https://doi.org/10.1016/j.jenvman.2017.10.060.

Richardson, D.M., Pysek, P., Rejmanek, M., Barbour, M.G., Panetta, F.D., West, C.J., 2000. Naturalization and invasion of alien plants: concepts and definitions. Divers. Distrib. 6, 93–107. https://doi.org/10.1046/j.1472-4642.2000.00083.x.

Rousselet, J., Imbert, C.-E., Dekri, A., Garcia, J., Goussard, F., Vincent, B., Denux, O., Robinet, C., Dorkeld, F., Roques, A., Rossi, J.-P., 2013. Assessing Species Distribution Using Google Street View: A Pilot Study with the Pine Processionary Moth. PLoS One 8, e74918. https://doi.org/10.1371/journal.pone.0074918.

Rundle, A.G., Bader, M.D.M., Richards, C.A., Neckerman, K.M., Teitler, J.O., 2011. Using google street view to audit neighborhood environments. Am. J. Prev. Med. 40, 94–100. https://doi.org/10.1016/j.amepre.2010.09.034.

Schwartz, M.W., Thorne, J.H., Viers, J.H., 2006. Biotic homogenization of the California flora in urban and urbanizing regions. Biol. Conserv. 127, 282–291. https://doi.org/10.1016/J.BIOCON.2005.05.017.

Shine, C., Kettunen, M., Genovesi, P., Essl, F., Gollasch, S., Rabitsch, W., Scalera, R., Starfinger, U., ten Brink, P., 2010. Assessment to support continued development of the EU Strategy to combat invasive alien species. Final Report for the European Commission. Brussels, Belgium.

Skórka, P., Lenda, M., Tryjanowski, P., 2010. Invasive alien goldenrods negatively affect grassland bird communities in Eastern Europe. Biol. Conserv. 143, 856–861. https://doi.org/10.1016/J.BIOCON.2009.12.030.

Skórka, P., Settele, J., Woyciechowski, M., 2007. Effects of management cessation on grassland butterflies in southern Poland. Agric. Ecosyst. Environ. 121, 319–324. https://doi.org/10.1016/j.agee.2006.11.001.

Speziale, K.L., di Virgilio, A., Lescano, M.N., Pirk, G., Franzese, J., 2018. Synergy between roads and disturbance favour Bromus tectorum L. invasion. PeerJ 2018. https://doi.org/10.7717/peerj.5529.

Steinmetz-Wood, M., Velauthapillai, K., O'Brien, G., Ross, N.A., 2019. Assessing the micro-scale environment using Google Street View: The Virtual Systematic Tool for Evaluating Pedestrian Streetscapes (Virtual-STEPS). BMC Public Health 19. https://doi.org/10.1186/s12889-019-7460-3.

Tokarska-Guzik, B., Dajdok, Z., Zając, M., Zając, A., Urbisz, A., Danielewicz, W., Hołdyński, C., 2012. Rośliny obcego pochodzenia w Polsce ze szczególnym uwzględnieniem gatunków inwazyjnych - Alien plants in Poland with particular reference to invasive species. Generalna Dyrekcja Ochrony Środowiska, Warszawa.

Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Rejmanek, M., Westbrooks, R.G., 1997. Introduced species: A significant component of human-caused global change. N. Z. J. Ecol. 21, 1–16.

Weber, E., 2017. Invasive plant species of the world : a reference guide to environmental weeds. Cabi.

Weber, E., 2001. Current and Potential Ranges of Three Exotic Goldenrods (*Solidago*) in Europe. Conserv. Biol. 15, 122–128. https://doi.org/10.1111/j.1523-1739.2001.99424.x.

Wickham, H., 2018. httr: Tools for Working with URLs and HTTP. R package version 1.4.0. https://CRAN.R-project.org/package=httr.

Wilcove, D.S., Rothstein, D., Dubow, J., Phillips, A., Losos, E., 1998. Quantifying Threats to Imperiled Species in the United States. Bioscience 48, 607–615. https://doi.org/10.2307/1313420.

Wood, S.N., 2017. Generalized Additive Models : an introduction with R (2nd edition). Chapman and Hall/CRC.