Ecological Informatics 48 (2018) 281–289

Contents lists available at ScienceDirect

Ecological Informatics

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

Mapping uncertainty of ICP-Forest biodiversity data: From standard treatment of diffusion to density-equalizing cartograms

Marta Galluzzi[a,*], Duccio Rocchini[b,c,d], Roberto Canullo[e], Ronald E. McRoberts[f], Gherardo Chirici[a]

[a] University of Florence, Department of Agricultural and Forest Economics, Engineering, Sciences and Technologies, Via San Bonaventura 13, 50145 Florence, Italy
[b] University of Trento, Center Agriculture Food Environment, Via E. Mach 1, 38010, S. Michele all'Adige (TN), Italy
[c] University of Trento, Centre for Integrative Biology, Via Sommarive 14, 38123 Povo (TN), Italy
[d] Fondazione Edmund Mach, Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Via E. Mach 1, 38010, S. Michele all'Adige (TN), Italy
[e] University of Camerino, School of Biosciences and Veterinary Medicine, Plant Diversity and Ecosystems Management Unit, Camerino, MC, Italy
[f] Northern Research Station, U.S. Forest Service, 1992 Folwell Avenue, Saint Paul, MN 55108, USA

ARTICLE INFO

Keywords:
Density-equalizing cartograms
Uncertainty map
Plant species richness
European biodiversity inference

ABSTRACT

Data uncertainty due to spatial gaps and heterogeneity is a fundamental problem in conservation and environmental planning. Thus, investigation of issues related to data uncertainty contributes to more efficient conservation plans. We evaluated the uncertainty of data related to forest diversity descriptors using a diffusion-based cartogram approach that visually displays how data information change in function with respect to degree of uncertainty. We used ground vegetation data for 3093 plots collected as part of the BioSoil project through the ICP Forests Level I network and stored in the LI-BioDiv database. For each plot, we assigned an uncertainty value based on the survey season and the mean monthly temperature for the survey period. The density-equalizing map or cartogram highlights that data collected in Spain, the United Kingdom and the German federal states of Berlin and Brandenburg have smaller values of species richness corresponding to larger values of uncertainty. We found that an awareness of the negative relationship between the survey period and species richness can lead to improved data handling and analysis. We demonstrated that cartograms are efficient tools for evaluating and managing uncertainty and can strengthen the results of data analysis by providing alternative perspectives and interpretations of spatial phenomena.

1. Introduction

Rigorous vegetation science and efficient environmental and forest planning require careful assessment and management of data uncertainty from sources (Mason et al., 2015; Meyer et al., 2016) that include systematic observer error (Hall and Okali, 1978), purposive or subjective rather than random sampling (Lepš and Hadincová, 1992), temporal change in plant communities (Kopecký and Macek, 2015), and the dynamic nature of species distributions (Rocchini et al., 2011). The adverse effects of data uncertainty are often underrated (Walther and Moore, 2005) but extend to predictive models, estimates of key biodiversity indicators (Mason et al., 2015) and the reliability of conservation strategy outcomes (Hortal et al., 2008; Meyer et al., 2016). Hence, understanding data uncertainty is necessary to avoid biased ecological inferences (Meyer et al., 2016).

Maps are a central element for describing phenomena and for understanding patterns beyond the data. In particular, the spatial aspects of data uncertainty could be represented by a” map of ignorance” that depicts uncertainty in a spatially explicit manner (Rocchini et al., 2011). For this purpose, density-equalizing maps or cartograms are increasingly used to represent aspects of spatial variation for multiple variables simultaneously (Hennig, 2014). Cartograms are maps that vary the size and the shape of geographic regions proportionally with respect to a selected variable (Gastner and Newman, 2004; Tobler, 2004). Specifically, Gastner and Newman (2004) proposed a diffusion-based method for producing density-equalizing maps by first defining a starting map density $\rho(r)$ (where r represents geographic position) as the basis for constucting the cartogram. Traditionally, cartograms have been used to represent economic and political data but more recently have also been used to display statistical information for purposes of

* Corresponding author.
E-mail addresses: marta.galluzzi@unifi.it (M. Galluzzi), duccio.rocchini@unitn.it (D. Rocchini), roberto.canullo@unicam.it (R. Canullo), rmcroberts@fs.fed.us (R.E. McRoberts), gherardo.chirici@unifi.it (G. Chirici).

https://doi.org/10.1016/j.ecoinf.2018.06.005
Received 5 March 2018; Received in revised form 19 June 2018; Accepted 23 June 2018
Available online 28 June 2018
1574-9541/

identifying the potential spread of disease (Kronenfeld and Wong, 2017) and quantifying the effects of non-random sampling on species occurrence estimates (Rocchini et al., 2017). Conventional maps use two features, color and geometry, to describe attributes, a practice that may be less well-suited for heterogeneous information and big-data. A cartogram uses three types of information: geographic information (distorted), statistical information (depending on area size) and additional related information (defined by color). Specifically, a cartogram depicts the values of the attribute of interest by changing area sizes and shapes to reflect different levels of importance.

Knowing the spatial distribution of a plant species contributes to understanding the spatio-temporal ecological processes and ecosystem functions that provide benefit for other species, physical phenomena and human well-being (Meyer et al., 2016). Describing and assessing how vegetation biodiversity patterns change in response to anthropogenic pressures is a great challenge whose solution is limited by the lack of biodiversity information (Mihoub et al., 2017). Existing sources of biodiversity information include the Global Biodiversity Information Facility (GBIF; http://www.gbif.org/) that provides world-scale biological data and the European Vegetation Archive (EVA; http://euroveg.org/), a centralized database of European vegetation plot data. Although these are important sources of data, they are limited by the different survey designs, operators and sampling protocols used to acquire the data. For Europe, data collected between 2005 and 2008 by the BioSoil-Biodiversity project as part of the International Cooperative Programme on Assessment and Monitoring of Air pollution effects on Forests (ICP Forests) represent a unique example of a pan-European forest and biodiversity database obtained using statistically rigourous systematic surveys (16 × 16 km grid: Level I network). Specifically, ground vegetation (GVG) data for 3093 plots distributed across the forests of 19 European countries are available through the LI-BioDiv database (Canullo, 2016).

Despite the above rigorous sampling design, the GVG data are still subject to uncertainty due to the different years and seasons in which the surveys were conducted, to different operators, and to different country or organization protocols. The main GVG concern regarding the LI-BioDiv data relates to differences in the years and months in which the data were collected. Plant species identification is related to morphometric characters such as leaves and flower structure. (Cope et al., 2012). Vegetation development is further affected by a complex interaction between abiotic and biotic factors throughout the year (Larcher, 2003; Rocchini et al., 2017). The optimal condition for plant growth occurs when metabolic and hormonal factors including radiation, temperature, and chemical condition are in a synergistic relationship (Larcher, 2003). A specific thermal range is a prerequisite for the life cycle with the rate of plant germination increasing as temperature increases after reaching the minimum threshold temperature (Hatfield and Prueger, 2015; Larcher, 2003). Temperature is a crucial factor that drives morphometric plant character, and multiple studies have suggested a significant relationship between mean monthly air temperature and flowering (Fitter and Fitter, 2002; Hatfield and Prueger, 2015). Therefore, the ICP Forests Manual for Ground Vegetation assessment

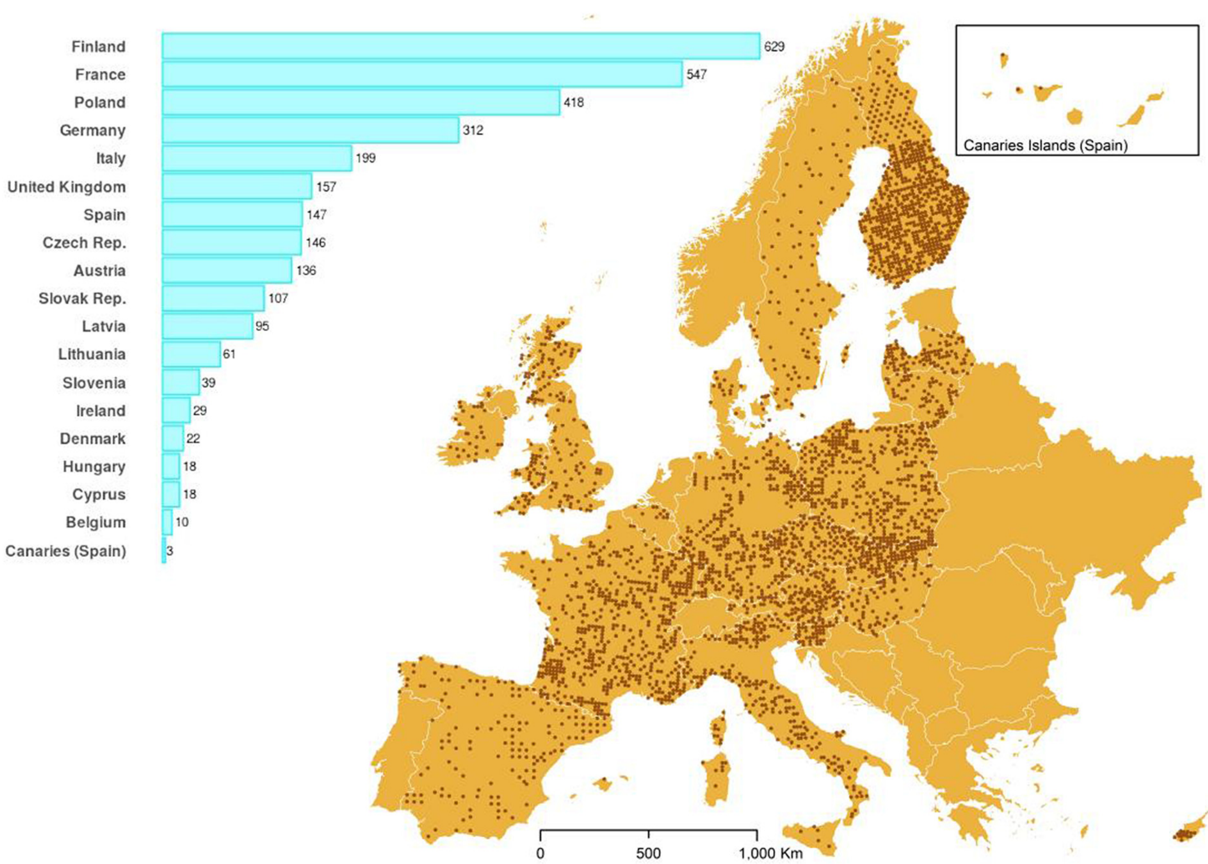

Fig. 1. Ground vegetation plots (3093) and number of plots by country included in the LI-BioDiv dataset (after BioSoil-Biodiversity project).

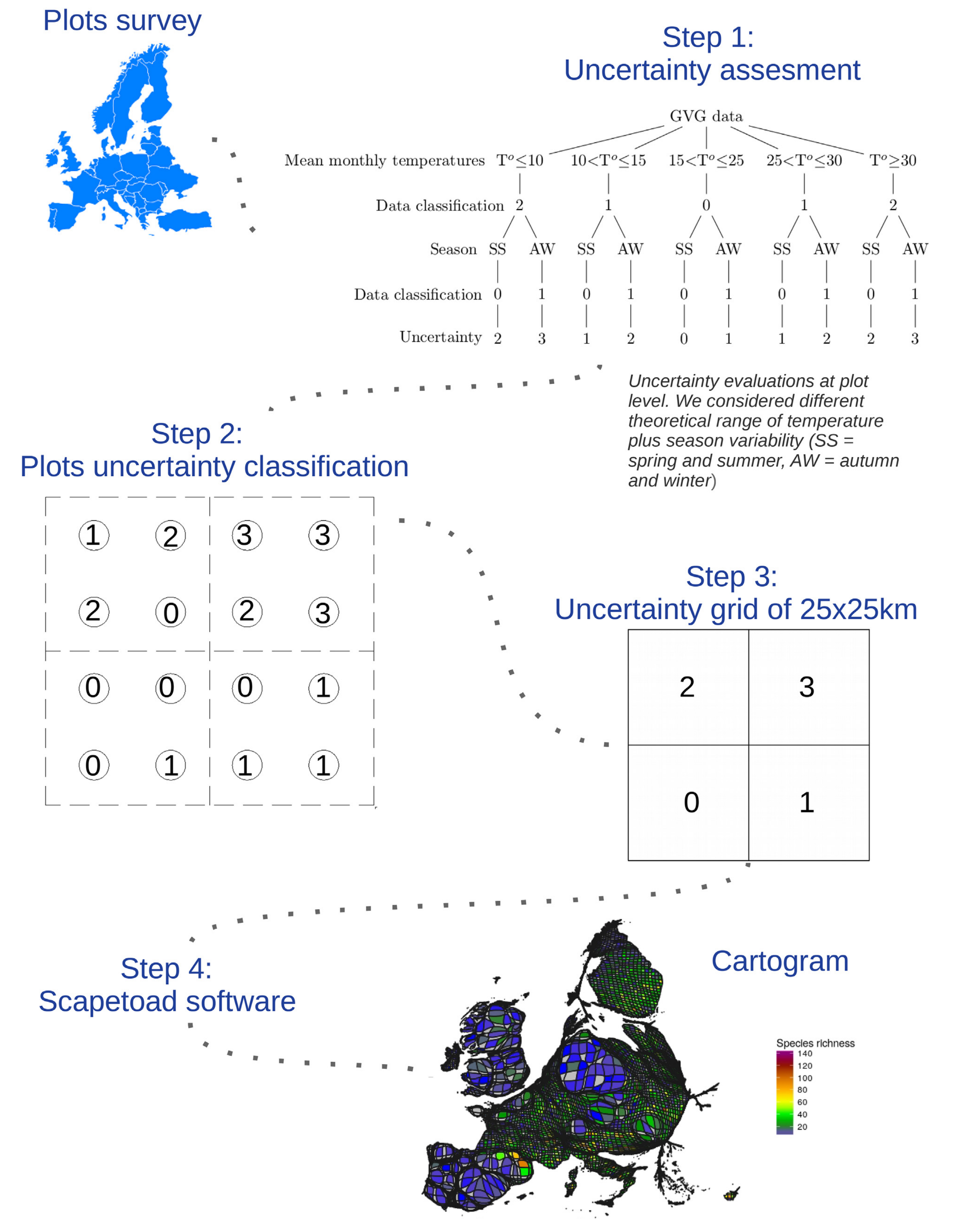

Fig. 2. Flowchart of steps to accomplish uncertainty cartograms. Step 1: Tree diagram explaining the uncertainty classification based i) on mean monthly temperature (increasing value of uncertainty moving from a theoretical optimum) and ii) on month of survey, with 0 uncertainty of plots surveyed between 1 April and 30 September (spring and summer). Step 2: Based on step 1, each plot was classified into different uncertainty values. To implement the cartogram, uncertainty values must be linked to a polygonal object, which in this study was represented by a grid cell with a resolution of 25-km × 25-km. Step 3: Each cell of the grid was associated with the uncertainty value of the plot with the greatest uncertainty record. Step 4: Starting from the 25-km × 25-km grid, the cartogram was produced through the open source software ScapeToad (https://scapetoad.choros.ch/).

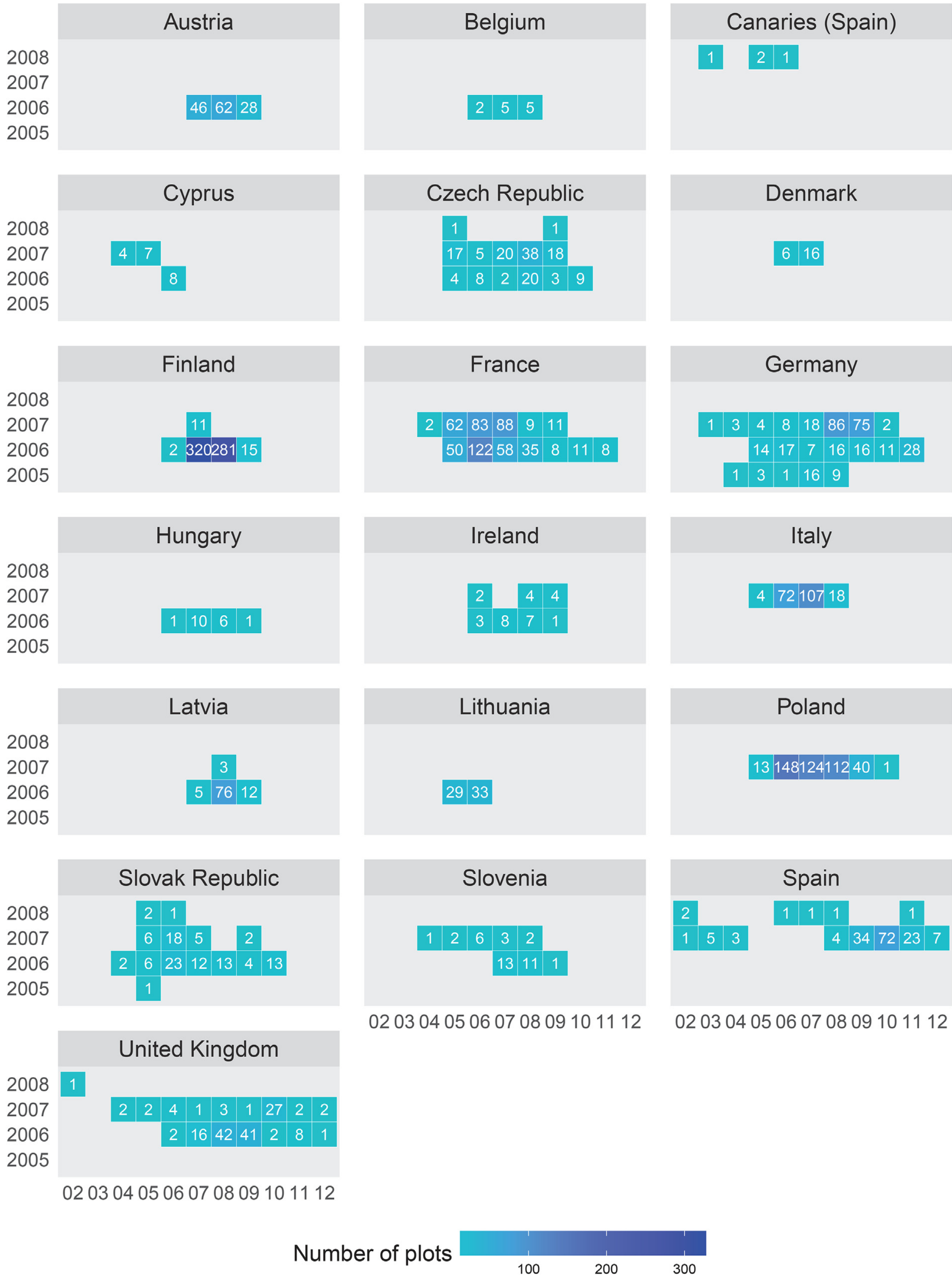

Fig. 3. Number of plot per month and year surveyed by each country for Ground vegetation assessment across the ICP Forests Level I network.

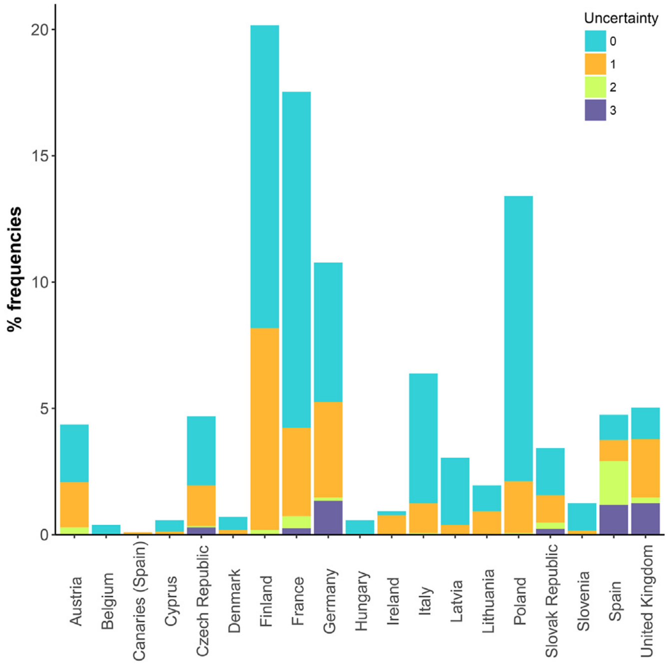

Fig. 4. Plots uncertainty percentage frequencies by country of GVG data included in the ICP Forests Level I network.

recommends that surveys be conducted when plants have the maximum biomass and when the maximum number of species can be correctly and readily identified (Canullo et al., 2013), although specific periods for specific bioregions are not prescribed (WGFB, 2011).

Explicitly describing and quantifying the uncertainty inherent in data recorded under different environmental conditions is the first step in clarifying how data uncertainty can be managed. Construction of maps that simultaneously depict observed data and their associated uncertainties is a challenge that has rarely been addressed. The aim of this study was to develop a procedure for mapping data uncertainty resulting from different GVG survey periods as recorded for the ICP Forests Level I network. The underlying purpose was to better manage and compare uncertainties when assessing European forest biodiversity. In particular, the study had three technical objectives (i) to develop a method for characterizing uncertainty based on the theoretical temperature range for a plant's life cycle, (ii) to map uncertainty using a diffusion-based method, and (iii) to develop a method for assessing biodiversity data quality using an uncertainty-biodiversity relationship. The results will be useful for achieving satisfactory accuracy of European forest biodiversity estimates, for correctly interpreting the estimates, and for facilitating comparable use of the data. Finally, we propose an alternative method for investigating and displaying data uncertainty (Fig. 5).

2. Materials and methods

Based on the LI-BioDiv dataset (UN⸌ECE ICP Forests PCC Collaborative Database; www.icp-forests.org) data from 19 countries (Austria, Belgium, Cyprus, Czech Republic, Denmark, Finland, France, Germany, Hungary, Italy, Ireland, Latvia, Lithuania, Poland, Slovak Republic, Slovenia, Spain, Sweden, United Kingdom), were assessed (Fig. 1). On each of the 3093 plots, all vascular plants were sampled in a surface of 400 m^2 between 2005 and 2008, following a standard manual with not strictly defined survey period (WGFB, 2011). The resulting species lists have been used to estimate “plant species richness” as a measure of European-level plant diversity. Two types of information were used to asses the uncertainty of the data for each plot: the mean monthly temperature and the season of the survey. The mean monthly temperature was estimated from *Climatologies at high resolution for the earth's land surface areas* (CHELSA) (Karger et al., 2017a, 2017b). We assigned an increasing value for uncertainty as mean monthly temperature deviated from the theoretical optimum of 15°C $< T \leq$ 25°C (Step 1 in Fig. 2) (Larcher, 2003; Rocchini et al., 2017). In a second step, we assigned an uncertainty value of 0 to plots surveyed between 1 April and 30 September (spring and summer, when plants have the maximum biomass, and the maximum number of species can be readily assessed (Canullo et al., 2013)) and an uncertainty value of 1 to plots surveyed between 1 October and 31 March (Fig. 2). Taking into account these two information items, an uncertainty value ranging between 0 (no seasonal uncertainty) and 3 (maximum seasonal uncertainty) was assigned to each plot as explained in the tree diagram in Fig. 2.

Maximum uncertainty was depicted for 25-km × 25-km grid cells encompassing the entire European continent. Cartograms were constructed using the open source software ScapeToad (https://scapetoad.choros.ch/). Distortions in the shapes of cells were determined by the uncertainty values, while color represented species richness values. Linear regression models were used to estimate the relationship between species richness and the uncertainty values. Finally an uncertainty versus latitude profile was developed for assessing species richness change across Europe.

3. Results and discussion

The number of GVG plots surveyed per month and year by each country was displayed using a heat map whereby values stored in a matrix were represented by colors (Wilkinson and Friendly, 2009) (Fig. 3). Spain, the United Kingdom and Germany had the greatest variability with resect to survey months followed by the Slovak Republic and France. During the autumn-winter months, Spain surveyed 111 of 147 plots, the United Kingdom surveyed 43 of 157 plots, Germany surveyed 133 of 312 plots, France surveyed 19 of 547 plots, and the Slovak Republic surveyed 13 of 107 plots.

The results of the uncertainty assessments showed that Spain, the United Kingdom and Germany had the most plots with the largest uncertainty value of 3 (Fig. 4). In particular, Spain had only a few plots with small uncertainty values. France, the Slovak Republic, Finland, Poland and the Czech Republic had both the most plots with small uncertainty values and fewest plots with large uncertainty values (Fig. 4). For the other countries, the plot uncertainty values were mostly small or absent (Fig. 4). Overall, 63% of plots had uncertainty values of 0, 28% of plots had uncertainty values of 1, and 3% and the 4% of plots had uncertainty values of 2 and 3, respectively.

The cartograms depicted the distribution of GVG uncertainty due to survey period. Fig. 5(a) is a cartogram in which the cell sizes were scaled according to the maximum uncertainty value of 3, while in Fig. 5(b) cell sizes were scaled according to the second largest uncertainty value of 2. The maps exhibited considerable distortion due to the heterogeneity of the survey periods across Europe. The German federal states of Berlin and Brandenburg, Spain and the United Kingdom showed the greatest cell distortions Fig. 5(a). In the second cartogram (Fig. 5(b)) Spain still showed the greatest distortion while distortions for the United Kingdom and Berlin and Brandenburg were

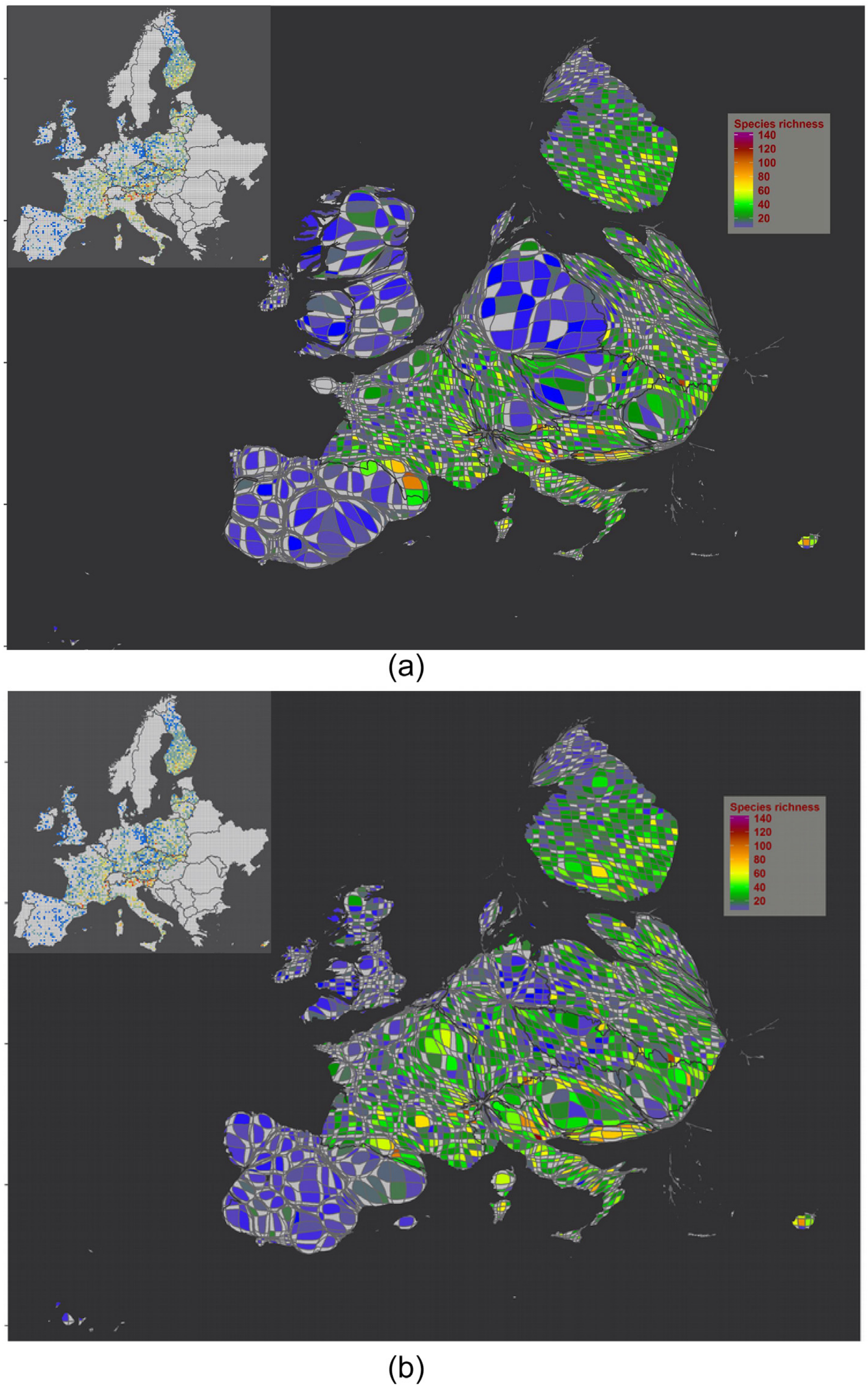

Fig. 5. Cartograms of uncertainty data-related (cell distortion): (a) proportional to maximum uncertainty value (equal to 3), (b) proportional to the second highest uncertainty value (equal to 2). Cells color range according to plant species richness values. In the upper left a reference undeformed map.

reduced. The cartogram showed that greater cell distortion was associated with smaller values of species richness, a result that was confirmed by the linear regression analyses ($F = 99.86$, $p < 0.001$) (Fig. 6).

The uncertainty versus latitude profile highlighted how species richness changed across latitudes by considering different uncertainty values (Fig. 7). Specifically, the vertical axis shows the latitude, the horizontal axis shows the uncertainty values while the color represents species richness values. Uniform species richness patterns corresponded to greater uncertainty values of 2 and 3, while heterogeneous patterns

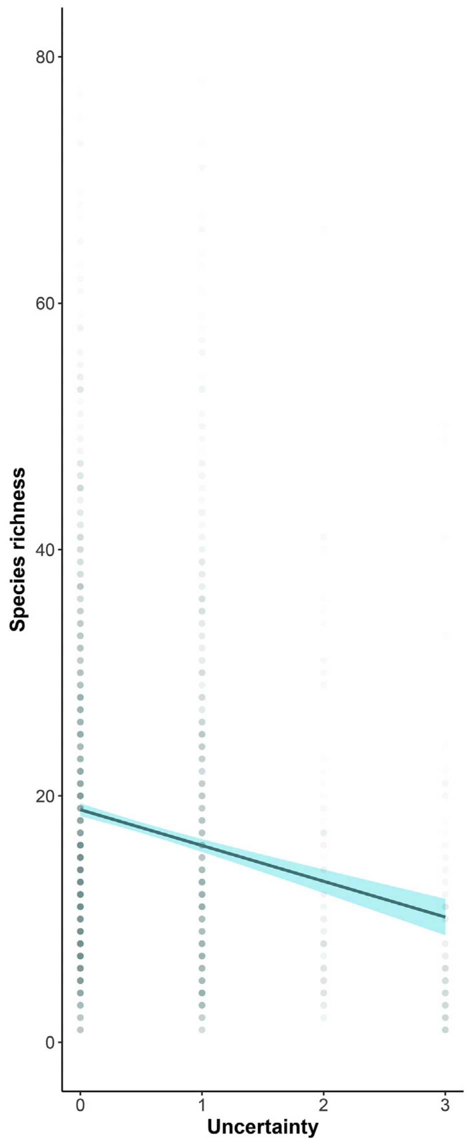

Fig. 6. Relation between uncertainty data values and plant species richness records (R^2 = 0.4, F = 99.86, p < 0.001, with 95% confidence interval) of GVG data included in the ICP Forests Level I network.

characterized smaller uncertainty values. The chart shows that when considering only smaller uncertainty values, species richness is greater for the lower latitude Mediterranean bioregion than for the higher latitude Boreal bioregion, (Fig. 7).

We have provided a different approach for visualizing and evaluating data uncertainty using diffusion-based cartograms. In particular, we have displayed how data information changes with respect to degree of uncertainty. For this study, area size distortions highlighted that data collected in Spain, the United Kingdom and the German federal states of Berlin and Brandenburg have uniformly large uncertainty values, based on survey period. Furthermore, such distortions corresponded to smaller species richness values as highlighted by grid color. The cartogram revealed the effects of lack of harmonized data collection protocols due to survey period and its effects on data quality. Plots surveyed in months with mean temperatures < 10 °C and during autumn and winter had very small numbers of species recorded. The cartogram visualizations gave different perspectives on phenomena, visually depicted spatial relations, and emphasized trends that occur in isolated areas (i.e. the German federal states of Berlin and Brandenburg). The two cartograms showed different distortions for the two greatest uncertainty levels. The United Kingdom and the German federal states of Berlin and Brandenburg were greatly distorted by considering only uncertainty level 3, while Spain is largely expanded in both maps. These patterns were not easy to be read in Fig. 4 that paradoxically contained more and complete information compared to a cartogram, but the use of both representations permits a clearer picture of the data pattern.

Understanding heterogeneity related to botanical data is a complex task that can change plans for managing data (Meyer et al., 2016). We found clear evidence of distinct species richness patterns across Europe, mainly due to non-random sampling (Rocchini et al., 2017) resulting from extreme differences in survey periods. Fig. 7 tracks data limitations and suggests positive prospects for using GVG data. The use of all GVG data can lead to underestimates of vascular plant biodiversity in term of species richness across Europe. The underestimation penalizes the Mediterranean bioregion (due to the uncertainty of the Spanish data) which is one of the most complex and biodiverse regions (Blondel and Aroson, 1999). We have demonstrated that it is possible to reduce the effects of underestimation by considering different levels of uncertainty. However, our results may be sensitive to the choice of uncertainty evaluation (Fig. 2). For example, different algorithms could be developed to assess uncertainty related to survey period (Rocchini et al., 2011), and it is possible that other solutions might produce additional information. Botanical uncertainty evaluation is complex because of environmental conditions and seasonal variability. In addition, because multiple factors affect both plant development and managment of big data, defining common criteria for assessing uncertainty is not easy. We think our solution leads to improved data interpretation as demonstrated by the results and the connections between uncertainty and vascular plant species richness records. In particular, the negative relationship between survey season and species richness is a clear indication of data misinterpretation of the field manual. GVG data from the systematic Forests (ex BioSoil-Biodiversity project) are a unique example of floristic data based on a rigourous, representative sampling design survey across Europe. Because data uncertainty could affect conservation strategies, our findings have important implications for assessing European forest biodiversity.

4. Conclusion

Density-equalizing maps or cartograms were used to depict data uncertainty. The cartograms allowed us to display geographic relationships among attributes and to highlight the effects of data uncertainty on data patterns. We highlighted data limitations and provided a method for understanding how different field data collection strategies may influence data analysis and statistical inference. In this case, data uncertainty led to underestimation of species richness, especially across Europe and for the Mediterranean bioregion. Botanical data are often heterogeneous depending on the observers, sampling designs and environment conditions. We demonstrated that cartograms are a useful tool for exploring and handling data, particularly when used together with other graphs and representations. Density-equalizing maps or cartograms give different representations of spatial phenomena that can be visualized and understood from different

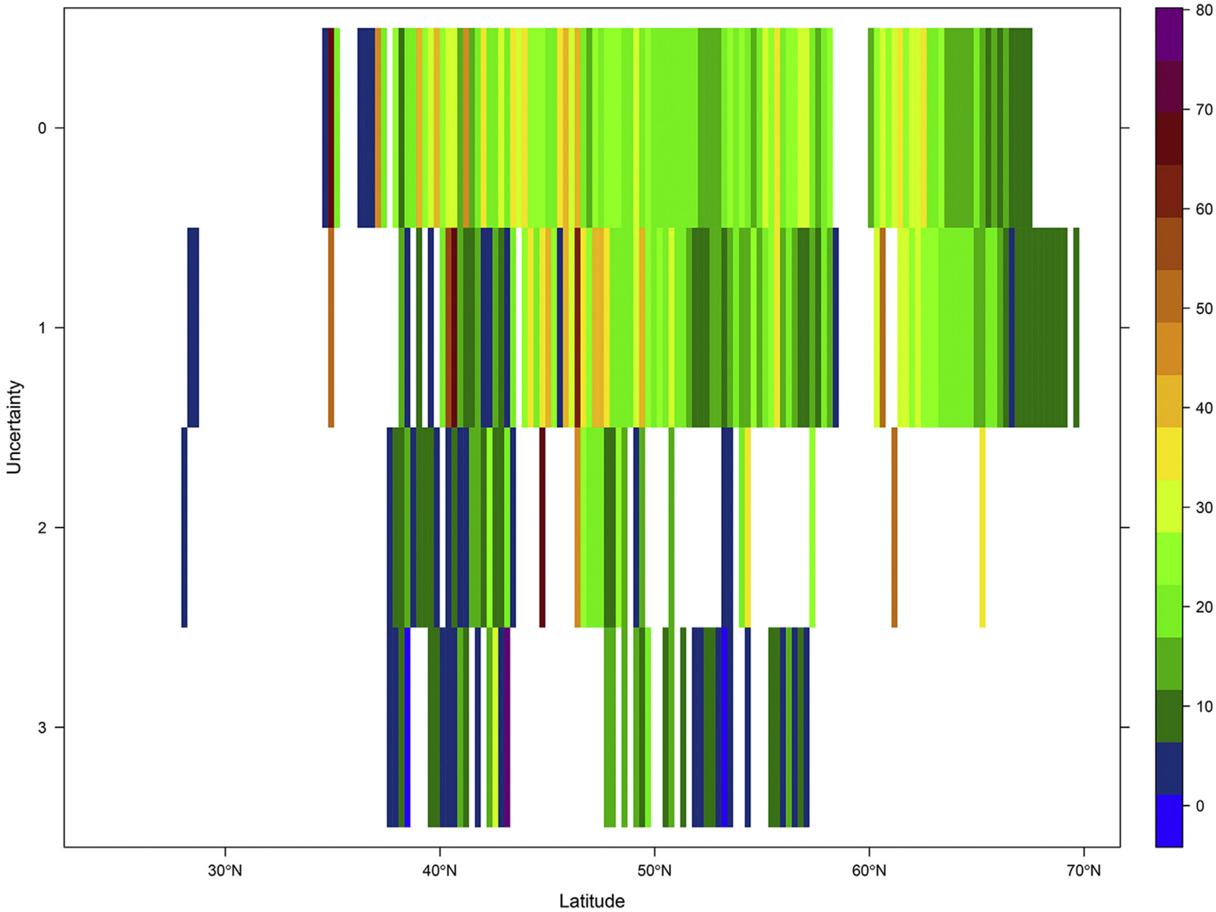

Fig. 7. Latitude profile of species richness according to the different values of data uncertainty (minimum: 0; maximum: 3) of GVG data included in the Li-BioDiv database under the ICP Level I network. The horizontal axis shows latitude degrees, and the vertical axis shows species richness average under different values of data uncertainty while the color represents species richness values. Lower plant species richness values occur at greater uncertainty values (3, on the lower part), while an increase of heterogeneity appears at smaller uncertainty values (0, at the bottom).

perspectives. Thus, effective alternative methods for evaluating and managing data uncertainty should be encouraged as a means of strengthening data analyses and their interpretation.

Acknowledgement

Research carried out in the frame of ICP Forests internal evaluation group “USPEX - Upscaling & Spatially explicit estimation of biophysical variables with remote sensing”, coordinated by Gherardo Chirici.

References

WGFB, 2011. The BioSoil Forest Biodiversity field manual, version 1.0/1.1/1.1a for the field assessment 2006-07. In: Joint Research Centre – Institute for Environment and Sustainability (Ed.), Evaluation of BioSoil Demonstration Project: Forest Biodiversity. Publications Office of the European Union, Luxembourg, pp. 81–102.
Blondel, J., Aroson, J., 1999. Biology and Wildlife of Mediterranean Region. Oxford University Press.
Canullo, R., 2016. The ICP forests Level I biodiversity data. A harmonized data source and baseline for plant species and structural diversity on european forest ecosystems. In: Michel, A.S.W. (Ed.), Forest Condition in Europe: 2016 Technical Report of ICP Forests. Report Under the UNECE Convention on Long-Range Transboundary Air Pollution (CLRTAP). BFW Austrian Research Centre for Forests, Vienna, pp. 89–105 BFW-Dokumentation 23.
Canullo, R., Starlinger, F., Giordani, P., 2013. Chapter 13 - Diversity and composition of plant and lichen species. In: Ferretti, M., Fischer, R. (Eds.), Forest Monitoring, Volume 12 of Developments in Environmental Science. Elsevier, pp. 237–250. https://doi.org/10.1016/B978-0-08-098222-9.00013-3.
Cope, J.S., Corney, D., Clark, J.Y., Remagnino, P., Wilkin, P., 2012. Plant species identification using digital morphometrics: a review. Expert Syst. Appl. 39 (8), 7562–7573. https://doi.org/10.1016/j.eswa.2012.01.073.
Fitter, A.H., Fitter, R.S.R., 2002. Rapid changes in flowering time in British plants. Science 296 (5573), 1689–1691. https://doi.org/10.1126/science.1071617.
Gastner, M.T., Newman, M.E.J., 2004. Diffusion-based method for producing density-equalizing maps. Proc. Natl. Acad. Sci. U. S. A. 101 (20), 7499–7504. https://doi.org/10.1073/pnas.0400280101.
Hall, J.B., Okali, D.U.U., 1978. Observer-Bias in a floristic survey of complex tropical vegetation. J. Ecol. 66 (1), 241–249. https://doi.org/10.2307/2259191.
Hatfield, J.L., Prueger, J.H., 2015. Temperature extremes: effect on plant growth and development. Weather Clim. Extremes 10 (Part A), 4–10. https://doi.org/10.1016/j.wace.2015.08.001.
Hennig, B.D., 2014. Gridded cartograms as a method for visualising earthquake risk at the global scale. J. Maps 10 (2), 186–194. https://doi.org/10.1080/17445647.2013.806229.
Hortal, J., Jiménez-Valverde, A., Gómez, J.F., Lobo, J.M., Baselga, A., 2008. Historical bias in biodiversity inventories affects the observed environmental niche of the species. Oikos 117 (6), 847–858. https://doi.org/10.1111/j.0030-1299.2008.16434.x.
Karger, D., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R., Zimmermann, N., Linder, H., Kessler, M., 2017a. Data from: Climatologies at High Resolution for the Earth's Land Surface Areas. https://doi.org/10.5061/dryad.kd1d4.
Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N., Linder, H.P., Kessler, M., 2017b. Climatologies at high resolution for the earth's land surface areas. Sci. Data 4 (170122). https://doi.org/10.1038/sdata.2017.122.
Kopecký, M., Macek, M., 2015. Vegetation resurvey is robust to plot location uncertainty. Divers. Distrib. 21 (3), 322–330. https://doi.org/10.1111/ddi.12299.
Kronenfeld, B.J., Wong, D.W.S., 2017. Visualizing statistical significance of disease clusters using cartograms. Int. J. Health Geogr. 16 (1). https://doi.org/10.1186/s12942-017-0093-9.
Larcher, W., 2003. Physiological Plant Ecology. Springer-Verlag, Berlin Heidelberg (vi edition).

Lepš, J., Hadincová, V., 1992. How reliable are our vegetation analyses? J. Veg. Sci. 3 (1), 119–124. https://doi.org/10.2307/3236006.

Mason, N.H.W., Holdaway, R.J., Richardson, S.J., 2015. Quantifying uncertainty in biodiversity data for monitoring and reporting indicators. In: Technical Report. Department of Conservation, 70 Moorhouse Avenue Christchurch 8011 New Zealand.

Meyer, C., Weigelt, P., Kreft, H., 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. Ecol. Lett. 19 (8), 992–1006. https://doi.org/10.1111/ele.12624.

Mihoub, J.-B., Henle, K., Titeux, N., Brotons, L., Brummitt, N.A., Schmeller, D.S., 2017. Setting temporal baselines for biodiversity: the limits of available monitoring data for capturing the full impact of anthropogenic pressures. Sci. Rep. 7. https://doi.org/10.1038/srep41591.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jiménez-Valverde, A., Ricotta, C., Bacaro, G., Chiarucci, A., 2011. Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. Prog. Phys. Geogr. 35 (2), 211–226. https://doi.org/10.1177/0309133311399491.

Rocchini, D., Garzon-Lopez, C.X., Marcantonio, M., Amici, V., Bacaro, G., Bastin, L., Brummitt, N., Chiarucci, A., Foody, G.M., Hauffe, H.C., He, K.S., Ricotta, C., Rizzoli, A., Rosà, R., 2017. Anticipating species distributions: handling sampling effort bias under a Bayesian framework. Sci. Total Environ. 584-585, 282–290. https://doi.org/10.1016/j.scitotenv.2016.12.038.

Tobler, W., 2004. Thirty five years of computer cartograms. Ann. Assoc. Am. Geogr. 94 (1), 58–73. https://doi.org/10.1111/j.1467-8306.2004.09401004.x.

Walther, B.A., Moore, J.L., 2005. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. Ecography 28 (6), 815–829. https://doi.org/10.1111/j.2005.0906-7590.04112.x.

Wilkinson, L., Friendly, M., 2009. The history of the cluster heat map. Am. Stat. 63 (2), 179–184. https://doi.org/10.1198/tas.2009.0033.