



Investigating the usability of databases with different spatial characteristics for the occurrence of invasive plant species: a Hungarian scale analysis

Georgina Veronika Vizsra¹, Kata Frei², Alida Anna Hábcenyus², Anna Soóky², Zoltán Bátori², Annamária Laborczi³, Nándor Csikós³, Gábor Szatmári³, Péter Szilassi¹

¹Department of Physical Geography and Geoinformatics, University of Szeged, Hungary

²Department of Ecology, University of Szeged, Hungary

³Department of Soil Mapping and Environmental Informatics, Institute for Soil Sciences, Centre for Agricultural Research, Hungary

*myosotis.sylvatica12@gmail.com

Nowadays, biological invasions have increasingly serious environmental consequences. The increasing spread of invasive plant species is the most serious environmental risk in protected areas. The removal of established populations requires a significant financial investment. To avoid conservation and economic damage, the spread of invasive species must be prevented in the first place. To identify potentially threatened habitats, databases are needed that reliably indicate the presence of invasive species. There are several free Citizen science-type databases available, however these data are sporadic and only provide information on the presence of plants, nevertheless, they may be useful in mapping invasive species.

In this study, we investigated the usability of three databases in relation to the occurrence of two invasive plant species (*Ailanthus altissima* and *Asclepias syriaca*). One of the three databases (EUROSTAT Land Use and Coverage Area Frame Survey (LUCAS)) is non-sporadic, it covers all areas of Hungary, and the data were collected by specialists. We examined two databases (Data Source Hungarian Ministry of Agriculture and gbif.org) of sporadic data nature against this database to determine the extent to which the sporadic distribution of the data affects the reliability of the results. The analysis was carried out according to the correlation between the distribution of the invasive species considered and the characteristics of the topsoil and certain hydrological and climatic parameters. For statistical analysis, we used the one-way ANOVA model with the function `aov`. The pairwise comparison of the values of the environmental parameters associated with the three different databases was performed with a Tukey post-hoc test with the function `TukeyHSD`.

The results showed that in many cases there are significant differences between the results from different databases. In most cases, a significant difference was found between the results of the non-sporadic database and results from one of the two sporadic databases.