Konrád K. D., Bede-Fazekas Á., Molnár Z. & Somodi I. (2022) Multilayer landscape classification based on potential vegetation. – Preslia 94: 631–650.

Supplementary Data S2. – Details of the methodology of the MPV estimation (based on Somodi et al. 2017)

The input data (MPV) of this study are either the result of previous predictions (Somodi et al. 2017) or of predictions created using the methodology published by Somodi et al. (2017). However, for the sake of clarity, the main steps of the previous analyses are summarised in this appendix.

The potential distribution (probability of survival) of 47 habitats (Bölöni et al. 2011, Appendix S3) was modelled, which are considered either potential natural vegetation (Tüxen 1956) or potential replacement vegetation (Chytrý 1998) currently in Hungary. Originally, 39 potential natural vegetation types were included (Somodi et al. 2017). The remaining 8 potential replacement vegetation types were added to complement the original estimation for this study. Input presence/absence vegetation data were derived from the Hungarian Actual Habitat Database (MÉTA, Molnár et al. 2007; Horváth et al. 2008), compiled by field surveys between 2003 and 2005 and providing data on the (semi)natural vegetation of Hungary in a hexagonal grid containing cells of approx. 700 m in diameter.

The explanatory variables considered were those of climate (CarpatClim-Hu; Szalai et al. 2013), soil (DOSoReMI.hu; Pásztor et al. 2015), topography (USGS 2004) and hydrology. All the non-climatic variables were aggregated for, or extracted to the centre of the hexagons in the MÉTA database. The exact environmental covariates were chosen by restricting pairwise Pearson correlation to 0.8, Variance Inflation Factor (VIF) to 50, and Condition Number (CN) to 30.

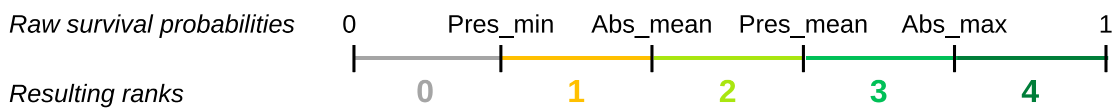
The estimation was carried out using gradient boosting models (GBMs; Elith et al. 2008), building a separate model for each vegetation type. Prior to modelling, hexagons with no recognizable (semi)natural vegetation were removed and the remaining data (approx. 36.46% of the country) were divided into two sets of the same size for training and evaluation. To allow the evaluation of model performance, the division was made with random sampling with prevalence stratification for each type separately.

Since the raw probability values given by the models are affected the specific characteristics of the given vegetation type, they are not directly comparable across vegetation types. In order to achieve comparability across vegetation types, the raw probabilities had to be the standardised for each vegetation type (Somodi et al. 2017). A rescaling procedure was used for this purpose, resulting in an ordinal scale of five ranks (0, 1, 2, 3, 4, indicating increasing survival probabilities). The 4 thresholds (Supplementary Table S2.1) were calculated separately for each vegetation type.

Supplementary Table S2.1. – Thresholds for the ranking procedure

Probability	**Abbreviation**	**Explanation**
Lowest	Pres_min	the minimal probability within hexagons with observed presence
Low	Abs_mean	the average probability within hexagons with observed absence
Medium	Pres_mean	the average probability within hexagons with observed presence
High	Abs_max	the highest probability within hexagons with observed absence

The ranking procedure can be visualised as follows:



The meaning of the categories is thus (based on Török et al. 2018):

- 0 – lower probabilities than the minimal probability within hexagons with observed presence
- 1 – higher probabilities than the minimal probability within hexagons with observed presence, but lower than the average probability within hexagons with observed absence
- 2 – higher probabilities than the average probability within hexagons with observed absence, but lower than the average probability within hexagons with observed presence
- 3 – higher probabilities than the average probability within hexagons with observed presence, but lower than the highest value within hexagons with observed absence
- 4 – higher probabilities than the highest value within hexagons with observed absence

This rescaling ensures that vegetation types with the same ranks are equally probable members of MPV in a given spatial unit (Somodi et al. 2017).

However, this rescaling was slightly modified , since previous experience showed that the difference between two lowest ranks has limited influence on the potential presence and can effectively be regarded as predicted absences. Therefore, at the beginning of the current analyses the two lowest ranks were combined to reduce the noise caused by distinguishing them. Thus, in the current study a four-level variant of the above ordinal scale was used for the MPNV data.

References

Bölöni J., Molnár Zs. & Kun A. (eds.) (2011) Magyarország élőhelyei. A hazai vegetációtípusok leírása és határozója. ÁNÉR 2011. [Habitats in Hungary. Description and identification guide of the Hungarian vegetation]. MTA ÖBKI, Vácrátót, Hungary. 441 p.

Chytrý M. (1998) Potential replacement vegetation: an approach to vegetation mapping of cultural landscapes. – Applied Vegetation Science 1: 177–188. https://doi.org/10.2307/1478947

Elith J., Leathwick J. R. & Hastie T. (2008) A working guide to boosted regression trees. – Journal of Animal Ecology 77: 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Hengl T. (2018) Global landform and lithology class at 250 m based on the USGS global ecosystem map (Version 1.0). Zenodo. – https://doi.org/10.5281/zenodo.1464846

Horváth F., Molnár Zs., Bölöni J., Pataki Zs., Polgár L., Révész A., Oláh K., Krasser D. & Illyés E. (2008) Fact sheet of the MÉTA database. – Acta Botanica Hungarica 50(Suppl.): 11–34. https://doi.org/10.1556/ABot.50.2008.Suppl.2

Molnár Zs., Bartha S., Seregélyes T., Illyés E., Botta-Dukát Z., Tímár G., Horváth F., Révész A., Kun A., Bölöni J., Bíró M., Bodonczi L., Deák József Á., Fogarasi P., Horváth A., Isépy I., Karas L., Kecskés F., Molnár Cs., Ortmann-né Ajkai A. & Rév Sz. (2007) A grid-based, satellite-image supported, multi-attributed vegetation maping method (MÉTA). – Folia Geobotanica, 42: 225–247. https://doi.org/10.1007/BF02806465

Pásztor L., Laborczi A., Takács K., Szatmári G., Dobos E., Illés G., Bakacsi Zs. & Szabó J. (2015) Compilation of novel and renewed, goal oriented digital soil maps using geostatistical and data mining tools. – Hungarian Geographical Bulletin 64(1):49–64. https://doi.org/10.15201/hungeobull.64.1.5

Somodi I., Molnár Zs., Czúcz B., Bede-Fazekas Á., Bölöni J., Pásztor L., Laborczi A. & Zimmermann N. E. (2017) Implementation and application of multiple potential natural vegetation models – a case study of Hungary. – Journal of Vegetation Science 28: 1260–1269. https://doi.org/10.1111/jvs.12564

Szalai S., Auer I., Hiebl J., Milkovich J., Radim T., Stepanek P., Zahradnicek P., Bihari Z., Lakatos M., Szentimrey T., Limanowka D., Kilar P., Cheval S., Deak Gy., Mihic D., Antolovic I., Mihajlovic V., Nejedlik P., Stastny P., Mikulova K., Nabyvanets I., Skyryk O., Krakovskaya S., Vogt J., Antofie T. & Spinoni J. (2013) Climate of the Greater Carpathian Region. Final Technical Report. URL: www.carpatclim-eu.org.

Török K., Csecserits A., Somodi I., Kövendi-Jakó A., Halász K., Rédei T. & Halassy M. (2018) Restoration prioritization for industrial area applying multiple potential natural vegetation modeling. – Restoration Ecology 26: 476–488. https://doi.org/10.1111/rec.12584

Tüxen R. (1956) Die heutige potentielle natürliche Vegetation als Gegenstand der Vegetationskartierung. Angewandte Pflanzensoziologie 13. Stolzenau, p. 4–42.

USGS (2004) Shuttle Radar Topography Mission, 3 Arc Second scene SRTM_u03_n045e016-SRTM_ff03_n048e022, Unfilled Unfinished 2.0, Global Land Cover Facility, February 2000. – University of Maryland, College Park, MD, USA