**Supplementary material**

**Environmental ranges discriminating between macrophytes groups in European rivers**

Table S1: Table of respective providers for the used dataset.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Country (abbreviated)** | **Name** | **Country (abbreviated)** |
| Agnieszka agnieszka Kolada | PL | Jens Kroker | DE |
| Alain Dohet | LU | Jochen Schaumburg | DE |
| An Leyssen | BE (Flanders) | Jürgen Böhmer | DE |
| Andreas Hoppe | DE | Kai Piirsoo | EE |
| Anik Schneiders | BE (Flanders) | Karin Pall | AT |
| Annette Baattrup-Pedersen | DK | Klaus Weyer | DE |
| Annette Baattrup-Pedersen | UK (GB) | Kris van Looy | BE (Flanders) |
| Annette Baattrup-Pedersen | DE | Krzysztof Szoszkiewicz | PL |
| Antun Alegro | HR | Lenka Tajmrova | CZ |
| Béla Csanyi | HU | Libuse Opatrilova | CZ |
| Brigitte Schmidt | AT | Livia Tothova | SK |
| Chris Burns | UK (NI) | Lukács András | HU |
| Christian Chauvin | FR | Maria Minicardi | IT |
| Christine Schranz | DE | Martin Mcgarrigle | IE |
| Christophe Laplace | FR | Matus Haviar | SK |
| Daniel Galoux | BE (Wallonia) | Nigel Willby | UK (GB) |
| Doris Stelzer | DE | Nora Welschbillig | LU |
| Eva Bellack | DE | Normunds Kadikis | LV |
| Gana Gecheva | BG | Peter Balazi | SK |
| Gorazd Urbanic | SI | Peter Rolauffs | DE |
| Henar Fraile | ES | Roelf Pot | NL |
| Hugo Coops | NL | Sebastian Birk | STAR (project) |
| Isabel Pardo | ES | Serban Iliescu | RO |
|  |  | Ulrich Schmieds | DE |
|  |  | Zofija Sinceviciene | LT |

Table S2: Performance measures of the model where missing values were replaced by the median and the model was assessed on the datasets excluding samples with missing values.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Out-Of-Bag performance | | | | Holdout performance | | | | 10 kfold-cross validation | | | |
| Accuracy | 50% (LCI=49%; HCI%=52%) | | | | 52% (LCI=50%; HCI=55%) | | | | 55% | | | |
| Cohen's kappa | 0.33 (LCI=0.32; HCI=0.35) | | | | 0.37 (LCI=0.33; HCI=0.40) | | | | 0.38 | | | |
| n | 6472 | | | | 1618 | | | | 8090 | | | |
| Groups | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Prevalence | 0.29 | 0.24 | 0.32 | 0.15 | 0.29 | 0.23 | 0.33 | 0.15 | 0.29 | 0.24 | 0.32 | 0.15 |

Table S3: Performance measures of the model where values were imputed with the missForest package at the location where no values were available.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Out-Of-Bag performance | | | | Holdout performance | | | | 10 kfold-cross validation | | | |
| Accuracy | 52% (LCI=51%; HCI%=53%) | | | | 51% (LCI=49%; HCI=54%) | | | | 54% | | | |
| Cohen's kappa | 0.35 (LCI=0.34; HCI=0.37) | | | | 0.34 (LCI=0.31; HCI=0.37) | | | | 0.37 | | | |
| N | 6472 | | | | 1618 | | | | 8090 | | | |
| Groups | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Prevalence | 0.29 | 0.24 | 0.32 | 0.15 | 0.29 | 0.23 | 0.33 | 0.15 | 0.29 | 0.24 | 0.32 | 0.15 |

Chart

Description automatically generated

Figure S1: Species occurrence along the NO3- gradient (sorted by decreasing 25th percentile values).