

RESEARCH ARTICLE

Countrywide classification of permanent grassland habitats at high spatial resolution

Nica Huber¹ , Christian Ginzler¹ , Robert Pazar¹ , Patrice Descombes^{2,3} ,
Andri Baltensweiler⁴ , Klaus Ecker⁵ , Eliane Meier⁶  & Bronwyn Price¹ 

¹Remote Sensing, Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8902, Birmensdorf, Switzerland

²Musée et Jardins botaniques cantonaux, Av. de Cour 14bis, 1007, Lausanne, Switzerland

³Department of Ecology and Evolution, University of Lausanne, Biophore, 1015, Lausanne, Switzerland

⁴GIS, Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8902, Birmensdorf, Switzerland

⁵Ecosystems Dynamics, Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8902, Birmensdorf, Switzerland

⁶Agricultural Landscapes and Biodiversity, Agroscope, Reckenholzstrasse 191, 8046, Zürich, Switzerland

Keywords

Bogs, distribution models, dry grasslands, fens, remote sensing, Sentinel, Switzerland

Correspondence

Nica Huber, Remote Sensing, Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8902 Birmensdorf, Switzerland. Tel: +41 44 739 28 19; Fax: +41 44 739 22 15. Email: nica.huber@wsl.ch

Editor: Mat Disney

Received: 11 March 2022; Revised: 8 June 2022; Accepted: 29 July 2022

doi: 10.1002/rse2.298

Remote Sensing in Ecology and Conservation 2023; **9** (1):133–151

Abstract

European grasslands face strong declines in extent and quality. Many grassland types are priority habitats for national and European conservation strategies. Countrywide, high spatial resolution maps of their distribution are often lacking. Here, we modelled the spatial distribution of 20 permanent grassland habitats at the level of phytosociological alliances across Switzerland at 10x10 m resolution. First, we applied ensemble models to provide distribution maps of the individual habitat types, using training data from various sources. Copernicus Sentinel satellite imagery and variables describing climate, soil and topography were used as predictors. The performance of these models was assessed based on the true skill statistics with a split-sampling of the data. Second, the individual maps were combined into countrywide maps of the most and second most likely habitat type, respectively, using an expert-based weighting approach. The performance of the combined map for the most likely habitat type was assessed via an independent testing dataset and a comparison of the predicted habitat-type proportions with extrapolations from field surveys. Most individual maps had useful to excellent predictive performance (TSS ≥ 0.6). For most grid cells in the combined maps, the most and second most likely habitat types were either ecologically closely related or representing two grassland types along a nutrient gradient. The same was true for omission errors. We found good agreement between the predicted and estimated proportions from field surveys. The area of raised bogs appears to be underestimated, while dry grasslands showed highest agreement. This work highlights the potential of earth observation data at fine spatial and temporal resolution to map habitats at broad scales, thereby providing the foundation for diverse conservation applications. A particular challenge remains in capturing the transition from nutrient-poor to nutrient-rich grasslands, which is highly important for biodiversity conservation.

Introduction

Permanent grasslands cover a substantial share of the present-day vegetation of Europe (17.4%; EUROSTAT, 2021) and are shaped by climatic, edaphic and topographic conditions as well as human management and disturbances (Davies et al., 2004; Devillers et al., 1991; Feurdean et al., 2018). Grasslands host a considerably high diversity of species,

particularly vascular plants (Boch et al., 2021; Willems et al., 1993), arthropods (Soderstrom et al., 2001), invertebrates (Standen, 2000) or grassland-adapted birds (Soderstrom et al., 2001). Yet, grassland habitats face strong declines in their extent and quality, predominantly due to land use change, agricultural intensification and abandonment, drainage or eutrophication (Bergamini et al., 2009; Gillet et al., 2016; Kuhn et al., 2021; Pepler-Lisbach et al., 2020).

Hence, many European grasslands are listed as priority habitats in the Habitats Directive of the European Union and the Swiss Ordinance on the Protection of Nature and Cultural Heritage, which both call for urgent protection.

The protection of grassland habitats in both the short and long term requires detailed mapping for effective maintenance and monitoring as well as for the establishment of buffer and connectivity zones (i.e., ecological networks; Jongman et al., 2011). This is particularly important in mountainous countries where highly fragmented and diverse habitat matrices result from high variability in site conditions as well as management practices. Switzerland features a diverse small-scale habitat structure and hosts a high number of grassland habitat types including many rare and endangered ones (Delarze et al., 2016). So far, information concerning the spatial distribution of vegetation types has primarily been provided by expert floristic field surveys (e.g., Eggenberg et al., 2001). However, such surveys are spatially restricted and highly time-consuming (Vanden Borre et al., 2011). Moreover, their informative value may be short-lived due to, for example, changes in land use, management practices or eutrophication (Bollens et al., 2001; Gillet et al., 2016; Graf et al., 2010; Rion et al., 2018). Habitat distribution models (Guisan & Zimmermann, 2000) have long been applied to analyze and predict habitat patterns, predominantly across large geographical extents. These models are primarily based on spatially explicit variables describing climate, geology or topography (e.g., Zimmermann & Kienast, 1999) and thus predict potential rather than existing vegetation. Yet, due to recent advances in remote sensing, high-resolution spatial, temporal and spectral information has become increasingly available for mapping and monitoring habitats (Corbane et al., 2015; Turner et al., 2003). The inclusion of remotely sensed variables of increasing resolution in addition to other environmental variables has shown great potential for refining distribution models and improving their performance (He et al., 2015).

Functional types of vegetation can be separated from each other relatively well with the help of remotely sensed data (Corbane et al., 2015). Yet, the distinction of phytosociological units, that is, plant alliances, is still challenging, particularly for grassland habitats. Challenges arise largely due to their spectral similarity (Ali et al., 2016; Wright & Wimberly, 2013), their spectral variability over time and space (Marcinkowska-Ochtyra et al., 2019; Tarantino et al., 2021), their potentially small spatial extent and their complex spatial structure (Diaz Varela et al., 2008; Küchler et al., 2004; Mehner et al., 2004), or the limited availability of high-quality field samples (Ali et al., 2016). Hence, at the level of phytosociological units, most recent studies have focused on the discrimination of few grassland habitat types in

protected areas (e.g., Natura 2000 sites; Tarantino et al., 2021, Marcinkowska-Ochtyra et al., 2019, Bhatnagar et al., 2020). Yet, focusing on protected areas alone does not provide a comprehensive overview of the distribution of ecologically valuable habitats and their remnants (Maiorano et al., 2007) and thus constrains further modelling and biodiversity conservation efforts.

Here, we map 20 permanent grassland habitat types (including fens, bogs and various meadows and pastures) across Switzerland at high spatial resolution (10 × 10 m). To cover a wide range of potential conservation applications, we provide a comprehensive set of mapping products, including (1) individual habitat maps, (2) combined grassland habitat maps, and (3) basic uncertainty maps. We chose a two-step approach (Fig. 1), where ensemble modeling outputs of the individual habitats were assembled into combined maps of the most and second most likely habitat type, respectively, using an expert-based weighting approach. The performance of the individual habitat maps was assessed based on the True Skill Statistic with a 10-fold split-sampling of the data. The performance of the ensemble maps of the individual habitat types as well as the performance of the combined map for the most likely habitat type were assessed via an independent testing dataset. In addition, the latter was evaluated via a comparison of the predicted proportions of each habitat type with extrapolations from field surveys.

Materials and Methods

Unless otherwise stated, all analyses were performed in the statistical software R (v4.0.3; R Core Team, 2018) and the geographic information system ArcGIS (v10.8.1, ESRI).

Data

Modelled habitat types

The Swiss habitat typology (TypoCH) of Delarze et al. (2015) underlies the national Red List of Habitats (Delarze et al., 2016) and is widely used to distinguish habitat types in Switzerland. The classification features nine habitat classes (1st level) subdivided into habitat groups (2nd level), further partitioned into habitat types (3rd level) typically corresponding to phytosociological alliances. According to this classification, grassland habitat types are either part of the habitat class 2 (Riparian areas, banks and wetlands) or 4 (Grasslands).

Here, we modelled 20 permanent grassland habitat types (3rd level; Table 1) belonging to the following habitat groups (2nd level): fens, wet meadows, raised bogs, reseeded and heavy fertilized grasslands, dry grasslands,

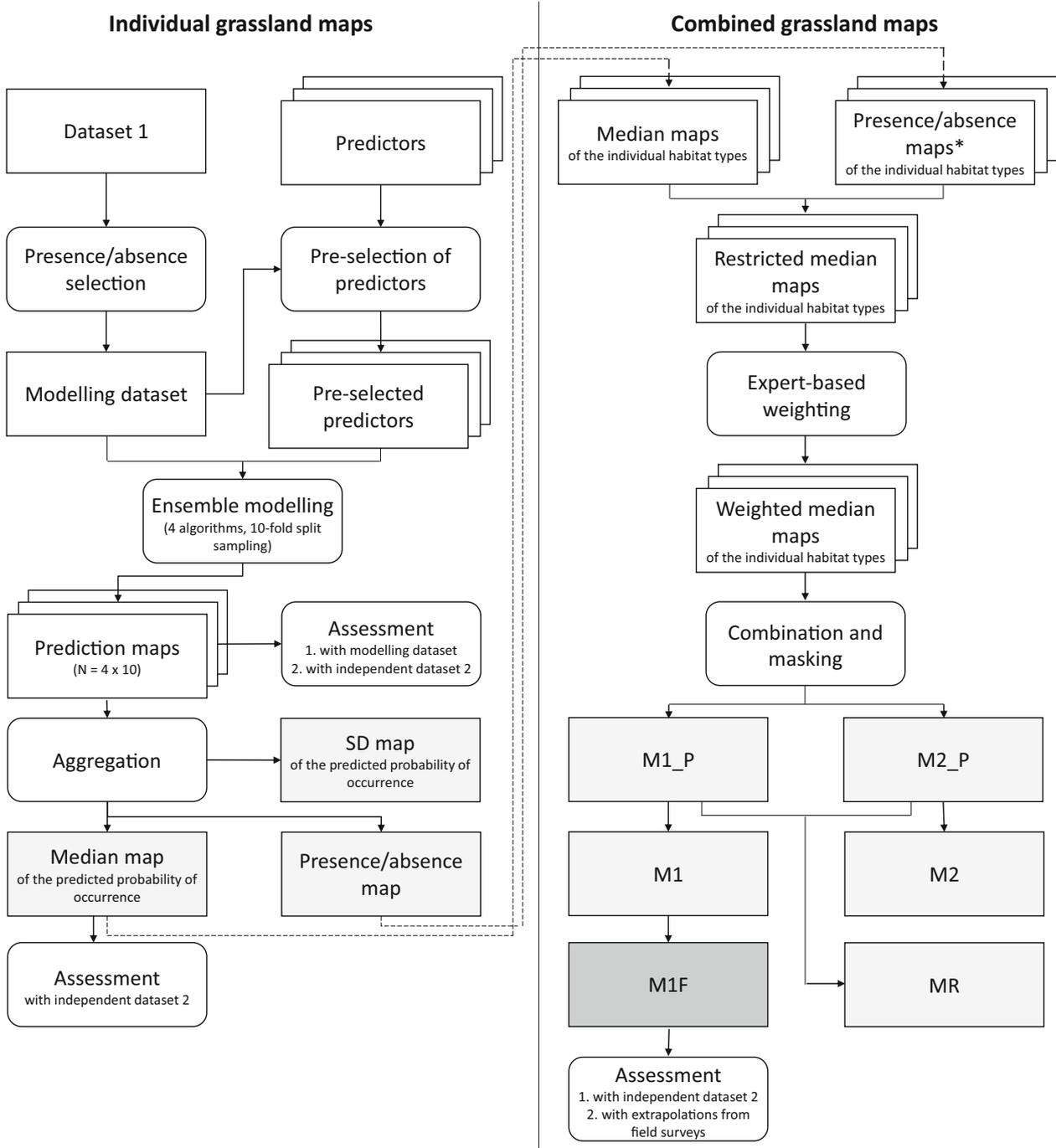


Figure 1. Overview of the modeling procedure. M1_P/M2_P: maps of the weighted median of the predicted probability of occurrence of the most/second most likely grassland habitat type, respectively; M1/M2: map denoting the most/second most likely habitat type, respectively; M1F: final combined map of the most likely habitat type (i.e., M1 after regional corrections); MR: map of the ratio of the probabilities of occurrence of the most and second most likely grassland habitat types. *For rare wet grassland habitat types (all except for 2.3.2/3), presence/absence maps derived from a threshold applied to the median maps of the predicted probability of occurrence were used in the combination procedure.

Table 1. Overview of the modelled permanent grassland habitat types. Abbreviation: abbreviation used throughout the manuscript to refer to the 20 modelled grassland habitat types. TypoCH: habitat typology according to Delarze et al. (2015); EUNIS: European Nature Information System taxonomies. This classification corresponds only partially to TypoCH, except if bold is used. NHV: Natural habitats worthy of protection in accordance with the Swiss Federal Ordinance on the Protection of Nature and Cultural Heritage. Short names for habitat groups (2nd level) used in the manuscript are given in round brackets.

Habitat types	Abbreviation	TypoCH	EUNIS	NHV
Fens				
<i>Magnocaricion elatae</i> W. Koch 26	2.2.1	2.2.1.1	C3.29, D5.21	x
<i>Cladietum marisci</i> All. 22 (<i>Magnocaricion</i> p.p)	2.2.1	2.2.1.2	C3.28, D5.24	x
<i>Caricion fuscae</i> W. Koch 26 em. Klika 34	2.2.2	2.2.2	D2.21, D2.22, D2.26	x
<i>Caricion davallianae</i> Klika 34	2.2.3	2.2.3	D4.1	x
Wet meadows				
<i>Molinion caerulea</i> W. Koch 26	2.3.1	2.3.1	E3.5	x
<i>Calthion palustris</i> Tüxen 37	2.3.2/3	2.3.2	D5.3, E3.4	x
<i>Filipenulion ulmariae</i> Lohmeyer 67	2.3.2/3	2.3.3	E5.42	x
Raised (open) bogs				
<i>Spagnion magellanici</i> Kästner et Flössner 33	2.4.1	2.4.1	D1.111, D1.12, X04	x
Re-seeded and heavy fertilized grasslands (Re-seeded grasslands)				
Ornamental lawns, football turfs, golf lawns, etc.	4.0.2	4.0.2		
Dry grasslands				
<i>Stipo-Poion carniolicae</i> Br.-Bl. 61 + <i>Stipo-Poion xerophilae</i> Br.-Bl. et Tüxen 43	4.2a	4.2.1.1	E1.24	x
<i>Cirsio-Brachypodium</i> Had. Et Klika 44	4.2a	4.2.1.2	E1.23	x
<i>Xerobromion</i> Br.-Bl. et Moor 38	4.2a	4.2.2	E1.27	x
<i>Diplachnion serotinae</i> Br.-Bl. 61	4.2a	4.2.3	E1.2	x
<i>Mesobromion</i> Br.-Bl. et Moor 38	4.2.4	4.2.4	E1.26	x
Nutrient-poor alpine and subalpine grasslands (Alpine/subalpine grasslands)				
<i>Seslerion caeruleae</i> Br.-Bl. 26	4.3.1	4.3.1	E4.431	
<i>Caricion firmae</i> Gams 36	4.3.2/4	4.3.2	E4.433	
<i>Caricion ferrugineae</i> Höhn 36	4.3.3	4.3.3	E4.41	x
<i>Elymion myosuroides</i> Gams 36	4.3.2/4	4.3.4	E4.42	x
<i>Nardion strictae</i> Br.-Bl. 26	4.3.5	4.3.5	E4.31	
<i>Festucion varia</i> Br.-Bl. 25	4.3.6	4.3.6	E4.33	
<i>Caricion curvulae</i> Br.-Bl. 26	4.3.7	4.3.7	E4.34	
Nutrient-rich pastures and meadows (Nutrient-rich pastures/meadows)				
<i>Arrhenatherion elatioris</i> W. Koch 26	4.5.1	4.5.1	E2.2	
<i>Polygono-Trisetion flavescens</i> Br.-Bl. et Tx 47	4.5.2	4.5.2	E2.3	
<i>Cynosurion cristati</i> Tüxen 47	4.5.3	4.5.3	E2.11, X09	
<i>Poion alpinae</i> Oberdorfer 50	4.5.4	4.5.4	E4.52	
Fallow grasslands				
<i>Convolvulo-Agropyron</i> Görs 66	4.6	4.6.1	-	
Abandoned grasslands with <i>Brachypodium pinnatum</i>	4.6	4.6.2	-	

nutrient-poor alpine and subalpine grasslands, nutrient-rich pastures and meadows as well as fallow grasslands (non-permanent types and those typically occurring only in scattered patches are excluded). In the following, the term 'wet grasslands' is used to refer to habitat types from the habitat class 2 (i.e., fens, wet meadows and raised bogs).

Data collection and preparation

In a two-step approach to data preparation, first, we collected various datasets from all over Switzerland, for which a habitat type classification was assigned by the data provider (for details, see Table 2). For the samples (i.e., point locations) of the 20 permanent grassland

habitat types to be modelled, we required a classification to the 3rd TypoCH level. In addition, we compiled samples from non-grassland habitat types, such as pioneer vegetation, shrubs or bushes, which served as further absences in the modelling of the individual grassland habitat types. For non-grassland samples, classification to the 3rd level was not a prerequisite. Overall, the compiled dataset (Fig. S2 in Supporting Information) included 50,096 samples, of which 35,566 and 14,530 were samples of grassland and non-grassland habitats, respectively.

In a second step, we performed a thorough selection of samples from the compiled dataset to derive both a training and a quasi-independent testing dataset, in the following referred to as dataset 1 and dataset 2, respectively.

Table 2. Datasets considered for the compilation of samples (compiled dataset). Data type: Data format provided by the data owner: cp: central coordinate of plot; pol: polygon; ps: (random) point sample (in polygon or without associated area). Pre-selection for this study: lo: last observation per plot/meadow for repeated surveys. For the testing data, only samples of the grassland habitats were considered.

Dataset	Full name	Focus habitat types; sampling design	Data type	Area [m ²]	Period of data collection	Pre-selection	Reference	Number of samples		
								Compiled dataset	Dataset 1	Dataset 2
BDM Z9	Swiss Biodiversity Monitoring program	all habitat types; regular grid across Switzerland	cp	10	2002-2018	lo	Weber et al. (2004)	1,505	1,417	25
ALL-EMA	Switzerland's farmland species and habitat monitoring program	agricultural landscape; 50x50 m grid in 1 km ² squares	cp	10	2015-2018	lo	www.allema.ch	24,719	1,248	742
TWW	National inventory of dry meadows and pastures	dry grasslands listed in the Federal Inventory; irregularly distributed across Switzerland	pol	326 - 508,966	1995-2014	Note S1 in Supporting Information	Eggenberg et al. (2001)	7,238	4,473	587
WBS /EKMS	Monitoring the effectiveness of habitat conservation in Switzerland / Swiss mire monitoring program	ferns and raised bogs listed in the Federal Inventory; irregularly distributed across Switzerland	cp / pol	10 / 8-22,451	2011-2017 / 1999-2011		Bergamini et al. (2013) / Grüning et al. (2005); Ecker et al. (2008)	1,887 / 9,167	407 / 602	145 / 428
BioAssembles / ModiPlant	Projects BioAssembles and ModiPlant	alpine ecosystems; vegetation plots in the Swiss Western Alps	cp / cp	4 / 4	2009 / 2002-2006		Dubuis et al. (2011) / Randin et al. (2006)	298 / 604	234 / 369	32 / 39
PermanentPlot	Swiss permanent vegetation plots	irregularly distributed across Switzerland	ps	1 - 800	1990-2019		Vittoz (2012)	958	276	60
AREA	Swiss land-use/land cover statistics	all habitat types; regular 100x100 m grid across Switzerland	ps	-	2012-2018	Note S1; Figure S1; Table S1	SFSO (2013)	2,253	1,987	208
Other		<i>Nardus</i> grasslands, hay meadows and other; irregularly distributed across Switzerland	ps	1 - 176,061	2010-2019	lo	Kurtogullari et al. (2020); van Klink et al. (2017); www.infofora.ch	1,467	650	15

Dataset 1 was used for the distribution modelling of the individual grassland habitat types, while dataset 2 was retained to assess the performance of the ensemble maps of the individual habitat types as well as the combined habitat map. To account for spatial autocorrelation (Dormann et al., 2007; Guisan et al., 2017; Ploton et al., 2020), we ensured that both datasets included only samples at least 300 m apart from each other across datasets. To this end, a data thinning routine was developed, iteratively selecting the most reliable samples based on a weighting following these criteria: (1) the number of samples within 300 m ('conflicting samples') of the respective sample, (2) year of observation, (3) habitat type, and, in the case of the TWW data (National inventory of dry meadows and pastures), (4) the compactness index. Thereby, samples featuring no or few conflicting samples were favored over samples with a high number of conflicting samples. Since habitat types may change over time, more recent samples were preferred over older samples. Further, rare grassland habitat types were favored over frequent ones to retain them in the thinned datasets (habitat type criterion). In the case of the TWW data, samples from compact polygons were given a higher weight in the selection routine than samples from narrow polygons because the first ones feature a purer remote sensing signal. For further details, see Figure S3.

The thinning routine was first applied to derive dataset 1, for which 8,140 samples of grassland habitats and 3,523 samples of non-grassland habitats were selected. The median year of observation was 2013 (Fig. S4). For dataset 2, only grassland samples were considered. Samples being part of dataset 1, and all samples of the same habitat type within a radius of 300 m of a dataset 1 sample were excluded (Ploton et al., 2020). Lastly, the thinning routine was applied to the remaining samples, selecting 2,281 samples with a median year of observation of 2012 for dataset 2.

Spatial predictors

We compiled 134 climatic, edaphic, topographic and Sentinel-based predictors covering various groups of ecophysiological essential parameters (Mod et al., 2016), from which, in a subsequent step, the most important were selected for each habitat type, excluding highly collinear variables (see *Pre-selection of spatial predictors*).

We prepared 17 climatic variables calculated by Zimmermann and Kienast (1999). In addition, we used continentality and maps of local soil properties (i.e., pH, nutrients, moisture, and soil moisture variability) generated by modelling ecological indicator values across Switzerland (Descombes et al., 2020). Topographic predictors included curvature, topographic position and

wetness indices, roughness as well as slope and aspect information (i.e., aspect sinus and cosinus; Baltensweiler et al., 2020).

From the Sentinel-1 Synthetic Aperture Radar (SAR) backscatter imagery (Torres et al., 2012), we applied two predictors based on composite γ^0 backscatter in vertical-vertical and vertical-horizontal polarization mode (Small et al., 2022; Waser et al., 2021).

From the Sentinel-2 imagery (Drusch et al., 2012), we used the orthorectified bottom-of atmosphere Level 2A data of the growing season (March–November) in the years 2017–2020. Processing was done in Google Earth Engine (Gorelick et al., 2017). We applied a two-step cloud masking procedure based on Sentinel-2 metadata and an additional cloud and shadow masking procedure based on a machine learning algorithm (s2cloudless algorithm) and sun elevation and azimuth, respectively (Pazúr et al., 2022; Zupanc, 2017). Further, we clipped each acquisition by an inside buffer of 1,000 m to avoid noise in the predictors caused by the overlapping acquisition pathways of Sentinel-2 (Fig. S5). We calculated a set of multi-annual predictors: (1) the medians of the individual bands (i.e., Blue, Green, Red, NIR, SWIR1 and SWIR2) over the whole studied time period (multiple growing seasons); (2) the medians, standard deviations, kurtosis, skewness and a set of percentiles of the NDVI and NDWI (normalized difference vegetation and water index, respectively) over the whole studied time period; and (3) seasonal medians and standard deviations of the NDVI and NDWI. The SWIR bands were resampled in Google Earth Engine with the default setting of nearest neighbor.

All non-Sentinel-2 predictors were bilinearly resampled to the 10×10 m spatial resolution of the Sentinel-2 predictors. For further details and a list of the predictors selected for the modeling, please see Table S2.

Maps of the individual grassland habitat types

Presence-absence data selection

For each habitat type, presences and absences were extracted from dataset 1 (Fig. 1). As presences, all samples allocated to the respective habitat type were used (Table S3). As absences, a set of samples was selected from all other samples (i.e., grassland and non-grassland samples) within dataset 1. To ensure a diverse set of absences with high discriminative power from both the realized and fundamental niche (Lobo et al., 2010), three types of absences were derived for each presence: (1) the three closest absences; (2) three absences at similar elevation; and (3) three random absences from all absences in dataset 1.

Pre-selection of spatial predictors

With the aim of achieving an accurate and robust spatial prediction per habitat type with a high discrimination between the grassland types, the 15 most important predictors were pre-selected for each habitat type (Table S4), avoiding multi-collinearity amongst the predictors in the model (Dormann et al., 2013). We followed a statistical approach (Guisan et al., 2017), where we assessed the predictive power of each predictor for each habitat type in a ten-fold split sampling of dataset 1 (training = 70%, testing = 30%), maintaining the ratio of presences and absences in each subset. Generalized linear models (GLMs) were fitted with a linear and a quadratic term for each predictor and the predictive power was assessed based on the average out-of-bag true skill statistic (TSS; Allouche et al., 2006). Predictors were iteratively selected starting with the best performing one and subsequently excluding highly correlated predictors (i.e., Pearson correlation coefficient $r > |0.7|$). In a second step, predictors with a variance inflation factor (VIF; Stine, 1995) > 5 were iteratively excluded (Guisan et al., 2017).

Habitat distribution modelling

To derive countrywide probability maps of the individual grassland habitat types, we applied an ensemble modeling approach (Araujo & New, 2007), which has been shown to improve robustness over individual algorithms (Hao et al., 2019; Seni & Elder, 2010). Here, we applied four common algorithms: random forest (RF; Breiman, 2001; using *randomForest*, Breimann et al., 2018), boosted regression trees (BRT; Elith et al., 2008; using *gbm*, Greenwell et al., 2020), generalized additive model (GAM; Hastie & Tibshirani, 1990; using *mgcv*, Wood, 2021), and GLMs (McCullagh & Nelder, 1989; using *MuMIn*, Barton, 2020). Presences and absences were assigned equal weights in model fitting (Barbet-Massin et al., 2012; Wisz & Guisan, 2009). For RFs, we used 5,000 trees and a stratified sampling to maintain the ratio of presences and absences. BRTs were run with 5,000 trees, up to 2-way interactions between predictors, at least 5 observations in the terminal node of the trees, and a shrinkage parameter of 0.01. For GAMs, a gamma value of 1.4 was applied. Binomial GLMs were fitted with a logit link and both a linear and quadratic term of each predictor. Thereby, we ensured a minimum number of at least 10 observations per predictor by restricting the number of pre-selected predictors if necessary to MaxP. Using the dredge function, all possible combinations were fitted with $0.5 * \text{MaxP} - \text{MaxP}$ predictors.

For each grassland habitat type (Fig. 1), we assessed model performance and derived countrywide ensemble

maps by ten-fold split-sampling of dataset 1 (training = 70%, testing = 30%), maintaining the ratio of presences and absences in each subset. For each algorithm and repeat, a model was fitted on the training data. To assess model performance, this model was used to predict the probability of occurrence for the held-out testing data. Subsequently, these probabilities were transformed into binary presence/absence predictions by optimizing the threshold maximizing TSS ($\text{threshold}_{\text{TSS_TestingData}}$). Lastly, TSS, sensitivity (proportion of correctly predicted presences) and specificity (proportion of correctly predicted absences) were calculated on the held-out testing data. On the other hand, the models fitted on the training data were used to calculate the predicted probability of occurrence across Switzerland. The resulting 40 (4 algorithms * 10 CV runs) maps were aggregated into (1) a map of the median of the predicted probability of occurrence, (2) a map of the standard deviation of the predicted probability of occurrence and (3) a binary presence/absence map. To derive aggregated presence/absence maps, each prediction map was first converted into an individual presence/absence map based on the $\text{threshold}_{\text{TSS_TestingData}}$. For the subsequent use of the aggregated map in the combination procedure, a presence was assigned to a specific pixel if more than half of the individual maps reported a presence, otherwise an absence was assigned. To assess the performance of the ensemble maps of the predicted probability of occurrence, and to compare it to that of the individual algorithms, TSS was calculated on dataset 2, which was not used for the modeling.

Combined grassland maps

Combination procedure

The 20 grassland habitat types were combined into two countrywide maps denoting the most likely (M1) and second most likely (M2) habitat type for each 10×10 m grid cell, respectively. For the combination procedure, we developed a weighted maximum probability approach based on iterative expert feedback (Figs. 1 and S6). In essence, for M1/M2, the habitat type with the highest/second highest weighted median of predicted probability of occurrence was assigned to each grid cell, respectively. The weights were assigned to each habitat type (Table S5) based on reviews by field experts. Thereby, the field experts visually inspected sections of the combined maps in an ArcMap Document in areas with which they were familiar. The weighting allowed increasing the abundance of underrepresented habitat types and decreasing the abundance of overrepresented ones. A habitat type was considered to be potentially present in a grid cell only if

the aggregated presence-absence map indicated a potential presence. For rare wet grassland habitat types, a stricter rule for potential presence was applied, based on a threshold applied to the median of the predicted probability of occurrence (Table S5). Lastly, regional adjustments were made to M1 in very wet and (rather) dry bioregions (BAFU, 2020; Swiss bioregions displayed in Figure 2) because large-scale patterns in soil moisture led to overestimations of wet habitats in wet bioregions and dry habitat types in dry bioregions (for details, see Note S2 and Fig. S7). The resulting final combined grassland map is referred to as M1F.

To analyze differences in the assignments to the grassland habitat types between M1 and M2, we further derived a map of the ratio of the probabilities of occurrence of the most (M1_P) and second most likely (M2_P) habitat types (MR). Values close to 1 indicate that the most and second most likely habitat types are almost equally likely to be present in the respective grid cell, while values clearly <1 indicate that the most likely habitat type is much more likely to be present than the second most likely one. A visualization of these combined maps is provided in Figures 2, 3 and S8.

In a last step, we masked out the following non-grassland habitats from all combined maps: forests (Waser et al., 2015), shrub forests (Rüetschi et al., 2021), hedges and groves, dwarf shrubs, crop rotation areas (Pazúr et al., 2022), areas with pioneer vegetation, settlements, roads, railways, areas with special use (e.g., allotment garden-areas), glaciers as well as stagnant and flowing waters (see Table S6 for mask dataset details).

Final combined map (M1F) evaluation

The accuracy of M1F was assessed in three ways. First, we compared the predicted habitat types with the observed habitat types of dataset 2, using confusion matrices. Sensitivities and specificities as well as geometric mean (G-Mean; squared root of the product of the sensitivity and specificity; Kubat et al., 1998) were derived for each habitat type, allowing an evaluation of the class-wise performance even if the distribution of the samples across the habitat types is skewed.

Second, we compared the predicted proportions of the habitat types with the estimated proportions from field

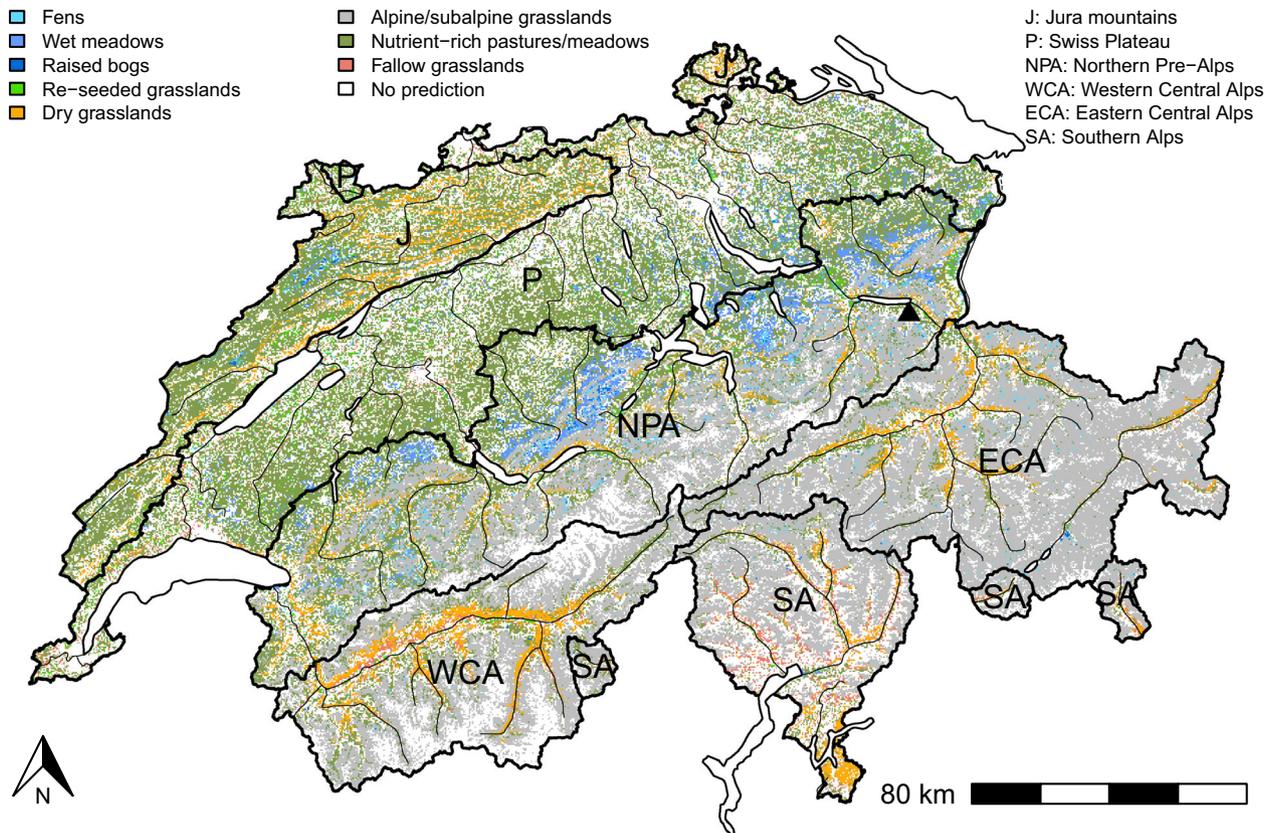


Figure 2. Habitat groups (2nd level) in the final combined map (M1F). For better visibility, the 20 grassland habitat types were colored according to their higher-level groups and only settlements were masked out. Biogeographical regions according to Gonseth et al. (2001) are shown (© FOEN). Lakes and rivers are mapped as white areas and grey lines, respectively. The black triangle marks the location of the small example map section visualized in Figure 3.

surveys in biotopes of national importance (WBS; “Wirkungskontrolle Biotopschutz Schweiz”; Monitoring the Effectiveness of Habitat Conservation in Switzerland; Bergamini et al., 2013). WBS monitors plant species on 10m² plots, which are randomly located within the biotopes of national importance. Data from the first survey run (i.e. 2011–2017: 799, 473 and 2,453 plots from fens, bogs and dry meadows and pastures, respectively) were used to estimate the proportions of 3rd level TypoCH habitat types within the national perimeters of respective fens, bogs and dry meadows and pastures using design based statistics (Tillé & Ecker, 2014).

Third, the predicted proportions of the habitat types were compared with their estimated Swiss-wide proportions based on data of the ALL-EMA program from 2015 to 2019 (Agricultural species and habitats’ monitoring program; www.allema.ch). ALL-EMA samples species and habitat data on 10 m² plots located on a 50 × 50m-grid in the agricultural area of 170 1 km² squares distributed across Switzerland. From this, Swiss-wide proportions were derived by first calculating the proportion of each habitat type with respect to all 361 plots per square, and then averaging this proportion across all squares, taking into account the different selection probabilities of the squares (Ecker et al., in prep).

Results

The resulting maps are available for download on the EnviDat platform under the following link (Huber et al. 2022): <https://doi.org/10.16904/envi.dat.341>.

Maps of the individual grassland habitat types

Pre-selected spatial predictors

Except for two habitat types, the predictors selected per habitat type covered all predictor groups (Fig. S9). The majority of the predictors originated from remote sensing (mean ± SD of the pre-selected predictors per habitat type: 6.9 ± 1.6) and topographic predictors (2.8 ± 2.2), while the latter were particularly relevant for wet and certain alpine/subalpine grasslands (Table S4). From the Sentinel-2 based indices, NDVI, NDWI, NIR and SWIR referring to the entire growing season were selected frequently (Table S4). Seasonal predictors were also frequently included in the modeling, particularly spring and summer medians and standard deviations of the NDVI, and were found to have high variable importance (assessed as the amount of accuracy the RF models lose by excluding the respective variable; Fig. S10). From the edaphic predictors, pH and soil nutrients were selected most often.

Habitat model performance

According to the 10-fold split-sampling of dataset 1, for most habitat types, the distribution models had useful to excellent predictive performance (i.e., median TSS ≥ 0.6; as in Coetzee et al., 2009; Fig. 4). The highest median TSS values were reached for raised bogs (2.4.1: 0.83), re-seeded (4.0.2: 0.85) and xeric grasslands (4.2a: 0.86). All of these habitats expressed both sensitivity and specificity values ≥ 0.88. Lowest median TSS values were found for fallow grasslands (4.6), nutrient-rich pastures/meadows, and two alpine/subalpine grasslands (i.e., 4.3.5 and 4.3.7).

The assessment based on the independent dataset 2 revealed a similar pattern (Fig. S11), with raised bogs (2.4.1: 0.86), re-seeded (4.0.2: 0.86) and one alpine/subalpine grassland (4.3.2/4: 0.94) reaching the highest ensemble TSS values. Lowest ensemble performance was found for fallow grasslands (4.6) and the nutrient-rich pasture 4.5.3. The best performing algorithm varied with habitat type (ensemble TSS value compared to the median TSS values of the individual algorithms), whereby the ensemble reached the highest or second highest performance among the algorithms for 15 of the 20 habitat types.

Countrywide maps for all modelled habitat types are shown in Figure S12.

Combined grassland maps

Large-scale patterns in the combined grassland map (M1F)

The final combined map featured distinct patterns across the biogeographical regions of Switzerland (Fig. 2). The Jura mountains and Swiss Plateau were predicted to be dominated by nutrient-rich pastures/meadows. Dry grasslands occurred more frequently in the Jura mountains than the Plateau. The mountainous regions were predicted to be dominated by alpine/subalpine grasslands either with nutrient-rich pastures/meadows or dry grasslands in the valley bottoms. Dry grasslands were predominantly present in dry and continental biogeographical regions. In contrast, the highest share of wet grasslands was predicted in the Northern Pre-Alps.

Comparison of most likely (M1) and second most likely (M2) habitat types

For most grid cells, the ratio of the predicted probability of occurrence of the most to the second most likely habitat type was high (i.e., mean of the medians = 0.895; Fig. S13). For three of the wet grassland types, the second most likely habitat type was most often another wet

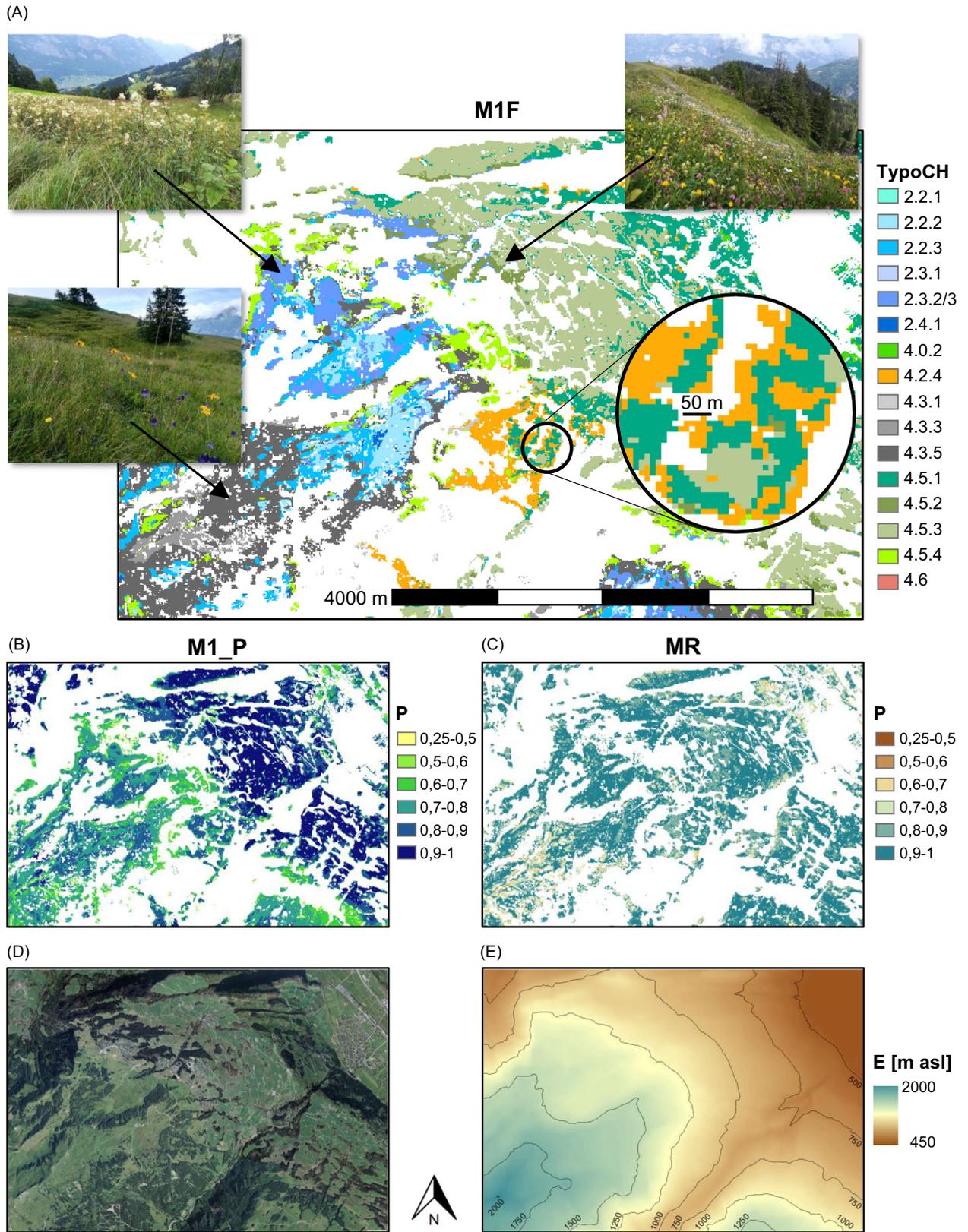


Figure 3. Visualization of three combined grassland mapping products for a small example map section. (A) Final combined map of the most likely habitat type (M1F). TypoCH: habitat typology according to Delarze et al. (2015). For further details, see Table 1. (B) Map of the weighted median of the predicted probability of occurrence (P) of the most likely grassland habitat type (M1_P). (C) Map of the ratio of the probabilities of occurrence of the most and second most likely grassland habitat types (MR). (D) Orthophoto (SWISSIMAGE 25) of the example map section Flumserberg © swisstopo. (E) Elevation (E; DHM25 L2 © swisstopo).

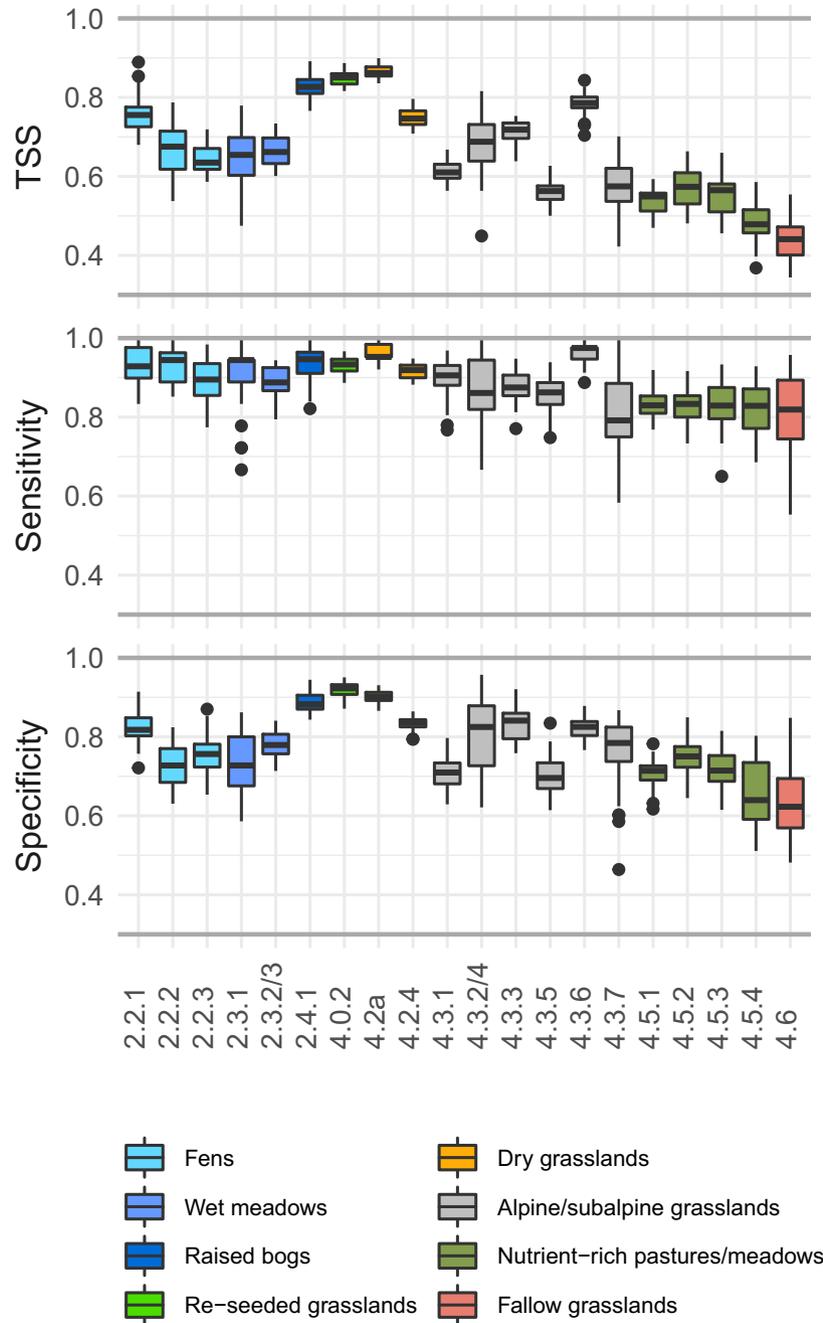


Figure 4. Model performance for the individual habitat types based on ensemble models with four algorithms and 10-fold split-sampling of data-set 1 (N = 4*10) assessed by True Skill Statistic (TSS), sensitivity and specificity. For an overview of the modelled habitat types, see Table 1. Colors according to the higher-level habitat groups (2nd level).

grassland (Fig. 5, Table S7). For other wet grasslands, alpine/subalpine grasslands (predominantly 4.3.5) as well as nutrient-rich pastures/meadows covered a high share of the second most likely habitat types. The second most likely habitat types of dry grasslands were mostly alpine/subalpine grasslands, nutrient-rich pastures/meadows or fallow grasslands, but rarely the other dry grassland type. For alpine/subalpine grasslands, the second most likely habitat type was mainly from the same group, except for 4.3.5, for which approximately half of the second most likely habitat and around 5% was assigned to the nutrient-rich pasture 4.5.4 and wet grasslands, respectively. In return, also for 4.5.4, around 50% of the second most likely habitat was assigned to 4.3.5. For the other nutrient-rich pastures/meadows, the second most likely habitat was most often another nutrient-rich pasture/meadow. The same was found for re-seeded grasslands (4.0.2). For fallow grasslands, the second most likely habitat types covered all groups except the ones belonging to wet grasslands.

Performance of the combined habitat map (M1F)

The highest G-Mean values were achieved for re-seeded grasslands as well as one dry and one alpine/subalpine grassland (Table 3: first column). Low G-Mean values below 0.5 were derived for three wet grasslands, one

alpine/subalpine grassland, two nutrient-rich pastures/meadows and particularly for fallow grasslands.

Many omission errors occurred within the same habitat group (2nd level) or at least the same habitat class (i.e., within wet grasslands; Table S8, Fig. S14). Omissions across groups mainly concerned the following habitat types: fen type 2.2.3, where more than half of the omissions were assigned to wet meadows, and wet-meadow type 2.3.2/3 and raised bogs, where omissions were often allocated to fens. Omissions of re-seeded grasslands (4.0.2) and the dry-meadow type 4.2.4 were predominantly assigned to nutrient-rich pastures/meadows, while those of the nutrient-rich pastures/meadows-type 4.5.4 were largely allocated to alpine/subalpine grasslands (particularly to 4.3.5).

We found good agreement between the predicted proportions of the habitat types in M1F and the estimated proportions from the WBS and ALL-EMA data ($\rho = 0.905, 0.856, 0.818$ and 0.846 for fens, bogs, dry meadows and pastures and the estimated Swiss-wide proportions, respectively; Table 3). The largest discrepancies occurred for raised bogs, which featured a distinct underestimation.

The comparison of M1F with the projections from the WBS for fens (Table 3: second column) indicated a slight underestimation of wet habitats in areas of national importance. The alpine/subalpine grassland 4.3.5 and the

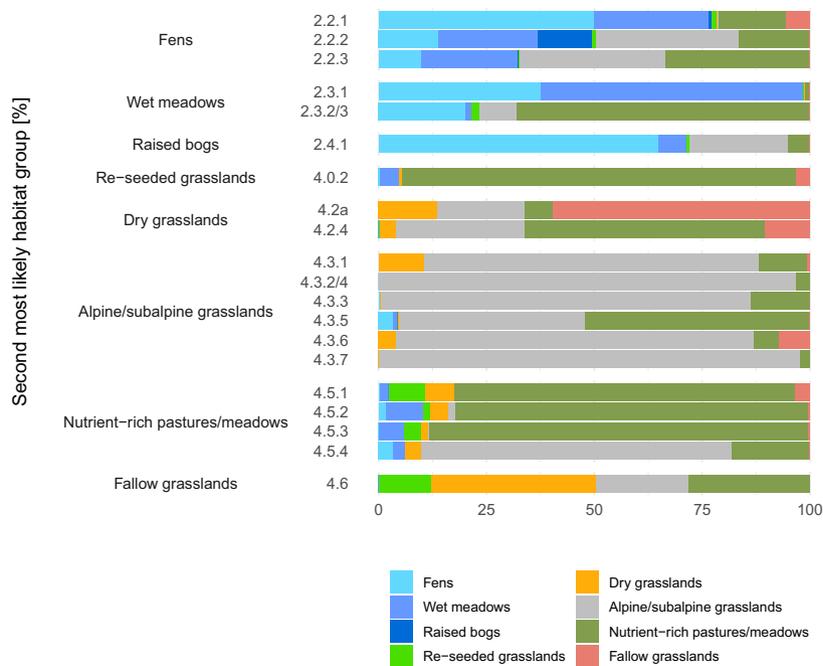


Figure 5. Comparison of the combined maps depicting the most and second most likely habitat type (M1/M2), respectively. Percent distributions of the second most likely habitat groups (2nd level; %, color) for each habitat type in M1 (rows). For an overview of the modelled habitat types, see Table 1.

Table 3. Evaluation of the final combined map (M1F) based on G-Mean (squared root of the product of the sensitivity and specificity for dataset 2) and a comparison with the projections for the biotopes of national importance (fens, bogs and dry meadows and pastures (TWW)) and the entire area of Switzerland. The projections are based on the WBS (Monitoring the Effectiveness of Habitat Conservation in Switzerland) and ALL-EMA (Switzerland's agricultural species and habitats' monitoring program) data. For the projections, the numbers are given in percentage of the respective area.

	G-Mean M1F	Fens M1F / WBS	Bogs M1F / WBS	TWW M1F / WBS	Switzerland M1F / ALL-EMA
Fens					
2.2.1	0.45	8.46 / 9.28	3.25 / 5.05	0.00 / 0.03	0.05 / 0.04
2.2.2	0.51	5.89 / 7.73	20.39 / 3.42	0.02 / 0.02	0.15 / 0.15
2.2.3	0.43	16.79 / 18.43	3.66 / 0.87	0.19 / 0.32	0.46 / 0.32
Wet meadows					
2.3.1	0.39	4.49 / 5.96	0.82 / 0.21	0.00 / 0.27	0.03 / 0.09
2.3.2/3	0.61	27.13 / 25.83	5.24 / 4.45	0.02 / 0.50	0.70 / 1.2
Raised bogs					
2.4.1	0.57	1.40 / 5.08	33.26 / 49.96	0.01 / 0.03	0.02 / 0.25
Re-seeded and heavy fertilized grassland					
4.0.2	0.88	0.31 / 0	0.31 / 0	0.11 / 0.00	0.55 / 0
Dry grasslands					
4.2a	0.53	0 / 0	0 / 0	3.65 / 4.14	0.12 / 0.04
4.2.4	0.74	0.36 / 0.21	0 / 0	25.49 / 26.65	1.03 / 0.97
Nutrient-poor alpine and subalpine grasslands					
4.3.1	0.47	0.56 / 0	0.03 / 0	15.74 / 1.38	1.43 / 1.31
4.3.2/4	0.66	0 / 0	0 / 0	0.04 / 0.29	0.70 / 0.17
4.3.3	0.61	0.11 / 0	0 / 0	9.24 / 9.75	2.61 / 0.66
4.3.5	0.51	9.23 / 3.34	1.77 / 0.47	6.33 / 7.21	3.77 / 2.74
4.3.6	0.54	0.21 / 0	0.07 / 0	6.83 / 1.61	1.01 / 0.18
4.3.7	0.51	0.08 / 0.07	0.01 / 0	0.09 / 0.07	1.32 / 0.66
Nutrient-rich pastures and meadows					
4.5.1	0.62	2.65 / 0.94	0.29 / 0	3.13 / 15.22	8.22 / 13.57
4.5.2	0.44	1.76 / 0.41	0.13 / 0	0.48 / 2.46	1.01 / 2.01
4.5.3	0.55	3.11 / 0.91	0.46 / 0	1.10 / 2.80	4.43 / 7.08
4.5.4	0.43	2.90 / 1.58	1.39 / 0	3.84 / 5.01	1.93 / 4.41
Fallow grasslands					
4.6	0.16	0.16 / 0.22	0.09 / 0	0.37 / 0.12	0.16 / 0.27
Other habitat types		14.41 / 20.01	28.83 / 35.57	23.32 / 22.81	70.27 / 63.88

nutrient-rich pastures/meadows showed an overestimation in M1F compared to the WBS projections, which was also found for the national perimeter of bogs (Table 3: third column). For bogs, it is further noticeable that the predicted share of the fen type 2.2.2 was distinctly overestimated (20.33% vs. 3.42%). For the national inventory of dry meadows and pastures (TWW; Table 3: fourth column), the two alpine/subalpine grasslands 4.3.1 and 4.3.6 featured strong overestimations. In contrast, the nutrient-rich pastures/meadows, particularly 4.5.1, appeared to be strongly underestimated.

The comparison with the Swiss-wide projections from the ALL-EMA data (Table 3: last column) revealed good agreement of the predicted shares for dry and most wet grasslands. The wet meadows showed an underestimation and most alpine/subalpine grasslands featured higher shares in M1F than the ALL-EMA projection. Yet, the most common 4.3.5 alpine/subalpine grassland showed

good agreement with the ALL-EMA projections (3.77% vs. 2.74%) as well as 4.3.1 (1.43% vs. 1.31%), which was strongly overestimated according to the WBS projections. The nutrient-rich pastures/meadows appeared to be underestimated.

Discussion

We successfully modelled the distribution of 20 permanent grassland habitat types at the level of phytosociological alliances across Switzerland using a large and unique high-quality field inventory dataset and spatial data, in particular, spatial and temporal indices from Copernicus Sentinel satellite imagery. Our study provides a comprehensive set of high spatial resolution (10x10 m) maps of the distribution of grassland habitats for a large area of heterogeneous terrain, climate and land use, offering a wide range of applications for biodiversity conservation.

The maps could (1) serve as input for further modelling, that is, in the context of species distribution models, ecosystem service mapping, the identification of transition zones between habitat types, or connectivity analyses; (2) be integrated into or combined with other habitat maps; (3) be used in monitoring projects, for example, to inform and plan field data collection campaigns; or (4) be applied for conservation planning, for example, to build ecological networks or extend conservation areas.

Spatial predictors

Predictors from all groups, that is, remote sensing, climate, soil and topography, were relevant for the vast majority of habitat types indicating that they capture complementary and ecophysiological aspects of habitat distributions (Mod et al., 2016). Satellite-based indices, especially those characterizing moisture content as well as vegetation greenness/density and its seasonality, were highly relevant for the modelling. The seasonal predictors likely allow a better discrimination of similar grassland types by capturing relevant differences in their phenology (Marcinkowska-Ochtyra et al., 2019; Tarantino et al., 2021). Topographic predictors turned out to be particularly relevant for wet grasslands and alpine/subalpine grasslands occurring on special terrain features, such as mountain ridges. These predictors were proxies for small- and large-scale patterns of soil water availability and topography-related energy (i.e., inducing erosion or accumulation processes), which are highly variable in landscapes with rugged topography (Scherrer & Guisan, 2019). From the edaphic predictors, pH and soil nutrients were selected most often and thus seem to be key to determining habitat distributions, particularly in areas with contrasting edaphic conditions (Dubuis et al., 2013; Scherrer & Guisan, 2019).

Combining habitat maps

We used a two-step approach, where the ensemble modeling outputs of the individual habitats were assembled into combined maps by integrating weights derived from an iterative expert elicitation process. This transparent and flexible approach has a number of advantages. First, in addition to the combined maps, maps of all individual habitat types are available and can be used, analyzed and updated independently. Furthermore, the combined maps could be updated when improved maps of individual habitat types become available. Third, expert weighting of the predicted probabilities of occurrence allowed finding more realistic thresholds between closely related habitat types and thus adjusting the abundance of habitat types in the combined map. Finally, we accounted for

uncertainties of the combination procedure by providing maps of the most and second most likely habitat type, respectively, as well as a map of the ratio of the probabilities of the two alternatives.

Performances of the grassland maps

Although, the best performing algorithm varied with habitat type, the ensemble represented the best or second best algorithm for the majority of the habitat types. This is in line with previous studies comparing different algorithms (e.g., Bouska et al., 2014; Norberg et al., 2019) and underlines the value of using ensembles.

For wet grasslands, most individual type maps performed well, while their relative performance to the other habitats dropped in the combined map M1F. Omissions predominantly occurred with other wet grassland habitats. This can be expected since wet habitats often occur in small-scale mosaics (Küchler et al., 2004). Of all wet grassland habitats, the individual map for raised bogs performed particularly well, but was challenging to integrate in the combined maps, especially regarding transitions to surrounding fens. Strict rules were required to avoid an overestimation of this rare and highly protected habitat. Moreover, a strict masking was applied to exclude shrub and treed bogs, which at least partly explains the underestimation of raised bogs in comparison to the WBS and ALL-EMA projections. Transitions between oligotrophic fens (particularly 2.2.3: *Caricion davallianae*) and nutrient-rich wet meadows (i.e., 2.3.2/3: *Calthion/Filipendulion*) were challenging to map but are particularly relevant for biodiversity conservation. Eutrophication represents one of the biggest current threats to wetlands (Rion et al., 2018), associated with transformations from fens towards nutrient-rich wet meadows (Bergamini et al., 2009).

For other permanent grasslands, performance was very high for both the individual and combined maps of reseeded and heavy fertilized grasslands (4.0.2). Omissions predominantly occurred with the mesic and nutrient-rich *Arrhenatherion* (4.5.1), indicating the greatest similarity with nutrient-rich meadows. The lowest performance was for fallow grassland maps, potentially because they represent a spectrum of grassland types with lack of management rather than an actual phytosociological unit (Delarze et al., 2015). Performance of dry grassland maps was medium to high. The four xeric types (modelled together as 4.2a) separated well from the semi-dry *Mesobromion* (4.2.4), which, however, often featured transitions to the nutrient-rich *Arrhenatherion* (4.5.1). The comparison with ALL-EMA indicated that some alpine/subalpine grasslands may be overestimated. However, this comparison is restricted due to the predominant

focus of ALL-EMA on the agricultural area. The high share of both mutual omissions and assignments to most and second most likely habitat types of the extensively managed alpine/subalpine grassland *Nardion* (4.3.5) and the nutrient-rich pasture *Poion alpinae* (4.5.4) indicates that the transition between these grassland types was difficult to map. On the one hand, *Poion alpinae* often co-occurs in small-scale mosaics with nutrient-poor alpine/subalpine grasslands and wetlands (Delarze et al., 2015), while, on the other hand, *Nardion* pastures gradually degrade into *Poion alpinae* when grazed intensively (Kurtogullari et al., 2020). In contrast, *Poion alpinae* was separated relatively well from the other nutrient-rich pastures/meadows, likely because it occurs at higher elevations. However, the other nutrient-rich pasture/meadow types were more difficult to distinguish from each other, likely because management drives their floristic composition. Yet, mixed uses, such as cutting followed by grazing, are common practice and may change between years, resulting in intermediate habitat types (Delarze et al., 2015).

Limitations

Prediction accuracies were assessed for all habitat types at the national extent using a comprehensive set of measures. Yet, accuracies likely differ across regions, particularly towards extreme conditions (i.e., dry, wet or high elevation). Additional measures, such as the area of applicability (Meyer & Pebesma, 2021), could be used to identify regions with particularly uncertain predictions. Region-specific modeling efforts may improve the maps provided here, however they likely require additional vegetation samples to better capture the habitat-environmental relationships. Alternatively, regional deficiencies (e.g., underestimations of wet habitats in dry bioregions) may represent niche truncations due to spatial sampling bias, which could be addressed with transnational vegetation samples (Chevalier et al., 2021).

As the vegetation samples were collected within different surveys, the assignment of vegetation plots to specific habitat types likely differed among data sources. This may, for instance, explain the contradictory assessment of the two projections (WBS, ALL-EMA) with regard to the percent area of the alpine/subalpine grassland *Seslerion* (4.3.1). Where available, homogenizing the assignment based on the species' presence and abundance could resolve this issue.

Acknowledgements

We thank all data providers, particularly Christoph Bühler, Jean-Yves Humbert, Pascal Vittoz and the Swiss Federal Statistical Office. We are grateful to Stefan

Eggenberg for his advice and expertise of the grassland habitat types and for his reviews of the mapping products. We also thank Céline Richter and the Pro Natura team as well as Ariel Bergamini for their field expert reviews. We thank Rafael Wüest Karpati for sharing data and R scripts and two anonymous reviewers for their helpful comments. Funding was provided by the Federal Office for the Environment.

References

- Ali, I., Cawkwell, F., Dwyer, E., Barrett, B. & Green, S. (2016) Satellite remote sensing of grasslands: from observation to management. *Journal of Plant Ecology*, **9**, 649–671.
- Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.
- Araujo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.
- BAFU. (2020) *Datenbeschreibung biogeographische regionen*. Bern: BAFU.
- Baltensweiler, A., Heuvelink, G.B.M., Hanewinkel, M. & Walthert, L. (2020) Microtopography shapes soil pH in flysch regions across Switzerland. *Geoderma*, **380**, 114663 (10 pp.).
- Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.
- Barton, K. (2020) *Package 'MuMIn' - Multi-Model Inference*. CRAN. Available at: <https://CRAN.Rproject.org/package=MuMIn> [Accessed 28nd September 2020].
- Bergamini, A., Ginzler, C., Schmidt, B., Kuchler, M. & Holderegger, R. (2013) Monitoring the effectiveness of habitat conservations: making changes visible. *Hotspot*, **28**, 18–19.
- Bergamini, A., Peintinger, M., Fakheran, S., Moradi, H., Schmid, B. & Joshi, J. (2009) Loss of habitat specialists despite conservation management in fen remnants 1995–2006. *Perspectives in Plant Ecology, Evolution and Systematics*, **11**, 65–79.
- Bhatnagar, S., Gill, L., Regan, S., Naughton, O., Johnston, P., Waldren, S. et al. (2020) Mapping vegetation communities inside wetlands using sentinel-2 imagery in Ireland. *International Journal of Applied Earth Observation and Geoinformation*, **88**, 102083 (13 pp.).
- Boch, S., Kurtogullari, Y., Allan, E., Lessard-Therrien, M., Rieder, N.S., Fischer, M. et al. (2021) Effects of fertilization and irrigation on vascular plant species richness, functional composition and yield in mountain grasslands. *Journal of Environmental Economics and Management*, **279**, 111629.
- Bollens, U., Gusewell, S. & Klotzli, F. (2001) Vegetation changes in two Swiss fens affected by eutrophication and desiccation. *Botanica Helvetica*, **111**, 121–137.

- Bouska, K.L., Whitley, G.W. & Lant, C. (2014) Development and evaluation of species distribution models for fourteen native central U.S. fish species. *Hydrobiologia*, **747**, 159–176.
- Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5–32.
- Breiman, L., Cutler, A., Liaw, A. & Wiener, M. (2020) *Package 'randomForest' - Breiman and Cutler's Random Forests for classification and regression*. CRAN. R package. Available at: <https://CRAN.Rproject.org/package=randomForest> [Accessed 28th September 2020].
- Chevalier, M., Broennimann, O., Cornuault, J. & Guisan, A. (2021) Data integration methods to account for spatial niche truncation effects in regional projections of species distribution. *Ecological Applications*, **31**, e02427.
- Coetzee, B.W.T., Robertson, M.P., Erasmus, B.F.N., VAN Rensburg, B.J. & Thuiller, W. (2009) Ensemble models predict Important Bird Areas in southern Africa will become less effective for conserving endemic birds under climate change. *Global Ecology and Biogeography*, **18**, 701–710.
- Corbane, C., Lang, S., Pipkins, K., Alleaume, S., Deshayes, M., García Millán, V.E. et al. (2015) Remote sensing for mapping natural habitats and their conservation status – New opportunities and challenges. *International Journal of Applied Earth Observation and Geoinformation*, **37**, 7–16.
- Davies, C.E., Moss, D. & Hill, M.O. (2004) EUNIS habitat classification revised 2004. In: *Report to the European Environment Agency*. Available at: <http://www.eea.europa.eu/data-and-maps/data/eunis-habitatclassification1/documentation/eunis-2004-report.pdf> [Accessed 21st October 2021].
- Delarze, R., Eggenberg, S., Steiger, P., Bergamini, A., Fivaz, F., Gonseth, Y. et al. (2016) Rote Liste Lebensräume der Schweiz. In: *Aktualisierte Kurzfassung zum technischen Bericht 2013 im Auftrag des Bundesamtes für Umwelt (BAFU)*. Bern: BAFU, p. 33.
- Delarze, R., Gonseth, Y., Eggenberg, S. & Vust, M. (2015) *Lebensräume der Schweiz. Ökologie - Gefährdung - Kennarten*, Bern, hep verlag ag. Bern: hep verlag, p. 456.
- Descombes, P., Walthert, L., Baltensweiler, A., Meuli, R.G., Karger, D.N., Ginzler, C. et al. (2020) Spatial modelling of ecological indicator values improves predictions of plant distributions in complex landscapes. *Ecography*, **43**, 1448–1463.
- Devillers, P., Devillers-Terschuren, J. & Ledant, J.-P. (1991) *CORINE biotopes manual. Habitats of the European Community*. Luxembourg: Commission of the European Communities.
- Diaz Varela, R.A., Ramil Rego, P., Calvo Iglesias, S. & Munoz Sobrino, C. (2008) Automatic habitat classification methods based on satellite images: a practical assessment in the NW Iberia coastal mountains. *Environmental Monitoring and Assessment*, **144**, 229–250.
- Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G. et al. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.
- Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carre, G. et al. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 27–46.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F. et al. (2012) Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, **120**, 25–36.
- Dubuis, A., Giovanettina, S., Pellissier, L., Pottier, J., Vittoz, P., Guisan, A. et al. (2013) Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables. *Journal of Vegetation Science*, **24**, 593–606.
- Dubuis, A., Pottier, J., Rion, V., Pellissier, L., Theurillat, J.-P. & Guisan, A. (2011) Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. *Diversity and Distributions*, **17**, 1122–1131.
- Ecker, K., Kuchler, M., Feldmeyer-Christe, E., Graf, U. & Waser, L. (2008) Predictive mapping of floristic site conditions across mire habitats: Evaluating data requirements. *Community Ecology*, **9**, 133–146.
- Eggenberg, S., Dalang, T., Dipner, M. & Mayer, C. (2001) *Kartierung und Bewertung der Trockenwiesen und -weiden von nationaler Bedeutung. Technischer Bericht. Schriftenreihe Umwelt*. Bern: Bundesamt für Umwelt, Wald und Landschaft BUWAL.
- Elith, J., Leathwick, J.R. & Hastie, T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802–813.
- EUROSTAT. (2021) *Land cover statistics [Online]*. Luxembourg: eurostat. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?eurostatitle=Land_cover_statistics#Land_cover_in_the_EU [Accessed 17th September 2021].
- Feurdean, A., Ruprecht, E., Molnar, Z., Hutchinson, S.M. & Hickler, T. (2018) Biodiversity-rich European grasslands: ancient, forgotten ecosystems. *Biological Conservation*, **228**, 224–232.
- Gillet, F., Mauchamp, L., Badot, P.M. & Mouly, A. (2016) Recent changes in mountain grasslands: a vegetation resampling study. *Ecology and Evolution*, **6**, 2333–2345.
- Gonseth, Y., Wohlgemuth, T., Sansonnens, B. & Buttler, A. (2001) *Die biogeographischen Regionen der Schweiz. Erläuterungen und Einteilungsstandard. Umwelt Materialien*. Bern: Bundesamt für Umwelt, Wald und Landschaft.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. & Moore, R. (2017) google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, **202**, 18–27.

- Graf, U., Wildi, O., Kuchler, M. & Ecker, K. (2010) Five-year changes in Swiss mire vegetation. *Botanica Helvetica*, **120**, 15–27.
- Greenwell, B., Boehemke, B., Cunningham, J. & Developers, G. (2020). Package 'gbm' - generalized boosted regression models. Available at: <https://CRAN.R-project.org/package=gbm> [Accessed 28th September 2020].
- Grünig, A., Steiner, G.M., Ginzler, C., Graf, U. & Kuchler, M. (2005) Approaches to Swiss Mire Monitoring. *Stapfia*, **85**, 435–452.
- Guisan, A., Thuiller, W. & Zimmermann, N.E. (2017) *Habitat suitability and distribution models*. Cambridge: Cambridge University Press.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Hao, T., Elith, J., Guillera-Arroita, G., Lahoz-Monfort, J.J. & Serra-Diaz, J. (2019) A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Diversity and Distributions*, **25**, 839–852.
- Hastie, T. & Tibshirani, R. (1990) *Generalized additive models*. London: Chapman & Hall.
- He, K.S., Bradley, B.A., Cord, A.F., Rocchini, D., Tuanmu, M.N., Schmidlein, S. et al. (2015) Will remote sensing shape the next generation of species distribution models? *Remote Sensing in Ecology and Conservation*, **1**, 4–18.
- Jongman, R.H.G., Bouwma, I.M., Griffioen, A., Jones-Walters, L. & VAN Doorn, A.M. (2011) The Pan European Ecological Network: PEEN. *Landscape Ecology*, **26**, 311–326.
- Kubat, M., Holte, R.C. & Matwin, S. (1998) Machine learning for the detection of oil spills in satellite radar images. *Machine Learning*, **30**, 195–215.
- Kuchler, M., Ecker, K., Feldmeyer-Christe, E., Graf, U., Kuchler, H. & Waser, L.T. (2004) Combining remotely sensed spectral data and digital surface models for fine-scale modelling of mire ecosystems. *Community Ecology*, **5**, 55–68.
- Kuhn, T., Domokos, P., Kiss, R. & Ruprecht, E. (2021) Grassland management and land use history shape species composition and diversity in Transylvanian semi-natural grasslands. *Applied Vegetation Science*, **24**, e12585 (12 pp).
- Kurtogullari, Y., Rieder, N.S., Arlettaz, R. & Humbert, J.Y. (2020) Conservation and restoration of *Nardus* grasslands in the Swiss northern Alps. *Applied Vegetation Science*, **23**, 26–38.
- Lobo, J.M., JIMENEZ-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103–114.
- Maiorano, L., Falcucci, A., Garton, E.O. & Boitani, L. (2007) Contribution of the Natura 2000 network to biodiversity conservation in Italy. *Conservation Biology*, **21**, 1433–1444.
- Marcinkowska-Ochtyra, A., Gryguc, K., Ochtyra, A., Kopeć, D., Jarocińska, A. & Sławik, Ł. (2019) Multitemporal hyperspectral data fusion with topographic indices—improving classification of natura 2000 grassland habitats. *Remote Sensing*, **11**, 2264 (22 pp.).
- Mccullagh, P. & Nelder, J.A. (1989) *Generalized linear models*. London: Chapman & Hall.
- Mehner, H., Cutler, M., Fairbairn, D. & Thompson, G. (2004) Remote sensing of upland vegetation: the potential of high spatial resolution satellite sensors. *Global Ecology and Biogeography*, **13**, 359–369.
- Meyer, H. & Pebesma, E. (2021) Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution*, **12**, 1620–1633.
- Mod, H.K., Scherrer, D., Luoto, M., Guisan, A. & Scheiner, S. (2016) What we use is not what we know: environmental predictors in plant distribution models. *Journal of Vegetation Science*, **27**, 1308–1322.
- Norberg, A., Abrego, N., Blanchet, F.G., Adler, F.R., Anderson, B.J., Anttila, J. et al. (2019) A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs*, **89**, e01370.
- Pazúr, R., Huber, N., Weber, D., Ginzler, C. & Price, B. (2022) AA national extent map of cropland and grassland for Switzerland based on Sentinel-2 data. *Earth System Science Data*, **14**, 295–305.
- Peppler-Lisbach, C., Stanik, N., Könitz, N., Rosenthal, G. & Vandvik, V. (2020) Long-term vegetation changes in *Nardus* grasslands indicate eutrophication, recovery from acidification, and management change as the main drivers. *Applied Vegetation Science*, **23**, 508–521.
- Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V. et al. (2020) Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature communications*, **11**, 1–11.
- Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M. & Guisan, A. (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography*, **33**, 1689–1703.
- R Core Team. (2018) *R: a language environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rion, V., Gallandat, J.-D., Gobat, J.-M., Vittoz, P. & Jurasinski, G. (2018) Recent changes in the plant composition of wetlands in the Jura Mountains. *Applied Vegetation Science*, **21**, 121–131.
- Rüetschi, M., Weber, D., Koch, T.L., Waser, L.T., Small, D. & Ginzler, C. (2021) Countrywide mapping of shrub forest using multi-sensor data and bias correction techniques. *International Journal of Applied Earth Observation and Geoinformation*, **105**, 102613 (10 pp.).
- Scherrer, D. & Guisan, A. (2019) Ecological indicator values reveal missing predictors of species distributions. *Scientific Reports*, **9**, 3061.

- Seni, G. & Elder, J. F. 2010. Ensemble methods in data mining: improving accuracy through combining predictions. *Synthesis Lectures on Data Mining and Knowledge Discovery*, 2(1), 1–126. <https://doi.org/10.2200/S00240ED1V01Y200912DMK002>
- SFSO. (2013) *Land use in Switzerland: results of the Swiss land use statistics*. Neuchâtel: Federal Statistical Office.
- Small, D., Rohner, C., Miranda, N., Ruetschi, M. & Schaepman, M.E. (2022) Wide-area analysis-ready radar backscatter composites. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–14.
- Soderstrom, B., Svensson, B., Vessby, K. & Glimskar, A. (2001) Plants, insects and birds in semi-natural pastures in relation to local habitat and landscape factors. *Biodiversity and Conservation*, 10, 1839–1863.
- Standen, V. (2000) The adequacy of collecting techniques for estimating species richness of grassland invertebrates. *Journal of Applied Ecology*, 37, 884–893.
- Stine, R.A. (1995) Graphical Interpretation of Variance Inflation Factors. *American Statistician*, 49, 53–56.
- Tarantino, C., Forte, L., Blonda, P., Vicario, S., Tomaselli, V., Beierkuhnlein, C. et al. (2021) Intra-annual sentinel-2 time-series supporting grassland habitat discrimination. *Remote Sensing*, 13, 277 (29 pp.).
- Tillé, Y. & Ecker, K. (2014) Complex national sampling design for long-term monitoring of protected dry grasslands in Switzerland. *Environmental and Ecological Statistics*, 21, 453–476.
- Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E. et al. (2012) GMES Sentinel-1 mission. *Remote Sensing of Environment*, 120, 9–24.
- Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E. & Steininger, M. (2003) Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution*, 18, 306–314.
- Van Klink, R., Boch, S., Buri, P., Rieder, N.S., Humbert, J.-Y. & Arlettaz, R. (2017) No detrimental effects of delayed mowing or uncut grass refuges on plant and bryophyte community structure and phytomass production in low-intensity hay meadows. *Basic and Applied Ecology*, 20, 1–9.
- Vanden Borre, J., Paelinckx, D., Mûcher, C.A., Kooistra, L., Haest, B., De Blust, G. et al. (2011) Integrating remote sensing in Natura 2000 habitat monitoring: prospects on the way forward. *Journal for Nature Conservation*, 19, 116–125.
- Vittoz, P. (2012) Permanent.Plot.ch – a database for Swiss permanent vegetation plots. *Biodiversity and Ecology*, 4, 337–337.
- Waser, L.T., Fischer, C., Wang, Z.Y. & Ginzler, C. (2015) Wall-to-wall forest mapping based on digital surface models from image-based point clouds and a NFI forest definition. *Forests*, 6, 4510–4528.
- Waser, L.T., Rüetschi, M., Psomas, A., Small, D. & Rehus, N. (2021) Mapping dominant leaf type based on combined Sentinel-1/-2 data – challenges for mountainous countries. *ISPRS Journal of Photogrammetry and Remote Sensing*, 180, 209–226.
- Weber, D., Hintermann, U. & Zangger, A. (2004) Scale and trends in species richness: considerations for monitoring biological diversity for political purposes. *Global Ecology and Biogeography*, 13, 97–104.
- Willems, J.H., Peet, R.K. & Bik, L. (1993) Changes in chalk-grassland structure and species richness resulting from selective nutrient additions. *Journal of Vegetation Science*, 4, 203–212.
- Wisz, M.S. & Guisan, A. (2009) Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, 9, 8.
- Wood, S. (2021). Package 'mgcv'. Available at: <https://CRAN.R-project.org/package=mgcv> [Accessed 28th September 2020].
- Wright, C.K. & Wimberly, M.C. (2013) Recent land use change in the Western Corn Belt threatens grasslands and wetlands. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 4134–4139.
- Zimmermann, N.E. & Kienast, F. (1999) Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *Journal of Vegetation Science*, 10, 469–482.
- Zupanc, A. (2017) Improving cloud detection with machine learning [Online]. In: *Sentinel Hub Blog*. Available at: <https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learningc09dc5d7cf13> [Accessed 9th November 2021].

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Note S1: Details on data preparation.

Note S2: Details on the regional adjustments of the countrywide classification map.

Table S1: Categories considered as non-grassland habitat types.

Table S2: Description of the predictors used by at least one model.

Table S3: Samples used for the habitat distribution modelling of the individual habitat types.

Table S4: Pre-selected predictors for the distribution modelling of the individual habitat types.

Table S5: Weights applied to the medians of the predicted probability of occurrence.

Table S6: Masks for non-grassland habitat types.

Table S7: Percent distribution of the second to the most likely habitat types.

Table S8: Confusion matrix and accuracy measures for dataset 2.

Figure S1: Arrangement of the non-grassland habitat-type samples derived from the land-use/land cover statistics.

Figure S2: Compiled dataset.

Figure S3: Workflow of the data preparation.

Figure S4: Year of observation.

Figure S5: Example of the noise in some predictors.

Figure S6: Weighted maximum probability approach.

Figure S7: Bioregions with regional adjustments.

Figure S8: Visualization of the two combined grassland mapping products missing in Figure 3.

Figure S9: Pre-selected predictors.

Figure S10: Variable importance.

Figure S11: Model performances for the individual grassland habitat types based on dataset 2.

Figure S12: Countrywide ensemble maps of the individual habitat types.

Figure S13: Ratio of the predicted probabilities of occurrence of the most to the second most likely habitat type.

Figure S14: Omission errors for dataset 2.