Mapping habitats sensitive to overgrazing in the Swiss Northern Alps using habitat suitability modeling

Yann Fragnière a,b,*, Jérôme Gremaud b, Elias Pesenti c, Sébastien Bétrisey a, Blaise Petitpierre d, Antoine Guisan e,f, Gregor Kozlowski a,g,h

a Department of Biology and Botanic Garden, University of Fribourg, Chemin du Musée 10, CH-1700 Fribourg, Switzerland.
b Atelier 11a, Independent Ecological Consultancy, Route du Vieux-Pont 57, CH-1630 Bulle, Switzerland
c Forest and Nature Service, State of Fribourg, Route du Mont Carmel 5, CH-1762 Givisiez, Switzerland
d Info Flora, c/o Conservatoire et Jardin botaniques de la ville de Genève, Chemin de l’Impératrice 1, CH-1292, Chambéry-Geneve, Switzerland
e Department of Ecology and Evolution, University of Lausanne, Biophore, CH-1015 Lausanne, Switzerland
f Institute of Earth Surface Dynamics, University of Lausanne, Geopolis, CH-1015 Lausanne, Switzerland
g Eastern China Conservation Centre for Wild Endangered Plant Resources, Shanghai Chenshan Botanical Garden, Shanghai 201602, China
h Natural History Museum Fribourg, Chemin du Musée 6, CH-1700 Fribourg, Switzerland

ARTICLE INFO

Keywords:
Alpine vegetation
Grazing pressure
Ecological niche modeling – ENM
Species distribution model – SDM
Sheep summering
Switzerland

ABSTRACT

Unsuitable livestock farming is considered as a main driver of biodiversity loss. In the high elevation areas across the world, many subalpine and alpine natural herbaceous communities are highly sensitive to sheep overgrazing. Such habitats of high biogeographic and conservation value are refugia for slow-growing, locally rare, and cold-adapted species. However, at a regional scale, no maps of these sensitive habitats exist that the authorities could refer to for regulating grazing. We therefore studied the possibility of using habitat suitability modeling (HSM) to map five selected habitats that are potentially threatened by overgrazing in the Northern Alps of the Swiss canton of Fribourg. For each habitat, we defined a group of characteristic species and gathered their occurrences from national data centers. These occurrences were pooled within each group and used as presences to calibrate the HSM. High-resolution predictors (10 m) were used to achieve fine-scale modeling. The models predicted accurate and ecologically relevant distributions for three of the habitats: alpine swards and heaths on exposed ridges and peaks, unstable calcareous screes and vegetated snow patches. These results show that habitats that are topographically and environmentally well-defined can be finely predicted by HSM using distributions of characteristic species for use in spatial conservation planning. In the context of summer pasture management, this helped us to translate the Swiss legal basis onto maps of authorized grazing pressure.

1. Introduction

Mountainous regions are hotspots of biodiversity: the topography, geochemistry, and climate form a great diversity of habitats with specialized flora (Körner, 1999). Land use, in particular logging and grazing, also has a major influence on plant distribution and richness in high elevation areas (Dirnböck et al., 2003). Agricultural activities have shaped alpine landscapes over centuries in many European countries, mainly through alpine transhumance. Seasonal migration to summer high pastures is still practiced today and remains culturally and economically very important (Bätzing, 2015). In general, the numbers of livestock in European alpine summer pastures have declined, with some exceptions, such as Switzerland (García-Gonzalez, 2008), where approximately 400,000 cattle (260,000 livestock units) and 200,000 sheep (23,000 livestock units) are brought each year to the summer pastures (Federal Office for Agriculture FOAG, 2020).

Large herbivores affect vegetation community patterns and ecosystem functioning through processes such as grazing, browsing, trampling, defecation and urination (Mysterud, 2006). More or less all landscapes from the subalpine to the subnival zone are grazed (Kaufmann et al., 2021). Sheep graze freely on most summer pastures. In free-grazing, sheep often explore large areas but prefer high elevations and reach steep slopes and crests. Cattle, on the other hand, tend not to move too far from shelters (Zabel, 2019; Kaufmann et al., 2021). This free-grazing management is often accompanied by a very unequal use of pastoral resources. Some lower areas are undergrazed, accompanied by...
a recolonization of the forest (Pellissier et al., 2013), while the sensitive areas in the alpine zone are frequently overgrazed (Meisser and Chate- lain, 2010).

At lower elevations and up to the subalpine level, low-intensity grazing on extensive pastures is often associated with a higher species richness compared to abandoned areas (Fischer et al., 2008). On the other hand, the literature shows rather inconsistent results on the effects of grazing at high elevations (Keller and Vittoz, 2013; Pardo et al., 2015; Kaufmann et al., 2021). Different experiments of grazing exclusion have been performed thus far, suggesting that the relationship between grazing intensity and plant diversity is not linear (Ning et al., 2004). The slow response of alpine species to grazing exclusion due to slow growth is difficult to study (Pardo et al., 2015). Abandonment of livestock grazing might decrease short-statured plants, stress-tolerant species of low or no nutritive value and mosses while promoting competitive species with high or medium nutritive value (Mayer et al., 2009; Vittoz et al., 2009). Some species of the highest alpine zone are adapted to low-intensive grazing, as typically ensured by ibex and chamois (Körner, 1999). However, an early ascent by sheep to the highest portions of the summer pasturelands can interrupt the growth and reproduction of many alpine plants of special interest through defoliation and trampling (García-González, 2008). Communities react individually depending on elevation and grazing animals, yet long-term grazing exclusion studies are insufficient (Mayer and Erschbamer, 2017).

Overgrazing is the use of pasture lands beyond the limit of their production capacity or an improper use in terms of grazing period and/or duration (Enne, 2004). Overgrazing is identified as one of the most important disturbances resulting in grassland degradation (Wang et al., 2019). It influences grassland conditions, which results in reduced vegetation cover, an increased abundance of unpalatable plant species and soil erosion (Wiesmair et al., 2017). It can lead to significant changes in biomass, floristic composition and diversity in an alpine context (Erschbamer et al., 2003). At the highest elevations, the communities, generally late-successional, are very stable and therefore particularly sensitive to overgrazing (Leuschner and Ellenberg, 2017). A literature review of the impacts of overgrazing is available in Appendix A (online supplementary material).

The mapping of habitats sensitive to overgrazing is therefore an important element for conservation and land management in alpine landscapes. The main issue is that such maps generally do not exist at a regional scale. Habitat suitability modeling (HSM), also called species distribution modeling or ecological niche modeling (Guisan et al., 2017), is increasingly used in conservation and land-use planning (Guisan et al., 2013; Muscatello et al., 2021). HSM is a method for predicting the suitability of a location for a biological entity (e.g., species, clade) based on its observed relationship with environmental conditions. Many modeling techniques and algorithms have been developed in recent decades (Guisan et al., 2017). HSM allows for the extrapolation of relatively few field samples to the entire potential range, which makes it a very useful tool for conservation, as maps over large regions are often necessary (Rodriguez et al., 2007). It is, for example, increasingly used to identify areas in need of restoration or preservation at the species level (Wilson et al., 2011; Guisan et al., 2013) and could also be used to identify alpine habitats sensitive to overgrazing, for instance through pooling or stacking multi-species data (Perrier and Guisan, 2006). However, to our knowledge, this has never been done.

In Switzerland, federal legal instruments that define which habitats are subject to grazing restrictions exist (Conseil fédéral, 2013), but the cantons responsible for implementing this legal framework do not have maps showing which areas should be grazed with reduced pressure or should be forbidden. Our aim was to translate the legal basis onto maps that could be used by public authorities for grazing management.

Accordingly, this study attempts to answer the following questions: (1) Can HSM be used for mapping habitats of high conservation value that can be locally threatened by overgrazing? (2) For which of the selected habitats are HSM predictions ecologically relevant? (3) Are the predicted maps usable for conservation purposes and land management, for example, for local authorities?

2. Materials and methods

2.1. Study area, habitats and species selection

The study area (Fig. 1) comprises the southern part of the canton of Fribourg, Switzerland (from 46.438° to 46.773° N and from 6.794° to 7.381° E), with an altitudinal gradient ranging from 530 to 2389 m a.s.l. (Vanil Noir). Approximately half of the area is covered by the Prealps.

Five habitats of high conservation value that could be locally threatened by overgrazing were identified as candidate units for modeling (Fig. 2, Table B1 in Appendix B): (1) Alpine alvars and heaths on exposed ridges and peaks. These areas face extremely harsh climate conditions, comparable to Arctic environments (García-González, 2008). They are relics of the last ice ages and serve as microrefugia for cold-adapted species (Delarze, 2015). (2) Unstable calcareous screes. A sparsely vegetated habitat, including highly specialized species adapted to moving screes and high limestone content (Delarze, 2015) (3) Dry alpine and subalpine calciphilous stepped grassland. A low productive and species-rich alpine habitats of high conservation value, (4) Subalpine mesophile calciphilous closed grasslands. These grasslands can maintain themselves naturally at the subalpine level in avalanche paths and are well known in the regions for their rich biodiversity (Gerber et al., 2010). However, these grasslands are more productive and probably less sensitive to overgrazing. (5) Vegetated snow patches, with a very short vegetation period (maximum 3 to 4 months), are unproductive and often moist (Delarze, 2015). The relationship with the habitat classification of the European Nature Information System, EUNIS (Davies et al., 2004), and of Switzerland, TYPO-CH (Delarze, 2015), is given in Table B1.

Each habitat was characterized by a list of species. We included mainly species of vascular flora but also some bryophytes and lichens. A species was selected in a given habitat if (1) it occurred in the study area and (2) it was considered a characteristic (diagnostic) or constant species in that habitat (Chytrý and Tichý, 2003). Characteristic species are concentrated in a particular habitat and are absent or rare in other habitats. As such, they are useful as positive indicators of the habitat. Constant species are those that occur frequently but not necessarily exclusively in a particular habitat (Chytrý et al., 2020). The habitat classification of Switzerland (Delarze, 2015), along with our field knowledge of the region, was used to select the species. For each habitat, all characteristic species were selected (if present in the study area if data were available), and constant species were selected when they were judged as informative (i.e., generally rare in other habitats in the region, according to local experts). All species occurrences were then pooled and modeled as one single entity for each habitat.

2.2. Species distribution datasets

Data of occurrences for all selected characteristic and constant species in the study region were gathered from the different Swiss national databases (Appendix B). Only data with a location accuracy lower than 20 m were retained. For each of the habitats, all occurrences of characteristic and constant species were merged. These pooled datasets were used as presence points for modeling.

Most presence-only habitat suitability models need some kind of absence or background data (also called pseudoabsences) to contrast the presence data. The selection of these background points (number, location) requires particular attention, as it significantly influences the model output (Cherfaoui and Lobo, 2008; Phillips et al., 2009; Barve et al., 2011; Barbet-Massin et al., 2012; Fourcade et al., 2014). The presence data, collected opportunistically by different people at different times, are spatially biased. Spatial sampling bias often occurs in the absence of a predefined sampling scheme because data are collected
in areas that are, for example, easily accessible or that have a higher potential of observations (Isaac and Pocock, 2015). To reduce this problem, we first eliminated redundant presence points (data thinning, Steen et al., 2021). Then, we sampled the background using an equal-stratified design to ensure their even distribution along the main environmental gradients (Hirzel and Guisan, 2002, Appendix B). We divided the study into 5 strata for the predictor elevation and 5 for the predictor topographic position index (in total: 25 strata). Since a large number of background points is generally advised (Barbet-Massin et al., 2012), we generated a total of 10,000 points with this strategy (400 per strata). In the absence of evidence of historical constraints to dispersal and accessibility, we sampled them across the whole study area.

2.3. Environmental predictors

We used 13 environmental predictors gathered from different sources (Table B2). Most predictors are based on the digital elevation model (DEM), calculated with different algorithms. They can be considered proxies at the local scale for temperature, insolation, runoff rate, soil water content, erosion potential, terrain morphometry, exposure to wind, soil thickness, etc. (Pouteau et al., 2012; Lannuzel et al., 2021), which all influence the vegetation (Wilson and Gallant, 2000). Other predictors were also used: an index indicating the abundance of visible stones and rock at the ground surface, the normalized difference vegetation index (NDVI) and land cover. More information about predictors can be found in Appendix B. All predictors were upsampled to 10 m (native resolution was between 0.25 and 10 m). Forested and constructed areas were excluded from the study area for modeling (Fig. 1).

Fig. 1. Map of the study region: southern part of the canton of Fribourg, Switzerland. a) Overall situation, b) topography, and c) land cover. Only white areas (open lands) were considered for modeling.

Multicollinearity is assumed to be an important source of model uncertainty, and it is safer to avoid selecting highly correlated predictors (Braunisch et al., 2013; Guisan et al., 2017). We looked at correlations between predictors (r-squared values). The variance inflation factor (VIF) was also computed using the USDM R package (Naimi et al., 2014). Some predictors were eliminated to obtain VIF scores below 10 and Pearson r values below 0.7, which are generally considered acceptable thresholds (Dormann et al., 2013; Guisan et al., 2017).
2.4. Models

Three different algorithms were used for habitat suitability modeling. The computation was performed in R (R Core Team, 2018): (1) general linear models (GLMs), (2) general additive models (GAMs), and (3) Maximum entropy (MaxEnt). The two first are standard popular regression techniques used since long for modeling species distributions (Guisan and Zimmermann, 2000). They are generalizations of ordinary linear regression that allow any probability function to be considered for the response and the linear predictor to be related to the response variable via a link function (Guisan et al., 2002; McCullagh and Nelder, 1999). To compute our GLM, we used linear and quadratic terms. GAMs are GLMs in which the polynomial functions between the response and covariates are replaced by several nonlinear smooth functions to model and capture the nonlinearities in the data (Guisan et al., 2002; Wood, 2017). GAMs were computed with the ‘gam’ R package (Hastie, 2019) using a smoothing spline with 4 degrees of freedom. For both regression techniques, the logistic function (logit) was used as the link function. Equal weighting for presence and background was used (Barbet-Massin et al., 2012). For each model, the relevant predictors were selected by a stepwise procedure based on the Akaike information criterion (Sakamoto et al., 1986). The third algorithm, MaxEnt, is often described as a machine-learning technique (Phillips et al., 2006; Elith et al., 2011), yet was shown to be a well-tuned and optimized GLM (and thus also a regression) with hinge functions and lasso penalization (Renner and Warton, 2013). It is available through the ‘dismo’ R package. MaxEnt has become a popular technique due to its excellent performance and accessibility (Elith et al., 2006). We used default settings (MaxEnt V.3.4.4), except that background points were sampled as described before instead of randomly.
2.5. Evaluation

In the absence of a truly independent presence-absence dataset, two different cross-validation methods were used to evaluate the models. First, a standard k-fold cross-validation was performed, with 5 folds (Guisan et al., 2017). The data were randomly divided into 5 partitions, with 20% of the data in each group (presence and background). Four partitions were used for model training, the left-out partition was used to test the model, and this procedure was repeated 5 times until each partition served for testing. This method is, however, known to generally lead to overoptimistic evaluations. This method is, however, known to generally lead to overoptimistic evaluations when the training and testing datasets are spatially correlated (Roberts et al., 2017). We therefore also used a spatially blocked strategy to create spatial partitions, thanks to the 'BlockCV' R package (Valavi, 2019). The study area and data (presence and background) were split into square blocks and distributed to create 5 partitions. Then, the procedure was the same as for the standard 5-fold cross-validation. The spatialAutoRange function of the 'BlockCV' R package was used to look at the existing autocorrelation in the predictors. As some important predictors showed a high spatial autocorrelation in the studied area (e.g., elevation), the blocks were dimensioned to be the largest as possible. We therefore used 10 km × 10 km blocks (study area of about 30 km × 40 km).

Different metrics were used to evaluate model discriminating capacity with cross-validation testing datasets (the reported value in the results is always the mean of the 5 folds): area under the receiving operating characteristic curve (AUC), maximum of the true skill statistic (Max-TSS), point biserial correlation (COR), sensitivity (Sens.), pseudospecificity (Spec.), symmetric extremal dependence index (SEDI), and Boyce index. The proportion of the area was computed for the Bioclim model, a simple envelope model that uses the spatial autocorrelation in the studied area (e.g., elevation), the blocks were dimensioned to be the largest as possible. We therefore used 10 km × 10 km blocks (study area of about 30 km × 40 km).

The performances of the different models were measured using a 5-fold block cross-validation (Table 1) or a standard 5-fold cross-validation (Table C1). Hereafter, we refer to the metrics obtained with the 5-fold block cross-validation strategy, which we judged as more relevant to assess true model performance, with less spatial sorting bias. Other information on the models can be found in Appendix C, in particular the R outputs of the different models, including for example the significance of the different predictors in GLM and GAM, or their contribution in MaxEnt (Fig. C1).

Overall, the evaluation metrics were close for GLM, GAM and MaxEnt and clearly better than geographic null models or the simple Bioclim models (Table 1). However, MaxEnt performed slightly better for all habitats, according to most metrics (AUC, Max-TSS, COR, and SEDI). Only rarely did the GAM outperform MaxEnt for some metrics, but the difference was minimal. The sensitivity of MaxEnt was generally good, similar to other models, but with a smaller predicted “suitable area”, which indicates that MaxEnt performed generally better in terms of accuracy and specificity. Only the Boyce index sometimes classified MaxEnt significantly worse than the GLM and GAM. Therefore, as the metrics for MaxEnt were overall better, hereafter we mainly report results with this technique.

### Table 1

<table>
<thead>
<tr>
<th>Habitat</th>
<th>Model</th>
<th>AUC</th>
<th>Max-TSS</th>
<th>COR</th>
<th>Sens.</th>
<th>Spec.</th>
<th>SEDI</th>
<th>Boyce index</th>
<th>Suitable area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpine swards and heaths on exposed ridges and peaks</td>
<td>NULL</td>
<td>0.49</td>
<td>0.01</td>
<td>NA</td>
<td>0.62</td>
<td>0.39</td>
<td>0.01</td>
<td>−0.49</td>
<td>0.69 %</td>
</tr>
<tr>
<td>(627 presence data).</td>
<td>Bioclim</td>
<td>0.77</td>
<td>0.47</td>
<td>0.15</td>
<td>0.77</td>
<td>0.70</td>
<td>0.62</td>
<td>−0.23</td>
<td>8.76 %</td>
</tr>
<tr>
<td>(193 presence data).</td>
<td>GAM</td>
<td>0.90</td>
<td>0.72</td>
<td>0.25</td>
<td>0.91</td>
<td>0.81</td>
<td>0.86</td>
<td>0.78</td>
<td>4.36 %</td>
</tr>
<tr>
<td>(627 presence data).</td>
<td>MaxEnt</td>
<td>0.90</td>
<td>0.71</td>
<td>0.30</td>
<td>0.89</td>
<td>0.81</td>
<td>0.85</td>
<td>0.81</td>
<td>3.38 %</td>
</tr>
<tr>
<td>Block 5-fold CV</td>
<td>NULL</td>
<td>0.59</td>
<td>0.19</td>
<td>NA</td>
<td>0.84</td>
<td>0.35</td>
<td>0.30</td>
<td>−0.46</td>
<td>0.73 %</td>
</tr>
<tr>
<td>Unstable calcareous screes</td>
<td>Bioclim</td>
<td>0.73</td>
<td>0.41</td>
<td>0.12</td>
<td>0.76</td>
<td>0.65</td>
<td>0.56</td>
<td>−0.37</td>
<td>15.42 %</td>
</tr>
<tr>
<td>(380 presence data).</td>
<td>GAM</td>
<td>0.86</td>
<td>0.63</td>
<td>0.20</td>
<td>0.88</td>
<td>0.75</td>
<td>0.78</td>
<td>0.62</td>
<td>7.93 %</td>
</tr>
<tr>
<td>(380 presence data).</td>
<td>MaxEnt</td>
<td>0.87</td>
<td>0.64</td>
<td>0.22</td>
<td>0.88</td>
<td>0.76</td>
<td>0.80</td>
<td>0.17</td>
<td>6.66 %</td>
</tr>
<tr>
<td>Dry alpine and subalpine calciphilous stepped grassland</td>
<td>NULL</td>
<td>0.65</td>
<td>0.30</td>
<td>0.13</td>
<td>0.46</td>
<td>0.83</td>
<td>0.44</td>
<td>−0.73</td>
<td>3.65 %</td>
</tr>
<tr>
<td>(472 presence data).</td>
<td>Bioclim</td>
<td>0.64</td>
<td>0.26</td>
<td>0.11</td>
<td>0.65</td>
<td>0.61</td>
<td>0.36</td>
<td>0.41</td>
<td>22.69 %</td>
</tr>
<tr>
<td>(470 presence data).</td>
<td>GAM</td>
<td>0.70</td>
<td>0.34</td>
<td>0.16</td>
<td>0.77</td>
<td>0.57</td>
<td>0.48</td>
<td>0.80</td>
<td>18.38 %</td>
</tr>
<tr>
<td>(470 presence data).</td>
<td>MaxEnt</td>
<td>0.79</td>
<td>0.48</td>
<td>0.24</td>
<td>0.83</td>
<td>0.64</td>
<td>0.64</td>
<td>0.91</td>
<td>19.85 %</td>
</tr>
<tr>
<td>Subalpine mesophile calciphilous closed grasslands</td>
<td>NULL</td>
<td>0.59</td>
<td>0.24</td>
<td>0.07</td>
<td>0.60</td>
<td>0.65</td>
<td>0.34</td>
<td>−0.61</td>
<td>2.69 %</td>
</tr>
<tr>
<td>(195 presence data).</td>
<td>Bioclim</td>
<td>0.73</td>
<td>0.38</td>
<td>0.18</td>
<td>0.76</td>
<td>0.61</td>
<td>0.52</td>
<td>−0.10</td>
<td>20.40 %</td>
</tr>
<tr>
<td>(195 presence data).</td>
<td>GAM</td>
<td>0.82</td>
<td>0.53</td>
<td>0.25</td>
<td>0.86</td>
<td>0.67</td>
<td>0.70</td>
<td>0.87</td>
<td>21.65 %</td>
</tr>
<tr>
<td>(195 presence data).</td>
<td>MaxEnt</td>
<td>0.84</td>
<td>0.55</td>
<td>0.30</td>
<td>0.89</td>
<td>0.65</td>
<td>0.72</td>
<td>0.80</td>
<td>21.98 %</td>
</tr>
<tr>
<td>Vegetated snow-patches</td>
<td>NULL</td>
<td>0.50</td>
<td>0.02</td>
<td>NA</td>
<td>0.64</td>
<td>0.38</td>
<td>0.03</td>
<td>−0.29</td>
<td>0.88 %</td>
</tr>
<tr>
<td>(477 presence data).</td>
<td>Bioclim</td>
<td>0.77</td>
<td>0.51</td>
<td>0.10</td>
<td>0.84</td>
<td>0.67</td>
<td>0.67</td>
<td>−0.62</td>
<td>8.18 %</td>
</tr>
<tr>
<td>(457 presence data).</td>
<td>GAM</td>
<td>0.88</td>
<td>0.68</td>
<td>0.18</td>
<td>0.90</td>
<td>0.77</td>
<td>0.83</td>
<td>0.78</td>
<td>6.12 %</td>
</tr>
<tr>
<td>(457 presence data).</td>
<td>MaxEnt</td>
<td>0.88</td>
<td>0.69</td>
<td>0.22</td>
<td>0.86</td>
<td>0.84</td>
<td>0.84</td>
<td>0.37</td>
<td>3.32 %</td>
</tr>
</tbody>
</table>

Characteristic and constant species were selected for each habitat and their occurrences pooled as presence points for modeling (Table B1). For example, 21 species were selected for vegetated snow-patches (193 presence data) and up to 57 species for dry alpine and subalpine calciphilous stepped grasslands (627 presence data).

The performances of the different models were measured using a 5-fold block cross-validation (Table 1) or a standard 5-fold cross-validation (Table C1). Hereafter, we refer to the metrics obtained with the 5-fold block cross-validation strategy, which we judged as more relevant to assess true model performance, with less spatial sorting bias. Other information on the models can be found in Appendix C, in particular the R outputs of the different models, including for example the significance of the different predictors in GLM and GAM, or their contribution in MaxEnt (Fig. C1).

Overall, the evaluation metrics were close for GLM, GAM and MaxEnt and clearly better than geographic null models or the simple Bioclim models (Table 1). However, MaxEnt performed slightly better for all habitats, according to most metrics (AUC, Max-TSS, COR, and SEDI).

For assessment, the metrics were calculated for geographic null models (Hijmans, 2012), a simple model based on the assumption that the closer to a known presence point, the more likely it is to find another presence point (inverse distance weighting). The Schoener’s D metrics (Warren et al., 2010) was calculated to quantify the degree of overlap of the different habitats. The value ranges from 0 (no overlap) to 1 (identical model projections).
The different metrics indicated that the habitat Alpine swards and heaths on exposed ridges and peaks can be modeled with a good degree of reliability and the higher reliability among the other habitats considered in this study. In the block cross-validation and using a threshold that maximizes TSS, the sensitivity reached 0.89, while the (pseudo) specificity was 0.81. The models of the two further habitats, unstable calcareous screes and subalpine mesophile calciphilous closed grasslands, which were determined as overall geographically and ecologically relevant. On the other hand, the models for the two habitats dry alpine and subalpine calciphilous stepped grassland and subalpine mesophile calciphilous closed grasslands seemed visually to lack accuracy, with approximately 20% of the total area predicted as suitable. Moreover, there was a large overlap between the predictions of these two habitats, with a Schoener’s D of 0.75 (Table C2).

4. Discussion

The present paper demonstrate the ability of HSM to be used for habitat mapping at high resolution for conservation purposes, and corroborates with recent studies carried out both in alpine and lowland areas (Vincent et al., 2019; Ramel et al., 2020; Lannuzel et al., 2021; Oliveira et al., 2021). The main novelty of our approach lies in the fact that we use modeling not with a single species, but with a group of characteristic species in order to map habitats of high conservation value. Among the five selected habitats, the rarest and probably the most sensitive to overgrazing were also the most reliably modeled in our study: alpine swards and heaths on exposed ridges and peaks, unstable calcareous screes and vegetated snow patches. For those, prediction maps can be used as a planning tool, and we hereafter give some recommendations on how to transmit these kinds of outputs to regional or political authorities, and the implications for practitioners to improve grazing management. It seems useful to discuss first, some aspects of modeling.

4.1. Pooled presence data

For all species selected as characters or constants for a habitat, the occurrence data were pooled together to serve as a single presence dataset for modeling. To our knowledge, this is the first time that this strategy has been used and evaluated to model alpine habitats at such a fine resolution (10 m). A similar approach was used to highlight areas with potential ecological quality according to different ecological groups at the national scale in Switzerland at a coarser resolution of 100 m (Petitpierre et al., 2021). HSM is generally used for single species, but the predictive approaches can be used to model any biological entity (Guisan et al., 2017), for example, in our case, a part of a community, including species that share part of their realized niches. It has been suggested that for functionally close species, expanding the niche concept to the whole group may be practically relevant (Hirzel and Lay, 2009), although formally an ensemble of species does not constitute a population, which is a key concept behind the niche theory (Pulliam, 2000). However, this approach is not without limitations. It reflects the current state of these plant communities and their species interactions and thus could be problematic in using for projections in the future where species interactions will likely change (Wisz et al., 2013; D’Amen et al., 2017; Trisos et al., 2020). Moreover, the species selection for a given habitat, based on literature and expert knowledge, cannot be fully objective. Finally, the ecological amplitude slightly varies among selected species and accordingly the realized niches does not overlap exactly. If some species with broader ecological niches are selected, the suitable areas for the entire habitat type could be overestimated, and the choice of species should therefore be done carefully. Nevertheless, an important advantage of this strategy is the greater sample size enabled by grouping several species, which is a very important condition for good model performance (Wisz et al., 2008).

4.2. Model performance

MaxEnt was the most high-performing algorithm, but the metrics and predictions of the GLM and GAM were overall close and could be used as an alternative, or an ensemble modeling approach could also be used (Meller et al., 2014). The default parametrization of MaxEnt was appropriate in our study, although it is worth to check if better model parametrization are possible by using automated protocols, for example the ENMeval or kuenm R package (Muscarella et al., 2014; Cobos et al., 2019).

We argue the importance of making cross-validations using spatial blocks, as standard cross-validation usually leads to overoptimistic results (Table C1). With data collected without a predefined sampling scheme, spatial autocorrelation between training and testing datasets generally occurs when a standard k-fold cross-validation is used, and therefore indicates high model performances, even for geographically null models. Cross validation using spatial blocks gives a better view of the true model performance (Roberts et al., 2017), yet it can sometimes also be problematic if occurrences are for example geographically restricted to a single area covering only a very few partitions, which was however not the case here.

Objectively assessing the performance of a presence-only model is difficult (Hirzel et al., 2006; Leroy et al., 2018), but the calculation of different metrics gave valuable insight in our case. For habitats that only occur over small areas, we found it particularly useful to compare the percentage of predicted “suitable area” and the sensitivity (true positive rate) with the block cross-validation. For example, considering the habitat Alpine swards and heaths on exposed ridges and peaks and the MaxEnt model, only 3.4% of the total area was projected as suitable, and on average, 89% of the presence points were correctly predicted in the testing block, which indicated good accuracy. On the other hand, the model for the habitat subalpine mesophile calciphilous closed grasslands predicted correctly 80% of the presence points in the testing blocks, which is correct, but lacks specificity with more than 20% of the total area predicted as suitable.

Most metrics (AUC, Max-TSS, COR, and SEDI) were largely correlated and consistent between them. The Boyce index was surprisingly worse for MaxEnt models in some cases. It is an interesting metric for presence-only models, but it must be interpreted with caution, as it seems very sensitive when testing presence is missing toward the higher prediction values. It is important to compare different metrics and, above all, to determine the ecological and geographic relevance of the model predictions.

4.3. Model predictions

The predictions of the different modeled habitats tended to overlap in some areas (Fig. 9). This could at first sight be seen as incoherent but in fact largely corresponds to the field reality. The different habitats can often change gradually or form mosaics. For example, in the study region, crests are home for species of alpine swards but generally together with species of alpine calciphilous stepped grassland or sometimes other habitats such as snow patches. From a conservation perspective, this overlap is not a problem because we do not want to map only the typical...
Fig. 3. MaxEnt model predictions for the five selected habitats: a) alpine swards and heaths on exposed ridges and peaks, b) unstable calcareous screes, c) dry alpine and subalpine calciphilous stepped grassland, d) subalpine mesophile calciphilous closed grasslands, and e) vegetated snow patches. Example in the mountain chain between Brenleire (Br.), Vanil Noir (V.N.), Vanil de l’Ecri (V.E.) and Pointe de Paray (P.P.), canton of Fribourg, Switzerland. The left panel shows 2D maps, and the right panel shows 3D views. The symbols in the maps show the position and angle of view for 3D views.
habitats (e.g., as defined in Delarze, 2015), and those forming transitional stages (i.e., in-between the typical ones) are also worth conservation.

The model for the habitat alpine swards and heaths on exposed ridges and peaks obtained the best evaluation metrics. Topographic position and elevation play a crucial role for that habitat (Delarze, 2015), it could therefore be well modeled mainly thanks to predictors like elevation and wind index. The projection seemed relevant to what can be expected in the field, pointing precisely exposed crests and summits at high elevation. Metrics and prediction for unstable calcareous screes and vegetated snow patches were also good. We noticed however a few problems with the first, as north exposed screes were projected as more suitable than south exposed screes, likely because our dataset was dominated by occurrences of cold adapted species, in particular Papaver occidentale (Fragnière et al., 2020). This highlights possible bias of such “pooling” method when a species dominates in the dataset. It is very conclusive to compare visually the projection of vegetated snow patches with late spring remaining snow patches (Fig. 4). We were able to find a recent orthophoto taken at the “right time”, which provided a further ground validation of our predictions. The projections of the two last habitats, dry alpine and subalpine calciphilous stepped grassland and subalpine mesophile calciphilous closed grasslands, were less accurate, with more than 20 % of the total study area predicted as suitable. Although these two habitats are more widespread, this nevertheless seems overoptimistic. Additional information is given in Appendix C.

Overall, habitats that were clearly linked with one or a few environmental predictors (e.g., elevation, topographic position, and stones on the soil surface) obtained both the best metrics and ecologically and geographically relevant predictions. Finding or developing appropriate ecologically-meaningful predictors (Mod et al., 2016) thus plays a crucial role for good model performance and transferability (Regos et al., 2019).

4.4. Perspectives and recommendations

In order to avoid the possible consequences of overgrazing, it is helpful to quantify the livestock carrying capacity of a specific area (García-González, 2008). The Swiss federal government has put in place several instruments to regulate land use, including the Ordinance on Direct Payments in Agriculture (Conseil fédéral, 2013), which specifies some grazing restrictions (maximum livestock load) in some habitats or areas by grazing management types (free, rotating or guarded) with the aim to favor biodiversity. The cantons are responsible for implementing this legal framework but need maps showing which areas do not have to be grazed or with a reduced load. Model predictions, along with other data such as elevation, were accordingly used to translate the federal legal basis onto maps for all open surfaces above 1500 m a.s.l. of the canton of Fribourg (Appendix D). Although habitat suitability models represent potentially powerful tools to support conservation decisions (Guisan et al., 2013), and especially land planning (Tulloch et al., 2016), we do not recommend providing spatial predictions for direct use by practitioners or decision-makers, but rather to first examine them and, if necessary, make appropriate improvements whenever possible or discard them otherwise until better predictions are obtained (e.g., when new predictors become available). Some small inconsistencies in model predictions are almost inevitable and can sometimes be simply adjusted by expert knowledge. Other approaches could also be used to model and spatially predict these plant communities, such as by stacking spatial predictions of individual species models (D’Amen et al., 2017).

To evaluate the model, we used the threshold that maximizes TSS, but it is not necessarily the best threshold for conservation purposes (Huggett, 2005; Liu et al., 2013) and threshold independent evaluations of plant communities are also possible (Scherer et al., 2020). If a threshold must be used to derive simpler predictions for practitioners, the optimal one should be discussed with the concerned parties or a way should be found to help them use the continuous prediction maps (e.g. habitat suitability) directly. In all cases, it is important to provide predictions maps that are understandable and easy to read and understand by end-users, e.g. with the gradient of habitat suitability divided into a few meaningful classes for their intended use.

5. Conclusion

Our study supports the ability of HSM to predict habitats of high
conservation value at a local scale using fine-resolution predictors and presence data from existing databases. With increasing access to precise species occurrence data and high-resolution environmental predictors, HSM could undoubtedly become a powerful tool in this context.

Plant communities in mountainous ecosystems are dependent on a grazing regime, both ecologically and economically sustainable. The methodology used in our paper could be replicated in other areas confronted with overgrazing and serve as basis for development of local maps of authorized grazing pressure and to apply management recommendations, for example of Natura 2000 habitats (European Commission, 2008). Fine-scale habitat modeling has at the same time an enormous potential in terms of nature conservation and landscape management in alpine areas, for example, regarding tourism (installations for skiing, trampling by hikers, etc.), or the planning of new natural reserves.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

YF, JG and EP developed the research idea; YF realized the modeling work and analyzed the data; BP and AG advised on the modeling work; GK and SB assessed the topics of grazing and impact on the alpine vegetation and analyzed the relevance of the results; YF, JG and GK led the writing of the manuscript; and all authors revised the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The occurrence data are to be requested from the Swiss national database (www.infospecies.ch). Model outputs and maps will be made available on request.

Acknowledgement

We thank the Forest and Nature Service of the canton of Fribourg for their initiative of considering summer pastures and nature conservation in the canton of Fribourg, which was the starting point of this work. We also thank all the local botanists who collected data in our region and made this kind of study possible. AG benefitted support from the ValPar.ch project funded by the Federal Office of the Environment. Finally, we thank reviewers for their constructive comments, which helped us to improve the manuscript.

Appendices. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biocon.2022.109742.

References


Hastie, T., 2019. gam: generalized additive models. Available at: https://cran.r-project.org/package=gam.


