UNIL | Université de Lausanne

Faculté de biologie et de médecine

Ecole de biologie

PREDICTING THE SPATIAL DISTRIBUTION OF MEDICINAL AND AROMATIC PLANTS AND THEIR ASSEMBLAGES IN THE ALPINE REGION

Travail de Maîtrise universitaire ès Sciences en comportement, évolution et conservation

Master Thesis of Science in Behaviour, Evolution and Conservation

par

Lou LESCUYER--DE DECKER

Directeur : Prof. Antoine Guisan
Superviseur (s) : Dr. Catherine Pfeifer, FiBL, Frick
Expert (s) : Anonymous
Département d'Ecologie et d'Evolution (DEE)

Janvier 2022

Abstract

Context: Increasing of modern and more specialised agriculture alters the Alpine landscape and its biodiversity. The strong ecological interest of medicinal and aromatic plants (MAPs), combined with their high economic value, could make the MAP cultivation a solution to promote more sustainable agriculture in the Alpine region.

Objectives: The first objective was to predict the distribution of some MAP species and their assemblages in the Alpine region, to identify most suitable areas for MAP cultivation. We then aimed to assess the impact of climate change on predicted suitable areas.

Methods: The study was conducted across the Alpine region. Species distribution modelling was used to predict suitable sites for MAP species and their assemblages. Then, richness and composition of MAP assemblages were assessed by combining the "sum of predicted probabilities" approach with the probability ranking rule.

Results: We found that, depending on environment, MAP assemblages are composed of different species and are of varying richness. Our results also showed that climate change would reshuffles MAP species in space in species-specific ways and affect both richness and composition of MAP assemblages.

Conclusions*:* With our findings, sustainable agriculture may be promoted by helping policy makers to support the emergence of MAPs production where it is most ecologically appropriate.

Keywords: medicinal and aromatic plants, species distribution modelling, Alpine region, biodiversity promotion, agricultural land planification, climate change

Résumé

Contexte : L'augmentation de l'agriculture moderne et plus spécialisée altère le paysage alpin et sa biodiversité. Le fort intérêt écologique des plantes aromatiques et médicinales (PAMs), allié à leur haute valeur économique, pourrait faire de la culture des PAMs une solution pour promouvoir une agriculture plus durable dans la région alpine.

Objectifs : Le premier objectif était de prédire la distribution de certaines espèces de PAMs et de leurs assemblages dans la région alpine, afin d'identifier les zones les plus appropriées pour la culture de PAMs. Nous avons ensuite cherché à évaluer l'impact du changement climatique sur les zones propices prédites.

Méthodes : L'étude a été menée dans toute la région alpine. La modélisation de la distribution des espèces a été utilisée pour prédire les sites appropriés aux espèces de PAMs et à leurs assemblages. Ensuite, la richesse et la composition des assemblages de PAMs ont été évaluées en combinant l'approche de la « somme des probabilités prédites » avec la règle de classement des probabilités.

Résultats : Nous avons constaté que, selon l'environnement, les assemblages de PAMs sont composés de différentes espèces et sont de richesse variable. Nos résultats ont également montré que le changement climatique entraînerait un remaniement des espèces de PAMs dans l'espace de manière spécifique à chaque espèce et affecterait à la fois la richesse et la composition des assemblages de PAMs.

Conclusions : Avec nos résultats, l'agriculture durable peut être promue en aidant les décideurs politiques à soutenir l'émergence de la production de PAMs là où elle est la plus écologiquement appropriée.

Abbreviations and acronyms

AUC: Area Under the Curve

EM: Ensemble Model

GAM: Generalised Additive Model

GBIF: Global Biodiversity Information Facility

GBM: Gradient Boosting Machine (i.e. boosted regression trees)

GLM: Generalised Linear Model

MAP: Medicinal and Aromatic Plant

PCA: Principal Component Analysis

PRR: Probability ranking rule

RF: Random Forest

ROC: Receiver Operating Characteristics

SDM: Species Distribution Modelling

TSS: True Skill Statistic

Table of contents

1. Introduction

Over the past decades, human activities such as agriculture have profoundly changed the Alpine landscape (Gurung et al., 2012; Price, 1999). Mountainous regions are today increasingly subject to intensive use (Spehn et al., 2010). In the Alpine valleys, agriculture has intensified in the most favourable areas, while more traditional forms of farming, where different agricultural products are produced over small surfaces, are wiped off the map as older generations of farmers disappear and remote mountain pastures are abandoned (Dobremez et al., 2015; Gellrich et al., 2007). With the increase of modern and more specialised farming, source of monocultures, biodiversity is decreasing (Chappell & LaValle, 2011; Firbank, 2005; Marini et al., 2011).

In the lowlands, as the Swiss Plateau, the negative impacts associated with intensive agriculture are mainly due to high fertilisation, massive pesticides use or even alteration of the initial nature of the environments, such as drain of wetlands (FOEN, 2017; Geiger et al., 2010). Resulting from large-scale over-fertilisation of ecosystems, vegetation is becoming increasingly homogeneous (FOEN, 2017). In Switzerland, the flora associated with cultivated land is now one of the most threatened plant groups, with 42% of these species assessed as vulnerable (Bornand et al., 2016). Likewise, intensive land use leads to generalised uniform ecological conditions and thus to a decrease in the diversity of habitats on these lands. (FOEN, 2017). A very high proportion of threatened environments is observed in habitats linked to agriculture (Delarze et al., 2016). Thus, unsustainable agriculture is one of the main culprits for the loss of biodiversity (Benton et al., 2021).

However, agricultural policies in Europe are currently changing from supporting production to developing ecological connectivity and thus to strengthen, improve and restore biodiversity (https://www.alpine-region.eu). As a result, there is a window of opportunity to seek to counter biodiversity loss by promoting agro-biodiversity in agricultural policies (EU, 2020). Emergence of more agro-biodiverse landscapes can be achieved through increased use of diverse mixed farming systems, generally support by organic farms (van Mansvelt et al., 1998). The adoption of more diversified practices, nevertheless, depends on the decision making of farmers, who will only adopt practices that are beneficial to them (Pfeifer et al., 2009).

Over the last two millennia, medicinal and aromatic plants (MAPs) have been widely used in Europe for different purposes, the main one being medicine (Dal Cero et al., 2014). Due to their use for therapeutic and culinary purposes, but also as components of cosmetics and other natural health products, MAPs are particularly implicated in human well-being and provide many ecosystem services (Millennium Ecosystem Assessment, 2005). Nowadays, many industries have developed an increasing interest in MAPs in order to use them as raw materials in the production of pharmaceutical, cosmetic and food products (Lubbe & Verpoorte, 2011). Consumers have also contributed to the growing demand for herbal extracts by paying more and more attention to natural ingredients produced in a sustainable manner. In the European market, Germany plays a dominant role in the production of MAPs (27%), followed by France (22%) and Italy (11%). In Germany, however, the current cultivation area for MAPs covers only 12% of the growing area required to meet the needs of industry (Argyropoulos, 2019).

As a result, there is an opportunity to expand the production of MAPs and many farmers might consider their cultivation as more profitable than other crops. Furthermore, the promotion of MAP crops could increase the value of biodiversity and promote its connectivity in monotone agriculturally shaped landscapes (Padulosi et al., 2002). Thus, the interest of the pharmaceutical and food industries for plant material, brought together with the need to protect plant biodiversity, creates an opportunity for farmers to diversify their production and improve their income while participating in a more sustainable agriculture (EIP-AGRI, 2020). MAPs however have different environmental requirements and can not grow everywhere. Accounting for these different requirements is particularly important for the organic production of MAPs, which tries to rely on ongoing ecological processes, instead of chemical input such as pesticides or synthetic fertilisers.

There is thus an increasing need to assess the potential for growing MAP crops at specific sites. One approach to this is to assess the environmental suitability of MAP species through the use of models and predict it at sites to be evaluated. Predictive species distribution modelling (SDM; Franklin, 2010; Guisan et al., 2017; Guisan & Zimmermann, 2000; Peterson et al., 2011) is a standard tool that can be used for this purpose, as for many other ones, including to support conservation decision making (Guisan et al., 2013; Tulloch et al., 2016). In addition, SDMs can produce future predictions under changing environmental conditions (Guisan et al., 2017), such as under climate change (Engler et al., 2011). Over the past two decades, SDMs have made great progress (Araújo et al., 2019; Guisan et al., 2017; Zurell et al., 2020), particularly in the establishment of ensemble modelling techniques (Hao et al., 2020). Ensemble modelling involves combining the projections of several different statistical techniques (or other aspects of the model parameterisation or of the data used) into a single projection (Araújo & New, 2007; Thuiller, 2004). As a result, this combined projection will have a smaller mean error than any of its individual components (Thuiller et al., 2009). Ensemble modelling has become increasingly used in SDM, although sometime challenged (Hao et al., 2020), and this approach will be used in this study as well.

The aims of the present study are, first, to predict the distribution of some characteristic MAP species in the Alpine region, to highlight areas that have the best potential for their cultivation, today and in a climatically changed future. Secondly, we aimed at identifying the most productive sites for MAP cultivation, i.e. where several of these species could grow together into MAP assemblages. To do this, we used the individual MAP predictions to assess the MAP richness and composition of the evaluated sites. We were particularly interested in 5 locations in Switzerland, where we know that existing farms grow MAPs species. We also investigated the impact of climate change on the predicted MAP assemblages. Based on previous work (Losapio et al., 2021), we assumed that due to global warming, we would observe a decline in the presence of MAPs in the Alpine zone, as well as changes in the composition of MAP assemblages. We also hypothesised that some species would colonise new areas in the future. In brief, this paper intends to answer the following questions: (1) How are MAP species distributed in the Alpine area and what are the environmental factors controlling them? (2) Where and how do MAP species assemble? (3) How will the distribution, richness and composition of medicinal and aromatic plant species assemblages change under climate change?

Here, we modelled more specifically the current and future spatial distribution of 27 commonly used medicinal and aromatic plant species in the Alpine region, at a 1 km x 1 km resolution. To this end, secondary data, including citizen science data as well as recent existing geographical product, such as soil, climatic or topographic maps were used. More precisely, we combined 4 different statistical modelling techniques in an ensemble modelling approach to obtain probabilistic distribution maps and associated uncertainties for each of our studied species. By summing predicted probabilities, the projections were then used to assess the richness per site of the species assemblages across the study area. The composition of the species assemblages was then identified using the probability ranking rule, combined with the previously used "sum of predicted probabilities" approach. Additionally, we examined species associations using multivariate ordination applied to the predictions obtained from our ensemble models. We finally investigated the effects of climate change on each of the elements mentioned above.

2. Methods

2.1 Overall approach

Figure 1 represent an overview of the major steps and research questions of this study. After preparing the various species and environmental predictor data (see Sections "2.3" to "2.5"), we built species distribution models (SDMs) for each species (see Sections "2.6" to "2.7"). The resulting spatial predictions were then used to several purposes, including mapping the species distribution and assessing richness and composition of species assemblages (see Section "2.8"). The whole process was also carried out for future forecasts using a climate change scenario. Codes produced during this study are available here: https://github.com/loulldd/Master-Thesis.git.

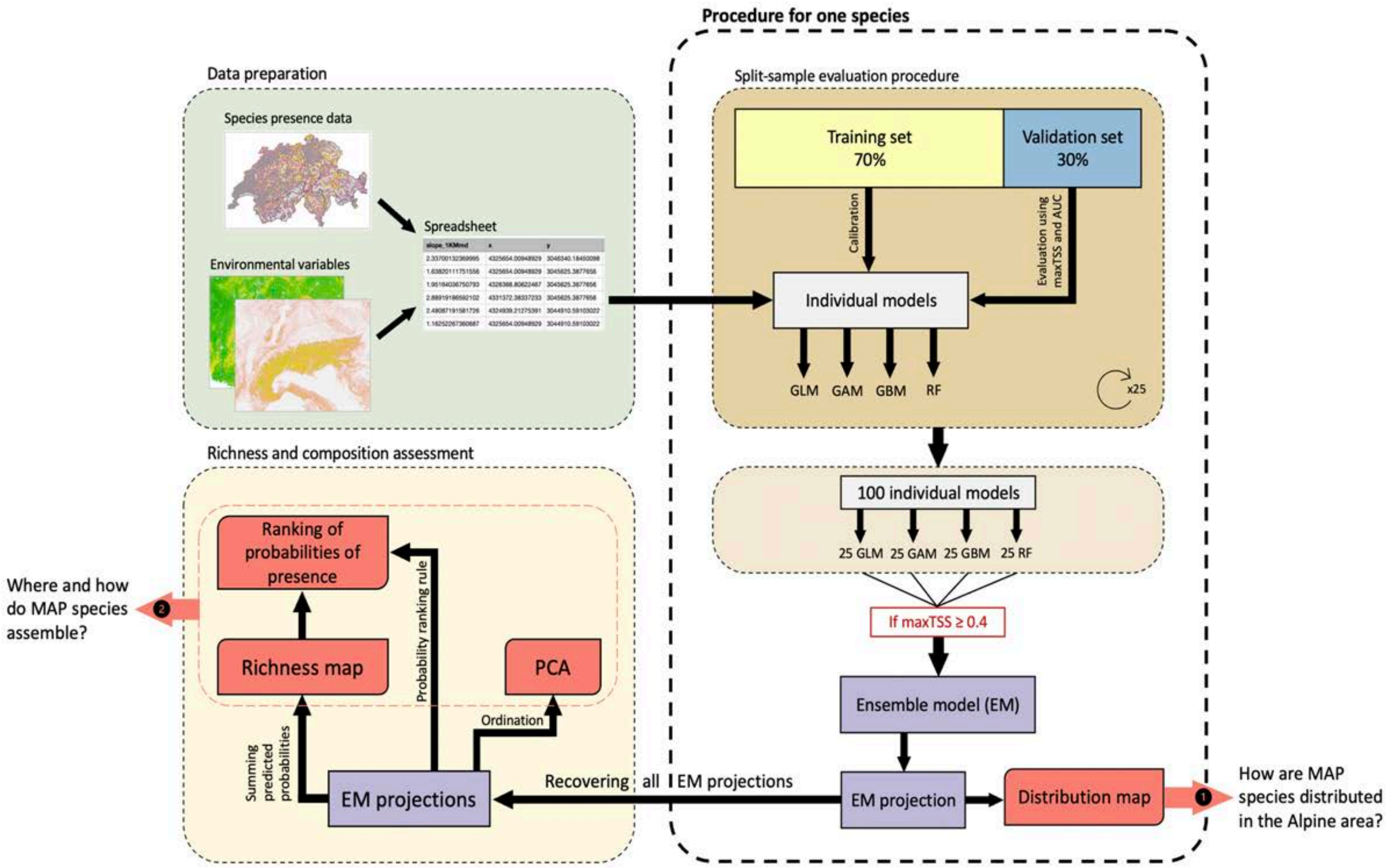


Figure 1. Overview of the major steps and research questions of this study, with a more detailed insight of the species distribution modelling (SDM) procedure used (on the right). The SDM procedure indicated here corresponds to the procedure applied for a species. It was therefore repeated 27 times, for each species studied. The whole process was also carried out for future forecasts using a climate change scenario, to answer the 3rd research question of this study not shown here: How will the distribution, richness and composition of medicinal and aromatic plant species assemblages change under climate change?

2.2. Study area

The study was conducted across the Alpine region, understood as the region extending from the Mediterranean coast to Vienna and defined by the EUSALP transboundary area (Figure 2; https://ec.europa.eu/regional_policy/index.cfm/en/policy/cooperation/macro-regional-strategies/alpine/). The study area spans seven countries: Austria, France, Germany, Italy, Liechtenstein, Slovenia, and Switzerland, with a total area of $\approx$ 441,006 km^2. Elevation ranges from -3 to 4809 m a.s.l. The area's climate is very diverse, including the Mediterranean, continental, and oceanic climates.

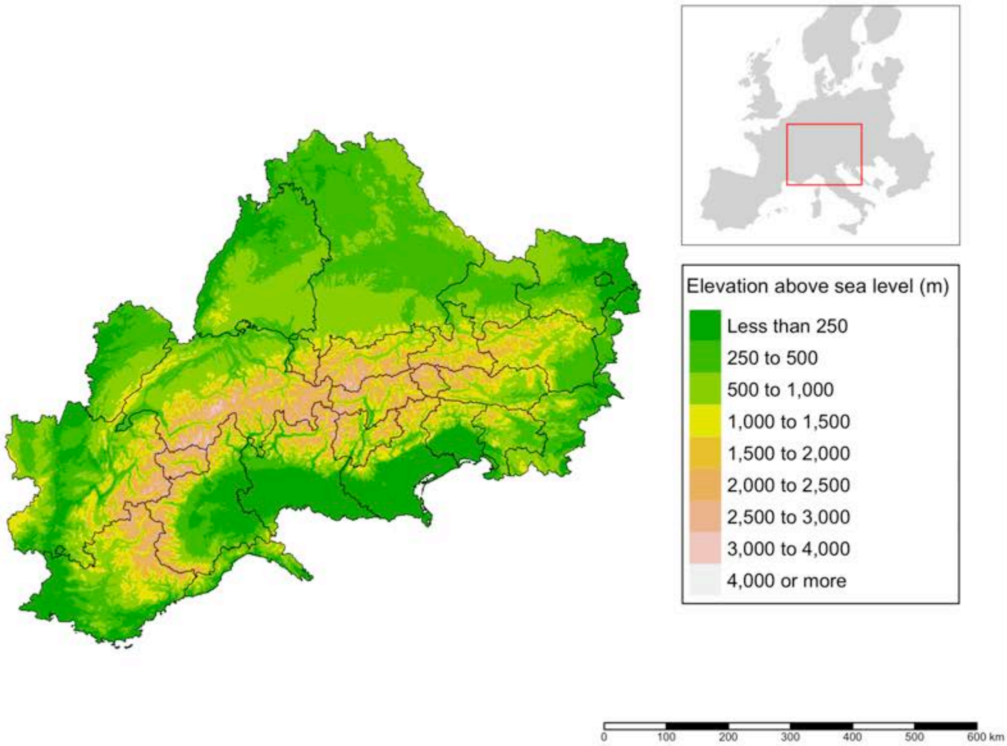


Figure 2. Elevation above sea level (in meters) of the EUSALP area.

2.3. Medicinal and aromatic plant data

We selected 27 medicinal and aromatic plant (MAP) species cultivated in the study area (Table 1). These species are used in food and pharmaceutical industries especially in Switzerland and were thus chosen because of their strong economic interest. They represent a diverse mix of plant families and growth altitude, with species growing from the colline (lowlands) to the alpine (highlands) belts. Elevation of occurrences ranges from -3 to 3745 m a.s.l. This selection also contains various plant growth forms (Raunkiaer et al., 2005) like herbaceous, woody plants, and shrubs.

Raw occurrence data for the 27 plant species were retrieved for the study area from the Global Biodiversity Information Facility (https://www.gbif.org). In addition, we received from Info Flora (https://www.infoflora.ch) the complete dataset of occurrences of the 27 MAP species at a 1 km x 1 km resolution, from 1980 to 2020 and for the whole territory of Switzerland.

For both GBIF and Info Flora datasets, data were then selected to cover a time range from 1980 to 2020 and match the temporal range of environmental variables (they correspond to averages or sums for this 40-year period). Initially, the observations coordinates were in the WGS84 coordinate system (i.e. World Geodetic System) for GBIF and CH1903+ (i.e. Swiss geodetic reference system) for Info Flora. We reprojected the coordinates of GBIF and Info Flora datasets in the EPSG:3035 projection system, which is the reference coordinate system for European Union (EU) countries and Europe in general. The two datasets were eventually bound before being aggregated so that there was only one observation per species per square kilometre. Finally, to ensure a minimum sample size sufficient for accurate model fitting (van Proosdij et al., 2016), only species with a minimum of 30 occurrences between 1980-2020 in the study area were considered.

Table 1. List of the twenty-seven medicinal and aromatic plant species studied

Scientific name	Common name	Abbreviation	Family	Growth altitude	GBIF DOI
Achillea millefolium L.	Yarrow	YARR	Asteraceae	Colline - subalpine belts	https://doi.org/10.15468/dl.ejt564
Alchemilla xanthochlora Rothm.	Lady's mantle	LAMA	Rosaceae	Meadows - montane belt	https://doi.org/10.15468/dl.q4xf6n
Althaea officinalis L.	Marsh mallow	MAMA	Malvaceae	Colline belt	https://doi.org/10.15468/dl.haaaqw
Artemisia umbelliformis Lam.	Wormwood	WORM	Asteraceae	Subalpine - alpine belts	https://doi.org/10.15468/dl.33mx5q
Artemisia vallesiaca All.	Mugwort from Valais	MUVA	Asteraceae	Colline - montane zone	https://doi.org/10.15468/dl.6a6umz
Cannabis sativa L.	Hemp	HEMP	Cannaba-ceae	Colline belt	https://doi.org/10.15468/dl.hpd7fb
Centaurea cyanus L.	Cornflower	CORN	Asteraceae	Colline - montane belts	https://doi.org/10.15468/dl.mwxrfd
Gentiana lutea L.	Yellow gentian	YEGE	Gentiana-ceae	Montane - subalpine belts	https://doi.org/10.15468/dl.42s8nt
Leontopodium alpinum Cass.	Edelweiss	EDEL	Asteraceae	Subalpine - alpine belts	https://doi.org/10.15468/dl.kanwb7
Malva sylvestris L.	Common mallow	MALL	Malvaceae	Colline - montane belts	https://doi.org/10.15468/dl.nwkdqe
Marrubium vulgare L.	White horehound	WHHO	Lamiaceae	Colline - montane belts	https://doi.org/10.15468/dl.xj8akp
Melissa officinalis L.	Lemon balm	LEBA	Lamiaceae	Colline belt	https://doi.org/10.15468/dl.5xptpz
Mentha $^{\times}$*piperita* L.	Peppermint	PEPP	Lamiaceae	×	https://doi.org/10.15468/dl.4eyhkd
Ocimum basilicum L.	Basil	BASI	Lamiaceae	Colline - montane belts	https://doi.org/10.15468/dl.658gc7
Origanum vulgare L.	Oregano	OREG	Lamiaceae	Colline - subalpine belts	https://doi.org/10.15468/dl.kncusq
Pimpinella saxifraga L.	Scarlet pimpernel	SCPI	Apiaceae	Colline - subalpine belts	https://doi.org/10.15468/dl.ymkntv
Plantago lanceolata L.	Ribwort plantain	RIPL	Plantagina-ceae	Colline - subalpine belts	https://doi.org/10.15468/dl.hdt9za
Primula veris L.	Cowslip	COWS	Primula-ceae	Meadows	https://doi.org/10.15468/dl.wudqvd
Rhodiola rosea L.	Golden root	GORO	Crassula-ceae	Subalpine - alpine belts	https://doi.org/10.15468/dl.sesqqn
Rosmarinus officinalis L.	Rosemary	ROSE	Lamiaceae	Colline belt	https://doi.org/10.15468/dl.ufpb7u
Salvia officinalis L.	Sage	SAGE	Lamiaceae	Colline belt	https://doi.org/10.15468/dl.jpzs6g
Sambucus nigra L.	Elderberry	ELBE	Caprifolia-ceae	Colline - montane belts	https://doi.org/10.15468/dl.bhdjmu
Saxifraga rotundifolia L.	Round-leaved saxifrage	RLSA	Saxifraga-ceae	Montane - subalpine belts	https://doi.org/10.15468/dl.bafj7c
Thymus vulgaris L.	Thyme	THYM	Lamiaceae	Colline - montane belts	https://doi.org/10.15468/dl.8p4f45
Vaccinium vitis-idea L.	Lingonberry	LIBE	Ericaceae	Montane - subalpine belts	https://doi.org/10.15468/dl.svmy95
Verbena officinalis L.	Common vervain	VERV	Verbena-ceae	Colline - montane belts	https://doi.org/10.15468/dl.z53cbq
Veronica officinalis L.	Common speedwell	SPEE	Scrophu-lariaceae	Montane - subalpine belts	https://doi.org/10.15468/dl.m7664z

2.4. Environmental variable preparation

Twenty-nine environmental variables - climatic, topographic or edaphic - were prepared at a 1 km × 1 km spatial resolution (Table 2). Current and future European bioclimatic data at a 30 arc-seconds resolution (~1 km^2) were recovered from CHELSA (Karger et al., 2017, 2018). To assess climate sensitivity, we worked with CMIP6 scenario SSP585 (see Appendix S1 for details), the worst case climatic scenario assuming a radiative forcing of 8.5 W/m^2. We retrieved bioclimatic variables representing the average of 5 different climatic models for the 2011-2040 time range. Slope, elevation, and aspect (sine and cosine) were upscaled from the existing 250 m resolution digital elevation model GMTED (Amatulli et al., 2018), using bilinear interpolation. Bilinear interpolation assigns a value at the output cell by taking the weighted average of the four nearest cell centres. The closer an input cell centre is to the output cell centre, the more influence its value will have on the output cell value. Likewise, edaphic variables were retrieved from SoilGrids (Poggio et al., 2021) and aggregated from 250 m resolution data to 1 km using bilinear interpolation. Finally, we recovered the WorldClim (Fick & Hijmans, 2017) solar radiation data at 30 arc-seconds resolution (~1 km^2) and computed the mean of the 12 months. For the future predictions, we assumed that the physical soil properties and topography would be unchanged by 2040.

All 29 variables were aligned to have the exact origin, with the CHELSA bioclimatic variables as reference. They were also reprojected in the projection system used in this study, EPSG:3035.

Table 2. List of environmental variables prepared at a 1 × 1 km m resolution and tested for multi-collinearity

Variable	Description and unit	Reference
bio1	Mean annual air temperature (°C)	CHELSA (Karger et al., 2017, 2018)
bio2	Mean diurnal air temperature range (°C)	
bio3[a]	Isothermality (i.e. ratio of diurnal variation to annual variation in temperatures) (°C)	
bio4[a]	Temperature seasonality (i.e. standard deviation of the monthly mean temperatures) (°C)	
bio5	Mean daily maximum air temperature of the warmest month (°C)	
bio6	Mean daily minimum air temperature of the coldest month (°C)	
bio7	Annual range of air temperature (i.e. the difference between the maximum temperature of warmest month and the minimum temperature of coldest month) (°C)	
bio8[a]	Mean daily mean air temperatures of the wettest quarter (°C)	
bio9[a]	Mean daily mean air temperatures of the driest quarter (°C)	
bio10	Mean daily mean air temperatures of the warmest quarter (°C)	
bio11	Mean daily mean air temperatures of the coldest quarter (°C)	
bio12	Annual precipitation amount (kg × m^{-2})	
bio13[a]	Precipitation amount of the wettest month (kg × m^{-2})	
bio14	Precipitation amount of the driest month (kg × m^{-2})	
bio15[a]	Precipitation seasonality (i.e. standard deviation of the monthly precipitation estimates) (kg × m^{-2})	
bio16	Mean monthly precipitation amount of the wettest quarter (kg × m^{-2})	
bio17	Mean monthly precipitation amount of the driest quarter (kg × m^{-2})	
bio18	Mean monthly precipitation amount of the warmest quarter (kg × m^{-2})	
bio19[a]	Mean monthly precipitation amount of the coldest quarter (kg × m^{-2})	
Slope[a]	Slope angle (°)	GMTED (Amatulli et al., 2018)
Elevation	Altitude (m a.s.l.)	
Aspect (sine)[a]	Slope direction (east-west)	
Aspect (cosine)	Slope direction (north-south)	
bulk[a]	Bulk density of the fine earth fraction (cg/cm^3)	SoilGrids (Poggio et al., 2021)
coarse[a]	Volumetric fraction of coarse fragments (> 2 mm) (vol‰)	
clay[a]	Proportion of clay particles (< 0.002 mm) in the fine earth fraction (g/kg)	
sand	Proportion of sand particles (> 0.05 mm) in the fine earth fraction (g/kg)	
silt[a]	Proportion of silt particles (≥ 0.002 mm and ≤ 0.05 mm) in the fine earth fraction (g/kg)	
Solar radiation[a]	Mean of monthly average of daily global solar radiation (kJ × m^{-2} × day^{-1})	WorldClim (Fick & Hijmans, 2017)

[a] Denotes variables that were kept for model calibration.

2.5. Environmental variable selection

To avoid multi-collinearity among the environmental variables, we performed bivariate Pearson's Correlation tests between all pairs of the 29 candidate variables. They were assessed sequentially in decreasing order of the Pearson Correlation coefficient values. The collinearity threshold was set at 0.70 in absolute value (Dormann et al., 2013). Thus, only variables "non-collinear" with any previously added variable were kept (i.e. pairs that had a correlation coefficient $r < |0.7|$). This strategy was applied one time to the whole set of occurrences of all species. As a result, out of the original 29 variables, 14 significant variables were kept for model calibration (Table 2).

2.6. Model calibration

For each species, models were calibrated with a set of explanatory variables corresponding to the previously selected variables. Using the biomod2 package (Thuiller et al., 2021; Thuiller et al., 2009) and the R v 4.0 software, probabilistic models were calibrated using four different statistical techniques: generalised linear models (GLM; McCullagh & Nelder, 2019), generalised additive models (GAM; Hastie & Tibshirani, 2017), boosted regression trees (GBM; Ridgeway, 1999), and random forest (RF; Breiman, 2001). These models associate the response to the explanatory variables to derive the probability of a pixel to host a given target species. While GLM is a traditional regression algorithm that allows the response variable to take several distributions and non-constant variance functions to be modelled, GAM is a semi-parametric extension of GLM that allows to implement non-parametric smoothers. Regarding GBM and RF, they are boosting and bagging tree-based (recursive partitioning) approaches respectively. Summarising, GLM and GAM are both regression approaches fitting a single model at a time, whereas GBM and RF are tree-based approaches that are averaging many models at each run. Working with four techniques within two different categories (regressions and trees) is an advantage as it allows comparing their predictions and assessing the associated uncertainty. Furthermore, these four techniques are among the most used and the choice of one over the others is not easy.

We finally built a consensual projection of potential medicinal and aromatic plant distribution by combining the results of the four different individual modelling techniques (GLM, GAM, GBM, and RF) into an ensemble model (EM; Araújo & New, 2007; Thuiller, 2004). Using ensemble modelling aims to reduce the error in the prediction. As long as the base models are diverse and independent, the prediction error decreases when the ensemble approach is used (Kotu & Deshpande, 2019). In the present case, we have decided to mix all models (i.e. all techniques and cross-validations runs; see Section ‘‘2.7’’) to produce our ensemble model. This one was obtained by averaging the individual model projections after having weighted them proportionally to the maximised value if the true skill statistic (maxTSS) score, which is a measure of model quality (Guisan et al., 2017; see Section “2.7”). This method has been shown to be particularly robust (Engler et al., 2011; Marmion et al., 2009).

All the mentioned modelling techniques require both presence and absence data. Yet, we only had species presence records available. We thus generated a set of 10'000 pseudo-absences (Barbet-Massin et al., 2012) using the biomod2 package (Thuiller et al., 2021, 2009), as now commonly done (Buisson et al., 2010). So that the models are not biased towards an overestimation of presence or pseudo-absence and to limit spatial autocorrelation, presence and pseudo-absence data were weighted equally during model calibration so that both had equal prevalence.

2.7. Model evaluation

The predictive power of individual models was evaluated through a repeated random data-splitting procedure. Each model was trained on 70% of the data (chosen randomly) and evaluated on the remaining 30% using two evaluation metrics: the maximisation of the true skill statistic (maxTSS; Allouche et al., 2006; see Guisan et al., 2017 for the maximisation procedure) and the area under the curve (AUC) of a receiver operating characteristics (ROC) plot (Fielding & Bell, 1997). TSS (i.e. its maximisation) and AUC are two complementary accuracy methods widely used to assess model performance (Shabani et al., 2018). TSS is calculated as specificity (fraction of correctly predicted presences) + sensitivity (fraction of correctly predicted absences) – 1 at every probability threshold (e.g. every 0.01 increment) and the maxTSS is obtained by taking the maximum TSS value across all. MaxTSS thus allows to evaluate the concordance between predicted and observed values independent of a specific threshold. The TSS metric varies between negative values (systematically wrong), 0 (random model) and 1 (perfect accordance). For example, a TSS of 0.5 means that the proportion of correctly predicted presences and absences is roughly 75% (0.75 + 0.75–1 = 0.5). As for the maxTSS metric, the AUC provides a single measure of overall accuracy that is not dependent upon a particular threshold (DeLeo, 1993), but is obtained through a measure of the area under the curve of a ROC plot obtained by plotting the combinations of sensitivity and [1 –

specificity] at all thresholds. Its value ranges between 0.5 (area under the diagonal line, meaning that the scores of the two groups do not differ) and 1 (no overlap in the distributions of the group scores). For instance, an AUC of 0.8 means that in 80% of the time, a random selection from the positive group (e.g. absences correctly classified as absences) will score higher than a random selection from the negative class (e.g. presences wrongly classified as absences).

The maximised TSS metric was further used as an evaluation reference to build our ensemble models. The quality threshold was set at 0.4, which is the threshold above which models are considered "useful" (Engler et al., 2013). Thus, only individual models with a maxTSS greater than or equal to this value were kept.

The entire split-sample evaluation procedure was repeated 25 times for each individual model. Therefore, for each species, 101 models were generated (25 resampling run models x 4 modelling techniques + 1 ensemble model). The final ensemble model was also assessed using the same metrics as in the previous steps.

2.8. Assessing the species richness and composition of medicinal and aromatic plant assemblages

2.8.1. Species richness maps

Projections from our ensemble models are continuous probability values between 0 and 1. To obtain an estimate of the medicinal and aromatic plant species richness in each site, we summed the projections that we got from our single species ensemble models (Dubuis et al., 2011). Summing predicted probabilities is a more advised approach than summing thresholded (i.e. binarised) individual species predictions (Calabrese et al., 2014), as it does not require a species-specific threshold but instead uses site-specific ecological constraints (i.e. through defining the individual species probabilities) to estimate richness in each site, which can then be also used to predict composition ("site-threshold"; Scherrer et al., 2018). Furthermore, it has been shown that summing predicted probabilities was a more robust approach as it yields an unbiased estimate of species richness at each prediction site (Calabrese et al., 2014; Dubuis et al., 2011), i.e., unlike summing binary predictions, it does not overpredict richness. Using this approach, we obtained current and future richness maps where each site was attributed an estimated number of species, ranging between 0 and 27, suitable to grow at this location. For example, with a value of 8.5, it is estimated that this location has the necessary conditions for the growth of about eight different species.

2.8.2. Ranking of probabilities of presence

In each study site, we wanted to assess the composition of the community of medicinal and aromatic species and know which species had the greatest probability of being present. We, therefore, used the probability ranking rule (PRR; D'Amen et al., 2015). In combination with summing predicted probabilities to estimate richness, the PRR allows identifying which species are the most likely to be present based on the ranking of the predicted probabilities of occurrence of all species in each site up to the (sum of probabilities-based) estimated richness (Scherrer et al., 2020). More specifically, the probabilities of occurrence calculated by the SDMs were ranked in decreasing order, the species with the highest probability of occurrence being classified first and the one with the lowest probabilities being last (in our case number 27). A number of species equal to the expected species richness previously determined (see Section "2.8.1") was then selected for each site. Using this approach, we obtained, for each species, current and future ranking maps where each site was attributed the rank of the species, ranging between 1 and 27. In addition, using pie charts, we observed in detail MAP species composition of 5 following Swiss farm locations: Attiswil (BE), Ebnat-Kappel (SG), Noble-Contrée (VS), Poschiavo (GR), Soral (GE).

2.8.3. Ordination for species associations analysis

We used ordination applied to the predictions obtained from our single species ensemble models to identify co-occurrence patterns (i.e. associations) in our medicinal and aromatic plant assemblages. The use of ordination allows summarising a multidimensional dataset so that, when it is projected onto a reduced dimensional space, any intrinsic pattern the data may possess becomes visually apparent (Pielou, 1984). Given the nature of our data (continuous probabilities), we had recourse to a Principal Component Analysis (PCA) to define our ordination space. This allowed us to uncover how the species were associated in current predicted plant assemblages and if potential changes in composition would happen in the future.

3. Results

3.1. Environmental variable selection

The 14 variables that were kept for model calibration had pairwise Pearson Correlation coefficient *r* smaller than |0.7| (Figure S2). Among them, half are bioclimatic variables. Three of them represent variations in climatic factors like bio3, bio4 and bio15. The others correspond to measurements of temperature (i.e. bio8 and bio9) and precipitation (i.e. bio13 and bio19) at different times of the year. All edaphic variables were selected except the proportion of sand in the fine earth fraction. Finally, the slope, slope direction (east-west) and solar radiation were also retained.

3.2. Model performance

All individual models obtained a TSS score greater than 0.4 which is the threshold above which models are considered “useful” (Engler et al., 2013; Figure S3). For this reason, all individual modelling techniques were considered for computing the ensemble models (EMs). The algorithm that obtained the best TSS and AUC scores across all species was RF.

When considering the EM projections, all species obtained elevated evaluation scores, with TSS values between 0.696 and 1 (Table 3). Such TSS values mean that, on average, presence and absence occurrences of a species were correctly predicted with a rate of 85–100%. *Artemisia vallesiaca* is the species that obtained the highest evaluation score, with a TSS value of 1. The majority of species achieved a TSS score greater than 0.85. Only 2 species had a score just below 0.7, *Achillea millefolium* and *Salvia officinalis*, with TSS values of 0.698 and 0.696 respectively, but in absolute terms their models remain very good. Regarding the AUC (= ROC), all species achieved values between 0.933 and 1, which are considered as good to excellent predictions (Guisan et al., 2017; Swets, 1988). More than two-thirds of the species even scored more than 0.95 for AUC, meaning that for more than 95% of the time a random selection from the positive group (e.g. absences correctly classified in absences) will have a score greater than a random selection from the negative class (e.g. presences wrongly classified as absences; DeLeo, 1993).

In view of the very good evaluation scores obtained, solely the results of the EM projections were used to produce the following results.

Table 3. TSS and ROC values for ensemble models of each species

Species	TSS	ROC
Achillea millefolium L.	0.698	0.933
Alchemilla xanthochlora Rothm.	0.857	0.985
Althaea officinalis L.	0.730	0.949
Artemisia umbelliformis Lam.	0.951	0.997
Artemisia vallesiaca All.	1	1
Cannabis sativa L.	0.938	0.992
Centaurea cyanus L.	0.808	0.971
Gentiana lutea L.	0.857	0.984
Leontopodium alpinum Cass.	0.932	0.996
Malva sylvestris L.	0.748	0.947
Marrubium vulgare L.	0.851	0.977
Melissa officinalis L.	0.738	0.945
Mentha ×*piperita* L.	0.929	0.991
Ocimum basilicum L.	0.999	1
Origanum vulgare L.	0.733	0.948
Pimpinella saxifraga L.	0.851	0.982
Plantago lanceolata L.	0.739	0.952
Primula veris L.	0.766	0.958
Rhodiola rosea L.	0.969	0.998
Rosmarinus officinalis L.	0.933	0.995
Salvia officinalis L.	0.696	0.934
Sambucus nigra L.	0.716	0.941
Saxifraga rotundifolia L.	0.855	0.983
Thymus vulgaris L.	0.849	0.980
Vaccinium vitis-idea L.	0.852	0.983
Verbena officinalis L.	0.797	0.967
Veronica officinalis L.	0.803	0.967

3.3. Predictions of current and future medicinal and aromatic plant species distributions

For each species, the maps of the predicted probabilities of presence in the present and future are available in Appendix S4.1-S4.27. Under current climate, some species are predicted to cover almost the entire EUSALP area, such as *Achillea millefolium* or *Origanum vulgare*. Others, like *Artemisia umbelliformis* or *Leontopodium alpinum*, are predicted to be only present at higher altitudes in the Alps. There are on the contrary species for which the predictions of their current distribution do not include the Alps, as *Althaea officinalis* or *Melissa officinalis*.

The maps of predicted future distributions reflect the impact of climate change on the species distributions. As these maps show, species would not react to climate change in the same way, rather experiencing substantial but different changes in their distributions. Two different types of changes are highlighted by our results (Table S5). While some species should persist in the same sites but with a reduced probability of occurrence, some should appear in new places where they were not previously observed. The latter, in addition to remaining present in the same places as now (although with a declining probability of presence), should also colonise a great part of the rest of the EUSALP area. Taking up some of the examples given above, A. *officinalis* and *M. officinalis* are two medicinal and aromatic plants which will, in the future, "go up" in the Alps and invade higher altitude areas. *A. umbelliformis* and *O. vulgare*, for their part, are two species that will see their probability of occurrence get reduced in the future and throughout the study area.

3.4. Location, richness, and composition of medicinal and aromatic plant assemblages

We computed the current and future species richness maps of medicinal and aromatic plants (MAPs) by summing the predicted probabilities of presence obtained from the ensemble models (EM) of all our studied species (Figure 3). Currently, the MAP assemblages with the highest species richness (i.e., the highest hotspots) are found in Switzerland, with values up to 18.2 (i.e. ca. 18 species; Figure 3A). As a general rule, we observe that the further away the assemblages are from Switzerland, the lower their richness. This is the case for instance in Germany, where plant assemblages are richer in the south (up to 8.5) than in the north (from 1.86 to 6.73). The south of France also contains many hotspots that can suit only around 10 of our selected species. By 2040, the EMs predict a drastic decrease in the assemblages of these MAPs (Figure 3B). However, the pattern observed in the future seems to remain the same as the one currently predicted. Indeed, Switzerland remains the area with the richest assemblages of these plants, this time with values reaching only around 8.5. For the rest of the EUSALP zone, the sum of probabilities does not exceed 5. Nonetheless, it seems that in certain regions, especially at the northern border of the EUSALP area, the hotspots have slightly increased. In the same way as we observed with the probability of occurrence maps for each species (see Section “3.3”), two distinct events occur in our future predictions: (1) the probability of presence decreases overall for all the species and (2) new areas are colonised by some of them.

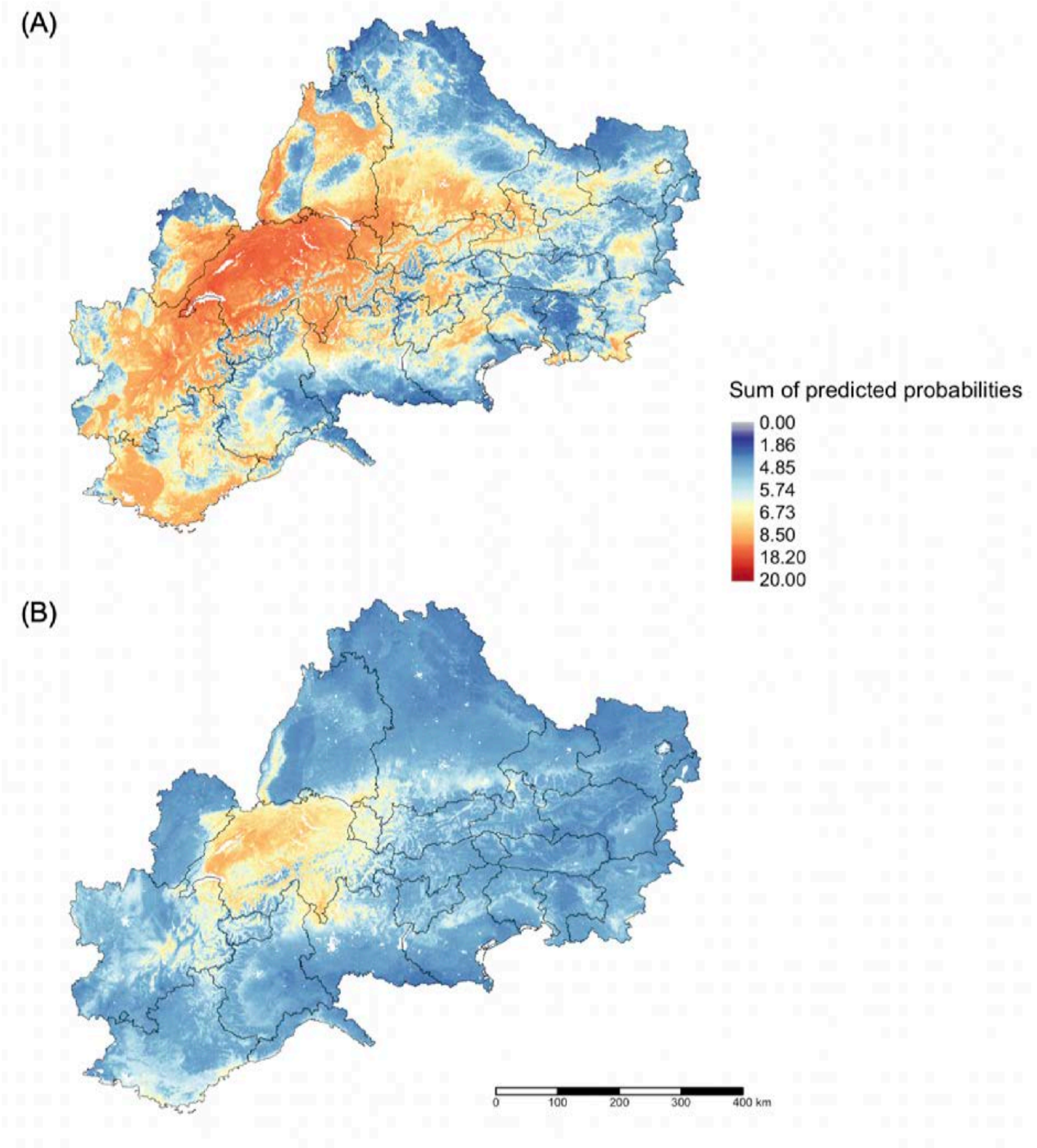


Figure 3. Estimated species richness (A) in the present and (B) by 2040 in the EUSALP region. The richness values were computed by summing the predicted probabilities of each species obtained from the ensemble models.

Regarding the composition of plant assemblages, we examined the ranking of the probabilities of occurrence for each species in the whole of the EUSALP area. The resulting maps for 3 of the species - *Althaea officinalis*, *Gentiana lutea* and *Thymus vulgaris* **-** are presented in this section (Figure 4) and the other maps can be found in Appendix S6.1-S6.24. The ranking of probabilities per site highlighted 3 types of patterns, represented by the 3 species present in Figure 4 (see Appendix S7 for classification of all species). By 2040, either (1) species will move up overall in the ranking (per site) over the entire area like *A. officinalis* (Figure 4; first row), (2) go down overall, as for *G. lutea* (Figure 4; second row), or (3) do both at different locations of the study area, as for *T. vulgaris* (Figure 4; third row). Therefore, climate change would not only impact species distribution in species-specific ways, but also the species per-site rankings. The fact that climate change can affect the ranking of the per-site species' probabilities of occurrence in multiple ways indicates that the composition of assemblages will suffer major modifications in the years to come.

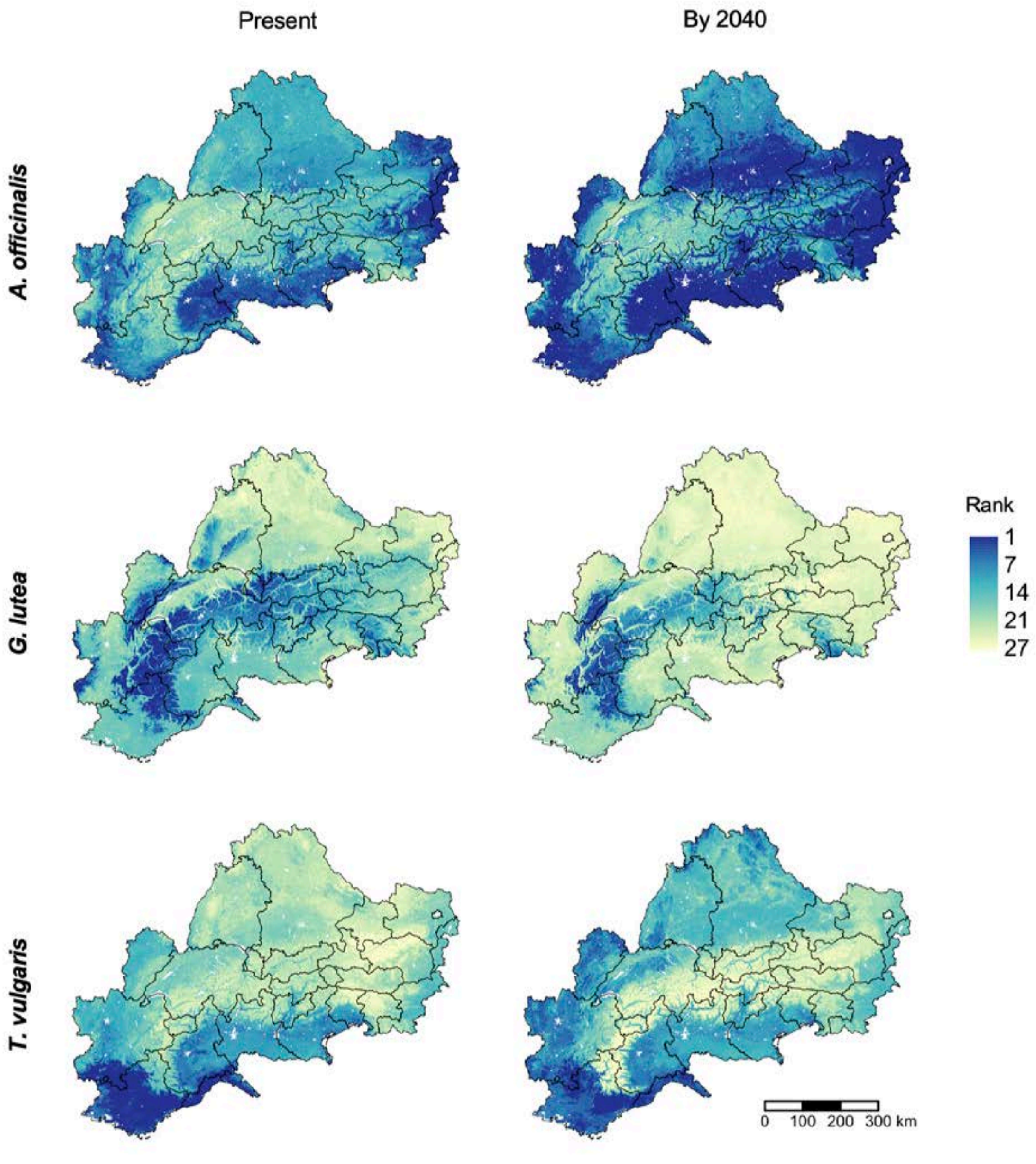


Figure 4. Ranking of occurrence probabilities of *Althaea officinalis* (first row), *Gentiana lutea* (second row) and *Thymus vulgaris* (third row) in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

We focus now more particularly on Switzerland, where we saw that the assemblages were the richest. To understand how the ranking influence the composition of medicinal and aromatic plants (MAPs) assemblage, we studied 5 locations, all situated in different cantons, where we know that existing farms grow MAPs species. Each farm location has a high species richness, ranging from around 14 to 16, except for Poschiavo (GR) whose species richness is 7 (Figure 5A). The assemblages are made up of various species and 7 of the 27 species studied, such as *Primula veris* and *Veronica officinalis*, can be found in each of the 5 farm locations (Figure 5B). However, all without exception experience a decrease in richness in the future, by at least a half for 4 out of 5 farm locations (Figure 5). *P. veris* and *V. officinalis* are expected to disappear from all assemblages in the future. In addition to have their richness greatly reduced, assemblages of these MAPs will thus also be subject to changes in their composition. For instance, according to our models, *Althaea officinalis* should in the future be present in Noble-Contrée (VS) and Poschiavo (GR), whereas it is currently absent.

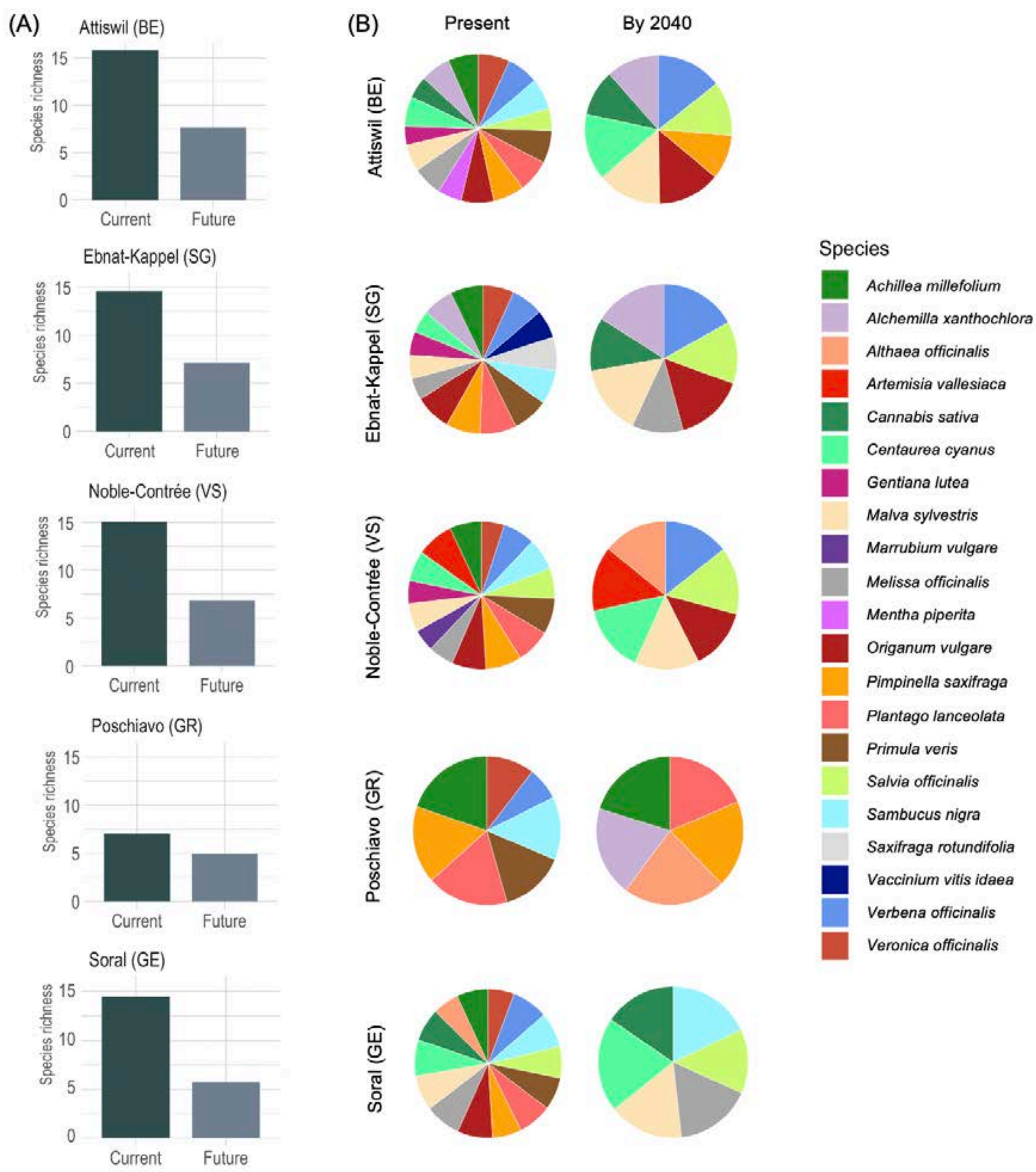


Figure 5. (A) Current and future estimated species richness (computed by summing the predicted probabilities) of medicinal and aromatic plants at five places in Switzerland. (B) Corresponding pie charts showing which species are predicted to be present in each location, in the present and by 2040. Pie charts were made using the probabilities of the highest ranked species for each location (depending on the specific richness at each location).

Changes in the composition of species assemblages can also be observed through the results of the Principal Component Analysis (PCA; Figure 6). Species are more sparsely distributed in the ordination space in the present (Figure 6A) than in the future (Figure 6B), where most species cluster along the x-axis. These species in particular will in the future have similar probabilities of presence. These shifts in the ordination space attest once again that the climate change will affect many factors, influencing the location and composition of medicinal and aromatic plant assemblages in the future.

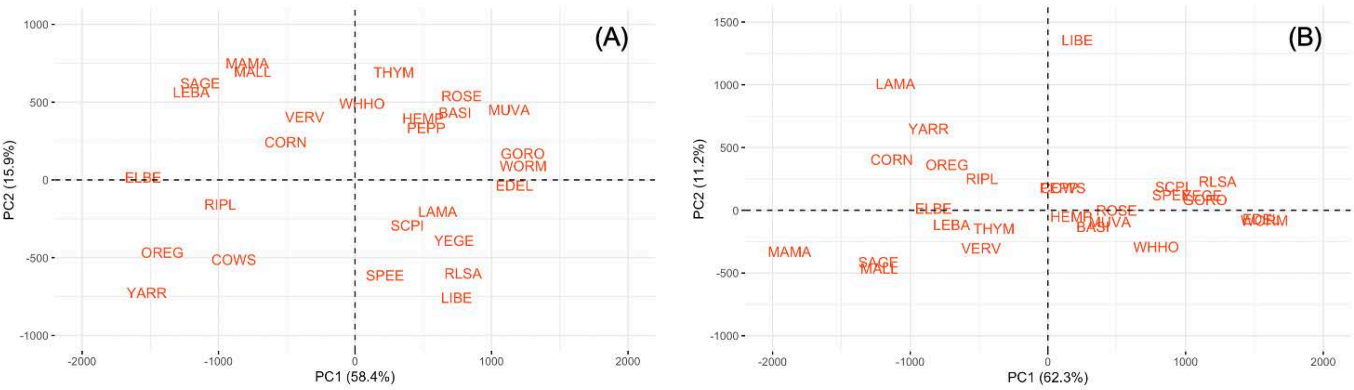


Figure 6. Principal component analysis results: (A) PCA score plot of the first two principal components (total explained variance equal to 74.3%) for the current predictions. (B) PCA score plot of the first two principal components (total explained variance equal to 73.5%) for the future predictions. The species are indicated by their abbreviations (see Table 1).

4. Discussion

In this study, we mapped the spatial distribution of 27 medicinal and aromatic plant (MAP) species and their assemblages across the Alpine region at a 1 km spatial resolution, to identify potential sites that would be most optimal for MAP cultivation. We thus also assessed the richness and composition of the species assemblages at each site across the study area. Lastly, we investigated the impact of climate change on the location, richness, and composition of the MAP assemblages. Our results show that the Alpine region has the potential to supports many MAP species, with Switzerland hosting the richest species assemblages of our selected MAPs. The composition of the MAP assemblages varies from site to site, as the latter do not supply the same environmental conditions. Our findings also demonstrated that, in the future, climate change will affect the richness and composition of MAP assemblages and, hence, which species would be the most suitable to ensure a productive yield. As we hypothesised, MAP diversity will decrease with global warming in the short term: in the Swiss Plateau, around half of MAP species will disappear locally in the future. Besides changes in species richness, we also observed changes in the way species would assemble with each other in the future. Nevertheless, a number of our studied species should still be present in the future and compose part of the MAP assemblages of the Alpine area. We could, however, not assess here which new MAP species could potentially colonise the Alpine zone in a warmer future, which would be an interesting perspective for future studies.

4.1. Comparison of model performance across modelling techniques and species

As we suspected, the ensemble modelling (i.e. combination of the results of the 4 different individual modelling techniques) showed greater accuracy than any of its individual components. Comparing the predictive power of the 5 different modelling techniques (i.e. GLM, GAM, GBM, RF and EM; see section “2.6”), we found that ensemble modelling represent overall the best performing approach (Table 3, Figure S3). The TSS values we obtained from our EM projections for each species ranged between ~0.7-1 (Table 3). On average, about 85-100% of the independent evaluation samples were correctly classified by the models, hence making the ensemble modelling technique the one that produced the highest rates of correctly classified samples. This result is consistent with those of previous studies using this same approach (Engler et al., 2011; Marmion et al., 2009).

When computing the EMs, as all individual models produced for each species obtained a TSS score greater than 0.4 (Figure S3), none of our selected species was removed from the study. Moreover, as said above, all species obtained very good to excellent scores, both for TSS and AUC values obtained from the EM projections (Table 3).

Thus, no major variation in accuracy was observed between the studied species (i.e. between their relative EM). The precision of the fitted models obtained from the ensemble modelling approach therefore indicates that reliable distribution maps of all our studied species can be produced and subsequently used for a range of analyses and research questions.

4.2. Accuracy of medicinal and aromatic plant habitat suitability forecasts

We were able to fairly accurately predict the distribution and assemblages of medicinal and aromatic plants (MAPs) across the Alpine region with niche-based environmental suitability models. Furthermore, the five locations studied in Switzerland constitute validation points for our forecasts, as they coincide with the location of actual farms cultivating MAPs.

The present environmental conditions hence largely contribute to explain the distribution of MAP species and their assemblages, as well as their response to future climate change. As a result, and as indicated by our findings, the predicted MAP assemblages are composed of different species and are of varying richness (Figure 3) depending on the environment (and sites). Several previous studies suggested that, in extreme regions as alpine environments, species distributions are mainly shaped by climate (Araújo & Guisan, 2006; Thuiller et al., 2004). The accuracy of environmental inputs (i.e. correctness of environmental predictors themselves) in species distribution modelling (SDM) is thus essential to fully understand how the environment influences some dimensions of biodiversity (here MAPs) and might cause or increase bias in the associated spatial predictions (Waltari et al., 2014). Studies carried out in mountainous regions, which show low bioclimatic congruence (i.e. degree of agreement in temperature and precipitation values between bioclimatic databases) due to lacking or unevenly distributed network of weather stations (Hijmans et al., 2005), are particularly prone to errors in spatial predictions (Beaumont et al., 2007; Soria-Auza et al., 2010). In regards to bioclimatic inputs, it seems that the CHELSA database, used here, is more suited than other global bioclimatic databases for terrestrial species (Morales-Barbero & Vega-Álvarez, 2019).

Likewise, especially in mountain ecosystems with complex topography, edaphic components are necessary for a good understanding of the distribution of biodiversity (Buri et al., 2020; Cianfrani et al., 2019; Mod et al., 2016). Soil is, like climate, one of the principal drivers of plant species distribution in the European Alps (Buri et al., 2020; Chauvier et al., 2021). This is why we chose here to include soil descriptors when predicting MAP species distributions in the Alpine region.

We used species data from two distributional databases: the Global Biodiversity Information Facility (GBIF) and Info Flora. Species presence data resulting from citizen science efforts are increasingly abundant and consequently, spatial predictions using these data are also increasingly used in research on conservation or climate change (Guisan & Thuiller, 2005; Guralnick & Hill, 2009; Jetz et al., 2012). Distributional databases, for which GBIF is the largest online provider of distribution records (Beck et al., 2013), however, can contain considerable biases that compromise the value of predictions (Fithian et al., 2015; Hefley et al., 2017; Robinson et al., 2018). In GBIF, a spatial bias due to an uneven sampling effort is particularly pronounced (Dorazio, 2014; Komori et al., 2020). Differences in funding and data sharing at the national scale indeed lead to huge spatial differences in contributions (Beck et al., 2013). SDMs trained on spatially biased data may reflect correlates of spatial autocorrelation rather than actual species distributions (Anderson & Gonzalez, 2011; Higa et al., 2014; Wisz et al., 2008). It is therefore often necessary to take certain measures to produce reliable results. One possible solution is then to spatially thin the data: i.e. filter data in the areas where data density is highest and keep data where the density is low (Boria et al., 2014; Steen et al., 2021) or try to correct the bias in the model (e.g. El-Gabbas & Dormann, 2018; Hefley et al., 2017; Phillips et al., 2009), although the latter can be difficult to implement without knowing the exact processes generating the biases. For presence-only data, however, the process of spatial thinning is relatively easy. Here, we more specifically chose to: (1) combine two databases: GBIF and Info Flora (see Section "2.3") to have the maximum number of observations in the study area and (2) aggregate data so that there was only one observation per species per square kilometre. This way, we ensured to avoid the over-representation of well-sampled locations. The approach of bias correction within the models could, however, be an interesting perspective for future studies.

4.3. Uneven impact of climate change on medicinal and aromatic plant species

The extreme scenario chosen in this study allows us to consider the worst-case scenario and to cover the other possible and less extreme cases. Our results demonstrated that, in some cases, climate warming will favour medicinal and aromatic plant (MAP) distribution for species like *Alchemilla xanthochlora*, *Centaurea cyanus* or *Mentha piperita*, which will benefit from higher temperatures by colonising new areas in the Alpine region. However, global warming will mainly reduce MAP species presence, possibly causing local extinctions at the end. These two different types of changes in the MAP distribution were highlighted by our EMs (Table S5). Besides, although some species should colonise new terrain, the whole Alpine region will not gain species since most species included here were already in the region and are accordingly rather expected to decline, as we observed with the 5 specific Swiss cases (Figure 5). Yet, we did not account for potential MAP species that could colonise the area in the future (e.g. from Southern regions in Europe) and change the MAP assemblages accordingly.

After accounting for the response of MAP species to climate change, we found strong changes in ranking and thus in MAP assemblages' composition. Our findings indeed emphasised three different patterns of change in the probability ranking (Table S7). Even if their overall presence probability across the Alpine region will decrease in the future, a few MAP species are predicted to gain in ranking. Two potential reasons can explain this phenomenon which seems contradictory: (1) species invade new sites where they are better adapted than other species and/or (2) their probabilities decrease, but less than the other species.

Therefore, global warming would reshuffles MAP species in space and time in species-specific ways, and thus make future assemblages not necessarily the same as today's ones, as already shown for entire plant communities (Huntley, 1991; Trivedi et al., 2008). Evidences of an uneven impact of climate change on plant species of the European Alps were already observed in previous studies (Klanderud, 2008; Losapio et al., 2021). In the near future, different combinations of MAP species, never observed before, could thus be observed. Taken together, the different types of change (in MAPs distributions and rankings) highlighted throughout this study reveal that there is no single factor, pattern or mechanism affecting the species distribution and assemblages. Instead, the species-specific impact of climate warming on MAP species biodiversity and composition will likely be mediated by complex and potentially novel ecological interactions (Alexander et al., 2015).

4.4. Limitations

This study predicted the spatial distribution of 27 medicinal and aromatic plant (MAP) species across the Alpine region in Europe at a 1 km spatial resolution. In mountain areas, however, habitats and species distributions can rapidly vary with altitude and across the complex topography, creating different exposures and a multitude of different microclimates. A spatial resolution of 1 km is therefore often considered low for modelling species distributions in alpine areas. However, several reasons contributed to justify our choice. First, many environmental variables are nowadays directly available at a resolution of 1 km across large areas, such as the entire Alpine region ($\approx$ 441,006 km^2 for the EUSALP area). In addition, a 1 km resolution represents a good compromise for optimising eco-informatic calculations given this very large extent considered. Our results at this resolution could however be re-used in future studies involving hierarchical spatial models (Chevalier et al., 2021; Mateo et al., 2019).

Although our results predict with overall good accuracy the distribution of the studied MAP species, our predictions are to be considered with care. Indeed, MAP cultivation involves factors other than the natural environment (Liliane & Charles, 2020), as technological (e.g. agricultural practices) and biological factors (e.g. diseases, insects, pests), which are not considered explicitly – i.e. biotic interactions are implicitly considered - in our models. Our findings are, however, providing complementary information on the factors driving the species' ecology and geography, which can then be used to improve the future planning of MAP cultivations while also accounting for local human-related factors.

Presence-only data were used in this study, mostly coming from wild MAP observations. With the ongoing MAPs breeding and domestication effort (Ekiert et al., 2021), it is likely that MAPs could thrive under a broader range of biophysical conditions than predicted in this study. As a consequence, our results probably underestimate the agronomic potential of MAPs. However, this study focused on MAPs as a solution to increase agro-biodiversity through a low-input (i.e. that minimises the use of production inputs such as purchased fertilisers and pesticides) and organic agricultural production (Alrøe & Kristensen, 2004). Relying on wild species distribution is therefore a good proxy of the context in which domesticated (i.e. agricultural) MAPs could evolve if they were introduced into sustainable agriculture.

4.5. Conclusion and potential implications for medicinal and aromatic plant cultivation

Here, we showed that the Alpine region was home to numerous medicinal and aromatic plant (MAP) species. In addition, we demonstrated that, without measures taken to mitigate global warming in the near future, the consequences on MAP diversity will be significant. In view of our results, addressing which and how species will distribute in the Alpine ecosystems while being impacted by global warming is of paramount importance, from both an ecological and economical point of view.

In terms of agricultural applications, mapping the distribution of MAP species and their assemblages may serve for planning of agricultural land. With the information gained from our results, sustainable agriculture may thus be promoted by helping policy makers to support the emergence of MAPs production where it is most ecologically appropriate.

In addition to their great economic value, due to their use as raw materials in many areas of production, MAPs provide ecosystem services essential to human health, livelihood and knowledge (Padulosi et al., 2002). Mapping of spatial distribution of MAPs species and their assemblages has, furthermore, already been used in previous studies as a way to estimate ecosystem services provided across a region (Cheminal et al., 2020; Vári et al., 2020). The same could be replicated in our case, using our results to assess and map the potential of MAP species as ecosystem service providers in the Alpine region. Thus, this would also participate in demonstrating that MAPs cultivation represents an ecologically and economically sustainable opportunity for agricultural areas in the Alpine region.

Cultivating MAPs in the Alpine region with climate change will however become more challenging, as MAP species suitable today may not be tomorrow. Besides, the combination of suitable MAP species is likely to change with global warming. To support the climate-change adaptation of farmers who will engage in MAP production in the upcoming years, there is a need to investigate how future combinations of MAPs can be integrated into farming systems. Therefore, in the face of global climate change, it will be more complex to plan the Alpine landscape to provide sufficient agricultural production, ecosystem services as well as conserve biodiversity.

Acknowledgements

I thank my supervisor, Catherine Pfeifer, for her help and support throughout this project. I also thank Pierre-Louis Rey for his advices on the conceptual part of this study, as well as Olivier Broennimann and Antoine Adde for their valuable help on computing aspects. I thank Prof. Antoine Guisan for providing useful insights throughout this Master, as well as for his comments on previous versions of the manuscript.

References

Alexander, J. M., Diez, J. M., & Levine, J. M. (2015). Novel competitors shape species' responses to climate change. *Nature*, *525*(7570), 515–518. doi: 10.1038/nature14952

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*(6), 1223–1232. doi: 10.1111/j.1365-2664.2006.01214.x

Alrøe, H. F., & Kristensen, E. S. (2004). Why have basic principles for organic agriculture? … and what kind of principles should they be? *Ecology & Farming*, *36*, 27–30.

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J., & Jetz, W. (2018). A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific Data*, *5*(1), 180040. doi: 10.1038/sdata.2018.40

Anderson, R. P., & Gonzalez, I. (2011). Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, *222*(15), 2796–2811. doi: 10.1016/j.ecolmodel.2011.04.011

Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., … Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*. (world). doi: 10.1126/sciadv.aat4858

Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, *33*(10), 1677–1688. doi: 10.1111/j.1365-2699.2006.01584.x

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, *22*(1), 42–47. doi: 10.1016/j.tree.2006.09.010

Argyropoulos, D. (2019). *EIP-AGRI Focus Group Plant-based medicinal and cosmetic products* [Starting paper]. European Commission.

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*(2), 327–338. doi: 10.1111/j.2041-210X.2011.00172.x

Beaumont, L., Pitman, A., Poulsen, M., & Hughes, L. (2007). Where will species go? Incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology, 13,* 1368–1385. doi: 10.1111/j.1365-2486.2007.01357.x

Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2013). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, *19*. doi: 10.1016/j.ecoinf.2013.11.002

Benton, T. G., Bieg, C., Harwatt, H., Pudasaini, R., & Wellesley, L. (2021). *Food system impacts on biodiversity loss* (p. 75). London: UN Environment Programme. Retrieved from UN Environment Programme website: http://www.unep.org/resources/publication/food-system-impacts-biodiversity-loss

Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling, 275*, 73–77. doi: 10.1016/j.ecolmodel.2013.12.012

Bornand, C., Gygax, A., Juillerat, P., Jutzi, M., Möhl, A., Rometsch, S., … Eggenberg, S. (2016). *Liste rouge Plantes vasculaires. Espèces menacées en Suisse.* (No. 1621; p. 178). Office fédéral de l'environnement, Berne et Info Flora, Genève.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5–32. doi: 10.1023/A:1010933404324

Buisson, L., Thuiller, W., Casajus, N., Lek, S., & Grenouillet, G. (2010). Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, *16*(4), 1145–1157. doi: 10.1111/j.1365-2486.2009.02000.x

Buri, A., Grand, S., Yashiro, E., Adatte, T., Spangenberg, J. E., Pinto-Figueroa, E., … Guisan, A. (2020). What are the most crucial soil variables for predicting the distribution of mountain plant species? A comprehensive study in the Swiss Alps. *Journal of Biogeography, 47*(5), 1143–1153. doi: 10.1111/jbi.13803

Calabrese, J. M., Certain, G., Kraan, C., & Dormann, C. F. (2014). Stacking species distribution models and adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography*, *23*(1), 99–112. doi: 10.1111/geb.12102

Chappell, M. J., & LaValle, L. A. (2011). Food security and biodiversity: Can we have both? An agroecological analysis. *Agriculture and Human Values*, *28*(1), 3–26. doi: 10.1007/s10460-009-9251-4

Chauvier, Y., Thuiller, W., Brun, P., Lavergne, S., Descombes, P., Karger, D. N., … Zimmermann, N. E. (2021). Influence of climate, soil, and land cover on plant species distribution in the European Alps. *Ecological Monographs, 91*(2), e01433. doi: 10.1002/ecm.1433

Cheminal, A., Kokkoris, I. P., Strid, A., & Dimopoulos, P. (2020). Medicinal and Aromatic Lamiaceae Plants in Greece: Linking Diversity and Distribution Patterns with Ecosystem Services. *Forests, 11*(6), 661. doi: 10.3390/f11060661

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: A Publication of the Ecological Society of America*, *31*(7), e02427. doi: 10.1002/eap.2427

Cianfrani, C., Buri, A., Vittoz, P., Grand, S., Zingg, B., Verrecchia, E., & Guisan, A. (2019). Spatial modelling of soil water holding capacity improves models of plant distributions in mountain landscapes. *Plant and Soil*, *438*(1), 57–70. doi: 10.1007/s11104-019-04016-x

Dal Cero, M., Saller, R., & Weckerle, C. S. (2014). The use of the local flora in Switzerland: A comparison of past and recent medicinal plant knowledge. *Journal of Ethnopharmacology*, *151*(1), 253–264. doi: 10.1016/j.jep.2013.10.035

D’Amen, M., Pradervand, J.-N., & Guisan, A. (2015). Predicting richness and composition in mountain insect communities at high resolution: A new test of the SESAM framework. *Global Ecology and Biogeography*, *24*(12), 1443–1453. doi: 10.1111/geb.12357

Delarze, R., Eggenberg, S., Steiger, P., Bergamini, A., Fivaz, F., Gonseth, Y., … Stucki, P. (2016). *Liste rouge des milieux de Suisse* (p. 33) [Abrégé actualisé du rapport technique 2013 sur mandat de l’Office fédéral de l’environnement (OFEV)]. Berne.

DeLeo, J. M. (1993). Receiver operating characteristic laboratory (ROCLAB): Software for developing decision strategies that account for uncertainty. *Proceedings of the Second International Symposium on Uncertainty Modelling and Analysis*, 318–325. College Park, MD: IEEE Computer Society Press. doi: 10.1109/ISUMA.1993.366750

Dobremez, L., Chazoule, C., Loucougaray, G., Pauthenet, Y., Nettier, B., Lavorel, S., … Fleury, P. (2015). Debate and controversy surrounding the intensification of grassland use in the Vercors region of France: The farming practices and concepts under discussion. *Fourrages, 2015*, 33–45.

Dorazio, R. M. (2014). Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. *Global Ecology and Biogeography*, *23*(12), 1472–1484. doi: 10.1111/geb.12216

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., … Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. doi: 10.1111/j.1600-0587.2012.07348.x

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*(6), 1122–1131. doi: 10.1111/j.1472-4642.2011.00792.x

EIP-AGRI. (2020). *EIP-AGRI Focus Group Plant-based medicinal and cosmetic products* [Final report]. European Commission.

Ekiert, H. M., Ramawat, K. G., & Arora, J. (2021). *Medicinal Plants: Domestication, Biotechnology and Regional Importance*. Springer Nature.

El-Gabbas, A., & Dormann, C. F. (2018). Improved species-occurrence predictions in data-poor regions: Using large-scale data and bias correction with down-weighted Poisson regression and Maxent. *Ecography*, *41*(7), 1161–1172. doi: 10.1111/ecog.03149

Engler, R., Randin, C. R., Thuiller, W., Dullinger, S., Zimmermann, N. E., Araújo, M. B., … Guisan, A. (2011). 21st century climate change threatens mountain flora unequally across Europe. *Global Change Biology*, *17*(7), 2330–2341. doi: 10.1111/j.1365-2486.2010.02393.x

Engler, R., Waser, L. T., Zimmermann, N. E., Schaub, M., Berdos, S., Ginzler, C., & Psomas, A. (2013). Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution. *Forest Ecology and Management*, *310*, 64–73. doi: 10.1016/j.foreco.2013.07.059

EU. (2020). *Farm to Fork Strategy: For a fair, healthy and environmentally-friendly food system*. European Union.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology, 37*(12), 4302–4315. doi: 10.1002/joc.5086

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38–49. doi: 10.1017/S0376892997000088

Firbank, L. G. (2005). Striking a new balance between agricultural production and biodiversity. *Annals of Applied Biology*, *146*(2), 163–175. doi: 10.1111/j.1744-7348.2005.040078.x

Fithian, W., Elith, J., Hastie, T., & Keith, D. A. (2015). Bias correction in species distribution models: Pooling survey and collection data for multiple species. *Methods in Ecology and Evolution*, *6*(4), 424–438. doi: 10.1111/2041-210X.12242

FOEN. (2017). *Biodiversity in Switzerland: Status and Trends* (No. 1630; p. 60). Bern: Federal Office for the Environment.

Franklin, J. (2010). *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511810602

Geiger, F., Bengtsson, J., Berendse, F., Weisser, W. W., Emmerson, M., Morales, M. B., … Inchausti, P. (2010). Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. *Basic and Applied Ecology*, *11*(2), 97–105. doi: 10.1016/j.baae.2009.12.001

Gellrich, M., Baur, P., Koch, B., & Zimmermann, N. (2007). Agricultural land abandonment and natural forest re-growth in the Swiss mountains: A spatially explicit economic analysis. *Agriculture, Ecosystems & Environment*, *118*, 93–108. doi: 10.1016/j.agee.2006.05.001

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*(9), 993–1009. doi: 10.1111/j.1461-0248.2005.00792.x

Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). *Habitat Suitability and Distribution Models: With Applications in R*. Cambridge: Cambridge University Press. doi: 10.1017/9781139028271

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., … Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424–1435. doi: 10.1111/ele.12189

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2), 147–186. doi: 10.1016/S0304-3800(00)00354-9

Guralnick, R., & Hill, A. (2009). Biodiversity informatics: Automated approaches for documenting global biodiversity patterns and processes. *Bioinformatics (Oxford, England)*, *25*(4), 421–428. doi: 10.1093/bioinformatics/btn659

Gurung, A. B., Wymann von Dach, S., Price, M. F., Aspinell, R., Balsiger, J., Baron, J. S., … Kohler, T. (2012). Global Change and the World’s Mountains-Research Needs and Emerging Themes for Sustainable Development A Synthesis From the 2010 Perth II Conference. *Mountain Research and Development*, *32*(1), S47–S54.

Hao, T., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2020). Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography*, *43*(4), 549–558. doi: 10.1111/ecog.04890

Hastie, T. J., & Tibshirani, R. J. (2017). *Generalized Additive Models*. Boca Raton: Routledge. doi: 10.1201/9780203753781

Hefley, T. J., Brost, B. M., & Hooten, M. B. (2017). Bias correction of bounded location errors in presence-only data. *Methods in Ecology and Evolution*, *8*(11), 1566–1573. doi: 10.1111/2041-210X.12793

Higa, M., Yamaura, Y., Koizumi, I., Yabuhara, Y., Senzaki, M., & Ono, S. (2014). Mapping large-scale bird distributions using occupancy models and citizen data with spatially biased sampling effort. *Diversity and Distributions*, *21*. doi: 10.1111/ddi.12255

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. doi: 10.1002/joc.1276

Huntley, B. (1991). How Plants Respond to Climate Change: Migration Rates, Individualism and the Consequences for Plant Communities. *Annals of Botany*, *67*, 15–22.

Jetz, W., McPherson, J. M., & Guralnick, R. P. (2012). Integrating biodiversity distribution knowledge: Toward a global map of life. *Trends in Ecology & Evolution*, *27*(3), 151–159. doi: 10.1016/j.tree.2011.09.007

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., … Kessler, M. (2017). Climatologies at high resolution for the earth’s land surface areas. *Scientific Data*, *4*(1), 170122. doi: 10.1038/sdata.2017.122

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., … Kessler, M. (2018). *Data from: Climatologies at high resolution for the earth's land surface areas* (Version 1, p. 7266827510 bytes) [Data set]. Dryad. doi: 10.5061/DRYAD.KD1D4

Klanderud, K. (2008). Species-specific responses of an alpine plant community under simulated environmental change. *Journal of Vegetation Science*, *19*(3), 363–372. doi: 10.3170/2008-8-18376

Komori, O., Eguchi, S., Saigusa, Y., Kusumoto, B., & Kubota, Y. (2020). Sampling bias correction in species distribution models by quasi-linear Poisson point process. *Ecological Informatics*, *55*, 101015. doi: 10.1016/j.ecoinf.2019.101015

Kotu, V., & Deshpande, B. (2019). Chapter 2—Data Science Process. In V. Kotu & B. Deshpande (Eds.), *Data Science (Second Edition)* (pp. 19–37). Morgan Kaufmann. doi: 10.1016/B978-0-12-814761-0.00002-2

Liliane, T. N., & Charles, M. S. (2020). Factors Affecting Yield of Crops. In *Agronomy—Climate Change & Food Security*. IntechOpen. doi: 10.5772/intechopen.90672

Losapio, G., Cerabolini, B. E. L., Maffioletti, C., Tampucci, D., Gobbi, M., & Caccianiga, M. (2021). The Consequences of Glacier Retreat Are Uneven Between Plant Species. *Frontiers in Ecology and Evolution*, *8*, 520. doi: 10.3389/fevo.2020.616562

Lubbe, A., & Verpoorte, R. (2011). Cultivation of medicinal and aromatic plants for specialty industrial materials. *Industrial Crops and Products*, *34*(1), 785–801. doi: 10.1016/j.indcrop.2011.01.019

Marini, L., Klimek, S., & Battisti, A. (2011). Mitigating the impacts of the decline of traditional farming on mountain landscapes and biodiversity: A case study in the European Alps. *Environmental Science & Policy*, *14*(3), 258–267. doi: 10.1016/j.envsci.2010.12.003

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R. K., & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, *15*(1), 59–69. doi: 10.1111/j.1472-4642.2008.00491.x

Mateo, R. G., Gastón, A., Aroca-Fernández, M. J., Broennimann, O., Guisan, A., Saura, S., & García-Viñas, J. I. (2019). Hierarchical species distribution models in support of vegetation conservation at the landscape scale. *Journal of Vegetation Science*, *30*(2), 386–396. doi: 10.1111/jvs.12726

McCullagh, P., & Nelder, J. A. (2019). *Generalized Linear Models* (2nd ed.). Boca Raton: Routledge. doi: 10.1201/9780203753736

Millennium Ecosystem Assessment. (2005). *Ecosystems and Human Well-being: Synthesis.* Presented at the Washington, DC. Washington, DC: Island Press.

Mod, H. K., Scherrer, D., Luoto, M., & Guisan, A. (2016). What we use is not what we know: Environmental predictors in plant distribution models. *Journal of Vegetation Science*. Retrieved from http://dx.doi.org/10.1111/jvs.12444

Morales-Barbero, J., & Vega-Álvarez, J. (2019). Input matters matter: Bioclimatic consistency to map more reliable species distribution models. *Methods in Ecology and Evolution*, *10*(2), 212–224. doi: 10.1111/2041-210X.13124

Padulosi, S., Leaman, D., & Quek, P. (2002). Challenges and Opportunities in Enhancing the Conservation and Use of Medicinal and Aromatic Plants. *Journal of Herbs, Spices & Medicinal Plants*, *9*(4), 243–267. doi: 10.1300/J044v09n04_01

Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). Ecological Niches and Geographic Distributions (MPB-49). In *Ecological Niches and Geographic Distributions (MPB-49)*. Princeton University Press. doi: 10.1515/9781400840670

Pfeifer, C., Jongeneel, R. A., Sonneveld, M. P. W., & Stoorvogel, J. J. (2009). Landscape properties as drivers for farm diversification: A Dutch case study. *Land Use Policy*, *26*(4), 1106–1115. doi: 10.1016/j.landusepol.2009.01.007

Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications: A Publication of the Ecological Society of America*, *19*(1), 181–197. doi: 10.1890/07-2153.1

Pielou, E. C. (1984). *The Interpretation of Ecological Data: A Primer on Classification and Ordination*. New York, NY: J. Wiley and Sons.

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *SOIL*, *7*(1), 217–240. doi: 10.5194/soil-7-217-2021

Price, M. F. (Ed.). (1999). *Global Change in the Mountains* (1st edition). New York: CRC Press.

Raunkiaer, C., Egerton, F. N., & Fausboll. (2005). *Life forms of plants and statistical plant geography*. London: Oxford University Press.

Ridgeway, G. (1999). The State of Boosting. *Computing Science and Statistics*, *31*, 172–181.

Robinson, O. J., Ruiz-Gutierrez, V., & Fink, D. (2018). *Correcting for bias in distribution modelling for rare species using citizen science data*. doi: 10.1111/DDI.12698

Scherrer, D., D'Amen, M., Fernandes, R. F., Mateo, R. G., & Guisan, A. (2018). How to best threshold and validate stacked species assemblages? Community optimisation might hold the answer. *Methods in Ecology and Evolution*, *9*(10), 2155–2166. doi: 10.1111/2041-210X.13041

Scherrer, D., Mod, H. K., & Guisan, A. (2020). How to evaluate community predictions without thresholding? *Methods in Ecology and Evolution*, *11*(1), 51–63. doi: 10.1111/2041-210X.13312

Shabani, F., Kumar, L., & Ahmadi, M. (2018). Assessing accuracy methods of species distribution models: AUC, Specificity, Sensitivity and the True Skill Statistic. *Acta Scientiarum Human and Social Sciences*.

Soria-Auza, R. W., Kessler, M., Bach, K., Barajas, P., Lehnert, M., Herzog, S., & Böhner, J. (2010). Impact of the quality of climate models for modelling species occurrences in countries with poor climatic documentation: A case study from Bolivia. *Ecological Modelling*, *221*, 1221–1229. doi: 10.1016/j.ecolmodel.2010.01.004

Spehn, E., Rudmann-Maurer, K., Körner, C., & Maselli, D. (2010). *Mountain Biodiversity and Global Change*. Basel: GMBA-DIVERSITAS.

Steen, V. A., Tingley, M. W., Paton, P. W. C., & Elphick, C. S. (2021). Spatial thinning and class balancing: Key choices lead to variation in the performance of species distribution models with citizen science data. *Methods in Ecology and Evolution*, *12*(2), 216–226. doi: 10.1111/2041-210X.13525

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science (New York, N.Y.)*, *240*(4857), 1285–1293. doi: 10.1126/science.3287615

Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, *10*(12), 2020–2027. doi: 10.1111/j.1365-2486.2004.00859.x

Thuiller, W., Araújo, M., & Lavorel, S. (2004). Do we need land-cover data to model species distribution in Europe? *Journal of Biogeography*, *31*, 353–361. doi: 10.1046/j.0305-0270.2003.00991.x

Thuiller, W., Georges, D., Gueguen, M., Engler, R., & Breiner, F. (2021). biomod2: Ensemble Platform for Species Distribution Modeling (Version 3.5.1). Retrieved from https://CRAN.R-project.org/package=biomod2

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, *32*(3), 369–373. doi: https://doi.org/10.1111/j.1600-0587.2008.05742.x

Trivedi, M., Morecroft, M., Berry, P., & Dawson, T. (2008). Potential effects of climate change on plant communities in three montane nature reserves in Scotland, UK. *Biological Conservation*, *141*, 1665–1675. doi: 10.1016/j.biocon.2008.04.008

Tulloch, A. I. T., Sutcliffe, P., Naujokaitis-Lewis, I., Tingley, R., Brotons, L., Ferraz, K. M. P. M. B., … Rhodes, J. R. (2016). Conservation planners tend to ignore improved accuracy of modelled species distributions to focus on multiple threats and ecological processes. *Biological Conservation*, *199*, 157–171. doi: 10.1016/j.biocon.2016.04.023

van Mansvelt, J. D., Stobbelaar, D. J., & Hendriks, K. (1998). Comparison of landscape features in organic and conventional farming systems. *Landscape and Urban Planning*, *41*(3), 209–227. doi: 10.1016/S0169-2046(98)00060-7

van Proosdij, A. S. J., Sosef, M. S. M., Wieringa, J. J., & Raes, N. (2016). Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, *39*(6), 542–552. doi: 10.1111/ecog.01509

Vári, Á., Arany, I., Kalóczkai, Á., Kelemen, K., Papp, J., & Czúcz, B. (2020). Berries, greens, and medicinal herbs—Mapping and assessing wild plants as an ecosystem service in Transylvania (Romania). *Journal of Ethnobiology and Ethnomedicine*, *16*(1), 13. doi: 10.1186/s13002-020-0360-x

Waltari, E., Schroeder, R., McDonald, K., Anderson, R. P., & Carnaval, A. (2014). Bioclimatic variables derived from remote sensing: Assessment and application for species distribution modelling. *Methods in Ecology and Evolution*, *5*(10), 1033–1042. doi: 10.1111/2041-210X.12264

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A., & Group, N. P. S. D. W. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*(5), 763–773. doi: 10.1111/j.1472-4642.2008.00482.x

Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., … Merow, C. (2020). A standard protocol for reporting species distribution models. *Ecography,* 1261–1277. doi: 10.1111/ecog.04960

Appendices

Appendix S1. About the CMIP6 scenario SSP585.

The CMIP6 scenario SSP585 represents the upper boundary of the range of scenarios described in the literature. In contrast to its predecessor, the CMIP5 scenario RCP8.5, it now incorporates socio-economic factors. The CMIP6 scenario SSP585 assumes that the world economy is growing. However, this social and economic development is based on an intensified exploitation of fossil fuel resources and an energy-intensive lifestyle worldwide.

Appendix S2. Supporting information for the Section “3.1”.

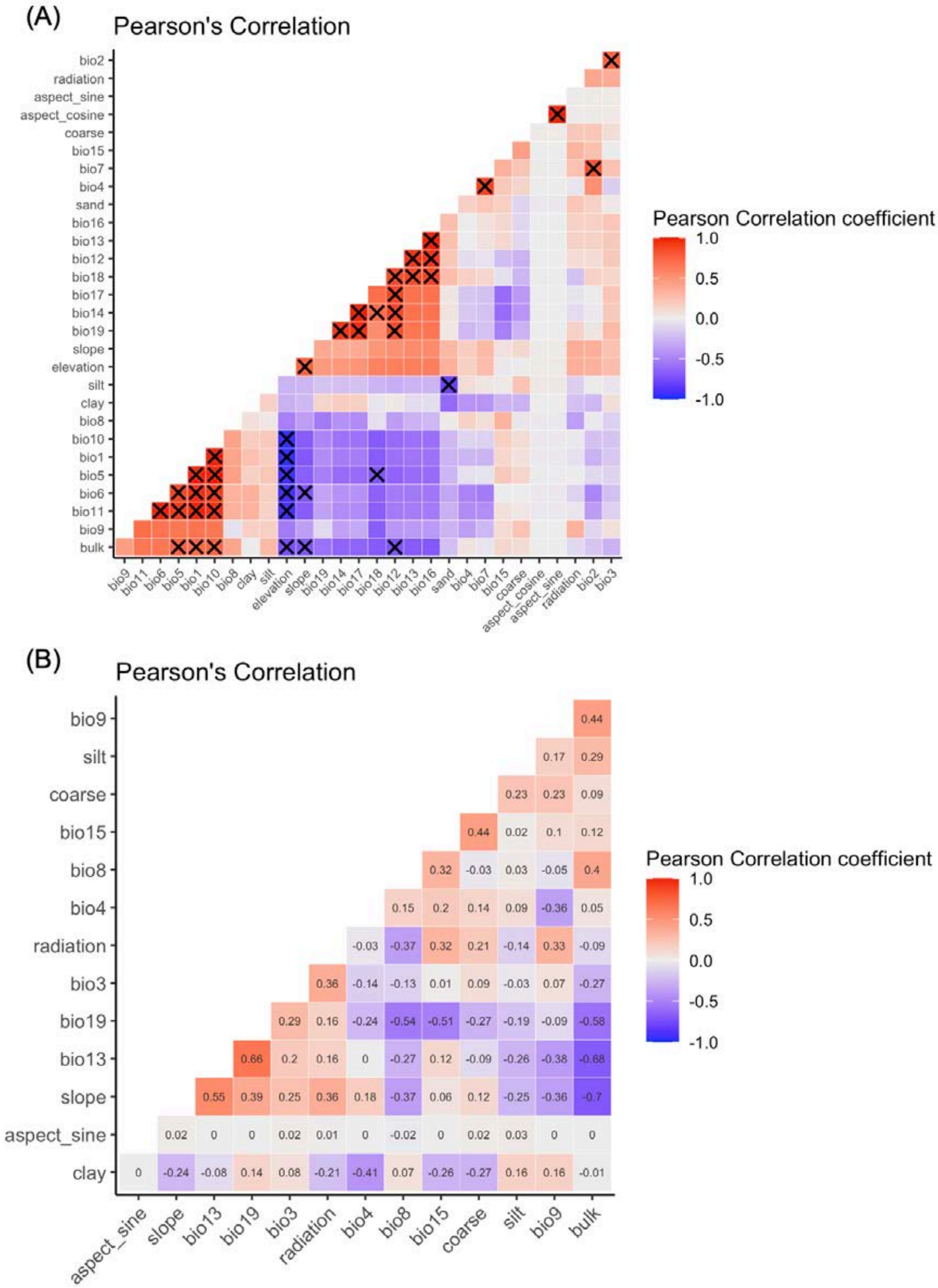
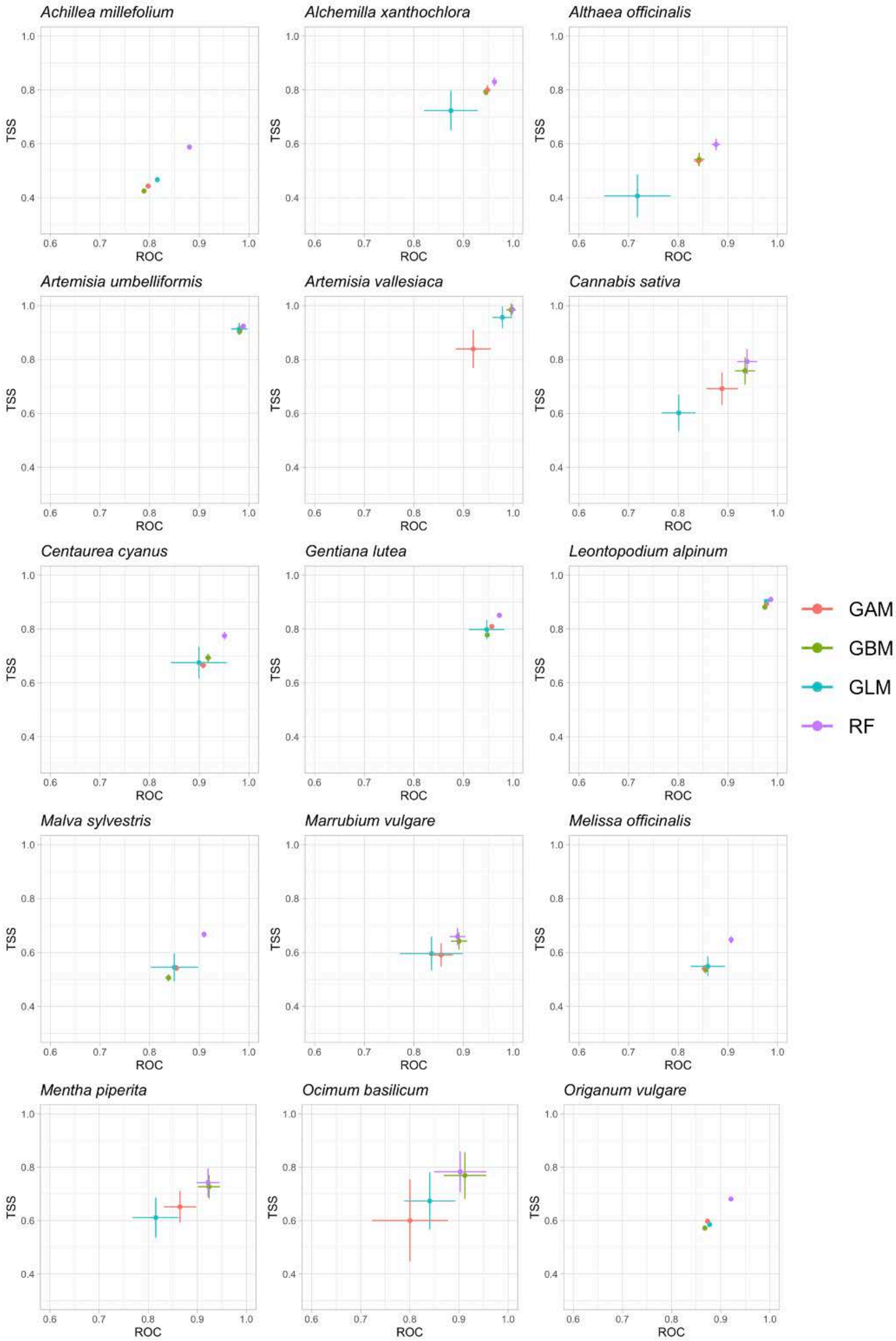


Figure S2. Correlation matrix of the Pearson correlation coefficients (A) between the original 29 variables and (B) between the 14 selected variables. In (A) crosses indicate values greater than |0.7|. In (B), the value of the correlation coefficient is written for each pair of variables.

Appendix S3. Comparison of model performance across modelling techniques and species.



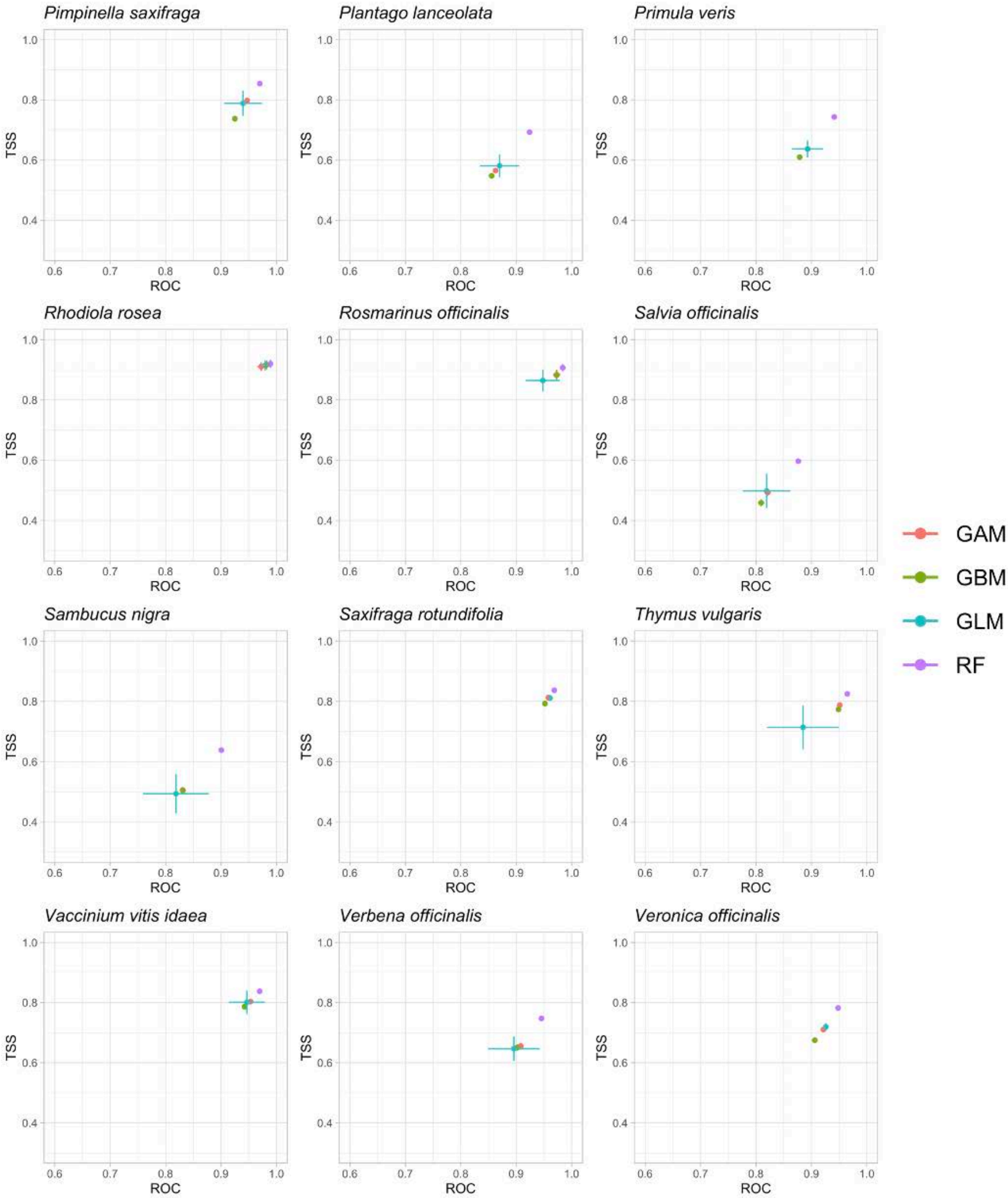


Figure S3. Visualisation of the predictive accuracy of each algorithm for each species. The points represent the mean of evaluation scores, TSS and ROC (= AUC) for a given modelling technique and lines represents associated standard deviations.

Appendix S4. Current and future maps of the predicted probabilities of the 27 studied species.

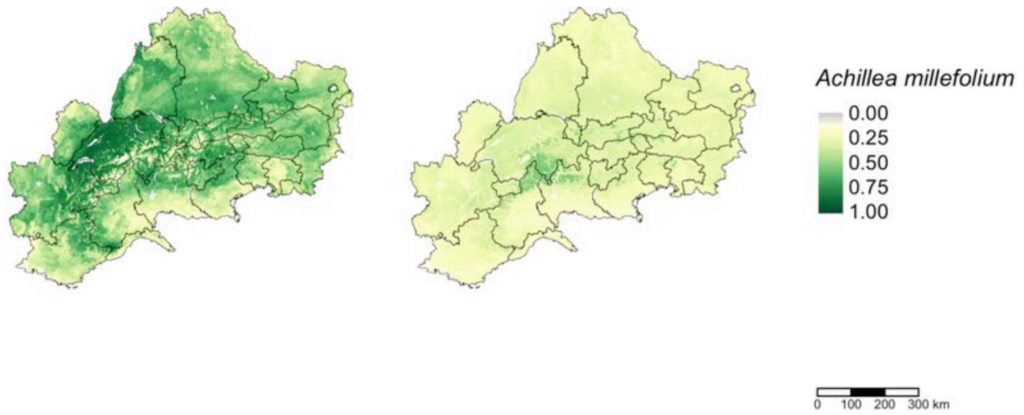


Figure S4.1. Maps of predicted probability of presence of *A. millefolium* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

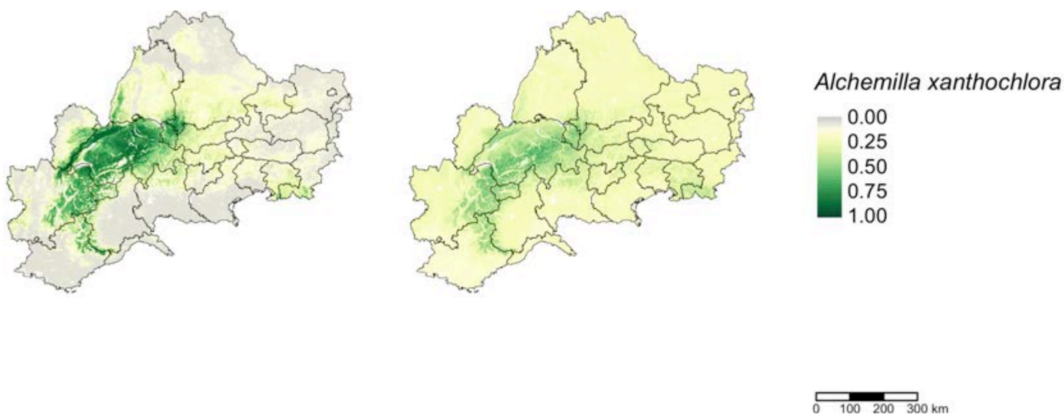


Figure S4.2. Maps of predicted probability of presence of *A. xanthochlora* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

Figure S4.3. Maps of predicted probability of presence of *A. officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

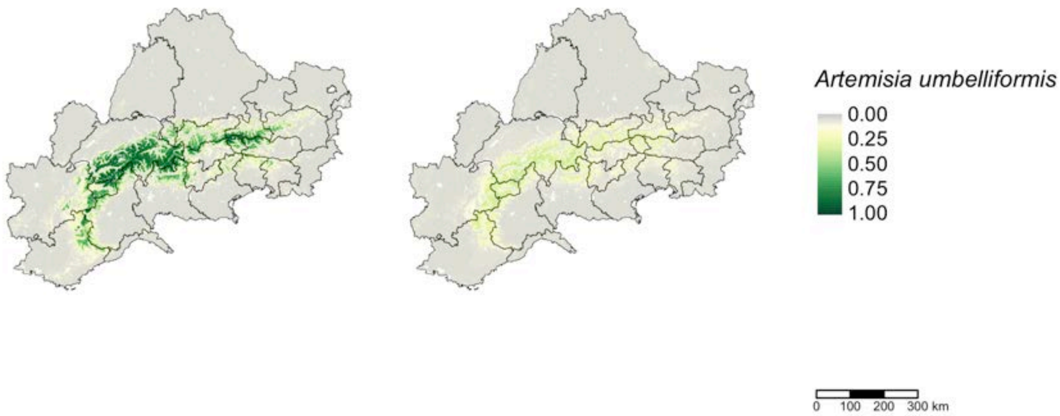


Figure S4.4. Maps of predicted probability of presence of *A. umbelliformis* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

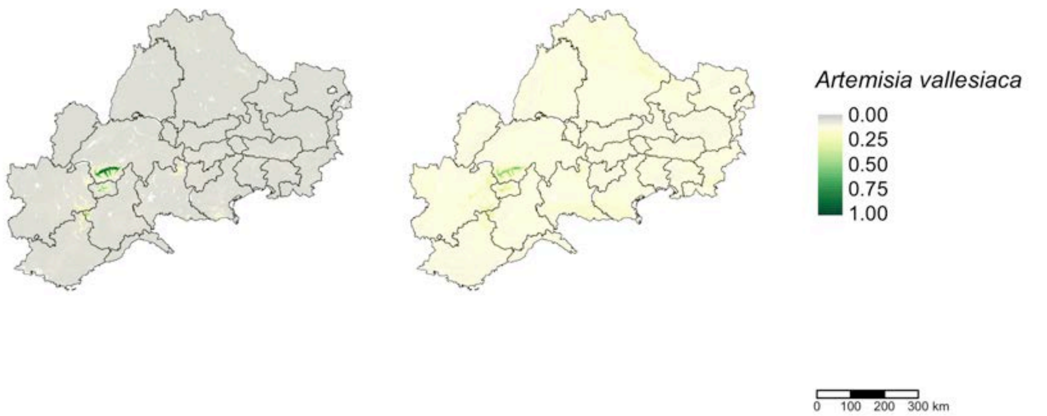


Figure S4.5. Maps of predicted probability of presence of *A. vallesiaca* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

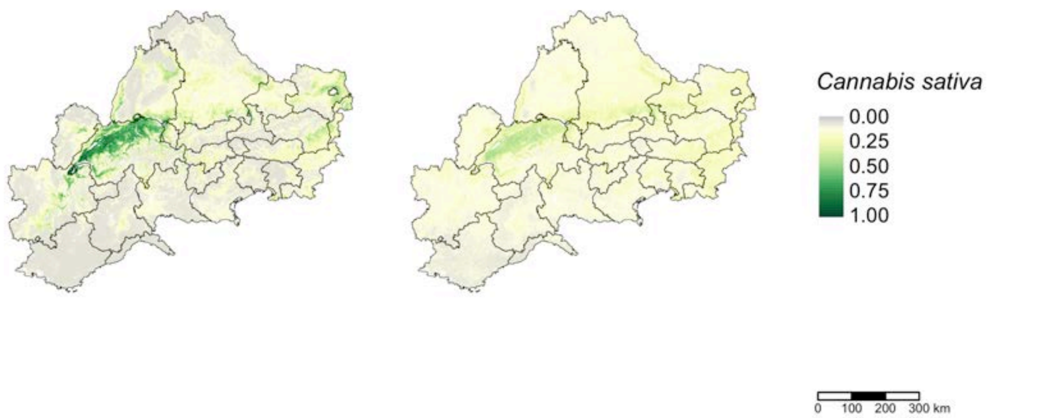


Figure S4.6. Maps of predicted probability of presence of *C. sativa* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

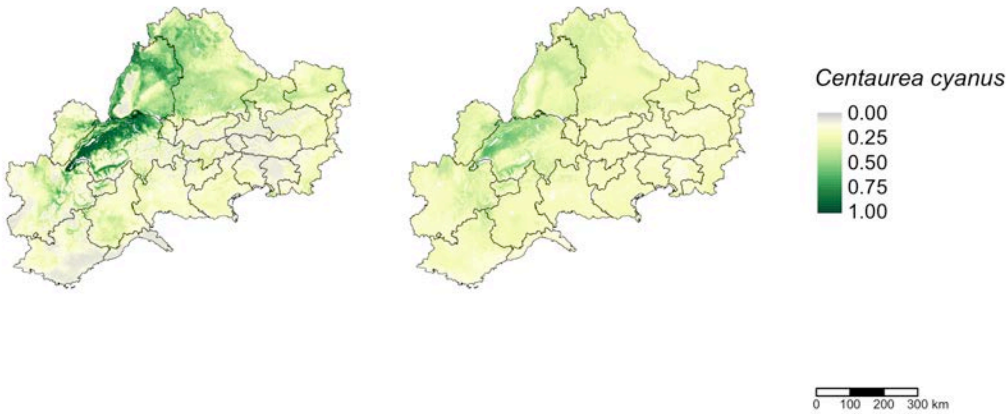


Figure S4.7. Maps of predicted probability of presence of *C. cyanus* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

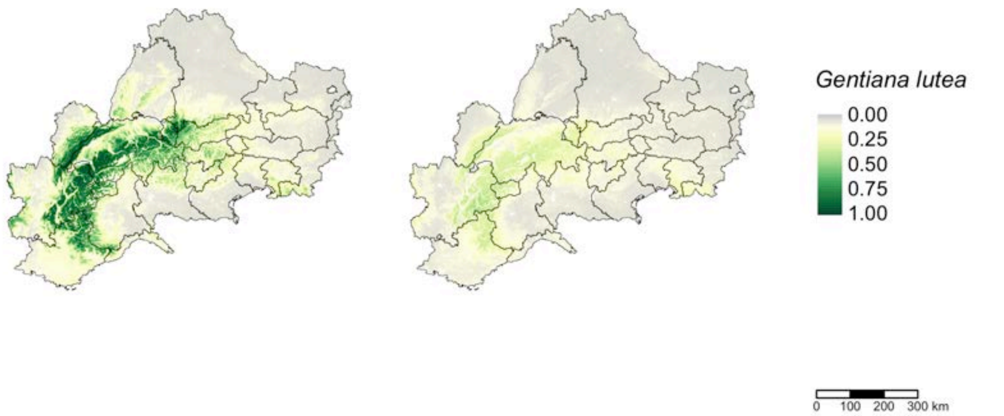


Figure S4.8. Maps of predicted probability of presence of *G. lutea* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

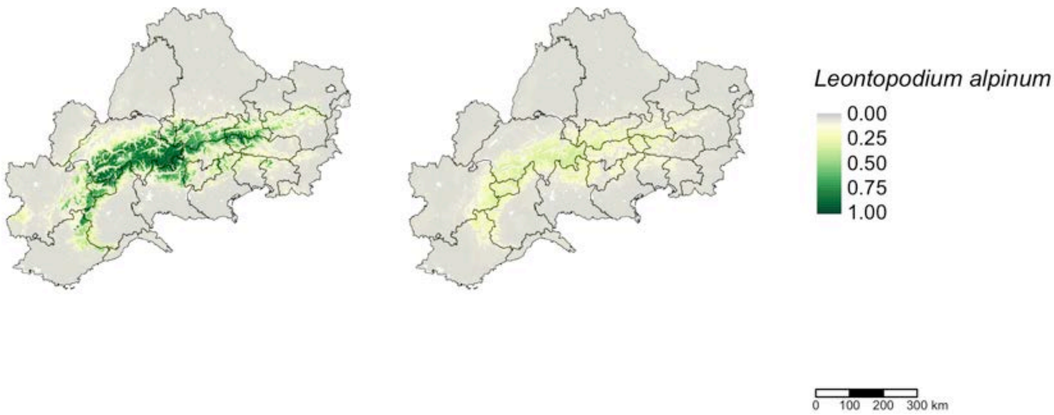


Figure S4.9. Maps of predicted probability of presence of *L. alpinum* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

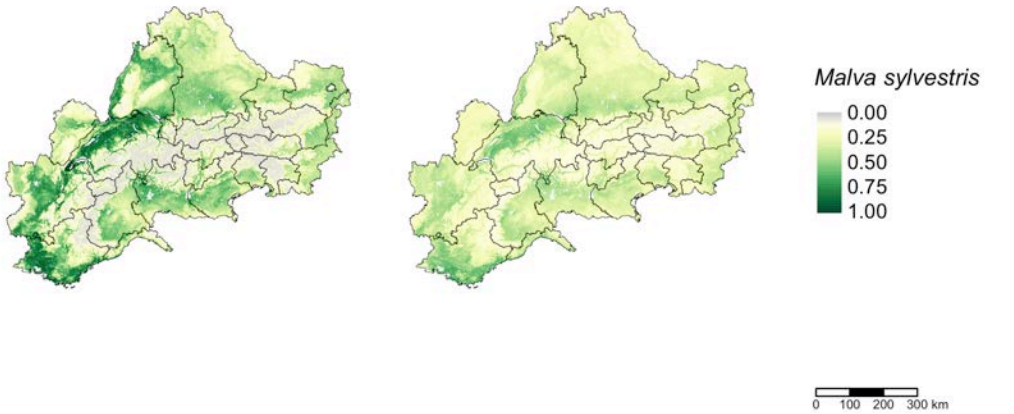


Figure S4.10. Maps of predicted probability of presence of *M. sylvestris* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

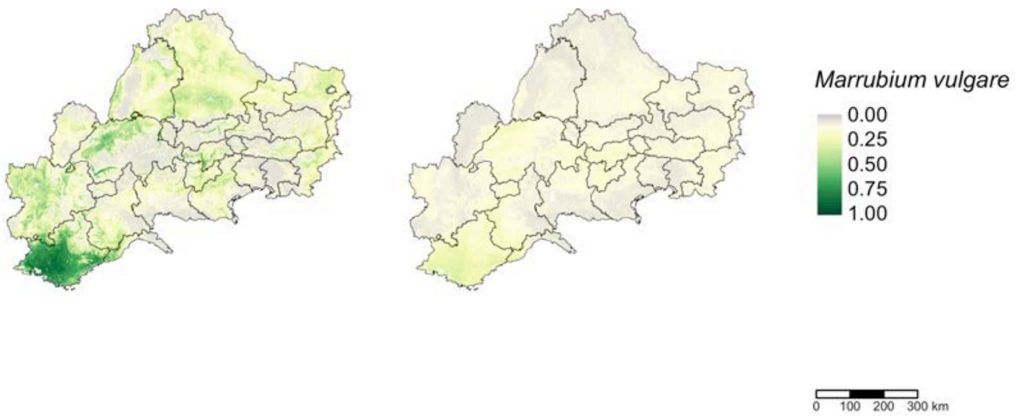


Figure S4.11. Maps of predicted probability of presence of *M. vulgare* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

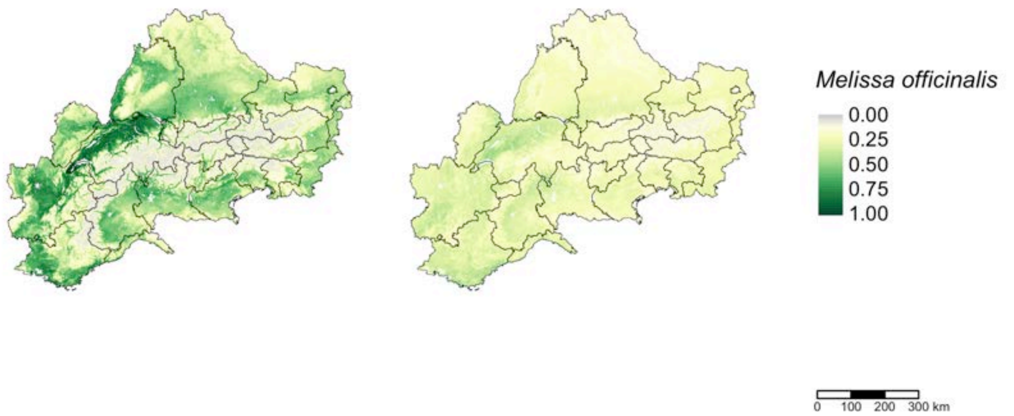


Figure S4.12. Maps of predicted probability of presence of *M. officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

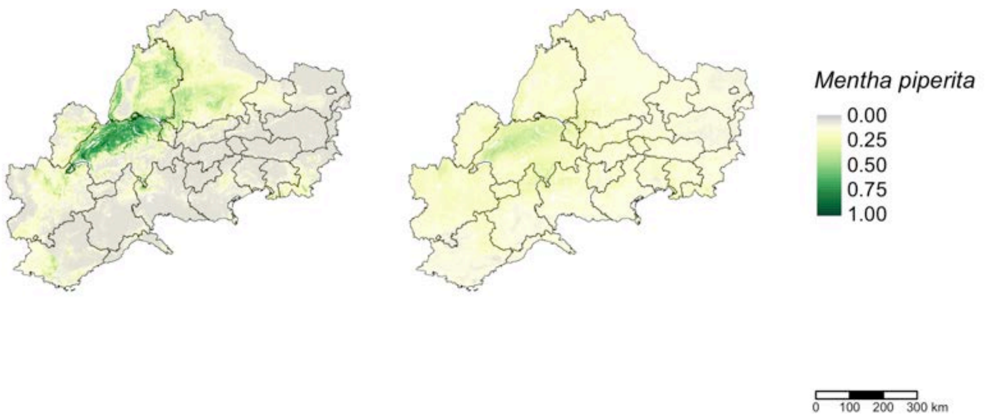


Figure S4.13. Maps of predicted probability of presence of *M. piperita* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

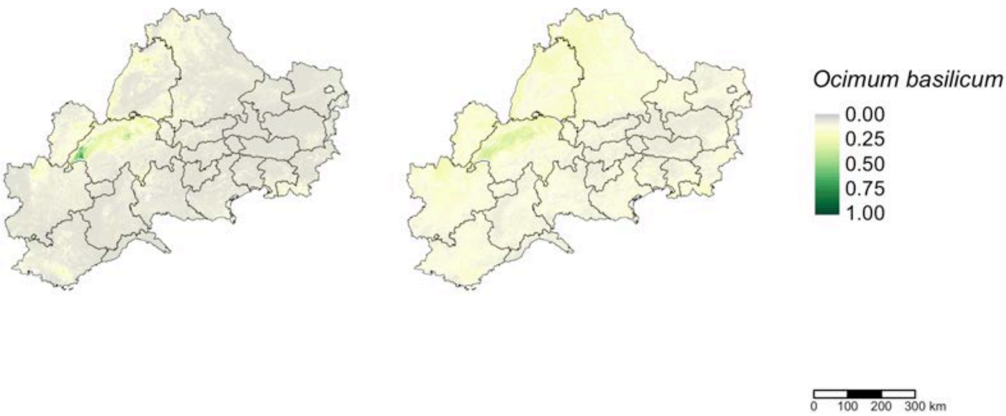


Figure S4.14. Maps of predicted probability of presence of *O. basilicum* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

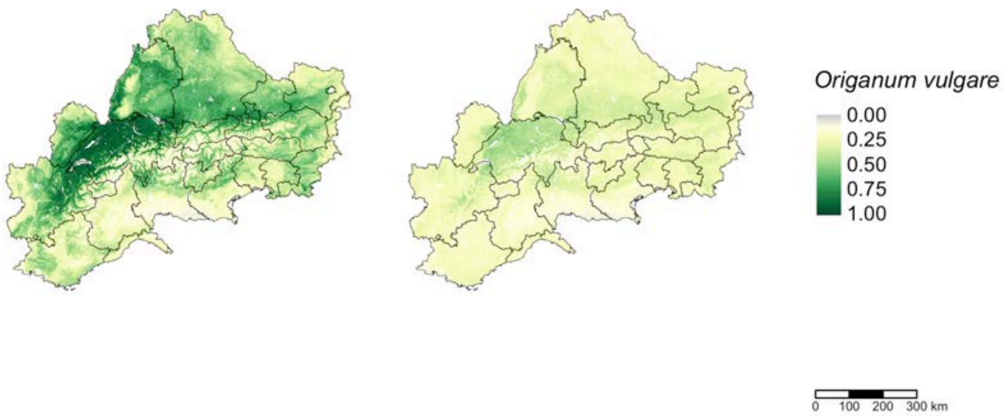


Figure S4.15. Maps of predicted probability of presence of *O. vulgare* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

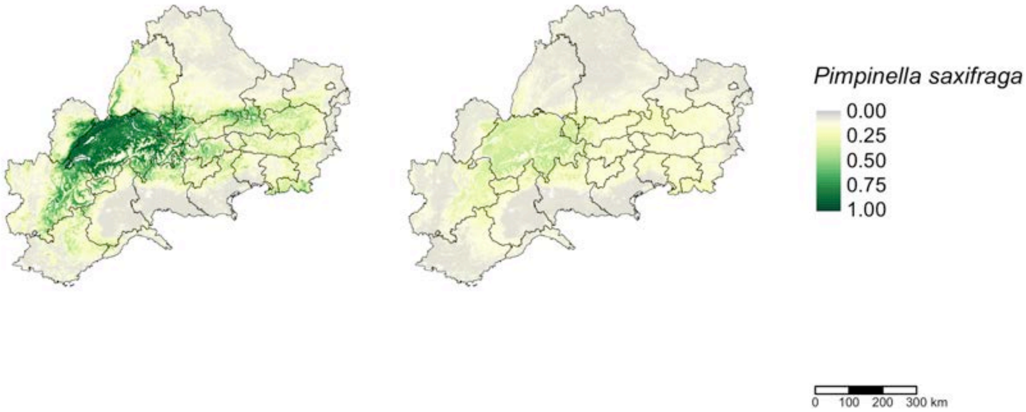


Figure S4.16. Maps of predicted probability of presence of *P. saxifraga* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

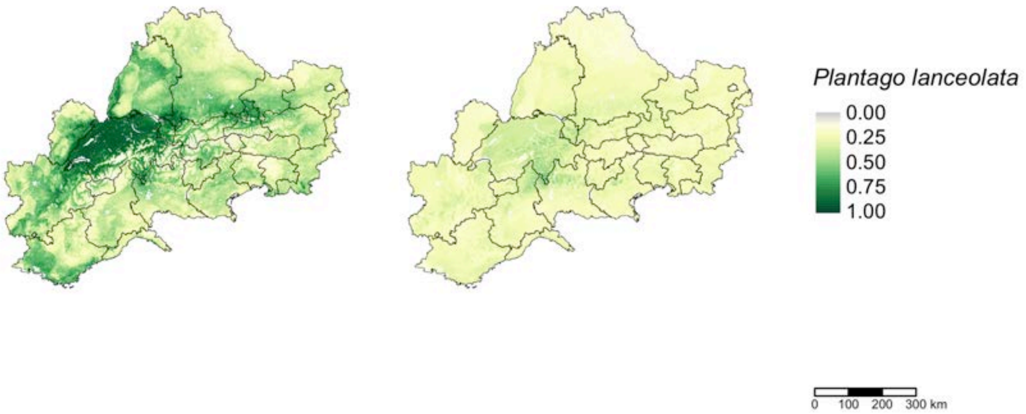


Figure S4.17. Maps of predicted probability of presence of *P. lanceolata* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

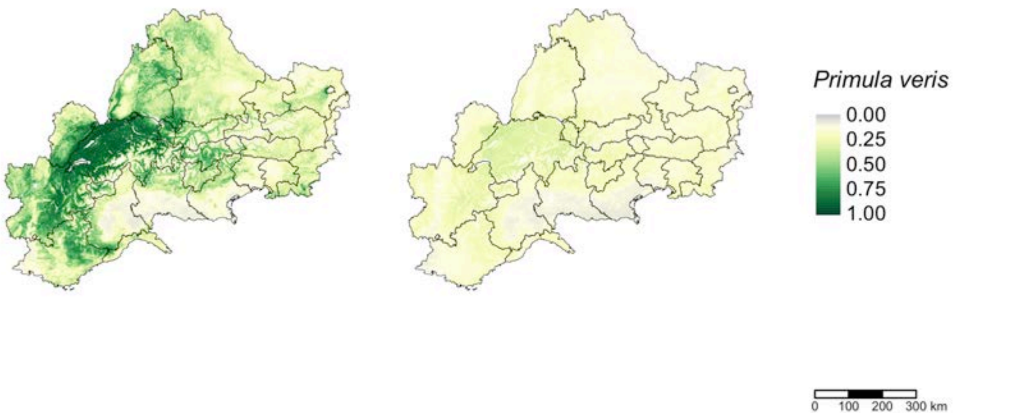


Figure S4.18. Maps of predicted probability of presence of *P. veris* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

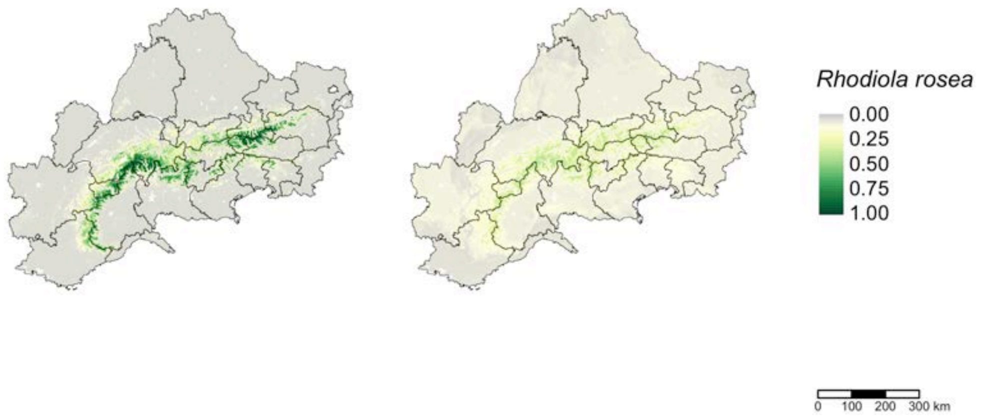


Figure S4.19. Maps of predicted probability of presence of *R. rosea* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

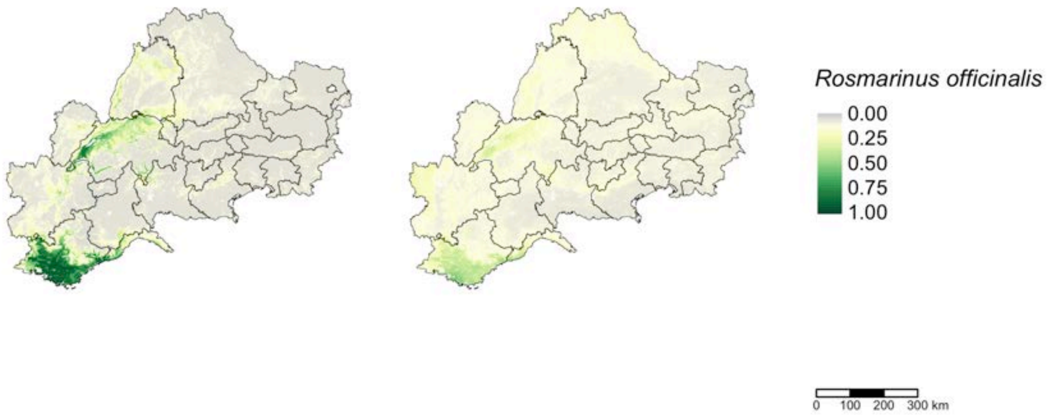


Figure S4.20. Maps of predicted probability of presence of *R. officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

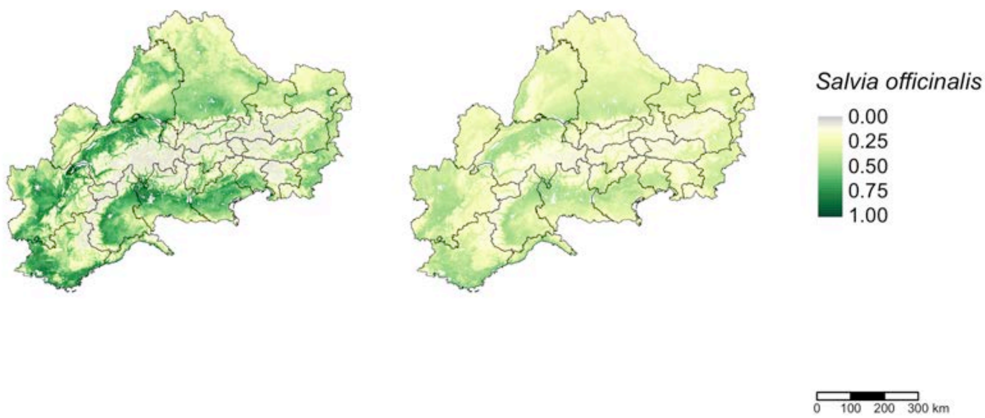


Figure S4.21. Maps of predicted probability of presence of *S. officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

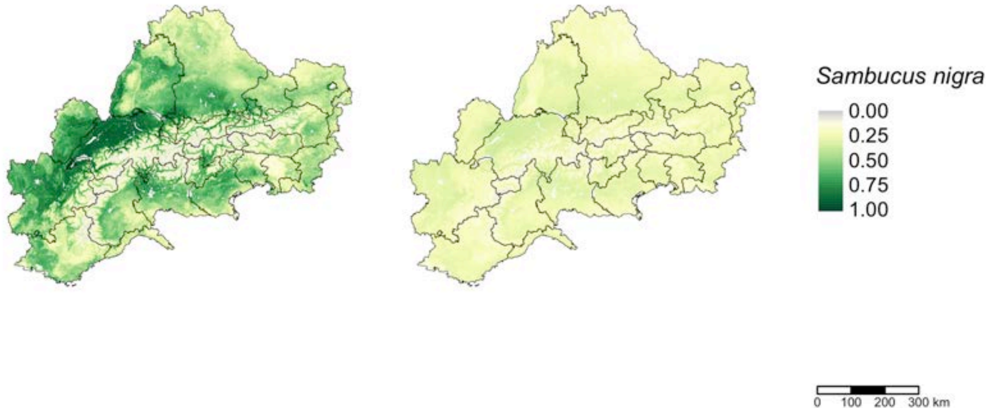


Figure S4.22. Maps of predicted probability of presence of *S. nigra* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

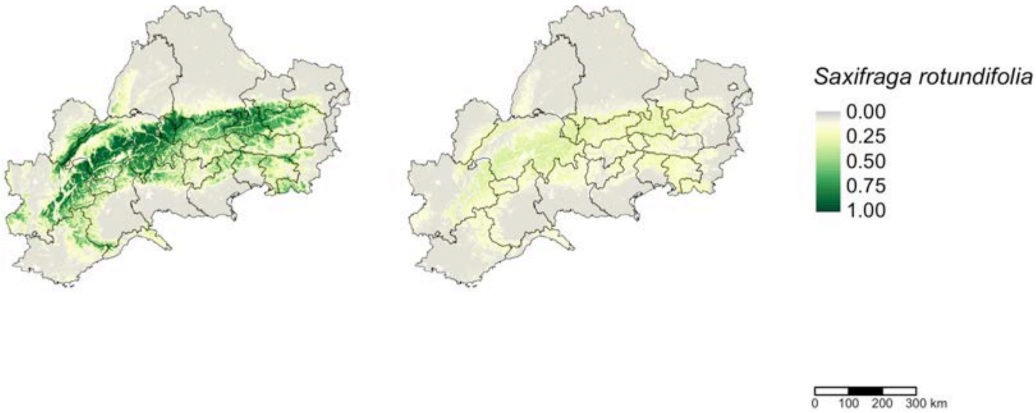


Figure S4.23. Maps of predicted probability of presence of *S. rotundifolia* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

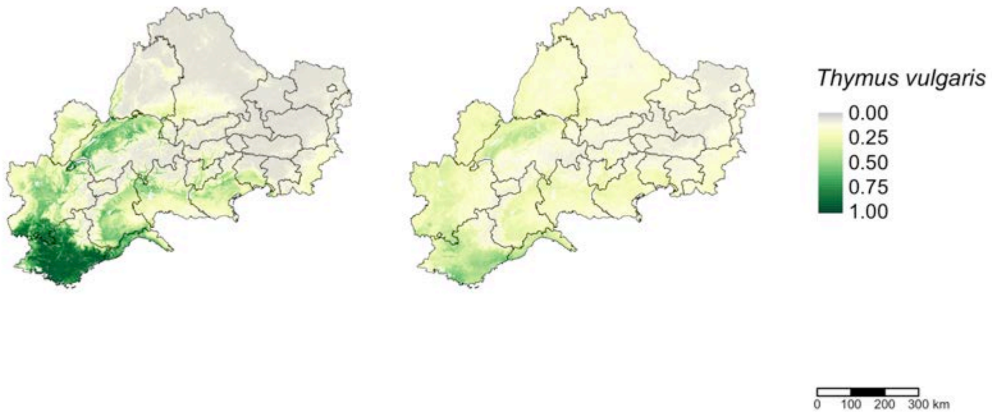


Figure S4.24. Maps of predicted probability of presence of *T. vulgaris* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

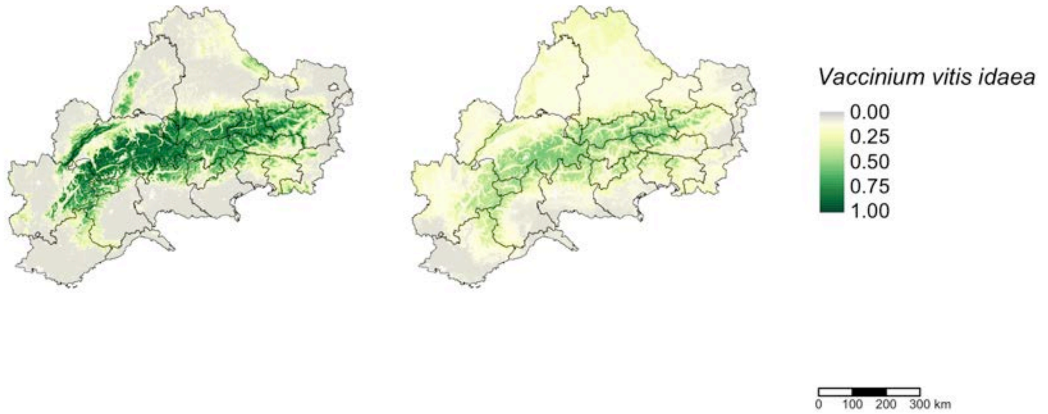


Figure S4.25. Maps of predicted probability of presence of *V. vitis idaea* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

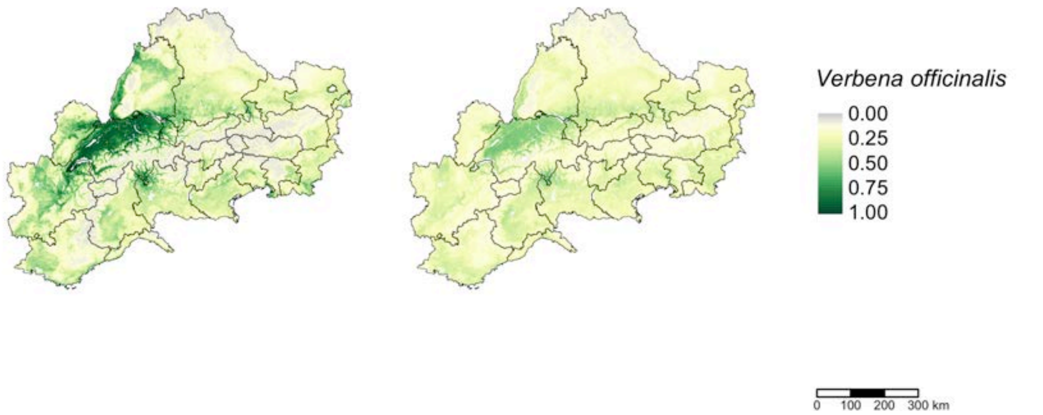


Figure S4.26. Maps of predicted probability of presence of *Verbena officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

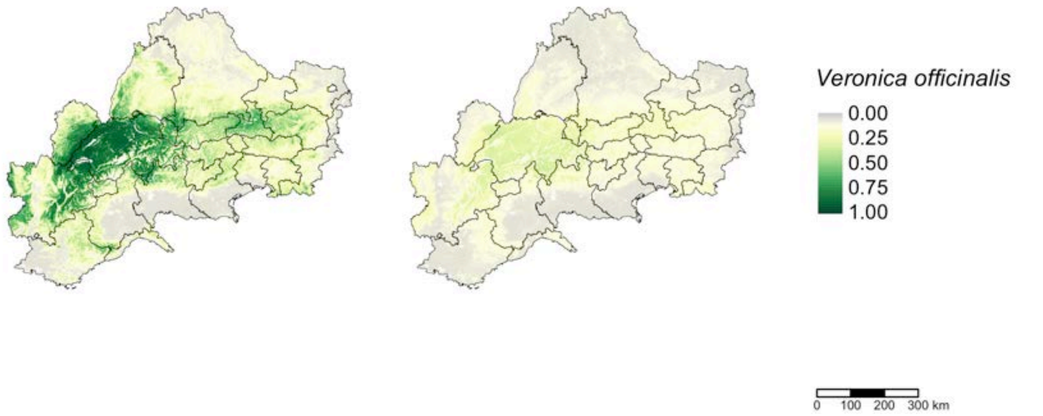


Figure S4.27. Maps of predicted probability of presence of *Veronica officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Probability is range from 0 to 1, 1 being the highest probability.

Appendix S5. Supporting information for the Section “3.3”.

Table S5. Classification of species according to expected changes in probabilities of occurrence in the future

Expected change	Species
Probability of occurrence reduced overall In the future, these species are predicted to be present in the same places as now but with their probability of presence reduced.	*Achillea millefolium* L. *Artemisia umbelliformis* Lam. *Gentiana lutea* L. *Leontopodium alpinum* Cass. *Marrubium vulgare* L. *Origanum vulgare* L. *Pimpinella saxifraga* L. *Plantago lanceolata* L. *Primula veris* L. *Saxifraga rotundifolia* L. *Veronica officinalis* L.
Colonisation of other sites In the future, these species are also predicted to be present in the same places as now, with their probability of presence reduced. Moreover, they should invade other sites with a low (but existing) probability of presence.	*Alchemilla xanthochlora* Rothm. *Althaea officinalis* L. *Artemisia vallesiaca* All. *Cannabis sativa* L. *Centaurea cyanus* L. *Malva sylvestris* L. *Melissa officinalis* L. *Mentha* ×*piperita* L. *Ocimum basilicum* L. *Rhodiola rosea* L. *Rosmarinus officinalis* L. *Salvia officinalis* L. *Sambucus nigra* L. *Thymus vulgaris* L. *Vaccinium vitis-idea* L. *Verbena officinalis* L.

Appendix S6. Current and future maps of the ranking of occurrence probabilities

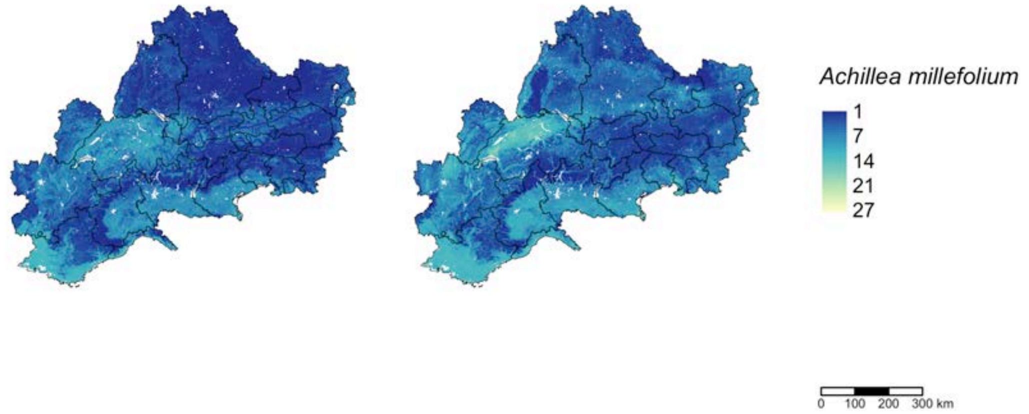


Figure S6.1. Maps of ranking of occurrence probabilities of *A. millefolium* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

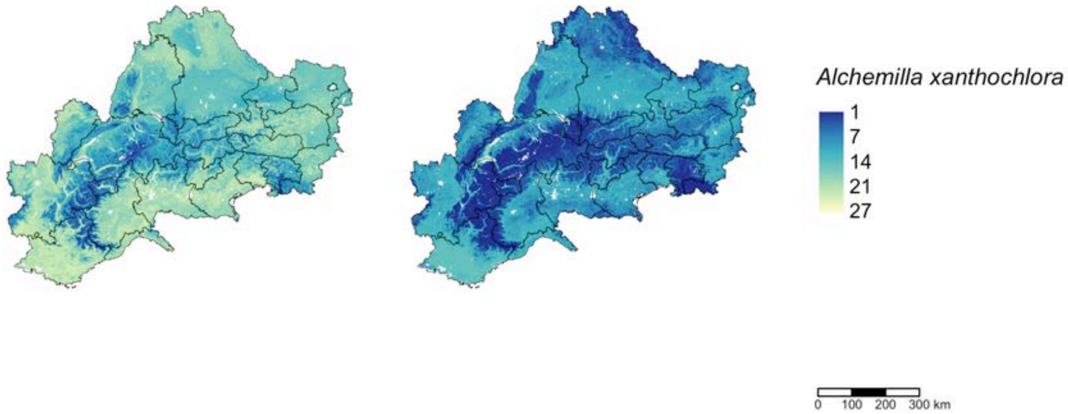


Figure S6.2. Maps of ranking of occurrence probabilities of *A. xanthochlora* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

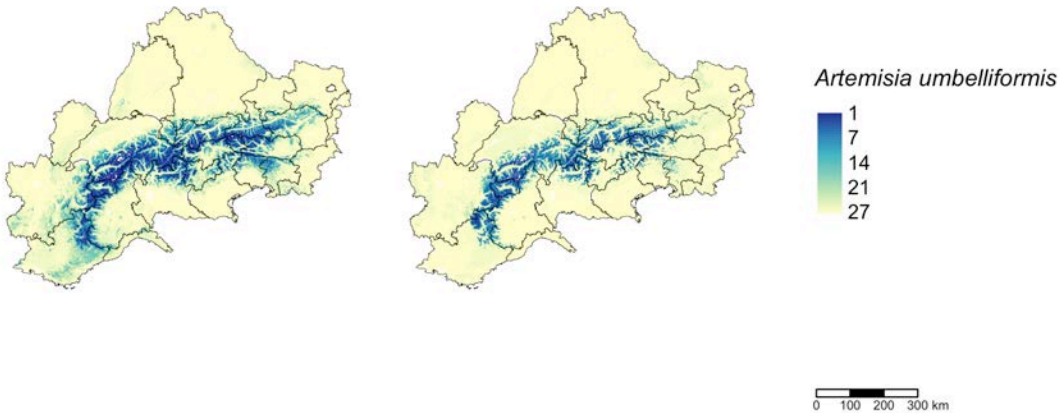


Figure S6.3. Maps of ranking of occurrence probabilities of *A. umbelliformis* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

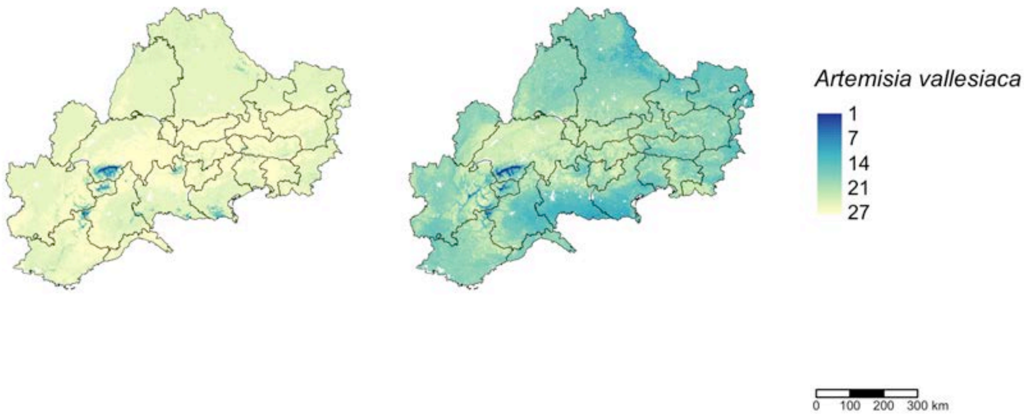


Figure S6.4. Maps of ranking of occurrence probabilities of *A. vallesiaca* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

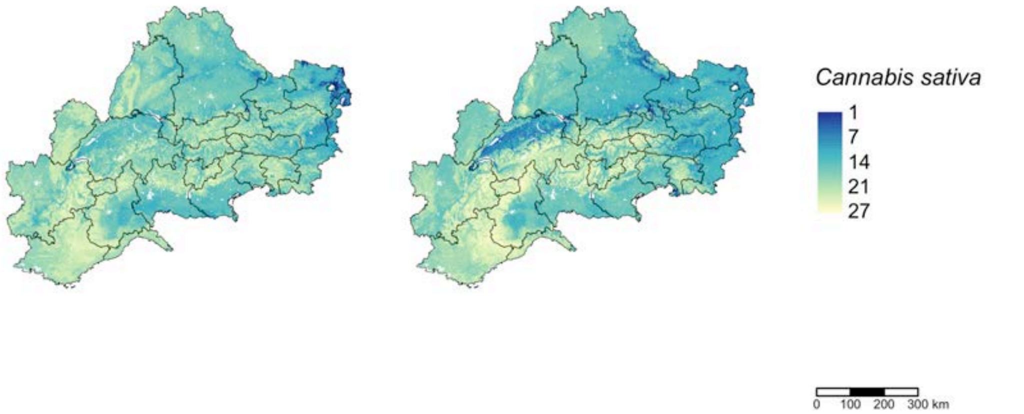


Figure S6.5. Maps of ranking of occurrence probabilities of *C. sativa* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

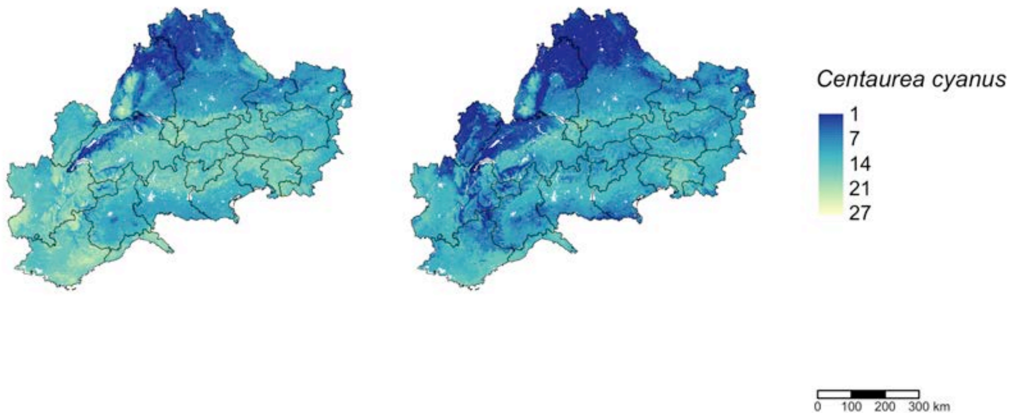


Figure S6.6. Maps of ranking of occurrence probabilities of *C. cyanus* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

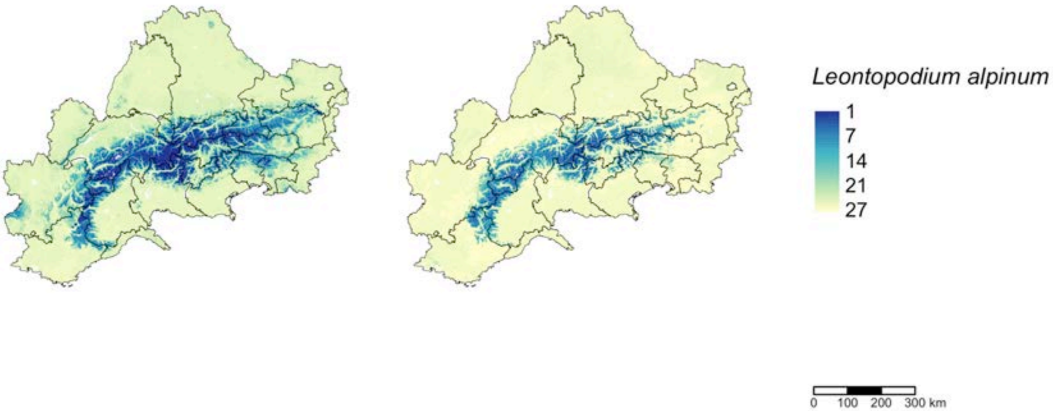


Figure S6.7. Maps of ranking of occurrence probabilities of *L. alpinum* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

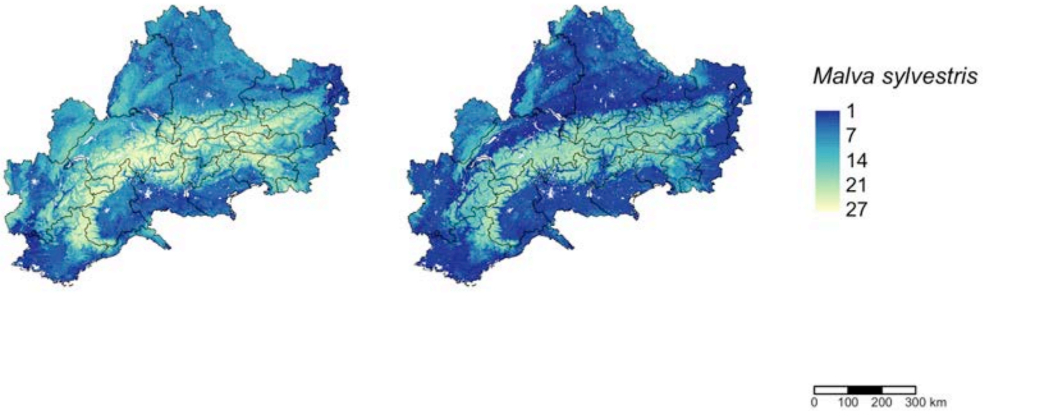


Figure S6.8. Maps of ranking of occurrence probabilities of *M. sylvestris* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

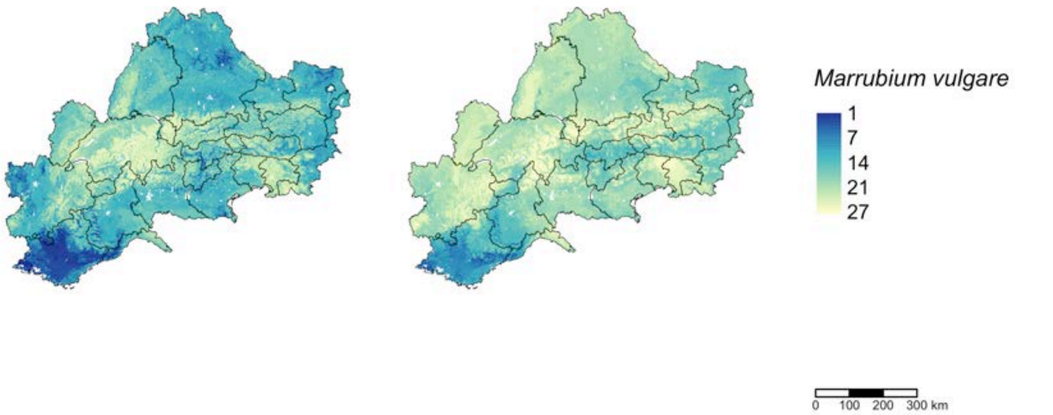


Figure S6.9. Maps of ranking of occurrence probabilities of *M. vulgare* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

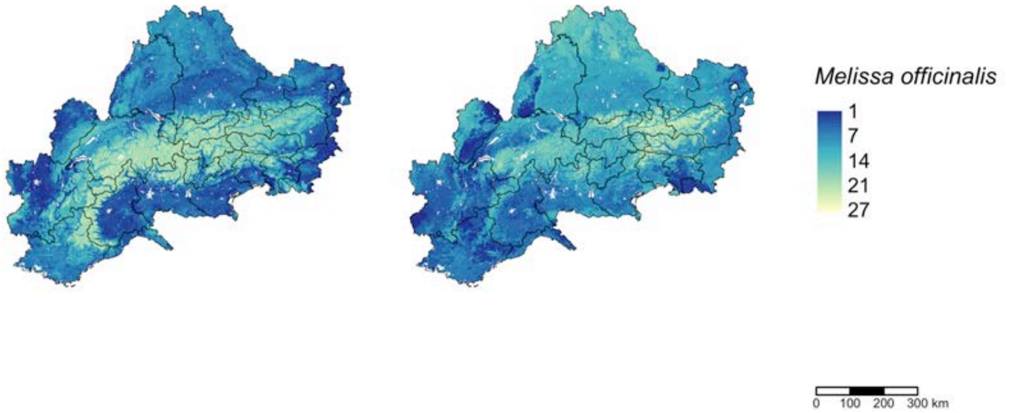


Figure S6.10. Maps of ranking of occurrence probabilities of *M. officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

Figure S6.11. Maps of ranking of occurrence probabilities of *M. piperita* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

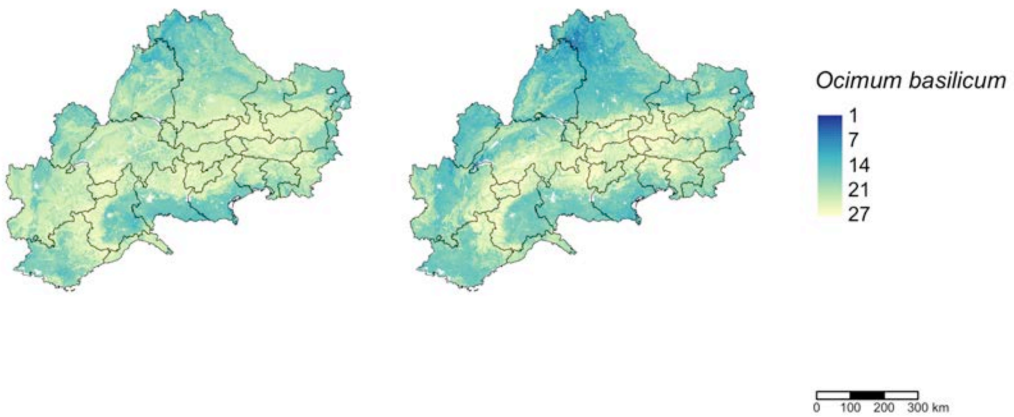


Figure S6.12. Maps of ranking of occurrence probabilities of *O. basilicum* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

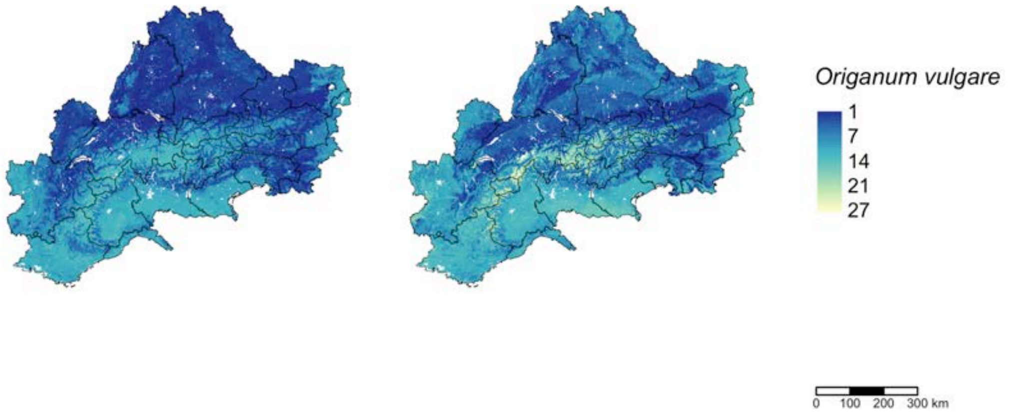


Figure S6.13. Maps of ranking of occurrence probabilities of *O. vulgare* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

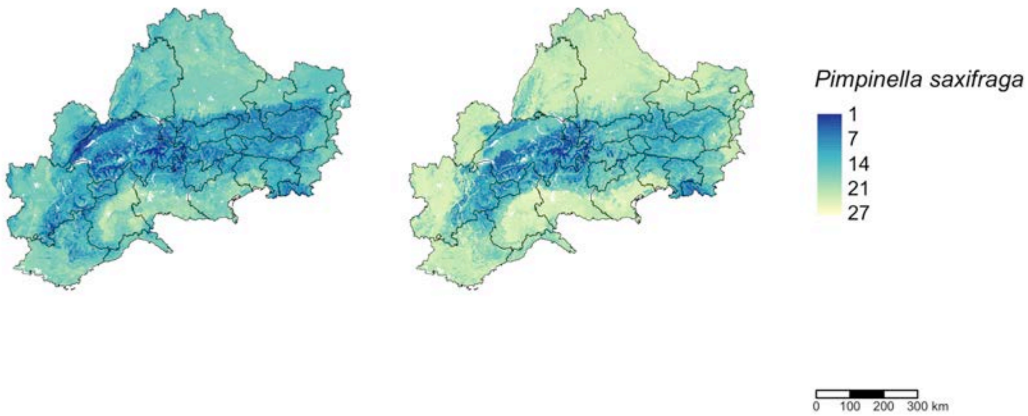


Figure S6.14. Maps of ranking of occurrence probabilities of *P. saxifraga* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

Figure S6.15. Maps of ranking of occurrence probabilities of *P. lanceolata* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

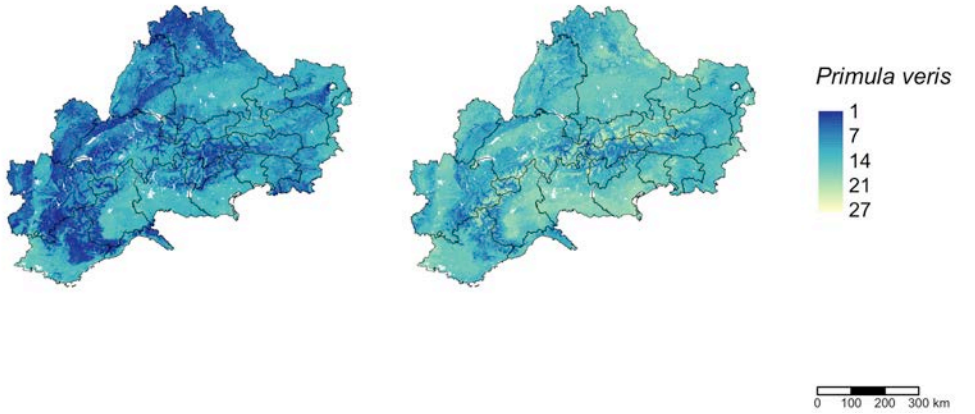


Figure S6.16. Maps of ranking of occurrence probabilities of *P. veris* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

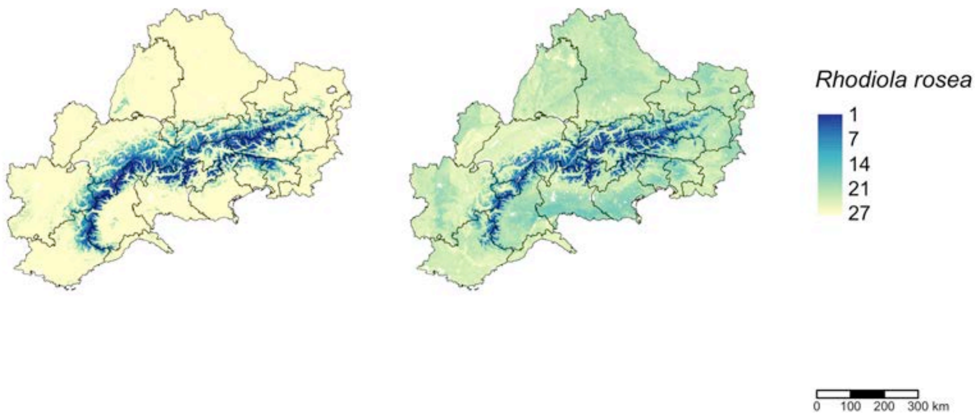


Figure S6.17. Maps of ranking of occurrence probabilities of *R. rosea* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

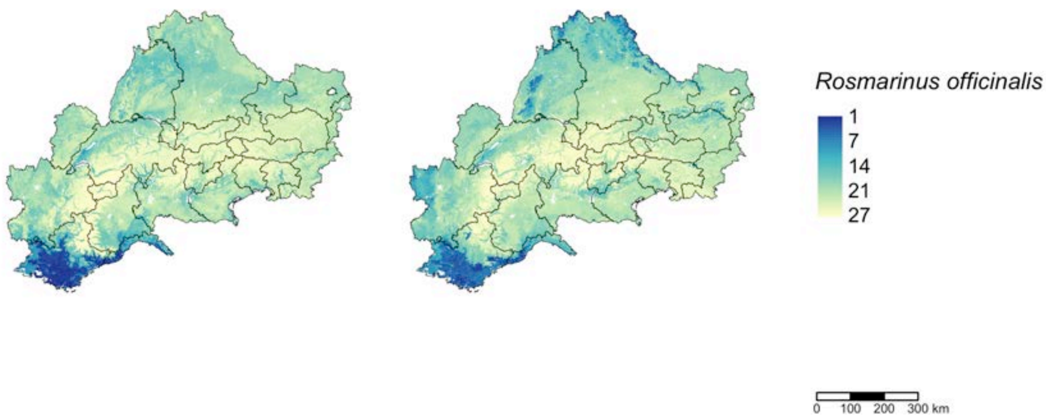


Figure S6.18. Maps of ranking of occurrence probabilities of *R. officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

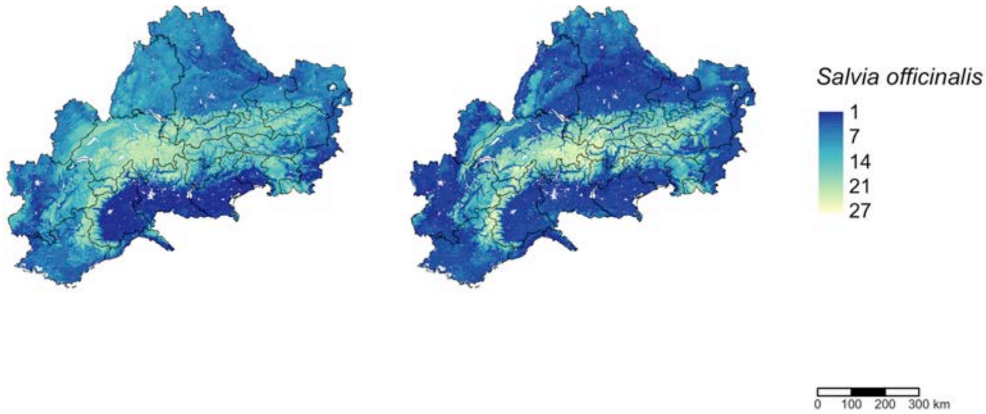


Figure S6.19. Maps of ranking of occurrence probabilities of *S. officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

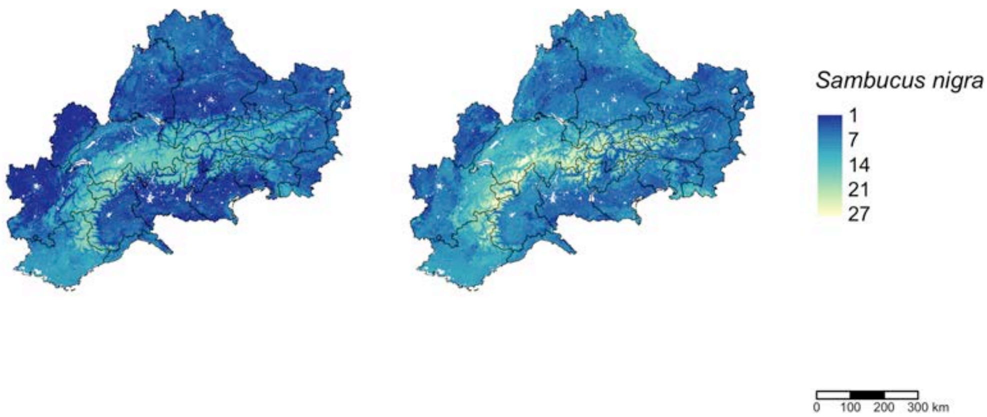


Figure S6.20. Maps of ranking of occurrence probabilities of *S. nigra* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

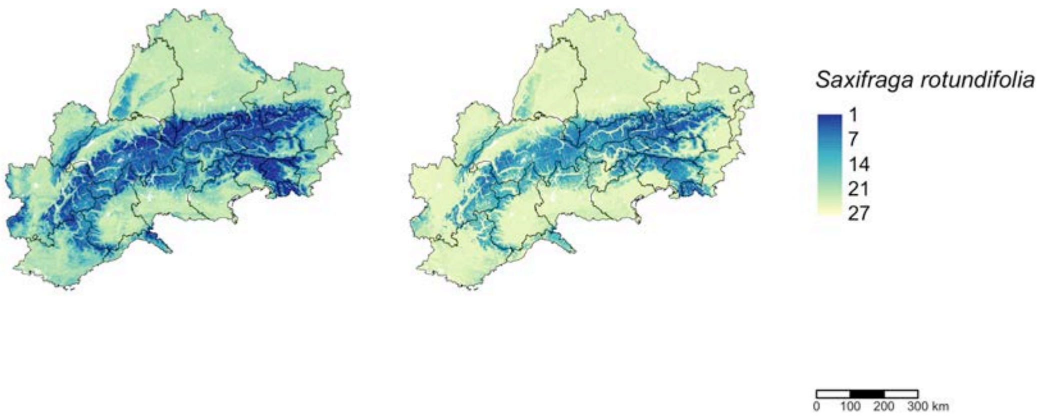


Figure S6.21. Maps of ranking of occurrence probabilities of *S. rotundifolia* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

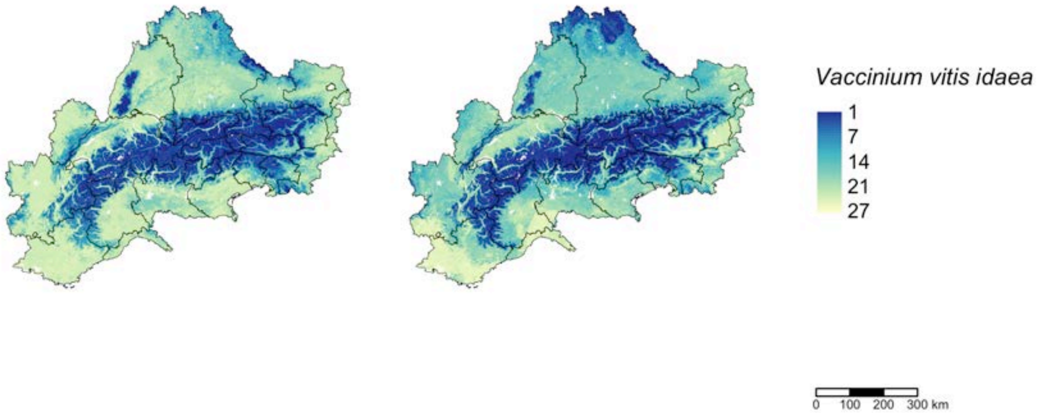


Figure S6.22. Maps of ranking of occurrence probabilities of *V. vitis idaea* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

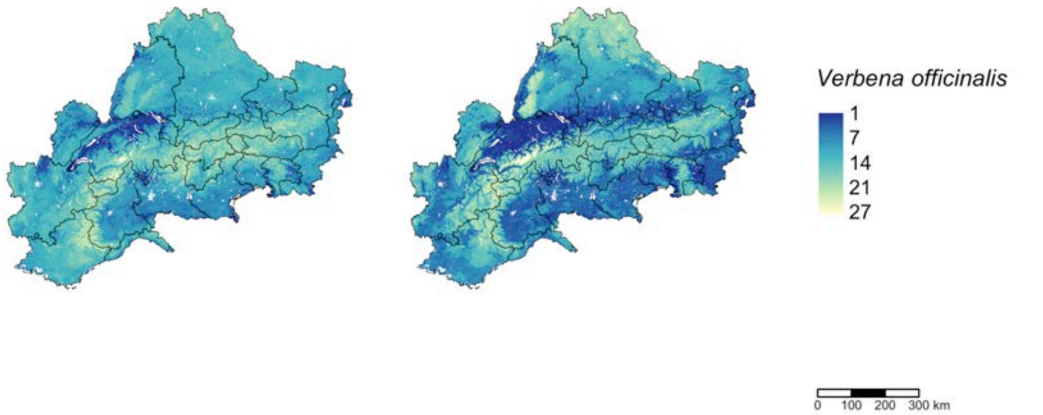


Figure S6.23. Maps of ranking of occurrence probabilities of *Verbena officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

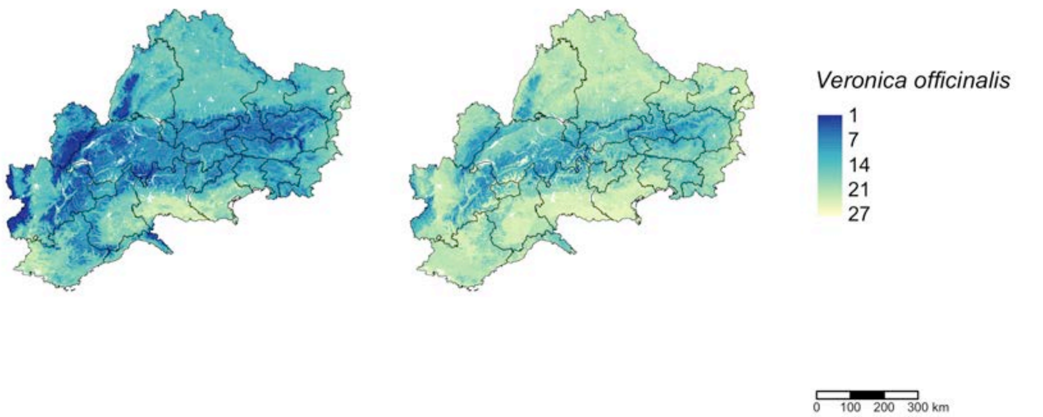


Figure S6.24. Maps of ranking of occurrence probabilities of *Veronica officinalis* in the EUSALP area, in the present (left) and by 2040 (right). Ranking is range from 1 to 27, 1 being the highest rank.

Appendix S6. Supporting information for the Section “3.4”.

Table S6. Classification of species according to expected changes in the ranking of probabilities in the future

Expected change	**Species**
Ranking decreases overall over the entire area	*Achillea millefolium* L. *Artemisia umbelliformis* Lam. *Gentiana lutea* L. *Leontopodium alpinum* Cass. *Marrubium vulgare* L. *Pimpinella saxifraga* L. *Primula veris* L. *Sambucus nigra* L. *Saxifraga rotundifolia* L. *Veronica officinalis* L.
Ranking increases overall over the entire area	*Alchemilla xanthochlora* Rothm. *Althaea officinalis* L. *Artemisia vallesiaca* All. *Centaurea cyanus* L. *Ocimum basilicum* L. *Rhodiola rosea* L. *Vaccinium vitis-idea* L.
Ranking increases or decreases according to the locations of the area	*Cannabis sativa* L. *Malva sylvestris* L. *Melissa officinalis* L. *Mentha* $^{\times}$*piperita* L. *Origanum vulgare* L. *Plantago lanceolata* L. *Rosmarinus officinalis* L. *Salvia officinalis* L. *Thymus vulgaris* L. *Verbena officinalis* L.