Environmental Modelling and Software 169 (2023) 105813

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft

ResNatSeed: An R package and shiny web app to predict the REStoration potential of NATive SEEDs using topographic factors

Davide Barberis [a,b], Marco Pittarello [c,*], Giampiero Lombardi [a,1], Michele Lonati [a,1]

[a] University of Torino, Department of Agriculture, Forest and Food Sciences, Largo Paolo Braccini 2, Grugliasco, Turin, 10095, Italy
[b] Ente di gestione Delle Aree Protette Delle Alpi Marittime, Piazza Regina Elena 30, 12010, Valdieri, Cuneo, Italy
[c] University of Torino, Department of Veterinary Sciences, Largo Paolo Braccini 2, Grugliasco, Turin, 10095, Italy

ARTICLE INFO

Keywords:
Restoration ecology
Donor grasslands
Topography
Native seeds
Suitability index

ABSTRACT

Grasslands ecological restoration relies on native seeds. The methods of assessment usually used are expert-based and focused on the similarity between the climatic and topographic factors of the two sites. The aim of the work was to develop a tool in R environment, named ResNatSeed, able to predict the suitability of a seed mixture for the receiving site, starting from some easily measurable topographic factors: elevation, slope, and aspect.

The modeling process used a training database containing vegetation and topographic information, allowing the modeling of each species abundance across the three factor gradients. A Suitability Index is then computed through the composition of a seed mixture (or that of the donor site) and on the topography of the restoration site, based on the previous models. An open access Shiny application was also set up to provide an easy-to-use tool for suitability modeling. This method can practically help practitioners during restoration programs using native seeds, reducing the subjectivity when choosing potential donor grasslands for valuable seeds.

1. Introduction

Native seeds are known to be the most suitable material for grasslands ecological restoration (Doherty et al., 2017; De Vitis et al., 2017; Mainz and Wieden, 2018) since native species have been selected by evolution to adapt to a specific environment. Moreover, commercial grass cultivars can cause genetic pollution of native populations and tend to have higher failures outside their optimal environments i.e., in extreme conditions (Kiehl et al., 2010). Indeed, Barni et al. (2007) suggests that the main reason for failures during revegetation above the timberline is the use of non-adapted commercially available seed mixtures.

Native seeds are mostly harvested from natural and semi-natural grasslands by hand collection, green hay, dry hay or by mechanical seed harvesting through vacuum, combined or brushing harvesters (Scotton et al., 2009). Different collection methods result in a different composition of harvested mixtures (Scotton and Ševčíková 2017). Among them, green hay is usually the most efficient method both in terms of number of collected species and quantity of collected seeds (Albert et al., 2019; Scotton, 2018a). The efficiency of the other methods widely varies depending on environment and species (Scotton, 2018a). Kiehl et al. (2010) found that 21–80% of the seeds produced by a grassland are harvested and transferred to the sowing sites, depending mostly on the environmental conditions and the vegetation composition of the donor grasslands. In the Alps, Barrel et al. (2015) reported values of 30–50%, while Scotton (2018a) and Scotton and Ševčíková (2017) reported values from 30 to 80%. These authors also highlighted that a fraction of the transferred seeds does not germinate readily but during the second year. This is due to the dormancy of the seeds of some species, which in temperate Europe is usually easily broken through the exposure to cold temperatures (Wagner et al., 2021). Other features influencing the harvest efficiency are seed weight and shape, which affects the ability of the different techniques to harvest the seeds (Wagner et al., 2021). By the way, the harvest of a given species is proportional to its abundance in the grassland (Scotton et al., 2009; Albert et al., 2019). The only species that cannot be harvested are the ones which have not developed seeds yet or have already shed their seeds (Scotton, 2018a; Scotton and Ševčíková, 2017). Thus, phenology is of great importance to choose the right timing and maximize the seed load and the number of species. For this reason, the number of species that can be collected is generally overestimated by donor grassland composition, although

* Corresponding author. Largo Paolo Braccini 2, Grugliasco, Turin, 10095, Italy.
E-mail addresses: davide.barberis@areeprotettealpimarittime.it (D. Barberis), marco.pittarello@unito.it (M. Pittarello), giampiero.lombardi@unito.it (G. Lombardi), michele.lonati@unito.it (M. Lonati).

[1] Lombardi Giampiero and Lonati Michele equally contributed to this work.

https://doi.org/10.1016/j.envsoft.2023.105813
Received 20 April 2023; Received in revised form 7 August 2023; Accepted 25 August 2023
Available online 26 August 2023
1364-8152/

useful to identify the possible species and their abundance. Being the best proxy of the seed mixture, the European directive 2010/60/UE – "Derogations for marketing of fodder plant seed mixtures intended for use in the preservation of the natural environment" allows reporting the grassland composition for the seed mixtures harvested directly on field.

Understanding the suitability of a seed mixture from a given donor site to a restoration site may be challenging. The methods of assessment usually used are expert-based and focused on the climatic similarity between the two sites. Little research has been done to model the seed mixtures suitability to a specific area, for instance by Doherty et al. (2017) and Shryock et al. (2022). The latter modeled the species distributions in the Mojave desert and divided it into homogeneous areas, which were then used to provide the most suitable seed mixture for a specific point in the studied area. This tool could potentially be applied to other sites but does not provide an index of suitability between a specific seed mixture and a potential restoration site. The identification of homogeneous genetic areas is commonly used also in forestry. However, analyses are always limited to single species (e.g. Falk and Mellert, 2011; Belletti et al., 2012). In the case of mixed grassland seeds or seed mixtures, instead, the approach should be more complex, using a multiple-species approach.

Doherty et al. (2017) focused on a climate similarity index to identify the possible areas where a specific seed mixture could be used according to macro-climatic conditions. However, the interest in macro-climatic similarity between sites is low, especially in Europe, which has laws and guidelines constraining the use of native seeds. The already cited EU Directive 2010/60/UE prescribes that a seed mixture must be used in the same homogeneous biogeographical area of the donor site where it was harvested. Thus, seeds can be used in areas already homogeneous at biogeographical, and thus macro-climatic, level. Consequently, differences among seed provenances should be found in micro-climate, which is mainly influenced by topography, in particular elevation, but also slope and aspect. They don't have a direct impact on plant physiology, but directly affect some variables which do. Elevation is a good proxy for temperature and rainfall, while aspect and slope are mainly proportional to the solar radiation reaching plants and soil (Körner, 2003; Austin, 2013; Boehm et al., 2021), which has several direct effects on evapotranspiration and the drought stress undergone by a plant (Austin, 2013). For these reasons, the three topographic variables mentioned above correlate well with the vegetation composition according to Austin (2013). Moreover, Boehm et al. (2021) also identified aspect, as a main factor affecting the seasonality of germination, and, consequently, the success of seedlings in case of extreme weather events or chill. Thus, topography could indirectly affect the outcome of a restoration process. As Austin (2013) pointed out, while indirect variables cannot provide explanation in terms of ecological processes, they could be useful for local predictions.

The lack of available research and modeling regarding species and habitat distribution along topographic gradients implies that the success of restoration is nowadays monitored mostly ex-post rather than at the planning stage (Gonzáles et al., 2014), and mainly through expert-based approaches. Anyway, many authors tried to find innovative solutions to this lack of available methods. For example, Jiménez-Alfaro et al. (2020) used Dexi, a Multi-Attribute Decision Making Program allowing the assessment of the suitability of different species for the restoration of agricultural habitats, by means of specific ecological and production traits.

A predictive model could help plan restoration according to the occurrence and abundance of species along ecological gradients. The distribution of species abundance along environmental gradients is a highly debated issue. Many authors support the "abundant-center hypothesis", stating that species abundance is expected to be higher at the center of their environmental niche and their performance declines outside optimal environmental conditions (Waldock et al., 2019; Dallas et al., 2020). However, other authors fully or partially reject this hypothesis (Sagarin et al., 2006; Santini et al., 2019; Dallas et al., 2020), because it is often not applicable to indirect variables like the topographic factors (Austin, 2013). While the abundant-center distribution can be easily modeled through a gaussian distribution, the modeling of abundance along indirect variables should make use of complex models that can fit the reality of the distribution independently from the shape of the relationship between abundance and the topographic variables (Austin et al., 2006). For instance, the potential natural vegetation (PNV) concept uses complex statistical models to predict the potential or original state of vegetation in a specific area, based on current vegetation distribution and environmental factors like climate, soil, and topography (Kowarik, 1987; Somodi et al., 2017).

This paper reports on the development of a tool in R environment (R Core Team, 2019) aimed to predict the suitability of a seed mixture to the receiving site, starting from some easily measurable topographic factors: elevation, slope, and aspect. The tool allows an objective assessment of such a suitability thanks to an R package, named `ResNatSeed`, and an easier to use Shiny web app. Shiny is a R interface that allows the generation of easy-to-use web applications (Chang et al., 2019), so that a user's knowledge of R is not required.

2. Package overview

2.1. Main features

`ResNatSeed` is a tool built with the R language that computes the suitability of a certain seed mixture with the conditions of the site where the seeds would be used for restoration purposes. The computation process uses the composition of the seed mixture, or the vegetation composition of the donor grassland where the mixture is harvested, and the topographic features of the restoration site (i.e. elevation, slope, and aspect). Based on statistical models, the expected abundance in the restoration site is calculated for each species, and then an index of suitability of the seed mixture (Suitability Index, SI) is predicted, as a function of elevation, slope, and aspect of the restoration site. The abundance of the species potentially occurring in the mixture is determined from a large database of vegetation surveys. Such a database can be either the default vegetation database, which is related to the Piedmont Region (NW Italy), or provided by the user to allow the usage of `ResNatSeed` for any geographical area. `ResNatSeed` is available both as a package running under R and as a Shiny app available offline and on the web either. Detailed tutorials for the use of `ResNatSeed` R package and its Shiny app are available at the package and Shiny app websites, respectively (links available at the 'Software and data availability' section).

2.2. Input files

Three input files are required to use `ResNatSeed`: a training database, the species composition of the seed mixture or of the corresponding donor grassland, and the topographic variables of the restoration site.

Training database. The training database originates from a set of located vegetation surveys and their corresponding topographic variables (elevation, slope, and aspect), named 'vegetation and topographical variables database'. `ResNatSeed` has a default training database of 4081 vegetation surveys carried out in natural and semi-natural grasslands in good conservation state in the whole alpine area of the Piedmont Region (North-Western Italy). Based on species frequency and abundance, it includes only species found in at least 50 surveys. The database includes 248 plant species to be used to define the seed mixture or the donor grassland composition. The surveys were spread along the main gradients of grasslands distribution, obtaining a highly reliable modeling for most species. Elevation ranged from 170 to 2912 m a.s.l., all the aspect range was covered, and slope ranged from 0 to 56°. In section "3.2" additional details on the surveying methodology are provided. The list of the plant species of the default training database is accessible by executing the command `data("cep.piem")`. Species names follow the Flora Alpina nomenclature (Aeschimann et al., 2004) and they are associated with the `cep.names` code, an eight-letter abbreviation of species names according to the Cornell Ecology Programs (CEP). However, to use the package in other biogeographic regions than the alpine one, users have the possibility to upload a 'vegetation and topographical variables database' with their own vegetation surveys and related topographic factors. A customized training database is generated afterwards using the `trainingDB` function by specifying a threshold of minimum frequency and minimum abundance of each species in the surveys. Species codes in CEP format are also generated at the same time. In Fig. 1 are shown, as an example, the structure of the input file 'vegetation and topographical variables database' and the output file 'training database' of the `trainingDB` function, respectively.

For reliable modeling, the custom database should include a large number of vegetation surveys distributed over sufficiently large geographic areas (e.g regional or national levels), to allow the species distribution modelling along environmental gradients. However, it is essential to consider that overly extensive areas (across different biogeographic regions or with large latitudinal and longitudinal gradients) can lead to excessive variability in ecological gradients. Therefore, the user must be able to assess the appropriate geographic extent to consider based on its purposes. Nowadays, a lot of large vegetation databases already exist for different geographical areas, e.g. VegBank for North America (http://vegbank.org/vegbank/index.jsp), European Vegetation Archive for Europe (http://euroveg.org/eva-database) and sPlot for the whole world (https://www.idiv.de/?id=176&L=0) (Wiser, 2016). These huge databases could potentially be used as input of `ResNatSeed` for specific areas of interest. However, considering the ecological aim of using native seeds for restoration, only data from natural and seminatural grasslands in a good conservation status and with a low presence of alien species should be used. If topographic variables are not available from databases, Digital Elevation Models (DEMs) are often made accessible for free by local administrations at a fine spatial grain of a few meters and, if not, 30-m grid DEMs covering the whole world are available too (e.g. Space Shuttle Radar Topography Mission, https://earthexplorer.usgs.gov and ASTER Global Digital Elevation Model, https://asterweb.jpl.nasa.gov/gdem.asp). From a DEM is possible to compute the elevation, slope, and aspect of each vegetation surveys with a Geographic Information System (GIS) software. The spatial accuracy of the DEM should be proportional to the accuracy of the vegetation data (Amatulli et al., 2018).

Seed mixture or donor grassland composition. The species composition of the seed mixture is the list of the species and their abundance in a seed batch whereas the donor grassland composition is the list of the species and their abundance surveyed in a grassland where native seeds are harvested. The list of such species should not include other species than those occurring in the training database, i.e. new species are not accepted, as SI is computed through statistical analyses using only the variables in the training database. Therefore, the more species in the seed mixture or donor grassland composition are present in the training database, the more reliable the SI calculated by `ResNatSeed` will be. With regards to the abundance of seeds in the batch, the amount of seeds per species is a good approximation of the germination potential of the different species, not considering any kind of dormancy. This approximation is possible only for areas with a low rate of dormancy, like temperate European areas, where dormancy is usually below 50% (e.g. Scotton, 2018b) and is mostly broken after the first cold season. In tropical areas, where dormancy rate is well above 50% (Kildisheva et al., 2020), this method could be highly unreliable, at least in the short-term. An additional approximation concerns the use of the donor grassland composition to estimate the seed mixture composition. The differences in the phenology of the various grassland species mean that when harvesting seeds, only the species with mature propagules will be included in the seed mixture, which will likely result in an overestimation of the number of species present in the batch. However,

vegetation and topographical variables database

Cod_ril	elevation	slope	aspect	Achillea millefolium aggr.	Anthoxanthum odoratum aggr.	Dactylis glomerata
R0001	646	2.8	74.0	20	20	72
R0002	630	3.1	87.3	8	4	72
R0003	623	4.6	81.0	8	64	16
R0004	652	0.8	121.4	24	0	4
R0005	665	2.6	160.5	12	4	12
[...]	[...]	[...]	[...]	[...]	[...]	[...]

↓

`ResNatSeed::trainingDB()`

↓

training database

Cod_ril	elevation	slope	southness	species	cep.names	abundance
R0001	646	2.8	74.0	Achillea millefolium aggr.	Achiaggr	20
R0005	665	2.6	160.5	Achillea millefolium aggr.	Achiaggr	12
R0002	630	3.1	87.3	Achillea millefolium aggr.	Achiaggr	8
R0003	623	4.6	81.0	Achillea millefolium aggr.	Achiaggr	8
R0004	652	0.8	121.4	Achillea millefolium aggr.	Achiaggr	24
R0004	652	0.8	121.4	Dactylis glomerata	Dactglom	4
R0001	646	2.8	74.0	Dactylis glomerata	Dactglom	72
R0005	665	2.6	160.5	Dactylis glomerata	Dactglom	12
R0002	630	3.1	87.3	Dactylis glomerata	Dactglom	72
R0003	623	4.6	81.0	Dactylis glomerata	Dactglom	16
[...]	[...]	[...]	[...]	[...]	[...]	[...]

Fig. 1. Process of the `trainingDB` function of `ResNatSeed` through which the input file 'vegetation and topographical variables database' is reshaped into the training database, taking into account the user-defined threshold of minimum frequency and abundance of each species.

if harvested when most of the dominant species are ripe, the abundance of the species in the grassland can represent most of the species and can be used as a proxy to calculate the potential adaptability of a seed mixture to a restoration site. In general, abundance must be a number bounded between 0 and 100. For the seed mixture it represents the proportion in weight of the seeds for each species, whereas for the donor grassland composition it can be the cover assessed visually (e.g. phytosociological surveys according to Braun-Blanquet, 1928) or the species relative abundance and cover reported with point-intercept methods (*sensu* Pittarello et al., 2016; Verdinelli et al., 2022). In case databases are characterized by botanical surveys conducted using different methodologies (e.g., point-intercept method, phytosociological surveys), it is advisable to standardize the abundance values to make them comparable. The seed mixture or donor grassland composition must be provided in a two-column database, where the first column contains CEP coded species names and the second with corresponding abundance.

Topographic variables. The topographic features of the restoration site are set from those of the potential receiving site where the restoration will occur. As for the training database, the selected variables are elevation (m a.s.l.), slope (°) and aspect (°N), three features easily extractable from the same abovementioned DEMs. We are aware that the factors influencing the distribution of plant species go well beyond these three considered (such as lithology, climate, water availability, management type, etc.), but our goal was to use a limited number of variables that are easily accessible to the user. Moreover, when the spatial scale is not extremely large, altitude is known to be correlated with vegetation composition as it is an indirect expression of other environmental variables with which it is correlated, such as temperature and rainfall regime (Austin, 2013).

2.3. Algorithm for SI computation

The core function of the `ResNatSeed` package is called `RestInd` and allows the computation of the Suitability Index (SI). The function algorithm (Fig. 2) consists in selecting the species listed in the composition of the seed mixture or of the donor grassland from the training database. Then, through a loop for each species, the following operations are carried out:

1) assignment to each survey in the training database of an altitude, slope and aspect class with a pace of 50 m, 5°, and 10°, respectively. Before such an assignment, aspect, being a circular variable, is linearized through its conversion to southness (southness = 180 − |aspect − 180|) to avoid circular variables issues (Chang et al., 2004). The same transformation is applied to the aspect of the restoration site.
2) selection, separately for each topographic variable and class, of the survey with the maximum abundance value. Then the surveys with the maximum abundance of the three topographic variables are merged excluding possible duplicates deriving from the selection from the different topographic variables (e.g. it is possible for a single case to report the maximum abundance of a species for both a slope and altitude class). Therefore, a database with unique surveys is created. The maximum abundance is useful to estimate the maximum potential of a species for the analyzed ecological range. Every survey thus represents the potential maximum abundance for that species, regardless of the abundance of other species.
3) Modeling of the species abundance in response to topographic factors through four Generalized Additive Models (GAM) fitted using the function "`gam`" of the "`mgcv`" package (Wood, 2017). GAMs are widely used for species distribution modelling as they are a non-parametric extension of Generalized Linear Models that use a data smoothing procedure and, therefore, with a great advantage that the shape of the species response curve does not have to be preliminary specified by a mathematical function (Austin, 2013). Before modeling, species abundance is converted to a 0–1 interval with the transformation proposed by Smithson and Verkuilen (2006) to allow the usage of the Beta distribution family. The four models are different according to the presence or absence of interaction among variables: i) without any interaction, ii) with interaction between elevation and southness, iii) with interaction between elevation and slope, and vi) with the interaction among all the three variables. The smoother basis set for single terms is a "thin plate regression spline" (`bs = 'tp'`), whereas the interaction between terms is set using a tensor product interaction `ti`

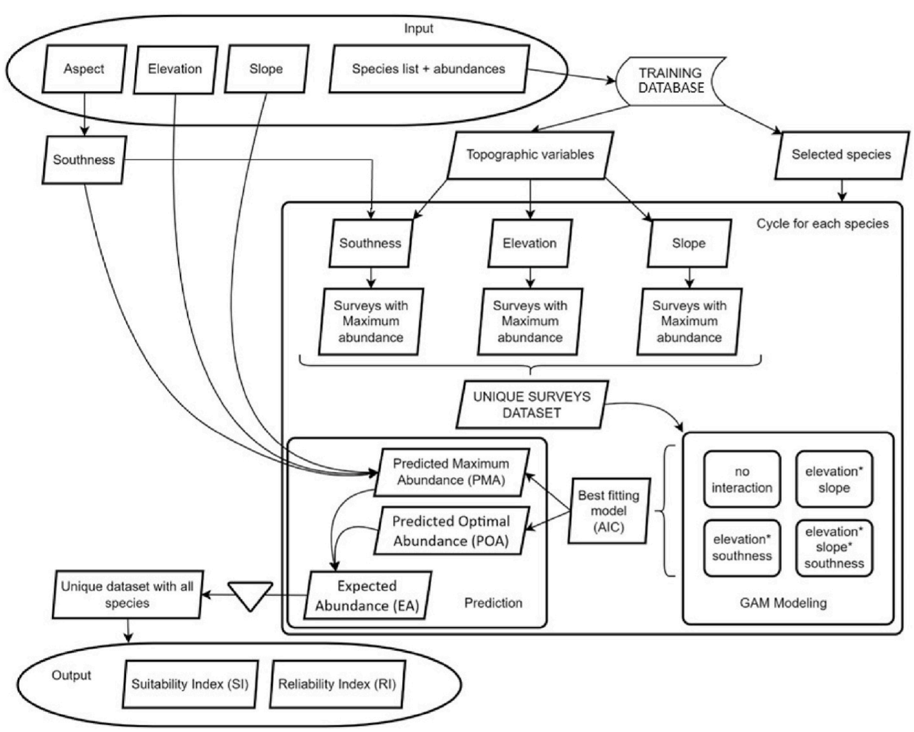

Fig. 2. Workflow of `RestInd` function algorithm of `ResNatSeed` R package.

`()`, as the main terms are also present. Different models are performed to find the best prediction of species abundance in relation to topographic variables. If no model fits, the species is discarded, otherwise the model with the lowest Akaike Information Criterion (AIC) is selected. Moreover, the Root Mean Squared Error (RMSE) and the adjusted R2 are computed to detect the goodness of model fitting. RMSE and adjusted R2 are computed with the function "`rmse`" and "`r2`", respectively, of the "`performance`" package (Lüdecke et al., 2021). In Table S1 the best models for each plant species listed in the default training database are reported, while the response curves of the modeled species across the three topographic variables are reported in Fig. S1.

4) Prediction of the maximum achievable species abundance in the restoration site by means of the best selected model and based on the values of elevation, slope and aspect set by the user. The computed value is named as 'Predicted Maximum Abundance' (PMA). The modeling process does not work if the site variables are beyond the range available in the training databae for each species to avoid extrapolation. However, to limit problems linked to this constraint, the algorithm allows to extrapolate the values if the lower limit of the slope is $< 5^\circ$ and if the lower and upper limits of southness are $<22.5^\circ$ and $>157.5^\circ$, respectively. In this way, if the input is 0° but the lower limit in the training database is for example 1°, the calculation is still performed.
5) Prediction of the 'Predicted Optimal Abundance' (POA), which is the maximum achievable abundance of a species in its optimal ecological condition, based on all possible combinations of elevation, slope and southness into the training database.
6) Computation of the ratio between the PMA and POA, which indicates how far (ratio = 0) or close (ratio = 1) a species is from its ecological optimum.
7) Computation of the 'Expected Abundance' (EA), which is the highest achievable abundance of a species in a restoration site, based on how far the species is from the ecological optimum, and considering a negligible dormancy rate. This index is computed from the multiplication of the ratio between the PMA and POA and the abundance of the species reported in the seed mixture or donor grassland composition.

Upon execution, the algorithm ultimately produces the Suitability Index (SI) and the Reliability Index (RI).

The SI represents the suitability of a seed mixture or donor grassland to restore a site with specific topographic characteristics. It is calculated dividing the sum of the predicted EAs of all modeled species by the sum of their abundances in the seed mixture or donor grassland and it ranges between 0 and 1. When SI tends to 0, the restoration site is totally beyond the optimal ecological ranges of all the species of the donor grassland or seed mixture, which is therefore not appropriate for the site restoration. Conversely, when SI tends to 1 the restoration site has the optimal ecological conditions for all the species of the seed mixture or donor grassland, which is therefore perfectly appropriate for that site restoration.

The RI is an index of the reliability of SI. It is computed by dividing the sum of the abundances in the seed mixture or donor grassland composition of modeled species by the sum of the original seed mixture or donor grassland composition abundances. This index accounts for the species present in the mixture but not in the training database, which are consequently excluded from the modeling. Also the RI ranges between 0 and 1. When RI is close to 0, few to none of the species contribute to the computation of the SI, whereas when RI is close to 1 the SI is computed with most to all the species. Therefore, the higher is the RI, the more reliable is the SI. Not all the species of the seed mixture and donor grassland composition are modeled as i) they can be missing from the training database or ii) the values of the topographic factors of the restoration site are beyond their ecological ranges (e.g. if the elevation of the restoration site is 250 m and a species has an elevation range bounded between 1000 and 3000 m, such a species cannot be modeled).

2.4. Output data

The `RestInd` function produces a list including three outputs:

- a table named '`DESCRIPTIVES`' containing descriptive information related to modeled plant species, such as the number of observations and the minimum, maximum, and mean values of the three topographical features (elevation, slope, and southness) of the surveys in which occurs each species;
- a table named '`SPECIES ABUNDANCES`' containing, for instance, the Predicted Maximum Abundance (PMA), Predicted Optimal Abundance (PMO), the Expected Abundance (EA) of each modeled plant species along with the performance parameters of the best Generalized Additive Model (i.e., adjusted R2 and the RMSE);
- a table names '`INDEXES`' containing the values of the Suitability Index (SI) and Reliability Index (RI).

Additional information related to the output data are retrievable from the help page of `RestInd` function (using `?RestInd` in the R console), from the package website (https://marcopittarello.github.io/ResNatSeed/reference/index.html), and from the "Instruction" section of the Shiny app (https://marco-pittarello.shinyapps.io/ResNatSeed_ShinyApp/).

3. Examples of application

3.1. Study area

The Piedmont training database originates from a dataset containing 4081 surveys derived from five research projects covering the whole region (I tipi pastorali del Piemonte - Cavallero et al. (2007), Ager I-Gral, H2020 Super-G, LIFE Xero-Grazing and RDP Regione Piemonte, 2014–2020 Filierba), accounting for 1114 species. Each survey contains the complete list of species and their abundance. Piedmont region is characterized by a temperate climate, varying from sub-mediterranean areas in the South, to sub-atlantic areas in the North. Moreover, it is surrounded on three sides by mountains, namely the Alps from the South-West to the North-East and the Apennines in the South-East. These highly diverse conditions make Piedmont the Italian region having the highest number of autochthonous plant species (Portale della flora d'Italia, 2022).

3.2. Methods

All the surveys were performed using the vertical point-quadrat method (Daget and Poissonet, 1971) along a 25 or 12.5 m long transect (depending on the project), with the indication of the complete list of species in a buffer of 1 m from each side of the transect. Nomenclature follows Aeschiman et al. (2004). The values of elevation, slope and aspect were extracted from a Digital Terrain Model with a grid of 25 m, available from the Piedmont Region cartographic portal (Regione Piemonte, 2011).

The selection of the species found in at least 50 surveys resulted in 248 species retained in the dataset. The frequency of occurrence of each plant species was converted to 100 measurements by multiplying such frequencies by 2 or 4 for transects with 50 or 25 observation points, respectively (Pittarello et al., 2016; Verdinelli et al., 2022). Such a conversion determines the species percentage cover (%SC), which is an estimate of species canopy cover. A value of 0.3 %SC was added for each additional species present in the 1-m buffer around the transect (Perotti et al., 2018).

3.3. Computation of suitability and reliability indices

Three hypothetical donor grasslands (A, B, C) were used to test the `RestInd` algorithm across three different hypothetical restoration sites (1, 2, 3) (Table 1).

Table 1
Topographic characteristics of the three hypothetical restoration sites used as an example for the modeling process.

Site	Elevation (m a.s.l.)	Slope (°)	Aspect (°)	Elevation Belt
1	250	5	200	Basal
2	850	25	180	Montane
3	1900	25	60	Alpine

The compositions of the three donor grasslands are reported in Table 2, along with the `ResNatSeed` results.

An example in R environment for the computation of the SI and RI for the donor grassland B at site 1 is shown below, whereas Fig. 3 is a screenshot of the application of the same example with the Shiny app.

Table 2
Example of the application of the `RestInd` function of `ResNatSeed` with the composition of three hypothetical donor grasslands (A, B, C) in three restoration sites (1, 2, 3) with different topographic characteristics (see Table 1). The percentage Species Cover (%SC) detected from a vegetation survey and the Expected Abundances (EAs) calculated by `ResNatSeed` for each site are reported for the species identified in the three donor grasslands. The Suitability (SI) and Reliability (RI) Indices are reported at the bottom of the table. Dashes indicate species which were absent in the training database, while NAs indicate species present in the training database but only in different topographic ranges compared to the ones covered in the training database.

Cep_names	Species	Donor grassland A				Donor grassland B				Donor grassland C			
		Data input	`ResNatSeed` output			Data input	`ResNatSeed` output			Data input	`ResNatSeed` output		
		%SC	EA (Site 1)	EA (Site 2)	EA (Site 3)	%SC	EA (Site 1)	EA (Site 2)	EA (Site 3)	%SC	EA (Site 1)	EA (Site 2)	EA (Site 3)
Achiaggr	*Achillea millefolium* aggr									6	1.38	1.92	3.36
Agroschr	*Agrostis schraderiana*	30	NA	NA	26.7								
Alchpent	*Alchemilla pentaphyllea*	2	NA	NA	0.36								
Anthaggr	*Anthoxanthum odoratum* aggr.	18	12.8	7.92	10.8	6	4.26	2.64	3.6				
Arrhelat	*Arrhenatherum elatius*					44	20.2	9.24	6.6				
Bracrupe	*Brachypodium rupestre*									11	NA	8.14	8.58
Bromerec	*Bromus erectus*									42	NA	38.2	30.7
Campbarb	*Campanula barbata*	4	NA	NA	3.04								
Carecary	*Carex caryophyllea*									1	NA	0.34	0.38
Carehirt	*Carex hirta*					2	–	–	–				
Carehumi	*Carex humilis*									9	NA	8.73	5.67
Caresemp	*Carex sempervirens*	46	NA	NA	39.6								
Centscab	*Centaurea scabiosa*									2	NA	0.92	1.58
Cerafont	*Cerastium fontanum*					2	0.74	0.26	0.1				
Dactglom	*Dactylis glomerata*					54	41.6	22.7	44.3				
Euphstri	*Euphrasia stricta*	4	NA	NA	2.72								
Festaggr	*Festuca ovina* aggr.									15	NA	6.75	13.1
Festrubr	*Festuca rubra*	40	NA	36.4	36.4	4	NA	3.64	3.64	2	NA	1.82	1.82
Festscab	*Festuca scabriculmis*	26	NA	NA	18.7								
Galiaggr.2	*Galium lucidum* aggr.									1	NA	0.09	0.17
Geummont	*Geum montanum*	20	NA	NA	9.6								
Helinumm	*Helianthemum nummularium*									2	NA	0.14	1.28
Helivers	*Helictotrichon versicolor*	2	NA	NA	0.8								
Hierpilo	*Hieracium pilosella*									4	NA	0.36	1.52
Hippcomo	*Hippocrepis comosa*									2	NA	0.3	0.64
Holclana	*Holcus lanatus*					14	13.4	10.2	4.76				
Leonhelv	*Leontodon helveticus*	42	NA	NA	21.8								
Leonhisp	*Leontodon hispidus*					24	3.84	1.2	13.7				
Lolipere	*Lolium perenne*					60	55.8	34.2	16.8				
Lotucorn	*Lotus corniculatus*	6	NA	0.6	3.72								
Nardstri	*Nardus stricta*	36	NA	NA	35.3								
Phlerhae	*Phleum rhaeticum*	4	NA	NA	3.56								
Pimpsaxi	*Pimpinella saxifraga*									1	NA	NA	0.65
Planlanc	*Plantago lanceolata*					8	5.92	3.52	1.92				
Poaalpi	*Poa alpina*	16	NA	NA	12.3								
Poaprat	*Poa pratensis*					2	1.12	0.34	1.1				
Poteneum	*Potentilla neumanniana*									1	NA	NA	0.55
Ranubulb	*Ranunculus bulbosus*									3	1.2	0.6	0.18
Rhodferr	*Rhododendron ferrugineum*	4	NA	NA	2.32								
Salvprat	*Salvia pratensis*									1	0.55	0.49	0.3
Sangmino	*Sanguisorba minor*									2	NA	0.32	0.8
Stacrect	*Stachys recta*									1	–	–	–
Stelmedi	*Stellaria media*					2	1.02	NA	NA				
Teuccham	*Teucrium chamaedrys*									9	NA	4.32	0.36
Thymaggr	*Thymus serpyllum* aggr.	2	NA	0.76	1.4					6	NA	2.28	4.2
Trifrepe	*Trifolium repens*					2	1.2	0.44	0.66				
Trisflav	*Trisetum flavescens*					4	NA	0.4	2.52				
Vaccgaul	*Vaccinium gaultherioides*	2	NA	NA	1.4								
Vaccmyrt	*Vaccinium myrtillus*	28	NA	NA	20.2								
	Suitability Index (SI)		0.71	0.69	0.76		0.68	0.40	0.44		0.31	0.64	0.63
	Reliability Index (RI)		0.05	0.20	1.00		0.96	0.98	0.98		0.08	0.98	0.99

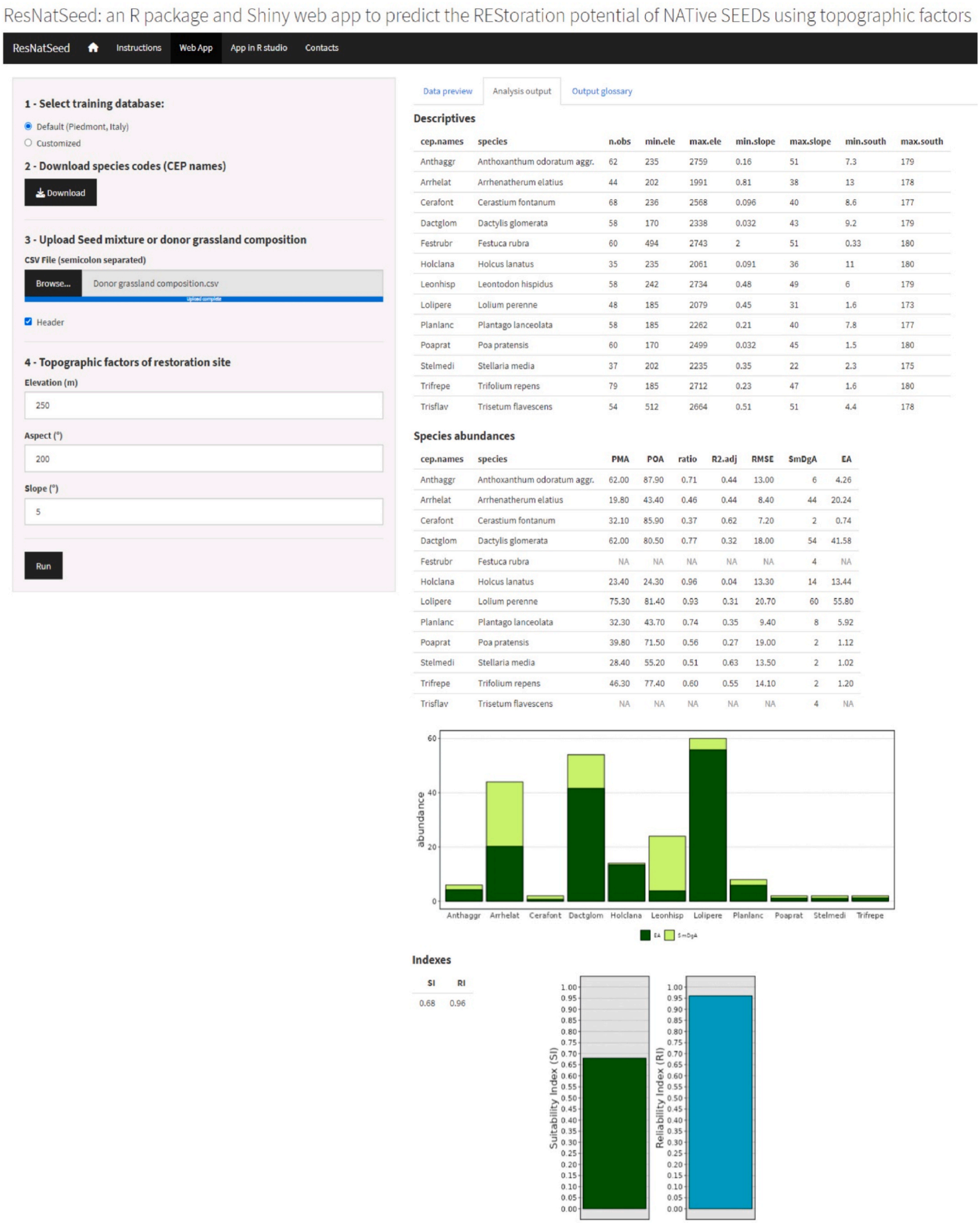

Fig. 3. Screenshot of the `ResNatSeed` usage via the Shiny app. On the left side (grey box) the user can set the input data, namely the donor grassland composition (grassland B in this example) and the topographical features of the restoration site (site 1 in this example). After clicking on the button 'Run', the output of the `RestInd` function will appear in the 'Analysis output' section of the right side of the screen.

```
library(ResNatSeed)
library(readxl)
```

Import of donor grassland B composition:

```
comp.B <- data.frame(read_excel("Donor grassland composition.xlsx",
                     sheet="B"))
comp.B

##    Cep_names  Abundance_%SC
## 1   Lolipere 60
## 2   Dactglom 54
## 3   Arrhelat 44
## 4   Leonhisp 24
## 5   Holclana 14
## 6   Planlanc  8
## 7   Anthaggr  6
## 8   Festrubr  4
## 9   Trisflav  4
## 10  Carehirt  2
## 11  Cerafont  2
## 12   Poaprat  2
## 13  Stelmedi  2
## 14  Trifrepe  2
```

Computation of the Suitability Index (SI) and Reliability index (RI) for the donor grassland B in the site 1:

```
results<-RestInd(trainingDB = NULL, # with 'NULL' the default training database will be
used
        composition=comp.B, # database with the composition of donor grassland B
        elevation=250, #elevation of site 1 (meters)
        slope=5, # slope of site 1 (degrees)
        aspect=200 # aspect of site 1 (North degrees)
)
```

The output of `RestInd`, stored in the object `results` are the following:

```
results

## $DESCRIPTIVES
##          cep.names                      species n.obs min.ele max.ele min.slope
## Anthaggr  Anthaggr Anthoxanthum odoratum aggr.    62     235    2759      0.16
## Arrhelat  Arrhelat       Arrhenatherum elatius    44     202    1991      0.81
## Cerafont  Cerafont          Cerastium fontanum    68     236    2568     0.096
## Dactglom  Dactglom          Dactylis glomerata    58     170    2338     0.032
## Festrubr  Festrubr               Festuca rubra    60     494    2743         2
## Holclana  Holclana             Holcus lanatus    35     235    2061     0.091
## Leonhisp  Leonhisp          Leontodon hispidus    58     242    2734      0.48
## Lolipere  Lolipere              Lolium perenne    48     185    2079      0.45
## Planlanc  Planlanc         Plantago lanceolata    58     185    2262      0.21
## Poaprat    Poaprat               Poa pratensis    60     170    2499     0.032
## Stelmedi  Stelmedi             Stellaria media    37     202    2235      0.35
## Trifrepe  Trifrepe             Trifolium repens    79     185    2712      0.23
## Trisflav  Trisflav         Trisetum flavescens    54     512    2664      0.51
##          max.slope min.south max.south
## Anthaggr        51       7.3       179
## Arrhelat        38        13       178
## Cerafont        40       8.6       177
## Dactglom        43       9.2       179
```

```
## Festrubr        51      0.33       180
## Holclana        36        11       180
## Leonhisp        49         6       179
## Lolipere        31       1.6       173
## Planlanc        40       7.8       177
## Poaprat         45       1.5       180
## Stelmedi        22       2.3       175
## Trifrepe        47       1.6       180
## Trisflav        51       4.4       178
##
## $SPECIES_ABUNDANCES
##    cep.names                    species  PMA  POA ratio R2.adj RMSE SmDgA
## 1   Anthaggr Anthoxanthum odoratum aggr. 62.0 87.9  0.71   0.44 13.0     6
## 2   Arrhelat       Arrhenatherum elatius 19.8 43.4  0.46   0.44  8.4    44
## 3   Cerafont          Cerastium fontanum 32.1 85.9  0.37   0.62  7.2     2
## 4   Dactglom          Dactylis glomerata 62.0 80.5  0.77   0.32 18.0    54
## 5   Festrubr               Festuca rubra   NA   NA    NA     NA   NA     4
## 6   Holclana              Holcus lanatus 23.4 24.3  0.96   0.04 13.3    14
## 7   Leonhisp          Leontodon hispidus 10.8 68.9  0.16   0.71  6.0    24
## 8   Lolipere              Lolium perenne 75.3 81.4  0.93   0.31 20.7    60
## 9   Planlanc         Plantago lanceolata 32.3 43.7  0.74   0.35  9.4     8
## 10   Poaprat               Poa pratensis 39.8 71.5  0.56   0.27 19.0     2
## 11  Stelmedi             Stellaria media 28.4 55.2  0.51   0.63 13.5     2
## 12  Trifrepe            Trifolium repens 46.3 77.4  0.60   0.55 14.1     2
## 13  Trisflav         Trisetum flavescens   NA   NA    NA     NA   NA     4
##       EA
## 1   4.26
## 2  20.24
## 3   0.74
## 4  41.58
## 5     NA
## 6  13.44
## 7   3.84
## 8  55.80
## 9   5.92
## 10  1.12
## 11  1.02
## 12  1.20
## 13    NA
##
## $INDEXES
##     SI   RI
## 1 0.68 0.96
```

. (*continued*).

3.4. Interpretation of the results

Donor grassland A was a typical alpine mesophilous grassland belonging to the phytosociological order *Caricetalia curvulae*, dominated by *Carex sempervirens* Vill., *Festuca rubra* aggr. and *Leontodon helveticus* (Mérat) Holub. The most suitable restoration site for this donor grassland was site 3, a steep NE-facing slope located at 1900 m a.s.l., for which the highest values of both SI and RI were retrieved. All the species were present in the training database, as indicated by the RI, which was 1, hence every species was successfully modeled. According to the RI, the modeling of the sites 1 and 2 accounted only for 5% (RI = 0.05) and 20% (RI = 0.20) of the abundance, respectively, thus the SI for those sites was highly unreliable.

Arrhenatherum elatius (L.) P.Beauv. ex J.Presl & C.Presl, *Dactylis glomerata* L. and *Lolium perenne* L. were the dominant species in donor grassland B, a lowland hay meadow, belonging to the phytosociological order *Arrhenatheretalia elatioris*. Only one species was absent in the training database, i.e. *Carex hirta* L., so the RI was high in all three restoration sites, being RI ≥ 0.96. Most of the species were potentially present in the three restoration sites, but only the first one had a SI higher than 0.5, being the most suitable site for this seed mixture.

Donor grassland C was a dry grassland of medium altitude dominated mainly by *Bromus erectus* (Huds.) Fourr. and secondarily by *Festuca ovina* aggr. and *Brachypodium rupestre* (Host) Roem. & Schult. This vegetation community belongs to the phytosociological order *Brometalia erecti*. Only *Stachys recta* L. was absent in the training database, explaining the maximum RI = 0.99. While a few species were adapted to the first restoration site, sites 2 and 3 were both sufficiently suitable (SI ≥ 0.63 and RI ≥ 0.98), despite their huge topographic differences. This is mainly due to the wide ecological optimum of the dominant species *Bromus erectus*. Indeed, according to the response curves derived from the GAMs (Fig. S1), this species showed a high plateau of %SC bounded between 800 m and 2000 m a.s.l. and between 60° and 180° of southness. Such opposite elevational and southness limits characterized the differences between the restoration sites 2 and 3, which however were both suitable for this seed mixture.

The three examples (Table 2) were highly representative of the function of `ResNatSeed`. Indeed, the three donor grasslands had a composition most compatible with the areas with their most similar ecological requirements. Moreover, the indices were useful to identify the level of compatibility. Indeed, the second donor grassland was a good example of the fact that the perfect donor grassland could be absent from the proposed ones, implying the need of further research of potential grasslands with a higher compatibility.

4. Conclusions

Restoration of degraded lands with native seeds is becoming widespread across the globe. `ResNatSeed` could help reduce the subjectivity during the assessment of ecological compatibility between donor grasslands and restoration sites. It can improve the choice of the best-suiting species or donor grasslands according to their ecological compatibility with the restoration site. However, as `ResNatSeed` cannot calculate the germinability of the mixture, it should be used by users as a source of general indication of the performance suitability of a seed mixture. Nowadays, `ResNatSeed` uses only three main topographic variables easily extractable from a DEM, but it could be further improved with other topographic, climatic or litho-pedological variables available for different areas or purposes. Moreover, nowadays several large vegetation databases are available for different geographical areas. These large databases could be used to train `ResNatSeed` for specific areas of interest, expanding the potential use of the tool for restoration ecology modeling. The implementation of `ResNatSeed` on an easy-to-use Shiny web app further extends the applicability of the tool, allowing professionals and non-experts of the software R to use this modeling method for restoration purposes.

Funding

The research whose results are reported in this paper were funded by European Union Horizon 2020 research and innovation programme, under grant agreement 774124, project SUPER-G (Developing Sustainable Permanent Grassland Systems and Policies), and Regione Piemonte RDP 2014–2020 programme, operation 16.1.1, project Prà da Smens.

Author contributions

M.L., G.L., M.P., D.B. conceived and designed the research; D.B., M.P. developed the tools; M.L., G.L., M.P., D.B. wrote and edited the manuscript.

Software and data availability

- R package
  - o Source code of `ResNatSeed` R package: https://github.com/MarcoPittarello/ResNatSeed
  - o Package website and vignettes: marcopittarello.github.io/ResNatSeed/
- Shiny app
  - o Source code of `Resnatseed` Shiny App: https://github.com/MarcoPittarello/ResNatSeed_ShinyApp
  - o Shiny app website: https://marco-pittarello.shinyapps.io/ResNatSeed_ShinyApp/. In addition to the `ResNatSeed` application, the website includes a series of tutorials for the correct use of the tool and a glossary of terms used.

Declaration of competing interest

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

Data availability

The code is freely available on github

Acknowledgements

First and foremost, authors acknowledge the numerous colleagues who carried out the surveys for various projects in the past years contributing to the implementation of the dataset used for the training of `ResNatSeed`.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2023.105813.

References

Aeschimann, D., Lauber, K., Moser, D.M., Theurillat, J.P., 2004. Flora Alpina. Belin & Zanichelli, Haupt.

Albert, Á.J., Mudrák, O., Jongepierová, I., Fajmon, K., Frei, I., Ševčíková, M., Klimešová, J., Doležal, J., 2019. Grassland restoration on ex-arable land by transfer of brush-harvested propagules and green hay. Agric. Ecosyst. Environ. 272, 74–82. https://doi.org/10.1016/j.agee.2018.11.008.

Amatulli, G., Domisch, S., Tuanmu, M., Parmentier, B., Ranipeta, A., Malczyk, J., Jetz, W., 2018. Data Descriptor: a suite of global, cross-scale topographic variables for environmental and biodiversity modeling. Sci. Data 5, 180040. https://doi.org/10.1038/sdata.2018.40.

Austin, M.P., 2013. Vegetation and environment: discontinuities and continuities. In: van der Maarel, E., Franklin, J. (Eds.), Vegetation Ecology, second ed. Wiley-Blackwell, pp. 71–106. https://doi.org/10.1002/9781118452592.ch3.

Austin, M.P., Belbin, L., Meyers, J.A., Doherty, M.D., Luoto, M., 2006. Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. Ecol. Model. 199, 197–216. https://doi.org/10.1016/j.ecolmodel.2006.05.023.

Barni, E., Freppaz, M., Siniscalco, C., 2007. Interactions between vegetation, roots, and soil stability in restored high-altitude ski runs in the Alps. Arctic Antarct. Alpine Res. 39, 25–33. https://doi.org/10.1657/1523-0430(2007)39[25:IBVRAS]2.0.CO;2.

Barrel, A., Bassignana, M., Curtaz, A., Huc, S., Koch, E.M., Spiegelberger, T., 2015. Native Seeds for the Ecological Restoration in Mountain Zone. Production and Use of Preservation Mixtures. Institut Agricole Régional, Aosta.

Belletti, P., Ferrazzini, D., Piotti, A., Monteleone, I., Ducci, F., 2012. Genetic variation and divergence in Scots pine (Pinus sylvestris L.) within its natural range in Italy. Eur. J. For. Res. 131, 1127–1138. https://doi.org/10.1007/s10342-011-0584-3.

Boehm, A.R., Hardegree, S.P., Glenn, N.F., Reeves, P.A., Moffet, C.A., Flerchinger, G.N., 2021. Slope and Aspect Effects on See db e d Microclimate and Germination Timing of Fall-Planted Seeds. Rangel. Ecol. Manag. 75, 58–67. https://doi.org/10.1016/j.rama.2020.12.003.

Braun-Blanquet, J., 1928. Pflanzensoziologie. Grundzüge der Vegetationskunde. Springer, Berlin, DE.

Cavallero, A., Aceto, P., Gorlier, A., Lombardi, G., Lonati, M., Martinasso, B., Tagliatori, C., 2007. I Tipi Pastorali Delle Alpi Piemontesi. Alberto Perdisa Editore, Bologna.

Chang, C.-R., Lee, P.-F., Bai, M.-L., Lin, T.-T., 2004. Predicting the geographical distribution of plant communities in complex terrain - a case study in Fushian Experimental Forest, northeastern Taiwan. Ecography 27, 577–588. https://doi.org/10.1111/j.0906-7590.2004.03852.x.

Chang, W., Cheng, J., Allaire, J.J., Xie, Y., McPherson, J., 2019. Shiny: Web Application Framework for R.

Daget, P., Poissonet, J., 1971. Une methode d'analyse phytologique des prairies. In: Annales Agronomiques, 48. CNRSB.P., Montpellier, France.

Dallas, T.A., Santini, L., Decker, R., Hastings, A., 2020. Weighing the evidence for the abundant-center hypothesis. Biodivers. Inf. 15, 81–91. https://doi.org/10.17161/bi.v15i3.11989.

De Vitis, M., Abbandonato, H., Dixon, K.W., Laverack, G., Bonomi, C., Pedrini, S., 2017. The European native seed industry: characterization and perspectives in grassland restoration. Sustainability 9 (10), 1682. https://doi.org/10.3390/su9101682.

Doherty, K.D., Butterfield, B.J., Wood, T.E., 2017. Matching seed to site by climate similarity: techniques to prioritize plant materials development and use in restoration. Ecol. Appl. 27 (3), 1010–1023. https://doi.org/10.1002/eap.1505.

Falk, W., Mellert, K.H., 2011. Species distribution models as a tool for forest management planning under climate change: risk evaluation of Abies alba in Bavaria. J. Veg. Sci. 22 (4), 621–634. https://doi.org/10.1111/j.1654-1103.2011.01294.x.

Gonzáles, E., Rochefort, L., Boudreau, S., Poulin, M., 2014. Combining indicator species and key environmental and management factors to predict restoration success of degraded ecosystems. Ecol. Indicat. 46, 156–166. https://doi.org/10.1016/j.ecolind.2014.06.016.

Jiménez-Alfaro, B., Frischie, S., Stolz, J., Gálvez-Ramírez, C., 2020. Native plants for greening Mediterranean agroecosystems. Nat. Plants 6, 209–214. https://doi.org/10.1038/s41477-020-0617-3.

Kiehl, K., Kirmer, A., Donath, T.W., Rasran, L., Hölzel, N., 2010. Species introduction in restoration projects – evaluation of different techniques for the establishment of semi-natural grasslands in Central and Northwestern Europe. Basic Appl. Ecol. 11, 285–299. https://doi.org/10.1016/j.baae.2009.12.004.

Kildisheva, O.A., Dixon, K.W., Silveira, F.A.O., Chapman, T., Di Sacco, A., Mondoni, A., Turner, S.R., Cross, A.T., 2020. Dormancy and germination: making every seed count in restoration. Restor. Ecol. 28 (S3), S256–S265. https://doi.org/10.1111/rec.13140.

Körner, C., 2003. Alpine Plant Life. Springer, Heidelberg, Germany. https://doi.org/10.1007/978-3-642-18970-8.

Kowarik, I., 1987. Kritische Anmerkungen zum theoretischen Konzept der potentiellen natürlichen Vegetation mit Anregungen zu einer zeitgemässen modifikation. Tuexenia 7, 53–67.

Lüdecke, D., Ben-Shachar, M.S., Patil, I., Waggoner, P., Makowski, D., 2021. Performance: an R package for assessment, comparison and testing of statistical models. J. Open Source Softw. 6 (60), 3139. https://doi.org/10.21105/joss.03139.

Mainz, A.K., Wieden, M., 2018. Ten years of native seed certification in Germany – a summary. Plant Biol. 21, 383–388. https://doi.org/10.1111/plb.12866.

Perotti, E., Probo, M., Pittarello, M., Lonati, M., Lombardi, G., 2018. A 5-year rotational grazing changes the botanical composition of sub-alpine and alpine grasslands. Appl. Veg. Sci. 21, 647–657. https://doi.org/10.1111/avsc.12389.

Piemonte, Regione, 2011. Digital terrain model, resolution 25m. Available from: https://www.geoportale.piemonte.it/cms/.

Pittarello, M., Probo, M., Lonati, M., Lombardi, G., 2016. Restoration of sub-alpine shrub-encroached grasslands through pastoral practices: effects on vegetation structure and botanical composition. Appl. Veg. Sci. 19, 381–390. https://doi.org/10.1111/avsc.12222.

Portale della flora d'Italia, 2022. Portal to the Flora of Italy. http://dryades.units.it/floritaly/index.php.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL. https://www.R-project.org/.

Sagarin, R.D., Gaines, S.D., Gaylord, B., 2006. Moving beyond assumptions to understand abundance distributions across the ranges of species. Trends Ecol. Evol. 21 (9), 524–530. https://doi.org/10.1016/j.tree.2006.06.008.

Santini, L., Pironon, S., Maiorano, L., Thuiller, W., 2019. Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses. Ecography 42, 696–705. https://doi.org/10.1111/ecog.04027.

Scotton, M., 2018a. Wild seed harvesting at mountainous species-rich grassland in calcareous Italian Alps. Rangel. Ecol. Manag. 71, 762–769. https://doi.org/10.1016/j.rama.2018.04.005.

Scotton, M., 2018b. Seed production in grassland species: morpho-biological determinants in a species-rich semi-natural grassland. Grass Forage Sci. 73 (3), 764–776. https://doi.org/10.1111/gfs.12359.

Scotton, M., Ševčíková, M., 2017. Efficiency of mechanical seed harvesting for grassland restoration. Agric. Ecosyst. Environ. 247, 195–204. https://doi.org/10.1016/j.agee.2017.06.040.

Scotton, M., Piccinin, L., Dainese, M., Sancin, F., 2009. Seed harvesting for ecological restoration: efficiency of haymaking and seed-stripping on different grassland types in the eastern Italian Alps. Ecol. Restor. 27 (1), 66–75. https://doi.org/10.3368/er.27.1.66.

Shryock, D.F., DeFalco, L.A., Esque, T.C., 2022. Seed Menus: an integrated decision-support framework for native plant restoration in the Mojave Desert. Ecol. Evol. 12, e8805 https://doi.org/10.1002/ece3.8805.

Somodi, I., Molnár, Z., Czúcz, B., Bede-Fazekas, Á., Bölöni, J., Pásztor, L., Laborczi, A., Zimmermann, N.E., 2017. Implementation and application of multiple potential natural vegetation models – a case study of Hungary. J. Veg. Sci. 28 (6), 1260–1269. https://doi.org/10.1111/jvs.12564.

Verdinelli, M., Pittarello, M., Caria, M.C., Piga, G., Roggero, P.P., Marrosu, G.M., Arrizza, S., Fadda, M.L., Lombardi, G., Lonati, M., Nota, G., Sitzia, M., Bagella, S., 2022. Congruent responses of vascular plant and ant communities to pastoral land-use abandonment in mountain areas throughout different biogeographic regions. Ecological Processes 11, 35. https://doi.org/10.1186/s13717-022-00379-9.

Wagner, M., Hulmes, S., Hulmes, L., Redhead, J.W., Nowakowski, M., Pywell, R.F., 2021. Green hay transfer for grassland restoration: species capture and establishment. Restor. Ecol. 29 (S1), e13259 https://doi.org/10.1111/rec.13259.

Waldock, C., Stuart-Smith, R.D., Edgar, G.J., Bird, T.J., Bates, A.E., 2019. The shape of abundance distributions across temperature gradients in reef fishes. Ecol. Lett. 22, 685–696. https://doi.org/10.1111/ele.13222.

Wiser, S.K., 2016. Achievements and challenges in the integration, reuse and synthesis of vegetation plot data. J. Veg. Sci. 27, 868–879. https://doi.org/10.1111/jvs.12419.

Wood, S.N., 2017. Generalized Additive Models: an Introduction with R, second ed. Chapman and Hall/CRC.