Driving forces of forest expansion dynamics across the Iberian Peninsula (1987–2017): A spatio-temporal transect

(A) Complementary information to the main text

2.2. Methodological framework

2.2.2. New forests' occurrence extraction, sampling on each BR, and filtering

The classified maps were overlaid to derive the land cover (LC) changes and temporal trajectories in the last thirty years. We selected forest locations in 2002 and 2017 (*needleleaf evergreen forests*, *broadleaf evergreen forests*, *broadleaf deciduous forests*) that were classified as crops categories, *grasslands*, or *shrublands* in the previous dates. Then, a set of sampling strategies were used for modeling focusing on (i) reducing the absence bias, (ii) applying uncertainty filtering, (iii) evaluating the presence-absence spatial autocorrelation, (iv) subsampling in bioclimatic regions (BRs), and finally, (v) equalizing the presence-absence locations (by prevalence). LC dynamics were classified in the form of binary presence-absence events. Thus, remaining stable locations in all periods were considered to be absences while changing locations towards new forest (NF) were considered to be presences.

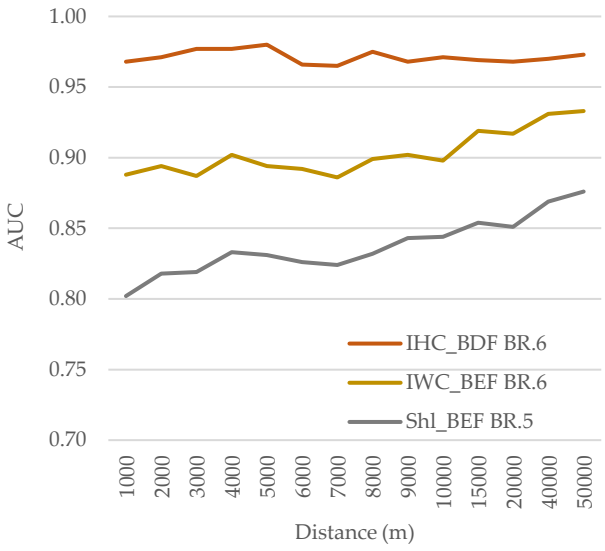
Commonly, presence and absence points are extracted randomly, regardless of the density of presences-absences in a local neighborhood. However, in the context of extensive LC change analyses (as in this research), considering random sampling over all the available absence points implies incorporating into the model locations with biophysical and socioeconomic attributes significantly distanced or without spatial relation to the presences, even within a BR (**absence sample bias**). A set of testbeds was performed to approximate an optimal distance to extract absence samples, which analyzed how the sampling distance to select absences according to presences influences the modeling results. This strategy allows for generating more spatially balanced presence-absence samples, fitting the number of absences according to the frequency of local presences. Three NF dynamics' (IHC to BDF, IWC to BEF, and Shl to BEF) sample locations and a predefined set of predictors were considered to evaluate the BRT models' performance (the area under the receiver operating curve, AUC) at specific sampling distances (Figure S1).

Figure S1. Square buffering distance testbeds for presence-absence balanced sampling. In (a), BRT models' performance results evaluated for each square buffer distance are shown, with their (b) graphical evolution.

Square buffer distance	IHC-BDF BR.6	IWC-BEF BR.6	Shl-BEF BR.5
1000 [1]	0.968 [2]	0.888 [2]	0.802 [2]
2000	0.971	0.894	0.818
3000	0.977	0.887	0.819
4000	0.977	0.902	0.833
5000	0.980	0.894	0.831
6000	0.966	0.892	0.826
7000	0.965	0.886	0.824
8000	0.975	0.899	0.832
9000	0.968	0.902	0.843
10000	0.971	0.898	0.844
15000	0.969	0.919	0.854
20000	0.968	0.917	0.851
40000	0.970	0.931	0.869
50000	0.973	0.933	0.876

[1] distance in meters. [2] Area under the curve (AUC).

(**a**)



(**b**)

Calculating the presence-absence density within a sampling distance required the considered BR area to be divided into equal-sized squares (as tiles), with a diagonal length (as the sampling dimension) ranging from 1,000 to 50,000 m. The numbers of presence and absence locations were estimated, and the minimum of the two values were calculated for each square. Then, absences in each square were

randomly sampled according to the minimum value calculated. This strategy permitted deriving a more spatially balanced presence-absence sample for modeling. The higher performance in the models was achieved by applying sampling distances over 20,000 m. This threshold distance value was used to derive absence samples for modeling.

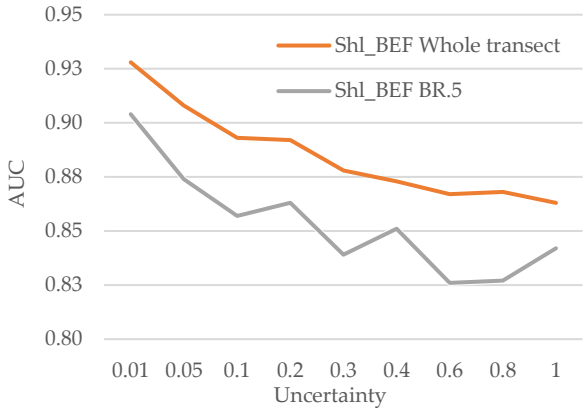
Each classified map had an associated **uncertainty** layer that specified the confidence of the classified pixels according to highly reliable training areas [54]. Using these uncertainty layers, the effect of different uncertainty restrictions on model performance and the number of pixels remaining for modeling were evaluated. As was expected, the global accuracy of the model increased when pixels with higher values of uncertainty were eliminated [57,58]. Therefore, due to the existing trade-off between uncertainty restrictions and an adequate number of pixels for modeling, two uncertainty levels were used: *uncertainty* < 0.1 was applied when there was a large sample size (e.g., NF from shrublands and grasslands that theoretically should provide ultra-high quality models), and *uncertainty* < 0.3 for classes with a limited number of pixels available for modeling (e.g., NF from crop categories). Figure S2 provides further details about the testbeds performed, and Table S6a details the remaining sample size after the uncertainty filtering.

Figure S2. Testbeds modeling using a set of *uncertainty* restrictions. In (a), BRT models' performance results are shown, with their (b) graphical correspondence.

Uncertainty	Shl to BEF Whole transect	Shl to BEF BR.5
0.01	0.928 [1]	0.904 [1]
0.05	0.908	0.874
0.1	0.893	0.857
0.2	0.892	0.863
0.3	0.878	0.839
0.4	0.873	0.851
0.6	0.867	0.826
0.8	0.868	0.827
1	0.863	0.842

[1] area under the curve (AUC).

(**a**)



(**b**)

The Moran index was calculated at discrete distances to **identify uncorrelated sampling points** of the dependent (i.e., presences-absences) and independent variables in the geographic space. The Moran index calculates the distance a point can be considered independent of another. Testbeds were performed, which found that at a distance between points of 140 m, the correlation coefficient dropped below 0.7, where the slope (a topoclimatic variable) was considered the independent variable serving to evaluate the autocorrelation test. We applied this minimum distance threshold to derive spatially uncorrelated presence-absence points for modeling.

The transect area was subdivided into **bioclimatic regions** (BRs) to analyze NF transition models in a more meaningful climatic way. The average of the mean temperature surfaces [59] comprising the 1987–2017 period and the BR's formulation [51], were used to subdivide the transect according to its climatic characteristics. A general overview of the BRs is detailed in Figure S4.

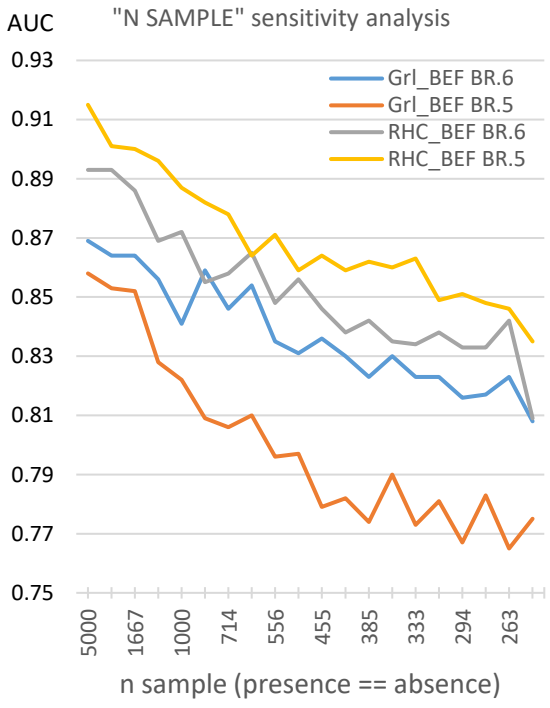
A **presence-absence balanced** sample was generated to be used for modeling, which assured the presence-absence sample's prevalence. In addition, when the number of presences-absences in the BRs exceeded 5 000 points, the points were randomly selected and limited to this value to reduce time processing. We admit that this sample restriction value could have compromised the extent to which the sample size might affect NF modeling. Several studies have evaluated the effect of sample size on species distribution model performance [60–63]. However, due to the lack of previous literature references regarding the optimal number of samples used for BRT modeling and the fact that diverse

sample size was assumed in other research [4,33,47,48], we performed a set of testbeds to investigate how a sampling size affected the model's performance and hence, established a minimum sampling size for modeling. The Grl to BEF and RHC to BEF dynamics samples in the Supramediterranean (BR.5) and Southern Mesomediterranean (BR.6) bioclimatic regions were selected because they provided a large initial sample size of around 5 000 presence-absence points (Figure S3). A first BRT model with the whole sample (5 000 points) served to derive the reference accuracy estimate. Then, we consecutively split the original sample from 2 to 20 parts (randomly and without replacement), generating k-subsamples used to evaluate the models' performance in each subsample. For instance, dividing the original sample into five parts would yield 5 BRT models that consider 1 000 presence-absence points each to evaluate each model's performance. As expected, the accuracy and AUC dropped when the sample size was reduced, increasing the models' performance dispersion (standard deviation) evaluated in each k-parts subsample. Therefore, we assumed an accuracy loss limit of 5% (maximum) from the initial model's performance estimated with the whole sample. Averaging the models' accuracy loss series, the accuracy loss limit percentage corresponds with 580 presence-absence points, a threshold used as a criterion to discard models (and therefore NF dynamics) with low sample representativeness in the BRs. Table S6b presents the number of pixels used for modeling, disaggregated by NF dynamics and BR.

Figure S3. Testbeds for the evaluation of the minimum sampling size through BRT modeling. In (a), the BRT models' performance results are shown, with their (b) graphical correspondence.

Sample size	Grl_BEF BR.6	Grl_BEF BR.5	RHC_BEF BR.6	RHC_BEF BR.5
5000	0.869	0.858	0.893	0.915
2500	0.864	0.853	0.893	0.901
1667	0.864	0.852	0.886	0.900
1250	0.856	0.828	0.869	0.896
1000	0.841	0.822	0.872	0.887
833	0.859	0.809	0.855	0.882
714	0.846	0.806	0.858	0.878
625	0.854	0.810	0.865	0.864
556	0.835	0.796	0.848	0.871
500	0.831	0.797	0.856	0.859
455	0.836	0.779	0.846	0.864
417	0.830	0.782	0.838	0.859
385	0.823	0.774	0.842	0.862
357	0.830	0.79	0.835	0.860
333	0.823	0.773	0.834	0.863
313	0.823	0.781	0.838	0.849
294	0.816	0.767	0.833	0.851
278	0.817	0.783	0.833	0.848
263	0.823	0.765	0.842	0.846
250	0.808	0.775	0.809	0.835

(**a**)



(**b**)

Lastly, once the presence-absence sample points were generated, they were intersected with the predictor variables, resulting in point features with data included in their attribute table.

2.2.3. Land Cover Change Modeling: Explanatory variables

The spatial determinants potentially driving the past NF transitions listed in Table 2 in the main text are described in the following paragraphs.

The **topography** defines physical characteristics that may have influenced NF occurrence. A 10 m per cell Digital Elevation Model (DEM) was used, derived from a 1:5 000 lidar dataset from the Spanish National Plan for Aerial Orthophotography 2010 (PNOA, https://pnoa.ign.es/el-proyecto-pnoa-lidar, last accessed on 1 July 2021) [71]. From that, several variables were derived:

- Slope, calculated in degrees.
- General curvature (dimensionless), estimated as the second derivative of the DEM, defines how convex or concave a surface is, which approximates soil moisture and the convergence/divergence of overland accumulated water flow [72]. A positive value denotes a convex surface, a negative value denotes a concave surface, and values close to 0 denote flat terrain landforms.
- Potential solar radiation (10 kJ / (m2 × day × μm)), evaluated on winter solstice dates following [73], indicates the available energy for NF growth in the most restrictive conditions.

Climatic conditions were characterized by the averaged mean temperature (d°C), the accumulated precipitation (dmm), and the Standardized Precipitation-Evapotranspiration Index (SPEI), extracted from the Topoclimatic Drought Atlas of the Spanish Iberian Peninsula [74–76]. The dataset includes monthly aggregates at a 100 m spatial resolution between 1950–2017, which was then aggregated annually. Each LC map produced was characterized by averaging the 15 years of climatic data prior to its reference date, i.e., 1972–1986 for LCM-1987, 1987–2001 for LCM-2002, and 2002–2016 for LCM-2017. The SPEI was considered a possible explanatory variable to evaluate drought impact on vegetation, given its multiscale analysis capacity for characterizing water availability, a driver that constrains primary and secondary forest growth [77]. Furthermore, SPEI indexes were calculated for 3, 6, 12, and 24 months using the 1950-2017 climate time series, providing a wide variety of drought aggregation to characterize the drought tolerance of vegetation (i.e., crops, natural/seminatural). Two temporal ranges were adapted to the analysis periods. For P1, SPEI indexes ranged from 1978–2002 and for P2, from 1993–2017, representing a buffer time of 25 years of drought characterization. Then, the number of drought episodes during which the drought intensity overcame -1 (SPEI < -1 as a measure of water deficit) and +1 (SPEI > +1 for water oversupply conditions) were counted, generating sixteen drought/humid intensity variables for each period.

To assess the influence of land accessibility and the topographic effect in the NF appearance, distance measures were calculated based on Euclidean distances and Cost distances. Euclidean distances were considered precise enough to avoid Geodetic distance computation, which would take into account the limited scopes considered and the characteristics of the cartographic projection used (UTM). The Euclidean distance variables used included the Euclidean distance to hydrography (first, second, and third stream order with perennial/intermittent streamflow regimes), the Euclidean distance to protected areas (national and natural parks), and the Euclidean distance to any forest patch (without discriminating between forest groups) with an area more than 120 × 120 m². Cost distance variables were used in two ways; first, using cost distance to urban areas and cost distance to provincial capitals, in both cases weighting the Euclidean distance with the slope and friction coefficients according to cell values (i.e., primary or secondary road network). Secondly, we estimated the cost distance to the main roads and the cost distance to secondary roads, using the slope as the cost surface. All distance variables were calculated using datasets from the Spanish National Geographic Institute (https://www.ign.es/web/ign/portal/cbg-area-cartografia, accessed on 01/07/2021). In addition, the LCM-1987 was used to derive the Euclidean distance to preexisting forests.

The lithological substrate may influence the nature of soils and, hence, the vegetation tolerance to acidic pH conditions (low pH, as in silicate rocks) or alkaline (high pH, as in carbonate rocks). As there is not a soil map for the whole area, we extracted the lithologic groups from the Geologic map of Spain at a scale of 1:1 000 000 (https://info.igme.es/catalogo/, accessed on 01/07/2021), and we reclassified them according to three categories: acidic, mixed, and alkaline.

Soil degradation is a severe environmental problem. Of several forms of soil erosion (water, wind, or landslide), water is the most influential on the fast degradation of natural ecosystems and land productivity. The European Union proposed a Directive (22/9/2006) that establishes a framework for soil protection. On this basis, the Spanish National Soil Erosion Inventory [78] provided a decennial

map (2002–2012) of the principal soil erosion agents at a 1:50 000 scale. We considered the water erosion agent ("sheet and rill erosion" product), estimated by applying the RUSLE (Revised Universal Soil Loss Equation) model [79]. This model is simple in its application and is widely disseminated and accepted by the scientific community. It predicts soil erosion based on six factors: the aggressiveness of the rain, the susceptibility of the soil to erosion, the length of the slope, the slope, the vegetation cover, and the soil conservation practices. Together, these factors predict the rate of soil loss (Mg/(ha × year)), which were aggregated into seven categorical levels (Table S5). Therefore, "sheet and rill erosion" is considered a driver that determines the local soil quality (carbon content, soil depth, fertility), which may affect forest regrowth, especially when related to cropland abandonment processes in marginal lands [80–82].

The level of protection in natural areas may have influenced forest regrowth associated with natural succession dynamics, in addition to agricultural abandonment in more inaccessible regions with less productive lands. The special status of protected areas with a historic long-term high-grade protection was considered, i.e., national parks and natural parks, according to the Spanish Natural Protected Areas database (https://www.redeuroparc.org/, accessed on 01/07/2021). The protected areas were split according to their initial protection date, adapted to the analysis periods (P1, P2) in the form of a dichotomous variable.

Population and agricultural censuses were collected to consider the potential repercussion of population characteristics in agrarian land-use change decisions and related forest regrowth dynamics. Datasets were collected at the municipality level by the Spanish National Statistics Institute (INE, https://www.ine.es/, accessed on 01/07/2021). We included the Agricultural Census 1989, 1999, and 2009, and the Population Census 1991, 2001, and 2011 datasets due to their temporal proximity to the LC maps' reference dates. The variables from the datasets allowed for evaluating population density and migration, the ageing structure, the economic activity and workforce composition, and the productivity of agricultural holdings.

Finally, fires directly impact forest ecosystems and land cover dynamics analyses, so we collected historical fire perimeter datasets through the Spanish regional forest management administrations (e.g., Andalucía, Aragón, Castilla-La Mancha, Castilla y León, Euskadi, Navarra, Madrid, Murcia) at the maximum temporal resolution. Therefore, burned areas were masked from the analysis, a process necessary to avoid matching artifacts between LCM temporal resolution (15 years), the year of the wildfire, and the time it takes for a forest to recover from fire disturbances, which depends on humidity and fire severity [83].

3.3. New forests' main drivers

The following paragraphs provide a detailed description of the results section, which serves as a basis for the discussion section detailed in the main text.

The NF driving forces were analyzed, focusing on the source categories (rainfed crop categories, irrigated crop categories, grasslands, and shrublands). The maps detailed in Figure 5a/b/c in the main text describe the drivers' temporal and spatial variation characteristics, detailed by period and BR. In this section, drivers are listed by their relative importance, accompanied by a symbol (in brackets) denoting a partial dependence main pattern, which is described in the main text section entitled 3.3 *New forests main drivers*. For instance, the (+grad) means that the transition probability increases with the variable; the (-grad) represents the opposite; and the (≠grad) means that there is not a clear positive or negative tendency. The (A-shape) indicates that the higher suitable conditions occur between two low values, while the (V-shape) denotes two high suitable conditions between a minimum. The (и-shape) depicts relatively high suitable conditions in low values that sharply decrease, then increase to a second peak, and finally decrease at larger values. Lastly, the (±grad) is used in the case of a specific driver depicts different patterns when aggregating BRs.

3.3.1. Drivers of new broadleaf deciduous forests (BDF)

Regarding rainfed (herbaceous and woody crops) source categories, according to Table 4, enough occurrences were detected to perform an analysis in the Supramediterranean BR for the RHC source category in the first period. The most important driver was the distance to forests (-grad), followed by precipitation (∩-shape), distance to hydrography network (-grad), distance to provincial capitals (+grad), and distance to urban settlements (-grad). Therefore, drivers suggest that the increase of deciduous forest from rainfed crops in the Supramediterranean BR was a process of forest completion in areas near forested areas, streams and torrents, and while far from capitals, relatively close to urban settlements.

In the second period, the number of occurrences increased in the Northern and Southern Mesomediterranean BRs (Figure 5a in the main text) for the RHC and RWC source categories. For RHC, the distance to the hydrographic network (-grad) was the main driver in the Northern Mesomediterranean BR, and in the Southern Mesomediterranean and Supramediterranean BRs (Figure S7a1), the distance to forests (-grad) was the main driver, followed by precipitation (±grad), distance to provincial capitals (V-shape) and urban settlements (-grad). Therefore, compared to the first period, the new BDF areas from rainfed crops appeared closer to rivers or streams, further from forests, further from provincial capitals, relatively close to local urban settlements and secondary roads, and with different precipitation patterns. In the case of RWC in the Southern Mesomediterranean BR, the number of holdings (-grad) was the main driver, followed by the distance to the hydrographic network (-grad), precipitation (-grad), and distance to forests (-grad).

For irrigated crops (herbaceous and woody crops) source categories during the first period, the most important drivers were the distance to the hydrographic network (-grad) followed by the distance to forests (-grad), which was clearly visible in Northern and Southern Mesomediterranean BRs (Table S7). Other primary drivers were the distance to provincial capitals (+grad) in the Southern Mediterranean BR, the slope (+grad), and the distance to main roads (+grad) in Coline and Southern and Northern Mesomediterranean BRs. Secondary drivers were the general curvature, temperature, distance to urban settlements, distance to secondary roads, soil erosion, and precipitation. Therefore, the drivers suggested that new BDF from irrigated crops was a process of forest aggregation or completion on lands closer to rivers, torrents, and streams (associated with riparian vegetation), with a steeper slope, and further from human activities, represented by primary and secondary roads, urban settlements (in less intensity), and provincial capitals.

In the second period, the most crucial drivers remained in the same BRs as in the first period but drivers appeared notably in the Thermomediterranean in this period. Other main drivers affecting the same BRs were temperature (±grad), distance to capitals (+grad), and slope (+grad for herbaceous and -grad for woody crops). Secondary drivers included the general curvature (V-shape), soil erosion (-grad), and with less importance, precipitation (+grad) and distance to urban settlements. Moreover, new drivers were identified in this period: solar radiation (V-shape) and population density (-grad). When comparing the temporal PDPs (Figure S7a2), the patterns observed are consistent with forest consolidation and completion occurring during the first period and forest expansion in the second period, the latter of which appeared within the limits of forested areas or at an even greater distance. Therefore, compared to the first period, the process of forest aggregation and completion was more clearly associated with riparian vegetation near rivers, streams, and torrents, a greater distance from provincial capitals and, in less intensity, with the distance to main roads and urban settlements, and population density. The relative importance of the distance to the hydrography network and the distance to forests decreased when comparing both periods (Table S7).

Concerning the grasslands source category during the first period, the main driver was the distance to forests (-grad) very clearly in the Montane and Supramediterranean BRs, though less important in the Coline and Southern Mesomediterranean BRs. Other main drivers were the temperature (-grad) and precipitation (+grad) in the Southern Mesomediterranean, versus soil erosion (-grad) and distance to provincial capitals (-grad) in the Coline BR. Secondary drivers were solar

radiation, slope, distance to urban settlements (-grad), number of holdings (-grad), and distance to main roads (-grad). Therefore, during the first period, drivers suggest that the new BDF from grasslands tended to be a forest completion and consolidation process in areas close to preexisting forested areas, between a range of temperature values, and with humid conditions, shady conditions, steeper slopes, and a lower number of holdings.

During the second period, the drivers in the Supramediterranean and Southern Mesomediterranean BRs followed a similar order of importance, with a slight increase in the number of holdings in the Southern Mesomediterranean (Figure S7a). Conversely, the importance of distance to forests decreased clearly in the Montane BR. The solar radiation driver disappeared (indicating no solar radiation constraints), and the importance of slope increased, with a higher probability of NF occurrence at lower slope values and lower probability at higher slope values (Figure S7a3). Moreover, the soil erosion driver disappeared in the Coline BR, while the importance of solar radiation increased. Therefore, NF continued expanding in the second period beyond the limits of previously forested areas, mainly on steeper slopes (at higher altitudes), close to provincial capitals with temperature-altitudinal restrictions, and shadier conditions.

For the shrublands source category in the Montane BR during the first period, the most important drivers were precipitation (+grad) and temperature (-grad), and secondarily the distance to forests (-grad), distance to urban settlements (+grad), and solar radiation (-grad). Conversely, the Supramediterranean BR denotes the opposite driver order, with the distance to forests (-grad) and to provincial capitals (V-shape) being the main drivers, followed by temperature (A-shape), precipitation (-grad), and drought episodes (-grad). Consequently, the drivers suggested that the new BDF derived from shrublands was a forest consolidation process principally located in humid, temperature-restricted areas with shady conditions.

In the second period, the main drivers in the Montane BR were similar to those in the first period, with temperature (-grad) and precipitation (+grad) increasing in their importance (Figure S7a4). However, two new drivers of lower importance were found: the distance to the hydrographic network (V-shape) and the distance to provincial capitals (V-shape). Conversely, distance to forests was not found to be a driver. Therefore, new BDF transitions were controlled mainly by water availability (with an increase in the mean value compared to the first period), temperature in locations far from capitals, and the hydrographic network. In the Supramediterranean BR, there was a significant change in the drivers' order, with the main drivers being the temperature (A-shape) (Figure S7a4), precipitation (A-shape) and solar radiation (-grad), and the secondary drivers being the distance to forests (-grad) and the distance to urban settlements (+grad). Thus, precipitation and temperature constraints (as their mean values increased in the second period) determined the new BDF appearance in areas close and far from urban settlements with shady conditions.

3.3.2. Drivers of new broadleaf evergreen forests (BEF)

For the rainfed crops' source categories (herbaceous and woody crops) during the first period when woody crops were the most important source category, the distance to forests (-grad) was the most important driver, followed by precipitation (±grad), slope (+grad), and soil erosion in the Supramediterranean and Southern Mesomediterranean BRs (Table 4 and Figure 5b in the main text). Secondary drivers were temperature, the number of workers in the building (construction) (+grad) and service (+grad) sectors, the distance to the hydrographic network (-grad), the distance to urban settlements (+grad), and the utilized agricultural area (-grad). Thus, the new BEF from rainfed crops was related to areas closer to forests, with higher water availability, steeper slopes, and lower soil erosion. It was also associated with higher rates of the population occupied in building (construction) and service activities. Socioeconomic drivers manifested in land cover change in the Supramediterranean BR, a mountainous area with lower farmland availability. In contrast, the distance and accessibility drivers predominated in the Southern Mesomediterranean BR, with more productive lowland agricultural areas, transitional zones near the border, and high relief nearby.

During the second period, the distance to forests (-grad) (Figure S7b1) and precipitation (±grad) remained as the main factors in the Supramediterranean and Southern Mesomediterranean BRs. Secondary drivers were temperature (-grad), slope (+grad) (Figure S7b1), distance to the hydrographic network (-grad), soil erosion (-grad), solar radiation (A-shape), population density (-grad), workers in the building (construction) sector (+grad), and drought episodes. Therefore, during the second period, new BEF was again associated with areas close to forested areas, with water resources, relatively close to the hydrographic network, and with lower temperature, a moderate slope, lower soil erosion values, and lower population density.

Concerning irrigated crops' source categories, only IWC in the Southern Mesomediterranean BR contained enough locations for modeling during both periods. In the first period, the most important driver was soil erosion (-grad), followed by the distance to forests (-grad), temperature (-grad), distance to secondary roads (+grad), and solar radiation (+grad). Thus, drivers suggest that the new BEF in this BR was a conversion process in areas close to previously forested areas, with lower soil degradation, lower temperature, higher solar radiation, and greater distance from secondary roads.

During the second period, the main driver was the temperature (-grad), which increased in importance compared to the first period, followed by the distance to forests (-grad) (Figure S7b2), slope (-grad), population density (+grad), and soil erosion (-grad), which decreased significantly in importance compared to the first period. Consequently, new BEF from IWC was associated with areas characterized by lower temperatures, proximity to forests, steeper slopes, more densely populated municipalities, and lower soil erosion rates.

Regarding the grasslands source category, during the first period, the distance to forests (-grad) was the main driver in the Supramediterranean and Southern Mesomediterranean BR, the temperature (+grad) in the Montane and the Supramediterranean BRs, and the soil erosion (-grad) and slope (-grad) in the Coline region. Secondary drivers were precipitation and solar radiation in all aforementioned regions, and distance to capitals and livestock units in specific bioclimatic regions. The drivers suggested that new BEF from grasslands occurred close to preexisting forested areas, with relatively low water availability (depending on the BR) and shady conditions (except in Coline), besides being temperature-constrained in higher slope areas farther from capitals.

During the second period, when not enough occurrences appeared in the Coline BR, distance to provincial capitals (-grad) (Figure S7b3) was the main driver in the Southern Mesomediterranean BR, precipitation (-grad) was the main driver in the Montane BR, and temperature (+grad) was the main driver in the Supramediterranean BR (Figure S7b3). In the last two regions, the distance to provincial capitals (±grad) was the second most important driver. Drivers with lower importance included slope (V-shape), distance to the hydrographic network (V-shape), solar radiation (-grad), the number of holdings (-grad), and the number of workers in the industry sector (-grad). Overall, the drivers suggest that the new BEF in the second period was mainly related to distance to provincial capitals, precipitation, temperature, and distance to forests, with different importances and patterns according to the BRs.

Concerning the shrublands source category, in the first period, the distance to forests (-grad) was the main driver in the Supramediterranean and Northern Mesomediterranean BRs, and the second driver after distance to urban settlements in the Montane BR (Figure S7b4). Solar radiation was the main driver in the Southern Mesomediterranean BR and the second in the Northern Mesomediterranean. Moreover, the temperature (±grad), precipitation (±grad), and slope (±grad) were the main secondary drivers in almost all of the BRs. Therefore, the drivers indicated that new BEF from shrublands was a process of forest consolidation in areas near previously forested areas, with shady conditions, within a range of values for temperature and precipitation (varying by BR), far from urban areas, and with lower slopes.

In the second period, temperature (±grad) was the main driver in the Montane and Southern Mesomediterranean BRs, distance to forests (V-shape) was the main driver in the Northern Mesomediterranean, and distance to provincial capitals (-grad) was the main driver in the

Supramediterranean BR. Precipitation (±grad) (Figure S7b4) and solar radiation (-grad) were also significant in all of the BRs. Secondary drivers were the slope (-grad), the distance to the hydrographic network (+grad), and the distance to secondary roads (+grad). Overall, during the second period, the new BEF from shrublands continued expanding in more distant locations from forests, with moderate precipitation, within a limited temperature range, with shady conditions, both close and far from capitals, and with lower slope values.

3.3.3. Drivers of new needleleaf evergreen forests (NEF)

Regarding the rainfed source categories (herbaceous and woody crops), enough occurrences to perform an analysis were located in the Southern Mesomediterranean BR for the RWC category during the first period (Table 4 and Figure 5c in the main text). The main driver was the distance to provincial capitals (+grad), followed by the distance to forests (-grad), the number of holdings (-grad), the distance to main roads (+grad), and soil erosion (-grad). Thus, the drivers suggested that the new NEF derived from rainfed woody crops was a conversion process in areas far from provincial capitals and main roads, close to forests, in areas with a lower number of holdings, and with lower soil erosion rates.

In the second period, new forested locations in the Northern and Southern Mesomediterranean BRs appeared. The main drivers in the Northern Mesomediterranean BR (related to RHC) were the distance to the hydrographic network (-grad) and the distance to provincial capitals (A-shape). In the Southern Mesomediterranean BR (related to RWC), the main drivers were the distance to provincial capitals (V-shape) and the distance to forests (-grad) (Figure S7c1). Secondary drivers were solar radiation (-grad), precipitation (-grad), distance to secondary (-grad), distance to main roads (+grad), and drought episodes (-grad). Consequently, in the second period, new NEF appeared in areas close and far from provincial capitals, near forests and the hydrographic network, and far from secondary roads, with shady conditions, low precipitation, and low drought recurrence events.

For irrigated crops' (herbaceous and woody crops) source categories, enough occurrences for analysis were detected in only the Southern Mesomediterranean BR for woody crops in the first period. In this context, the most relevant driver was soil erosion (-grad), followed by distance to main roads (+grad), the number of holdings (-grad), precipitation (V-shape), and livestock units (+grad). Therefore, the drivers suggest that the new NEF derived from irrigated woody crops was a conversion process in areas with lower soil erosion, greater distance from main roads (Figure S7c2), a lower number of holdings, precipitation constraints, and higher livestock units.

In the second period, new locations in the Northern and Southern Mesomediterranean BRs appeared. Precipitation (V-shape) was the main driver in the Southern Mesomediterranean BR and the secondary driver in the northern region. In the Northern Mesomediterranean BR, the main driver was the distance to provincial capitals (V-shape) in the case of woody crops and the distance to the hydrographic network (A-shape) for herbaceous crops. Secondary drivers were solar radiation, soil erosion, distance to forests, the number of holdings, livestock units, and distance to secondary roads. Consequently, new NEF from irrigated crops was associated with areas characterized by various precipitation patterns, far from provincial capitals, with a hydrographic network and preexisting forests, and with shady and sunny conditions.

Concerning the grasslands source category, in the first period, the distance to forests (-grad) was the main driver in the Montane and Supramediterranean BRs, and the distance to provincial capitals (V-shape) was the main driver in the Southern Mesomediterranean region. Secondary drivers were the distance to urban settlements (+grad), precipitation, solar radiation, slope, temperature, the utilized agriculture area, and protected areas. Therefore, the drivers suggest that new NEF from grasslands occurred close to preexisting forests, far from provincial capitals and urban settlements, and with different precipitation and temperature patterns in BRs, shady conditions, and steeper slopes.

In the second period, the slope (A-shape) (Figure S7c3) was the main driver in the Montane BR, the percentage of workers in the industry sector (A-shape) was the main driver in the Supramediterranean region, and precipitation (A-shape) in the Southern Mesomediterranean.

Furthermore, the distance to forests (±grad) (Figure S7c3) and temperature (±grad) were also significant in all BRs. Secondary drivers included the distance to provincial capitals (±grad), the utilized agriculture area, and solar radiation. Therefore, during the second period, the new NEF from grasslands continued expanding at more distant locations from forests, with moderate temperature and precipitation constraints, with steeper slopes up to a limit (Figure S7c3), relatively close and far from capitals, and with shady conditions.

For the shrublands source category, in the first period, the distance to forests (-grad) was the main driver in the Supramediterranean and Northern Mesomediterranean BRs, and the second driver after the solar radiation (-grad) in the Montane BR. Precipitation was the main driver in the Southern Mesomediterranean BR and the second driver in the Supramediterranean and Northern Mesomediterranean BRs. Secondary drivers included slope, temperature, distances to provincial capitals, urban settlements, the hydrographic network, the main roads, and soil erosion. Therefore, the drivers suggested that the increase of needleleaf evergreen forests from shrublands was a process of forest consolidation in locations close to previously forested areas, with different precipitation and humidity regimes (which varied by BR), with shady conditions (Figure S7c4), with lower slope values, within a range of temperature constraints (differing depending on BR), and far from urban areas.

In the second period, precipitation (±grad) was the main driver in the Southern Mesomediterranean and Thermomediterranean BRs, and the most important driver after distance to forests (-grad) (Figure S7c4) in the Northern Mesomediterranean, Supramediterranean, and Montane BRs. Secondary drivers included slope (-grad), distance to provincial capitals, distance to the hydrographic network, solar radiation (Figure S7c4), temperature, and drought or humid events. Therefore, during the second period, the increase of needleleaf evergreen forests from shrublands occurred in more distant locations from previously forested areas, within a wide variety of precipitation values (depending on the BR), with shady conditions, relatively far from provincial capitals and the hydrographic network, and with low drought recurrence and higher water oversupply events.

(B) Tables section

Table S1. Details of the imagery used for each land cover map (LCM). Dates are in day-month-year format.

LCM	200-030	200-031	200-032	200-033	200-034
1987			15/03/1988	15/03/1988	15/03/1988
	30/4/1987	14/4/1987	14/04/1987	14/04/1987	14/04/1987
	17/6/1987	17/6/1987	14/06/1986	17/06/1987	17/06/1987
	4/8/1987	4/8/1987	21/07/1988	21/07/1987	21/07/1987
	20/8/1987	20/8/1987	14/08/1985	20/08/1987	20/08/1987
	2/9/1986	2/9/1986	23/09/1988	02/09/1986	02/09/1986
2002			08/03/2000	08/03/2000	16/03/2000
	28/4/2001	28/4/2001	12/04/2001	28/04/2001	12/04/2001
	30/5/2001	30/5/2001	30/05/2001	30/05/2001	10/06/2002
	1/7/2001	1/7/2001	28/06/2000	01/07/2001	
	25/7/2001	25/7/2001	17/07/2001	25/07/2001	25/07/2001
			15/08/2000	15/08/2000	21/08/2002
	17/9/2003	11/9/2001	24/09/2000	24/09/2000	01/09/2003
2017	15/3/2017	15/3/2017	15/03/2017	15/03/2017	15/03/2017
	19/4/2018	19/4/2018	19/04/2018	19/04/2018	19/14/2018
	19/6/2017	19/6/2017	30/06/2015	30/06/2015	11/06/2020
	5/7/2017	5/7/2017	18/07/2016	18/07/2016	13/07/2020
	3/8/2016	3/8/2016			
	22/8/2017	22/8/2017	19/08/2016	19/08/2016	14/08/2020
	7/9/2017	7/9/2017	29/09/2019	29/09/2019	1/10/2020

Table S2. LCM-1987 confusion matrix for Landsat scenes 200–030 to 200–034.

	Ground truth samples															
Classified map	**NEF**	**BDF**	**BEF**	**Shl**	**Grl**	**BrS**	**Urb**	**WaB**	**IHC**	**RHC**	**IWC**	**RWC**	**Grh**	**Total**	**CE (%)**	**UA (%)**
(NEF) Needleleaf evergreen forest	**573 842**	996	5 940	8 600	126	80	1	4	3	6	219	21	0	589 839	2.7	97.3
(BDF) Broadleaf deciduous forest	269	**342 174**	1 866	1 443	449	0	0	0	217	20	31	0	0	346 469	1.2	98.8
(BEF) Broadleaf evergreen forest	14 153	3 367	**724 699**	31 566	4 440	0	5	0	9	25	2 434	332	0	781 031	7.2	92.8
(Shl) Shrublands	2 353	711	8 160	**1 224 637**	78 091	784	53	8	43	488	833	1 731	0	1 317 893	7.1	92.9
(Grl) Grasslands	33	3 884	789	65 427	**1 275 094**	14 347	642	2	176	17 068	207	7 291	0	1 384 959	7.9	92.1
(BrS) Bare soils	1	0	0	1 570	4 608	**187 122**	410	10	17	3 320	22	131	0	197 211	5.1	94.9
(Urb) Urban areas and Infrastructures	0	0	0	0	36	131	**31 614**	0	0	2	0	0	0	31 783	0.5	99.5
(WaB) Water bodies	0	2	0	0	0	10	0	**24 141**	7	1	0	0	0	24 161	0.1	99.9
(IHC) Irrigated herbaceous crops	0	287	0	133	245	0	3	1	**241 215**	877	1 206	252	0	244 219	1.2	98.8
(RHC) Rainfed herbaceous crops	0	4	0	3 146	9 275	10 979	196	89	1 983	**2 380 492**	501	14 319	0	2 420 986	1.7	98.3
(IWC) Irrigated woody crops	54	89	526	1 643	928	50	265	0	973	450	**279 723**	43 491	0	328 192	14.8	85.2
(RWC) Rainfed woody crops	0	2	31	1 105	5 362	9 561	2 847	2	137	27 225	11 007	**1 067 022**	2	1 124 302	5.1	94.9
(Grh) Greenhouses	0	0	0	0	0	20	77	0	0	2	0	1	**11 378**	11 478	0.9	99.1
NoData	6	0	0	199	6	80	0	0	0	0	0	0	0	291		
Total	590 711	351 518	742 010	1 339 469	1 378 659	223 165	36 114	24 257	244 780	2 429 976	296 185	1 134 592	11 380	8 802 814	OA = 95.0%	
OE (%)	2.9	2.7	2.3	8.6	7.5	16.2	12.5	0.5	1.5	2.0	5.6	6.0	0.0		OAw = 97.2%	
PA (%)	97.1	97.3	97.7	91.4	92.5	83.8	87.5	99.5	98.5	98.0	94.4	94.0	100.0		k = 0.9	

OE: omission error, CE: commission error, PA: producer's accuracy, UA: user's accuracy, *k*: kappa index of agreement, OA: overall accuracy, OAw: overall accuracy weighted by the ground truth area considering only classified pixels (unclassified pixels are not considered errors and are solved by spatial proximity). Figures in bold represent locations where the predicted and the true labels match.

Table S3. LCM-2002 confusion matrix for Landsat scenes 200–030 to 200–034.

	Ground truth samples																
Classified map	**NEF**	**BDF**	**BEF**	**Shl**	**Grl**	**BrS**	**Urb**	**WaB**	**IHC**	**RHC**	**IWC**	**RWC**	**RiC**	**Grh**	**Total**	**CE (%)**	**UA (%)**
(NEF) Needleleaf evergreen forest	**823 747**	3 903	19 314	24 536	3 720	0	2	0	16	27	742	45	0	0	876 052	6.0	94.0
(BDF) Broadleaf deciduous forest	5 863	**482 145**	24 327	13 142	10 049	0	0	0	833	13	97	0	0	0	536 468	10.1	89.9
(BEF) Broadleaf evergreen forest	76 891	7 260	**1 155 685**	207 405	36 961	0	5	0	170	597	7 014	2 032	0	0	1 494 020	22.6	77.4
(Shl) Shrublands	28 787	1 902	72 707	**1 460 560**	318 082	1 965	85	6	71	1 746	4 437	5 356	0	0	1 895 703	23.0	77.0
(Grl) Grasslands	2 738	2 008	12 320	164 349	**1 518 469**	33 801	1 333	0	352	93 428	432	21 685	0	1	1 850 916	18.0	82.0
(BrS) Bare soils	8	0	0	26 615	20 967	**199 711**	2 048	180	34	19 252	41	843	0	305	270 004	26.0	74.0
(Urb) Urban areas and Infrastructures	0	0	0	12	28	175	**50 776**	0	0	13	0	20	0	1	51 025	0.5	99.5
(WaB) Water bodies	0	2	0	0	28	28	0	**44 105**	28	10	0	0	0	0	44 201	0.2	99.8
(IHC) Irrigated herbaceous crops	2	1 111	44	128	708	6	0	0	**252 189**	1 333	1 449	2	127	0	257 097	1.9	98.1
(RHC) Rainfed herbaceous crops	63	56	87	5 865	44 110	29 305	2 056	55	2 747	**3 594 131**	102	11 939	6	30	3 690 554	2.6	97.4
(IWC) Irrigated woody crops	1 192	391	6 641	19 559	5 698	45	153	0	3 005	1 436	**378 139**	85 307	0	0	501 565	24.6	75.4
(RWC) Rainfed woody crops	1 114	11	796	8 990	45 651	11 544	13 893	0	3 167	91 239	42 435	**1 610 655**	0	20	1 829 515	12.0	88.0
(RiC) Rice crops	0	2	0	0	0	0	0	0	372	2	0	0	**6 685**	0	7 062	5.3	94.7
(Grh) Greenhouses	0	0	0	0	0	56	178	0	0	3	0	0	0	**29 640**	29 878	0.8	99.2
NoData	663	3 481	383	3 897	1 565	38 624	0	0	0	0	0	0	0	0	48 613		
Total	941 067	502 273	1 292 303	1 935 059	2 006 037	315 261	70 527	44 345	262 983	3 803 228	434 887	1 737 886	6 819	29 998	13 382 673	OA = 86.7%	
OE (%)	12.5	4.0	10.6	24.5	24.3	36.7	28.0	0.5	4.1	5.5	13.0	7.3	2.0	1.2		OAw = 92.2%	
PA (%)	87.5	96.0	89.4	75.5	75.7	63.3	72.0	99.5	95.9	94.5	87.0	92.7	98.0	98.8		k = 0.9	

OE: omission error, CE: commission error, PA: producer's accuracy, UA: user's accuracy, *k*: kappa index of agreement, OA: overall accuracy, OAw: overall accuracy weighted by the ground truth area considering only classified pixels (unclassified pixels are not considered errors and are solved by spatial proximity). Figures in bold represent locations where the predicted and the true labels match.

Table S4. LCM-2017 confusion matrix for Landsat scenes 200–030 to 200–034.

	Ground truth samples																
Classified map	**NEF**	**BDF**	**BEF**	**Shl**	**Grl**	**BrS**	**Urb**	**WaB**	**IHC**	**RHC**	**IWC**	**RWC**	**RiC**	**Grh**	**Total**	**CE (%)**	**UA (%)**
(NEF) Needleleaf evergreen forest	**715 557**	802	5 953	6 276	33	0	0	1	20	2	171	4	0	0	728 817	1.8	98.2
(BDF) Broadleaf deciduous forest	257	**424 599**	1 739	2 272	2 403	0	0	0	358	39	24	1	0	0	431 692	1.6	98.4
(BEF) Broadleaf evergreen forest	9 000	1 027	**866 533**	64 636	2 841	0	0	1	24	32	2 391	321	0	0	946 805	8.5	91.5
(Shl) Shrublands	4 483	595	28 744	**1 195 923**	76 336	489	8	1	24	193	1 712	954	3	0	1 309 465	8.7	91.3
(Grl) Grasslands	264	820	3 490	67 951	**1 138 744**	13 115	498	11	327	18 274	598	13 289	0	2	1 257 383	9.4	90.6
(BrS) Bare soils	0	15	0	1 050	2 086	**191 240**	835	17	20	5 101	19	415	0	55	200 853	4.8	95.2
(Urb) Urban areas and Infrastructures	0	0	0	41	22	2 955	**67 573**	0	0	52	1	36	0	1	70 680	4.4	95.6
(WaB) Water bodies	0	87	0	0	5	11	2	**30 809**	67	0	2	0	0	0	30 984	0.6	99.4
(IHC) Irrigated herbaceous crops	0	168	0	25	219	0	6	0	**149 625**	951	447	6	50	0	151 496	1.2	98.8
(RHC) Rainfed herbaceous crops	0	18	0	2 519	4 182	8 933	196	11	1 322	**2 382 902**	279	6 830	0	2	2 407 194	1.0	99.0
(IWC) Irrigated woody crops	22	138	656	1 304	657	45	28	0	967	282	**341 843**	19 136	0	0	365 078	6.4	93.6
(RWC) Rainfed woody crops	1	3	11	820	2 047	2 318	2 817	0	163	32 916	24 966	**1 213 489**	0	2	1 279 555	5.2	94.8
(RiC) Rice crops	0	18	0	0	0	0	0	0	130	1	2	0	**3 227**	0	3 377	4.5	95.5
(Grh) Greenhouses	0	0	0	0	0	11	117	0	0	3	0	0	0	**22 809**	22 940	0.6	99.4
NoData	199	45	84	74	44	80	0	0	0	0	0	0	0	0	525		
Total	729 781	428 335	907 212	1 342 891	1 229 620	219 199	72 080	30 852	153 045	2 440 747	372 455	1 254 480	3 279	22 870	9 206 845	OA = 95.0%	
OE (%)	1.9	0.9	4.5	10.9	7.4	12.8	6.3	0.1	2.2	2.4	8.2	3.3	1.6	0.3		OAw = 96.9%	
PA (%)	98.1	99.1	95.5	89.1	92.6	87.2	93.7	99.9	97.8	97.6	91.8	96.7	98.4	99.7		k = 0.9	

OE: omission error, CE: commission error, PA: producer's accuracy, UA: user's accuracy, *k*: kappa index of agreement, OA: overall accuracy, OAw: overall accuracy weighted by the ground truth area considering only classified pixels (unclassified pixels are not considered errors and are solved by spatial proximity). Figures in bold represent locations where the predicted and the true labels match.

Table S5. Soil loss' categorical value ranges and interpretation descriptions. Units of soil loss: Mg/(ha × year)

Soil loss	Interpretation
0	Areas not susceptible to the erosive process, such as urban spaces, roads, reservoirs, etc.
0 – 5	Areas with very low levels of erosion and tolerable soil loss. There is no net erosion.
5 – 10	Areas with low levels of erosion and soil loss that may be tolerable. There is probably no net erosion.
10 – 25	Areas with mild erosive processes. There is erosion, although it is not visible to the naked eye.
25 – 50	Areas with moderate erosive processes. There is erosion, although it may not be visible to the naked eye.
50 – 100	Areas with severe erosive processes. There is erosion, and it is visible to the naked eye.
100 – 200	Areas with very serious erosive processes. There is erosion, and it is evident to the naked eye.
> 200	Areas with extreme erosive processes. There is erosion, and it is obvious to the naked eye.

Table S6a. The number of presences after applying uncertainty restrictions. The bioclimatic regions' (BR) numbers assignation is detailed in Figure S4.

Bio. region	BR.1		BR.2		BR.3		BR.4		BR.5		BR.6		BR.7		BR.8	
NF dynamic	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
RHC-BDF [1]	0	0	204	2716	1	339	0	0	1977	12646	725	3986	2	51	173	1226
RHC-BEF [1]	0	0	84	308	2	42	0	0	2063	16155	1663	15504	12	5	79	359
RHC-NEF [1]	0	0	97	184	1	389	0	0	52	448	239	776	0	22	141	1365
RWC-BDF [1]	0	0	0	1	0	0	0	0	150	1327	371	2780	16	215	7	261
RWC-BEF [1]	0	0	0	0	0	0	0	0	2620	7439	21615	72939	71	170	10	53
RWC-NEF [1]	0	0	0	0	0	3	0	0	66	94	1049	2368	213	214	59	312
IHC-BDF [1]	0	0	308	122	3360	1486	0	0	3271	3658	5307	8300	85	337	1290	4557
IHC-BEF [1]	0	0	23	33	118	11	0	0	121	307	1364	1589	20	0	75	257
IHC-NEF [1]	0	0	25	8	231	1859	0	0	13	14	328	879	18	5	49	2548
IWC-BDF [1]	0	0	170	114	1109	953	0	0	1185	2273	3967	9632	493	8250	4222	7810
IWC-BEF [1]	0	0	98	115	50	26	0	0	7336	4545	65198	52404	646	321	553	342
IWC-NEF [1]	0	0	6	22	45	223	0	0	167	47	3819	4138	674	532	382	1569
Grl-BDF [2]	29	275	31328	37624	11612	3505	136	77	14310	35806	1067	3323	0	35	326	476
Grl-BEF [2]	3	80	2061	1054	1293	302	389	194	45600	66619	72322	157122	131	37	169	200
Grl-NEF [2]	218	499	6391	2871	4819	8607	316	73	1127	1686	2010	4452	206	296	166	481
Shl-BDF [2]	0	0	2397	3465	275	538	4	1	21458	20335	841	417	1	1	196	418
Shl-BEF [2]	0	2	11311	16234	130	83	57	325	334428	292334	129296	84458	165	164	3204	4905
Shl-NEF [2]	1	0	3974	6028	508	1760	355	599	74017	65752	26391	44419	760	2384	12198	27249

[1] Presences given an *uncertainty* < 0.3 (crop categories). [2] Presences given an *uncertainty* < 0.1 (Shl and Grl).

Table S6b. The number of presence-absence locations used for modeling after applying uncertainty and spatial autocorrelation restrictions. Presence-absences were extracted within BRs, using absence points with a steady thematic temporal pattern. Grey highlights indicate which NF dynamics were analyzed. The figures refer to the number of sites (equal number of presence and absence) available for analysis. The bioclimatic regions' (BR) numbers assignation is detailed in Figure S4.

Bio. region	BR.1		BR.2		BR.3		BR.4		BR.5		BR.6		BR.7		BR.8	
NF dynamic	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
RHC-BDF [1]	0	0	61	316	1	170	0	0	1059	4807	371	1730	1	26	94	613
RHC-BEF [1]	0	0	42	154	1	21	0	0	1168	5000	853	5000	6	3	40	180
RHC-NEF [1]	0	0	49	92	1	195	0	0	26	224	120	388	0	11	69	600
RWC-BDF [1]	0	0	0	1	0	0	0	0	15	97	198	1315	8	108	4	131
RWC-BEF [1]	0	0	0	0	0	0	0	0	450	921	5000	5000	36	85	5	27
RWC-NEF [1]	0	0	0	0	0	2	0	0	33	47	503	1141	107	107	30	156
IHC-BDF [1]	0	0	154	61	562	706	0	0	159	225	1733	2197	43	169	489	1691
IHC-BEF [1]	0	0	12	17	59	6	0	0	61	154	477	545	10	0	38	129
IHC-NEF [1]	0	0	13	4	51	201	0	0	7	7	164	440	9	3	25	847
IWC-BDF [1]	0	0	85	57	555	477	0	0	86	253	1808	4329	207	1972	1421	3239
IWC-BEF [1]	0	0	49	58	25	13	0	0	342	313	5000	5000	323	161	277	171
IWC-NEF [1]	0	0	3	11	23	112	0	0	84	24	1536	2052	337	266	152	789
Grl-BDF [2]	15	138	4729	5000	950	1119	68	39	5000	5000	1025	3132	0	18	163	238
Grl-BEF [2]	2	40	1867	910	1228	294	195	97	5000	5000	5000	5000	66	19	85	100
Grl-NEF [2]	109	250	2129	2207	411	346	158	37	1113	1644	1928	4179	103	148	83	241
Shl-BDF [2]	0	0	1202	1611	138	269	2	1	5000	5000	421	209	1	1	98	209
Shl-BEF [2]	0	1	5000	5000	65	42	29	163	5000	5000	5000	5000	83	82	3169	4876
Shl-NEF [2]	1	0	3739	5000	51	65	178	300	5000	5000	5000	5000	378	1079	5000	5000

[1] Presences given an *uncertainty* of < 0.3 (crop categories). [2] Presences given an *uncertainty* of < 0.1 (Shl and Grl).

Table S7. Models' results by source category and forest group. The main drivers' relative importances and related PDP patterns are provided. For the sake of simplicity, only BDF modeling results are shown. The bioclimatic regions' (BR) number equivalence is described in Figure S4.

BDF - Period 1	RHC						RWC		IHC						IWC						Grl								Shl			
Variable \ BR.	BR.5		BR.6		BR.8		BR.6		BR.3		BR.6		BR.8		BR.6		BR.7		BR.8		BR.2		BR.3		BR.5		BR.6		BR.2		BR.5	
Slope									13.4	+			8.3	+	17.6	V					28.3	A	19.3	+								
General Curvature									28	-									6.2	-												
Pot_Rad_Wint																					10.9	-	19.6	-	20.8	-			11.2	-		
Ac_Rain	21.9	и									13.4	A									10	-			17	+	26.7	+	38.5	+	17.7	A
Av_Me_Temp															25.1	-					20.4	A			11.7	A	37.5	-	19.2	-	18.2	A
DE6_S6																															12.9	-
DE9_S6																																
DE3_S12																																
DE3_S24																																
HE3_S12																																
Eu_Dist_Hyd	20.6	-									17	-	56.8	-	25	-			49.2	-												
Co_Dist_M_Roads											13.8	+	9	=	12.6	+											9.7	-				
Co_Dist_S_Roads													9.9	=					7.3	+												
Co_Dist_Urb	12.7	-							22.1	+															17.5	-			15.3	+		
Co_Dist_Cap	18.7	+									36.3	+							18.5	V			22.2	-							18.9	-
Eu_Dist_Forest	26	-							20.4	-	19.6	-	16	-	19.6	-			18.9	-	30.4	-	15.3	-	33	-	11.3	-	15.8	-	32.3	-
Soil_Erosion									16.1	-													23.6	-								
Protected_Areas																																
Pop_Density																																
W_Building																																
W_Services																																
W_Industry																																
Num_Hold																											14.8	-				
LSU																																
UAA																																
Model accuracy	0.84								0.83		0.88		0.86		0.85				0.85		0.85		0.81		0.79		0.94		0.87		0.8	
BDF - Period 2	RHC						RWC		IHC						IWC						Grl								Shl			
Variable \ BR.	BR.5		BR.6		BR.8		BR.6		BR.3		BR.6		BR.8		BR.6		BR.7		BR.8		BR.2		BR.3		BR.5		BR.6		BR.2		BR.5	
Slope									14.4	+			8	+	22.1	-	7.1	-			22.9	A	20.1	V								
General Curvature			8.1	V					32.5	V			6.3	V					6	-												
Pot_Rad_Wint									17.9	V													26.2	-	21.2	-			11.1	-	18.1	-
Ac_Rain	18.9	N	23.3	+			15.4	-			17.6	+					5.7	=			19.9	-			18.8	A			24.7	+	22.7	A
Av_Me_Temp			15.2	V			12.8	-							19.8	-	73.3	+			19.9	A					40.8	-	35	-	26.9	A
DE6_S6																																
DE9_S6																																
DE3_S12																																
DE3_S24																																
HE3_S12																																
Eu_Dist_Hyd	18.4	-	22.9	-	38.7	-	18.7	-			18.1	-	46.7	-	26.8	V	9.4	-	41.1	-					18.1	-	15.6	-	14.9	V		
Co_Dist_M_Roads											12.9	A																				
Co_Dist_S_Roads					12.3	-																										
Co_Dist_Urb	14.9	-			20.1	-			21.1	+													18.8	+			12.2	-			15.6	V
Co_Dist_Cap	22.1	V			16	+					34.2	+	20.7	+					19.8	+	21.1	+	20.6	-	19.5	-			14.4	+		
Eu_Dist_Forest	25.6	-	30.6	-	12.9	-	14.4	-			17.2	-	18.4	V	17.2	-	4.5	=	21.6	-	16.2	-			22.4	-	11.8	-			16.7	-
Soil_Erosion									14.3	-					14.2	-							14.3	-								
Protected_Areas																																
Pop_Density																			11.4	-												
W_Building																																
W_Services																																
W_Industry																																
Num_Hold							38.7	-																			19.7	-				
LSU																																
UAA																																
Model accuracy	0.79		0.83		0.85		0.9		0.77		0.77		0.8		0.78		0.95		0.79		0.75		0.77		0.73		0.9		0.85		0.73	

(C) Figures section

Figure S4. Climatic segmentation of the Iberian Peninsula in bioclimatic regions (BRs) and the spatial context of the transect. Due to the reduced spatial representativeness of the transect, the Alpine and Subalpine regions were aggregated into region 1, and the Criomediterranean and Oromediterranean regions in the Sierra Nevada mountain range were aggregated into region 4. Furthermore, the extensive Mesomediterranean region was subdivided into two parts, separated by the Iberian System mountain range; thus, the Southern part is represented by region 6 while the Northern part is represented by region 8.

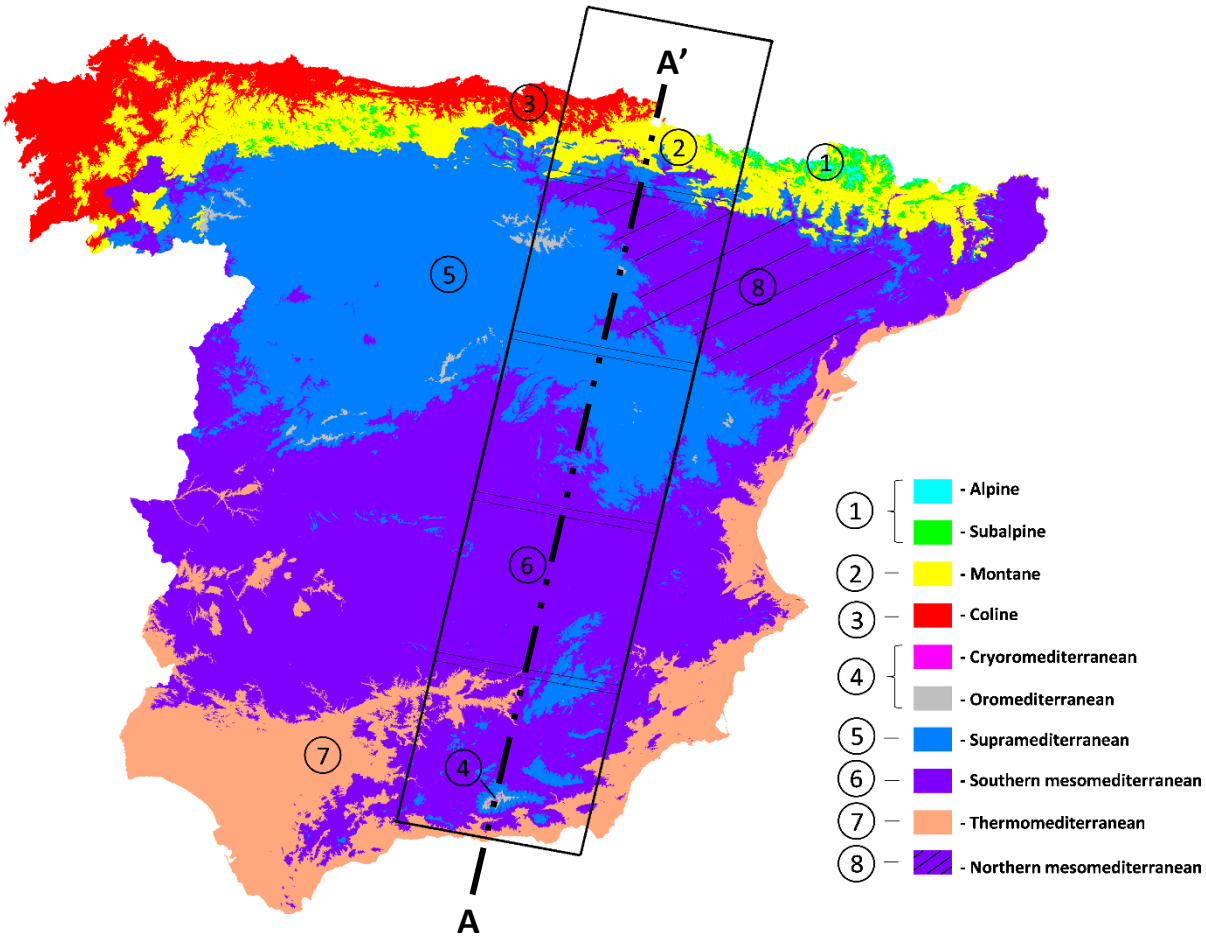


Figure S5. Profiles of altitude, average mean temperature, average accumulated annual precipitation, and potential solar radiation, as measured along the transect line (**A-A'**) depicted in Figure S4. Note the well-known and high correlation between altitude and average mean temperature.

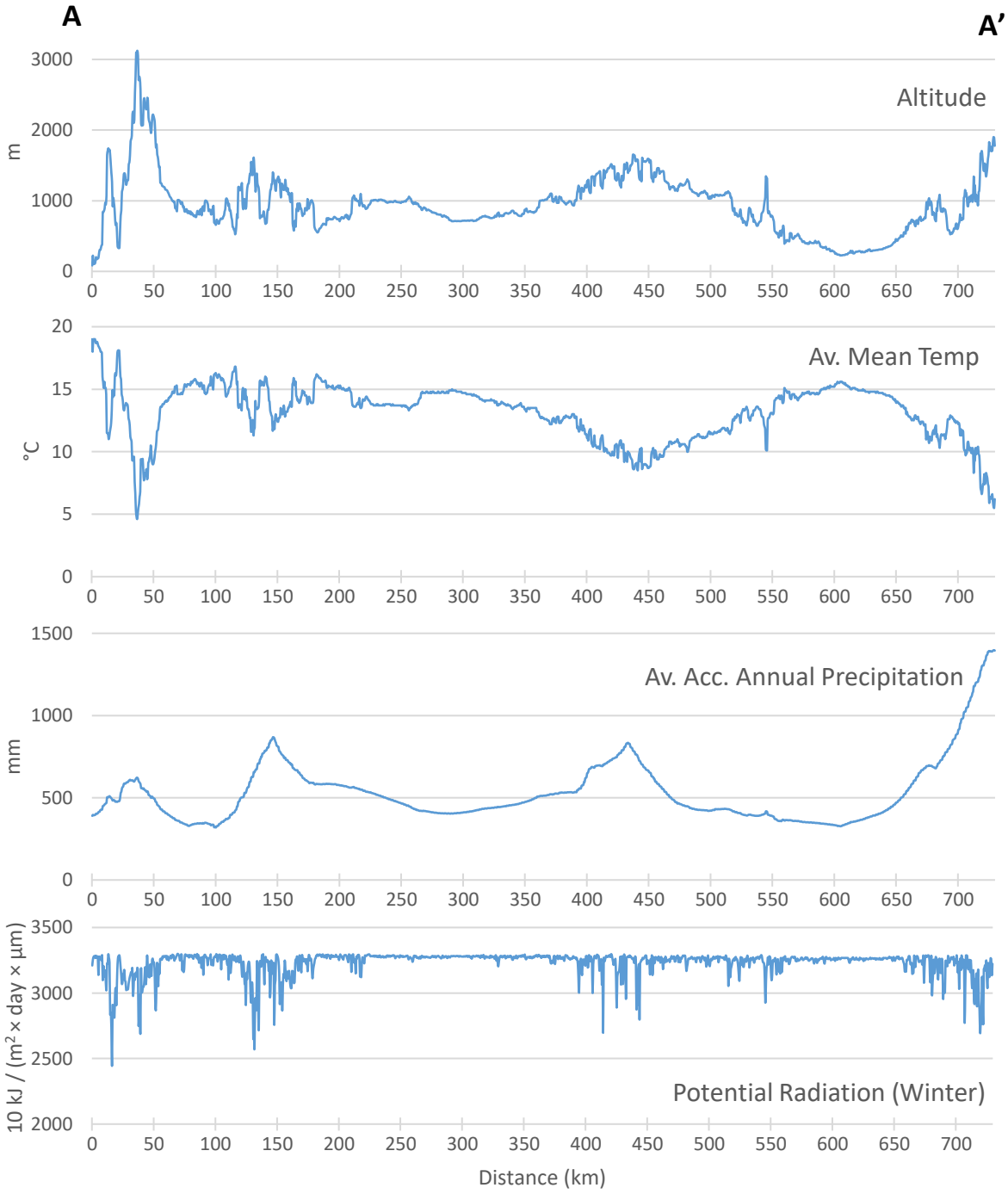


Figure S6. New forest dynamics disaggregated by bioclimatic region (BR). In (a), a histogram of the frequency of the categories evolving to NF in three temporal moments is shown. The chord diagram between 1987–2002 (b) and between 2002–2017 (c) are represented. The numbers on the chord diagrams represent the absolute frequency of each category disaggregated by BR.

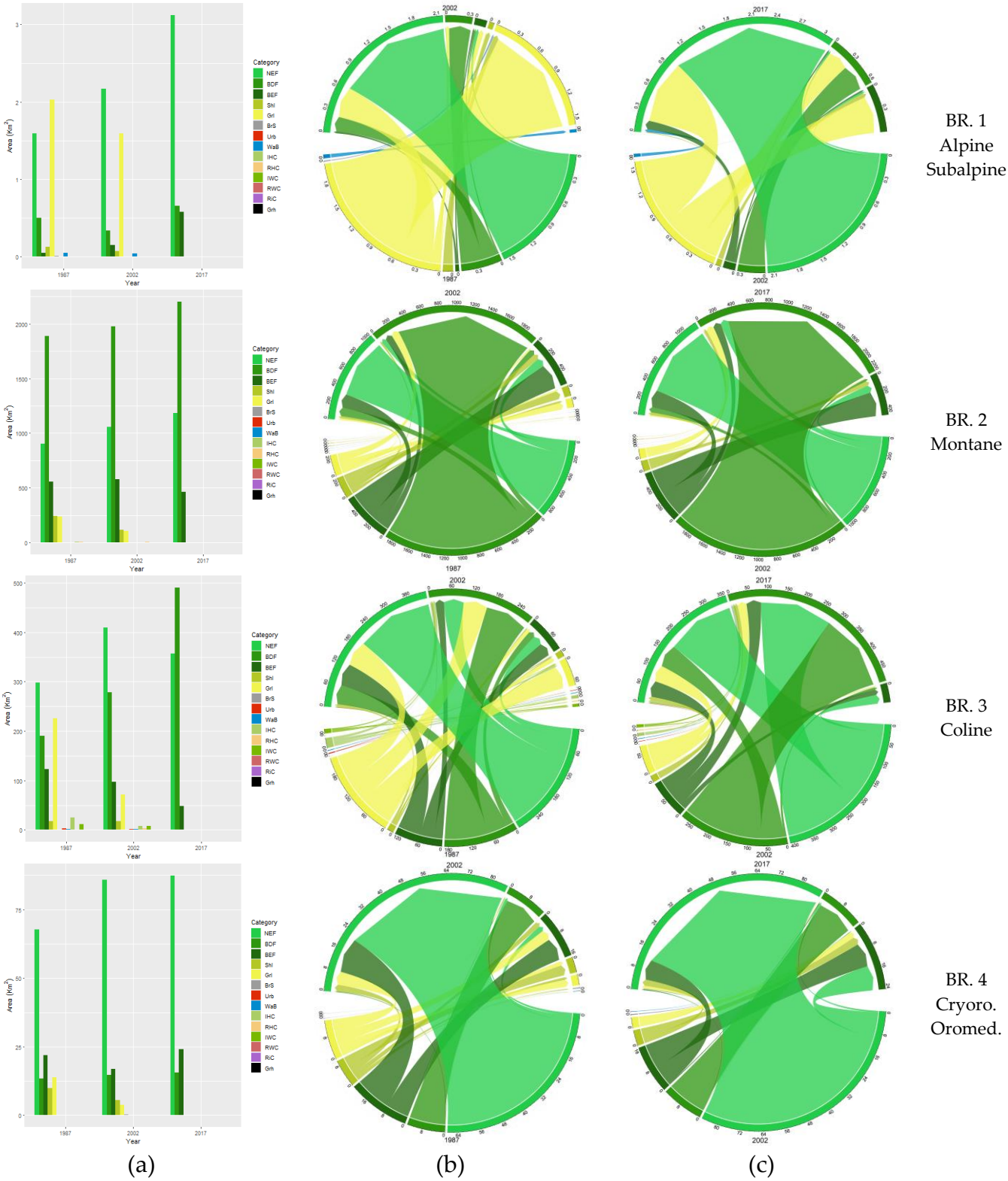


Figure S6. New forest dynamics disaggregated by BR (Figure continued).

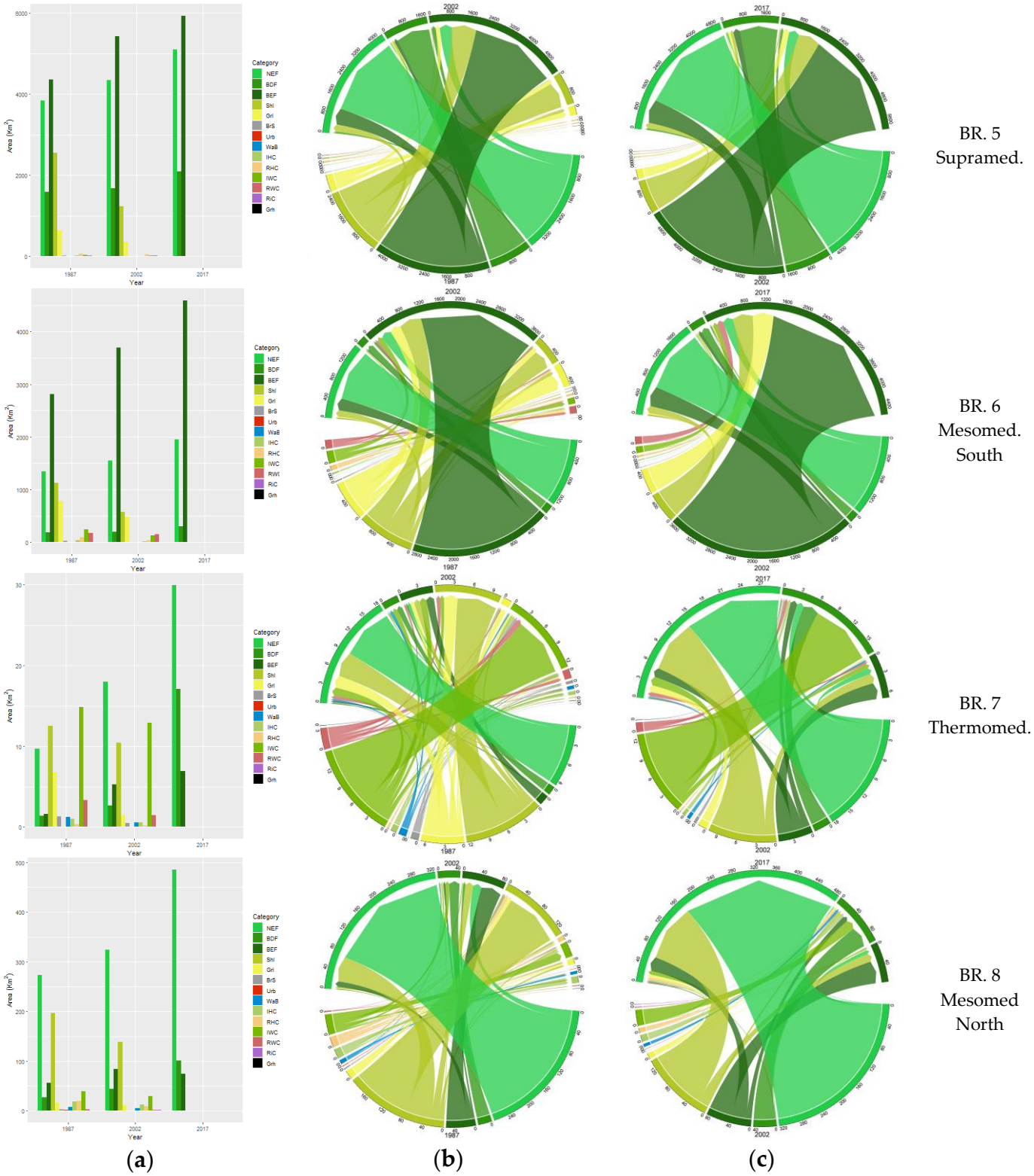


Figure S7a. Examples of representative dependence plots for the new BDF source categories. Blue and red colors correspond to the first (1987–2002) and second (2002–2017) periods. The density of occurrences is represented by vertical lines on the x-axes, with black rug lines denoting the 5%, 50%, and 95% percentiles.

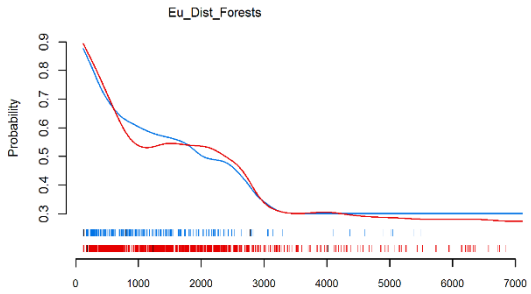
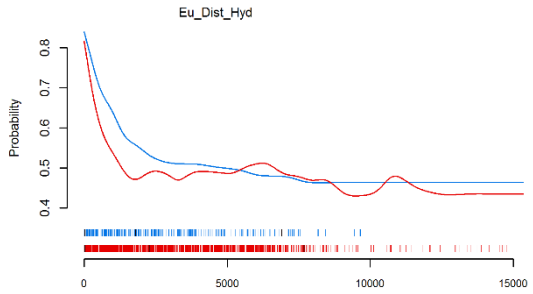
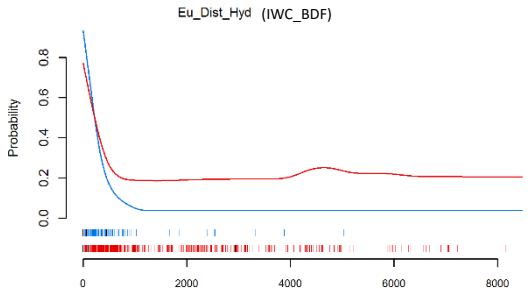
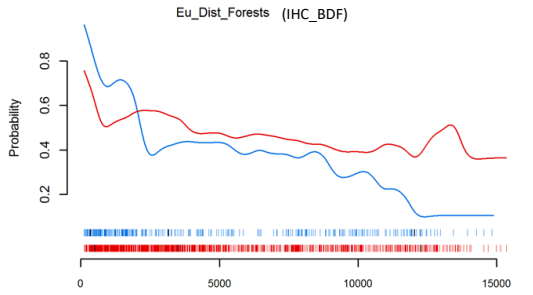
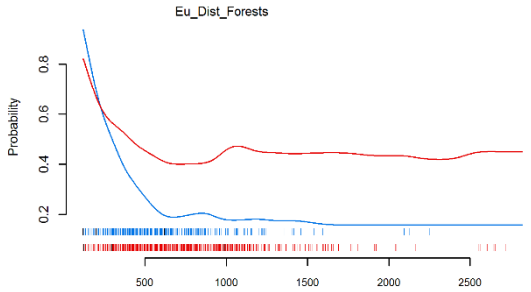
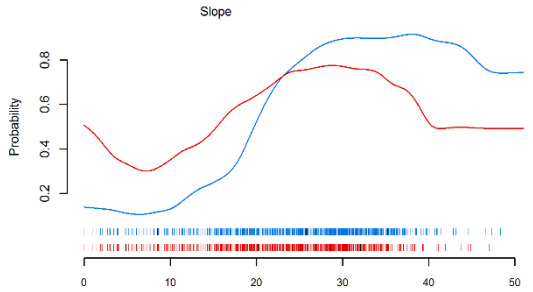
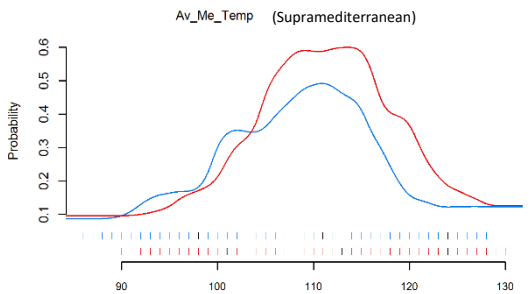
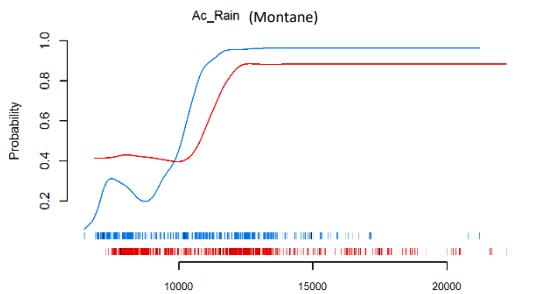
NF dynamic	Dependence plots	
1 RHC-BDF Supramediterranean		
2 IWC/IHC-BDF Northern Mesomediterranean		
3 Grl-BDF Montane		
4 Shl-BDF Montane/Supramediterranean		

Figure S7b. Examples of representative dependence plots for the new BEF source categories. Blue and red colors correspond to the first (1987–2002) and second (2002–2017) periods. The density of occurrences is represented by vertical lines on the x-axes, with black rug lines denoting the 5%, 50%, and 95% percentiles.

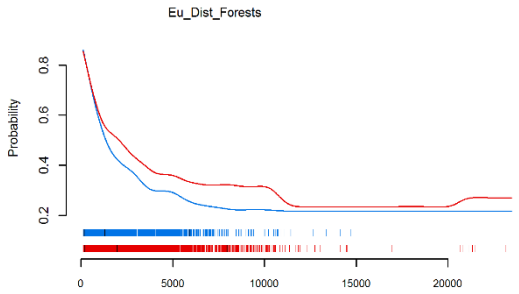
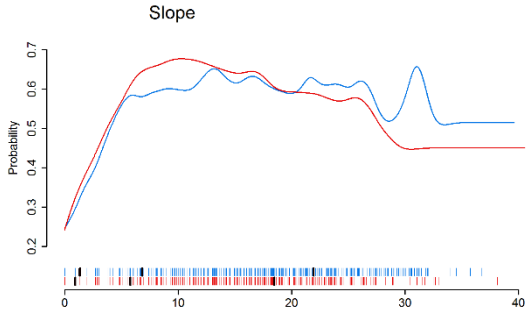
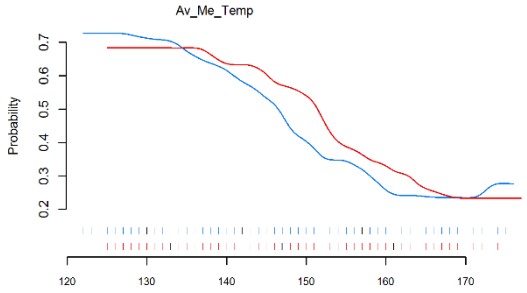
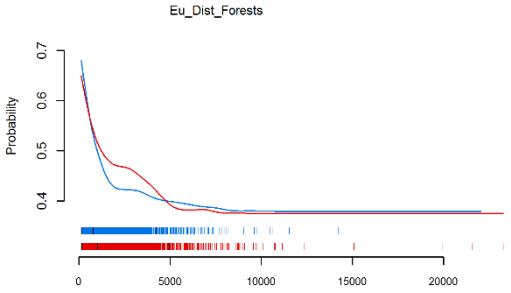
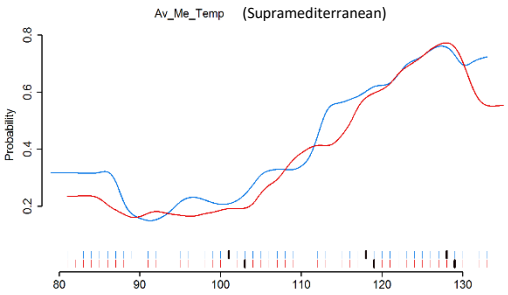
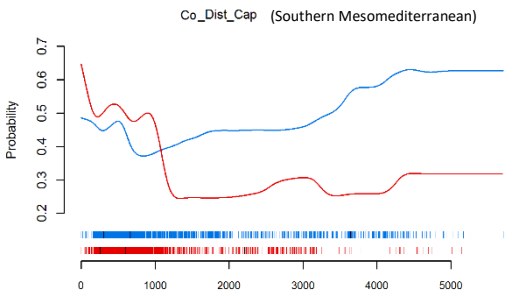
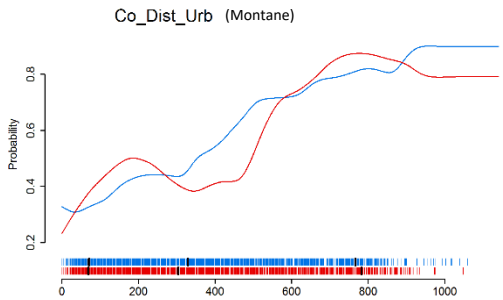
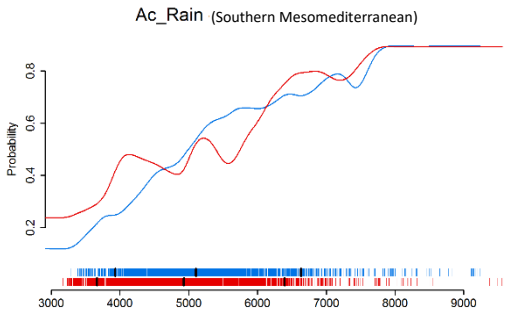
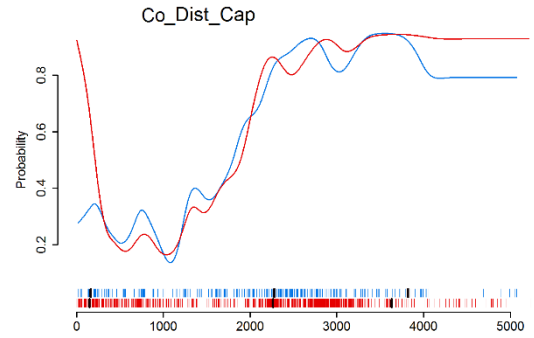
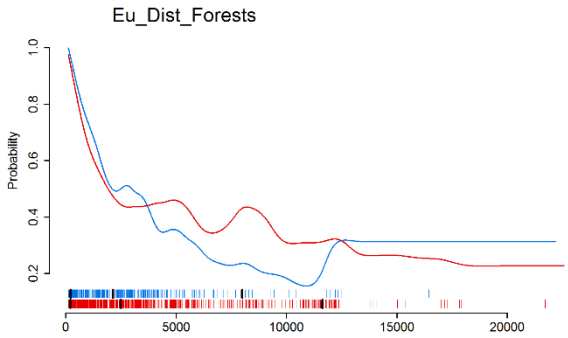
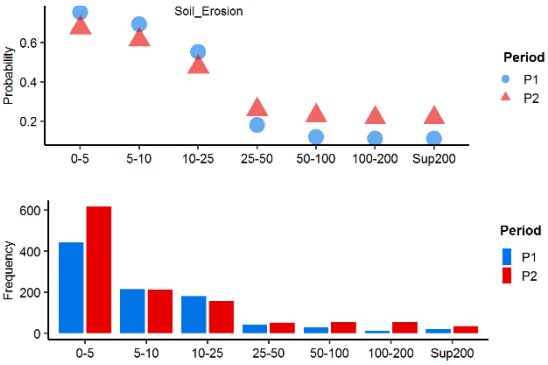
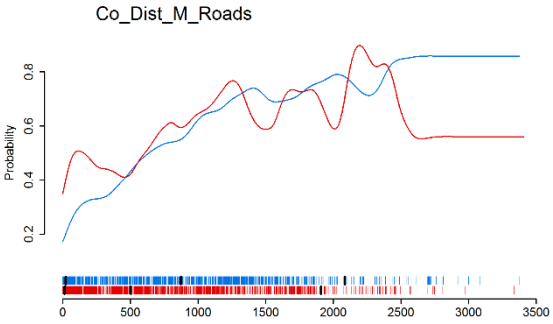
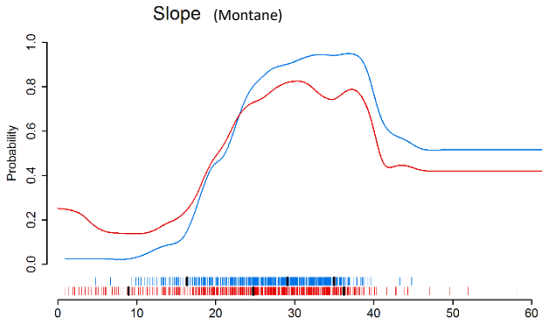
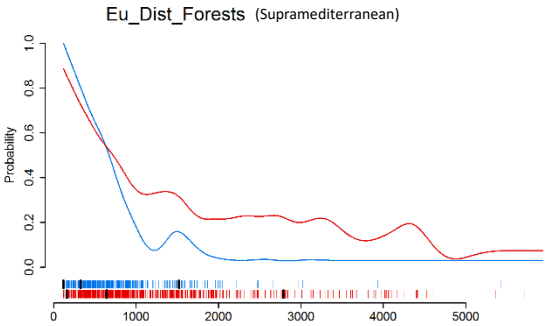
NF dynamic	Dependence plots	
1 RWC-BEF Southern Mesomediterranean		
2 IWC-BEF Southern Mesomediterranean		
3 Grl-BEF Supramediterranean/Southern Mesomediterranean		
4 Shl-BEF Montane/Southern Mesomediterranean		

Figure S7c. Examples of representative dependence plots for the new NEF source categories. Blue and red colors correspond to the first (1987–2002) and second (2002–2017) periods. The density of occurrences is represented by vertical lines on the x-axes, with black rug lines denoting the 5%, 50%, and 95% percentiles.

NF dynamic	Dependence plots	
1 RWC-NEF Southern Mesomediterranean		
2 IWC-NEF Southern Mesomediterranean		
3 Grl-NEF Montane/ Supramediterranean		
4 Shl-NEF Montane/ Supramediterranean