Ghallab A, Boubekraoui H, Laaribya S, Ben-Said M (2024). **Potential natural vegetation pattern based on major tree distribution modelling in the western Rif of Morocco**
iForest – Biogeosciences and Forestry – doi: 10.3832/ifor4602-017

Supplementary Material

Tab. S1 - The seventeen environmental variables used in the species distribution modelling in the present study. DEM: digital elevation model.

Group	Code	Variables	Range and unit	Data resolution	References / creation
Bioclimatic variables	**Pmoy**	Mean annual precipitation	272 - 1852 mm	50 m	Ghallab & Taiqui (2015)
	Tmoy	Mean annual temperature	12.72 - 19.28 °C	50 m	Ghallab & Taiqui (2015)
	Tmin	Mean minimum temperature of the coldest month	(-1.84) - 10.32 °C	50 m	Ghallab & Taiqui 2015)
	Tmax	Mean maximum temperature of the hottest month	28.43 - 37.59 °C	50 m	Ghallab & Taiqui (2015)
	Q2	Emberger's pluviothermic quotient	66 - 189	50 m	Derived using the formula: $2000 \cdot Pmoy/(Tmax^2-Tmin^2)$
		Soil variables			
	Litho	Lithology	Siliceous, argillaceous carbonates, and basic	50 m	Derived from geological maps
	TVDI	Temperature Vegetation Dryness Index	0 - 1	50 m	Derived using Sandholt et al. (2002) empirical parameterization (Landsat 8 image)
Topographic variables	**Alti**	Altitude	0 - 2140 m	50 m	Derived from DEM raster
	pente	Slope	0 - 82 %	50 m	Created from DEM (ArcGIS-spatial analyst tools)
	expos	Aspect	0° - 360 °	50 m	Created from DEM (ArcGIS-spatial analyst tools)
	Courb	Curvature	(-2) - 2	50 m	Created from DEM (ArcGIS-spatial analyst tools)
Luminosity and NDVI	**RadSola**	Solar radiation	1160 to 2140 kWH m^{-2} y^{-1}	50 m	Created from DEM (ArcGIS-spatial analyst tools) McCune & Keon (2002)
	NDVI	Normalized Difference Vegetation Index	(-0.6) - 0.7	50 m	Derived from Landsat 8 image
Geographic variables	**coordX**	Longitude	422931 - 585878 m	50 m	Generated using ArcGIS Add Data from the XY Coordinates tool
	coordY	Latitude	434431 - 591183 m	50 m	Generated using ArcGIS Add Data from the XY Coordinates tool
	DistMed	Distance from the Mediterranean Sea	0 - 99424 m	50 m	Generated using ArcGIS Spatial Analyst tool
	DistOce	Distance from the Atlantic Ocean	0 - 150090 m	50 m	Generated using ArcGIS Spatial Analyst tool

Tab. S2 - Pearson correlation matrix for the ten variables selected in the modeling process. Strong correlations (> 0.70) are shown in bold. See Tab. 2 for abbreviations.

	Alti	**coordx**	**coordy**	**Courb**	**DistMed**	**DistOce**	**Expos**	**Litho**	**NDVI**	**Pente**	**Pmoy**	**Q2**	**RadSola**	**Tmin**	**Tmoy**	**TVDI**	**Tmax**
Alti	1.00	0.62	-0.23	0.06	-0.23	0.61	0.13	0.05	0.34	0.54	**0.80**	**0.75**	0.39	**-0.75**	**-0.82**	-0.26	-0.22
coordx		1.00	-0.27	0.00	-0.44	**0.96**	0.12	0.03	0.00	0.55	0.33	0.44	0.07	-0.28	**-0.70**	0.10	-0.46
coordy			1.00	0.00	**-0.73**	-0.53	-0.03	-0.23	0.09	-0.07	-0.25	0.05	-0.15	0.45	0.07	0.09	-0.65
Courb				1.00	0.00	0.00	0.00	0.00	-0.02	0.03	0.04	-0.06	0.10	-0.37	-0.01	0.06	0.00
DistMed					1.00	-0.17	-0.05	0.19	-0.06	-0.32	0.02	-0.35	0.08	-0.23	0.49	-0.15	**0.95**
DistOce						1.00	0.12	0.10	-0.03	0.50	0.36	0.37	0.11	-0.37	-0.63	0.06	-0.21
Expos							1.00	0.038	0.05	0.20	0.12	0.11	0.03	-0.10	-0.07	-0.22	-0.04
Litho								1.00	-0.08	0.01	0.17	0.08	0.04	-0.11	0.05	-0.01	0.20
NDVI									1.00	0.26	0.42	0.37	0.04	-0.28	-0.18	-0.43	0.03
Pente										1.00	0.37	0.41	-0.17	-0.30	-0.53	-0.10	-0.28
Pmoy											1.00	**0.89**	0.45	-0.69	-0.43	-0.32	0.05
Q2												1.00	0.38	-0.43	-0.53	-0.21	-0.35
RadSola													1.00	-0.38	-0.15	-0.04	0.06
Tmin														1.00	0.45	0.28	-0.22
Tmoy															1.00	0.10	0.50
TVDI																1.00	-0.17
Tmax																	1.00

Tab. S3 - Relative contributions (%) of environmental variables in the construction of species distribution models. Note that the values of the most important variables are shown in bold, i.e., those which made the greatest contribution to the model. Am: *A. marocana*, Ca: *C. atlantica*, Oe: *O. sylvestris*, Ph: *P. halepensis*, Pib: *P. iberica*, Pmag: *P. maghrebiana*, Pn: *P. mauretanica*, Qc: *Q. faginea*, Qcc: *Q. coccifera*, Ql: *Q. lusitanica*, Qp: *Q. pyrenaica*, Qr: *Q. rotundifolia*, Qs: *Q. suber*, Ta: *T. Articulata.*

Contrib.	Variable	Am	Ca	Oe	Ph	Pib	Pmag	Pn	Qc	Qcc	Ql	Qp	Qr	Qs	Ta
Percent contribution	Tmoy	**65.5**	**62.5**	**37**	**41.5**	8.1	**62.6**	**75.6**	4.9	**23.8**	**20.3**	16.5	**81.3**	0.8	**36.5**
	Tmax	**9**	2.7	**11.4**	**16.5**	**32.4**	**9.6**	1.4	**9.7**	**39.6**	**32**	6.1	**5.2**	**9.5**	**32.4**
	Pmoy	**13.6**	**23.9**	**28.2**	1.2	2	2.5	0.3	**32.3**	**11.7**	10.2	**25.7**	2.2	**21.7**	**17.4**
	Litholo	6.4	0.9	7.5	10.9	**14.7**	**16.6**	**12.3**	1.3	3.7	**21.8**	2.7	2.5	**60.3**	2.6
	TVDI	5.3	**5.7**	2.3	6	12.6	7	**5.6**	**51.1**	3.6	10.5	**17.2**	0.6	7.3	0.3
	Pente	0	0.3	2.5	**20.9**	**19.7**	0.6	1.1	0.1	8.4	2	2.4	**4.4**	0.1	9.8
	Tmin	0.1	3.2	3.9	0.1	6.9	0	2.7	0.1	1.8	2.1	**28.5**	2.5	0.2	0.2
	Expos	0.1	0.1	3.6	2.2	2	0.1	0.1	0.3	0.2	0.2	0.1	0.7	0.1	0.3
	RadSola	0	0.6	1.2	0.4	0.1	0.7	0.2	0.1	6.2	0.2	0.8	0.1	0.1	0.1
	Courb	0	0.1	2.4	0.1	1.5	0.2	0.7	0	1	0.6	0	0.4	0	0.4
Permutation importance	Tmoy	**60.2**	**67.7**	**28.4**	**49.7**	6.4	**81.1**	**97.7**	**12.6**	**26.8**	**18.9**	**38.2**	**71.6**	1.5	**27**
	Tmax	**18.9**	**6.5**	**17.1**	**21.5**	**63.6**	**7.6**	**1.2**	**23.4**	**28.2**	**49.4**	12.8	**14.8**	**12.5**	**37.2**
	TVDI	6.6	4.2	4.9	**9.4**	**8.2**	1.3	0.4	**52.7**	4.5	3.4	**18.2**	0.8	6.3	0.5
	Pmoy	**9.4**	**14.5**	**27.9**	2.5	1.5	1.1	0.1	9.4	**12.4**	7.6	4.7	0.3	**14.4**	**21.5**
	Litho	2.4	4.5	5.5	4	2.5	**7.1**	**0.4**	0.5	2.7	**16.9**	1.3	2.4	**63.5**	1
	Pente	1.6	0.8	1.8	6.6	6.9	1.1	0	0.4	8.4	0.4	4	**7.9**	0.4	10.4
	Tmin	0.5	0.6	5.3	0.3	**9**	0.1	0	0.3	6.6	3.1	**20.3**	1	0.5	0.8
	Expos	0.3	0.3	5.7	5.7	0.3	0.2	0	0.7	0.4	0.3	0.1	0.5	0.4	1.1
	RadSola	0	0.5	1.8	0.3	0	0.2	0	0	7.1	0.1	0.5	0.2	0.4	0.2
	Courb	0.1	0.4	1.5	0	1.5	0.1	0.1	0	2.9	0	0	0.6	0	0.4

Tab. S4 - Summary of the predicted potential and current areas of the fourteen taxa studied. The current area is based on the National Forest Inventory (NFI).

Species	Potential area (ha)			Current area (NFI)		
	Including unforested lands (1)	**Excluding unforested lands** (2)	**Unforested land (%)** ((2)-(1))/(2)	**(ha)**	**%**/(1)	**%**/(2)
A. marocana	22429	18920	16	3000	13	16
C. atlantica	36198	29736	18	4192	12	14
P. maghrebiana	60727	47883	21	9597	16	20
O. sylvestris	326562	57876	82	-		
P. halepensis	65834	47902	27	1382	2	3
P. iberica	66285	42736	36	1229	2	3
P. mauretanica	10658	9623	10	501	5	5
Q. faginea	54279	48298	11	10790	20	22
Q. coccifera	351736	127020	64	17722	5	14
Q. lusitanica	65254	35847	45	-		
Q. pyrenaica	45744	39164	14	2087	5	5
Q. rotundifolia	238457	124481	48	42336	18	34
Q. suber	316874	173641	45	113330	36	65
T. articulata	188800	76071	60	21205	11	28

Tab. S5 - Results of the binomial test and the Kappa statistical index obtained after comparison of the developed PNV model with the real distribution of the studied species.

Category	Binomial Test			Agreement relevance	
	Observed Proportion	**Test proportion**	**Signification exact (unil.)**	**Kappa measure**	**Approximate Significance**
1 (concordance)	0.76	0.76	0.249	0.670	0.000
0 (no concordance)	0.24				
Total	1.00				

Fig. S1 - Omission rate and predicted area for broadleaved (a) and coniferous (b) species.

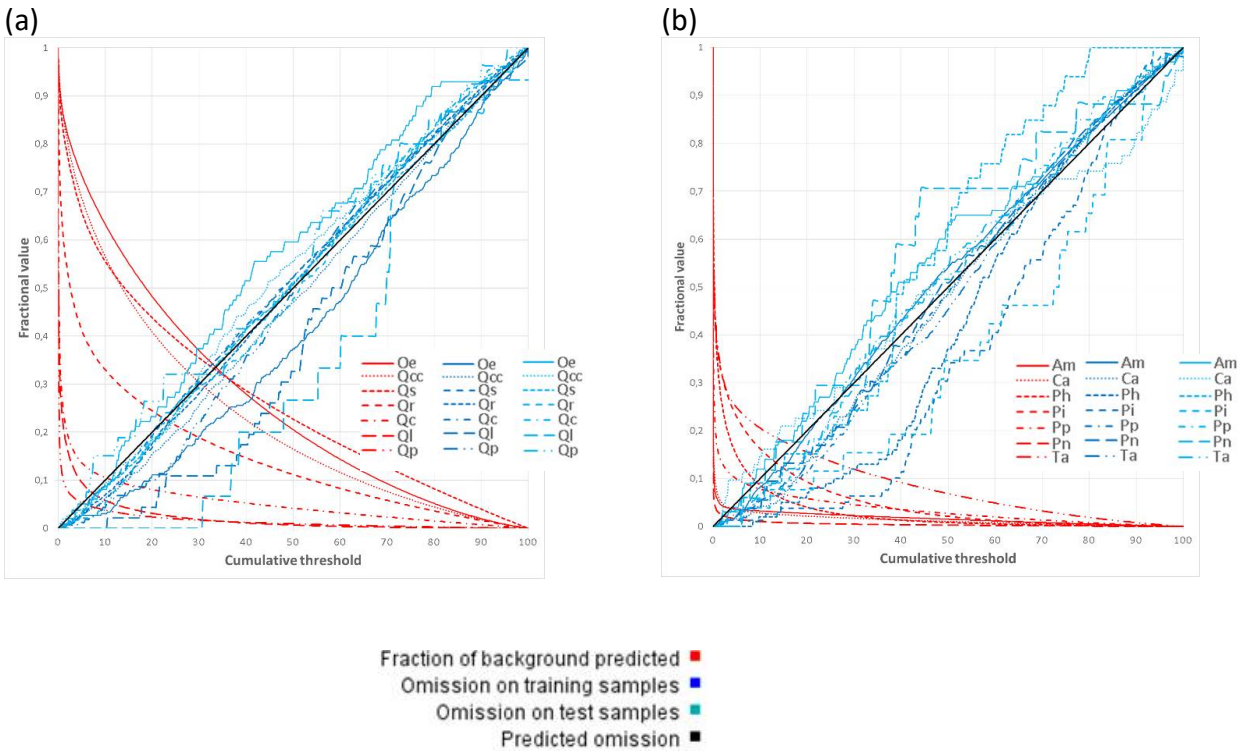


Fig. S2 - Receiver operating characteristic (ROC) curve for broadleaved (a) and coniferous (b) species.

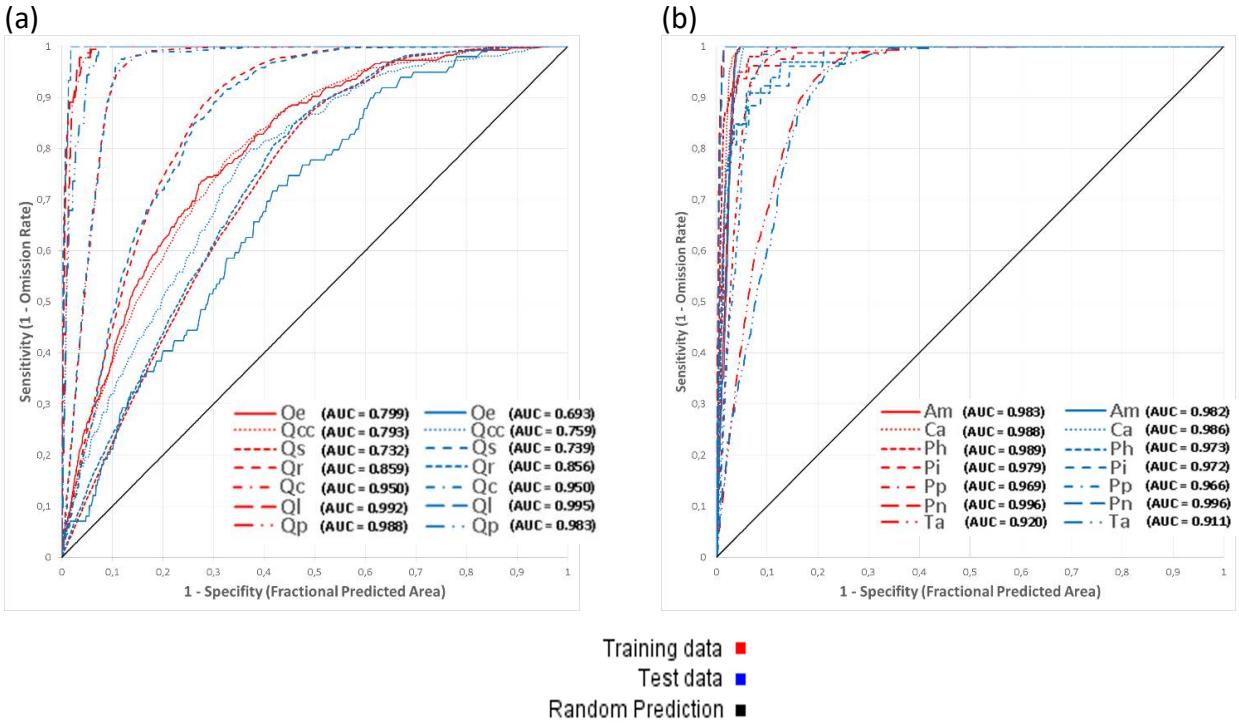
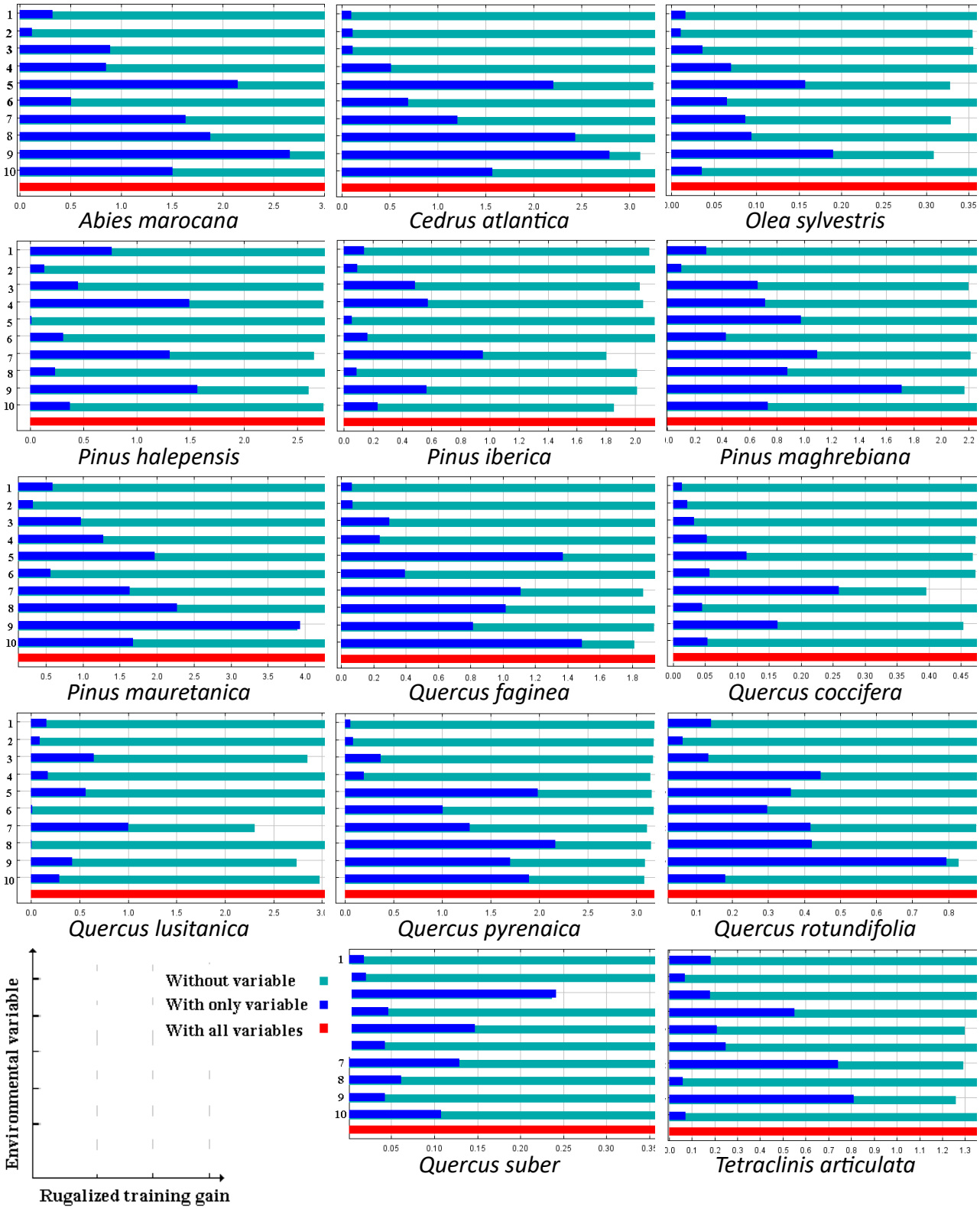


Fig. S3 - Results of the jackknife test showing the importance of the ten studied variables in determining the species potential range using training gain. 1 : Curvature, 2 : Aspect, 3 : Lithology, 4 : Slope, 5 : Pmoy, 6 : Solar radiation, 7 : Tmax, 8 : Tmin, 9 : Tmoy, 10 : Soil humidity.



Appendix 1 – Rationale of the jackknife test applied.

The jackknife test facilitates the assessment of the individual contributions of each variable to the overall model. It provides comparative results regarding the performance of the model under three different conditions: i) using each variable alone, ii) including one variable at a time, and iii) using all variables together.

The jackknife test reveals that for *A. marocana, Cedrus atlantica, P. maghrebiana, O. sylvestris, P. halepensis, P. mauretanica, Q. rotundifolia,* and *T. articulata*, Tmoy is the environmental variable that contributes the most information to their SDMs. The gain is more significant when this variable is the only explanatory variable used and results in the most substantial loss of model performance when omitted. For *P. iberica, Q. coccifera,* and *Q. lusitanica*, Tmax provides the most significant gain when used solely, as it is the variable that most affects the quality of the model when neglected. For *Q. faginea*, soil moisture (TVDI) contains the most significant information compared to the other variables, thus improving its SDM. For *Q. pyrenaïca*, Tmin provides the greatest improvement in terms of the model's quality. When considered independently, lithology emerges as the most significant factor influencing the *Q. suber* SDM. Neglecting the impact of this variable also compromises the accuracy of the predictive outcomes.

With the exception of some species such as *P. mauretanica, Q. lusitanica*, and *Q. suber*, omitting each variable in turn did not significantly reduce the training gain (lighter blue bars). Therefore, no single variable contained a substantial amount of useful information in addition to that provided by the other remaining variables. However, for all the species analyzed, the SDMs developed by excluding variables one at a time (light blue bars) or with subsets of the remaining variables (dark blue bars) consistently performed worse than models that used all variables together (red bars). This result underscores the importance of including the complete set of variables to improve model effectiveness in the study area.