

Supplementary Material

1 Supplementary Figures and Tables

1.1 Supplementary Tables

Supplementary Table 1. SDMs compared for each Joshua tree species and for the rangewide model. Each of the eight models was evaluated across 25 cross-validation runs and two algorithms (GAM and random forest). See Table 1 in main text for definition of terms.

Model	Terms*
1	AHM+STMX+WTMN+AMPn+MAXn+HLI+PCV+Pratio+Sand +Slope+TPI+TSD+Trange
2	AHM+MAT+AMPn+MAXn+HLI+PCV+Pratio+Sand +Slope+TPI+TSD+Trange
3	CMD+SP+WP+WTMN+AMPn+MAXn+HLI+PCV+Pratio+Sand +Slope+TPI+TSD+Trange
4	CMD+MAP+WTMN+AMPn+MAXn+HLI+PCV+Pratio+Sand +Slope+TPI+TSD+Trange
5	MAP+MAT+AMPn+MAXn+HLI+PCV+Pratio+Sand +Slope+TPI+TSD+Trange
6	MAP+STMX+WTMN+AMPn+MAXn+HLI+PCV+Pratio+Sand +Slope+TPI+TSD+Trange
7	WP+SP+WTMN+STMX+AMPn+MAXn+HLI+PCV+Pratio+Sand +Slope+TPI+TSD+Trange
8	WP+SP+MAT+AMPn+MAXn+HLI+PCV+Pratio+Sand +Slope+TPI+TSD+Trange

Supplementary Table 2. Changes made to initial satellite Joshua tree identifications based on field surveys and secondary satellite surveys by experienced observers. Numbers are grid cell counts from the study region (see main text for details).

Field Validation Surveys					
Status Change	AZ	CA	NV	UT	Total
?->A	11	0	0	0	11
?->P	13	3	97	0	113
A->D	0	5	10	1	16
A->P	221	795	2282	139	3437
D->P	19	8	12	23	62
P->A	58	1	0	0	59
P->D	0	0	1	0	1
C->A	3	0	0	0	3
C->P	1	0	0	0	1
No Status Change	5027	9037	10908	375	25347
Total Changed	326	812	2402	163	3703
Grand Total	5353	9849	13310	538	29050

Secondary Satellite Surveys					
Status Change	AZ	CA	NV	UT	Total
?->A	227	185	296	0	708
?->D		1	16	0	17
?->P	63	157	884	0	1104
A->C	0	248	0	0	248
A->D	35	515	1145	121	1816
A->P	1142	11695	13747	91	26675
D->A	28	18	22	1	69
D->C	0	1	0	0	1
D->P	31	201	158	42	432
P->A	611	3895	1074	18	5598
P->D	2	4	21	3	30
P->C		6	0	0	6
C->A	36	28	0	0	64
C->P	3	84	2	0	89
No Status Change	5043	24122	10071	485	39721
Total Changed	2178	17038	17365	276	36857
Grand Total	7221	41160	27436	761	76578

Definitions: P=present; A=absent; C=cultivated yard tree; D=dead; ?=no status assigned

Supplementary Table 3. Model performance and relative variable importance.

A. *Yucca jaegeriana* models (sorted by average AIC for GAM, and AUC for random forest).

GAM

Model performance					Variable relative importance																	
Model	AIC _{avg}	AUC	TSS	R ²	AHM	AMP	CMD	HLI	MAP	MAT	MAXN	PCV	Pratio	Sand	Slope	SMT	SP	TPI	Trange	TSD	WMT	WP
3	69858	0.796	0.462	0.388		2.47	8.21	0.38			4.82	11.18	14.25	7.08	1.55		15.21	3.94	6.84	7.79	13.18	3.09
7	70524	0.798	0.465	0.382		2.13		0.45			7.14	11.98	18.15	9.50	2.38	5.62	17.92	3.87	1.79	10.03	5.06	3.97
8	71029	0.799	0.469	0.377		2.68		0.53		13.02	6.98	12.85	17.39	8.74	2.49		14.90	3.79	3.73	9.83		3.07
1	71326	0.798	0.457	0.375	33.80	2.05		0.20			6.72	9.17	11.11	7.86	2.49	6.19		3.26	2.50	10.14	4.51	
4	71685	0.785	0.444	0.372		2.25	7.97	0.24	26.83		5.48	9.95	6.97	6.58	1.65			3.50	6.47	8.42	13.69	
2	71766	0.799	0.467	0.371	31.65	2.62		0.27		12.02	6.71	10.11	10.83	6.84	2.67			3.16	3.60	9.53		
6	72348	0.787	0.444	0.366		2.16		0.25	31.87		7.68	10.58	9.30	9.06	2.48	5.49		3.45	1.54	10.78	5.38	
5	72571	0.792	0.455	0.364		2.40		0.29	28.22	15.01	7.45	10.97	9.39	7.62	2.50			3.16	3.18	9.81		

Random forest

Model performance					Variable relative importance																	
Model	AIC _{avg}	AUC	TSS	R ²	AHM	AMP	CMD	HLI	MAP	MAT	MAXN	PCV	Pratio	Sand	Slope	SMT	SP	TPI	Trange	TSD	WMT	WP
2	-	0.859	0.567	0.832	0.17	0.05		0.02		0.16	0.08	0.10	0.11	0.09	0.04			0.05	0.07	0.10		
1	-	0.858	0.563	0.835	0.17	0.05		0.02			0.08	0.10	0.10	0.09	0.04	0.12		0.05	0.07	0.10	0.10	
5	-	0.858	0.566	0.833		0.05		0.02	0.16	0.17	0.08	0.10	0.11	0.09	0.04			0.05	0.07	0.10		
8	-	0.857	0.563	0.838		0.05		0.02		0.16	0.07	0.10	0.10	0.09	0.04		0.14	0.05	0.07	0.09		0.12
7	-	0.857	0.563	0.841		0.05		0.02			0.07	0.09	0.10	0.09	0.04	0.12	0.14	0.05	0.06	0.09	0.10	0.12
6	-	0.856	0.560	0.837		0.05		0.02	0.16		0.08	0.10	0.11	0.09	0.04	0.12		0.05	0.07	0.10	0.10	
4	-	0.855	0.557	0.843		0.05	0.11	0.02	0.16		0.07	0.10	0.11	0.09	0.04			0.05	0.07	0.10	0.11	
3	-	0.855	0.555	0.846		0.05	0.10	0.02			0.07	0.09	0.10	0.09	0.04		0.14	0.05	0.07	0.09	0.10	0.13

B. *Yucca brevifolia* models (sorted by average AIC for GAM, and AUC for random forest).

GAM

Model performance					Variable relative importance																	
Model	AICavg	AUC	TSS	R ²	AHM	AMP	CMD	HLI	MAP	MAT	MAXN	PCV	Pratio	Sand	Slope	SMT	SP	TPI	Trange	TSD	WMT	WPI
7	65640	0.455	0.980	0.570		2.27		1.24			10.63	15.66	9.53	2.67	2.14	7.55	14.24	4.73	3.46	13.56	9.19	3.15
3	66062	0.452	0.983	0.558		1.95	5.54	1.11			10.07	16.37	10.22	2.02	1.47		15.60	5.27	5.39	12.06	9.36	3.57
8	67156	0.443	0.978	0.561		2.02		1.31		12.48	9.89	15.47	10.33	2.31	2.18		15.93	4.64	5.58	14.38		3.47
1	67678	0.438	0.934	0.557	8.58	4.30		1.47			11.71	13.62	9.70	3.29	2.30	5.08		6.24	4.48	14.46	14.77	
6	68392	0.432	0.946	0.550		3.80		1.43	5.07		12.72	12.92	9.29	3.18	2.24	10.48		5.61	4.76	13.87	14.65	
4	68518	0.431	0.967	0.537		2.92	8.56	1.16	8.22		10.90	12.41	9.49	2.22	1.22			5.65	9.75	10.11	17.41	
2	68946	0.428	0.946	0.553	10.78	4.08		1.47		16.09	10.97	12.84	9.68	3.20	2.45			5.70	6.96	15.78		
5	69843	0.420	0.929	0.539		3.41		1.45	6.94	22.93	11.79	12.29	9.32	2.77	2.32			5.10	7.06	14.61		

Random Forest

Random Forest																						
Model performance					Variable relative importance																	
Model	AIC _{avg}	AUC	TSS	R ²	AHM	AMP	CMD	HLI	MAP	MAT	MAXN	PCV	Pratio	Sand	Slope	SMT	SP	TPI	Trange	TSD	WMT	WP
4	-	0.878	0.600	0.819		0.05	0.10	0.02	0.14		0.08	0.11	0.12	0.04	0.04			0.04	0.08	0.13	0.11	
3	-	0.877	0.597	0.821		0.05	0.10	0.02			0.07	0.11	0.11	0.04	0.03		0.08	0.04	0.08	0.12	0.10	0.13
6	-	0.876	0.594	0.811		0.05		0.02	0.12		0.08	0.11	0.11	0.04	0.04	0.12		0.04	0.07	0.12	0.10	
1	-	0.875	0.593	0.811	0.13	0.05		0.02			0.08	0.11	0.11	0.04	0.04	0.11		0.04	0.07	0.12	0.10	
5	-	0.875	0.592	0.810		0.05		0.03	0.13	0.14	0.08	0.12	0.12	0.05	0.04			0.04	0.08	0.13		
7	-	0.875	0.590	0.814		0.05		0.02			0.07	0.11	0.11	0.04	0.04	0.12	0.08	0.04	0.07	0.12	0.10	0.12
2	-	0.875	0.590	0.809	0.14	0.05		0.03		0.14	0.08	0.12	0.11	0.05	0.04			0.04	0.08	0.13		
8	-	0.874	0.590	0.812		0.05		0.02		0.14	0.08	0.11	0.11	0.04	0.04		0.09	0.04	0.08	0.12		0.13

C. Rangewide random forest model (sorted by AUC).

Model performance					Variable relative importance																	
Model	AIC _{avg}	AUC	TSS	R ²	AHM	AMP	CMD	HLI	MAP	MAT	MAXN	PCV	Pratio	Sand	Slope	SMT	SP	TPI	Trange	TSD	WMT	WP
7	-	0.869	0.577	0.813		4.58		1.88			6.48	9.67	9.86	5.39	3.53	10.48	10.96	3.75	6.00	8.50	8.26	10.65
8	-	0.869	0.576	0.809		4.84		2.04		12.82	6.79	10.43	10.53	5.76	3.75		11.76	4.00	6.69	9.23		11.37
1	-	0.868	0.579	0.806	14.65	5.11		2.11			7.41	10.50	10.89	5.88	3.99	10.76		4.23	6.63	9.10	8.73	
2	-	0.868	0.576	0.802	15.99	5.50		2.28		13.16	7.74	11.25	11.64	6.30	4.27			4.53	7.43	9.91		
6	-	0.868	0.576	0.808		4.93		2.09	13.82		7.29	10.49	10.84	5.89	3.90	11.55		4.15	6.71	9.17	9.17	
3	-	0.868	0.574	0.820		4.55	8.30	1.79			6.34	9.84	9.94	5.69	3.36		11.01	3.61	6.57	8.69	8.77	11.55
4	-	0.868	0.575	0.816		4.91	9.38	1.98	14.41		7.09	10.74	10.93	6.25	3.76			4.00	7.42	9.32	9.81	
5	-	0.868	0.575	0.805		5.31		2.28	15.13	14.19	7.68	11.33	11.63	6.35	4.20			4.48	7.48	9.96		

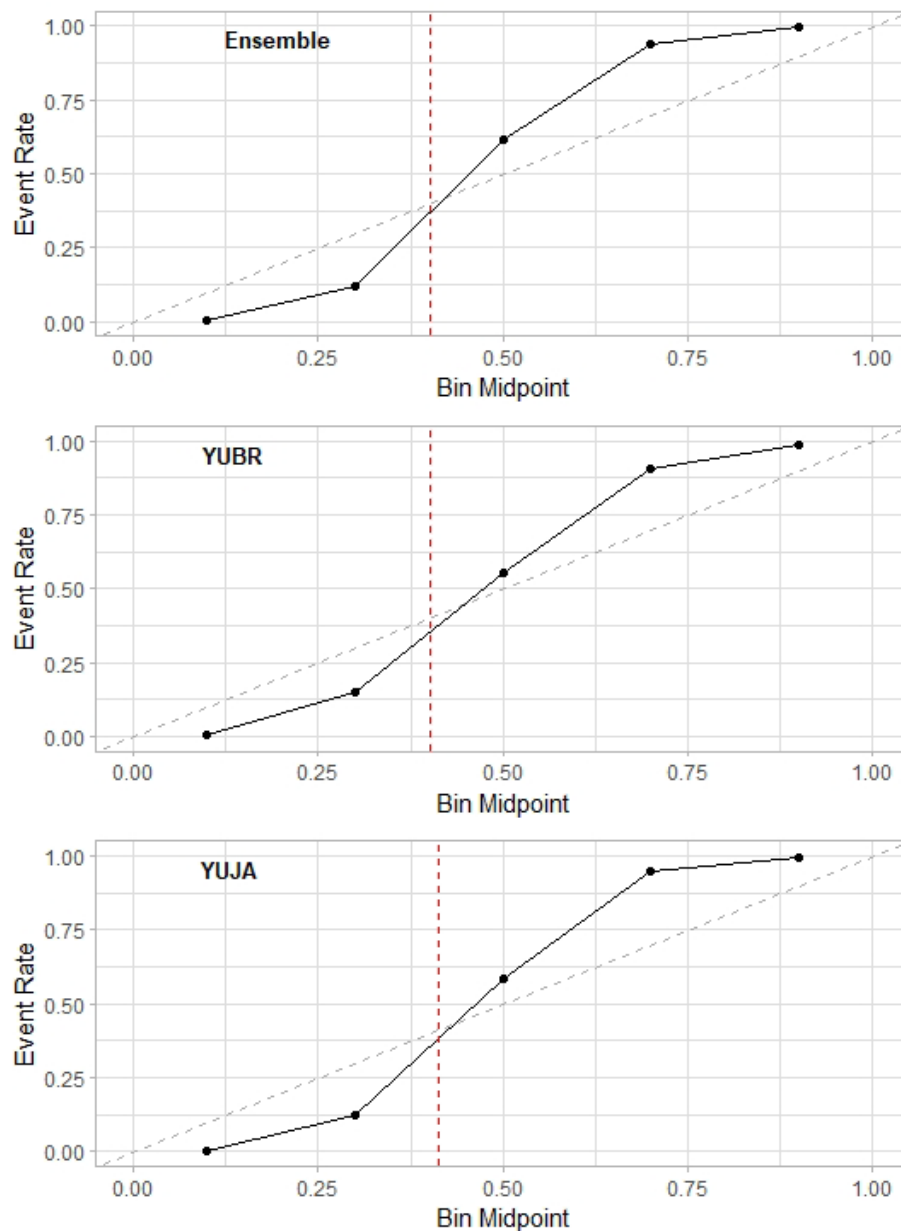
Supplementary Table 4. Comparison of historic models with occupied habitat (current geographical distribution). Total area of habitat is excluding the area of obscured imagery. These comparisons include habitat for both *Yucca brevifolia* and *Yucca jaegeriana* for comparison with historic maps that did not distinguish two species.

	Area (km ²) and % Overlap*			
	Rowlands	Godsoe et al.	Cole et al.	Wilkening et al.
Overlap with current distribution, km ² (%)	15,145 (46.4%)	15,380 (47.1%)	16,223 (49.7%)	26,026 (79.7%)
Non-overlap with Rangewide Model, km ²	10,509	8,643	54,495	17,398
Non-overlap with habitat, km ²	17,492	17,258	16,414	6,612
Total Area of Models, km ²	25,654	24,023	70,718	43,422
Correctly ID'd Present	59.0%	64.0%	22.9%	59.9%
% Area of Current Distribution Presence missed	53.6%	52.9%	50.3%	20.3%

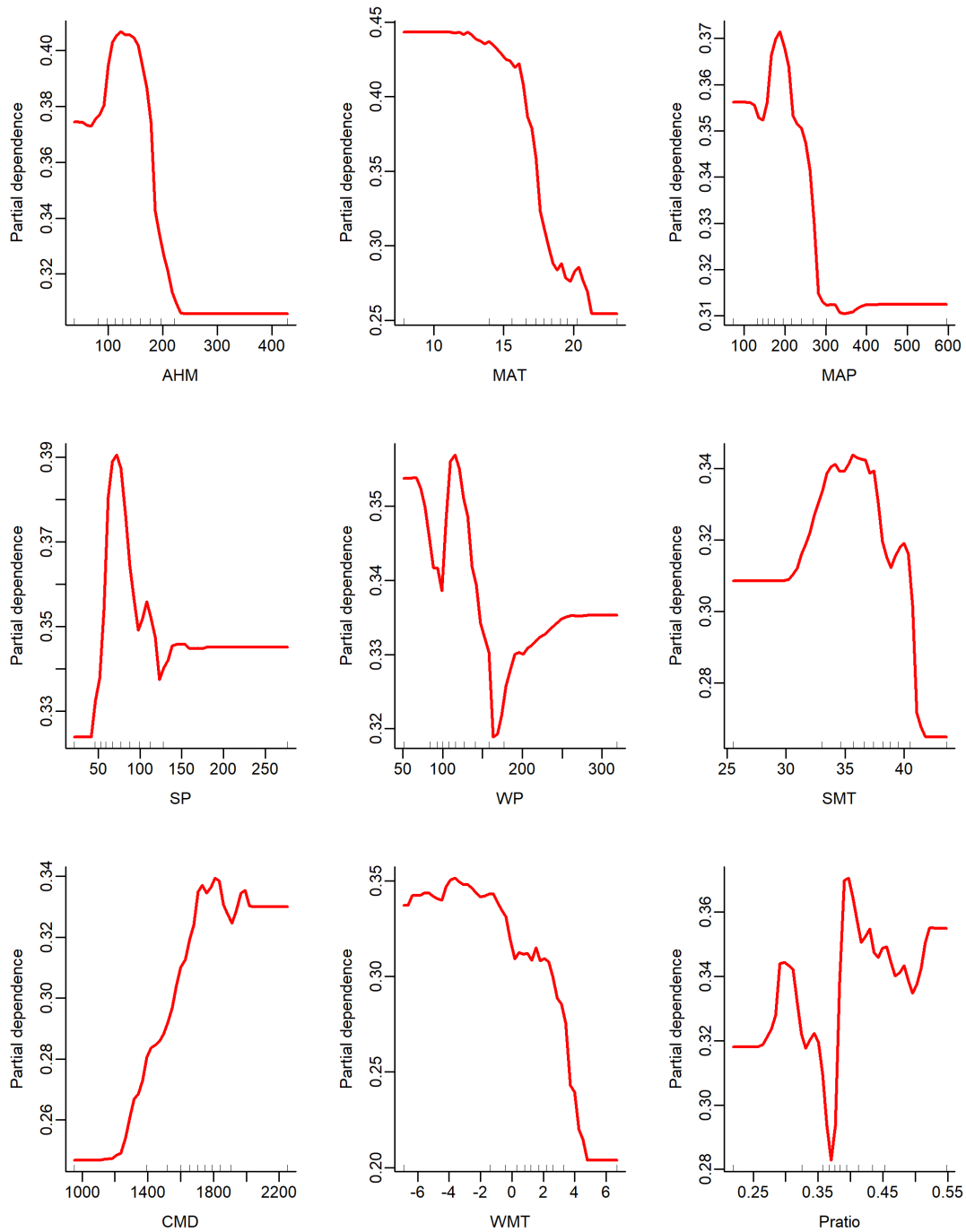
***Methods:** Final habitat map compared with previously published Joshua tree models, including shapefiles of previous models (Godsoe et al. 2009, Wilkening et al. 2020) or by digitizing distributions from historical publications (Rowlands 1978, Cole et al. 2011). Only those portions of the models that coincided with the empirically derived portion of the distribution map were compared. They were compared using the editing, clipping, and calculate geometry tools in ArcGIS (ESRI 10.8.1). For each comparison we calculated the total presence area (km²) that overlapped with our habitat map, the area of the model that did not overlap with our habitat map (non-overlap: model), and the area of our habitat that did not overlap with the model (non-overlap: USGS empirical).

1.2 Supplementary Figures

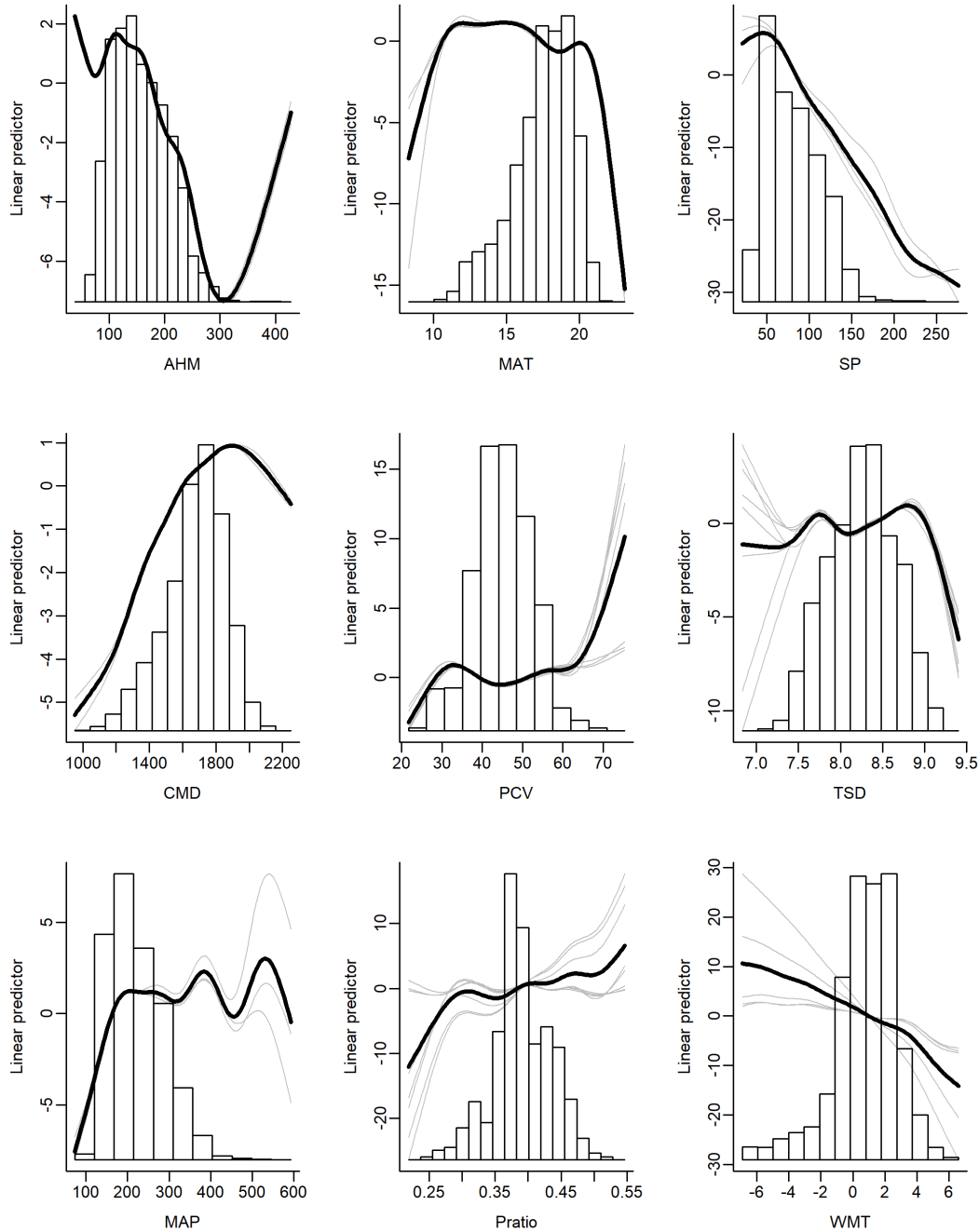
Supplementary Figure 1. Model calibration curves derived by binning the predicted habitat probabilities into 5 probability classes of width 0.2 (e.g., 0-0.2, 0.2-0.4, etc...) and comparing the midpoints of each bin with the observed frequency of presences within that subset of data. Panels display curves for the ensemble rangewide random forest model ('Ensemble'), the *Yucca brevifolia* GAM / random forest ensemble (YUBR), and the *Yucca jaegeriana* GAM / random forest ensemble (YUJA). Vertical red lines indicate the optimal model threshold based on maximizing the sum of model sensitivities and specificities (i.e., the threshold for classifying habitat versus non-habitat).



Supplementary Figure 2. Partial response curves from random forest SDM for *Y. jaegeriana*. Response curves depict the marginal influence of each term on model predictions. Curves were generated using the R package ‘pdp’ (Greenwell 2017). The top nine terms based on permutation importance are presented.

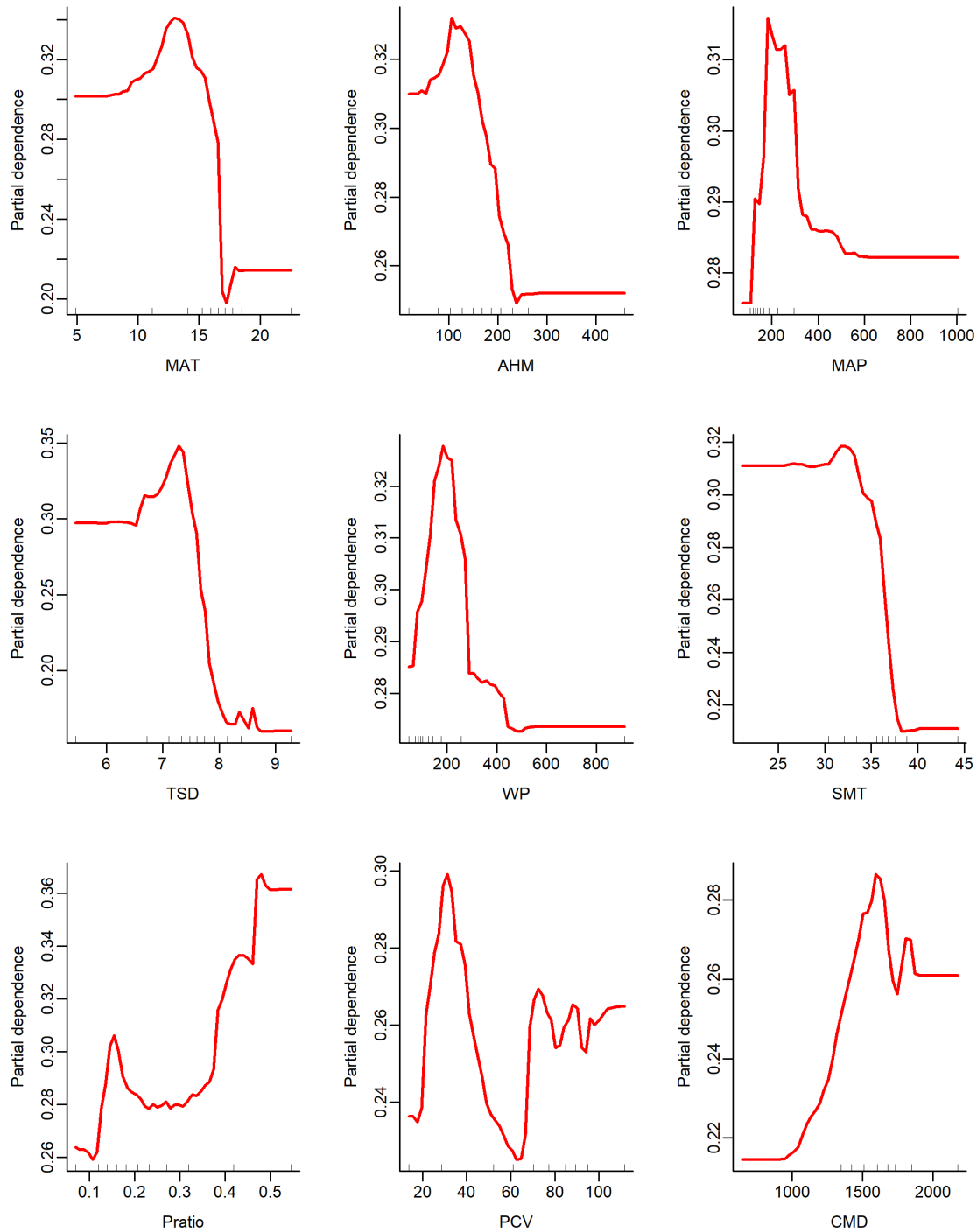


Supplementary Figure 3. Partial response curves from the GAM SDM for *Y. jaegeriana*. Response curves depict the marginal influence of each term on predicted probabilities. Gray lines in each panel show response curves from individual candidate models, while the thick black lines show the model-averaged response curve. Histograms depict a random sample of background points from the study region. The nine most important terms are displayed.

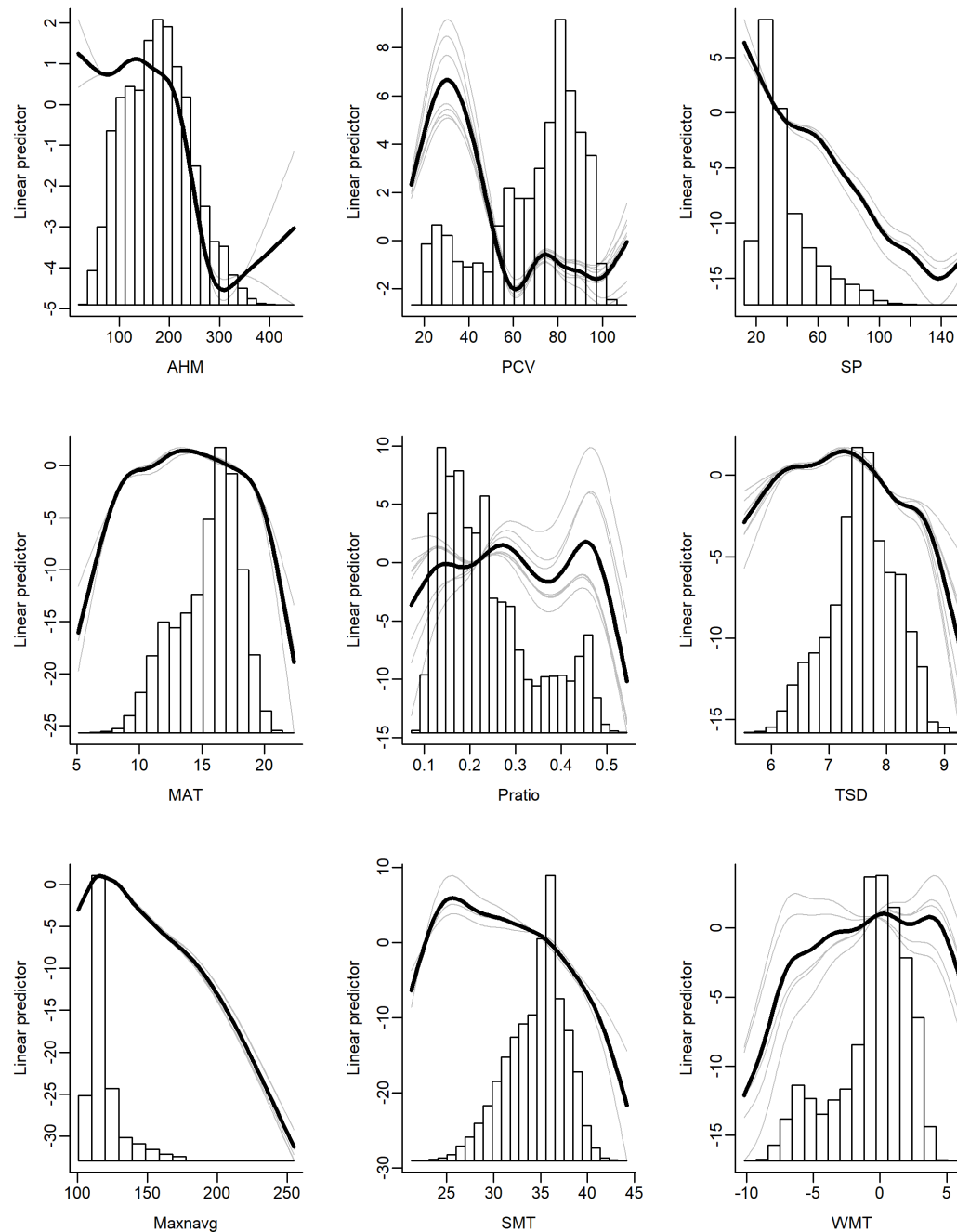


Supplementary Figure 4. Partial response curves from random forest SDM for *Y. brevifolia*. Response curves depict the marginal influence of each term on model predictions. Curves were

generated using the R package ‘pdp’ (Greenwell 2017). The top nine terms based on permutation importance are presented.



Supplementary Figure 5. Partial response curves from the GAM SDM for *Y. brevifolia*. Response curves depict the marginal influence of each term on predicted probabilities. Gray lines in each panel show response curves from individual candidate models, while the thick black lines show the model-averaged response curve. Histograms depict a random sample of background points from the study region. The nine most important terms are displayed.



Supplementary Figure 6. Partial response curves from the range-wide random forest SDM for the combined *Y. brevifolia* and *Y. jaegeriana* datasets. Response curves depict the marginal influence of each term on model predictions. Curves were generated using the R package ‘pdp’ (Greenwell 2017) for all terms.

