7. SUPPLEMENTAL MATERIAL:

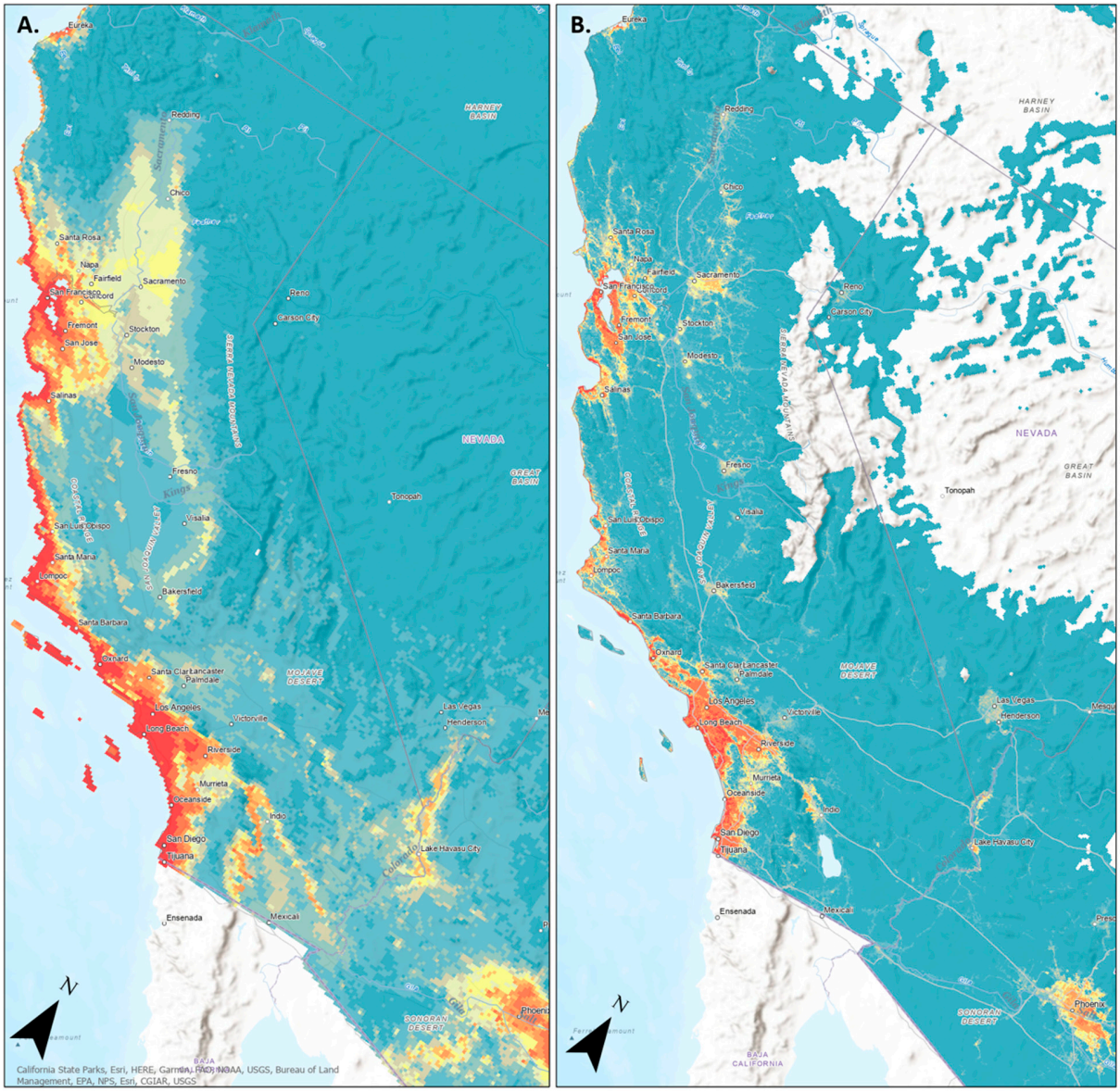

Figure S1. FR (A) and COR (B) predicted suitability estimates in Central and Southern California, and parts of Arizona, with major cities and roads overlayed for reference.

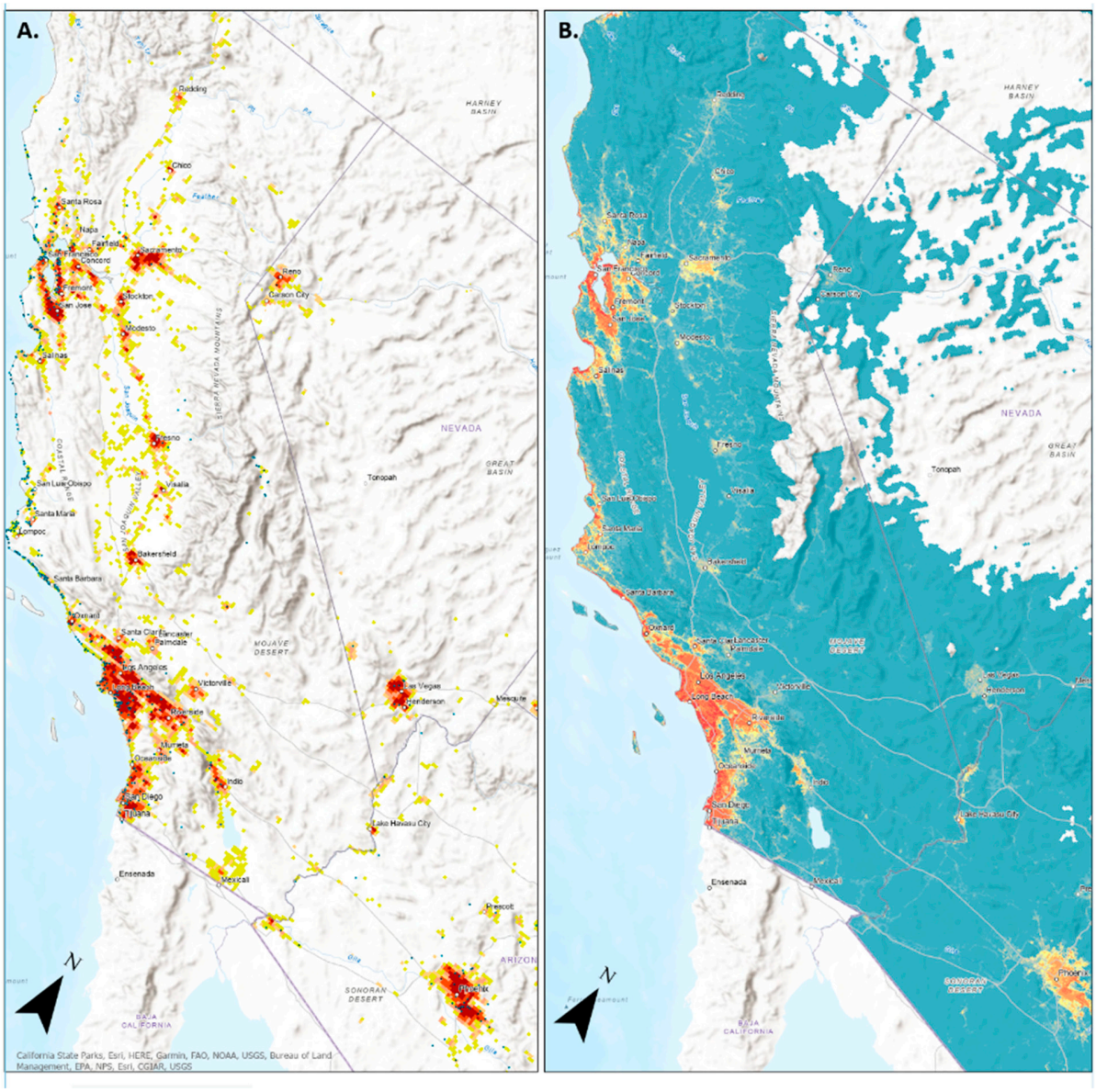

Figure S2. Percent Cover Medium Development (A) versus COR predictive output (B). (A) Percent cover medium development is depicted from low to high values, respectively colored as yellow to red. Blue points represent presence points used in generating the COR model. (B) Shows COR predicted probability of presence cloglog output. Warmer colors indicate higher probabilities of presence. This figure demonstrates the high likelihood of COR being too heavily weighted with the positive association of medium development and monarch presence points.

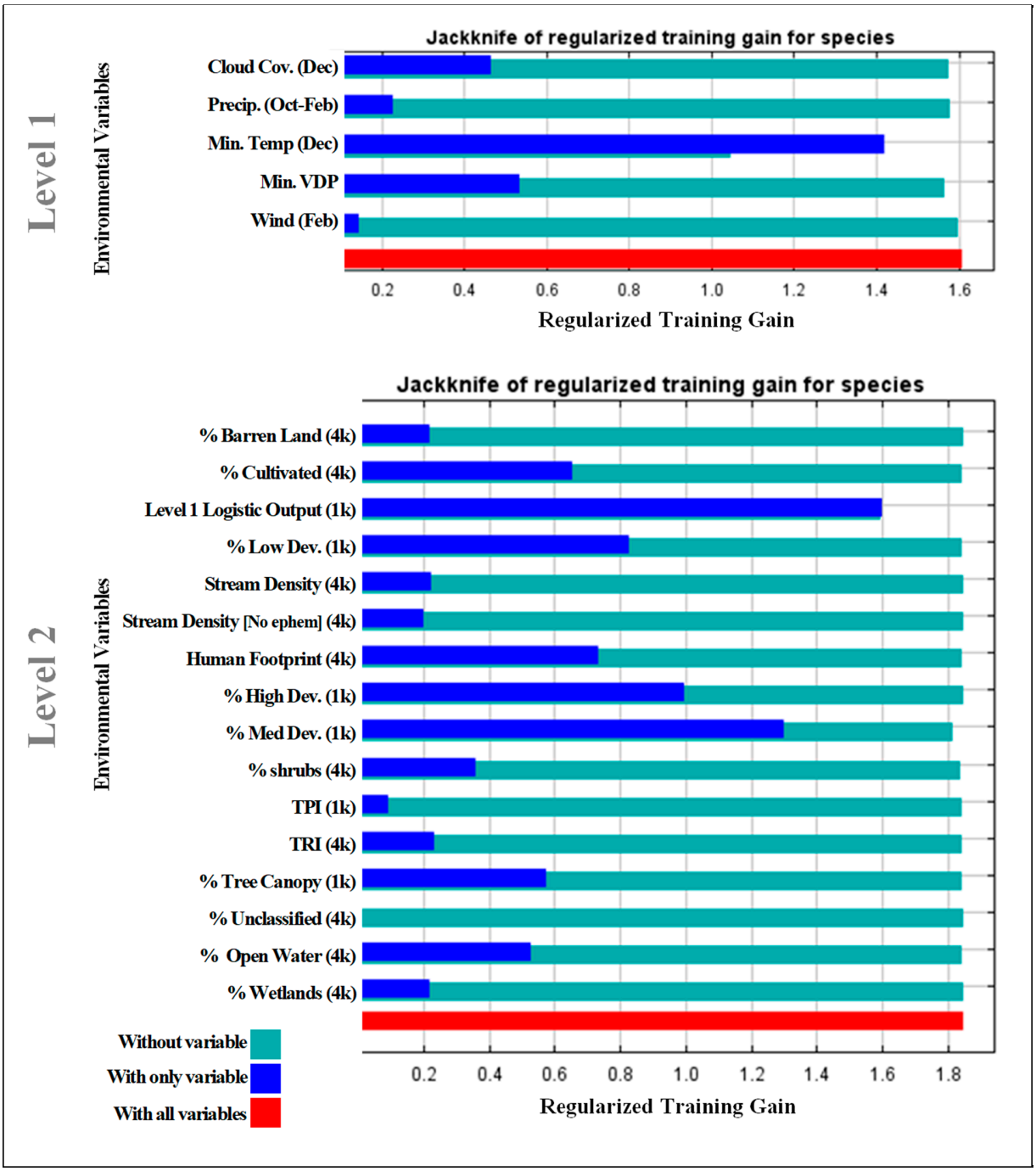

Figure S3. Jackknife Variable Importance under optimal settings for FR and COR. Blue bars show the regularized training gain of a model generated with the single predictor, whereas the teal bars show the regularized training gain of a model generated with all predictors, minus one. The red bar shows the regularized training gain of the model with all predictors. COR results are for LT model only, as LQT model had nearly identical results.

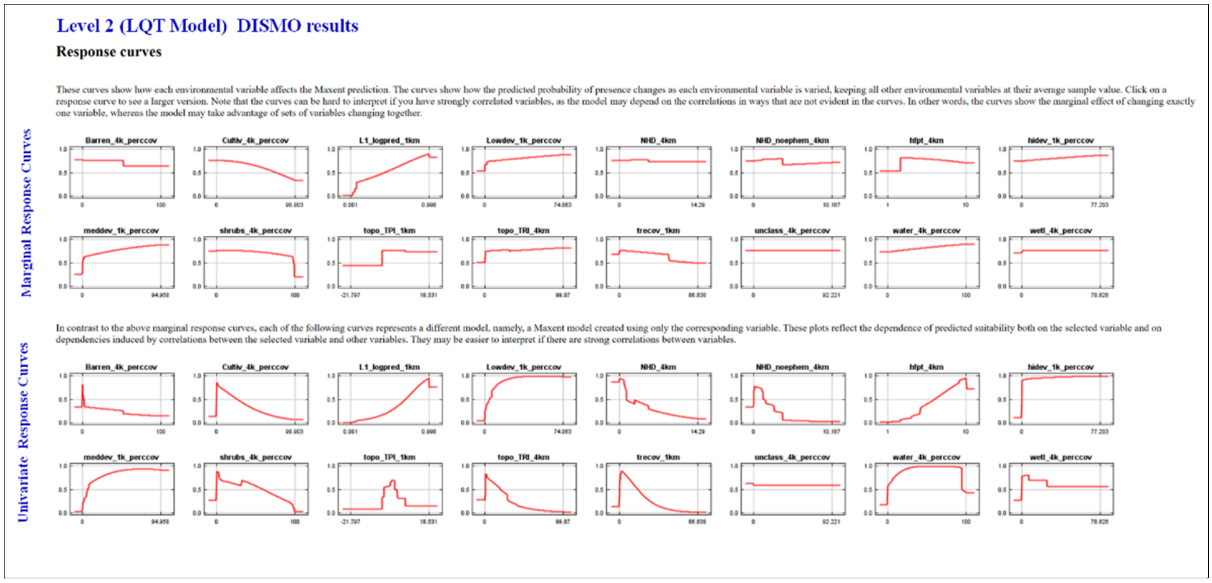

Figure S4. Marginal and Univariate Response Curves from the peak overwintering range (COR) Model, run with L Q and T feature classes which had an AICc score comparable to the LT Run shown in Figure 7.

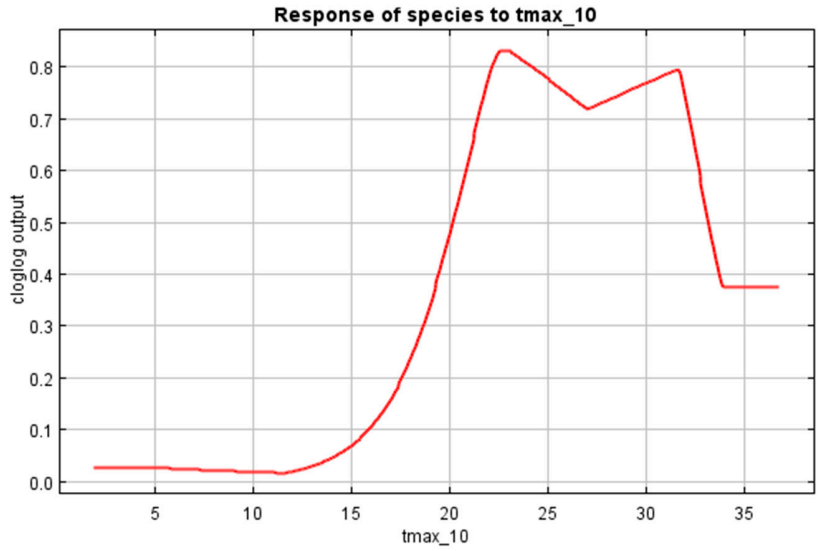

Figure S5. Univariate Response Curve for maximum temperature in the hottest month, October (tmax_10) from the preliminary model run with all candidate variables for the population wide model (FR) included, listed in Table 1.

Variable	Percent contribution
L1_logpred_1km	75.5
meddev_1k_perccov	18.2
Lowdev_1k_perccov	1.8
Cultiv_4k_perccov	1.2
topo_TRI_4km	0.9
shrubs_4k_perccov	0.7
water_4k_perccov	0.7
wetl_4k_perccov	0.3
hidev_1k_perccov	0.3
trecov_1km	0.2
topo_TPI_1km	0.1
NHD_noephem_4km	0.1
NHD_4km	0
Barren_4k_perccov	0
unclass_4k_perccov	0

Figure S6. Percent contribution of COR predictors in the run with Human footprint used as a bias file instead of a predictor, in an effort to ameliorate the apparent overfitting to developed areas.

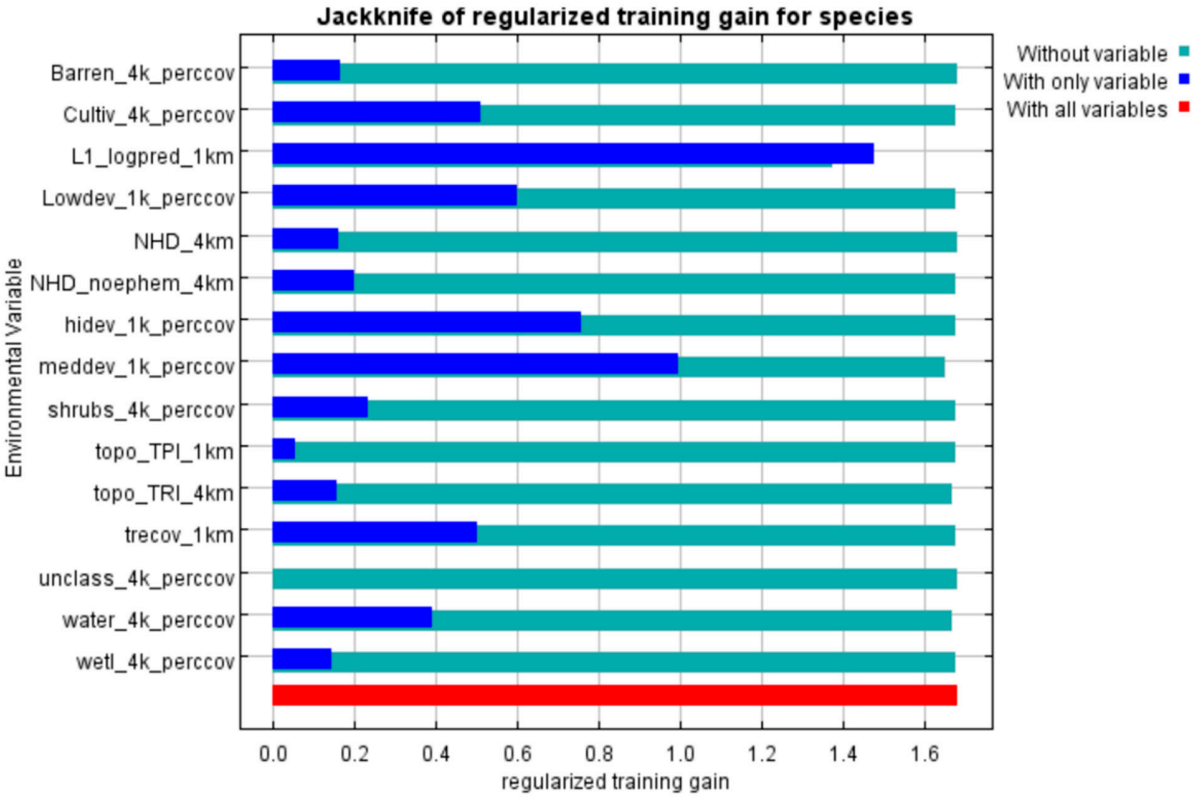

Figure S7. Jackknife Variable importance results of COR predictors run with Human footprint used as a bias file instead of a predictor, to ameliorate the apparent overfitting to developed areas.

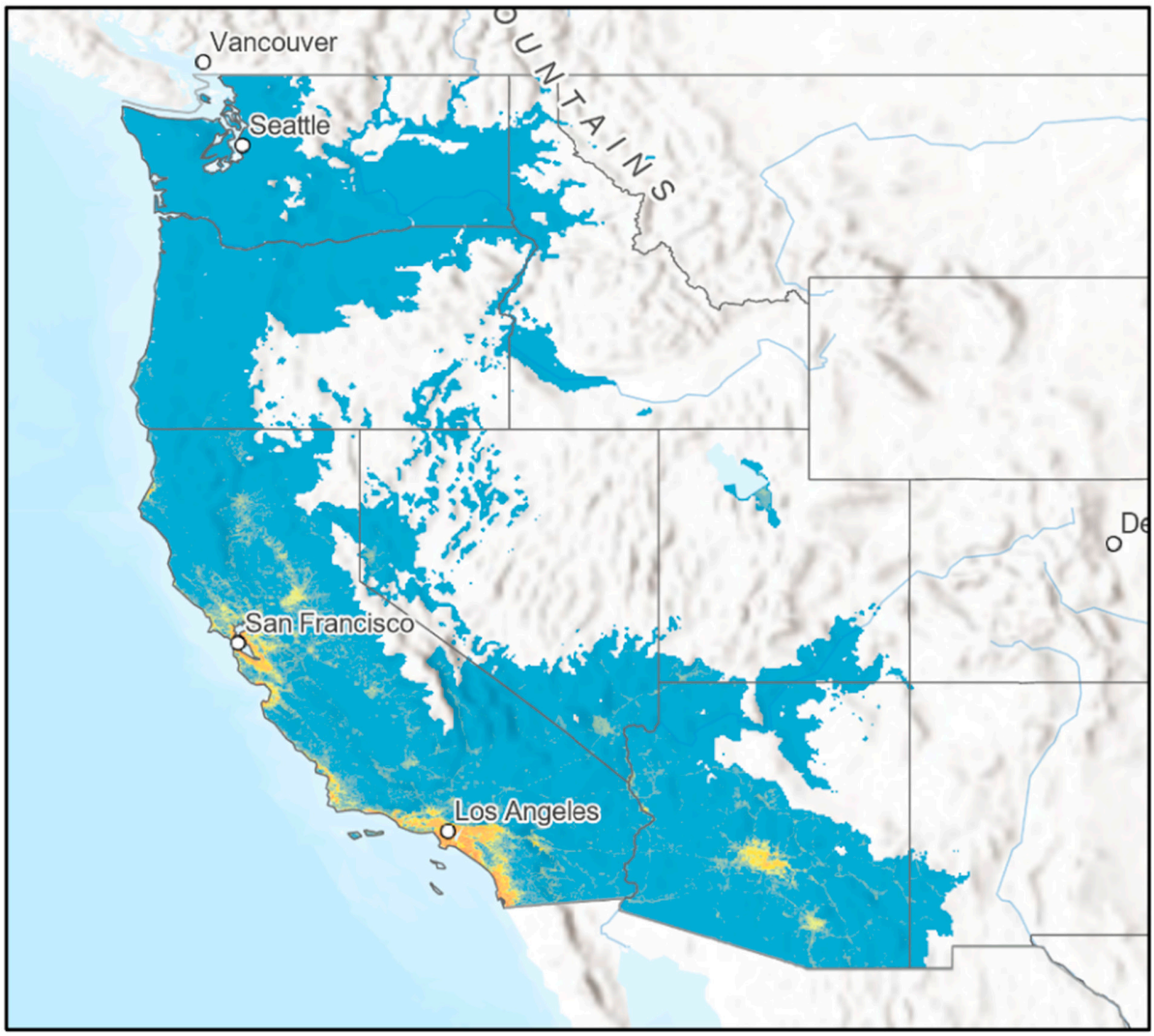

Figure S8. Cloglog Predictive output of COR predictors run with Human footprint used as a bias file instead of a predictor, in an effort to ameliorate the apparent overfitting to developed areas.

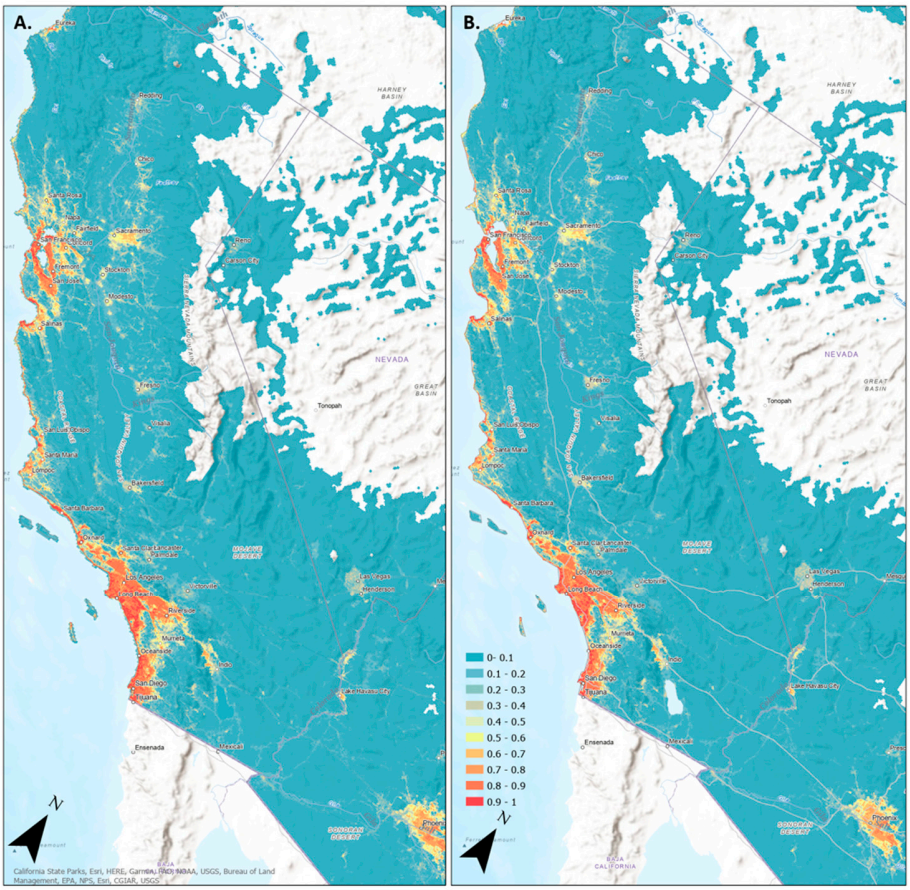

Figure S9. Comparison of output from the two top performing models from the COR Model. (A) shows output from the model run with Linear and Threshold features, with a regularization multiplier of 1.5 (LT_RM1.5 run). (B) shows output from the model run with linear, threshold, and quadratic features with a regularization multiplier of 1.5 (LQT_RM1.5 run).

7.1 Methods Supplemental Materials:

7.1.1 Presence Data Cleaning & Processing

Journey North presence was cleaned using R (R Core Team, 2013). Records missing Latitude or Longitude, and low precision records (< 3 decimal degree places ~111m accuracy) were removed. Observations were filtered to FR and COR extents and to October-February and November-February (respectively). Vague or imprecise grove locations from Xerces were removed (see Fisher et al*., in prep* for details; Ch.1) before analysis, leaving 304 overwintering site records. These were converted from polygons to central points for analysis using ArcMap 10.6 and combined with the JN points (ESRI, 2018). Combined, FR had 782 occurrence records and COR had 755. Last, the Maxent.jar algorithm was used to thin points to one per raster cell to reduce spatial autocorrelation (Phillips et al., 2017) such that the final number of occurrence points in FR was 490 and 637 for COR.

The extent of both FR and COR were truncated to exclude Baja California Norte, Mexico, due to data limitations. However, there are known overwintering sites in northern Baja, and in future research a climatic model including Baja, given data is available, could help identify overwintering regions at the

south end of the population's extent. But unfortunately, predictive data layers for smaller scale, landscape-level analyses, are currently unavailable or sparse, and accuracy of Baja overwintering site locations were not yet refined at the time of these analyses.

Table S1. Presence data provided, and subsequently removed through cleaning processes described in the methods.

Number of Presences	**FR**	**COR**
Cleaned and Run through Maxent	787	755
Removed due to NA predictor Values (by Maxent)	-5	-7
Removed for Spatial Thinning (by Maxent)	-292	-111
Final # Occurrences Used to Fit Model:	490	637

7.1.2 Environmental Predictors Selecting Appropriate Resolution:

Ensuring an appropriate level of resolution was attempted by estimating movement rates of monarchs during their late migration (FR), and peak overwintering periods (COR). Migratory monarchs were marked in Arizona, and resighted in Pismo, California, and Mexican overwintering sites (Billings, 2019). Across all recaptures, the daily average distance traveled was 4.74km (2.95 mi). Therefore, a resolution of 4km was selected for FR predictors.

To estimate an appropriate level for which monarchs could select from for the COR (peak overwintering season) model, we referred to the available data from the literature. Two mark-resight studies have looked at overwintering monarch movement *between* sites during peak overwintering. Griffiths (2014) demonstrated overwintering monarchs moved between groves that were 0.3-1.9 miles apart. James & Kappen (2021) report individuals moved between groves, which ranged from 1.5-2.6 miles apart. These distances were larger than individual groves, and monarchs moved as far as these studies could have detected based on the location of monitored groves.

Predictive Data Layers:

FR data was acquired from the following sources. Monthly climatic data for temperature, VPD, dew point, and precipitation was sourced from PRISM, as monthly (average daily) values of 30-year normals from 1981-2010, and a resolution of ~4km (PRISM Climate Group, 2014). Monthly average Wind speed ($m\ s^{-1}$) data was sourced from WORLDCLIM 2.1, with data from 1970-2000 (Fick & Hijmans, 2017) with a resolution of 2.5mins (~5km). Average monthly cloud cover with ~1 km resolution was sourced from EarthENV, which generated the layer from remote sensing data sampled twice daily from 2000-2014 (Wilson & Jetz, 2016). Formatting of all spatial data was done using R and ArcMap 10.8 (ESRI, 2018; R Core Team, 2013). Raw predictor data was resampled in R bilinearly with the aggregate and disaggregate functions from the raster package to match the 4 km resolution of interest (Hijmans, 2022).

The percentage coverage for 10 different vegetation classifications come from the 2016 National Land Cover Database (NLCD) (Dewitz, 2019). The original data consists of 16 land cover classifications with 30 m^2 pixels (detailed descriptions of each of these classes can be found here). Categories were combined to compute percent coverage for 10 separate land classes at the 1km and 4km scales to reduce the number of predictors (Table S3). The final categories were calculated by counting the number of 30 m^2 pixels of each combined class within a 1 km and 4 km pixel across the COR extent in R and included percent coverage of: open water, naturally barren land, cultivated lands, low development, medium

development, high development, wetlands, shrubs and grassland, forest, and unclassified lands. Percent tree canopy was similarly calculated at 1 and 4 km resolutions from the NLCD 2016 USFS Tree Canopy Cover data (U.S. Forest Service, 2019).

Human footprint data were sourced from Leu et al., (2008), where impacts such as fragmentation, exotic plant invasion risk, etc. were combined and ranked over the western United States on a scale of 1-10, at a resolution of 4 km (Leu et al., 2008).

Topographic complexity was included in three different manners: Vector Ruggedness Measure (VRM), Topographic Ruggedness Index (TRI) and Topographic Position Index (TPI). Each was created by EarthENV by aggregating 250 m pixel cells in a moving window analysis to a 1 km resolution. Aggregated 1 km pixel values were calculated from the average of the surrounding cells for TRI and VRM, whereas the median value was used to estimate TPI for each raster cell (Amatulli et al., 2018). All layers were formatted using R (R Core Team, 2013).

Stream density was included as a predictor in two different formats: using all forms of surface water sources (including all stream types and stormwater drainages), and by including only non-ephemeral water sources likely to be available the entire overwintering season (excluding drainage lines, ephemeral streams, stormwater, and canal ditches). National Hydrography Dataset stream data or (NHD) (Moore et al., 2019; U.S. Geological Survey, 2020) was used to calculate density in ArcMap 10.6.

Table S2. Source and resolution of raw predictor data.

Raw Predictor Data	**Source**	**Original Spatial Resolution**	**Temporal Resolution**
Monthly Avg/min/max Temperature Monthly Min/Max VPD Monthly Avg Dew Point Monthly/Annual Precipitation	PRISM	~4km	30-year normals (1981-2010)
Monthly Average Wind Speed	WORLDCLIM 2.1	2.5 mins (~5km)	1970-2000
Monthly Average Cloud Cover	EarthENV	~1km	2000-2014
All Land Classifications (See Table S3) USFS Tree Canopy Cover	NLCD	30m2	2016
Streams	NHD	Vector Data	1970-2000
Topographic Complexity (TRI, TPI, & VRM)	EarthENV	~1km	NA

Table S3. Reclassification for National Landcover Database (NLCD) categories into our 10 Percent Cover Categories.

New Class/Group Name	Original NLCD Class Name
Open Water	Open Water
Barren Land; Natural	Perennial snow/ice
	Barren Land
Cultivated Land	Cultivated Crops
	Hay/ Pasture
Low Development	Developed, Open Space
Medium Development	Developed, Low Intensity
	Developed, Medium Intensity
High Development	Development, High Intensity
Wetlands	Woody Wetlands
	Emergent/Herbaceous Wetlands
Shrub and Grassland	Herbaceous (Grassland)
	Shrub/Scrub
Forrest	Deciduous Forest
	Evergreen Forest
	Mixed Forest
Unclassified	Unclassified

Table S4. Optimal Model Settings and Performance – performance of top five model setting combinations for FR and for COR candidate models. Model settings altered between runs were the Regularization Multiplier (from 1-5 in 0.5 increments) and Feature Classes including Linear (L), Quadratic (Q), Threshold (T), and Hinge (H). Average validation AUC for each Model is reported. Optimal model settings, denoted with a (*) were those with the lowest AICc score, or within 2 points of the lowest AICc score.

Model	Feature Classes	Regularization Multiplier	AICc	Δ AICc	AUC
FR	LT	1	9434.424	0.000*	0.9253
	LQT	1	9442.425	8.001	0.9252
	QT	1	9448.113	13.689	0.9190
	LQHT	1.5	9459.907	25.483	0.9257
	T	1	9460.546	26.122	0.9185
COR	LT	1.5	14742.142	0.000*	0.9624
	LQT	1.5	14743.562	1.421*	0.9619
	LQT	1	14750.800	8.659	0.9613
	LT	2	14751.521	9.380	0.9632
	LQT	2	14754.549	12.407	0.9624