Supplementary Material for "Host plants and climate structure habitat associations of the western monarch butterfly," by Dilts et al.

<u>Content included in supplementary material:</u>

- Additional methods (description of environmental covariates)
- Additional results (background selection, parameterization, variable reduction, and model validation)
- Supplementary Tables 1-7
- Supplementary Figure 1
- Supplementary Figure 2
- Supplementary Figure 3

Additional methods – Detailed description of environmental covariates

A key aspect of all species distribution models is the selection and preparation of environmental predictors that may influence the distribution of the species in question. Such variables typically include both biotic and abiotic layers such as land cover type, climatic variables, soil type, and topographic variables. We identified a set of 25 variables of potential ecological importance to milkweeds and/or monarch butterflies in the western U.S. From these, for each species, we selected a smaller subset of variables that were not correlated with each other, and used these in the final habitat suitability models (Steele et al., 2016).

Little has been published concerning environmental factors that influence the distribution of milkweeds, so most variables were selected based on expert ecological knowledge and data availability. We favored proximal variables (e.g., minimum temperature of the coldest month) rather than distal variables (e.g., elevation) that affect organisms less directly (Merow et al., 2013). The final set of candidate variables is provided in **Table 2** and is available upon request from the Xerces Society (monarchs@xerces.org). A few variables were more complex to calculate or require explanation as to their potential ecological significance, and these are described further here.

Climate data (PRISM)

Annual precipitation, maximum temperature of warmest month, mean annual temperature, mean temperature of warmest month, mean temperature of wettest month, minimum temperature of the coldest month, precipitation of coldest season, precipitation of the warmest season, temperature range were downloaded from the PRISM climate website (http://www.prism.oregonstate.edu/normals/). The coldest quarter was December through February and the warmest quarter was June through August. July was the warmest month and December was the coldest month.

Climate data (WORLDCLIM)

Precipitation seasonality and temperature seasonality were downloaded from the WORLDCLIM website (http://www.worldclim.org/current) as version 1.4.

Climate data (WNA)
Number of warming degree days was downloaded from the Climate WNS (Western North America) website (https://sites.ualberta.ca/~ahamann/data/climatewna.html).

Actual evapotranspiration
Stephenson (1998) defined actual evapotranspiration as “evaporative water loss from a site covered by a hypothetical standard crop, given the prevailing water availability”. This definition can be thought of as the potential for plant productivity given the simultaneous availability of both water and energy. The AET layer used in this study was developed by Dobrowski et al. (2013) and was downloaded from the AdaptWest website.

Mean climatic water deficit
Mean climatic water deficit represents the unmet atmospheric potential for evapotranspiration and can be used as a proxy for drought (Dobrowski et al., 2013). Stevens and Frey (2010) found that the Palmer Drought Severity Index is a major constraint on late-season monarch butterfly breeding distribution. Our use of mean climatic water deficit represents the hypothesis that aridity may be a major constraint on the geographic distribution of monarchs or milkweeds. CWD was downloaded from the AdaptWest website.

Number of arming degree days
Warming degree-days is a metric of heat accumulation through time that is often used in phenology analyses. Use of this metric was inspired by Stevens and Frey (2010), who included a map layer based on the minimum degree day accumulation necessary for monarch butterfly larval development for the last summer generation (August-September) to determine which areas would reliably produce the adults that migrate to the overwintering sites.

Precipitation seasonality
Precipitation seasonality is one of the original bioclimatic variables defined as used in Nix and Busby (1986) and Hijmans et al. (2005) and is defined as the coefficient of variation of monthly precipitation.

Temperature seasonality
Also an original bioclimatic variable, temperature seasonality is defined as the standard deviation of monthly temperature.

Soil variables
Percent clay, percent sand, percent silt, pH, and soil bulk density were included as candidate covariates in the models. We used the POLARIS soil dataset (Chaney et al., 2016) which is a gridded soil product that is derived from models based upon USDA SSURGO soil data. The SSURGO (Soil Survey Geographic Database) data product contains large gaps in the western U.S. making it unsuitable for regional modeling. In contrast, the STATSGO dataset (State Soil Geographic Database) is available across all of the western U.S. states but has a very large minimum mapping unit. Our use of POLARIS as the soil dataset is a change from Steele et al. (2016) who used STATSGO in their modeling effort.

Compound Topographic Index
The Compound Topographic Index, sometimes also referred to as the Topographic Wetness Index, is a measure of relative soil moisture potential based upon the upslope drainage area and slope (Beven and Kirkby, 1979) and is calculated as ln(upslope area / tan (slope)). We calculated it using scripts from The Nature Conservancy (Evans et al., 2014) and a 90 m digital elevation model. In arid regions of the west, *A. speciosa* has been thought to be generally limited to relatively moist areas such as riparian areas or ditches.

Distance to water
Like the Compound Topographic Index, distance to water was included because of the observed association of *A. speciosa* and relatively moist areas. It was calculated from the medium resolution National Hydrography Dataset (U.S. Geological Survey 2015) because this resolution is more consistent across the west than the high resolution National Hydrography Dataset and because the medium resolution data was adequate to our 270 m analytical scale. The current modeling effort differed from Steele et al. (2016) in that we considered distance to both perennial water and distance to intermittent water separately. Both datasets were constructed from the USGS National Hydrography Dataset stream data with perennial being defined using the FCode 46006. For water bodies, intermittent was defined using 39001 and 36100. Distance to perennial water was the distance to the nearest source of perennial water, either stream or polygonal water body. Distance to intermittent water was similarly constructed.

Reclassified LANDFIRE Existing Vegetation Type Layer (i.e. Land Cover)
We simplified the categorical LANDFIRE layer depicting Existing Vegetation Type (LANDFIRE 2013) by cross walking it to fewer, lumped classes deemed to be of likely ecological significance for milkweeds and monarch butterflies. **Supplementary Table 3** shows which land cover types were grouped together.

Milkweed models as covariates for the monarch butterfly breeding model
We used our final, calibrated milkweed habitat suitability models as inputs to the model of monarch butterfly breeding habitat suitability, along with the environmental variables listed in Table 2. Milkweed presence is a critical variable in monarch butterfly breeding habitat suitability, and the maps that we produced are the best available west-wide estimate of milkweed distribution.

Resolution of environmental covariates
The native resolution of the candidate environmental rasters varied from 30 m pixels to 1 km pixels. To use MaxEnt, it was necessary to aggregate or downsample all layers until they had the same grain size. Merow et al. (2013) recommend that resolutions should be chosen that provide data from proximal rather than distal variables. In our case, we wanted to balance several factors: 1) the ecological scale of processes that determine suitability for these species, 2) the spatial accuracy of our species presence locations, 3) the native resolution of environmental variables, and 4) computational constraints (finer-grained rasters greatly increase processing time). We conducted a set of initial experiments using three different resolutions: 90 m, 270 m, and 900 m, and ultimately found that 270 m provided the best compromise between the factors. All data sets were downloaded and projected into the same geographic coordinate system. To attain a standard resolution, variables with a native resolution that was coarser than 270 m were downsampled

with the ArcMap Resample tool using the nearest neighbor setting, to avoid interpolation and false precision. Continuous rasters with a native resolution of 30 m or 90 m were aggregated to 270 m pixels using the mean value. The reclassified LANDCOVER layer was aggregated 270 m resolution using the majority option. Once all the layers were aggregated or downsampled to 270 m, we used ArcMap to snap them together and clip them to the same extent.

REFERENCES

Beven, K.J., & Kirkby, M.J. (1979) A physically based, variable contributing area model of basin hydrology, *Hydrolological Sciences Bulletin*, 24, 43–69.

Chaney, N.W., Wood, E.F., McBratney, A.B., Hempel, J.W., Nauman, T.W., Brungard, C.W., & Odgers, N.P. (2016) POLARIS: A 30-meter probabilistic soil series map of the contiguous United States. *Geoderma*, 274, 54-67.

Dobrowski, S.Z., Abatzoglou, J., Swanson, A.K., Greenberg, J.A., Mynsberge, A R., Holden, Z.A., & Schwartz, M.K. (2013) The climate velocity of the contiguous United States during the 20th century. *Global Change Biology*, 19(1), 241-251.

Evans, J.S., Oakleaf, J., Cushman, S.A., & Theobald, D. (2014) An ArcGIS Toolbox for Surface Gradient and Geomorphometric Modeling, version 2.0-0. Available: http://evansmurphy.wix.com/evansspatial. Accessed: 2015 Dec 2nd.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965-1978.

Nix, H. A., & Busby, J. (1986). BIOCLIM, a bioclimatic analysis and prediction system. Division of Water and Land Resources: Canberra.

Rollins, M.G. (2009) LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. *International Journal of Wildland Fire*, 18(3), 235-249.

Steele, M., Engler, J.D., Cruz, L., Jepsen, S., Black, S., Taylor, A., & Jones, A. (2016) Phase I: Western Monarch and Milkweed Habitat Suitability Modeling Project- MaxEnt Model Outputs. Available: https://ecos.fws.gov/ServCat/Reference/Profile/56630. Accessed: 2017 Jan 19

Stephenson, N. (1998). Actual evapotranspiration and deficit: biologically meaningful correlates of vegetation distribution across spatial scales. *Journal of Biogeography*, 25(5), 855-870.

Stevens, S.R., & Frey, D.F. (2010) Host plant pattern and variation in climate predict the location of natal grounds for migratory monarch butterflies in western North America. *Journal of Insect Conservation*, 14(6), 731-744.

Additional results – From background selection, model parameterization, variable reduction, and model validation steps

Based upon the recommendations of Anderson and Gonzalez (2011) we chose to perform species-specific optimization of a number of parameters including background area, regularization, and feature type. Background area is important in a presence-only modeling framework like Maxent because presences are contrasted with random background points and changing the size of the background can change the outcome of the model in sometimes dramatic ways (Anderson and Raza, 2010; Merow et al., 2013). Following the approaches of VanDerWal et al. (2009) we generated models at varying buffer sizes and then following Iturbide et al. (2015) sought to determine a saturation threshold using a Michaelis-Menten function that simultaneously optimizes both model fit while minimizing the background extent. In our study, species tended to have optimal background areas that varied from 270 to 300 km buffer sizes (**Supplementary Table 1**). At small buffer sizes (e.g. 10 to 100 km), models tended to have suboptimal performance. At larger buffer sizes models performed increasingly well, but the increase in improvement for each subsequently diminished. We did not find a true plateau in AUC with increasing buffer size, which we attribute to using a randomly-withheld validation dataset. Hence we used the 95% confidence interval of the saturation level to determine our optimal buffer.

After determining an optimal background size we tuned our Maxent models assessing five different feature types (linear, linear+quadratic, linear+quadratic+product, linear+quadratic+product+threshold, hinge) and five different levels of regularization (smoothing) in a factorial manner resulting in twenty-five models per species (Anderson et al., 2000; Merow et al., 2013; Radosavljevic et al., 2014). Models were assessed using validation AUC, AUCdiff (training AUC – validation AUC), and a new metric that combines Validation AUC and AUCdiff introduced here called penalized AUC or pAUC. pAUC is defined as (validation AUC - (training AUC – validation AUC)). pAUC is based on the assumption that model overfit (as measured by AUCdiff) is equal in proportion to model fit. Hence a difference in pAUC between two models can either be driven by higher overfit or lower model fit. In contrast to approaches such as Akaike's Information Criterion and Bayesian Information Criterion there is no parameter for sample size.

Our results found that the following milkweeds had the best models *A. asperula*, *A. subulata*, *A. eriocarpa*, *A. californica*, and *A. speciosa* (**Supplementary Figure 1**). The monarch model tended to rank in the middle with decent fit and minimal overfit. On the other end of the spectrum, species such as *A. cryptoceras*, *A. incarnata*, *A. erosa*, and *A. viridiflora* had the poorest models in terms of fit and overfit. The use of pAUC further separated the better milkweed models, which tended to be built upon a greater number of training points, from the poorer performing milkweed models, which tended to have few training points, because the poorer performing models often tended to be quite overfit. When comparing best models, models that used validation AUC differed dramatically from those chosen using pAUC. When using pAUC as the criterion for selecting a model, the best model varied among a range of feature types and regularization parameters (**Supplementary Table 4**). In contrast, had validation AUC been used to select models then the outcome would have been very different with all models having a regularization of either 1 or 2 and all but two models having a regularization value of 1.

Our findings are in congruence with the many authors who have called attention to the tendency for Maxent to produce overfit models and have suggested evaluating models in a fashion that incorporates penalties for overfitting (Warren and Seifert, 2011; Merow et al., 2013; Shcheglovitova and Anderson, 2013).

Our final step in the model building process was to reduce the number of parameters in order to aid in model interpretation. We employed an iterative approach in which variables were removed if they contributed less than 3% as measured by the Maxent permutation importance of if they had a Pearson's correlation greater than 0.7 with another higher ranking variable. We initially built a model using all variables and then re-built models after removing variables that met the above criteria. We repeated this process but used 10% permutation importance as the final cutoff to ensure highly interpretable and parsimonious models. Typically, it took between one and five iterations to get to a final model (**Supplementary Table 5**).

Model overfit as measured by AUCdiff was highest for *A. cryptoceras*, *A. incarnata*, *A. tuberosa*, and *A. asperula*, all of which had AUCdiff greater than 0.05. AUCdiff was strongly correlated with sample size. pAUC, which is introduced in this paper, represents a novel method for combining validation AUC and AUCdiff into a single measure. Although pAUC is functionally related to both validation AUC and AUCdiff, when the models were ranked by validation AUC and pAUC there were important differences. pAUC ranked *A. incarnata, A. cryptoceras*, *A. tuberosa*, *A. cordifolia*, and *A. asperula* lower than they were ranked based on validation AUC. These were all models with small sample sizes suggesting that pAUC, by incorporating AUCdiff, accounts functionally for sample size without explicitly incorporating a sample size penalty in the way that the Akaike or Bayesian Information Criterion would (Burnham and Anderson, 2003).

REFERENCES

Anderson, R.P., & Gonzalez Jr, I. (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. *Ecological Modelling*, *222*(15), 2796-2811.

Anderson, R.P., & Raza, A. (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus Nephelomys) in Venezuela. *Journal of Biogeography*, *37*(7), 1378-1393.

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., & Gutiérrez, J.M. (2015) A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*, *312*, 166-174.

Merow, C., Smith, M.J., & Silander Jr, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, *36*(10), 1058-1069.

Radosavljevic, A., & Anderson, R.P. (2014) Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography*, *41*(4), 629-643.

Shcheglovitova, M., & Anderson, R.P. (2013) Estimating optimal complexity for ecological niche models: a jackknife approach for species with small sample sizes. *Ecological Modelling*, *269*, 9-17.

Steele, M., Engler, J.D., Cruz, L., Jepsen, S., Black, S., Taylor, A., & Jones, A. (2016) Phase I: Western Monarch and Milkweed Habitat Suitability Modeling Project- MaxEnt Model Outputs. Available: https://ecos.fws.gov/ServCat/Reference/Profile/56630. Accessed: 2017 Jan 19

VanDerWal, J., Shoo, L.P., Graham, C., & Williams, S.E. (2009) Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological Modelling*, *220*(4), 589-594.

Warren, D.L., & Seifert, S.N. (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, *21*(2), 335-342.

Supplementary Table 1: Sources of the occurrence data used in the study.

Source	Count
Southwest Monarch Study	5678
Arizona State University, SEINet	5515
United States Geological Service; BISON	4897
Xerces Society	4732
California Consortium	3283
U.S. Fish and Wildlife Service	3114
University of California, Davis	1951
College of Western Idaho	1398
iNaturalist	1341
Journey North	1318
Global Biodiversity Information Facility	937
University of California, Berkley	739
CalFlora	617
Washington Department of Fish and Wildlife	590
Idaho Department of Fish and Game	537
Oregon State University	488
National Phenology Network	415
U.S. Forest Service	410
Boise Parks & Recreation	314
Monarch Larva Monitoring Project	304
U.S. National Park Service	293
PNW Consortium	265
USDA PLANTS Database	257
private citizens	192
Flickr	185
University of Washington	162
California State University, Chico	154
Wyoming Biodiversity Citizen Science Initiative	78
Utah State University	76
Iowa State University, Department of Entomology	58
Bureau of Land Management	44
Western Monarch listserv	44
Friends of Red Rock Canyon	41
University of British Columbia	29
U.S. Navy	29
California Polytechnic State University	26
University of Montana	19
Humble Roots Nursery	17
University of Arizona Herbarium	16

University of Alabama	12
City of Eugene Parks and Open Space	10
University of Alberta Museums	10
University of Nevada Herbarium	8
Rombough Biological	7
U.S. Bureau of Land Management	7
Missouri Botanical Garden	6
State of Oregon	6
Washington Butterfly Association	5
Natural Resources Conservation Service	4
University of Connecticut	3
University of Kansas Biodiversity Institute	2
DPLEX listserv	2
The Nature Conservancy	1

Supplementary Table 2: Model name, number of records prior to applying geographic thinning, number of records after applying geographic thinning, and the ratio of pre to post-thinning records. In addition to the thirteen milkweeds that were modeled there were 44 additional milkweed species in the database that were not included in the model due to small samples sizes.

Model	Pre-thinning records	Post-thinning records	Ratio
Danaus plexippus - all records	8427	732	11.5
Danaus plexippus - adult records	5236	924	5.7
Danaus plexippus - breeding records	1498	635	2.4
Danaus plexippus - breeding records without *Ascelpias curassavica*	n/a	586	
Ascelpias speciosa	9256	1219	7.6
Ascelpias fascicularis	3260	226	14.4
Ascelpias subulata	609	47	13.0
Ascelpias eriocarpa	838	42	20.0
Ascelpias californica	637	40	15.9
Ascelpias asperula	1511	39	38.7
Ascelpias tuberosa	695	37	18.8
Ascelpias viridiflora	246	35	7.0
Ascelpias erosa	532	34	15.6
Ascelpias subverticillata	1448	33	43.9
Ascelpias cordifolia	758	30	25.3
Ascelpias cryptoceras	640	24	26.7
Ascelpias incarnata	266	21	12.7
Species not modeled			
Asclepias involucrata	410		
Asclepias linaria	386		
Asclepias latifolia	378		
Asclepias albicans	373		
Asclepias nyctaginifolia	294		
Asclepias incarnata	266		
Asclepias viridiflora	246		
Asclepias pumila	223		
Asclepias engelmanniana	212		
Asclepias curassavica	196		
Asclepias vestita	175		
Asclepias brachystephana	151		
Asclepias macrosperma	148		
Asclepias labriformis	136		
Asclepias solanoana	136		
Asclepias nummularia	125		
Asclepias macrotis	124		
Asclepias hallii	120		
Asclepias uncialis	111		

Asclepias hypoleuca	95
Asclepias elata	85
Asclepias arenaria	83
Asclepias lemmonii	81
Asclepias verticillata	71
Asclepias rusbyi	70
Asclepias michauxii	65
Asclepias oenotheroides	50
Asclepias quinquedentata	43
Asclepias glaucescens	42
Asclepias stenophylla	28
Asclepias welshii	24
Asclepias eastwoodiana	21
Asclepias syriaca	18
Asclepias sanjuanensis	17
Asclepias cutleri	12
Asclepias sperryi	10
Asclepias linearis	2
Asclepias ovalifolia	2
Asclepias emoryi	1
Asclepias lanceolata	1
Asclepias purpurascens	1
Asclepias scaposa	1
Asclepias sullivantii	1
Asclepias viridis	1

Supplementary Table 3: Landfire land cover classes as reclassified into forty-six cover types for this study. Value refers to the value field in the Landfire dataset.

Value	Lumped Class Name	Lumped ID	Value	Lumped Class Name	Lumped ID
3969	Agricultural-Aquaculture	1	3184	Exotic Herbaceous	17
3989	Agricultural-Aquaculture	1	3182	Exotic Herbaceous-Upland	18
3968	Agricultural-Graminoid	2	3183	Exotic Herbaceous-Upland	18
3978	Agricultural-Graminoid	2	3181	Exotic Herbaceous-Upland	18
3988	Agricultural-Graminoid	2	3259	Exotic Tree-Shrub	19
3966	Agricultural-Graminoid	2	3180	Exotic Tree-Shrub	19
3986	Agricultural-Graminoid	2	3134	Grassland	20
3967	Agricultural-Graminoid	2	3142	Grassland	20
3977	Agricultural-Graminoid	2	3130	Grassland	20
3987	Agricultural-Graminoid	2	3135	Grassland	20
3961	Agricultural-high structure	3	3147	Grassland	20
3981	Agricultural-high structure	3	3503	Grassland	20
3962	Agricultural-high structure	3	3133	Grassland	20
3982	Agricultural-high structure	3	3256	Grassland	20
3960	Agricultural-high structure	3	3141	Grassland	20
3980	Agricultural-high structure	3	3132	Grassland	20
3965	Agricultural-Row Crop	4	3148	Grassland	20
3975	Agricultural-Row Crop	4	3149	Grassland	20
3985	Agricultural-Row Crop	4	3150	Grassland	20
3964	Agricultural-Row Crop	4	3195	Grassland	20
3984	Agricultural-Row Crop	4	3143	Grassland-Alpine	21
3963	Agricultural-Row Crop	4	3136	Grassland-Alpine	21
3983	Agricultural-Row Crop	4	3144	Grassland-Alpine	21
3294	Barren	5	3171	Grassland-Alpine	21
3260	Deciduous closed tree canopy	6	3068	Grassland-Alpine	21
3264	Deciduous closed tree canopy	6	3071	Grassland-Alpine	21
3266	Deciduous closed tree canopy	6	3067	Grassland-Alpine	21
3262	Deciduous closed tree canopy	6	3070	Grassland-Alpine	21
3009	Deciduous open tree canopy	7	3131	Grassland-Coastal	22
3011	Deciduous open tree canopy	7	3129	Grassland-Coastal	22
3012	Deciduous open tree canopy	7	3138	Grassland-Montane	23
3013	Deciduous open tree canopy	7	3139	Grassland-Montane	23
3008	Deciduous open tree canopy	7	3146	Grassland-Montane	23
3201	Deciduous open tree canopy	7	3137	Grassland-Subalpine	24
3237	Deciduous open tree canopy-Montane	8	3145	Grassland-Subalpine	24
3236	Deciduous open tree canopy-Subalpine	9	3140	Grassland-Subalpine	24
3112	Deciduous sparse tree canopy	10	3261	Mixed evergreen-deciduous closed tree canopy	25
3295	Developed	11	3265	Mixed evergreen-deciduous closed tree canopy	25
3298	Developed	11	3267	Mixed evergreen-deciduous closed tree canopy	25
3296	Developed	11	3263	Mixed evergreen-deciduous closed tree canopy	25
3297	Developed	11	3063	Mixed evergreen-deciduous closed tree canopy	25
3299	Developed	11	3061	Mixed evergreen-deciduous open tree canopy	26
3903	Developed	11	3062	Mixed evergreen-deciduous open tree canopy	26
3908	Developed	11	3157	Mixed evergreen-deciduous open tree canopy	26
3913	Developed	11	3156	Mixed evergreen-deciduous open tree canopy-Lowland	27
3924	Developed	11	3158	Mixed evergreen-deciduous open tree canopy-Montane	28
3929	Developed	11	3154	Mixed evergreen-deciduous open tree canopy-Montane	28
3934	Developed	11	3159	Mixed evergreen-deciduous open tree canopy-Montane	28
3904	Developed	11	3255	Mixed evergreen-deciduous shrubland	29
3909	Developed	11	3251	Mixed evergreen-deciduous shrubland-Montane	30
3914	Developed	11	3170	Mixed evergreen-deciduous sparse tree canopy	31
3923	Developed	11	3113	Mixed evergreen-deciduous sparse tree canopy	31
3928	Developed	11	3118	Mixed evergreen-deciduous sparse tree canopy	31
3900	Developed	11	3120	Mixed evergreen-deciduous sparse tree canopy-Upland	32
3910	Developed	11	3292	Open Water	33
3920	Developed	11	3488	Riparian	34
3925	Developed	11	3495	Riparian	34
3940	Developed	11	3504	Riparian	34
3945	Developed	11	3163	Riparian	34
3901	Developed	11	3164	Riparian	34
3911	Developed	11	3254	Riparian	34
3921	Developed	11	3257	Riparian	34
3926	Developed	11	3253	Riparian	34
3941	Developed	11	3258	Riparian	34
3946	Developed	11	3385	Riparian	34
3902	Developed	11	3162	Riparian	34
3907	Developed	11	3151	Riparian	34
3912	Developed	11	3155	Riparian	34
3922	Developed	11	3152	Riparian-Montane	35
3927	Developed	11	3252	Riparian-Subalpine	36
3942	Developed	11	3160	Riparian-Subalpine	36
3947	Developed	11	3123	Shrubland	37
3177	Evergreen closed tree canopy	12	3212	Shrubland	37
3037	Evergreen closed tree canopy	12	3125	Shrubland	37

3039	Evergreen closed tree canopy	12
3018	Evergreen closed tree canopy	12
3047	Evergreen closed tree canopy	12
3166	Evergreen closed tree canopy	12
3052	Evergreen closed tree canopy	12
3045	Evergreen closed tree canopy	12
3051	Evergreen closed tree canopy	12
3041	Evergreen closed tree canopy	12
3058	Evergreen closed tree canopy	12
3055	Evergreen closed tree canopy	12
3056	Evergreen closed tree canopy	12
3178	Evergreen closed tree canopy	12
3043	Evergreen closed tree canopy	12
3230	Evergreen closed tree canopy	12
3035	Evergreen closed tree canopy	12
3028	Evergreen closed tree canopy	12
3208	Evergreen closed tree canopy	12
3232	Evergreen closed tree canopy	12
3231	Evergreen closed tree canopy	12
3050	Evergreen closed tree canopy	12
3167	Evergreen closed tree canopy	12
3205	Evergreen closed tree canopy	12
3036	Evergreen closed tree canopy	12
3042	Evergreen closed tree canopy	12
3174	Evergreen closed tree canopy	12
3014	Evergreen open tree canopy	13
3034	Evergreen open tree canopy	13
3200	Evergreen open tree canopy	13
3060	Evergreen open tree canopy	13
3206	Evergreen open tree canopy	13
3172	Evergreen open tree canopy	13
3173	Evergreen open tree canopy	13
3027	Evergreen open tree canopy	13
3203	Evergreen open tree canopy	13
3017	Evergreen open tree canopy	13
3202	Evergreen open tree canopy	13
3023	Evergreen open tree canopy	13
3049	Evergreen open tree canopy	13
3019	Evergreen open tree canopy	13
3016	Evergreen open tree canopy	13
3025	Evergreen open tree canopy	13
3059	Evergreen open tree canopy	13
3053	Evergreen open tree canopy	13
3054	Evergreen open tree canopy	13
3179	Evergreen open tree canopy	13
3032	Evergreen open tree canopy	13
3048	Evergreen open tree canopy	13
3161	Evergreen open tree canopy	13
3029	Evergreen open tree canopy	13
3030	Evergreen open tree canopy-Montane	14
3114	Evergreen open tree canopy-Montane	14
3024	Evergreen open tree canopy-Montane	14
3026	Evergreen open tree canopy-Montane	14
3227	Evergreen open tree canopy-Montane	14
3234	Evergreen open tree canopy-Montane	14
3235	Evergreen open tree canopy-Montane	14
3031	Evergreen open tree canopy-Montane	14
3228	Evergreen open tree canopy-Montane	14
3021	Evergreen open tree canopy-Montane	14
3022	Evergreen open tree canopy-Montane	14
3233	Evergreen open tree canopy-Subalpine	15
3020	Evergreen open tree canopy-Subalpine	15
3057	Evergreen open tree canopy-Subalpine	15
3038	Evergreen open tree canopy-Subalpine	15
3044	Evergreen open tree canopy-Subalpine	15
3015	Evergreen open tree canopy-Subalpine	15
3046	Evergreen open tree canopy-Subalpine	15
3229	Evergreen open tree canopy-Subalpine	15
3033	Evergreen open tree canopy-Subalpine	15
3165	Evergreen sparse tree canopy	16
3115	Evergreen sparse tree canopy	16
3119	Evergreen sparse tree canopy	16
3116	Evergreen sparse tree canopy	16
3117	Evergreen sparse tree canopy	16
3220	Shrubland	37
3080	Shrubland	37
3085	Shrubland	37
3210	Shrubland	37
3078	Shrubland	37
3105	Shrubland	37
3110	Shrubland	37
3097	Shrubland	37
3099	Shrubland	37
3103	Shrubland	37
3104	Shrubland	37
3108	Shrubland	37
3214	Shrubland	37
3215	Shrubland	37
3216	Shrubland	37
3101	Shrubland	37
3074	Shrubland	37
3087	Shrubland	37
3217	Shrubland	37
3065	Shrubland	37
3082	Shrubland	37
3127	Shrubland	37
3211	Shrubland	37
3090	Shrubland	37
3091	Shrubland	37
3109	Shrubland	37
3100	Shrubland	37
3076	Shrubland	37
3121	Shrubland	37
3122	Shrubland	37
3153	Shrubland	37
3213	Shrubland	37
3064	Shrubland	37
3072	Shrubland	37
3079	Shrubland	37
3124	Shrubland	37
3095	Shrubland	37
3204	Shrubland	37
3250	Shrubland	37
3075	Shrubland	37
3081	Shrubland	37
3088	Shrubland	37
3066	Shrubland	37
3093	Shrubland	37
3094	Shrubland	37
3077	Shrubland	37
3096	Shrubland-Coastal	38
3128	Shrubland-Coastal	38
3092	Shrubland-Coastal	38
3107	Shrubland-Montane	39
3126	Shrubland-Montane	39
3098	Shrubland-Montane	39
3168	Shrubland-Montane	39
3083	Shrubland-Montane	39
3084	Shrubland-Montane	39
3086	Shrubland-Montane	39
3106	Shrubland-Montane	39
3169	Shrubland-Subalpine	40
3186	Shrubland-Upland	41
3293	Snow-Ice	42
3001	Sparsely Vegetated	43
3002	Sparsely Vegetated	43
3003	Sparsely Vegetated	43
3004	Sparsely Vegetated	43
3007	Sparsely Vegetated	43
3218	Sparsely Vegetated	43
3219	Sparsely Vegetated	43
3221	Sparsely Vegetated	43
3223	Sparsely Vegetated	43
3006	Sparsely Vegetated-Alpine	44
3222	Sparsely Vegetated-Alpine	44
3944	Undeveloped Ruderal Grassland	45
3949	Undeveloped Ruderal Grassland	45
3943	Undeveloped Ruderal Shrubland	46
3948	Undeveloped Ruderal Shrubland	46

Supplementary Table 4: Optimal buffer distance derived from the Michaelis-Menten function.

Species	**Buffer Distance (km)**	**AUC**
A.subverticillata	170	0.848
A.asperula	180	0.962
A.incarnata	190	0.843
A.cryptoceras	200	0.838
A.erosa	210	0.801
A.viridflora	240	0.904
A.cordifolia	250	0.889
A.eriocarpa	270	0.944
A.californica	270	0.828
A.fascicularis	280	0.917
A.subulata	300	0.947
A.tuberosa	300	0.921

Supplementary Table 5: Feature type and regularization parameters of the highest penalized AUC model and model with the highest validation AUC. Feature types are (L = linear, LQ = linear+quadratic, LQP = linear+quadratic+product, LQPT = linear+quadratic+product_threshold, H = hinge)

Model	pAUC type	pAUC reg.	vAUC type	vAUC reg.
Danaus plexippus – all	LQP	2	LQPT	1
Danaus plexippus – adults	LQPT	5	LQPT	1
Danaus plexippus – breeding	H	5	LQPT	1
Danaus plexippus – breeding w/o *A. curassavica*	LQPT	4	LQPT	1
Asclepias speciosa	H	3	H	1
Asclepias fascicularis	LQP	5	H	1
Asclepias subulata	H	5	LQPT	2
Asclepias eriocarpa	LQP	1	LQP	1
Asclepias californica	H	4	H	1
Asclepias asperula	LQPT	1	LQPT	1
Asclepias tuberosa	LQP	2	LQPT	1
Asclepias viridiflora	LQP	1	LQPT	1
Asclepias erosa	LQP	2	H	1
Asclepias subverticillata	H	1	H	1
Asclepias cordifolia	H	3	LQPT	1
Asclepias cryptoceras	H	1	H	1
Asclepias incarnata	H	2	H	2

Supplementary Table 6: Number of iterations in which the variable was retained.

	Danaus plexippus – all	*Danaus plexippus* – adults	*Danaus plexippus* – breeding	*Danaus plexippus* – breeding w/o *A. curassavica*	*Asclepias speciosa*	*Asclepias fascicularis*	*Asclepias subulata*	*Asclepias eriocarpa*	*Asclepias californica*	*Asclepias asperula*	*Asclepias tuberosa*	*Asclepias viridiflora*	*Asclepias erosa*	*Asclepias subverticillata*	*Asclepias cordifolia*	*Asclepias cryptoceras*	*Asclepias incarnata*
actual evapotranspiration								3		2	4						
annual precipitation						3									2		
climatic water deficit		3	1		2			3	3							1	
maximum temperature of warmest month					2												
mean annual temperature												2					
mean temperature of the warmest quarter		3	4				2										
mean temperature of the wettest month																	
minimum temperature of the coldest	4			2						2	4		2				
number of degree days						2											
precipitation of coldest season			2						3				2	1			1
precipitation of warmest season	2			2	2	5		3			3	2	2				
precipitation seasonality								3	2	2							
temperature range						1				2							
temperature seasonality	2	1	1			4		1				2		3			1
percent clay																	
percent sand										2							
percent silt																	
pH																	
soil bulk density																	
aspect							1							2	1		2
compound topographic index					1							2		3			3
distance to intermittent water	1		1					2			1		2				
distance to perennial water						5		1		1				2	1		4
land cover		3	4		2		2		1	2	4		1		2	1	4
slope												1	2	3	2	1	4
Asclepias speciosa	4		2	2													
Asclepias fascicularis	4	3	4	2													
Asclepias subulata																	
Asclepias eriocarpa				1													
Asclepias californica																	
Asclepias asperula																	
Asclepias tuberosa																	
Asclepias viridiflora	1																
Asclepias erosa																	
Asclepias subverticillata	4		4	2													
Asclepias cordifolia																	
Asclepias cryptoceras																	
Asclepias incarnata																	

Supplementary Table 7: Equal specificity vs. sensitivity threshold obtained from validation data for dividing relative habitat suitability maps into binary suitable and non-suitable areas.

Danaus plexippus – all records	0.35
Danaus plexippus – adults only	0.34
Danaus plexippus – breeding only	0.33
Danaus plexippus – breeding w/o *A. curassavica*	0.34
Asclepias speciosa	0.39
Asclepias fascicularis	0.40
Asclepias subulata	0.50
Asclepias eriocarpa	0.58
Asclepias californica	0.56
Asclepias asperula	0.42
Asclepias tuberosa	0.32
Asclepias viridiflora	0.54
Asclepias erosa	0.50
Asclepias subverticillata	0.54
Asclepias cordifolia	0.54
Asclepias cryptoceras	0.27
Asclepias incarnata	0.26