Workflow	Category	Necessary Reporting	Method Used	Details	Citation
(A) Obtaining and processing occurrence data	metadata	(A1) source of occurrence data	GBIF	We utilized the Global Biodiversity Information Facility (GBIF) to assess the current and predicted future distribution of E. rosea in the United States. A total of 1,879 records were downloaded on November 12, 2023.	GBIF.org (12 November 2023) GBIF Occurrence Download https://doi.org/10.15468/dl.9th8ab
		(A2) download date; version of data source	creation date (11/12/2023 2:11:28 AM), simple tab-separated values (TSV) download format		
		(A3) basis of records	Human Observation (presence only)	These records were checked for accuracy and quality and are considered reliable observations in the large majority of cases (Maldonado et al. 2015).	Maldonado C, Molina CI, Zizka A, Persson C, Taylor CM, Albán J, Chilquillo E, Rønsted N, Antonelli A (2015) Estimating species diversity and distribution in the era of Big Data: To what extent can we trust public databases? Global Ecology and Biogeography 24: 973–984. https://doi.org/10.1111/geb.12326
		(A4) spatial extent	Contiguous United States	Our environmental datasets spanned the entirety of North America. Occurances outside the extent of our environmental layers are irrelivant for the purpose of this study which aims to investigate E.rosea's potential range expansion via natural dispersal.	Barve, N., V. Barve, A. Jiménez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberón, and F. Villalobos. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modelling 222: 1810–1819.
		(A5) temporal range	2011-2023		
	processing	(A6-1) duplicate coordinates	Deleted	447 duplicated points were removed. (1,432 total)	Radosavljevic, A., and R. P. Anderson. 2014. Making better Maxent models of species distributions: complexity, overfitting and evaluation. Journal of Biogeography 41: 629–643.
		(A6-2) spatial/environmental outlier; error	Outliers manually removed	E. rosea observations were plotted in ArcGIS pro. Any unisual points falling far outside the population cluster were investigated. Two occurance records were located in commercial garden centers in Ohio and New York and removed from our datatset. (1,430 total)	

		(A6-3) spatial/coordinate uncertainty	1000 meters or fewer uncertanty	The recorded coordinates of a data point may not necessarily correspond to its exact collection location due to differences in specificity levels. To ensure accuracy, we removed all data points with coordinate specificity greater than 1000 meters. In total, 383 occurance record were removed due to inadequate spesicifty. (1,047 total)	
		(A7-1) sampling bias	Spatial Thinning	In order to mitigate the impact of sampling bias, we applied a spatial filter to the occurrence dataset to ensure that no two locations were within a 10 km radius of each other. 473 localities shared the same grid cell. (574 total)	
(B) Obtaining and processing environmental data	metadata & processing	(B1) source	AdaptWest Project. 2022. Gridded current and projected climate data for North America at 1km resolution, generated using the ClimateNA v7.30 software (T. Wang et al., 2022). Available at adaptwest.databasin.org	Current and projected climate data were acquired from the Adapt West Project (AdaptWest Project 2022), comprising 33 parameters evaluated for their relevance in predicting E. rosea presence.	AdaptWest Project. 2022. Gridded current and projected climate data for North America at 1km resolution, generated using the ClimateNA v7.30 software (T. Wang et al., 2022). Available at adaptwest.databasin.org
		(B2) download date; version of data source	11/11/2023		
		(B3) spatial resolution	1km resolution	(0.015485, 0.015485)(x,y)	
		(B4) temporal range	2000-2020		
			2021-2040		
(C) Model calibration	data input	(C1) number of background data	10000	For each geographical background we selected 10,000 random cells that did not hold a species presence record.	Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231-259.

		(C2) sampling method for background data	Random Seed, crossvalidate	By selecting "random seed" a different random sample will be used to for valudating the model	Jarnevich, C. S., Talbert, M., Morisette, J., Aldridge, C., Brown, C. S., Kumar, S., Manier, D., Talbert, C., & Holcombe, T. (2017). Minimizing effects of methodological decisions on interpretation and prediction in species distribution studies: An example with background selection. Ecological Modelling, 363, 48–56. https://doi.org/10.1016/j.ecolmodel.2017.08.017
		(C3) variable selection	Bio26, Bio27, Bio33	All 33 bioclimatic layers were analyzed for their relatedness using the R-package "ENMeval" and a correlation matrix was generated using the function "raster.cor.matrix" (Kass et al. 2021). The results of this matrix allowed us to determine which variables could be disregarded because they contributed mostly redundant data to our model and could lead to overfitting.	Kass, J. M., Muscarella, R., Galante, P. J., Bohl, C. L., Pinilla-Buitrago, G. E., Boria, R. A., Soley-Guardia, M., & Anderson, R. P. (2021). ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. Methods in Ecology and Evolution, 12(9), 1602–1608. https://doi.org/10.1111/2041-210X.13628
	algorithm	(C4) name	Maximum Entropy Modeling of Species Geographic Distributions	ENM was performed using the maximum entropy approach as implemented in MAXENT 3.4.4 (Phillips, Anderson, & Schapire, 2006).	Steven J. Phillips, Miroslav Dudík, Robert E. Schapire. [Internet] Maxent software for modeling species niches and distributions (Version 3.4.1). Available from url: http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed on 2023-11-17.
		(C5) version of algorithm and software	MaxEnt 3.4.4		
		(C6) parameterization	ENMevaluate	The optimal model parameters were tuned using the function ENMevaluate in the package ENMeval for R. Within AENMevaluate we evaluated models with the following feature classes: linear, quadratic, product and, and hinge and regularization multipliers between 0.5 and 3.	Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models . Methods in Ecology and Evolution, 5(11), 1198–1205. https://doi.org/10.1111/2041-210x.12261

(D) Model transfer and evaluation	evaluation	(D1) evaluation index	AUC	We evaluated the performance of the models by two different methods using a randomized subset of occurrences for model evaluation: (a) an omission error test and (b) the receiver operating characteristic (ROC) curve	Boyce, M.S., P.R. Vernier, S.E. Nielsen and F.K.A. Schmiegelow. 2002. Evaluating resource selection functions. Ecol. Model., 157, 281-300. Hirzel, A.H., G. Le Lay, V. Helfer, C. Randin and A. Guisan. 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecol. Model., 199, 142-152.
		(D2) threshold for evaluation index	10% train/test omission rate.	Threshold at 10% to create the binary suitability map, we expect the omission rate in the test data to be 10%. Higher omission in the testing data reflects over-fitting (noise and/or bias).	Radosavljevic, A., and R. P. Anderson. 2014. Making better Maxent models of species distributions: complexity, overfitting and evaluation. Journal of Biogeography 41: 629–643.
		(D3) dataset used to evaluate models	random partitioning	default in MaxEnt	
	output	(D4) format/transformation	logistic	we used the logistic output format	
		(D5) threshold	10 percentile training presence	we repeated this procedure 10 times for each algorithm and used the Lowest Present Threshold values (Pearson et al., 2007) to transform each map in binary.	
	extrapolation	(D6) novelty of projected environments compared with training environments	MaxEnt(MESS)	MaxEnt conducted multivariate environmental similarity surfaces (MESS) analysis to assess the effect of model extrapolation on values of predictor variables lying outside the training range, that is, projecting models on non-analogous climates, following Elith et al. (2011)	
		(D7) collinearity shift between training and projected environments	NA	NA	
		(D8) extrapolation strategy	Extrapolate and no clamping	ten replicates of each model were conducted by extrapolation and no clamping.	