SA. Supplementary Information: Ecological Niche Modeling and Spatio-Temporal Dynamics Workflow

This supplementary document details the computational workflow, parameterization, and spatial data processing implemented for modeling the distribution and kinematics of *Calotropis procera* in South America. The complete analytical pipeline was executed in the R programming environment [1].

SA1 Data Acquisition and Spatial Filtering

Occurrence Data: Primary occurrence records for *C. procera* were aggregated from the Global Biodiversity Information Facility (GBIF) alongside regional datasets from Paraguay field campaigns, Brazil (from speciesLink network, 2024-Set-30, `specieslink.net/search`), and Argentina [2]. Data from these three sources were combined using tidy data criteria [3] into an initial dataset. Duplicate records and occurrences lacking precise geographic coordinates were removed.

Spatial Autocorrelation Mitigation: To minimize environmental over-representation and spatial autocorrelation caused by sampling bias, we applied a spatial thinning algorithm using the `spThin` R package [4]. Nearest Neighbor Analysis (NNA) was conducted using `spatstat` [5] to quantify baseline spatial aggregation. Based on species dispersal capacity and environmental grid resolution, records were thinned to ensure a minimum separation distance of 10 km.

Background Sampling: We constrained the calibration area to limit the inclusion of environmentally suitable but historically inaccessible regions. A 900 km buffer was generated around the thinned occurrence records using the `sf` package [6] (projected in Eckert IV to ensure equal-area operations before reverting to WGS84). Within this calibration extent, 10,000 background points were randomly sampled.

SA2 Environmental Predictors and Dimensionality Reduction

Baseline climate data (1970–2000) were obtained from WorldClim v2.1 [7] at a spatial resolution of 2.5 arc-minutes ($\approx$ 5 km). To prevent multicollinearity from inflating model variance, we conducted a rigorous collinearity analysis. Highly correlated bioclimatic variables (Pearson's $|r| > 0.90$) or those exhibiting high Variance Inflation Factors (VIF) were systematically excluded. The final retained predictor set comprised 10 non-collinear bioclimatic variables (e.g., bio_13, bio_14, bio_15, bio_18, bio_19, bio_2, bio_3, bio_5, bio_7, bio_8).

SA3 Model Tuning and Selection (ENMeval)

Ecological niche models were fitted using the `maxnet` algorithm [8] via the `ENMeval` 2.0 package [9]. To ensure robust evaluation under spatially structured environments, we partitioned occurrences into testing and training sets using the 'block' spatial cross-validation method.

We systematically tuned model complexity by evaluating a comprehensive grid of hyperparameters:

- **Feature Classes (FC):** Linear (L), Linear-Quadratic (LQ), and Linear-Quadratic-Hinge (LQH).
- **Regularization Multipliers (RM):** Ranging from 0.5 to 8.0 in increments of 0.5.

Composite Model Selection: Rather than relying on a single metric, we implemented a normalized composite scoring system to select the optimal model, balancing discrimination capacity and overfitting risk. The composite score (C) was calculated as:

$$C = w_{\text{auc}}\hat{A}_{\text{val}} + w_{\text{or}}(1 - \hat{O}_{10}) + w_{\text{diff}}(1 - \hat{\Delta}_{\text{auc}}) + w_{\text{aicc}}(1 - \hat{\text{AICc}}) \quad (1)$$

where $\hat{A}_{\text{val}}$ is the normalized validation AUC, $\hat{O}_{10}$ is the normalized 10% omission rate, $\hat{\Delta}_{\text{auc}}$ is the normalized difference between training and validation AUC, and $\hat{\text{AICc}}$ is the normalized Akaike Information Criterion corrected for small sample sizes. The weights were assigned as $w_{\text{auc}} = 0.40, w_{\text{or}} = 0.20, w_{\text{diff}} = 0.20$, and $w_{\text{aicc}} = 0.20$. Based on this rigorous evaluation, the `fc.LQH_rm.1.5` model was selected as optimal.

SA4 Future Projections and Kinematic Expansion Extraction

CMIP6 Projections: The optimized model was projected onto future climate scenarios derived from the Coupled Model Intercomparison Project Phase 6 (CMIP6). Future projections were rigorously subjected to a quality control (QC) pipeline to ensure strict topological congruence (extent, resolution, and CRS) with the baseline data, executing bilinear resampling where necessary via the `terra` package [10]. Projections were mapped using the *cloglog* link function, approximating habitat suitability bounds $[0, 1]$.

Ensemble and Spatio-Temporal Dynamics: For each projected period, the ensemble median and Interquartile Range (IQR) were computed pixel-by-pixel across the selected General Circulation Models (GCMs). The temporal evolution of the suitability grids was analyzed using discrete-time finite difference operators to extract the spatial kinematics:

1. **Suitability Velocity (V):** The rate of change in habitat suitability over time (first derivative).
2. **Suitability Acceleration (A):** The change in velocity over time (second derivative), capturing the non-linear intensification or deceleration of the invasion front.

All mapping and spatial visualizations were generated utilizing a combination of `ggplot2` [11], `tidyterra`, and `patchwork` [12], prioritizing perceptually uniform color palettes (e.g., Viridis/Magma) to ensure high-contrast readability.

SA5 Climate Models and Scenario Interpretation

To robustly assess the potential expansion and contraction of *C. procera*, we integrated baseline observational data with future climate projections derived from the Coupled Model Intercomparison Project Phase 6 (CMIP6) [13].

Baseline Climate: The historical reference climate was established utilizing WorldClim v2.1 [7], representing the average climatological conditions for the 1970–2000 period. This dataset interpolates global climate surfaces based on an extensive network of meteorological station records, serving as the "current" reference state for suitability modeling.

General Circulation Models (GCMs): To account for the inherent epistemic uncertainty in future climate simulations, we utilized a multi-model ensemble approach. The ensemble incorporates projections from distinct GCMs participating in CMIP6, selected based on their availability in the WorldClim database at the required spatial resolution. The specific GCMs included in our ensemble median and Interquartile Range (IQR) kinematics calculations are detailed in Table SA1.

Future Climate Scenarios: Future projections for the periods centered around 2050 (2041–2060) and 2070 (2061–2080) were driven by the SSP2-4.5 pathway (Shared Socioeconomic Pathway 2,

Table SA1: Global Climate Models (GCMs) from CMIP6 included in the ensemble projections.

GCM Identifier	Originating Institution / Consortium
ACCESS-CM2	CSIRO and ARCCSS, Australia
MPI-ESM1-2-HR	Max Planck Institute for Meteorology, Germany
MRI-ESM2-0	Meteorological Research Institute, Japan
UKESM1-0-LL	Met Office Hadley Centre and NERC, UK

Radiative Forcing 4.5 W/m^2) [14]. This "Middle of the Road" scenario represents intermediate levels of mitigation and adaptation challenges, serving as a robust benchmark for climate-driven ecological modeling.

SA6 *Calotropis procera* in the western region of Paraguay



Figure SA1: Photograph of a record of the occurrence of Calotropis procera in the western region of Paraguay.

SA7 Performance metrics for best model selection

SA8 New record of *Calotropis procera* for Paraguay

Table SA2: Coordenadas geográficas para nuevos registros de *Calotropis procera*

Código	Latitud	Longitud	Colector/a
CP1	-19.30555	-59.75386	Feltes, O.
CP2	-22.46639	-59.63278	Feltes, O.
CP3	-22.06333	-59.68306	Feltes, O.

(continúa en la próxima página)

(continuación)

Código	Latitud	Longitud	Colector/a
CP4	-22.44194	-60.10833	Feltes, O.
CP5	-22.43139	-61.04611	Feltes, O.
CP6	-22.30111	-60.02639	Feltes, O.
CP7	-19.7875	-59.77667	Feltes, O.
CP8	-21.4362222	-61.4059444	Pérez de Molas, L.
CP9	-21.3965833	-61.4611944	Pérez de Molas, L.
CP10	-21.4493056	-61.3887778	Pérez de Molas, L.
CP11	-21.5048889	-61.3127222	Pérez de Molas, L.
CP12	-21.2581722	-61.6291889	Pérez de Molas, L.
CP13	-21.3741417	-61.4915278	Pérez de Molas, L.
CP14	-20.7835194	-60.5407778	Pérez de Molas, L.
CP15	-22.5231667	-56.9469444	Pérez de Molas, L.
CP16	-22.5734167	-60.0982222	Pérez de Molas, L.
CP17	-22.5756944	-60.0966111	Pérez de Molas, L.
CP18	-22.5755833	-60.0967222	Pérez de Molas, L.
CP19	-22.5755556	-60.0966389	Pérez de Molas, L.
CP20	-22.5758056	-60.0966111	Pérez de Molas, L.
CP21	-22.005074	-60.6240885	Pérez de Molas, L.
CP22	-22.131914	-57.7172753	Mereles, F.
CP23	-21.9511111	-61.6930556	Mereles, F.
CP24	-21.8930556	-61.6227778	Mereles, F.
CP25	-21.8452778	-61.6535556	Mereles, F.
CP26	-20.0723056	-61.5479444	Mereles, F.
CP27	-20.0648889	-61.8829444	Mereles, F.
CP28	-19.9841111	-58.9560000	Mereles, F.
CP29	-19.9841389	-59.5696389	Mereles, F.
CP30	-19.9828889	-60.0075000	Mereles, F.
CP31	-19.9836111	-59.7340278	Mereles, F.
CP32	-19.6341667	-60.6447222	Mereles, F.
CP33	-19.6622222	-60.8344444	Mereles, F.
CP34	-21.7708333	-60.9455556	Mereles, F.
CP35	-21.6616667	-61.0944444	Mereles, F.
CP36	-21.2036111	-61.6636111	Mereles, F.
CP37	-21.4105556	-61.4225000	Mereles, F.
CP38	-21.3088889	-61.8388889	Mereles, F.
CP39	-21.5050556	-61.6519444	Mereles, F.
CP40	-21.6131389	-62.0296667	Mereles, F.
CP41	-21.2358333	-60.9936111	Mereles, F.
CP42	-21.1019444	-60.6719444	Mereles, F.
CP43	-21.1386111	-60.4333333	Mereles, F.
CP44	-23.8127778	-60.7791667	Mereles, F.
CP45	-23.0747222	-61.0735278	Mereles, F.
CP46	-20.7175000	-59.7159722	Mereles, F.
CP47	-22.6127778	-61.3361111	Mereles, F.
CP48	-22.0311111	-61.2413889	Mereles, F.
CP49	-21.4344444	-61.8838889	Mereles, F.
CP50	-21.6600000	-61.8788889	Mereles, F.
CP51	-23.4836111	-60.7580556	Mereles, F.
CP52	-23.6729722	-60.1611111	Mereles, F.

(continúa en la próxima página)

(continuación)

Código	Latitud	Longitud	Colector/a
CP53	-23.5104722	-58.1798889	Mereles, F.
CP54	-22.6094444	-58.2868611	Mereles, F.
CP55	-22.8825000	-58.8261111	Mereles, F.
CP56	-22.2408889	-58.4161111	Mereles, F.; Brusquetti, A.
CP57	-22.3967500	-58.9022222	Mereles, F.; Brusquetti, A.
CP58	-21.3500000	-59.7172222	Mereles, F.; Brusquetti, A.
CP59	-23.1494444	-59.6122222	Mereles, F.; Brusquetti, A.
CP60	-23.1527778	-59.6561111	Mereles, F.; Brusquetti, A.
CP61	-22.8600000	-59.7497222	Mereles, F.; Brusquetti, A.
CP62	-22.9141667	-59.8133333	Mereles, F.; Brusquetti, A.
CP63	-22.8680556	-60.0302778	Mereles, F.; Brusquetti, A.
CP64	-22.1088889	-59.2758333	Mereles, F.; Brusquetti, A.
CP65	-22.5119444	-59.1980556	Mereles, F.; Brusquetti, A.
CP66	-22.9466667	-59.6516667	Mereles, F.; Brusquetti, A.
CP67	-22.6500000	-61.2436111	Mereles, F.; Brusquetti, A.
CP68	-22.6127778	-61.3361111	Mereles, F.; Brusquetti, A.
CP69	-22.6155556	-61.4519444	Mereles, F.; Brusquetti, A.
CP70	-22.7061111	-61.1469444	Mereles, F.; Brusquetti, A.
CP71	-22.0755556	-61.2402778	Mereles, F.; Brusquetti, A.
CP72	-23.3713889	-59.5352778	Mereles, F.; Brusquetti, A.
CP73	-23.6305556	-59.5444444	Mereles, F.; Brusquetti, A.
CP74	-23.6241667	-59.5447222	Mereles, F.; Brusquetti, A.
CP75	-23.7225000	-59.8597222	Mereles, F.; Brusquetti, A.
CP76	-23.6941667	-59.8672222	Mereles, F.; Brusquetti, A.
CP77	-23.4991667	-59.8980556	Mereles, F.; Brusquetti, A.
CP78	-22.3761111	-62.5194444	Mereles, F.; Brusquetti, A.
CP79	-22.2433333	-62.4611111	Mereles, F.; Brusquetti, A.
CP80	-22.0288889	-61.8244444	Mereles, F.; Brusquetti, A.
CP81	-21.9805556	-62.3069444	Mereles, F.; Brusquetti, A.
CP82	-21.9983333	-61.6919444	Mereles, F.; Brusquetti, A.
CP83	-21.8930556	-61.6227778	Mereles, F.; Brusquetti, A.
CP84	-21.7025000	-61.0402778	Mereles, F.; Brusquetti, A.
CP85	-21.7708333	-60.9455556	Mereles, F.; Brusquetti, A.
CP86	-21.4105556	-61.4225000	Mereles, F.; Brusquetti, A.
CP87	-21.6935000	-61.8441667	Mereles, F.; Brusquetti, A.
CP88	-29.6601667	-61.8788889	Mereles, F.; Brusquetti, A.
CP89	-21.5055278	-61.6532222	Mereles, F.; Brusquetti, A.
CP90	-21.5747222	-61.8233333	Mereles, F.; Brusquetti, A.
CP91	-21.6131389	-62.0296667	Mereles, F.; Brusquetti, A.
CP92	-23.6558333	-60.6722222	Mereles, F.; Brusquetti, A.
CP93	-23.4836111	-60.7580556	Mereles, F.; Brusquetti, A.
CP94	-23.4582778	-60.7772500	Mereles, F.; Brusquetti, A.
CP95	-23.4808333	-60.7316667	Mereles, F.; Brusquetti, A.
CP96	-23.5108333	-60.6169444	Mereles, F.; Brusquetti, A.
CP97	-23.6729722	-60.1612778	Mereles, F.; Brusquetti, A.
CP98	-23.0747222	-61.0735278	Mereles, F.; Brusquetti, A.
CP99	-23.0736111	-61.0035000	Mereles, F.; Brusquetti, A.
CP100	-23.0725000	-60.9533333	Mereles, F.; Brusquetti, A.
CP101	-23.4152778	-60.7938889	Mereles, F.; Brusquetti, A.

(continúa en la próxima página)

(continuación)

Código	Latitud	Longitud	Colector/a
CP102	-23.2955556	-60.7994444	Mereles, F.; Brusquetti, A.
CP103	-23.1394444	-61.3919444	Mereles, F.; Brusquetti, A.
CP104	-23.2586111	-61.4208333	Mereles, F.; Brusquetti, A.
CP105	-23.2491667	-61.4161111	Mereles, F.; Brusquetti, A.
CP106	-19.9833333	-59.5138889	Mereles, F.; Brusquetti, A.
CP107	-19.9828889	-60.0075000	Mereles, F.; Brusquetti, A.
CP108	-19.9844444	-60.0545056	Mereles, F.; Brusquetti, A.
CP109	-19.8394444	-59.9374167	Mereles, F.; Brusquetti, A.
CP110	-19.9825000	-59.8225000	Mereles, F.; Brusquetti, A.

References

[1] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2023). URL `https://www.R-project.org/`.

[2] Powell, P. A. & Spagarino, C. R. ¿ calotropis procera (apocynaceae) se expande en el chaco argentino? *Boletín de la Sociedad Argentina de Botánica* **60**, 351 (2025).

[3] Wickham, H., Grolemund, G. *et al. R for data science*, vol. 2 (O'Reilly Sebastopol, CA, 2017).

[4] Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B. & Anderson, R. P. spthin: an r package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38**, 541–545 (2015).

[5] Baddeley, A., Rubak, E. & Turner, R. *Spatial point patterns: methodology and applications with R* (CRC press, 2015).

[6] Pebesma, E. Simple features for r: Standardized support for spatial vector data. *The R Journal* **10**, 439–446 (2018).

[7] Fick, S. E. & Hijmans, R. J. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology* **37**, 4302–4315 (2017).

[8] Phillips, S. J., Anderson, R. P., Dud'ık, M., Schapire, R. E. & Blair, M. E. Opening the black box: an open-source release of maxent. *Ecography* **40**, 887–893 (2017).

[9] Kass, J. M. *et al.* Enmeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. *Methods in Ecology and Evolution* **12**, 1602–1608 (2021).

[10] Hijmans, R. J. *terra: Spatial Data Analysis* (2023). URL `https://CRAN.R-project.org/package=terra`. R package version 1.7-55.

[11] Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York, 2016).

[12] Pedersen, T. L. *patchwork: The Composer of Plots* (2022). URL `https://CRAN.R-project.org/package=patchwork`. R package version 1.1.2.

[13] Eyring, V. *et al.* Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development* **9**, 1937–1958 (2016).

[14] O'Neill, B. C. *et al.* The scenario model intercomparison project (scenariomip) for cmip6. *Geoscientific Model Development* **9**, 3461–3482 (2016).

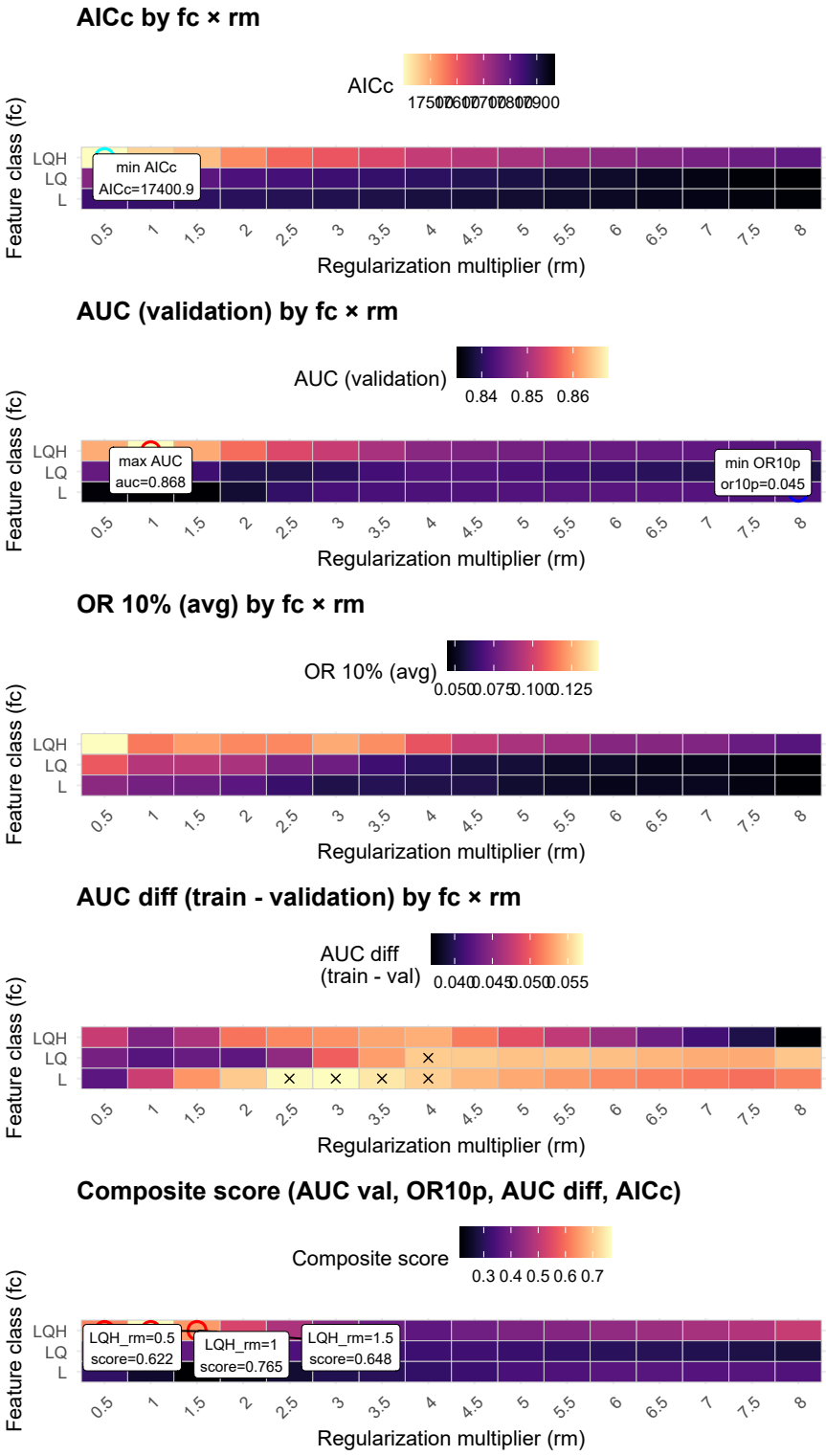

AICc by fc × rm
AICc
Feature class (fc)
LQH
LQ
L
min AICc
AICc=17400.9
0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8
Regularization multiplier (rm)
AUC (validation) by fc × rm
AUC (validation)
0.84 0.85 0.86
max AUC
auc=0.868
min OR10p
or10p=0.045
OR 10% (avg) by fc × rm
OR 10% (avg)
0.050 0.075 0.100 0.125
AUC diff (train - validation) by fc × rm
AUC diff
(train - val)
0.040 0.045 0.050 0.055
Composite score (AUC val, OR10p, AUC diff, AICc)
Composite score
0.3 0.4 0.5 0.6 0.7
LQH_rm=0.5
score=0.622
LQH_rm=1
score=0.765
LQH_rm=1.5
score=0.648

Figure SA2: .