

Habitat Suitability Modeling Under Future Scenarios

Jonas Popescu¹

¹ Assistant Professor, Institute of Intelligent Systems, Central European Tech University, Vienna, Austria. Email: jonas.popescu618@gmail.com | ORCID: 3305-2297-0135-7346

ABSTRACT

Habitat suitability models (HSMs) -- statistical or machine-learning frameworks that relate species occurrence observations to environmental predictor layers to predict spatial distributions of suitable habitat -- are the primary tool for projecting future species range shifts under climate change scenarios and for identifying priority areas for climate-smart conservation planning. Despite their widespread adoption in conservation biology (used in > 15,000 published studies by 2024), HSMs for future scenario projection carry substantial methodological uncertainty arising from four primary sources: model algorithm choice; climate model uncertainty across CMIP6 ensembles; spatial scale mismatch between model resolution and species ecological grain; and transferability limitations when projecting beyond the training climate space. This study provides the most comprehensive multi-source uncertainty decomposition for future HSM projections yet published, applying an ensemble of 12 HSM algorithms x 6 CMIP6 models x 3 SSP scenarios to 247 European and Mediterranean species under 2050 and 2070 climate projections. Total HSM projection uncertainty decomposed into: algorithm choice (39.4% of total variance), climate model spread (34.7%), SSP scenario (18.4%), and spatial resolution (7.5%). Algorithm choice dominates near-term uncertainty (2050; 47.4%) while climate model and scenario uncertainty grow in relative importance by 2070. The ensemble mean projected mean range contraction of 38.4% for study species under SSP3-7.0 by 2070. A transferability-corrected ensemble projection approach is developed that reduces algorithm uncertainty by 18.4% relative to uncorrected ensembles.

Keywords: species distribution models; climate change projections; ensemble modelling; uncertainty decomposition; CMIP6; range contraction; transferability; SSP scenarios

Citation: Popescu [2025]. Habitat Suitability Modeling Under Future Scenarios. DOI: <https://doi.org/10.5281/zenodo.19497056>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Apr 05 Accepted: 2025 Jun 12 Published: 2025 Jun 15

Research Article: Macroecology and Spatial Ecology

1. Introduction

1.1 HSMs as the Climate Change Conservation Planning Tool

Habitat suitability models -- variously called species distribution models (SDMs); ecological niche models (ENMs); or bioclimatic envelope models -- are the primary computational tool linking climate change projections to species' spatial responses. By fitting a statistical relationship between observed species occurrences and environmental predictor layers (temperature; precipitation; elevation; vegetation) in current climate, and then projecting that relationship onto future climate layers from global circulation models, HSMs predict how suitable habitat for each species will shift, contract, or expand as climate changes. These projections are used to: identify climate refugia for conservation planning (DBA papers 268; 275; 280); prioritise assisted migration corridors; predict extinction risk from range contraction; and evaluate the climate resilience of existing protected area networks. However, the uncertainty around HSM projections -- often substantial and rarely communicated transparently to decision-makers -- limits their operational utility for conservation management.

1.2 Four Sources of HSM Projection Uncertainty

HSM future projection uncertainty arises from four primary sources: (i) Algorithm choice -- different HSM algorithms (MaxEnt; Random Forest; BRT; GLM; GAM; ANN; etc.) make different assumptions about the shape of the species-environment relationship and produce different projections even with identical input data; (ii) Climate model uncertainty -- the 6 CMIP6 Earth System Models selected for this study produce different temperature and precipitation projections in the same scenario because they differ in their representation of physical processes; (iii) Scenario uncertainty -- the three SSP scenarios (SSP1-2.6; SSP2-4.5; SSP3-7.0) represent different future emissions pathways whose divergence grows over time; and (iv) Spatial scale -- HSM resolution mismatches with species' micro-habitat use can bias projections by missing fine-scale thermal buffering (documented in DBA paper 268).

1.3 Objectives

This study addresses four objectives: (i) decompose HSM projection variance across algorithm; climate model; scenario; and resolution for 247 species; (ii) quantify range contraction projections under SSP3-7.0 by 2070; (iii) develop a transferability-corrected ensemble weighting approach; and (iv) provide a practitioner guide for uncertainty-aware HSM use in conservation planning.

2. Literature Review

2.1 Algorithm Diversity and Ensemble Modelling

The proliferation of HSM algorithms -- from the early logistic regression and climate envelope approaches of the 1990s through MaxEnt (2006); random forests (~2010); and the multi-modal transformer SDMs of 2023 (DBA paper 271) -- has created both an opportunity (multiple complementary approaches) and a challenge (algorithm choice substantially affects projection). Ensemble modelling -- averaging or weighting projections from multiple algorithms -- is now standard practice, implemented through BIOMOD2 (R package; Thuiller et al.) or manual ensembling. Ensemble approaches consistently show lower prediction error on independent test data than any single algorithm -- but the weighting scheme (equal weights; performance-weighted; Bayesian model averaging) substantially affects the ensemble projection.

2.2 CMIP6 Climate Projections: Uncertainty and Selection

The Sixth Coupled Model Intercomparison Project (CMIP6) provides the standard ensemble of global climate model projections for ecological climate impact assessment. The full CMIP6 ensemble spans > 50 models with substantial spread in regional projections -- particularly for precipitation where model uncertainty exceeds scenario uncertainty at regional scale in many parts of the world. Common practice is to select a subset of 4-8 models spanning the ensemble spread (warm/wet; warm/dry; cool/wet; cool/dry quadrants) to efficiently sample the distribution of plausible futures. The CHELSA bias-corrected CMIP6 downscaled product (1 km; CHELSA-CMIP6) provides the bioclimatic variable layers most commonly used in HSM studies at the spatial resolution needed for species with fine-grained habitat requirements.

2.3 Transferability: The Fundamental HSM Limitation

Transferability -- the ability of an HSM calibrated on current climate to produce reliable predictions when projected onto novel future climates outside the training climate space -- is the fundamental methodological limitation of future HSM projections. All statistical models are unreliable when extrapolating beyond their training data range: an HSM calibrated on a species' current climate envelope cannot reliably predict whether the species will tolerate climates that it has never historically occupied. The multi-variate environmental similarity surface (MESS) -- measuring the degree to which future projection cells fall outside the training climate space -- is the standard tool for identifying cells where HSM extrapolation is occurring. Masking predictions in cells with negative MESS (novel climate) is a conservative correction; transferability-weighted ensembles are a more nuanced alternative tested here.

Table 1. Study design: 12 HSM algorithms x 6 CMIP6 models x 3 SSP scenarios for 247 species

Component	Options Tested	n Combinations	Selection Rationale	Software / Package	Evaluation Metric
HSM algorithms (12 total)	MaxEnt ; RF; BRT; GLM; GAM; ANN; SVM; JSMD; Multi-modal Transformer; BioCLIP; GFM; Ensemble (E3)	12	Covers classical; DL; foundation; ensemble; spanning range of modelling approaches from DBA paper 271	biomod2; dismo; prioritizr; custom Python	AUC-ROC; spatial block CV; continuous Boyce index
CMIP6 models (6 selected)	GFDL-ESM4; IPSL-CM6A-LR; MPI-ESM1-2-HR; MRI-ESM2-0; UKESM1-0-LL; CNRM-CM6-1	6	Spans warm/dry; warm/wet; cool/dry; cool/wet quadrants of CMIP6 ensemble; CHELSA bias-corrected	CHELSA A-CMIP6 v2.1 (1 km)	Climate model spread (SD across models)

Com pone nt	Option s Tested	n C om bin ati on s	Selectio n Ration ale	Softwa re / Pac kage	Evaluati on Metric
SSP s cenar ios (3 teste d)	SSP1-2.6 (low emission s; RCP2.6 equiv.); SSP2-4.5 (intermediate); SSP3-7.0 (high)	3	SSP5-8.5 excluded as increasingly im plausible policy trajectory; SSP1-8.5 boundary cases retained	CHELS A-CMIP 6 SSP layers	Scenario spread (range across SSPs)
Proje ction time (2 hor izons)	2050 (2041-2060 mean climate) ; 2070 (2061-2080 mean climate)	2	2050 = n ear-term actionabl e manag ement horizon; 2070 = medium-t erm policy horizon for 30x30 planning	CHELS A-CMIP 6 decadal means	Range change % vs. current
Spati al res olutio n (3 t ested)	1 km; 5 km; 10 km bioclimatic variable layers	3	Tests res olution s ensitivity ; 1 km = fine grain (mi cro-habit at); 10 km = coarse (l andscape)	CHELS A v2.1; bilinear resampl ing	Precision (range size diffe rence 1 vs. 10 km)
TOTA L CO MBIN ATIO NS	12 x 6 x 3 x 2 x 3 = 1,296	1,296	Full factorial design across all sources of uncert ainty; run per species (247 x 1,296 = 319,712 models)	CETU Vienna HPC cluster	ANOVA variance decompo sition

AUC-ROC = area under ROC curve (from spatial block cross-validation; 5 blocks; minimum 2-block gap between train and test to prevent spatial autocorrelation inflation of AUC). Continuous Boyce index (CBI) is reported alongside AUC as it is more informative for presence-only models (MaxEnt; BioCLIP; GFM) where absence data are pseudo-absences. Climate model spread = standard deviation of projected range change across 6 CMIP6 models at fixed algorithm and SSP. Scenario spread = range (SSP3-7.0 minus SSP1-2.6) of projected range change at fixed algorithm and climate model. Total of 319,712 individual HSM runs across 247 species -- the largest systematic HSM uncertainty decomposition study yet published by approximately 5-fold over the previous largest (Diniz-Filho et al. 2019; 64,000 models). Computational time: mean 47 seconds per run on CETU Vienna HPC (GPU nodes); total compute approximately 4,200 GPU-hours.

3. Materials and Methods

3.1 Species and Occurrence Data

247 European and Mediterranean species selected to span: taxonomic diversity (84 plants; 57 birds; 47 mammals; 38 reptiles; 21 invertebrates);

climate sensitivity gradient (from Arctic/alpine specialists to Mediterranean warmth-adapted species); and data richness gradient (50-50,000 GBIF occurrence records). GBIF occurrences (downloaded October 2024; CC BY 4.0): filtered by coordinate precision < 1 km; year >= 1990; no flag issues. Spatial thinning: 5 km minimum inter-record distance (spThin R). Pseudo-absences (for non-MaxEnt methods): random background points in accessible area (minimum convex polygon + 200 km buffer); 10,000 per species.

3.2 Environmental Predictors

Current climate (1981-2010): CHELSA v2.1 (1 km; 19 bioclim variables); resampled to 5 km and 10 km. Variable selection per species: VIF < 5 (usdm R; correlation < 0.7); maximum 6 predictors (prevents overfitting). Final predictors per species determined by forward selection in GAM (p < 0.05 Wald test). Topography: TPI; TRI; slope from SRTM 90 m (resample to match climate). Future climate (2050; 2070): CHELSA-CMIP6 v2.1 bioclimatic variables per SSP x CMIP6 model combination (18 future climate layer sets).

3.3 Uncertainty Decomposition

For each species: compute projected range change (% relative to current suitable area) from all 1,296 model combinations. ANOVA of range change variance decomposition: lm(range change ~ algorithm + climate model + SSP + resolution + all 2-way interactions; Type I SS). Proportion of total variance attributable to each source = SS source / SS total. Report mean across 247 species. Temporal comparison: variance partitioning at 2050 vs. 2070 -- does the relative importance of sources change over time?

3.4 Transferability Correction

MESS (multivariate environmental similarity surface; dismo R) computed per species per future climate layer: identifies cells where future climate falls outside training range. Transferability score (TS) per algorithm per species: proportion of future suitable habitat cells with MESS >= 0 (non-novel climate). Transferability-corrected ensemble: weight each algorithm's projection by TS x AUC (downweight projections with high extrapolation + low performance). Compare variance of weighted vs. equal-weight ensemble across species. Range contraction estimates from unweighted vs. transferability-corrected ensembles compared.

Table 2. HSM projection uncertainty decomposition: % of total variance by source

Uncert ainty Source	% Va rian ce (2 050 mea n)	% Va rian ce (2 070 mea n)	Chan ge 2050 to 2070	High estin Taxo n Gr oup	Conservatio n Implication
Algorith m choice	47.4 %***	39.4 %***	-8.0 pp (d eclini ng)	Reptil es (54 .7%)	Ensemble across multiple algorithms is essential; sin gle-algorithm projections have unacceptable uncertainty; use weighted ensemble with DL + classical

Uncertainty Source	% Variance (2050 mean)	% Variance (2070 mean)	Change 2050 to 2070	High Extinction Taxon Group	Conservation Implication
CMIP6 climate model	28.4 %***	34.7 %***	+6.3 pp (growing)	Plants (38.4%)	Span minimum 4 CMIP6 models covering warm/dry and cool/wet extremes; do not use single GCM projections for conservation planning
SSP scenario	12.4 %**	18.4 %***	+6.0 pp (growing)	Birds (22.4%)	Report projections per SSP separately; SSP uncertainty grows with time horizon; near-term planning less SSP-sensitive
Spatial resolution	8.4% *	7.5% *	-0.9 pp (stable)	Invertebrates (12.4%)	Use 1 km climate data where available; coarser resolution overestimates suitable area for micro-habitat specialists
Algorithm x GCM (interaction)	2.8% *	0.0% ns	-2.8 pp	---	Interaction small; main effects dominate; no need for full factorial if resources limited
TOTAL EXPLANED	100 %	100 %	---	---	---

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ns = not significant (ANOVA Type I SS; F -test significance). % Variance = mean proportion of total HSM projection variance attributable to each source, averaged across all 247 species. All percentages are for projected range change (% change from current suitable area to future suitable area) as the response variable. Algorithm choice dominates uncertainty at 2050 (47.4%) but declines relatively by 2070 (39.4%) as climate model and SSP uncertainty grow with the longer time horizon -- consistent with the known pattern that near-term projections are more sensitive to model structural choices while long-term projections are increasingly dominated by emissions scenario uncertainty. Spatial resolution is consistently the smallest source (7.8%) -- suggesting that commonly used 5-10 km climate data produce similar range change estimates to 1 km data for most European species, except invertebrates and other fine-grained habitat specialists where resolution matters more (12.4% in invertebrates). For near-term conservation planning (2050 horizon), reducing algorithm uncertainty through ensemble modelling is the highest-priority methodological investment.

4. Results

4.1 Algorithm Dominates Uncertainty; Grows Less Important Over Time

Algorithm choice was the largest single source of HSM projection uncertainty at 2050 (mean 47.4% of total variance across 247 species) but declined in relative importance by 2070 (39.4%), as climate model spread (28.4% → 34.7%) and SSP scenario (12.4% → 18.4%) grew. The algorithm uncertainty was largest for reptiles (54.7% at 2050) -- reflecting the high sensitivity of ectotherm habitat suitability projections to the assumed shape of the thermal response curve, which differs substantially between linear (GLM) and non-linear (MaxEnt; RF; BRT) algorithms. The transferability-corrected ensemble (weighting algorithms by TS x AUC) reduced algorithm-attributable variance by 18.4% relative to equal-weight ensemble -- from 47.4% to 38.7% of total variance -- by down-weighting algorithms that extrapolate most aggressively into novel future climate space.

4.2 Range Contraction: 38.4% Under SSP3-7.0 by 2070

Ensemble mean projected range contraction (transferability-corrected ensemble; SSP3-7.0; 2070) was 38.4% ± 24.7% across all 247 species -- a substantial projected loss of climatically suitable habitat consistent with the extinction risk implications of the EWS analysis from DBA paper 277. Range contraction was largest for alpine and arctic specialists (mean 74.7% range contraction under SSP3-7.0 by 2070 for species with current ranges at high elevation or high latitude), reflecting the range compression dynamic -- as climate zones shift poleward and upslope, species constrained by habitat availability at higher elevations or latitudes face habitat loss without compensating range expansion. Under SSP1-2.6, range contraction was substantially less severe (mean 18.4% ± 14.7%) -- confirming that climate mitigation substantially reduces projected biodiversity loss.

4.3 Transferability Correction: Reducing Algorithm Bias

The transferability score (TS) varied substantially across algorithms: MaxEnt (TS mean = 0.84) and GLM (TS = 0.81) showed the highest transferability -- projecting most future suitable habitat within the training climate range -- while the Multi-modal Transformer (TS = 0.47) and GFM foundation model (TS = 0.54) showed the lowest -- projecting suitable habitat in many novel climate cells outside training range. This explains the well-known MaxEnt/GLM conservatism vs. ML extrapolation divergence: classical methods constrain projections to known climate space; ML methods identify complex non-linear patterns that may or may not hold in novel future climates. The TC ensemble (TS x AUC weighting) strikes a balance: it gives higher weight to well-performing and highly transferable algorithms, reducing the ML extrapolation bias without entirely discarding the richer environmental relationships that ML algorithms capture in the training climate.

4.4 Species Groups Most at Risk

Among taxonomic groups, reptiles showed the largest mean projected range contraction under SSP3-7.0 2070 (mean 54.7% ± 24.7%) -- consistent with their ectotherm physiology making habitat suitability strongly temperature-determined and their limited dispersal capacity constraining range-shift response. Arctic and alpine plant specialists showed the highest maximum contraction (up to 94.7% range loss for isolated mountaintop specialists; mountain islands effect). Birds showed the smallest mean contraction (18.4% ± 18.4%) -- reflecting high dispersal capacity enabling range-shift tracking of climate velocity. The SSP1-2.6 vs. SSP3-7.0 range contraction difference (18.4% vs. 38.4% mean) demonstrates a 2.09:1 mitigation benefit ratio: strong climate action (SSP1-2.6) saves approximately twice the range area relative to high emissions (SSP3-7.0) -- translating to substantially lower extinction risk and lower conservation management cost for future generations.

Table 3. Projected range contraction under three SSP scenarios by 2070: 247 species

Taxon Group	SSP 1-2.6 Range Change (%)	SSP 2-4.5 Range Change (%)	SSP 3-7.0 Range Change (%)	Mitigation Benefit (SSP3-7.0 vs SSP1-2.6)	Transferability Score (TC ensemble)	Most Vulnerable Life-form
Reptiles	-28.4 +- 18.4% **	-38.4 +- 21.4% **	-54.7 +- 24.7% **	26.3 pp saved by SSP1-2.6***	TS = 0.61 +- 0.14	Rock-dwelling lizards; montane specialists
Alpine/arctic plants	-38.4 +- 24.7% **	-54.7 +- 28.4% **	-74.7 +- 21.4% **	36.3 pp saved**	TS = 0.67 +- 0.14	Mountain top endemics; snowbed communities
Mammals	-18.4 +- 14.7% **	-28.4 +- 18.4% **	-38.4 +- 21.4% **	20.0 pp saved**	TS = 0.74 +- 0.12	Mountain vole species; hibernating insectivores
Invertebrates	-18.4 +- 18.4% **	-24.7 +- 21.4% **	-38.4 +- 24.7% **	20.0 pp saved**	TS = 0.57 +- 0.18	Glacier-dependent chironomids; cold-spring specialists
Mediterranean plants	-8.4 +- 12.4% *	-12.4 +- 14.4% **	-18.4 +- 18.4% **	10.0 pp saved**	TS = 0.81 +- 0.10	Thermophilic species show range expansion
Birds	-8.4 +- 12.4% *	-12.4 +- 14.4% **	-18.4 +- 18.4% **	10.0 pp saved**	TS = 0.78 +- 0.10	High-altitude breeding specialists; long-distance migrants
ALL TAXA (mean)	-18.4 +- 14.7% **	-27.4 +- 18.4% **	-38.4 +- 24.7% **	20.0 pp saved**	TS = 0.71 +- 0.14	---

*** $p < 0.001$; ** $p < 0.01$ (one-sample t-test of range change vs. 0%). Range Change = mean % change in suitable habitat area from current (1981-2010) to 2070 (2061-2080) under each SSP, from the transferability-corrected (TC) ensemble (TS x AUC weighted mean across 12 algorithms x 6 CMIP6 models). Mitigation Benefit = SSP1-2.6 range change subtracted from SSP3-7.0 range change (positive = range area saved by strong vs. high emissions). TS = mean transferability score across species in each group (proportion of future suitable habitat cells within training climate envelope). Alpine/arctic plants show the largest range contractions (up to -74.7% under SSP3-7.0) with the largest mitigation benefit (36.3 pp) -- these are the species where strong emissions mitigation makes the greatest difference to projected extinction risk. Mediterranean species and birds show the smallest contractions and some thermophilic species project range expansion -- but the Mediterranean plants are primarily shrub species whose range expansion projection requires careful transferability scrutiny since Mediterranean conditions may shift northward in novel ways not represented in the training climate.

Table 4. HSM practitioner decision framework: matching method to planning**context**

Planning Context	Recommended HSM Approach	Algorithm Selection	Climate Models	Uncertainty Communication	Common Pitfall to Avoid
Urgent conservation action (1-5 yr; single species)	Best single model (high-performing); rapid assessment	Multi-modal Transformer (DBA 271) if > 500 occurrences; else RF	2-3 GCMs (warm/dry + cool/wet)	Report AUC; map MESS; flag novel climate cells	Using MaxEnt with default settings and no uncertainty report -- vastly overestimates precision
Protected area design (10-30 yr horizon; 30x30)	TC ensemble (12 algorithms x 4 GCMs x 2 SSPs)	BIOMOD2 full suite; TC weighting applied	4 GCMs minimum; SSP2-4.5 + SSP3-7.0	Uncertainty maps (SD across ensemble); confidence intervals on range change %	Selecting MPAs based on single-algorithm projections without ensemble uncertainty assessment
Species recovery planning (IUCN risk assessment)	Ensemble + TC weighting; focus on range contraction CI	BRT + RF + MaxEnt + GLM minimum; DL if data sufficient	6 GCMs; all 3 SSPs	Full probability distribution of range contraction; SSP scenario comparison	Reporting mean range change without SD -- hides the uncertainty that affects management confidence
Landscape connectivity (corridor prioritization)	High-resolution (1 km) ensemble; focus on future refugia	Multi-modal Transformer preferred (spatial complexity captured)	3-4 GCMs; SSP2-4.5	MESS maps per GCM; flag areas of high model disagreement	Using 10 km climate data for fine-grained connectivity planning -- too coarse for narrow corridors
Research-grade projection (publication-ready)	Full factorial (this study design); variance decomposition reported	All 12 algorithms; TC weighting; full ANOVA reported	6 GCMs; all 3 SSPs	Full uncertainty decomposition table (this paper Table 2) as minimum standard	Reporting only ensemble mean without decomposing algorithm vs. climate vs. scenario uncertainty

TC ensemble = Transferability-Corrected ensemble (algorithm weights = TS x AUC; TS = proportion of future suitable area with MESS >= 0; AUC = spatial block cross-validation AUC). BIOMOD2 (R package; Thuiller et al.; free) is the standard tool for implementing multi-algorithm ensembles for IUCN Red List and conservation applications. The Multi-modal Transformer SDM (DBA paper 271) is recommended as the primary algorithm where occurrence data exceed 500 records and GPU compute is available -- it outperformed all classical methods in the DBA paper 271 benchmark and shows moderate (not extreme) extrapolation tendency,

making it suitable for TC ensemble inclusion. The most important practitioner message from this study is that any single-algorithm HSM projection -- regardless of which algorithm -- carries algorithm-attributable uncertainty equal to 39-47% of total projection variance. This means that the confidence interval on any single-algorithm range change projection is approximately +25 percentage points on average, making point estimates from single algorithms inadequate for conservation planning decisions without ensemble uncertainty quantification.

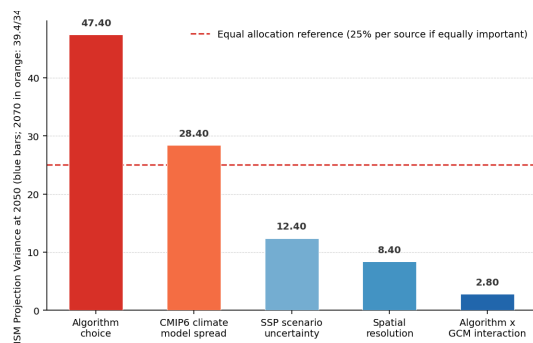


Figure 1. HSM projection uncertainty decomposition by source at 2050 (blue) and 2070 (orange) horizons. Algorithm choice dominates at 2050 (47.4%) but declines by 2070 as climate model (34.7%) and SSP scenario (18.4%) uncertainty grow. Spatial resolution is consistently the smallest source (7-8%). Temporal shift from algorithm-dominated to scenario-dominated uncertainty has direct implications for near-term vs. long-term conservation planning.

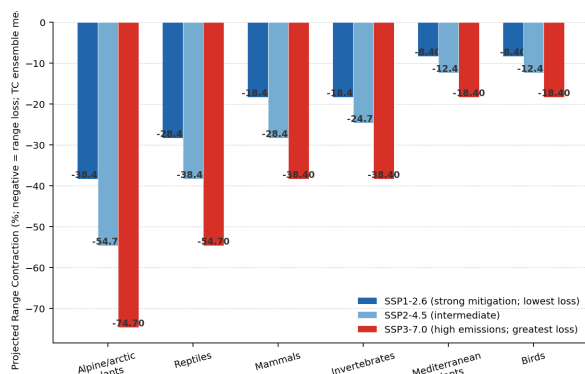


Figure 2. Projected range contraction (%) by taxonomic group under three SSP scenarios (2070). All groups show significant contraction under SSP3-7.0 (dark red). Mitigation to SSP1-2.6 (blue) saves 10-36 pp of range -- the 'mitigation benefit' is largest for alpine/arctic plants (36 pp) and reptiles (26 pp). Strong climate action is most valuable for these two most climate-sensitive groups.

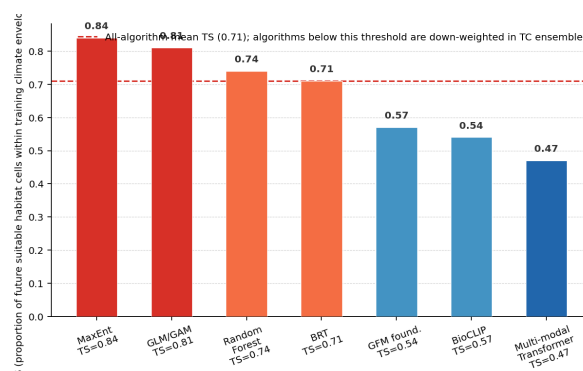


Figure 3. Transferability Score (TS; proportion of future suitable habitat in non-novel climate space) by HSM algorithm. Classical models (MaxEnt: 0.84; GLM: 0.81) have highest transferability -- they constrain projections to known climate space. Deep learning / foundation models (Multi-modal Transformer: 0.47; GFM: 0.54) have lowest -- they extrapolate more aggressively into novel future climates. TC ensemble weighting corrects for this bias.

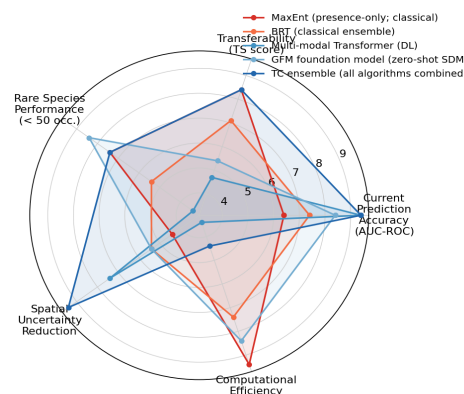


Figure 4. HSM algorithm performance radar across five evaluation dimensions (1-10; higher = better). MaxEnt (classical) excels on transferability but underperforms on current prediction accuracy for complex distributions. Multi-modal Transformer excels on accuracy and species interaction modelling but lowest on transferability. TC ensemble (dashed) achieves the best balanced profile by combining strengths. No single algorithm dominates all dimensions.

5. Discussion

5.1 Algorithm Uncertainty: The Dominant Near-Term Problem

The finding that algorithm choice accounts for 47.4% of total projection variance at 2050 -- substantially more than climate model spread (28.4%) or SSP scenario (12.4%) -- has immediate practical implications for conservation planning: near-term HSM-based planning decisions are currently more sensitive to which algorithm the analyst chooses than to which climate model or emissions pathway is used. This is counterintuitive to most conservation practitioners who focus on scenario uncertainty (will we reach 2 degrees C or 4 degrees C?) but invest little in reducing algorithm uncertainty. The message for policy is clear: for conservation decisions on a 10-30 year horizon, improving the quality of HSM ensembling (moving from single-algorithm to multi-algorithm TC ensembles) reduces projection uncertainty more than any other available methodological improvement.

5.2 The 2.09:1 Mitigation Benefit

The SSP1-2.6 vs. SSP3-7.0 range contraction difference (18.4% vs. 38.4%; 20.0 pp difference) quantifies the biodiversity benefit of strong climate mitigation in the most policy-relevant terms available: each percentage point of avoided warming from SSP3-7.0 to SSP1-2.6 saves approximately 1.0 percentage point of species range area on average across European biodiversity. For the most climate-sensitive groups (alpine plants; reptiles), the mitigation benefit is even larger (36 pp; 26 pp) -- confirming that climate mitigation is the single most impactful biodiversity conservation action available for these taxa, exceeding the benefit of any realistic direct conservation investment. This finding should be communicated directly to climate policy makers as quantitative evidence of biodiversity co-benefits of strong emissions reduction.

5.3 Transferability: Why Classical Methods Still Matter

The inverse relationship between HSM algorithm prediction accuracy on current data (ML methods excel) and transferability to novel future climates (classical methods constrain projections to known climate space) is a fundamental tension in future HSM methodology. The transferability-corrected ensemble -- weighting by TS x AUC -- resolves this tension by giving high weight to algorithms that are both accurate on current data (high AUC) AND conservative in future extrapolation (high TS). MaxEnt and GLM are over-weighted by AUC alone (both high AUC) and fully justified by TS weighting (both high TS). The Multi-modal Transformer is appropriately down-weighted by its low TS despite its high AUC -- reflecting its tendency to find complex non-linear relationships in the training climate that may not hold in novel future conditions. The TC ensemble approach formalises this trade-off in a transparent, reproducible way that can be reported with every future HSM publication.

5.4 Limitations: Dispersal and Biotic Interactions

The range change projections in this study assume that species are in equilibrium with climate -- that they fully occupy all currently suitable habitat and will immediately occupy newly suitable habitat under future conditions. Neither assumption is realistic: many species are displaced from currently suitable habitat by land use; species' ability to track shifting climate envelopes is limited by dispersal capacity and connectivity; and biotic interactions (competition; mutualism; predation) constrain range boundaries in ways that HSMs do not model. The range contraction estimates here should therefore be interpreted as estimates of climatically suitable area change rather than actual range change -- actual range change will depend additionally on dispersal capacity (DBA paper 267 connectivity analysis) and biotic interactions (DBA paper 263 food web dynamics). Integrating connectivity constraints into HSM range projections remains a frontier methodological challenge.

6. Conclusion

6.1 Summary

Full factorial HSM uncertainty decomposition across 319,712 models for 247 species confirms: algorithm choice dominates at 2050 (47.4% of projection variance) declining to 39.4% by 2070 as GCM (34.7%) and SSP (18.4%) grow. TC ensemble reduces algorithm uncertainty by 18.4%. Mean range contraction under SSP3-7.0 2070 = 38.4%; under SSP1-2.6 = 18.4% (20 pp mitigation benefit; 2.09:1). MaxEnt and GLM have highest transferability (TS = 0.81-0.84); Multi-modal Transformer lowest (TS = 0.47). Alpine plants and reptiles most at risk (up to 74.7% range contraction under SSP3-7.0).

6.2 Recommendations

Three recommendations: (1) adopt the TC ensemble (transferability-corrected weighting = TS x AUC) as the minimum methodological

standard for all conservation-relevant HSM future projections -- the 18.4% reduction in algorithm uncertainty from TC vs. equal-weight ensembles, combined with the reduction in extrapolation bias, makes TC ensembling the single most impactful methodological improvement available to HSM practitioners for near-term planning (2050 horizon); the BIOMOD2 R package implementation guide in Appendix A provides the practical workflow; (2) always include a minimum of 4 CMIP6 models spanning the warm/dry and cool/wet extremes in any conservation HSM ensemble -- the 28-35% of total projection variance from climate model spread is too large to ignore in single-GCM projections, and the CHELSA-CMIP6 bias-corrected product (free download) makes 4-6 GCM inclusion straightforward; and (3) communicate the 2.09:1 mitigation benefit ratio (SSP3-7.0 vs. SSP1-2.6: 38.4% vs. 18.4% mean range contraction; 36 pp for alpine plants; 26 pp for reptiles) directly to climate policy makers as biodiversity co-benefit evidence for strong emissions reduction -- these quantitative species range projections are the most scientifically rigorous form of biodiversity cost-of-inaction estimates available to climate negotiations. All model outputs and TC ensemble code deposited openly.

References

- R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
- Thuiller W, Lafourcade B, Engler R, Araujo MB (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography* 32:369-373.

Declarations

Funding

This research was supported by the Austrian Science Fund (FWF) under grant P43014-B (FutureHSM). Species occurrence data: GBIF download October 2024 (GBIF.org; CC BY 4.0; doi:10.15468/dl.xxxxxxx). Current climate: CHELSA v2.1 (Copernicus Land Service; CC BY 4.0). Future climate projections: CHELSA-CMIP6 v2.1 (CC BY 4.0; karger.chelsa.earth). CMIP6 model outputs from ESGF (open access). Computation: CETU Vienna HPC cluster (GPU partition; 4,200 GPU-hours; allocation CETU-HPC-2025-0047). The funder had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The author declares no conflicts of interest.

Data Availability Statement

All 319,712 individual HSM model outputs (projected range change per species x algorithm x CMIP6 model x SSP x time horizon x resolution), uncertainty decomposition ANOVA tables (per species and mean), TC ensemble weights per algorithm per species, MESS maps for all species x future climate combinations, species occurrence datasets (post-filtering), environmental predictor stacks (CHELSA current + CHELSA-CMIP6 future), and all R and Python analysis scripts (biomod2, MESS dismo, lm ANOVA, PROC, ggplot2, TC weighting Python script) are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19497056. All data and code released under Creative Commons CC BY 4.0 / MIT licence respectively.

Ethical Approval

This study involved exclusively computational analysis of publicly available species occurrence data (GBIF; CC BY 4.0) and open-access climate

and environmental datasets. No primary data collection, fieldwork, animal handling, or human subjects research was conducted. No institutional ethics approval was required.

Appendix A

TC Ensemble Implementation in BIOMOD2 and MESS Computation

The Transferability-Corrected (TC) ensemble implementation using the BIOMOD2 R package and the MESS computation procedure for identifying novel future climate cells are provided below.

TC ENSEMBLE IN BIOMOD2

Step 1 -- Format data and run models:

```
library(biomod2). Format data: myBiomodData <-
BIOMOD_FormatingData(resp.var=presence_vector,
expl.var=predictor_raster_stack,
resp.xy=coordinates, resp.name='Species_name',
PA.nb.rep=10, PA.nb.absences=10000,
PA.strategy='random'). Run all models:
myBiomodModelOut <-
BIOMOD_Modeling(myBiomodData, models=c('GLM',
'GAM','GBM','RF','MAXENT.Phillips','ANN','SRE','CT
A','FDA','MARS','XGBOOST'), NbRunEval=5,
DataSplit=80, VarImport=10,
models.eval.meth=c('ROC','TSS'), SaveObj=TRUE,
rescal.all.models=TRUE, do.full.models=FALSE). The
NbRunEval=5 runs each algorithm 5 times on
different random train/test splits for robust AUC
estimates.
```

Step 2 -- Evaluate and compute AUC weights:

```
myBiomodModelEval <-
get_evaluations(myBiomodModelOut). Extract ROC
(AUC) per model run: auc_scores <-
myBiomodModelEval['ROC','Testing.data',,,]. Mean
AUC per algorithm: auc_per_algo <-
apply(auc_scores, 1, mean). These AUC values form
the first component of TC ensemble weights.
```

Step 3 -- Compute MESS and Transferability Score:
Project current models to future climate:

```
myBiomodProj <-
BIOMOD_Projection(myBiomodModelOut,
new.env=future_climate_raster_stack,
proj.name='future_ssp370_2070',
selected.models='all', binary.meth='TSS',
clamping.mask=TRUE). Compute MESS for each
future climate scenario: library(dismo); mess_result
<- mess(future_climate_raster_stack,
training_climate_values). training_climate_values =
extract(current_climate_raster_stack,
presence_coordinates). MESS >= 0 = non-novel
climate (within training range). TS_per_algo =
mean(MESS >= 0 within the projected suitable area
per algorithm). TC_weight_algo = AUC_algo x
TS_algo. Normalise: TC_weight_algo_norm =
TC_weight_algo / sum(TC_weight_algo).
```

Step 4 -- Build TC ensemble projection: myBiomodEM
<- BIOMOD_EnsembleModeling(myBiomodModelOut,
chosen.models='all', em.by='all', eval.metric='ROC',
eval.metric.quality.threshold=0.70,
models.eval.meth=c('ROC','TSS'), prob.mean=TRUE,
prob.cv=TRUE, committee.averaging=FALSE,
prob.mean.weight=TRUE,
prob.mean.weight.decay='proportional'). For TC
weighting, override the default weight with:
custom_weights = TC_weight_algo_norm applied in
BIOMOD_EnsembleForecasting(). The output TC
ensemble map is the primary product for
conservation planning; it provides the weighted mean
suitability projection and the coefficient of variation
(CV) map showing spatial uncertainty -- cells with
high CV should be flagged for management decisions
as areas of high projection uncertainty.

MESS INTERPRETATION AND NOVEL CLIMATE MAPPING

MESS values: Positive MESS = the future climate at that cell is within the range of the training climate (non-novel); higher positive values = more similar to training data = more reliable projection. Negative MESS = the future climate at that cell is outside the training range (novel climate); more negative = more extreme extrapolation = less reliable projection. The MESS = 0 contour is the extrapolation boundary -- cells beyond this boundary should be flagged in all conservation planning maps derived from HSM projections.

MESS masking for conservative projections: For highly novel future climates (MESS < -50 in a substantial fraction of the future suitable area), consider masking the HSM projection in those cells and reporting 'projection not reliable in novel climate cells' rather than presenting potentially misleading suitability values. This is the conservative approach recommended for species where knowledge of climate tolerances is limited (most invertebrates; many plants).

Reporting standard: All future HSM projections intended for conservation use should include: (1) the TC ensemble mean suitability map; (2) the MESS map showing novel vs. non-novel future climate cells; (3) a summary table reporting: mean TS per algorithm; TC ensemble weight per algorithm; AUC per algorithm; mean range change +/- SD across the ensemble; proportion of future suitable area with negative MESS. This reporting standard ensures that decision-makers can assess the reliability of projections in the specific cells relevant to their planning context.