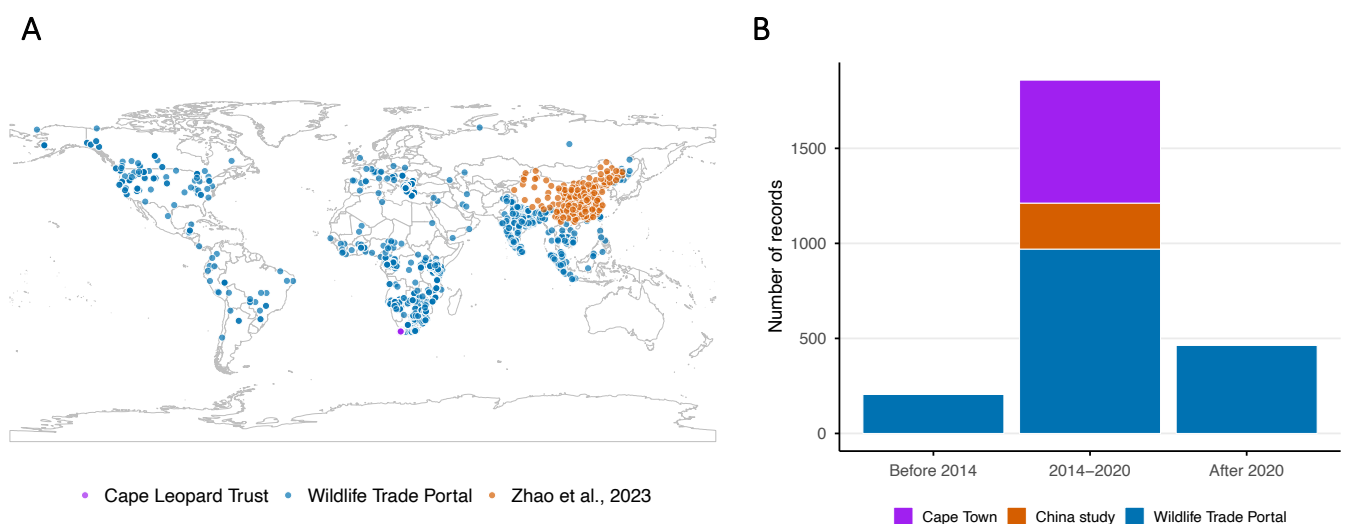


Supplementary Information

Data sources and coverage of poaching records

To achieve broader geographic and taxonomic coverage than any single data source alone, poaching occurrence records were compiled from three complementary sources: the Wildlife Trade Portal (WTP) incident database, a species-level poaching dataset for China spanning 2014-2020 from Zhao et al., 2023, and a locally compiled dataset for the Cape Leopard Trust spanning 2019-2020. The WTP database alone substantially underrepresented East and Southeast Asia, a region of known high poaching pressure, and the China-specific dataset was incorporated specifically to address this gap, while the Cape Leopard Trust dataset was incorporated as it represented snare incident data, an additional spatial information of poaching occurrence. Supplementary Fig. S1a shows the spatial distribution of records contributed by each source and Supplementary Fig. S1b shows their distribution across the study period. Because the China and Cape Town datasets are documented only as spanning a known date range rather than providing individual record-level dates, their temporal resolution is coarser than that of the WTP data; this is reflected in Supplementary Fig. S1b, where records from these two sources are shown as falling entirely within their known time block rather than being broken down by individual year.



Supplementary Figure S1. Spatial and temporal distribution of poaching occurrence records by data source.

a, Global distribution of poaching records prior to grid-cell aggregation, coloured by source: the Wildlife Trade Portal incident database (blue), the Zhao et al., 2023 poaching dataset (orange) and the Cape Leopard Trust dataset (purple). **b,** Number of records per data source across three time blocks (before 2014, 2014–

2020, after 2020). Because the China and Cape Town datasets are documented only as spanning 2014-2020 rather than by individual record year, all records from these two sources are shown within that block; only the Wildlife Trade Portal dataset has per-record year information and is broken down across all three blocks.

Full cross-validation results by pseudo-absence replicate

Model performance was evaluated using sensitivity (the proportion of observed poaching locations correctly identified) and the Boyce index (which evaluates whether predicted suitability at observed presences is consistently higher than expected under a random distribution), under both standard random 5-fold cross-validation and the more conservative spatial block cross-validation, which tests transferability to geographically distinct regions withheld entirely during training (see Methods). These metrics were calculated independently for each of the five pseudo-absence replicates, since each replicate represents a different random draw of background points and could in principle yield different performance. Supplementary Table S1 reports these replicate-level values in full, allowing the consistency of model performance across replicates to be assessed directly; the mean $\pm$ standard deviation values reported in the main text (Results: Model performance) were calculated across the five pseudo-absence replicates means shown here. The relatively small standard deviations across replicates for both models and both cross-validation strategies indicate that model performance was not strongly sensitive to the pseudo-absence replicates used.

Supplementary Table S1. Model performance metrics for each pseudo-absence replicate. Sensitivity and Boyce index are reported for GAM and RF under both standard 5-fold and spatial block cross-validation, for each of the five pseudo-absence replicates individually (see Methods). Values are mean $\pm$ standard deviation across the five folds within each replicate and cross-validation strategy. The mean $\pm$ standard deviation values reported in the main text (Results: Model performance) were calculated across these five replicate-level means.

Model	Dataset	5-fold Cross Validation		Spatial Block Cross Validation	
		Sensitivity	Boyce Index	Sensitivity	Boyce Index
GAMs	Dataset 1	0.806 $\pm$ 0.03	0.970 $\pm$ 0.02	0.796 $\pm$ 0.04	0.967 $\pm$ 0.01
	Dataset 2	0.791 $\pm$ 0.05	0.961 $\pm$ 0.02	0.785 $\pm$ 0.04	0.963 $\pm$ 0.01
	Dataset 3	0.799 $\pm$ 0.01	0.966 $\pm$ 0.00	0.792 $\pm$ 0.03	0.947 $\pm$ 0.03
	Dataset 4	0.8 $\pm$ 0.02	0.966 $\pm$ 0.02	0.789 $\pm$ 0.06	0.958 $\pm$ 0.02
	Dataset 5	0.808 $\pm$ 0.02	0.967 $\pm$ 0.01	0.795 $\pm$ 0.04	0.947 $\pm$ 0.05

RF	Dataset 1	0.966 ± 0.002	0.832 ± 0.006	0.811 ± 0.04	0.968 ± 0.02
	Dataset 2	0.967 ± 0.003	0.854 ± 0.007	0.808 ± 0.03	0.962 ± 0.01
	Dataset 3	0.967 ± 0.002	0.842 ± 0.009	0.797 ± 0.03	0.972 ± 0.01
	Dataset 4	0.964 ± 0.005	0.84 ± 0.013	0.805 ± 0.04	0.967 ± 0.01
	Dataset 5	0.961 ± 0.002	0.83 ± 0.01	0.809 ± 0.04	0.967 ± 0.02

Assessment of predictor collinearity

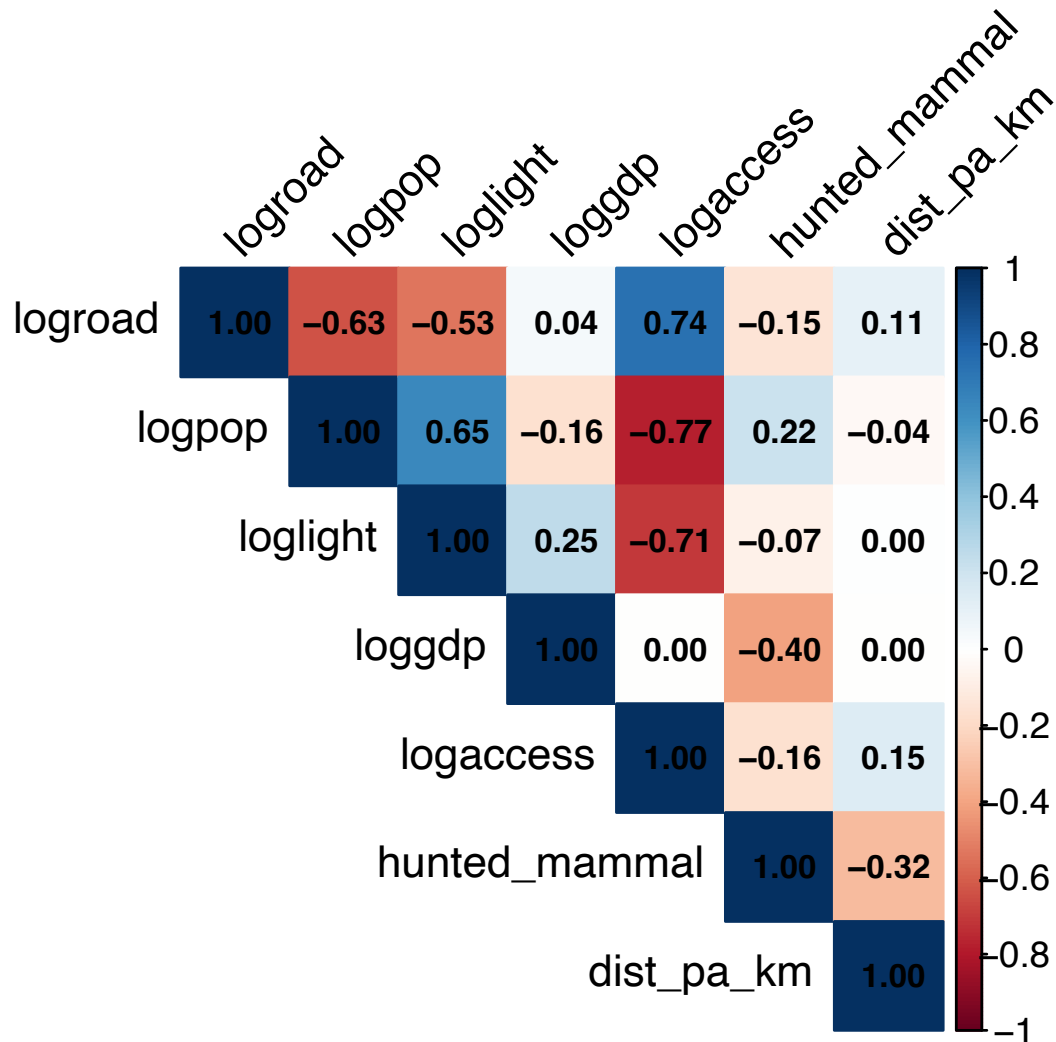
Population density was initially considered as a candidate predictor of poaching probability but was excluded prior to final model fitting due to collinearity with accessibility (travel time to nearest city). To assess this collinearity, population density values were extracted at all presence and pseudo-absence point locations from a gridded mean population density raster, obtained by calculating the mean of all population density values from 2010 to 2020 within each cell. Presence points (identical across all five pseudo-absence replicates) were included once, and pseudo-absence points from all five replicates were pooled, yielding a combined dataset for collinearity assessment. Population density, along with road density, night-time light intensity, GDP, and accessibility, was log transformed ($\log_{10} + 1$) prior to correlation analysis, consistent with the transformation applied to these predictors in the final GAM.

Pairwise Pearson correlations among all candidate predictors are shown below in Supplementary Fig. S2. Population density (log-transformed) was strongly correlated with accessibility (log-transformed; Pearson $r = -0.77$), supporting its exclusion from the final predictor set. Accessibility was retained over population density because it provides a more direct proxy for connectivity to urban markets, a recognised primary driver of illegal wildlife harvest.

Accessibility also showed moderate correlation with road distance ($r = 0.74$) and night-time light intensity ($r = -0.71$). These predictors were retained despite this correlation because they capture conceptually distinct mechanisms: accessibility reflects overall connectivity and travel cost to urban markets, road distance reflects local infrastructure and ease of physical movement, and night-time light intensity reflects local human activity and settlement intensity rather than distance-based connectivity.

To further assess multicollinearity, variance inflation factors (VIF) were calculated for all candidate predictors (Supplementary Table S2). All values in the final model were below the conventional threshold of 5 (range: 1.17-3.41). Accessibility's VIF declined from 4.14 to 3.41 following the

exclusion of population density, confirming that this exclusion resolved the principal source of multicollinearity in the predictor set.



Supplementary Figure S2. Pairwise Pearson correlation among candidate predictors of poaching probability, computed on log-transformed scale (except hunted mammal richness and distance to protected areas, which were not log-transformed) using presence points ($n=1,168$) combined with pooled pseudo-absence points from all five replicates. Colour and printed values indicate the Pearson correlation coefficient; blue indicates positive correlation; red indicates negative correlation.

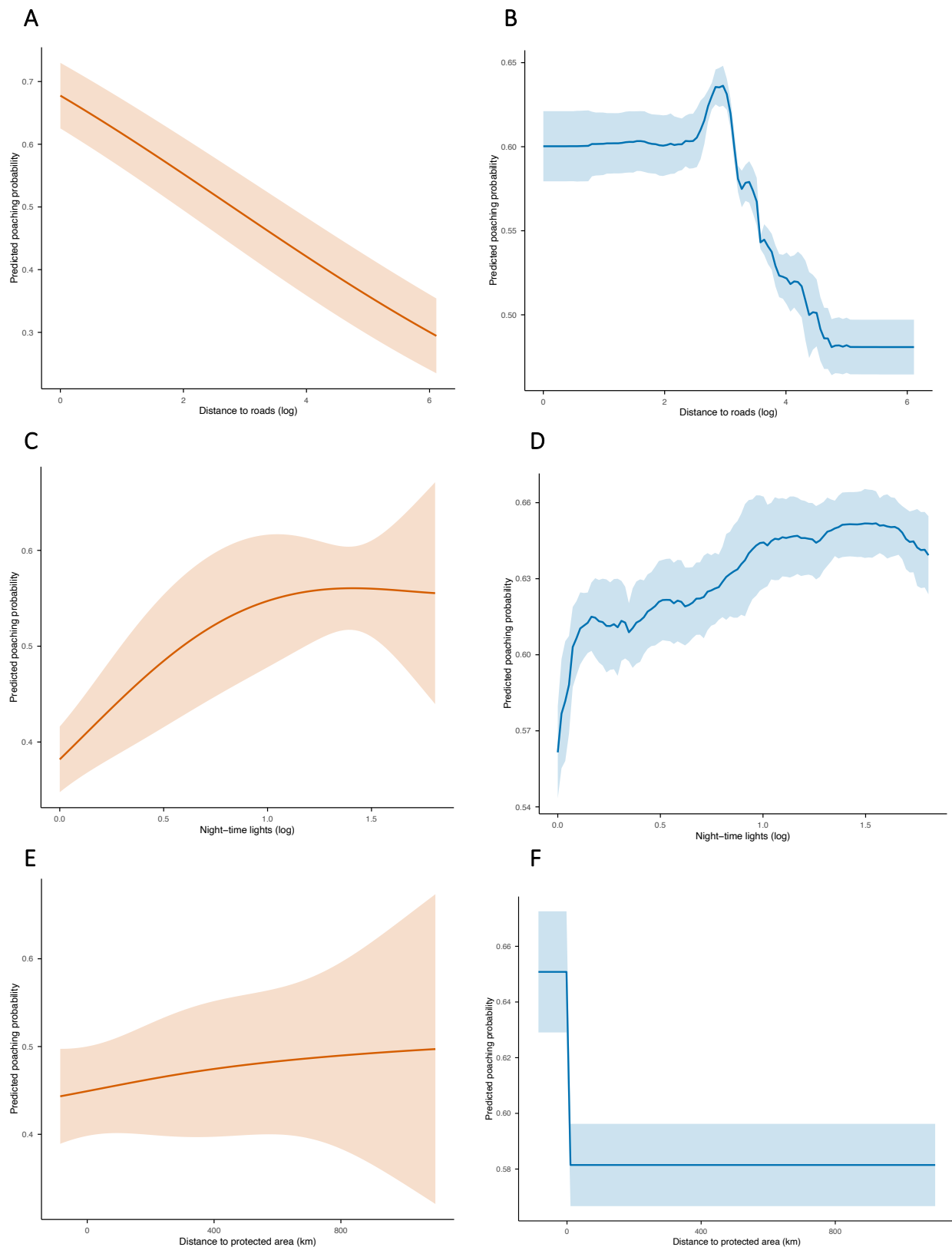
Supplementary Table S2. Variance inflation factors (VIF) for candidate predictors of poaching probability.

Predictor	VIF (with population density)	VIF (final model)
Accessibility (log)	4.14	3.41

Population density (log)	3.16	– (excluded)
Night-time lights (log)	2.69	2.36
Road distance (log)	2.29	2.25
GDP (log)	1.48	1.34
Hunted mammal richness	1.42	1.39
Distance to protected areas	1.18	1.17

GAM and RF response curves for secondary predictors

Response curves for the three weakest predictors in the final models (distance to roads, night-time light intensity, and distance to protected areas) are presented here rather than in the main text to keep Figure 3 focused on the predictors with the strongest and most interpretable effects.

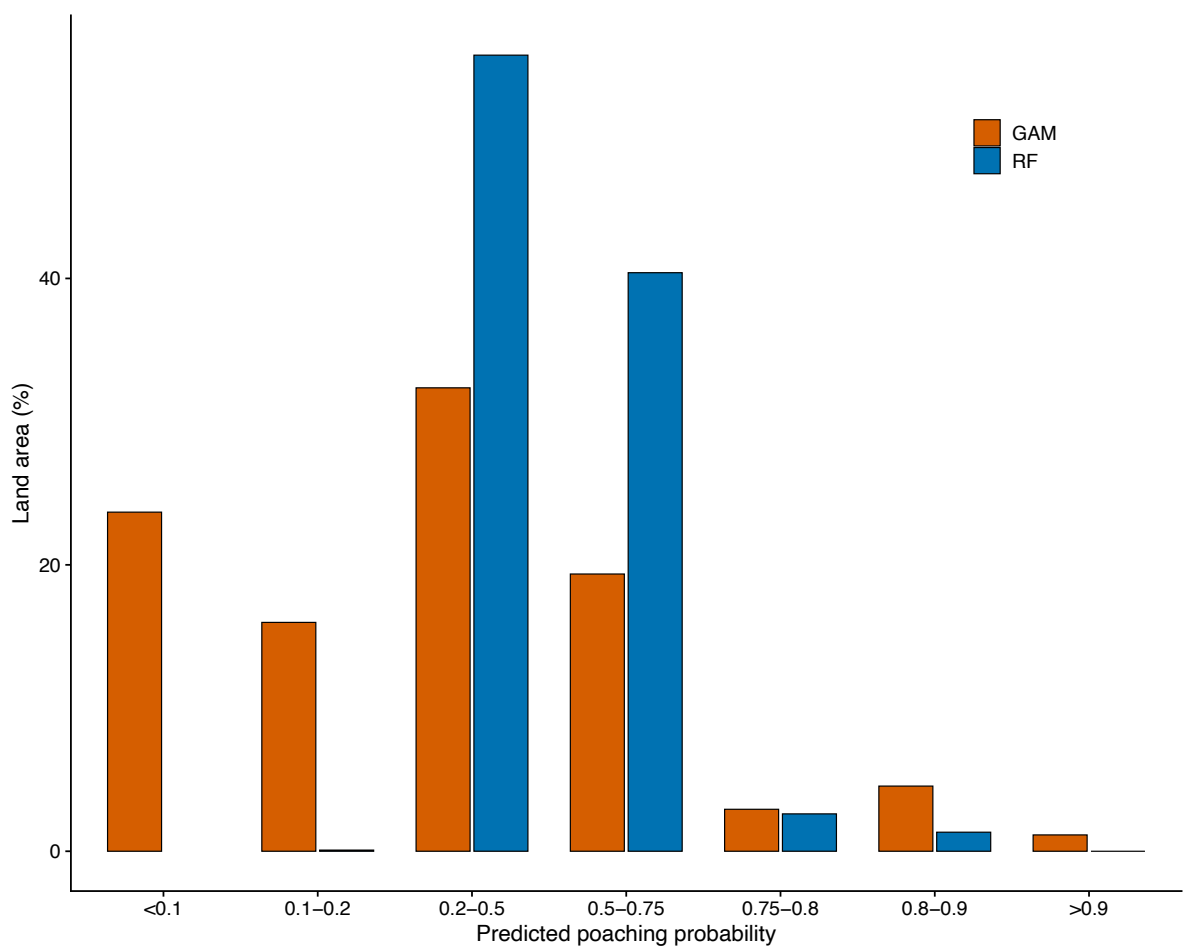


Supplementary Figure S3. GAM and RF response curves for the weakest predictors. **a**, GAM-predicted effect of distance to roads (log). **b**, Pooled RF partial dependence for distance to roads. **c**, GAM-predicted effect of night-time light intensity (log). **d**, Pooled RF partial dependence for night-time intensity. **e**, GAM-predicted effect of distance to protected areas. **f**, Pooled RF partial dependence for distance to protected areas. Solid

lines represent mean prediction across five pseudo-absence replicates, shaded areas indicate $\pm$ SD across replicates.

Land area distribution across predicted poaching-probability categories

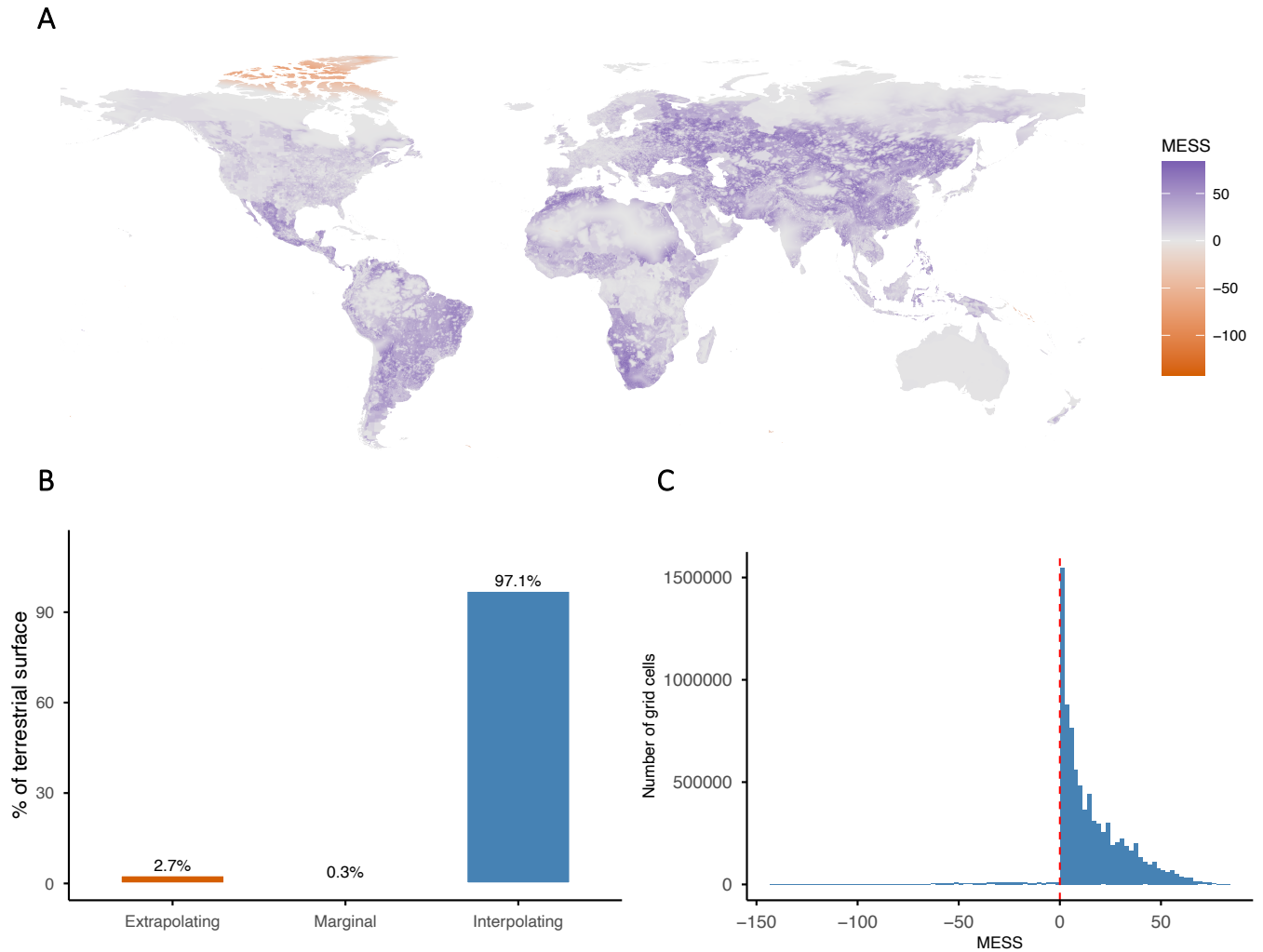
The percentage of the global terrestrial surface falling within each predicted-probability interval is shown separately for GAM and RF in Supplementary Fig. S4, offering a more granular view of the distributional contrast described in the main text. GAM predictions span a wider range of the [0,1] probability scale, identifying both extensive low-risk areas and a distinct, if small, high-risk tail, whereas RF predictions are concentrated within a narrower intermediate range and rarely approach extremes. This reflects the general tendency of random forests to average across many decision trees and shrink predictions toward the centre of the response range, in contrast to the smooth parametric functions underlying the GAM, which can extend closer to the full range of possible values.



Supplementary Figure S4. Land area classified within each predicted poaching-probability interval, compared between GAM and RF. Bars show the percentage of the global terrestrial surface falling within each predicted-probability bin, for the GAM (orange) and RF (blue) ensemble mean surface.

Assessment of spatial extrapolation

To assess the reliability of spatial extrapolation across the prediction domain, we computed the Multivariate Environmental Similarity Surface (MESS; see Methods), which quantifies, for each grid cell, how similar its environmental conditions are to those represented in the training data. Positive MESS values indicate that a cell's conditions fall within the training range (interpolation); negative values indicate that at least one predictor falls outside that range (extrapolation). Supplementary Fig. S5 shows the resulting global MESS surface, alongside the percentage of land area falling into each similarity category and the full distribution of MESS values across all grid cells. Most of the terrestrial surface (97.1%) fell within the training data's environmental envelope, with true extrapolation confined to a small proportion of land area (2.7%), concentrated along the high-Arctic margins of northern Canada and Greenland.



Supplementary Figure S5. Environmental extrapolation across the prediction domain, assessed using the Multivariate Environmental Similarity Surface (MESS). **a**, Global MESS map, computed from raw (untransformed) predictor values using one presence dataset combined with all five pseudo-absence replicates as the reference distribution (see Methods). Positive values (blue) indicate grid cells with environmental conditions similar to those in the training data, negative values (orange) indicate extrapolation beyond the training data's range. **b**, Percentage of the terrestrial surface classified as extrapolating (MESS < 0), marginal (MESS = 0, at the exact boundary of the training range), or interpolating (MESS > 0). **c**, Distribution of MESS values across all grid cells; the dashed red line marks MESS = 0.