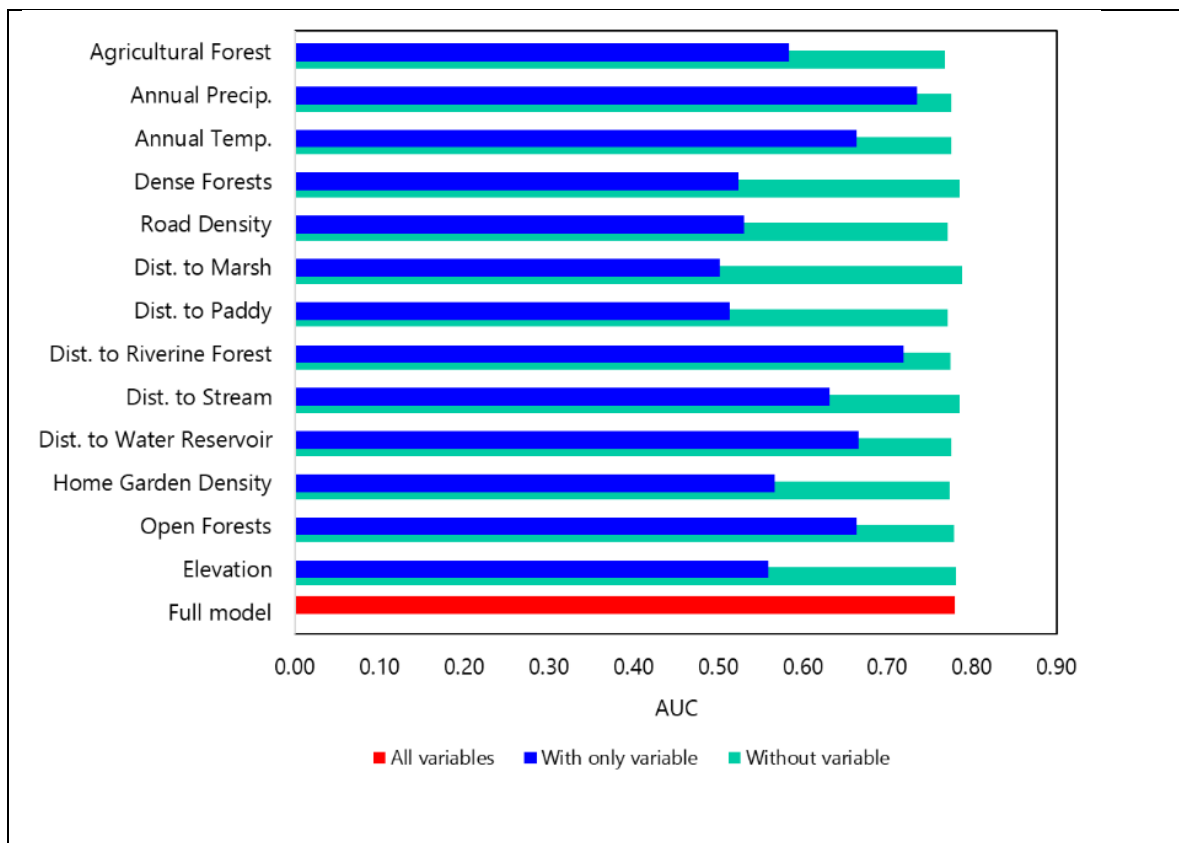


## Using citizen science to study a mesocarnivore: the jungle cat *Felis chaus* in Sri Lanka

SRIYANIE MITHTHAPALA, JEREMY DERTIEN, NIROSHA LIYANAGE, NIROSHAN MIRANDO, ANYA AVANTHI WEERAWARDANA RATNAYAKA, ASHAN THUDUGALA, DARSHANI WIJESINGHE and SAMPATH DE ALWIS GOONATILAKE

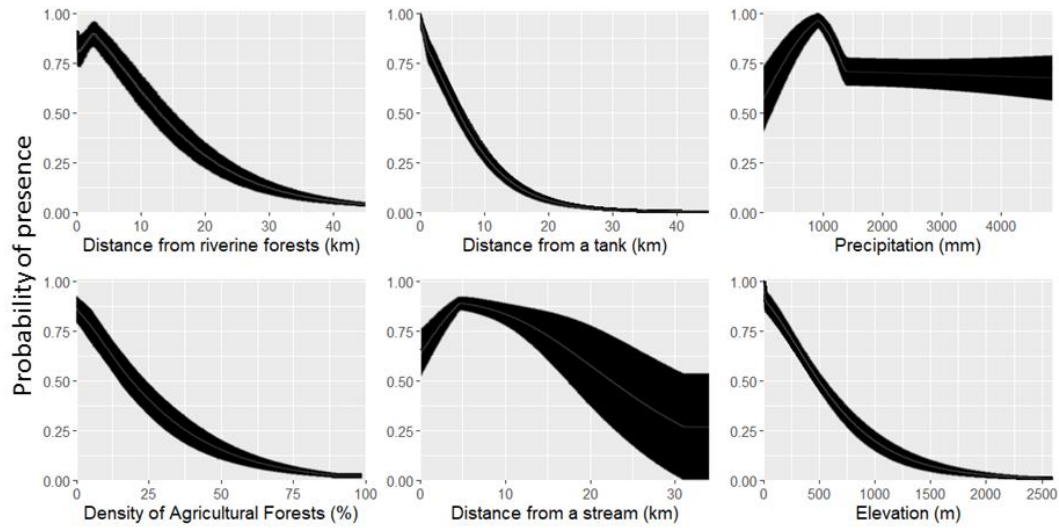
SUPPLEMENTARY FIG. 1 Jack-knife of highest mean area under the receiver operator characteristic curve (AUC).

Jackknife analysis estimates the bias and standard error (variance) of a parameter. This analysis is based on leaving out an observation from the sample set and resampling. Therefore, if there are  $n$  data points in the sample, the jackknife sampling technique will consist of  $n$  samples each with  $n-1$  data points in each sample subset analysis. The estimates of all models are then aggregated into a single estimate of the parameter, providing an estimate for bias and variance (Kissell & Poserina, 2017). When normalized units are used, the area under the curve (often referred to as simply the AUC) is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one.



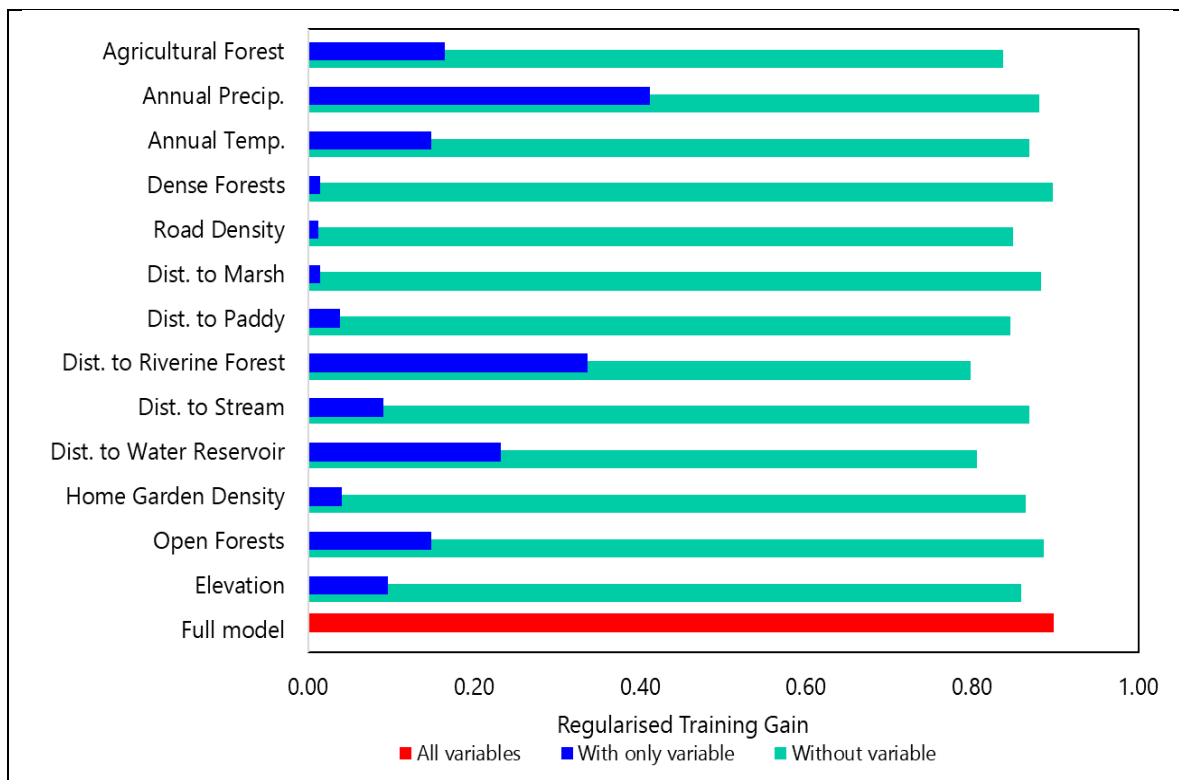
SUPPLEMENTARY FIG. 2 The top six covariates from the habitat suitability model of jungle cats.

The probability of jungle cat presence decreased with the increasing distance from riverine forests and from a water reservoir. In addition, increasing density of agricultural forests and increasing elevation were negatively correlated with jungle cat habitat suitability. Covariates are graphed by holding all other covariates within the model at their respective means and are presented with 95% confidence intervals.



### SUPPLEMENTARY FIG. 3 Jack-knife of regularised training gain.

Regularised training gain for each variable when habitat suitability was modelled with only that variable (dark blue) and the full model without that variable (turquoise). Regularised training gain is a measure of the increased model fit of the Maxent model distribution to the species locations versus the fit of a uniform distribution to the species locations, with a penalty factor. The exponential of the gain is the average ratio of the sample likelihood at an observed versus a background location. Hence, higher values of regularised training gain for a model with one specific variable is an indication that that variable is of greater importance in modelling habitat suitability. Annual precipitation, distance to riverine forests, and distance to water reservoirs had the highest regularised training gains when held alone, indicating that they are the most important variables for explaining jungle cat habitat suitability.



### References

KISSEL, R. & POSERINA, J. (2017) Advanced Math and Statistics. In *Optimal Sports Math, Statistics, and Fantasy* pp 103-135 (ed. R. Kissel & J. Poserina) Academic Press Books – Elsevier, Cambridge, Massachusetts.