Supplementary methods

Statistical models

Generalized additive models of forest presence or absence against climate, human influence and topographic variables were performed in R version 4.2.0 using the package mgcv version 1.8-40. Models utilized a binomial distribution with the default logit link, and four knots were defined for each variable to allow for simple non-linear effects of predictor variables. A global model with all variables, models including climate with just topography or just human influence, a model with only climate variables, a model without any climate variables, and a null model were run, defined by the following code. Variables defined are:

```
mod_global <- gam(code ~ s(bc_12,k=4)+s(bc_15,k=4)+s(TPI,k=4)+s(LATW,k=4),
data=data,family=binomial)

mod_TPI <- gam(code ~ s(bc_12,k=4)+s(bc_15,k=4)+s(TPI,k=4),
data=data,family=binomial)

mod_LATW <- gam(code ~ s(bc_12,k=4)+s(bc_15,k=4)+s(LATW,k=4),
data=data,family=binomial)

mod_climate <- gam(code ~ s(bc_12,k=4)+s(bc_15,k=4), data=data,family=binomial)

mod_no_climate <- gam(code ~ s(TPI,k=4)+s(LATW,k=4), data=data,family=binomial)

mod_null <- gam(code ~ 1,data=data,family=binomial)
```

bc_12 - Mean Annual Precipitation (mm) from Worldclim2 (Fick et al. 2017)

bc_15 – Precipitation Seasonality from from Worldclim2 (Fick et al. 2017)

LATW – Human Influence Score from Last of the Wild v2 (Wildlife Conservation Society et al. 2005)

TPI – Topographic Position Index (Amatulli et al., 2018)

The *gam.check* function was used to check model convergence, and model convergence was achieved for all models. Checking of patterns in model residuals indicated the number of knots (k) may have been too low for the data, so the global model was re-run with 6 and 8 knots. Doing so did not improve the *gam.check* diagnostics around knot number, indicating a low number of knots was not driving the residual pattern. In addition, increasing the number of knots only made minor improvements to the proportion deviance explained and AIC scores of the model, as in table below. For this reason, 4 knots for each variable were retained.

Knots	%DE	AIC
4	44.1	12236
6	44.6	12146
8	45.1	12037

Model concurvity for the global model was tested for the global model using the *concurvity* function. The worst, observed and estimate concurvity statistics were close to zero, indicating model concurvity was not a problem.

A binomial generalized linear model was tested as an alternative to the generalized additive model. This model had a deviance explained of 42% and an AIC score of 12510, indicating a worse fit than the generalized additive model with splines, so the spline model was retained for the analysis.

Table S1. Top candidate models of binomial generalized additive models that predict global forest probability.

Region	MAP	Seasonality	Human influence index	Topographic index	df	AICc	ΔAICc	*%DE*	*w*
Global (n=18,489)									
	+	+	+	+	12.2	12,820	0	47.9	1
	+	+	+		9.3	12,883	63	47.6	0
	+	+		+	9.6	13,036	217	47.0	0
	+	+			6.9	13,076	256	47.0	0
			+	+	6.9	23,696	10,876	3.5	0
Null					1.0	24,549	11,729	0	0

AICc refers to Akaike Information criterion corrected for small sample size, ΔAICc to the difference between each model's AICc and the minimum AICc found, *w* to Akaike weights

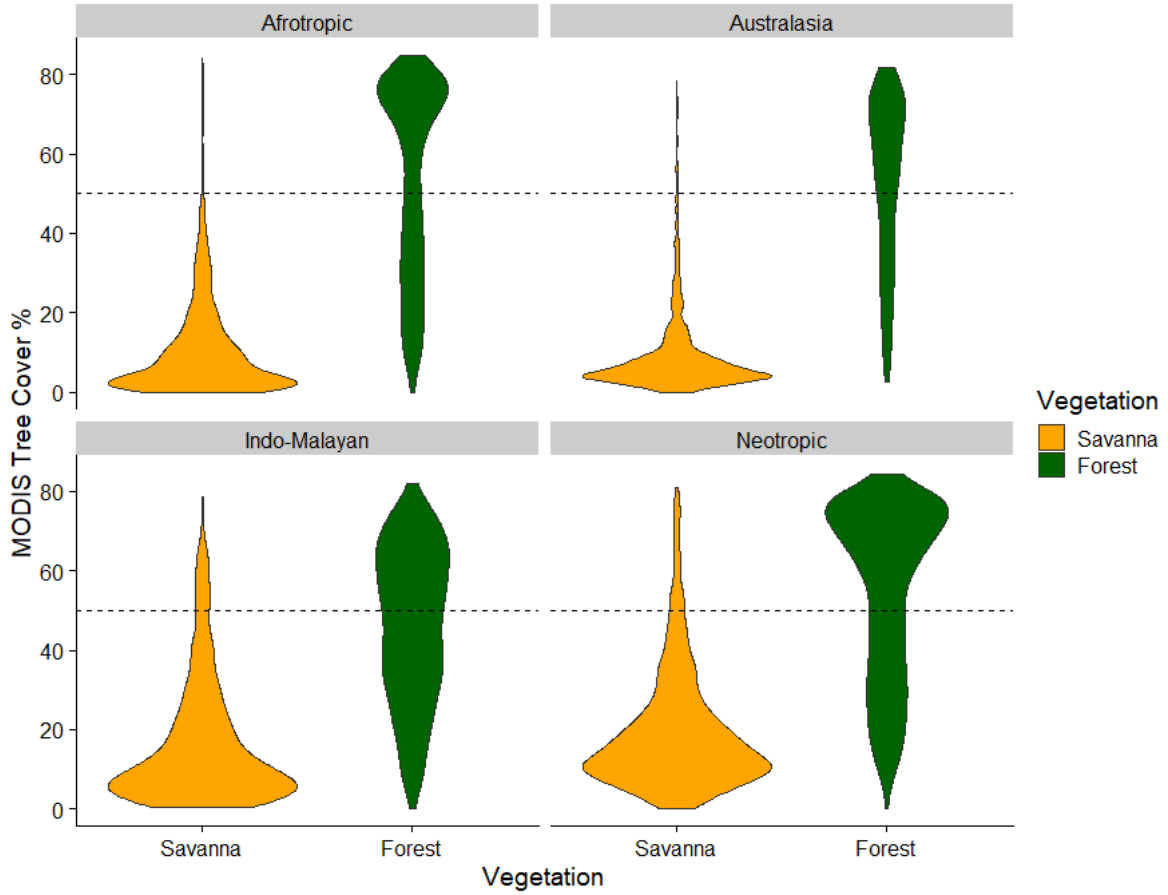

Fig. S1. Violin plots of point classification plotted against MODIS tree cover percentage, with 50% cover indicated by dashed line, showing that the majority of classified points fall within the MODIS tree cover percentages.

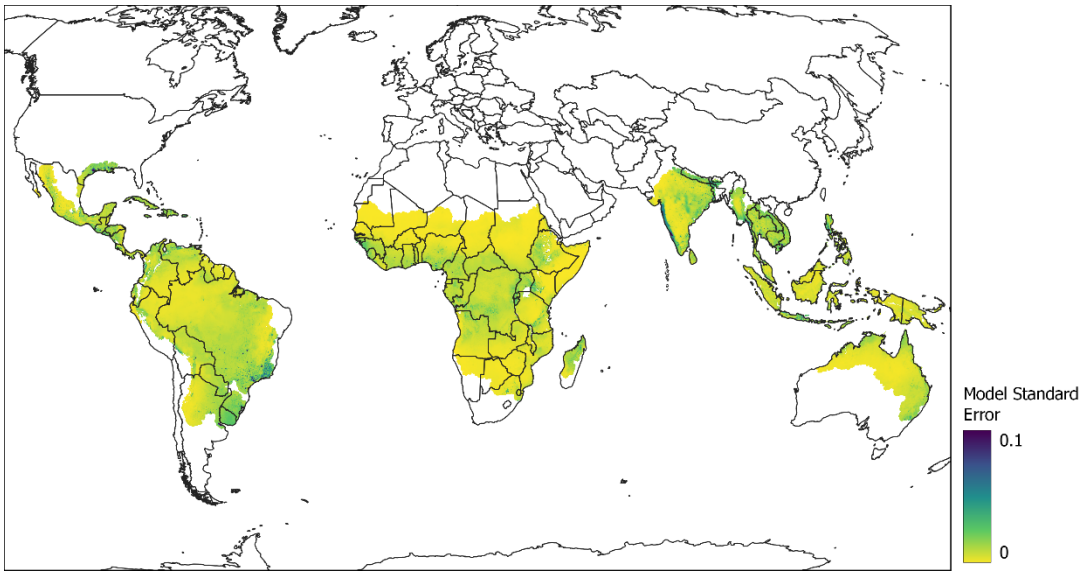

Fig. S2. Map of the model standard errors. Yellow areas have low standard errors, while areas tending towards green to blue have higher standard errors and therefore variance in model prediction.