

RESEARCH ARTICLE

Seasonal and inter-annual drivers of yellow fever transmission in South America

Arran Hamlet¹*, Katy A. M. Gaythorpe¹, Tini Garske¹, Neil M. Ferguson

MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom

* arran.hamlet14@imperial.ac.uk

OPEN ACCESS

Citation: Hamlet A, Gaythorpe KAM, Garske T, Ferguson NM (2021) Seasonal and inter-annual drivers of yellow fever transmission in South America. *PLoS Negl Trop Dis* 15(1): e0008974. <https://doi.org/10.1371/journal.pntd.0008974>

Editor: Andrew S. Azman, Johns Hopkins Bloomberg School of Public Health, UNITED STATES

Received: October 25, 2019

Accepted: November 11, 2020

Published: January 11, 2021

Copyright: © 2021 Hamlet et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: We acknowledge joint Centre funding from the UK Medical Research Council and Department for International Development. (Grant reference: MR/R015600/1, Affiliation: MRC Centre for Global Infectious Disease Analysis) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

In the last 20 years yellow fever (YF) has seen dramatic changes to its incidence and geographic extent, with the largest outbreaks in South America since 1940 occurring in the previously unaffected South-East Atlantic coast of Brazil in 2016–2019. While habitat fragmentation and land-cover have previously been implicated in zoonotic disease, their role in YF has not yet been examined. We examined the extent to which vegetation, land-cover, climate and host population predicted the numbers of months a location reported YF per year and by each month over the time-period. Two sets of models were assessed, one looking at interannual differences over the study period (2003–2016), and a seasonal model looking at intra-annual differences by month, averaging over the years of the study period. Each was fit using hierarchical negative-binomial regression in an exhaustive model fitting process. Within each set, the best performing models, as measured by the Akaike Information Criterion (AIC), were combined to create ensemble models to describe interannual and seasonal variation in YF. The models reproduced the spatiotemporal heterogeneities in YF transmission with coefficient of determination (R^2) values of 0.43 (95% CI 0.41–0.45) for the interannual model and 0.66 (95% CI 0.64–0.67) for the seasonal model. For the interannual model, EVI, land-cover and vegetation heterogeneity were the primary contributors to the variance explained by the model, and for the seasonal model, EVI, day temperature and rainfall amplitude. Our models explain much of the spatiotemporal variation in YF in South America, both seasonally and across the period 2003–2016. Vegetation type (EVI), heterogeneity in vegetation (perhaps a proxy for habitat fragmentation) and land cover explain much of the trends in YF transmission seen. These findings may help understand the recent expansions of the YF endemic zone, as well as to the highly seasonal nature of YF.

Author summary

Yellow fever (YF) is a viral haemorrhagic fever found in tropical South America and Africa that affects both humans and non-human primates (NHPs). Despite a long-standing recognition of YF as a pathogen of significant public health concern, not much is known about why cases are reported in some years and not others, or what drives the strong seasonal trends that are observed in South America.

Competing interests: The authors have declared that no competing interests exist.

Using a combination of statistical techniques, the authors have looked at the relationship between different types of land-use, vegetation, climate, human and NHP and the reporting of YF across all of South America and both inter-annual and seasonally over the period 2003–2016. The authors have found that both are highly influenced by changes in vegetation, however the inter-annual is additionally influenced by land-cover and the seasonal climatic factors.

The authors have found that the differing drivers of seasonal and inter-annual transmission, though both are predominantly shaped by changes in vegetation, highlighting the role of land-cover in influencing the inter-annual and climate the seasonal.

This research enhances our understanding of YF transmission in South America, describing the geographic and temporal distribution of cases in relation to vegetation, land-cover and climate. This research is of particular importance, given the recent large-scale outbreaks of YF across the continent, global YF vaccination shortages and the exportation of cases globally.

Introduction

Disease transmission is influenced by both intra- and interannual variations in weather and the environment, particularly for vector-borne pathogens [1–3]. These deviations may be relatively short in duration, due to seasonal changes in weather [2] or phenomena such as El Niño [4], or they may represent a more persistent change, due to climate change [5] or alterations to land-cover [6]. While climate change is likely to alter both the distribution and intensity of a number of diseases [7–9], this process takes place over a substantially longer period of time than anthropogenic land conversion which can completely change large swathes of natural habitat in a few years [10,11]. Rapid habitat change is often associated with disease occurrence [12], especially of zoonotic infections [13], potentially due to an increased interaction between sylvatic reservoirs and humans, expressly at intermediate levels of transformation [14].

Yellow fever (YF) is a zoonotic disease caused by the yellow fever virus (YFV), a flavivirus arbovirus infecting both humans and non-human primates (NHPs) [15]. Originating in Africa, YF spread to South America with the slave trade [16] and is currently endemic in 34 countries in Africa and 13 in South America [17]. In South America, YFV transmission occurs in two cycles, the sylvatic and urban. In the former, transmission is maintained by sylvatic mosquito species of the *Haemagogus* and *Sabethes* genera between NHPs, with humans considered incidental hosts. If the virus establishes itself in the domestic *Aedes aegypti*, also a vector of both dengue and Zika viruses, transmission can be sustained in the absence of a NHP reservoir. This can cause large and explosive outbreaks, the latest being the 2015–2016 outbreak in Angola and the Democratic Republic of the Congo, the largest in the past 30 years [18].

In South America the sylvatic cycle has accounted for almost all cases since 1942 [19], and has historically been confined to Amazonian regions. However, over the past 20 years the area where YF is endemic in NHPs has seen rapid geographic expansion. In Brazil this has resulted in 5 reassessments of the zone of YF endemicity since 2000, with the latest update in 2018 including the entire country [20,21]. The reasons for this expansion are unknown. Furthermore, outbreaks in Brazil's South-East Atlantic forest in 2016–2017 and 2017–2018 have been the largest ever recorded in the country, in humans and NHPs, with cases reported in states that have never previously recorded YF [22]. While there is no evidence of urban transmission in these outbreaks, confirmed human and NHP cases in the vicinity of Brazil's largest cities—areas with high *Aedes aegypti* density—is a cause for concern [23,24]. In addition to a lack of

understanding of these interannual drivers of transmission, there is a general dearth of understanding of the seasonality of YF. Despite the seasonality of YF having long been well established [25,26], there has been little research on quantifying the associated environmental and climatic drivers of seasonality in South America

We investigated the drivers of YF transmission both intra and interannually, using temporally varying covariates related to climate, land-cover, vegetation and human and NHP demographics. Covariates related to these were selected based on their previously demonstrated roles in vector-biology, increased suitability for disease transmission or relationships with YF [27–30]. In particular, we examined the role of vegetation cover and its heterogeneity/fragmentation. While the influence of habitat fragmentation in YFV transmission has previously been postulated [14,31,32], detailed research into the role habitat fragmentation and land-cover play in YFV sylvatic spillover is absent.

By utilising two model structures, we investigated the differential drivers of seasonal, and interannual variation in YF incidence. These models were fit to the number of months reporting YF at each administrative level 1 geographic unit (for example, the province or state), using exhaustive model fitting in hierarchical negative-binomial regressions. The best performing models, as defined by the Akaike Information Criterion (AIC) were weighted and combined using Akaike weights to produce an ensemble model [33]. Model robustness was confirmed using spatial block bootstrapping.

Materials and methods

YF data

Reports of YF cases in humans were assembled from various sources, including the Weekly Epidemiological Record [34], Disease Outbreak News [35], and the Pan American Health Organization [36] for the period 2003–2016. Only reports where the month of symptom onset was recorded were included (823 of the original 1073 reports), and these were geo-located to the first sub-national administrative level, here termed province.

Two datasets were derived from our report database. For each, we classified each month (over the 14 years of the data) for each province as a report month if one or more YF cases had onset dates in that month. This resulted in 397 report months over 165 unique provinces. For the interannual analysis, numbers of report months were summed within each year for each province—giving a dataset of the number of report months for each province for each of the 14 years considered. For the seasonal dataset, numbers of report months were summed over years for each province and each of the 12 months of the calendar year.

The inter-annual dataset, which uses the calendar year is a simplification of long term (multi-year) transmission patterns. Disease transmission likely does not conform fully to these demarcations of years, and so may not be fully captured by our usage of the calendar year for inter-annual transmission. However, in the absence of previously defined seasonal patterns for each administrative location (which will likely change across the region of study) we have defaulted to the current World Health Organization/Pan American Health Organization format which uses the simple calendar year [36].

Covariates

In total, 19 covariates were considered (Table 1). These were selected based on knowledge of the biology and distributions of vector species, host dynamics, inferences from the role of land-cover change and vegetation heterogeneity and the epidemiology of yellow fever in South America [2,30,37,38]. For the temporally changing covariates in the seasonal model, the values of the first year were subtracted from the final year of study (2003 and 2016 respectively), for

Table 1. Table of classification types and the covariates included.

Covariate	Classification	Temporal resolution	Contributing covariate types	Description	Reference
Monkey species (count)	Host population	Static		The number of monkey species	(41)
Logarithm of human population	Host population	Static		The logarithm of human population	(39)
Change in logarithm of human population		Static		The change in the logarithm of human population from 2003 to 2016	
Enhanced Vegetation Index (EVI) (0–1)	Vegetation	Annual/ monthly		A vegetation index designed to improve sensitivity in high biomass regions	(43)
EVI amplitude	Vegetation	Annual/ monthly		The amplitude of the EVI	
Day temperature (°C)	Climate	Annual/ monthly		The day temperature	(42)
Day temperature amplitude (°C)	Climate	Annual/ monthly		The amplitude of the day temperature	
Rainfall	Climate	Annual/ monthly		The rainfall	(44)
Rainfall amplitude	Climate	Annual/ monthly		The amplitude of the rainfall	
Forest cover	Land-cover	Annual	MCD12Q1 –Value 1	The proportion of the administrative unit covered by Evergreen Needleleaf forest	(43)
		Annual	MCD12Q1 –Value 2	The proportion of the administrative unit covered by Evergreen Broadleaf forest	
		Annual	MCD12Q1 –Value 3	The proportion of the administrative unit covered by Deciduous Needleleaf forest	
		Annual	MCD12Q1 –Value 4	The proportion of the administrative unit covered by Deciduous Broadleaf forest	
		Annual	MCD12Q1 –Value 5	The proportion of the administrative unit covered by Mixed forest	
Savanna cover	Land-cover	Annual	MCD12Q1 –Value 8	The proportion of the administrative unit covered by Woody savanna	
		Annual	MCD12Q1 –Value 9	The proportion of the administrative unit covered by Savanna	
Cropland/natural vegetation mosaic cover	Land-cover	Annual	MCD12Q1 –Value 14	The proportion of the administrative unit covered by cropland/natural vegetation mosaic	
Urban cover	Land-cover	Annual	MCD12Q1 –Value 13	The proportion of the administrative unit covered by urban areas	
Cropland cover	Land-cover	Annual	MCD12Q1 –Value 12	The proportion of the administrative unit covered by cropland	
Forest cover temporal change	Land-cover change	Annual		Interannual model: The current year landcover–previous year landcover value Seasonal model: The final year (2016)–the first year (2003) landcover value	
Savanna cover temporal change	Land-cover temporal change	Annual			
Natural vegetation/cropland mosaic cover temporal change	Land-cover temporal change	Annual			
Urban cover temporal change	Land-cover temporal change	Annual			
Cropland cover temporal change	Land-cover temporal change	Annual			
Vegetation heterogeneity	Vegetation heterogeneity	Annual/ monthly		This is the standard deviation of the EVI at a 1x1km resolution within the administrative unit	
Vegetation heterogeneity temporal change	Vegetation heterogeneity temporal change	Annual/ monthly		Interannual model: The current year landcover–previous year vegetation heterogeneity value Seasonal model: The final year (2016)–the first year (2003) vegetation heterogeneity	

<https://doi.org/10.1371/journal.pntd.0008974.t001>

the interannual covariates, we subtracted the value of the one year from the next year (i.e. 2003 land cover– 2002 land cover produced the 2003 land cover temporal change).

Covariates were standardised to facilitate comparison through the following formula,

$$z = \frac{x - \mu}{\sigma},$$

where z , is the standardised value, x , the pre-standardised value, μ , the mean of the pre-standardised values and σ , the standard deviation of the values. Standardised coefficient values for the ensemble models of the interannual and seasonal models are found in [Table 1](#).

The variable importance refers to how a measurement score decreases when a feature is not available. Initially the full model with the initial dataset is fit, and the R^2 value calculated. Then the model is refit to a modified initial dataset, where the variable of interest within the dataset been assigned the mean value of that covariate. This is done to produce “dummy data” which creates a covariate that does not provide any useful information. The R^2 is then calculated from this refit, and the variable importance for variable, i , calculated through,

$$\text{variable importance}_i = 1 - \frac{\text{refit } R^2_i}{\text{original } R^2_i}.$$

This provides a measure of the feature importance with respect to the original R^2 .

Host population data

Country and year specific human population sizes were obtained from the UN World Population Prospects [39] and averaged over the study period to obtain average population sizes. Province level estimates of population were obtained by disaggregating this data by using LandScan 2015 [40] population estimates with a 1/120 degree resolution to calculate the proportion of the national population within each province. The mean logarithm of human population over the time-period was used in all seasonal and interannual models. In addition, the relative change in the human population over the 14-year time period was also tested as a covariate (defined as $\log(\text{population in 2016}/\text{population in 2003})$).

Information on NHP species distribution was obtained through distribution maps of mammals in the western hemisphere [41]. These data were available as demarcations of distribution, which was geo-located to the province level. This was used to calculate the number of NHP species present in each province.

Climate, vegetation and vegetation heterogeneity data

Datasets, 2003–2016, for temperature [42], enhanced vegetation index (EVI) [43] and rainfall [44] were aggregated to the administrative unit 1 level from their original resolutions (of between 1/120 and 1/12 degree) by calculating population-weighted means, based on the population distribution provided by LandScan 2015 [40]. Here the climate and vegetation data is weighted by the population present, provided by LandScan 2015, and aggregated up to the administrative unit level, this weighting is used to provide climate/vegetation data that is representative of human interaction. The amplitude of the annual cycle Fourier component of these variables was also calculated, to account for the impact of seasonal variation on reports.

Spatial heterogeneity in vegetation was assessed by evaluating the standard deviation of the enhanced vegetation index (EVI) at its original 1/120 degree resolution within an administrative unit.

These covariates were averaged over time using different methods for the interannual and seasonal model datasets. For the seasonal dataset, monthly covariates were provided by taking

the mean covariate value in a month across all years 2003–2016 to provide the monthly average over this time-period. For the inter-annual dataset, the mean value of the covariate in a year was used.

Land-cover data

Land-cover was provided by the MODIS dataset [45], which characterises the dominant land-cover type, 1 of 17, at a grid resolution of 0.8333° globally. This information was aggregated to the province level and the proportion of the province area occupied by each land-cover type calculated. Forest, savanna and shrub land types were summed to provide overall forest and savanna cover. For the seasonal model, the mean land-cover proportions for an administrative unit across the study period (2003–2016) were used. For the inter-annual dataset, land-cover was provided for each year. The inter-annual dataset is included with this submission as [S1 Data](#) and the seasonal dataset as [S2 Data](#).

Regression models

Following initial covariate exploration, a list of covariates identified as relevant to YFV transmission were considered, with log of human population and the fractional change in logarithm human populations included in every model. By considering an exhaustive combination of all 19 covariates, we had 524,288 model structures for the interannual and seasonal frameworks, for a total of 1,048,576 models.

These were fit to either the number of months reporting yellow fever each year (interannual model) or the sum across years of the number of yellow fever reports in each calendar month (seasonal model) using hierarchical negative binomial regression models [46]. A negative binomial model was used due to its appropriateness for measuring count data, and its suitability for considering the overdispersion of the data.

Conceptually, hierarchical models are similar to running a standard regression where each row in the dataset refers to an administrative location and a time point (month or year depending on the model structure). By utilising a hierarchical structure however, we can allow parameters to vary between administrative location to avoid introducing biases that arise from treating temporally varying covariates within a location as independent [47]. Here we allow the intercept to vary by administrative location to account for this. These models are shown through the following equations [48,49],

$$Y_i | E_i \sim \text{Poisson}(E_i)$$

$$E_i = \mu_i$$

$$\ln(\mu_i) = \beta_0 + \beta_1 X_i + \beta_2 X_i + \dots + \beta_m X_{mi} + \varepsilon_i$$

$$E_i \sim \text{Gamma}(\lambda_i, K_i)$$

Where Y_i is the report months of YF in a province, and X_i the explanatory covariates and ε_i represents the random intercept as defined by the province. E_i , λ_i and K_i are the distribution parameters where E_i has a Gamma distribution with parameter λ_i and K_i with the negative binomial distribution, the mean and variance are

$$E(Y_i) = \mu_i$$

$$\text{Var}(Y_i) = \mu_i + \frac{\mu_i^2}{K_i}.$$

Models were then ranked based on their Akaike Information Criterion (AIC) and those with an AIC within 3 of the best performing model, as defined as the model with the lowest AIC value, were combined using Akaike weights [33]. To do so, the relative differences in AIC are calculated by,

$\Delta_i = AIC_i - \min(AIC)$ and this is used to obtain an estimated relative likelihood of model, i , in proportion to the other models, $k = 1 \dots K$, included through,

$$w_i = \frac{\exp\{-\frac{1}{2}\Delta_i\}}{\sum_{k=1}^K \exp\{-\frac{1}{2}\Delta_k\}}.$$

The product of each of these model specific weights, w_i , and their corresponding model specific predicted values, p_i , are summed to generate a single set of weighted predictions, p_A ,

$$p_A = \sum_{k=1}^K w_k p_k.$$

Out-of-sample performance was ascertained using a stringent method of cross-validation called spatial block bootstrapping (See [S1 Text](#) and [S3 Fig](#)).

Results

Geographical, seasonal and interannual heterogeneities in YF reports

We identified 397 unique months with a report of YF, hereby termed report months (defined spatiotemporally by the administrative unit and month), for the period 2003–2016, in 432 level 1 administrative units across 8 countries ([Fig 1](#)). Peru, Colombia and Brazil accounted for 79% of all report months, with Peru alone accounting for 39% ([Figs 1A and 2](#)). Within countries, report months show substantial spatial heterogeneity, with a notable clustering in Amazonian regions of Brazil, eastern Peru and Northern Bolivia. States in the South-East Atlantic coast of Brazil have also recorded large numbers of report months ([Fig 1B](#)) (21).

The frequency of report months was relatively stable and high during 2003–2008, after which numbers fell, then plateaued until 2015, when they dipped to the lowest level seen with only 1 reported event ([Fig 1A](#)). It should be emphasised that report months are a presence/absence indicator and not a proxy for infection incidence. Throughout the endemic zone, YF follows highly seasonal patterns. At the continent scale, transmission is highest from December to February, before dropping to a relatively low level over June to September, and a period of minimal occurrence in October and November ([Fig 2](#)). However, this pattern varies slightly by country and latitude (see [S2 Text](#)).

Geographic distributions of model predictions

In total 46 inter-annual and 246 seasonal models had AIC values within 3 of the best (lowest AIC) performing model and so were included in ensemble models.

The ensemble interannual and seasonal models accurately approximate spatiotemporal heterogeneities in YF reports, with coefficient of determination (R^2) values of 0.43 (95% CI 0.41–0.45) for the interannual and 0.66 (95% CI 0.64–0.67) for the seasonal ensemble predictions.

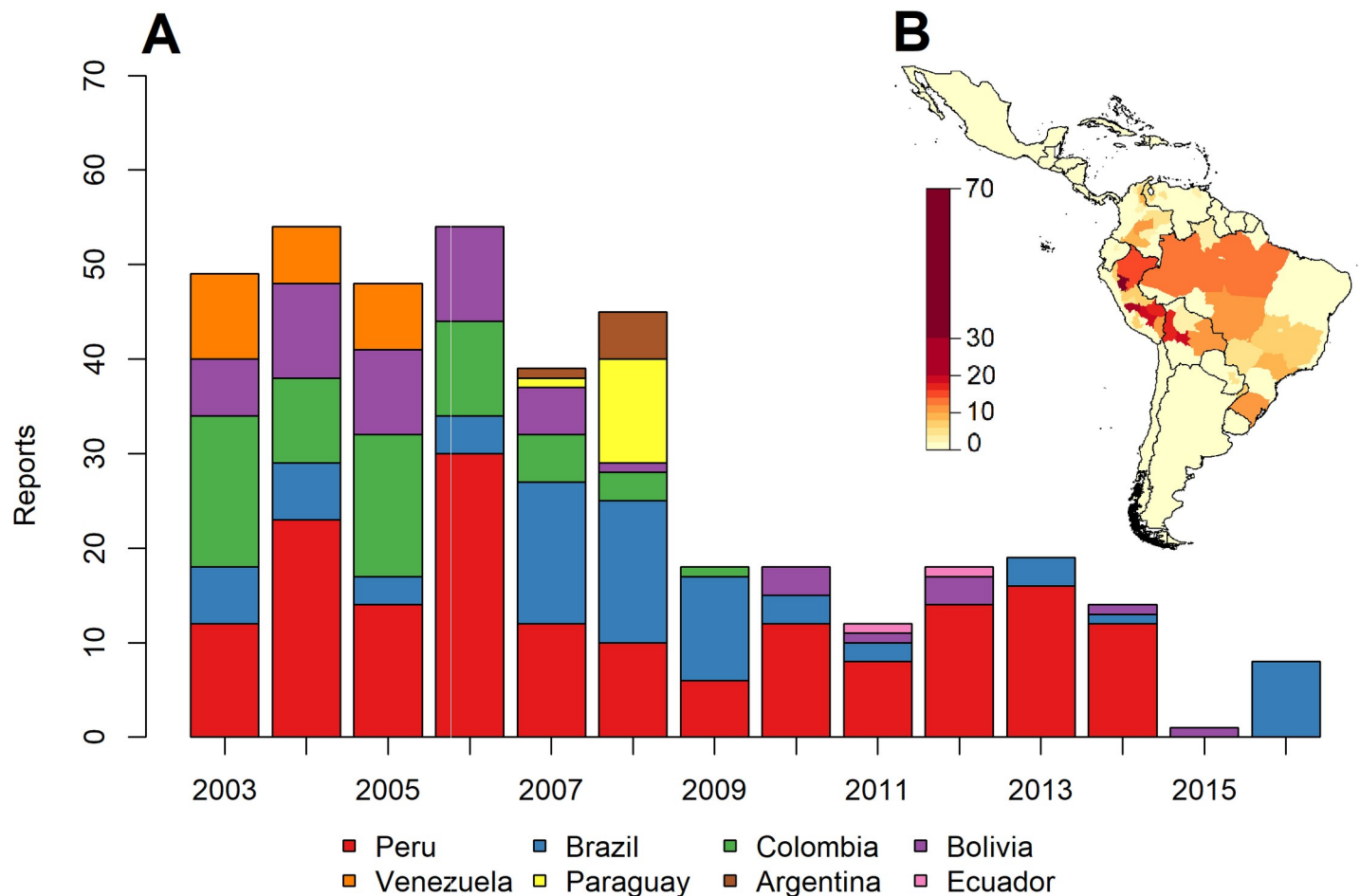


Fig 1. (A) Number of yellow fever report months over time (2003–2016) by country. (B) Total number of yellow fever reports by province (2003–2016) across South America. Figs were produced using the programming language R version 3.5.1 and used publicly available data gathered from the Weekly Epidemiological Record published by the WHO [34].

<https://doi.org/10.1371/journal.pntd.0008974.g001>

Due to the additional rigour of using spatial block bootstrapping compared with using an entirely random validation set (See SI), out-of-sample prediction R^2 values were lower at 0.31 (95% CI 0.28–0.34) for the interannual model and 0.45 (95% CI 0.44–0.48) for the seasonal model.

Model predictions were summed over time for each model to facilitate visual comparison with the data (Fig 3A and 3B vs Fig 1B). Both models reproduce the observed geographic distributions of reports well, though the aggregate ensemble seasonal model predictions give a better fit to the data (Fig 3). Differences between the ensemble interannual model predictions and the data range from -4.35 to +3.08. The model over-predicts reports for much of Eastern Peru and the North-West of Brazil, and predicts fewer reports than observed for Rio Grande do Sul in Brazil, and Misiones province in Argentina. There is additionally a cluster of lower than observed predictions on the Colombian/Venezuelan border. Ensemble seasonal model predictions showed deviations from the data an order of magnitude smaller than seen for the interannual model. The seasonal model slightly underpredicts YF reports, with only Brazilian states in the Amazon, Rio de Janeiro, and the East/North-East of the country predicted as having more reports than observed.

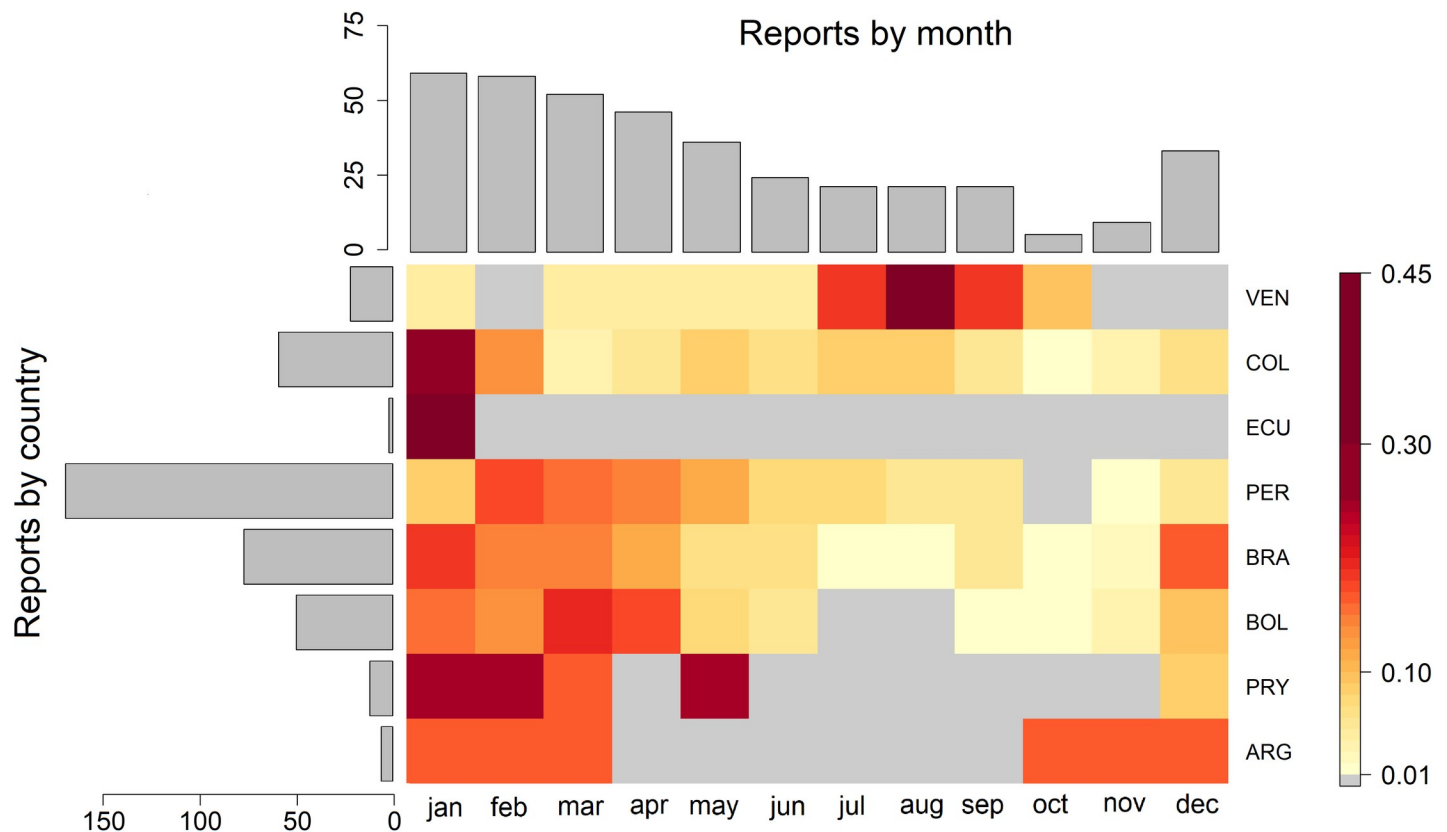


Fig 2. Yellow fever reports by country and month. The heatmap shows the proportion of reports in a country by calendar month, the bar chart on the left-hand side shows the total number of reports by country and the bar chart above shows the total number of months reporting cases by month. Countries are ordered by latitude.

<https://doi.org/10.1371/journal.pntd.0008974.g002>

Temporal distributions of model predictions

In addition to representing geographic variation, the models also consider temporal heterogeneity in YF incidence (Fig 4). The interannual model and the seasonal model fit temporal trends with the in-sample R^2 values of 0.43 (95% CI 0.41–0.45) and 0.66 (95% CI 0.64–0.67) respectively.

At the continent level, the inter-annual and seasonal model replicate the trends, but not the overall magnitude of temporal variation in report months. In the Inter-annual model, report months are underpredicted until 2009, after which they are slightly over-predicted. In the seasonal model, the model underestimates the data until the 5th month, then over-predicts later months (Fig 4A and 4B). The accuracy of the models at the country level varies (see SI). When years and months are ranked by the number of report months, there is a high degree of concordance between predictions and the data. This is shown in the high Pearson correlation coefficient values between the predicted and actual rank of years and months, at 0.926 for the interannual model, and 0.873 for the seasonal model (Fig 4C and 4D).

Drivers of seasonal, annual and long-term yellow fever transmission

The interannual and seasonal ensemble models showed both similarities and differences in the predictors found to be most significant (Table 2). For both, the covariate grouping relating to host demographics were the most important, with log of human population explaining the most variance in both model sets. The number of NHP species present also had a smaller but significant contribution for each. Both demographic predictors were positively associated with

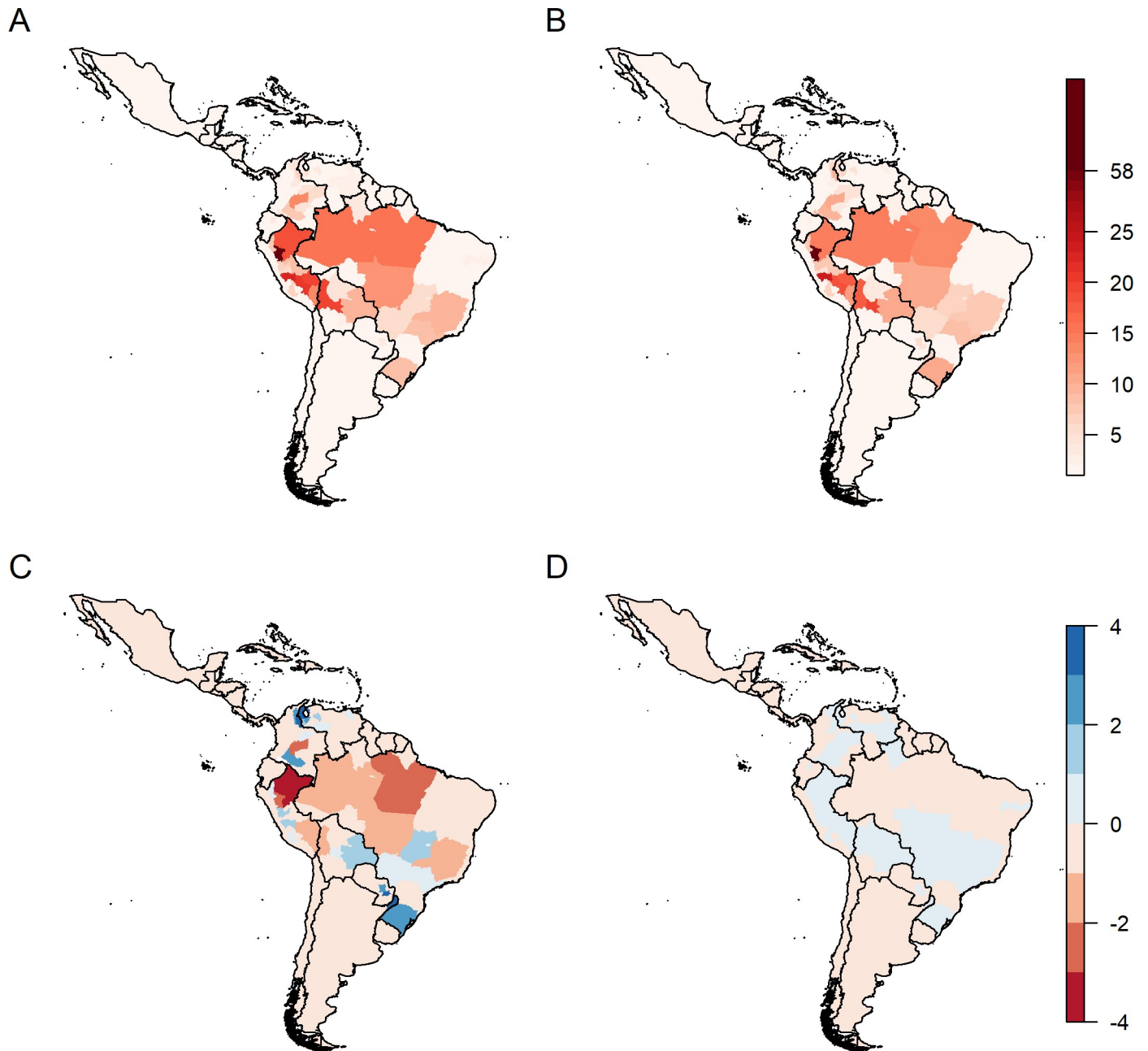


Fig 3. Ensemble model predictions of the number of YF report months for the (A) interannual model and the (B) seasonal model. (C) and (D) show the differences between these predictions and the data for the interannual model and the seasonal model, respectively. Figs were produced using the programming language R version 3.5.1 and the data was generated by the authors.

<https://doi.org/10.1371/journal.pntd.0008974.g003>

YF reports. The Enhanced Vegetation Index (EVI) was the second most important predictor for both models, again positively associated with YF reports. Other predictors differed between models, likely reflecting that the interannual model selected predictors best able to reflect long-term trends in YF reports, while the seasonal model selected those able to reproduce intra-annual seasonal patterns.

For the interannual model, landcover (cropland and savannah being negatively associated with YF) and vegetation heterogeneity (the standard deviation of EVI) were the next most

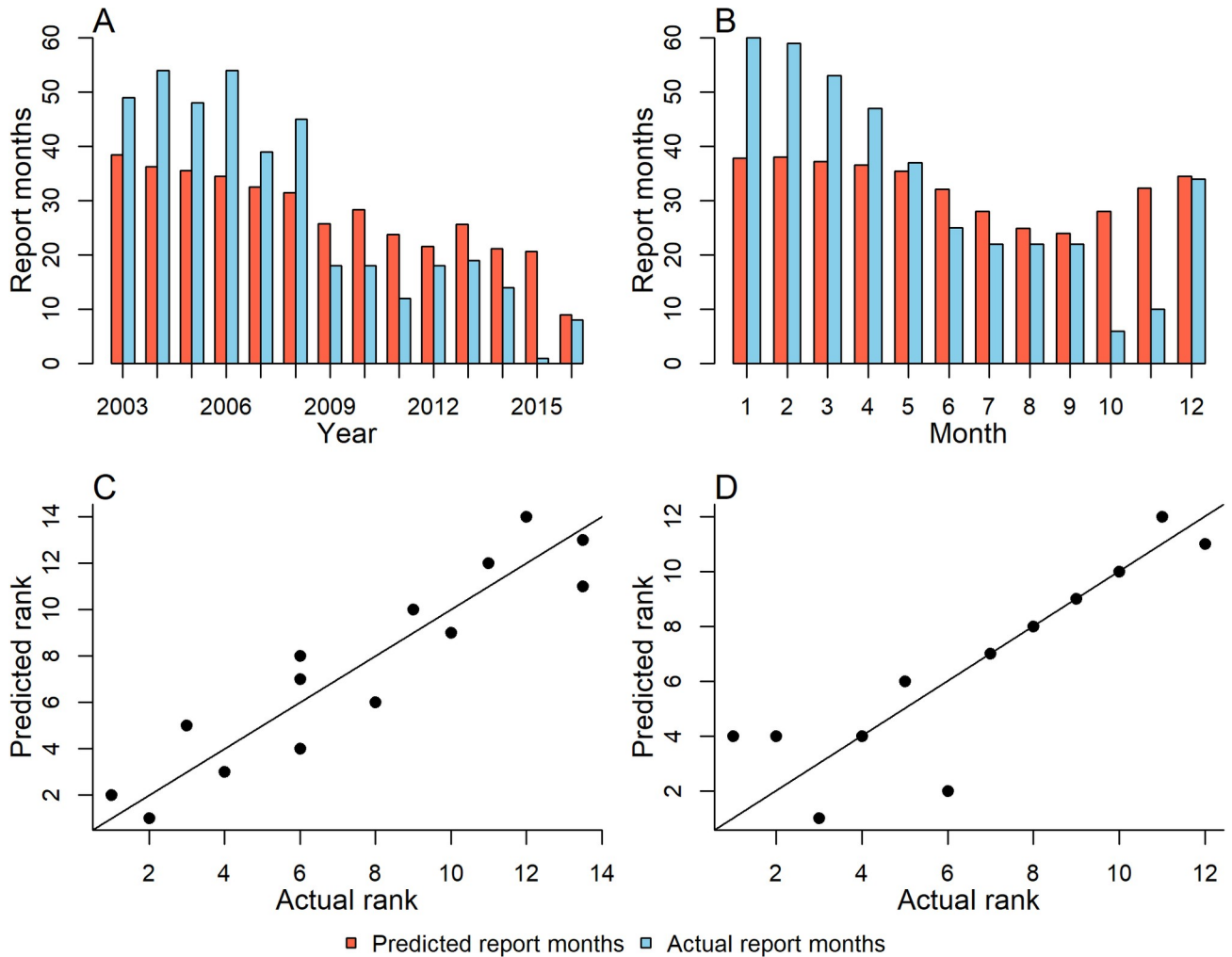


Fig 4. Summed ensemble model predictions (points) for (A) each year for the interannual model (A), and for each month for the seasonal model (B), contrast against the actual summed report months (lines) for each year or month. Yearly (C) and monthly (D) predictions ranked against the actual report months for the interannual and seasonal models, respectively (lines show predicted = actual).

<https://doi.org/10.1371/journal.pntd.0008974.g004>

important predictor groupings. Temporal changes between the current and previous year in vegetation and land-cover were significant predictors but made relatively small contributions to model fit. No climate coefficients were significant in the interannual model.

For the seasonal model, mean monthly day temperature and mean monthly rainfall amplitude (see [Materials and Methods](#) for definitions) were the other significant predictors, both negatively associated with YF reports.

Significant covariates were found in all (or almost) of the best performing models, with all covariate groupings found in interannual models except climate, and in the seasonal models only climate, vegetation and host demographics were found in the best performing models. Variable importance was highest in the EVI for both inter-annual and seasonal models, with the vegetation heterogeneity of a similar level of importance in the inter-annual model, and the number of NHP species and logarithm of human population slightly, but still important in the seasonal model. Despite the significance of the mean day temperature in the seasonal model, it was found to have an almost negligible variable importance—indicating it did not particularly contribute to predictive accuracy.

Table 2. Table of the permutation importance of different covariate groups, and individual covariates as well as standardised coefficient values. Only covariates that were significant in at least one of the model sets are shown. (A) Refers to the inter-annual model, and (B) the seasonal model.

Covariate groupings	% of models covariate group found in		Covariate	% of models covariate found in		Variable importance		Coefficient values	
	A	B		A	B	A	B	A	B
Climate	0	100	Mean day temperature	0	99.5	0	0.01		-2.71 (95% CI: -5.1 - -0.33)
			Mean rainfall amplitude	0	100	0	0.25		-0.35 (95% CI: -0.64 - -0.06)
Vegetation	100	100	EVI	100	100	0.82	0.84	6.42 (95% CI: 3.09–9.76)	6.21 (95% CI: 3.89–8.53)
			EVI amplitude	100	0	0.19	0	1.17 (95% CI: 0.33–2.01)	
Vegetation heterogeneity	100	0	Vegetation heterogeneity	100	0	0.79	0	3.34 (95% CI: 1.77–4.92)	
Vegetation heterogeneity temporal change	100	0	Vegetation heterogeneity temporal change	100	0	0.14	0	-0.39 (95% CI: -0.6 - -0.17)	
Land-cover	100	0	Cropland cover	100	0	0.21	0	-2.47 (95% CI: -4.27 - -0.66)	
			Savanna cover	100	0	0.5	0	-2.27 (95% CI: -3.73 - -0.8)	
Land-cover temporal change	100	0	Savanna cover temporal change	100	0	0	0	0.32 (95% CI: 0.1–0.54)	
Host demographics	100	100	Number of NHP species	100	100	0.48	0.74	1.34 (95% CI: 0.75–1.92)	1.68 (95% CI: 0.9–2.47)
			Log of human population	100	100	0.24	0.66	6.89 (95% CI: 0.62–13.16)	13.65 (95% CI: 5.19–22.11)

<https://doi.org/10.1371/journal.pntd.0008974.t002>

Discussion

In this study we have described the geographic, seasonal and interannual trends in YF reports in Latin America from 2003–2016, using publicly available data. We used hierarchical negative binomial regression models to create ensemble models predicting interannual and seasonal variation in YF transmission with a series of climatic, land-cover, vegetation and host demographic covariates. Our models explained a substantial amount of the observed variation, with R² values of 0.43 (95% CI 0.41–0.45) for the interannual and 0.66 (95% CI 0.64–0.67) for the seasonal model.

The geographic distribution of reports highlights “hotspots” for YF transmission, in Eastern Peru, North Western Peru and South Eastern Brazil (Fig 1B). The seasonal model reproduced these geographic trends more accurately than the interannual model. Continental-level inter-annual and seasonal trends in the data were also well-reproduced by the respective models, though both models captured geographic variation (e.g. at the country level) in these temporal trends less well (Figs 4 and S1 and S2)—albeit numbers of report months were often low when stratified by country. While at this level, the magnitude of temporal trends in report months are not fully captured, the relative ranking of years is and therefore model results can shed some light on what is associated with increased, or decreased, YF reporting in particular years and months.

While differing covariates are important for driving interannual and seasonal changes in YF transmission, vegetation (EVI) is highly influential for both models. This has been previously highlighted as a predictor of seasonal YF transmission [2], and potentially acts as a proxy for the interaction of rainfall and temperature, both important for arboviral transmission, while also taking into account a more complex interaction than is captured by either covariate alone. The potential additional complexity is highlighted through the absence of substantial correlations between either covariate and EVI. In both the interannual and the seasonal models, the log of human population was the most important predictor. This is not unexpected—larger populations give more opportunity for spillover, and since a report month is a month where one or more human YF cases are reported, larger populations are more likely to accumulate 1 or more cases in any one month even with a spatially invariant per-capita risk of YF.

While there is no detected relationship between EVI and population at this spatial and temporal timescale, there is potentially an interaction of population and EVI, with anthropogenic pressures having long-term consequences for the EVI. However, at this spatial and temporal scale these changes in relation to YFV transmission are hard to disentangle.

For the interannual model, landcover and heterogeneity in vegetation were also influential covariates in explaining interannual variation in YF reports. While cropland and savanna cover are negatively associated with YF reports, vegetation heterogeneity is positively associated. The heterogeneity covariate we adopted maybe acting as a proxy for habitat fragmentation. Fragmentation may affect sylvatic hosts in a number of ways, such as increasing their exposure to human contacts via modified behaviours [50,51] or increased susceptibility to infection due to a stress-weakened immune system [52]. Furthermore, vegetation heterogeneity may alter vector dynamics and predispose greater rates of spillover either through increased human-sylvatic cycle contact or favouring of more anthropophilic vector species in fragmented habitats [53]. These effects have previously been suggested to affect zoonotic disease transmission, but until now had not been statistically implicated in YF emergence [14,31].

While we have explained a substantial proportion of the seasonal and inter-annual variation in YF reporting across South America (2003–2016) (Interannual model: 0.43 (95% CI 0.41–0.45), seasonal model: 0.66 (95% CI 0.64–0.67)), this still means that, respectively, 67% and 34% of this variation is unexplained. This, in part, may be due to the spatial resolution at which the study was carried out. Due to data limitations in the reporting of YF cases, we may not have fully captured the relationship between climate and environment with YF spill over at the local or individual level. This may explain why some covariates that may be expected to be associated with increased spillover, such as forest cover and change in forest cover, have not been found to be significant. Furthermore, these covariate changes may actually occur, and remain, over several years. By solely investigating year to year variation in the inter-annual model, we may not be accurately capturing the importance of these covariates by failing to find significant effects to what may be a significant relationship. Additionally, the usage of the calendar year, rather than a disease specific “transmission” based description of the year may lead to us to unable to find these associations of covariates with transmission. To account for this, future modelling work should take place at a higher spatial resolution and considering the role of multi-year variation in covariates, though the trade-off between the availability and quality of data with a potentially furthered understanding should be thoroughly explored.

While climatic and landcover fluctuations both inter- and intra-annually lead to changes that can lead to increased disease transmission, they do not represent the whole picture of spillover. In order for YF to enter human populations it has to be both circulating within the NHP reservoir, and there has to be human exposure to the sylvatic cycle. Across South America (2000–2014), 60% of human cases of YF were in people employed in farming, hunting or fishing—highly seasonal activities [36]. This changing risk of exposure is likely to account for a proportion of the temporal and spatial reporting of YF. In order to better capture these relationships with YF spillover into human populations across South America, future modelling exercises should endeavour to capture both the underlying suitability to disease transmission, and these correlates and determinants of exposure.

This analysis uses 397 months of YF report months, where we only included publicly available case reports [34,36] which had a confirmed onset date and which could be geolocated to at least the province level. Due to missing data, 23% of case reports were excluded from our analysis. In addition, due to the remote locations that sylvatic YF is often found in and the non-specific symptoms many cases show, it is likely that substantial numbers of YF cases are never recorded [15,54]. Underreporting in rural areas may lead us to underestimate YF risk in those locations. However, surveillance and data quality issues affect estimation of absolute case

incidence, report months (presence/absence of cases in a specific administrative unit in a particular month) is likely to be more robust to under-ascertainment, as it only takes one reported case to be classified as YF positive. We are unable to identify whether the predictors of YF transmission we have identified affect sylvatic transmission or human exposure, given we have only analysed reports of human cases here. Data on NHP cases of YF across the continent are limited however, and their omission is a permissible oversight given this.

While expansion of the endemic zone is occurring, increases in population-level vaccination coverage in the endemic zones, where the majority of transmission is predicted, has precluded much of the human population from infection. This is in contrast with areas outside of this zone—where YF vaccination is either not usually necessary or not prioritised, and where spillover is more likely given the available of susceptible humans. This may go some way to explaining the decrease in report months over the time period (Fig 1).

In conclusion this body of work represents an important quantification of both the seasonality and interannual transmission of YF across South America (2003–2016). By identifying covariates, and their statistical relationship, with report months of YF, the work presented here may be used to highlight areas that have an increased probability for transmission. This may then allow for the targeting of surveillance in areas that have a higher risk of YF reporting, based on their climate and environment, without currently reported cases. This application could have substantial public health value, in a context where the geographic range of YF is changing and vaccine stocks are still limited.

Supporting information

S1 Fig. Inter-annual model predictions for the 8 countries reporting YF over the study period (2003–2016). The blue line indicates the data and the red dots the model predictions at the time-point.

(TIF)

S2 Fig. Seasonal model predictions for the 8 countries reporting YF over the study period (2003–2016). The blue line indicates the data and the red dots the model predictions at the time-point.

(TIF)

S3 Fig. A) The grid of 5° x 5° longitude of latitude with provinces assigned and colour coded by the grid point closest to their centroid coordinates. B) Examples of the training (blue) and validation (red) datasets as chosen by random sampling of grid points.

(TIF)

S1 Text. Country-level inter-annual and seasonal model ensemble predictions.

(DOCX)

S2 Text. Out-of-sample validation: Spatial block bootstrapping.

(DOCX)

S1 Data. Dataset used for inter-annual models.

(CSV)

S2 Data. Dataset used for seasonal models.

(CSV)

Author Contributions

Conceptualization: Arran Hamlet, Tini Garske, Neil M. Ferguson.

Data curation: Arran Hamlet, Tini Garske.

Formal analysis: Arran Hamlet.

Funding acquisition: Tini Garske, Neil M. Ferguson.

Investigation: Arran Hamlet.

Methodology: Arran Hamlet, Katy A. M. Gaythorpe, Tini Garske, Neil M. Ferguson.

Project administration: Tini Garske, Neil M. Ferguson.

Resources: Arran Hamlet, Katy A. M. Gaythorpe, Tini Garske, Neil M. Ferguson.

Software: Arran Hamlet.

Supervision: Arran Hamlet, Tini Garske, Neil M. Ferguson.

Validation: Arran Hamlet.

Visualization: Arran Hamlet.

Writing – original draft: Arran Hamlet.

Writing – review & editing: Arran Hamlet, Katy A. M. Gaythorpe, Tini Garske, Neil M. Ferguson.

References

1. Muturi EJ. Larval rearing temperature influences the effect of malathion on *Aedes aegypti* life history traits and immune responses. *Chemosphere*. 2013; 92(9):1111–6. <https://doi.org/10.1016/j.chemosphere.2013.01.055> PMID: 23419321.
2. Hamlet A, Jean K, Perea W, Yactayo S, Biey J, Van Kerkhove M, et al. The seasonal influence of climate and environment on yellow fever transmission across Africa. *PLoS neglected tropical diseases*. 2018; 12(3):e0006284. Epub 2018/03/16. <https://doi.org/10.1371/journal.pntd.0006284> PMID: 29543798; PubMed Central PMCID: PMC5854243.
3. Craig MH, Snow RW, le Sueur D. A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitology today*. 1999; 15(3):105–11. [https://doi.org/10.1016/s0169-4758\(99\)01396-4](https://doi.org/10.1016/s0169-4758(99)01396-4) PMID: 10322323.
4. Fuller DO, Troyo A, Beier JC. El Nino Southern Oscillation and vegetation dynamics as predictors of dengue fever cases in Costa Rica. *Environ Res Lett*. 2009; 4(1). Artn 01401110.1088/1748-9326/4/1/014011. WOS:000265878500012. <https://doi.org/10.1088/1748-9326/4/1/014011> PMID: 19763186
5. Lyon B, Dinku T, Raman A, Thomson MC. Temperature suitability for malaria climbing the Ethiopian Highlands. *Environ Res Lett*. 2017; 12(6). ARTN 06401510.1088/1748-9326/aa64e6. WOS:000403667800002. <https://doi.org/10.1088/1748-9326/aa7146> PMID: 30344619
6. Bauch SC, Birkenbach AM, Pattanayak SK, Sills EO. Public health impacts of ecosystem change in the Brazilian Amazon. *Proceedings of the National Academy of Sciences of the United States of America*. 2015; 112(24):7414–9. <https://doi.org/10.1073/pnas.1406495111> PMID: 26082548; PubMed Central PMCID: PMC4475939.
7. Curto de Casas SI, Carcavallo RU. Climate change and vector-borne diseases distribution. *Soc Sci Med*. 1995; 40(11):1437–40. Epub 1995/06/01. [https://doi.org/10.1016/0277-9536\(95\)00040-e](https://doi.org/10.1016/0277-9536(95)00040-e) PMID: 7667648.
8. Astrom C, Rocklov J, Hales S, Beguin A, Louis V, Sauerborn R. Potential Distribution of Dengue Fever Under Scenarios of Climate Change and Economic Development. *Ecohealth*. 2012; 9(4):448–54. <https://doi.org/10.1007/s10393-012-0808-0> WOS:000317970500011. PMID: 23408100
9. Artzy-Randrup Y, Alonso D, Pascual M. Transmission intensity and drug resistance in malaria population dynamics: implications for climate change. *PloS one*. 2010; 5(10):e13588. <https://doi.org/10.1371/journal.pone.0013588> PMID: 21060886; PubMed Central PMCID: PMC2965653.
10. Warren-Thomas EM, Edwards DP, Bebbler DP, Chhang P, Diment AN, Evans TD, et al. Protecting tropical forests from the rapid expansion of rubber using carbon payments. *Nature communications*. 2018; 9(1):911. <https://doi.org/10.1038/s41467-018-03287-9> PMID: 29500360; PubMed Central PMCID: PMC5834519.

11. Jokar Arsanjani J. Characterizing, monitoring, and simulating land cover dynamics using GlobeLand30: A case study from 2000 to 2030. *J Environ Manage.* 2018; 214:66–75. <https://doi.org/10.1016/j.jenvman.2018.02.090> PMID: 29518597.
12. Patz JA, Daszak P, Tabor GM, Aguirre AA, Pearl M, Epstein J, et al. Unhealthy landscapes: Policy recommendations on land use change and infectious disease emergence. *Environ Health Perspect.* 2004; 112(10):1092–8. <https://doi.org/10.1289/ehp.6877> PMID: 15238283; PubMed Central PMCID: PMC1247383.
13. Allen T, Murray KA, Zambrana-Torrel C, Morse SS, Rondinini C, Di Marco M, et al. Global hotspots and correlates of emerging zoonotic diseases. *Nature communications.* 2017; 8(1):1124. Epub 2017/10/27. <https://doi.org/10.1038/s41467-017-00923-8> PMID: 29066781; PubMed Central PMCID: PMC5654761.
14. Faust CL, McCallum HI, Bloomfield LSP, Gottdenker NL, Gillespie TR, Torney CJ, et al. Pathogen spill-over during land conversion. *Ecol Lett.* 2018; 21(4):471–83. Epub 2018/02/22. <https://doi.org/10.1111/ele.12904> PMID: 29466832.
15. Barrett AD, Higgs S. Yellow fever: a disease that has yet to be conquered. *Annu Rev Entomol.* 2007; 52:209–29. <https://doi.org/10.1146/annurev.ento.52.110405.091454> PMID: 16913829.
16. Bryant JE, Holmes EC, Barrett AD. Out of Africa: a molecular perspective on the introduction of yellow fever virus into the Americas. *PLoS pathogens.* 2007; 3(5):e75. <https://doi.org/10.1371/journal.ppat.0030075> PMID: 17511518; PubMed Central PMCID: PMC1868956.
17. Jentes ES, Pomeroy G, Gershman MD, Hill DR, Lemarchand J, Lewis RF, et al. The revised global yellow fever risk map and recommendations for vaccination, 2010: consensus of the Informal WHO Working Group on Geographic Risk for Yellow Fever. *Lancet Infect Dis.* 2011; 11(8):622–32. [https://doi.org/10.1016/S1473-3099\(11\)70147-5](https://doi.org/10.1016/S1473-3099(11)70147-5) PMID: 21798462.
18. World Health Organization. Yellow fever in Africa and the Americas, 2016. *Wkly Epidemiol Rec.* 2017; 92(32):442–52. PMID: 28799735.
19. Johansson MA, Arana-Vizcarrondo N, Biggerstaff BJ, Gallagher N, Marano N, Staples JE. Assessing the risk of international spread of yellow fever virus: a mathematical analysis of an urban outbreak in Asuncion, 2008. *The American journal of tropical medicine and hygiene.* 2012; 86(2):349–58. <https://doi.org/10.4269/ajtmh.2012.11-0432> PMID: 22302873; PubMed Central PMCID: PMC3269406.
20. Chaves T, Orduna T, Lepetic A, Macchi A, Verbanaz S, Riskey A, et al. Yellow fever in Brazil: Epidemiological aspects and implications for travelers. *Travel Med Infect Dis.* 2018; 23:1–3. Epub 2018/05/12. <https://doi.org/10.1016/j.tmaid.2018.05.001> PMID: 29751132.
21. Romano AP, Costa ZG, Ramos DG, Andrade MA, Jayme Vde S, Almeida MA, et al. Yellow Fever outbreaks in unvaccinated populations, Brazil, 2008–2009. *PLoS neglected tropical diseases.* 2014; 8(3):e2740. Epub 2014/03/15. <https://doi.org/10.1371/journal.pntd.0002740> PMID: 24625634; PubMed Central PMCID: PMC3953027.
22. Rezende IM, Sacchetto L, Munhoz de Mello E, Alves PA, Iani FCM, Adelino TER, et al. Persistence of Yellow fever virus outside the Amazon Basin, causing epidemics in Southeast Brazil, from 2016 to 2018. *PLoS neglected tropical diseases.* 2018; 12(6):e0006538. Epub 2018/06/05. <https://doi.org/10.1371/journal.pntd.0006538> PMID: 29864115; PubMed Central PMCID: PMC6002110.
23. Dorigatti I, Hamlet A, Aguas R, Cattarino L, Cori A, Donnelly CA, et al. International risk of yellow fever spread from the ongoing outbreak in Brazil, December 2016 to May 2017. *Euro Surveill.* 2017; 22(28). <https://doi.org/10.2807/1560-7917.ES.2017.22.28.30572> PMID: 28749337.
24. Massad E, Amaku M, Coutinho FAB, Struchiner CJ, Lopez LF, Coelho G, et al. The risk of urban yellow fever resurgence in Aedes-infested American cities. *Epidemiology and infection.* 2018:1–7. Epub 2018/05/31. <https://doi.org/10.1017/S0950268818001334> PMID: 29843824.
25. Kumm H. Seasonal variations in rainfall: Prevalence of Haemagogus and incidence of jungle yellow fever in Brazil and Colombia. *Transactions of the Royal Society of Tropical Medicine and Hygiene.* 1950; 43(6):673–82.
26. Monath TP, Vasconcelos PF. Yellow fever. *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology.* 2015; 64:160–73. <https://doi.org/10.1016/j.jcv.2014.08.030> PMID: 25453327.
27. Althouse BM, Hanley KA, Diallo M, Sall AA, Ba Y, Faye O, et al. Impact of climate and mosquito vector abundance on sylvatic arbovirus circulation dynamics in Senegal. *The American journal of tropical medicine and hygiene.* 2015; 92(1):88–97. <https://doi.org/10.4269/ajtmh.13-0617> PMID: 25404071; PubMed Central PMCID: PMC4347398.
28. Hamlet A, Jean K, Ferguson N, Van Kerkhove MD, Yactayo S, Perea W, et al. The seasonal influence of climate and environment on yellow fever transmission across Africa. 2017.

29. Wee LK, Weng SN, Raduan N, Wah SK, Ming WH, Shi CH, et al. Relationship between Rainfall and Aedes Larval Population at Two Insular Sites in Pulau Ketam, Selangor, Malaysia. *Se Asian J Trop Med*. 2013; 44(2):157–66. WOS:000327171400004. PMID: [23691624](#)
30. Hamrick PN, Aldighieri S, Machado G, Leonel DG, Vilca LM, Uriona S, et al. Geographic patterns and environmental factors associated with human yellow fever presence in the Americas. *PLoS neglected tropical diseases*. 2017; 11(9):e0005897. Epub 2017/09/09. <https://doi.org/10.1371/journal.pntd.0005897> PMID: [28886023](#); PubMed Central PMCID: PMC5607216.
31. Bicca-Marques JC, de Freitas DS. The role of monkeys, mosquitoes, and humans in the occurrence of a yellow fever outbreak in a fragmented landscape in south Brazil: protecting howler monkeys is a matter of public health. *Tropical Conservation Science*. 2010; 3(1):78–89.
32. de Almeida MAB, dos Santos E, Cardoso JD, da Silva LG, Rabelo RM, Bicca-Marques JC. Predicting Yellow Fever Through Species Distribution Modeling of Virus, Vector, and Monkeys. *Ecohealth*. 2019; 16(1):95–108. <https://doi.org/10.1007/s10393-018-1388-4> WOS:000462144000007. PMID: [30560394](#)
33. Wagenmakers EJ, Farrell S. AIC model selection using Akaike weights. *Psychon B Rev*. 2004; 11(1):192–6. <https://doi.org/10.3758/bf03206482> WOS:000220674200028. PMID: [15117008](#)
34. World Health Organization. The Weekly Epidemiological Record (WER). World Health Organisation.
35. World Health Organization. Disease Outbreak News (DON). World Health Organisation.
36. Pan American Health Organization. YELLOW FEVER: Number of Confirmed Cases and Deaths by Country in the Americas, 1960–2015 http://ais.paho.org/phis/viz/ed_yellowfever.asp: Pan American Health Organization.; 2017 [15/08/2017].
37. Kay RF, Madden RH, VanSchaik C, Higdon D. Primate species richness is determined by plant productivity: Implications for conservation. *Proceedings of the National Academy of Sciences of the United States of America*. 1997; 94(24):13023–7. <https://doi.org/10.1073/pnas.94.24.13023> WOS: A1997YJ45600060. PMID: [9371793](#)
38. Alencar J, Serra-Friere NM, Marcondes CB, Silva JD, Correa FF, Guimaraes AE. Influence of Climatic Factors on the Population Dynamics of Haemagogus Janthinomys (Diptera: Culicidae), a Vector of Sylvatic Yellow Fever. *Entomol News*. 2010; 121(1):45–52. WOS:000289595200007.
39. United Nations DoEaSA, Population Division, Population Estimates and Projections Section. World Population Prospects: The 2016 Revision 2016 [08 July 2016]. Available from: <http://esa.un.org/unpd/wpp/>.
40. Bright EA, Rose AN, Urban ML. LandScan 2015. In: Laboratory ORN, editor. <http://web.ornl.gov/sci/landscan/>: Oak Ridge National Laboratory; 2016.
41. Patterson BD, Ceballos G, Sechrest W, Tognelli MF, Brooks T, Luna L, et al. Digital Distribution Maps of Mammals of the Western Hemisphere, version 3.0. In: NatureServe, editor. Arlington, Virginia, USA2007.
42. Garske T, Ferguson NM, Ghani AC. Estimating air temperature and its influence on malaria transmission across Africa. *PLoS one*. 2013; 8(2):e56487. <https://doi.org/10.1371/journal.pone.0056487> PMID: [23437143](#); PubMed Central PMCID: PMC3577915.
43. NASA. Land Processes Distributed Active Archive Centre (LP DAAC) Vegetation Indices 16-Day L3 Global 1 km (13 A2) Sioux Falls, South Dakota: USGS/Earth Resources Observation and Science (EROS) Center; [13 July 2012]. Available from: http://lpdaac.usgs.gov/get_data.
44. Joyce R, Janowiak J, Arkin P, Xie P. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J Hydromet*. 2004; 5:487–503.
45. Friedl M, Sulla-Menashe D, Tan B, Schneider A, Ramankutty N, Sibley A, et al. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens Environ*. 2010; 114(1):168–82.
46. Seagroatt V. An introduction to medical statistics, 2nd edition—Bland, M. *J Psychosom Res*. 1996; 41(5):495–6. [https://doi.org/10.1016/S0022-3999\(96\)00098-0](https://doi.org/10.1016/S0022-3999(96)00098-0) WOS:A1996VZ06300012.
47. Gelman A, Hill J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press; 2006.
48. Hilbe JM. *Negative Binomial Regression*: Cambridge University Press; 2007.
49. Tran P, Waller L. Variability in results from negative binomial models for lyme disease measured at different spatial scales. *Environ Res*. 2015; 136:373–80. <https://doi.org/10.1016/j.envres.2014.08.041> WOS:000346755000046. PMID: [25460658](#)
50. Gottdenker NL, Chaves LF, Calzada JE, Saldana A, Carroll CR. Host Life History Strategy, Species Diversity, and Habitat Influence Trypanosoma cruzi Vector Infection in Changing Landscapes. *PLoS neglected tropical diseases*. 2012; 6(11). ARTN e188410.1371/journal.pntd.0001884. WOS:000311889000113. <https://doi.org/10.1371/journal.pntd.0001884> PMID: [23166846](#)

51. Goldberg TL, Gillespie TR, Rwego IB, Estoff EL, Chapman CA. In Forest Fragmentation as Cause of Bacterial Transmission among Primates, Humans, and Livestock, Uganda (vol 14, pg 1375, 2008). *Emerging infectious diseases*. 2008; 14(11):1825–. WOS:000260617000041.
52. Seltmann A, Czirjak GA, Courtiol A, Bernard H, Struebig MJ, Voigt CC. Habitat disturbance results in chronic stress and impaired health status in forest-dwelling paleotropical bats. *Conserv Physiol*. 2017; 5(1):cox020. Epub 2017/04/20. <https://doi.org/10.1093/conphys/cox020> PMID: 28421138; PubMed Central PMCID: PMC5388297.
53. Burkett-Cadena ND, Vittor AY. Deforestation and vector-borne disease: Forest conversion favors important mosquito vectors of human pathogens. *Basic Appl Ecol*. 2018; 26:101–10. <https://doi.org/10.1016/j.baae.2017.09.012> WOS:000427070700010.
54. Johansson MA, Vasconcelos PF, Staples JE. The whole iceberg: estimating the incidence of yellow fever virus infection from the number of severe cases. *Trans R Soc Trop Med Hyg*. 2014; 108(8):482–7. <https://doi.org/10.1093/trstmh/tru092> PMID: 24980556; PubMed Central PMCID: PMC4632853.