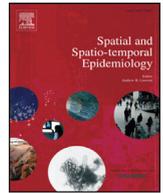


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Spatial and Spatio-temporal Epidemiology

journal homepage: [www.elsevier.com/locate/sste](http://www.elsevier.com/locate/sste)

## Improving disaggregation models of malaria incidence by ensembling non-linear models of prevalence

Tim C.D. Lucas<sup>a,\*</sup>, Anita K. Nandi<sup>a</sup>, Suzanne H. Keddie<sup>a</sup>, Elisabeth G. Chestnutt<sup>a</sup>, Rosalind E. Howes<sup>a</sup>, Susan F. Rumisha<sup>a,d</sup>, Rohan Arambepola<sup>a</sup>, Amelia Bertozzi-Villa<sup>b</sup>, Andre Python<sup>a</sup>, Tasmin L. Symons<sup>a</sup>, Justin J. Millar<sup>a</sup>, Punam Amratia<sup>a</sup>, Penelope Hancock<sup>a</sup>, Katherine E. Battle<sup>a</sup>, Ewan Cameron<sup>a</sup>, Peter W. Gething<sup>c,d</sup>, Daniel J. Weiss<sup>a</sup>

<sup>a</sup> Malaria Atlas Project, Big Data Institute, University of Oxford, Oxford, UK<sup>b</sup> Institute for Disease Modeling, Bellevue, WA, USA<sup>c</sup> Telethon Kids Institute, Perth Childrens Hospital, Perth, Australia<sup>d</sup> Curtin University, Perth, Australia

## ARTICLE INFO

## Article history:

Received 10 January 2020

Revised 13 April 2020

Accepted 18 June 2020

Available online 4 July 2020

## Keywords:

Spatial statistics

Disaggregation regression

Stacking

Surveillance data

## ABSTRACT

Maps of disease burden are a core tool needed for the control and elimination of malaria. Reliable routine surveillance data of malaria incidence, typically aggregated to administrative units, is becoming more widely available. Disaggregation regression is an important model framework for estimating high resolution risk maps from aggregated data. However, the aggregation of incidence over large, heterogeneous areas means that these data are underpowered for estimating complex, non-linear models. In contrast, prevalence point-surveys are directly linked to local environmental conditions but are not common in many areas of the world. Here, we train multiple non-linear, machine learning models on *Plasmodium falciparum* prevalence point-surveys. We then ensemble the predictions from these machine learning models with a disaggregation regression model that uses aggregated malaria incidences as response data. We find that using a disaggregation regression model to combine predictions from machine learning models improves model accuracy relative to a baseline model.

© 2020 The Author(s). Published by Elsevier Ltd.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

## 1. Introduction

High-resolution maps of malaria risk are vital for control and elimination (Battle et al., 2019; Weiss et al., 2019). However, mapping malaria in lower burden countries presents new challenges as traditional mapping of prevalence from cluster-level surveys (Battle et al., 2019; Bhatt et al., 2017; 2015; Weiss et al., 2019) is often not effective for two reasons. Firstly, so few individuals are infected that most surveys will detect zero positives (Sturrock et al., 2016). Secondly, there is a lack of nationally representative prevalence surveys in low burden countries (Sturrock et al., 2016; 2014). Routine surveillance data of malaria case counts, often aggregated over administrative regions defined by geographic polygons, is becoming more reliable and more widely available (Sturrock et al., 2016) and recent work has focussed on methods for estimating high-resolution malaria risk from these data (Johnson et al., 2019;

Law et al., 2018; Li et al., 2012; Sturrock et al., 2014; Taylor et al., 2017; Wilson and Wakefield, 2018). However, the aggregation of cases over space means that the data may be spatially uninformative, especially if the case counts are aggregated over large or heterogeneous areas, because it is unclear where within the polygon, and in which environments, the cases occurred. This data is therefore often under-powered for fitting flexible, non-linear models as is required for accurate malaria maps (Bhatt et al., 2017; 2015). A method that combines prevalence point-surveys and aggregated surveillance data, and therefore leverages the strength of both, has great potential.

Here we propose a two-stage method. In the first stage we train a suite of machine learning models, using point-level, binomial prevalence data and environmental covariates. In the second stage we combine predictions from these models by using them as covariates in a polygon-level, disaggregation regression model that uses malaria incidence (aggregated to administrative units) as the response. Unlike joint likelihood models (Wang et al., 2018), this method does not combine both prevalence and incidence data within one model. Instead the aim is to use the prevalence data

\* Corresponding author.

E-mail address: [timcdlucas@gmail.com](mailto:timcdlucas@gmail.com) (T.C.D. Lucas).

to find useful non-linear transformations of the environmental covariates which are then subsequently used in the disaggregation regression models.

Stacking, or stacked generalization, uses a second-stage model to combine predictions from a number of models by training the second-stage model using out-of-samples predictions from the first-stage models (Wolpert, 1992). The modelling scheme proposed here has similarities to stacking methods used for malaria mapping (Bhatt et al., 2017) and elsewhere (Breiman, 1996b; Hao et al., 2019; Sill et al., 2009; Wolpert, 1992). However, as the response data in the machine learning models and the disaggregation regression models are on different scales (prevalence is a proportion while incidence is a rate) we cannot simply take a weighted average of the predictions from the machine learning models as in a standard stacking scheme (Hao et al., 2019; Sill et al., 2009). Instead the predictions need to be transformed to the incidence scale with a separately fitted model (Cameron et al., 2015). Applications in other disease contexts have used a similar stacking scheme where data from vector or wild-animal host species are used to train models, the predictions from which are then used as covariates in a final model (Pigott et al., 2014; Shearer et al., 2016). In such applications we would always expect to need additional covariates as well as the modelled distributions of hosts or vectors. However, in the case examined here, both sets of data are direct measures of some aspect of malaria transmission rate, and therefore it is possible, though not guaranteed, that we would not need any further covariates.

Model stacking (Wolpert, 1992) has proven effective in many realms (Bhatt et al., 2017; Breiman, 1996b; Hao et al., 2019; Sill et al., 2009). Stacking improves predictions by controlling bias and variance; as long as suitably diverse models are averaged, they will have different biases while high variance in models should be averaged out. This understanding of how stacking improves model performance indicates that diversity in models is important for stacking to be effective. Diversity in models is typically created in two ways: by using diverse training datasets (Breiman, 1996a) (as in Random Forests for example) and by using functionally different models (for example by averaging tree based models and neural networks) (Breiman, 1996b). One important trade-off in spatial modelling is whether to use local data (with a smaller sample size but that is likely to be representative of the area of study) or global data that have a larger sample size but a less close association with the areas of study. For the application of malaria mapping, we can think about diversity of training data in this context and expect that stacking separate models trained on local and global data will also increase the diversity of predictions in a useful way.

To test the effectiveness of the proposed approach we used data from four countries with relatively complete surveillance data: Madagascar, Colombia, Indonesia and Senegal. We focused our analysis on comparing the predictive performance of disaggregation regression when given different sets of covariates. Therefore we keep the structure of the disaggregation regression model the same and only vary the covariates provided to the model. In each country we fitted stage 1 machine learning models trained on prevalence data and raw environmental covariates. We made new covariates using predictions from these models. We then tested whether stage 2 disaggregation regression models with these new covariates performed better than a baseline disaggregation regression model that directly used the raw environmental covariates. We tested this approach using machine learning models trained on local prevalence data as well as models trained on a global prevalence dataset. While there was no consistently best model we found that, in most cases, the two stage method worked better than the single stage baseline disaggregation regression models. Using predictions from machine learning models, trained on local prevalence data as covariates improved the performance of disaggregation regression models relative to the disaggregation regression models that only used the raw environmental covariates.

regression models relative to the disaggregation regression models that only used the raw environmental covariates. In contrast, using predictions from machine learning models trained on the global prevalence dataset rarely improved predictive performance.

## 2. Methods

### 2.1. Epidemiological data

We used two data sources that reflect *P. falciparum* malaria transmission; point-prevalence surveys and polygon-level, aggregated incidence data. We selected Madagascar, Colombia, Indonesia and Senegal as case examples as they all have fairly complete, publicly available, surveillance data at a finer geographical resolution than administrative level one (i.e. higher resolution than state or province). The prevalence survey data were extracted from the Malaria Atlas Project prevalence survey database using only data from 1990 onwards (Bhatt et al., 2015; Guerra et al., 2007; Pfeffer et al., 2018). While the data covered a large time period, we did not model time explicitly as we are here focussed on spatial, rather than temporal modelling. Although we have not accounted for time in the models, as long as sampling in space is independent of time, the correct relationships should be recovered. The prevalence points were then standardised to an age range of 2–10 using the model from (Smith et al., 2007). This data was used as both a global dataset and as regional subsets. The global dataset contains 55,914 surveys in 44,842 distinct locations and represents samples from 5,687,304 individuals. As there were relatively few surveys in Colombia we used all points from South America (7,719 individuals from 522 locations) while for the other countries we used only data from that country (Madagascar: 89,381 individuals from 1505 locations. Indonesia: 1,512,888 individuals from 4778 locations. Senegal: 80,896 individuals from 1762 locations).

The polygon incidence data (i.e. malaria incidence aggregated to administrative units) were collected from government reports (Colombian National Institute of Health, 2016; Indonesia Ministry of Health, 2013; Rakotorahalahy, 2009; Senegal Ministry of Health, 2009) and standardised using methods defined in Cibulskis et al. (2011). This standardisation step accounts for missed cases due to lack of treatment seeking, missing case reports, and cases that sought medical attention outside the public health systems (Battle et al., 2016). For reports where cases were not reported at the species level, national estimates of the ratio between *P. falciparum* and *Plasmodium vivax* cases from the World Malaria Report were used to calculate *P. falciparum* only cases (World Health Organization, 2016). For incidence rates we divide by 1000 to give the Annual Parasite Index (API). To keep the analysis focused on spatial estimates we selected one year of surveillance data for each country. We used annual surveillance data from 2013 for Madagascar (110 districts), 2015 for Colombia (952 municipalities), 2013 for Indonesia (244 regencies and cities) and 2009 for Senegal (34 departments). These years were selected as they had the most complete data in each case.

Raster surfaces (i.e. population gridded to  $5 \times 5$  km pixels) of population for the years 2005, 2010 and 2015, were created using data from WorldPop (Gaughan et al., 2013; Linaud et al., 2012; Sorichetta et al., 2015) and from GPWv4 (NASA, 2018) where WorldPop did not have values. Population rasters for the remaining years were created by linear interpolation.

### 2.2. Raw environmental covariates

We considered a suite of environmental and anthropological covariates, at a resolution of approximately  $5 \times 5$  kilometres that included the annual mean and log standard deviation of land surface

temperature, enhanced vegetation index, malaria parasite temperature suitability index, elevation, tasseled cap wetness, log accessibility to cities and log night lights (Gething et al., 2011; Weiss et al., 2014; 2015; 2018). All covariates were aligned in their native resolution (500 m or 1 km) and then aggregated to 5 km resolution, therefore there should be minimal effects from spatial misalignment. The covariates were standardised and centered to have a mean of zero and a standard deviation of one. We refer to this set of transformed variables as the raw environmental covariates (even though some of the covariates are anthropogenic rather than environmental) to distinguish them from other covariates created from predictions from stage 1 machine learning models. The raw environmental variables were used as covariates in the stage 1 machine learning models as well as being used directly as covariates in the baseline stage 2 disaggregation regression models.

### 2.3. Stage 1 machine learning models

For each country specific dataset and for the global dataset we fitted 5 stage 1 models via *caret* (Kuhn et al., 2017): elastic net (Zou and Hastie, 2012), Random Forest (Wright and Ziegler, 2015), projection pursuit regression (Friedman and Stuetzle, 1981), neural networks (Venables and Ripley, 2002) and boosted regression trees (gradient boosted models, subsequently GBM) (Ridgeway, 2017). These models were fitted to both the full malaria prevalence dataset and to the regional subsets of the data. Our response variable was prevalence and we weighted the data by sample size (i.e. the number of people tested for malaria in each survey). We used the raw environmental covariates described above as covariates in these machine learning models. This process can therefore be seen as creating non-linear transformations of the raw covariates that are hopefully better correlated with malaria incidence than the raw environmental covariates are. For each model we ran five-fold cross-validation to select hyperparameters using random search for Random Forest and boosted regression trees and grid search for the other models. Root mean square error (RMSE) was used to select the best performing model. We note that spatial or random cross-validation could have been used in this step. The choice is less critical than for the cross-validation scheme used to test model performance but using random cross-validation might select for hyperparameters giving more complex or flexible models. To an extent, when ensembling models, high variance is better than high bias as the variance gets averaged out (Breiman, 1996b).

Predictions from these models were then made across each country respectively. These predictions were empirical logit transformed so that they were on the linear predictor scale of the disaggregation regression model. An empirical logit was used rather than a standard logit as there were many predictions of exactly zero. These predicted surfaces were subsequently used as covariates in the stage 2 disaggregation regression models (Fig. 1). Plots of the correlation matrices for all covariates can be seen in S46S49. The correlation between covariates varied from country to country. In general there is little correlation between global and local machine learning predictions. Accessibility is often strongly correlated with predictions from the machine learning models. In Senegal there is a lot of correlation between variables. However, we note that collinearity between variables is not as problematic in a predictive context as it is when interpreting regression coefficients is the aim of the analysis, though it is still potentially a waste of degrees of freedom. The experiments that follow in this paper assess the performance of disaggregation regression models when using the predictions from these machine learning models as covariates as compared to a baseline disaggregation regression model using only the raw environmental covariates. See the supplementary material for plots of the grid search hyperparameter

performance, out-of-sample scatter plots and plots of the predicted surfaces.

### 2.4. Stage 2 disaggregation regression

The model fitted to aggregated incidence data was a disaggregation regression model (Law et al., 2018; Li et al., 2012; Sturrock et al., 2014; Taylor et al., 2017; Wilson and Wakefield, 2018). The models were implemented and fitted using Template Model Builder (Kristensen et al., 2016) in R (R Core Team, 2018) and we note that these models cannot be fitted using INLA (Lindgren and Rue, 2015) as we are not using a linear link function. This model is defined by a likelihood at the level of the polygon with covariates and a spatial random field at the pixel-level. Values at the polygon-level are given the subscript  $a$  while pixel level values are indexed with  $b$ .

The aggregated incidence count data,  $y_a$  is given a Poisson likelihood

$$y_a \sim \text{Pois}(i_a \text{pop}_a)$$

where  $i_a$  is the estimated polygon incidence rate and  $\text{pop}_a$  is the observed polygon population-at-risk. This polygon-level likelihood is linked to the pixel-level incidence and prevalence by

$$i_a = \frac{\sum (i_b \text{pop}_b)}{\sum \text{pop}_b}$$

$$i_b = 2.616p_b - 3.596p_b^2 + 1.594p_b^3$$

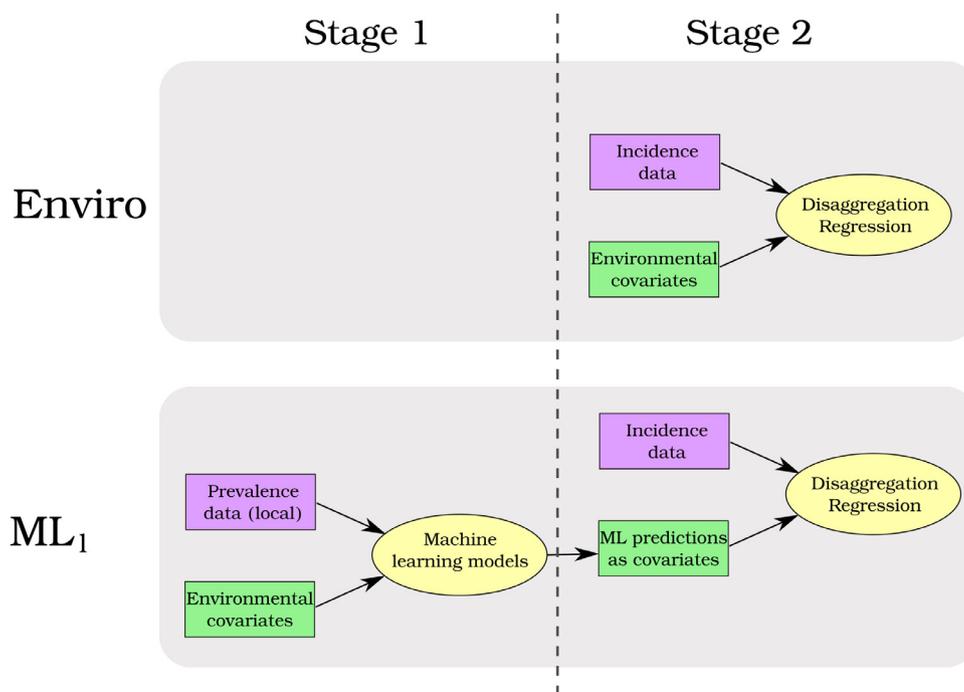
where the polynomial is a function from a previously published model (Cameron et al., 2015). The fact that the model is explicit about the relationship between prevalence and incidence has two advantages. Firstly, predictions of prevalence can be easily made directly from the linear predictor of the model. Secondly, it means that the logit-transformed predictions from the machine learning models are correctly scaled. The linear predictor of the model is related to prevalence by a typical logit link function and includes: an intercept,  $\beta_0$ ; covariates,  $X$ , with a vector of regression parameters,  $\beta$ ; a spatial, Gaussian, random field,  $u_b(\rho, \sigma_u)$ ; and an *iid* random effect,  $v_a(\sigma_v)$ .

$$p_b = \text{logit}^{-1}(\beta_0 + \beta X + u_b(\rho, \sigma_u) + v_a(\sigma_v))$$

The Gaussian spatial effect has a Matérn covariance function and two hyperparameters:  $\rho$ , the nominal range (beyond which correlation is  $< 0.1$ ) and  $\sigma_u$ , the marginal standard deviation. The *iid* random effect models both unobserved explanatory factors and extra-Poisson sampling error. As described in more detail below we do not vary the structure of this model in the methodological comparison. The matrix of covariates,  $X$ , is the only component that varies and is made up of various combinations of raw environmental covariates and predictions from the machine learning models.

Finally, we complete the model by setting priors on the parameters  $\beta_0$ ,  $\beta$ ,  $\rho$  and  $\sigma_u$  and  $\sigma_v$ . We assigned  $\rho$  and  $\sigma_u$  a joint penalised complexity prior (Fugstad et al., 2018) such that  $P(\rho < 1) = 0.0001$  (except for Indonesia where we set  $P(\rho < 3) = 0.0001$  due to its much larger size) and  $P(\sigma_u > 1) = 0.0001$ . This prior encoded our *a priori* preference for a simpler, smoother random field. We set this prior such that the random field could explain most of the range of the data if required.

We assigned  $\sigma_v$  a penalised complexity prior (Simpson et al., 2017) such that  $P(\sigma_v > 0.05) = 0.0001$ . This was based on a comparison of the variance of Poisson random variables, with rates given by the number of polygon-level cases observed, and an independently derived upper and lower bound for the case counts using the approach defined in (Cibulskis et al., 2011). We found that an *iid* effect with a standard deviation of 0.05 would be able



**Fig. 1.** Schematic of the baseline disaggregation regression model (Enviro) and the two stage method (ML<sub>1</sub>). Models are shown in yellow ovals, malaria data is shown in purple rectangles and covariates are shown in green rectangles. The baseline model (Enviro) uses aggregated incidence data and raw environmental covariates in a disaggregation regression model. In the two stage method (ML<sub>1</sub>), new covariates are created in stage 1 by training machine learning models on prevalence data. Predictions from these machine learning models are used as covariates in the stage 2 disaggregation regression. Only one of the two stage models (ML<sub>1</sub>) is shown for simplicity. If ML<sub>g</sub> was included as well for example, it would look the same as ML<sub>1</sub> except that the prevalence data (pink box in stage 1) would have the global database of prevalence surveys. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to account for the discrepancy between the assumed Poisson error and the independently derived error.

Finally, we set different priors on the regression coefficients depending on which covariates were used. When raw environmental covariates or a mix of raw environmental covariates and predictions from machine learning models were used we set the prior to be weakly regularising,  $\beta_i \sim \text{Norm}(0, 0.4)$ , such that it was unlikely that any single covariate could explain the full range of the response data. When only machine learning model predictions were used we set  $\beta_i \sim \text{Norm}(\frac{1}{M}, 0.4)$  where  $M$  is the number of machine learning models being used. This prior sets our *a priori* expectation that all the machine learning prediction models are positively and equally correlated with incidence i.e. this prior encodes standard model averaging. It is important to note that this setup does not constitute true stacking in which we would enforce  $\sum_i \beta_i = 1$  (Bhatt et al., 2017). In a preliminary analysis we tested the case where we force  $\beta_i > 0$  which in practice is largely the same as the  $\sum_i \beta_i = 1$  (Breiman, 1996b) but allows a small amount of flexibility to handle misspecification in p2i. This analysis did not show any benefits to this approach so it was not considered further.

## 2.5. Experiments

All experiments involve comparing predictive performance of stage 2 disaggregation regression models when different combinations of covariates are used. After model fitting we made predictions over the study areas and reaggregated back to the administrative level of the surveillance data. Our primary performance metric was correlation between observed and aggregated predictions. We also examined the calibration of models by calculating the proportion of held out data that were within their 80% credible intervals. We used the raw environmental covariates, centered and standardised, as a baseline stage 2 model (Fig. 1). We subsequently refer to this model as Enviro. We then performed two

experiments. In the first (experiment 1) we tested whether using predictions from stage 1 machine learning models trained on local (i.e. within-country) prevalence data improved predictions (Fig. 1). In the second (experiment 2) we tested whether using predictions from stage 1 machine learning models trained on global prevalence data improved predictions.

In experiment 1 we compared the baseline to two models that used locally trained machine learning models (Table 1). Firstly, we combined the predictions from the machine learning models and the environmental covariates in one model (subsequently called Enviro + ML<sub>1</sub>). In this model therefore, the environmental covariates are effectively used twice, once in their raw form and once in transformed form (i.e. the predictions from the machine learning models). Secondly, we used only the predictions from the machine learning models (Fig. 1) trained on local data (subsequently called ML<sub>1</sub>). As the environmental covariates are used as covariates in the machine learning models, this model is still ultimately driven by the values of the raw environmental covariates. However, in this model we are only using the environmental covariates in the transformed space learned by the machine learning models.

In experiment 2 we compared the baseline model to two models that used predictions from machine learning models trained on the global dataset of prevalence surveys (Table 1). In the first, we used only predictions from the machine learning models trained on the global data (ML<sub>g</sub>). In the second, we combined predictions from the machine learning models trained on regional data and predictions from the machine learning models trained on global data (ML<sub>1</sub> + ML<sub>g</sub>).

In each experiment we used two cross-validation schemes. In the first, polygon incidence data was randomly split into six cross-validation folds. In the second, polygon incidence data was split spatially into  $k$  folds (via  $k$ -means clustering on the polygon centroids). We set  $k$  as 3 for Madagascar and Colombia. Due to its large size we set  $k$  as 7 for Indonesia. Due to the small sample

**Table 1**

Summary of stage 2 models and the experiments they are grouped into. All models are disaggregation regression models fitted to aggregated incidence data. The only difference between the models is which covariates are being used. Environmental covariates includes the full set of eight variables. Local ML covariates are the predictions from stage 1 machine learning models, trained with local prevalence data and the eight environmental covariates as inputs. The local prevalence datasets are data from within each country except for Colombia where local prevalence refers to South American data. Global ML covariates are the predictions from stage 1 machine learning models, trained with global prevalence data (i.e. the full prevalence database) and the eight environmental covariates as inputs. Experiment 1 tests whether including Local ML covariates improves predictive performance while Experiment 2 tests whether including Global ML covariates improves performance.

Model	Environmental covariates	Local ML covariates	Global ML covariates
<b>Experiment 1</b>			
Enviro	✓		
Enviro + ML <sub>l</sub>	✓	✓	
ML <sub>l</sub>		✓	
<b>Experiment 2</b>			
Enviro	✓		
ML <sub>g</sub>			✓
ML <sub>l</sub> + ML <sub>g</sub>		✓	✓

**Table 2**

Pearson correlations between observed and predicted values for experiment 1.

Country	Cross-validation	Enviro	Enviro + ML <sub>l</sub>	ML <sub>l</sub>
Colombia	Random	0.59	<b>0.61</b>	0.59
Colombia	Spatial	0.12	0.25	<b>0.33</b>
Indonesia	Random	0.52	<b>0.59</b>	0.48
Indonesia	Spatial	0.45	<b>0.51</b>	0.44
Madagascar	Random	<b>0.70</b>	0.69	0.68
Madagascar	Spatial	0.22	0.18	<b>0.55</b>
Senegal	Random	<b>0.58</b>	0.57	0.51
Senegal	Spatial	0.63	0.63	0.51

**Table 3**

Pearson correlations between observed and predicted values for experiment 2.

Country	Cross-validation	Enviro	ML <sub>g</sub>	ML <sub>l</sub> + ML <sub>g</sub>
Colombia	Random	<b>0.59</b>	0.55	0.58
Colombia	Spatial	0.12	0.12	<b>0.33</b>
Indonesia	Random	<b>0.52</b>	0.32	0.46
Indonesia	Spatial	0.45	0.41	0.45
Madagascar	Random	<b>0.70</b>	0.67	0.68
Madagascar	Spatial	0.22	0.51	<b>0.55</b>
Senegal	Random	<b>0.58</b>	0.50	0.49
Senegal	Spatial	<b>0.63</b>	0.55	0.52

size, we set k as 5 for Senegal. This spatial cross-validation scheme is testing the ability of the different models to make predictions far from data where the spatial random field is not informative.

### 3. Results

Table 2 gives the correlation between observed and held out data (under random and spatial cross-validation) for experiment 1 (models Enviro, ML<sub>l</sub> and Enviro + ML<sub>l</sub>). Many of the differences in performance are rather marginal. Enviro was the best performing model in two cases (random cross-validation in Madagascar and Senegal). In one case, Enviro performed equally well as another model; under spatial cross-validation in Senegal Enviro and Enviro + ML<sub>l</sub> performed equally well. Of the remaining five cases, in two cases ML<sub>l</sub> performed best and in three cases Enviro + ML<sub>l</sub> performed best. The greatest benefits to using prediction from machine learning models instead of, or in combination with, the raw environmental variables occurred under spatial cross-validation and in the cases when the Enviro model was particularly poor.

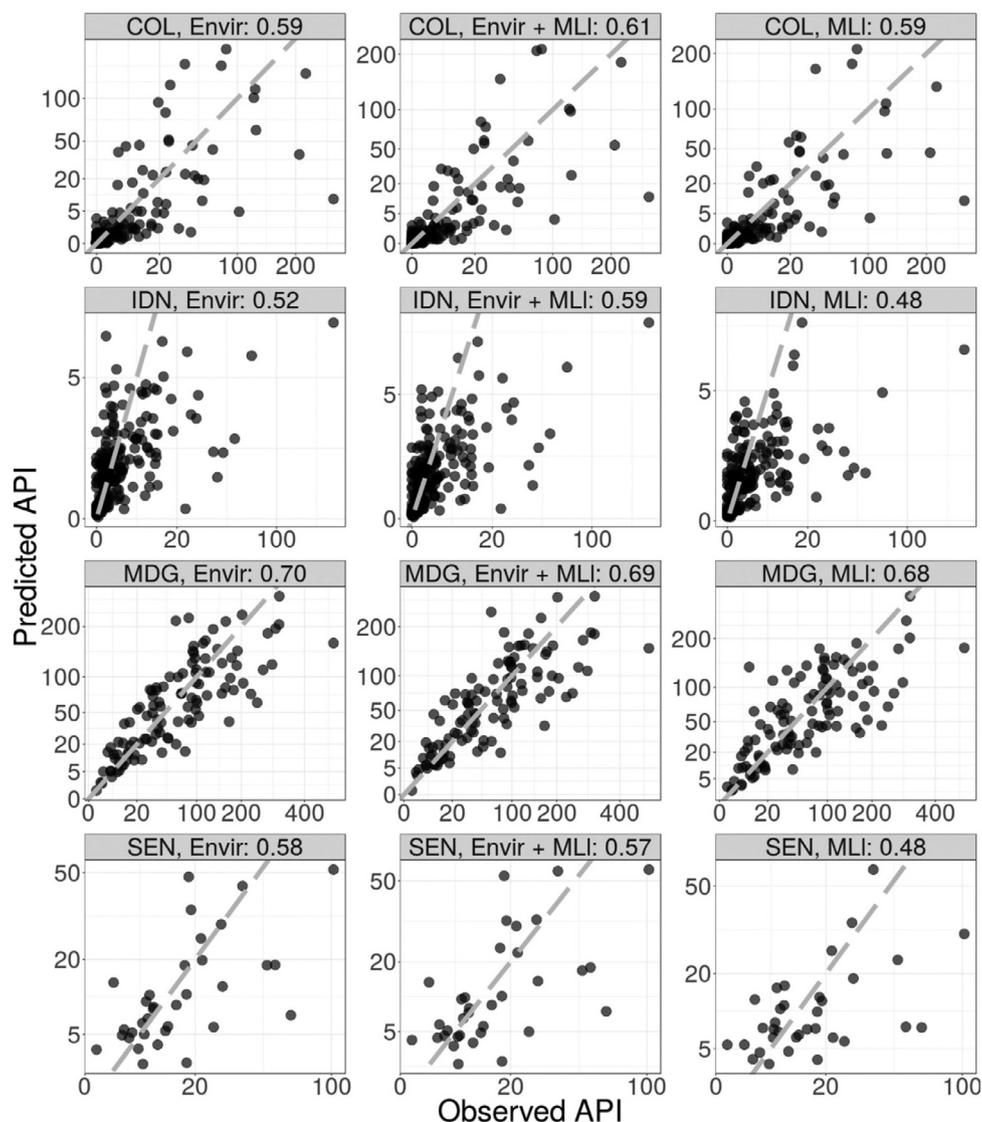
Fig. 2 shows scatter plots of the model performance under random cross-validation for experiment 1 while Fig. 3 shows scatter plots of the model performance under spatial cross-validation. It can be seen that without environmental covariates, the models in Madagascar fail to predict very high or very low values correctly. Fig. 4 shows the input data and spatial out-of-sample predictions of the Enviro model and ML<sub>l</sub> model in Colombia.

Table 3 gives the correlation between observed and held out data (under random and spatial cross-validation) for experiment 2 (models Enviro, ML<sub>g</sub> and ML<sub>l</sub> + ML<sub>g</sub>). In six cases, Enviro was the best or tied best performing model. The ML<sub>g</sub> model was never the best performing model and only outperforms Enviro in one case (spatial cross-validation in Madagascar). In two cases ML<sub>l</sub> +

ML<sub>g</sub> was the best performing model (spatial Colombia and spatial Madagascar). Comparing across Tables 2 and 3 we can see that ML<sub>g</sub> outperforms ML<sub>l</sub> once (spatial cross-validation in Senegal). In only two cases (Spatial Senegal and Spatial Indonesia) did ML<sub>l</sub> + ML<sub>g</sub> outperform ML<sub>l</sub>.

Table 4 shows the out-of-sample coverage of the 80% credible intervals for all models, countries and cross-validation schemes. The coverage in Colombia was very poor with no models achieving a coverage above 0.4. The coverage was acceptable in the other three countries with most values lying between 0.7 and 0.9. Overall there was a general tendency for models to be slightly over confident.

We can examine the relationship between the RMSE of the machine learning models to their fitted regression coefficients (weights). These values are given in Table 5. In all five sets of machine learning models (four sets trained on local data and one set trained on global data), Random Forest performs the best. We have not forced  $\beta_i > 0$  but we have set the priors for these coefficients with a positive mean. In nearly all cases the fitted values are positive. If the prevalence data and incidence data are not biased relative to each other we would expect the models with the lowest RMSE to also have the biggest regression coefficient. This occurs in three cases where Random Forest has the lowest RMSE and the biggest coefficient. In a further two cases, Random Forest has the lowest RMSE but GBM has the highest coefficient. The predictions from Random Forest and GBM are highly correlated in part because they both perform well and in part because they are both tree based models. Finally, it can be seen that the relationship between RMSE and regression coefficients was much weaker in Indonesia. For the models trained on local data a neural network has the highest fitted coefficient while for the models trained on global data an elastic net has the highest fitted coefficient.



**Fig. 2.** Observed data against predictions for random cross-validation hold-out samples on a square root transformed scale. There are 12 cases composed of 4 countries (COL.: Colombia, IDN: Indonesia, MDG: Madagascar, SEN: Senegal) and three sets of covariates (Envir: raw environmental covariates only, Enviro + ML<sub>l</sub>: raw environmental covariates and machine learning covariates trained on local prevalence data combined, ML<sub>l</sub>: Machine learning models trained on local prevalence data only).

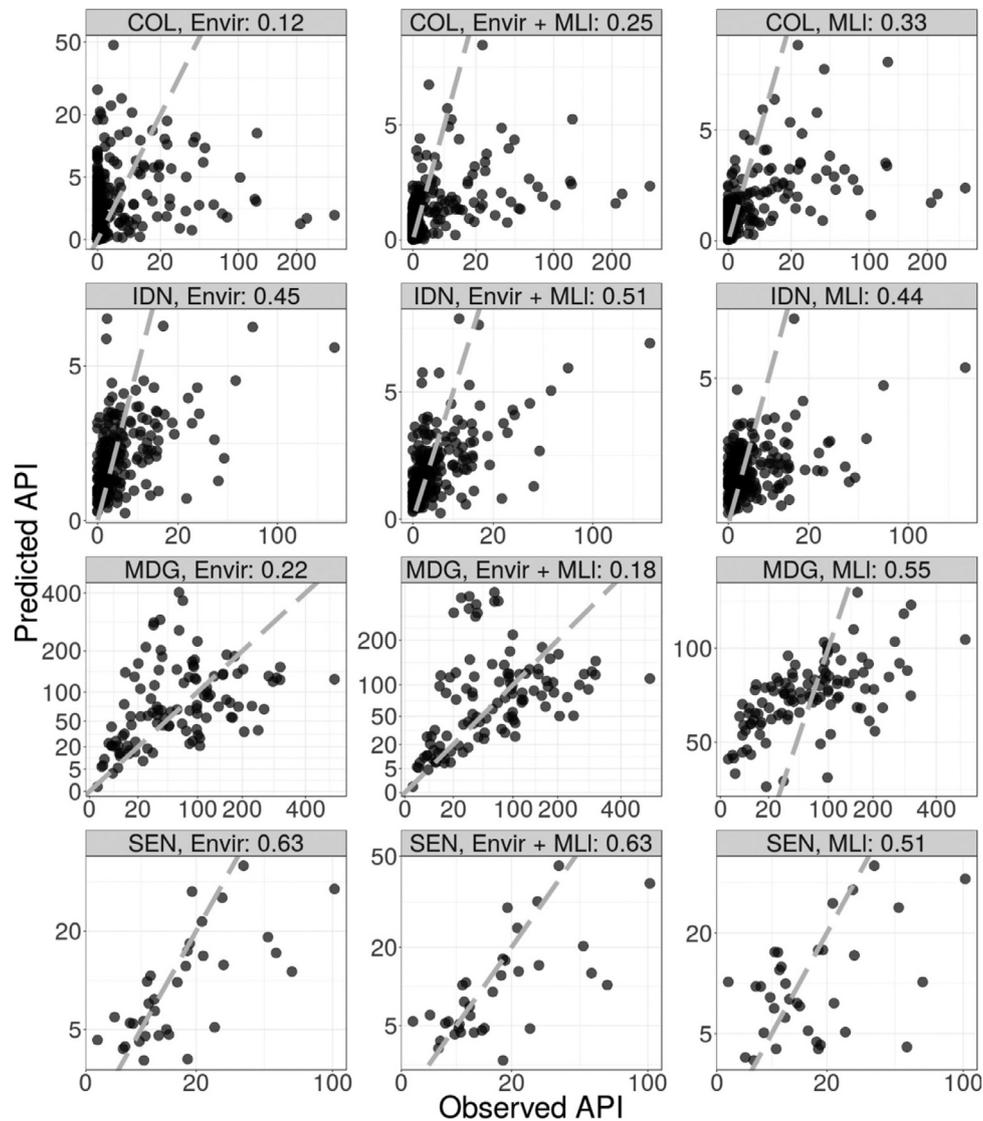
**Table 4**  
Coverage of 80% credible intervals. Values outside 0.7–0.9 are shown in bold.

Country	CV	Enviro	ML <sub>l</sub>	Enviro + ML <sub>l</sub>	ML <sub>g</sub>	ML <sub>l</sub> + ML <sub>g</sub>
Colombia	Random	<b>0.28</b>	<b>0.28</b>	<b>0.29</b>	<b>0.28</b>	<b>0.30</b>
Colombia	Spatial	<b>0.30</b>	<b>0.33</b>	<b>0.33</b>	<b>0.33</b>	<b>0.34</b>
Indonesia	Random	0.80	0.81	0.78	0.79	0.77
Indonesia	Spatial	0.80	0.78	0.78	0.76	0.75
Madagascar	Random	0.80	0.77	0.75	0.77	0.76
Madagascar	Spatial	<b>0.65</b>	0.74	0.70	0.75	0.76
Senegal	Random	0.79	0.79	0.79	0.79	0.82
Senegal	Spatial	0.85	<b>0.91</b>	0.85	<b>0.94</b>	0.85

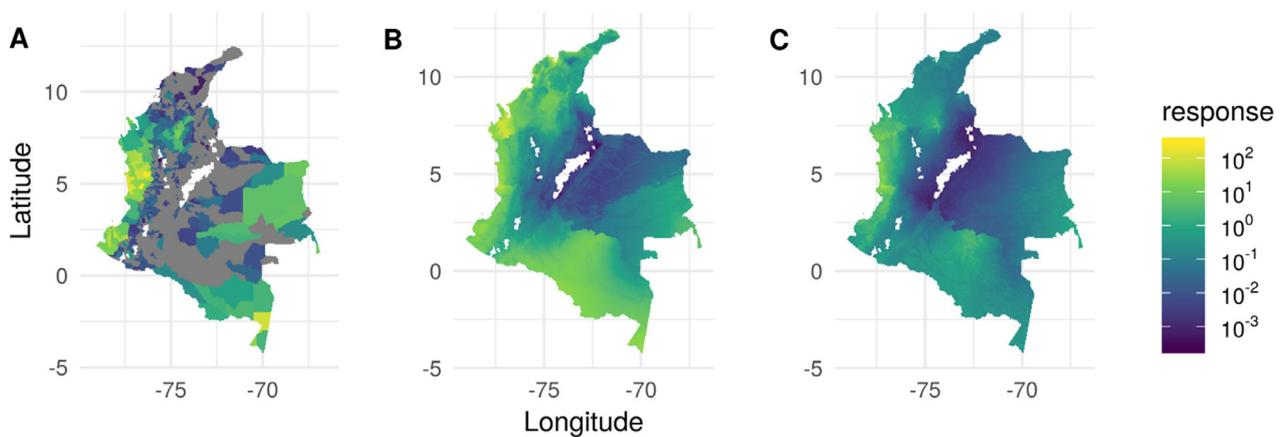
#### 4. Discussion

We have studied the predictive performance of disaggregation regression of malaria incidence when provided with different sets of covariates. In experiment 1 we compared a baseline model that used only raw environmental covariates (Enviro) to two models that used predictions from machine learning models trained on local prevalence points (ML<sub>l</sub> and Enviro + ML<sub>l</sub>). Overall, experiment 1 suggests that the predictions from the disaggregation mod-

els were better when using covariates created using predictions from machine learning models (trained on local prevalence points) than when using raw environmental covariates. This increased performance comes despite the prevalence data being on a different scale (a proportion instead of a rate) and being measurements of a different aspects of malaria transmission (prevalence rather than incidence) as well as the fact that the model we have used to translate between the two scales is imperfect. However, there was no clear best model between ML<sub>l</sub> and Enviro + ML<sub>l</sub>. Therefore,



**Fig. 3.** Observed data against predictions for spatial cross-validation hold-out samples on a square root transformed scale. There are 12 cases composed of 4 countries (COL.: Colombia, IDN: Indonesia, MDG: Madagascar, SEN: Senegal) and three sets of covariates (Envir: raw environmental covariates only, Envir + MLI: raw environmental covariates and machine learning covariates trained on local prevalence data combined, MLI: Machine learning models trained on local prevalence data only).



**Fig. 4.** A) Observed data for Colombia (grey for zero incidence). B) Out-of-sample predictions for the spatial cross-validation, environmental covariates only model. C) Out-of-sample predictions for the spatial cross-validation, local machine learning only model. For each cross-validation fold, predictions are made for the held out data which are then combined to make a single surface.

**Table 5**

Machine learning model results and fitted parameters (i.e. model weights) of the machine learning predictions only models (i.e. ML<sub>l</sub> local only and ML<sub>g</sub> global only). For each dataset (the country and whether the data was local or global) the best model (lowest RMSE) is shown in bold. Similarly, the largest coefficient within each disaggregation model is shown in bold.

Country	Model	RMSE <sub>l</sub>	$\beta_l$	RMSE <sub>g</sub>	$\beta_g$
Colombia	RF	<b>0.068</b>	0.625	<b>0.169</b>	0.180
Colombia	GBM	0.073	<b>0.952</b>	0.178	-0.218
Colombia	enet	0.070	0.219	0.233	0.183
Colombia	nnet	0.070	0.129	0.220	0.527
Colombia	ppr	0.070	0.667	0.205	<b>0.546</b>
Indonesia	RF	<b>0.081</b>	0.447	<b>0.169</b>	0.178
Indonesia	GBM	0.085	0.357	0.178	0.289
Indonesia	enet	0.091	0.303	0.233	<b>0.526</b>
Indonesia	nnet	0.089	<b>0.506</b>	0.220	0.316
Indonesia	ppr	0.089	0.364	0.205	0.089
Madagascar	RF	<b>0.100</b>	0.538	<b>0.169</b>	<b>0.529</b>
Madagascar	GBM	0.105	<b>0.570</b>	0.178	0.432
Madagascar	enet	0.116	0.301	0.233	0.262
Madagascar	nnet	0.113	0.033	0.220	0.364
Madagascar	ppr	0.109	0.469	0.205	0.403
Senegal	RF	<b>0.092</b>	<b>0.339</b>	<b>0.169</b>	<b>0.425</b>
Senegal	GBM	0.099	0.261	0.178	0.408
Senegal	enet	0.103	0.344	0.233	0.205
Senegal	nnet	0.099	0.254	0.220	0.190
Senegal	ppr	0.098	0.268	0.205	0.126

when using these methods, both of these models should be fitted and the model with the best predictive performance for a given dataset selected.

Furthermore, many of the performance improvements are rather marginal. However, in a few cases such as Colombia and Madagascar under spatial cross-validation the performance boost is large. For example, using just environmental covariates in Madagascar under spatial cross-validation gives a correlation between observed and predicted data of 0.22. Such a model would be unusable for any applied or policy work. In contrast, using the predictions from the machine learning models trained on local data gives a correlation value of 0.55, which while still relatively poor is possibly a useful model.

It is also of note that in two cases, the simplest model (Enviro) performed the best. Given that Enviro + ML<sub>l</sub> includes the same covariates (with some extra) this is likely explained by increased variance. The model is trying to estimate a number of regression parameters, a random field and hyperparameters so removing extraneous covariates should help parameter estimation. This is particularly true in Senegal which only has 34 datapoints and higher correlation between covariates (S49). Future work should consider joint models which use the prevalence and incidence data as response data so that the disaggregation has more degrees of freedom. This would hopefully allow the accurate estimation of more regression parameters. The prevalence to incidence relationship used in this paper could be used to link the two data types. Furthermore, if the parameters in the prevalence incidence relationship were treated as unknown parameters, with informative priors, this would allow three data to inform the relationship and for the uncertainty in the original model to be propagated properly.

As expected, the model performance was generally worse under spatial cross-validation than under random cross-validation. This implies that the models are still relying heavily on the spatial Gaussian random field. Furthermore, the difference between random and spatial cross-validation is usually bigger than the difference between different models. This suggests that better data coverage is more important than which specific model is used and that the models are still relying quite strongly on the random field.

Overall, the models fitted using predictions from machine learning models trained on the global database of prevalence point-surveys were worse than those using either environmental data alone or than those using predictions from machine learning models trained on local data. This in itself is not particularly surprising. However, it does indicate that continental scale models should consider using a mosaic of locally trained machine learning models for example. What is more surprising is that using both the ML<sub>l</sub> and ML<sub>g</sub> covariates together did not improve performance, especially considering the relatively low correlation between the local and global predictions (Figure S46–49). Given that this approach of creating diverse predictions is a core element of stacking methodology we would expect this set of covariates to perform as well or better than just the ML<sub>l</sub> covariates.

We have demonstrated that using prevalence data alongside aggregated incidence data can improve predictive performance. Unlike the maps presented by Weiss et al. (2019), the estimates presented in this paper are intended for methodological research, not directly for policy decisions. However, they do further demonstrate that aggregated incidence data can be used to create high resolution estimates of malaria incidence. Furthermore, in countries such as Senegal, the incidence data is much coarser than intervention implementation units, so disaggregated maps like the ones here have clear policy uses. It is important, however, to note that we are using out-of-sample prediction of aggregated polygon data as our performance metric while the true target for prediction is the high resolution risk surface. In rare cases there is both aggregate and unaggregated data available so that the high resolution accuracy can be tested as well as the aggregate level performance (Sturrock et al., 2014). However, this case is rare. Some simulation studies have been performed (Johnson et al., 2019; Law et al., 2018; Wilson and Wakefield, 2018) but these are often with few covariates and on correctly specified models. This is an important area for future research.

While the approach presented here is related to stacking, it differs in that we have not constrained the regression parameters to be positive nor included a sum-to-one constraint, i.e. the result is not simply a weighted average of the level zero model predictions. We did not include these constraints because the first stage and second stage models were trained on response data on different scales. However, given our priors, nearly all the fitted regression coefficients in models with only machine learning predictions were positive. Therefore, in practice these models are working in a way similar to standard stacking. However, the coefficients certainly do not sum to one, and fitted intercepts are negative to account for this.

One drawback of using predictions from machine learning models as covariates is that the uncertainties in the predictions are not propagated properly through to the final predictions. This could be handled with an appropriate error model (Richardson and Gilks, 1993). Where the machine learning models explicitly provide estimates of uncertainty, these could be used to inform such models. In the absence of such individual-prediction estimates of uncertainty, the cross-validation error could instead be used to inform priors for error models. It is however worth noting that the environmental covariates used here are also modelled and therefore similar care would ideally be taken in characterising the uncertainty in their values.

## 5. Conclusions

Overall we find that including predictions from machine learning models trained on prevalence point-surveys can improve the performance of disaggregation regression models for malaria incidence relative to using raw environmental covariates. This extra modelling step can be seen as finding useful, non-linear transfor-

mations of the raw environmental covariates. This view is important for understanding how the model will predict in areas with no incidence data; in this situation the data cannot inform the Gaussian random field and predictions are driven by this non-linear transformation of the raw environmental covariates. Training the machine learning models on local data (i.e. prevalence data from the same country or region as the incidence data) shows much better performance than when training the machine learning models on the full, global, dataset of prevalence point-surveys. More countries, particularly those with medium or low malaria burdens, are providing timely and accurate routine surveillance data of malaria cases. Therefore it may be expected that disaggregation regression may become more popular and operationally relevant. Indeed, methods similar to  $ML_g$  are already being used for global mapping in malaria (Battle et al., 2019; Weiss et al., 2019) though the results here suggest that future global mapping efforts should allow local prevalence data to have a stronger influence on the estimates. We have here presented a method for improving the predictive performance of these models by using ancillary, prevalence point-survey data.

## Acknowledgments

The authors acknowledge the National Malaria Control Programme of Madagascar for sharing their routine case data for this analysis. This work was funded by the Bill and Melinda Gates Foundation.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.sste.2020.100357.

## References

- Battle, K.E., Bisanzio, D., Gibson, H.S., Bhatt, S., Cameron, E., Weiss, D.J., Mappin, B., Dalrymple, U., Howes, R.E., Hay, S.I., et al., 2016. Treatment-seeking rates in malaria endemic countries. *Malar. J.* 15 (1), 20.
- Battle, K.E., Lucas, T.C.D., Nguyen, M., Howes, R.E., Nandi, A.K., Twhog, K.A., Pfeiffer, D.A., Cameron, E., Rao, P.C., Casey, D., et al., 2019. Mapping the global endemicity and clinical burden of *Plasmodium vivax*, 2000–17: a spatial and temporal modelling study. *Lancet* 394.
- Bhatt, S., Cameron, E., Flaxman, S.R., Weiss, D.J., Smith, D.L., Gething, P.W., 2017. Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *J. R. Soc. Interface* 14 (134), 20170520.
- Bhatt, S., Weiss, D., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K., Moyes, C., Henry, A., Eckhoff, P., et al., 2015. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature* 526 (7572), 207.
- Breiman, L., 1996. Bagging predictors. *Mach. Learn.* 24 (2), 123–140.
- Breiman, L., 1996. Stacked regressions. *Mach. Learn.* 24 (1), 49–64.
- Cameron, E., Battle, K.E., Bhatt, S., Weiss, D.J., Bisanzio, D., Mappin, B., Dalrymple, U., Hay, S.I., Smith, D.L., Griffin, J.T., et al., 2015. Defining the relationship between infection prevalence and clinical incidence of *Plasmodium falciparum* malaria. *Nat. Commun.* 6.
- Cibulskis, R.E., Aregawi, M., Williams, R., Otten, M., Dye, C., 2011. Worldwide incidence of malaria in 2009: estimates, time trends, and a critique of methods. *PLoS Med.* 8 (12), e1001142.
- Colombian National Institute of Health, 2016. Sistema Nacional de Vigilancia en Salud Pública. online data resource.
- Friedman, J.H., Stuetzle, W., 1981. Projection pursuit regression. *J. Am. Stat. Assoc.* 76 (376), 817–823.
- Fuglstad, G.-A., Simpson, D., Lindgren, F., Rue, H., 2018. Constructing priors that penalize the complexity of Gaussian random fields. *J. Am. Stat. Assoc.* 1–8.
- Gaughan, A.E., Stevens, F.R., Linard, C., Jia, P., Tatem, A.J., 2013. High resolution population distribution maps for Southeast Asia in 2010 and 2015. *PLoS ONE* 8 (2), e55882.
- Gething, P.W., Van Boeckel, T.P., Smith, D.L., Guerra, C.A., Patil, A.P., Snow, R.W., Hay, S.I., 2011. Modelling the global constraints of temperature on transmission of *Plasmodium falciparum* and *P. vivax*. *Parasites Vectors* 4 (1), 92.
- Guerra, C.A., Hay, S.I., Lucio-Parades, L.S., Gikandi, P.W., Tatem, A.J., Noor, A.M., Snow, R.W., 2007. Assembling a global database of malaria parasite prevalence for the malaria atlas project. *Malar. J.* 6 (1), 17.
- Hao, T., Elith, J., Guillera-Aroita, G., Lahoz-Monfort, J.J., 2019. A review of evidence about use and performance of species distribution modelling ensembles like biomod. *Divers. Distrib.* 25 (5), 839–852.
- Indonesia Ministry of Health, 2013. State health reports.
- Johnson, O., Diggle, P., Giorgi, E., 2019. A spatially discrete approximation to log-Gaussian Cox processes for modelling aggregated disease count data. *Stat. Med.* 38 (24), 4871–4887. doi:10.1002/sim.8339.
- Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., Bell, B.M., 2016. TMB: Automatic differentiation and Laplace approximation. *J. Stat. Softw.* 70 (5), 1–21. doi:10.18637/jss.v070.i05.
- Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., the R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., Hunt, T., 2017. caret: Classification and regression training. R package version 6.0-76.
- Law, H.C., Sejdinovic, D., Cameron, E., Lucas, T.C., Flaxman, S., Battle, K., Fukumizu, K., 2018. Variational learning on aggregate outputs with gaussian processes. *Advances in Neural Information Processing Systems*, pp. 6081–6091.
- Li, Y., Brown, P., Gesink, D.C., Rue, H., 2012. Log Gaussian Cox processes and spatially aggregated disease incidence data. *Stat. Methods Med. Res.* 21 (5), 479–507.
- Linard, C., Gilbert, M., Snow, R.W., Noor, A.M., Tatem, A.J., 2012. Population distribution, settlement patterns and accessibility across Africa in 2010. *PLoS ONE* 7 (2), e31743.
- Lindgren, F., Rue, H., 2015. Bayesian spatial modelling with R-INLA. *J. Stat. Softw.* 63 (19), 1–25.
- NASA, 2018. Gridded population of the world (GPW), v4.
- Pfeiffer, D.A., Lucas, T.C.D., May, D., Harris, J., Rozier, J., Twhog, K.A., Dalrymple, U., Guerra, C.A., Moyes, C.L., Thorn, M., et al., 2018. malariaAtlas: an R interface to global malariometric data hosted by the Malaria Atlas Project. *Malar. J.* 17 (1), 352.
- Piggott, D.M., Golding, N., Mylne, A., Huang, Z., Henry, A.J., Weiss, D.J., Brady, O.J., Kraemer, M.U., Smith, D.L., Moyes, C.L., et al., 2014. Mapping the zoonotic niche of Ebola virus disease in Africa. *Elife* 3, e04395.
- R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.
- Rakotrahahaly, A. J., 2009. NMCP Madagascar monthly case data 2013–2014. Pers. Comms. Programme National de Lutte contre le Paludisme.
- Richardson, S., Gilks, W.R., 1993. Conditional independence models for epidemiological studies with covariate measurement error. *Stat. Med.* 12 (18), 1703–1722.
- Ridgeway et al., 2017. gbm: Generalized boosted regression models. R package version 2.1.3.
- Senegal Ministry of Health, 2009. Senegal Statistical Yearbook 2009.
- Shearer, F.M., Huang, Z., Weiss, D.J., Wiebe, A., Gibson, H.S., Battle, K.E., Piggott, D.M., Brady, O.J., Putaporntip, C., Jongwutiwes, S., et al., 2016. Estimating geographical variation in the risk of zoonotic *Plasmodium knowlesi* infection in countries eliminating malaria. *PLoS Negl. Trop. Dis.* 10 (8), e0004915.
- Sill, J., Takács, G., Mackey, L., Lin, D., 2009. Feature-weighted linear stacking. arXiv:0911.0460.
- Simpson, D., Rue, H., Riebler, A., Martins, T.G., Sørbye, S.H., et al., 2017. Penalising model component complexity: a principled, practical approach to constructing priors. *Stat. Sci.* 32 (1), 1–28.
- Smith, D.L., Guerra, C.A., Snow, R.W., Hay, S.I., 2007. Standardizing estimates of the *Plasmodium falciparum* parasite rate. *Malar. J.* 6 (1), 131.
- Sorichetta, A., Hornby, G.M., Stevens, F.R., Gaughan, A.E., Linard, C., Tatem, A.J., 2015. High-resolution gridded population datasets for Latin America and the Caribbean in 2010, 2015, and 2020. *Sci. Data* 2, 150045.
- Sturrock, H.J., Bennett, A.F., Midekisa, A., Gosling, R.D., Gething, P.W., Greenhouse, B., 2016. Mapping malaria risk in low transmission settings: challenges and opportunities. *Trends Parasitol.* 32 (8), 635–645.
- Sturrock, H.J., Cohen, J.M., Keil, P., Tatem, A.J., Le Menach, A., Ntshalintshali, N.E., Hsiang, M.S., Gosling, R.D., 2014. Fine-scale malaria risk mapping from routine aggregated case data. *Malar. J.* 13 (1), 421.
- Taylor, B.M., Andrade-Pacheco, R., Sturrock, H.J., 2017. Continuous inference for aggregated point process data. *J. R. Stat. Soc. Ser. A (Statistics in Society)*.
- Venables, W.N., Ripley, B.D., 2002. *Modern Applied Statistics with S*, fourth ed. Springer, New York. ISBN 0-387-95457-0
- Wang, C., Puhan, M.A., Furrer, R., Group, S.S., et al., 2018. Generalized spatial fusion model framework for joint analysis of point and areal data. *Spat. Stat.* 23, 72–90.
- Weiss, D.J., Bhatt, S., Mappin, B., Van Boeckel, T.P., Smith, D.L., Hay, S.I., Gething, P.W., 2014. Air temperature suitability for *Plasmodium falciparum* malaria transmission in Africa 2000–2012: a high-resolution spatiotemporal prediction. *Malar. J.* 13 (1), 171.
- Weiss, D.J., Lucas, T.C.D., Nguyen, M., Nandi, A.K., Bisanzio, D., Battle, K.E., Cameron, E., Twhog, K.A., Pfeiffer, D.A., Rozier, J.A., et al., 2019. Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum*, 2000–17: a spatial and temporal modelling study. *Lancet* 394.
- Weiss, D.J., Mappin, B., Dalrymple, U., Bhatt, S., Cameron, E., Hay, S.I., Gething, P.W., 2015. Re-examining environmental correlates of *Plasmodium falciparum* malaria endemicity: a data-intensive variable selection approach. *Malar. J.* 14 (1), 68.
- Weiss, D.J., Nelson, A., Gibson, H.S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., et al., 2018. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*.
- Wilson, K., Wakefield, J., 2018. Pointless spatial modeling. *Biostatistics*.
- Wolpert, D.H., 1992. Stacked generalization. *Neural Netw.* 5 (2), 241–259.
- World Health Organization, 2016. World Malaria Report 2016. World Health Organization.
- Wright, M.N., Ziegler, A., 2015. Ranger: a fast implementation of random forests for high dimensional data in C++ and R. arXiv preprint arXiv:1508.04409.
- Zou, H., Hastie, T., 2012. elasticnet: Elastic-Net for sparse estimation and sparse PCA. R package version 1.1.

**Update**

**Spatial and Spatio-temporal Epidemiology**

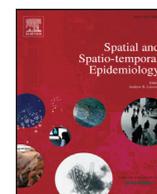
Volume 36, Issue , February 2021, Page

DOI: <https://doi.org/10.1016/j.sste.2020.100399>



Contents lists available at ScienceDirect

# Spatial and Spatio-temporal Epidemiology

journal homepage: [www.elsevier.com/locate/sste](http://www.elsevier.com/locate/sste)

## Erratum regarding missing Declaration of Competing Interest statements in previously published articles



Declaration of Competing Interest statements were not included in the published version of the following articles that appeared in previous issues of Spatial and Spatio-temporal Epidemiology

The appropriate Declaration/Competing Interest statements, provided by the Authors, are included below.

1. Examining the role of a retail density ordinance in reducing concentration of tobacco retailers (Spatial and Spatio-temporal Epidemiology, 2019; 32C) <https://doi.org/10.1016/j.sste.2019.100307> The authors were contacted after publication to request a Declaration of Interest statement.
2. Improving disaggregation models of malaria incidence by ensembling non-linear models of prevalence (Spatial and Spatio-temporal Epidemiology, 2020; <https://doi.org/10.1016/j.sste.2020.100357>) The authors were contacted after publication to request a Declaration of Interest statement.
3. Measuring neighbourhood social dimensions using individual responses: An application of multilevel factor analysis and econometrics (Spatial and Spatio-temporal Epidemiology, 2019; 32C) <https://doi.org/10.1016/j.sste.2019.100318> The authors were contacted after publication to request a Declaration of Interest statement.
4. A Bayesian spatio-temporal analysis on racial disparities in hypertensive disorders of pregnancy in Florida, 2005–2014 (Spatial and Spatio-temporal Epidemiology, 2019; 29C 43–50) <https://doi.org/10.1016/j.sste.2019.03.002> The authors were contacted after publication to request a Declaration of Interest statement.
5. Quantifying geographic regions of excess stillbirth risk in the presence of spatial and spatio-temporal heterogeneity (Spatial and Spatio-temporal Epidemiology, 2019; 29C 97–109) <https://doi.org/10.1016/j.sste.2019.01.00> The authors were contacted after publication to request a Declaration of Interest statement.
6. NIMBLE for Bayesian Disease Mapping (Spatial and Spatio-temporal Epidemiology, 2020; 33C) <https://doi.org/10.1016/j.sste.2020.100323> The authors were contacted after publication to request a Declaration of Interest statement.
7. Geography and patient history in long-term lipid lowering medication adherence for primary prevention of cardiovascular disease (Spatial and Spatio-temporal Epidemiology, 2019; 29C 13–29) <https://doi.org/10.1016/j.sste.2018.12.001> The authors were contacted after publication to request a Declaration of Interest statement.
8. Spatial analysis of COVID-19 clusters and contextual factors in New York City (Spatial and Spatio-temporal Epidemiology, 2020; 34C) <https://doi.org/10.1016/j.sste.2020.100355> The authors were contacted after publication to request a Declaration of Interest statement.
9. A data-driven approach for estimating the change-points and impact of major events on disease risk (Spatial and Spatio-temporal Epidemiology, 2019; 29C 111–118) <https://doi.org/10.1016/j.sste.2018.08.005> The authors were contacted after publication to request a Declaration of Interest statement.
10. Where did I get dengue? Detecting spatial clusters of infection risk with social network data (Spatial and Spatio-temporal Epidemiology, 2018; 29C 163–175) <https://doi.org/10.1016/j.sste.2018.11.005> The authors were contacted after publication to request a Declaration of Interest statement.
11. Pure spatial and space-time clusters of self-harm in Kwai Tsing 2004 to 2012 (Spatial and Spatio-temporal Epidemiology, 2018; 27C 1–9) <https://doi.org/10.1016/j.sste.2018.07.002> The authors were contacted after publication to request a Declaration of Interest statement.
12. Analysis of the spatial distribution of scientific publications regarding vector-borne diseases related to climate variability in South America (Spatial and Spatio-temporal Epidemiology, 2018; 26C 35–93) <https://doi.org/10.1016/j.sste.2018.04.003> The authors were contacted after publication to request a Declaration of Interest statement.
13. Regional variation in lung and bronchus cancer survival in the US using mortality-to-incidence ratios (Spatial and Spatio-temporal Epidemiology, 2018; 26C 107–112) <https://doi.org/10.1016/j.sste.2018.06.004> The authors were contacted after publication to request a Declaration of Interest statement.

DOL of original article: [10.1016/j.sste.2020.100323](https://doi.org/10.1016/j.sste.2020.100323)

14. Dynamics of dengue disease with human and vector mobility (*Spatial and Spatio-temporal Epidemiology*, 2018; 25C 57–66) <https://doi.org/10.1016/j.sste.2018.03.001> The authors were contacted after publication to request a Declaration of Interest statement.
15. Climate variability and dengue fever in Makassar, Indonesia: Bayesian spatio-temporal modelling (*Spatial and Spatio-temporal Epidemiology*, 2020; 33C) <https://doi.org/10.1016/j.sste.2020.100335> The authors were contacted after publication to request a Declaration of Interest statement.