1 **Supplementary appendix**

2 **Supplement to: Global malaria infection risk from climate**

3 **change**

4

5 **Authors**

6 Chao Li[1], Shunsuke Managi*[1]

7

8 **Affiliations**

9 1 Urban Institute, Kyushu University, Japan

10 * Correspondent to: Shunsuke Managi, managi@doc.kyushu-u.ac.jp, Kyushu University

11 744 Motooka, Nishi-ku, Fukuoka 819-0395 Japan

12

13

Materials

Plasmodium falciparum Parasite Rate (PfPR)

The Malaria Atlas Project offers malaria infection prevalence annually, termed PfPR, in 2-10 olds globally from 2000 to 2019 [1] (https://malariaatlas.org/explorer/#/). This data set is a series of spatial grid data with a 5-km resolution, covering the regions from 60 ^{o} N to 60° S. Malaria is widely infected in Africa, though the infection prevalence has apparently decreased in the past 20 years [1]. To make the spatial resolution of PfPR concordant with the climatic data set, we upscale the spatial data set to 0.25 arc degree by the average method. **Figure S1** shows the spatial distribution of the temporally averaged value of PfPR in 2-10 olds (PfPR$_{2-10}$) from 2000 to 2019. In those past 20 years, although the coverage of the contemporary malaria control inventions was skyrocketing [1], the public health regarding malaria infection in developing countries, especially in Sub-Saharan African countries, was still remarkably weak. For example, even in 2019, the PfPR$_{2-10}$ is still over 60% in many Sub-Sahara African regions [1].

Temperature

Temperature data are extracted from the United States National Aeronautics and Space Administration (NASA) data set, NASA Global Land Data Assimilation System Version 2 (GLDAS-2.1). The GLDAS-2.1 system is based on a combination of model and observation data from 2000 to the present, whose products are a variety of monthly average climatic variables with a 0.25-arc-degree spatial resolution [2]. Firstly, we obtain several climatic variables from GLDAS-2.1, including temperature, air pressure, absolute

36 humidity, precipitation and wind speed. These variables are all monthly average values and

37 have been converted into annual average values. However, the correlations among these

38 climatic variables are too high, which would lead to multicollinearity in the regressions.

39 For example, the correlation coefficient between annual average air pressure and

40 temperature is 0.745 (p-value $< 0.1\%$). Additionally, another critical independent variable,

41 vegetation index, is significantly correlated with absolute humidity (correlation coefficient:

42 0.807, p-value $< 0.1\%$), precipitation (0.805, p-value $< 0.1\%$) and wind speed (-0.716, p-

43 value $< 0.1\%$). Therefore, the only annual average temperature is kept among those five

44 climatic variables. Secondly, evidence shows that malaria transmission through *Anopheles*

45 *gambiae* peaks at 25℃ [3]. Hence, the relationship between malaria infection rate and

46 temperature is non-linear and similar to the U-shape [4, 5, 6]. We take the temperature square

47 into account. Furthermore, extreme low and high temperatures are unsuitable for

48 transmission vector survival. In this way, the standard deviation of monthly average

49 temperature in a specific year would also affect the PfPR. In terms of climate, three

50 variables, involving annual average temperature, annual average temperature square, and

51 standard deviation of monthly average temperature in a specific year, are incorporated into

52 the analysis.

53

54 ***Vegetation Index***

55   Local vegetation affects malaria vector population size, survival, biting, among

56 others [7, 8, 9]. Therefore, the local vegetation status should be considered in the analysis. The

57 normalized difference vegetation index (NDVI) is a widely used indicator to assess how

58 much the observed grid contains live green vegetation [10, 11]. The NDVI ranges from -100%

59   (no live green vegetation) to 100% (full of live green vegetation). The moderate-resolution

60   imaging spectroradiomenter Terra and Aqua satellites provide monthly NDVI data from

61   2000 to the present with a 0.05-arc-degree spatial resolution. The MODIS vegetation index

62   products used in the analysis are MOD13C2 and MYD13C2 [12]. To link the NDVI data set

63   to the temperature and $PfPR_{2-10}$ data set, we upscale the NDVI data set to 0.25 arc degree

64   by the average method. Then, these data are further converted into annual average values.

65

Population Density and GDP per Capita

67   The population density data are obtained from the WorldPop Project. The

68   WorldPop Project provides spatial grid data set of population count from 2000 to 2020

69   with a 1-km resolution [13, 14]. Because the length of the latitudes decreases with the

70   increasing latitude, the conversion of population count from 1-km to 0.25-arc-degree

71   resolution is not perfect. At least, we cannot use the direct sum method. To solve this issue,

72   we regard the population count in every grid as the population density, and the unit of this

73   data set is regarded as capita per square kilometer (Cap/km^2). Here, we upscale the data

74   from 1-km to 0.25-arc-degree resolution by the average method. The data set eventually

75   illustrates the annual population density with a 0.25-arc-degree resolution.

76   GDP per capita is associated with development level and public health investment.

77   The higher development level and more public health investment, including a wide range

78   of international aid, are linked with a relatively lower malaria infection rate [1, 15]. However,

79   the spatial grid data set about GDP per capita is unavailable. We employ the country-level

80   GDP per capita from the World Bank. The unit of GDP per capita is the current U.S. dollar.

81   We assume that the GDP per capita in every grid is the same in a country.

82

Climate Change Scenarios and Temperature Prediction

84   Malaria infection rate is associated with temperature [3, 4, 6]. The different climate

85   change scenarios definitely lead to diverse malaria infection rates, assuming other

86   conditions are the same. The Intergovernmental Panel on Climate Change (IPCC) relies on

87   the Coupled Model Intercomparison Projects (CMIP) to predict the future temperature

88   based on several future emission scenarios [16, 17]. The 2013 IPCC fifth assessment report

89   (AR5) features four Representative Concentration Pathways (RCPs) according to various

90   possible future greenhouse gas emissions [17]. Based on the RCPs and the CMIP5 model, the

91   IPCC projects four scenarios, specifically RCP2.6, RCP4.5, RCP6.0, and RCP8.5, which

92   are well-known. In 2021, the IPCC sixth assessment report (AR6) is released. In the IPCC

93   AR6, updated emission scenarios are derived from different socioeconomic assumptions,

94   called the Shared Socioeconomic Pathways (SSPs) [16]. These scenarios are the updated

95   version of the four scenarios in the IPCC AR5, so the new scenarios listed in the IPCC

96   AR6 are named SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 and are projected by the

97   CMIP6 model [16]. Except SSP3-7.0, the other three scenarios are advanced from the

98   previous scenarios in the AR5. The SSP1-2.6 scenario predicts that the temperature

99   increase stabilizes at about 1.8°C by the end of the century. In the SSP2-4.5 scenario, the

100 temperatures rise 2.7°C by 2100. According to the SSP3-7.0 scenario, the temperatures go

101 up around 3.6°C by the end of the century. Finally, the SSP5-8.5 scenario predicts that the

102 temperatures grow 4.4°C or higher by 2100. The IPCC projects global temperatures in four

103 periods, including 2021-2040 (near term), 2041-2060 (medium term), and 2081-2100 (long

104 term), based on the different scenarios. In each period, the data set contains 12 grid data of

105   monthly average temperature. The IPCC provides open access data based on over 30-model

106   stimulation with a 1-arc-degree resolution (https://interactive-atlas.ipcc.ch/). Furthermore,

107   the IPCC also offers the predicted population density grid data produced by the models'

108   stimulation with a 1-arc-degree resolution. We obtain the population density data under

109   SSP2-4.5, SSP3-7.0, and SSP5-8.5, during three periods.

110

111   *Grids*

112   We integrate the abovementioned data set to build a long panel spatial grid data set.

113   Because we only focus on the endemic areas, the grids are dropped, whose $PfPR_{2-10}$ is

114   always 0. Additionally, every grid is required to have at least two-year complete records.

115   Without enough records, the grid is also removed. Accordingly, 65,705 grids are kept. The

116   coverage of our study exceeds 40 million km^2 based on the assumption that every grid

117   covers 770 km^2. According to statistical tests and data availability, our analysis includes

118   temperature, NDVI, population density, and GDP per capita. Previous studies indicate that

119   climate change affects vector-borne diseases [18, 19, 20], so temperature-related variables and

120   NDVI are considered. Moreover, population density and prevention also play a critical role

121   in malaria transmission [1, 21]. Although detailed prevention investments are unavailable, we

122   replace them with GDP per capita.

123

Methods

Spatially Stationary Model

126 At first, we assume that relationships between $PfPR_{2-10}$ and other selected

127 independent variables are the same in each grid. For example, the marginal effects of a 1%

128 increase in the NDVI in Africa and South America are the same, even though they are

129 thousands of kilometers away. Without considering the spatial contexts, those relationships

130 do not vary spatially [22, 23]. To estimate the links among variables, we apply the panel

131 regression method. Because our data set is a panel data set, the hypotheses on the time-

132 fixed effects of each grid should be carefully tested. Three typical panel regression models

133 based on different time-fixed effect hypotheses, including fixed effects model (FEM),

134 random effects model (REM), and pooled ordinary least square (POLS), are widely used.

135 We perform several statistical tests to select the reasonable spatially stationary model.

136 Firstly, we execute the F test for individual effects [24, 25]. The significant result indicates that

137 the time-fixed effects exist since the null hypothesis is that no time-fixed effects are needed

138 [25]. Hence, the POLS is rejected because this model assumes that no time-fixed effects exist.

139 Then, we implement the Hausman test to detect that either the REM or the FEM is

140 reasonable [26]. The null hypothesis of the Hausman test that the preferred model is the REM

141 is rejected due to the significant test result. Therefore, the FEM is preferred in the spatially

142 stationary analysis.

143 Here, the FEM is illustrated as follows:

$$PfPR_{it} = \boldsymbol{\beta} \boldsymbol{X'}_{it} + \alpha_i + \mu_{it} \qquad (1)$$

144   where $PfPR_{it}$ denotes the annual PfPR2-10 (ranging from 0 to 1) in grid i in t year, $\boldsymbol{X}_{it}$

145   denotes a matrix of independent variables, including annual average temperature (℃),

146   annual average temperature square (℃^{2}), the standard deviation of monthly average

147   temperature in t year (℃), NDVI, population density (Cap/km^2), and GDP per capita

148   (USD/Cap) in measurement i in t year, α_i denotes the time-fixed effects, μ_{it} denotes an

149   idiosyncratic error, and $\boldsymbol{\beta}$ is a vector of parameters to be estimated.

150

Spatially Non-stationary Model

152   Obviously, the marginal effects of the climate variable on malaria infection rate are

153   not always, even never, the same spatially in the real world. For instance, Giesen et al.

154   summarize the impacts of climate change on malaria in Africa in their review, and the

155   marginal effects of climate change are different country by country [27]. Two possible

156   reasons flash. Firstly, the relationship between climate change and the malaria infection

157   rate is non-linear [3, 4, 5, 6]. To solve this issue, we employ the square temperature term.

158   Secondly, the association of malaria infection rate with temperature spatially varies [28]. The

159   advanced spatially non-stationary model should be employed to grasp this spatial

160   variability.

161   Geographically weighted panel regression (GWPR) is an improvement of

162   geographically weighted regression, which allows the regression coefficients to vary

163   spatially [29, 30] and is widely used in spatial analyses. Basically, the GWPR divides the total

164   data set into a large number of sub-datasets according to the optimal bandwidth. The

165   bandwidth of the GWPR is a threshold distance to judge whether two grids have a spatial

166   relationship [22, 29, 31]. The number of sub-datasets equals the number of grids. Every sub-

167   dataset takes one grid as the center and the bandwidth as the radius to select the other grids.

168   Of note, every grid could be utilized several times in the dividing process. The bandwidth

169   calibration is the essential step in the analysis. The mean square prediction error of the

170   regression is the critical index to calibrate the bandwidth. The lowest mean square

171   prediction error means the highest goodness of fit [32]. In the GWPR bandwidth calibration,

172   the mean square prediction error is calculated as follows:

$$MSPE(b) = \frac{m\sum_j[y_j - \widehat{y_j(b)}]^2}{(m-p+1)^2} \tag{2}$$

173   where $MSPE(b)$ is the mean square prediction error when the bandwidth is b, m is the data

174   size, y_j is the dependent variable of the jth record, $\widehat{y_j(b)}$ is the predicted value of the jth

175   record when the bandwidth is b, and p is the number of the parameters in the analysis. In

176   the GWR, the basic version of the GWPR, in a way, previous studies generally assume that

177   **Equation 2** is a U-shape function without statistical or mathematical evidence [22, 23, 32]. This

178   assumption helps to reduce calculating time but also brings enormous risks to the analysis.

179   We apply the step increment selection without assuming that **Equation 2** is a U-

180   shape function. The bandwidth selection extent is from 0.25 to 20 arc degrees, and the

181   examined bandwidth increases 0.25 arc degree every time because the spatial resolution of

182   the grid data is 0.25 arc degree. The optimal bandwidth is 4.25 arc degrees (**Figure S2**).

183   The mean square prediction error reaches the minimum when the bandwidth equals 0.75,

184   but we reject it. According to the grid data set spatial resolution, every sub-dataset at most

185   has 25 grids, but some of them might only have one grid. If the sub-datasets have only one

186   grid, they do not fit FEM's statistical requirements. Moreover, some grids might not be

187   spatially linked with other grids with a 0.75-arc-degree bandwidth, which leads to a not

188   robust GWPR result. Therefore, the optimal bandwidth, 4.25 arc degrees, is calibrated.

189   The GWPR is performed after the optimal bandwidth has been calibrated. The

190   GWPR model is written as follows:

$$PfPR_{it} = \boldsymbol{\beta}_i \boldsymbol{X'}_{it} + \alpha_i + \epsilon_{it} \tag{3}$$

191   where $\boldsymbol{\beta}_i$ denotes a vector of parameters in the regression with the sub-dataset taking the

192   grid i as the center. There are 65,705 sub-datasets, so the number of regressions on sub-

193   datasets is also 65,705, and every estimation is based on the sub-dataset's spatial weight

194   vector. For this reason, the parameters, $\boldsymbol{\beta}_i$, vary spatially. The spatial weight vector is

195   calculated as follows:

$$\boldsymbol{W}_i = \begin{cases} [1 - (\dfrac{\boldsymbol{d}}{b})^2]^2, d_k \in \boldsymbol{d} \ \ if \ d_k \leq \ b \\ \ \ \ 0, \ \ \ \ \ \ d_k \notin \boldsymbol{d} \ \ if \ d_k > b \end{cases} \tag{4}$$

196   where $\boldsymbol{W}_i$ denotes a vector of spatial weights between the grid i and its neighbors, $\boldsymbol{d}$

197   represents a vector of distances between the grid i and its neighbors, b is the calibrated

198   optimal bandwidth, and d_k is the distance between the grid i and the grid k. Then, the

199   GWPR estimates the parameters based on the spatial weight vectors. Because the time-

200   fixed effects term (α_i) in is unknown in **Equation 3** [24], we transform that equation as

201   follows:

$$PfPR_{it} - \overline{PfPR_\iota} = \boldsymbol{\beta}_i (\boldsymbol{X}_{it} - \overline{\boldsymbol{X}_\iota})' + (\epsilon_{it} - \overline{\epsilon_\iota}) \tag{5}$$

202   where $\overline{PfPR_\iota}$ denotes the mean of the PfPR$_{2\text{-}10}$ of the grid i, $\overline{\boldsymbol{X}_\iota}$ denotes a vector of means

203   of independent variables of the grid i, and $\overline{\epsilon_\iota}$ denotes the mean of the error of the grid i. To

204   simplify **Equation 5**, we define the matrix of transformed independent variables $\boldsymbol{x}_{it}$ and

205   the transformed dependent variable $pfpr_{it}$ as follows:

$$x_{it} = X_{it} - \overline{X_\iota} \tag{6}$$

$$pfpr_{it} = PfPR_{it} - \overline{PfPR_\iota} \tag{7}$$

$$\sigma_{it} = \epsilon_{it} - \overline{\epsilon_\iota} \tag{8}$$

206   In the light of **Equations 6-8**, **Equation 3** is rewritten as follows:

$$pfpr_{it} = \boldsymbol{\beta}_i \boldsymbol{x'}_{it} + \sigma_{it} \tag{9}$$

207   The parameter estimation could be displayed as follows:

$$\boldsymbol{\beta}_i = [\boldsymbol{x}_{it}{}^T \boldsymbol{W}_i \boldsymbol{x}_{it}]^{-1} \boldsymbol{x}_{it}{}^T \boldsymbol{W}_i pfpr_{it} \tag{10}$$

208

209   ***Statistical Indicators and 10-fold Cross-Validation***

210   Because we need to make the prediction based on the results of IPCC stimulations,

211   we must confirm the model's accuracy. Several statistical indicators, including R^2, root

212   mean square error (RMSE), mean absolute error (MAE), correlation coefficient (*r*) between

213   observed and predicted values, and regression coefficients between observed and predicted

214   values (intercept α and slope β), are applied to depict the accuracy of the model. R^2 is a

215   critical statistical indicator describing the goodness of fit, which is expressed as follows:

$$R^2 = 1 - \frac{\sum_{k=1}^{n}(OPfPR_k - PPfPR_k)^2}{\sum_{k=1}^{n}(OPfPR_k - \overline{OPfPR})^2} \tag{11}$$

216   where n represents the observation number in the whole data set, $OPfPR_k$ represents the

217   kth record of the observed PfPR$_{2\text{-}10}$, $PPfPR_k$ represents the kth record of the predicted

218   PfPR$_{2\text{-}10}$, and $\overline{OPfPR}$ represents the mean of the observed PfPR$_{2\text{-}10}$. The RMSE is sensitive

219   to both systematic error and random error, expressed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(OPfPR_k - PPfPR_k)^2} \qquad (12)$$

220   The MAE is computed as follows:

$$MAE = mean(|OPfPR_k - PPfPR_k|) \qquad (13)$$

221   In the analysis, the MAE is expected to be lower. Additionally, the regression coefficients

222   between observed and predicted values are estimated as follows:

$$OPfPR_k = \alpha + \beta PPfPR_k + \delta_k \qquad (14)$$

223   where α is the intercept of the regression, the ideal value of α is 0, β is the slope, its ideal

224   value is 1, and δ_k is a random error term.

225   To check the reliability of the GWPR model, we perform the 10-fold cross-

226   validation. In the 10-fold cross-validation, the whole data set is randomly divided into ten

227   subsets. In every individual fold cross-validation, nine subsets are used to train the GWPR

228   model, while the left one subset is employed to test the accuracy and ability of prediction.

229   All statistical indicators, involving R^2, RMSE, MAE, r, α, and β, are reported based on the

230   training and testing results.

231

Effects of Temperature Change on PfPR$_{2-10}$

233   The effects of temperature change on PfPR$_{2-10}$ cannot directly be estimated because

234   other necessary variables, such as NDVI, population, and GDP per capita, cannot be

235   precisely predicted. The errors from the other variable predictions would drastically affect

236   the estimation of PfPR$_{2-10}$. However, if we assume that other variables are the same in the

237   various scenarios, we could extract the difference among the impacts of temperature

238   change in different periods. The difference between the two temperature scenarios is

239   calculated as follows:

$$\Delta PfPR_{it} = S2PfPR_{it} - S1PfPR_{it} \tag{15}$$

240   where $\Delta PfPR_{it}$ denotes the difference between the estimation of PfPR$_{2\text{-}10}$ with the

241   different temperature scenarios in the grid i during the period t, $S1PfPR_{it}$ denotes the

242   estimation of PfPR$_{2\text{-}10}$ based on the temperature scenario $S1$ in the grid i during the period

243   t, and $S2PfPR_{it}$ denotes the estimation of PfPR$_{2\text{-}10}$ based on the temperature scenario $S2$

244   in the grid i during the period t. We hope to obtain the impacts of the higher greenhouse

245   gas emission on PfPR$_{2\text{-}10}$, so the SSP1-2.6 scenario is always set as the temperature

246   scenario $S1$ and the temperature scenario $S2$ is one of the other three scenarios, including

247   SSP2-4.5, SSP3-7.0, and SSP5-8.5.

248   To estimate $S1PfPR_{it}$ and $S2PfPR_{it}$, we expand **Equation 3** as follows:

$$S1PfPR_{it} = \boldsymbol{\beta}_{1i}\boldsymbol{TeS1}'_{it} + \boldsymbol{\beta}_{2i}\boldsymbol{TeS1^{2}}'_{it} + \boldsymbol{\beta}_{3i}\boldsymbol{TSDS1}'_{it} + \boldsymbol{\beta}_{4i}\boldsymbol{NDVI}'_{it} \tag{16}$$
$$+ \boldsymbol{\beta}_{5i}\boldsymbol{POP}'_{it} + \boldsymbol{\beta}_{6i}\boldsymbol{GDP}'_{it} + \alpha_i$$

$$S2PfPR_{it} = \boldsymbol{\beta}_{1i}\boldsymbol{TeS2}'_{it} + \boldsymbol{\beta}_{2i}\boldsymbol{TeS2^{2}}'_{it} + \boldsymbol{\beta}_{3i}\boldsymbol{TSDS2}'_{it} + \boldsymbol{\beta}_{4i}\boldsymbol{NDVI}'_{it} \tag{17}$$
$$+ \boldsymbol{\beta}_{5i}\boldsymbol{POP}'_{it} + \boldsymbol{\beta}_{6i}\boldsymbol{GDP}'_{it} + \alpha_i$$

249   where $\boldsymbol{\beta}_{1i}, \boldsymbol{\beta}_{2i}, \boldsymbol{\beta}_{3i}, \boldsymbol{\beta}_{4i}, \boldsymbol{\beta}_{5i}$, and $\boldsymbol{\beta}_{6i}$ are parts of the estimated GWPR parameters $\boldsymbol{\beta}_i$

250   from **Equation 10**, $\boldsymbol{TeS1}_{it}$ denotes the stimulated annual average temperature under the

251   SSP1-2.6 scenario in the grid i during the period t, $\boldsymbol{TSDS1}_{it}$ denotes the standard

252   deviation of stimulated monthly average temperature under the SSP1-2.6 scenario in the

253   grid i during the period t, $\boldsymbol{TeS2}_{it}$ denotes the stimulated annual average temperature under

254   another scenario in the grid i during the period t, $\boldsymbol{TSDS2}_{it}$ denotes the standard deviation

255   of stimulated monthly average temperature under another scenario in the grid i during the

256 period t, and $\textbf{NDVI}_{it}$, $\textbf{POP}_{it}$, and $\textbf{GDP}_{it}$ denote the NDVI, population density, and GDP

257 per capita in the grid i during the period t, respectively.

258 In view of **Equations 16** and **17**, **Equation 15** is rewritten and simplified as follows:

$$\Delta PfPR_{it} = \boldsymbol{\beta}_{1i}(\boldsymbol{TeS2}_{it} - \boldsymbol{TeS1}_{it})' + \boldsymbol{\beta}_{2i}(\boldsymbol{TeS2^2}_{it} - \boldsymbol{TeS1^2}_{it})' \qquad (18)$$
$$+ \boldsymbol{\beta}_{3i}(\boldsymbol{TSDS2}_{it} - \boldsymbol{TSDS1}_{it})'$$

259

Difference in Infection Cases between Different Scenarios

261 We estimate the infection case changes due to the shifts in development scenarios

262 based on the predicted differences of $PfPR_{2\text{-}10}$ between different scenarios. In this study,

263 the predicted dependent variable is the $PfPR_{2\text{-}10}$, but the IPCC does not provide the

264 population in 2-10 olds. Hence, we assume the PfPR in all ages is the same as $PfPR_{2\text{-}10}$ to

265 predicate the infection case changes. The infection cases changes can be estimated as

266 follows:

$$\Delta ICC_{itk} = \Delta PfPR_{itk} \times Population_{itk} \qquad (19)$$

267 where ΔICC_{itk} denotes the infection case change due to the k shift of development

268 scenarios in the grid i during the period t, $\Delta PfPR_{itk}$ denotes the difference due to the k

269 shift of development scenarios in the grid i during the period t estimated by **Equation 18**,

270 and $Population_{itk}$ denotes the IPCC predicted population in the grid i during the period

271 t under the higher-temperature scenario in the k shift. The shifts of development scenarios

272 include three types: from SSP1-2.6 to SSP2-4.5, from SSP1-2.6 to SSP3-7.0, and from

273 SSP1-2.6 to SSP5-8.5.

274

Model Performance

The R^2 of the GWPR model is 94.75%, indicating the high accuracy of the estimation. The outcomes of 10-fold cross-validation are listed in **Table S1**. In the 10-fold cross-validation, the R^2s of the training and testing data sets are over 94%, which hints at the model's stability. Besides, the RMSEs, MEAs, αs, βs, and *r*s are close to their ideal values, further proposing the high reliability of the estimated impacts of temperature change on PfPR$_{2\text{-}10}$.

Estimated Effects of Temperature Change on PfPR2-10

Figures S3-S5 illustrate the difference between the estimation of PfPR$_{2\text{-}10}$ with the different temperature scenarios during 2021 – 2040, including the difference between SSP2-4.5 and SSP1-2.6 ($\Delta PfPR_{2021-2040}^{SSP245-SSP126}$), the difference between SSP3.7-0 and SSP1-2.6 ($\Delta PfPR_{2021-2040}^{SSP370-SSP126}$), and the difference between SSP5.8-5 and SSP1-2.6 ($\Delta PfPR_{2021-2040}^{SSP585-SSP126}$). **Figure S6** demonstrates the statistical distribution of $\Delta PfPR_{2021-2040}^{SSP245-SSP126}$, assuming that all conditions are the same except temperatures. The higher temperature scenario SSP2-4.5 positively impacts the malaria infection rate in many grids, resulting in a potential fall in PfPR$_{2\text{-}10}$ during 2041 - 2060. Nevertheless, most grids show an increasing trend. Manifest in Africa, the rising trends in Congo (CD), Central African Republic (CF), Cameroun (CM), Burkina Faso (BF), Nigeria (NG), etc., mark the adverse effects of climate change. Since the difference in temperature change between SSP2-4.5 and SSP1-2.6 is tiny, the impacts of climate change are not potent, but they appear in the worst malaria-affected areas (**Figure S1**). According to **Figures S7** and **S8**

297   that illuminate the statistical distributions of $\Delta PfPR_{2021-2040}^{SSP370-SSP126}$ and

298      $\Delta PfPR_{2021-2040}^{SSP585-SSP126}$, respectively, the influences of climate change shift partially. In

299   **Figure S7**, the impacts of temperature change from SSP1-2.6 to SSP3-7.0 on PfPR$_{2-10}$

300   range from -1.68% to 2.66%, and the grid number of positive impacts outstrips the number

301   of grids with negative effects. **Figure S8**'s distribution is similar to **Figure S7**'s. Moreover,

302   we compare the geographical distributions of average PfPR$_{2-10}$ and temperature change's

303   impacts. Notably, compared with other countries, CD, CF, CM, BF, and NG are the worst

304   hit by malaria in the past 20 years. Worse, they might face the most harmful impacts of

305   climate change on the malaria infection rates during 2021 - 2040.

306   With further increase in temperature during 2041-2060, the distribution and

307   intention of negative impacts shift. **Figures S9-S11** show the difference between SSP2-4.5

308   and SSP1-2.6 ($\Delta PfPR_{2041-2060}^{SSP245-SSP126}$), the difference between SSP3.7-0 and SSP1-2.6

309   ($\Delta PfPR_{2041-2060}^{SSP370-SSP126}$), and the difference between SSP5-8.5 and SSP1-2.6

310   ($\Delta PfPR_{2041-2060}^{SSP585-SSP126}$) during 2041 – 2060. $\Delta PfPR_{2041-2060}^{SSP245-SSP126}$ becomes minus in most

311   areas of CD, CF, CM, BF, and NG. The positive impacts of the increase in temperature on

312   PfPR$_{2-10}$ also rise, according to $\Delta PfPR_{2041-2060}^{SSP370-SSP126}$ and $\Delta PfPR_{2041-2060}^{SSP585-SSP126}$ because in

313   SSP3.7-0 and SSP5-8.5, the predicted temperatures are higher than the scenario SSP2-4.5.

314   The negative impacts mainly appear in Gabon (GA), Angola (AO), and Zambia (ZM),

315   which are at the boundaries of high malaria-affected regions. The possible result of climate

316   change, especially in Sub-Saharan Africa, is that the increased temperature would spur the

317   high-risk area to spread to the higher latitude regions and coastlines. Moreover, the original

318   high-risk areas are no longer suitable for human living, owing to extreme temperatures.

319   **Figures S12-S14** demonstrate the statistical distribution of $\Delta PfPR_{2041-2060}^{SSP245-SSP126}$,

320     $\Delta PfPR_{2041-2060}^{SSP370-SSP126}$, and $\Delta PfPR_{2041-2060}^{SSP585-SSP126}$, respectively. Compared with differences

321   among the scenarios during 2021- 2040, the values of $\Delta PfPR_{2041-2060}^{SSP245-SSP126}$,

322     $\Delta PfPR_{2041-2060}^{SSP370-SSP126}$, and $\Delta PfPR_{2041-2060}^{SSP585-SSP126}$ are absolutely larger. Moreover, the

323   impacts of the increase in temperature are positive in most grids of Africa, Asia, and South

324   America.

325   By the end of the 21st century, the four scenarios end in different temperature

326   predictions. **Figures S15-S17** demonstrate the difference between SSP2-4.5 and SSP1-2.6

327   ($\Delta PfPR_{2081-2100}^{SSP245-SSP126}$), the difference between SSP3.7-0 and SSP1-2.6

328   ($\Delta PfPR_{2081-2100}^{SSP370-SSP126}$), and the difference between SSP5.8-5 and SSP1-2.6

329   ($\Delta PfPR_{2081-2100}^{SSP585-SSP126}$) during 2081 – 2100. Because the difference in temperatures of the

330   four scenarios becomes more significant, the impacts of temperature change on PfPR$_{2\text{-}10}$

331   are intensified. The spatial distributions of $\Delta PfPR_{2081-2100}^{SSP245-SSP126}$, $\Delta PfPR_{2081-2100}^{SSP370-SSP126}$,

332   and $\Delta PfPR_{2081-2100}^{SSP585-SSP126}$ are similar to the situation of $\Delta PfPR_{2081-2100}^{SSP585-SSP126}$, but the

333   values of the difference become larger during 2080 – 2100. **Figures S18-S20** propose the

334   statistical distribution of $\Delta PfPR_{2081-2100}^{SSP245-SSP126}$, $\Delta PfPR_{2081-2100}^{SSP370-SSP126}$, and

335     $\Delta PfPR_{2081-2100}^{SSP585-SSP126}$. In fact, $\Delta PfPR_{2081-2100}^{SSP585-SSP126}$ ranges -77.20% to 101.54%.

336   **Figure S21** illustrates the shapes of the relationship between annual average

337   temperature and PfPR$_{2\text{-}10}$ in each grid. The basic shapes of the relationship are two types,

338   specifically, U-shape and inverted U-shape. The coefficient of the square temperature term

339   decides the basic shape of the relationship. Additionally, together with the coefficient of

340   temperature term, the symmetry lines of the relationship of each grid can be pinned down.

341   In **Figure S21**, we set the temperature range interested from 0°C to 50°C. If the symmetry

342 lines of the relationships of some grids do not belong to the range interested, we consider

343 those relationships are either monotonically decreasing or monotonically increasing in the

344 range, mathematically. Therefore, the other four shapes are derived from the basic types,

345 including increase based on inverted U-shape (Increase Inverted U), decrease based on U-

346 shape (Decrease U), decrease based on inverted U-shape (Decrease Inverted U), and

347 increase based on U-shape (Increase U). Previous studies generally put forward that the

348 relationship at national or regional levels is an inverted U-shape, and the temperature

349 impacts would peak at a specific temperature. However, the relationship at the grid level is

350 obviously more complicated and various. Hence, we must treat the impacts of climate

351 change carefully. Although we accept the assumption that the relationship between

352 temperature and malaria infection rates is non-linear, we did not force to set the malaria-

353 transmission peaking temperature to 25°C or other values, as the previous studies

354 mentioned [4, 6]. In our results, most grids' transmission peaking temperatures are not 25°C.

355 We employ the annual average temperature to probe the relationship, whereas previous

356 studies use laboratory environments and real-time or shorter-term temperatures. As

357 mentioned before, even though the annual average temperature is 25°C, how many days

358 and hours are at that temperature remains unknown. For this reason, in many grids, the

359 transmission peaking temperatures are not 25°C. The shapes in each grid are even different,

360 though either U-shape or inverted U-shape. Furthermore, real-world environments are

361 significantly more complex than laboratory environments. The ignored factors by the

362 laboratory environments are also related to the malaria-transmission peaking temperature.

363 These ignored variables lead to the difference in the transmission peaking temperatures

364 between our result and previous studies.

365

Estimated Difference in Infection Cases between Different Scenarios

368   **Figures S22-S24** demonstrate the spatial distributions of the estimated differences

369 in infection cases between different scenarios ($\Delta DIC_{2021-2040}$) during $2021 - 2040$. Under

370 three scenario shifts, including from SSP1-2.6 to SSP2-4.5, from SSP1-2.6 to SSP3-7.0,

371 and from SSP1-2.6 to SSP5-8.5, the $\Delta DIC_{2021-2040}$ are expressed as $\Delta DIC_{2021-2040}^{SSP126-SSP245}$,

372 $\Delta DIC_{2021-2040}^{SSP126-SSP370}$, and $\Delta DIC_{2021-2040}^{SSP126-SSP585}$, respectively. The spatial distribution of

373 $\Delta DIC_{2021-2040}$ is affected by the population density pattern and the difference between

374 estimated PfPR$_{2\text{-}10}$ with different temperature scenarios. Hence, the PfPR$_{2\text{-}10}$ most-

375 increased regions might not have the largest infection case increment. However, infection

376 cases are predicted to increase most in NG, CD, and BF, if the scenario changes from SSP1-

377 2.6 to SSP2-4.5 in the near term. In some areas of these three countries, the infection cases

378 increase by over 10 cases/km^2 (**Figure S22**). Slight rising temperatures aggravate the

379 situation in the countries severely hit by malaria. Under the shifts from SSP1-2.6 to SSP2-

380 4.5 and SSP5-8.5, NG, CD, and BF are still the most affected, but the impact degree

381 decreases (**Figures S23** and **S24**). This reduction in the negative impacts of temperatures

382 is because the temperatures in these countries are close to the most suitable value for

383 malaria transmission. **Figures S25-S27** illustrate the statistical distribution of

384 $\Delta DIC_{2021-2040}^{SSP126-SSP245}$, $\Delta DIC_{2021-2040}^{SSP126-SSP370}$, and $\Delta DIC_{2021-2040}^{SSP126-SSP585}$, respectively. In Africa,

385 most grids suffer from the negative impact of climate change on malaria transmission in

386 the near term.

387   **Figures S28-S30** demonstrate the spatial distributions of $\Delta DIC_{2041-2060}^{SSP126-SSP245}$,

388    $\Delta DIC_{2041-2060}^{SSP126-SSP370}$, and $\Delta DIC_{2041-2060}^{SSP126-SSP585}$, respectively. During 2041 - 2060, the

389   impacts of temperature increase become positive in most regions because the temperatures

390   in those areas exceed the most suitable value for malaria transmission. For example, in the

391   Southern part of NG, $\Delta DIC_{2041-2060}$ is less than -20 case/km^2 under all three scenario

392   shifts, which indicates a decrease in malaria infection cases due to temperature increase.

393   Compared with **Figures S22-S24**, the temperature impacts show an opposite direction in

394   most regions in **Figures S28-S30. Figures S31-S33** illustrate the statistical distribution of

395    $\Delta DIC_{2041-2060}^{SSP126-SSP245}$, $\Delta DIC_{2041-2060}^{SSP126-SSP370}$, and $\Delta DIC_{2041-2060}^{SSP126-SSP585}$, respectively. In most

396   regions, the temperature increase positively impacts the public health issue.

397   With the temperature increase during 2081 - 2100, the negative impacts of the

398   temperature increase move to high-latitude and coastline areas. **Figures S34-S36**

399   demonstrate the spatial distributions of $\Delta DIC_{2081-2100}^{SSP126-SSP245}$, $\Delta DIC_{2081-2100}^{SSP126-SSP370}$, and

400    $\Delta DIC_{2081-2100}^{SSP126-SSP585}$, respectively. The most negative impacts locate in severely infected

401   borders, as shown in **Figure S1**. **Figures S37-S39** show the statistical distributions of

402    $\Delta DIC_{2081-2100}^{SSP126-SSP245}$, $\Delta DIC_{2081-2100}^{SSP126-SSP370}$, and $\Delta DIC_{2081-2100}^{SSP126-SSP585}$. Although the ΔDIC

403   becomes negative in most areas, high-risk regions shift to areas that are less severely

404   infected during 2000 - 2019. Therefore, the effect of rising temperatures must be taken into

405   account.

406

Data Availability

All data sources used in the analyses, along with fully reproducible code, are publicly available at https://github.com/MichaelChaoLi-cpu/Malaria_And_Climate_Change.

412   **Figure:**

413
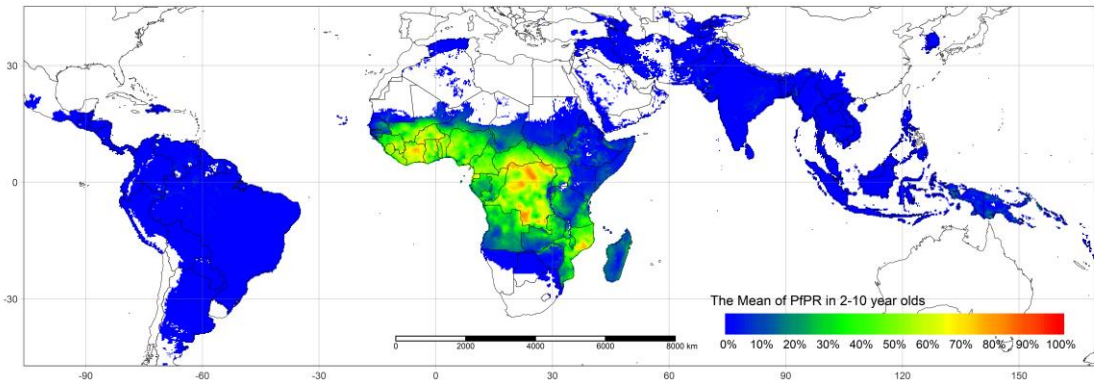
414   **Figure S1: The mean of PfPR$_{2\text{-}10}$ from 2000 to 2019**

415   (Note: if the PfPR$_{2\text{-}10}$ in a grid is always 0 from 2000 to 2019, the grids are blank)

416

417
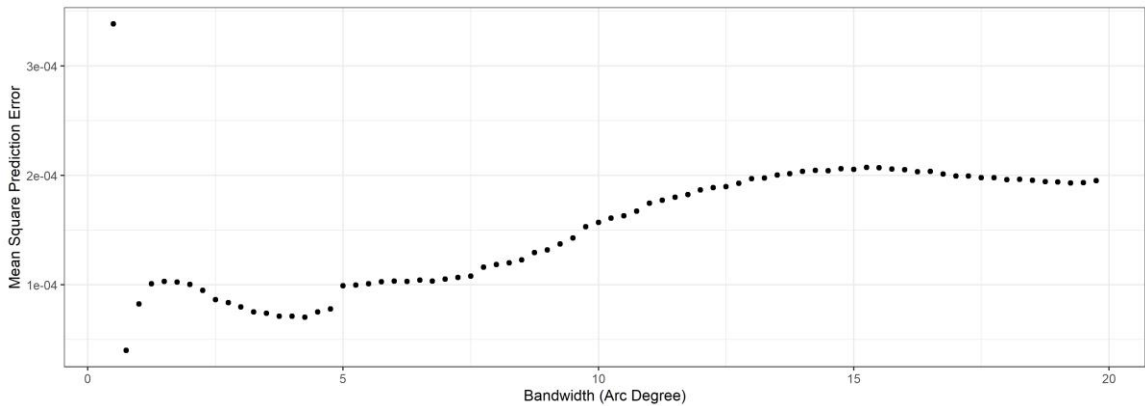
418   **Figure S2: Step Increment Bandwidth Selection**

419
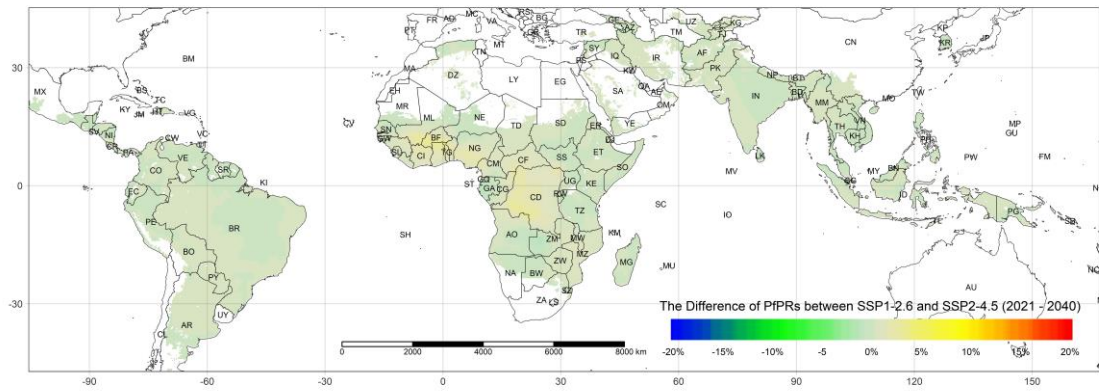
Figure S3: The Difference in PfPR$_{2-10}$ between SSP1-2.6 and SSP2-4.5 (2021 - 2040)

421
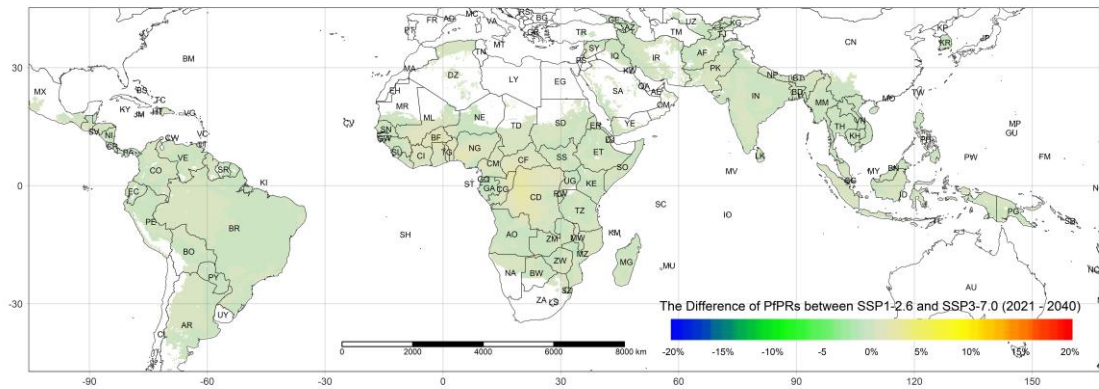
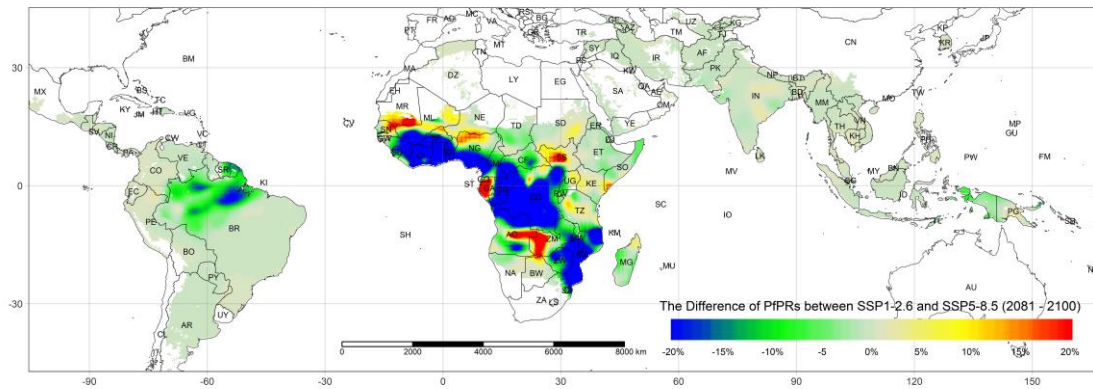**Figure S4: The Difference in PfPR$_{2-10}$ between SSP1-2.6 and SSP3-7.0 (2021 - 2040)**

423
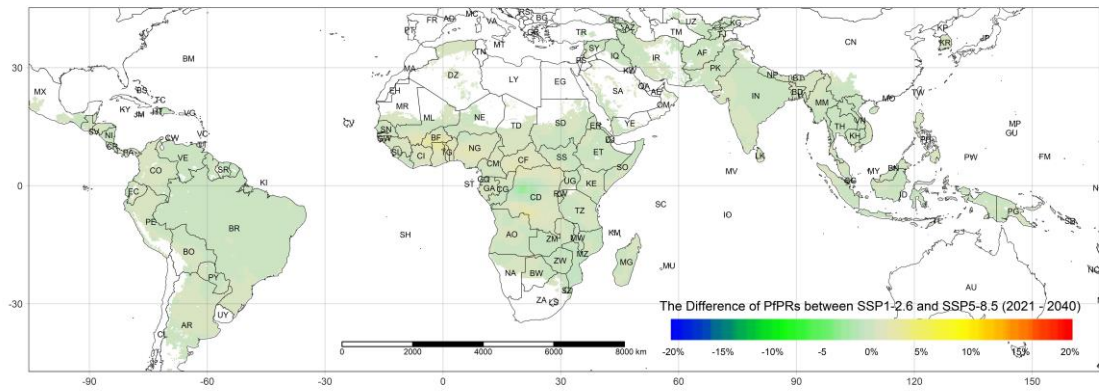
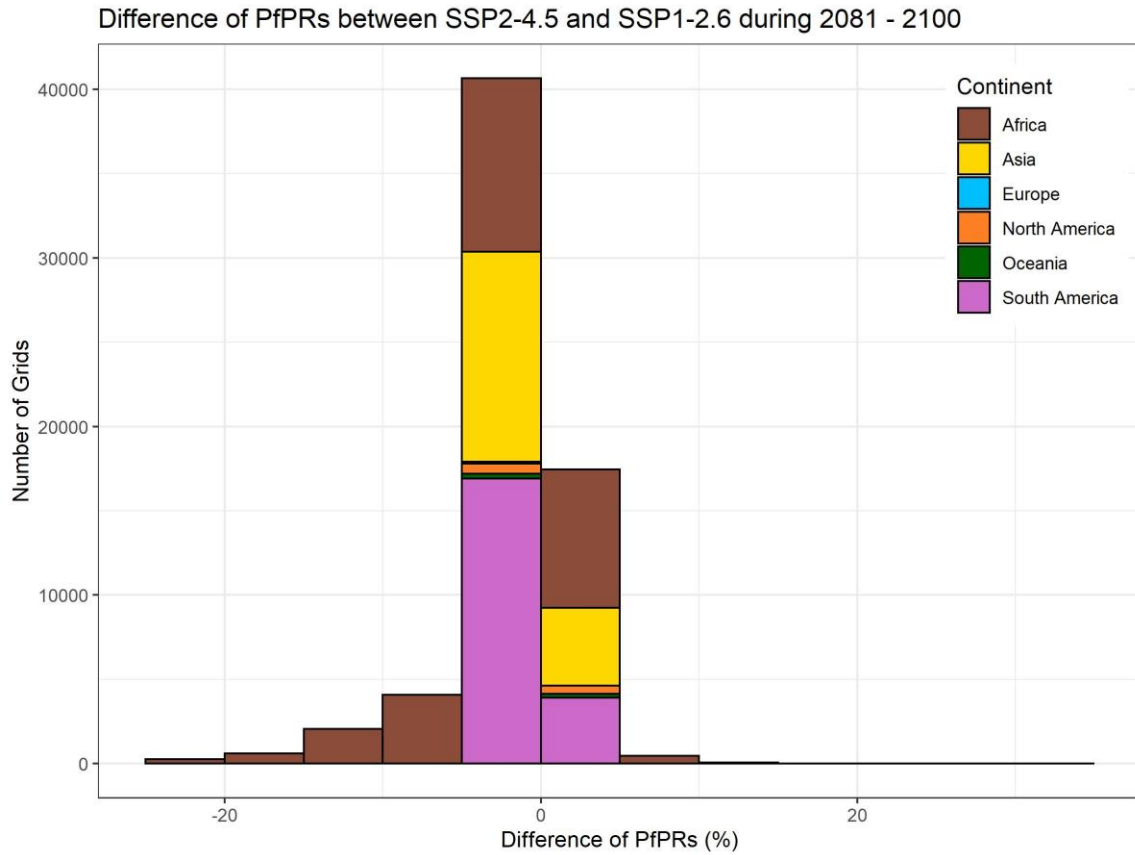**Figure S5: The Difference in PfPR$_{2-10}$ between SSP1-2.6 and SSP5-8.5 (2021 - 2040)**

425

426

427

Difference of PfPRs between SSP2-4.5 and SSP1-2.6 during 2021 - 2040

428
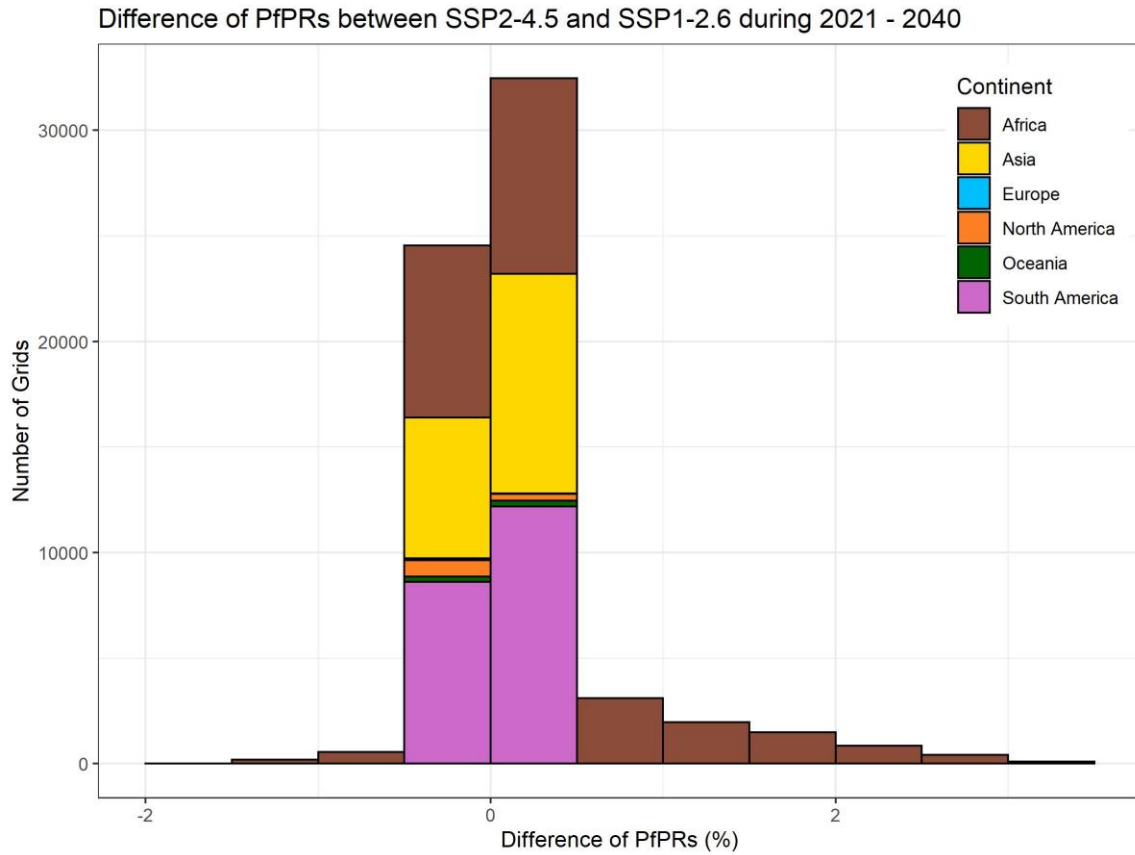
Figure S6: The Distribution of the Difference in PfPR$_{2-10}$ between SSP1-2.6 and

SSP2-4.5 (2021 - 2040)

Difference of PfPRs between SSP3-7.0 and SSP1-2.6 during 2021 - 2040

431
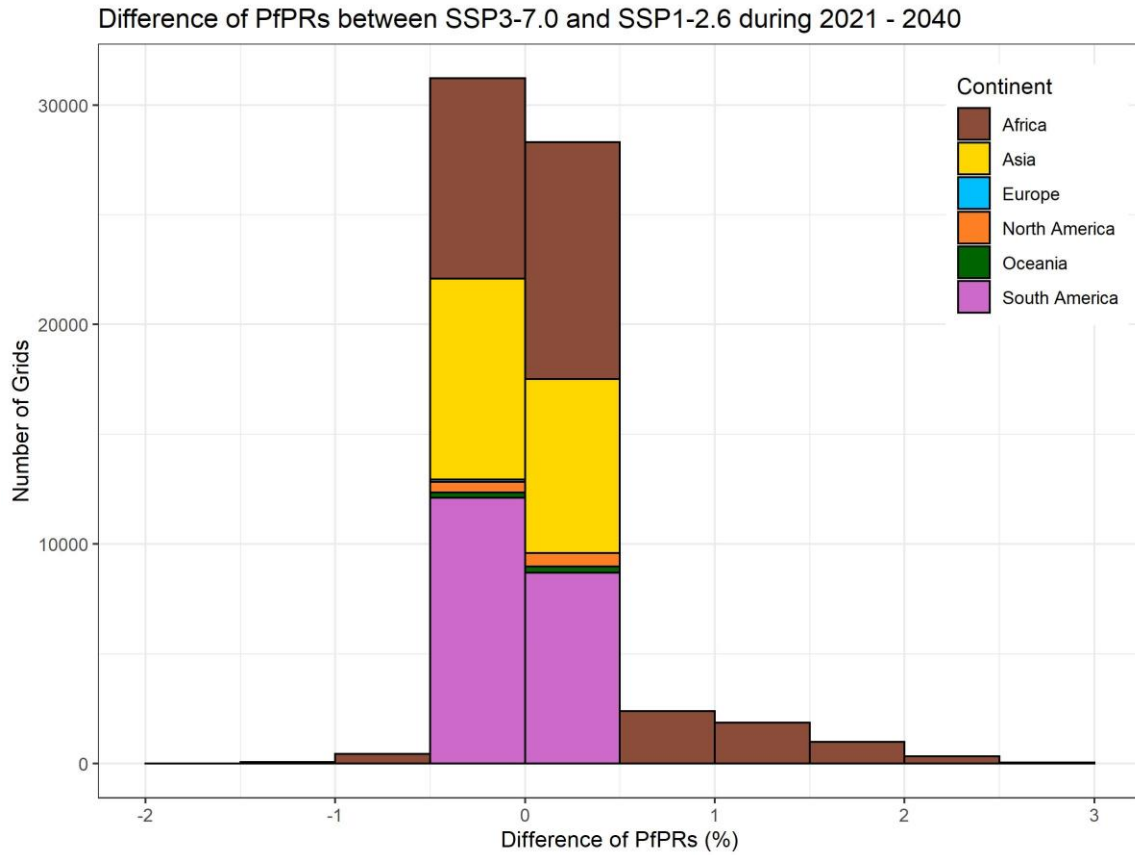
Figure S7: The Distribution of the Difference in PfPR$_{2\text{-}10}$ between SSP1-2.6 and

SSP3-7.0 (2021 - 2040)

Difference of PfPRs between SSP5-8.5 and SSP1-2.6 during 2021 - 2040

434
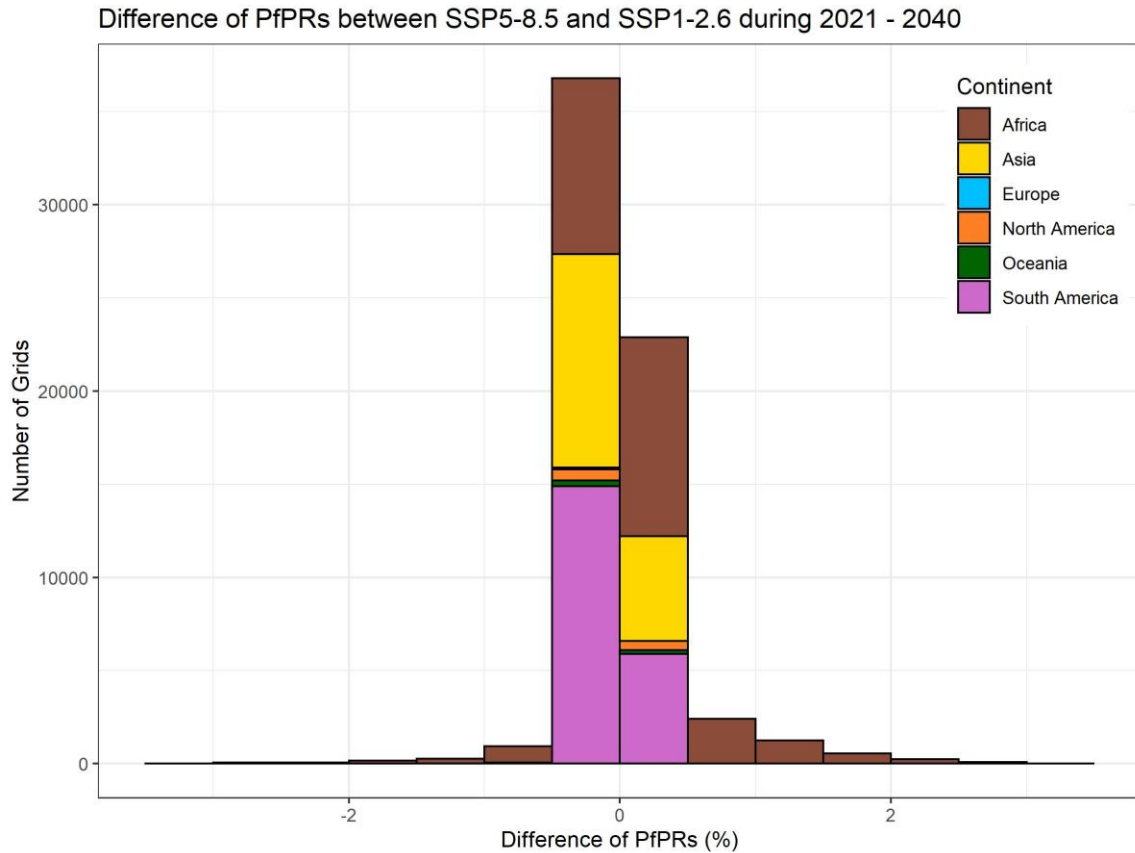
Figure S8: The Distribution of the Difference in PfPR$_{2\text{-}10}$ between SSP1-2.6 and

SSP5-8.5 (2021 - 2040)

435

436

437

438

439
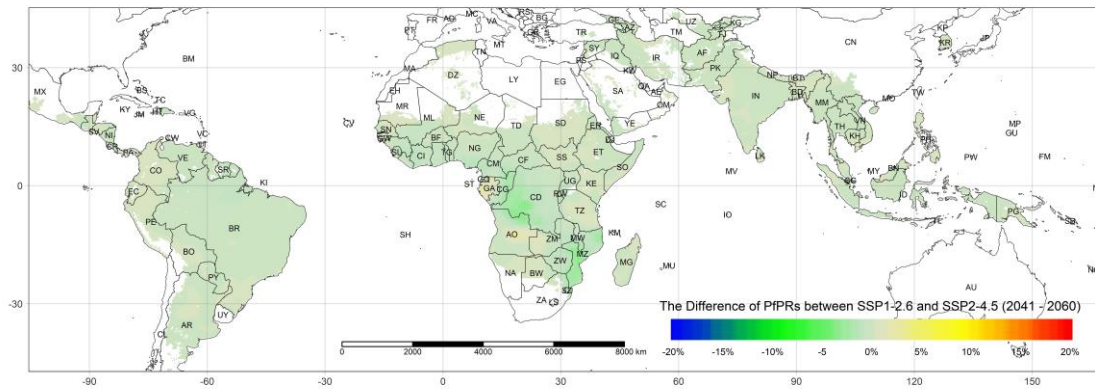
Figure S9: The Difference in PfPR$_{2\text{-}10}$ between SSP1-2.6 and SSP2-4.5 (2041 - 2060)

441
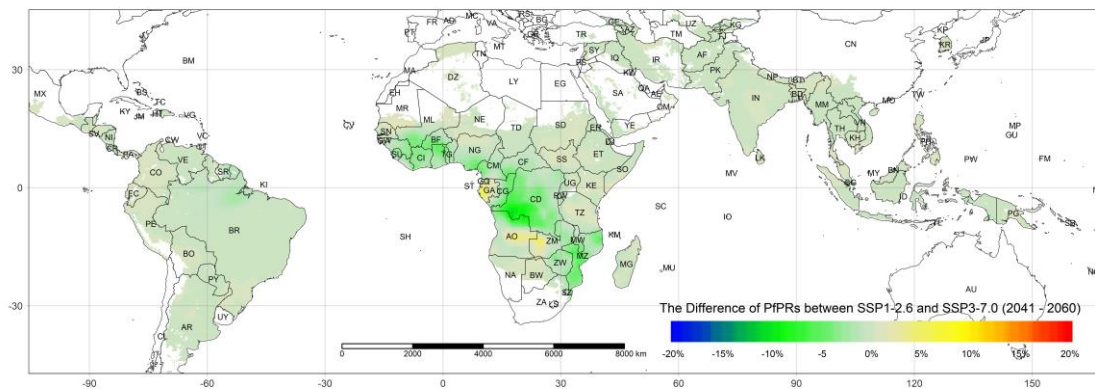
Figure S10: The Difference in PfPR2-10 between SSP1-2.6 and SSP3-7.0 (2041 -

2060)

444
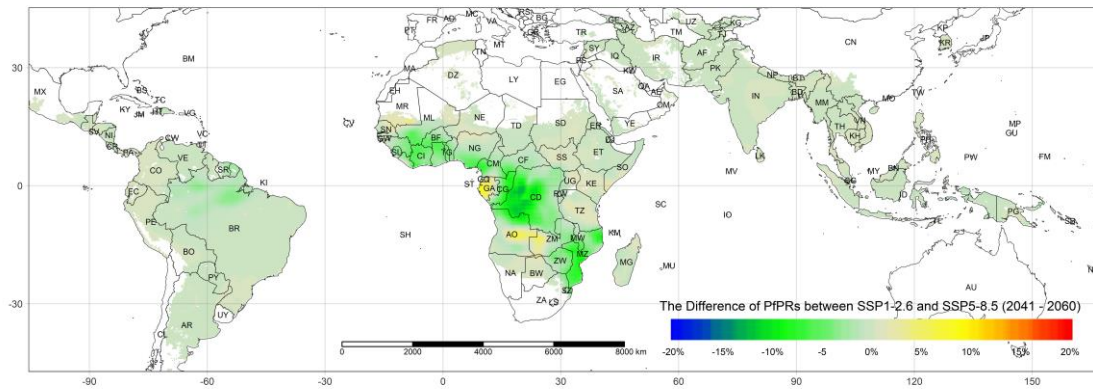
Figure S11: The Difference in PfPR2-10 between SSP1-2.6 and SSP5-8.5 (2041 -

2060)

447

448

449

Difference of PfPRs between SSP2-4.5 and SSP1-2.6 during 2041 - 2060

450
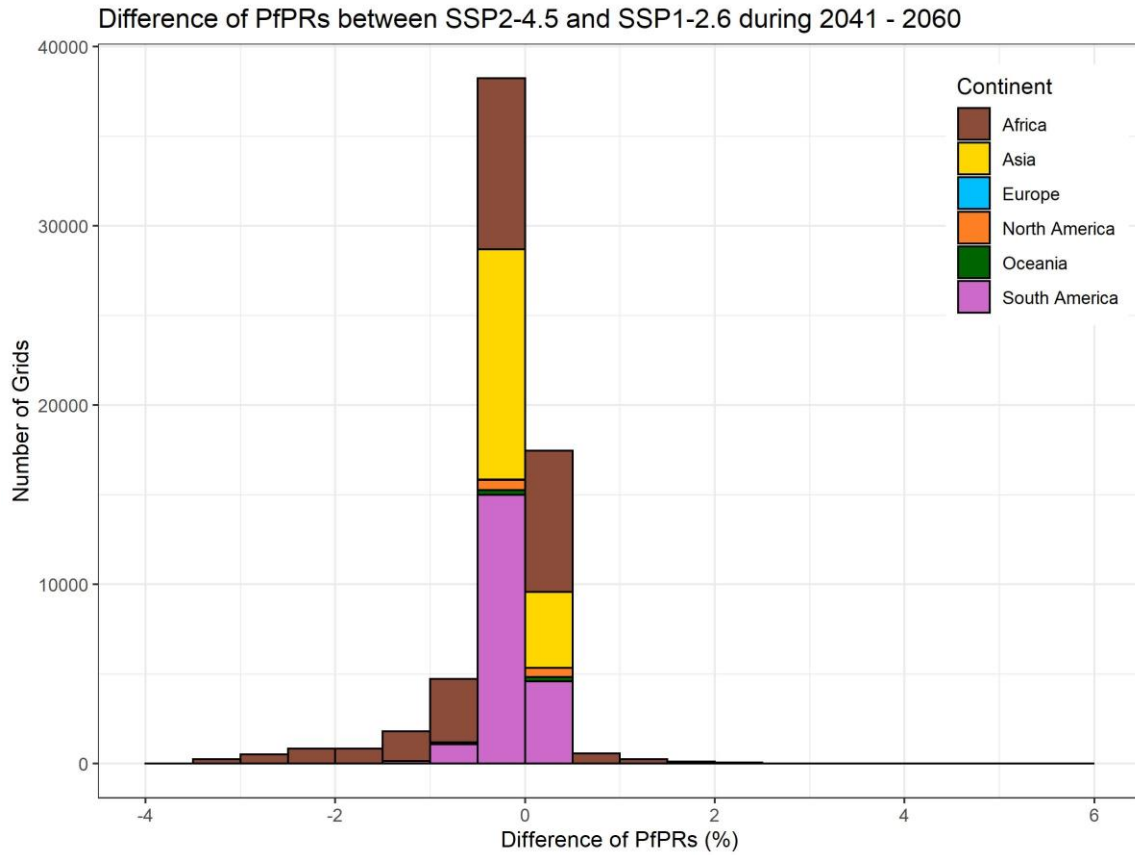
Figure S12: The Distribution of the Difference in PfPR2-10 between SSP1-2.6 and

451

SSP2-4.5 (2041 - 2060)

452

Difference of PfPRs between SSP3-7.0 and SSP1-2.6 during 2041 - 2060

453

Figure S13: The Distribution of the Difference in PfPR2-10 between SSP1-2.6 and

SSP3-7.0 (2041 - 2060)

454

455

Difference of PfPRs between SSP5-8.5 and SSP1-2.6 during 2041 - 2060
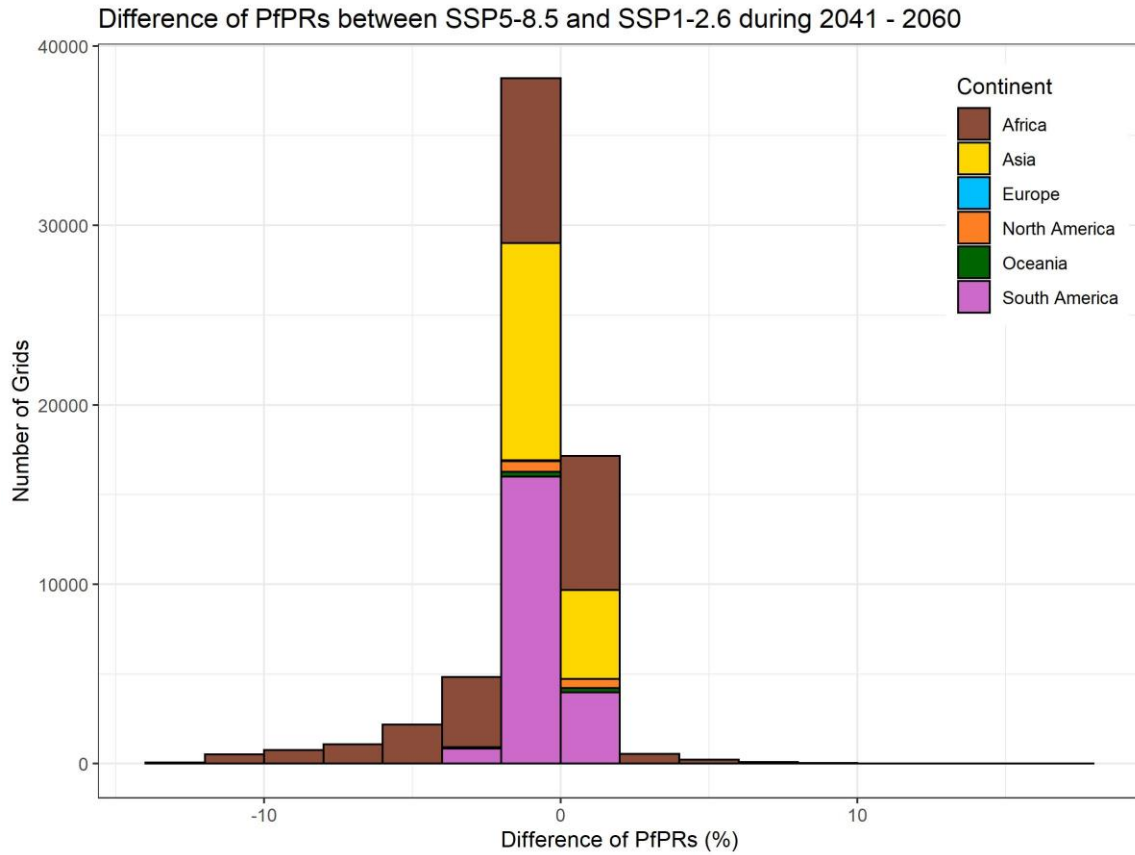
Figure S14: The Distribution of the Difference in PfPR2-10 between SSP1-2.6 and SSP5-8.5 (2041 - 2060)

456

457

458

459

460

461
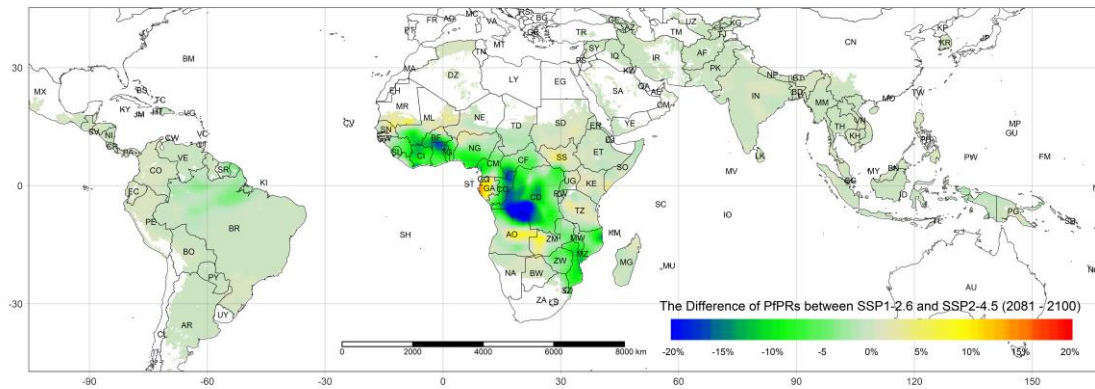
Figure S15: The Difference in PfPR2-10 between SSP1-2.6 and SSP2-4.5 (2081 -

2100)

464
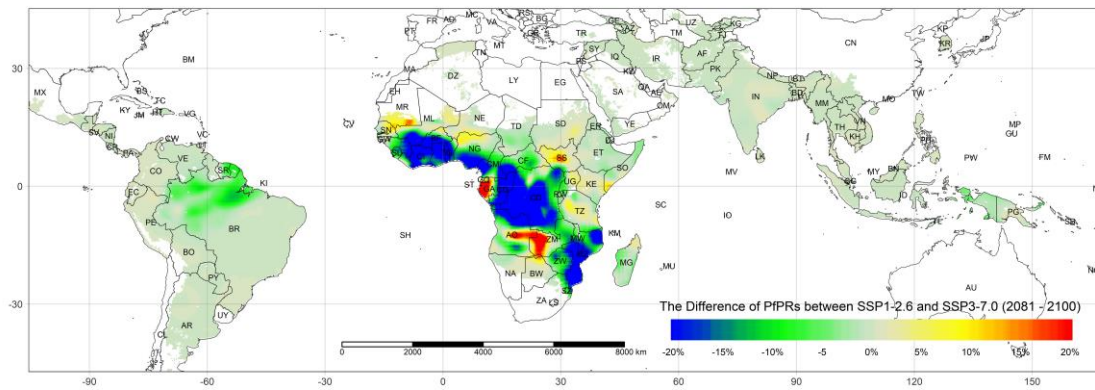
Figure S16: The Difference in PfPR2-10 between SSP1-2.6 and SSP3-7.0 (2081 -

2100)

467

Figure S17: The Difference in PfPR2-10 between SSP1-2.6 and SSP5-8.5 (2081 -

2100)

470

471

472

Difference of PfPRs between SSP2-4.5 and SSP1-2.6 during 2081 - 2100

473

Figure S18: The Distribution of the Difference in PfPR2-10 between SSP1-2.6 and

SSP2-4.5 (2081 - 2100)

Difference of PfPRs between SSP3-7.0 and SSP1-2.6 during 2081 - 2100

476
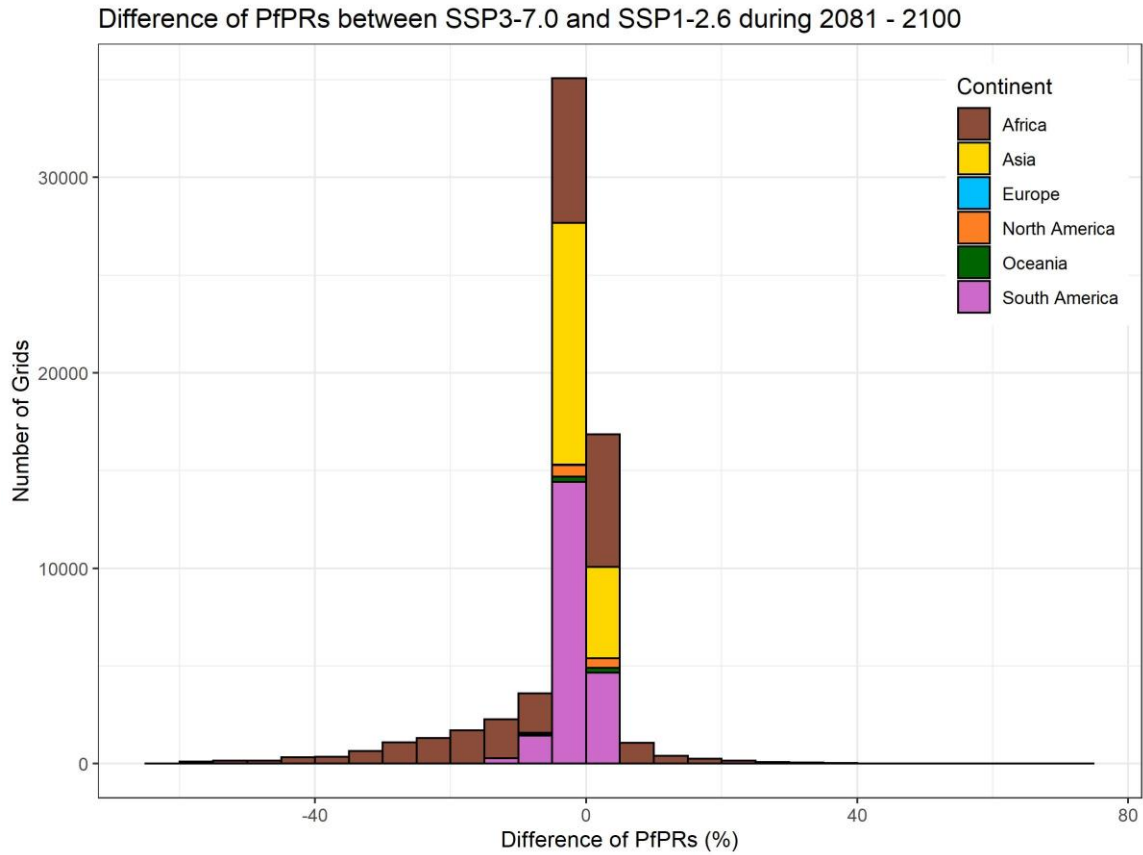
Figure S19: The Distribution of the Difference in PfPR2-10 between SSP1-2.6 and

SSP3-7.0 (2081 - 2100)

Difference of PfPRs between SSP5-8.5 and SSP1-2.6 during 2081 - 2100

479
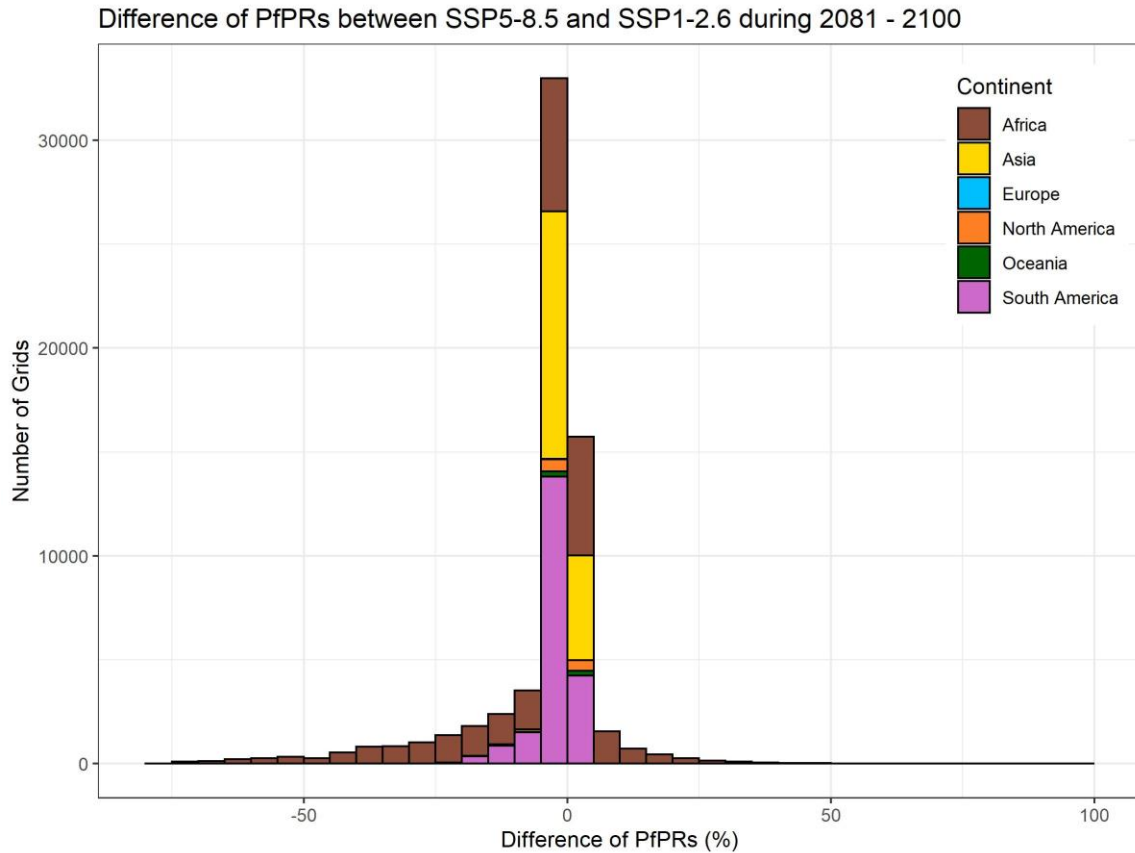
Figure S20: The Distribution of the Difference in PfPR2-10 between SSP1-2.6 and

SSP5-8.5 (2081 - 2100)

482
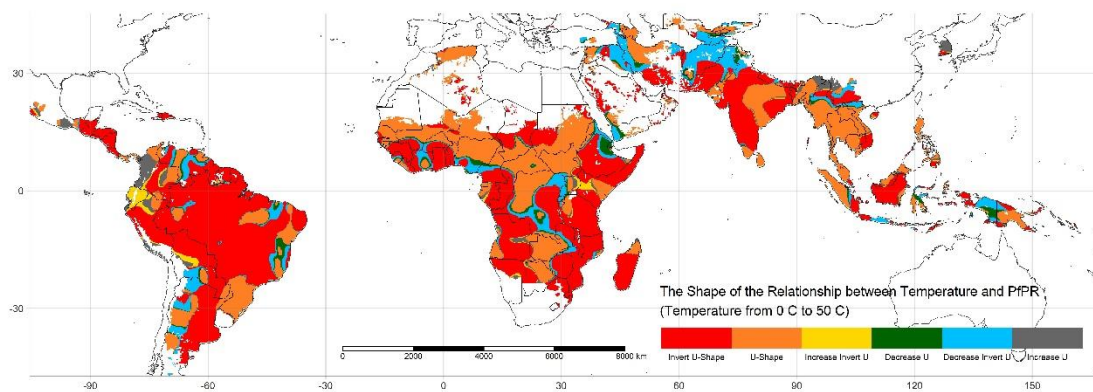
Figure S21: The Shape of Relationship between Temperature and PfPR2-10
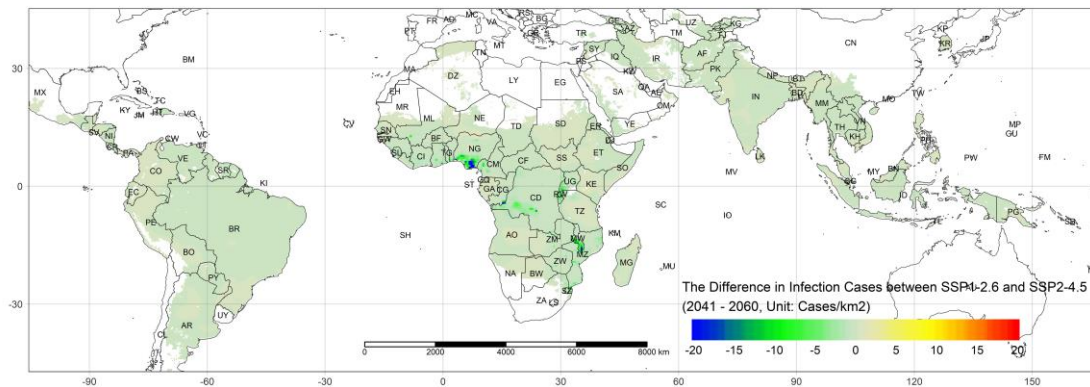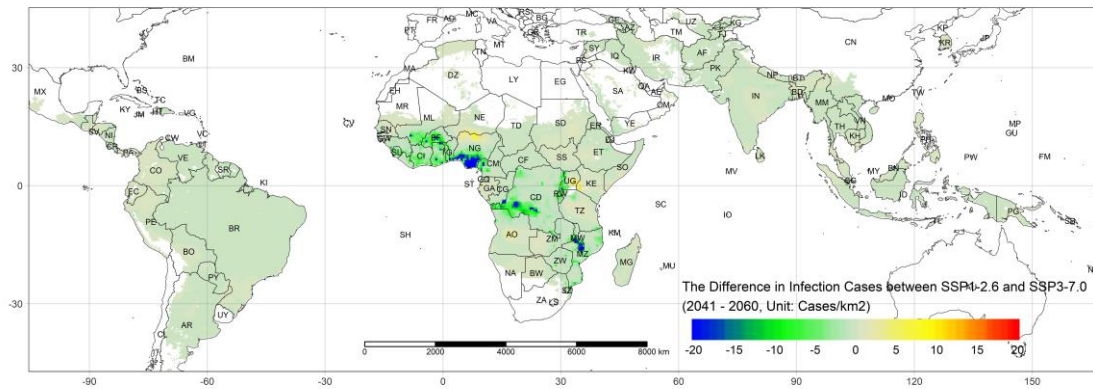
484
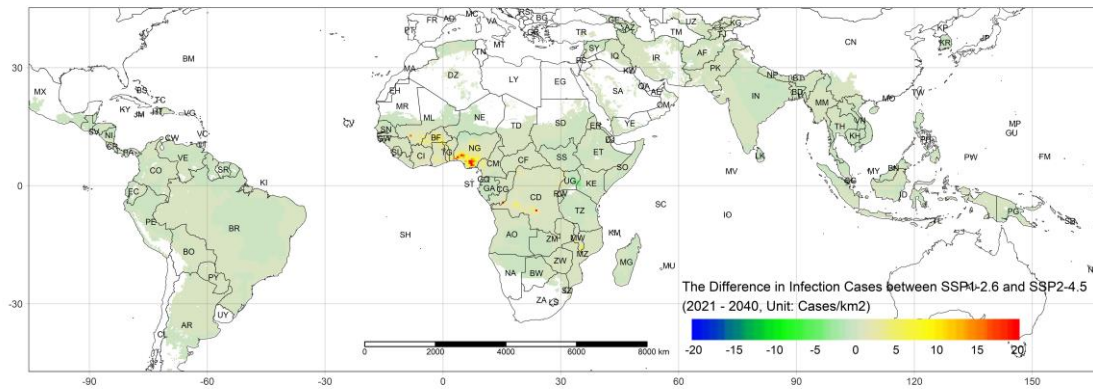
Figure S22: The Difference in Infection Cases between SSP1-2.6 and SSP2-4.5

during 2021 - 2040

487
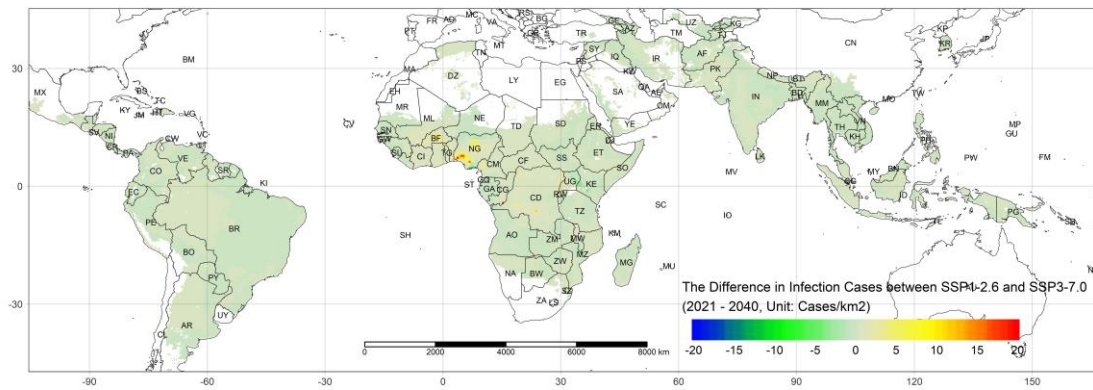
Figure S23: The Difference in Infection Cases between SSP1-2.6 and SSP3-7.0

during 2021 - 2040

490
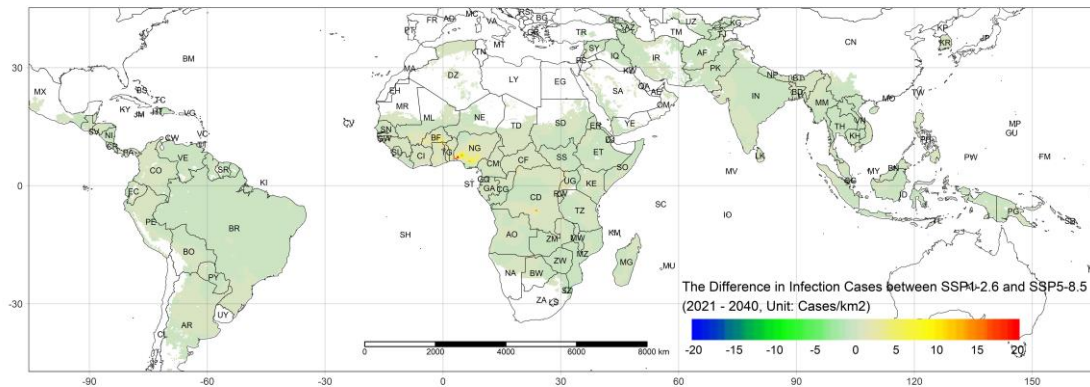
Figure S24: The Difference in Infection Cases between SSP1-2.6 and SSP5-8.5

during 2021 - 2040

493
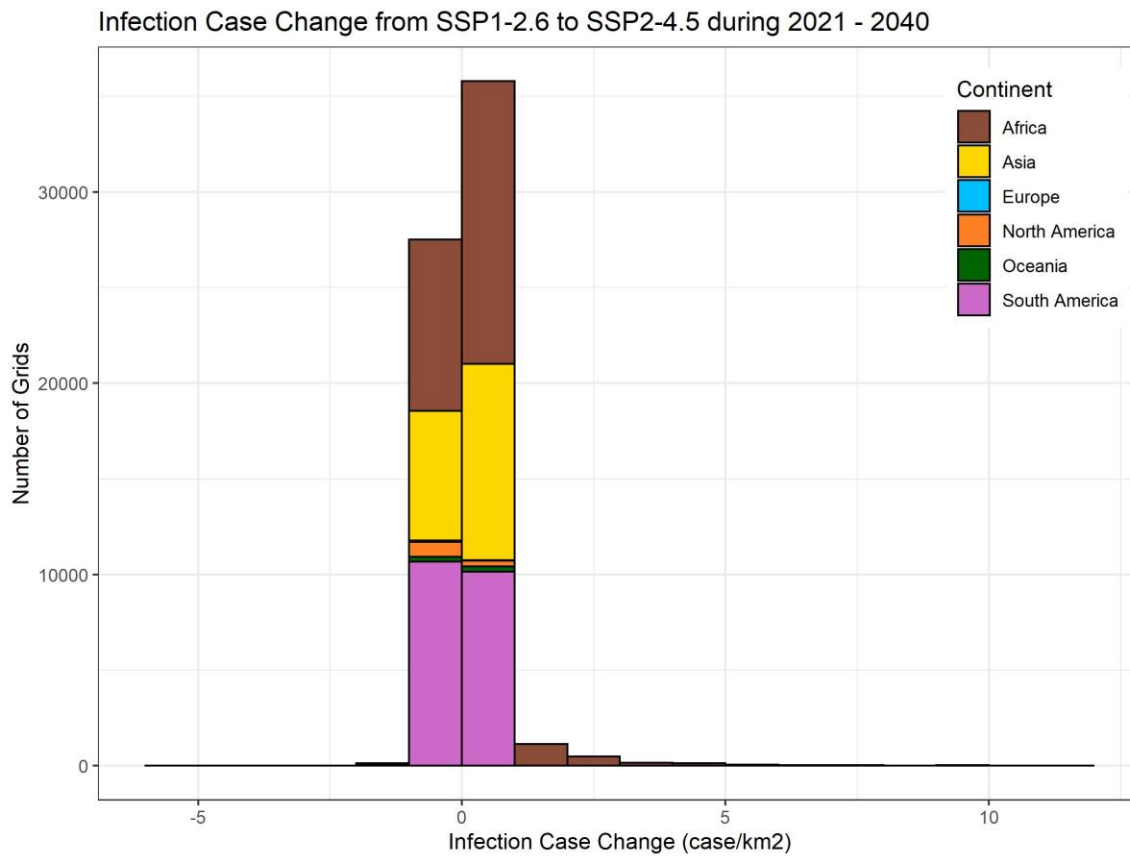
Figure S25: Infection Case Change from SSP1-2.6 to SSP2-4.5 during 2021 – 2040

495
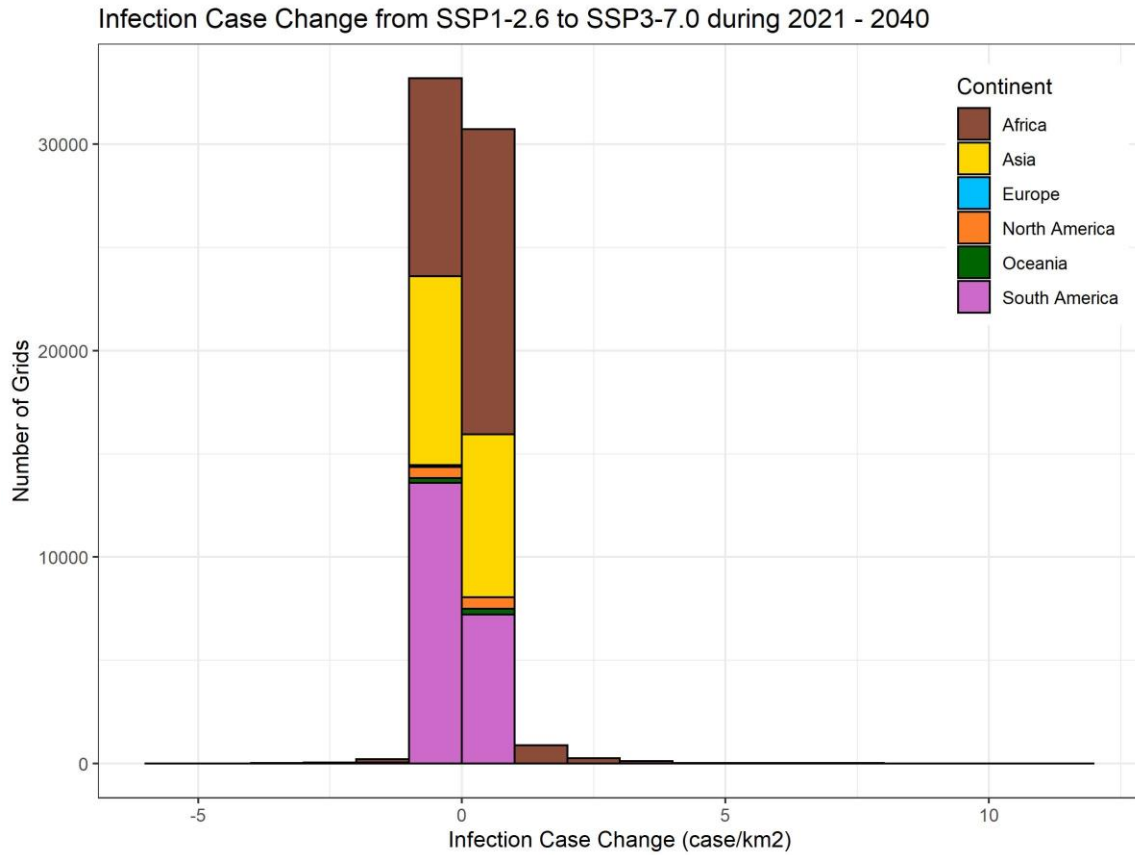
Figure S26: Infection Case Change from SSP1-2.6 to SSP3-7.0 during 2021 – 2040
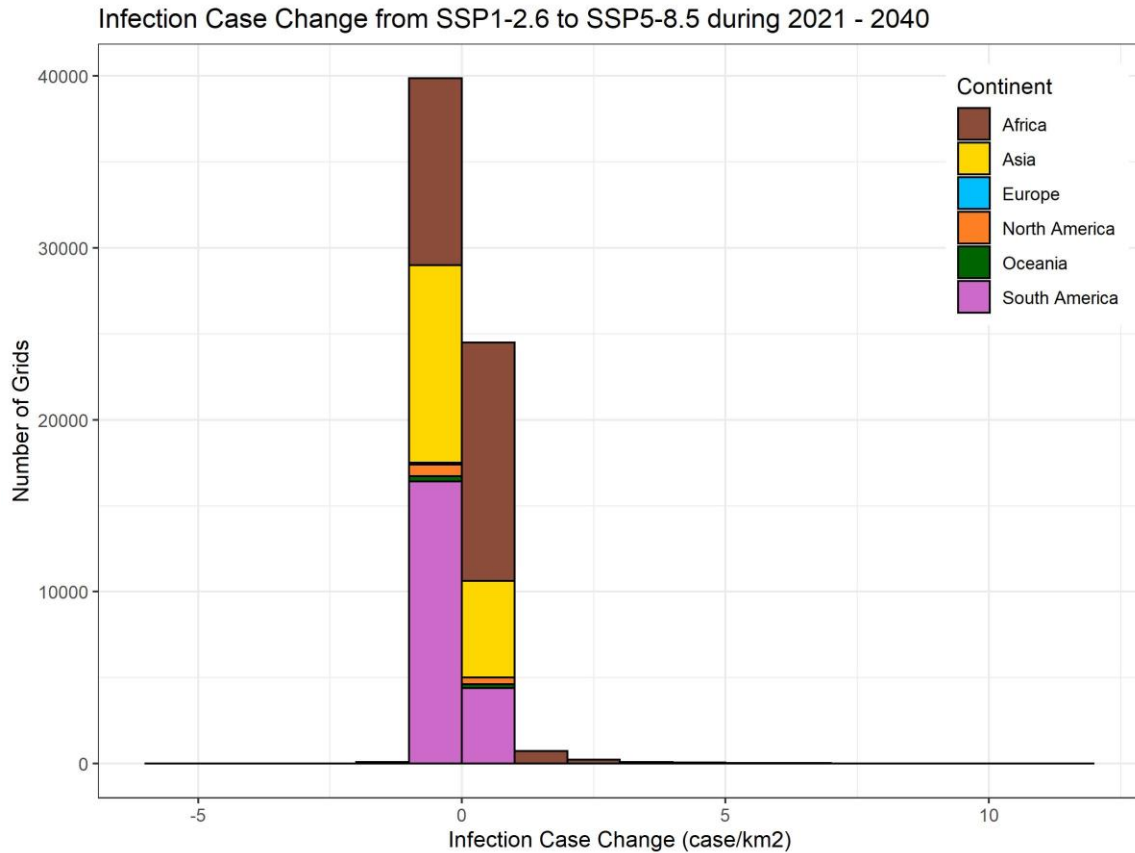
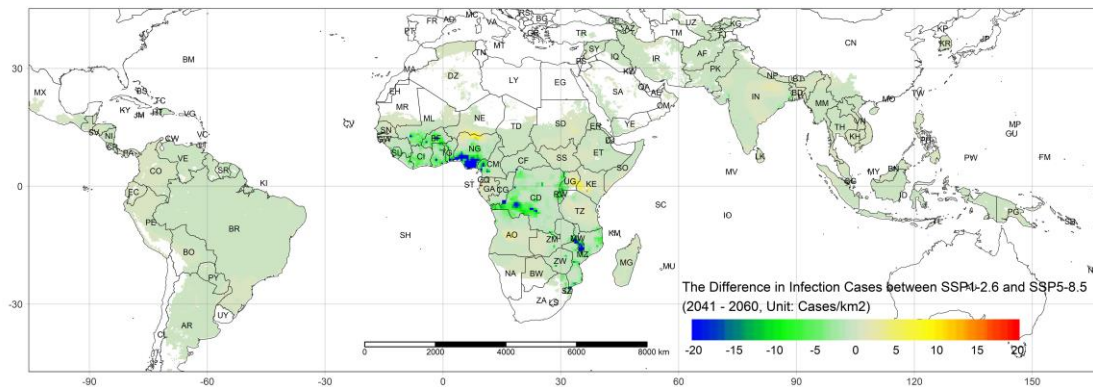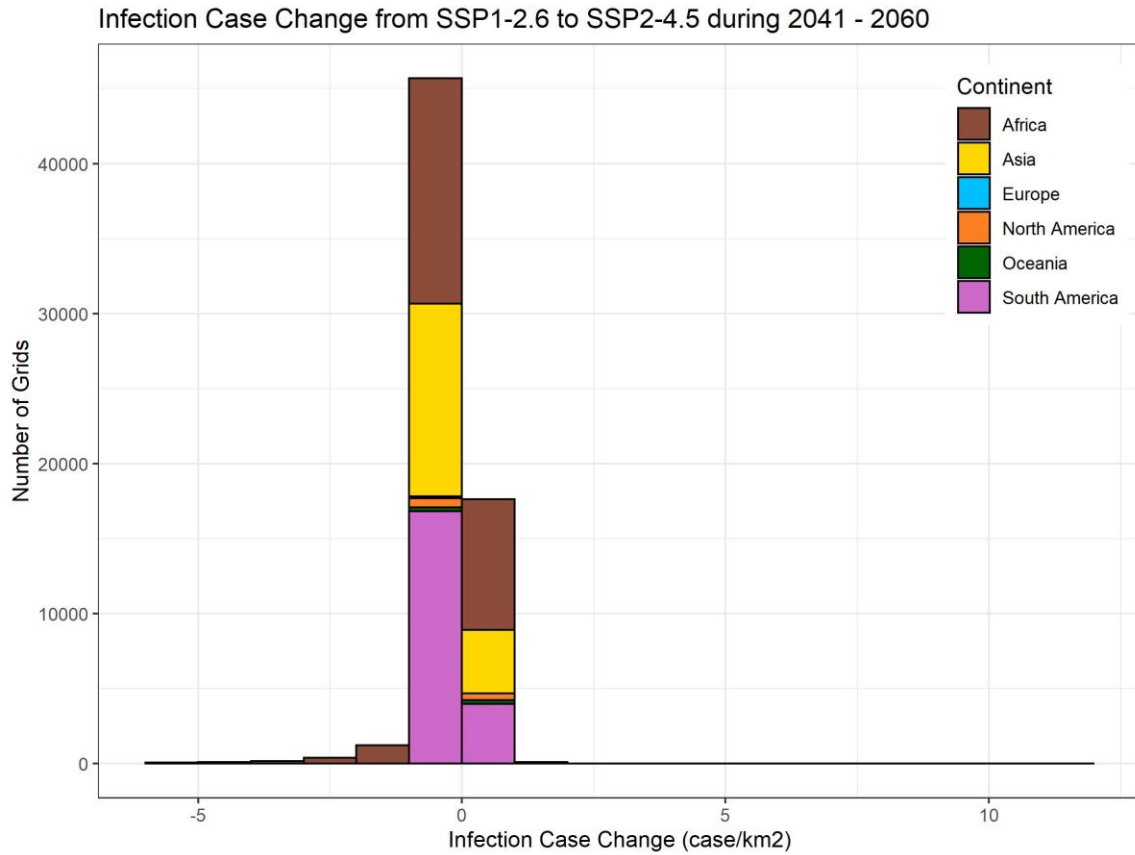Figure S27: Infection Case Change from SSP1-2.6 to SSP5-8.5 during 2021 – 2040

499

Figure S28: The Difference in Infection Cases between SSP1-2.6 and SSP2-4.5

during 2041 - 2060

502

Figure S29: The Difference in Infection Cases between SSP1-2.6 and SSP3-7.0

during 2041 - 2060

505

Figure S30: The Difference in Infection Cases between SSP1-2.6 and SSP5-8.5

during 2041 - 2060

Figure S31: Infection Case Change from SSP1-2.6 to SSP2-4.5 during 2041 – 2060

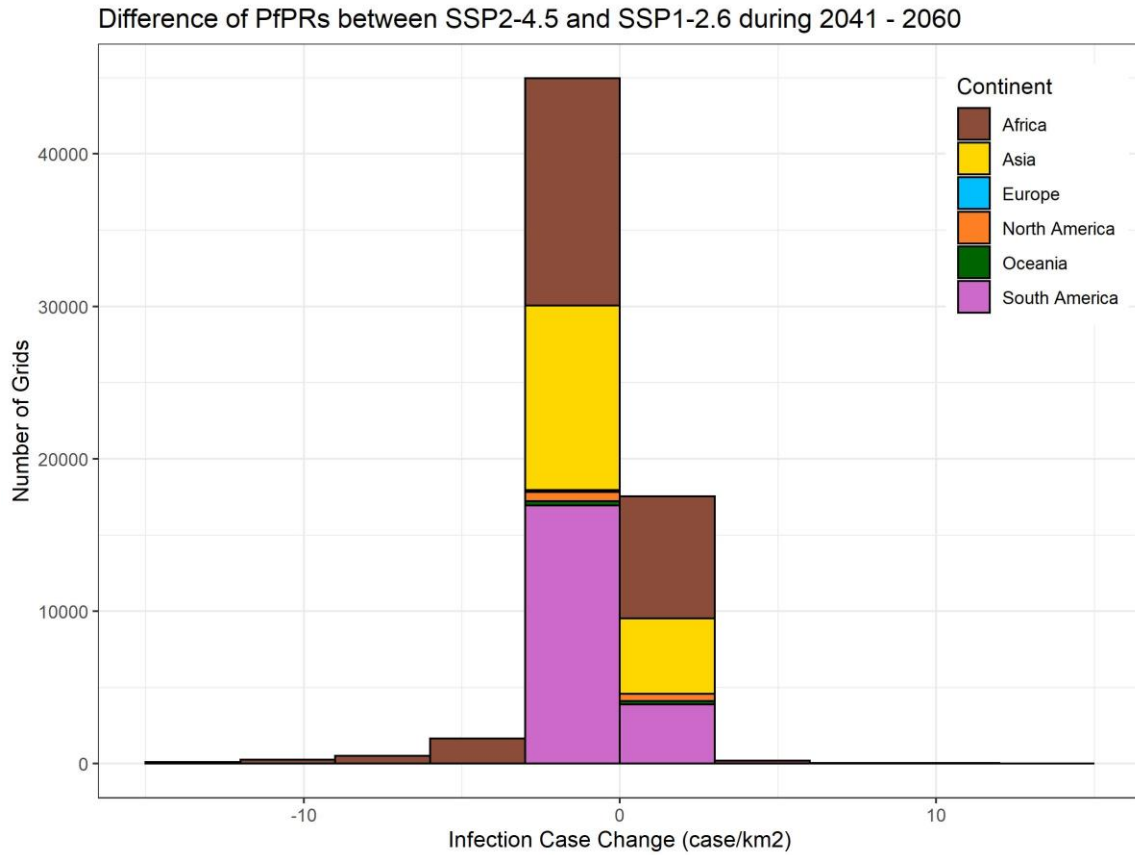Difference of PfPRs between SSP2-4.5 and SSP1-2.6 during 2041 - 2060

510

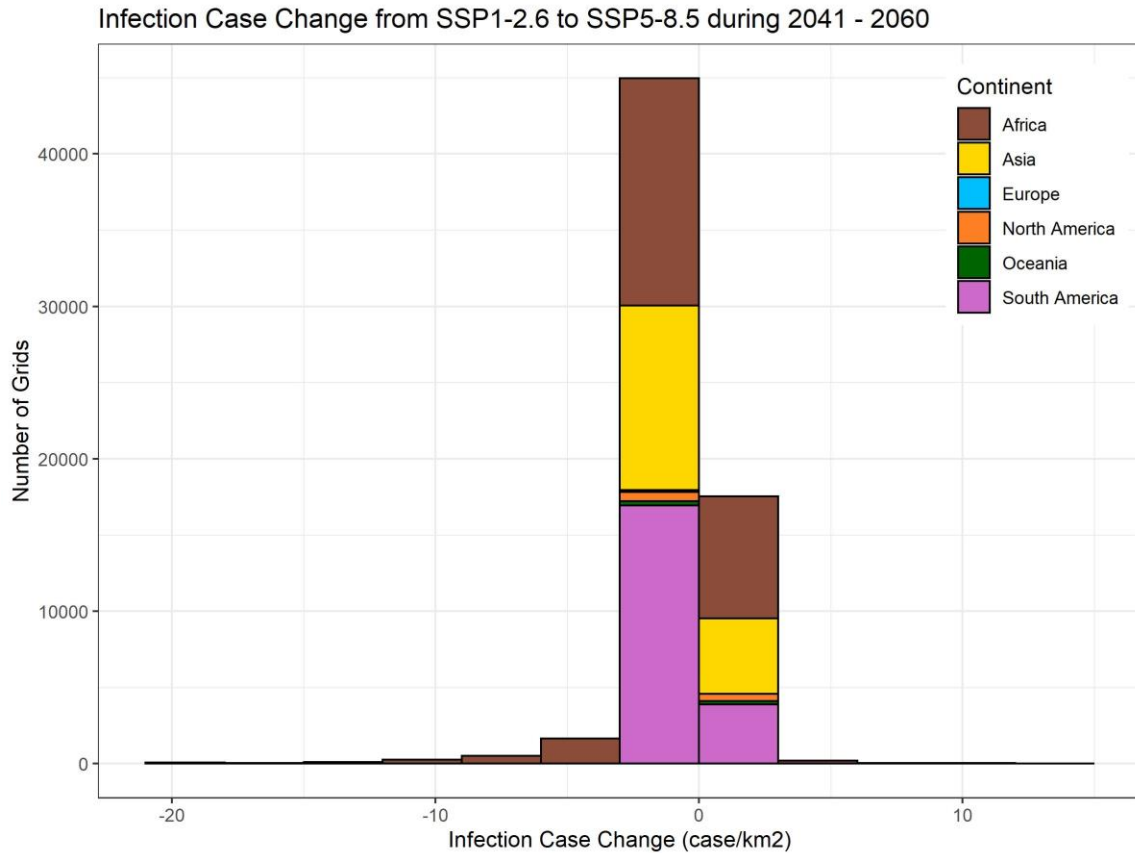511   **Figure S32: Infection Case Change from SSP1-2.6 to SSP3-7.0 during 2041 – 2060**

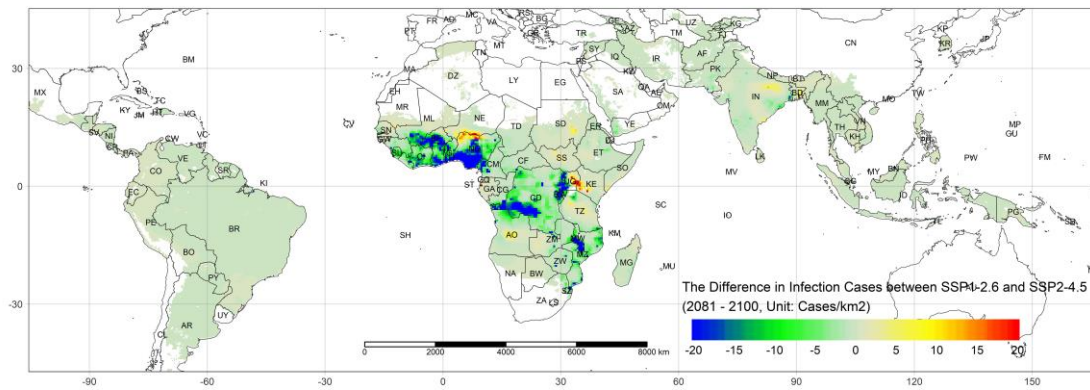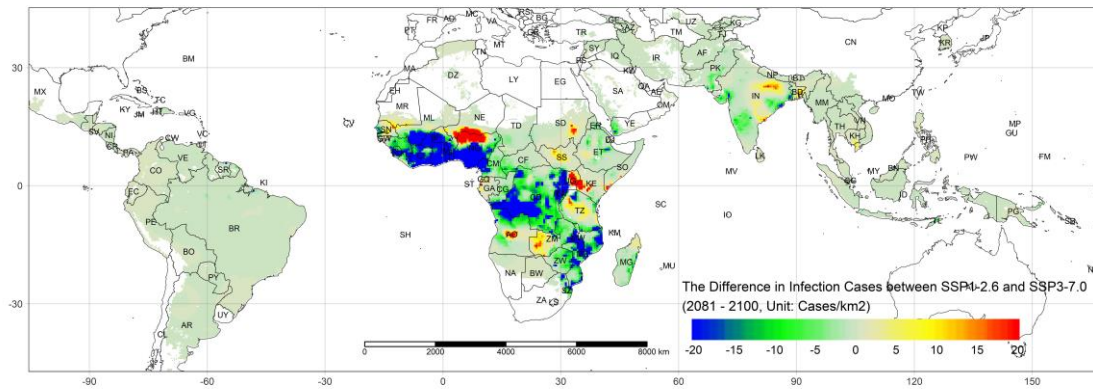Figure S33: Infection Case Change from SSP1-2.6 to SSP5-8.5 during 2041 – 2060
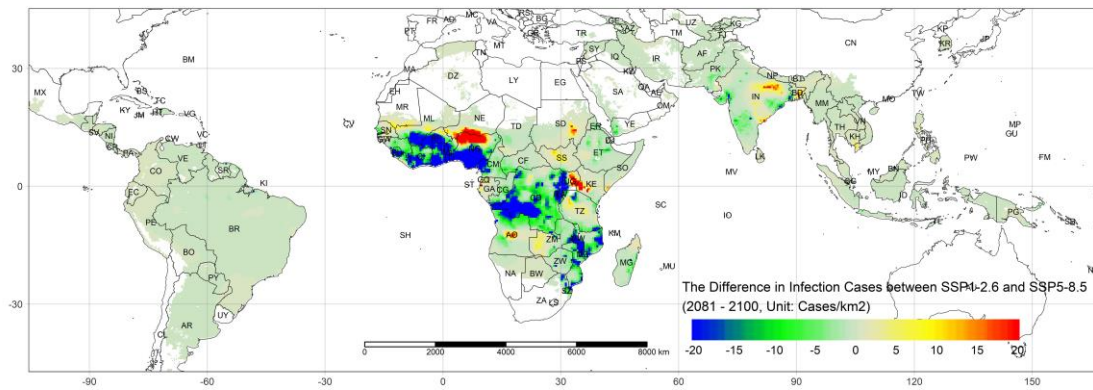

Figure S34: The Difference in Infection Cases between SSP1-2.6 and SSP2-4.5

during 2081 - 2100

Figure S35: The Difference in Infection Cases between SSP1-2.6 and SSP3-7.0

during 2081 - 2100

Figure S36: The Difference in Infection Cases between SSP1-2.6 and SSP5-8.5

during 2081 - 2100

Infection Case Change from SSP1-2.6 to SSP2-4.5 during 2081 - 2100
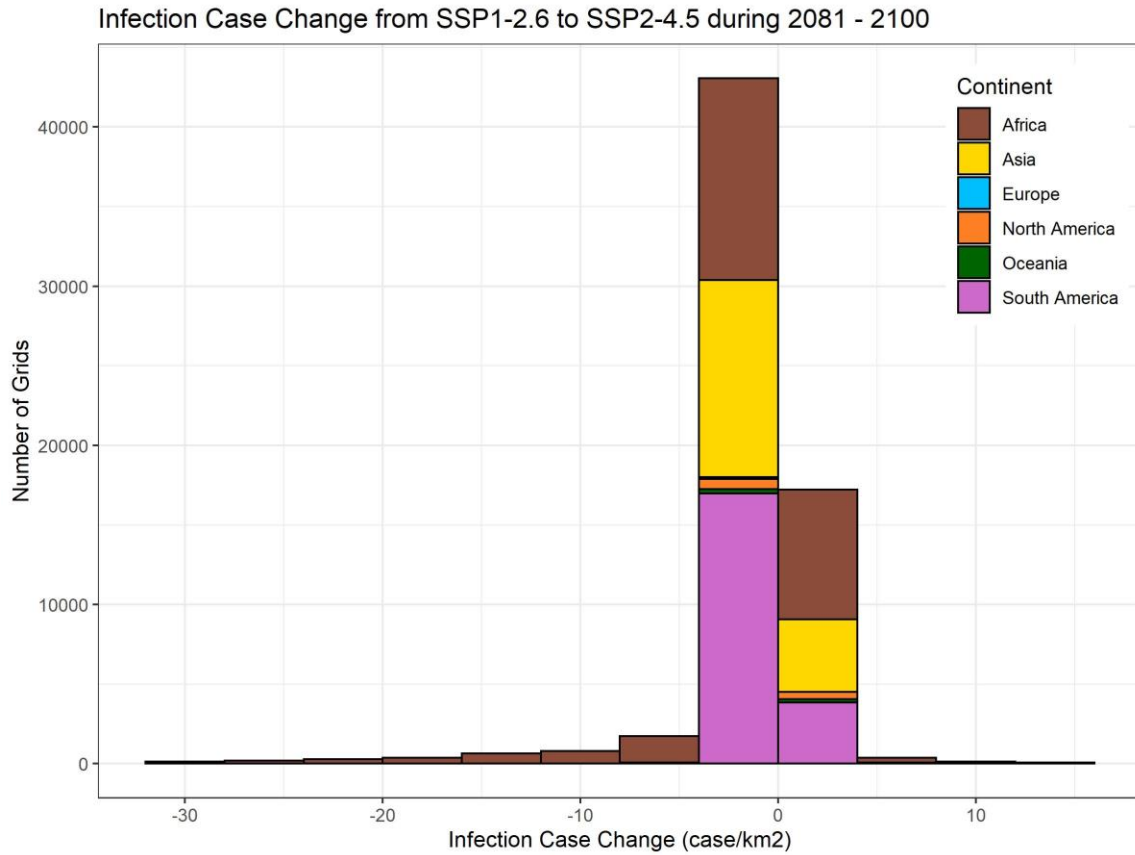
523

Figure S37: Infection Case Change from SSP1-2.6 to SSP2-4.5 during 2081 – 2100

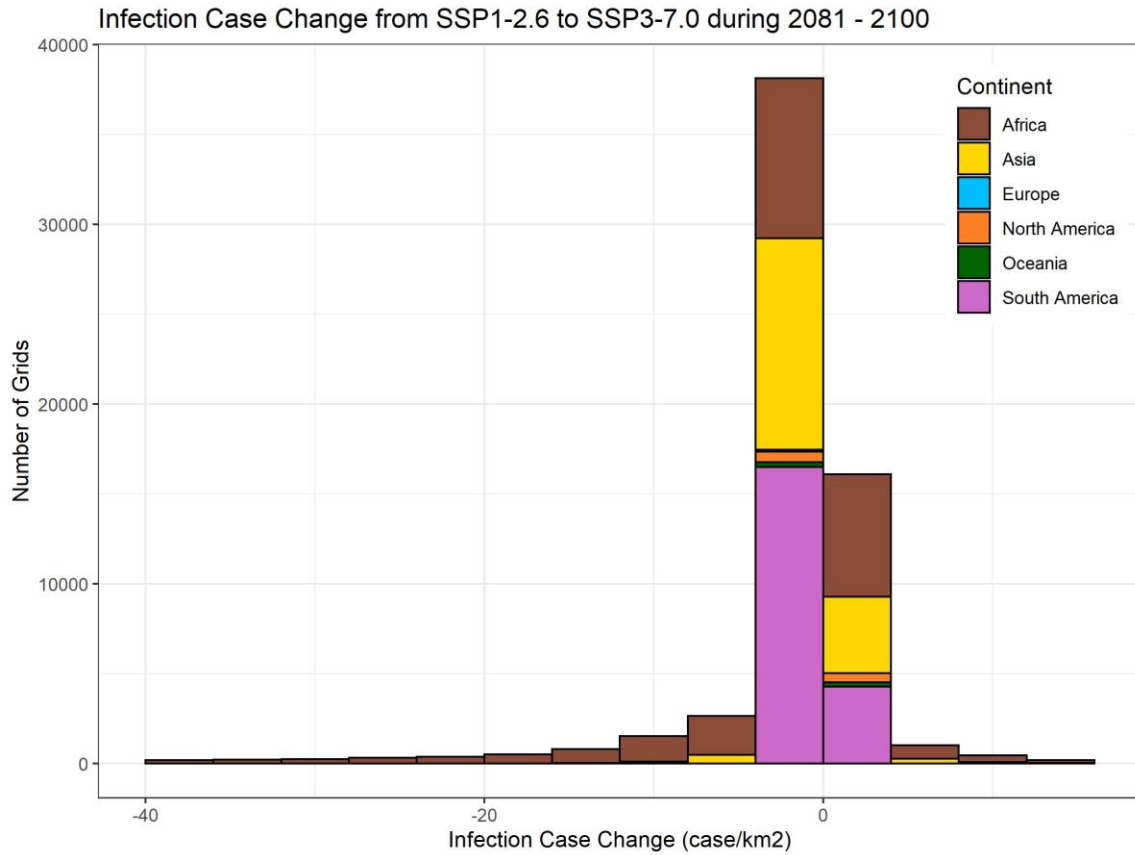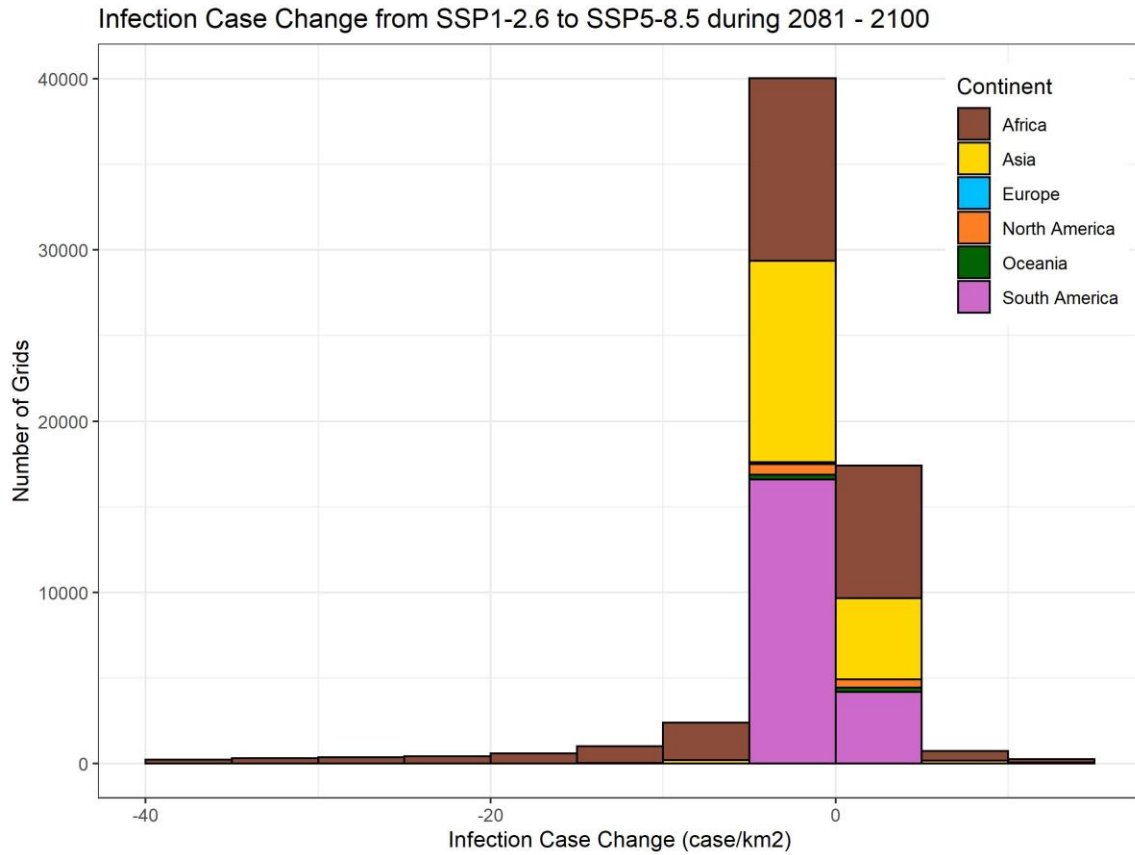**Figure S38: Infection Case Change from SSP1-2.6 to SSP3-7.0 during 2081 – 2100**

Infection Case Change from SSP1-2.6 to SSP5-8.5 during 2081 - 2100

527

Figure S39: Infection Case Change from SSP1-2.6 to SSP5-8.5 during 2081 – 2100

529

530

Table:

Table S1: Statistical Indicators Summary of 10-Fold Cross-Validation

	Statistical Indicator of Testing							Statistical Indicator of Training						
	N	R^2	RMSE(%)	MAE(%)	α	β	r	N	R^2	RMSE(%)	MAE(%)	α	β	r
1	1158867	94.75%	4.45%	1.94%	0.006	0.944	0.973	128763	94.67%	4.49%	1.95%	0.006	0.944	0.973
2	1158867	94.74%	4.45%	1.94%	0.006	0.944	0.973	128763	94.78%	4.44%	1.93%	0.006	0.943	0.974
3	1158867	94.75%	4.45%	1.94%	0.006	0.943	0.973	128763	94.72%	4.44%	1.94%	0.006	0.944	0.973
4	1158867	94.75%	4.45%	1.94%	0.006	0.944	0.973	128763	94.76%	4.46%	1.94%	0.006	0.944	0.973
5	1158867	94.75%	4.45%	1.94%	0.006	0.943	0.973	128763	94.69%	4.46%	1.94%	0.006	0.945	0.973
6	1158867	94.74%	4.45%	1.94%	0.006	0.943	0.973	128763	94.75%	4.45%	1.94%	0.006	0.945	0.973
7	1158867	94.75%	4.45%	1.94%	0.006	0.944	0.973	128763	94.65%	4.48%	1.94%	0.006	0.942	0.973
8	1158867	94.74%	4.45%	1.94%	0.006	0.944	0.973	128763	94.82%	4.45%	1.95%	0.006	0.943	0.974
9	1158867	94.75%	4.45%	1.94%	0.006	0.943	0.973	128763	94.73%	4.46%	1.93%	0.006	0.945	0.973
10	1158867	94.74%	4.45%	1.94%	0.006	0.944	0.973	128763	94.75%	4.45%	1.94%	0.006	0.942	0.973

Note: The Ideal value of β is 1, and the ideal value of α is 0. When the values of statistical indicators are close between training results and testing result, the predictions of model are reliable.

Reference:

1. Bhatt S, *et al.* The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature* **526**, 207-211 (2015).

2. Rodell M, *et al.* The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society* **85**, 381-394 (2004).

3. Mordecai EA, Ryan SJ, Caldwell JM, Shah MM, Labeaud AD. Climate change could shift disease burden from malaria to arboviruses in Africa. *The Lancet Planetary Health* **4**, e416-e423 (2020).

4. Mordecai EA, *et al.* Optimal temperature for malaria transmission is dramatically lower than previously predicted. *Ecology Letters* **16**, 22-30 (2013).

5. Mordecai EA, *et al.* Thermal biology of mosquito-borne disease. *Ecology Letters* **22**, 1690-1708 (2019).

6. Johnson LR, *et al.* Understanding uncertainty in temperature effects on vector-borne disease: a Bayesian approach. *Ecology* **96**, 203-213 (2015).

7. Lourenço PM, Sousa CA, Seixas J, Lopes P, Novo MT, Almeida APG. Anopheles atroparvus density modeling using MODIS NDVI in a former malarious area in Portugal. *Journal of Vector Ecology* **36**, 279-291 (2011).

8. Reiner RC, Geary M, Atkinson PM, Smith DL, Gething PW. Seasonality of Plasmodium falciparum transmission: a systematic review. *Malaria Journal* **14**, (2015).

9. Brugueras S, *et al.* Environmental drivers, climate change and emergent diseases transmitted by mosquitoes and their vectors in southern Europe: A systematic review. *Environ Res* **191**, 110038 (2020).

10. Amadi JA, *et al.* Sensitivity of vegetation to climate variability and its implications for malaria risk in Baringo, Kenya. *PLoS One* **13**, e0199357 (2018).

11. Midekisa A, Senay G, Henebry GM, Semuniguse P, Wimberly MC. Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia. *Malaria Journal* **11**, 165 (2012).

571

572   12.   Didan K, Munoz AB, Solano R, Huete A, University of Arizona: Vegetation Index
573   Phenology Lab. MODIS vegetation index user's guide (MOD13 series).) (2015).

574

575   13.   Sorichetta A, Hornby GM, Stevens FR, Gaughan AE, Linard C, Tatem AJ. High-
576   resolution gridded population datasets for Latin America and the Caribbean in 2010,
577   2015, and 2020. *Scientific Data* **2**, 150045 (2015).

578

579   14.   Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ. Population Distribution,
580   Settlement Patterns and Accessibility across Africa in 2010. *PLoS One* **7**, e31743
581   (2012).

582

583   15.   World Health Organization. World malaria report 2021.) (2021).

584

585   16.   IPCC. Climate change 2022: impacts, adaptation and vulnerability.) (2022).

586

587   17.   IPCC. Climate change 2014 synthesis report.). IPCC (2014).

588

589   18.   Rogers DJ, Randolph SE. Climate Change and Vector-Borne Diseases. In:
590   *Advances in Parasitology* (eds Hay SI, Graham A, Rogers DJ). Academic Press
591   (2006).

592

593   19.   Altizer S, Ostfeld Richard S, Johnson Pieter TJ, Kutz S, Harvell CD. Climate
594   Change and Infectious Diseases: From Evidence to a Predictive Framework.
595   *Science* **341**, 514-519 (2013).

596

597   20.   Buchwald AG, Hayden MH, Dadzie SK, Paull SH, Carlton EJ. Aedes-borne disease
598   outbreaks in West Africa: A call for enhanced surveillance. *Acta Tropica* **209**,
599   105468 (2020).

600

601   21.   Nyaruaba R, Mwaliko C, Mwau M, Mousa S, Wei H. Arboviruses in the East
602   African Community partner states: a review of medically important mosquito-
603   borne Arboviruses. *Pathogens and Global Health* **113**, 209-228 (2019).

604

605   22.   Brunsdon C, Fotheringham S, Charlton M. Geographically Weighted Regression.
606   *Journal of the Royal Statistical Society: Series D (The Statistician)* **47**, 431-443
607   (1998).

608

609   23.   Fotheringham A, Brunsdon C, Charlton M. *Geographically Weighted Regression:
610   The Analysis of Spatially Varying Relationships*. John Wiley & Sons (2002).

611
612 24. Croissant Y, Millo G. Panel Data Econometrics in R: The plm Package. *J Stat Softw*
613 **27**, (2008).

614
615 25. Breusch TS, Pagan AR. The Lagrange Multiplier Test and its Applications to
616 Model Specification in Econometrics. *The Review of Economic Studies* **47**, 239
617 (1980).

618
619 26. Kang S. A note on the equivalence of specification tests in the two-factor
620 multivariate variance components model. *Journal of Econometrics* **28**, 193-203
621 (1985).

622
623 27. Giesen C*, et al.* The impact of climate change on mosquito-borne diseases in Africa.
624 *Pathogens and Global Health* **114**, 287-301 (2020).

625
626 28. Kulkarni MA, Duguay C, Ost K. Charting the evidence for climate change impacts
627 on the global spread of malaria and dengue and adaptive responses: a scoping
628 review of reviews. *Globalization and Health* **18**, (2022).

629
630 29. Brunsdon C, Fotheringham AS, Charlton ME. Geographically Weighted
631 Regression: A Method for Exploring Spatial Nonstationarity. *Geographical*
632 *Analysis* **28**, 281-298 (2010).

633
634 30. Fotheringham AS, Oshan TM. Geographically weighted regression and
635 multicollinearity: dispelling the myth. *Journal of Geographical Systems* **18**, 303-
636 329 (2016).

637
638 31. Beenstock M, Felsenstein D. *The econometric analysis of non-stationary spatial*
639 *panel data*. Springer (2019).

640
641 32. Gollini I, Lu B, Charlton M, Brunsdon C, Harris P. GWmodel: An R package for
642 exploring spatial heterogeneity using geographically weighted models. *J Stat Softw*
643 **63**, (2015).

644

645