

1 Probabilistic seasonal dengue 2 forecasting in Vietnam using 3 superensembles

4 **Felipe J Colón-González**^{1,6,9,10*§}, **Leonardo Soares Bastos**^{1,6,8}, **Barbara Hofmann**²,
5 **Alison Hopkin**², **Quillon Harpham**², **Tom Crocker**³, **Rosanna Amato**³, **Iacopo**
6 **Ferrario**², **Francesca Moschini**², **Samuel James**², **Sajni Malde**², **Eleanor Ainscoe**²,
7 **Vu Sinh Nam**⁴, **Dang Quang Tan**⁵, **Nguyen Duc Khoa**⁵, **Mark Harrison**³, **Gina**
8 **Tsarouchi**², **Darren Lumbroso**², **Oliver Brady**^{1†}, **Rachel Lowe**^{1,6,7†}

*For correspondence:

felipe.colon@lshtm.ac.uk (Felipe J Colón-González)

†These authors contributed equally to this work

§Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Public Health, London School of Hygiene & Tropical Medicine, London, UK

1 Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, United Kingdom; ²HR Wallingford, United Kingdom; ³Met Office, United Kingdom; ⁴National Institute of Hygiene and Epidemiology, Vietnam; ⁵General Department of Preventive Medicine, Vietnam; ⁶Centre on Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London, UK; ⁷Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain; ⁸Scientific Computing Program, Oswaldo Cruz Foundation (Fiocruz), Rio de Janeiro, Brazil; ⁹School of Environmental Sciences, University of East Anglia, Norwich, UK; ¹⁰Tyndall Centre for Climate Change Research, University of East Anglia, Norwich, UK

19 **Abstract** Timely information is key for decision-making. The ability to predict dengue
20 transmission ahead of time would significantly benefit planners and decision-makers. Dengue is
21 climate-sensitive. Monitoring climate variability could provide advance warning about dengue risk.
22 Multiple dengue early warning systems have been proposed. Often, these systems are based on
23 deterministic models that have limitations for quantifying the probability that a public health event
24 may occur. We introduce an operational seasonal dengue forecasting system where Earth
25 observations and seasonal climate forecasts are used to drive a superensemble of probabilistic
26 dengue models to predict dengue risk up to six months ahead. We demonstrate that the system
27 has skill and relative economic value at multiple forecast horizons, seasons, and locations. The
28 superensemble generated, on average, more accurate forecasts than those obtained from the
29 models used to create it. We argue our system provides a useful tool for the development and
30 deployment of targeted vector control interventions, and a more efficient allocation of resources in
31 Vietnam.

33 Introduction

34 Dengue is a mosquito-transmitted viral infection spread by *Aedes* mosquitoes in urban and peri-
35 urban environments in tropical and subtropical countries (*Powell and Tabachnick, 2013; Li et al.,*
36 *2014; Kraemer et al., 2015*). About half of the global population is at risk of dengue transmission
37 (*Brady et al., 2012; Bhatt et al., 2013*). There is no specific antiviral treatment for dengue, and
38 vaccination is restricted to seropositive individuals (*World Health Organization, 2018*). Dengue
39 prevention relies on mosquito control measures which are primarily insecticide-based (*Dusfour*

40 *et al., 2019*). The increasing resistance to insecticides highlights the need for targeted and effective
41 interventions (*Moyes et al., 2017*).

42 Vietnam is particularly affected by dengue with an estimated burden of about two million yearly
43 infections (*Bhatt et al., 2013; Shepard et al., 2016*), although, on average, only 95,000 cases have
44 been reported annually to the Ministry of Health over the period 2002–2020. The economic impact
45 of dengue in Vietnam is estimated to be USD 30–95 million per annum (*Shepard et al., 2014, 2016;*
46 *Hung et al., 2018*). Dengue in Vietnam is primarily spread by *Ae. aegypti* and, to a lesser extent, by
47 *Ae. albopictus* (*Tsunoda et al., 2014; Thi et al., 2017*).

48 In Vietnam, dengue is characterized by strong seasonality and substantial inter-annual and
49 spatial variability (Supplementary Figure 1). Dengue exhibits different behaviour in different parts
50 of the country. In the north, where temperatures are lower than in the rest of the country, most
51 provinces have few or no cases. An exception is Hanoi, which has reported, on average, about
52 8,700 dengue cases per year over the past ten years. In central and southern provinces where
53 temperatures are warmer, many provinces report thousands of cases annually albeit with large
54 interannual variation.

55 Dengue control in the country primarily involves community engagement and mobilisation
56 to reduce breeding sites, and outdoor low-volume insecticide spraying in the vicinity of reported
57 dengue cases to kill adult mosquitoes (*Cuong et al., 2013*). One limitation of dengue control
58 measures is that they are essentially reactive, meaning they take place after cases have occurred.
59 This situation hampers the ability of public health professionals to reduce the magnitude and
60 severity of outbreaks. Dengue surveillance is mostly passive, relying on clinical cases reported
61 by patients seeking healthcare (*Cuong et al., 2013*). The diagnosis and reporting of dengue cases
62 typically suffers delays which vary across time and space (*Bastos et al., 2019*). This situation hinders
63 the timely generation and communication of information on when and where transmission occurs,
64 limiting the ability of health professionals to plan and execute control measures.

65 If accurate predictions are available, public health decision-makers and planners could design,
66 implement and target interventions to the most at-risk places in a timely fashion. Disease mod-
67 els driven by Earth observations have been valuable for predicting dengue risk ahead of time,
68 supporting decision-making in multiple settings (*Lowe et al., 2016, 2017; Colón-González et al.,*
69 *2018a*). Climate variation is one of the main drivers of dengue ecology. Temperature, for example,
70 regulates the development, biting, and reproduction rates, and the spatial distribution of *Aedes*
71 mosquitoes (*Gage et al., 2008; Reinhold et al., 2018; Kraemer et al., 2019*). The concentration and
72 replication of dengue viruses within mosquitoes is also temperature dependent (*Watts et al., 1987;*
73 *Gage et al., 2008*). Rising temperatures increase dengue transmission to an optimum range of
74 26–29°C (*Mordecai et al., 2017*). Large diurnal temperature ranges (> 20°C) reduce transmission
75 and increase mosquito mortality (*Lambrechts et al., 2011*). Precipitation modulates the creation
76 or flush away of mosquito breeding sites (*Stewart Ibarra et al., 2013*). Rising humidity increases
77 dengue risk (*Colón-González et al., 2017*) as relative humidity levels of at least 50–55% prolong
78 mosquito survival (*Simon-Oke and Olofintoye, 2015*). Wind speed reduces the biting activity of
79 mosquitoes reducing dengue risk (*Sedda et al., 2018*). Delayed effects of climate on dengue have
80 been demonstrated and correspond to climatic influences on dengue through their effect on the
81 life cycle of both mosquitoes and the dengue virus (*Naish et al., 2014*). Consequently, the climatic
82 conditions during the low-dengue season may be indicative of dengue incidence in the following
83 high-dengue season (*Lauer et al., 2018; Lowe et al., 2018*).

84 Multiple studies have highlighted the potential usefulness of seasonal-climate-driven epidemi-
85 ological surveillance for decision-making and planning (*Lowe et al., 2014; Lauer et al., 2018; Tomp-*
86 *kins and Di Giuseppe, 2015; Tompkins et al., 2019*). These studies have used subseasonal (i.e.
87 between two weeks and two months ahead) forecasts to inform disease models, and compute
88 predictions of dengue risk. There has been limited progress in using subseasonal-to-seasonal
89 climate forecasting to compute prospective forecasts on a routine basis. There are several chal-
90 lenges for implementing operational and sustainable subseasonal (henceforth seasonal) early

91 warning systems (*Thomson et al., 2014*). Some of these challenges include the lack of multi-decadal
92 health data sets with which to train and validate seasonal-climate-driven early warning systems,
93 the common mismatch of scales between climate data outputs and data used for decision-making,
94 and a general lack of consensus as to how to communicate uncertainties to users (*Tompkins et al.,*
95 *2019*).

96 Dengue early warning systems driven by Earth observations and seasonal climate forecasts
97 have been proposed using a range of modelling approaches (*Yamana et al., 2016*), including auto-
98 regressive integrative moving average (ARIMA) (*Eastin et al., 2014*), deterministic regression (*Hii*
99 *et al., 2013; Lauer et al., 2018*), spatio-temporal Bayesian hierarchical models (*Lowe et al., 2016,*
100 *2017*), least absolute shrinkage and selection operator (LASSO) regression (*Chen et al., 2018*), and
101 machine learning (*Stolerman et al., 2019*). Often, models are validated using block cross-validation
102 to select the model specification with the lowest out-of-sample predictive error (*Lowe et al., 2016;*
103 *Lauer et al., 2018*). This approach takes advantage of all available data to make repeated out-
104 of-sample model predictions, which increases the robustness of model adequacy and skill score
105 statistics. One drawback of this method is that it does not preserve the time ordering of the data.
106 Also, predictions are computed for some time periods using a model trained on data from a later
107 time period (*Bergmeir and Benítez, 2012*).

108 Previous dengue risk prediction studies have relied on outputs from one or two competing
109 models (e.g. *Lowe et al., 2016; Lauer et al., 2018*). However, combining forecasts from multiple
110 competing models into a superensemble can result in more accurate predictions than those from
111 any individual model (see, for example, *Yamana et al., 2016; Johansson et al., 2019*). The use of
112 model superensembles for the development of dengue early warning systems has been seldom
113 explored. Moreover, predictions are typically made for a selected year or month (*Lowe et al., 2014,*
114 *2016; Lauer et al., 2018*) rather than for a series of lead times into the future, or a whole season
115 (*Lowe et al., 2017*). In some cases, systems are designed exclusively for research purposes in
116 isolation from relevant stakeholders who may become potential users.

117 Typically, dengue early warning systems are based on deterministic models (*Hii et al., 2013;*
118 *Eastin et al., 2014; Lauer et al., 2018*) which may under-represent heterogeneity and stochastic
119 cessation of transmission. However, decision-makers are increasingly interested in understand-
120 ing the uncertainties related to the models used to develop decision-support tools, and in the
121 probabilities that an event of public health concern may or may not take place (*Lowe et al., 2014;*
122 *Colón-González et al., 2018b; Lake et al., 2019*). Spatio-temporal probabilistic models have the
123 advantage of being able to quantify the probability that an event (e.g. an outbreak) may occur at
124 specific times and for specific locations. Public health officials may be more inclined to take action if
125 the probability of observing an outbreak exceeds a certain value (*Lowe et al., 2016*). Both modellers
126 and decision-makers should pay attention to, and agree on the definition of outbreaks thresholds
127 so that model predictions are a useful guide for planning and decision-making.

128 In several countries, including Vietnam, outbreaks are defined using a so called *endemic channel*
129 (*Badurdeen et al., 2013; Runge Ranzinger et al., 2014*) which corresponds to the mean number
130 of cases per month or season over a long-term period (*Brady et al., 2015*). In Vietnam, endemic
131 channels are defined for each province using the last five years of dengue surveillance data.
132 Outbreak years are removed from the computation of the endemic threshold. When dengue cases
133 exceed the mean plus two standard deviations, an *outbreak* is declared. One limitation of this
134 approach is that often outbreak years are removed arbitrarily or quasi-quantitatively to increase the
135 sensitivity of the outbreak threshold (*Brady et al., 2015*). In areas where dengue incidence is typically
136 low (e.g. < 10 cases per month), the endemic channel may be frequently exceeded generating
137 statistical alarms of little public health importance (*Noufaily et al., 2019*). Despite these limitations,
138 endemic channels are widely used for dengue control decision-making in a variety of countries
139 and provide a practical decision point around which forecasts can be targeted (*Hussain-Alkhateeb*
140 *et al., 2018; Olliaro et al., 2018*).

141 Here, we introduce a superensemble of probabilistic spatio-temporal hierarchical dengue models

142 driven by Earth observations and seasonal climate forecasts. The model framework was co-
143 designed with stakeholders from the World Health Organization, the United Nations Development
144 Programme, the Vietnamese Ministry of Health, the Pasteur Institute Ho Chi Minh City, the Pasteur
145 Institute Nha Trang, the Institute of Hygiene and Epidemiology Tay Nguyen (TIHE), and the National
146 Institute of Hygiene and Epidemiology (NIHE). The system is designed to generate monthly estimates
147 of dengue risk across Vietnam (331,210 km²) at the province level ($n = 63$) in *near-real time* (i.e. current
148 time minus processing time). The superensemble is used in an expanding window time series
149 cross-validation framework (*Hyndman and Athanasopoulos, 2014*) to generate probabilistic dengue
150 forecasts, which allow us to calculate the probability of exceeding pre-defined dengue outbreak
151 thresholds for a forecast horizon (i.e. lead time) of one to six months.

152 The superensemble constitutes the dengue fever component of a forecasting system called
153 *D-MOSS* (i.e. Dengue forecasting MOdel Satellite-based System). The system operates using a suite
154 of Earth observation data sources from satellites, and has its first implementation in Vietnam.
155 Vietnam is divided into 63 provinces which are subdivided into 713 districts and has an estimated
156 population of 95.5 million people. The D-MOSS system produces results at the province level and
157 is accompanied by a skill assessment that is applied consistently across the whole of Vietnam.
158 The intention is to give an overall evaluation of the performance of our method, rather than an
159 assessment that pertains to the characteristics of any particular province.

160 Results

161 We fitted a total of 14,592 different models (128 unique model specifications across 114 forecast
162 months, i.e. January 2007 to December 2016). Dengue data was modelled using Earth observations
163 of mean monthly values of minimum temperature, maximum temperature, precipitation amount
164 per day, specific humidity, diurnal temperature range, wind speed, and sea surface temperature
165 anomalies for the Niño 3.4 Region (i.e. 5°S–5°N and 170°–120°W) as explanatory variables. Each
166 variable was included in isolation, as well as with all possible combinations (see Methods and
167 Materials). Models included the proportion of people living in urban, periurban and rural areas
168 (henceforth land-cover data). Land-cover data were included in all models as they change annually
169 and not monthly (see Methods and Materials). Predictions up to six months ahead were made at
170 each time step starting in January 2007 using seasonal climate hindcasts (retrospective forecasts)
171 data from the UK Met Office Global Seasonal Forecasting System version 5 (*MacLachlan et al., 2015*;
172 *Scaife et al., 2014*).

173 Model superensemble

174 Models were evaluated against five verification metrics: continuous rank probability score (CRPS),
175 root mean squared error (RMSE), mean absolute error (MAE), deviance information criterion (DIC),
176 and Watanabe-Akaike information criterion (WAIC). We selected the best two models within each
177 verification metric to create a model superensemble. Verification metrics were computed for each
178 model specification at each time step and averaged across the whole time series. For all verification
179 metrics, lower values indicate a smaller difference between the forecasts and the observations.
180 The five models with the lowest values for the selected verification metrics (i.e. best performing)
181 were selected to generate a superensemble (Table 1). Combining all best performing models into a
182 model superensemble led to a lower CRPS and MAE than any of the five competing models. The
183 RMSE of the superensemble, however, was higher than two of the five models.

184 When stratified by forecast horizon time, the predictive ability of the superensemble was greater
185 than or comparable to that of the best performing competing models (Figure 1). Across all metrics,
186 the predictive ability of the superensemble deteriorated as the forecast horizon increased from one
187 to six months ahead. This situation is also evident where the forecast ensemble mean and its 95%
188 credible interval (i.e. the interval in the domain of the posterior probability distribution) are plotted
189 against the observed number of dengue cases (Supplementary Figure 2). Notice that the accuracy
190 of the predictions worsens as the forecast horizon expands. We noted that the credible intervals

Table 1. Verification metrics and seasonal climate predictors of the model superensemble, and the best performing models.

Model	CRPS	RMSE	MAE	DIC	WAIC	Seasonal climate predictors
Superensemble	78.7	110.9	80.9	-	-	Models 1:5
1	79.8	112.0	81.6	45866.7	46077.6	SH, WS, DTR, SST
2	80.2	110.2	81.1	45865.1	46076.4	SH, DTR
3	82.1	110.6	82.6	45862.1	46071.6	WS
4	112.0	146.3	115.7	45854.9	46065.7	Tmin, Tmax
5	115.9	154.0	120.2	45855.6	46066.2	Tmin, DTR, SST

Table 1-source data 1. Tmin: mean monthly minimum temperature (°C), Tmax: mean monthly maximum temperature (°C), DTR: diurnal temperature range (°C), SH: specific humidity (dimensionless), WS: wind speed (m s^{-1}), SST: sea surface temperature anomalies in the Niño region 3.4 (°C). CRPS: continuous ranked probability score, RMSE: root mean squared error, MAE: mean absolute error, DIC: deviance information criterion, WAIC: Watanabe–Akaike information criterion.

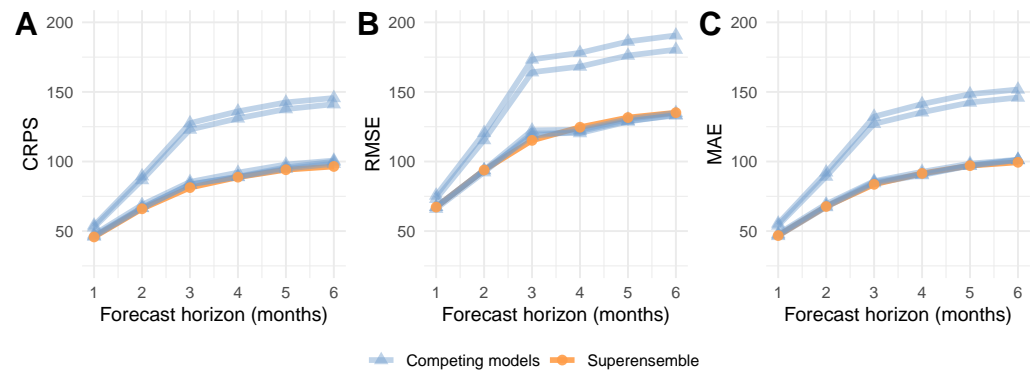


Figure 1. Forecasting metrics of five competing models (blue) compared against a model superensemble (orange). Metrics shown are (A) continuous rank probability score (CRPS), (B) root mean squared error (RMSE), and (C) mean absolute error (MAE). All metrics are shown as a function of the forecast horizon.

191 of the predictions gradually became narrower as the number of months used to train the models
 192 increased (Supplementary Figure 2).

193 The predictive ability of the model also varied with the month of the year (Figure 2). Overall,
 194 larger discrepancies between observed and predicted values were observed between July and
 195 December, when typically more cases are reported. Similar patterns were observed for the CRPS,
 196 RMSE and MAE (Supplementary Figure 3).

197 The skill of the forecast showed significant spatiotemporal variation. Skill was evaluated using
 198 the continuous rank probability skill score, CRPSS (*Bradley and Schwartz, 2011*) comparing the
 199 CRPS of the superensemble to that of a baseline model (see Methods and Materials). CRPSS values
 200 larger than zero indicate a better predictive skill than that of the baseline model. Figure 2 shows
 201 that the CRPSS was consistently better than the baseline across the whole country for the period
 202 December to July. From August to November the skill is reduced for selected provinces in the central
 203 and southern regions characterised by larger dengue incidence variability.

204 Supplementary Figure 4 shows the observed and posterior predictive mean dengue cases
 205 across the 63 provinces computed one month ahead using the model superensemble. The model
 206 superensemble is able to reproduce the spatiotemporal dynamics of dengue fever with reasonable
 207 skill although predictions tend to underestimate the number of cases particularly during large
 208 outbreaks.

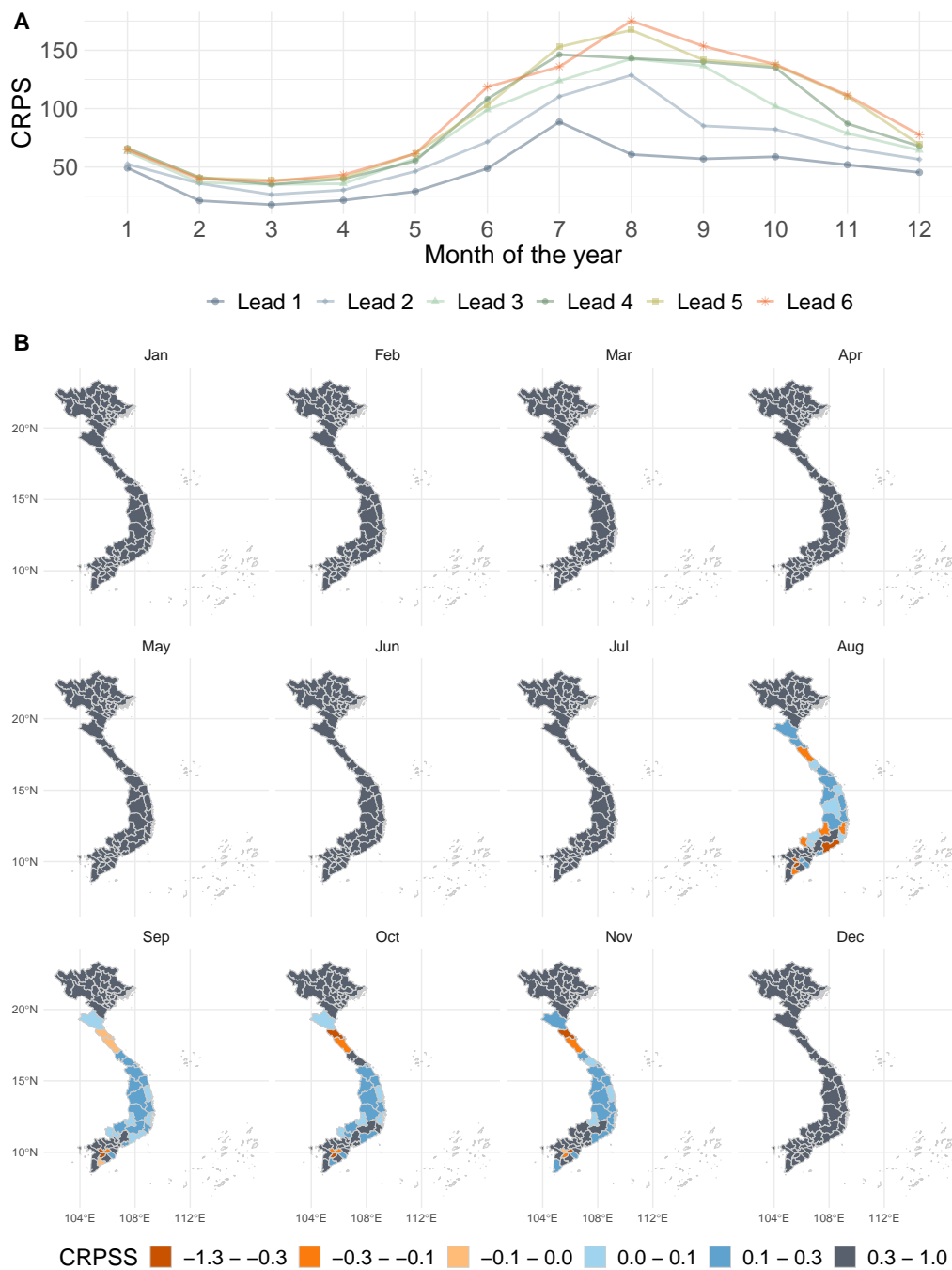


Figure 2. (A) Continuous rank probability score (CRPS) of the model superensemble arranged by month and forecast horizon, averaged across Vietnam. (B) Spatiotemporal variation of the continuous rank probability skill score (CRPSS) of the model superensemble. Gray and blue shaded areas indicate a better performance of the superensemble compared to a baseline model. Orange areas indicate a lower performance of the superensemble compared to the baseline model.

209 **Outbreak detection**

210 The skill of the model for outbreak detection was evaluated using the Brier score (*Brier, 1950*).
211 For the Brier score, smaller values are better. Four moving outbreak thresholds were defined to
212 evaluate the skill of the superensemble for outbreak prediction: i) the endemic channel plus one
213 standard deviation, ii) the endemic channel plus two standard deviations, iii) the 75th percentile of
214 the distribution of dengue cases, and iv) the 95th percentile of the distribution of dengue cases. The
215 endemic channel was calculated as the number of dengue cases per month and per province over
216 the previous five years in agreement with current practice at the Vietnamese Ministry of Health.
217 The 75th and 95th percentiles were calculated over the whole observational period at each time step
218 (see Methods and Materials). We then calculated the probability of exceeding the moving outbreak
219 threshold based on the posterior marginal distribution of the predicted number of cases.

220 Predictive skill was evaluated by comparing the predicted probability of exceeding the moving
221 outbreak threshold to observed outbreaks. Observed outbreak months were defined as months
222 where the number of recorded dengue cases exceeded the moving outbreak thresholds. Figure
223 3 shows the observed outbreaks and their corresponding predicted probabilities based on the
224 marginal posterior distribution of the model output. Figure 4 depicts the outbreak detection skill of
225 the superensemble across the forecast horizon (top) and time of the year (bottom).

226 As expected, the highest skill was achieved at a lead time of one month, after which the skill of
227 the superensemble gradually declines. Across the forecast horizon, the highest skill was observed
228 when using an outbreak threshold based on the 95th percentile of the distribution of dengue cases,
229 followed by the endemic channel plus two standard deviations. Stratified by month of the year, the
230 skill of the superensemble was generally greater between April and October. We noted that using
231 the endemic plus one or two standard deviations results in a significant decrease in skill between
232 June and July. This situation is not observed when using percentiles. When stratified by month, skill
233 was larger using the 95th percentile of the distribution of dengue cases followed by the endemic
234 channel plus two standard deviations.

235 There was significant spatial variation in the outbreak detection skill of the superensemble.
236 Figure 5 shows that across all moving outbreak thresholds, skill was greater in the northern
237 provinces compared to the central and southern provinces. Note that using the 95th percentile as a
238 threshold results in a larger number of provinces with a low Brier score (< 0.1).

239 Public health officials may be more likely to take preventive action if the probability of observing
240 an outbreak exceeds a certain value (*Lowe et al., 2016*). We investigated different probability
241 thresholds to define the optimal cut-off value that maximised the sensitivity and specificity of the
242 superensemble. As a metric of the accuracy of the forecasts, we computed the area under the
243 receiver operating characteristic curve (AUC) where an AUC value of 1 represents perfect skill, and a
244 value of 0.5 represents no better skill than a random guess. Supplementary Figure 5 indicates that
245 the forecasts always performed better than randomly guessing. As expected, the AUC decreased as
246 the forecast horizon increased. The AUC values ranged between 0.91 and 0.72.

247 **Relative economic value**

248 A cost-loss analysis of the *relative economic value* of the superensemble was undertaken using the
249 value index (*Richardson, 2000; Thornes and Stephenson, 2001*). The value index ranges between
250 zero and one, with one indicating a perfect forecast. The relative economic value of the superensem-
251 ble can be interpreted as the cost of using the system relative to the cost of either never preventing
252 outbreaks or always taking action. Location- and time-specific costs and losses may be difficult to
253 quantify. Therefore, our figures are only illustrative.

254 We defined a range of theoretical epidemic thresholds ranging between the 51st and the 99th
255 percentiles of the distribution of dengue cases for the whole time series. Outbreak thresholds were
256 province- and month-specific. Figure 6 shows that the superensemble has a theoretical relative
257 economic value in multiple provinces across a range of cost-loss ratios (i.e. the ratio of the expenses
258 derived of taking preventive action to the potential losses averted) and outbreak thresholds. As

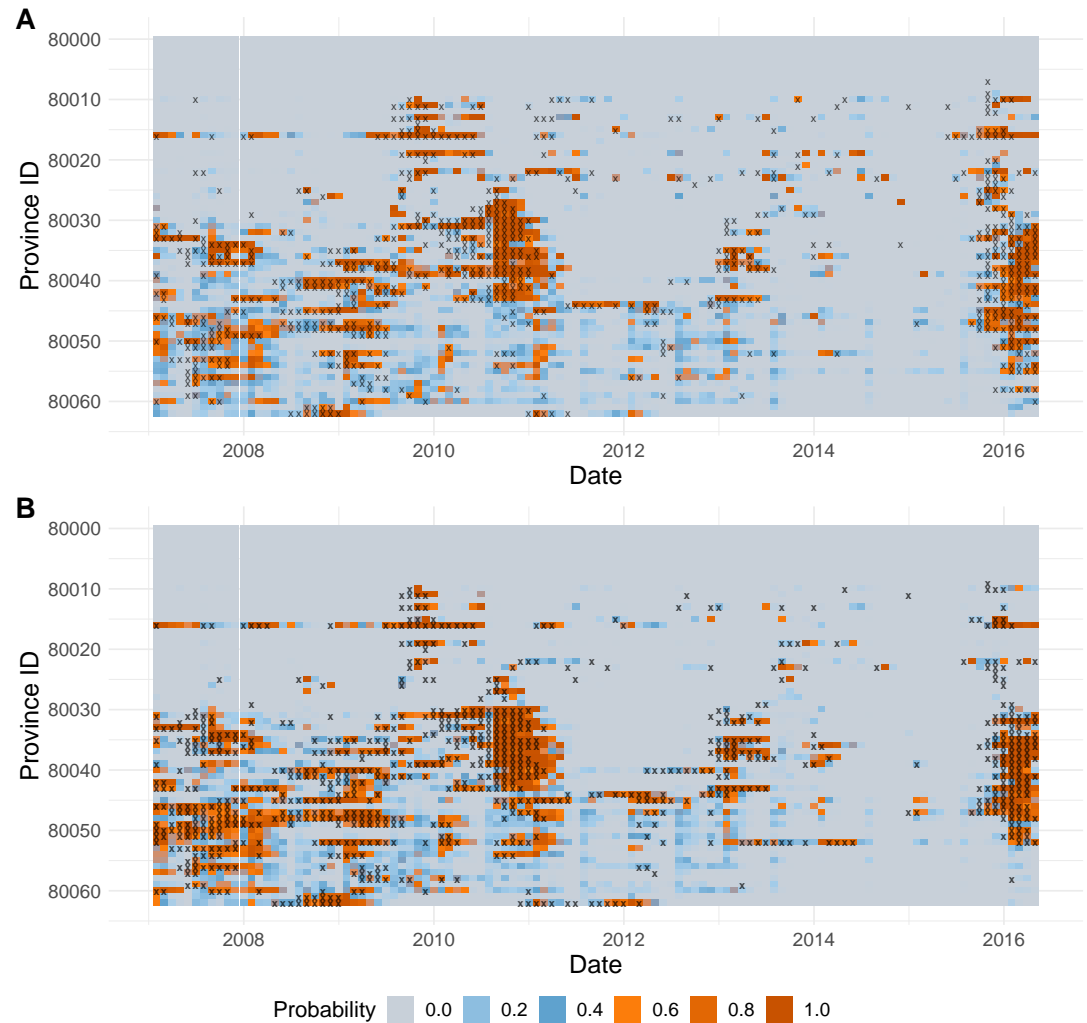


Figure 3. Predicted probability of exceeding the (A) mean plus two standard deviations, and (B) 95th percentile moving outbreak thresholds one month ahead. The X axis represents each of the forecast months. The Y axis indicates each of the provinces arranged from north (top) to south (bottom). Observed outbreaks (i.e. observed dengue cases above the outbreak threshold) are marked with a cross. Darker red colours represent a higher probability of exceeding the outbreak threshold.

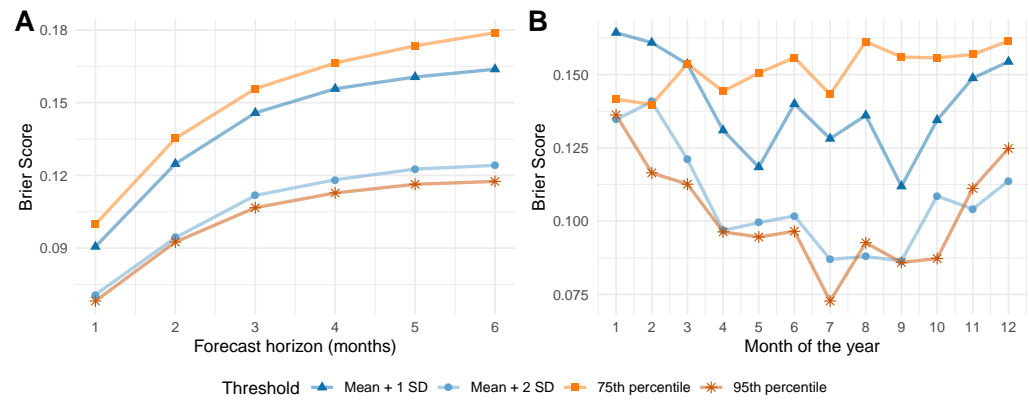


Figure 4. (A) Variation in the Brier score of the superensemble predictions for a forecast horizon (lead time) from one to six months calculated for four different outbreak thresholds. (B) Variation in the Brier score of the superensemble predictions per calendar month, and for four different outbreak thresholds. Scores correspond to their mean values for the period January 2007 to December 2016 ($n = 114$ months), and across 63 Vietnamese provinces. Lower scores indicate a greater accuracy for detecting outbreaks.

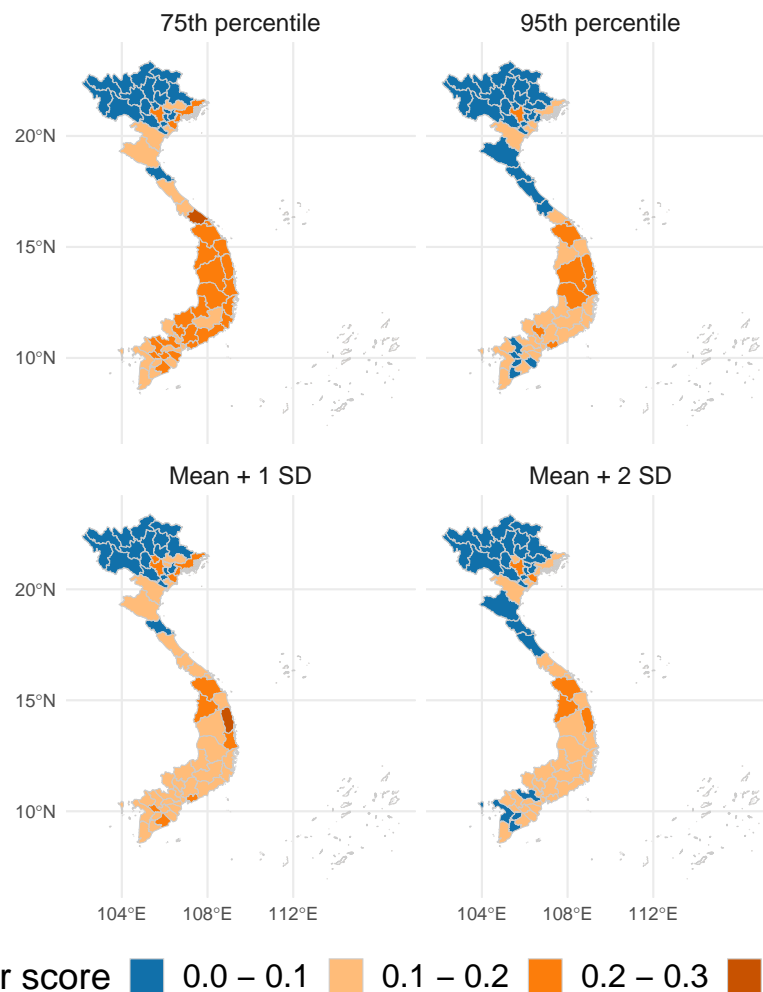


Figure 5. Spatial variation of the mean Brier score including all forecast horizons (lead times) from one to six months calculated for four different outbreak thresholds for the period January 2007 to December 2016 ($n = 114$) for each of the 63 Vietnamese provinces. Lower Brier scores (in blue) indicate a greater accuracy for detecting outbreaks.

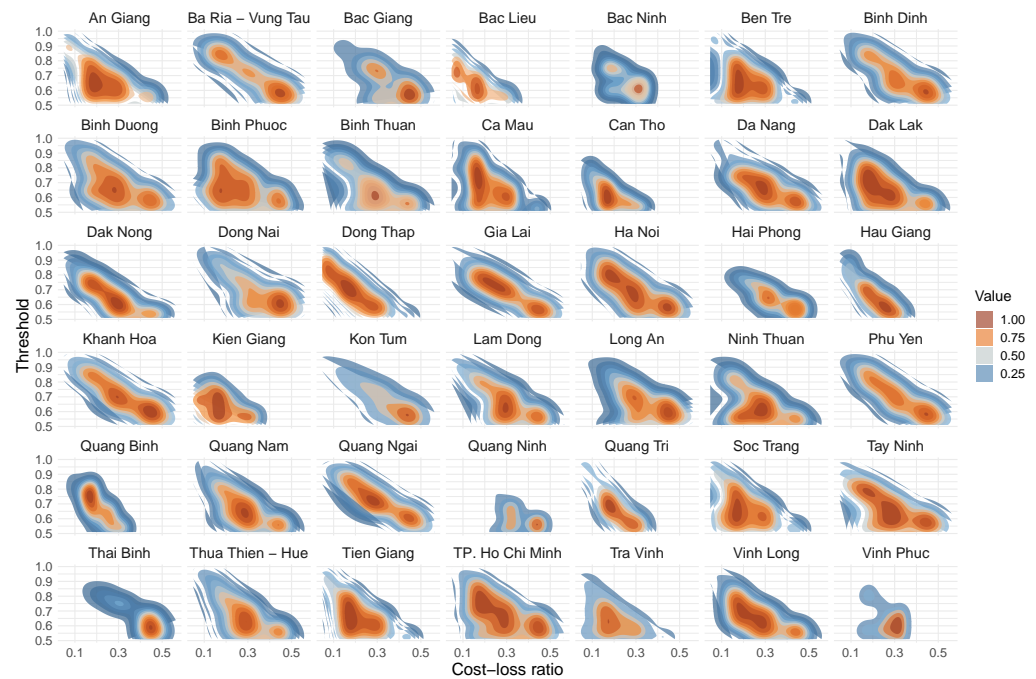


Figure 6. Relative economic value of using the dengue forecasting superensemble one month ahead under a range of cost-loss ratios (X axis) and outbreak thresholds (Y axis) for 42 Vietnamese provinces where the superensemble showed value at for multiple cost-loss ratios. Orange colours indicate a greater economic value, and blue colours a lower value.

259 the outbreak threshold increases and outbreaks become rarer, the superensemble has relative
 260 economic value at lower cost-loss ratios. Overall, larger values were observed when the cost-loss
 261 ratio was between 0.2 and 0.3 (i.e. the cost of taking preventive action is between 1/3 and 1/5 of the
 262 potential losses caused by not taking action). Supplementary Figure 6 shows the spatiotemporal
 263 variation in relative economic value across Vietnam for the forecast horizon of one to six months.
 264 Note, the superensemble has relative economic value in areas where dengue is typically endemic
 265 such as the central and southern provinces. The superensemble had no relative economic value
 266 for northern provinces where dengue is typically absent. The relative economic value gradually
 267 declined from an average value of 0.31 one month ahead across 75% ($n=47$) of the provinces, to an
 268 average value of 0.13 six months ahead across 62% ($n=39$) of the provinces .

269 **Prospective predictions of dengue risk**

270 The model superensemble was driven by seasonal climate forecasts to generate dengue forecasts
 271 for the period April to September 2020 using near-real time seasonal climate forecast data from
 272 the UK Met Office's Global Seasonal forecasting system version 5 (GloSea 5) (*MacLachlan et al.,*
 273 *2015; Scaife et al., 2014*). The posterior mean of the predicted values for each ensemble member
 274 are shown on Figure 7 (solid lines) for all Vietnamese provinces along with the corresponding 95%
 275 credible interval (dashed lines). The upper bound of the shaded areas represent percentile-based
 276 moving outbreak thresholds computed using observed dengue case data over the period August
 277 2002 to March 2020.

278 There is little spread between the 42 ensemble members indicating little between-member
 279 variability. It is noted that, as expected, the credible intervals increase as the forecast horizon
 280 increases, reflecting the uncertainties associated with the seasonal climate models used to generate
 281 the forecasts. For multiple provinces, the predicted mean number of dengue cases is above the
 282 75th percentile outbreak threshold (i.e. yellow and orange shaded regions). This suggests that he
 283 April to September transmission season may be above normal conditions in multiple areas across

284 Vietnam and particularly in Binh Doung, Dong Nai and Ho Chi Minh, all of which are located in the
285 southern part of the country, characterised by endemic transmission.

286 Discussion

287 We introduced a probabilistic dengue early warning system based on a model superensemble,
288 formulated using Earth observations, and driven by seasonal climate forecasts. The system is
289 able to generate accurate probabilistic forecasts of dengue metrics that could guide policy- and
290 decision-making processes. We demonstrate that using a model superensemble results in better
291 forecasts than using individual models.

292 Deciding which predictive model is the best from a suite of competing models is not a straightfor-
293 ward task. Each model carries somewhat different information of the modelled processes. Here, we
294 present a method for reconciling between-model disagreements while improving forecast accuracy.
295 The combination of models into a superensemble helps offset individual model biases across
296 time and space (*Krishnamurti et al., 2016*). Superensembles were initially developed for climate
297 modelling (*Krishnamurti et al., 1999*), and have gradually gained popularity in disease modelling
298 (see for example *Yamana et al., 2016, 2017; Johansson et al., 2019*).

299 Our novel dengue early warning system relies on probabilistic models to properly reflect forecast
300 uncertainty, and to explicitly assign probabilities to outcomes (*Held et al., 2017*). The system has
301 been developed to aid policy- and decision-making processes in Vietnam with the guidance of key
302 stakeholders in the Ministry of Health of Vietnam, the World Health Organization regional office, the
303 Pasteur Institutes in Nha Trang and Ho Chi Minh, Province-level Ministries of Health, the Vietnamese
304 National Institute of Hygiene and Epidemiology, and the Tay Nguyen Institute of Hygiene and
305 Epidemiology. A range of stakeholder engagement workshops, face-to-face meetings with users,
306 and surveys were conducted to tailor the system to the users' needs. Our results demonstrate that
307 our spatio-temporal superensemble could guide changes to the current practice in dengue control
308 towards a more preventative approach allowing bespoke and targeted public health interventions,
309 and a more efficient allocation of scarce resources.

310 We demonstrate that the superensemble outperforms the predictive ability of any individual
311 probabilistic model from a suite of top performing models in line with previous research (*Krish-
312 namurti et al., 1999; Yamana et al., 2016, 2017; Johansson et al., 2019*). Model performance was
313 assessed using a range of proper verification metrics across time and space to ensure the quality
314 of our probabilistic predictions. To our knowledge, this is one of the first early warning systems
315 informed by Earth observations to demonstrate skill for prospective year-round dengue prediction
316 in a robust out-of-sample framework. It is also one of the first prototypes for routine dengue early
317 warning at multiple time leads.

318 We found that the performance of the superensemble varied with geographic location, forecast
319 horizon, and time of the year. The system showed skill in predicting spatiotemporal variations in
320 dengue cases and outbreak occurrence at forecast horizons of up to six months ahead. Predictions
321 deteriorated, and uncertainty increased as the forecast horizon expanded, in line with previous
322 research (*Reich et al., 2016; Chen et al., 2018; Funk et al., 2019; Tompkins et al., 2019*). Forecasts
323 improved, and credible intervals decreased as time progressed and dengue data increased likely due
324 to an improvement in the associations learned by the superensemble. Forecast skill is lower for the
325 onset and peak of the transmission season due to substantial interannual variation (Supplementary
326 Figure 7).

327 Relative to a baseline seasonal predictive model (see Methods and Materials), the superensem-
328 ble made, on average, more accurate predictions across most provinces, and for most of the
329 year in agreement with previous studies (*Lowe et al., 2016; Chen et al., 2018*). This observation,
330 however, conflicts with the results obtained by *Lauer et al. (2018)* and *Johansson et al. (2019)* for
331 whom climate-naïve models had better skill than seasonal-climate-informed ones. *Johansson et al.
332 (2019)* noted that incorporating seasonal climate data into predictive models may increase model
333 complexity at the expense of lower out-of-sample forecast skill.

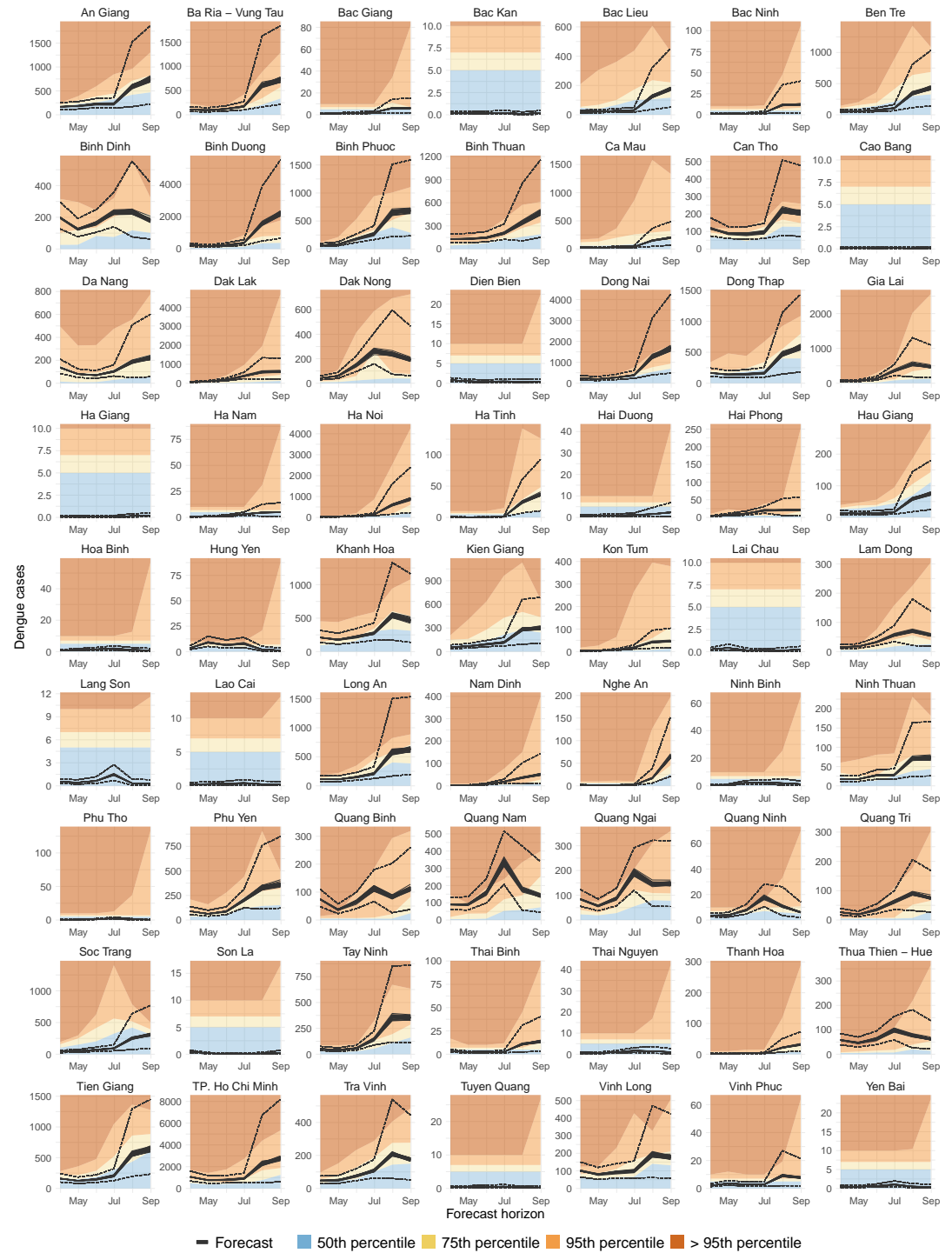


Figure 7. Predicted dengue cases for the period April to September 2020 for all Vietnamese provinces using a model superensemble. The forecast was issued on 5 April 2020. The X axis indicates the forecast horizon. The Y axis indicates the predicted number of dengue cases. Solid lines indicate the mean estimate for each of 42 forecast ensemble members. The dashed lines indicate the upper and lower bounds of the 95% credible intervals. The upper bound of the shaded areas indicates the month- and province-specific percentiles based on dengue data for the period August 2002 to March 2020.

334 Outbreaks are difficult to predict, even more so at forecast horizons of several months ahead.
335 Nevertheless, our superensemble demonstrated skill for outbreak detection up to a lead time of six
336 months, evaluated using proper scores over a suite of moving outbreak thresholds. One of the best
337 performing outbreak thresholds was the endemic channel plus two standard deviations, computed
338 using data from the previous five years. This method is currently in use across Vietnam (*Badurdeen*
339 *et al., 2013; Runge Ranzinger et al., 2014*) although with the difference of excluding outbreak years
340 from the computation of the threshold. Following our results, however, we recommend the use
341 of the 95% percentile as a threshold for outbreak detection as it gave better and more consistent
342 results, likely due to its ability for detecting very large outbreaks. Recognizing the limitations of the
343 province-level data, it is encouraging that our predictions are skillful in most provinces up to six
344 months in advance.

345 In disease forecasting, each decision (e.g. to prevent or not to prevent an outbreak) has an
346 associated cost that will lead to a benefit or a loss depending on the outcome. Decision-makers have
347 the task of selecting the action that minimises potential losses. We used a simple cost-loss analysis
348 (*Thornes and Stephenson, 2001*) to assess the relative economic value of the superensemble. Our
349 figures are only illustrative. Still, they highlight that using a dengue early warning system has relative
350 economic value compared to not using a forecast. The superensemble had considerable value
351 across most provinces. However, in northern provinces, where dengue is essentially absent, the
352 forecast is predicted to have limited relative economic value compared to always preparing for an
353 outbreak or never preparing for it. The assessment of the relative economic value of the dengue
354 early warning system may help stakeholders justify public investment in the development and
355 generation of seasonal forecasts, or to help raise awareness of their potential value.

356 Although our proposed early warning system provides useful information for public health
357 preparedness and response, it has some limitations worth mentioning. First, whilst our modelling
358 framework incorporates important determinants of dengue occurrence such as climate and urban-
359 isation, it does not explicitly incorporate, at this stage, the potential effects of other important
360 determinants of disease such as the deployment of mosquito control interventions, vector indices,
361 serotype-specific circulation, herd-immunity, and the mobility of people and goods all of which
362 may lead to significant changes in the level of risk experienced locally (*Reiter, 2001*). Stakeholders
363 in Vietnam have recently started collecting supporting data on vector indices and dengue virus
364 serotypes. However, it is noted that there is a lack of publicly available, continuous, and long-term
365 data sets that could be used to inform modelling efforts. In this study, we account for some of the
366 variation that might be attributed to these factors by using spatio-temporal random effects in each
367 of the models included in the superensemble. Future developments of the system may incorporate
368 some of these factors if data are made available. Second, the quality and consistency of the dengue
369 data are affected by the limited confirmation of suspected cases through laboratory diagnostic test,
370 leading to large uncertainties that are difficult to quantify. Third, our computations of dengue risk
371 do not take into consideration uncertainties due to the potential under- or misreporting of dengue
372 cases. Consequently, our model superensemble forecasts may underestimate the real number of
373 cases occurring at any given time.

374 **Conclusion**

375 We have demonstrated that a probabilistic dengue early warning model, formulated using Earth
376 observations and driven by seasonal climate forecasts, to produce probabilistic predictions of
377 exceeding pre-defined outbreak-detection thresholds. A theoretical cost-loss analysis showed that
378 the system has relative economic value for a range of cost-loss ratios across most of the country,
379 indicating that such a system may be useful for decision-making processes. The dengue forecasting
380 system presented here has been rolled out across Vietnam and could be tailored for other dengue-
381 endemic countries. Further work may include investigating the feasibility of producing probabilistic
382 forecasts with sufficient skill at the district or commune levels.

383 **Methods and Materials**

384 **Dengue surveillance data**

385 Monthly dengue cases were obtained from the Vietnamese Ministry of Health. Data were retrieved
386 for the period August 2002 to December 2019 at the province level ($n = 63$). Data comprised
387 suspected and confirmed dengue cases although there was no indication as to how many cases fell
388 within each category. The data set did not contain serotype-specific information.

389 **Historical Earth observation data**

390 Minimum, maximum and mean air temperature at 2 metres above ground ($^{\circ}\text{C}$) were derived from
391 MODIS daily L3 global land surface temperature products (*Wan et al., 2015*) with a spatial resolution
392 of 1 km^2 . Precipitation amount per day (mm day^{-1}), was initially retrieved from from the Tropical
393 Rainfall Measurement Mission (*Hijmans, 2011*) at a spatial resolution of 25 km^2 up to April 2014.
394 After April 2014, precipitation data were obtained from the Global Precipitation Mission (*Huffman*
395 *et al., 2019*) at a spatial resolution of 10 km^2 . Daytime specific surface humidity (kg kg^{-1}) was
396 calculated using the daytime MODIS L2 water vapour near infrared MOD 5 products (*Gao, 2015*)
397 with a spatial resolution of 1 km^2 . Average daily wind speed at 10 metres above ground (m s^{-1})
398 was retrieved from the European Centre for Medium-range Weather Forecasts (ECMWF) ERA-5
399 reanalysis (*Copernicus, 2019*) for the period 2002-2011, at a spatial resolution of 31 km^2 . After 2011,
400 wind speed data were obtained from the NOAA Climate Forecast System (*Saha et al., 2014*) at a 20
401 km^2 resolution. Monthly sea surface temperature anomalies for the Niño region 3.4 (5°S – 5°N , and
402 170° – 120°W) were obtained from the NOAA Center for Weather and Climate Prediction, Climate
403 Prediction Center (*NOAA, 2020*) for the period 2002 to 2020. Population weighted monthly averages
404 for each climate variable were calculated for each province.

405 Population data from the Worldpop project (*WorldPop, 2018*) at a 100 m^2 spatial resolution was
406 used to calculate annual gridded weights for each province. At the time of processing, Worldpop
407 data were only available for the years, 2009, 2010, 2015, 2020. Intervening years were generated
408 using linear interpolation.

409 **Historical demographic and land-cover data**

410 Total population per province were retrieved from the Socioeconomic Data and Applications Center
411 (SEDAC) Gridded Population of the World project version 4.11 (*CIESIN, 2019*) at a 1 km^2 resolution,
412 5-yearly for the period 2000 to 2020. Intervening years were generated using linear interpolation.
413 The percentage of urban, peri-urban and rural land-cover per province for the period 2002-2019
414 was derived from the ESA CCI Land-cover project (*ESA, 2017; Copernicus, 2019*) which describes the
415 land surface into 22 classes at a spatial resolution of 0.00277 degrees.

416 **Seasonal climate forecasts**

417 Seasonal climate forecasts of minimum temperature, maximum temperature, daily precipitation,
418 specific humidity, wind speed and sea surface temperature anomalies for the Niño region 3.4
419 were obtained from the UK Met Office Global Seasonal Forecasting System version 5 (GloSea5)
420 (*MacLachlan et al., 2015; Scaife et al., 2014*). GloSea5 comprises 42 ensemble members built
421 around a high resolution climate prediction model (HadGEM3). Ensemble members differ due
422 to small stochastic physics perturbations provided by the Stochastic Kinetic Energy Backscatter
423 v2 (*Bowler et al., 2009*). GloSea5 has a resolution of 0.83 degrees in latitude and 0.56 degrees
424 in longitude for the atmosphere, and 0.25×0.25 degrees for the ocean. GloSea5 has two major
425 components, the forecast itself and the associated hindcasts or *historical re-forecasts*, which are
426 used for calibration and skill assessment.

427 **Seasonal climate hindcasts**

428 Hindcast data (i.e. historical forecasts) were retrieved from the Copernicus Climate Data Store
429 (*Copernicus, 2019*) at monthly time steps for each of the 28 ensemble members. Data were obtained

430 using the GloSea5 system (*MacLachlan et al., 2015; Scaife et al., 2014*) across a forecast horizon
431 of one to six months ahead for the period January 2012 to December 2016. At the time of the
432 computations, data for the period May to October 2016 were unavailable.

433 **Modelling approach**

434 Let $Y_{i,t}$ be the number of dengue cases for province $i = 1, \dots, I$ at time $t = 1, \dots, T$ be modelled using
435 Bayesian generalised linear mixed models (GLMM). Models were fitted using a negative binomial
436 specification to account for potential over-dispersion in the data. The general algebraic definition
437 of the models is given by:

$$Y_{i,t} \mid \mu_{i,t}, \phi \sim \text{NegBin}(\mu_{i,t}, \phi),$$

438 where $\mu_{i,t}$ is the expected number of dengue cases for province i and time t , and $\phi > 0$ is the
439 negative binomial dispersion parameter. A logarithmic link function of the expected number of
440 cases is modelled as:

$$\log(\mu_{i,t}) = \alpha + \log(P_{i,a[t]}) + \rho \log(Y/P_{i,t-1}) + \sum_k \beta_k X_{i,t,k} + \gamma_{i,a[t]} + \delta_{i,m[t]} + u_i + v_i,$$

441 where α corresponds to the intercept; $\log(P_{i,a[t]})$ denotes the logarithm of the population at risk
442 for province i and year $a[t]$, included as an offset to adjust case counts by population; ρ is the
443 auto-regressive coefficient; $\log(Y/P_{i,t-1})$ is the logarithm of the observed dengue incidence rate in
444 the previous month with regression coefficient ρ ; X is a matrix of k seasonal Earth observation
445 (meteorological and land-cover type) explanatory variables with regression coefficients β . Long-
446 term trends are modelled using province-specific unstructured random effects for each year
447 ($\gamma_{i,a[t]}$). Seasonality is accounted for by using province-specific structured random effects for each
448 calendar month ($\delta_{i,m[t]}$) with first order auto-regressive prior to allow each month to depend on
449 the previous one. Unknown confounding factors, such as interventions and spatial dependency
450 structures representing, for example, human mobility, were incorporated using structured (v_i) and
451 unstructured (u_i) random effects for each province i . Spatial random effects were specified using a
452 Besag-York-Mollie model (*Besag et al., 1991*) which incorporates a spatial effect with a Gaussian
453 exchangeable prior to account for unstructured variation, and a spatial effect with an intrinsic
454 conditional auto-regressive prior to account for spatially-structured variability. Delayed effects
455 of meteorological factors on dengue were accounted for by incorporating the moving average
456 of temperature, precipitation, specific humidity, and diurnal temperature range (DTR) all of them
457 lagged zero to two months, and sea surface temperature anomalies in the Niño region 3.4 lagged
458 three months based on previous research (e.g. (*Petrova et al., 2019; Colón-González et al., 2018a;*
459 *Lowe et al., 2018*) and exploratory analyses. No delayed effects were considered for wind speed.
460 Flat priors were set to regression coefficients (α, ρ, β) and penalising complexity priors were assumed
461 for both the dispersion parameters and the precision for all random effects (*Simpson et al., 2017*).
462 Models were fitted in R version 3.6.1 using the INLA package (*Rue et al., 2009*). The relevant R code
463 is available at https://github.com/FelipejColon/paper_dengue_superensemble.

464 **Model selection**

465 The best subset of seasonal climate predictors leading to the lowest observed-prediction discrepan-
466 cies for a given model was obtained using an expanding window time series cross-validation (TSCV)
467 algorithm (*Hyndman and Athanasopoulos, 2014*). Land-cover variables were included in all models
468 as they varied annually. We iteratively fitted all possible models containing one seasonal climate
469 predictor at the time, then two seasonal climate predictors, and so on, until all seasonal climate
470 variables were included in a full model (*Colón-González et al., 2018a*). The predictive ability of each
471 model was evaluated using the continuous rank probability score (CRPS), root mean squared error
472 (RMSE), mean absolute error (MAE) deviance information criterion (DIC) and the Watanabe-Akaike

473 information criterion (WAIC). Verification metrics were computed in R using the `SpecsVerification`
474 and `ModelMetrics` packages (*Siegert, 2017; Hunt, 2018*).

475 TSCV was implemented using an expanding window approach dividing the data set into multiple
476 training and testing sets. The initial training set comprised data from August 2002 to December
477 2006. Each time step (k), a further month of data was added to the training set until the training set
478 contained $n - 6$ observations. The testing set comprised the climate hindcast data for the six months
479 immediately after the last observation in the training set for each geographical area. Seasonal
480 climate hindcast data comprised 28 ensemble members.

481 We calculated the mean CRPS, RMSE, MAE, DIC and WAIC for each model specification across all
482 time steps and ensemble members. The best two performing models for each verification metric
483 were selected to create a superensemble.

484 **Accounting for autocorrelation in disease transmission**

485 The number of dengue cases occurring at time t is directly dependent on the number of cases
486 that occurred in the recent past. Previous research suggests that including the logarithm of the
487 number of cases in the previous month $\log(Y_{t-1})$ helps accounting for such temporal correlation in
488 disease transmission (*Imai et al., 2015*). Additionally, incorporating $\log(Y_{t-1})$ as a covariate improves
489 the predictive ability of seasonal-climate-informed disease models by reducing residual dispersion
490 (*Imai et al., 2015*). One complication of accounting for the number of cases in the previous month
491 in an operational system is that $\log(Y_{t-1})$ in the temporal window of the forecast can only be known
492 up to time $t + 1$. More specifically, to generate dengue forecasts one month ahead, $\log(Y_{i,t-1})$ for
493 time $t + 1$ corresponds to the logarithm of the number of dengue cases at time t . To generate the
494 forecasts two months ahead we first fitted a climate naïve model with the following specification:

$$\log(\mu_{it}) = \alpha + \log(P_{i,a[t]}) + \rho \log(Y/P_{i,t-1}) + \gamma_{i,a[t]} + \delta_{i,m[t]} + u_i + v_i,$$

495 with α as the intercept; $\log(P_{i,a[t]})$ as the population at risk in province i at time $a[t]$, included as
496 an offset; $\log(Y/P_{i,t-1})$ is the logarithm of the observed dengue incidence rate in the previous
497 month with regression coefficient ρ ; $\gamma_{i,a[t]}$ as province-specific unstructured random effects; $\delta_{i,m[t]}$
498 as province-specific structured random effects with an AR1 auto-correlation term; and u_i and v_i as
499 province-specific structured and unstructured random effects. We then predicted the number of
500 dengue cases for time $t + 1$. The logarithm of the predicted number of cases at time $t + 1$ was then
501 used as $\log(Y_{i,t-1})$ for time $t + 2$. We repeated these steps for each lead time in the forecast. Given
502 the high number of zero counts in the set, we used the logarithm of the number of dengue cases
503 lagged 1 month plus one.

504 **Baseline model**

505 A baseline model was developed for comparison purposes. The algebraic definition of the model is
506 given by:

$$\log(\mu_{it}) = \alpha + \log(P_{i,a[t]}) + \gamma_{i,a[t]} + \delta_{i,m[t]} + u_i + v_i,$$

507 where α is the intercept; $\log(P_{i,a[t]})$ is the population at risk in province i at time $a[t]$, included as an
508 offset; $\gamma_{i,a[t]}$ are province-specific unstructured random effects; $\delta_{i,m[t]}$ are province-specific structured
509 random effects with an AR1 auto-correlation term; and u_i and v_i as province-specific structured and
510 unstructured random effects.

511 **Model superensemble**

512 Given a number of competing models, we generated a model superensemble (*Yamana et al., 2016*)
513 using Bayesian model averaging (BMA) implemented in the `INLABMA` package (*Bivand et al., 2015*).
514 BMA relies on the weighted sum of the conditional marginals obtained from a group of competing
515 models (*Gómez-Rubio et al., 2019*). The posterior distribution given data D is estimated as follows:

$$\pi(\Delta|D) \simeq \sum_{i=1}^K \pi(\Delta|D, M_k)w_i$$

516 with Δ as the quantity of interest; M_k indicates the $M_1 \dots M_K$ competing models that will form the
517 superensemble; and the weights w_i computed based on the maximum likelihood (ML) and deviance
518 information criterion (DIC) of the competing models. The weights w_i were calculated as follows:

$$w_i = \frac{\exp(X_k - \max(X))}{\sum_{i=1}^K \exp(X_k - \max(X))} \times 0.5,$$

519 where X is a vector of ML or DIC values for each of the competing models k . The vector of ML
520 and DIC weights were multiplied by 0.5 so that the resulting superensemble weights would sum to
521 one, and then added together into a single vector w_i . A new set of superensemble weights were
522 computed each time a forecast was issued using training data from all previous years.

523 The predictive ability of the superensemble was evaluated using the CRPS, RMSE and MAE. In
524 addition, we evaluated the skill of the obtained forecasts using the continuous rank probability skill
525 score (CRPSS). CRPSS is defined as:

$$CRPSS = 1 - \frac{CRPS_f}{CRPS_b},$$

526 where $CRPS_f$ is the continuous rank probability score (CRPS) of the forecast and $CRPS_b$ is the
527 CRPS of a baseline model used as a reference (*Bradley and Schwartz, 2011*).

528 **Outbreak detection evaluation**

529 The skill of the model for detecting dengue outbreaks was evaluated using the Brier (*Brier, 1950*) and
530 logarithmic scores (*Good, 1952*) both of which are strictly proper scoring rules based on probability
531 densities. The Brier score was calculated as follows:

$$BS = \frac{1}{N} \sum_{i=1}^N (f_i - O_i)^2,$$

532 where N is the number of predictions; f_i is the forecast probability that an outbreak may happen;
533 and O_i takes the value of one if there was an outbreak, or zero if there was no outbreak. The
534 logarithmic score was calculated as follows:

$$LS = \frac{1}{N} \sum_{i=1}^N \log(p_i),$$

535 where p_i denotes the probability assigned to the observed outcome over N predictions. Scores
536 were computed in R using the `scoring` package (*Merkle and Steyvers, 2013*).

537 **Relative economic value**

538 Unlike skill, the relative economic value (V) of the model superensemble depends on requirements
539 set by the user (*Thornes and Stephenson, 2001*). Typically, V is evaluated in monetary terms and is
540 particularly useful when the probability of occurrence of an adverse event (e.g. a major outbreak)
541 is known. If the probability of occurrence of an outbreak is greater than the ratio of the cost of
542 taking preventive action divided by the loss incurred by not taking action (C/L ratio) and an outbreak
543 occurs, then it will pay off to take action. If the probability of occurrence of an outbreak is lower
544 than the C/L ratio, then it does not pay off to take preventive action. If the probability is equal to
545 the C/L ratio, it does not matter if action is taken or not.

546 V was estimated for a range of theoretical C/L ratios following (*Thornes and Stephenson, 2001*)
547 by comparing the mean cost of using the forecasting system for outbreak detection compared
548 to the mean expense incurred by either never preventing outbreaks or, on the contrary, taking
549 preventive action every month of the year. V takes a value of one if the forecast is perfect, and

550 a value of zero if it is no better than the default action plan. If V is negative, it indicates that the
551 forecast is so poor that it would be more cost effective not to use it. V is represented as follows:

$$V = \frac{E(S) - E(A)}{E(S) - E(P)},$$

552 where $E(S)$ is the expense incurred by taking preventive action each month of the year or the losses
553 incurred by no taking action at all even when an outbreak occurred, whichever is the cheapest
554 method when not using the forecast; $E(A)$ represents the total cost of the forecast calculated as
555 the cost of type 1 (false positive) and type 2 (false negative) errors plus the cost of acting when the
556 forecast was correct; and $E(P)$ indicates the cost incurred with a perfect forecast (i.e action is taken
557 only when there is an outbreak).

558 Funding

559 UK Space Agency Dengue forecasting MOdel Satellite-based System (D-MOSS)

- 560 • Felipe J Colón-González
- 561 • Leonardo Soares Bastos
- 562 • Barbara Hofmann
- 563 • Alison Hopkin
- 564 • Quillon Harpham
- 565 • Tom Crocker
- 566 • Rosanna Amato
- 567 • Iacopo Ferrario
- 568 • Francesca Moschini
- 569 • Samuel James
- 570 • Sajni Malde
- 571 • Eleanor Ainscoe
- 572 • Mark Harrison
- 573 • Gina Tsarouchi
- 574 • Darren Lumbroso
- 575 • Oliver Brady
- 576 • Rachel Lowe

577 The funders had no role in study design, data collection and interpretation, or the decision to
578 submit the work for publication.

579 Acknowledgments

580 We would like to thank Dr Satoko Otsu, and Dr Trang Cong Dai Dai from the World Health Organiza-
581 tion Office in Vietnam, and Mr Dao Khanh Tung from the United Nations Development Programme
582 Office in Vietnam for the helpful discussions and support for the development of this project. We
583 would also like to acknowledge the support and insights of the Pasteur Institute Ho Chi Minh City,
584 the Pasteur Institute Nha Trang, the Institute of Hygiene and Epidemiology Tay Nguyen (TIHE), and
585 the National Institute of Hygiene and Epidemiology (NIHE).

586 References

- 587 **Badurdeen S**, Valladares D, Farrar J, Gozzer E, Kroeger A, Kuswara N, Runge Ranzinger S, Tinh H, Leite P,
588 Mahendradhata Y, Skewes R, Verall A. Sharing experiences: towards an evidence based model of dengue
589 surveillance and outbreak response in Latin America and Asia. *BMC Public Health*. 2013; 13:607.
- 590 **Bastos LS**, Economou T, Gomes MFC, Villela DAM, Coelho FC, Cruz OG, Stoner O, Bailey T, Codeço CT. A
591 modelling approach for correcting reporting delays in disease surveillance data. *Statistics in Medicine*. 2019;
592 38:4363–4377.

- 593 **Bergmeir C**, Benítez J. On the use of Cross-Validation for Time Series Predictor Evaluation. *Inf Sci.* 2012;
594 191:192–213.
- 595 **Besag J**, York J, Mollie A. Bayesian image restoration, with two applications in spatial statistics. *An I Stat Math.*
596 1991; 43:1–20.
- 597 **Bhatt S**, Gething P, Brady O, Messina J, Farlow A, Moyes C, Drake J, Brownstein J, Hoen A, Sankoh O, Myers
598 M, George D, Jaenisch T, Wint G, Simmons C, Scott T, Farrar J, Hay S. The global distribution and burden of
599 dengue. *Nature.* 2013; 496(7446):504–507.
- 600 **Bivand RS**, Gómez-Rubio V, Rue H. Spatial Data Analysis with R-INLA with Some Extensions. *Journal of Statistical*
601 *Software.* 2015; 63(20):1–31. <http://www.jstatsoft.org/v63/i20/>.
- 602 **Bowler NE**, Arribas A, Beare SE, Mylne KR, Shutts GJ. The local ETKF and SKEB: upgrades to the MOGREPS
603 short-range ensemble prediction system. *Q J R Meteorol Soc.* 2009; 41:767–776.
- 604 **Bradley AA**, Schwartz SS. Summary Verification Measures and Their Interpretation for Ensemble Forecast. *Mon*
605 *Weather Rev.* 2011; 139:3075–3089.
- 606 **Brady O**, Smith D, Scott, Hay S. Dengue disease outbreak definitions are implicitly variable. *Epidemics.* 2015;
607 11:91–102.
- 608 **Brady O**, Gething P, Bhatt S, Messina J, Brownstein J, Hoen A, Moyes C, Farlow A, Scott T, Hay S. Refining the
609 Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. *PLoS Negl Trop Dis.* 2012;
610 6(8).
- 611 **Brier GW**. Verification of forecasts expressed in terms of probability. *Mon Weather Rev.* 1950; 78(1):1–3.
- 612 **Chen Y**, Ong JHY, Rajarethinam J, Yap G, Ng LC, Cook AR. Neighbourhood level real-time forecasting of dengue
613 cases in tropical urban Singapore. *BMC Medicine.* 2018; 16:129.
- 614 **CIESIN**, Gridded Population of the World (GPW), v4; 2019. <https://sedac.ciesin.columbia.edu/data/collection/gpw-v4>.
615
- 616 **Colón-González FJ**, Harris I, Osborn T, Steiner São Bernardo C, Peres C, Hunter P, Warren R, van Vuuren D, Lake
617 I. Limiting global-mean temperature increase to 1.5–2 °C could reduce the incidence and spatial spread of
618 dengue fever in Latin America. *Proc Natl Acad Sci U S A.* 2018; 115:6243–6248.
- 619 **Colón-González FJ**, Lake IR, Morbey RA, Elliot AJ, Pebody R, Smith GE. A Methodological Framework for the
620 Evaluation of Syndromic Surveillance Systems: A Case Study of England. *BMC Public Health.* 2018; 18:544.
- 621 **Colón-González FJ**, Peres C, Steiner São Bernardo C, Hunter P, Lake I. After the epidemic: Zika virus projections
622 for Latin America and the Caribbean. *PLoS Neglected Tropical Diseases.* 2017; 11(11). www.scopus.com.
- 623 **Copernicus**, Copernicus Climate Data Store; 2019. <https://cds.climate.copernicus.eu/#/home>.
- 624 **Cuong HQ**, Vu NT, Cazelles B, Boni MF, Thai KT, Rabaa MA, Quang LC, Simmons CP, Huu TN, Anders KL. Spa-
625 tiotemporal Dynamics of Dengue Epidemics, Southern Vietnam. *Emerg Infect Dis.* 2013; 19:945–953.
- 626 **Dusfour I**, Vontas J, David J, Weetman D, Fonseca D, Corbel V, Kamaraju R, Mamadou B, Ademir J, Shinji K, Fabrice
627 C. Management of insecticide resistance in the major *Aedes* vectors of arboviruses: Advances and challenges.
628 *PLoS Negl Trop Dis.* 2019; 13:e0007615.
- 629 **Eastin M**, Delmelle E, Casas I, Wexler J, C S. Intra- and interseasonal autoregressive prediction of dengue
630 outbreaks using local weather and regional climate for a tropical environment in Colombia. *Am J Trop Med*
631 *Hyg.* 2014; 91:598–610.
- 632 **ESA**, Land cover CCI Product User Guide Version ; 2017. http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf.
633
- 634 **Funk S**, Camacho A, Kucharski AJ, Lowe R, Eggo RM, Edmunds WJ. Assessing the performance of real-time
635 epidemic forecasts: A case study of Ebola in the Western Area region of Sierra Leone. *PLoS Comput Biol.*
636 2019; 15(2):e1006785.
- 637 **Gage K**, Burkot T, Eisen R, , Hayes E. Dengue disease outbreak definitions are implicitly variable. *Am J Prev Med.*
638 2008; 25:436–450.

- 639 **Gao BC**, MODIS Atmosphere L2 Water Vapor Product; 2015. https://atmosphere-imager.gsfc.nasa.gov/MOD05_L2/index.html.
640
- 641 **Gómez-Rubio V**, Bivand RS, Rue H. Bayesian model averaging with the integrated nested Laplace approximation).
642 *Econometrics*. 2019; Submitted.
- 643 **Good IJ**. Rational decision. *J Roy Stat Soc*. 1952; 14A:107–114.
- 644 **Held L**, Meyer S, Bracher J. Probabilistic forecasting in infectious disease epidemiology: the 13th Armitage
645 lecture. *Stats Med*. 2017; 36(22):3443–3460.
- 646 **Hii Y**, Zhu H, Ng N, Ng L, Rocklöv J. Forecast of dengue incidence using temperature and rainfall. *PLoS Negl Trop*
647 *Dis*. 2013; 6:e1908.
- 648 **Hijmans R**, TRMM (TMPA) Rainfall Estimate L3 3 hour 0.25 degree x 0.25 degree V7; 2011. https://disc.gsfc.nasa.gov/datasets/TRMM_3B42_7/summary.
649
- 650 **Huffman GJ**, Stocker EF, Bolvin DT, Nelkin EJ, Tan J, GPM IMERG Final Precipitation
651 L3 Half Hourly 0.1 degree x 0.1 degree V06; 2019. <https://catalog.data.gov/dataset/gpm-imerg-final-precipitation-l3-half-hourly-0-1-degree-x-0-1-degree-v06-gpm-3imerghh-at-g>.
652
- 653 **Hung T**, Clapham H, Bettis A, Cuong H, Thwaites G, Wills B, Boni M, Turner H. The Estimates of the Health and
654 Economic Burden of Dengue in Vietnam. *Trends Parasitol*. 2018; 34:904–918.
- 655 **Hunt T**. ModelMetrics: Rapid Calculation of Model Metrics; 2018, [https://CRAN.R-project.org/package=](https://CRAN.R-project.org/package=ModelMetrics)
656 [ModelMetrics](https://CRAN.R-project.org/package=ModelMetrics), r package version 1.2.2.
- 657 **Hussain-Alkhateeb L**, Kroeger A, Oliario P, Rocklöv J, Sewe MO, Tejeda G, Benítez D, Gill B, Hakim SL, Carvalho
658 RG, Bowman L, Petzold M. Early warning and response system (EWARS) for dengue outbreaks: Recent
659 advancements towards widespread applications in critical settings. *PLoS One*. 2018; 13:e0196811.
- 660 **Hyndman R**, Athanasopoulos G, *Forecasting: Principles and Practice*. OTexts; 2014. Accessed 12/08/2014.
661 <https://www.otexts.org/book/fpp>.
- 662 **Imai C**, Armstrog B, Chalabi Z, Mangtani P, Hashizume M. Time series regression model for infectious disease
663 and weather. *Environ Res*. 2015; 142:319–327.
- 664 **Johansson MA**, Apfeldorf KM, Dobson S, Devita J, Buczak AL, Baugher B, Moniz LJ, Bagley T, Babin SM, Guven E,
665 et al. An open challenge to advance probabilistic forecasting for dengue epidemics. *Proc Natl Acad Sci U S A*.
666 2019; 116(48):24268–24274.
- 667 **Kraemer MU**, Sinka ME, Duda KA, Mylne AQ, Shearer FM, Barker CM, Moore CG, Carvalho RG, Coelho GE, Bortel
668 WV, Hendrickx G, Schaffner F, Elyazar IR, Teng HJ, Brady OJ, Messina JP, Pigott DM, Scott TW, Smith DL, Wint GW,
669 et al. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *eLife*. 2015; 4:e08347.
- 670 **Kraemer M**, Reiner J RC, Brady O, Messina J, Gilbert M, Pigott D, Yi D, Johnson K, Earl L, Marczak L, Shirude S,
671 Davis Weaver N, Bisanzio D, Perkins T, Lai S, Lu X, Jones P, Coelho G, Carvalho R, Van Bortel W, et al. Past
672 and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nature Microbiology*. 2019;
673 4(5):854–863. www.scopus.com.
- 674 **Krishnamurti TN**, Kishtawal CM, Zhang Z, LaRow T, Bachiochi D, Williford E, Gadgil S, Surendran S. Improved
675 weather and seasonal climate forecasts from multimodel superensemble. *Science*. 1999; 285:1548–1550.
- 676 **Krishnamurti TN**, Kumar V, Bhardwaj ASA, Ghosh T, Ross R. A review of multimodel superensemble forecasting
677 for weather, seasonal climate, and hurricanes. *Rev of Geophys*. 2016; 54(2):336–377.
- 678 **Lake IR**, Colón-González FJ, Barker G, Morbey RA, Smith GE, Elliot AJ. Machine learning to refine decision making
679 within a syndromic surveillance service. *BMC Pub Health*. 2019; 19:559.
- 680 **Lambrechts L**, Paaijmans K, Fansiri T, Carrington L, Kramer L, Thomas M, TW S. Impact of daily temperature
681 fluctuations on dengue virus transmission by *Aedes aegypti*. *Proc Natl Acad Sci U S A*. 2011; 108:7460–7465.
- 682 **Lauer S**, Sakrejda K, Ray E, Keegan L, Bi Q, Suangtho P, Hinjoy S, Iamsirithaworn S, Suthachana S, Laosiritaworn
683 Y, Cummings D, Lessler J, Reich N. Prospective forecasts of annual dengue hemorrhagic fever incidence in
684 Thailand, 2010–2014. *Proc Natl Acad Sci U S A*. 2018; 115:E2175–E2182.

- 685 **Li Y**, Kamara F, Zhou G, Puthiyakunnon S, Li C, Zhoi Y, Yao L, Yang G, Chen X. Urbanization increases *Aedes*
686 *albopictus* larval habitats and accelerates mosquito development and survivorship. *PLoS Negl Trop Dis*. 2014;
687 8:e3301.
- 688 **Lowe R**, Barcellos C, Coelho CAS, Bailey TC, Coelho GE, Graham R, Jupp T, Ramalho WM, Carvalho MS, Stephenson
689 DB, Rodó X. Dengue outlook for the World Cup in Brazil: an early warning model framework driven by real-time
690 seasonal climate forecasts. *Lancet Plan Health*. 2014; 14:619–626.
- 691 **Lowe R**, Coelho C, Barcellos C, Sá Carvalho M, De Castro Catão R, Coelho G, Ramalho M, Bailey T, Stephenson D,
692 Rodó X. Evaluating probabilistic dengue risk forecasts from a prototype early warning system for Brazil. *eLife*.
693 2016; 5:e11285.
- 694 **Lowe R**, Gasparrini A, Van Meerbeeck C, Lippi C, Mahon R, Trotman A, Rollock L, Himds A, Ryan S, Stewart-Ibarra
695 A. Nonlinear and delayed impacts of climate on dengue risk in Barbados: A modelling study. *PLoS Med*. 2018;
696 15:e1002613.
- 697 **Lowe R**, Stewart-Ibarra AM, Petrova D, García-Díez M, Borbor-Cordova MJ, Mejía R, Regato M, Rodó X. Climate
698 services for health: predicting the evolution of the 2016 dengue season in Machala, Ecuador. *Lancet Plan*
699 *Health*. 2017; 1:142–151.
- 700 **MacLachlan C**, Arribas A, Peterson KA, Maidens A, Fereday D, Scaife AA, Gordon M, Vellinga M, Williams A,
701 Comer RE, Camp J, Xavier P. Description of GloSea5: the Met Office high resolution seasonal forecast system.
702 *Q J R Met Soc*. 2015; 141:1072–01084.
- 703 **Merkle EC**, Steyvers M. Choosing a Strictly Proper Scoring Rule. *Decision Analysis*. 2013; 10:292–304.
- 704 **Mordecai EA**, Cohen JM, Evans MV, Gudapati P, Johnson LR, Lippi CA, Miazgowicz K, Murdock CC, Rohr JR, Ryan
705 SJ, Savage V, Shocket MS, Stewart-Ibarra AM, Thomas MB, Weikel DP. Detecting the impact of temperature on
706 transmission of Zika, dengue, and chikungunya using mechanistic models. *PLoS Neglected Tropical Diseases*.
707 2017; 11:e0005568.
- 708 **Moyes CL**, Vontas J, Martins AJ, Ng LC, Koou SY, Dusfour I, Raghavendra K, ao Pinto J, Corbel V, David JP, Weetman
709 D. Contemporary status of insecticide resistance in the major *Aedes* vectors of arboviruses infecting humans.
710 *PLoS Negl Trop Dis*. 2017; 11:e0005625.
- 711 **Naish S**, Dale P, Mackenzie JS, McBride J, Mengersen K, Tong S. Climate change and dengue: a critical and
712 systematic review of quantitative modelling approaches. *BMC Inf Dis*. 2014; 14:167.
- 713 **NOAA**, Monthly Atmospheric and SST Indices; 2020. <https://www.cpc.ncep.noaa.gov/data/indices/sstoi.indices>.
- 714 **Noufaily A**, Morbey RA, Colón-González FJ, Elliot AJ, Smith GE, , Lake IR, McCarthy N. Comparison of statistical
715 algorithms for daily syndromic surveillance aberration detection. *Bioinformatics*. 2019; 17:3110–3118.
- 716 **Olliaro P**, Fouque F, Kroeger A, Bowman L, Velayudhan R, Santelli AC, García D, Ramm RS, Sulaiman LH, Sánchez-
717 Tejada G, Correa-Morales F, Gozzer E, Basso-Garrido C, Quang LC, Gutiérrez G, Yadon ZE, Runge-Ranzinge
718 S. Improved tools and strategies for the prevention and control of arboviral diseases: A research-to-policy
719 forum. *PLoS Negl Trop Dis*. 2018; 12:e0005967.
- 720 **Petrova D**, Lowe R, Stewart-Ibarra A, Ballester J, Koopman S, Rodó X. Sensitivity of large dengue epidemics in
721 Ecuador to long-lead predictions of El Niño. *Clim Services*. 2019; In Press:1.
- 722 **Powell J**, Tabachnick W. History of domestication and spread of *Aedes aegypti* a review. *Memorias do Instituto*
723 *Oswaldo Cruz*. 2013; 108:11–17.
- 724 **Reich NG**, Lauer SA, Sakrejda K, Iamsirithaworn S, Hinjoy S, Suangtho P, Suthachana S, Clapham HE, Salje H,
725 Cummings DAT, Lessler J. Challenges in Real-Time Prediction of Infectious Disease: A Case Study of Dengue in
726 Thailand. *PLoS Negl Trop Dis*. 2016; 10(6):e0004761.
- 727 **Reinhold J**, Lazzari R, Lahondère C. Effects of the Environmental Temperature on *Aedes aegypti* and *Aedes*
728 *albopictus* Mosquitoes: A Review. *Insects*. 2018; 8:158.
- 729 **Reiter P**. Climate Change and Mosquito-Borne Disease. *Environ Health Perspect*. 2001; 109:141–161.
- 730 **Richardson DS**. Skill and relative economic value of the ECMWF ensemble prediction system. *Q J R Meteorol*
731 *Soc*. 2000; 126(563):649–667.

- 732 **Rue H**, Martino S, Chopin N. Approximate Bayesian Inference for Latent Gaussian Models Using Integrated
733 Nested Laplace Approximations (with discussion). *J R Stat Soc Series B Stat Methodol.* 2009; 71:319–392.
- 734 **Runge Ranzinger S**, McCall P, Kroeger A, Horstick O. Dengue disease surveillance: an updated systematic
735 literature review. *Trop Med Int Health.* 2014; 19:1116–1160.
- 736 **Saha S**, Moorthi S, Wu X, Wang J, Nadiga S, Tripp P, Behringer D, Hou YT, ya Chuang H, Iredell M, Ek M, Meng J,
737 Yang R, na Mendez MP, van den Dool H, Zhang Q, Wang W, Chen M, Becker E. The NCEP Climate Forecast
738 System Version 2. *J Climate.* 2014; 27:2185–220.
- 739 **Scaife A**, Arribas A, Blockley E, Brookshaw A, Clark RT, Dunstone N, Eade R, Fereday D, Folland CK, Gordon M,
740 Hermanson L, Knight JR, Lea DJ, MacLachlan C, Maidens A, Martin M, Peterson AK, Smith D, Vellinga M, Wallace
741 E, et al. Description of GloSea5: the Met Office high resolution seasonal forecast system. *Geophys Res Lett.*
742 2014; 41:2514–2519.
- 743 **Sedda L**, Vilela A, Aguiar E, Gaspar C, Gonçalves A, Olmo R, Silva A, De Cássia Da Silveira L, Eiras A, Drumond B,
744 Kroon E, Marques J. The spatial and temporal scales of local dengue virus transmission in natural settings: A
745 retrospective analysis. *Parasites and Vectors.* 2018; 11(1). www.scopus.com.
- 746 **Shepard D**, Undurraga E, Betancourt-Cravioto M, Guzmán MG, Halstead S, Harris E. Approaches to Refining
747 Estimates of Global Burden and Economics of Dengue. *PLoS Negl Trop Dis.* 2014; 8:e3306.
- 748 **Shepard D**, Undurraga E, Halasa Y, Stanaway J. The global economic burden of dengue: a systematic analysis.
749 *The Lancet Inf Dis.* 2016; 16:935–941.
- 750 **Siebert S**. SpecsVerification: Forecast Verification Routines for Ensemble Forecasts of Weather and Climate;
751 2017, <https://CRAN.R-project.org/package=SpecsVerification>, r package version 0.5-2.
- 752 **Simon-Oke IA**, Olofintoye LK. The Effect of Climatic Factors on the Distribution and Abundance of Mosquito
753 Vectors in Ekiti State. *J Biol Agric Healthc.* 2015; 5:142–146.
- 754 **Simpson D**, Rue H, Riebler A, Martins TG, Sørbye SH, et al. Penalising model component complexity: A principled,
755 practical approach to constructing priors. *Statistical science.* 2017; 32(1):1–28.
- 756 **Stewart Ibarra A**, Ryan S, Beltrán E, Mejía R, Silva M, Muñoz A. Dengue Vector Dynamics (*Aedes aegypti*)
757 Influenced by Climate and Social Factors in Ecuador: Implications for Targeted Control. *PLoS One.* 2013;
758 8:e78263.
- 759 **Stolerman LM**, Maia PD, Kutz JN. Forecasting dengue fever in Brazil: An assessment of climate conditions. *PLoS*
760 *One.* 2019; 14:e0220106.
- 761 **Thi KLP**, Briant L, Gavotte L, Labbe P, Perriat-Sanguinet M, Cornillot E, Vu TD, Nguyen TY, Tran VP, Nguyen VS,
762 Devaux C, Afelt A, Tran CC, Phan TN, Tran ND, Frutos R. Incidence of dengue and chikungunya viruses in
763 mosquitoes and human patients in border provinces of Vietnam. *Parasites & Vectors.* 2017; 10:556.
- 764 **Thomson MC**, Mason S, Platzer B, Mihretie A, Omumbo J, Mantilla G, Ceccato P, Jancloues M, Connor S. Climate
765 and health in Africa. *Earth Perspectives.* 2014; 1:17.
- 766 **Thornes JE**, Stephenson DB. How to judge the quality and value of weather forecast products. *Meteorol Appl.*
767 2001; 8:307–314.
- 768 **Tompkins A**, Colón-González F, Di Giuseppe F, Namanya D. Dynamical Malaria Forecasts Are Skillful at Regional
769 and Local Scales in Uganda up to 4 Months Ahead. *Geohealth.* 2019; 3:58–66.
- 770 **Tompkins A**, Di Giuseppe F. Potential Predictability of Malaria in Africa Using ECMWF Monthly and Seasonal
771 Climate Forecasts. *J Appl Meteor Climatol.* 2015; 54:521–540.
- 772 **Tsunoda T**, Cuong T, Dong T, Yen N, Le N, Phong T, Minakawa N. Winter Refuge for *Aedes aegypti* and *Ae.*
773 *albopictus* Mosquitoes in Hanoi during Winter. *PLoS One.* 2014; 9:e95606.
- 774 **Wan Z**, Hook S, Hulley G, MYD11A1 MODIS/Aqua Land Surface Temperature/Emissivity Daily L3 Global 1km SIN
775 Grid V006; 2015. <https://lpdaac.usgs.gov/products/myd11a1v006/>.
- 776 **Watts D**, Burke D, Harrison B, Whitmire R, Nisalak A. Effect of temperature on the vector efficiency of *Aedes*
777 *aegypti* for dengue 2 virus. *Amer J of Trop Med Hyg.* 1987; 36(1):143–152. www.scopus.com.
- 778 **World Health Organization**. Background Paper on dengue vaccines. Geneva, Switzerland: World Health
779 Organization; 2018.

- 780 **WorldPop**, The spatial distribution of population in Vietnam; 2018. <https://cds.climate.copernicus.eu/#!/home>.
- 781 **Yamana T**, Kandula S, Shaman J. Superensemble forecasts of dengue outbreaks. *J R Soc Interface*. 2016;
782 13:20160410.
- 783 **Yamana T**, Kandula S, Shaman J. Individual versus superensemble forecasts of seasonal influenza outbreaks in
784 the United States. *PLoS Comput Biol*. 2017; 13(11):e1005801.

785 **Appendix 1**

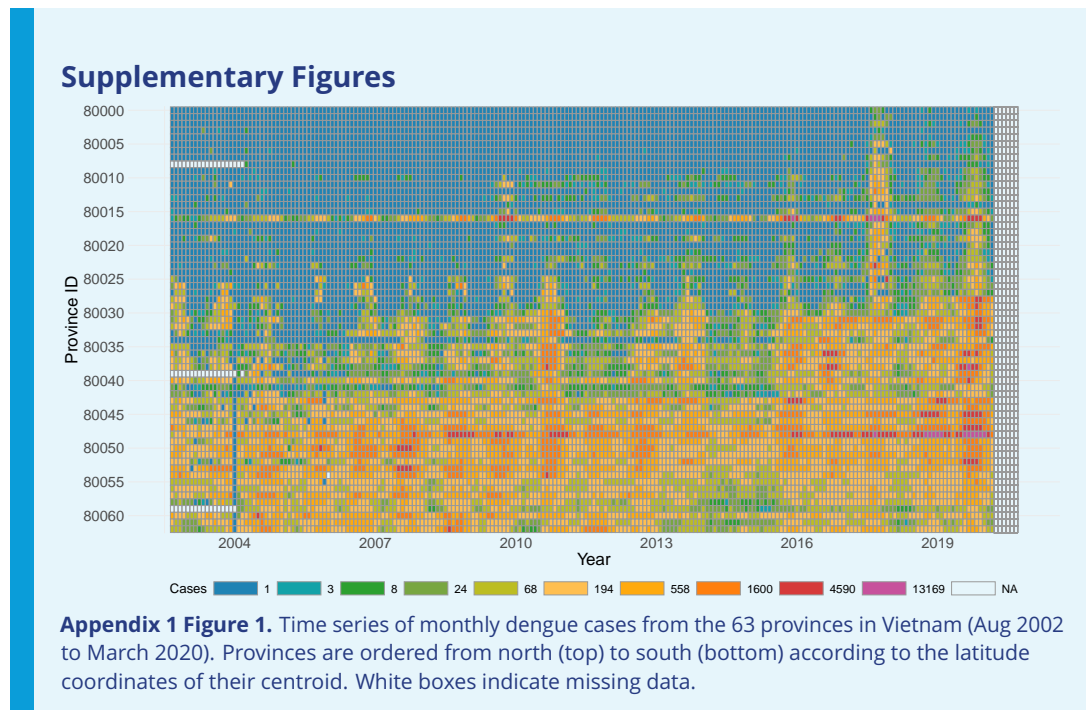
786

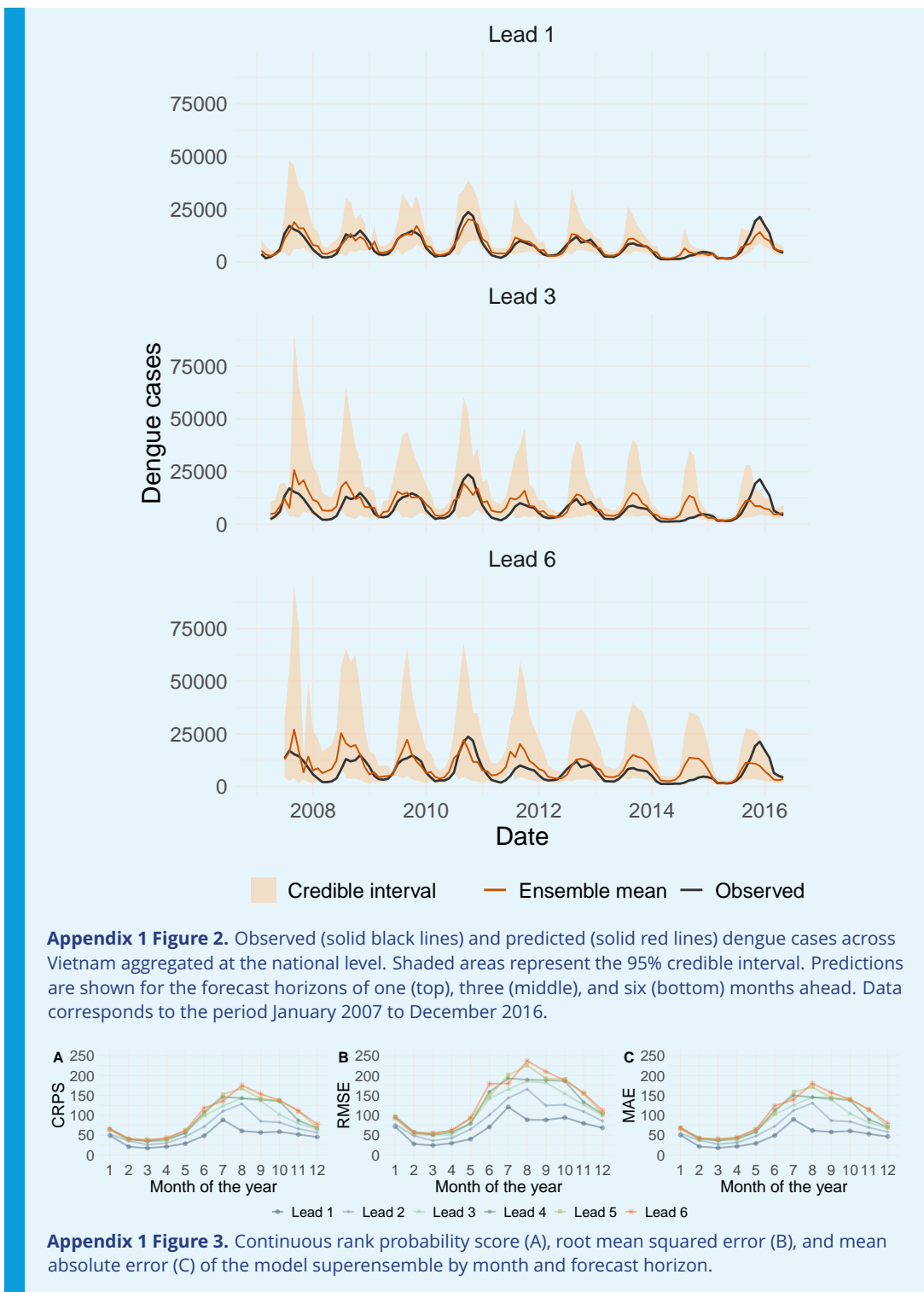
787

788

789

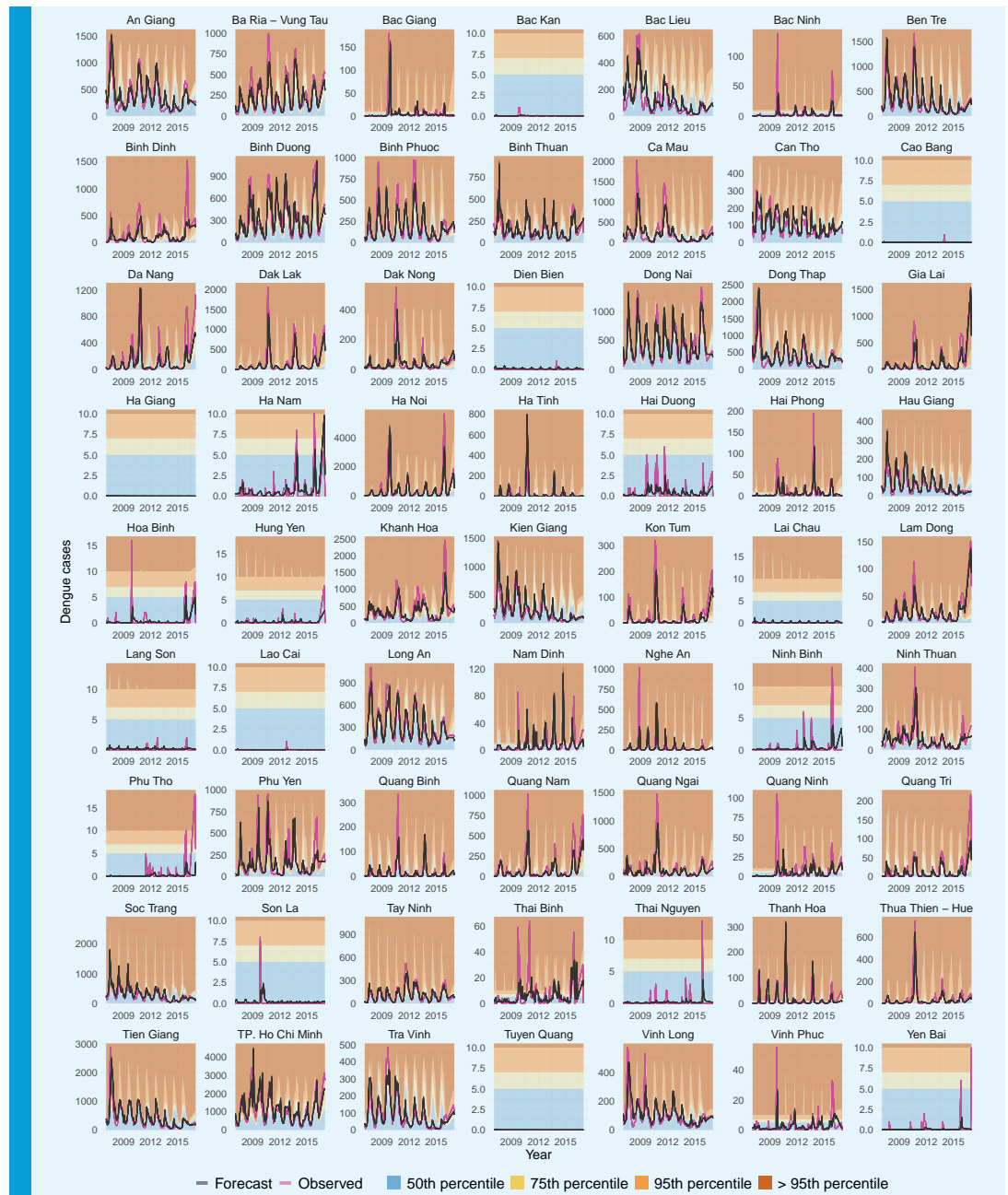
790





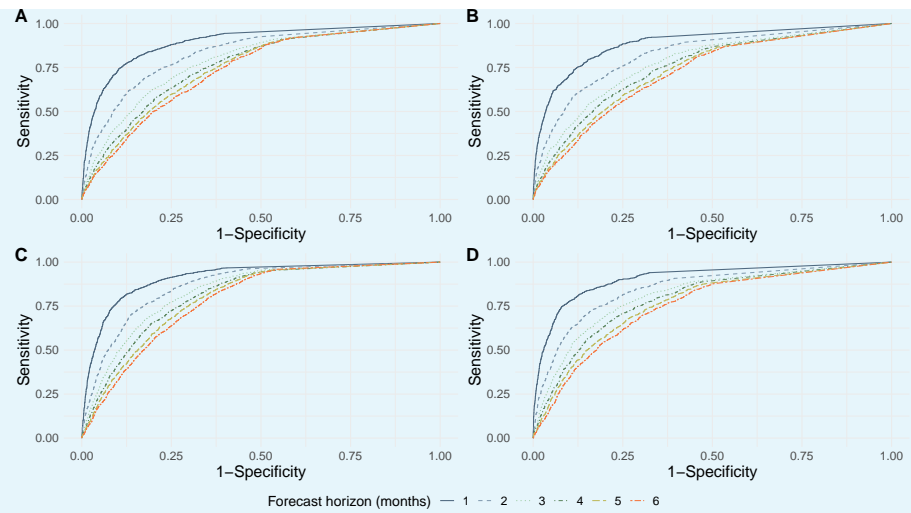
792
793
794
795
796

798
799
800



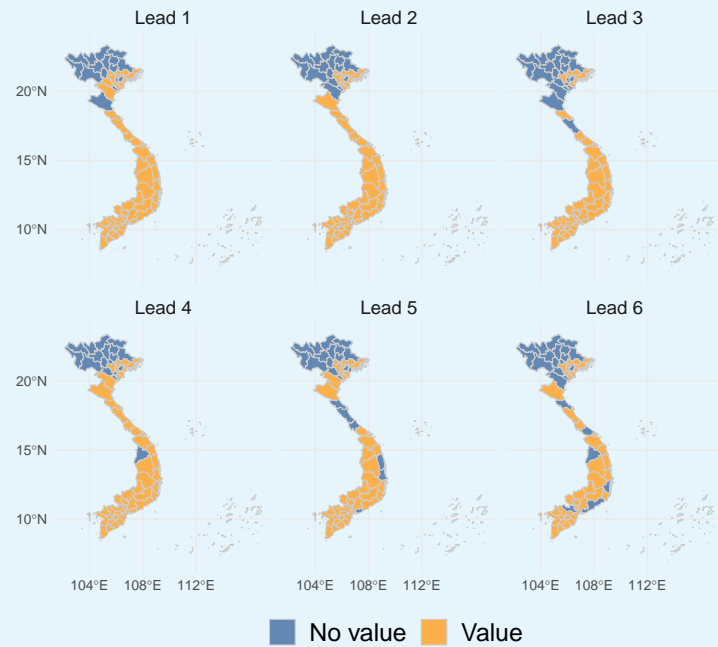
802
803
804
805
806

Appendix 1 Figure 4. Observed (solid black lines) and predicted (solid blue lines) dengue cases across 63 Vietnamese provinces computed one month ahead. Data corresponds to the period January 2007 to December 2016. Shaded areas represent a range of month- and province-specific percentiles to guide public health decision-making.



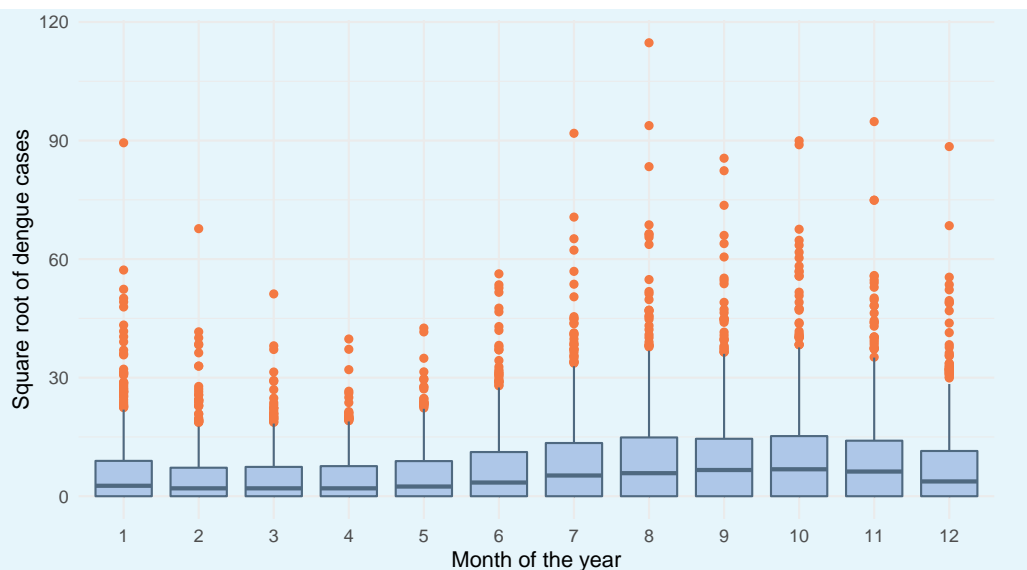
808
809
810
812

Appendix 1 Figure 5. Area under the curve (AUC) for different probabilities of exceeding four different moving outbreak thresholds: one (A) and two (B) standard deviations above the mean, 75th (C) and 95th (D) percentiles.



813
814
815
816
818

Appendix 1 Figure 6. Spatial variation of the relative economic value of the model superensemble over the forecast horizon of one to six months. Orange shaded areas indicate provinces where there is relative economic value (based on a range of theoretical cost-loss ratios and outbreak thresholds). Blue shaded areas indicate provinces where the superensemble had no relative economic value.



819
820
821
822
823
824
825
826

Appendix 1 Figure 7. Month-specific variability in dengue cases across Vietnam. The X axis indicates the month of the year. The Y axis indicates increases in the number of dengue cases (square root transformed). The upper and lower limits of each box represent the inter-quartile range of the distribution of dengue cases for each month. The middle solid line indicates the median value. The upper and lower whiskers indicate the maximum and minimum values of the dengue case distribution (excluding outliers which are indicated in orange circles). Outliers are values beyond ± 1.5 times the inter-quartile range.