Spatial temporal distribution of COVID-19 risk during the early phase of the pandemic in Malawi

Alfred Ngwira Corresp., 1 , **Felix Kumwenda** 1 , **Eddons Munthali** 1 , **Duncan Nkolokosa** 1

1 Basic Sciences Department, Lilongwe University of Agriculture and Natural Resources, Lilongwe, Malawi

Corresponding Author: Alfred Ngwira
Email address: alfngwira@yahoo.com

Background: COVID-19 has been one of the greatest challenges the world has faced since the second world war.This study aimed at investigating the distribution of COVID-19 in both space and time in Malawi. **Methods:** The study used publicly available data of COVID-19 cases for the period from April 2 to October 28, 2020. Semiparametric spatial temporal models were fitted to the number of monthly confirmed cases as an outcome data, with time and district as independent variables, where district was the spatial unit, while accounting for sociodemographic factors. **Results:** The study found significant effects of location and time, with the two interacting. The spatial distribution of COVID-19 risk showed major cities being at greater risk than rural areas. Over time, the COVID-19 risk was increasing then decreasing in most districts with the rural districts being consistently at lower risk. High proportion of elderly people was positively associated with COVID-19 risk ($\beta = 1.272$, 95% CI= 0.171; 2.370) than low proportion of elderly people. There was negative association between poverty incidence and COVID-19 risk ($\beta = -0.100$, 95% CI = -0.136; -0.065). **Conclusion:** Future or present strategies to limit the spread of COVID-19 should target major cities and the focus should be on time periods that had shown high risk. Furthermore, the focus should be on elderly and rich people.

1 **Spatial temporal distribution of COVID-19 risk during the early**
2 **phase of the pandemic in Malawi**

3 Alfred Ngwira[1], Felix Kumwenda[1], Eddons Munthali[1], Duncan Nkolokosa[1]

4 [1]Basic Sciences Department, Lilongwe University of Agriculture and Natural Resources,
5 Lilongwe, Malawi.

6 Corresponding Author:

7 Alfred Ngwira[1],

8 Bunda College, Lilongwe, Malawi.

9 Email: alfngwira@yahoo.com

11 **Abstract**

12 **Background:** COVID-19 has been one of the greatest challenges the world has faced since the
13 second world war. This study aimed at investigating the distribution of COVID-19 in both space
14 and time in Malawi.

15 **Methods:** The study used publicly available data of COVID-19 cases for the period from April 2
16 to October 28, 2020. Semiparametric spatial temporal models were fitted to the number of
17 monthly confirmed cases as an outcome data, with time and district as independent variables,
18 where district was the spatial unit, while accounting for sociodemographic factors.

19 **Results:** The study found significant effects of location and time, with the two interacting. The
20 spatial distribution of COVID-19 risk showed major cities being at greater risk than rural areas.
21 Over time, the COVID-19 risk was increasing then decreasing in most districts with the rural
22 districts being consistently at lower risk. High proportion of elderly people was positively
23 associated with COVID-19 risk (β = 1.272, 95% CI= 0.171; 2.370) than low proportion of
24 elderly people. There was negative association between poverty incidence and COVID-19 risk (β
25 = -0.100, 95% CI = -0.136; -0.065).

26 **Conclusion:** Future or present strategies to limit the spread of COVID-19 should target major
27 cities and the focus should be on time periods that had shown high risk. Furthermore, the focus
28 should be on elderly and rich people.

29   **Key words:** spatial epidemiology, spatial risk, COVID-19 spread, spatiotemporal modeling

30   Introduction

31   COVID-19 is a corona virus disease, which was first reported in Wuhan, China, in 2019. It is

32   characterized by severe acute respiratory syndrome (SARS), hence also known as SARS-CoV-2

33   (WHO, 2020). It was declared a pandemic by WHO on March 11, 2020, and it was first reported

34   in Africa, in Egypt on February 14, 2020. The first case in Malawi was reported on April 2,

35   2020. Since its onset, COVID-19 has been one of the greatest disease pandemics of recent times.

36   It has not only caused a great loss to life, but has also impacted negatively to the global

37   economy, through severe disruption of international travel, tourism and trade (Pak et al., 2020).

38   COVID-19 is highly transmissible, but less deadly. For example, compared to human

39   immunodeficiency virus (HIV), as of June 29, 2020, 6 months after COVID-19 emerged in

40   December, 2019, there were close to 10 million confirmed cases, and at least 500 thousand

41   deaths worldwide (Petersen et al., 2020), compared to 3.3 million new HIV annual cases, and 1.8

42   million deaths during the peak periods in 1997 and 2005 (GBD 2015 HIV Collaborators, 2016).

43   Based on the reproductive rate (R_0), transmissibility of COVID-19 is also higher than other

44   viruses, where the reproductive rate (R_0) of COVID-19 is estimated at 2.5 (1.8-3.6) compared to

45   2-3 for SARS-CoV and 1918 influenza pandemic, 0.9 for middle east respiratory syndrome

46   coronavirus (MERS-CoV), and 1.5 for 2009 influenza pandemic (Petersen et al., 2020).

47   The spread of COVID-19 is driven by environmental factors like temperature and

48   humidity, where cases spike with low temperature and humidity (Mecenas et al., 2020; Islami &

49   Jalili, 2020; Bherwani et al., 2020). Demographic and socioeconomic factors like population

50   density, age distribution, income, number of tourists, also play a role in the spread of COVID-19

51   cases, where high cases are observed with high density, age, income and large number of tourists

52   (Tantrakarnapa et al., 2020). Density has also been found not to be correlated with COVID-19

53   cases (Hamid et al., 2020), contrary to the belief of many that it is positively correlated with

54   COVID-19 cases. Exposure to air pollution is another possible driver of COVID-19 risk

55   distribution (Sarmadi et al., 2020), with increased cases and deaths, where there is high air

56   pollution.

57   Understanding disease space and time dynamics is important for the epidemiologists as

58   with space distribution, the hot spot areas are marked for intervention (Snow, 1855). In addition,

59 possible drivers of the epidemic in those hot spots are suggested for further etiological

60 investigation. Regarding temporal distribution, times with high disease risk are also identified

61 which gives crew to possible causes including, in particular, seasonal changes. A number of

62 studies on spatial temporal distribution of COVID-19 have been conducted (Chen et al., 2020;

63 Yang et al., 2020; Gayawan et al., 2020; Martellucci et al., 2020; Adekunle et al., 2020; Arashi

64 et al., 2020; Kumar et al., 2020; Briz-Redón & Serrano-Aroca, 2020; Diop et al., 2020; Jia et al.,

65 2020; Xie et al., 2020; Likassa, 2020; Sarfo & Karuppannan, 2020; Daw et al., 2020; Ye & Hu,

66 2020). The majority of these though, have used the geographical information system (GIS)

67 technology as compared to statistical modelling using spatial temporal models. A few studies

68 that have used the statistical approach of spatial temporal analysis to my knowledge are

69 Gayawan et al (2020), who used the Possion hurdle model to take into account excess zero

70 counts of COVID-19 cases, Briz-Redon & Serrano Aroca (2020), who used the separable

71 random effects model with structured and unstructured area and time effects, and Chen et al.

72 (2020), who used the inseparable spatial temporal model. In addition, in Africa, spatial temporal

73 analysis of COVID-19 cases has been limited (Gawayan et al., 2020; Arashi et al., 2020;

74 Adekunle et al., 2020). In Malawi, at the time of this study, no study on spatial temporal

75 distribution of COVID-19 cases had been spotted. Only one study that focused on prediction of

76 COVID-19 cases using mathematical models was seen (Kuunika, 2020). The aim of this study

77 was to determine the spatial temporal trends of COVID-19 confirmed cases in Malawi while

78 using the spatial temporal statistical models. The objectives of the study were:

79 - To establish the estimated or predicted risk trend by geographical location.

80 - To estimate the temporal risk trend of COVID-19 by geographical location.

81   The article has been organized as follows. First, study methods are described in terms of

82 data collection and statistical analysis. Thereafter, results and discussion of results are presented.

83 Finally, conclusions and implications regarding the findings of the study are made.

84

85 **Materials and Methods**

86 **Data**

87 The study used the publicly available COVID-19 monthly confirmed cases data for Malawi,

88 which were extracted from the Public Health Institute of Malawi situation reports

89 (https://malawipublichealth.org/index.php/resources/covid-19-sitrep-updates/detail). The study

90   also collected data on demographic and socioeconomic indicators of each district from the

91   Malawi data portal (https://malawi.opendataforafrica.org/lilehtd/covid-19), which were used as

92   independent variables. The following were the demographic and socioeconomic indicators

93   included in the study, whose choice was mainly based on the literature (e.g Pequeno et al., 2020;

94   Hamidi et al., 2020): total population, population density, percentage of people with running

95   water, percentage of people with no toilet, percentage of people who are aged 65 years and

96   above, poverty incidence and percentage of people with no radio. The population size for each

97   district was used as the expected number of people to be infected in each district. The study

98   period run from April 2, the day the first case was recorded to October 28, the last day in the

99   month of October, the COVID-19 situation report was released. The study period was divided

100   into 6 months as follows: April (02/04/20 – 30/04/20), May (01/05/20 – 29/05/20), June

101   (30/05/20 – 30/06/20), July (01/07/20 – 18/07/20), August (19/07/20 – 28/08/20), September

102   (29/08/20 – 30/09/20) and October (01/10/20-28/10/20).

103   **Statistical analysis**

104   Descriptive analysis involved the time series plot of the cumulative confirmed cases and those

105   who had died of COVID-19 for the whole country from the first month the epidemic began to the

106   month of October, 2020. Multiple variable spatial temporal models were then fitted in R using

107   the Bayesian approach while using the integrated nested laplace approximations (INLA). INLA

108   estimation of Bayesian hierarchical models is proved to be faster than the use of Markov Chain

109   Monte Carlo (MCMC) methods. A spatial temporal model assuming linear trend of cases over

110   time as proposed by Bernardinelli et al. (1995) was first fitted. That is, let y_{it} be the observed

111   disease cases in region i and at time t, then the observed disease cases have the Poisson

112   distribution with the data model defined as $y_{it} \sim Poisson\ (e_{it}r_{it})$, where $e_{it}r_{it}$ is the mean of

113   disease cases, e_{it} and r_{it} are the expected disease cases and relative risk respectively. The model

114   of the relative risk is then defined as: $\log\ (r_{it}) = \eta_{it}$, where η_{it} is the predictor specified as:

115   $\eta_{it} = \mu + u_i + v_i + (\emptyset + \varphi_i).$ (1)

116   The model with predictor (1) will be denoted by A. In (1), μ is the overall disease relative risk, u_i

117   is the area level unstructured random effect where area level unit was district, v_i is the area level

118   structured random effect, $\emptyset$ is overall time trend and φ_i is the area specific time trend. The area

119   level random effects represent the proxy of unmeasured or unobserved risk factors of COVID-19

120   in each area, for example, number of lock-downs and average temperature of each area. To

121   incorporate the observed covariates like population density in each area in the model, (1) can be

122   written as:

123    $\eta_{it} = \mu + \omega^T\theta + u_i + v_i + (\emptyset + \varphi_i)t.$ (2)

124   In (2), ω is a vector of the observed covariates and θ are the corresponding coefficients. The

125   unstructured area level effects were modelled by the independent normal distribution with zero

126   mean, that is, $u_i \sim N(0, \sigma_u^2)$, and the structured random effects were assigned the intrinsic

127   conditional auto-regressive (ICAR) according to Besag (1991), that is,

128    $v_i|v_{je\theta_i} \sim N\left(\frac{\Sigma_{je\theta_i} v_j w_{ij}}{\Sigma_{je\theta_i} w_{ij}}, \frac{\sigma_v^2}{\Sigma_{je\theta_i} w_{ij}}\right).$ The weakness of model A is the linearity assumption on the effect

129   of time on the relative risk of the disease. To take a more flexible approach on the effect of time,

130   nonlinear spatial temporal models were also explored. The predictor, η_{it}, or the nonlinear spatial

131   temporal model for the time effect is specified as follows:

132    $\eta_{it} = \mu + \omega^T\theta + u_i + v_i + \gamma_t + \beta_t + \delta_{it}.$ (3)

133   The model with predictor (3) will be denoted by B. The u_i and v_i in the model are still the area

134   level unstructured and structured effects respectively as defined in (1), and γ_t and β_t are the

135   unstructured and structured temporal effects. The unstructured time effects were modelled by the

136   independent normal distribution with zero mean, that is, $\gamma_t \sim N(0, \sigma_\gamma^2)$, and the structured temporal

137   effects were assigned the first order random walk prior distribution defined as:

138    $\pi(\beta|\sigma_\beta^2) \propto exp\left(-\frac{1}{2\sigma_\beta^2}\Sigma_{t=2}^T(\beta_t - \beta_{t-1})^2\right).$ A second order random walk was also explored in

139   case the data would show a more pronounced linear trend. The second order random walk is

140   defined as: $\pi(\beta|\sigma_\beta^2) \propto exp\left(-\frac{1}{2\sigma_\beta^2}\Sigma_{t=2}^T(\beta_t - 2\beta_{t-1} + \beta_{t-2})^2\right).$ The last term in (3), δ_{it},

141   represents the interaction between area and time. Four forms of interaction between space and

142   time are possible according to Knorr-Held (2000). The first form of interaction assumes

143   interaction between the unstructured region effect, u_i, and the unstructured temporal effect, γ_t,

144   we denote it by model B 1, and in this case the interaction effect is assigned the independent

145   normal distribution, that is, $\delta_{it} \sim N(0, \sigma_\delta^2)$. The second type of interaction, is the interaction of

146   structured area effect, v_i, and the unstructured temporal effect, γ_t, we denote it by model B 2.

147   This form of interaction assumes conditional intrinsic (CAR) distribution for the areas for

148   each time independently from all the other times. The third is the interaction between

149   unstructured area random effect, u_i, and the structured temporal effect, β_t, we denote this by

150   model B 3. The prior distribution for each area is assumed to be a second order random walk

151   across time. The last possible space time interaction is that between area structured effect, v_i,

152   and the structured time effect, β_j, we call it model B 4. In this case, the second order random

153   walk prior that depends on neighboring areas was assigned for each area. The prior for the

154   variance parameters was the log-Gamma (1, 0.0005), which is considered as the default

155   minimal informative prior (Blangiardo et al., 2013). The fixed effects were assigned the default

156   normal, that is, N (0, 0). The model choice was by the deviance information criteria (DIC) as

157   proposed by Spiegelhalter et al. (2002), where a smaller DIC means a better model in terms of

158   fit and complexity. It is the sum of the measure of model fit, called the deviance, denoted by

159     $\overline{D}$ and the effective number of parameters denoted by $P_{\overline{D}}$. Since the fitted model was fully

160   Bayesian, the selected model was checked for sensitivity by checking the sensitivity of the

161   hyper-parameters for changes in their prior distributions (Roos & Held, 2011). The log-

162   Gamma (1, 0.01) and log-Gamma (0.5, 0.001) priors were the alternative priors in the

163   sensitivity analysis of the hyper-parameters. The selected model was then used to estimate the

164   relative risk, r_{it}

Results

166   Figure 1 shows the graph of cumulative confirmed and dead cases of COVID-19 from the

167   time the first case was reported, April 2, to October 28, 2020. There were 5904 confirmed

168   cases and 184 deceased cases as of October 28, 2020. Generally, there were low total cases of

169   those who had died of COVID-19 as compared to those who were confirmed.

170   Table 1 shows the DIC of the fitted spatial temporal models defined in the methods

171   section. Model B 11 had smallest DIC than the rest of the models, though it was not

172   significantly different from model B 12, since the DIC difference was not greater than 3

173   (Spiegelhalter et al., 2002). Sensitivity analysis of model B11 showed that the hyper-

174   parameters posterior distributions were generally stable for small changes in their prior

175   distributions. For example, Figure 2 shows similar posterior distributions of the variance of

176   unstructured effect of area, and variance of interaction effect of area and time for different

177   prior distributions: log-Gamma (1, 0.0005), log-Gamma (1, 0.01) and log-Gamma (0.5,

178   0.001). The first two priors resulted into posteriors which were almost super-imposing, that is,

179   one laid over the other. The results of model B11 are therefore presented and discussed. Table

180   2 presents results of model B 11. The proportion of elderly people (65 + years) was positively

181   associated with COVID-19 (β = 1.272, 95 % CI= 0.171; 2.370). The positive coefficient of

182   elderly people means that high proportion of elderly people compared to low proportion of

183   elderly people is associated with increased risk of COVID-19. Poverty incidence was negatively

184   associated with COVID-19 risk (β = -0.100, 95 % CI = -0.136; -0.065). Negative coefficient of

185   poverty incidence means that high poverty incidence is associated with reduced risk of being

186   diagnosed with COVID-19 as compared to low poverty incidence. All other sociodemographic

187   factors were not significant determinants of COVID-19 confirmed cases, since the p-values

188   were greater than 0.05. All the random effects, that is, the effect of time, the effect of location,

189   including the interaction effect of location and time, were significant predictors of the risk of

190   contracting COVID-19, since the estimated variance parameters were significantly greater

191   than zero, considering that the confidence intervals excluded zero. Area level spatial effects

192   and the effects of time, modelled by the second order random walk prior (RW2), were highly

193   significant as evidenced from the large variance parameters.

194   The spatial temporal distribution of overall fitted risk (Figure 3 and Figure 4), shows

195   that by space (Figure 3), areas in and around the cities like Mzimba in the northern region,

196   Lilongwe in the center and Blantyre in the south, were at increased risk of being confirmed of

197   COVID-19 compared to the rest of the areas. Over time, from April to October (Figure 4), the

198   risk in all districts was increasing from April to August, where it started to decline to October.

199   Blantyre was consistently at high risk to COVID-19 over time compared to other areas. Most

200   of the districts in the rural areas, were consistently at low risk of contracting COVID-19 over

201   time compared to urban districts. Figure 5 presents the spatial risk of contracting COVID-19.

202   Spatial risk represents the residual risk due to unobserved or unmeasured factors of COVID-

203   19. In general, the spatial risk looks high in many areas in June, July, August and a little bit

204   high in September compared to April, May and October.

Discussion

205

206   The study looked at spatial temporal distribution of COVID-19 in Malawi from the time the

207   first case was recorded, April 2, to October 28, 2020, while using the inseparable statistical

208   spatial temporal model. The use of inseparable model, allowed the investigation of the joint or

209   interaction effect of time and location on the risk of being confirmed of COVID-19. The use

210   of nonparametric model for time effect (RW2), also enabled the capturing of the subtle

211   influences of time on the risk of contracting COVID-19.

212   The study finds significant effect of poverty incidence and proportion of elderly (65 +

213   years) people on COVID-19 risk. The results on effect of poverty on COVID-19 risk are

214   consistent with the literature (Tantrakarnapa et al., 2020; Sarmadi et al., 2020), where it was

215   found that high income compared to low income was positively associated with COVID-19

216   cases. Low risk of COVID-19 among poor people than rich people may be due to the fact that

217   such people are less likely to travel to cities and abroad where they can be exposed to the

218   disease. In addition, low COVID-19 risk among poor people may be due to the fact that in

219   Malawi and in poor countries in general, disease monitoring or surveillance inform of testing

220   is very minimal, which results in less people to be diagnosed of the disease. The higher risk of

221   COVID-19 for higher income than lower income entails the same relationship between age

222   and COVID-19 since wealth is a proxy for age as wealth is accumulated over time. The

223   increased risk of COVID-19 for elderly people than young people is due to the fact that

224   elderly people have a weaker immune system than the youths and hence, they are susceptible

225   to infection. Of special interest, the results on population density show negative correlation

226   with COVID-19 infection rate and that the association is not significant. This is contrary to

227   the common understanding that a virus is positively associated with density. Of course, this is

228   not strange, considering that previous studies have found mixed results, with some finding

229   negative association (Carozzi et al., 2020), and others finding positive association (Hamid et

230   al., 2020; Pequeno et al., 2020; Sarmadi et al., 2020), and that in many cases, there has been

231   no significant association. It is argued that the negative effect of density on COVID-19

232   infection rate or lack of significant positive association may be attributed to some mediating

233   factors that offset the usual positive association. For example, high density areas may attract

234   young folks who are usually resistant to COVID-19 infection. In addition, high density

235   populations are positively correlated with proxies of social distancing (Carozzie et al., 2020)

236   due to the increased consciousness of being at risk, which in turn reduces COVID-19 cases.

237   The spatial distribution of COVID-19 risk in Malawi at the given time period, shows

238   the cities and the surrounding areas being at increased risk than the rural areas. The

239   explanation to the observed spatial gradient is a matter of conjecture. One possible factor

240   driving the observed spatial pattern would be the population size and population density. The

241   cities have higher population density than the rural, and COVID-19 is therefore more likely to

242   spread fast through the movement and frequent contact between people. Case comparison

243   investigations have found positive correlation of population density and COVID-19 (Penerliev &

244   Petkov, 2020), where for example, in Italy, Lombardy, the population density is three times

245   higher than in Piedmont and the incidence rate was also over three times higher than in

246   Piedmont. Evidence of high population density as a risk factor of disease transmission has

247   been seen in India, where influenza transmission rates have been found to increase above a

248   population density of 282 people per square kilometer (African Centre of Strategic Studies,

249   2020).

250   The other possible contributor to the observed rural city spatial gradient of COVID-19

251   risk in Malawi, would be international exposure. In this case, cities have higher international

252   exposure than the rural through international flights among others, which would mean more

253   imported cases. Evidence of international exposure as a risk factor of COVID-19 transmission

254   has been observed in Africa as a whole, where countries with high international exposure like

255   South Africa, Nigeria, Morocco, Egypt, and Algeria have had higher COVID-19 cases than

256   their counterparts. International exposure as a fuel of COVID-19 transmission has also been

257   documented in Brazil where it was found that cases increased with increase in international

258   flights jetting into the country (Pequeno et al., 2020). The ability to testing may be another

259   catalyst of COVID-19 risk distribution in Malawi, where in the urban centers, a large number of

260   people are tested, and hence, more people are reported to have COVID-19 than in the rural

261   setting. COVID-19 testing rate has been confirmed to be positively correlated with virus

262   infection rate (Hamid et al., 2020).

263   Regarding the temporal distribution, a spike of the disease risk from June to August,

264   might be attributed to the effects of unrestricted presidential election political rallies which were

265   being held in most cases, in urban centers. The presidential election was held on June 23, 2020.

266   In addition, the rise in COVID-19 cases during this time would be attributed to the decreasing

267   temperatures at this time of the year as this is the time of cold season. The decline of disease risk

268   from August to October may be due to the increasing temperatures, as this time, marks the

269 beginning of hot season. Negative correlation between COVID-19 cases and temperature has
270 been documented (Pequeno et al., 2020).
271   The study did not go without weaknesses. The first weakness was that, due to the absence
272 of population size for each area at each time point, the base population at risk for each area was
273 assumed to be constant across time which was not practically valid. The other weakness was that
274 the study did not take into account the effect of other important factors of COVID-19 like the
275 number of lock-downs and average temperature of each district, while investigating the spatial
276 temporal distribution of the risk. Nevertheless, the results of the study are still valid, as all
277 unmeasured factors were captured through the area level and time random effects which were
278 considered as proxies for such factors.

279 **Conclusion**
280 The study found a significant effect of both location and time on COVID-19 risk, and the
281 effect of either of the two depended on the other, that is, interaction. The risk of COVID-19
282 for major cities and the surrounding areas was high compared to the rural districts and that
283 over time, the risk for rural areas remained relatively lower than in cities. The risk of
284 contracting COVID-19 was increasing from June to August and started to decline thereafter.
285 Elderly people were positively correlated with COVID-19 infection rate, and poverty
286 incidence had a negative effect on COVID-19 risk. The implications of the study are that
287 future interventions to halt the disease transmission, should target the major cities like
288 Blantyre, Zomba, Mangochi, Lilongwe and Mzuzu, and that by time, attention should be paid
289 to the month of June, July and partly August, when it is very cold. Furthermore, attention
290 should be paid to elderly people and rich people.

291

292 **Acknowledgements**
293 None

294

295 **References**
296 Adekunle IA, Onanuga AT, Akinola OO, Ogunbanjo OW. 2020. Modelling spatial variations of
297   coronavirus disease (COVID-19) in Africa. *Science of the Total Environment* 729:
298   138998.

299   African Centre for Strategic Studies. 2020. Mapping Risk Factors for the Spread of COVID-19
300   in Africa. Available at: https://africacenter.org/spotlight/mapping-risk-factors-spread-
301   covid-19-africa/

302   Arashi M, Bekker A, Salehi M, Millard S, Erasmus B, Cronje T, Golpaygani M. 2020. Spatial
303   analysis and prediction of COVID-19 spread in South Africa after lockdown. Available
304   at: https://arxiv.org/ftp/arxiv/papers/2005/2005.09596.pdf.

305   Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M. 1995. Bayesian analysis of
306   space-time variation in disease risk. *Statistics in Medicine* 14: 2433–2443.

307   Besag J, York J, Mollié A. 1991. Bayesian image restoration with two applications in spatial
308   statistics. *Annals of the Institute of Statistical Mathematics* 43(1): 1–59.

309   Bherwani H, Gupta A, Anjum S , Anshul A, Kumar R. 2020. Exploring dependence of COVID-
310   19 on environmental factors and spread prediction in India. *npj Climate and Atmospheric*
311   *Science* 3:38.

312   Blangiardo M, Cameletti M, Baio G, Rue H. 2013. Spatial and spatio-temporal models with R-
313   INLA. *Spatial and Spatio-temporal Epidemiology* 7: 39 – 55.

314   Briz-Redón A, Serrano-Aroca A. 2020. A spatio-temporal analysis for exploring the effect of
315   temperature on COVID-19 early evolution in Spain. *Science of the Total Environment*
316   728: 138811.

317   Carozzi F, Provenzano S, Roth S. 2020. Urban Density and COVID-19. IZA – Institute of Labor
318   Economics. IZA DP No. 13440. Available at: https://www.iza.org/publications/dp/13440.

319   Chen Z, Zhang Q, Lu Yi, Guo Z, Zhang X, Zhang W, Guo C, Liao C, Li Q, Han X, Lu J. 2020.
320   Distribution of the COVID-19 epidemic and correlation with population emigration from
321   Wuhan, China. *Chinese Medical Journal* 133(9)

322   Daw MA, El-Bouzedi AH, Ahmed MO, Cheikh Y. 2020. Spatial Distribution and Geographic
323   Mapping of COVID-19 in Northern African Countries; A Preliminary Study. *J Clin*
324   *Immunol Immunother* 6: 032.

325   Diop BZ, Ngom M, Biyong CP, Biyong JNP. 2020. The relatively young and rural population
326   may limit the spread and severity of COVID-19 in Africa: a modelling study. *BMJ*
327   *Global Health* 5: e002699.

328   Eslami H, Jalili M. 2020. The role of environmental factors to transmission of SARS-CoV-2
329   (COVID-19) *AMB Expr* 10:92

330   Gayawan E, Awe OO, Oseni BM, Uzochukwu IC, Adekunle A, Samuel G, Eisen DP, Adegboye
331   OA. 2020. The spatiotemporal epidemic dynamics of COVID-19 outbreak in Africa.
332   *Epidemiology and Infection* 148: 1–11.

333   GBD 2015 HIV Collaborators. 2016. Estimates of global, regional, and national incidence,
334   prevalence, and mortality of HIV, 1980–2015: the Global Burden of Disease Study 2015.
335   *Lancet HIV* 3: e361–87

336   Hamidi S, Sabouri S, Ewing R. 2020. Does Density Aggravate the COVID-19 Pandemic?
337   *Journal of the American Planning Association*. DOI: 10.1080/01944363.2020.1777891.

338   Jia1 JS, Lu X, Yuan Y, Xu G, Jia J, Christakis NA. 2020. Population flow drives spatio-temporal
339   distribution of COVID-19 in China. *Nature* 582: 389.

340   Knorr-Held L. 2000. Bayesian modelling of inseparable space-time variation in disease risk.
341   *Statistics in Medicine* 19: 2555-2568.

342   Kumar J, Sahoo S, Bharti BK, Walker S. 2020. Spatial distribution and impact assessment of
343   COVID-19 on human health using geospatial technologies in India. *International Journal*
344   *of Multidisciplinary Research and Development* 7(5): 57-64

345   Kuunika. 2020. Mathematical Modeling of COVID-19 in Malawi. Quantifying the Potential
346   Burden of Novel Coronavirus. Available at: https://thanzi.org/wp-
347   content/uploads/Kuunika-Modeling_COMThinkTank_21May2020-1-1-1.pdf.

348   Likassa HT. 2020. The Impacts of Covariates on Spatial Distribution of Corona Virus 2019
349   (COVID-19): What Do the Data Show through ANCOVA and MANCOVA? *EJMO*
350   4(2):141–148

351   Martellucci CA, Sah R, Rabaan AA, Dhama K, Casalone C, Arteaga-Livias K, Sawano T, Ozaki
352   A, Bhandari D, Higuchi A, Kotera Y, Fathah Z, Roy N, Rahman MAU, Tanimoto T,
353   Rodriguez-Morales AJ. 2020. Changes in the Spatial Distribution of COVID-19
354   Incidence in Italy using GIS-based maps. *Ann Clin Microbiol Antimicrob* 19:30.

355   Mecenas P, Bastos RTdRM, Vallinoto ACR, Normando D. 2020. Effects of temperature and
356   humidity on the spread of COVID-19: A systematic review. *PLoS ONE* 15(9): e0238339.

357   Pak A, Adegboye OA, Adekunle A, Rahman KM, McBryde ES, Eisen DP. 2020. Economic
358   consequences of the COVID-19 outbreak: the need for epidemic preparedness. *Frontiers*
359   *in Public Health* 8, 1

360   Penerliev M, Petkov V. 2020. Geodemographic Aspects of Covid-19. *Espaço e Economia* (18).

361   Available at: URL: http://journals.openedition.org/espacoeconomia/13444atial

362   Pequeno P, Mendel B, Rosa C, Bosholn M, Souza JL, Baccaro F, Barbosa R, Magnusson W.

363   2020. Air transportation, population density and temperature predict the spread of

364   COVID-19 in Brazil. *PeerJ* 8:e9322. DOI 10.7717/peerj.9322.

365   Petersen E, Koopmans M, Go U, Hamer DH, Petrosillo N, Castelli F, Storgaard M, Khalili SA,

366   Simonsen L. 2020. Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics.

367   *Lancet Infect Dis* 20: e238–44

368   Roos M, Held L. 2011. Sensitivity analysis in Bayesian generalized linear mixed models for

369   binary data. *Bayesian Analysis* 6 (2): 259-278.

370   Sarfo AK, Karuppannan S. 2020. Application of Geospatial Technologies in the COVID-19

371   Fight of Ghana. *Transactions of the Indian National Academy of Engineering* 5:193–204

372   Sarmadi M, Moghanddam VK, Dickerson AS, Martelletti L. 2020. Association of COVID-19

373   distribution with air quality,sociodemographic factors, and comorbidities: an ecological

374   study of US states. Air Quality, Atmosphere & Health. Available at:

375   https://doi.org/10.1007/s11869-020-00949-w

376   Snow J. 1855. On the Mode of Communication of Cholera. Available at:

377   http://www.ph.ucla.edu/epi/snow/snowbook.html

378   Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. 2002. Bayesian measures of model

379   complexity and fit (with discussion). *Journal of the Royal Statistical Society - Series B*

380   64(4): 583–639.

381   Tantrakarnapa K, Bhopdhornangkul B, Nakhaapakorn K. 2020. Influencing factors of COVID-

382   19 spreading: a case study of Thailand. *J Public Health* (Berl.): From Theory to Practice

383   WHO. Coronavirus disease (COVID-19): Situation Report – 203, 2020. Available at:

384   https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200810-covid-

385   19-sitrep-203.pdf?

386   Xie X, Naminse EY, Liu S, Yi Q. 2020. The spatial and temporal pattern of COVID-19 and its

387   effect on human development in China. *Global J. Environ. Sci. Manage.* 6(SI):107-118.

388   Yang W, Deng M, Li C, Huang J. 2020. Spatio-Temporal Patterns of the 2019-nCoV Epidemic

389   at the County Level in Hubei Province, China. *Int. J. Environ. Res. Public Health*

390   17:2563

391   Ye L, Hu L. 2020. Spatiotemporal distribution and trend of COVID-19 in the Yangtze River

392   Delta region of the People's Republic of China. *Geospatial Health* 15:889.

393   **Table 1: DIC of all fitted models**

394

395   **Table 2: Spatial temporal model of COVID-19 monthly confirmed cases (model B11)**

396

397   **Figure 1:** Cumulative cases of confirmed and those who had died of COVID-19 in Malawi

398   from April 2 to October 28, 2020. (CC) Confirmed cases. (DC) deceased cases.

399

400   **Figure 2:** Sensitivity analysis of the hyper-parameters. (A) Posterior distribution of

401   unstructured area variance and (B) Posterior distribution of area × time variance. (Dashed)

402   log-Gamma (1, 0.0005), (Dotted) log-Gamma (1, 0.01) and (Solid) log-Gamma (0.5, 0.001).

403

404   **Figure 3:** Evolution of overall predicted risk by district over time. (A) April, (B) May, (C)

405   June, (D) July, (E) August, (F) September and (G) October. Green (low risk), yellow (medium

406   risk) and red (high risk).

407   **Figure 4:** Evolution of relative risk by district over time. (DZ) Dedza, (DA) Dowa, (KU)

408   Kasungu, (LL) Lilongwe, (MC) Mchinji, (KK) Nkhota-Kota, (NU) Ntcheu, (NS) Ntchisi,

409   (SA) Salima, (CP) Chitipa, (KA) Karonga, (MZ), Mzimba/Mzuzu, (CK) Chikwawa, (CZ)

410   Chiradzulu, (MHG) Machinga, (MH) Mangochi, (MJ) Mulanje, (MN) Mwanza, (NE) Nsanje,

411   (TO) Thyolo, (ZA) Zomba, (PE) Phalombe and (BK) Balaka.

412   **Figure 5:** Evolution of spatial risk of COVID-19 over time (Interaction effect of location and

413   time on exponential scale). (A) April, (B) May, (C) June, (D) July, (E) August, (F) September

414   and (G) October. Green (low risk), yellow (medium risk) and red (high risk).

415

416

DIC of the fitted models

1   **Table 1: DIC of all fitted models**

Model code	Description	DIC
A	Parametric linear time trend	13675.03
B 11	Non parametric time trend, type 1 interaction, RW2.	1052.041
B 12	Non parametric time trend, type 1 interaction, RW1.	1052.071
B 21	Non parametric time trend, type 2 interaction, RW2.	1094.861
B 22	Non parametric time trend, type 2 interaction, RW1.	1094.922
B 31	Non parametric time trend, type 3 interaction, RW2.	1107.471
B 32	Non parametric time trend, type 3 interaction, RW1.	1081.685
B 41	Non parametric time trend, type 4 interaction, RW2.	1098.379
B 42	Non parametric time trend, type 4 interaction, RW1.	1073.298

Table 2: Spatial temporal model of COVID-19 monthly confirmed cases (model B 11)

1   **Table 2: Spatial temporal model of COVID-19 confirmed cases (model B11)**

Variable	Coefficient (95 % confidence interval)
Fixed effects	
Population density	-0.003 (-0.007, 0.001)
Proportion of households with no toilet	0.089 (-0.071, 0.250)
Proportion of households with running water	-0.020 (-0.073, 0.033)
Proportion of households with no radio	-0.095 (-0.216, 0.025)
Proportion of people with age 65+ years	1.272 (0.171, 2.370)
Poverty incidence	-0.100 (-0.136, -0.065)
Variance parameters	
Area (i.i.d)	1.240 (0.621, 2.170)
Area (spatial)	1870.72 (128.861, 6839.73)
Time (i.i.d)	108.64 (0.331, 689.52)
Time (RW2)	1.45 (0.020, 8.29)
Area x Time (i.i.d)	3.30 (2.252, 4.68)

Figure 1

Cumulative cases of confirmed and those who had died of COVID-19 in Malawi from April 2 to October 28, 2020.

(CC) Confirmed cases. (DC) deceased cases.
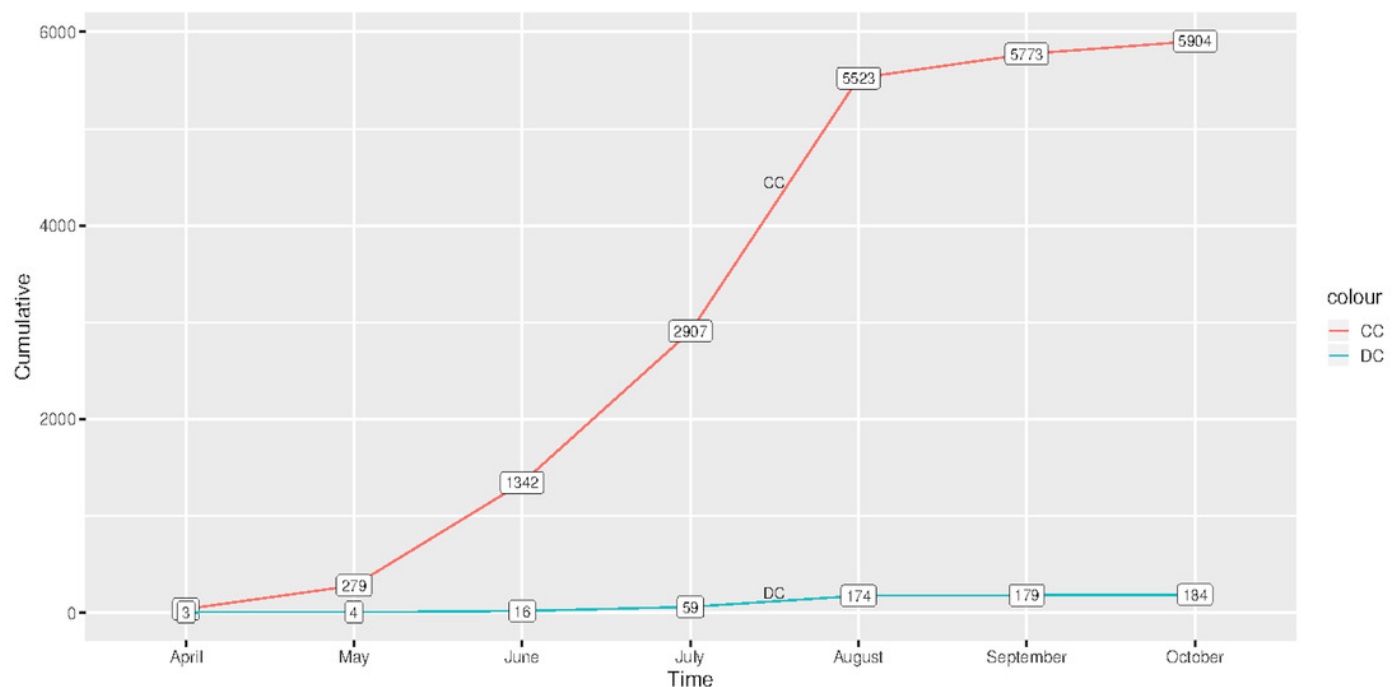
Figure 2

Sensitivity analysis of the hyper-parameters.

(A) Posterior distribution of unstructured area variance and (B) Posterior distribution of area × time variance. (Dashed) log-Gamma (1, 0.0005), (Dotted) log-Gamma (1, 0.01) and (Solid) log-Gamma (0.5, 0.001).
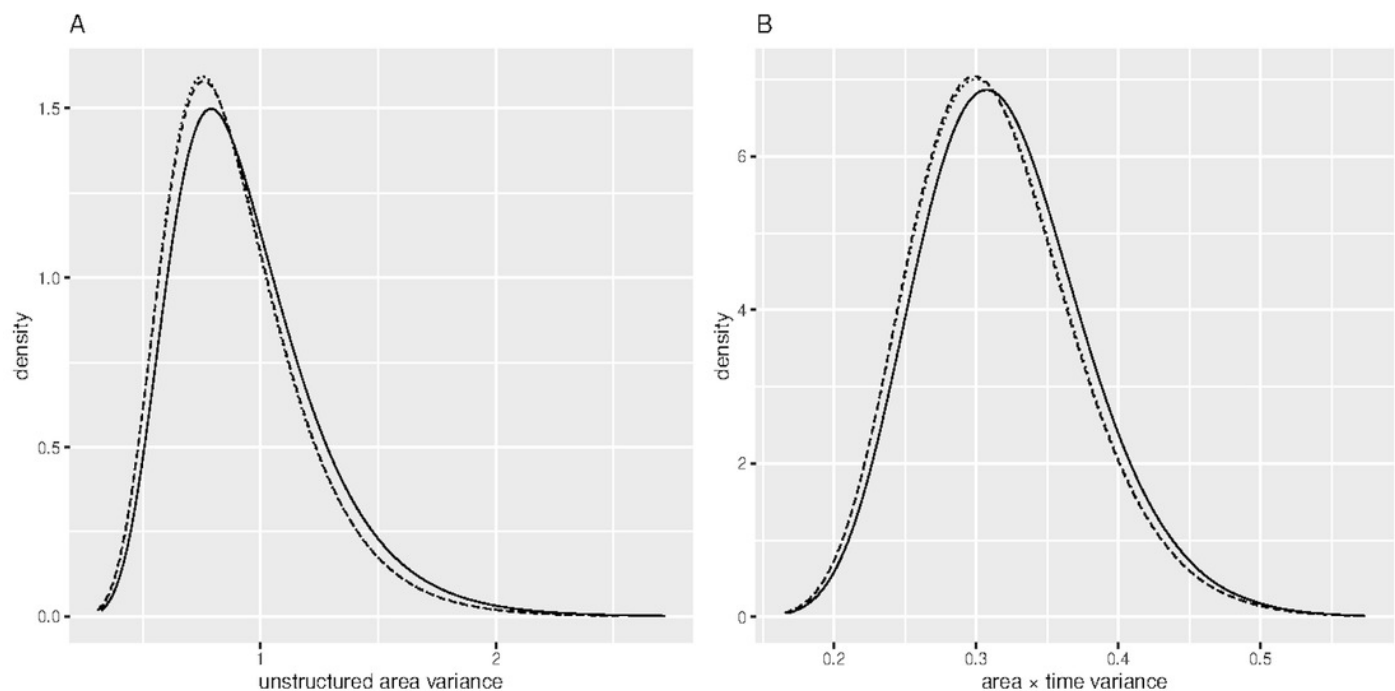
Figure 3

Evolution of overall predicted risk by district over time

(A) April, (B) May, (C) June, (D) July, (E) August, (F) September and (G) October. Green (low risk), yellow (medium risk) and red (high risk).
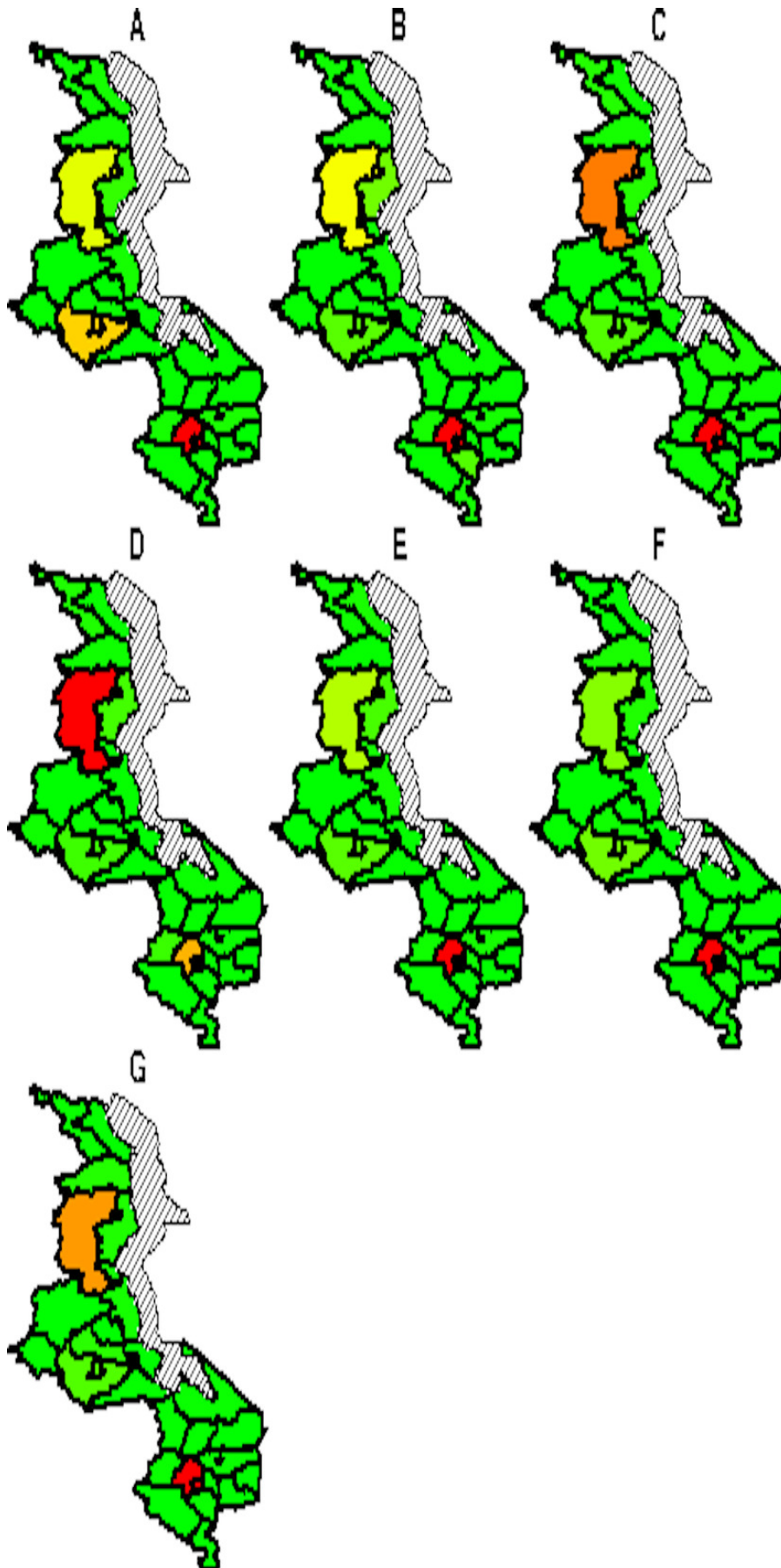
Figure 4

Evolution of relative risk by district over time.

(DZ) Dedza, (DA) Dowa, (KU) Kasungu, (LL) Lilongwe, (MC) Mchinji, (KK) Nkhota-Kota, (NU) Ntcheu, (NS) Ntchisi, (SA) Salima, (CP) Chitipa, (KA) Karonga, (MZ), Mzimba/Mzuzu, (CK) Chikwawa, (CZ) Chiradzulu, (MHG) Machinga, (MH) Mangochi, (MJ) Mulanje, (MN) Mwanza, (NE) Nsanje, (TO) Thyolo, (ZA) Zomba, (PE) Phalombe and (BK) Balaka.
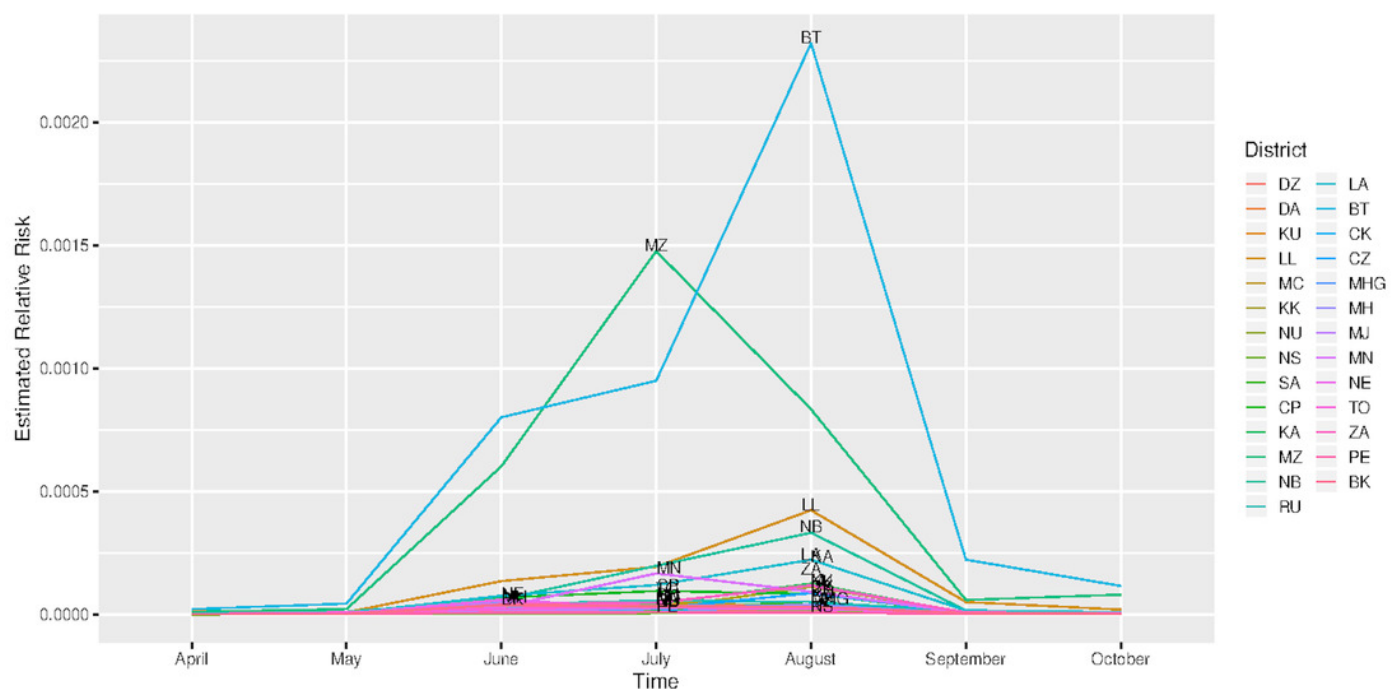
Figure 5

Evolution of spatial risk of COVID-19 over time (Interaction effect of location and time on exponential scale).

(A) April, (B) May, (C) June, (D) July, (E) August, (F) September and (G) October. Green (low risk), yellow (medium risk) and red (high risk).