Evaluation of the impacts of air quality on the risk of COVID-19 in England

Yingfei Zha

10/12/2023

COVID-19 in England

Introduction

In this study, the data including COVID-19 cases up to 15 October 2020 from 149 upper tier local authorities (i.e. counties) in England are analyzed to investigate whether the exposure to ambient air pollution makes individuals more susceptible to COVID-19. Additionally, socio-economic factors such as unemployment rate and ethnic minorities proportion in each county are also taken into account as important confounders. We want to explore whether a high unemployment rate and a high ethnic minorities proportion result in a higher risk of COVID-19 incidence. For the above hypotheses, we use a spatial ecological design to estimate the risk of COVID-19 incidence under the impact of above covariates.

Method

The model fitted using Bayesian inference is shown below. The number of COVID-19 cases Y_i is said to follow a Poisson distribution, where E_i is an offset that accounts for the expected number of COVID-19 cases based on population and known incidence rates. λ_i is the risk of COVID-19 incidence relative to what we would expect in a county. The covariates X_i are vectors of proportion of ethnic minorities, concentrations of fine particulate matter (PM 2.5) and unemployment rate in each county. The prior of β is set to the default prior, which is a Normal distribution prior with mean equals to 0 and standard deviation equals to 1000. U_i is a spatial random effect which handles variations that the current covariates cannot explain. This random effect is fitted using a reparametrized Besag, York and Mollie (BYM) model, where σ is the standard deviation of spatial effect and τ is the independent standard deviation. θ_1 and θ_2 are integrated parameters of the BYM model. θ_1 is the overall standard deviation of U_i and θ_2 is the spatial proportion. If θ_1 and θ_2 are zero then there will be no spatial dependency and the spatial random effect will be constant. The penalized complexity prior of θ_1 is an Exponential distribution prior with a median of 0.5. It implies that 1 standard deviation increasing results in 1.65 times the risk. The prior distribution of θ_2 has a prior median of 0.5, which implies that there is a certain amount of spatial dependency between counties. This prior is chosen because 0.5 is a good middle ground from 0 to 1.

$$
\begin{aligned}
Y_i &\sim \text{Poisson}\left(E_i \lambda_i\right) \\
\log\left(\lambda_i\right) &= \mu + \beta_1 X_{Ethnicity} + \beta_2 X_{PM2.5} + \beta_3 X_{Unemployment} + U_i \\
U_i &\sim \text{BYM}\left(\sigma^2, \tau^2\right) \\
\theta_1 &= \sqrt{\sigma^2 + \tau^2} \\
\theta_2 &= \sigma / \sqrt{\sigma^2 + \tau^2}
\end{aligned}
$$

Results

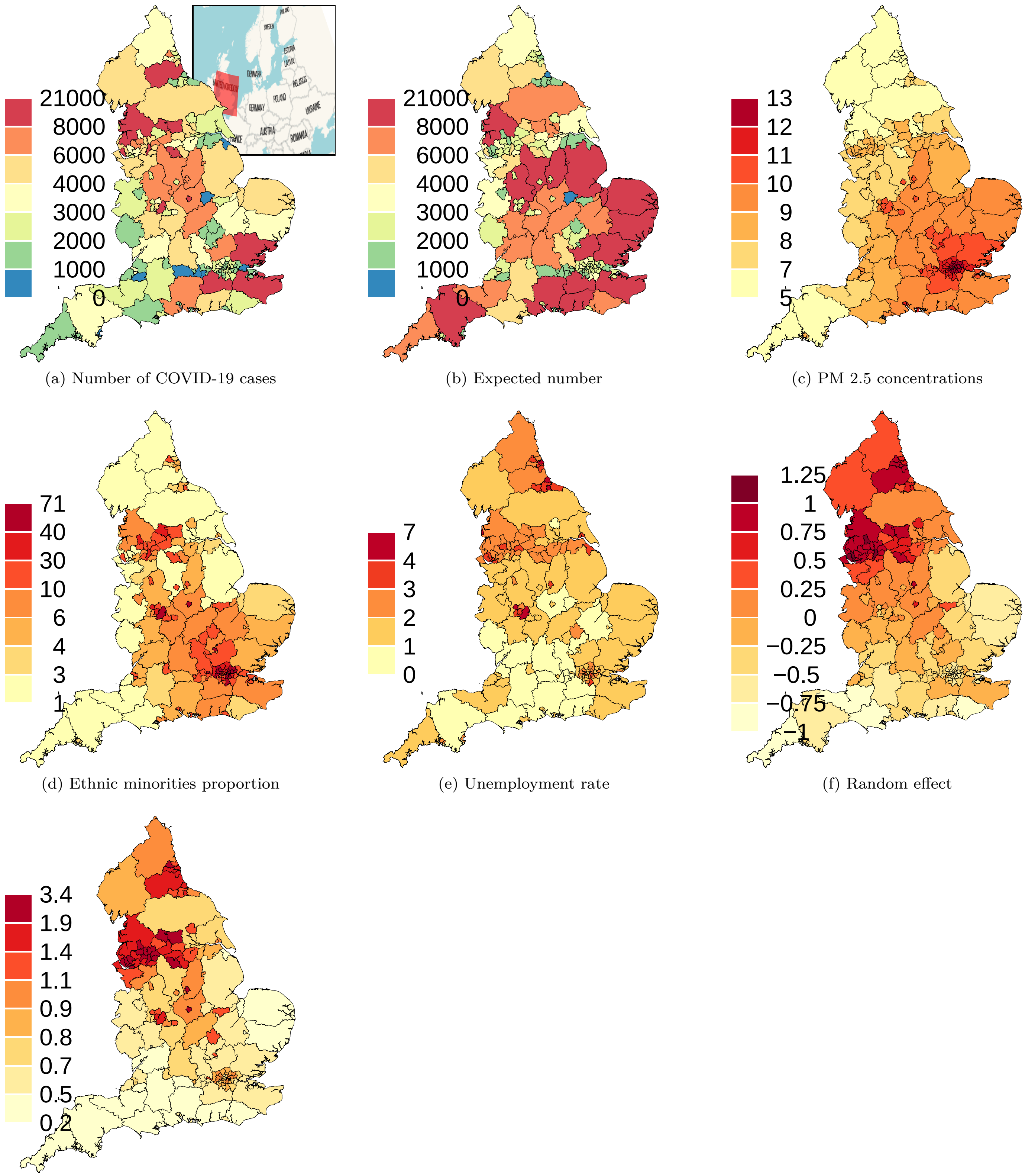
The statistical results from our model are shown in **Table 2**. The proportion of ethnic minorities and unemployment rate both show significant influence on the risk of COVID-19 incidence, however, the concentrations of PM2.5 do not show significant influence because the credible intervals of PM2.5 concentrations include 1. Additionally, in plot (c) of **Figure 3**, air pollution is more concerning to the southeast of England. However, comparing to plot (g), there are not very high risk of COVID-19 incidence so exposure to ambient air pollution should not be influential to the number of COVID-19 cases. Moreover, the unemployment rate seems to have 10 times more influence on the risk of COVID-19 incidence than the ethnic minorities proportion. For each unit increase in the unemployment rate, the risk of COVID-19 incidence increases by 12%. For each unit increase in the ethnic minorities proportion, the risk of COVID-19 incidence increases by 1.2%. The results also suggest that the residual standard deviation is 0.293 indicating that there are some effects that do not explained by the covariates. The spatial proportion is 0.907 which is not 0, so counties with higher risk of COVID-19 incidence tend to be surrounded by counties with higher risk. In plot (f) of random effect in **Figure 3**, we can also find out that the model overestimates the risk of COVID-19 incidence in the north but underestimates the risk in the south of England. To be more specifically, the risk of COVID-19 incidence is very high in the north as shown in plot (g), but none of the covariates explain this phenomenon. The ethnic minorities proportion in the north is pretty low shown in plot (d). Also for the counties with high risk of COVID-19 incidence (1.4-3.4 times more) in the north, not all of them have the highest unemployment rate (3-7) as shown in plot (e). To the southwest of England, the random effect is negative meaning that there are some factors causing the risk is lower than what we expect. Although given the fact that the unemployment rate and ethic minorities proportion are low there, the population is very high. Thus there must be some factors that are not in our model causing this very low risk of COVID-19 incidence in these counties.

Summary

This study analyzes COVID-19 cases up to 15 October 2020 from 149 counties of England to look for indicators that correlate with inceases in the risk of COVID-19 incidence. The statistical results indicate that the air quality does not have influence on the risk of COVID-19 incidence and the counties with high air pollution do not show high risk. Additionally, increasing in unemployment rate and ethnic minorities proportion can increase the risk of COVID-19 incidence. Unemployment is causing more impact on the risk of COVID-19 incidence than having more ethnic minorities in a county. However, we also find out that the high risk of COVID-19 incidence in some northern counties and the low risk in the south are not explained by the above covariates. Therefore our model should be expanded to include more potential factors.

Table 1: Estimated coefficients, SD (log scale) and spatial proportion (log scale)

	0.5quant	0.025quant	0.975quant
Intercept	0.365	0.218	0.610
Ethnic monority	1.012	1.008	1.016
PM2.5	1.057	0.996	1.123
Unemployment	1.120	1.059	1.184
sd	0.293	0.259	0.336
propSpatial	0.907	0.768	0.975



(a) Number of COVID-19 cases

(b) Expected number

(c) PM 2.5 concentrations

(d) Ethnic minorities proportion

(e) Unemployment rate

(f) Random effect

(g) Risk of COVID-19 incidence

Figure 1: Plots of England COVID-19 data

Appendix

```
knitr::opts_chunk$set(echo = FALSE, warning = FALSE, message = FALSE)

############COVID-19################
(load("~/Downloads/England_shp.RData"))
UK_shp$logExpected = log(UK_shp$E)
# remove an island
UK2 = UK_shp[grep("Wight", UK_shp$Name, invert = TRUE),]
(load("~/Downloads/englandRes.RData"))
rn=c("Intercept","Ethnic monority","PM2.5", "Unemployment","sd","propSpatial")
sd = englandRes$parameters$summary[, paste0(c(0.5, 0.025, 0.975), "quant")][5:6,]
k = exp(englandRes$parameters$summary[, paste0(c(0.5, 0.025, 0.975), "quant")][1:4,])
t= data.frame(rbind(k,sd), row.names = rn)
knitr::kable(t, digits = 3, col.names = c("0.5quant", "0.025quant", "0.975quant"), caption = "Estimated
casesCol = mapmisc::colourScale(UK2$cases, dec = -3, breaks = 12,
col = "Spectral", style = "quantile", rev = TRUE)
Ecol = mapmisc::colourScale(UK2$E, breaks = casesCol$breaks,
col = casesCol$col, style = "fixed")
pmCol = mapmisc::colourScale(UK2$modelledpm25, breaks = 9,
dec = 0, style = "quantile")
ethCol = mapmisc::colourScale(UK2$Ethnicity, breaks = 9,
digits = 1, style = "quantile")
uCol = mapmisc::colourScale(UK2$Unemployment, breaks = 12,
dec = 0, style = "quantile")
rCol = mapmisc::colourScale(englandRes$data$random.mean,
breaks = 12, dec = -log10(0.25), style = "quantile")
fCol = mapmisc::colourScale(englandRes$data$fitted.exp,
breaks = 9, dec = 1, style = "quantile")
insetEngland1 = mapmisc::openmap(UK2, zoom = 3, fact = 4,
path = "waze", crs = CRS("+init=epsg:3035"))
library("raster")
insetEngland = raster::crop(insetEngland1, extend(extent(insetEngland1),
-c(25, 7, 4, 9.5) * 100 * 1000))
library("sp")
mapmisc::map.new(UK2)
mapmisc::insetMap(UK_shp, "topright", insetEngland, width = 0.4)
plot(UK2, col = casesCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", casesCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = Ecol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", casesCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = pmCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", pmCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = ethCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", ethCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = uCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", uCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = rCol$plot, add = TRUE, lwd = 0.2)
```

```r
mapmisc::legendBreaks("left", rCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = fCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", fCol, bty = "n")
```