OPEN

Conditional Poisson Regression with Random Effects for the Analysis of Multi-site Time Series Studies

Jose Barrera-Gómez,[a,b,c] *Xavier Puig,*[d] *Josep Ginebra,*[d] *and Xavier Basagaña*[a,b,c]

Abstract: The analysis of time series studies linking daily counts of a health indicator with environmental variables (e.g., mortality or hospital admissions with air pollution concentrations or temperature; or motor vehicle crashes with temperature) is usually conducted with Poisson regression models controlling for long-term and seasonal trends using temporal strata. When the study includes multiple zones, analysts usually apply a two-stage approach: first, each zone is analyzed separately, and the resulting zone-specific estimates are then combined using meta-analysis. This approach allows zone-specific control for trends. A one-stage approach uses spatio-temporal strata and could be seen as a particular case of the case–time series framework recently proposed. However, the number of strata can escalate very rapidly in a long time series with many zones. A computationally efficient alternative is to fit a conditional Poisson regression model, avoiding the estimation of the nuisance strata. To allow for zone-specific effects, we propose a conditional Poisson regression model with a random slope, although available frequentist software does not implement this model. Here, we implement our approach in the Bayesian paradigm, which also facilitates the inclusion of spatial patterns in the effect of interest. We also provide a possible extension to deal with overdispersed data. We first introduce the equations of the framework and then illustrate their application to data from a previously published study on the effects of temperature on the risk of motor vehicle crashes. We provide R code and a semi-synthetic dataset to reproduce all analyses presented.

Keywords: Epidemiologic methods; Multi-site; One-stage; Spatial structure; Time series

(Epidemiology 2023;34: 873–878)

Submitted September 20, 2022; accepted August 2, 2023

From the [a]ISGlobal, Barcelona, Spain; [b]Universitat Pompeu Fabra (UPF), Barcelona, Spain; [c]CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain; and [d]Department of Statistics and Operations Research, Universitat Politècnica de Catalunya, Barcelona, SpainThe authors report no conflicts of interest.

The data used is not owned by the authors and cannot be shared. A semi-synthetic dataset and the computing code to replicate the results are provided as Supplemental Digital Content.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Xavier Basagaña, ISGlobal, Dr. Aiguader 88, 08003 Barcelona, Spain. E-mail: xavier.basagana@isglobal.org.

ISSN: 1044-3983/23/346-873878

DOI: 10.1097/EDE.0000000000001664

The analysis of time series studies linking a health indicator's daily counts with environmental variables (e.g., mortality or hospital admissions with air pollution concentrations or temperature; or motor vehicle crashes with temperature) is usually conducted with Poisson regression models that control for long-term and seasonal trends. Trends can be controlled by using splines of time or temporal strata, for example, strata defined by the combination of year, month, and day of the week. The latter approach is equivalent to a time-stratified case–crossover analysis.[1] When the study includes multiple zones (e.g., cities), analysts usually apply a two-stage approach, in which each zone is analyzed separately (first stage) and the zone-specific results of the first stage are then combined using univariate or multivariate meta-analysis (second stage). This approach allows a zone-specific control for trends.

A one-stage approach to analyze this kind of data is to pool all data together and define strata as the combinations of zone, year, month, and day of the week. This approach has been called space–time-stratified case–crossover.[2] Such a model could be seen as a particular case of the case–time series framework recently proposed.[3,4] In that framework, the temporal trends are captured using temporal strata that may or may not vary by zone. Depending on how the temporal strata are defined, the number of strata can escalate very rapidly in a long time series with many zones. A computationally efficient alternative is to fit a conditional Poisson regression model, which uses the likelihood conditional to the sum of events within each stratum. This method avoids the need to estimate the strata parameters, leading to a multinomial model that provides the same estimates for the remaining parameters.[1]

The case–time series approach allows estimating zone-specific effects. For example, if one is interested in the effect of air pollution, one can include interaction terms between zone or zone-specific variables and the air pollution terms.[4,5] Another approach would be assuming zone-specific random effects for the air pollution terms (i.e., random slope). However, the available (frequentist) software for conditional Poisson regression does not allow including random effects.

Here, we show how to fit a conditional Poisson regression model with a random slope using a Bayesian approach. This is an approach similar to the conditional linear mixed models derived for the case of linear regression.[6] The Bayesian paradigm also facilitates the inclusion of spatial structure in the random effects. We first introduce the equations of the modeling framework and then apply them to an illustrative example using a previously published study on the effects of temperature on the risk of motor vehicle crashes.[7] We provide the R code used and a semi-synthetic data set to reproduce similar results to those presented in the example.

MODELING FRAMEWORK

Consider three-level nested data in which level 1 identifies I geographical zones, that is, cities or climatic regions; level 2 identifies J_i time strata, that is, combinations of year, month, and day of week, and level 3 identifies K_{ij} daily observations. Let Y_{ijk} be the outcome (e.g., motor vehicle crash count) observed in the k-th day, $k = 1, 2, \ldots, K_{ij}$, within the j-th time stratum, $j = 1, 2, \ldots, J_i$, within the i-th zone, $i = 1, 2, \ldots, I$. Let $X_{lijk}, l = 1, 2, \ldots, p$, be a set of variables potentially associated with Y_{ijk}, where $l = 1$ corresponds to the exposure of interest (e.g., temperature) and $l = 2, 3, \ldots, p$ correspond to additional time-varying confounder variables in the relationship of interest (e.g., daily mean levels of humidity or rainfall in an analysis of temperature). We consider the Generalized Linear Mixed Model (GLMM) framework under the following assumptions: (1) The outcome count follows a Poisson distribution; (2) As usual in practice, the relationship between the mean of the outcome and a linear combination of the explanatory variables is exponential (i.e., log–linear), which provides estimations of the association of interest in terms of risk ratio (RR); (3) Baseline risk varies by zone and time strata; and (4) There is a potential geographical-specific association between the outcome and the exposure of interest.

POISSON REGRESSION MIXED MODEL

The modeling framework just introduced above is compatible with the following Poisson regression mixed model:

$$\begin{cases} Y_{ijk} \sim \text{Pois}(\lambda_{ijk}) \\ log(\lambda_{ijk}) = \kappa_{ij} + \beta_{1i} X_{1ijk} + \sum_{l=2}^{p} \beta_l X_{lijk} \\ \beta_{1i} = \beta_1 + u_{1i} \\ u_{1i} \sim \mathcal{N}(0, \sigma_1) \end{cases} . \quad (1)$$

Model (1) is a mixed-effects (or simply mixed) model, with fixed effects in the intercept at level 2 (e.g., zone-time level) and random effects in the slope at level 1 (e.g., zone level). Specifically, the interpretation of terms in model (1) is as follows:

(1) κ_{ij} is related to the baseline risk (e.g., mean number of crashes) in time stratum j in zone i. That is, the set of κ_{ij} parameters capture variations in average mortality by zone and zone-specific time trends and seasonality in mortality. κ_{ij} parameters are treated as fixed effects. Thus, the number of parameters $\kappa_{ij}, \sum_{i=1}^{I} J_i$, is usually large (e.g., in the study on motor vehicle crashes mentioned above, $\sum_{i=1}^{I} J_i = 7,056$). Further, κ_{ij} parameters also capture any characteristic that could be *ij*-stratum specific, that is, within *ij*-stratum invariant but between *ij*-strata varying. As a particular case, κ_{ij} parameters also capture any characteristic which is *i*-stratum specific (e.g., percentage of homes with air conditioning or percentage of individuals with high or low education).

(2) β_{1i} measures the association, in zone i, between the exposure of interest, X_1(e.g., daily ambient temperature), and the outcome,Y. β_{1i} comprises two contributions: (1) β_1, which is the parameter of main interest, is related to the overall (i.e., population-level, common to all zones) adjusted effect of the exposure on the outcome; and (2) u_{1i}, a random disturbance of that association at zone level (i.e., level 1), capturing the distinctive feature of each zone in the association of interest. u_{1i} is modeled as a random effect in the slope of interest (β_1), following a normal distribution with mean zero and standard deviation σ_1, which models the variability of the association of interest across the different zones.

(3) β_l ($l = 2, 3, \ldots, p$), is a set of parameters related to additional time-varying adjusting covariates $X_l, l = 2, 3, \ldots, p$ (e.g., daily mean levels of humidity or rainfall), assumed to be the same in all zones.

One could try to estimate model (1) using a frequentist or Bayesian approach. Next, we describe both separately.

In the frequentist approach, in the absence of the random slope (u_{1i}), one can obtain a consistent estimation of β_1 using maximum likelihood.[8,9] Standard statistical software provides such an estimation (e.g., glm function in R). However, if the number of nuisance parameters κ_{ij} is high enough, computational issues can arise (e.g., in the study on crashes, R freezes on a typical computer normally when trying to fit (1) with no random effect in the slope). In our experience, when trying to fit model (1), including both fixed effects at level 2 (κ_{ij}) and random slope at level 1 (u_{1i}), with usual software for mixed models (e.g., glmer function in the R package lme4) the software freezes or the algorithm does not converge, even in small data sets.

In the Bayesian approach, using flat priors, as usual in practice, for parameters κ_{ij}, model (1) may provide estimations for the parameters of interest that are not consistent as a consequence of the incidental parameters problem, even in the case of not having a random slope.[9,10]

An efficient alternative to overcome those problems is to work with the conditional likelihood, which avoids the need to estimate the nuisance parameters κ_{ij}, as described next.

CONDITIONAL POISSON REGRESSION MIXED MODEL

To avoid the problems derived from the presence of a high number of nuisance parameters κ_{ij} in the model (1), described above, we can make inference based on the conditional likelihood, obtained by conditioning on a sufficient statistic of the nuisance parameters, as is the total outcome count in each zone-time stratum ij(less formally, making an inference based on within-strata variability of the data). As a result, the conditional likelihood does not depend on the nuisance parameters. Specifically, based on conditional likelihood, model (1) becomes (eAppendix, Section 1; http://links.lww.com/EDE/C64):

$$\begin{cases} \mathbf{Y}_{ij}|S_{ij}, X \sim \text{Multinom}(S_{ij}, \pi_{ij1}, \pi_{ij2}, \ldots, \pi_{ijK_{ij}}) \\ S_{ij} := \sum_{k=1}^{K_{ij}} Y_{ijk} \\ \pi_{ijk} = g_{ijk}/g_{ij\cdot} \\ g_{ijk} := exp(u_{1i}X_{1ijk} + \sum_{l=1}^{p} \beta_l X_{lijk}) \\ g_{ij\cdot} := \sum_{k=1}^{K_{ij}} g_{ijk} \\ u_{1i} \sim \mathcal{N}(0, \sigma_1) \end{cases} \quad (2)$$

In model (2), the vector of the K_{ij} observations (i.e., days) for the outcome count in the zone-time stratum ij, conditional to the total number of outcome counts in that stratum,S_{ij}, follows a multinomial distribution with K_{ij} probability parameters. These probability parameters depend only (via g_{ijk}) on the β coefficients and the random component in the slope u_{1i}. Hence, in model (2) only the β coefficients and σ_1 need to be estimated to obtain zone-specific associations between the exposure of interest and the outcome, adjusted for the potential confounders included in the model.

Model (2) provides consistent inference for the parameters of interest. Briefly, it can be shown (eAppendix, Section 2; http://links.lww.com/EDE/C64) that likelihood of model (1),$\mathcal{L}$, factors as

$$\mathcal{L}(\beta, \sigma_1, \kappa|\mathbf{y}) = \mathcal{L}_1(\beta, \sigma_1, \kappa|S)\mathcal{L}_2(S, \beta, \sigma_1|\mathbf{y}). \quad (3)$$

This result implies that $\mathcal{L}_2$ provides consistent inference for parameters β and σ_1.[9,10] Precisely, $\mathcal{L}_2$ is the likelihood associated with model (2) (eAppendix, Section 2; http://links.lww.com/EDE/C64). Using $\mathcal{L}_2$ to make inference is sometimes referred to as partial likelihood estimation.[9,11,12] In case of not having the random slope, model (2) can be estimated, in the frequentist approach, with standard software (e.g., gnm package in R).[1] To the extent of our knowledge, frequentist estimation of model (2) is not implemented in standard statistical software in the presence of a random slope.

We propose fitting model (2) using a Bayesian approach. Such an approach also facilitates the inclusion of spatial structure in the random slope in the model. Specifically, eAppendix Section 3; http://links.lww.com/EDE/C64, illustrates how to fit both cases of independent random effects or spatially patterned random effects using the Gaussian conditional autoregressive distribution, which allows spatial correlation between neighboring zones.[13] This Bayesian approach has two advantages. First, as indicated above, it provides consistent estimations for the parameters of interest. Second, setting priors for κ_{ij} is not necessary because those nuisance parameters are not involved in the model.

Fitting model (2) requires some data manipulation from the original time series format. Specifically, one needs to aggregate the data at level 2, that is, for each zone-time stratum, to provide the counts of each multinomial distribution, whose length may vary between strata. We provide an R code to automatically convert the data set to the required format (eAppendix, Section 3.3.1; http://links.lww.com/EDE/C64). The Bayesian approach requires setting prior distributions for the parameters in the model, that is, the effect of interest,β_1, the coefficient associated with the confounders,$\beta_2, \ldots, \beta_p$, and the standard deviation of the normal distribution of the random effect in the slope of interest,σ_1. For all these parameters, we set noninformative priors, which we specifically describe in the illustrative example below. All credible intervals derived from Bayesian analyses presented in this work are equal-tailed.

OVERDISPERSION

Overdispersion is not a rare situation when modeling data sets in the kind of study analyzed here. Dealing with overdispersion in conditional model (2) can be challenging. In the case of overdispersion, following other authors,[14] model (1) could be extended by adding an error term, ε_{ijk}, in the linear predictor at the observation level (e.g., day). Then, after conditioning, the expression for g_{ijk} in model (2) becomes $g_{ijk} = exp(u_{1i}X_{1ijk} + \sum_{l=1}^{p} \beta_l X_{lijk} + \varepsilon_{ijk})$, where $\varepsilon_{ijk} \sim \mathcal{N}(0, \sigma_{e_i})$, which allows the magnitude of overdispersion being different in each zone, even allowing for zones with no overdispersion. To decide if the error term ε_{ijk} should be included in the model, both models, with and without the error term, can be fitted. Then, model selection can be decided according to some criteria such as the Deviance Information Criteria.[15] In eAppendix, Section 3.4; http://links.lww.com/EDE/C64, we provide more details as well as R code and an illustrative example.

ILLUSTRATIVE EXAMPLE

We illustrate our methodology using a previously published study on the effects of temperature on the risk of motor vehicle crashes.[7] Data includes all crashes involving driver performance factors that occurred in the warm season (May 15–October 15) during the years 2000–2011 in 14 meteorologic regions of Catalonia. We define strata as unique combinations of region, year, month, and day of the week, leading to 7,056 unique strata. For these data, in models (1) and (2), $p = 4$: X_1, as the exposure of interest, for maximum daily temperature, and X_2, X_3, and X_4, as confounders, for binary indicators of the holiday, the day being the first or the last day in a holiday period, and precipitation, respectively. In the Bayesian models, we set noninformative priors as follows: For each of the β coefficients, a normally distributed prior with mean 0 and precision 10^{-6}(i.e., variance 10^6). For σ_1, we defined $\sigma_1 = 1/\sqrt{\tau_1}$, where τ_1 followed a gamma distribution with mean 1 and variance 10^3. Figure 1 shows the results obtained using the usual two-stage approach (frequentist analysis), and the one-stage Bayesian approach using independent or spatially patterned random effects. The frequentist analysis estimated a relative risk (RR) of 1.051 (95% confidence interval = 0.999, 1.106) for a 5 °C increase in temperature, with heterogeneity between zones $I^2 = 70\%$. The Bayesian models estimated RRs of 1.051 (95% credible interval = 0.976, 1.131) and 1.044 (95% credible interval = 0.995, 1.093) when including independent or spatial random effects, respectively. The three analyses led to similar conclusions, all of them capturing a between-zone variability of the association of interest. Regarding effect size, zone 5 showed the highest association consistently in the three analyses and the rest of the zones having similar estimates in all three approaches, with only some small differences observed in zone 14.

DISCUSSION

We present an approach for one-stage analysis of multi-city case–crossover analyses that allows estimating city-specific effects using random effects and a Bayesian implementation. This approach allows incorporating the spatial structure of the data to borrow strength from neighboring regions in the estimation process. We provide reproducible code and a data set to facilitate the implementation of this methodology in new studies.

Our approach is based on the mixed-effects Poisson regression model framework to model data nested in two levels, with level 1 and level 2 being cities and spatio-temporal strata, respectively. We model baseline risk spatio-temporal variations (at level 2) with fixed effects in the model intercept while allowing estimation of heterogeneous effects (at level 1, between-city) by including a random slope associated with the exposure of interest. This approach has been used in linear models.[6] For the Poisson case, such a model cannot be fitted in the frequentist paradigm with the usual mixed-models software.

The kind of two-level nested data presented here is usually analyzed with a two-stage approach. In the first stage, data from each city (level 1) are modeled independently. In the second stage, zone-specific measures of the association of interest estimated in the first stage are pooled using appropriate fixed effects or random effects meta-analysis.,[7,16,17] This two-stage approach naturally incorporates zone-specific control for trends and estimation of zone-specific effects of the exposure of interest. The latter can be directly obtained from

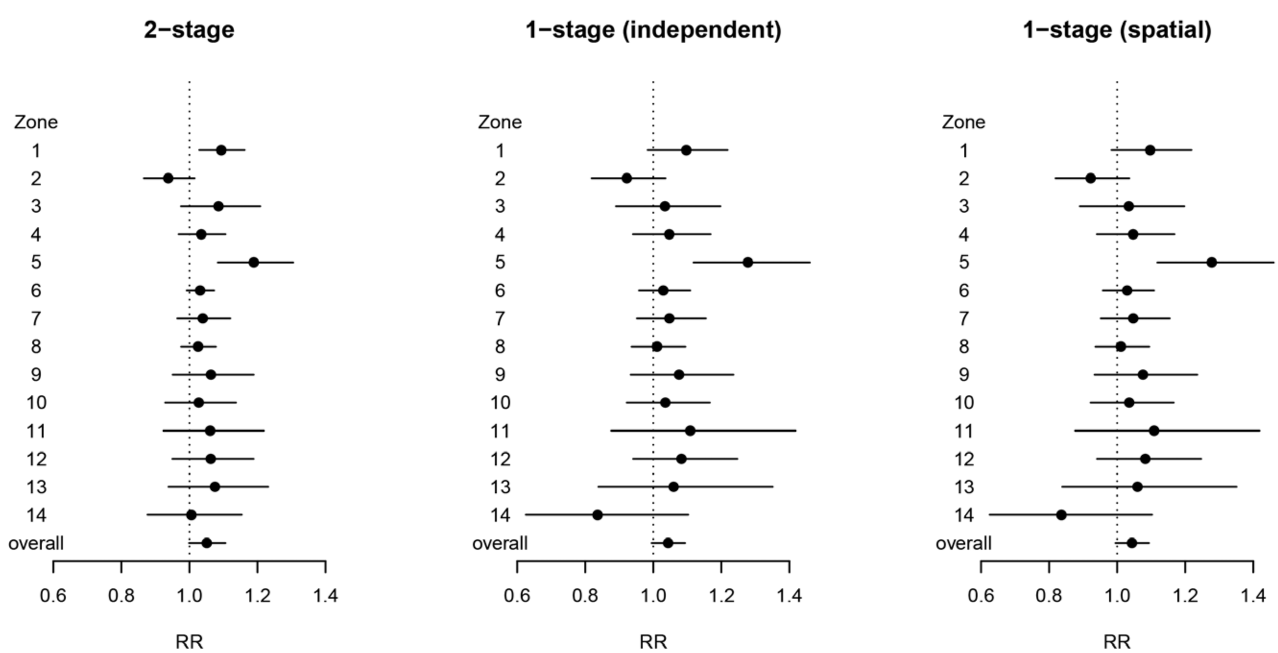

FIGURE 1. Association (relative risk [RR] and 95% confidence or credible interval) between the daily number of motor vehicle crashes involving driver performance and a 5°C increase in maximum temperature, by zone. Models were adjusted for indicator variables for precipitation, holidays, and the day being the first or the last day in a holiday period. The left panel shows the best linear unbiased predictions (BLUPs) of the zone-specific RRs, obtained from a meta-analysis of zone-specific estimates (two-stage frequentist analysis); the middle panel shows the Bayesian one-stage analysis using conditional Poisson regression with independent random effects; and the right panel shows the Bayesian one-stage analysis using conditional Poisson regression with spatially patterned random effects, using a Gaussian conditional autoregressive (CAR) distribution.

the analysis of each city, or instead, one can calculate the best linear unbiased predictions from a meta-analysis. This is the approach that we adopt in our example, as it is consistent with the mixed-model approach we applied. The best linear unbiased predictions combine the zone-specific estimate with the average estimate, thus borrowing strength from the other zones to obtain an estimate closer to the underlying true value.[18]

Other studies have used a similar approach to the one presented here when analyzing individual-level case–crossover data.[19,20] In that case, they used a Poisson regression model similar to (1), with individual data, including strata as fixed effects and allowing for spatial structure, using a Bayesian approach. This approach can theoretically lead to inconsistent estimates due to the incidental parameter problem.[9,10] Using the multinomial model avoids the incidental data problem and potential discrepancies of results when using different prior distributions.

We provide a way to implement a one-stage analysis of the data based on mixed models. However, one could also fit a one-stage analysis by including exposure by zone interaction terms in a space–time-stratified case–crossover analysis, or in an individual-level case–crossover analysis. Using random slopes, as in our analysis, is more parsimonious and introduces shrinkage in the estimates, as when computing the best linear unbiased predictions in a two-stage analysis. The model with interaction terms would produce estimates more similar to the first-stage estimates in a two-stage approach.

Two-stage and one-stage approaches often provide similar results. However, in some cases, both approaches could lead to different results, triggered most likely by discrepant modeling assumptions and specifications or by using different techniques for obtaining estimates and their uncertainty.[21] Hence, in the example presented here, it is expected that two-stage and one-stage approaches produce similar results because both include the same variables in the model. However, confounders in the model could lead to differences in results between one-stage and two-stage approaches. In the two-stage approach, the effect of each confounder can be different in each zone (e.g., city) while, in the one-stage approach, it is assumed to be the same in all zones. This can be relaxed by including in the model confounder by zone interactions or random slopes for the confounders. Future work can examine if there are other situations in which the two approaches lead to different results. However, one advantage of the Bayesian implementation we propose here is easily incorporating prior information (e.g., on the expected direction of estimates) or other spatial correlation patterns, among other features. Regarding the precision of the estimates in our illustrative example, confidence intervals obtained in the two-stage approach are similar to credible intervals obtained in the one-step approach, although in some scenarios the one-stage approach could perform better in terms of coverage.[22] Nevertheless, in general, comparing frequentist and Bayesian results, as in our example, should be carried out carefully since confidence and credible intervals have different meanings.

Using spatial structure is useful for capturing the potential spatial correlation of associations. This correlation often occurs in environmental epidemiologic studies, as associations tend to be more similar in zones close to each other than in others that are farther apart. In addition, spatial structure can lead to more reliable estimates in small zones in which the relatively low amount of information is enriched by borrowing information from surrounding zones. The one-stage approach we implement can easily accommodate the incorporation of spatial structure. We provide an example and code to implement that. Two-stage approaches also allow the incorporation of spatial structures. For instance, using Bayesian techniques, one can implement the second stage assuming a spatial correlation as a function of the distance between zones.[23] Another example allows for hierarchical geographic structures (e.g., cities within countries), although it cannot accommodate neighborhood-based spatial structures.[17]

In our example, the parameter of interest was one-dimensional. Studies on temperature or air pollution often use cross-bases to model lagged effects in short-term association studies, which implies a multidimensional parameterization. Our framework could be extended to allow for distributed lag effects, implementing the underlying algebra described in the seminal paper on distributed lag nonlinear models,[24] although it would require some considerations, for example, on how to incorporate the spatial structure with multidimensional parameters. This can be a topic for future research.

Regarding potential overdispersion, in the data of the study on crashes, the crude ratio variance to the mean of the daily number of crashes was 6.0. Stratifying by zone, it was higher than 1.5 only in zones 6, 8, and 4 (ratios 2.7, 2.0, and 1.6, respectively). After adjusting for the exposure and confounders, those ratios were all below 1.2. Overdispersion was not an important problem in the study on crashes, but, in general, the presence of overdispersion should be considered. In cases where overdispersion is present, our basic methodology would underestimate the variation of the estimates. Extensions of the Bayesian conditional model we present to account for overdispersion are not straightforward.[8] In case of overdispersion, an option consists of adding an error term in the model to allow for extra variation, in a similar way to what other authors have suggested.[14] In the eAppendix Section 3.4; http://links.lww.com/EDE/C64, we provide an illustration on how to implement such an approach.

Data on crashes included only 14 zones, while, ideally, a higher number of zones would be more appropriate to capture spatial patterns. Despite this, we consider a spatially patterned structure for the random effects to allow other researchers to adapt our code in case their data expand to a higher, more suitable, number of zones.

Another possible drawback is related to computation. In some settings, such as including a high number of zones and a higher number of parameters or incorporating the method we suggest to take into account potential overdispersion, our approach could be computationally demanding in comparison to the frequentist two-stage approach. Future research could work on the implementation of more efficient algorithms considering alternative software (e.g., Stan or INLA). Incorporation of autocorrelation could also be a topic of future research.

In conclusion, we propose a one-stage Bayesian estimation process that can be fitted in a context with a large number of spatio-temporal strata while allowing the estimated effect of interest to vary by zone. We provide R code and a semi-synthetic to reproduce the analyses, which are available at https://dataverse.csuc.cat/dataset.xhtml?persistentId=doi:10.34810/data235.

ACKNOWLEDGMENTS

We acknowledge support from the grant CEX2018-000806-S funded by MCIN/AEI/ 10.13039/501100011033, support from Ministry of Research and Universities of the Government of Catalonia (2021 SGR 01563) and support from the Generalitat de Catalunya through the CERCA Program.

REFERENCES

1. Armstrong BG, Gasparrini A, Tobias A. Conditional poisson models: a flexible alternative to conditional logistic case cross-over analysis. *BMC Med Res Methodol*. 2014;14:122.
2. Wu Y, Li S, Guo Y. Space-time-stratified case-crossover design in environmental epidemiology study. *Health Data Sci*. 2021;2021:9870798.
3. Gasparrini A. The case time series design. *Epidemiology*. 2021;32:829–837.
4. Gasparrini A. A tutorial on the case time series design for small-area analysis. *BMC Med Res Methodol*. 2022;22:129.
5. Xu Y, Dadvand P, Barrera-Gómez J, et al. Differences on the effect of heat waves on mortality by socio-demographic and urban landscape characteristics. *J Epidemiol Community Health*. 2013;67:519–525.
6. Verbeke G, Spiessens B, Lesaffre E. Conditional linear mixed models. *Am Stat*. 2001;55:25–34.
7. Basagaña X, Escalera-Antezana JP, Dadvand P, et al. High ambient temperatures and risk motor vehicle crashes in Catalonia, Spain (2000-2011): a time-series analysis. *Environ Health Perspect*. 2015;123:1309–1316.
8. Allison PD, Waterman RP. Fixed-effects negative binomial regression models. *Sociol Methodol*. 2002;32:247–265.
9. Lancaster T. The incidental parameter problem since 1948. *J Econometrics*. 2000;95:391–413.
10. Lancaster T. Orthogonal parameters and panel data. *Rev Econ Stud*. 2002;69:647–666.
11. Wooldridge JM. *Econometric Analysis of Cross Section and Panel Data*, 2nd Edition. Cambridge, MA: The MIT Press; 2010.
12. Christensen R. *Log-Linear Models and Logistic Regression*. New York, NY: Springer; 1997.
13. Besag JE, Kooperberg C. On conditional and intrinsic autoregressions. *Biometrika*. 1995;82:733–746.
14. Breslow NE. Extra-Poisson variation in log-linear models. *J R Stat Soc Ser C Appl Stat*. 2018;33:38–44.
15. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit (with discussion). *J R Stat Soc Series B Stat Methodol*. 2002;64:583–616.
16. Basagaña X, Pedersen M, Barrera-Gómez J, et al; ESCAPE Birth Outcomes working group. Analysis of multicentre epidemiological studies: contrasting fixed or random effects modelling and meta-analysis. *Int J Epidemiol*. 2018;47:1343–1354.
17. Sera F, Gasparrini A. Extended two-stage designs for environmental research. *Environ Health*. 2022;21:41.
18. van Aert RCM, Schmid CH, Svensson D, Jackson D. Study specific prediction intervals for random-effects meta-analysis: a tutorial. *Res Synth Methods*. 2021;12:429–447.
19. Bennett JE, Blangiardo M, Fecht D, Elliott P, Ezzati M. Vulnerability to the mortality effects of warm temperature in the districts of England and Wales. *Nat Clim Change*. 2014;4:269–273.
20. Konstantinoudis G, Minelli C, Vicedo-Cabrera AM, Ballester J, Gasparrini A, Blangiardo M. Ambient heat exposure and COPD hospitalisations in England: a nationwide case-crossover study during 2007–2018. *Thorax*. 2022;77:1098–1104.
21. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med*. 2017;36:855–875.
22. Kontopantelis E. A comparison of one-stage vs two-stage individual patient data meta-analysis methods: a simulation study. *Res Synth Methods*. 2018;9:417–430.
23. Dominici F, Peng RD. *Statistical Methods for Environmental Epidemiology with R*. New York, NY: Springer; 2008.
24. Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Stat Med*. 2010;29:2224–2234.

