SPATIAL AND TEMOPRAL VARIABILITY IN SOCIAL, EMOTIONAL AND BEHAVIOURAL DEVELOPMENT OF CHILDREN IN PRESCHOOL

Samantha Ofili, Sarah Barry
University of Strathclyde, Samantha.Ofili@strath.ac.uk;

Background

- Child development is influenced by multiple different contexts such as school and neighbourhoods[1]. Therefore, variation in development can be attributed to individual and contextual differences.
- Observations taken from individuals in the same cluster are expected to be more similar than those taken from individuals in different clusters. The similarities between subjects in the same cluster violate the assumption of independence described in the standard linear model, requiring a different approach.
- The multilevel model is also known as a mixed effect, variance components, random effect, hierarchical or nested model. Multilevel modelling explicitly allows the effects of multiple contexts to be modelled at the same time and the estimation of context specific effects via a small number of parameters using the covariance structure of random effects.
- Previous research using the population mental health dataset Child in Mental Health (ChiME)[2] has shown that after adjustment for demographics, there was electoral geographical clustering of likely mental health difficulties for preschool children in the northeast electoral wards of Glasgow between 2010 to 2012[3].

Aims

Using ChiME data, this study aimed to investigate:

1) How does preschool mental health vary over the years?
2) Is the spatial pattern evident at lower spatial scales?
3) What is the magnitude of the contextual effect for preschool and area on individual outcomes?

Methods

Study Population: 36,674 children attending a local authority or partnership preschool in Glasgow (Fig 1) between 2010 and 2017.

Outcome: The main outcome was total social, emotional and behavioural difficulties measured through the Strengths and Difficulties Questionnaire[5] (SDQ) which ranged from 0-40.

Data Structure: In a cross-classified structure (Fig 2), at each year, children were nested within school and nested in small areas defined by CATTs (consistent areas through time) to overcome boundary changes[6]

Fig 1. Map of Glasgow City Council[4]

Statistical Methods: The SDQ was assumed to have a zero-inflated negative binomial distribution which was modelled by fitting generalised linear mixed spatial convolution model as specified by Barry et al[3].

For years *t*,..,8 with, *i*…36674 individuals, in *k*,…181 preschools and *j*,..,1120 CATT areas:

– **response distribution**: $\mathrm{Y_{ijkt}} \sim ZINB(\lambda, \theta)$ where $\lambda = \frac{\mathrm{r(1-p)}}{\mathrm{p}}$ (1)

– **zero component**: $\mathrm{logit}(\theta) = \gamma_0$ (2)

– **count component**: $\log(\lambda_{ijkt}) = \beta_1 \mathrm{Sex}_{ijkt} + \beta_2 \mathrm{Age}_{ijkt} + \beta_3 \mathrm{Deprivation}_{ijkt} + \beta_4 \mathrm{Year}_t + \boldsymbol{u}_j + v_j + \alpha_k + \delta_t$ (3)

$$\boldsymbol{u} \sim \mathrm{MVN}(0, \Omega_{CAR});\ \text{where } \Omega_{CAR} = \tau^2(\boldsymbol{D} - \boldsymbol{W})^{-1} (4); \quad v_j \sim N(0, \sigma_v^2)\ (5); \quad \alpha_k \sim N(0, \sigma_\alpha^2)\ (6); \quad \delta_t \sim N(0, \boldsymbol{\Omega}_t)\ (7)$$

- Spatial autocorrelation is modelled through the random effect **u** which is given a conditional autoregressive (CAR) distribution where D is the number of neighbours and W is a weighted adjacency matrix (eq. 4). This allows the j^{th} region to deviate from the overall log mean, and is common to all individuals within the region.
- Unstructured spatial effects (eq. 5) and a random intercept pre-school establishment (eq. 6) were modelled through independent and identically distributed Gaussian random effects.
- Temporal correlation was included through a growth curve random intercept with the covariance modelled using a random walk of order 1 (eq. 7).
- Random effects were incrementally added to the model and model fit was assessed using DIC. Eq. 3 shows the linear predictor for the model with the lowest DIC. The Intra-Class correlation Coefficient (ICC) was measured using the Nakagawa et al log-normal approximation approach[7] example shown for preschool in eq. 8.

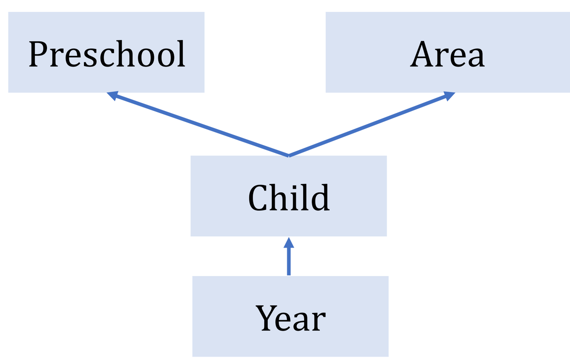


Fig 2. Data Structure

$$\mathrm{ICC_{NB-ln}} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_v^2 + \Omega_{CAR} + \boldsymbol{\Omega}_t + \ln(1 + \frac{1}{\lambda} + \frac{1}{r})} \quad (8)$$

Following a Bayesian approach, where random effect precisions (where precision p= 1/variance) were treated as hyperparameters with uninformative prior distributions the analyses were carried out in R version 4.1.0[8] and INLA[9] .

Results

- Overall, spatial autocorrelation measured through Moran's I was stronger at CATT level (I=0.11, p<0.001) than ward level (I=0.03, p=0.28)
- Table 1 shows total SDQ scores were higher in boys, and increased with age and deprivation.
- Figure 3 shows the distribution of the posterior precision of the random effects.
- Overall, the general contextual effect of area, school and year were low. Table 2 shows the correlation between children in the same context.
- The temporal effect shown is shown Figure 4
- Figure 5 shows combined structured and unstructured effects that was common to all years.

Table 1. Summary of fitted values (on the response scale)

Fixed Effects	Value	95%CI
Intercept		4.85 (4.56-5.04)
Sex	Male	1.37 (1.34-1.40)
Age (years)	4.5-5	0.89 (0.87-0.90)
	5-5.5	1.18 (1.13- 1.24)
	5.5+	1.83 (1.50- 2.26)
Deprivation quintile	4	1.09 (1.05-1.13)
	3	1.15 (1.11-1.20)
	2	1.19 (1.15 -1.24)
	1 = 20% MOST deprived	1.21 (1.16- 1.26)

Table 2. ICC for each random effect

Child	CATT*	School	Year	ICC
i≠i'	j=j'	k≠k'	t≠t'	$6.42e^{-56}$
i≠i'	j≠j	k=k'	t≠t'	1.90^{e-07}
i≠i'	j≠j	k≠k'	t=t'	5.83^{e-103}

**For CATT, the structured and unstructured effects were combined*

Fig 3. Posterior density of the random effect precisions

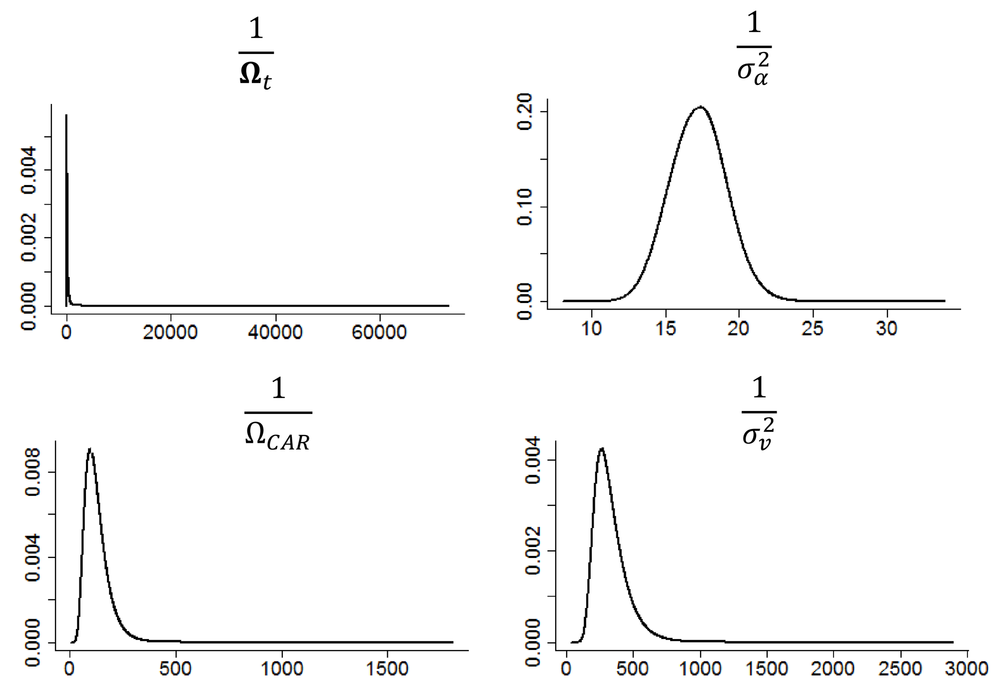


Fig 4. Posterior mean temporal effect with 95% CI

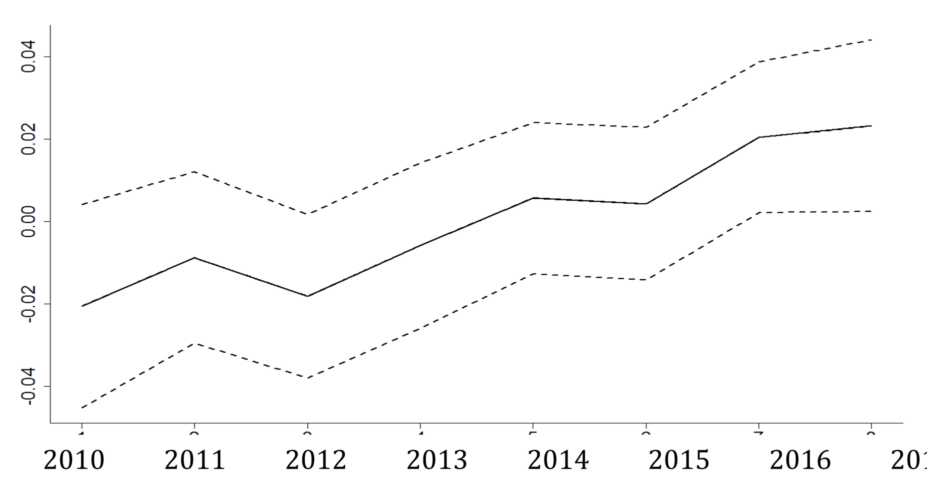
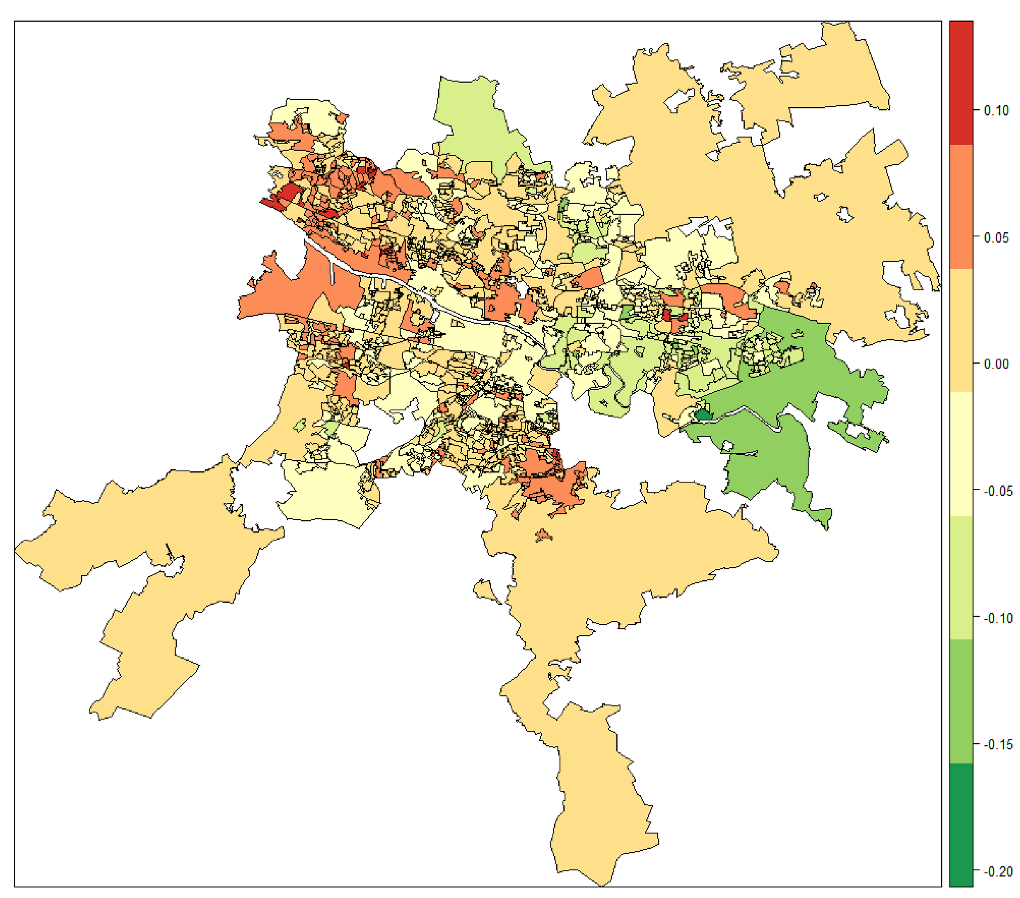


Fig 5. Posterior mean spatial effects



Discussion

- This study aimed to investigate the role of contextual effects on individual differences in development at a finer spatial resolution than previously reported and to explore the temporal structure of the data.
- As in the previous study, there was evidence that boys and those who were more deprived had higher scores. Unlike previously report, here it was found that increasing age led to higher scores.
- While there was an increasing linear trend common across all areas, the size of the effect was small.
- Spatially, the structure differed to that seen at ward level, with living in several areas of the north east of the city having lowering effects on SDQ scores on average.
- Overall there was low correlation between individuals in the same school, year or CATT. Contextual variances were high suggesting there may be shrinkage.
- Future research will explore the sensitivity of the precision terms in the model to the specification of the priors.

References

1. Bronfenbrenner, U. and Morris, P.A., 1998. The ecology of developmental processes.
2. Barry, S.J.E., Marryat, L., Thompson, L., Ellaway, A., White, J., McClung, M. and Wilson, P., 2015. Mapping area variability in social and behavioural difficulties among G lasgow pre-schoolers: linkage of a survey of pre-school staff with routine monitoring data. *Child: care, health and development*, *41*(6), pp.853-864.
3. https://www.gla.ac.uk/researchinstitutes/healthwellbeing/research/mrccsosocialandpublichealthsciencesunit/programmes/settings/educationalsettings/chime/
4. https://maps-glasgow.com/glasgow-city-council-map
5. Goodman, R., 1997. The Strengths and Difficulties Questionnaire: a research note. *Journal of child psychology and psychiatry*, *38*(5), pp.581-586.
6. Exeter, D., Boyle, P.J., Feng, Z., Flowerdew, R. and Scheirloh, N., 2005. The creation of'consistent areas through time'(CATTs) in Scotland, 1981-2001.
7. Nakagawa, S., Johnson, P.C. and Schielzeth, H., 2017. The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, *14*(134), p.20170213.
8. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
9. Rue, H., Martino, S. and Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, *71*(2), pp.319-392.