Survey Methodology, December 1990
Vol. 16, No. 2, pp. 255–269
Statistics Canada

Spatial-Temporal Modelling of Spatially Aggregate Birth Data

DAVID R. BRILLINGER[1]

ABSTRACT

Births by census division are studied via graphs and maps for the province of Saskatchewan for the years 1986-87. The goal of the work is to see how births are related to time and geography by obtaining contour maps that display the birth phenomenon in a smooth fashion. A principal difficulty arising is that the data are aggregate. A secondary goal is to examine the extent to which the Poisson-lognormal can replace for data that are counts, the normal regression model for continuous variates. To this end a hierarchy of models for count-valued random variates are fit to the birth data by maximum likelihood. These models include: the simple Poisson, the Poisson with year and weekday effects and the Poisson-lognormal with year and weekday effects. The use of the Poisson-lognormal is motivated by the idea that important covariates are unavailable to include in the fitting. As the discussion indicates, the work is preliminary.

KEY WORDS: Aggregate data; Borrowing strength; Contouring; Extra-Poisson variation; Locally-weighted analysis; Maps; Periodogram; Poisson distribution; Poisson-lognormal distribution; Random effects; Spatial data; Time series; Unmeasured covariates.

1. INTRODUCTION

The concern of this work is spatial-temporal data, that is quantities recorded as functions of space and time. The analysis of such data should be "easy" because of the graphing possibilities, *e.g.* rate versus time or effect versus geography, in the manner of residual plots so often employed in regression analysis; however in the present case the aggregation of basic elements leads to substantial difficulties.

The specific data studied consists of daily births for the calendar years 1986 and 1987 to women aged 25-29 for each of the 18 census divisions of the province of Saskatchewan. The corresponding population sizes, as determined in the 1986 Census, are also employed in order to compute rates. The reason that Saskatchewan was selected for this pilot study is that it is moderate sized and its boundaries and those of its census divisions are fairly regular. (The latter was important at the early stages of the work because computer based maps were then unavailable). Women aged 25-29 were selected because that was the 5 year age group with most births. These data were provided to the author by Statistics Canada. They are characterized by being aggregate, by being non Gaussian and by being non stationary in space and time.

It is wished to understand the relationship of births to time and geography, specifically to allow temporal and spatial patterns of fertility and possible surprises to show themselves. There are two central aspects to the study; a locally-weighted analysis of aggregate data is developed and random effects models are set down and fit to handle extra-Poisson variation. The latter part may be viewed as an inquiry into the flexibility of the Poisson-lognormal to handle unmeasured covariates and errors. The locally weighted analysis proceeds by developing weights, $w_i(x,y)$, that are meant to reflect the influence of the i-th census division (an aggregate) on the point location with coordinates (x,y). Given census division data, these

[1] David R. Brillinger, Statistics Department, University of California, Berkeley, CA, 94720, U.S.A.

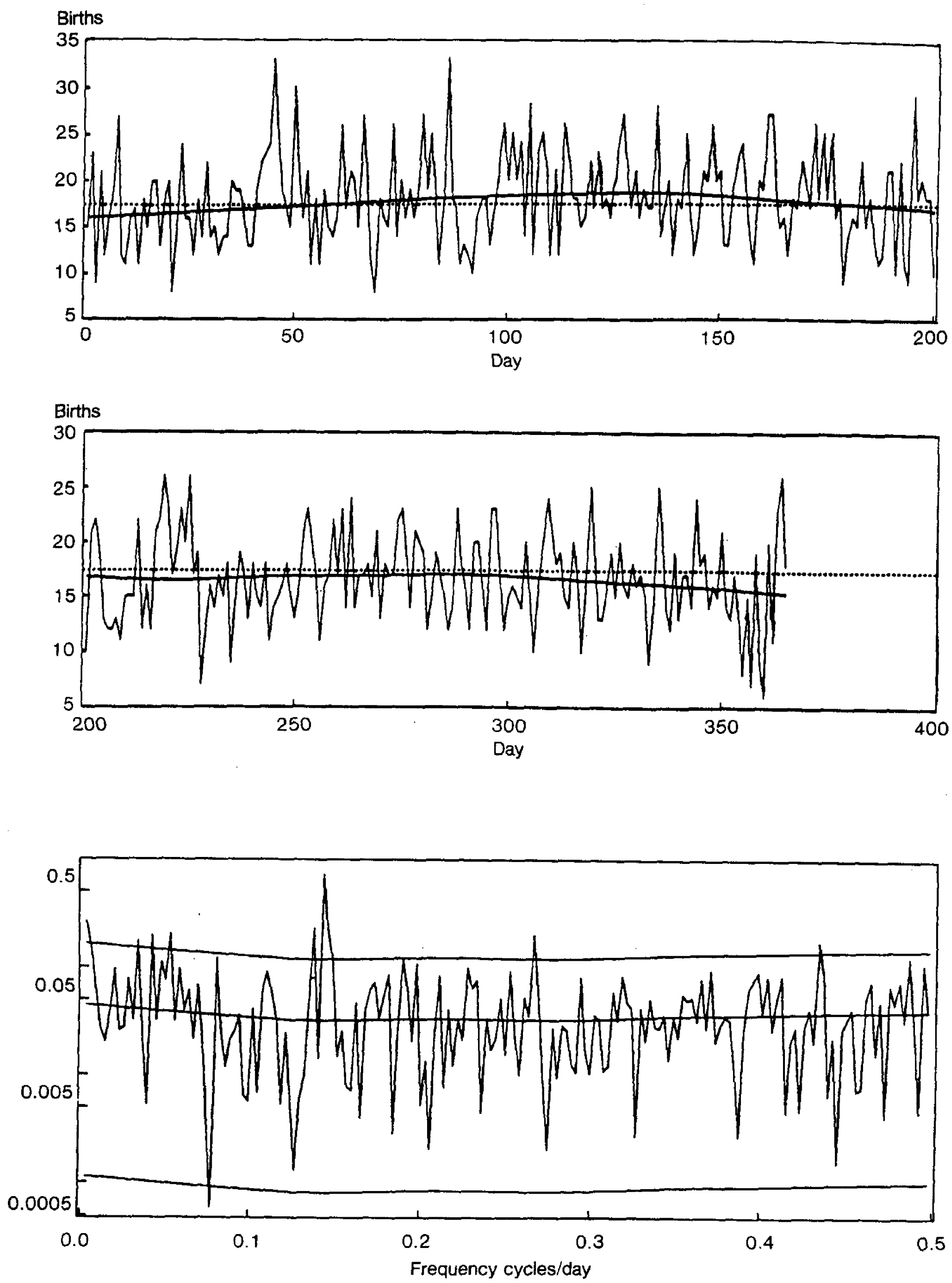

Figure 1. Top: Time series of annual births to women aged 25-29 in 1986 for the Province of Saskatchewan. Bottom: Periodogram of the square roots of the count graphed above. The solid lines provide approximate 95% marginal confidence limits. The peak corresponds to a period of 7 days.

weights are then applied to individual terms of the log-likelihood or corresponding estimation equations and parameter estimates evaluated.

It is to be emphasized that this is a preliminary report on work in progress. For example the fine structure of the data is not taken advantage of and no measures of uncertainty of the various estimates have been provided. The expressions employed for the weights, in this present work, are naive and bound to change form with further study, but the character of the analysis may be anticipated to remain of some interest.

The companion paper Brillinger (1990) considers some aspects of the spatial case alone.

2. BIRTHS AS A TIME SERIES

The top graph of Figure 1 provides the total number of births in Saskatchewan for each day of 1986. The dashed line is the 1986 mean level. The solid line is the result of heavily smoothing the series and is meant to highlight any trend. This graph does not, with casual inspection, provide striking evidence of any special phenomenon. However when the periodogram of the square root of the counts is computed, see bottom graph of Figure 1, something of interest appears. (The square root is employed to make the series more nearly symmetrical and normal). The upper and lower solid lines on the graph provide approximate 95% marginal confidence limits about a heavily smoothed version. A peak is apparent at a frequency of .143 cycles/day corresponding to a period of 7 days. This periodic phenomenon is well known in the analysis of birth data, see *e.g.* Cohen (1983) and Miyaoka (1989) and references therein. It is usually ascribed to doctors intervening in the natural process of labour and inducing births particularly on weekdays.

3. BIRTHS AS A SPATIAL PROCESS

Figure 2 provides, for each census division, and for women aged 25 to 29 the annual rate of births for the years 1986 and 1987 combined. One sees the highest rate of .208 births per woman per year to occur in the northern half of the province while the two lowest rates appear in the census divisions containing Regina and Saskatoon.

Figure 3 provides the numerical difference between the annual rate for 1987 and that for 1986 for each of the 18 census divisions. (Note that the 1986 census population has been taken as the divisor in each case). The differences are scattered around 0. It is to be noted that these rates are, however, based on fairly widely varying population sizes.

In the previous section the presence of a phenomenon of period 7 days was noted. Figure 4 presents the difference between the average weekday rate and the average weekend rate, (weekdays meaning Monday through Friday) for each census division. In all but one census division, the weekday rate is higher. This is consistent with various other studies and, as suggested in Section 2, is very possibly due to doctors inducing labor on weekdays (to avoid births on weekends).

The various rates presented in Figures 2, 3, 4 are average values for individual census divisions.

4. PROBLEMS ARISING

Maps of most quantities of direct interest that assign average values to the wholes of counties thereby lie, lie, lie.

With these graphic words Tukey (1979) deplores the use of maps such as those of Figures 2, 3, 4 that are constant across geographic divisions. Indeed examination of Figure 2, as does common knowledge, suggests that the birth phenomenon quite likely varies smoothly across census division boundaries. A principal concern of this work is to develop contour maps displaying smooth variation. It is hoped that such maps will prove useful in the discovery of general stochastic descriptions of the phenomenon and will allow insightful exploratory analyses.

A second concern of this work is with the statistical distribution of the counts themselves. A natural special stochastic model to employ is the Poisson. Yet in past studies the birth process has been found to relate to many socio-economic quantities, *e.g.* diet, lifestyle, weather, environment, weekday, holidays, age structure. Further the population of the various census divisions has varied around the Census Day values throughout 1986-87 and lastly the women's ages are scattered from 25 to 29. In summary it seems necessary to employ a more flexible model than the Poisson, specifically a model able to handle omitted covariates. The Poisson-lognormal will be employed in this work. As a sideline, due to the presence of the standard deviation parameter in the Poisson-lognormal, there will be a borrowing of strength that takes place in combining the data values, in the manner described by Mallows and Tukey (1982). (The term "borrowing strength" is employed, rather than for example "empirical Bayes" as some might prefer, because it has been in use for a substantial time period and because of its broader implications). Dean *et al.* (1989) is another recent reference concerned with handling extra-variation.

5. LOCALLY-WEIGHTED ANALYSIS

In the case of nonaggregate data, locally-weighted fitting is a convenient fashion by which to estimate smoothly varying quantities. Suppose one has a variate Y with probability distribution $p(Y \mid \Theta)$ depending on the finite dimensional parameter Θ. Suppose one wishes an estimate of Θ particular to the location with coordinates (x,y). Suppose the datum Y_i is available for location (x_i,y_i). Let $W_i(x,y)$ be a weight dependent on the distance of (x_i,y_i) to (x,y).

Consider estimating Θ by maximizing the weighted log-likelihood

$$\sum_i W_i(x,y) \log p(Y_i \mid \Theta) \tag{1}$$

or (often equivalently) by solving the system of estimating equations

$$\sum_i W_i(x,y) \Psi(Y_i \mid \hat{\Theta}) = 0 \tag{2}$$

with $\Psi(Y \mid \Theta) = \partial \log p / \partial \Theta$, the score function.

To illustrate the technique consider an elementary case, specifically take Y to be normal with mean μ and variance σ^2. The locally weighted estimate of μ at (x,y) results from minimizing

$$\sum_i W_i(x,y) [Y_i - \mu]^2$$

and is given by

$$\hat{\mu}(x,y) = \sum_i W_i(x,y) Y_i \Big/ \sum_i W_i(x,y),$$

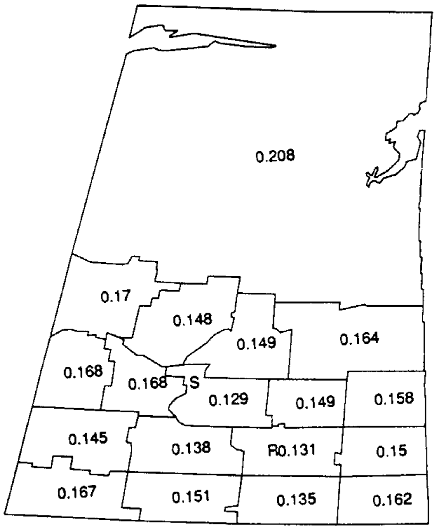

Figure 2. The average annual birth rate for women aged 25 to 29 for the years 1986 and 1987, plotted above census divisions. "R" and "S" indicate the locations of Regina and Saskatoon respectively.

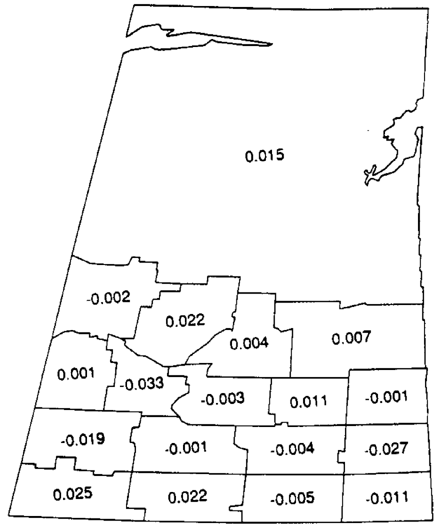

Figure 3. The 1987 rate minus the 1986 rate for the same data as Figure 2.

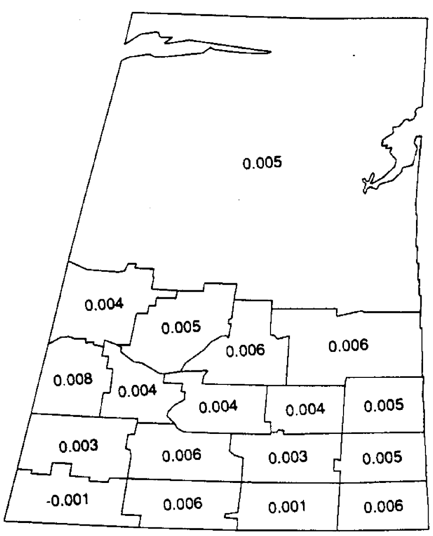

Figure 4. The average weekday rate minus the average weekend rate for the same data as Figure 2.

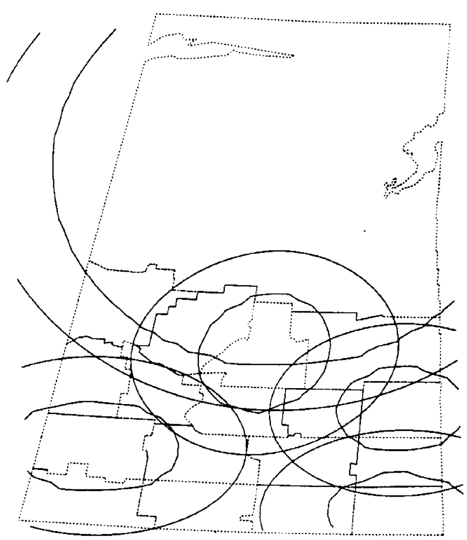

Figure 5. The weights, $W_i(x,y)$ applied in equations (1) or (2), computed via expression (4), for four of the census divisions. The weights are not shown for all the divisions in the interests of clarity. The contours at levels .50 and .99 are shown.

an expression with intuitive appeal. It is to be noted that such formulas are commonly used in computer graphics as interpolation procedures, see for example Franke (1982).

Among references we may mention Gilchrist (1967) concerned with "discounting", Pelto *et al.* (1968), concerned with least squares, Cleveland and Kleiner (1975), who suggested the use of moving midmeans and Stone (1977) focusing on regression. In the discussion of Stone's paper, Brillinger (1977) suggested the form (2) for a general distribution and justified it as a Bayes' rule. Specifically consider the loss function

$$L(Y \mid \Theta) = -\log p(Y \mid \Theta).$$

Suppose an estimate is desired at $\mathbf{r} = (x,y)$. The Bayes' risk may be written

$$E\{L(Y \mid \Theta_{\mathbf{r}})\} = E\{E\{L(Y \mid \Theta_{\mathbf{r}}) \mid \mathbf{r}\}\}.$$

Bayes' rule seeks

$$\min_{\Theta} E\{L(Y \mid \Theta) \mid \mathbf{r}\}.$$

With data Y_i, $\mathbf{r}_i$, and $W_i(\mathbf{r})$ a kernel centred at $\mathbf{r}_i$, one approximates the conditional expected value here by

$$E\{\log p(Y \mid \Theta) \mid \mathbf{r}\} \approx \sum_i W_i(\mathbf{r}) \log p(Y_i \mid \Theta)$$

and so is led to expression (1).

Tibshirani and Hastie (1987) develop an equi-weighted local likelihood estimation procedure. Cleveland and Devlin (1988) develop the least squares approach in real detail. Staniswalis (1989) studies and implements the general p case. Advantages of the locally-weighted technique include: no "hidden model" distribution assumption, the possibility of discerning non-additivity, variants for resistance and influence, simple additivity of the observation component, and no matrix inversion (as, for example, kriging requires).

The birth data of concern in this work is aggregate (or grouped) totals over census divisions. The procedure of the preceding section cannot therefore be employed directly. The problem is that of obtaining appropriate weights $w_i(x,y)$ evidencing the effect of the census division i on the location (x,y). Suppose $|R_i|$ denotes the area of census division i. Then the naive weight function is

$$w_i(x,y) = \frac{1}{|R_i|} \quad \text{for } (x,y) \text{ in } R_i$$

and equal 0 otherwise. In this work functions of the essential form

$$w_i(x,y) = \frac{1}{|R_i|} \int_{R_i} W(x-u, y-v)\,du\,dv \tag{3}$$

will be employed where $W(\cdot)$ is a kernel appropriate for the nonaggregate case as for example studied in Cleveland and Devlin (1988). The formula (3) may be motivated by consideration of the Poisson point process case, see Appendix II. Estimates will be determined via the criteria (1) or (2) with W_i replaced by w_i.

The specific weights employed at $\mathbf{r} = (x,y)$ in this preliminary work are

$$w_i(\mathbf{r}) = \exp\{-(1-\rho)^2\|\mathbf{r} - \mathbf{r}_i\|^2/2\tau^2\} \qquad (4)$$

outside the ellipse $(\mathbf{r}_0 - \bar{\mathbf{r}}_i)\mathbf{S}_i^{-1}(\mathbf{r}_0 - \mathbf{r}_i)' = d_0^2 = 5.991$ and equal 1 inside. Here $\|\mathbf{r}\|^2 = x^2 + y^2$, $\rho = d_0/\sqrt{(\mathbf{r} - \bar{\mathbf{r}}_i)\mathbf{S}_i^{-1}(\mathbf{r} - \bar{\mathbf{r}}_i)'}$ and $\tau = .025$, where $\bar{\mathbf{r}}_i = E\,U_i$ and $\mathbf{S}_i = \operatorname{var} U_i$ with U_i a variate uniformly distributed within R_i. This choice of ρ makes the weight function continuous. The logic is that the census divisions are approximated by ellipses with the same mean and variance-covariance matrix. (The specific values were chosen after a bit of experimentation, in part to make the area in the initial ellipse about .95 of the division's). One could have employed other shapes than ellipses, *e.g.* rectangles, but this is preliminary work and it is anticipated that later work will employ weights of the form (3).

Figure 5 displays the .50 and .99 contours of the $w_i(x,y)$ plotted for several of the census divisions. The contours are seen to follow the general shapes of the census divisions. The jaggedness in some of the contours results from the discreteness of the 40×40 grid employed in the computations.

Other weight functions constructed with somewhat similar problems in mind may be found in Tobler (1979) and Dyn and Wahba (1982). Advantages of the present approach, as listed for the nonaggregate case above include: the terms in (1) or (2) are additive and do not interact, no matrix inversion is needed, and resistance to outliers is easily built in.

Cliff and Ord (1975) Section 5.1, discusses measures of the influence of counties on other counties. The concern of this present paper however is the influence of a "county" on a point location. It is to be remarked that perhaps the weight, providing the influence, should depend on some covariates, *e.g.* county population.

6. A POISSON FIT

Throughout the analysis, the female population aged 25-29 and births to its members will be considered. Let $i = 1, \ldots, 18$ index census division. Let N_i denote the census count of the women in the i-th division. (These are the counts for Census Day, 3 June 1986). Let B_i denote the total number of births to women aged 25-29 in the two years 1986-87.

Suppose that the probability distribution $p(\cdot)$ of Section 5 is that B_i is Poisson with mean $2N_i\mu$. (The presence of the multiplier 2 is so the parameter μ is an annual birth rate). One logic for the Poisson assumption comes from the idea that birthdays are random, see Brillinger (1986).

With the Poisson assumption, the locally weighted estimate of the annual birth rate at location (x,y) is given by

$$\hat{\mu}(x,y) = \sum_i w_i(x,y)\,B_i \Big/ 2\sum_i w_i(x,y)N_i. \qquad (5)$$

These values are computed for (x,y) on a 40 by 40 grid and the corresponding contour plot is given in Figure 6. The contours are seen to vary smoothly. This (smoothed) rate varies from .14 to .20, with the higher values in the upper half and the lower centred around the Province's most urban part.

As indicated previously, the data under study has important temporal characteristics. Models need to take this into account. In particular the weekly periodicity needs to be handled as well as possible trends in population sizes. The following model seems worth considering. Let j be an indicator variable with $j = 1$ if the count is for a weekday and $j = 2$ if the count is for

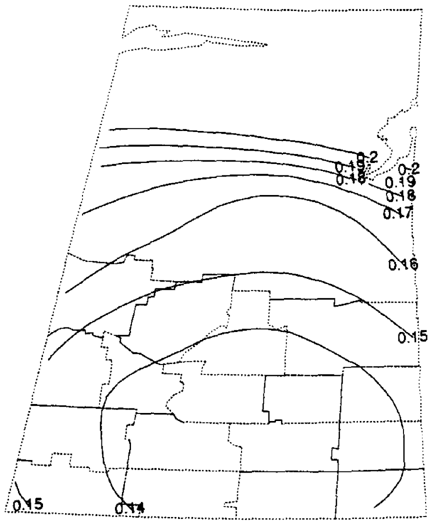

Figure 6. Expression (5) graphed for the weights of (4) with B_i the count of births in census division i during 1986-87 and N_i the corresponding population count of women aged 25-29.

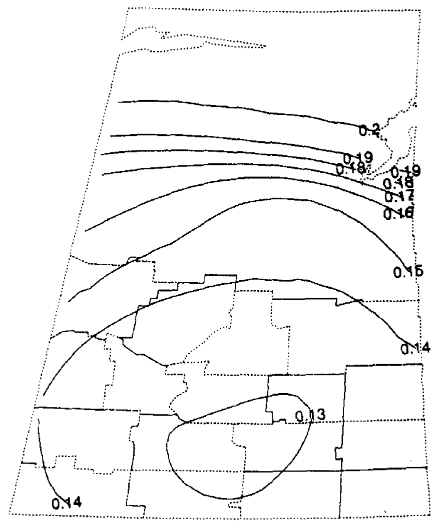

Figure 7. The estimated birth rate $\exp\{\hat{\alpha}\}$ obtained by locally weighted fitting assuming that the number of births, B, given the population at risk, N, is Poisson with mean $N \exp\{\alpha \pm \beta \pm \gamma\}$ with the first $\pm$ sign plus for weekdays and minus for weekends and the second $\pm$ plus for 1986 and minus for 1987.

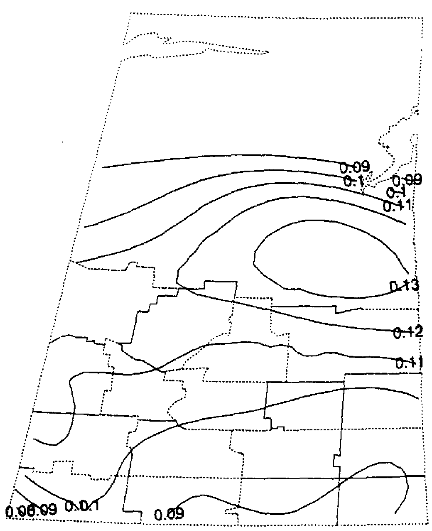

Figure 8. Plot of the estimated weekday effect $\hat{\beta}(x,y)$ obtained as per Figure 7.

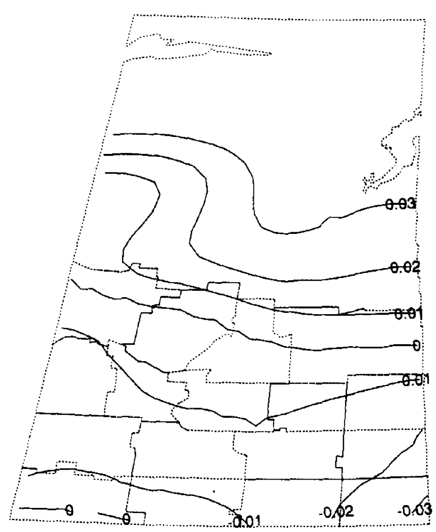

Figure 9. The estimated year effect $\hat{\gamma}(x,y)$ as per Figure 7.

a weekend. Let k be a second indicator variable with $k = 1$ for 1986 and $k = 2$ for 1987. Let B_{ijk} denote the corresponding number of births in census division i. Suppose that B_{ijk} given N_i is Poisson with mean $N_i \exp\{\alpha + \beta_j + \gamma_k\}$. β_j is the weekday effect, γ_k the year effect and it will be assumed that $\beta_1 + \beta_2, \gamma_1 + \gamma_2 = 0$ to make the model identifiable. If there is no weekday effect, then $\beta_1, \beta_2 = 0$. If there is no year effect, then $\gamma_1, \gamma_2 = 0$. Now, via locally-weighted analysis presented in Section 5, one can obtain estimates of α, β and γ as functions of location (x,y). (For simple balance in the computations, only the first $364 = 7 \times 52$ days of each year have been employed).

Figure 7 provides the estimate $\exp\{\hat{\alpha}(x,y)\}$ obtained of the annual birth rate. It is interesting to note that, relative to the constant rate Poisson model, the contours have expanded out somewhat from the urban areas. Figure 8 provides the estimated weekday effect, $\hat{\beta}_1(x,y)$, obtained. In its case there is bulge to the east. These values are quite a different representation from that of the naive differences of Figure 4. In particular, now there is a reflection of the differing population sizes. The order of magnitude of the $\hat{\beta}$'s is .08 to .13 while $\hat{\alpha}$ is order -2.1 to -1.6. Figure 9 provides the estimated year effect, $\hat{\gamma}_1(x,y)$. Its values vary from $-.03$ to .03. Numerically, the weekday-weekend effect is the larger.

The just preceding analysis suggests that there are basic variables that can affect birth rates and that modelling and analysis needs to take this circumstance into account.

7. POISSON-LOGNORMAL FITS

With a multi-dimensional explanatory variable $\mathbf{x}$ in hand, a Poisson model that has B of mean $N\exp\{\mathbf{x}\Theta\}$ might do a good job of explaining the data. Examples of explanatory variables include: diet, lifestyle, weather, environment, holidays, population change, age structure, vagaries of boundaries. In the present situation, these variables are not at hand. The omitted variables in the model will be assumed specifically accumulated into an error variable. It will be assumed that, given ϵ, the variate B is Poisson with mean $N\mu\exp\{\epsilon\}$ and that ϵ is normal with mean 0 and variance σ^2. In the case of this model B is said to have a Poisson-lognormal distribution. Some information on this distribution may be found in Shaban (1988). Sometimes ϵ enters directly from the problem context, see Brillinger and Preisler (1983) for one example, but in the present case it is simply assumed present.

A critical difficulty, that arises in working with a Poisson-lognormal model, is that closed expressions do not exist for the probability function. Yet the model is clearly flexible for introducing effects and handling unavailable variables. Following the work of Bock and Lieberman (1970), Pierce and Sands (1975) and Hinde (1982), one can proceed via numerical quadrature. The probability function may be written

$$p(Y) = \frac{1}{Y!}\int (ve^{\sigma z})^Y \exp\{-ve^{\sigma z}\}\phi(z)dz$$

with ϕ the standard normal density, with Y corresponding to B and with v corresponding to $N\mu$. To proceed with a data analysis the integral is approximated by a finite number of terms involving nodes, z_l, and weights, w_l,

$$p(Y) \approx \frac{1}{Y!}\sum_{l=1}^{L} (ve^{\sigma z_l})^Y \exp\{-ve^{\sigma z_l}\}\, w_l.$$

Listings of nodes and weights may be found in Abramowitz and Stegun (1964) for example.

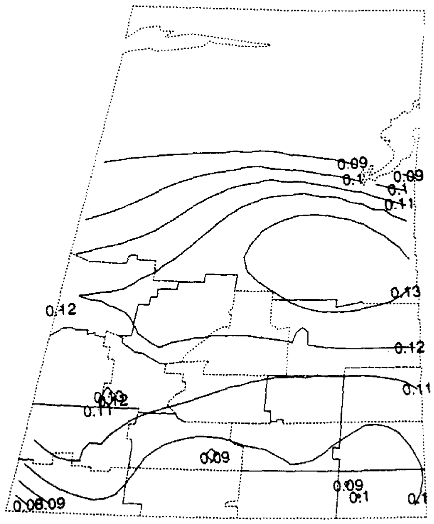

Figure 10. A plot comparable to Figure 7, except that now a normal error term is added to the linear predictor.

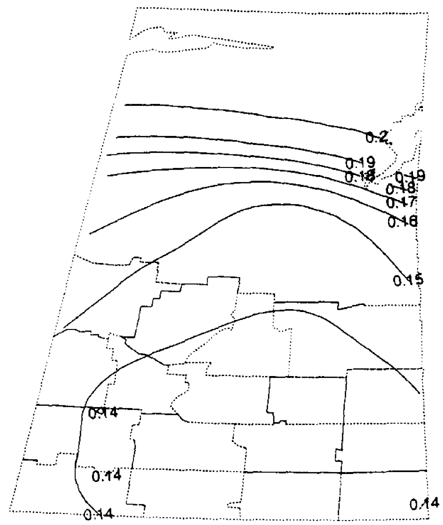

Figure 11. A plot comparable to Figure 8, except now (as in Figure 10) a normal error term has been added to the linear predictor.

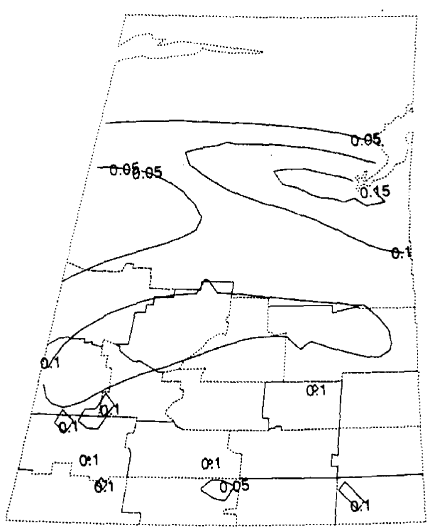

Figure 12. A plot comparable to Figure 9, except now (as in Figure 10) a normal error term has been added to the linear predictor.

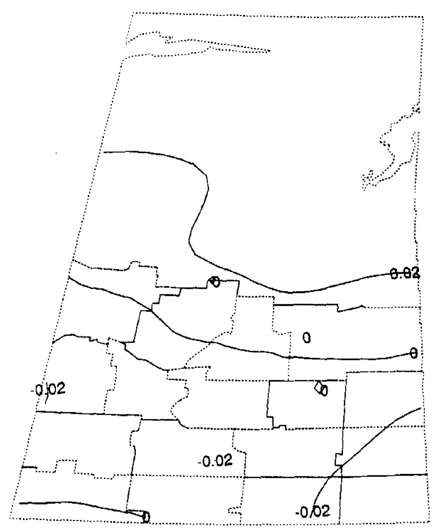

Figure 13. The estimated standard error, $\hat{\sigma}(x,y)$, of the normal term added to the linear predictor.

Figures 10, 11, 12, 13 provide the results of fitting the Poisson-lognormal model including weekday and year effects and employing $L = 5$ nodes. The model assumes B_{ijk} given N_i and Z is Poisson with mean

$$N_i \exp\{\alpha + \beta_j + \gamma_k + \sigma Z\}$$

Z denoting a standard normal deviate and further assumes the separate Z's independent. Here i indexes census division, j weekday or not and k year. Figure 10, a contour plot of $\exp\{\hat{\alpha}(x,y)\}$, again shows a dip around the urban region as in Figure 7. The irregularity in the figure suggests that in one case perhaps the estimation procedure converged to a local extremum. Figures 11 and 12 similarly provide $\hat{\beta}(x,y)$ and $\hat{\gamma}(x,y)$. There are again suggestions of local extrema. Figure 13, a contour plot of $\hat{\sigma}(x,y)$, is not easily described. It suggests that the estimate, $\hat{\sigma}$, is fairly variable. The estimate is seen to be of order of magnitude .1 and so comparable to the weekday effect of Section 6.

All the work on estimation with the Poisson-lognormal, that we know about, involves some form of approximation. For example Clayton and Kaldor (1987) approximate the conditional Poisson log-likelihood by a quadratic and Aitchison and Ho (1989) also employ numerical integration, albeit after a transformation of the parameters. A new type of approximation has recently been proposed in Crouch and Spiegelman (1990). Its effectiveness for the Poisson-lognormal remains to be studied.

8. DISCUSSION

Locally-weighted analysis and random effect models appear to provide a flexible means of dealing with a broad class of problems involving geographic data. The random effect terms have two important roles: handling omitted effects and borrowing strength for improved estimates of the principal parameters. For the Poisson alone, naive totals are efficient, yet there exists extra-Poisson variability due to omitted variables in the present case.

The approach is computer intensive, because of the numerical integration and the maximum likelihood estimation at many points on a grid, but proved quite manageable on the Berkeley network of Sun 3/50's.

Much future work remains including: tools for assessing fit, uncertainty computation and display, weight function choice (particularly choice of τ in (4)), analyses for other age groups and provinces, and appropriate asymptotics. Further understanding needs to be gained as to why with nearby initial values the optimizing routine sometimes converged to somewhat distant estimates. An advantage of the present circumstance is that there exists immense amounts of other data to be made use of as work progresses. Examination of Figures 6 on shows an important limitation of the technique – it is providing too much fine detail in the northern half of the province.

Other recent papers devoted to the analysis of vital statistics rates are: Cressie and Read (1989), Clayton and Kaldor (1987), Tsutakawa (1988) and Manton *et al.* (1989). These papers are however not directed at the problem of obtaining a smooth surface, which is the concern of this work.

It is amusing to note that the presence of the weekly period in the phenomenon allowed the author to deduce early on in the work that a confusion had arisen over which data set was to be supplied. When the days of fewest births were determined for the initial data set supplied, the days were found to be (apparently) Friday and Saturday. This was because the year 1987 had been supplied, and not the desired 1986.

After the analyses were completed it was learned that the birth counts were based on 1981 census divisions, while the population counts were based on 1986. Luckily the boundaries have not changed much, but this circumstance provides yet more reason for wanting a procedure that can handle extra-variation.

9. ADDENDUM

In the paper a case has been made for the inclusion of an error term, ϵ, to reflect pertinent covariates that were unavailable for the analysis. This led to the employment of the Poisson-lognormal distribution. In Tukey (1990) an index of urbanicity of a census division is constructed. It is based on the populations of the three largest places in the division. The values, x_i, of the index are given in Figure 14 and are seen to be lowest in the census divisions containing Regina and Saskatoon.

The table below gives the results of employing Glim to fit the successive Poisson models for B_{ijk} given N_i: (i) $N_i \exp\{\alpha + \beta_j + \gamma_k\}$, (ii) $N_i \exp\{\alpha + \beta_j + \gamma_k + \delta x_i\}$, and (iii) $N_i \exp\{\alpha + \beta_j + \gamma_k + \delta_1 x_i + \delta_2 x_i^2\}$.

Variables	Deviance	d.f.	*p*-value
weekday, year	227.3	69	
+ urbanicity	86.69	68	
+ urbanicity**2	83.13	67	.088

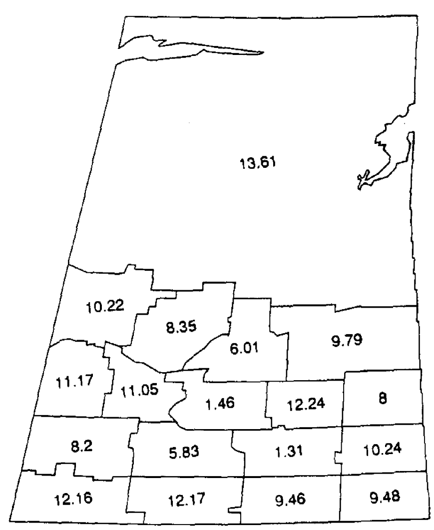

Figure 14. The values of the Tukey index of urbanicity.

By bringing in this urbanicity variable, x_i, now a Poisson model is satisfactory for the circumstance.

Finally the Referee made some comments that spell out quite specifically the assumptions and limitations of this present study. The work is continuing and the intention is to address these comments. Rather than paraphrasing, it seems more sensible to provide the referee's own words.

"The choice of weights is *ad hoc* and requires more thought. If one had two divisions, both of the same area but with vastly different populations N_i, should the weighting be the same? It depends on whether area or population density is thought to be more important. Use of the latter may remove the spurious fine detail in the northern half of the province."

"There are traps with N_i's, which the author appears to be aware of, but I think the reader needs extra warning. It might help to have approximate measures of uncertainty ([Section 1] promises none). Figure 3 cannot really be interpreted, since positive or negative values may be due to random fluctuations about zero. The contours in Figure 6 are calculated with vastly different precision, and in some respects are incomparable. And, [in Section 6], upon estimating α, β and γ, it would be tempting (but unwise) to assume that such values are significant."

"All random variables in sight are assumed independent. Another way to motivate these weighted models is to assume a multivariate distribution, with the property that the conditional mean at (x,y), given the surrounding data, is a weighted combination of those data. Then the joint distribution exhibits dependence."

ACKNOWLEDGEMENTS

The author would like to thank G. Brackstone, R. Gussela, R. Raby, B. Sander, P. Spector, M. Subhani, R. Villani for assistance in obtaining the data and maps, for help with computational geometry and with parallel computing. John Tukey, Rob Tibshirani and the Referee made very helpful comments on the first draft. The research was supported by National Science Foundation Grant DMS-8900613.

APPENDIX I

In this Appendix a few computing details are provided. The census divisions and the province boundaries are specified as polygons. To compute the weights $w_i(x,y)$ an algorithm was required to check whether a given point was inside a given polygon. To compute the mean and variance of a random point inside a given polygon, an algorithm for breaking a polygon up into triangles was required. Such algorithms are discussed in Preparata and Shamos (1985) for example. The approximate likelihood was maximized via the Harwell FORTRAN routine va09a. For the parallel computations the 40 by 40 grid was broken up into 20 disjoint segments and the computations thence carried out on 20 separate work stations. As in Brillinger and Preisler (1983), factors were introduced into the likelihood to stabilize the computations. Miyaoka (1989) found that the computations could be sensitive to the number of nodes employed. In the present series of computations, the number was increased until the results did not change much. There is also the problem of selecting inital values. Here they were taken to be the method of moment estimates, although these are perhaps too inefficient.

APPENDIX II

For simplicity, consider the case of a point process $\{x_j\}$ with rate function ν on the line. The local weighted log likelihood for a Poisson process is, up to a constant,

$$\sum_j W(x - x_j) \log \nu(x_j) - \int W(x - u)\nu(u)du.$$

So, the locally weighted estimate of the rate is

$$\hat{\nu}(x) = \sum_j W(x - x_j) \Big/ \int W(x - u)du,$$

the usual form of estimate. Suppose now the line is broken into intervals R_i, and the aggregate count available is $N(R_i)$. One desires

$$\sum_{x_j \in R_i} W(x - x_j).$$

If this last is to be approximated by $\Theta N(R_i)$, then the method of moments leads to

$$\Theta = \int_{R_i} W(x - u)du \Big/ |R_i|$$

and thence to expression (3).

REFERENCES

ABRAMOWITZ, M., and STEGUN, I.A. (1964). *Handbook of Mathematical Functions.* National Bureau of Standards, Washington.

AITCHISON, J., and HO, C.H. (1989). The multivariate Poisson-log normal distribution. *Biometrika,* 76, 643-653.

BOCK, R.D., and LIEBERMAN, M. (1970). Fitting a response model for *n* dichotomously scored items. *Psychometrika,* 35, 179-197.

BRILLINGER, D.R. (1977). Discussion of Stone (1977). *The Annals of Statistics,* 5, 622-623.

BRILLINGER, D.R. (1986). The natural variability of vital rates and associated statistics. *Biometrics,* 42, 693-734.

BRILLINGER, D.R. (1990). Mapping aggregate birth data. Technical report, Statistics Department, University of California, Berkeley.

BRILLINGER, D.R., and PREISLER, H.K. (1983). Maximum likelihood estimation in a latent variable problem. *Studies in Econometrics, Time Series and Multivariate Statistics,* 31-65. New York: Acadamic Press.

CLAYTON, D., and KALDOR, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics,* 43, 671-681.

CLEVELAND, W.S., and DEVLIN, S.J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association,* 83, 596-610.

CLEVELAND, W.S., and KLEINER, B. (1975). A graphical technique for enhancing scatterplots with moving statistics. *Technometrics,* 17, 447-454.

CLIFF, A.D., and ORD, J.K. (1975). Model building and the analysis of spatial pattern in human geography. *Journal of the Royal Statistical Society*, 37, 297-348.

COHEN, A. (1983). Seasonal daily effect on the number of births in Israel. *Applied Statistics*, 32, 228-235.

CRESSIE, N., and READ, T.R.C. (1989). Spatial data analysis of regional counts. *Biometrical Journal*, 31, 699-719.

CROUCH, E.A.C., and SPIEGELMAN, D. (1990). The evaluation of integrals of the form $\int_{-\infty}^{\infty} f(t) \exp\{-t^2\} dt$: application to logistic-normal models. *Journal of the American Statistical Association*, 85, 464-469.

DEAN, C., LAWLESS, J.F., and WILLMOT, G.E. (1989). A mixed Poisson-inverse-Gaussian regression model. *Canadian Journal of Statistics*, 17, 171-182.

DYN, N., and WAHBA, G. (1982). On the estimation of functions of several variables from aggregated data. *SIAM Journal on Mathematical Analysis*, 13, 134-152.

FRANKE, R. (1982). Scattered data interpolation: tests of some methods. *Mathematics of Computation*, 38, 181-200.

GILCHRIST, W.G. (1967). Methods of estimation involving discounting. *Journal of the Royal Statistical Society*, 29, 355-369.

HINDE, J. (1982). Compound regression models. GLIM 82 (Ed. R. Gilchrist), 109-121. *Lecture Notes in Statistics*, 14. New York: Springer-Verlag.

MALLOWS, C., and TUKEY, J.W. (1982). An overview of techniques of data analysis emphasizing its exploratory aspects. *Some Recent Advances in Statistics* (Eds. Tiago de Oliveira, J. *et al.*), 111-172. London: Academic.

MANTON, K.G., WOODBURY, M.A., STALLARD, E., RIGGAN, W.B., CREASON, J.P., and PELOM, A.C. (1989). Empirical Bayes procedures for stabilizing maps of U. S. cancer mortality rates. *Journal of the American Statistical Association*, 84, 637-650.

MIYAOKA, E. (1989). Application of mixed Poisson-process models to some Canadian birth data. *Canadian Journal Statistics*, 17, 123-140.

PELTO, C.R., ELKINS, T.A., and BOYD, H.A. (1968). Automatic contouring of irregularly spaced data. *Geophysics*, 33, 424-430.

PIERCE, D.A., and SANDS, B.R. (1975). Extra-binomial variation in binary data. Technical Report 46, Statistics Department, Oregon State University.

PREPARATA, F.P., and SHAMOS, I. (1985). *Computational Geometry*. New York: Springer-Verlag.

SHABAN, S.A. (1988). Poisson-lognormal distributions. *Lognormal Distributions*, 195-210 (Eds. E.L. Crow and K. Shimizu). New York: Marcel Dekker.

STANISWALIS, J.G. (1989), The kernel estimate of a regression function in likelihood-based models. *Journal of the American Statistical Association*, 84, 276-283.

STONE, C.J. (1977). Consistent nonparametric regression. *The Annals of Statistics*, 5, 595-620.

TIBSHIRANI, R., and HASTIE, T. (1987). Local likelihood estimation. *Journal of the American Statistical Association*, 82, 559-567.

TOBLER, W.R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Assocation*, 74, 519-536.

TSUTAKAWA, R.K. (1988). Mixed model for analyzing geographic variability in mortality rates. *Journal of the American Statistical Association*, 83, 37-42.

TUKEY, J.W. (1979). Statistical mapping: what should not be plotted. Proc. 1976 Workshop on Automated Cartography. DHEW Publication No. (PHS) 79-1254, 18-26. Included in The Collected Works of J.W. Tukey, Vol. 5 (1988), (Ed. W.S. Cleveland). Pacific Grove: Wadsworth.

TUKEY, J.W. (1990). Graphical diplays of: Are the (x,y) pairs compatible with a linear dependence? Technical Report No. 301, Princeton University.