Bayesian modelling with applications to sheep pox epidemic data
Malesios C.\textsuperscript{1}, Demiris N.\textsuperscript{2}, Ntzoufras I.\textsuperscript{2}, Abas Z.\textsuperscript{1} Kalogeropoulos K.\textsuperscript{3} and Koutroumanidis T.\textsuperscript{1}
\textsuperscript{1} Department of Agricultural Development, Democritus University of Thrace, Greece
\textsuperscript{2} Department of Statistics, Athens University of Economics and Business, Greece
\textsuperscript{3} Department of Statistics, London School of Economics, UK

Features of epidemic data
- Spatio-temporal nature.
- Environmental noise.
- Multicolinearity issues.
- Presence of excess zeros.

Motivating example
- Sheep pox is a highly contagious viral infection of sheep that can have devastating consequences.
- Major sheep pox epidemic in Evros Prefecture, northeastern Greece during 1994-98.
- Included 35,440 dead animals and 250 infected farms.

Bayesian regression modelling
- We adopt the Bayesian paradigm for modelling the weekly frequencies of disease occurrence incorporating various factors such as spatial information, environmental covariates and autoregressive (O-U type) processes.
- We ameliorate the existing methodologies by improving accuracy of predictions through the incorporation of spatial information, implementing suitable zero-inflated distributions, dealing with correlated data and introducing an epidemic interpretation via a branching process approximation.

Let \( Y_{ij} \) be the number of occurrences for the ith year and the jth week (i=0,1,2,3,4; j=0,1,...,52). Poisson and ZIP models are special cases of:

\[ Y_{ij} \sim g(y_{ij} | \theta_{ij}; p_{ij}) \]
\[ g(y_{ij} | \theta_{ij}; p_{ij}) = p_{ij}I_{[y_{ij}=0]} + (1-p_{ij})f(y_{ij} | \theta_{ij}) \]
\[ \theta_{ij} = h(\mu_{ij}) = \log(\mu_{ij}) \]
\[ \log(\mu_{ij}) = X_{ij} \theta + b_{ij} + U_{ij} + \theta_{x} * y_{ij-1} + K(\Theta, \kappa) \]

where \( \Theta = (\Theta_{\beta}, \Theta_{\kappa}) \) comprises of fixed-effects parameters \( \beta \), AR(1) parameter associated with number of cases in previous week (\( \theta_{\kappa} \)) and parameters of transmission kernel functions, \( b_{ij} \) yearly random effects and \( U_{ij} \) an O-U process adjusting for serial correlation.

Epidemic control
- We connect the Poisson component of our model to a suitable branching process.
- Using this representation we obtain the probability of epidemic extinction by solving \( \exp(g(\lambda)) = q \exp(\lambda) \)
- For ZIP model: \( \text{prob}(\text{extinction}) = 1 \land (q(\lambda_{ij}) + p_{ij}) \)
- We decompose the average rate of infection to its endemic and epidemic components:

\[ \lambda = \lambda_{\text{endemic}} \cdot \lambda_{\text{epidemic}} = \exp(\Theta_{\text{endemic}} + \Theta_{\text{epidemic}}) \]

Table 1: Estimated average extinction probabilities (q) for different covariate combinations

<table>
<thead>
<tr>
<th></th>
<th>min</th>
<th>average</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>distance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>min</td>
<td>q=0.78</td>
<td>q=0</td>
<td>q=1</td>
</tr>
<tr>
<td>median</td>
<td>q=0.66</td>
<td>q=0.66</td>
<td>q=0.66</td>
</tr>
<tr>
<td>max</td>
<td>q=0.41</td>
<td>q=1</td>
<td>q=0</td>
</tr>
</tbody>
</table>

Table 2: Endemic/epidemic decomposition of \( \lambda \)

<table>
<thead>
<tr>
<th></th>
<th>min</th>
<th>median</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_{\text{endemic}} )</td>
<td>0.17</td>
<td>2.84</td>
<td>121.6</td>
</tr>
<tr>
<td>( \lambda_{\text{epidemic}} )</td>
<td>1.27</td>
<td>1.005</td>
<td>3.19</td>
</tr>
</tbody>
</table>

Discussion
We have proposed a unified approach for effectively dealing with frequent problems of epidemic data.
Our modelling framework extends similar approaches (Choi et al., 2012; Branscum et al., 2008) in various ways:
- Incorporation of spatial information.
- Adopt g-prior-based approaches for coping with multicollinearity problems between covariates, especially those of environmental and/or meteorological data nature.
- Use of suitable distributions for modelling excess zeros datasets, particularly the zero-inflated Poisson (ZIP) distribution.
- Combine branching process theory with Bayesian regression models to identify epidemic potential.

Future directions
- Model selection has been solely based on standard (deviance-based) predictive approaches. The temporal nature of this type of data suggests a prequential approach.
- Explore continuous (DeWijs type) spatio-temporal models.