Use of space-time models to investigate the stability of patterns of disease

Juan Jose Abellan¹, Sylvia Richardson², Nicky Best²

¹ Small Area Health Statistics Unit (SAHSU)
² Department of Epidemiology and Public Health
Imperial College London

Correspondence author: Juan Jose Abellan
Small Area Health Statistics Unit
Department of Epidemiology & Public Health
Imperial College London

Mail address: St Mary’s campus
School of Medicine
Norfolk Place
London W2 1PG
Telephone: +44 (0)20 7594 3302
Fax: +44 (0)20 7402 2150
e-mail: j.abellan@imperial.ac.uk
Running title: Use of space-time models.

Key words: Bayesian hierarchical models, disease mapping, space-time interactions, congenital anomalies, mixtures models, stable disease patterns.

Acknowledgements: We thank Mireille Toledano and Catherine Keshishian for comments on the congenital anomalies data, and Noël Cressie and Alex Lewin for insightful discussions about the mixture model and classification rules.

This work has been partially funded by contract 200-2005-13328 from the US Centers for Disease Control and Prevention.

SAHSU is funded by a grant from the Department of Health, Department for Environment, Food and Rural Affairs, Environment Agency, Scottish Executive, Welsh Assembly Government and Northern Ireland Department of Health, Social Services and Public Safety.

Juanjo Abellan acknowledges partial support from grant MTM2004-03290 from the Spanish Ministry of Education and Science.

Abbreviations: • ICD: International Classification of Diseases.
• ICD9: International Classification of Diseases, 9th revision.
• ICD10: International Classification of Diseases, 10th revision.
• MCMC: Markov Chain Monte Carlo.
• ROC: Receiver-operating characteristic.
Contents
Abstract

In this paper we show how the inclusion of a time dimension in disease mapping models strengthens the epidemiological interpretation of the overall pattern of risk. We discuss a class of Bayesian hierarchical models in which the stable spatial and time patterns as well as departures from these stable components are characterised and estimated simultaneously. We show how useful rules for classifying areas as stable can be constructed based on the posterior distribution of the space-time interactions. We carry out a simulation study to investigate the sensitivity and specificity of the decision rules we propose and finally we illustrate our approach in a case study of congenital anomalies in England. Our results confirm that extending hierarchical disease mapping models to models that simultaneously consider space and time leads to a number of benefits in terms of interpretation and potential for detection of localized excesses.
1 Introduction

The fields of geographical epidemiology and public health surveillance have benefited from combined advances in the statistical modelling of spatial data and in geographical information systems. Exploring and characterising a variety of spatial patterns of diseases at a fine geographical resolution has become possible (Banerjee et al. 2004) and the use of hierarchical models estimated in a Bayesian framework to account for different levels of variability of such data is now well established (Richardson 2003). Inference on the relative risks of interest is usually obtained through the implementation of Bayesian computations which allow to output the posterior distribution of the relative risks in each area. Based on these posterior distributions, decision rules to detect areas of increased risks have been calibrated by Richardson et al. (2004). Insight into the sensitivity of the resulting inference to the choice of the structure of the different components of the hierarchical model has been gained through the use of simulation studies (Best et al. 2005) and numerous case studies.

Most studies consider data aggregated over a period of time and because of this cannot address important epidemiological questions about the stability of the estimated patterns of disease. Indeed, two quite different situations can give rise to the same accumulated number of cases in an area over a set time period: (i) the accumulation in any subinterval of time is roughly proportional to the length of the subinterval or (ii) there is substantial variability in the rate of accumulation over time. The epidemiological interpretation of these two situations is quite different. Patterns corresponding to (i) occur in a ‘repeatable manner’ over time and hence could be induced by environmental or socio-demographic risk factors that act in a stable way throughout the whole period. In contrast, situation (ii) exhibits substantial variability of the risk over time, pointing to potentially emerging short latency risk factors that would create a high excess of cases in a few short time intervals, or alternatively to artefactual variations possibly due to abrupt changes in recording practices in some areas. Hence the epidemiological interpretation of overall patterns of risk and in particular of high risk areas is considerably strengthened if the full space-time profile of the risks is uncovered. This leads naturally to the use of space-time models for analysing small area disease variability. In these models, predictable spatial and time patterns are characterised and specific departure from these predictable components are estimated simultaneously.

There are however a number of statistical issues that arise as a consequence of using space-time models. Possibly the most important of these is the increase of the sparseness in the counts. Even in pure spatial analyses where data are aggregated at a certain spatial level over a given time period, counts can be small if working with small areas and/or rare diseases. Hence the benefit of disaggregating over time and modelling in the time dimension with space-time interaction parameters need to be carefully calibrated. The aim of our study is to demonstrate how it is possible to exploit the rich output of space-time model fitting to gain interpretability without loosing power.

Throughout, the epidemiological context that we are considering is the analysis of the geographical and temporal variations of the risk of non-chromosomal congenital anomalies in England over a sixteen year period. These variations can potentially be associated both to socio-economic and environmental factors.
such as proximity to landfill sites (Elliott et al 2001) or to heterogeneity in recording practises (Boyd et al 2005) and there is great interest in understanding if any estimated spatial pattern is artefactual or occurs in a stable manner over time.

2 Material and Methods

2.1 Hierarchical space-time models for count data

The data consist of observed number of congenital anomalies \( Y_{it} \) and total number of births \( n_{it} \) for area \( i = 1, \ldots, I \) and year \( t = 1, \ldots, T \). Because the rate of congenital anomalies is approximately 1 per 1000 births, a binomial model for the counts is more appropriate than the usual Poisson model. Without loss of generality, the space-time models that we will formulate will thus be within the binomial framework with logistic link, extending these models to the Poisson case is straightforward.

In the Bayesian hierarchical framework that we consider, the binomial likelihood of the data is considered as the first level of the model, i.e. for modelling the within area variability of the counts conditional on unknown risk parameters. These parameters in turn are given prior distributions at a second level of the model where the space-time structure is specified. The aim of the second level model is to characterise space-time patterns that are predictable from the data over the whole time period and to uncover atypical departure from these patterns if any. Different sets of parameters are thus introduced to represent respectively predictable global space-time structures and specific space-time interactions.

At the first level, we define a binomial model for the within-area variability of the counts:

\[
Y_{it} \sim \text{Binomial}(n_{it}, \pi_{it})
\]  

where \( \pi_{it} \) is the risk of, say, congenital malformation in area \( i \) and year \( t \).

At the second level of the model, the risks \( \pi_{it} \) are split on the logit scale into an overall risk \( \alpha \), main spatial effects \( \lambda_i \), main temporal effects \( \xi_t \), and space-time interaction terms \( \nu_{it} \). All these effects are treated as random variables and given prior distributions which specify how information can be borrowed across space or time in order to better capture the underlying structure of the risks. The main spatial and temporal effects, \( \lambda_i \) and \( \xi_t \), are the ‘predictable’ parts and we extend the spatial model introduced by Besag et al. (1991) to model their distribution. The spatial dependence is represented by means of a prescribed neighbourhood graph that defines for each area \( i \) its set of neighbours (e.g. adjacent areas) denoted by \( \partial_i \). Associated with this graph is an adjacency \( I \times I \) matrix, say \( W = (w_{jk}) \), where \( w_{jk} = 1 \) if area \( j \) and \( k \) are considered adjacent, and 0 otherwise. Similarly, time neighbours are simply defined by the two adjacent time points, with associated time adjacency matrix \( Q \).

The most commonly used parametric model to express spatial dependence is the conditional autoregressive model (CAR). Given an adjacency matrix \( W \), this specifies the conditional distribution of a set of parameters \( \mu_i \), by:

\[
p(\mu_i | \mu_j, j \neq i) \sim N(\overline{\mu}_i, \sigma^2_{\mu_i}/k_i)
\]  

where \( \sigma^2_{\mu_i} \) is the variance of the spatial effect.

6
where $\sigma^2$ is an unknown variance parameter and $\pi_i = \sum_{j \in \partial i} \mu_j / k_i$, $k_i$ the number of neighbours of area $i$. Thus, the value of a parameter in one area is influenced by the average value of its neighbours, with additional variability quantified by a conditional variance $\sigma^2 / k_i$. We will use the notation,

$$\mathbf{\mu} \sim \text{CAR}(\mathbf{W}, \sigma^2_{\mu})$$

to denote the conditional autoregressive process specified in (??), where $\mathbf{\mu}$ is the vector $(\mu_1, \mu_2, \ldots, \mu_I)$.

This CAR model assumes a strong dependence and has only one free parameter linked to the conditional variance $\sigma^2_{\mu}$. To increase flexibility, it is recommended to use as spatial prior the sum of a CAR process and an unstructured exchangeable normal component with mean 0 and variance $\sigma^2_{\lambda}$. We will refer to this model as the convolution BYM model to acknowledge its introduction by Besag, Yorke and Mollić in 1991. It can be written in compact form as

$$\lambda_i \sim \text{Normal}(\mu_i, \sigma^2_{\lambda}), \ i = 1, \ldots, I$$

$$\mathbf{\mu} \sim \text{CAR}(\mathbf{W}, \sigma^2_{\mu}).$$

We have introduced the CAR and BYM model in the spatial context, but the definition is equally applicable to model temporal structure.

The second level of our model is thus given by:

$$\logit(\pi_{it}) = \alpha + \lambda_i + \xi_t + \nu_{it}, \ i = 1, \ldots, I; \ t = 1, \ldots, T$$

$$\lambda_i \sim \text{Normal}(\mu_i, \sigma^2_{\lambda}), \ i = 1, \ldots, I$$

$$\mathbf{\mu} \sim \text{CAR}(\mathbf{W}, \sigma^2_{\mu})$$

$$\xi_t \sim \text{Normal}(\gamma_t, \sigma^2_{\xi}), \ t = 1, \ldots, T$$

$$\gamma \sim \text{CAR}(\mathbf{Q}, \sigma^2_{\gamma}) \quad (3)$$

In (??), besides separable spatial and temporal BYM structures for the logit risk, we have introduced space-time interactions parameters $\{\nu_{it}, i = 1, \ldots, I; t = 1, \ldots, T\}$ which capture any departure from predictable patterns based on the overall time trend and the overall spatial risk surface. These space-time interaction parameters are thus key for characterising the stability of the underlying spatial patterns, large fluctuations of $(\nu_{it}, t = 1, \ldots, T)$ indicating instability of risk in area $i$. How to specify a prior distribution for the $(\nu_{it})$ which helps to distinguish stable predictable patterns from atypical ones will be discussed in detail in the next section.

As in any Bayesian analysis, a third level of the model is defined so that the variance parameters that are involved in the second level equations (??) are themselves treated as unknown and given (hyper)prior distributions. We chose inverse gamma with parameters 0.5 and 0.0005 following Wakefield et al (2000).

### 2.2 Characterising patterns of space-time interactions

Several considerations have to be taken into account when specifying a prior structure for the interactions $\{\nu_{it}, i = 1, \ldots, I; t = 1, \ldots, T\}$. For non-infectious health outcomes, one would expect that the overall space and time components capture adequately most of the structure and that substantial space-time interactions
are not common. Hence, some smoothing of the $\nu_{it}$ parameters is necessary to ensure that the model is not overparametrised and that noisy small space-time interaction parameters are shrunk towards zero. On the other hand, we also want to allow the possibility that for a few areas, there are ‘true’ departures from the overall stable model. This leads naturally to choose a mixture model for the distribution of the $\nu_{it}$ with two components: the first one models small $\nu_{it}$ parameters that reflect only residual noise and are not of epidemiological interest, whereas the second one captures ‘true’ departures from the space and time main effects. Mixture models are typically used in Bayesian analysis when heterogeneity is suspected as they give a flexible prior structure that can be used for classification (Richardson and Green, 1997). In our case, it is heterogeneity of the space-time interactions that we are concerned with. Specifically, we consider:

$$\nu_{it} \sim p \text{Normal}(0, \tau_1^2) + (1 - p) \text{Normal}(0, \tau_2^2).$$ \hspace{1cm} (4)

The prior for $p$ is uniform on $[0, 1]$ and half-normal hyperprior distributions for the standard deviations $\tau_k, k = 1, 2$ are specified to reflect that $\tau_1$ has to be small to effect shrinkage, whereas the prior for $\tau_2$ allows a large range of values for this parameter:

$$\tau_1 \sim \text{Normal}(0, 0.01) \cdot I_{(0, +\infty)}$$

$$\tau_2 \sim \text{Normal}(0, 100) \cdot I_{(0, +\infty)}$$ \hspace{1cm} (5)

where $I$ denotes the indicator function.

As usual in the Bayesian mixture model formulation, latent allocation variables $z_{it}$ are defined that take the value 0 or 1 if $\nu_{it}$ comes from the $\text{Normal}(0, \tau_1^2)$ or the $\text{Normal}(0, \tau_2^2)$ components respectively. These allocation variables are used to compute the posterior probabilities $p_{it}$ that each space-time interaction parameter $\nu_{it}$ comes from, say, the second component: $p_{it} \equiv P(z_{it} = 1 \mid \text{data})$. In turn these posterior probabilities are employed to classify the areas into two subsets $C_1$ and $C_2$ corresponding to ‘stable’ and ‘unstable’ risk patterns, respectively. Areas in $C_1$ would be those showing small non interpretable departures from main spatial and temporal effects. In contrast, areas classified in $C_2$ will be those with higher levels of temporal variability in the risk, not due to chance. In the following, we consider two decision rules for defining ‘stable’ and ‘unstable’ risk patterns:

**Rule 1:** an area $i$ is in $C_2$ (‘unstable’) if $p_{it} > p_{cut}$ for at least one $t, t = 1, \ldots, T$, where $p_{cut}$ is a threshold probability to be defined.

**Rule 2:** an area $i$ is in $C_2$ (‘unstable’) if the average of the three highest posterior probabilities $p_{it}$ is above a threshold $p_{cut}$.

Rule 1 is less restrictive than Rule 2 which looks for more evidence of variability by averaging over three time points. We have chosen three as corresponding to roughly 20% of the 16 time points in our study. Neither rules make any assumption about the specific shape of the pattern of the space-time interactions but a variety of rules akin to Rule 2 could be usefully constructed to look for variability in consecutive time
The two rules defined above will be calibrated in a simulation study and illustrated on our case study.

2.3 Congenital anomalies data

We consider the analysis of the geographical and temporal variation of risk of congenital anomalies in England between 1983 and 1998. We obtained data on live and still births from the births registry; we considered all non-chromosomal congenital anomalies combined (ICD9 740-759; ICD10 Q00-Q99) from the National Congenital Anomalies System. Both registers are postcoded (ca. 1.5 million postcodes in England, 15 households on average) and maintained by the Office for National Statistics. A copy of the data is held by the UK Small Area Health Statistics Unit (www.sahsu.org).

We divided England into a grid of around 5500 5x5 km squares based on the UK National Grid. We compiled all data in a Geographical Information System (GIS), based on a conformal projection (Universal Transverse Mercator), with a notional resolution of 1 metre. For each grid square we estimated the number of births and the number of congenital anomalies from the residential postcode locations. Since this led to very sparse data for some years and squares, it was decided to aggregate some of the squares, in order to have no more than a four to five fold interquartile range in the number of births per square over England. Thus the geography used in our analysis divides England into a grid of 970 variable size squares (Figure ??). The annual number of both congenital anomalies and births were aggregated for each square in the new grid. A descriptive summary of the data is shown in Table ??.

2.4 Data Generation

Besides the case study, the performance of our space-time model formulation and in particular the classification of areas into ‘stable’ and ‘unstable’ will be evaluated in a comprehensive set of simulations. In order to be realistic in terms of hypothesized spatial variability of risk and number of births, we will base our simulations on the congenital malformation set up but with a reduced set of 309 areas in order to ease the computational burden (see Figure ??).

To be precise, we use the posterior medians of the spatial and temporal main effects, $\lambda^*_i$ (range, on the odds-ratio scale, between 0.3 and 2.1) and $\xi^*_t$ (from 0.62 to 1.79 on the odds-ratio scale), derived from fitting model (??) to the congenital malformation data (as well as $\alpha^* = \log(0.01)$) to generate the predictable spatial and temporal patterns for the data replicas in our simulation study. For a subset of squares $\mathcal{M}$ we also add space-time interaction with different variances. Specifically, we define the set of simulated risks $\pi^*_it$ as:

$$
\pi^*_it = \expit(\alpha^* + \lambda^*_i + \xi^*_t + \nu^*_it)
$$

where $\expit(x) = \exp(x)/(1 + \exp(x))$, and
\[ \nu_{it}^* = \begin{cases} 
0 & \text{if } i \notin \mathcal{M} \\
\sim N(\log 2, \kappa^2) & \text{if } i \in \mathcal{M} 
\end{cases} \]

where \( \mathcal{M} \) is a subset of squares for which the risk is modified and perturbed by noise. We consider three different cases for the amount of noise, a reference case with no noise (\( \kappa^2 = 0 \)), one with a medium amount of noise (\( \kappa^2 = (0.5)^2 \)), and one with a high level of noise (\( \kappa^2 = (1.5)^2 \)), referred respectively to as medium and high variance cases. Values for \( \nu_{it}^*, i \in \mathcal{M} \) were generated just once. Profiles of the risk patterns (\( \pi_{it}^* \)) for each variance type case are displayed in Figure 2. We can see that between the reference case (green line) and the medium case (blue line), the risks vary about two-fold. We do not expect the variability generated for the high variance case (red line) to be realistic, rather we use this case as a benchmark.

Using the simulated risks \( \pi_{it}^* \), we generated 50 data replicas, one for each variance type scenario. This is done separately for each square, using a multinomial distribution, based on the total number of observed cases \( Y_i = \sum_{t=1}^{T} Y_{it} \) in each square. This ensures that the total number of cases in the replica was the same as in the original data. In other words, the \( r \)th data replica for area \( i \), \( Y_i^{(r)} = (Y_{i1}^{(r)}, \ldots, Y_{iT}^{(r)}) \), was generated from

\[ Y_i^{(r)} \sim \text{multinomial}(Y_i, \pi_i^*) \]

Note that as the aggregated number of cases in each square is the same for all replicas, a pure spatial model fitted to the time aggregated data will give exactly the same results regardless the variance type scenario.

Finally, besides the different variance cases for the variability of the space-time interactions, we also consider three different scenarios for the proportion of modified squares. Indeed the fitting of model (2) and in particular its ability to separate effectively the predictable part from the space-time interactions will be influenced by the overall number of modified (\( i.e. \) unstable) squares. Thus we consider the three cases of 20\%, 8\%, and 1\% of modified squares in \( \mathcal{M} \), which we denote hereafter as \( \mathcal{M}_{20}, \mathcal{M}_{8}, \mathcal{M}_{1} \), respectively. The exact number of squares were \( |\mathcal{M}_{20}| = 59, |\mathcal{M}_{8}| = 26 \) and \( |\mathcal{M}_{1}| = 3 \) (see Figure ?? in supplementary material which shows the squares selected for each scenario).

### 3 Results

#### 3.1 Model implementation and convergence issues

Bayesian inference is based on the joint distribution of all parameters (\( e.g. \lambda_i, \xi_t, \nu_{it} \)) given the data. In our case, this joint distribution is intractable analytically and is instead simulated using the framework of Markov chain Monte Carlo (MCMC) algorithms (Gilks et al 1996) that is now commonly used for Bayesian inference in a wide variety of applications. The free software WinBUGS (Spiegelhalter et al. 2003), based on MCMC algorithms, was used to implement all models (see Supplementary material for the Winbugs code).

For the simulation study, the results are based on a thinned (every 10th) sample of 2000 observations from the posterior distribution of the parameters, after discarding the first 10000 as burn-in. We used a longer run for the case study (30000 iterations after a 20000 burn-in).
It is well known that mixture models are difficult to estimate and that for example, multimodal likelihood and label switching problems can occur (Richardson and Green 1997). Here label switching issues are circumvented by the strong constraint on $\tau_1$. Whether there is enough information in a data set to estimate a two-component mixture for the variance of the interaction parameters is an issue. In any application of our model to real data, sizable residual variability is expected once the predictable part is fitted because the structure of the predictable part is deliberately kept simple. Hence, the space-time interaction parameters will absorb this residual variability and there will be enough heterogeneity to support the estimation of the two mixture components for their variance. In our reference simulation scenario however, the estimation of a 2-component mixture is artificial and substantial overlap between the two components is to be expected. This is confirmed by the large posterior variability of the mixture parameters in the reference case (see Table ??).

Convergence of chains was judged on visual checks by running two chains with different starting values. The ratio of the Monte Carlo error to the posterior standard deviation was roughly kept under 5% as is usually recommended. Figure ?? in the Supplementary material displays graphs of the posterior densities of the mixture parameters in the case study for two separate chains and illustrate how two chains that started at widely different starting points led to the same posterior distributions. Figure ?? shows the time series plots of the last 49,500 simulations for each of the parameters in the mixture component.

### 3.2 Estimation of spatial patterns

We first want to show that in cases where the spatial structure is stable, the additional complexity of the space-time analysis does not perturb the estimation of the stable spatial patterns. Indeed in cases where there is a stable pattern, the introduction of space-time parameters over-parametrises the model with unnecessary parameters and hence could lead to a loss of precision. We estimate a pure spatial model on the time aggregated counts $Y_i$, where (??) is replaced by

$$
\begin{align*}
\logit(\pi_i) &= \alpha + \lambda_i \\
\lambda_i &\sim \text{Normal}(\mu_i, \sigma_\lambda^2) \\
\mu &\sim \text{CAR}(W, \sigma_\mu^2)
\end{align*}
$$

(7)

For this comparison, we have used the reference case of the $M_{20}$ scenario described in Section ???. The comparison of the estimated spatial risks between the spatial only model (??) and the spatio-temporal model (??) reveals that the estimated spatial risks $\lambda_i$ are almost identical between the pure spatial model and the space-time model despite the increase in the sparseness and the inclusion of many more parameters in the space-time analysis (see Figure ??).

Besides the estimation of the $\lambda_i$, it is also of interest to look at the posterior probabilities $Pr(\exp(\lambda_i) > 1 \mid \text{data})$ as it has been shown in a previous study (Richardson et al, 2004) that these can be used to pinpoint areas of elevated risks. In particular, a rule that thresholds these posterior probabilities above 80% was suggested as giving a good compromise between sensitivity and specificity for detecting areas of increased
risks. The comparison of the posterior probabilities estimated with the pure spatial versus the space-time model shows little discrepancy (Figure ??) and hence no loss of power when using a space-time model instead of a pure spatial model. Table ?? numerically confirms this; out of the 147 areas with true spatial excess risk of at least 10% (including both modified and unmodified areas), 110 areas would be detected by such a rule in the pure spatial case (same number in all replicas by construction), and 108 on average (over all replicas) when using the space-time model. Therefore, the sensitivity of this rule is around 74%, which is adequate and similar for both models.

3.3 Estimation of the space-time interactions

As a first check on the performance of our space-time model with mixture prior on the interactions, we computed the empirical standard deviation of the $\nu_{it}$: $SD(\nu_{it}) = \left(\frac{1}{15} \sum_{t=1}^{16} (\nu_{it} - \bar{\nu}_i)^2\right)^{0.5}$ where $\bar{\nu}_i$ is the average of the $\nu_{it}$ over the 16 time points. Boxplots of the $SD(\nu_{it})$ are shown in Figure ?? for the 3 variance cases (reference, medium and high) and the different scenarios $M_{20}, M_{8}$, and $M_{1}$.

The reference case where all patterns are stable shows as expected that the $\nu_{it}$ are small and identical between the modified and unmodified areas in all scenarios. Comparison between the medium variance boxplots (blue) and the high variance boxplots (red) for the modified areas shows that the model captures well the increase of variability of the space-time interactions, and this irrespectively of the number of modified areas. The comparison between the 20%, 8% and 1% scenarios shows that the increasing number of modified areas has some influence on the overall fit of the model and hence the space-time patterns in the unmodified areas. Indeed in the 1% scenario, the $SD(\nu_{it})$ of the unmodified areas are small and fairly similar for the reference, medium and high variance cases. As the number of modified areas increases, $SD(\nu_{it})$ for the unmodified areas also increase somewhat, specially for the high variance case. This is to be expected since the large fraction (20%) of unstable areas in this scenario renders the ‘stable’ pattern less typical and blurs the distinction between stable and unstable cases.

We have also displayed in Figure ?? (bottom right) the equivalent boxplots that would be obtained in the 20% case if instead of a mixture model (??) for the space-time interactions, we had simply assumed that these interactions came from an exchangeable distribution with common variance, $\nu_{it} \sim \text{Normal}(0, \tau^2)$. From this plot, we can clearly see that an inappropriate exchangeable assumption leads to a substantial blur between the estimated variability of the interaction parameters of modified and unmodified areas, and that it would be difficult to interpret any patterns of the variability of the interactions as corresponding to ‘true variability’. Hence the specification of the mixture model (??) is key for the good performance of our approach.

Table ?? summarizes the posterior estimates of the mixture parameters for all three scenarios and variance cases. The parameter $p$ estimates the proportion of space-time ‘pixels’ showing departure from main effects. Because of the design of our simulation set-up, we would roughly expect $p$ to reflect the proportion of modified areas, i.e. 20%, 8%, and 1% respectively. In both the medium and high variance cases, this proportion is reasonably well estimated, with relatively narrow 95% credibility intervals. As expected, the values of $\tau_1$ are

12
quite small in all variance cases and scenarios, since the predictable part is the true model for the unmodified areas. Values of $\tau_2$ are well estimated in the high variance case, the true value 1.5 being within the average 95% credibility interval. However, in the medium variance case, $\tau_2$ is a little overestimated (estimated value around 0.7 whilst the true value is 0.5). This could be due to the prior on $\tau_1$ being a little too restrictive and thus biasing some interactions towards the second component.

In the reference case where there is no true variability over time, the mixture artificially separates the $\nu_{it}$ into 2 groups, which would be misleading to interpret. However the difference between $\tau_1$ and $\tau_2$ is quite small, thus indicating that the second component quantifies only a modest increased variability of the risks over time. Moreover, the credibility intervals for $\tau_2$ and $p$ are large, pointing to lack of identifiability as could be expected in this case (see Section ??). Other pointers that would suggest that, in this case, the second component is indeed not capturing real evidence of variability of the risks are discussed in the next section.

3.4 Performance of the classification rules

Since we are in a simulation set-up with ‘truly’ stable and unstable areas, it is possible to assess the sensitivity and specificity of the Rules 1 and 2 for classifying areas as ‘unstable’. Figure ?? shows average ROC curves (sensitivity versus 1-specificity for different cut-off values $p_{cut}$) for each variance case (medium in blue and high in red) for the $\mathcal{M}_8$ and $\mathcal{M}_{20}$ scenarios corresponding to Rule 1 (top) and Rule 2 (bottom). As expected the sensitivity is higher in both rules for the easier high variance case (red line). For Rule 1, one can see that for a specificity of 10%, the power is around 70% in both cases which makes this rule useful. In order to better interpret the ROC curves, complementary plots of the observed fraction of false positives and of false negative versus cut-off values are given in Figure ???. In the realistic medium variance case, the lines cross around $p_{cut} = 0.2$ for the $\mathcal{M}_8$ scenario, and $p_{cut} = 0.4$ for the $\mathcal{M}_{20}$. These cut-off values offer a compromise between the two types of errors, that are balanced at around 20%. In our context however, the two types of errors are not symmetric and it is typically more important to achieve high specificity. If we wish to control the false positive rate at around 10%, say, then we should consider higher values for the cut-off, a value of $p_{cut} = 0.5$ guarantees good specificity in each scenario, with around 20% false negative rates.

Performance of Rule 2 is very similar to that of Rule 1 (Figures ?? and ?? – bottom). The false positive and false negative curves cross at $p_{cut} = 0.2$ for both $\mathcal{M}_{20}$ and $\mathcal{M}_8$ scenarios in the medium variance case. By construction Rule 2 is more conservative than Rule 1 and will require somewhat lower thresholds for the same specificity.

Once the areas are classified, interpretation of the variability can be gained by looking at the time profiles of the estimated interactions. Figure ?? (bottom row) shows a sample of profiles of the space-time interactions estimated for the areas classified as ‘unstable’ according to Rule 1 and a cut-off value $p_{cut} = 0.5$. By comparing to the simulated pattern (top left of Figure ??), it is clear that the simulated pattern is well recovered. Note that in general, one would not expect, as in our simulated case, the same time profile of variability for all areas. As this does not influence the classification results (the $\nu_{it}$ are modelled independently over $i$), we chose to do this in our simulation set-up in order to be able to display the results
better. In our case study, we show how exploratory clustering of the time patterns for the areas declared as ‘unstable’ is necessary to confer additional interpretability.

On Figure ??, we have also shown a sample of time profiles for the reference case. Using Rule 1 and a cut-off of 0.5, on average over the 50 replicas no areas would be classified as ‘unstable’. Looking at the profiles of the 10 areas with highest space-time variability among the replica with highest values of $SD(\nu_{it})$, we can see even in these ‘extreme’ realizations the variability is small, especially compared to the medium and high variance cases. Notice the different scale of the y-axis in the plot of the reference case as opposed to the other two plots. In fact, if plotted in the same scale, the space-time profiles are hardly distinguishable from the stable profile of the medium and high variance cases.

Finally, we investigated the combination of Rule 1 with the rule $Pr(\exp(\lambda_i) > 1 | data)$. Out of the 250 unmodified areas in the $M_{20}$ scenario, 91 had $\lambda^*_i > 1.1$; of those 84 and 85 had $Pr(\exp(\lambda_i) > 1) > 0.8$ in the medium and high variance cases, respectively, and all of them were classified as stable using Rule 1. Analogously, out of the 283 unmodified squares in $M_8$, 98 had $\lambda^*_i > 1.1$; of those 94 and 93 had $Pr(\exp(\lambda_i) > 1) > 0.8$ in the medium and high variance cases, respectively, and again all of them were considered as stable by Rule 1. This shows that our modelling approach combined with appropriate classification rules can effectively separate out the areas where elevated risk is occurring in a repeatable way.

4 Case study: Risk of non-chromosomal congenital anomalies in England

We have estimated the hierarchical space-time model (equations (??) and (??)) on the congenital malformation data. Spatial and temporal patterns are shown in Figure ??). Note that there is a clear drop in 1990 and subsequent decrease in 1991 and 1992, after which the time trend stabilizes. In 1990, some minor anomalies were reclassified and excluded from the combined congenital anomalies. See OPCS (1995, page 12), for details on the ICD codes excluded.

There is evidence of spatial heterogeneity of the relative risks with higher values estimated for regions of the North, North East and North West, some central areas, and in the South West (Figure ??). In the densely populated Greater London area, it is difficult to see the pattern because the grid squares are small but there are also higher risks in the London area and some parts of the South and South East. Maternal age and deprivation are known risk factors for some non-chromosomal congenital anomalies. It is likely that they contribute to the geographical pattern of the overall risk that is shown in Figure ??.

Posterior medians and 95% credibility intervals for the mixture parameters are shown in Table ?? (bottom row). In contrast to the simulation results, the mean posterior standard deviation of the first component, $\tau_1$ is larger. This is unsurprising. In our simulation set-up, the risk in the unmodified areas was totally predictable from the separable space and time components. As a consequence the first component was just capturing noise created by the over-parametrisation. In contrast, for any real data set, some departure from the predictable pattern will be observed for almost all areas, the extent of which is then modeled by the
mixture model. The first mixture component will thus summarize a small lack of fit of $\pi_{it}$ from the predicted risk built by $\lambda_i$ and $\xi_t$, that is not noteworthy of further investigation. There was no indication of lack of convergence in the visual checks (see Supplementary Figure ??) and the credibility intervals for the mixture parameters were narrow.

Using Rule 1 and a cut-off of 0.5, 125 of the 970 (13%) variable size grid squares were classified as exhibiting some instability. These squares are highlighted on the map in Figure ?? (squares with borders in blue). They do not appear to be specially linked to a particular region. The proportion of squares classified as ‘unstable’ is close to the estimation of $(1-p)$, which would be 16% and represents the number of space-time ‘pixels’ in the second component of the mixture.

To interpret the variability shown in the squares classified as ‘unstable’, it is useful to attempt to cluster the time profiles of the interactions. Using simple hierarchical clustering (Mardia et al 1979) on these time profiles, we found 5 main profile patterns that are displayed in Figure ?? (left). Four subgroups exhibit smooth-like trends (increasing or decreasing) over time, indicating that the interactions terms are used to ‘adjust’ these areas to the general time trend. Indeed the implementation of the recommended reclassification of some minor anomalies did not happen at the same speed over the UK and thus shift from the overall time trends are to be expected. Besides these 4 groups, one small subgroup (5 squares) shows a sudden high peak in the year 1997 (blue line). This unexpected and large time variation in the risk could correspond to an abrupt change of local recording practise or to a real excess of cases, either situations warranting further investigation.

Finally, we found a close correspondence between the areas classified as ‘unstable’ by Rule 1 and by Rule 2. The top 125 areas that would be considered ‘unstable’ by Rule 2 (ranked by their score for Rule 2) would contain 104 of the areas similarly classified by Rule 1. This leads of course to similar clusters of time profiles, Figure ?? (right).

5 Discussion

The use of Bayesian hierarchical spatial models has become widespread in disease mapping and ecological studies of health-environment associations. In most of the studies carried out, the data are typically aggregated over an extensive time period, typically more than a decade. Extension of hierarchical spatial models to space-time modelling of one or several diseases has been discussed by a number of authors (Bernardinelli et al. 1995, Waller et al. 1997, Knorr-Held and Besag 1998, Knorr-Held 2000, Richardson et al. 2006). The purpose of these authors was to propose and investigate the fit of a variety of space-time model formulations. In this paper, we build on these extensions but our purpose is different. We aim to distinguish in a statistically informed way areas where the risk is ‘predictable’ by a simple combination of overall spatial pattern and time trends – from ‘atypical’ areas. Our purpose is dual: to strengthen the interpretation of the geographical patterns of risk that occur in a ‘repeatable’ way and to pinpoint ‘atypical’ or unstable areas showing evidence of unusual variability in the time pattern of the risk.
The epidemiological interpretation of ‘repeatable’ versus ‘atypical’ patterns has to be done with respect to the health outcome investigated. In the case of a disease where short latency effects are plausible, like congenital malformations, unusual variability, in particular excess of risk, is important to investigate further as it may point to the emergence of an environmental hazard. On the other hand, stable spatial patterns are more likely to be linked to recurrent socio-demographic and life style risk factors. For other health outcomes, ‘atypical’ risk patterns over time may point towards a local change in recording practice or health care. For example, a sudden increase or decrease in the risk of ‘avoidable deaths’ in some areas might give an early warning that health care is deteriorating or improving in those areas. Hence, routine evaluation of space-time patterns could be built into a surveillance system.

Due to the complex dependence patterns over space and over time of the occurrence of many chronic health outcomes, and the inherent large stochastic variability due to rare events, estimating separately time trends in each area will not be efficient as it will be difficult to establish a baseline pattern separately for each area. In our Bayesian approach, we use the power of hierarchical modelling to borrow information over space and time in order to estimate typical predictable patterns for each area. We further strengthen the inference by using a joint mixture model for the space-time interactions which again borrows information across all the time points and areas to improve inference, while at the same time explicitly recognizing their heterogeneity. Prior knowledge is used to specify meaningful priors for the mixture model. In particular for the sake of identifiability, it is important to make sensible prior assumptions for the variance of the component of the mixture that captures the idiosyncratic space-time interactions (in our formulation the first component) so that the mixture achieves a meaningful separation between small non-interpretable and ‘truly’ large space-time interactions.

The space-time hierarchical model that we have formulated is easily implemented by freely available WinBUGS software and thus provides a useful tool for health-environment investigations and health-practice surveillance. We tested the performance of this novel formulation in several simulation scenarios and demonstrated that by post-processing the rich output of the Bayesian space-time model, we can build classification rules that have good operational characteristics to detect ‘atypical’ areas. Our simulations also showed that with respect to estimating the overall spatial pattern, using a more sophisticated model that includes a time element and space-time interactions does not damage the estimation of the pure spatial pattern. Rather, it helps to characterize spatial excess risks that are stable over time. Our simulations were performed within the framework of binomial variability, with expected median number of events per area per year between 3 and 10. Extension of our simulation set-up to the Poisson case and to smaller number of expected events per area would be interesting and will be considered in the future.

Clearly, once a reasonable classification of areas is achieved, it is of interest to explore further the time pattern of the risks in ‘atypical’ areas. In our simulation set-up, we deliberately chose not to introduce any specific time patterns, for example step changes that could signal, for example, the occurrence of a new environmental hazard. If detection of such situations were of particular epidemiological interest, modified rules to declare areas ‘atypical’ could easily be defined and calibrated. Moreover, instead of using rules based
on the $p_{it}$, which are marginal posterior probabilities for each $i, t$, we could investigate how to exploit the joint distribution of the allocations $z_i = (z_{i1}, \ldots, z_{iT})^t$.

In our analysis of congenital malformations in England, we employed standard hierarchical clustering of the time profiles of ‘atypical’ squares to uncover interesting patterns. How to improve classification rules as well as tailoring statistical models to cluster the time profiles efficiently is an interesting extension of our approach that we plan to report on in future work.

In conclusion, we introduced a novel class of space-time models and demonstrated how epidemiological interpretation of risk patterns is considerably strengthened by the inclusion of the time dimension.
List of Tables
<table>
<thead>
<tr>
<th>Year</th>
<th>5th Centile</th>
<th>1st Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quartile</th>
<th>95th Centile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Congenital anomalies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>13.67</td>
<td>18</td>
<td>43</td>
</tr>
<tr>
<td>1984</td>
<td>0</td>
<td>4</td>
<td>9</td>
<td>13.67</td>
<td>18</td>
<td>40</td>
</tr>
<tr>
<td>1985</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>12.99</td>
<td>18</td>
<td>41</td>
</tr>
<tr>
<td>1986</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>12.85</td>
<td>17</td>
<td>40.6</td>
</tr>
<tr>
<td>1987</td>
<td>0</td>
<td>4</td>
<td>9</td>
<td>13.39</td>
<td>18</td>
<td>41.6</td>
</tr>
<tr>
<td>1988</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>12.75</td>
<td>17</td>
<td>41</td>
</tr>
<tr>
<td>1989</td>
<td>0</td>
<td>3</td>
<td>8</td>
<td>12.21</td>
<td>16</td>
<td>39</td>
</tr>
<tr>
<td>1990</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>8.007</td>
<td>11</td>
<td>25.6</td>
</tr>
<tr>
<td>1991</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>6.916</td>
<td>10</td>
<td>21</td>
</tr>
<tr>
<td>1992</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>5.994</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>1993</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5.638</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>1994</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5.52</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>1995</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5.451</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>1996</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5.485</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>1997</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5.305</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>1998</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>5.249</td>
<td>7</td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Births and Still births</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1983</td>
<td>43</td>
<td>220</td>
<td>503</td>
<td>615.6</td>
<td>896</td>
<td>1420.2</td>
</tr>
<tr>
<td>1984</td>
<td>43</td>
<td>222</td>
<td>503</td>
<td>623.2</td>
<td>906</td>
<td>1435.6</td>
</tr>
<tr>
<td>1985</td>
<td>46</td>
<td>233</td>
<td>522</td>
<td>642.7</td>
<td>929</td>
<td>1493.6</td>
</tr>
<tr>
<td>1986</td>
<td>44</td>
<td>233</td>
<td>528</td>
<td>647.5</td>
<td>945</td>
<td>1504</td>
</tr>
<tr>
<td>1987</td>
<td>46.8</td>
<td>245</td>
<td>542</td>
<td>667.3</td>
<td>979</td>
<td>1522.6</td>
</tr>
<tr>
<td>1988</td>
<td>46</td>
<td>253</td>
<td>560</td>
<td>678.4</td>
<td>1005</td>
<td>1559.4</td>
</tr>
<tr>
<td>1989</td>
<td>49</td>
<td>251</td>
<td>552</td>
<td>673.5</td>
<td>981</td>
<td>1480.6</td>
</tr>
<tr>
<td>1990</td>
<td>46</td>
<td>252</td>
<td>562</td>
<td>691.2</td>
<td>1018</td>
<td>1558.2</td>
</tr>
<tr>
<td>1991</td>
<td>46</td>
<td>253</td>
<td>558</td>
<td>684.9</td>
<td>1001</td>
<td>1528.4</td>
</tr>
<tr>
<td>1992</td>
<td>45</td>
<td>252</td>
<td>553</td>
<td>678.4</td>
<td>984</td>
<td>1537.4</td>
</tr>
<tr>
<td>1993</td>
<td>47</td>
<td>243</td>
<td>544</td>
<td>660.3</td>
<td>962</td>
<td>1492.8</td>
</tr>
<tr>
<td>1994</td>
<td>46</td>
<td>241</td>
<td>537</td>
<td>654.1</td>
<td>945</td>
<td>1475</td>
</tr>
<tr>
<td>1995</td>
<td>43</td>
<td>237</td>
<td>530</td>
<td>638</td>
<td>915</td>
<td>1444.8</td>
</tr>
<tr>
<td>1996</td>
<td>45.8</td>
<td>241</td>
<td>528</td>
<td>639.2</td>
<td>906</td>
<td>1445.2</td>
</tr>
<tr>
<td>1997</td>
<td>46</td>
<td>236</td>
<td>522</td>
<td>632.7</td>
<td>891</td>
<td>1446.4</td>
</tr>
<tr>
<td>1998</td>
<td>45.4</td>
<td>237</td>
<td>524</td>
<td>626.5</td>
<td>886</td>
<td>1452.8</td>
</tr>
</tbody>
</table>

Table 1: Descriptive summary of the case study data: distribution of cases and denominator counts across areas
\( \Pr(\exp(\lambda_i) > 1 \mid \text{data}) > 0.8 \)

<table>
<thead>
<tr>
<th>True spatial risk</th>
<th>( \Pr(\exp(\lambda_i) &gt; 1 \mid \text{data}) &gt; 0.8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spatial</td>
</tr>
<tr>
<td>&gt; 1.1</td>
<td>37</td>
</tr>
</tbody>
</table>

Table 2: Areas with true spatial excess risk above 10% that are declared as having an increased risk using the rule \( \Pr(\exp(\lambda_i) > 1 \mid \text{data}) > 0.8 \) with the spatial and spatiotemporal models.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>( \tau_1 )</th>
<th>( \tau_2 )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( M_{20} ) Reference</td>
<td>0.04 (0.02, 0.06)</td>
<td>0.16 (0.05, 3.50)</td>
<td>0.75 (0.36, 0.89)</td>
</tr>
<tr>
<td>Medium</td>
<td>0.05 (0.04, 0.07)</td>
<td>0.72 (0.63, 0.82)</td>
<td>0.89 (0.85, 0.91)</td>
</tr>
<tr>
<td>High</td>
<td>0.04 (0.02, 0.06)</td>
<td>1.39 (1.27, 1.52)</td>
<td>0.81 (0.79, 0.83)</td>
</tr>
<tr>
<td>( M_8 ) Reference</td>
<td>0.04 (0.02, 0.06)</td>
<td>0.10 (0.05, 2.76)</td>
<td>0.76 (0.36, 0.91)</td>
</tr>
<tr>
<td>Medium</td>
<td>0.04 (0.03, 0.07)</td>
<td>0.75 (0.64, 0.91)</td>
<td>0.94 (0.92, 0.96)</td>
</tr>
<tr>
<td>High</td>
<td>0.04 (0.02, 0.06)</td>
<td>1.43 (1.27, 1.63)</td>
<td>0.92 (0.90, 0.93)</td>
</tr>
<tr>
<td>( M_1 ) Reference</td>
<td>0.03 (0.02, 0.06)</td>
<td>0.10 (0.05, 2.15)</td>
<td>0.73 (0.35, 0.89)</td>
</tr>
<tr>
<td>Medium</td>
<td>0.04 (0.03, 0.06)</td>
<td>0.70 (0.46, 1.30)</td>
<td>0.99 (0.97, 1.00)</td>
</tr>
<tr>
<td>High</td>
<td>0.05 (0.03, 0.06)</td>
<td>1.32 (0.96, 1.95)</td>
<td>0.99 (0.98, 0.99)</td>
</tr>
<tr>
<td>Case Study</td>
<td>0.30 (0.27, 0.32)</td>
<td>0.59 (0.54, 0.68)</td>
<td>0.84 (0.76, 0.91)</td>
</tr>
</tbody>
</table>

Table 3: Posterior medians and 95% credibility intervals of the standard deviations of the two mixture components and the mixing proportion (all averaged over the 50 replicas), obtained under each scenario.
List of Figures
Figure 1: Grid of variable size squares covering England (shaded in lilac), and squares considered for the simulation study (shaded in green).
Figure 2: Top left: risk profiles of the unmodified squares compared to those modified under each variance type case, represented on the probability scale. The other 3 plots are represented are odds ratio on the logarithmic scale. Top right: 10 most ‘atypical’ estimated profiles in the reference case (blown up scale). Bottom row: estimated risk profiles for 10 randomly selected areas classified as ‘atypical’ using rule 1 in the medium (left) and high (right) variance cases for the $M_8$ scenario.
Figure 3: Scatterplots of the posterior medians of the main spatial effects and associated posterior probabilities obtained from the pure spatial model versus those in the spatiotemporal one in the reference variance case of the $M_{20}$ scenario.
Figure 4: Boxplots of the posterior median of the empirical standard deviations of the space time interactions when the $\nu_{it}$ are modelled as a mixture of two normals: $\mathcal{M}_{20}$ (top left), $\mathcal{M}_{8}$ (top right), and $\mathcal{M}_{1}$ (bottom left) scenarios. The bottom right boxplot corresponds to the empirical standard deviations using a exchangeable normal model for the space-time interactions.
Figure 5: ROC curves associated with Rules 1 (top row) and 2 (bottom row) for the $M_{20}$ (1st column) and $M_{8}$ (2nd column) scenarios.
Figure 6: Curves of false negative and false positive rates associated with Rules 1 (top row) and 2 (bottom row) for the $M_{20}$ (1st column) and $M_{8}$ (2nd column) scenarios.
Figure 7: Top: map of estimated spatial risks for the case study. Overimposed with blue borders are the areas classified as 'atypical' using Rule 1. Bottom: estimated main time trend in the case study.
Figure 8: Estimated risk profiles of the clusters of ‘atypical’ areas according to Rules 1 (left) and 2 (right) in the case study.
Appendix

Supplementary figures
Figure 1: Subset of 309 grid squares (gray) and squares selected (red) to ‘bump’, in the three different scenarios: $M_{20}$ (top), $M_8$ (bottom left), and $M_1$ (bottom right).
Figure 2: Posterior distribution of the standard deviations of the mixture component. 2 chains with overdispersed initial values.
Figure 3: Time series plots of the simulated 50,000 values (excluding the first 500) for the mixture components in the case study. 2 chains with overdispersed initial values.
WinBUGS model

This is the model with two mixture components for the variances of the space-time interactions terms as implemented in WinBUGS.

model { 

# The model

for (i in 1:N) {
    for (j in 1:T) {
        cm[i,j] ~ dbin(pi[i,j], births[i,j])
        logit(pi[i,j]) <- inter+lambda[i]+xi[j]+nu[i,j]
        OR[i,j] <- exp(lambda[i]+xi[j]+nu[i,j])
    }
}

for (i in 1:N){
    ORlambda[i] <- exp(lambda[i])
    prob.lambda[i] <- step(lambda[i])
}

for (j in 1:T){
    ORxi[j] <- exp(xi[j])
    prob.xi[j] <- step(xi[j])
}

# Prior distributions:

# - Intercept:

alpha ~ dnorm(0, 0.001)
ORalpha <- exp(inter)

# - Space:
for (i in 1:N){
    lambda[i] ~ dnorm(mu[i], tau.lambda)
}
mu[1:N] ~ car.normal(adj[], weights[], num[], tau.mu)

# - Time:

for(j in 1:T){
    xi[j] ~ dnorm(gamma[j], tau.xi)
}
gamma[1:T] ~ car.normal(adj.t[], weights.t[], num.t[], tau.gamma)

# - Interaction:

for(i in 1:N){
    for(j in 1:T){
        nu[i,j] ~ dnorm(0, tau.nu[ind[i,j]])
        ORnu[i,j] <- exp(nu[i,j])
        prob.nu[i,j] <- step(nu[i,j])
    }
}

# Calculating empirical variance in time

for(i in 1:N){ sd.nu[i] <- sd(nu[i, 1:T]) }

# Hyperpriors:

tau.lambda ~ dgamma(0.5, 0.0005)
tau.xi ~ dgamma(0.5, 0.0005)
tau.mu ~ dgamma(0.5, 0.0005)
tau.gamma ~ dgamma(0.5, 0.0005)

sigma.lambda <- 1/sqrt(tau.lambda)
sigma.xi <- 1/sqrt(tau.xi)
sigma.mu <- 1/sqrt(tau.mu)
sigma.gamma <- 1/sqrt(tau.gamma)
for(i in 1:N){
    for(j in 1:T){
        ind[i,j] ~ dcat(P[])
    }
}

sigma.nu[1] ~ dnorm(0, 100)I(0.0,)
kappa ~ dnorm(0, 0.01)I(0.0,)

P[1:2] ~ ddirch(alpha[])

for(i in 1:2){
    tau.nu[i] <- pow(sigma.nu[i], -2)
    alpha[i] <- 1
    alpha[2] <- 1
}

# Weights for adjacency matrices in space and time, respectively
for(k in 1:2210){
    weights[k] <- 1
}
for(k in 1:30){
    weights.t[k] <- 1
}
}