



HAL
open science

A new centered spatio-temporal autologistic regression model with an application to local spread of plant diseases

Anne Gégout-Petit, Lucia Guérin-Dubrana, Shuxian Li

► **To cite this version:**

Anne Gégout-Petit, Lucia Guérin-Dubrana, Shuxian Li. A new centered spatio-temporal autologistic regression model with an application to local spread of plant diseases. *Spatial Statistics*, 2019, 31, pp.100361. 10.1016/j.spasta.2019.100361 . hal-01926115

HAL Id: hal-01926115

<https://inria.hal.science/hal-01926115v1>

Submitted on 25 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 A new centered spatio-temporal autologistic regression
2 model with an application to local spread of plant
3 diseases

4 Anne Gégout-Petit^a, Lucia Guérin-Dubrana^{b,c}, Shuxian Li^b

5 ^a*Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France,*
6 *anne.gegout-petit@univ-lorraine.fr*

7 ^b*Université de Bordeaux, ISVV, UMR-1065 INRA, France*

8 ^c*Bordeaux Sciences Agro, Gradignan, FRANCE*

9 **Abstract**

We propose a new spatio-temporal autologistic centered model for binary data on a lattice. Centering allows the self-regression coefficients to be interpreted by separating the large-scale structure from the small-scale structure. One of the coefficients determines the overall level (or average) of the process, the second determines the spatial autocorrelation. We discuss the existence of the joint distribution of the process and carry out numerical studies to highlight the interest of this type of centering. We suggest using the estimator that maximises the Pseudo Likelihood (denoted Maximum Pseudo Likelihood Estimator (MPLE) in the following) and we give a method for choosing the neighbourhood structure. We run simulations studies that show that the estimation method and model selection method work well. The method is applied to model and fit epidemiological data on Esca disease in a vineyard in the Bordeaux region.

10 *Keywords:* spatial-temporal modeling, large-scale model structure, binary
11 response, Maximum Pseudo-Likelihood Estimation, Autologistic Model

12 **1. Introduction**

13 Since spatial and spatio-temporal data are commonly present in nature,
14 the modelling of such data has increased the interest of many scientists from
15 various fields such as ecology, epidemiology and image analysis. Binary data
16 is of particular interest for modelling the occurrence of an event such as dis-
17 ease or death. Spatio-temporal and spatial binary data models are quite
18 adequate to study the evolution of a known plant disease on a grid if propa-
19 gation and interaction between neighbours is suspected.

20 40 years ago, Besag (1974) firstly presented an autologistic model for spa-
21 tial binary data, assuming a simple dependence on surrounding neighbours.
22 This model has been shown to be useful and then was extended by Gumpertz
23 et al. (1997); Huffer and Wu (1998) to take into account the effects of some
24 covariates.

25 More recently, Zhu et al. (2008) and Zheng and Zhu (2008) generalised au-
26 tologistic regression models to simultaneously account for covariates, spatial
27 covariates, and time dependence for binary data that are measured repeat-
28 edly over time on a grid.

29 However it is difficult to interpret the parameters in a non-centered autol-
30 ogistic regression model. This problem has been first pointed out by Caragea
31 and Kaiser (2009) who presented a centered parameterisation for an autolo-
32 gistic spatial regression model to overcome interpretation problems. Follow-
33 ing this work, Hughes et al. (2011) discussed in more detail the estimation of
34 parameters and random field simulations from an centered autologistic model
35 on a lattice. Then, Wang and Zheng (2013) presented a centered spatio-
36 temporal autologistic regression model. They used and compared several

37 methods for estimating model parameters and coefficients. The drawback
38 of this spatio-temporal model is that the temporal dependency is not causal
39 and the state of a point at time t is linked with its state at time $t - 1$ and
40 $t + 1$. This model has good mathematical properties but is not useful for
41 interpretation by the practitioner. A review of the literature on autologistic
42 models applied to binary data was recently given by Zhu and Zheng (2016).

43 In this paper, we present a new centered spatio-temporal autologistic
44 model which depends only on the past. We will show, on simulated data, the
45 advantage of this kind of centering over other autologistic models; indeed,
46 we will see that, unlike the other models, it gives the expected average of
47 the process. Since the joint distribution of such models is very complex to
48 write, we recommend an estimation method based on the Maximisation of
49 the Pseudo-Likelihood (MPL). We present a method for choosing the neigh-
50 bourhood structure by using the value of the Pseudo-Likelihood. We apply
51 the model for a better understanding of the spread of esca grapevine disease
52 in a vineyard. We used leaf symptom data recorded in a vineyard in the
53 Bordeaux region from 2004 to 2017.

54 The paper is organised as follows: in this section we first give the formal-
55 ism to define random field, neighbours structure and autologistic models and
56 we review the literature on spatial and spatio-temporal autologistic mod-
57 els. In Section 2 we present our new centered autologistic model and discuss
58 the existence of the joint distribution of the spatio-temporal process. Then,
59 in Section 3, we present several simulations results to compare the spatio-
60 temporal autologistic model depending on the past under different centered
61 parametrisations. In section 4, we present an inference algorithm to calcu-

62 late the Maximum Pseudo-Likelihood Estimator of our model. We evaluate
63 it by calculating parameter estimates on simulated data. We also present a
64 method for choosing between possible neighbourhood structures. We study
65 its performance by comparing the selected neighbourhood structures on sim-
66 ulated data with the actual structure. Section 5 presents an analysis of real
67 data on the spread of a disease in a vineyard. In the last section, we discuss
68 the interest of the methodology and give some perspectives on this study.

69 *1.1. Spatial autologistic models*

70 Let $[Z]$ denote the distribution of random variable Z . Let $\mathbf{Z} = \{Z_i : i =$
71 $1, \dots, n\}$ be the random field where $Z_i \in \{0, 1\}$ represents the state at the i th
72 point s_i of a lattice $S = \{s_1, \dots, s_n\}$. The distribution $[Z]$ is given by the
73 conditional laws

$$[Z_i | Z_j, j \neq i] \sim \text{Binary}(p_i),$$

74 where for $1 \leq i \leq n$, $p_i = \mathbb{P}(Z_i = 1 | Z_j, j \neq i)$.

75 In addition, we have a neighbourhood structure: we assume that each
76 position s_i is associated with a set N_i containing the neighbours of s_i . We
77 suppose that this relation is symmetric that is $s_j \in N_i \Leftrightarrow s_i \in N_j$. This
78 neighbourhood structure defines a non-oriented graph whose nodes are the
79 locations s_i 's and there is an edge between s_i and s_j if $s_j \in N_i$. From now
80 on, we will use an notation abuse and confuse i and s_i in the formulae by
81 using " $j \in N_i$ " instead of " j such that $s_j \in N_i$ ".

If moreover we assume that the conditional distributions of the random

field \mathbf{Z} satisfy the following Markov property:

$$[Z_i|Z_j, j \neq i] = [Z_i|Z_j, j \in N_i] \quad \text{for all } 1 \leq i \leq n, \quad (1)$$

82 then \mathbf{Z} is said to be a Markov random field associated with the neighbour
 83 structure given by the N_i 's. The conditional binary probability can be ex-
 84 pressed in exponential family form:

$$\mathbb{P}(Z_i|Z_j, j \in N_i) = \frac{\exp(Z_i A_i(Z_j, j \in N_i))}{1 + \exp(A_i(Z_j, j \in N_i))} \quad (2)$$

85 where A_i is called a natural parameter function. When the Z_i 's take values
 86 in $\{0, 1\}$ and that will be the case in this paper, the use of the logit function
 87 instead of the A_i to model the p_i is very common; it is defined by the following
 88 equation:

$$A_i(Z_j, j \in N_i) = \text{logit}(p_i) = \log\left(\frac{p_i}{1 - p_i}\right) = \log\left(\frac{\mathbb{P}(Z_i = 1|Z_j, j \in N_i)}{\mathbb{P}(Z_i = 0|Z_j, j \in N_i)}\right).$$

89 Besag (1974) showed that the natural parameter functions must be of form:

$$A_i(Z_j, j \in N_i) = \text{logit}(p_i) = \alpha_i + \sum_{j \in N_i} \rho_{ij} Z_j, \quad (3)$$

90 with α_i a leading constant. Note that the regression coefficients ρ_{ij} may
 91 depend on site s_i and its spatial neighbour s_j . The variables Z_i 's are inde-
 92 pendent from each other if for each $1 \leq i \leq n$, $\mathbb{P}(Z_i = 1|Z_j, j \neq i)$ does
 93 not depend on the Z_j . From Equations 2 and 3, we can easily see that it
 94 is equivalent to $\rho_{ij} = 0 \quad \forall (i, j) \in S \times S$ with $S = \{1, \dots, n\}$. It means
 95 that the parameters ρ_{ij} 's reflect the dependences within the lattice. An ob-
 96 vious question is the existence of the joint distribution of the spatial process

97 $\{Z_i : i = 1, \dots, n\}$ that is only defined through the conditional probabilities
 98 $[Z_i|Z_j, j \in N_i]$. Coefficients ρ_{ij} must satisfy certain restrictions for a joint
 99 distribution of the Z_i 's to exist (Gaetan and Guyon, 2008); in particular,
 100 Besag (1974) showed that the symmetry condition $\rho_{ij} = \rho_{ji}$ is necessary for
 101 the joint distribution to exist.

102 The modelling can be generalized to include covariates $\mathbf{X} = \{\mathbf{X}_i, i = 1, \dots, n\}$.
 103 In this case, the natural parameter function is given in Caragea and Kaiser
 104 (2009) by:

$$\text{logit}(p_i) = \mathbf{X}_i^T \boldsymbol{\beta} + \sum_{j \in N_i} \rho_{ij} Z_j, \quad (4)$$

105 with $\boldsymbol{\beta}$ is a k -vector of regression parameters. Caragea and Kaiser (2009)
 106 discussed the difficulties in interpreting model parameters for a non-centered
 107 autologistic model given by Equation 4. Indeed, let $p_i = \mathbb{P}(Z_i = 1|Z_j, j \in$
 108 $N_i, \mathbf{X}_i)$ so that the odds that $Z_i = 1$ in model (4) is $p_i/(1 - p_i)$. Let $c_i =$
 109 $\frac{\exp(\mathbf{X}_i^T \boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i^T \boldsymbol{\beta})}$ being the probability of occurrence under the spatial independence
 110 model (when all the ρ_{ij} 's equal 0), the odds that $Z_i = 1$ is $c_i/(1 - c_i)$. Then
 111 the log odds ratio for model (4) relative to the independence model is:

$$\log \left(\frac{p_i/(1 - p_i)}{c_i/(1 - c_i)} \right) = \sum_{j \in N_i} \rho_{ij} Z_j$$

112 In this case, the odds of $Z_i = 1$ in model (4) relative to the independence
 113 model increases for any nonzero neighbours, and can never decrease. This is
 114 not reasonable if most of neighbours are zeros and could bias the realisations
 115 towards 1.

116 To overcome these difficulties of interpretation, a centered spatial autologistic

117 model was introduced by Caragea and Kaiser (2009). In this model, the value
 118 of the Z_j 's in the regression is centered by their expected "large-scale" value
 119 and the p_i 's are given by:

$$\text{logit}(p_i) = \mathbf{X}_i^T \boldsymbol{\beta} + \sum_{j \in N_i} \rho_{ij} \left(Z_j - \frac{\exp(\mathbf{X}_j^T \boldsymbol{\beta})}{1 + \exp(\mathbf{X}_j^T \boldsymbol{\beta})} \right). \quad (5)$$

120 The authors of Caragea and Kaiser (2009) pointed out that model (5) is
 121 similar to the parametrisation customarily used for auto-Gaussian models.
 122 They show that the alternative (centered) parameterization overcomes the
 123 difficulty of interpretation of the non centered model.

124 In the model given by Equation 5, the term $\mathbf{X}_i^T \boldsymbol{\beta}$ (called "large-scale
 125 model component by Caragea and Kaiser (2009)), determines the expected
 126 value of p_i . On average over the entire field, $\frac{\exp(\mathbf{X}_j^T \boldsymbol{\beta})}{1 + \exp(\mathbf{X}_j^T \boldsymbol{\beta})}$ corresponds to the
 127 proportion of Z_i that are equal to 1. Variances, covariances and other high-
 128 order portions of the data structure are determined by the second term of
 129 Equation 5, Caragea and Kaiser (2009) called it the small-scale model com-
 130 ponent. About the modelling of Markov random field and the interpretation
 131 of the coefficients of $\text{logit}(p_i)$ in term of small or large scale-structure, see
 132 Kaiser and Cressie (2000) or Cressie (1993) p.114. In Caragea and Kaiser
 133 (2009), the parameters were estimated by MPL.

134 Hughes et al. (2011) focused on the methods for estimating the coefficients
 135 and parameters in the centered autologistic model. They used MPLE and
 136 also parametric bootstrap, Monte Carlo Maximum Likelihood (MCML) and
 137 MCMC Bayesian approaches to infer the parameters. They also discussed
 138 ways to optimise the effectiveness of their algorithms. They also compared
 139 the performance of the three approaches in an in-depth simulation study.

140 They found that “inference for regression parameters in the centered model
 141 is reliable only for reasonably large lattices ($n > 900$) and no more than
 142 moderate spatial dependence”. They recommended the MPLE for its easier
 143 implementation and much faster execution. A package for the free software R
 144 is available estimating the centered spatial models parameters (Hughes, 2014).
 145 More recently, Wolters (2017) discussed coding and centering in spatial au-
 146 tologistic models and he recalled the good properties of MPLE for estimating
 147 non-centered models parameters. It is not always true in the centered case
 148 because PL can exhibit multiple local optima.

149 *1.2. Spatio-temporal autologistic models*

150 Now let \mathbf{Z}_t to denote a random field indexed by discrete time t . Z_{it} for
 151 $i = 1, \dots, n$ and $t \in \mathbb{Z}$ is a random binary variable indexed by position s_i and
 152 time t . And the covariates \mathbf{X} are k -vectors indexed by i and t .

153 Zhu et al. (2005) generalised the autologistic regression models to account
 154 for covariates, spatial dependence, and temporal dependence simultaneously.
 155 The model specifies the joint distribution of $\{\mathbf{Z}_t : t \in \mathbb{Z}\}$ by a family of
 156 conditional distributions:

$$\mathbb{P}(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_2} | \mathbf{Z}_t; t \neq t_1, \dots, t_2) \tag{6}$$

$$\propto \exp \left\{ \sum_{t'=t_1}^{t_2} \left[\sum_{i=1}^n \mathbf{x}_{i,t'}^T \boldsymbol{\beta} Z_{i,t'} + \frac{1}{2} \sum_{i=1}^n \sum_{j \in N_i} \rho_1 Z_{i,t'} Z_{j,t'} \right] + \sum_{t'=t_1}^{t_2+1} \sum_{i=1}^n \rho_2 Z_{i,t'} Z_{i,t'-1} \right\}.$$

157 for all $t_1, t_2 \in \mathbb{Z}^2$ such that $t_1 < t_2$, where $\mathbf{X}_{i,t}$ is the k -vector of the covariates
 158 at site s_i and time t . Note that the specification is consistent for all $t_1 < t_2$,
 159 and the joint distribution of $\{\mathbf{Z}_t : t \in \mathbb{Z}\}$ can be shown to exist by Theorem
 160 2.1.1 of Guyon (1995). If $N_{i,t} = \{(j, t) : j \in N_i\} \cup \{(i, t-1), (i, t+1)\}$

161 denotes the neighbourhood set for the position s_i and the t th time point, the
 162 full conditional distribution of the model is

$$[Z_{i,t}|Z_{i',t'} : (i', t') \neq (i, t)] = [Z_{i,t}|Z_{i',t'} : (i', t') \in N_{i,t}] \quad \text{and}$$

$$\begin{aligned} & \text{logit}(\mathbb{P}(Z_{i,t} = 1|Z_{i',t'} : (i', t') \in N_{i,t}; \mathbf{X})) & (7) \\ & = \mathbf{X}_{i,t}^T \boldsymbol{\beta} + \sum_{j \in N_i} \rho_1 Z_{j,t'} + \rho_2 (Z_{i,t-1} + Z_{i,t+1}). \end{aligned}$$

163 Note that in model (6), the coefficients corresponding to the ρ_{ij} of Equations
 164 (3), (4), (5), are all the same and they equal ρ_1 meaning that all the
 165 spatial neighbour's relations have the same intensity. The coefficient ρ_2 cor-
 166 responds to a "temporal" autoregression, which is the same regardless of
 167 spatial position and time. On the other hand, Zheng and Zhu (2008) pointed
 168 out that one drawback of the parameter estimation for the spatio-temporal
 169 autologistic regression model presented by Zhu et al. (2005), was based on
 170 MPLE whose statistical efficiency is not well established in the centered case.
 171 Zheng and Zhu (2008) suggest a fully Bayesian approach and compared it
 172 to estimation via MPL or MCMC Maximum Likelihood. Another drawback
 173 is the difficulty of using model (7) for a real application. Indeed it seems
 174 unrealistic to model the probability of an event occurring in terms of the
 175 future.

176 It is probably why Zhu et al. (2008) developed a spatio-temporal autol-
 177 ogistic regression model which depends only on the past. They suggested
 178 to infer the model parameters by maximum Likelihood estimation. On one
 179 hand, they assume that the temporal dependencies satisfy the following prop-

erty $[\mathbf{Z}_t | \mathbf{Z}_{t'}, t' = t - 1, t - 2, \dots] = [\mathbf{Z}_t | \mathbf{Z}_{t'}, t' = t - 1, \dots, t - \tau]$ i.e. the model is
a Markov model of order τ (the distribution of Z_t depends to the past only
through the τ last times). On the other hand, for a given time t , the spatial
field is a spatial Markov random field and the conditional probabilities are
given by

$$\begin{aligned} \text{logit} [\mathbb{P}(Z_{it} = 1 | Z_{jt}, j \in N_i; \mathbf{Z}_{t'}, t' = t - 1, \dots, t - \tau; \mathbf{X})] & \quad (8) \\ & = \mathbf{X}_{i,t'}^T \boldsymbol{\beta} + \sum_{j \in N_i} \rho_1 Z_{j,t} + \sum_{s=1}^{\tau} \rho_{1+s} Z_{i,t-s}. \end{aligned}$$

Here the model allows the temporal autoregression to be of order greater
than 1. And the coefficients ρ_{1+s} depend on the order s of the autoregres-
sion. For instance for the spread of an illness, we can expect that all these
coefficients are positive which means that the probability increases with the
age of the symptoms. However, nothing is said about the existence of the
joint distribution of such a process in the paper.

To overcome the difficulties of interpretation identified by Caragea and
Kaiser (2009) in the spatial case, Wang and Zheng (2013) were the first to
develop a centered parametrisation version of Equation 8 in a spatio-temporal
framework. They propose the following modelling:

$$\begin{aligned} \text{logit}(\mathbb{P}(Z_{i,t} = 1 | Z_{i',t'} : (i', t') \in N_{i,t}; \mathbf{X})) & \quad (9) \\ & = \mathbf{X}_{i,t}^T \boldsymbol{\beta} + \sum_{j \in N_i} \rho_1 Z_{j,t}^* + \rho_2 (Z_{i,t-1}^* + Z_{i,t+1}^*), \end{aligned}$$

$$\text{with } Z_{i,t}^* = Z_{i,t} - \frac{\exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta})}{1 + \exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta})}. \quad (10)$$

195 They proposed Expectation-Maximisation (EM) algorithm to maximise
 196 the Pseudo-Likelihood and Monte Carlo Expectation-Maximisation Likeli-
 197 hood, as well as consider Bayesian inference to obtain the estimates of model
 198 parameters. They found that Monte Carlo Expectation-Maximisation Like-
 199 lihood algorithm is optimal taking into account the criteria of calculation
 200 time and accuracy of the estimate. Further, they compared the statistical
 201 efficiency of these approaches.

202 In the next section we propose a new centered spatio-temporal model and
 203 we believe that such model is better adapted for parameter interpretation in
 204 a spatio-temporal context.

205 **2. A new centered spatio-temporal autologistic model**

206 *2.1. Model specification*

207 In this section, we propose a new spatio-temporal autologistic model,
 208 specified by Markov field Markov chain (Gaetan and Guyon, 2008). It can
 209 include spatio-temporal covariates. In order to avoid bias and interpretation
 210 problems caused by spatial self-regression, we propose to center the corre-
 211 sponding covariates term. We define the model as follows. First, we assume
 212 that, conditionally to the covariates, $\{\mathbf{Z}_t, t = 1, 2, \dots\}$ is a Markov chain:

$$[\mathbf{Z}_t | \mathbf{Z}_{t-1}, \mathbf{Z}_{t-2}, \dots, \mathbf{X}] = [\mathbf{Z}_t | \mathbf{Z}_{t-1}, \mathbf{X}_t]$$

213 where \mathbf{X} is the spatio-temporal process of vector of the covariates $(X_{i,t})_{1 \leq i \leq n, 1 \leq t \leq T}$.
 214 Moreover, we assume that \mathbf{Z}_t is a Markov random field conditional on \mathbf{Z}_{t-1}
 215 with spatial neighbour structure N_i , that means

$$[Z_{i,t}|Z_{j,t}, j \neq i; \mathbf{Z}_{t-1}, \mathbf{X}_t] = [Z_{i,t}|Z_{j,t}, j \in N_i, \mathbf{Z}_{t-1}, \mathbf{X}_t].$$

216 More precisely, we define the conditional distribution of $Z_{i,t}$ by:

$$\text{logit}(\mathbb{P}(Z_{i,t} = 1|Z_{j,t}; j \neq i, \mathbf{Z}_{t-1}, \mathbf{X}_t)) = \mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_1 \sum_{j \in N_i} Z_{j,t}^{**} + \rho_2 Z_{i,t-1}, \quad (11)$$

$$\text{where } Z_{i,t}^{**} = Z_{i,t} - \frac{\exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_2 Z_{i,t-1})}{1 + \exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_2 Z_{i,t-1})}. \quad (12)$$

217 We will discuss the interest of this centering in the next section but we
 218 can note that this new centered specification looks like a hierarchical model
 219 with a latent auto-regressive model of order 1 given by:

$$\begin{aligned} \text{logit}(\mathbb{P}(\mathbf{Z}_{i,t} = 1|\xi_{i,t}, \mathbf{X}_{i,t})) &= \mathbf{X}_{i,t}^T \boldsymbol{\beta} + \xi_{i,t}, \\ \xi_{i,t} &= \rho_2 \xi_{i,t-1} + \omega_{i,t}, \quad \text{with} \\ Z_{i,t-1} &\approx \xi_{i,t-1} \quad \text{and} \\ \sum_{j \in N_i} Z_{j,t}^{**} &= \sum_{j \in N_i} \left(Z_{j,t} - \frac{\exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 Z_{j,t-1})}{1 + \exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 Z_{j,t-1})} \right) \approx \omega_{i,t}. \end{aligned}$$

220 2.2. Model Interpretation

221 There are two main differences between one-step centered model of Wang
 222 and Zheng (2013) given by Equation 10 and our new centered model given by
 223 Equation 12. The first one is that we do not center de temporal term $Z_{i,t-1}$.
 224 Indeed, there is no difficulty for the interpretation of ρ_2 because unlike the
 225 $\sum_{j \in N_i} Z_{j,t}$, $Z_{i,t-1}$ is known at time t and can be treated like the other covari-
 226 ates. The second one is that the $Z_{j,t}$'s in the term of spatial autoregression,

227 are centered differently: the former is centered with $\frac{\exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta})}{1 + \exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta})}$, and the lat-
 228 ter is centered with $\frac{\exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 Z_{j,t-1})}{1 + \exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 Z_{j,t-1})}$. It has to be noted that due to the
 229 link function, the expectation $\mathbb{E}(Z_{it} | \mathbf{Z}_{t-1}, \mathbf{X}_t)$ equals $\frac{\exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_2 Z_{i,t-1})}{1 + \exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_2 Z_{i,t-1})}$. The
 230 construction of this new centered autologistic model is explained as follows;
 231 the aim of the modelling is to separate the large- and small-scale structures.
 232 The parameters of spatio-temporal dependence ρ_1, ρ_2 can be interpreted and
 233 they have a practical interpretation if the event modelled is an illness:

- 234 • Instantaneous spatial dependence ρ_1 . It quantifies the spatial autocor-
 235 relation between neighbours for the occurrence of the event at each time.
 236 To model illness, we expect some kind of common sensibility quantified
 237 by $\rho_1 \geq 0$. Strong spatial dependence indicates a highly aggregated
 238 spatial structure, which makes it possible to identify and monitor the
 239 aggregated zones with high infection possibility.
- 240 • Temporal dependence ρ_2 . It quantifies the temporal dependence on the
 241 previous year's status. Again, we expect that the illness likely remains
 242 at a place such that $0 \leq \rho_2$. ($\rho_2 < 0$ would indicate a temporal evolu-
 243 tion with high frequency at 2-year cycle, this is not adapted for most
 244 of the biological processes); ρ_2 is a term of autoregressive regression.
 245 Strong temporal dependence can be interpreted as a smooth temporal
 246 evolution. If the external effects (covariates) are constant, the individ-
 247 uals have a tendency to keep their status. For instance, if we consider
 248 annual data (meaning that t refers to the year), this may indicate no
 249 need to monitor two consecutive years if the exterior factors are the
 250 same between two years.

251 *2.3. Existence of the joint distribution*

252 Many authors have discussed the existence of a joint multivariate distri-
 253 bution defined by a set of univariate conditional distributions for auto-models
 254 that include the autologistic models. Hammersley & Clifford were the first
 255 to work on this subject in 1971 (unpublished manuscript!) and the results
 256 were written and generalised for instance in Grimmett (1973); Kaiser and
 257 Cressie (2000); Gaetan and Guyon (2008). It is probably not possible to
 258 show the existence of the distribution of the whole spatio-temporal processes
 259 defined by Equation 11 for all years t and sites of the lattice s_i . In our case,
 260 the functionals that appear in the right side of Equation 11 are not invari-
 261 ant by permutating the temporal indices. Note that the model of Zhu et al.
 262 (2005) given by Equation 6 satisfies this invariance necessary for the joint
 263 distribution to exist, but it raises the problem of using the future ($Z_{i,t+1}$) to
 264 model the present ($Z_{i,t}$). To prove the existence of our process, we consider
 265 the framework of Markov chain of Markov fields presented in Guyon and
 266 Hardouin (2002). Using the Hammersley-Clifford theorem given for instance
 267 in Gaetan and Guyon (2008), we can show the existence of the joint law of
 268 spatial process \mathbf{Z}_t for a fixed t given the past of \mathbf{Z}_t and the current informa-
 269 tion about the covariates and derive an expression of this conditional joint
 270 spatial distribution. Thus, the existence of the distribution of the whole
 271 spatio-temporal process, is trivial by recursivity. In addition, we can also
 272 obtain the formula of the conditional transition probabilities of the spatial
 273 Markov chain. We have the following result.

274 **Theorem 2.1.** *Let $(\mathbf{Z}_t)_{(0 \leq t \leq T)}$ be the spatio-temporal process defined by (11),*
 275 *let us denote $\mathcal{F}_t^X = \sigma\{\mathbf{X}_{i,s}, 1 \leq i \leq n, s \leq t\}$ and $\mathcal{F}_t^Z = \sigma\{Z_{i,s}, 1 \leq i \leq$*

276 $n, s \leq t\}$ the σ -algebra generated by the covariates and the process of interest
 277 respectively.

278 Given $\mathcal{F}_{t-1}^Z \wedge \mathcal{F}_t^X$ the conditional joint law of \mathbf{Z}_t denoted by $\pi_t(\cdot \mid \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z)$
 279 is well defined. Moreover, for $\mathbf{z} = (z_1, \dots, z_n) \in \{0, 1\}^n$, the spatial condi-
 280 tional joint law is of the form

$$\begin{aligned} \pi_t(\mathbf{z} \mid \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) &= C(\mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) \exp\left(\sum_{i \in S} \Phi_i(z_i) + \sum_{\{i,j\}} \Phi_i(z_i, z_j)\right) \quad \text{with} \\ \Phi_i(z_i, \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) &= z_i \left(\mathbf{X}_{i,t}^T \boldsymbol{\beta} - \rho_1 \sum_{j \in N_i} \frac{\exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 Z_{j,t-1})}{1 + \exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 Z_{j,t-1})} + \rho_2 Z_{i,t-1} \right) \\ \Phi_i(z_i, z_j) &= \rho_1 \mathbb{1}_{\{j \in N_i\}} z_i z_j \end{aligned}$$

281 and $C(\mathcal{F}_t^X, \mathcal{F}_{t-1}^Z)$ can be considered as constant if the past of process \mathbf{Z}_t and
 282 the current values at time t of the covariates are known.

283

284 The transition probabilities of the Markov chain are

$$\begin{aligned} \mathbb{P}(\mathbf{y}, \mathbf{z} \mid \mathcal{F}_t^X) &= C(y, \mathcal{F}_t^X) \exp\left(\sum_{i \in S} (z_i \mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_1 \sum_{j \in N_i} z_i z_j)\right) \\ &\times \exp\left(\sum_{i \in S} z_i (\rho_2 y_i - \rho_1 \sum_{j \in N_i} \frac{\exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 y_j)}{1 + \exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 y_j)} - \rho_2 y_i)\right) \end{aligned}$$

285 **Proof:** We use the Hammersley-Clifford theorem given for instance in
 286 Gaetan and Guyon (2008) in a conditional form. With the above notation,
 287 let us define two assumptions given by (13) and (14) that says that if

$$\begin{aligned} \text{logit}(\mathbb{P}(Z_{i,t} = z_i \mid Z_t^i = z^i, \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z)) & \tag{13} \\ &= A_i(z^i, \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) B_i(z_i, \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) + C_i(z_i, \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) + D_i(z^i, \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) \end{aligned}$$

and if $\forall i \neq j$, it exists α_i and $\rho_{ij} = \rho_{ji}$ such that

$$A_i(z^i, \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) = \alpha_i(\mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) + \sum_{j \neq i} \rho_{ij} B_j(z_j, \mathcal{F}_t^X, \mathcal{F}_{t-1}^Z) \quad (14)$$

288 The conditional laws satisfying (13) and (14) are consistent with a joint
 289 distribution that is a Markov random field. If we rewrite (11) in the following
 290 form, it is easy to see that our model satisfies the required conditions:

$$\begin{aligned} & \text{logit}(\mathbb{P}(Z_{i,t} = 1 | Z_{j,t}; j \neq i, \mathbf{Z}_{t-1}, \mathbf{X}_t)) \\ &= \underbrace{\mathbf{X}_{i,t}^T \boldsymbol{\beta} - \rho_1 \sum_{j \in N_i} \frac{\exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 Z_{j,t-1})}{1 + \exp(\mathbf{X}_{j,t}^T \boldsymbol{\beta} + \rho_2 Z_{j,t-1})}}_{=\alpha_i(\mathcal{F}_t^X, \mathcal{F}_{t-1}^Z)} + \rho_2 Z_{i,t-1} + \sum_{j \neq i} \underbrace{\rho_1 \mathbb{1}_{j \in N_i}}_{\rho_{ij}} \underbrace{Z_{j,t}}_{B_j(z_j)}. \end{aligned}$$

291 The expression of π comes directly from Hammersley-Clifford theorem and
 292 the transitions probabilities from Guyon and Hardouin (2002). \square

293 3. Comparative simulation study

294 3.1. Simulation study objective

295 The idea of proposing the new centered model is to make an agreement
 296 between large-scale model and data structure. In particular, if parameters are
 297 intended to reflect an overall mean or the effect of covariates, then they should
 298 have a constant interpretation across varying levels of statistical dependence.
 299 In this section, we compare three models: the model without centering that
 300 Caragea and Kaiser called “traditional” and one-step centered as well as the
 301 new centered model, defined again below. We want to verify if the marginal
 302 structure of data reflects the large-scale structure. The three studied models
 303 are given by the general following equation differing according to the kind of
 304 centering of $Z_{i,t}^{**}$.

$$\text{logit}(p_{i,t}) = \mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_1 \sum_{j \in N_i} Z_{j,t}^{**} + \rho_2 Z_{i,t-1},$$

305

$$Z_{i,t}^{**} = Z_{i,t} \quad \text{traditional model} \quad (15)$$

$$Z_{i,t}^{**} = Z_{i,t} - \frac{\exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta})}{1 + \exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta})} \quad \text{one-step model} \quad (16)$$

$$Z_{i,t}^{**} = Z_{i,t} - \frac{\exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_2 Z_{i,t-1})}{1 + \exp(\mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_2 Z_{i,t-1})} \quad \text{new model} \quad (17)$$

306 The agreement between spatial large-scale model structures and marginal
 307 data structures have been already examined by Caragea and Kaiser (2009)
 308 for both centered and traditional spatial autologistic regression models. They
 309 showed that the realised trajectories from the traditional autologistic regres-
 310 sion model can not reflect the large-scale structure, and this difficulty can be
 311 alleviated by centered parametrisation.

312 In this paper, we focus on examining the time variation of the large-scale
 313 model structure and marginal data structure for the three spatio-temporal
 314 autologistic models mentioned above in the case when the covariates depend
 315 only on the time t but not on the location on the grid. Note that Caragea and
 316 Kaiser (2009) has already carried out a simulation study in the purpose to
 317 verify the agreement between the spatial model and spatial data structures
 318 according to the different values of the spatial covariates.

319

320 Therefore, we have chosen to simulate trajectories of a dynamic process
 321 of Markov random field that have a temporal large-scale structure with a
 322 deterministic tendency. For site s_i at year t , we define a large model structure
 323 with one temporal covariate: $\text{logit}(p_{i,t}) = \beta_0 + \beta_1 X_{it} + \rho_1 \sum_{j \in N_i} Z_{j,t}^{**} + \rho_2 Z_{i,t-1}$,

324 where $X_{it} = X_t$ is a temporal covariate, that is spatial constant at year t .

325 Thus the average large-scale model at year t is:

$$L_t = \frac{1}{n} \sum_{i=1}^n \frac{\exp(\beta_0 + \beta_1 X_t)}{1 + \exp(\beta_0 + \beta_1 X_t)} = \frac{\exp(\beta_0 + \beta_1 X_t)}{1 + \exp(\beta_0 + \beta_1 X_t)}. \quad (18)$$

326 Moreover, we can also compute an average scale conditional to the past
327 defined by:

$$C_t = \frac{1}{n} \sum_{i=1}^n \frac{\exp(\beta_0 + \beta_1 X_t + \rho_2 Z_{i,t-1})}{1 + \exp(\beta_0 + \beta_1 X_t + \rho_2 Z_{i,t-1})}. \quad (19)$$

328 It can be noted that this process is not deterministic but specific to each
329 realisation of the field \mathbf{Z}_{t-1} .

330 To represent marginal data structure, the marginal empirical data mean
331 of the Markov random field at year t is computed from a simulated field at
332 time t as:

$$D_t = \frac{1}{n} \sum_{i=1}^n Z_{it}^{sim}. \quad (20)$$

333 The objective of the simulation studies presented here is to compare the
334 behaviour of L_t , C_t and that of D_t for the three different models differing by
335 the kind of centering of the spatial autoregression.

336 3.2. Sampling Algorithms

337 Hughes et al. (2011) proposed to use perfect sampling to generate Markov
338 Random Field (MRF) samples. The advantage of the perfect sampling com-
339 pared to Gibbs sampling is that we don't need the burn-in step, nor do we
340 need to decide the spacing numbers. It gives us the exact draws from a given

341 target distribution when the algorithm completes, details are given in Kendall
342 (2005) Kendall (2005). However, its algorithm running time is random even
343 is still finite. We do not know at which moment the lower chain and the
344 upper chain coalesce. So the number of repetitions is random. In our case,
345 we have to generate a Markov chain Markov random field, the computation
346 time of such a algorithm is quite difficult to control.

347 Here we use Gibbs sampler but start at a "perfect simulated" sample, we
348 call it PGS sampling. It is less time consuming than perfect sampling, and do
349 not need to decide burn-in and spacing when compared with Gibbs sampling.
350 The PGS sampling was often used to generate the simulated trajectories of
351 autologistic model (Zhu et al., 2008; Zheng and Zhu, 2008; Wang and Zheng,
352 2013).

353 *3.3. Simulated data*

354 We focus on two types of large-scale model structures, the first one with-
355 out covariate and the second one with increasing temporal tendency given
356 by a covariate depending only on the time. Both models were for data on a
357 20 by 20 lattice for 50 time units; the details being given in the two following
358 sections 3.3.1 and 3.3.2. We simulated data according to these structures
359 but with different values of "auto-regression parameter" (ρ_1, ρ_2) , in order
360 to evaluate the joint effects of (ρ_1, ρ_2) to the agreement between large-scale
361 structure models and data structures. One trajectory of each of the three
362 models is drawn for each configuration of (ρ_1, ρ_2) . To study the dispersion
363 of the empirical large scale structure and confirm the possible tendencies ex-
364 hibited by the trajectories, we have performed 100 independent realisations
365 of process D_t for each of the three models (traditional, one-step and new

366 centered) and specified by different values of (ρ_1, ρ_2) and then computed and
 367 drawn the empirical fluctuation dynamical intervals. Simulations results are
 368 presented in the following sections.

369 3.3.1. Model 1

370 We first consider a model without covariates (that is with only an inter-
 371 cept). It is given by $\text{logit}(p_{it}) = \beta_0 + \rho_1 \sum_{j \in N_i} Z_{j,t}^{**} + \rho_2 Z_{i,t-1}$, there is no
 372 temporal covariate in this model apart the term of temporal auto-regression.
 373 The baseline level of infection is chosen via β_0 such that $\frac{\exp(\beta_0)}{1+\exp(\beta_0)} = 0.2$ lead-
 374 ing to a constant expected model structure constant $L_t = 0.2$. The initial
 375 field is generated by independent Bernoulli variables with parameter $p = 0.2$.
 376 To see the effect of the intensity of autocorrelation given by the two param-
 377 eters (ρ_1, ρ_2) on the difference between the models, we generate simulated
 378 data and draw one trajectory from each model with different values of pa-
 379 rameters $(\rho_1, \rho_2) \in \{0.3, 0.5, 0.7\}^2$. The empirical confidence intervals were
 380 computed with 100 realisations of independent trajectories for each models
 381 and for different values of parameters $(\rho_1, \rho_2) \in \{0.5, 0.7\}^2$.

382 3.3.2. Model 2

383 We consider here a model with a temporal trend. We consider large-
 384 scale structures with one temporal covariate: $\beta_0 + \beta_1 X_t$. As said above, X_t
 385 is constant spatially for each year but show monotonic increasing with time:
 386 $X_t = t$.

387 We choose β_0 such that $\frac{\exp(\beta_0)}{1+\exp(\beta_0)} = 0.1$ and the initial field is generated by
 388 independent Bernoulli variables with parameter $p = 0.1$. With this model,
 389 we have $L_t = \frac{\exp(\beta_0 + \beta_1 t)}{1 + \exp(\beta_0 + \beta_1 t)}$, a monotonic increasing function. With the chosen



Figure 1: Comparison between large-scale model structure L_t (represented by black line), the expected means C_t according to the past (lines) and empirical mean of data structures D_t for the traditional (blue dashed), the one-step (red dot-dashed) and new (violet dots) centered models for different values of auto-regression parameters (ρ_1, ρ_2) . The grid is 20×20 , and $0 \leq t \leq 50$, baseline infection $\frac{\exp(\beta_0)}{1+\exp(\beta_0)} = 0.2$.

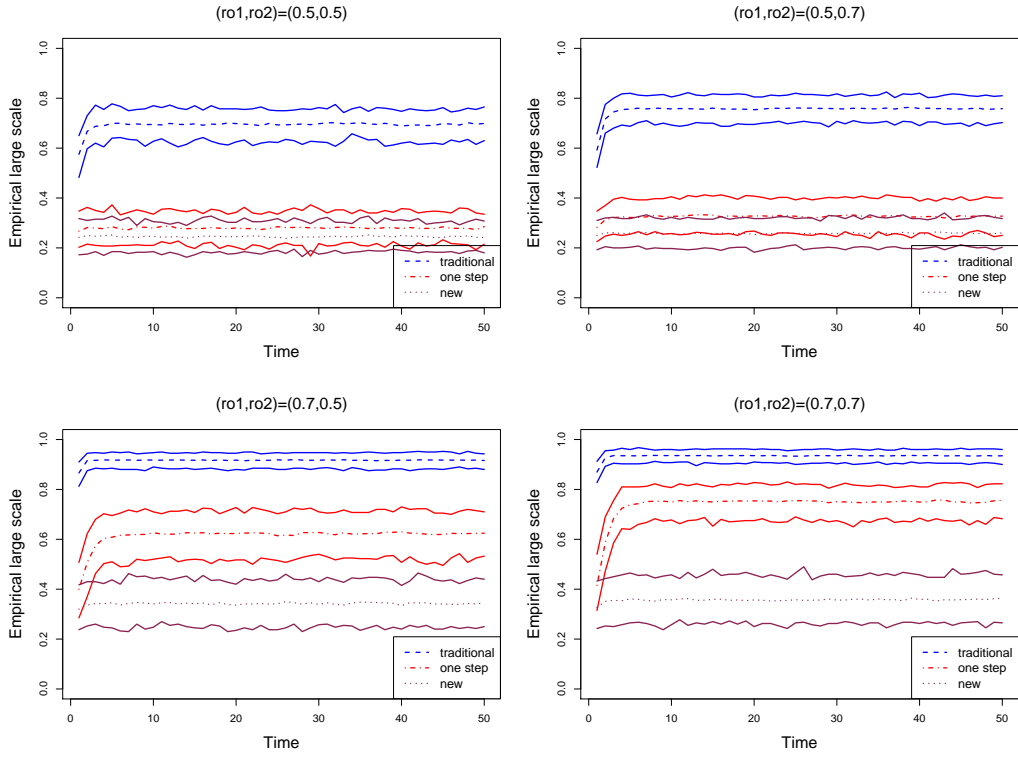


Figure 2: Empirical confidence curves for the large scale structures D_t for the model without covariate. The grid is 20×20 , $t = 50$, baseline infection $\frac{\exp(\beta_0)}{1+\exp(\beta_0)} = 0.2$. About the centering, traditional (resp. one-step and new centered) is drawn in blue dashed, (resp. red dot-dashed and violet dots).

390 coefficients, L_t increases from 0.1 at time one to 0.94 at time 50. Again, for
391 a trajectory, we compute the conditional mean given the past. Unlike the
392 model without covariates, we expect an empirical large-scale structure that
393 increases with time.

394 3.3.3. Results

395 From Figure 2 and Figure 4, we see that the spread of the realisations of
396 the empirical mean D_t is not very large and that for the given value of the
397 parameter (ρ_1, ρ_2) , the models may differ more or less. We see that scatter
398 of the empirical average of the spatial field for each t is low enough to trust
399 and interpret the difference between the individual curves shown in Figure 1.

400 In Figure 1 (resp. Figure 3) we draw one trajectory from each model
401 without covariate (resp. with covariates) for different values of ρ_1, ρ_2 to
402 study the difference between the models according to these parameters that
403 reflect the dependence level between the $Z_{i,t}$. We see that the empirical mean
404 D_t for the traditional model is always greater than the expected large scale
405 structure and than the realisation of D_t for the one-step centered and new
406 centered models. The value of D_t for the traditional model is not sensible to
407 the value of the temporal autoregression parameter ρ_2 but it is very sensible
408 to the spatial auto-regression parameter ρ_1 and it increases as ρ_1 increases.

409 Regarding the centered models, they show the same large scale behaviour
410 when parameter ρ_1 equals 0.3 or 0.5. For $\rho_1 = 0.7$, the two centered models
411 show different large scale behaviour but the D_t for new centered model agrees
412 with the expected mean behaviour. The difference between them increases
413 as ρ_2 increases.

414 In summary, the difference between data simulated according to the one-

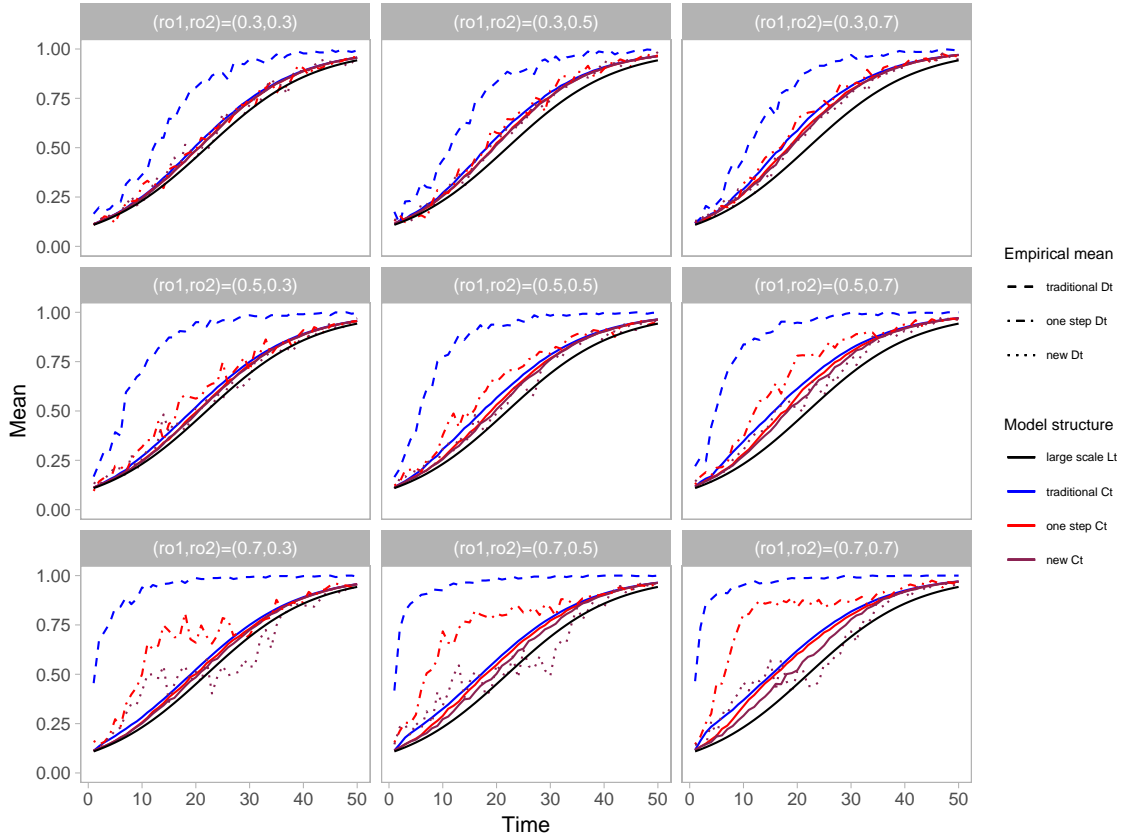


Figure 3: Comparison between large-scale model structure L_t (represented by black line), the expected means C_t according to the past (lines) and empirical mean of data structures D_t for the traditional (blue dashed), the one-step (red dot-dashed) and new (violet dots) centered models for different values of auto-regression parameters (ρ_1, ρ_2) . The grid is 20×20 , and $0 \leq t \leq 50$, baseline infection $\frac{\exp(\beta_0)}{1+\exp(\beta_0)} = 0.1$, and covariate coefficient $\beta_1 = 0$, and covariate $X_t = t$.

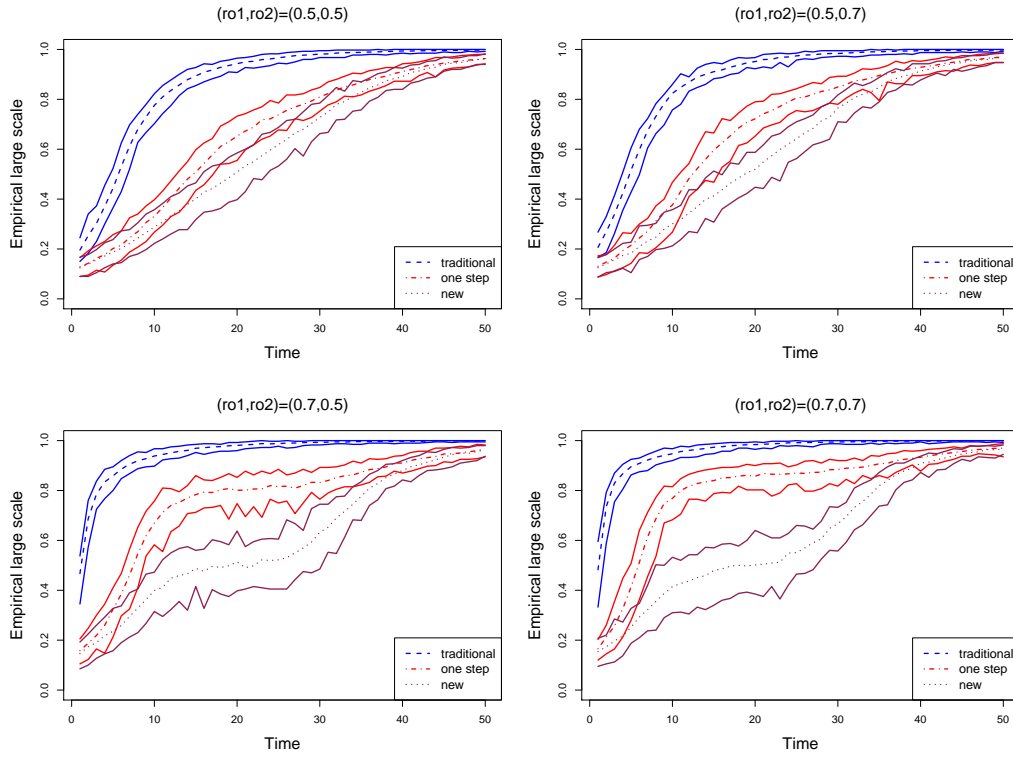


Figure 4: Empirical confidence curves for the large scale structures D_t for the model with temporal covariate $X_t = t$. The grid is 20×20 , $t = 50$, baseline infection $\frac{\exp(\beta_0)}{1+\exp(\beta_0)} = 0.1$, and covariate coefficient $\beta_1 = 0$, and covariate $X_t = t$. About the centering, traditional (resp. one-step and new centered) is drawn in blue dashed, (resp. red dot-dashed and violet dots).

415 step model and the new centered autologistic model are small except when
416 both spatial and temporal dependence are relatively strong. It can be seen
417 from the simulated data that the traditional or one-step centered model over-
418 represents the large-scale structure and incorrectly differs from the expected
419 structure of the model. As a result, interpretations are complex and can lead
420 to erroneous conclusions.

421 **4. Estimation**

422 From now on, we only consider the new centered model. We propose here
423 a method of estimation by Pseudo-Likelihood Maximisation (MPL) and a
424 method for selecting the most likely neighbourhood structure of the data set.
425 We show their performances by testing them on simulated data.

426 *4.1. Estimation by Pseudo-Likelihood Maximisation*

427 Since the structure of the new centered autologistic model is more com-
428 plicated than the traditional or the one-step centered one, both Monte Carlo
429 Maximum Likelihood Estimation (MCML) and Bayesian methods can be
430 very heavy and sophisticated to implement. We propose to estimate the pa-
431 rameters of our model using the estimator that maximises Pseudo-Likelihood.
432 It is very easy to implement. Some authors have studied the mathematical
433 properties such as convergence of this estimator under the constraint of ho-
434 mogeneity and ergodicity of the Markov Random Field Markov Chain and
435 other required assumptions can be found in Guyon and Hardouin (2002).
436 The imbrication of the parameters in the definition of the centered variables
437 $Z_{i,t}^{**}$'s (Equation 17) and the presence of covariates make these conditions
438 hard to verify, that is why, in the following, we only look at the empirical

439 behaviour of this estimator. The Pseudo-Likelihood for our model is given
 440 by the following formula:

$$\begin{aligned}
 \mathcal{PL}(\boldsymbol{\beta}, \rho_1, \rho_2) &= \prod_{t=1}^T \left(\prod_{1 \leq i \leq n} p_{it} \right) & (21) \\
 &= \prod_{t=1}^T \left(\prod_{1 \leq i \leq n} \frac{\exp \left(\mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_1 \left(\sum_{j \in N_i} Z_{j,t}^{**} \right) + \rho_2 Z_{i,t-1} \right)}{1 + \exp \left(\mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_1 \left(\sum_{j \in N_i} Z_{j,t}^{**} \right) + \rho_2 Z_{i,t-1} \right)} \right).
 \end{aligned}$$

441 MPL Estimator is the vector of parameters $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \rho_1, \rho_2\}$ that maximises
 442 $\mathcal{PL}(\boldsymbol{\beta}, \rho_1, \rho_2)$. We see that the spatial auto-regression covariate $Z_{j,t}^{**}$ imbricated
 443 the couple of parameters (ρ_1, ρ_2) themselves so that it is not possible
 444 to consider $Z_{j,t}^{**}$ as a common “external covariate”. For this reason, the max-
 445 imisation has to be made by an Expectation-Maximisation algorithm. We
 446 give the details of this algorithm in the next section.

447 4.1.1. Algorithm

448 We applied the EMPL (Expectation Maximisation Pseudo-Likelihood) al-
 449 gorithm, the principle is the same as described in Zheng and Zhu (2008), but
 450 with two iteration steps to accelerate the numerical algorithm/calculation.

451

452 The steps are as follows:

- 453 • Initialisation: to obtain the estimation of $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}, \rho_2)$, denoted by
 454 $(\tilde{\boldsymbol{\beta}}, \tilde{\rho}_2)$, from model $\text{logit}(p_{it}) = \mathbf{X}_{i,t}^T \boldsymbol{\beta} + \rho_2 Z_{i,t-1}$, we maximise the cor-
 455 responding log Pseudo(partial)-Likelihood by Quasi-Newton.
- 456 • Step 2: to obtain the estimation of $\boldsymbol{\theta} = (\boldsymbol{\beta}, \rho_1, \rho_2)$, denoted by $\check{\boldsymbol{\theta}} =$
 457 $(\check{\boldsymbol{\beta}}, \check{\rho}_1, \check{\rho}_2)$, for the new centered autologistic model:

- 458 1. Initialization: Set initial values: $\boldsymbol{\theta}^0 = (\tilde{\boldsymbol{\beta}}, 1, \tilde{\rho}_2)$
- 459 2. Expectation: Given $\boldsymbol{\theta}^{l-1}$, compute the $Z_{j,t}^{**}$'s by removing the cor-
460 responded trend.
- 461 3. Maximisation: Obtain $\boldsymbol{\theta}^l$ by maximising the log Pseudo-Likelihood
462 by Quasi-Newton.
- 463 4. Go to 2
- 464 • Obtain estimates $(\check{\boldsymbol{\beta}}, \check{\rho}_1, \check{\rho}_2)$.

465 *4.1.2. Variance of the MPLE*

466 Even if the convergence properties of the Maximum Pseudo-Likelihood
467 Estimator is well known under specified conditions discussed above, the
468 variance of the estimator has to be carefully estimated. For the variance-
469 covariance matrix of the coefficients, we propose to compute the matrix
470 $\mathbf{U}'\mathbf{W}\mathbf{U}$ as if we were in the case of a classic variance in the case of Maximum
471 Likelihood in the logistic case. The matrix \mathbf{U} is a $nT \times p$ matrix defined by
472 theses rows $\mathbf{U}_{it.} = (1, X_i^T, \sum_{j \in \mathcal{N}_i} Y_{jt}, Y_{i,t-1})$ for each (i, t) , $1 \leq t \leq T$ and
473 $1 \leq i \leq n$. \mathbf{W} is the diagonal $nT \times nT$ matrix with coefficients being equal
474 to $\check{p}_{it}(1 - \check{p}_{it})$ that depends on the estimation parameters $(\check{\boldsymbol{\beta}}, \check{\rho}_1, \check{\rho}_2)$.

475 We compared the variances of the estimators with ones computed by
476 “bootstrap” on simulated data.

477 *4.2. Model choice*

478 Although the estimation method is efficient for a given neighbourhood
479 structure, it is necessary to select the best one. Indeed, in the context of
480 modelling the occurrence of a disease, learning the structure of the neigh-
481 bourhood is a way to understand the mechanisms of disease spread.

482 Firstly, we had the idea of adapting the ABC method (for Approximate
483 Bayesian Computation) proposed by Grelaud et al. (2009) in order to choose
484 the best model for the neighbourhood. ABC is a Likelihood-free inference
485 method in the bayesian framework that is very convenient when the Likeli-
486 hood is not available in a closed form. First introduced by Pritchard et al.
487 (1999) and expanded in Beaumont et al. (2002) and Marjoram et al. (2003),
488 ABC method was adapted by Grelaud et al. (2009) for model choice in Gibbs
489 Random Fields (GRF). We first thought about using this method because
490 our model is not so far from a GRF. But because of the centering parametri-
491 sation, it is not possible to produce a sufficient statistics in order to compute
492 a simple distance between the simulated field and the observed one. We can
493 see in equations of the spatial joint laws of Theorem 2.1, that the parameters
494 (β, ρ_1) are nested in the potentials depending on the spatio-temporal field \mathbf{Z} .
495 We tried this method with different statistics without success. However, this
496 sophisticated method is not essential here because we observed on simulated
497 data that the Log-Pseudo-Likelihood value is a very simple and performant
498 indicator to choose the model of a given data set.

499 The approach we propose is the following: we use “experts opinions” to
500 determine a set of possible neighbours structures to consider. For each struc-
501 ture, we estimate the parameters and simply choose the structure that opti-
502 mises Log-Pseudo-Likelihood.

503 *4.3. Simulation study*

504 *4.3.1. Model 1*

505 The first configuration was again $20 * 20$ grid for 15 years without covari-
506 ate. The initial field is generated by Bernoulli distribution with parameter

507 0.1 and model parameters are $\beta_0 = -1.4, \beta_1 = 0, \rho_1 = 0.5, \rho_2 = 0.5$.

508 We standardise the neighbourhood structure here: we suppose the points
509 s_i 's are on a grid, and we assume their spatial distribution is like a matrix –
510 each point is located on the intersection of a row and a column. All models
511 considered in this section to generate simulated data have the same structure
512 of neighbourhood: each point has four neighbours on the same row (the
513 four nearest that is in our case the two nearest on each side) and also two
514 neighbours in the same column. Points of the first row (resp. second row)
515 have only two (resp. three) neighbours on the same row and points on the
516 first and second columns have only one neighbour on the same column (and
517 so on for the last row or column). This configuration is denoted $v_r = 2$ (for
518 2 neighbours on each side in the row) and $v_c = 1$ (resp. 1 in the column).
519 Figure 7 show such kind of neighbourhood for $v_r = 4$ and $v_c = 2$.

520 For our simulation study to infer the neighbour's structure, we assume
521 that all the neighbourhood structures considered are regular that is, the
522 definition of the neighbourhood is the same for all the points of the grid.
523 And we only consider the neighbourhood structures defined by the number
524 of neighbours on each side in the same row v_r (or in the same column v_c).
525 We do not consider the possibility of having neighbours on the diagonal.

526 4.3.2. Model 2

527 Again a $20 * 20$ grid for 15 years with one temporal covariate with large
528 variation first increasing from 1 to 8 the 8 first years and next decreasing
529 by 1 until the year 15. $x(t) = t$ for $1 \leq 8$ and $x(t) = 14 - t$ for $1 \leq 7$,
530 $\beta_0 = -2.8, \beta_1 = 0.1, \rho_1 = 0.5, \rho_2 = 0.5$. The structure of neighbourhood is
531 given by $v_r = 2$ and $v_c = 1$.

532 *4.3.3. Estimation results*

533 We used again PGS sampling methods described in section 3.2 to simulate
 534 trajectories of the process on a 20 * 20 grid for 15 years in different config-
 535 urations. The purpose is to study the performance of the estimations and
 536 of the model selection procedure. We compute the estimations via EMPL
 537 algorithm detailed in section 4.1.

538 For the variance, we compared the values estimated by the method ex-
 539 plained above on a sample, with the experimental value of the variance when
 540 we estimate the parameters of $B = 500$ independent repetitions of the same
 541 model. Results of the inference for Model 1 (resp. Model 2) is given in table
 542 1 (resp. 2).

543 Figures 5 and 6 show the dispersion of $B = 500$ estimations of B inde-
 544 pendent simulations of each model.

	β_0	β_1	ρ_1	ρ_2
mean	-1.47(-1.4)	0.003(0)	0.519(0.5)	0.560(0.5)
est st.deviation	0.066	0.014	0.028	0.071
boot. st.deviation	0.083	0.018	0.034	0.068

Table 1: Maximum Pseudo-Likelihood estimation for Model 1 without covariate, true values are in brackets. Variance estimators is computed by our proposed method (est st.estimation) and by repetitions method (boot st.estimation)

545 All these results show the good performance of the Maximum Pseudo-
 546 Likelihood Estimator and its standard deviation. We have to note that the
 547 method is very easy to implement and the results are available almost in-
 548 stantly while it would be not the case with MCMC or Bayesian methods. It

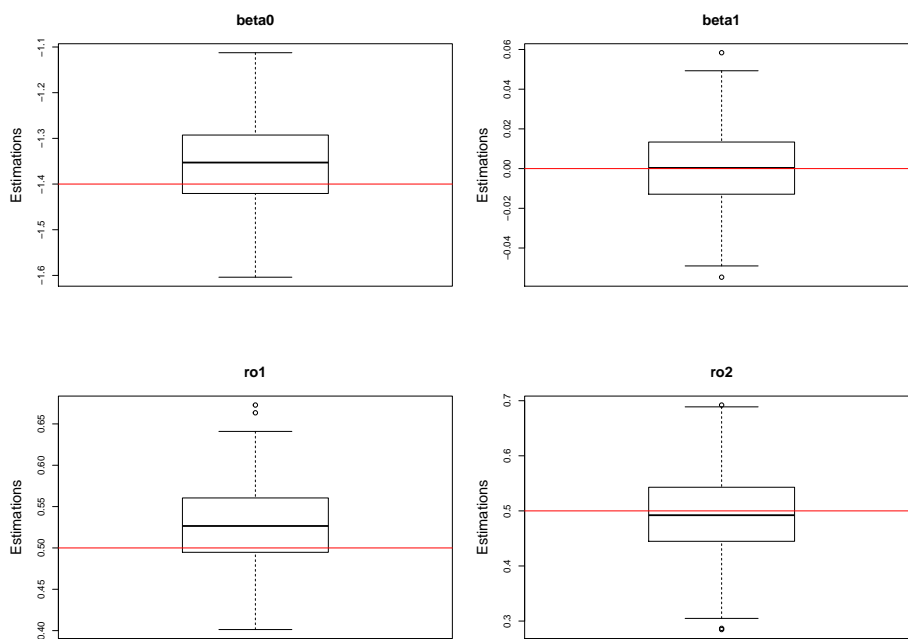


Figure 5: Box-plot of $B = 500$ estimations of the four parameters in Model 1. Red lines show the true values of parameters.

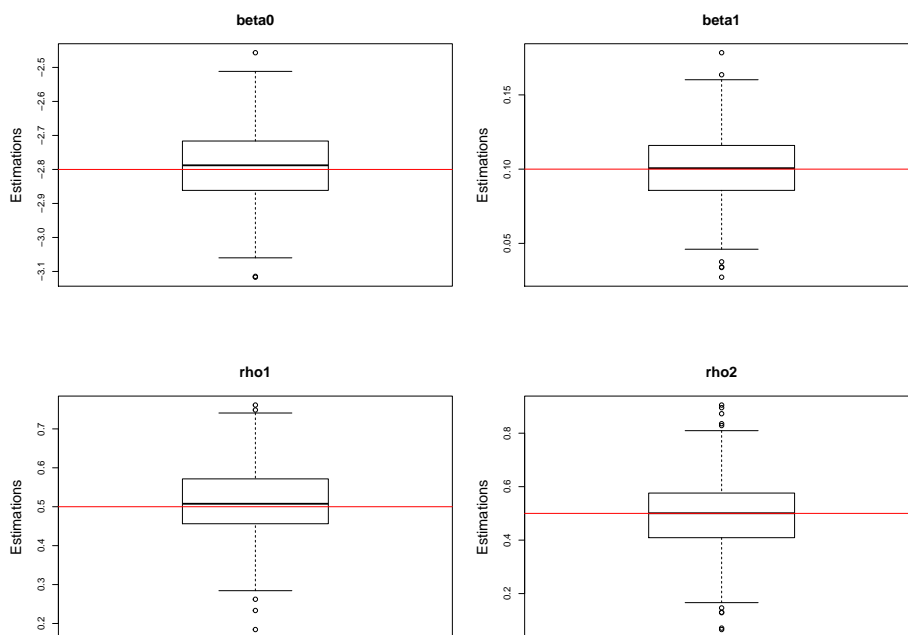


Figure 6: Box-plot of $B = 500$ estimations of the four parameters in Model 2. Red lines show the true values of parameters

	β_0	β_1	ρ_1	ρ_2
mean	-2.757(-2.8)	0.094(0.1)	0.488(0.5)	0.486(0.5)
st.deviation	0.097	0.021	0.042	0.130
boot. st.deviation	0.108	0.022	0.073	0.130

Table 2: Maximum Pseudo-Likelihood estimation for Model 2 with temporal covariate, true values are in brackets. Variance estimators is computed by our proposed method (est st.estimation) and by repetitions method (boot st.estimation).

549 should also be noted that we have made estimates on simulated data gener-
550 ated with different parameter values and that the method’s performance has
551 remained as good as above.

552 4.3.4. Model choice results

553 We first show the effects of different choices of neighbourhood graphs on
554 the estimation of the spatial auto-regression parameter ρ_1 of Equation 11.
555 For a given simulated dataset, we perform estimations of the parameters for
556 different graphs of neighbourhood. Results are shown in Table 3 that con-
557 firms the intuitive result that the estimation of the spatial autoregressive
558 parameter ρ_1 decreases with the number of neighbours. We have to notice
559 that this decrease is not proportional to the number of neighbours of each
560 points of the grid. Moreover estimation with a wrong neighbourhood struc-
561 ture does not affect the estimation of the other parameters of the models
562 (regression on the past parametrised by ρ_2 and on the covariate by β).

563 To see the good performance of the model choice rule, we simulated 500
564 independent realisations of different kinds of models under three different
565 neighbourhood structures, without (resp. with a temporal covariate) and for

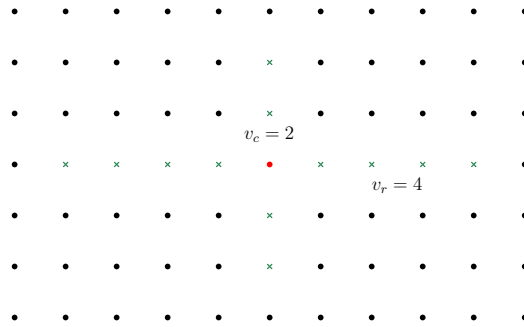


Figure 7: Structure of neighbourhood used for the simulations. Green crosses are the neighbours of the red point.

566 three different values of ρ_1 . We estimated the parameters under six differ-
 567 ent neighbours structures and choosed the model with the biggest Pseudo-
 568 Likelihood. Results are shown in Table 4 (resp. Table 5).

569 We see that the ability of the rule to detect the true neighbours structure
 570 is globally good. However, it has better performance if the model does not
 571 include covariates and almost perfect while $\rho_1 \geq 0.3$. Note that if the value of
 572 ρ_1 is low, it means that the spatial autocorrelation with the neighbours is low
 573 and thus it is not so important to infer the neighbourhood structure properly.
 574 The performance of the rule is degraded by the presence of a covariate and
 575 again when the relative weight of the spatial auto-correlation in the model is
 576 lower.

577 5. Application to local spread of plant diseases

578 5.1. Context and modelling

579 The spectrum of applications of spatio-temporal autologistic models is
 580 large. Thus for instance it was applied for the inference of networks in order

Model	v_r	v_c	β_0	β_1	ρ_1	ρ_2
1	1	1	-1.434	0.005	0.604	0.539
2	2	1	-1.470(-1.4)	0.003(0)	0.519(0.5)	0.560(0.5)
3	2	2	-1.467	0.003	0.420	0.560
4	3	1	-1.468	0.004	0.427	0.563
5	3	2	-1.475	0.004	0.368	0.563
6	3	3	-1.464	0.004	0.317	0.567

Table 3: Estimation by Pseudo-Likelihood in different models for neighbourhood. v_r (resp. v_c) is the number of neighbours on each side of a point on the row (resp. on the column). Estimations for the true model are in red and the true values of the parameters are in brackets.

581 to study competition in financial markets in Betancourt et al. (2018). In our
582 paper, the model was built with application to plant epidemiology. Indeed,
583 the purpose here is to analyse the spread of the esca grapevine trunk disease
584 over a 14-year period in a vineyard of Bordeaux in France, including 1 980
585 vines in a block of 30 rows and 66 columns. Esca is a grapevine trunk disease
586 that remains poorly understood but causing extensive damage in vineyard
587 worldwide and resulting in major economic losses (Bertsch et al., 2013; Mug-
588 nai et al., 1999). Grapevine esca is a complex dieback disease associated with
589 pathogenic fungi that degrade the woody part of the vine. It exhibits dis-
590 coloured foliar symptoms (Mugnai et al., 1999; Surico et al., 2008; Lecomte
591 et al., 2012). Leaf symptoms are erratic in the extent that they appear during
592 late spring or summer, however they can appear on year and not the following
593 one. The disease leads to a decrease in wine quality, and worse, to vine de-

True Model	Selected Model																	
	$\rho_1 = 0.3$						$\rho_1 = 0.4$						$\rho_1 = 0.5$					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	474	15	5	4	1	1	495	3	2	0	0	0	500	0	0	0	0	0
2	13	451	21	13	1	1	0	486	5	9	0	0	0	499	1	0	0	0
3	0	13	470	0	15	2	0	0	498	0	2	0	0	0	500	0	0	0

Table 4: Model choice by maximising the Pseudo-Likelihood for a model on a $20 * 20$ grid for 15 years without covariate. $\beta_0 = -1.4$, $\rho_2 = 0.5$ and three different values of ρ_1 .

True Model	Selected Model																	
	$\rho_1 = 0.3$						$\rho_1 = 0.4$						$\rho_1 = 0.5$					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	357	60	27	24	15	17	401	42	32	14	2	9	452	27	7	9	2	3
2	60	287	48	73	15	17	40	344	45	52	9	10	18	424	24	31	1	2
3	18	51	314	12	66	39	15	38	390	2	34	21	4	31	438	1	25	1

Table 5: Model choice by maximising the Pseudo-Likelihood for a model on a $20 * 20$ grid for 15 years with a covariate. $\beta_0 = -2.8$, $\beta_1 = 0.1$, $\rho_2 = 0.5$ and three different values of ρ_1 . $X_t = t$ for $t \leq 8$ and $16 - t$ for $8 \leq t \leq 15$.

594 cline and death at long term. In order to better understand the factors that
595 drive the esca spread, several studies based on spatio-temporal mathematical
596 modelling have been used. Some of them have focused on the contagiousness
597 of symptomatic vines in order to improve the prophylactic control of esca. In
598 Stefanini et al. (2000), a non- centered auto-logistic multinomial statistical
599 model with autoregression on the past and on the neighbourhood was used
600 to study the spatio-temporal dynamics of the esca at the scale of a vineyard.
601 However, the modelling and the inference method were not discussed. More
602 recently, a non-centered autologistic model with Bayesian inference was also
603 used to analyse a 17-year dataseries resulting from the monitoring of one

604 vineyard since planting (Zanzotto et al., 2013). In Li et al. (2016), we used
605 join count procedures to analyse aggregation and spread of the esca over an
606 eight-year period.

607 The lattice data came from the esca disease monitoring for 14 years be-
608 tween 2004 and 2017 in a commercial vineyard of the Bordeaux region (vine-
609 yard 13 in Li et al. (2016)), which was planted in 1989 with the cultivar
610 Cabernet Sauvignon (*Vitis vinifera*), and trained in accordance using the
611 Guyot trellising system. Within the surveyed plot, the distances between
612 rows and between vines were respectively 1.4 m and 1.2 m respectively. The
613 foliar esca expression of contiguous vines was recorded each year at the end
614 of August according to methods described in Li et al. (2016). The mean
615 annual esca prevalence was 9.5%, varying annually between 1.8 and 16.8%.

616 We propose to analyse our data with the objectives to understand: (i)
617 the effect of the status of a vine (symptomatic or not) the year preceding
618 the observation; (ii) the effect of the frequencies of infected plants among the
619 vine’s neighbours the year preceding the observation on the occurrence of the
620 symptom for the given vine. Moreover we want to capture the instantaneous
621 spatial correlation between vines in the same year. We want to choose the
622 best neighbourhood structure for the previous year’s effect and the instantane-
623 ous effect. We are clearly in a context of selecting the best suited models
624 in terms of past and instantaneous neighbourhood structures.

625 According to physiopathologists, the neighbourhood structures are el-
626 lipses shaped (see Figure 8). Indeed, vines are row planted leading to pos-
627 sible anisotropy. The neighbourhood of a vine i located in s_i is given by all
628 vines included in a ellipse defined by its semi-major axis and semi-major axis.

629 These quantities are denoted v_r and v_c for the instantaneous neighbourhood
630 \mathcal{N}_i (resp. p_r and p_c for the past one \mathcal{N}_i^p). The model is the following. For
631 a vine i , located in s_i at time t , the probability to present the symptom
632 according to the history of the vineyards and the neighbourhood is given by:

$$\text{logit}(p_{it}) = \beta_0 + \beta_1 \left(\sum_{j \in \mathcal{N}_i^p} Z_{j,t-1} \right) + \rho_1 \left(\sum_{j \in \mathcal{N}_i} Z_{j,t}^{**} \right) + \rho_2 Z_{i,t-1} \quad (22)$$

633 Twenty five neighbourhood structures were tested for the two auto-regressions
634 (instantaneous and on the past) corresponding to 625 different models. A
635 neighbourhood was defined by an ellipse around the vine with a number of
636 neighbours v_r (resp. v_c) in the direction of the row (resp. column). v_r and v_c
637 vary independently from 1 to 5 neighbours leading to 25 possible structures.
638 We also defined p_l and p_c the corresponding parameters for the neighbour-
639 hood concerning the past.

640 5.2. Results

641 The most important result is that the value of the estimation of coefficient
642 ρ_2 is very robust whatever the chosen neighbourhood structure, ($\rho_2 = 2.28$).
643 It means that the risk for a vine that has already expressed the symptoms
644 the previous year to express them again is multiplied by $\exp(2.28) = 9.7$
645 comparing to the same risk for a vine without expression the previous year.
646 This result corroborates that of Guérin-Dubrana et al. (2013). These authors
647 have shown that a declining vine has expressed esca symptoms on average
648 two to three times the years before. The re-expression of leaf symptoms is
649 frequent and reveals the advanced state of internal infection (Maher et al.,
650 2012).

651 With respect to other terms of the regression, another important result is
652 that the choice of the neighbourhood for the instantaneous correlation is more
653 important than the one for the past regression. The 50 best models according
654 to the PL are those with the instantaneous neighbourhood defined by an
655 ellipse (that could also be a circle) with radius of 5 in a row direction and
656 4 in the other one. This effect is more important than the effect of the past
657 that covers all the possibilities for the neighbourhood. The interpretation of
658 this instantaneous autocorrelation is likely to be found in local environmental
659 effects such as soil properties. This effect is not due to a spread of the illness.
660 Note that it captures a little anisotropy showing more effect along the row,
661 that is a standard result in vineyards because the vines are nursed along the
662 row. The estimated coefficients and associated standard deviation for the
663 best model are given in Table Table 6.

v_r	v_c	p_r	p_c	β_0	β_1	ρ_1	ρ_2
5	4	1	1	-3.04 (0.035)	0.178 (0.034)	0.135 (0.006)	2.28 (0.05)

Table 6: Instantaneous and past neighbour structures and estimated coefficients for the best model on the real data. Standard deviation of the parameters are in bracket.

664 We have already commented the estimation of autoregression ρ_2 . The
665 value of the one for β_0 indicates a spontaneous level of infection equal to
666 $\exp(-3,04) = 0.05$. β_1 is the coefficient of regression that quantifies the
667 spread of the illness, like in the logistic model, when the level of the illness is
668 still low, the presence of a one more vine with symptoms in the neighbour-
669 hood at a time multiplies by $\exp(0.178) = 1.19$ the risk to present symptoms
670 the following year. This result shows the small role of the neighbouring

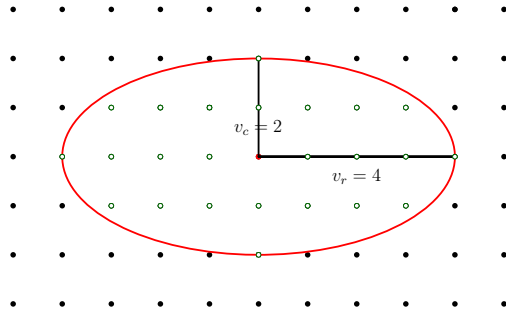


Figure 8: Structure of neighbourhood used for esca data. Green circles are the neighbours of the red point.

671 symptomatic vines on the disease occurrences. It confirms those of Li et al.
 672 (2016) which suggested a limited potential of secondary local spread from
 673 neighbouring symptomatic vines.

674 6. Discussion and conclusion

675 In this paper, we have proposed a new spatio-temporal model for the
 676 study of binary data evolving with time on a lattice. At each time t , the spa-
 677 tial covariates are centered at its expected value which depends on the value
 678 of the covariates and also on the values of the field in the past. Simulations
 679 studies in section 3, show the interest of this new centering to demonstrate
 680 that what is naturally taken to represent large-scale model structure in the
 681 traditional parameterization of an autologistic model does not necessarily re-
 682 flect marginal structure in the data it generates. This new model allows the
 683 practitioner to make a good interpretation of the spatial regression parameter
 684 that was not possible in previous models. We have shown the ability of the
 685 Maximum Pseudo-Likelihood Estimator to infer quickly the value of the pa-

686 rameters. Maximising Pseudo-Likelihood also allows us to choose efficiently
687 between models with different with different neighbourhood structures. Even
688 if the whole spatio-temporal joint distribution of the process is not proved
689 to exist, we still discuss the existence of the spatial joint law of the process
690 at any time given the covariates and the past of the process. This approach
691 seems coherent with the recursive construction of the process along the time.

692 The model and the method presented in this paper are very suitable and
693 efficient for modelling the evolution of an illness on a lattice taking into
694 account covariates and spatial auto-correlation. They allow to measure and
695 quantify effects of the neighbourhood in the past on the occurrence of the
696 illness at a given time. It should be noted that although we have thought this
697 model by thinking of a repartition of plants on a lattice, it could be applied
698 to other spatial repartitions provided that the structure of neighbours is well
699 defined. Moreover, for reasons of sparsity we have chosen auto-regression
700 coefficients ρ_1 and ρ_2 which do not depend on the site s_i or the neighbour
701 s_j but the model can be easily complicated depending on the purpose of
702 the modelling (for instance we can distinguish several kind of neighbours).
703 The purpose of the application was to show that our methodology is easy
704 to implement, the data at our disposal were very simple without covariates
705 other than data from the past and the state of the neighbours. But the next
706 step is the acquisition of spatial or spatio-temporal covariates (soil properties,
707 vigour of the plant, water stress..) to better understand their effect on leaf
708 symptoms. Temporal covariates such as weather information would also be
709 interesting to incorporate. We also plan to develop a free software package
710 for the R software that would be available for the analysis of spatio-temporal

711 binary data.

712 **Acknowledgements**

713 We wish to thank Avner Bar Hen and Cécile Hardouin for precious discus-
714 sions at the beginning of this work. We wish to acknowledge the vine-grower
715 who participated in this study and also Sylvie Bastien and David Morais for
716 their excellent technical assistance. This research was supported by Bordeaux
717 Sciences Agro, the Regional Council of Aquitaine, the JEAN POUPELAIN
718 Foundation, the French Ministry of Agriculture and the Food-processing in-
719 dustry and Forest (CASDAR V1303).

720 **References**

- 721 M. A. Beaumont, W. Zhang, and D. J. Balding. Approximate bayesian
722 computation in population genetics. *Genetics*, 162(4):2025–2035, 2002.
- 723 C. Bertsch, M. Ramírez-Suero, M. Magnin-Robert, P. Larignon, J. Chong,
724 E. Abou-Mansour, A. Spagnolo, C. Clément, and F. Fontaine. Grapevine
725 trunk diseases: complex and still poorly understood. *Plant pathology*, 62
726 (2):243–265, 2013.
- 727 J. Besag. Spatial interaction and the statistical analysis of lattice systems.
728 *Journal of the Royal Statistical Society. Series B (Methodological)*, pages
729 192–236, 1974.
- 730 B. Betancourt, A. Rodríguez, and N. Boyd. Investigating competition in
731 financial markets: a sparse autologistic model for dynamic network data.
732 *Journal of Applied Statistics*, 45(7):1157–1172, 2018.

- 733 P. C. Caragea and M. S. Kaiser. Autologistic models with interpretable pa-
734 rameters. *Journal of agricultural, biological, and environmental statistics*,
735 14(3):281–300, 2009.
- 736 N. Cressie. *Statistics for Spatial Data: Wiley Series in Probability and Statis-*
737 *tics*. Wiley-Interscience New York, 1993.
- 738 C. Gaetan and X. Guyon. *Modélisation et statistique spatiales*, volume 63.
739 Springer, 2008.
- 740 A. Grelaud, C. P. Robert, J.-M. Marin, F. Rodolphe, J.-F. Taly, et al. Abc
741 likelihood-free methods for model choice in gibbs random fields. *Bayesian*
742 *Analysis*, 4(2):317–335, 2009.
- 743 G. R. Grimmett. A theorem about random fields. *Bulletin of the London*
744 *Mathematical Society*, 5(1):81–84, 1973.
- 745 L. Guérin-Dubrana, A. Labenne, J. C. Labrousse, S. Bastien, R. Patrice, and
746 A. Gégout-Petit. Statistical analysis of grapevine mortality associated with
747 esca or eutypa dieback foliar expression. *Phytopathologia Mediterranea*, 52
748 (2):276–288, 2013.
- 749 M. L. Gumpertz, J. M. Graham, and J. B. Ristaino. Autologistic model
750 of spatial pattern of phytophthora epidemic in bell pepper: effects of soil
751 variables on disease presence. *Journal of Agricultural, Biological, and En-*
752 *vironmental Statistics*, pages 131–156, 1997.
- 753 X. Guyon. *Random fields on a network: modeling, statistics, and applica-*
754 *tions*. Springer Science & Business Media, 1995.

- 755 X. Guyon and C. Hardouin. Markov chain markov field dynamics: models
756 and statistics. *Statistics: A Journal of Theoretical and Applied Statistics*,
757 36(4):339–363, 2002.
- 758 F. W. Huffer and H. Wu. Markov chain monte carlo for autologistic regression
759 models with application to the distribution of plant species. *Biometrics*,
760 pages 509–524, 1998.
- 761 J. Hughes. `ngspatial`: A package for fitting the centered autologistic and
762 sparse spatial generalized linear mixed models for areal data. *R Journal*,
763 6(2), 2014.
- 764 J. Hughes, M. Haran, and P. C. Caragea. Autologistic models for binary
765 data on a lattice. *Environmetrics*, 22(7):857–871, 2011.
- 766 M. S. Kaiser and N. Cressie. The construction of multivariate distributions
767 from markov random fields. *Journal of Multivariate Analysis*, 73(2):199–
768 220, 2000.
- 769 W. S. Kendall. Notes on perfect simulation. *Markov Chain Monte Carlo:*
770 *Innovations and Applications*, 7, 2005.
- 771 P. Lecomte, G. Darrietort, J.-M. Liminana, G. Comont, A. Muruamendi-
772 araz, F.-J. Legorburu, E. Choueiri, F. Jreijiri, R. El Amil, and M. Fermaud.
773 New insights into esca of grapevine: the development of foliar symptoms
774 and their association with xylem discoloration. *Plant Disease*, 96(7):924–
775 934, 2012.

- 776 S. Li, F. Bonneu, J. Chadoeuf, D. Picart, A. Gégout-Petit, and L. Guerin-
777 Dubrana. Spatial and temporal pattern analyses of esca grapevine disease
778 in vineyards in france. *Phytopathology*, 107(1):59–69, 2016.
- 779 N. Maher, J. Piot, S. Bastien, J. Vallance, P. Rey, and L. Guérin-Dubrana.
780 Wood necrosis in esca-affected vines: types, relationships and possible links
781 with foliar symptom expression. *OENO One*, 46(1):15–27, 2012.
- 782 P. Marjoram, J. Molitor, V. Plagnol, and S. Tavaré. Markov chain monte
783 carlo without likelihoods. *Proceedings of the National Academy of Sciences*,
784 100(26):15324–15328, 2003.
- 785 L. Mugnai, A. Graniti, G. Surico, et al. Esca (black measles) and brown
786 wood-streaking: two old and elusive diseases of grapevines. *Plant disease*,
787 83(5):404–418, 1999.
- 788 J. K. Pritchard, M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman.
789 Population growth of human y chromosomes: a study of y chromosome
790 microsatellites. *Molecular biology and evolution*, 16(12):1791–1798, 1999.
- 791 F. M. Stefanini, G. Surico, G. Marchi, et al. Longitudinal analysis of symp-
792 tom expression in grapevines affected by esca. *Phytopathologia Mediter-
793 ranea*, 39(1):225–231, 2000.
- 794 G. Surico, L. Mugnai, and G. Marchi. The esca disease complex. In *Integrated
795 management of diseases caused by fungi, phytoplasma and bacteria*, pages
796 119–136. Springer, 2008.

- 797 Z. Wang and Y. Zheng. Analysis of binary data via a centered spatial-
798 temporal autologistic regression model. *Environmental and Ecological*
799 *Statistics*, 20(1):37–57, 2013.
- 800 M. Wolters. Better autologistic regression. *Front. Appl. Math. Stat.*, pages
801 3–24, 2017.
- 802 A. Zanzotto, M. Gardiman, S. Serra, D. Bellotto, F. Bruno, F. Greco, and
803 C. Trivisano. The spatiotemporal spread of esca disease in a cabernet
804 sauvignon vineyard: a statistical analysis of field data. *Plant Pathology*,
805 62(6):1205–1213, 2013.
- 806 Y. Zheng and J. Zhu. Markov chain monte carlo for a spatial-temporal
807 autologistic regression model. *Journal of Computational and Graphical*
808 *Statistics*, 17(1), 2008.
- 809 J. Zhu and Y. Zheng. Autologistic regression models for spatio-temporal
810 binary data. *Handbook of Discrete-Valued Time Series*, pages 367–387,
811 2016.
- 812 J. Zhu, H.-C. Huang, and J. Wu. Modeling spatial-temporal binary data
813 using markov random fields. *Journal of Agricultural, Biological, and En-*
814 *vironmental Statistics*, 10(2):212–225, 2005.
- 815 J. Zhu, Y. Zheng, A. L. Carroll, and B. H. Aukema. Autologistic regression
816 analysis of spatial-temporal binary data via monte carlo maximum likeli-
817 hood. *Journal of agricultural, biological, and environmental statistics*, 13
818 (1):84–98, 2008.