DEM Discussion Papers

Department of Economics and Management

2013/12

Approximate Maximum Likelihood Estimation of the Autologistic Model

Marco Bee, Diego Giuliani, Giuseppe Espa

ISSN 2282-2801 DEM discussion papers [online]

UNIVERSITÀ DEGLI STUDI
DI TRENTO

DEM Discussion Paper

ISSN 2282-2801 DEM Discussion Papers [online]
Università degli Studi di Trento

Department of Economics and Management, University of Trento, Italy.

Editors

Luciano ANDREOZZI luciano.andreozzi@unitn.it

Roberto GABRIELE roberto.gabriele@unitn.it

Technical officer

Marco TECILLA marco.tecilla@unitn.it

Guidelines for authors
Papers may be written in Italian or in English. Faculty members of the Department must submit to one of the editors in pdf format. Management papers should be submitted to R. Gabriele. Economics Papers should be submitted to L. Andreozzi. External members should indicate an internal faculty member that acts as a referee of the paper.

Typesetting rules:

1. papers must contain a first page with title, authors, abstract, keywords and codes. Page numbering starts from the following page;
2. a template is available upon request from the managing editors.

Approximate Maximum Likelihood Estimation of the Autologistic Model

Marco Bee
Department of Economics and Management, University of Trento

Giuseppe Espa
Department of Economics and Management, University of Trento

Diego Giuliani
Department of Economics and Management, University of Trento

November 11, 2013

1 Introduction

Point estimation of spatial models is well known to be a difficult issue. One general reason is that, from a probabilistic point of view, a spatial model is a random field, typically characterized by a complex dependence structure, of which only a single realization is available for estimation. More specifically, what makes estimation overly complicated is the computational intractability of the normalizing constant of the joint density, even for moderate lattice sizes. This problem is particularly serious for Maximum Likelihood Estimation (MLE) procedures, as the normalization constant depends on the parameters of the model, and thus cannot be ignored in the maximization of the likelihood function.

On the other hand, the conditional distributions at single sites, given values at neighboring locations, usually admit simple representations. Exploiting this idea, Besag (1975) developed the so-called Maximum Pseudo-Likelihood Estimation (MPLE) method, which is still very popular in practical applications.

MPLE is the earliest and simplest approach to estimation of the parameters of the most important model for spatially dependent binary random variables, i.e. the autologistic model. As will be made clear in Sect. 2, MPLE is based on the pseudo-likelihood function, defined by the product of the conditional distributions at all locations given the values at neighboring locations. To obtain the estimators, the pseudo-likelihood is maximized with respect to the parameters as if it were a likelihood, i.e. by means of standard logistic regression estimation techniques. However, the two functions coincide exactly if the observations at different locations are independent,

a condition that is only satisfied in trivial cases. As a result, PMLEs are consistent and asymptotically normal (Geman and Graffigne, 1987; Comets, 1992; Guyon and Künsch, 1992) but not efficient, with a loss of efficiency positively related to the (absolute) value of the spatial dependence parameter.

Given that the difficulties are mostly caused by the normalization constant, research has focused on methods of evaluating it, possibly in an approximate way. The pioneering work by Ogata and Tanemura (1984) develops various techniques for approximating the likelihood function. Moyeed and Baddeley (1991) use an iterative stochastic approximation approach. Huffer and Wu (1998) employ Markov Chain Monte Carlo methods. Gu and Zhu (2001) compute MLEs by combining Markov Chain Monte Carlo and stochastic approximation methods. Huang and Ogata (2002) propose a generalization of Maximum Pseudo-Likelihood. Friel and Pettitt (2004) develop a method for exact MLE of the autologistic model that is simulation-free for lattices with smallest row or column not larger than 10 and can be extended to larger sizes by means of Monte Carlo techniques. Hughes et al. (2011) propose an approach, called Monte Carlo maximum likelihood, that maximizes numerically an approximation of the likelihood function. Finally, Wang and Zheng (2013) use expectation-maximization pseudo-likelihood and Monte Carlo expectation-maximization likelihood.

Recently, MLE for intractable likelihoods (see Murray et al., 2006, for a useful classification) has received some attention in the literature, mostly because various simulation-based approaches allow to approximate the likelihood in these setups. In this paper we propose to extend to the autologistic model the Approximate MLE (AMLE) method developed by Rubio and Johansen (2013). In short (see Sect. 3 below for details), the most appealing feature of this approach is that it allows to obtain MLEs without performing a formal maximization of the likelihood function. Thus, it is the ideal candidate for the computation of MLE of spatial models, because the problem of evaluating the normalizing constant is bypassed. Moreover, as long as we are able to sample the model to be estimated, its implementation is straightforward, and can be extended to other spatial models with essentially no modifications. Finally, as the simulation of the autologistic model is usually based on the Metropolis algorithm, which generates in a sequential manner a single random variable for each location, the method works well even for very large dimensions, as the only limit is the machine's physical memory. The two last features are non-negligible strengths with respect to existing approaches, often characterized by involved implementation and/or poor performances for large lattice sizes.

The rest of this paper is organized as follows: Section 2 reviews the autologistic model, Section 3 introduces the AMLE methodology and develops its implementation to the autologistic model, Section 4 presents first the results of extensive Monte Carlo experiments aiming at a comparison

of AMLE, MLE and MPLE in terms of Mean Squared Error and then a real-data application. Finally, Section 5 concludes.

2 The autologistic model

By spatial model we mean a statistical model for a spatial pattern of data $\boldsymbol{y} = (y_i \in A \subset \mathbb{R}^s, i = 1, \ldots, C$, where $\mathbb{R}^s$ is the s-dimensional Euclidean space) having density

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = \frac{e^{-Q(\boldsymbol{y},\boldsymbol{\theta})}}{Z(\boldsymbol{\theta})}. \tag{1}$$

The normalizing constant is given by $Z(\boldsymbol{\theta}) = \int_{A^C} e^{-Q(\boldsymbol{y},\boldsymbol{\theta})}\mu(d\boldsymbol{y})$, where $\mu(d\boldsymbol{y})$ is Dirac's delta measure $\delta_{\boldsymbol{y}}(d\boldsymbol{y})$ in the discrete case and $d\boldsymbol{y}$ in the continuous case.

The autologistic model is a special case of (1). The joint distribution of $\tilde{\boldsymbol{y}} = (\tilde{y}_1, \ldots, \tilde{y}_L)'$, where $\tilde{\boldsymbol{y}} = \text{vec}(\tilde{\boldsymbol{Y}})$, $\tilde{\boldsymbol{Y}} = (\tilde{y}_{i,j})$ $(i = 1, \ldots, N; j = 1, \ldots, M$, $L = MN)$ and vec is the operator that stacks the columns of a matrix on top of one another, is given by (Strauss, 1992; Casella and Robert, 2004, Example 5.8; Arbia, 2006, Sect. 2.4.2.3)

$$p(\tilde{\boldsymbol{y}}) = \frac{1}{Z(J,H)} \exp\left\{ -J \sum_{(i,j)\in\mathcal{N}} \tilde{y}_i\tilde{y}_j - H\sum_i \tilde{y}_i \right\}, \tag{2}$$

where $J \in \mathbb{R}$ and $H \in \mathbb{R}$ are parameters, $\mathcal{N}$ is a prespecified neighborhood equivalence relation, $\tilde{y}_i \in \{-1, 1\}$ and $Z(J,H) = \sum_{i,j} \exp\{-J\sum_{(i,j)\in\mathcal{N}} \tilde{y}_i\tilde{y}_j - H\sum_i \tilde{y}_i\}$ is the normalization constant. The model is of paramount importance in statistical mechanics, where it is known as Ising model of ferromagnetism (Cipra, 1987).

Let now $y_i = (\tilde{y}_i + 1)/2$. The conditional representation of (2) is given by

$$P(y_i = 1|y_j, j \neq i) = \frac{\exp\{2(H + J\sum_j y_j)\}}{1 + \exp\{2(H + J\sum_j y_j)\}} = \frac{\exp\{\alpha + \beta\sum_j y_j\}}{1 + \exp\{\alpha + \beta\sum_j y_j\}}. \tag{3}$$

If the y_js were independent, this would be a logistic regression model. Unfortunately, as they are dependent, MLE must be based on the joint distribution (2), whose normalizing constant $Z(J,H)$ becomes rapidly intractable from the computational point of view as L increases.

3 AMLE of spatial models

3.1 A review of AMLE

Given a sample $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) \in \mathbb{R}^{q \times n}$ from a distribution with cumulative distribution function $F_{\boldsymbol{Y}}(\boldsymbol{y}; \boldsymbol{\theta})$, let the likelihood function be denoted by $L(\boldsymbol{\theta}; \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$, where $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p$ is a vector of parameters. Assume for a moment a Bayesian setup, such that $\pi(\boldsymbol{\theta})$ is the prior distribution of $\boldsymbol{\theta}$ and $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ is the posterior, given by

$$\pi(\theta|\boldsymbol{y}) = \frac{f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\boldsymbol{\Theta}} f(\boldsymbol{y}|\boldsymbol{t})\pi(\boldsymbol{t})d\boldsymbol{t}}. \tag{4}$$

Suppose that we can construct the following approximation of the likelihood function:

$$\hat{f}_\epsilon(\boldsymbol{y}|\boldsymbol{\theta}) = \int_{\mathbb{R}^n} K_\epsilon(\boldsymbol{y}|\boldsymbol{z}) f(\boldsymbol{z}|\boldsymbol{\theta}) d\boldsymbol{z}, \tag{5}$$

where $K_\epsilon(\boldsymbol{y}|\boldsymbol{z})$ is a normalized Markov kernel and ϵ is a scale parameter. Plugging (5) into (4) we obtain an approximation of the posterior:

$$\hat{\pi}_\epsilon(\theta|\boldsymbol{y}) = \frac{\hat{f}_\epsilon(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\boldsymbol{\Theta}} \hat{f}_\epsilon(\boldsymbol{y}|\boldsymbol{t})\pi(\boldsymbol{t})d\boldsymbol{t}}.$$

Note that, for a uniform prior on a suitable set $\boldsymbol{D} \subset \mathbb{R}^p$, the maximization of the likelihood function and the maximization of the posterior density are equivalent.

The typical kernel $K_\epsilon(\boldsymbol{y}|\boldsymbol{z})$ is defined as follows:

$$K_\epsilon(\boldsymbol{y}|\boldsymbol{z}) \propto \begin{cases} 1 & \rho(\eta(\boldsymbol{y}), \eta(\boldsymbol{z})) < \epsilon, \\ 0 & \text{otherwise}, \end{cases} \tag{6}$$

where $\eta : \mathbb{R}^{q \times n} \to \mathbb{R}^l$ is a summary statistic, $\rho : \mathbb{R}^{q \times n} \times \mathbb{R}^{q \times n} \to \mathbb{R}^+$ is a metric and $\epsilon > 0$. The simplest version uses $\eta(\boldsymbol{y}) = \boldsymbol{y}$, in which case we obtain the Approximate Bayesian Computation (ABC) algorithm introduced by Pritchard et al. (1999). For reasons that will be outlined below, when employing this technique in the AMLE setup we shall use the sufficient statistics of the model, if available.

The preceding discussion motivates the following algorithm:

Algorithm 1 *(AMLE)*

1. *Obtain a sample* $\boldsymbol{\theta}^*_\epsilon = (\boldsymbol{\theta}^*_{\epsilon,1}, \ldots, \boldsymbol{\theta}^*_{\epsilon,m})'$ *from the approximate posterior* $\hat{\pi}_\epsilon(\theta|\boldsymbol{y})$*;* m *is commonly called ABC sample size;*

2. *Use this sample to construct a nonparametric estimator* $\hat{\phi}$ *of the density* $\hat{\pi}_\epsilon(\theta|\boldsymbol{y})$*;*

3. *Compute the maximum of* $\hat{\phi}$, $\tilde{\boldsymbol{\theta}}_{m,\epsilon}$. *This is an approximation of the MLE* $\hat{\boldsymbol{\theta}}$.

Step 1 can be carried out using any algorithm, as it is not required that the sample be independent. The most common solution is the simple ABC algorithm.

Algorithm 2 *(simple ABC)*

1. *Simulate* $\boldsymbol{\theta}'$ *from the prior distribution* $\pi(\cdot)$;
2. *Generate* $\boldsymbol{y} = (y_1, \ldots, y_n)'$ *from* $f(\cdot|\boldsymbol{\theta}')$;
3. *Accept* $\boldsymbol{\theta}'$ *with probability* $\propto K_\epsilon(\boldsymbol{y}|\boldsymbol{z})$, *otherwise return to Step 1.*

Rubio and Johansen (2013) study the asymptotic properties of the estimator, which are particularly important because they can give some insight into the choice of the input parameters. The crucial result is that, under a mild condition about $K_\epsilon(\boldsymbol{x}|\boldsymbol{y})$, $\hat{\pi}_\epsilon(\boldsymbol{\theta}|\boldsymbol{x})$ converges pointwise to $\pi(\boldsymbol{\theta}|\boldsymbol{x})$ as $\epsilon \to 0$, for any $\boldsymbol{\theta} \in \boldsymbol{D}$.

A corollary of the preceding result suggests how to choose the summary statistic: if η is a sufficient statistic for $\boldsymbol{\theta}$, the ABC approximation converges pointwise to the posterior distribution. It is therefore clear that, if a sufficient statistic is available for $\boldsymbol{\theta}$, one should use it in the algorithm.

Finally, under the additional condition of equicontinuity of $\hat{\pi}_\epsilon(\cdot|\boldsymbol{y})$ on $\boldsymbol{D}$, it is possible to show that $\lim_{\epsilon \to 0} \hat{\pi}_\epsilon(\tilde{\boldsymbol{\theta}}|\boldsymbol{y}) = \pi(\tilde{\boldsymbol{\theta}}|\boldsymbol{y})$, where $\tilde{\boldsymbol{\theta}}$ is the unique maximizer of $\pi(\cdot|\boldsymbol{y})$.

Suppose now to have a simple random sample $\boldsymbol{\theta}^*_\epsilon = (\boldsymbol{\theta}^*_{\epsilon,1}, \ldots, \boldsymbol{\theta}^*_{\epsilon,m})'$ from the approximate posterior $\hat{\pi}_\epsilon(\cdot|\boldsymbol{x})$ with mode $\tilde{\boldsymbol{\theta}}_\epsilon$ and an estimator $\tilde{\boldsymbol{\theta}}_{m,\epsilon}$ of $\tilde{\boldsymbol{\theta}}_\epsilon$ obtained from $\boldsymbol{\theta}^*_\epsilon$ and such that $\tilde{\boldsymbol{\theta}}_{m,\epsilon} \to \tilde{\boldsymbol{\theta}}_\epsilon$ almost surely when $m \to \infty$. From these results it follows that, for any $\gamma > 0$, there exists $\epsilon > 0$ such that $\lim_{m \to \infty} |\hat{\pi}_\epsilon(\tilde{\boldsymbol{\theta}}_{m,\epsilon}|\boldsymbol{y}) - \pi(\tilde{\boldsymbol{\theta}}|\boldsymbol{y})| \leq \gamma$ almost surely.

It is worth noting that $\tilde{\boldsymbol{\theta}}_{m,\epsilon}$ is an approximation of the MLE, with asymptotic variance related to the numerical value of ϵ: in particular, depending on ϵ, the estimator may be more or less efficient than the MLE. Ionides (2005) has indeed shown that the maximization of a smoothed approximation of the likelihood such as (5) may be preferable to the maximization of the likelihood itself.

In step 2, Rubio and Johansen (2013) suggest to use Kernel Density Estimation (KDE). The performance of KDE deteriorates as the dimension of $\boldsymbol{\theta}$ increases. Intractable likelihoods often have few unknown parameters, as is the case of the autologistic likelihood considered in this paper, which has only two unknown parameters. Nevertheless, an increase in the number of parameters is likely to cause higher computational costs.

3.2 AMLE of the autologistic model

Let $\boldsymbol{Y} = (y_{1,1}, \ldots, y_{N,M})$ be the $N \times M$ matrix containing the realization of an autologistic model. From now on, we will use the notation $\boldsymbol{y} = \text{vec}(y_{1,1}, \ldots, y_{N,M})$. The implementation of AMLE requires specific versions of Algorithms 1 and 2 and appropriate choices of the numerical values of the input parameters. We consider Algorithm 2 first.

One of the assumptions of AMLE is that the prior is uniform. When the support is a compact subset of $\mathbb{R}^p$ it may make sense to use the full support, but, when the domain is the p-dimensional Euclidean space, we need starting values and intervals around them, sufficiently wide to contain the true parameter values. If the distribution of the initial parameters is known, the widths can be approximately computed analytically, otherwise Monte Carlo techniques can be easily implemented.

For the autologistic model, the natural choice of the starting value is the MPLE estimator $\hat{\boldsymbol{\theta}}^{MPLE}$. We then sample candidate values of the parameters from the $U[\hat{\theta}_i^{MPLE} + d_i, \hat{\theta}_i^{MPLE} + u_i]$ distributions, where $d_i \leq 0$ and $u_i \geq 0$ $(i = 1, 2)$ are such that the intervals $[\hat{\theta}_i^{MPLE} + d_i, \hat{\theta}_i^{MPLE} + u_i]$ contain with very high probability the true values of the parameters.

The choice of d_i and u_i can be based on a simulation experiment, consisting in simulating B times the autologistic model and computing 99.9% confidence intervals. The parameters d_i and u_i are then given by the bounds of the confidence intervals. In the present context, we use $B = 100$.

We know from theory that, if η in (6) is a sufficient statistic for $\boldsymbol{\theta}$, the AMLE algorithm has favorable convergence properties. Sufficient statistics for the autologistic model (2) are given by $\sum_i \tilde{y}_i$ and $\sum_{(i,j)\in\mathcal{N}} \tilde{y}_{ij}$. In terms of (3), they can be conveniently rewritten as $S_1 = \sum_{i=1}^{L} y_i$ and $S_2 = \sum_{i=1}^{M-1} \sum_{j=1}^{N} \mathbb{I}_{\{y_{i,j}=y_{i+1,j}\}} + \sum_{i=1}^{M} \sum_{j=1}^{N-1} \mathbb{I}_{\{y_{i,j}=y_{i,j+1}\}}$, where $\mathbb{I}$ is the indicator function.

We are now in a position to detail Algorithm 2 in the autologistic case:

Algorithm 3

1. *Simulate* $\alpha' \sim U[\hat{\alpha}^{MPLE}+d_1, \hat{\alpha}^{MPLE}+u_1]$, $\beta' \sim U[\hat{\beta}^{MPLE}+d_2, \hat{\beta}^{MPLE}+u_2]$*;*

2. *Simulate a realization* $\boldsymbol{y}'$ *of the autologistic model with parameters* α' *and* β'*;*

3. *Compute the sufficient statistics* S_1' *and* S_2' *using* $\boldsymbol{y}'$*. If the Euclidean distance* $||\boldsymbol{S}-\boldsymbol{S}'|| = \sqrt{(S_1 - S_1')^2 + (S_2 - S_2')^2}$ *is smaller than* ϵ*, where* $\boldsymbol{S} = (S_1, S_2)$*,* $\boldsymbol{S}' = (S_1', S_2')$ *and* S_1 *and* S_2 *are the sufficient statistics computed with the observed data, accept* α' *and* β'*; otherwise, return to Step 1.*

Table 1: Bias, variance and MSE of classical MLE and AMLE of p for $X \sim \text{Bin}(n, p)$ with $n = 100$ and $p = 0.5$. The results shown are averages of 100 Monte Carlo replications. The remaining parameters are $m = 1000$, $D = [0, 1]$.

	MLE	AMLE		
		$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.5$
bias	$4.10 \cdot 10^{-3}$	$2.83 \cdot 10^{-4}$	$1.72 \cdot 10^{-3}$	$1.56 \cdot 10^{-2}$
variance	$2.23 \cdot 10^{-3}$	$5.49 \cdot 10^{-5}$	$2.98 \cdot 10^{-4}$	$5.64 \cdot 10^{-2}$
MSE	$2.25 \cdot 10^{-3}$	$5.50 \cdot 10^{-5}$	$3.01 \cdot 10^{-4}$	$5.66 \cdot 10^{-2}$
rejection rate		0.990	0.792	0.020

The simulation of the autologistic model at Step 2 is carried out by means of the Metropolis algorithm (Metropolis et al., 1953; Gu and Zhu, 2001, p. 346; Huang and Ogata, 2002, p. 6) with 500 Monte Carlo steps. Step 3 is based on standard kernel density estimation methods: analogously to Rubio and Johansen (2013), we use the `kde` command of the `ks` package of `R` to compute the nonparametric estimator $\hat{\phi}$ of the density $\hat{\pi}_\epsilon(\theta|\boldsymbol{y})$. Finally, we compute $\hat{\boldsymbol{\theta}} = \arg\max(\hat{\phi})$.

Besides $\boldsymbol{D}$, in order to use the algorithm one has to set the parameters m and ϵ. The choice is problem-dependent and will be examined thoroughly in the next section. Before focusing on the autologistic model, we show the results of a toy simulation experiment that illustrates how the variance of the estimator depends on ϵ.

Example 1. We simulate $B = 100$ times a binomial random variable $X \sim \text{Bin}(n, p)$ with $n = 100$ and $p = 0.5$. At each replication, we estimate p by means of the MLE $\hat{p} = x/n$ and by means of AMLE, using an ABC sample size $m = 1000$, $D = [0, 1]$ and $\epsilon \in \{0.01, 0.05, 0.1\}$. It is well-known from classical MLE asymptotic theory that $\text{var}(\hat{p}) \approx p(1-p)/n = 0.0025$. Table 1 gives the bias, the variance and the MSE of the two estimators. AMLE is clearly preferable to MLE, in terms of variance and MSE, for $\epsilon \in \{0.01, 0.05\}$, whereas MLE is better for $\epsilon = 0.5$. The rejection rate is given by $B(1 - m/\sum_{i=1}^{B} T_i)$, where T_i is the number of replications of the ABC algorithm needed to generate m observations from $\hat{\pi}_\epsilon(\theta|\boldsymbol{y})$ at the i-th replication of Algorithm 3. Whereas with $\epsilon = 0.5$ almost all observations are accepted, the computational cost associated to the case $\epsilon = 0.01$ is much larger, because approximately 99% of the simulated parameter values are rejected.

4 Numerical experiments

We now turn to the simulation experiments concerning the use of AMLE for estimating the autologistic model. As said above, three parameters need to be set as input of the algorithm. In doing this, we have to face a trade-off between computational burden and precision of the estimators. In particular, the precision increases as ϵ gets smaller and m gets larger. $\boldsymbol{D}$ is less important because, provided it includes the true parameter value, it mainly affects the rejection rate (Rubio and Johansen, 2013).

We note first that there is a negative relationship between ϵ and the rejection rate rr, because rr is given by $rr \stackrel{\text{def}}{=} \#\{\boldsymbol{y}' : |\eta(\boldsymbol{y}) - \eta(\boldsymbol{y}')| > \epsilon\}/B$, where B is the number of replications, and

$$\lim_{B\to\infty} rr = P(|\eta(\boldsymbol{y}) - \eta(\boldsymbol{y}')| > \epsilon). \tag{7}$$

Application of AMLE to the autologistic model is quite heavy from the computational point of view. Suppose indeed to be interested in implementing Algorithm 3 with $m = 100$ and to use a value of ϵ such that the rejection rate is $rr = 0.99$. Then the total number of autologistic models that need to be simulated is 10 000. Using the Metropolis algorithm with 500 steps, the total number of random variables to be simulated is $500 \cdot M \cdot N \cdot 10\,000$.

The following simulation experiments are based on setups similar to Gu and Zhu (2001), Huang and Ogata (2002) and Friel and Pettitt (2004), so that the results obtained with AMLE are comparable to exact MLEs produced by their procedures. In detail, the three experiments carried out in this paper are organized as follows:

1. Simulate an $N \times M$ autologistic model with $N = 125$ and $M = 12$. This is the dimension of the real-data analysis carried out in this paper as well as in Gu and Zhu (2001) with the Wiebe's wheat data (see below). The hyperparameters are $\alpha = 0$ and $\beta \in \{-0.4, -0.2, 0, 0.2, 0.4\}$. α is treated as a known parameter, so that we only estimate β.

2. Same as 1, but now the lattice has $M = N = 64$, as in Gu and Zhu (2001)'s Monte Carlo investigations.

3. Analogously to Friel and Pettitt (2004), simulate a $N \times M$ autologistic model with $N = 12$ and $M = 100$. The hyperparameters are $\alpha \in \{-0.3, 0, 0.3\}$ and $\beta \in \{-0.3, -0.1, 0, 0.1, 0.3\}$. Both parameters are considered unknown and therefore estimated from the simulated data.

All the simulations are repeated $B = 15$ times. In order to compare the estimators, we use the Root-Mean-Square-Error (RMSE), given by $\text{RMSE}(\hat{\theta}) = \sqrt{b(\hat{\theta})^2 + \text{var}(\hat{\theta})}$, where $b(\hat{\theta}) = (1/B)\sum_{i=1}^{B} \hat{\theta}_i - \theta$ is the bias,

Table 2: Bias, standard deviation, Root-Mean-Squared Error and relative efficiency of $\hat{\beta}^{AMLE}$ vs $\hat{\beta}^{MPLE}$ for various values of β, with $N = 125$, $M = 12$ and α known. Relative efficiency is defined as RelEff = RMSE($\hat{\beta}^{AMLE}$)/RMSE($\hat{\beta}^{MPLE}$).

		Bias ($\times 10^{-2}$)	Sd ($\times 10^{-2}$)	RMSE ($\times 10^{-2}$)	RelEff
$\beta = -0.4$	MPLE	3.39	0.67	3.45	0.43
	AMLE	1.20	0.89	1.49	
$\beta = -0.2$	MPLE	3.06	0.64	3.13	0.57
	AMLE	−1.51	0.95	1.79	
$\beta = 0$	MPLE	−0.37	0.61	0.71	1.04
	AMLE	0.39	0.63	0.74	
$\beta = 0.2$	MPLE	−2.21	0.74	2.34	0.34
	AMLE	−0.16	0.77	0.78	
$\beta = 0.4$	MPLE	−2.78	0.55	2.83	0.23
	AMLE	−0.22	0.62	0.65	

$\text{var}(\hat{\theta}) = (1/B)\sum_{i=1}^{B}(\hat{\theta}_i - \hat{\bar{\theta}})^2$ is the variance and $\hat{\bar{\theta}} = (1/B)\sum_{i=1}^{B}\hat{\theta}_i$ is the sample mean.

Before carrying out the experiments outlined above, we perform AMLE with $\alpha = 0$ and $\beta = 0.4$ with various values of ϵ, in order to get some insight into the impact of ϵ on the performance of the algorithm. Fig. 1 shows the results for $\epsilon \in \{5, \ldots, 15\}$ and $m \in \{200, 400\}$. For comparison purposes, we also report the RMSE of the MPLE estimator.

The graph suggests that the RMSE is essentially the same for all values of ϵ, as the differences seem to be random and mainly related to sampling variability. In the following we use $\epsilon = 9$, as it is the numerical value corresponding to the smallest RMSE, and $m = 200$, as $m = 400$ doubles computing time with no significant precision improvement.

Tables 2 and 3 give numerical values of bias, standard deviation, RMSE and relative efficiencies, measured as $RMSE(\hat{\beta}^{AMLE})/RMSE(\hat{\beta}^{MPLE})$, respectively for $N = 125$, $M = 12$ and for $N = M = 64$. Figure 2 shows relative efficiencies of $\hat{\beta}^{AMLE}$ with respect to $\hat{\beta}^{MPLE}$ both for $N = 125$, $M = 12$ and for $N = M = 64$.

From the tables and the figure, at least three facts emerge clearly. First, the relative efficiency of AMLE gets larger as β increases in absolute value; this is unsurprising, as MPLE corresponds to MLE when $\beta = 0$, and the loss of efficiency is an increasing function of $|\beta|$. Second, the improvement in relative efficiency brought by AMLE seems to be different for positive and negative spatial dependence. This outcome may be partly explained by the sampling variability related to the small number of replications, but not

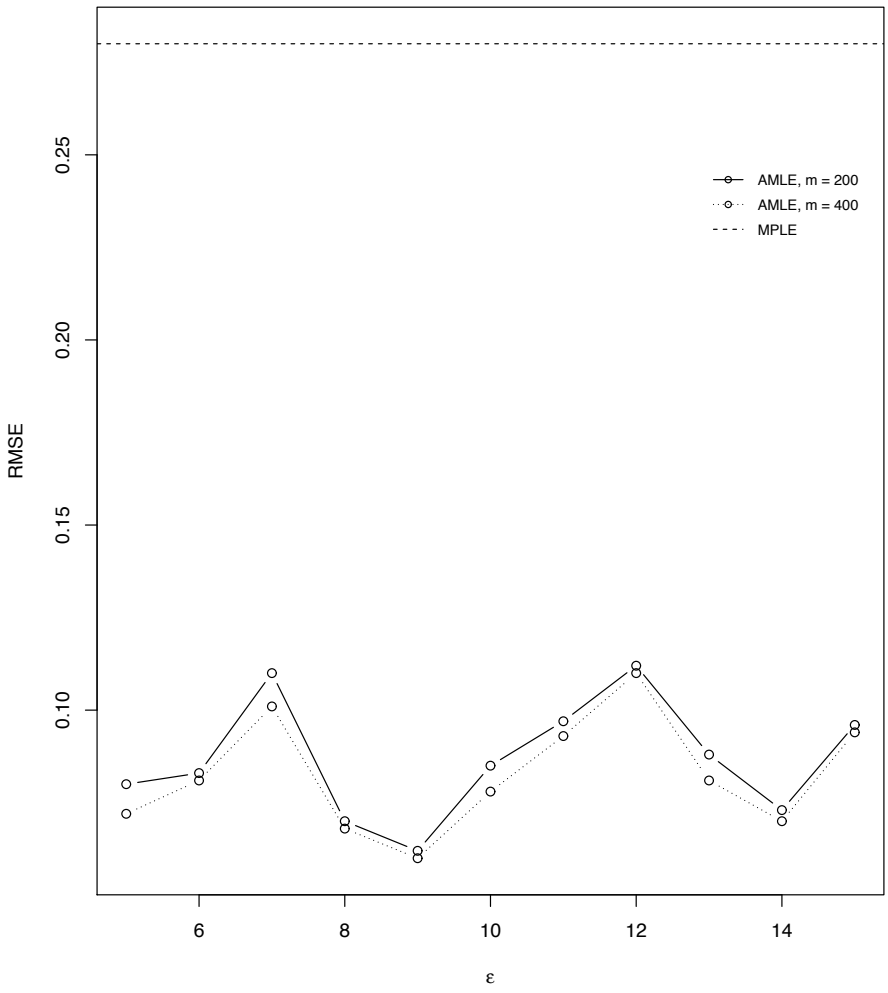

Figure 1: RMSE of three estimators of β (univariate case, hyperparameters $\alpha = 0, \beta = 0.4$, sample sizes $m \in \{200, 400\}$).

Table 3: Bias, standard deviation, Root-Mean-Squared Error and relative efficiency of $\hat{\beta}^{AMLE}$ vs $\hat{\beta}^{MPLE}$ for various values of β, with $N = M = 64$. Relative efficiency is defined as RelEff = RMSE($\hat{\beta}^{AMLE}$)/RMSE($\hat{\beta}^{MPLE}$).

		Bias ($\times 10^{-2}$)	Sd ($\times 10^{-2}$)	RMSE ($\times 10^{-2}$)	RelEff
$\beta = -0.4$	MPLE	3.30	0.62	3.36	0.19
	AMLE	0.27	0.58	0.64	
$\beta = -0.2$	MPLE	1.50	0.37	1.54	0.33
	AMLE	0.09	0.49	0.50	
$\beta = 0$	MPLE	0.11	0.43	0.45	1.07
	AMLE	0.04	0.48	0.48	
$\beta = 0.2$	MPLE	1.34	0.56	1.45	0.61
	AMLE	0.45	0.75	0.88	
$\beta = 0.4$	MPLE	2.85	0.52	2.89	0.24
	AMLE	0.43	0.56	0.70	

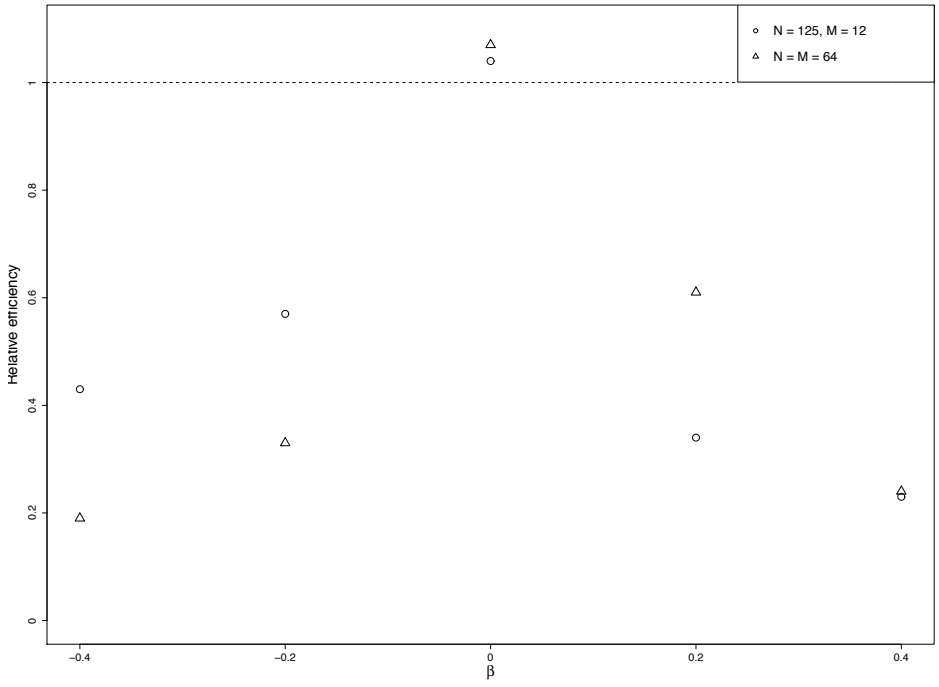

Figure 2: Relative efficiency of y of $\hat{\beta}^{AMLE}$ with respect to $\hat{\beta}^{MPLE}$, measured as $RMSE(\hat{\beta}^{AMLE})/RMSE(\hat{\beta}^{MPLE})$ in the univariate case for $\alpha = 0$ and $\beta \in \{-0.4, -0.2, 0, 0.2, 0.4\}$, with $m = 200$.

new in the spatial statistics literature (see, e.g., Schabenberger and Gotway, 2002; Griffith and Arbia, 2010; Arbia et al., 2011; Arbia et al., 2013), where "asymmetric" results corresponding to setups with spatial dependence of the same magnitude but different sign are well known. Finally, the smaller MSE of AMLE is entirely due to a smaller bias, whereas the standard deviation of the two estimators is approximately the same or slightly larger for AMLE, probably because of some extra Monte Carlo sampling variability.

For comparison purposes, we recall that for model (2) with $\alpha = 0$ on a 64×64 grid, Huang and Ogata (2002) find via simulation that the MLE of J has efficiencies of 0.983, 0.963, 0.897, and 0.750 with respect to PMLE for $J \in \{0.1, 0.2, 0.3, 0.4\}$[1]. For the same model on a 125×12 grid, Gu and Zhu (2001) find efficiencies of 0.173, 0.770, 1.106, 0.635 and 0.180 with respect to PMLE for $J \in \{-0.4, -0.2, 0, 0.2, 0.4\}$.

The rejection rate is between 0.991 and 0.993 for the experiments in Table 2 and between 0.992 and 0.994 for those in Table 3. Thus, to obtain a sample of 200 parameter values, the number of simulations of the autologistic model is approximately equal to 25 000 in the first case and to 28 000 in the second case. The amount of time needed ranges between about two and two and a half hours per replication on a 3.16 GHz machine.

We now turn to the bivariate case, using first $\epsilon = 9$ and $m = 200$ as in the univariate case. The results in Table 4 show that overall AMLE performs better than MPLE, but the difference is smaller than in the preceding tables. One reason may be that AMLE needs a smaller ϵ and/or a larger m to reach convergence. However, before considering this issue, it is worth noting that this result is also related to a peculiar feature of the MPLE estimators.

Similarly to the univariate cases, the improvement is mainly due to a smaller bias component. For some parameter combinations (for example, when $\alpha = 0$ and $\beta = -0.3$), MPLE is approximately unbiased, so that the two estimators have a similar MSE. In these instances, concluding that MPLE performs well seems more appropriate than concluding that AMLE performs poorly, as $\hat{\alpha}^{MPLE}$ is approximately unbiased. Continuing the analysis of the same example, when $\alpha = 0$ and $\beta = -0.3$, $\hat{\beta}^{AMLE}$ has a smaller RMSE than $\hat{\beta}^{MPLE}$, because the latter shows a systematic bias. The other cases where the difference is small mostly correspond to $|\beta| = 0.1$ and are therefore not surprising, because, when $|\beta| \to 0$, MPLEs converge to MLEs.

As of the convergence issue, note that the need of a smaller ϵ and/or a larger m for reaching convergence in the bivariate case may be justified by the fact that kernel density estimation works worse when the dimension of the parameter space is high. Thus, we repeated some of the experiments with $\epsilon = 5$ and $m = 500$. Detailed results are displayed in Table 5 and

[1]Note that Huang and Ogata (2002) measure the efficiency as the ratio of the MSEs. In order to be able to carry out meaningful comparisons, we have taken the squared root of the efficiencies reported in their paper, which corresponds to the ratio of the RMSEs.

Table 4: Bias, standard deviation, Root-Mean-Squared Error and relative efficiency of $\hat{\alpha}^{AMLE}$ vs $\hat{\alpha}^{MPLE}$ and $\hat{\beta}^{AMLE}$ vs $\hat{\beta}^{MPLE}$ for various values of α and β, with $N = 12$, $M = 100$, $\epsilon = 9$ and $m = 200$. Relative efficiency is defined as RelEff = RMSE($\hat{\theta}^{AMLE}$)/RMSE($\hat{\theta}^{MPLE}$), where θ is equal to α or β.

		Bias ($\times 10^{-2}$)	Sd ($\times 10^{-2}$)	RMSE ($\times 10^{-2}$)	RelEff
$\alpha = -0.3, \beta = -0.3$	$\hat{\alpha}^{MPLE}$	1.895	2.395	3.054	0.819
	$\hat{\alpha}^{AMLE}$	−0.079	2.373	2.502	
	$\hat{\beta}^{MPLE}$	3.167	0.986	3.318	0.324
	$\hat{\beta}^{AMLE}$	−0.002	1.076	1.076	
$\alpha = -0.3, \beta = -0.1$	$\hat{\alpha}^{MPLE}$	1.988	1.473	2.474	0.791
	$\hat{\alpha}^{AMLE}$	0.931	1.723	1.958	
	$\hat{\beta}^{MPLE}$	0.256	0.857	0.894	0.988
	$\hat{\beta}^{AMLE}$	−0.234	0.851	0.883	
$\alpha = -0.3, \beta = 0.1$	$\hat{\alpha}^{MPLE}$	1.919	1.583	2.488	0.806
	$\hat{\alpha}^{AMLE}$	1.091	1.682	2.005	
	$\hat{\beta}^{MPLE}$	1.650	1.375	2.148	0.824
	$\hat{\beta}^{AMLE}$	1.087	1.396	1.769	
$\alpha = -0.3, \beta = 0.3$	$\hat{\alpha}^{MPLE}$	−1.493	1.735	2.288	0.665
	$\hat{\alpha}^{AMLE}$	−0.035	1.592	1.592	
	$\hat{\beta}^{MPLE}$	−2.684	0.959	2.850	0.349
	$\hat{\beta}^{AMLE}$	−0.068	0.993	0.995	
$\alpha = 0, \beta = -0.3$	$\hat{\alpha}^{MPLE}$	−0.040	1.860	1.860	1.006
	$\hat{\alpha}^{AMLE}$	−0.206	1.859	1.871	
	$\hat{\beta}^{MPLE}$	1.362	0.835	1.598	0.689
	$\hat{\beta}^{AMLE}$	−0.466	0.998	1.101	
$\alpha = 0, \beta = -0.1$	$\hat{\alpha}^{MPLE}$	0.181	1.806	1.815	1.000
	$\hat{\alpha}^{AMLE}$	0.123	1.812	1.816	
	$\hat{\beta}^{MPLE}$	−0.085	1.075	1.079	0.992
	$\hat{\beta}^{AMLE}$	0.019	1.070	1.070	
$\alpha = 0, \beta = 0.1$	$\hat{\alpha}^{MPLE}$	−0.449	1.376	1.448	0.824
	$\hat{\alpha}^{AMLE}$	−0.454	1.103	1.193	
	$\hat{\beta}^{MPLE}$	0.534	0.850	1.004	0.897
	$\hat{\beta}^{AMLE}$	0.047	0.900	0.901	
$\alpha = 0, \beta = 0.3$	$\hat{\alpha}^{MPLE}$	1.503	1.621	2.211	0.498
	$\hat{\alpha}^{AMLE}$	0.213	1.079	1.100	
	$\hat{\beta}^{MPLE}$	−3.013	1.432	3.336	0.518
	$\hat{\beta}^{AMLE}$	−1.167	1.274	1.728	
$\alpha = 0.3, \beta = -0.3$	$\hat{\alpha}^{MPLE}$	−1.935	1.819	2.656	0.832
	$\hat{\alpha}^{AMLE}$	−0.779	2.067	2.209	
	$\hat{\beta}^{MPLE}$	1.753	0.812	1.932	0.814
	$\hat{\beta}^{AMLE}$	−1.331	0.837	1.572	
$\alpha = 0.3, \beta = -0.1$	$\hat{\alpha}^{MPLE}$	−0.419	1.396	1.457	0.992
	$\hat{\alpha}^{AMLE}$	0.544	1.390	1.446	
	$\hat{\beta}^{MPLE}$	2.264	1.169	2.548	0.852
	$\hat{\beta}^{AMLE}$	1.489	1.354	2.171	
$\alpha = 0.3, \beta = 0.1$	$\hat{\alpha}^{MPLE}$	2.467	1.381	2.827	0.512
	$\hat{\alpha}^{AMLE}$	0.525	1.348	1.446	
	$\hat{\beta}^{MPLE}$	−1.964	0.783	2.115	0.538
	$\hat{\beta}^{AMLE}$	−0.776	0.831	1.137	
$\alpha = 0.3, \beta = 0.3$	$\hat{\alpha}^{MPLE}$	2.624	1.389	2.969	0.780
	$\hat{\alpha}^{AMLE}$	1.681	1.594	2.317	
	$\hat{\beta}^{MPLE}$	−2.676	0.639	2.751	0.769
	$\hat{\beta}^{AMLE}$	−1.703	1.256	2.116	

Table 5: Bias, standard deviation, Root-Mean-Squared Error and relative efficiency of $\hat{\alpha}^{AMLE}$ vs $\hat{\alpha}^{MPLE}$ and $\hat{\beta}^{AMLE}$ vs $\hat{\beta}^{MPLE}$ for various values of α and β, with $N = 12$, $M = 100$, $\epsilon = 5$ and $m = 500$. Relative efficiency is defined as RelEff = RMSE($\hat{\theta}^{AMLE}$)/RMSE($\hat{\theta}^{MPLE}$), where θ is equal to α or β.

		Bias ($\times 10^{-2}$)	Sd ($\times 10^{-2}$)	RMSE ($\times 10^{-2}$)	RelEff
$\alpha = -0.3, \beta = -0.3$	$\hat{\alpha}^{MPLE}$	1.106	2.117	2.388	1.013
	$\hat{\alpha}^{AMLE}$	−0.729	2.308	2.420	
	$\hat{\beta}^{MPLE}$	2.991	0.849	3.109	0.305
	$\hat{\beta}^{AMLE}$	−0.036	0.948	0.949	
$\alpha = 0, \beta = -0.3$	$\hat{\alpha}^{MPLE}$	−1.512	1.500	2.130	1.006
	$\hat{\alpha}^{AMLE}$	−1.387	1.788	2.143	
	$\hat{\beta}^{MPLE}$	2.042	0.782	2.186	0.432
	$\hat{\beta}^{AMLE}$	−0.046	0.943	0.945	
$\alpha = 0, \beta = -0.1$	$\hat{\alpha}^{MPLE}$	−0.552	1.864	1.944	0.951
	$\hat{\alpha}^{AMLE}$	−0.467	1.788	1.848	
	$\hat{\beta}^{MPLE}$	1.086	1.192	1.612	0.968
	$\hat{\beta}^{AMLE}$	−0.021	1.561	1.561	
$\alpha = 0, \beta = 0.1$	$\hat{\alpha}^{MPLE}$	1.004	1.475	1.784	0.831
	$\hat{\alpha}^{AMLE}$	0.185	1.471	1.483	
	$\hat{\beta}^{MPLE}$	−1.821	1.208	2.185	0.919
	$\hat{\beta}^{AMLE}$	−0.962	1.764	2.009	
$\alpha = 0, \beta = 0.3$	$\hat{\alpha}^{MPLE}$	−1.778	1.270	2.185	0.689
	$\hat{\alpha}^{AMLE}$	−0.975	1.148	1.592	
	$\hat{\beta}^{MPLE}$	−2.507	1.094	2.735	0.413
	$\hat{\beta}^{AMLE}$	0.034	1.130	1.130	
$\alpha = 0.3, \beta = 0.3$	$\hat{\alpha}^{MPLE}$	2.633	1.636	3.100	0.757
	$\hat{\alpha}^{AMLE}$	1.708	1.611	2.348	
	$\hat{\beta}^{MPLE}$	−2.770	0.964	2.933	0.349
	$\hat{\beta}^{AMLE}$	−0.348	0.962	1.023	

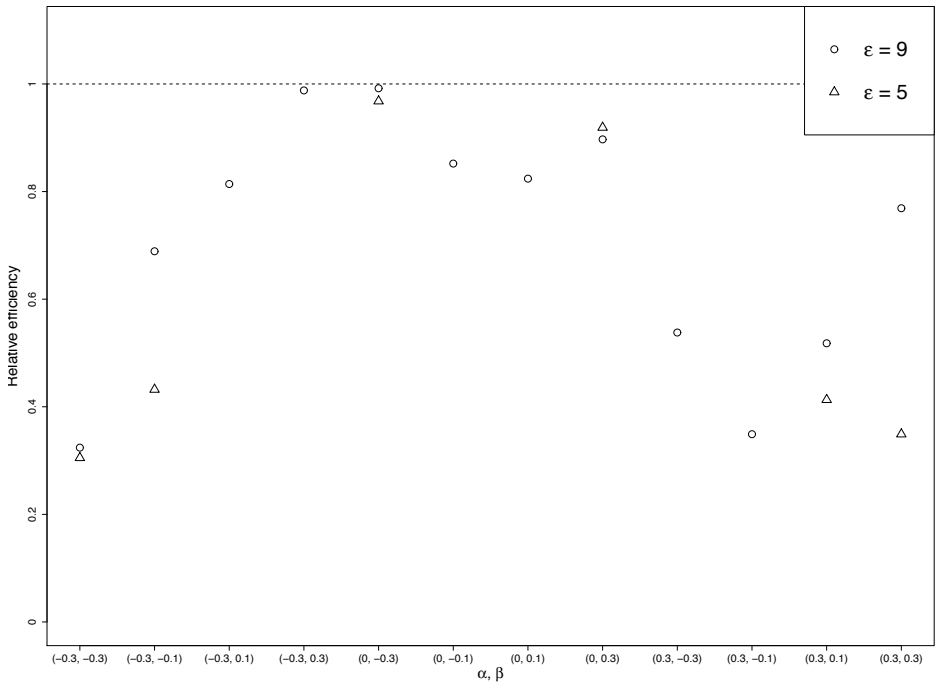

Figure 3: RMSE of the estimator of β in the bivariate case for various parameter values, $\epsilon \in \{5, 9\}$, $m = 500$).

relative efficiencies are shown in Fig. 3.

Now the improvement obtained with AMLE is more significant, but mostly for $\hat{\beta}$. Moreover, the gain in efficiency of AMLE is much larger when $|\beta|$ is larger. Similarly to the univariate case, the relative efficiency of the estimators of the intercept in cases with $\beta = 0.3$ and $\beta = -0.3$ is quite different.

Analogously to the single-parameter case, for all the experiments in Table 4 the rejection rate is between 0.991 and 0.993, so that, for simulating $m = 200$ pairs of parameters, the computing times are about the same. As for the setup of Table 5, the rejection rate is between 99.6 and 99.8%, so that, in order to get 500 simulated parameter values, one needs to generate between approximately 130 000 and 230,000 autologistic models. The necessary amount of time is between about 12 and 22 hours respectively.

4.1 Application: Wiebe's wheat data

We apply the AMLE method to the same dataset used in Gu and Zhu (2001), now available in the `R` package `agridat` (Wright, 2013). The data are the results of a uniformity trial of 1500 plots of wheat conducted in Idaho in

1927, and refer to a rectangular lattice of dimension 125×12. Values larger than the mean are set equal to 1, the remaining values are equal to 0, so that the variable of interest is binary. We use AMLE with $m = 500$ and $\epsilon = 9$, obtaining $\hat{\beta}^{AMLE} = 0.376$. The result is in good agreement with the MLE computed by Gu and Zhu (2001), equal to 0.372 with an estimated standard deviation of 0.012.

5 Conclusion

In this paper we have estimated the parameters of the autologistic model via Approximate Maximum Likelihood. The finite sample properties of the estimators depend on the parameters of the algorithm, so that we ran simulation experiments aiming at measuring the precision of the estimators and its relationship to the input parameters.

The choice of the scaling factor ϵ and of the ABC sample size m seems to have a rather limited impact on the performance of the algorithm in the single-parameter case, but is more important in the two-parameter model, probably because kernel density estimation is less effective when the dimension of the space increases.

Overall, the performance of the algorithm is excellent in the univariate case, where relative efficiency of AMLE in terms of RMSE, ranges between 19% and 43% when $|\beta| = 0.4$ and between 33% and 61% when $|\beta| = 0.2$. In the bivariate case, and for different parameter values, the big picture is similar, but only under more stringent conditions about ϵ and m, and especially as concerns the estimator of β: the relative efficiency of the estimators of β is between 0.305 and 0.432 for $|\beta| = 0.3$, and between 0.919 and 0.968 for $|\beta| = 0.1$. It is worth noting explicitly that in the two-parameter example our Monte Carlo evidence suggests that reducing ϵ and/or increasing m gives estimators with smaller RMSE. Therefore these outcomes, which are indeed essentially in line with MLE results obtained in the literature in analogous frameworks, can probably be improved, at the cost of an increased computational burden.

The AMLE method has two major advantages. First, it can be used for any dimension of the lattice, as long as it can be simulated; basically, this means that the only limit is the computer's physical memory. Second, its implementation is easy. The computational cost is rather large, but this is a serious problem mostly in simulation experiments, where the model needs to be estimated several times. In real-data point estimation problems, where the procedure has to be run only once, setting a small ϵ and a large ABC sample size is usually not a major difficulty.

We mention two important issues that require further research. First, as pointed out above, the method can be easily extended to other classical spatial models for which MLE is difficult. The most obvious example, but

certainly not the only one, is the autonormal model. Moreover, the approach can probably be used in more complicated spatial models where MLE is not available at all, for example when data are not only spatially dependent, but also clustered in groups, and nested or crossed effects are of specific interest (see e.g. the model proposed by Corrado and Fingleton, 2012).

References

Arbia, G. (2006). *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*. Advances in Spatial Science. Springer.

Arbia, G., Bee, M., and Espa, G. (2011). Aggregation of regional economic time series with different spatial correlation structures. *Geographical Analysis*, 43:78–103.

Arbia, G., Bee, M., and Espa, G. (2013). Testing isotropy in spatial econometric models. *Spatial Economic Analysis*, 8:228–240.

Besag, J. E. (1975). Statistical analysis of non-lattice data. *The Statistician*, 24:179–195.

Casella, G. and Robert, C. (2004). *Monte Carlo Statistical Methods*. Springer, second edition.

Cipra, B. (1987). An introduction to the Ising model. *American Mathematical Monthly*, 94:937–959.

Comets, F. (1992). On consistency of a class of estimators for exponential families of markov random fields on the lattice. *Annals of Statistics*, 20:455–468.

Corrado, L. and Fingleton, B. (2012). Where is the economics in spatial econometrics? *Journal of Regional Science*, 52:210–239.

Friel, N. and Pettitt, A. N. (2004). Likelihood estimation and inference for the autologistic model. *Journal of Computational and Graphical Statistics*, 13:232–246.

Geman, S. and Graffigne, G. (1987). Markov random field image models and their applications to computer vision. In Gleason, A. M., editor, *Proceedings of the International Congress of Mathematics*. American Mathematical Society.

Griffith, D. A. and Arbia, G. (2010). Detecting negative spatial autocorrelation in georeferenced random variables. *International Journal of Geographical Information Science*, 24:417–437.

Gu, M.-G. and Zhu, H.-T. (2001). Maximum likelihood estimation for spatial models by Markov chain Monte Carlo stochastic approximation. *Journal of the Royal Statistical Society B*, 63:339–355.

Guyon, X. and Künsch, H. R. (1992). *Asymptotic Comparison of Estimators in the Ising Model*, pages 177–198. Number 74 in Stochastic Models, Statistical Methods, and Algorithms in Image Analysis, Lecture Notes in Statistics. Springer.

Huang, F. and Ogata, Y. (2002). Generalized pseudo-likelihood estimates for markov random fields on lattice. *Annals of the Institute of Statistical Mathematics*, 54:1–18.

Huffer, F. and Wu, H. (1998). Markov Chain Monte Carlo for autologistic regression models with application to the distribution of plant species. *Biometrics*, 54:509–524.

Hughes, J., Haran, M., and Caragea, P. C. (2011). Autologistic models for binary data on a lattice. *Environmetrics*, 22:857–871.

Ionides, E. L. (2005). Maximum smoothed likelihood estimation. *Statistica Sinica*, 15:1003–1014.

Metropolis, N., Rosenbluth, A., Rosenbluth, N., Teller, A., and Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092.

Moyeed, R. A. and Baddeley, A. J. (1991). Stochastic approximation of the MLE for a spatial point pattern. *Scandinavian Journal of Statistics*, 18:39–50.

Murray, I., Ghahramani, Z., and MacKay, D. (2006). Mcmc for doubly-intractable distributions. In *Proceedings of the 22nd Annual Conference on Uncertainty in Artificial Intelligence UAI06*, pages 359–366.

Ogata, Y. and Tanemura, M. (1984). Likelihood analysis of spatial point patterns. *Journal of the Royal Statistical Society. Series B*, 46:496–518.

Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A., and Feldman, M. T. (1999). Population growth of human y chromosomes: A study of y chromosome microsatellites. *Molecular Biology and Evolution*, 16:1791–1798.

Rubio, F. J. and Johansen, A. M. (2013). A simple approach to maximum intractable likelihood estimation. *Electronic Journal of Statistics*, 7:1632–1654.

Schabenberger, O. and Gotway, C. (2002). *Statistical Methods For Spatial Data Analysis*. Chapman and Hall.

Strauss, D. (1992). The many faces of logistic regression. *The American Statistician*, 46:321–327.

Wang, Z. and Zheng, Y. (2013). Analysis of binary data via a centered spatial-temporal autologistic regression model. *Environmental and Ecological Statistics*, 20:37–57.

Wright, K. (2013). *Agridat: Agricultural datasets*. R package version 1.8.

List of recent papers published un the same series

2013/01 A regular multidistance among fuzzy numbers, Franco Molinari

2013/02 Employer moral hazard and wage rigidity. The case of worker-owned and investor-owned firms, Marina Albanese, Cecilia Navarra and Ermanno Tortia

2013/03 Origins and prospects of the Euro existential crisis, Luigi Bonatti and Andrea Fracasso

2013/04 Uses And Motivations For Credit Derivatives: An Empirical Investigation Into Italian Banks, Eleonora Broccardo, Maria Mazzuca and Elmas Yaldiz

2013/05 At the core of the international financial system, Valentina Feroldi, Edoardo Gaffeo

2013/06 Reassessing the spatial determinants of the growth of Italian SMEs, Roberto Gabriele, Diego Giuliani, Marco Corsino, Giuseppe Espa

2013/07 A spatial and sectoral analysis of firm demography in Italy, Giuseppe Espa, Danila Filipponi, Diego Giuliani, Davide Piacentino

2013/08 Using information markets in grantmaking. An assessment of the issues involved and an application to Italian banking foundations, Edoardo Gaffeo

2013/09 Interbank contagion and resolution procedures: inspecting the mechanism, Edoardo Gaffeo e Massimo Molinari

2013/10 Transatlantic contagion 2010-…, Roberto Tamborini

2013/11 Do middle managers matter?, Elena Feltrinelli, Roberto Gabriele, Sandro Trento