Markovian auto-models with mixed states

Cécile Hardouin[1] and Jian-Feng Yao[2]

[1] SAMOS - Université Paris 1
90 rue de Tolbiac,
75634 Paris Cedex 13, France
(e-mail: hardouin@univ-paris1.fr)

[2] IRMAR - Université de Rennes 1
Campus de Beaulieu,
35042 Rennes Cedex, France
(e-mail: Jian-Feng.Yao@univ-rennes1.fr)

Abstract. We present a new class of Markovian auto-models with a mixed state space $E = \{0\} \cup]0; +\infty[$ involving both discrete and continuous states. We first introduce an extension of the Besag's auto-models to the multivariate case ; then we define the specific Markovian random field defined on a lattice S, whose components are valued in E with conditional distribution belonging to an exponential family. We study two particular examples, based on the use of the exponential distribution and the Gaussian positive distribution, and look for the admissibility conditions for such models. Last, we present briefly some experimental results obtained for the analysis of motion measurements of video sequences.
Keywords: Auto-models, Mixed states.

1 Besag auto-models : multivariate extension

We consider a set of sites $S = \{1, ..n\}$, a measurable space $(E, \mathcal{E})$ (usually a subset of $\mathbb{R}^d$) equipped with the measure ν. The product space is $(\Omega, \mathcal{O}) = (E^S, \mathcal{E}^{\otimes S})$ with the product measure $\nu^S = \nu^{\otimes S}$. A random field is a probability measure μ over $(\Omega, \mathcal{O})$; we assume that μ admits a probability density f everywhere positive w.r.t ν^S.

The set of sites S is equipped with a graph structure $\mathcal{G}$, symmetrical and reflexive called the neighborhood graph; $\langle i, j \rangle$ denotes that i and j are neighbors, for $i \neq j$. A non empty set $C \subseteq S$ is a clique if C is a single point or if all elements of C are pairwise neighbors.

The field is Markovian if all the conditional distributions on the outer configurations depend on the configurations on the neighborhoods.

Let us note $\mathbf{0}$ a reference layout of Ω ($\mathbf{0}$ is 0 when $E = \mathbb{N}$, $\mathbb{R}$ or $\mathbb{R}_+$), then we can write

$$\mu(dx) = f(x)\nu^S(dx), \qquad f(x) = f(\mathbf{0}) \exp U(x)$$

with $U(\mathbf{0}) = 0$. Moreover, the energy U is the sum of potentials ϕ_A, $A \in \mathcal{C}$ the set of cliques.

According to Besag's definition ([Besag, 1974]), a real-valued field X is an auto-model if its distribution μ can be written as

$$U(x) = \sum_{i \in S} \phi_i(x_i) + \sum_{\{i,j\}} \beta_{ij} x_i x_j \ ,$$

with $\beta_{ij} = \beta_{ji}$. Thus an auto-model is a Markovian field with cliques of at most two points and linear pairwise interaction potentials.

We denote the conditional law on a site i by $f_i(x_i|.)$. The following result characterizes Besag's auto-models in the $d-$dimensional case.

Theorem 1 *We assume that for each site i, the conditional density belongs to a multi-parameter exponential family:*

$$\ln f_i(x_i|.) = \langle A_i(.), B_i(x_i) \rangle + C_i(x_i) + D_i(.) \ , \ A_i \in \mathbb{R}^d \ , \ B_i(x_i) \in \mathbb{R}^d \ . \quad (1)$$

with $B_i(0) = C_i(0) = 0$ for $0 \in E$. And that the family of sufficient statistics $\{B_i(x_i)\}$ is regular in the sense that

$$\text{for all } i \in S, \quad Span\{B_i(x_i), \ x_i \in E\} = \mathbb{R}^d \ .$$

Then there exist for all $i, j \in S, \ i \neq j$, a family of vectors $\alpha_i \in \mathbb{R}^d$ and a family of $d \times d$ matrices satisfying $\beta_{ij} = \beta_{ji}^t$ such that

$$A_i(.) = \alpha_i + \sum_{j \neq i} \beta_{ij} B_j(x_j) \ . \quad (2)$$

Consequently the set of potentials is given by

$$\phi_i(x_i) = \langle \alpha_i, B_i(x_i) \rangle + C_i(x_i) \ , \quad (3)$$

and

$$\phi_{ij}(x_i, x_j) = \phi_{ij}(x_i, x_j) = B_i^t(x_i) \beta_{ij} B_j(x_j) \ . \quad (4)$$

See [Hardouin and Yao, 2004] for the proof.

Conversely, a Gibbs distribution with potentials (3) and (4) has conditional distributions given by (1) and (2) as soon as the energy U is admissible, i.e. $\int_\Omega \exp U(x) \nu^S(dx) < \infty$.

2 Random variable with mixed states

2.1 Distribution of mixed exponential family $\mathcal{L}(p, \xi)$:

We consider X which takes values in $E = \{0\} \cup]0, +\infty[$, equipped with the measure

$$\nu(dx) = \delta(dx) + \lambda(dx) \quad (5)$$

where δ is the Dirac measure at 0, and λ is the Lebesgue measure on $\mathcal{B}(]0, +\infty[)$.

We define the random variable X with mixed exponential family distribution on E. Let $p \in]0,1[$; then $X = 0$ with probability p, and with probability $1-p$, $X > 0$ follows a distribution which belongs to an exponential family, with the probability density :

$$g_\xi(x) = G(\xi)\ \exp\langle \xi, T(x)\rangle\ ,\ x > 0$$

where T is defined such as $T(0) = 0$. The probability density of X on E is (w.r.t. ν) :

$$\begin{aligned} f_\theta(x) &= p\delta(x) + (1-p)g_\xi(x) \\ &= p\ \exp\left\{(1-\delta(x))\ln\frac{(1-p)G(\xi)}{p} + \langle \xi, T(x)\rangle\right\} \\ &= Z^{-1}(\theta)\ \exp\langle \theta, B(x)\rangle \end{aligned}$$

where $\theta = (\theta_1, \theta_2)^t = (\ln\frac{(1-p)G(\xi)}{p}, \xi)^t$ and $B = (\delta^*, T^t)^t$ where we set $\delta^* = 1-\delta$ in order to have $B(0) = 0$.

We denote this mixed distribution by $\mathcal{L}(p,\xi)$. Let us precise two particular cases of further use.

Mixed exponential distribution $\mathcal{E}(p,\lambda)$:

Let $g_\xi(x) = \lambda\exp\{-\lambda x\}\ ,\ x > 0$. Then

$$f(x) = p\exp\{\delta^*(x)\ln\frac{(1-p)\lambda}{p} - \lambda x\} = Z^{-1}(\theta)\ \exp\langle\theta, B(x)\rangle$$

Here $\theta = (\theta_1,\theta_2)^t = (\ln\frac{(1-p)\lambda}{p}, \lambda)^t$ and the sufficient statistics is $B(x) = (\delta^*(x), -x)^t$. Conversely we have $\lambda = \theta_2$ and $p = \frac{\theta_2}{\theta_2 + \exp\theta_1}$.

Mixed positive Gaussian distribution $G(p,\sigma^2)$

With probability $1-p$, X=$|Z|$ where Zsim$N(0,\sigma^2)$. The probability density of X is given by $f(x) = Z^{-1}(\theta)\ \exp\langle\theta, B(x)\rangle$ with $\theta = (\theta_1,\theta_2)^t = (\ln\frac{2(1-p)}{p\sigma\sqrt{2\pi}}, \frac{1}{2\sigma^2})^t$ and $B(x) = (\delta^*(x), -x^2)^t$. We get also $\sigma^2 = \frac{1}{2\theta_2}$ and $p = \frac{2}{2+\sqrt{2\pi\theta_2}\exp\theta_1}$.

3 Markovian auto-models with mixed states

We now consider a random field X on $S = \{1,2,\cdots,n\}$, $X = (X_1, X_2, \cdots, X_n)$, in $F = E^S = (\{0\}\cup]0,+\infty[)^S$.

We assume that the family of the conditional distributions $f_i(x_i|.)$ belongs to the family of mixed distributions $\mathcal{L}(p_i(.),\xi_i(.))$ described previously. In other words, we can write (**??**) with

$$\ln f_i(x_i|.) = \mathcal{L}(p_i(.),\xi_i(.)) = \langle A_i(.), B_i(x_i)\rangle + C(x_i) + D_i(.)\ ,$$

with $B_i(x_i) = (\delta^*(x_i), T_i^t(x_i))^t$. Theorem 1 ensures that there exists vectors $\alpha_i \in \mathbb{R}^2$ and $2\times 2-$matrices β_{ij} verifying $\beta_{ij} = \beta_{ji}^t$ such that

$A_i(.) = \theta_i(.) = \alpha_i + \sum_{j \neq i} \beta_{ij} B_j(x_j)$
and the potentials of the joint energy are given by (3) and (4).

Let us specify the resulting auto-models when we take for the density g_ξ of the positive component in each site first the exponential distribution and next the positive Gaussian distribution. For each example, we give conditions ensuring the admissibility of the models; then we specify them to the four nearest neighbors system, with or without isotropy. We further use the resulting models in two different contexts: we look for a "good" suitable set of parameters of the auto-exponential model in the rainfall framework, and apply the positive Gaussian auto-model to motion measurements of video sequences.

3.1 Mixed auto-exponential models

We suppose that the conditional distributions are in the family of mixed exponential distributions $\mathcal{E}(p_i(.), \lambda_i(.))$. Then, there exist $\alpha_i = (a_i, b_i)^t$, $\beta_{ij} = \begin{pmatrix} c_{ij} & d^*_{ij} \\ d_{ij} & e_{ij} \end{pmatrix}$ verifying $c_{ij} = c_{ji}$, $e_{ij} = e_{ji}$ and $d_{ij} = d^*_{ji}$ such that we can write the global energy as:

$$U(x) = \sum_{i \in S} \alpha_i^t B(x_i) + \sum_{(i,j):\langle i,j \rangle} B^t(x_i) \beta_{ij} B(x_j) \tag{6}$$

$$\begin{aligned} U(x) = \textstyle\sum_{i \in S} a_i \delta^*(x_i) - \sum_{i \in S} b_i x_i + \sum_{\langle i,j \rangle} c_{ij} \delta^*(x_i)\delta^*(x_j) \\ - \textstyle\sum_{(i,j):\langle i,j \rangle} d_{ij} x_i \delta^*(x_j) + \sum_{\langle i,j \rangle} e_{ij} x_i x_j \end{aligned} \tag{7}$$

We note that potential $\phi(x_i, x_j) = x_i \delta(x_j)$ is not symmetric in (x_i, x_j).

Proposition 1 *We assume that U satisfies the following condition* **(A)** *:*

$$(\mathbf{A}) : \begin{cases} \forall i \in S, \forall A \subset \partial i, \quad b_i + \sum_{j \in A} d_{ij} > 0 \\ \forall i, j \in S, \quad e_{ij} \leq 0 \end{cases} \tag{8}$$

Then the energy U is admissible.

Proof: see [Hardouin and Yao, 2004].

Under condition **(A)**, the model with density defined by $f(x) = Z^{-1} \exp U(x)$ where U satisfies (6) or (7) is called mixed exponential auto-model.

Conditional distributions:
By construction, for each i, $f_i(x_i|.)$ sim $\mathcal{E}(p_i(.), \lambda_i(.))$.
$f(x_i|\, x^i) = Z^{-1}(\theta, x^i) \exp\{\theta_1(x^i)\delta^*(x_i) - \theta_2(x^i)x_i\}$, where

$\theta_1(x^i) = a_i + \sum_{j:\langle i,j\rangle} \{c_{ij}\delta^*(x_j) - d^*_{ij}x_j\}$ and $\theta_2(x^i) = b_i + \sum_{j:\langle i,j\rangle} \{d_{ij}\delta^*(x_j) - e_{ij}x_j\}$

Example 1 : Mixed exponential auto-model with the 4 nearest neighbors.

We consider $S = [1, M] \times [1, N]$, and suppose that the energy is isotropic. Then we can write the energy depending on 5 parameters $\theta = (a, b, c, d, e)$:

$$U(x) = \sum_{i\in S}(a\delta^*(x_i) - bx_i) + \sum_{\langle i,j\rangle}\{c\delta^*(x_i)\delta^*(x_j) + ex_ix_j\} - d \sum_{(i,j):\langle i,j\rangle} x_i\delta^*(x_j)$$

(**A**) : $b > 0$, $b + 4d > 0$ and e ≤ 0

Conditional distribution is defined by :

$f(x_i|x^i) = Z^{-1}(\theta, x^i)\exp U_i(x_i|x^i)$, where $U_i(x_i|x^i) = \theta_1(x^i)\delta^*(x_i) - \theta_2(x^i)x_i$

$$\text{with} : \begin{cases} \theta_1(x^i) = a + c(4 - v_i(0)) - dv_i(+) \\ \theta_2(x^i) = b + d(4 - v_i(0)) - ev_i(+) \\ v_i(0) = \sum_{j:\langle i,j\rangle} \delta(x_j) \text{ and } v_i(+) = \sum_{j:\langle i,j\rangle} x_j \end{cases}$$

Particularly, $\left(X_i \mid x^i,\, X_i > 0\right) \mathrm{sim}\mathcal{E}xp(\theta_2(x^i))$ and
$P\left(X_i = 0 \mid x^i\right) = \frac{\theta_2(x^i)\exp\{-\theta_1(x^i)\}}{1+\theta_2(x^i)\exp\{-\theta_1(x^i)\}}$.

Application: We now assume that the context is rainfall data. We note $x_i = 0$ when it does not rain at the site i, and $x_i > 0$ otherwise. The model should satisfy conditions such that the rain increases with $v_i(+)$, and is decreasing w.r.t $v_i(0)$, where $v_i(+)$ and $v_i(0)$ are the cumulated height of rainfall on the neighbor sites and the number of neighbor sites where it does not rain. This implies the following constraints on the parameters:

$a \in \mathbb{R}$, $c > 0$, $d \leq 0$, $b > -4d$, $e = 0$. We remark here that $e = 0$; we then propose other models involving $e \neq 0$, which induces cooperation. One solution is to consider a censored or a truncated exponential distribution on the positive component, i.e the state space is $E = \{0\} \cup]0, K]$ where K is a fix positive constant. This model is then admissible without any condition on the parameters and therefore permits to introduce cooperation between neighbor sites, via parameter $e \neq 0$. Another solution which we propose in the following example is to apply the mixed auto-model feature.

Example 2 : Double mixed exponential auto-model:

$E = \{0\} \cup]0, K[\cup \{K\}$. We are in the context of a 3-dimensional variable: let p, $q \in]0, 1[$; we set $X = 0$ with probability p, $X = K$ with probability q, and $X \in]0, K[$ with probability $1 - p - q$, according to an exponential distribution on this interval. Again, Theorem 1 ensures the model is well defined; moreover, the model is admissible and allows cooperation between neighbor sites.

3.2 Gaussian positive auto-model

We now suppose that the conditional distributions belong to the family of positive Gaussian mixed-state distribution $G(p_i(.), \sigma_i^2(.))$ given above. Then, there exist a family of vectors $\alpha_i = (a_i, b_i)^t$, and matrices $\beta_{ij} = \begin{pmatrix} c_{ij} & d^*_{ij} \\ d_{ij} & e_{ij} \end{pmatrix}$ verifying $c_{ij} = c_{ji}$, $e_{ij} = e_{ji}$ and $d_{ij} = d^*_{ji}$ such that we can write the global energy as:

$$\begin{aligned} U(x) = \textstyle\sum_{i\in S} a_i\delta(x_i) - \sum_{i\in S} b_i x_i^2 + \sum_{\langle i,j\rangle} c_{ij}\delta(x_i)\delta(x_j) \\ - \textstyle\sum_{(i,j):\langle i,j\rangle} d_{ij} x_i^2 \delta(x_j) + \sum_{\langle i,j\rangle} e_{ij} x_i^2 x_j^2 \end{aligned} \tag{9}$$

Let us describe in more details the local distributions. By construction, in each site i, the conditional distribution is $G(p_i(.), \sigma_i^2(.))$ with parameters

$\theta_{i,1}(.) = a_i + \sum_{j\neq i}[c_{ij}\delta(x_j) - d^*_{ij}x_j^2]$

$\theta_{i,2}(.) = b_i + \sum_{j\neq i}[d_{ij}\delta(x_j) - e_{ij}x_j^2]$

Particularly, $\theta_{i,2}(.) = \frac{1}{2\sigma_i^2(.)}$ et $p_i(.) = \frac{2\exp\theta_{i,1}(.)}{\sqrt{\pi/\theta_{i,2}(.)}+2\exp\theta_{i,1}(.)}$. It follows that necessarily for all i and its possible neighboring configuration $(.) = (x_j,\ j \neq i)$, the variance parameter of the Gaussian component must be positive i.e. $\frac{1}{2\sigma_i^2(.)} > 0$.

Proposition 2 *We assume that U satisfies the following condition* **(B)** *:*

$$(\mathbf{B}) : \begin{cases} \forall i \in S, \forall A \subset S \setminus i,\ b_i + \sum_{j\in A} d_{ij} > 0 \\ \forall i, j \in S,\ e_{ij} \leq 0 \end{cases} \tag{10}$$

Then the energy U is admissible. Consequently, the associated positive Gaussian auto-model is well defined.

See [Bouthemy et al., 2004] for the proof.

Let us now describe the particular model using the four nearest neighbors system ; we denote here by $\{i \pm (1,0),\ i \pm (0,1)\}$ the four neighbors of i ; furthermore, we assume that the field is homogeneous in space, i.e. the parameters are the same for all sites. Moreover, we will allow possible anisotropy between the horizontal and vertical directions. Under all these considerations and by the previous results, there exist a vector $\alpha = (a, b)$ and two 2×2 matrices

$$\beta^{(k)} = \begin{pmatrix} c_k & d^*_k \\ d_k & e_k \end{pmatrix} \qquad ,\quad k = 1, 2$$

such that $\forall i,\ \ \alpha_i = \alpha$, $\forall\{i,j\},\ \ \beta_{ij} = 0$ unless i and j are neighbors where

$$\beta_{ij} = \beta^{(1)} \text{ for } j = i \pm (1,0), \qquad \beta_{ij} = \beta^{(2)} \quad \text{for } j = i \pm (0,1)$$

We need further to set parameters d^*_1, d^*_2, e_1, e_2 to zero, since otherwise we get a repulsive field with neighbor sites in competition which is not suited to the homogeneous motion textures we intend to analyze below.The model has then 6 parameters $(a,\ b,\ c_1,\ c_2,\ d_1, d_2)$.

Now we come for an application to video sequences. Temporal textures (or dynamic textures) designate video contents involving natural (almost stationary) dynamic phenomena such as rivers, sea waves, moving foliage, etc. Mixed state auto-models allow us to specify non linear models, to take into account the spatial context and to introduce both symbolic information (no motion) and continuous motion values, which is of great interest to handle dynamic pictures; we do not model the time-varying intensity function but the motion measurements themselves.

In order to evaluate the performance of the proposed modeling, we examine if the introduced auto-models can realize two fundamental characteristics of a homogeneous texture, namely spatial isotropy and spatial stationarity. For the positive Gaussian auto-models used here, isotropy occurs if (and only if) $c_1 = c_2$ and $d_1 = d_2$. The admissibility condition given in the former result is then reduced to the unique simple condition $b > 0$.

In each experiment, we estimate the parameters by the usual pseudo-likelihood method; this method has good consistency properties for classical one-parameter auto-model and we conjecture that it is still the case for the multi-parameters auto-models considered here. The full description and discussion of the empirical results can be found in [Bouthemy et al., 2004].

The first experiment is to consider motion from trees, which is believed to be spatially isotropic, and close-up shots of a moving escalator, which is clearly anisotropic (vertical motion). In the first case, we fit both the 6-parameter (a, b, c_1, c_2, d_1, d_2) anisotropic (positive Gaussian) auto-model and the 4-parameter (a, b, c, d) isotropic one. The obtained estimates of c_1 and c_2 in one hand, and of d_1 and d_2 in another hand are almost identical, and are moreover very close to the estimated values obtained for c and d in the istropic feature. While for the moving escalator, we get significant differences between c_1 and c_2 as well as between d_1 and d_2.

The second experiment was conducted to analyze spatial sationarity. For a given texture, we divide the motion map into 12 blocks of the same size and fit an anisotropic positive gaussian auto-model to each block. This has been applied to sea-waves images and to a river motion texture. The obtained results show that the 12 sets of the estimated parameters for the sea waves texture are nearly the same, reflecting the expected spatial stationarity; while they are significantly different for the river, which confirms the assumption of non spatial stationarity for this kind of motion texture.

4 Conclusion

We have introduced a new class of random field models, namely mixed state auto-models. This approach is made possible by extending Besag's one parameter auto-models to the multi-parameter case. We provide a construction of these models and show via the given examples how useful and promising these mixed state auto-models can be; we point out for instance their performance to realize some fundamental characteristics of an homogeneous dynamic motion texture. We are currently developing other applications of these new auto-models, namely fitting pluviometric measures; there are many other possible applications in various domains, as soon as the data involves both discrete and continuous components.

There are still several questions which need further investigations; first, the convergence of the pseudo-likelihood has to be established; also, some efficient Monte Carlo simulation algorithms have to be designed for these mixed auto-models.

References

[Allcroft and Glasbey, 2003]D.J. Allcroft and C. Glasbey. A latent gaussian Markov random field model for spatio-temporal rainfall disagregation. *Applied Statistics 52*, pages 487-498, 2003.

[Arnold et al., 1999]B.C. Arnold, E. Castillo, JM Sarabia. *Conditional specification of statistical models*, Springer verlag, New York,1999.

[Besag, 1974]J. Besag. Spatial interactions and the statistical analysis of lattice systems. *JRSS B*, 148**,** pages 1-36, 1974.

[Bouthemy et al., 2004]P. Bouthemy, C. Hardouin, G. Piriou, J. Yao. Auto-models with mixed states and analysis of motion textures. *Preprint IRISA n° 1682,* 2005.

[Guyon, 1995]X. Guyon. *Random fields on a network: Modeling, Statistics and applications.* Springer-Verlag, New York, 1995.

[Guyon and Hardouin, 2002]X. Guyon, C. Hardouin. Auto modèles markoviens à états mixtes, *Annales des journées de statistique SFDS*, *Bruxelles, 2002.*

[Hardouin and Yao, 2004]C. Hardouin and J. Yao. Markovian auto-models with mixed states. *Technical report, IRMAR, Université de Rennes 1*, 2004.

[Whittle, 1963]P. Whittle. Stochastic processes in several dimensions. *Bull. Inst. Statist. Inst. 40,* pages 974-994, 1963.