# Multivariate auto-models

# Application to mixed states data

C. Hardouin, J-F Yao

SAMOS - Université Paris 1

IRMAR Université Rennes 1

$X = \{X_i,\ i \in S\}$ champ sur $S = \{1, ..., n\}$

Distributions conditionnelles $p_i(x_i \mid\ .) = p_i(x_i \mid x_j,\ j \neq i)$

Pb : spécifier les $\{p_i\}$ pour qu'il existe une loi jointe $P$ compatible.

On s'intéresse au cas des auto-modèles.

- Auto-modèles multivariés

- Application à des conditionnelles Beta, modèles coopératifs

- Application à des variables à états mixtes; $E = \{0\} \cup ]0; +\infty[$.

- Expérimentation sur des mesures de mouvement de séquences vidéo.

# 1. Besag auto-models : multivariate extension

Finite set of sites $S = \{1, ..n\}$,

$\mathcal{G}$ : a symmetric graph on $S$, $\langle i, j \rangle$ a pair of neighbouring sites.

$X = \{X_i,\, i \in S\}$ is a Markovian field on $\Omega = E^S$, $E \subset \mathbb{R}^d$

$X$ is specified by a probability distribution $\mu$ on $\Omega$, $\mu$ has an everywhere positive density $P$ .

$$P(x) = Z^{-1} \exp Q(x) \;, \tag{1}$$

where $Z$ is a normalization constant.

Hammersley-Clifford : $Q$ is a sum of potentials $G$ deduced from a set of cliques.

For a site $i$, let $p_i(x_i|\cdot) = \mu_i(x_i|x_j, j \neq i) = p_i(x_i \mid x_j,\ j \in \partial i)$ ,be the probability (positive) density of $X_i$ given $\{X_j = x_j, j \neq i\}$.

Besag's auto-models are constructed under two assumptions:

**[B1]** pairwise-only dependence.

$$Q(x) = \sum_{i \in S} G_i(x_i) + \sum_{\{i,j\}} G_{ij}(x_i, x_j) \ .$$

Reference configuration $\tau = (\tau_i) \in \Omega$. Identifiability conditions: for all $i$, $j \in S$ and $v \in E$,

$$G_i(\tau_i) = G_{ij}(\tau_i, v) = G_{ij}(v, \tau_j) = 0 \ .$$

**[b2]** $\log p_i(x_i|\cdot) = A_i(\cdot)B_i(x_i) + C_i(x_i) + D_i(\cdot)$ ,

et pour chaque $i$, $B_i$ linéaire en $x_i$

Extensions :

- Lee, Kaiser, Cressie 2001 : Multiway dependence in exponential family conditional distributions *JMVA 79*

- Kaiser, Cressie 2000, The construction of multivariate distributions from MRF *JMVA 73*

- Paramètre multidimensionnel pour la famille exponentielle

**[B2]** $\log p_i(x_i|\cdot) = \langle A_i(\cdot), B_i(x_i)\rangle + C_i(x_i) + D_i(\cdot)$ , $A_i(\cdot) \in \mathbb{R}^d$, $B_i(x_i) \in \mathbb{R}^d$.

with the normalization conditions: for all $i \in S$, $\quad B_i(\tau_i) = C_i(\tau_i) = 0$

**Theorem**

*Assume conditions [B1], [B2] and the following regularity condition*

*[C]* *For all* $i \in S$, *Span* $\{B_i(x_i),\ x_i \in E\} = \mathbb{R}^d$.

*Then there exist a family of* $d$*-dim vectors* $\{\alpha_i,\ i \in S\}$ *and a family of* $d \times d$ *matrices* $\{\beta_{ij},\ i, j \in S,\ i \neq j\}$ *satisfying* $\beta_{ij}^T = \beta_{ji}$ *such that*

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

*Consequently the potentials are given by*

$$G_i(x_i) = \langle \alpha_i, B_i(x_i) \rangle + C_i(x_i) \qquad (3)$$

$$G_{ij}(x_i, x_j) = B_i^T(x_i) \beta_{ij} B_j(x_j) \qquad (4)$$

Le thm détermine la forme des paramètres canoniques i.e des combinaisons linéaires des stats exhaustives pour reconstruire $P$.

Rmq : Pas d'hypothèse de symétrie sur le modèle, on peut avoir des graphes orientés ou non.

Champ spatialement symétrique : les matrices $\beta_{ij}$ sont symétriques

**Proposition**

*Assume that the function $Q$ is defined in [B1] with potentials $G_i$ and $G_{ij}$ given in (3) - (4), and that it is admissible i.e $\int_\Omega \exp Q(x)\nu(dx) < \infty$. Then the family of conditional distributions $p_i(x_i \,|\, .)$ belongs to an exponential family of type [B2], with $A_i(.)$ verifying (2).*

Rmq : Conditions d'admissibilité : au cas par cas

## Application to a pair of Gaussian and Gamma variables

$X_1 \mid X_2 \sim \; Gamma$ on $E_1 = [0, \infty[$

$X_2 \mid X_1 \sim \; Gaussian$ on $E_2 = \mathbb{R}$

Gamma distribution on $E_1$ with $a, b > 0$ ;
$f_{\theta_1}(x) = \exp\{\langle \theta_1, B_1(x) \rangle - \psi_1(\theta_1)\}$,
where $\theta_1 = (b, a-1)^t$ and $B_1(x) = (-x+1, \log x)^t$.
We choose $\tau_1 = 1$, $B_1(\tau_1) = 0$.

Gaussian density on $E_2$: $g_{\theta_2}(x) = \exp\langle \theta_2, B_2(x) \rangle - \psi_2(\theta_2)$
where $\theta_2 = (m/\sigma^2, -(2\sigma^2)^{-1})$ and $B_2(x) = (x, x^2)^t$.
We choose $\tau_2 = 0$, $B_2(\tau_2) = 0$.

We consider now $(X_1, X_2)$ valued in $E_1 \times E_2$. The reference configuration is $\tau = (1, 0)$.

$$\log p_1(x_1|x_2) = \log f_{\theta_1(x_2)}(x_1) = \langle A_1(x_2), B_1(x_1) \rangle - D_1(x_2) \; ,$$
$$\log p_2(x_2|x_1) = \log g_{\theta_2(x_1)}(x_2) = \langle A_2(x_1), B_2(x_2) \rangle - D_2(x_1) \; .$$

From Theorem, there exists two vectors $\alpha_1,\ \alpha_2$ and one $(2 \times 2)$-matrix $\beta$ such that

$$A_1(x_2) = \alpha_1 + \beta B_2(x_2)\ , \quad A_2(x_1) = \alpha_2 + \beta^t B_1(x_1)$$

The joined density is $P(x_1, x_2) = P(\tau) \exp Q(x_1, x_2)$ with

$$Q(x_1, x_2) = \langle \alpha_1, B_1(x_1) \rangle + \langle \alpha_2, B_2(x_2) \rangle + B_1(x_1)^t \beta B_2(x_2)\ .$$

$\beta$ is not necessarly symmetrical. The model contains 8 parameters.

Explicit conditions on these parameters can be obtained in a straightforward way to ensure admissibility of $Q$.

Rmq : on retrouve très simplement un résultat de Arnold, Castillo, Sarabia.

## 2. Cooperative auto-models with Beta conditional distributions

Several common auto-models imply competition between neighbouring sites.

Auto-Poisson: $Q(x) = \sum_i \{\alpha_i x_i - \log(x_i!)\} + \sum_{i,j} \beta_{ij} x_i x_j$

NCS for admissibility: $\beta_{ij} \leq 0 \rightarrow$ the interactions between $i$ and $j$ are competitive.

$X_i \mid \{X_j,\ j \neq i\} \sim Beta(p_i, q_i)$ on $[0, 1]$

$p_i(x_i \mid x_j,\ j \neq i) = \kappa(p_i, q_i)\ x_i^{p_i - 1}\ (1 - x_i)^{q_i - 1}$

$p_i(x_i \mid x_j,\ j \neq i) = \exp\{\langle \theta_i, B(x_i) \rangle - \psi_i(\theta_i)\}$

where $\theta_i = (p_i - 1, q_i - 1)^T$, $B(x) = (\log 2x, \log 2(1 - x))^T$

and $\psi_i(\theta_i) = (p_i + q_i - 2) \log 2 + \log \kappa(p_i, q_i)$.

$X$ random field with such Beta conditionals ; theorem implies there exists vectors $\alpha_i = (a_i, b_i)^T$, matrices $\beta_{ij} = \begin{pmatrix} c_{ij} & d_{ij} \\ d^*_{ij} & e_{ij} \end{pmatrix}$ with $d_{ij} = d^*_{ji}$ such that

$$A_i(.) = \begin{pmatrix} a_i + \sum_{j \neq i} \{c_{ij} \log 2x_j + d_{ij} \log 2(1 - x_j)\} \\ b_i + \sum_{j \neq i} \{d^*_{ij} \log 2x_j + e_{ij} \log 2(1 - x_j)\} \end{pmatrix}$$

and $Q(x_1, ..x_n) = \sum_i \langle \alpha_i, B(x_i) \rangle + \sum_{i,j} B(x_i)^T \beta_{ij} B(x_j)$

The reference configuration is $\tau = (\frac{1}{2}, ..., \frac{1}{2})$.

**Proposition** *This model is well defined and admissible if*

*(i) for all* $i, j$ , $\beta_{ij} = \begin{pmatrix} c_{ij} & d_{ij} \\ d^*_{ij} & e_{ij} \end{pmatrix} \leq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

*(ii) for all* $i$, $1 + a_i + \log 2 \sum_{j \neq i} \{c_{ij} + d_{ij}\} > 0$ *and* $1 + b_i + \log 2 \sum_{j \neq i} \{d^*_{ij} + e_{ij}\} > 0$.

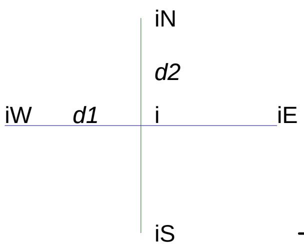
*Spatial cooperation versus competition*

$$E[X_i \mid .] = \frac{1 + A_{i,1}(.)}{2 + A_{i,1}(.) + A_{i,2}(.)}.$$

Cooperation: $c_{ij} = e_{ij} = 0$ Competition: $d_{ij} = d^*_{ij} = 0$

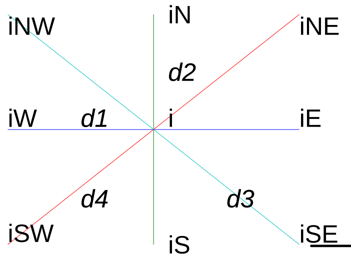
*A cooperative model with the four and eight nearest neighbours system:*

We assume spatial symmetry, spatial stationarity and spatial cooperation




$\longrightarrow$ 4 parameters $(a, b, d_1, d_2)$ with conditions

$$d_1 \leq 0,\ d_2 \leq 0,\ 1 + a + 2(d_1 + d_2)\log 2 > 0,\ 1 + b + 2(d_1 + d_2)\log 2 > 0$$




$\longrightarrow$ 8 parameters $(a, b, d_1, d_2, d_3, d_4)$ with conditions

$d_k \leq 0$ $(k = 1, 4)$ , $1 + a + 2(d_1 + d_2 + d_3 + d_4) \log 2 > 0$, $1 + b + 2(d_1 + d_2 + d_3 + d_4) \log 2 > 0$

Estimation by the Pseudo likelihood method: $L(x; \phi) = \prod_i p_i(x_i \mid x_j,\ j \neq i)$

100 simulations of 600 scans of the Gibbs sampler on a square lattice 64×64,

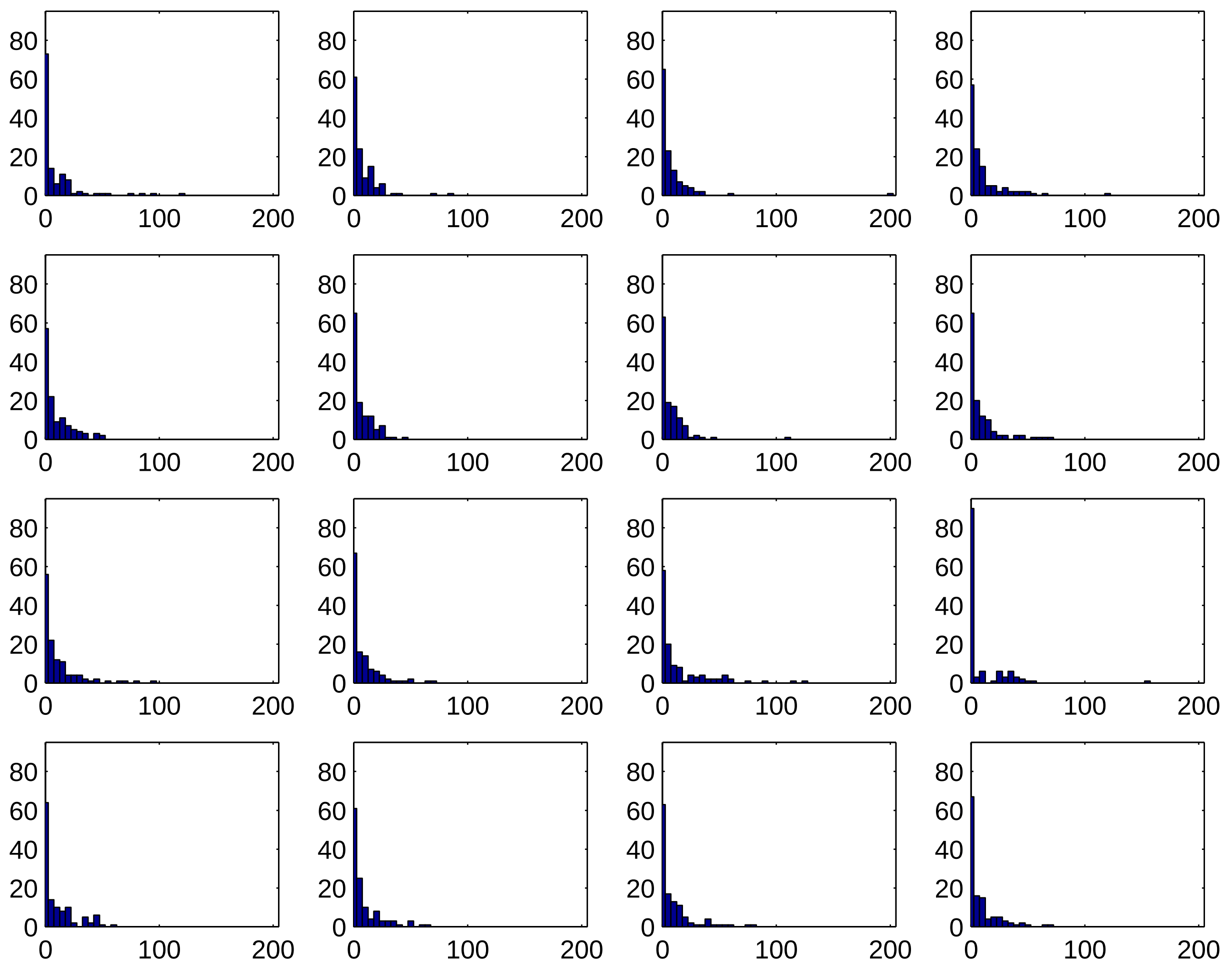
Valeurs : Cressie, Kaiser 2202 : valeurs pour des arbres malades, modèle hiérarchique avec $d_1 = d_2$

| Parameter | $a$ | $b$ | $d_1$ | $d_2$ |
|---|---|---|---|---|
| True values | 16.6 | 18.9 | -4.5 | -4.5 |
| Mean<br>std dev. | 16.6004<br>(0.5847) | 19.0062<br>(0.5872) | -1.4725<br>(0.2742) | -4.5093<br>(0.3153) |

Valeurs arbitraires

| Parameter | $a$ | $b$ | $d_1$ | $d_2$ | $d_3$ | $d_4$ |
|---|---|---|---|---|---|---|
| True values | 12.0 | 16.0 | -1.0 | -3.0 | -0.5 | -2.0 |
| Mean | 12.0353 | 16.0319 | -0.9803 | -3.0295 | -0.5263 | -1.9818 |
| std dev. | (0.5847) | (0.5872) | (0.2742) | (0.3153) | (0.2362) | (0.2602) |

# 3. Markovian Auto-models with mixed states



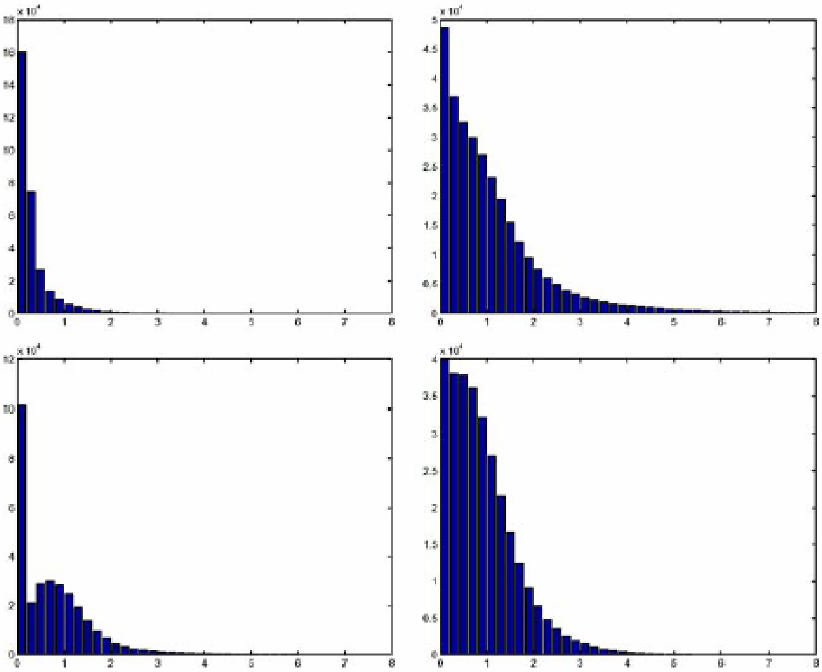


*Figure 4*. Sample histograms of motion measures $\{v_{res}\}$. Top to bottom, and left to right: grass, foliage, trees and sea-waves.

1. **Random variable with mixed states**; **Mixed exponential family** $\mathcal{L}(p, \xi)$ **:**

$$X \in E = \{0\} \cup ]0, +\infty[,$$

$$m(x) = \delta(x) + \lambda(dx) \tag{5}$$

where $\delta$ is the Dirac measure at 0, and $\lambda$ is the Lebesgue measure.

Let $\gamma \in ]0, 1[$ ; then $X = 0$ with probability $\gamma$,

and with $1 - \gamma$, $X > 0$ follows a distribution which belongs to an exponential family, with the probability density :

$$g_\xi(x) = H(\xi) \exp\langle \xi, T(x) \rangle$$

$T$ is defined such as $T(0) = 0$.

We define $\delta^*(x) = 1 - \mathbf{1}_{\{0\}}(x)$.

The probability density of $X$ on $E$ is (w.r.t. $\nu$) :

$$\begin{aligned} f_\theta(x) &= \gamma \mathbf{1}_{\{0\}}(x) + (1-\gamma)\delta^*(x) g_\xi(x) \\ &= Z^{-1}(\theta) \exp\langle \theta, B(x) \rangle \end{aligned}$$

where $\theta = (\theta_1, \theta_2)^T = (\ln \frac{(1-p)H(\xi)}{p}, \xi)^T$

and $B = (\delta^*, T^T)^T$.

One-to-one correspondence between the natural parameter $\theta$ and the original parameters $\xi$ and $\gamma$:

$$\xi = \theta_2 \ , \quad \gamma = \frac{H(\xi)}{H(\xi) + e^{\theta_1}} \ .$$

If $\xi, \ T(x) \in \mathbb{R}^s$, then **dim**(exponential family)$=s+1$.

2. **Generalization :** $E = \{e_1\} \cup \{e_2\} \cup ... \cup \{e_k\} \cup ]-\infty; e_1[\cup]e_1; e_2[\cup]e_k; \infty[$

## 3. Auto-models for mixed state data

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

We assume that the family of the conditional distributions $p_i(x_i|.)$ belongs to the family of mixed distributions $\mathcal{L}(\gamma_i(.), \xi_i(.))$, i.e.

$$\ln p_i(x_i|.) = \langle \theta_i(\cdot), B(x_i) \rangle - D_i(\cdot)$$

with $B(x_i) = (\delta^*(x_i), T(x_i)^T)^T$.

Theorem 1 $\Rightarrow$ there exists a family of $(s+1)-$vectors $\{\alpha_i\}$ and $(s+1) \times (s+1)$ matrices $\{\beta_{ij}\}$ satisfying $\beta_{ij} = \beta_{ji}^T$, such that

$$A_i(.) = \theta_i(.) = \alpha_i + \sum_{j \neq i} \beta_{ij} B(x_j)$$

and the energy function is

$$Q(x) = \sum_{i \in S} \alpha_i^T B(x_i) + \sum_{(i,j):<i,j>} B^T(x_i) \beta_{ij} B(x_j)$$

For each specific family $g_\xi$, we have to determine conditions ensuring the admissibility.

# 4. Application to motion analysis from video sequences

**1. Motion measurements from video sequences**

Let $\{I_i(t)\}$ be an image sequence, where $i = (i_1, i_2)$ are the pixel locations and $t = 1, .., T$ time instants in the sequence.

A motion map at time $t$ is $X(t) = \{X_i(t)\} = \{\| v_i(t) \|\}$ where the motion field $\{v_i(t)\}$ is estimated by a regularised minimisation of $\sum[I_{i+v_i(t)}(t+1) - I_i(t)]^2$.

We consider video sequences of natural scenes.

**2.Gaussian positive auto-model.**

Mixed positive Gaussian distribution $\mathcal{G}(\gamma, \sigma^2)$

With probability $\gamma$, $X = 0$ and with $1 - \gamma$, $X = |Z|$ where $Z \sim N(0, \sigma^2)$.

$$f(x) = \exp\{\langle \theta, B(x) \rangle + \log \gamma\}$$

with $\theta = (\theta_1, \theta_2)^T = (\log \frac{2(1-\gamma)}{\gamma\sigma(2\pi)^{1/2}}, \frac{1}{2\sigma^2})^T$ and $B(x) = (\delta^*(x), -x^2)^T$.

We get also $\sigma^2 = \frac{1}{2\theta_2}$ and $\gamma = \frac{2}{2+(2\pi\theta_2)^{1/2}\exp\theta_1}$.

Auto-model with mixed positive Gaussian distributions

$X \in \Omega = (\{0\} \cup ]0, +\infty[)^S$.

$p_i(x_i|.) \in \mathcal{G}(\gamma_i(.), \sigma_i^2(.))$.

Theorem 1 $\Rightarrow \exists \alpha_i = (a_i, b_i)^t$ , $\beta_{ij} = \begin{pmatrix} c_{ij} & d_{ij}^* \\ d_{ij} & e_{ij} \end{pmatrix}$ with $d_{ji} = d_{ij}^*$,

$$Q(x) = \sum_{i \in S} \left( a_i\delta^*(x_i) - b_i x_i^2 \right) + \sum_{\{i,j\}} \left( c_{ij}\delta^*(x_i)\delta^*(x_j) - d_{ij}x_i^2\delta^*(x_j) + e_{ij}x_i^2x_j^2 \right)$$

$A_i(.) = \theta_i(.) = \alpha_i + \sum_{j \neq i} \beta_{ij} B(x_j),$

$\theta_{i,1}(.) = a_i + \sum_{j \neq i} \left( c_{ij} \delta^*(x_j) - d_{ij} x_j^2 \right)$

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

Admissibility : necessarly, for all $i$, $\sigma_i^2(.) > 0$ and $p_i(.) \in ]0; 1[$.

Proposition *Under the following condition* $(A)$*, the energy* $Q$ *is admissible :*

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

## 3. Experimental results

Analysis of motion measurements $X(t) = \{X_i(t)\} = \{\| v_i(t) \|\}$

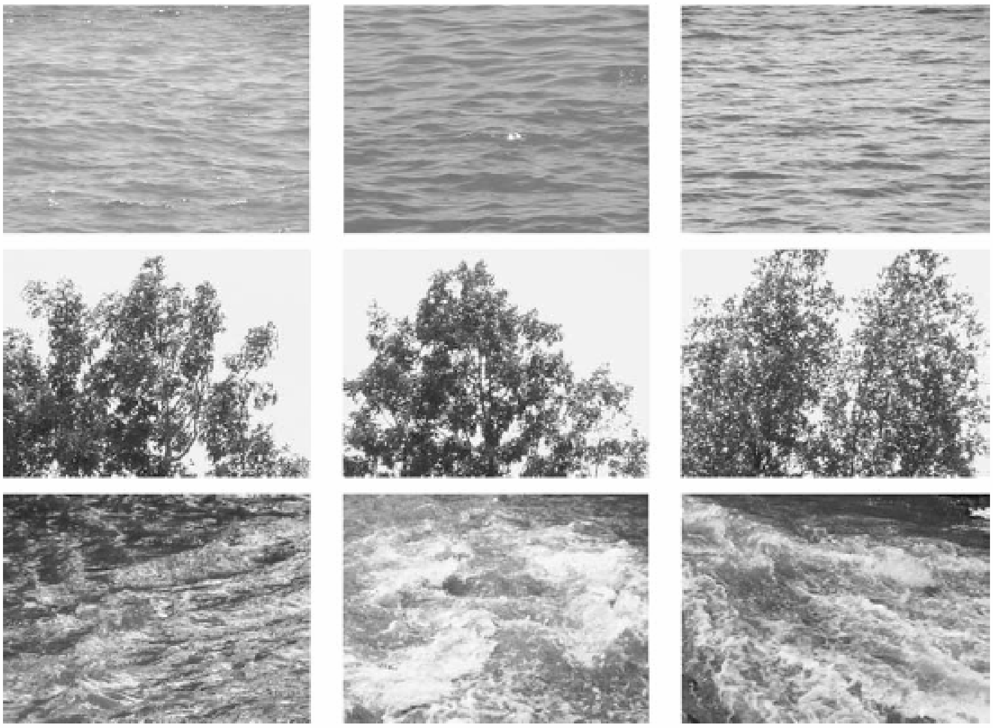
**Model:** Positive Gaussian mixed state auto-model,

$+$ four nearest neighbours system,

$+$ homogeneity in space, spatial symmetry, cooperation, possible anisotropy between the two directions.

Parameter $\phi = (a, b, c_1, c_2)$; admissibility condition: $b > 0$;

Estimation by pseudo likelihood



*gure 2*. Sample images from different videos of natural dynamic scenes. Top to bottom: moving grass, foliage, sea-waves, trees and rivers.

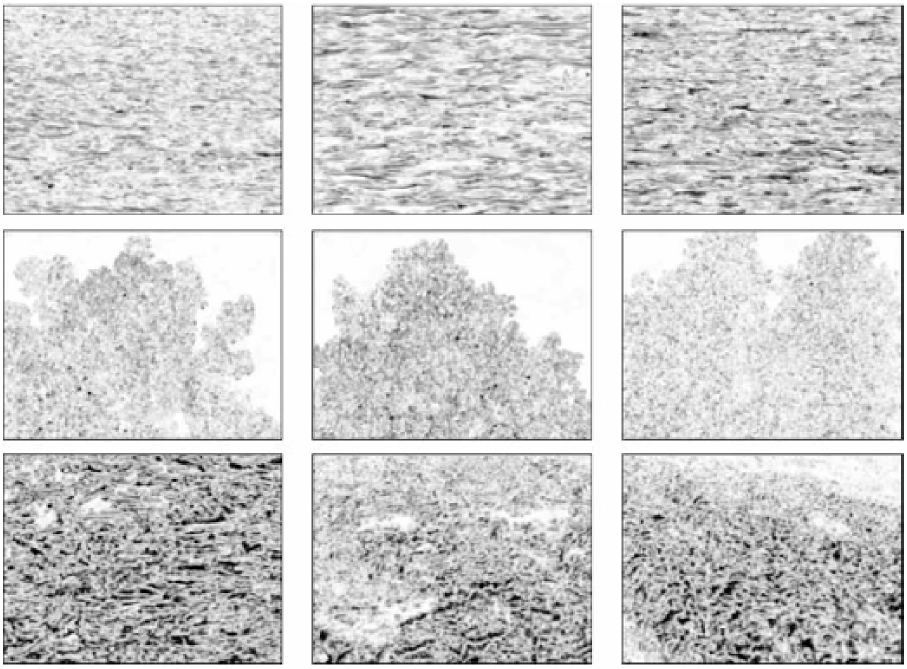


*Figure 3*. Sample motion measures $\{v_{res}\}$ from the videos of Fig. 2. Top to bottom: grass, foliage, sea-waves, trees and rivers (white = 0; black = maximum value).

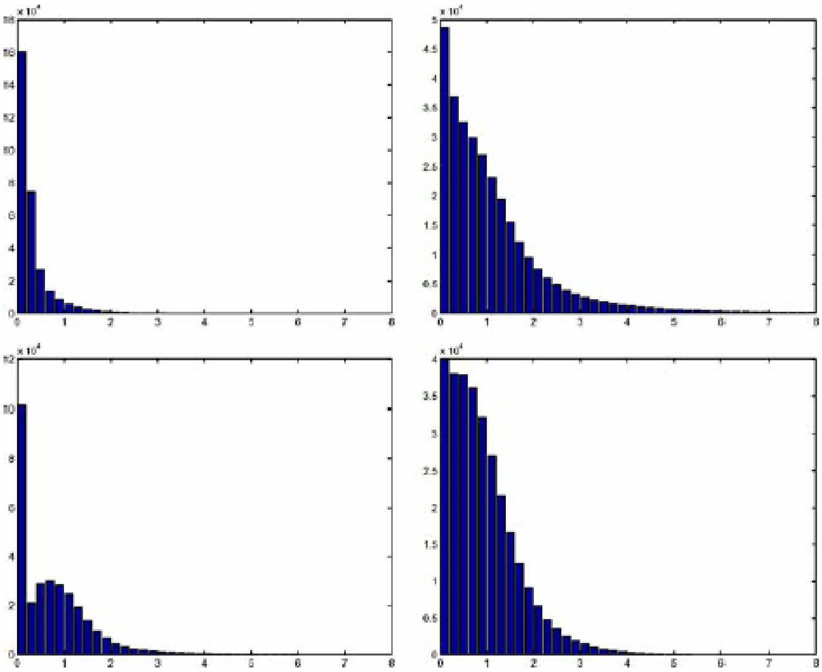


*Figure 4*. Sample histograms of motion measures $\{v_{res}\}$. Top to bottom, and left to right: grass, foliage, trees and sea-waves.

1. Spatial isotropy occurs if $c_1 = c_2$.

Motion from trees :

a typical set is $\hat{\phi} = (\hat{a}, \hat{b}, \hat{c}_1, \hat{c}_2) = (, , , )$.

| Full model | $a$ | $b$ | $c_1$ | $c_2$ |
|---|---|---|---|---|
| | $-5.805$ | 3.044 | 3.057 | 2.954 |
| Isotropic model | $a$ | $b$ | $c$ | |
| | $-5.781$ | 3.044 | 3.000 | |

Motion from sea waves :

| Full model | $a$ | $b$ | $c_1$ | $c_2$ |
|---|---|---|---|---|
| | $-7.941$ | 3.370 | 5.792 | 1.422 |

2. Spatial stationarity feature.

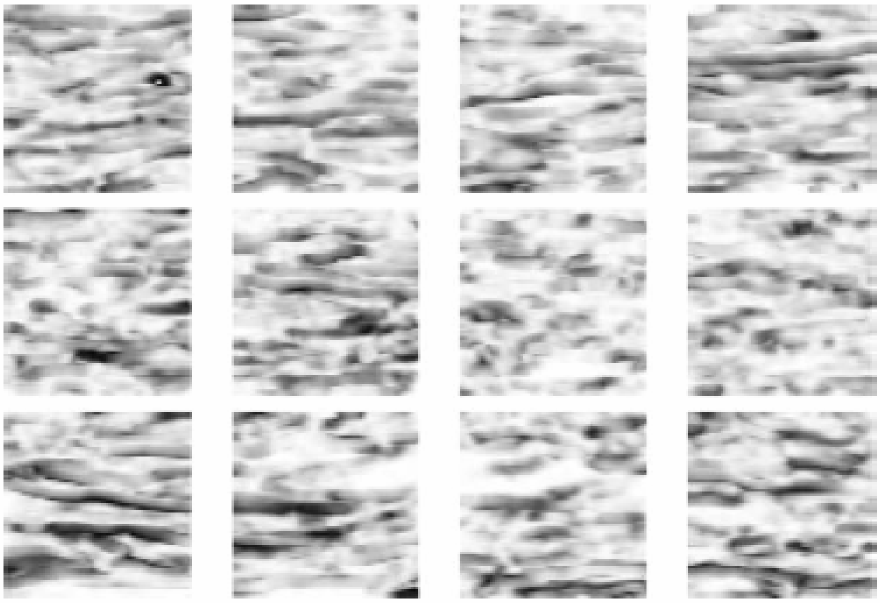


*Figure 11*. Sea-waves sequence: Set of disjoint blocks of local motion measures of size $65 \times 65$ at a given time instant of the sequence. To
to bottom and left to right: $B_1 \ldots B_{12}$.

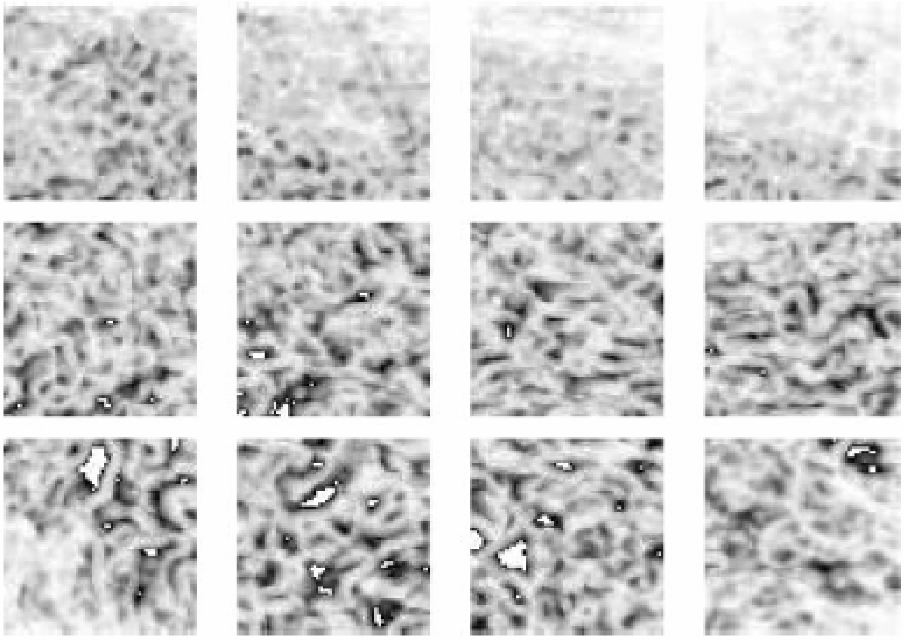


*Figure 12*. River sequence: Set of disjoint blocks of local motion measures of size $65 \times 65$ at a given time instant of the sequence. Top t
bottom and left to right: $B_1 \ldots B_{12}$.

The motion map is divided into 12 disjoint blocks and the model is fitted to each block.

Motion from sea-waves :

| | $a$ | $b$ | $c_1$ | $c_2$ |
|---|---|---|---|---|
| $B_1$ | -9.3021 | 0.2969 | 5.8790 | 2.1368 |
| $B_2$ | -9.3995 | 0.3288 | 5.4686 | 2.7248 |
| $B_3$ | -9.0482 | 0.3415 | 7.2315 | 1.1405 |
| $B_4$ | -9.6020 | 0.3290 | 7.3358 | 1.3610 |
| $B_5$ | -8.9100 | 0.3710 | 5.6541 | 2.0467 |
| $B_6$ | -7.3573 | 0.3996 | 5.7413 | 1.1077 |
| $B_7$ | -7.5743 | 0.4395 | 5.2463 | 1.7163 |
| $B_8$ | -7.4782 | 0.5879 | 5.0888 | 1.8579 |
| $B_9$ | -8.3047 | 0.3627 | 6.3699 | 1.1809 |
| $B_{10}$ | -7.6136 | 0.3017 | 6.4159 | 0.7092 |
| $B_{11}$ | -8.8630 | 0.2863 | 7.5933 | 0.6516 |
| $B_{12}$ | -8.8784 | 0.3287 | 5.8394 | 1.8503 |
| St D | 0.8220 | 0.0830 | 0.8403 | 0.6203 |

Motion from river:

| | $a$ | $b$ | $c_1$ | $c_2$ |
|---|---|---|---|---|
| $B_1$ | -10.0262 | 0.3619 | 7.7689 | 3.4854 |
| $B_2$ | -8.3930 | 0.4458 | 5.7136 | 1.9323 |
| $B_3$ | -7.1043 | 0.6909 | 3.5173 | 3.4653 |
| $B_4$ | -5.6905 | 0.8777 | 3.2295 | 2.4021 |
| $B_5$ | 17.1302 | 0.1196 | 21.9513 | 10.8203 |
| $B_6$ | 8.1796 | 0.1142 | 13.3851 | 7.9190 |
| $B_7$ | 8.2705 | 0.1067 | 13.1800 | 8.3103 |
| $B_8$ | 8.1845 | 0.1281 | 13.5695 | 8.0117 |
| $B_9$ | -11.7890 | 0.1120 | 9.7920 | 1.0696 |
| $B_{10}$ | 8.0416 | 0.0751 | 13.6809 | 7.8909 |
| $B_{11}$ | -3.5098 | 0.0994 | 11.5152 | -4.1550 |
| $B_{12}$ | 12.9198 | 0.1130 | 4.7471 | 5.1304 |
| St D | 10.1118 | 0.2690 | 5.6741 | 4.1530 |

# Conclusions

- Consistency

- Mixed state auto-models

- Applications to image analysis, epidemiology, pluviometry....

- Dynamics