Statistical Models for Road Traffic Forecasting

J-M. Loubes
Collaboration with: Guillaume Allain, Thibault Espinasse, Fabrice Gamboa, Jean-Noël Kien

INRIA-12th of June

Overview

Overview

The actors of the collaboration

- Mediamobile Vtrafic
  Mediamobile ensures the production and broadcasting of reliable and pertinent real-time traffic information. Founded in 1996, Mediamobile originated from a partnership between TDF Group (European leader in media content broadcast) and the automotive manufacturer Renault in the framework of a European Program for Research and Development of Intelligent Transportation.
- Institut de Mathématiques de Toulouse
  The Toulouse Mathematics Institute, CNRS Research Laboratory, federates the mathematics community of the Toulouse area. One of the biggest mathematical team in France (around 400 people)

People involved in the collaboration

Six years collaboration leading to three patents. Actual people involved

- Mediamobile Vtrafic
    - Philippe Goudal head of the prediction department
    - Guillaume Allain Engineer has been Engineer/CIFRE Ph. D Student of the project
    - Jean-Noël Kien Engineer/CIFRE Ph. D Student
- Institut de Mathématiques de Toulouse
    - Fabrice Gamboa Professor
    - Jean-Michel Loubes Professor
    - Elie Maza AssistantProfessor
    - Thibault Espinasse AssistantProfessor
    - Jean-Noël Kien Engineer/CIFRE Ph. D Student

Overview

Mediamobile's task

- **Gathering** raw traffic information
- **Processing** and agregating
- **Broadcasting** (radio, www, mobile device...)

$\Rightarrow$ Fancy new services : **forecasting** and **dynamic routine**

Industrial constraints :

- **coverage** $\begin{cases} \text{each road of the network} \\ \text{from real time to long run} \end{cases}$
- **quality/accuracy** $\begin{cases} \text{controlled speed prediction error} \\ \text{controlled jam prediction error} \end{cases}$
- **user friendly** $\begin{cases} \text{automatable} \\ \text{adaptative} \\ \text{easy to update} \end{cases}$

Road traffic data-Road network

What is a road network ?

- Graph composed of a set of pair (**edges**,**vectices**)
- **Complexity** of the graph $\rightarrow$ *Functional Road Classes (FRC)*
- **FRC** $\rightarrow$ road type classification (arterial, collector, local road...)

FRC	**Number of edges**	$\sum \mathrm{L[km]}$
0	46 175	22 580
1	232 572	42 793
2	462 907	75 453
3	998 808	175 790
{0,1,2,3}	1 740 462	316 616

Tab: Number of edges by FRC

- Network coverage depends on the FRC

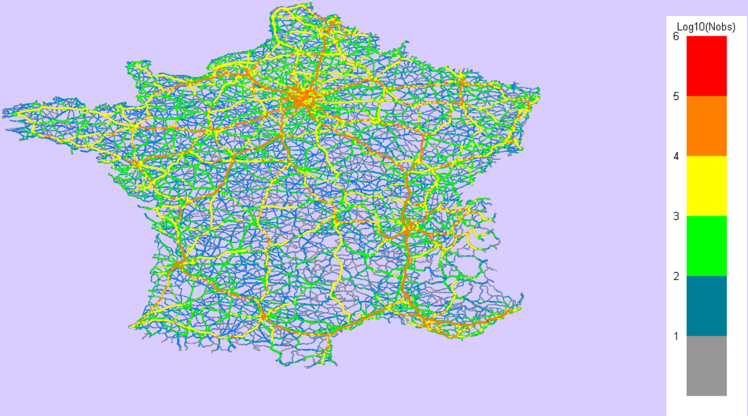
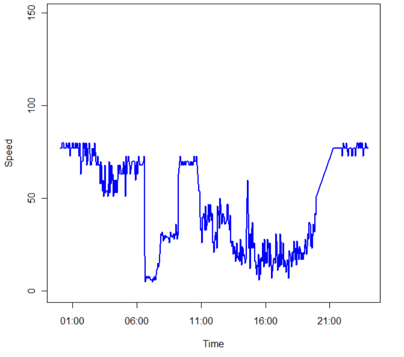


Fig: Network coverage by all FRC {0,1,2,3} from 03/01/2009 to 05/31/2009

Speed data

What is a speed data ?
Loop sensor

- speed calculated from flow and density (conservation law)



Pros

- More accurate
- 3min constant frequency

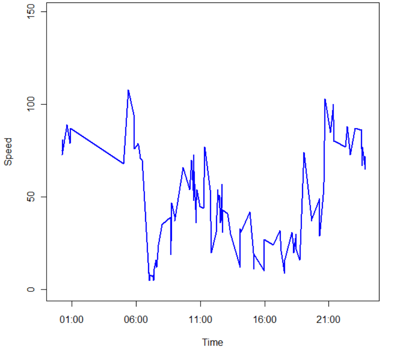
Cons

- Located only in main roads
- Thresholded at national speed limits

Speed data

GPS sensor : Floating Car Data

- positions are mapped on a graph → building speeds



Pros

- Can potentially cover all the graph
- Raw source of data

Cons

- Less accurate → GPS and map-matching error
- More variable → outlier emergence
- Random frequency → user feedback

Overview

Local road trafficking forecasting with ℓ^1

Our Goal

- Appoach the road traffic dynamic with local statistical models

$$V(s_q, t_{p+h}) = F(Q(s,t), \rho(s,t)\ldots) \rightarrow V_{q,p+h} = g_{q,p,h}(\underbrace{\{V_{i,k}; i \in G, k \in T\}}_{X})$$

Problems

- High dimension of X
- All $V_{i,k}$ not influent

Solution

- Regularization
- Selection

Modelizing traffic dynamic with significative effects **only**

$$V_{q,p+h} = g_{q,p,h}(V_{i,k}) \rightarrow V_{q,p+h} = \sum_{i \in G, k \in T} \beta_{i,k}.V_{i,k}$$

where

$$\widehat{\beta_{i,k}} = \underbrace{K((i,k),(q,p+h))}_{\textbf{Kernel}}$$

Kernel selection : fit road traffic dynamic

- learning a sparse set of influence parameters

$$\widehat{\beta} = \arg\min_{\beta} \left(\|V_{q,p+h} - \sum_{i \in G, k \in T} \beta_{i,k}.V_{i,k}\|^2 + \lambda \sum |\beta_{i,k}| \right)$$

Conclusion

- Short run local model
- Forecast and complete missing data
- Time and spatial road traffic dynamic used
- Exists block version to privilegiate certain axis

Improve accuracy of short/long run predictions with weather data

Partnership between Mediamobile and Meteo-France

Rupture model

$$V(x, t_1) = V(x, t_0) + C(.) \times \mathbb{1}_{M(x,t_1) \neq M(x,t_0)} \quad \text{with } t_1 - t_0 < \tau_{sta}$$

- τ_{sta} : timespan for a stationary traffic flow

$C(.)$ correction term can depend on :

- edge x : road specifications, geographical areas
- nature and intensity of the weather evolution
- traffic state at t_0 : $V(x, t_0)$

Model selection based on C(.) structure

Linear thresholded biais model

If $V(x, t_0) \geqslant \alpha$,

$$V(x, t_1) = V(x, t_0) - \underbrace{\beta}_{\text{correction term}} .(V(x, t_0) - \underbrace{\alpha}_{\text{break parameter}})$$

Or else,

$$V(x, t_1) = V(x, t_0)$$

Advantages

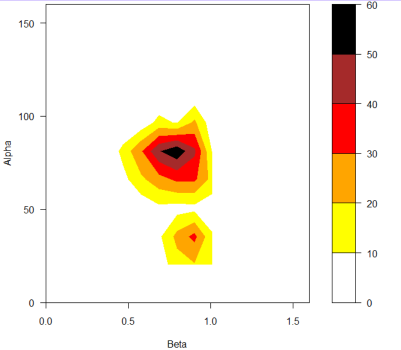
- takes traffic state into consideration
- thresholded model yields interpretable model

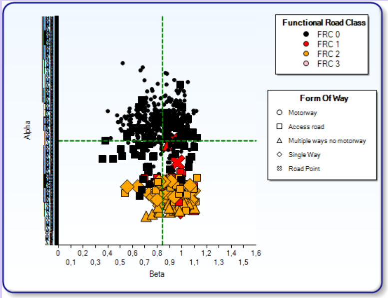
Drawback

- edge by edge model

Network generalization

Repartition of (α, β) parameters





Results

- β can be generalized
- Repartition of α depends on the FRC

Results with 10 000 random edges FRC 0

Weather condition	α **(% of FreeFlowSpeed)**	β	**# obs**
Low rain	93%	0,95	7236
Medium and Strong rain	90%	0.95	3316
Freezing rain	NA	NA	0
Rain and Snow	94%	0.97	1011
Snow	83%	0.96	2621
Hail	NA	NA	0
Drizzle	89%	0.90	615

For instance, let the free flow speed equals 100 km/h : a car travels at 130 km/h on a freeway and strong rain appears.
Since $130 > 90\%.100$, car speed **decreases** to $130 - 95\%.(130 - 90) =$ **92 km/h**.

Model the relationship between speeds and calendar

How it is used :

- → $D \neq$ day of the prediction
- → the speed curve is not observed

$\Big\}$ $\forall p, h \;\; X = C$

- → « Inboard configuration » $\Rightarrow$ low **complexity**

mathematical model : linear model with k fixed

$$g(t_k, x) = \beta_0 + \left.\begin{array}{l} \beta_1 \mathbb{1}_{\{c=\text{Monday}\}} + \beta_2 \dots \\ \beta_8 \mathbb{1}_{\{c=\text{January}\}} + \beta_{19} \dots \\ \beta_{20} \mathbb{1}_{\{c=\text{Hollidays}\}} + \dots \end{array}\right\} \quad \textbf{one oder effects}$$

$$+ \left.\begin{array}{l} \beta_{1,8} \mathbb{1}_{\{c=\text{Monday}\cap\text{January}\}} + \dots \\ \beta_{1,20} \mathbb{1}_{\{c=\text{Monday}\cap\text{Holli}\}} + \dots \\ \dots \end{array}\right\} \quad \textbf{Second order effects}$$

2

Drawbacks :

- → Functionnal aspects are lost
- → $(N+1) \times K$ effects

Model by classification

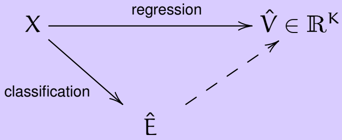
- **Functional mixture model**
- speed curve V is represented as a **finite number of patterns** :

$$f_1, \dots, f_i, \dots, f_m \text{ avec } f_i \in \mathbb{R}^K$$

$$V = \sum_{i=1}^{m} \mathbb{1}_{E=i}\, f_i + \epsilon_i \text{ et } f^\star = f_E$$

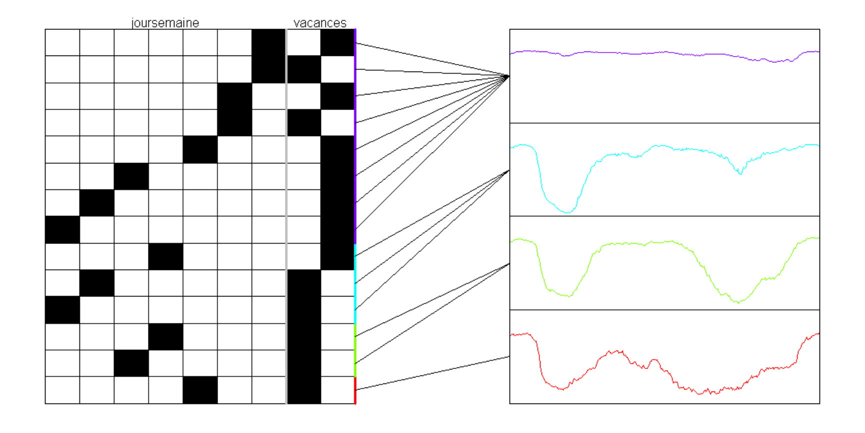
$$\left.\begin{array}{l} E \in \{1, \dots, m\} \text{ i.i.d. hidden R.V.} \\ \epsilon_i \in \mathbb{R}^K,\ \epsilon_i \sim \mathcal{N}(0, \Sigma_i \in \mathcal{M}_{K,K}) \end{array}\right\} \mathbb{E}[V|E=i] = f_i\ ,\ \text{Var}[V|E=i] = \Sigma_i$$

- **Classification** of E then prediction of V by $f^\star$:



The classification model

$X = \{\text{Day of the week, Hollidays}\}$ and $m = 4$



Model the information contained in the speed of the day

Frame :

$\rightarrow$ Prediction **in the day** D
$\rightarrow$ Spedds V^p are **known**

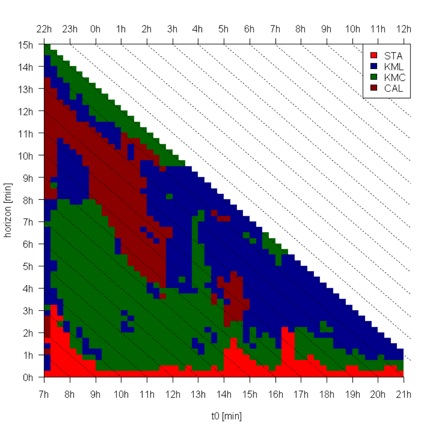
$\Bigg\}$ p fixed, $X = (V^p, C)$

- X Time series
- How many patterns ?
  - h big et p small :
    $\Rightarrow m$ small
  - $\rightarrow$ h small and p big :
    $\Rightarrow m$ **big**

Overcoming non stationarity

Restriction of the forecast profile :

- STA : $g(V^p) = V_p$
- KMC10 : $X = V^p$, $m = 10$ et $\tau = 1$h
- KML4 : $m = 4$ et $\tau = \infty$
- CAL4 : $X = C$, $m = 4$



Avantages :

- High stability
- Small processing time

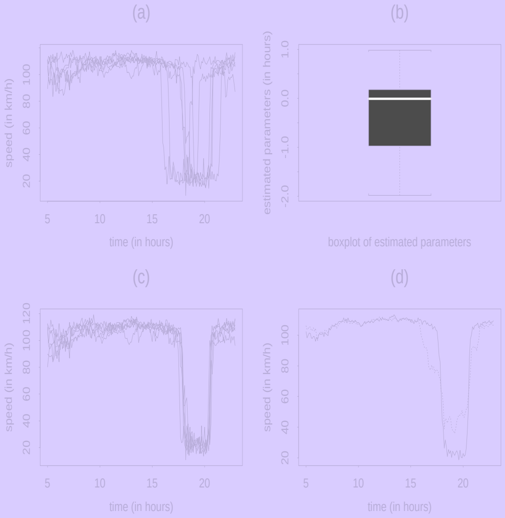
Prediction of the travel time

Example for a travel with $h = 1$ (forecast at one hour)

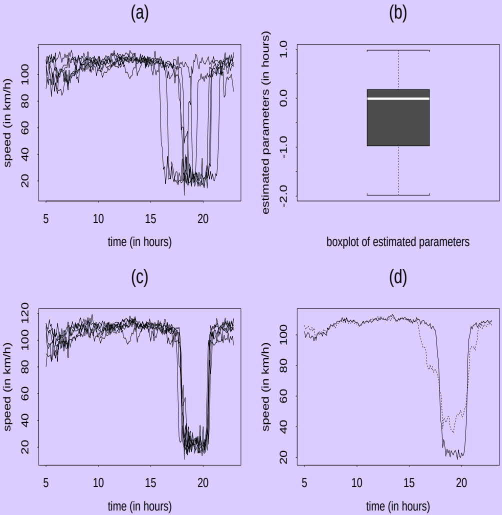
	Mean of de the relative error [%]					
	REF	STA	C10	L4	CAL4	BP
BPI (14km)	32.3	21.2	**14.2**	15.3	17.5	14.4
BPE (21km)	41.8	24.6	17.6	18.9	21.6	**17.1**
A86ES (22km)	20.4	14.7	15.4	13.2	12.6	**10.1**
N118W (26km)	25.4	16.7	9.6	9.8	14.3	**9**
A4W (35km)	21.3	17.5	12.8	13	15	**11.8**

Overview

Shift on traffic jams



Shift on traffic jams



A more general model : Shape invariant model (SIM)

$$Y_{ij} = f_j^*(x_i) + \varepsilon_{ij} \quad i = 1 \ldots n_j,\ j = 1 \ldots J.$$

- ε is as before a Gaussian white noise with variance σ^2
- $\exists f^* : \mathbb{R} \to \mathbb{R}$ with

$$f_j^*(\cdot) = a_j^* f^*(\cdot - \theta_j^*) + \upsilon_j^* \quad (\theta_j^*, a_j^*, \upsilon_j^*) \in \mathbb{R}^3,\ \forall j = 1 \ldots J.$$

A more general model : Shape invariant model (SIM)

$$Y_{ij} = f_j^*(x_i) + \varepsilon_{ij} \quad i = 1 \ldots n_j,\ j = 1 \ldots J.$$

- ε **is as before a Gaussian white noise with variance** σ^2
- $\exists f^* : \mathbb{R} \to \mathbb{R}$ with

$$f_j^*(\cdot) = a_j^* f^*(\cdot - \theta_j^*) + \upsilon_j^* \quad (\theta_j^*, a_j^*, \upsilon_j^*) \in \mathbb{R}^3,\ \forall j = 1 \ldots J.$$

A more general model : Shape invariant model (SIM)

$$Y_{ij} = f_j^*(x_i) + \varepsilon_{ij} \quad i = 1 \ldots n_j,\ j = 1 \ldots J.$$

- ε is as before a Gaussian white noise with variance σ^2
- $\exists f^* : \mathbb{R} \to \mathbb{R}$ **with**

$$f_j^*(\cdot) = a_j^* f^*(\cdot - \theta_j^*) + \upsilon_j^* \quad (\theta_j^*, a_j^*, \upsilon_j^*) \in \mathbb{R}^3,\ \forall j = 1 \ldots J.$$

A more general model : Shape invariant model (SIM)

$$Y_{ij} = f_j^*(x_i) + \varepsilon_{ij} \quad i = 1 \ldots n_j, \ j = 1 \ldots J.$$

- ε is as before a Gaussian white noise with variance σ^2
- $\exists f^* : \mathbb{R} \to \mathbb{R}$ with

$$f_j^*(\cdot) = a_j^* f^*(\cdot - \theta_j^*) + \upsilon_j^* \quad (\theta_j^*, a_j^*, \upsilon_j^*) \in \mathbb{R}^3, \ \forall j = 1 \ldots J.$$

Some references on SIM

- Lawton, W.M., Sylvestre, E.A. and Maggio, M.G. (1972) Introduce SIM and an estimation method (SEMOR).
- Kneip, A. and Gasser, T.(1988) consistency of the SEMOR method.
- Hardle, W. and Marron, J.S.(1990) build an asymptotic normal estimator using a kernel estimator for f^*.
- Gamboa, F., Loubes, J.M. and Maza, E. (2004) build an easy computable asymptotically normal estimator for translations based on DFT.
- Vimond, M. (2005-2007) Efficient estimation in SIM and more general models using DFT and profile likelihood. Ph D Thesis
- Bigot, J. , Loubes, J.M. and Vimond, M.. (2010) Rigid deformations on compact Lie Groups.

Some references on SIM

- **Lawton, W.M., Sylvestre, E.A. and Maggio, M.G. (1972) Introduce SIM and an estimation method (SEMOR).**
- Kneip, A. and Gasser, T.(1988) consistency of the SEMOR method.
- Hardle, W. and Marron, J.S.(1990) build an asymptotic normal estimator using a kernel estimator for f^*.
- Gamboa, F., Loubes, J.M. and Maza, E. (2004) build an easy computable asymptotically normal estimator for translations based on DFT.
- Vimond, M. (2005-2007) Efficient estimation in SIM and more general models using DFT and profile likelihood. Ph D Thesis
- Bigot, J. , Loubes, J.M. and Vimond, M.. (2010) Rigid deformations on compact Lie Groups.

Some references on SIM

- Lawton, W.M., Sylvestre, E.A. and Maggio, M.G. (1972) Introduce SIM and an estimation method (SEMOR).
- **Kneip, A. and Gasser, T.(1988) consistency of the SEMOR method.**
- Hardle, W. and Marron, J.S.(1990) build an asymptotic normal estimator using a kernel estimator for f^*.
- Gamboa, F., Loubes, J.M. and Maza, E. (2004) build an easy computable asymptotically normal estimator for translations based on DFT.
- Vimond, M. (2005-2007) Efficient estimation in SIM and more general models using DFT and profile likelihood. Ph D Thesis
- Bigot, J. , Loubes, J.M. and Vimond, M.. (2010) Rigid deformations on compact Lie Groups.

Some references on SIM

- Lawton, W.M., Sylvestre, E.A. and Maggio, M.G. (1972) Introduce SIM and an estimation method (SEMOR).
- Kneip, A. and Gasser, T.(1988) consistency of the SEMOR method.
- **Hardle, W. and Marron, J.S.(1990) build an asymptotic normal estimator using a kernel estimator for f^*.**
- Gamboa, F., Loubes, J.M. and Maza, E. (2004) build an easy computable asymptotically normal estimator for translations based on DFT.
- Vimond, M. (2005-2007) Efficient estimation in SIM and more general models using DFT and profile likelihood. Ph D Thesis
- Bigot, J. , Loubes, J.M. and Vimond, M.. (2010) Rigid deformations on compact Lie Groups.

Some references on SIM

- Lawton, W.M., Sylvestre, E.A. and Maggio, M.G. (1972) Introduce SIM and an estimation method (SEMOR).
- Kneip, A. and Gasser, T.(1988) consistency of the SEMOR method.
- Hardle, W. and Marron, J.S.(1990) build an asymptotic normal estimator using a kernel estimator for f^*.
- **Gamboa, F., Loubes, J.M. and Maza, E. (2004) build an easy computable asymptotically normal estimator for translations based on DFT.**
- Vimond, M. (2005-2007) Efficient estimation in SIM and more general models using DFT and profile likelihood. Ph D Thesis
- Bigot, J. , Loubes, J.M. and Vimond, M.. (2010) Rigid deformations on compact Lie Groups.

Some references on SIM

- Lawton, W.M., Sylvestre, E.A. and Maggio, M.G. (1972) Introduce SIM and an estimation method (SEMOR).
- Kneip, A. and Gasser, T.(1988) consistency of the SEMOR method.
- Hardle, W. and Marron, J.S.(1990) build an asymptotic normal estimator using a kernel estimator for f^*.
- Gamboa, F., Loubes, J.M. and Maza, E. (2004) build an easy computable asymptotically normal estimator for translations based on DFT.
- **Vimond, M. (2005-2007) Efficient estimation in SIM and more general models using DFT and profile likelihood. Ph D Thesis**
- Bigot, J. , Loubes, J.M. and Vimond, M.. (2010) Rigid deformations on compact Lie Groups.

Some references on SIM

- Lawton, W.M., Sylvestre, E.A. and Maggio, M.G. (1972) Introduce SIM and an estimation method (SEMOR).
- Kneip, A. and Gasser, T.(1988) consistency of the SEMOR method.
- Hardle, W. and Marron, J.S.(1990) build an asymptotic normal estimator using a kernel estimator for f^*.
- Gamboa, F., Loubes, J.M. and Maza, E. (2004) build an easy computable asymptotically normal estimator for translations based on DFT.
- Vimond, M. (2005-2007) Efficient estimation in SIM and more general models using DFT and profile likelihood. Ph D Thesis
- **Bigot, J. , Loubes, J.M. and Vimond, M.. (2010) Rigid deformations on compact Lie Groups.**

Model

Recall the model

$$Y_{i,j} = f^*(x_i - \theta_j^*) + \varepsilon_{ij}, i = 1, \dots, N, j = 1, \dots, J. \quad (1)$$

- f^* is an unknown T-periodic function
- $(\theta_j^*)_{j=1\dots J}$ is an unknown parameter of $\mathbb{R}^J$
- The design is uniform : $x_i = 2i\pi/T, i = 1, \dots N$
- (ε_{ij}) is a Gaussian white noise with variance σ^2

The model is not well posed. Identifiability problem

Model

Recall the model

$$Y_{i,j} = f^*(x_i - \theta_j^*) + \varepsilon_{ij}, i = 1, \dots, N, j = 1, \dots, J. \quad (1)$$

- **f^* is an unknown T-periodic function**
- $(\theta_j^*)_{j=1\dots J}$ is an unknown parameter of $\mathbb{R}^J$
- The design is uniform : $x_i = 2i\pi/T, i = 1, \dots N$
- (ε_{ij}) is a Gaussian white noise with variance σ^2

The model is not well posed. Identifiability problem

Model

Recall the model

$$Y_{i,j} = f^*(x_i - \theta_j^*) + \varepsilon_{ij}, i = 1, \ldots, N, j = 1, \ldots, J. \quad (1)$$

- f^* is an unknown T-periodic function
- $(\theta_j^*)_{j=1\ldots J}$ **is an unknown parameter of** $\mathbb{R}^J$
- The design is uniform : $x_i = 2i\pi/T, i = 1, \ldots N$
- (ε_{ij}) is a Gaussian white noise with variance σ^2

The model is not well posed. Identifiability problem

Model

Recall the model

$$Y_{i,j} = f^*(x_i - \theta_j^*) + \varepsilon_{ij}, i = 1, \ldots, N, j = 1, \ldots, J. \quad (1)$$

- f^* is an unknown T-periodic function
- $(\theta_j^*)_{j=1\ldots J}$ is an unknown parameter of $\mathbb{R}^J$
- **The design is uniform :** $x_i = 2i\pi/T$**,** $i = 1, \ldots N$
- (ε_{ij}) is a Gaussian white noise with variance σ^2

The model is not well posed. Identifiability problem

Model

Recall the model

$$Y_{i,j} = f^*(x_i - \theta_j^*) + \varepsilon_{ij}, i = 1, \ldots, N, j = 1, \ldots, J. \quad (1)$$

- f^* is an unknown T-periodic function
- $(\theta_j^*)_{j=1\ldots J}$ is an unknown parameter of $\mathbb{R}^J$
- The design is uniform : $x_i = 2i\pi/T$, $i = 1, \ldots N$
- (ε_{ij}) **is a Gaussian white noise with variance** σ^2

The model is not well posed. Identifiability problem

Model

Recall the model

$$Y_{i,j} = f^*(x_i - \theta_j^*) + \varepsilon_{ij}, i = 1, \dots, N, j = 1, \dots, J. \qquad (1)$$

- f^* is an unknown T-periodic function
- $(\theta_j^*)_{j=1\dots J}$ is an unknown parameter of $\mathbb{R}^J$
- The design is uniform : $x_i = 2i\pi/T, i = 1, \dots N$
- (ε_{ij}) is a Gaussian white noise with variance σ^2

The model is not well posed. Identifiability problem

Identifiability

Set $\alpha_j^* = \frac{2\pi}{T}\theta_j^*$.

Replacing

- α^* by
$$\alpha^* + c\mathbf{1} + 2k\pi \quad (c \in \mathbb{R}, k \in \mathbb{Z}^J) \tag{2}$$
- f^* by $f^*(\cdot - c)$

the observation equation remains invariant

Identifiability constraints

- Parameter set A is compact
- $\alpha^* \in A$
- If $\alpha \in A$ and $\alpha =$(2) holds then $\alpha = \alpha^*$

Examples

$A_1 = \{\alpha \in [-\pi, \pi[^J : \alpha_1 = 0\}$

$A_2 = \{\alpha \in [-\pi, \pi[^J : \sum \alpha_j = 0$ and $\alpha_1 \in [0, 2\pi/J]\}$

Identifiability

Set $\alpha_j^* = \frac{2\pi}{T}\theta_j^*$.

Replacing

- α^* by
$$\alpha^* + c\mathbf{1} + 2k\pi \quad (c \in \mathbb{R}, k \in \mathbb{Z}^J) \tag{2}$$
- f^* by $f^*(\cdot - c)$

the observation equation remains invariant

Identifiability constraints

- Parameter set A is compact
- $\alpha^* \in A$
- If $\alpha \in A$ and $\alpha =$(2) holds then $\alpha = \alpha^*$

Examples

$A_1 = \{\alpha \in [-\pi, \pi[^J \colon \alpha_1 = 0\}$

$A_2 = \{\alpha \in [-\pi, \pi[^J \colon \sum \alpha_j = 0 \text{ and } \alpha_1 \in [0, 2\pi/J]\}$

Identifiability

Set $\alpha_j^* = \frac{2\pi}{T}\theta_j^*$.

Replacing

- α^* by
$$\alpha^* + c\mathbf{1} + 2k\pi \quad (c \in \mathbb{R}, k \in \mathbb{Z}^J) \tag{2}$$
- f^* by $f^*(\cdot - c)$

the observation equation remains invariant

Identifiability constraints

- Parameter set A is compact
- $\alpha^* \in A$
- If $\alpha \in A$ and $\alpha =$(2) holds then $\alpha = \alpha^*$

Examples

$A_1 = \{\alpha \in [-\pi, \pi[^J \colon \alpha_1 = 0\}$

$A_2 = \{\alpha \in [-\pi, \pi[^J \colon \sum \alpha_j = 0 \text{ and } \alpha_1 \in [0, 2\pi/J]\}$

Identifiability

Set $\alpha_j^* = \frac{2\pi}{T}\theta_j^*$.

Replacing

- α^* by
$$\alpha^* + c\mathbf{1} + 2k\pi \quad (c \in \mathbb{R}, k \in \mathbb{Z}^J) \tag{2}$$
- f^* by $f^*(\cdot - c)$

the observation equation remains invariant

Identifiability constraints

- Parameter set A is compact
- $\alpha^* \in A$
- If $\alpha \in A$ and $\alpha =$(2) holds then $\alpha = \alpha^*$

Examples

$A_1 = \{\alpha \in [-\pi, \pi[^J : \alpha_1 = 0\}$

$A_2 = \{\alpha \in [-\pi, \pi[^J : \sum \alpha_j = 0 \text{ and } \alpha_1 \in [0, 2\pi/J]\}$

Estimation procedure

Main simple idea

For any $c \in \mathbb{R}$ the shift operator T_c defined on T-periodic functions

$$T_c(f) = f(\cdot - c)$$

has common **eigenvectors**

$$T_c[\exp(2i\pi/T\cdot)] = \exp(-2i\pi c/T)\exp(2i\pi/T\cdot)$$

More generally on a general group (here the torus), Fourier transform diagonalizes any translation operators acting on functions on the group (forward to extensions)

Rewrite the regression model using the eigenvectors

Estimation procedure

Main simple idea

For any $c \in \mathbb{R}$ the shift operator T_c defined on T-periodic functions

$$T_c(f) = f(\cdot - c)$$

has common **eigenvectors**

$$T_c[\exp(2i\pi/T\cdot)] = \exp(-2i\pi c/T)\exp(2i\pi/T\cdot)$$

More generally on a general group (here the torus), Fourier transform diagonalizes any translation operators acting on functions on the group (forward to extensions)

Rewrite the regression model using the eigenvectors

Estimation procedure

Main simple idea

For any $c \in \mathbb{R}$ the shift operator T_c defined on T-periodic functions

$$T_c(f) = f(\cdot - c)$$

has common **eigenvectors**

$$T_c[\exp(2i\pi/T\cdot)] = \exp(-2i\pi c/T)\exp(2i\pi/T\cdot)$$

More generally on a general group (here the torus), Fourier transform diagonalizes any translation operators acting on functions on the group (forward to extensions)

Rewrite the regression model using the eigenvectors

Estimation procedure

Main simple idea

For any $c \in \mathbb{R}$ the shift operator T_c defined on T-periodic functions

$$T_c(f) = f(\cdot - c)$$

has common **eigenvectors**

$$T_c[\exp(2i\pi/T\cdot)] = \exp(-2i\pi c/T)\exp(2i\pi/T\cdot)$$

More generally on a general group (here the torus), Fourier transform diagonalizes any translation operators acting on functions on the group (forward to extensions)

Rewrite the regression model using the eigenvectors

Rewriting the model in terms of the Fourier transform

Taking the DFT and neglecting the (deterministic) error between the DFT and the Fourier transform. The model may be rewritten as (N is odd)

$$d_{jl} = e^{-il\alpha_j^*} c_l(f^*) + w_{jl}, l = -(N-1)/2, \dots, (N-1)/2, j = 1, \dots, J$$

- $c_l(f^*)$ is the Fourier coefficient of f^*
- (w_{jl}) is a complex Gaussian white noise with variance σ^2/N

Rewriting the model in terms of the Fourier transform

Taking the DFT and neglecting the (deterministic) error between the DFT and the Fourier transform. The model may be rewritten as (N is odd)

$$d_{jl} = e^{-il\alpha_j^*} c_l(f^*) + w_{jl}, l = -(N-1)/2, \dots, (N-1)/2, j = 1, \dots, J$$

- $c_l(f^*)$ is the Fourier coefficient of f^*
- (w_{jl}) is a complex Gaussian white noise with variance σ^2/N

Building a M-function

- Re phased Fourier coefficients

$$\tilde{c}_{jl}(\alpha) = e^{il\alpha_j} d_{jl} \quad (\alpha \in A)$$

- Mean of Re phased Fourier coefficients

$$\hat{c}_l(\alpha) = \frac{1}{J} \sum_{j=1}^{J} \tilde{c}_{jl}(\alpha)$$

$\tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*} w_{jl}$ and $\hat{c}_l(\alpha^*) = c_l(f^*) + 1/J \sum_{j=1}^{J} e^{il\alpha_j^*} w_{jl}$

Building a $\mathcal{M}$-function

- **Re phased Fourier coefficients**

$$\tilde{c}_{jl}(\alpha) = e^{il\alpha_j} d_{jl} \quad (\alpha \in A)$$

- Mean of Re phased Fourier coefficients

$$\hat{c}_l(\alpha) = \frac{1}{J} \sum_{j=1}^{J} \tilde{c}_{jl}(\alpha)$$

$\tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*} w_{jl}$ and $\hat{c}_l(\alpha^*) = c_l(f^*) + 1/J \sum_{j=1}^{J} e^{il\alpha_j^*} w_{jl}$

Building a $\mathcal{M}$-function

- Re phased Fourier coefficients

$$\tilde{c}_{jl}(\alpha) = e^{il\alpha_j} d_{jl} \quad (\alpha \in A)$$

- **Mean of Re phased Fourier coefficients**

$$\hat{c}_l(\alpha) = \frac{1}{J} \sum_{j=1}^{J} \tilde{c}_{jl}(\alpha)$$

$\tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*} w_{jl}$ and $\hat{c}_l(\alpha^*) = c_l(f^*) + 1/J \sum_{j=1}^{J} e^{il\alpha_j^*} w_{jl}$

Building a $\mathcal{M}$-function

- Re phased Fourier coefficients

$$\tilde{c}_{jl}(\alpha) = e^{il\alpha_j} d_{jl} \quad (\alpha \in A)$$

- Mean of Re phased Fourier coefficients

$$\hat{c}_l(\alpha) = \frac{1}{J} \sum_{j=1}^{J} \tilde{c}_{jl}(\alpha)$$

$\tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*} w_{jl}$ and $\hat{c}_l(\alpha^*) = c_l(f^*) + 1/J \sum_{j=1}^{J} e^{il\alpha_j^*} w_{jl}$

The M-function

Idea : The deviation $\tilde{c}_{jl}(\alpha) - \hat{c}_l(\alpha)$ should be small for $\alpha = \alpha^*$

$$M_n(\alpha) := \frac{1}{J} \sum_{j=1}^{J} \sum_{-(N-1)/2}^{(N-1)/2} \delta_l^2 |\tilde{c}_{jl}(\alpha) - \hat{c}_l(\alpha)|^2$$

- (δ_l) is l^2 sequence of weights discussed later

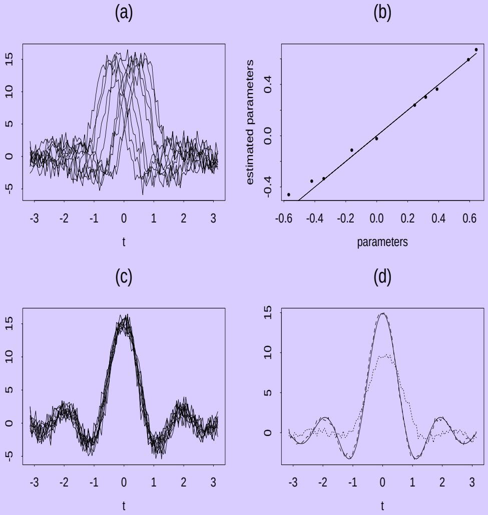
The M-function

Idea : The deviation $\tilde{c}_{jl}(\alpha) - \hat{c}_l(\alpha)$ should be small for $\alpha = \alpha^*$

$$M_n(\alpha) := \frac{1}{J} \sum_{j=1}^{J} \sum_{-(N-1)/2}^{(N-1)/2} \delta_l^2 |\tilde{c}_{jl}(\alpha) - \hat{c}_l(\alpha)|^2$$

- (δ_l) is l^2 sequence of weights discussed later

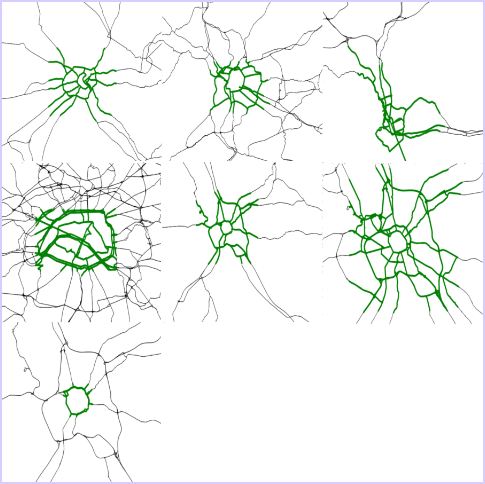
An artificial data example



Overview

Graph of roads network

Modeling : Random process $(X_i^{(n)})_{n\in\mathbb{Z},i\in G}$



Graph of roads network

Modeling : Random process $(X_i^{(n)})_{n \in \mathbb{Z}, i \in G}$

- Indexed by (discrete) time $\mathbb{Z}$ and the **graph** G of the road traffic network
- Gaussian
- Centered
- "Stationary"
- Extension of classical tools from time series to **graphs**

Graph of roads network

Modeling : Random process $(X_i^{(n)})_{n \in \mathbb{Z}, i \in G}$

- Indexed by (discrete) time $\mathbb{Z}$ and the **graph** G of the road traffic network
- Gaussian
- Centered
- "Stationary"
- Extension of classical tools from time series to **graphs**

Graph of roads network

Modeling : Random process $(X_i^{(n)})_{n \in \mathbb{Z}, i \in G}$

- Indexed by (discrete) time $\mathbb{Z}$ and the **graph** G of the road traffic network
- Gaussian
- Centered
- "Stationary"
- Extension of classical tools from time series to **graphs**

Graph of roads network

Modeling : Random process $(X_i^{(n)})_{n\in\mathbb{Z},i\in G}$

- Indexed by (discrete) time $\mathbb{Z}$ and the **graph** G of the road traffic network
- Gaussian
- Centered
- “Stationary“
- Extension of classical tools from time series to **graphs**

Graph of roads network

Modeling : Random process $(X_i^{(n)})_{n\in\mathbb{Z},i\in G}$

- Indexed by (discrete) time $\mathbb{Z}$ and the **graph** G of the road traffic network
- Gaussian
- Centered
- “Stationary“
- Extension of classical tools from time series to **graphs**

Graph of roads network

Modeling : Random process $(X_i^{(n)})_{n \in \mathbb{Z}, i \in G}$

- Indexed by (discrete) time $\mathbb{Z}$ and the **graph** G of the road traffic network
- Gaussian
- Centered
- “Stationary“
- Extension of classical tools from time series to **graphs**

Objective : Yield a parametric model $(\mathcal{K}_\theta)_{\theta \in \Theta}$ for covariance operators of X

Gaussian Process on Graph : Origin of the Problem

Trafic : Predict the speed of the vehicles with missing values

For now : Spatial dependency is not exploited

Aims

- Give a model that uses spatial dependency
- Estimate the spatial correlation
- Spatial filtering

Gaussian Process on Graph : Origin of the Problem

Trafic : Predict the speed of the vehicles with missing values

For now : Spatial dependency is not exploited

Aims

- Give a model that uses spatial dependency
- Estimate the spatial correlation
- Spatial filtering

Gaussian Process on Graph : Origin of the Problem

Trafic : Predict the speed of the vehicles with missing values

For now : Spatial dependency is not exploited

Aims

- Give a model that uses spatial dependency
- Estimate the spatial correlation
- Spatial filtering

Graph

Model : Speed process $(X_i)_{i \in G}$ indexed by the vertices G of a graph $\mathbf{G}$.

Definition (Unoriented weigthed graph)

$\mathbf{G} = (G, W)$ *:*

- G *set of vertices (infinite countable)*
- $W \in [-1, 1]^{G \times G}$ *Weigthed adjacency operator (symmetric)*

Neighbors : $i \sim j$ if $W_{ij} \neq 0$
Degree of a vertex : $D_i = \sharp\{j, i \sim j\}$.
H_0

- $D := \sup_{i \in G} D_i < +\infty$, $\mathbf{G}$ has bounded degree
- $\forall i \in G, \sum_{j \in G} |W_{ij}| \leqslant 1$ even renormalize

Graph

Model : Speed process $(X_i)_{i \in G}$ indexed by the vertices G of a graph $\mathbf{G}$.

Definition (Unoriented weigthed graph)

$\mathbf{G} = (G, W)$:

- G *set of vertices (infinite countable)*
- $W \in [-1, 1]^{G \times G}$ *Weigthed adjacency operator (symmetric)*

Neighbors : $i \sim j$ if $W_{ij} \neq 0$

Degree of a vertex : $D_i = \sharp\{j, i \sim j\}$.

H_0

- $D := \sup_{i \in G} D_i < +\infty$, $\mathbf{G}$ has bounded degree
- $\forall i \in G, \sum_{j \in G} \left|W_{ij}\right| \leqslant 1$ even renormalize

Graph

Model : Speed process $(X_i)_{i \in G}$ indexed by the vertices G of a graph $\mathbf{G}$.

Definition (Unoriented weigthed graph)

$\mathbf{G} = (G, W)$ *:*

- G *set of vertices (infinite countable)*
- $W \in [-1, 1]^{G \times G}$ *Weigthed adjacency operator (symmetric)*

Neighbors : $i \sim j$ if $W_{ij} \neq 0$
Degree of a vertex : $D_i = \sharp\{j, i \sim j\}$.

H_0

- $D := \sup_{i \in G} D_i < +\infty$, $\mathbf{G}$ has bounded degree
- $\forall i \in G, \sum_{j \in G} |W_{ij}| \leqslant 1$ even renormalize

Graph

Model : Speed process $(X_i)_{i \in G}$ indexed by the vertices G of a graph $\mathbf{G}$.

Definition (Unoriented weigthed graph)

$\mathbf{G} = (G, W)$ *:*

- G *set of vertices (infinite countable)*
- $W \in [-1, 1]^{G \times G}$ *Weigthed adjacency operator (symmetric)*

Neighbors : $i \sim j$ if $W_{ij} \neq 0$
Degree of a vertex : $D_i = \sharp\{j, i \sim j\}$.
H_0

- $D := \sup_{i \in G} D_i < +\infty$, $\mathbf{G}$ has bounded degree
- $\forall i \in G, \sum_{j \in G} \left|W_{ij}\right| \leqslant 1$ even renormalize

Problem

Remark :

- Our work robust to renormalization
- For $\mathbb{Z}$, for instance : $W_{ij}^{(\mathbb{Z})} = \frac{1}{2}1_{|i-j|=1}$

W acts on $l^2(G)$:

$$\forall u \in l^2(G), \forall i \in G, (Wu)_i := \sum_{j \in G} W_{ij} u_j$$

Under H_0

W bounded as operator of $B_G := l^2(G) \to l^2(G)$:

$$\|W\|_{2,op} \leqslant 1$$

H_0' : The entries of W belongs to a finite set

Problem

Remark :

- Our work robust to renormalization
- For $\mathbb{Z}$, for instance : $W_{ij}^{(\mathbb{Z})} = \frac{1}{2}1_{|i-j|=1}$

W acts on $l^2(G)$:

$$\forall u \in l^2(G), \forall i \in G, (Wu)_i := \sum_{j \in G} W_{ij} u_j$$

Under H_0

W bounded as operator of $B_G := l^2(G) \to l^2(G)$:

$$\|W\|_{2,op} \leqslant 1$$

H_0' : The entries of W belongs to a finite set

Problem

Remark :

- Our work robust to renormalization
- For $\mathbb{Z}$, for instance : $W_{ij}^{(\mathbb{Z})} = \frac{1}{2}\mathbb{1}_{|i-j|=1}$

W acts on $l^2(G)$:

$$\forall u \in l^2(G), \forall i \in G, (Wu)_i := \sum_{j \in G} W_{ij} u_j$$

Under H_0

W bounded as operator of $B_G := l^2(G) \to l^2(G)$:

$$\|W\|_{2,op} \leqslant 1$$

H_0' : The entries of W belongs to a finite set

Problem

Remark :

- Our work robust to renormalization
- For $\mathbb{Z}$, for instance : $W_{ij}^{(\mathbb{Z})} = \frac{1}{2}1_{|i-j|=1}$

W acts on $l^2(G)$:

$$\forall u \in l^2(G), \forall i \in G, (Wu)_i := \sum_{j \in G} W_{ij} u_j$$

Under H_0

W bounded as operator of $B_G := l^2(G) \to l^2(G)$:

$$\|W\|_{2,op} \leqslant 1$$

H_0' : The entries of W belongs to a finite set

General approach

Observation : Correlations are independent of the position and the orientation

Aim : Propose a *stationary* and *isotropic* model for covariances
$(X_i)_{i\in G}$ Gaussian, zero-mean, with covariance $K \in \mathbb{R}^{G\times G}$:
$\Rightarrow$ Characterized by K

Aim : Extension of time series
$\Rightarrow$ Construction MA with adjacency operator
$+$ *isotropic* modification of the graph

General approach

Observation : Correlations are independent of the position and the orientation

Aim : Propose a *stationary* and *isotropic* model for covariances

$(X_i)_{i \in G}$ Gaussian, zero-mean, with covariance $K \in \mathbb{R}^{G \times G}$:

$\Rightarrow$ Characterized by K

Aim : Extension of time series

$\Rightarrow$ Construction MA with adjacency operator

$+$ *isotropic* modification of the graph

General approach

Observation : Correlations are independent of the position and the orientation
Aim : Propose a *stationary* and *isotropic* model for covariances
$(X_i)_{i \in G}$ Gaussian, zero-mean, with covariance $K \in \mathbb{R}^{G \times G}$:
$\Rightarrow$ Characterized by K

Aim : Extension of time series
$\Rightarrow$ Construction MA with adjacency operator
$+$ *isotropic* modification of the graph

General approach

Observation : Correlations are independent of the position and the orientation

Aim : Propose a *stationary* and *isotropic* model for covariances

$(X_i)_{i \in G}$ Gaussian, zero-mean, with covariance $K \in \mathbb{R}^{G \times G}$:

$\Rightarrow$ Characterized by K

Aim : Extension of time series

$\Rightarrow$ Construction MA with adjacency operator

$+$ *isotropic* modification of the graph

For $\mathbb{Z}$: $(\epsilon_n)_{n \in \mathbb{Z}}$ white noise

$$X_n = \sum_{k \in \mathbb{N}} a_k \epsilon_{n-k}$$

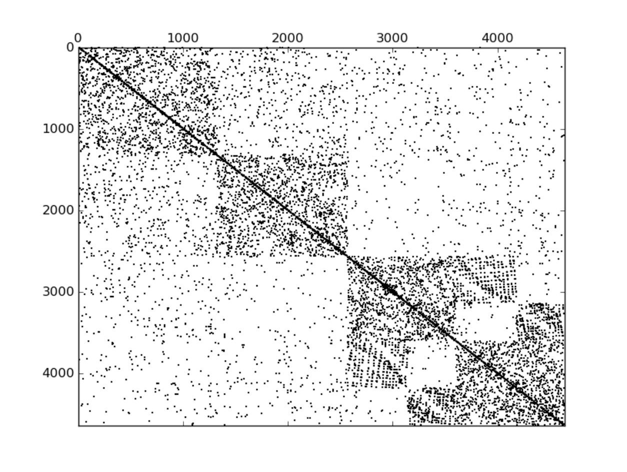
General approach

Observation : Correlations are independent of the position and the orientation

Aim : Propose a *stationary* and *isotropic* model for covariances
$(X_i)_{i \in G}$ Gaussian, zero-mean, with covariance $K \in \mathbb{R}^{G \times G}$:
$\Rightarrow$ Characterized by K

Aim : Extension of time series
$\Rightarrow$ Construction MA with adjacency operator
$+$ *isotropic* modification of the graph

Aim : Maximum Likelihood Estimation
$\Rightarrow$ Generalize Whittle's approximation


0
1000
2000
3000
4000
0
1000
2000
3000
4000

A few bibliography

Spectral representation of stationary processes :

- $\mathbb{Z}^{\mathrm{d}}$: X. Guyon
- Homogeneous tree : J-P. Arnaud
- Distance-transitive graphs : H. Heyer

Maximum Likelihood

- $\mathbb{Z}$: here R. Azencott et D. Dacunha-Castelle
- $\mathbb{Z}^{\mathrm{d}}$: X. Guyon, R. Dahlhaus

A few bibliography

Spectral representation of stationary processes :

- $\mathbb{Z}^{\mathrm{d}}$: X. Guyon
- Homogeneous tree : J-P. Arnaud
- Distance-transitive graphs : H. Heyer

Maximum Likelihood

- $\mathbb{Z}$: here R. Azencott et D. Dacunha-Castelle
- $\mathbb{Z}^{\mathrm{d}}$: X. Guyon, R. Dahlhaus

Example $G = \mathbb{Z}$

X Gaussian centered process with covariance K is stationary if

$$\exists (r_k)_{k\in\mathbb{N}}, K_{ij} = r_{|i-j|}$$

Spectral density

If $r \in l^1, \exists f, K_{ij} = \frac{1}{2\pi}\int_{[0,2\pi]} f(t)\cos((j-i)t)\,dt := (T(f))_{ij}$

Let $g, f(t) = g(\cos(t))$, As

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})}\right)^k\right)_{ij} = \frac{1}{2\pi}\int_{[0,2\pi]} \cos(t)^k \cos((j-i)t)\,dt,$$

Operator representation

We have $K = g(W^{(\mathbb{Z})})$

Example $G = \mathbb{Z}$

X Gaussian centered process with covariance K is stationary if

$$\exists (r_k)_{k \in \mathbb{N}}, K_{ij} = r_{|i-j|}$$

Spectral density

If $r \in l^1, \exists f, K_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} f(t) \cos((j-i)t)\, dt := (T(f))_{ij}$

Let $g, f(t) = g(\cos(t))$, As

$$\forall i,j,k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})} \right)^k \right)_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} \cos(t)^k \cos((j-i)t)\, dt,$$

Operator representation

We have $K = g(W^{(\mathbb{Z})})$

Example $G = \mathbb{Z}$

X Gaussian centered process with covariance K is stationary if

$$\exists (r_k)_{k \in \mathbb{N}}, K_{ij} = r_{|i-j|}$$

Spectral density

If $r \in l^1$, $\exists f$, $K_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} f(t) \cos((j-i)t)\, dt := (T(f))_{ij}$

Let g, $f(t) = g(\cos(t))$, As

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})} \right)^k \right)_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} \cos(t)^k \cos((j-i)t)\, dt,$$

Operator representation

We have $K = g(W^{(\mathbb{Z})})$

Example $G = \mathbb{Z}$

X Gaussian centered process with covariance K is stationary if

$$\exists (r_k)_{k\in\mathbb{N}}, K_{ij} = r_{|i-j|}$$

Spectral density

If $r \in l^1, \exists f, K_{ij} = \frac{1}{2\pi}\int_{[0,2\pi]} f(t)\cos((j-i)t)\,dt := (T(f))_{ij}$

Let $g, f(t) = g(cos(t))$, As

$$\forall i,j,k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})}\right)^k\right)_{ij} = \frac{1}{2\pi}\int_{[0,2\pi]} \cos(t)^k \cos((j-i)t)\,dt,$$

Operator representation

We have $K = g(W^{(\mathbb{Z})})$

Example $G = \mathbb{Z}$

We can also write

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})}\right)^k\right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

where T_k is the k-th Tchebychev's polynomial

$$\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

- g polynomial of degree q : MA_q
- $\frac{1}{g}$ polynomial of degree p : $AR_p \cdots$

Aim : Generalize this kind of representation

Example $G = \mathbb{Z}$

We can also write

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})}\right)^k\right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

where T_k is the k-th Tchebychev's polynomial

$$\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

- g polynomial of degree q : MA_q
- $\frac{1}{g}$ polynomial of degree p : AR_p $\cdots$

Aim : Generalize this kind of representation

Example $G = \mathbb{Z}$

We can also write

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})}\right)^k\right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

where T_k is the k-th Tchebychev's polynomial

$$\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

- g polynomial of degree q : MA_q
- $\frac{1}{g}$ polynomial of degree p : $AR_p \cdots$

Aim : Generalize this kind of representation

Example $G = \mathbb{Z}$

We can also write

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})} \right)^k \right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

where T_k is the k-th Tchebychev's polynomial

$$\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

- g polynomial of degree q : MA_q
- $\frac{1}{g}$ polynomial of degree p : $AR_p \cdots$

Aim : Generalize this kind of representation

Example $G = \mathbb{Z}$

We can also write

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})}\right)^k\right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

where T_k is the k-th Tchebychev's polynomial

$$\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1-\lambda^2}} d\lambda$$

- g polynomial of degree q : MA_q
- $\frac{1}{g}$ polynomial of degree p : AR_p $\cdots$

Aim : Generalize this kind of representation

Identity resolution

Spectral decomposition

$$\exists E, \mathcal{M}, W = \int_{\mathcal{M}} \lambda dE(\lambda)$$

Definition (Identity resolution)

$\mathcal{M}$ *Sigma-algebra* $E : \mathcal{M} \to B_G$ *such that* $\forall \omega, \omega' \in \mathcal{M}$,

1) $E(\omega)$ *self-adjoint operator.*
2) $E() = 0, E(\Omega) = I$
3) $E(\omega \cap \omega') = E(\omega)E(\omega')$
4) *Si* $\omega \cap \omega' =$ *, then* $E(\omega \cup \omega') = E(\omega) + E(\omega')$

$\forall i, j \in G, \forall \omega \in \mathcal{M}, \mu_{ij}(\omega) = E_{ij}(\omega)$

Identity resolution

Spectral decomposition

$$\exists \mathrm{E}, \mathcal{M}, W = \int_{\mathcal{M}} \lambda d\mathrm{E}(\lambda)$$

Definition (Identity resolution)

$\mathcal{M}$ *Sigma-algebra* $\mathrm{E} : \mathcal{M} \to \mathrm{B}_G$ *such that* $\forall \omega, \omega' \in \mathcal{M}$,

1) $\mathrm{E}(\omega)$ *self-adjoint operator.*
2) $\mathrm{E}() = 0, \mathrm{E}(\Omega) = \mathrm{I}$
3) $\mathrm{E}(\omega \cap \omega') = \mathrm{E}(\omega)\mathrm{E}(\omega')$
4) *Si* $\omega \cap \omega' =$ *, then* $\mathrm{E}(\omega \cup \omega') = \mathrm{E}(\omega) + \mathrm{E}(\omega')$

$\forall i, j \in G, \forall \omega \in \mathcal{M}, \mu_{ij}(\omega) = \mathrm{E}_{ij}(\omega)$

Identity resolution

Spectral decomposition

$$\exists E, \mathcal{M}, W = \int_{\mathcal{M}} \lambda dE(\lambda)$$

Definition (Identity resolution)

$\mathcal{M}$ *Sigma-algebra* $E : \mathcal{M} \to B_G$ *such that* $\forall \omega, \omega' \in \mathcal{M}$,

1) $E(\omega)$ *self-adjoint operator.*
2) $E() = 0, E(\Omega) = I$
3) $E(\omega \cap \omega') = E(\omega)E(\omega')$
4) *Si* $\omega \cap \omega' =$, *then* $E(\omega \cup \omega') = E(\omega) + E(\omega')$

$\forall i, j \in G, \forall \omega \in \mathcal{M}, \mu_{ij}(\omega) = E_{ij}(\omega)$

Extension to a graph

Definition

$(X_i)_{i \in G}$ *Gaussian field with covariance* K.

$$\textit{If } K = \int_{\mathrm{Sp}(W)} g(\lambda) dE(\lambda),$$

- g *polynomial :* $MA_q^{(W)}$
- $\frac{1}{g}$ *polynomial :* $AR_p^{(W)} \cdots$

Remarks :

- Conditions about g
- Equivalence with $\mathbb{Z}$
- $K = g(W)$, with normal convergence of the series
- Dependency on W

Extension to a graph

Definition

$(X_i)_{i \in G}$ *Gaussian field with covariance* K.

$$\textit{If } K = \int_{Sp(W)} g(\lambda) dE(\lambda),$$

- g *polynomial :* $MA_q^{(W)}$
- $\frac{1}{g}$ *polynomial :* $AR_p^{(W)}$ $\cdots$

Remarks :

- Conditions about g
- Equivalence with $\mathbb{Z}$
- $K = g(W)$, with normal convergence of the series
- Dependency on W

Extension to a graph

Definition

$(X_i)_{i \in G}$ *Gaussian field with covariance* K.

$$\textit{If } K = \int_{\mathrm{Sp}(W)} g(\lambda) dE(\lambda),$$

- g *polynomial :* $MA_q^{(W)}$
- $\frac{1}{g}$ *polynomial :* $AR_p^{(W)}$ $\cdots$

Remarks :

- Conditions about g
- Equivalence with $\mathbb{Z}$
- $K = g(W)$, with normal convergence of the series
- Dependency on W

Extension to a graph

Definition

$(X_i)_{i \in G}$ *Gaussian field with covariance* K.

$$\textit{If } K = \int_{\mathrm{Sp}(W)} g(\lambda) dE(\lambda),$$

- g *polynomial :* $MA_q^{(W)}$
- $\frac{1}{g}$ *polynomial :* $AR_p^{(W)}$ $\cdots$

Remarks :

- Conditions about g
- Equivalence with $\mathbb{Z}$
- $K = g(W)$, with normal convergence of the series
- Dependency on W

Extension to a graph

Definition

$(X_i)_{i \in G}$ *Gaussian field with covariance* K.

$$\textit{If } K = \int_{\mathrm{Sp}(\mathcal{W})} g(\lambda) dE(\lambda),$$

- g *polynomial :* $MA_q^{(\mathcal{W})}$
- $\frac{1}{g}$ *polynomial :* $AR_p^{(\mathcal{W})}$ $\cdots$

Remarks :

- Conditions about g
- Equivalence with $\mathbb{Z}$
- $K = g(W)$, with normal convergence of the series
- Dependency on W

Extension to a graph

Definition

$(X_i)_{i \in G}$ *Gaussian field with covariance* K.

$$\textit{If } K = \int_{Sp(W)} g(\lambda) dE(\lambda),$$

- g *polynomial :* $MA_q^{(W)}$
- $\frac{1}{g}$ *polynomial :* $AR_p^{(W)}$ $\cdots$

Remarks :

- Conditions about g
- Equivalence with $\mathbb{Z}$
- $K = g(W)$, with normal convergence of the series
- Dependency on W

Extension to a graph

Definition

$(X_i)_{i \in G}$ *Gaussian field with covariance* K.

$$\textit{If } K = \int_{Sp(W)} g(\lambda) dE(\lambda),$$

- g *polynomial :* $MA_q^{(W)}$
- $\frac{1}{g}$ *polynomial :* $AR_p^{(W)}$ $\cdots$

Remarks :

- Conditions about g
- Equivalence with $\mathbb{Z}$
- $K = g(W)$, with normal convergence of the series
- Dependency on W

Problem :

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(\mathbf{G}_n)_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z}$: $G_n = [1, n]$.
- $\theta_0 \in \mathring{\Theta}$, $\mathbf{X} \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction X_n of $\mathbf{X}$ to $\mathbf{G}_n$, cov : $K_n(f_\theta)$
- $m_n = \sharp G_n$

Aim : Estimate θ_0 by maximum likelihood :

$$L_n(\theta) := -\frac{1}{2}\left(m_n \log(2\pi) + \log\det\left(K_n(f_\theta)\right) + X_n^T \left(K_n(f_\theta)\right)^{-1} X_n\right)$$

Problem :

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(\mathbf{G}_n)_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z}$: $G_n = [1, n]$.
- $\theta_0 \in \mathring{\Theta}$, $\mathbf{X} \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction X_n of $\mathbf{X}$ to $\mathbf{G}_n$, cov : $K_n(f_\theta)$
- $m_n = \sharp G_n$

Aim : Estimate θ_0 by maximum likelihood :

$$L_n(\theta) := -\frac{1}{2}\left(m_n \log(2\pi) + \log\det\left(K_n(f_\theta)\right) + X_n^T \left(K_n(f_\theta)\right)^{-1} X_n\right)$$

Problem :

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(\mathbf{G_n})_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z}$: $G_n = [1, n]$.
- $\theta_0 \in \mathring{\Theta}$, $\mathbf{X} \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction X_n of $\mathbf{X}$ to $\mathbf{G}_n$, cov : $K_n(f_\theta)$
- $m_n = \sharp G_n$

Aim : Estimate θ_0 by maximum likelihood :

$$L_n(\theta) := -\frac{1}{2}\left(m_n \log(2\pi) + \log\det\left(K_n(f_\theta)\right) + X_n^T \left(K_n(f_\theta)\right)^{-1} X_n\right)$$

Problem :

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(\mathbf{G_n})_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z}$: $G_n = [1, n]$.
- $\theta_0 \in \mathring{\Theta}$, $\mathbf{X} \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction X_n of $\mathbf{X}$ to $\mathbf{G}_n$, cov : $K_n(f_\theta)$
- $m_n = \sharp G_n$

Aim : Estimate θ_0 by maximum likelihood :

$$L_n(\theta) := -\frac{1}{2}\left(m_n \log(2\pi) + \log\det\left(K_n(f_\theta)\right) + X_n^T \left(K_n(f_\theta)\right)^{-1} X_n\right)$$

Problem :

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta\in\Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(\mathbf{G_n})_{n\in\mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z}$: $G_n = [1, n]$.
- $\theta_0 \in \mathring{\Theta}$, $\mathbf{X} \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction X_n of $\mathbf{X}$ to $\mathbf{G_n}$, cov : $K_n(f_\theta)$
- $m_n = \sharp G_n$

Aim : Estimate θ_0 by maximum likelihood :

$$L_n(\theta) := -\frac{1}{2}\left(m_n \log(2\pi) + \log\det\left(K_n(f_\theta)\right) + X_n^T \left(K_n(f_\theta)\right)^{-1} X_n\right)$$

Problem :

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(\mathbf{G_n})_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z}$: $G_n = [1, n]$.
- $\theta_0 \in \mathring{\Theta}$, $\mathbf{X} \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction X_n of $\mathbf{X}$ to $\mathbf{G_n}$, cov : $K_n(f_\theta)$
- $m_n = \sharp G_n$

Aim : Estimate θ_0 by maximum likelihood :

$$L_n(\theta) := -\frac{1}{2}\left(m_n \log(2\pi) + \log\det\left(K_n(f_\theta)\right) + X_n^T \left(K_n(f_\theta)\right)^{-1} X_n\right)$$

Problem :

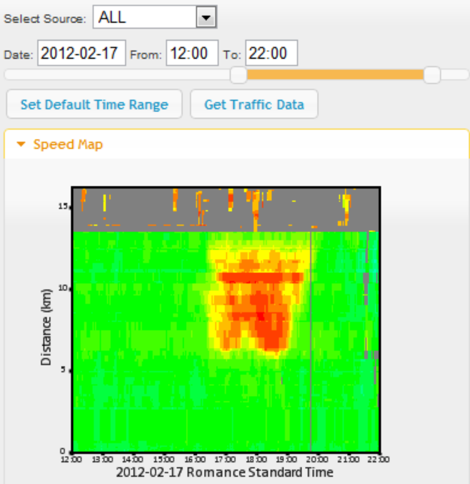
- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta\in\Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(\mathbf{G}_n)_{n\in\mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z}$: $G_n = [1, n]$.
- $\theta_0 \in \mathring{\Theta}$, $\mathbf{X} \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction X_n of $\mathbf{X}$ to $\mathbf{G}_n$, cov : $K_n(f_\theta)$
- $m_n = \sharp G_n$

Aim : Estimate θ_0 by maximum likelihood :

$$L_n(\theta) := -\frac{1}{2}\left(m_n \log(2\pi) + \log\det\left(K_n(f_\theta)\right) + X_n^T \left(K_n(f_\theta)\right)^{-1} X_n\right)$$

The concrete problem

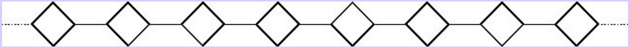
The concrete problem



Overview

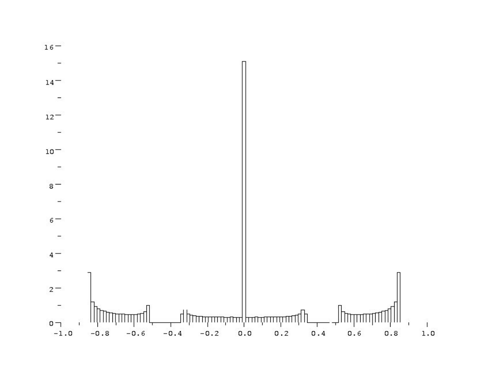
Applications

Fig: Graphe G



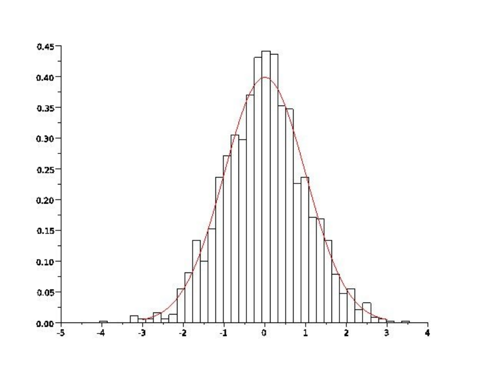
Applications

Fig: Empirical spectral measure

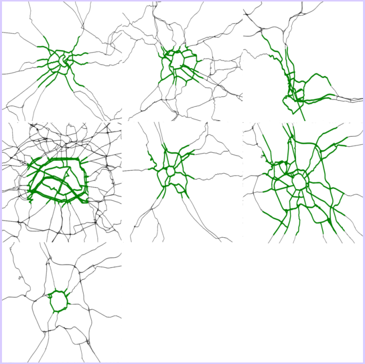


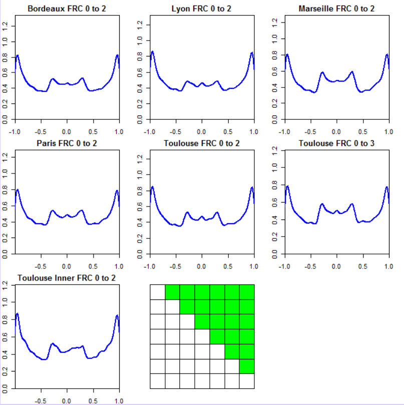
Applications

Fig: Empirical distribution of estimation error



Spectrum of the road network





Real datas

Aim : Predict missing values on FRC 0 in Toulouse

Real datas

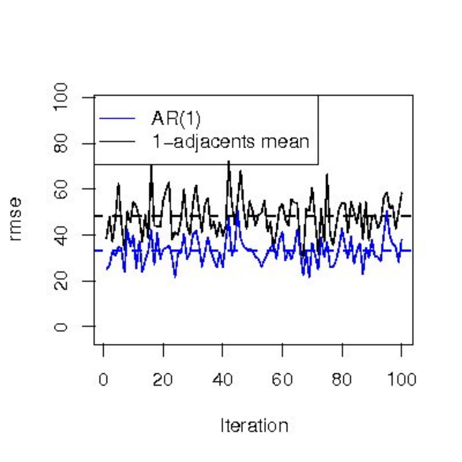
Aim : Predict missing values on FRC 0 in Toulouse

Protocol :

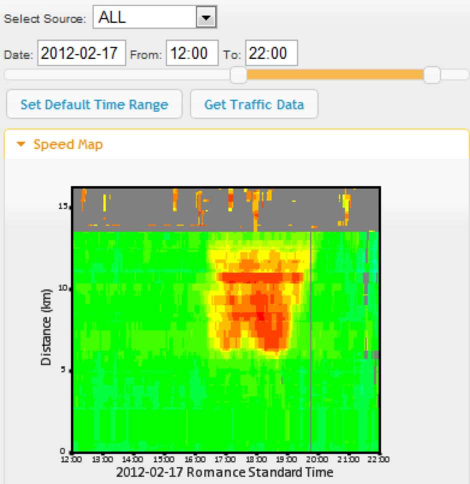
- 10% of datas hidden to test the quality of the prediction
- Model : AR_1

Real datas

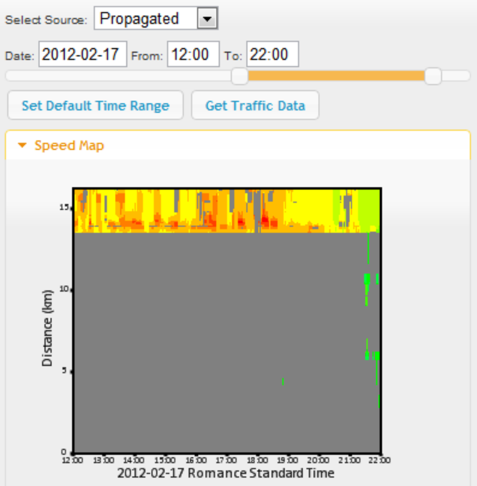
Aim : Predict missing values on FRC 0 in Toulouse



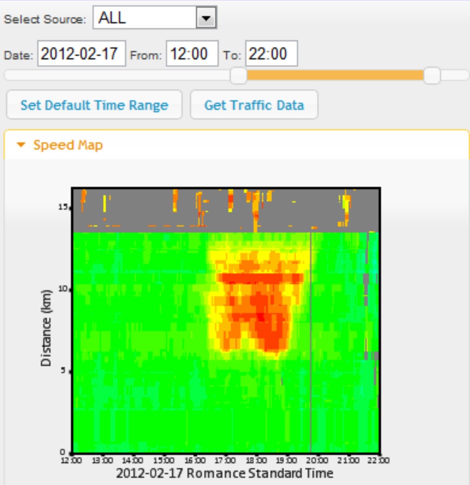
A concrete problem

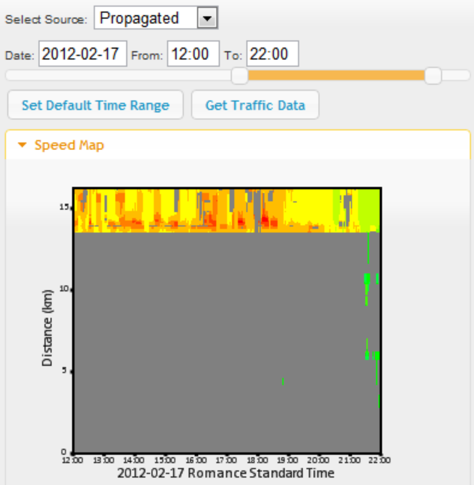


A solution ?


Select Source: Propagated
Date: 2012-02-17 From: 12:00 To: 22:00
Set Default Time Range
Get Traffic Data
Speed Map
Distance (km)
15
10
5
0
12:00 13:00 14:00 15:00 16:00 17:00 18:00 19:00 20:00 21:00 22:00
2012-02-17 Romance Standard Time

Let's compare





An alternative : prediction

Observations of the process on a growing sequence of subgraphs of G, $\text{with missing values}$.
Let (O, M) be a partition of G. The set O will denote the asymptotic for observed values index set, and B the "blind" missing values index set (finite).
Let $(G_N)_{N\in\mathbb{N}}$ be a growing sequence of induced subgraphs of G. From now on, we assume that N is large enough to ensure $B \subset G_N$. The observation index set will be denoted $O_N := O \cap G_N$. We consider the restriction $X_{O_N} := (X_i)_{i\in O_N}$, and assume from now on, that we dispose of a consistent estimation procedure $\hat{f}_N$ for f, such that there exists (r_N) such that

- $\mathbb{E}\left[\left\|\hat{f}_N - f\right\|_\infty^2\right]^{\frac{1}{2}} \leqslant r_N.$
- $\mathbb{E}\left[\left\|\hat{f}_N - f\right\|_\infty^4\right]^{\frac{1}{4}} \leqslant r_N.$

Linear Prediction : Kriging

Recall that the best linear predictor of Z_B (this is also the best predictor in the Gaussian case) can be written

$$\bar{Z}_B = P_{[X_{O_N}]}(f) Z_B := a_B^T{}_{MO_N}(f) \left({}_{O_N}(f) \right)^{-1} X_{O_N}.$$

Then, remark that we asymptotically observe X_O and introduce the best linear prediction of Z_B knowing X_O :

$$\tilde{Z}_B := P_{[X_O]}(f) Z_B := a_B^T{}_{MO}(f) \left({}_{O}(f) \right)^{-1} X_O.$$

The blind problem can be formulated as following :

- Estimation step : Estimate $P_{[X_{O_N}]}(f)$ by $\hat{P}_{[X_{O_N}]}(f) := P_{[X_{O_N}]}(\hat{f})$
- Prediction step : Build $\hat{Z}_B := P_{[X_{O_N}]}(\hat{f}) Z_B$

Extension to graph of Kriging method

Under the assumption that There exists $m, M > 0$ such that

$$\forall t \in \mathrm{Sp}(A), m \leqslant f(t) \leqslant M.$$

Risk :

$$\mathbf{R}_{K,N} = \sup_{Z_B \in [X_B] \mathrm{var}(Z_B)=1} \mathbb{E}\left[\left(Z_B - \hat{Z}_B\right)^2\right]^{\frac{1}{2}}.$$

$$\mathcal{R}_N = \sup_{Z_B \in [X_B] \mathrm{var}(Z_B)=1} \mathbb{E}\left[\left(\tilde{Z}_B - \bar{Z}_B\right)^2\right]^{\frac{1}{2}} + \sup_{Z_B \in [X_B] \mathrm{var}(Z_B)=1} \mathbb{E}\left[\left(\bar{Z}_B - \hat{Z}_B\right)\right.$$

Result :

$$\mathcal{R}_N \leqslant \frac{\sqrt{M}(m+M)}{m^2} r_N + \frac{1}{m^2}\left(\frac{M^{\frac{5}{2}}}{m} + M^{\frac{3}{2}}\right) \sum_{k \geqslant d_G(B,(G \setminus G_N))} \left|\left(\frac{1}{f}\right)_k\right|.$$

Thank you for your Attention