

Learning the Smoothness of Weakly Dependent Functional Times Series

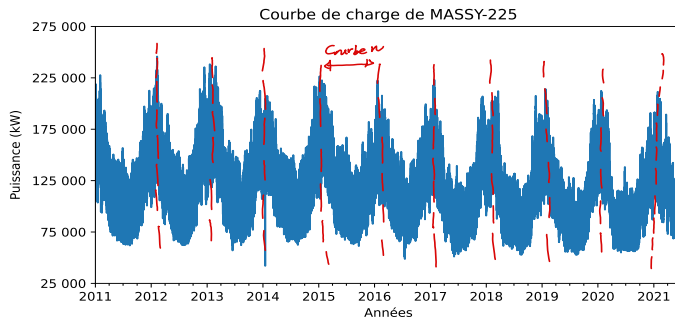
Hassan MAISSORO, Valentin PATILEA and Myriam VIMOND

54^{es} JOURNÉES DE STATISTIQUE DE LA SFdS, 3-7 JULY 2023

Introduction (1/3)

Example of a connection point for the extraction and injection of electricity

- ▶ A set of N time-dependent curves, $X_n : [0, 1] \rightarrow \mathbb{R}$, $n = 1 \dots N$.



- ▶ The trajectories are **irregular**.
- ▶ We observe each curve **every 10 mins** + **measurement errors**.
- ▶ **Regularity** and **final goal** should be considered in reconstruction.

Introduction (2/3)

Observation scheme

For $n = 1, \dots, N$, X_n is measured with error at discrete, randomly sampled points :

$$Y_{n,k} = X_n(T_{n,k}) + \varepsilon_{n,k}, \quad 1 \leq k \leq M_n,$$

- ▶ $\{X_n\}$ is a stationary process of $\mathcal{H} = \mathbb{L}^2[0, 1]$,
- ▶ $M_1, \dots, M_N \stackrel{i.i.d.}{\sim} M$ with expectation λ ,
- ▶ the observation times $T_{n,k} \sim T$ are i.i.d.,
- ▶ $\varepsilon_{n,k} \sim \epsilon$ are independent centered errors,
- ▶ $\{X_n\}$, $\{M_n\}$, $\{\varepsilon_{n,k}\}$, and $\{T_{n,k}\}$ are mutually independent.

Introduction (3/3)

Motivation

We aim to estimate the **local regularity parameters** of the trajectories for **FTS** in the context of **weak dependency**.

Using dependent curves measured with noise at random discrete points, our goal is to perform **adaptive estimation** of :

- ▶ mean and autocovariance kernel functions,
 - ▶ depth functions, etc.
-
- ▶ The concept of **local regularity** was considered by GOLOVKINE ET AL., (2022) for **i.i.d. functional data**.
 - ▶ For FTS, mean and autocovariance estimators have already been considered by RUBÌN AND PANARETOS (2020) under the hypothesis that these functions admit at least one derivative.
 - ▶ We extend the results of GOLOVKINE ET AL., (2022) to FTS to perform estimates that adapt to the local regularity.

Outline

- 1 Introduction
- 2 Local regularity parameters
 - Definition and estimation
 - Weak dependency assumption
 - Concentration bounds
 - Application
- 3 Perspectives
- 4 Take home message

Local regularity parameters (1/5)

Definition and estimation

Definition. The process X admits a *local regularity* at $t \in I$, with *local exponent* $H_t \in (0, 1)$ and *Hölder constant* $L_t > 0$, if

$$\mathbb{E} \left[(X(u) - X(v))^2 \right] \approx L_t^2 |u - v|^{2H_t},$$

for all u, v satisfying $t - \Delta/2 \leq u \leq t \leq v \leq t + \Delta/2$ for some $\Delta > 0$.

Local regularity parameters (1/5)

Definition and estimation

Definition. The process X admits a *local regularity* at $t \in I$, with *local exponent* $H_t \in (0, 1)$ and *Hölder constant* $L_t > 0$, if

$$\mathbb{E} [(X(u) - X(v))^2] \approx L_t^2 |u - v|^{2H_t},$$

for all u, v satisfying $t - \Delta/2 \leq u \leq t \leq v \leq t + \Delta/2$ for some $\Delta > 0$.

Estimation. We use some nonparametric estimates \tilde{X}_n to recover the X_n 's. For any u, v close to t , let

$$\hat{\theta}(u, v) = \frac{1}{N} \sum_{n=1}^N \left\{ \tilde{X}_n(v) - \tilde{X}_n(u) \right\}^2.$$

Our estimators of H_t and L_t^2 are defined as empirical counterparts of their respective definition. Let $t_1 = t - \Delta/2$, $t_3 = t + \Delta/2$. The estimators of H_t and L_t^2 are

$$\hat{H}_t = \frac{\log(\hat{\theta}(t_1, t_3)) - \log(\hat{\theta}(t_1, t))}{2 \log(2)} \quad \text{and} \quad \hat{L}_t^2 = \frac{\hat{\theta}(t_1, t_3)}{\Delta^{2\hat{H}_t}}.$$

Local regularity parameters (2/5)

Weak dependency assumption

Let $\{X_n\}_{n \in \mathbb{Z}}$ be a stationary FTS, with **continuous paths**, on $I = [0, 1]$:

- ▶ $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$: space of square integrable functions ;
- ▶ $(\mathcal{C}, \|\cdot\|_{\infty})$: space of continuous functions on I .

The space $\mathbb{L}_{\mathcal{C}}^p$ is the space of \mathcal{C} -valued random element X such that

$$\nu_p(\|X\|_{\infty}^p) = (\mathbb{E}[\|X\|_{\infty}^p])^{1/p} < \infty.$$

Weak dependency assumption : $\{X_n\}_n$ is $\mathbb{L}_{\mathcal{C}}^p$ – **m-approximable**.

- ▶ \mathbb{L}^p – **m-approximation** for \mathcal{H} -valued functional data was introduced by HÖRMANN and KOKOSZKA (2010).
- ▶ We need a dependency type of $\{X_n\}$ that can be inherited by $\{X_n(t)\}$ because we are studying $\{X_n\}$ locally at $t \in I$ and such we use $\|\cdot\|_{\infty}$ instead of $\|\cdot\|_{\mathcal{H}}$.

Example. $FAR(1)$ is $\mathbb{L}_{\mathcal{C}}^p$ – *m-approximable*.

Local regularity parameters (3/5)

Concentration bounds

- ▶ Let $\{X_n\}$ be \mathbb{L}_C^4 – **m-approximable**.
- ▶ Assume that the \mathbb{L}^2 -risk of smoothing is suitably bounded.

Then, for some $\varphi, \psi \in (0, 1)$ such that

$$6L_t^2 \Delta^{-2\varphi} \varphi |\log \Delta| < \psi,$$

and for λ large enough, we have :

$$\mathbb{P} \left(\left| \widehat{H}_t - H_t \right| > \varphi \right) \leq \frac{f_1}{N \varphi^2 \Delta^{4H_t}} + 4b \exp \left(-f_2 N \varphi^2 \Delta^{4H_t} \right),$$

$$\begin{aligned} \mathbb{P} \left(\left| \widehat{L}_t^2 - L_t^2 \right| > \psi \right) &\leq \frac{g_1}{N \psi^2 \Delta^{4H_t+4\varphi}} + \frac{f_1}{N \varphi^2 \Delta^{4H_t}} \\ &\quad + 4b \exp \left(-f_2 N \varphi^2 \Delta^{4H_t} \right) + 2b \exp \left(-g_2 N \psi^2 \Delta^{4H_t+4\varphi} \right). \end{aligned}$$

where $b > 0$ is a constant and $f_1, f_2, g_1, g_2 > 0$ are also constants depending on the dependence measure.

Local regularity parameters (4/5)

Application : sample paths of a FAR(1)

We simulate a FAR(1) where $\{\xi_n\}$ are i.i.d. 'tied-down' *multifractional Brownian motion* (see STOEV and TAQQU (2006)) paths with :

- ▶ a logistic H_t function and $L_t^2 = 4$,
- ▶ and a kernel $\beta(s, t) = \alpha st$, with $\alpha = 9/4$.

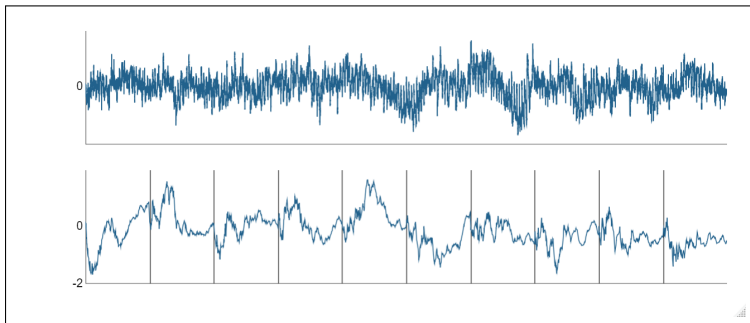


Figure – Time series of $N = 250$ observations of a simulated FAR(1) without error. The last ten functions are shown in the bottom graph.

Local regularity parameters (5/5)

Application : estimation of local regularity parameters

Estimation of H_t and L_t^2 at $t = 1/2$ using the previous FAR(1) and taking $\epsilon \sim \mathcal{N}(0, \sigma^2 = 0.04)$.

► Obtained reasonably good results :

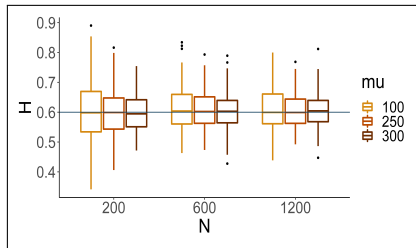


Figure – Estimates of \hat{H}_t . The line is the true $H_t = 0.6$.

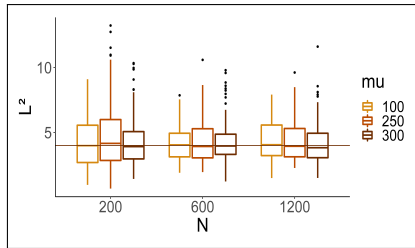


Figure – Estimates of \hat{L}_t^2 . The line is the true $L_t^2 = 4$.

Perspectives

- 1 Introduction
- 2 Local regularity parameters
- 3 Perspectives**
- 4 Take home message

Perspectives

Let

$$\mu(t) = \mathbb{E}(X_n(t)) \text{ and } \gamma_\ell(s, t) = \mathbb{E}(X_n(s)X_{n+\ell}(t)) - \mu(s)\mu(t)$$

be the mean and autocovariance functions of the stationary process $\{X_n\}$ at times t and (s, t) .

Adaptive mean function estimation. A naive estimator of $\mu(t)$ is

$$\hat{\mu}_N(t, h) = N^{-1} \sum_{n=1}^N \hat{X}_n(t, h),$$

where $\hat{X}_n(t, h)$ is a nonparametric estimator of X_n , and h a bandwidth.

- **The objective :** to estimate $\mu(t)$ by selection of h according to the local regularity of $\{X_n\}$ at time t , and by selecting the relevant curves of the sample.

Adaptive autocovariance kernel estimation.

- **The objective :** to estimate $\gamma_\ell(s, t)$ by selection of h according to the local regularity of $\{X_n\}$ at time t , and by selecting the relevant curves of the sample

Take home message

① Estimation of local regularity for FTS.

- Local regularity parameters are : **exponent** H_t and **Hölder constant** L_t^2 .
- Exponential bound for the concentration of the estimators of H_t and L_t^2 under \mathbb{L}_C^4 – **m-approximation**.
- The simulations show that \hat{H}_t and \hat{L}_t^2 give satisfactory results.

② Adaptive estimation of the mean and autocovariance kernel functions.

- Optimal smoothing parameter used to reconstruct curves depends on the final goal.
- Simulations show satisfactory results for the mean function.

► Perspectives :

- adaptive estimators for anomaly detection,
- Robust prediction model, etc.

Thanks for your attention !