



HAL
open science

Quality of Experience's Prediction in Mobile Networks : The case of VoIP service

Jean Steve TAMO TCHOMGUI, Julien Jacques, Guillaume Fraysse, Vincent Barriac, Stéphane Chrétien

► **To cite this version:**

Jean Steve TAMO TCHOMGUI, Julien Jacques, Guillaume Fraysse, Vincent Barriac, Stéphane Chrétien. Quality of Experience's Prediction in Mobile Networks : The case of VoIP service. Spring school "Statlearn: challenging problems in statistical learning", Apr 2022, Cargèse, France. hal-04023153

HAL Id: hal-04023153

<https://hal.science/hal-04023153>

Submitted on 10 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Quality of Experience's Prediction in Mobile Networks : The case of VoIP service

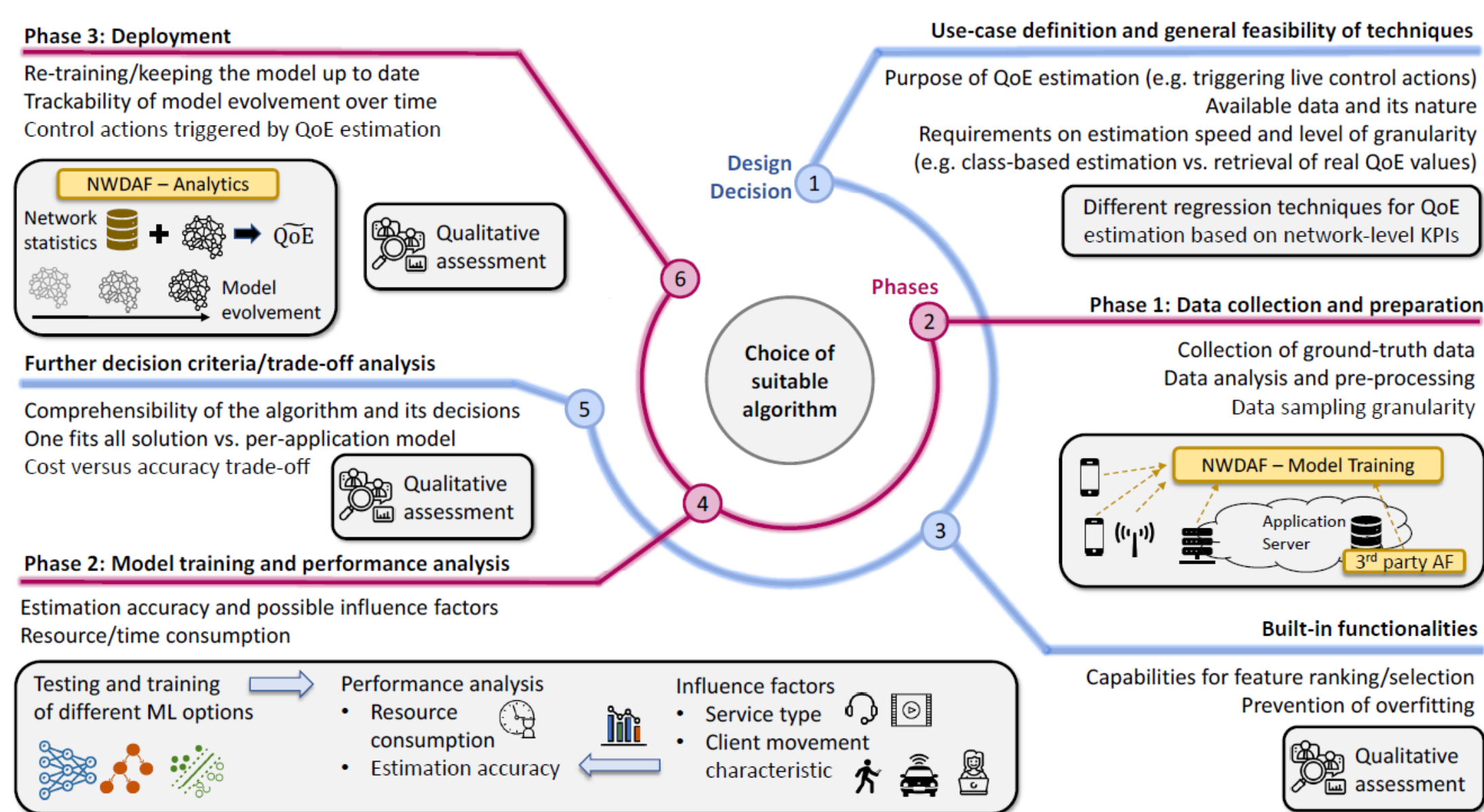
Jean Steve TAMO TCHOMGUI, Julien JACQUES, Stéphane CHRETIEN,
Guillaume FRAYSSE, Vincent BARRIAC
{jean-steve.tamo-tchomgui, julien.jacques, stephane.chretien}@univ-lyon2.fr
{guillaume.fraysse, vincent.barriac}@orange.com

Abstract

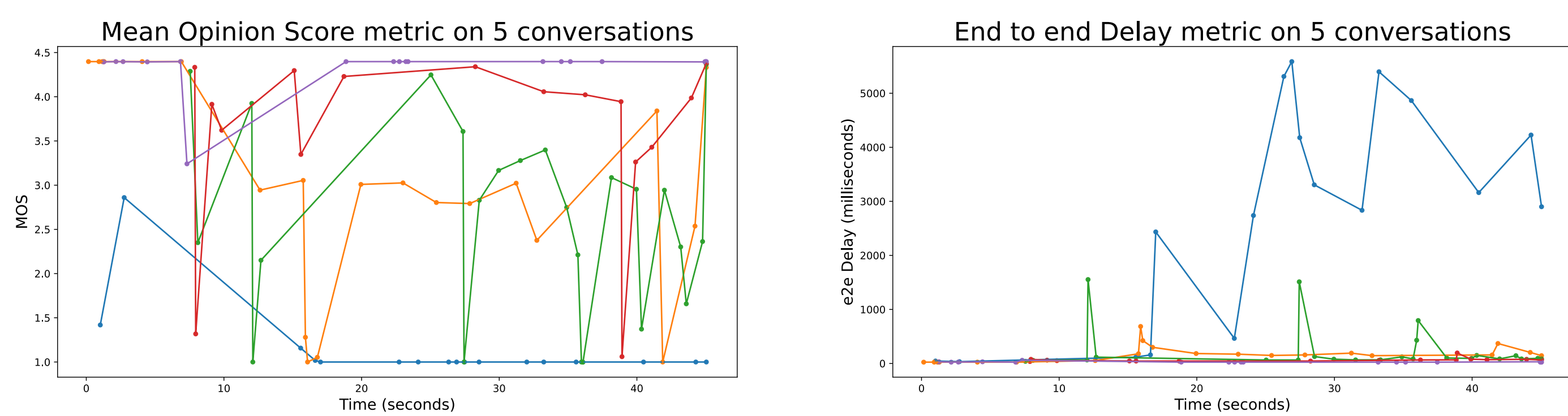
We present a comparative study of the functional data analysis approach and the approach based on the theory of signatures for a prediction problem in a Quality of Experience (QoE) monitoring projet. The problem we consider is a functional linear regression problem where both the dependent variables and the independent variables are functional variables. Our experiments demonstrate the efficiency of the signature approach for prediction in a difficult functional regression problem.

Introduction

The **Quality of Experience (QoE)**, a measure that reflects the end-user's perception of a service, as well as the factors that can influence it, is measured at high and irregular frequency over time. The problem we are considering therefore finds its natural formulation in the context of functional linear regression. Our goal is to approximate this QoE by ML models for the Voice over IP (VoIP) service in order to make recommendations for corrective actions on the network.



As we can see in the graph below, QoE in terms of Mean Opinion Score (MOS) and one factor (end-to-end delay) that can affect it are measured at high and irregular frequency.



Basis expansion

Having n realizations \mathcal{X}_i of a one variable, the goal is to construct \mathbf{X}_i curves that fit well to the data and have a regular shape such that :

$$\mathcal{X}_i = \{(x_{i,1}, t_{i,1}), (x_{i,2}, t_{i,2}), \dots, (x_{i,m_i}, t_{i,m_i})\}, 1 \leq i \leq n \text{ et } 1 \leq j \leq m_i$$

$$x_{i,j} = \mathbf{X}_i(t_{i,j}) + \varepsilon_{i,j} = \text{Snapshot of the } \mathbf{X}_i \text{ curve at } t_{i,j}.$$

A common way to reach at goal is to assume that the curves belong to a finite dimensional space spanned by a basis of functions (B-Splines, Fourier, Wavelets...) such that :

$$\mathbf{X}_i(t) = \sum_{k=1}^K \theta_{i,k} \phi_k(t) \quad [1]$$

The red line in plot is a B-splines cubic smoother with penalty on second derivative and for $K = 7$. The blue dots is the observed data.

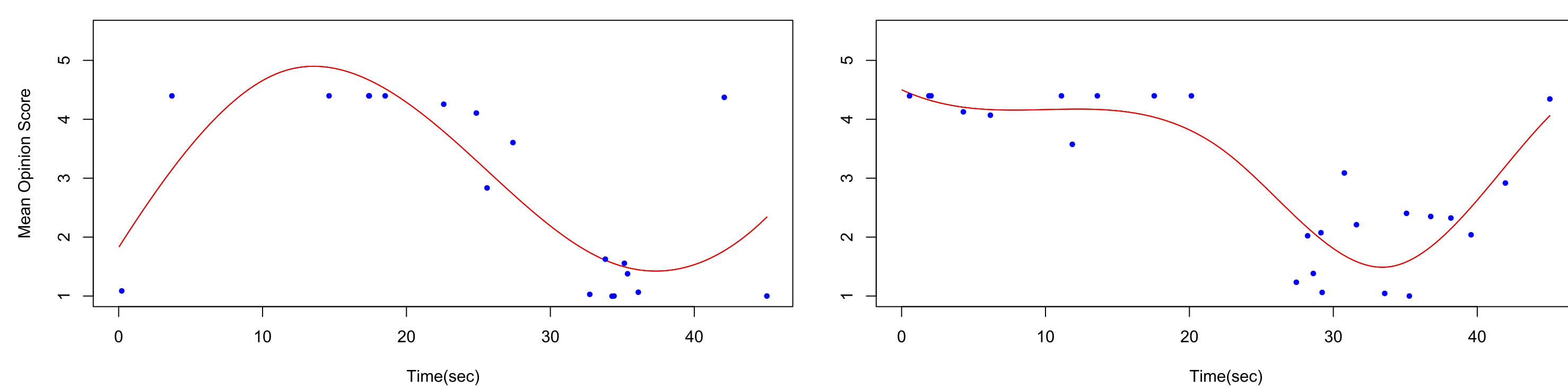


Fig. Fit data using B-Splines basis function

Model 1 : Traditionnal fonctionnal regression model

We perform the regression of a Functional variable Y representing the QoE on Functional variables $(X^k)_{1 \leq k \leq d}$ the explicative metrics.

$$Y_i(t_{i,j}) = \beta_0(t_{i,j}) + \sum_{k=1}^d \int_T \beta_k(s, t_{i,j}) X_i^k(s) ds + \varepsilon_i(t_{i,j}).$$

Signatures

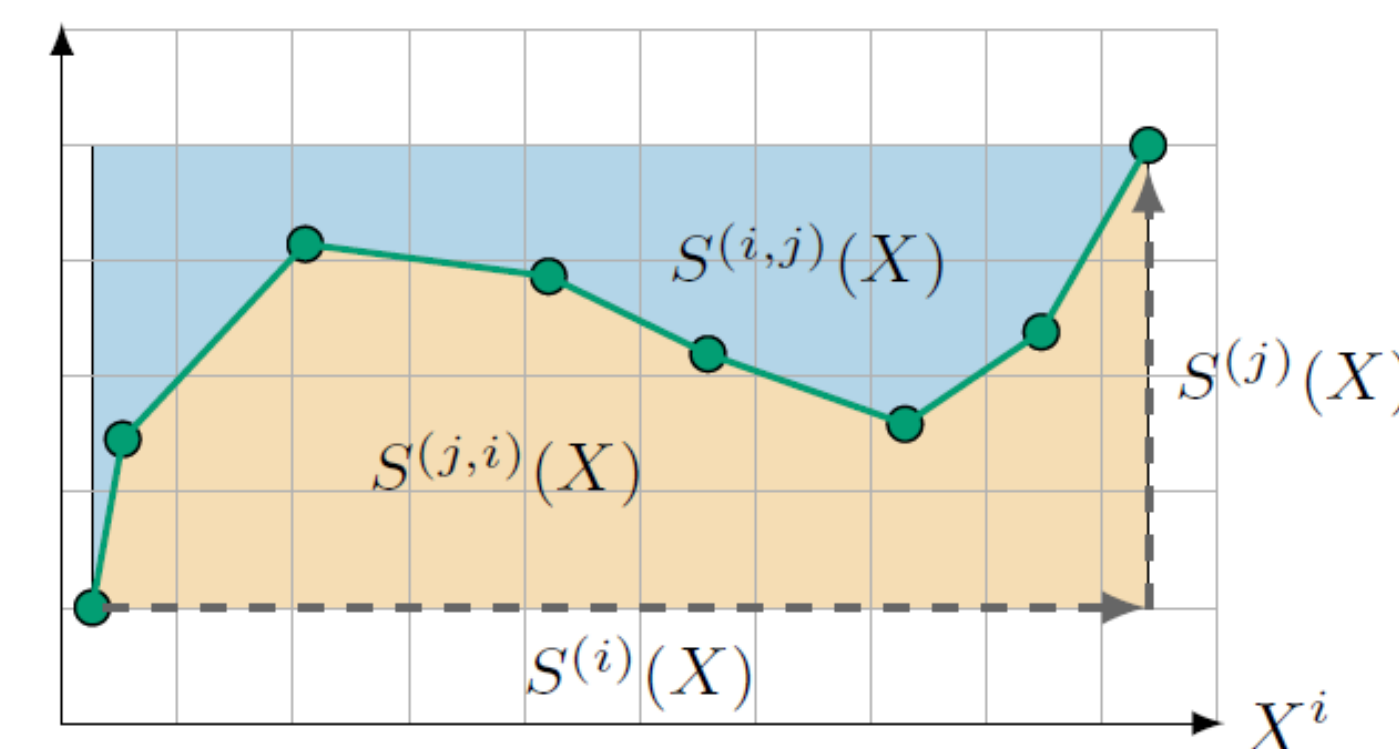
The **Signature** [2] of a (multidimensionnal) function $X : T = [a, b] \rightarrow \mathbb{R}^d$ with $d \geq 2$ is the infinite sequence given by

$$S(X)_{a,b} = \{1; S(X)_{a,b}^1; \dots; S(X)_{a,b}^d; S(X)_{a,b}^{1,1}; \dots; S(X)_{a,b}^{d,d}; \dots\},$$

$$S(X)_{a,b}^i = \int_{a < s < b} dX^i(s) = X^i(b) - X^i(a), \quad 1 \leq i \leq d,$$

$$S(X)_{a,b}^{i_1, i_2, \dots, i_k} = \int_{a < s < b} S(X)_{a,s}^{i_1, i_2, \dots, i_{k-1}} dX^{i_k}(s), \quad 1 \leq i_1, \dots, i_k \leq d.$$

X^j Geometric interpretation of signature coefficients



The terms $S^{(i)}(X)$ and $S^{(j)}(X)$ are the increments of the the coordinates i and j of the function X respectively.

The terms $S^{(i,j)}(X)$ and $S^{(j,i)}(X)$ correspond to the areas outlined by the path (the blue and orange regions respectively).

To build our model-based signature terms, we use the **Signature of X truncated at order m** , denoted $S^m(X)$, which is sequence containing all the the signature coefficients of order lower than or equal to m .

Model 2 : Linear functional regression with truncated signatures [3]

We perform here a classical multi-output regression where the features are the signature terms of X truncated at order m . The objective variable is given by the coefficients of

the functional representation of $Y : Y_i(t) = \sum_{j=1}^K \theta_{ij} \varphi_j(t) = \Theta_i \varphi(t)$ with $1 \leq i \leq n$.

$$\mathbb{E}[\Theta | X] = \langle \beta_m^*, S^m(X)_{a,b} \rangle \quad \text{et} \quad \text{Var}(\Theta | X) = \sigma^2 \leq \infty.$$

Predictive quality

To compare the two models, we use the **RMSPE** (Root Mean Square Prediction Error) :

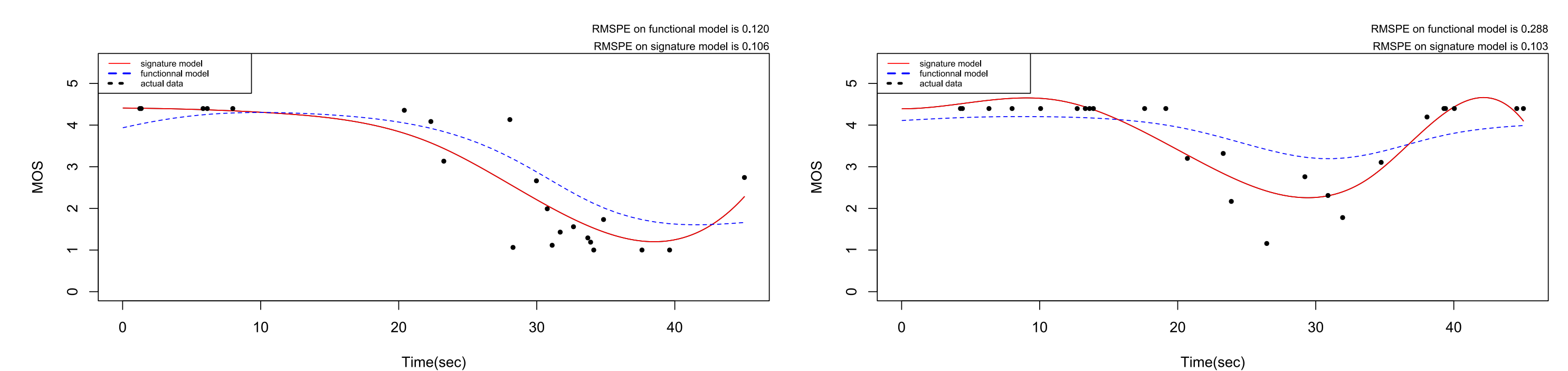
$$\text{RMSPE} = \left[\frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \left(\frac{1}{m_i} \sum_{j=1}^{m_i} \left(Y_i^{\text{test}}(t_{i,j}) - \hat{Y}_i^{\text{test}}(t_{i,j}) \right)^2 \right) \right]^{1/2}$$

As the following table shows, RMSPE value are in favor of the signature approach :

	Traditionnal Fonctionnal	Fonctionnal with Signatures
RMSPE	0.306	0.169

Tab. the test set predictive error of our two models

The prediction curves given by the studied models on two samples are represented in the graph below. We also noted the RMSPE for each sample :



Conclusion and future works

We have mainly noticed that signatures have the advantage of avoiding the step of representing the functions of the ADF, whereas this step remains unavoidable in the classical functional approach. The prediction results suggest that signatures are a very good approach to consider for functional prediction problems with functional covariates. Signatures also have the advantage of allowing the analysis of irregularly sampled data within the same individual. One of the challenges to be faced in our future work is to take advantage of the structure of the data to be analysed to understand how to scale up when we have to deal with many cohorts conjointly.

Next steps :

- Mixture of fonctionnal regression models ;
- Examination of how uncertainty propagates with the order of the signatures.

References

- [1] J. O. Ramsay and B. W. Silverman. Functional Data Analysis 2nd ed., 2005. ser. Springer Series in Statistics. New York: Springer-Verlag.
- [2] I. Chevyrev and A. Kormilitzin. A primer on the signature method in machine learning, 2016. <http://arxiv.org/abs/1603.03788> arXiv :1603.03788.
- [3] A. Fermanian. Learning time-dependent data with the signature transform, PhD Thesis, 2021, Sorbonne Université, Paris.