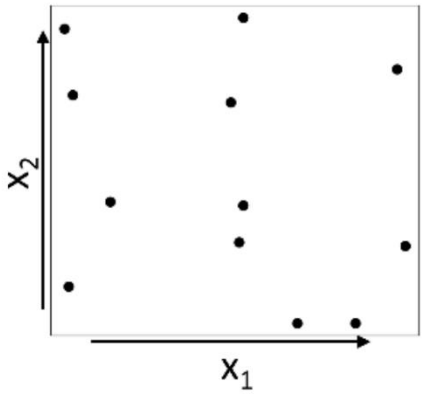
Bases de données simulées pour l'apprentissage des métamodèles

J. Betancourt** , J. Rohmer*

Séminaire de restitution du projet ANR RISCOPE, 1-2 février 2022

Approche par substitution : **métamodélisation**

Design of Experiments (DoE)
Points to perform simulations



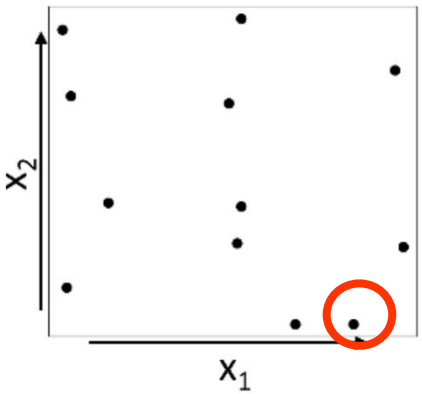
X1, X2 : paramètres d'entrée du simulateur numérique

Approche par substitution : **métamodélisation**

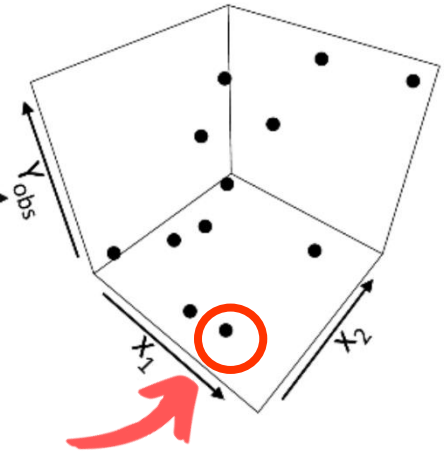
Cout de temps de calcul!

Design of Experiments (DoE)
Points to perform simulations

Simulation
Performing simulations for the points of the DoE

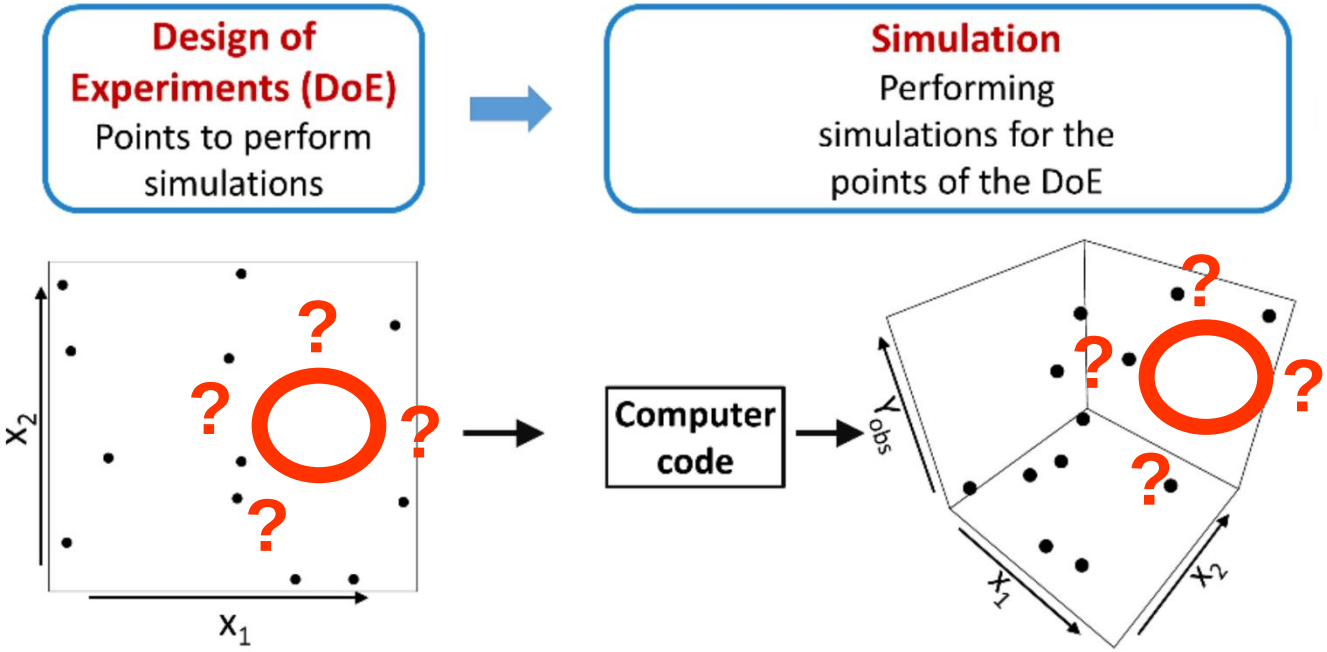


Computer code



Chaque point =
un résultat de simulation couteuse pour **X1,X2** donné

Approche par substitution : **métamodélisation**

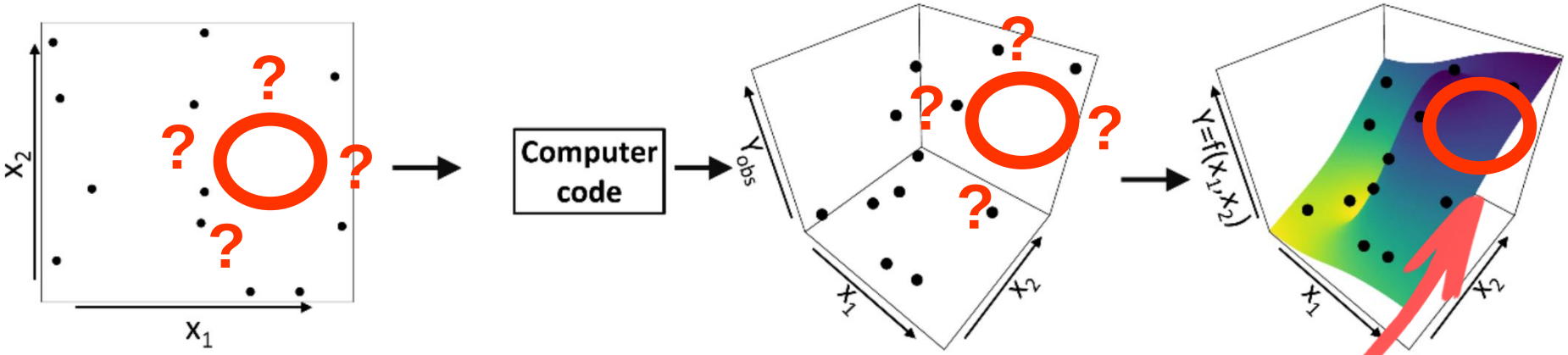


Approche par substitution : **métamodélisation**

Design of Experiments (DoE)
Points to perform simulations

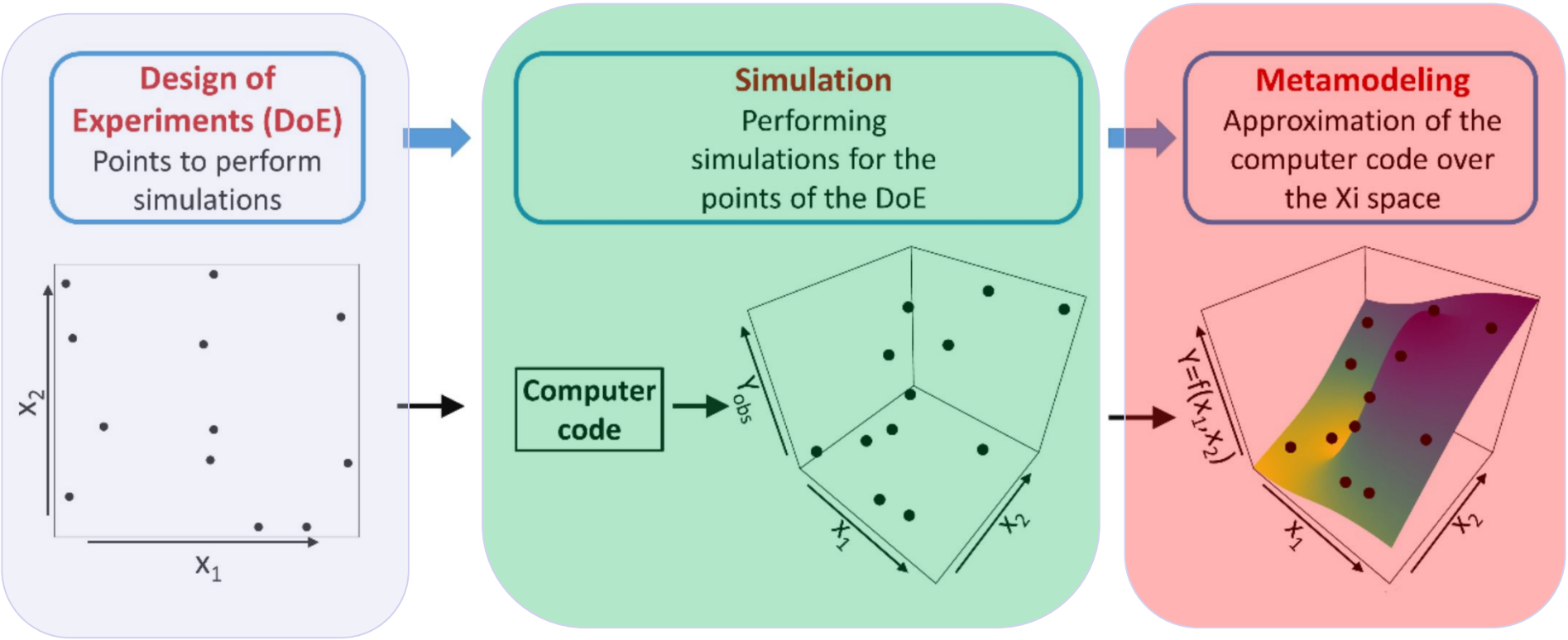
Simulation
Performing simulations for the points of the DoE

Metamodeling
Approximation of the computer code over the Xi space



Remplir les 'trous'
avec un modèle statistique
(=métamodèle)

Approche par substitution : **métamodélisation**

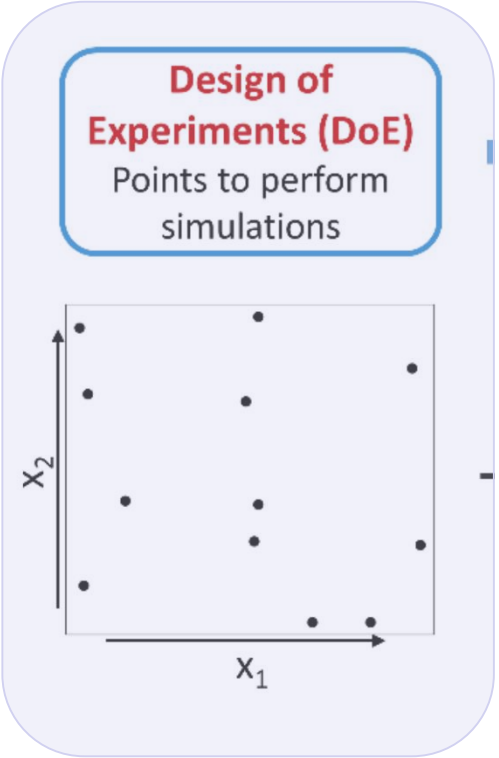


Cette présentation

Présentation de Rodrigo Pedreros (14h30, 01/02/22)

Présentation de demain (10h45, 02/02/22)

Définition du **plan d'expérience**

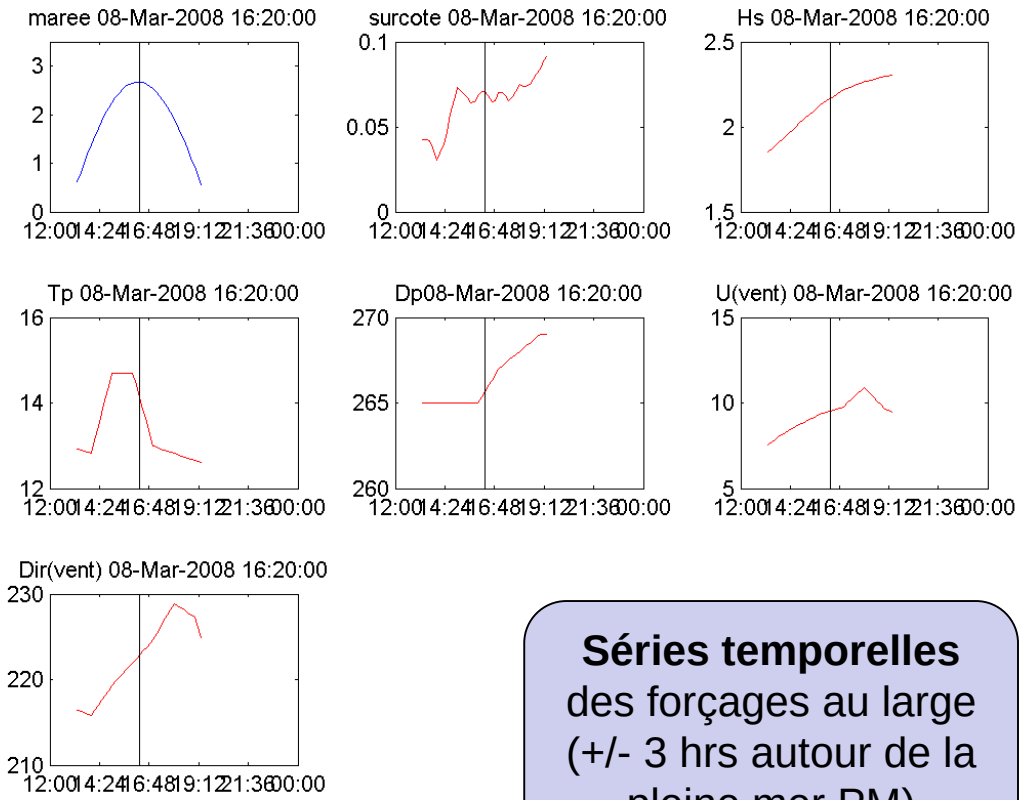


Cette présentation

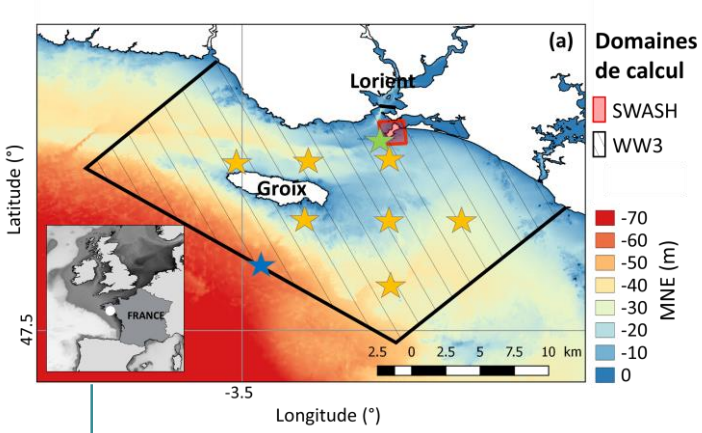
→ Comment choisir un **nombre limité (100-200)** de valeurs **X1,X2,X3,…,Xn**?

→ Comment faire cette sélection lorsque **X1,X2,X3,…,Xn** sont des **séries temporelles**?

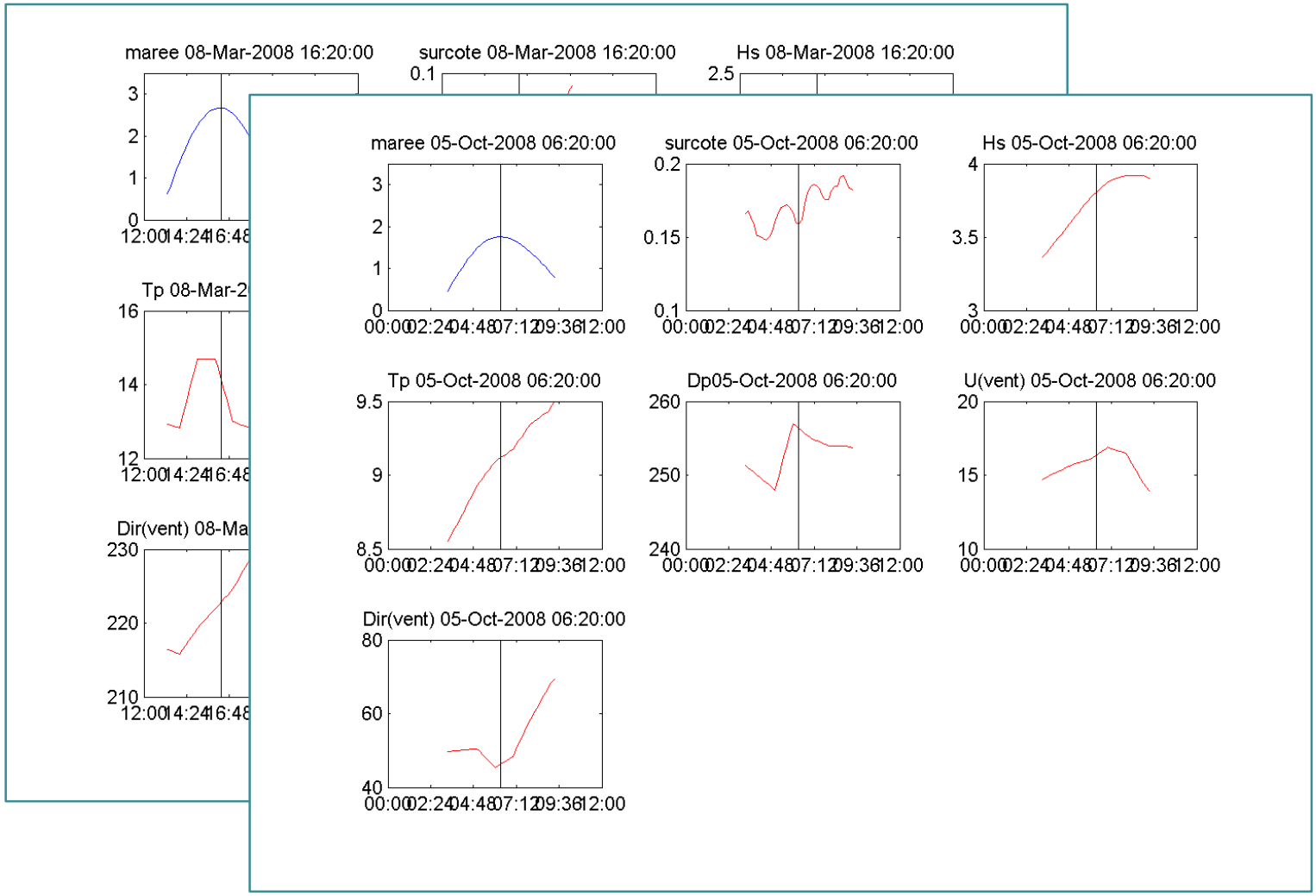
Exemple d'entrées du simulateur



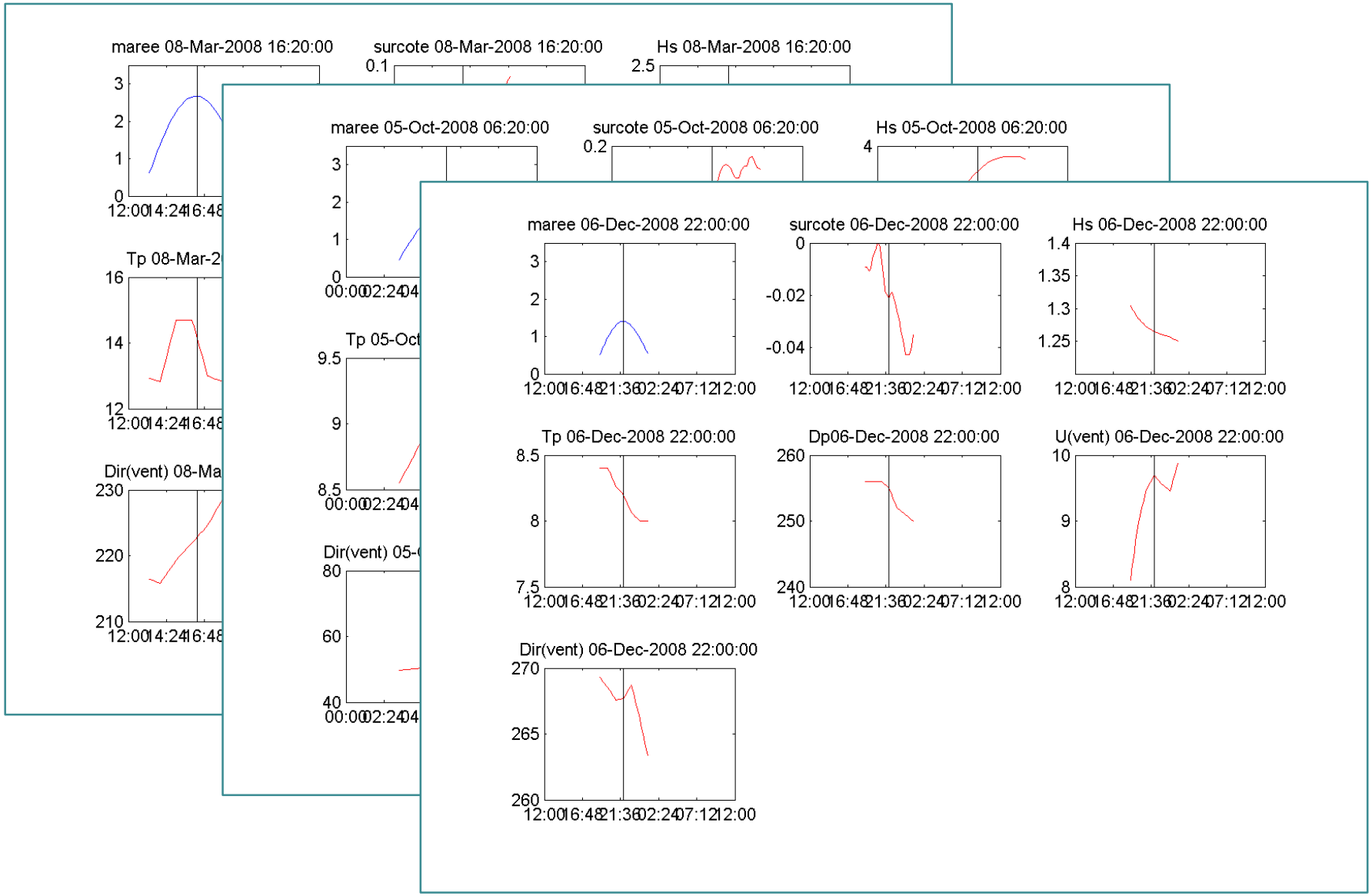
Séries temporelles des forçages au large (+/- 3 hrs autour de la pleine mer PM)



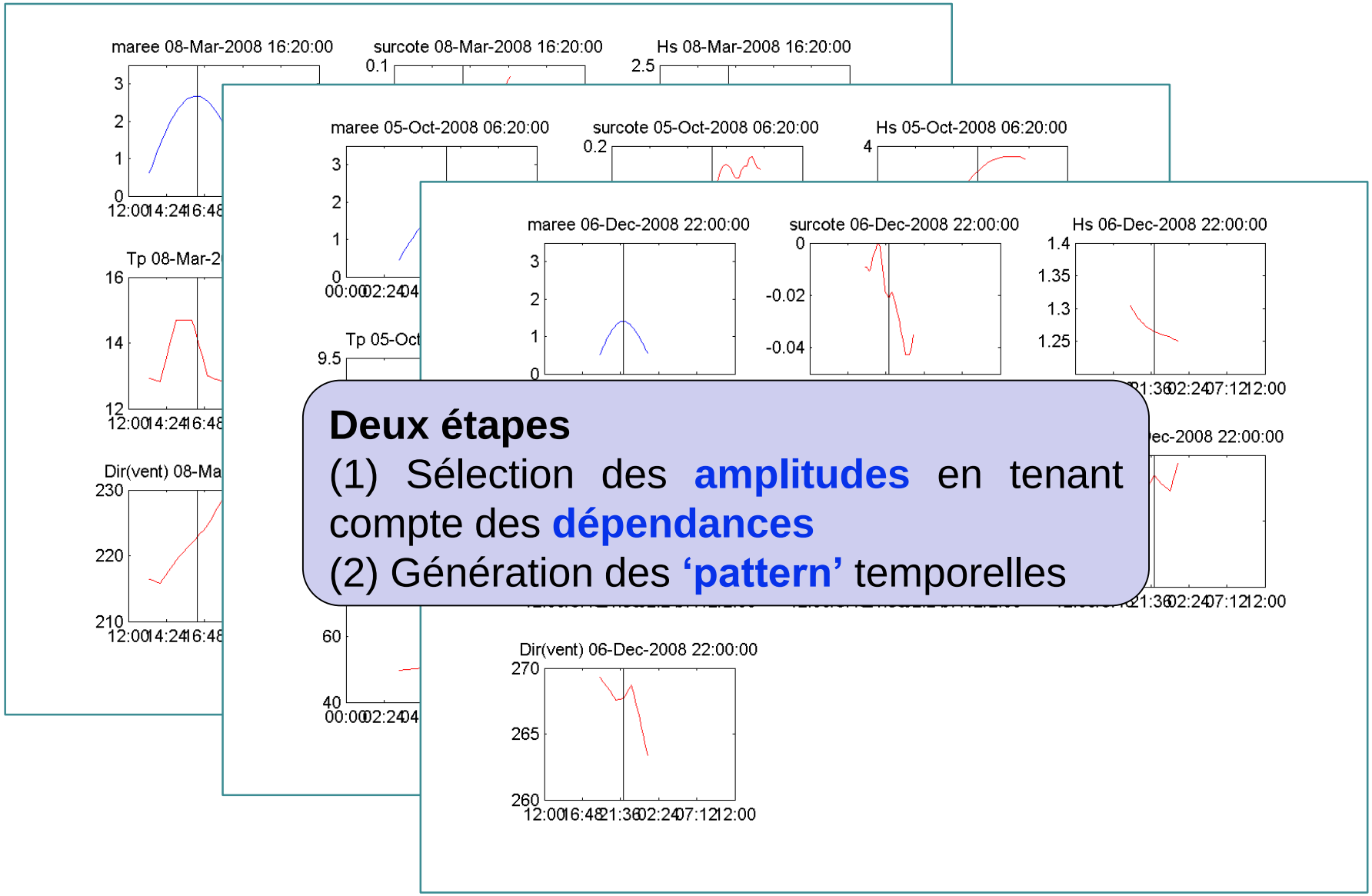
Exemple d'entrées du simulateur



Exemple d'entrées du simulateur



Exemple d'entrées du simulateur



Deux étapes

(1) Sélection des **amplitudes** en tenant compte des **dépendances**

(2) Génération des **'pattern'** temporelles

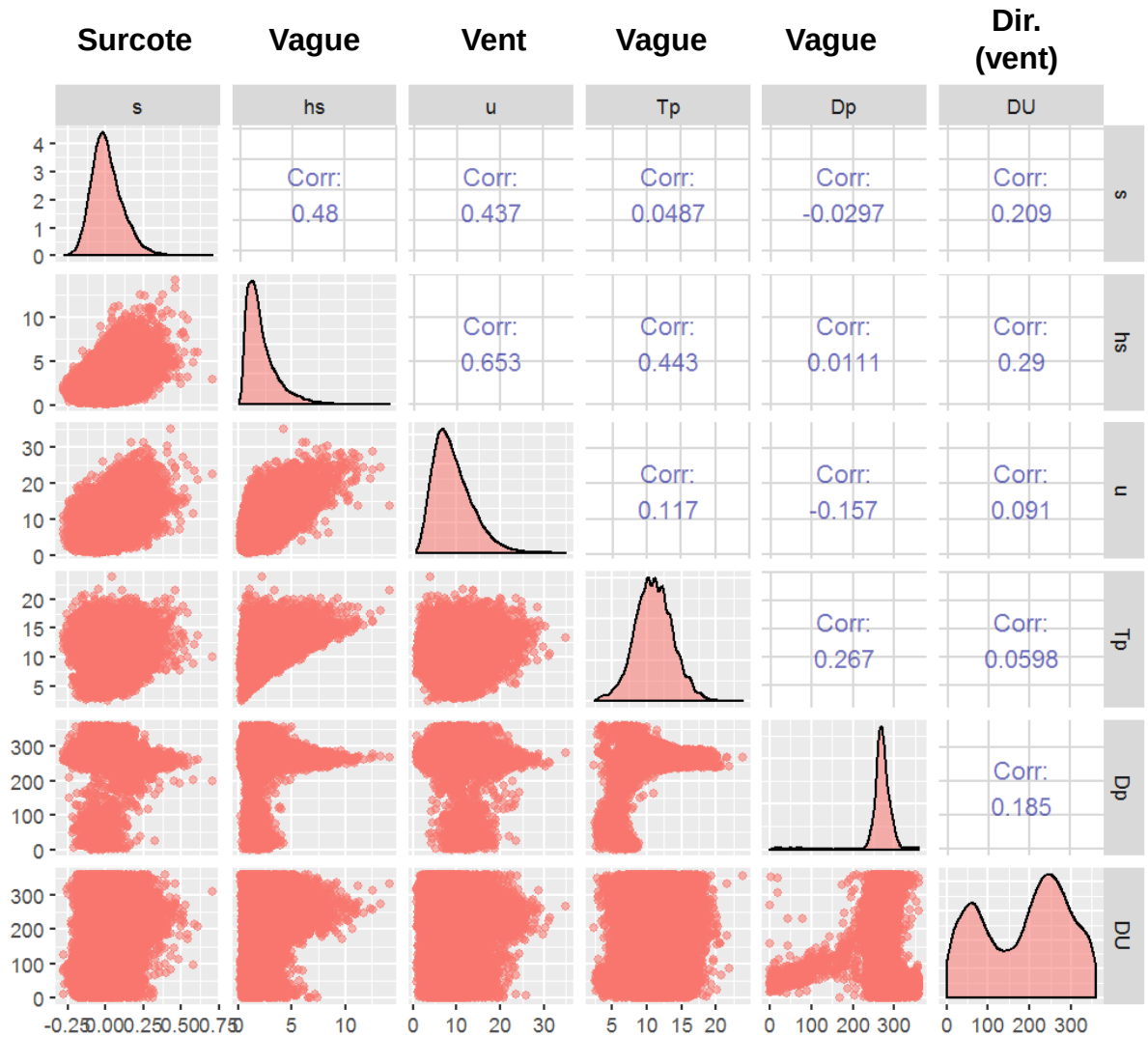
Etape (1) sélection des amplitudes

Sélection des amplitudes

- **Option 1:** sélectionner parmi les **évènements passés**

Sélection des amplitudes

- **Option 1:** sélectionner parmi les **évènements passés**



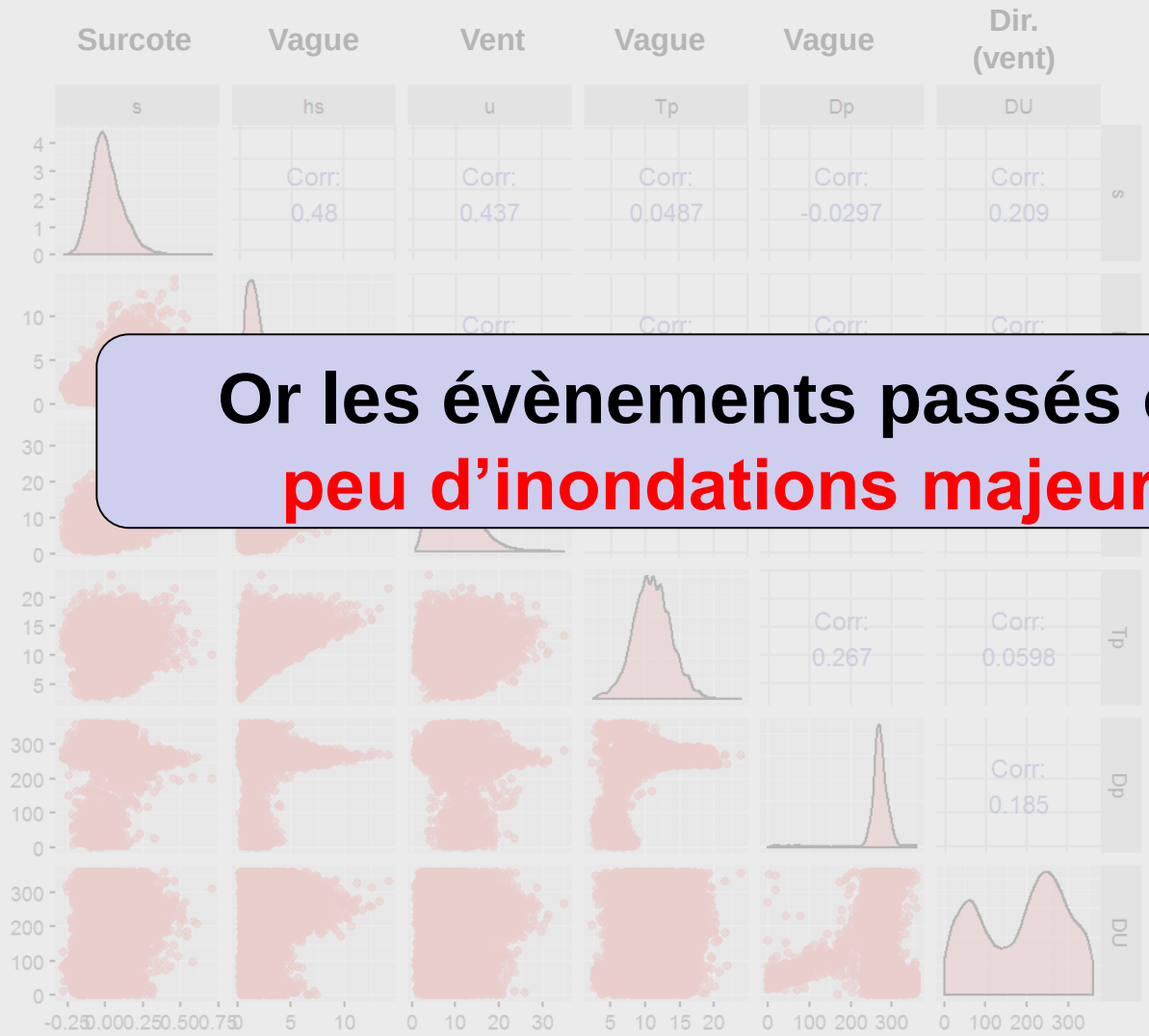
Extraction des amplitudes des séries temporelles **1900-2016 reconstituées** (et corrigées) **[1]**

- Sur -3/+3 hrs autour PM
  - Valeur max (s;hs;u)
  - Tp(à max hs)
  - Dp(à max hs)
  - DU(à max u)
- Avec PM > 2.342m (IGN69)
- Nombre total = 20557

[1] Idier et al. (2020), présentation Déborah Idier, 15h10, 01/02/22

Sélection des amplitudes

- **Option 1:** sélectionner parmi les **évènements passés**



Extraction des amplitudes des séries temporelles **1900-2016 reconstituées** (et corrigées) **[1]**

69)

Or les évènements passés ont menés à peu d'inondations majeures (<10)…

[1] Idier et al. (2020), présentation Déborah Idier, 15h00, 01/02/22

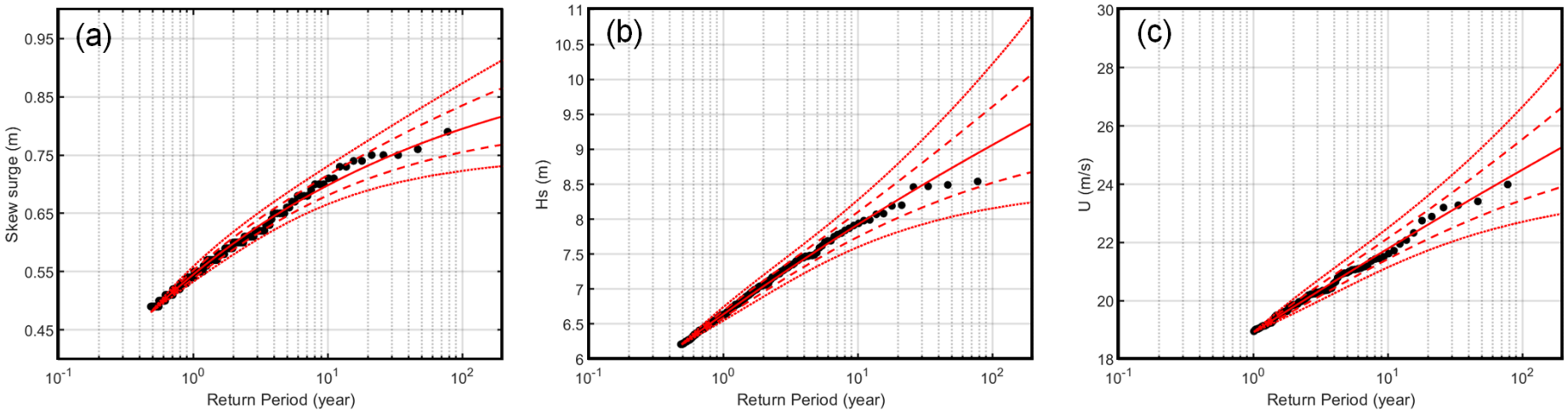
Sélection des amplitudes

- **Option 2:** générer des **évènements synthétiques extrêmes**

[1] Coles et al. (2001)

Sélection des amplitudes

- **Option 2:** générer des **évènements synthétiques extrêmes**
  - **Extrapolation des valeurs extrêmes** pour (S, Hs, U) via la théorie des valeurs extrêmes **[1]**

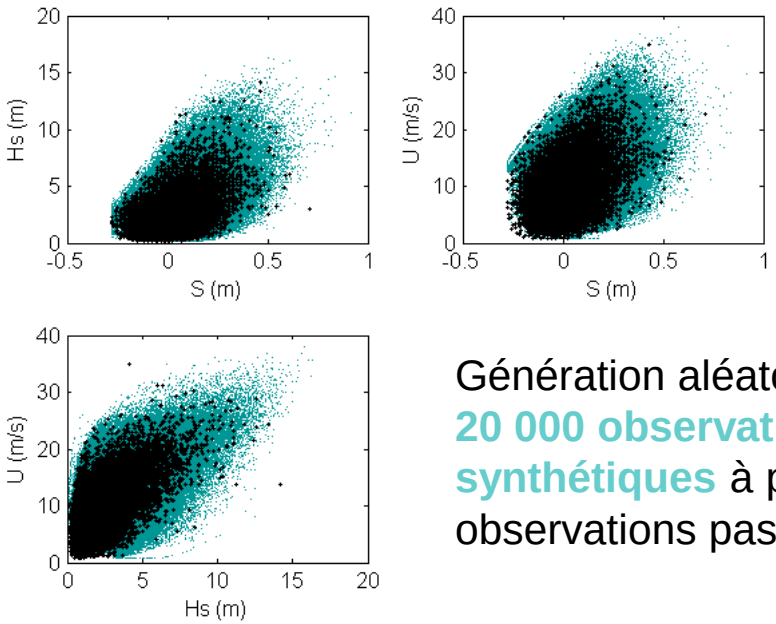


Ajustement des marginales via une loi GPD

[1] Coles et al. (2001)

Sélection des amplitudes

- **Option 2:** générer des évènements synthétiques extrêmes
  - Extrapolation des valeurs extrêmes pour (S, Hs, U) via la théorie des valeurs extrêmes **[1]**
  - **Génération aléatoire de triplets** intégrant **la dépendance conjointe** via une analyse multivariée des valeurs extrêmes **[2]**



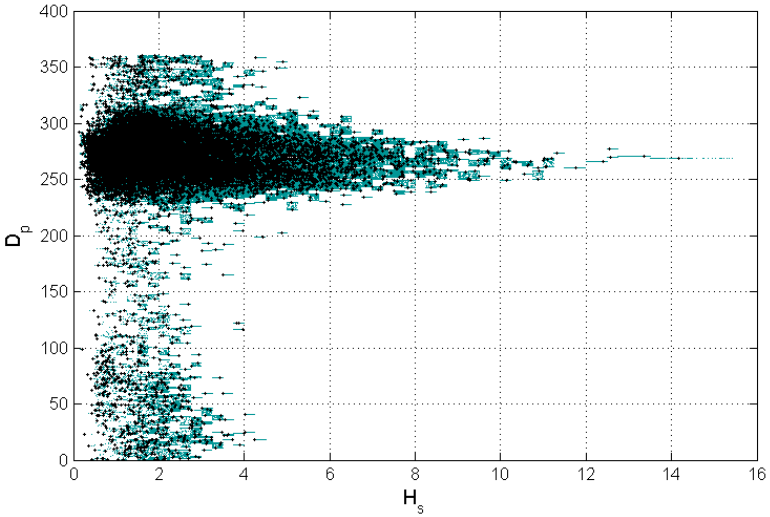
Génération aléatoire de **20 000 observations synthétiques** à partir des observations passées

[2] Heffernan and Tawn (2004)

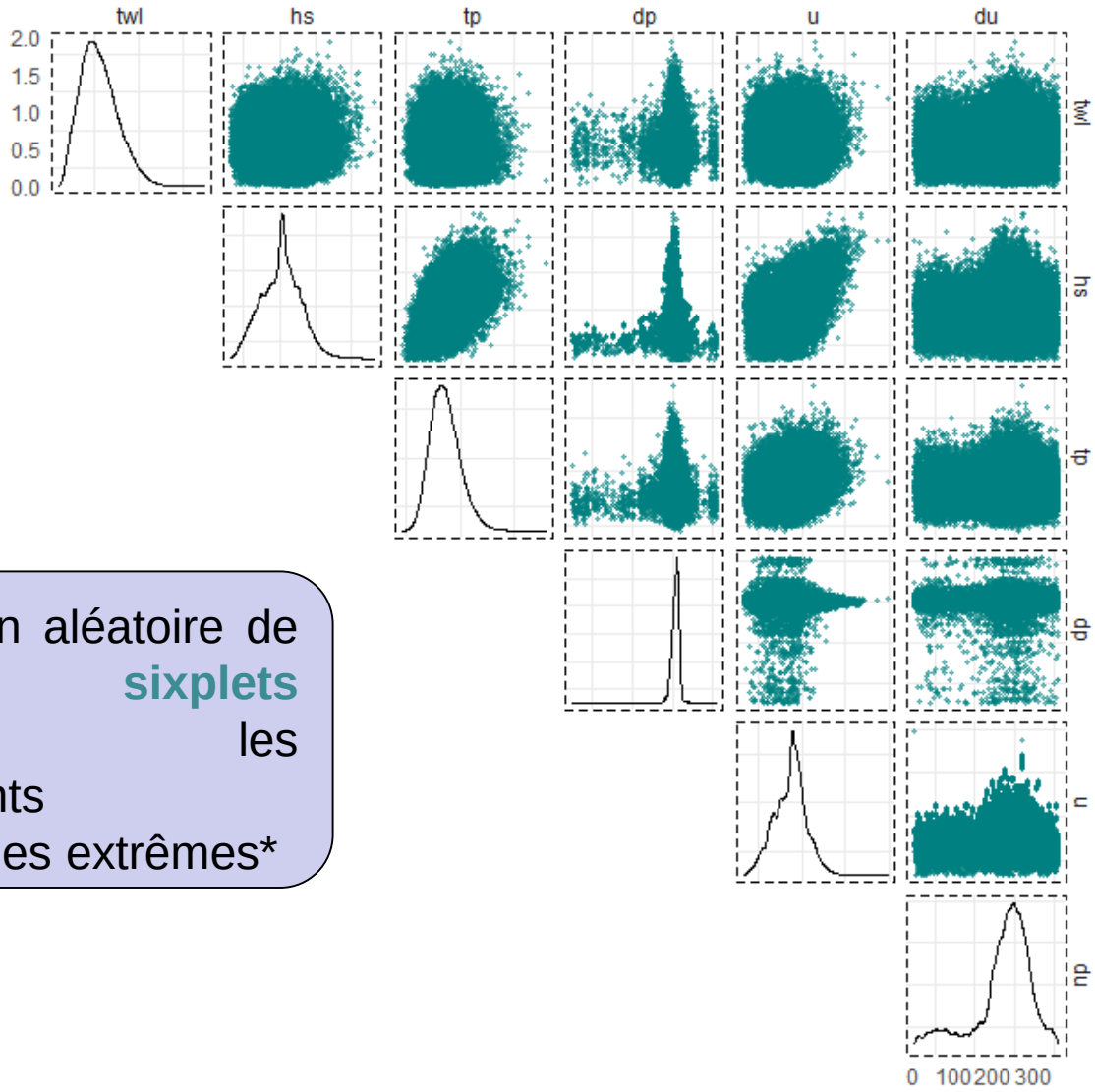
Sélection des amplitudes

- **Option 2:** générer des évènements synthétiques extrêmes
  - Extrapolation des valeurs extrêmes pour S, Hs, U via la théorie des valeurs extrêmes **[1]**
  - Génération aléatoire de triples intégrant l'interdépendance via une analyse multivariée des valeurs extrêmes **[2]**
  - Intégration de la **dépendance avec les covariables** (Tp, Dp ,Du) via une approche empirique **[3]**

Exemple dépendance (Hs,Dp)

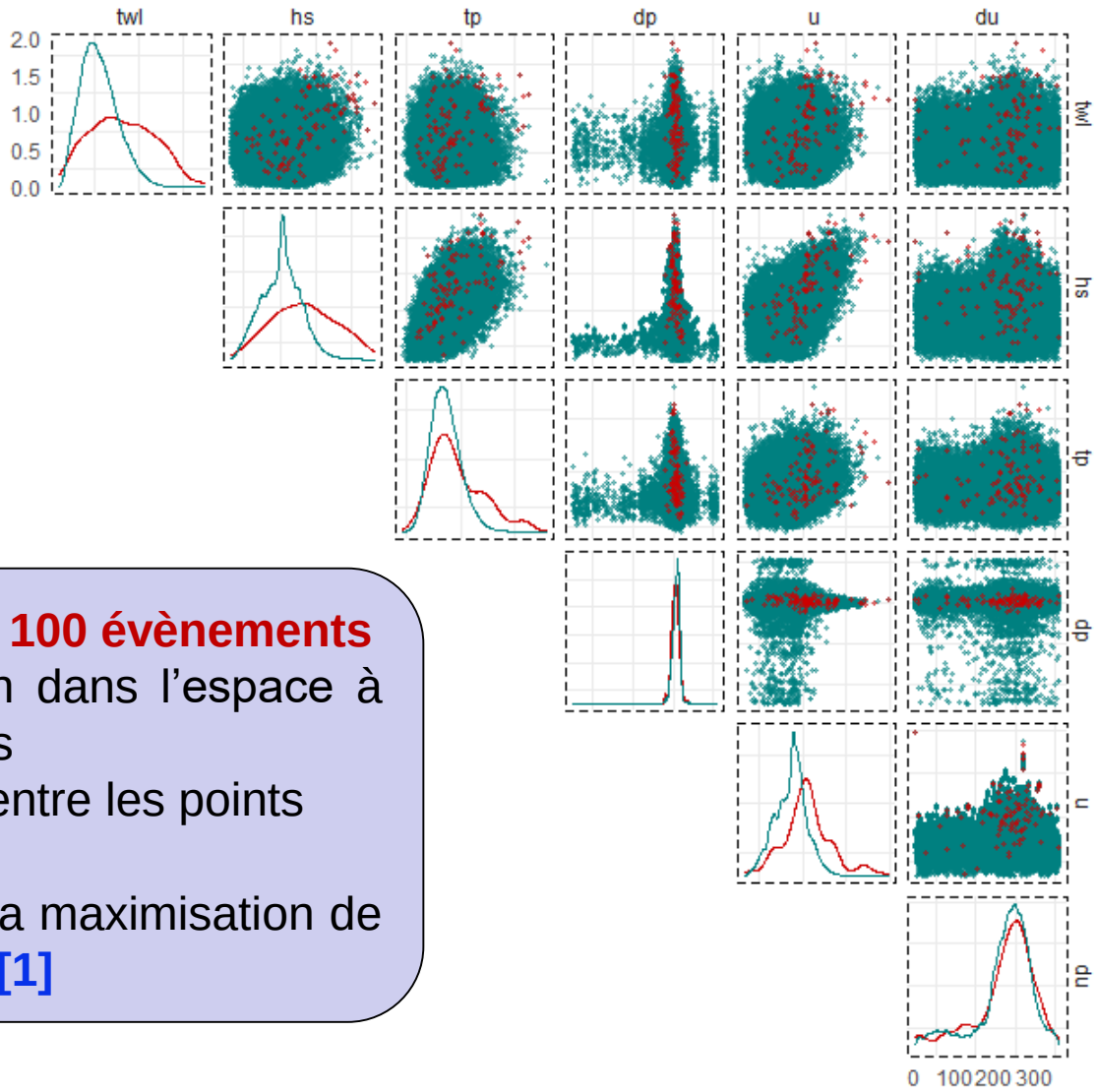


[3] Gouldby et al. (2014)



Génération aléatoire de **20000 sixplets** décrivant les évènements synthétiques extrêmes*

*5000 affichés



Sélection de **100 évènements**
→ répartition dans l'espace à 6 dimensions
→ distance entre les points
Méthode:
Clustering via maximisation de dissimilarité **[1]**

[1] Camus et al. (2010)

Etape (2) génération des séries temporelles

Plan d'expériences fonctionnel - **Objectif**

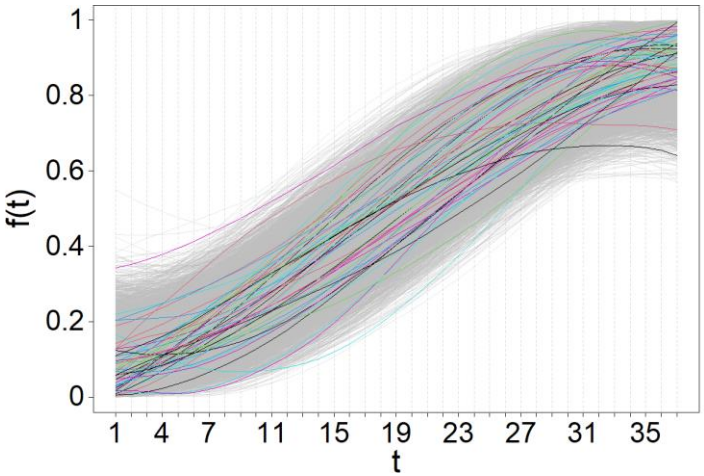
Extension aux entrées fonctionnelles

- Transformer points scalaires extremes en points fonctionnels extremes
- Très important:
  - générer des courbes avec **sens physique**
  - gardant caractéristiques statistiques et géométriques historiques

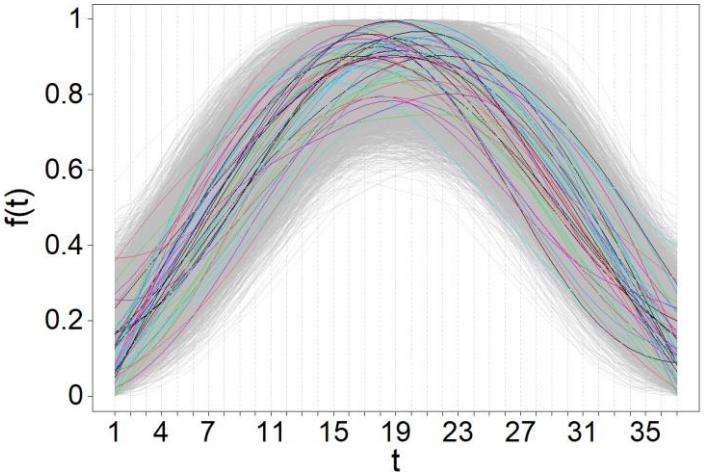
Plan d'expériences fonctionnel - **Exemple**

Deux entrées fonctionelles

Entrée 1 →



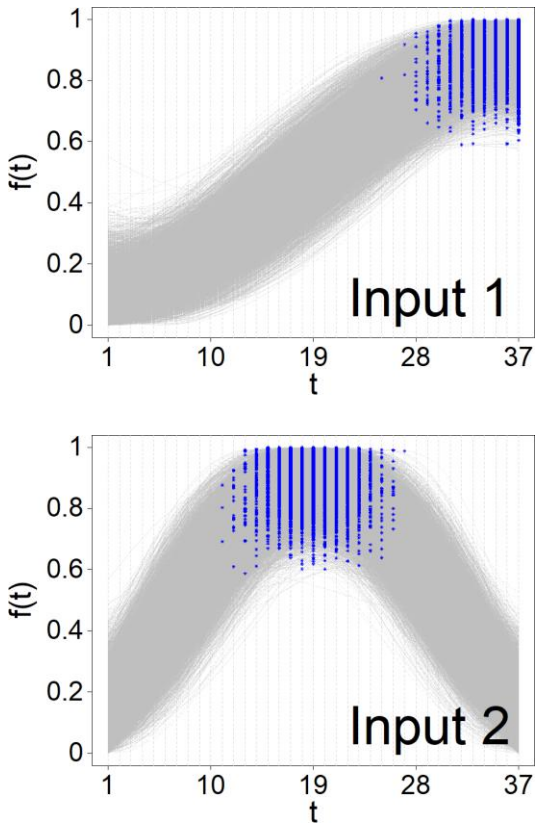
Entrée 2 →

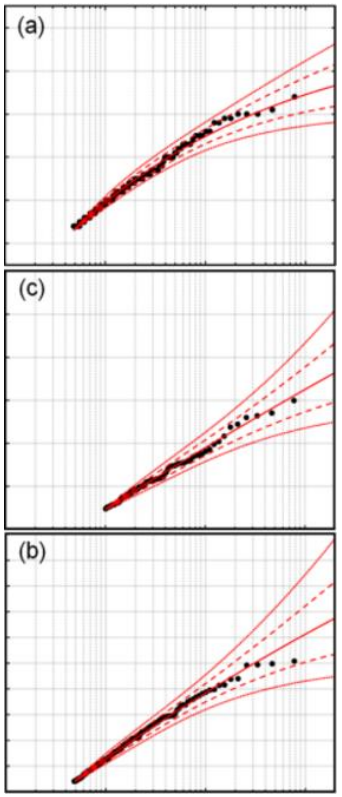


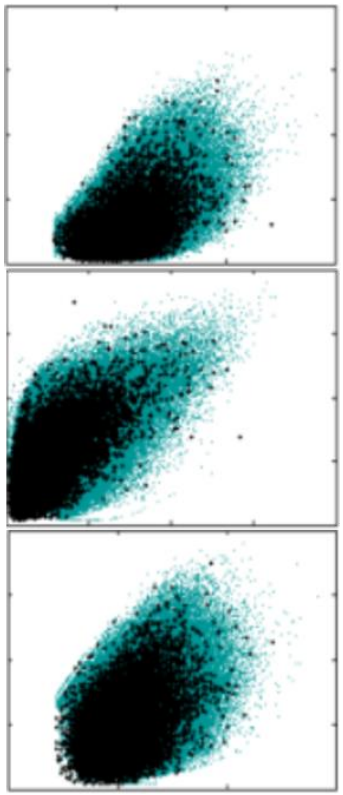
Plan d'expériences fonctionnel – Étapes précédentes

Generation des points scalaires extrêmes

1. Extraire représentation scalaire des courbes historiques
2. Apprendre le comportement des représentations
3. Extrapoler pour obtenir des points extrêmes





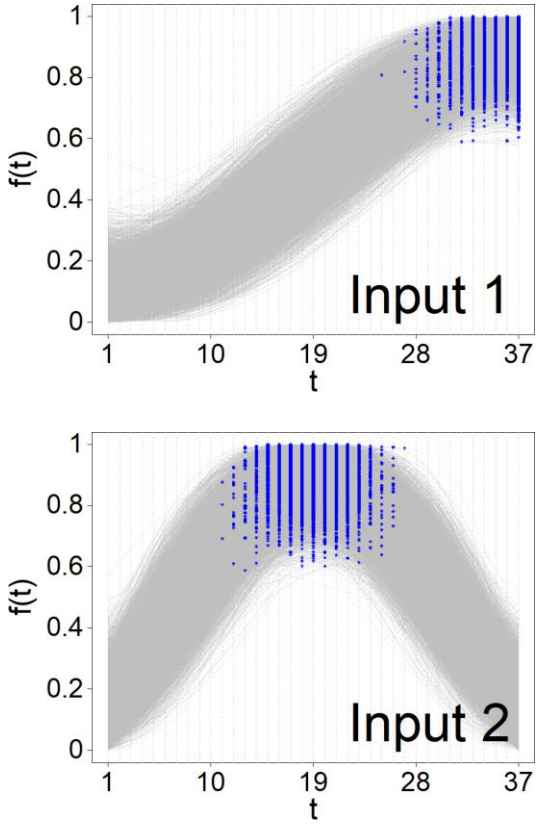


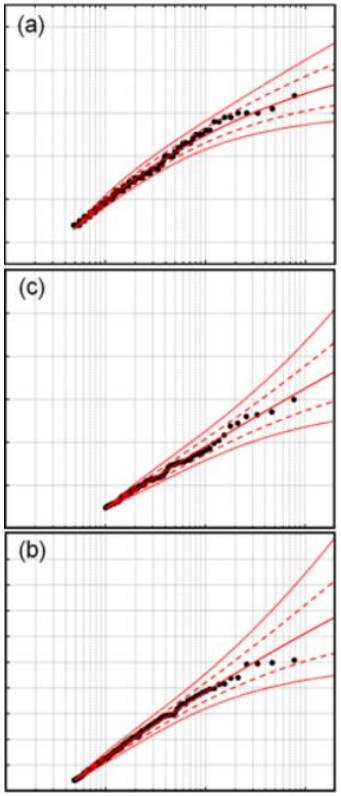
Plan d'expériences fonctionnel – **Étapes précédentes**

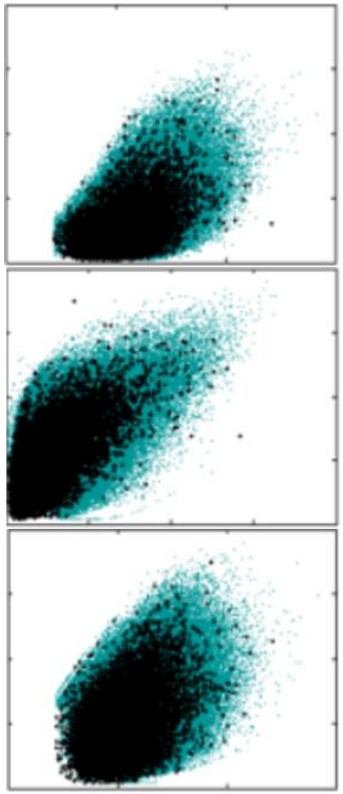
Generation des points scalaires extrêmes

1. Extraire représentation scalaire des courbes historiques
2. Apprendre le comportement des représentations
3. Extrapoler pour obtenir des points extrêmes

Important: ces valeurs extrêmes n'ont pas un instant associé







Plan d'expériences fonctionnel - **Procédure**

Obtention des points fonctionnels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
    - localisation temporelle de la représentation scalaire
    - forme typique des courbes
2. Localiser chaque valeur extrême scalaire dans un pas de temps
3. Échantillonner les 36 valeurs restants de chaque courbe

Plan d'expériences fonctionnel - **Procédure**

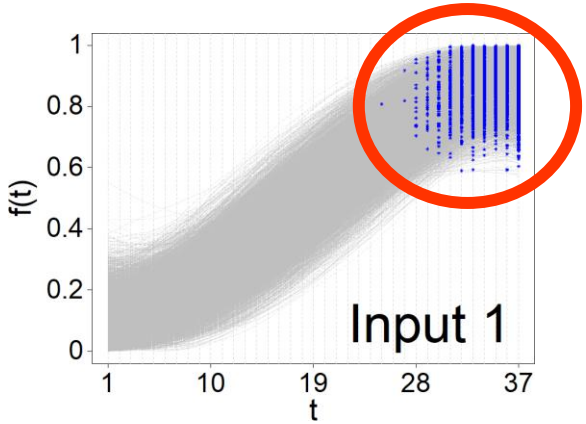
Obtention de points fonctionels extrêmes

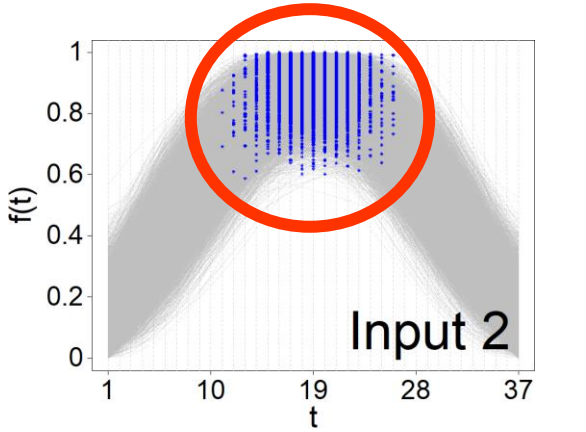
1. Apprendre le comportement de chaque entrée comme fonction
    - localisation temporelle de la représentation scalaire
    - forme typique des courbes

Plan d'expériences fonctionnel - **Procédure**

Obtention de points fonctionels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





Classificateur probabiliste – Quadratic Discriminant Analysis

Entrainement

- Utilise représentations scalaires historiques
- Apprend distribution du pas de temps conditionné a la valeur de la représentation

Classification

- Reçoit une valeur scalaire: ***s***
- Retourne un vecteur de probabilités:

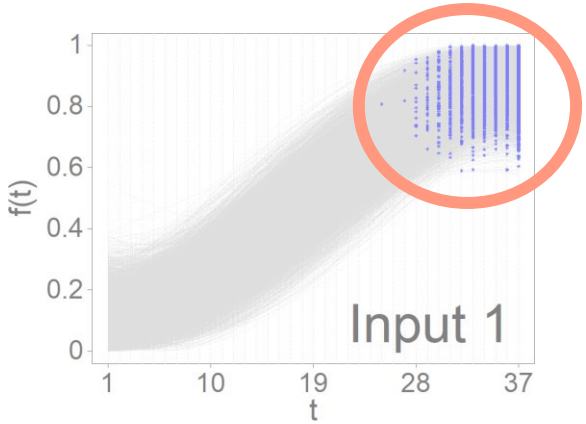
$$P(t'=t \mid f(t') = s), \quad \forall\, t$$

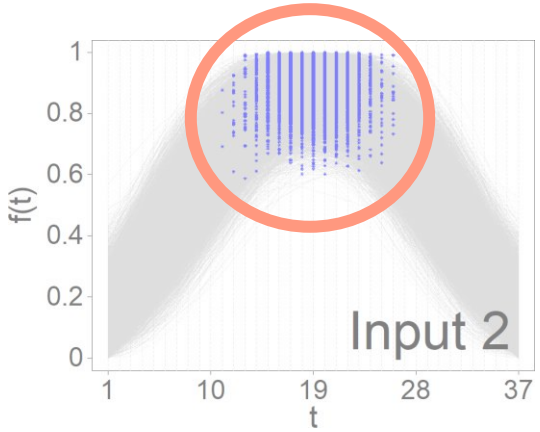
Package R MASS (Venables & Ripley, 2022)

Plan d'expériences fonctionnel - **Procédure**

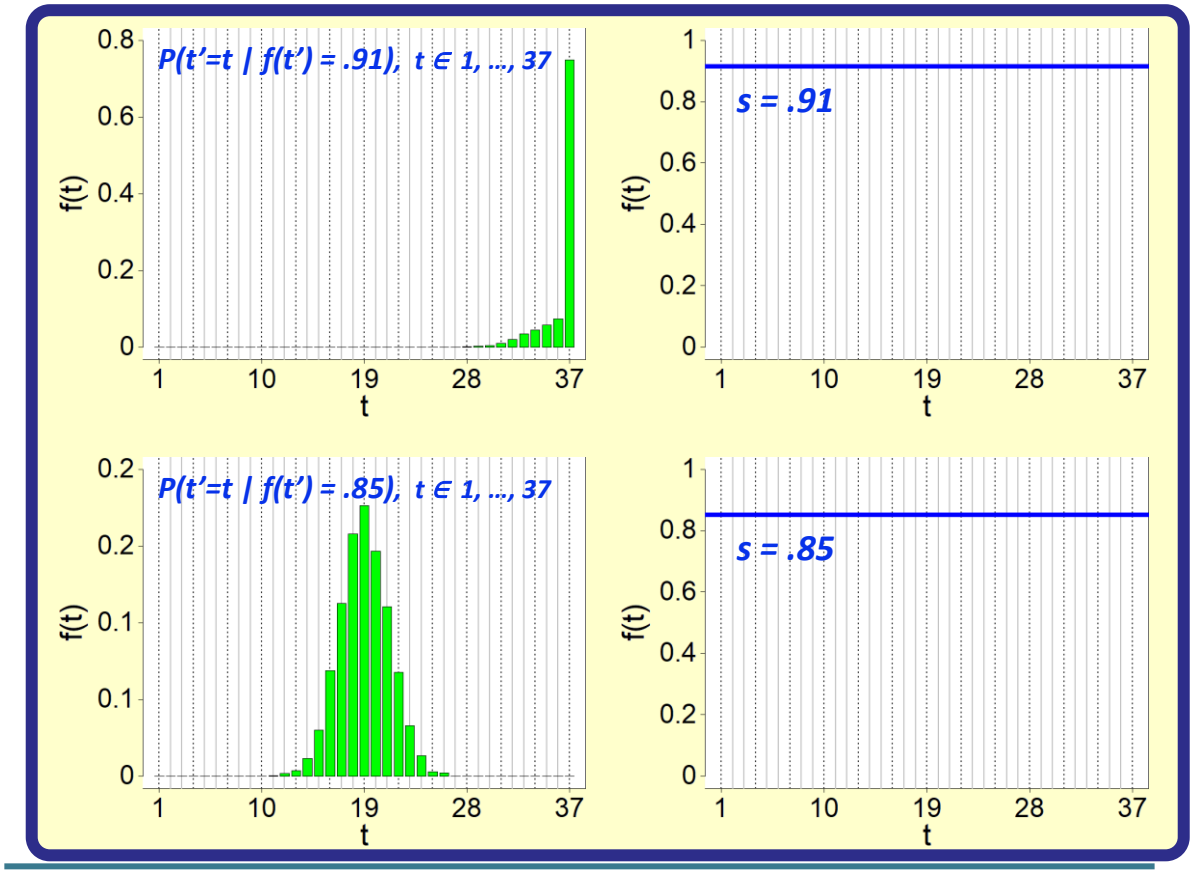
Obtention de points fonctionels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





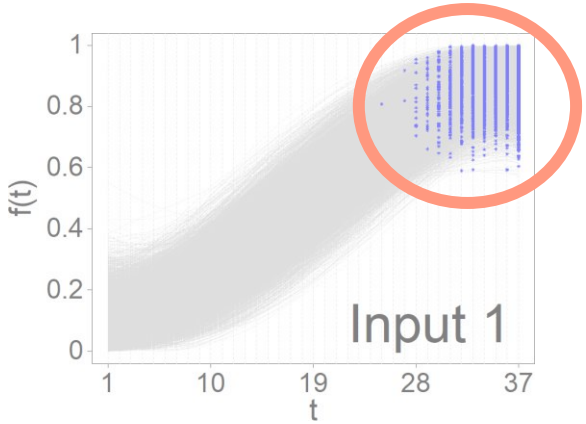
Classificateur probabiliste – Quadratic Discriminant Analysis

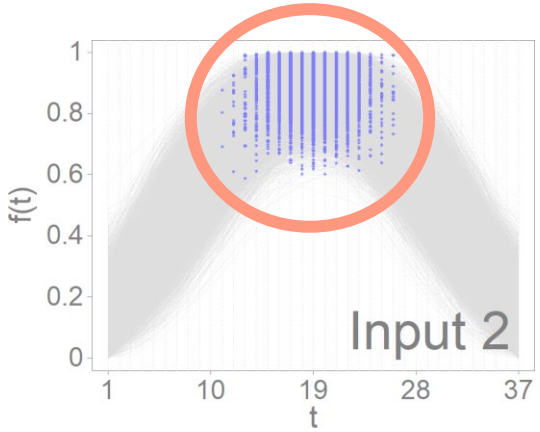


Plan d'expériences fonctionnel - **Procédure**

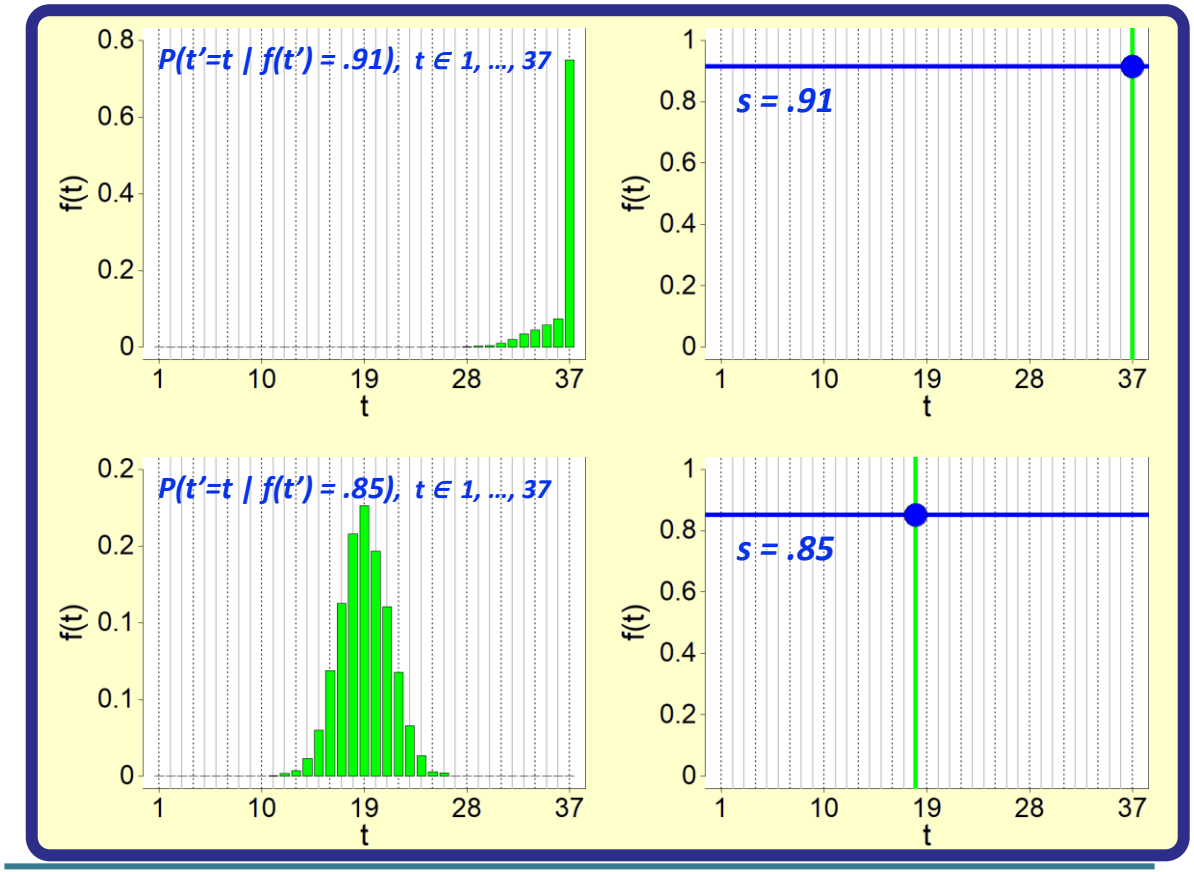
Obtention de points fonctionels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





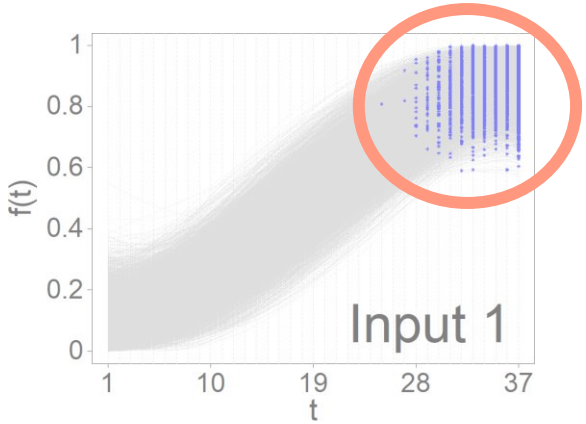
Classificateur probabiliste – Quadratic Discriminant Analysis

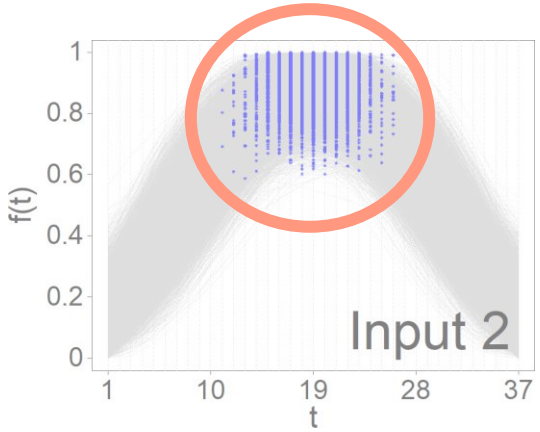


Plan d'expériences fonctionnel - **Procédure**

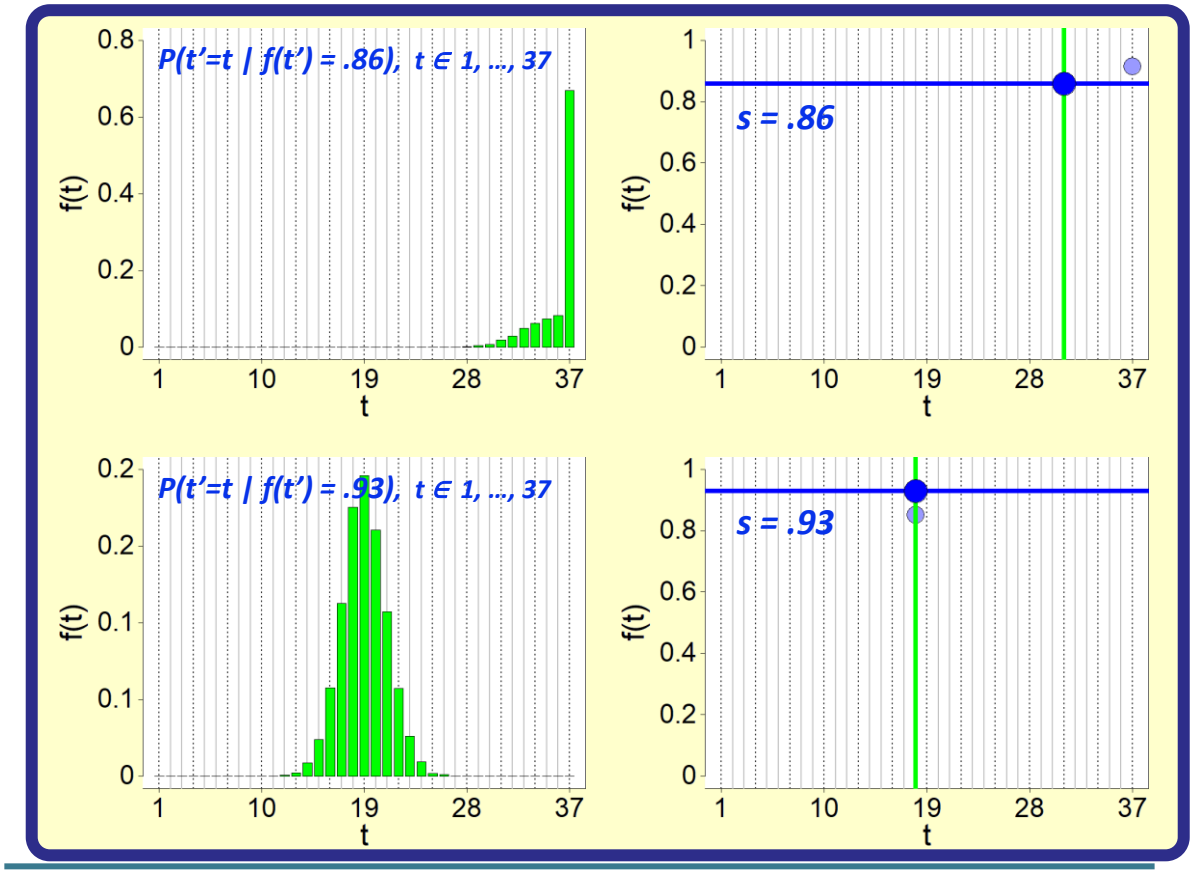
Obtention de points fonctionels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





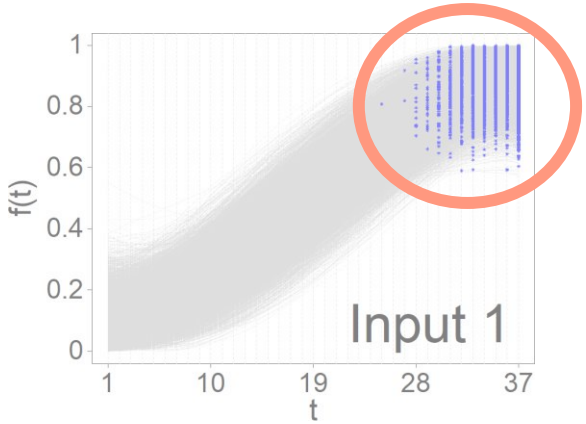
Classificateur probabiliste – Quadratic Discriminant Analysis

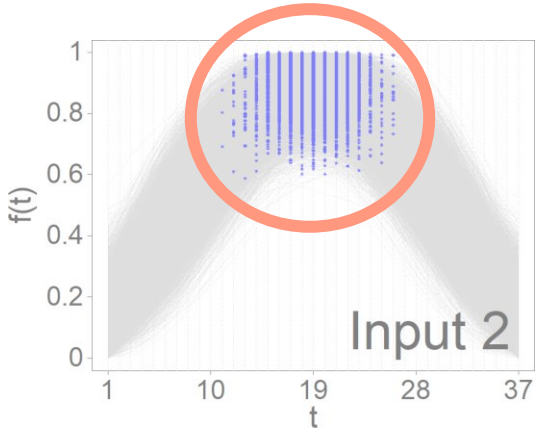


Plan d'expériences fonctionnel - **Procédure**

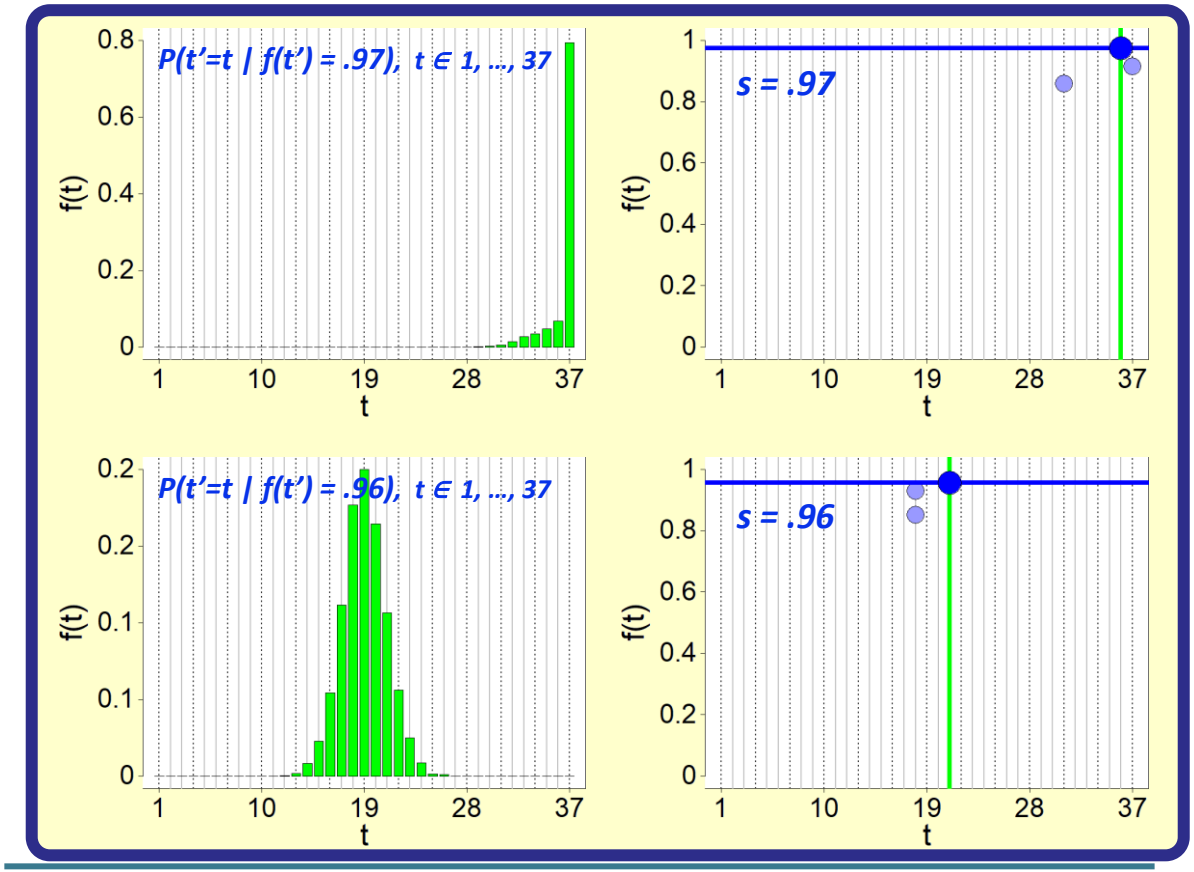
Obtention de points fonctionels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





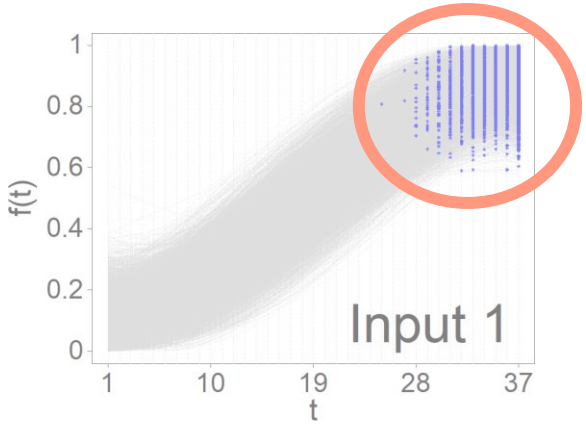
Classificateur probabiliste – Quadratic Discriminant Analysis

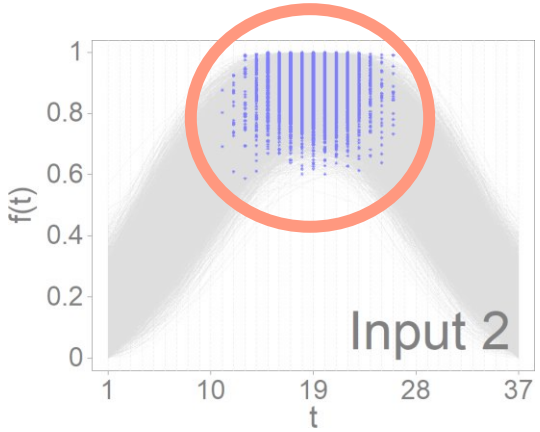


Plan d'expériences fonctionnel - **Procédure**

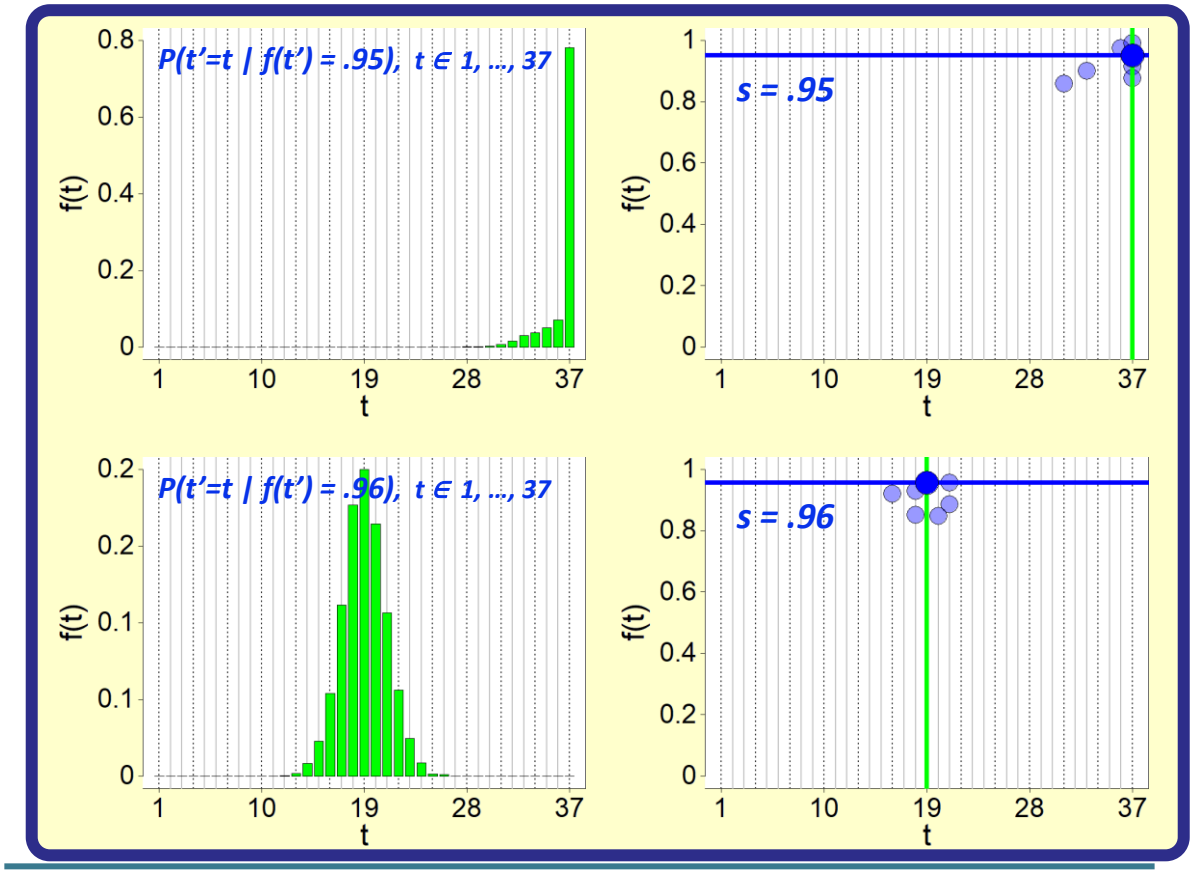
Obtention de points fonctionels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





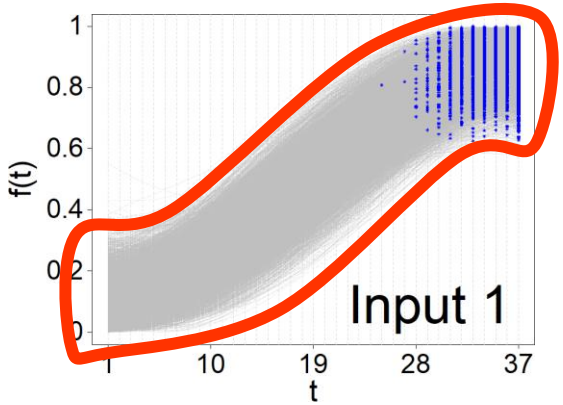
Classificateur probabiliste – Quadratic Discriminant Analysis

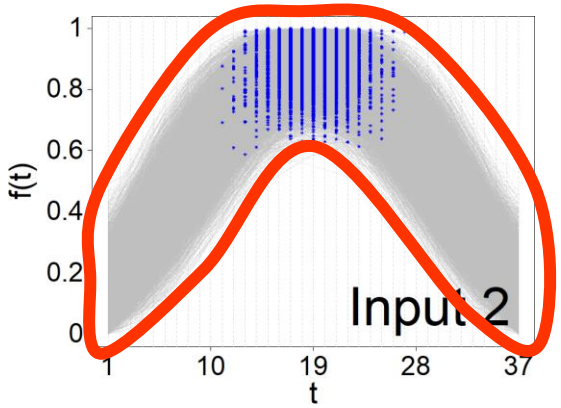


Plan d'expériences fonctionnel - **Procédure**

Obtention de points fonctionels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





Distribution Multivarié Gaussienne

Entrainement

- Utilise courbes historiques
- Apprend distribution de ***f(t)*** pour tout ***t***
- Apprend corrélation ***c(f(t), f(t'))*** pour tout ***(t, t')***

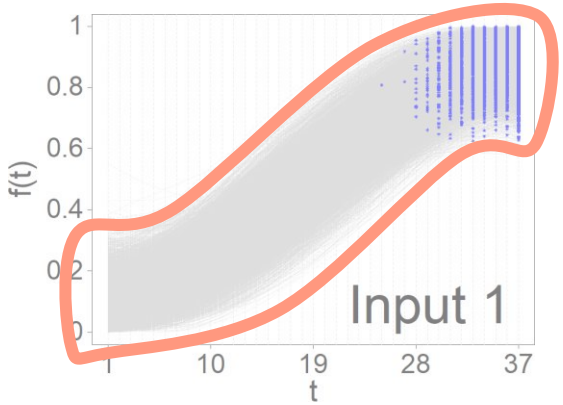
Echantillonnage

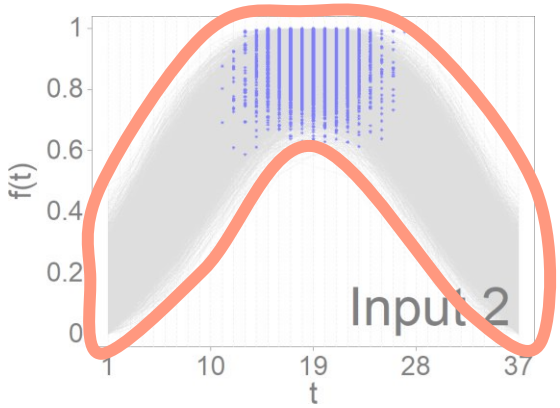
- Vecteur aléatoire des valeurs ***f(t)*** pour tout ***t***
- Chaque vecteur aléatoire *une courbe*

Plan d'expériences fonctionnel - **Procédure**

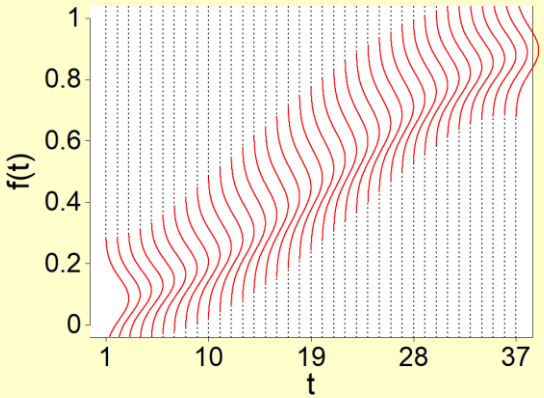
Obtention de points fonctionels extrêmes

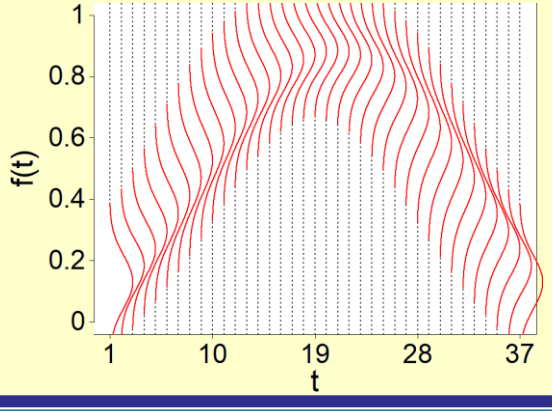
1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





Distribution Multivarié Gaussienne

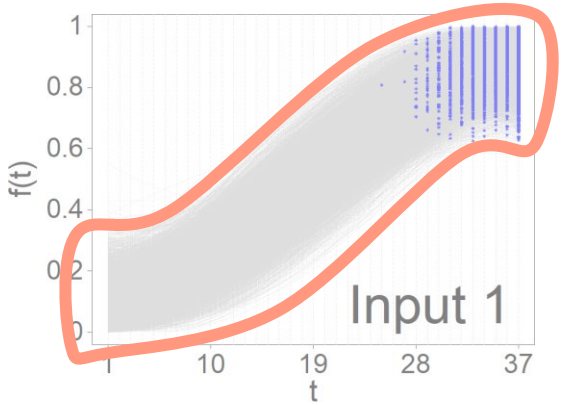


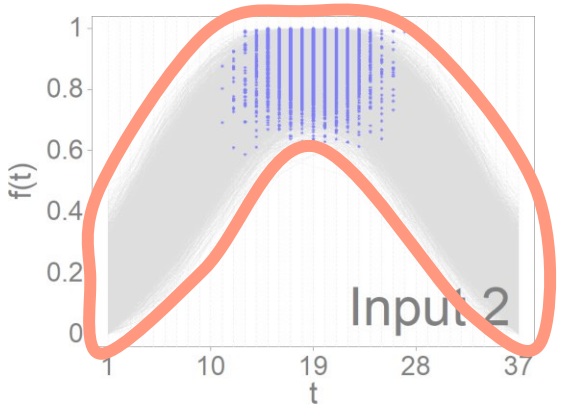


Plan d'expériences fonctionnel - **Procédure**

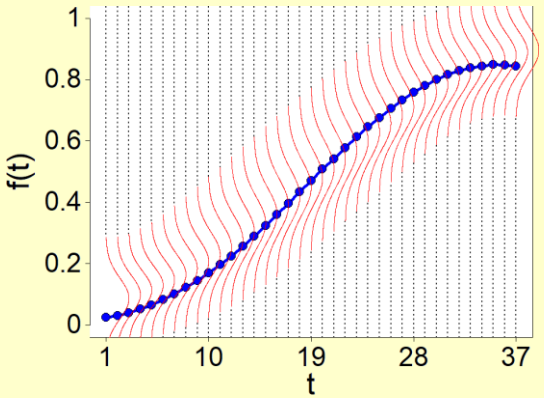
Obtention de points fonctionels extrêmes

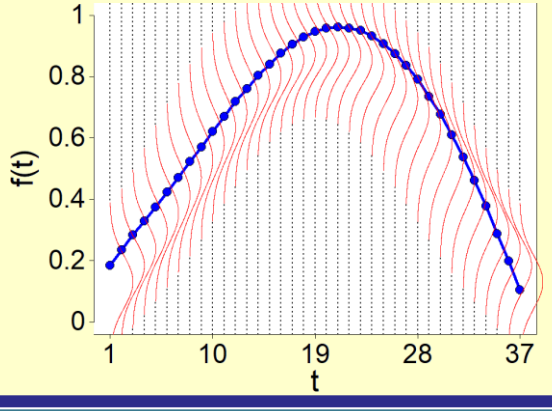
1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





Distribution Multivarié Gaussienne

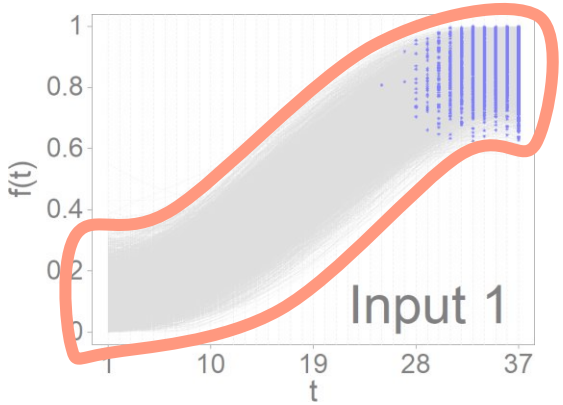


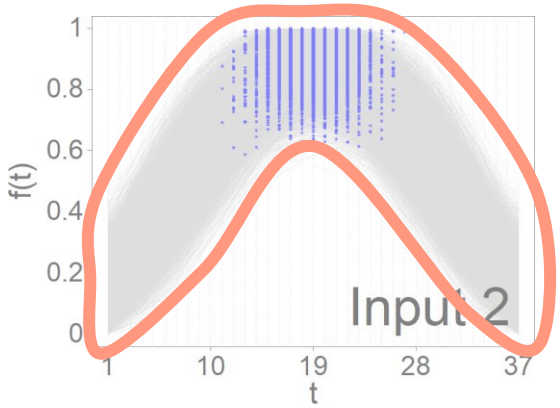


Plan d'expériences fonctionnel - **Procédure**

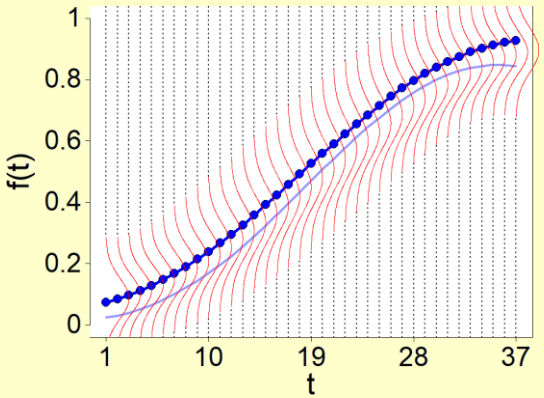
Obtention de points fonctionels extrêmes

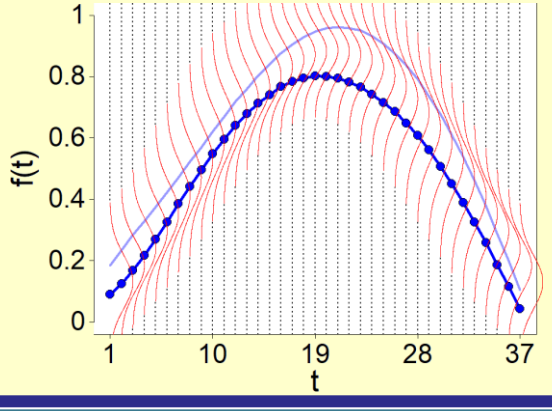
1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





Distribution Multivarié Gaussienne

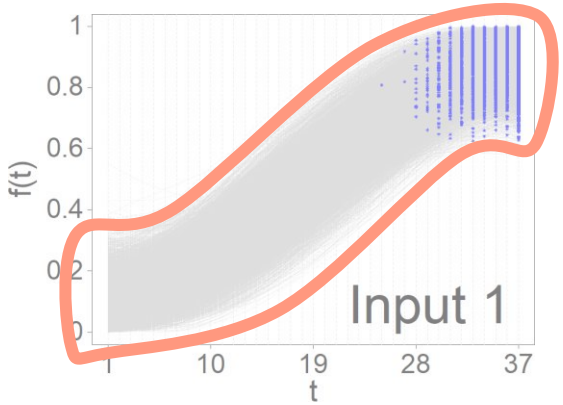


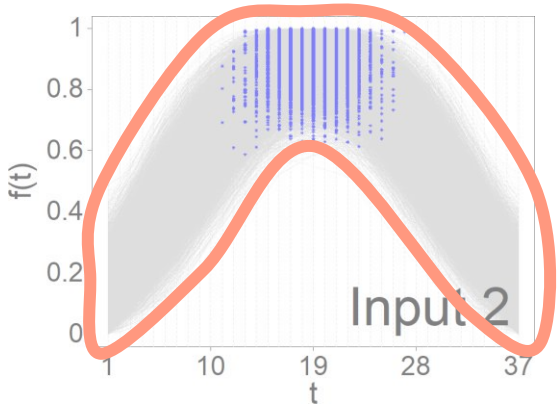


Plan d'expériences fonctionnel - **Procédure**

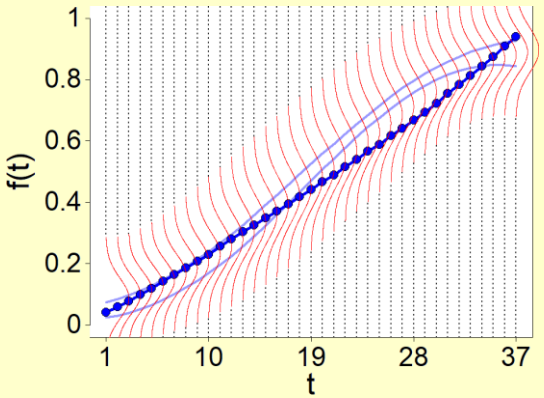
Obtention de points fonctionels extrêmes

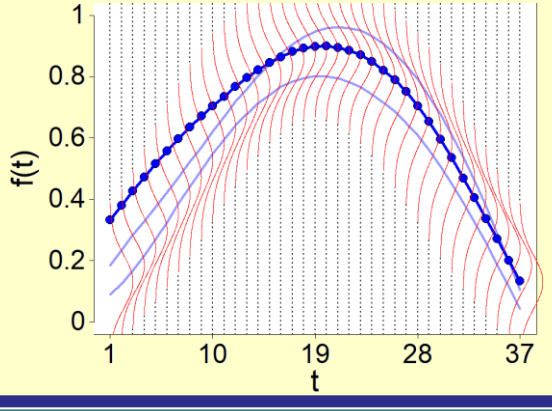
1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





Distribution Multivarié Gaussienne

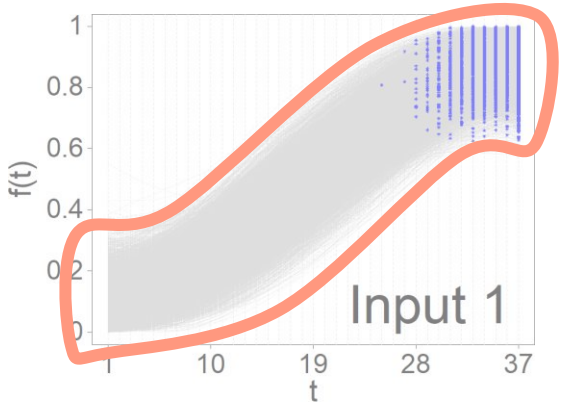


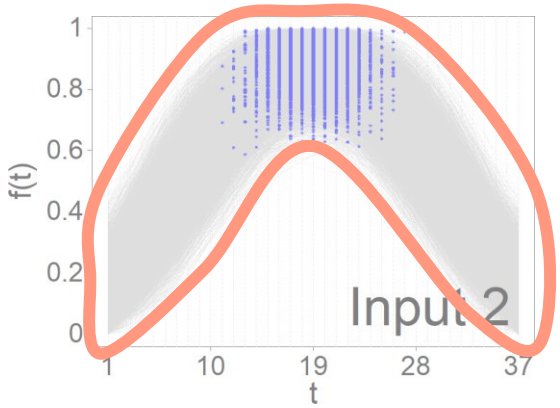


Plan d'expériences fonctionnel - **Procédure**

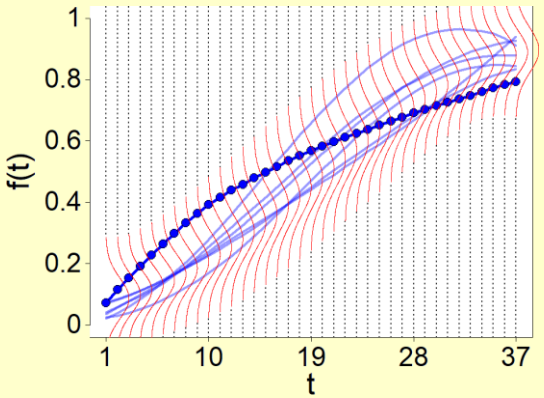
Obtention de points fonctionels extrêmes

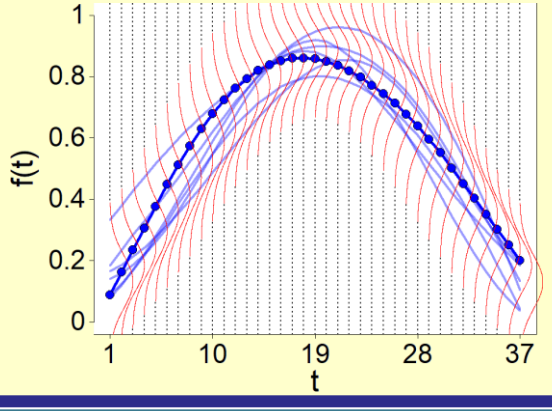
1. Apprendre le comportement de chaque entrée comme fonction
   - localisation temporelle de la représentation scalaire
   - forme typique des courbes





Distribution Multivarié Gaussienne





Plan d’expériences fonctionnel - **Procédure**

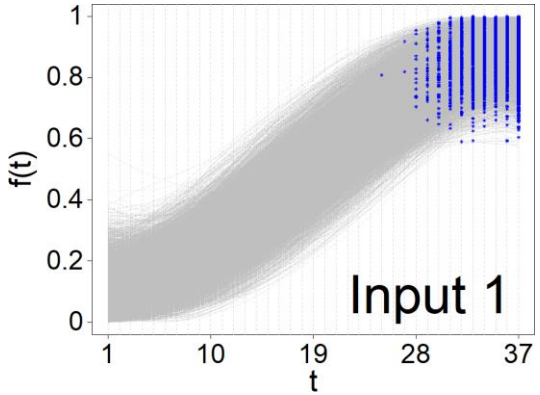
Obtention de points fonctionels extrêmes

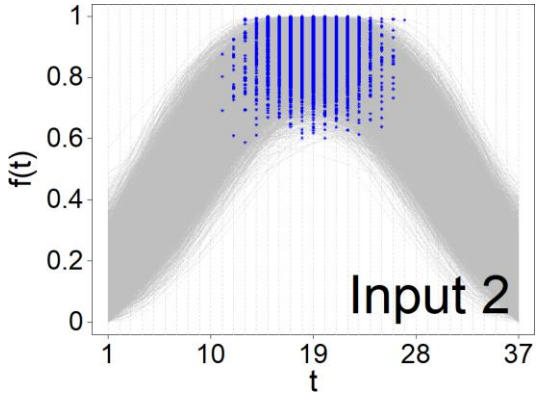
1. Apprendre le comportement de chaque entrée comme fonction
2. Localiser chaque valeur extrême scalaire dans un pas de temps
3. Échantillonner les 36 valeurs restants de chaque courbe

Plan d'expériences fonctionnel - **Procédure**

Obtention de points fonctionels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
2. Localiser chaque valeur extrême scalaire dans un pas de temps
3. Échantillonner les 36 valeurs restants de chaque courbe

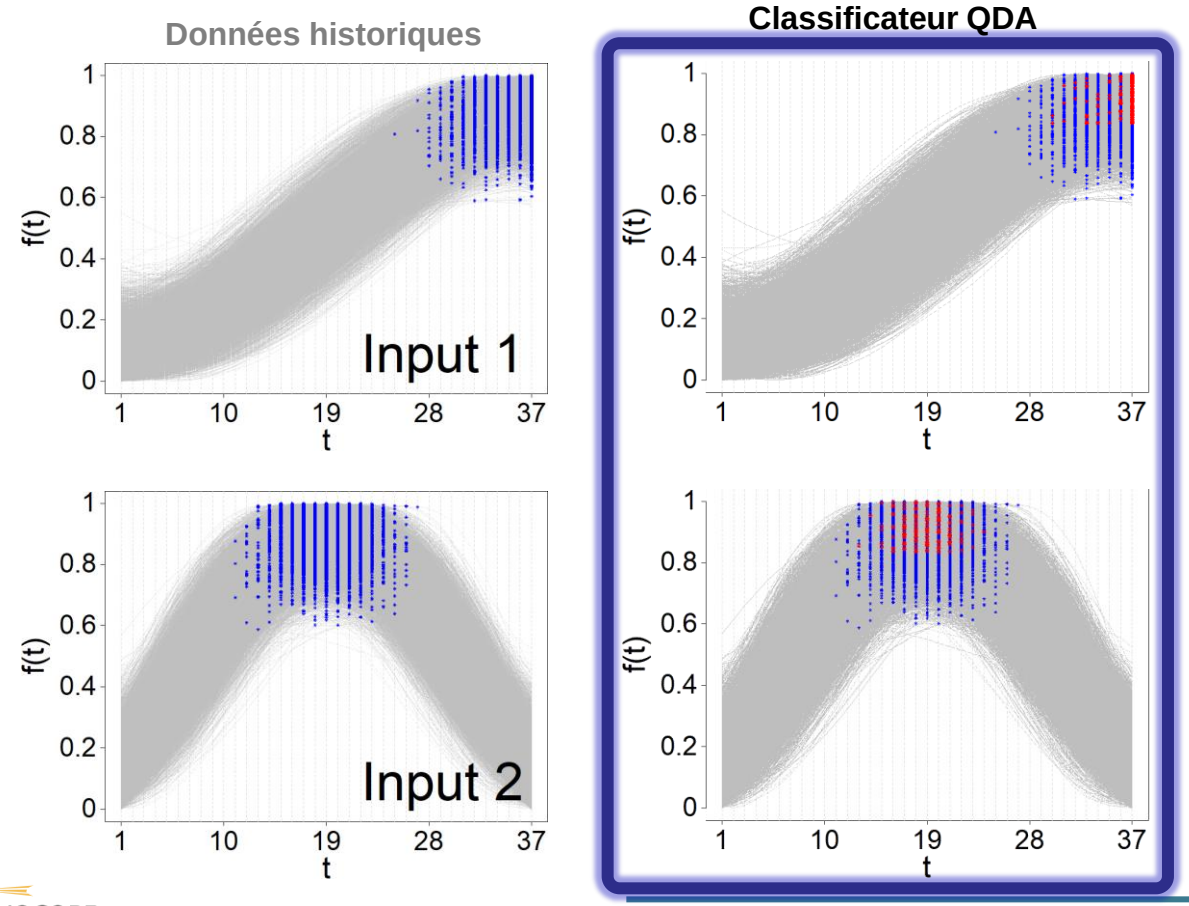




Plan d'expériences fonctionnel - **Procédure**

Obtention de points fonctionels extrêmes

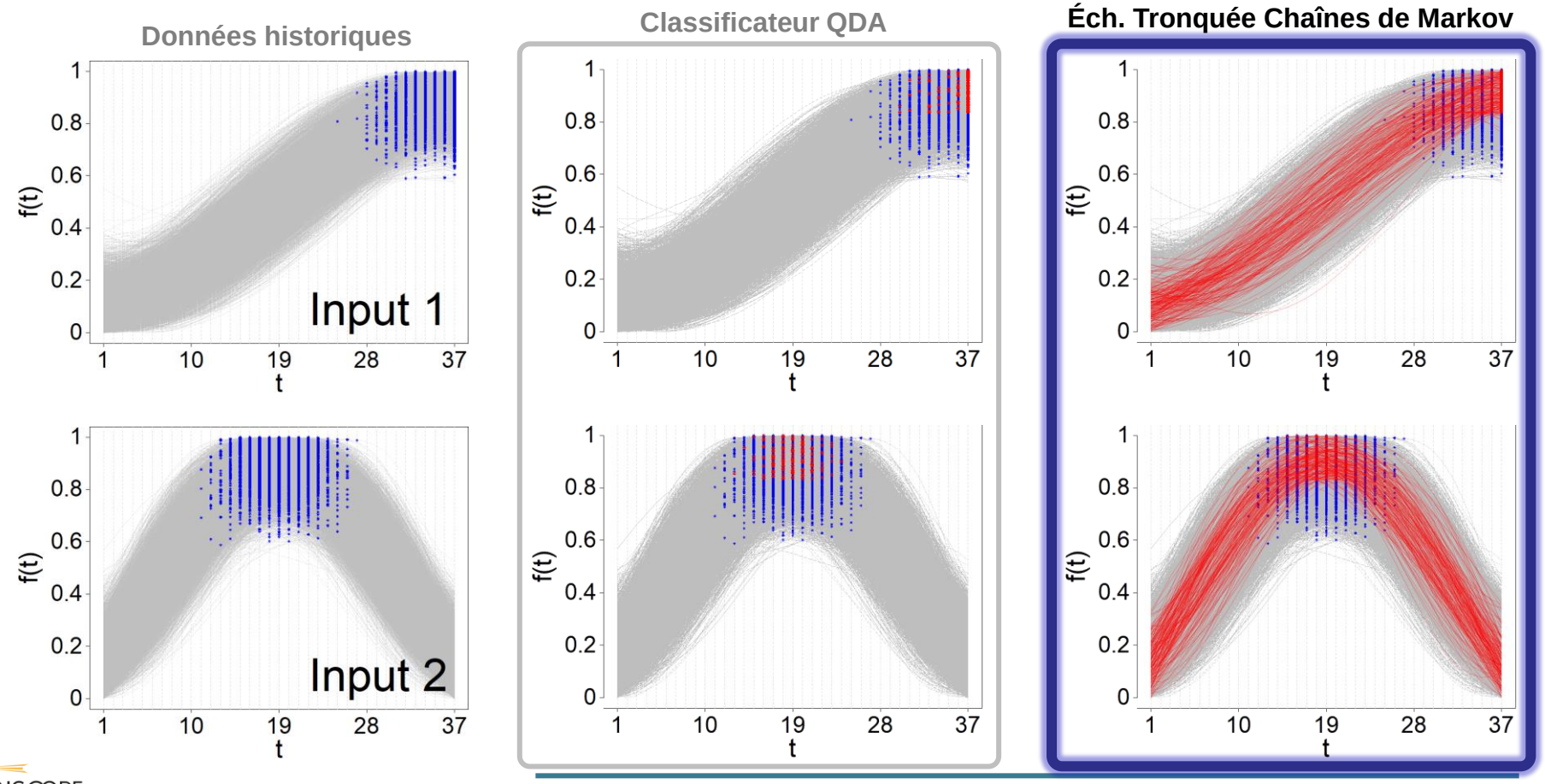
1. Apprendre le comportement de chaque entrée comme fonction
2. Localiser chaque valeur extrême scalaire dans un pas de temps
3. Échantillonner les 36 valeurs restants de chaque courbe



Plan d'expériences fonctionnel - **Procédure**

Obtention de points fonctionels extrêmes

1. Apprendre le comportement de chaque entrée comme fonction
2. Localiser chaque valeur extrême scalaire dans un pas de temps
3. Échantillonner les 36 valeurs restants de chaque courbe

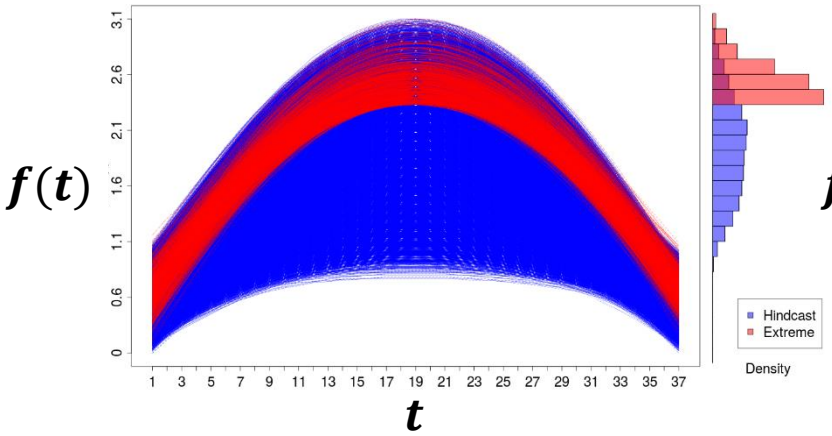


Package R tmg (Pakman, 2015)

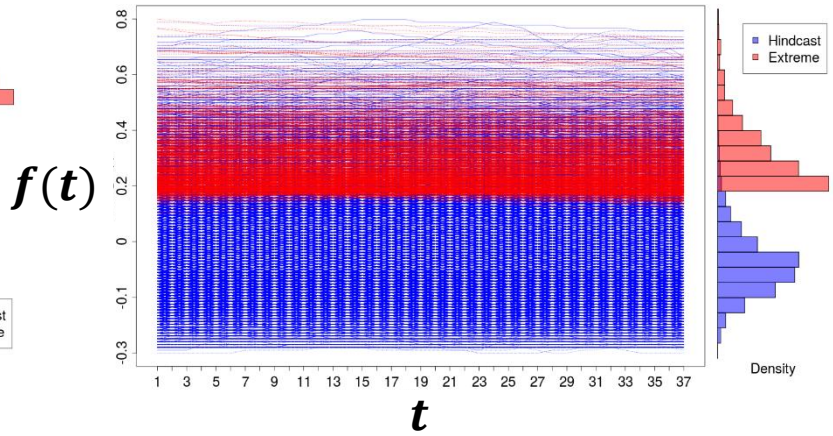
Résultats RISCOPE

Plan d'expériences fonctionnel – **Résultats RISCOPE**

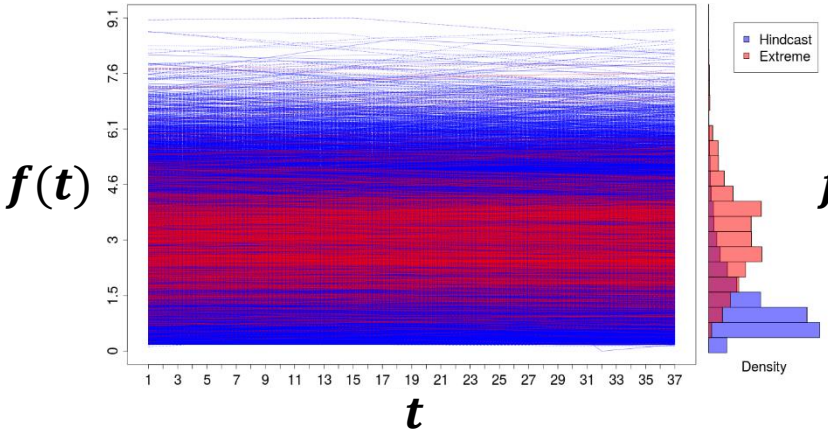
Marée



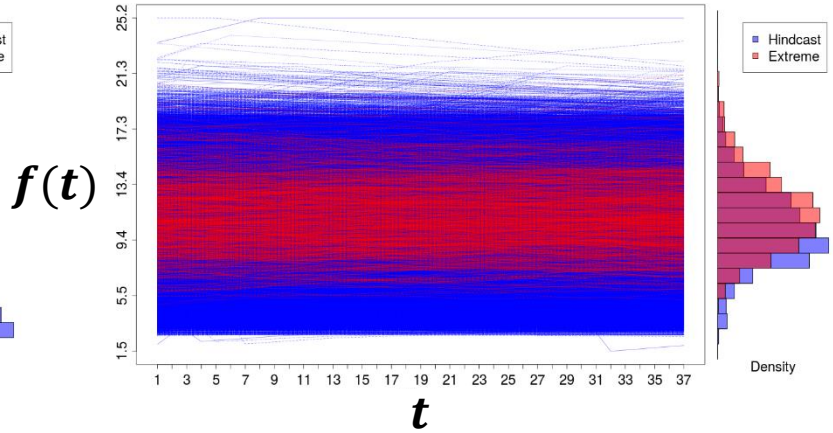
Surcote



Hauteur significative

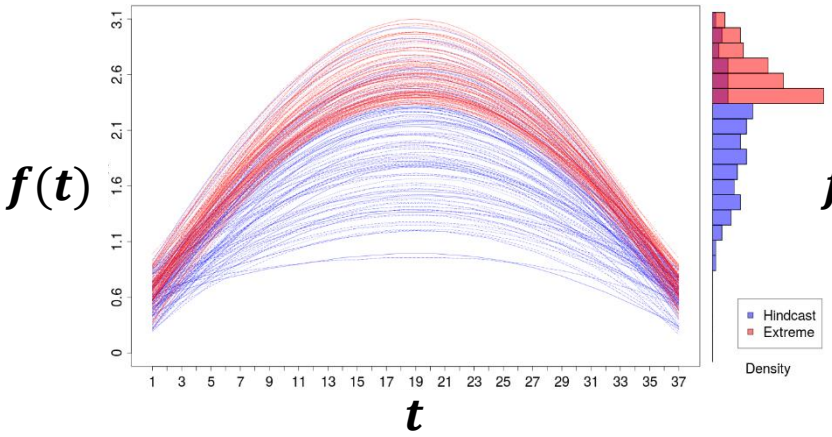


Période Pic

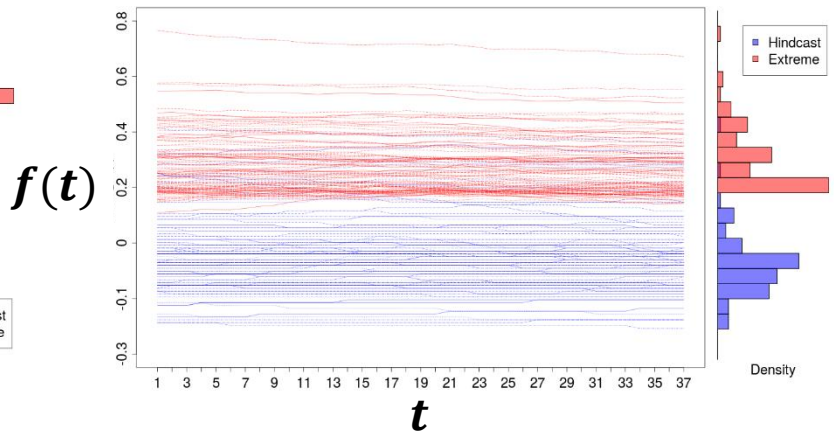


Plan d'expériences fonctionnel – **Résultats RISCOPE**

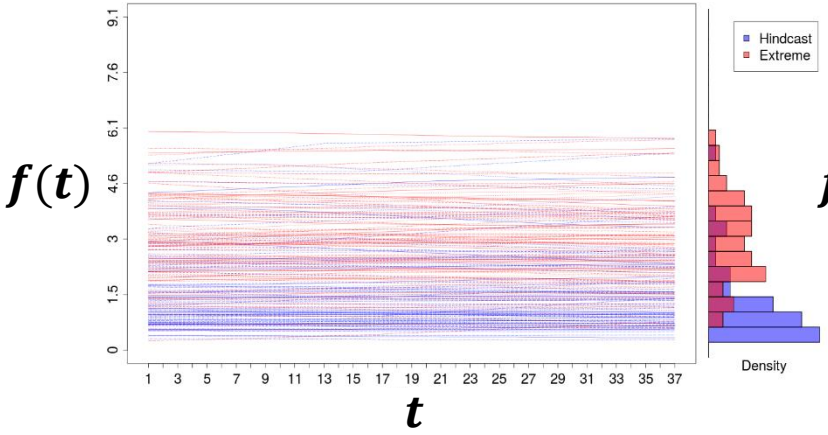
Marée



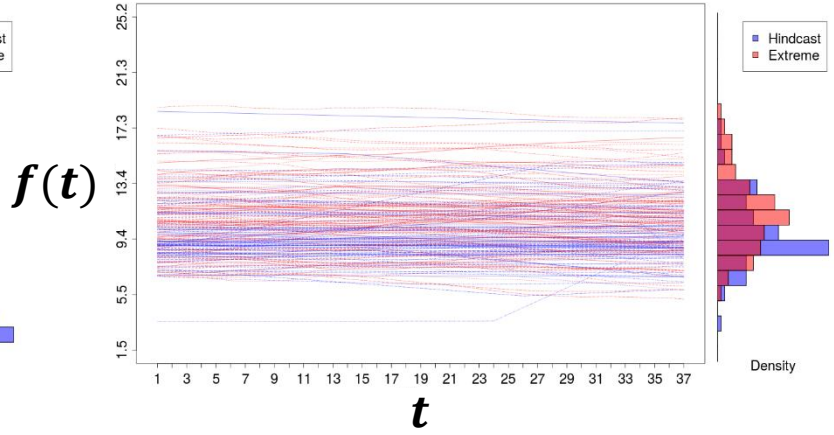
Surcote



Hauteur significative

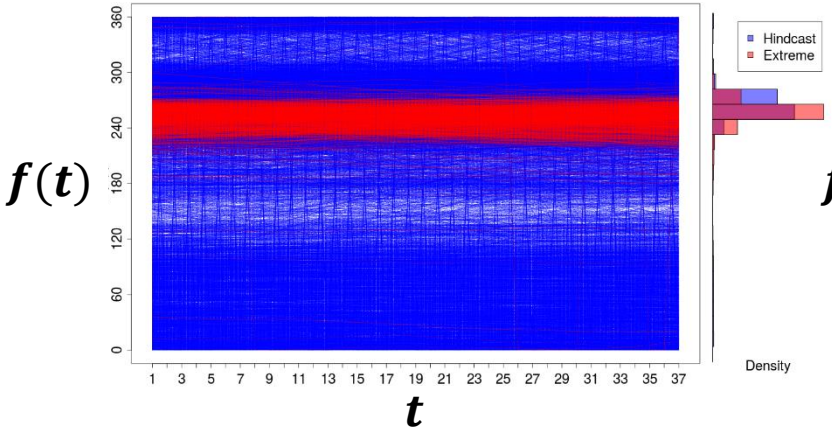


Période Pic

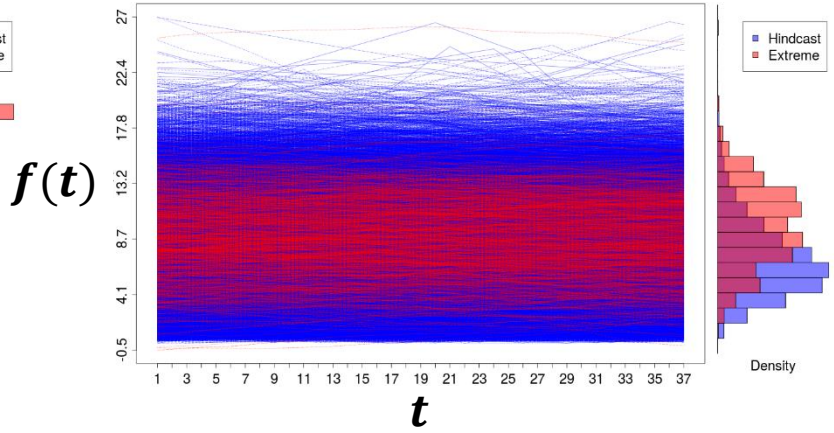


Plan d'expériences fonctionnel – **Résultats RISCOPE**

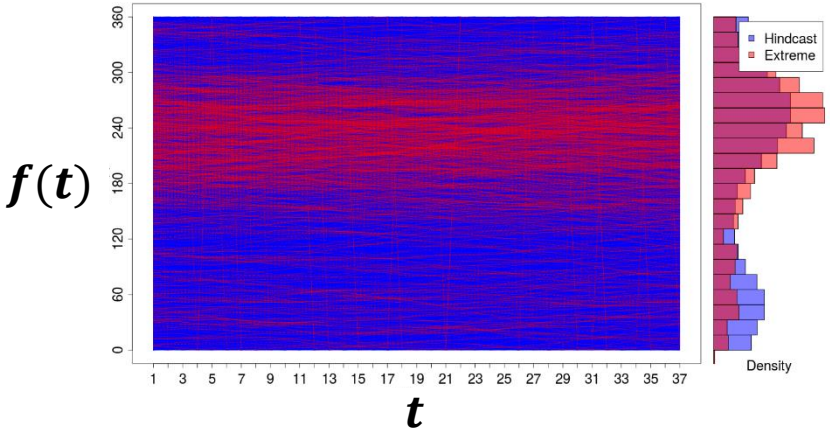
Direction pic



Vent

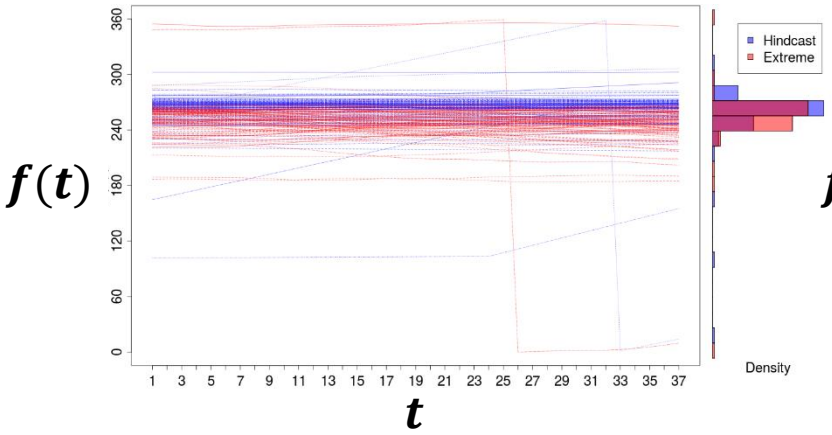


Direction du vent

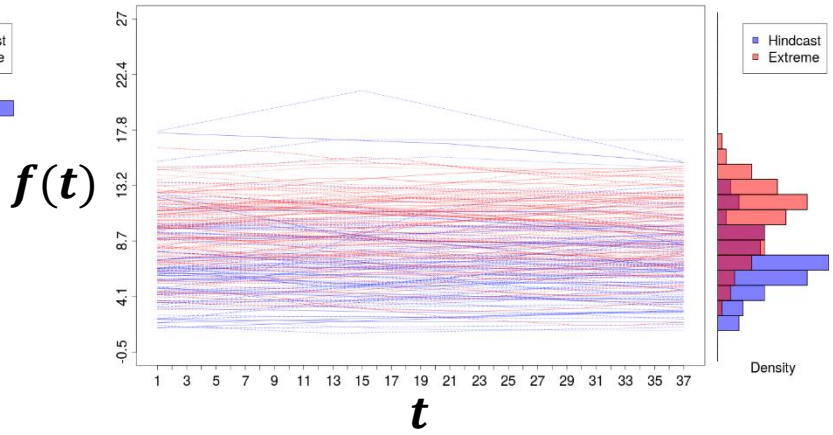


Plan d'expériences fonctionnel – **Résultats RISCOPE**

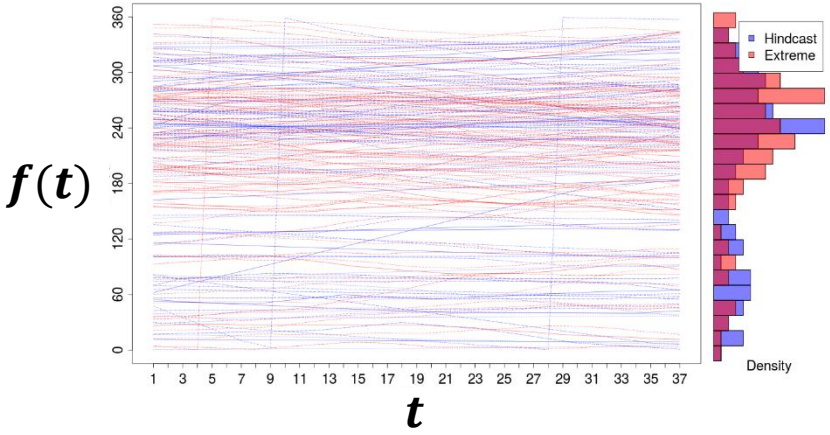
Direction pic



Vent



Direction du vent



Bilan

Bilan

Base d'apprentissage contient:

- **100 évènements synthétiques** en utilisant l'approche décrite
- **21 évènements historiques** (9 avec et 12 sans inondation; cf. [1])
- 16 scenarios avec **petites variations de 9 évènements** historiques.

Total: **174 scenarios.**

[1] Idier et al. (2020)