Exposé et témoignage
Post-doctorante

Journée Maths en Herbe

Margaux Zaffran

21 janvier 2026

Parcours jusqu'ici

- ENSTA Paris (2016-2020)
- M1 Mathématiques Appliquées (2017-2018)
- Année de césure : stages EDF R&D et GE Healthcare
- M2 Data Sciences

Parcours jusqu'ici

- ENSTA Paris (2016-2020)
- M1 Mathématiques Appliquées (2017-2018)
- Année de césure : stages EDF R&D et GE Healthcare
- M2 Data Sciences
- Thèse en statistiques au Centre de Mathématiques Appliquées de l'Ecole polytechnique, à l'Inria et en CIFRE avec EDF R&D (2024)

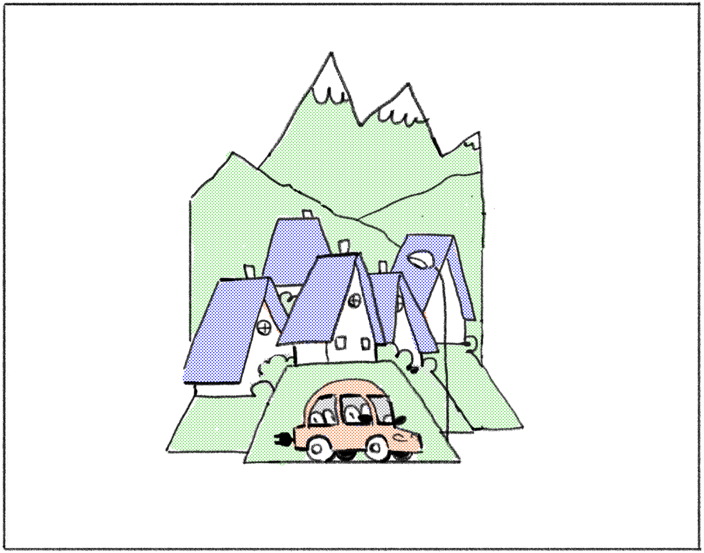
Parcours jusqu'ici

- ENSTA Paris (2016-2020)
- M1 Mathématiques Appliquées (2017-2018)
- Année de césure : stages EDF R&D et GE Healthcare
- M2 Data Sciences
- Thèse en statistiques au Centre de Mathématiques Appliquées de l'Ecole polytechnique, à l'Inria et en CIFRE avec EDF R&D (2024)
- Post-doctorante à UC Berkeley (2024-2025)

Parcours jusqu'ici

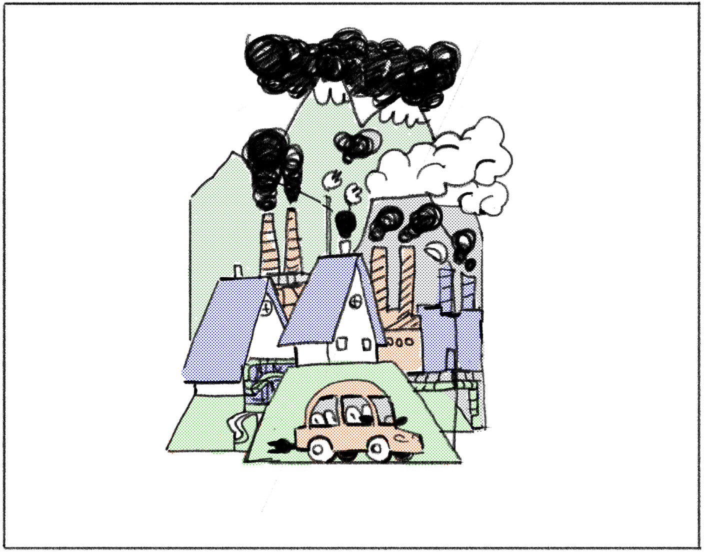
- ENSTA Paris (2016-2020)
- M1 Mathématiques Appliquées (2017-2018)
- Année de césure : stages EDF R&D et GE Healthcare
- M2 Data Sciences
- Thèse en statistiques au Centre de Mathématiques Appliquées de l'Ecole polytechnique, à l'Inria et en CIFRE avec EDF R&D (2024)
- Post-doctorante à UC Berkeley (2024-2025)
- Post-doctorante au Laboratoire de Mathématiques d'Orsay depuis cette année

Électrification, électrification renouvelable



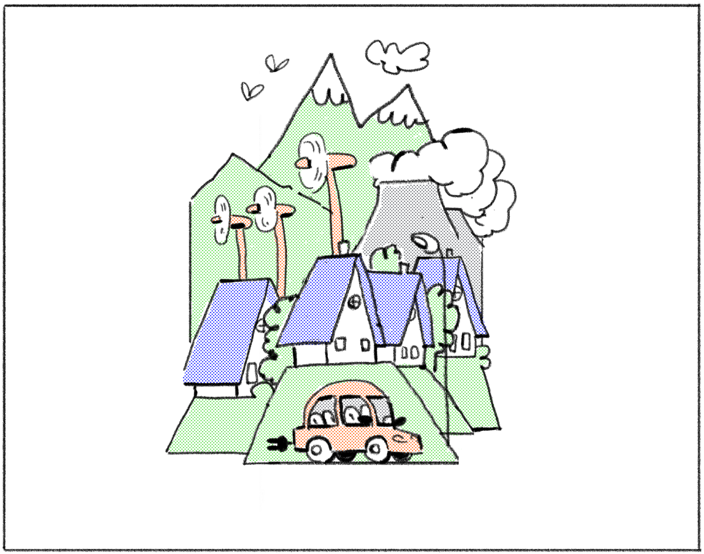
Illustrations @theoremlinger

Électrification, électrification renouvelable



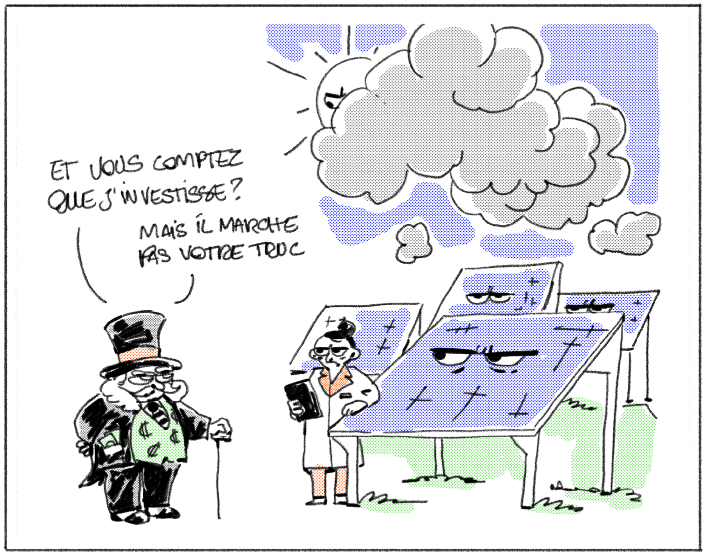
Illustrations @theoremlinger

Électrification, électrification renouvelable

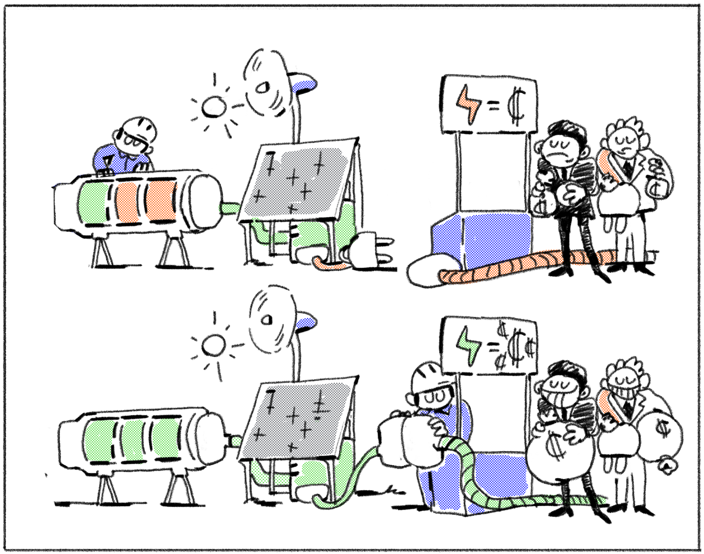


Illustrations @theoremlinger

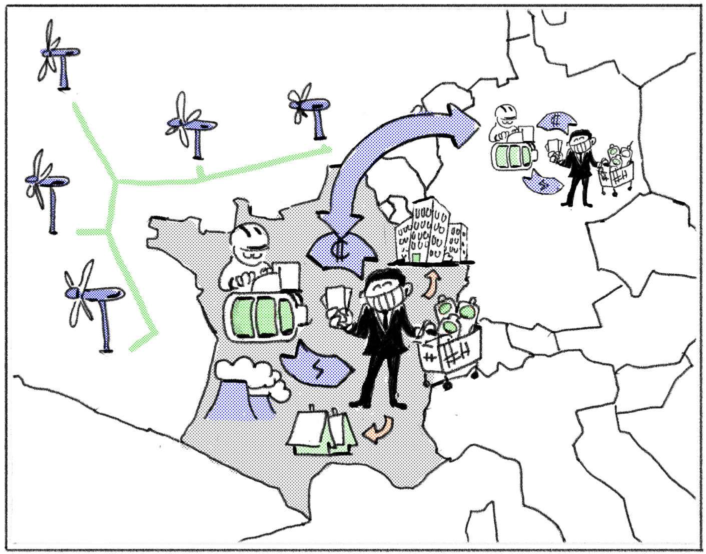
Une situation paradoxale



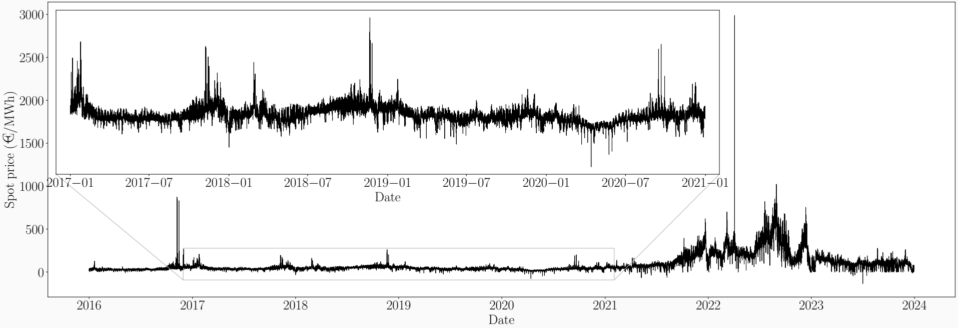
Vendre et stocker au bon moment !



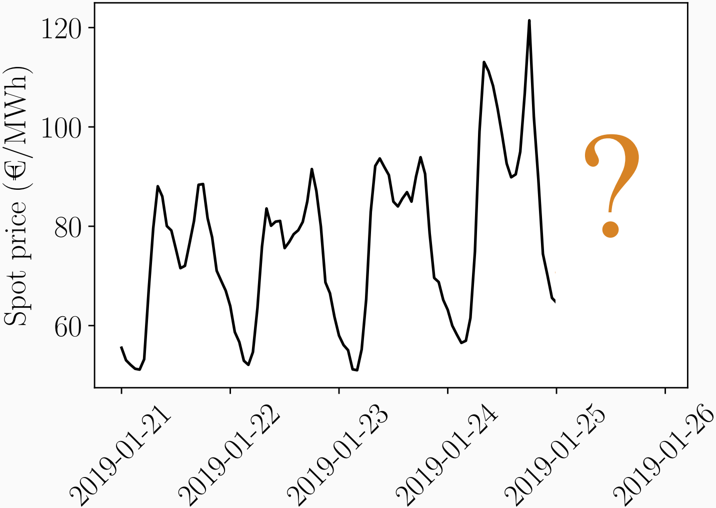
Mais au fait, à qui vendre ? Et qui vend ?



Prix spot français de l'électricité de 2016 à 2023



Objectif de prévision



Prévoir à partir de variables descriptives

Jour et heure	Prix	Prix J-1	Prix J-7	Conso.	Jour
11/01/16 0PM	21.95	15.58	13.78	58800	Lundi
11/01/16 1PM	20.04	19.05	13.44	57600	Lundi
⋮	⋮	⋮	⋮	⋮	⋮
12/01/16 0PM	21.51	21.95	25.03	61600	Mardi
12/01/16 1PM	19.81	20.04	24.42	59800	Mardi
⋮	⋮	⋮	⋮	⋮	⋮
18/01/16 0PM	38.14	37.86	21.95	70400	Lundi
18/01/16 1PM	35.66	34.60	20.04	69500	Lundi
⋮	⋮	⋮	⋮	⋮	⋮

Prévoir à partir de variables descriptives

Jour et heure	Prix	Prix J-1	Prix J-7	Conso.	Jour
11/01/16 0PM	**21.95**	15.58	13.78	58800	Lundi
11/01/16 1PM	**20.04**	19.05	13.44	57600	Lundi
⋮	⋮	⋮	⋮	⋮	⋮
12/01/16 0PM	**21.51**	21.95	25.03	61600	Mardi
12/01/16 1PM	**19.81**	20.04	24.42	59800	Mardi
⋮	⋮	⋮	⋮	⋮	⋮
18/01/16 0PM	**38.14**	37.86	21.95	70400	Lundi
18/01/16 1PM	**35.66**	34.60	20.04	69500	Lundi
⋮	⋮	⋮	⋮	⋮	⋮

- $y_t \in \mathbb{R}$

Prévoir à partir de variables descriptives

Jour et heure	Prix	Prix J-1	Prix J-7	Conso.	Jour
11/01/16 0PM	**21.95**	15.58	13.78	58800	Lundi
11/01/16 1PM	**20.04**	19.05	13.44	57600	Lundi
⋮	⋮	⋮	⋮	⋮	⋮
12/01/16 0PM	**21.51**	21.95	25.03	61600	Mardi
12/01/16 1PM	**19.81**	20.04	24.42	59800	Mardi
⋮	⋮	⋮	⋮	⋮	⋮
18/01/16 0PM	**38.14**	37.86	21.95	70400	Lundi
18/01/16 1PM	**35.66**	34.60	20.04	69500	Lundi
⋮	⋮	⋮	⋮	⋮	⋮

- $y_t \in \mathbb{R}$
- $x_t \in \mathbb{R}^d$

$$\rightarrow y_t = f(x_t) + \varepsilon_t$$

Prévoir à partir de variables descriptives

Jour et heure	Prix	Prix J-1	Prix J-7	Conso.	Jour
11/01/16 0PM	**21.95**	15.58	13.78	58800	Lundi
11/01/16 1PM	**20.04**	19.05	13.44	57600	Lundi
⋮	⋮	⋮	⋮	⋮	⋮
12/01/16 0PM	**21.51**	21.95	25.03	61600	Mardi
12/01/16 1PM	**19.81**	20.04	24.42	59800	Mardi
⋮	⋮	⋮	⋮	⋮	⋮
18/01/16 0PM	**38.14**	37.86	21.95	70400	Lundi
18/01/16 1PM	**35.66**	34.60	20.04	69500	Lundi
⋮	⋮	⋮	⋮	⋮	⋮

- $y_t \in \mathbb{R}$
- $x_t \in \mathbb{R}^d$

$$\rightarrow y_t = f(x_t) + \varepsilon_t$$

Ex. : $y_1 = 21.95$ et $x_1 = (01, 0PM, 15.58, 13.78, 58800, \text{Lundi})$

Utiliser le passé pour évaluer les modèles

	Jour et heure	Prix	Prix J-1	Prix J-7	Conso.	Jour
→	11/01/16 0PM	**21.95**	15.58	13.78	58800	Lundi
→	11/01/16 1PM	**20.04**	19.05	13.44	57600	Lundi
	⋮	⋮	⋮	⋮	⋮	⋮
→	12/01/16 0PM	**21.51**	21.95	25.03	61600	Mardi
→	12/01/16 1PM	**19.81**	20.04	24.42	59800	Mardi
	⋮	⋮	⋮	⋮	⋮	⋮
→	18/01/16 0PM	**38.14**	37.86	21.95	70400	Lundi
→	18/01/16 1PM	**35.66**	34.60	20.04	69500	Lundi
	⋮	⋮	⋮	⋮	⋮	⋮

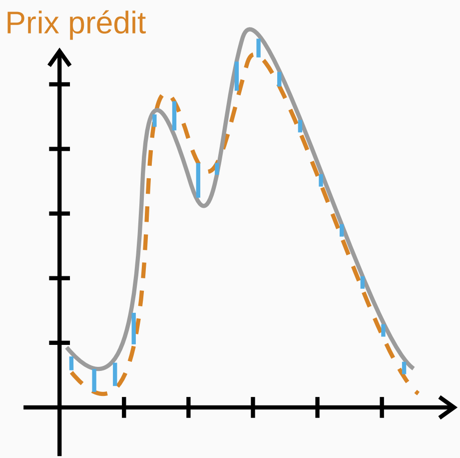
Utiliser le passé pour évaluer les modèles

	Jour et heure	Prix	Prix J-1	Prix J-7	Conso.	Jour
$\rightarrow$	11/01/16 0PM	?	15.58	13.78	58800	Lundi
$\rightarrow$	11/01/16 1PM	?	19.05	13.44	57600	Lundi
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\rightarrow$	12/01/16 0PM	?	21.95	25.03	61600	Mardi
$\rightarrow$	12/01/16 1PM	?	20.04	24.42	59800	Mardi
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\rightarrow$	18/01/16 0PM	?	37.86	21.95	70400	Lundi
$\rightarrow$	18/01/16 1PM	?	34.60	20.04	69500	Lundi
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

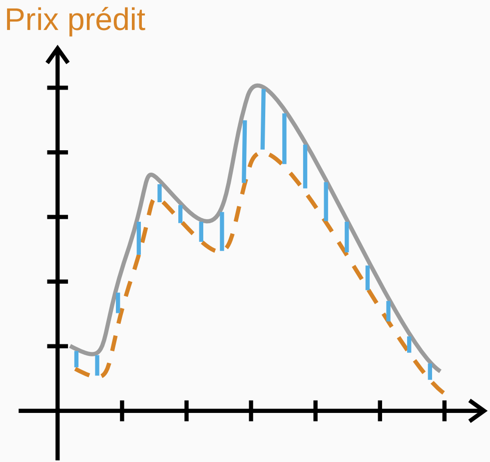
Utiliser le passé pour évaluer les modèles

	Jour et heure	Prix	Prix J-1	Prix J-7	Conso.	Jour
$\rightarrow$	11/01/16 0PM	**21.95**	15.58	13.78	58800	Lundi
$\rightarrow$	11/01/16 1PM	**20.04**	19.05	13.44	57600	Lundi
	⋮	⋮	⋮	⋮	⋮	⋮
$\rightarrow$	12/01/16 0PM	**21.51**	21.95	25.03	61600	Mardi
$\rightarrow$	12/01/16 1PM	**19.81**	20.04	24.42	59800	Mardi
	⋮	⋮	⋮	⋮	⋮	⋮
$\rightarrow$	18/01/16 0PM	**38.14**	37.86	21.95	70400	Lundi
$\rightarrow$	18/01/16 1PM	**35.66**	34.60	20.04	69500	Lundi
	⋮	⋮	⋮	⋮	⋮	⋮

Évaluer les prédictions d'un modèle



...



erreur = | prix réel − prix prédit |

Évaluer les fonctions f

Pour chaque modèle ou fonction f, on calcule l'**erreur moyenne empirique** :

$$\text{erreur empirique } (f) = \frac{1}{N} \sum_{i=1}^{N} \text{erreur de } f \text{ sur l'exemple i}.$$

Évaluer les fonctions f

Pour chaque modèle ou fonction f, on calcule l'**erreur moyenne empirique** :

$$\text{erreur empirique }(f) = \frac{1}{N}\sum_{i=1}^{N} \text{erreur de } f \text{ sur l'exemple i}.$$

En pratique, on restreint notre recherche à $f \in \mathcal{F}$.

Évaluer les fonctions f

Pour chaque modèle ou fonction f, on calcule l'**erreur moyenne empirique** :

$$\text{erreur empirique }(f) = \frac{1}{N}\sum_{i=1}^{N} \text{erreur de } f \text{ sur l'exemple i}.$$

En pratique, on restreint notre recherche à $f \in \mathcal{F}$.

Modèle linéaire

$$\mathcal{F} := \{ f \text{ telles que } f(x) = \gamma + \beta^T x, \\ \text{avec } (\gamma, \beta) \in \mathbb{R} \times \mathbb{R}^d \}$$

Trouver la meilleure fonction f parmi un groupe de candidates

Modèle linéaire

$$\mathcal{F} := \Big\{ f \text{ telles que } f(x) = \gamma + \beta^T x, \\ \text{avec } (\gamma, \beta) \in \mathbb{R} \times \mathbb{R}^d \Big\}$$

On cherche donc les **meilleures valeurs de** β **et** γ pour que f donne des prédictions qui ressemblent au mieux à nos valeurs à prédire.

Trouver la meilleure fonction f parmi un groupe de candidates

Modèle linéaire

$$\mathcal{F} := \Big\{ f \text{ telles que } f(x) = \gamma + \beta^T x, \text{ avec } (\gamma, \beta) \in \mathbb{R} \times \mathbb{R}^d \Big\}$$

On cherche donc les **meilleures valeurs de β et γ** pour que f donne des prédictions qui ressemblent au mieux à nos valeurs à prédire.

$\Rightarrow$ On cherche les valeurs de β et γ telles que f donne des prédictions de **plus petite erreur moyenne empirique possible**.

Trouver la meilleure fonction f parmi un groupe de candidates

Modèle linéaire

$$\mathcal{F} := \Big\{ f \text{ telles que } f(x) = \gamma + \beta^T x, \\ \text{avec } (\gamma, \beta) \in \mathbb{R} \times \mathbb{R}^d \Big\}$$

On cherche donc les **meilleures valeurs de β et γ** pour que f donne des prédictions qui ressemblent au mieux à nos valeurs à prédire.

$\Rightarrow$ On cherche les valeurs de β et γ telles que f donne des prédictions de **plus petite erreur moyenne empirique possible**.

$$\boxed{f^* = \underset{f \in \mathcal{F}}{\operatorname{argmin}} \{\text{erreur empirique } (f)\}}$$

Trouver la meilleure fonction f parmi un groupe de candidates

Modèle linéaire

$$\mathcal{F} := \Big\{ f \text{ telles que } f(x) = \gamma + \beta^T x, \\ \text{avec } (\gamma, \beta) \in \mathbb{R} \times \mathbb{R}^d \Big\}$$

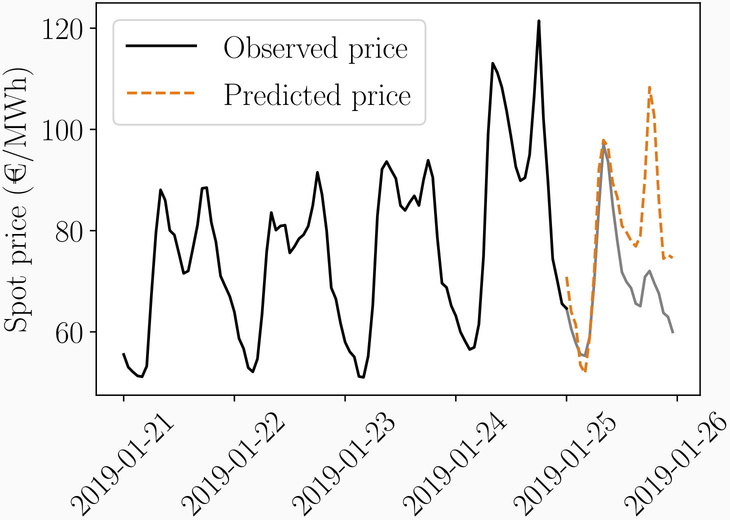
On cherche donc les **meilleures valeurs de** β **et** γ pour que f donne des prédictions qui ressemblent au mieux à nos valeurs à prédire.

$\Rightarrow$ On cherche les valeurs de β et γ telles que f donne des prédictions de **plus petite erreur moyenne empirique possible**.

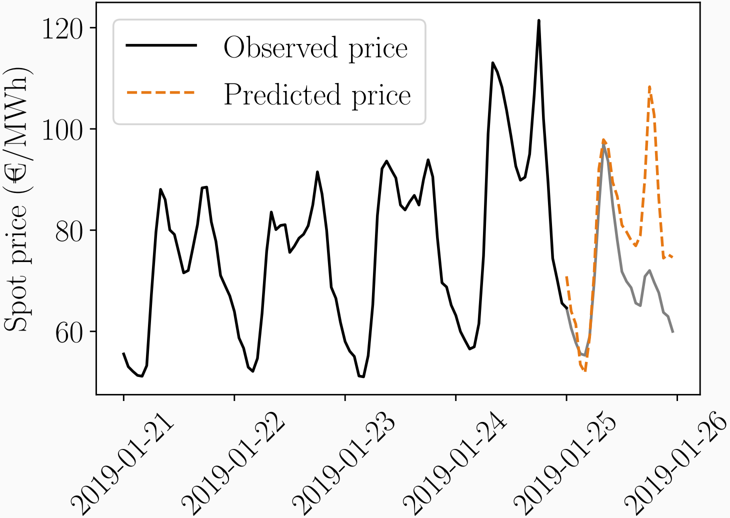
$$\boxed{f^* = \underset{f \in \mathcal{F}}{\operatorname{argmin}} \{\text{erreur empirique } (f)\}}$$

$\hookrightarrow$ Réseau de neurones $\longrightarrow$ $\mathcal{F}$ plus riche

Qu'est-ce que cela donne sur les prix de l'électricité ?



Qu'est-ce que cela donne sur les prix de l'électricité ?

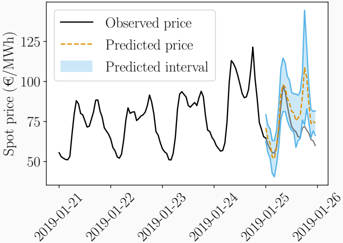


↪ Les meilleurs modèles de prévision de prix font 10% d'erreur relative

Et la fiabilité dans tout ça ?

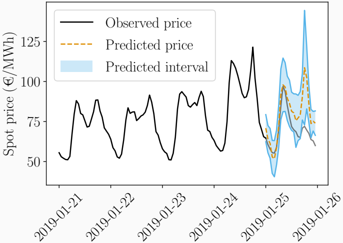
Prévoir avec confiance : région prédictive

Nouvel objectif :



Prévoir avec confiance : région prédictive

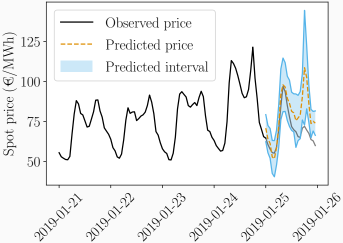
Nouvel objectif :



Quantifier l'incertitude prédictive avec :

Prévoir avec confiance : région prédictive

Nouvel objectif :

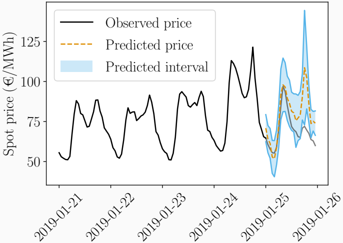


Quantifier l'incertitude prédictive avec :

- **Des méthodes théoriquement valides**

Prévoir avec confiance : région prédictive

Nouvel objectif :

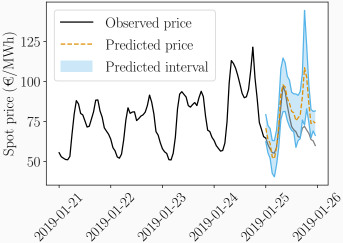


Quantifier l'incertitude prédictive avec :

- Des méthodes théoriquement valides
- **Peu d'hypothèses sur la distribution des données**

Prévoir avec confiance : région prédictive

Nouvel objectif :

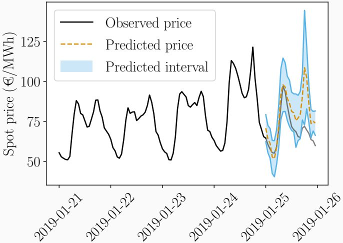


Quantifier l'incertitude prédictive avec :

- Des méthodes théoriquement valides
- Peu d'hypothèses sur la distribution des données
- **Des garanties agnostiques au modèle de prédiction**
  $\rightsquigarrow$ Approche post-hoc (i.e. s'encapsuler en bout de chaîne autour de n'importe quel algorithme opérationnel pré-existant)

Prévoir avec confiance : région prédictive

Nouvel objectif :



Quantifier l'incertitude prédictive avec :

- Des méthodes théoriquement valides
- Peu d'hypothèses sur la distribution des données
- **Des garanties agnostiques au modèle de prédiction**

  $\rightsquigarrow$ Approche post-hoc (i.e. s'encapsuler en bout de chaîne autour de n'importe quel algorithme opérationnel pré-existant)

Formalisation de la quantification d'incertitude prédictive

L'utilisateur choisit un **niveau de confiance** $1-\alpha$.

Par exemple 0.9 (autrement dit 90%) correspond à $\alpha = 0.1$.

Formalisation de la quantification d'incertitude prédictive

L'utilisateur choisit un **niveau de confiance** $1 - \alpha$.
Par exemple 0.9 (autrement dit 90%) correspond à $\alpha = 0.1$.

Notons $\mathcal{C}_\alpha(x)$ une région prédictive de niveau $1 - \alpha$ en un point x.

Formalisation de la quantification d'incertitude prédictive

L'utilisateur choisit un **niveau de confiance** $1 - \alpha$.
Par exemple 0.9 (autrement dit 90%) correspond à $\alpha = 0.1$.

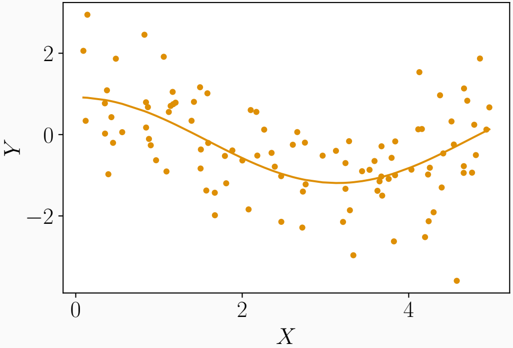
Notons $\mathcal{C}_\alpha(x)$ une région prédictive de niveau $1 - \alpha$ en un point x.

Celle-ci doit vérifier la propriété suivante :

$$\boxed{\mathbb{P}\left\{Y \in \mathcal{C}_\alpha\left(X\right)\right\} \geq 1 - \alpha,}$$

tout en étant la plus petite possible.

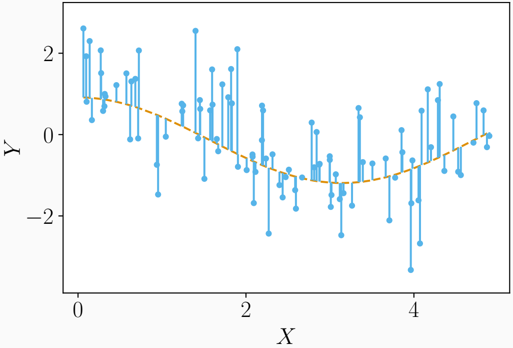
Prédictions conformes par partition[1,2,3] : phase d'entraînement



- Apprendre f

[1] ?, *Algorithmic Learning in a Random World*
[2] ?, *Inductive Confidence Machines for Regression*, ECML
[3] ?, *Distribution-Free Predictive Inference for Regression*, JRSS B

Prédictions conformes par partition[1,2,3] : phase de calibration



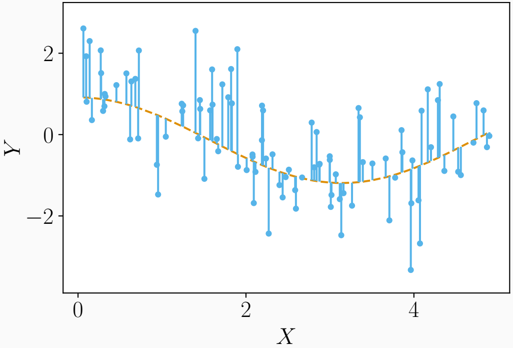
- Prédire avec f

[1] ?, *Algorithmic Learning in a Random World*

[2] ?, *Inductive Confidence Machines for Regression*, ECML

[3] ?, *Distribution-Free Predictive Inference for Regression*, JRSS B

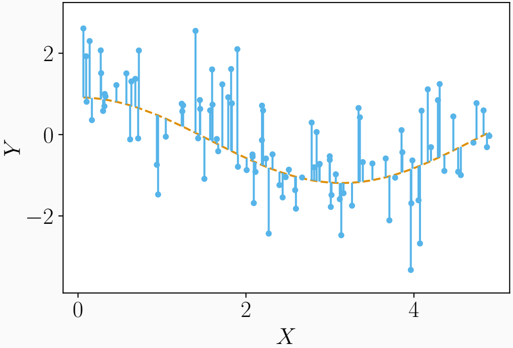
Prédictions conformes par partition[1,2,3] : phase de calibration



- Prédire avec f
- Obtenir les erreurs

[1] ?, *Algorithmic Learning in a Random World*
[2] ?, *Inductive Confidence Machines for Regression*, ECML
[3] ?, *Distribution-Free Predictive Inference for Regression*, JRSS B

Prédictions conformes par partition[1,2,3] : phase de calibration



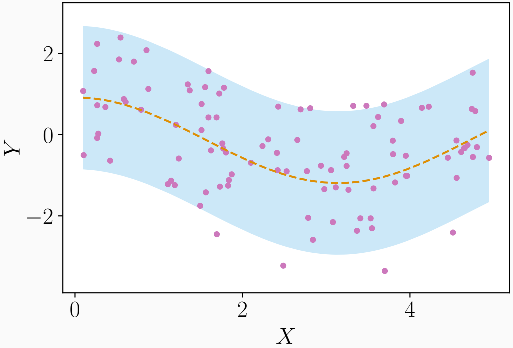
- Prédire avec f
- Obtenir les erreurs
- Trouver le quantile de niveau 0.9 de ces erreurs (la $0.9 \times n$-ème plus grande erreur), notée $q_{0.9}$ (erreurs)

[1] **?**, *Algorithmic Learning in a Random World*

[2] **?**, *Inductive Confidence Machines for Regression*, ECML

[3] **?**, *Distribution-Free Predictive Inference for Regression*, JRSS B

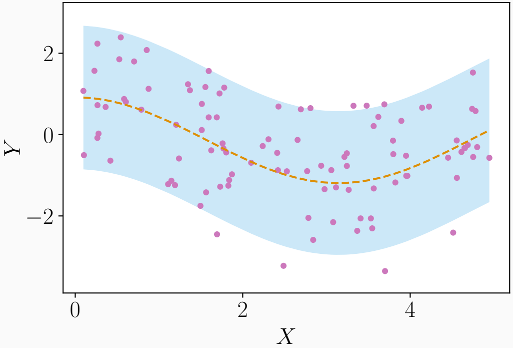
Prédictions conformes par partition[1,2,3] : phase de prédiction



- Prédire avec *f*

[1] **?**, *Algorithmic Learning in a Random World*
[2] **?**, *Inductive Confidence Machines for Regression*, ECML
[3] **?**, *Distribution-Free Predictive Inference for Regression*, JRSS B

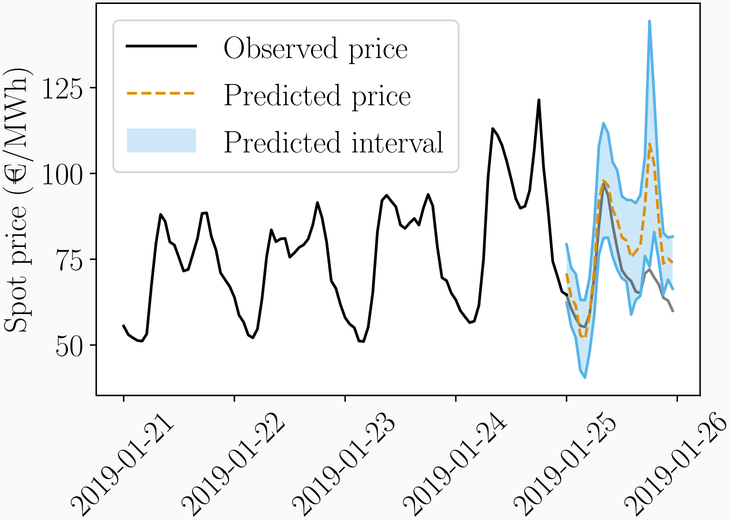
Prédictions conformes par partition[1,2,3] : phase de prédiction



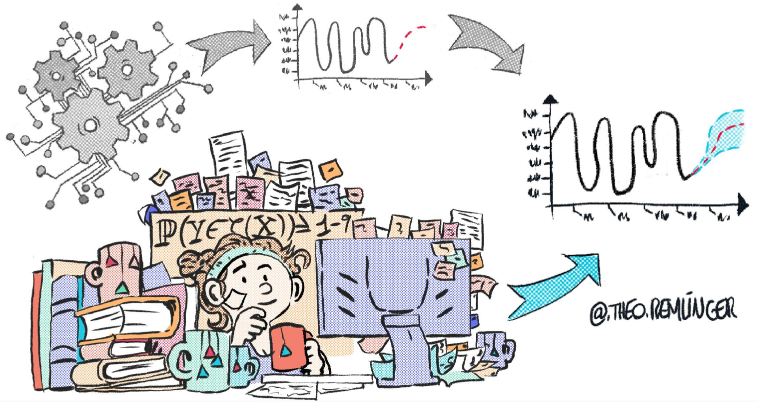
- Prédire avec f
- Construire $\mathcal{C}(x)$:

$$[f(x) - q_{0.9}(\text{erreurs});\; f(x) + q_{0.9}(\text{erreurs})]$$

[1] **?**, *Algorithmic Learning in a Random World*
[2] **?**, *Inductive Confidence Machines for Regression*, ECML
[3] **?**, *Distribution-Free Predictive Inference for Regression*, JRSS B


Observed price
Predicted price
Predicted interval
Spot price (€/MWh)
125
100
75
50
2019-01-21
2019-01-22
2019-01-23
2019-01-24
2019-01-25
2019-01-26

Résumons



Des idées (biaisées) d'approfondissement

- Phénomènes temporels (cycles, tendances, dépendence aux jours d'avant ...) mettent en défaut l'hypothèse des prédictions conformes.

Des idées (biaisées) d'approfondissement

- Phénomènes temporels (cycles, tendances, dépendence aux jours d'avant ...) mettent en défaut l'hypothèse des prédictions conformes. **Extension de la méthode aux séries temporelles**.

Des idées (biaisées) d'approfondissement

- Phénomènes temporels (cycles, tendances, dépendence aux jours d'avant ...) mettent en défaut l'hypothèse des prédictions conformes. **Extension de la méthode aux séries temporelles**.
- La garantie n'est qu'en moyenne. La méthode fait plus d'erreurs sur les week-ends que la semaine par exemple.

Des idées (biaisées) d'approfondissement

- Phénomènes temporels (cycles, tendances, dépendence aux jours d'avant ...) mettent en défaut l'hypothèse des prédictions conformes. **Extension de la méthode aux séries temporelles**.
- La garantie n'est qu'en moyenne. La méthode fait plus d'erreurs sur les week-ends que la semaine par exemple.

$$\mathbb{P}\left(Y \in \mathcal{C}_{1-\alpha}(X) | X = x\right) \geq 1 - \alpha$$

Des idées (biaisées) d'approfondissement

- Phénomènes temporels (cycles, tendances, dépendence aux jours d'avant ...) mettent en défaut l'hypothèse des prédictions conformes. **Extension de la méthode aux séries temporelles**.
- La garantie n'est qu'en moyenne. La méthode fait plus d'erreurs sur les week-ends que la semaine par exemple.

$$\cancel{\mathbb{P}\left(Y \in \mathcal{C}_{1-\alpha}(X) | X = x\right) \geq 1 - \alpha}$$

Des idées (biaisées) d'approfondissement

- Phénomènes temporels (cycles, tendances, dépendence aux jours d'avant ...) mettent en défaut l'hypothèse des prédictions conformes. **Extension de la méthode aux séries temporelles**.
- La garantie n'est qu'en moyenne. La méthode fait plus d'erreurs sur les week-ends que la semaine par exemple.

$$\cancel{\mathbb{P}\left(Y \in \mathcal{C}_{1-\alpha}(X) | X = x\right) \geq 1 - \alpha}$$

Obtention de garanties plus conditionnelles.

Des idées (biaisées) d'approfondissement

- Phénomènes temporels (cycles, tendances, dépendence aux jours d'avant ...) mettent en défaut l'hypothèse des prédictions conformes. **Extension de la méthode aux séries temporelles**.
- La garantie n'est qu'en moyenne. La méthode fait plus d'erreurs sur les week-ends que la semaine par exemple.

$$\cancel{\mathbb{P}\left(Y \in \mathcal{C}_{1-\alpha}(X) | X = x\right) \geq 1 - \alpha}$$

  Obtention de garanties plus conditionnelles.
- Que se passe-t-il en présence de données manquantes ? Le modèle de prévision de consommation n'a pas pu prédire (panne de courant par ex.) mais nous voulons quand même prédire le prix.

Des idées (biaisées) d'approfondissement

- Phénomènes temporels (cycles, tendances, dépendence aux jours d'avant ...) mettent en défaut l'hypothèse des prédictions conformes. **Extension de la méthode aux séries temporelles**.
- La garantie n'est qu'en moyenne. La méthode fait plus d'erreurs sur les week-ends que la semaine par exemple.

$$\cancel{\mathbb{P}\left(Y \in \mathcal{C}_{1-\alpha}(X) | X = x\right) \geq 1 - \alpha}$$

Obtention de garanties plus conditionnelles.

- Que se passe-t-il en présence de données manquantes ? Le modèle de prévision de consommation n'a pas pu prédire (panne de courant par ex.) mais nous voulons quand même prédire le prix. **Quantification d'incertitudes en présence de données manquantes**.

Messages à retenir

- Les mathématiques peuvent permettre de faire des prédictions

Messages à retenir

- Les mathématiques peuvent permettre de faire des prédictions
- Applicable pour de nombreux problèmes réels

Messages à retenir

- Les mathématiques peuvent permettre de faire des prédictions
- Applicable pour de nombreux problèmes réels
- Ne jamais interpréter des prédictions ponctuelles

Messages à retenir

- Les mathématiques peuvent permettre de faire des prédictions
- Applicable pour de nombreux problèmes réels
- Ne jamais interpréter des prédictions ponctuelles
- Offrir des garanties aux citoyens

Messages à retenir

- Les mathématiques peuvent permettre de faire des prédictions
- Applicable pour de nombreux problèmes réels
- Ne jamais interpréter des prédictions ponctuelles
- Offrir des garanties aux citoyens
- Transversalité et universalité des mathématiques

References i