

Capsule 4 - Comparer des prévisions de référence

MQT-2101 - Analyse et modélisation des données

Pourquoi ça compte

Une méthode élaborée fait-elle mieux qu'une règle simple?

Un modèle peut très bien reproduire le passé et prévoir moins bien que la dernière valeur connue.

Objectifs de la capsule

- Séparer apprentissage et validation.
- Calculer les deux références naïves.
- Distinguer ajustement et prévision.

Bien ajuster n'est pas bien prévoir

APPRENTISSAGE

paramètres et structure

VALIDATION

erreurs sur le futur simulé



La séparation du cas guidé

60

mois d'apprentissage : 2020 à 2024

12

mois de validation : 2025

T

origine : décembre 2024

$T = 12 / 2024.$

Le naïf à origine fixe

dropeau: prédiction
dernière valeur dispo

$$\hat{y}_{T+h|T} = y_T, \quad h \geq 1.$$

$T+h \rightarrow h$ pas de temps après T

Une origine en décembre 2024. La dernière consommation mensuelle connue est répétée pour tous les mois de 2025.

prédiction au temps $T+h =$ valeur au temps T

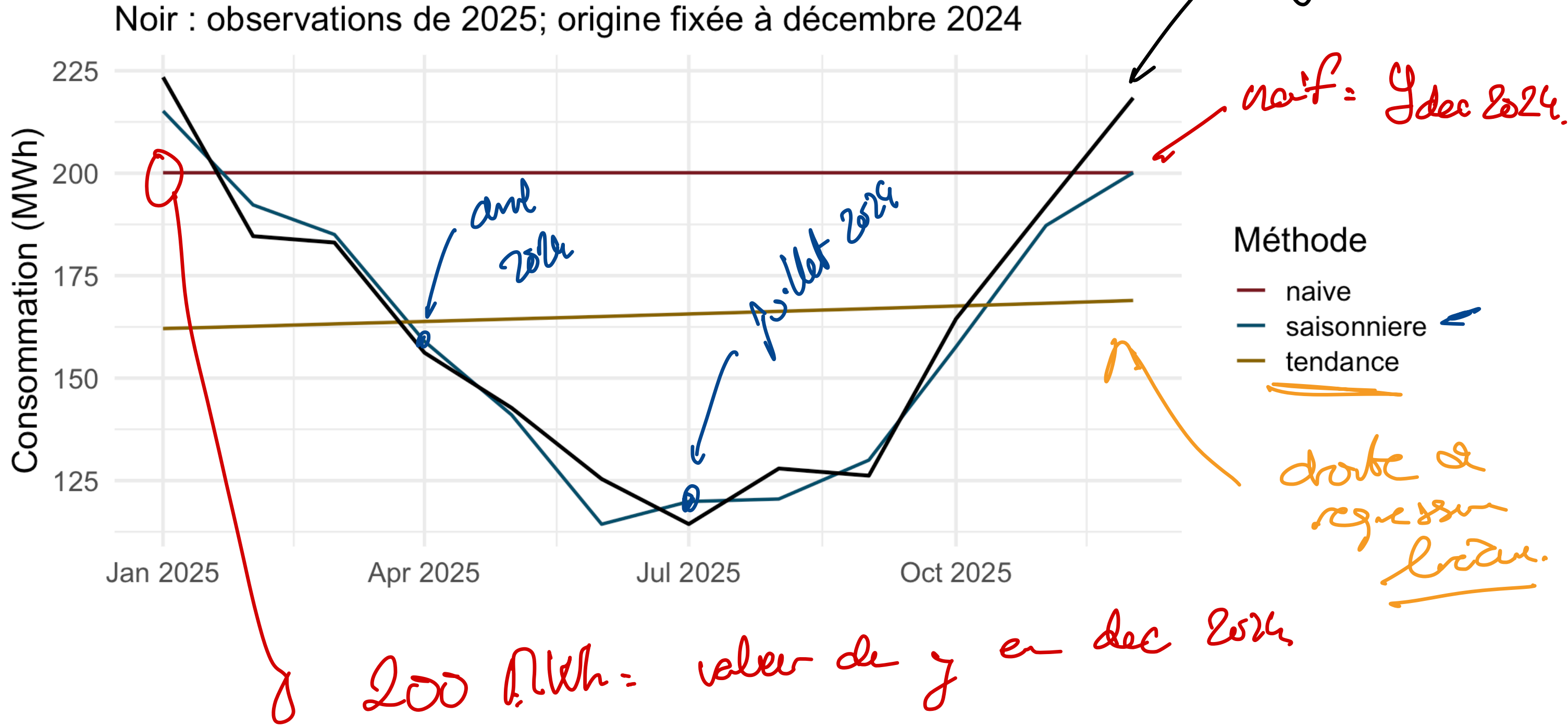
Le naïf saisonnier

$$\hat{y}_{T+h|T} = y_{T+h-m}, \quad 1 \leq h \leq m.$$

$m = 12$: reprendre les douze mois de la dernière année.

prédiction au temps $T+h$ = valeur de y au temps.
 $T+h-m$,
réviser un an en arrière
prédire Mars 2025 $\rightarrow$ juillet Mars 2024

Comparer sur les mêmes mois



Du code à la référence

```
1 refs <- prevoir_references(apprentissage$consommation_kwh, horizon = 12, periode = 12)
2 validation |> mutate(naive = refs$naive,
3                     saisonniere = refs$saisonniere)
```



Choisir la période selon l'horizon

UN MOIS

répéter plusieurs origines à $h = 1$

$T = \text{jan } 2025$
 $T = \text{jan } 2026$

UN AN

évaluer aussi les horizons jusqu'à 12

Réajuster avant la vraie prévision

1

Évaluer sur 2025

valider

2

Fixer le choix

valider
modèle

3

Réajuster sur 2020
à 2025

modèle choisi
news avec nouvel
 apprentissage : 2020 à 2025

4

Prévoir 2026

prévision

À retenir

Même origine

même information disponible

Référence

naïve et saisonnière

Séparation

comparer avant de réajuster

Production autonome 6.4

Produisez les douze prévisions naïves et saisonnières à la fin de 2024. Superposez-les à la consommation de 2025. Expliquez pourquoi une colonne `lag(consumption_kwh)` sur tout le fichier répond à un autre scénario.

Auto-vérification

- Mes deux références partagent l'origine de décembre 2024, le calendrier et les douze horizons.
- Je n'ai utilisé aucune valeur de 2025 pour calculer les prévisions à origine fixe.
- J'ai expliqué pourquoi un décalage sur tout le fichier correspond à un autre scénario.