

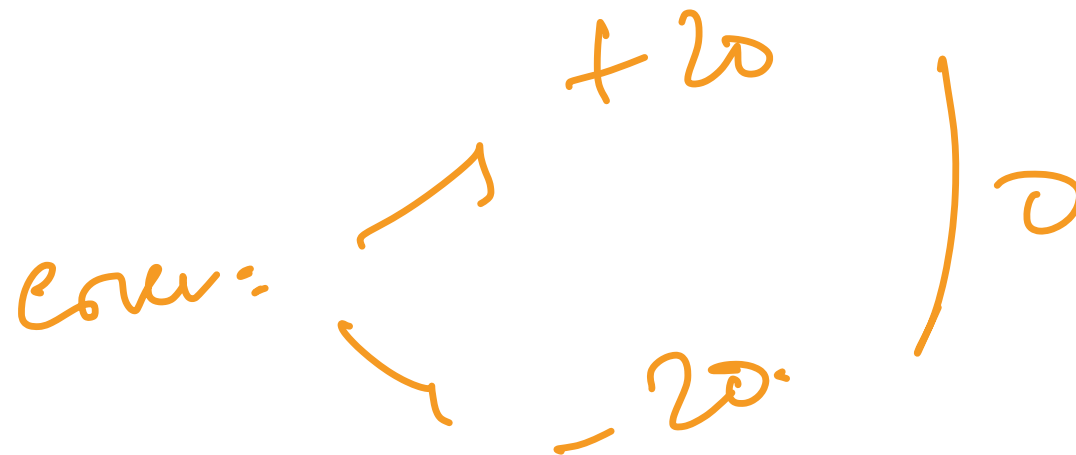
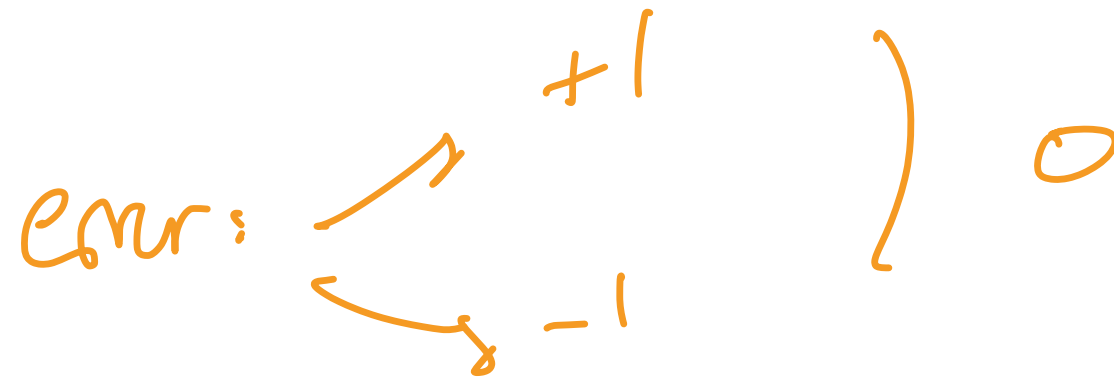
Capsule 5 - Mesurer et lire les erreurs

MQT-2101 - Analyse et modélisation des données

Pourquoi ça compte

Une erreur moyenne proche de **zéro** suffit-elle?

Les sur-prévisions et les sous-prévisions peuvent s'annuler. Leur ampleur et leur coût restent présents.



Objectifs de la capsule

- Calculer les cinq mesures d'erreur.
- Interpréter le signe et les unités.
- Examiner les erreurs au-delà d'un score.

Une convention de signe

residu

$$\underline{e_{T+h}} = \underbrace{y_{T+h}}_{\text{vrai}} - \underbrace{\hat{y}_{T+h|T}}_{\text{prédict}}$$

$E > 0$

sous-prévision

$E < 0$

sur-prévision

ME, MAE et RMSE

$$\text{ME} = \frac{1}{n} \sum e_i,$$

$$\text{MAE} = \frac{1}{n} \sum |e_i|.$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum e_i^2}.$$

absolue.
valeur absolue de e_i

Même unité que la cible. MAE est aussi appelée MAD.

Les erreurs peuvent s'annuler

-20, +20

deux erreurs, en unités

ME = 0

aucun biais moyen ici

MAE = 20

erreur absolue persistante

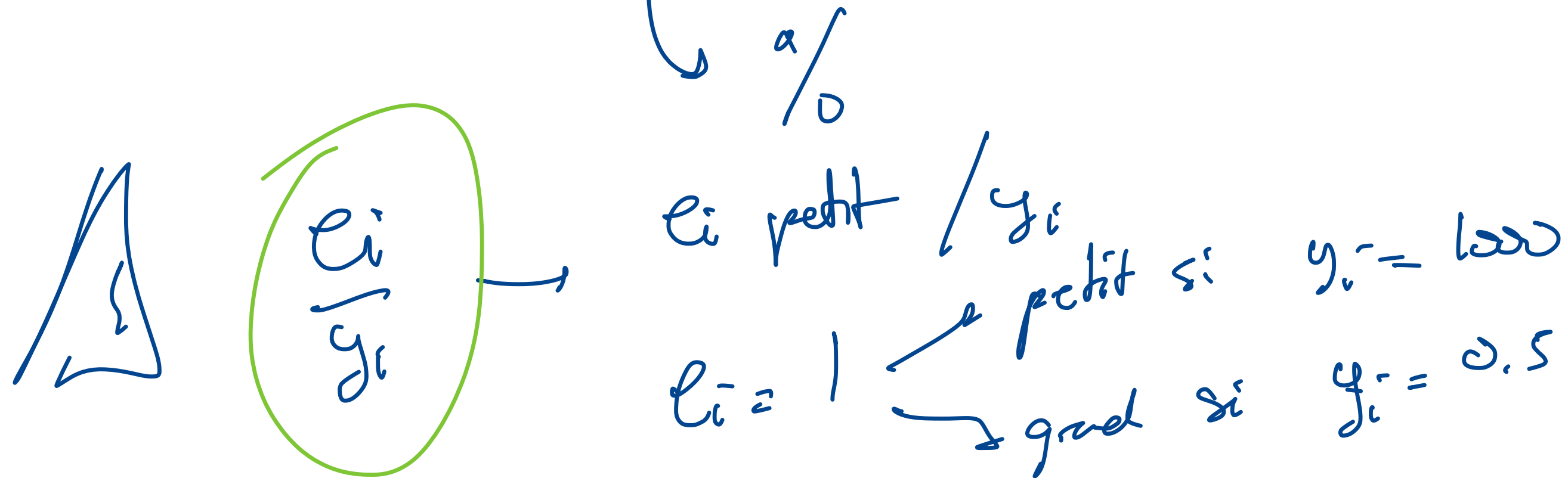
RMSE = 20 également dans cet exemple.

Le MAPE et ses limites

$$\text{MAPE} = \frac{100}{n} \sum \left| \frac{e_i}{y_i} \right|$$

Annotations:
- $\frac{100}{n}$: moyenné
- $\left| \frac{e_i}{y_i} \right|$: valeur absolue
- $\frac{e_i}{y_i}$: propre erreur sur vraie valeur

Zéro : indéfini. Valeur proche de zéro : instabilité. Zéro arbitraire : pourcentage inadapté.



La MASE : une échelle d'apprentissage

$$q_m = \frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|, \quad \text{MASE} = \frac{\text{MAE}_{\text{validation}}}{q_m}.$$

Handwritten notes:
- Above the sum: *erreur naïf*
- A red arrow points from the q_m in the denominator to the $m=1$ in the text below.
- A blue arrow points from the $\text{MAE}_{\text{validation}}$ to the handwritten note on the right.

*MAE validation
erreur validation*

$m = 1$: référence naïve. $m = 12$: référence saisonnière mensuelle.

Que signifie $MASE < 1$?

INTERPRÉTATION

MAE plus petite que q_m

LIMITE

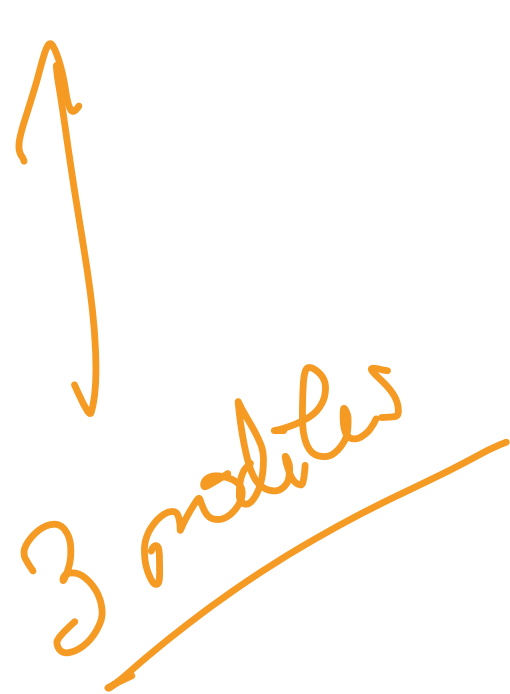
pas une victoire garantie sur la validation

$$MASE < 1 \rightarrow MAE_{val} < q_m$$

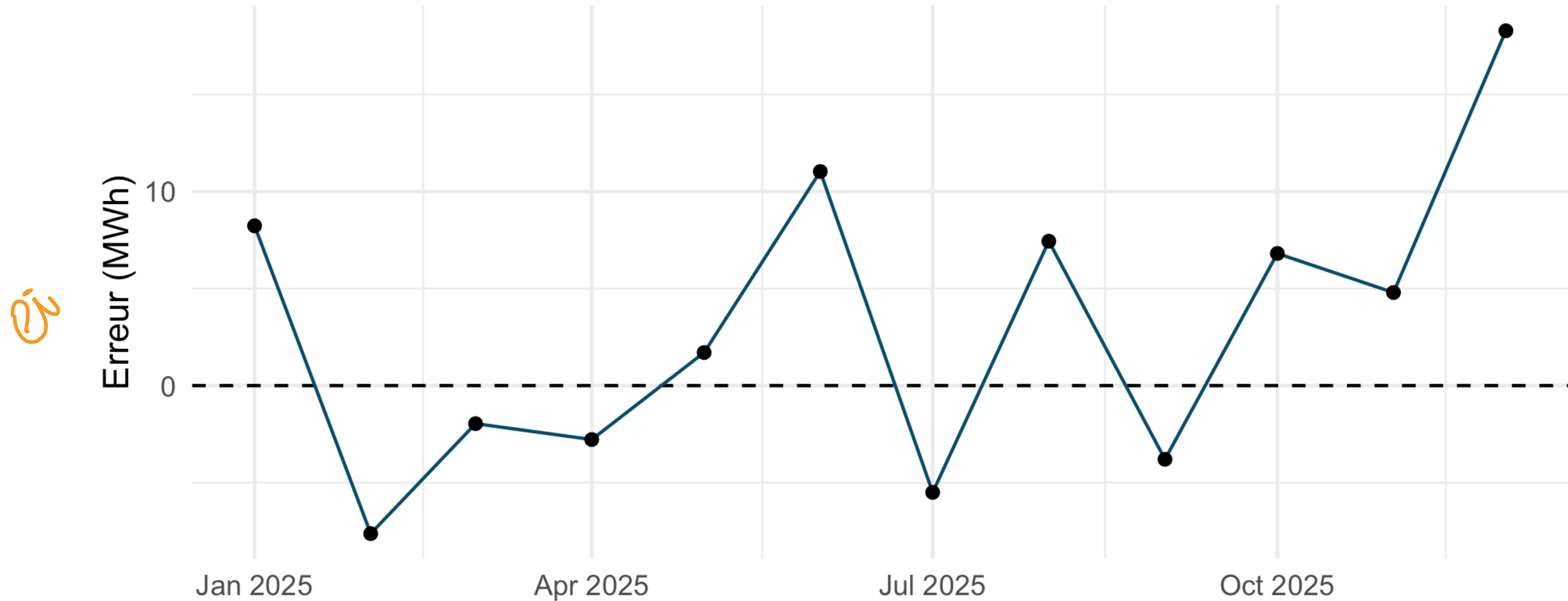
Le résultat du cas guidé

Méthode	ME (kWh)	MAE (kWh)	RMSE (kWh)	MAPE (%)	MASE
naive	-36 878,67	43 810,17	51 298,07	31,69	6,71
saisonniere	3 052,50	6 662,33	7 987,13	4,11	1,02
tendance	-2 172,18	31 573,64	35 965,65	20,50	4,84

3 modules



Lire le temps et la distribution



À retenir

Signe

sous-prévision ou sur-prévision

Ampleur

MAE, RMSE, MAPE, MASE

Diagnostic

temps et distribution

Production autonome 6.5

Calculez ME, MAE, RMSE, MAPE et MASE des trois candidats. Interprétez les unités et le signe de ME. Comparez directement leurs MAE puis lisez le graphique et l'histogramme des erreurs.

Auto-vérification

- J'ai interprété les unités des mesures et le signe de ME.
- L'échelle de ma MASE utilise uniquement les données d'apprentissage.
- J'ai signalé les limites du MAPE en présence de zéros sans les supprimer.
- J'ai distingué les erreurs de validation des résidus d'apprentissage.