

Prévision de la ressource solaire intermittente

Cyril Voyant, professeur — Mines Paris – PSL, Centre O.I.E., Sophia Antipolis

UE14 — Cours ingénieur civil des Mines

Projet de recherche appliquée : modéliser et prévoir le rayonnement solaire global

MINES PARIS – PSL

CENTRE O.I.E.

SOPHIA ANTIPOLIS



Plan du cours

01

Pourquoi prévoir le solaire ?

Enjeux pour le réseau, unités, données d'Ajaccio

02

Le rayonnement solaire

Direct, diffus, atmosphère, position du Soleil

03

Séries temporelles et prévision

Composantes, horizon, premier modèle AR(1)

04

Ciel clair et indice de ciel clair

Modèle Solis, CSI, stationnarité

05

Modèle AR et moindres carrés

Construire X et y, moindres carrés, choix de p

06

Évaluation, ELM et apprentissage en ligne

Persistence, métriques, ELM, fenêtre glissante, frugalité

Annexes : stationnarité au sens large, espérance conditionnelle, critères AIC/BIC. Glossaire et références en fin de document.

Objectifs du projet

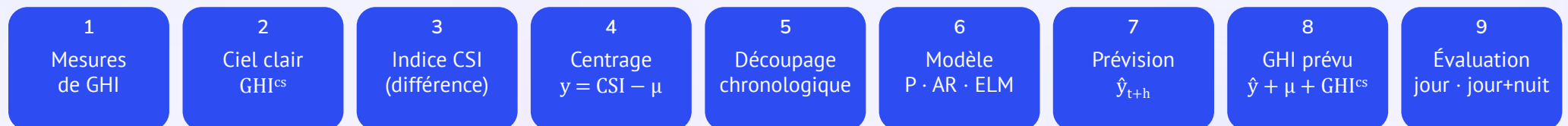
Contexte

- Le rayonnement solaire est **intermittent** (nuages).
- Le réseau électrique doit anticiper la production photovoltaïque.
- Besoin : des prévisions **fiables à court terme** (ici 1 h).

À la fin, vous saurez

- calculer un **indice de ciel clair** (CSI) ;
- construire et estimer un modèle **AR(p)** par moindres carrés ;
- construire un **ELM** et en mesurer la variabilité ;
- comparer des modèles à la **persistance** (MAE, RMSE, score) ;
- mettre à jour un modèle **en ligne** et juger sa **frugalité**.

La chaîne complète que nous allons construire



Cette chaîne revient à chaque chapitre ; l'étape en cours y est surlignée.

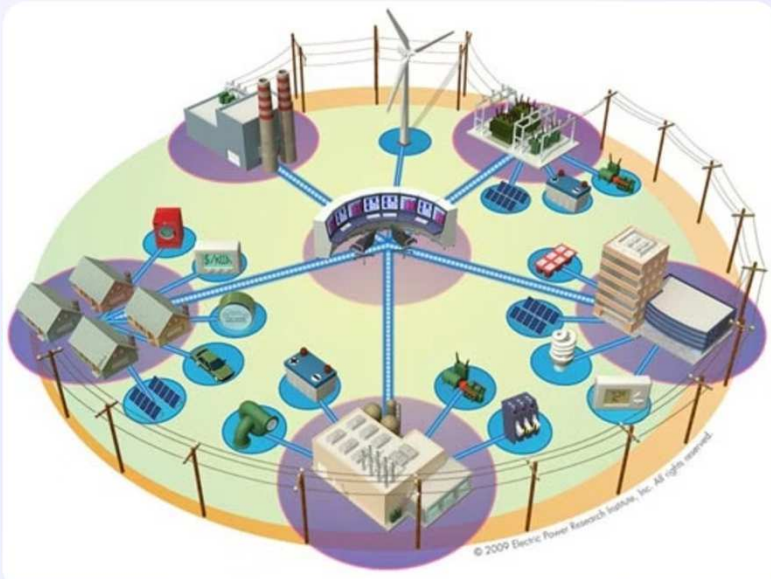
Notations (valables pour tout le cours)

Symbole	Signification
GHI_t	rayonnement global horizontal mesuré à l'heure t (Wh/m^2)
GHI_t^{cs}	GHI par ciel clair (modèle Solis)
CSI_t	indice de ciel clair : $GHI_t - GHI_t^{cs}$ (Wh/m^2)
y_t	CSI centré, la série modélisée : $y_t = CSI_t - \mu$
$\hat{y}_{t+h}$	prévision de y faite à t pour $t + h$
h	horizon, en nombre de pas (1 pas = 1 h)
p	nombre de retards (valeurs passées utilisées)
$GHI_{t+h}^{prévu}$	$\hat{y}_{t+h} + \mu + GHI_{t+h}^{cs}$; 0 la nuit ou si négatif
e_t	erreur de prévision : $GHI_{t+h}^{prévu} - GHI_t$
μ	moyenne du CSI (période d'apprentissage)

Symbole	Signification
$\mathbf{X}, \mathbf{y}$	matrice des entrées, vecteur des cibles
$\mathbf{w}$	coefficients du modèle AR
ε_t	erreur (partie imprévisible)
$\mathbf{W}_{in}, \mathbf{b}$	poids et biais aléatoires de l'ELM
$\mathbf{H}, \boldsymbol{\beta}, L$	sorties cachées, poids de sortie, nb de neurones
$\theta_z, \delta, \omega, \varphi$	angle zénithal, déclinaison, angle horaire, latitude
$\alpha ; W$	régularisation ridge ; longueur de la fenêtre glissante
N	nombre d'exemples (lignes de $\mathbf{X}$)
Δt	pas de temps (ici 1 h)

Convention : **gras minuscule** = vecteur, **gras majuscule** = matrice, maigre = nombre. Le temps est toujours en indice : y_t, y_{t-1}, CSI_{t+h} .

Pourquoi prévoir ? Réseau et production solaire



Réseau électrique intelligent. Figure : © EPRI, 2009.

La production solaire varie avec les nuages. Pour l'intégrer au réseau, il faut **l'anticiper**.

Stabilité du réseau

Anticiper les baisses de production pour garder la fréquence stable.

Stockage

Charger ou décharger les batteries au bon moment.

Équilibre offre / demande

Ajuster les autres moyens de production (hydraulique, STEP*, gaz...).

Valeur économique

Vendre la production au meilleur prix et réduire les coûts d'exploitation.

* STEP : station de transfert d'énergie par pompage (barrage réversible).

Irradiance, irradiation et production photovoltaïque

Irradiance (W/m²)

Puissance reçue par m² à un instant donné. Exemple : 1 000 W/m² en plein soleil à midi.

Irradiation (Wh/m²)

Énergie reçue par m² pendant une durée. Sur une heure : irradiation horaire en Wh/m² = irradiance moyenne en W/m².

Nos données

GHI **horaire** en Wh/m². Les fichiers Météo-France sont en J/cm² : 1 J/cm² = 2,78 Wh/m², donc Wh/m² = J/cm² / 0,36.

De l'irradiance à la puissance électrique

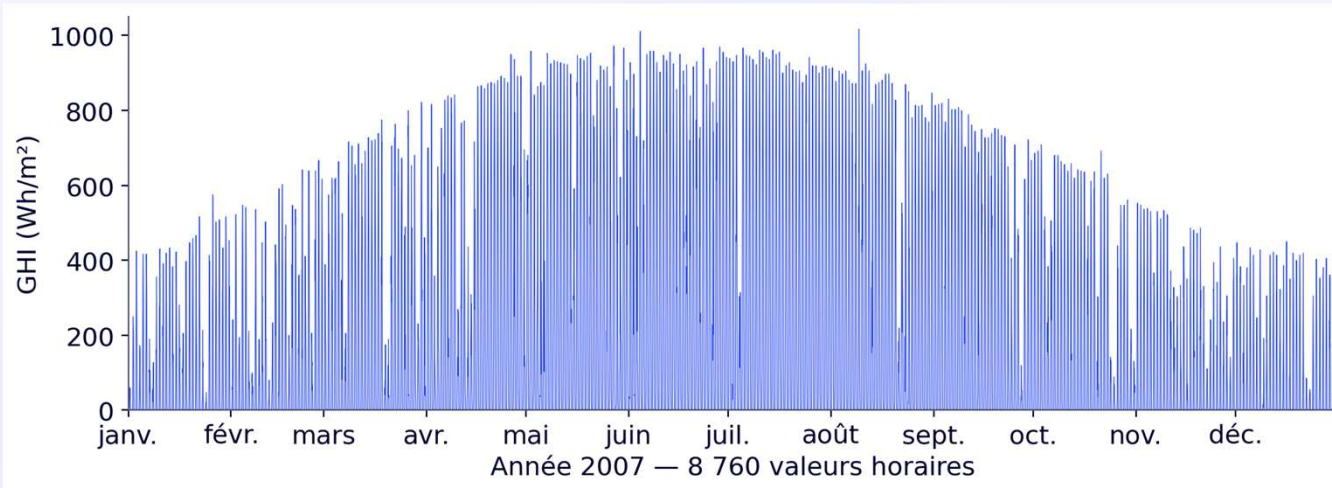
$$P \approx \eta \cdot S \cdot G$$

η : rendement du module ($\approx 20\%$), S : surface (m²), G : irradiance sur le module (W/m²).

Exemple : 10 m² sous 800 W/m² $\rightarrow P \approx 0,20 \times 10 \times 800 = 1,6$ kW.

La production est proportionnelle à l'irradiance : **prévoir le GHI, c'est prévoir la production.**

Le jeu de données : Ajaccio, 1998-2007

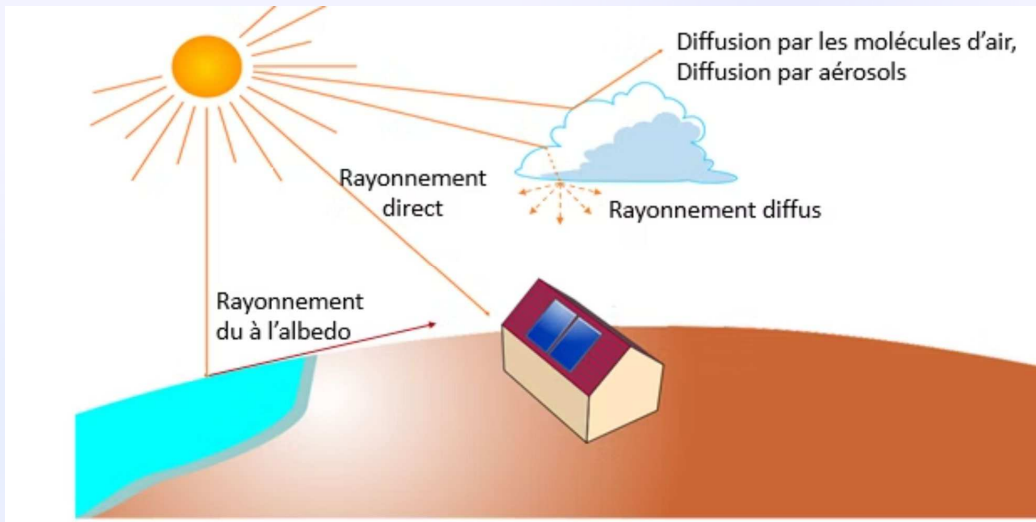


GHI horaire mesuré à Ajaccio en 2007 (Météo-France).

- **Site** : Ajaccio (Corse), 41,9° N, 8,8° E, 4 m.
- **Grandeur** : GHI horaire (Wh/m²).
- **Période** : 1998-2007, soit 87 600 heures.
- **Maximum** : 1 025 Wh/m².
- **Moyenne de jour** : 368 Wh/m².

10 ans de mesures : 8 ans pour **apprendre**, 1 an pour **régler** le modèle, 1 an pour le **tester** (chapitre 5).

Composition du rayonnement global



Direct (DNI)

Reçu directement du Soleil, mesuré sur une surface perpendiculaire aux rayons (*direct normal*). Dominant par ciel clair.

Diffus (DHI)

Diffusé par l'air, les aérosols et les nuages. Dominant par ciel couvert.

Réfléchi (albédo)

Renvoyé par le sol. Ne compte que pour une **surface inclinée**, pas pour le GHI.

Sur une surface horizontale

$$\text{GHI} = \text{DNI} \cdot \cos \theta_z + \text{DHI}$$

θ_z : angle zénithal (angle entre le Soleil et la verticale). GHI = *Global Horizontal Irradiance*.

La traversée de l'atmosphère

En traversant l'atmosphère, le rayonnement est atténué par quatre mécanismes.

Absorption

Vapeur d'eau, ozone, CO_2 : une partie de l'énergie est absorbée.

Diffusion de Rayleigh

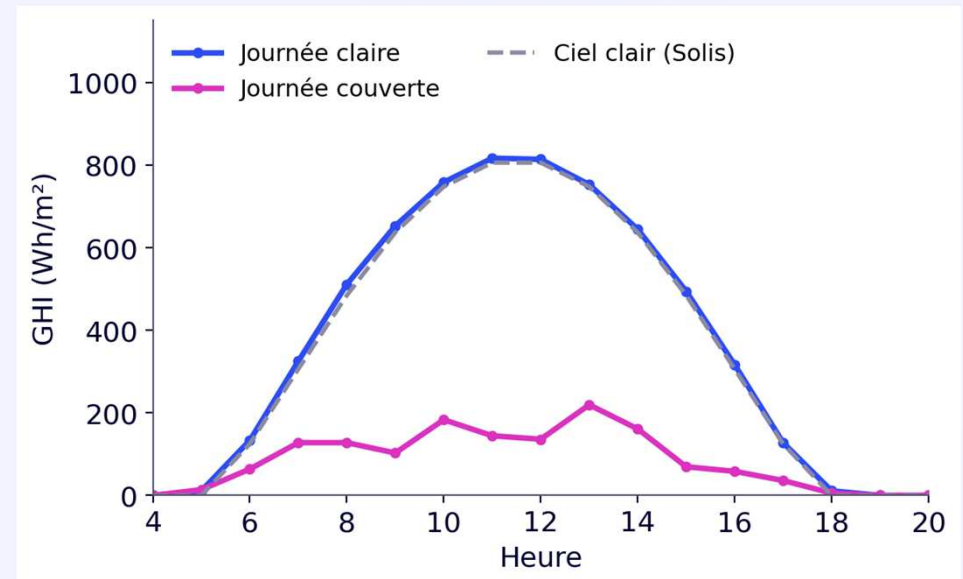
Par les molécules d'air : c'est elle qui rend le ciel bleu.

Diffusion de Mie

Par les aérosols (poussières, sel marin, pollution) : dépend de la turbidité.

Nuages

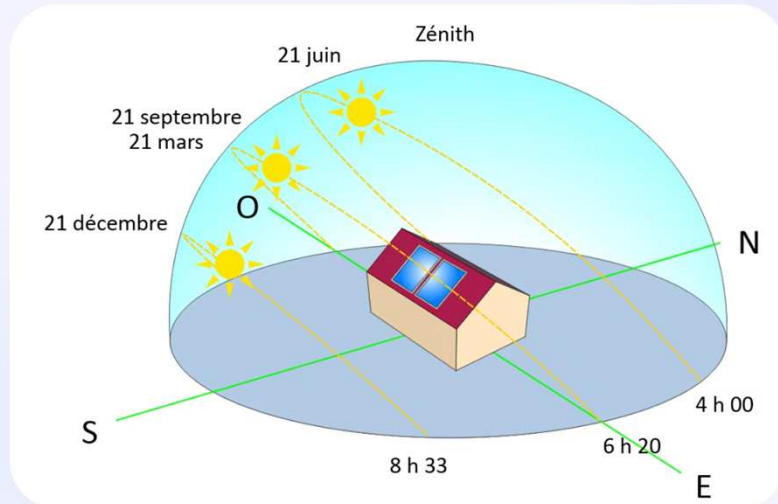
Réflexion, diffusion et absorption : **principale cause d'intermittence.**



Ajaccio, 2007 : une journée claire et une journée couverte.

Par ciel couvert, le GHI tombe ici à **environ 20 %** de sa valeur ciel clair.

La position du Soleil



Course du Soleil selon la saison.

- **Angle zénithal** θ_z : entre le Soleil et la verticale. Plus il est grand, plus l'atmosphère traversée est épaisse.
- **Déclinaison** δ : angle entre la direction du Soleil et le plan de l'équateur ; de $-23,45^\circ$ à $+23,45^\circ$ au fil de l'année.
- **Angle horaire** ω : 15° par heure, nul au midi solaire. **Latitude** φ : $41,9^\circ$ à Ajaccio.

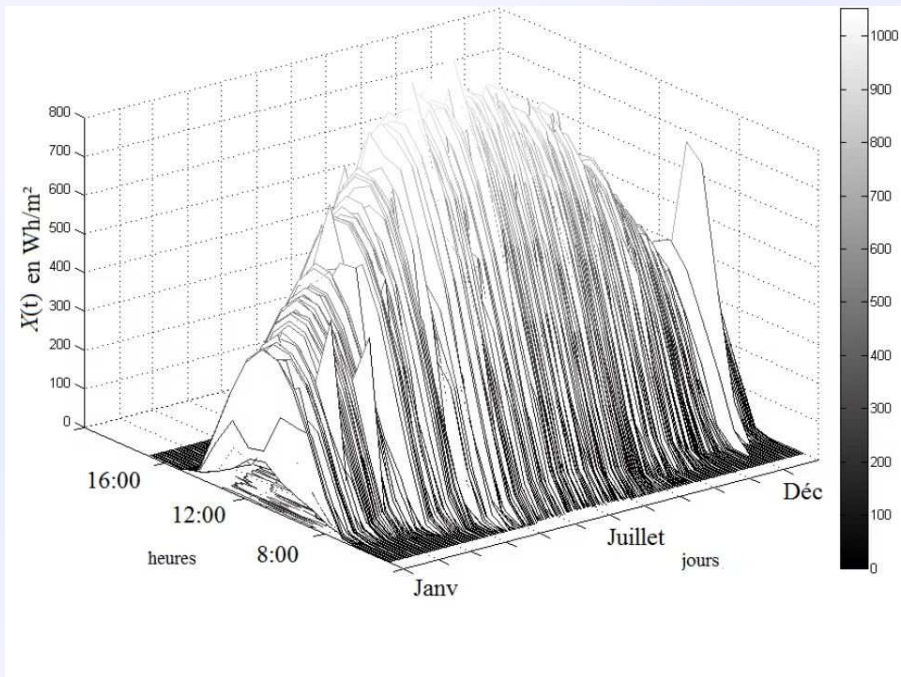
$$\cos \theta_z = \sin \varphi \cdot \sin \delta + \cos \varphi \cdot \cos \delta \cdot \cos \omega$$

$$G_0 = I_0 \cdot (d_0/d)^2 \cdot \cos \theta_z$$

G_0 : rayonnement hors atmosphère ; $I_0 \approx 1\,361 \text{ W/m}^2$ (constante solaire ; 1 367 dans les codes plus anciens).

La distance Terre-Soleil d varie de $\pm 1,7\%$, donc $(d_0/d)^2$ varie de $\pm 3,3\%$.

La double saisonnalité du GHI



GHI horaire sur une année : heures du jour $\times$ jours de l'année.

Saisonnalité journalière

Nul la nuit, maximal vers midi solaire : un cycle de **24 h** bien décrit par le modèle ciel clair.

Saisonnalité annuelle

L'amplitude du cycle journalier varie avec la saison (déclinaison, durée du jour) : un cycle de **365 jours**.

Le GHI brut **n'est pas stationnaire** : sa moyenne dépend de l'heure et du mois. Il faut retirer ces deux cycles avant de modéliser.

Qu'est-ce qu'une série temporelle ?

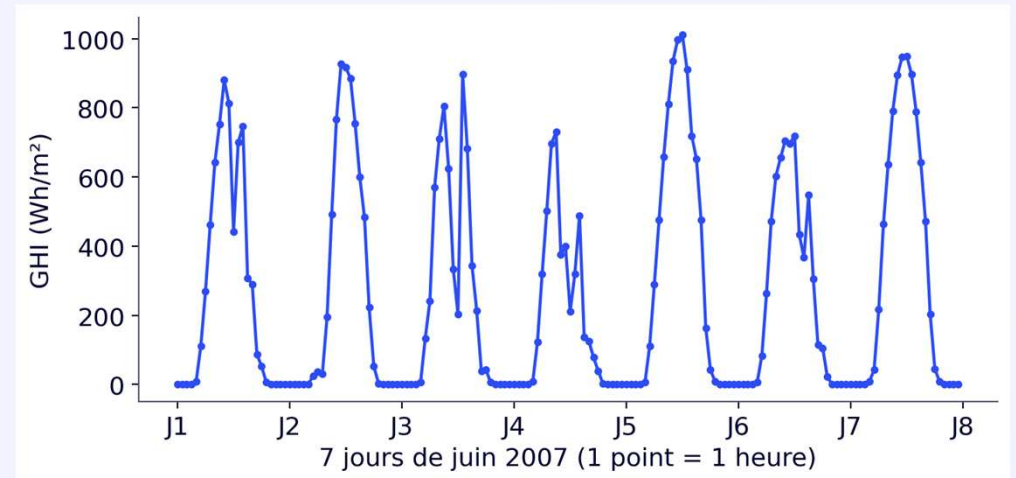
Une suite de mesures $y_1, y_2, \dots, y_N$ d'une même grandeur, prises à **pas de temps constant** Δt . Ici $\Delta t = 1$ h.

Trois propriétés

- **Régularité** : les mesures sont espacées du même pas.
- **Dépendance** : le passé renseigne sur le futur.
- **Aléa** : chaque valeur contient une part imprévisible.

Ce que l'on souhaite pour modéliser

- **Stationnarité** : moyenne et variance stables dans le temps.
- **Série complète** : pas de valeurs manquantes.



Ajaccio : une semaine de GHI horaire.

Le GHI brut n'est **pas** stationnaire (cycles jour/nuit et saisons) : c'est tout l'objet du chapitre 4.

Les composantes d'une série temporelle

Tendance T_t

Évolution lente sur le long terme : croissante, décroissante ou nulle.

Saisonnalité S_t

Cycles réguliers et reproductibles (jour, année). Très marquée pour le GHI.

Cycle C_t

Oscillations longues, non strictement périodiques (phénomènes climatiques).

Résidu R_t

Partie aléatoire non expliquée. Idéalement un **bruit blanc** (valeurs indépendantes, de moyenne nulle).

Pour le GHI : la saisonnalité domine, avec deux cycles (journalier et annuel).

Décomposer une série temporelle

Modèle additif

$$y_t = T_t + S_t + C_t + R_t$$

Quand l'amplitude des fluctuations reste constante.

Modèle multiplicatif

$$y_t = T_t \times S_t \times C_t \times R_t$$

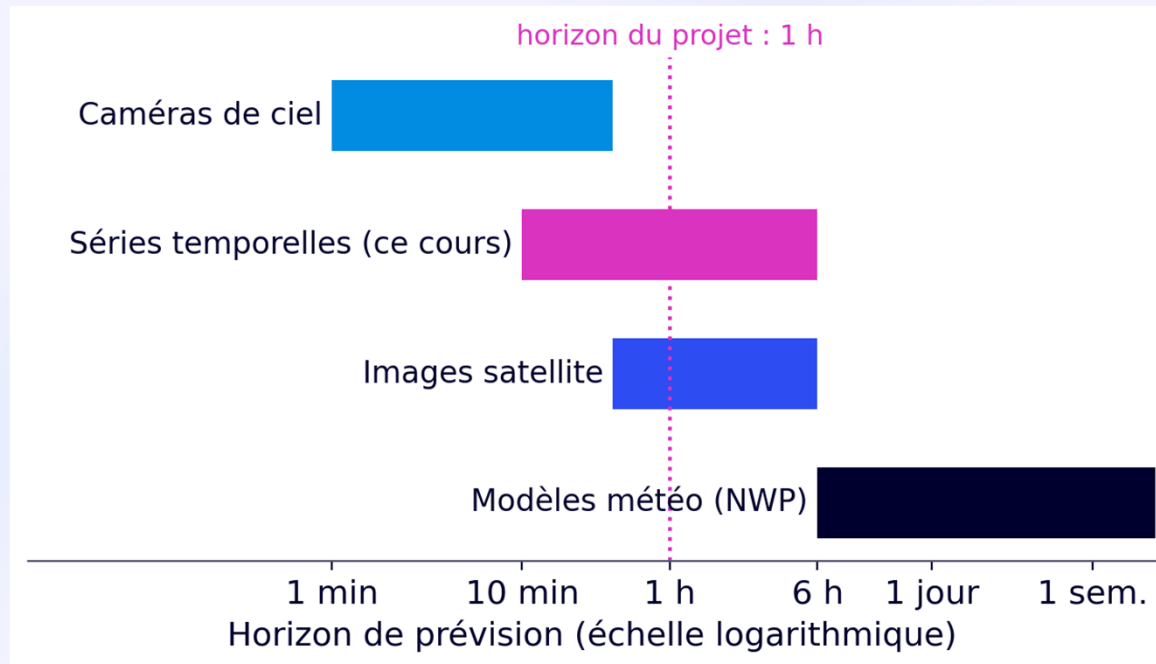
Quand l'amplitude grandit avec le niveau (cas du GHI : faible l'hiver, forte l'été).

Pourquoi décomposer ?

- Retirer la tendance et la saisonnalité (partie **prévisible sans effort**) ;
- garder un résidu plus facile à modéliser ;
- vérifier que ce résidu ressemble à un bruit blanc.

Lien avec la suite : soustraire le GHI ciel clair, $CSI_t = GHI_t - GHI_t^{cs}$, c'est appliquer le **modèle additif** : on retire d'un coup les deux cycles. Le rapport (modèle multiplicatif) existe aussi : chapitre 4.

Horizon, pas de temps et types de prévision



Méthodes usuelles selon l'horizon (ordres de grandeur).

- **Pas de temps** $\Delta t = 1 \text{ h}$; **horizon** h en nombre de pas : $h = 1 \rightarrow 1 \text{ h}$, $h = 3 \rightarrow 3 \text{ h}$.
- **Prévision directe** : un modèle entraîné pour chaque horizon h .
- **Prévision réursive** : on prévoit $t + 1$, puis on réinjecte cette prévision pour $t + 2$, etc.

Dans ce cours : horizon 1 h, prévision directe.

Prévoir : une partie prévisible et une erreur

$$y_{t+1} = f(y_t, y_{t-1}, \dots, y_{t-p+1}) + \varepsilon_{t+1}$$

La partie prévisible

f résume ce que le passé nous apprend sur le futur. La prévision est $\hat{y}_{t+1} = f(y_t, \dots, y_{t-p+1})$.

L'erreur

$\varepsilon_{t+1} = y_{t+1} - \hat{y}_{t+1}$ est la part imprévisible. Si le modèle est bon, ε ressemble à un bruit blanc.

L'objectif

Choisir f pour que les erreurs soient **les plus petites possible** en moyenne quadratique.

Le formalisme convient au GHI : mesures régulières et forte dépendance d'une heure à l'autre (inertie des nuages). Pour aller plus loin : annexe B.

Premier exemple : le modèle AR(1), chiffré

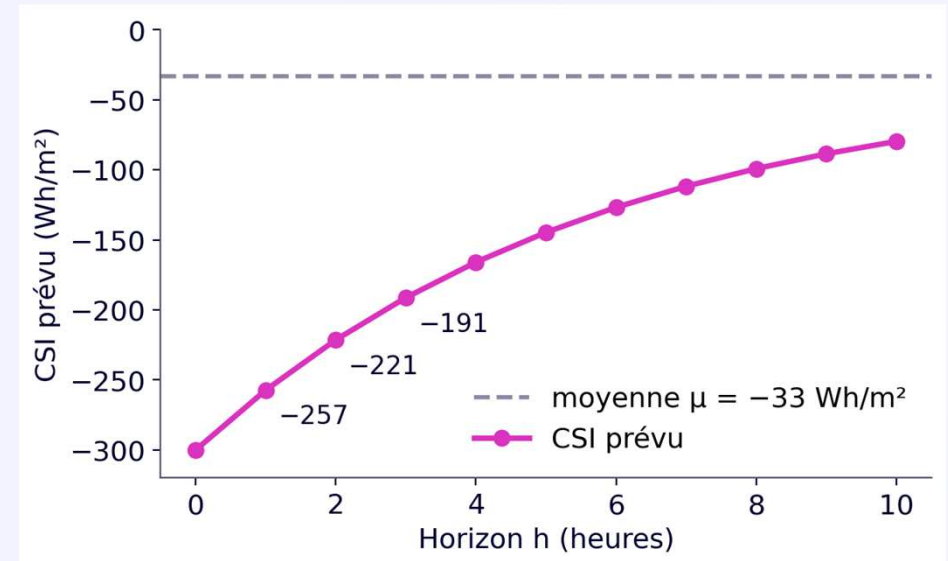
$$y_{t+1} = w \cdot y_t + \varepsilon_{t+1}$$

- $y_t = \text{CSI}_t - \mu$: écart du CSI à sa moyenne ; w : poids du dernier écart.
- $|w| < 1$: la série revient toujours vers sa moyenne.

$$\hat{y}_{t+h} = w^h \cdot y_t$$

Ajaccio : $\mu \approx -33 \text{ Wh/m}^2$ et $w \approx 0,84$. Nuage : $\text{CSI}_t = -300 \text{ Wh/m}^2$, donc $y_t = -267$.

$\hat{y}_{t+1} = 0,84 \times (-267) = -224$: CSI prévu = -257 Wh/m^2 ; puis -221 et -191 à 2 h et 3 h.



Le CSI prévu retourne vers sa moyenne quand h augmente.

Le modèle ciel clair

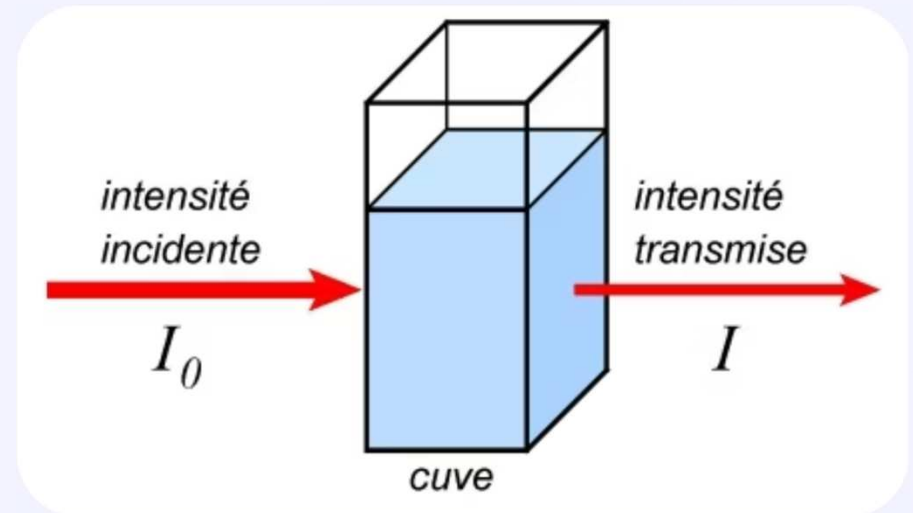
Qu'est-ce que c'est ?

Le GHI que l'on mesurerait **en l'absence de nuages**, compte tenu de la position du Soleil et de l'atmosphère claire.

Entrées

- position du site (latitude, longitude, altitude) ;
- position du Soleil (θ_z) ;
- vapeur d'eau et aérosols (AOD : épaisseur optique des aérosols).

Rôle : GHI_t^{cs} sert de référence pour retirer les cycles jour/nuit et saisons.



Loi de Beer-Lambert : l'intensité décroît exponentiellement avec l'épaisseur traversée.

Le modèle Solis

Solis (Ineichen, 2008) est une version simplifiée d'un modèle de transfert radiatif. L'idée de départ est la loi de Beer-Lambert :

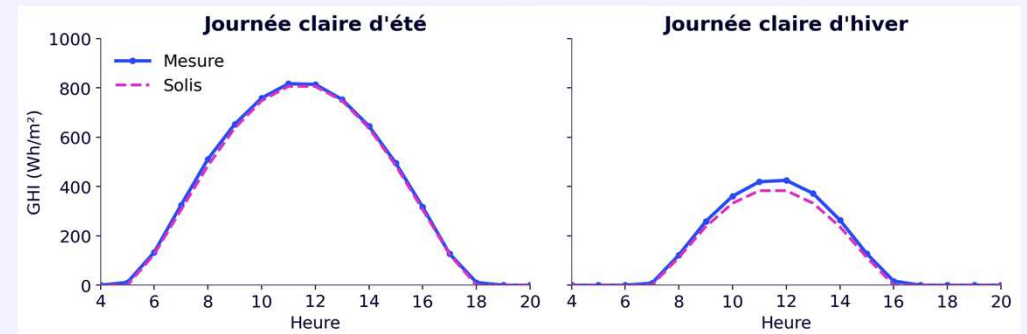
$$GHI^{cs} \approx I_0 \cdot \exp(-\tau / \cos \theta_z) \cdot \cos \theta_z$$

- τ : épaisseur optique (air, vapeur d'eau, aérosols) ;
- Solis ajuste τ et l'exposant selon la vapeur d'eau, les aérosols et l'altitude.

Réglages : zone rurale, colonne d'eau 3,79 cm, AOD 0,16, $I_0 = 1\,367 \text{ W/m}^2$.

La mesure horaire est la **moyenne de l'heure écoulée** : on calcule le Soleil **au milieu de l'heure**, $\omega = (i - 12,5) \times 15^\circ$.

Écart moyen par ciel clair à Ajaccio : $\approx 32 \text{ Wh/m}^2$.



Ajaccio, 2007 : Solis et mesure par ciel clair.

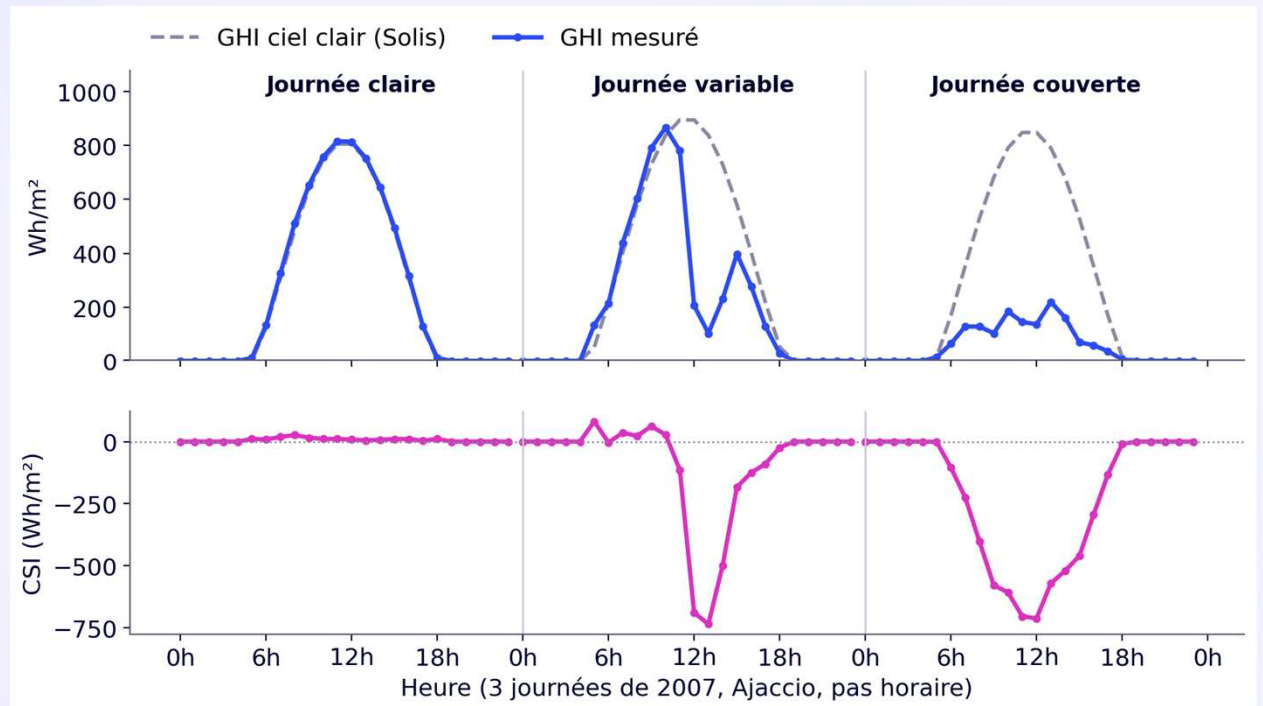
Stationnariser : l'indice de ciel clair (CSI)

$$CSI_t = GHI_t - GHI_t^{cs}$$

- ≈ 0 par ciel clair, et **0 la nuit** ;
- **négatif** sous les nuages (jusqu'à $-GHI_t^{cs}$) ;
- moyenne $\mu \approx -33 \text{ Wh/m}^2$: on **centre**, $y_t = CSI_t - \mu$.

La démarche

1. calculer GHI_t^{cs} puis le CSI ;
2. centrer : $y_t = CSI_t - \mu$;
3. prévoir $\hat{y}_{t+h}$;
4. reconstruire $GHI_{t+h}^{prévu} = \hat{y}_{t+h} + \mu + GHI_{t+h}^{cs}$ (0 la nuit, jamais négatif).



Ajaccio, 2007 : en haut le GHI, en bas le CSI. Le cycle du jour a disparu du CSI.

Deux façons de stationnariser

Le modèle Solis permet de retirer les cycles de deux manières :

Différence (modèle additif) — ce cours

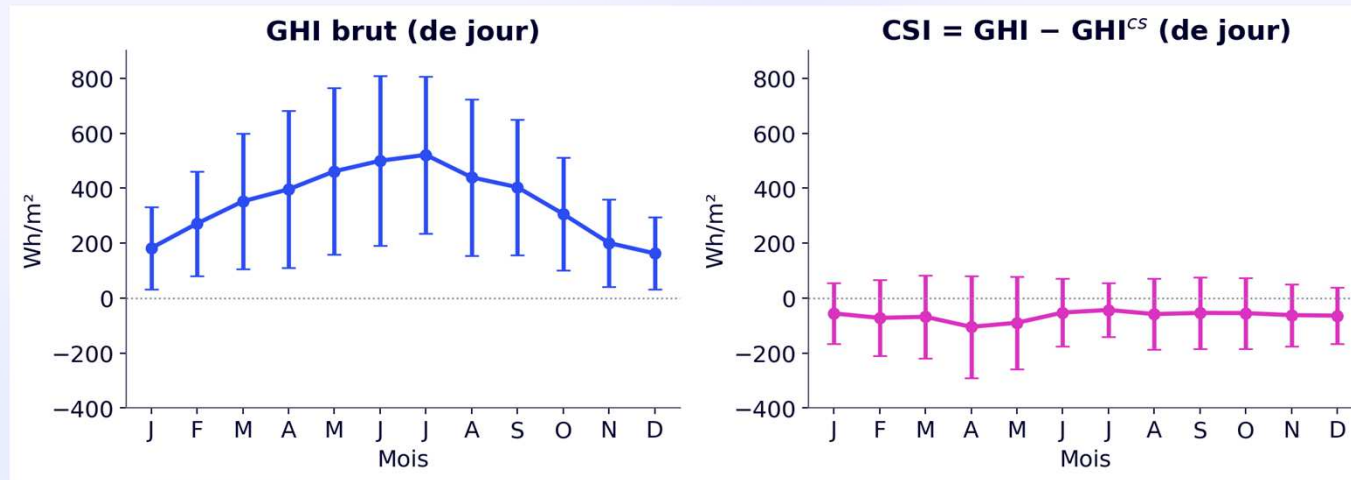
- $CSI_t = GHI_t - GHI_t^{cs}$, en Wh/m^2 ;
- pas de division : valable **de jour comme de nuit** (la nuit, $0 - 0 = 0$) ;
- on garde toutes les heures, sans filtrage.

Rapport (modèle multiplicatif)

- $k_t = GHI_t / GHI_t^{cs}$, sans unité (≈ 1 par ciel clair) ; très utilisé dans la littérature ;
- la nuit : **0 / 0**, indéfini ; au lever et au coucher, $GHI^{cs} \approx 0$: valeurs qui explosent ;
- oblige à **filtrer sur la hauteur solaire** (garder par exemple $h > 10^\circ$).

Pour des données horaires avec nuits, le rapport est **pas ou peu adapté** : nous utilisons la **différence**. Le rapport est aussi très sensible aux petits décalages d'horodatage et aux seuils de filtrage choisis (Voyant et al., 2025a ; 2026).

Le CSI est-il stationnaire ?



Moyenne $\pm$ écart-type par mois, heures de jour, Ajaccio 1998-2007.

- **GHI** : la moyenne va de 163 à 521 Wh/m² selon le mois.
- **CSI** : la moyenne reste entre -105 et -43 Wh/m², proche de 0.
- L'écart-type du CSI varie encore (97 à 185 Wh/m²) : plus de nuages au printemps.

Le CSI est **approximativement stationnaire** : beaucoup plus que le GHI, sans être parfait. Définition formelle : annexe A.

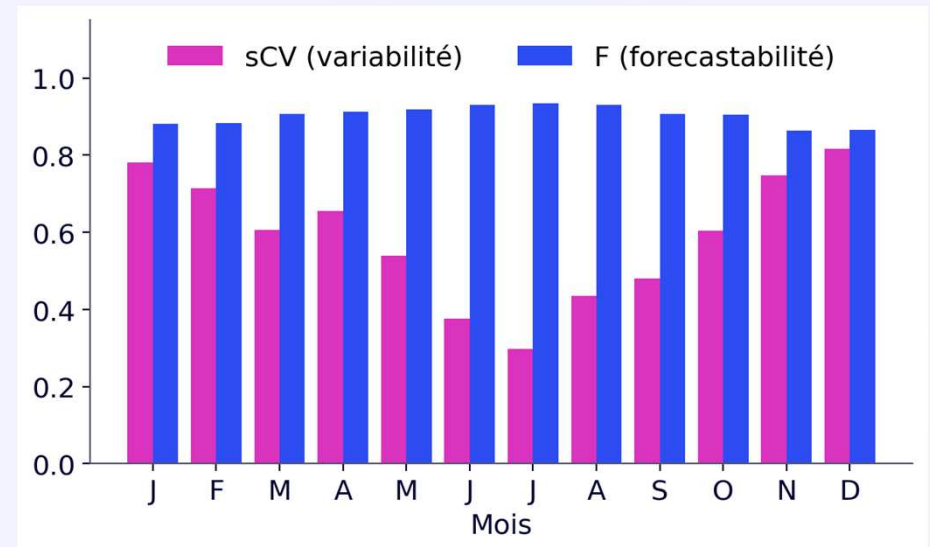
Variabilité d'un site : le sCV

Combien le GHI s'écarte-t-il, en moyenne, du ciel clair ? On mesure l'écart $\varepsilon_t = \text{GHI}_t - \text{GHI}^{\text{cs}}_t$ (c'est le CSI) :

$$\text{sCV} = (\sqrt{2} / (1 - d)) \cdot \sqrt{[\sum_t (\text{GHI}_t - \text{GHI}^{\text{cs}})^2 / \sum_t (\text{GHI}^{\text{cs}})^2]}$$

- **Numérateur** : la taille typique des écarts au ciel clair (les nuages).
- **Dénominateur** : la taille typique du GHI ciel clair.
- $\sqrt{2} / (1 - d)$ ramène le pire cas à 1 : ciel clair, puis très couvert, une heure sur deux ($d = 0,2$: même sous un ciel couvert, il reste au moins 20 % du ciel clair).

sCV = 0 : toujours ciel clair ; sCV = 1 : variabilité maximale. **Il ne dépend pas de l'horizon** : c'est une propriété du site.



Ajaccio 1998-2007 : sCV et F à 1 h, mois par mois (Voyant et al., 2026).

Ajaccio : sCV = 0,52 sur 1998-2007.

Hiver : 0,8 (ciel changeant) ; juillet : 0,3 (ciel souvent clair).

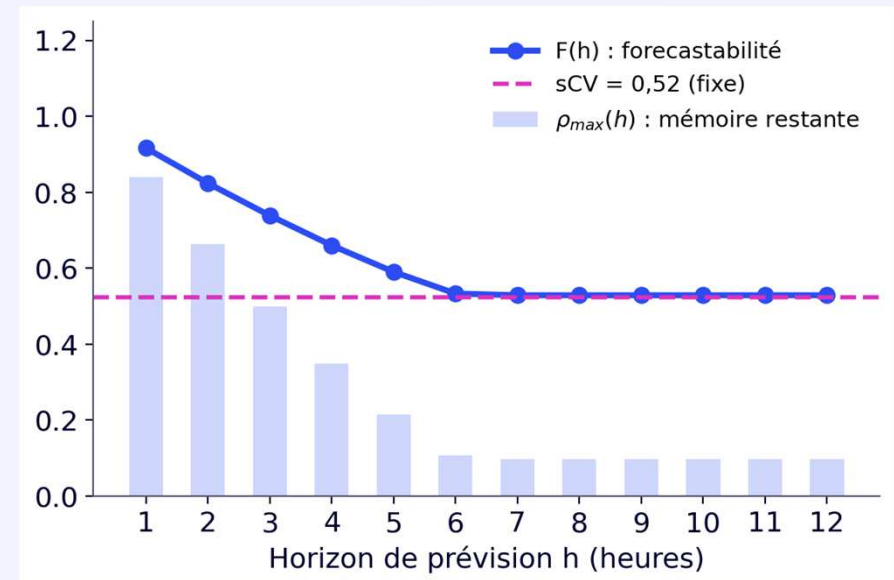
Forecastabilité : F dépend de l'horizon

Une série variable n'est pas forcément imprévisible : si les écarts au ciel clair **durent**, on peut les anticiper.

- $\rho(k)$: corrélation entre l'écart au ciel clair maintenant, ε_t , et k heures plus tard, ε_{t+k} (la **mémoire** des nuages).
- $\rho_{\max}(h)$: la plus forte de ces corrélations pour k entre h et 12 h : ce qu'il reste de mémoire quand on prévoit à l'horizon h.

$$F(h) = 1 - sCV \cdot (1 - \rho_{\max}(h))$$

- mémoire parfaite ($\rho_{\max} = 1$) : $F = 1$, tout est prévisible ;
- aucune mémoire ($\rho_{\max} = 0$) : $F = 1 - sCV$.

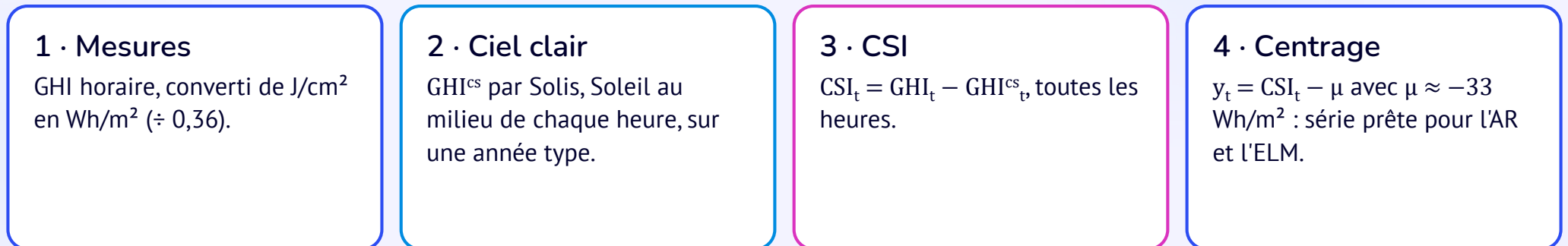
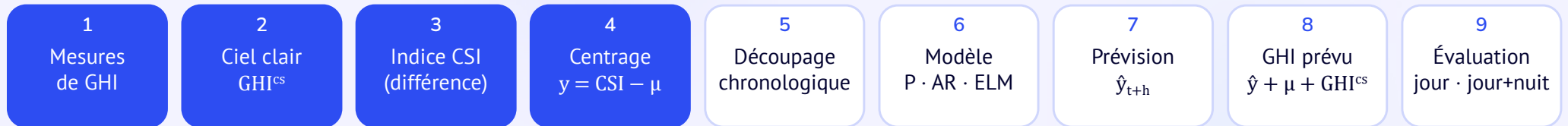


Ajaccio 1998-2007 : F et mémoire restante selon l'horizon (Voyant et al., 2026).

F = 0,92 à 1 h, 0,74 à 3 h, 0,53 à 6 h.

L'atmosphère est **chaotique** : la mémoire des nuages s'efface en quelques heures. sCV, lui, ne bouge pas.

Le pré-traitement pas à pas



Avec la différence, les nuits ne posent aucun problème ($0 - 0 = 0$) : **pas de filtrage sur la hauteur solaire**, contrairement au rapport.

Du modèle AR à la régression linéaire

Un modèle **AR(p)** prévoit la prochaine valeur à partir des p dernières :

$$y_{t+1} = w_1 y_t + w_2 y_{t-1} + \dots + w_p y_{t-p+1} + \varepsilon_{t+1}$$

Écriture matricielle

$$\mathbf{y} = \mathbf{X} \mathbf{w} + \boldsymbol{\varepsilon}$$

- $\mathbf{y}$ ($N \times 1$) : les valeurs à prévoir ;
- $\mathbf{X}$ ($N \times p$) : une ligne par exemple, avec les p retards ;
- $\mathbf{w}$ ($p \times 1$) : les coefficients à estimer ;
- $\boldsymbol{\varepsilon}$ ($N \times 1$) : les erreurs.

C'est une **régression linéaire multiple**. Comme y est centré, **pas besoin de constante**.

Construire $\mathbf{X}$ et $\mathbf{y}$: un exemple chiffré

Cinq valeurs de CSI centré (Wh/m^2) : $y_1 = -10$; $y_2 = -25$; $y_3 = -240$; $y_4 = -180$; $y_5 = -30$. On prend $p = 2$.

Ligne	y_t	y_{t-1}	Cible y_{t+1}
$t = 2$	-25	-10	-240
$t = 3$	-240	-25	-180
$t = 4$	-180	-240	-30

- Les deux colonnes du milieu forment $\mathbf{X}$ (3×2), la dernière forme $\mathbf{y}$ (3×1).
- Chaque ligne décale la précédente d'un pas.

En pratique : on garde **toutes les heures**, nuits comprises (la nuit, $y = -\mu \approx 33$). À Ajaccio : 8 ans, soit environ 70 000 lignes pour l'apprentissage.

Les moindres carrés ordinaires (MCO)

On cherche $\mathbf{w}$ qui rend la somme des erreurs au carré minimale :

$$J(\mathbf{w}) = \|\mathbf{y} - \mathbf{X}\mathbf{w}\|^2 = \sum_i (y_i - \mathbf{x}_i^T \mathbf{w})^2$$

Au minimum, la dérivée est nulle :

$$-2 \mathbf{X}^T(\mathbf{y} - \mathbf{X}\mathbf{w}) = 0 \Rightarrow \mathbf{X}^T\mathbf{X} \mathbf{w} = \mathbf{X}^T\mathbf{y}$$

- Conditions : beaucoup plus d'exemples que de coefficients ($N \gg p$) ;
- colonnes de $\mathbf{X}$ non proportionnelles entre elles.

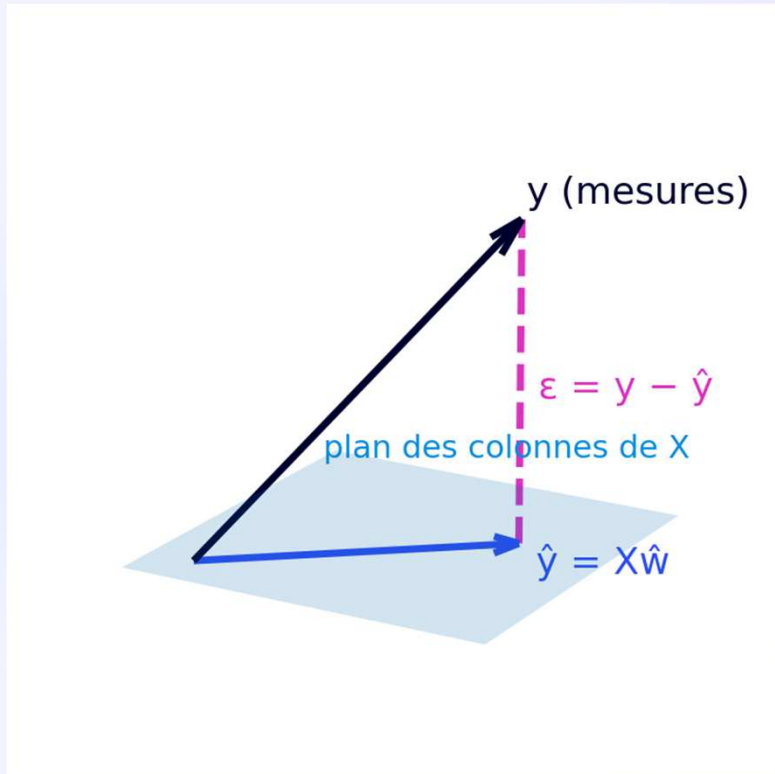
$$\hat{\mathbf{w}} = (\mathbf{X}^T\mathbf{X})^{-1} \mathbf{X}^T\mathbf{y}$$

Solution exacte, en une seule fois : pas d'itération, un seul minimum.

Ajaccio, AR(2) : $\hat{\mathbf{w}} = (0,97 ; -0,15)$, calculé en 6 ms.

Le dernier écart y_t pèse de loin le plus.

Interprétation géométrique



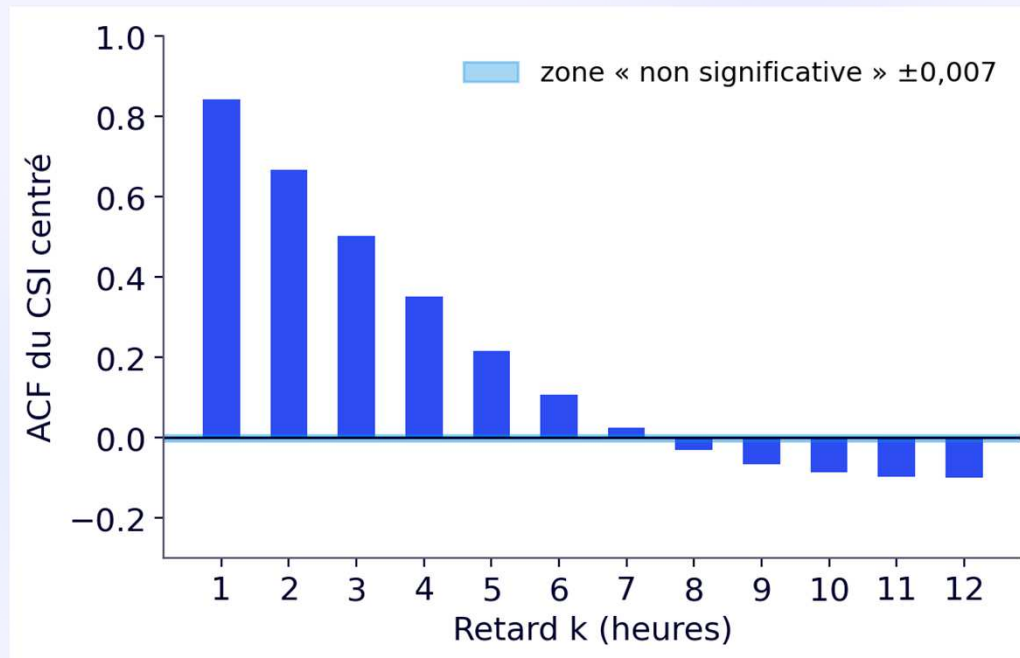
- $X\hat{w}$ est le point du « plan des colonnes de X » le **plus proche** de y : c'est sa projection orthogonale.
- L'erreur $\epsilon = y - X\hat{w}$ est **perpendiculaire** à ce plan.
- Pythagore : $\|y\|^2 = \|\hat{y}\|^2 + \|\epsilon\|^2$. C'est la base du coefficient R^2 .

Les MCO ne « devinent » rien : ils projettent les mesures sur ce que les entrées permettent d'expliquer.

Apprentissage, validation, test



Choisir p (1/3) : autocorrélation (ACF)

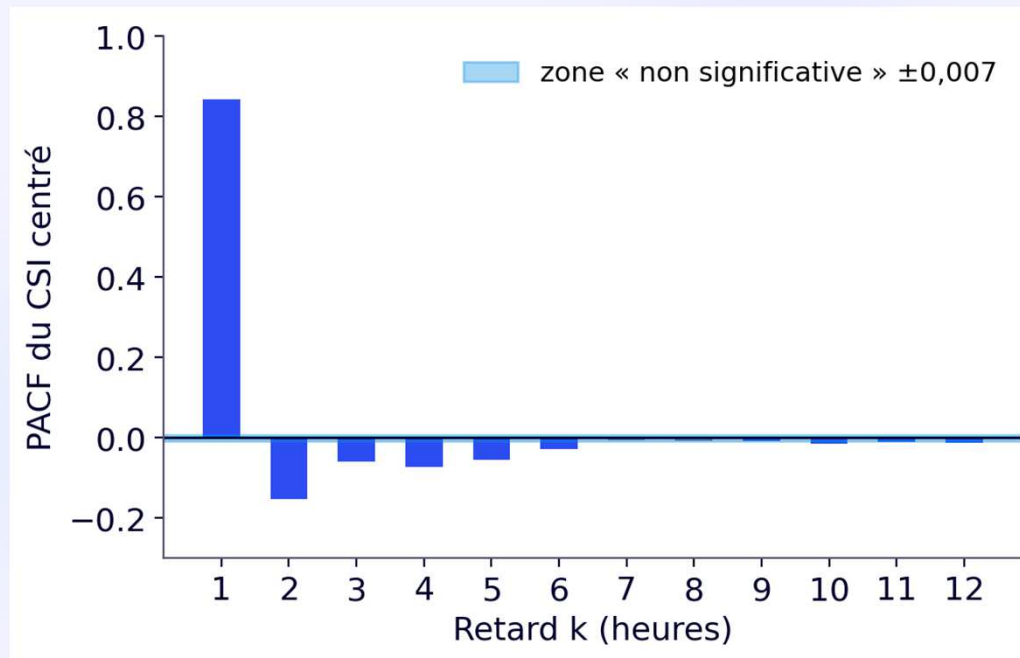


ACF du CSI centré, Ajaccio 1998-2005, toutes les heures.

- **ACF(k)** : corrélation entre y_t et y_{t-k} , la valeur k heures avant.
- Ici $ACF(1) = 0,84$, $ACF(2) = 0,67$, $ACF(3) = 0,50$: elle décroît lentement.
- Mais cette corrélation se **transmet de proche en proche** : si y_t dépend de y_{t-1} , qui dépend de y_{t-2} , alors y_t et y_{t-2} sont corrélés même sans lien direct.

L'ACF ne dit donc pas combien de retards sont **vraiment utiles**. Il faut la PACF.

Choisir p (2/3) : autocorrélation partielle (PACF)



PACF du CSI centré, Ajaccio 1998-2005, toutes les heures.

- **PACF(k)** : une fois connus $y_{t-1}, \dots, y_{t-k+1}$, que **rajoute** y_{t-k} ? C'est le dernier coefficient w_k d'un AR(k).
- Une barre **hors de la zone bleue** : ce retard apporte une information directe.
- Pour un vrai AR(p), la PACF s'annule après le retard p.

Ici PACF(1) = **0,84**, PACF(2) = -0,15, puis des valeurs sous 0,08 : la PACF suggère **p ≈ 2**.

Mais avec $N \approx 70\,000$ points, le seuil $\pm 0,007$ est minuscule : presque tout est « significatif ». Significatif ≠ utile.

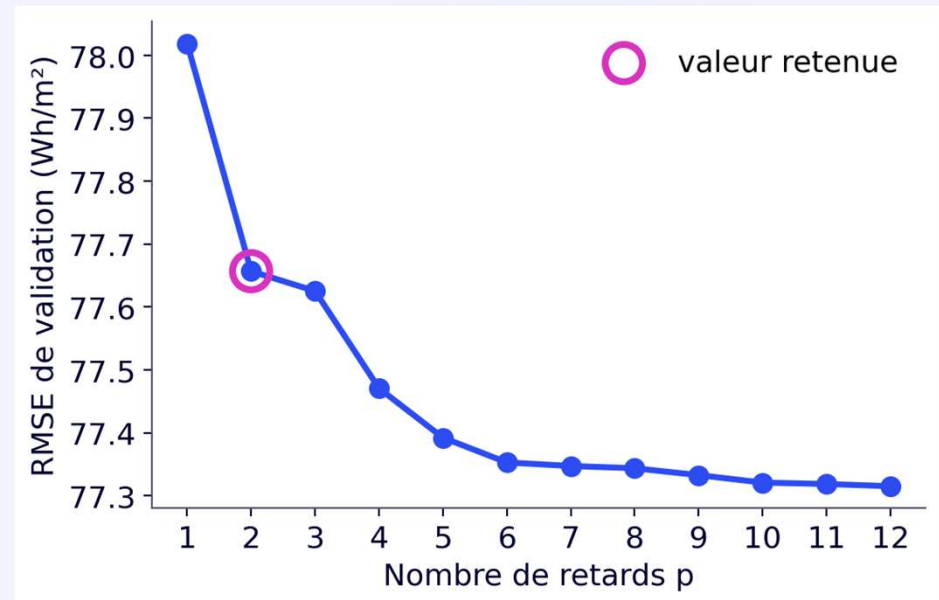
Choisir p (3/3) : approche filtre ou approche enveloppe

Filtre (filter)

- on regarde **les données seules**, avant tout modèle : PACF, critères AIC / BIC (annexe C) ;
- rapide, mais indirect.

Enveloppe (wrapper)

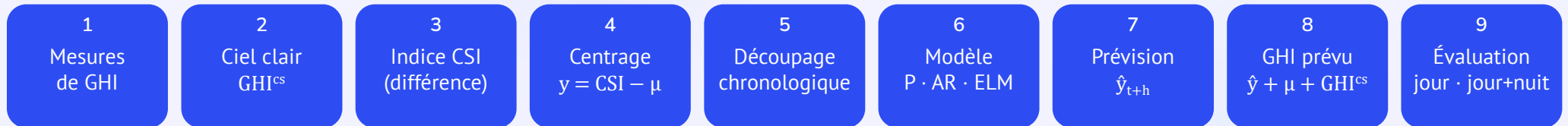
- on **entraîne** un $AR(p)$ pour $p = 1$ à 12 sur 1998-2005 ;
- on mesure **l'erreur de prévision** sur la validation (2006) ;
- guidé par l'objectif réel, mais plus coûteux.



Approche enveloppe : RMSE de validation (2006, heures de jour) selon p .

$p = 1$: 78,0 ; $p = 2$: 77,7 ; $p = 12$: 77,3 Wh/m^2 . **Parcimonie** : le plus petit p à moins de 0,5 % du meilleur, soit **$p = 2$** .

Récapitulatif : prévoir le GHI avec un AR(p)



1. Calculer GHI^{cs} (Solis), puis $CSI_t = GHI_t - GHI^{cs}_t$; le centrer : $y_t = CSI_t - \mu$.
2. Découper : apprentissage / validation / test.
3. Choisir p sur la validation.
4. Construire $\mathbf{X}$ et $\mathbf{y}$; estimer $\hat{\mathbf{w}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$.
5. Prévoir $\hat{y}_{t+1} = \hat{\mathbf{w}}^T \mathbf{x}_t$ avec $\mathbf{x}_t = (y_t, \dots, y_{t-p+1})$.
6. Reconstruire $GHI^{prévu}_{t+1} = \hat{y}_{t+1} + \mu + GHI^{cs}_{t+1}$; mettre **0 la nuit** (Soleil sous l'horizon) et **0 si la prévision est négative**.
7. Évaluer sur 2007 : de jour, puis jour + nuit.

Simple, rapide, interprétable : c'est la **référence linéaire** à battre.

Les modèles de référence

Tout modèle doit être comparé à une référence simple. S'il ne fait pas mieux, il n'apporte rien.

Persistence simple

« Il fera comme maintenant. »

$$GHI^{\text{prévu}}_{t+1} = GHI_t$$

Ignore la course du Soleil.

Ciel clair

« Il fera beau. »

$$GHI^{\text{prévu}}_{t+1} = GHI^{\text{cs}}_{t+1}$$

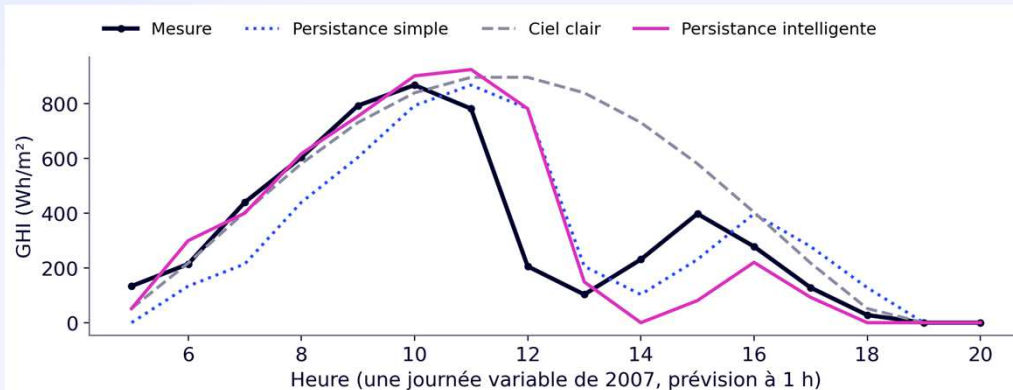
Ignore les nuages.

Persistence intelligente

« Le ciel restera le même. »

$$GHI^{\text{prévu}}_{t+1} = GHI_t - GHI^{\text{cs}}_t + GHI^{\text{cs}}_{t+1}$$

Garde le même écart au ciel clair.



Persistence intelligente = AR(1) avec $w = 1$:
la référence du score SS.

Deux règles, pour tous les modèles : 1) la nuit (Soleil sous l'horizon, $\theta_z > 90^\circ$), prévision = 0 ; 2) une prévision négative est remplacée par 0.

Mesurer l'erreur

$$e_t = \text{GHI}^{\text{prévu}}_t - \text{GHI}_t \quad (\text{erreur de prévision})$$

$$\text{nMBE} = [(1/N) \sum e_t] / \text{moyenne}(\text{GHI}_t)$$

$$\text{MAE} = (1/N) \sum |e_t|$$

$$\text{RMSE} = \sqrt{[(1/N) \sum e_t^2]} ; \text{nRMSE} = \text{RMSE} / \text{moyenne}(\text{GHI}_t)$$

- **nMBE** (biais) : > 0 si le modèle **sur-estime** en moyenne, < 0 s'il **sous-estime**. Les erreurs de signes opposés se compensent : ce n'est **pas** un score pour départager les modèles.
- **MAE** : erreur moyenne, en Wh/m², facile à interpréter.
- **RMSE** : punit davantage les grosses erreurs (nuage soudain). Toujours $\geq$ MAE.
- **nRMSE** : RMSE en % du GHI moyen, pour comparer des sites.

Condition essentielle : si $|\text{nMBE}| > 1 \%$, il y a sans doute un problème, ou un très mauvais modèle. On peut retrancher le biais, à condition de le tester ; nous ne le faisons pas ici.

Erreurs calculées sur le **GHI reconstruit**, sur l'année de **test** (2007) : d'abord aux **heures de jour**, puis **jour + nuit**. Toujours dire lesquelles !

Comparer à une persistance : SS et NICE Σ

$$SS = 1 - \text{RMSE}_{\text{modèle}} / \text{RMSE}_{\text{persistance intelligente}}$$

- **Score de compétence** : > 0 si l'on bat la persistance intelligente, < 0 sinon.
- Ne regarde que le RMSE.

$$\text{NICE}^k = L^k(\text{modèle}) / L^k(\text{persistance simple})$$

$$L^k = ((1/N) \sum |e_t|^k)^{1/k} ; \text{NICE}^\Sigma = (\text{NICE}^1 + \text{NICE}^2 + \text{NICE}^3) / 3$$

$e_t = \text{GHI}^{\text{prévu}}_t - \text{GHI}_t$: l'erreur de prévision (diapo précédente).

- $k = 1$: erreurs moyennes (MAE) ; $k = 2$: RMSE ; $k = 3$: grosses erreurs ;
- **NICE $^\Sigma < 1$** : le modèle bat la persistance simple sur les trois normes à la fois.

Deux références différentes : SS se compare à la persistance **intelligente**, NICE $^\Sigma$ à la persistance **simple** (Voyant et al., 2025b).

Résultats à Ajaccio : prévision à 1 h

Modèle	RMSE jour	nRMSE jour	nRMSE jour + nuit	nMBE jour	SS	NICE ^Σ
Persistence simple	135	36,4 %	51,5 %	-0,9 %	-0,66	1,00
Ciel clair	148	39,9 %	56,4 %	+16,4 %	-0,82	1,06
Persistence intelligente	82	22,0 %	31,0 %	+1,3 %	0,00	0,59
AR(2)	79	21,3 %	30,1 %	+2,0 %	+0,03	0,58
ELM (L = 50, meilleur tirage)	76	20,5 %	28,9 %	+1,1 %	+0,07	0,55

Test 2007. Jour : 4380 heures, GHI moyen 372 Wh/m². Jour + nuit : 8759 heures, GHI moyen 186 Wh/m². RMSE en Wh/m² ; SS et NICE calculés de jour.

La persistance intelligente divise le RMSE de la persistance simple par 1,7. L'AR gagne **3 %** sur elle, l'ELM **7 %**.

Jour seulement ou jour + nuit ?

79 → 56

RMSE de l'AR (Wh/m²)

372 → 186

GHI moyen (Wh/m²)

21 → 30 %

nRMSE de l'AR

- La nuit (Soleil sous l'horizon), tous les modèles prévoient 0 par construction et ne se trompent pas : ajouter ces heures **fait baisser le RMSE** ($\approx \div \sqrt{2}$).
- Mais la moyenne du GHI est **divisée par 2** : le nRMSE **augmente**.
- Le classement des modèles, lui, ne change pas.

Un chiffre d'erreur sans la liste des heures utilisées ne veut rien dire : **toujours préciser jour ou jour + nuit**.

Le biais (nMBE) : effet des deux règles

Règle 1 (nuit) : quand le Soleil est sous l'horizon ($\theta_z > 90^\circ$), la prévision vaut 0. **Règle 2** : une prévision négative est remplacée par 0.

nMBE : biais moyen, en % du GHI moyen (voir « Mesurer l'erreur »). Test 2007.

Modèle	nMBE sans règles, jour	nMBE sans règles, jour + nuit	nMBE avec règles, jour	nMBE avec règles, jour + nuit
Persistence simple	-0,88 %	0,00 %	-0,88 %	-1,02 %
Ciel clair	+16,44 %	+16,28 %	+16,44 %	+16,28 %
Persistence intelligente	+0,35 %	0,00 %	+1,25 %	+1,11 %
AR(2)	+1,49 %	-0,24 %	+2,02 %	+1,87 %
ELM	+0,69 %	-0,38 %	+1,13 %	+0,99 %

Comment lire ce tableau

- **Sans règles, jour + nuit** : les deux persistances ont un nMBE **exactement nul**.
- **Règle 2** : mettre à 0 les prévisions négatives **remonte** la moyenne (persistence intelligente : +0,35 % → +1,25 % de jour).
- **Règle 1** : le nMBE de la persistence simple passe de 0 à **-1,02 %** sur jour + nuit (diapo suivante).
- **Ciel clair** : +16 % : il prévoit toujours du soleil.

Un nMBE dépend de la façon de le calculer : **toujours préciser** jour ou jour + nuit, avec ou sans règles.

Pourquoi la persistance passe de 0 à -1 % ?

	avant-dernière heure de jour	dernière heure de jour	première heure de nuit
Mesure	150	43	0
Persistance sans règles	...	150	43
Avec la règle 1 (nuit)	...	150	0

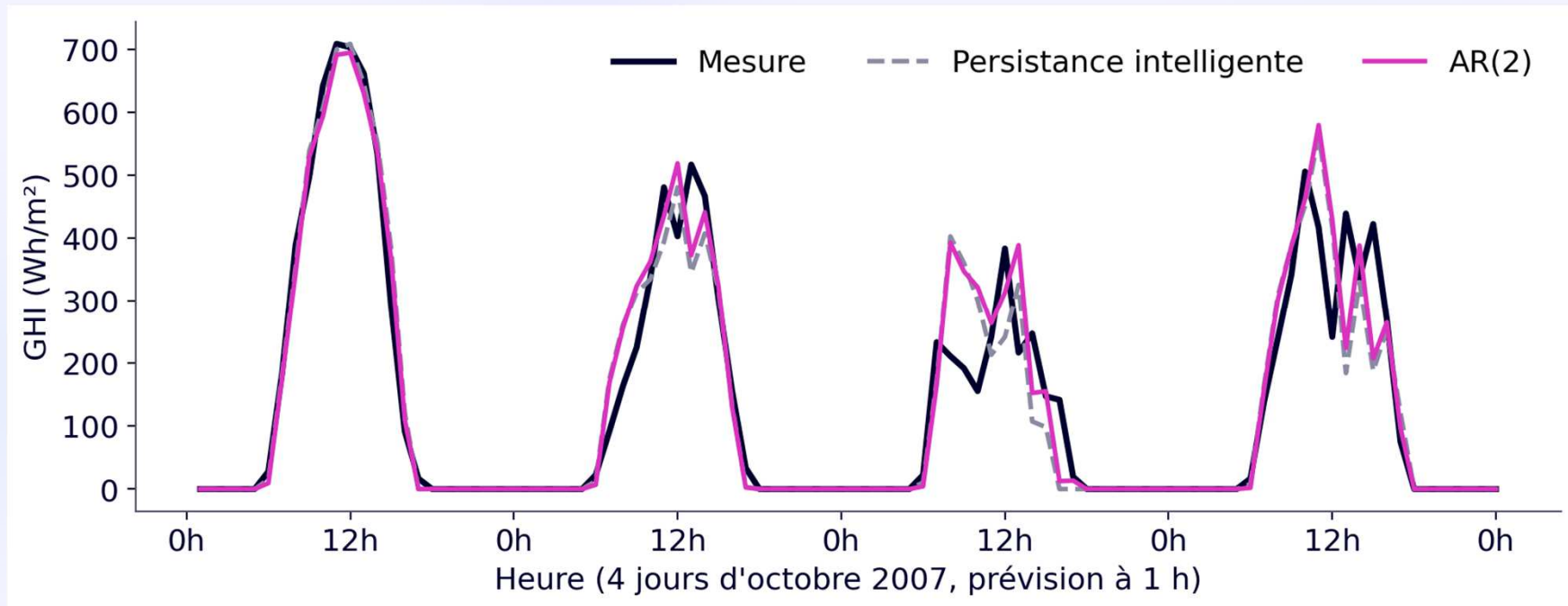
Valeurs en Wh/m² ; 43 Wh/m² = moyenne 2007 de la dernière heure de jour.

- **Sans règles** : la persistance recopie la mesure de l'heure d'avant. Sur l'année, chaque mesure sert une fois de mesure et une fois de prévision : les deux sommes sont égales, **nMBE = 0**.
- **Avec la règle 1 (nuit)** : à la première heure de nuit, le Soleil est sous l'horizon, donc la prévision 43 est remplacée par 0. Chaque jour, une prévision d'environ 43 Wh/m² disparaît.
- Sur l'année : $365 \times 43 \approx 15\,642$ Wh/m², plus 1 008 Wh/m² de petites valeurs mesurées de nuit : **16 650 Wh/m²** manquent.
- Rapporté au GHI mesuré sur l'année (1 630 161 Wh/m²) : **1,02 %**, d'où un nMBE jour + nuit de -1,02 %.

Une seule heure par jour, mais elle pèse lourd : la moyenne jour + nuit ne vaut que 186 Wh/m², car elle compte les heures de nuit à zéro.

La règle est physiquement juste (pas de Soleil la nuit), mais elle change le biais.

Prévision et mesure, heure par heure



Ajaccio, octobre 2007 : prévision à 1 h.

Les deux modèles suivent bien les journées régulières et ratent les changements brusques : un nuage qui arrive ne se voit pas dans le passé.

Les limites du modèle AR

L'AR(p) suppose une relation **linéaire** entre le CSI et ses retards. C'est souvent suffisant à court terme, mais pas toujours.

Situations difficiles

- passages rapides ciel clair / ciel couvert (fronts nuageux) ;
- effets de bord de nuage ($GHI > GHI^{cs}$) ;
- ciel très variable (cumulus) ;
- liens non linéaires entre retards.

Sous-ajustement

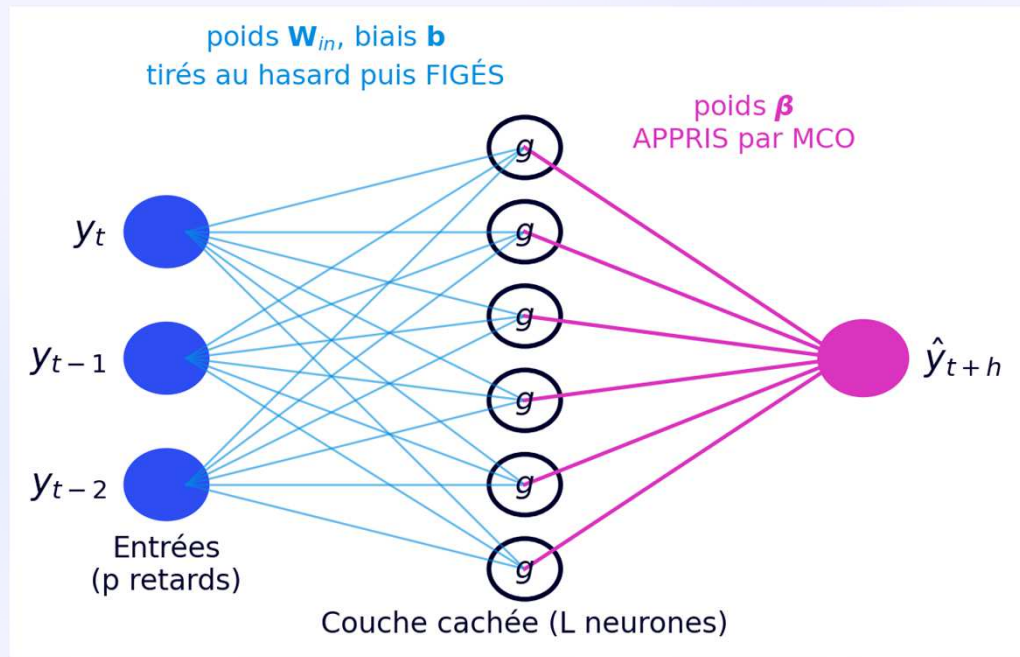
Un modèle linéaire ne capte pas une relation courbe, même avec beaucoup de retards.

Saturation

Au-delà d'un certain p, ajouter des retards n'améliore plus rien (diapo « Choisir p »).

Idée : transformer les entrées de façon non linéaire, puis garder les moindres carrés. C'est l'**ELM**.

Extreme Learning Machine (ELM) : le principe



ELM à une couche cachée (Huang et al., 2006).

- Un réseau de neurones à **une seule couche cachée**.
- Les poids d'entrée $\mathbf{W}_{in}$ et les biais $\mathbf{b}$ sont **tirés au hasard puis figés**.
- Chaque neurone applique une fonction non linéaire g (tanh, sigmoïde...).
- Seuls les poids de sortie β sont appris, **par moindres carrés**.

Pas de rétropropagation du gradient : l'apprentissage se fait **en une seule résolution**.

ELM : la formulation

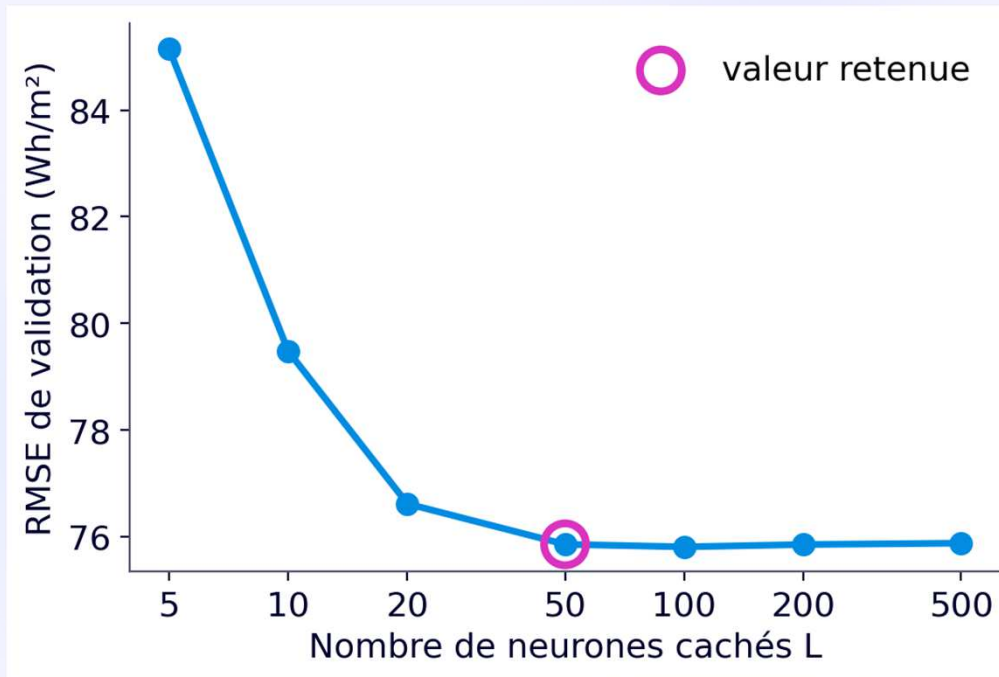
Les p retards de chaque ligne passent par la couche cachée :

$$\mathbf{H} = g(\mathbf{X} \mathbf{W}_{\text{in}} + \mathbf{b}) \quad (N \times L)$$

	Modèle	Solution par moindres carrés	Colonnes utilisées
AR	$\mathbf{y} = \mathbf{X}\mathbf{w} + \boldsymbol{\varepsilon}$	$\hat{\mathbf{w}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$	les retards du CSI centré
ELM	$\mathbf{y} = \mathbf{H}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$	$\boldsymbol{\beta} = (\mathbf{H}^T\mathbf{H} + \alpha\mathbf{I})^{-1}\mathbf{H}^T\mathbf{y}$	les sorties des L neurones

Même équation, seules les colonnes changent. Le petit terme $\alpha\mathbf{I}$ (régression *ridge*) stabilise le calcul quand L est grand ; à Ajaccio, $\alpha = 10^{-3}$.

ELM : choisir le nombre de neurones L



RMSE de validation (2006, heures de jour) selon L, pour $p = 2$.

- $L = 5$: 85,2 Wh/m² → **sous-ajustement**.
- $L = 50$: 75,9 Wh/m².
- Au-delà, plus aucun gain : 75,8 à $L = 200$, 75,9 à $L = 500$ (risque de **surapprentissage**).

Comme pour p : tester quelques valeurs (5 à 500) sur la validation et garder la plus petite presque optimale. Ici **L = 50**.

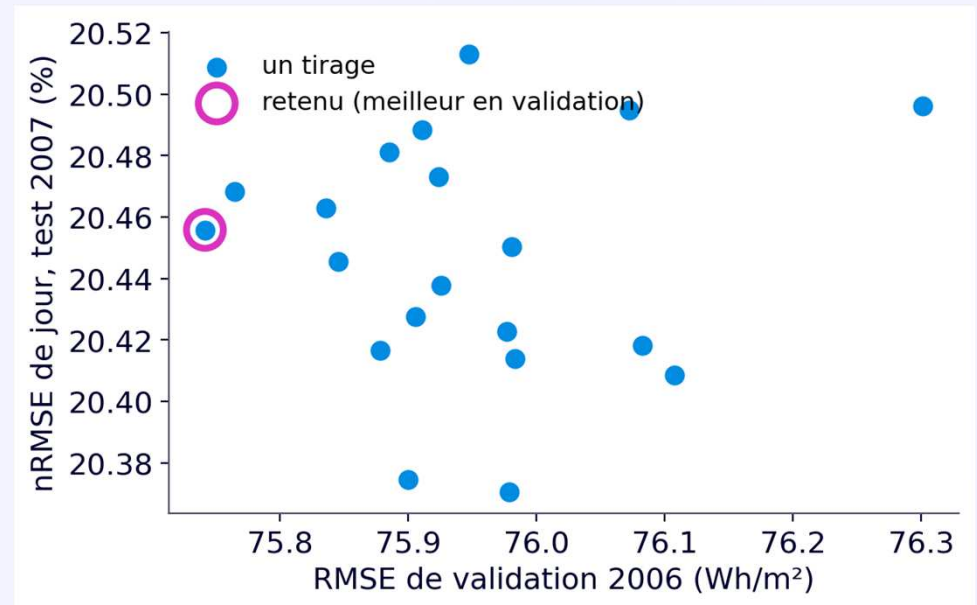
ELM : un résultat qui dépend du tirage

- Les poids d'entrée sont aléatoires : deux entraînements donnent deux modèles différents.
- **Recherche de candidat** : on fait 20 tirages et on garde celui qui a la plus petite erreur sur la **validation** (2006).
- Le test (2007) ne sert qu'à la fin, pour le tirage retenu.

Tirage retenu : nRMSE test **20,46 %**.

Les 20 tirages : de 20,37 à 20,51 % (moyenne 20,45 %).

Le meilleur en validation n'est pas forcément le meilleur en test : ici l'écart est faible, mais il faut le vérifier.



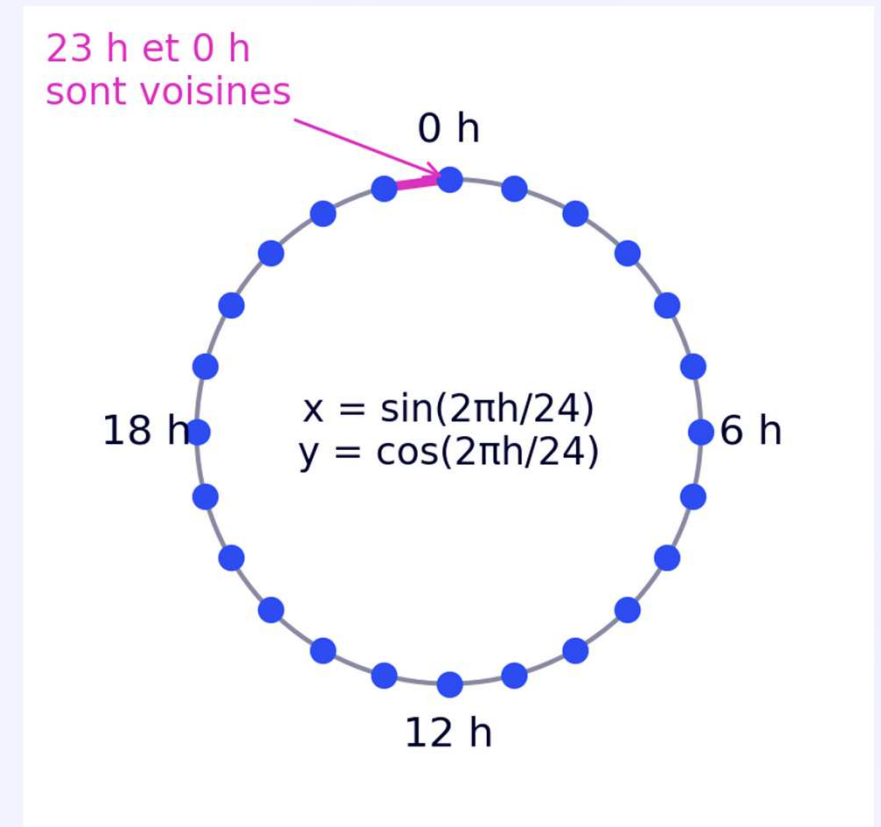
Chaque point : un tirage. Validation 2006 (horizontal), test 2007 (vertical).

Donner l'heure au modèle : encodage sin / cos

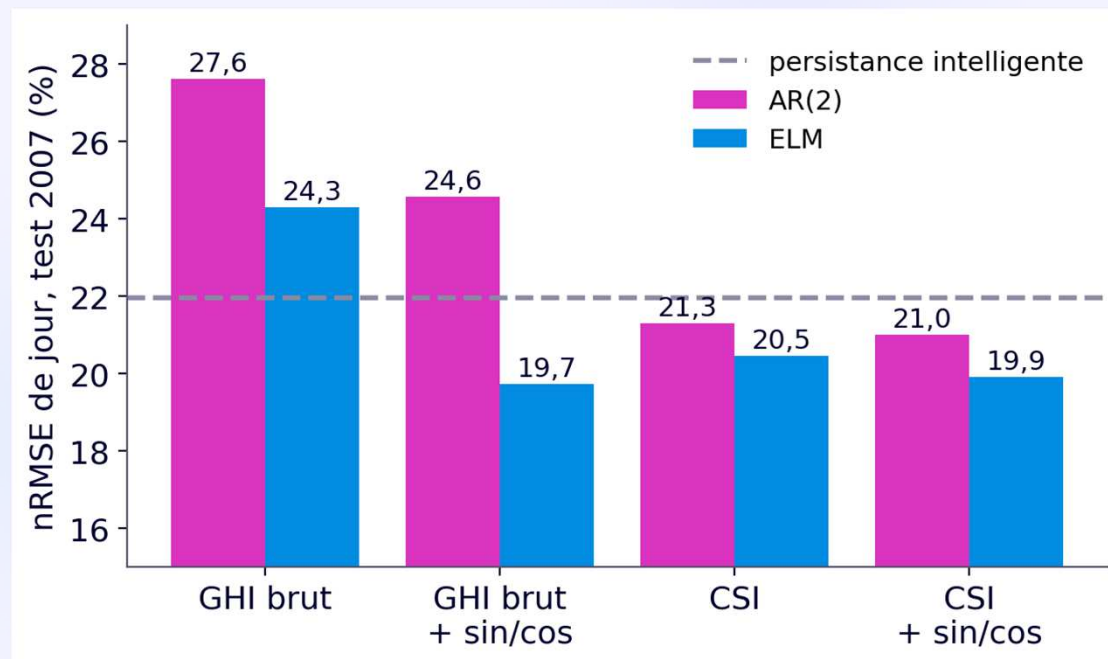
- Idée : ajouter aux retards des **variables de calendrier** connues à l'avance : l'heure h et le jour de l'année j .
- Problème : codée par un nombre, 23 h est « loin » de 0 h, alors que ce sont deux heures voisines.
- Solution : placer l'heure sur un **cercle**, avec deux variables :

$$(\sin(2\pi h / 24) ; \cos(2\pi h / 24)) \text{ et } (\sin(2\pi j / 365) ; \cos(2\pi j / 365))$$

4 colonnes de plus dans **X**. Le modèle « sait » alors quelle heure et quelle saison il est (Voyant et al., 2025a).



Stationnariser ou donner l'heure ? AR et ELM



nRMSE de jour, test 2007, même protocole pour les 8 variantes. Approche « sans ciel clair » : Voyant et al., 2025a.

Entrées	AR(2) nRMSE	AR NICE ^Σ	ELM nRMSE	ELM NICE ^Σ
GHI brut	27,6 %	0,76	24,3 %	0,66
GHI + sin/cos	24,6 %	0,68	19,7 %	0,54
CSI	21,3 %	0,58	20,5 %	0,55
CSI + sin/cos	21,0 %	0,57	19,9 %	0,54

AR (linéaire) : a besoin du CSI. Sur le GHI brut, même avec sin/cos, il fait moins bien que la persistance intelligente.

ELM (non linéaire) : avec sin/cos, il apprend lui-même le cycle du Soleil, à partir du GHI brut.

nRMSE et NICE^Σ donnent le même classement.

AR ou ELM ? Le bilan

Atouts de l'ELM

- apprentissage rapide, une seule résolution ;
- capte des relations non linéaires ;
- même méthode (MCO) que l'AR.

Limites de l'ELM

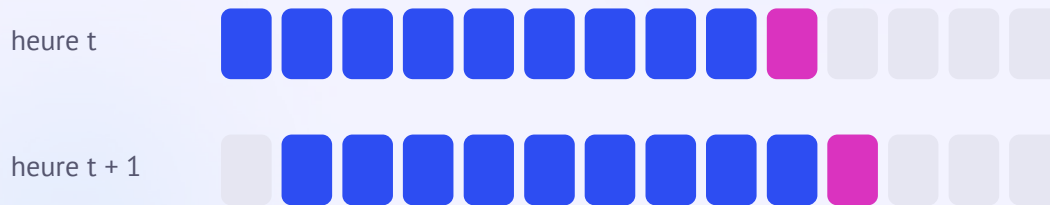
- choix de L à faire ;
- résultat qui varie selon le tirage ;
- moins interprétable que l'AR ;
- surapprend si L devient grand ;
- demande beaucoup de données.

À Ajaccio, à 1 h

- AR(2) : 2 coefficients ; nRMSE 21,3 %.
- ELM : 51 appris + 150 aléatoires ; nRMSE 20,5 %.
- Score SS : +0,03 (AR) contre +0,07 (ELM).

Environ **100 fois plus de coefficients** (appris + aléatoires) pour 4 points de score en plus. Le gain justifie-t-il la complexité ?

L'apprentissage en ligne (1/2) : la fenêtre glissante



bleu : les W heures d'apprentissage ; rose : l'heure à prévoir

- La matrice $\mathbf{X}$ garde **toujours W lignes** : les W dernières heures.
- Chaque heure, on **ajoute** la ligne la plus récente et on **retire** la plus ancienne, puis on ré-estime par ridge :

$$\hat{\mathbf{w}} = (\mathbf{X}^T \mathbf{X} + \alpha \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}$$

- Astuce : pas besoin de tout recalculer. $\mathbf{X}^T \mathbf{X} \leftarrow \mathbf{X}^T \mathbf{X} + \mathbf{x}_{\text{nouveau}} \mathbf{x}_{\text{nouveau}}^T - \mathbf{x}_{\text{ancien}} \mathbf{x}_{\text{ancien}}^T$.
- Même chose pour l'ELM, avec $\mathbf{H}$ à la place de $\mathbf{X}$.

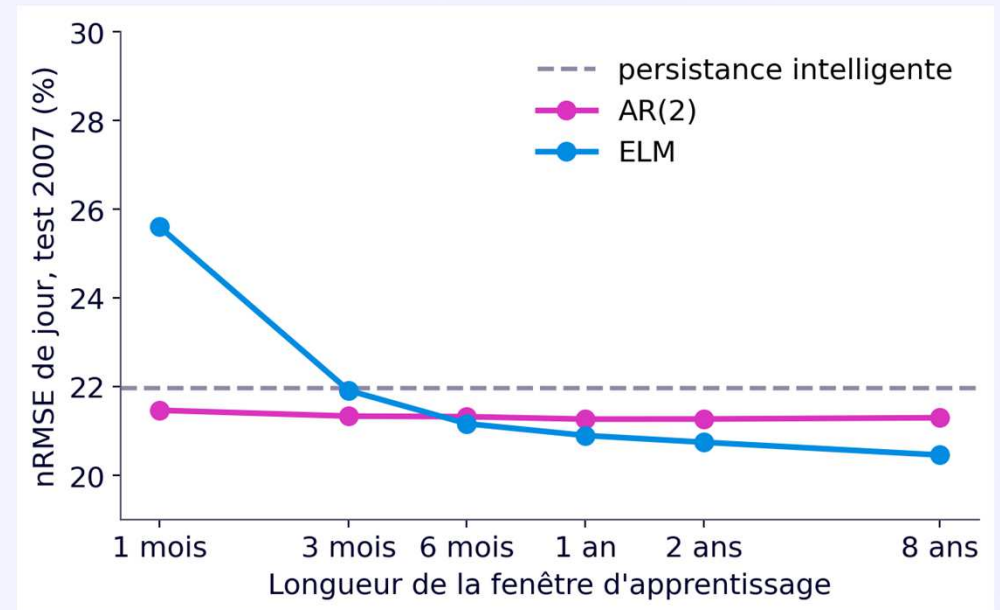
Mise à jour : 11 μs pour l'AR, 38 μs pour l'ELM, par heure.

Utile quand le système **change** (capteur qui dérive, panneau encrassé, nouveau site) : le modèle oublie le passé lointain.

L'apprentissage en ligne (2/2) : quelle longueur de fenêtre ?

- Test 2007 : chaque heure, ridge ré-estimé sur les W dernières heures, de 1 mois à 2 ans.
- Fenêtre courte : suit les changements récents... mais apprend sur peu de données.

Fenêtre W	AR(2)	ELM (L = 50)
1 mois	21,5 %	25,6 %
3 mois	21,3 %	21,9 %
6 mois	21,3 %	21,2 %
1 an	21,3 %	20,9 %
2 ans	21,3 %	20,7 %
8 ans (fixe)	21,3 %	20,5 %



nRMSE de jour sur 2007 selon la longueur de la fenêtre.

L'AR (2 coefficients) marche même sur 1 mois. L'ELM (51 coefficients) **se dégrade nettement** sur fenêtre courte : plus un modèle a de paramètres, plus il lui faut de données.

La frugalité

Un modèle **frugal** obtient une bonne précision avec peu de calcul, peu de données et peu de maintenance.

Modèle	Coefficients appris	Calcul d'apprentissage	Données nécessaires	nRMSE jour (test)
Persistance intelligente	0	aucun	1 heure	22,0 %
AR(2)	2	6 ms	1 mois suffit	21,3 %
ELM (L = 50)	51 (+ 150 aléatoires)	58 ms par tirage	plusieurs années	20,5 %
AR, fenêtre glissante 1 mois	2	11 μ s par heure	720 heures	21,5 %

Temps mesurés sur un ordinateur portable : seuls les ordres de grandeur comptent.

La bonne question n'est pas « quel est le meilleur modèle ? » mais « **le gain justifie-t-il le coût ?** » (calcul, données, maintenance, surveillance).

Merci à vous de jouer : le projet

Réseau et unités

Prévoir le GHI, c'est prévoir la production.

Rayonnement

Direct, diffus ; position du Soleil ; nuages.

Séries temporelles

Composantes, horizon, AR(1).

Ciel clair et CSI

Solis ; CSI = GHI – GHI^{cs}, centré.

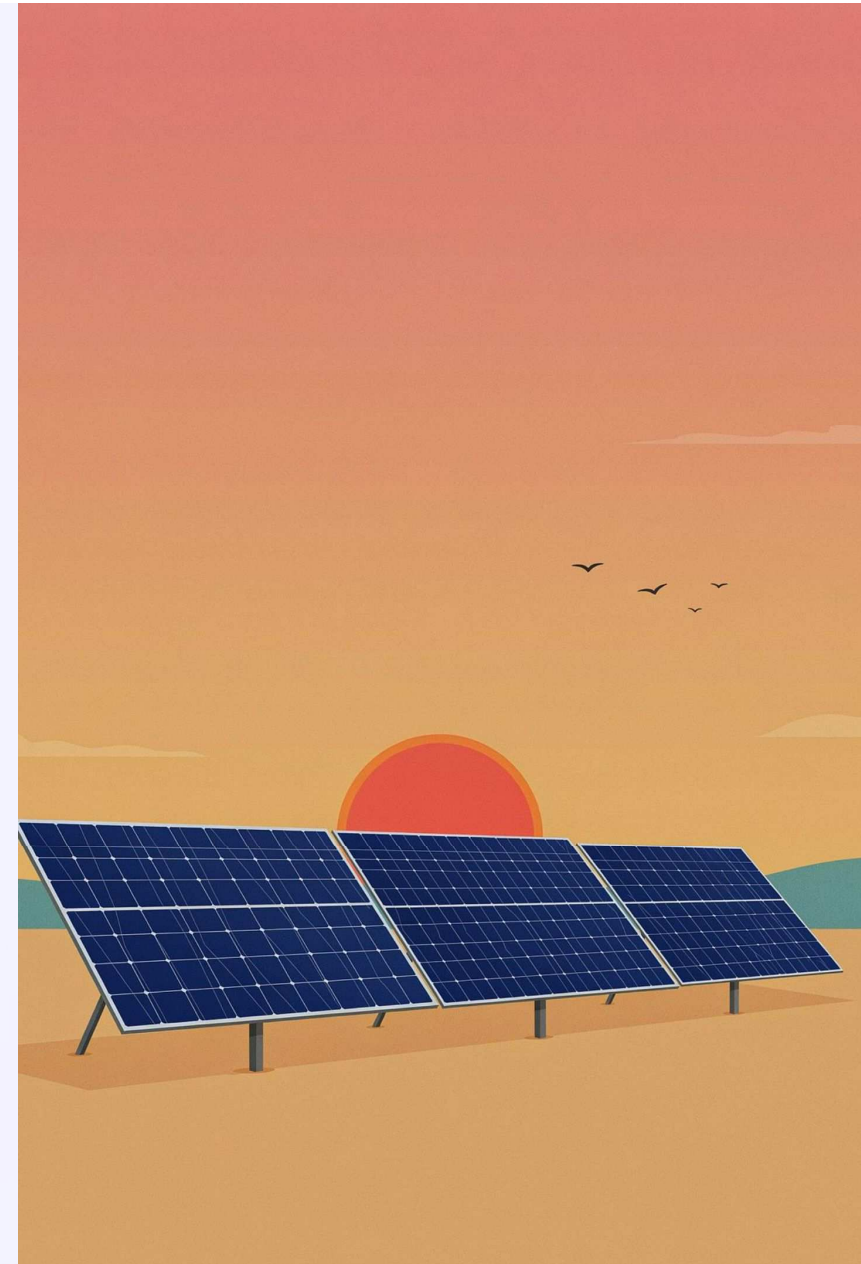
AR et MCO

$\hat{w} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$, choix de p.

Évaluation, ELM, en ligne

Jour / jour + nuit, fenêtre glissante, frugalité.

Cyril Voyant – cyril.voyant@minesparis.psl.eu – Mines Paris – PSL, Centre O.I.E.



Glossaire

Sigle	Signification
AR(p)	modèle autorégressif d'ordre p
ACF / PACF	autocorrélation / autocorrélation partielle
AIC / BIC	critères d'information d'Akaike / bayésien
AOD	épaisseur optique des aérosols
CSI	indice de ciel clair (clear-sky index)
DHI	rayonnement diffus horizontal
DNI	rayonnement direct normal
ELM	Extreme Learning Machine
GHI	rayonnement global horizontal
MAE	erreur absolue moyenne
MCO	moindres carrés ordinaires (OLS en anglais)

Sigle	Signification
NWP	prévision numérique du temps (modèle météo)
nRMSE	RMSE normalisé par le GHI moyen
RMSE	racine de l'erreur quadratique moyenne
MBE	erreur moyenne (biais)
sCV	coefficient de variation stochastique
F	forecastabilité (prévisibilité)
NICE	score relatif à la persistance simple
sin / cos	encodage circulaire de l'heure et du jour
SS	score de compétence (skill score)
STEP	station de transfert d'énergie par pompage

Références

- Huang G.-B., Zhu Q.-Y., Siew C.-K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, 70, 489–501. doi:10.1016/j.neucom.2005.12.126
- Ineichen P. (2008). A broadband simplified version of the Solis clear sky model. *Solar Energy*, 82(8), 758–762. doi:10.1016/j.solener.2008.02.009
- Voyant C. et al. (2017). Machine learning methods for solar radiation forecasting: a review. *Renewable Energy*, 105, 569–582. doi:10.1016/j.renene.2016.12.095
- Voyant C. et al. (2022). Benchmarks for solar radiation time series forecasting. *Renewable Energy*, 191, 747–762. doi:10.1016/j.renene.2022.04.065
- Voyant C. et al. (2026). Stochastic coefficient of variation: assessing the variability and forecastability of solar irradiance. *Renewable Energy*, 256, 123913. doi:10.1016/j.renene.2025.123913
- Voyant C., Despotovic M., Notton G., Saint-Drenan Y.-M., Asloune M., Garcia-Gutierrez L. (2025a). On the importance of clearsky model in short-term solar radiation forecasting. *Solar Energy*, 289, 113490. doi:10.1016/j.solener.2025.113490
- Voyant C. et al. (2025b). NICEk metrics: unified and multidimensional framework for evaluating deterministic solar forecasting accuracy. *Sustainable Energy Technologies and Assessments*, 81, 104588. doi:10.1016/j.seta.2025.104588
- Hyndman R. J., Athanasopoulos G. (2021). *Forecasting: Principles and Practice*, 3e éd. OTexts. otexts.com/fpp3
- Box G. E. P., Jenkins G. M., Reinsel G. C., Ljung G. M. (2015). *Time Series Analysis: Forecasting and Control*, 5e éd. Wiley.
- Données : Météo-France, station d'Ajaccio. Figure du réseau intelligent : © EPRI, 2009.

Annexe A — Stationnarité au sens large

Un processus (z_t) est **stationnaire au sens large** si :

$$E[z_t] = \mu \text{ (constante)} \quad \text{et} \quad \text{Cov}(z_t, z_{t+k}) = \gamma(k) \text{ (ne dépend que du décalage } k)$$

En pratique

- la moyenne est constante dans le temps ;
- la variance est finie et constante ;
- la dépendance entre deux instants ne dépend que de leur écart.

Vérification simple : tracer la moyenne et l'écart-type mois par mois (diapo « Le CSI est-il stationnaire ? »). Les tests statistiques existent mais ne sont pas nécessaires ici.

Annexe B — La meilleure prévision possible

$$\hat{y}_{t+h} = E[y_{t+h} \mid y_t, y_{t-1}, \dots, y_{t-p+1}]$$

- Parmi toutes les prévisions construites à partir du passé, celle qui minimise l'erreur quadratique moyenne est l'**espérance conditionnelle** : la moyenne des futurs possibles, sachant le passé.
- On écrit alors $y_{t+h} = \hat{y}_{t+h} + \varepsilon_{t+h}$: partie prévisible + erreur, avec une erreur de moyenne nulle.
- Un modèle AR(p) est une approximation **linéaire** de cette espérance ; l'ELM, une approximation **non linéaire**.

Si les erreurs du modèle forment un bruit blanc (ACF des erreurs non significative), il ne reste plus d'information exploitable dans le passé.

Annexe C — Choisir p avec AIC ou BIC

$$\text{AIC} \approx N \cdot \ln(\hat{\sigma}^2) + 2(p+1)$$

$$\text{BIC} \approx N \cdot \ln(\hat{\sigma}^2) + (p+1) \cdot \ln N$$

- $\hat{\sigma}^2$: variance des erreurs sur l'apprentissage ;
- on garde le p qui donne le plus petit critère.

- Le premier terme baisse quand le modèle s'ajuste mieux ; le second **pénalise** chaque coefficient ajouté.
- BIC pénalise plus fort qu'AIC : il choisit des modèles plus petits.
- Avantage : pas besoin de période de validation. Inconvénient : suppose des erreurs gaussiennes.