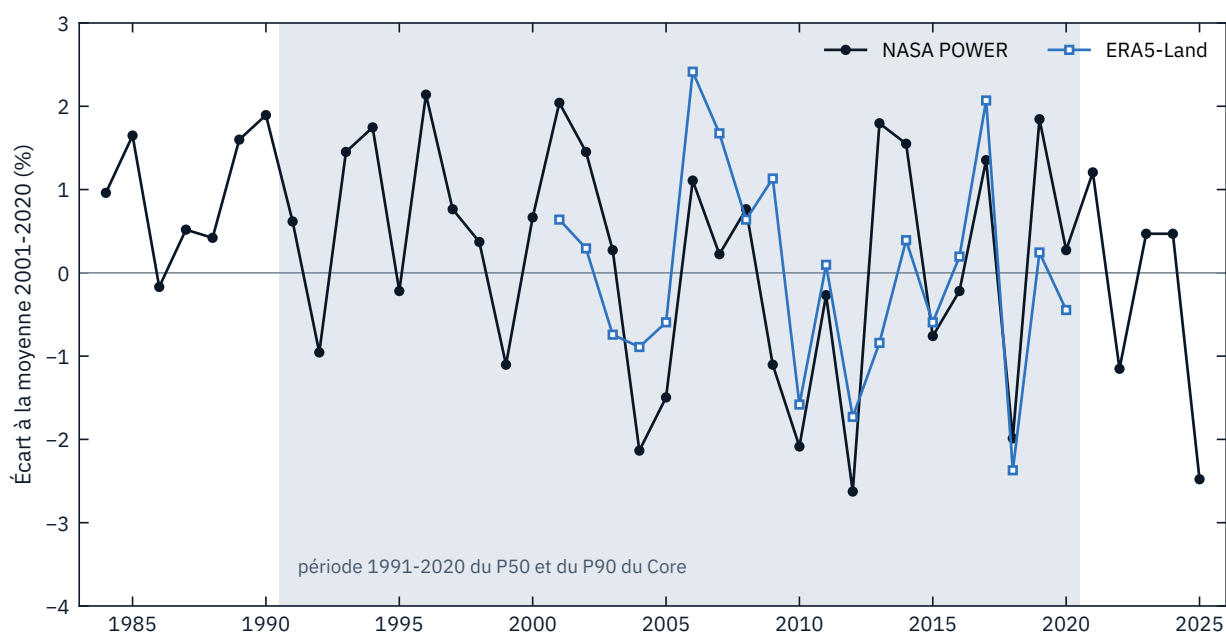




L'incertitude

Contrôler les séries, distinguer écart et erreur, et chiffrer ce qu'un productible peut manquer



Rayonnement global annuel à Kankan, écart à la moyenne 2001-2020, selon NASA POWER (1984-2025) et ERA5-Land (2001-2020). Section 4.

Siba Kalivogui

Ingénieur efficacité énergétique

Énergie solaire | Octobre 2026 | v2.0 | CC BY-NC-SA 4.0

Données : RDS Guinée 2.0.1 et ESMAP/WAPP

kumascience.com

Trois sources, trois chiffres

Sur les 23 mois complets mesurés à Kankan entre octobre 2021 et octobre 2023, NASA POWER donne un rayonnement global supérieur de 3,0 % à celui de la station, ERA5-Land de 3,9 % et SARAH-3 de 11,2 % (guide 3). À Tarambaly, sur 23 mois aussi, les mêmes sources donnent +0,3, +0,6 et +7,0 %. Une série décalée de 11 % passe pourtant le contrôle de qualité presque sans alerte (section 1). La station elle-même n'est exacte qu'à 3 % près. Ce guide montre ce que le contrôle de qualité garantit, ce qu'un écart entre sources indique, et comment chiffrer de combien un productible peut se tromper.

Ce guide suppose acquis les guides 1 à 6. Données : stations ESMAP/WAPP de Kankan et de Tarambaly [4], niveau de confiance **A**; NASA POWER, ERA5-Land et SARAH-3 par le RDS Guinée 2.0.1 [6], niveau **B**. Notations et unités : guide 0, section 9. Licence CC BY-NC-SA 4.0.

Table des matières

1 Le contrôle qualité	3
2 Les niveaux de confiance	4
3 Écart entre sources et erreur	5
4 La variabilité interannuelle	5
5 Le budget d'incertitude	6
6 Exercices	9
7 Résumé	10
Références	11

1 Le contrôle qualité écarte les valeurs impossibles, pas les valeurs biaisées

Les séries horaires du Core et du RDS passent un contrôle de qualité automatique, faute de pouvoir relire des dizaines de millions de valeurs. Les tests viennent du réseau de référence BSRN [1]. Le premier borne le rayonnement global par deux courbes qui dépendent du cosinus de l'angle zénithal μ_0 et de l'éclairement extraterrestre normal du jour S_a (guide 3) :

$$\text{limite physiquement possible : } 1,5 S_a \mu_0^{1,2} + 100 \quad (1)$$

$$\text{limite extrêmement rare : } 1,2 S_a \mu_0^{1,2} + 50 \quad (2)$$

en W/m^2 . Une valeur au-dessus de la première est rejetée ; au-dessus de la seconde, elle est signalée pour examen. D'autres tests comparent les composantes entre elles (la fermeture du guide 3) et cherchent les valeurs figées ou les sauts. Dans le RDS, 9 201 valeurs horaires sur 63,5 millions ont été rejetés, soit 0,014 % [6].

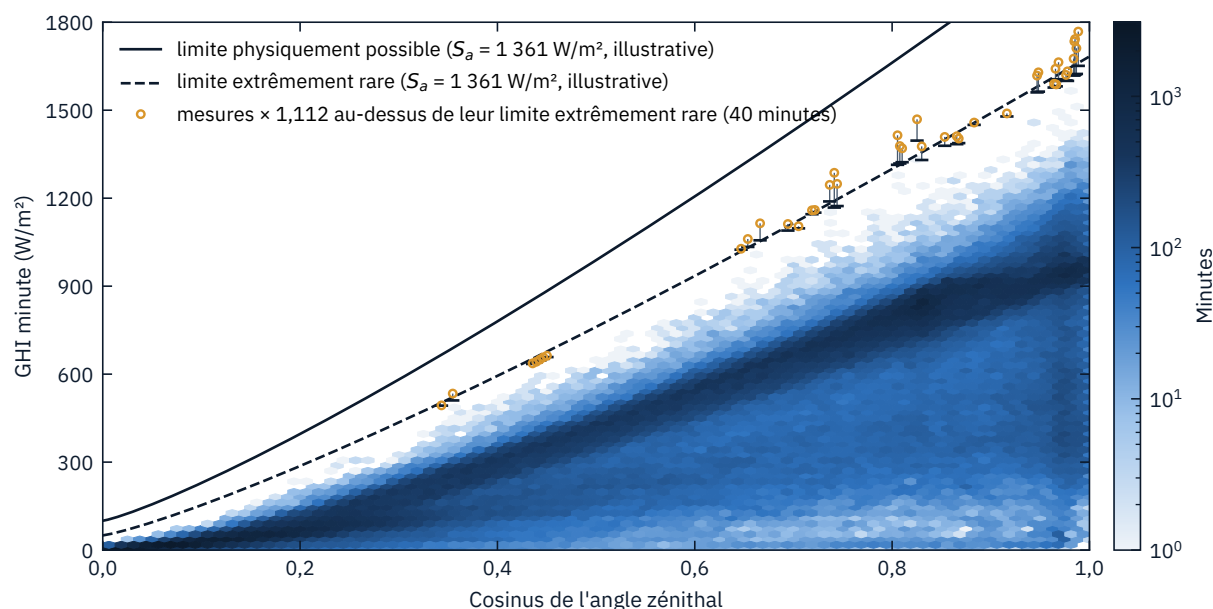


Figure 1 : Rayonnement global minute mesuré à Kankan (densité, 526 052 minutes de jour, octobre 2021 à octobre 2023). Les deux courbes, tracées pour $S_a = 1\,361 \text{ W/m}^2$, sont illustratives ; le comptage utilise S_a de chaque jour. Cercles : minutes qui dépassent la limite extrêmement rare quand toutes les mesures sont multipliées par 1,112 ; le trait sous chaque cercle marque la limite propre à cette minute.

Bornes BSRN sur les mesures et sur une série biaisée

À Kankan, aucune minute ne dépasse la limite extrêmement rare ; le maximum atteint $1\,590 \text{ W/m}^2$. À Tarambaly, 11 minutes la dépassent, lors d'épisodes de renforcement par les nuages (guide 4) ; aucune ne dépasse la limite physiquement possible. Multiplions toutes les mesures de Kankan par 1,112, l'écart de SARAH-3 à la station : 40 minutes sur 526 052 dépassent alors la limite extrêmement rare (0,008 %), aucune la limite physiquement possible. À Tarambaly, la même opération en signale 164 (0,031 %).

Source : fichiers ESMAP/WAPP, minutes de jour | Période : oct. 2021 à oct. 2023 | Niveau de confiance **A**

Une série décalée de 11 % passe donc le contrôle presque sans alerte. Les bornes sont larges parce qu'elles doivent laisser passer les vrais pics comme celui du 18 septembre 2022 à Kankan. Le test de fermeture ne détecte pas mieux ce décalage quand les trois composantes viennent du même modèle, qui les produit cohérentes entre elles.

À retenir

Le contrôle de qualité dit qu'une valeur est possible. Pour savoir si une série est juste, il faut la comparer à une référence mesurée.

2 Un niveau de confiance classe l'origine d'une donnée, pas son erreur

Le Core attribue à chaque valeur un niveau A, B ou C selon la façon dont elle a été obtenue. Le contrôle de qualité automatique ne fait jamais monter une valeur au niveau A.

Niveau	Définition du Core	Exemples de cette collection
A	mesure directe par instrument de référence	stations ESMAP/WAPP de Kankan et de Tarambaly
B	donnée modélisée ou interpolée selon une méthode documentée	NASA POWER, ERA5-Land, SARAH-3, CAMS; séries horaires contrôlées du RDS
C	estimation d'expert sans validation instrumentale	pertes d'onduleur, de câbles et de disponibilité du guide 6

Table 1 : Niveaux de confiance du Core et exemples.

Le niveau ne dit pas de combien une valeur se trompe. Au niveau B, NASA POWER s'écarte de 0,3 % de la station de Tarambaly en moyenne mensuelle, et SARAH-3 de 11,2 % de celle de Kankan. Au niveau A, la mesure a elle aussi une incertitude.

Incertitude des stations

L'exploitant estime l'incertitude de mesure à $\pm 3,0$ % pour le rayonnement global quand l'angle zénithal est compris entre 30 et 60°, et à $\pm 2,0$ % pour le rayonnement direct normal, avec un intervalle de confiance de 95 %. À Kankan, après environ un an, une comparaison sur site avec des instruments de référence a donné des écarts moyens de 0,4 % sur le global et de 0,8 % sur le direct [4].

Source : rapports de mesure de CSP Services | Période : oct. 2021 à oct. 2023 | Niveau de confiance **A**

3 Aux deux stations, l'écart entre NASA POWER et SARAH-3 vient surtout de SARAH-3

Sans station, on compare les sources entre elles. Le guide 3 a cartographié l'écart entre NASA POWER et SARAH-3 sur les localités du RDS : de 2,3 à 12,8 %, 6,3 % en médiane. Aux deux stations, on peut vérifier ce que cet écart représente.

	Kankan	Tarambaly (maille de Labé)
NASA POWER face à la station	+3,0 %	+0,3 %
ERA5-Land face à la station	+3,9 %	+0,6 %
SARAH-3 face à la station	+11,2 %	+7,0 %
SARAH-3 face à NASA POWER	+7,9 %	+6,7 %

Table 2 : Rayonnement global moyen des mois à 28 jours complets ou plus (23 mois par station, octobre 2021 à octobre 2023), rapporté à la station.

Aux deux stations, l'écart entre les deux sources est porté pour l'essentiel par SARAH-3. Si l'on avait pris cet écart pour l'incertitude de NASA POWER, on l'aurait surestimée ; si l'on avait choisi SARAH-3, on se serait trompé de 7 à 11 %. Deux stations ne suffisent pas pour dire que NASA POWER est la meilleure source partout en Guinée : elles le désignent pour ces deux sites, et une station dans une autre région reste nécessaire pour le vérifier.

Erreur fréquente : l'écart pris pour une incertitude

Un écart entre sources signale les sites où il faut se méfier. Il ne dit pas laquelle se trompe, ni de combien : deux sources peuvent partager le même biais et s'accorder à tort. Seule une mesure de référence chiffre une erreur.

4 D'une année à l'autre, le rayonnement annuel varie d'environ 1,5 % (écart-type)

La variabilité interannuelle se calcule sur les totaux annuels d'une longue série. Le Core en tire un P50 et un P90 de l'irradiation annuelle, sur les 30 années 1991-2020 de NASA POWER. Le PX est la valeur dépassée X % des années : le P90 est le 10^e percentile.

Variabilité interannuelle du rayonnement global

Sur les 34 localités du RDS et la période 2001-2020, le coefficient de variation annuel va de 1,21 à 1,65 % avec NASA POWER (médiane 1,50 %) et de 1,14 à 1,87 % avec ERA5-Land (médiane 1,51 %). La corrélation entre les anomalies des deux sources, année par année, va de 0,38 à 0,78 (médiane 0,62 ; 0,61 à Kankan). À Kankan sur 1991-2020 avec NASA POWER : moyenne 2 039 kWh/m², écart-type 28,1 kWh/m² (1,38 %), P50 2 042 et P90 1 995 kWh/m², soit 2,3 % sous le P50 ; l'année la plus faible, 2012, donne 1 982 kWh/m².

Source : RDS Guinée 2.0.1, séries mensuelles | Période : 1984-2025 (NASA), 2001-2020 (ERA5) | Niveau de confiance **B**

Les deux sources s'accordent sur l'amplitude de la variabilité, pas toujours sur les années faibles. Le P90 empirique repose sur les trois ou quatre années les plus faibles

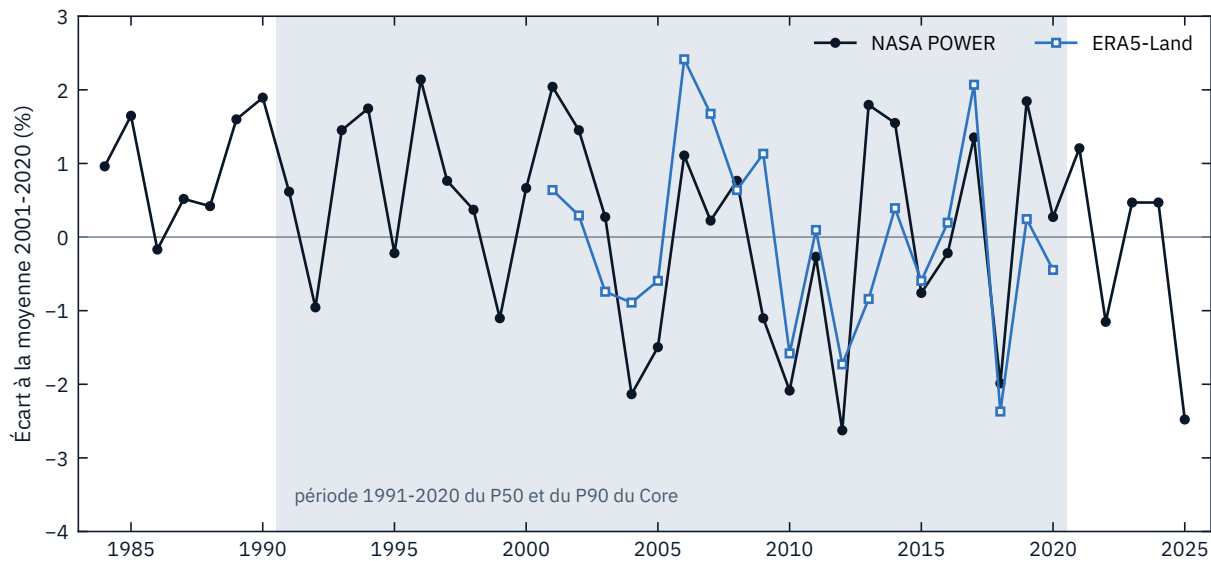


Figure 2 : Rayonnement global annuel à Kankan, écart à la moyenne 2001-2020 de chaque source. Totaux annuels pondérés par la longueur des mois.

des trente : en rééchantillonnant ces trente années (bootstrap), il se situe entre 1 991 et 2 014 kWh/m² dans 90 % des cas. La moyenne 2001-2020 de NASA POWER est aussi inférieure de 0,7 % en médiane à celle de 1984-2000, période couverte par des satellites plus anciens [6]; l'analyse ne permet pas de dire si ce décalage est climatique ou instrumental, mais il est du même ordre que l'écart-type interannuel.

Erreur fréquente : le P90 interannuel pris pour un P90 complet

Le P90 du Core ne contient que la variabilité d'une année à l'autre. Il ne contient pas l'erreur de la source, qui touche toutes les années de la même façon. Ce n'est pas le P90 qu'exige le financement d'un projet [3].

5 Sans station au sol, la ressource domine le budget d'incertitude du productible de Kankan

5.1 Combiner les composantes

Quand les composantes d'erreur sont indépendantes, leurs écarts-types relatifs se combinent en quadrature [2] :

$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2} \quad (3)$$

Si l'erreur totale suit une loi normale, le P90 d'une année vaut

$$P_{90} = P_{50} (1 - 1,282 \sigma) \quad (4)$$

Les composantes ne se comportent pas de la même façon dans le temps. L'erreur de la source, du modèle thermique ou de la transposition se répète chaque année : elle est systématique. La variabilité du climat change d'une année à l'autre ; sur la moyenne de N années, son écart-type est divisé par $\sqrt{N}$.

5.2 Le champ de Kankan

On part du productible du guide 6 pour un champ au sol ventilé nettoyé une fois par mois : $P_{50} = 1\,548 \text{ kWh/kWc/an}$. Le tableau 3 donne l'écart-type retenu pour chaque composante, sa justification et son statut. Deux cas sont comparés : un bureau d'études qui n'a que NASA POWER, et un bureau d'études qui dispose de deux ans de station sur le site.

Les deux cas partent du même P50, calculé au guide 6 avec le rayonnement de la station. Ce choix isole l'effet de l'incertitude à valeur centrale fixée. Il ne reproduit pas le P50 qu'aurait établi un bureau d'études disposant du seul NASA POWER. Sur les mêmes jours, le GHI de NASA POWER dépasse celui de la station de 2,7 % à Kankan (guide 6) ; à pertes égales, ce P50 aurait été plus élevé d'environ autant.

Composante	σ	Justification	Statut
Ressource, sans station	4,0 %	hypothèse de travail. Repères : écarts de NASA POWER aux deux stations de +0,3 et +3,0 % ; pour son propre modèle, le Global Solar Atlas annonce une incertitude attendue pouvant atteindre ± 8 % sur le GHI dans certains contextes [5]	hypothèse
Ressource, avec station	1,8 %	± 3 % à 95 % pour la mesure, et deux années pour représenter le long terme ($1,38/\sqrt{2}$)	estimé
Variabilité interannuelle	1,38 %	NASA POWER 1991-2020 à Kankan	calculé (B)
Encrassement	2,0 %	perte annuelle observée de 8,0 % puis 4,8 % d'une année à l'autre, au même rythme de nettoyage	estimé
Autres postes	1,5 %	plages usuelles des hypothèses du guide 6	hypothèse
Transposition	1,0 %	écart isotrope/Perez de 0,9 à 1,2 %, effet d'Erbs de 0,3 % (guide 5)	estimé
Modèle thermique	1,0 %	perte de 8,3 à 9,9 % selon le modèle pour une même pose (guide 6)	estimé

Table 3 : Budget d'incertitude du productible annuel à Kankan. Les écarts-types sont relatifs au productible.

Le P90 du champ de Kankan sans station au sol

Étape 1, les carrés. On élève au carré chaque écart-type du tableau 3, en pour cent : ressource sans station $4,0^2 = 16,00$; variabilité interannuelle $1,38^2 = 1,90$; encrassement $2,0^2 = 4,00$; autres postes $1,5^2 = 2,25$; transposition $1,0^2 = 1,00$; modèle thermique $1,0^2 = 1,00$.

Étape 2, la somme et la racine. La somme vaut 26,15. Selon l'équation (3), $\sigma = \sqrt{26,15} = 5,11$ %.

La ressource pèse $16,00/26,15 = 61$ % de cette somme : c'est sa part de la variance.

Étape 3, le P90 d'une année. Selon l'équation (4) :

$$P_{90} = 1\,548 \times (1 - 1,282 \times 0,0511) = 1\,548 \times 0,9345 = 1\,447 \text{ kWh/kWc/an.}$$

Étape 4, le P90 de la moyenne de dix ans. Seule la variabilité interannuelle se moyenne : son carré devient $1,90/10 = 0,19$. Les autres termes restent, $26,15 - 1,90 = 24,25$, d'où $\sigma = \sqrt{24,25 + 0,19} = 4,94$ % et $P_{90} = 1\,548 \times (1 - 1,282 \times 0,0494) = 1\,450 \text{ kWh/kWc/an.}$

P50 et P90 du champ de Kankan

P50 commun aux deux cas : 1 548 kWh/kWc/an (guide 6). Sans station, $\sigma = 5,11\%$ pour une année : le P90 d'une année vaut 1 447 kWh/kWc/an, et celui de la moyenne de dix ans 1 450. Avec deux ans de station, $\sigma = 3,65\%$: P90 de 1 475 pour une année et de 1 480 pour la moyenne de dix ans. La composante systématique vaut 4,92 % sans station et 3,38 % avec.

Source : chiffres des guides 5 et 6, composantes du tableau 3 | Période : oct. 2021 à oct. 2023 | Niveau de confiance 

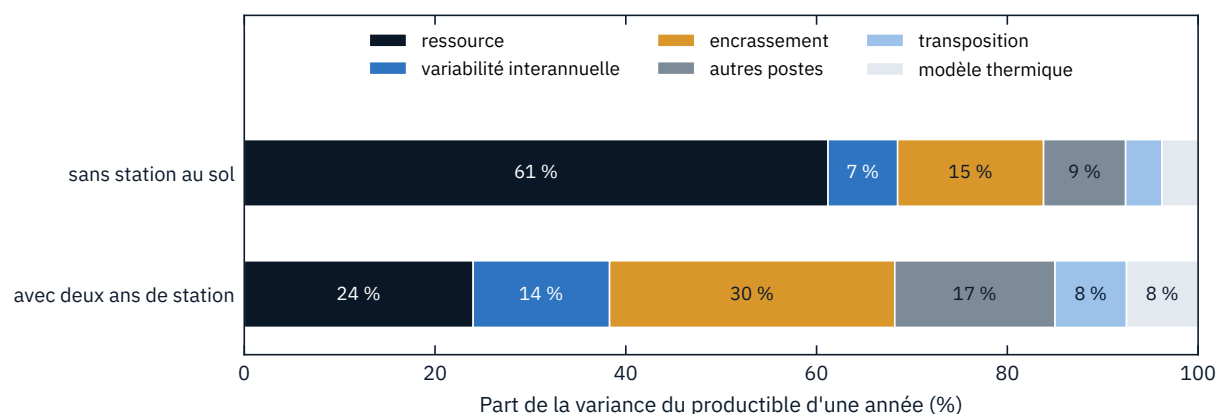


Figure 3 : Part de chaque composante dans la variance du productible d'une année à Kankan, selon le tableau 3.

Sans station, la ressource représente 61 % de la variance ; aucun autre travail sur le modèle ne réduit beaucoup l'incertitude tant qu'elle reste à ce niveau. Avec deux ans de station, elle tombe à 24 %, et l'encrassement devient le premier poste (30 %). Le P90 de dix ans est à peine plus élevé que celui d'une année, parce que la part systématique domine : le climat de Kankan varie peu, alors que l'erreur de la source se répète chaque année. La variabilité interannuelle seule aurait placé le P90 d'une année à 1,8 % sous le P50 (2,3 % avec le percentile empirique du Core), contre 6,5 % avec toutes les composantes.

Ces chiffres dépendent des hypothèses du tableau, en particulier de l'écart-type de la ressource sans station. Ils montrent la structure de l'incertitude et l'ordre des priorités ; ils ne remplacent pas une étude de ressource réalisée selon les règles de l'art [3].

En pratique

Pour chaque composante d'un budget, écrire sa valeur, d'où elle vient et son statut (mesurée, estimée ou supposée). Réduire d'abord la composante dominante : sans station, c'est la ressource ; à Kankan avec une station, c'est le plan de nettoyage.

6 Exercices

Exercice 1 (application)

Pour $\mu_0 = 0,9$ et $S_a = 1\,361 \text{ W/m}^2$, calculer les deux bornes BSRN. Une minute mesurée à $1\,450 \text{ W/m}^2$ est-elle rejetée ?

Corrigé

Étape 1, la puissance : $0,9^{1,2} = e^{1,2 \ln 0,9} = e^{-0,126} = 0,881$.

Étape 2, la limite extrêmement rare : $1,2 \times 1\,361 \times 0,881 = 1\,439$, plus 50, soit $1\,489 \text{ W/m}^2$.

Étape 3, la limite physiquement possible : $1,5 \times 1\,361 \times 0,881 = 1\,799$, plus 100, soit $1\,899 \text{ W/m}^2$.

Étape 4, le verdict : $1\,450 \text{ W/m}^2$ est sous les deux limites. La valeur n'est ni rejetée ni signalée.

Exercice 2 (application)

Retrouver l'écart-type total du cas « avec station » du tableau 3, puis le P90 d'une année.

Corrigé

On suit la méthode de l'exemple, en remplaçant la ressource sans station (4,0 %) par la ressource avec station. Le tableau l'arrondit à 1,8 %; sa valeur non arrondie vaut $\sqrt{1,5^2 + (1,38/\sqrt{2})^2} = \sqrt{2,25 + 0,95} = 1,79 \%$, en prenant pour la mesure la moitié de $\pm 3 \%$ à 95 %.

Étape 1, les carrés : $1,79^2 = 3,20$; $1,38^2 = 1,90$; $2,0^2 = 4,00$; $1,5^2 = 2,25$; $1,0^2 = 1,00$; $1,0^2 = 1,00$.

Étape 2, la somme et la racine : $3,20 + 1,90 + 4,00 + 2,25 + 1,00 + 1,00 = 13,35$, et $\sigma = \sqrt{13,35} = 3,654 \%$, arrondi à 3,65 %.

Étape 3, le P90 : $1,282 \times 0,03654 = 0,04684$, et $P_{90} = 1\,548 \times (1 - 0,04684) = 1\,548 \times 0,95316 = 1\,475 \text{ kWh/kWc/an}$.

Exercice 3 (application)

Avec un écart-type interannuel de 1,38 %, de combien le P90 est-il sous le P50 pour une année, puis pour la moyenne de dix ans, si l'on ignore les autres composantes ?

Corrigé

Une année : avec une seule composante, $\sigma = 1,38 \%$, et l'écart entre P50 et P90 vaut $1,282 \times 1,38 = 1,77 \%$.

Dix ans : l'écart-type de la moyenne de dix années vaut $1,38/\sqrt{10} = 1,38/3,16 = 0,44 \%$, et l'écart entre P50 et P90, $1,282 \times 0,44 = 0,56 \%$. On obtient le même résultat en divisant directement 1,77 par $\sqrt{10}$.

Exercice 4 (application)

Deux sources donnent pour un site 1 950 et 2 080 kWh/m² par an. Peut-on en déduire l'incertitude de la première ?

Corrigé

Non. L'écart de 6,7 % signale un site à examiner. Il ne dit pas laquelle des deux sources se trompe ; à Kankan et à Tarambaly, c'est surtout celle qui donnait le chiffre le plus élevé.

Exercice 5 (Python)

Avec le fichier `nasa-power/mensuel_1981_2025.csv` du RDS, calculer les totaux annuels 1991-2020 d'une localité, son P50 et son P90 (`numpy.percentile` à 50 et à 10), puis l'intervalle du P90 par bootstrap sur 5 000 tirages.

Corrigé

Pondérer chaque moyenne mensuelle par le nombre de jours du mois. À Kankan : P50 2 042 et P90 1 995 kWh/m²; l'intervalle à 90 % du P90 va de 1 991 à 2 014 kWh/m².

7 Résumé

1. Le contrôle de qualité écarte les valeurs impossibles; une série décalée de 11 % le passe presque sans alerte.
2. Un écart entre sources signale un site à examiner; à Kankan et à Tarambaly, il venait surtout de SARAH-3.
3. La variabilité interannuelle est faible (écart-type d'environ 1,5 %); le P90 du Core, qui ne contient qu'elle, n'est pas un P90 complet.
4. À Kankan sans station, la ressource fait 61 % de la variance du productible; avec deux ans de station, l'encrassement passe en tête.

Le guide 8 montre comment une mesure au sol réduit la composante ressource, et jusqu'où ce gain se transporte loin de la station.

Références

- [1] C. N. Long, E. G. Dutton, *BSRN Global Network recommended QC tests, V2.0*, Baseline Surface Radiation Network, 2010.
- [2] JCGM, *Évaluation des données de mesure : guide pour l'expression de l'incertitude de mesure*, JCGM 100 :2008, BIPM, 2008.
- [3] M. Sengupta et al., *Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications*, 3^e édition, rapport NREL/TP-5D00-77635, 2021.
- [4] World Bank, ESMAP, WAPP, *Solar resource measurement campaign in West Africa*, stations de Kankan et de Tarambaly (Guinée), données à la minute et rapports de mesure de CSP Services (2023), licence CC BY 4.0. <https://energydata.info>
- [5] World Bank, Solargis, *Global Solar Atlas, Accuracy*, page d'aide consultée en septembre 2026. <https://globalsolaratlas.info/support/accuracy>
- [6] Kuma Science, *RDS Guinée 2.0.1*, référentiel de données solaires, limites documentées (CAVEATS.md), 2026. <https://doi.org/10.5281/zenodo.21051754>



Kuma Science

Kuma Science construit les référentiels de données énergétiques de l'Afrique de l'Ouest, en commençant par la Guinée. Chaque chiffre publié porte sa source et son niveau de confiance.