

Surveillance photovoltaïque basée sur les données météorologiques : comment BDPV évolue sur AWS

par Lior Perez et David Trebosc | le 15 janvier 2025

L'adoption de l'énergie solaire connaît une croissance sans précédent à l'échelle mondiale, avec un ajout remarquable de 346 GW en 2023, ce qui représente une augmentation de 74 % par rapport à l'année précédente. Cet essor pose un défi crucial aux producteurs d'énergie solaire résidentiels : comment surveiller efficacement leurs installations pour s'assurer d'un rendement optimal et d'un retour sur investissement ? ASSO BDPV (Base de Données PhotoVoltaïque), un organisme français à but non lucratif, a relevé ce défi en créant une plateforme basée sur AWS qui surveille plus de 3 000 installations solaires en temps quasi réel. Leur solution combine des données météorologiques avancées et des algorithmes sophistiqués pour surveiller et optimiser la production d'énergie solaire pour les petits producteurs.

Dans cet article, nous explorons comment BDPV a utilisé les services AWS pour créer une architecture de surveillance évolutive, fiable et rentable. En tant qu'organisation gérée par des bénévoles, elle fait face à des défis techniques uniques. Nous examinons comment ils ont stratégiquement combiné plusieurs services AWS et l'impact réel de leur solution sur la communauté solaire en pleine croissance en France.

Surveillance des installations solaires pour les producteurs individuels

ASSO BDPV a été créée en 2016 pour soutenir les producteurs photovoltaïques résidentiels qui n'ont souvent pas l'expertise nécessaire pour surveiller correctement leurs installations solaires. S'appuyant sur une plateforme Web initialement créée par des bénévoles en 2008, l'organisation a permis aux producteurs de suivre leurs données de production solaire en temps quasi réel et de comparer leurs performances à celles des installations voisines. Cette approche comparative fournit aux producteurs des informations précieuses sur la santé et l'efficacité de leur installation.

Cependant, la production d'énergie solaire est intrinsèquement variable, influencée par de multiples facteurs, notamment les conditions météorologiques et les obstacles physiques qui créent des ombres. Pour améliorer la précision de la surveillance et automatiser la détection des anomalies, BDPV a récemment mené des recherches qui ont mené au développement d'un algorithme sophistiqué. L'algorithme calcule la production théorique à une résolution temporelle fine d'intervalles de 15 minutes pour chaque installation en utilisant des mesures d'irradiance satellitaires, la réanalyse météorologique et la modélisation numérique du terrain. La comparaison de cette production théorique avec la production mesurée réelle permet au système d'identifier automatiquement les anomalies et de suivre la dégradation des panneaux solaires au fil du temps.

L'algorithme constitue la pierre angulaire des rapports de performance complets de BDPV, qui sont générés toutes les deux semaines pour plus de 3 000 installations. Ces rapports fournissent aux producteurs individuels des informations détaillées sur les performances de leur installation et les avisent automatiquement de toute anomalie détectée. Cette approche de surveillance proactive renforce la confiance des producteurs dans la technologie solaire tout en permettant des réponses rapides aux problèmes potentiels, minimisant ainsi les pertes de production.

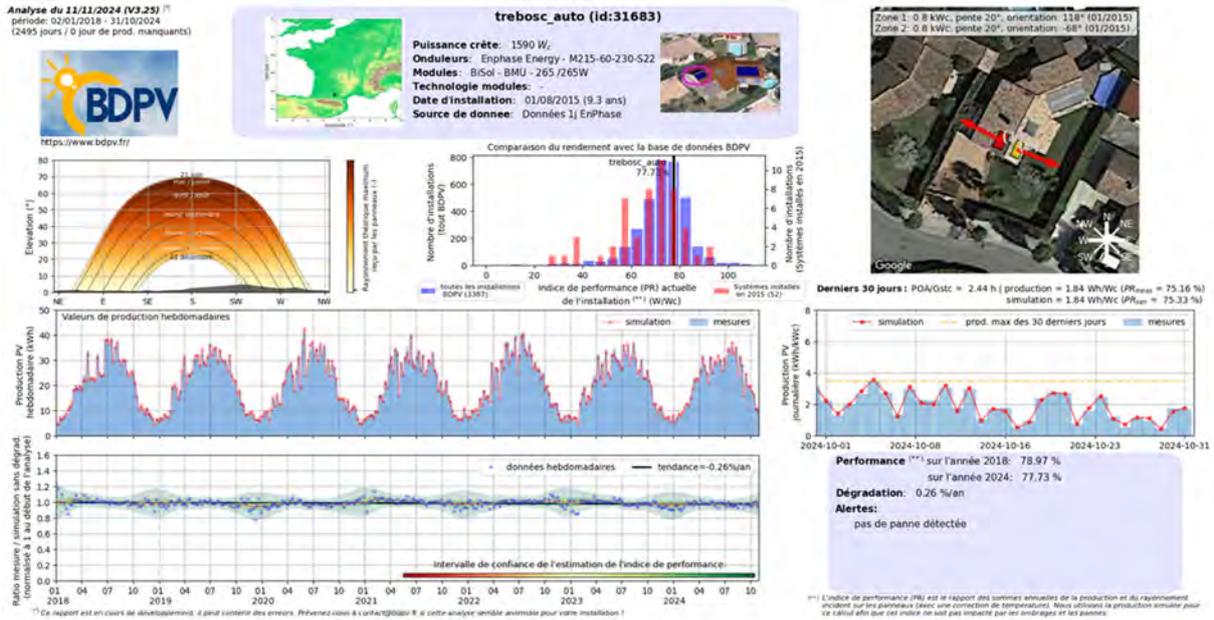


Figure 1. Rapport de performance pour une installation d'énergie solaire

Un regard plus attentif sur les rapports de performance de BDPV

Les rapports de performance fournissent à chaque producteur des informations complètes sur la santé de son installation grâce à plusieurs visualisations et mesures clés :

1. Analyse du parcours solaire et de l'ombrage

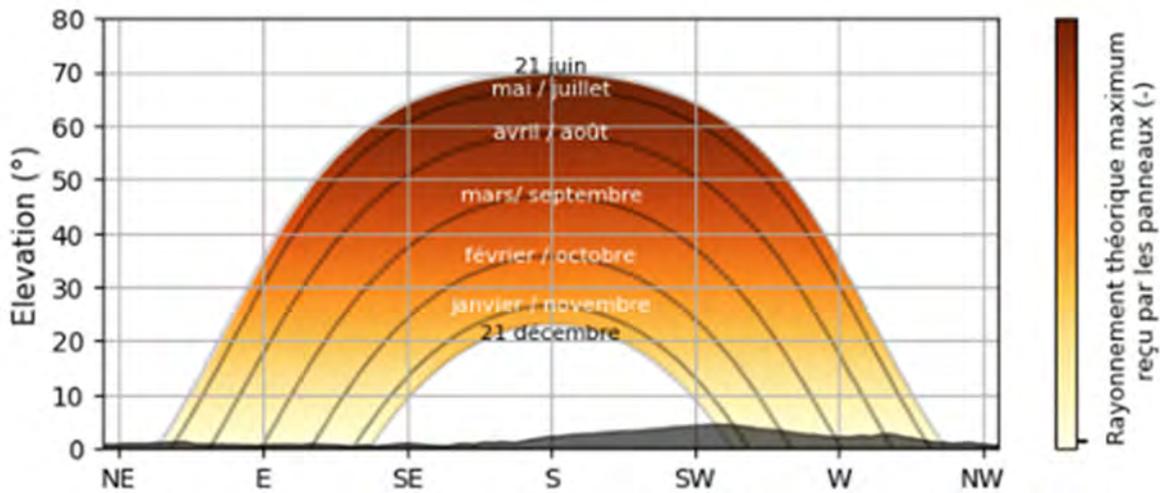


Figure 2. Analyse de l'ombrage montrant les trajectoires solaires au fil des saisons, l'intensité de la couleur indiquant les niveaux de rayonnement

Le rapport comprend une analyse de l'ombrage qui superpose la trajectoire du soleil tout au long de l'année avec les obstructions potentielles. À l'aide de dégradés de couleurs allant du jaune à l'orange foncé, il montre l'intensité maximale théorique du rayonnement, aidant ainsi les propriétaires de panneaux solaires à comprendre comment les ombres des bâtiments ou des éléments de terrain voisins ont un impact sur leur production à différents moments de la journée et de l'année.

2. Suivi de la performance de la production

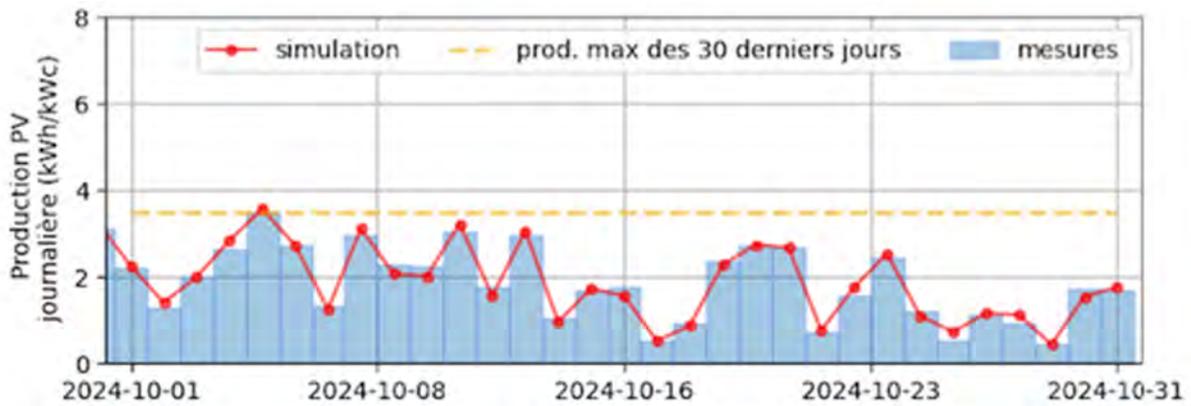


Figure 3. Production solaire quotidienne au cours des 30 derniers jours : mesures réelles (barres bleues) comparées à la production simulée (ligne rouge), avec une production maximale sur la période (ligne pointillée orange)

Le rapport de surveillance de la production offre une comparaison visuelle claire entre la production réelle et la production simulée. Les barres bleues représentent la production mesurée, tandis qu'une ligne rouge indique la performance attendue basée sur la simulation. Cette simulation tient compte des conditions météorologiques réelles à l'aide de données satellitaires, ce qui explique pourquoi la ligne rouge suit les mêmes variations météorologiques que la production mesurée. Le rapport affiche ces données à intervalles hebdomadaires et quotidiens, fournissant ainsi des tendances à long terme et des informations immédiates sur les performances.

3. Évolution de la performance à long terme

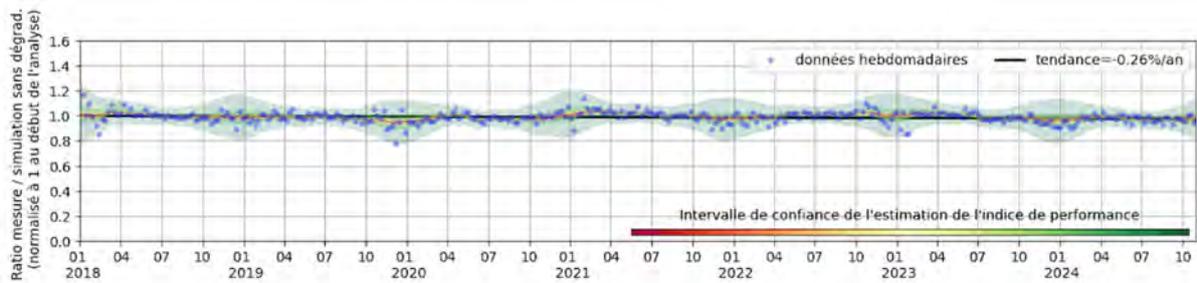


Figure 4. Ratio de performance au fil du temps : ratio hebdomadaire de la production réelle par rapport à la production prévue simulée (points bleus), avec ligne de tendance (noire)

Le rapport suit l'efficacité de l'installation au fil du temps grâce à une analyse du ratio de performance. Cette visualisation présente des points de performance hebdomadaires, un ratio comparant la production mesurée aux attentes simulées, avec une ligne de tendance noire issue d'une régression robuste. Les intervalles de confiance à code couleur en vert indiquent la plage de performance attendue, tandis que les zones rouges mettent en évidence les périodes de sous-performance importante qui peuvent nécessiter une attention particulière. Le système calcule les taux de dégradation annuels à partir de la pente de régression, aidant ainsi les producteurs à comprendre le processus de vieillissement de leurs panneaux et à le comparer aux normes de l'industrie.

Ces analyses automatisées aident les producteurs à maintenir une performance optimale de leurs installations. Une section récapitulative dans le coin inférieur droit du rapport affiche les mesures clés, y compris les taux de dégradation à long terme et les alertes système, ce qui permet aux utilisateurs non techniques d'évaluer rapidement la santé de leur système.

Le défi : mise à l'échelle de la génération de rapports de performance

Bien que BDPV ait réussi à gérer sa communauté de milliers de producteurs solaires pendant des années grâce à une simple analyse comparative, l'introduction de son système d'analyse de performance sophistiqué a apporté « une fusée en termes de complexité », se souvient David Trebosc, président de BDPV. « Notre système initial fonctionnait bien pour comparer les installations, mais la génération de ces nouveaux rapports de performance très détaillés à grande échelle nécessitait une refonte complète de notre approche. »

L'ampleur de l'opération était impressionnante : plus de 3 000 installations solaires transmettant quotidiennement des données de production, chacune ayant besoin de rapports de performance complets toutes les deux semaines. Derrière ces rapports se trouvent des couches de traitement de données complexes. Le système devait ingérer et analyser de grandes quantités de données météorologiques, en extrayant des séries chronologiques pour chaque installation à partir de plusieurs images satellites. Pour chaque installation, l'algorithme devait traiter ces données météorologiques ainsi que des modèles numériques de terrain pour calculer la production théorique à intervalles de 15 minutes. Ensuite, le système devait comparer cette production théorique aux données de production mesurées réelles pour détecter les anomalies et calculer les mesures de performance. Enfin, ces données analysées devaient être transformées en rapports visuels détaillés avec plusieurs graphiques et mesures, y compris l'analyse du parcours solaire, les comparaisons de production et les tendances de dégradation à long terme. Tout cela devait être généré et distribué à des milliers d'utilisateurs.

Le fonctionnement en tant qu'association à but non lucratif a ajouté une autre couche de complexité au défi. Avec des ressources financières limitées et s'appuyant sur le soutien de bénévoles plutôt que sur du personnel technique à temps plein, BDPV avait besoin d'une solution à la fois évolutive et rentable. « Nous devons être créatifs, » dit David Trebosc. « Nous ne pouvions pas jeter de l'argent sur le problème : nous avons besoin de solutions intelligentes qui maximiseraient nos ressources limitées. »

Le flux de travail de traitement nécessitait une coordination minutieuse de plusieurs étapes interconnectées. Le système devait gérer de nombreuses tâches de traitement de données, de l'ingestion de données satellitaires à la génération finale de rapports. Certaines d'entre elles s'exécutaient en séquence, tandis que d'autres étaient en parallèle, tout en maintenant de solides capacités de surveillance et de gestion des erreurs.

« Ce dont nous avons besoin, c'était d'une architecture capable de gérer des calculs scientifiques complexes à grande échelle, tout en restant gérable pour une organisation gérée par des bénévoles. Trouver cet équilibre était notre plus grand défi », explique Yves-Marie Saint-Drenan, le chercheur à l'origine du nouvel algorithme de BDPV.

Une architecture composite, combinant plusieurs services AWS

Pour relever ce défi, BDPV a recherché une solution soigneusement architecturée utilisant les services cloud. L'objectif était de sélectionner les services en fonction de leurs forces et de leur

rentabilité spécifiques, en créant un système capable de fournir des performances et une fiabilité élevées dans les contraintes de l'organisation.

AWS offre un éventail diversifié de services qui répondent à diverses exigences de traitement. L'architecture développée par BDPV utilise ces services pour optimiser les performances et la rentabilité tout au long de son pipeline :

- Amazon S3 sert de solution de stockage principale, hébergeant à la fois l'imagerie satellitaire brute et les données traitées. Son efficacité en termes de coûts et sa capacité à paralléliser les opérations de lecture et d'écriture l'ont rendu parfait pour l'ensemble de données volumineux de données météorologiques historiques.
- AWS Step Functions orchestre l'ensemble du flux de travail, gérant des séquences de traitement complexes et permettant une exécution parallèle lorsque cela est possible. Ce service s'est avéré précieux pour coordonner les différents composants tout en maintenant la visibilité sur le processus.
- Les fonctions AWS Lambda gèrent les tâches de courte durée et parallélisables, telles que l'extraction de pixels spécifiques à partir d'images satellites. Leur nature sans serveur signifie que vous ne payez que pour le temps de calcul réel, ce qui les rend rentables.
- Amazon ECS (Elastic Container Service) est utilisé pour les tâches qui nécessitent des bibliothèques spécialisées ou des temps de traitement plus longs que ce que Lambda permet. Il offre la flexibilité d'exécuter des charges de travail conteneurisées tout en maintenant la rentabilité.
- AWS Glue gère les processus d'extraction, de transformation et de chargement (ETL), en particulier pour la concaténation des données quotidiennes et la création de séries chronologiques. Sa nature gérée réduit les frais généraux opérationnels.

Un regard plus attentif sur le flux de travail de génération de rapports de performance

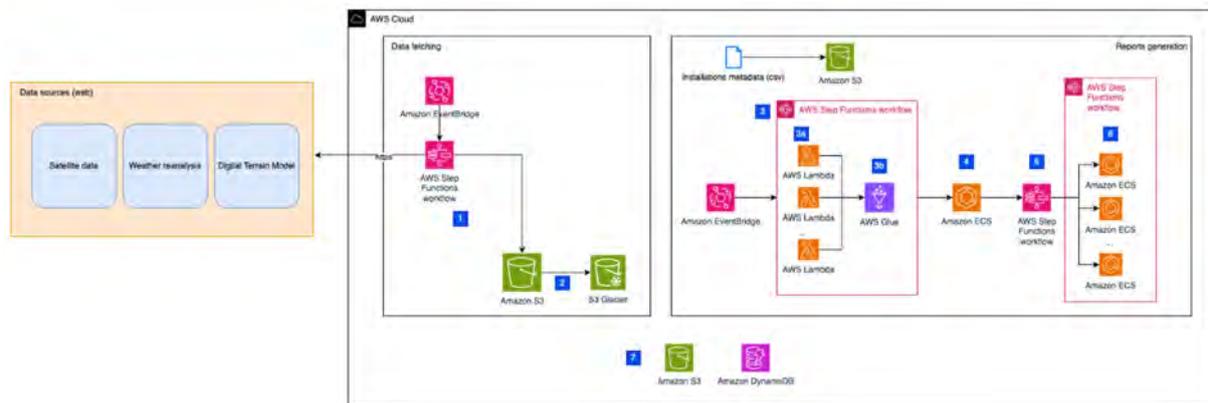


Figure 5 : Flux de travail de génération de rapports de performance

La génération de rapports de performance implique plusieurs étapes, orchestrées par le biais des flux de travail Step Functions. Dans cette section, nous examinons chaque étape de ce processus :

1. Traitement des données satellitaires

Le flux de travail, déclenché quotidiennement par Amazon EventBridge selon un calendrier, commence par la récupération et le traitement des données satellitaires. Cette étape utilise Amazon ECS en raison de l'exigence de bibliothèques spécialisées qui dépassent les limitations de taille de Lambda. Les images satellites brutes sont traitées et recadrées aux zones géographiques pertinentes.

2. Optimisation des coûts avec Amazon S3 Glacier

Pour optimiser les coûts de stockage, les images satellites LSA-SAF (Land Surface Analysis Satellite Application Facility) originales sont automatiquement déplacées vers S3 Glacier par le biais de politiques de cycle de vie. Cette approche réduit considérablement les coûts de stockage tout en maintenant l'accès aux données historiques en cas de besoin.

3. Génération de séries chronologiques de points de grille

Le système génère des séries chronologiques pour chaque point de grille (représentant un emplacement géographique) en utilisant une combinaison de fonctions Lambda pour le traitement parallèle et AWS Glue pour l'agrégation des données. Ce processus a été détaillé dans un article précédent, démontrant comment gérer efficacement ce traitement de données à grande échelle. À l'étape 3a, chaque fonction Lambda traite une journée de données. Ensuite, à l'étape 3b, les données de séries chronologiques sont restructurées avec AWS Glue, changeant la clé de partition de basée sur la date à basée sur le point de grille. Cette transformation optimise la structure des données pour les étapes de traitement suivantes.

4. Intégration des données historiques

Les données nouvellement extraites sont concaténées avec les données précédemment traitées des exécutions antérieures, assurant ainsi un ensemble de données continu et complet pour l'analyse.

5. Prétraitement des données

Un flux de travail Step Functions gère la phase de prétraitement en orchestrant les fonctions Lambda et les tâches ECS qui préparent les données pour la génération de rapports. À la fin de cette étape, toutes les données nécessaires sont organisées par point de grille sur Amazon S3, prêtes pour l'étape de traitement finale.

6. Génération de rapports parallèles

L'étape finale consiste à générer des rapports de performance pour chaque installation. Pour optimiser le temps de traitement, le flux de travail parallélise la génération entre les points de grille en utilisant plusieurs tâches ECS. Cette approche réduit considérablement le temps de traitement global pour les plus de 3 000 installations.

7. Stockage des données et gestion de l'état :

Amazon DynamoDB suit l'état du traitement et stocke les paramètres de calcul, en particulier ceux utilisés pour comparer les installations entre elles. Amazon S3 stocke les données générées à chaque étape du processus, fournissant ainsi une solution de stockage fiable, rentable et évolutive.

Ce flux de travail démontre comment différents services AWS peuvent être combinés pour créer un pipeline de traitement robuste et efficace. L'utilisation de Step Functions pour l'orchestration assure l'exécution et la surveillance fiables de l'ensemble du processus, tandis que la combinaison de divers services de calcul (Amazon ECS, Lambda et AWS Glue) offre le juste équilibre entre performance et rentabilité pour différents types de tâches de traitement.

Impact réel

Depuis la mise en œuvre de la solution AWS, BDPV a réalisé des améliorations significatives de ses capacités de surveillance :

- Capacité de traitement accrue pour gérer plus de 3 000 installations avec des rapports bimensuels
- Réduction de 90 % des coûts de calcul et de stockage grâce à l'utilisation de technologies sans serveur et de conteneurs, et à l'utilisation optimisée d'Amazon S3 et de S3 Glacier
- Détection précoce des problèmes de performance, aidant ainsi les producteurs à minimiser les pertes de production
- Évolution d'une simple analyse comparative à une surveillance précise des performances

« Les services AWS nous ont permis de construire quelque chose d'sans précédent dans l'industrie solaire : une plateforme de surveillance gratuite qui combine des données satellitaires avec des algorithmes sophistiqués pour fournir une analyse de performance détaillée à des milliers de petites installations », dit David Trebosc. « Ce qui est remarquable, c'est que nous ayons réussi cela en tant qu'organisation gérée par des bénévoles. Bien que nous n'ayons aucune expérience préalable en matière d'informatique en nuage, la facilité de mise en œuvre d'AWS et la richesse des ressources en ligne ont fait qu'il a été relativement simple pour nous de donner vie à cette solution. »

Pour l'avenir, BDPV continue de se concentrer sur l'amélioration de ses capacités de surveillance. L'équipe poursuit actuellement des recherches qui démontrent l'engagement de BDPV envers l'amélioration continue du soutien à la communauté croissante des propriétaires de maisons solaires.

Conclusion

Dans cet article, nous avons exploré comment BDPV a utilisé plusieurs services AWS pour créer une architecture évolutive pour la surveillance des installations photovoltaïques. La combinaison d'Amazon S3, d'AWS Step Functions, d'AWS Lambda, d'AWS Glue et d'Amazon ECS leur a permis de créer une solution qui traite des données satellitaires complexes et génère des rapports de performance détaillés pour des milliers d'installations.

Cette architecture démontre comment les organisations peuvent :

- Utiliser une approche composite combinant différents services AWS pour une performance et un coût optimaux
- Utiliser des technologies sans serveur pour minimiser les frais généraux opérationnels
- Mettre à l'échelle des flux de travail de traitement de données complexes tout en maintenant la rentabilité

- Créer des solutions de surveillance sophistiquées avec des ressources limitées

Article original en Anglais : <https://aws.amazon.com/fr/blogs/industries/weather-based-photovoltaic-monitoring-how-bdpv-scales-on-aws/>