Document de travail · Vérification interne | v1.0 | 2026-11-20

# Prévision à 24 heures de la charge horaire du réseau électrique dominicain

Évaluation d’un modèle fixe sur les 8 784 heures de 2024

Adrian Dunkley | info@caribbeanaiassociation.com

## Résumé

Nous avons comparé une régression linéaire à deux règles qui reprennent des mesures passées pour prévoir la puissance active sur le réseau interconnecté dominicain. Le modèle a été ajusté sur 26 087 heures exploitables antérieures à 2024. Les trois méthodes ont été évaluées sur les mêmes 8 784 heures de 2024. La régression a ramené l’erreur absolue moyenne de 123,42 MW pour la règle de la veille à 105,39 MW, soit une baisse de 14,6 %. Elle a sous-estimé la demande de 30,91 MW en moyenne. Le gain de précision s’accompagne donc d’un biais plus marqué. [1, 2]

## Question de recherche

Une régression combinant le calendrier et la demande passée prévoit-elle mieux la charge à 24 heures que la mesure de la veille à la même heure ?

## Données et méthodes

Nous avons utilisé la version 5 des données du SENI de Rafael Uceta et la variable active_power_mw. Le fichier contient 35 039 mesures horaires du 2 janvier 2021 à 1 h au 31 décembre 2024 à 23 h. Nous avons comparé chaque horodatage aux colonnes du mois et de l’heure, vérifié l’absence de doublons et reconstruit les décalages en heures réelles. Les 168 premières mesures n’avaient pas de valeur antérieure d’une semaine. [1, 2]

Pour l’heure cible t, nous avons utilisé la demande aux heures t−24, t−48 et t−168, une constante et des indicateurs d’heure, de jour de semaine et de mois. La mesure à t−24 doit être accessible au moment de prévoir. Nous avons ajusté les coefficients par moindres carrés ordinaires jusqu’à fin 2023, puis les avons gardés fixes en 2024. Les deux règles de comparaison reprenaient les mesures à t−24 ou t−168. Toutes disposaient du même historique à chaque prévision.

Nous avons calculé l’erreur absolue moyenne en mégawatts et l’erreur signée comme la prévision moins la mesure. Pour estimer l’incertitude sur le gain par rapport à la veille, nous avons moyenné par jour les différences appariées d’erreur absolue. Avec la graine 20261120, nous avons tiré 5 000 échantillons circulaires de blocs de sept jours consécutifs, chacun ramené à 366 jours. Nous avons répété le calcul avec des blocs de 14 jours.

## Résultats

L’erreur absolue moyenne s’élevait à 123,42 MW pour la veille, à 162,97 MW pour la semaine précédente et à 105,39 MW pour la régression. Cette dernière a devancé la règle de la veille chacun des 12 mois. Son erreur la plus forte a atteint 130,38 MW en juillet. [1]

Le gain moyen sur la règle de la veille était de 18,04 MW. L’intervalle percentile à 95 % allait de 11,40 à 24,31 MW avec les blocs de sept jours et de 12,20 à 24,12 MW avec ceux de 14 jours. L’erreur signée moyenne était de −30,91 MW pour la régression et de −0,39 MW pour

la veille. Les racines des erreurs quadratiques moyennes étaient respectivement de 137,42 et 170,29 MW. [1]

## Taille moyenne de l’erreur horaire en 2024




Une valeur faible correspond à une erreur plus petite. Les trois méthodes portent sur les mêmes 8 784 heures de test. Source [1], calculs de l’auteur.

## Discussion

Sur l’année testée, un planificateur soucieux de la taille habituelle des écarts préférerait la régression. La sous-estimation moyenne de 30,91 MW justifie aussi une étude des erreurs aux heures de pointe. Une correction étalonnée sur les erreurs de 2024 devrait être évaluée sur des observations ultérieures.

## Limites

Ce test rétrospectif porte sur une année complète de données préparées par les auteurs du jeu de données. Il suppose des mesures disponibles à temps et ne reconstitue pas les délais ni les révisions. Nous avons exclu la météo mesurée à l’heure cible, indisponible au moment de prévoir. Nous n’avons étudié ni la demande non satisfaite, ni les réserves, ni les coûts de programmation des centrales, ni la performance de l’exploitant. Les intervalles décrivent la variabilité de l’année testée avec le modèle ajusté.

## Reproduire l’analyse

Télécharger datos_modelo.csv dans la version 5 du jeu [1]. Lire les dates dans l’ordre jour/mois/année et vérifier les colonnes du mois et de l’heure.

Associer chaque mesure aux valeurs des 24, 48 et 168 heures précédentes. Ajuster la régression avant 2024 en prenant la première catégorie de chaque variable du calendrier comme référence.

Calculer la valeur absolue de l’erreur pour chaque heure de 2024, puis sa moyenne. Pour agréger les résultats mensuels, pondérer par le nombre d’heures de chaque mois.

## Tableau de données

| Mois de 2024 | Heures | Erreur absolue moyenne, veille (MW) | Erreur absolue moyenne, semaine précédente (MW) | Erreur absolue moyenne, régression (MW) |
|---|---|---|---|---|
| 1 | 744 | 102.297 | 125.613 | 92.504 |
| 2 | 696 | 126.808 | 178.018 | 102.88 |
| 3 | 744 | 108.949 | 158.052 | 95.925 |
| 4 | 720 | 130.14 | 207.691 | 112.416 |
| 5 | 744 | 132.782 | 230.034 | 116.888 |
| 6 | 720 | 120.838 | 172.054 | 98.426 |
| 7 | 744 | 163.895 | 180.52 | 130.382 |
| 8 | 744 | 126.179 | 117.589 | 104.585 |
| 9 | 720 | 112.595 | 164.216 | 103.157 |
| 10 | 744 | 117.224 | 122.403 | 107.985 |
| 11 | 720 | 120.623 | 163.38 | 105.494 |
| 12 | 744 | 118.673 | 138.836 | 93.773 |

## Références

[1] Uceta, R. (2025). Short-Term Forecasting Dataset SENI Dominican Rep, version 5. Kaggle. File: datos_modelo.csv.

[2] Uceta-Acosta et al. (2025). A dataset of exogenous variables and historical electricity demand for short-term load forecasting of the national interconnected electric system (SENI) in the Dominican Republic from 2021 to 2024. Data in Brief 63, 112057.

## Crédits des données

Uceta, SENI v5. Original hourly data: CC BY-NC-SA 4.0. CAIA evaluation results.