

Analyse de la consommation électrique par la data science

Présentée par:

Alisson NUNES BONNATO
João Victor FARIA DE SOUZA
Younes OUARHIM
Alexandre CHOUERI



Sommaire :

1. Introduction
2. Traitement des données
3. Choix des variables
4. Modèle de régression multilinéaire
5. Réseau de neurones LSTM
6. Modèle de forêts aléatoires
7. Conclusion et ouverture

Introduction:

Données initiales: 2 bases de données au format CSV. L'une avec des données temporelles de consommation énergétique moyenne et l'autre avec des données climatiques.

Modèle: Proposer un algorithme de prédiction basé sur les concepts de Machine Learning pour estimer la consommation énergétique.

Rendu: Une base de données au format CSV avec des données de consommation estimées pour les mois suivants.

Pré-traitement des données

Database donées énergétiques

	time	mean	temp
0	2009-07-15 00:00:00	0.341659	11.6
1	2009-07-15 00:30:00	0.274463	11.5
2	2009-07-15 01:00:00	0.231365	11.4
3	2009-07-15 01:30:00	0.202708	11.3
4	2009-07-15 02:00:00	0.182256	11.2

Database donées
climatiques

	date	ind	rain	ind.1	temp	ind.2	wetb	dewpt	vappr	rhum	...	ind.3	wdsp	ind.4	wddir	ww	w	sun	vis	clht	clamt
0	2009-07-15 00:00:00	3	0.0	0	11.6	0	11.4	11.3	13.3	98	...	2	6	2	310	2	81	0.0	25000	999	1
1	2009-07-15 01:00:00	3	0.0	0	11.4	0	11.1	10.8	13.0	96	...	2	6	2	310	2	11	0.0	30000	999	1
2	2009-07-15 02:00:00	3	0.1	0	11.2	0	10.9	10.6	12.8	96	...	2	7	2	290	2	11	0.0	30000	999	1
3	2009-07-15 03:00:00	3	0.0	0	11.1	0	10.8	10.5	12.7	96	...	2	8	2	300	2	11	0.0	30000	999	1
4	2009-07-15 04:00:00	3	0.0	0	11.4	0	11.1	10.9	13.0	96	...	2	7	2	310	2	11	0.0	30000	999	1



Traitement des données

Données ajoutées:

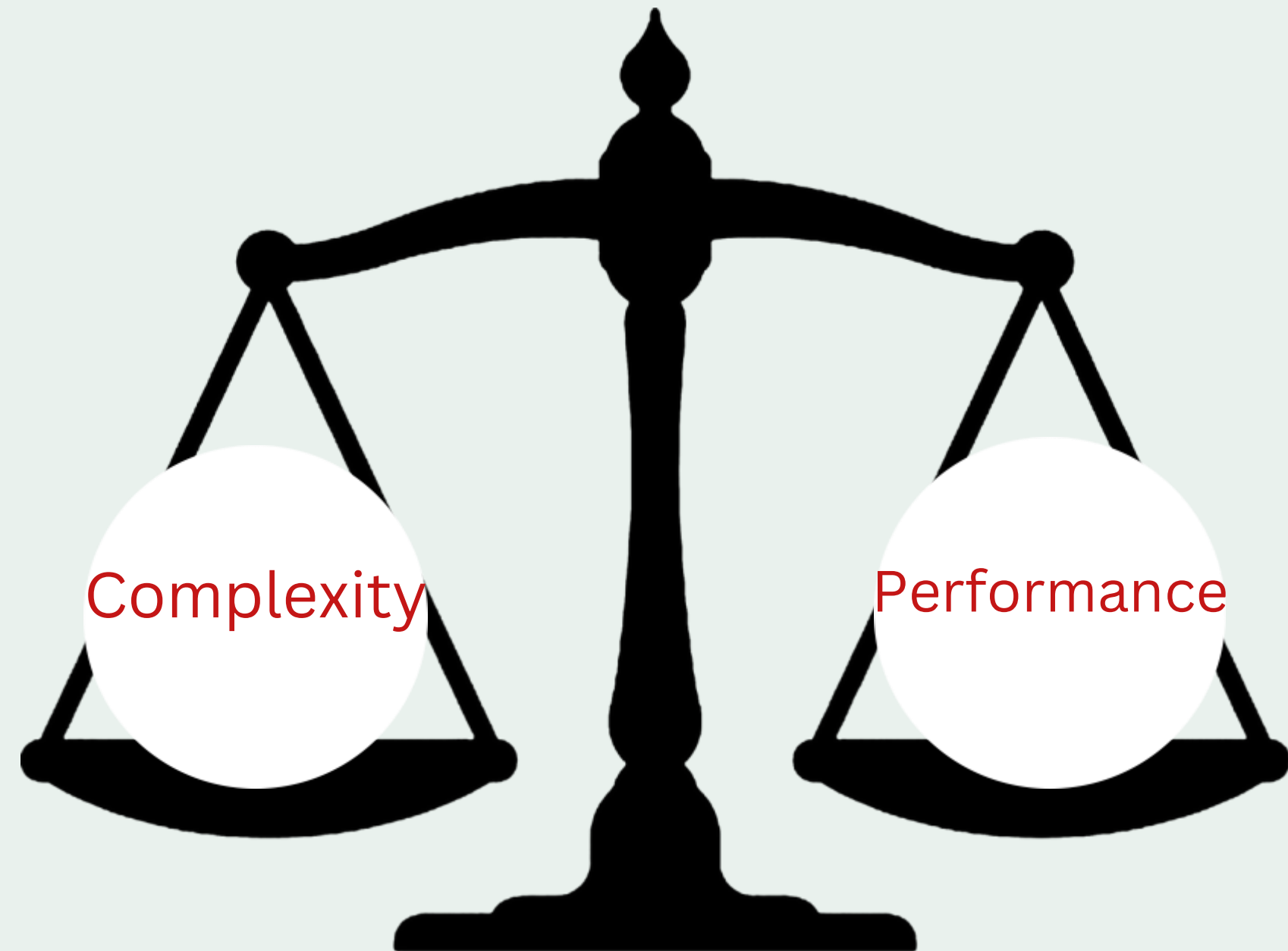
	ind	rain	ind.1	temp	ind.2	wetb	dewpt	vappr	rhum	msl	...	clamt	weekend	season	work_time	year	month	day	hour	minute	weekofyear
date																					
2009-07-15 00:00:00	3.0	0.00	0.0	11.6	0.0	11.40	11.30	13.30	98.0	1002.80	...	1.0	0	2	0	2009	7	15	0	0	29
2009-07-15 00:30:00	3.0	0.00	0.0	11.5	0.0	11.25	11.05	13.15	97.0	1003.15	...	1.0	0	2	0	2009	7	15	0	30	29
2009-07-15 01:00:00	3.0	0.00	0.0	11.4	0.0	11.10	10.80	13.00	96.0	1003.50	...	1.0	0	2	0	2009	7	15	1	0	29
2009-07-15 01:30:00	3.0	0.05	0.0	11.3	0.0	11.00	10.70	12.90	96.0	1003.95	...	1.0	0	2	0	2009	7	15	1	30	29
2009-07-15 02:00:00	3.0	0.10	0.0	11.2	0.0	10.90	10.60	12.80	96.0	1004.40	...	1.0	0	2	0	2009	7	15	2	0	29

Choix des variables

$$AIC = 2K - 2\ln(L)$$

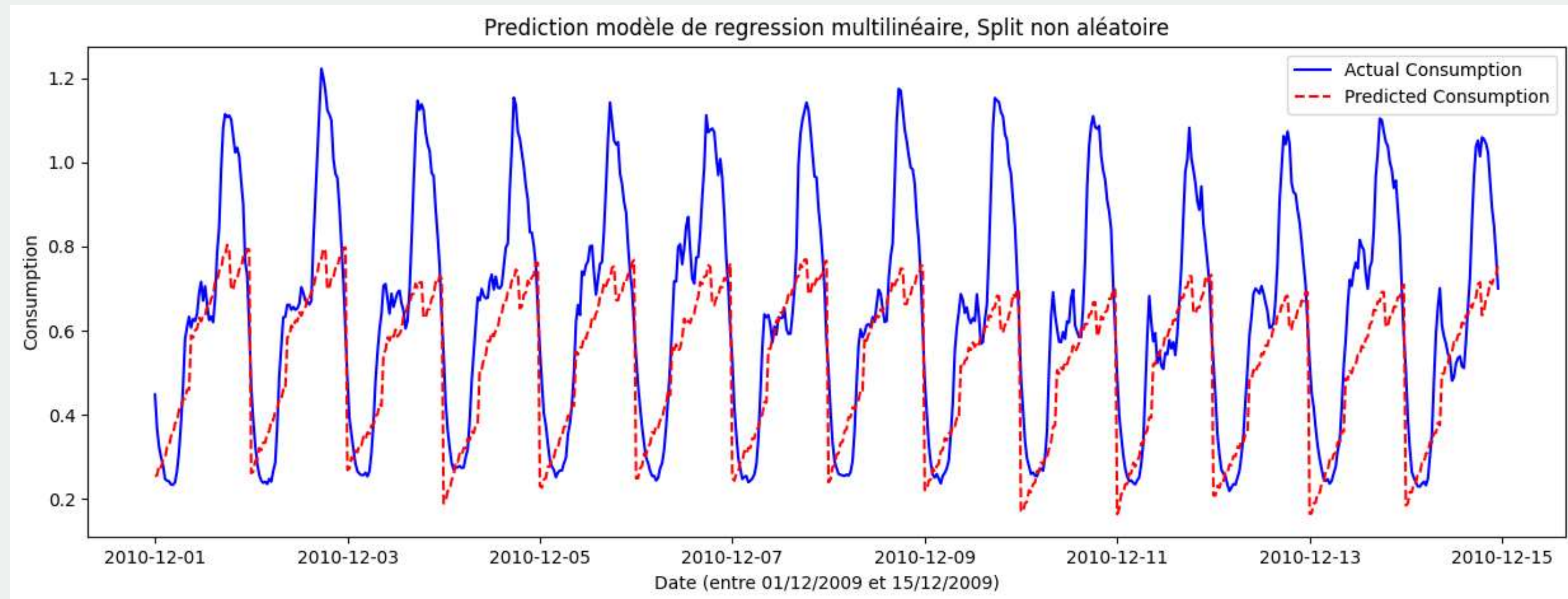
K: Nombre de variables indépendantes

L: Estimateur du maximum de vraisemblance



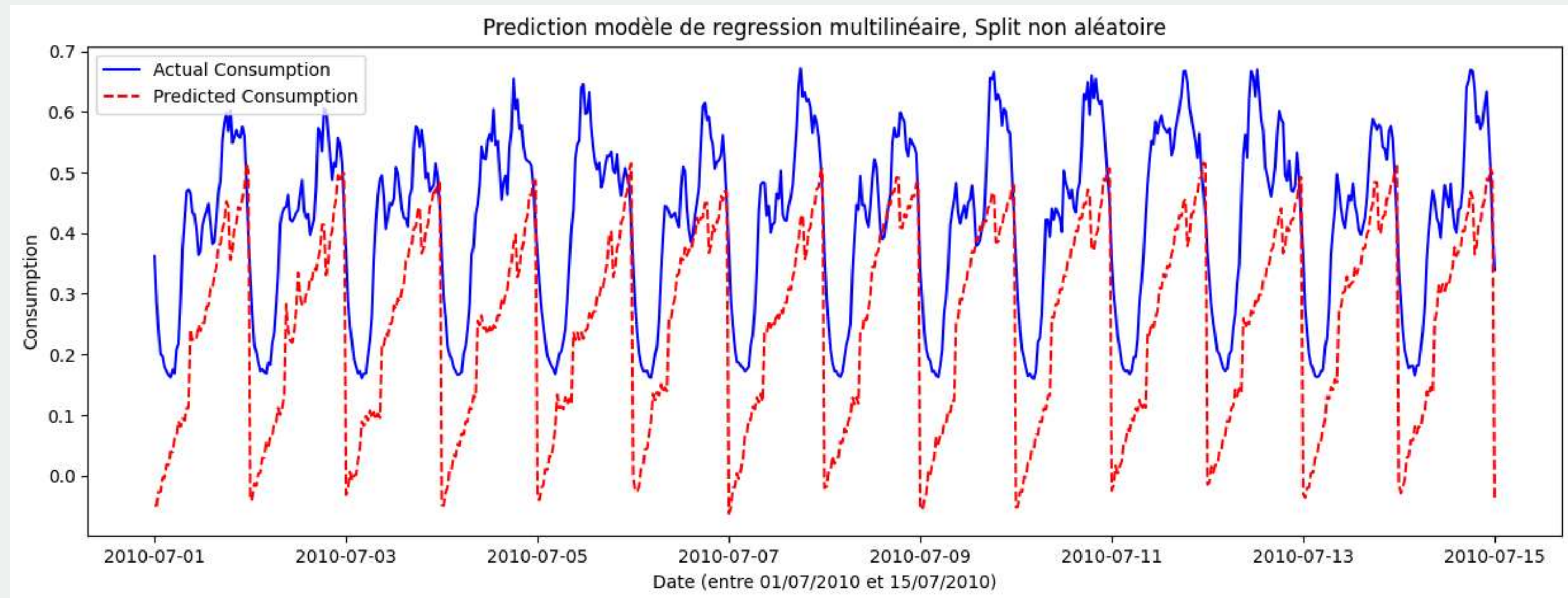
Modèle de régression multilinéaire

Prévision d'hiver:



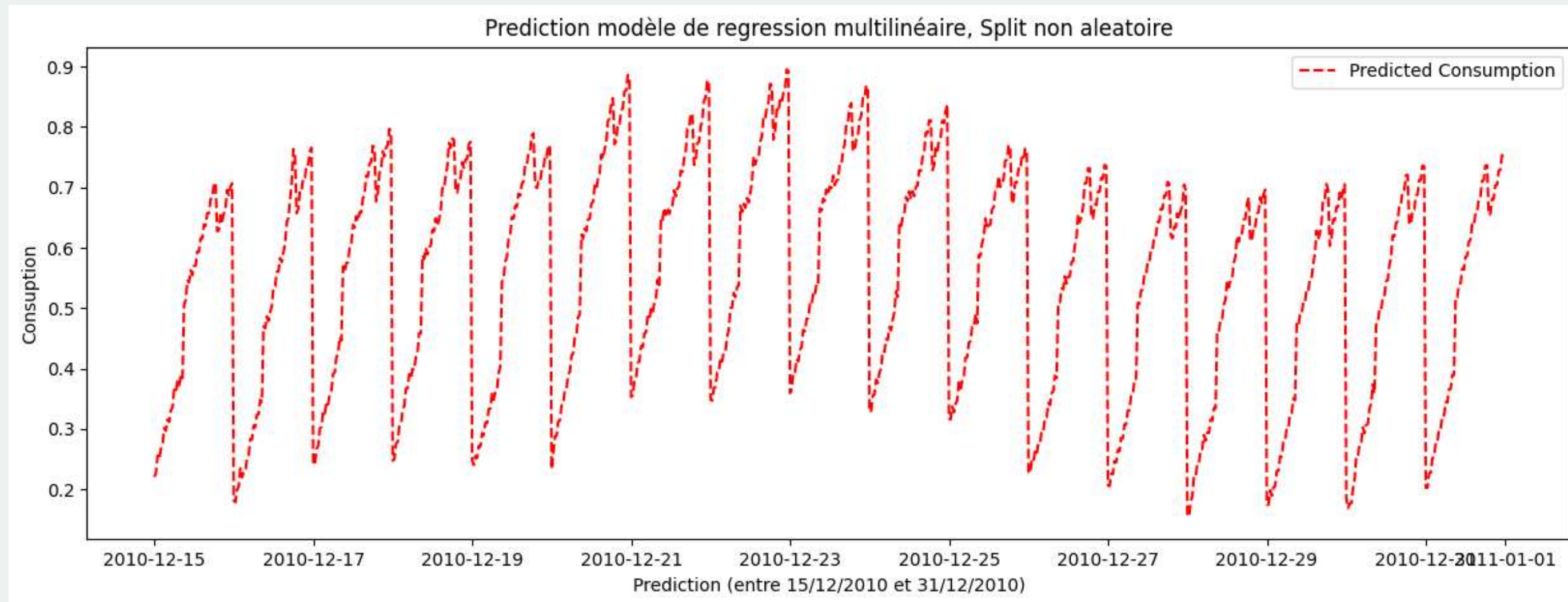
Modèle de régression multilinéaire

Prévision d'été:



Modèle de régression multilinéaire

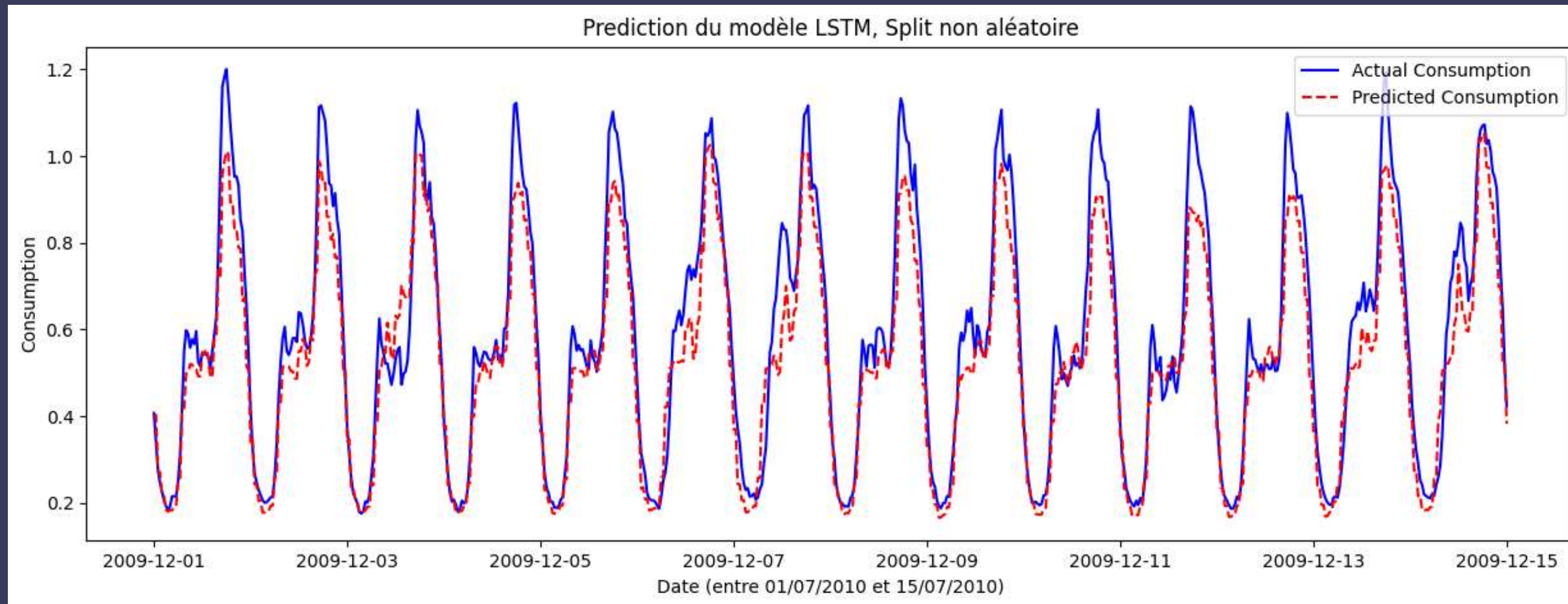
Résultats de prévision:



MAPE: 23,86%

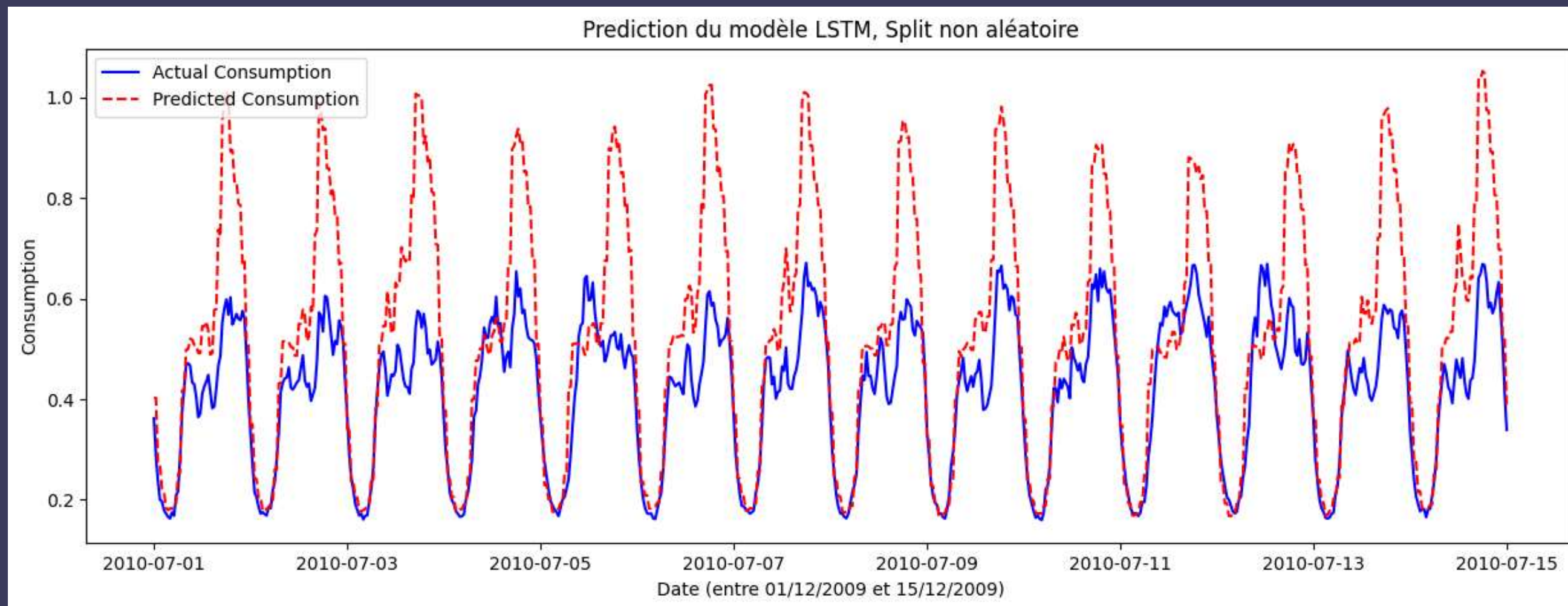
Modèle LSTM

Prévision d'hiver:



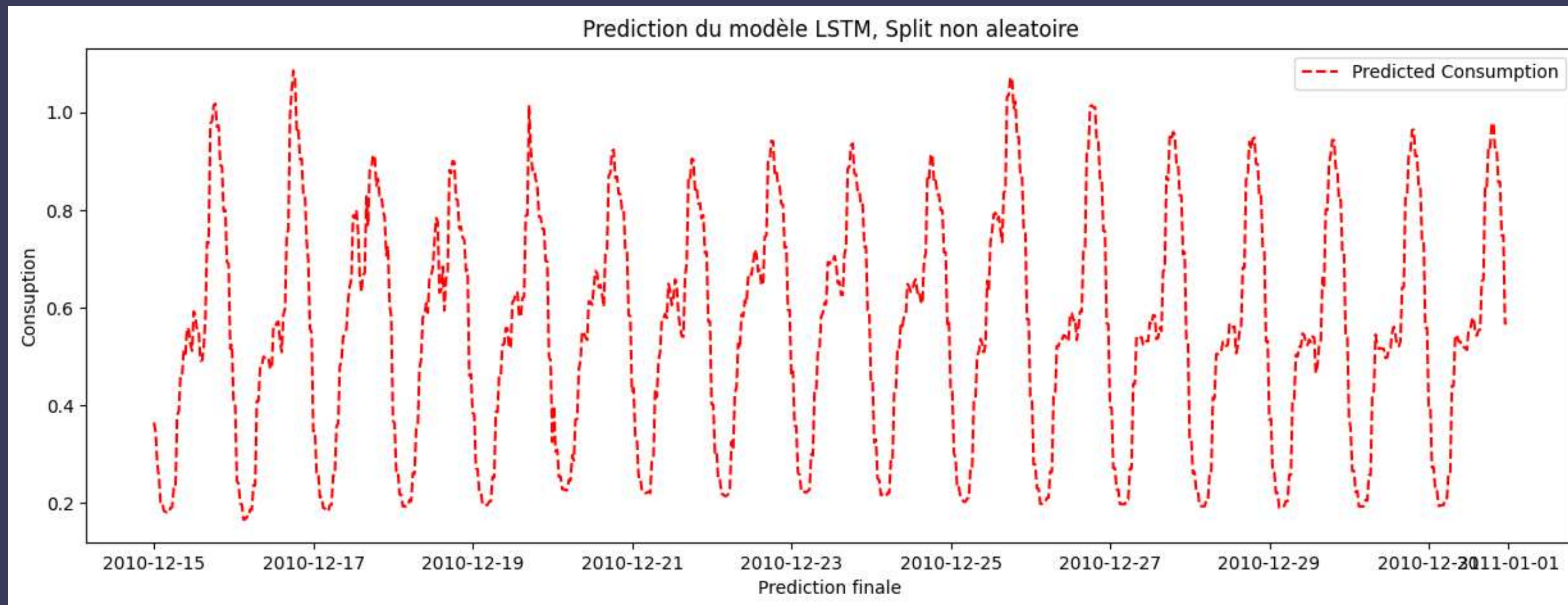
Modèle LSTM

Prévision d'été:



Modèle LSTM

Résultats de prévision:



MAPE: 13,42%

Modèle forêts aléatoires

Entraînement: Split aléatoire et temporelle

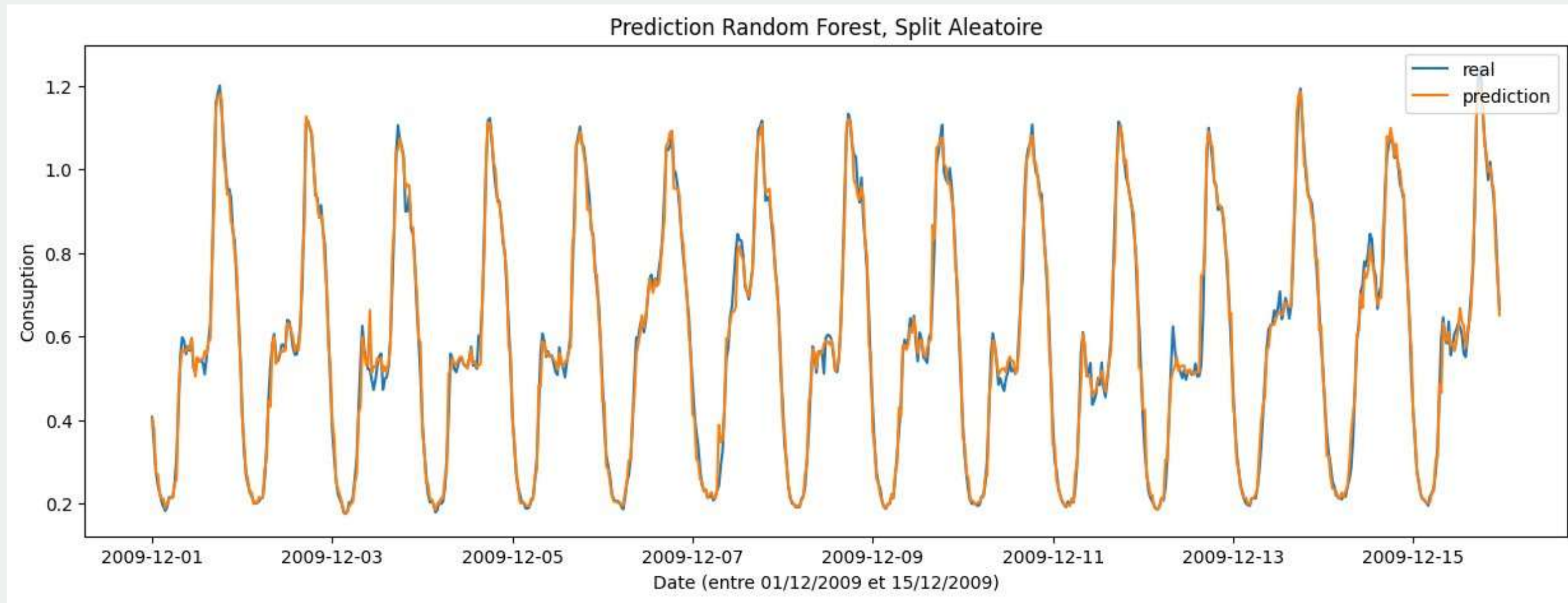
Hyperparamètre:

- RandomSearch et après GridSearch autour de le point trouvé par RandomSearch.
- TimeSeriesSplit.
- Scoring: Neg. MAPE

Resultats : Split aléatoire a un MAPE plus petit que le split temporelle.

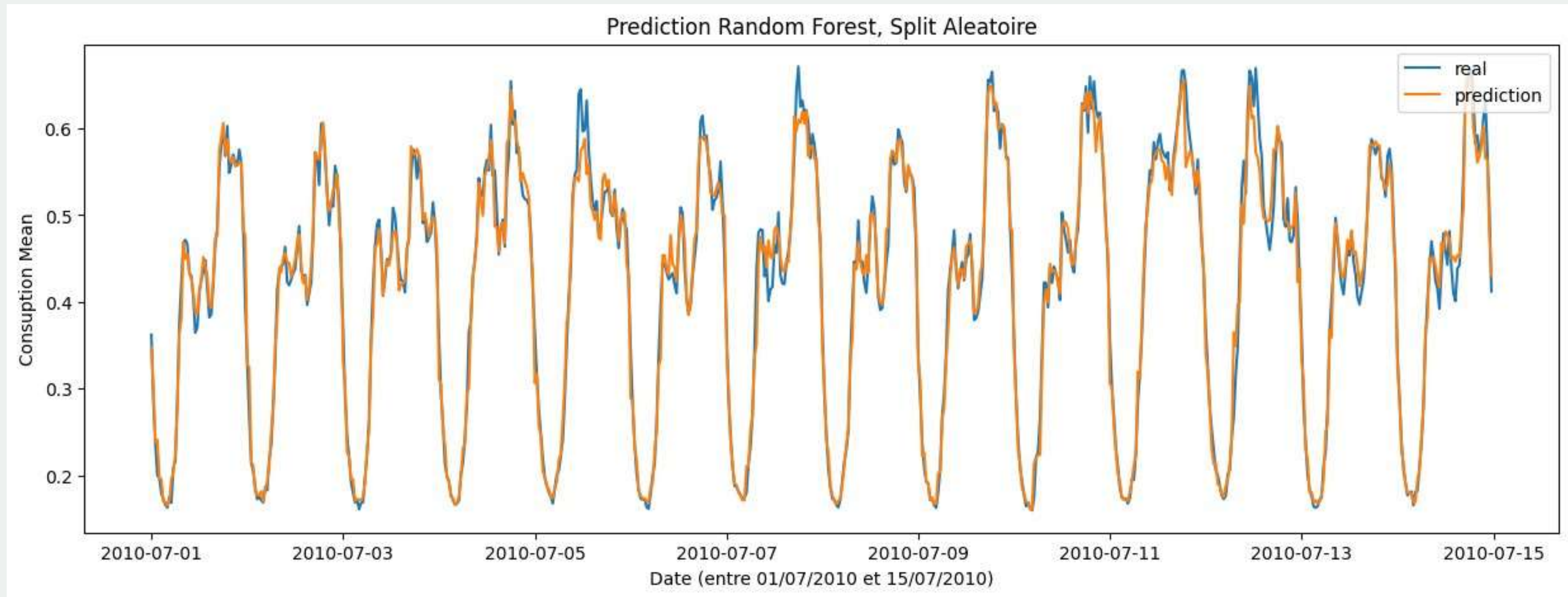
Modèle forêts aléatoires

Prévision d'hiver:



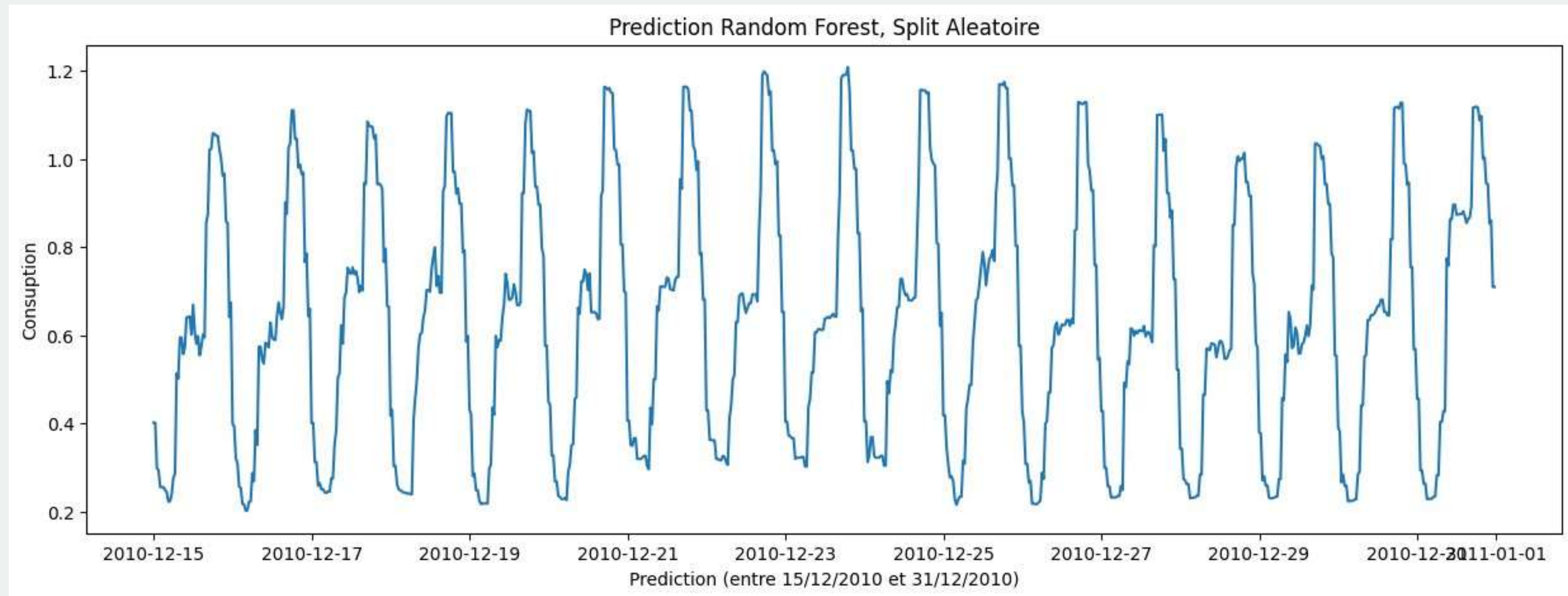
Modèle forêts aléatoires

Prévision d'été:



Modèle forêts aléatoires

Résultats de prévision:



Modèle forêts aléatoires

Résultats de prévision:

	Split Aléatoire	Split Temporelle
MAPE Test Data	7,61%	11,12%
MAPE Fin Dec 2009	3,78%	6,85%
Hyperparamètres	Criterion : poisson n_estimators : 10	Criterion : squared error n_estimators : 50 max_depth : 10

Perspectives:

- Appliquer le modèle LSTM aux erreurs de prédiction pour mieux comprendre et anticiper ces erreurs
- Combiner les modèles pour obtenir une meilleure prédiction
 - Optimisation du réseau de neurones LSTM

Conclusion :

- Un modèle plus convaincant que les autres : le modèle de forêts aléatoires
- Modèle logiquement utilisé pour compléter les données manquantes
- Confiance en les prédictions de ces données au vu des résultats obtenues avec ce modèle



Merci Pour votre attention

Projet prévision de consommation électrique

