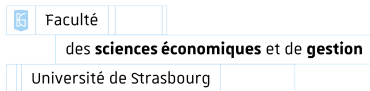


Evolution du prix des matières énergétiques sur le marché international: cas du prix du pétrole

DJAFON Kokouvi Joseph ,RAZAFIMANANTENA Iriantsoal
,KAPO Din

Faculté des sciences économiques et de gestion de Strasbourg



① Introduction:

② Estimations

③ Conclusion

1 Introduction:

2 Estimations

3 Conclusion

Contexte et enjeux

- Enjeux de la transition écologique: Etats et industries de l'automobile
- Énergies fossiles VS Energies renouvelables
- Enjeux économiques et modèles de prévision

Le prix du pétrole

- La variable d'intérêt dans le cadre de l'analyse des matières premières énergétiques est leur prix en \$, Les données ont été récupérées sur le site de la Federal Reserve
- Nous devons nous demander tout d'abord de quelle manière nous allons modéliser les prix. Ici, nous sortons du cadre classique de l'économétrie, où le but est d'analyser un effet causal. Ici, nous sommes en data science : nous voulons faire des prédictions.
 - Les méthodes classiques de prévision en séries temporelles consisteraient à utiliser des processus stochastiques, par exemple : $X_t = X_{t-1} + \epsilon$ AR(1)
 - Nous allons donc nous tourner plutôt vers les modèles de machine learning qui, du fait de leur grande adaptabilité, vont permettre de réajuster le modèle avec le temps.

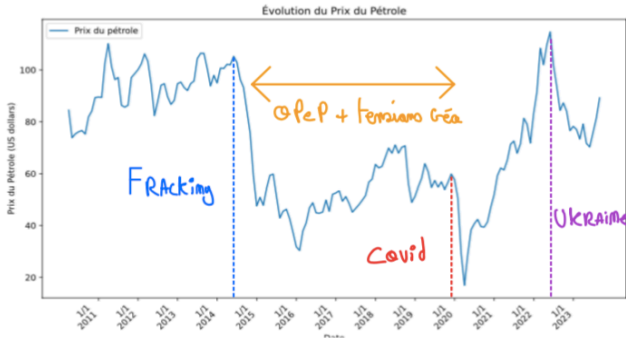


Figure 1: évolution du prix du pétrole

comment modéliser le prix du pétrole ?

- Sur ce graphique, les variations majeures du prix du pétrole sont expliquées par des chocs qui sont, par définition, imprévisibles et difficiles à mesurer
 - Souci : en économétrie, ainsi qu'en data science, on ne peut inclure comme variable explicative que des variables que l'on peut bien mesurer et prédire, ce qui n'est pas le cas des chocs.
 - Solutions : nous allons utiliser le prix des autres matières premières énergétiques pour modéliser le prix du pétrole. Nous disposerons de 162 données mensuelles entre 2010 et 2023.

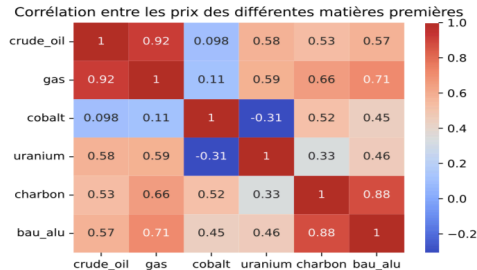


Figure 2: Corrélation entre les prix des différentes matières premières

- Pour avoir une première idée de nos interactions entre les prix des différentes matières premières énergétiques et celui du pétrole, nous utiliserons une matrice de corrélation.

Stat	Cr. Oil	Gas	Cobalt	Uran.	Charb.	B. Alu
Count	162.00	162.00	162.00	162.00	162.00	162.00
Mean	71.47	3.07	40.10	35.90	121.57	154.42
Std	22.34	0.63	16.22	10.63	87.63	34.56
Min	16.98	1.87	21.82	18.57	51.38	115.00
25%	51.64	2.54	29.53	27.39	74.09	131.62
50%	71.46	2.94	33.05	35.19	96.16	143.50
75%	91.95	3.61	50.12	41.70	123.77	162.20
Max	114.68	5.03	93.55	65.00	467.78	269.40

Table 1: Statistique Descriptive

- Données en unités variées, rendant la comparaison de variabilité complexe. L'usage de mesures normalisées est recommandé. Notablement, cet ensemble réel n'a ni outliers ni données manquantes.

Spécification

- Comme nous avons pu le constater, certaines variables sont fortement corrélées avec le prix du pétrole, mais d'autres le sont moins. Nous allons tout de même les inclure dans notre modèle et utiliser la régularisation pour diminuer l'importance des variables les moins pertinentes dans notre modèle
- Notre modèle est : $\log(PY_t) = \beta_0 + \sum_{j=1}^5 \log(PX_{tj})\beta_j$
- note: $\log(PY_t)$ est le logarithme du prix du pétrole et $\log(PX_t)$ Représente le logarithme du prix de la jème matière énergétique, et nous prenons le logarithme pour avoir des élasticités et des résultats dans des échelles comparables. Car, avec les élasticités, nous n'avons plus de problèmes d'unité.

① Introduction:

② Estimations

③ Conclusion

Les problèmes de minimisation

- Nous allons procéder en découpant notre échantillon en un training set (\tilde{T}) de 80% et un test (T) set de 20%.
- Nos problèmes de minimisation sont les suivants :
- $\hat{\beta}_{\text{RIDGE}} = \arg \min_{\beta} \left\{ \sum_{i=1}^{126} \left(\log(PY_t) - \beta_0 - \sum_{j=1}^5 \log(PX_{tj}) \beta_j \right)^2 + \lambda \sum_{j=1}^5 \beta_j^2 \right\}$
- $\hat{\beta}_{\text{LASSO}} = \arg \min_{\beta} \left\{ \sum_{i=1}^{126} \left(\log(PY_t) - \beta_0 - \sum_{j=1}^5 \log(PX_{tj}) \beta_j \right)^2 + \lambda \sum_{j=1}^5 |\beta_j| \right\}$
- La méthode Random Forest implique de déterminer le nombre optimal d'arbres qui minimise l'erreur de prédiction

Paramètres optimaux

Méthode	Paramètre Optimal
Ridge	$\lambda_{\text{Ridge}}^{\text{opt}} = 0.0085$
Lasso	$\lambda_{\text{Lasso}}^{\text{opt}} = 0.000094$
Random Forest	Nombre d'arbres : 10

Table 2: Paramètres optimaux

- On remarque que $\lambda_{\text{Ridge}}^{\text{opt}}$ et $\lambda_{\text{Lasso}}^{\text{opt}}$ sont très bas, ce qui suggère que l'importance de la régularisation est légère ; donc, on peut conclure que nos données ne souffrent (ou très peu) de multicollinéarité et d'overfitting.
- Nous avons seulement besoin de 10 arbres ce qui veut dire qu'un petit nombre d'arbres permet de capturer efficacement les tendances sous-jacentes sans nécessiter une forêt plus large.

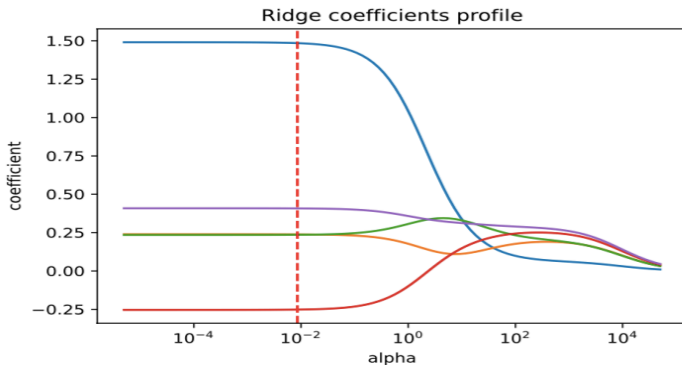


Figure 3: Coefficients de Ridge en fonction de la régularisation

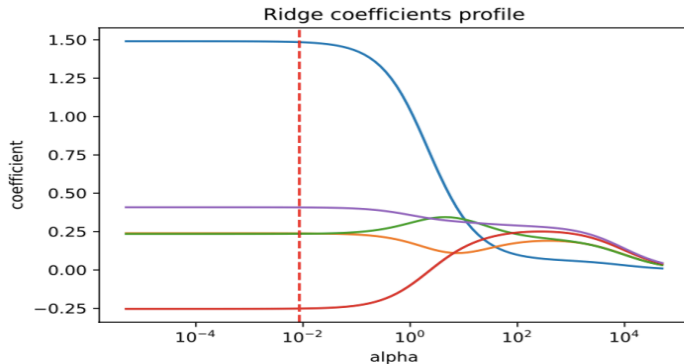


Figure 4: Coefficients de Lasso en fonction de la régularisation

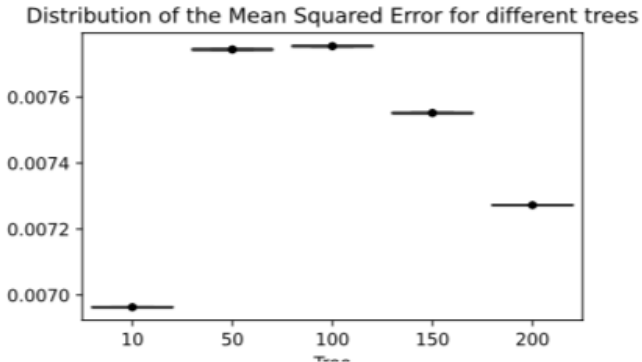


Figure 5: Distribution de l'EQM pour chaque nombre d'arbres

VARIABLES	MÉTHODES		
	RIDGE	LASSO	R-F
GAZ	1.738	1.737	0.916
COBALT	0.194	0.190	0.015
URANIUM	0.198	0.193	0.026
CHARBON	-0.056	-0.056	0.020
BAU-ALU	-0.413	-0.404	0.021
MSE TRAINING	0.02	0.02	0.00
MSE TEST	0.01	0.01	0.01

Figure 6: Résultats des régressions

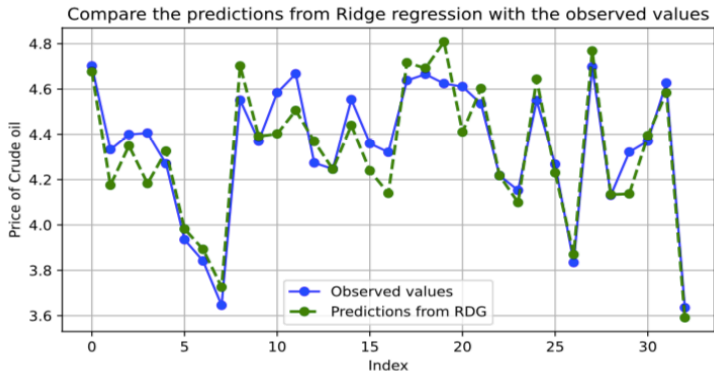


Figure 7: Prédictions versus vraies valeurs (RIDGE)

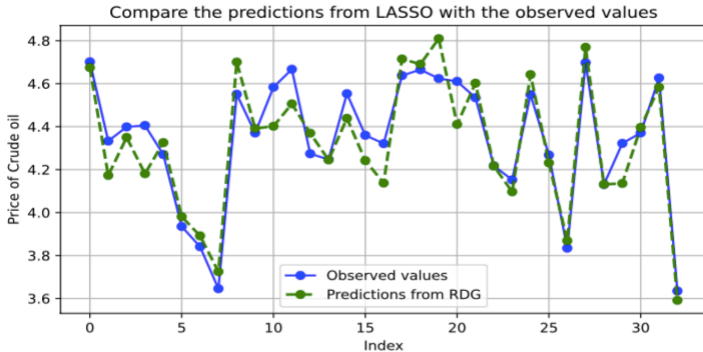


Figure 8: Prédictions versus vraies valeurs (LASSO)

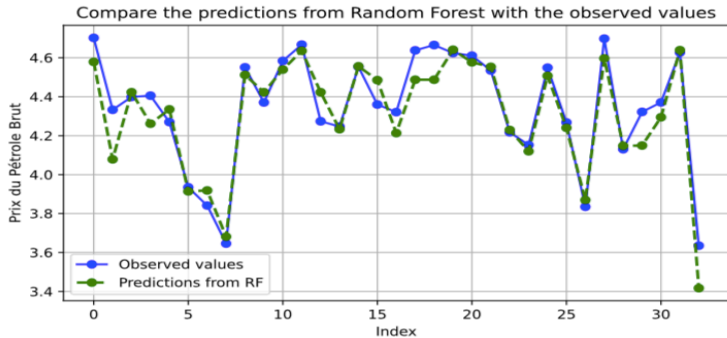


Figure 9: Prédictions versus vraies valeurs (RANDOM FOREST)

① Introduction:

② Estimations

③ Conclusion

Conclusion

- Plus de complémentarité que de substituabilité entre les énergies fossiles et renouvelables
- Des interventions nécessaires des Etats. Notamment des subventions pour rendre plus attractif l'utilisation des énergies renouvelables
- Les limites de notre spécification et de nos méthodes d'estimation peuvent limiter la portée de nos résultats
- D'autres modèles pour tester la robustess de nos résultats ; ARMA, MGARCH, ARDL.