Investigate the determinants of the European Carbon market: a nonparametric approach

Cristiano Salvagnin
c.salvagnin@unibs.it

University of Brescia
Department of Economics and Management

A.M.A.S.E.S. XLVII
Milan, September 21, 2023

UNIVERSITÀ
DEGLI STUDI
DI BRESCIA

Co-authors:

Aldo Glielmo
Bank of Italy

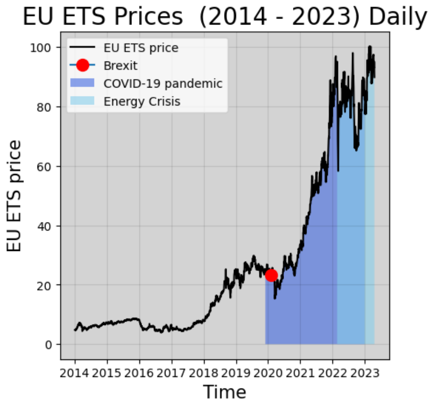
Maria Elena De Giuli
University of Pavia
Department of Economics and Management

Antonietta Mira
Università della Svizzera italiana (USI)
Faculty of Economics
University of Insubria, Como

Acknowledgment to **Antonio Di Noia** (USI) for the suggestion on the use of Gaussian Process regression for the imputation and aggregation of the data.

Outline

Motivation



The EU ETS price shows two different volatility patterns before and after 2019. The major increase in the EU ETS price can be seen during the COVID-19 pandemic when the carbon prices start to rise steeply. The energy crisis started in 2022, gives an important contribution increasing the volatility of the permits prices. The overlaying area shows the joint effect of the pandemic and the energy crisis.

Goals

Our goals:

- What are the EU ETS price determinants?
- Our non-parametric approach is a viable methodology for this market?
- Which category of variables are the most informative respect to the EU ETS price?
- There is a time scale that is more informative than the others?
- What are the performances of the Information Imbalance selection in the Nowcast and Forecast of EU ETS price?

Data

ID	Category	Variables	Database
0	T	EU ETS (EUA)	Bloomberg®
1	UNC	GPR	GPR website
2	UNC	VSTOXX (V2X)	Bloomberg®
3	UNC	Uncertainty EUR/USD (CAFZUUEU)	Bloomberg®
4	UNC	Uncertainty EUR/JPY (CAFZUEJP)	Bloomberg®
5	UNC	Uncertainty EUR/GBP (CAFZUEGB)	Bloomberg®
6	UNC	Uncertainty EUR/CHF (CAFZUECH)	Bloomberg®
7	COM	ICE Dutch TTF Natural Gas (TTF0NXHR)	Bloomberg®
8	COM	Electricity Prices Spain (OMLPDAHD)	Bloomberg®
9	COM	Electricity Prices Germany (EXAPBDHD)	Bloomberg®
10	COM	Electricity Prices Italy (ELIODAHD)	Bloomberg®
11	COM	Electricity Prices France (PWNXFRAV)	Bloomberg®

Category T: Target; **Category UNC**: Uncertainty variables; **Category COM**: Commodity related variables;

ID	Category	Variables	Database
12	COM	ICE Brent oil futures (CO1 Comdty)	Bloomberg®
13	COM	ICE Coal Rotterdam futures (TMA Comdty)	Bloomberg®
14	COM	Gold (GCZ3 Comdty)	Bloomberg®
15	ER	EUR/USD spot (EUR/USD)	Eikon Refinitiv®
16	ER	EUR/JPY spot (EUR/JPY)	Eikon Refinitiv®
17	ER	EUR/GBP spot (EUR/GBP)	Eikon Refinitiv®
18	ER	EUR/CHF spot (EUR/CHF)	Eikon Refinitiv®
19	ENR	Bloomberg Energy price return index (EUNRJP)	Bloomberg®
20	ENR	Solactive ESG Fossil Eurozone 50 index (S0ESG50N)	Bloomberg®
21	ENR	S&P Eurozone 50 Environmental index (SPEENDET)	Bloomberg®
22	ENR	MSCI Europe Energy Sector index (MXEU0EN)	Bloomberg®
23	ENR	ERIX index	Bloomberg®
24	ENR	EUROSTOXX Electricity index (SXEELC)	Bloomberg®
25	CTRY	EUROnext100 (N100)	Bloomberg®

Category ER: Exchange rates; **Category ENR**: Energy indexes; **Category CTRY**: Country indexes;

ID	Category	Variables	Database
26	CTRY	IBEX35 (IBEX)	Eikon Refinitiv®
27	CTRY	DAX	Eikon Refinitiv®
28	CTRY	CAC	Eikon Refinitiv®
29	CTRY	FTSE Mib	Eikon Refinitiv®
30	MACRO	Euro-area 3-month bond yield	Bloomberg®
31	MACRO	Euro-area 10-year bond yield	Bloomberg®
32	MACRO	Euro-area inflation (HICP)	Eurostat
33	MACRO	Euro-area GDP (current value)	Eurostat

Category CTRY: Country indexes; **Category MACRO**: Macroeconomic variables.

Database characteristics: Start to end: January 2014 - April 2023
Time scales: all variables are Daily, Inflation is Monthly and GDP is Quarterly.

Shifting time scales

The first problem found was the heterogeneity of the predictors time scale, we use Gaussian Processes (GP) to solve it.

Gaussian Process equation:

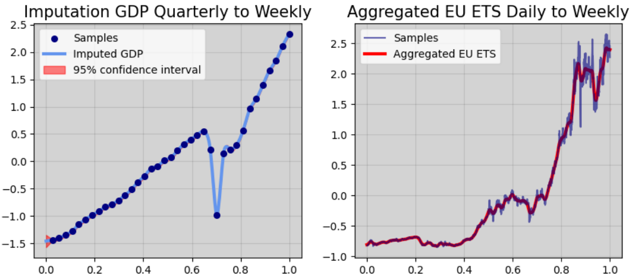
$$f(x) \sim GP(m(x), k(x, x'))$$

Selected kernel function: **matern**

$$k(x, x') = \frac{1}{\Gamma(\nu)2^{1-\nu}} \left(\sqrt{2\nu}\frac{\|x - x'\|}{l}\right)^{\nu} K_{\nu}\left(\sqrt{2\nu}\frac{\|x - x'\|}{l}\right)$$

The *matern kernel* is selected after testing other kernels: Rational Quadratic, RBF, Additive, Multiplicative, with and without constant terms.

Imputation and Aggregation examples



The above plot shows an example of the **imputation process** for the GDP (left side) from *quarterly* to *weekly*, and the **aggregation process** for the EU ETS price (right side) from *daily* to *weekly*.

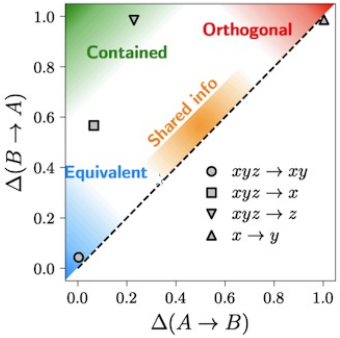
The Information Imbalance

What are the EU ETS price determinants?

We use the **Information Imbalance**: a measure to compute the information content of a space A with respect to a space B.

$$\Delta(A \to B) = \frac{2}{N}\mathbb{E}\left[r_B \mid r_A = 1\right]$$

This plot introduces the **Information Imbalance plane** for a 3D Gaussian generated dataset with small z-variance. The identified areas show the different relationships captured.



Information Imbalance for EUA

We define two spaces

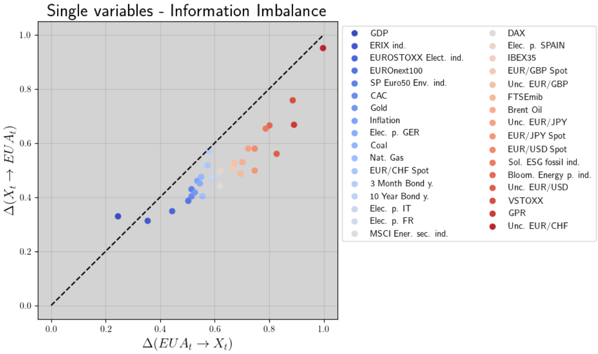
- $\text{EUA}_{t+\delta t}$ = target set,
- X_t = predictors set.

We answer the following question: *which are the most informative variables of the predictors set* (X_t) *with respect to the target set* $(EUA_{t+\delta t})$*?*

$$\Delta(X_t \to \text{EUA}_{t+\delta t})$$

The Information Imbalance identifies informative variables in the presence of any non-linearity.

Information Imbalance analysis

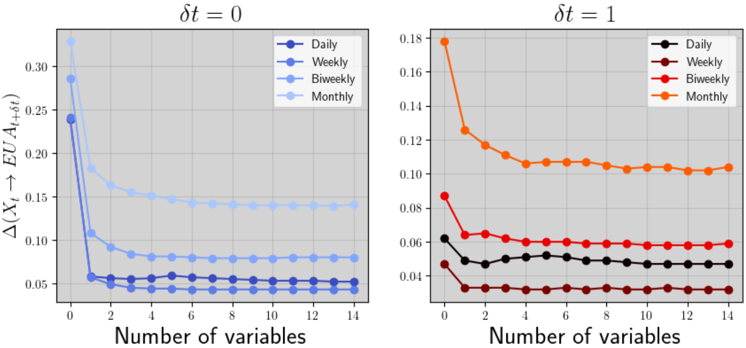


By looking at the imbalances between the target and each variables of the informative pool, the GDP is by far the most informative single variable, followed by the ERIX index and the EUROSTOXX Electricity price return index. We can see that there is no clear visual about the possible relationship between the target and the informative pool.

The Greedy optimization

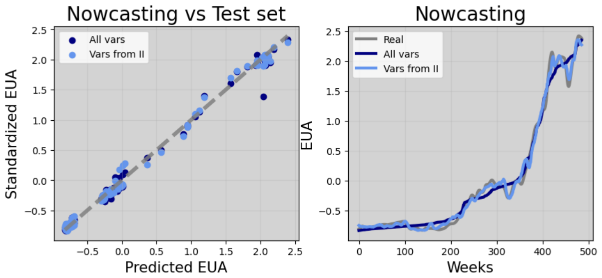
The second step of our analysis involves application the greedy optimization algorithm of the information imbalance. The first step is to identify the most informative single-variable subset. Then, the most informative two-variable subset is identified, and so on.
The results show the Macroeconomic and Commodity variables are again the most informative ones. In particular GDP, Inflation, Gold and Spanish electricity prices are persistent in each subset.

The most informative time scale



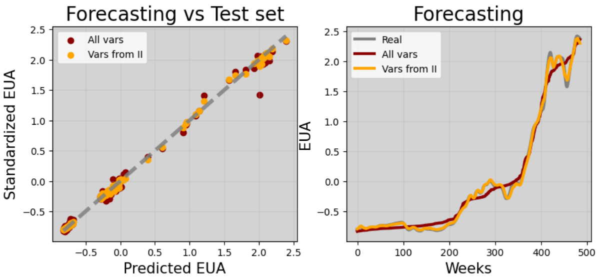
Another question that arises through the research was the use of the Information Imbalance to find also the most informative time scale. The imbalance analysis shows that the most informative time scale is the weekly. We implement also a time scale analysis with a rolling window of 1 (day, week, biweek, month) showing the same preference.

Nowcast the EU ETS price



We compute two predictions: one with only the 4-tuple resulting from the Imbalance analysis and one with all the variables in the dataset.
The prediction is carried out by using a GP with a *kfold cross validation* with $k = 5$ and random splits.
The results show that the performances of model tested are: $R^2_{II} = 0.99$ and $MSE_{II} = 0.005$, while the all variable model show $R^2 = 0.98$ and $MSE = 0.007$.

Forecast the EU ETS price



The same methodology is used for the 1-week ahead Forecast, k-fold cross validation assumes $k = 5$ folds, again with random splits.
The scatter plot shows that the orange points (II subset) are closer to the dashed line respect to the dark red ones. In the second plot is easier to see that the orange model follows the real data better than the all variable model. The II subset is again preferred with $R^2_{II} = 0.99$ and $MSE_{II} = 0.001$ compared to $R^2 = 0.98$ and $MSE = 0.007$.

Key takeaways

- Macroeconomic and Commodities variables affect carbon price.
- Financial factors are less important.
- Gaussian Process predicts carbon price with low error.
- Information Imbalance allows for more accurate estimations.
- Weekly time scale is the most informative and provides the lowest mean squared error for carbon price estimation.

Future research

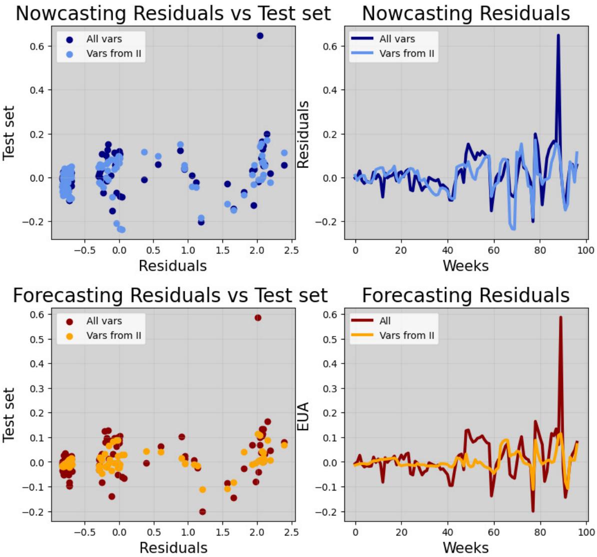
- Jack-knife error estimation (in progress).
- Analyse carbon price volatility determinants in a non-parametric framework.
- Develop a non-parametric model for predicting carbon price volatility.
- Explore a new multi-set Information Imbalance approach.

References

1. Glielmo, Aldo, et al. "Ranking the information content of distance measures." PNAS nexus 1.2 (2022).
2. Del Tatto, Vittorio, et al. "Robust inference of causality in high-dimensional dynamical processes from the Information Imbalance of distance ranks." arXiv preprint (2023).
3. Lovcha, Yuliya, Alejandro Perez-Laborda, and Iryna Sikora. "The determinants of CO2 prices in the EU emission trading system." Applied Energy 305 (2022).
4. Aller, Carlos, Lorenzo Ductor, and Daryna Grechyna. "Robust determinants of CO2 emissions." Energy Economics 96 (2021).
5. Adekoya, Oluwasegun B. "Predicting carbon allowance prices with energy prices: A new approach." Journal of Cleaner Production 282 (2021).

Thank You

Residual analysis



Residual information content - Information Imbalance