

Effects of CO₂ Emissions on Car Sales

Danmeng Yang
Maria Aba
Matteo Posenato

In a world increasingly affected by climate change, machines are an important factor.

It is estimated that the number of cars in the world is around 1.2 billion.

Car manufacturers have to be ready for a change in the production of more and more sustainable cars, given the increasing number of emission-conscious consumers.

- ❖ Analyze and forecast time series sales
- ❖ Understand the relationship between CO2 emissions and sales of different types of vehicles

We considered two datasets, both retrieved from **ourworlddata**. The first dataset consists of the sales of different types of vehicles, including, petroleum, diesel, full hybrid, plug-in electric hybrids, and fully electric battery vehicles for some European countries (Italy, France, Spain, Belgium,...).

The second one represents the average emission of CO₂ (in grams) per kilometer of each of the previously mentioned vehicle types in each country.

For convenience we merged the two into one dataset, aligning them by country and year from 2001 to 2019.

	Year	Entity	battery_electric_number	plugin_hybrid_number	full_mild_hybrid_number	petrol_number	diesel_gas_number	co2_per_km
4	2001	France	407.000000	0.0	84.000000	986491.000	1.267750e+06	160.096853
5	2001	Germany	56.000050	0.0	606.000544	2185246.962	1.155300e+06	178.923476
8	2001	Italy	147.000000	0.0	23.000000	1557121.000	8.880280e+05	158.485545
10	2001	Netherlands	13.000711	0.0	383.020946	408614.346	1.212766e+05	173.746087
14	2001	United Kingdom	29.000000	0.0	719.000000	2019996.000	4.376280e+05	177.881718

The merged dataset now contains 309 observations and 8 variables.

The variables are as follow:

- ❖ **Year** : from 2001 to 2019
- ❖ **Entity**: the different countries
- ❖ **The number of sales of the different types of cars**
 - Battery_electric_number
 - Plugin_hybrid_number
 - Full_mild_hybrid_number
 - Petrol_number
 - Diesel_gas_number
- ❖ **Co2_per_km**: the emission of carbon in grams per kilometer

- ❖ Remove the column with the country code
- ❖ Remove countries with fewer data
- ❖ Remove NA values
- ❖ Scale data with the MinMax scaler

EDA

we chose five countries for which we have data for each type of car since 2001:

- ❖ France
- ❖ Netherlands
- ❖ Germany
- ❖ United Kingdom
- ❖ Italy

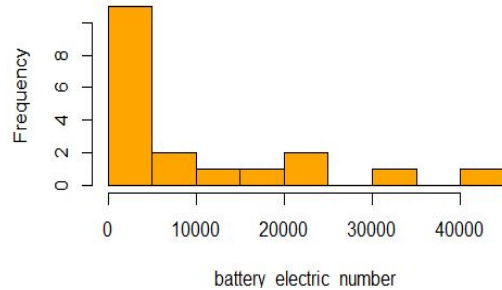
TYPES OF CARS: France

Distribution

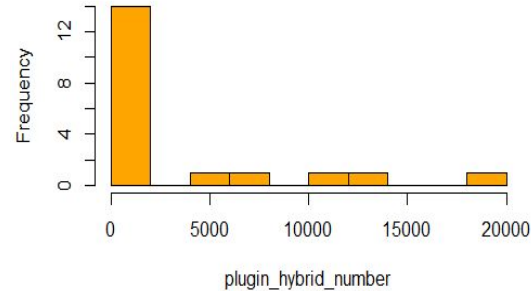


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

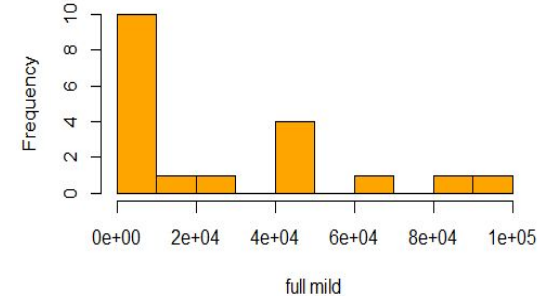
Histogram of battery(France)



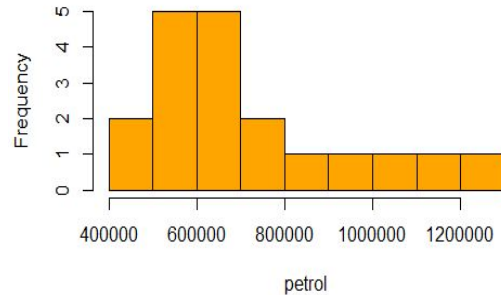
Histogram of plugin hybrid(France)



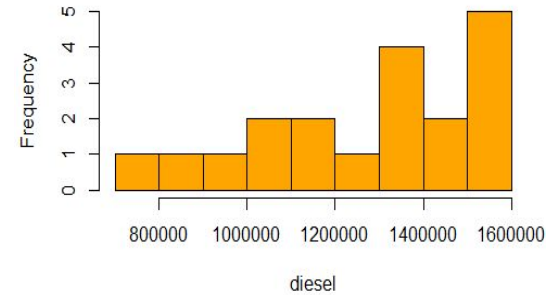
Histogram of full mild hybrid(France)

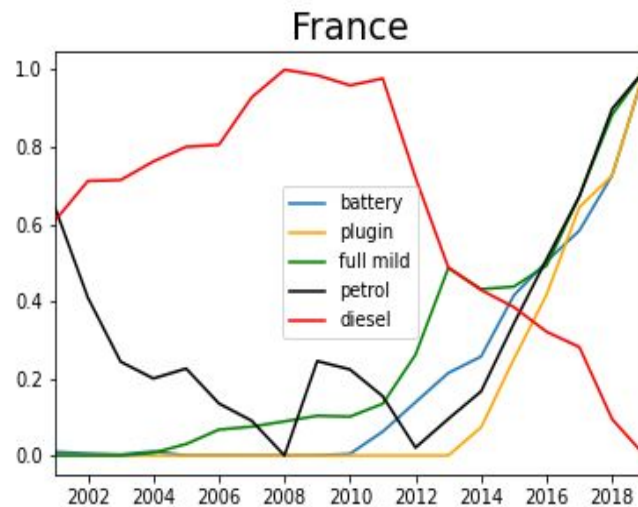


Histogram of petrol(France)

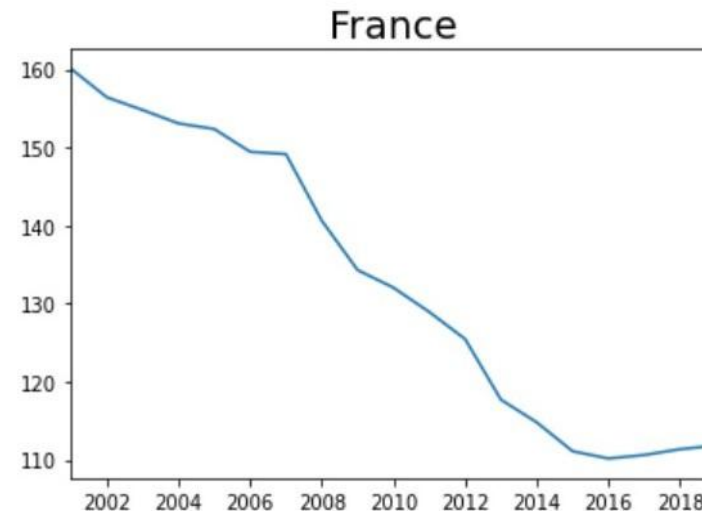


Histogram of diesel(France)





- All the new types of vehicles follows similar behaviour
- Diesel and Petrol cars have different trend



- CO2 emissions decrease until 2016, after which, in contrast to other countries, they seem to stabilise

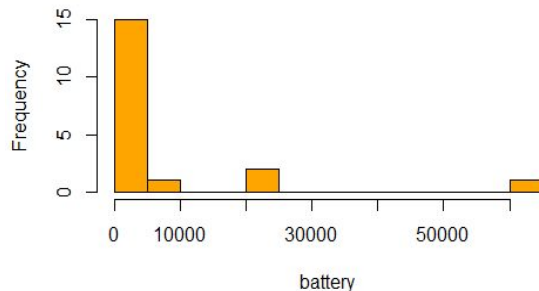
TYPES OF CARS: Netherlands

Distribution

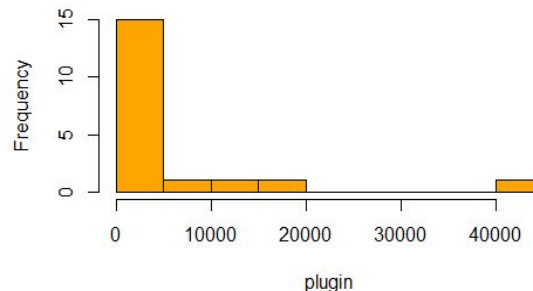


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

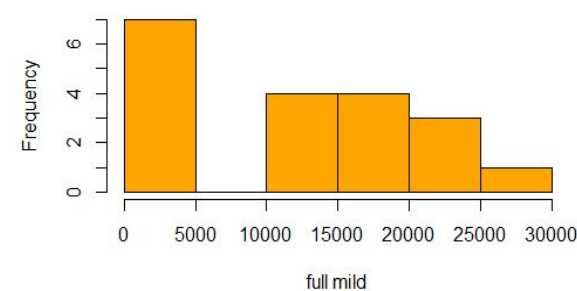
Histogram of battery(Netherlands)



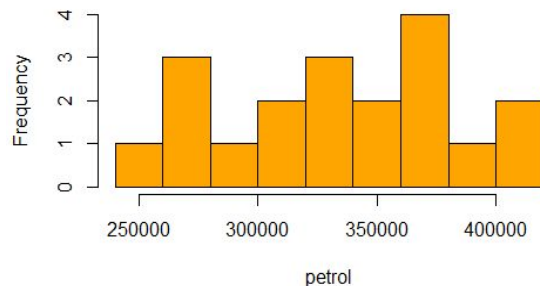
Histogram of plugin hybrid(Netherlands)



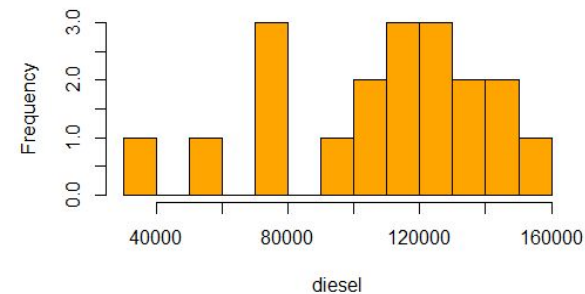
Histogram of full mild hybrid(Netherlands)

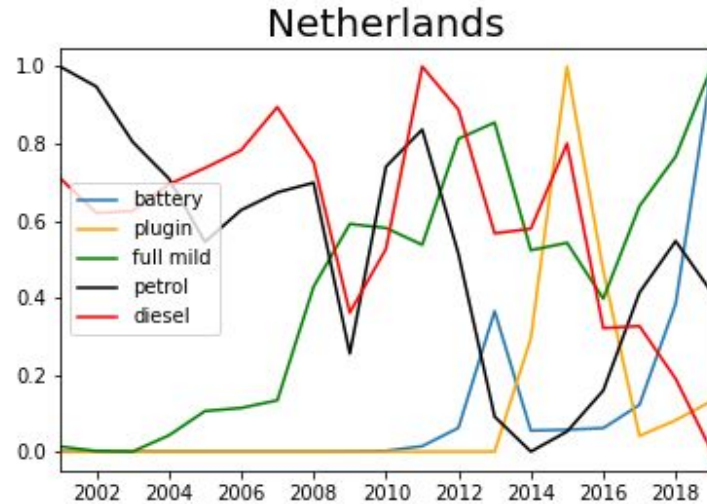


Histogram of petrol(Netherlands)

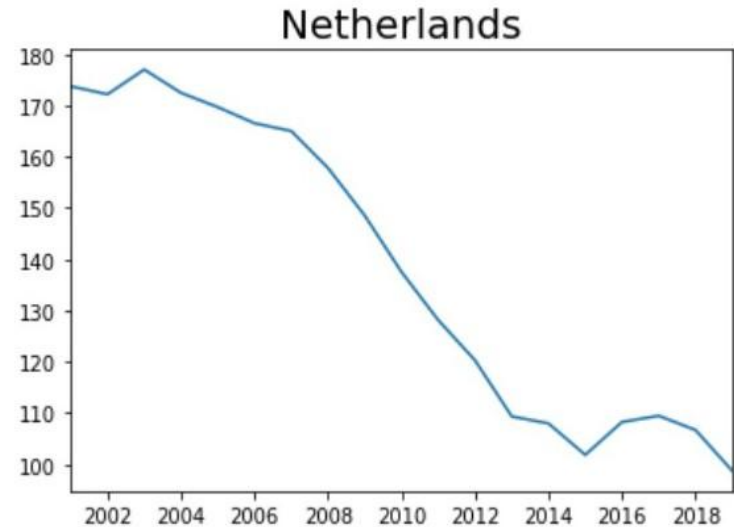


Histogram of diesel(Netherlands)





→ The different types of cars have an obvious main peak in their behaviour



→ CO2 emissions have a decreasing trend, but after 2015 until 2017 the trend changed and then went down again

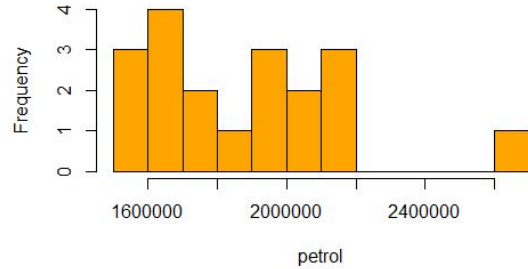
TYPES OF CARS: Germany

Distribution

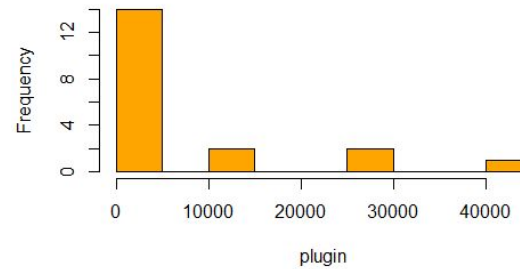


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

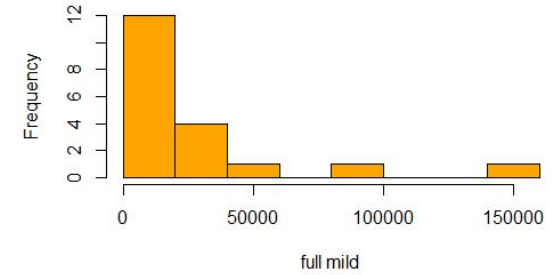
Histogram of petrol(Germany)



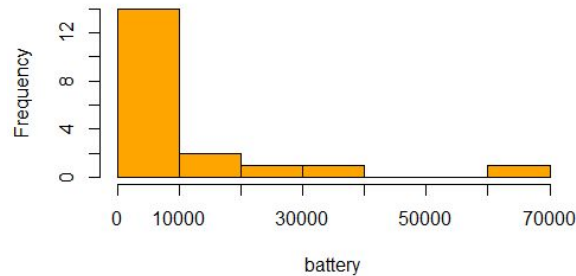
Histogram of plugin hybrid(Germany)



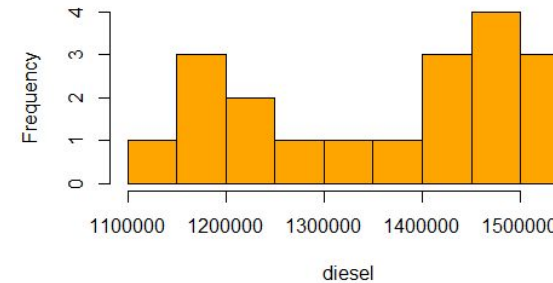
Histogram of full mild hybrid(Germany)

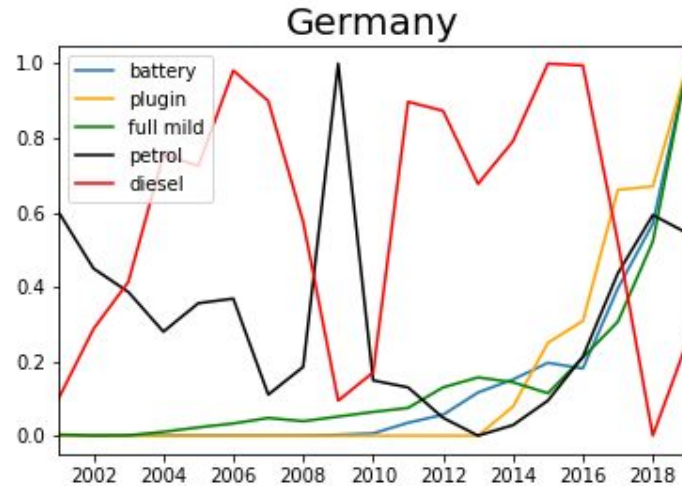


Histogram of battery(Germany)

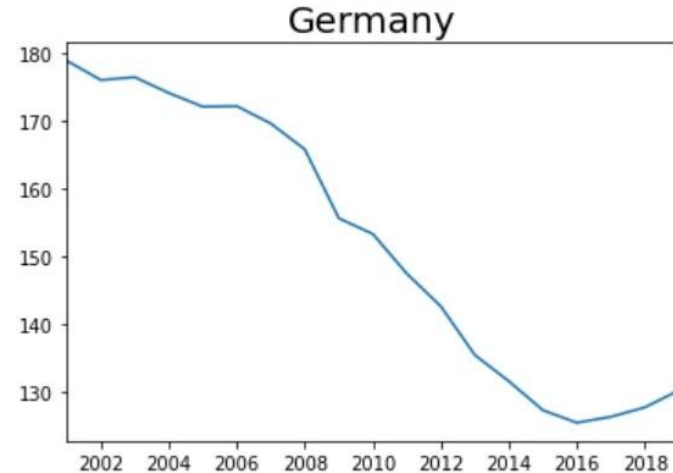


Histogram of diesel(Germany)





- After the global crisis of 2008, purchases of diesel cars plummeted, in contrast, petrol cars grew instantly
- After dieselgate (2015/16) we have another collapse in diesel car purchases



- CO2 emissions fall until 2016 where they start to rise again

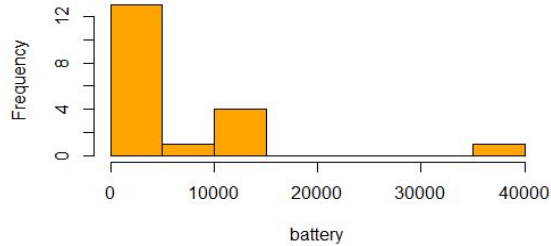
Types of cars: United Kingdom

Distribution

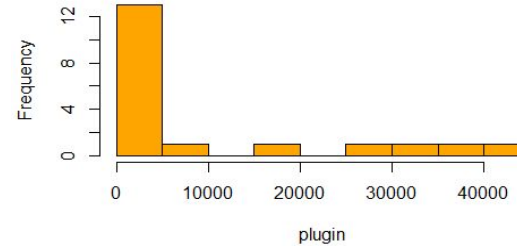


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

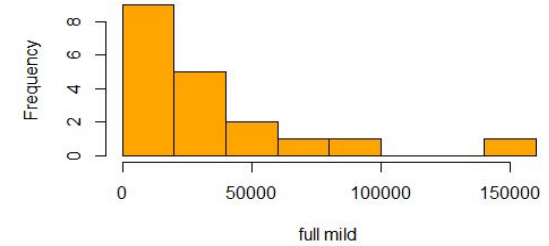
Histogram of battery(United Kingdom)



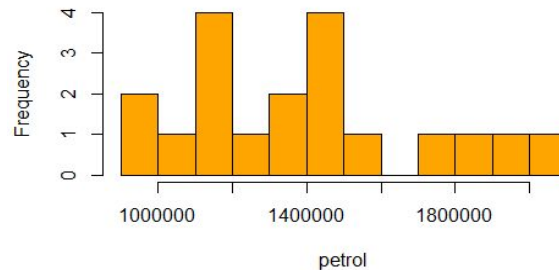
Histogram of plugin hybrid(United Kingdom)



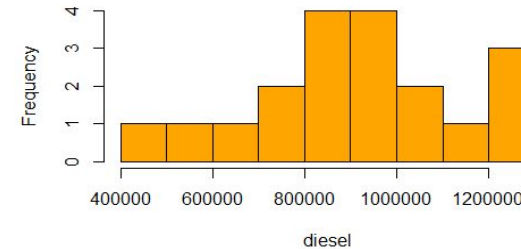
Histogram of full mild hybrid(United Kingdom)



Histogram of petrol(United Kingdom)



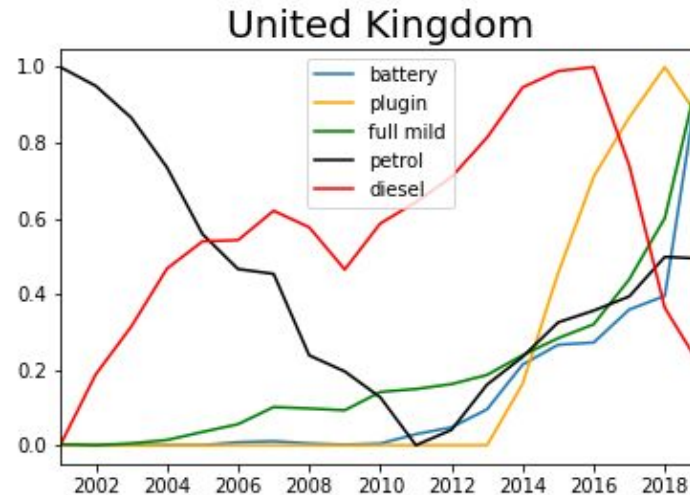
Histogram of diesel(United Kingdom)



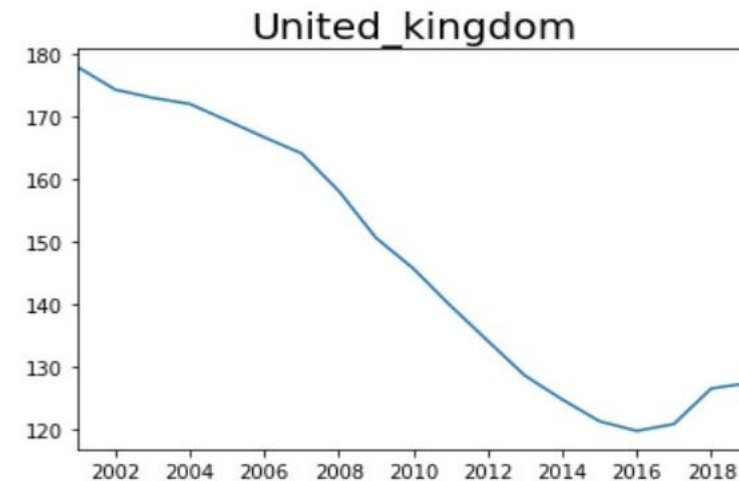
TYPE OF CARS AND CO2 EMISSION THROUGH TIME: UNITED KINGDOM



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



- The new types of cars have similar behaviour
- After 2011 the trend of petrol purchasing changed
- Diesel cars grows until the dieselgate, after that drop definitely down



- CO2 emissions fall until 2016 where they start to rise again, in the last three years seems to have stabilised

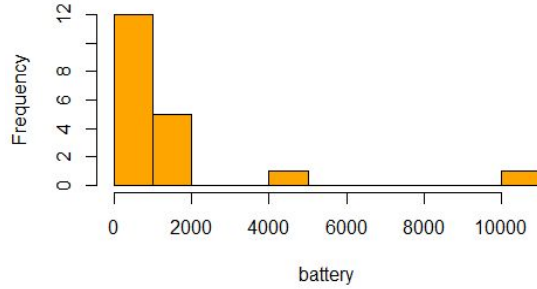
Types of cars: Italy

Distribution

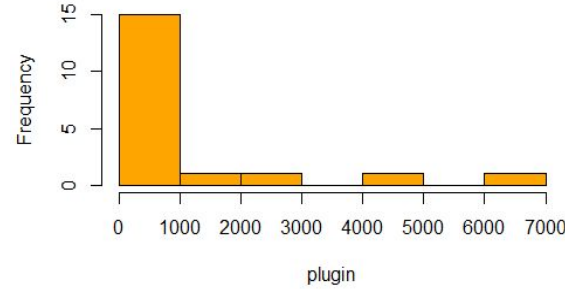


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

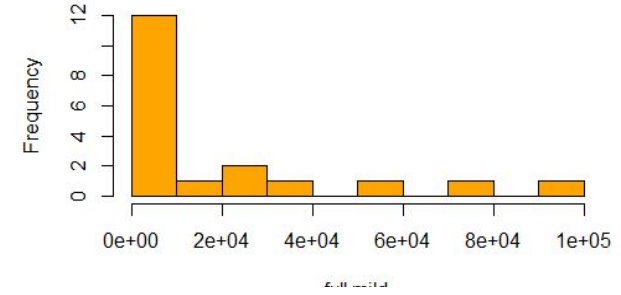
Histogram of battery(Italy)



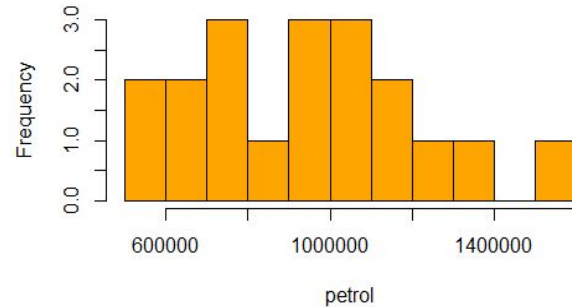
Histogram of plugin hybrid(Italy)



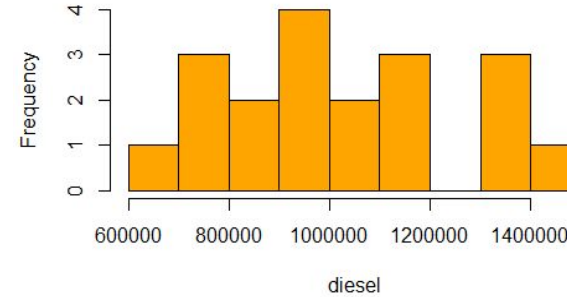
Histogram of full mild hybrid(Italy)

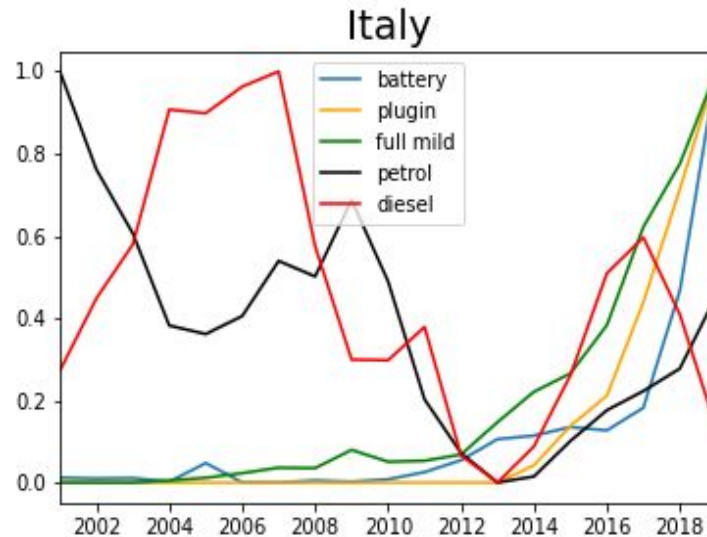


Histogram of petrol(Italy)

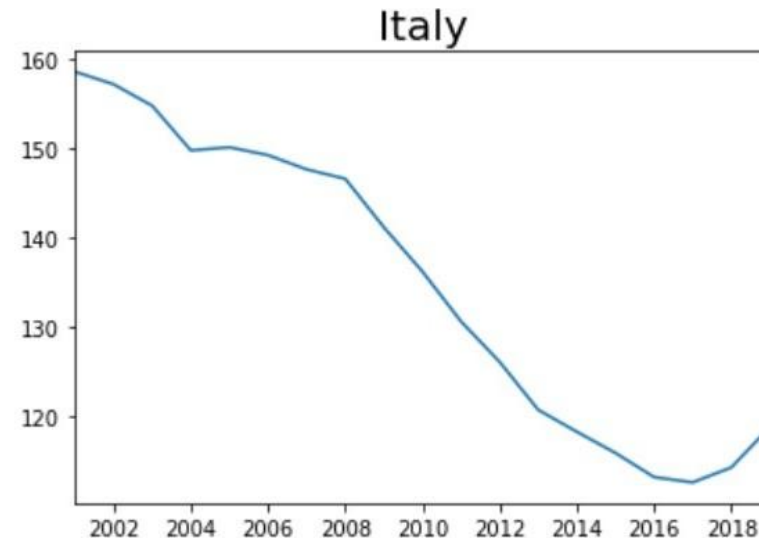


Histogram of diesel(Italy)





- The new types of cars have similarly behaviour
- Diesel have two main pick, the first decrease around 2007 and the second one around 2017



- Like other countries, CO2 emissions decrease until 2016/2017 and then start to increase again

- All the variables in every country have non-linear trend
- We do not have a seasonality phenomenon
- CO2 emissions for all countries have the same trend except for the last three years
- The two phenomena that had the greatest impact on car purchases were the global crisis of 2008 and the dieselgate of 2015

- ❖ Bass Model
- ❖ GGM
- ❖ ARIMA
- ❖ GAM

Bass Model & GGM

FRANCE: Battery



UNIVERSITA
DEGLI STUDI
DI PADOVA

We developed a diffusion model for each car type for different European countries. Approximate parameter values for the BASS models (m , p , q) can be observed in the table below.

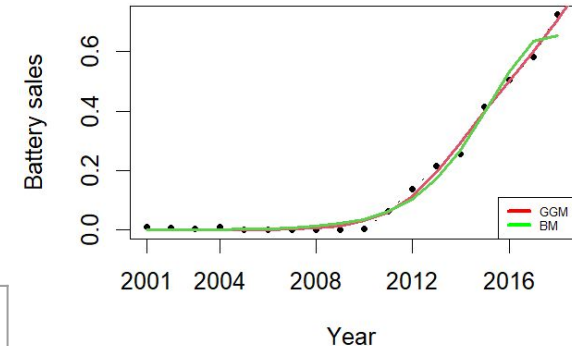
	m	p	q
BM	4.753751e+00	3.864789e-05	5.571326e-01

To get better results, we also developed a GGM model and approximate parameter values can be observed in the table below

	k	pc	qc	ps	qs
GGM	1.180291e+01	6.478135e-06	4.795843e-01	2.374147e-04	5.794605e-01

All parameters resulted significant for the simple BASS model.

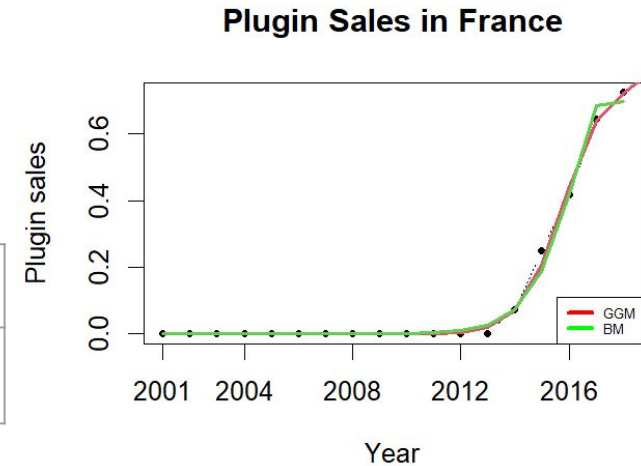
Battery Sales in France



All parameters resulted significant for the simple BASS model.

	m	p	q
BM	2.859422e+00	1.630577e-08	1.055894e+00

	k	pc	qc	ps	qs
GGM	7.772111e+00	1.234704e-06	5.957730e-01	3.646618e-08	1.094724e+00



Bass Model & GGM

FRANCE: Full Mild

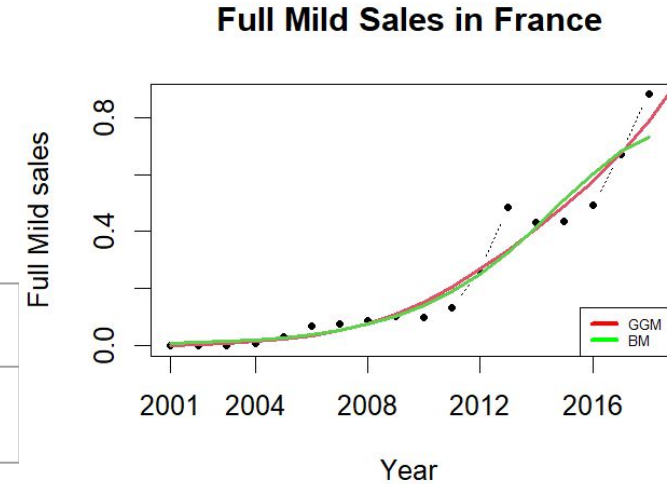


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

All parameters resulted significant for the simple BASS model.

	m	p	q
BM	8.8522320272	0.0007085541	0.3363436357

	k	pc	qc	ps	qs
GGM	91.43147770	0.002681304	0.415894221	0.000332252	0.187997569

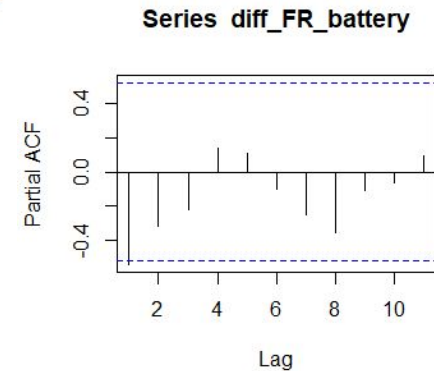
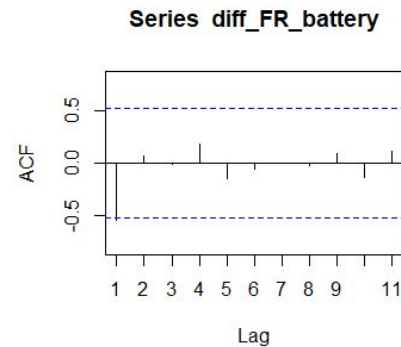
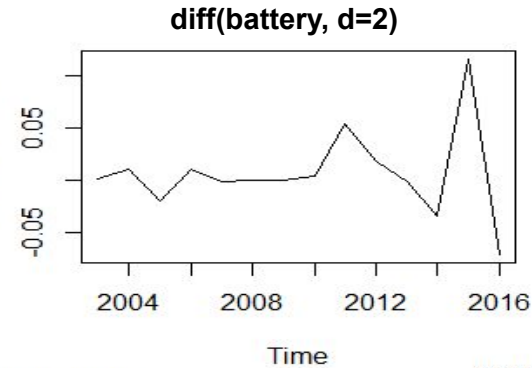
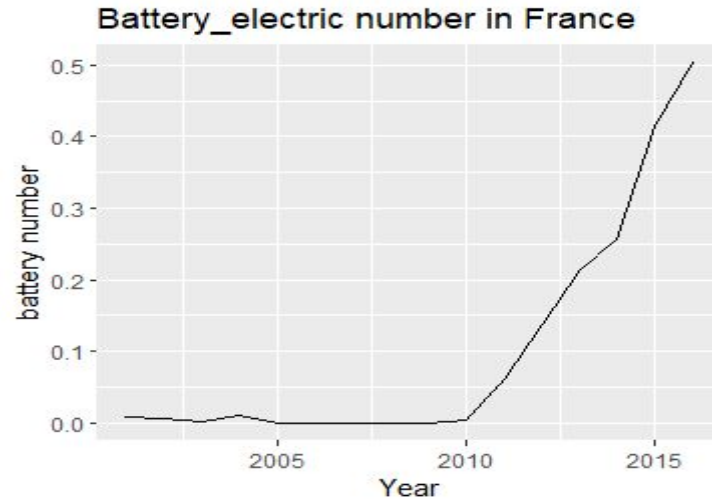


ARIMA: Time Series

FRANCE : Battery

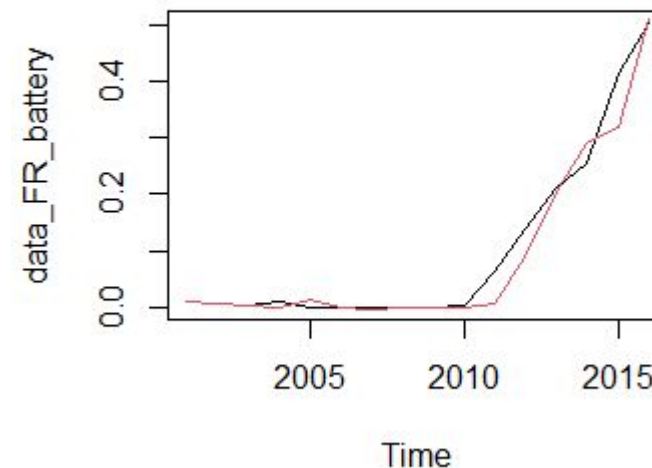


UNIVERSITÀ
DEGLI STUDI
DI PADOVA



Arima with $d = 2$: looking for best p and q .

Model	AICc
ARIMA(0,2,0)	-47.49646
ARIMA(1,2,0)	-49.17481
ARIMA(0,2,1)	-48.27529
ARIMA(2,2,0)	-45.88223
ARIMA(1,2,1)	-45.87817



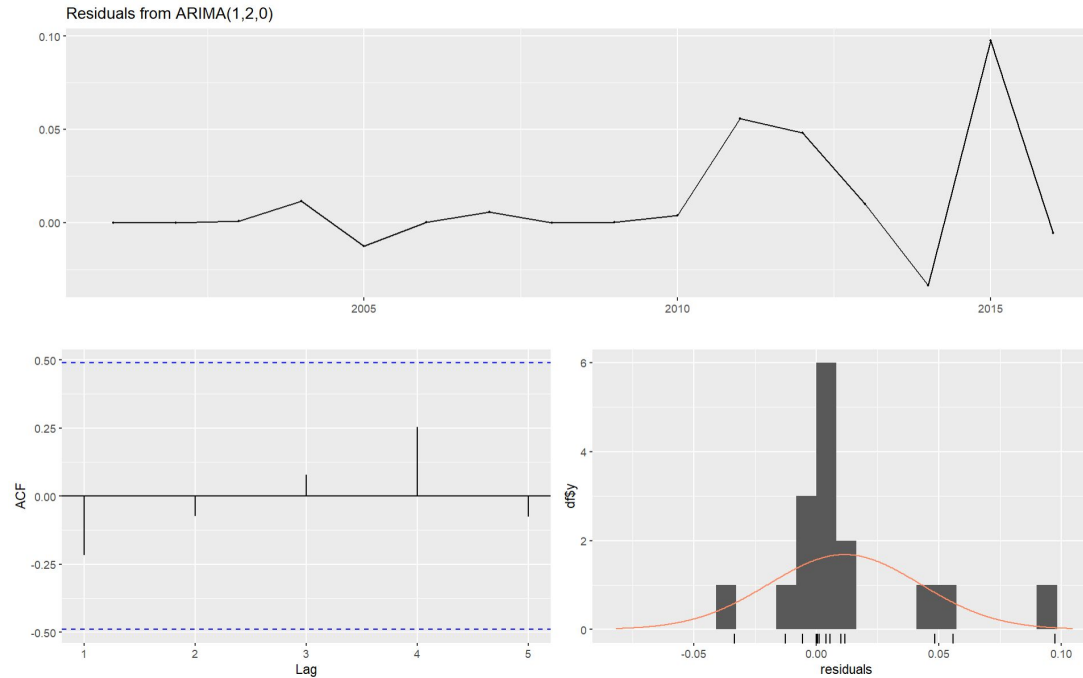
We compared some handpicked ARIMA models to those suggested by `auto.arima` in R on the basis of the AIC.

ARIMA: Residuals

FRANCE : Battery

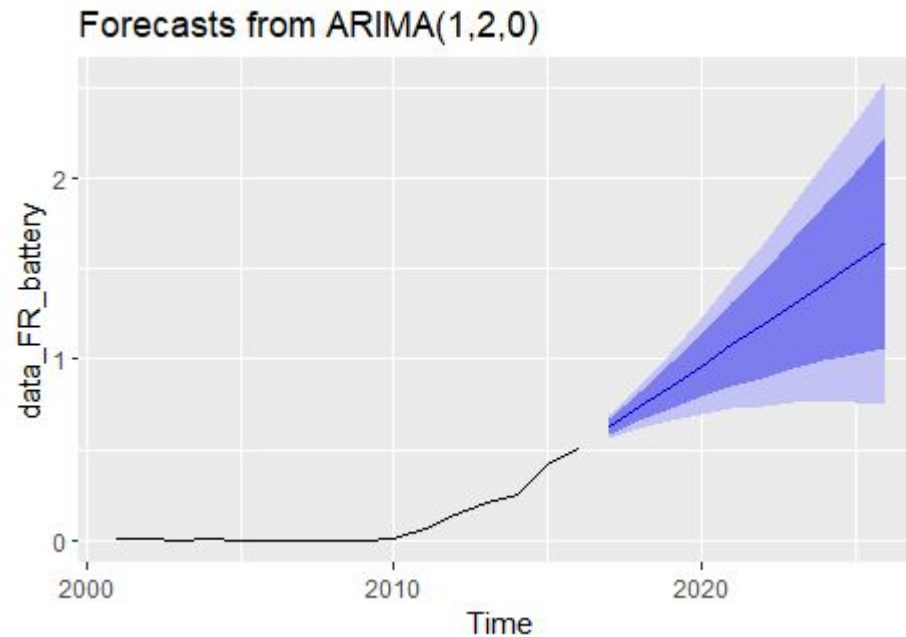


UNIVERSITÀ
DEGLI STUDI
DI PADOVA



→ Residuals are completely white noise

Forecast with the selected model



Comparison of accuracy on training and test set: We considered 16 years forms 2001 to 2016 as our training dataset and the last three years as a test set.

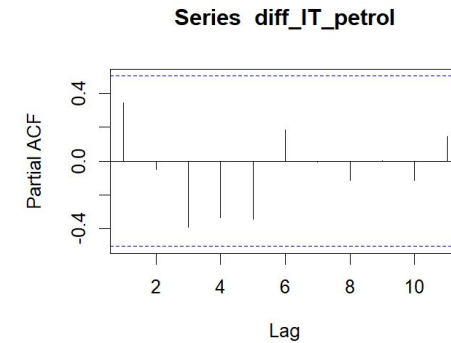
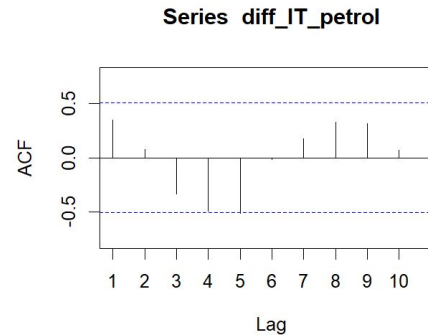
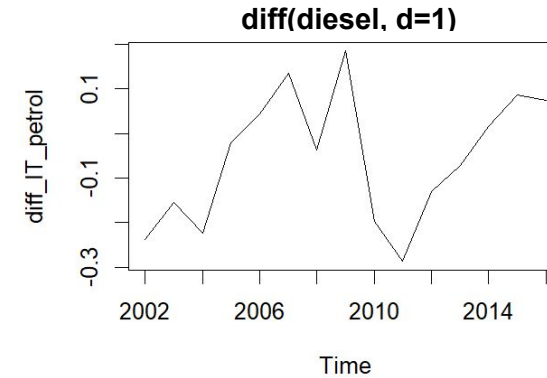
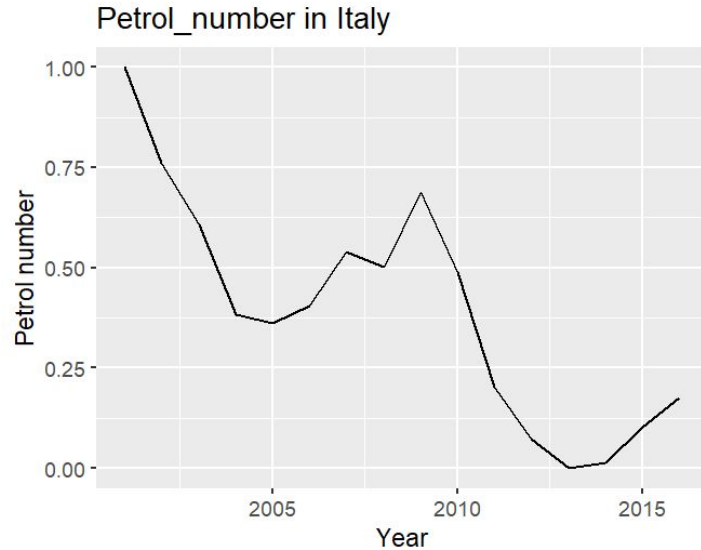
	RMSE	MAE
Train Set	0.03215906	0.01789065
Test Set	0.08881124	0.06797979

ARIMA: Time Series

Italy : Petrol

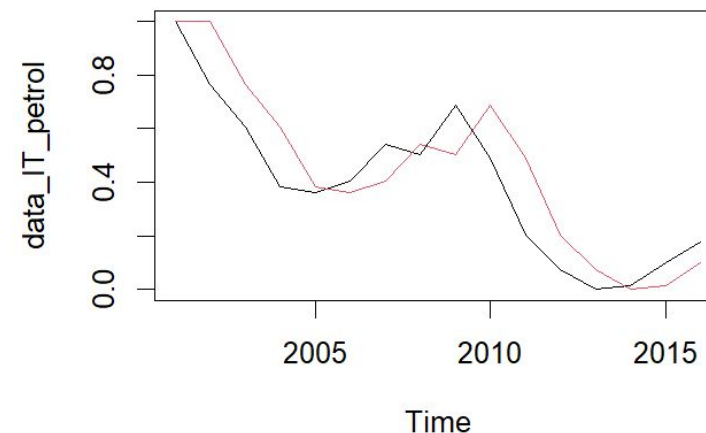


UNIVERSITÀ
DEGLI STUDI
DI PADOVA



Arima with $d = 1$: looking for best p and q .

Model	AICc
ARIMA(0,1,0)	-11.15857
ARIMA(1,1,0)	-10.14887
ARIMA(0,1,1)	-9.695889
ARIMA(1,1,1)	-6.334253



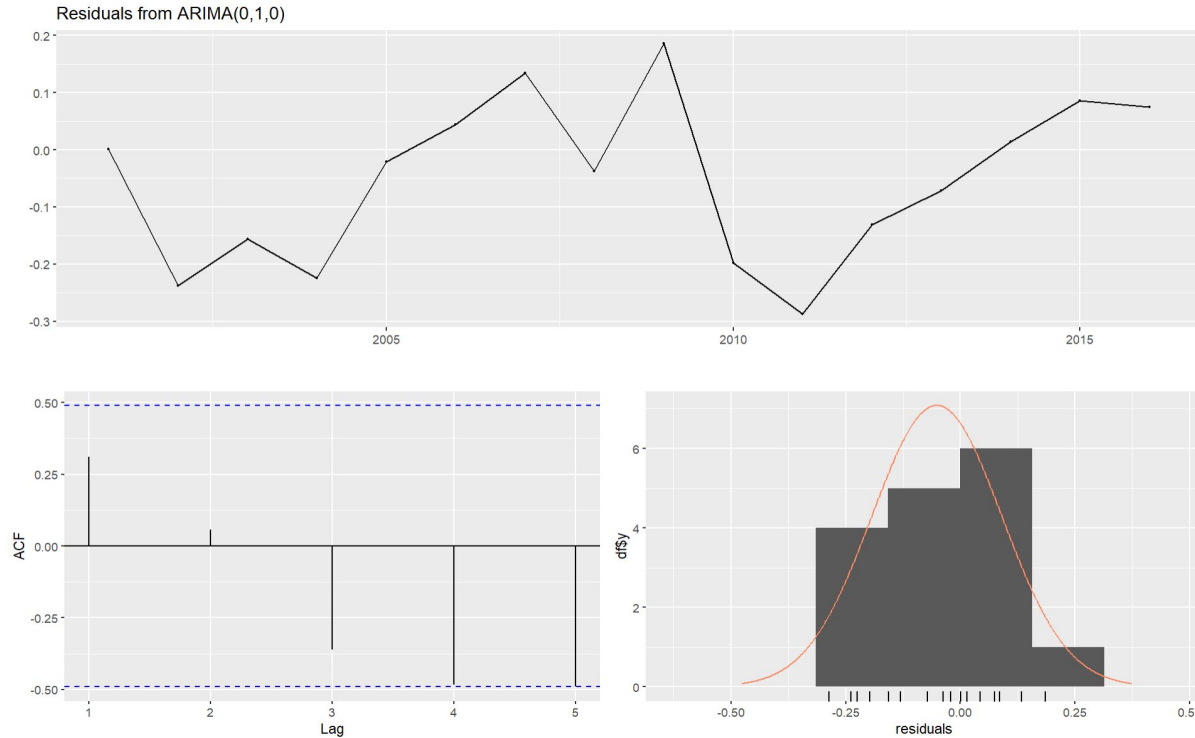
We compared some handpicked ARIMA models to those suggested by `auto.arima` in R on the basis of the AIC.

ARIMA: Residuals

ITALY: Petrol

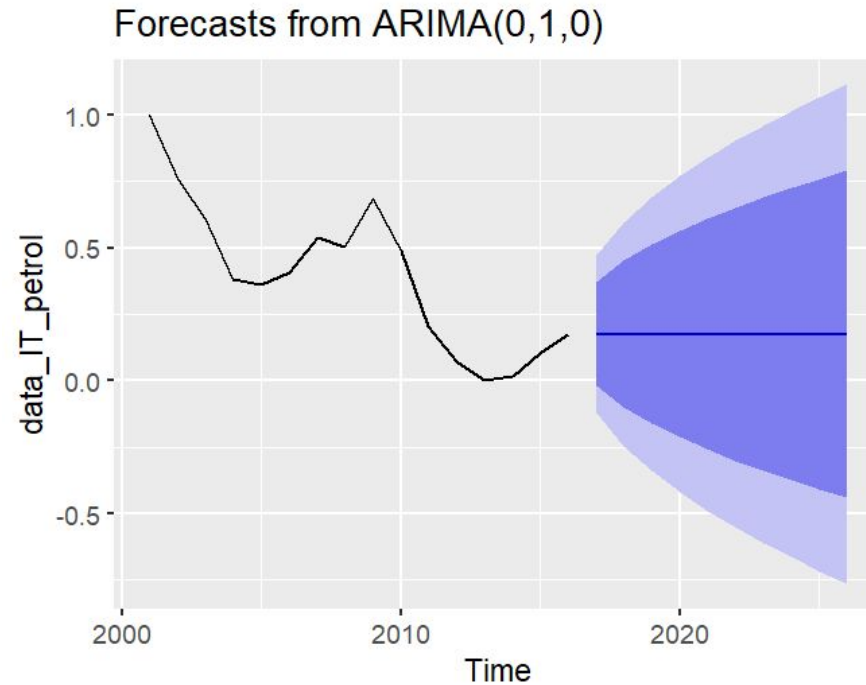


UNIVERSITÀ
DEGLI STUDI
DI PADOVA



→ Residuals are completely white noise

Forecast with the selected model



Comparison of accuracy on training and test set:

	RMSE	MAE
Train Set	0.1467028	0.1190060
Test Set	0.1734884	0.1423246

Further Analysis: CO₂ Emission

GAM

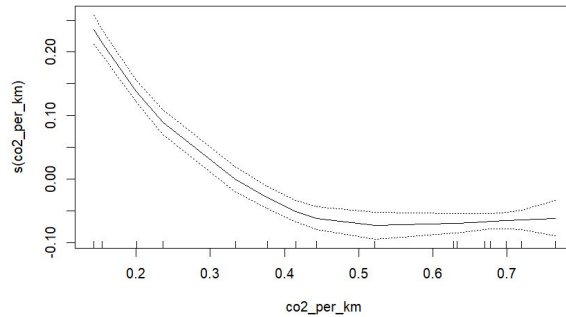
GAM: with spline

France

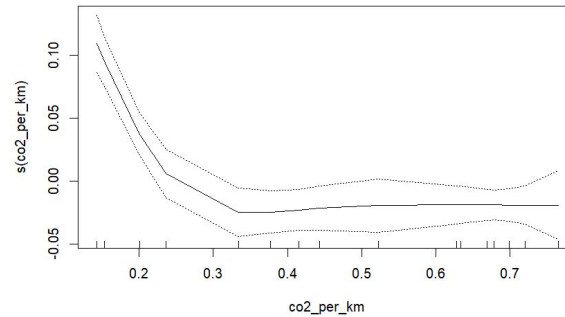


UNIVERSITA
DEGLI STUDI
DI PADOVA

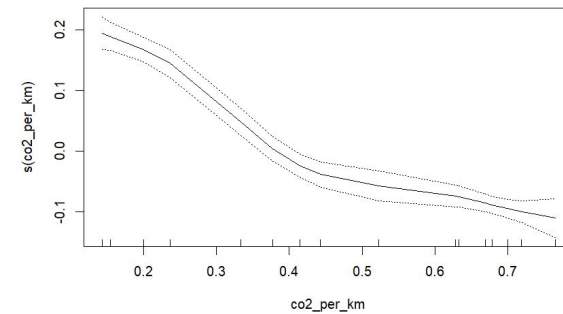
Battery Electric GAM



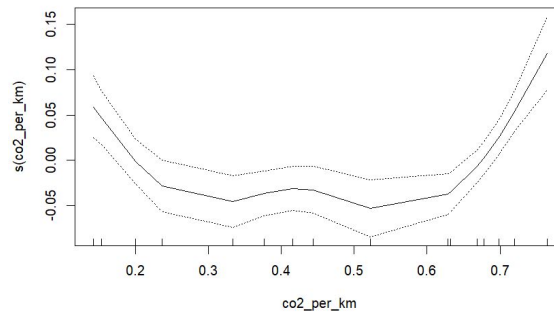
Plugin Hybrid GAM



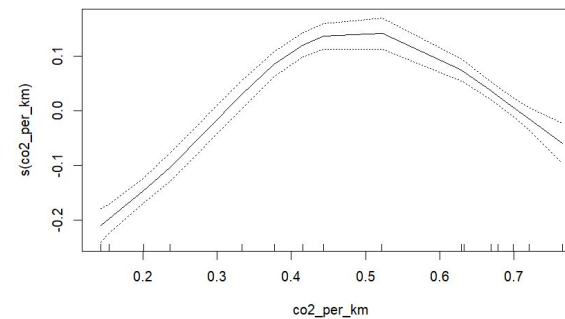
Full Mild GAM



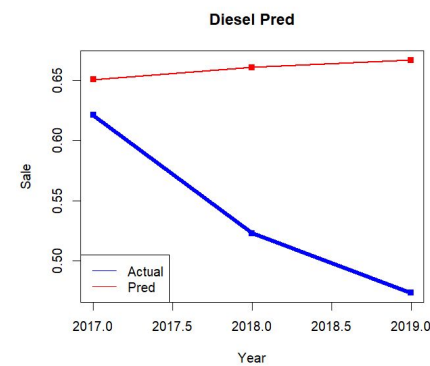
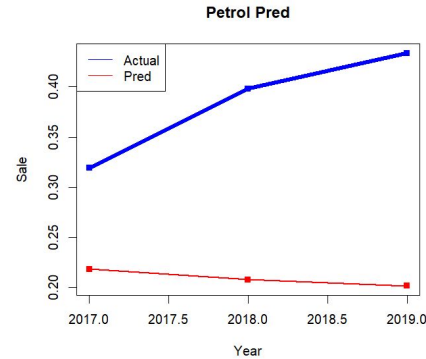
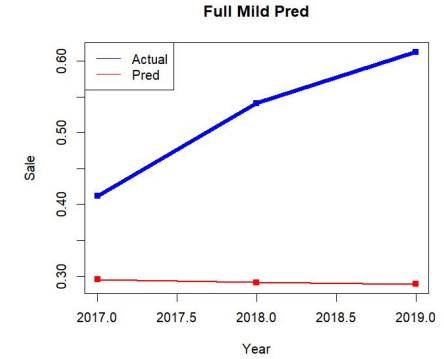
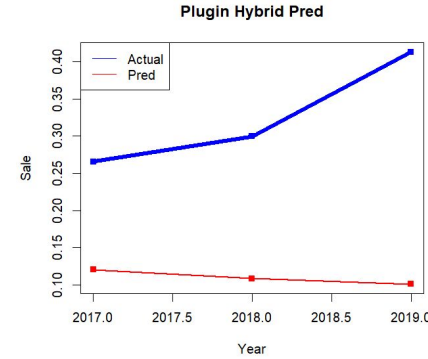
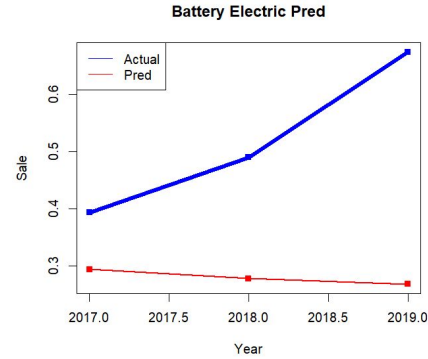
Petrol GAM



Diesel Gas GAM



	GAM with Spline on Test Set	
Car Type	DEV	RMSE
Battery Electric	0.221	0.271
Plugin Hybrid	0.221	0.227
Full Mild Hybrid	0.181	0.246
Petrol	0.101	0.183
Diesel Gas	0.057	0.138



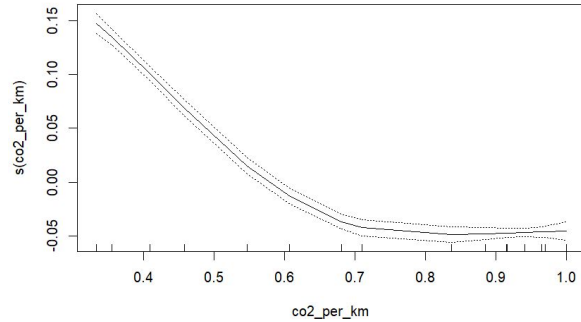
GAM: with spline

Germany

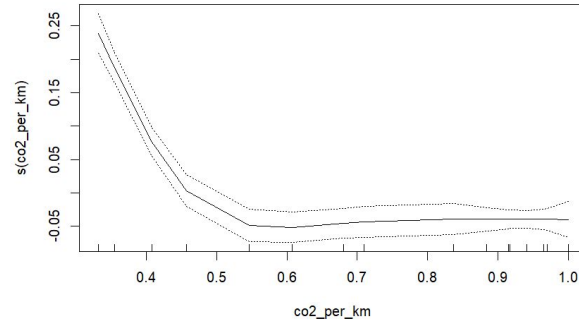


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

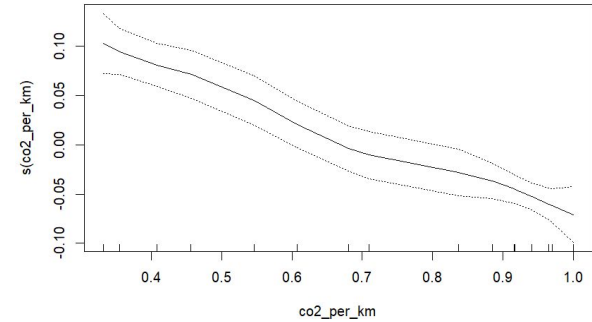
Battery Electric GAM



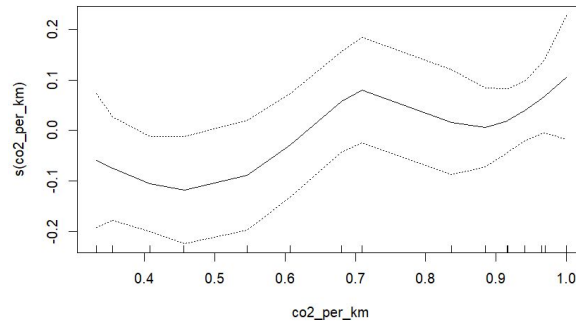
Plugin Hybrid GAM



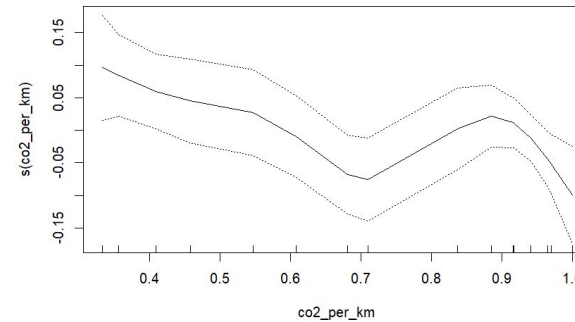
Full Mild GAM



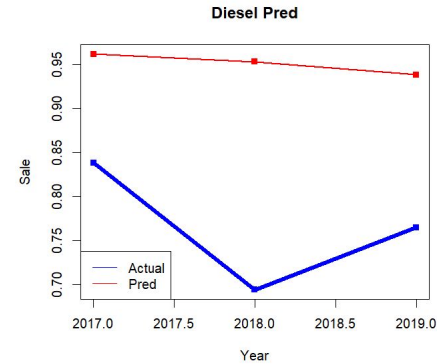
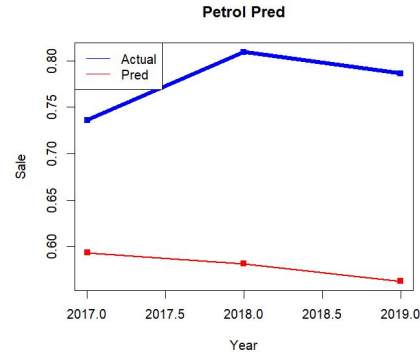
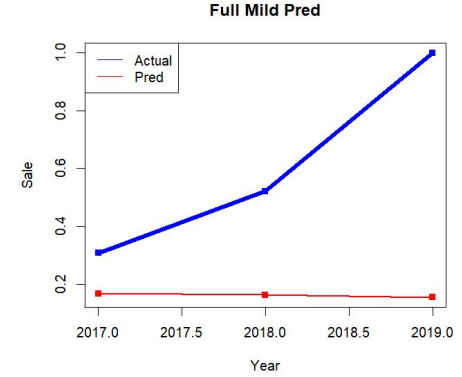
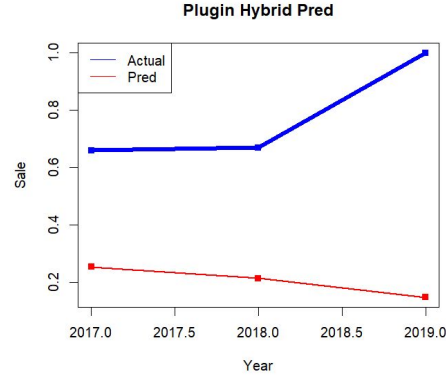
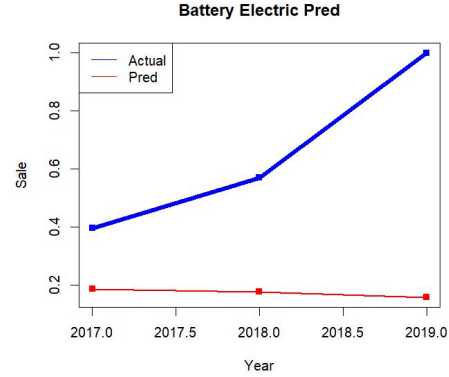
Petrol GAM



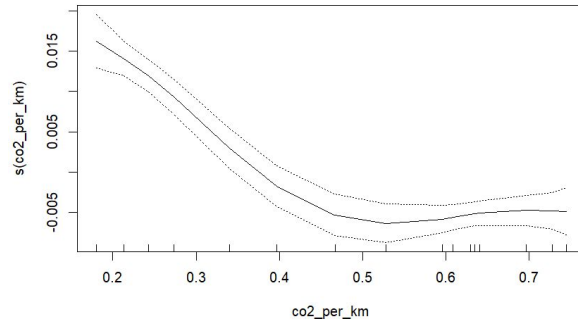
Diesel Gas GAM



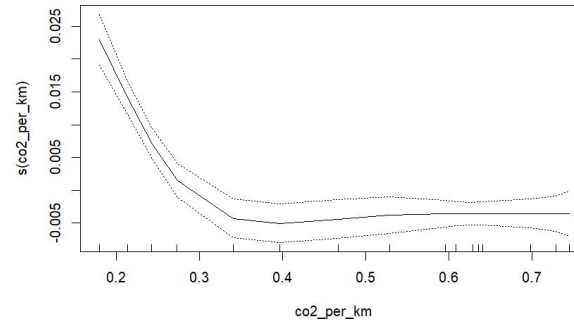
	GAM with Spline on Test Set	
Car Type	DEV	RMSE
Battery Electric	0.906	0.550
Plugin Hybrid	1.098	0.605
Full Mild Hybrid	0.864	0.537
Petrol	0.122	0.202
Diesel Gas	0.112	0.193



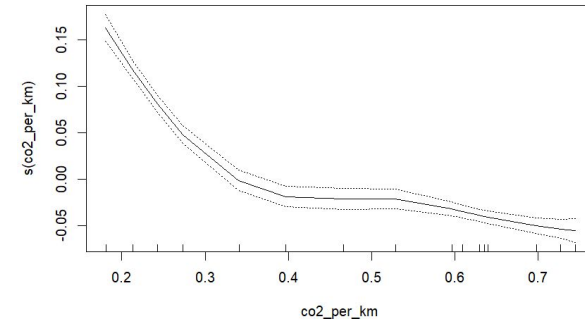
Battery Electric GAM



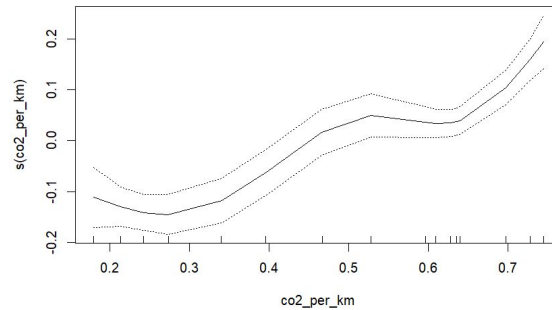
Plugin Hybrid GAM



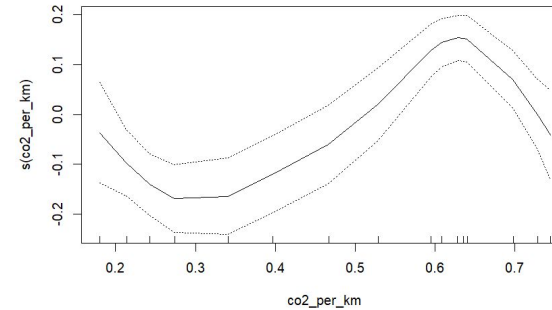
Full Mild GAM



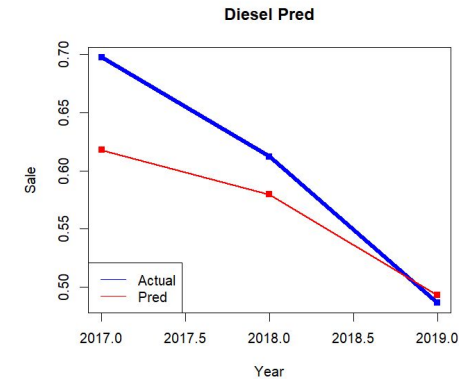
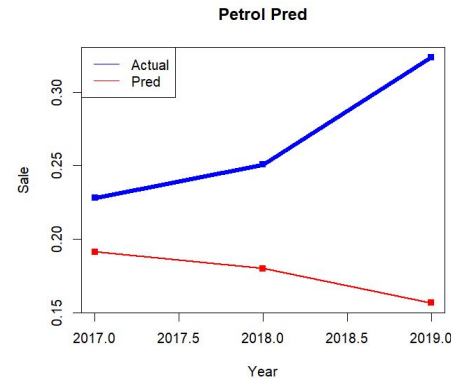
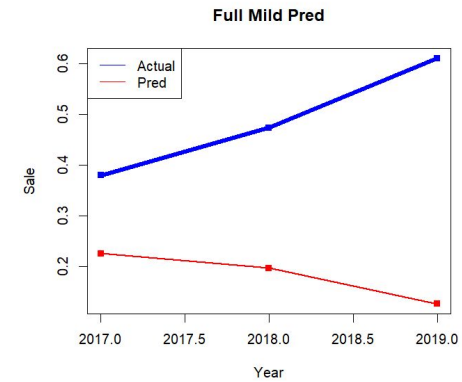
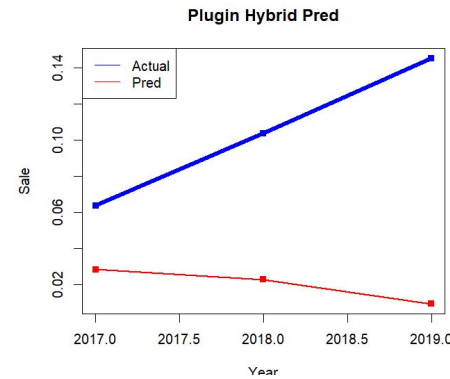
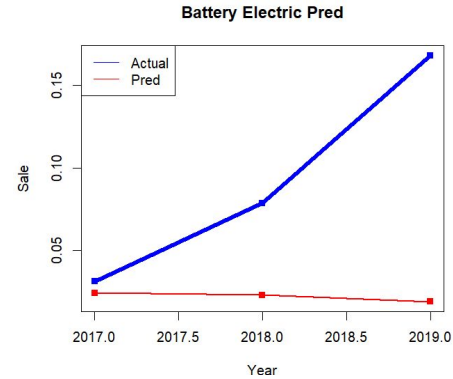
Petrol GAM



Diesel Gas GAM



	GAM with Spline on Test Set	
Car Type	DEV	RMSE
Battery Electric	0.026	0.092
Plugin Hybrid	0.026	0.094
Full Mild Hybrid	0.335	0.334
Petrol	0.034	0.107
Diesel Gas	0.007	0.050



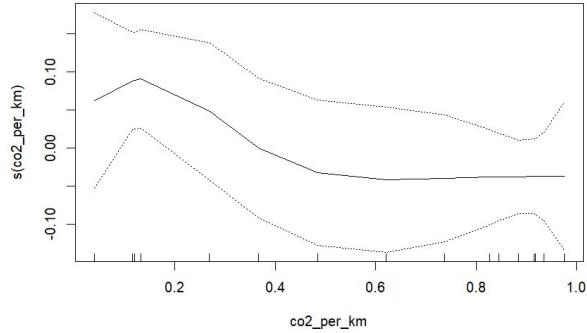
GAM: with spline

Netherlands

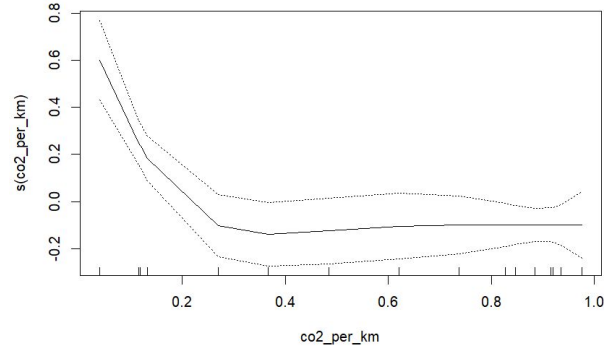


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

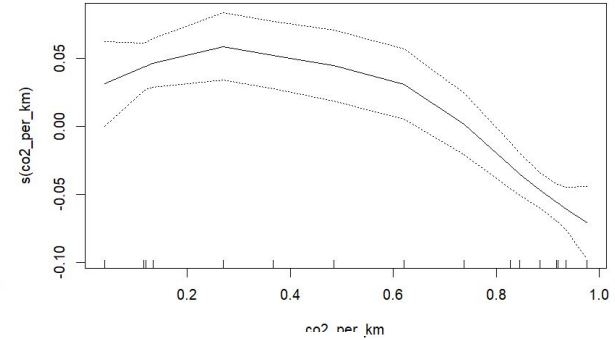
Battery Electric GAM



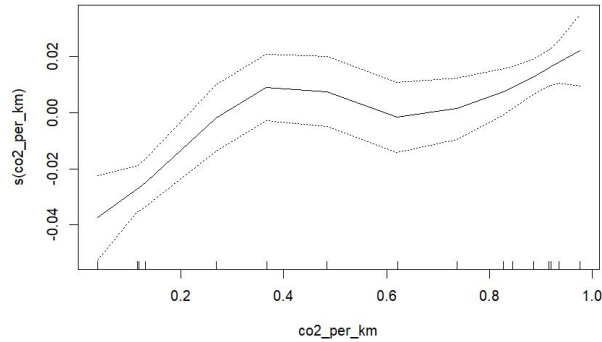
Plugin Hybrid GAM



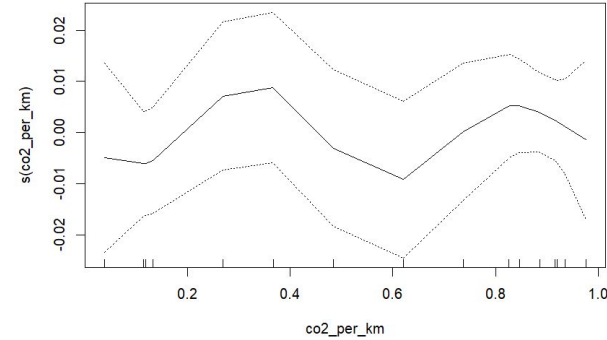
Full Mild GAM



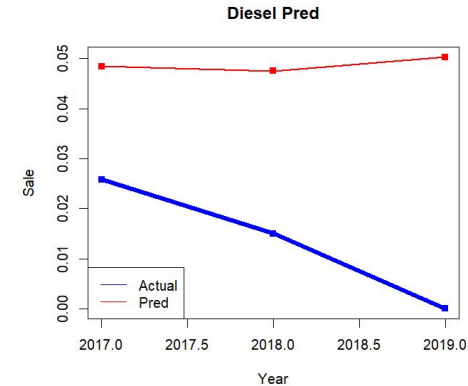
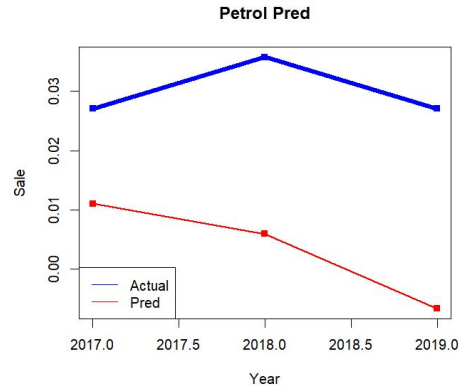
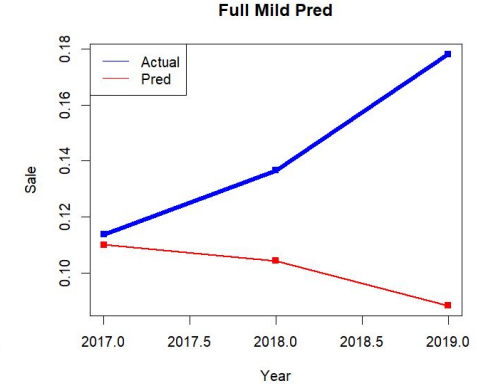
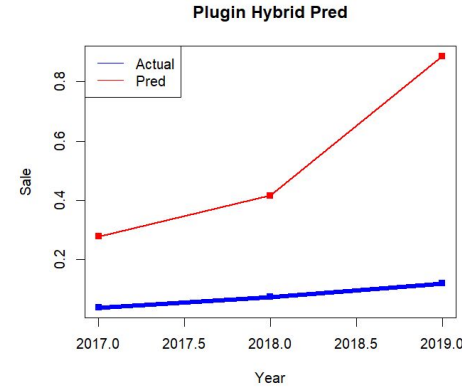
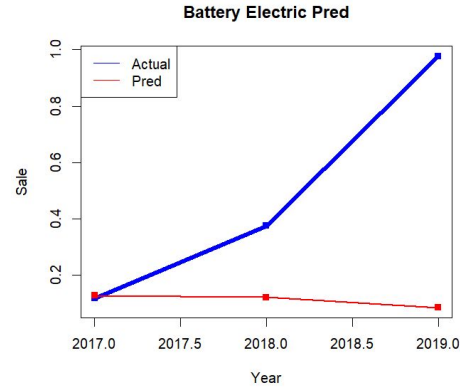
Petrol GAM



Diesel Gas GAM



	Test	
Car Type	DEV	RMSE
Battery Electric	0.862	0.536
Plugin Hybrid	0.764	0.505
Full Mild Hybrid	0.009	0.055
Petrol	0.002	0.028
Diesel Gas	0.004	0.370



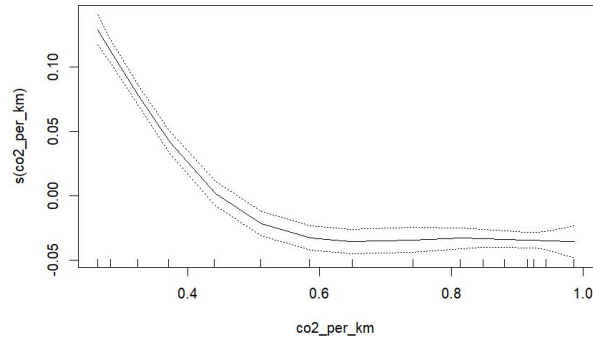
GAM: with spline

United Kingdom

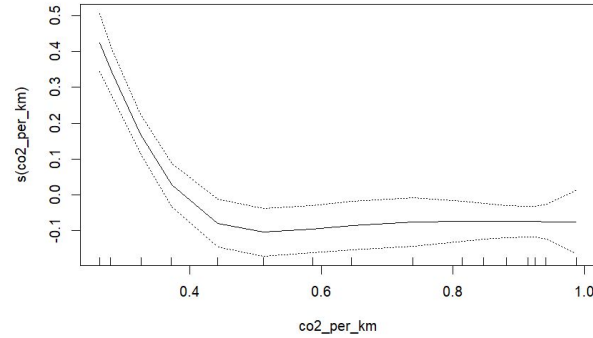


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

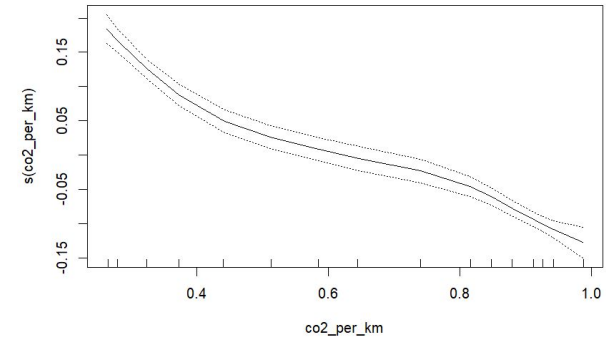
Battery Electric GAM



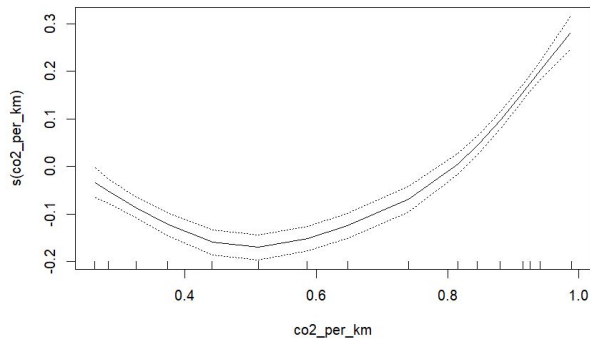
Plugin Hybrid GAM



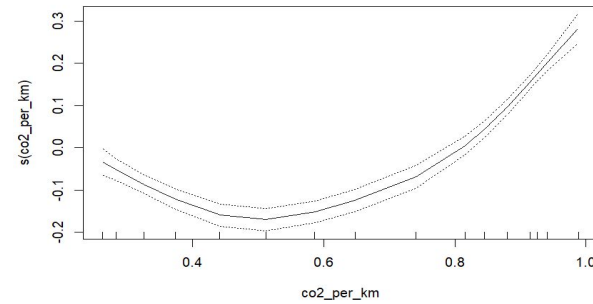
Full Mild GAM



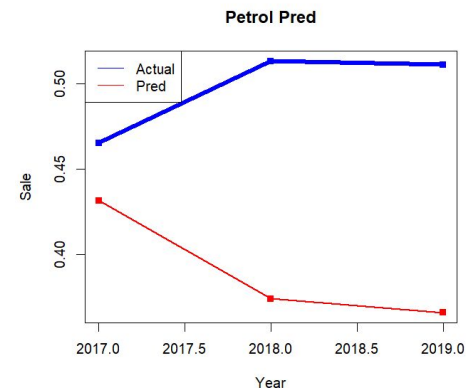
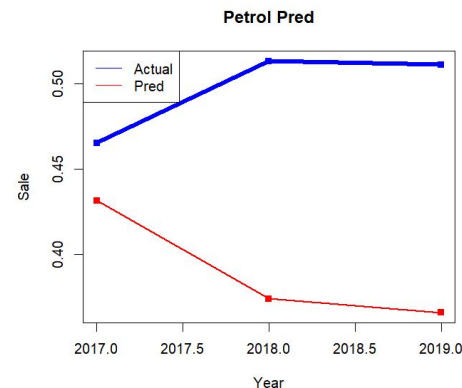
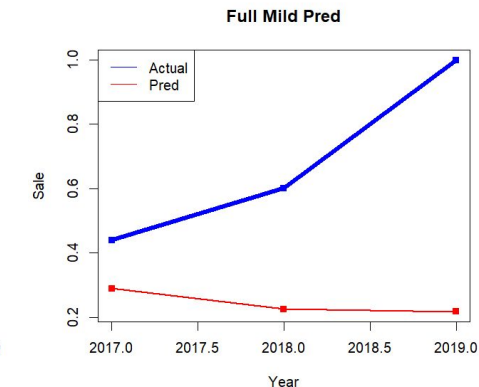
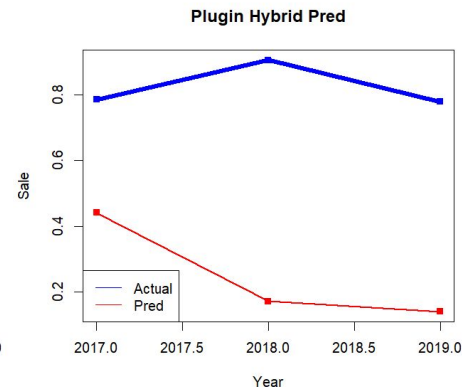
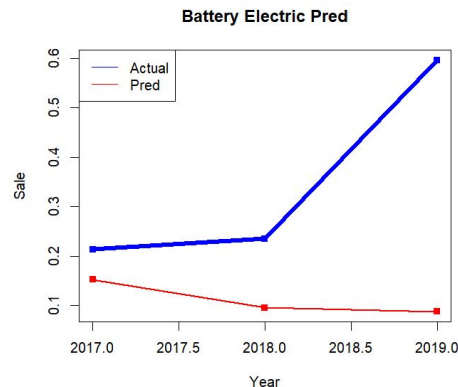
Petrol GAM



Petrol GAM

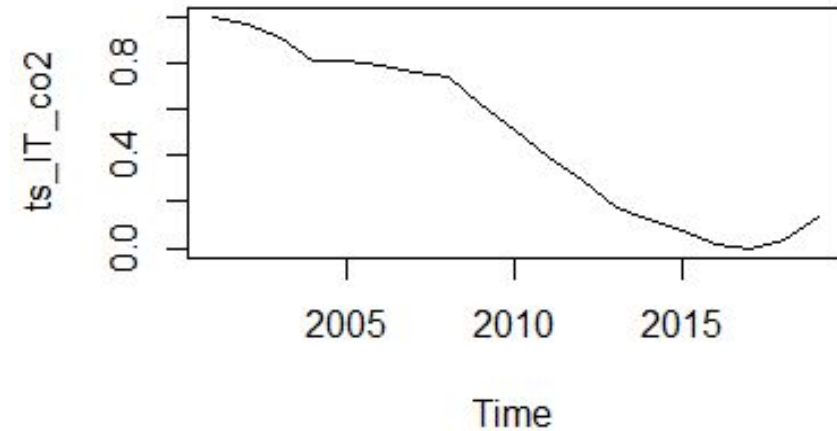
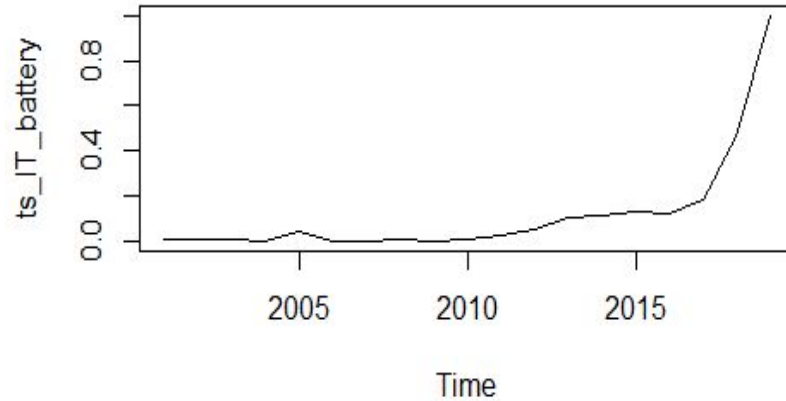


	Test	
Car Type	DEV	RMSE
Battery Electric	0.282	0.307
Plugin Hybrid	1.062	0.595
Full Mild Hybrid	0.773	0.508
Petrol	0.042	0.118
Diesel Gas	0.220	0.271



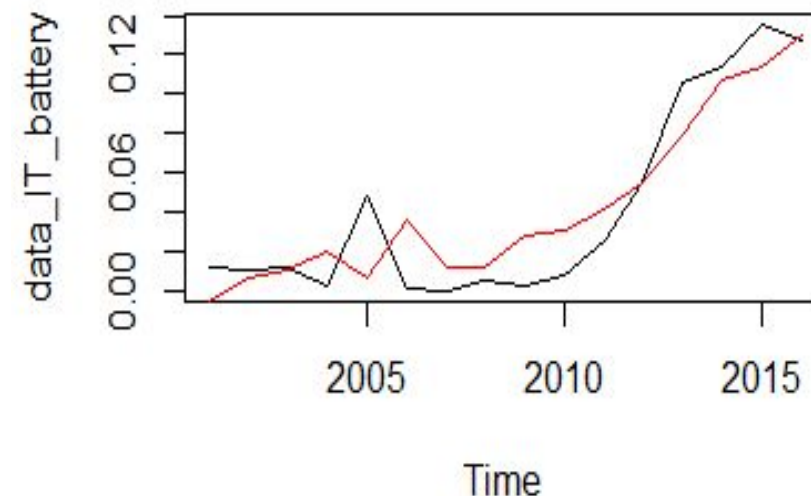
ARIMA

- First, we analysed the time series to understand the possible relationships.
- We used the `auto.arima` model to find the best p , d and q parameters.
- We applied the arima model to see how co_2 emissions affect car purchases, using emissions as an external regressor via the `xreg` parameter.
- Once the model was trained, we analysed the residuals to observe its behaviour.
- Finally, We tested the model with the test set to evaluate its performance.

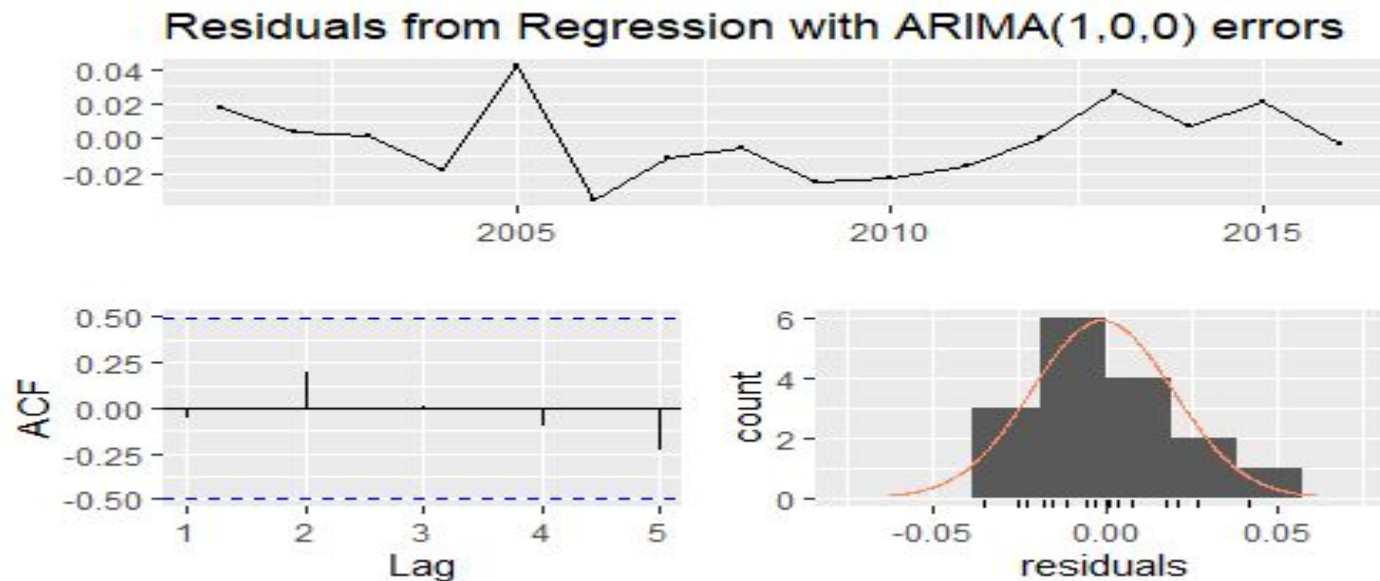


- The two time series are co2 emissions and purchases of electric cars

MODEL	AICc
ARIMA(1,0,0)	-67,921260
ARIMA(0,0,0)	-65,290870
ARIMA(0,0,1)	-65,653510
ARIMA(2,0,0)	-63,779640



→ The best model is with $p=1$, $d=0$ and $q=0$



→ The residual is completely white noise

We used Root Mean Square Error and Mean Absolute Error values to evaluate performance in the training and test set.

	RMSE	MAE
Training Set	0,019883	0,015924
Test Set	0,644481	0,549354

OTHER RESULTS: FRANCE



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

VARIABLES	MODEL	AICc	RMSE training set	RMSE test set	MAE training set	MAE test set
Electric battery	ARIMA(2,0,0)	-43,027640	0.0347278	0.607405	0.020603	0.579018
Plugin hybrid	ARIMA(1,0,0)	-31,186100	0.61179	0.704783	0.043007	0.686985
Full mild hybrid	ARIMA(0,0,1)	-39,720710	0.047494	0.882553	0,037607	0,874089
Petrol	ARIMA(1,0,0)	-9,677453	0,132587	0,241916	0,109864	0,218217
Diesel	ARIMA(2,0,0)	-18,973320	0,083481	0,566007	0,062650	0,561493

OTHER RESULTS: NETHERLANDS



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

VARIABLES	MODEL	AICc	RMSE training set	RMSE test set	MAE training set	MAE test set
Electric battery	ARIMA(0,0,0)	-30,461870	0,072739	0,636299	0,040811	0,516937
Plugin hybrid	ARIMA(0,0,1)	2,172297	0,176123	0,229892	0,115489	0,222016
Full mild hybrid	ARIMA(1,0,0)	-9,781617	0,121634	0,842795	0,084614	0,832427
Petrol	ARIMA(0,0,0)	-0,286537	0,205617	0,388654	0,158554	0,380093
Diesel	ARIMA(0,0,0)	-2,088918	0,176538	0,553477	0,138877	0,536718

OTHER RESULTS: UNITED KINGDOM



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

VARIABLES	MODEL	AICc	RMSE training set	RMSE test set	MAE training set	MAE test set
Electric battery	ARIMA(2,0,0)	-58,399710	0,021785	0,665337	0,014762	0,582284
Plugin hybrid	ARIMA(1,0,1)	-23,069610	0,062417	0,583190	0,046476	0,579584
Full mild hybrid	ARIMA(0,0,1)	-66,754450	0,020508	0,704948	0,016775	0,664023
Petrol	ARIMA(2,0,0)	-22,702930	0,076185	0,536578	0,062663	0,516520
Diesel	ARIMA(1,0,0)	-20,513070	0,086338	0,305006	0,068497	0,227025

OTHER RESULTS: GERMANY



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

VARIABLES	MODEL	AICc	RMSE training set	RMSE test set	MAE training set	MAE test set
Electric battery	ARIMA(2,0,0)	-73,898560	0,013614	0,724602	0,010090	0,673741
Plugin hybrid	ARIMA(1,0,0)	-40,781990	0,045519	0,752507	0,032307	0,734197
Full mild hybrid	ARIMA(0,0,0)	-71,045820	0,020463	0,670305	0,015447	0,605565
Petrol	ARIMA(0,0,0)	-1,810891	0,219545	0,123738	0,135715	0,100861
Diesel	ARIMA(0,0,1)	9,412217	0,220612	0,292694	0,147396	0,243226

- We can see that the best model for understanding sales is the GGM model, this is especially true for new technologies as we are still in a phase of information dissemination, which pushes imitator customers to follow innovator consumers, thus not following the values of the past making the arima model less performing
- CO2 emissions influence purchasing trends, this is due to the fact that the higher the emissions, the more information reaches the public, thus prompting consumers to choose less polluting cars

END