An exploratory data analysis of energy consumption of WWTP. Influencing factors and possible methods for benchmarking

<u>Miguel Mauricio-Iglesias</u>, Stefano Longo, Juan M. Lema, Almudena Hospido

Department of Chemical Engineering
Universidade de Santiago de Compostela (Spain)

BioGroup

www.usc.es/biogrup

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Introduction. Why and how energy efficiency in WWTP ?

Water and wastewater treatment are large energy consumers

In Germany [1], WWTP consume around 1% of the electricity consumption

In Italy [2] WWTP consume around 1% of the electricity consumption

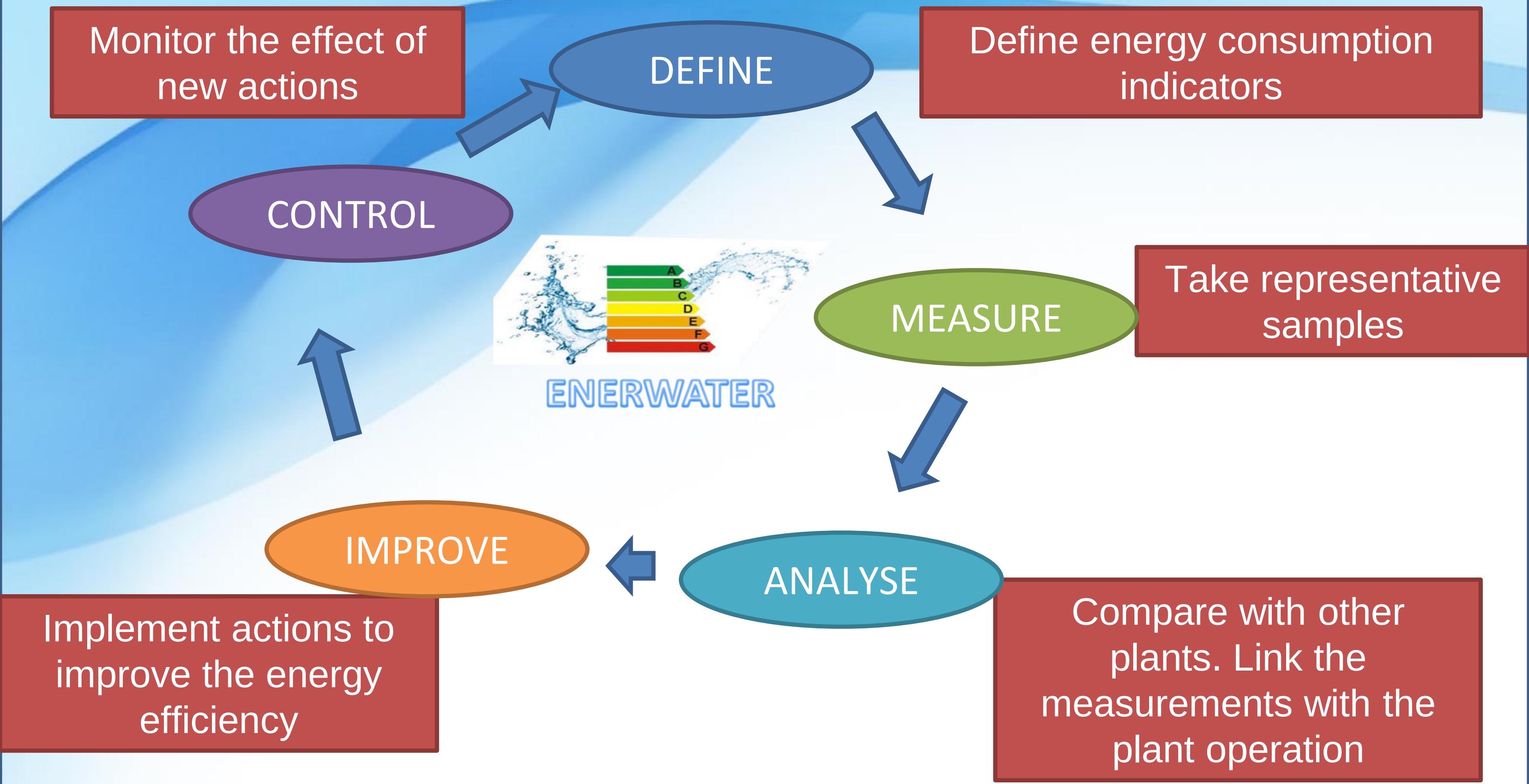
In Spain, domestic and industrial water cycles account for 2-3% of total electric energy consumption. Including water management and agricultural demand, could reach 4-5% [3].

[1] Reinders M, et al (2012) Solution approaches for energy optimization in the water sector. IWA World Congress on Water, Climate and Energy

[2] Foladori P, Vaccari M, Vitali F.(2012) Water Sci Technol;72(6):1007-1015.

[3] Fundación OPTI. Estudio de Prospectiva. Consumo energético en el sector del agua [Prospective studies. Energy consumption in water sector] 2012

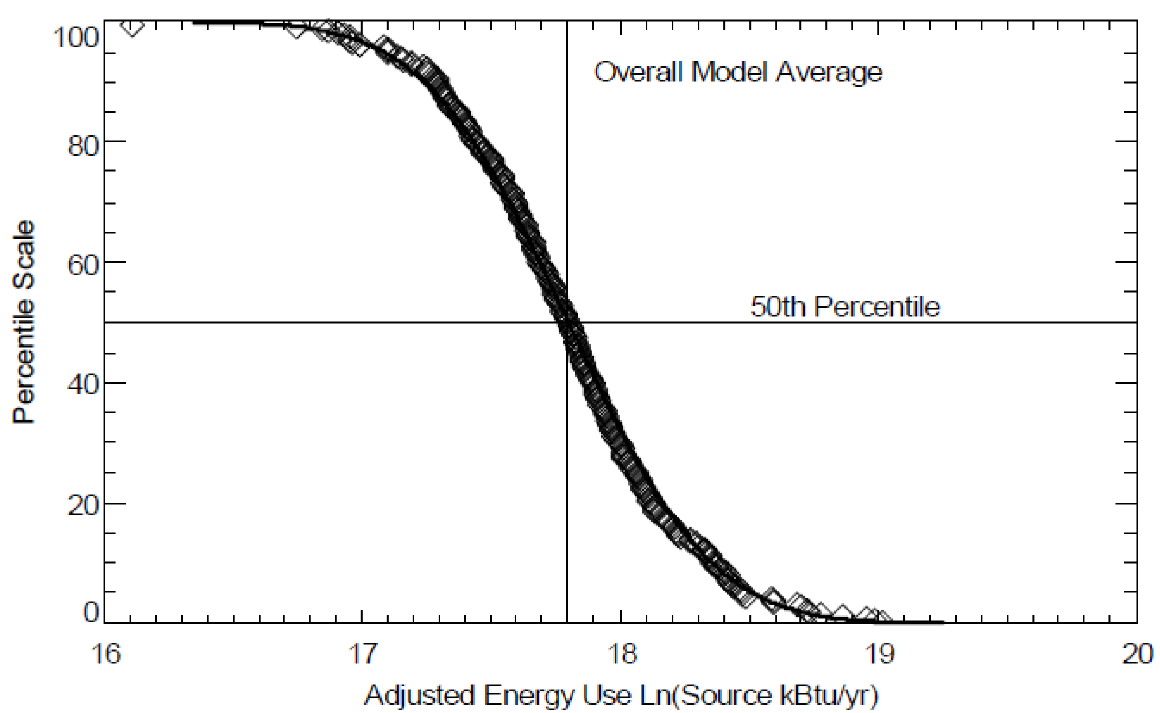
Introduction. Why and how energy efficiency in WWTP ?



Introduction. Why and how energy efficiency in WWTP ?

Previous work: ENERGY STAR initiative

$$\log(Energy)=12.5+0.89\log(F)+0.49\log(BOD_{inf})-0.20\log(BOD_{eff})-0.43\log(PLF)-0.33\,TF+0.16\,Nut$$



BioGroup

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Outline

- Introduction. Why and how energy efficiency in WWTP
- Regression methods
- Example of analysis. WWTP and country location
- Conclusions and perspectives for benchmarking

Regression method. Sample and variate characteristics

Linear regression

$$Y = a + Xb + \varepsilon$$

- *X (n x m)* represents the inputs, covariates or independent variables
- *Y (n x 1)* is the output
- *b (m x 1)* are regression coefficients
- *m* represents the number of inputs
- *n* represents the number of plants
- *a* is the intercept
- *ε* is the error, assumed to be normally distributed: *N(0,σ)*

Provided that *b* is a constant *(m x 1)* vector, the regression is said to be linear

Regression method. Sample and variate characteristics

Data from Data from 601 WWTPs were inventoried

- population equivalent (PE) load basis, both the designed value and the actually served value;
- flow rate (design and average);
- influent and effluent wastewater characteristics, i.e. chemical oxygen demand (COD), biochemical oxygen demand (BOD), total suspended solids (TSS), total nitrogen (TN) and total phosphorus (TP).
- The energy consumption of major pieces of equipment, such as blowers, mixers, pumps, aeration systems and filters was found in a number of cases. Additionally, more general data on energy consumed by the buildings for lighting and heating were also reported.

In this analysis only the plants for which all the data are provided were retained: 185 plants from Germany, Spain and France covering a total of 9.86 million PE

Regression method. Sample and variate characteristics

Variable	Abbreviation	Median	Range	Quantitative	Units
Country	*Count*		'France'; 'Spain'; 'Germany'	No	-
Secondary treatment	*2treat*		'BNR'; 'MBR'; 'Unspecified Secondary Treatment'; 'CAS'; 'Extended Aeration'; 'Other'	No	-
Design size	*DS*	20 000	150 - 833 333	Yes	PE
Actual size	*AS*	12 000	19 – 500 118	Yes	PE
Average flowrate	*F*	3643	6 - 126082	Yes	m^3/d
Plant load factor	*PLF*	59.3	1 - 492	Yes	%
Specific flowrate	*SFl*	230.7	69 - 990	Yes	L/(PE·d)
Influent COD concentration	*CODinf*	560	121 - 1945	Yes	mgCOD/L
Effluent COD concentration	*CODeff*	29	3 - 1022	Yes	mgCOD/L
Influent N concentration	*Ninf*	52.8	6 – 151	Yes	mgN/L
Effluent N concentration	*Neff*	9	0 – 76	Yes	mgN/L
Influent P concentration	*Pinf*	7.1	1 – 19	Yes	mgP/L
Effluent P concentration	*Peff*	1.1	0 – 8	Yes	mgP/L
Removal rate of COD[#]	*CODrem*	1463	3 – 58318	Yes	kg COD/d
Removal rate of N[#]	*Nrem*	116	0 – 4227	Yes	kg N/d
Removal rate of P[#]	*Prem*	15.7	0 -705	Yes	kg N/d
Energy consumption	*W*	1521	3 – 36653	Yes	kWh/d

Regression method. Sample and variate characteristics

Step 1. Data Inspection

Inspection of data to remove missing data (final sample 185) and check scatterplot matrix to check that the curvature is taken into account (log transformation of W, DS, AS, F, COD_{rem}, N_{rem}, P_{rem}). Check relationship with dependent variable

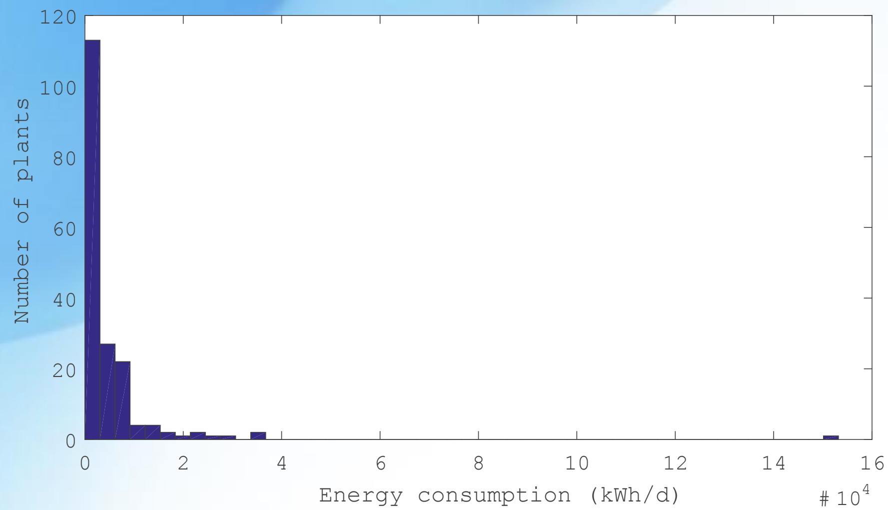
BioGroup

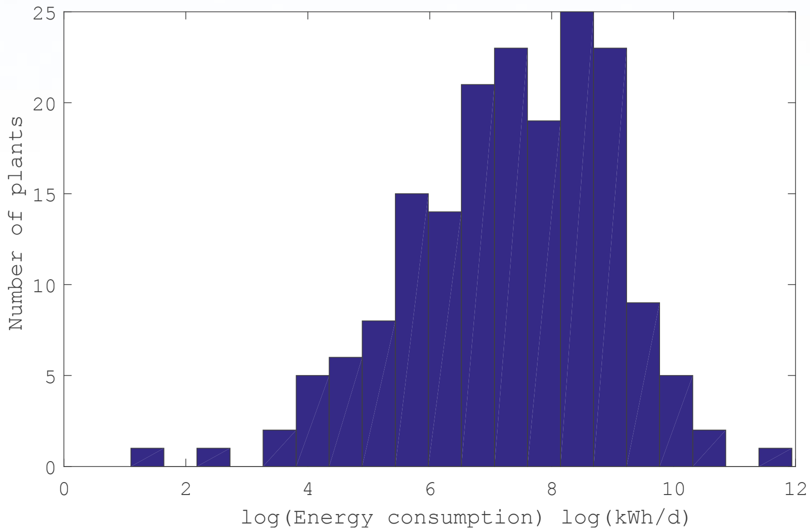
USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Regression method. Sample and variate characteristics

Step 1. Data Inspection

Inspection of data to remove missing data (final sample 185) and check scatterplot matrix to check that the curvature is taken into account (log transformation of W, DS, AS, F, COD_{rem}, N_{rem}, P_{rem}). Check relationship with dependent variable

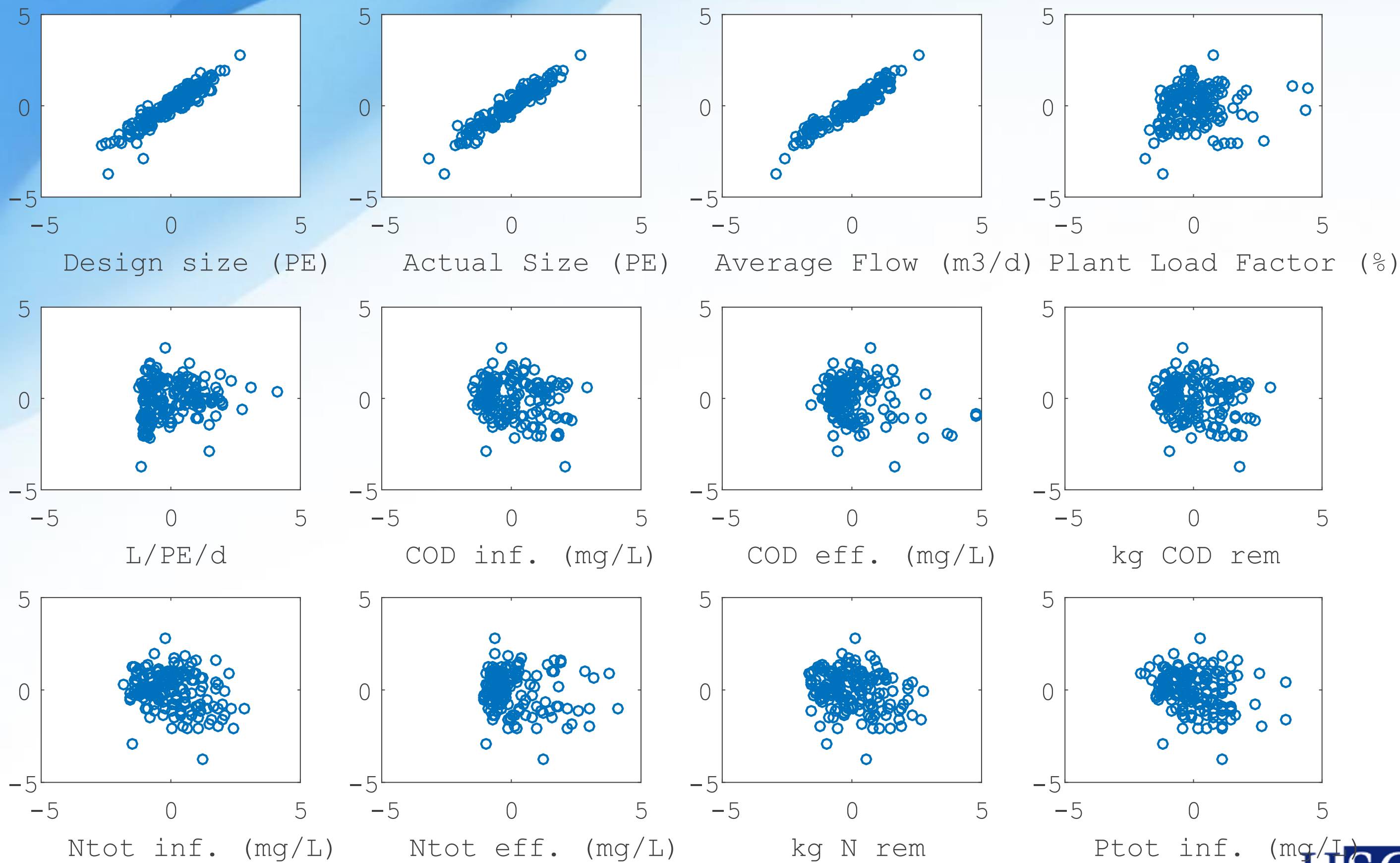




Regression method. Sample and variate characteristics

Step 1. Data Inspection

Inspection of data to remove missing data (final sample 185) and check scatterplot matrix to check that the curvature is taken into account (log transformation of W, DS, AS, F, COD_{rem}, N_{rem}, P_{rem}). Check relationship with dependent variable



Regression method. Sample and variate characteristics

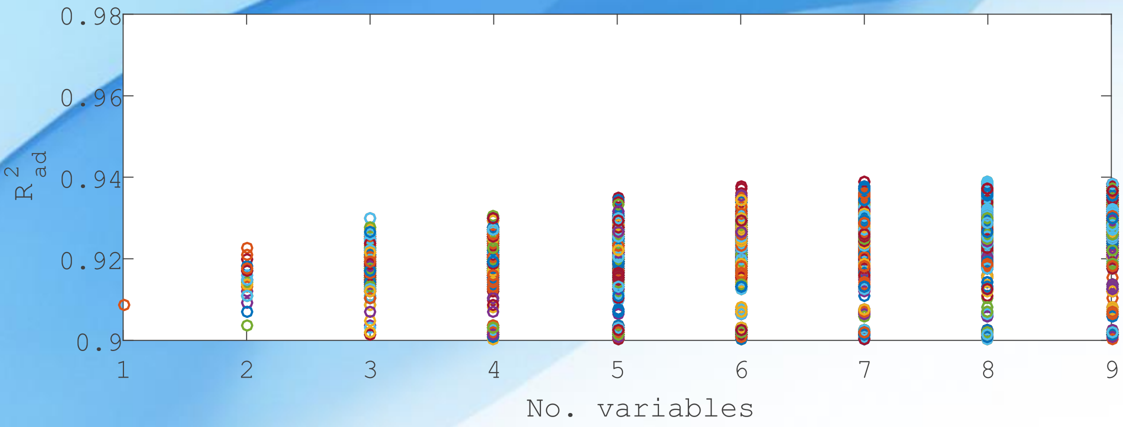
Step 2. Variable selection

Build all the combinations of the input variables check model selection criteria

Regression method. Sample and variate characteristics

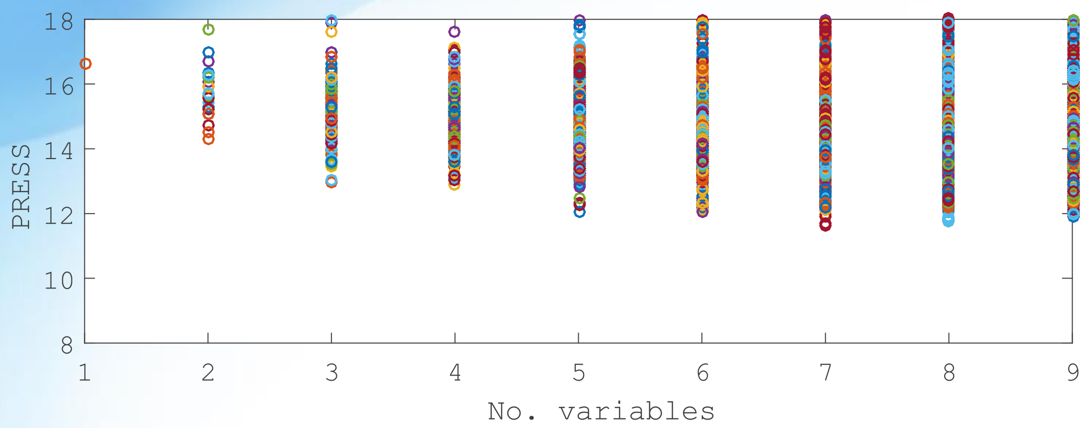
Step 2. Variable selection

Build all the combinations of the input variables check model selection criteria



Adjusted R^2

$$R^2_{adj} = 1 - \frac{N-1}{N-p}\frac{SSE}{SSTO}$$



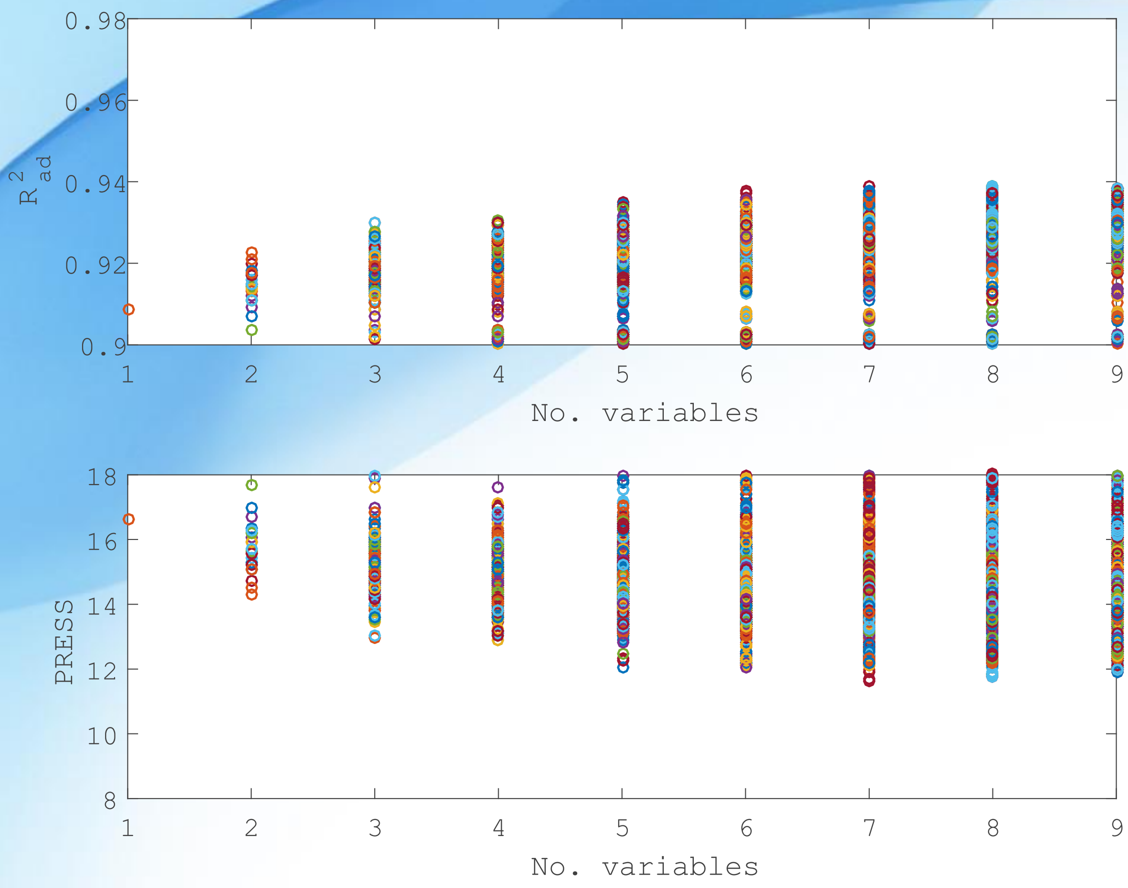
Prediction sum of squares (PRESS)

$$PRESS = \sum_{i=0}^{N} \left(Y_i - \hat{Y}_{i(i)}\right)^2$$

Regression method. Sample and variate characteristics

Step 2. Variable selection

Build all the combinations of the input variables check model selection criteria



Best sets:

Actual Size
Average Flow rate
Plant Load Factor
Influent COD
Effluent COD
Effluent N

Regression method. Sample and variate characteristics

<u>Step 2b. Variable selection with composed covariates</u>

For the best set of variables, build covariates that stand for mixed effects, squared effects, etc.

Candidates:

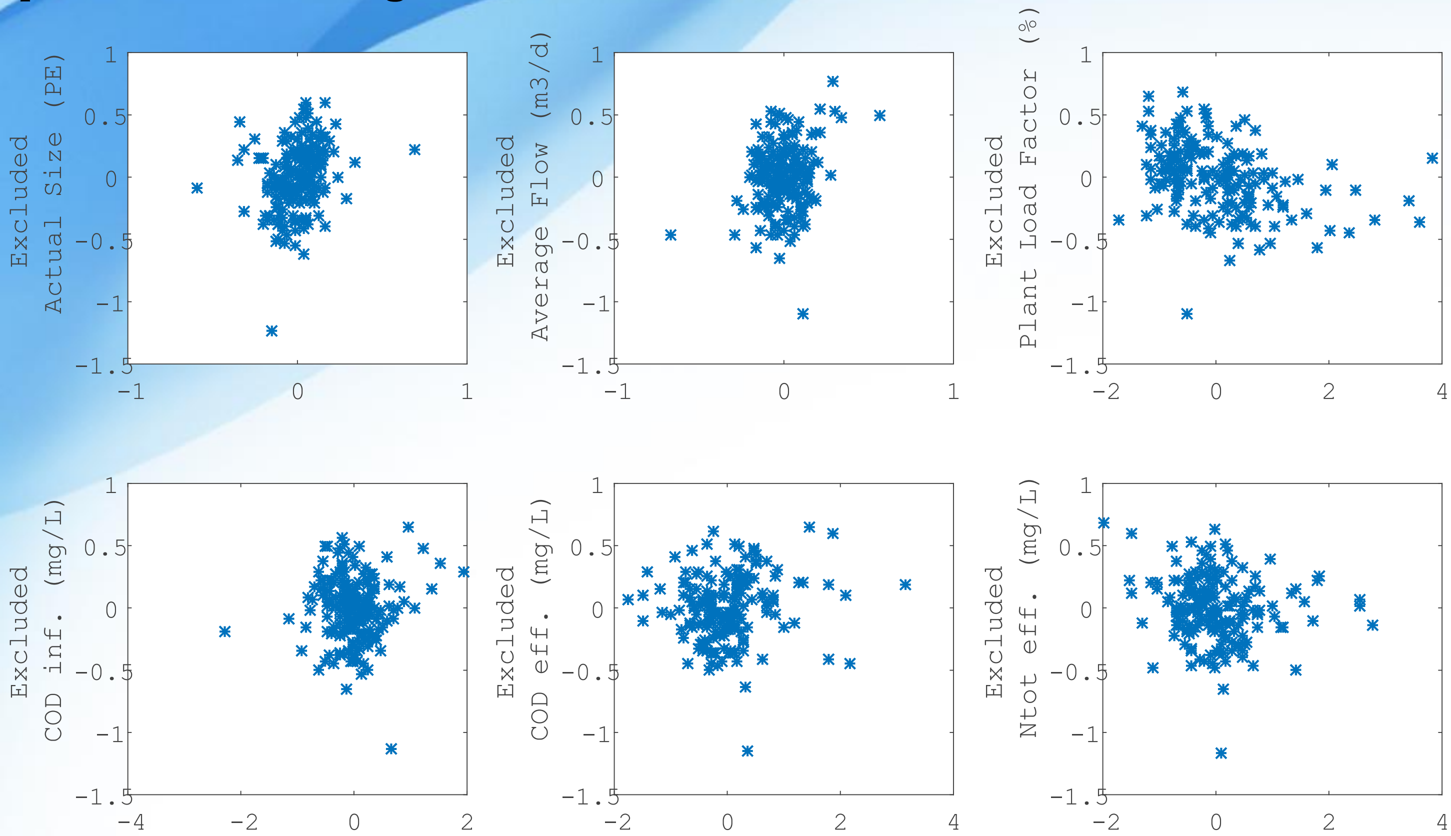
All the variables squared

COD load = Average flowrate · Infuent COD
N load = Average flowrate · Infuent N
Percentage of COD removed = 100 x effluent COD/influent COD
Percentage of N removed = 100 x effluent N/influent N

Regression method. Sample and variate characteristics

Step 3. Model validation and refinement

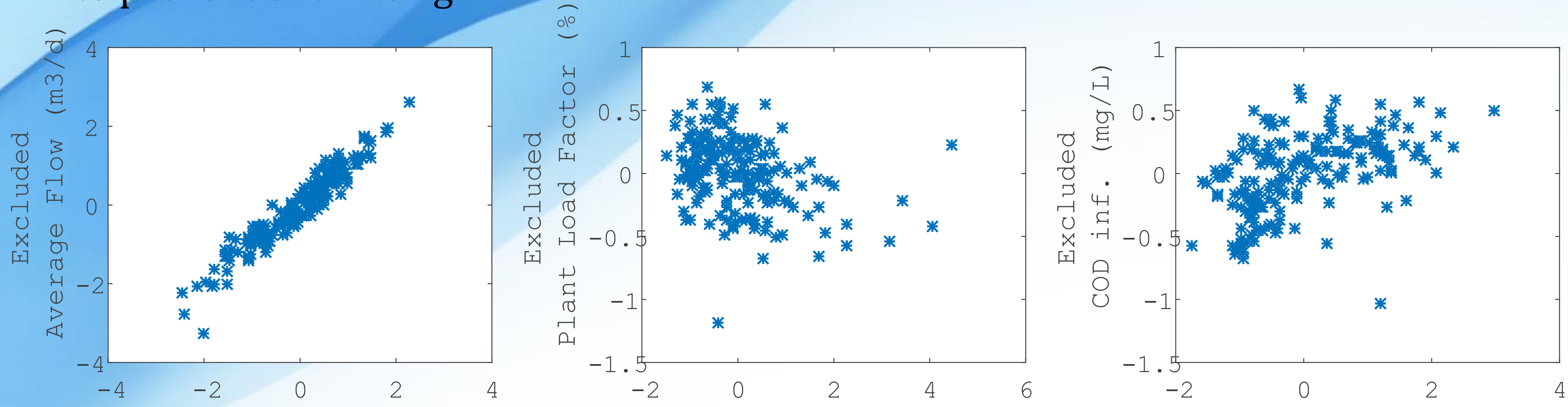
Check simplifications of the model that lead to similar fitting. Divide the data into a training dataset (2/3 of the original data) and a validation dataset (remaining 1/3) to prevent overfitting.



Regression method. Sample and variate characteristics

Step 3. Model validation and refinement

Check simplifications of the model that lead to similar fitting. Divide the data into a training dataset (2/3 of the original data) and a validation dataset (remaining 1/3) to prevent overfitting.



$$\log(W) = 0.67 + 0.82\log(F) - 4.20 \cdot 10^{-3} PLF + 9.12 \cdot 10^{-4} CODinf$$

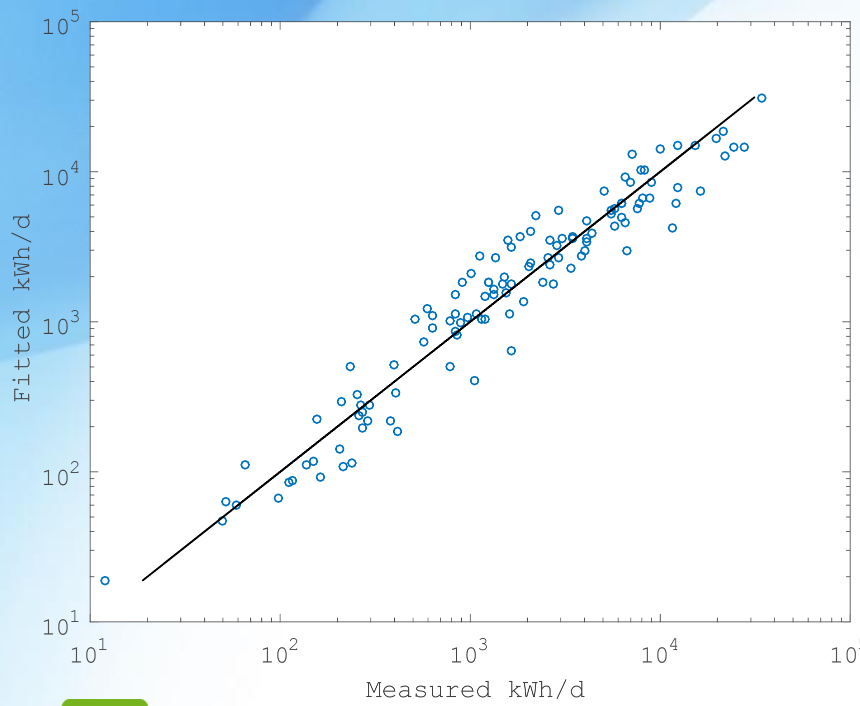
Regression method. Sample and variate characteristics

Step 3. Model validation and refinement

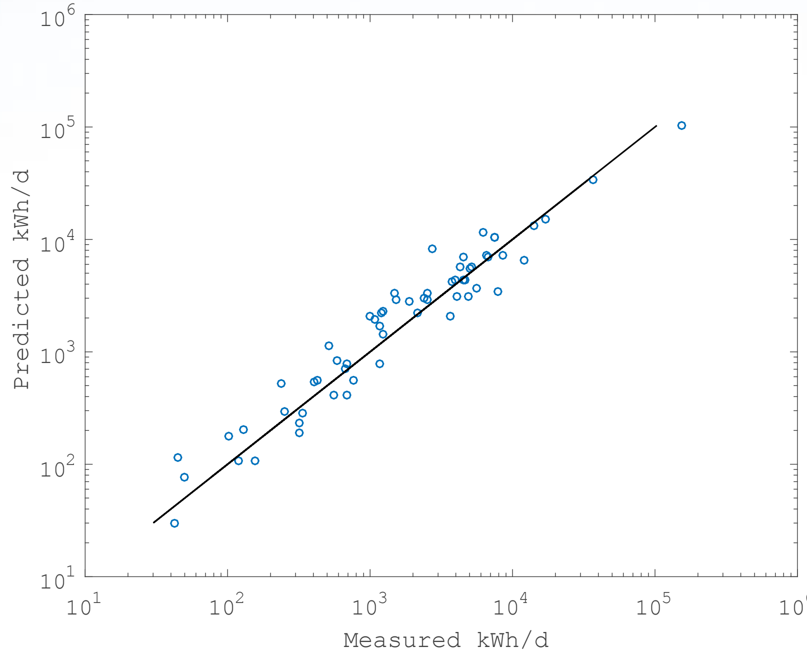
Check simplifications of the model that lead to similar fitting. Divide the data into a training dataset (2/3 of the original data) and a validation dataset (remaining 1/3) to prevent overfitting.

$$\log(W) = 0.67 + 0.82\log(F) - 4.20 \cdot 10^{-3} PLF + 9.12 \cdot 10^{-4} CODinf \qquad R^2 = 0.942$$

Calibration

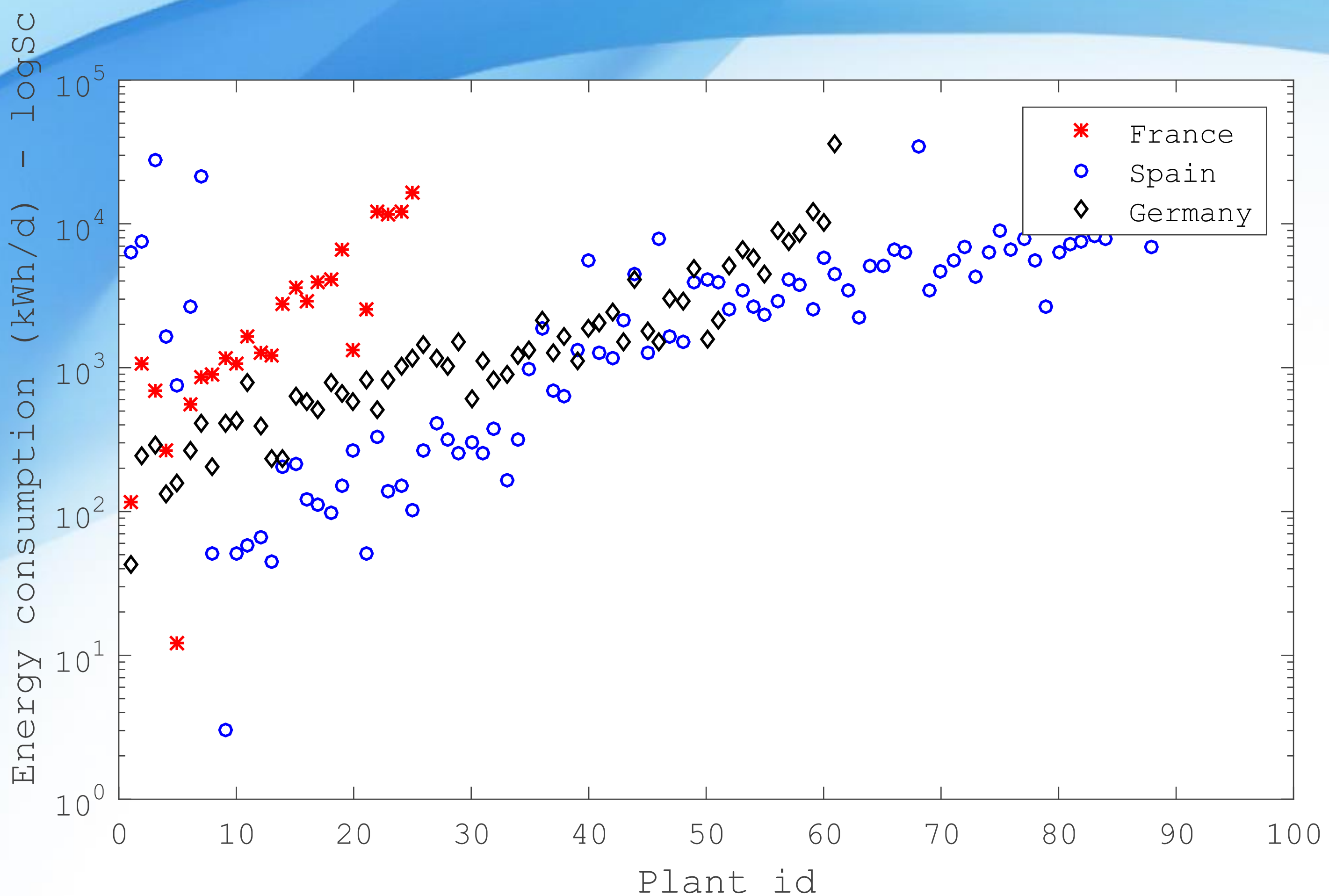


Validation



WWTP characteristics & energy consumption. Country location

Does the country location impact the WWTP energy consumption?



BioGroup

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

WWTP characteristics & energy consumption. Country location

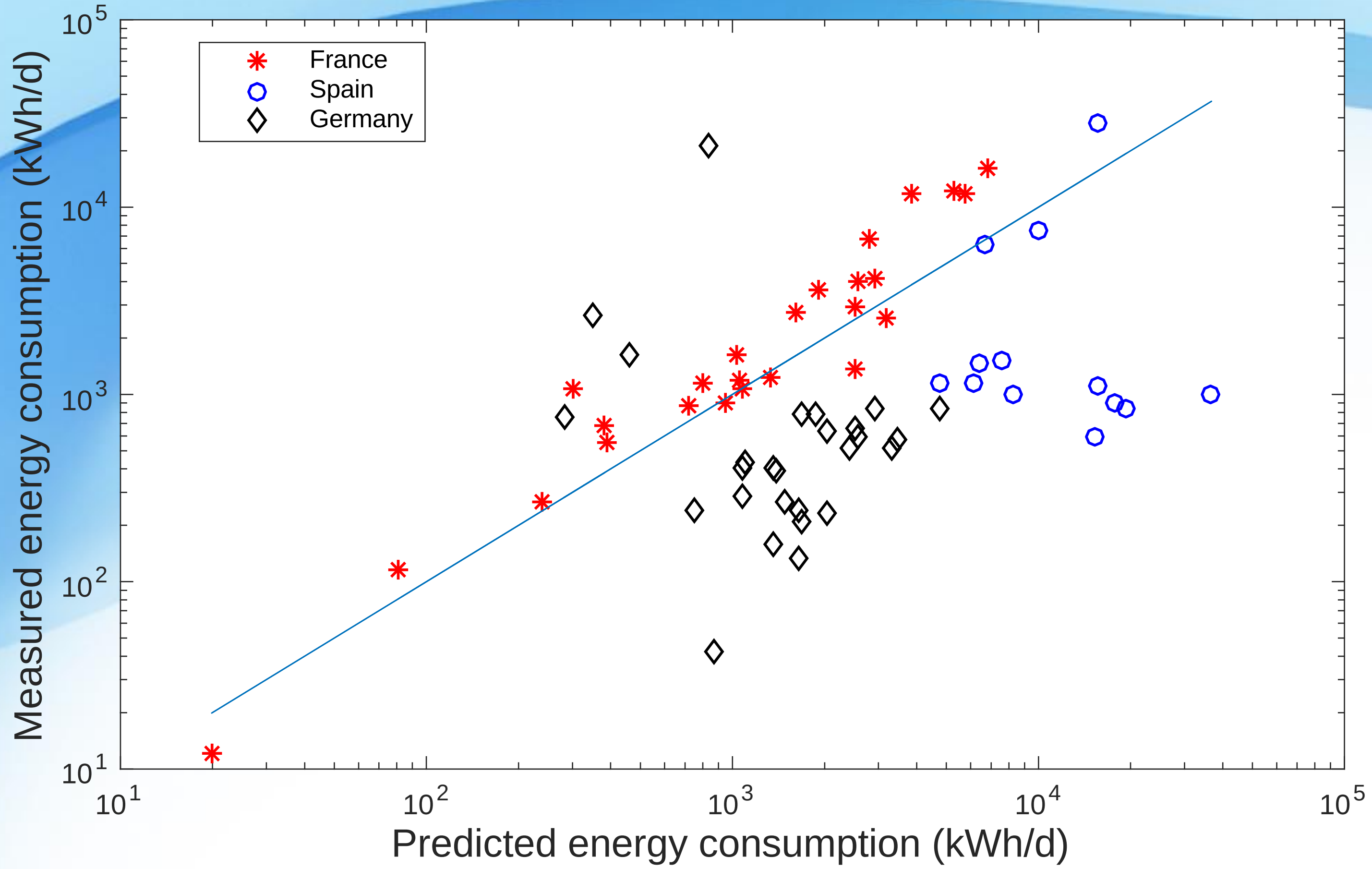
Does the country location impact the WWTP energy consumption?

	Germany		France		Spain	
No control	-0.497***	(0.084)	0.614***	(0.157)	0.797***	(0.110)
Log(F)	-0.425***	(0.059)	0.541***	(0.110)	0.680***	(0.077)
Log(F), CODinf	-0.300***	(0.065)	0.609***	(0.107)	0.416***	(0.099)
Log(F), CODinf, PLF	-0.270***	(0.063)	0.528***	(0.104)	0.378***	(0.095)
Log(F), CODinf, PLF, 2treat	-0.133	(0.155)	0.840***	(0.114)	0.003	(0.235)

The differences persist as we control for more covariates although they become smaller. They are no longer significant between Germany and Spain when only secondary treatment based on biological nutrient removal is kept

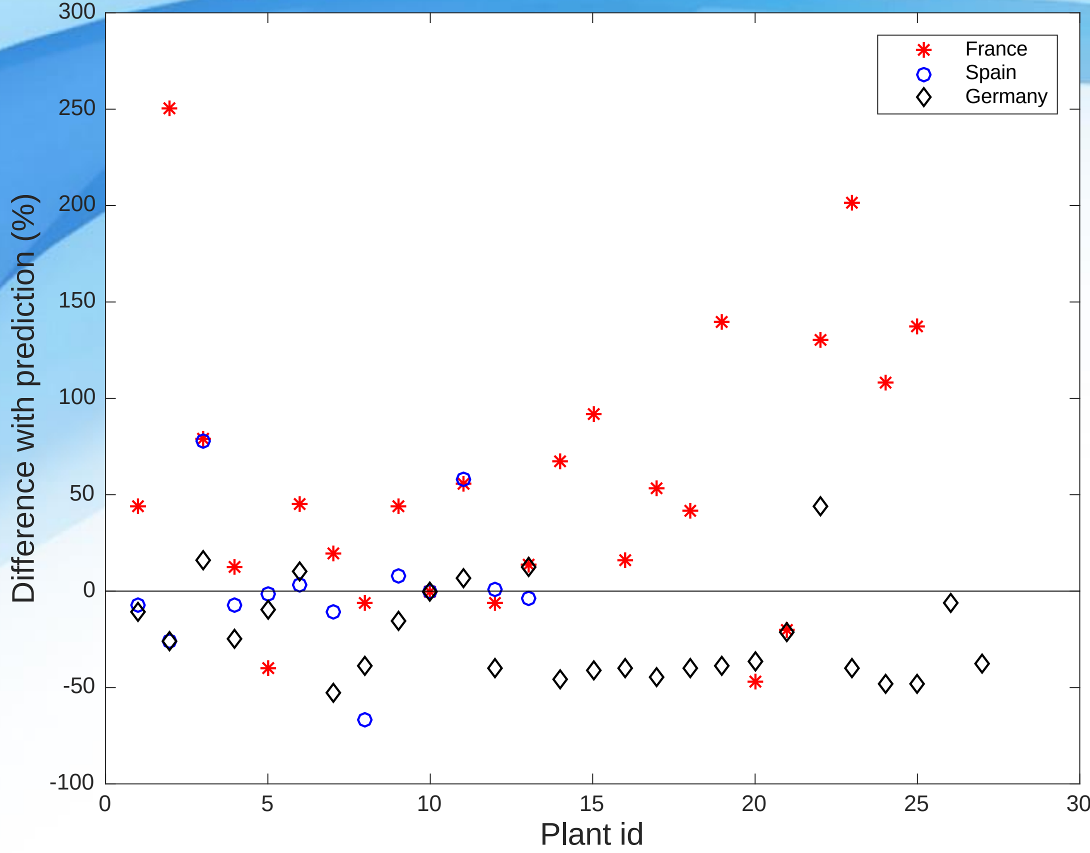
WWTP characteristics & energy consumption. Country location

Does the country location impact the WWTP energy consumption?



WWTP characteristics & energy consumption. Country location

Does the country location impact the WWTP energy consumption?



French WWTPs consume around 50% more energy than comparable German and Spanish WWTPs

Conclusions

- Regression methods are appropriate tools for benchmarking WWTP energy consumption
- As many variables are related to each other, it is possible to use just a few covariates to cover a large part of the variability
- Extreme care must be taken when making comparisons: are the WWTPs really comparable? Is the sample for comparison random?

Future perspectives

- WWTPs are composed of many processes. Overall benchmarking is not enough for diagnosis

Acknowledgements

This work is funded by:

- Marie Curie IEF FP7 project GREENCOST
- CSA H2020 ENERWATER

Thanks to

María Loureiro - Universidade de Santiago de Compostela

To know more

Come to the ENERWATER workshop today at 17:40 at Recital Room!

Check poster P26 on Wednesday session

(D1-S3-58) 17.20 at Lecture Room 2
Improving Energy Efficiency In Wastewater Treatment. Identification Of Key Parameters And Key Performance Indicators (KPIs) by Pablo Campo

@EnerwaterPro

Enerwater Project

www.enerwater.eu

Miguel Mauricio-Iglesias, Stefano Longo, Juan M. Lema, Almudena Hospido

miguel.mauricio@usc.es

Department of Chemical Engineering
Universidade de Santiago de Compostela (Spain)