Norwegian University of Science and Technology

Department of Electric Power Engineering

TK8117: Multivariate Data Analysis

PhD Project: Energy Management and Control of Offshore Platforms Integrating Renewable Energy

VISTA scholar – PhD Candidate

Spyridon Chapaloglou

Phone number: +30 6979471675

Mail: spyridon.chapaloglou@ntnu.no

Contents

1. PhD project introduction

PhD Project

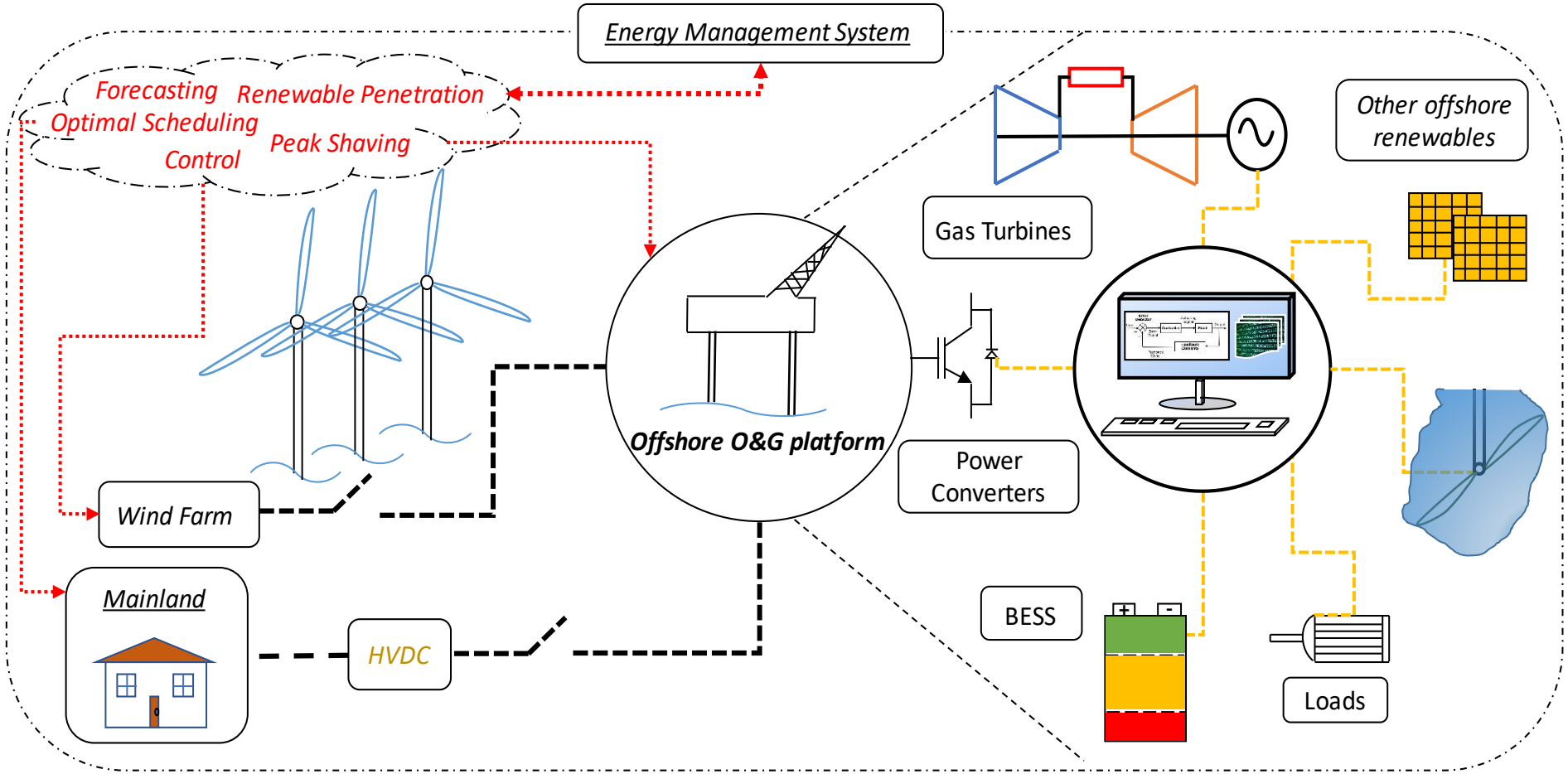
*RES: **R**enewable **E**nergy **S**ources*
*BESS: **B**attery **E**nergy **S**torage **S**ystem*
*GT: **G**as **T**urbines*

Aim

1. Integration of *RES* & *BESS* to isolated O&G platforms
2. Comparison to grid interconnected O&G operation

Targets

- ✓ Operational costs & Emission ↓
- ✓ Operation flexibility (GT+RES+BESS+Grid) ↑
- ✓ RES dependency/reliability (storage system) ↑



2. Data Description

Weather Data Description

Hourly resolution – 3 years (2016-2018)

Considered Dataset (reanalysis)

Institution | Model

- NASA | MERRA-2

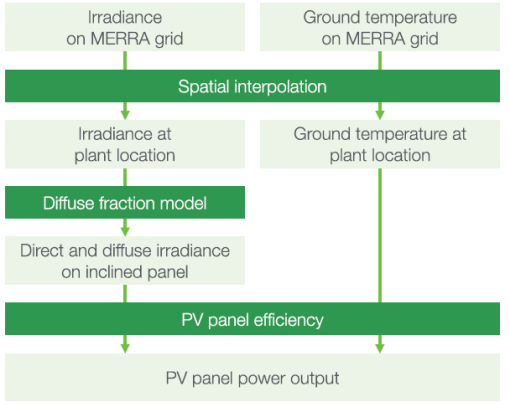
Location Coordinates (Oslo)

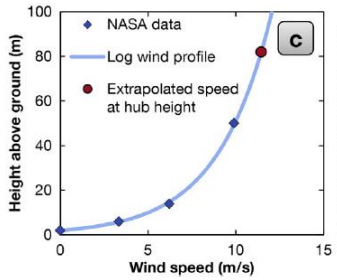
1. Latitude : 59.9133 °
2. Longitude : 10.7390 °

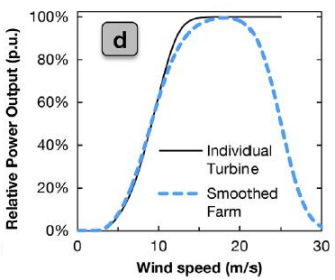
#	***Weather Variable (X)***	***Units***	***Explanation***
1	Wind Speed	$\left(\frac{m}{s}\right)$	*@ the turbine's hub height (100 m above ground)*
2	Air Temperature	(°C)	*2 m above ground level*
3	Precipitation	$\left(\frac{mm}{h}\right)$	*Total precipitation, over land only*
4	Snowfall	$\left(\frac{mm}{h}\right)$	*Total precipitation in the form of snow, over land only*
5	Snow Mass	$\left(\frac{kg}{m^2}\right)$	*Amount of snow per land area*
6	Air density	$\left(\frac{kg}{m^3}\right)$	*@ ground level*
7	G-L Solar Irradiance	$\left(\frac{W}{m^2}\right)$	*Ground-level incident shortwave radiation flux (cloud cover + aerosols)*
8	TOA Solar Irradiance	$\left(\frac{W}{m^2}\right)$	*Top of atmosphere incident shortwave radiation flux (before cloud cover + aerosols)*
9	Direct Irradiance	$\left(\frac{kW}{m^2}\right)$	*Direct radiation flux to the solar panel plane*
10	Diffuse Irradiance	$\left(\frac{kW}{m^2}\right)$	*Diffuse radiation flux to the solar panel plane*
11	Cloud cover fraction	-	*Averaged over grid cell and summed over all height above ground*

RES power generation (*Y*)

1. Wind Power (2 MW)
2. PV Power (1 MW)

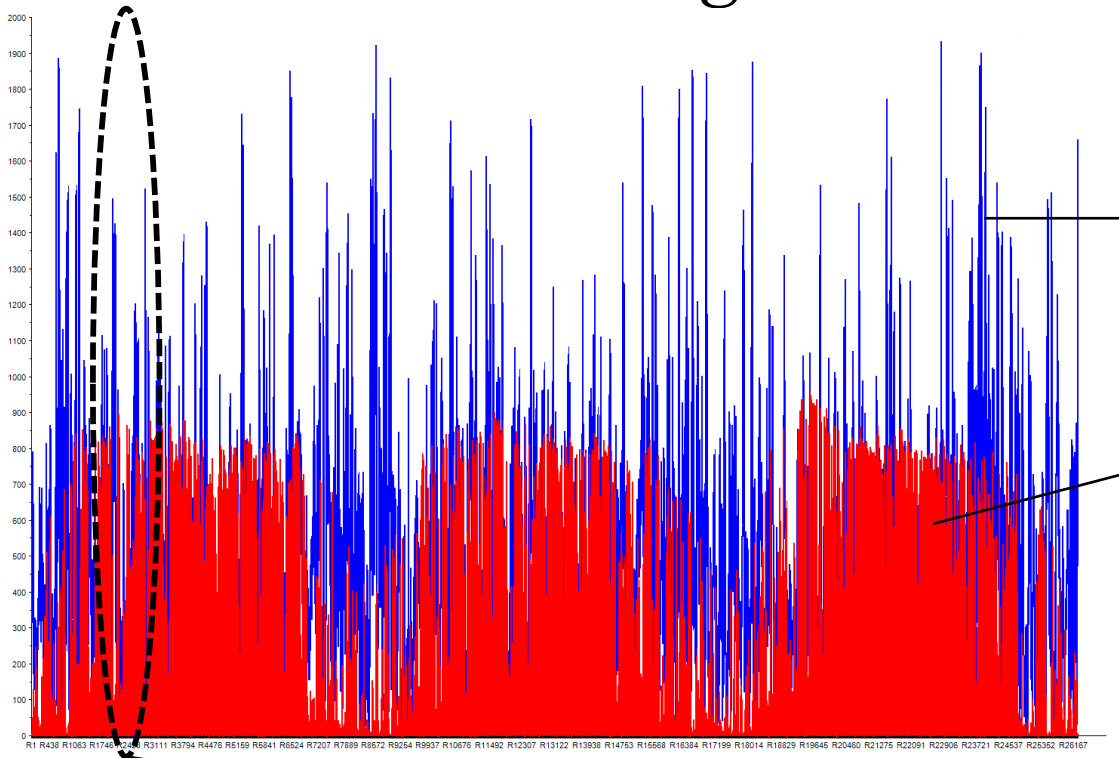






- Wind Turbine: Vestas V90 2000, 100 m hub, cut-in: 4 m/s
- Solar Panels: 35° tilt, 180° azimuth, 2-axis tracking

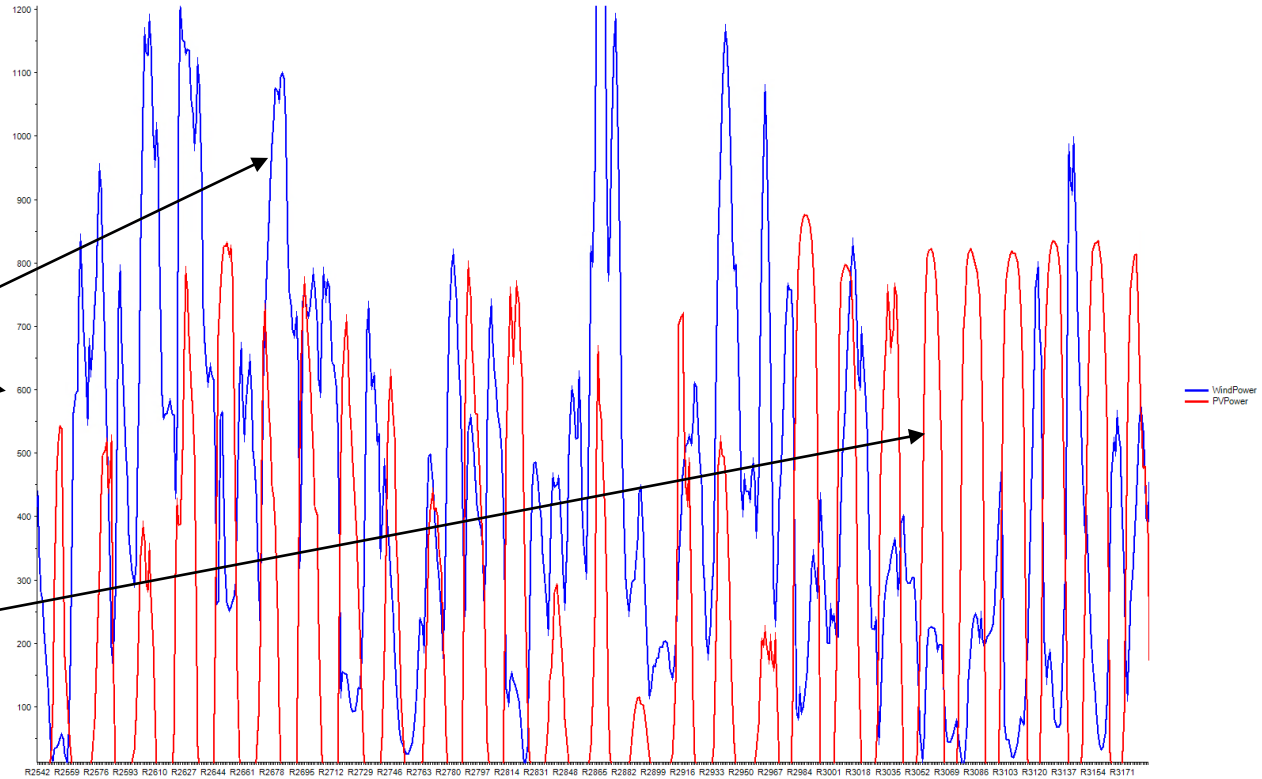
RES Data Plotting



Wind power → No clear pattern

PV power → Seasonal pattern

Zoom area



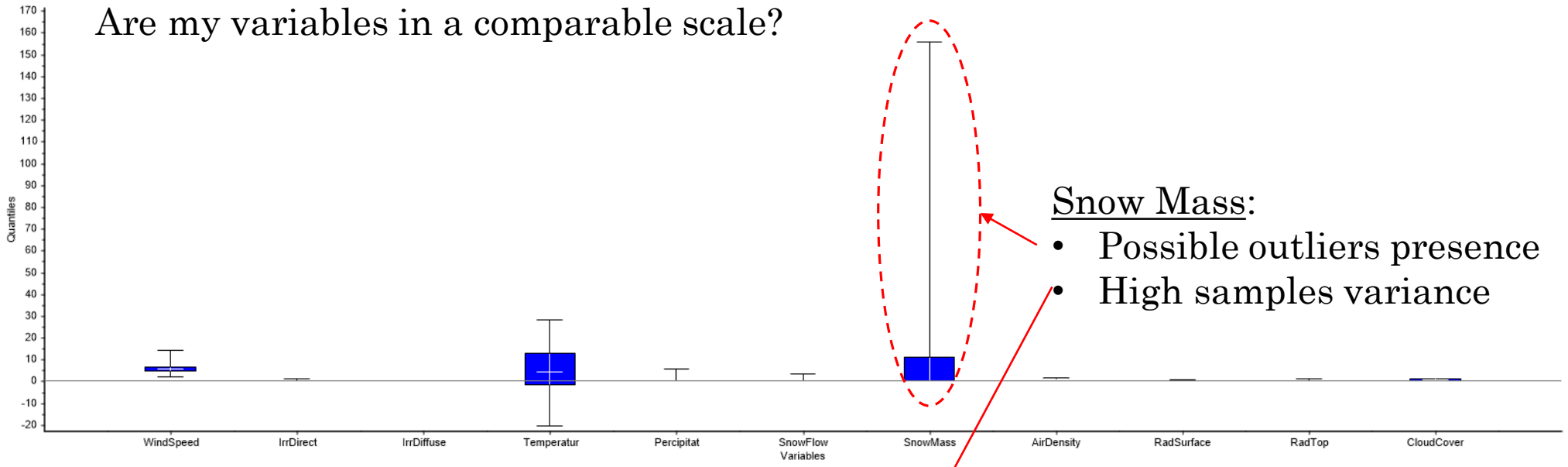
- **Large variations**
- **No clear pattern**

- **Smoother profile**
- **Clear intraday pattern**

3. Principal Component Analysis

Principal Component Analysis (1st approach)

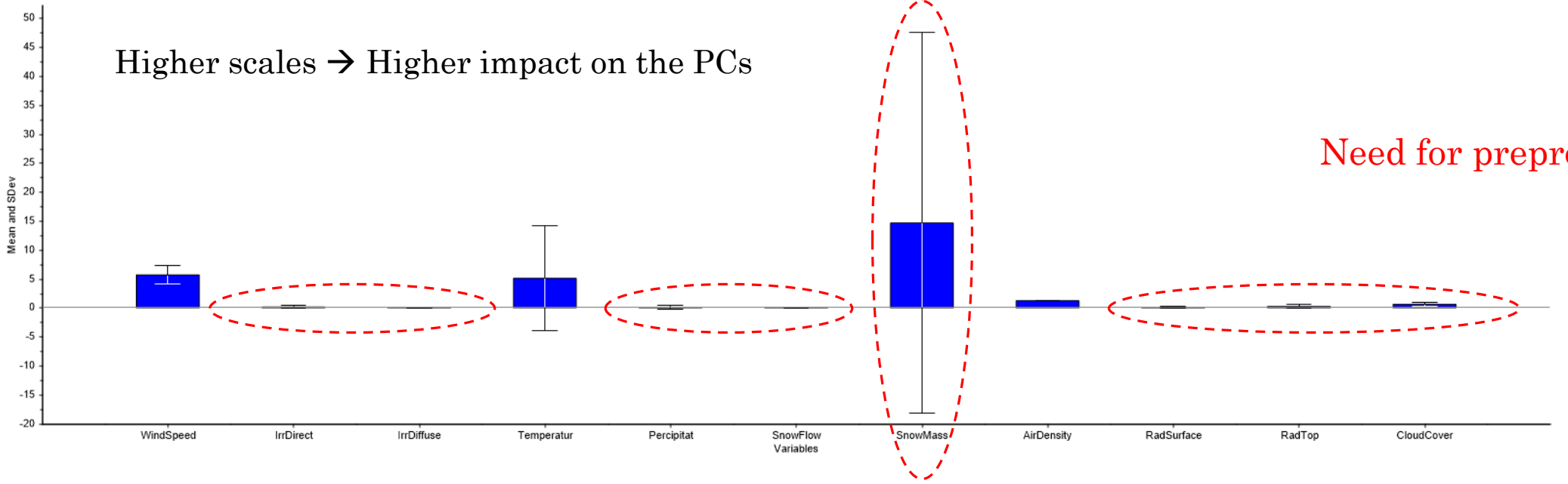
Are my variables in a comparable scale?



Snow Mass:

- Possible outliers presence
- High samples variance

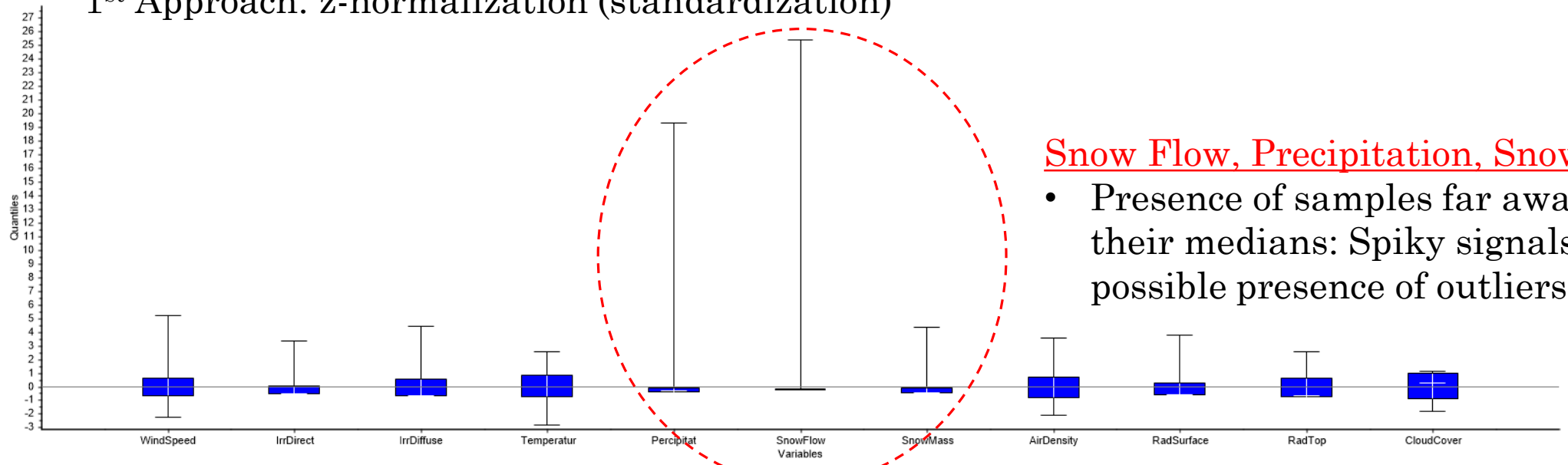
Higher scales → Higher impact on the PCs

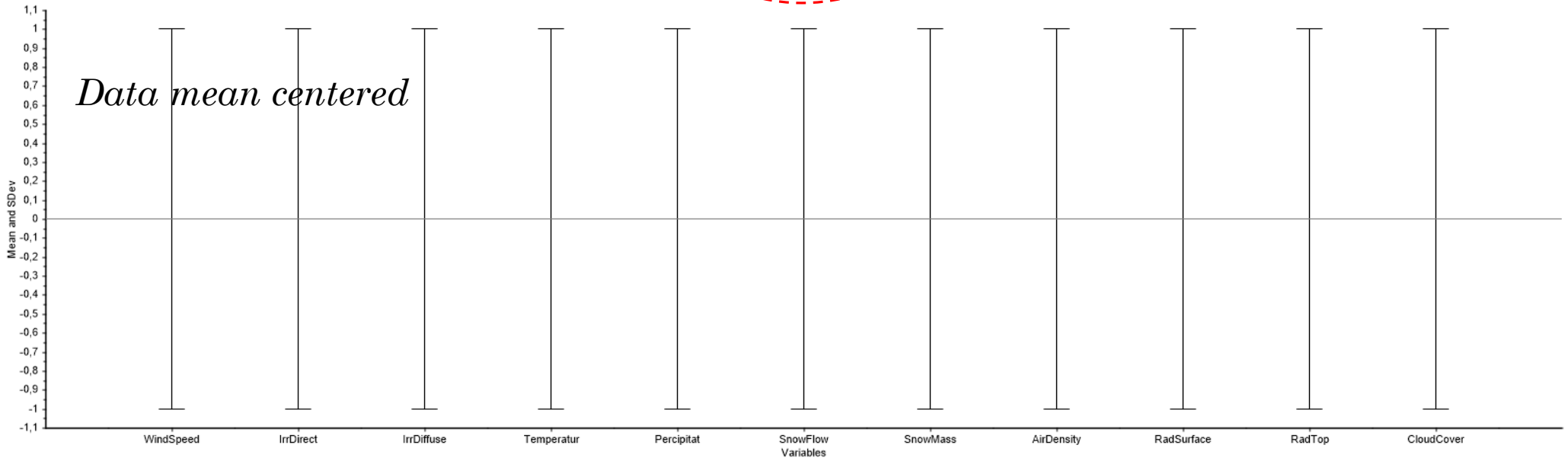


Need for preprocessing!

Principal Component Analysis (1st approach)

1st Approach: z-normalization (standardization)





Data mean centered

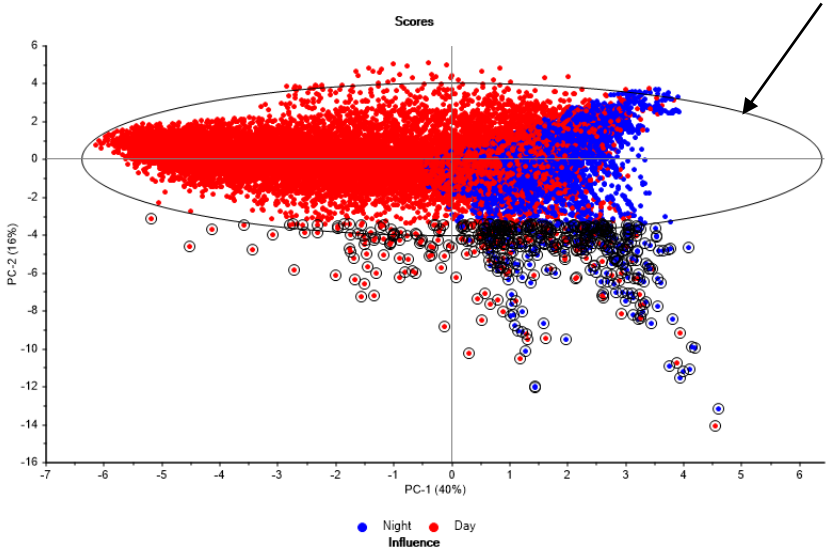
Snow Flow, Precipitation, Snow Mass:

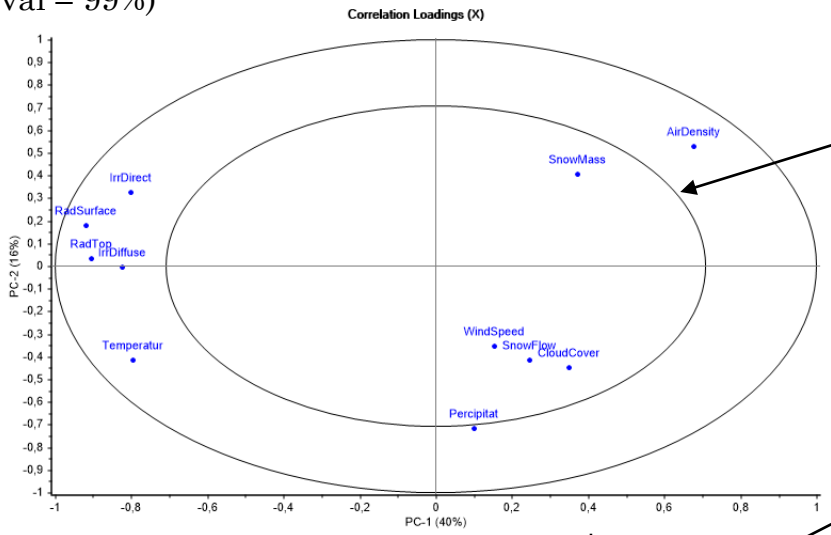
- Presence of samples far away from their medians: Spiky signals with possible presence of outliers

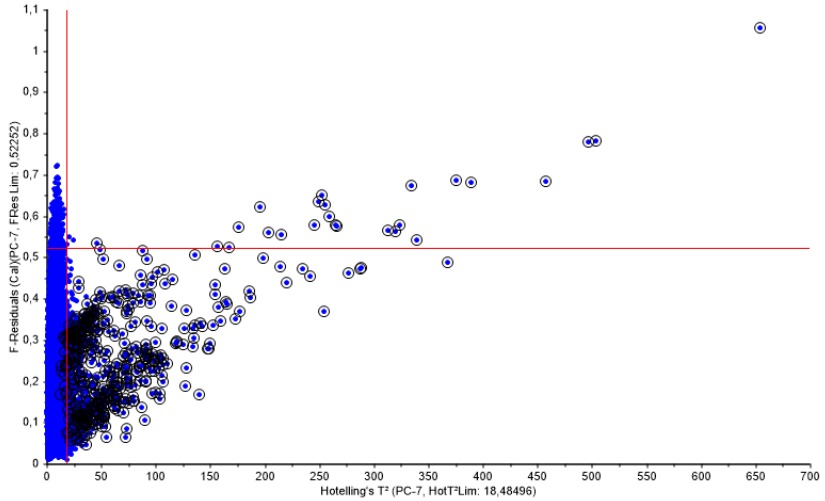
Principal Component Analysis (1st approach)

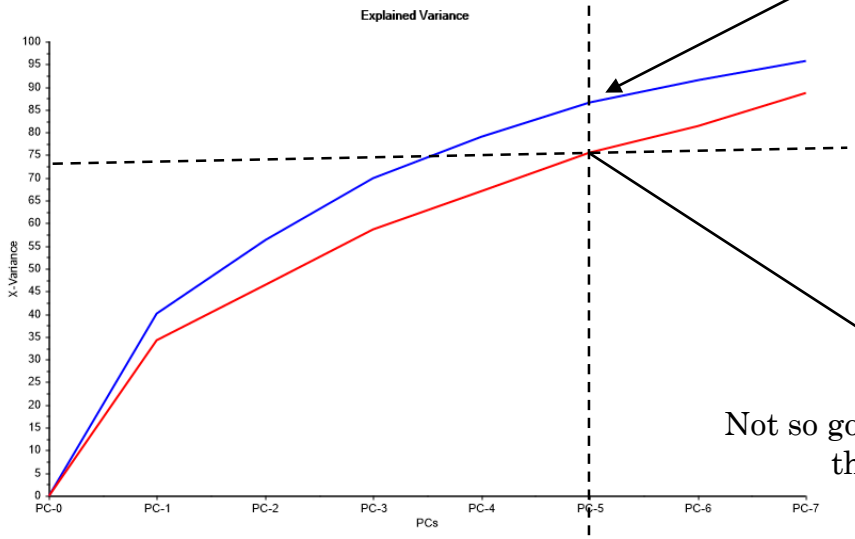
PCA algorithm: SVD

Selected confidence level for the Hotelling's ellipsoid (confidence interval = 99%)









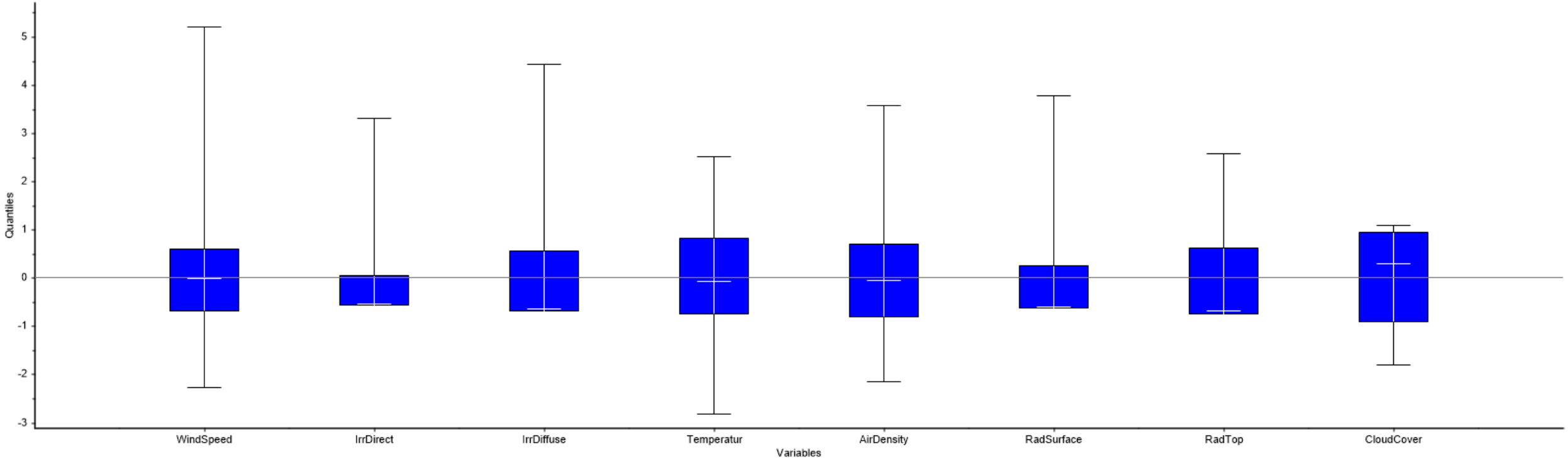
- Variables with < 50% contribution to the PCs: *Snow Mass, Snow Flow, Cloud Cover, Wind Speed*
- Lots of samples beyond the critical limits (leverage and F-statistic residuals)
- Relatively low explained variance – small improvement with more PCs
- High residual values
- Day vs Night patterns seem that can be better separated

Not so good predictive model for the weather state

Principal Component Analysis (2nd approach)

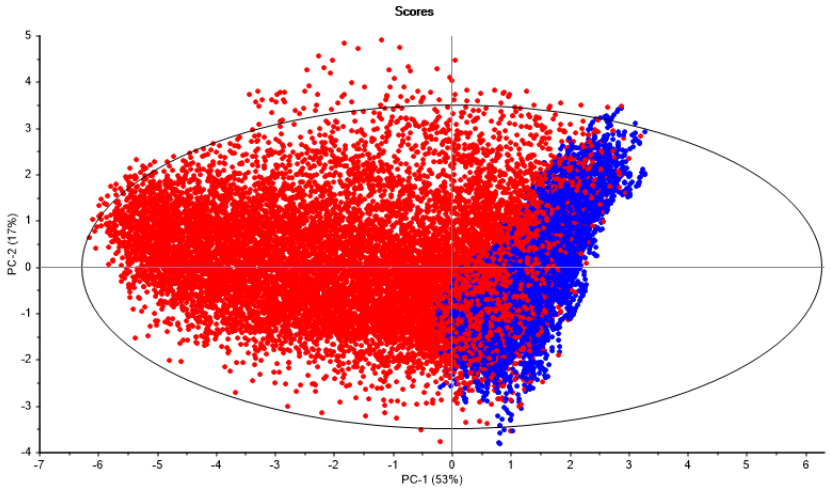
2nd Approach: standardization without Snow Flow, Precipitation, Snow Mass

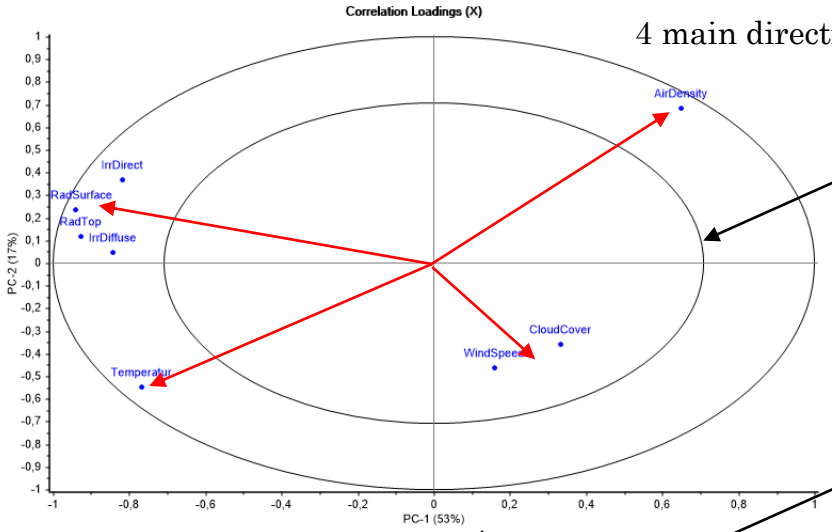
Better data pre processing for PCA

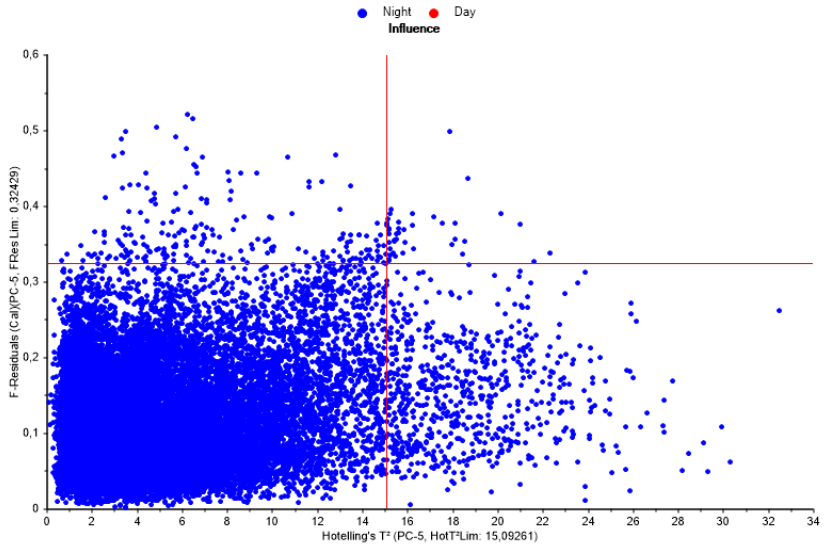


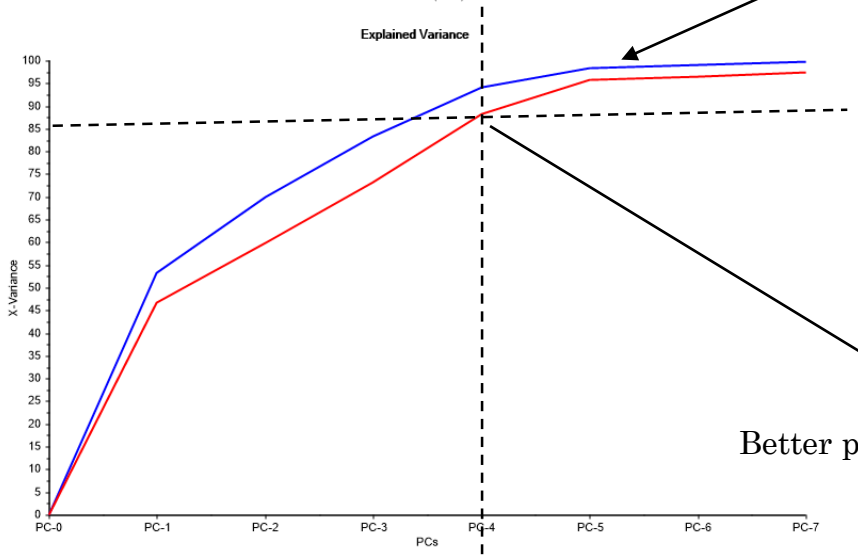
Principal Component Analysis (2nd approach)

PCA algorithm: SVD









4 main directions describe most of our data

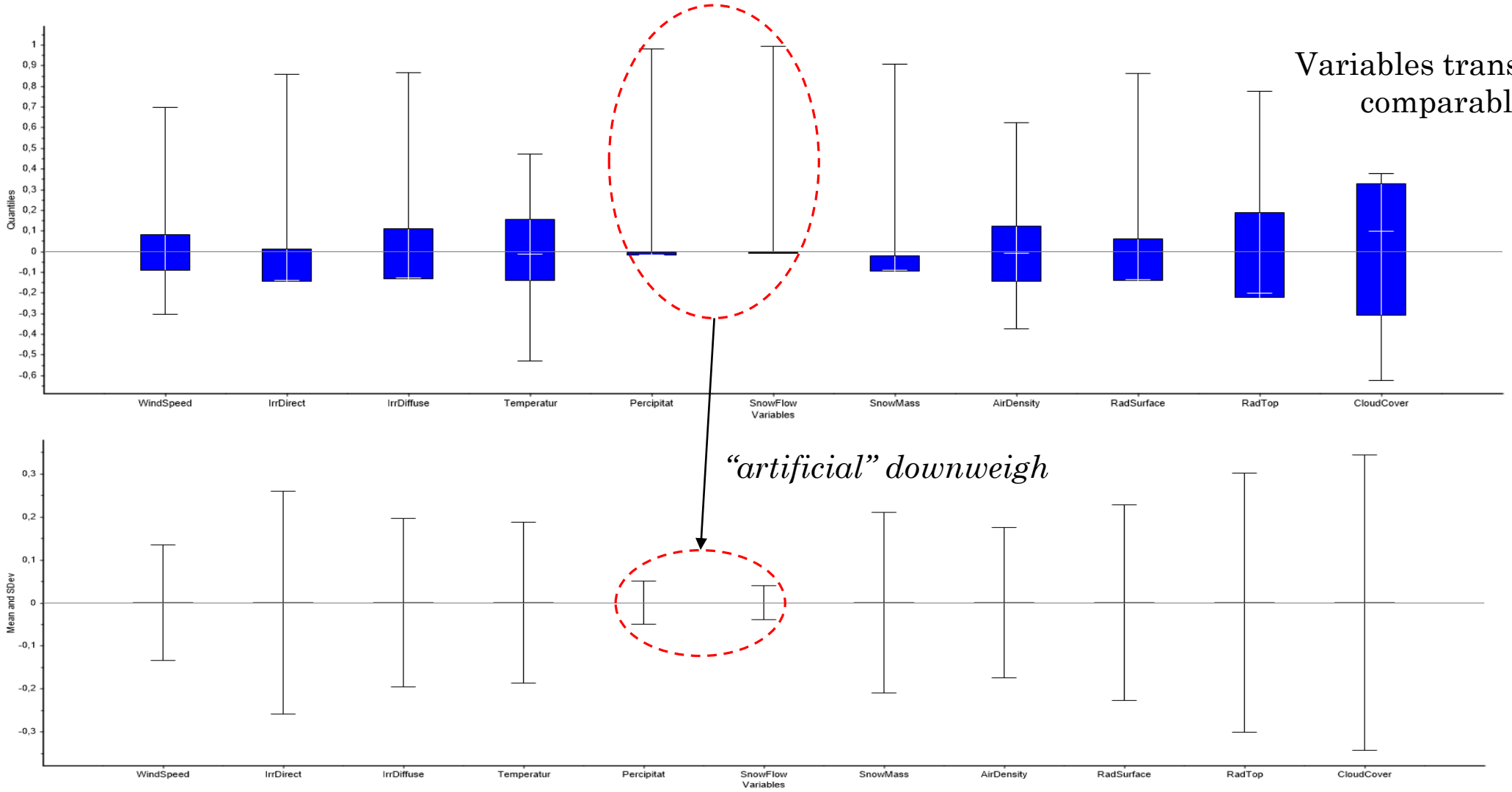
- Variables with < 50% contribution to the PCs: *Cloud Cover, Wind Speed*
- Fewer samples beyond the critical limits (leverage and F-statistic residuals)
- Higher explained variance – small improvement after 5 PCs
- Less samples to "suspect" as outliers
- Day vs Night patterns better separated

Better predictive performance with less PCs

Principal Component Analysis (3rd approach)

3rd Approach: mean centering and range normalization

each row is divided by its range, i.e. max value - min value

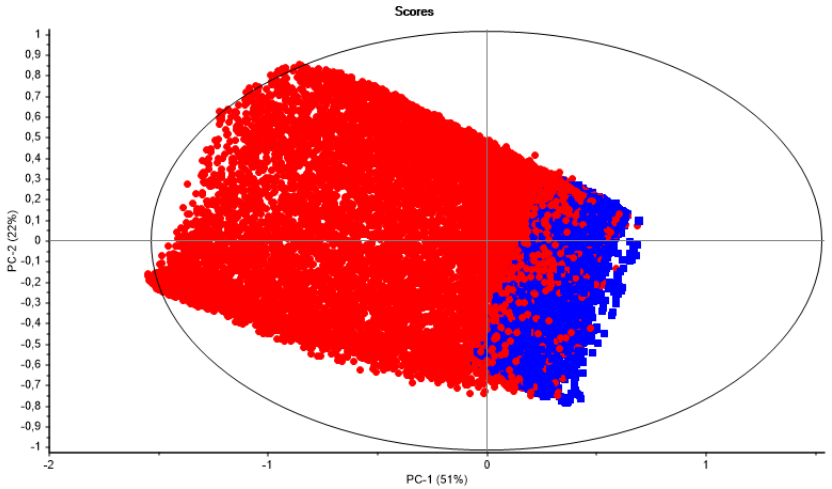


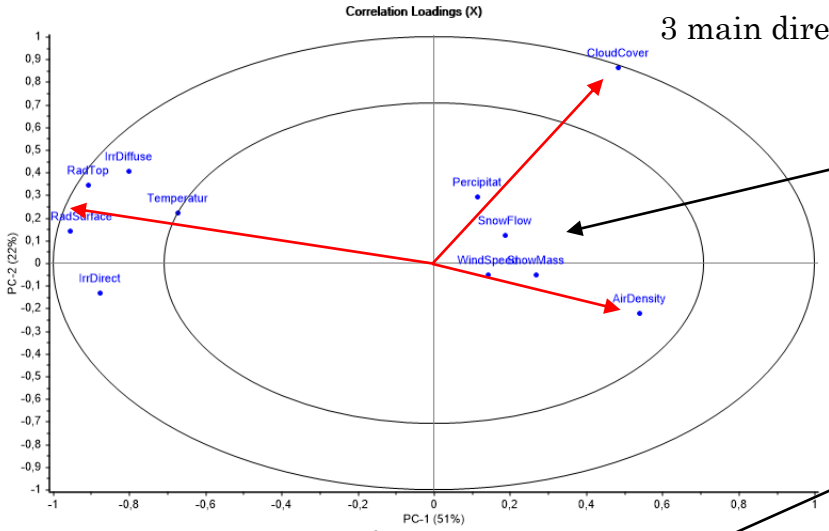
Variables transformed in a comparable range

"artificial" downweigh

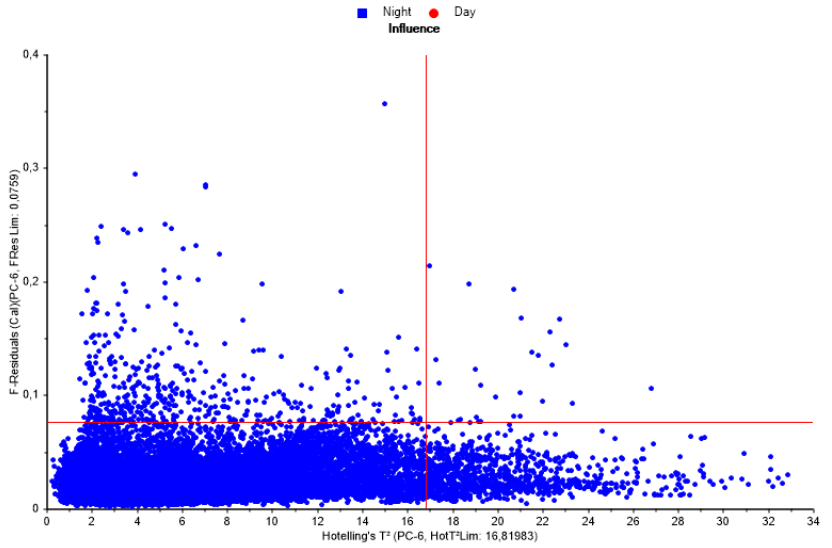
Principal Component Analysis (3rd approach)

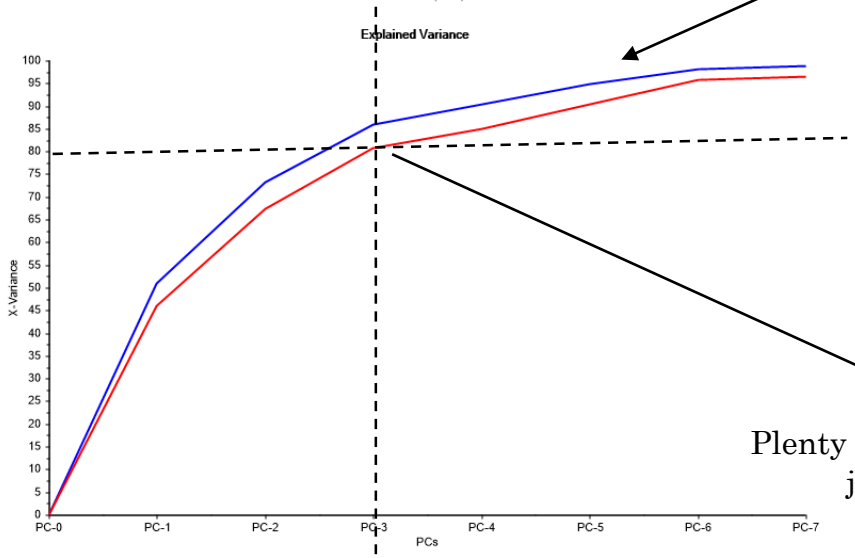
PCA algorithm: SVD





3 main directions describe most of our data



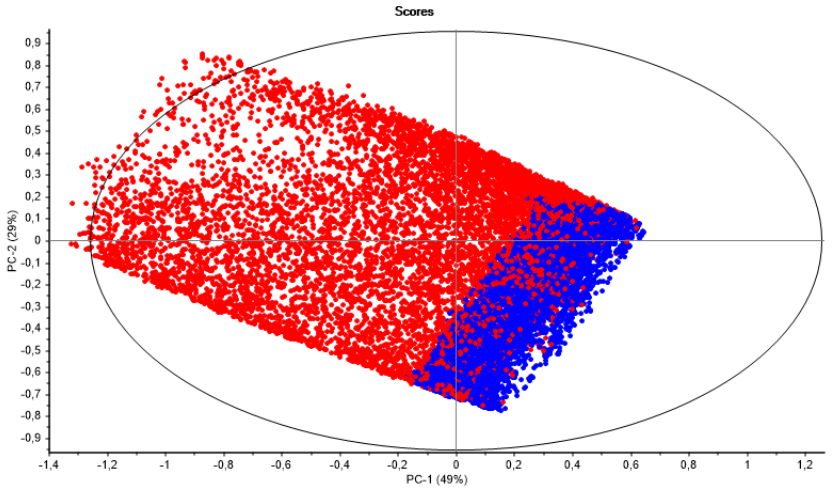


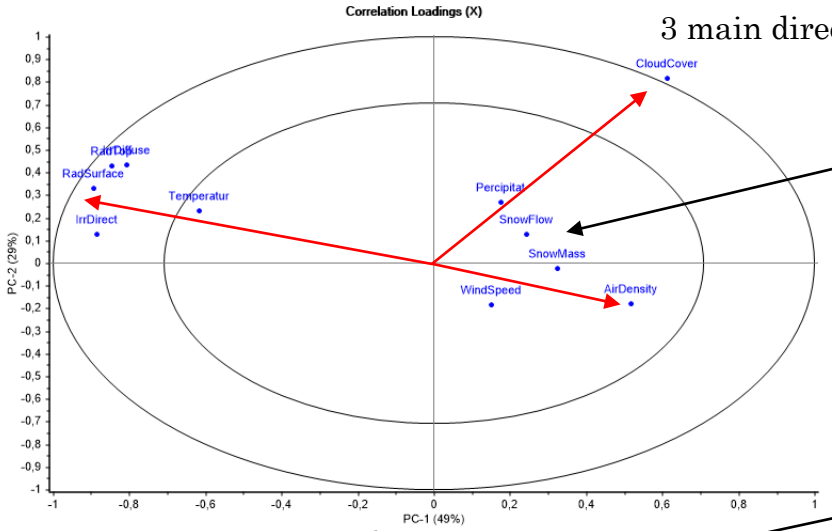
- Variables with less contribution to the PCs: *Precipitation, Wind Speed, Snow Mass, Snow Flow, Air Density*
- Fewer samples beyond the critical limits (leverage and F-statistic residuals)
- Higher explained variance at lower PCs number
- Many samples with high leverage but smaller residuals
- Clear separation of Day and Night

Plenty of explained variance just from 3 PCs

Principal Component Analysis (4th approach)

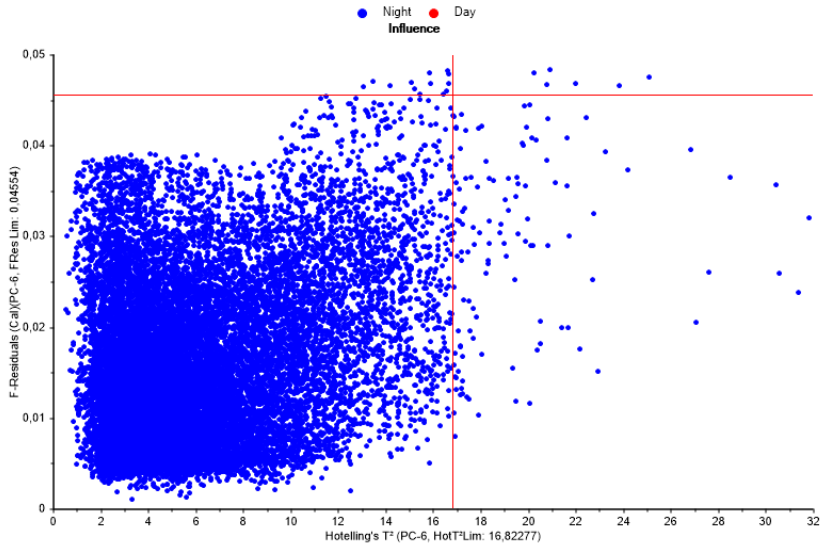
PCA algorithm: SVD, Outliers Removed

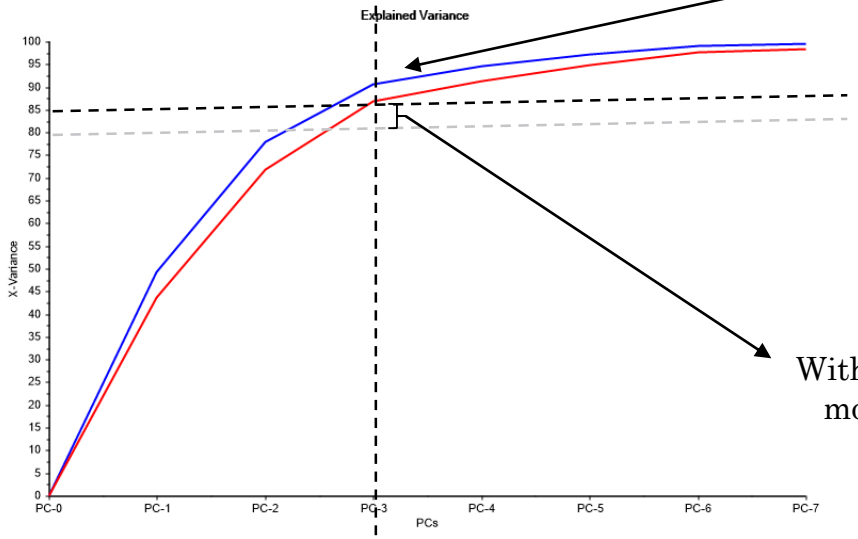




3 main directions describe most of our data

- Variables with less contribution to the PCs: *Precipitation, Wind Speed, Snow Mass, Snow Flow, Air Density, Temperature*
- Fewer outliers to suspect
- Best model of X variance
- Good separation of Day and Night





With several outlier removed, the model explains more variance

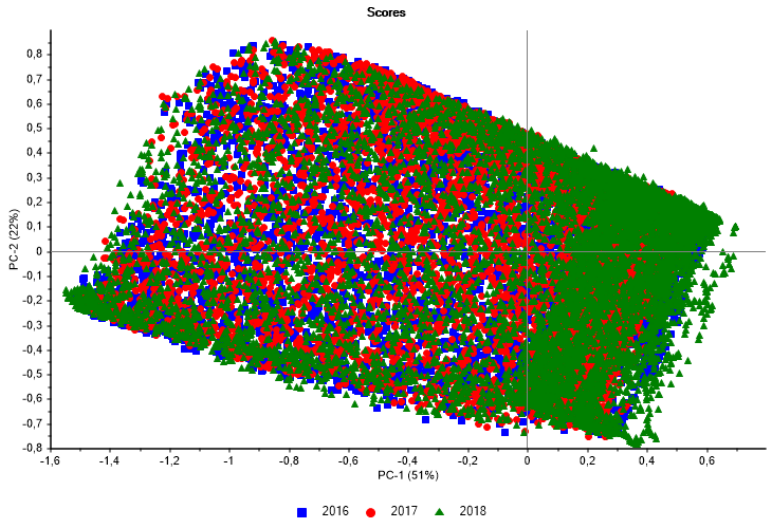
Where the removed samples correspond to ?

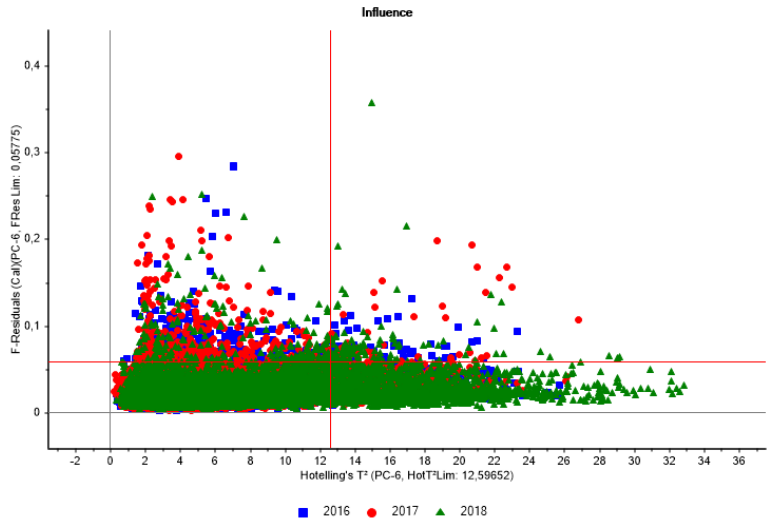
Can they indicate any anomaly ?

Irregular Samples (outlier detection)

Yearly Grouping

PCA: Approach 3

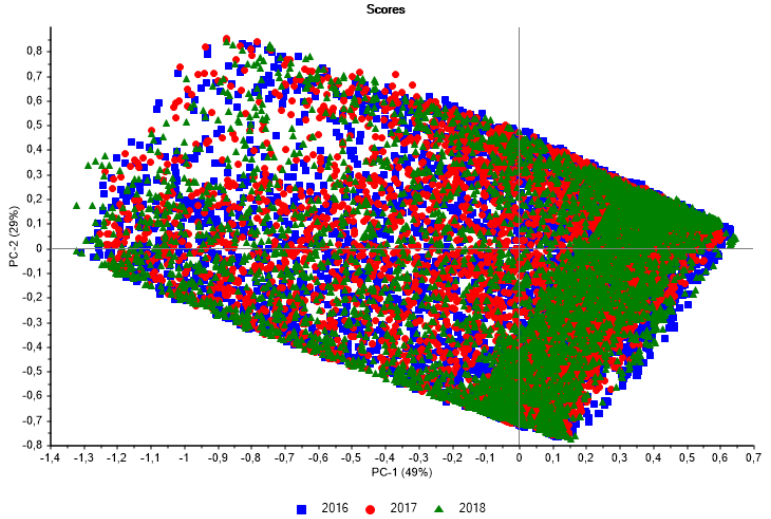


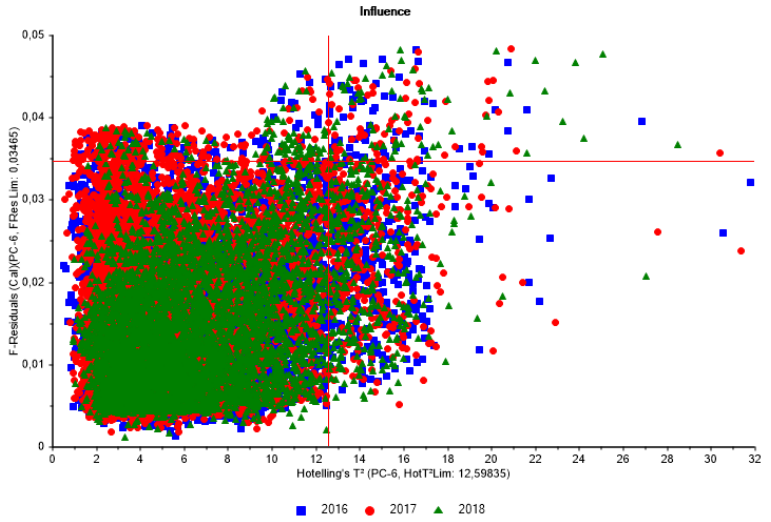


Green samples (= 2018) changed behavior when outliers removed

Suspect Year **2018**

PCA: Approach 4

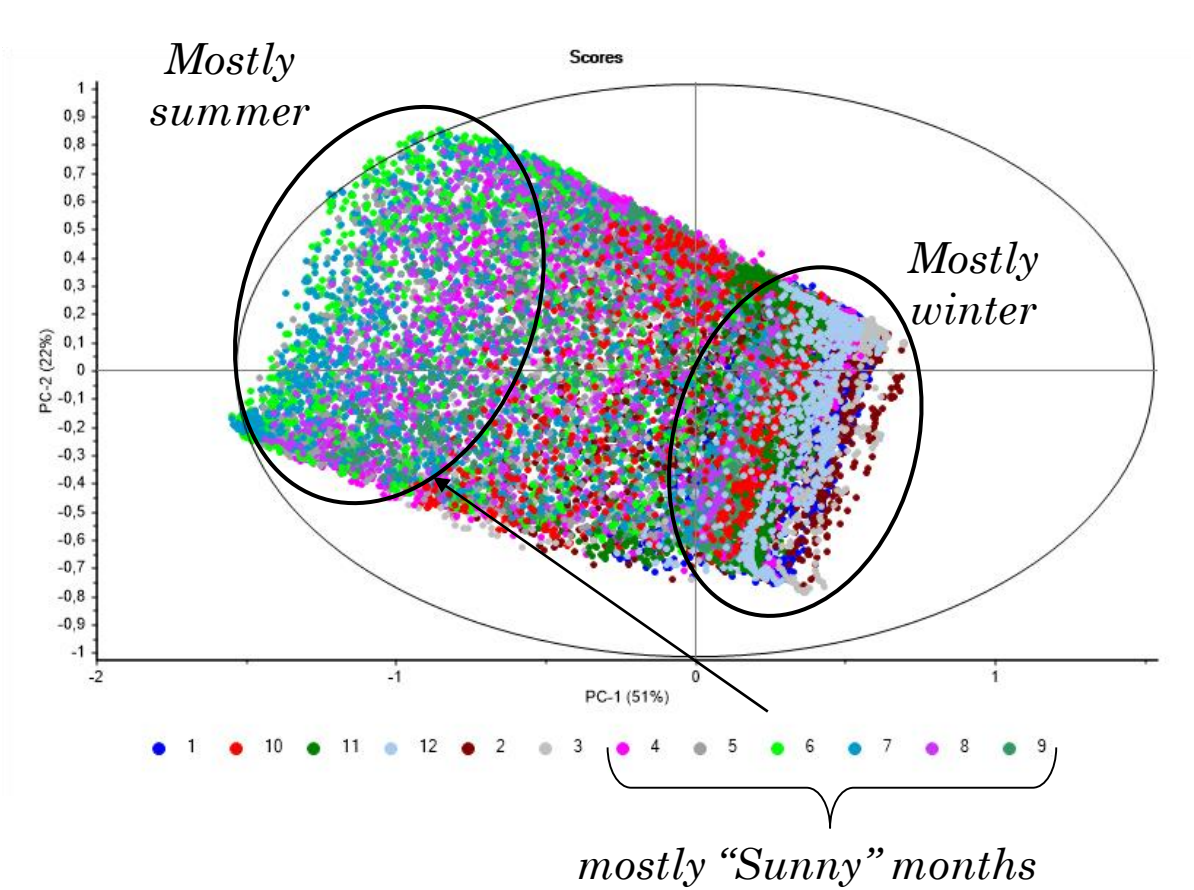




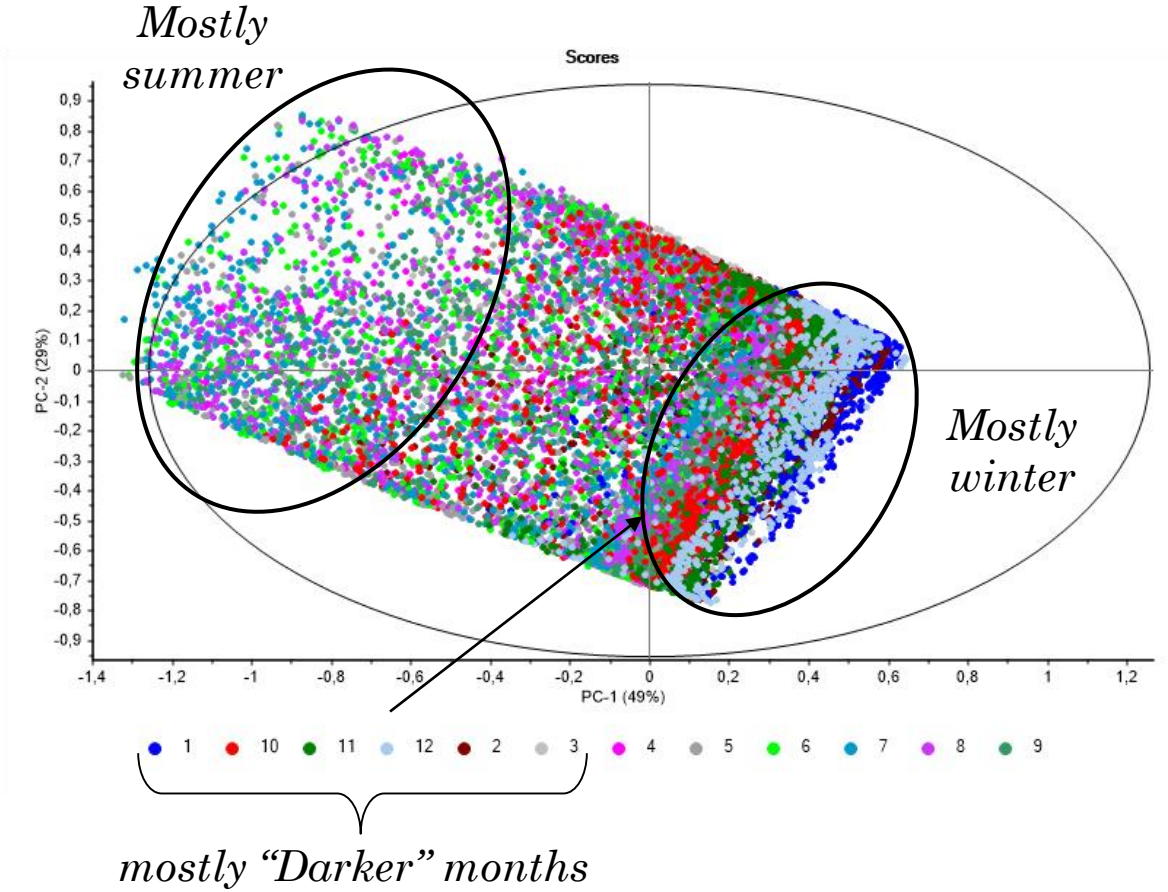
Irregular Samples (outlier detection)

Monthly Grouping

PCA: Approach 3



PCA: Approach 4

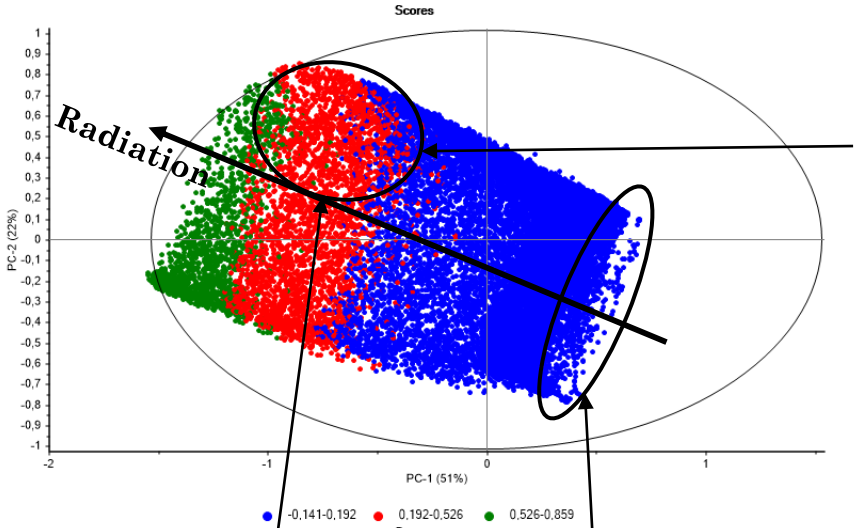


Suspect Year **2018 &**
"mostly sunny" months

Irregular Samples (outlier detection)

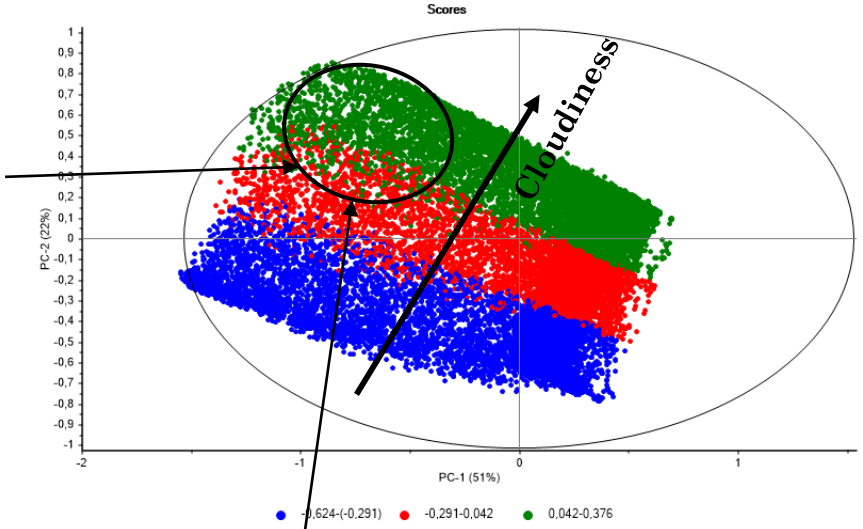
Approach 3

Grouping Variable: Surface **Solar Radiation**



Samples of *Medium to High* Cloudiness **and** Radiation

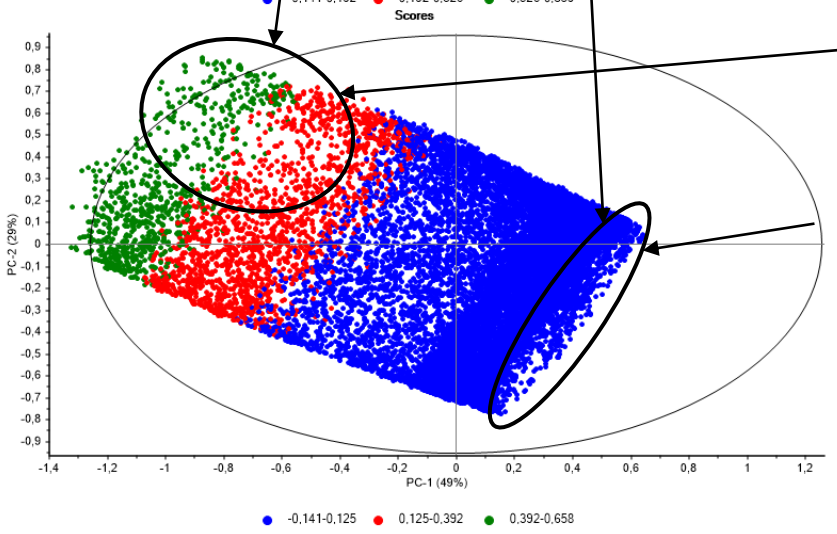
Grouping Variable: **Cloud Coverage**

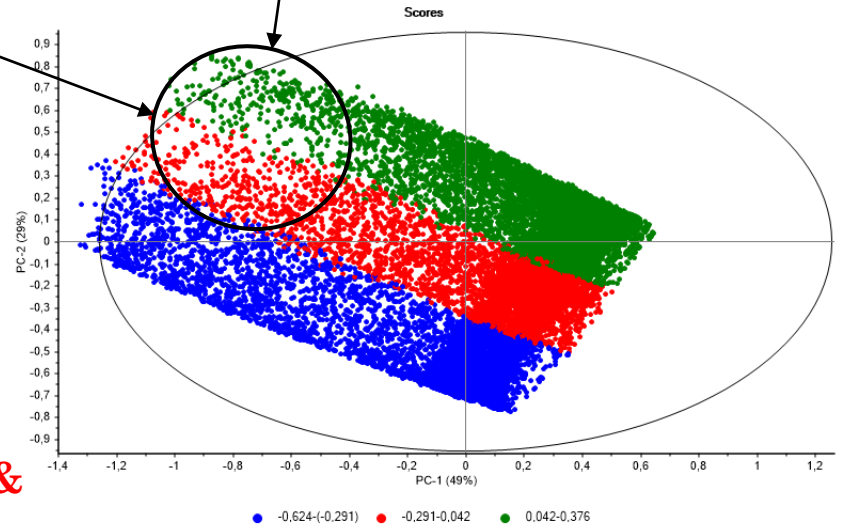


Approach 4

Such samples were among the ones removed from the

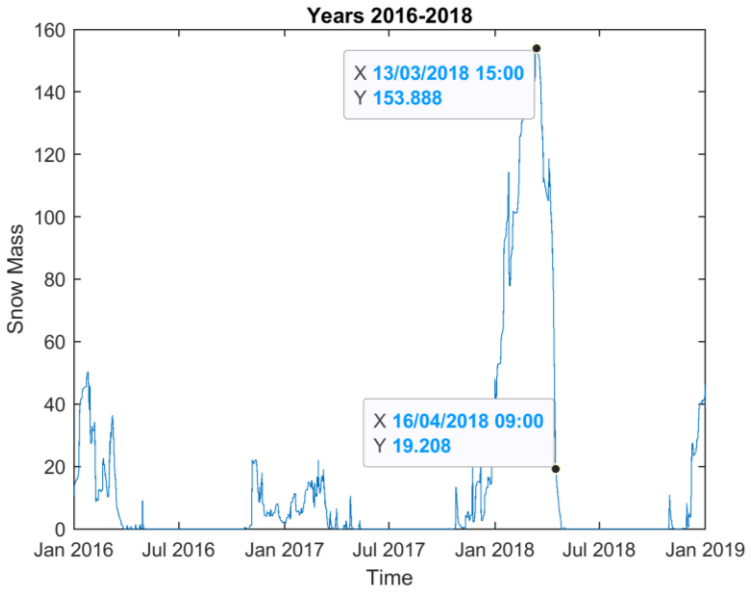
Better winter months clustering



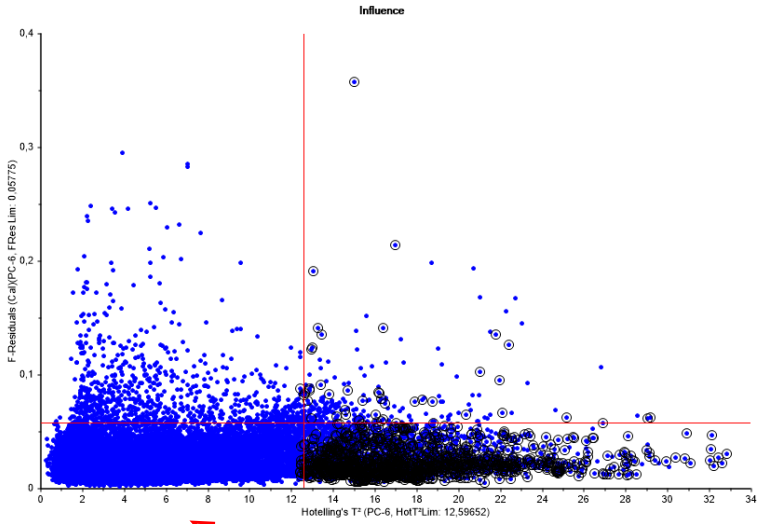


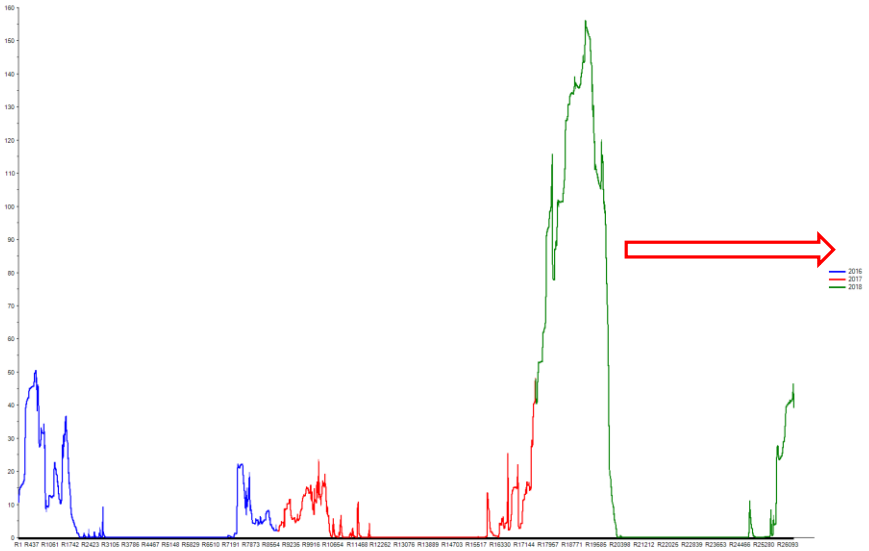
Suspect Year **2018 &** ***"mostly Summer"*** months **& Possibly Spring** months

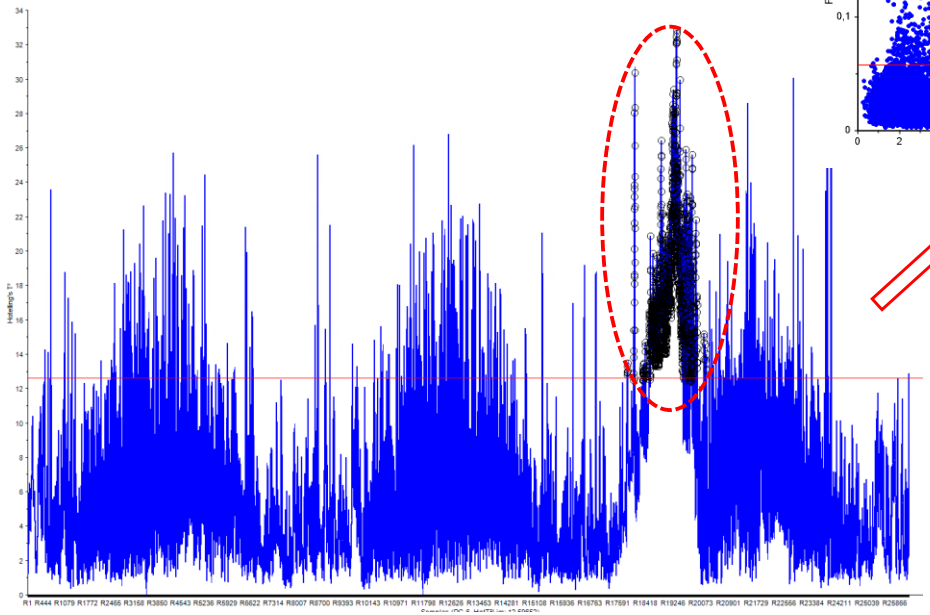
Irregular Samples (outlier detection)



- Samples with high leverage (*i.e. high* T^2) affect the model's explained variance (*PCA* 3^{rd} vs 4^{th})
- In 2018 there was significant Snow Mass during the first months
- Irregularity identified at year **2018,** on months **March-April**





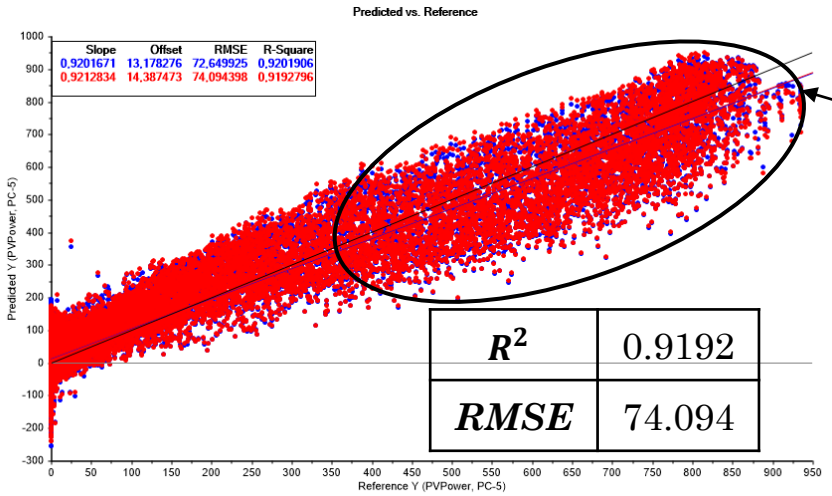


4. Regression Analysis

Principal Component Regression (PCR-1)

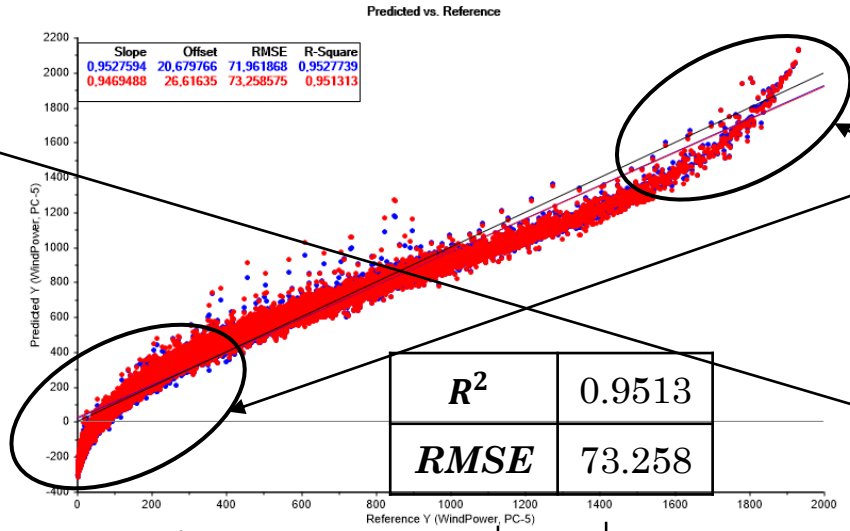
PCR algorithm: PCA (1st approach) +MLR (least squares to the PCs)

PV Power

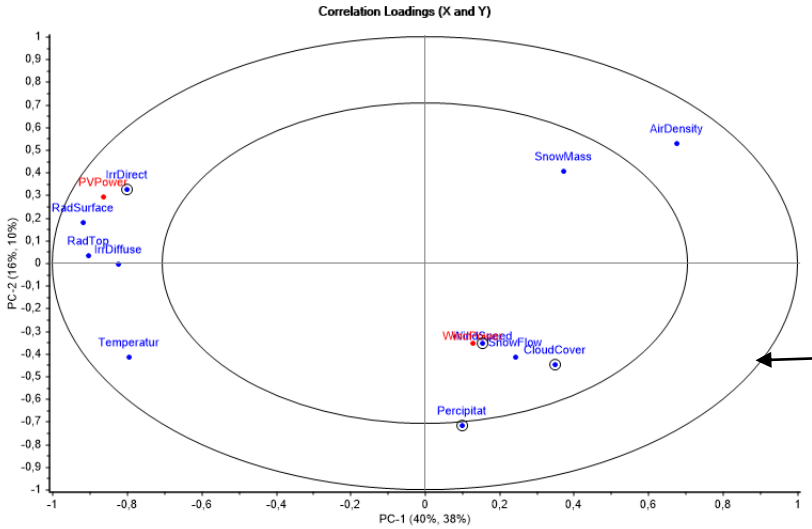


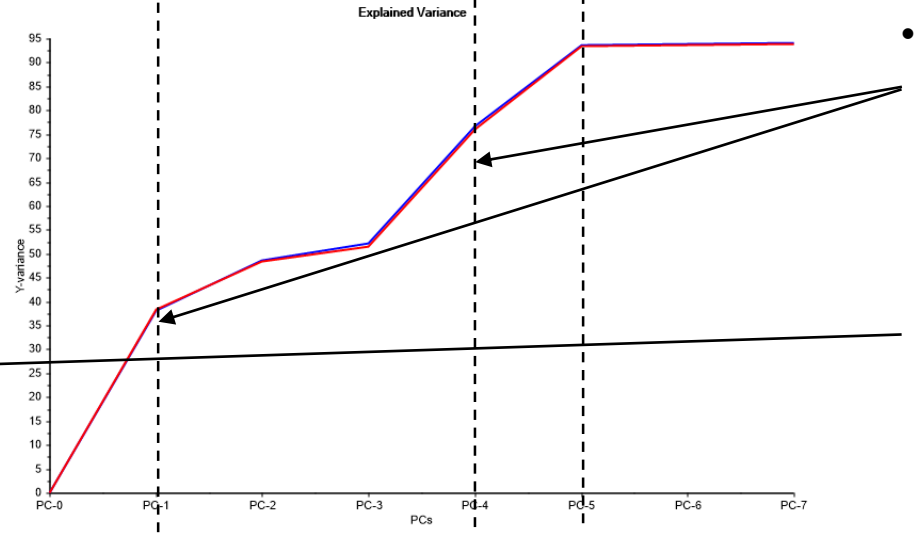
R^2	0.9192
RMSE	74.094

Wind Power



R^2	0.9513
RMSE	73.258





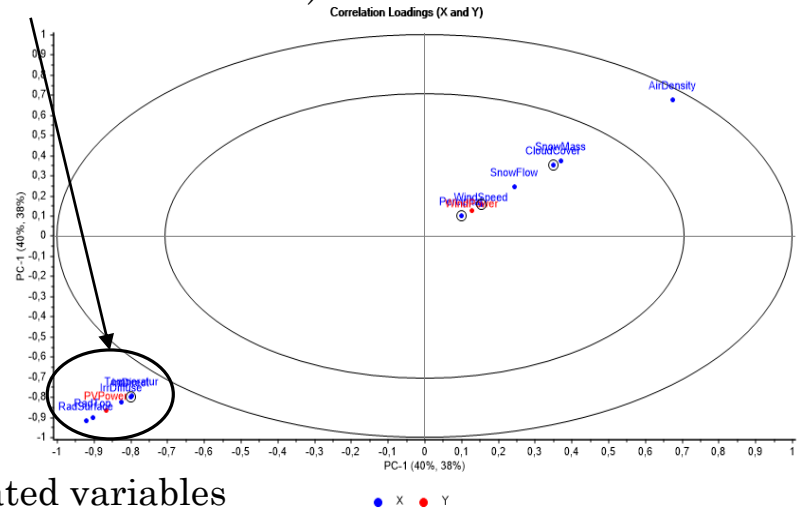
*X: **5 PCs***
Y: 2 outputs (PV & Wind Power)

- The upper and lower parts are not well explained → *Non-linear Wind Turbine Power curve*
- Higher spread for higher PV power values → Possibly due to increased T (*non-linear efficiency effect*)
- Explained variance of the X dataset does not correspond to best Y explanation → Stepwise improvement for: *PC1, PC4, PC5*

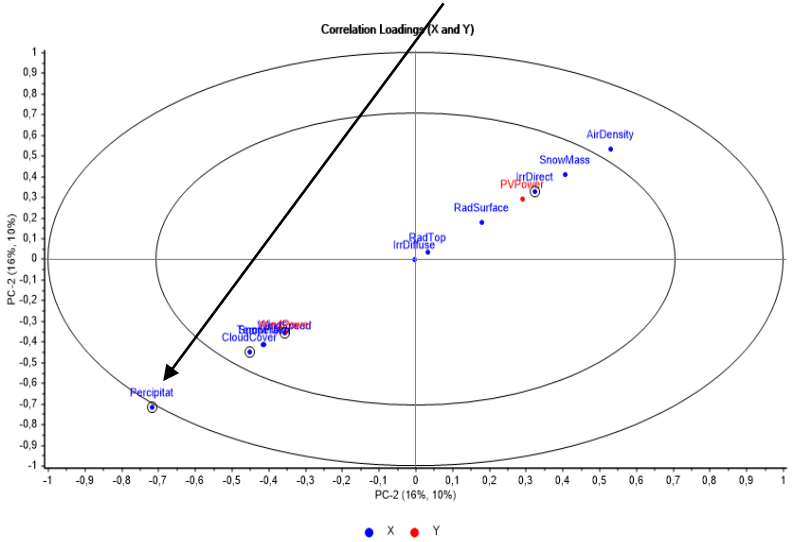
Principal Component Regression (PCR-1)

Contribution of *PCs* to *Y* predictions

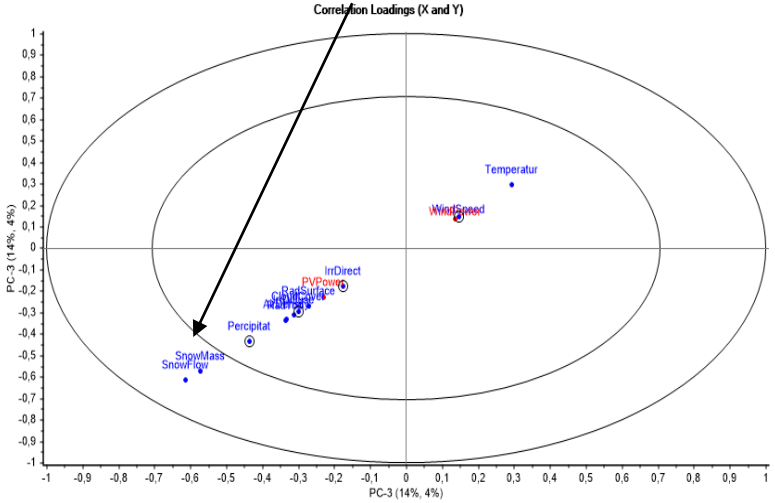
PC 1 → PV Power related variables
(*solar radiation related variables*)



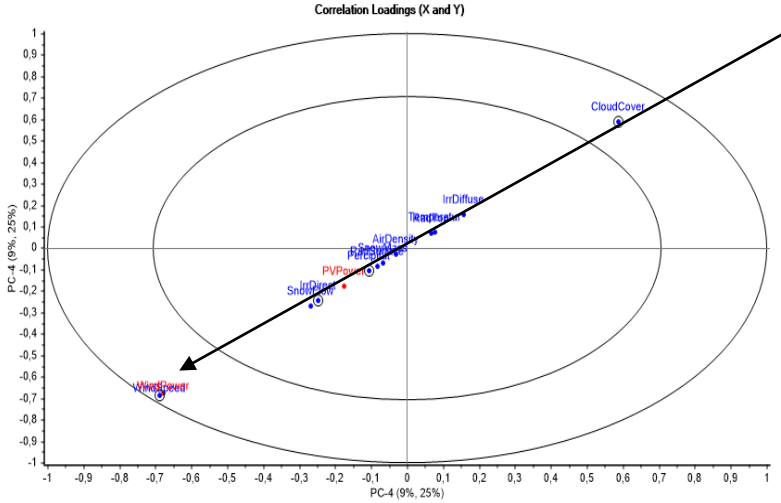
PC 2 → Precipitation (*Does not explain well a Y variable*)

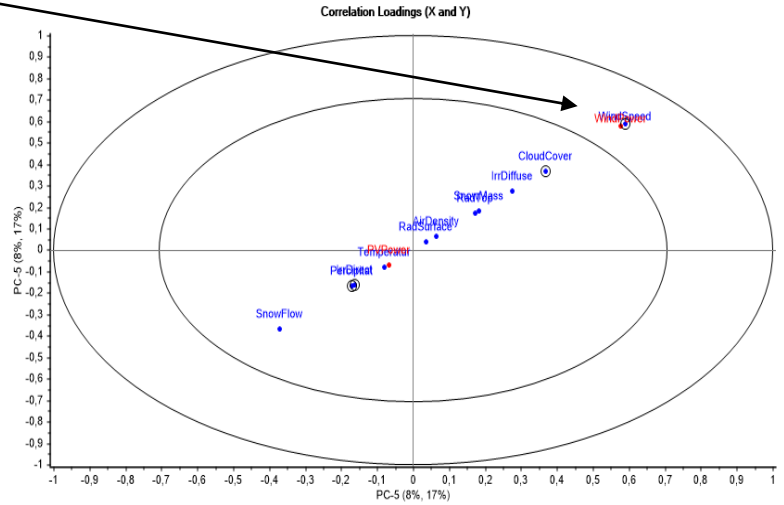


PC 3 → Snow related variables
(*Do not explain well Y variables*)



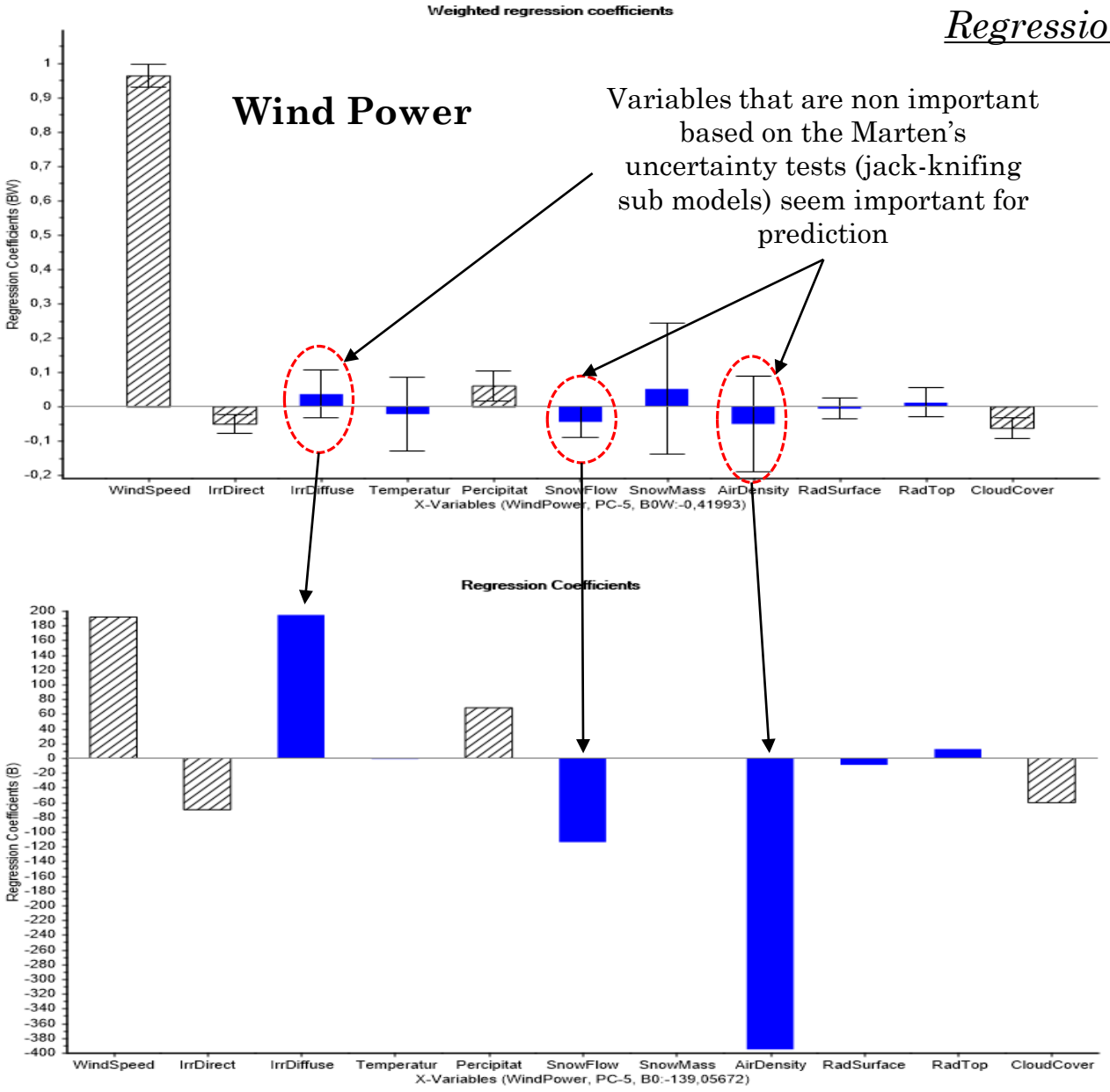
PC 4 & PC 5→ Wind Power related variable (*Wind Speed*)

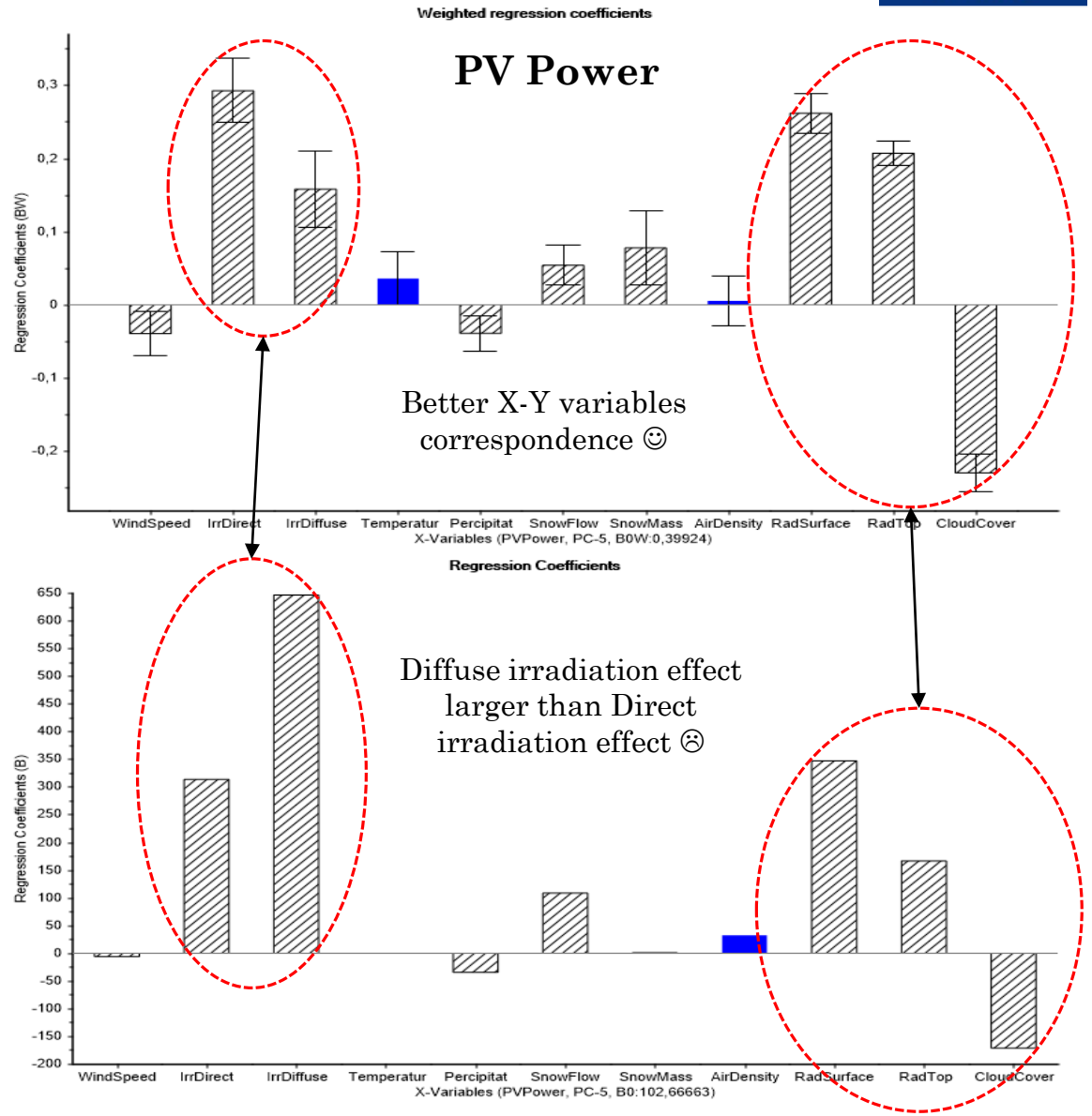




Principal Component Regression (PCR-1)

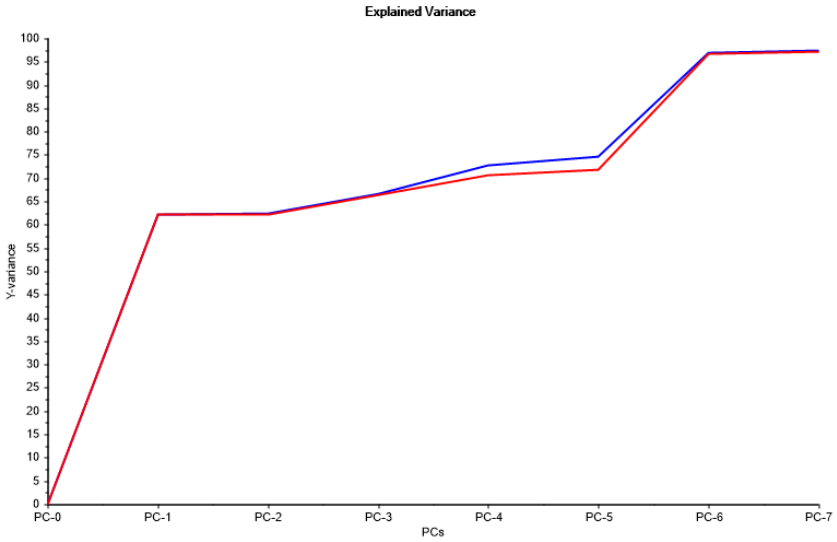
Regression Coefficients

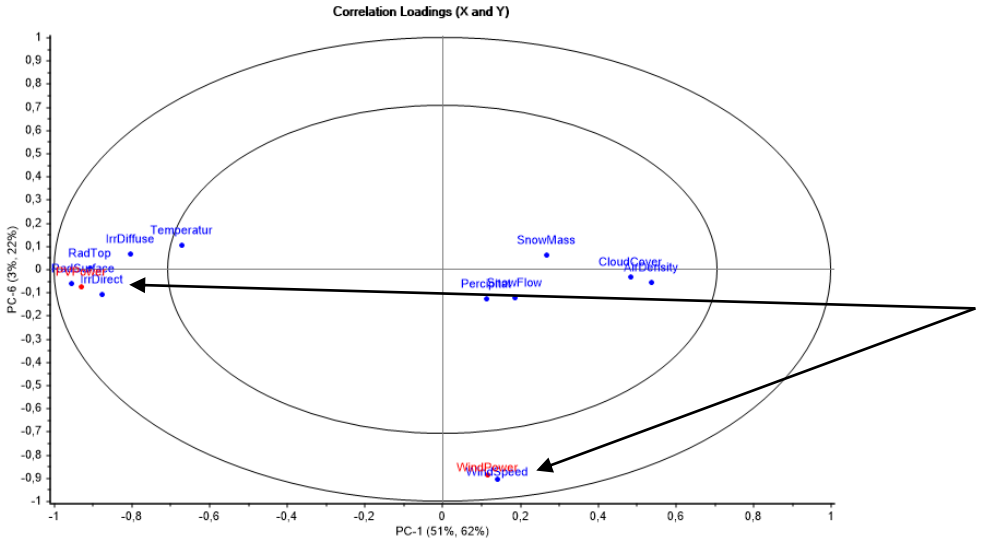




Principal Component Regression (PCR-2)

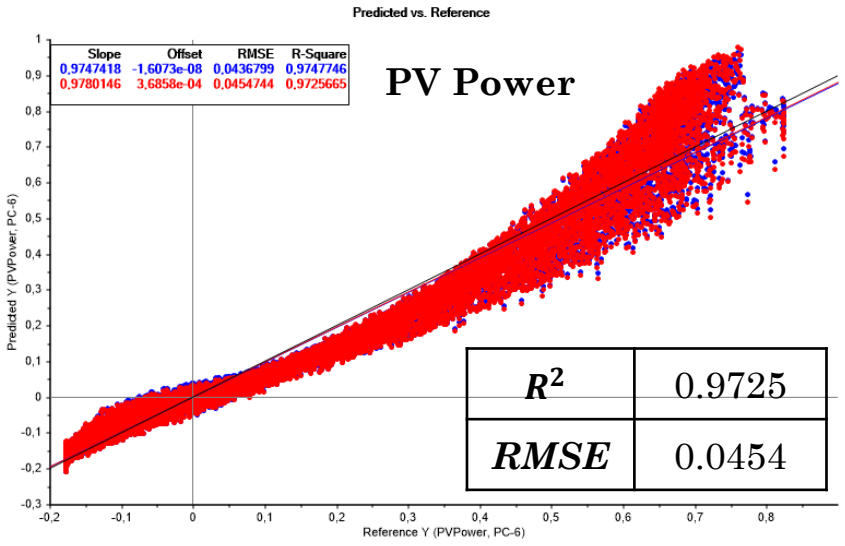
PCR algorithm: PCA (3rd approach) +MLR (least squares to the scores)



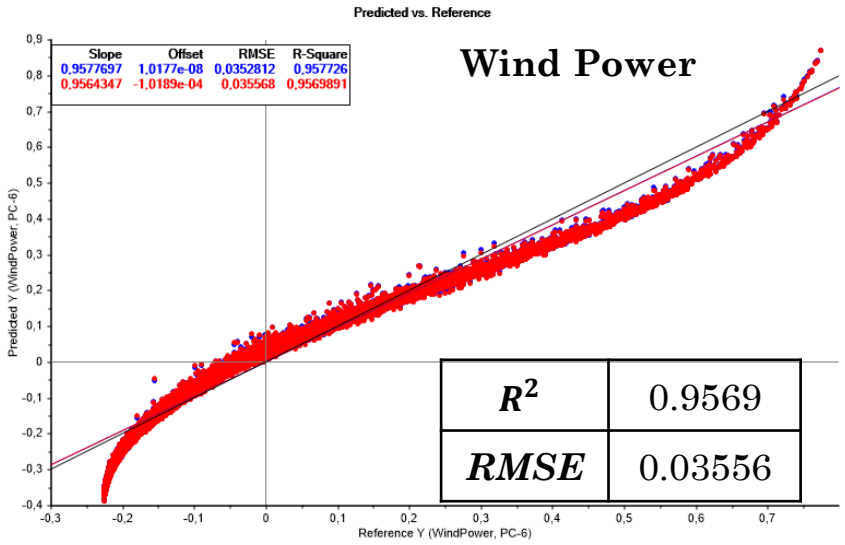


*X: **6 PCs***
Y: 2 outputs (PV & Wind Power)

- Best result → 6 PCs
- PCs that contribute more to explain Y are associated with 2 particular PCs (*PC1, PC6*)
- Same trend in *Y* variables prediction



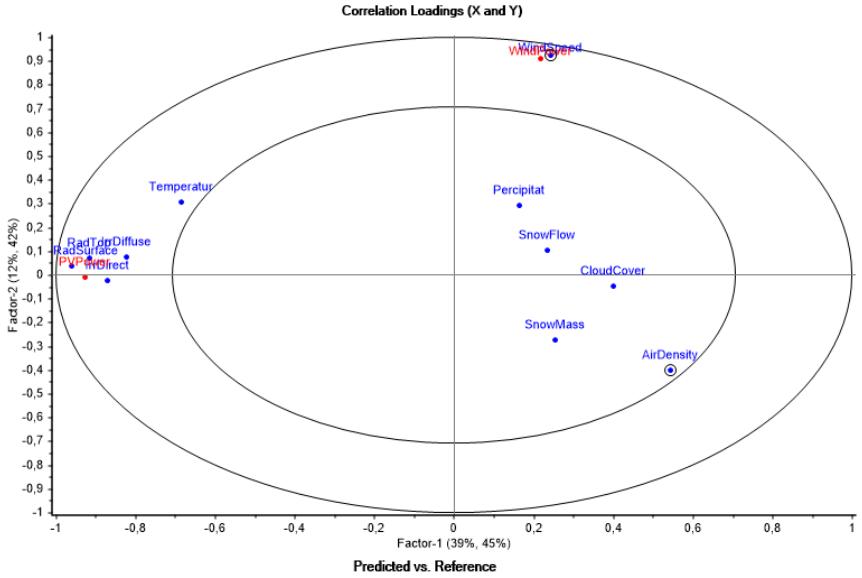
R^2	0.9725
RMSE	0.0454

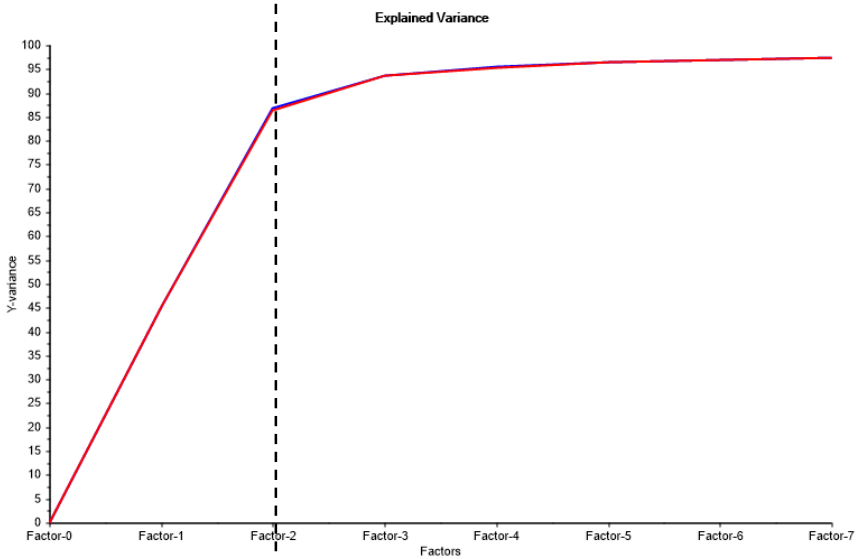


R^2	0.9569
RMSE	0.03556

Partial Linear Squares Regression (PLSR-1)

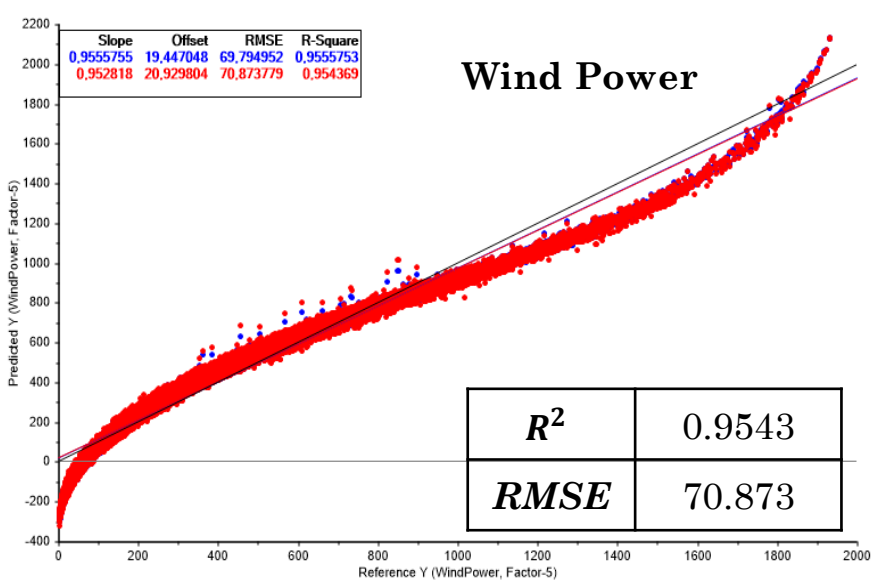
PLSR algorithm: NIPALS



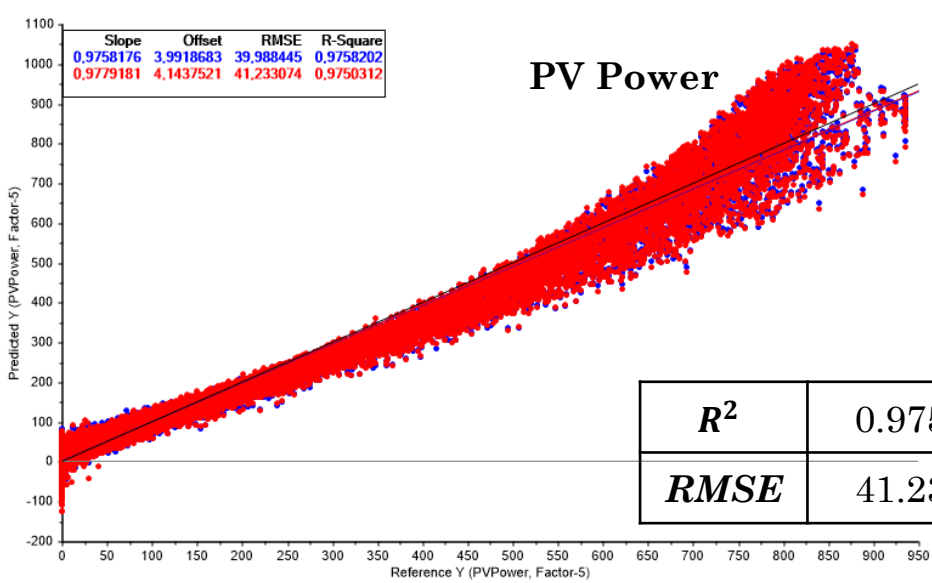


*X: **5 Factors***
Y: 2 outputs (PV & Wind Power)

- Two most important Factors directly identified
- Y variance plot better than PCR
- Same trend in *Y* variables prediction



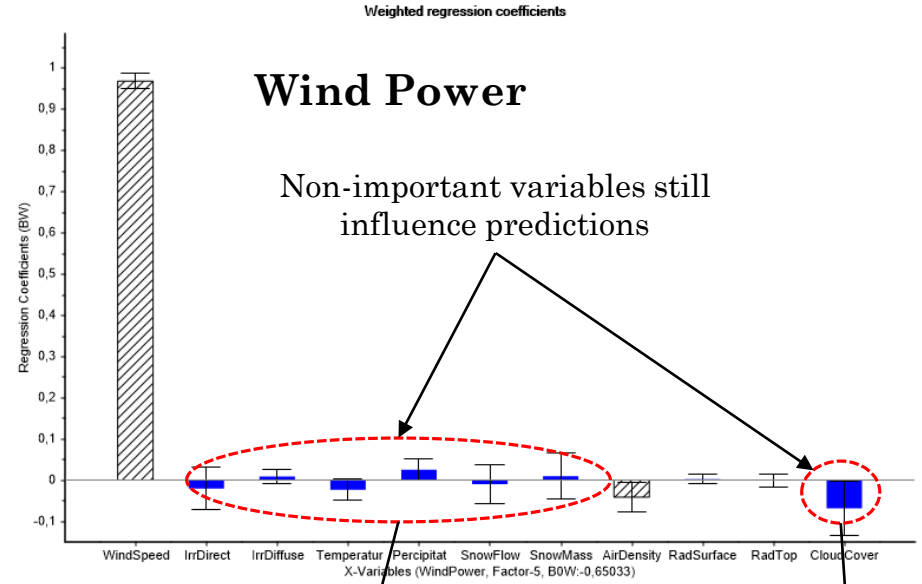
R^2	0.9543
RMSE	70.873



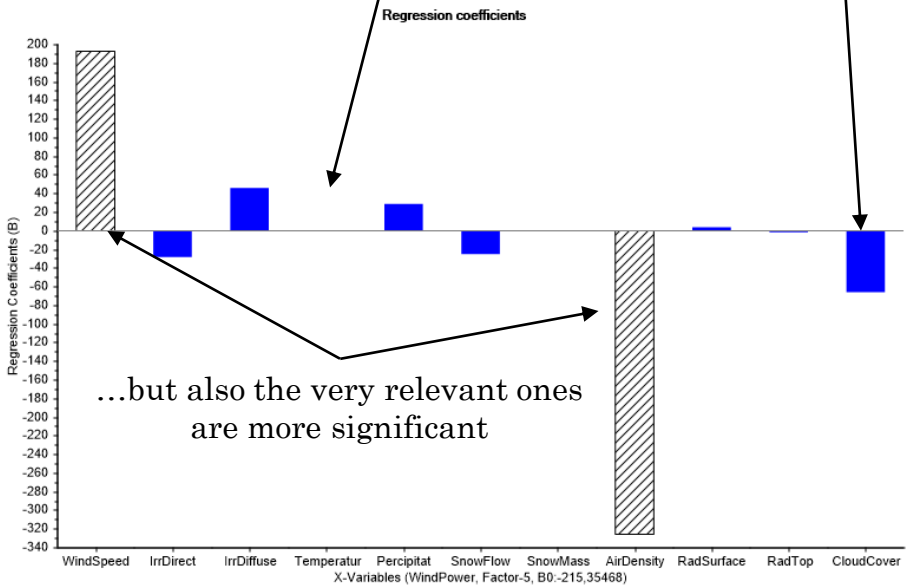
R^2	0.9750
RMSE	41.233

Partial Linear Squares Regression (PLSR-1)

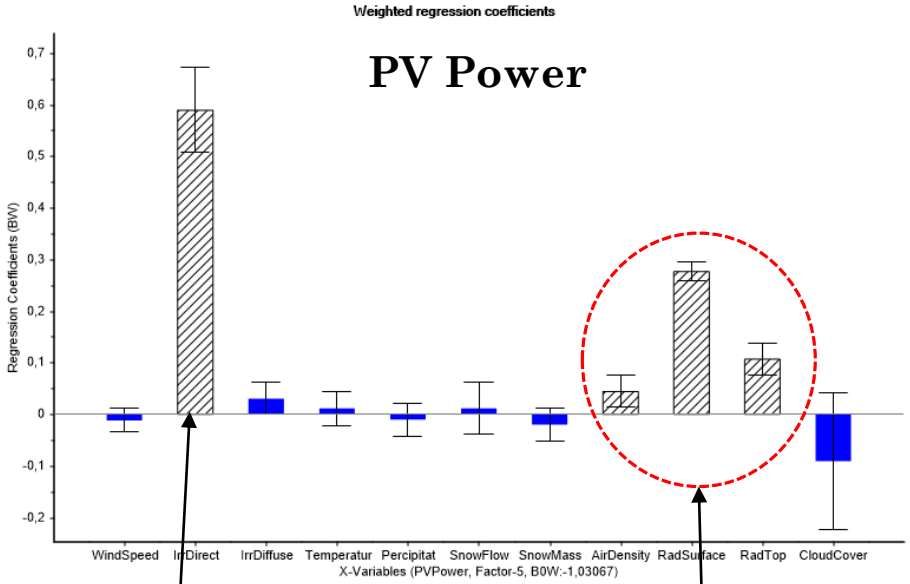
Regression Coefficients

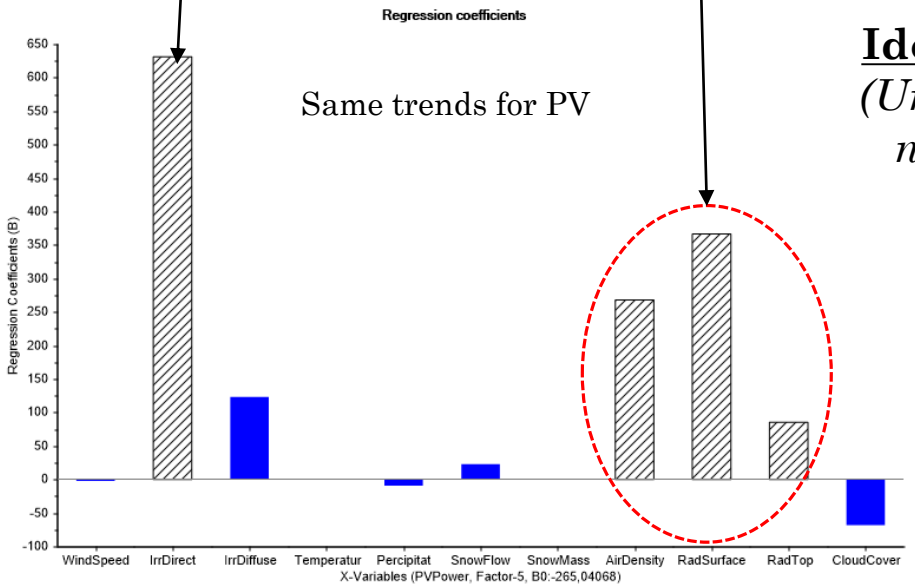


Non-important variables still influence predictions



…but also the very relevant ones are more significant



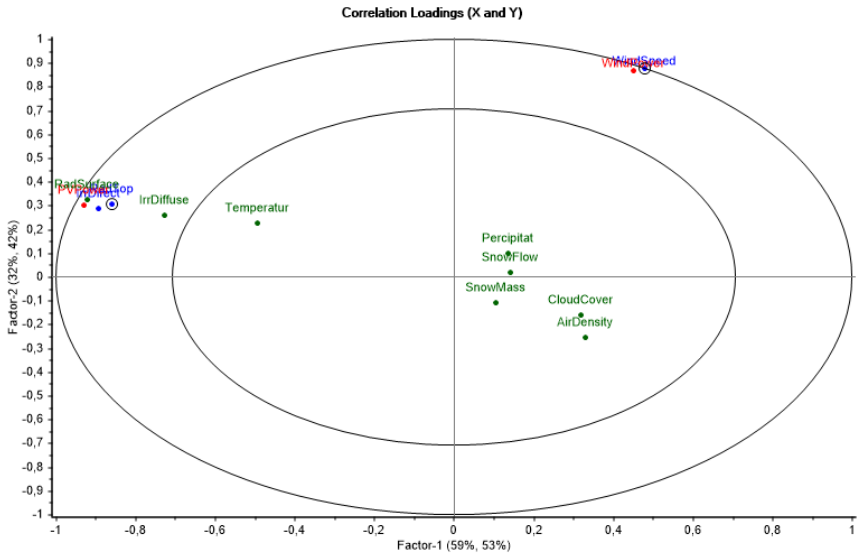


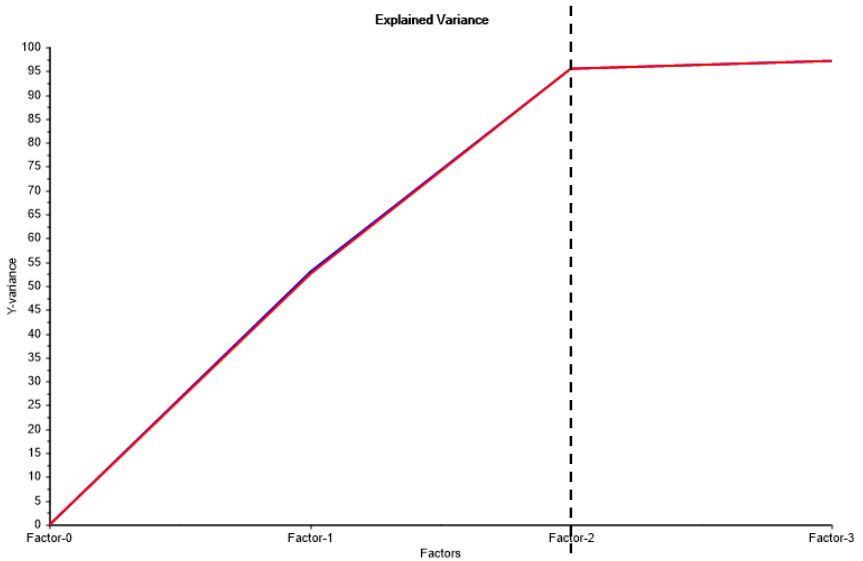
Same trends for PV

Idea: *Downweigh (Unscrambler) the non-important variables !*

Partial Linear Squares Regression (PLSR-2)

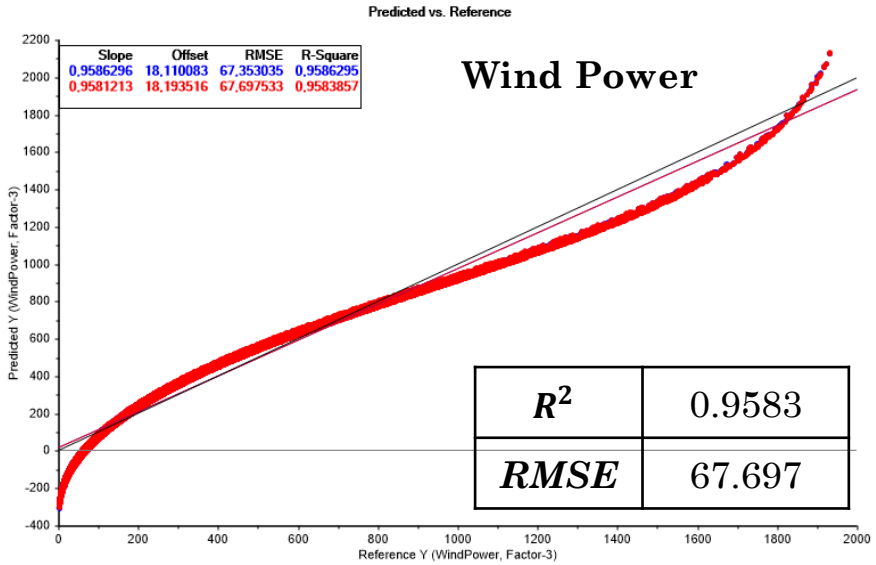
PLSR algorithm: NIPALS



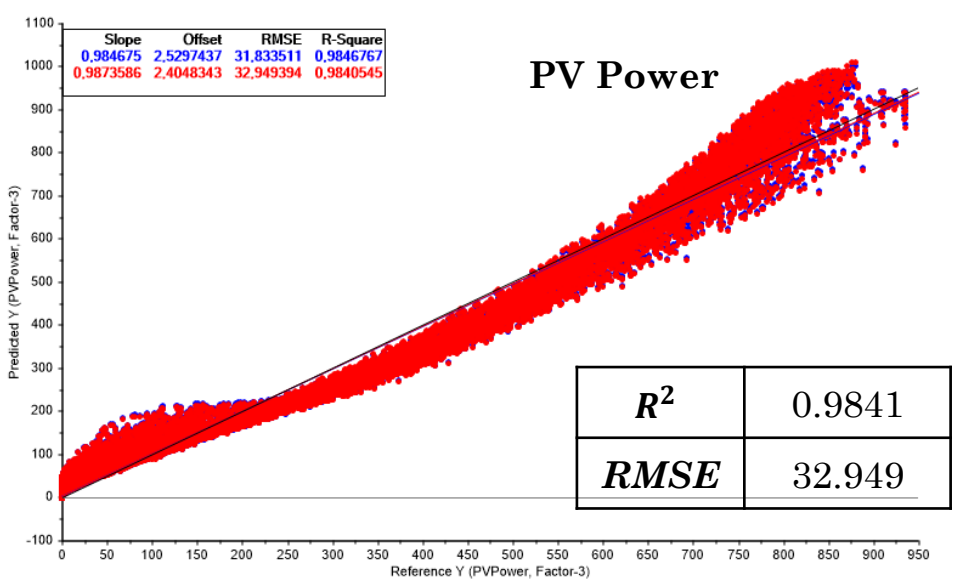


X: ***3 Factors***
Y: 2 outputs (PV & Wind Power)

> 90% Y-variance just with 2 factors !



R^2	0.9583
RMSE	67.697

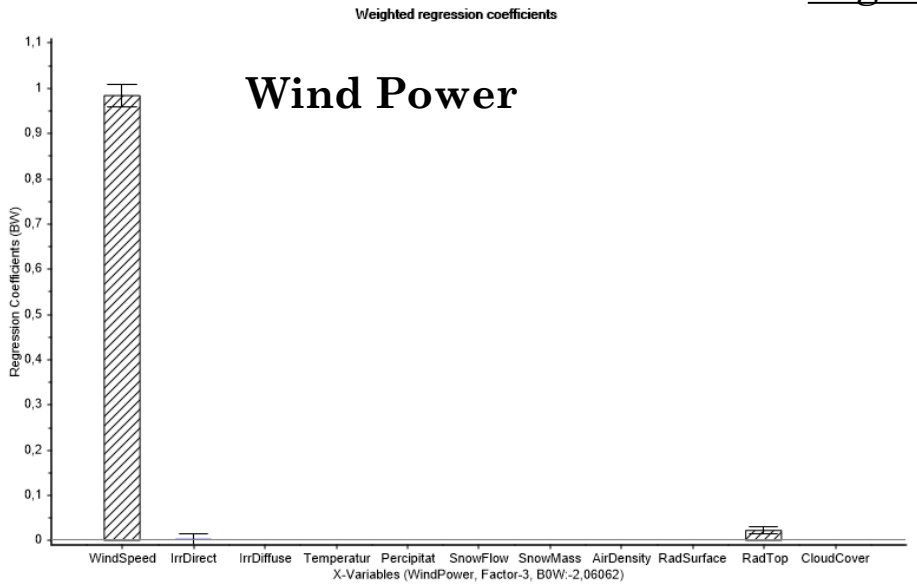


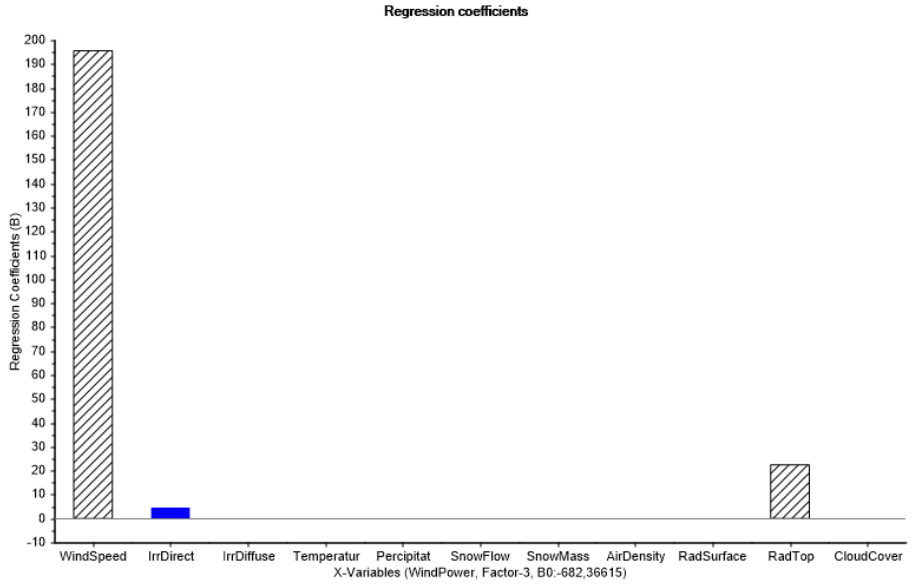
R^2	0.9841
RMSE	32.949

Partial Linear Squares Regression (PLSR-2)

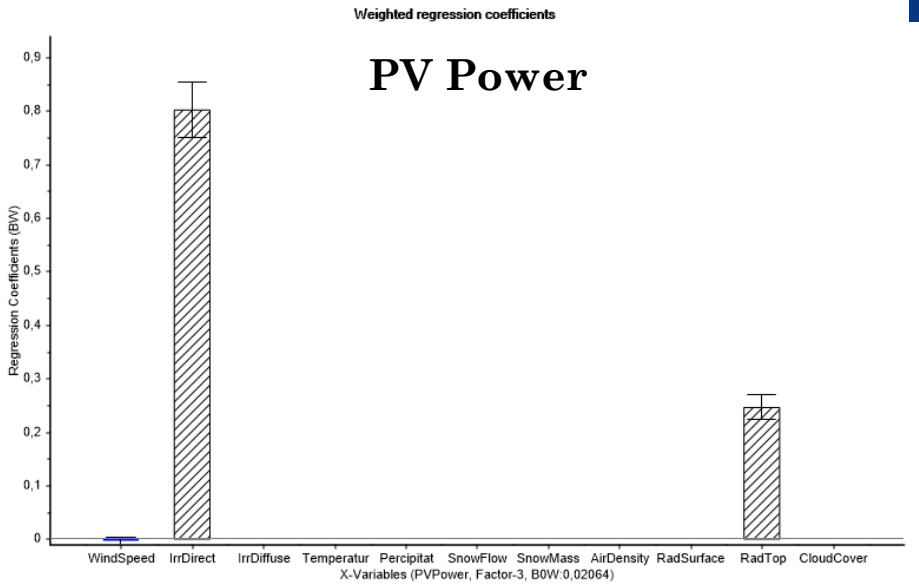
Regression Coefficients

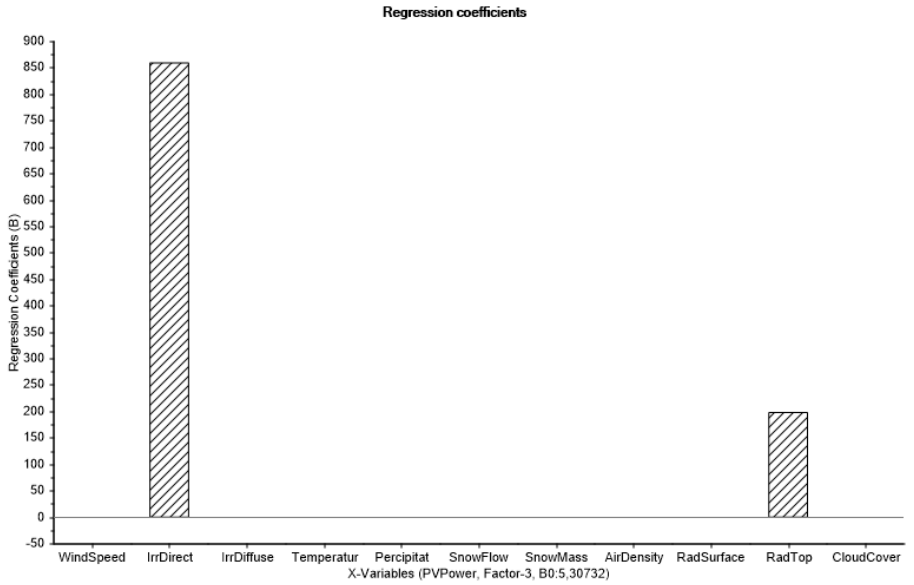
Wind Power





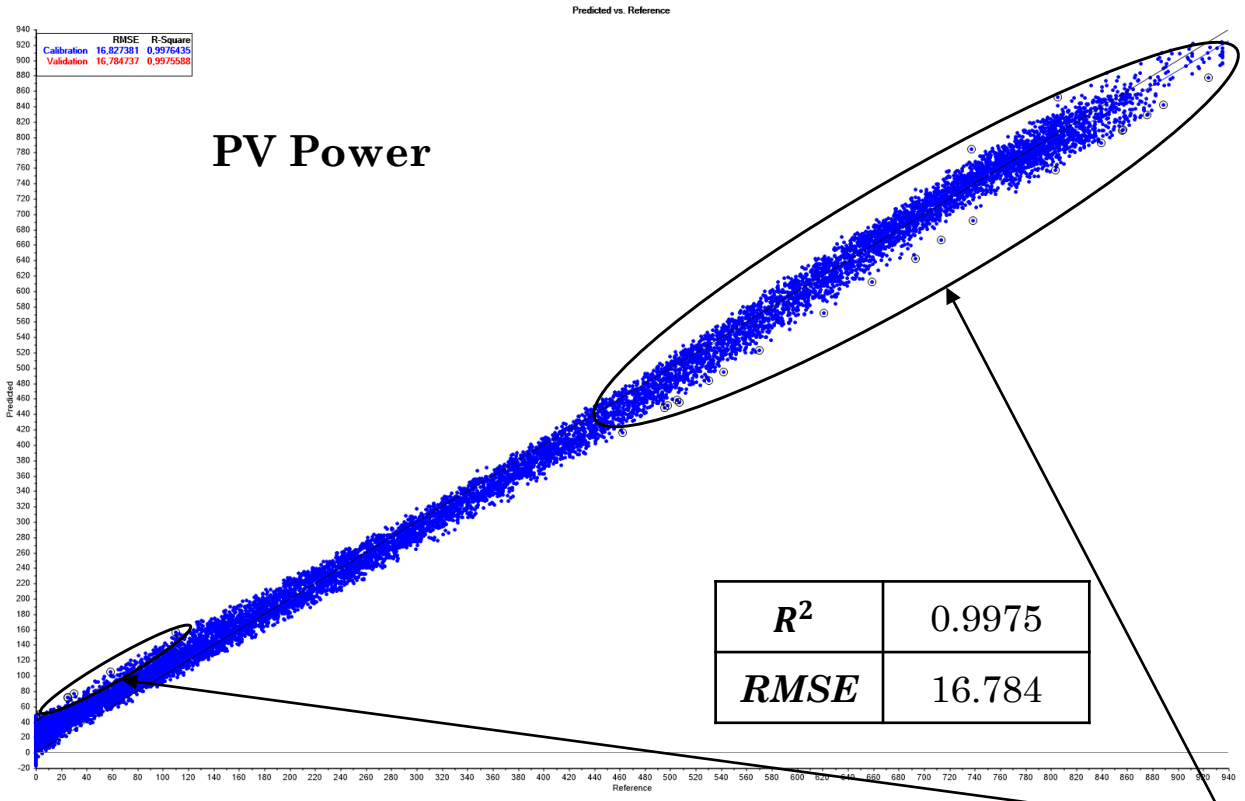
PV Power





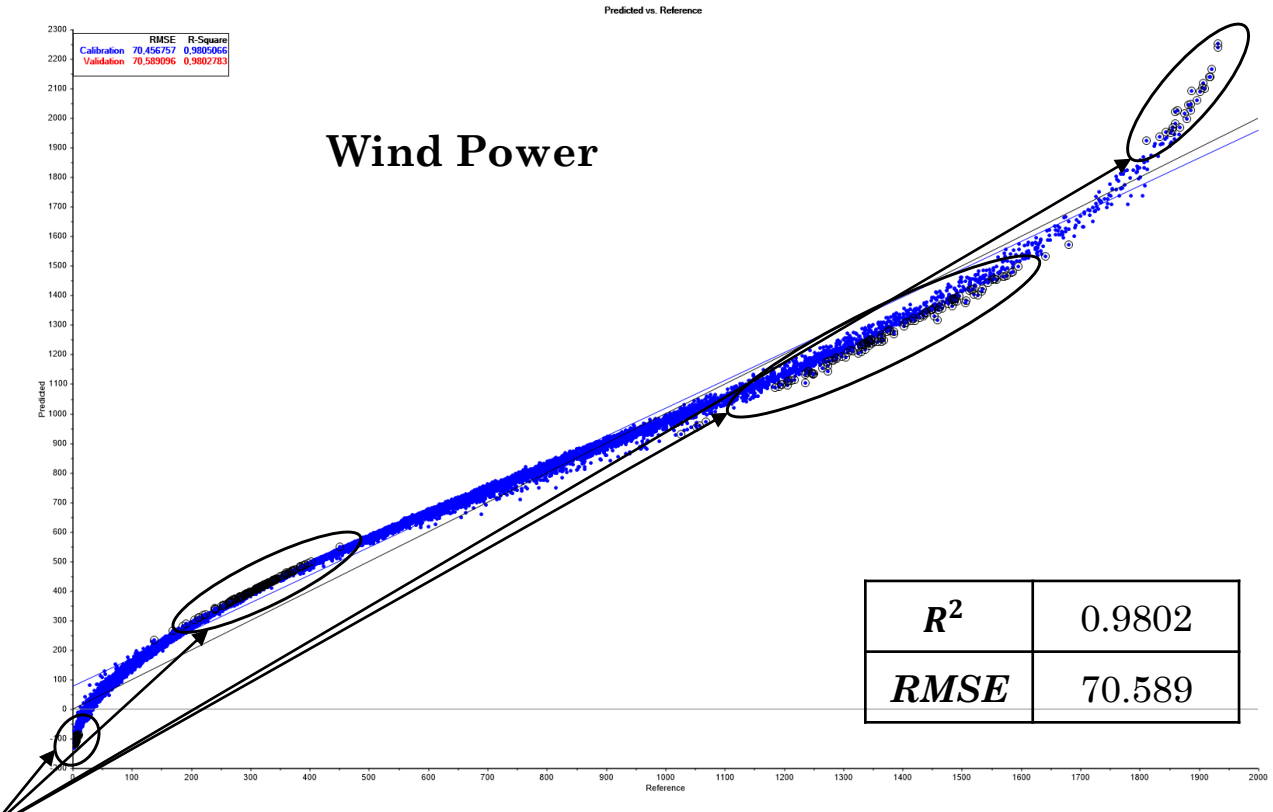
Support Vector Regression (SVR)

- ε-SVR algorithm
- Radial Basis Kernel (non-linear mapping to infinite dimensions)
- Data scaling [-1,1]



PV Power

R^2	0.9975
RMSE	16.784



Wind Power

R^2	0.9802
RMSE	70.589

Many support vectors in these regions →
These regions were hard to be modeled with Linear Regression methods → Possibly those are important non –linear regions

Regression models comparison

***Y(1)* – PV Power**					
	PCR-1st	*PCR-2nd*	*PLSR-1*	*PLSR-2*	*SVR*
Factors/PCs	5	6	5	3	-
R^2	0.9192	0.9725	0.9750	0.9841	0.9975
RMSE	74.094	0.0454	41.233	32.949	16.784
***Y(2)* – Wind Power**					
R^2	0.9513	0.9569	0.9543	0.9583	0.9802
RMSE	73.258	0.03556	70.873	67.697	70.589

Same pre-processing (z-score)

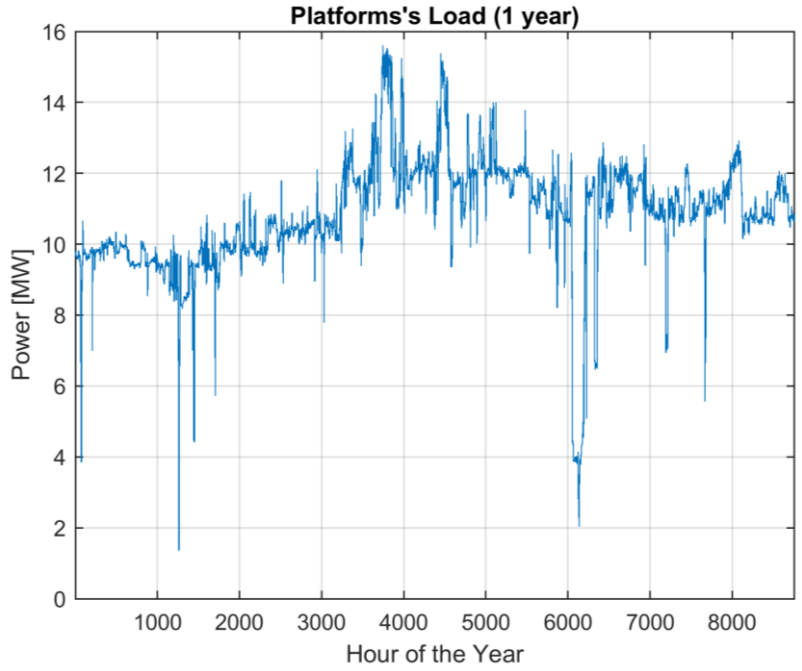
Generally, in terms of R^2 it was found: *SVR* > *PLSR* > *PCR*

5. Clustering Analysis

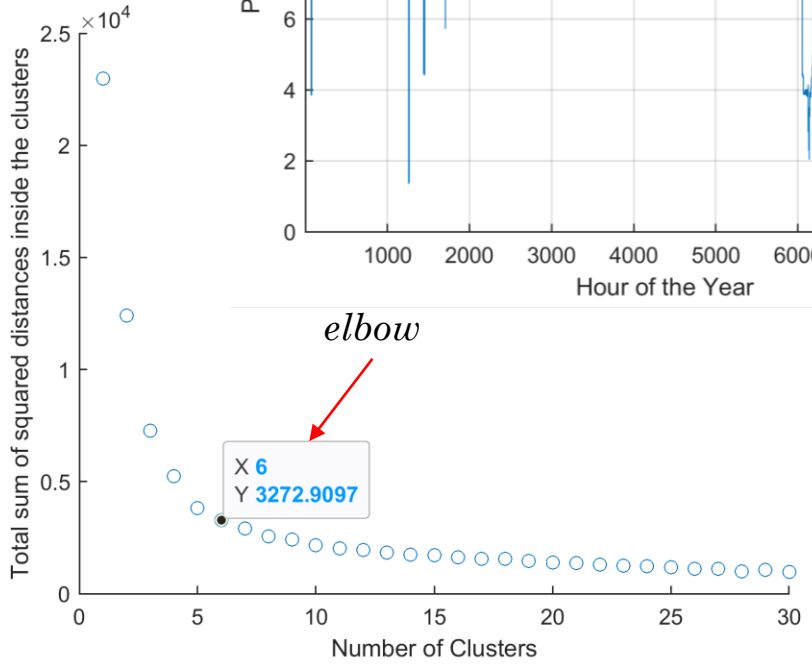
K-means Clustering

- **Data**: Platform's Daily Load Curve *(24 dimensional vectors)*

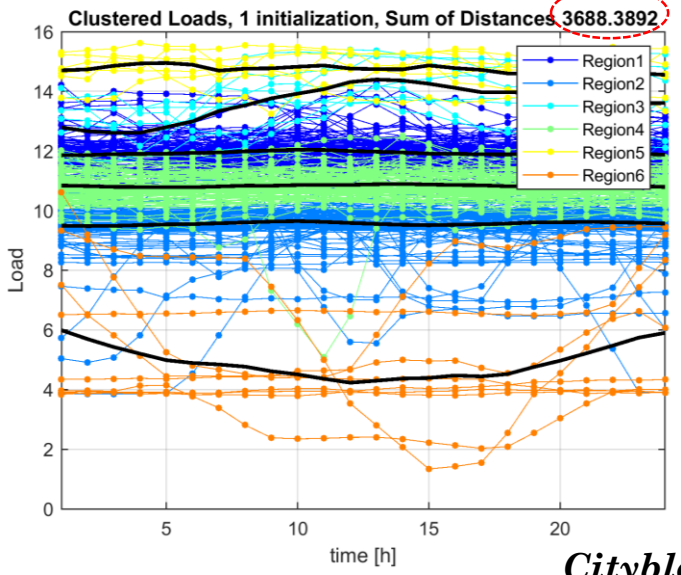
- **Distance** : Squared Euclidean / Cityblock
- **Initializations** : 10000

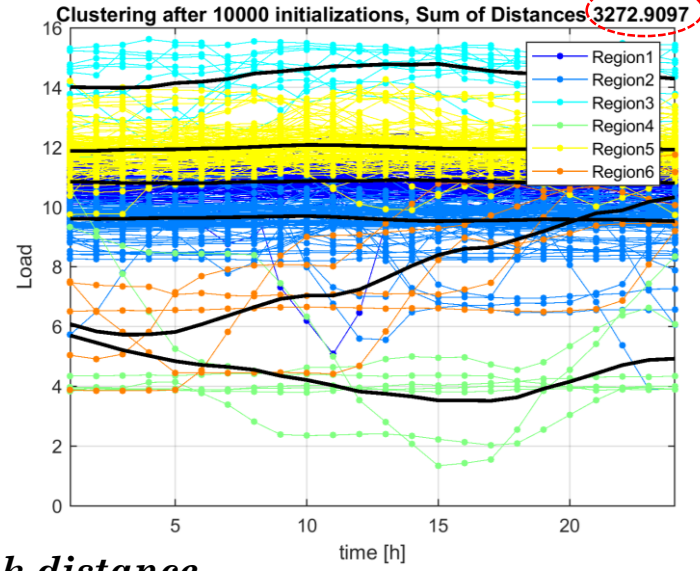


Euclidean distance

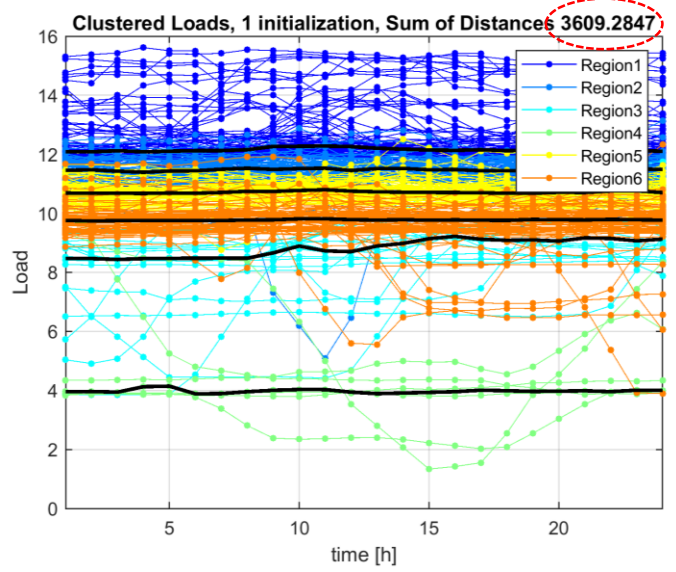


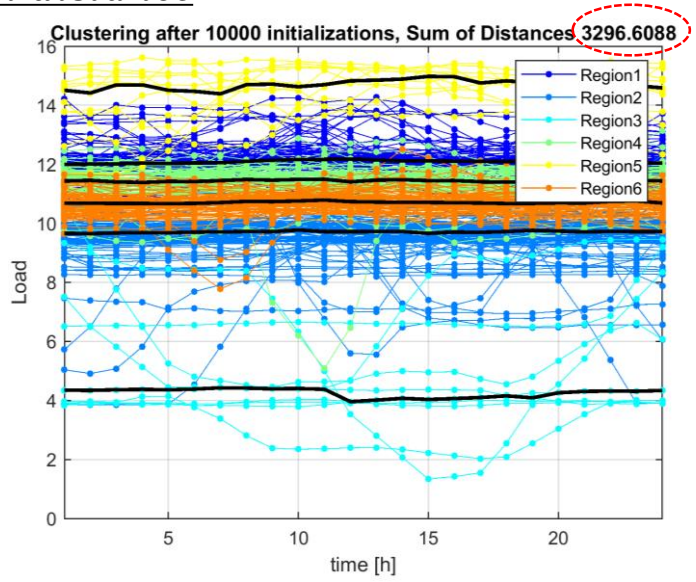
Sq. Euclidean distance



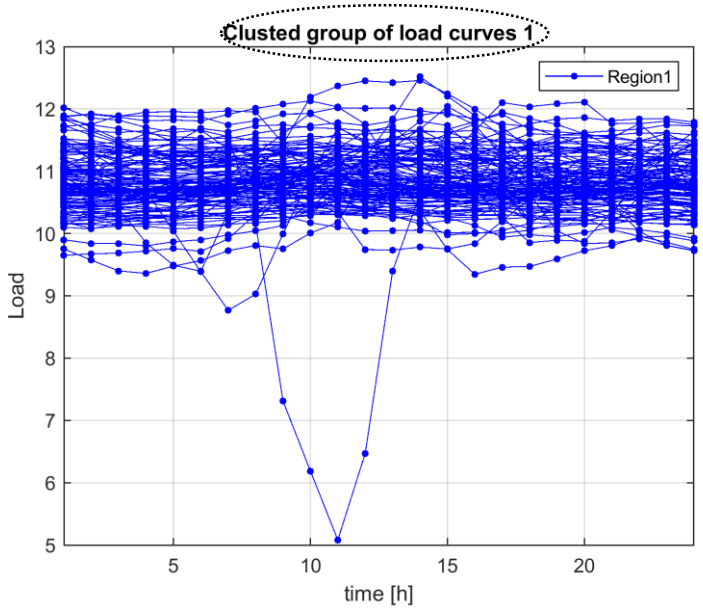


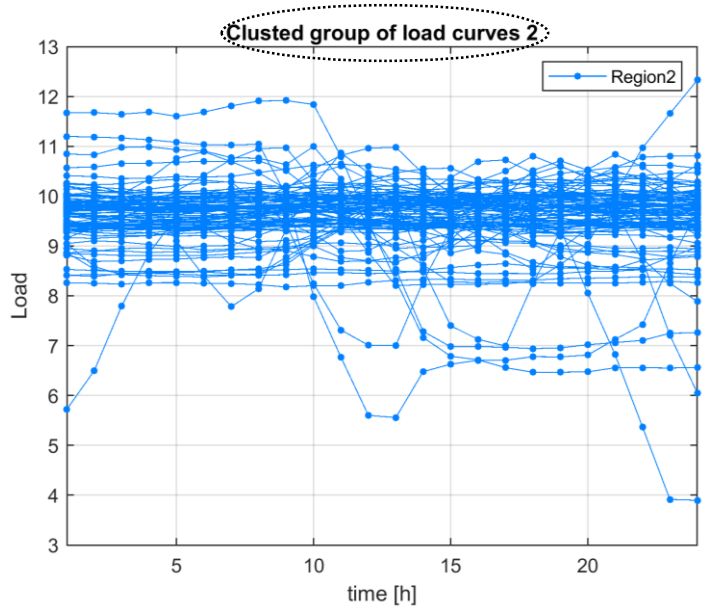
Cityblock distance

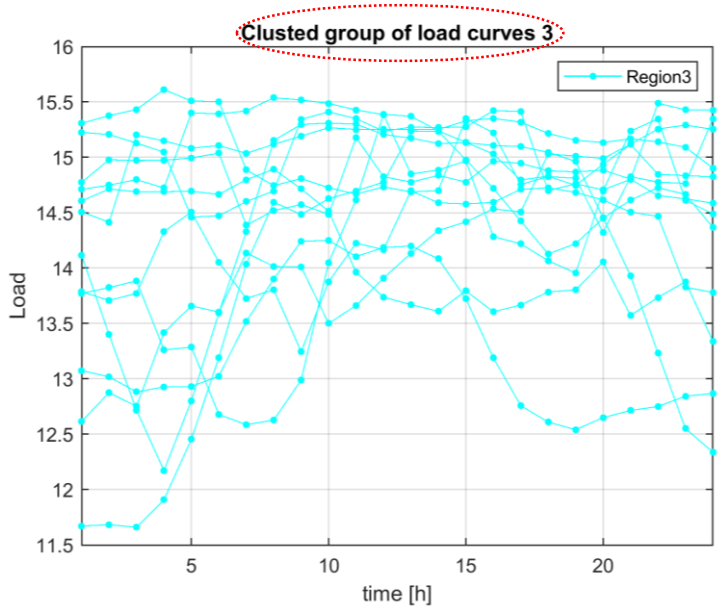


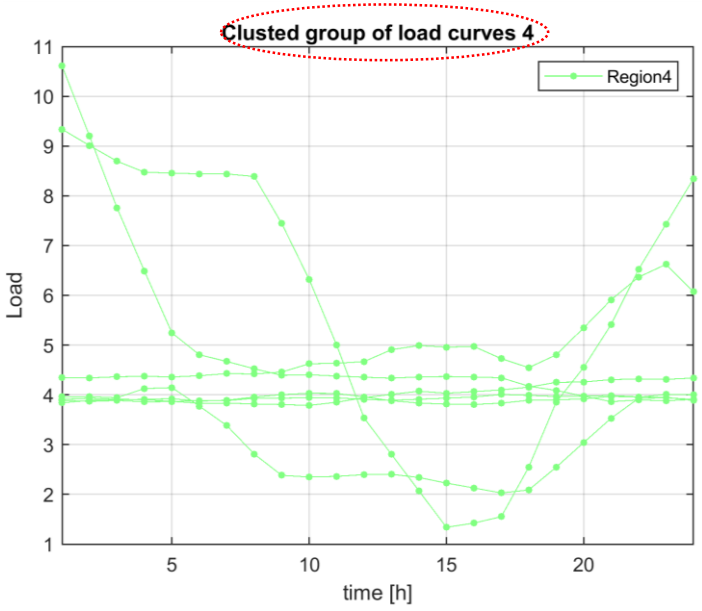


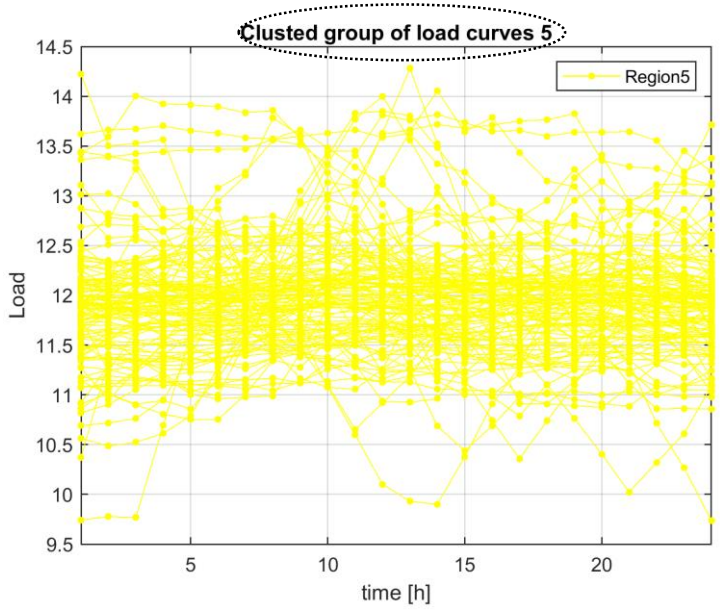
K-means Clustering

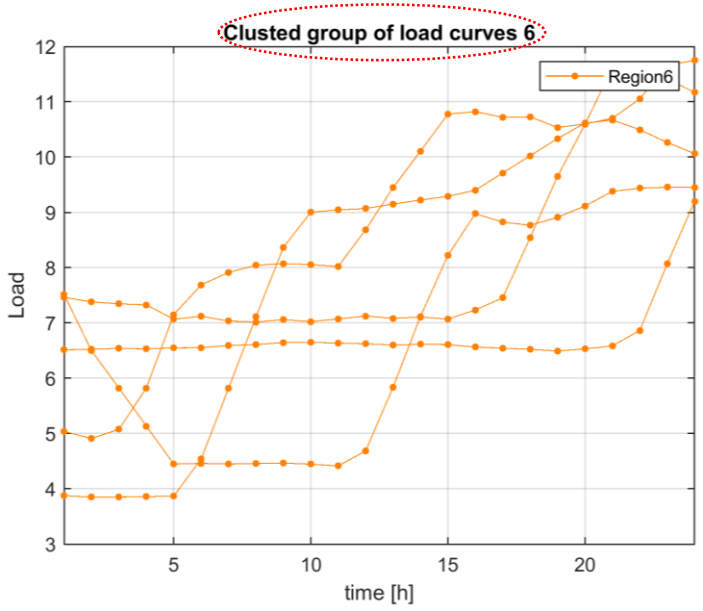






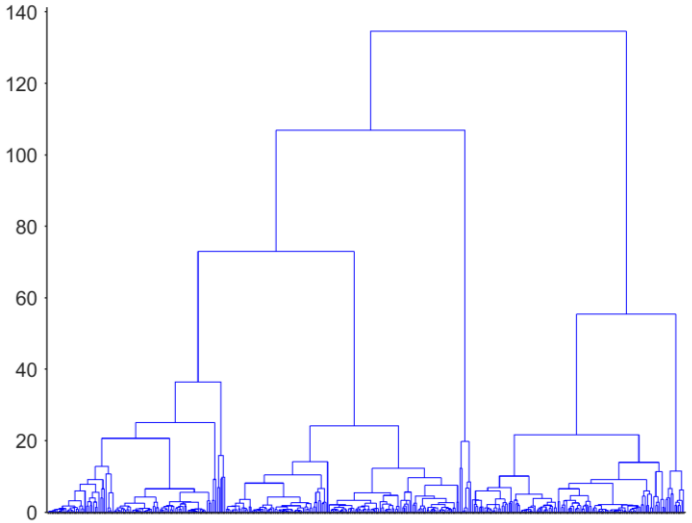






Hierarchical Clustering (*Agglomerative*)

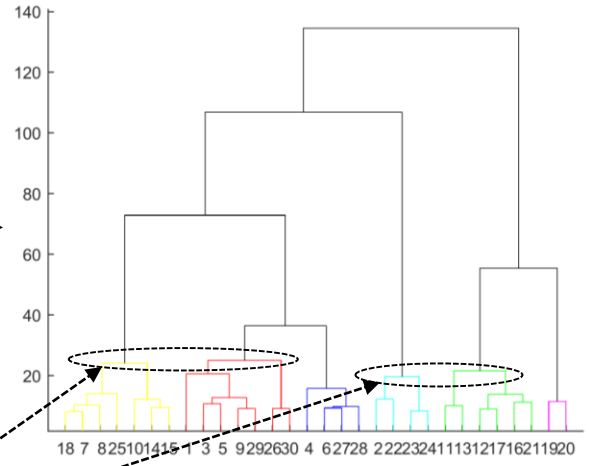
- **Data**: Platform's Daily Load Curve *(24 dimensional vectors)*
- **Linkage** : Ward's method *(how much the sum of squares will increase when we merge 2 selected clusters)*
- **Distance**: Euclidean

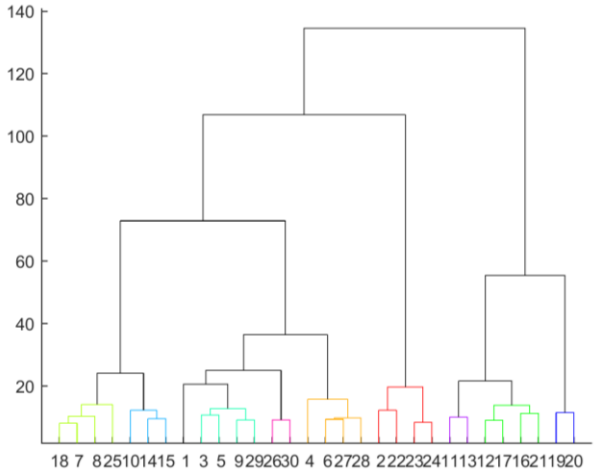


All-data dendrogram

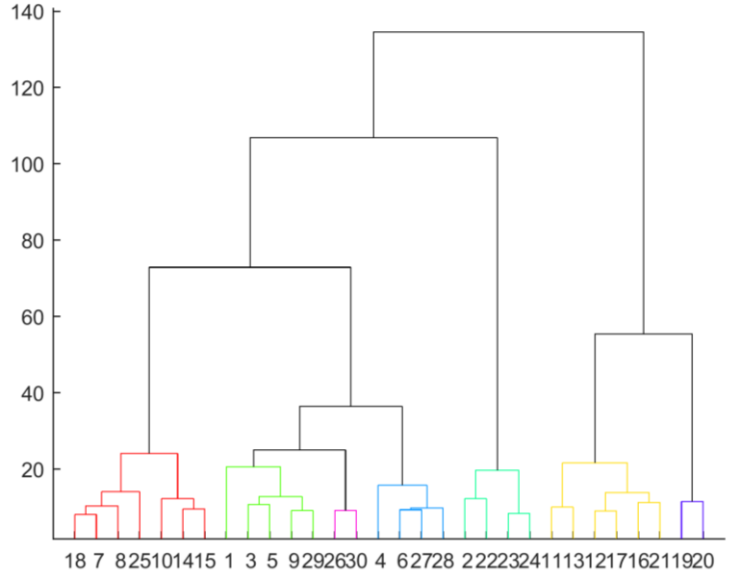
Linkage height corresponds to clusters distance

k=6 *(from previous analysis)* → similar clustering results





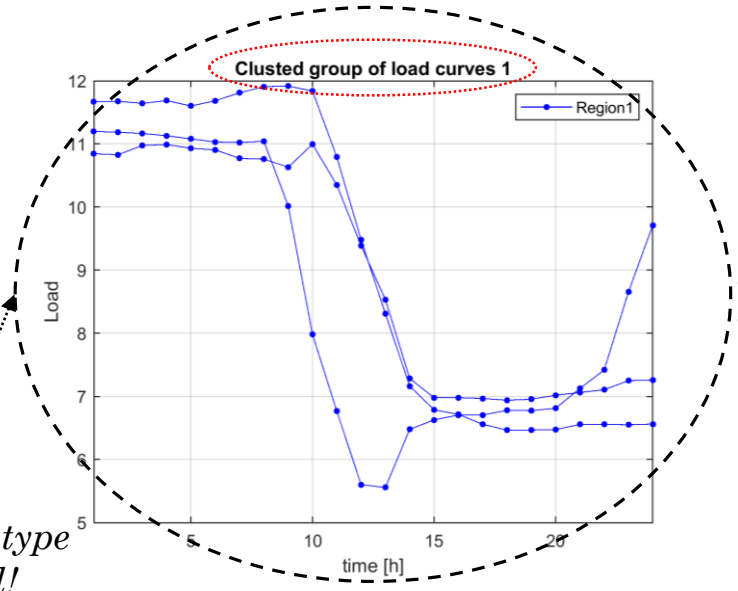
Clusters 1-2 and 4-5 have relatively low distance among them



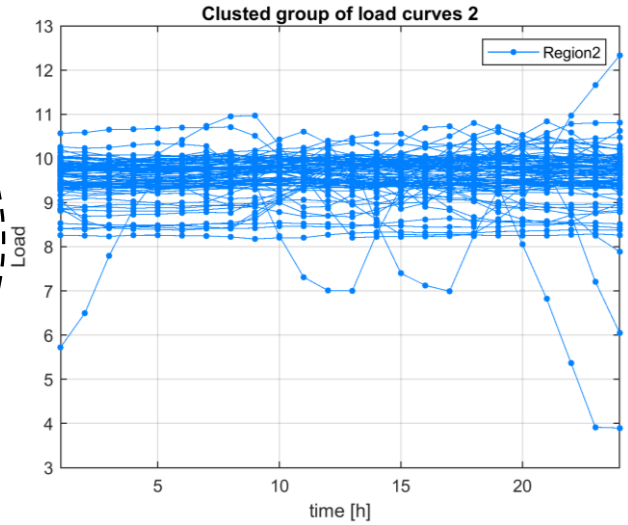
k=9 (many clusters with similar height)

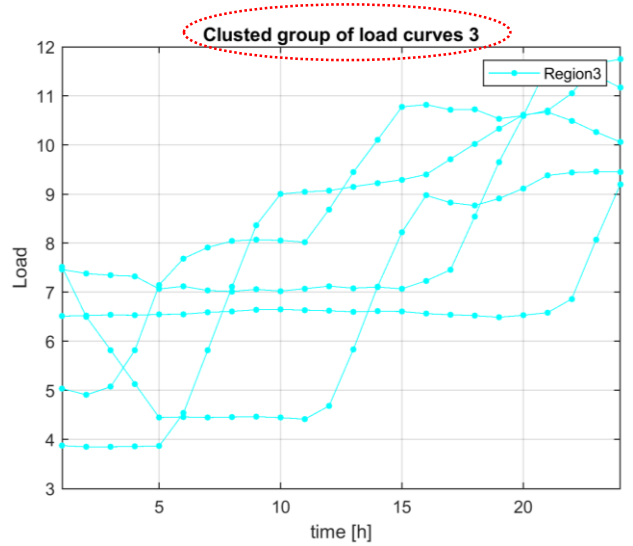
k=7 (best)

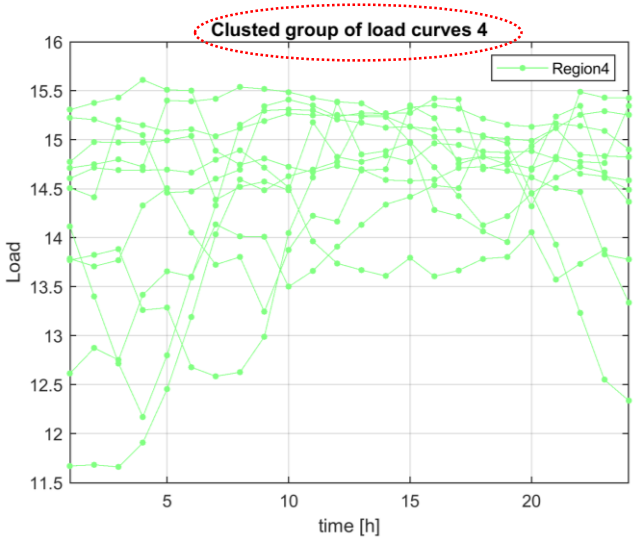
Hierarchical Clustering (*Agglomerative*)

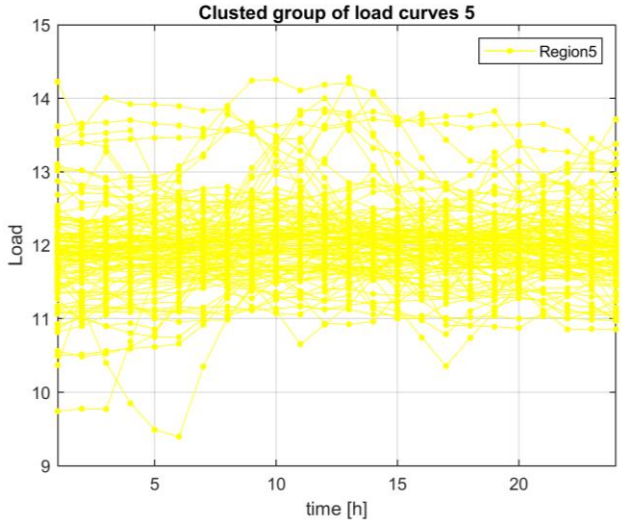


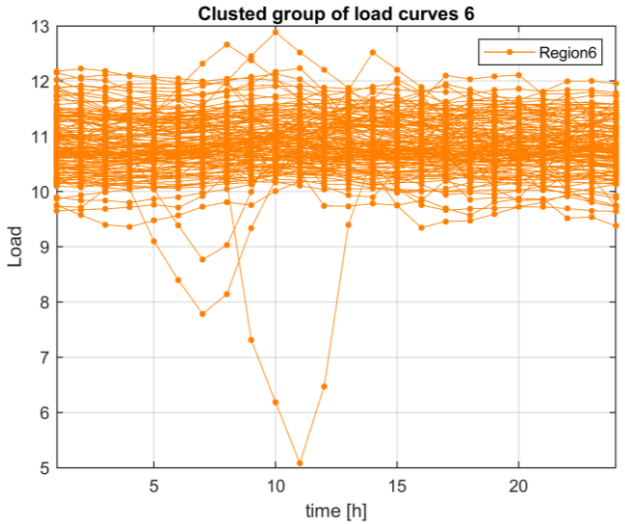
New cluster type identified!

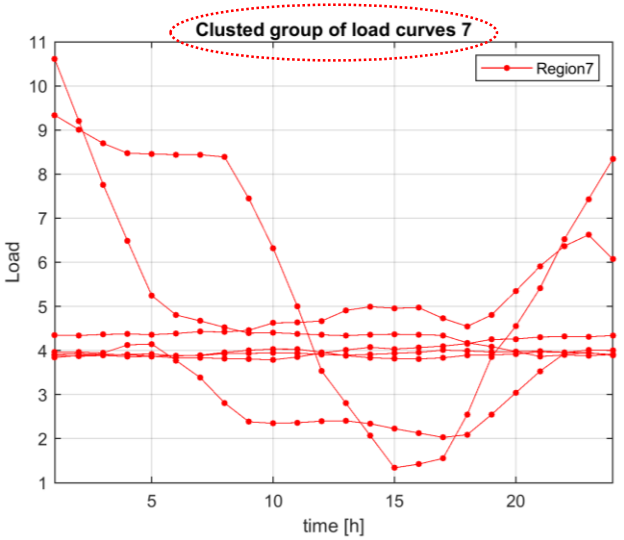












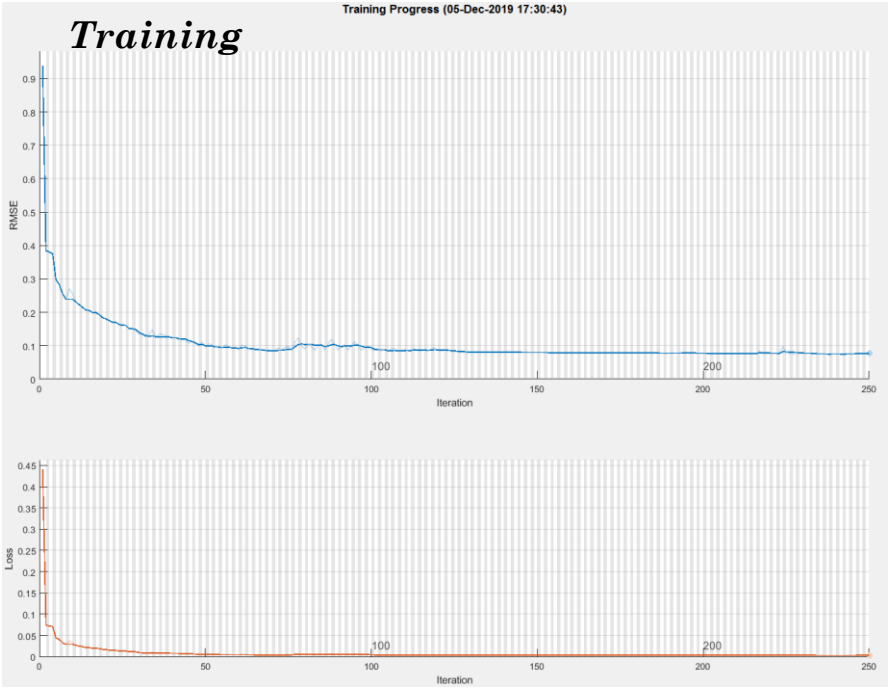
6. Deep Learning application

LSTM Load Forecasting

- **Data**: Platform's annual Load Timeseries *(8760 steps)*
- **Train** : 90% of data, 250 epochs, *adam (adaptive moment estimation) SGD*, minibatch=128, $a_0 = 0.005$
- **Architecture**: 1 Input Layer – 1 LSTM layer (200 hidden blocks) – 1 Output Layer (*fully connected*)

RNN → Vanishing gradients ☹
LSTM → Model short and long dependencies ☺

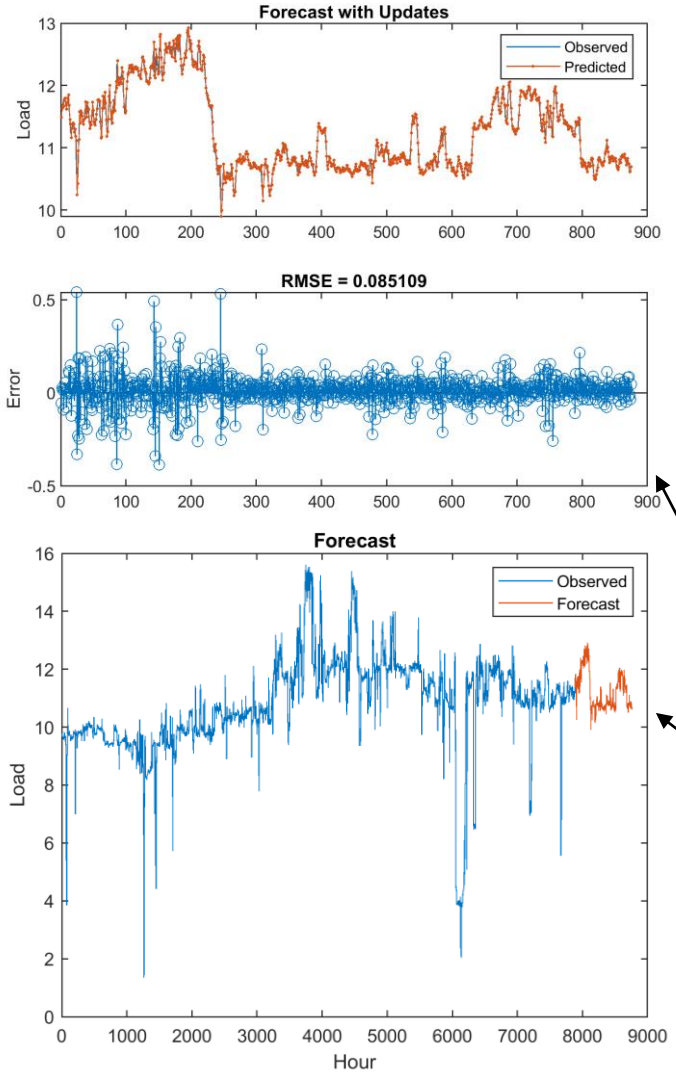
Training



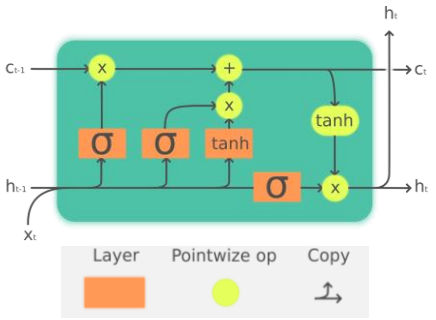
Responses (across all observations)

$$Loss: mse = \frac{1}{2N}\sum_{i=1}^{M}(X_i - T_i)^2$$

Observations — *Network responses* — *Target values*



Hidden block



1. Input gate
2. Forget gate
3. Cell update
4. Output gate

$\sigma(x) = (1 + e^{-x})^{-1}$

Gating: acts like a selector
(0 → discard, 1→keep)

tanh: scaling to [-1,1]
(*values don't explode*)

Very good performance for 1-step ahead predictions !

Possibility to efficiently model Wind Power !

Thank you for your attention

Questions