Pollution Modeling for the Energy Industry

(Report of a work in progress)

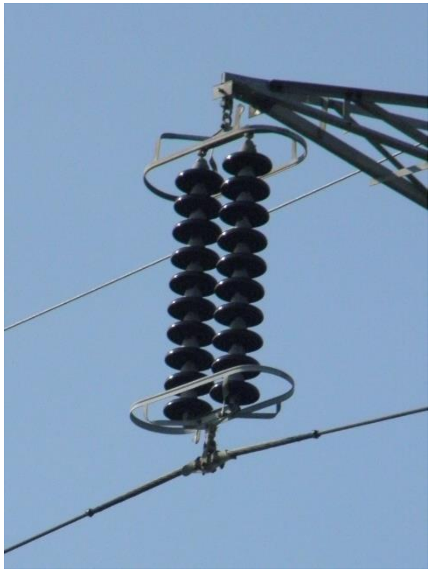
<u>Marcel Kvassay</u>, *P. Krammer, R. Forgáč, M. Očkay, J. Mojžiš*

<u>National project APVV-20-0548 ARIEN (2021-2023)</u>

Institute of Informatics, Slovak Academy of Sciences
VUJE, a.s. (https://www.vuje.sk/home)

27. 11. 2023

Problem: insulator pollution potentially causing discharge and short-circuiting (e.g. in wet conditions)



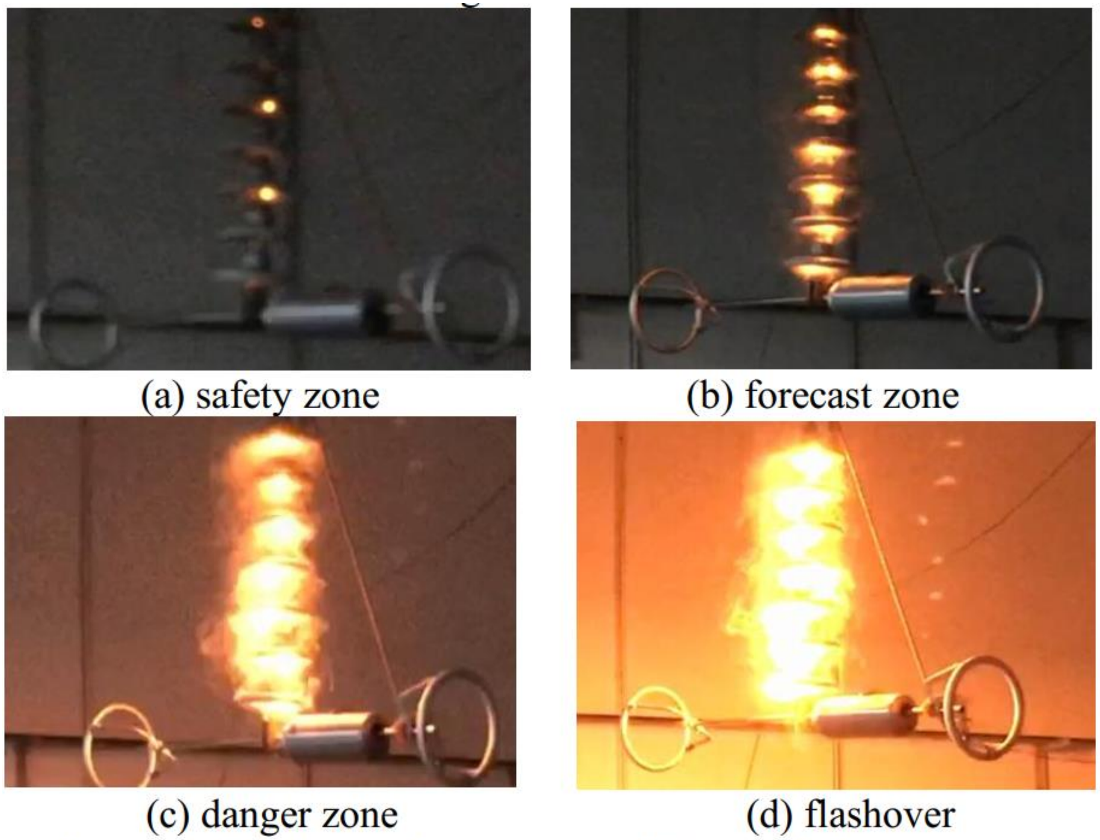


Figure 3. Discharge phenomena at different discharge stages

Source: Zeng et al. (2020). Experimental Study on High Frequency Pulse Current Variation Characteristics of Pollution Discharge of Insulators. In IOP Conference Series: Earth and Environmental Science (Vol. 446, No. 4, p. 042006). IOP Publishing.

Sample of field-measured data (VUJE, STN 33 0405)

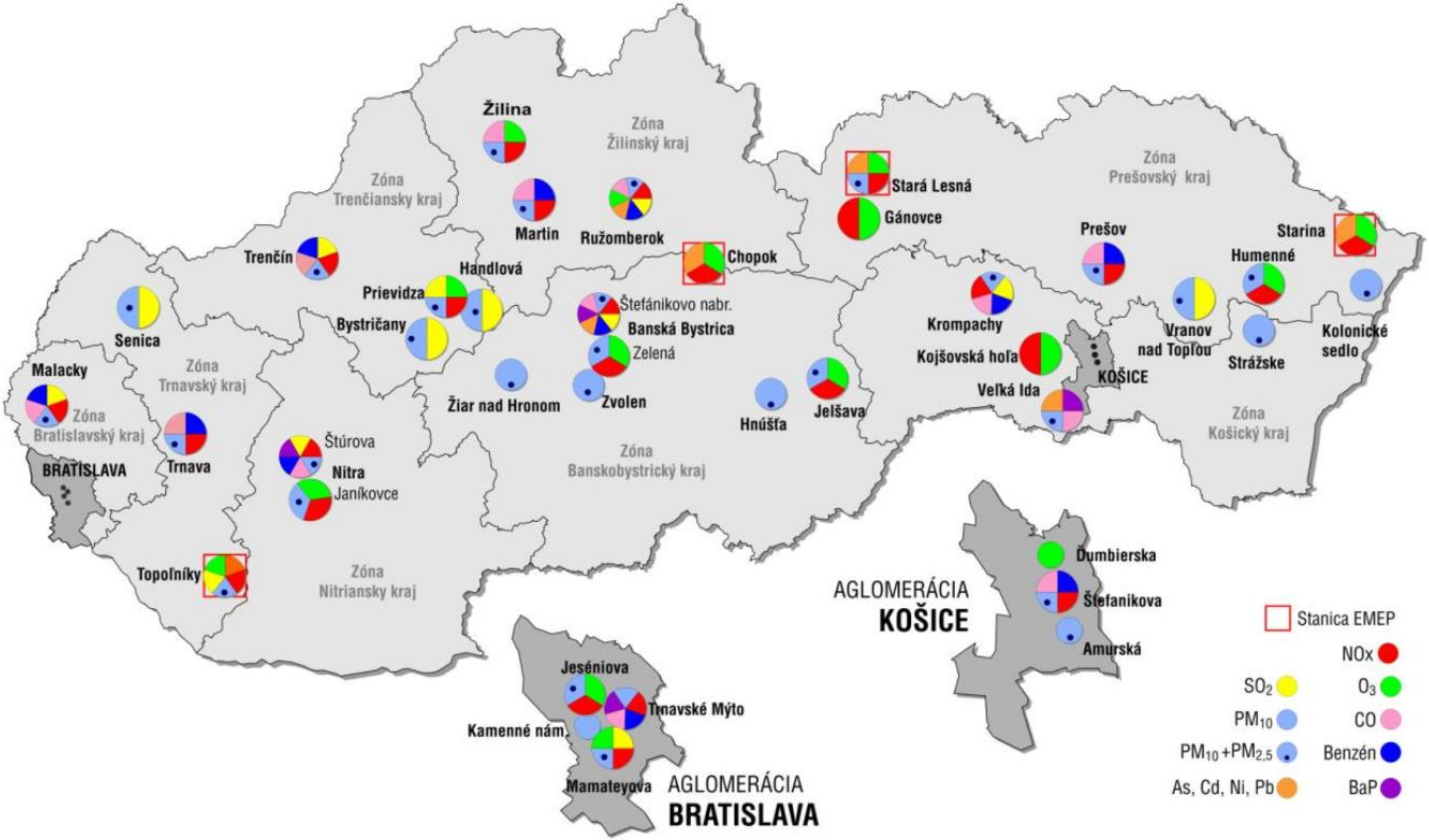
Cumbersome field measurements according to Technical standard STN 33 0405

New Idea:
Replace them with Artificial Intelligence

Target:
Upper bounds of S, Sr, g0.2

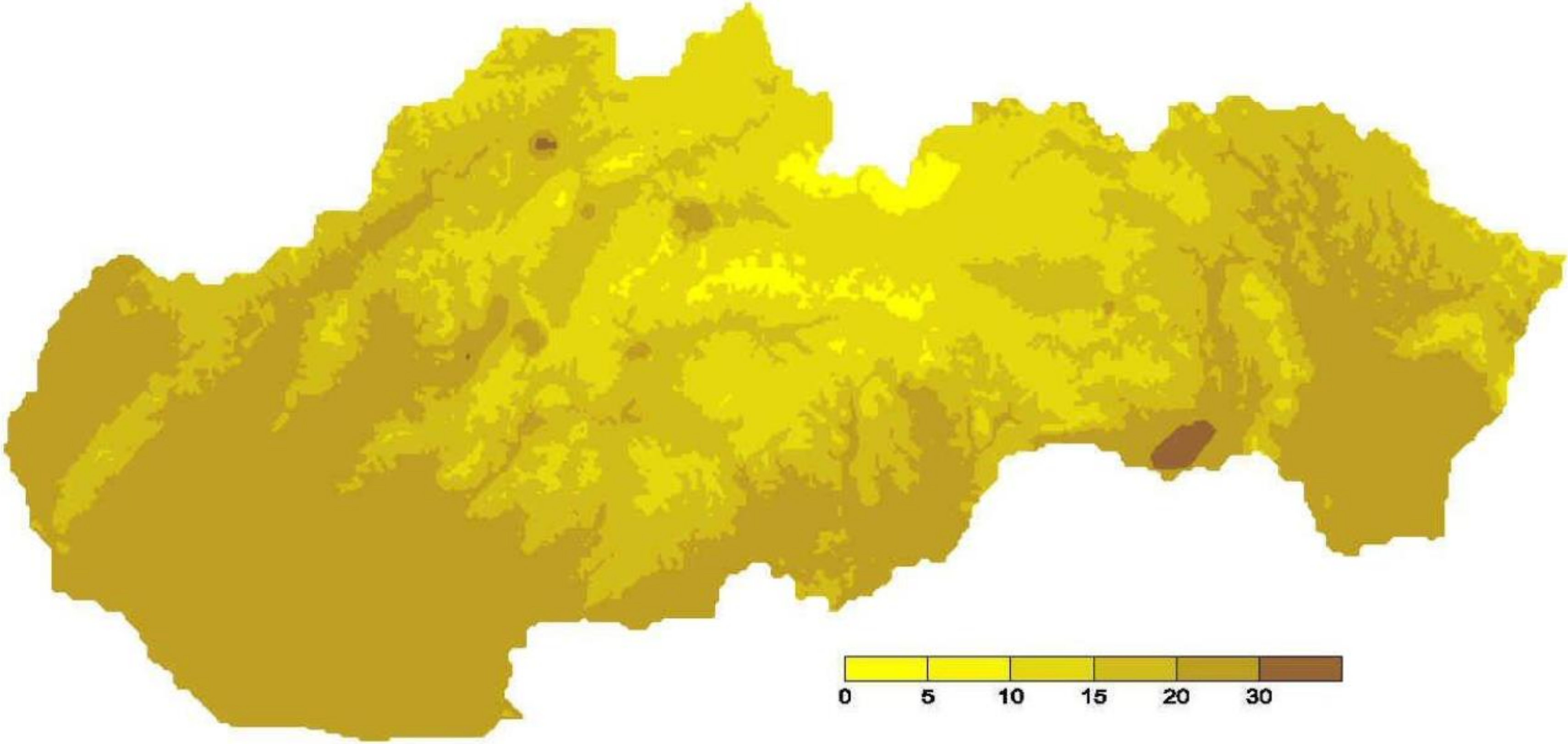
	A	E	F	G	H	I	O	P	Q	W	X	Y	AE	AF	AG	AH	AI
1				Namerané hodnoty									Ročné výsledky				
2				S: Total deposit (mg/day*cm²)			Sr: Soluble fraction			g0.2: Conductivity			Upper bounds Q99.5% S	Sr	g0.2	p (súčin)	Stupeň znečistenia
3		Doba merania		Číslo zberu			Číslo zberu			Číslo zberu							
4	Stanice	od	do	1	2	3	1	2	3	1	2	3					
5	Podunajské biskupice	13/05/2008	14/04/2009	0.00532007	0.01402077	0.01094552	0.00182158	0.00391103	0.00392174	1256.2	1411.6	1254.5	0.021485180	0.005984409	2192.68	0.28	I.
6	Senica 220 kV	13/05/2008	14/04/2009	0.00717915	0.01112768	0.0138279	0.00179479	0.00199302	0.00126439	1277.7	953.3	1069.6	0.024007898	0.002585061	2135.60	0.13	I.
7	Bošaca 400 kV	13/05/2008	14/04/2009	0.00589869	0.01066157	0.00720594	0.00214839	0.00288773	0.0012001	1475.9	1485.7	1211.6	0.017694859	0.003666455	2058.38	0.13	I.
8	Lemešany 400 kV	14/05/2008	15/04/2009	0.01019011	0.00708271	0.00557723	0.0024377	0.00124831	0.00051968	1402.1	1502.7	1261.6	0.013381296	0.004684182	2070.61	0.13	I. **)
9	Lemešany 220 kV	14/05/2008	15/04/2009	0.01070443	0.00688984	0.00468252	0.00252342	0.00156977	0.00087864	1333.9	1258.9	1178.4					
10	Rimavská Sobota 400 kV	14/05/2008	16/04/2009	0.00825066	0.00808458	0.00432356	0.00230911	0.00209481	0.00075542	1200.0	1220.5	1182.1	0.014782757	0.003673845	1949.54	0.11	I.
11	Levice 400 kV	15/05/2008	16/04/2009	0.00568439	0.00879714	0.00286094	0.00158584	0.00171978	0.00056255	1350.7	971.1	1007.1	0.016208572	0.004855633	2240.28	0.18	I.
12	Veľký Ďur 400 kV	15/05/2008	16/04/2009	0.00336456	0.01136877	0.0034985	0.00064291	0.00302167	0.00061612	1126.2	1033.9	1193.8	0.016471624	0.004889243	2484.27	0.20	I.
13	Bystričany 220 kV	15/05/2008	17/04/2009	0.00967578	0.01278853	0.0066166	0.00298441	0.00412533	0.00197159	1416.6	1297.3	1431.3	0.017324492	0.005810816	1992.46	0.20	I.
14	Križovany 400 kV	15/05/2008	17/04/2009	0.01148737	0.00796135	0.00451643	0.00342946	0.00235733	0.00092686	1310.7	1262.5	1311.6	0.016561182	0.004818413	2222.45	0.18	I. **)
15	Križovany 220 kV	15/05/2008	17/04/2009	0.00960248	0.00675054	0.00379852	0.0028797	0.00243234	0.00081435	1237.5	987.5	1078.6					
16	V427	05/05/2011	05/04/2012	0.0097	0.0190	0.0442	0.0031	0.0046	0.0180	1200.89	1268.75	821.43	0.106559	0.037356	1470.006	5.85	III.
17	V427	05/05/2011	05/04/2012	0.0153	0.0455	0.0153	0.0028	0.0104	0.0045	1291.07	984.82	929.46	0.113406	0.02115	1915.027	4.59	II.
18	V496	12/05/2010	20/04/2011	0.0177	0.0249	0.0359	0.0073	0.0081	0.0068	1041.1	1258.9	1192	0.078765523	0.054579695	2551.631659	10.97	IV.
19	Poznámky:																
20	**) - hodnoty boli vypočítane pre rozvodňu ako celok																

New idea: Rely on Data from the Network of Slovak air-pollution measuring stations (SHMÚ)



Numerical algorithms simulate spread of pollution over Slovakia (we started with yearly averages)

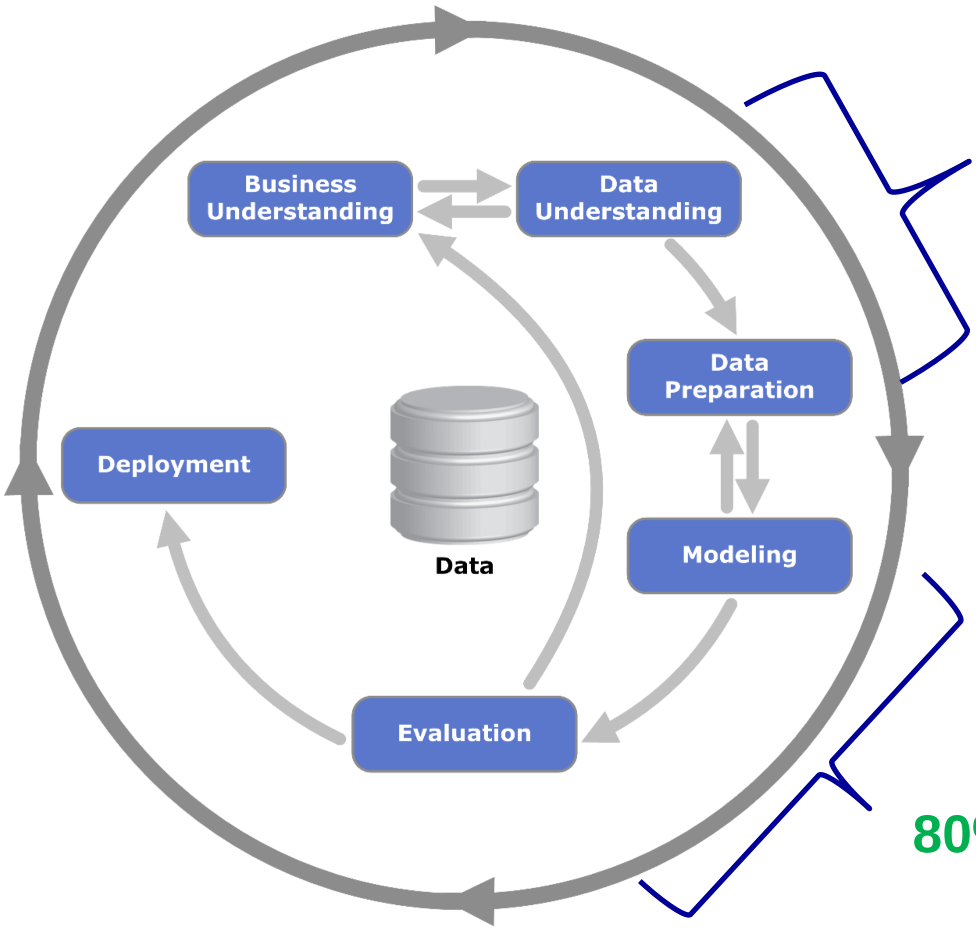
Figure: Local yearly averages of $PM_{10}\ [\mu g.m^{-3}]$, 2016.



Data from VUJE+SHMÚ for our feasibility study

	Input attributes							Output (target) attributes		
meranie	lat	lon	PM10	PM2.5	NO2	SO2	O3	S	Sr	g0.2
1	48.11267	17.21781	21.00	14.70	7.40	10.12	47.86	0.02150	0.00600	2192.70
2	48.30581	17.03103	21.06	14.74	6.90	4.61	57.75	0.02440	0.00400	2310.90
3	48.66647	17.39286	21.82	15.27	6.30	4.09	60.56	0.02400	0.00260	2135.60
4	48.79436	17.83872	21.00	14.70	6.30	4.27	48.70	0.01770	0.00370	2058.40
5	49.11592	18.42628	17.68	15.03	7.00	4.02	45.47	0.01960	0.00440	2002.50
6	49.21431	18.86353	29.33	24.93	5.40	3.65	48.50	0.01100	0.00610	1697.20
7	49.09164	18.97614	29.50	25.08	5.60	4.00	48.88	0.02130	0.00530	2598.90
8	49.08581	19.49019	25.18	21.40	6.60	4.41	61.13	0.01310	0.00340	2251.80
9	48.79694	19.36706	21.95	18.65	5.60	3.60	44.86	0.01290	0.00440	2170.30
10	48.96456	20.53486	17.68	15.03	6.70	4.46	61.12	0.01590	0.00500	1924.20
11	48.85797	21.25386	27.45	19.21	6.90	4.66	55.35	0.01340	0.00470	2070.60
12	48.56881	22.04022	21.00	14.70	7.20	4.68	51.77	0.01660	0.00420	2459.40
13	48.85622	21.84814	24.70	17.29	7.50	4.57	53.96	0.02750	0.00440	2264.30
14	48.60181	21.04483	31.73	21.97	6.60	4.39	43.46	0.01850	0.00820	1652.90
15	48.38006	20.04481	22.55	15.78	6.80	4.22	47.70	0.01480	0.00370	1949.50
16	48.23336	18.60417	21.00	14.70	8.00	4.50	56.36	0.01620	0.00490	2240.30
17	48.20564	18.42531	21.00	14.70	6.10	4.43	59.00	0.01650	0.00490	2484.30
19	48.55900	18.74136	22.56	19.18	5.60	4.62	37.68	0.02000	0.00500	1990.00
20	48.66664	18.51206	28.54	19.01	6.70	19.40	48.96	0.01730	0.00580	1992.50
21	48.33717	17.65869	21.42	15.00	7.00	4.68	59.82	0.01660	0.00480	2222.50

CRISP-DM Proces (1) – our effort expectations

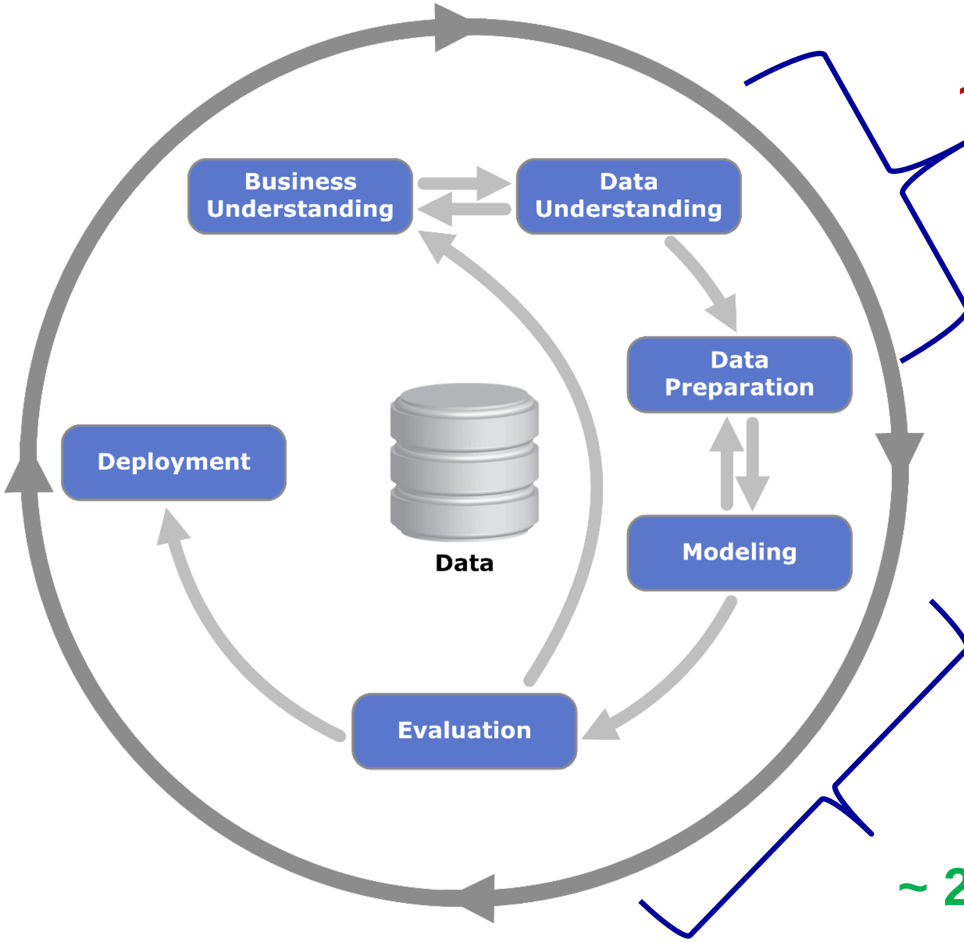


20% effort
(„preparation“)

80% effort („real work“)

- Figure: By Kenneth Jensen - Own work based CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=24930610

CRISP-DM Proces (2) – reality:



~ 80% effort and time:

Dates:

4/2020 – first data samples

... – data understanding

1.10.2020 – start feasibility study

... – exploring, checking, correcting and pre-processing data

8.3.2021 – final correct data supplied

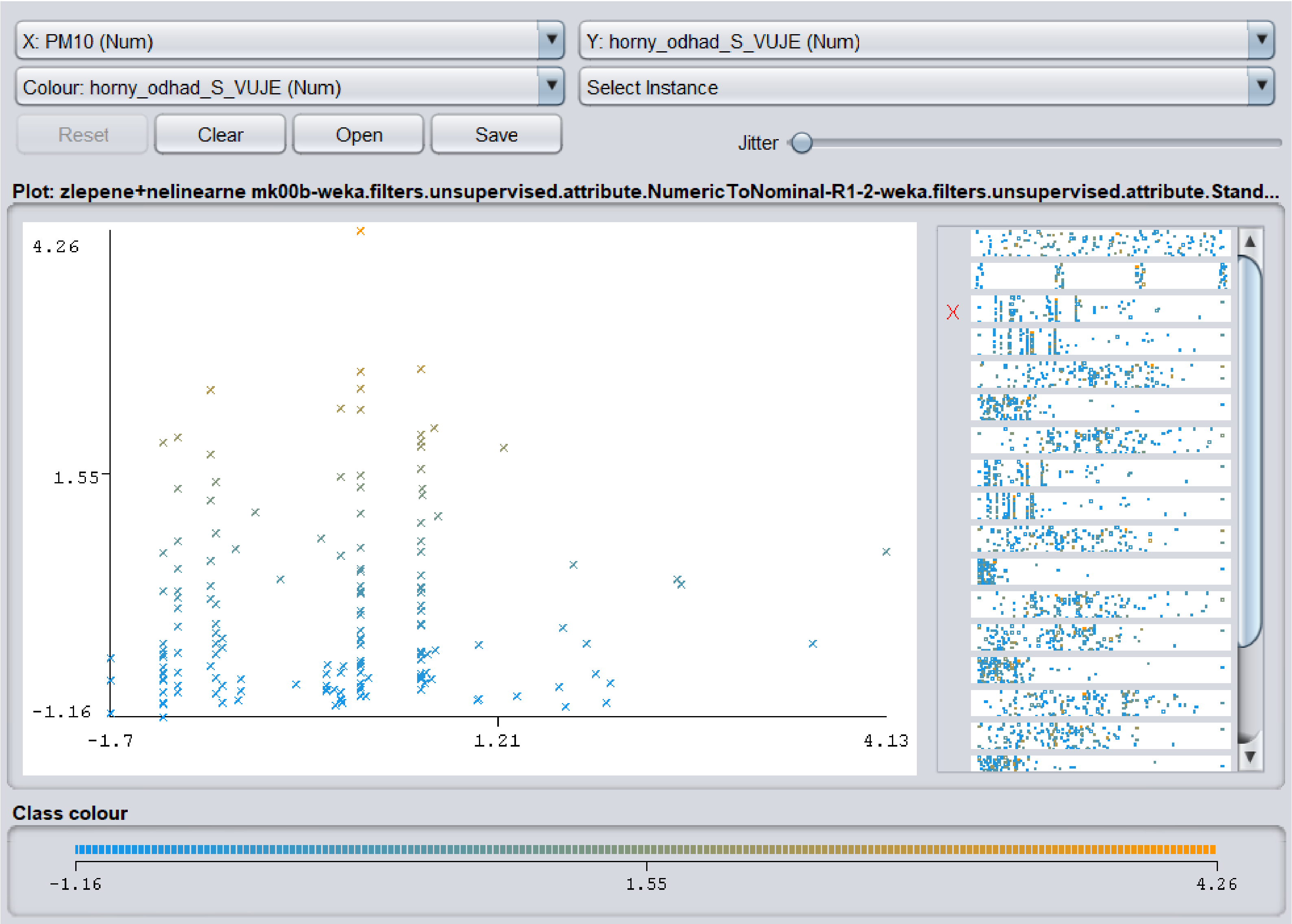
26.3.2021 – final presentation

31.3.2021 – end of feasibility study final report

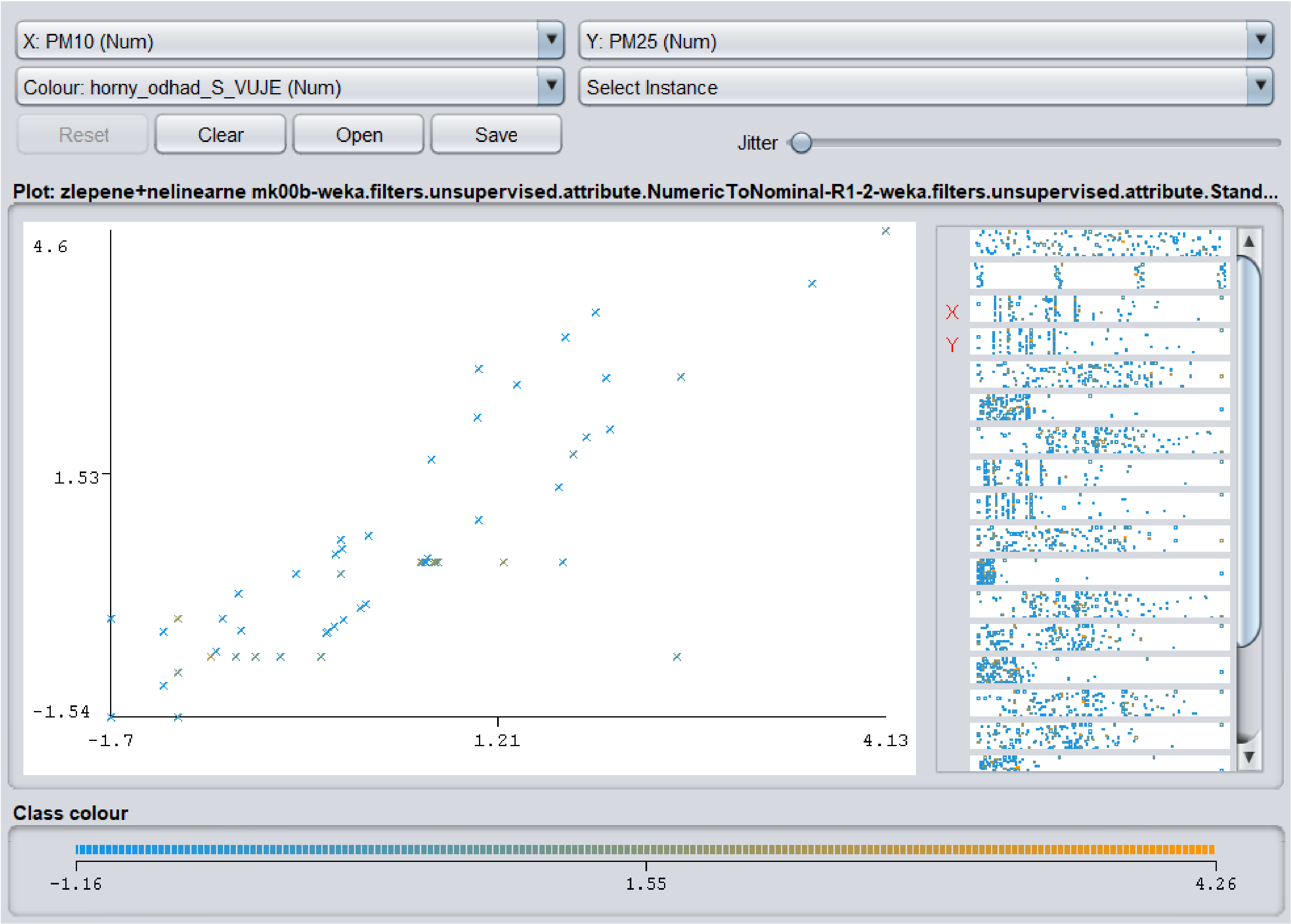
~ 20% effort and time

- Figure: By Kenneth Jensen - Own work based CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=24930610

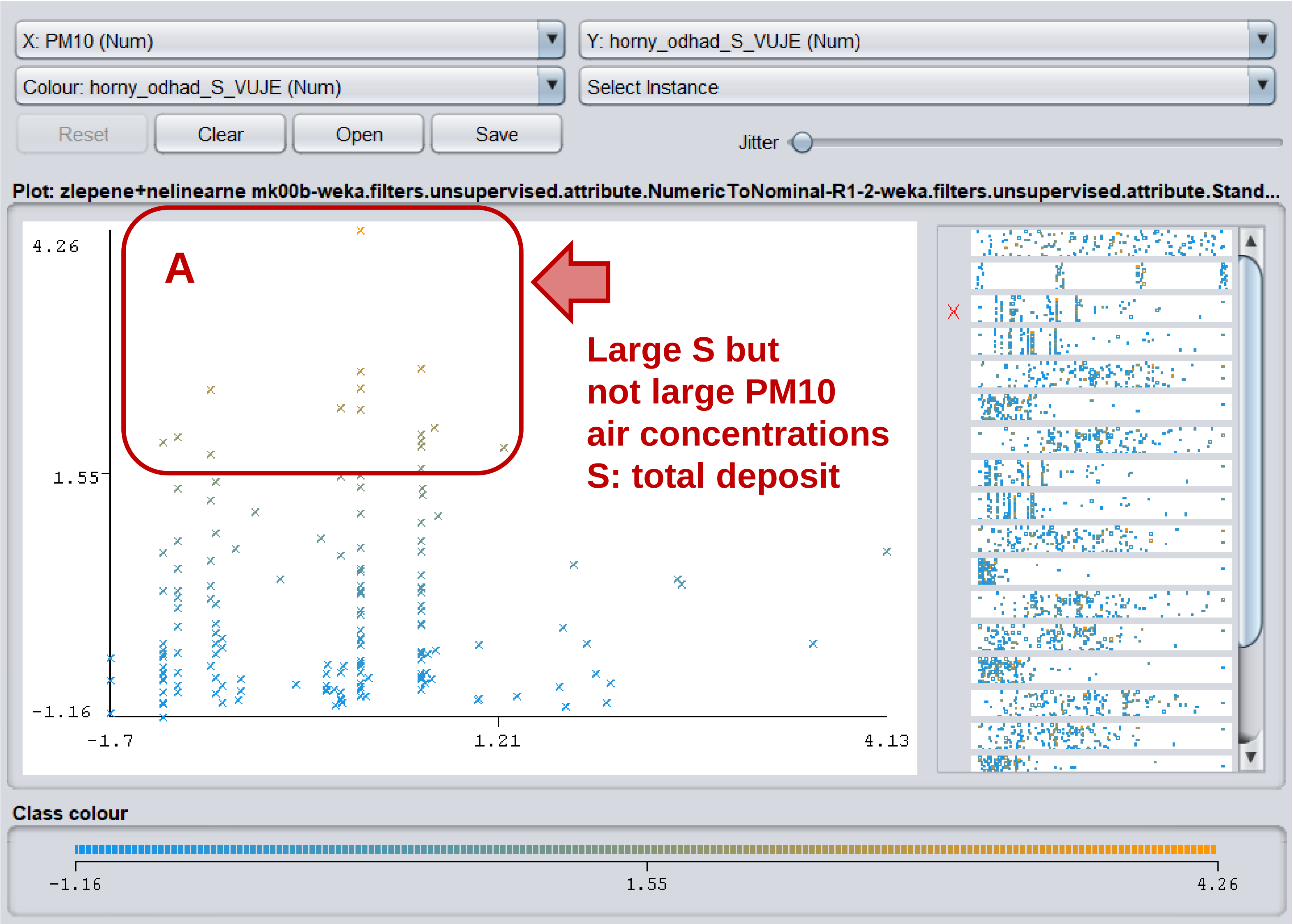
Problem: no correlation between PM10 and S_VUJE:



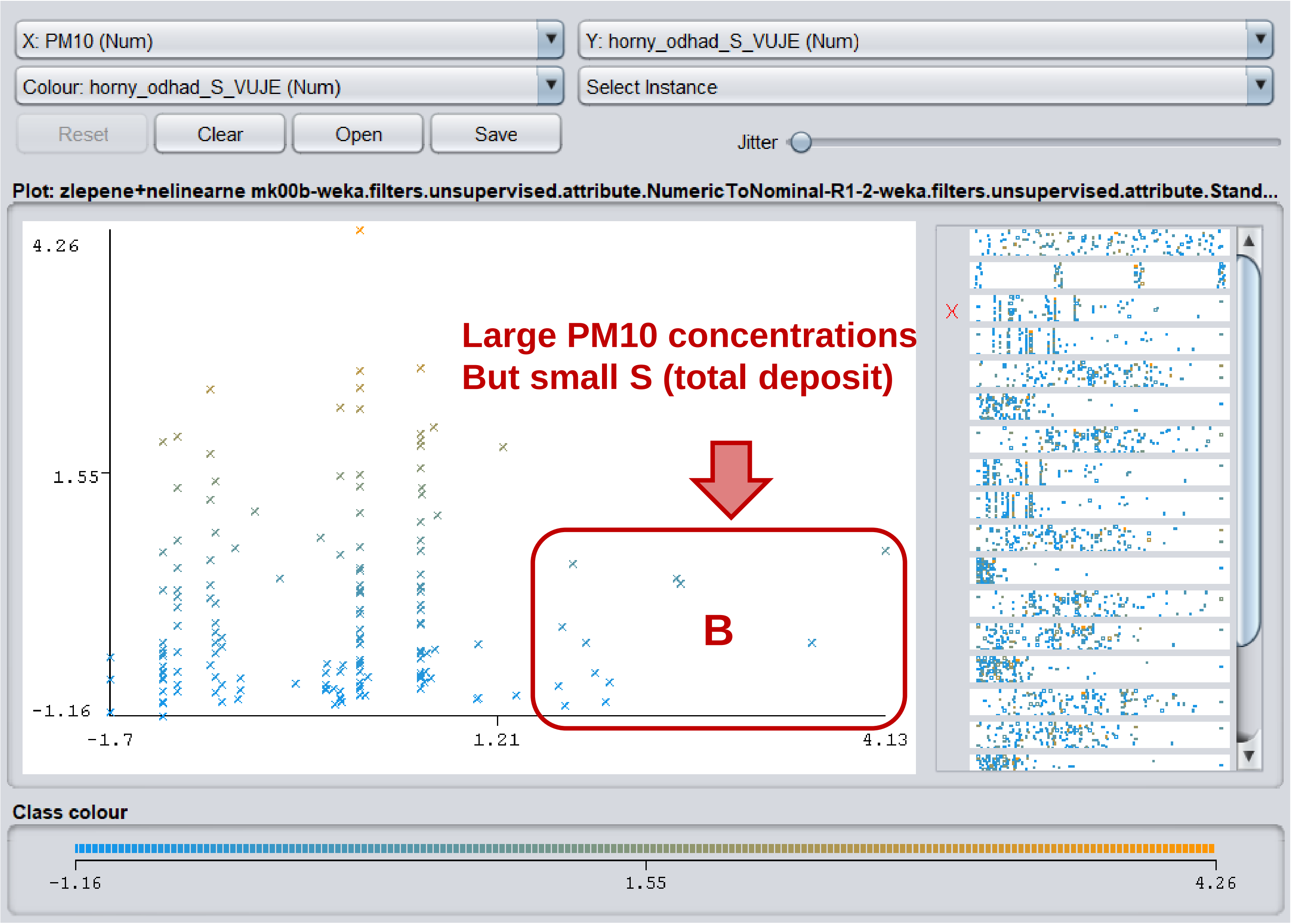
What we expected (Example):
Nice relation between PM10 vs. PM2.5:



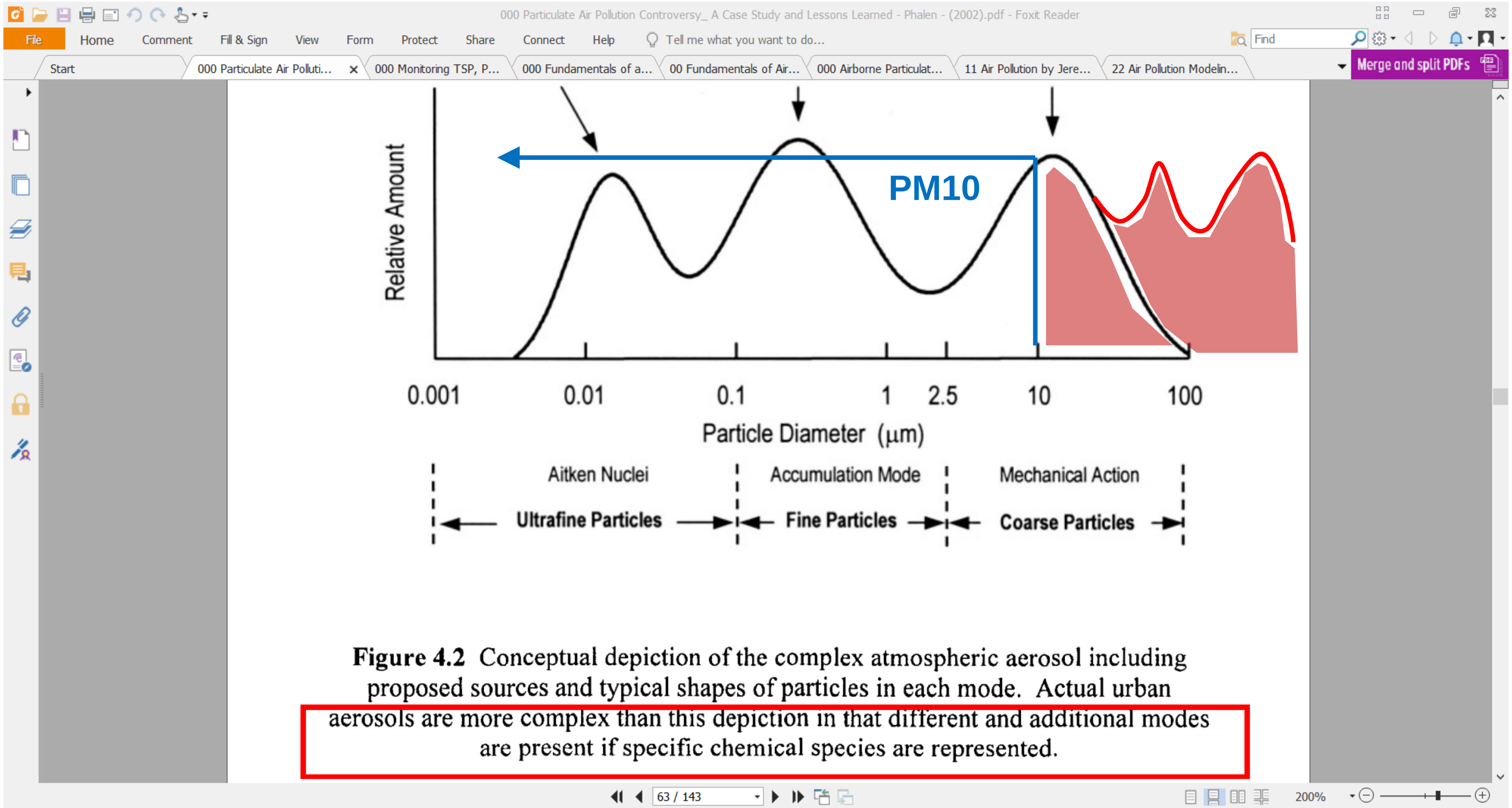
Problem: symptom A



Problem: symptom B



Search for possible causes of the problem (1):



Credit: Phalen, R.F., 2002. *The particulate air pollution controversy: A case study and lessons learned*. Springer Science & Business Media.

Search for possible causes of the problem (2)

Semi-remote location in northern Italy:

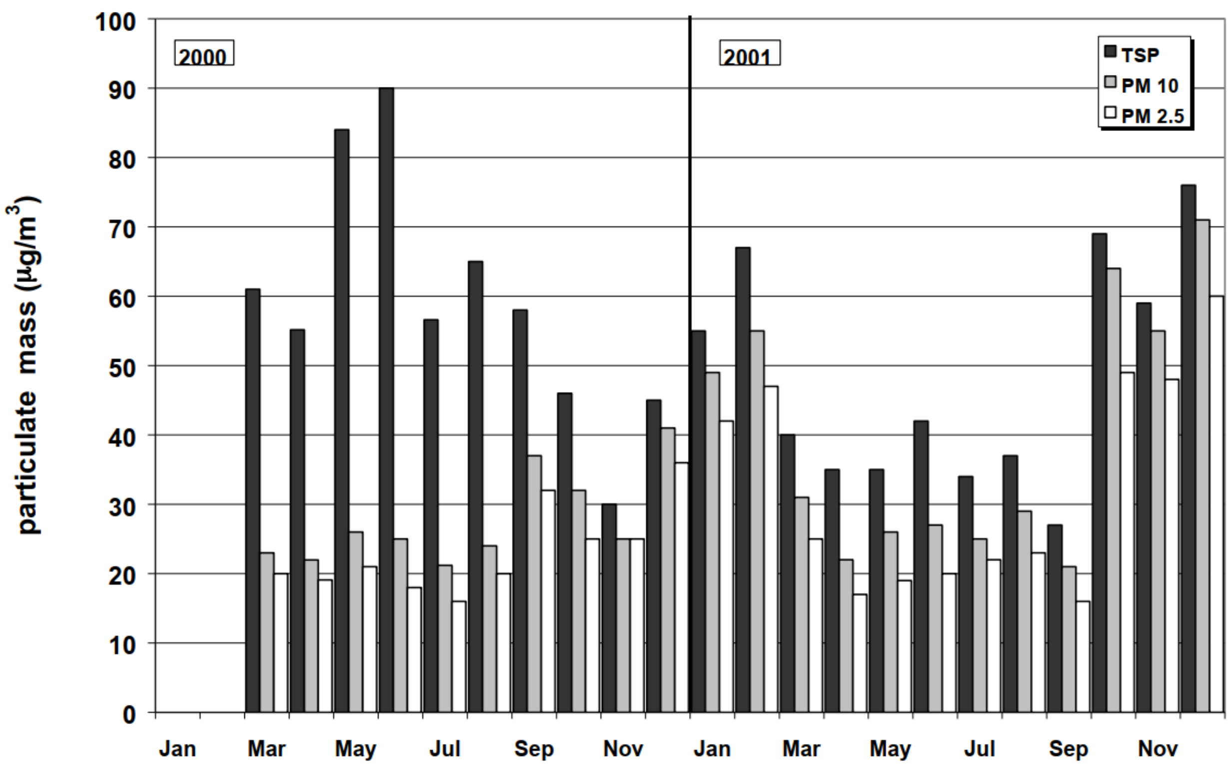
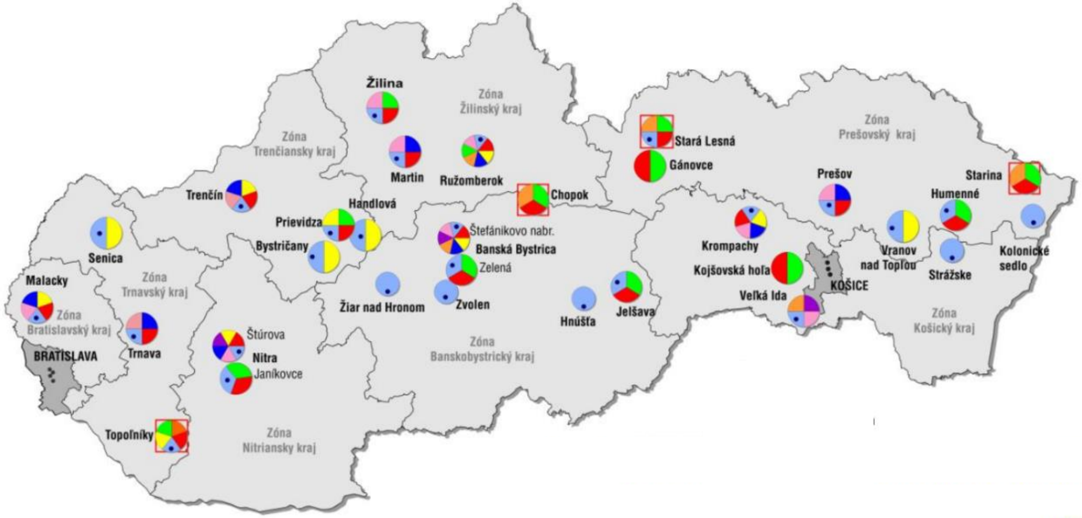


FIGURE 1 - **The monthly average concentrations of TSP, PM10 and PM2.5, calculated from 24h daily collected samples, show greater annual changes for TSP during the measurement period than for PM10 and PM 2.5.**

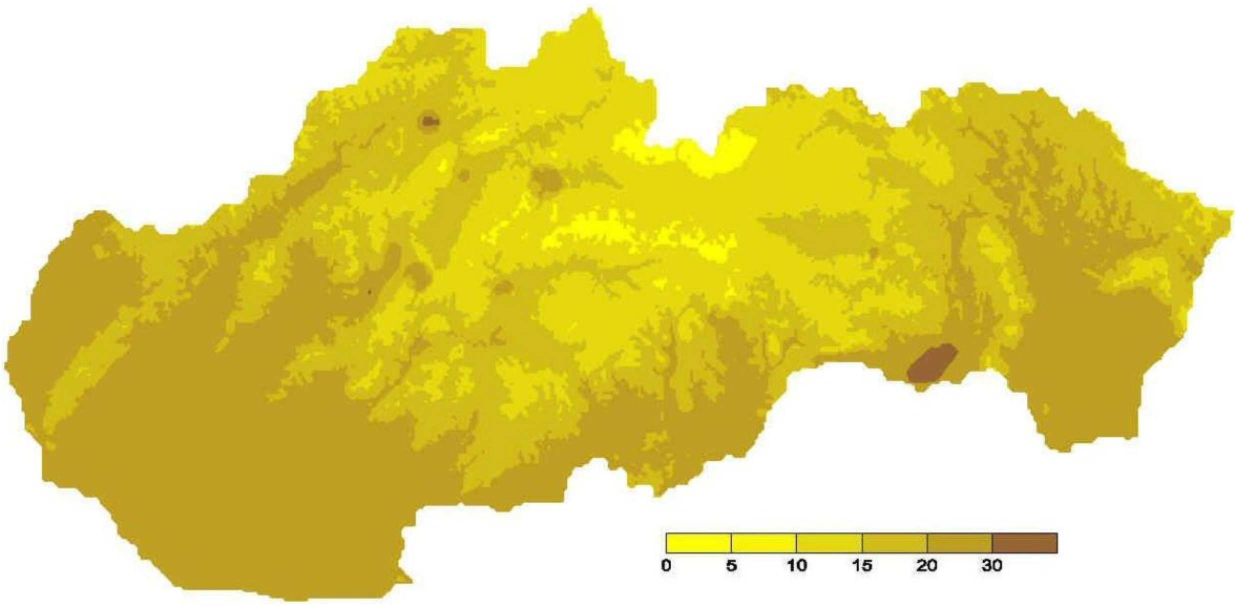
Credit: Rembges, D. and Kotzias, D., 2003. Monitoring TSP, PM 10 and PM 2.5 at a semi-remote area in northern Italy. *Fresenius Environmental Bulletin*, 12(5), pp.402-405.

Tentative Answer: Incompleteness of SHMÚ data: Smaller local temporary pollution sources missing



+ NEIS database of major polluters

For a location far away from known pollution sources, low PM10 values will be predicted, ignoring the possibility of intense local (though perhaps temporary) pollution sources of TSP



Tentative explanation of symptom A (1):
(after consultations with SHMÚ)

Local source of pollution (TSP) unknown to SHMÚ (NEIS)

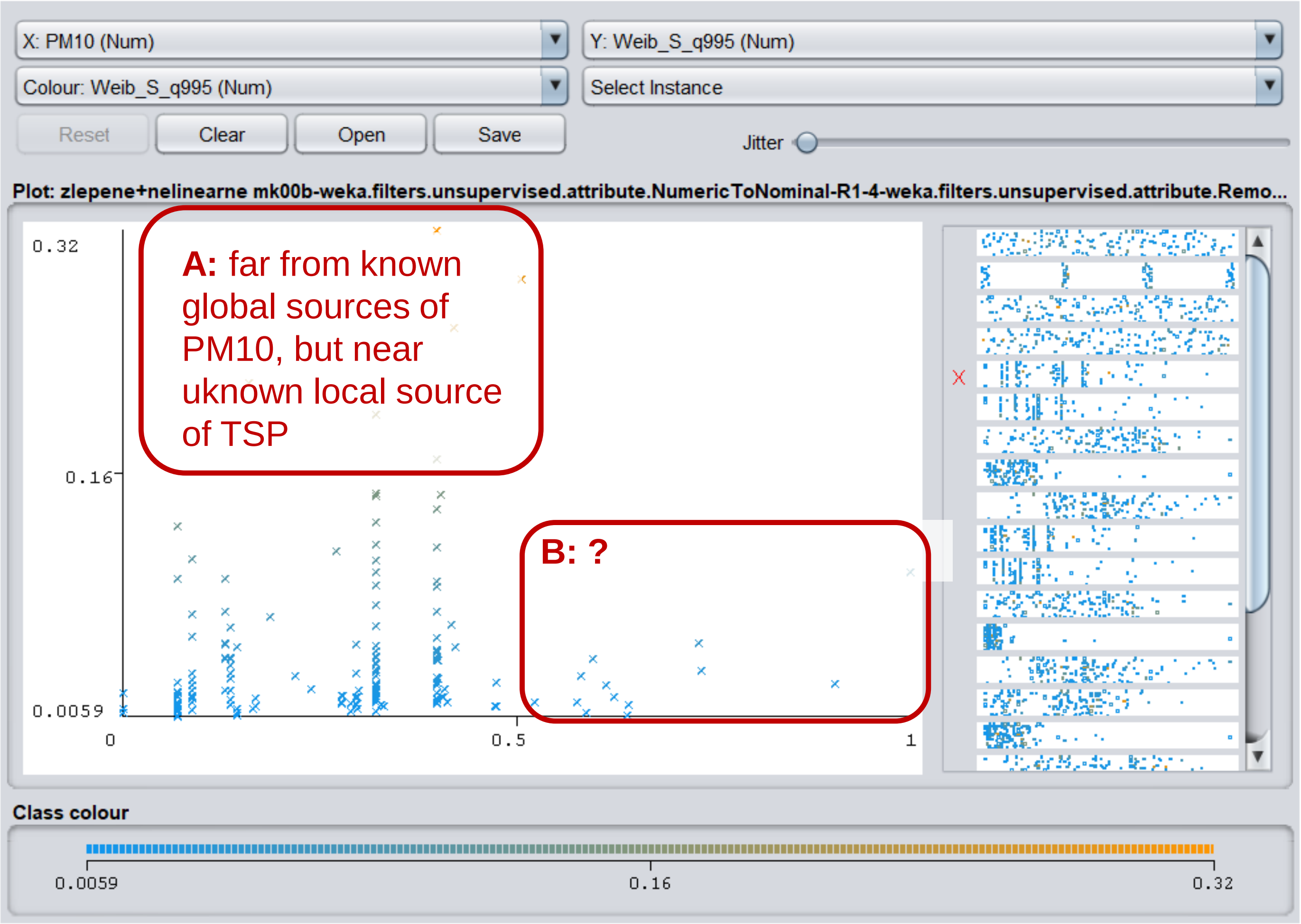
(The source may be temporary)

Consequence: Symptom A

⇒ low PM10 air concentrations
⇒ large total deposit S

Tentative explanation of symptom A (2):

(after consultations with SHMÚ)



Tentative explanation of symptom B (1):
(after consultations with SHMÚ)

Near a big global PM10 source known to SHMÚ (NEIS), but without any strong local source of TSP (e.g. US Steel in Košice-Šaca):

Underlying Hypothesis:

PM10 particles float like fog (negligible effect of gravitation):

Apart from rain + dew, virtually no PM10 deposits

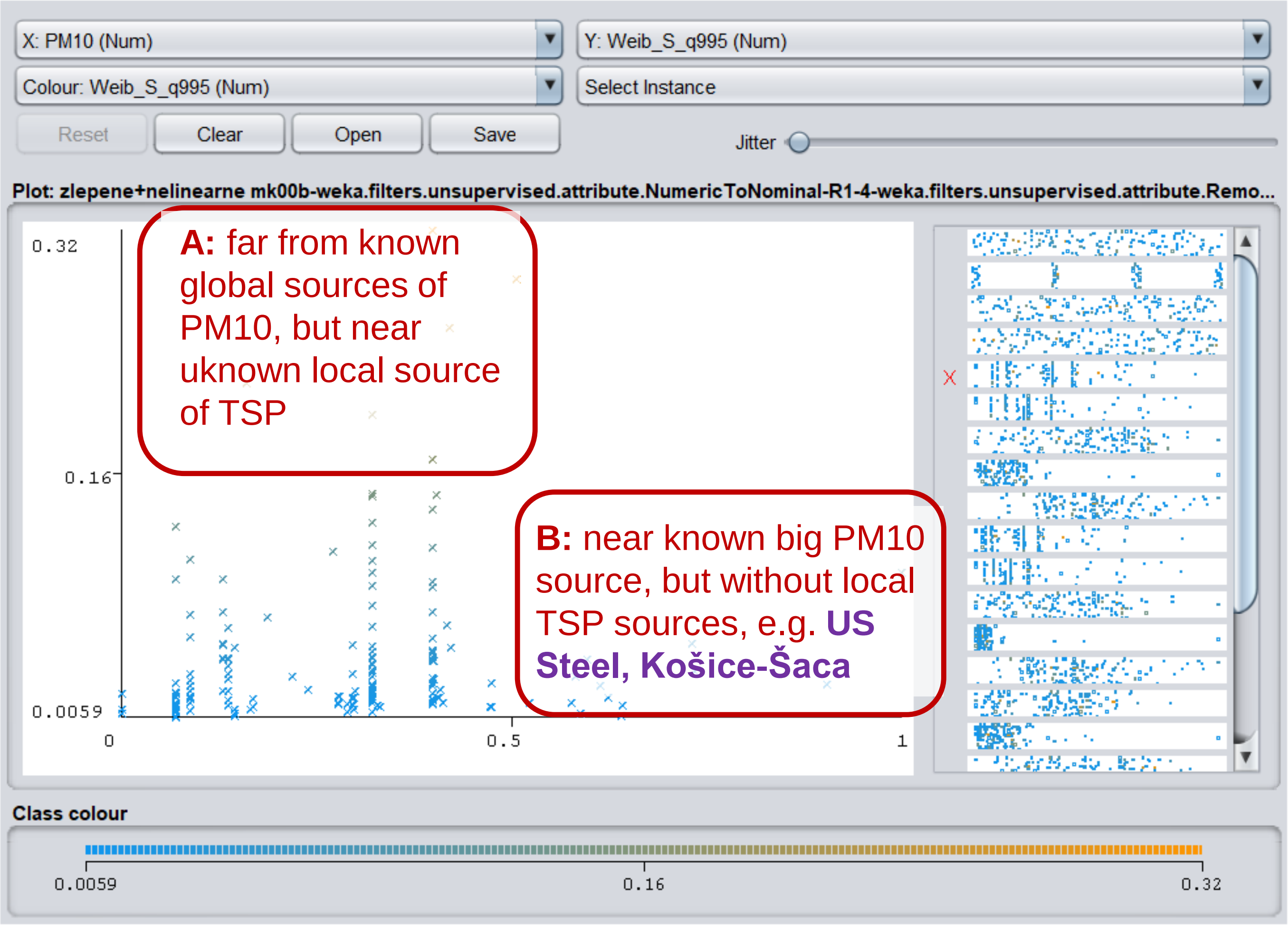
Consequence: Symptom B

⇒ high PM10 air concentrations

⇒ small total deposit S

Corollary:
total deposit S ≈ TSP - PM10

Tentative explanation of symptom B (2):
(after consultations with SHMÚ; see also [ref.1])



Conclusion: **PM10 and PM2.5 cannot reliably predict total deposit S,**
NEW TYPES OF INPUT ATTRIBUTES NEEDED

New Types of Data Considered
(most of them dynamic, hourly or weekly)

- *Types of **agricultural crops** cultivated across Slovakia*
- Database of ***major air polluters*** *in Slovakia (**NEIS**)*
- ***Satellite imagery*** *(Sentinel-2)*
- Hourly ***precipitation*** *totals across Slovakia*
- ***Hourly air concentrations*** *of major air pollutants:* ***PM10, PM2.5,*** NO_2, SO_2, O_3
- ***Hourly estimated wet deposition*** *of major wet pollutants*

- ***Thus far we tested the easiest to process ones;***
- ***In the process, we slightly re-defined our task:***

Changed target attributes: Field-measured 6-Wk values

(no longer their estimated upper bounds)

	A	E	F	G	H	I	O	P	Q	W	X	Y	AE	AF	AG	AH	AI
1							Namerané hodnoty							Ročné výsledky			
2				S: Total deposit (mg/day*cm²)			Sr: Soluble fraction			g0.2: Conductivity			Upper bounds Q99.5% S	Sr	g0.2	p (súčin)	Stupeň znečistenia
3		Doba merania		Číslo zberu			Číslo zberu			Číslo zberu							
4	Stanice	od	do	1	2	3	1	2	3	1	2	3					
5	Podunajské biskupice	13/05/2008	14/04/2009	0.00532007	0.01402077	0.01094552	0.00182158	0.00391103	0.00392174	1256.2	1411.6	1254.5	0.021485180	0.005984409	2192.68	0.28	I.
6	Senica 220 kV	13/05/2008	14/04/2009	0.00717915	0.01112768	0.0138279	0.00179479	0.00199302	0.00126439	1277.7	953.3	1069.6	0.024007898	0.002585061	2135.60	0.13	I.
7	Bošaca 400 kV	13/05/2008	14/04/2009	0.00589869	0.01066157	0.00720594	0.00214839	0.00288773	0.0012001	1475.9	1485.7	1211.6	0.017694859	0.003666455	2058.38	0.13	I.
8	Lemešany 400 kV	14/05/2008	15/04/2009	0.01019011	0.00708271	0.00557723	0.0024377	0.00124831	0.00051968	1402.1	1502.7	1261.6	0.013381296	0.004684182	2070.61	0.13	I. **)
9	Lemešany 220 kV	14/05/2008	15/04/2009	0.01070443	0.00688984	0.00468252	0.00252342	0.00156977	0.00087864	1333.9	1258.9	1178.4					
10	Rimavská Sobota 400 kV	14/05/2008	16/04/2009	0.00825066	0.00808458	0.00432356	0.00230911	0.00209481	0.00075542	1200.0	1220.5	1182.1	0.014782757	0.003673845	1949.54	0.11	I.
11	Levice 400 kV	15/05/2008	16/04/2009	0.00568439	0.00879714	0.00286094	0.00158584	0.00171978	0.00056255	1350.7	971.1	1007.1	0.016208572	0.004855633	2240.28	0.18	I.
12	Veľký Ďur 400 kV	15/05/2008	16/04/2009	0.00336456	0.01136877	0.0034985	0.00064291	0.00302167	0.00061612	1126.2	1033.9	1193.8	0.016471624	0.004889243	2484.27	0.20	I.
13	Bystričany 220 kV	15/05/2008	17/04/2009	0.00967578	0.01278853	0.0066166	0.00298441	0.00412533	0.00197159	1416.6	1297.3	1431.3	0.017324492	0.005810816	1992.46	0.20	I.
14	Križovany 400 kV	15/05/2008	17/04/2009	0.01148737	0.00796135	0.00451643	0.00342946	0.00235733	0.00092686	1310.7	1262.5	1311.6	0.016561182	0.004818413	2222.45	0.18	I. **)
15	Križovany 220 kV	15/05/2008	17/04/2009	0.00960248	0.00675054	0.00379852	0.0028797	0.00243234	0.00081435	1237.5	987.5	1078.6					
16	V427	05/05/2011	05/04/2012	0.0097	0.0190	0.0442	0.0031	0.0046	0.0180	1200.89	1268.75	821.43	0.106559	0.037356	1470.006	5.85	III.
17	V427	05/05/2011	05/04/2012	0.0153	0.0455	0.0153	0.0028	0.0104	0.0045	1291.07	984.82	929.46	0.113406	0.02115	1915.027	4.59	II.
18	V496	12/05/2010	20/04/2011	0.0177	0.0249	0.0359	0.0073	0.0081	0.0068	1041.1	1258.9	1192	0.078765523	0.054579695	2551.631659	10.97	IV.
19	Poznámky:																
20	**) - hodnoty boli vypočítane pre rozvodňu ako celok																

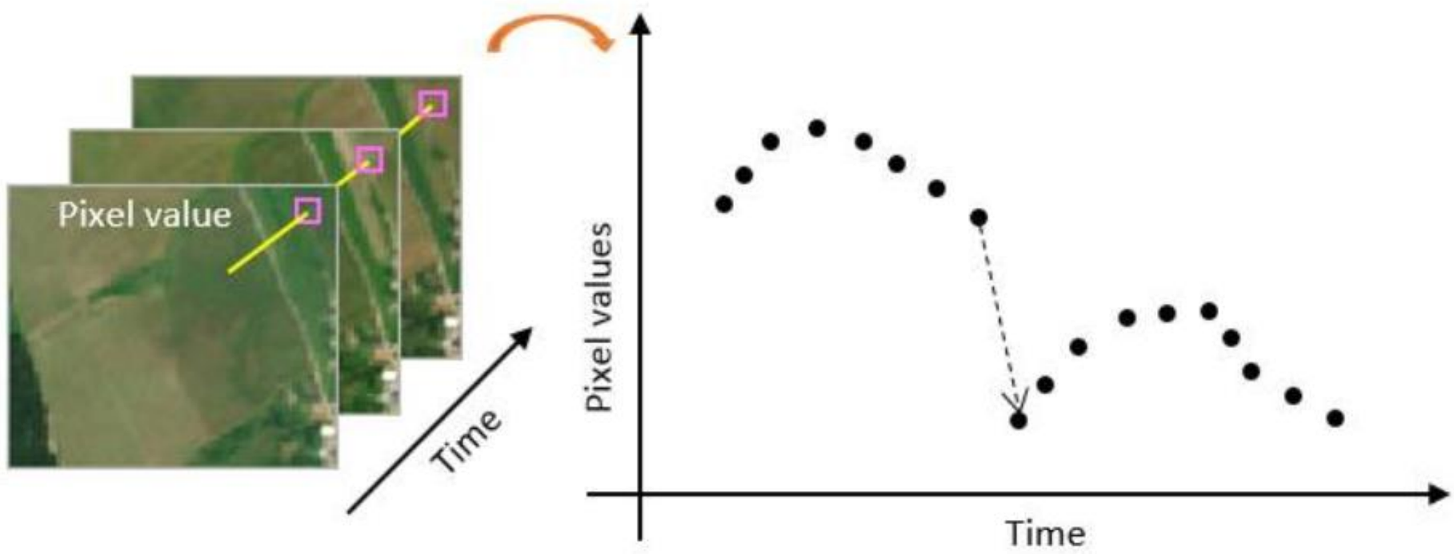
Satellite data and attributes

Calculation of satellite attributes (1)

Table 1. List of used satellite spectral bands including normalized difference indices.

Name	Scale	Pixel Size	Wavelength	Description
AOT	0.001	10 m		Aerosol optical thickness
B1	0.0001	10 m	443.9 nm	Aerosols
B2	0.0001	10 m	496.6 nm	Blue
B3	0.0001	10 m	560.0 nm	Green
B4	0.0001	10 m	664.5 nm	Red
B6	0.0001	20 m	740.2 nm	Red Edge 2
B8	0.0001	10 m	835.1 nm	NIR
L1 B10 cir	0.0010	60 m	1373.5 nm	Cirrus
B11	0.0001	20 m	1613.7 nm	SWIR 1
NDVI (normalized difference vegetation index)	0.0001	10 m		NDVI = (B8 − B4)/(B8 + B4)
NDWI (normalized difference water index)	0.0001	10 m		NDWI = (B3 − B8)/(B3 + B8)
NDSI (normalized difference soil index)	0.0001	20 m		NDSI = (B3 − B11)/(B3 + B11)
Moisture index	0.0001	20 m		moisture index = (B8 − B11)/(B8 + B11)

Calculation of satellite attributes (2)



<u>13 spectral bands or common indices:</u> **AOT, B1, B2 ... B11, NDWI, Moist**

<u>5 types of time series</u>: **Val, roz, dif (roz/days), absRoz, absDif**

- In total => 5 × 13 = ***65 time series per pixel***

<u>Double statistical aggregation</u>:

1. **Temporal** (per pixel): **Min, max, avg, P25, P50, P75**
2. **Spatial** (per measurement location) : **Min, max, avg, P25, P50, P75**

<u>Example satellite attribute names:</u>

- **B10_roz_avgT_Q50S** ,
- **NDWI_absdif_maxT_Q50S**

B8 = NIR, B10 = cirrus, B11 = SWIR

Satellite Attribute Results Summary *(see [ref.2])*

For individual target variables S, Sr, and g02, before the introduction of the new input attributes (i.e., without the measurement date and the satellite-derived ones), the correlation of their estimated values with the true ones was 0.23, 0.23, and 0.12, respectively. After adding the satellite-derived attributes, these correlations rose to 0.50, 0.33, and 0.46, respectively. The maximum correlations were achieved after the measurement date (Day of Year, number of the middle day) was also added; the correlations then rose to 0.54, 0.40, and 0.64, respectively. Among the three modeled target variables, the highest accuracy was achieved for g02, where the relative absolute error was 0.762 (down from original 1.04) and the correlation coefficient was 0.64 (up from original 0.12). This striking difference can also be observed in Figure 5: on the left are the results of our earlier attempts to estimate the long-term upper bounds of g02 purely from the SHMU input attributes [15], and on the right, the results of our present attempt to directly estimate the field-measured values of g02 from satellite-derived attributes and the measurement date. In both cases, the ideal situation in which the estimated values equal the predicted ones is represented by the dashed diagonal line, while the solid green line represents the actual linear regression

Satellite Attribute Results Summary (2): g02

Correlation (no satellites): 0.12

Correlation (with satellites): 0.64

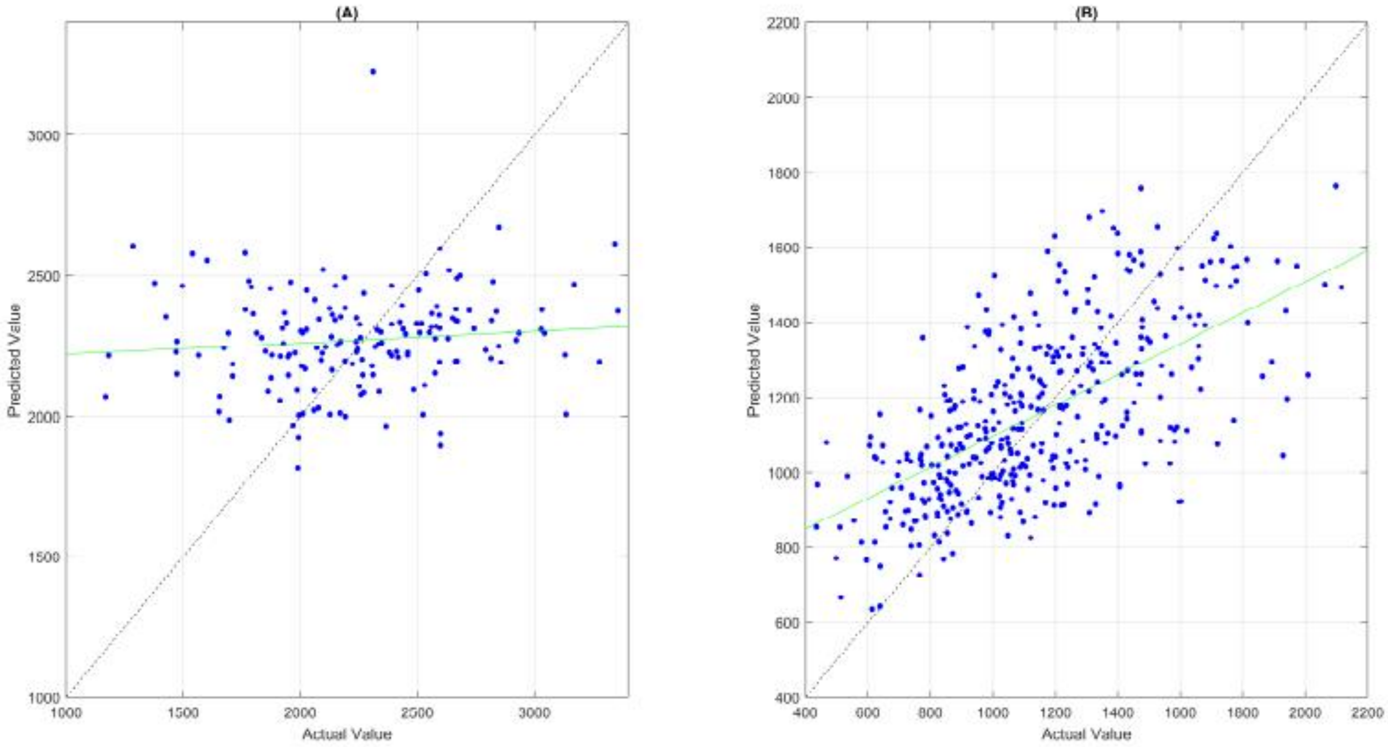


Figure 5. Correlation of actual values (horizontal axis) with predicted values (vertical axis) for (**A**) long-term upper bounds of g02 estimated from the SHMU input attributes, and (**B**) field-measured values of g02 estimated from the satellite-derived attributes and the measurement date.

Conclusion: Satellite data helped increase the correlation of predicted vs measured values from low (cca 0.2-0.3) to medium (cca 0.5-0.6) levels

Adding more new types of attributes

- **Hourly precipitations**
  2020-21 (radar products of SHMÚ)
- **Hourly air concentrations of PM10, PM2.5, O3, SO2, NO2 (N)**
  2020-21 (SHMÚ, RIO+CMAQ)
- **Hourly wet depositions of PM10, PM2.5, O3, SO2, NO2 (N)**
  2021 (SHMÚ, CMAQ)

Hourly precipitations 2020-21
(Aggregated daily and six-weekly)

- ***We expected smooth sailing to high correlations > 75% ...***
- ***... but were sadly disappointed:***

- ***The largest improvement was achieved for Sr:***
- ***Correlation 46%*** *(previously 40%)*
- ***(6% increase, 13 input attributes, better than pairwise interactions)***

- ***No tangible improvement even after adding attributes derived from hourly concentrations of the main pollutants and their wet depositions***

How to Break the Deadlock?
Where is the problem?

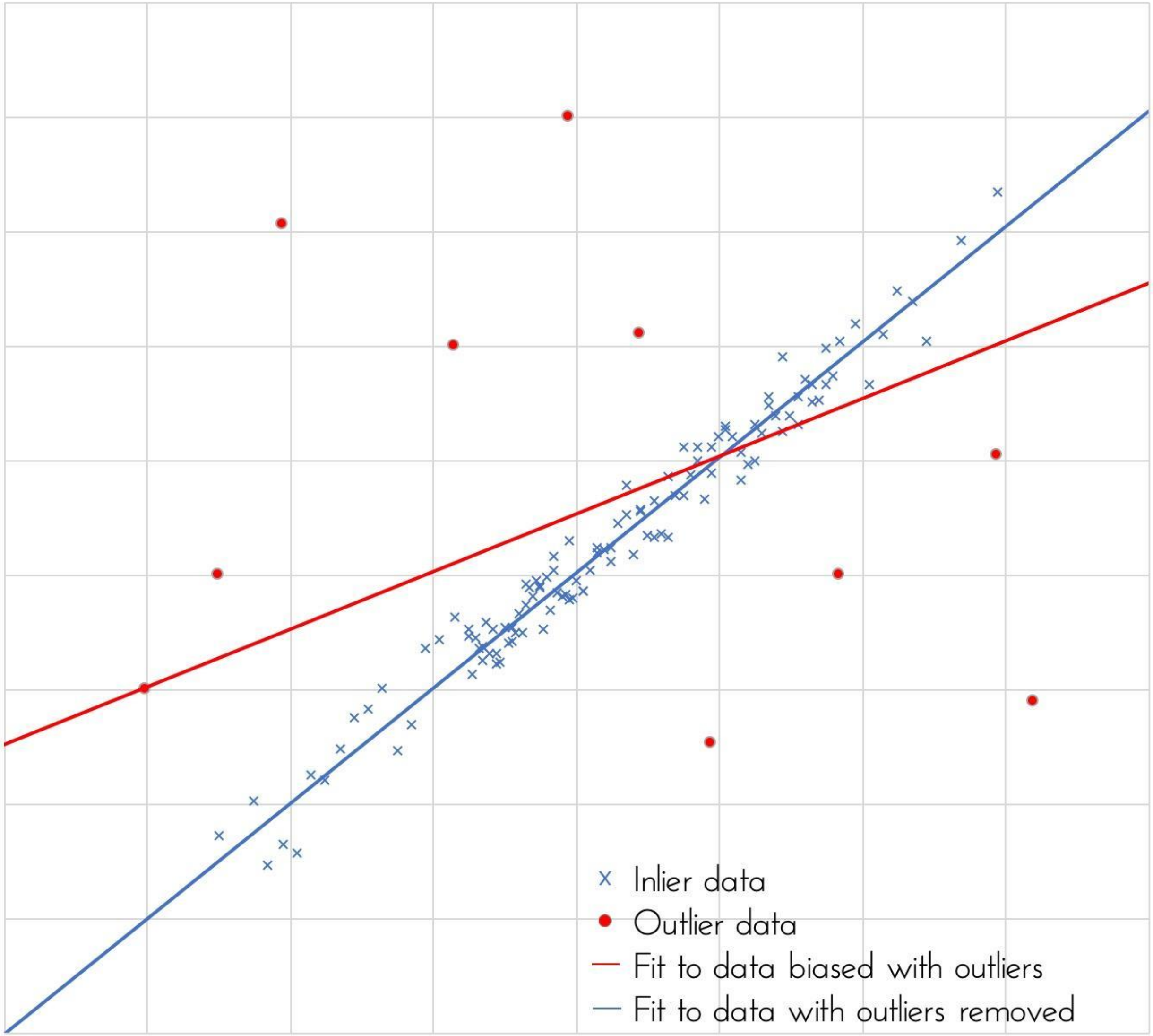
- Wrong Data Types?
  - We do not think so...
- Small Data Quantity?
  - Perhaps, but cannot do much about it...
- Poor Data Quality?
  - Most readily available option: How to deal with noisy data?

How to Deal with Noisy Data?
Tentative answer: Robust Statistics

The essence of robust statistics is...

- **... best shown by an illustrative example:**

Robust statistics – an illustrative example

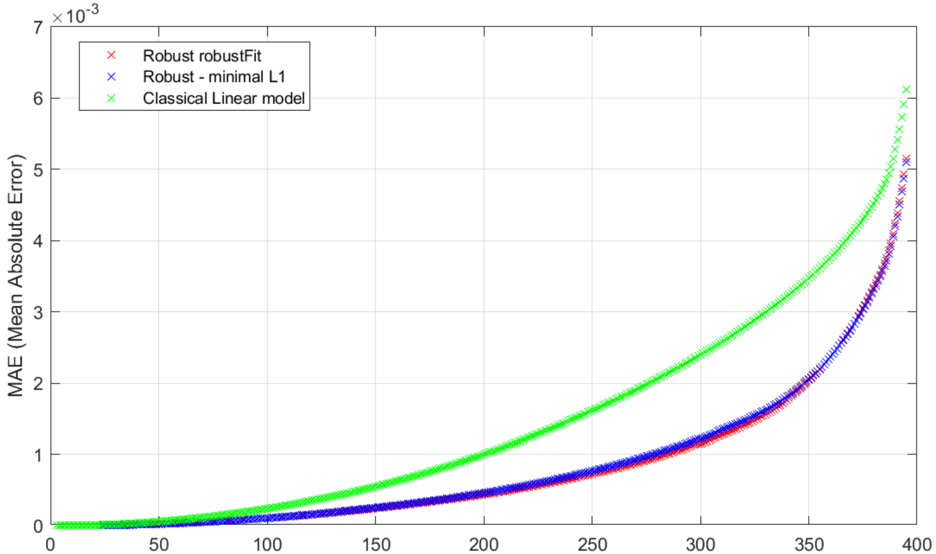


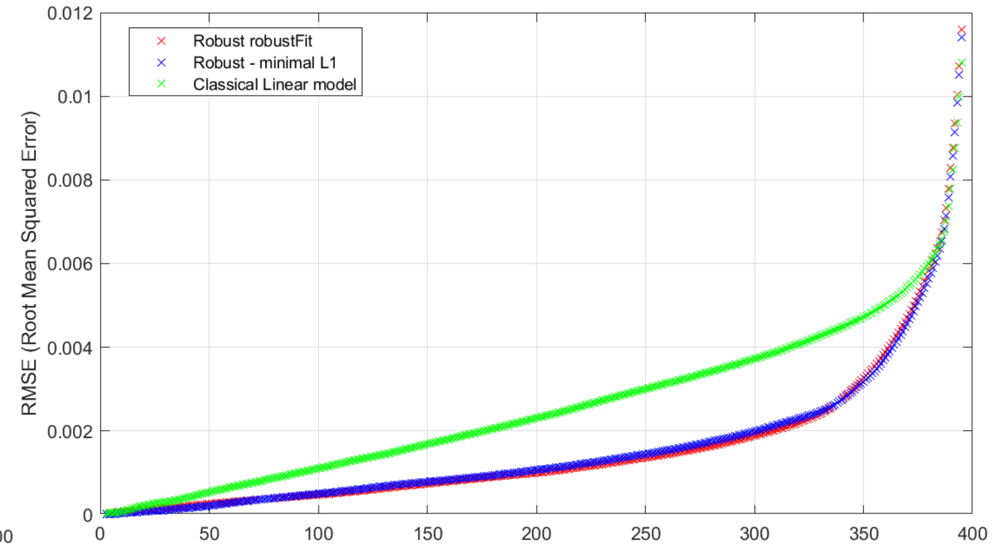
Robust strategies:

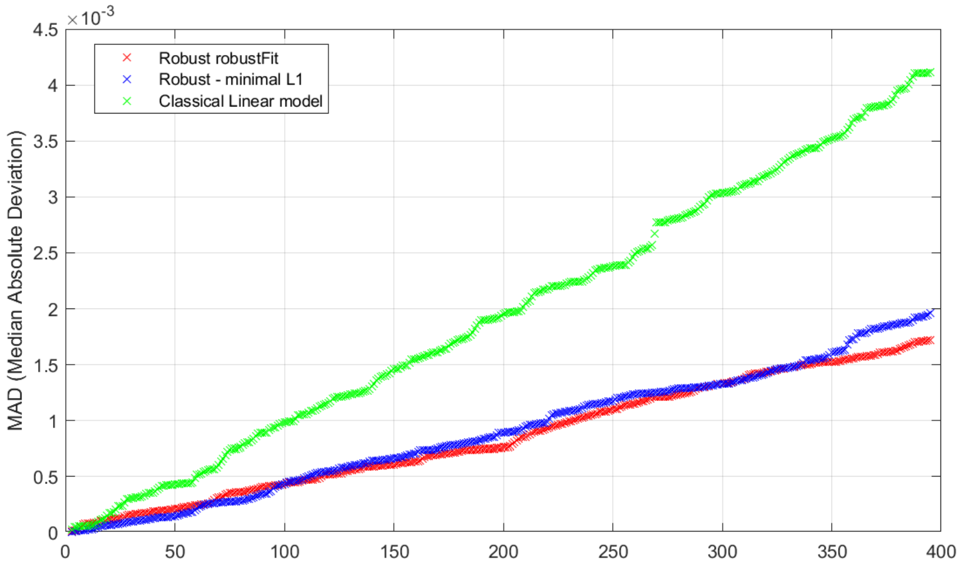
1. Remove outliers and do ordinary statistics (least squares)
2. Use L1 (MAE) instead of L2 (MSE)
3. L2-based iterative re-weighting schemes for outlier treatment

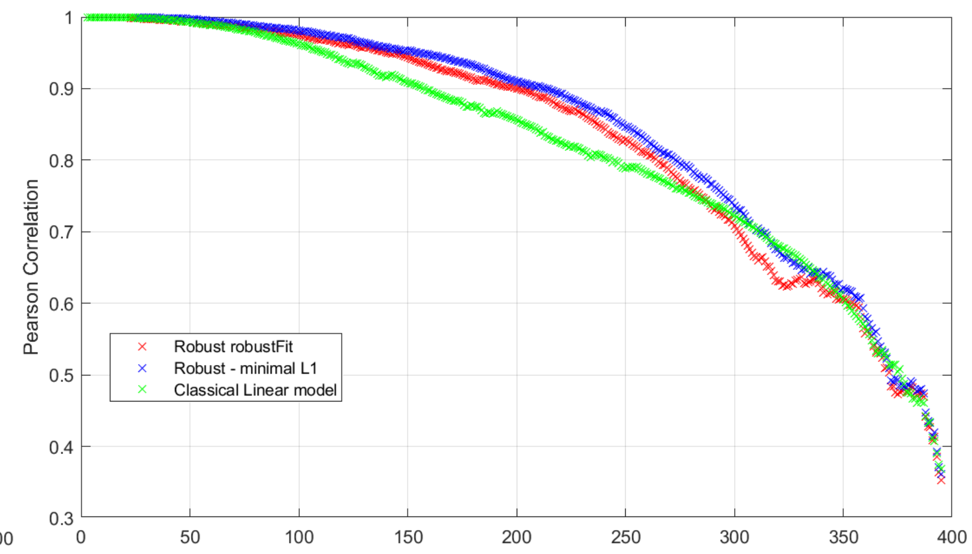
Credit: https://www.adelaide.edu.au/aiml/our-research/machine-learning/robust-statistics

Robust statistics – application to our data (Residual error analysis for total deposit, S)









Ongoing process of evaluation against real field-measured data (VUJE, STN 33 0405):

Ambition to replace **cumbersome field measurements** with
Artificial intelligence + robust statistics
Preliminary achievement: prediction errors reduced by cca 50% for 75% of 6-week field measurements
Evaluation procedure: (summary)

	A	E	F	G	H	I	O	P	Q	W	X	Y	AE	AF	AG	AH	AI
1							Namerané hodnoty							Ročné výsledky			
2				S: Total deposit (mg/day*cm²)			Sr: Soluble fraction			g0.2: Conductivity			Upper bounds Q99.5% S	Sr	g0.2	p (súčin)	Stupeň znečistenia
3		Doba merania		Číslo zberu			Číslo zberu			Číslo zberu							
4	Stanice	od	do	1	2	3	1	2	3	1	2	3					
5	Podunajské biskupice	13/05/2008	14/04/2009	0.00532007	0.01402077	0.01094552	0.00182158	0.00391103	0.00392174	1256.2	1411.6	1254.5	0.021485180	0.005984409	2192.68	0.28	I.
6	Senica 220 kV	13/05/2008	14/04/2009	0.00717915	0.01112768	0.0138279	0.00179479	0.00199302	0.00126439	1277.7	953.3	1069.6	0.024007898	0.002585061	2135.60	0.13	I.
7	Bošaca 400 kV	13/05/2008	14/04/2009	0.00589869	0.01066157	0.00720594	0.00214839	0.00288773	0.0012001	1475.9	1485.7	1211.6	0.017694859	0.003666455	2058.38	0.13	I.
8	Lemešany 400 kV	14/05/2008	15/04/2009	0.01019011	0.00708271	0.00557723	0.0024377	0.00124831	0.00051968	1402.1	1502.7	1261.6	0.013381296	0.004684182	2070.61	0.13	I. **)
9	Lemešany 220 kV	14/05/2008	15/04/2009	0.01070443	0.00688984	0.00468252	0.00252342	0.00156977	0.00087864	1333.9	1258.9	1178.4					
10	Rimavská Sobota 400 kV	14/05/2008	16/04/2009	0.00825066	0.00808458	0.00432356	0.00230911	0.00209481	0.00075542	1200.0	1220.5	1182.1	0.014782757	0.003673845	1949.54	0.11	I.
11	Levice 400 kV	15/05/2008	16/04/2009	0.00568439	0.00879714	0.00286094	0.00158584	0.00171978	0.00056255	1350.7	971.1	1007.1	0.016208572	0.004855633	2240.28	0.18	I.
12	Veľký Ďur 400 kV	15/05/2008	16/04/2009	0.00336456	0.01136877	0.0034985	0.00064291	0.00302167	0.00061612	1126.2	1033.9	1193.8	0.016471624	0.004889243	2484.27	0.20	I.
13	Bystričany 220 kV	15/05/2008	17/04/2009	0.00967578	0.01278853	0.0066166	0.00298441	0.00412533	0.00197159	1416.6	1297.3	1431.3	0.017324492	0.005810816	1992.46	0.20	I.
14	Križovany 400 kV	15/05/2008	17/04/2009	0.01148737	0.00796135	0.00451643	0.00342946	0.00235733	0.00092686	1310.7	1262.5	1311.6	0.016561182	0.004818413	2222.45	0.18	I. **)
15	Križovany 220 kV	15/05/2008	17/04/2009	0.00960248	0.00675054	0.00379852	0.0028797	0.00243234	0.00081435	1237.5	987.5	1078.6					
16	V427	05/05/2011	05/04/2012	0.0097	0.0190	0.0442	0.0031	0.0046	0.0180	1200.89	1268.75	821.43	0.106559	0.037356	1470.006	5.85	III.
17	V427	05/05/2011	05/04/2012	0.0153	0.0455	0.0153	0.0028	0.0104	0.0045	1291.07	984.82	929.46	0.113406	0.02115	1915.027	4.59	II.
18	V496	12/05/2010	20/04/2011	0.0177	0.0249	0.0359	0.0073	0.0081	0.0068	1041.1	1258.9	1192	0.078765523	0.054579695	2551.631659	10.97	IV.
19	Poznámky:																
20	**) - hodnoty boli vypočítane pre rozvodňu ako celok																

References *(work in progress, more to be expected)*

References:

- **Ref.1:** Krammer, P., Kvassay, M., Forgáč, R., Očkay, M., Skovajsová, L., Hluchý, L., Skurčák, Ľ., Pavlov, Ľ. Regression analysis and modeling of local environmental pollution levels for the electric power industry needs. *Computing and informatics*, 2022, vol. 41, no. 3, p. 861-884. URL: https://doi.org/10.31577/cai_2022_3_861

- **Ref.2:** Krammer, P., Kvassay, M., Mojžiš, J., Kenyeres, M., Očkay, M., Hluchý, L., Pavlov, Ľ. and Skurčák, Ľ., 2022. Using satellite imagery to improve local pollution models for high-voltage transmission lines and insulators. *Future Internet*, 14(4), p.99. URL: https://www.mdpi.com/1999-5903/14/4/99

- **Ref.3:** Krammer, P., Kvassay, M., Habala, O., Mojžiš, J., Hluchý, L., Pavlov, L. and Skurčák, L., 2023. Refinement of an Environmental Pollution Model for the Needs of the Electric Power Industry by Addition of Precipitation Attributes. In *SACI 2023* (pp. 611-616). IEEE.

Main avenues for future work:

- **Domain-motivated (interpretable) input atributes – precondition for robustness**
- **There is a well-defined robust statistical methodology that should be observed:**

Hoaglin, D.C., Mosteller, F. and Tukey, J.W. eds., 2000. Understanding robust and exploratory data analysis (Vol. 76). John Wiley & Sons.

Thank You
Any questions?