

Dataset (IoT) V1

This data concerns readings from 4 base stations (particles and NO2) and weather supplied by NTNU and TELENOR

Traffic data was downloaded from a public website
See further details in the next slides.

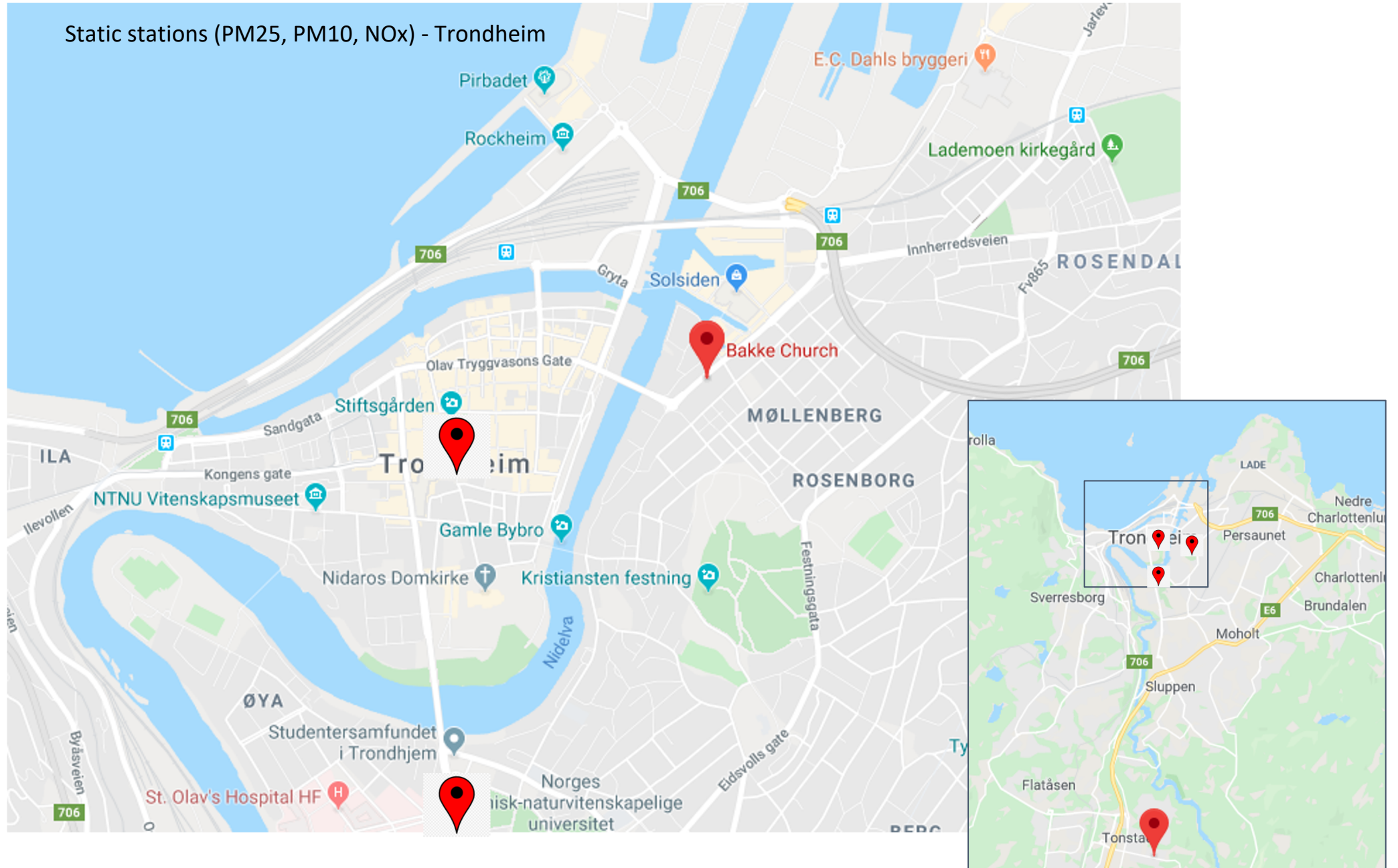
The data is saved in file DataloTPhysicalAI_v1.zip

Contacts (search AI4EU members) :

Dataset prepared by João Paulo Costeira

Owners Sigmund Akelsen, Kerstin Bach

Trondheim Pollution sensing infrastructure

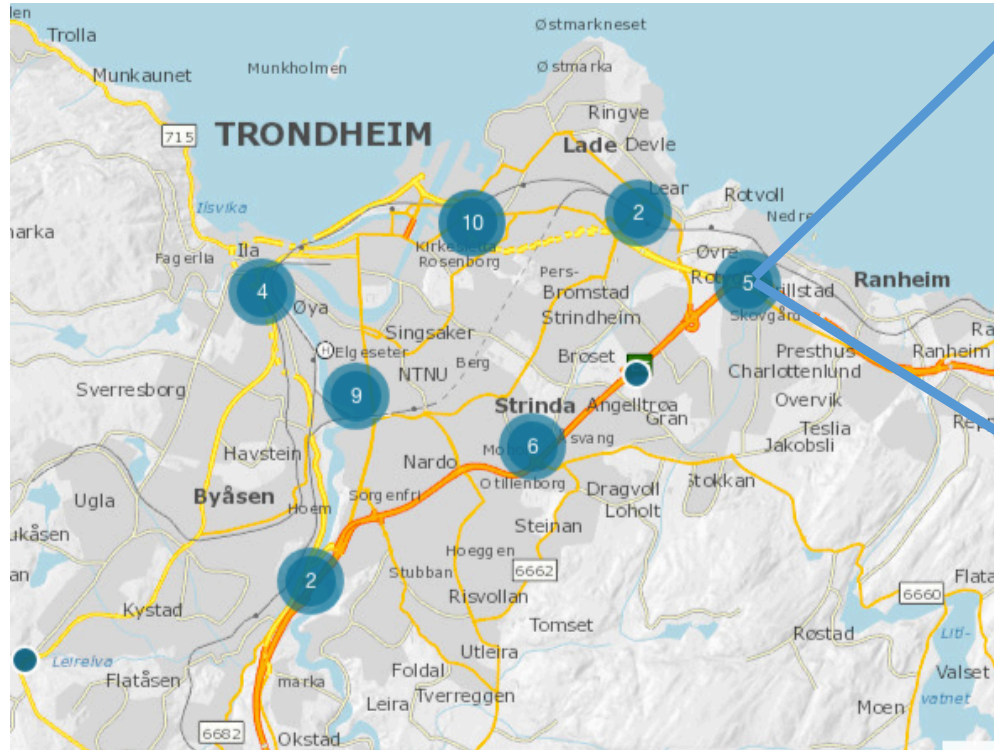


TELENOR/NTNU dataset

	NILU_Torvet	NILU_Torvet	NILU_Torvet	NILU_Elgeseter	NILU_Elgeseter	NILU_Elgeseter	NILU_Bakke kirke	NILU_Bakke kirke	NILU_Bakke kirke	NILU_E6-Tiller	NILU_E6-Tiller	NILU_E6-Tiller	MET_SN68860	MET_SN68860	MET_SN68860	MET_SN68860	MET_SN68860	MET_SN68860
timestamp	pm25	pm10	no2	no2	pm25	pm10	pm25	pm10	no2	pm25	pm10	no2	temperature	rain	pressure	humidity	wind_from_direction	wind_speed
02/06/2019 02:00	2.9	4.1	1.8	8.499642999999999	3.5376269999999996	8.495968	2.06546	4.816458	4.271076	1.225017	3.85926	5.5837	7.0	0.0	997.2	82.0	225.33333333333331	1.4666666666666668
02/06/2019 03:00	2.9	4.2		3.9199589999999995	1.2460149999999999	3.899939	0.0	0.532546	3.860525	0.413038	1.534032	5.124034	6.9	0.0	997.4	82.0	239.66666666666666	1.7333333333333334
02/06/2019 04:00	1.6	2.9	2.1	4.919745	1.214907	2.1575450000000003	0.0	0.934321	2.1577509999999999	0.0	0.908867	4.485728	7.2	0.0	997.6	80.0	225.33333333333331	1.4666666666666668
02/06/2019 05:00	1.3	1.9	6.1	11.650747	1.221206	3.4818559999999996	0.0	1.8199400000000001	6.076667	0.013578	1.044722	6.9230589999999999	7.6	0.0	997.6	78.0	218.16666666666666	1.2500000000000002
02/06/2019 06:00	1.4	2.1	8.2	7.141941	0.0	1.978441	1.23475	4.643394	1.400892	2.906365	4.0044379999999995	8.11656	7.2	0.0	997.8	79.0	181.16666666666666	0.7333333333333334
02/06/2019 07:00	1.5	1.9	6.2	14.654339000000002	1.729712	3.249278	0.012589	1.6142379999999998	6.629839	3.065448	4.2686199999999999	10.121107	7.6	0.0	998.1	79.0	292.16666666666667	0.5666666666666665
02/06/2019 08:00	1.4	1.7	4.8	12.878681	0.92719	2.279957	0.344179	1.764687	37.749658000000000	2.35584	3.516135	8.162879	8.6	0.0	998.1	75.0	131.33333333333334	0.7000000000000001
02/06/2019 09:00	1.5	1.7	10.4	25.226328	0.562619	2.1951400000000003		12.703332000000000	0.0	1.056363	9.208029	8.2	0.0	998.4	75.0	53.83333333333334	1.0333333333333334	
02/06/2019 10:00	1.7	2.1	23.6	20.909237	3.7561379999999995	13.09742	0.0	1.9969259999999998	9.908355	0.065737	2.435348	11.696416000000001	10.0	0.0	998.3	71.0	32.833333333333336	2.3333333333333333
02/06/2019 11:00	2.3	3.0	11.0	16.807358999999998	2.295492	9.152674000000001	1.874298	5.402904	6.488517	0.067347	4.256745	18.581031	11.2	0.0	998.0	62.0	37.333333333333336	2.9166666666666665
02/06/2019 12:00	2.2	2.8	8.5	22.327594	2.5315630000000002	11.330327	0.590106	4.87355	8.475633	1.2383309999999998	7.854373	25.265857	10.4	0.0	998.1	62.0	27.83333333333333	2.45
02/06/2019 13:00	1.4	1.8	9.5	22.677704000000002	1.952062	15.232548000000001	2.844785	6.29967	11.174353	2.558217	12.2676509999999999	24.407697	10.3	0.0	997.8	66.0	36.166666666666666	2.2833333333333333
02/06/2019 14:00	1.5	2.0	10.7	21.622743	2.866227	18.177683	5.707433	10.903429	10.750735	0.8821540000000001	11.986153	21.915302	11.6	0.0	997.5	63.0	31.833333333333333	2.9500000000000006
02/06/2019 15:00	1.7	2.3	11.5	17.092315	3.78632	20.720544	2.465564	6.046378	6.684139	0.490626	11.630504	28.147491	12.2	0.0	996.9	63.0	32.166666666666666	2.45
02/06/2019 16:00	1.7	2.4	11.5	22.555768	5.086475	19.913829999999997	2.040394	5.777467	14.571556	2.070141	11.667173	34.691654	11.7	0.0	996.3	64.0	38.5	2.8833333333333333
02/06/2019 17:00	2.5	3.5	10.0	15.589997	3.082189	11.889532	5.94747	12.2087	20.362251	1.176245	11.963697	29.968734	11.9	0.0	995.7	65.0	44.5	2.6833333333333336
02/06/2019 18:00	3.4	4.3	6.4	30.448653999999998	1.674411	7.634742	4.174035	12.547701	24.673407	3.465649	13.687917	31.302127000000002	11.9	0.0	994.9	67.0	39.5	2.7833333333333333
02/06/2019 19:00	3.9	4.9	9.3	16.411777	3.949785	14.696712	1.602508	10.784259	18.386236	2.658786	13.393331	30.959532	11.6	0.0	993.9	69.0	45.0	3.1333333333333333
02/06/2019 20:00	4.3	5.9	12.9	22.271142	2.275123	9.610968	6.0514730000000005	13.410557999999998	16.743631	3.8100400000000003	13.45584	22.628498	11.3	0.0	993.1	74.0	39.333333333333336	2.25
02/06/2019 21:00	3.5	7.3	15.6	18.584232	4.545344	16.125815	0.5823	9.1668729999999999	14.141863	3.263505	9.580097	18.20737	10.9	0.0	992.1	79.0	32.0	1.3333333333333333
02/06/2019 22:00	3.6	10.2	15.6	22.924073	3.6849160000000003	11.948322000000001	0.0	4.053516	11.783389	3.211709	9.711449	17.570470999999998	10.2	0.0	991.1	83.0	41.666666666666666	1.2
02/06/2019 23:00	3.5	9.2	12.5	12.724366999999999	1.599129	8.757839	4.483975	9.123535	10.171109	2.441758	7.6357539999999999	15.396094	12.4	0.0	989.9	69.0	159.5	0.9666666666666668
03/06/2019 00:00	3.1	7.0	10.0	10.231104	2.426734	5.6566160000000005	6.0202050000000025	8.426522	6.609430000000001	1.731366	3.999994	3.8716690000000002	11.4	0.0	989.3	77.0	210.0	2.1333333333333333
03/06/2019 01:00	3.0	6.3	5.5	6.395125	1.638333	3.3837260000000002	0.0	1.0502129999999998	4.088285	0.940249	1.768879	2.642702	13.0	0.0	988.5	72.0	196.5	2.6333333333333333
03/06/2019 02:00	2.4	4.2	4.2	3.7477440000000004	1.4511129999999999	2.476466	0.0	0.0	2.145491	2.893695	3.8298080000000003	3.9566879999999998	12.7	0.0	988.0	76.0	200.5	3.5833333333333333
03/06/2019 03:00	1.6	2.2	2.6	5.353004	3.6180220000000003	4.068568	0.0	0.0	6.872272	2.667265	3.768705	8.205467	12.8	0.0	987.3	77.0	191.5	4.1333333333333333
03/06/2019 04:00	1.2	1.5	4.3	13.897020000000001	0.89305	1.294103	3.808619	4.284639	7.169898	0.968968	3.60625	29.474759000000002	13.9	0.1	986.6	71.0	186.16666666666666	3.7333333333333334
03/06/2019 05:00	1.0	1.1	9.8	16.437314999999998	0.0734980000000001	1.103647	2.020492	2.887871	10.424985000000000	2.0462	6.855034	28.405911	14.4	0.0	986.0	67.0	194.66666666666666	4.216666666666667
03/06/2019 06:00	1.0	1.3	14.9	20.669467	0.0	3.17703	4.432632	9.800211000000001	16.453139	2.794245	9.132866	15.188278	14.5	0.0	985.4	69.0	190.33333333333331	5.266666666666667
03/06/2019 07:00	1.4	4.7	14.7	19.789704	3.853572	7.42085	3.5231879999999998	14.274239999999999	11.430937	2.1783900000000003	9.739873	18.462193	14.1	0.0	985.3	72.0	190.66666666666666	5.15
03/06/2019 08:00	1.9	8.1	13.6	18.215688	0.125748	1.9599119999999999	0.0	2.915773	6.343013	2.2876380000000003	5.560121	11.391485000000001	14.0	0.1	985.1	76.0	192.16666666666666	4.8166666666666666
03/06/2019 09:00	2.0	6.3	11.8	23.020447	0.0	1.555179	0.0	0.0	22.236809	2.293939	12.933831	24.257724	16.0	0.0	984.8	68.0	193.83333333333331	3.9666666666666667

PM10, PM25, NO2 in the for stations – long series, hourly sampled (several years)
 Meteo (particular station) – temperature, rain, pressure, humidity, wind direction, wind speed

Traffic data (public website)



Trafikkregistri Navn	Vegreferansé Fra	Til	Dato	Fra tidspunkt	Til tidspunkt	Volum	Dekningsgrai	Antall timer	Antall timer	Antall timer	Ikke gyldig le	Lengdekvallit	Felt	Felt gyldig fr	Felt gyldig til	< 5,6m	>= 5,6m	5,6m - 7,6m	7,6m - 12,5m	12,5m - 16,0	16,0m - 24,0	>= 24,0m
Trafikkregistri Navn	Vegreferansé from	to	date	fromtime	to time	Volum	coverage	Antall timer	Antall timer	Antall timer	not valid len	LENGTH QU	Felt	Felt gyldig fr	Felt gyldig til	< 5,6m	>= 5,6m	5,6m - 7,6m	7,6m - 12,5m	12,5m - 16,0	16,0m - 24,0	>= 24,0m
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T00:00+02:00	20/09/2019	00:00	01:00	30	100	1	1	0	9	70	1	2017-01-04T18:06:00.209	2017-01-04T18:06:00.209	17	4	2	0	1	1	0
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T00:00+02:00	20/09/2019	00:00	01:00	12	100	1	1	0	5	58,33	2	2017-01-04T18:06:00.209	2017-01-04T18:06:00.209	2	5	0	0	1	4	0
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T00:00+02:00	20/09/2019	00:00	01:00	16	100	1	1	0	5	68,75	3	2017-01-04T18:06:00.209	2017-01-04T18:06:00.209	8	3	0	0	2	0	1
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T00:00+02:00	20/09/2019	00:00	01:00	20	100	1	1	0	9	55	4	2017-01-04T18:06:00.209	2017-01-04T18:06:00.209	11	0	0	0	0	0	0
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T00:00+02:00	20/09/2019	00:00	01:00	3	100	1	1	0	1	66,67	5	2017-01-04T18:06:00.209	2017-01-04T18:06:00.209	0	2	0	1	0	1	0
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T01:00+02:00	20/09/2019	00:00	01:00	29	100	1	1	0	12	58,62	6	2017-01-04T18:06:00.209	2017-01-04T18:06:00.209	11	6	0	0	1	2	3
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T00:00+02:00	20/09/2019	00:00	01:00	61	100	1	1	0	26	57,38	Totalt i retnii	-	-	24	11	0	0	2	6	3
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T00:00+02:00	20/09/2019	00:00	01:00	49	100	1	1	0	15	69,39	Totalt i retnii	-	-	25	9	2	1	3	2	1
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T01:00+02:00	20/09/2019	00:00	01:00	110	100	1	1	0	41	62,73	Totalt	-	-	49	20	2	1	5	8	4
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T01:00+02:00	20/09/2019	01:00	02:00	25	100	1	1	0	4	84	1	2017-01-04T18:06:00.209	2017-01-04T18:06:00.209	17	4	1	2	0	1	0
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T01:00+02:00	20/09/2019	01:00	02:00	8	100	1	1	0	3	62,5	2	2017-01-04T18:06:00.209	2017-01-04T18:06:00.209	2	3	1	0	1	1	0
17961V7281	Holtermanns FV6690 HP 3	2019-09-20T01:00+02:00	20/09/2019	01:00	02:00	8	100	1	1	0	2	75	3	2017-01-04T18:06:00.209	2017-01-04T18:06:00.209	4	2	1	0	1	0	0

Hourly countings, type of vehicle



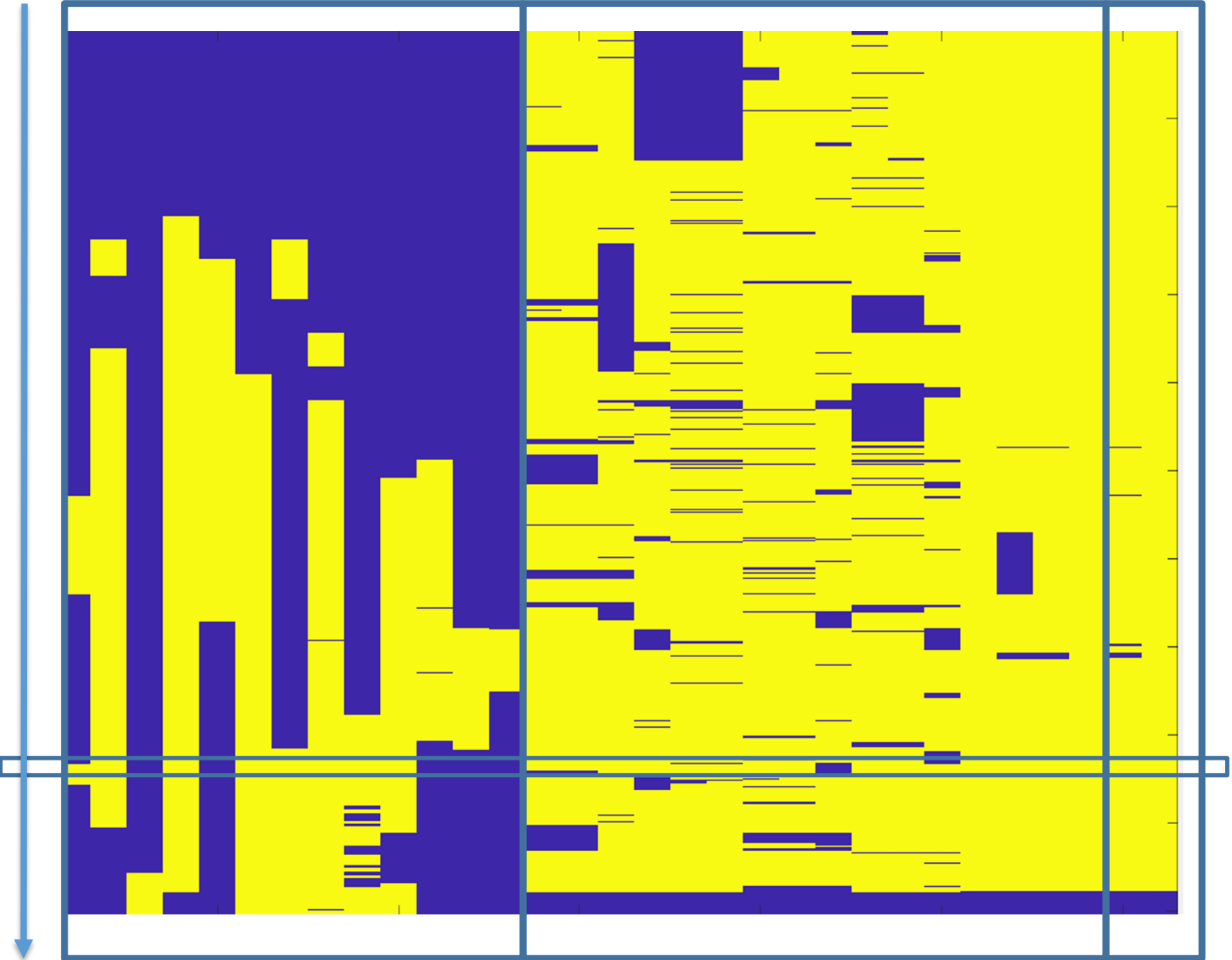
Missing



Available Data

Data with outliers (0's, NaN, negative values)
Missing values
Heterogeneous

Time



Each measurement is a multidimensional
Heterogeneous "point"

Traffic (13 sensors)

Pollution

Meteo

File dataset_Trondheim_pollutionv1.mat

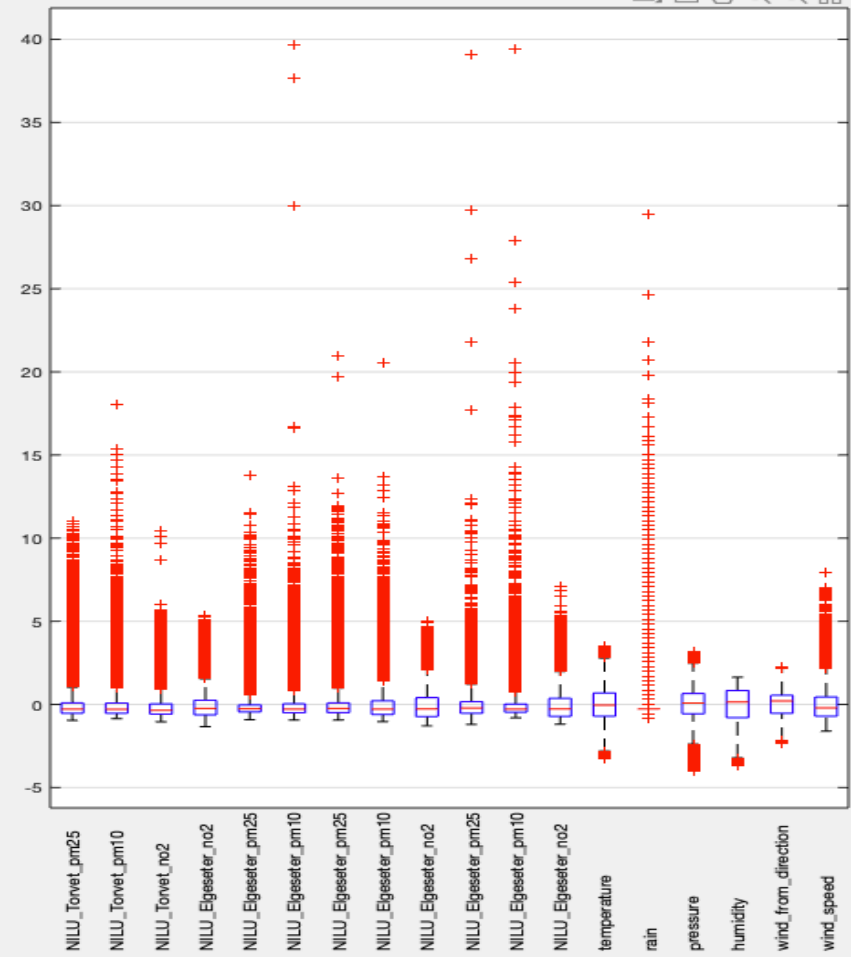
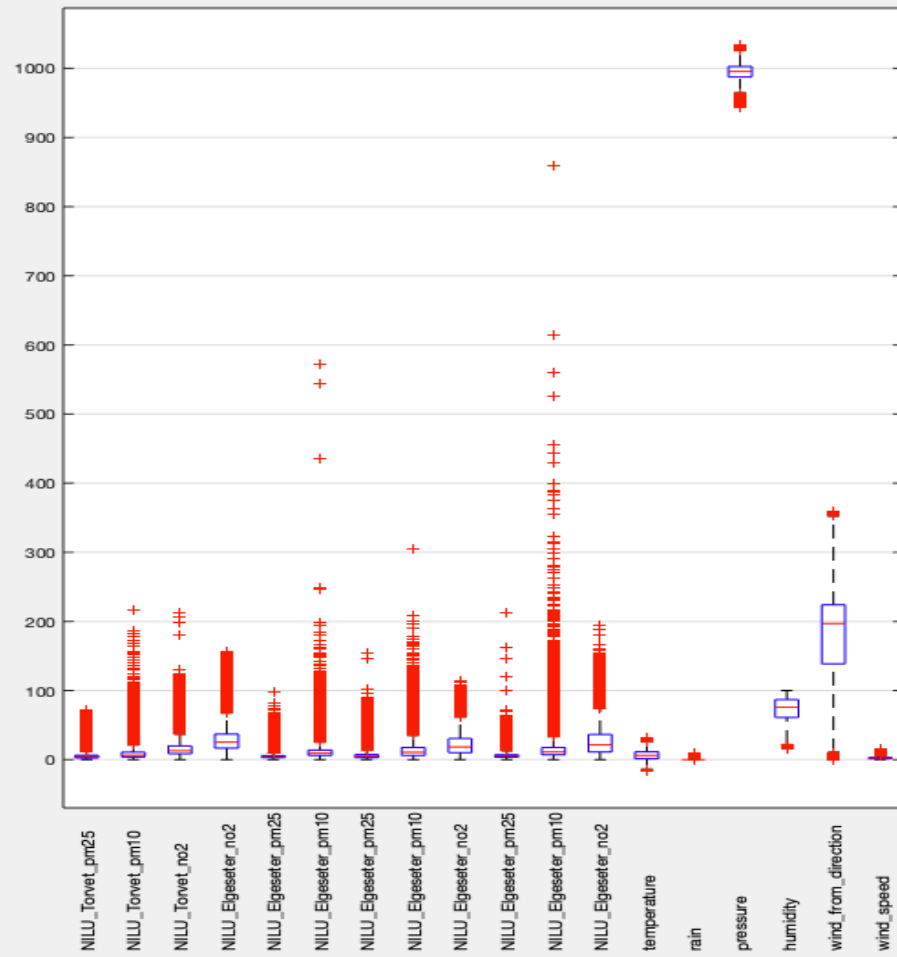
Name	Size	Bytes	Class	Attributes
datafilled	48864x18	7036416	double	
datafilled_byday	432x2036	7036416	double	
labels	1x19	2696	cell	
meanstdmedian	3x18	432	double	
missing	48866x18	7036704	double	
missing_byday	432x2036	7036416	double	
origdata	48866x18	7036704	double	
origdata_table	48867x19	117857500	table	
timestamps	48864x1	7427328	cell	
timestamps_days	2036x1	309472	cell	

Data is ordered by timestamp one sample every hour, from Jan 1, 2014 00:00 (timestamps{1}) until July 29 23:00 (timestamps{end})
Origdata has the hourly data order per rows. Each column is a variable (see next slide) and there are 18 categories of data.

Variables “by day” are just the “stack” of measurements for each day (each column has 18*24 measurements).

There is a difference of 2 measurementes because the original data ends at 1AM July 30 !

For the location of the sensors see slide in the beginning of this powerpoint

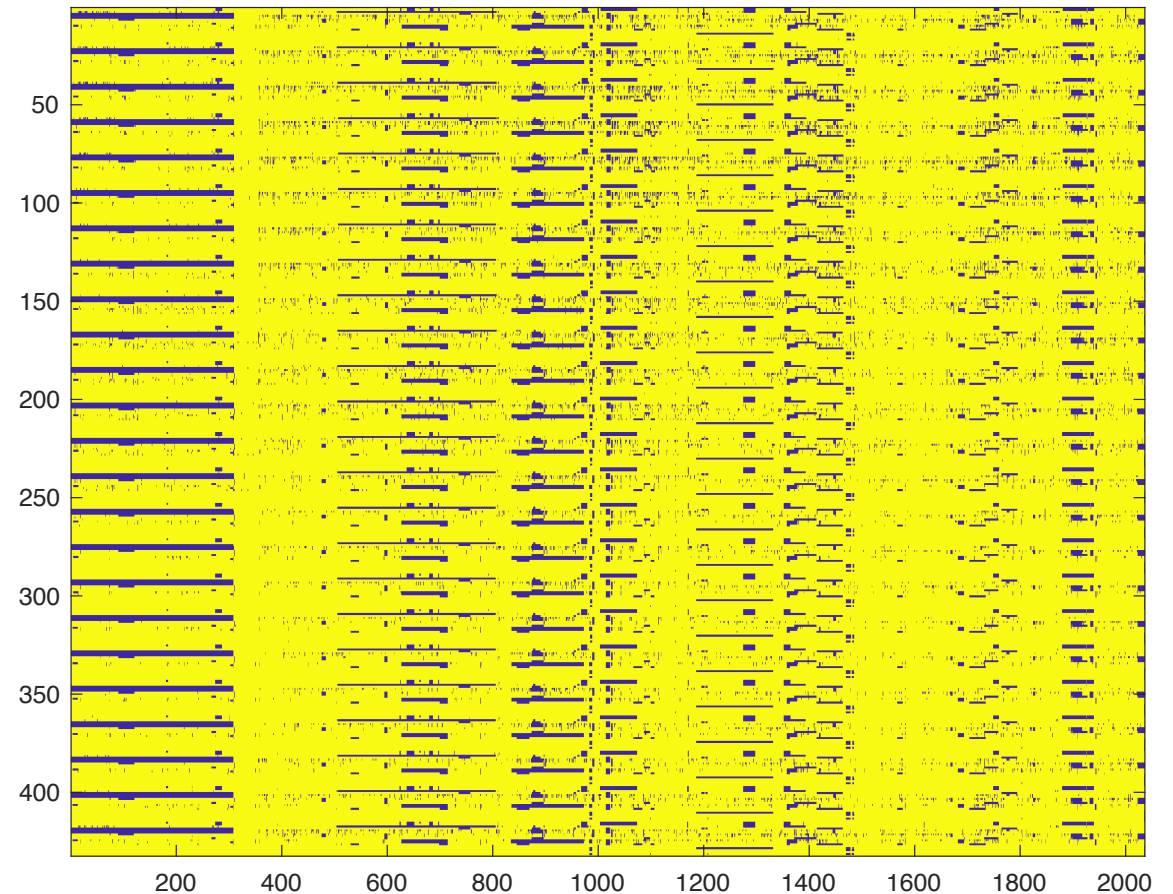


Distribution of the data. Boxplots show median, 25% and 75% quartiles. On the left original data, on the right normalized data $1/\sigma(x - \text{mean})$.

Variable are 1-12 pollution (pm10, pm2.5, no2) 4 stations and meto (rain, temperature, pressure wind_direction wind_speed)

Missing data (each column is one day of data)

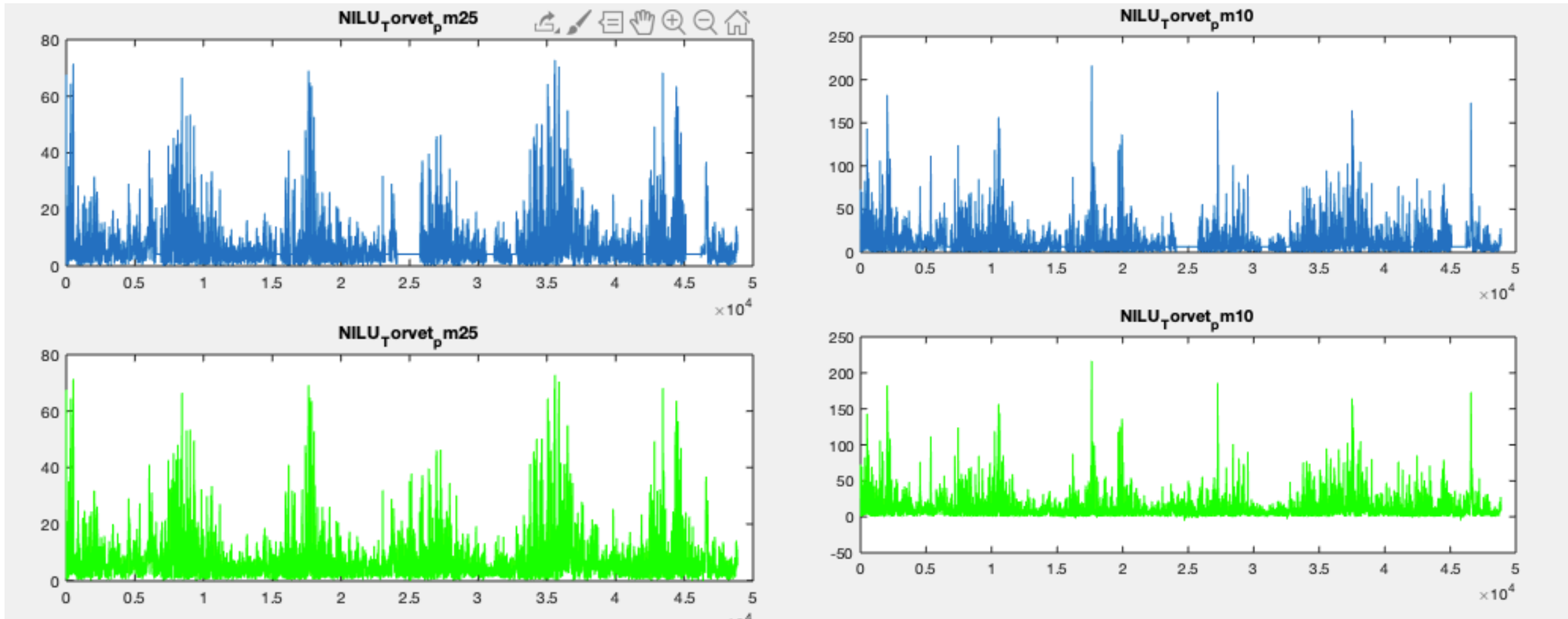
- Original data has missing values represented by NaN in the original data.
- The pattern of missing values are in blue.
- The missing data (11%) is scattered.

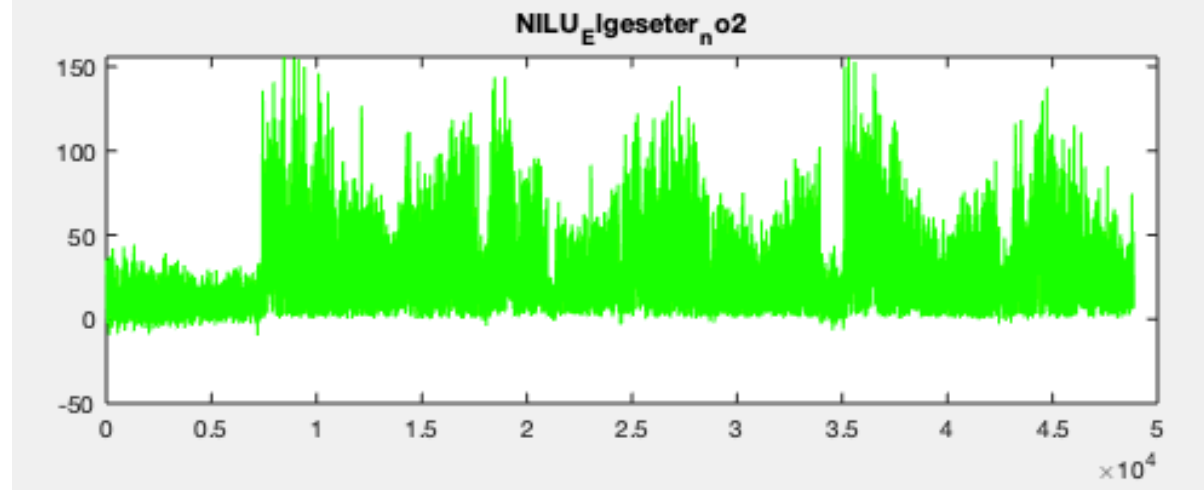
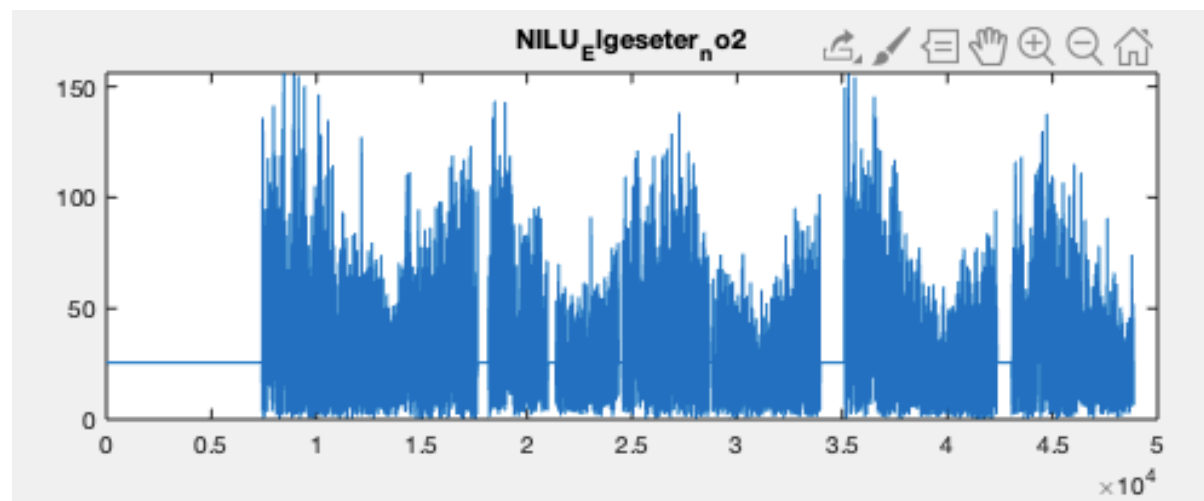
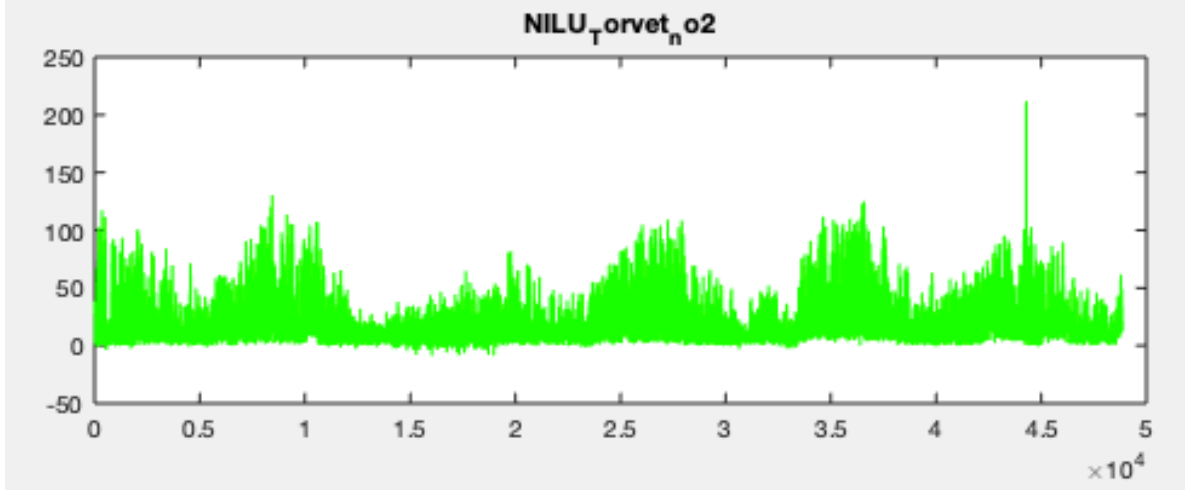
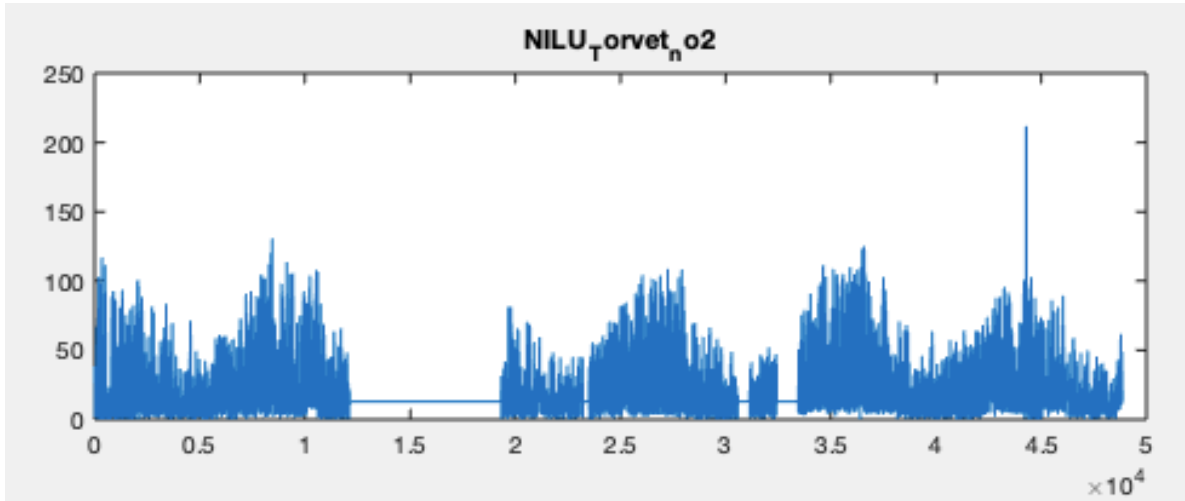


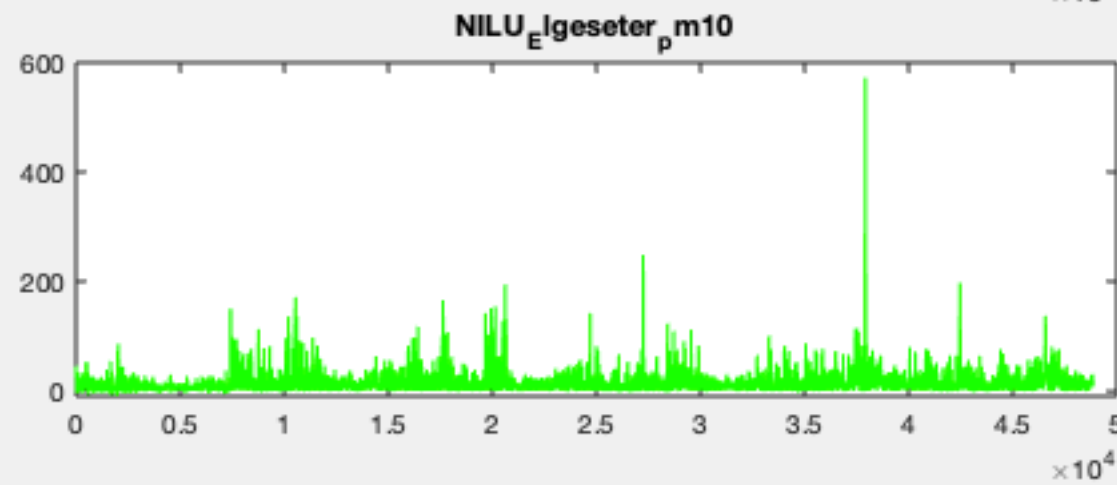
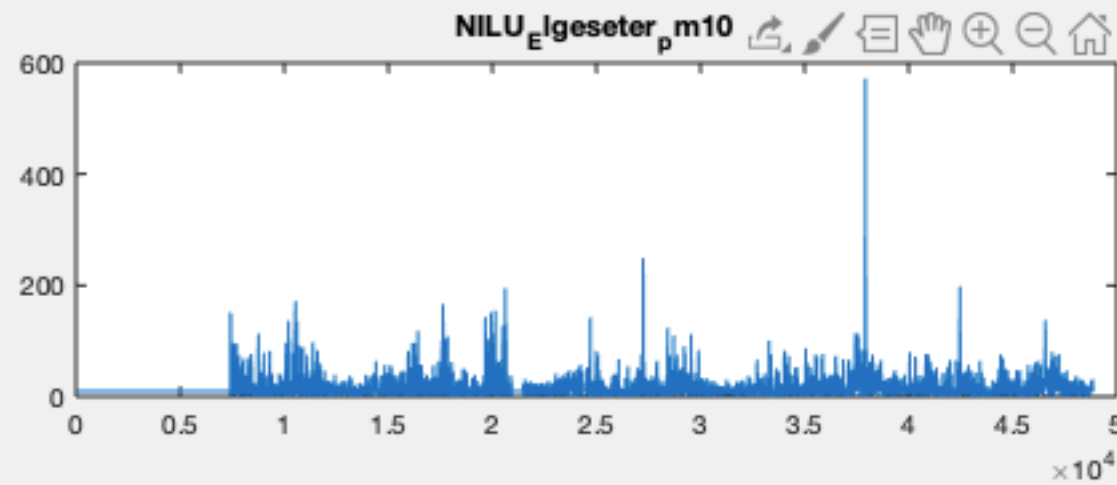
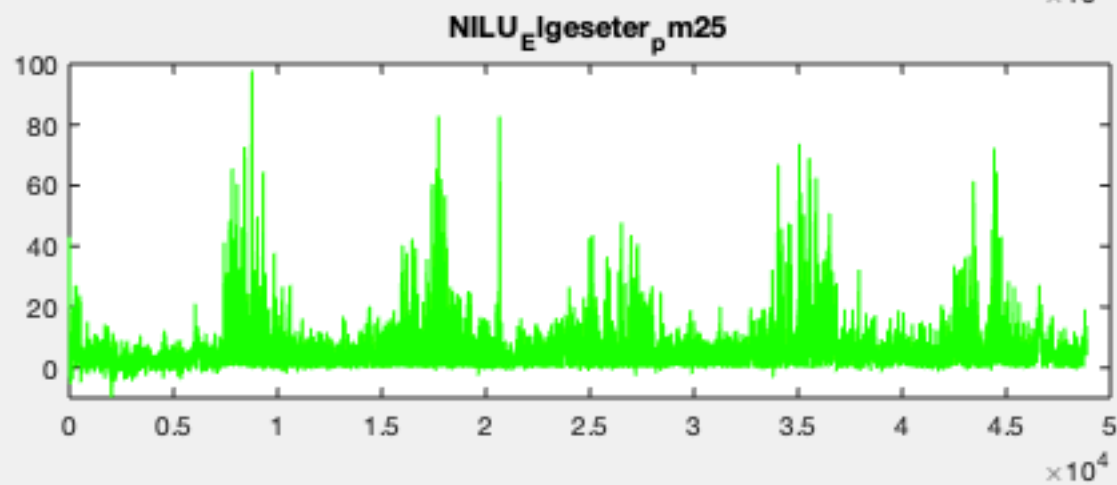
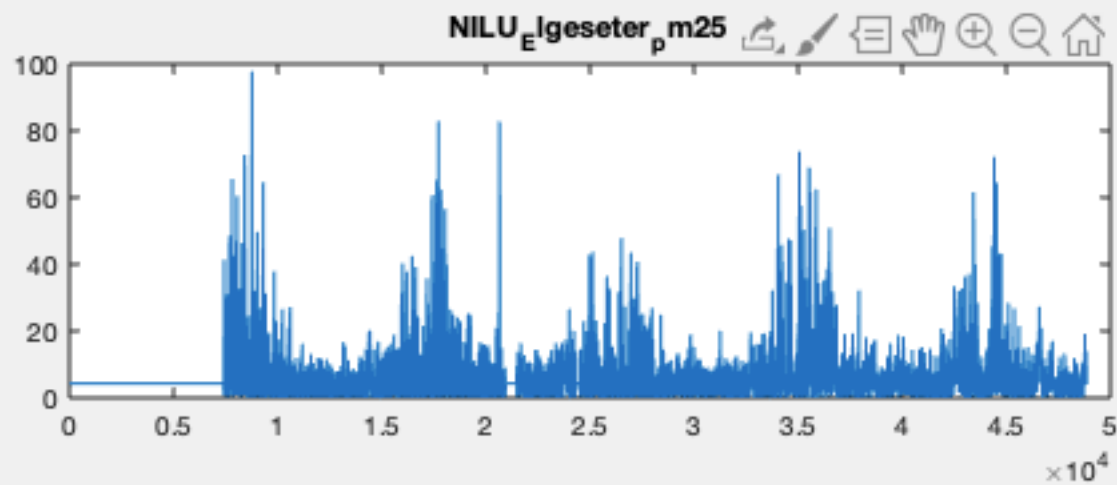
We run a “completion” algorithm just to “give it a shot” and check if the redundancy and regularity is enough to complete it. Of course it is not (missing is not random, we do not enforce positivity and did not normalize the data). In a rush to have complete time series ☺

Consider you have 18 variables for 3 years with 11% outliers ... for now!

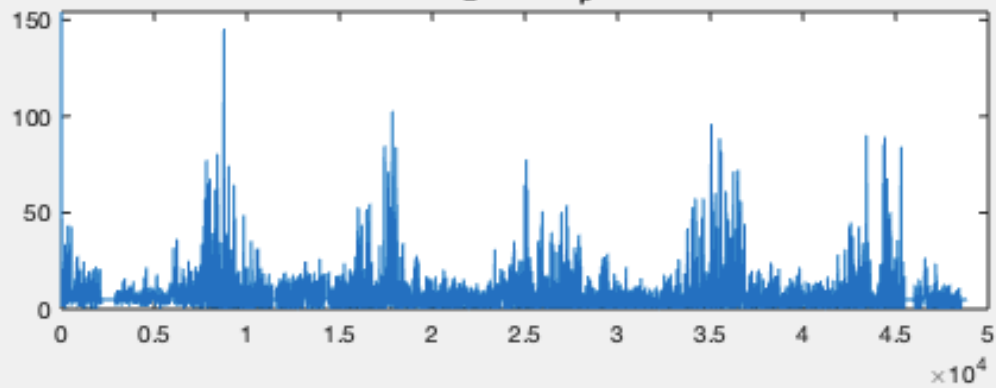
A glimpse on the data (blue) and the completion (green, var datafilled,datafilled_byday)



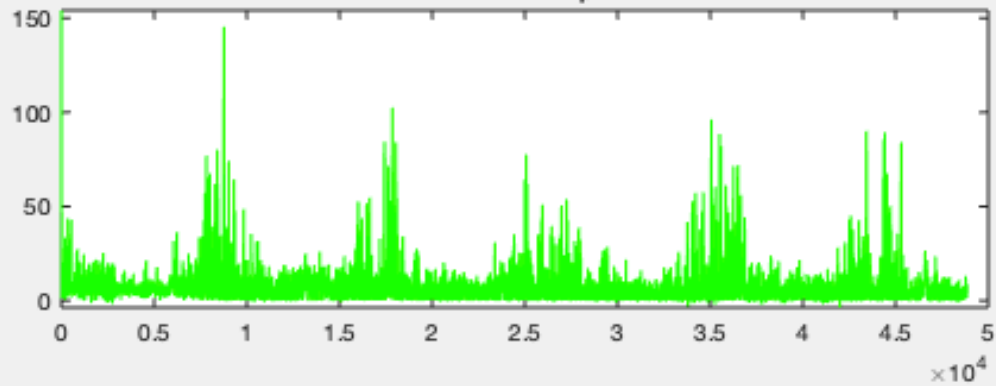




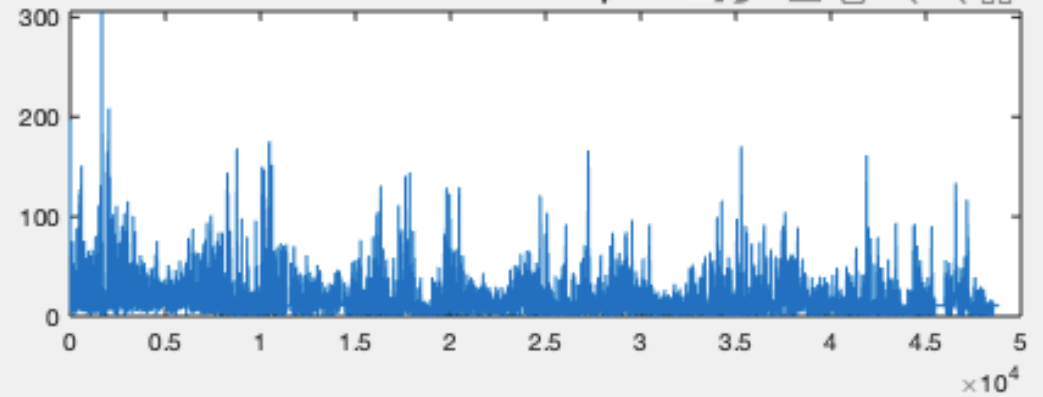
NILU_EIgeseter_p_m25



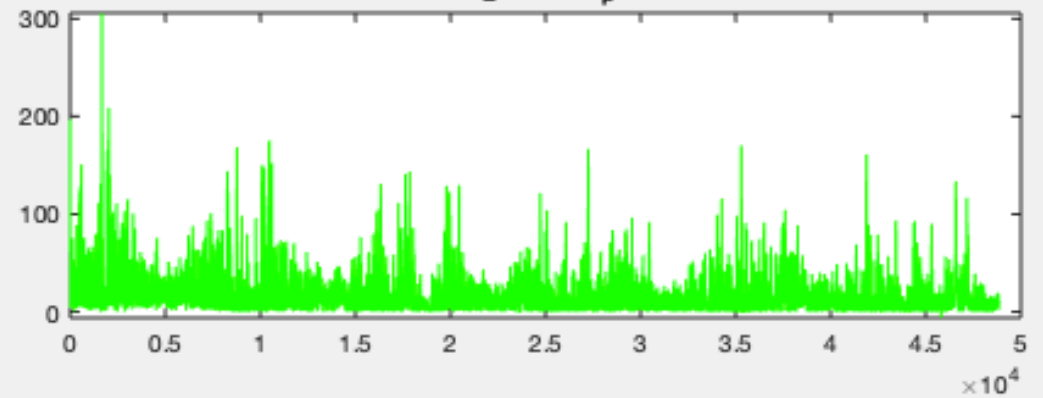
NILU_EIgeseter_p_m25



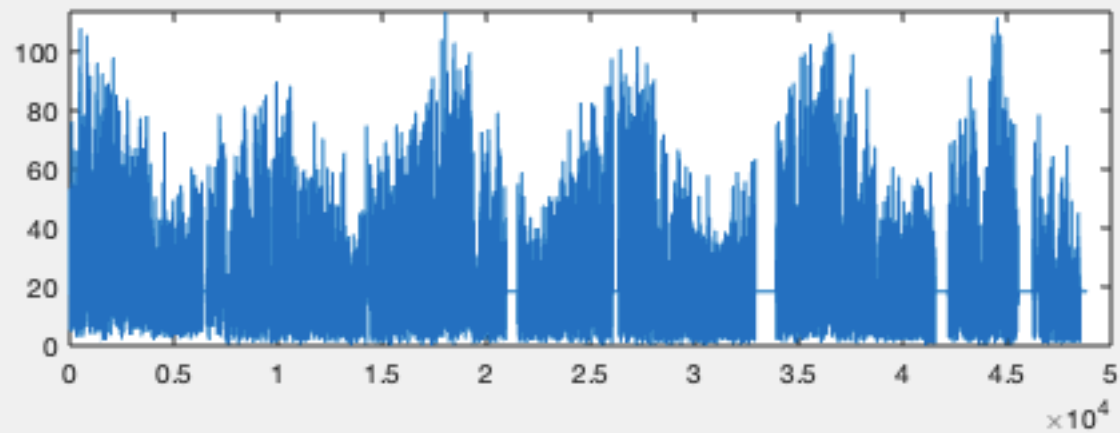
NILU_EIgeseter_p_m10



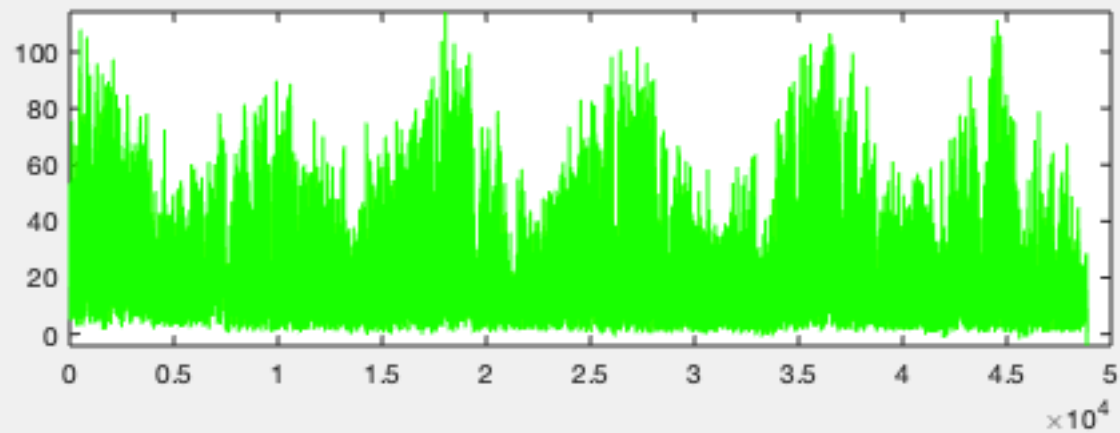
NILU_EIgeseter_p_m10



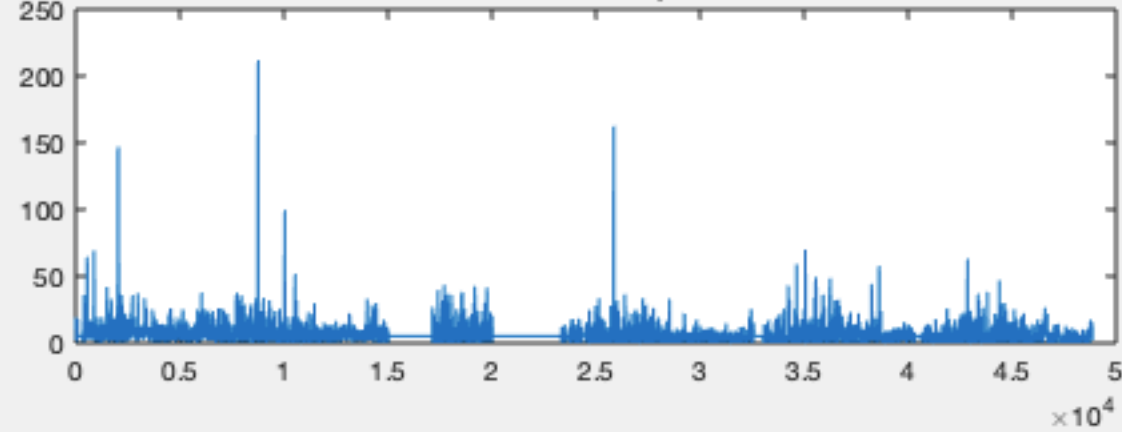
NILU_EIgeseter_n_o2



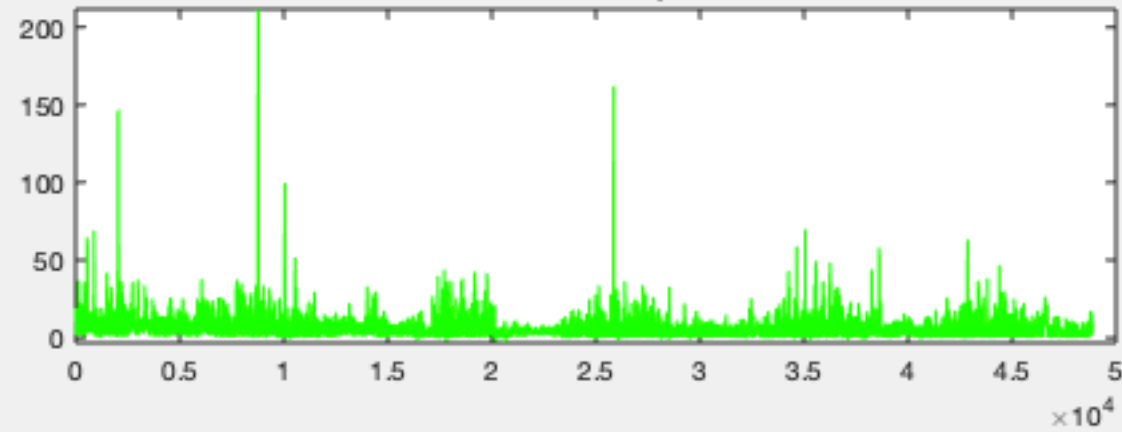
NILU_EIgeseter_n_o2

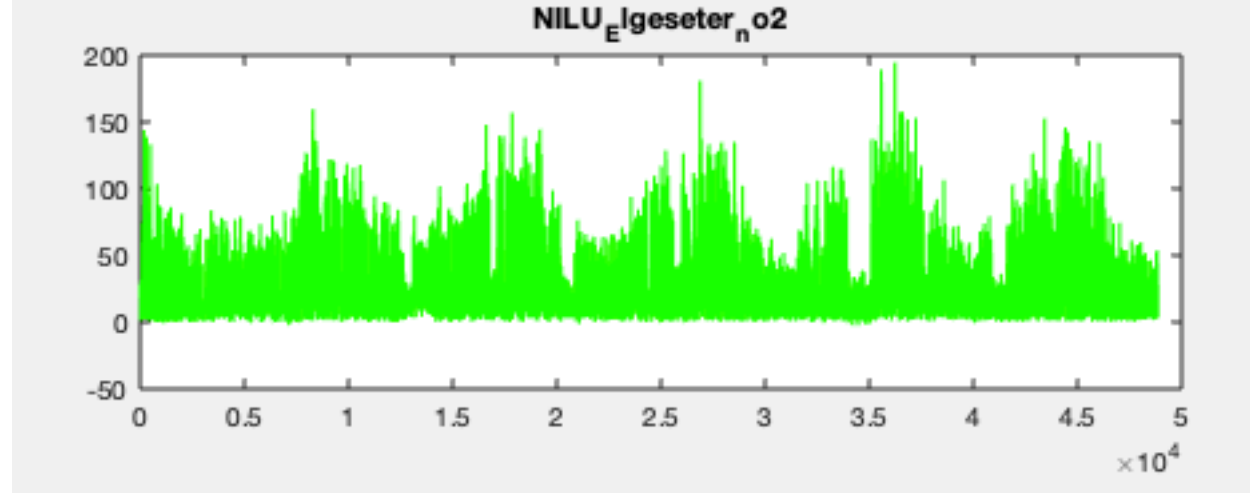
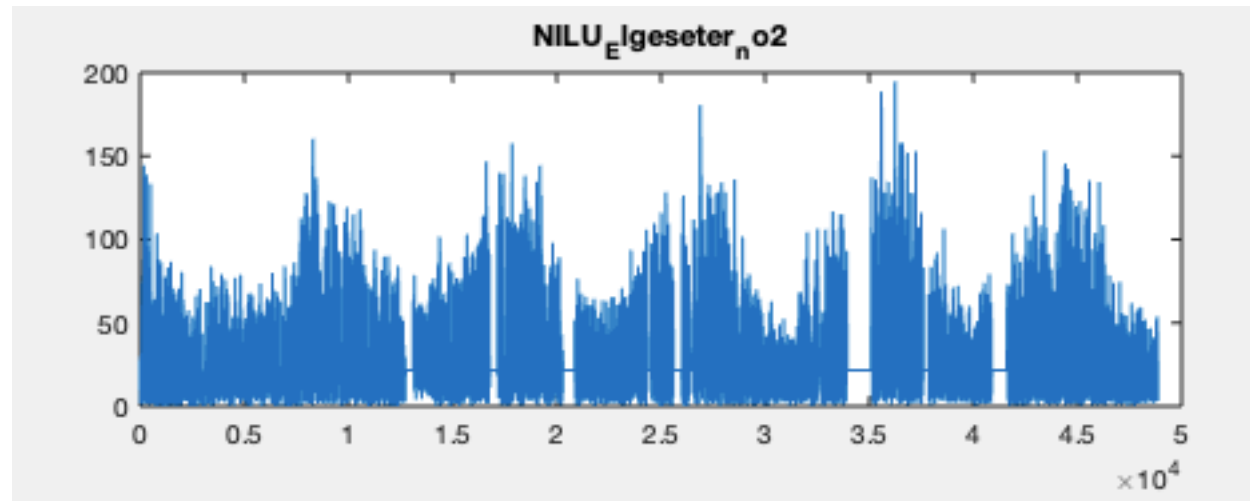
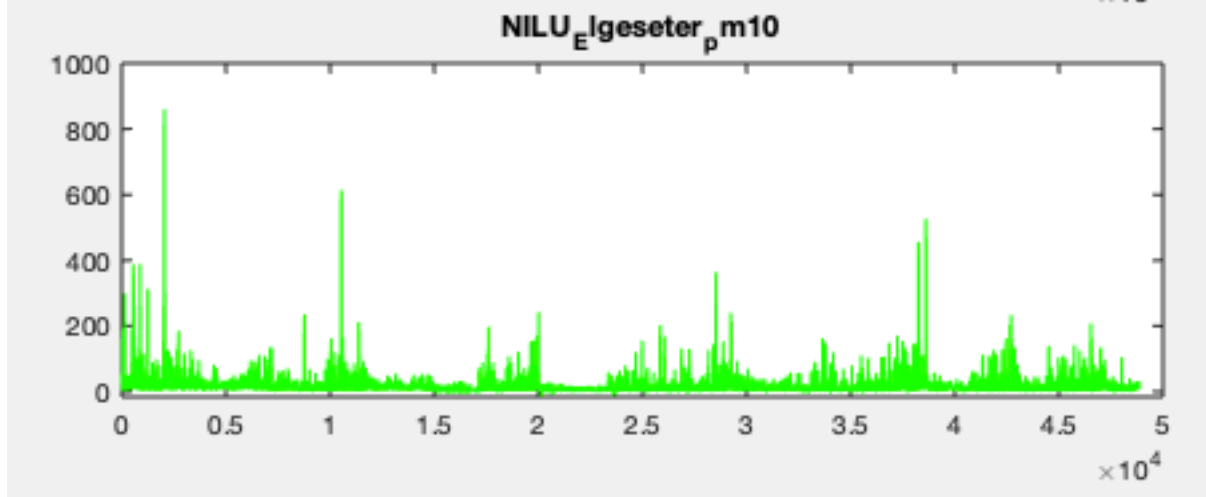
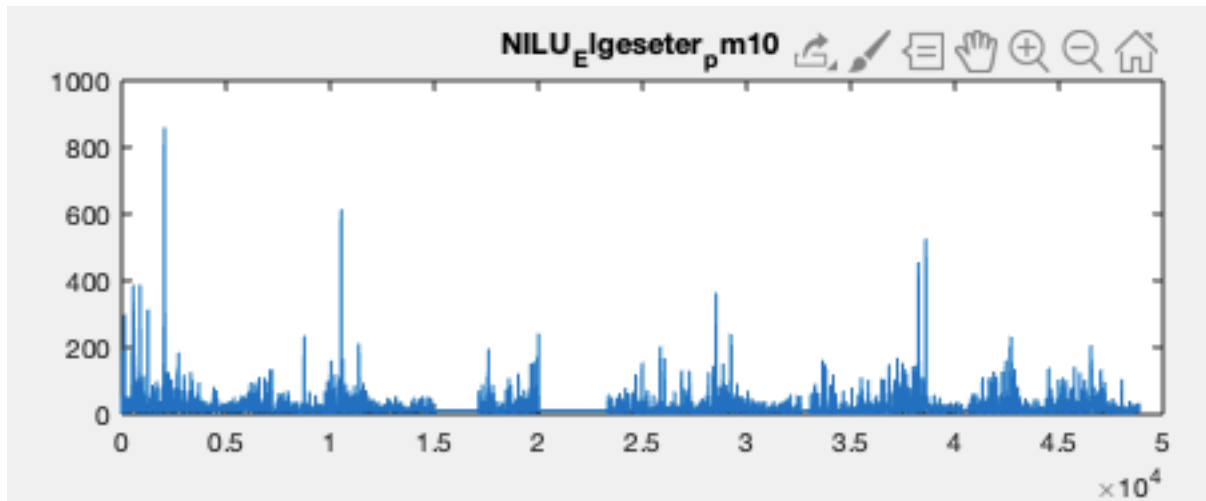


NILU_EIgeseter_p_m25

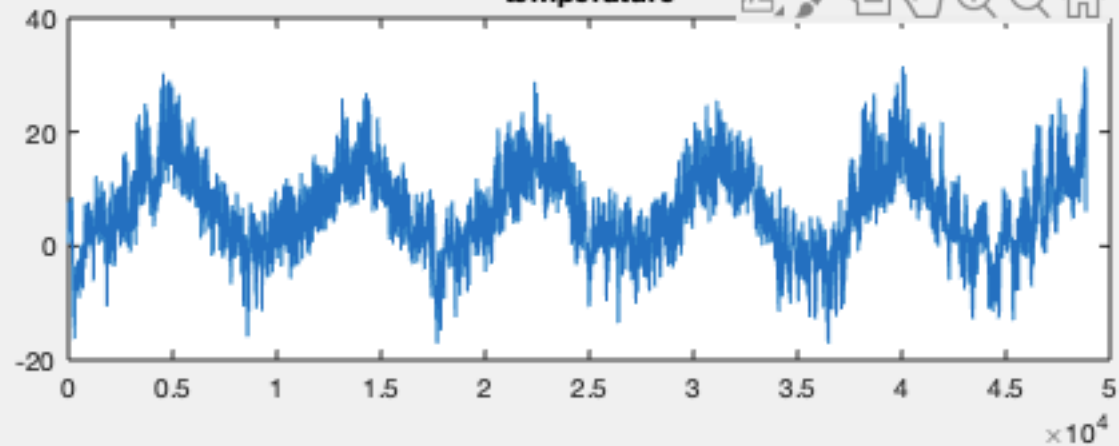


NILU_EIgeseter_p_m25

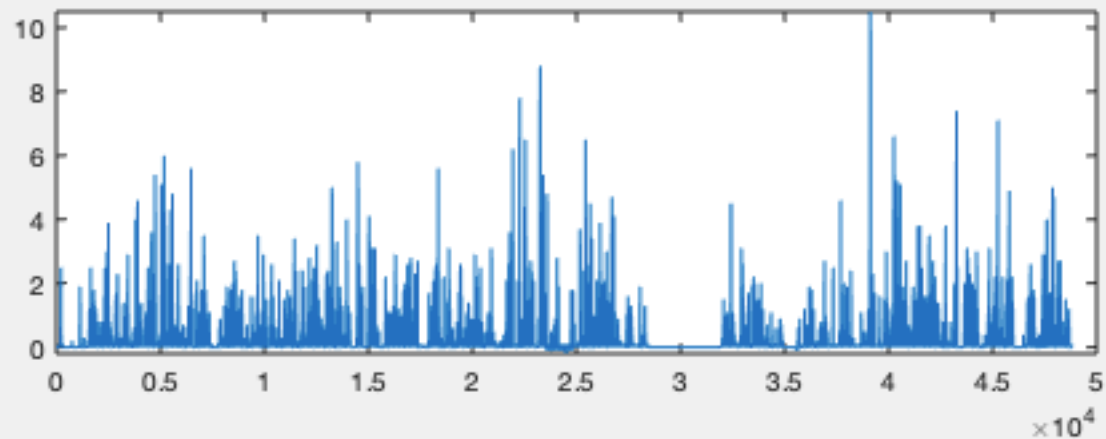




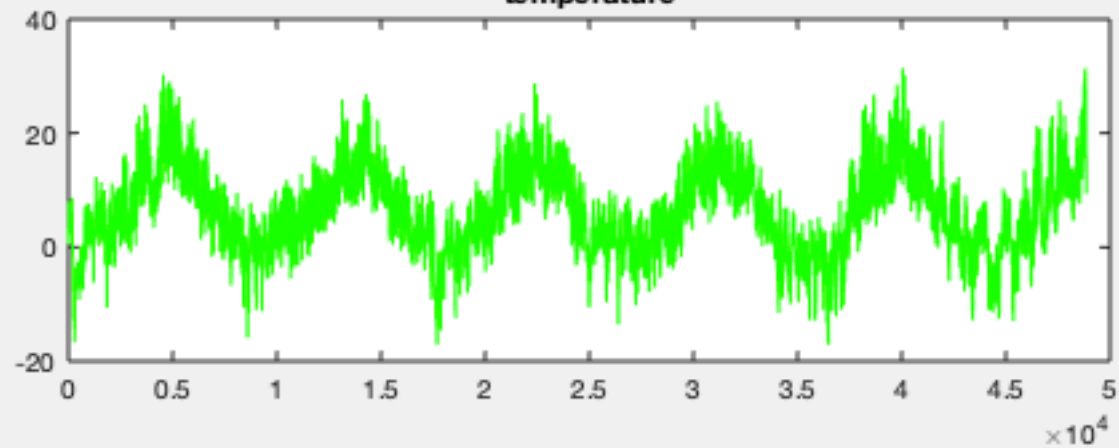
temperature



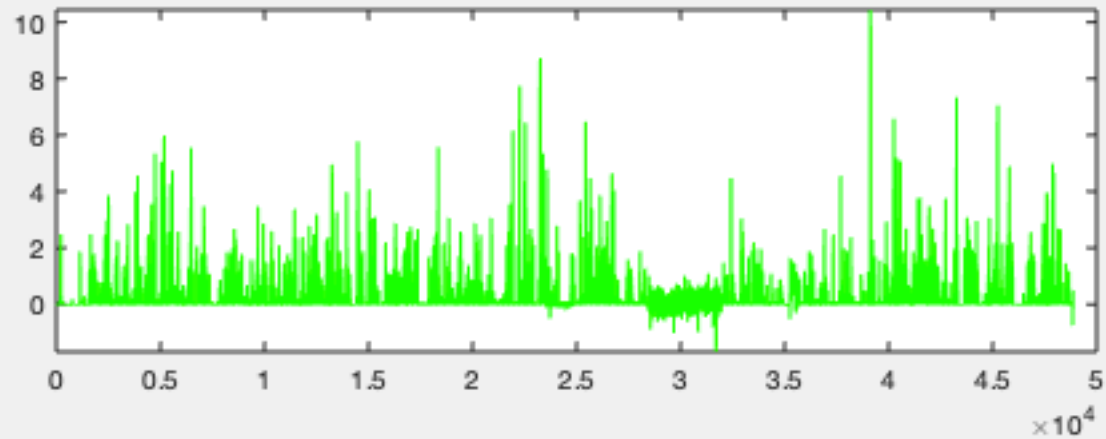
rain

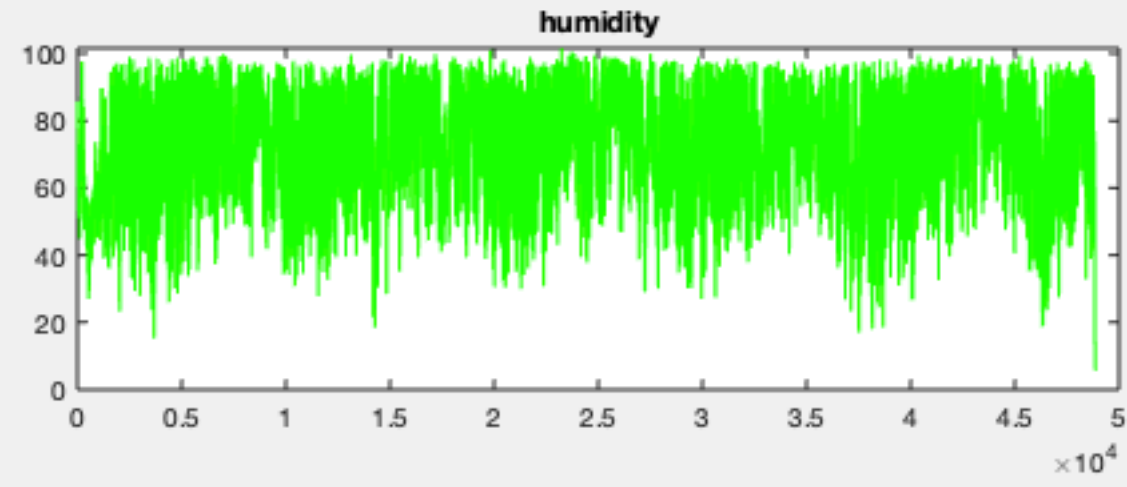
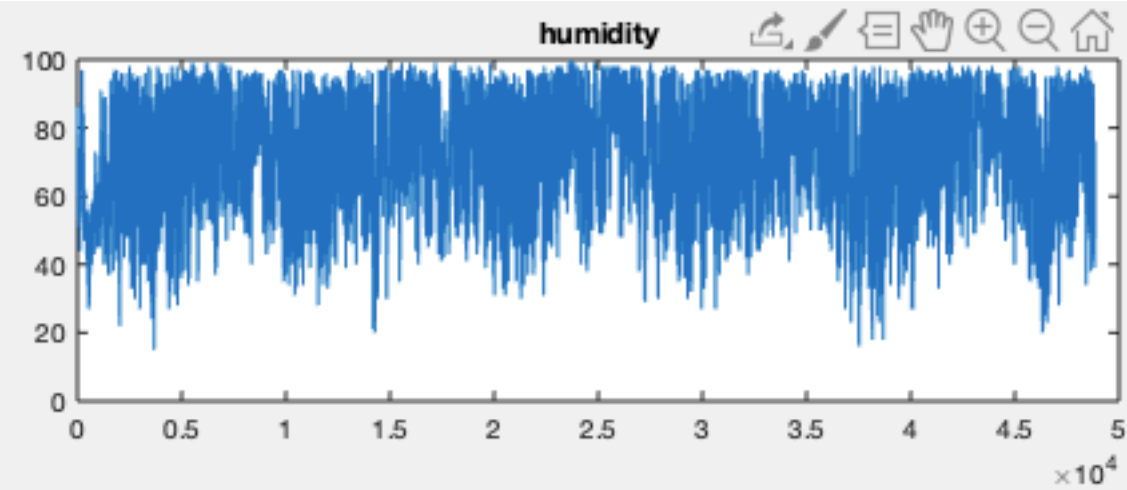
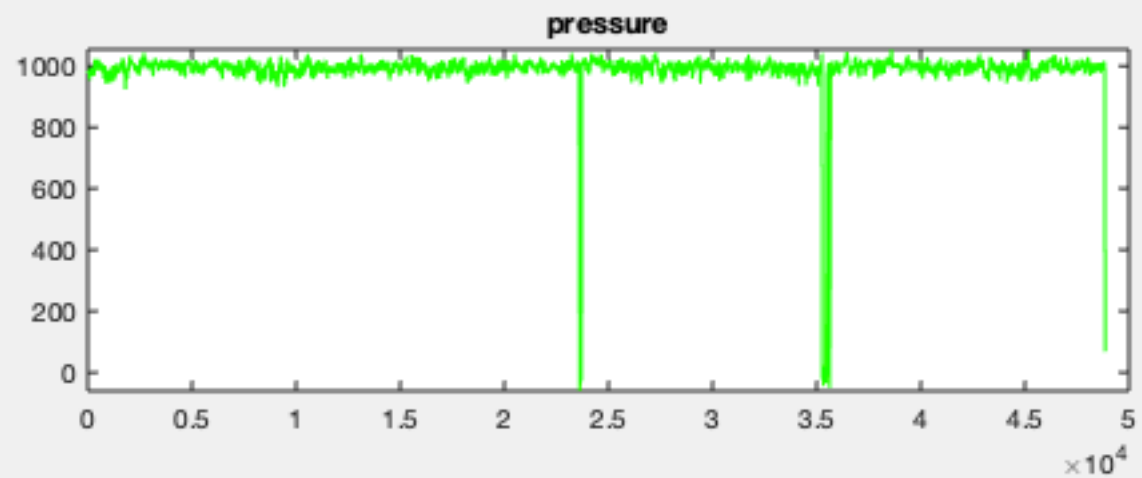
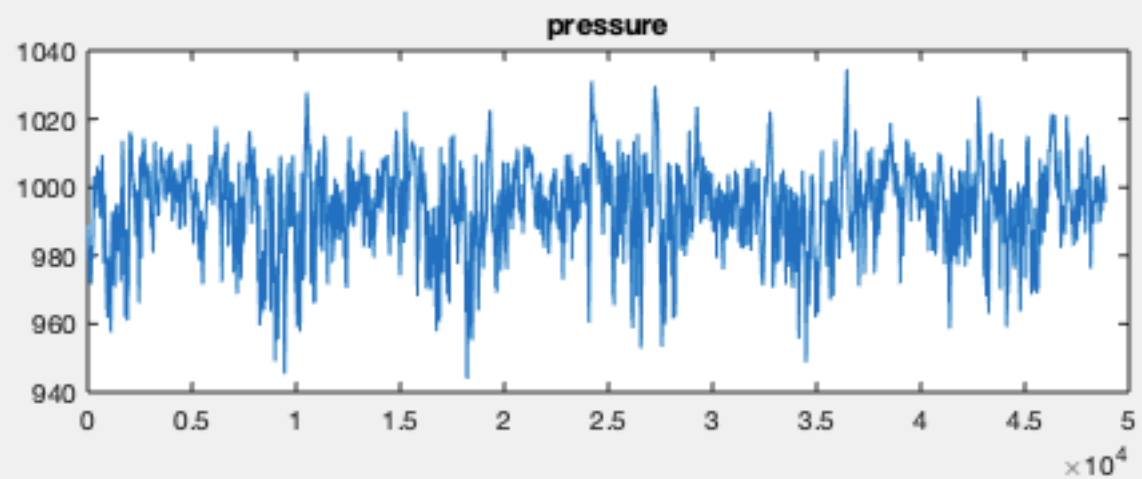


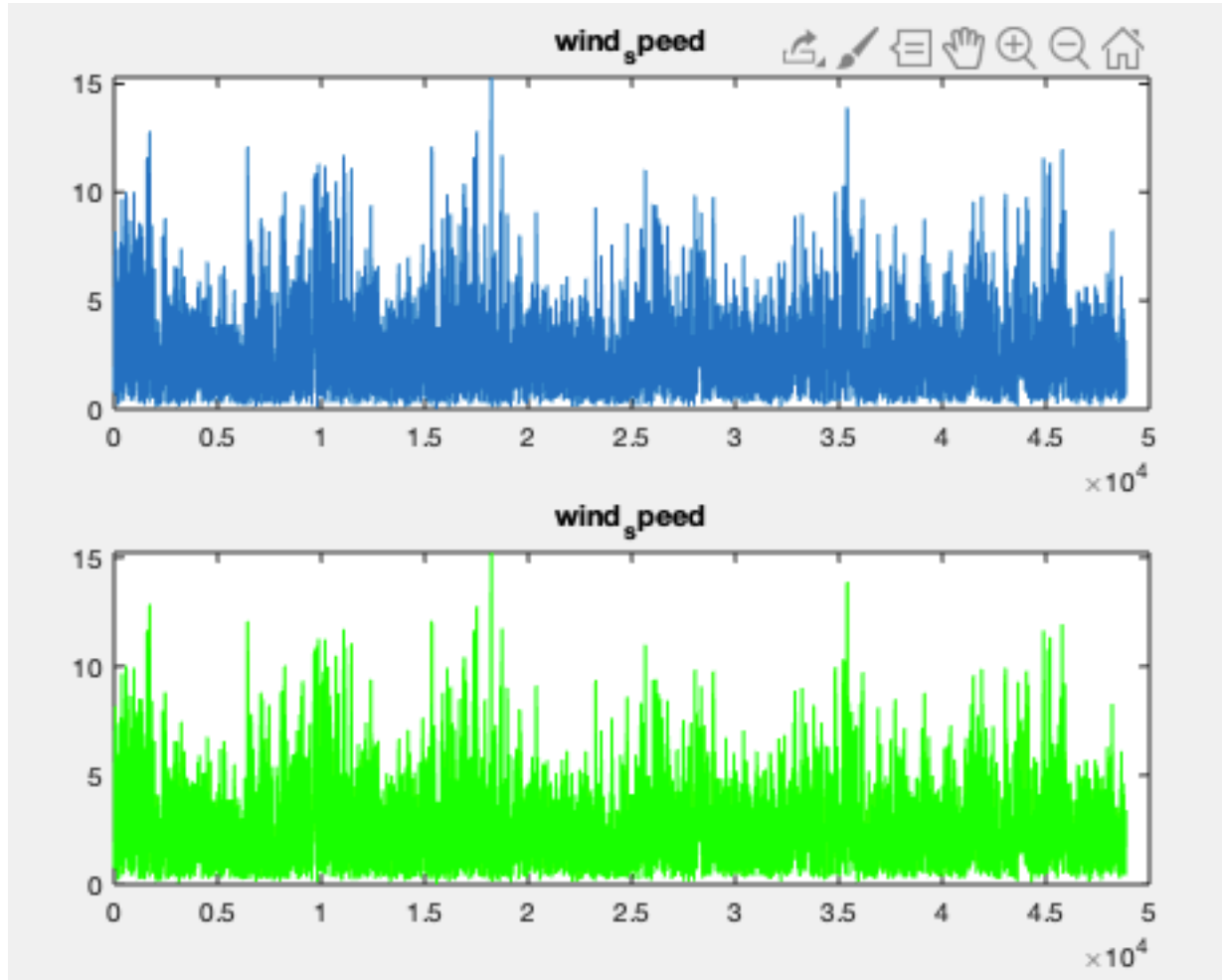
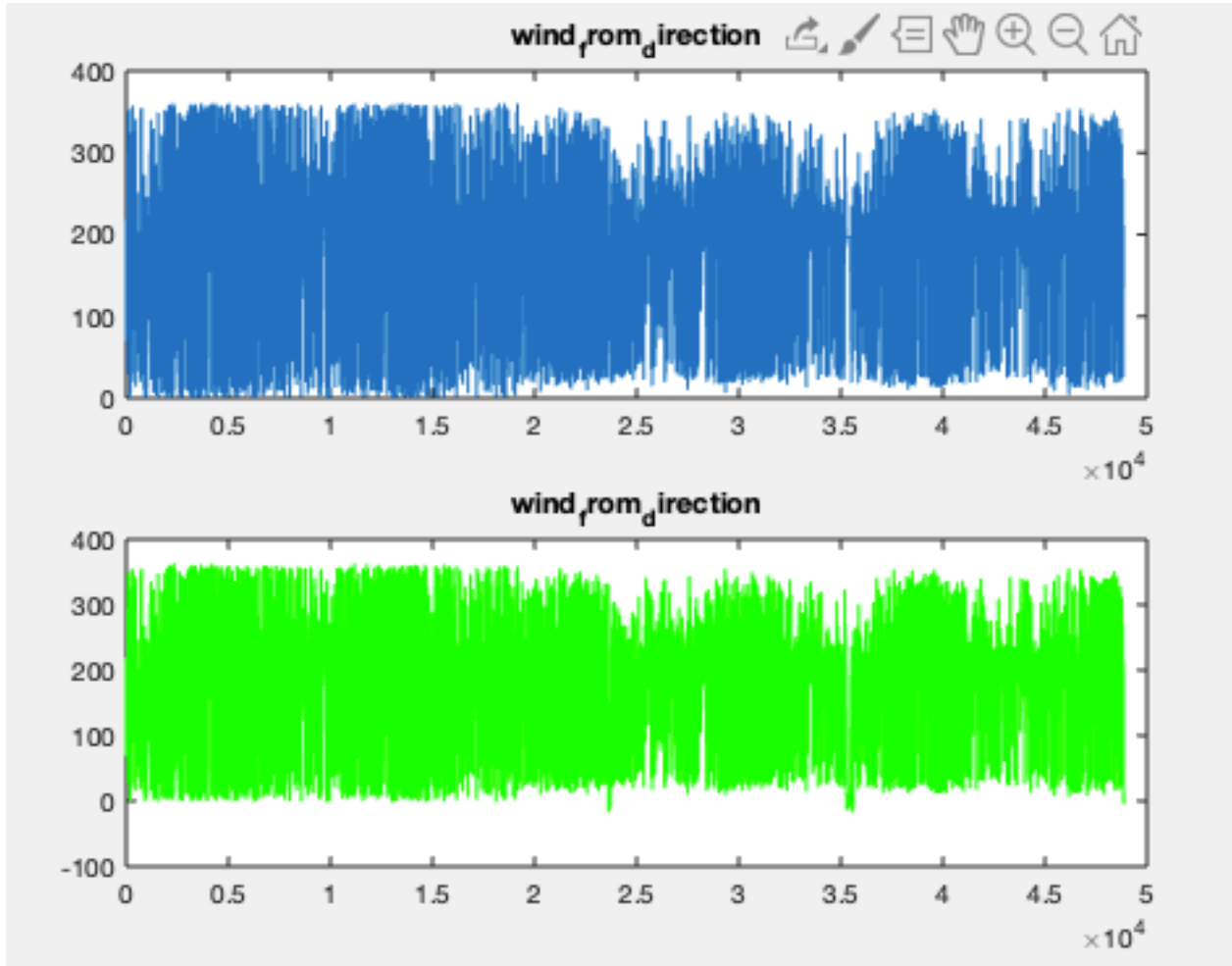
temperature



rain







Traffic

- In file traffic.xlsx we stored traffic data for 2019. It is still raw, will be distilled soon

Buses Trondheim

<https://github.com/mpolden/atbapi>

<https://github.com/norrs/busbuddy>

City Bikes Trondheim

<https://trondheimbysykkel.no/en/open-data>

Turn off for: English

<http://flow.dai-labor.de/datasets/>

Information collected during the proof-of-concept installation at Ernst-Reuter-Platz can be downloaded as an archive.

metadata_json.txt (sensor labels, positions, etc.)

sample.zip (data sample over approx. one day)

2018-02-20_combined.zip (Collection of Traffic, People, Wifi Count and Parking Space Survey Jan/Feb 2018)

traffic_data_1.pickle.tar.gz (the same data as above (traffic only!) as .pickle)

2018-07-11_combined (collection of traffic, person, Wifi and parking space counts until July 2018)