



X5.96 Bridging Automatic and Manual Pollen Monitoring: A Path Towards Homogenized Long-Term Time Series

Maria Lbadaoui-Darvas, Laura Sukienė,
José Oteros, Evgeny Kadantsev, Ingrida
Saulienė and Regula Gehrig

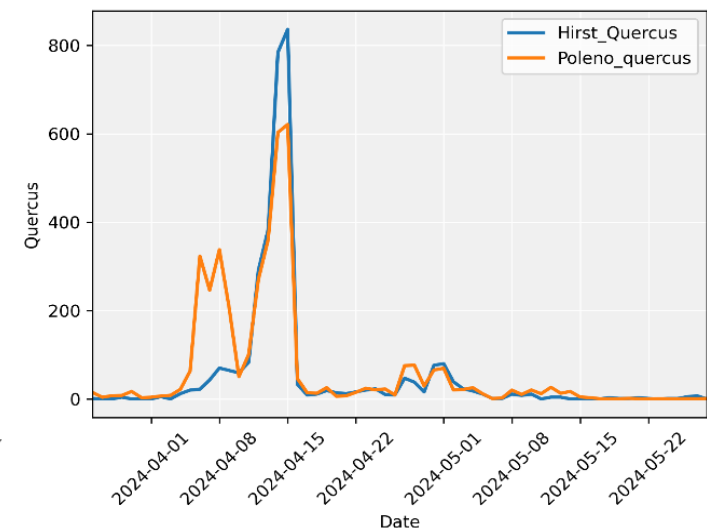
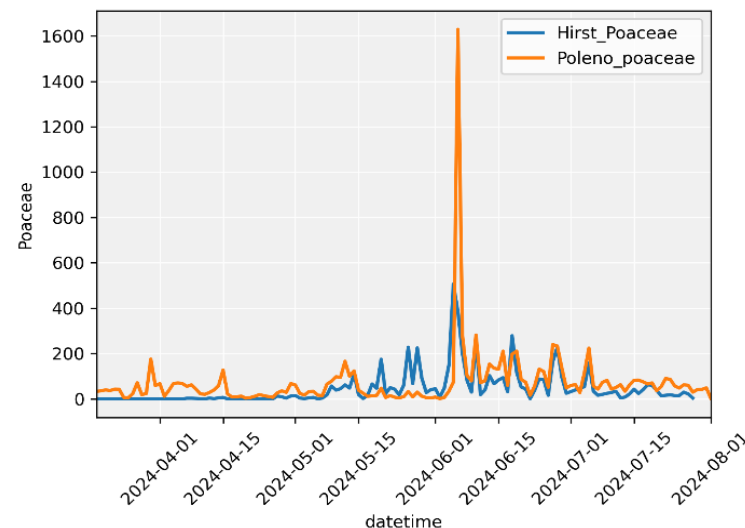
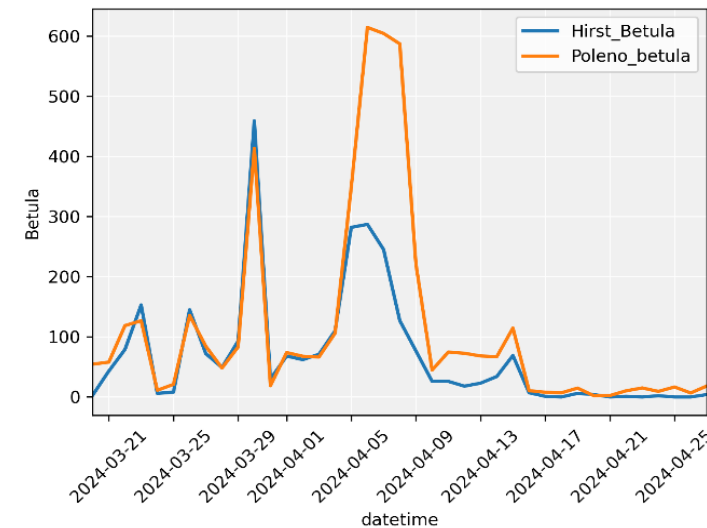
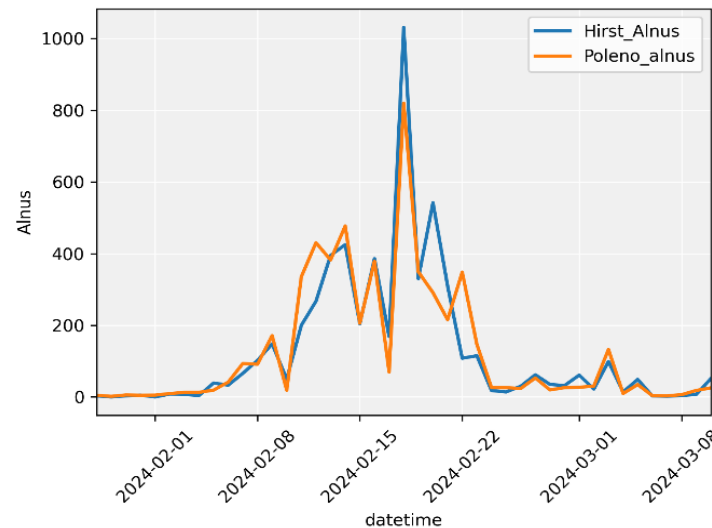
30.4.2025



Selected target taxa, seasons and algorithm

Basel - 2024

- **Target taxa: Alnus, Betula, Poaceae and Quercus**
 - **Selection criteria for the target taxa:**
 - Reliable poleno and hirst datasets
 - Relevant for all the regions in the project
 - **Selection criteria for the season:**
 - Lower threshold: earliest date at which the 2.5%SPI is reached for any station per country per taxon
 - Upper threshold: latest date at which 97.5%SPI is reached for any station per country per taxon
 - Modifications for alnus, poaceae
 - **Selection criteria for the recognition algorithm:**
 - Commonly available for each participant (MCH 2022 no supervisor)
 - Possibility to re-train the transfer operator on reanalysed data once a new model is ready



Features: meteorological data

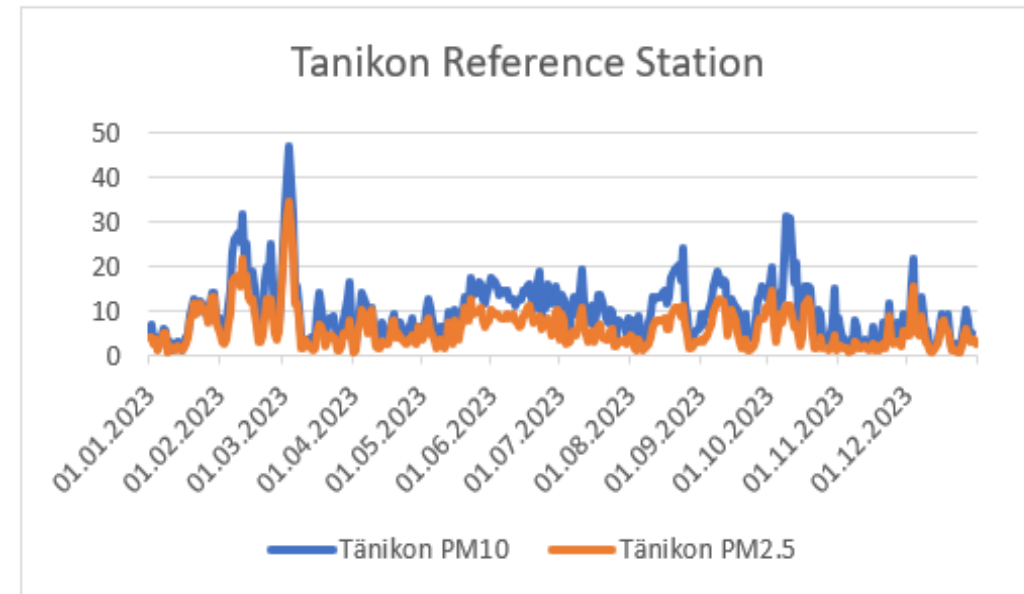
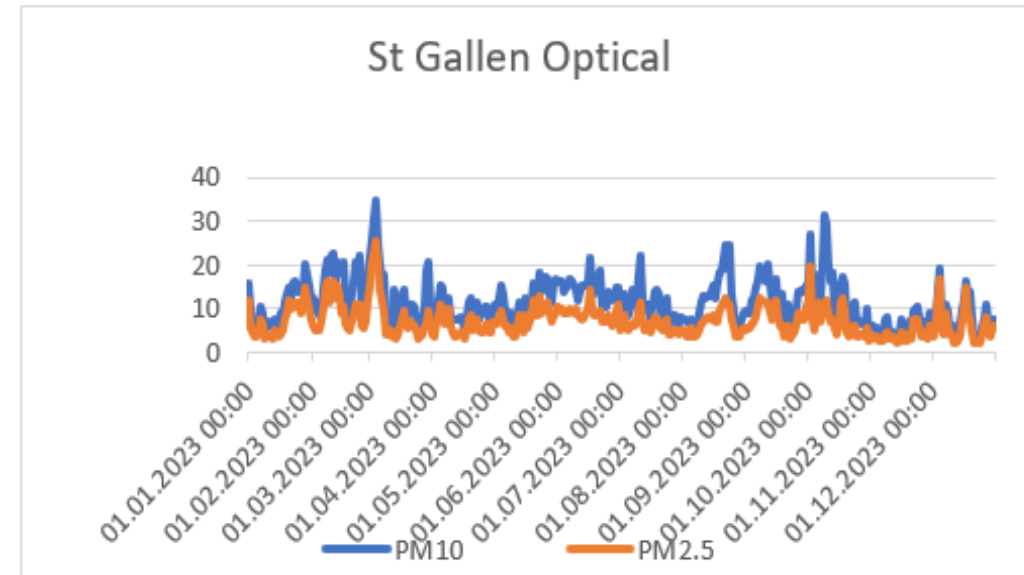
- **Daily data codes and variable names in datafile:**
 - Daily mean temperature (C): tre200d0 (mean_temp)
 - Daily minimum temperature: tre200dn (min_temp)
 - Daily maximum temperature: tre200dx (max_temp)
 - Daily sum precipitation (mm): rka150d0 (sum_prec)
 - Daily mean scalar wind speed (m/s): fkl010d0 (mean_wind_vel)
 - Daily max windspeed: fkl010d1 (max_wind_vel)
 - Daily mean wind direction: dkl010d0 (mean_wind_dir)
 - Daily mean RH (%) : ure200d0 (mean_RH)
 - Daily minimum RH : ure200dn (min_RH)
 - Daily maximum RH: ure200dx (max_RH)
 - Daily duration of sunshine (min): sre00d0 (tot_sun)
 - Daily mean pressure at the station (): prestad0 (mean_pres)
- **Hourly data codes and variable names in the datafile:**
 - Hourly mean temperature (C): tre200h0 (hmean_temp)
 - Hourly sum precipitation (mm): rre150h0 (hsum_prec)
 - Hourly mean scalar wind speed (m/s): fkl010h0 (hmean_wind_vel)
 - Hourly max scalar windspeed: fkl010h1 (hmax_wind_vel)
 - Hourly mean wind direction: dkl010h0 (hmean_wind_dir)
 - Hourly mean RH (%) : ure200h0 (hmean_RH)
 - Hourly duration of sunshine (min): sre00h0 (htot_sun)
 - Hourly mean pressure at the station (): prestah0 (hmean_pres)

6 maximum values
averaged to derive daily
aggregated values



Features: air pollution data

- PM data:
 - Daily data are used (hourly are not available for each station)
 - Data availability:
<https://www.bafu.admin.ch/bafu/en/home/topics/air/state/data/data-query-nabel.html> (Basel, Lugano, Locarno, Payerne, Zurich),
<https://www.ostluft.ch/index.php?id=datenabfragen> (St.Gallen/Buchs), Neuchatel, Luzern and Geneva are from cantonal resources (inluft)
 - Sites with no reference data:
 - Use data from the closest reference monitor
 - Exception: large difference in land use or altitude, then local OPC data is preferred



Dataset format

A	B	C	D	E	F	G	H	I	J	K
STA	datetime	Alnus	Betula	Quercus	Poaceae	alnus	betula	carpinus	corylus	cupressus
Basel	2023-01-01	NaN	NaN	NaN	NaN	16.94444444	6.669444444	0.3638888889	35.72048611	4.122222222
Basel	2023-01-02	NaN	NaN	NaN	NaN	9.722222222	7.384027778	0.9097222222	51.4375	4.416666667
Basel	2023-01-03	NaN	NaN	NaN	NaN	0.2777777778	2.381944444	0	0	2.944444444
Basel	2023-01-04	1	0	0	1	3.333333333	4.2875	0.7277777778	9.144444444	38.27777778
Basel	2023-01-05	1	0	0	0	2.222222222	2.381944444	2.729166667	6.286805556	3.533333333
Basel	2023-01-06	1	0	0	0	0.8333333333	0.7145833333	0.3638888889	0.5715277778	0.5888888889
Basel	2023-01-07	7	0	0	0	12.08333333	10.00416667	2.183333333	57.43854167	6.183333333
Basel	2023-01-08	0	0	0	0	2.222222222	1.905555556	0.7277777778	9.715972222	3.827777778
Basel	2023-01-09	4	0	0	0	1.805555556	2.14375	1.455555556	7.429861111	2.797222222
Basel	2023-01-10	0	0	0	0	2.916666667	3.334722222	0.5458333333	7.144097222	0.7361111111
Basel	2023-01-11	1	0	0	0	2.361111111	3.096527778	0.7277777778	9.144444444	1.325

Site

Time

Target taxa (Hirst)

Poleno taxa

Feature space
36 features /
target taxon

V	W	X	Y	Z	AA	AB	AC	AD	AE	AF
PM2.5	PM10	mean_wind_dir	mean_wind_vel	max_wind_vel	mean_pres	sum_prec	tot_sun	mean_temp	min_temp	max_temp
20.5	21.9	97	2.8	9.5	983.8	0	354	11.4	7.9	15.4
6	9.4	129	2.1	14.1	985.5	3.1	118	10.1	5.7	14.8
4.7	5.8	149	0.8	3.9	992.8	0.1	14	6.9	2.6	13.5
6.9	8	120	1.7	9.2	991.6	0	0	6.4	2	12.1
1.6	2.1	243	2.5	11.5	988.4	0	0	11	6.5	14.2
2.4	3.1	207	1.5	4.9	986.5	0	6	8.5	4.7	13.8
8.8	9.2	102	2.5	8.3	979.4	0	396	7.8	4.6	13.2
10	10	104	2.2	7.4	970.1	10.7	0	6.4	4.4	12.9

Air pollution

Meteorological data



Funded by
the European Union

SYLVA^{air}
Observing life in air



Exploratory Data Analysis

30.4.2025



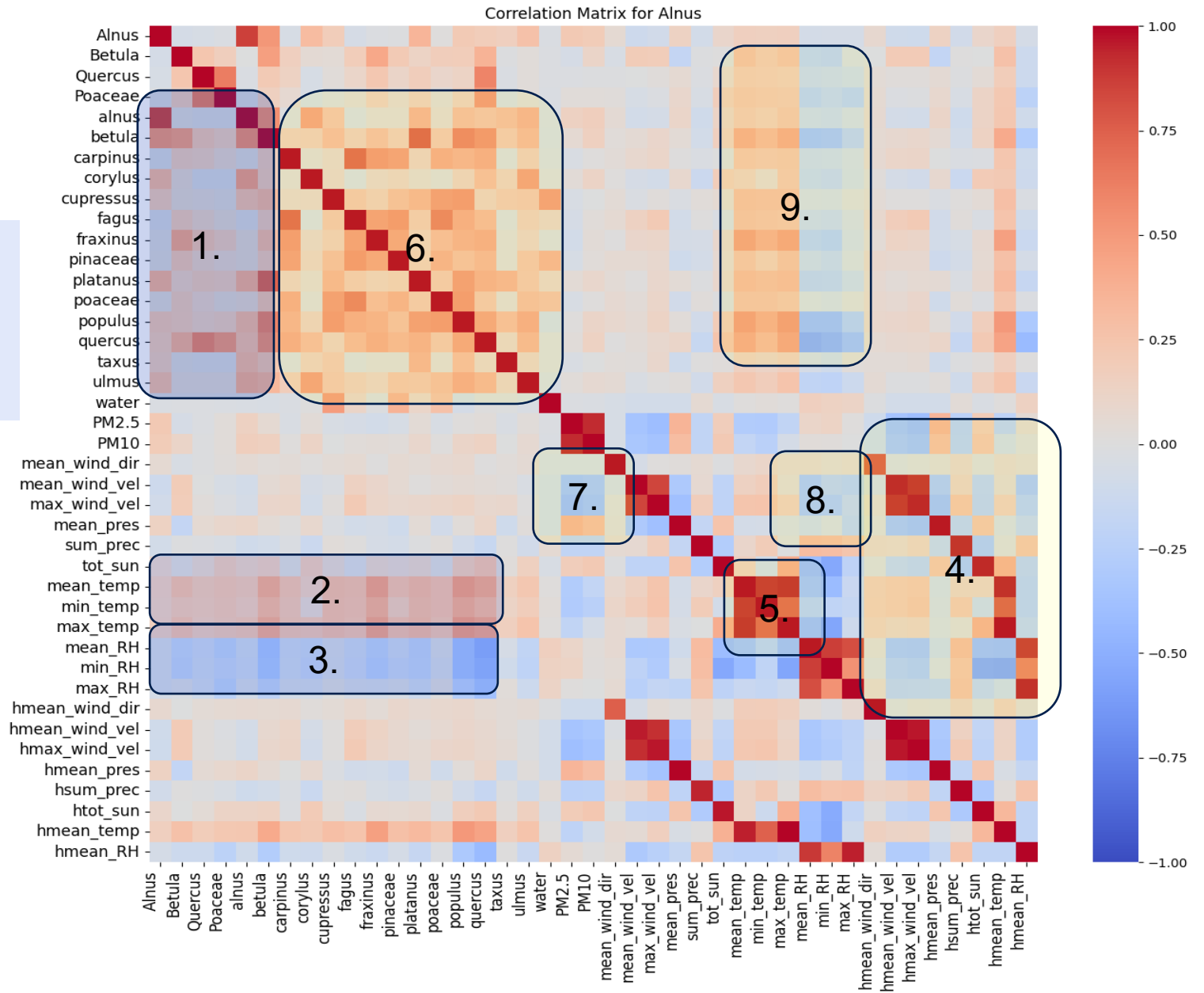
Pearson correlations

Linear correlations of the target taxa:

1. Poleno taxa
2. Temperature
3. Relative humidity

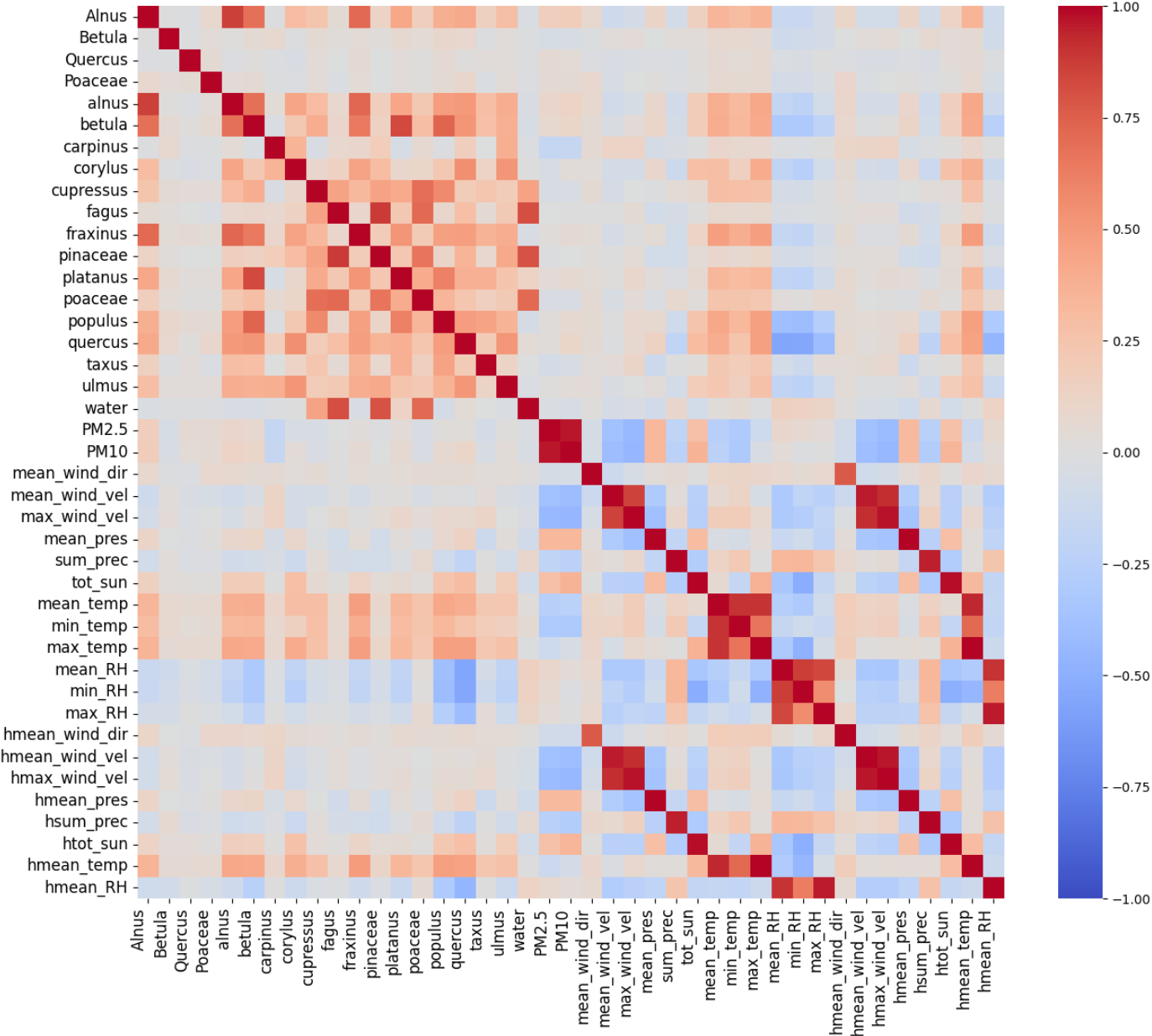
Interference between features:

4. Hourly and daily data
5. Temperature and humidity
6. Poleno taxa
7. Wind and PM
8. Wind and RH
9. Poleno taxa and temperature/RH

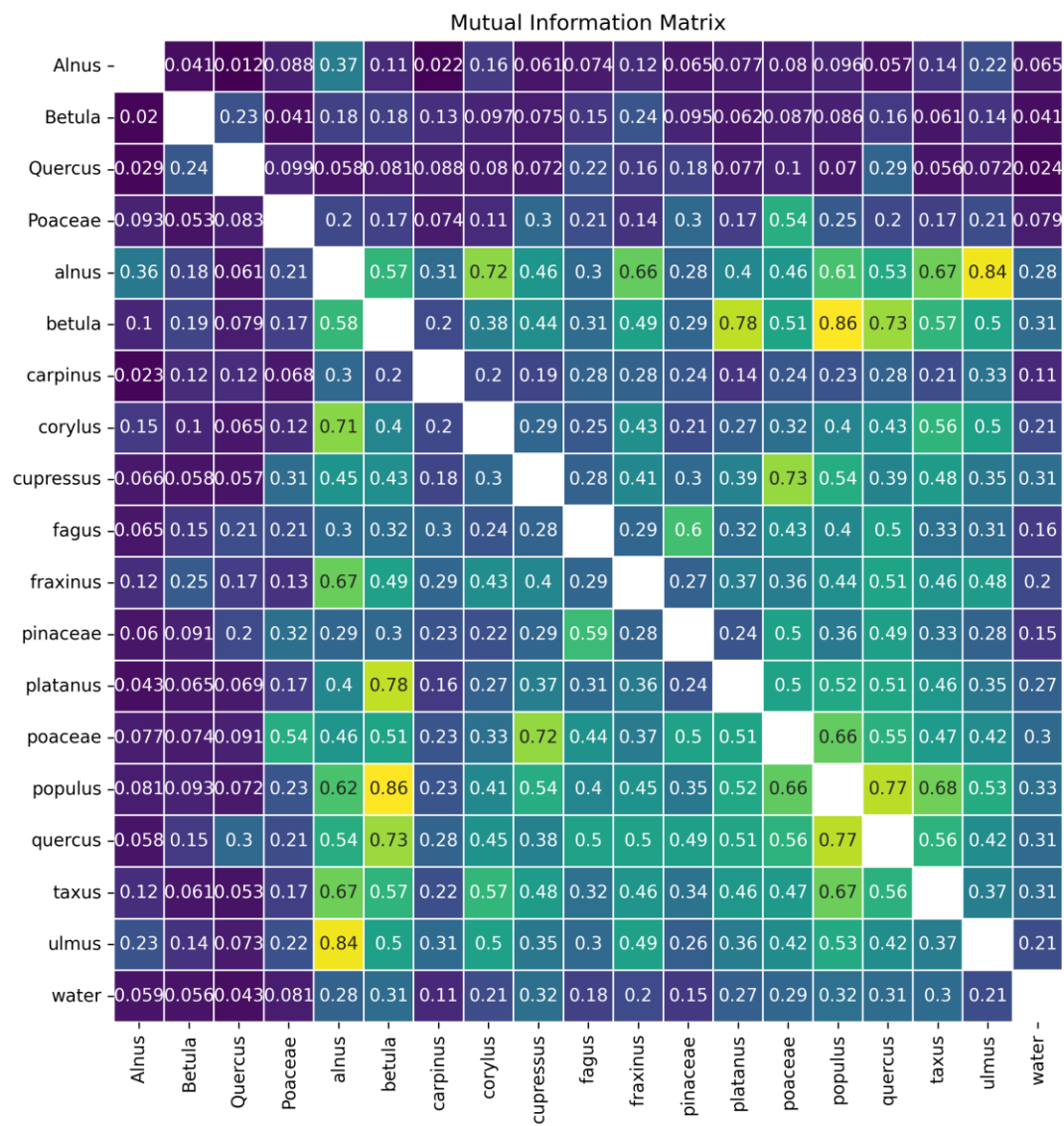


Spearman correlations

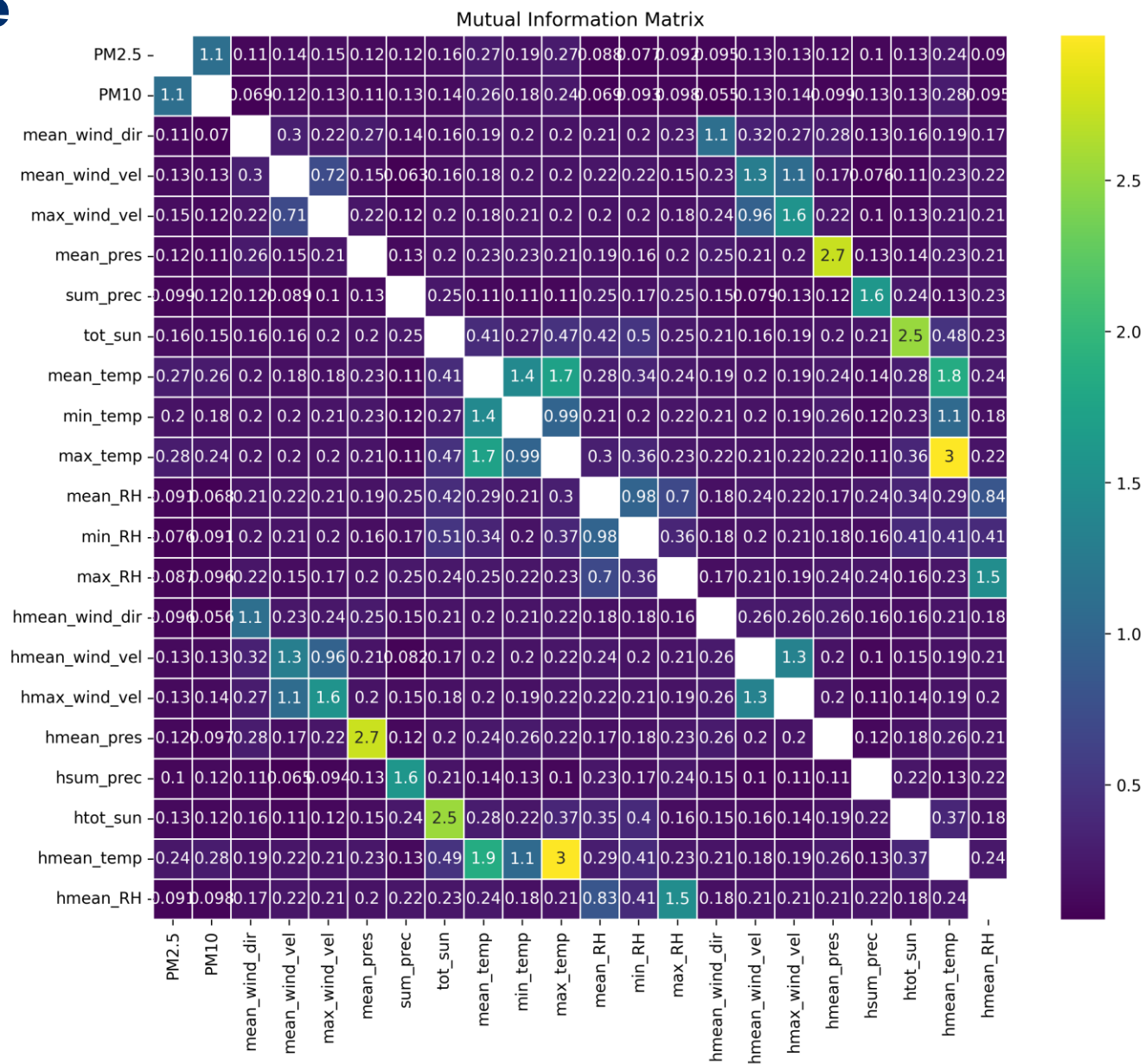
Qualitatively similar to Pearson



Mutual information: feature space



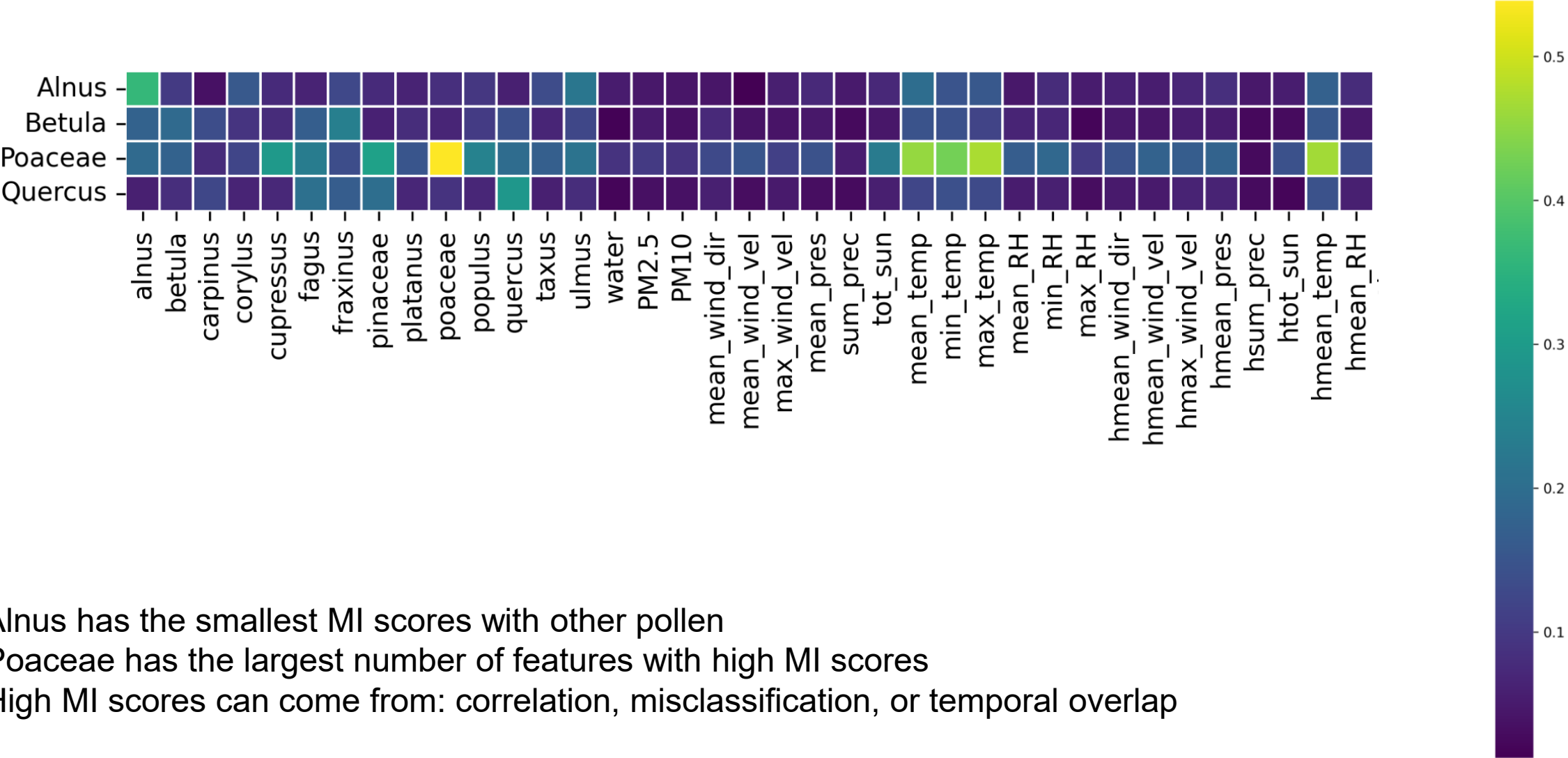
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1



2.5
2.0
1.5
1.0
0.5

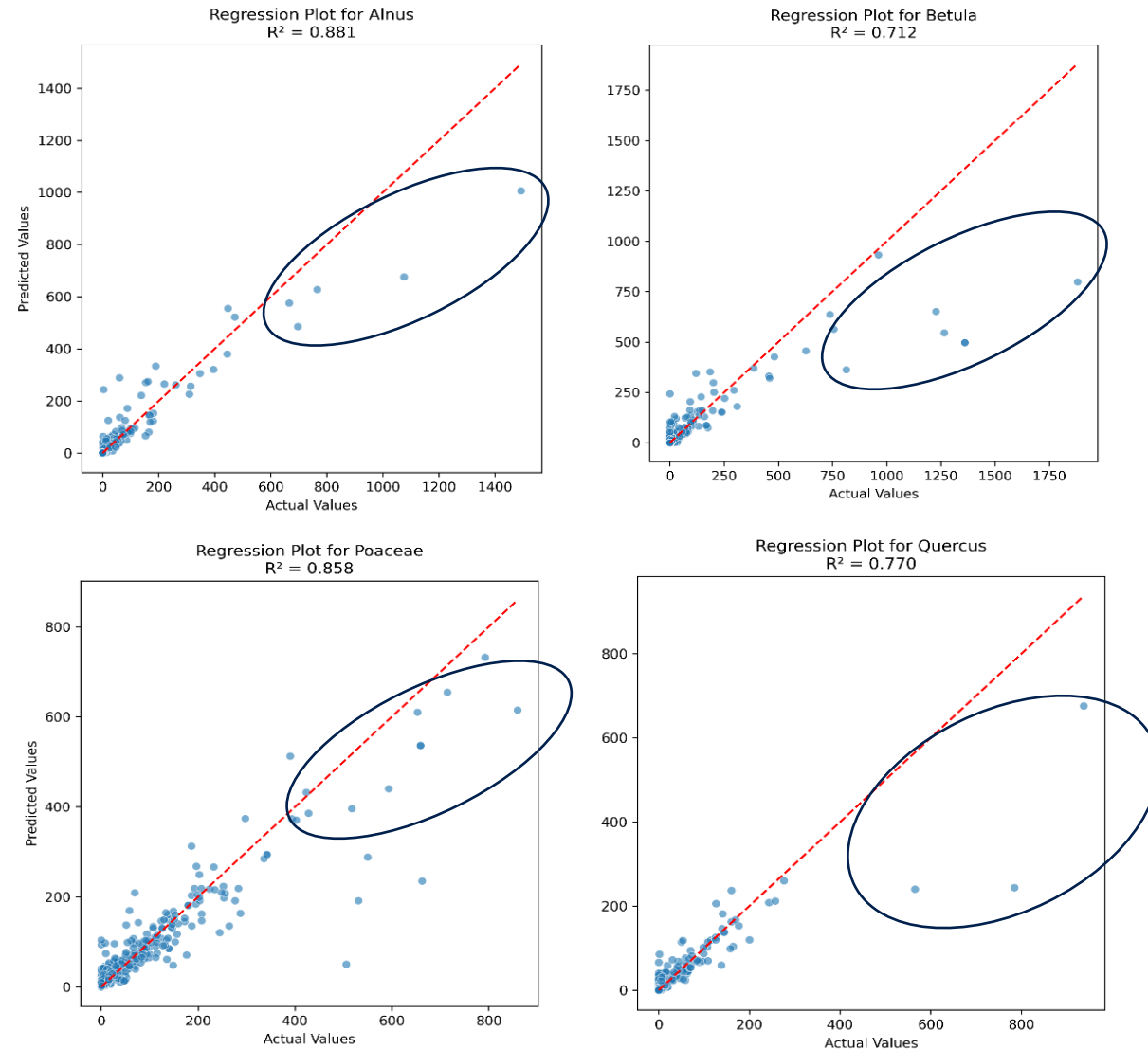


Mutual information: Target vs. feature



- Alnus has the smallest MI scores with other pollen
- Poaceae has the largest number of features with high MI scores
- High MI scores can come from: correlation, misclassification, or temporal overlap

Random Forest regression plots without data augmentation



Circled areas show the underestimation of high manual pollen concentrations by the random forest regression without weighted data augmentation

Statistical metrics of the regression methods tested

Taxon	Model	Validation	R^2	RMSE
Alnus	gam	Internal	0.07	3
Alnus	gam	External	0.07	2.36
Betula	gam	Internal	0.01	1.1
Betula	gam	External	0	0.93
Poaceae	gam	Internal	0.05	2.8
Poaceae	gam	External	0.05	2.72
Quercus	gam	Internal	0.15	0.54
Quercus	gam	External	0	0.36
Alnus	gbm	External	0.61	1.33
Alnus	gbm	Internal	0.37	2.47
Betula	gbm	Internal	0.05	1.08
Betula	gbm	External	0.01	0.92
Poaceae	gbm	Internal	0.07	2.76
Poaceae	gbm	External	0.05	2.72
Quercus	gbm	Internal	0.24	0.51
Quercus	gbm	External	0	0.42
Alnus	pls	Internal	0.54	2.12
Alnus	pls	External	0.12	2.02
Betula	pls	External	0.01	0.93
Betula	pls	Internal	0	1.1
Poaceae	pls	External	0.02	2.76
Poaceae	pls	Internal	0.01	2.84
Quercus	pls	Internal	0.03	0.57
Quercus	pls	External	0	0.35
Alnus	rf	Internal	0.69	1.76
Alnus	rf	External	0.26	1.7
Betula	rf	Internal	0.38	0.9
Betula	rf	External	0.02	0.94
Poaceae	rf	Internal	0.2	2.58
Poaceae	rf	External	0.03	2.76
Quercus	rf	Internal	0.59	0.4
Quercus	rf	External	0.01	0.46
Alnus	svmRadial	Internal	0.58	2.55
Alnus	svmRadial	External	0.37	1.75
Betula	svmRadial	Internal	0.11	1.09
Betula	svmRadial	External	0	0.94
Poaceae	svmRadial	Internal	0.14	2.77
Poaceae	svmRadial	External	0.03	2.82
Quercus	svmRadial	Internal	0.25	0.52
Quercus	svmRadial	External	0	0.38