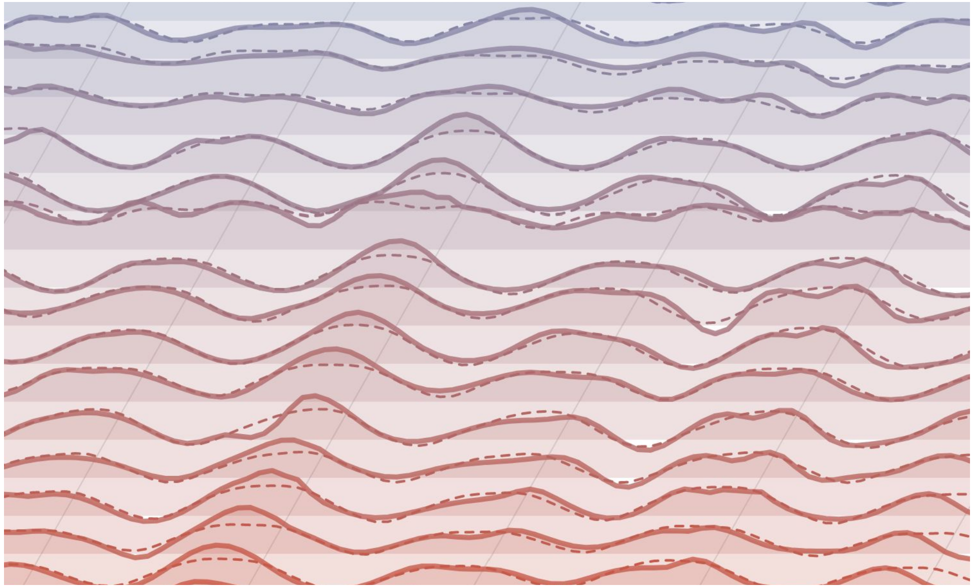
Investigating air quality in Amsterdam during the 2020 COVID-19 lockdowns

Presentation by **Marnix Hamelberg** about his research internship at

Geodan



Supervised by **Tom van Tilburg**

The research internship is part of the Master's Geo-Information Science curriculum at Wageningen University & Research supervised by **Jan Clevers**

Relevant links:

Source code + report:
https://github.com/MarnixHamelberg/geodan

Blog: https://www.cartoma.info/blog

Introduction

Introduction | Keywords

Air quality

Open data

Predictions (machine learning)

Local scale

Road traffic / weather data

Variable importance

COVID-19 lockdowns

Preprocessing

Trend analysis

Differences?

Linear (long / short term) correlations

New insights?

Introduction | Context

Air quality -> **Air pollutants** -> **Nitrogen dioxide** (NO2)

Impacting **climate** & **health** + **Reflection of human behavior?**

Potential factors: **road traffic + weather** -> **their effects**

Open data available -> **Ground sensors** (+ **Satellites**)

Gain **new insights** -> Philosophy of *Geodan*

Local scale: **Amsterdam**

COVID-19 lockdowns -> **Unique opportunity** to understand the **human influence**

Currently **limited research**

Introduction | Context

Air quality -> **Air pollutants** -> **Nitrogen dioxide** (NO2)

Impacting **climate** & **health** + **Reflection of human behavior?**

Potential factors: **road traffic + weather** -> **their effects**

Open data available -> **Ground sensors** (+ **Satellites**)

Gain **new insights** -> Philosophy of *Geodan*

Local scale: **Amsterdam**

COVID-19 lockdowns -> **Unique opportunity** to understand the **human influence**

Currently **limited research**

Needs

- **Investigate** air quality & other factors on a local scale in relation to the COVID-19 lockdowns
- Demand for **additional use cases** and **derivatives** of available open data
- Applying **machine learning**

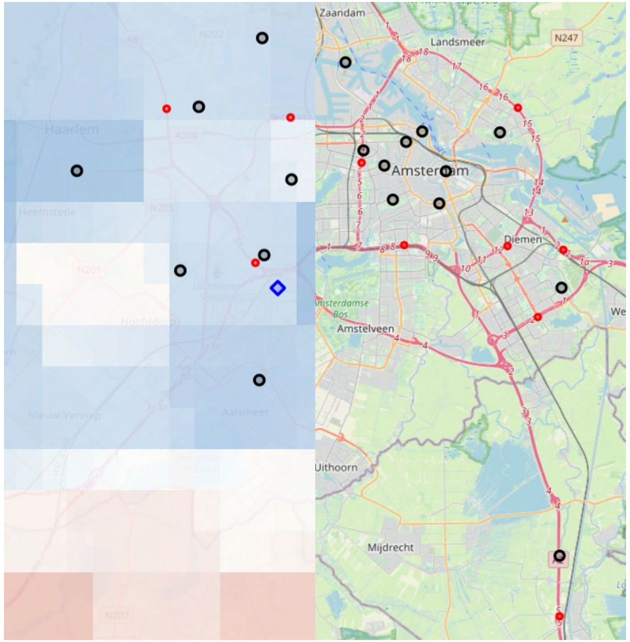
Introduction | Data

Variable of interest, i.e. response variable

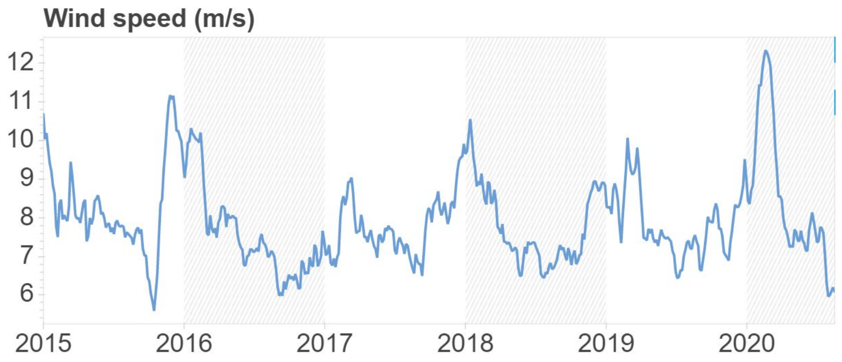
- **Landelijk Meetnet Luchtkwaliteit** (**LML**) -> NO2 data

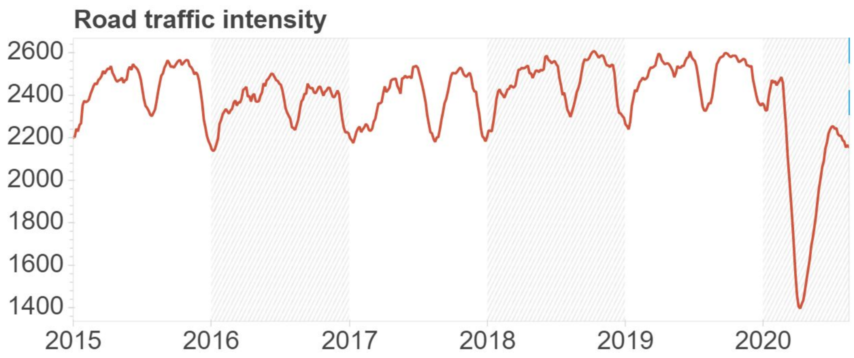
Factors that are (potentially) related to the variable of interest. i.e. explanatory variables

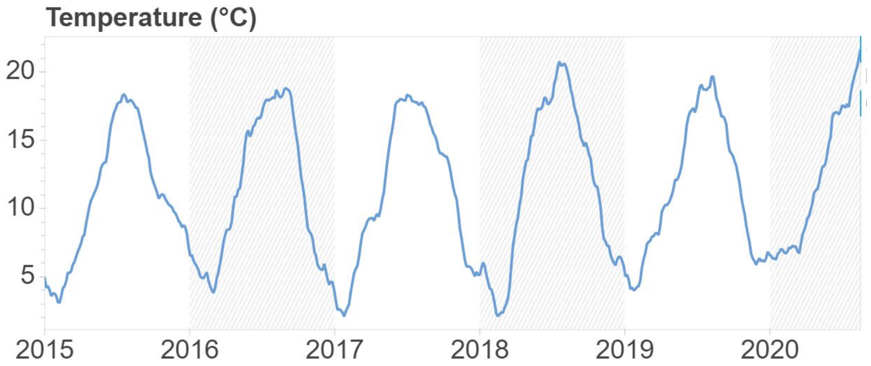
- **Sentinel-5 Precursor** satellite (**S5P**) -> NO2 data
- **Nationale Databank Wegverkeersgegevens** (**NDW**) -> Road traffic data
    - Road traffic intensity / speed
- **Koninklijk Nederlands Meteorologisch Instituut** (**KNMI**) -> Weather data
    - Wind speed / direction, temperature, etc.

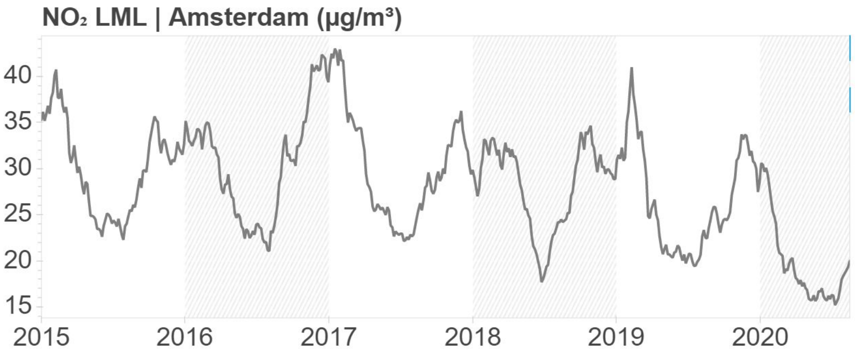


Introduction | Data preview

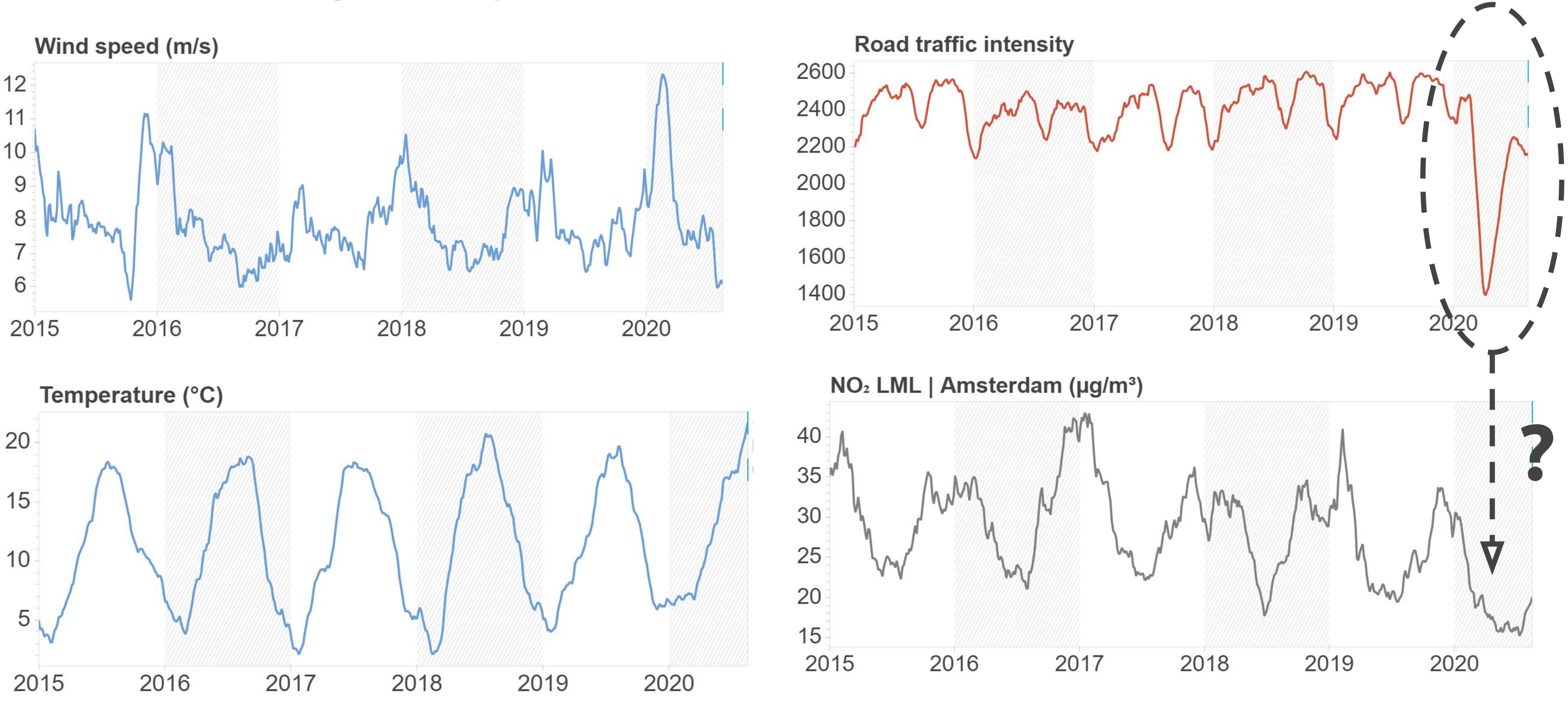








Introduction | Data preview



Introduction | Research questions / objectives

1) **Preparation**
   a) **Align** datasets **-** > Data **preprocessing**
2) **Exploration**
   a) Relation to NO2 -> **Linear relationships**
3) **Prediction**
   a) Predicting NO2 using **machine learning** -> **Prediction accuracy?**
   b) Variable **impact** on NO2 -> **Variable importance?**
   c) Effects **COVID-19 lockdowns**? -> Does 2020 data **disrupt** prediction accuracy and variable importance?
      i) What can this tell us about the **human influence**?
   d) Satellite data **contribution** [extra]
4) **Trend analysis**
   a) **Changes** during COVID-19 lockdowns?
      i) What more can this tell us about the human influence?

Methodology

Methodology | Preparation

Python coding environment -> ***Jupyter*** notebook

Geodan **servers** and ***Google Earth Engine***

Literature research -> **specifications**

Preprocess steps

- **Load** 2015-2020 data -> aggregate on **hourly intervals**
- **Filter outliers** and **smoothen** data
- **Log transform** exponential NO2 data
- **Convert** to sine and cosine (thanks for the tip Brian!)

Methodology | Preparation

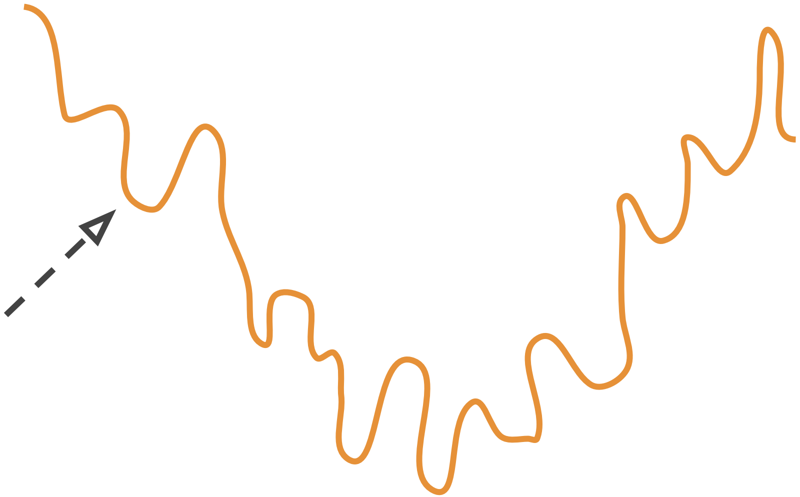
Python coding environment -> ***Jupyter*** notebook

Geodan **servers** and ***Google Earth Engine***

Literature research -> **specifications**

Preprocess steps

- **Load** 2015-2020 data -> aggregate on **hourly intervals**
- **Filter outliers** and **smoothen** data
- **Log transform** exponential NO2 data
- **Convert** to sine and cosine (thanks for the tip Brian!)



Methodology | Preparation

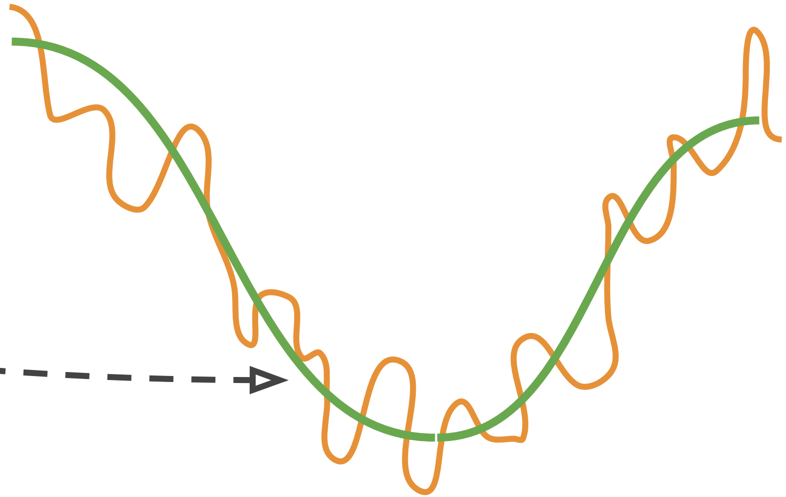
Python coding environment -> ***Jupyter*** notebook

Geodan* servers** and ***Google Earth Engine

Literature research -> specifications

Preprocess steps

- **Load** 2015-2020 data -> aggregate on **hourly intervals**
- **Filter outliers** and **smoothen** data
- **Log transform** exponential NO2 data
- **Convert** to sine and cosine (thanks for the tip Brian!)



Methodology | Preparation

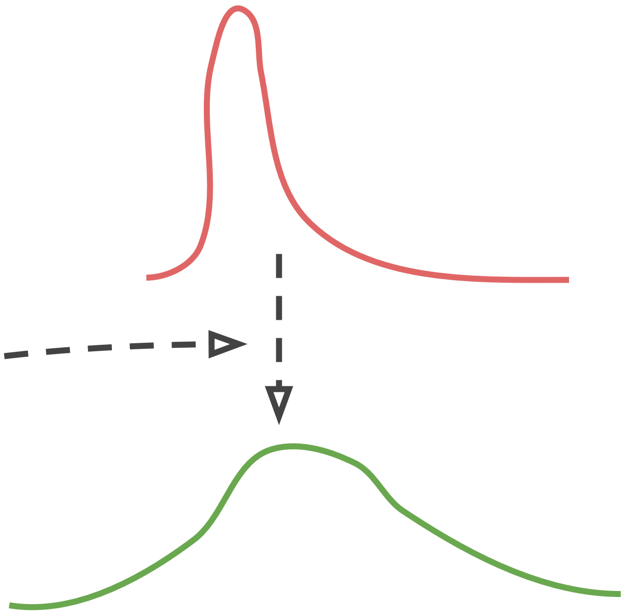
Python coding environment -> ***Jupyter*** notebook

Geodan **servers** and ***Google Earth Engine***

Literature research -> specifications

Preprocess steps

- **Load** 2015-2020 data -> aggregate on **hourly intervals**
- **Filter outliers** and **smoothen** data
- **Log transform** exponential NO2 data
- **Convert** to sine and cosine (thanks for the tip Brian!)



Methodology | Preparation

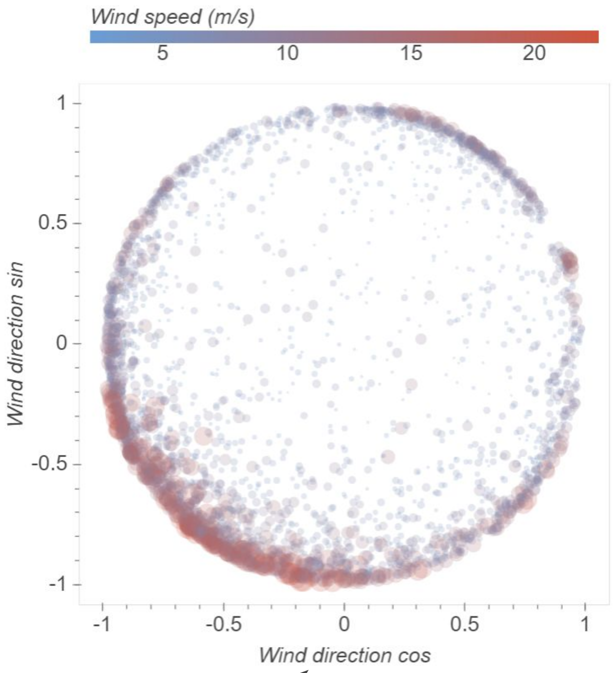
Python coding environment -> ***Jupyter*** notebook

Geodan **servers** and ***Google Earth Engine***

Literature research -> **specifications**

Preprocess steps

- **Load** 2015-2020 data -> aggregate on **hourly intervals**
- **Filter outliers** and **smoothen** data
- **Log transform** exponential NO2 data
- **Convert** to sine and cosine (thanks for the tip Brian!)



Methodology | Investigation

Relationships

- **Linear correlations** & **explainable spread**

Predictions

- **Random forest regressor** machine learning model -> **predict** NO2 data -> **cross validation** -> **prediction accuracy**
- **Variable importance** -> remove **redundant** variables
- Repeat: 2020 COVID-19 data is **excluded** & S5P data is **included**

Trend analysis

- Long term **seasonal trends** -> linear **harmonic model**
- Temporal **overlap comparison**

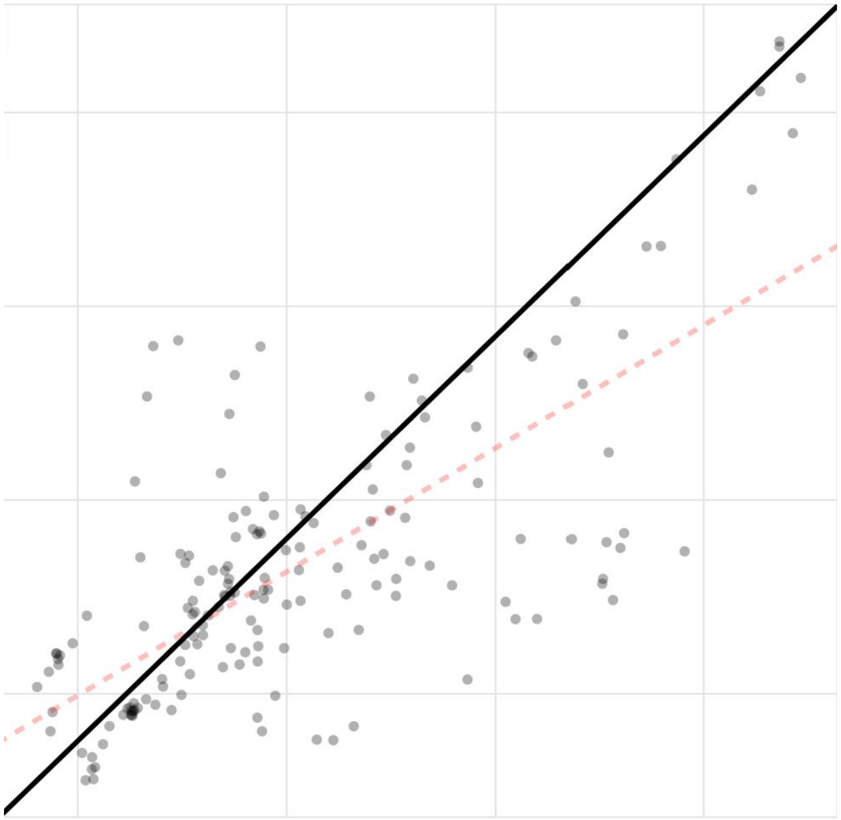
Methodology | Statistics

Pearson correlation coefficient: **r**

- Measures the **linear correlation** between two variables

Coefficient of determination: **r^2**

- Measures the proportion of the **spread** in the response variable (i.e. NO2) that is predictable from the explanatory variable (i.e. road traffic/weather)



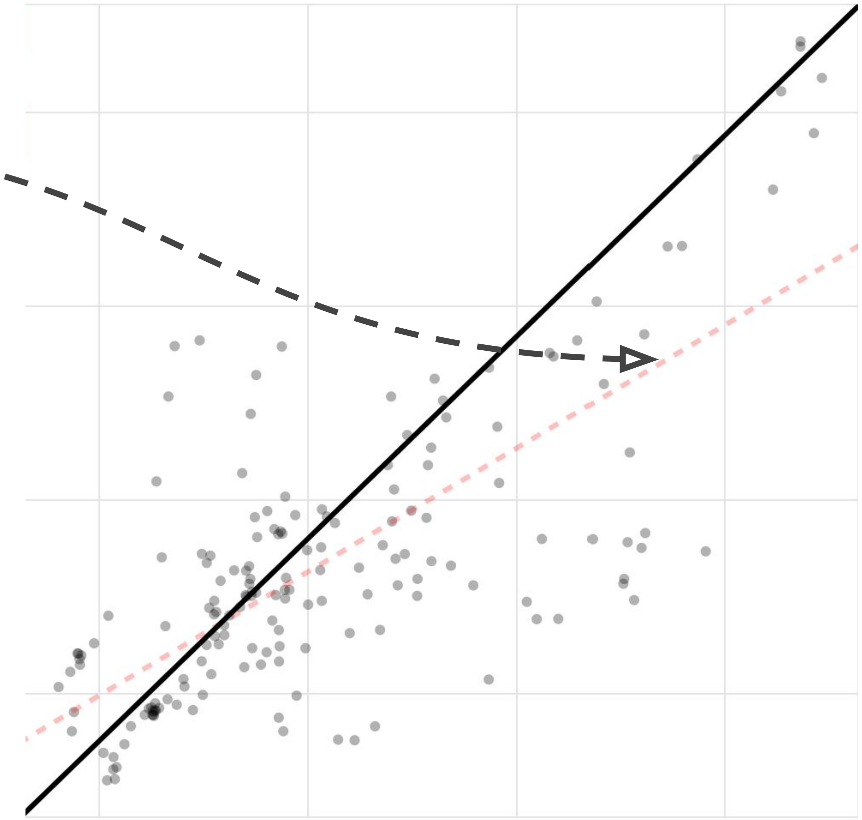
Methodology | Statistics

Pearson correlation coefficient: **r**

- Measures the **linear correlation** between two variables

Coefficient of determination: $\mathbf{r^2}$

- Measures the proportion of the **spread** in the response variable (i.e. NO2) that is predictable from the explanatory variable (i.e. road traffic/weather)



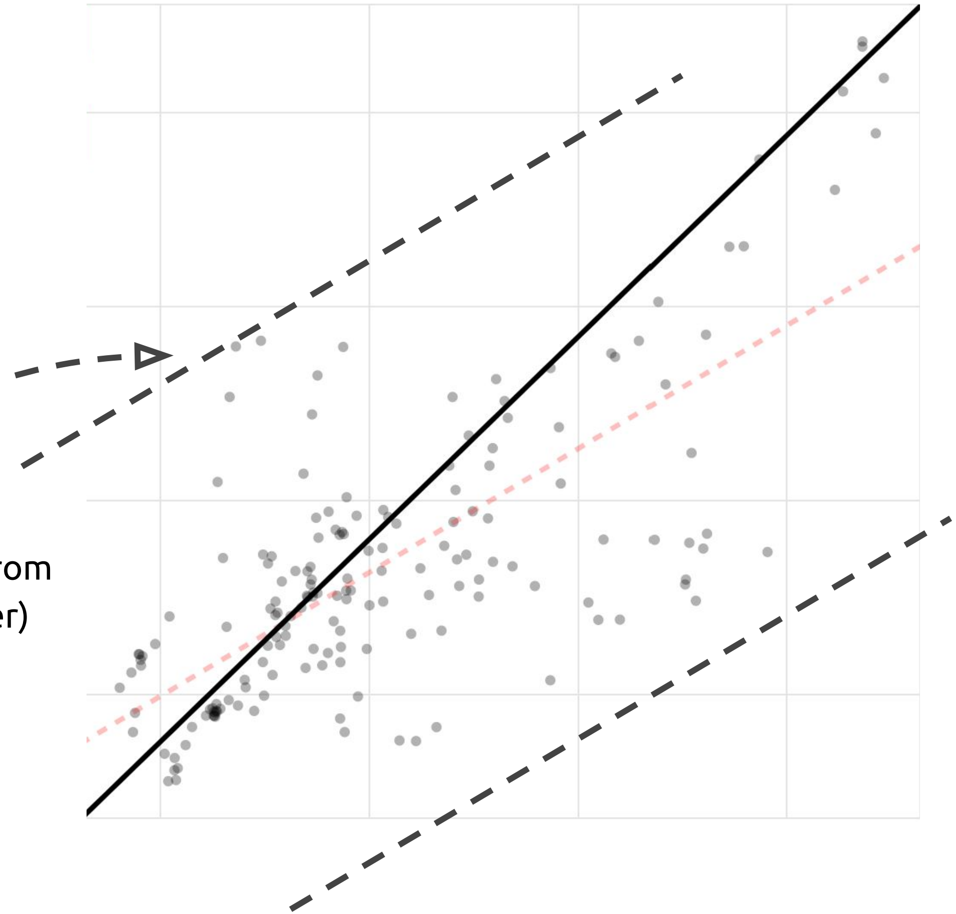
Methodology | Statistics

Pearson correlation coefficient: r

- Measures the **linear correlation** between two variables

Coefficient of determination: r^2

- Measures the proportion of the **spread** in the response variable (i.e. NO2) that is predictable from the explanatory variable (i.e. road traffic/weather)



Methodology | Investigation

Relationships

- **Linear correlations** & **explainable spread**

Predictions

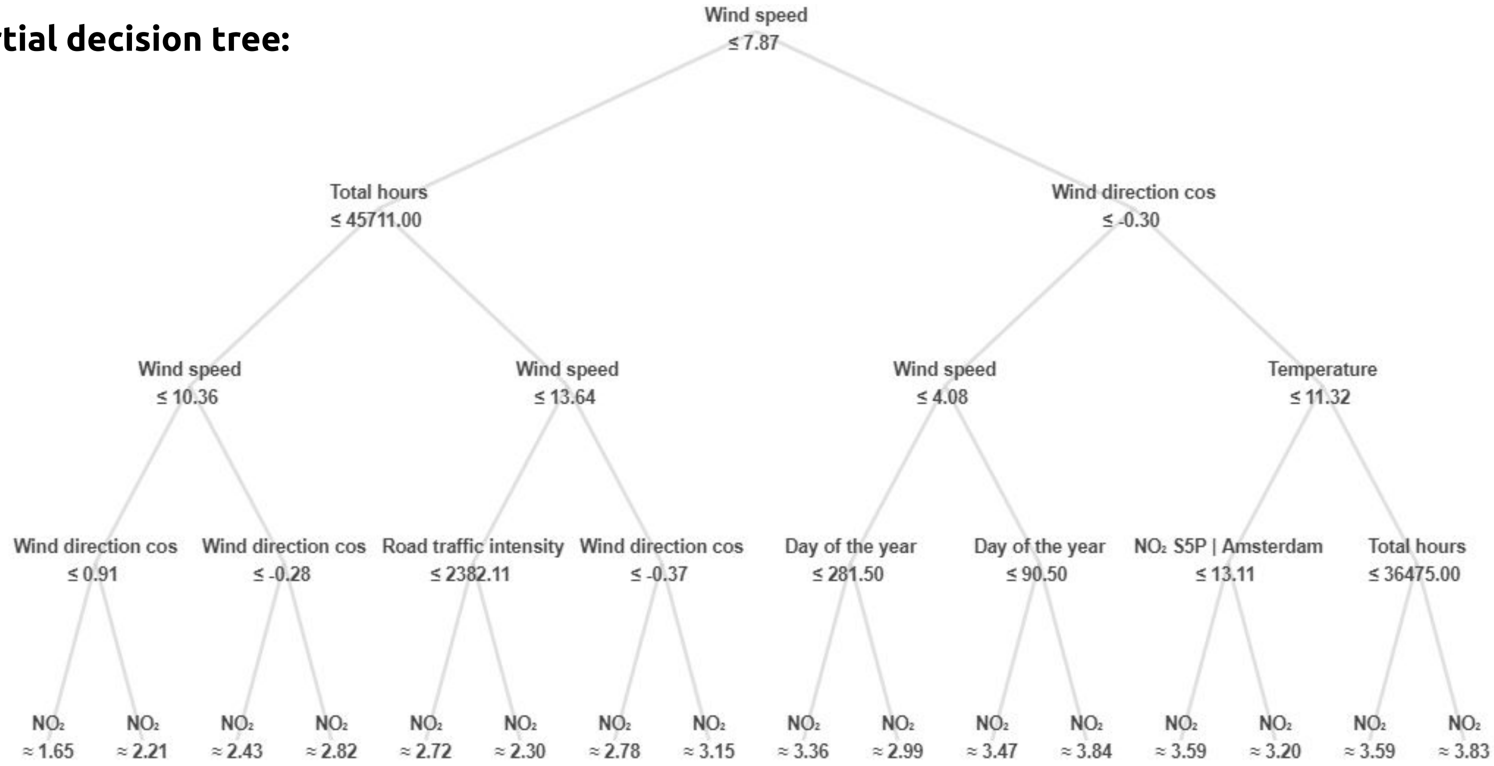
- **Random forest regressor** machine learning model -> **predict** NO2 data
  -> **cross validation** -> **prediction accuracy**
- **Variable importance** -> remove **redundant** variables
- Repeat: 2020 COVID-19 data is **excluded** & S5P data is **included**

Trend analysis

- Long term **seasonal trends** -> linear **harmonic model**
- Temporal **overlap comparison**

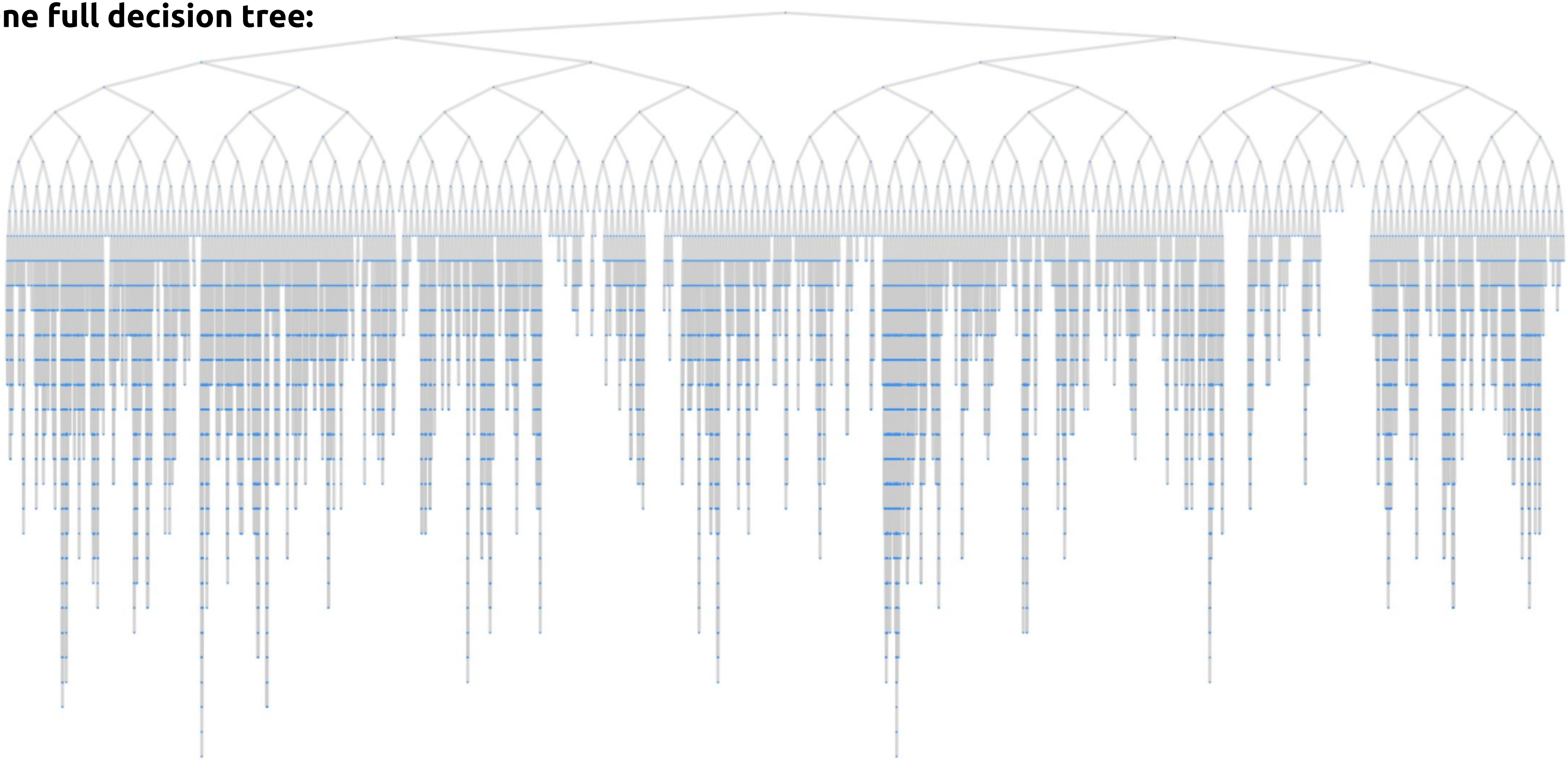
Methodology | Random forest regressor

One partial decision tree:



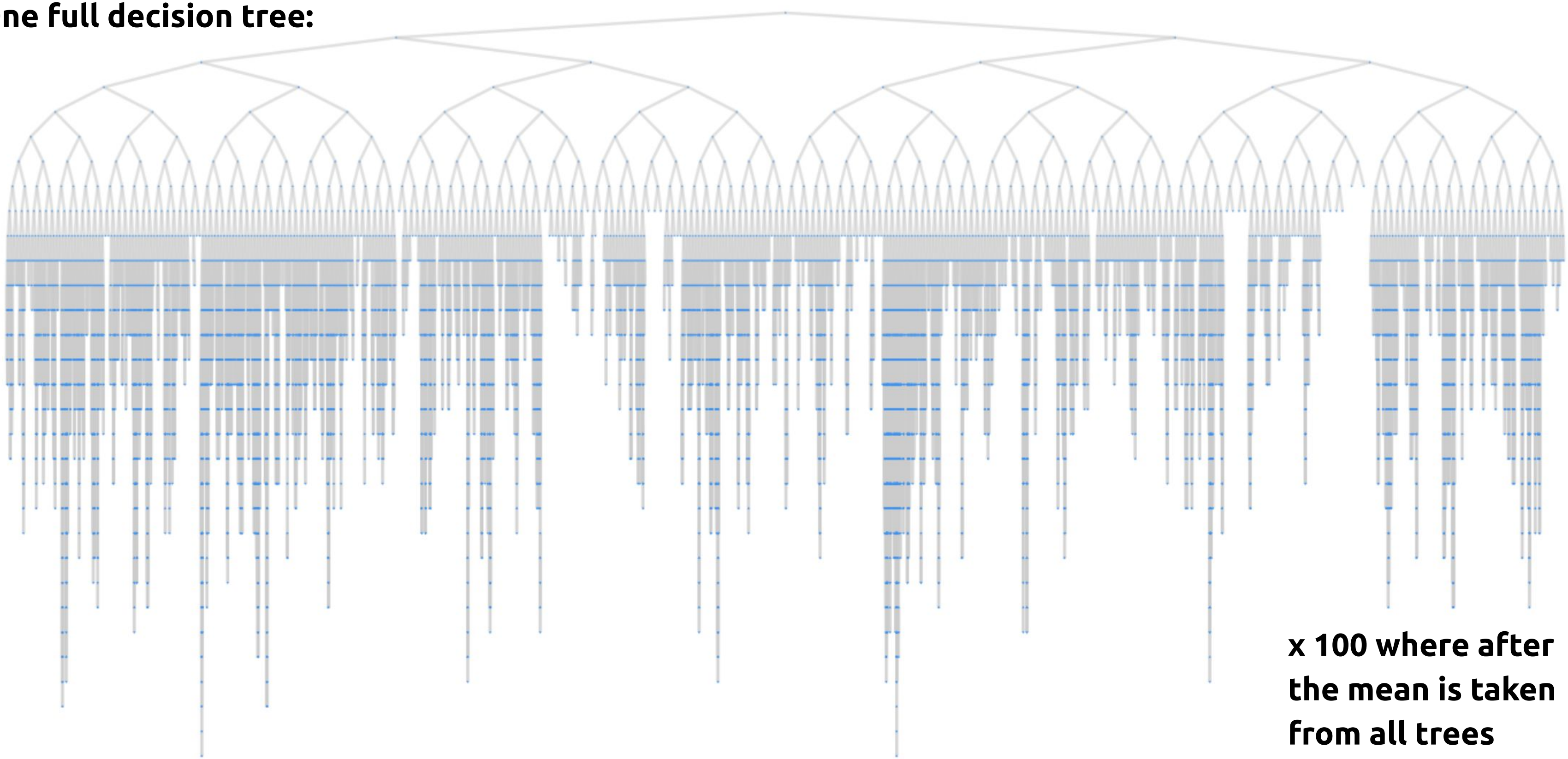
Methodology | Random forest regressor

One full decision tree:



Methodology | Random forest regressor

One full decision tree:



x 100 where after the mean is taken from all trees

Methodology | Investigation

Relationships

- **Linear correlations** & **explainable spread**

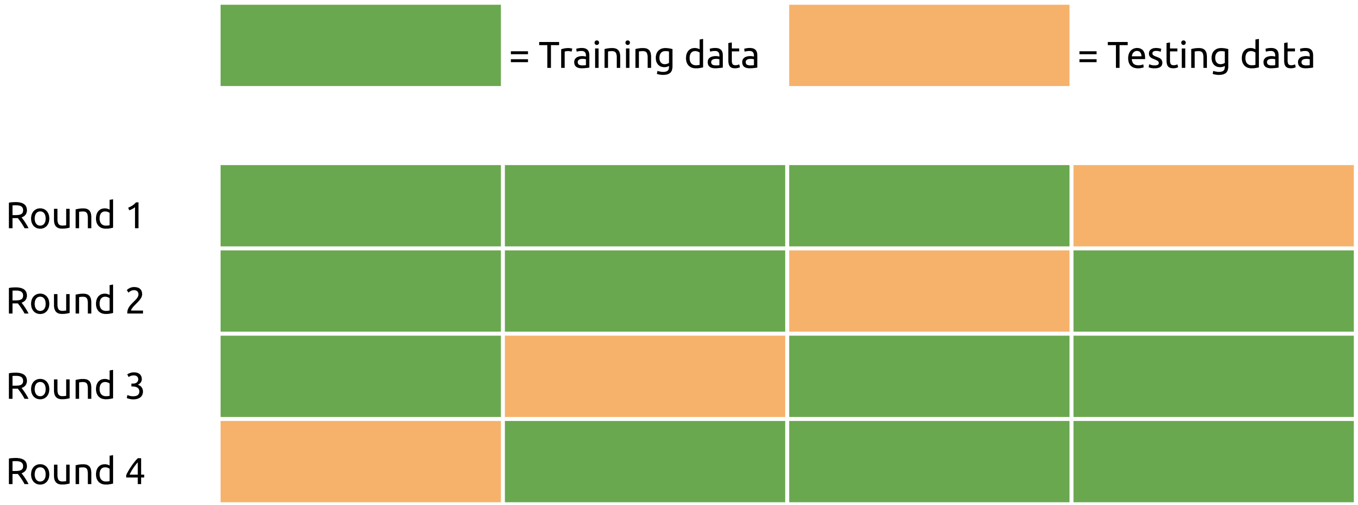
Predictions

- **Random forest regressor** machine learning model -> **predict** NO2 data
  -> **cross validation** -> **prediction accuracy**
- **Variable importance** -> remove **redundant** variables
- Repeat: 2020 COVID-19 data is **excluded** & S5P data is **included**

Trend analysis

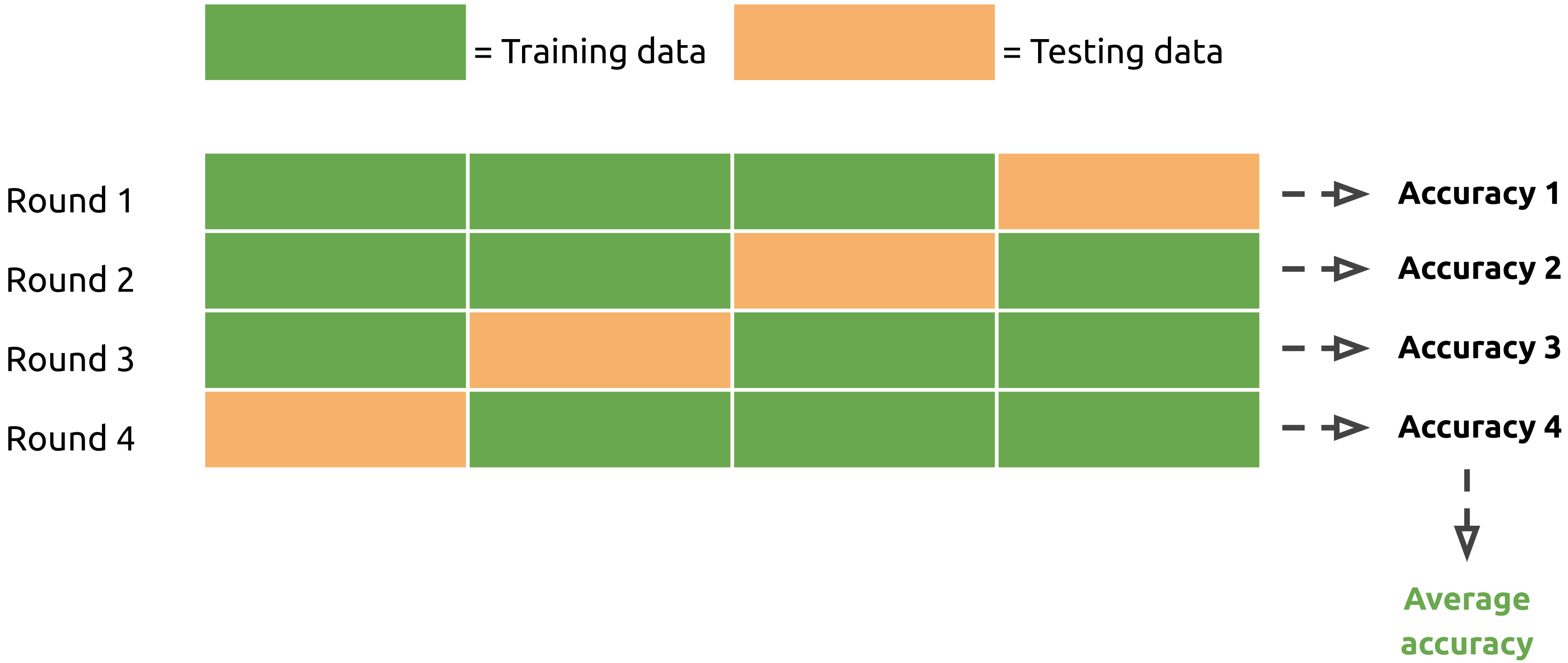
- Long term **seasonal trends** -> linear **harmonic model**
- Temporal **overlap comparison**

Methodology | Cross validation



Methodology | Cross validation

Methodology | Cross validation



Methodology | Investigation

Relationships

- **Linear correlations** & **explainable spread**

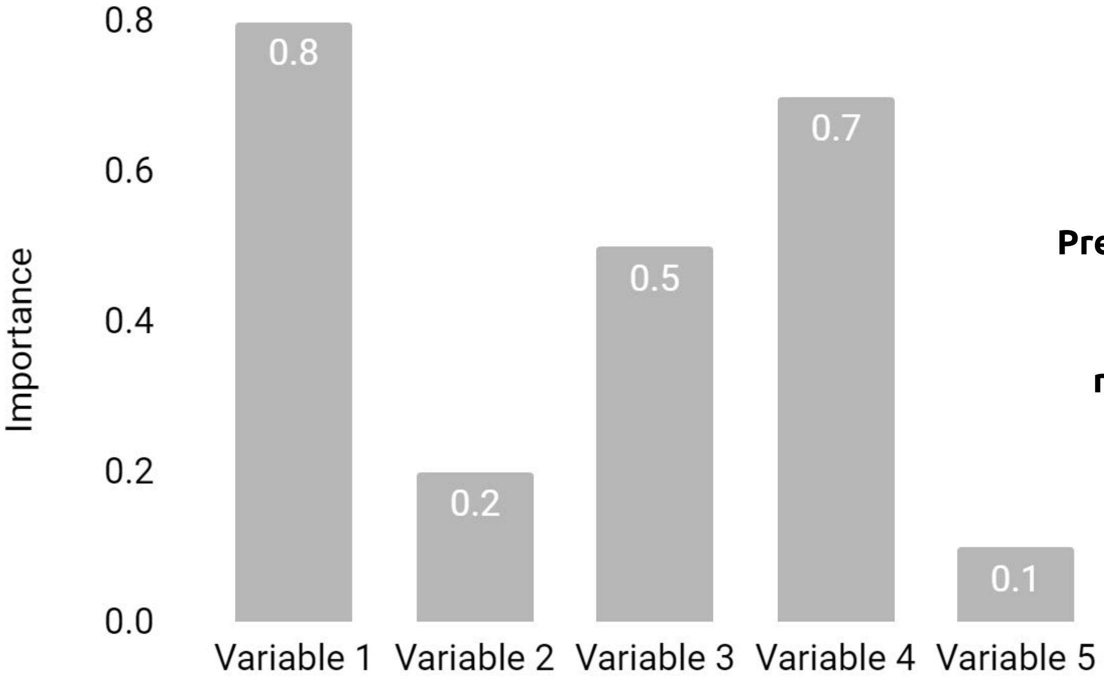
Predictions

- **Random forest regressor** machine learning model -> **predict** NO2 data -> **cross validation** -> **prediction accuracy**
- **Variable importance ->** remove **redundant** variables
- Repeat: 2020 COVID-19 data is **excluded** & S5P data is **included**

Trend analysis

- Long term **seasonal trends** -> linear **harmonic model**
- Temporal **overlap comparison**

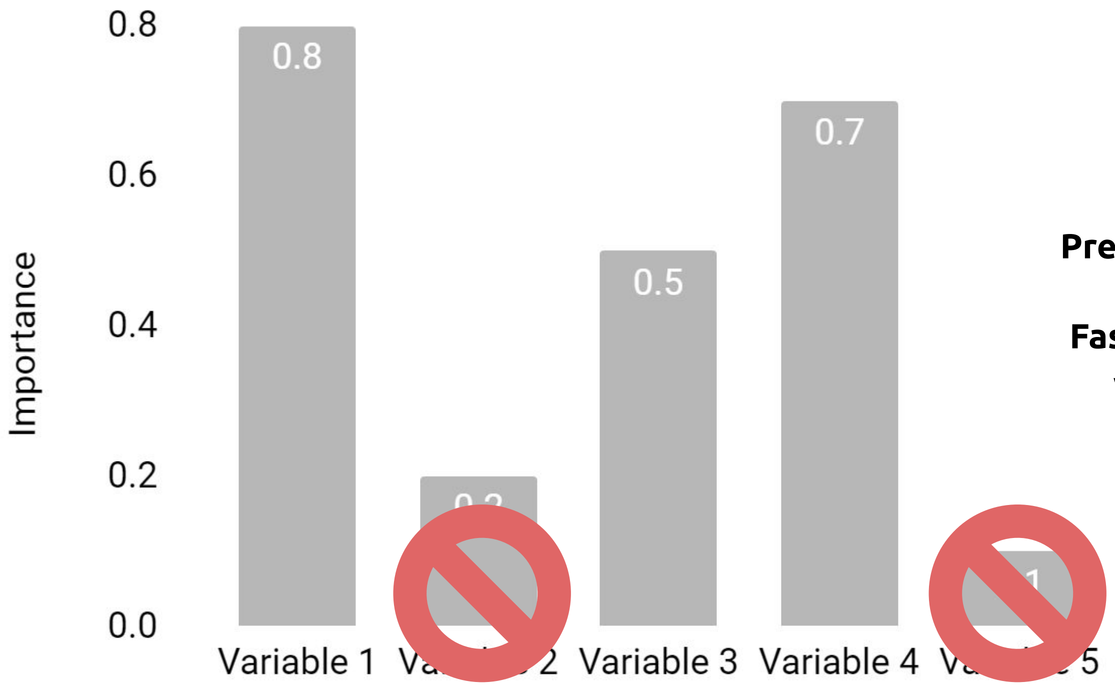
Methodology | Variable Importance



Prediction accuracy 90%

Slow training -> redundant variables

Methodology | Variable Importance



Prediction accuracy 88%

Fast training -> optimal variables - > focus

Methodology | Investigation

Relationships

- **Linear correlations** & **explainable spread**

Predictions

- **Random forest regressor** machine learning model -> **predict** NO_2 data -> **cross validation** -> **prediction accuracy**
- **Variable importance** -> remove **redundant** variables
- Repeat: 2020 COVID-19 data is **excluded** & S5P data is **included**

Trend analysis

- Long term **seasonal trends** -> linear **harmonic model**
- Temporal **overlap comparison**

Methodology | Investigation

Relationships

- **Linear correlations** & **explainable spread**

Predictions

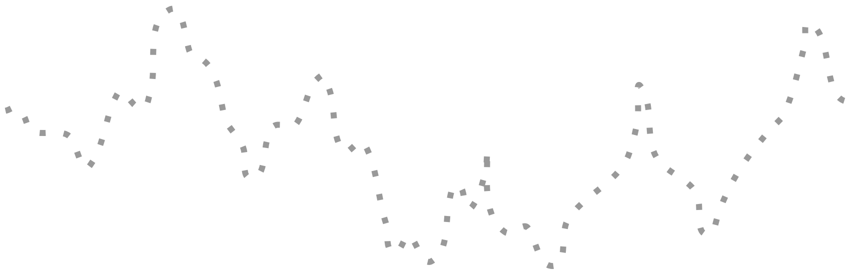
- **Random forest regressor** machine learning model -> **predict** NO2 data -> **cross validation** -> **prediction accuracy**
- **Variable importance** -> remove **redundant** variables
- Repeat: 2020 COVID-19 data is **excluded** & S5P data is **included**

Trend analysis

- Long term **seasonal trends** -> linear **harmonic model**
- Temporal **overlap comparison**

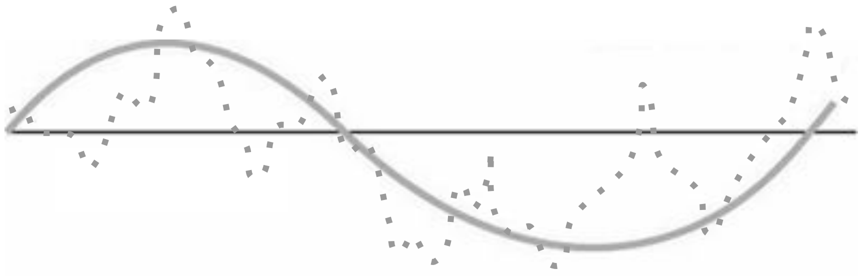
Methodology | Harmonic model

Raw data:



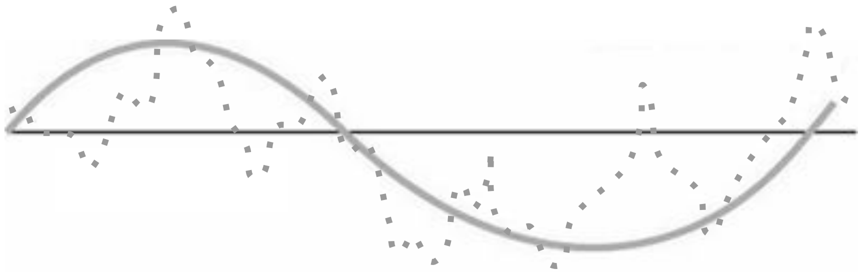
Methodology | Harmonic model

Yearly seasonal wave:

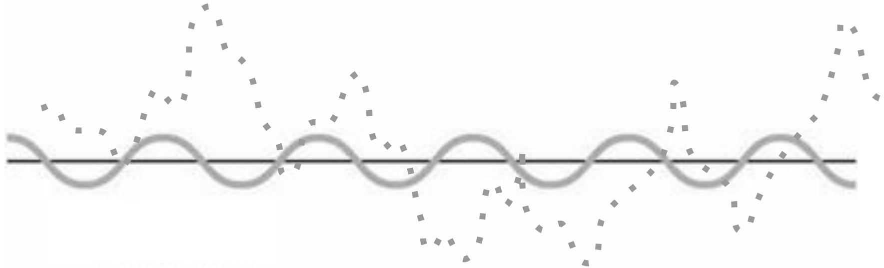


Methodology | Harmonic model

Yearly seasonal wave:

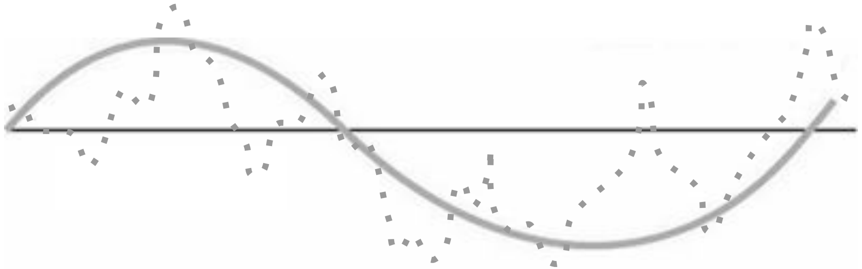


Quarterly seasonal wave:

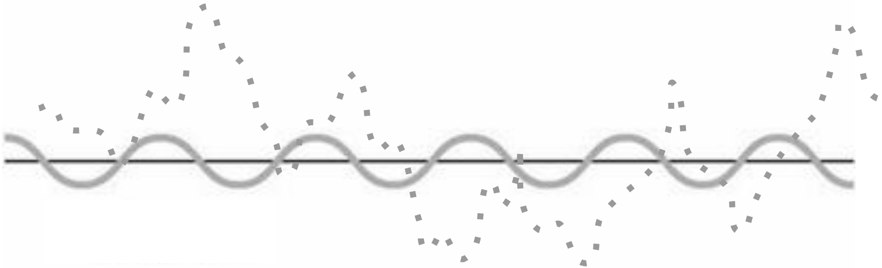


Methodology | Harmonic model

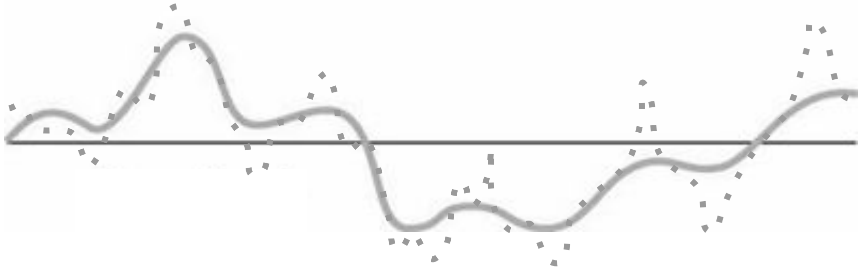
Yearly seasonal wave:



Quarterly seasonal wave:



Merged seasonal waves:



Methodology | Investigation

Relationships

- **Linear correlations** & **explainable spread**

Predictions

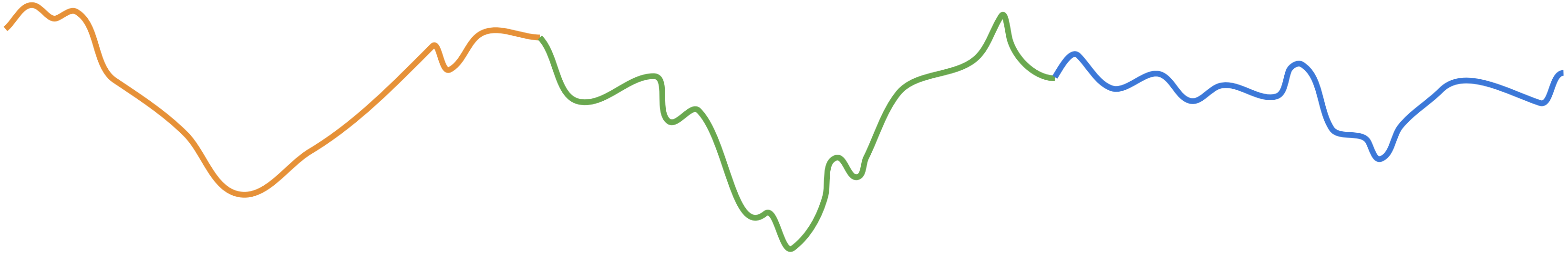
- **Random forest regressor** machine learning model -> **predict** NO2 data -> **cross validation** -> **prediction accuracy**
- **Variable importance** -> remove **redundant** variables
- Repeat: 2020 COVID-19 data is **excluded** & S5P data is **included**

Trend analysis

- Long term **seasonal trends** -> linear **harmonic model**
- Temporal **overlap comparison**

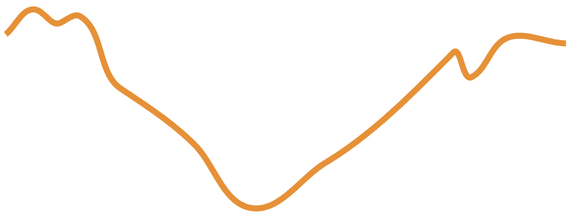
Methodology | Overlap comparison

Year 1: **Year 2:** **Year 3:**

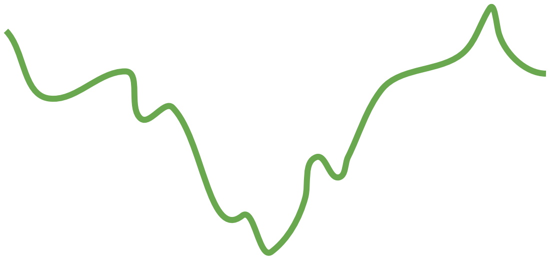


Methodology | Overlap comparison

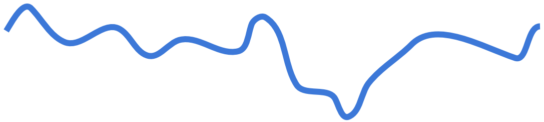
Year 1:



Year 2:



Year 3:

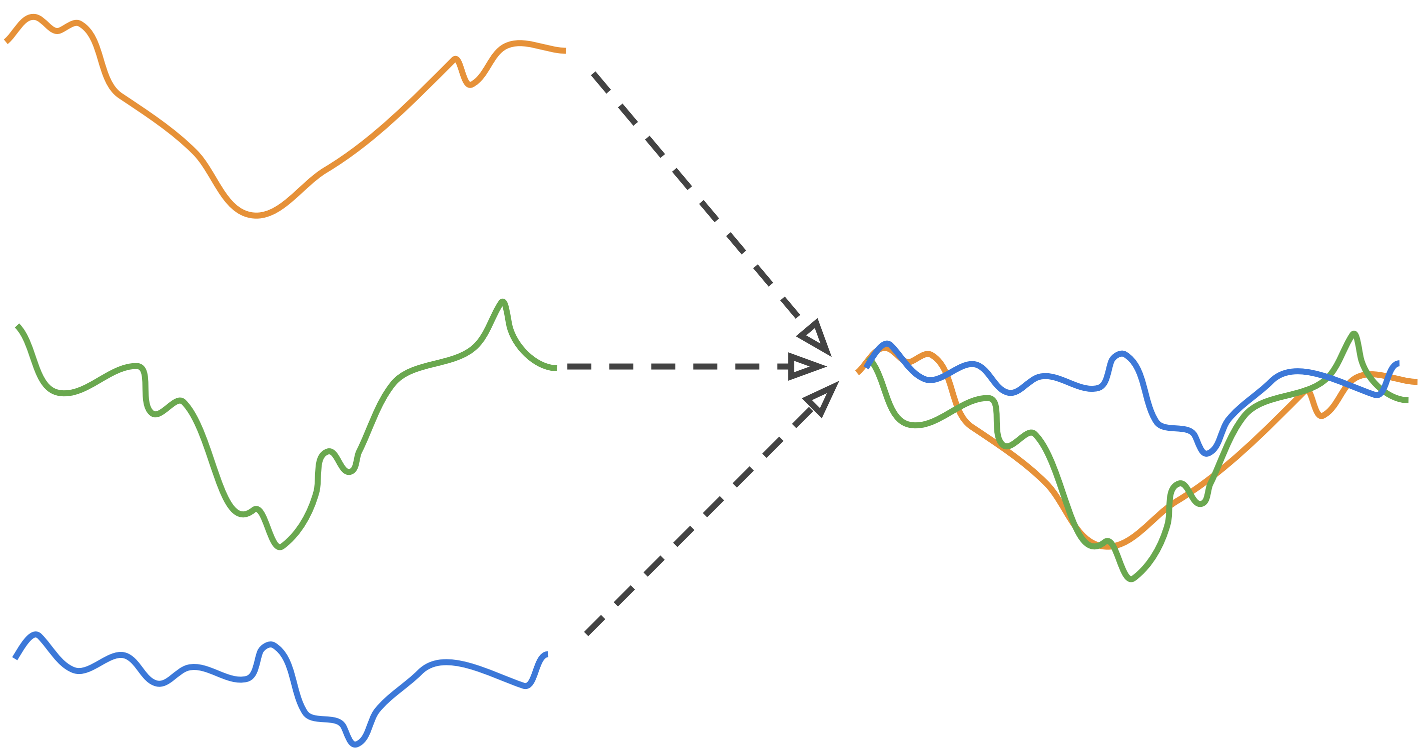


Methodology | Overlap comparison

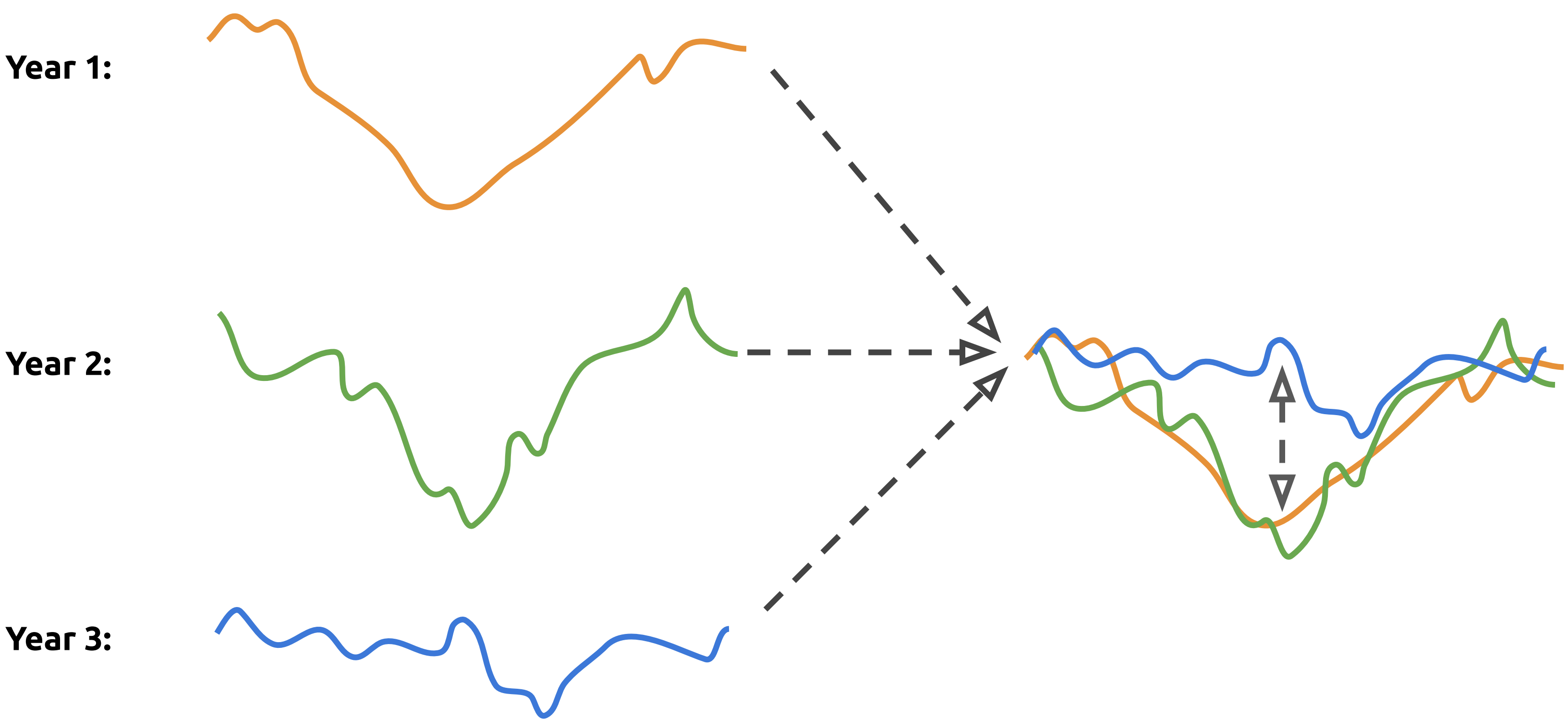
Year 1:

Year 2:

Year 3:

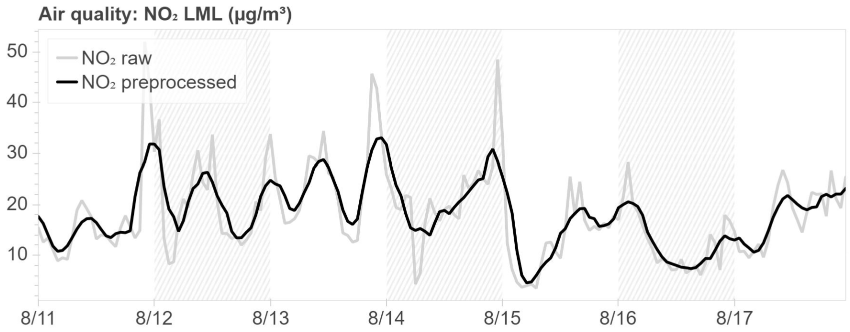


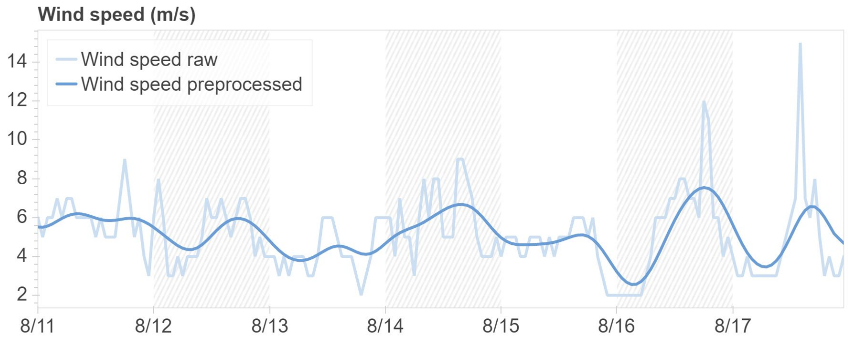
Methodology | Overlap comparison

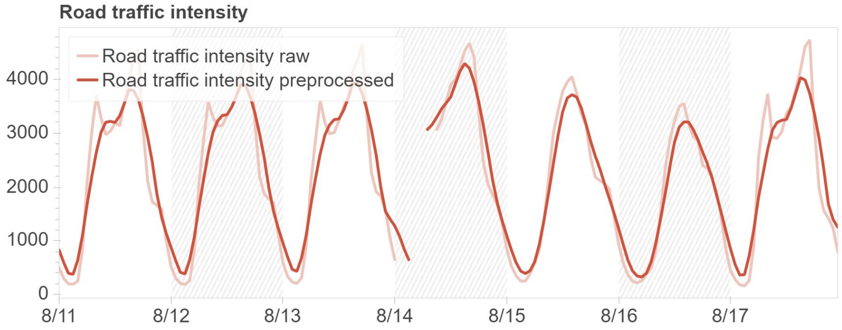


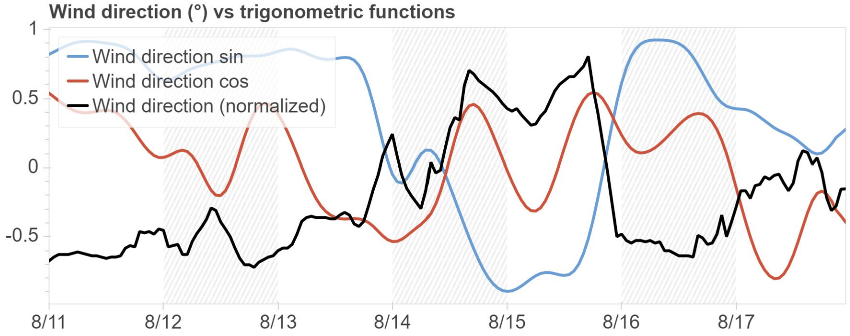
Results

Results | Preparation





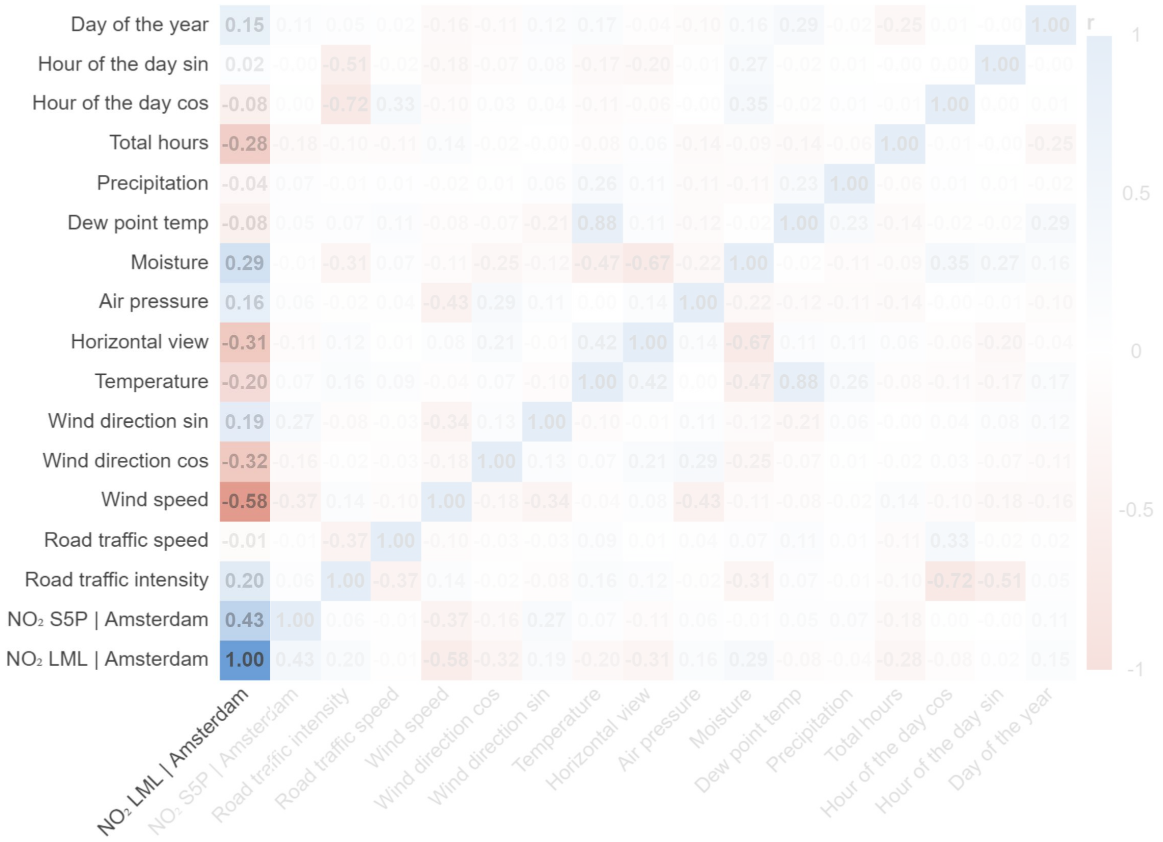




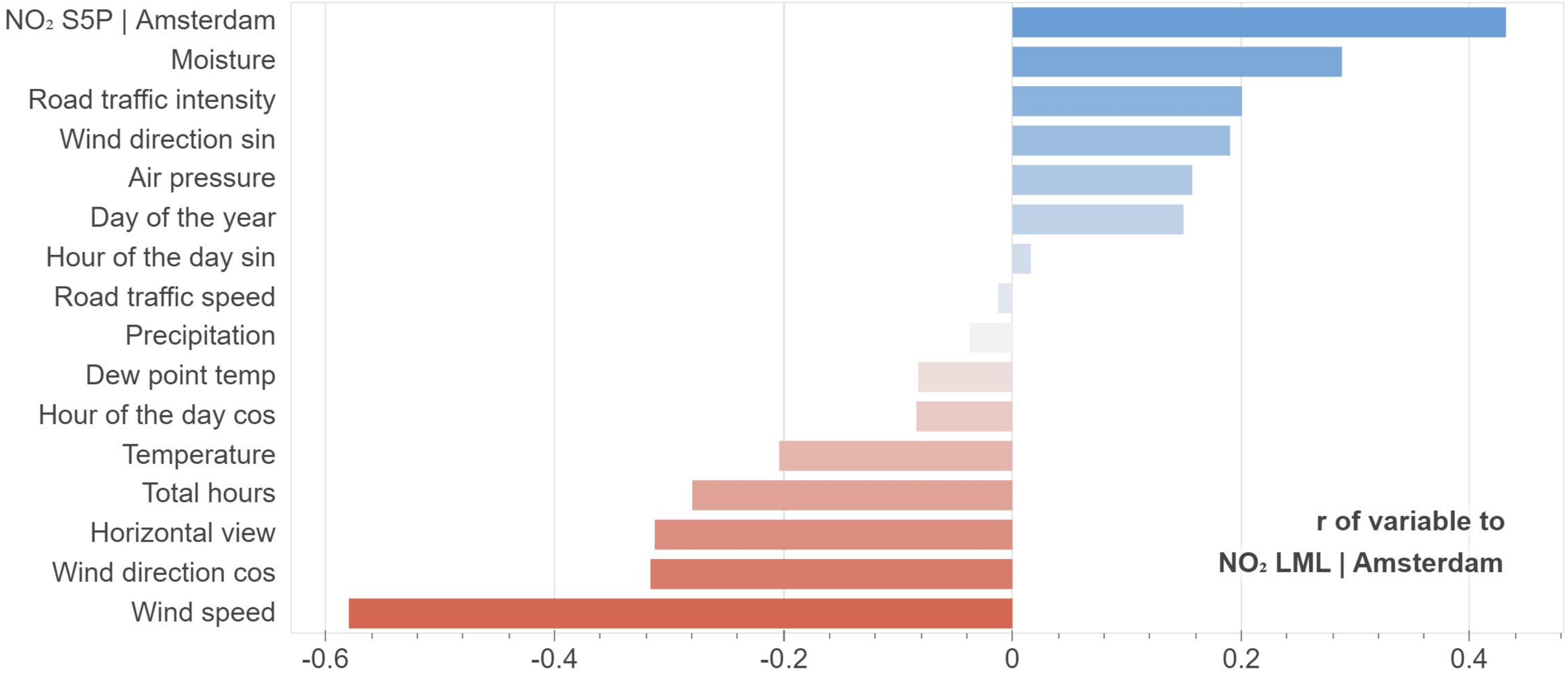
Results | Exploration - Linear correlations

	NO_2 LML | Amsterdam	NO_2 S5P | Amsterdam	Road traffic intensity	Road traffic speed	Wind speed	Wind direction cos	Wind direction sin	Temperature	Horizontal view	Air pressure	Moisture	Dew point temp	Precipitation	Total hours	Hour of the day cos	Hour of the day sin	Day of the year
Day of the year	0.15	0.11	0.05	0.02	-0.16	-0.11	0.12	0.17	-0.04	-0.10	0.16	0.29	-0.02	-0.25	0.01	-0.00	1.00
Hour of the day sin	0.02	-0.00	-0.51	-0.02	-0.18	-0.07	0.08	-0.17	-0.20	-0.01	0.27	-0.02	0.01	-0.00	0.00	1.00	-0.00
Hour of the day cos	-0.08	0.00	-0.72	0.33	-0.10	0.03	0.04	-0.11	-0.06	-0.00	0.35	-0.02	0.01	-0.01	1.00	0.00	0.01
Total hours	-0.28	-0.18	-0.10	-0.11	0.14	-0.02	-0.00	-0.08	0.06	-0.14	-0.09	-0.14	-0.06	1.00	-0.01	-0.00	-0.25
Precipitation	-0.04	0.07	-0.01	0.01	-0.02	0.01	0.06	0.26	0.11	-0.11	-0.11	0.23	1.00	-0.06	0.01	0.01	-0.02
Dew point temp	-0.08	0.05	0.07	0.11	-0.08	-0.07	-0.21	0.88	0.11	-0.12	-0.02	1.00	0.23	-0.14	-0.02	-0.02	0.29
Moisture	0.29	-0.01	-0.31	0.07	-0.11	-0.25	-0.12	-0.47	-0.67	-0.22	1.00	-0.02	-0.11	-0.09	0.35	0.27	0.16
Air pressure	0.16	0.06	-0.02	0.04	-0.43	0.29	0.11	0.00	0.14	1.00	-0.22	-0.12	-0.11	-0.14	-0.00	-0.01	-0.10
Horizontal view	-0.31	-0.11	0.12	0.01	0.08	0.21	-0.01	0.42	1.00	0.14	-0.67	0.11	0.11	0.06	-0.06	-0.20	-0.04
Temperature	-0.20	0.07	0.16	0.09	-0.04	0.07	-0.10	1.00	0.42	0.00	-0.47	0.88	0.26	-0.08	-0.11	-0.17	0.17
Wind direction sin	0.19	0.27	-0.08	-0.03	-0.34	0.13	1.00	-0.10	-0.01	0.11	-0.12	-0.21	0.06	-0.00	0.04	0.08	0.12
Wind direction cos	-0.32	-0.16	-0.02	-0.03	-0.18	1.00	0.13	0.07	0.21	0.29	-0.25	-0.07	0.01	-0.02	0.03	-0.07	-0.11
Wind speed	-0.58	-0.37	0.14	-0.10	1.00	-0.18	-0.34	-0.04	0.08	-0.43	-0.11	-0.08	-0.02	0.14	-0.10	-0.18	-0.16
Road traffic speed	-0.01	-0.01	-0.37	1.00	-0.10	-0.03	-0.03	0.09	0.01	0.04	0.07	0.11	0.01	-0.11	0.33	-0.02	0.02
Road traffic intensity	0.20	0.06	1.00	-0.37	0.14	-0.02	-0.08	0.16	0.12	-0.02	-0.31	0.07	-0.01	-0.10	-0.72	-0.51	0.05
NO_2 S5P | Amsterdam	0.43	1.00	0.06	-0.01	-0.37	-0.16	0.27	0.07	-0.11	0.06	-0.01	0.05	0.07	-0.18	0.00	-0.00	0.11
NO_2 LML | Amsterdam	1.00	0.43	0.20	-0.01	-0.58	-0.32	0.19	-0.20	-0.31	0.16	0.29	-0.08	-0.04	-0.28	-0.08	0.02	0.15

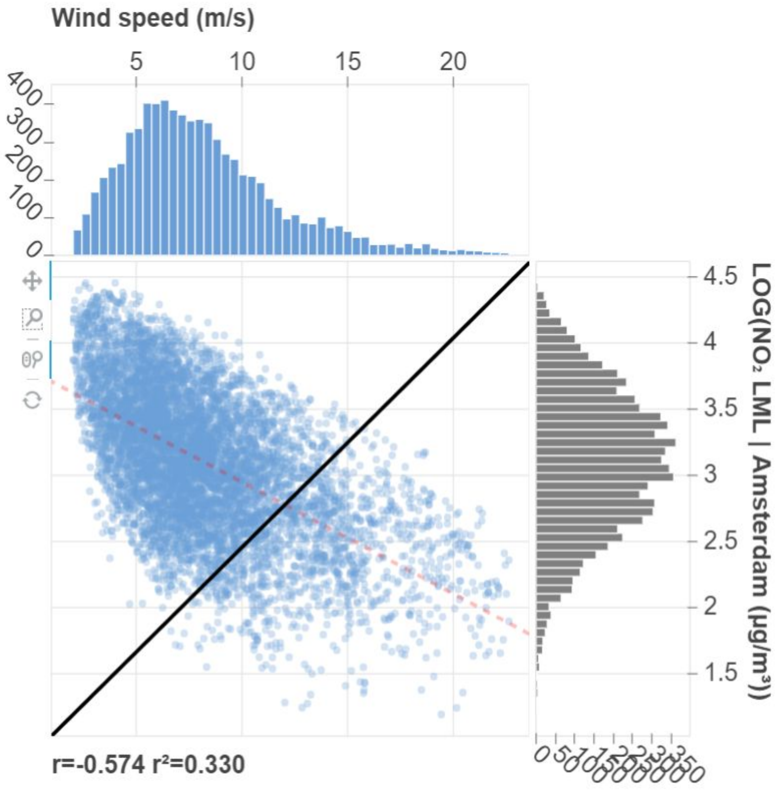
Results | Exploration - Linear correlations

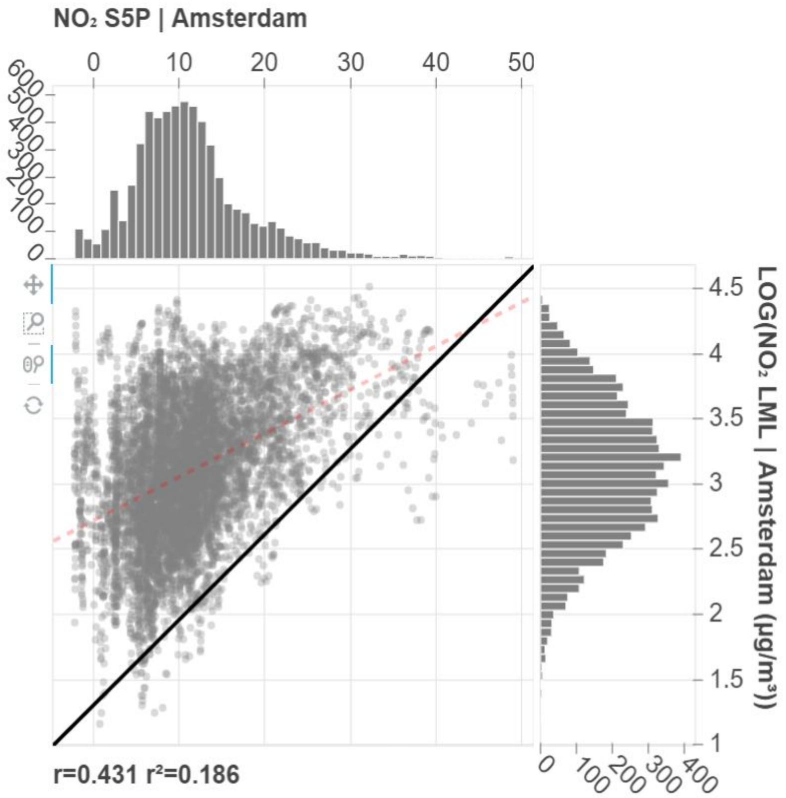


Results | Exploration - Linear correlations

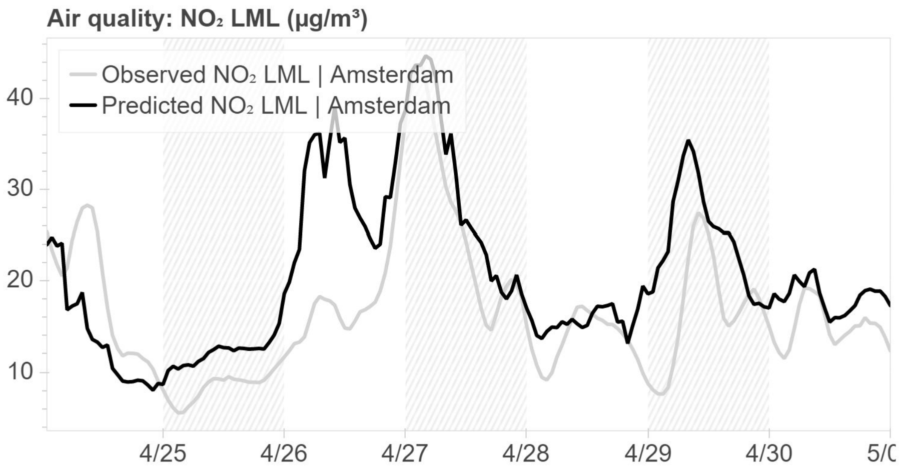


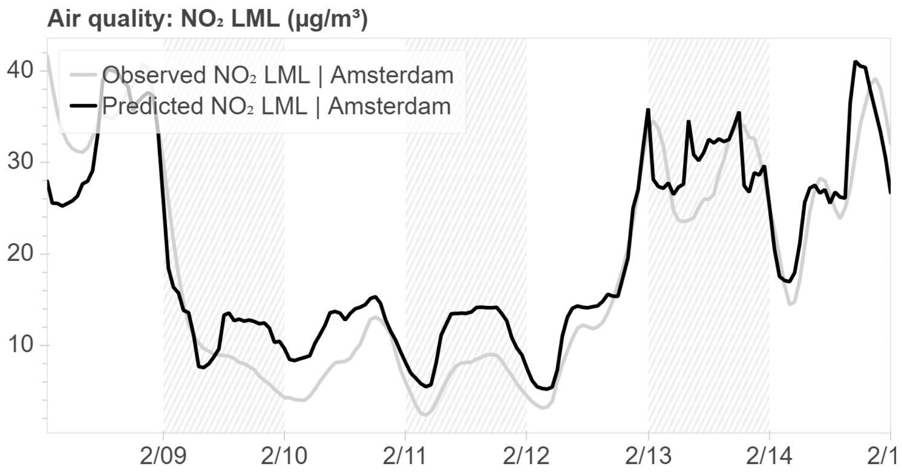
Results | Exploration - Linear correlations



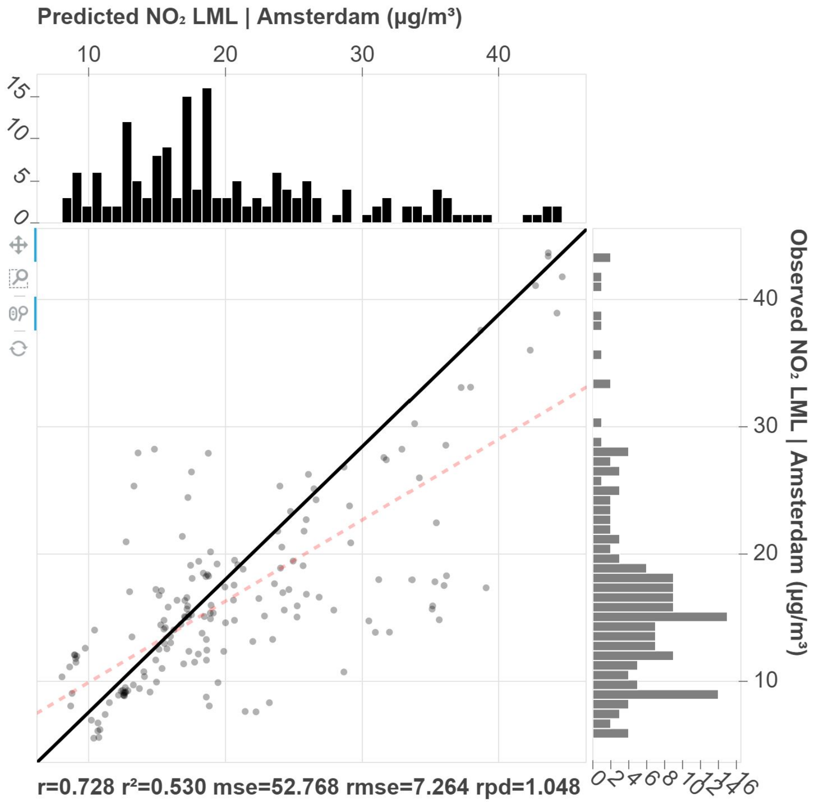


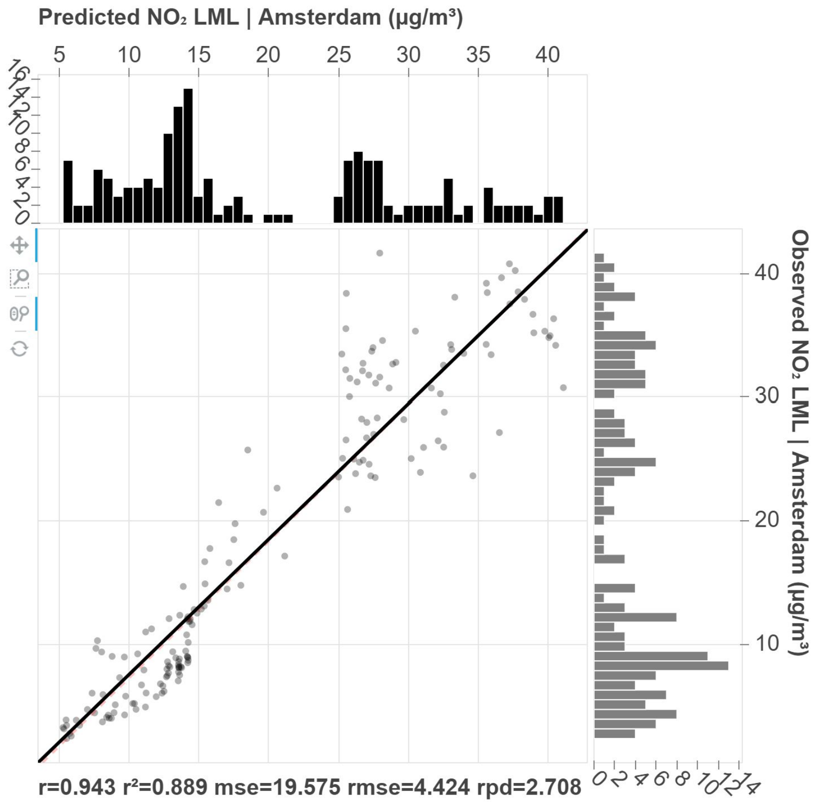
Results | Prediction





Results | Prediction





Results | Prediction accuracy

Location	r	-/+ 95% CI	r²	-/+ 95% CI	MSE	-/+ 95% CI	RMSE	-/+ 95% CI
NO_2 LML | Amsterdam	0.854	0.107	0.732	0.172	0.063	0.050	0.248	0.084
NO_2 LML | Amsterdam-Oude Schans	0.777	0.168	0.610	0.241	0.112	0.106	0.327	0.142
NO_2 LML | Amsterdam-Stadhouderskade	0.766	0.170	0.593	0.242	0.085	0.076	0.285	0.112
NO_2 LML | Amsterdam-Van Diemenstraat	0.854	0.126	0.733	0.197	0.080	0.082	0.277	0.118
NO_2 LML | Amsterdam-Haarlemmerweg	0.828	0.092	0.688	0.150	0.079	0.056	0.277	0.097
NO_2 LML | Amsterdam-Jan van Galenstraat	0.844	0.112	0.716	0.174	0.072	0.062	0.264	0.095
NO_2 LML | Amsterdam-Vondelpark	0.772	0.165	0.603	0.237	0.117	0.091	0.337	0.117
NO_2 LML | Amsterdam-Nieuwendammerdijk	0.794	0.117	0.634	0.179	0.168	0.151	0.401	0.170
NO_2 LML | Amsterdam-Einsteinweg	0.891	0.062	0.794	0.109	0.069	0.057	0.258	0.094
NO_2 LML | Amsterdam-Ookmeer	0.815	0.104	0.667	0.164	0.195	0.152	0.434	0.164
NO_2 LML | Zaanstad-Hemkade	0.842	0.133	0.713	0.204	0.174	0.312	0.393	0.275
NO_2 LML | Amsterdam -Kantershof	0.805	0.087	0.650	0.139	0.129	0.037	0.359	0.051
NO_2 LML | Badhoevedorp-Sloterweg	0.835	0.092	0.700	0.149	0.113	0.075	0.332	0.104
NO_2 LML | Zaanstad-Hoogtij	0.868	0.060	0.755	0.101	0.140	0.049	0.373	0.062
NO_2 LML | Spaarnwoude-Machineweg	0.843	0.049	0.712	0.082	0.160	0.055	0.399	0.067
NO_2 LML | Hoofddorp-Hoofdweg	0.803	0.126	0.648	0.192	0.193	0.087	0.436	0.095
NO_2 LML | Oude Meer-Aalsmeerderdijk	0.806	0.100	0.652	0.155	0.124	0.047	0.350	0.062
NO_2 LML | Haarlem-Schipholweg	0.853	0.069	0.728	0.114	0.088	0.036	0.296	0.058
Total mean	**0.825**	**0.108**	**0.685**	**0.167**	**0.120**	**0.088**	**0.336**	**0.109**

Results | Prediction accuracy

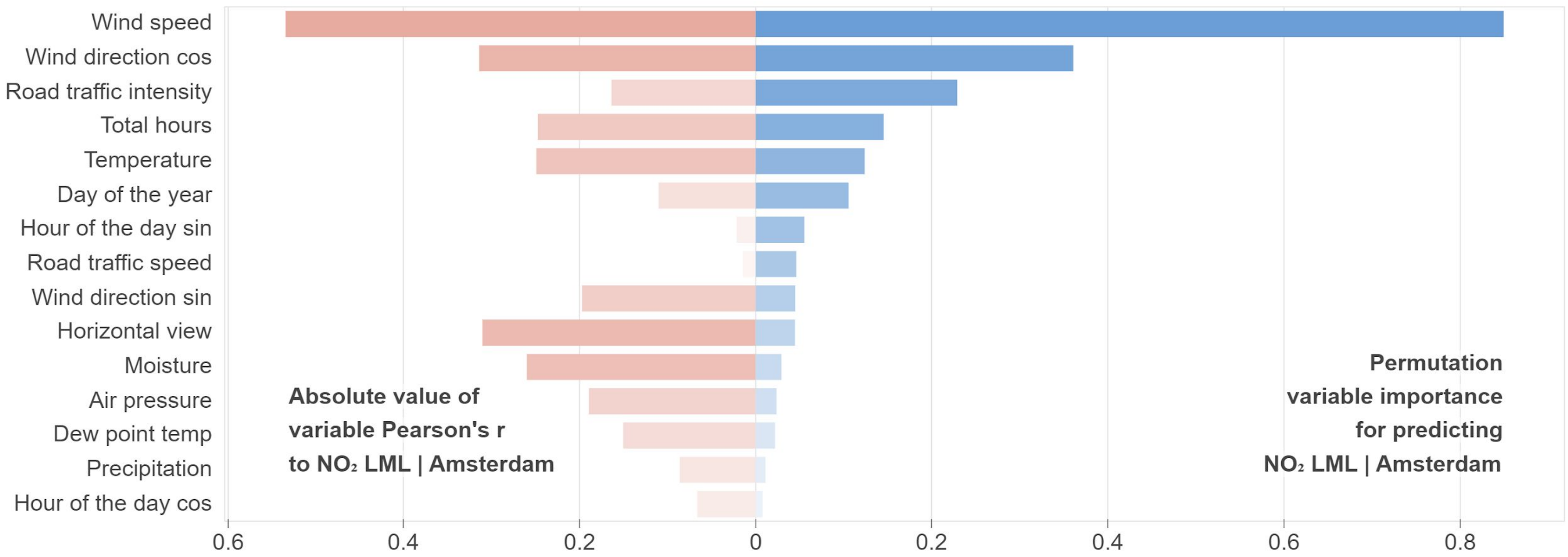
Location	r	-/+ 95% CI	r^2	-/+ 95% CI	MSE	-/+ 95% CI	RMSE	-/+ 95% CI
NO_2 LML | Amsterdam	0.854	0.107	0.732	0.172	0.063	0.050	0.248	0.084
NO_2 LML | Amsterdam-Oude Schans	0.777	0.168	0.610	0.241	0.112	0.106	0.327	0.142
NO_2 LML | Amsterdam-Stadhouderskade	0.766	0.170	0.593	0.242	0.085	0.076	0.285	0.112
NO_2 LML | Amsterdam-Van Diemenstraat	0.854	0.126	0.733	0.197	0.080	0.082	0.277	0.118
NO_2 LML | Amsterdam-Haarlemmerweg	0.828	0.092	0.688	0.150	0.079	0.056	0.277	0.097
NO_2 LML | Amsterdam-Jan van Galenstraat	0.844	0.112	0.716	0.174	0.072	0.062	0.264	0.095
NO_2 LML | Amsterdam-Vondelpark	0.772	0.165	0.603	0.237	0.117	0.091	0.337	0.117
NO_2 LML | Amsterdam-Nieuwendammerdijk	0.794	0.117	0.634	0.179	0.168	0.151	0.401	0.170
NO_2 LML | Amsterdam-Einsteinweg	0.891	0.062	0.794	0.109	0.069	0.057	0.258	0.094
NO_2 LML | Amsterdam-Ookmeer	0.815	0.104	0.667	0.164	0.195	0.152	0.434	0.164
NO_2 LML | Zaanstad-Hemkade	0.842	0.133	0.713	0.204	0.174	0.312	0.393	0.275
NO_2 LML | Amsterdam -Kantershof	0.805	0.087	0.650	0.139	0.129	0.037	0.359	0.051
NO_2 LML | Badhoevedorp-Sloterweg	0.835	0.092	0.700	0.149	0.113	0.075	0.332	0.104
NO_2 LML | Zaanstad-Hoogtij	0.868	0.060	0.755	0.101	0.140	0.049	0.373	0.062
NO_2 LML | Spaarnwoude-Machineweg	0.843	0.049	0.712	0.082	0.160	0.055	0.399	0.067
NO_2 LML | Hoofddorp-Hoofdweg	0.803	0.126	0.648	0.192	0.193	0.087	0.436	0.095
NO_2 LML | Oude Meer-Aalsmeerderdijk	0.806	0.100	0.652	0.155	0.124	0.047	0.350	0.062
NO_2 LML | Haarlem-Schipholweg	0.853	0.069	0.728	0.114	0.088	0.036	0.296	0.058
Total mean	**0.825**	**0.108**	**0.685**	**0.167**	**0.120**	**0.088**	**0.336**	**0.109**

Results | Prediction accuracy

r	r^2
0.825	**0.685**

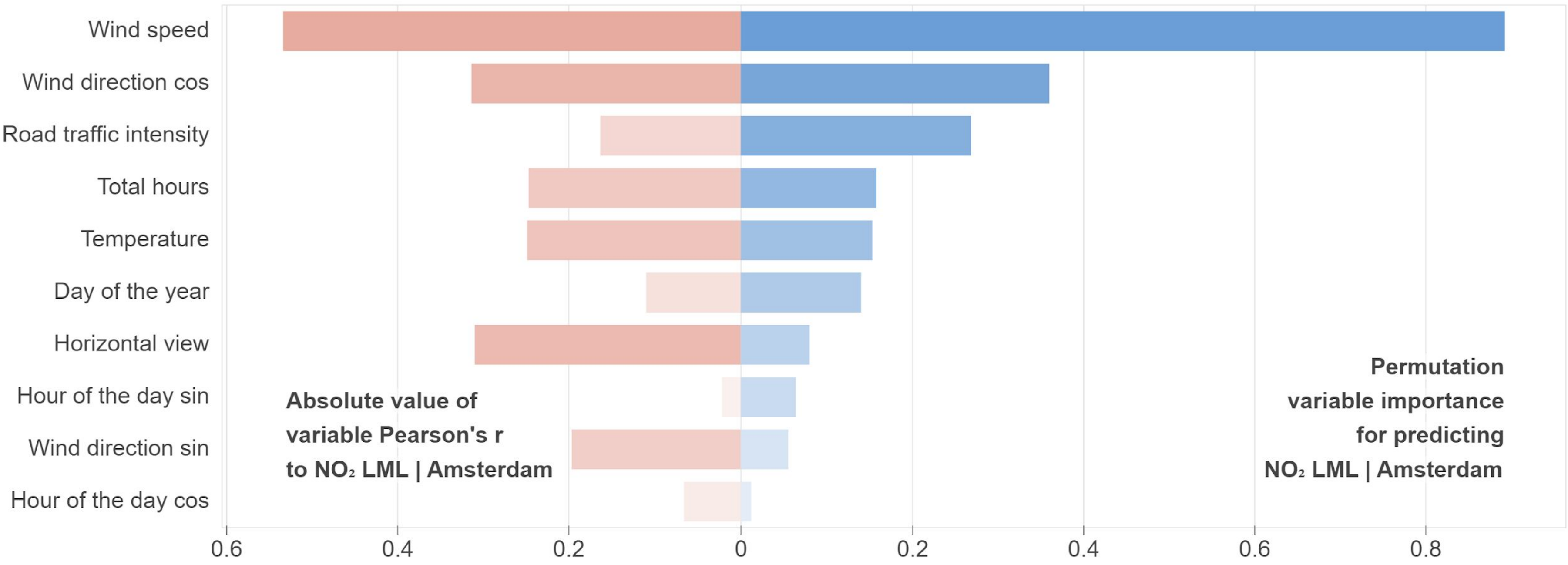
This means that about **~70%** of hourly NO2 data can be accurately predicted

Results | Prediction - Variable importance



Results | Prediction - Variable importance

Removing redundant variables



Results | Prediction accuracy - selected variables

Prediction accuracy after **removing** redundant variables

r	r^2
0.824	**0.683**
-0.11%	*-0.27%*

Accuracy stays roughly the same

Results | Prediction accuracy - no COVID-19 data

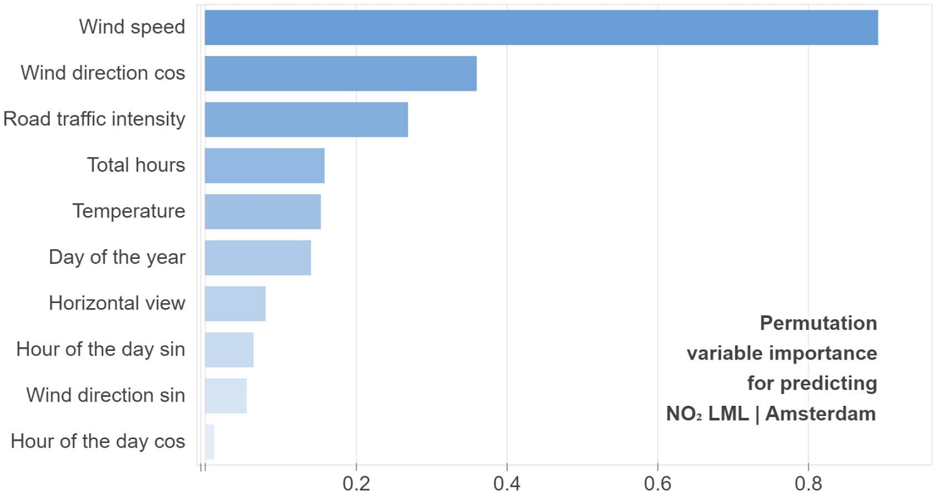
Prediction accuracy after **removing** COVID-19 training & testing data

r	r^2
0.845	**0.715**
+2.51%	*+4.70%*

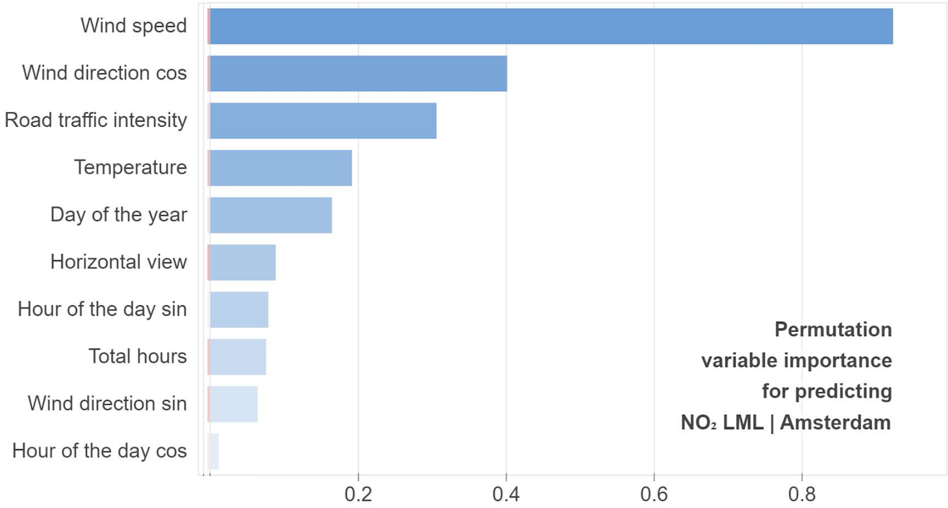
Accuracy improves. This suggest that the data found during COVID-19 lockdowns **disrupt normal 'expected' patterns**

Results | Prediction - Variable importance

Including COVID-19 data



Excluding COVID-19 data

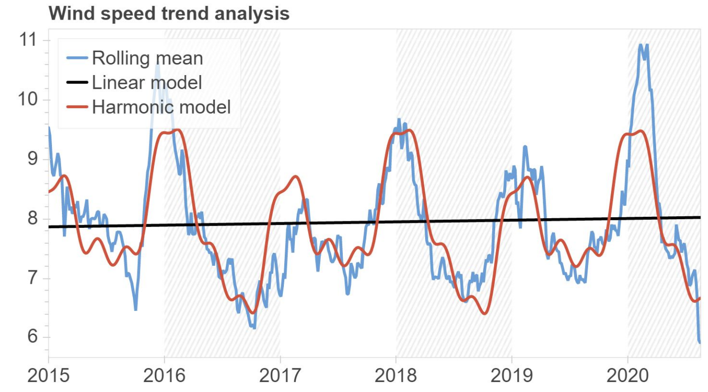


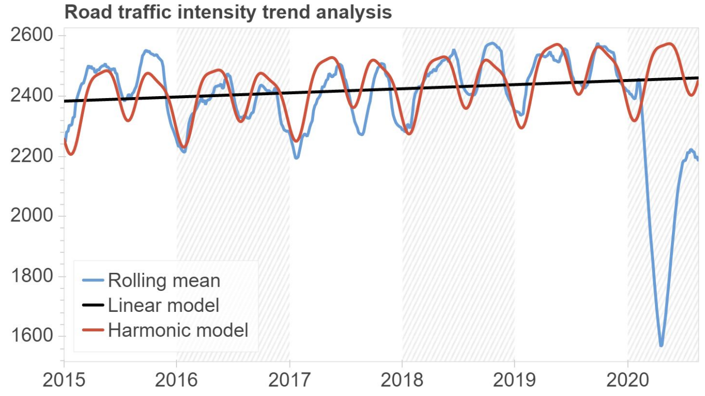
Road traffic intensity is **not** becoming more important during COVID-19 lockdowns in predicting NO2 -> Its causality remains unclear

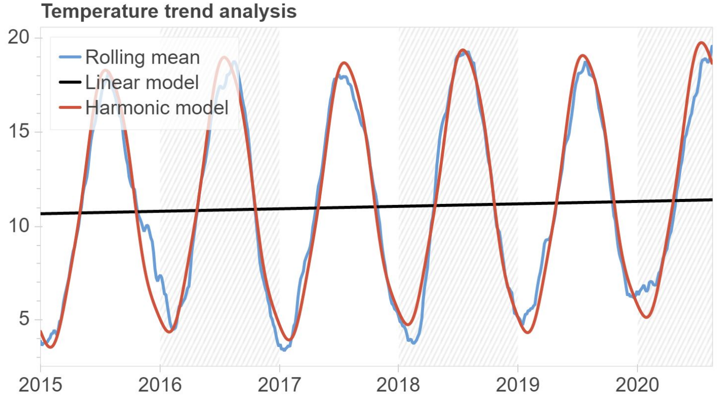
Results | Prediction accuracy - satellite data [extra]

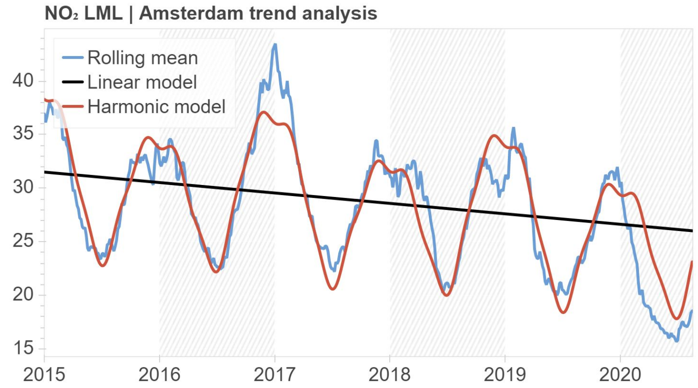
r	r^2
+0.98%	*+1.60%*

Results | Trend analysis - Harmonic model

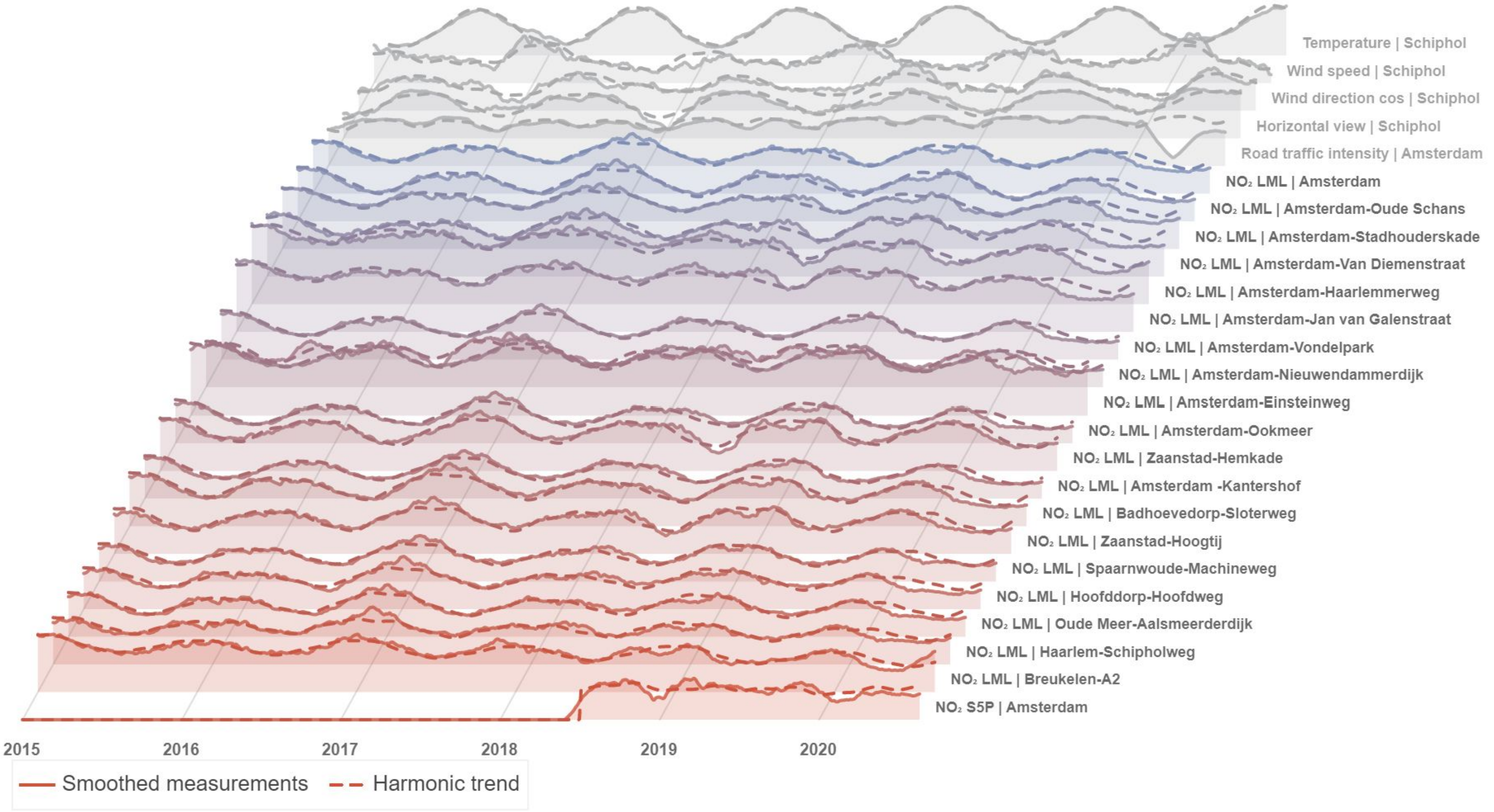




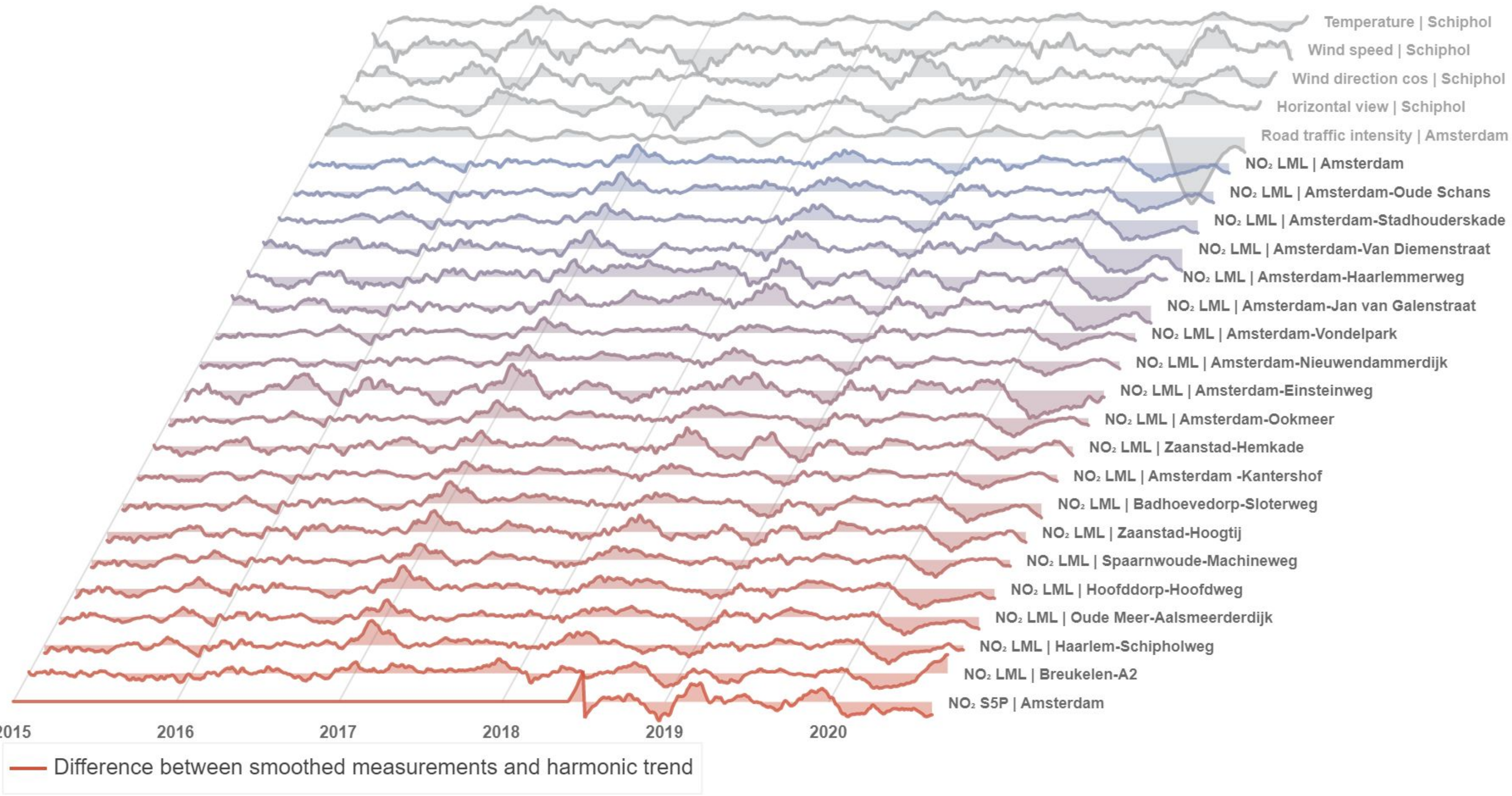




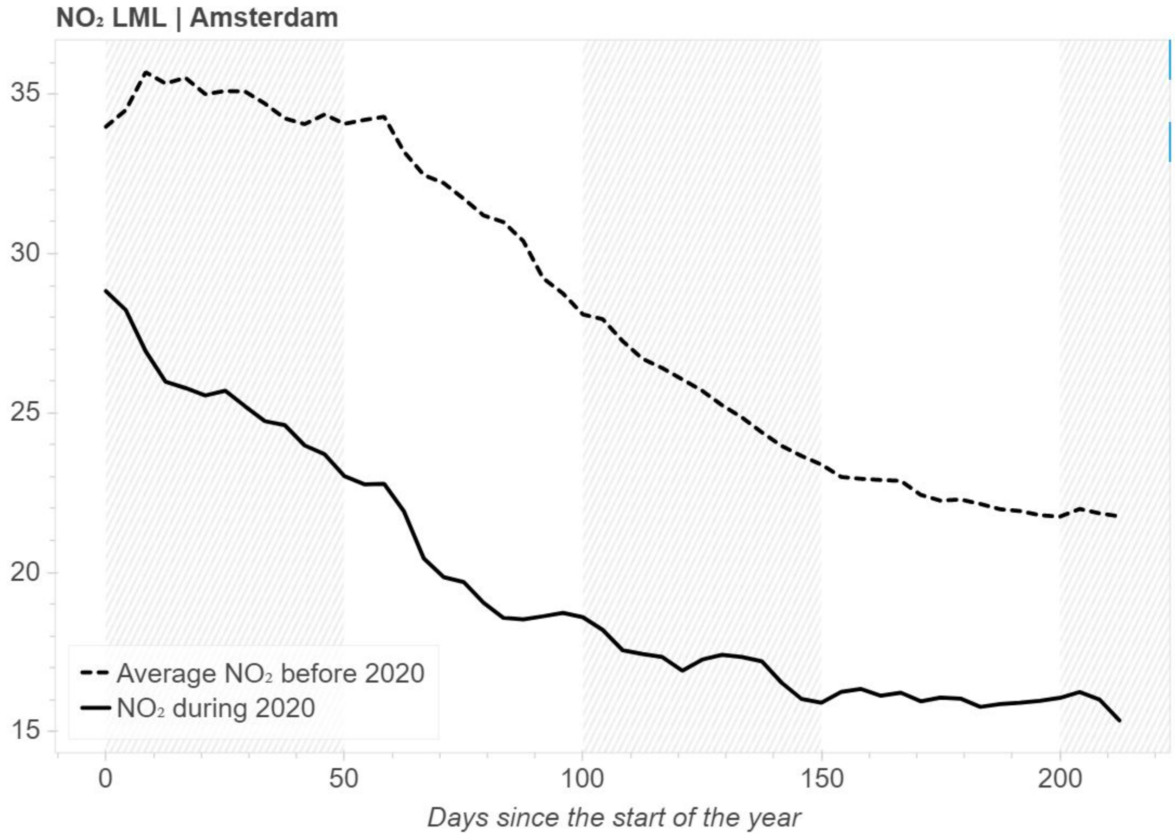
Results | Trend analysis - Harmonic model



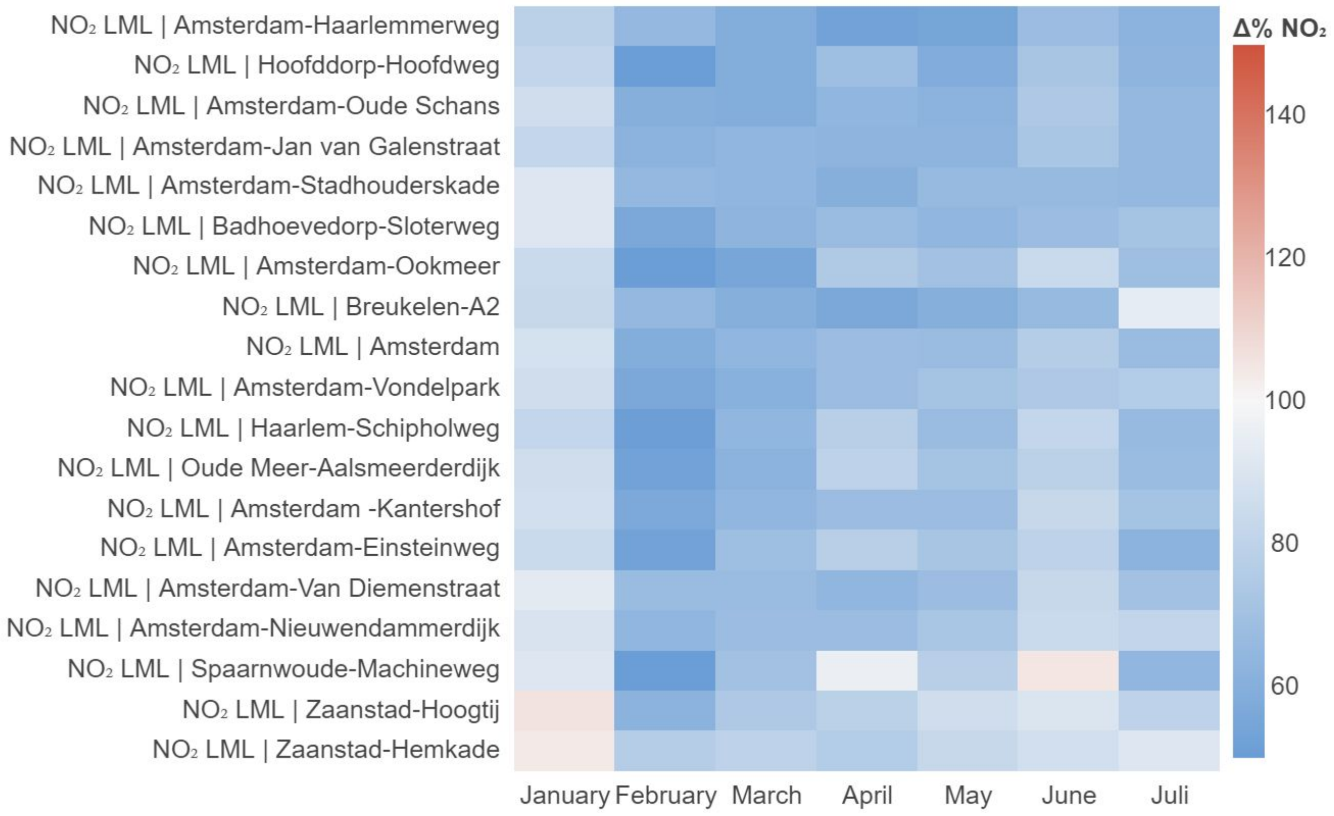
Results | Trend analysis - Harmonic model



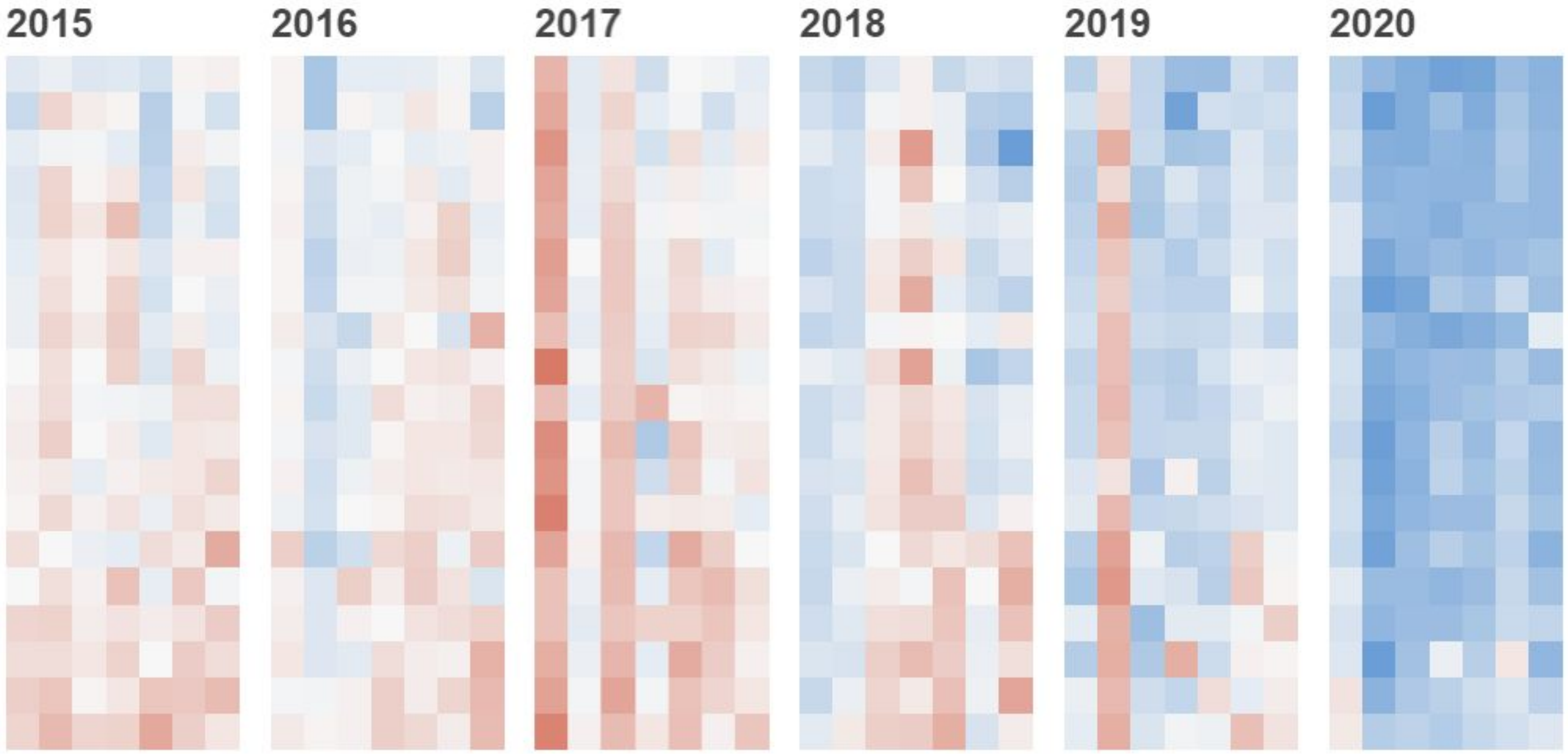
Results | Trend analysis - Overlap comparison



Results | Trend analysis - Overlap comparison



Results | Trend analysis - Overlap comparison



Results | Trend analysis - Overlap comparison

Location	January	February	March	April	May	June	Juli	Total mean
NO_2 LML | Amsterdam-Haarlemmerweg	79.34%	65.22%	59.39%	53.11%	54.96%	68.60%	62.28%	**63.27%**
NO_2 LML | Hoofddorp-Hoofdweg	81.10%	47.12%	59.66%	69.86%	58.85%	72.80%	63.01%	**64.63%**
NO_2 LML | Amsterdam-Oude Schans	86.05%	60.49%	59.56%	64.48%	62.91%	74.69%	65.21%	**67.63%**
NO_2 LML | Amsterdam-Jan van Galenstraat	82.51%	62.69%	64.01%	63.59%	63.75%	73.09%	65.54%	**67.88%**
NO_2 LML | Amsterdam-Stadhouderskade	91.99%	65.42%	64.95%	60.28%	66.32%	66.49%	65.81%	**68.75%**
NO_2 LML | Badhoevedorp-Sloterweg	91.27%	56.13%	63.34%	67.13%	64.08%	68.98%	71.16%	**68.87%**
NO_2 LML | Amsterdam-Ookmeer	84.72%	46.27%	55.10%	75.07%	70.46%	84.37%	69.52%	**69.36%**
NO_2 LML | Breukelen-A2	83.66%	65.25%	60.10%	56.22%	60.80%	66.42%	94.88%	**69.62%**
NO_2 LML | Amsterdam	88.12%	59.91%	64.87%	68.43%	67.27%	77.10%	67.70%	**70.49%**
NO_2 LML | Amsterdam-Vondelpark	86.03%	56.44%	61.68%	68.74%	71.88%	74.14%	76.41%	**70.76%**
NO_2 LML | Haarlem-Schipholweg	82.77%	51.45%	64.84%	78.78%	67.71%	82.96%	66.99%	**70.78%**
NO_2 LML | Oude Meer-Aalsmeerderdijk	86.85%	53.17%	62.03%	80.09%	71.92%	79.31%	67.57%	**71.56%**
NO_2 LML | Amsterdam -Kantershof	87.19%	57.90%	64.28%	68.37%	68.53%	83.43%	71.66%	**71.62%**
NO_2 LML | Amsterdam-Einsteinweg	84.61%	53.97%	69.48%	78.46%	72.61%	80.75%	62.26%	**71.73%**
NO_2 LML | Amsterdam-Van Diemenstraat	93.06%	67.45%	67.39%	65.00%	68.48%	83.92%	70.14%	**73.64%**
NO_2 LML | Amsterdam-Nieuwendammerdijk	89.81%	64.24%	68.46%	68.15%	73.18%	84.19%	81.38%	**75.63%**
NO_2 LML | Spaarnwoude-Machineweg	91.15%	45.50%	70.13%	96.13%	78.45%	105.05%	64.75%	**78.74%**
NO_2 LML | Zaanstad-Hoogtij	106.00%	62.87%	74.88%	79.53%	86.24%	90.02%	80.02%	**82.79%**
NO_2 LML | Zaanstad-Hemkade	104.93%	77.84%	80.36%	76.59%	83.37%	87.62%	91.94%	**86.09%**
NO_2 | Amsterdam	**88.48%**	**58.91%**	**64.97%**	**70.42%**	**69.04%**	**79.16%**	**71.48%**	**71.78%**
Road traffic intensity | Amsterdam	*109.65%*	*104.26%*	*78.32%*	*59.91%*	*73.08%*	*84.06%*	*93.44%*	**86.10%**
Wind speed | Schiphol	*99.10%*	*158.98%*	*116.60%*	*96.21%*	*104.46%*	*96.76%*	*111.72%*	**111.97%**

Results | Trend analysis - Overlap comparison

Location	January	February	March	April	May	June	Juli	Total mean
NO_2 LML | Amsterdam-Haarlemmerweg	79.34%	65.22%	59.39%	53.11%	54.96%	68.60%	62.28%	**63.27%**
NO_2 LML | Hoofddorp-Hoofdweg	81.10%	47.12%	59.66%	69.86%	58.85%	72.80%	63.01%	**64.63%**
NO_2 LML | Amsterdam-Oude Schans	86.05%	60.49%	59.56%	64.48%	62.91%	74.69%	65.21%	**67.63%**
NO_2 LML | Amsterdam-Jan van Galenstraat	82.51%	62.69%	64.01%	63.59%	63.75%	73.09%	65.54%	**67.88%**
NO_2 LML | Amsterdam-Stadhouderskade	91.99%	65.42%	64.95%	60.28%	66.32%	66.49%	65.81%	**68.75%**
NO_2 LML | Badhoevedorp-Sloterweg	91.27%	56.13%	63.34%	67.13%	64.08%	68.98%	71.16%	**68.87%**
NO_2 LML | Amsterdam-Ookmeer	84.72%	46.27%	55.10%	75.07%	70.46%	84.37%	69.52%	**69.36%**
NO_2 LML | Breukelen-A2	83.66%	65.25%	60.10%	56.22%	60.80%	66.42%	94.88%	**69.62%**
NO_2 LML | Amsterdam	88.12%	59.91%	64.87%	68.43%	67.27%	77.10%	67.70%	**70.49%**
NO_2 LML | Amsterdam-Vondelpark	86.03%	56.44%	61.68%	68.74%	71.88%	74.14%	76.41%	**70.76%**
NO_2 LML | Haarlem-Schipholweg	82.77%	51.45%	64.84%	78.78%	67.71%	82.96%	66.99%	**70.78%**
NO_2 LML | Oude Meer-Aalsmeerderdijk	86.85%	53.17%	62.03%	80.09%	71.92%	79.31%	67.57%	**71.56%**
NO_2 LML | Amsterdam -Kantershof	87.19%	57.90%	64.28%	68.37%	68.53%	83.43%	71.66%	**71.62%**
NO_2 LML | Amsterdam-Einsteinweg	84.61%	53.97%	69.48%	78.46%	72.61%	80.75%	62.26%	**71.73%**
NO_2 LML | Amsterdam-Van Diemenstraat	93.06%	67.45%	67.39%	65.00%	68.48%	83.92%	70.14%	**73.64%**
NO_2 LML | Amsterdam-Nieuwendammerdijk	89.81%	64.24%	68.46%	68.15%	73.18%	84.19%	81.38%	**75.63%**
NO_2 LML | Spaarnwoude-Machineweg	91.15%	45.50%	70.13%	96.13%	78.45%	105.05%	64.75%	**78.74%**
NO_2 LML | Zaanstad-Hoogtij	106.00%	62.87%	74.88%	79.53%	86.24%	90.02%	80.02%	**82.79%**
NO_2 LML | Zaanstad-Hemkade	104.93%	77.84%	80.36%	76.59%	83.37%	87.62%	91.94%	**86.09%**
NO_2 | Amsterdam	**88.48%**	**58.91%**	**64.97%**	**70.42%**	**69.04%**	**79.16%**	**71.48%**	**71.78%**
Road traffic intensity | Amsterdam	*109.65%*	*104.26%*	*78.32%*	*59.91%*	*73.08%*	*84.06%*	*93.44%*	**86.10%**
Wind speed | Schiphol	*99.10%*	*158.98%*	*116.60%*	*96.21%*	*104.46%*	*96.76%*	*111.72%*	**111.97%**

Results | Trend analysis - Overlap comparison

Location	January	February	March	April	May	June	Juli
NO_2 | Amsterdam	**88.48%**	**58.91%**	**64.97%**	**70.42%**	**69.04%**	**79.16%**	**71.48%**
Road traffic intensity | Amsterdam	*109.65%*	*104.26%*	*78.32%*	*59.91%*	*73.08%*	*84.06%*	*93.44%*
Wind speed | Schiphol	*99.10%*	*158.98%*	*116.60%*	*96.21%*	*104.46%*	*96.76%*	*111.72%*

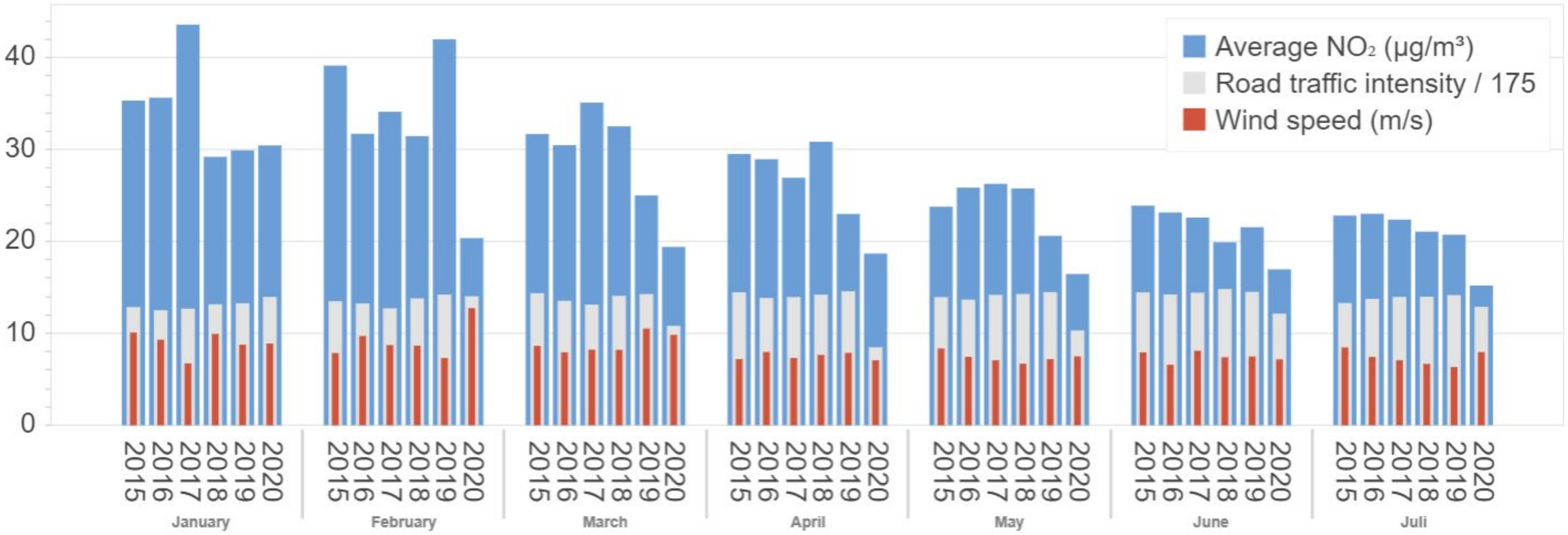
Results | Trend analysis - Overlap comparison

Location	January	February	March	April	May	June	Juli
NO_2 | Amsterdam	**88.48%**	**58.91%**	**64.97%**	**70.42%**	**69.04%**	**79.16%**	**71.48%**
Road traffic intensity | Amsterdam	*109.65%*	*104.26%*	*78.32%*	*59.91%*	*73.08%*	*84.06%*	*93.44%*
Wind speed | Schiphol	*99.10%*	*158.98%*	*116.60%*	*96.21%*	*104.46%*	*96.76%*	*111.72%*

Results | Trend analysis - Overlap comparison

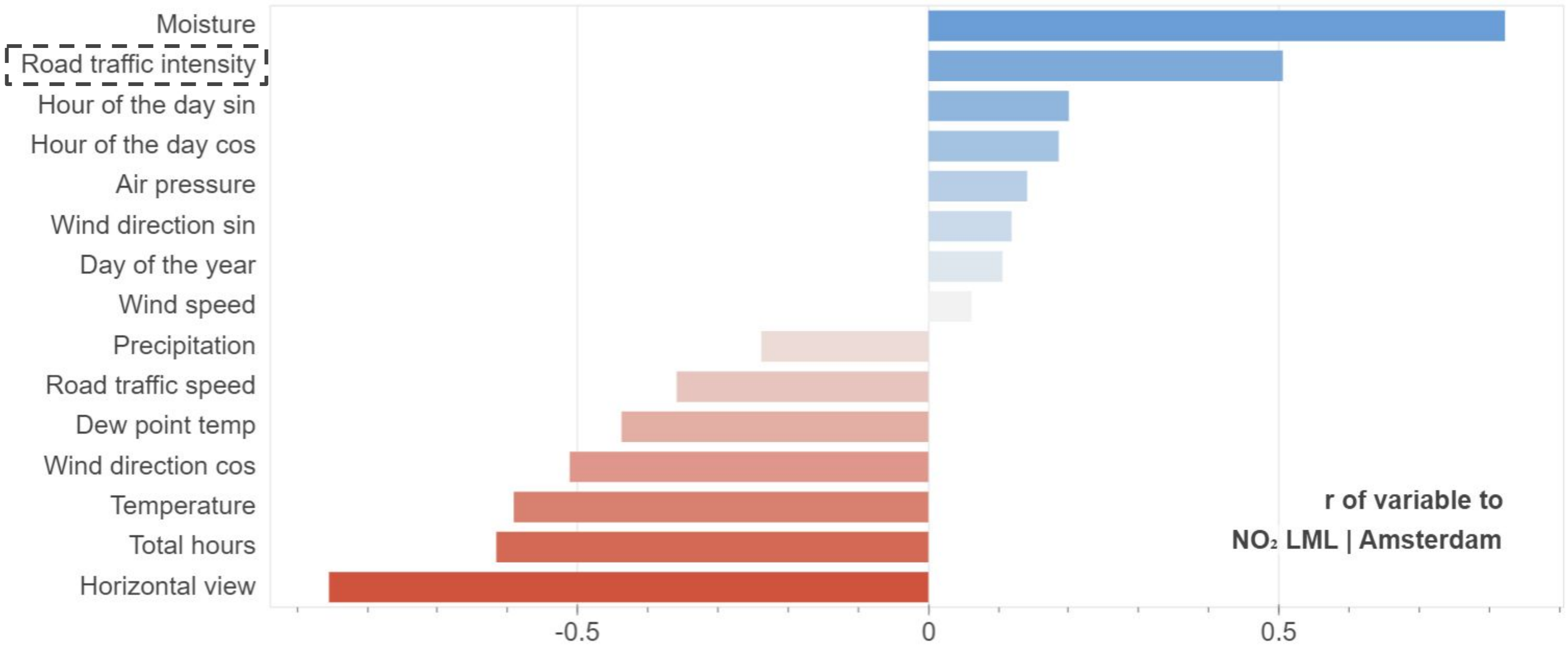
Location	January	February	March	April	May	June	Juli
NO_2 | Amsterdam	**88.48%**	**58.91%**	**64.97%**	**70.42%**	**69.04%**	**79.16%**	**71.48%**
Road traffic intensity | Amsterdam	*109.65%*	*104.26%*	*78.32%*	*59.91%*	*73.08%*	*84.06%*	*93.44%*
Wind speed | Schiphol	*99.10%*	*158.98%*	*116.60%*	*96.21%*	*104.46%*	*96.76%*	*111.72%*

Results | Trend analysis - Long term



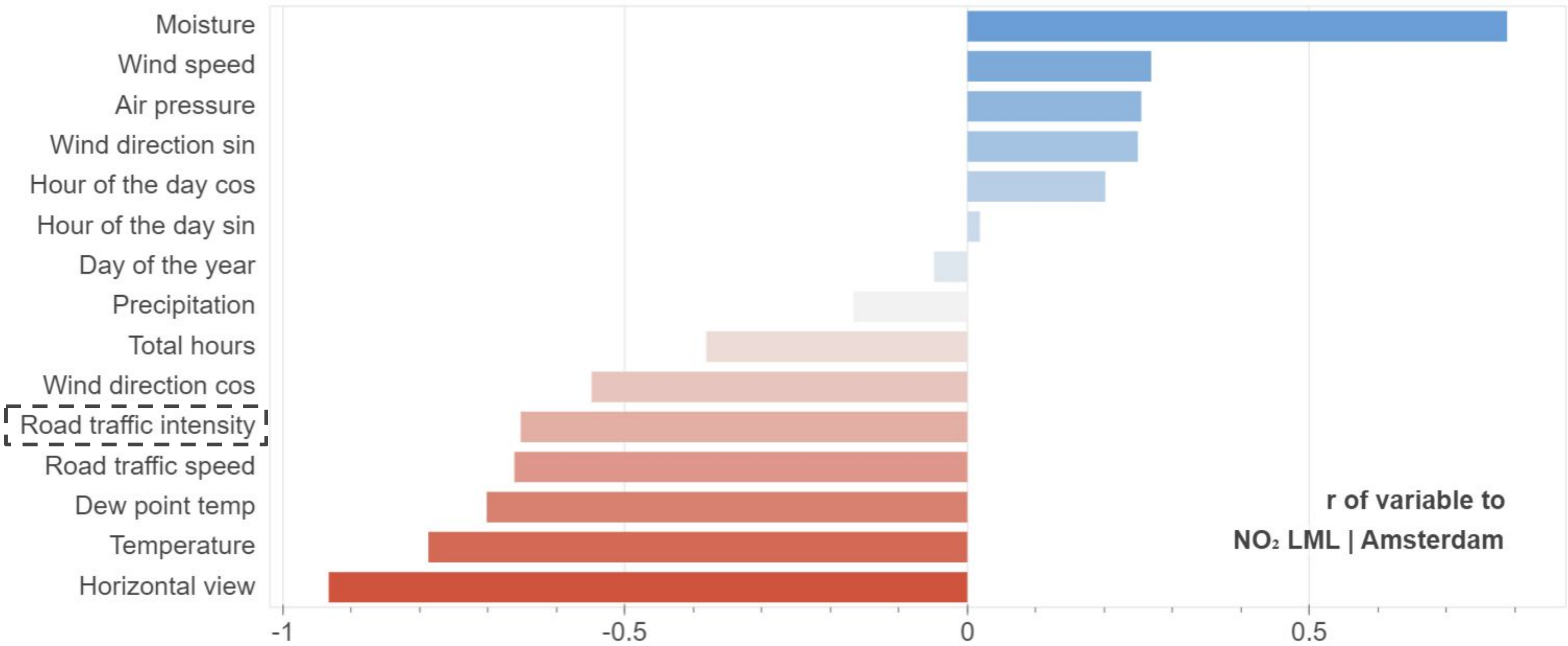
Results | Trend analysis - Long term

Including COVID-19 data



Results | Trend analysis - Long term

Excluding COVID-19 data



Results | Summary

- Preprocessing **helps**!
- **Weak** short term linear correlations
- Predictions are **relatively accurate** & provide **more insight** in the weather + road traffic influence
    - **Strong importance** weather -> wind speed / direction
    - Road traffic intensity has some importance, but hard to **verify causality**
    - Excluding COVID-19 data slightly **improves** prediction accuracy
        - But does **not** alter road traffic intensity importance -> Partially verifying **lack** of causality
    - Satellite data slightly improves prediction accuracy [extra]
- NO2 and road traffic are **lower** than the seasonal trend during COVID-19 lockdowns
- **Stronger** long term linear correlations
    - Road traffic intensity has a **negative** correlation before COVID-19 lockdowns
    - But a **positive** correlation during COVID-19 lockdowns
        - What can this tell us??
            - **Independent** behavior?
            - Needs **further research** -> Sub populations + Accounting for residuals

Discussion + Conclusion

NO2 was lower than normal during COVID-19 lockdowns in Amsterdam

So was the road traffic intensity, but it is still the question if this caused the lower NO2

NO2 and road traffic intensity may have independently reacted to other (human) factors

This is supported by the fact that the weather did not change from the normal behavior

Further research should look into these potential other factors

But also further investigate the road traffic effects on NO2 as this is easily quantifiable

These findings may advance the additional usefulness of these open data sources

And hopefully contribute to a cleaner air by improving local decision making

Questions?