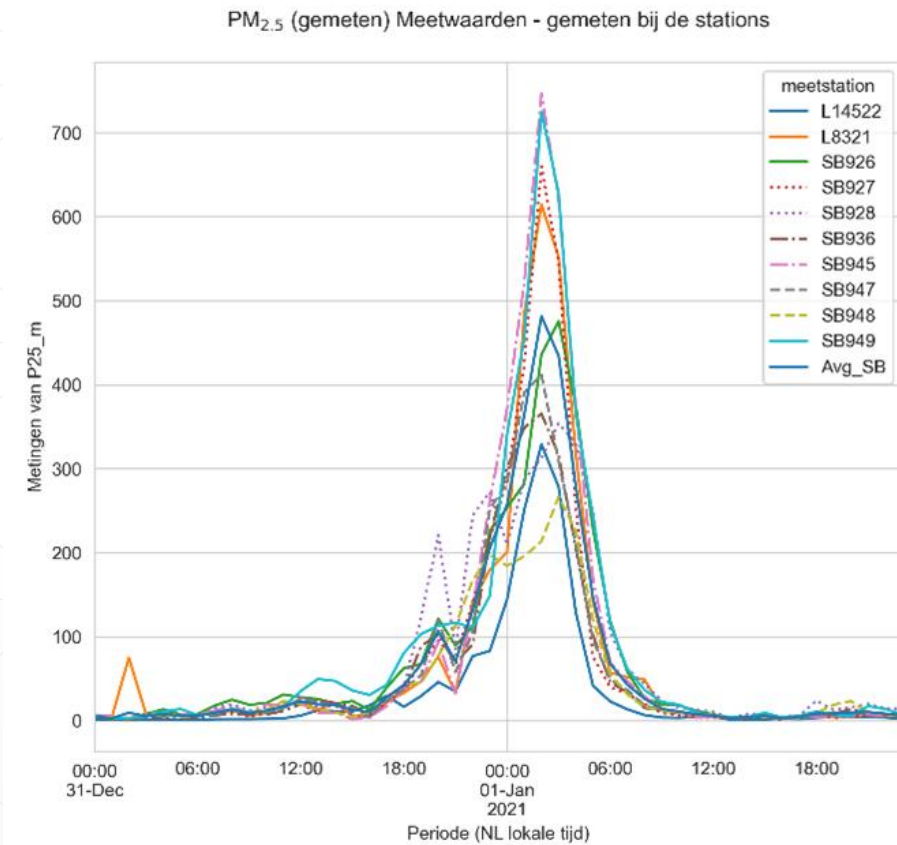ALVA –
software for display and analysis of air pollution data

LV2
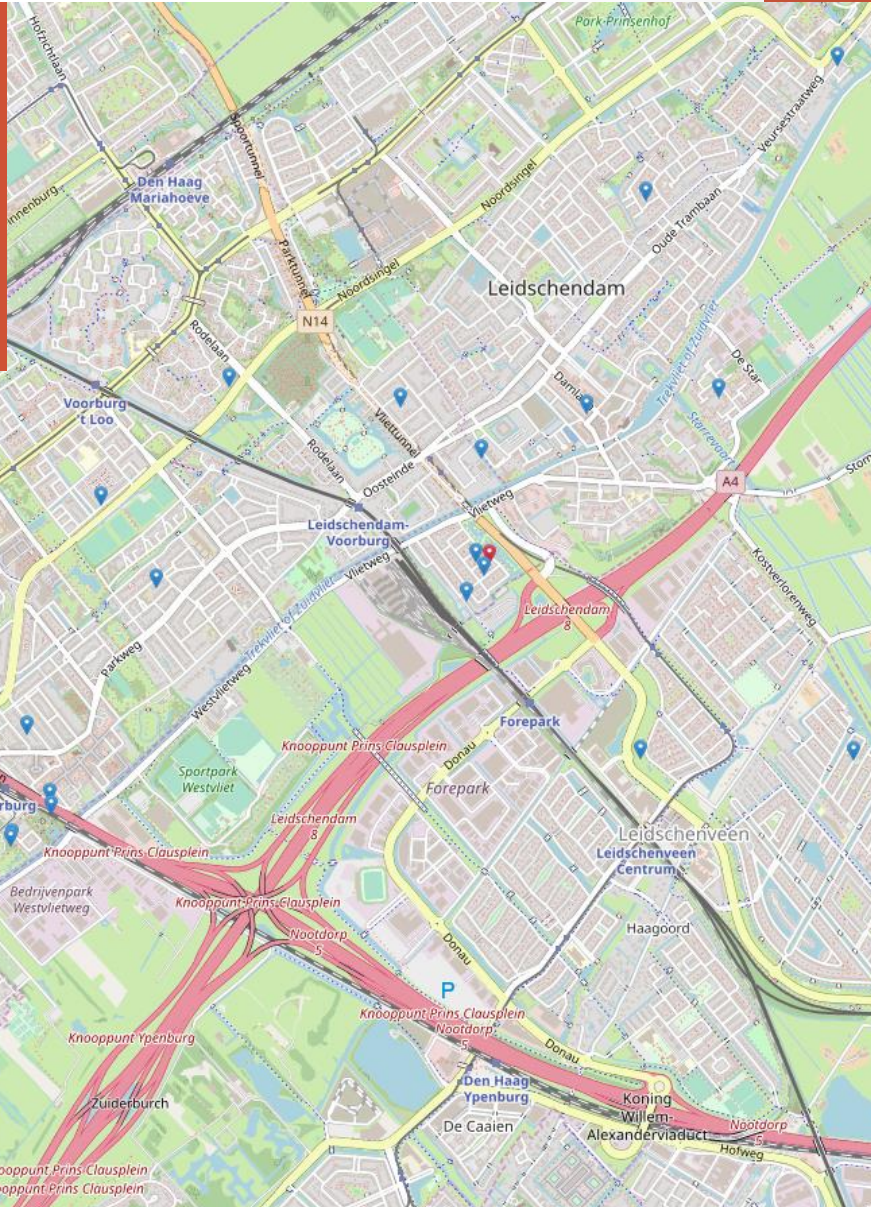
Lucht voor Leidschendam-Voorburg   Frans Kets

PM₂.₅ (gemeten) Meetwaarden - gemeten bij de stations

2.600.000 carkilometers per day
in the municipality,
79% is transit traffic

A4 and A12 responsible for 57%
of the CO_2
(source – MRDH)

**Trigger (2017):
construction
Rotterdamse Baan**

**"*Nederland wordt schoner*"
Also valid for
Leidschendam-Voorburg?**

Pollution Dome The Hague

Construction ongoing in
Binckhorst, CID

Proximity A4

Roads into The Hague ("inprikkers")

    A12 – Utrechtse Baan

    A13 – Rotterdamse Baan

    N14

    S100 - Supernovaweg

Levels of Citizen Science

Level 4 'Extreme'	• Collaborative Science – problem definition, data collection and analysis
Level 3 'Participatory science'	• Participation in problem definition and data collection
Level 2 'Distributed Intelligence'	• Citizens as basic interpreters
Level 1 'Crowdsourcing'	• Citizens as sensors

COMMUNITY PARTICIPATION

Source: Presentation Claire Ellul, University College London
"Citizen Science and the Science of Cities"

Motivation of the Citizen Scientist?

Short term
- Action oriented

Long term
- Learning oriented
  - Pollution types
  - Atmospheric chemistry
  - Health effects
  - *Relationships between these three*
- Essential: Growpath towards the long term goal

Levels of Citizen Science

Level 4 'Extreme'	• Collaborative Science – problem definition, data collection and analysis
Level 3 'Participatory science'	• Participation in problem definition and data collection
Level 2 'Distributed Intelligence'	• Citizens as basic interpreters
Level 1 'Crowdsourcing'	• Citizens as sensors

COMMUNITY PARTICIPATION

Source: Presentation Claire Ellul, University College London
"Citizen Science and the Science of Cities"

Can we observe changes in traffic patterns?

OUR
Problem Definition

Motivation of the Citizen Scientist?

Short term
• Action oriented

Long term
• Learning oriented
    • Pollution types
    • Atmospheric chemistry
    • Health effects
    • *Relationships between these three*
• Essential: Growpath towards the long term goal

Levels of Citizen Science

Level 4 'Extreme'	• Collaborative Science – problem definition, data collection and analysis
Level 3 'Participatory science'	• Participation in problem definition and data collection
Level 2 'Distributed Intelligence'	• Citizens as basic interpreters
Level 1 'Crowdsourcing'	• Citizens as sensors

COMMUNITY PARTICIPATION

Source: Presentation Claire Ellul, University College London "Citizen Science and the Science of Cities"

OUR
Problem Definition

Can we observe changes in traffic patterns?

Where do peaks in pollution come from?

Citizen Scientist?

Short term
• Action oriented

Long term
• Learning oriented
    • Pollution types
    • Atmospheric chemistry
    • Health effects
    • *Relationships between these three*
• Essential: Growpath towards the long term goal

Levels of Citizen Science

Level 4 'Extreme'	• Collaborative Science – problem definition, data collection and analysis
Level 3 'Participatory science'	• Participation in problem definition and data collection
Level 2 'Distributed Intelligence'	• Citizens as basic interpreters
Level 1 'Crowdsourcing'	• Citizens as sensors

COMMUNITY PARTICIPATION

Source: Presentation Claire Ellul, University College London "Citizen Science and the Science of Cities"

OUR Problem Definition

Can we observe changes in traffic patterns?

What do peaks in pollution come from?

Citizen Scientist?

Can we reliably check on adherence to EU and WHO bounds?

- Action oriented

Long term

- Learning oriented
  - Pollution types
  - Atmospheric chemistry
  - Health effects
  - *Relationships between these three*
- Essential: Growpath towards the long term goal

Levels of Citizen Science

Level 4 'Extreme'	• Collaborative Science – problem definition, data collection and analysis
Level 3 'Participatory science'	• Participation in problem definition and data collection
Level 2 'Distributed Intelligence'	• Citizens as basic interpreters
Level 1 'Crowdsourcing'	• Citizens as sensors

COMMUNITY PARTICIPATION

Source: Presentation Claire Ellul, University College London "Citizen Science and the Science of Cities"

OUR
Problem Definition

Can we observe changes in traffic patterns?

What do peaks in pollution come from?

Citizen Scientist?

Can we reliably check on adherence to EU and WHO bounds?

• Action oriented

Do we trust the measurements and the calibration procedures applied?

Learning oriented

• Pollution types

• Atmospheric chemistry

• Health effects

• *Relationships between these three*

• Essential: Growpath towards the long term goal

Levels of Citizen Science

Level 4 'Extreme'	• Collaborative Science – problem definition, data collection and analysis
Level 3 'Participatory science'	• Participation in problem definition and data collection
Level 2 'Distributed Intelligence'	• Citizens as basic interpreters
Level 1 'Crowdsourcing'	• Citizens as sensors

COMMUNITY PARTICIPATION

Source: Presentation Claire Ellul, University College London "Citizen Science and the Science of Cities"

Can we observe changes in traffic patterns?

OUR Problem Definition

What do peaks in pollution come from?

Citizen Scientist?

Can we reliably check on adherence to EU and WHO bounds?

• Action oriented

Do we trust the measurements and the calibration procedures applied?

Learning oriented

Can we separate source and propagation / distribution effects ?

• Health effects

• *Relationships between these three*

• Essential: Growpath towards the long term goal

Levels of Citizen Science

Level 4 'Extreme'	• Collaborative Science – problem definition, data collection and analysis
Level 3 'Participatory science'	• Participation in problem definition and data collection
Level 2 'Distributed Intelligence'	• Citizens as basic interpreters
Level 1 'Crowdsourcing'	• Citizens as sensors

COMMUNITY PARTICIPATION

Source: Presentation Claire Ellul, University College London
"Citizen Science and the Science of Cities"

OUR
Problem Definition

Can we observe changes in traffic patterns?

What do peaks in pollution come from?

Citizen Scientist?

Can we reliably check on adherence to EU and WHO bounds?

• Action oriented

Do we trust the measurements and the calibration procedures applied?

Learning oriented

Can we separate source and propagation / distribution effects ?

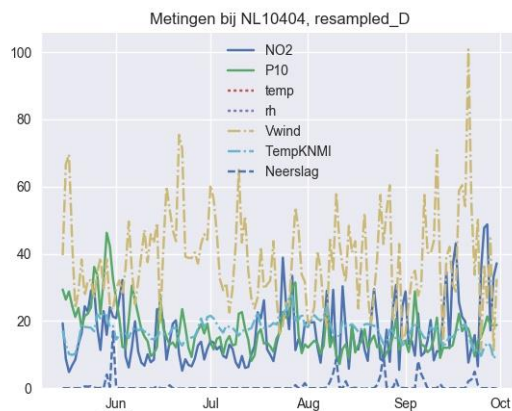• Health effects

What do directional plots tell us?

three

• Essential: Growpath towards the long term goal

Levels of Citizen Science

Level 4 'Extreme'	• Collaborative Science – problem definition, data collection and analysis
Level 3 'Participatory science'	• Participation in problem definition and data collection
Level 2 'Distributed Intelligence'	• Citizens as basic interpreters
Level 1 'Crowdsourcing'	• Citizens as sensors

COMMUNITY PARTICIPATION

Source: Presentation Claire Ellul, University College London "Citizen Science and the Science of Cities"

OUR Problem Definition

Can we observe changes in traffic patterns?

What do peaks in pollution come from?

Citizen Scientist?

Can we reliably check on adherence to EU and WHO bounds?

• Action oriented

Do we trust the measurements and the calibration procedures applied?

Learning oriented

Can we separate source and propagation / distribution effects ?

• Health effects

What do directional plots tell us?

three

Can we see on our data: effects covid lockdown; fireworks ban?

The prototype I

samenmeten.rivm.nl

csv files

knmi.nl

Python, pandas, seaborn

Metingen bij NL10404, resampled_D

Vergelijking NO2 per D bij verschillende stations

Gemiddelde verontreiniging per uur gedurende dag @NL10404

Vergelijking temp / rh; samples per D

September 2018

Gemiddelde verontreiniging gedurende week @SPPS014
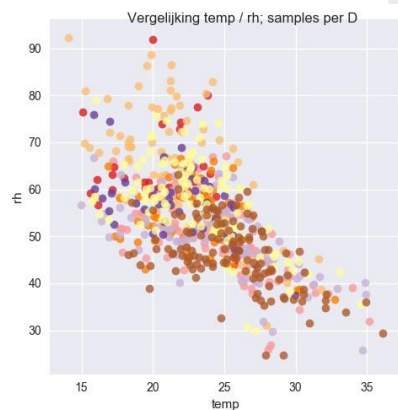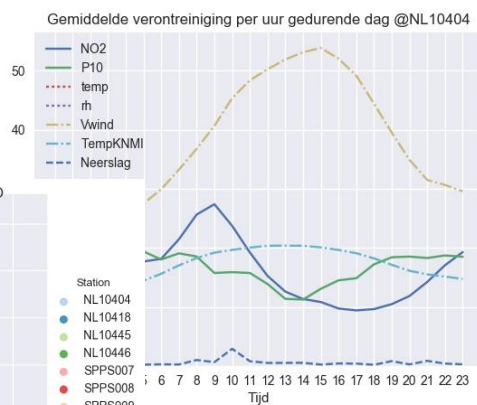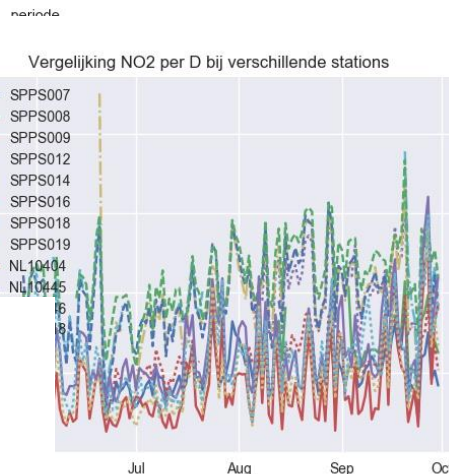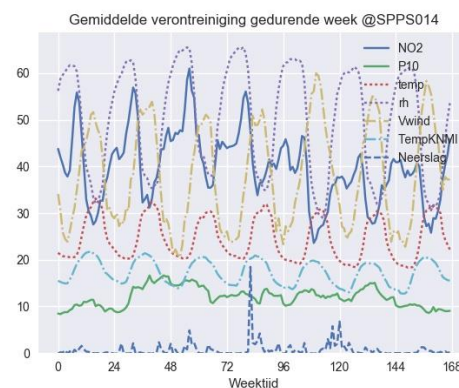
First approach

Many displays at once

Are measurements valid?

Do different stations give similar messages?

Is there a functional dependence of pollution measurements with time of the day, season?

Issues

Prototype

datastructure reordered by sorting operations, given different demands

No flexibility in sequence of operations

Every new demand added complexity --------→ maintainability

Little or no numerical output

Lots of csv files to be read

Carddeck for user input

The prototype II

September 2019

```
FILE LTD 5 2019 10 12
FILE LTD 6 2019 10 12
FILE LTD 7 2019 10 12
FILE LTD 8 2019 10 12
FILE LTD 9 2019 10 12
FILE LTD 10 2019 10 12
FILE LTD 11 2019 10 12
FILE LTD 12 2019 10 12

RANGE 1/10/2019 31/10/2019 Periode - Oktober -  2019
GRENSWAARDEN 0 120 NaN
KNMI
RANGE4STATION CSLV 12 5/8/2018
RANGE4STATION CSLV 10 5/8/2018 24/01/2019
RANGE4STATION CSLV 29 25/1/2019
RANGE4STATION CSLV 31 25/1/2019
RANGE4STATION CSLV 34 1/9/2019
RANGE4STATION CSLV 36 15/7/2019
RANGE4STATION CLSV 40 15/6/2019
RANGE4STATION CSLV 43 15/5/2019
RANGE4STATION CSLV 52 1/6/2019
# SELECTPOLL_SNGL NO2 P10 P25 rh
SELECTPOLL_SNGL NO2 P10 P25 rh te
STATIONCOMPARE P10 NO2 P25 ALL
AVERAGE _NL0_2 _Andere RIVM 0 RIV
SCATTER NO2 P10 CSLV 7 CSLV 8 CSL
FUNCTIE UpD
FUNCTIE DpW
FUNCTIE Windkracht
```

Python, pandas, seaborn

Metingen p. windsnelheid @NO2
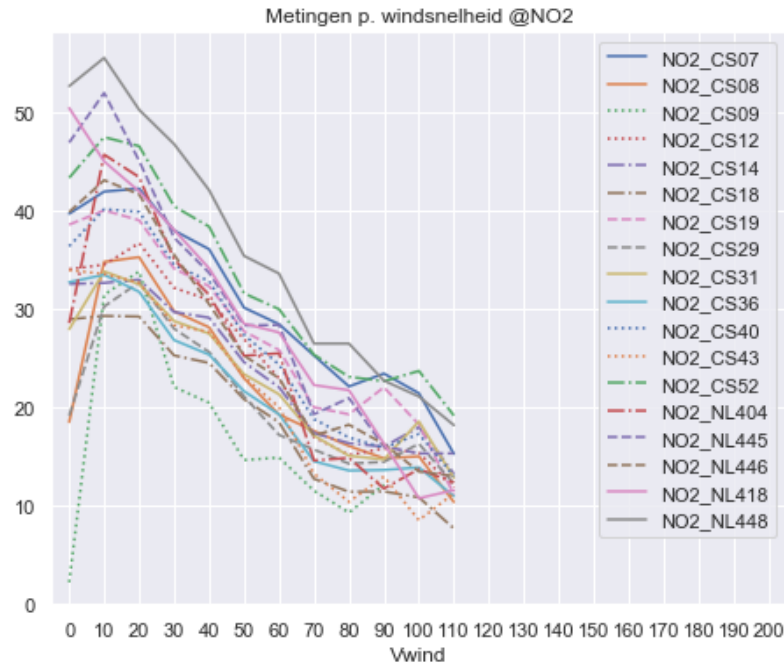
Periode - November - 2019
Metingen p. windsnelheid @NO2   Periode: 1/11/2019 - 30/11/2019
File:NO2_Windsnelheid   wk 44 - 48

Metingen	Gemiddeld	Maximum	Loc Max	Minimum	Loc Min	Mediaan	Aantal
NO2_CS07	30.3	42.2	(20)	15.3	(110)	29.3	12
NO2_CS08	21.9	35.3	(20)	10.4	(110)	18.9	12
NO2_CS09	17.3	33.8	(20)	2.2	(0)	14.8	10
NO2_CS12	24.4	36.7	(20)	11.9	(110)	24.9	12
NO2_CS14	23.7	33.0	(20)	13.3	(110)	23.3	12
NO2_CS18	19.2	29.3	(10)	7.7	(110)	19.7	12
NO2_CS19	27.4	40.0	(10)	11.2	(110)	26.8	12
NO2_CS29	20.6	32.7	(20)	12.2	(110)	18.2	12
NO2_CS31	22.8	33.8	(10)	12.8	(110)	22.4	12
NO2_CS36	21.5	33.5	(10)	11.0	(110)	20.5	12
NO2_CS40	26.4	40.1	(10)	13.0	(110)	25.9	12
NO2_CS43	21.3	33.9	(0)	8.5	(100)	21.6	12
NO2_CS52	32.6	47.5	(10)	19.2	(110)	30.8	12
O2_NL404	25.2	45.7	(10)	11.7	(90)	25.4	12
O2_NL445	29.9	52.0	(10)	15.3	(100)	28.3	12
O2_NL446	26.4	43.1	(10)	13.0	(110)	24.1	12
O2_NL418	29.0	50.4	(0)	10.8	(100)	28.1	12
O2_NL448	35.9	55.5	(10)	18.2	(110)	34.5	12

Expanding

Add statistical data to displays
mean
min
max
median
number of points displayed

Issues

Prototype

Little possibility for:
editing of time series
derivation of trends
manipulation of time series
adding time series

Always many displays generated

Harder and harder to maintain

Names of measurements hard coded

Carddeck for user input

ALVA 2 design principles

- **Datastructure**
  - Backbone of time (local time linked to UTC) + derived quantities + meteorological data (if supplied)
  - Sampling per hour in the backbone
  - Per workstation:
    - Measurements as function of time, and derived quantities
    - Double index per workstation and per measurement type
  - Datastructure expands and reduces according to data fill
    - Operations can add pseudo workstations, and can add time series to a workstation
  - Workstation metadata and measurement type (mostly) separate
    - Exception: routines for reading (some) csv files
- **Dataflow**
  - Checking of user input at the start
  - Order of operations as by the card sequence, but analysis operations, plotting and function derivation always at the end
    - Plug-in principle
  - Application domain can vary per card

Maintain audit trail
Generate numbers

ALVA 2.2 functionality

Selection, Manipulation, Analysis, Plotting

Data selection
 Selecting a subset of data from csv files
 New RIVM API supported (sw CS Gouda)

Data manipulation
 Mathematical manipulation,
 Selection of data based on logical conditions
 Filtering
 Counting hours/days of excess pollution

Add pseudo stations
 Results of time shift, mean, maths

Analysis and plotting
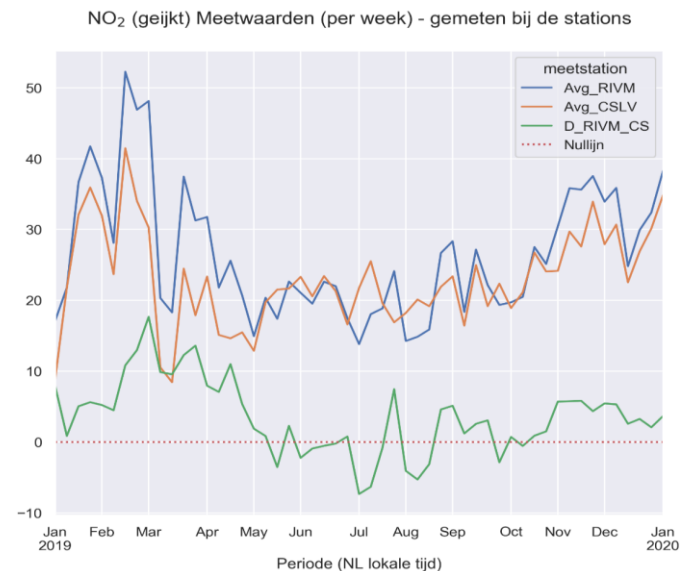 Direct comparisons,
 Crossplots
 Correlations
 1-D and 2-D function plots

ALVA 2.2 application 1


```
Bereken_stat D_RIVM_CS Avg_RIVM  Avg_CSLV -
Bereken_stat Nullijn 0
```


How do "paddenstoelen" compare with RIVM stations?

Combine the measurements of the paddenstoelen and store into one pseudostation

Combine the RIVM stations into another pseudostation

Subtract the values of the two pseudostations (math operations available [RPN])

Add zero line

Plot

HP35 (RPN calculator)
Fig.from Wikipedia

ALVA 2.2 application 2
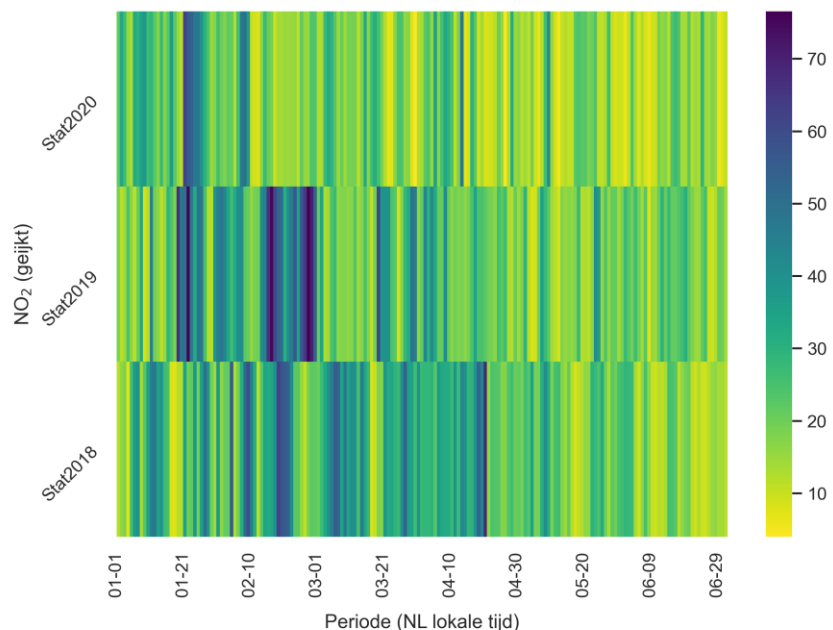
Can we observe Corona lockdown effects in our measurements?

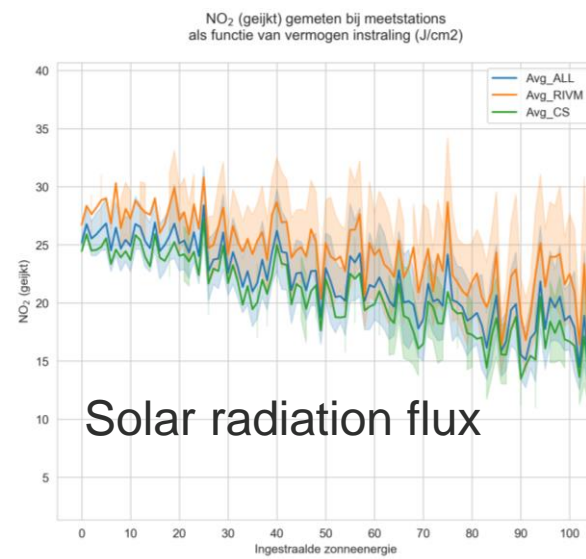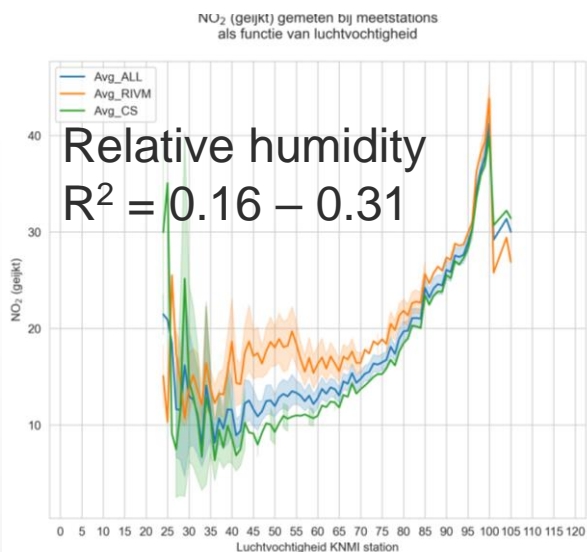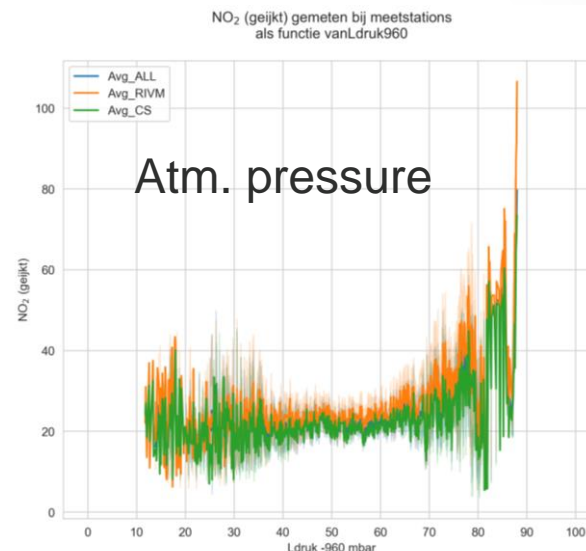*Comparison of 2018, 2019, 2020 NO$_2$ data from 3 RIVM stations in The Hague*
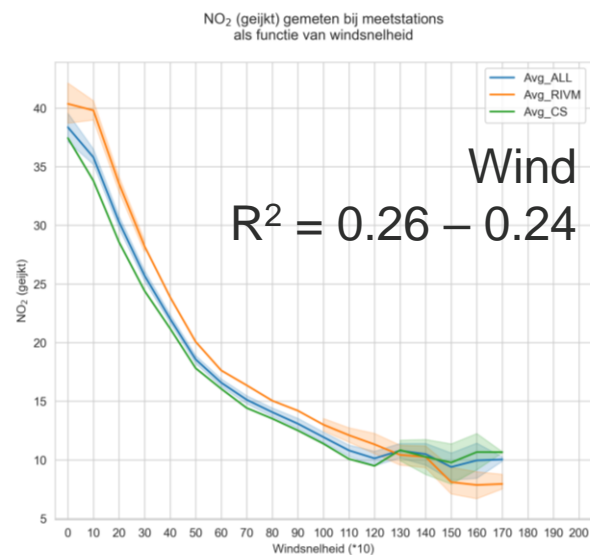
Processing

➢ Select period 2018 – 2020

➢ Calculate pseudo station: average of all curves of the three stations (Stat2020)

➢ Time shift the data forward by 2 respectively 1 year

➢ Create pseudostations Stat2018 and Stat2019

➢ Display for the period 1/1 – 30/6

➢ different sampling intervals (day and week)

Lockdown date was 15 March.
No significant change at that date

ALVA 2.2 application 3 - statistics

$$NO_{2,observed} = NO_{2,emitted} * f(Var\ 1)$$

NO₂ (geijkt) gemeten bij meetstations als functie van windsnelheid

Wind
R² = 0.26 − 0.24

NO₂ (geijkt) gemeten bij meetstations als functie vanLdruk960

Atm. pressure

NO₂ (geijkt) gemeten bij meetstations als functie van luchtvochtigheid

Relative humidity
R² = 0.16 − 0.31

NO₂ (geijkt) gemeten bij meetstations als functie van vermogen instraling (J/cm2)

Solar radiation flux

Can we observe Corona lockdown effects in our measurements?

Q1:
What is the relation between meteorological quantities and the observed NO₂ signal?

Assume:
NO₂ produced in large time period independent of the weather circumstances

Can one observe functions such as:
$$NO_{2,observed} = NO_{2,emitted} * f(Var\ 1, Var\ 2, ..)$$

for which we can invert?

Relationships are presumably complex, with issues such as collinearity, and time lags between effects.
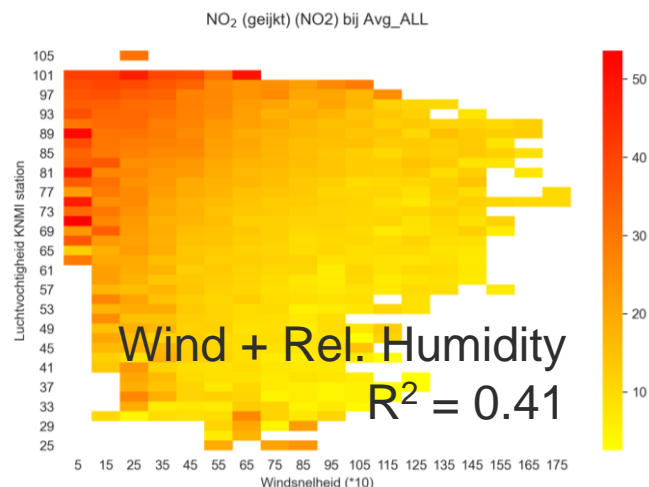
Limitation of scope:
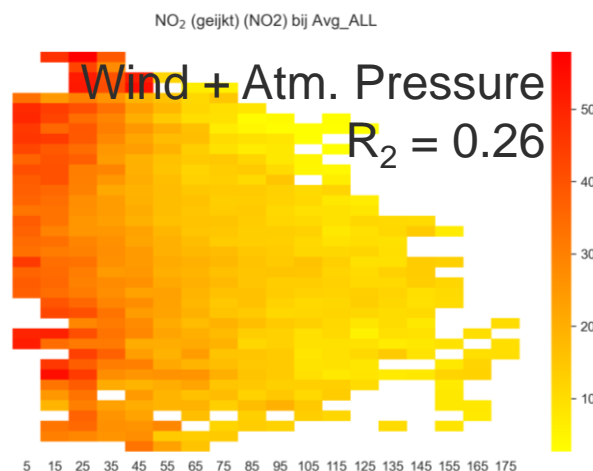*What are the most significant factors?*

ALVA 2.2 application 3 - statistics

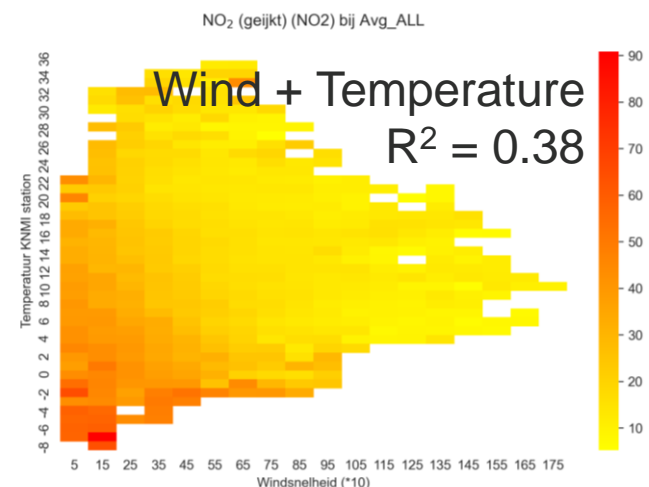$$NO_{2,observed} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Gemiddelde NO₂ (geijkt) per Luchtvochtigheid KNMI station / Windsnelheid (*10)
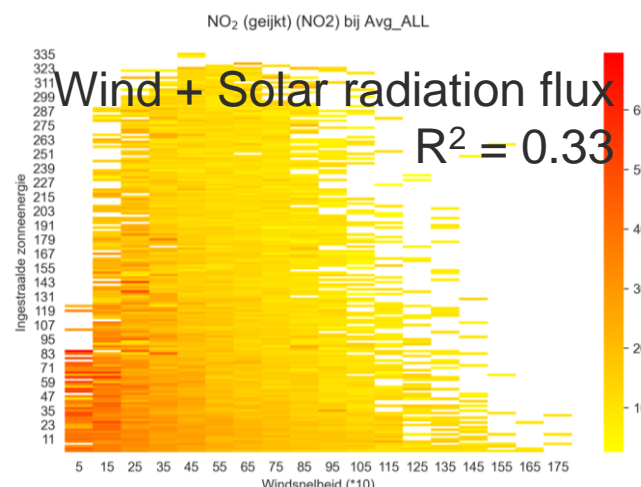
NO₂ (geijkt) (NO2) bij Avg_ALL

Wind + Rel. Humidity
$R^2 = 0.41$

Gemiddelde NO₂ (geijkt) per Ldruk -960 mbar / Windsnelheid (*10)

NO₂ (geijkt) (NO2) bij Avg_ALL

Wind + Atm. Pressure
$R_2 = 0.26$

Gemiddelde NO₂ (geijkt) per Temperatuur KNMI station / Windsnelheid (*10)

NO₂ (geijkt) (NO2) bij Avg_ALL

Wind + Temperature
$R^2 = 0.38$

Gemiddelde NO₂ (geijkt) per Ingestraalde zonneenergie / Windsnelheid (*10)

NO₂ (geijkt) (NO2) bij Avg_ALL

Wind + Solar radiation flux
$R^2 = 0.33$

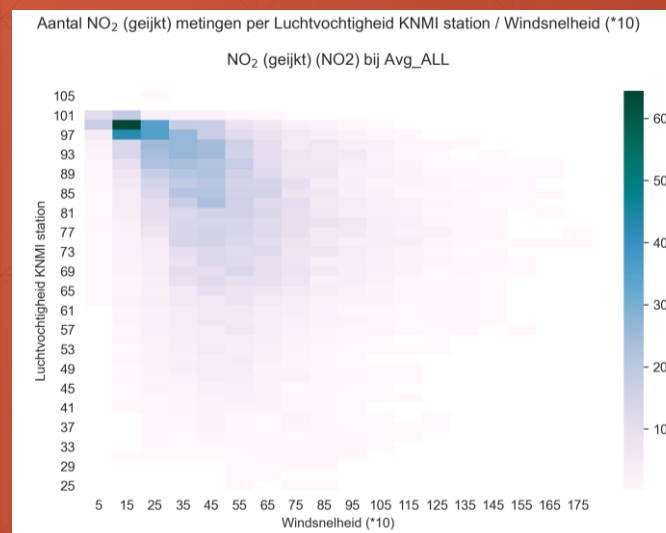## Can we observe Corona lockdown effects in our measurements?

Q2:
What is the relation between the observed NO₂ signal and two of the parameters?

R2 relates to a linear function of two parameters
Statistical weight:
# of measurement points per bin:

Aantal NO₂ (geijkt) metingen per Luchtvochtigheid KNMI station / Windsnelheid (*10)

NO₂ (geijkt) (NO2) bij Avg_ALL
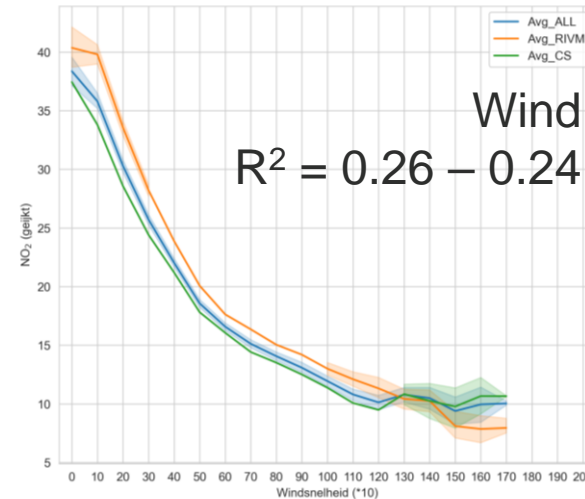
ALVA 2.2 application 3 - statistics
Inversion strategies

$$NO_{2,observed} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Gemiddelde NO$_2$ (geijkt) per Luchtvochtigheid KNMI station / Windsnelheid (*10)

NO$_2$ (geijkt) (NO2) bij Avg_ALL

Wind + Rel. Humidity
R^2 = 0.41

Luchtvochtigheid KNMI station

Windsnelheid (*10)

NO$_2$ (geijkt) gemeten bij meetstations als functie van windsnelheid

Wind
R^2 = 0.26 − 0.24

NO$_2$ (geijkt)

Windsnelheid (*10)

— Avg_ALL
— Avg_RIVM
— Avg_CS

Gemiddelde NO$_2$ (geijkt) per Temperatuur KNMI station / Windsnelheid (*10)

NO$_2$ (geijkt) (NO2) bij Avg_ALL

Wind + Temperature
R^2 = 0.38
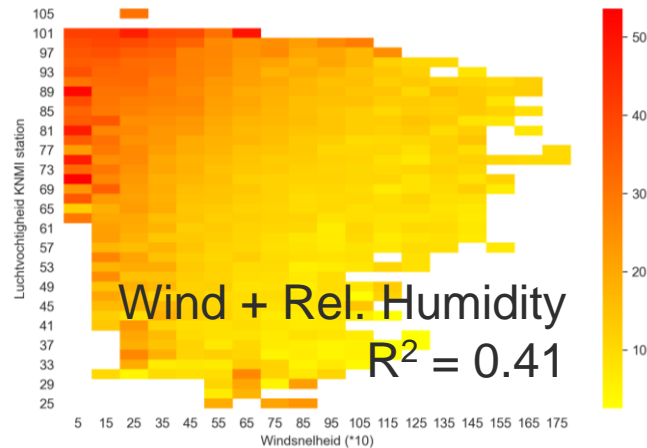
Temperatuur KNMI station

Windsnelheid (*10)
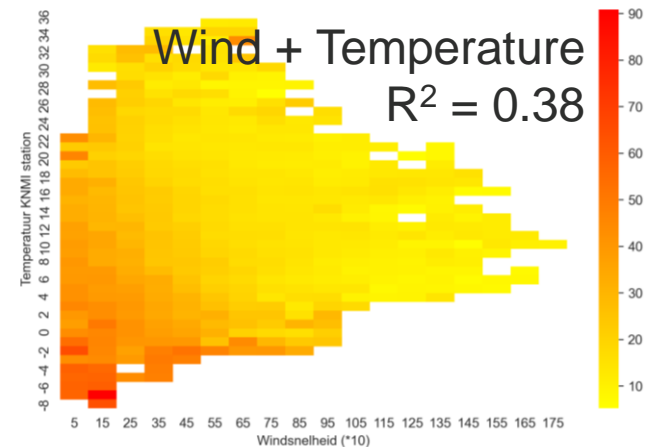
```
Bereken_attr    factor1    1 22 2000 / Vwindm *
Bereken_attr    factor2    24.25 40 / 8 64 / 40
Bereken_attr    factor3    .25 factor2 max ALL
Bereken_attr    NO2cor1    NO2med factor3 / ALL
Bereken_attr    expfactor 0.05 0.94 10 Vwindm −
Bereken_attr    NO2cor2    NO2med expfactor / AL
Bereken_attr    vochtfactor 0.73 0.38 LVochtm *
Bereken_attr    NO2cor3    NO2med vochtfactor /
Bereken_attr    tempfactor  42.7 0.76 Tempm *
Bereken_attr    NO2cor4    NO2med tempfactor / A
```

Can we observe Corona lockdown effects in our measurements?

Q3:
What is the inversion strategy?

$$NO_{2,obs} = NO_{2,emitted} * f(Var\ 1)$$

1: Relation described by three straight lines

2: Relation given by a function
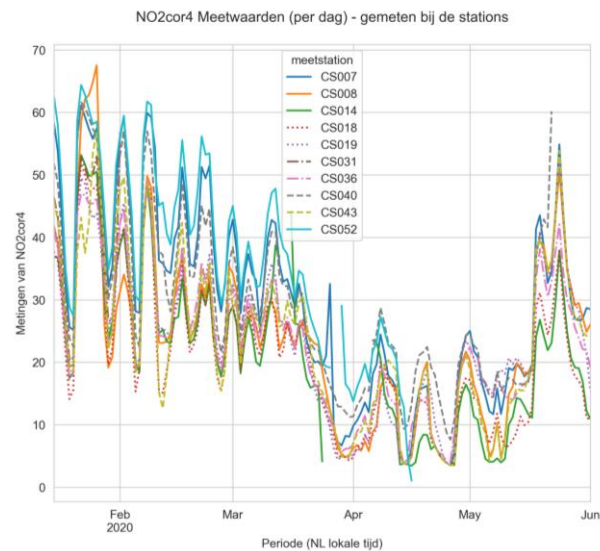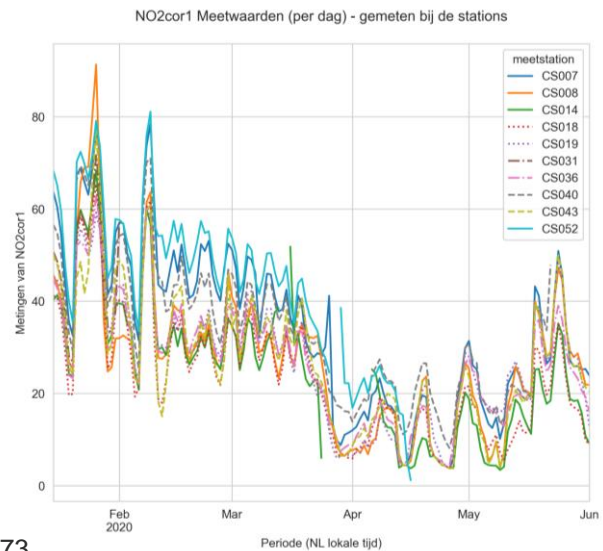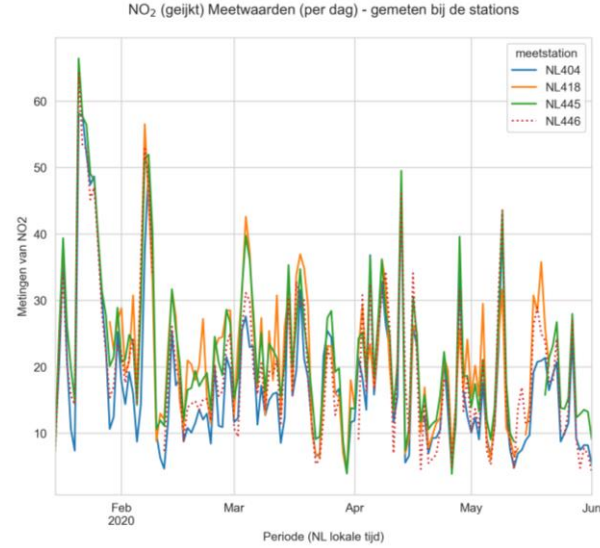
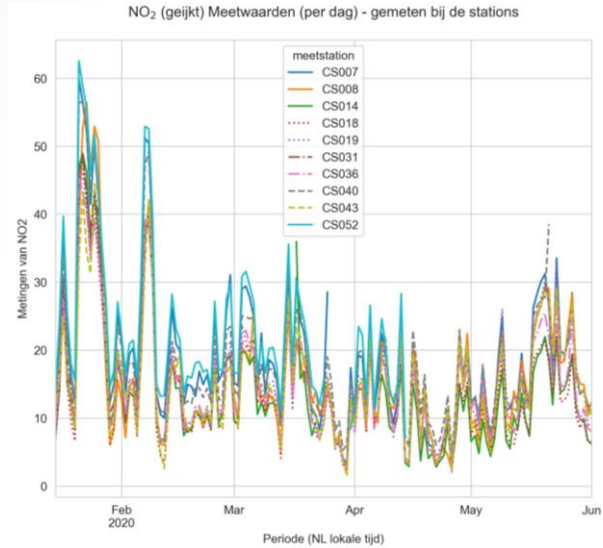$$NO_{2,obs} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Linear relation with Wind + Var 2 ?

3: Invert for Vwind & Rel Humidity

4: Invert for Vwind &Temperature

ALVA 2.2 application 3
Inversion results

Can we observe Corona lockdown effects in our measurements?

Q4:
Results?

$$NO_{2,obs} = NO_{2,emitted} * f(Var\ 1)$$

1: Relation described by three straight lines

2: Relation given by a function

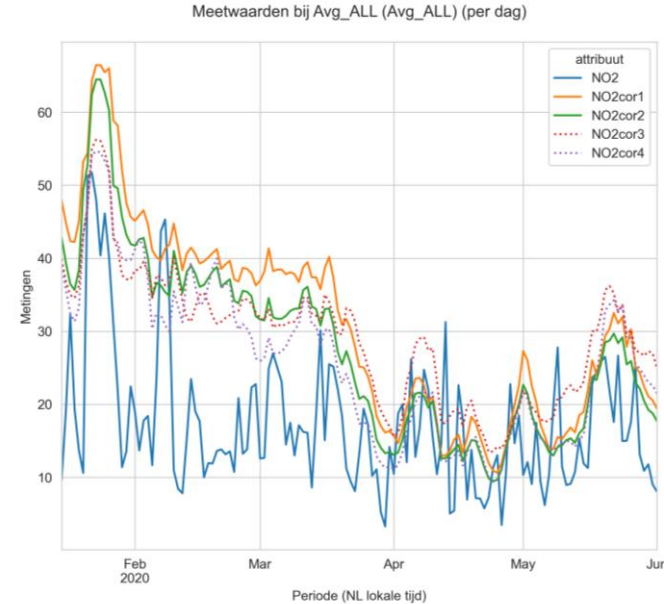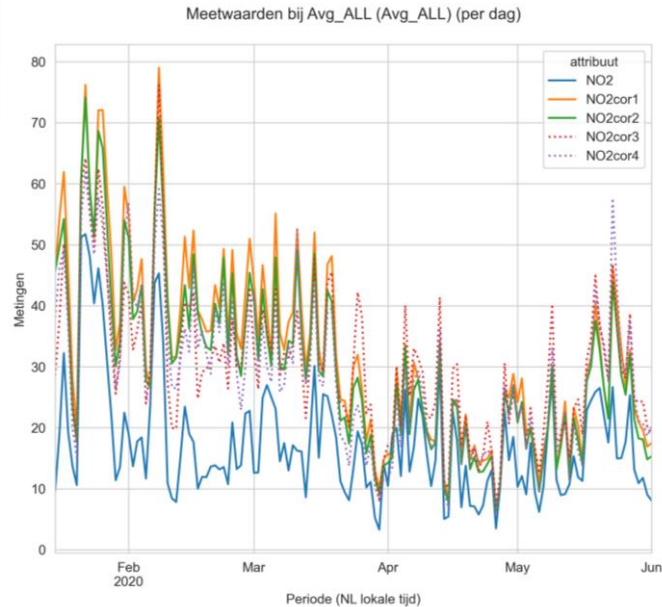$$NO_{2,obs} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Linear relation with Wind + Var 2 ?

3: Invert for Vwind & Rel Humidity

4: Invert for Vwind & Temperature

ALVA 2.2 application 3
Inversion results

Inversion operators are "smoothed" in time

169

Can we observe Corona lockdown effects in our measurements?

Q4:
Results?

$$NO_{2,obs} = NO_{2,emitted} * f(Var\ 1)$$

1: Relation described by three straight lines

2: Relation given by a function

$$NO_{2,obs} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Linear relation with Wind + Var 2 ?

3: Invert for Vwind & Rel Humidity

4: Invert for Vwind &Temperature

Conclusions

- **Relation pollution and meteorology complex**
  - Dispersion effects and pollution build-up are function of time
  - Relations between meteorological quantities

- **Reasonable proxy for NO$_2$ pollution dispersion: windspeed**

- **"a bit of tinkering / smoothing" needed**
  - Software and approach have some limitations

- **More complex statistical approaches:**

  - Willers, S., S. van den Elshout. 'KUNNEN WE MET DE RANDOM FOREST METHODE DOOR DE BOMEN HET BOS WEER ZIEN?' Lucht Juli 2020 (2020): 13–17.

  - Velders, G.J.M., S.M. Willers, J.Wesseling, S. van den Elshout, E. van der Swaluw, D. Mooibroek, S. van Ratingen. 'Improvements in Air Quality in the Netherlands during the Corona Lockdown Based on Observations and Model Simulations'. Atmospheric Environment 247 (februari 2021): 118158.

Future of ALVA

- **Principle:**
  - **Do not try to emulate capabilities of professional software (samenmeten.rivm.nl; samen-analyseren.nl)**
  - **Focus on extensions for own needs, own research questions & own ideas**

- **Easy changes:**
  - Incorporate Palmes chemical tube measurements
  - Incorporate other sources of meteorologic information
  - Better visualisation of 2D function plots (3D displays)

- **Difficult**
  - Graphical frontend & use of Geographic information
  - Currently limited to 1 hour sampling (Luftdaten access gives higher sampling rates)

- **Make system available as open source**

Maintain audit trail
Generate numbers

Conclusions

ALVA

- **ALVA is meant to be complementary to developments such as samenmeten.rivm.nl and samen-analyseren.rivm.nl**

- **Non-standard analysis**

- **Possibility to explore data relationships in detail**

- **Retain capability to see what is happening in your own code**

- **Expand when and where needed; "tinker"**

- **Requires some computer literacy**

- **Objective: open source**
  - Currently distribution to CS groups in Zuid-Holland and on request (LV2@kpnmail.nl)

Supplementary material

LV2 Lucht voor Leidschendam-Voorburg

ALVA 2.2 application 3
Inversion strategies

$$NO_{2,observed} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Gemiddelde Avg_RIVM_NO2 per windrichting/sterkte
NO₂ (geijkt) (NO2)

Gemiddelde NO₂ (geijkt) per Windsnelheid (*10) / Windrichting

Gemiddelde NO₂ (geijkt) per Windsnelheid (*10) / Windrichting

NO₂ (geijkt) (NO2) bij Avg_RIVM

Wind + WindDirection
R^2 = 0.33
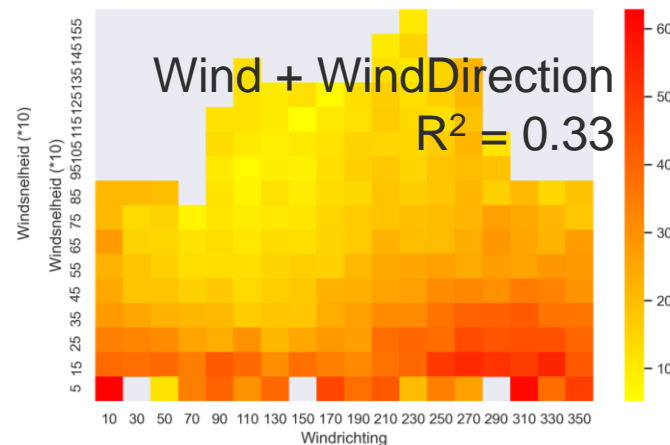
```
Jaaroverzicht 2019
Straal = wind (M/s), kleur = Av(Avg_RIVM_NO2)      Periode:   1/1/2019 - 31/12/2019
File:Avg_RIVM_NO2_Av_wroos_(VIII)                             wk  1 - 1

Station/Meting Gemiddeld  Maximum  Loc Max  Minimum  Loc Min   Mediaan    Aantal
vg_RIVM NO2       26.3     128.6    0121:10    2.3    0930:03     21.8       8732
KNMI215 Wind       4.4      15.0    0307:17    0.00   0120:22      4.0       8760


Gemiddelde(NO₂ (geijkt) (NO2))
Maximum:   68.4 bij Windrichting 310    ZO- Zuidoost          Windsnelheid    0
Zwaartepunt     bij Windrichting 228    ZW- Zuidwest          Windsnelheid  1.1

Conventie windrichting:  ---> O= 0  ---> N= 90  ---> W=180  ---> Z=270  ---> O=360
```
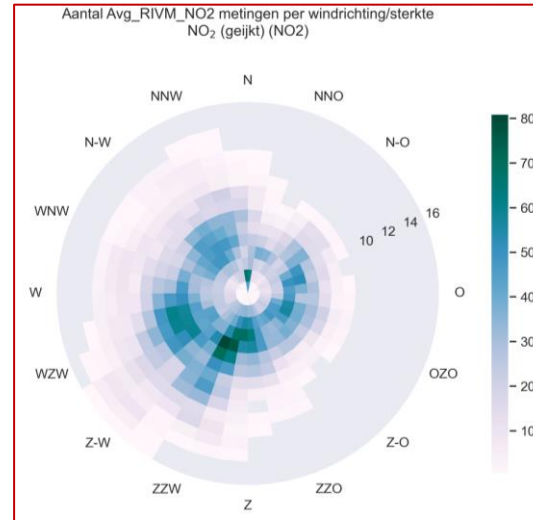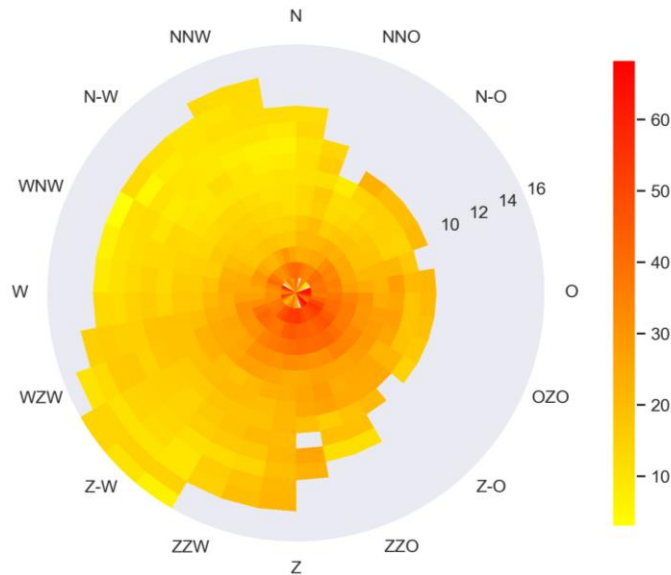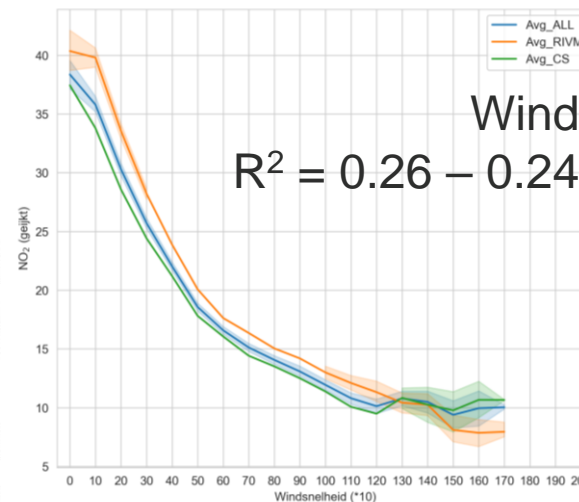
Can we observe Corona lockdown effects in our measurements?

Q5:
What do we miss?

$$NO_{2,obs} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Winddirection

High average NO2 for low-speed eastern winds

➢ Temperature

➢ High pressure regions

ALVA 2.2 application 3
Inversion strategies

$$NO_{2,observed} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Can we observe Corona lockdown effects in our measurements?

Q5:
What do we miss?

$$NO_{2,obs} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Winddirection
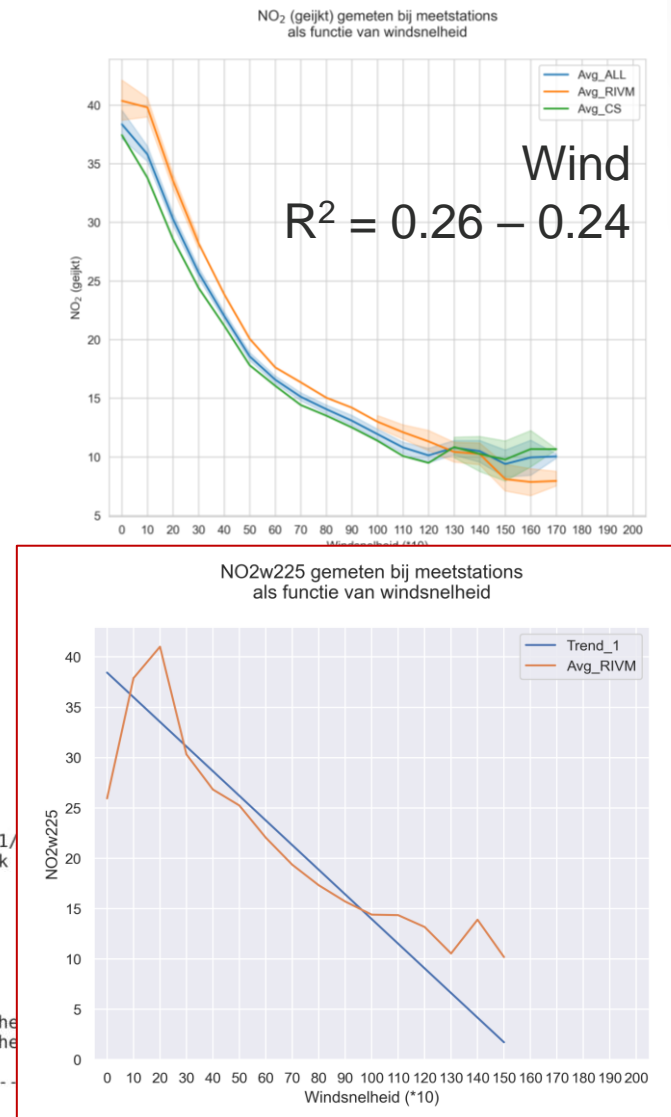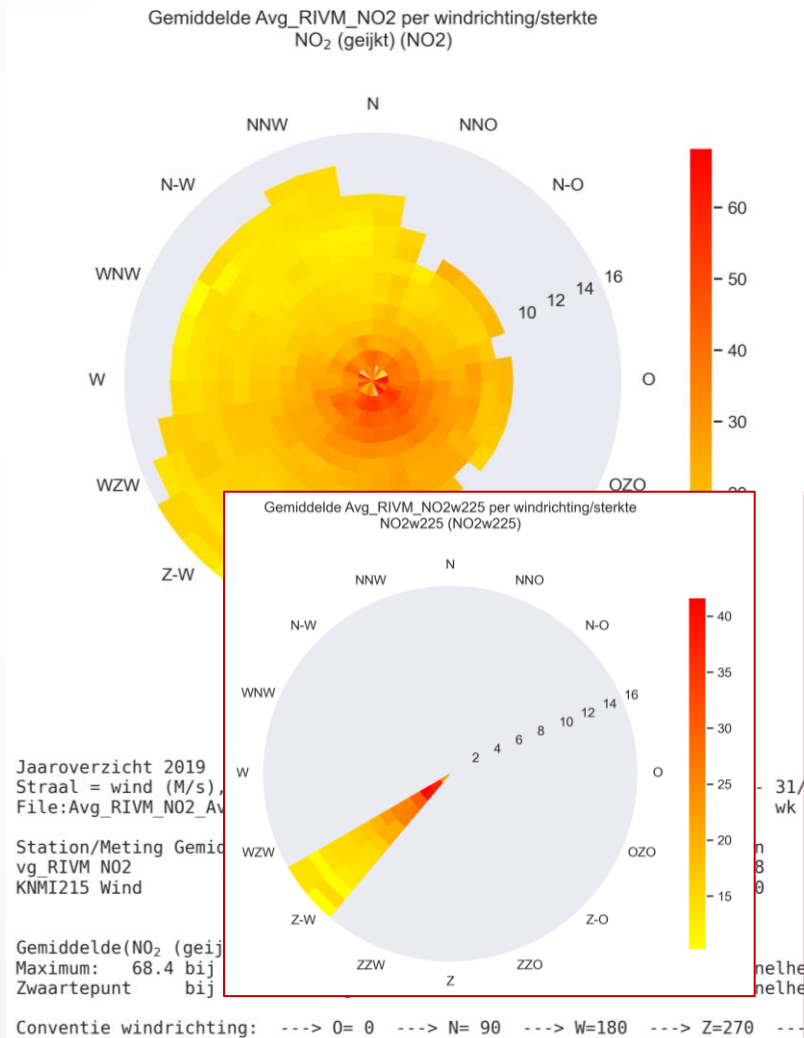
High average NO_2 for low-speed eastern winds

➢ Temperature

➢ High pressure regions

Large windspeeds with South-west winds, missing in Eastern winds

Hence: no uniform sampling for windforce – magnitudes.

ALVA 2.2 application 3
Inversion strategies

$$NO_{2,observed} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Gemiddelde Avg_RIVM_NO2 per windrichting/sterkte
NO₂ (geijkt) (NO2)

NO₂ (geijkt) gemeten bij meetstations als functie van windsnelheid

Wind
$R^2 = 0.26 - 0.24$

Gemiddelde Avg_RIVM_NO2w225 per windrichting/sterkte
NO2w225 (NO2w225)

NO2w225 gemeten bij meetstations als functie van windsnelheid

Can we observe Corona lockdown effects in our measurements?

Q4:
What do we miss?

$$NO_{2,obs} = NO_{2,emitted} * f(Var\ 1, Var\ 2)$$

Winddirection

High average NO2 for low-speed eastern winds

➢ Temperature

➢ High pressure regions

Large windspeeds with South-west winds, missing in Eastern winds

Hence: no uniform sampling for windforce – magnitudes.