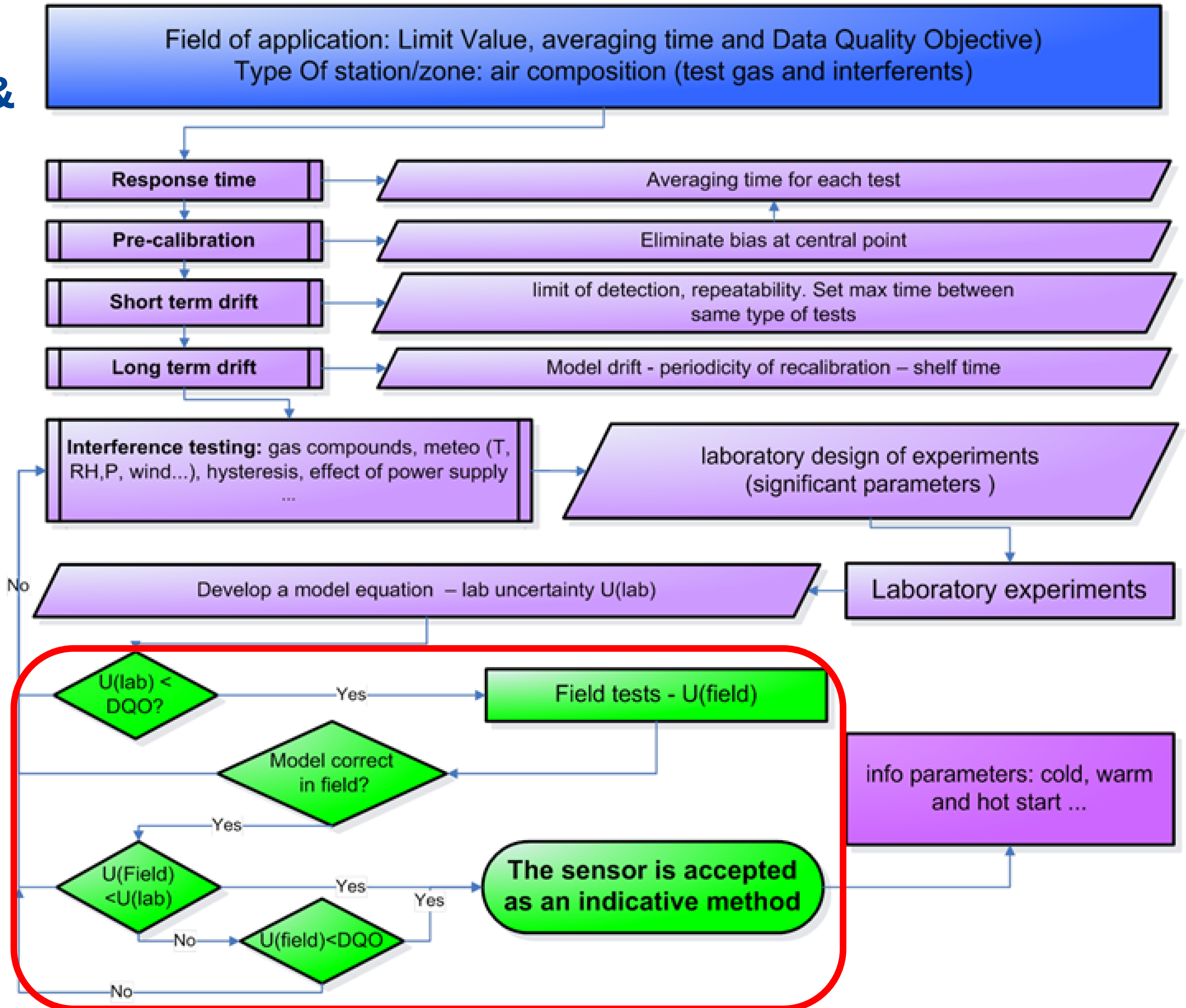
# Calibration of a cluster of low-cost sensors for the measurement of air pollution in ambient air

**L. Spinelle[1], M. Gerboles[1], M.G. Villani[2], Manuel Aleixandre[3] and F. Bonavitacola[4]**
[1]European Commission, JRC, Ispra (VA), Italy
[2]ENEA, Ispra (VA), Italy
[3]Instituto de Física Aplicada, Madrid, Spain
[4]Phoenix Sistemi & Automazione s.a.g.l., Muralto (TI), Switzerland

**IEEE SENSORS 2014 – Valencia, Spain**
**November 2-5, 2014**

European Commission
Joint Research Centre
MACPoll

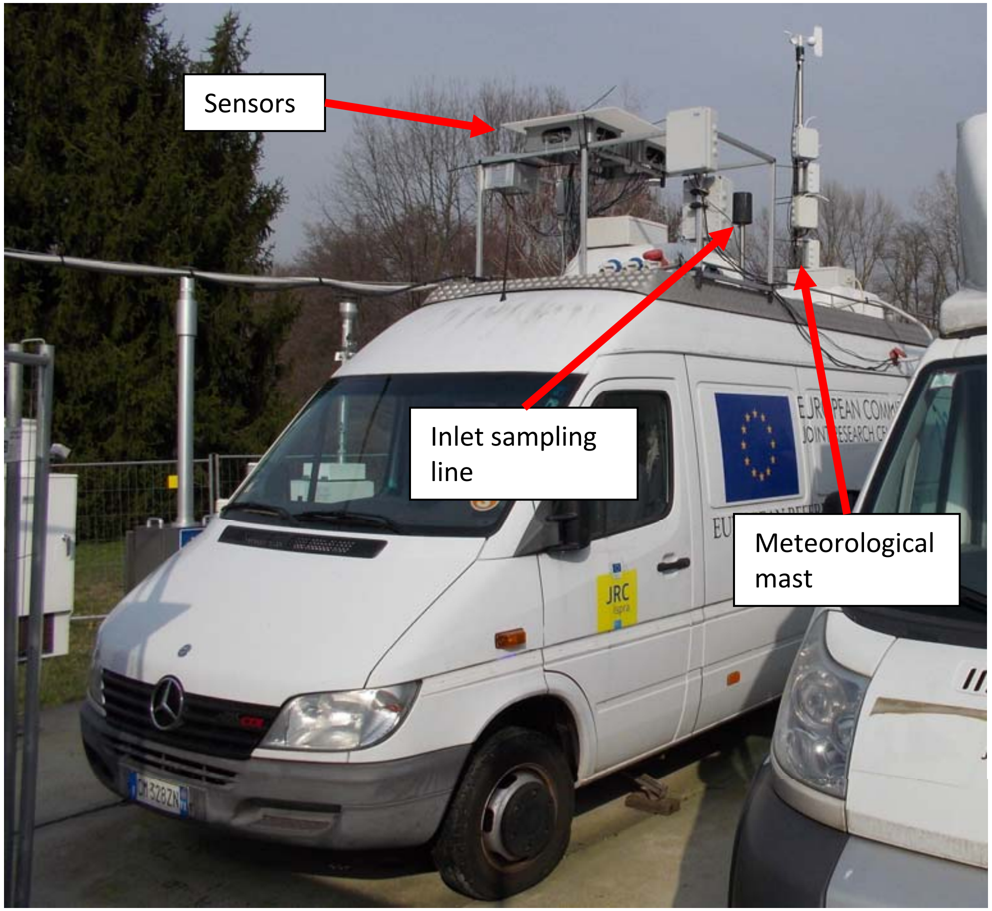
# Evaluation & Validation Protocol

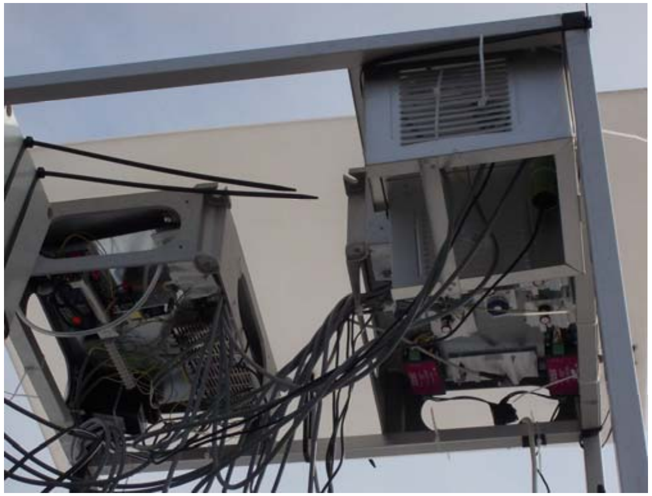



Spinelle L, Aleixandre M, Gerboles M. Protocol of evaluation and calibration of low-cost gas sensors for the monitoring of air pollution. EUR 26112. Luxembourg (Luxembourg): Publications Office of the European Union; 2013. JRC83791

MACPoll

# Field calibration of a cluster of sensors



# Linear regression and multilinear regression

| Manufacturer | Model | $R^2$ of linear regression | Multivariate linear model | $R^2$ |
|---|---|---|---|---|
| αSense | O3 sensors B4 | 0.07 | $O_3 = \frac{Rs - bNO_2 - cNO_2.H_2O - d}{a}$ | 0.49 |
| **Citytech** | **O3_3E1F** | **0.87** | $\boldsymbol{O_3 = \frac{Rs - bNO_2 - c}{a}}$ | **0.91** |
| CairPol | CairclipO3/NO2 | Unknown | $O_3 = \frac{Rs - bNO_2 - c}{a}$ | Unknown |
| αSense | NO2-B4 | 0.06 | $NO_2 = \frac{Rs - bO_3 - cT - dRH - e}{a}$ | 0.56 |
| Citytech | NO2 3E 50 | 0.01 | $NO_2 = \frac{Rs - bO_3 - cT - dRH - e}{a}$ | 0.63 |
| Citytech | NO 3E 100 | 0.05 | Unknown | Unknown |
| **e2V** | **2710 sensor** | **0.31** | $\boldsymbol{NO_2 = \frac{Rs - bO_3 - cT - d}{a}}$ | **0.36** |
| **e2V** | **4514 sensor** | **0.34** | $\boldsymbol{NO_2 = \frac{Rs - bO_3 - cNO - dT - e}{a}}$ | **0.42** |
| CairPol | CairClip NO2 | 0.37 | $NO_2 = \frac{Rs - bO_3 - c}{a}$ | 0.74 |
| Figaro | 5042 sensor | 0.17 | $CO = \frac{Rs - bT - cRH - d}{a}$ | 0.23 |
| e2V | 4514 sensor | 0.56 | | 0.58 |
| Edinburgh Sensors | Gascard NG | 0.14 | $CO_2 = \frac{Rs - bT - cRH - d}{a}$ | 0.47 |
| ELT Sensors | S-100H | 0.58 | | 0.62 |

## Linear regression

- ≠ sensors
- depend on the exposure conditions
- include all interfering effects

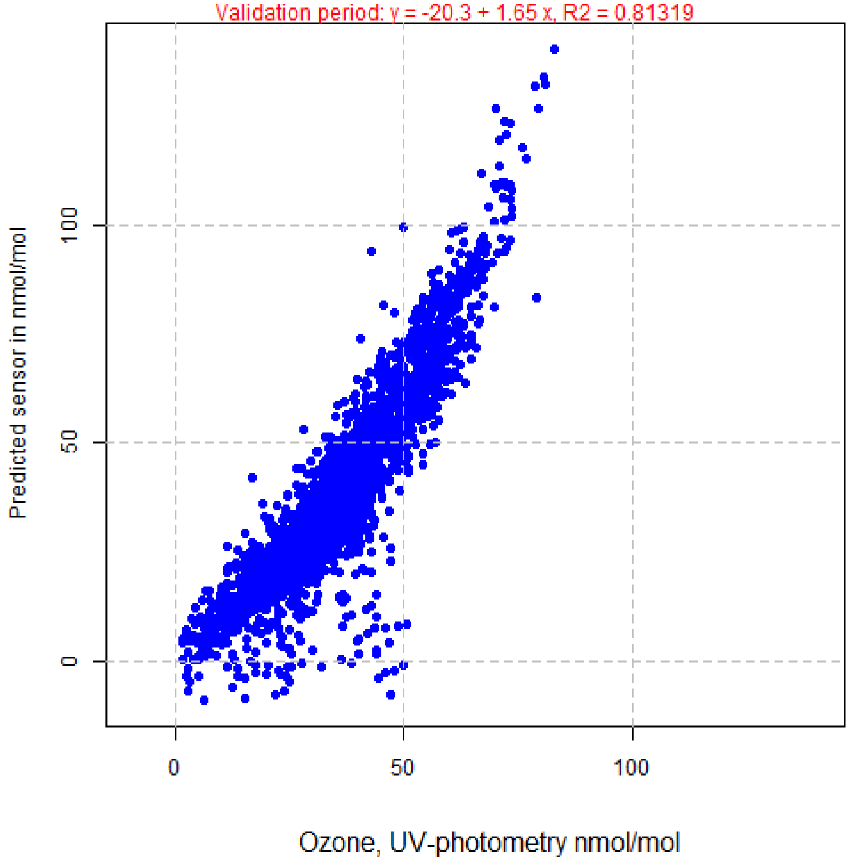
## Multilinear regression

based on laboratory experiments

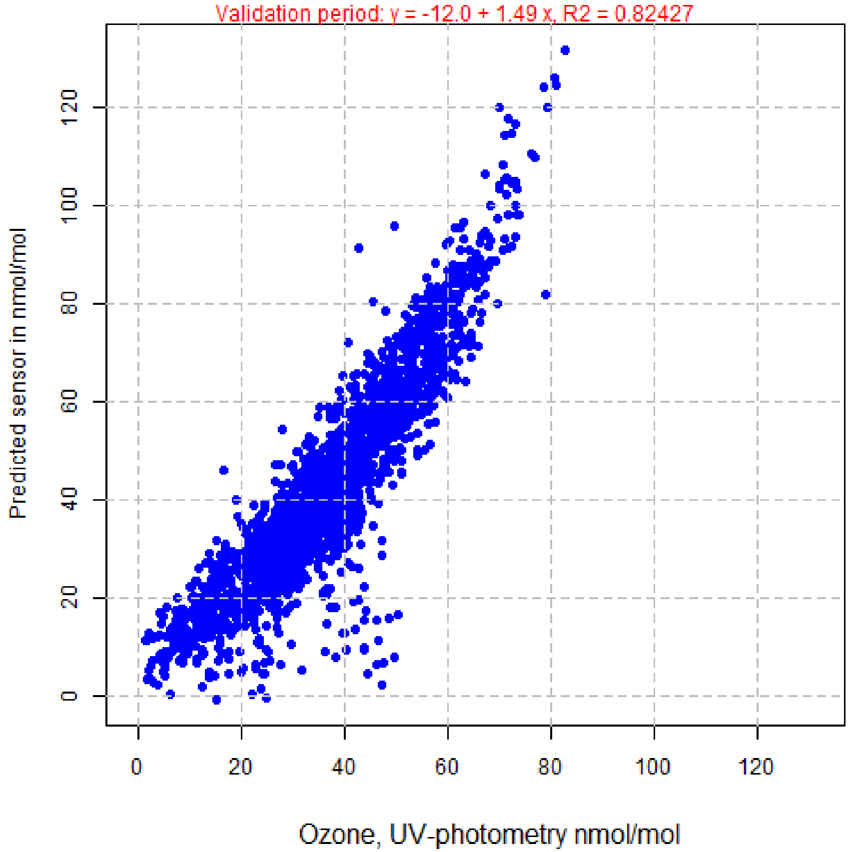
- improve the quality of the data
- needs other variables (gaseous compounds, temperature, humidity…)

# Linear regression and multilinear regression

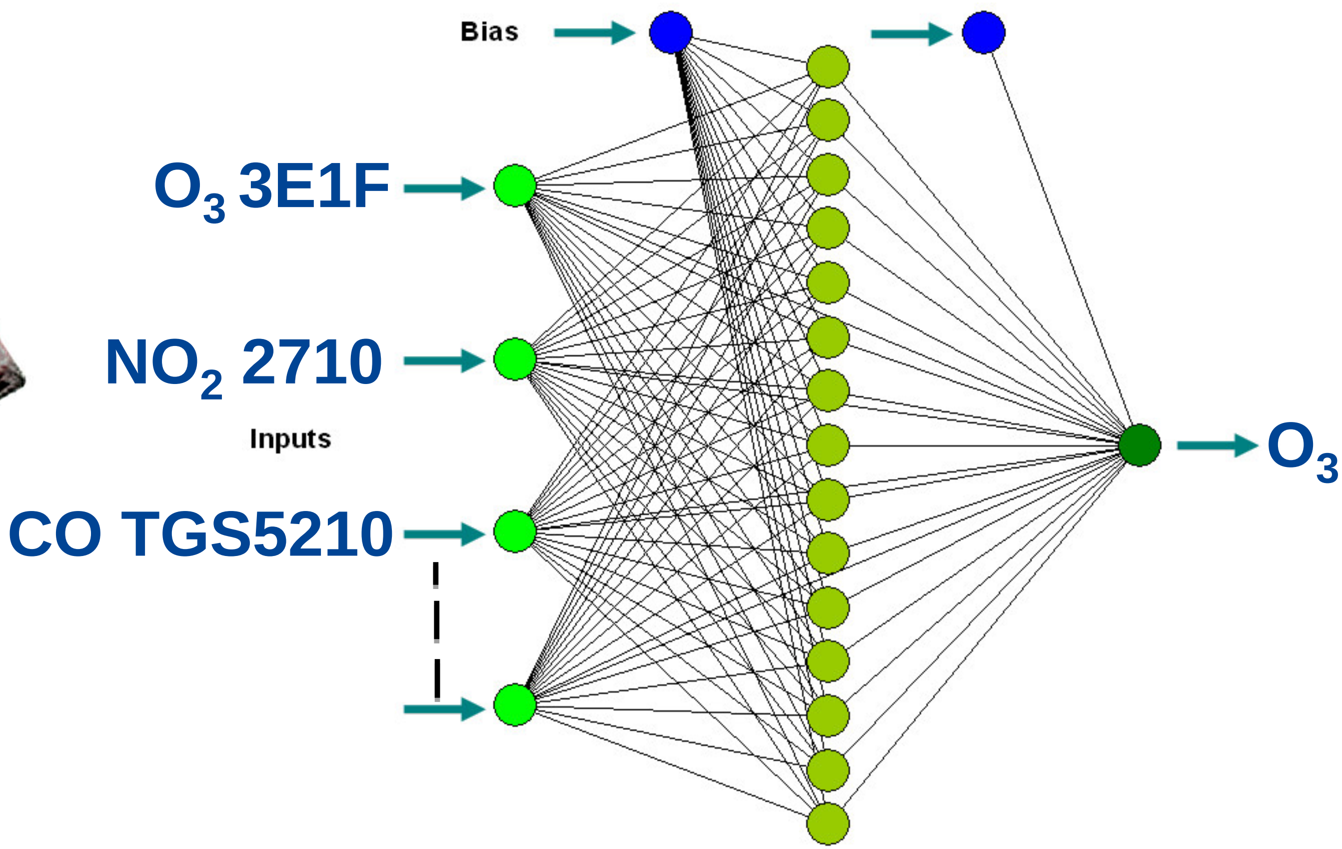
**Linear Regression**
O3_3E1F = f(O3)




**Multi-Linear Regression**
O3_3E1F = f(O3, NO2)



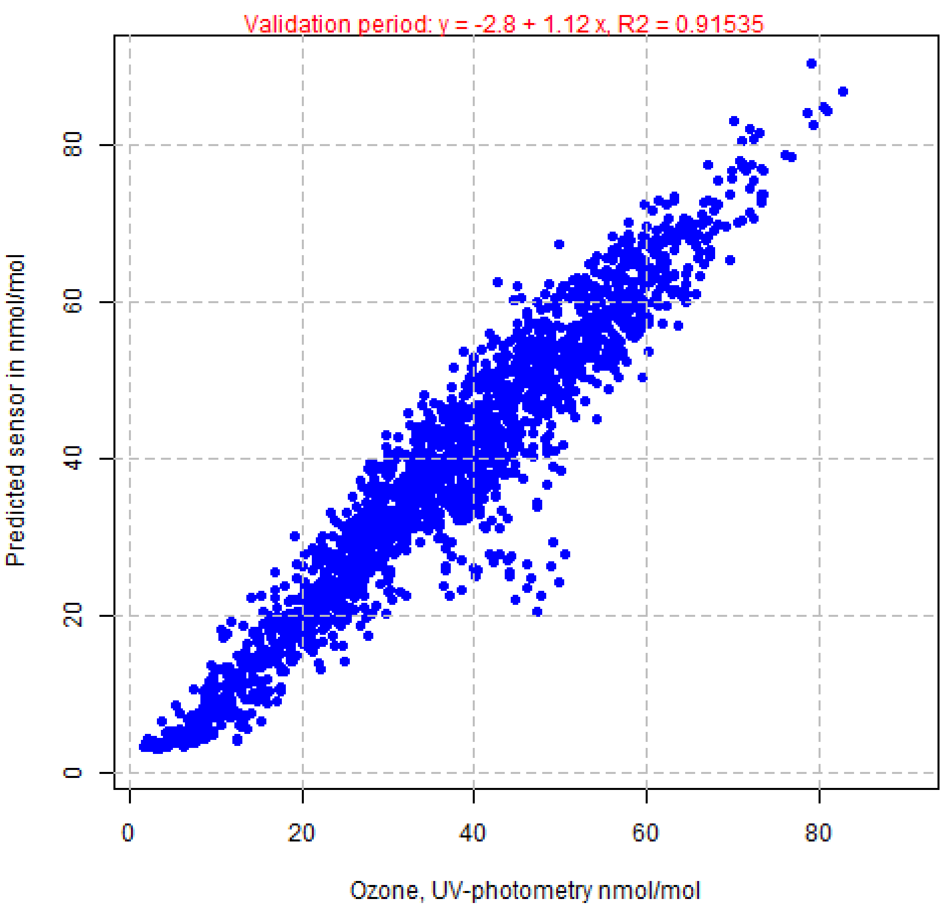
# Artificial Neural Network



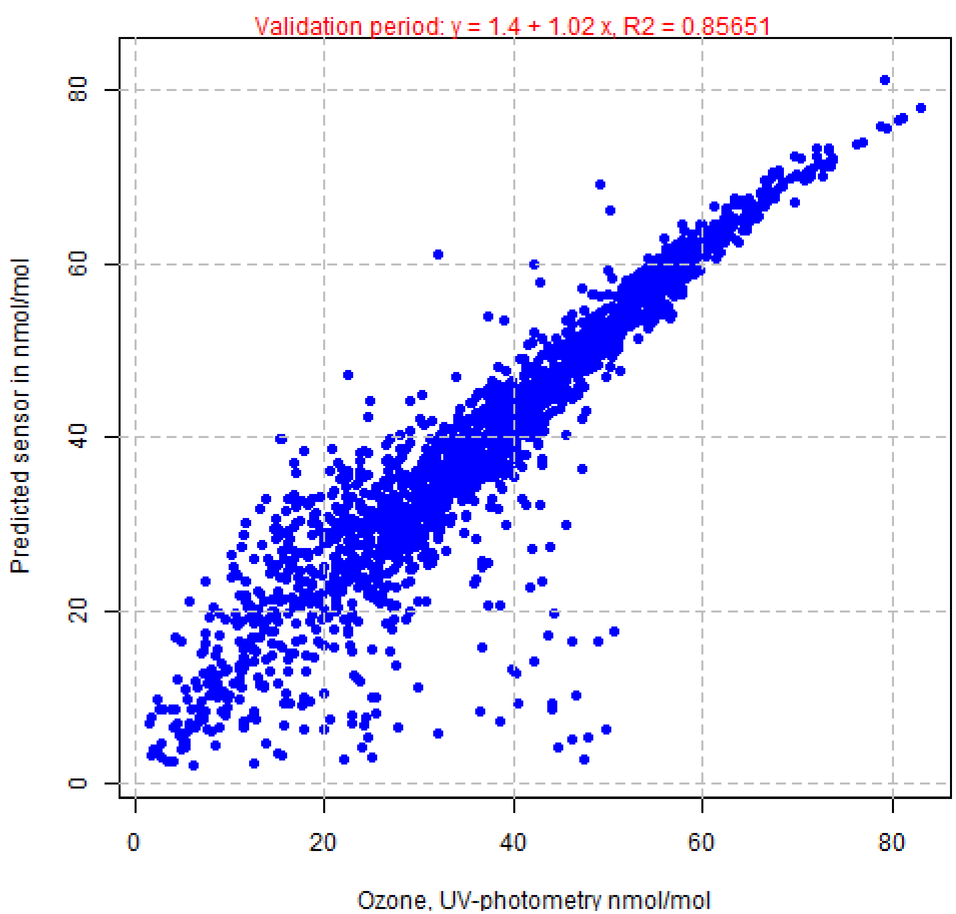

http://midsizeinsider.com/en-us/article/googles-neural-network-makes-advances-i

# Artificial Neural Network

Art. Neural Network, raw sensor values




Art. Neural Net, calibrated sensor values

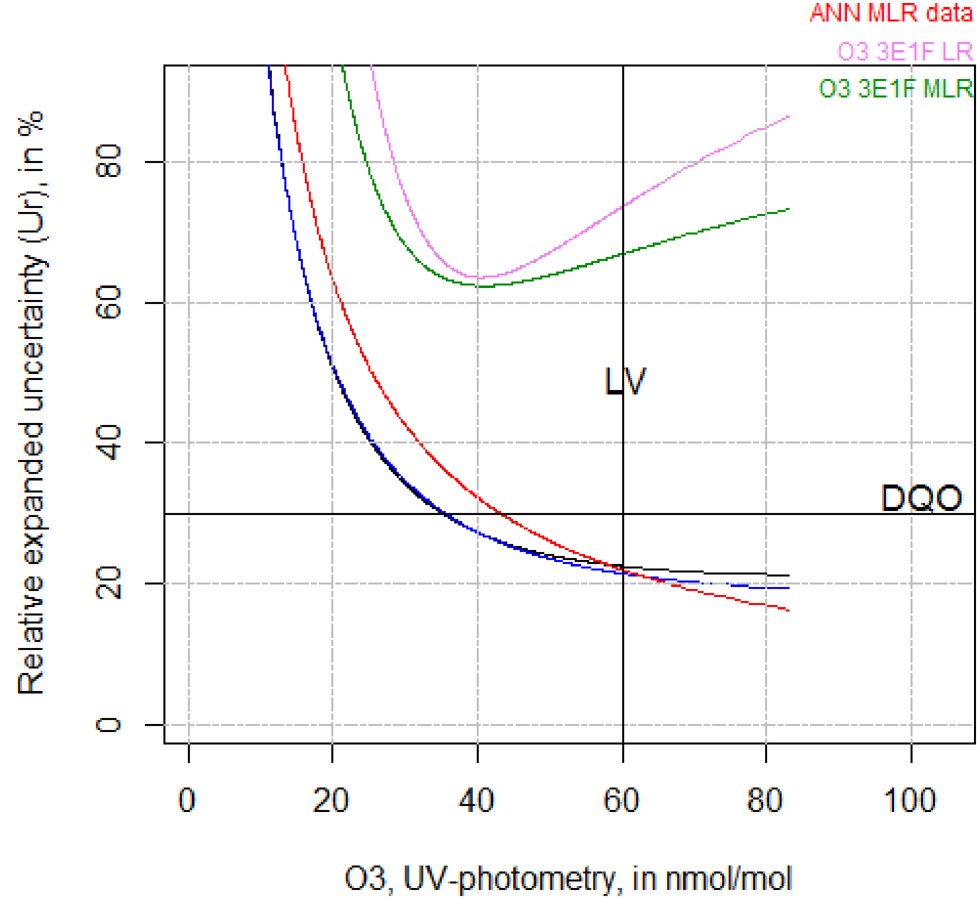


# Model Uncertainty

## $O_3$ $NO_2$ NO CO $CO_2$

| Algorithms | Ambient parameters | Inputs |
|---|---|---|
| LM | No | Sensor |
| MLR | No | Sensor + Reference |
| ANN | No | Sensors |
| ANN+Std | No | Sensors |
| ANN+MLR | No | Sensors + Reference |

$$U_r(y_i) = 2\left(\sqrt{\frac{RSS}{(n-2)} - u^2(x_i) + [a + (b-1)\cdot x_i]^2}\right) / y_i$$

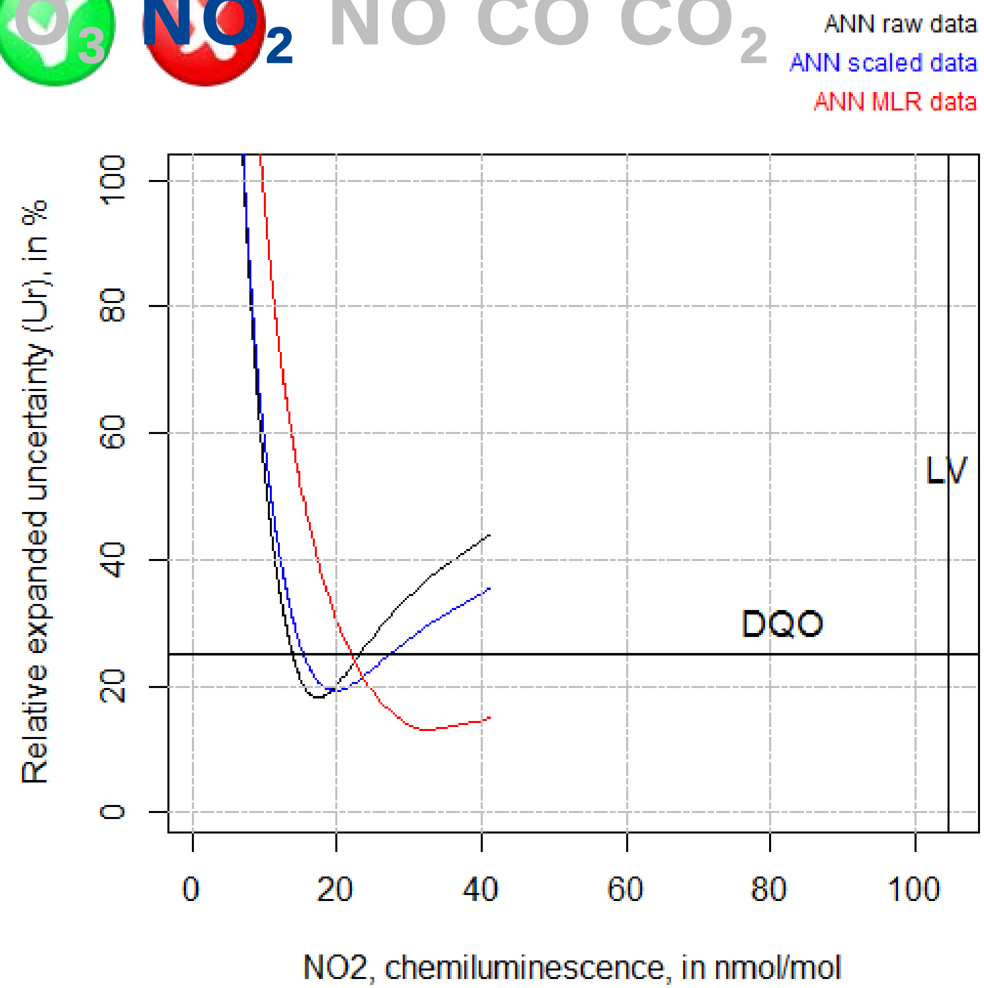


# Model Uncertainty

$O_3$ $NO_2$ NO CO $CO_2$

| Algorithms | Ambient parameters | Inputs |
|---|---|---|
| ~~LM~~ | ~~No~~ | ~~Sensor~~ |
| ~~MLR~~ | ~~No~~ | ~~Sensor + Reference~~ |
| ANN | No | Sensors + Abs. Hum. |
| ANN+Std | No | Sensors + Abs. Hum. |
| ANN+MLR | No | Sensors + Reference |

$$U_r(y_i) = 2\left(\sqrt{\frac{RSS}{(n-2)} - u^2(x_i) + [a + (b-1)\cdot x_i]^2}\right) / y_i$$

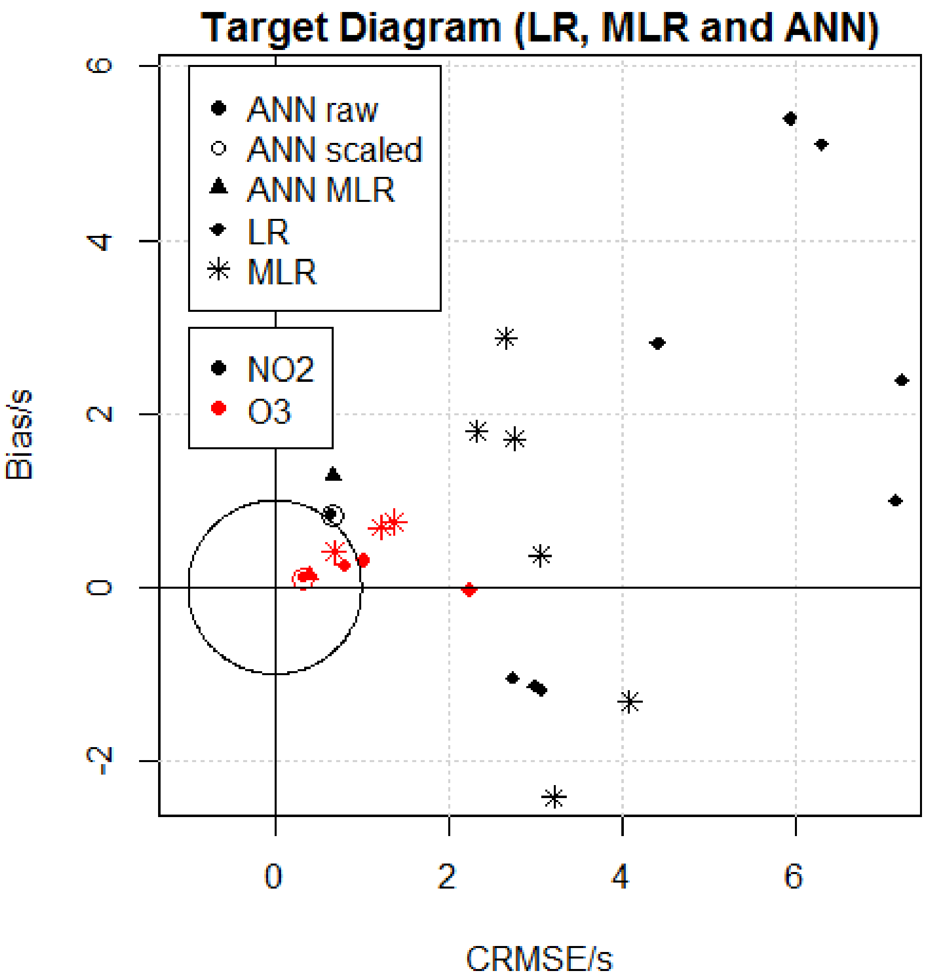


# Model Uncertainty - Target Diagram

- **target cycle** = model results are within the observation uncertainty range

- **symbols out** of the target circle = RMSE > s (standard deviation of reference measurements)

- **ANN** show a lower unbiased RMSE (called centered root-mean-square error, CRMSE) and a lower bias



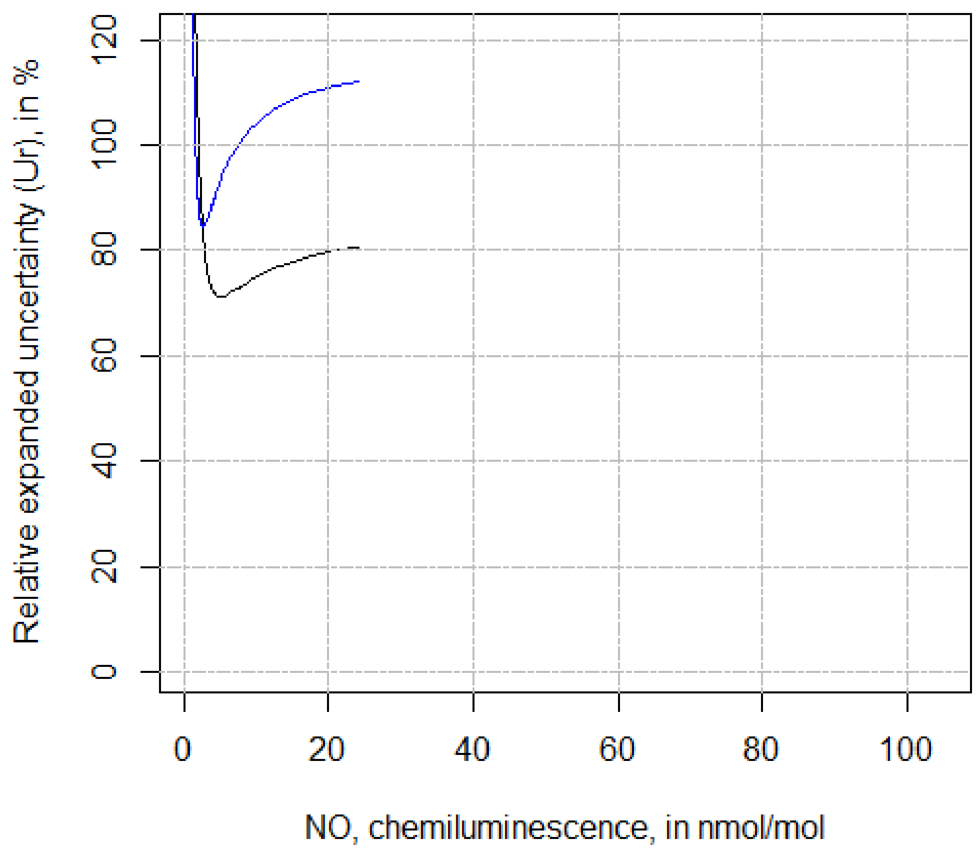
# Model Uncertainty

$O_3$ $NO_2$ NO CO $CO_2$

ANN raw data
ANN scaled data
ANN MLR data

| Algorithms | Ambient parameters | Inputs |
|---|---|---|
| ~~LM~~ | ~~No~~ | ~~Sensor~~ |
| ~~MLR~~ | ~~No~~ | ~~Sensor + Reference~~ |
| ANN | No | Sensors + T. + Hum. |
| ANN+Std | No | Sensors + T. + Hum. |
| ~~ANN+MLR~~ | ~~No~~ | ~~Sensors + Reference~~ |

$$U_r(y_i) = 2\left( \sqrt{\frac{RSS}{(n-2)} - u^2(x_i) + [a + (b-1)\cdot x_i]^2} \right) / y_i$$

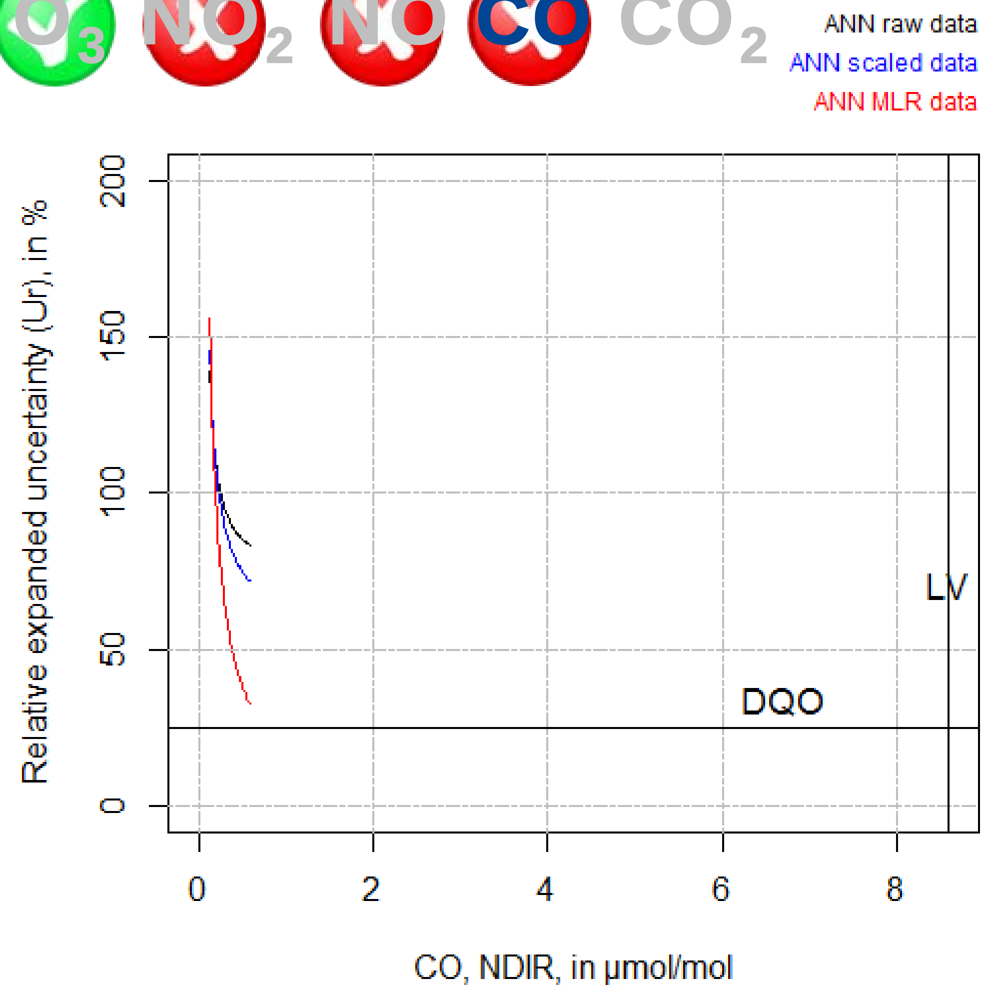


# Model Uncertainty

$O_3$ $NO_2$ NO CO $CO_2$

| Algorithms | Ambient parameters | Inputs |
|---|---|---|
| ~~LM~~ | ~~No~~ | ~~Sensor~~ |
| ~~MLR~~ | ~~No~~ | ~~Sensor + Reference~~ |
| ANN | No | Sensors + T. + Hum. |
| ANN+Std | No | Sensors + T. + Hum. |
| ANN+MLR | No | Sensors + Reference |

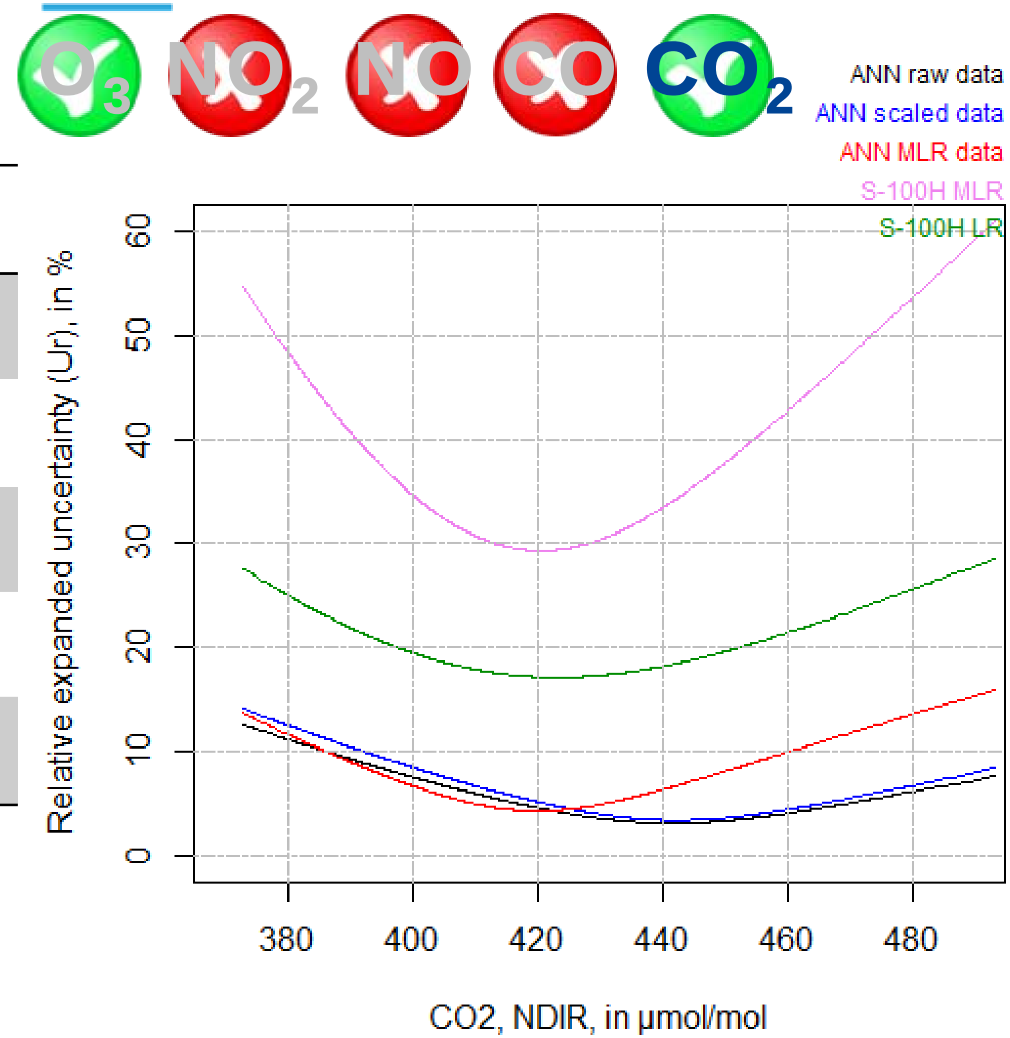
$$U_r(y_i) = 2\left(\sqrt{\frac{RSS}{(n-2)} - u^2(x_i) + [a + (b-1)\cdot x_i]^2}\right) / y_i$$



# Model Uncertainty

| Algorithms | Ambient parameters | Inputs |
|---|---|---|
| LM | No | Sensor |
| MLR | No | Sensor + Reference |
| ANN | No | Sensors |
| ANN+Std | No | Sensors |
| ANN+MLR | No | Sensors + Reference |

$$U_r(y_i) = 2\left(\sqrt{\frac{RSS}{(n-2)} - u^2(x_i) + [a+(b-1)\cdot x_i]^2}\right)/y_i$$



# Conclusion calibration methods for the whole cluster of sensors

- The DQO for indicative methods can be met for $O_3$, likely for $NO_2$. High uncertainty for NO and CO (>75%). For $CO_2$, low uncertainty down to about 5%.
- Linear and Multilinear regression gives the highest U.
- ANN methods: higher $R^2$ and lower CRMSE -> lower U; lower bias to reference data (slopes and intercept nearer to 1 and 0, respectively).
- Reference data (meteo / gas) does decrease measurement uncertainty for the ANN methods.
- ANN can solve cross sensitivity issues from which suffers the major part of sensors (gaseous interference, temperature/humidity dependence).

# Thank You…

M. Gerboles, M. Aleixandre, F. Lagler, M. G. Villani,
F. Bonavitacola, M. Penza

N. R. Jensen, A. Dell'Acqua, C. Gruening, G. Manca,
S. Martins Dos Santos

Reports at:
ftp://ftp_erlap_ro:3rlapsyst3m@s-jrciprvm-ftp-ext.jrc.it/ERLAPDownload.htm

Or send a mail at

laurent.spinelle@jrc.ec.europa.eu
michel.gerboles@jrc.ec.europa.eu