# Myair Toolkit for Model Evaluation

**Amy Stidworthy[1], David Carruthers[1], Jenny Stocker[1], Sam Royston[1], Dimitris Balis[2], Eleni Katragkou[2], Jaakko Kukkonen[3]**

[1]Cambridge Environmental Research Consultants (CERC)
[2]Aristotle University of Thessaloniki (AUTH)
[3]Finnish Meteorological Institute (FMI)

**15th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes**

Madrid, Spain 6 – 9 May 2013

## Summary

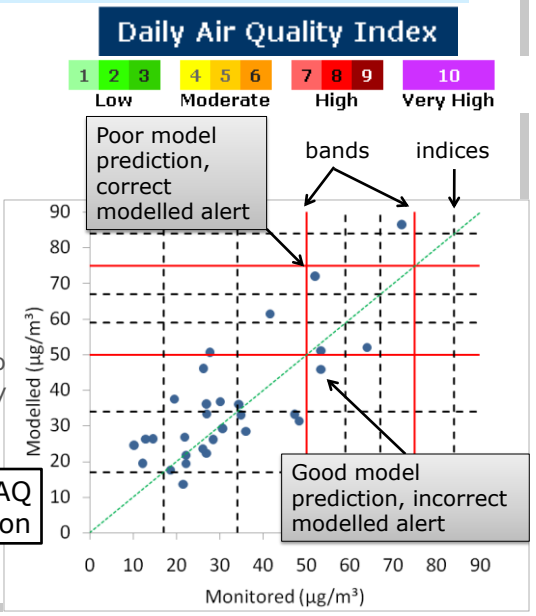
- Background
- Existing tools and methodologies
- Myair Toolkit capabilities
- Practicalities
- Summary

## Background
## PASODOBLE project

- PASODOBLE is the Copernicus (GMES) downstream service project, producing **local-scale air quality services for Europe** under the nam
- Local for
  package
  evaluate
  forecasts
  **Evaluat**


## Background
## Why AQ forecast models need special tools

- Air quality (AQ) forecasting systems predict air quality in terms of bandings.
- Forecasts aim to get the band correct (low, moderate etc).
- An alert is issued by the forecasting system if a moderate, high or very high band is forecast
- Therefore, validating a forecasting system is different to validating concentrations directly output from an AQ model.




Scatter plot for AQ forecast system validation

## Background
## Development procedure

- The Myair Toolkit validates air quality model output, focussing on requirements for **standardised evaluation of local air quality forecast** models.
- The Toolkit was designed following an extensive review (2010-11) of the state-of-the-art in air quality model evaluation.
- Toolkit builds on existing tools and initiatives.

## Existing tools and methodologies

- AQ model validation tools and methodologies:
  - Model Validation Kit
  - ASTM model evaluation methodology
  - FAIRMODE Delta tool
- Meteorological forecasting models:
  - Event-based statistics used at ECMWF for validating the accuracy of Numerical Weather Predictions can be used to validate AQ forecasting systems.
- openair data analysis tools:
  - UK project for the air pollution community
  - Free, open-source, innovative data analysis tools.

## Myair Toolkit capabilities
## What can Myair Toolkit do?

- Assess your model's forecast skill
- Assess your model's concentration predictions
- Help you investigate model performance at individual stations using openair graphs
- Easily import a wide range of gridded and point modelled data formats
- Download and import in situ monitoring data for the UK (also CSV files)
- Save graphical and statistical output to your computer
- Run in batch mode, for easy automation

## Myair Toolkit capabilities
## What can Myair Toolkit do?

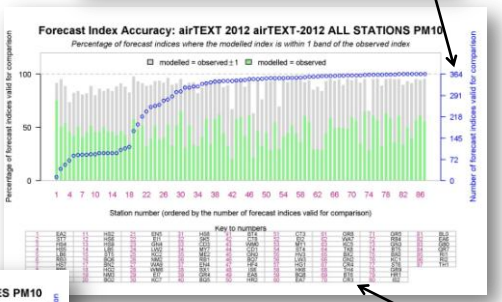
- **Assess your model's forecast skill**
- Assess your model's concentration predictions
- Help you investigate model performance at individual stations using openair graphs
- Easily import a wide range of gridded and point modelled data formats
- Download and import in situ monitoring data for the UK (also CSV files)
- Save graphical and statistical output to your computer
- Run in batch mode, for easy automation

## Myair Toolkit capabilities
## Assess your model's forecast skill

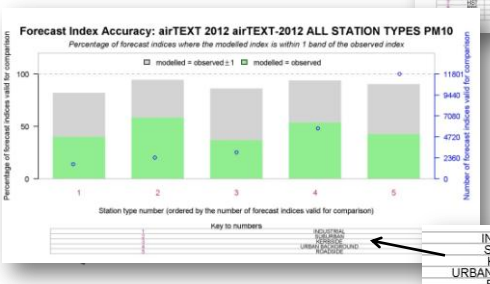
**Look at the percentage of forecast indices within one of observed (should be close to 100%) for each pollutant, grouped by station…**

- modelled = observed ± 1
- modelled = observed

Number of forecast indices valid for comparison




Station number




INDUSTRIAL
SUBURBAN
KERBSIDE
URBAN BACKGROUND
ROADSIDE

**… or grouped by station type (e.g. roadside, urban background, rural etc).**

## Myair Toolkit capabilities
## Assess your model's forecast skill

**Look at model's skill at predicting alert threshold exceedences (i.e. pollution episodes) in different ways:**

| | | Alert modelled? | |
|---|---|---|---|
| | | **Yes** | **No** |
| **Alert observed?** | **Yes** | a | b |
| | **No** | c | d |

a, b, c and d are counts of the number of days where alerts were or were not modelled and were or were not observed

$$\text{Odds Ratio Skill Score (ORSS)} = \frac{ad - bc}{ad + bc}$$

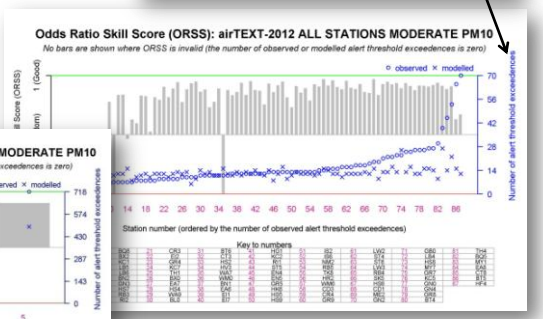
ORSS gives equal weighting to correct non-prediction and to correct prediction

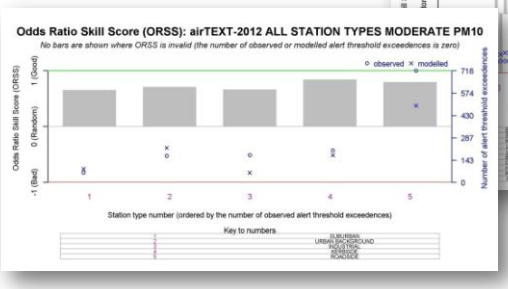
| | | |
|---|---|---|
| Perfect score: | b=c=0 | ORSS=1 |
| Good score: | ad>bc | ORSS>0 |
| Bad score: | bc>ad | ORSS<0 |
| Fail score: | a=d=0 | ORSS=-1 |

## Myair Toolkit capabilities
## Assess your model's forecast skill

ORSS grouped by station...

Number of forecast indices valid for comparison




... or grouped by station type

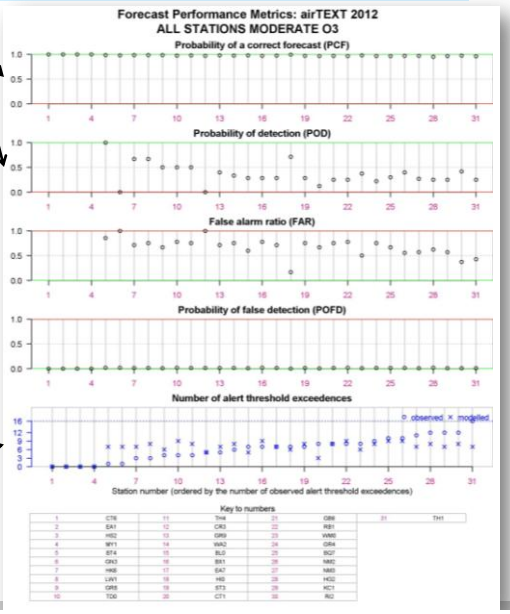
*ORSS is a good measure if a lot of episodes are measured, but note that it's easy to get a good score if there are few episodes compared to the number of forecasts because d will be high*

## Myair Toolkit capabilities
## Assess your model's forecast skill

Probability

Using the Toolkit you can also look at other measures of model skill, for example the 'probability of detection' and the 'false alarm ratio' for different alert thresholds...

Number of alerts



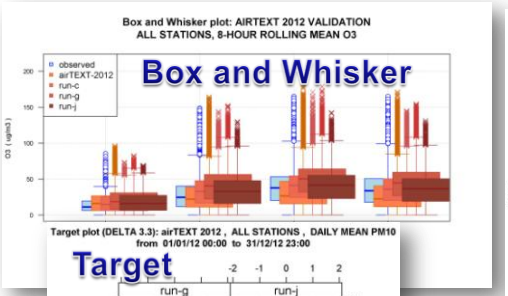
## Myair Toolkit capabilities
# What can Myair Toolkit do?

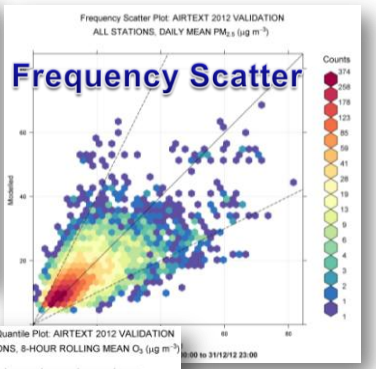
- Assess your model's forecast skill
- **Assess your model's concentration predictions**
- Help you investigate model performance at individual stations using openair graphs
- Easily import a wide range of gridded and point modelled data formats
- Download and import in situ monitoring data for the UK (also CSV files)
- Save graphical and statistical output to your computer
- Run in batch mode, for easy automation

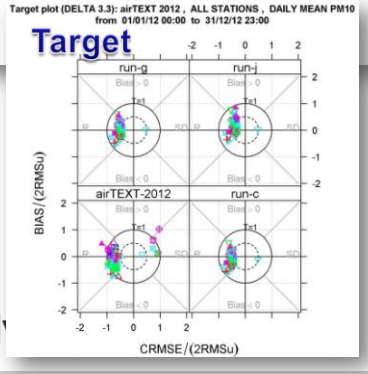
Myair Pasodoble

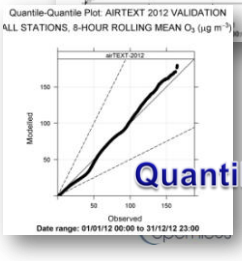
Copernicus

## Myair Toolkit capabilities
# Assess your model's concentration predictions



## Myair Toolkit capabilities
# Assess your model's concentration predictions

Statistical output includes standard results such as mean, bias, standard deviation and more (this table is an extract from the User Guide)

| Name | Description | Equation |
|---|---|---|
| Num.valid.values | Number of values | |
| obs.mean<br>mod.mean | Mean | $1/n \sum C$ |
| SDO<br>SDM | Standard Deviation | $\sqrt{1/n \sum (C - \overline{C})^2}$ |
| MB | Mean Bias | $\overline{(C_p - C_o)}$ |
| NMSE | Normalised Mean-Square-Error | $\overline{(C_p - C_o)^2} / \overline{C_o C_p}$ |
| R | Pearson's Correlation Coefficient | $\mathrm{cov}(C_p, C_o) / \sigma_{C_p} \sigma_{C_o}$ |
| Fac2 | Factor of 2 | Fraction of data where $0.5 \leq C_p / C_o \leq 2$ (when $C_o = 0$, $C_p / C_o \rightarrow \infty$ and the data pair is not counted) |
| Fb | Fractional Bias | $(\overline{C_p} - \overline{C_o}) / 0.5(\overline{C_o} + \overline{C_p})$ |
| Fs | Fractional Standard Deviation | $(\sigma_{C_p} - \sigma_{C_o}) / 0.5(\sigma_{C_o} + \sigma_{C_p})$ |
| obs.max<br>mod.max | Maximum<br>Robust highest concentration | $\max C$<br>$\chi(n) + (\chi - \chi(n)) \ln\left(\frac{3n-1}{2}\right)$, where $n$ is the number of values used to characterise the upper end of the concentration distribution, $\chi$ is the average of the $n-1$ largest values, and $\chi(n)$ is the $n^{th}$ largest value; $n$ is taken to be 26. |
| obs.RHC<br>mod.RHC | | |

Myair Pasodoble

Copernicus

EUROPEAN COMMISSION

## Myair Toolkit capabilities
# What can the Myair Toolkit do?

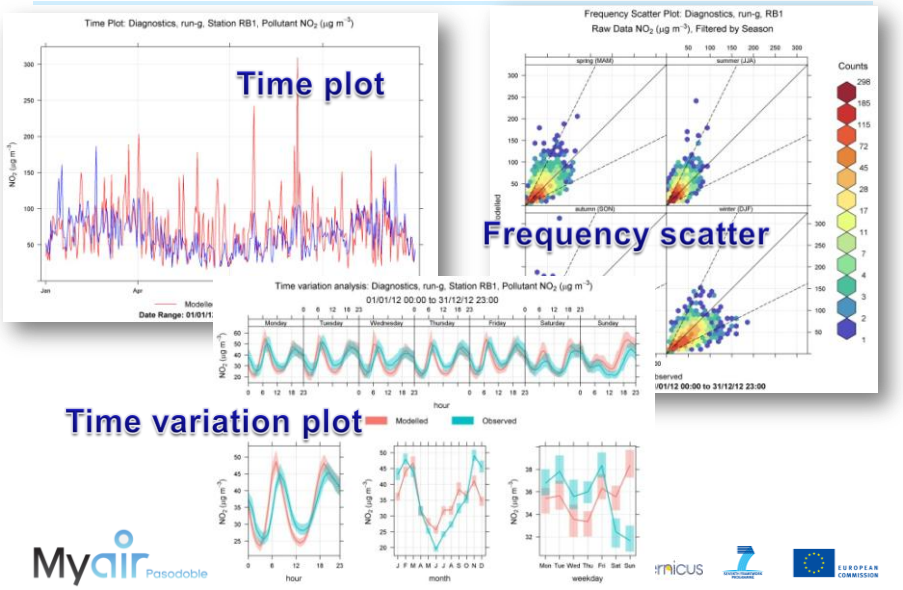
- Assess your model's forecast skill
- Assess your model's concentration predictions
- **Help you investigate model performance at individual stations using openair graphs**
- Easily import a wide range of gridded and point modelled data formats
- Download and import in situ monitoring data for the UK (also CSV files)
- Save graphical and statistical output to your computer
- Run in batch mode, for easy automation

Myair Pasodoble

Copernicus

EUROPEAN COMMISSION

## Myair Toolkit capabilities
## Investigate model performance at individual stations

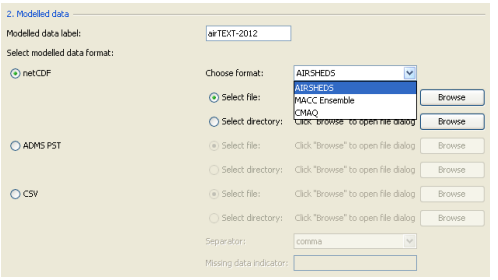



## Myair Toolkit capabilities
## What can the Myair Toolkit do?

- Assess your model's forecast skill
- Assess your model's concentration predictions
- Help you investigate model performance at individual stations using openair graphs
- **Easily import a wide range of gridded and point modelled data formats**
- Download and import in situ monitoring data for the UK (also CSV files)
- Save graphical and statistical output to your computer
- Run in batch mode, for easy automation

## Myair Toolkit capabilities
# Import gridded and point modelled data formats

- Supported modelled data formats:
  - Gridded netCDF
    - AIRSHEDS
    - MACC Ensemble
    - CMAQ
  - Point data
    - ADMS PST
    - Generic CSV



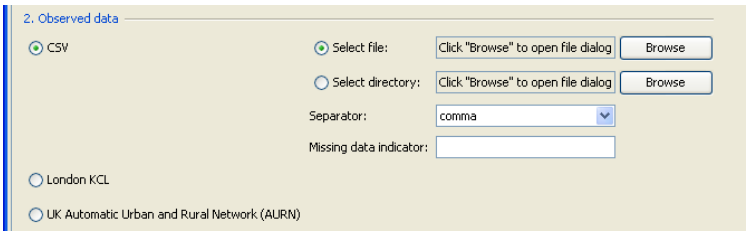

- The Toolkit interpolates gridded data to station locations
- You can import a single file or a whole directory of files

## Myair Toolkit capabilities
# What can the Myair Toolkit do?

- Assess your model's forecast skill
- Assess your model's concentration predictions
- Help you investigate model performance at individual stations using openair graphs
- Easily import a wide range of gridded and point modelled data formats
- **Download and import in situ monitoring data for the UK (also CSV files)**
- Save graphical and statistical output to your computer
- Run in batch mode, for easy automation

## Myair Toolkit capabilities
# Download and import monitoring data

- In situ observed data for 2 UK networks can be downloaded and imported automatically (London KCL and UK AURN)
- Observed data in a generic CSV format can be imported from a single file or directory of files




## Myair Toolkit capabilities
# What can the Myair Toolkit do?

- Assess your model's forecast skill
- Assess your model's concentration predictions
- Help you investigate model performance at individual stations using openair graphs
- Easily import a wide range of gridded and point modelled data formats
- Download and import in situ monitoring data for the UK (also CSV files)
- **Save graphical and statistical output to your computer**
- Run in batch mode, for easy automation

## Myair Toolkit capabilities
# Save graphical and statistical output

- Saves graphs as image files (JPG, PNG) or PDFs for importing into documents
- Saves data (raw, processed and statistics) in CSV files, to provide an audit trail and for further analysis

*One of the CSV files output by the Toolkit*

airTEXT 2012_Validation_forecast_alert_stats.csv

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | pollutant | type | model | alert.name | num.obs.alerts | a | b | c | d |
| 2 | pm10 | suburban | airTEXT-2012 | moderate | 64 | 9 | 77 | 55 | 2266 |
| 3 | pm10 | suburban | airTEXT-2012 | high | 5 | 5 | 30 | 0 | 2372 |
| 4 | pm10 | urban background | airTEXT-2012 | moderate | 171 | 30 | 192 | 141 | 5328 |
| 5 | pm10 | urban background | airTEXT-2012 | high | 12 | 11 | 80 | 1 | 5599 |
| 6 | pm10 | kerbside | airTEXT-2012 | moderate | 205 | 66 | 109 | 139 | 2681 |
| 7 | pm10 | kerbside | airTEXT-2012 | high | 38 | 22 | 36 | 16 | 2921 |
| 8 | pm10 | industrial | airTEXT-2012 | moderate | 175 | 22 | 41 | 153 | 1425 |
| 9 | pm10 | industrial | airTEXT-2012 | high | 32 | 10 | 15 | 22 | 1594 |
| 10 | pm10 | roadside | airTEXT-2012 | moderate | 718 | 157 | 337 | 561 | 10746 |
| 11 | pm10 | roadside | airTEXT-2012 | high | 88 | 53 | 143 | 35 | 11570 |

airTEXT 2012_Validation_forecas

Myair Pasodoble

Copernicus

## Myair Toolkit capabilities
# What can the Myair Toolkit do?

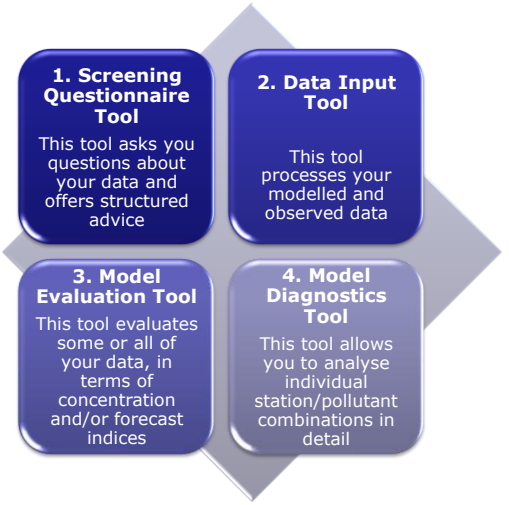
- Assess your model's forecast skill
- Assess your model's concentration predictions
- Easily import a wide range of gridded and point modelled data formats
- Download and import in situ monitoring data for the UK (also CSV files)
- Save graphical and statistical output to your computer
- Help you investigate model performance at individual stations using openair graphs
- **Run in batch mode, for easy automation**

Batch mode allows easy integration of model evaluation into automatic processes, and also easy re-generation of results with new data

Myair Pasodoble

## Practicalities
## What do you get?

- 4 tools
- Runs on most commonly-used platforms, including Windows, Linux, Mac
- Requires you to download and install some free software, which only takes a few minutes
- Comprehensive User Guide included




Myair Pasodoble

Copernicus

## Summary

- The Myair Toolkit for Model Evaluation is a powerful new tool for the evaluation of air quality forecasting models
- The Toolkit was developed building on existing tools and methodologies
- You can download the **free** toolkit from http://www.cerc.co.uk/environmental-software/myair-toolkit.html

Myair Pasodoble

Copernicus

## Acknowledgements

The authors wish to thank:

- PASODOBLE project partners, in particular the users in the 'Local forecast model evaluation support' work package and those involved in evaluating the AIRSHEDS products, for their valuable feedback during development and demonstration of the Myair Toolkit.
- EU FP7 programme for funding the 3-year PASODOBLE project (completed April 2013)