Toolkit for Evaluating Regional and Local Air Quality Models with Observations

Amy Stidworthy, CERC
M. Oades, D. Connolly, F. Otu-Larbi, O. Wild, S. Jain, R. Doherty, J. Zhong, W.J. Bloss, D.J. Carruthers

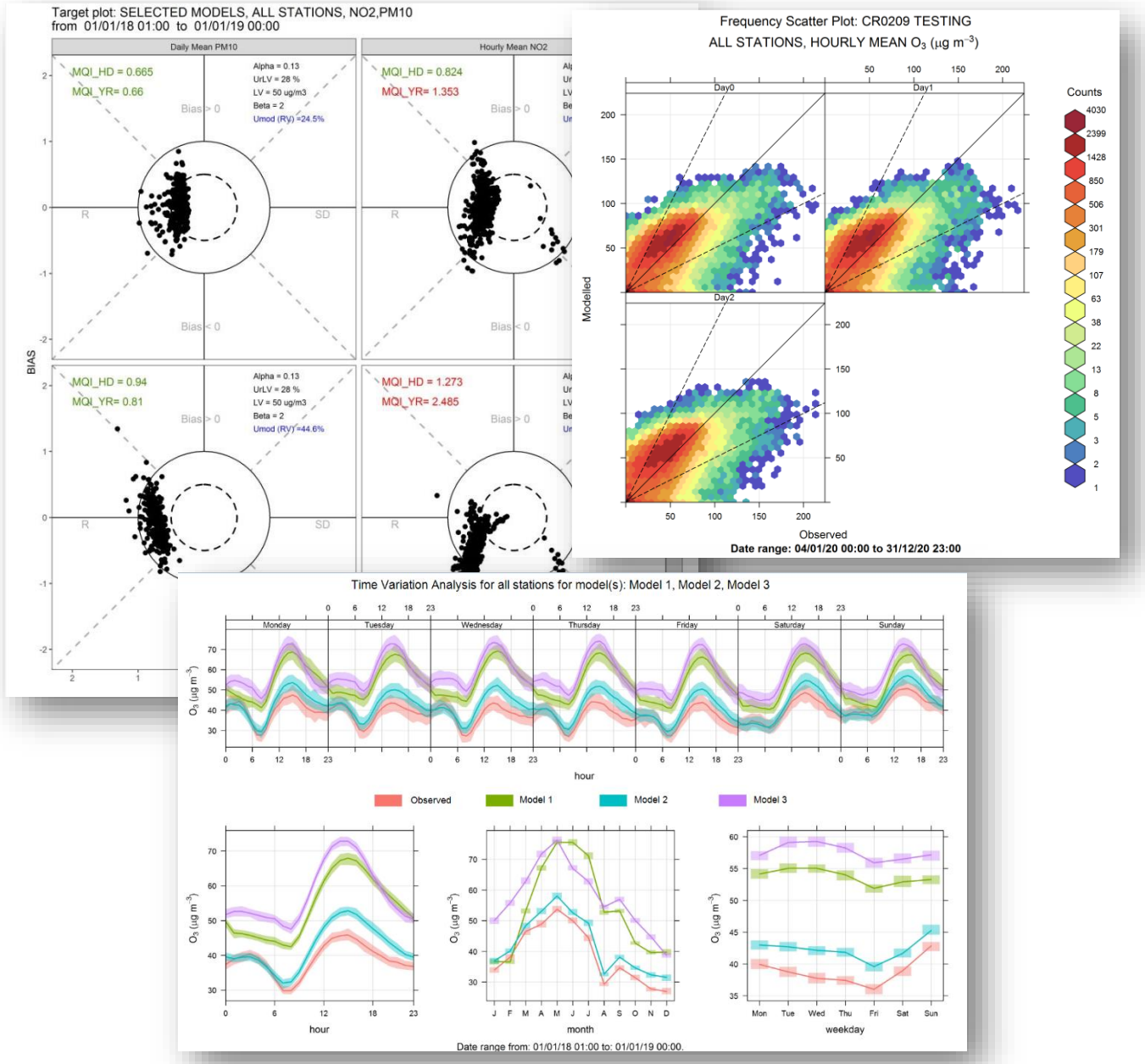
MAQS-Health

Multi-Model Air Quality System for Health Research

CERC · The University of Edinburgh · University of Birmingham · University of Hertfordshire UH · Lancaster University · Clean Air Programme

Introduction

- Motivation
  - An essential component of any application of **MAQS-Health** is validation of concentration outputs from both the regional model and the coupled system against in situ observed data
- CERC's Model Evaluation Toolkit has been developed to:
  - provide an automated, standardised method of comparing MAQS-Health coupled system output with in situ observed data
  - be compatible with all modelled data formats supported by the MAQS-Health coupled system
  - provide easy access to online measured data
  - enable easy model inter-comparison
  - output a wide range of statistics and publication-ready graphs
- The Toolkit has been used
  a) By CERC and MAQS-Health System Beta Testers for MAQS-Health
  b) By CERC on behalf of UK Defra for the recent Defra Model Inter-Comparison Exercise (4 models, 415 sites, 1 year of hourly data, 5 pollutants)



MAQS-Health

Background

CERC's Model Evaluation Toolkit was first developed by CERC during the FP7 PASODOBLE project to provide tools for **evaluating local and regional air quality forecasting models.**

The Toolkit has been **maintained and developed by CERC** since and is available to download from the CERC website with comprehensive documentation.

For the MAQS-Health project, a tool was needed for **comprehensive, consistent and repeatable inter-comparison** between regional model outputs, MAQS-Health coupled system outputs and in situ observed data.

The Toolkit has been developed to provide a **verification system for MAQS-Health** based on requirements collected in March 2020 at the MAQS-Health Stakeholder Requirements Workshop.

Modelled Data

Wide range of compatible **modelled concentration formats**: MAQS-Health, ADMS, CMAQ, WRF-CHEM, EMEP, CAMx, CHIMERE, CAMS + Generic

Supports species mapping for **composite pollutants** (e.g. NOx)

Flexible **averaging time**: n-hourly / daily / weekly / monthly / yearly data

Flexible **averaging statistic**: mean, maximum, rolling mean

Accounts for **time zone differences** between modelled and observed data

Modelled data can be provided in a **single file or multiple files**, with a date-dependent template path if required

Multiple modelled datasets can be assessed

Observed Data

Two options:

1. **Automatically access** hourly observations from 7 monitoring networks

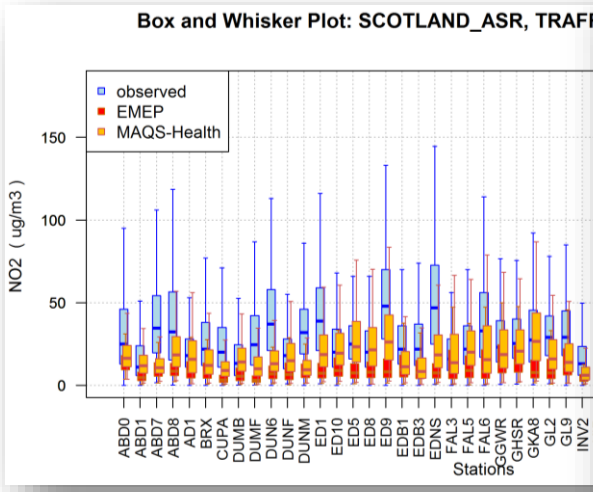
and/or

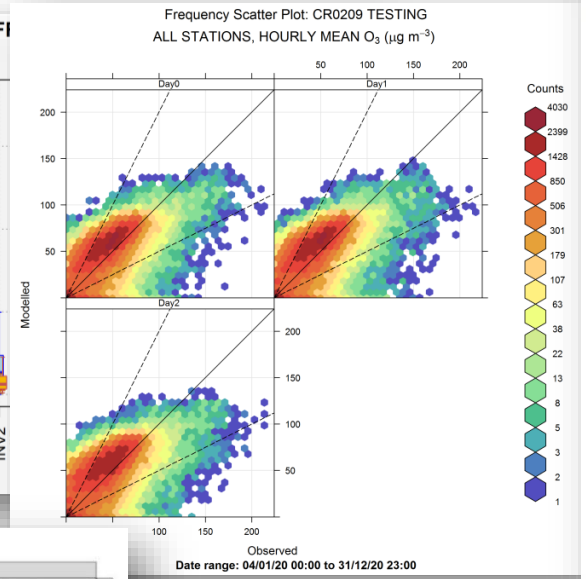
2. Provide a **file** of observed data in plain text format
   - Flexible averaging time: n-hourly / daily / weekly / monthly / yearly data
   - Flexible averaging statistic: mean, maximum, rolling mean

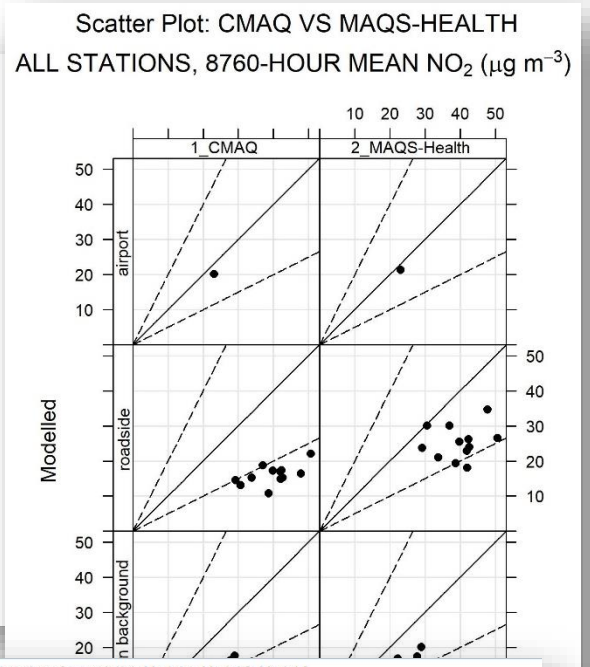
Network	Coverage	Description
EUROPE	Europe	EEA's AirBase database of European monitoring sites
AURN	UK	UK Automatic Urban and Rural Network
SAQN	Scotland	Air Quality in Scotland network
WAQN	Wales	Welsh Air Quality Network
NI	Northern Ireland	Northern Ireland Air Network
AQE	England	Air Quality England network
ERG	UK	Networks operated by the Environmental Research Group (ERG) at Imperial College London, including the London Air Network

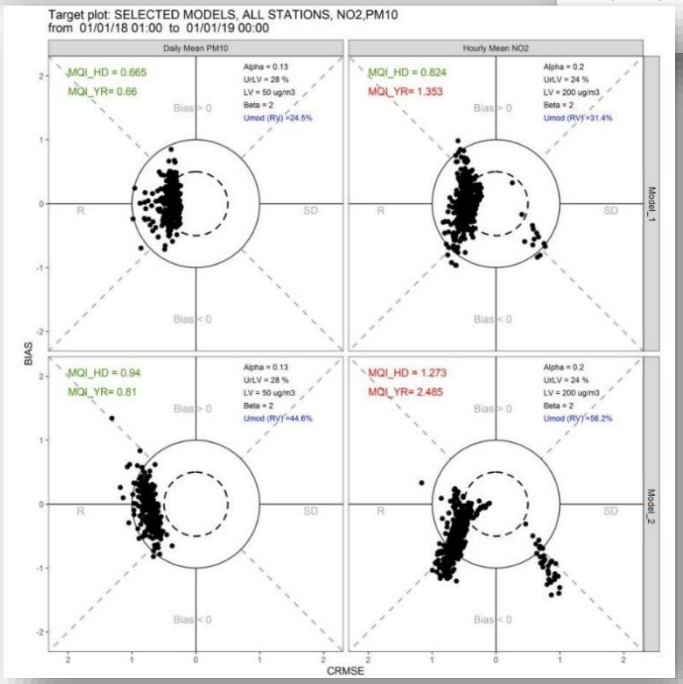
Toolkit outputs: Graphs

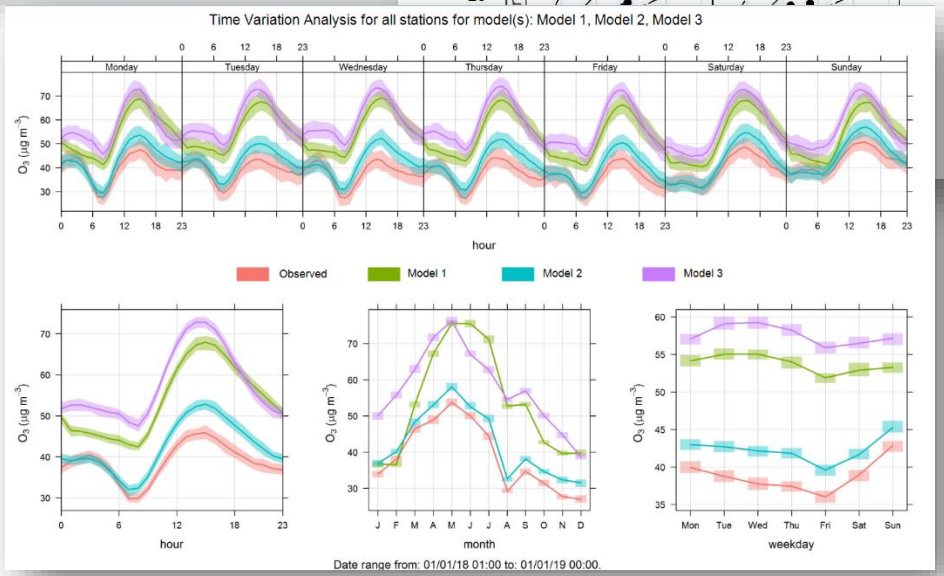
- Box plot
- Scatter plot (conventional, frequency and statistical)
- Time variation plot
- FAIRMODE Target plot, which accounts for measurement uncertainty
- Quantile-Quantile plot
- Time plot (per station and averaged over all stations)
- Polar plot (requires met data)
- Forecast index accuracy
- Forecast performance metrics
- Odds ratio skill score











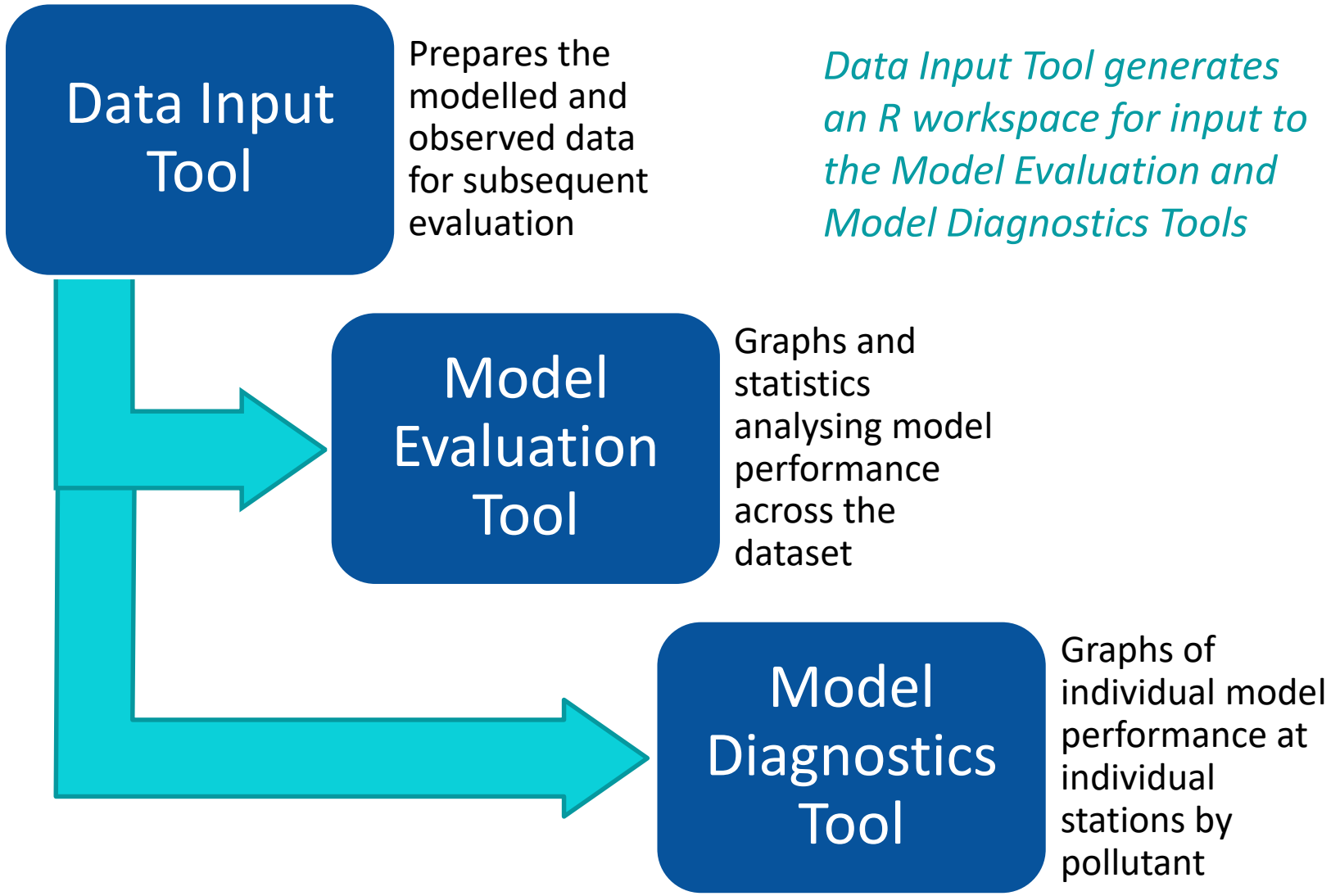
MAQS-Health

Toolkit outputs: Statistics

- The Model Evaluation Toolkit output includes CSV files that contain all the statistics shown on the graphs, grouped and filtered according to user requirements
- Statistics include:
  - N, Mean, Max, SD, MB, NMSE, R, FAC2, FB, FS, RMSE, MSE, MSE_{bias}, MSE_{var}, mMSE, NMSD, RHC, CRMSE, MGE, NMGE, COE, IOA + Box Plot statistics
  - Ozone statistics: AOT40 (forests), AOT40 (crops), SOMO35
  - FAIRMODE Target Plot statistics: MQI, MPI_{bias}, MPI_{r}, MPI_{sd}, RMSUM, per station and also 90^{th} percentile summary statistics (including MQI_{90})
  - Event detections and misses, POD, FAR, OR, ORSS, AQ Index, Episodes
    - Important for forecasting in terms of alert thresholds, but also for AQD exceedances

Running the Toolkit

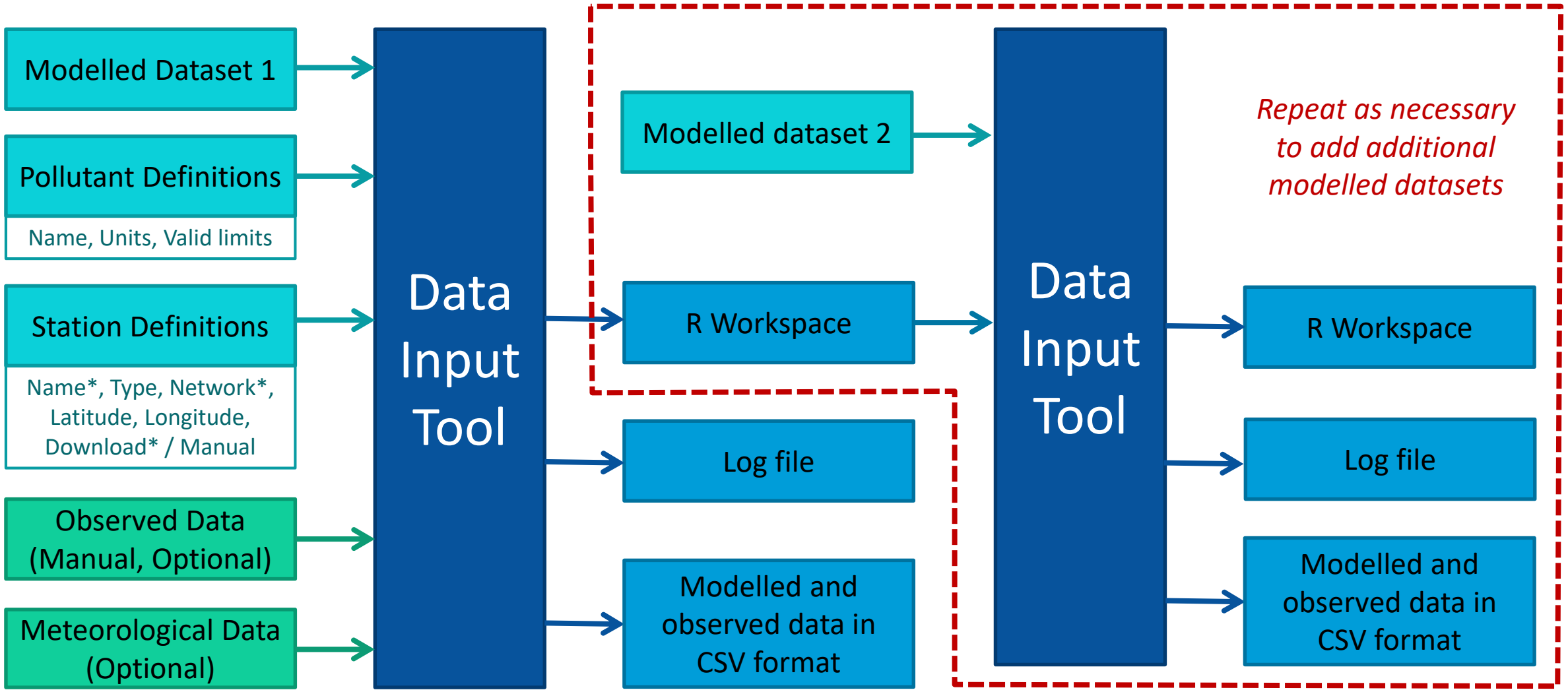
- Runs on Linux-based HPC systems and Windows
  - User Interface provided for Windows users
  - Comprehensive documentation provided for both environments
- Open source and repeatable: tools consist of an R script (utilizing many openair functions) with an input parameter file specifying settings and input file paths
- Output includes a comprehensive log file for full traceability



Data Input Tool generates an R workspace for input to the Model Evaluation and Model Diagnostics Tools

Stage 1: Data Input Tool

Prepares the modelled and observed data for subsequent evaluation, exporting the mod-obs data pairs to an R workspace



* Observed data for can be automatically downloaded for the same period as the modelled dataset for recognised station names and recognised online networks

MAQS-Health

Stage 2: Model Evaluation Tool

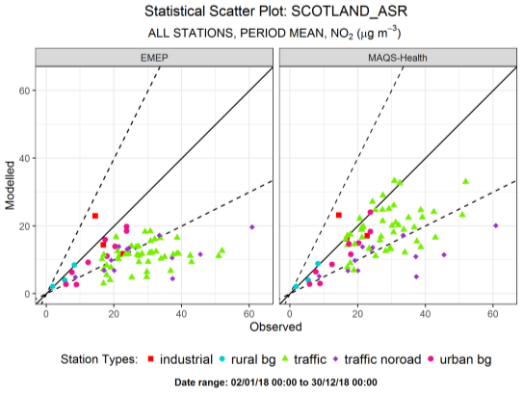
Analyses model performance across the whole dataset (or subsets of it)

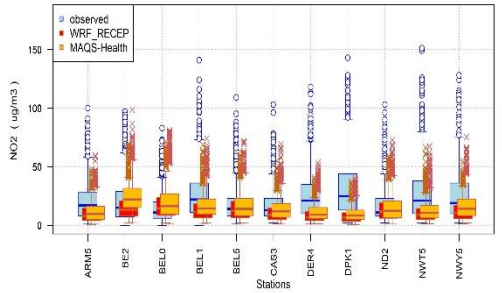
- R Workspace
- Averaging Times / Statistics
- Data Capture Thresholds
- **Subsetting**
  - Subset by Modelled datasets, Pollutants, Stations, Networks, Station Types, Time Period
- **Grouping**
  - Group by Station, Station Type, Model or Pollutant
- **Filtering**
  - Filter by Station Type, Model and / or Pollutant
- Composite pollutants
- Model colours

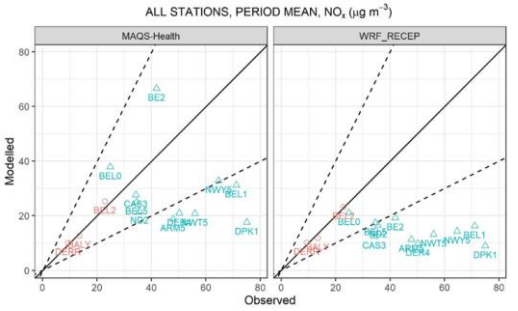
Model Evaluation Tool

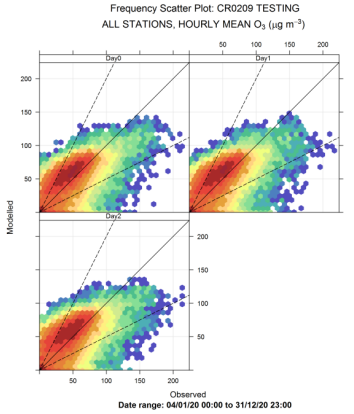
Only assesses averaging periods which have valid data for the selected modelled datasets and the observed dataset

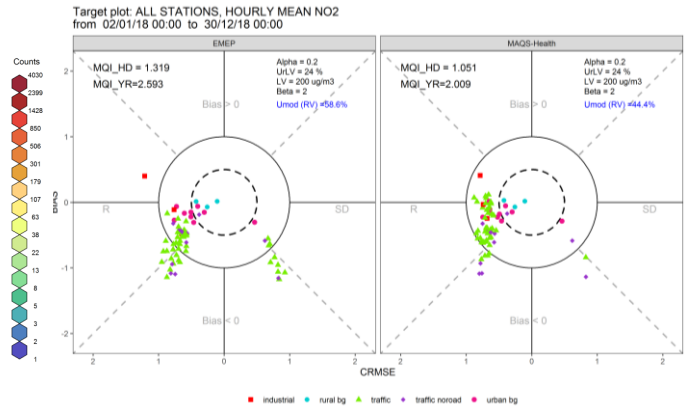
- Log file
- **Statistics**
  - CSV
- **Graphs**
  - PNG, JPG or PDF











MAQS-Health

Stage 3: Model Diagnostics Tool

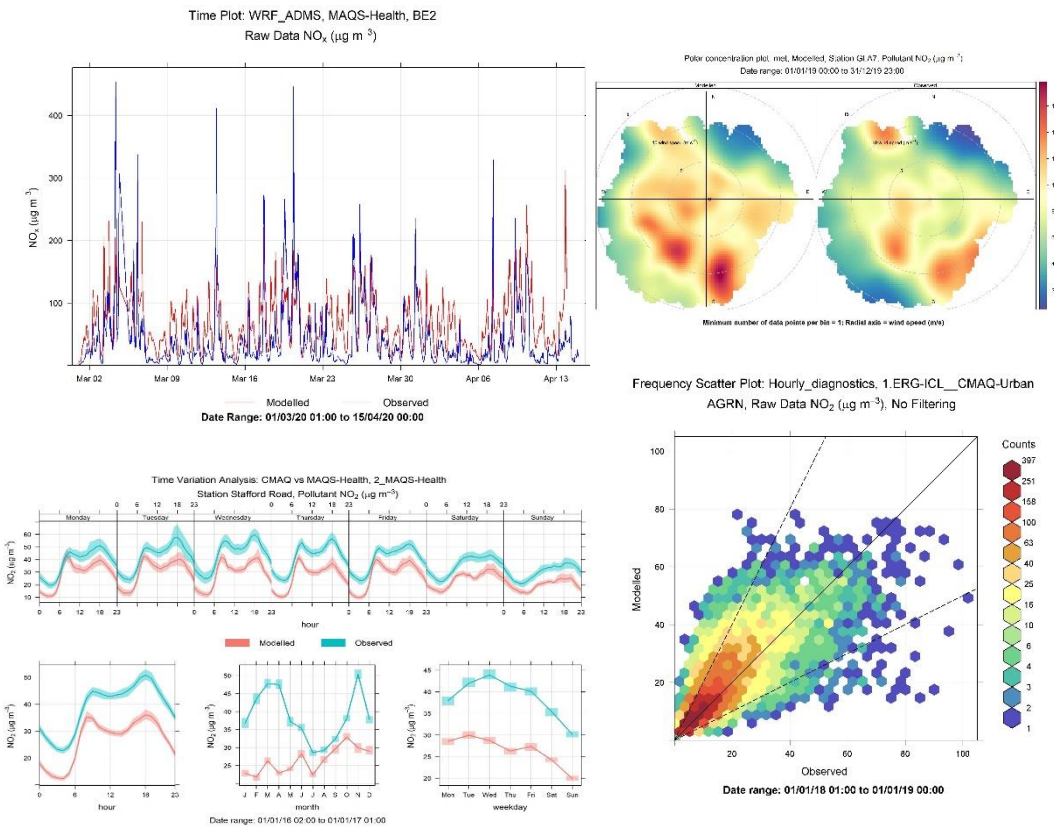
Produces graphs of individual model performance at individual stations by pollutant

- R Workspace
- Averaging Times / Statistics
- Data Capture Thresholds
- Subsetting
  - Subset by Modelled datasets, Pollutants, Stations, Time Period

Model Diagnostics Tool

Only assesses averaging periods which have valid data for the selected modelled datasets and the observed dataset

- Log file
- Graphs
  - PNG, JPG or PDF



MAQS-Health

Summary

- The Model Evaluation Toolkit provides open and intuitive tools for evaluating regional and local air quality model concentrations with observations.
- The Toolkit has been developed and tested during the Multi-model Air Quality System for Health Research (MAQS-Health) project and enables straightforward and consistent inter-comparison between regional model outputs and coupled system results from the MAQS-Health system.
- Features include automatic access to online measured data and a wide range of graphical and statistical output.
- The Toolkit is currently being prepared for general release and will be available to download in the near future.

Work supported by the UK Government's Strategic Priorities Fund (SPF) Clean Air Program, administered by the Met Office (DN424739)

amy.stidworthy@cerc.co.uk