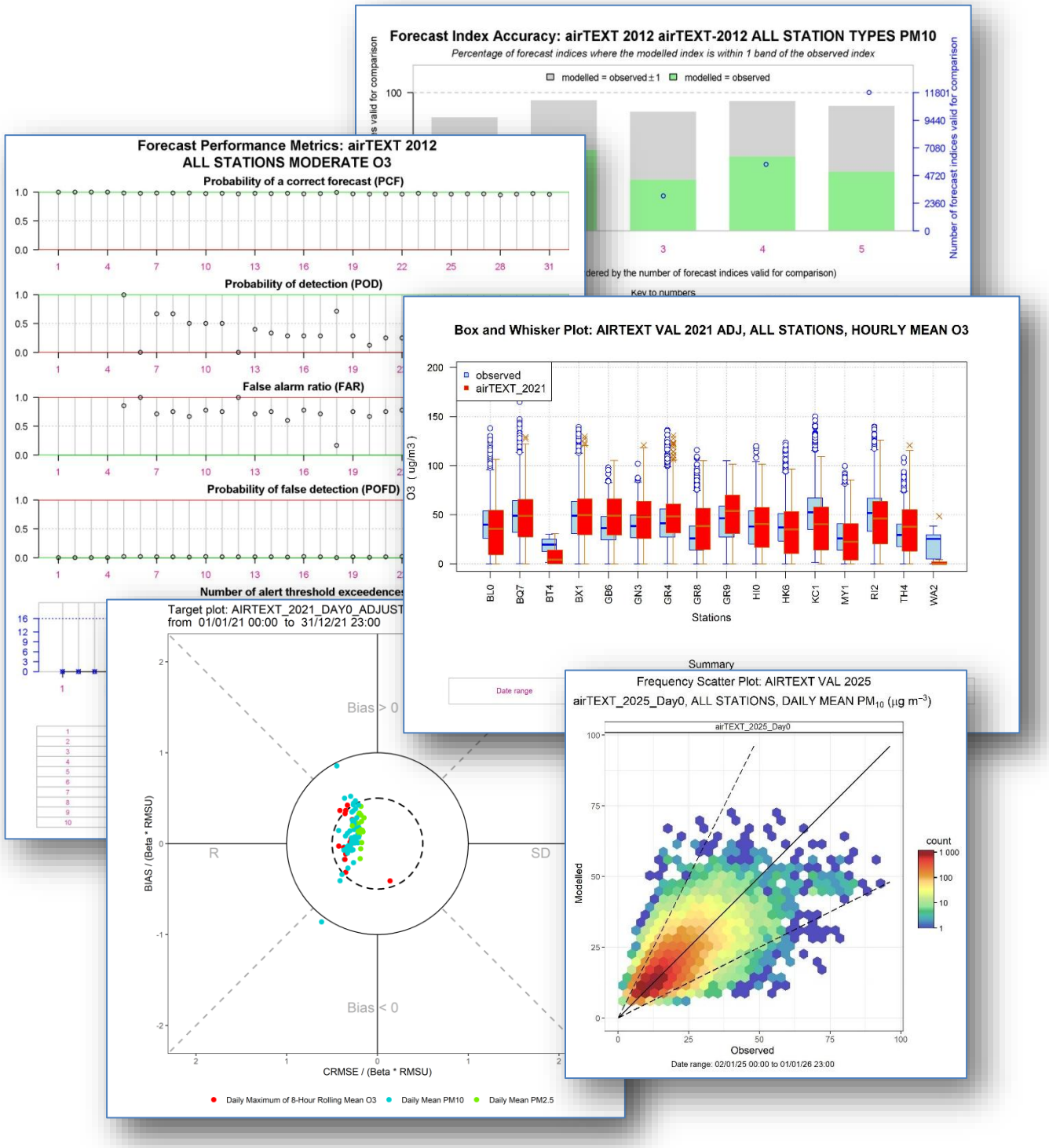
Model Evaluation Toolkit
User Guide
Windows Operating Systems
Version 5.4, June 2026



Support
The website for the Model Evaluation Toolkit is www.cerc.co.uk/ModelEvaluationToolkit. To report issues or suggest improvements, please send an email to help@cerc.co.uk.

Contents

1 Introduction

Regional and municipal governments are increasingly interested in providing services to assess and forecast local- and city-level air quality. Air quality forecasts on these scales can be disseminated to health services and the public in terms of air quality alerts, to inform and warn 'at-risk' groups about impending pollution episodes and provide advice to mitigate the risk. Local air quality modelling is critical in assessment of air quality against the EU air quality directive as it can provide high resolution maps of concentration where the population is most dense and allows the investigation of proposed mitigation measures on short or long time scales. Understanding of the benefits, limitations and performance of individual models, the input data required and the extent of the options available is often lacking. Setting standard evaluation criteria and comparing model capabilities in a structured way is therefore a crucial task.

The Model Evaluation Toolkit was built and developed under the local forecast model evaluation support work package of the EU's 7th Framework, PASODOBLE project. It was enhanced as an initiative of the MAQS-Health project (Multi-Model Air Quality System for Health Research), focused on developing a system for high resolution prediction capability for outdoor air quality[1]. The Toolkit draws on existing best practice such as the EU Joint Research Council's (JRC) FAIRMODE initiative on model evaluation [1] and the openair project tools [2,3]. Additional statistical metrics were added as part of a review of methods used to assess the performance of atmospheric dispersion models, on behalf of the Atmospheric Dispersion Modelling Liaison Committee (ADMLC) [4].

The Toolkit is a simple-to-install, user-friendly environment that guides the user through the process of evaluating model predictions of local air quality and investigating the model performance. It runs on both Windows and Linux operating systems. This guide gives instructions for running the Model Evaluation Toolkit on Windows.

The Toolkit can take modelled data from regional-scale or local-scale models as input. Observed data are *in situ* time series data. Missing data are handled if they are indicated by a standard value. As output, the Toolkit creates report-ready plots of the model performance in predicting concentrations and predicting alerts with respect to defined thresholds, for single or multiple sites, single or multiple pollutants, single or multiple modelled datasets, single or multiple site types and single or multiple monitoring networks. Results can be classified by the type of monitoring site and the pollutant for each modelled dataset. The diagnosis of model performance for individual sites and individual pollutants produces time series plots, scatter plots, polar plots and analyses with respect to month, day of the week and hour of the day. All of the plotted data are also exported to data files to provide an audit trail and make the data available for further analysis and visualisation.

[1] https://www.ukcleanair.org/projects/maqs-health-multi-model-air-quality-system-for-health-research/

2 Getting started

The Toolkit can be used on Windows and Linux operating systems and does not require any software to be purchased. Before using the Toolkit you will need to install R [5] and the Toolkit itself; this will just take a few minutes. Detailed installation instructions are given in Sections 2.1 and 2.2.

2.1 Install R

The Model Evaluation Toolkit is compatible with R version 4.5.3 or later.

Follow these step-by-step instructions to download and install R from the internet:

1. Go to http://www.r-project.org/
2. From the left-hand menu, under **Downloads**, select **CRAN**
3. Choose a CRAN mirror for your locality (in the UK, choose the mirror for the University of Bristol) and click on the link
4. Under **Download and install R** click on the link for your operating system
5. Click on **base**
6. Click to download the install program
7. Run the install program, taking care to install R in a folder where you have write privileges.

Note: If you do not have direct access to the internet from your computer, for example you access the internet through a university network, then when you install R, instead of accepting all the defaults, at the Setup screen choose not to accept all the defaults and when offered, choose 'Internet2' as the internet option. This will force R to use the same proxy settings used by Internet Explorer. The defaults for all other options can be accepted.

2.2 Install the Model Evaluation Toolkit

Follow these step-by-step instructions to install the Model Evaluation Toolkit:

1. Log on as Local Administrator for the PC.
2. The Model Evaluation Toolkit installation files will have been supplied by download link. Unzip the downloaded *.zip* file to a local folder.
3. In Explorer, browse to this folder and double-click on the file *setup.exe*. The **Welcome** window in Figure 2-1 will be launched.

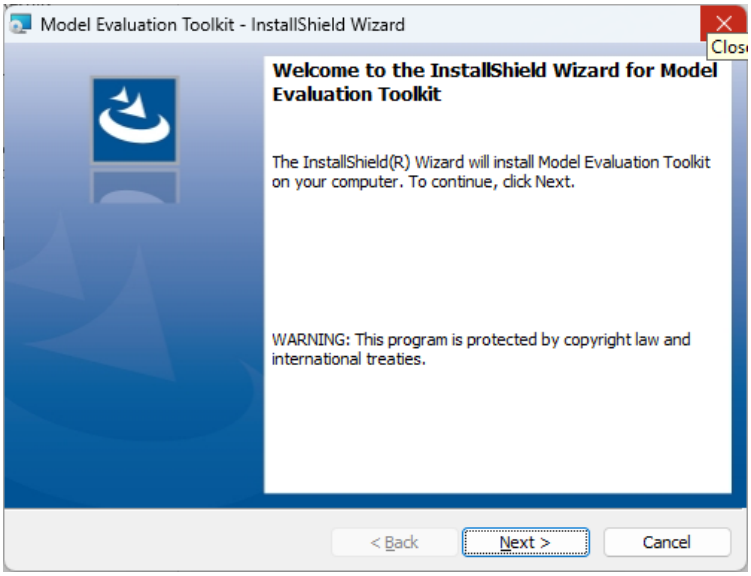


Figure 2-1 The Model Evaluation Toolkit *Welcome* screen

4. Click 'Next' on the **Welcome** screen and then enter your Customer Information as shown in Figure 2-2. Click 'Next' to proceed to the **Destination Folder** screen, as shown in Figure 2-3.

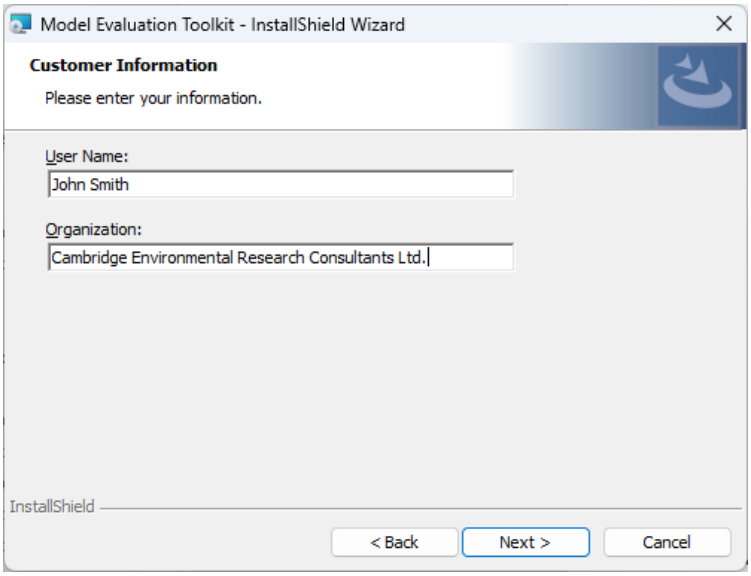


Figure 2-2 The Model Evaluation Toolkit *Customer Information* screen

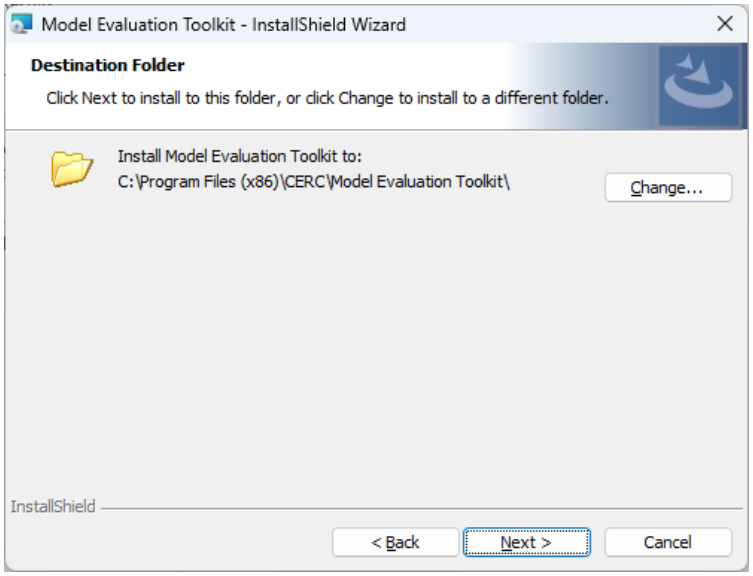


Figure 2-3 The Model Evaluation Toolkit *Destination Folder* screen

5. The default installation folder is *C:\Program Files (x86)\CERC\Model Evaluation Toolkit*. If required, use the **Change...** button to select another destination folder. Click **OK** to return to the 'Destination Folder' screen.
6. Click **Next >** to display the screen shown in Figure 2-4.

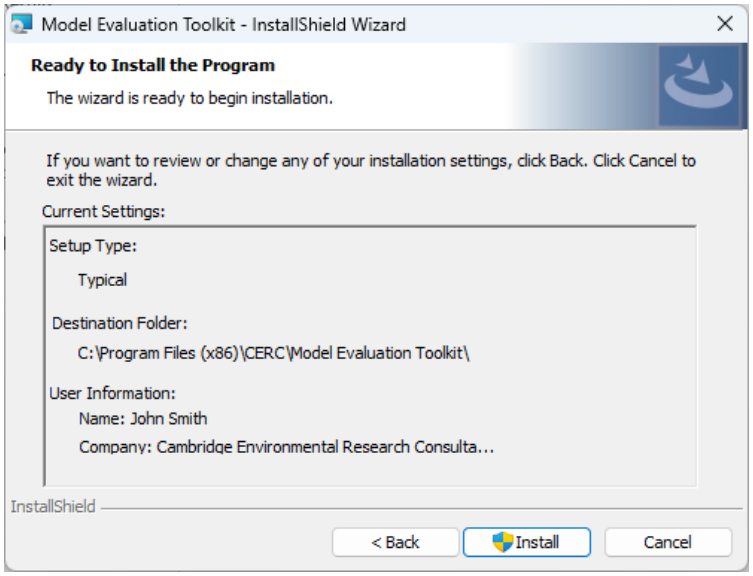


Figure 2-4 The Model Evaluation Toolkit *Ready to Install* screen

If you wish to amend any details, press the **< Back** and **Next >** buttons as appropriate.

7. Click **Install**. When the Model Evaluation Toolkit files have been successfully installed, the installer may prompt a restart, as shown in Figure 2-5. The restart can be delayed by clicking **No** but it is required for proper use of the Model Evaluation Toolkit. A final screen will now appear, as shown in Figure 2-6.

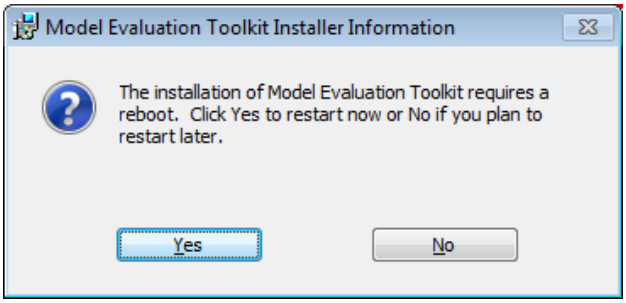


Figure 2-5 The installer may prompt you for a restart

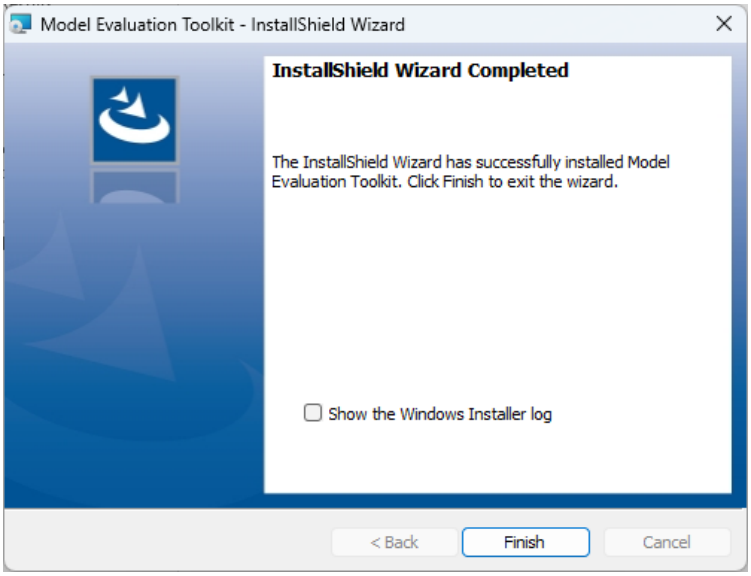


Figure 2-6 The Model Evaluation Toolkit *InstallShield Wizard Completed* screen

8. Click **Finish** to complete the installation. The install wizard automatically creates three shortcuts in the **Programs** menu under **Model Evaluation Toolkit**; one each for the **Data Input Tool**, **Model Diagnostics Tool** and **Model Evaluation Tool**.

9. Open the Data Input Tool and select **Update monitoring station metadata** from the **Tools** menu (Figure 2-7). This will open a processing window and download all relevant data for the station networks AQE, AURN, EUROPE, KCL, NI, SAQN and WAQN.

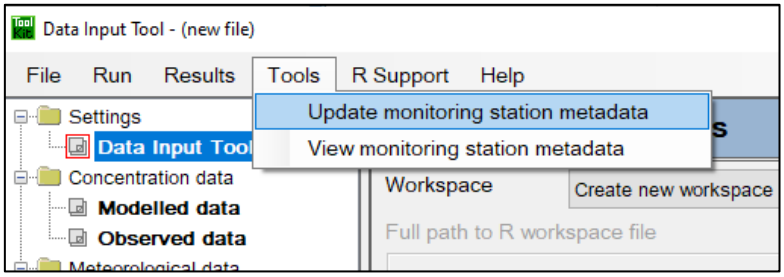


Figure 2-7 Downloading monitoring station metadata

Once downloaded, you can view the station data files by clicking **Tools | View monitoring station metadata**. A file is produced for each network.

3 Using the Toolkit

The Model Evaluation Toolkit consists of three Tools:

1. Data Input

   This Tool processes the modelled and observed concentration data, saving the processed data in an R workspace and (optionally) a CSV file.

2. Model Evaluation

   This Tool processes all or some of the modelled and observed data in the R workspace file, producing graphical outputs and (optionally) CSV files.

3. Model Diagnostics

   This Tool provides further investigation of the performance of a model at a particular monitoring station(s) for a particular pollutant(s).

This section of the User Guide gives step-by-step instructions for using each Tool in the Toolkit.

Tip: Various supplementary data input files are required by the Model Evaluation Toolkit, all of which are described in this User Guide. The \Data sub-folder of the Model Evaluation Toolkit installation folder includes examples of these files.

3.1 Data Input Tool

The Data Input Tool processes modelled and observed data, saving it in an R workspace file which later can be imported into the Model Evaluation and Model Diagnostics Tools for analysis.

The Tool supports both modelled data for specific station locations ('point data') and gridded modelled data. Gridded data is converted into point data for station locations by one of two methods:

i. Interpolation of gridded data to station locations

ii. Taking the value at the grid point nearest to the station location

Observed data can either be automatically downloaded from the internet and imported (UK only) or input using simple-format CSV files.

The Toolkit supports the evaluation of multiple modelled datasets, either from multiple models or from multiple runs of the same model, over the same time period. Initially, a new R workspace is created with one modelled dataset and an associated observed dataset. Further modelled datasets can then be added to the existing R workspace. There is no limit to the number of datasets that can be added to an R workspace, but memory issues may be encountered if the number and scope of the datasets are large.

Modelled data is assumed to use the hour-ending convention for hourly data, i.e. hourly data is labelled with the time at the end of the hour. Hourly observed data can be either hour-ending or hour-ending.

This section gives a step-by-step guide to using the Data Input Tool.

Step 1: **Open the Data Input Tool**

Figure 3-1 shows the opening screen of the Data Input Tool interface.

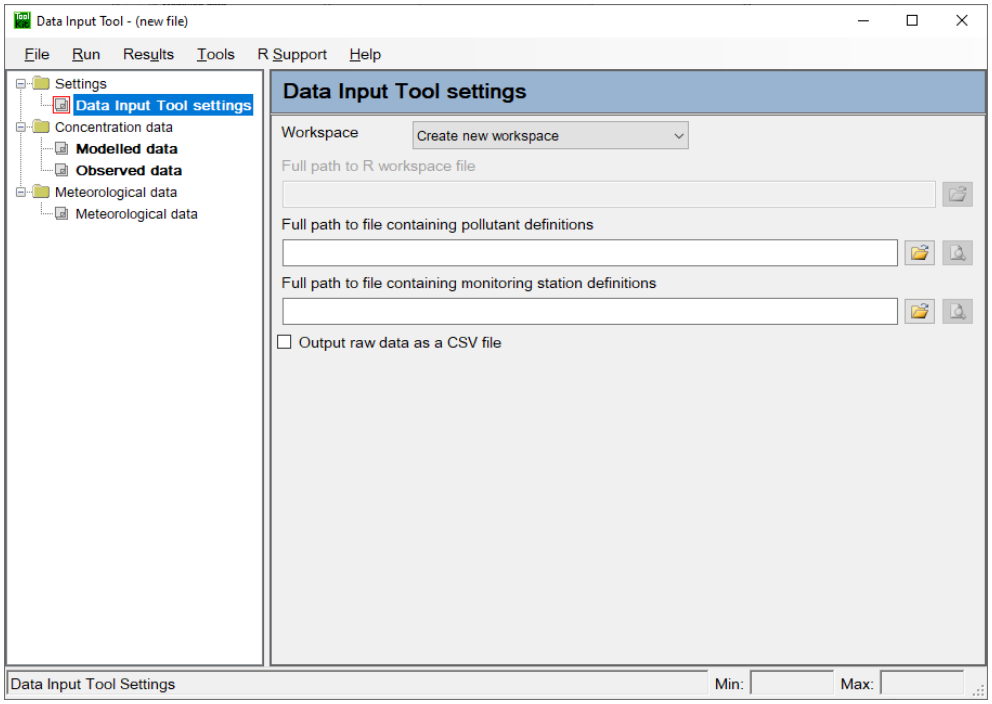


Figure 3-1 The Data Input Tool Settings screen

Step 2: **Select whether you wish to create a new workspace or add modelled data to an existing workspace**

You can choose to either create a new R workspace or to add a modelled dataset to an existing R workspace created with the Data Input Tool.

Note:

- *When creating a new workspace, the data range of the modelled dataset should cover the whole time period you wish to evaluate, including the start and end times.*
- *When adding a modelled dataset to an existing R workspace, the R workspace must have been created by the same version of the Data Input Tool as you are currently using.*
- *When adding multiple modelled datasets to a workspace, the modelled data averaging time and statistic must be the same for all modelled datasets. Refer to Table 3-1 for details.*

Step 3: **<u>If creating a new workspace</u>, browse to select a pollutant definitions file**

The pollutant definitions file is a CSV (comma-separated) file, containing a list of all the pollutants for which you wish to process data. For each pollutant, a set of parameters must be set, as described in Table 3-1. **It is very important these**

parameters are set correctly for your data. An example pollutant definitions file, *pollutants.csv*, is included in the *Data* folder, and is shown in Figure 3-2.

	A	B	C	D	E	F	G	H
1	pollutant	output.units	conv.ugm3.ppb	avg.time	statistic	min.allowed	max.allowed	ozone
2	no2	ug/m3	0.52	1 hour	mean	0	1000	no
3	o3	ug/m3	0.5	1 hour	mean	0	1000	yes
4	pm10	ug/m3	1	1 hour	mean	0	1000	no
5	pm2.5	ug/m3	1	1 hour	mean	0	1000	no
6								

Figure 3-2 Example pollutant definitions file *pollutants.csv*

Column header	Description	Allowed values
pollutant	Name of the pollutant to be used in all output	*n/a*
output.units	Concentration units to be used in all output	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
conv.ugm3.ppb	Conversion factor from µg/m^3 to ppb (used for unit conversions)	Any numeric value
avg.time	Averaging time for output	Integer value and unit (e.g. '1 day', '8 hours') or just unit (e.g. 'hour', 'year')
statistic	Statistic to be output	'max', 'mean' or 'rolling mean'
min.allowed	Minimum allowed concentration value	Any numeric value
max.allowed	Maximum allowed concentration value	Any numeric value
ozone	Indicates which pollutant name corresponds to ozone (if any), to enable ozone-specific statistics to be output	yes, no

Table 3-1 Data included in the pollutant definitions file

Step 4: **<u>If creating a new workspace</u>, browse to select a stations definitions file**

The stations definitions file is a CSV file containing a list of all the stations for which you wish to process data, and some associated parameters, as described in Table 3-2.

Column header	Description	Allowed values
Name	Station name or code. This is the value that will be output on graphics and statistic files.	Any string
Type	Station type, for example urban background, rural or roadside. This will enable filtering or grouping by station type if required.	Any string
Network	Station network e.g. AURN, KCL, CSV.	String. Must be one of: KCL, AURN, AQE, WAQN, SAQN, NI, EUROPE if auto.download is yes.
Latitude	Latitude of station position (only required if gridded modelled data are used)	Numeric value (°)
Longitude	Longitude of station position (only required if gridded modelled data are used)	Numeric value (°)
Auto.download	Specifies whether observed data for each station should be automatically downloaded or whether it will be supplied as a CSV file	yes, no

Table 3-2 Data included in the station definitions file

This data can be found in the station metadata files, which can be found in the *install\Monitoring_stations_metadata* folder. You can open this folder directly from the interface by selecting **Tools | View monitoring station metadata**.

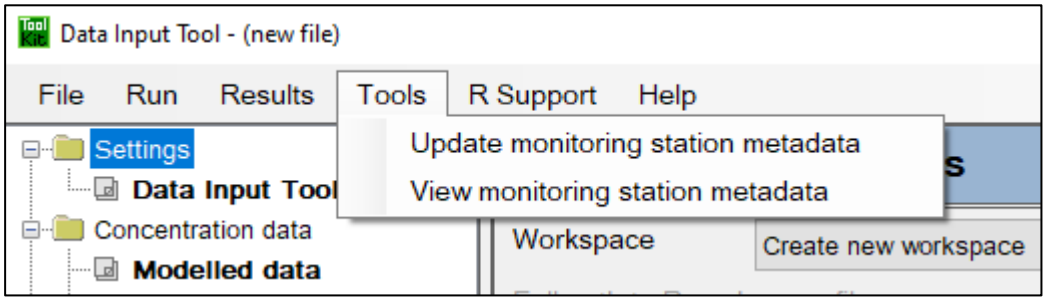


Figure 3-3 Selecting the option to view monitoring station metadata

The *\Data* folder includes the example station definition files *stations.csv* (for use with modelled data at station locations) and *stations_latlon.csv* (for use with gridded modelled data). *stations_latlon.csv* is shown in Figure 3-4.

	A	B	C	D	E	F	G
1	name	type	latitude	longitude	network	auto.download	
2	ABD	Urban Background	57.15736	-2.09428	AURN	yes	
3	ABD9	Urban Background	57.1574	-2.09477	AURN	yes	
4	ABD7	Urban Traffic	57.14456	-2.10647	AURN	yes	
5	ABD8	Urban Traffic	57.13389	-2.0942	AURN	yes	
6	ARM6	Urban Traffic	54.35373	-6.65456	AURN	yes	
7	AH	Rural Background	52.50385	-3.03418	AURN	yes	
8	ACTH	Rural Background	55.79216	-3.2429	AURN	yes	
9	BAAR	Urban Traffic	54.85149	-6.27496	AURN	yes	
10	BALM	Urban Background	54.8616	-6.25087	AURN	yes	
11	BARN	Urban Background	53.58005	-1.47587	AURN	yes	
12	BAR2	Urban Background	53.55593	-1.48515	AURN	yes	
13	BAR3	Urban Background	53.56292	-1.51044	AURN	yes	
14	BPLE	Urban Traffic	51.07479	-4.04192	AURN	yes	
15	BHA4	Urban Traffic	51.39092	-2.35503	AURN	yes	
16	BATH	Urban Traffic	51.39113	-2.35416	AURN	yes	
17	BEL2	Urban Background	54.59965	-5.92883	AURN	yes	
18	BEL4	Suburban Background	54.59126	-5.89546	AURN	yes	
19	BEL	Urban Background	54.58653	-5.90167	AURN	yes	

Figure 3-4 Example station definitions file for use with gridded modelled data, *stations_latlon.csv*

Note: If you are using automatically-downloaded data, the station names must match the official station codes.

Step 5: **<u>If adding data to an existing workspace,</u> browse to find the existing workspace**

If you have selected the option to **Add modelled data to an existing workspace**, click the icon next to the **Full path to R workspace file** box and browse to the required workspace.

Step 6: **Select the option to output raw data to a CSV file if required**

The processed data generated by the Data Input Tool will always be output as an R workspace for analysis in the Model Evaluation Tool and/or Model Diagnostics Tool.

In addition, there is an option to output the processed data as a CSV file, which can be useful for further analysis using other programs such as Excel. Select this option if required.

Step 7: **The Modelled data screen**

Move to the **Modelled data** screen (Figure 3-5) by clicking on the title in the menu on the left.

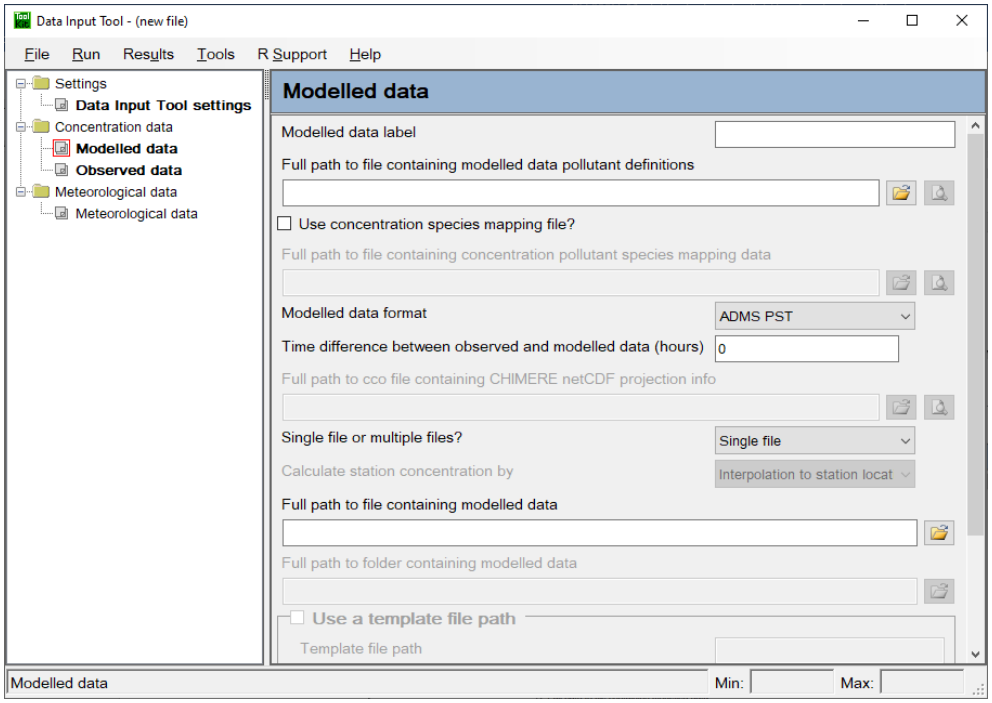


Figure 3-5 The Data Input Tool Modelled data screen

Step 8: **Supply a label for the modelled data**

Supply a label for the set of modelled data you are using. This label will be used in the output produced by the Toolkit to identify the modelled dataset. The label should be alpha-numeric and not include any special characters such as '&', '!' or commas, and should not be "mod". An error will be shown in the log file if any of these values are used.

Step 9: **Select the modelled data pollutant definitions file**

The modelled data pollutant definitions file must be in CSV format. It contains a list of all the pollutants for which you wish to process data, and some associated parameters for each pollutant. **It is very important these parameters are set correctly for your data.** Refer to Table 3-3 for details of the information required.

An example modelled data pollutant definitions file, *pollutants_mod.csv*, can be found in the *Data* folder, and is shown in Figure 3-6.

	A	B	C	D	E	
1	pollutant	mod.alias	mod.units	mod.avg.time	mod.statistic	
2	no2	no2_conc	ug/m3	1 hour	mean	
3	o3	o3_conc	ug/m3	1 hour	mean	
4	pm10	pm10_conc	ug/m3	1 hour	mean	
5						

Figure 3-6 Example modelled data pollutant definitions file *pollutants_mod.csv*

Column header	Description	Allowed values
pollutant	Name of the pollutant to be used in all output. If a species mapping file is used this should match the species mapping file row pollutants.	*n/a*
mod.alias	Pollutant name as it appears in the modelled data. Not required if species mapping file is used.	*n/a*
mod.units	Units that apply to the modelled data. Not required if species mapping file is used.	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
mod.avg.time	Averaging time of the modelled data, in hours, days, months or years (minimum 1 hour)	Integer value and unit (e.g. '1 day', '8 hours') or just unit (e.g. 'hour', 'year')
mod.statistic	Statistic that applies to the modelled data	'max', 'mean' or 'rolling mean'

Table 3-3 Data included in the modelled data pollutant definitions file

Note: The 'mod.alias' settings will be the same as 'pollutant' unless the pollutant names within the modelled data are different to 'pollutant', in which case use the aliases to make sure the correct fields are extracted from the data.

Step 10: Select whether species mapping is required

Adding a species mapping file facilitates the possibility of evaluating composite pollutants. Species mapping allows pollutants such as NO_x and PM_{10} to be evaluated against measured data for regional models where only components are output, for example NO and NO_2 or constituent PM particles.

Selecting **Use concentration species mapping file** enables you to specify the path to the file.

Step 11: Provide the file path to the species mapping file

To perform species mapping, a species mapping file must be supplied, containing conversion factors of each component species to a pollutant in ug m^{-3}.

An example species mapping file, *species_mapping.csv*, can be found in the *\Data* folder, and is shown in Figure 3-7. The file is in CSV format. The first 4 rows contain:

- The keyword SPECIESMAPVERSION2
- The keyword ROWSTOCOLUMNS, indicating that there is a row of data for each modelled pollutant, and a column for each observed pollutant
- Number of columns of pollutant data
- Number of rows of pollutant data

There then follows a table of conversion factors for converting concentrations of the modelled pollutants to concentrations (in ug m^{-3}) of the observed pollutants.

So, if the example file shown in Figure 3-7 were used, the modelled CO concentration to be compared with observed values would be calculated by multiplying the model output value by 1160, and the modelled PM10 concentration would be calculated by summing the concentrations of the constituent pollutants PNO3, PS04, …BSOA.

	A	B	C	D	E	F	G	H	I	J
1	SPECIESMAPVERSION2									
2	ROWSTOCOLUMNS									
3	7									
4	19									
5	Variables	O3	NO2	NOX	SO2	CO	PM10	PM2.5		
6	O3	2000.0	0.0	0.0	0.0	0.0	0.0	0.0		
7	NO2	0.0	1900.0	1900.0	0.0	0.0	0.0	0.0		
8	NO	0.0	0.0	1900.0	0.0	0.0	0.0	0.0		
9	SO2	0.0	0.0	0.0	2700.0	0.0	0.0	0.0		
10	CO	0.0	0.0	0.0	0.0	1160.0	0.0	0.0		
11	PNO3	0.0	0.0	0.0	0.0	0.0	1.0	1.0		
12	PSO4	0.0	0.0	0.0	0.0	0.0	1.0	1.0		
13	PNH4	0.0	0.0	0.0	0.0	0.0	1.0	1.0		
14	POA	0.0	0.0	0.0	0.0	0.0	1.0	1.0		
15	PEC	0.0	0.0	0.0	0.0	0.0	1.0	1.0		
16	FPRM	0.0	0.0	0.0	0.0	0.0	1.0	1.0		
17	FCRS	0.0	0.0	0.0	0.0	0.0	1.0	1.0		
18	CPRM	0.0	0.0	0.0	0.0	0.0	1.0	0.0		
19	CCRS	0.0	0.0	0.0	0.0	0.0	1.0	0.0		
20	NA	0.0	0.0	0.0	0.0	0.0	1.0	0.0		
21	PCL	0.0	0.0	0.0	0.0	0.0	1.0	0.0		
22	PH2O	0.0	0.0	0.0	0.0	0.0	0.0	0.0		
23	ASOA	0.0	0.0	0.0	0.0	0.0	1.0	1.0		
24	BSOA	0.0	0.0	0.0	0.0	0.0	1.0	1.0		
25										

Figure 3-7 Example species mapping file *species_mapping.csv*

Step 12: Select the modelled data format

The formats of modelled data supported by the Data Input Tool are shown in Table 3-4.

Modelled data format	Description
ADMS PST	Point receptor output format from the ADMS suite of atmospheric dispersion models.
ADMS output netCDF	Gridded output format from the ADMS suite of atmospheric dispersion models.
CMAQ netCDF as I/O API	Gridded modelled data from the CMAQ model.
CAMx netCDF as I/O API	Gridded modelled data from the CAMx model.
WRF-Chem netCDF	Gridded modelled data from the WRF-Chem model.
EMEP netCDF	Gridded modelled data from the EMEP model.
CHIMERE netCDF	Gridded modelled data from the CHIMERE model.
Generic netCDF	Generic gridded modelled data format used as input in the Multi-Model Air Quality System (MAQS).
CAMS Regional ADS netCDF	Gridded data from the Copernicus Atmosphere Monitoring Service (CAMS) regional (European) ensemble forecast/analysis air quality product, extracted using the CAMS Atmosphere Data Store (ADS).
CAMS Regional Reanalysis netCDF	Gridded data from the Copernicus Atmosphere Monitoring Service (CAMS) regional (European) reanalysis product, downloaded from regional.atmosphere.copernicus.eu
Defra Modelled Background Maps CSV	Data output from UK DEFRA's Modelling of Ambient Air Quality project
Generic CSV	User-supplied CSV file containing modelled data. Refer to Section 8 for details of the required format.

Table 3-4 Description of modelled data formats supported by the Model Evaluation Toolkit

Step 13: Enter the time difference between the observed data and the modelled data

Enter the number of hours' difference between the time zone of the observed data and that of the modelled data. The difference should be positive if the observed data time zone is ahead of the modelled data time zone, and negative if the observed data time zone is behind the modelled data time zone.

Step 14: <u>If you are using data from CHIMERE</u>, provide the path to the netDCF projection information (.cco) file

CHIMERE netCDF output files do not contain any information on the projection of the file. An accompanying text file (with extension .cco) contains the projection information (projection type and attributes). Enter the full path to the .cco file. You can use the folder icon to navigate to the file path.

Step 15: Select whether the modelled data are in a single file or multiple files

You can select either a single file or a whole folder of data files (for example, you may have a separate modelled data file for each month of your analysis period).

Step 16: <u>If the modelled data format is gridded</u>, select whether the station concentration should be calculated by interpolation or grid value

If the selected modelled data is in gridded format, the option to select a method for calculating the station concentration will be enabled. The concentration at each station can be calculated either by interpolation or by taking the gridded value nearest to the station.

Step 17: Enter the file path to the modelled data

If the modelled data is contained in a single file:

- Enter the full path to the file containing modelled data. You can use the folder icon to navigate to the file path.

If the modelled data is contained in multiple files:

- Enter the full path to the folder containing the modelled data. You can use the folder icon to navigate to the folder path.
- There is the option to specify a template file path. This option can be used if the individual modelled data files have names that include date and/or time information. The following placeholders can be used as part of the folder and/or file names:
    - **%Y** – 4-digit year
    - **%M** – 2-digit month
    - **%D** – 2-digit day of the month

- **%J** – 3-digit day of the year
- **%h** – 2-digit hour

So if, for example, your modelled data files were named *AllReceptors_2022_05_07.pst* (containing data for 7 May 2022) and so on, you would enter the following template file path:

AllReceptors_%Y_%M_%D.pst

- If you have specified a template file path, enter the start and end date and time of the modelled dataset, in the format **YYYYMMDDHH**.

Step 18: <u>If the modelled data format is Generic CSV</u>, specify the data separator

Use the drop-down menu to select the appropriate separator for your data.

Step 19: <u>If the modelled data format is Generic CSV</u>, specify the missing data indicator

In the text box enter the value that is used in the modelled data file to indicate missing data. This is commonly -999 in ADMS files, or NA in other files.

Step 20: The Observed data screen

Move to the **Observed data** screen (Figure 3-8) by clicking on the title in the menu on the left.

If you selected the option to add data to an existing workspace on the **Data Input Tool settings** screen, there is no need to specify the details of the observed data, as they are already included in the workspace. Therefore, de-select the **Observed data** option and skip to Step 27:.

If you selected the option to create a new workspace on the **Data Input Tool settings** screen, then leave the **Observed data** option selected and continue.

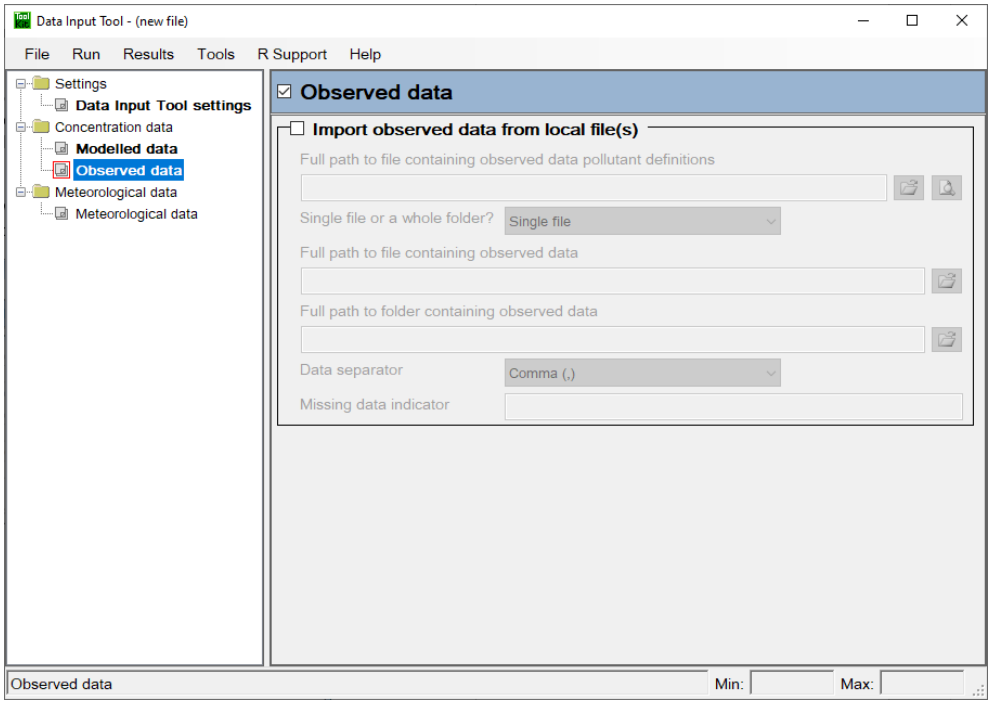


Figure 3-8 The Data Input Tool Observed data screen

Step 21: Select whether to import observed data from local CSV file(s)

Observed data can be auto-downloaded from a selection of networks during the Data Input Tool run, or supplied by the user in one or more CSV files. Details of the networks from which data can be auto-downloaded are given in Section 9.

If you will be using auto-downloaded data, skip to Step 27:.

If you will be supplying observed data in CSV file(s), select the option to **Import observed data from local file(s)** and continue with the steps below. See Section 8 for details of the required CSV file format.

Step 22: Select the observed data pollutant definitions file

The observed data pollutant definitions file must be in CSV format with a comma separator. It contains a list of all the pollutants for which you wish to process data, and some associated parameters for each pollutant. Entries must be included for any network containing one or more stations for which data will not be auto-downloaded (that is, stations for which **auto.download** is set to **no** in the station definitions file). **It is very important these parameters are set correctly for your data.** Refer to Table 3-5 for details of the information required.

An example observed data pollutant definitions file, *pollutants_obs.csv*, can be found in the *Data* folder, and is shown in Figure 3-9.

	A	B	C	D	E	F	G
1	pollutant	obs.alias	obs.units	obs.avg.time	obs.statistic	network	hour.ending
2	no2	no2	ug/m3	1 hour	mean	csv	no
3	o3	o3	ug/m3	1 hour	mean	csv	no
4	pm10	pm10	ug/m3	1 hour	mean	csv	no
5	pm2.5	pm25	ug/m3	1 hour	mean	csv	no
6							

Figure 3-9 Example observed data pollutant definitions file, *pollutants_obs.csv*

Column	Description	Allowed values
pollutant	Name of the pollutant to be used in all output	*n/a*
obs.alias	Pollutant name as it appears in the observed data	*n/a*
obs.units	Units that apply to the observed data	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
obs.avg.time	Averaging time of the observed data, in hours, days, months or years (minimum 1 hour)	Integer value and unit (e.g. '1 day', '8 hours') or just unit (e.g. 'hour', 'year')
obs.statistic	Statistic that applies to the observed data	'max', 'mean' or 'rolling mean'
network	Name of observation network	*n/a*
hour.ending	Indicates how data for a particular hour are labelled	'yes' if labelled with time at end of hour, 'no' if labelled with time at start of hour

Table 3-5 Data included in the observed data pollutant definitions file

Step 23: Select whether the observed data will be supplied in a single file or all files in a folder

You can select either a single file or a whole folder of data files (for example, you may have a separate observed data file for each month of your analysis period).

Step 24: Enter the file path for the observed data

If the observed data is contained in a single file:

- Enter the full path to the file containing the observed data. You can use the folder icon to navigate to the file path.

If the observed data is contained in multiple files:

- Enter the full path to the folder containing the observed data. You can use the folder icon to navigate to the file path.

Step 25: Specify the observed data separator

Use the drop-down menu to select the appropriate separator for your data.

Step 26: Specify the missing data indicator

In the text box enter the value that is used in the observed data file to indicate missing data.

Step 27: The Meteorological data screen

Move to the **Meteorological data** screen (Figure 3-10). Here, there is the option to add meteorological data to your workspace. This data can be output from the Data Input Tool in ADMS *.met* format, and is used by the Model Diagnostic Tool when creating polar plots.

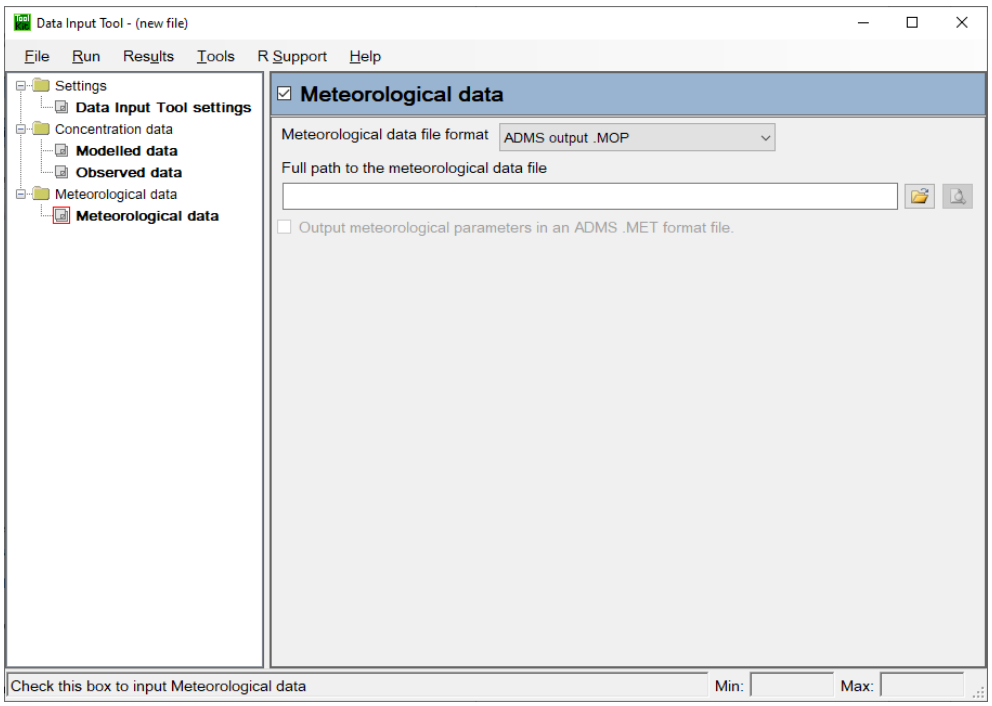


Figure 3-10 The Data Input Tool Meteorology data screen

If you wish to add meteorological data to your workspace, select the **Meteorological data** option.

The formats of meteorological data supported by the Data Input Tool are shown in Table 3-6.

Meteorological data format	Description
ADMS output .MOP	ADMS meteorological output file
ADMS input/output .MET	ADMS meteorological input file
ADMS output .PRO	ADMS boundary layer profile output file
ADMS output .MIO	MAQS meteorological output file
WRF gridded netCDF	Meteorological data from the WRF model
ADMS generic gridded netCDF	Gridded meteorological data file in netCDF format used in MAQS

Table 3-6 Types of meteorological data supported by the Data Input Tool

Select the format of your data from the drop-down list and browse to select the file. Finally, select whether to output the data in *.met* format.

Step 28: Save the settings file and run the Data Input Tool

To save the settings file, select **Save** or **Save As...** from the **File** menu.

To run the Tool, select **Run** from the toolbar and then **Data Input Tool** (Figure 3-11).

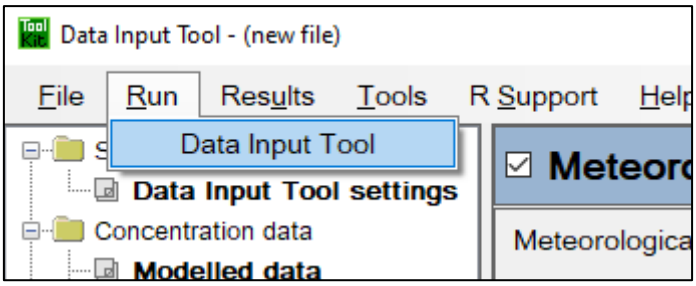


Figure 3-11 Running the Data Input Tool

Step 29: Check the log file

The Data Input Tool generates a log file when it runs; this records all the processing messages that are shown in the run screen as well as any error and warning messages issued. It is good practice to check the log file for any problems that may have occurred.

To see the log file, select **View log** from the **Results** menu. This option will open the log file in your preferred text editor (Notepad by default).

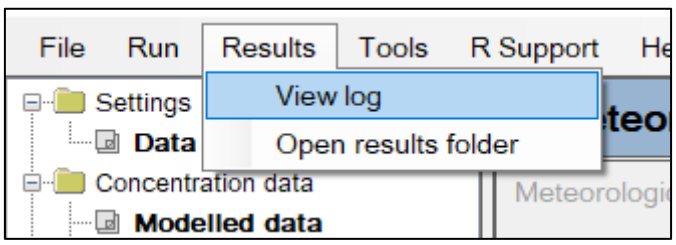


Figure 3-12 Viewing the log file

*Tip: If **View log** is not available to select, no log file exists for the currently loaded .tki file.*

Step 30: View results folder

To open the results folder for the currently loaded *.tki* file in Explorer, select **Open results folder** from the **Results** menu (Figure 3-13).

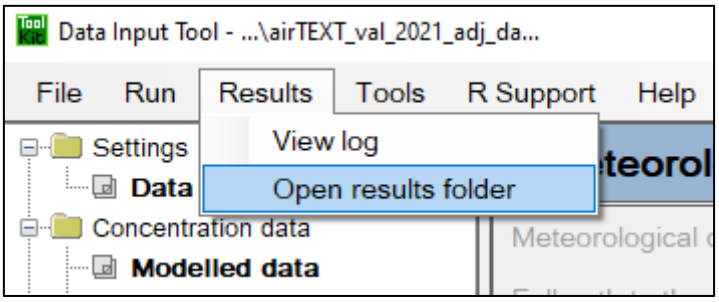


Figure 3-13 Opening the results folder

*Tip: If **Open results folder** is not available to select, this is because the .tki file has not yet been run and therefore no results are available to view.*

Refer to Section 4.1 for details of the output from the Data Input Tool.

3.2 Model Evaluation Tool

The Model Evaluation Tool analyses data from the R workspace file created by the Data Input Tool. Results can be output graphically and numerically. Full details of the output available can be found in Section 4.2.

This section gives a step-by-step guide to using the Model Evaluation Tool.

Step 1: Open the Model Evaluation Tool

Figure 3-14 shows the opening screen of the Model Evaluation Tool interface.

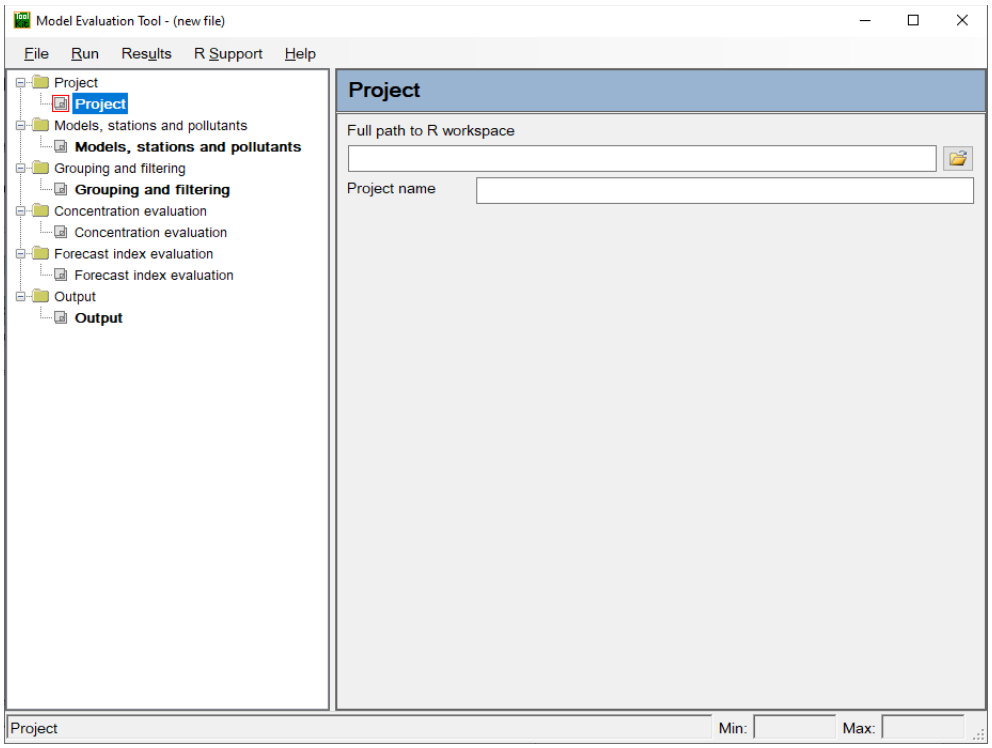


Figure 3-14 The Model Evaluation Tool Project screen

Step 2: Select a workspace

Browse to select a workspace created by the Data Input Tool.

Step 3: Enter a project name

The project name will appear on the titles of graphs. The project name is limited to 20 characters.

Step 4: The Models, stations and pollutants screen

Move to the **Models, stations and pollutants** screen (Figure 3-15) by clicking on the title in the menu on the left. Here you select which data to analyse.

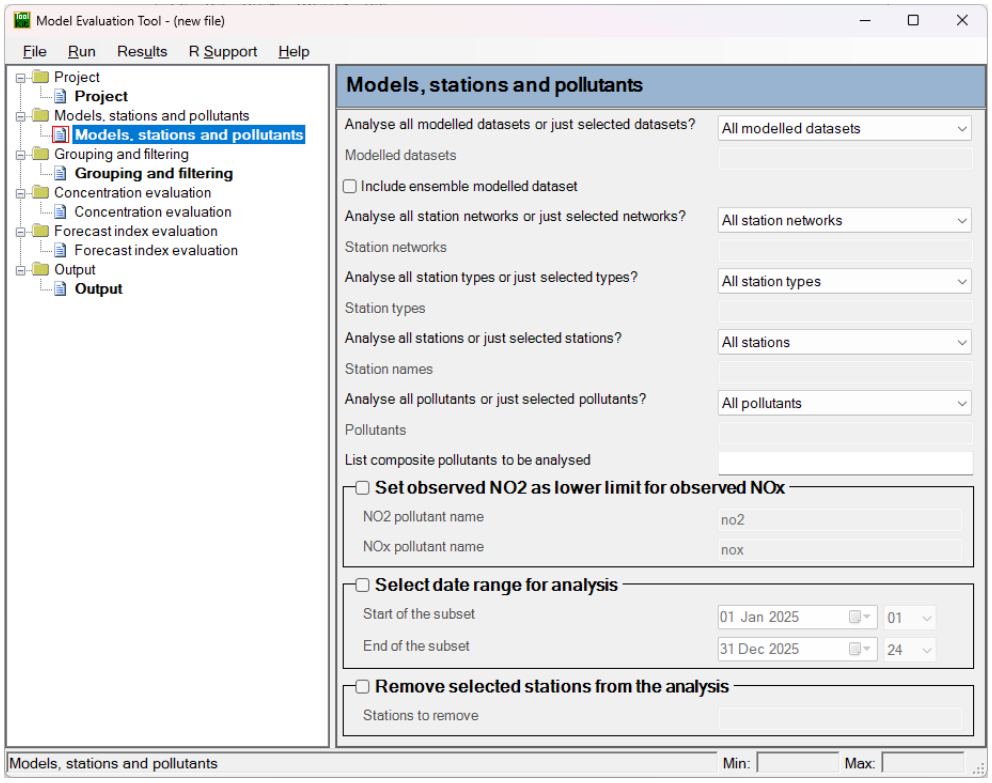


Figure 3-15 The Model Evaluation Tool Models, stations and pollutants screen

Step 5: Select which modelled dataset(s) to analyse

Choose to analyse all modelled datasets or selected model datasets. If you choose to analyse selected modelled datasets, enter either the label for a single dataset or a comma-separated list of labels of the datasets to be analysed, for example “My Model 1, My Model 2, My Model 3”.

Tip: If you don’t know the labels of the modelled datasets that are available, type any text into the box. When you run the Tool it will declare that no data are available for that modelled dataset and give you a list of the available labels.

Step 6: Select whether to include the ensemble modelled dataset

If this option is selected, an additional modelled dataset will be created by collating all the modelled datasets together into one “ensemble”, calculated by taking the median value of all modelled datasets, for each time period, for each pollutant and station.

Step 7: Select which station network(s) to analyse

You can either analyse data from all station networks, or one or more selected station networks. If you choose to analyse data from selected station networks, enter a comma-separated list of networks (for example “AURN, KCL, SAQN”) in the **Station networks** box.

Tip: If you don't know which station networks are available, type any text into the box. When you run the Tool it will declare that no data are available for that station network and give you a list of the available station networks.

Step 8: Select which station type(s) to analyse

You can either analyse data from all station types, or one or more selected station types. If you choose to analyse data from selected station types, enter a comma-separated list of station types (for example "roadside, rural background, urban background") in the **Station types** box.

Tip: If you don't know which station types are available, type any text into the box. When you run the Tool it will declare that no data are available for that station type and give you a list of the available station types.

Step 9: Select which station(s) to analyse

You can either analyse data from all stations, or one or more selected stations. If you choose to analyse data from selected stations, enter a comma-separated list of stations (for example "KCL1, KCL2") in the **Stations** box.

Tip: If you don't know which stations are available, type any text into the box. When you run the Tool it will declare that no data are available for that station and give you a list of the available stations.

Step 10: Select which pollutant(s) to analyse

You can either analyse data for all pollutants, or one or more selected pollutants. If you choose to analyse data for selected pollutants, enter a comma-separated list of pollutants (for example "NO2, NOx, CO2") in the **Pollutants** box.

Tip: If you don't know which pollutants are available, type any text into the box. When you run the Tool it will declare that no data are available for that pollutant and give you a list of the available pollutants.

Step 11: Specify composite pollutants to be analysed

The Toolkit enables the analysis of 'composite' pollutants. These are pollutants for which the concentration can be calculated by combining concentrations of other pollutants for which data are available in your workspace. For example, if NO and NO_2 data were available, it would also be possible to analyse NO_X (= NO + NO_2) as a composite pollutant. Or, if NO_2 and NO_X were available, the ratio of NO_2 to NO_X (= NO_2/ NO_X) could be analysed.

To analyse a composite pollutant:

- Ensure the individual pollutants required are present as output pollutants from the Data Input Tool (available pollutants to analyse in the workspace).

- Enter the equation to calculate the composite pollutant in the **List composite pollutants to be analysed** box using common operators e.g. +(add), -(subtract), /(divide), *(multiply).

- If you wish to analyse multiple composite pollutants, list each equation separated by a comma for example “NO+NO2, NO2/NOx”

Step 12: Set observed NO2 as the lower limit for observed NOx

As NO_X concentrations include NO_2, the NO_2 concentration should never be greater than the NO_X concentration. However, in observed data this is not always the case, and in some instances where different measuring methods are used, NO_2 values may be greater than NO_X values.

If the **Set observed NO2 as the lower limit for observed NOx** is selected, the concentration of NO_X will be set equal to the concentration of NO_2 for observed data for any instances where the observed NO_2 is greater than the observed NO_X.

If you select this option, you must then enter the pollutant names corresponding to NO_2 and NO_X in your workspace (for most cases this is expected to be just “NO2” and “NOx”).

Step 13: Select date range for analysis

By default, data for the whole time period covered by the modelled data will be analysed. To analyse data for a shorter period, select this option and select the start and end date and time.

Step 14: Remove selected stations from the analysis

You may wish to remove certain stations from the analysis, if, for example, you are aware of measurement problems there. If so, select this option, and in the text box specify stations to remove as a comma-separated list (for example “MY1, BCH1”).

Step 15: The Grouping and filtering screen

Move to the **Grouping and filtering** screen (Figure 3-16).

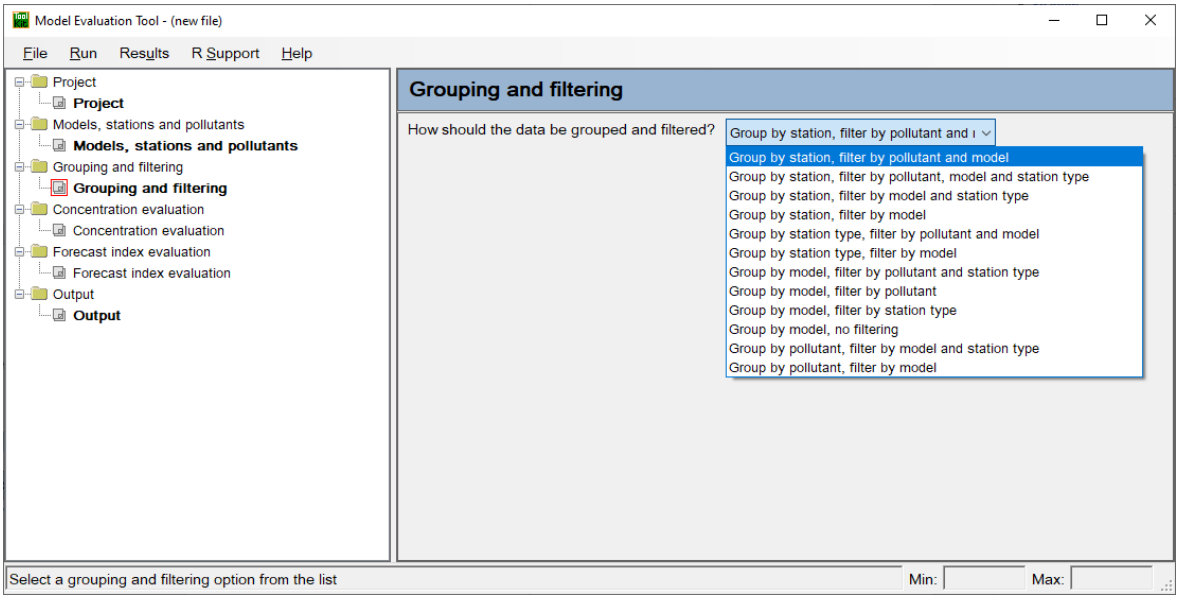


Figure 3-16 The Model Evaluation Tool Grouping and filtering screen

Here you can select how the data should be 'grouped' and 'filtered' in the analysis.

'Grouping' describes what each data point on a plot represents. For example, if the data are grouped by station, each data point on the plot represents a mean value over all the stations included in the analysis.

'Filtering' describes how the data is to be split across plots; for example, if the option to filter by station type is selected then data for different station types will be distinguished in the plots, either through creating separate plots for each station type, or using separate colours and symbols for each station type on the same plot.

The data can be grouped and filtered by:

- Station
- Station type
- Pollutant
- "Model" (i.e. modelled dataset)

The default option is to group by station and filter by pollutant and model. Data must always be either grouped by or filtered by model because the datasets must always be evaluated separately against the observed data.

Step 16: The Concentration evaluation screen

Move to the **Concentration evaluation** screen (Figure 3-17). Select this option to produce output comparing the modelled concentrations to observed concentrations. On this screen you are able to select from a range of output plot types, and specify what data will be shown in the plots.

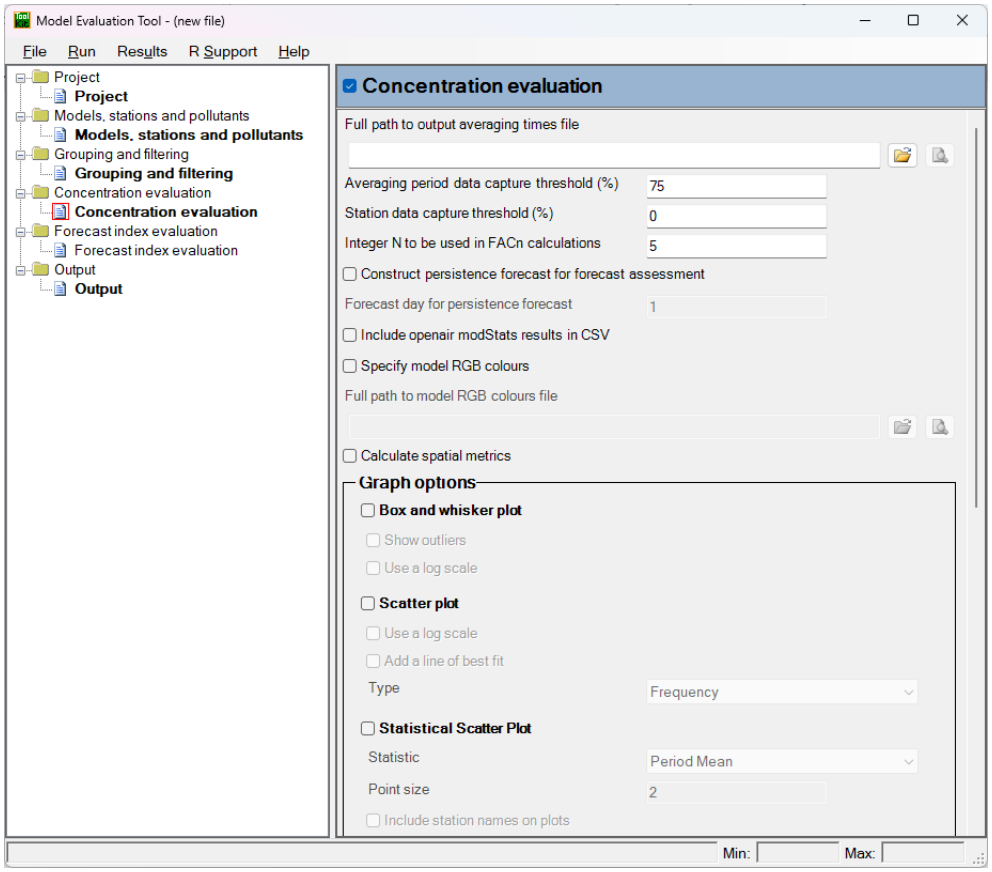


Figure 3-17 The Model Evaluation Tool "Concentration evaluation" screen

Step 17: Specify the output averaging times file

If you are performing the concentration evaluation, you must supply an output averaging times file. The output averaging times file is a comma-separated file containing one row per pollutant. For each pollutant, a number of parameters must be set, as described in Table 3-7. An example output averaging times file, *output_averaging_hourly.csv*, is included in the *\Data* folder and is shown in Figure 3-18.

	A	B	C	D
1	pollutant	output.avg.time	output.statistic	output.daily.max
2	no2	1 hour	mean	no
3	o3	1 hour	mean	no
4	pm10	1 hour	mean	no
5	pm2.5	1 hour	mean	no
6				

Figure 3-18 Example output averaging times file, *output_averaging_hourly.csv*

Column header	Description	Allowed values
pollutant	Name of the pollutant as defined in the pollutant data file used in the Data Input Tool	*n/a*
output.avg.time	Averaging time to be applied to the concentration output, in hours, days, months or years (minimum 1 hour)	Integer value and unit (e.g. '1 day', '8 hour') or just unit (e.g. 'hour', 'year')
output.statistic	Statistic that applies to the concentration output	'max', 'mean' or 'rolling mean'
output.daily.max	Should output be based on the daily maximum of the calculated values?	'yes' or 'no'

Table 3-7 Data included in the output averaging times file

Some additional columns in the output averaging times file are optional, but required for certain statistical metrics to be calculated. These are described below in Table 3-8.

Column header	Description	Allowed values
ae.band	Concentration value (AE_band_conc) used in the calculation of AE_band. If not supplied then AE_band will not be calculated.	Any numerical value. Assumed to be in the concentration units given in the pollutant definitions file used in the Data Input Tool.
ct	Concentration threshold value C_t used in the calculation of FMS. This is required if the option to calculate spatial metrics is selected on the Concentration Evaluation screen.	Any numerical value. Assumed to be in the concentration units given in the pollutant definitions file used in the Data Input Tool.
gamma.fn	Weighting factor γ_{FN} applied to AFN in the calculation of 1DMOE. This is required if the option to calculate spatial metrics is selected on the Concentration Evaluation screen.	Any numerical value.
gamma.fp	Weighting factor γ_{FP} applied to AFP in the calculation of 1DMOE. This is required if the option to calculate spatial metrics is selected on the Concentration Evaluation screen.	Any numerical value.
percentile	Percentile to calculate and plot if the Statistical Scatter Plot is selected with the 'Period percentile' option.	Any numerical value between 0 and 100

Table 3-8 Optional columns in the output averaging times file

Browse to select your output averaging times file.

Step 18: Specify the averaging period data capture threshold

Enter a data capture threshold to apply to the output averaging process. For example, for an 8-hour rolling mean, a 75% data capture threshold means that at least 6 hours of data must be valid in each period for the averaged data for this period to be considered valid and used in the evaluation.

Step 19: Specify the station data capture threshold

Enter a data capture threshold to apply to stations for the analysis. The data capture of a station must exceed this threshold in order for the station to be included in the analysis. For example, if the station data capture threshold is 50, any station for which there are valid observed/modelled data pairs for less than 50% of the analysis time period will be excluded from the analysis.

Step 20: Specify the integer N to be used in FACn calculations

Specify the integer N to be used to calculate the fraction of model predictions within a factor of N of observations. This is set to N = 5 as default and requires a minimum of N = 3. Note that FACn is output in addition to the standard FAC2 metric.

Step 21: Specify whether to construct a persistence forecast for forecast assessment

Select this option to construct a persistence forecast, which is a forecast that uses the observed concentrations from one day as the forecast concentrations for a number of days in the future. For example, if this option is selected, and the **Forecast day for persistence forecast** is set to 1 (the default) then the observations from one day will be used as the persistence forecast for the next day. If the **Forecast day for persistence forecast** is 2, then the observations from one day will be used as the persistence forecast for the day after the next day.

If this option is selected, a number of additional statistical metrics are calculated which benchmark modelled concentrations against the persistence forecast; these are described briefly in Table 4-7.

Step 22: Specify whether to output openair modStats results

If this option is selected, then a number of further additional columns of data are added to the output CSV file; these are described briefly in Table 4-2. For more information about the statistics generated by the modStats function, please refer to openair documentation [3].

Step 23: Select whether to specify model RGB colours

Selecting this option allows you to specify the colours used to represent each dataset (observed and modelled) on box and whisker plots when filtered by model, and on time variation plots when grouped by model.

To use this option, an additional input file is required, in which the colours to be used are specified. This file, referred to as the 'model RGB colours file', is in CSV format. An example, *model_rgb.csv*, can be found in the *\Data* folder and is shown in Figure 3-19. The file contains a row for the observed dataset and for each modelled dataset, in which the RGB components of the colours to be used are specified (in the range 0 – 255). If this example file were used, the observed data would be represented in black.

	A	B	C	D	E
1	model	red	green	blue	
2	obs	0	0	0	
3	Model1	146	206	81	
4	Model2	81	97	40	
5	Model3	15	57	65	
6					

Figure 3-19 Example model RGB colours file, *model_rgb.csv*

If you have selected the option to specify the colours, browse to select the model RGB colours file.

Step 24: Specify whether to calculate spatial metrics

Select this option to additionally calculate a range of statistical metrics based on the areas in which models and/or observations indicate exceedance of a particular threshold value. These are described in Table 4-3.

Step 25: Select graph options

The Model Evaluation Tool can produce a wide range of graphical output, as listed below. Select the types of plot you wish to output. More information about the different types of graph can be found in Section 4.2.1.

a. Box and whisker plot

- Select **Show outliers** if you would like 'outliers' to be plotted on the graph.
- Select **Use a log scale** if you would like the plot axis to be on a logarithmic scale.

b. Scatter plot

- Select **Use a log scale** if you would like the plot axis to be on a logarithmic scale.
- Select **Add a line of best fit** if you would like a line of best fit and the corresponding equation to be plotted on the graph.

- Select the type of scatter plot you wish to output – a frequency scatter plot or conventional scatter plot. A frequency scatter plot shows the frequency of occurrence of each data point whereas a conventional scatter plot shows one point per pairwise modelled-observed data point.

  Tip: The frequency scatter plot is better suited to larger sets of data, the conventional scatter plot to smaller sets

c. Statistical Scatter plot

- Select the statistic to be displayed. Note that the AOT40 Crops[2], AOT40 Forests[3] and SOMO35[4] statistics are only suitable for an ozone pollutant, and Period Percentile requires an additional 'Percentile' column in the output averaging times file.

- Enter your chosen size for the points on the scatter plot

- Select **Include station names on plots** if you would like each point on the plot to be labelled with the corresponding station name.

- Select **Filter station type by colour** if the grouping and filtering option includes filter by station type and you wish the station types to be distinguished by colour.

- Select **Filter station type by symbol** if the grouping and filtering option includes filter by station type and you wish the station types to be distinguished by symbol.

- Select **Filter station type by graph** if the grouping and filtering option includes filter by station type and you wish a separate plot for each station type.

- Select **Plot linear reference lines given by A*x** if you wish to include linear reference lines on the plot (e.g. y=x). For each reference line, the value of A and the line type should be entered, with the settings for each line separated by a comma. The available line types are solid, dashed and dotted. If this option is selected the default lines plotted are y=x (solid line), y=0.5x (dashed line) and y=2x (dashed line).

d. Time Variation Analysis plot

- Select the type of time variation analysis. You should consider the time period covered by your data – for example, if your data doesn't cover a complete year, it may not be appropriate to analyse the data by the month of the year.

- Select **Show individual modelled data sets on separate plots** to produce a separate plot for each set of modelled data. If this option is not selected, data for all modelled datasets will be displayed on a single graph, using a

[2] AOT40 Crops - Accumulated dose of ozone Over a Threshold of 40 ppb, across the crop growing season of May-July. Source: https://www.emep.int/mscw/definitions.pdf

[3] AOT40 Forests – Accumulated dose of ozone Over a Threshold of 40 ppb, across the forests growing season of April-September. Source: https://www.emep.int/mscw/definitions.pdf

[4] SOMO35 – Sum of Ozone Means Over 35 ppb. Source: https://www.emep.int/mscw/definitions.pdf

separate line for each modelled dataset. It is recommended that this option is selected if your workspace contains multiple modelled datasets.

e. Quantile-Quantile plot

- Select **Use a log scale** if you would like the plot axis to be on a logarithmic scale.

f. Time plot averaged over all stations

g. Target plot

Note: The target plot can only be produced for the following statistics:

- *Daily maximum 8-hour rolling mean O_3*
- *Hourly mean NO_2*
- *Daily mean PM_{10} and $PM_{2.5}$*

- Select whether to include station names on the plot
- Select how the filtering should be represented. For example, if filtering by station type, a plot can be produced for each station type, or all the station types can be represented using different colours or symbols on the same plot.
- Select whether station types should be distinguished when not filtering by station type. If this is selected, then when not filtering by station type, the station types will be shown in different colours and/or symbols.
- Provide the file path to the target plot uncertainty values files.

  The uncertainty values file is a comma-separated CSV file containing one row of parameters per pollutant. For each pollutant, the coefficients of the measurement uncertainty given by the FAIRMODE DELTA methodology [1] must be specified. Table 3-9 describes these parameters. An example file, *pol_measurement_uncertainties.csv*, can be found in the *Data* folder, and is shown in Figure 3-20.

	A	B	C	D	E	F	G	H	I	J	K	L
1	pollutant	Beta	UrLV	LV	alpha	Np	Nnp	target.units	target.avg.time	target.statistic	target.daily.max	
2	NO2	2	0.24	200	0.2	5.2	5.5	ug/m3	1 hour	mean	no	
3	O3	2	0.18	120	0.79	11	3	ug/m3	8 hour	rolling mean	yes	
4	PM10	2	0.28	50	0.25	20	1.5	ug/m3	day	mean	no	
5	PM2.5	2	0.36	25	0.5	20	1.5	ug/m3	day	mean	no	
6												

Figure 3-20 Example pollutant measurement uncertainty values file *pol_measurement_uncertainties.csv*

Column header	Description	Allowed values
pollutant	Name of the pollutant as defined in the pollutant data file used in the Data Input Tool	*n/a*
Beta	Coverage factor, *k*.	Refer to the DELTA methodology [1]
UrLV	Measurement uncertainty, u_r^{LV}	Refer to the DELTA methodology [1]
LV	Limit value (or reference value)	Refer to the DELTA methodology [1]
alpha	Proportion of the measurement uncertainty that is independent of the limit value.	Refer to the DELTA methodology [1]
Np, Nnp	Coefficients used for annual averages to account for the compensation of errors due to random noise, periodic re-calibration of the instruments etc	Refer to the DELTA methodology [1]
target.units	Units that apply to the measurement uncertainty	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
target.avg.time	Averaging time that applies to the measurement uncertainty, which should be equal to that applied to the concentration evaluation output (minimum 1 hour)	Integer value and unit (e.g. '1 day', '8 hour') or just unit (e.g. 'hour', 'year')
target.statistic	Statistic that applies to the measurement uncertainty, which should be equal to the concentration evaluation output	'max', 'mean' or 'rolling mean'
target.daily.max	Is the target plot based on the daily maximum of the calculated statistics?	'yes' or 'no'

Table 3-9 Data included in the uncertainty values file

Note that some of the plots can only be produced for particular grouping and filtering options. Table 3-10 presents the available paired options for grouping and filtering for each plot.

Plot	Allowed Group	Allowed Filter
Box and whisker plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
	Station type	Model and Pollutant
	Model	Pollutant and Station type
Scatter plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
Statistical scatter plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
Time variation analysis plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
	Station type	Model and Pollutant
	Model	Pollutant
	Model	Pollutant and Station type
Quantile-quantile plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
Time plot	Station	Model and Pollutant
Target plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
Forecast index plots	Station	Model
	Station	Model and Pollutant
	Station	Model and Station type
	Station	Model and Pollutant and Station type
	Station type	Model
	Station type	Model and Pollutant
	Pollutant	Model
	Pollutant	Model and Station type
	Model	-
	Model	Pollutant
	Model	Station type
	Model	Pollutant and Station type

Table 3-10 Grouping and filtering options available for each plot type output from the Model Evaluation Tool

Note: If you select a plot that cannot be produced, a message will be given in the log file indicating why the plot has not been produced, for example a lack of data, or incompatibility with the grouping and filtering option selected.

Step 26: The Forecast index evaluation screen

Move to the **Forecast index evaluation** screen (Figure 3-21). This option converts both the observed and modelled concentrations to 'forecast indices' and performs an evaluation based on these forecast indices. The user specifies the concentration levels to be represented by each forecast index. If you do not wish to use this option, skip to Step 33:. Otherwise, select the option and continue with the steps below.

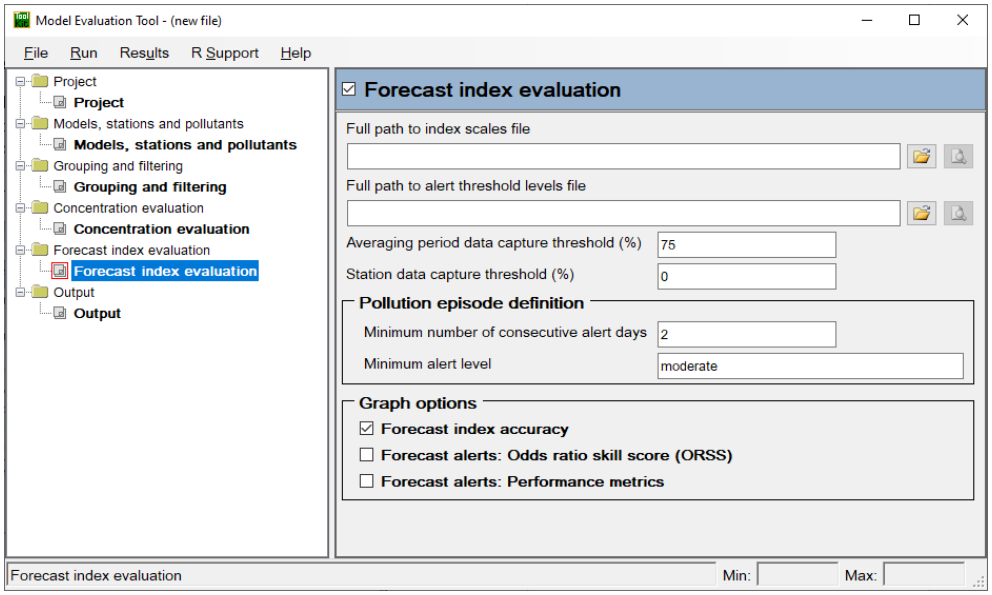


Figure 3-21 The Model Evaluation Tool Forecast index evaluation screen

Step 27: Supply the path of the index scales file

The index scales file is a comma-separated CSV file containing one row per pollutant. An example, *index_scales.csv*, can be found in the *Data* folder and is shown in Figure 3-22. For each pollutant, a number of index-related parameters must be set. Table 3-11 describes these parameters. At least one index threshold concentration must be defined.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	pollutant	index.units	index.avg	index.statistic	index.daily.max	i1	i2	i3	i4	i5	i6	i7	i8	i9	i10
2	no2	ug/m3	1 hour	mean	yes	0	68	135	201	268	335	401	468	535	601
3	o3	ug/m3	8 hour	rolling mean	yes	0	34	67	101	121	141	161	188	214	241
4	pm10	ug/m3	day	mean	yes	0	17	34	51	59	67	76	84	92	101
5	pm2.5	ug/m3	day	mean	yes	0	12	24	36	42	48	54	59	65	71
6															

Figure 3-22 Example index scales file, *index_scales.csv*

Column header	Description	Allowed values
pollutant	Name of the pollutant as defined in the pollutant data file used in the Data Input Tool	*n/a*
index.units	Units in which the index threshold concentrations are given	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
index.avg.time	Averaging time that applies to the index threshold concentrations (minimum 1 hour)	Integer value and unit (e.g. '1 day', '8 hour') or just unit (e.g. 'hour', 'year')
index.statistic	Statistic that applies to the index threshold concentrations	'max', 'mean' or 'rolling mean'
index.daily.max	Is the forecast given as the daily maximum of the calculated indices?	'yes' or 'no'
i1	Threshold concentration for index level 1	Any numeric value
i2	Threshold concentration for index level 2	Any numeric value
i{n}	Threshold concentration for index level n	Any numeric value

Table 3-11 Data included in the index scales file

Step 28: Supply the path of the alert threshold levels file

The alert threshold levels file is a comma-separated CSV file containing one row per alert. Each alert should have a name and an alert threshold defined in terms of the indices in the index scales file. Table 3-12 describes the parameters in the file. An example alert thresholds file, *alert_thresholds.csv*, can be found in the *\Data* folder, and is show in Figure 3-23.

	A	B
1	index	name
2	4	moderate
3	7	high

Figure 3-23 Example alert thresholds file, *alert_thresholds.csv*

Column header	Description	Allowed values
index	Threshold index for this alert	Any integer value
name	The name to give to this alert in all output	*n/a*

Table 3-12 Data included in the alert threshold levels file

Step 29: Specify the averaging period data capture threshold

Enter a data capture threshold to apply to the output averaging process. For example, for an 8-hour rolling mean, a 75% data capture threshold means that at least 6 hours of data must be valid in each period for the averaged data for this period to be considered valid and used in the analysis. The threshold will also be applied to each day if index.daily.max is set to "yes" in the index scales file.

Step 30: Specify the station data capture threshold

Enter a data capture threshold to apply to stations for the analysis. The data capture of a station must exceed this threshold in order for the station to be included in the analysis. For example, if the station data capture threshold is 50%, any station for which there are valid observed/modelled data pairs for less than 50% of the analysis time period will be excluded from the analysis.

Step 31: Specify the pollution episode definition

A pollution episode is defined as a particular alert level being reached on a particular number of consecutive days. Specify the alert level and number of days you wish to use.

Step 32: Select the graph output options

There are three graph options:

- Forecast index accuracy
- Forecast alerts: Odds ratio skill score (ORSS)
- Forecast alerts: Performance metrics

For more information about this output, please see Section 4.2.2.

Select the graph output you require.

Step 33: The Output screen

Move to the Output screen (Figure 3-24). Here you can select the type of output files to be produced.

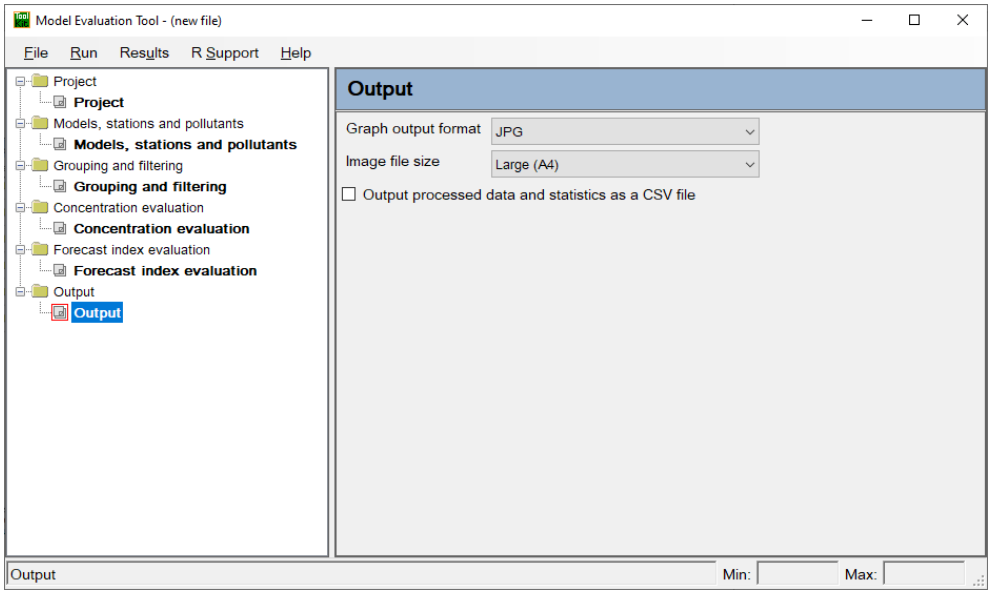


Figure 3-24 The Model Evaluation Tool Output screen

Step 34: Select the graph output format and image file size

There are three graph output options: JPG, PNG and PDF. The first two options produce image files that can be imported into other documents; one image file is produced for each graph. The PDF option produces one PDF file (size A4) per graph type.

The image file size is used for JPG and PNG output. There are three options: Large (A4), Medium and Small. Reducing the image size will increase the proportional size of any text in the image file. These options provide flexibility to produce graphs for reports or presentations.

Step 35: Choose whether to output the processed data and statistics as a CSV file

This is a very useful option that produces a CSV file containing all of the numerical data used to create the graphs, as well as some statistics not shown on the graphs. Refer to Section 4.2.1.8 for details of the contents of this file.

Step 36: Save the settings file and run the Model Evaluation Tool

To save the settings file, select **Save** or **Save As...** from the **File** menu. Model Evaluation Tool settings files are saved with the file extension *.tkm*. To run the Tool, select **Model Evaluation Tool** from the **Run** menu.

Step 37: Check the log file

The Model Evaluation Tool generates a log file when it runs; this records all the processing messages that are shown in the run screen as well as any error and warning messages issued. It is good practice to check the log file for any problems that may have occurred.

To see the log file, select **View log** from the **Results** menu. This option will open the log file in your preferred text editor (Notepad by default).

Step 38: View results folder

To open the results folder for the currently loaded *.tkm* file in Explorer, select **Open results folder** from the **Results** menu.

Refer to Section 4.2 for details of the output from the Model Evaluation Tool.

3.3 Model Diagnostics Tool

The Model Diagnostics Tool enables more detailed investigation of the performance of a model. Each plot produced by the Tool contains data for one pollutant from one station.

Step 1: Open the Model Diagnostics Tool

Figure 3-25 shows the opening screen of the Model Diagnostics Tool

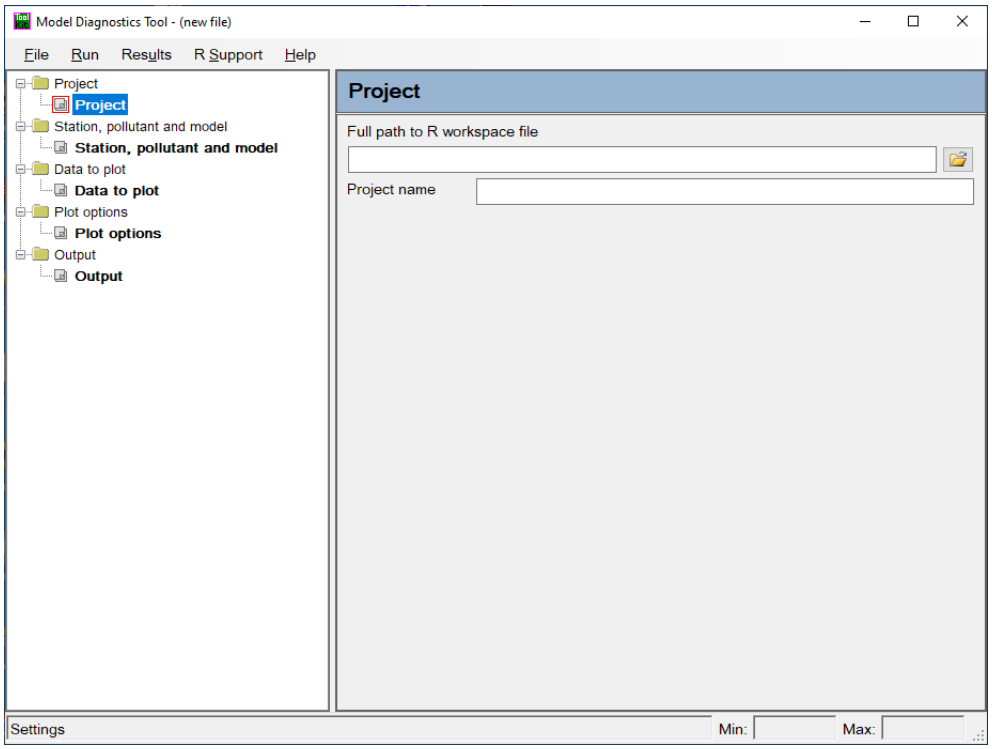


Figure 3-25 The Model Diagnostics Tool Project screen

Step 2: Select the R workspace file

Browse to the select the R workspace file (created by the Data Input Tool) containing the data to be analysed.

Step 3: Enter a project name

The project name should identify the project or data being analysed. The name will appear in the title in graphs output from the Tool.

Step 4: **The Station, pollutant and model screen**

Move to the **Station, pollutant and model** screen (Figure 3-26). Here you can select the combination of station, pollutant and model for which to produce output, and select a date range for analysis.

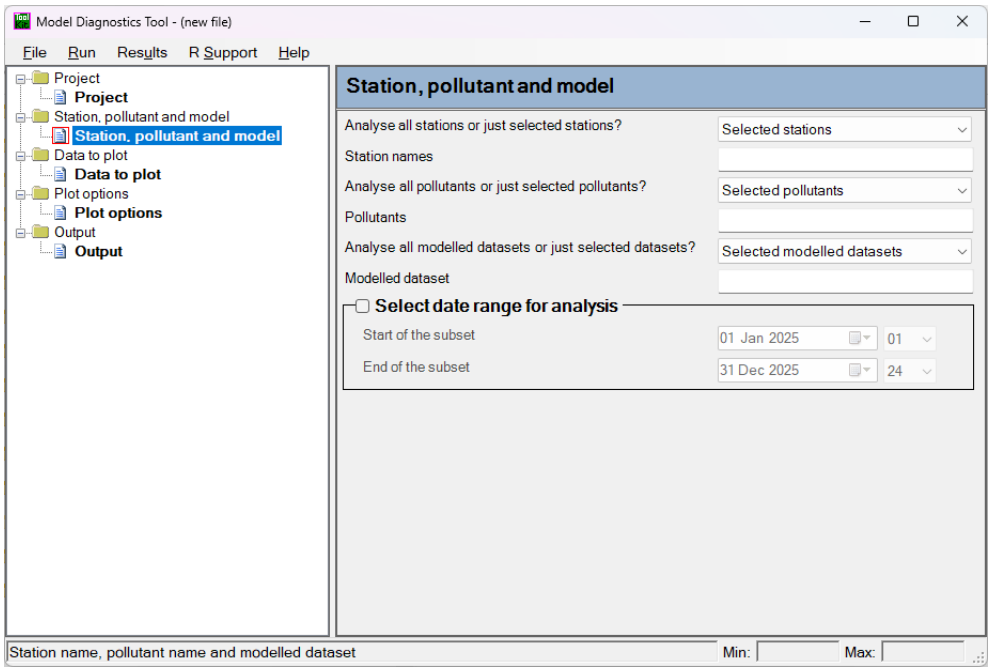


Figure 3-26 The Model Diagnostics Tool Station, pollutant and model screen

Step 5: **Select which stations to analyse**

You can select whether to analyse all stations or just selected stations. If analysing selected stations, enter the names of the stations to be analysed, separated by a comma (for example “MY1, BTH”).

Step 6: **Select which pollutants to analyse**

Select whether to analyse all pollutants or selected pollutants. If analysing selected pollutants, enter the names of the pollutants separated by a comma (for example “NO2, NOx, CO2”).

Step 7: **Select which modelled datasets to analyse**

Select whether to analyse all modelled datasets or selected modelled datasets. If analysing selected modelled datasets, enter the names of the modelled datasets separated by a comma (for example “Model1, Model2”).

Step 8: **Select the date range for analysis**

Select this option if you wish to analyse data for a subset of the time period covered by the workspace. Then select the start and end date and time of the period you wish to analyse.

Step 9: **The Data to plot screen**

Move to the **Data to plot** screen (Figure 3-27).

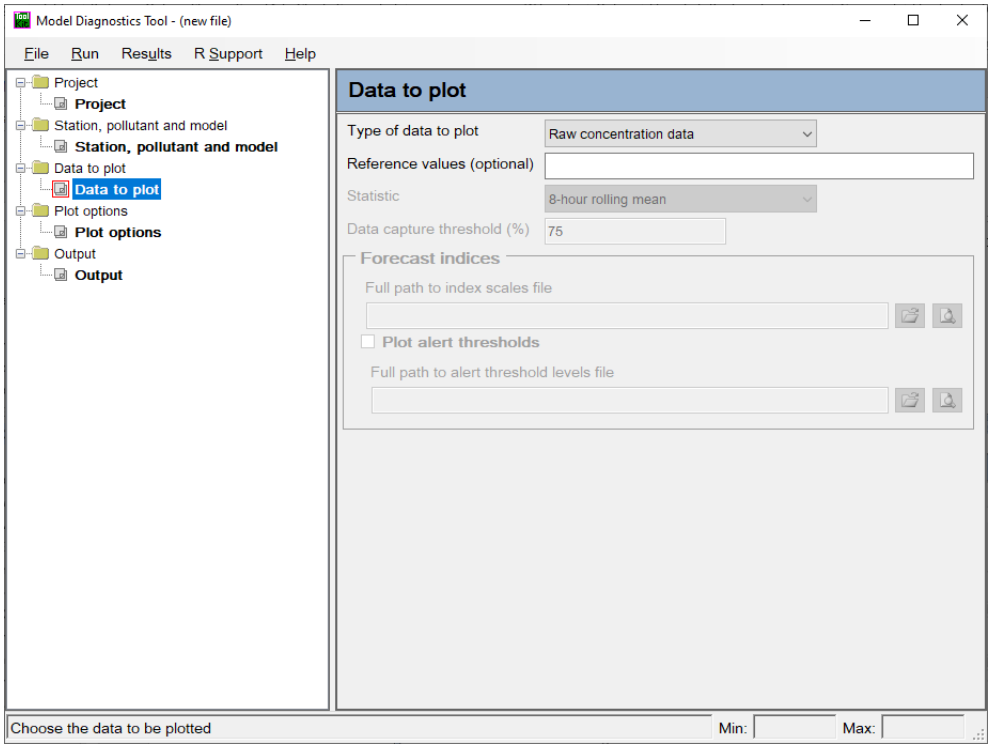


Figure 3-27 The Model Diagnostics Tool Data to plot screen

Step 10: **Select the type of data to plot**

Select the type of data to plot. There are three options:

1. Plot raw concentration data

   This option plots the data as input, using the averaging time of the modelled data.

2. Plot calculated statistics

   This option produces plots of one statistic. Select the statistic from the drop-down list in the **Statistic** box.

3. Plot calculated forecast indices

   This option converts both the modelled and observed data to indices (as defined in the index scales file described in Section 3.2, and detailed in Table 3-11) and plots the index values.

Step 11: **Select whether to include reference values**

If the option to plot **Raw concentration data** or **Calculated statistics** is selected, then up to ten reference values may also be plotted on the graph. To enter more than one reference value, separate the values with a comma, for example “50, 100”.

Step 12: **Enter data capture threshold**

If the option to plot **Calculated statistics** or **Forecast indices** is selected, specify a data capture threshold to apply to the averaging process. For example, for an 8-hour rolling mean, a 75% data capture threshold means that at least 6 hours of data must be valid in each period for the averaged data for that period to be valid and used in the graph.

Step 13: **Select whether to plot alert threshold levels**

If the option to plot **Forecast indices** is selected, there is the option to plot alert thresholds levels on the graphs as reference lines. In order to use this option, you will need to specify an alert threshold levels file (described in Section 3.2).

Step 14: **The Plot options screen**

Move to the **Plot options** screen (Figure 3-28). Here, you can select which types of plot to output.

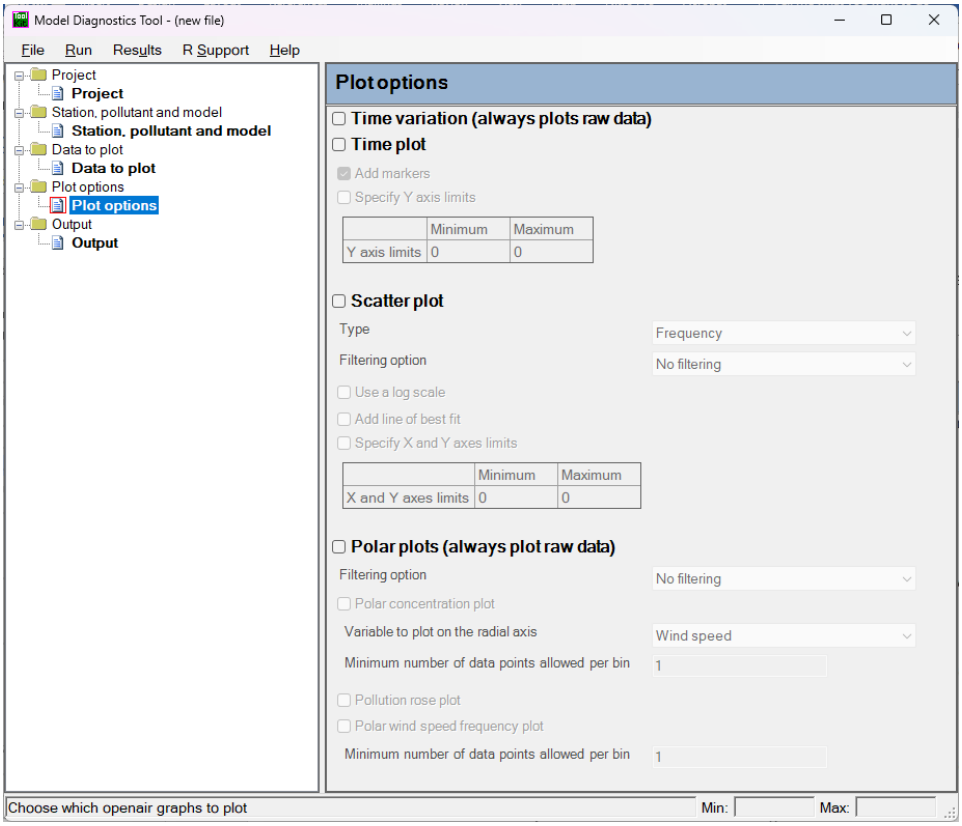


Figure 3-28 The Model Diagnostics Tool Plot options screen

Four types of plot can be produced (for full details, please see Section 4.3):

1. Time variation

   Averages by hour of the day, day of the week and month are plotted. The settings on the **Data to plot** screen do not apply to this graph; the raw data are always plotted.

2. Time plot

   This is a time series plot of both modelled and observed data.

   - By default, markers are included on the plot to represent each data point. Un-select **Add markers** to remove markers from plots.

   - By default, Y axis limits are determined by the value range in the data to be plotted. Select **Specify Y axis limits** to over-ride this and set the minimum and maximum Y axis limits. Note that these limits will be applied to all stations, pollutants and models plotted, so this option is best suited to situations where a single pollutant has been selected.

3. Scatter plot

   This option plots modelled against observed data.

   - Select which type of scatter plot you require; a frequency scatter plot shows the frequency of occurrence of each data point whereas a conventional scatter plot shows one point per pairwise modelled-observed data point.

     If you have selected the option to plot forecast indices, the type of scatter plot produced will always be a frequency plot.

     Tip: The frequency scatter plot is better suited to larger sets of data, the conventional scatter plot to smaller sets.

   - Select a filtering option if required. The data can be filtered by weekday, month or season.

   - Select whether to plot the data on a logarithmic scale.

   - Select whether to add a line of best fit and the corresponding equation to the scatter plot.

   - By default, X and Y axes limits are determined by the value range in the data to be plotted. Select **Specify X and Y axes limits** to over-ride this and set the minimum and maximum X and Y axes limits. Note that these limits will be applied to all stations, pollutants and models plotted, so this option is best suited to situations where a single pollutant has been selected.

4. Polar plots

   The polar plots illustrate the variation in concentration with wind direction. The settings on the **Data to plot** screen do not apply to this graph; the raw data are always plotted. To produce polar plots, meteorological data is

required. This is input via the Data Input Tool, and can be supplied in a number of formats – please see Section 3.1 for details.

- Select a filtering option if required. The data can be filtered by weekday, month or season.
- Select which type(s) of polar plot to output:
  - Polar concentration plot. If selected, the following formatting options must be specified:
    - Variable to plot on the radial axis. The data can be banded according to the wind speed, hour of the day or month of the year.
    - Minimum number of data points per band
  - Pollution rose plot – this shows, for each wind direction, the frequency that the concentration is in a particular range.
  - Polar wind speed frequency plot – this shows wind speed-direction frequencies. If selected, the minimum number of data points per band must be specified.

Step 15: The Output screen

Move to the **Output** screen (Figure 3-29). Here you can select the output file format.

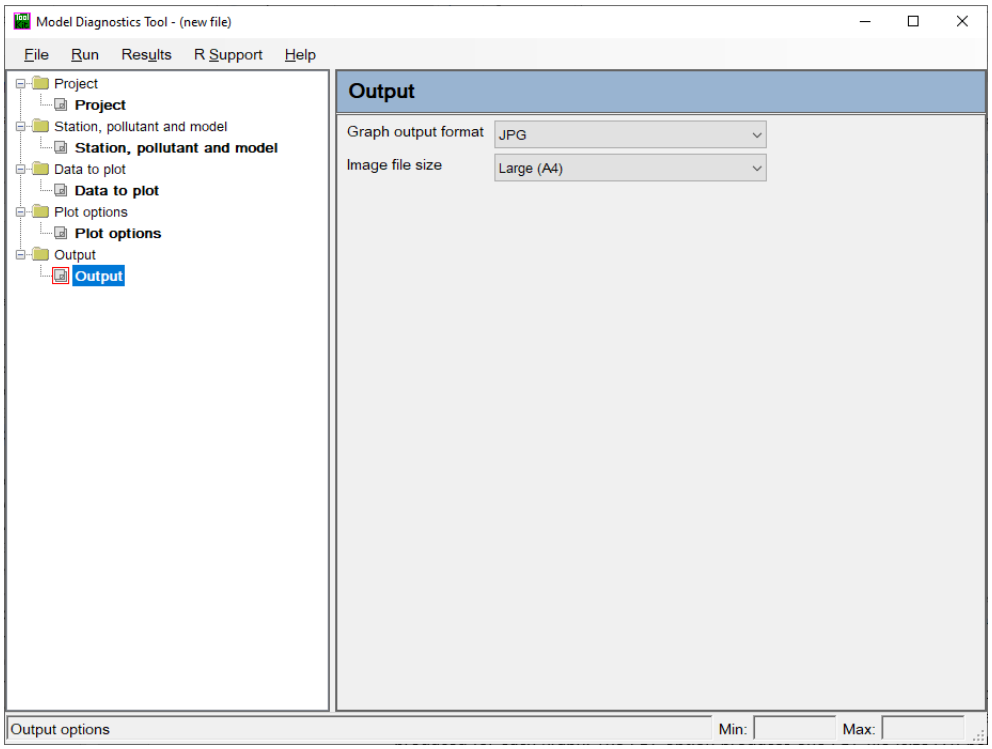


Figure 3-29 The Model Diagnostics Tool Output screen

There are three graph output options: JPG, PNG and PDF. The first two options produce image files that can be imported into other documents; one image file is

produced for each graph. The PDF option produces one PDF file (size A4) per graph type.

The image file size is used for JPG and PNG output. There are three options: Large (A4), Medium and Small. Reducing the image size will increase the proportional size of any text in the image file. These options provide flexibility to produce graphs for reports or presentations.

Step 16: Save the settings file and run the Model Diagnostics Tool

To save the settings file, select **Save** or **Save As...** from the **File** menu. Model Diagnostics Tool settings files are saved with the file extension *.tkd*. To run the Tool, select **Model Diagnostics Tool** from the **Run** menu.

Step 17: Check the log file

The Model Diagnostics Tool generates a log file when it runs; this records all the processing messages that are shown in the run screen as well as any error and warning messages issued. It is good practice to check the log file for any problems that may have occurred.

To see the log file, select **View log** from the **Results** menu. This option will open the log file in your preferred text editor (Notepad by default).

Step 18: View results folder

To open the results folder for the currently loaded *.tkd* file in Explorer, select **Open results folder** from the **Results** menu. Refer to Section 4.3 for details of the output from the Model Diagnostics Tool.

4 Output

All three Tools generate output files that are written to the same folder as the input *.tki*, *.tkm* or *.tkd* file, with the name of the input file as a prefix, for easy identification. In addition to Tool-specific output, all three Tools produce a *.log* file that records messages written to the output screen, including any error or warning messages.

4.1 Data Input Tool Output

4.1.1 R Workspace

This is an R workspace containing all the data imported by the Data Input Tool, processed and ready for import into the Model Evaluation and Model Diagnostic Tools.

This workspace can also be loaded into R to explore the data further.

4.1.2 CSV file

This contains the time series of concentrations (in output units as defined in the pollutants information file) for every station and pollutant for which both monitored and modelled data are available, with the averaging time and statistic specified in the pollutant definitions file.

4.2 Model Evaluation Tool Output

4.2.1 Concentration Evaluation Output

4.2.1.1 Box and Whisker plot

The Box and Whisker plot shows the following data for each station, for each of the modelled and observed datasets:

1. The lower whisker
2. The 25th percentile (the lower quartile – the lower end of the box)
3. The 50th percentile (the median – the horizontal line inside the box)
4. The 75th percentile (the upper quartile – the upper end of the box)
5. The upper whisker

The inter-quartile range (IQR) is defined as the 75th percentile minus the 25th percentile, i.e. the length of the box. The lower whisker is defined as the lowest concentration value still within 1.5xIQR of the lower quartile. The upper whisker is defined as the highest concentration value still within 1.5xIQR of the upper quartile. Optionally, the outliers lying outside of the upper and lower whisker can also be plotted, and the plot can be displayed on a $\log_{10}$ scale. An example box and whisker plot is shown in Figure 4-1.

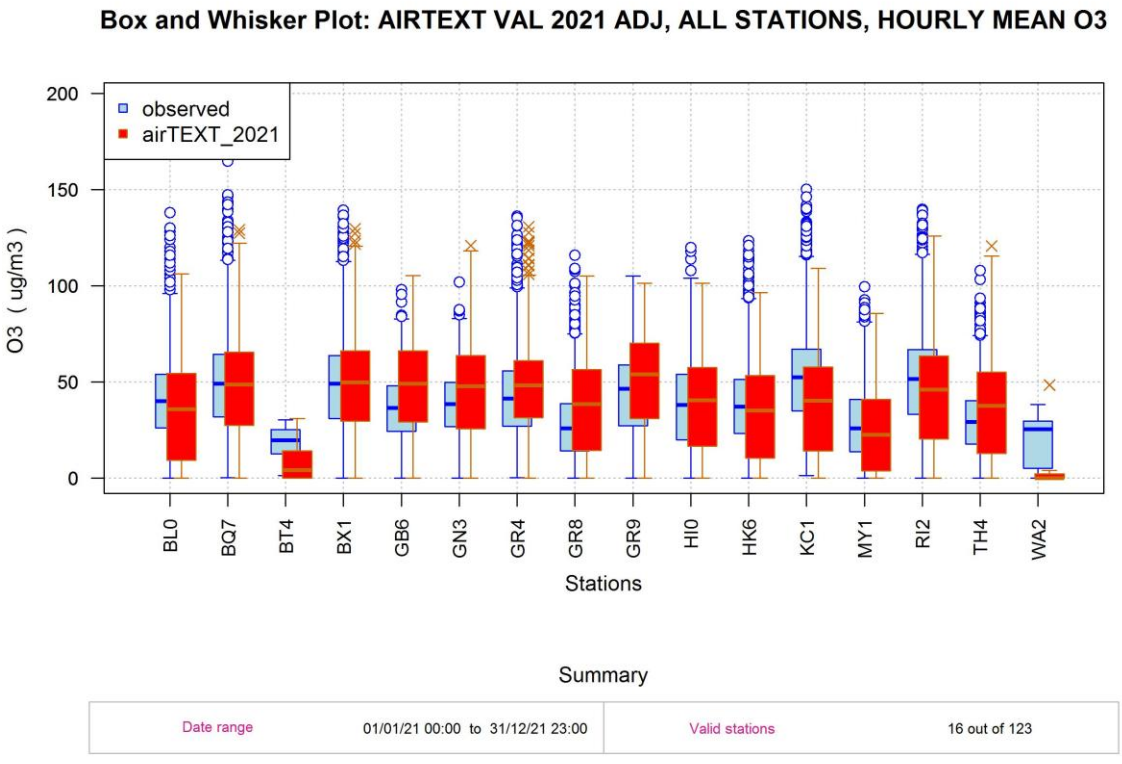


Figure 4-1 Example box and whisker plot output from Model Evaluation Tool

4.2.1.2 Scatter Plot

This plot compares the modelled and observed concentrations on a scatter plot. The frequency scatter plot shows the frequency of occurrence of each point whereas the conventional scatter plot shows one point per pairwise modelled-observed data point. The data can optionally be plotted on a $\log_{10}$ scale. The line of exact correspondence between observed and modelled concentration is shown as a solid black line. Points within the area bounded by the two dotted lines indicate that the modelled concentration and the observed concentration are within a factor of two. Figure 4-2 presents an example frequency scatter plot and Figure 4-3 presents an example conventional scatter plot.

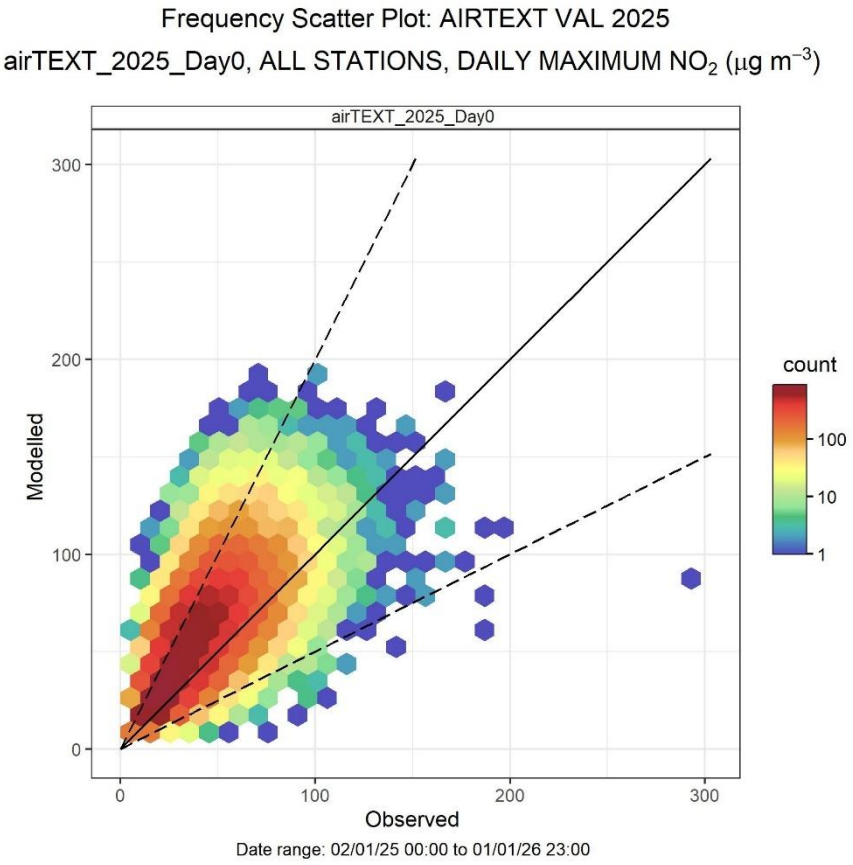


Figure 4-2 Example frequency scatter plot output from the Model Evaluation Tool

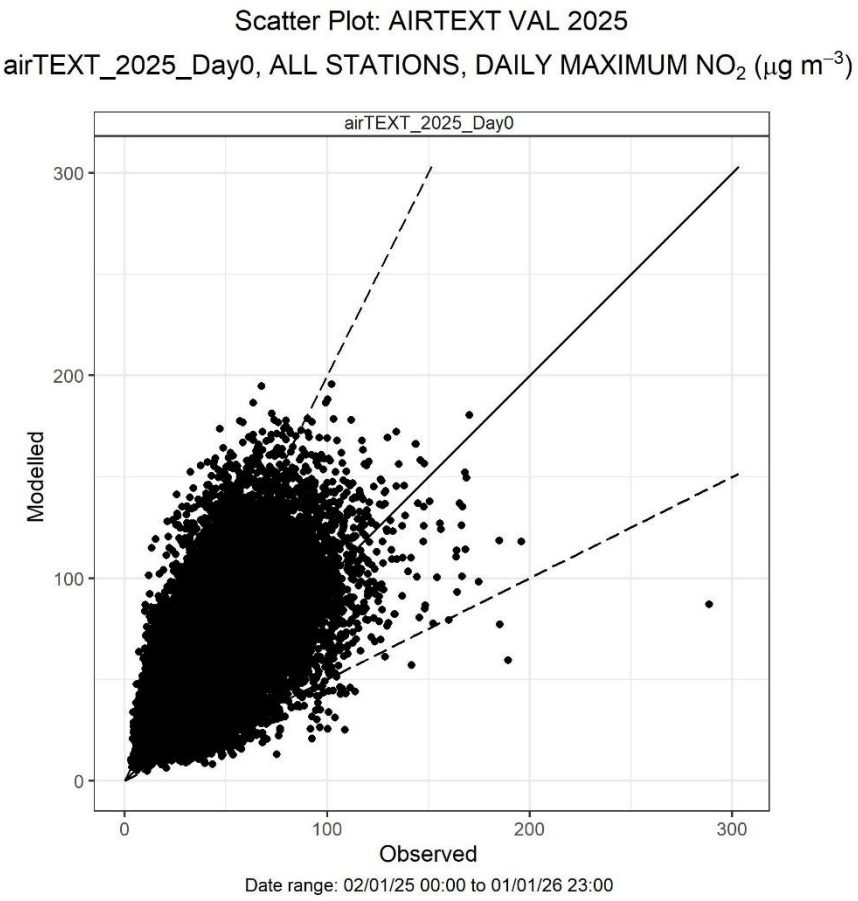


Figure 4-3 Example conventional scatter plot output from the Model Evaluation Tool

4.2.1.3 *Statistical Scatter Plot*

The statistical scatter plot compares calculated statistics of modelled and observed data. The statistics are calculated from the whole period of data being analysed. The plot shows a point per station, with the calculated statistic for the modelled data on the y-axis and the calculated statistic for the observed data on the x-axis. An example is shown in Figure 4-4.

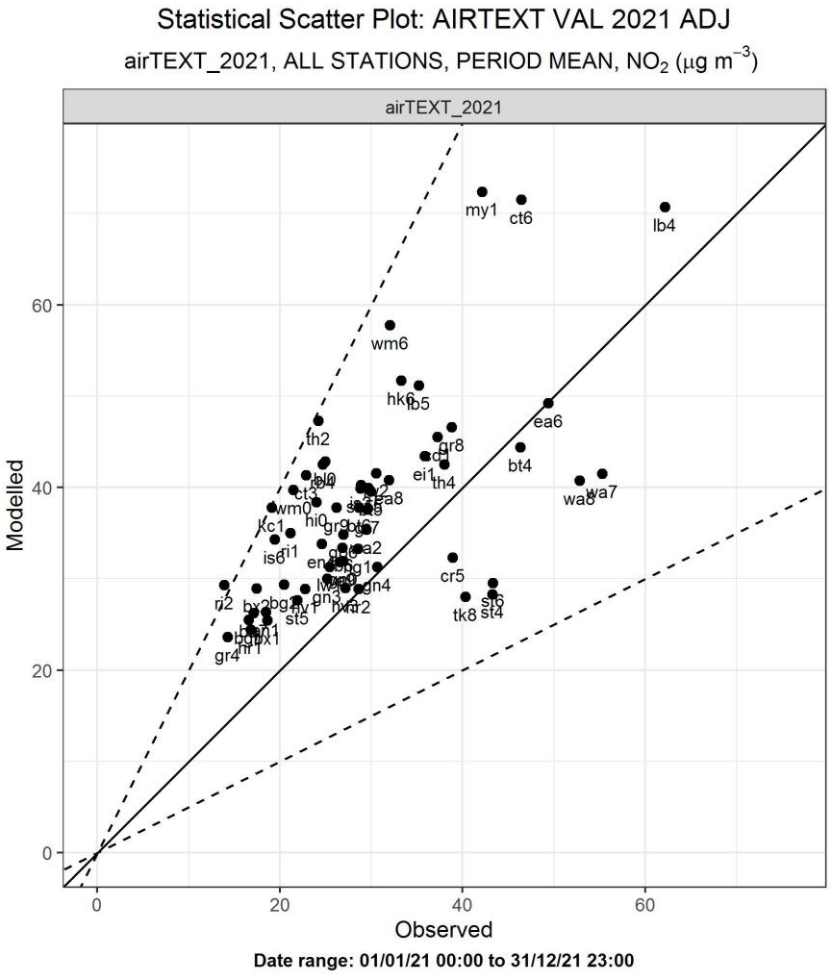


Figure 4-4 Example statistical scatter plot output from the Model Evaluation Tool

4.2.1.4 *Time Variation Analysis Plot*

The time variation analysis plot illustrates how the mean concentration varies by hour of the day, day of the week and month of the year. The shaded areas indicate a 95% confidence interval. An example is shown in Figure 4-5.

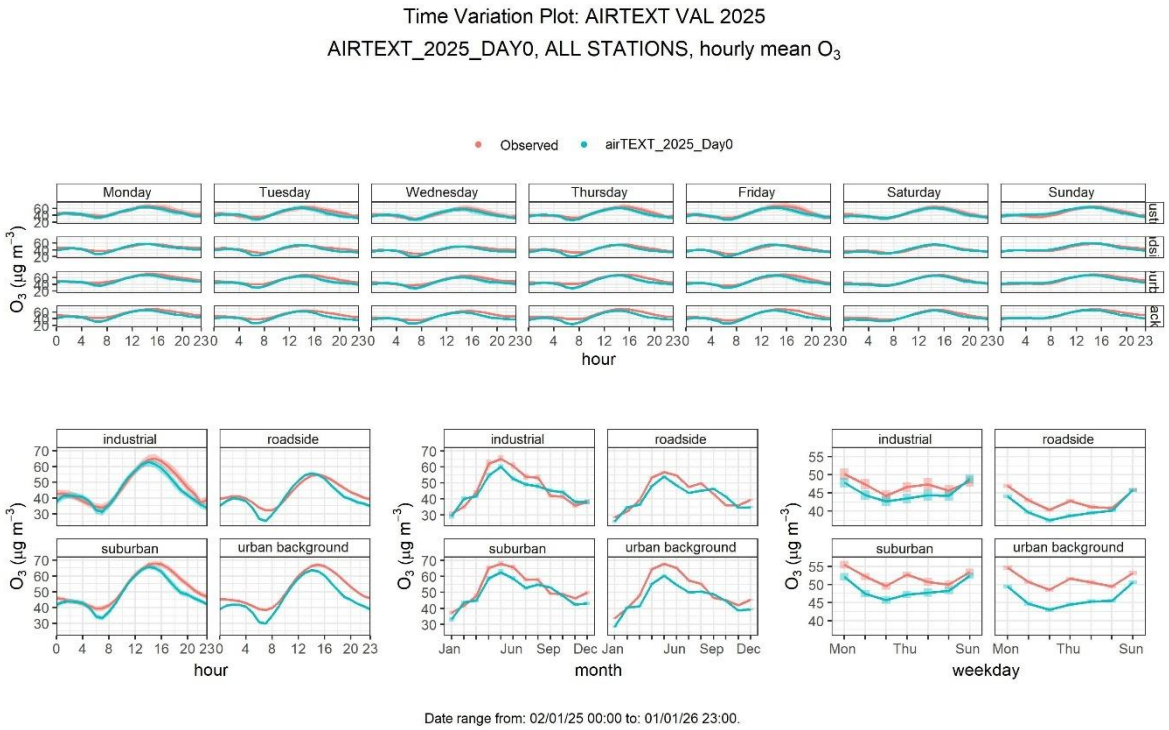


Figure 4-5: Example time variation plot output from the Model Evaluation Tool

4.2.1.5 *Quantile-Quantile Plot*

This plot compares the modelled and observed concentrations ordered independently from lowest to highest concentration, as a quantile-quantile plot. The data can optionally be plotted on $\log_{10}$ scales. The line of exact correspondence between observed and modelled concentration is shown as a solid black line. Points within the area bounded by the two dotted lines indicate that the modelled concentration and the observed concentration are within a factor of two. Figure 4-6 presents an example quantile-quantile plot.

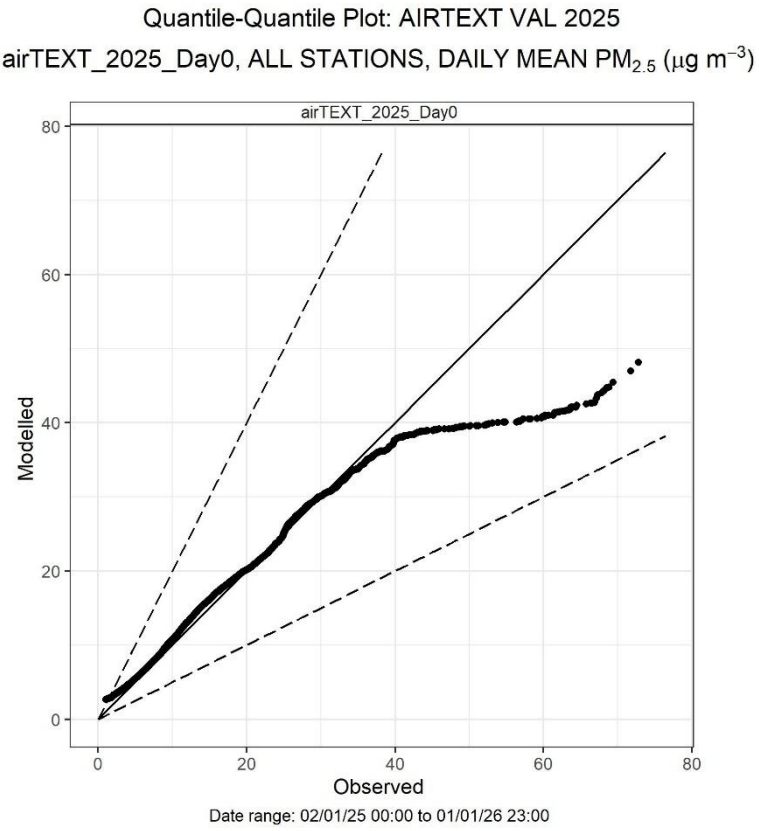


Figure 4-6 Example quantile-quantile plot output from the Model Evaluation Tool

4.2.1.6 *Time plot averaged over all stations*

The time plot compares the modelled to the observed concentrations for each pollutant as time series plots, averaged over all stations. If a forecast assessment is also being carried out, then the plot includes alert thresholds. Figure 4-7 shows an example time plot.

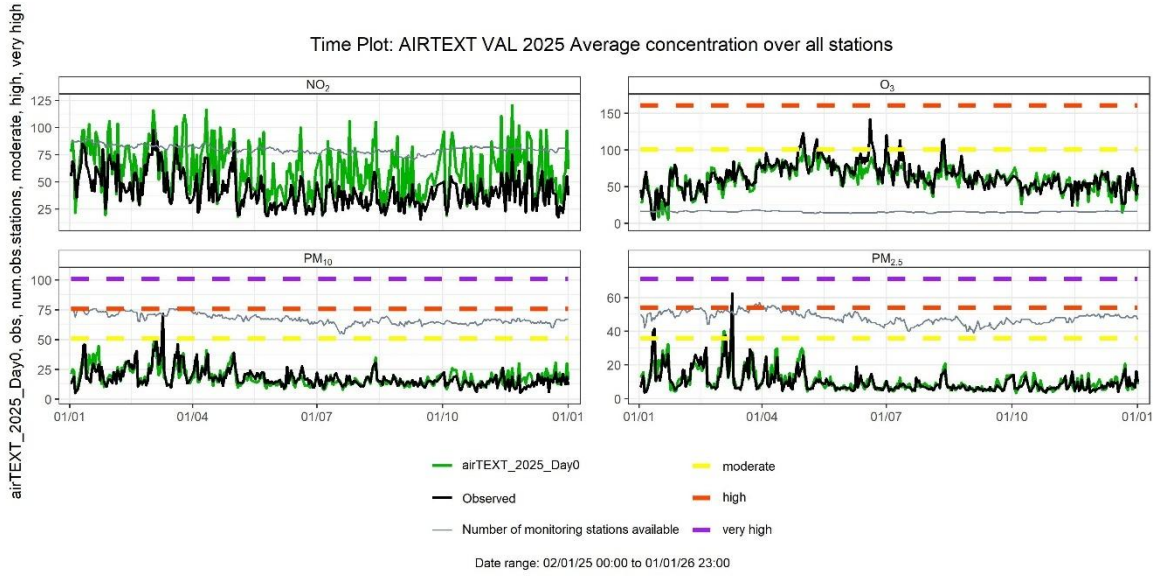


Figure 4-7 Example time plot output from the Model Evaluation Tool

4.2.1.7 *Target Plot*

The target plots produced by the Model Evaluation Tool are consistent with those produced by version 7.2 of the FAIRMODE DELTA Tool. An example is shown in Figure 4-8.

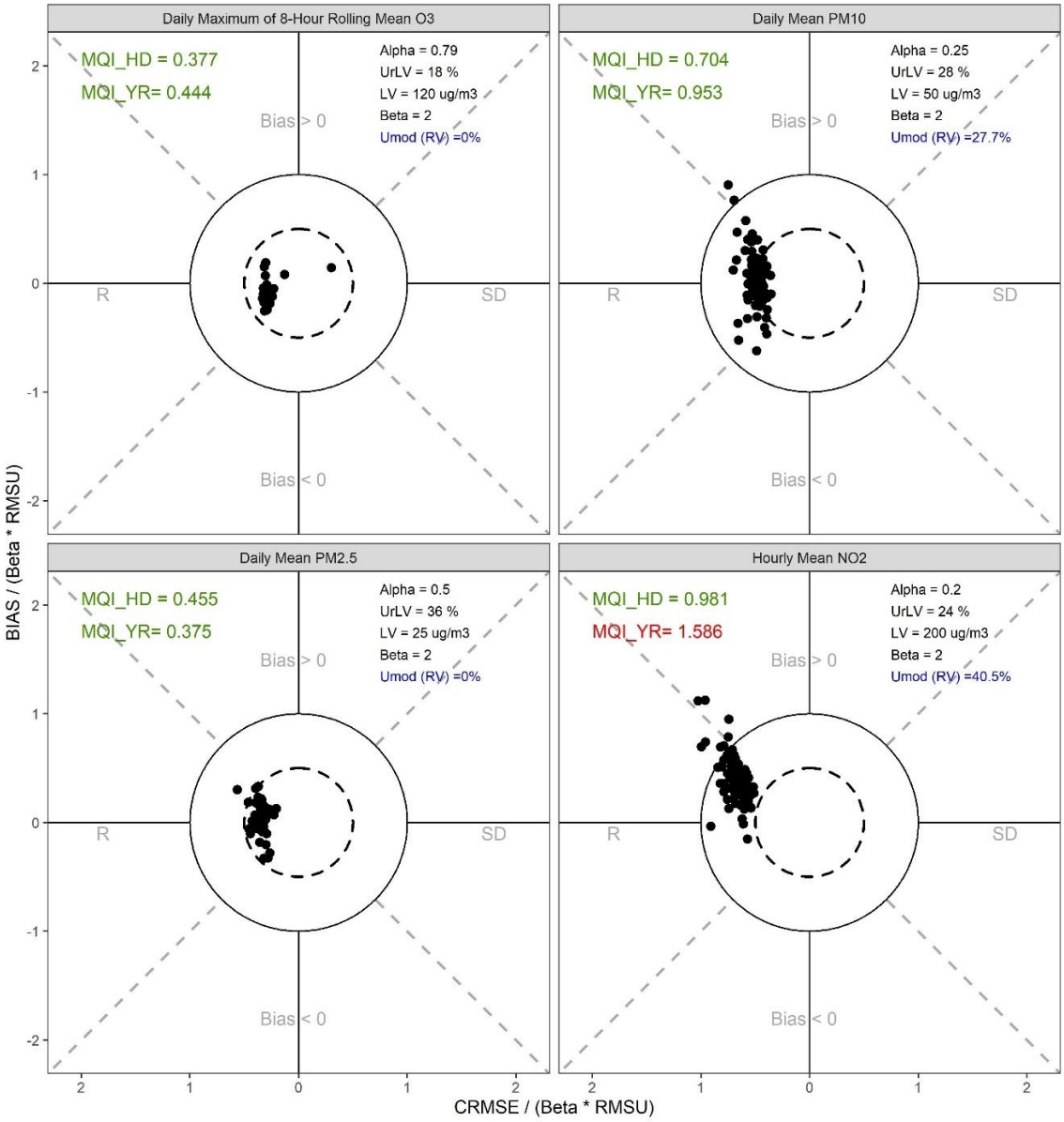


Figure 4-8 Example target plot output from the Model Evaluation Tool, for four pollutants and one modelled dataset.

The metrics shown on the target plot for each monitoring station and pollutant are listed below. In the equations, C represents concentration, and σ the standard deviation; observed values are indicated by the subscript 'O' and modelled values by the subscript 'P'.

Centralised root-mean-square error (CRMSE):

$$\mathrm{CRMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left((C_{Pi} - \overline{C_P}) - (C_{Oi} - \overline{C_O})\right)^2}$$

$\overline{C_P}$ is the mean modelled concentration, $\overline{C_O}$ is the mean observed concentration.

Mean bias (BIAS):

$$\text{BIAS} = \frac{1}{N}\sum_{i=1}^{N}(C_{Pi} - C_{Oi})$$

Root-mean-square of the expanded measurement uncertainty (RMSU):

$$\text{RMSU} = U_r(LV)\sqrt{(1-\alpha^2)\left(\overline{C_O}^2 + {\sigma_O}^2\right) + \alpha^2 LV^2}$$

where σ_o^2 is the standard deviation of the observed data, *LV* is the limit value of interest and the other coefficients ($U_r(LV)$ and α) are derived for a specific pollutant from measurement data, as described in the FAIRMODE DELTA methodology [1] and associated papers [6,7].

The target plot shows $\frac{\text{BIAS}}{\beta\text{RMSU}}$ against $\frac{\text{CRMSE}}{\beta\text{RMSU}}$, where β=2. However, because CRMSE is always positive, only the right-hand side of the diagram would be needed. The negative X axis section is therefore used to provide additional information, indicating whether the error is dominated by correlation error or standard deviation error. This information is obtained using the relationship between CRMSE, correlation, R, and standard deviation σ (see Table 4-1 for definitions of R and σ): $CRMSE^2 = {\sigma_O}^2 + {\sigma_P}^2 - 2\sigma_O\sigma_P R$

The sign of CRMSE when plotted on the target plot is determined by the ratio of the value of CRMSE assuming a perfect correlation (R=1) to the value of CRMSE assuming a perfect standard deviation ($\sigma_O = \sigma_P$):

$$\left(\frac{CRMSE\ (R=1)}{CRMSE(\sigma_O = \sigma_P)}\right)^2 = \frac{(\sigma_O - \sigma_P)^2}{2\sigma_O\sigma_P(1-R)}$$

If the ratio is greater than 1 then standard deviation error dominates and the point is plotted on the positive part of the X axis. If the ratio is less than 1 then correlation error dominates and the point is plotted on the negative part of the X axis.

The **Modelling Quality Indicator (MQI)** is defined as a combination of BIAS and CRMSE:

$$MQI = \frac{\text{RMSE}}{\beta \times RMSU} = \sqrt{\left(\frac{BIAS}{\beta RMSU}\right)^2 + \left(\frac{CRMSE}{\beta RMSU}\right)^2}$$

This represents the distance from the origin on the Target Plot. The outer solid black circle represents MQI=1; the inner dotted black circle represents MQI=0.5. Models meet the **Modelling Quality Objective (MQO)** if 90% of points have MQI less than 1, i.e. they lie inside the outer circle. The 90th percentile MQI value is given on the plot as 'MQI_HD' and coloured green if less than 1; red otherwise.

The **Modelling Quality Indicator for annual means ($\text{MQI}_{\text{annual}}$)** is based solely on the difference between the modelled and observed annual means, as a proportion of the measurement uncertainty for annual means, UOmean**:**

$$UOmean = U_r(LV)\sqrt{\frac{(1-\alpha^2)}{N_p}\overline{C_O}^2 + \frac{\alpha^2 \times LV^2}{N_{np}}}$$

$$MQI_{annual} = \frac{|\overline{C_O} - \overline{C_P}|}{\beta \times UOmean}$$

Models meet the **Modelling Quality Objective for annual means** (**MQO_{annual}**) if 90% of points have MQI_{annual} less than 1. The 90th percentile MQI_{annual} value is given on the plot as 'MQI_YR' and coloured green if less than 1; red otherwise.

The target plot includes a representation of **Model Uncertainty (**$RMSUM_{perc}$**)**, expressed in relation to the measurement uncertainty and the relevant limit value.

$$RMSUM = RMSU\sqrt{\left(\frac{RMSE}{RMSU}\right)^2 - 1}$$

$$RMSUM_{perc} = \frac{RMSUM}{RMSU} \times LV \times 100$$

If the root-mean square error (RMSE) is less than the measurement uncertainty (RMSU) then $RMSUM_{perc}$ cannot be calculated and is given as 0%, with the rationale that (by this measure, at least) no meaningful improvement of the model can be made. The 90th percentile value of $RMSUM_{perc}$ is given on the plot as 'Umod(RV)' and coloured blue. The origins of this metric are explained in [1].

The equivalent values for annual means are $RMSUM_{annual}$ and $RMSUM_{annual\_perc}$.

$$RMSUM_{annual} = UOmean\sqrt{\left(\frac{\overline{C_O} - \overline{C_P}}{UOmean}\right)^2 - 1}$$

$$RMSUM_{annual\_perc} = \frac{RMSUM_{annual}}{UOmean} \times LV \times 100$$

A number of FAIRMODE **Modelling Performance Indicators** are also calculated for each station and each pollutant.

The following three statistical metrics are temporal metrics calculated using hourly values:

- The **Modelling Performance Indicator – Bias (**MPI_{Bias}**)**:

$$MPI_{Bias} = \frac{|\overline{C_O} - \overline{C_P}|}{\beta \times RMSU}$$

  The Modelling Performance Criterion – Bias (MPC_{Bias}) is the 90th percentile of all valid values of the MPI_{Bias}.

- The **Modelling Performance Indicator – correlation (**MPI_R**)**:

$$MPI_R = \frac{1 - R}{0.5\beta^2 \frac{RMSU^2}{\sigma_O \sigma_P}}$$

  The Modelling Performance Criterion – Correlation (MPC_R) is the 90th percentile of all valid values of the MPI_R.

- The **Modelling Performance Indicator – standard deviation (**MPI_{SD}**)**:

$$MPI_{SD} = \frac{|\sigma_P - \sigma_O|}{\beta \times RMSU}$$

  The Modelling Performance Criterion – Correlation (MPC_{SD}) is the 90th percentile of all valid values of the MPI_{SD}.

The following two statistical metrics are spatial metrics calculated using annual values:

- The **Spatial Modelling Performance Indicator – correlation (Spatial_MPI$_R$)**:

$$Spatial\_MPI_R = \frac{1 - R}{0.5\beta^2 \frac{UOmean^2}{\sigma_O \sigma_P}}$$

- The **Spatial Modelling Performance Indicator – standard deviation (Spatial_MPI$_{SD}$)**:

$$Spatial\_MPI_{SD} = \frac{|\sigma_P - \sigma_O|}{\beta \times UOmean}$$

4.2.1.8 CSV output files

If the option to **Output processed data and statistics as CSV** is selected, statistics are calculated for each pollutant (in output units) and are output in a CSV file, named **_conc_stats.csv*. The statistics are calculated for the variable by which the data is grouped, and additionally over all data, filtered by the filters selected. For example, if data is grouped by station and filtered by model and pollutant, the statistics are output for each station, and for all valid stations for each model and pollutant.

A number of statistics are always included in this file; others are only included if certain options are selected. The statistics that are always included are listed in Table 4-1. Where 'obs' or 'mod' occur in variable names, these indicate observed or modelled values respectively. In the equations, C represents concentration, and σ the standard deviation; observed values are indicated by the subscript 'o' and modelled values by the subscript 'p'.

Name	Description	Equation
num.valid.values	Number of values	
obs.mean mod.mean	Mean	$1/n \sum C$
SDO SDM	Standard Deviation of observed and modelled data	$\sqrt{1/n \sum (C - \overline{C})^2}$
MB	Mean Bias	$\overline{(C_p - C_o)}$
R	Pearson's Correlation Coefficient	$\mathrm{cov}(C_p, C_o)/\sigma_{C_p}\sigma_{C_o}$
R2	Square of Pearson's Correlation Coefficient	R^2
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	$(\sigma_{C_p} - \sigma_{C_o})/0.5(\sigma_{C_o} + \sigma_{C_p})$
RMSE	Root Mean-Square-Error	$\sqrt{\frac{1}{n}\sum (C_p - C_o)^2}$
NMSD	Normalised Mean Standard Deviation	$(\sigma_{C_p} - \sigma_{C_o})/\sigma_{C_o}$
ME	Mean Error	$\overline{\lvert C_p - C_o \rvert}$
MSE	Mean Square Error	$\overline{(C_o - C_p)^2}$ which can also be expressed as $MSEs + MSE.cor + mMSE$
MSEs	Systematic component of MSE	$(\overline{C_p} - \overline{C_o})^2$
MSEu	Unsystematic component of MSE	$MSE - MSEs = MSE.cor + mMSE$
MSE.cor	Variance component of MSE	$(\sigma_{C_p} - R\sigma_{C_o})^2$
mMSE	Minimum achievable MSE	${\sigma_{C_o}}^2(1 - R^2)$
obs.max mod.max	Maximum	$\max C$
obs.RHC mod.RHC	Robust Highest Concentration	$\chi(n) + (\chi - \chi(n))\ln\left(\frac{3n-1}{2}\right)$, where n is the number of values used to characterise the upper end of the concentration distribution, χ 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.AOT40.forests mod.AOT40.forests	Accumulated dose of ozone over a threshold of 40 ppb, across the forests growing season of April-September	

Name	Description	Equation
obs.AOT40.crops mod.AOT40.crops	Accumulated dose of ozone over a threshold of 40 ppb, across the crop growing season of May-July	
obs.SOMO35 mod.SOMO35	Sum of ozone means over 35 ppb	
obs.lower.whisker mod.lower.whisker	Lower whisker	
obs.lower.quartile mod.lower.quartile	Lower quartile	
obs.median mod.median	Median	
obs.upper.quartile mod.upper.quartile	Upper quartile	
obs.upper.whisker mod.upper.whisker	Upper whisker	
CRMSE	Centred Root Mean Square Error	$\sqrt{\overline{[(C_p - \overline{C_p}) - (C_o - \overline{C_o})]^2}}$
B_NMBF	Normalised Mean Bias Factor	$\frac{\overline{C_p}}{\overline{C_o}} - 1$ if $\overline{C_p} \geq \overline{C_o}$ $1 - \frac{\overline{C_o}}{\overline{C_p}}$ if $\overline{C_p} < \overline{C_o}$
MAE	Mean Absolute Error	$\lvert\overline{C_o} - \overline{C_p}\rvert$
NAE	Normalised Absolute Error	$\frac{\overline{\lvert C_o - C_p\rvert}}{0.5(\overline{C_o} + \overline{C_p})}$
MFE	Mean Fractional Error	$\overline{\left(\frac{\lvert C_p - C_o\rvert}{0.5(C_p + C_o)}\right)}$
NRMSE	Normalised Root Mean Square Error	$\sqrt{\frac{\overline{(C_o - C_p)^2}}{\overline{C_o}\ \overline{C_p}}}$
NCRMSE	Normalised Centred Root Mean Square Error	$\frac{\sqrt{\overline{[(C_p - \overline{C_p}) - (C_o - \overline{C_o})]^2}}}{\sigma_{C_o}}$
tau	Peak Time Error	$time\left(C_{p}{}^{U}_{peak}\right) - time\left(C_{0_{peak}}\right)$ where $C_{o_{peak}}$ is the peak observed value $C_{p}{}^{U}_{peak}$ is the unpaired peak predicted value
epsilon	Peak Magnitude Error	$C_{p}{}^{U}_{peak} - C_{o_{peak}}$ where $C_{o_{peak}}$ is the peak observed value $C_{p}{}^{U}_{peak}$ is the unpaired peak predicted value

Name	Description	Equation
NMSE	Normalised Mean-Square-Error	$\overline{(C_p - C_o)^2} / \overline{C_o C_p}$ which can also be written as $NMSEs + NMSE.cor + mNMSE$
NMSEs	Component of NMSE due to systematic errors	Equivalent expression: $\frac{4FB^2}{4 - FB^2}$
NMSEu	Component of NMSE due to unsystematic errors	$NMSE - NMSE_s = NMSE.cor + mNMSE$
NMSE.cor	Variance component of NMSE	
mNMSE	Minimum Normalised Mean-Square-Error	
FAC5	Fraction of model predictions within a factor of 5 of observations	Fraction of data satisfying $\frac{1}{5} \leq \frac{C_p}{C_o} \leq 5$
AE_band_conc	Concentration band used in the calculation of AE_band	
AE_band	Percentage of predictions that fall within a given absolute band around observations	$percentage(\lvert C_p - C_o \rvert < AE\_band\_conc)$ Only calculated if AE_band_conc is included in the output averaging times file (column ae.band).
rrank	Spearman Correlation Coefficient (or Ranked Correlation)	Rank values of C_o and C_p separately, then calculate as for R
NSD	Normalised Standard Deviation	$\frac{\sigma_p}{\sigma_o}$
theta	Linear Least-Squares Slope Parameter	$\frac{\overline{(C_p - \overline{C_p})(C_o - \overline{C_o})}}{\overline{(C_p - \overline{C_p})^2}}$
AU	Accuracy of Unpaired Peak	$\frac{C_{p_{peak}}^{U} - C_{o_{peak}}}{C_{o_{peak}}}$ where $C_{o_{peak}}$ is the peak observed value $C_{p_{peak}}^{U}$ is the unpaired peak predicted value
AP	Accuracy of Paired Peak	$\frac{C_{p_{peak}} - C_{o_{peak}}}{C_{o_{peak}}}$ where $C_{o_{peak}}$ is the peak observed value $C_{p_{peak}}$ is the corresponding (paired peak) predicted value
FMT	Figure of Merit in Time	$100 \frac{\sum_i min\{C_{pi}, C_{oi}\}}{\sum_i max\{C_{pi}, C_{oi}\}}$ where i denotes time i

Name	Description	Equation
HH	Symmetrically Normalised Root Mean-Square-Error	$\sqrt{\frac{\overline{(C_p - C_o)^2}}{\overline{C_p C_o}}}$
KS	Kolmogorov-Smirnov Parameter (Maximum absolute difference between the two cumulative distributions of the observed and predicted values)	$N \times \max_x \lvert prob(C_p \le x) - prob(C_o \le x) \rvert$ where N is the number of observed/predicted pairs x is an incrementally increasing concentration threshold The predicted and observed distributions are defined to differ "at level α" if $KS > N\sqrt{-\ln\left(\frac{\alpha}{2}\right)}$ where α is typically between 0.001 and 0.2
MSI	Mielke Square Index	$1 - \frac{N^{-1}\sum_{i=1}^{N}\lvert C_{oi} - C_{pi}\rvert^2}{\sigma_X^2 + \sigma_Y^2 + \left(\overline{C_o} - \overline{C_p}\right)^2}$ where σ_X^2 and σ_Y^2 are population variances
MMSI	Modified Mielke Index	$\lambda = \frac{2}{\frac{\sigma_X}{\sigma_Y} + \frac{\sigma_Y}{\sigma_X} + \frac{\left(\overline{C_o} - \overline{C_p}\right)^2}{\sigma_X \sigma_Y}} R$ where σ_X^2 and σ_Y^2 are population variances

Table 4-1 Details of the statistics always included in the Model Evaluation Tool concentration statistics output file

If the **Include openair modStats results in CSV file** option is selected, then a number of further additional columns of data are added to the CSV file; these are described briefly in Table 4-2. For more information about the statistics generated by the modStats function, please refer to openair documentation [3]. Each of these statistics is labelled with the prefix "modstats."

Statistic	Description
n	The number of complete pairs of data
FAC2	Fraction of predictions within a factor of two
MB	The mean bias
MGE	The mean gross error
NMB	The normalised mean bias
NMGE	The normalised mean gross error
RMSE	The root mean squared error
r	The Pearson correlation coefficient
P	The P value of the correlation coefficient

COE	The *Coefficient of Efficiency* based on Legates and McCabe (1999, 2012) which spans from -1 to +1 with values approaching +1 representing better model performance
IOA	The Index of Agreement based on Willmott et al. (2011), which spans between -1 and +1 with values approaching +1 representing better model performance

Table 4-2 Details of the statistics used by the openair modStats package, included in the Model Evaluation Tool concentration statistics output file

If the option to calculate spatial metrics is selected, the additional parameters listed in Table 4-3 are included in the concentration statistics CSV file. These parameters are based on the areas in which models and/or observations indicate exceedance of a particular threshold value, which must be entered in the 'ct' column in the output averaging times file.

Name	Description
AFN	Area of False Negatives For a particular threshold value, AFN is the area over which exceedance is observed but not predicted
AFP	Area of False Positives For a particular threshold value, A_{FP} is the area over which exceedance is predicted but not observed
FMS	Figure of Merit in Space (Threat Score) For a particular threshold value, FMS is the area over which exceedance is correctly predicted as a fraction of the area where exceedance is predicted and/or observed, at a fixed time
1DMOE	User-oriented 1-dimensional Measure of Effectiveness More generalised form of FMS where user can specify relative importance of false negative and false positive predictions
2DMOE.x	2-dimensional Measure of Effectiveness For a particular threshold value, 2DMOE.x is the area where exceedance is correctly predicted as a fraction of the area where exceedance is observed (measure of underprediction)
2DMOE.y	2-dimensional Measure of Effectiveness For a particular threshold value, 2DMOE.y is the area where exceedance is correctly predicted as a fraction of the area where exceedance is predicted (measure of overprediction)
RANK	Combined ranking metric Which combines values of correlation, fractional bias, figure of merit in space and Kolmogorov-Smirnov parameter.

Table 4-3 Additional parameters included in the Model Evaluation Tool concentration statistics output file when the option to calculate spatial metrics is selected

Mathematical expressions and equations for these parameters are given in Table 4-4. Here A_o is the area where observations indicate exceedance, and A_p is the area where predictions indicate exceedance. Values are calculated using a 'pairwise' method (using data summation as an estimate for area).

Name	Expression	Equation
AFN	$A_o - (A_p \cap A_o)$	$\frac{1}{2}\sum_i [\lvert C_{oi} - C_{pi}\rvert + (C_{oi} - C_{pi})]$

Name	Expression	Equation
AFP	$A_p - (A_p \cap A_o)$	$\frac{1}{2}\sum_i [\lvert C_{oi} - C_{pi}\rvert + (C_{pi} - C_{oi})]$
FMS	$\frac{A_p \cap A_o}{A_p \cup A_o}$	$\frac{\frac{1}{2}\sum_i [(C_{oi} + C_{pi}) - \lvert C_{oi} - C_{pi}\rvert - C_t}{A_p \cap A_o + A_{FN} + A_{FP}}$
1DMOE	$\frac{A_p \cap A_o}{A_p \cap A_o + \gamma_{FN} AFN + \gamma_{FP} AFP}$ where γ_{FN} and γ_{FP} are weighting factors determined by the user	
2DMOE.x	$\frac{A_p \cap A_o}{A_o}$	$MOE_{FN} = \frac{2 - FB_{FN} - FB_{FP}}{2 + FB}$ where FB_{FN} and FB_{FP} are the underpredicting (false-negative) and overpredicting (false-positive) components of the fractional bias
2DMOE.y	$\frac{A_p \cap A_o}{A_p}$	$MOE_{FP} = \frac{2 - FB_{FN} - FB_{FP}}{2 - FB}$ where FB_{FN} and FB_{FP} are the underpredicting (false-negative) and overpredicting (false-positive) components of the fractional bias
RANK	$R^2 + \left(1 - \left\lvert\frac{FB}{2}\right\rvert\right) + FMS + \left(1 - \frac{KS}{100N}\right)$	Where parameters are as defined in previous tables and N is the number of observed/predicted pairs

Table 4-4 Expressions for parameters included in the Model Evaluation Tool concentration statistics output file when the option to calculate spatial metrics is selected

If the option to create a target plot is selected, the additional parameters listed in Table 4-5 are included in the concentration statistics CSV file. Equations for these parameters are given in Section 4.2.1.7.

Name	Description
RMSU	Root mean square of the expanded measurement uncertainty
NMB	Normalised mean bias
MQI	Modelling Quality Indicator
MPI.bias	Modelling Performance Indicator – Bias
RMSUM	Model uncertainty expressed in concentration units
RMSUM.perc	Model uncertainty expressed as a percentage
MPI.r	Modelling Performance Indicator – Correlation
MPI.sd	Modelling Performance Indicator – Standard deviation
CRMSE.sign	Equal to CRMSE multiplied by 1 or -1
UOmean	Annual mean equivalent of RMSU
MQI.annual	Modelling Quality Indicator for annual means
RMSUM.annual	Annual mean equivalent of RMSUM
RMSUM.annual.perc	Annual mean equivalent of RMSUM.perc

Name	Description
Spatial_MPI_corr	Spatial Modelling Performance Indicator based on station annual means – Correlation
Spatial_MPI_sd	Spatial Modelling Performance Indicator based on station annual means – Standard deviation

Table 4-5 Additional parameters included in the Model Evaluation Tool concentration statistics output file when target plot output is selected

Some further summary statistics (listed in Table 4-6) are also produced for target plots. These can be found in a separate CSV output file named **_conc_target_stats.csv*.

Name	Description
MQI.90	90th percentile of all valid values of MQI
MPI.bias.90	90th percentile of all valid values of MPI.bias
MPI.r.90	90th percentile of all valid values of MPI.r
MPI.sd.90	90th percentile of all valid values of MPI.sd
RMSUM.90	90th percentile of all valid values of RMSUM
RMSUM.perc.90	90th percentile of all valid values of RMSUM.perc
MAQI.annual.90	90th percentile of all valid values of MQI.annual
RMSUM.annual.90	90th percentile of all valid values of RMSUM.annual
RMSUM.annual.perc.90	90th percentile of all valid values of RMSUM.annual.perc
output.plot	TRUE if a plot has been produced for this averaging time and pollutant, FALSE otherwise. Depends on the averaging time used in the air quality directive for the pollutant.

Table 4-6 Parameters included in the target plot statistics file output from the Model Evaluation Tool

If the option to calculate persistence is selected, the additional parameters listed in Table 4-7 are included in the concentration statistics CSV file. These parameters use a 'persistence' model (ps), which uses the observed concentrations from one day as the forecast concentrations for a number of days in the future. For example, if the *Forecast day for persistence forecast* is set to 1 (the default) then the observations from one day will be used as the persistence forecast for the next day. If the *Forecast day for persistence forecast* is 2, then the observations from one day will be used as the persistence forecast for the day after the next day.

Name	Description	Equation
MFEp	Mean Fractional Error for 'persistence' (*ps*) model (i.e. assuming no change in concentration)	$\overline{\left(\frac{\lvert C_{ps} - C_o \rvert}{0.5(C_{ps} + C_o)}\right)}$
MPI_FP	FAIRMODE Model Performance Indicator (MPI) for forecast modelling: Compares the error in predictions to (i) error in 'persistence' model predictions (*ps*), and (ii) measurement uncertainty. Fulfilled if ≤1.	$\frac{MFE}{MFE_{ps}}$

Table 4-7 Additional parameters included in the Model Evaluation Tool concentration statistics output file when the option to calculate persistence is selected

If the **Time plot averaged over all stations** option is selected, then an additional CSV file named **_conc_agg.csv* is created containing the aggregated values plotted on the graph. More details are given in Table 4-8.

Name	Description
date	The date and time of each data point
pollutant	Name of the pollutant measured
obs	The measured concentration of a pollutant averaged over all relevant observed stations and averaging time determined by the *output averaging times* file. The units are determined by the *pollutants* file
[Model Name]	There will be a column per modelled dataset. The modelled concentration of a pollutant averaged over all relevant modelled stations and averaging time determined by the *output averaging times* file. The units are determined by the *pollutants* file
num.obs.stations	Number of observation stations that measure each pollutant. The *obs* column values are averaged over these stations
num.[Model Name].stations	Number of modelled stations that measure each pollutant. The [Model Name] column values are averaged over these stations. There will be a column per model.
moderate.threshold	The moderate concentration threshold as defined in the *index scales* and *alert threshold* files. Only included when forecast index evaluation is performed.
high.threshold	The high concentration threshold as defined in the *index scales* and *alert threshold* files. Only included when forecast index evaluation is performed.

Table 4-8 Details of the headers in the CSV file produced when the 'Time plot averaged over all stations' option is selected

4.2.2 Forecast Index Evaluation Output

The forecast index evaluation option produces three types of output; these are described in this section.

All three types of graph are based on a 'forecast index', which is calculated from the modelled and observed data according to the index threshold definitions in the index scales file (described in Section 3.2, step 9).

4.2.2.1 Forecast index accuracy

This graph shows the performance of the model's forecast index predictions against forecast indices calculated from observed concentrations.

The graph is a stacked bar chart that shows, for each station, the percentage of calculated forecast indices valid for comparison where the modelled index was equal to the observed index (green) and where the modelled index was equal to the observed index plus or minus one band (grey). Only forecast periods for which both modelled and observed indices can be calculated are included in the assessment. The stations are sorted by the number of indices valid for comparison, which is also shown on the chart by the blue circles and the right-hand y-axis. A key to the stations is given below the graph. Refer to Figure 4-9 for an example.

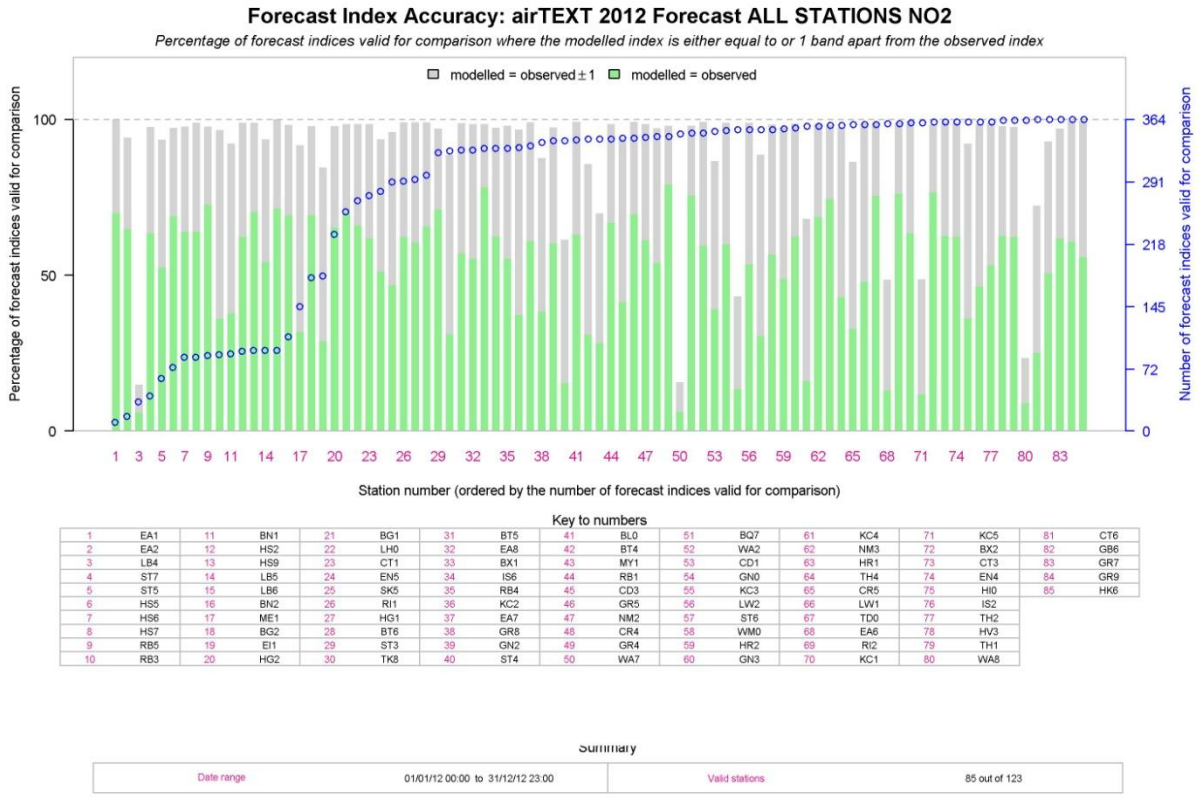


Figure 4-9 Example of the forecast index accuracy graph (grouped by station and filtered by model and pollutant; displaying model 'Forecast' and pollutant NO_2)

4.2.2.2 *Forecast alert accuracy*

It is usual in operational pollution forecasting to use pollution bandings to help communicate pollution levels to the public. For example, a common set of bandings for a 1 to 10 forecast index scale is shown in Table 4-9.

Band	Forecast index range	Alert threshold
LOW	1 to 3	n/a
MODERATE	4 to 6	4
HIGH	7 to 9	7
VERY HIGH	10	10

Table 4-9 Example set of bandings, with associated alert thresholds

Depending on the system, a forecast index in the MODERATE, HIGH or VERY HIGH range may trigger an alert to the public; it is therefore important for system operators to understand whether the system issues these alerts correctly.

The assessment of forecast alerts is carried out by calculating metrics for each monitoring station based on considering the exceedence of an alert threshold as an 'event'. The number of events observed and modelled, modelled but not observed, observed but not modelled, not modelled and not observed are summed to get the parameters *a*, *b*, *c* and *d* respectively. This is summarised in Table 4-10.

		Event observed	
		Yes	No
Event modelled	Yes	*a*	*b*
	No	*c*	*d*

Table 4-10 Definition of the forecast alert parameters

The forecast index evaluation includes two sets of graphical output for the assessment of the accuracy of forecast alerts:

1. **Odds ratio skill score (ORSS)**

   The odds ratio skill score (ORSS) is calculated from the alert metrics as follows:

$$\text{Odds ratio (OR)} = \frac{ad}{bc}$$

$$\text{ORSS} = \frac{OR - 1}{OR + 1}$$

   A perfect system will have *b* and *c* equal to zero, which means $\text{OR} \to \infty$, which means $\text{ORSS} \to 1$.

   A poor system will have *a* and *d* equal to zero, which means $\text{OR} = 0$, which means $\text{ORSS} = -1$.

   The odds ratio is a good metric for determining if a model is good at correctly issuing and not issuing alerts. It gives equal weighting to the correct prediction of an alert and to the correct non-prediction of a non-alert. If no alerts are observed or no alerts are forecast then ORSS is invalid. The graph shows ORSS for each station, where the stations are ordered by the number of observed alert threshold exceedences. The number of observed and modelled alerts for each station is also plotted in blue (right-hand y-axis). A key to the stations is shown below the graph. Figure 4-10 shows an example of an ORSS graph.

Figure 4-10 Example of a ORSS graph produced by the forecast index evaluation (grouped by pollutant and filtered by model; displaying model 'Forecast')

2. Performance metrics

In an operational pollution forecasting system, it is important not to issue alerts when there should not be an alert, but it is arguably *more* important to accurately issue an alert when an alert should be issued.

The following performance metrics give information about the skill of a model in terms of its ability to issue accurate alerts, and any tendency towards 'false alarms'.

$$\text{Probability of a correct forecast (PCF)} = \frac{a+d}{a+b+c+d}$$

$$\text{Probability of detection (POD)} = \frac{a}{a+c}$$

$$\text{False alarm ratio (FAR)} = \frac{b}{a+b}$$

$$\text{Probability of false detection (POFD)} = \frac{b}{b+d}$$

Each of these metrics lies in the range 0 to 1. A good score for PCF and POD is 1; a good score for FAR and POFD is 0. These metrics are combined to produce a 'total forecast skill':

$$\text{Total forecast skill (total.skill)} = \frac{1}{4}\big(PCF + POD + (1-FAR) + (1-POFD)\big)$$

A good score for the total forecast skill is 1.

The frequency bias score, defined below, is greater than 1 if the system has a tendency to issue 'false alarms', and less than 1 if the system has a tendency to fail to issue alerts when they should be issued.

$$\text{Frequency bias score (FBS)} = \frac{a+b}{a+c}$$

The graphical output from the performance metrics option is a series of five graphs, showing the PCF, POD, FAR, POFD and number of observed and modelled alert threshold exceedences for each station. Again, the stations are sorted by the number of observed alert threshold exceedences and a key to the stations is shown below the graphs. Figure 4-11 and Figure 4-12 show examples of this output grouped and filtered by different variables.

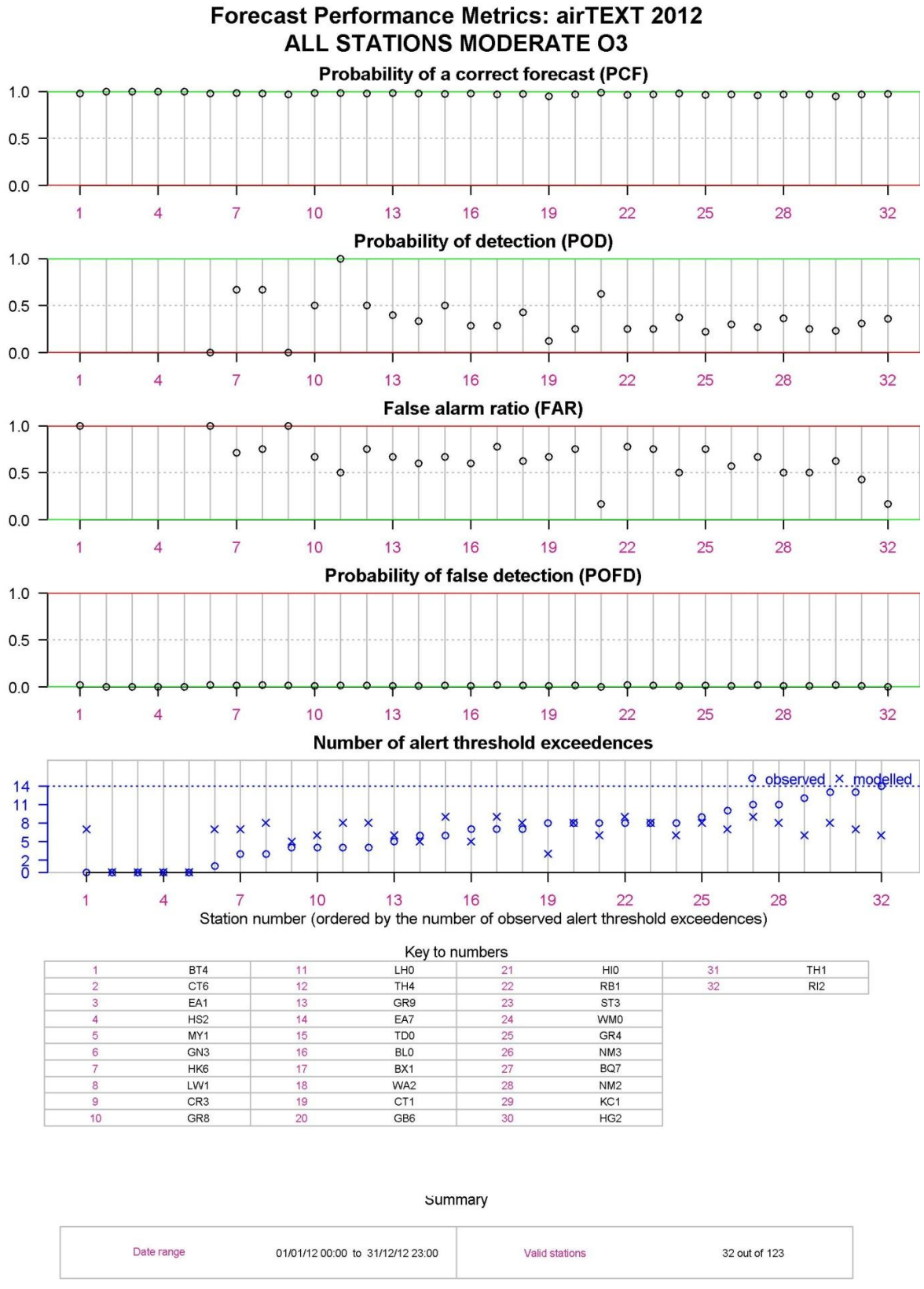


Figure 4-11 Example of a forecast performance metrics graph produced by the forecast index evaluation (grouped by station and filtered by model and pollutant; displaying pollutant O3)

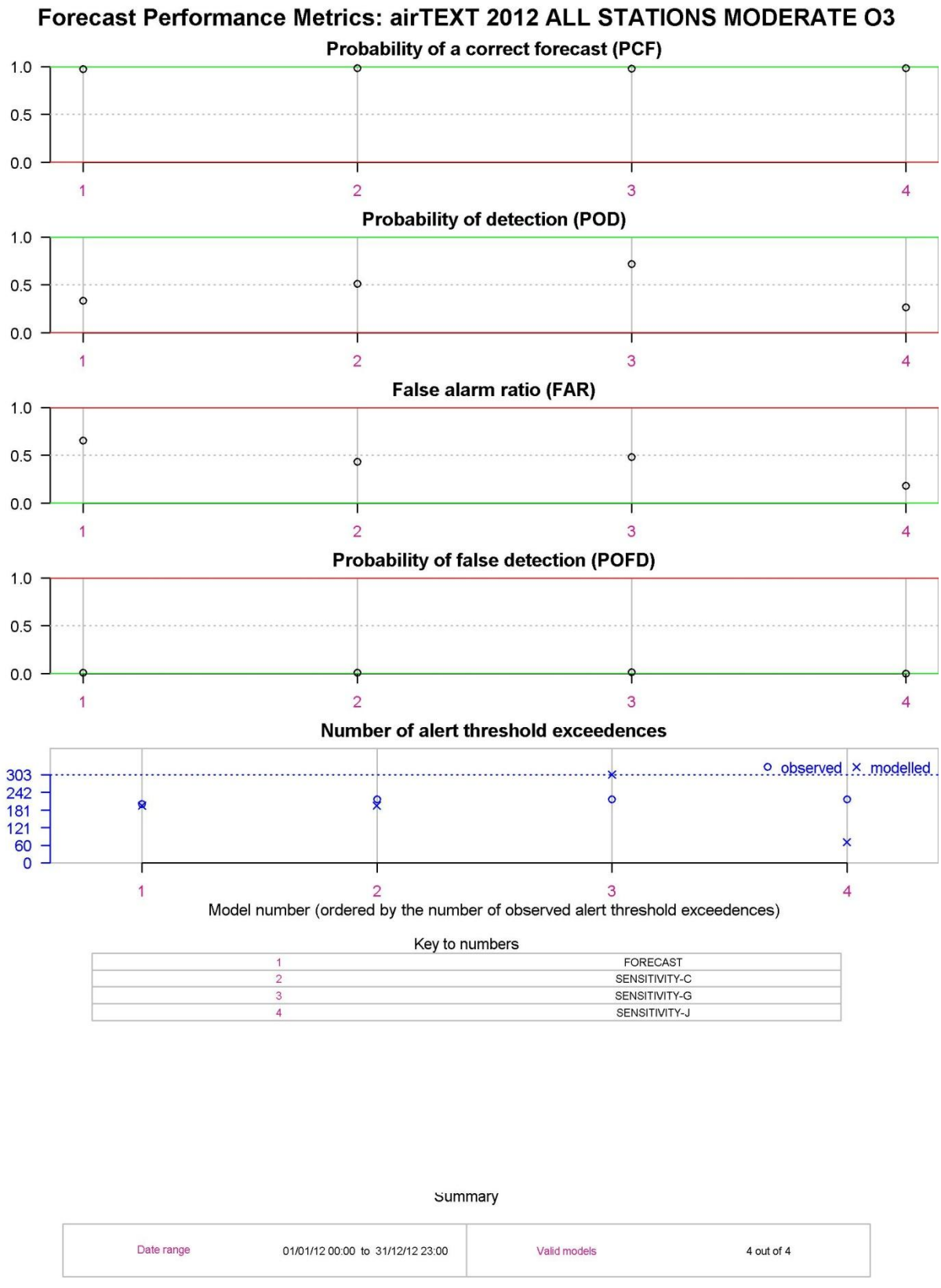


Figure 4-12 Example of a forecast performance metrics graph produced by the forecast index evaluation (grouped by model and filtered by pollutant; displaying pollutant O3)

4.2.2.3 CSV output files

The CSV files included in the forecast index evaluation output, and the data included in these files, are as follows:

**_forecast_index_data.csv*:

The data included in the **_forecast_index_data.csv* file are described in Table 4-11.

Name	Description
obs.output.units mod.output.units	concentrations for index averaging time and statistic, in output units
obs.index.units mod.index.units	concentrations for index averaging time and statistic, in index units
index.obs index.mod	index
index.absdiff	absolute difference between observed and modelled indices
alert.level.obs alert.level.mod	alert level

Table 4-11 Data included in the **_forecast_index_data.csv* output file

**_forecast_index_stats.csv*:

The data included in the **_forecast_index_stats.csv* file are described in Table 4-12.

Name	Description
num.valid.indices	number of valid calculated index values
perc.indices.same	percentage of indices where the modelled index was correct
perc.indices.oneapart	percentage of indices where the modelled index was one band either above or below the correct index

Table 4-12 Data included in the **_forecast_index_stats.csv* output file

_forecast_alert_stats.csv:

The data included in the *_forecast_alert_stats.csv* file (for each station, pollutant and alert threshold) are listed in Table 4-13. Full definitions of these parameters are given in Section 4.2.2.2.

Name	Description
num.obs.alerts	number of observed alerts
a	number of exceedences of alert thresholds that are both observed and modelled
b	number of exceedences of alert thresholds that are modelled but not observed
c	number of exceedences of alert thresholds that are observed but not modelled
d	number of exceedences of alert thresholds that are not observed and not modelled
FBS	Frequency Bias Score
POD	Probability Of Detection
PCF	Probability of a Correct Forecast
FAR	False Alarm Ratio
POFD	Probability Of False Detection
total.skill	Total Forecast Skill
OR	Odds Ratio
ORSS	Odds Ratio Skill Score
TS	Threat Score
GSS	Gilbert Skill Score
HSS	Heidke Skill Score
PSS	Peirce Skill Score
SR	Success Ratio
FAC2	Fraction of predictions within a factor of two (as for the concentration statistics)
RANK2	Combined ranking metric including forecasting components. $FA2 + POD + \gamma(1 - FAR)$ where $\gamma = \begin{cases} 0 \; if \; FAR = 0 \\ 1 \; if \; FAR \neq 0 \end{cases}$

Table 4-13 Data included in the *_forecast_alert_stats* output file

_forecast_episodes.csv:

The data included in the *_forecast_episodes.csv* file are described in Table 4-14.

Name	Description
start.date	episode start date
end.date	episode end date
num.days	duration of episode (days)
num.alerts.same	number of days during episode when observed and modelled alert levels are the same
max.alert.obs max.alert.mod	maximum alert level during episode
min.alert.obs min.alert.mod	minimum alert level during episode

Table 4-14 Data included in the *_forecast_episodes.csv* output file

**_forecast_episode_data.csv*:

The **_forecast_episode_data.csv* file contains the underlying daily data for the episodes listed in the **_forecast_episodes.csv* file. The data included in the **_forecast_episode_data.csv* file are described in Table 4-15.

Name	Description
obs.output.units mod.output.units	concentrations for index averaging time and statistic, in output units
obs.index.units mod.index.units	concentrations for index averaging time and statistic, in index units
index.obs index.mod	index reached
index.absdiff	absolute difference between observed and modelled indices
alert.level.obs alert.level.mod	alert level
episode.id	episode identification number

Table 4-15 Data included in the **_forecast_episode_data.csv* output file

4.3 Model Diagnostics Tool Output

The Model Diagnostics Tool can output a range of different plot types, as described in Section 3.3. Each plot shows the results for a single pollutant, station and modelled dataset. Plots are produced for all combinations of the pollutants, stations and datasets selected by the user. This can result in a large number of plots, so the plots are organised into sub-folders, which are named according to the type of plot they contain.

Examples of the different types of plot are shown in the following sections.

4.3.1 Time Variation Plot

The time variation plot compares modelled and observed concentrations by (clockwise from top) hour and day of the week, day of the week, month of the year and hour of the day. The shaded area indicates the 95% confidence interval in the mean. For an example, refer to Figure 4-13.

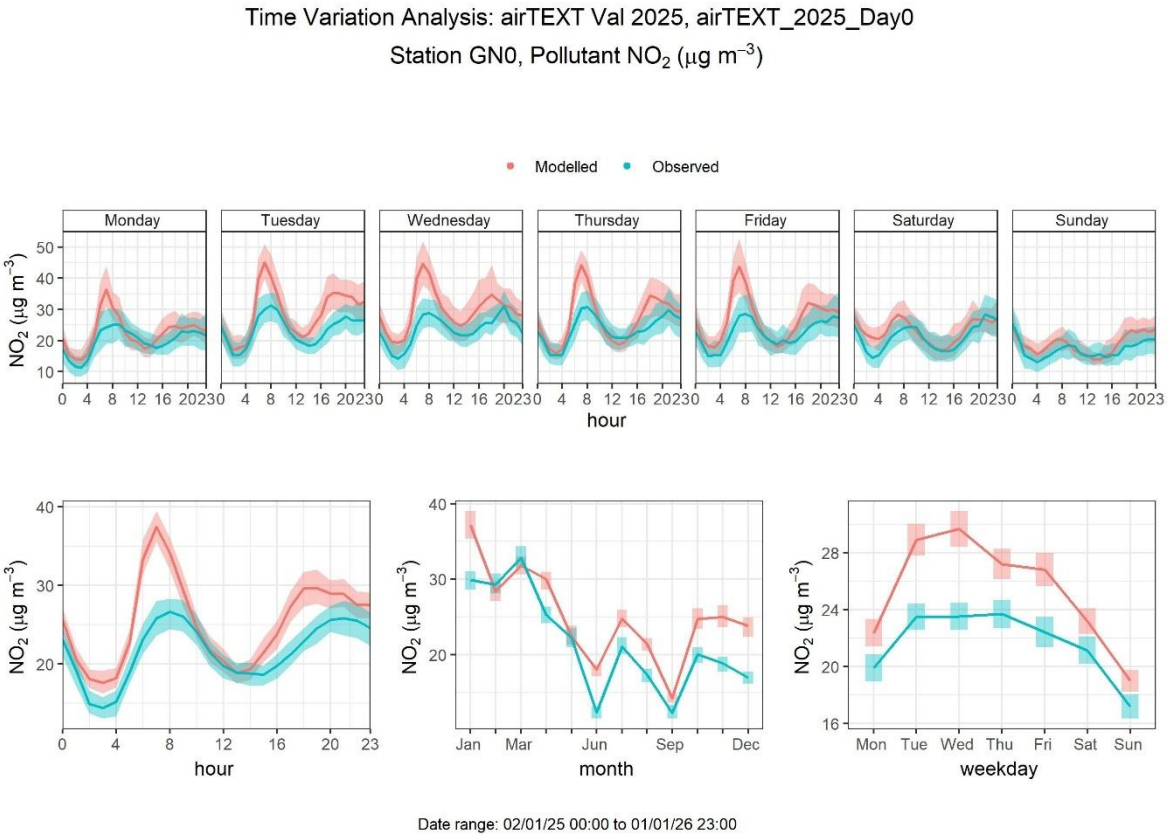


Figure 4-13 Example of a time variation plot from the Model Diagnostics Tool

4.3.2 Scatter Plot

The scatter plot shows a direct comparison of modelled and observed concentrations, optionally filtered by weekday, month or season. The frequency scatter plot shows the frequency of occurrence of each point whereas the conventional scatter plot shows one point per pairwise modelled-observed data point. The line of exact correspondence between observed and modelled concentration is shown as a solid black line. Points within the area bounded by the two dotted lines indicate that the modelled concentration and the observed concentration are within a factor of two. If the option to add a line of best fit is selected, this line and its equation are included on the plot. Figure 4-14 to Figure 4-16 show examples of frequency and conventional scatter plots.

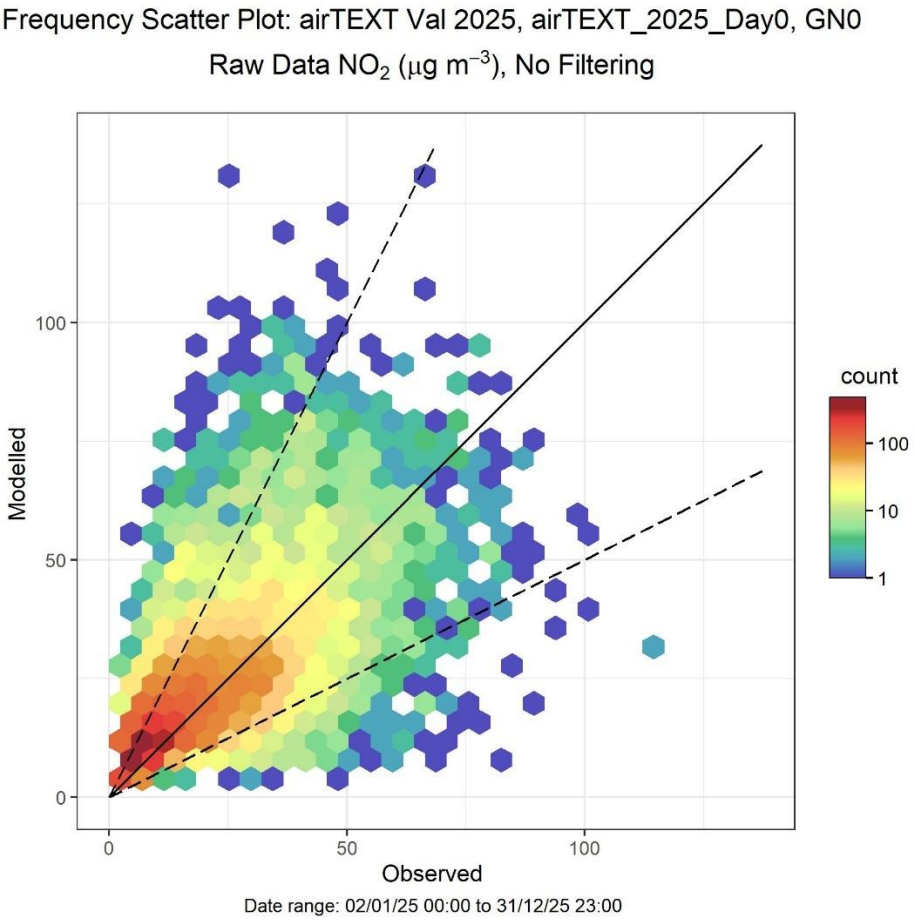


Figure 4-14 Example frequency scatter plot output from the Model Diagnostics Tool; raw data with no filtering

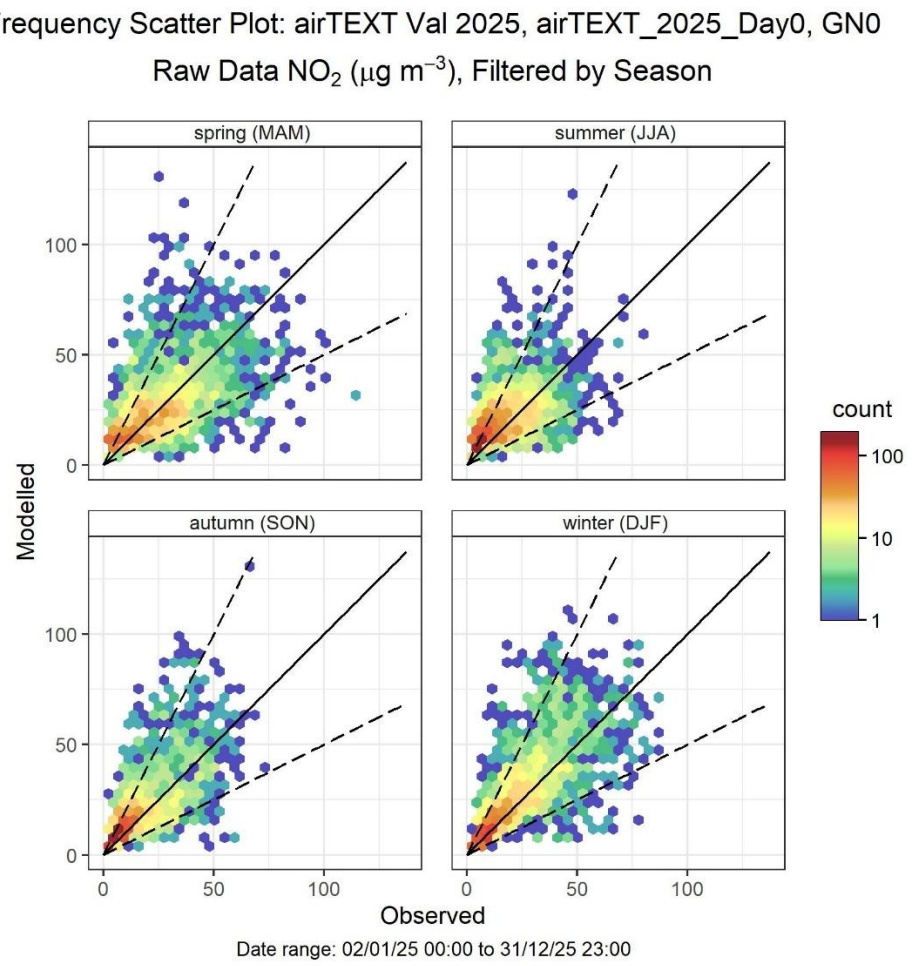


Figure 4-15 Example frequency scatter plot output from the Model Diagnostics Tool; raw data filtered by season

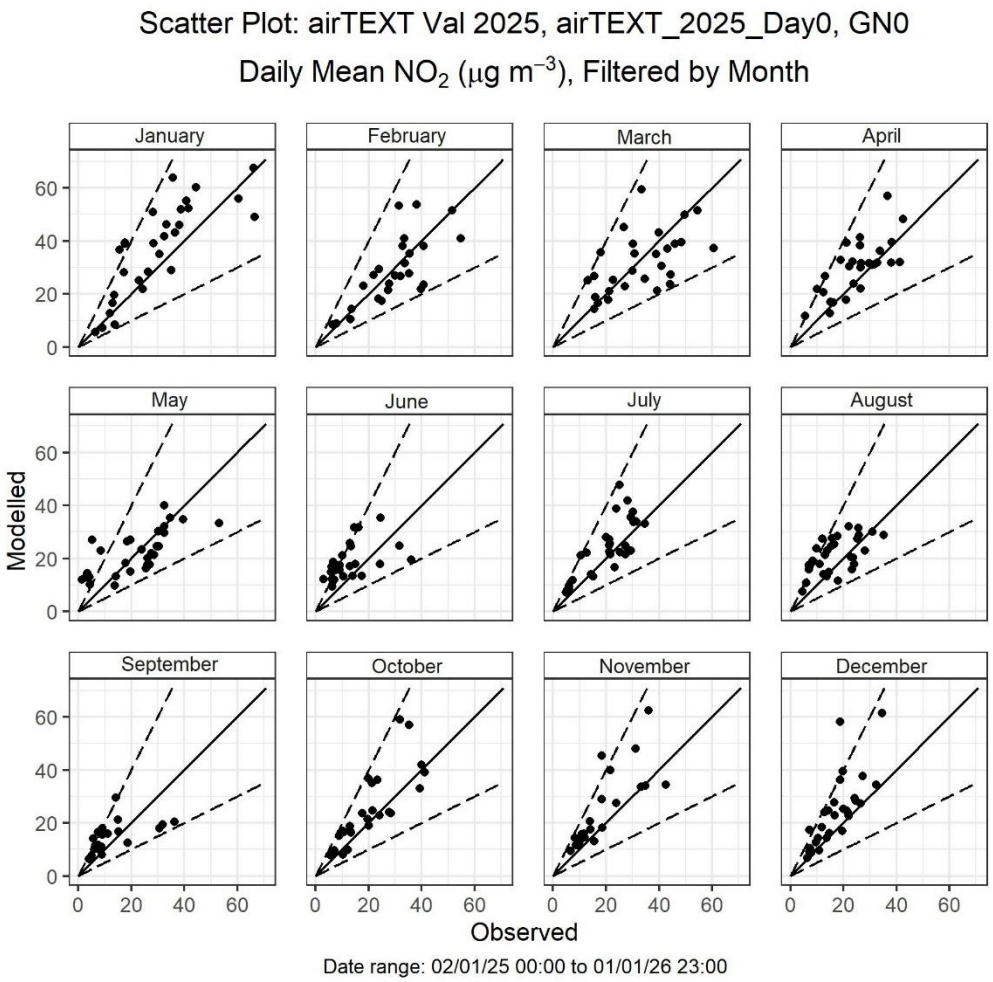


Figure 4-16 Example conventional scatter plot output from the Model Diagnostics Tool; calculated statistics (daily maximum NO_2) filtered by month

4.3.3 Time Plot

The time plot shows a time series of modelled and observed concentrations. Figure 4-17, Figure 4-18 and Figure 4-19 show example time plots.

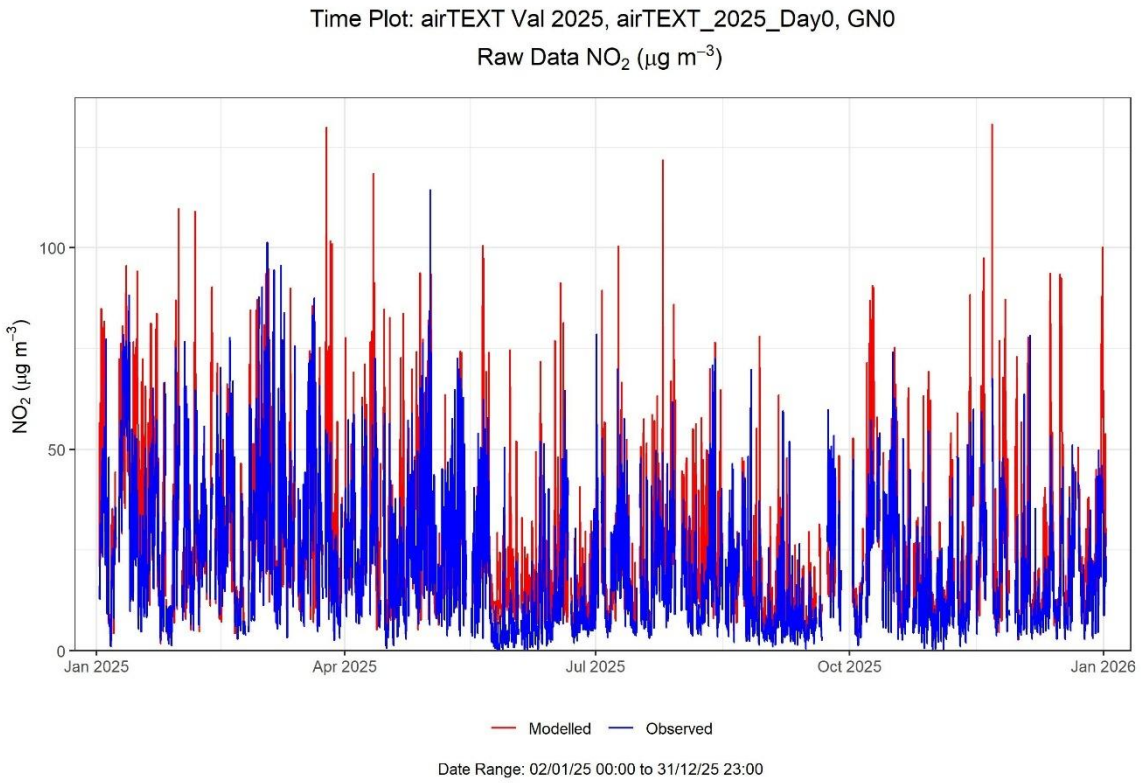


Figure 4-17 Example time plot output from the Model Diagnostics Tool; raw data

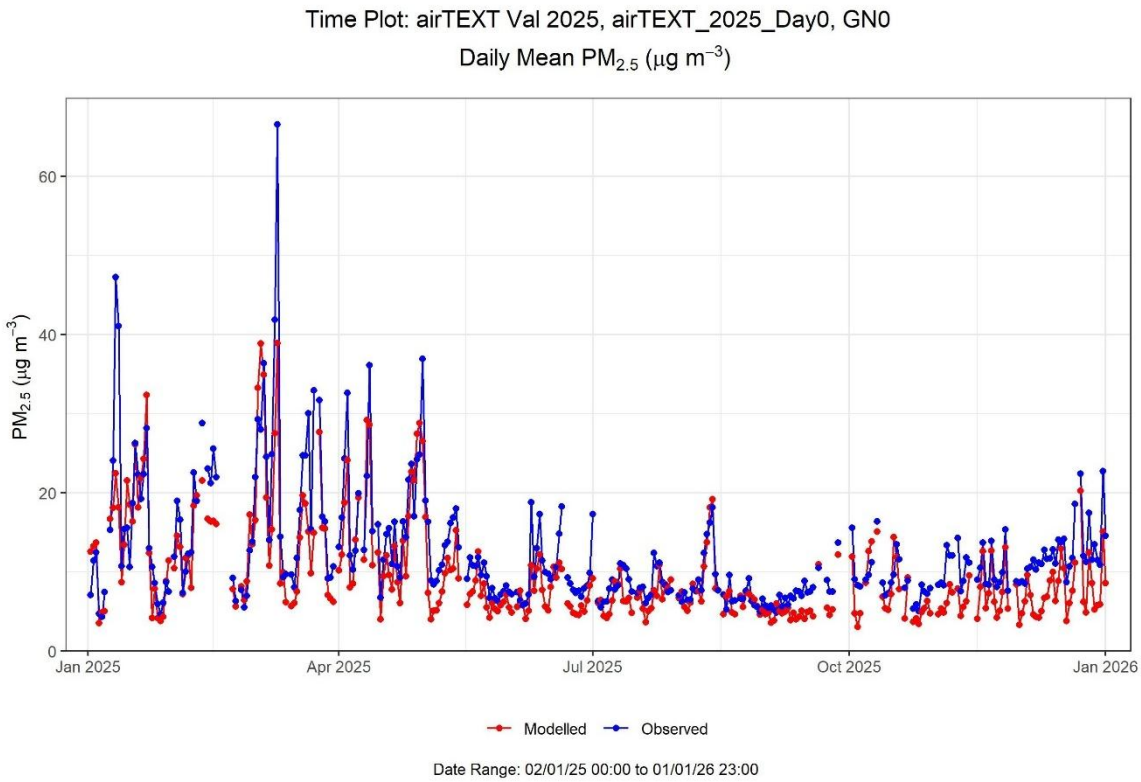


Figure 4-18 Example time plot output from the Model Diagnostics Tool; daily mean, with markers

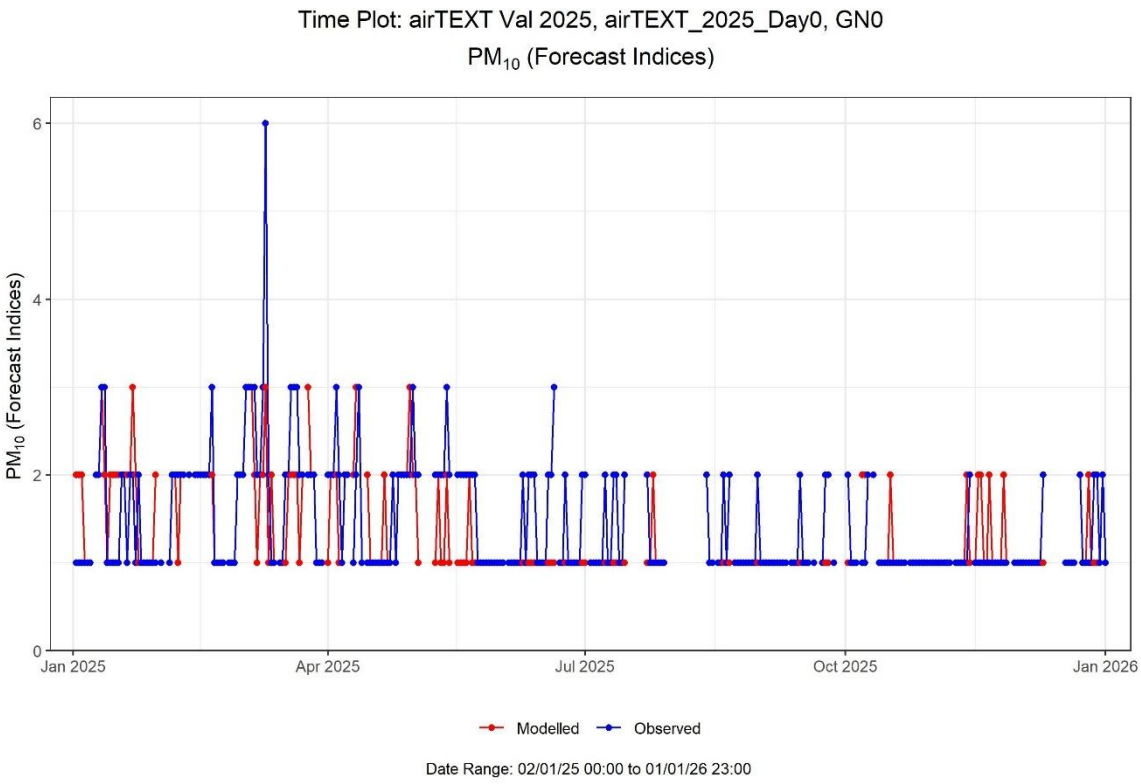


Figure 4-19 Example time plot output from the Model Diagnostics Tool; forecast indices, with markers

4.3.4 Polar Plots

4.3.4.1 Polar concentration plot

The polar concentration plot shows the variation in concentration with wind direction. The concentration data can be banded according to the wind speed, hour of the day or month of the year. Examples are shown in Figure 4-20 and Figure 4-21.

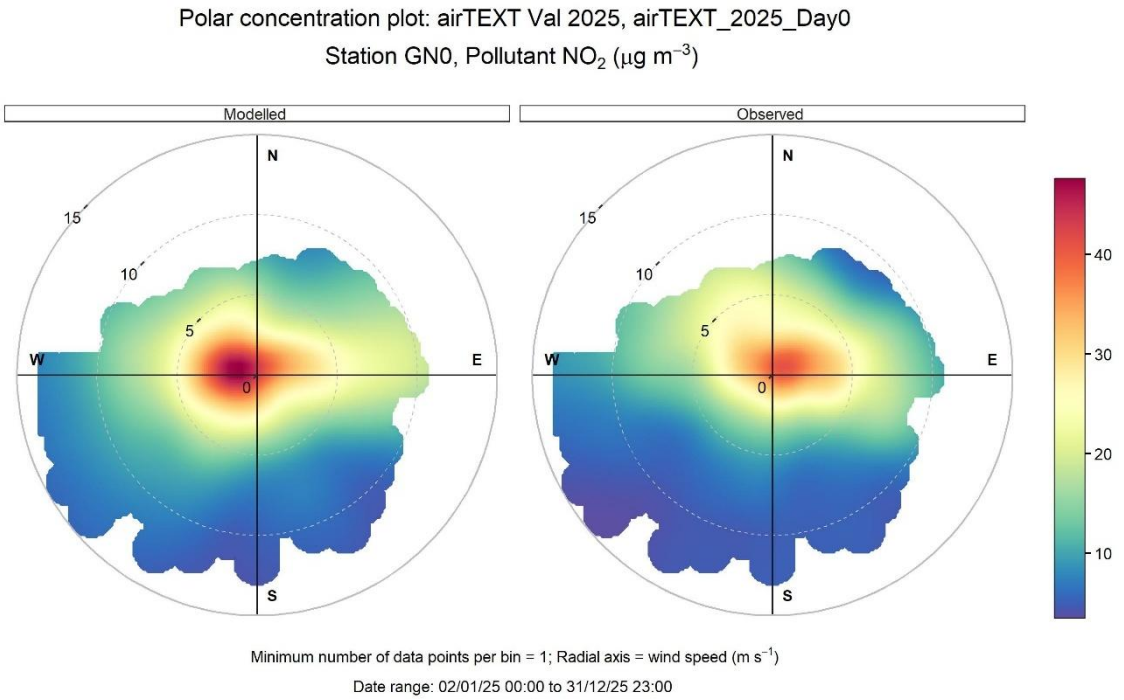


Figure 4-20 Example polar concentration plot output from the Model Diagnostics Tool; data banded according to wind speed

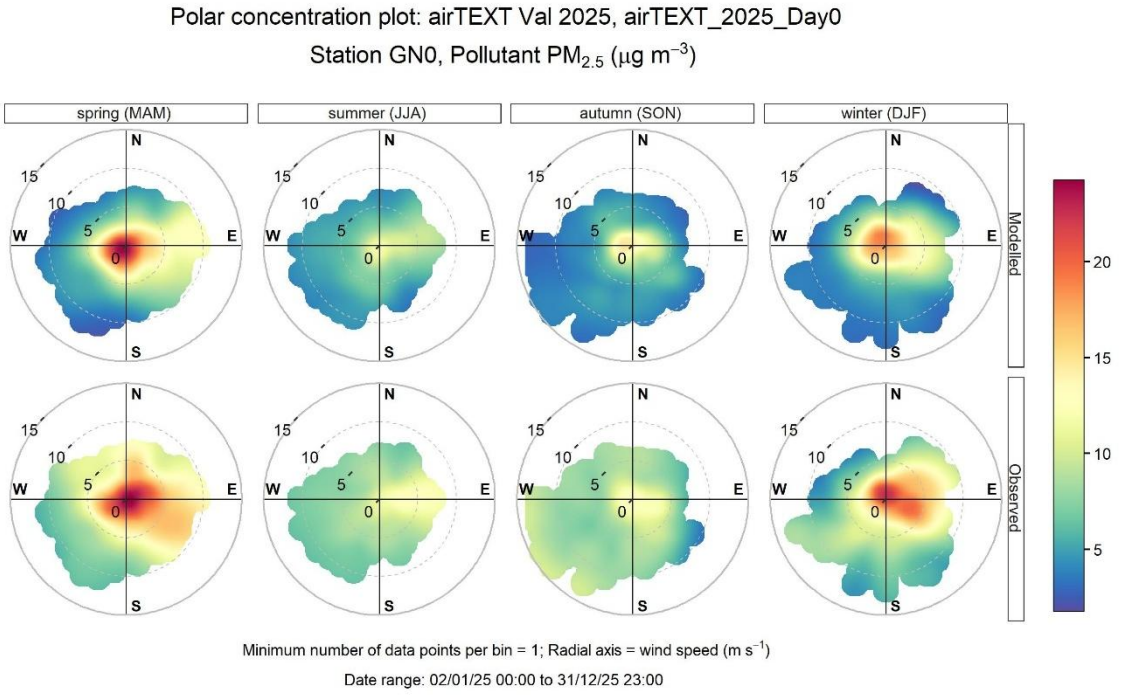


Figure 4-21 Example polar concentration plot output from the Model Diagnostics Tool; data banded according to wind speed, filtered by season

4.3.4.2 Pollution rose plot

The pollution rose plot shows, for each wind direction, the frequency that the concentration is in a particular range. An example is shown in Figure 4-22.

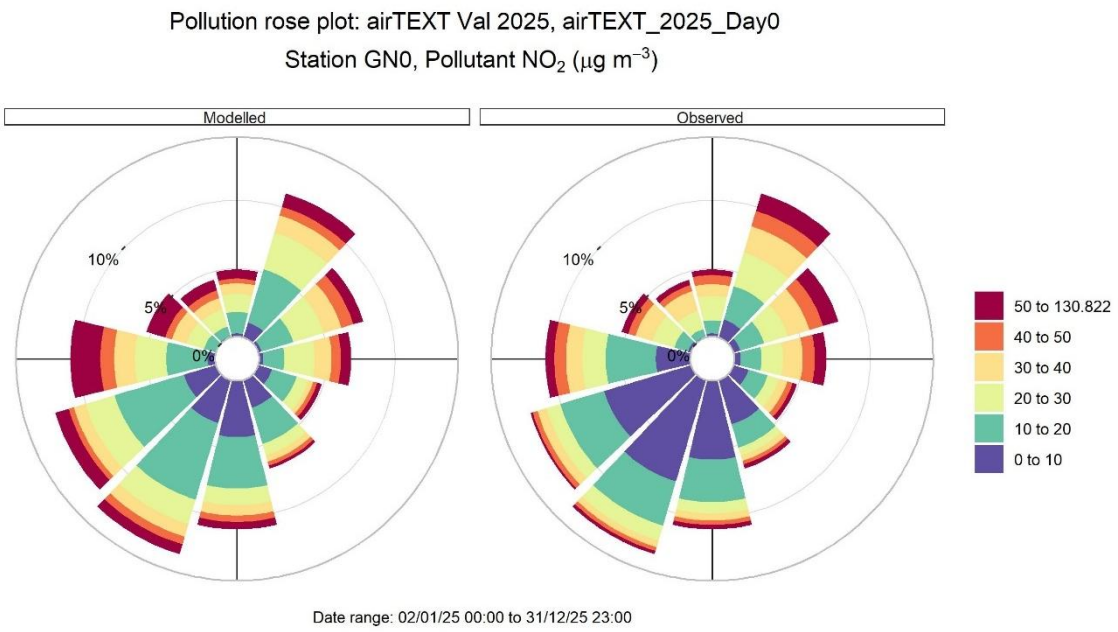


Figure 4-22 Example pollution rose plot output from the Model Diagnostics Tool

4.3.4.3 Polar wind speed frequency plot

The polar wind speed frequency plot shows the variation in wind speed with wind direction. An example is shown in Figure 4-23.

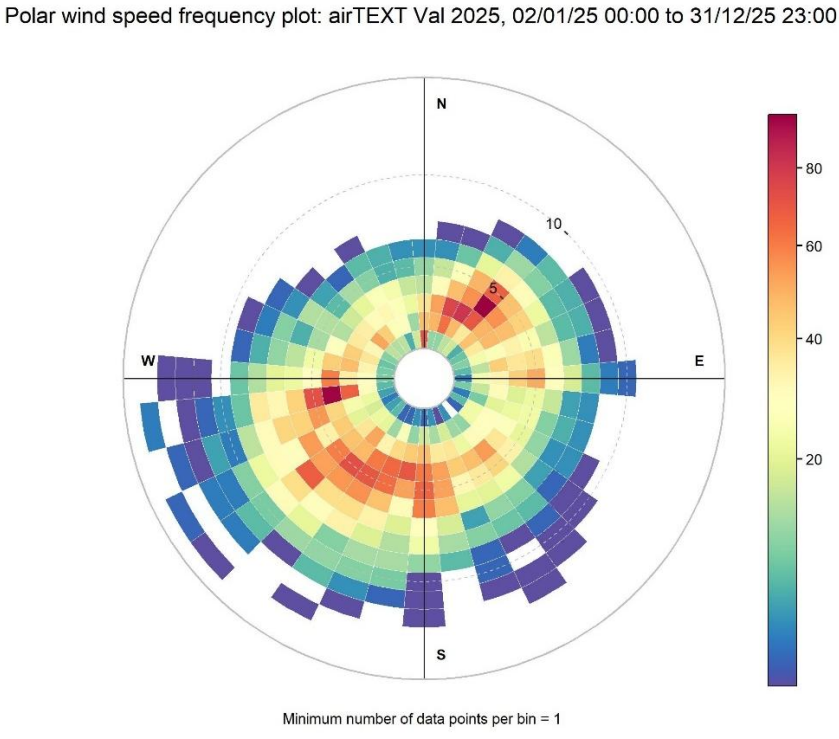


Figure 4-23 Example polar wind speed frequency plot output from the Model Diagnostics Tool

5 R Support

The three Model Evaluation Toolkit Tools each contain a number of utilities for managing the interaction of the Toolkit with R. Each Tool has an **R Support** menu, as shown in Figure 5-1.

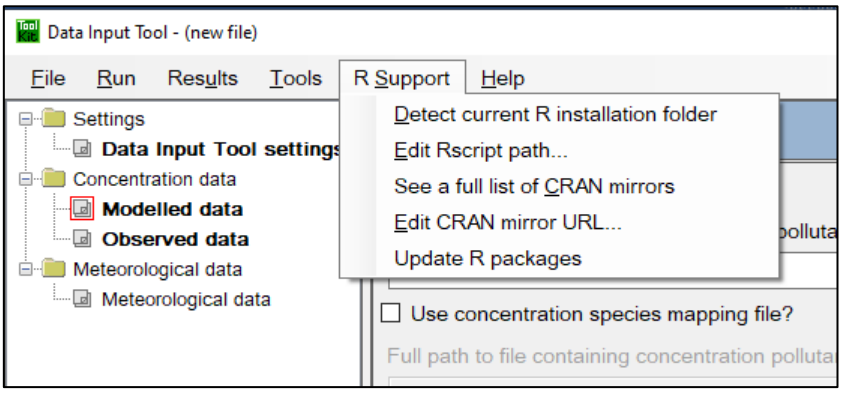


Figure 5-1 R Support menu

5.1 Detect current R installation folder

The first time you run any Model Evaluation Toolkit Tool, the Toolkit will automatically set itself to use the *version of R that was most recently installed*. If you subsequently install a different version of R and wish to update the Toolkit to use this newly-installed version of R, select **Detect current R installation folder**.

Note: The version of R used is recorded in the .log file produced each time a Tool is run.

5.2 Edit Rscript path...

This option displays the path where the version of R currently being used by the Toolkit is installed (Figure 5-2). If you wish the Toolkit to use a different installed version of R, edit the path as required and save the file.

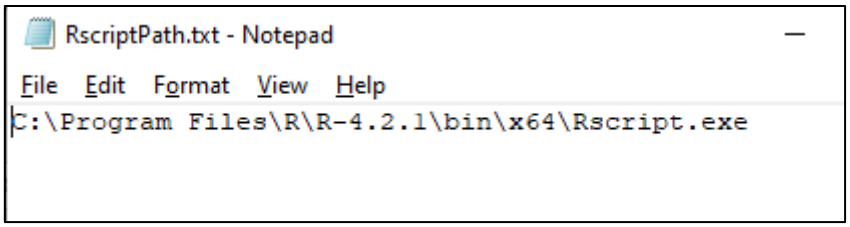


Figure 5-2 The *RscriptPath.txt* file

5.3 See a full list of CRAN mirrors

Select this option to view a list of the web addresses of all CRAN mirrors, on the R project website.

5.4 Edit CRAN mirror URL...

The web address of your preferred CRAN mirror must be specified. The mirror most relevant for your geographical location should be chosen. The default is suitable for the UK; for other locations select **See a full list of CRAN mirrors** from the R Support menu to see a list of the web addresses of all CRAN mirrors. Copy the link that applies to your location, then select **Edit CRAN mirror URL...**, replace the

link in the text file with your copied link and save the file. To ensure this CRAN mirror is used as the default mirror in future, see Section 6.

5.5 Update R packages

Selecting this option will update your installed R packages to the latest version, according to the CRAN mirror specified. It is good practice to update your R packages from time to time, to ensure the version you are using contains the latest functionality and improvements.

6 Changing the default settings using a template file

You may wish to make changes to the default settings used by one or more of the Tools in the Toolkit. For example, if you are outside the UK you may wish to use a non-UK CRAN mirror for R. Changes to the default settings can be made using a 'template' file, as follows:

Step 1: Open the Tool concerned.

Step 2: Change the settings to your new default values.

Step 3: Save the file.

Step 4: From the **File** menu, select **Preferences...** and then select the **Template** tab (**Figure 6-1**).

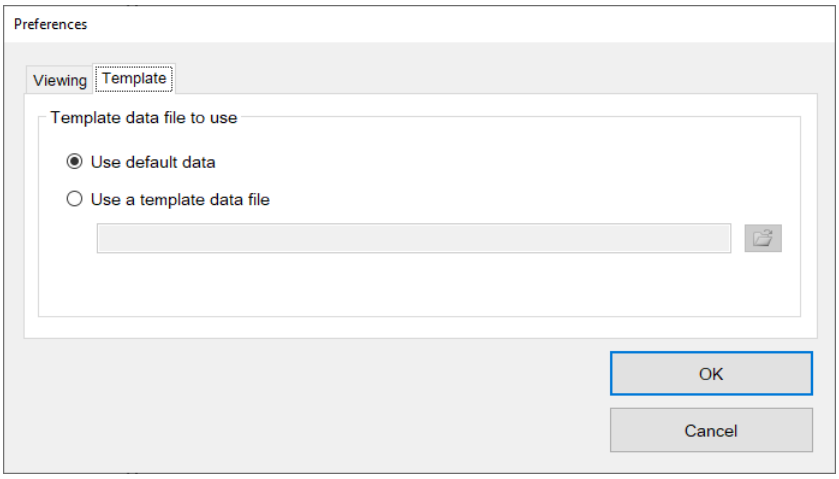


Figure 6-1 Preferences screen showing the Template tab

Step 5: Select **Use a template data file** and browse to select the file you saved in Step 3.

Step 6: Click OK.

Step 7: If necessary (for instance to change the default CRAN mirror) repeat Steps 1 to 6 for the other Tools.

In future, any new input files will use the settings in your saved template files as the default settings. If you wish to revert to the standard default settings, simply select **Use default data** on the **Preferences** screen.

7 Batch mode facility

The Data Input, Model Evaluation and Model Diagnostics Tools can be run from the command-line, making it possible to automate tasks, for example in a batch file.

The procedure is the same for each of the Tools:

1. Open the Tool, enter your settings and save the settings file
2. On the command-line (or in a batch file) type

   {R's Rscript.exe path} {Path of the Tool's .r file} {settings file path}

More information is given below about the three components of the instruction:

{R's Rscript.exe path}: This is the path of the *Rscript.exe* file installed with R. A typical path on a Windows PC (where R was installed in *C:\Program Files\R\R-4.5.3*) would be

```
"C:\Program Files\R\R-4.5.3\bin\x64\Rscript.exe"
```

{Path of the Tool's .r file}: This is the path of the R-script for the Tool you want to run. For example, for the Data Input Tool, if the Toolkit is installed at *C:\ModelEvaluationToolkit* then the path would be

```
"C:\ModelEvaluationToolkit\DataInput.r"
```

{settings file path}: This is the path of the settings file that you saved. For example, if you saved a Data Input Tool settings file to the folder *C:\Toolkit_output* with the filename *example.tki* then the settings file path would be

```
"C:\Toolkit_output\example.tki"
```

8 Generic CSV file format

The Generic CSV file format should be used:

- to supply observed data, if these data are not auto-downloaded or supplied from an existing workspace
- to supply modelled data, if the **Generic CSV** modelled data format is selected

An example file in Generic CSV format, *BN1_BN2.csv*, can be found in the *Data* folder, and is shown in Figure 8-1. A description of the data columns to be included in the file is given in Table 8-1.

	A	B	C	D	E	F	G	H
1	year	month	day	hour	station	NO2	PM10	
2	2008	1	1	1	BN1	-999	103.1	
3	2008	1	1	2	BN1	59.1	37.6	
4	2008	1	1	3	BN1	61.8	24.6	
5	2008	1	1	4	BN1	59.2	25.8	
6	2008	1	1	5	BN1	65.4	24.5	
7	2008	1	1	6	BN1	74.3	23.7	
8	2008	1	1	7	BN1	77.1	23.1	
9	2008	1	1	8	BN1	79.2	20.8	
10	2008	1	1	9	BN1	78.2	20.7	
11	2008	1	1	10	BN1	83.1	20.6	
12	2008	1	1	11	BN1	73.5	20.3	
13	2008	1	1	12	BN1	86.6	23	
14	2008	1	1	13	BN1	94.2	24.9	
15	2008	1	1	14	BN1	85.3	23.1	
16	2008	1	1	15	BN1	93.8	27.8	
17	2008	1	1	16	BN1	97.3	28	

Figure 8-1 Example file in Generic CSV format

Column header	Description	Allowed values
year	Year	Numeric
month	Month of the year	Numeric
day	Day of the month	Numeric
hour	Hour of the day	Numeric
station	Station name; this should match the station name in the stations definitions file.	Any string value e.g. “MY1”, “BN1”, “Station1”
(pollutant name)	Pollutant concentration Repeat this column for each pollutant you wish to include in the file	Column header: any string value, e.g. “NO2”. Pollutant concentration: numeric

Table 8-1 Data included in a file in ‘Generic CSV’ format

Note:

- *If a single file of modelled data is supplied, it must contain data for each pollutant to be analysed. If multiple files of modelled data are supplied, each pollutant to be analysed must be present in at least one file.*
- *Observed data files do not have to include data for all pollutants to be analysed.*

9 Observation networks supported for auto-download

The Data Input Tool allows observed data from a number of networks to be automatically downloaded during the Data Input Tool run. If this option is used, data for all required stations and pollutants will be downloaded and imported for a time period matching the modelled data.

This option requires internet access.

The following observation networks are supported for auto-download. Details of the stations included in each network can be found in the *Monitoring_stations_metadata* sub-folder of your Model Evaluation Toolkit installation folder.

- Imperial College London (KCL)

  The Environmental Research Group (ERG) at Imperial College London (formerly at King's College London) maintains a network of monitoring sites in Greater London.

- UK Automatic and Rural Network (AURN)

  The UK AURN is the national network of automatic monitoring stations around the UK.

- Air Quality England (AQE)
- Welsh Air Quality Network (WAQN)
- Scottish Air Quality Network (SAQN)
- Northern Ireland (NI)
- Europe (EUROPE)

10 References

1. Thunis, P. and C. Cuvelier (2022). **DELTA version 7.0: Concepts / User's guide / Diagrams.** http://aqm.jrc.ec.europa.eu/index.aspx
2. Carslaw, D.C. and K. Ropkins, (2012). **openair — an R package for air quality data analysis**. Environmental Modelling & Software. Volume 27-28, pp. 52–61. https://doi.org/10.1016/j.envsoft.2011.09.008
3. Carslaw, D.C. and Davison, J. (2026). **The openair book — A Guide to the Analysis of Air Pollution Data**. https://openair-project.github.io/book/
4. Stidworthy, A., Hamilton, V., Stocker, J., Scratcherd, E., Hood, C., Connolly, D., Johnson, K. and Carruthers, D. (2026). **A review of methods used to assess the performance of atmospheric dispersion models**. Report for the UK Atmospheric Dispersion Modelling Liaison Committee https://admlc.com/publications/ (accessed 5 June 2026)
5. R Development Core Team (2010). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.
6. Thunis, P., D. Pernigotti and M. Gerboles (2013). **Model quality objectives based on measurement uncertainty. Part I: Ozone.** Atmospheric Environment, Volume 79, Pages 861-868. https://doi.org/10.1016/j.atmosenv.2013.05.018
7. Pernigotti, D., M. Gerboles, C. Belis and P. Thunis (2013). **Model quality objectives based on measurement uncertainty. Part II: PM10 and NO2.** Atmospheric Environment, Volume 79, Pages 869-878. https://doi.org/10.1016/j.atmosenv.2013.07.045