



Italian National Agency for New Technologies,  
Energy and Sustainable Economic Development



WELCOME

*Fairmode WG3 training session  
Bologna and online, 31 May 2023*

Antonio Piersanti, Alexandra Monteiro  
Lina Vitali, Kees Cuvelier, Philippe Thunis



## morning

afternoon

### **10.30 – 11.00 Welcome and Theoretical Session**

*Rationale, main features and formulations of the methodology for Forecast validation included within the Delta Tool – Alexandra and Antonio*

### **11.00 – 12.30 Training Session**

*11.00 – 11.20 Overview of the available Diagrams for Forecast validation – Antonio*

*11.20 – 11.30 How to prepare a new database – Kees*

*11.30 – 12.30 Showcase of examples of Applications – Kees*

### **12.30 – 13.30 Lunch Break**

### **13.30 – 14.40 Breakout groups working session**

*13.30 – 13.40 Description of the breakout groups working session and delivering of the working materials: both the material used during the morning and a proposal for some new exercises, together with all the expected outcomes. A list of recommended exercises, chosen from the delivered working materials, will be provided to the attendees according to their declared Delta Tool user level: beginner, intermediate, advanced – Lina*

*13.40 – 14.40 Attendees' exercises. A trainer will be present in each breakout group in case issues arise.*

### **14.40 – 15.30 Questions & Answers session – closing**

*Questions, issues and ideas raised from attendees' personal work and/or from previous own applications.*

### **15.30 – 16.00 Issues on preparation of individual data (optional)**

*Questions and issues on preparation of own input data.*

# Our history

2020-...

- the proposed indicators were **tested by CT3 community** P. Durka (IEP), L. Vitali (ENEA), A. Monteiro, C. Gama (UniAveiro), G. Giovannini, M. Stortini, R. Amorati, G. Verratti (ARPAE), A. Tanzarella (Arpa Puglia), E. Grøtting Wærsted (MetNorway), A. Bartocha (ATMOTERM), S. Vranckx and A. D'Ausilio (VITO) and CAMS Regional

2021

- feedback of the users was collected and discussed during hackathons & FAIRMODE Technical Meetings
- consensus was reached on the final current formulation

2022

- a new version of the DELTA Tool (7.0) was developed including the new indicators (available download at <https://aqm.jrc.ec.europa.eu/index.aspx>) → thanks to Kees Cuvelier
- **FAIRMODE Guidance Document on Modelling Quality Objectives and Benchmarking** was produced including the new formulation (<https://publications.jrc.ec.europa.eu/repository/handle/JRC129254>) → thanks to Philippe Thunis

2023

- CT3 → the new WG3 (FAIRMODE roadmap 2023-2025)
- paper submitted to Geoscientific Model Development → in revision
- first Training Workshop (Bologna & online) → we are making history

# FORECAST Modelling Quality Objective: DELTA tool proposal

The forecast MQO should test 3 different features of a forecast model:

**Detection of the start / end of an episode**  
(sudden changes in the concentration)



use of “persistence model” as a benchmark

**Threshold exceedances**  
(as trigger for short term action plans)



use of standard threshold indicators

**Air Quality Index forecast**  
(public information and AQ Directive )



use of AQI agreement/comparison

# FORECAST Modelling Quality Objective: DELTA tool proposal

Detection of the start / end of an episode  
(sudden changes in the concentration)



use of “persistence model” as a benchmark

Target for the forecast model M is to do better than the “persistence” model P

$$MQI_{forecast} = \sqrt{\frac{\frac{1}{N} \sum_{i=1}^N (M_i - O_i)^2}{\frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2}}$$

$MQO_{forecast}$  is fulfilled if  $MQI_{forecast} \leq 1$ ,

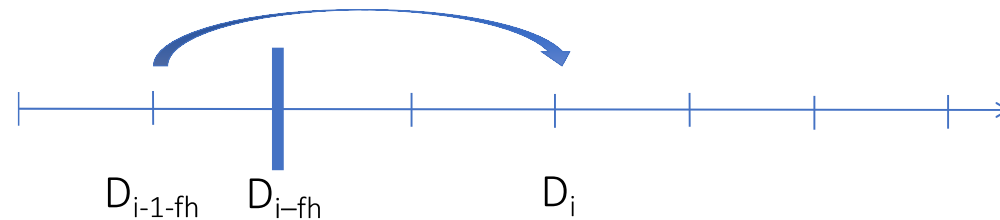
# FORECAST Modelling Quality Objective: DELTA tool proposal

Detection of the start / end of an episode  
(sudden changes in the concentration)



use of “persistence model” as a benchmark

- For day  $D_i$  forecast was made at day  $D_{i-\text{forecast horizon}}$ . ( $P_i = O_{i-1-\text{forecast horizon}}$ )
- At day of the forecast, only observation from the previous day are available
- The persistence model uses these observation for all forecast horizons



Forecast horizon  $fh = 0, +1, +2, \dots +n$

# FORECAST Modelling Quality Objective: DELTA tool proposal

Detection of the start / end of an episode  
(sudden changes in the concentration)



use of “persistence model” as a benchmark

Target for the forecast model M is to do better than the “persistence” model P, **taking into account the uncertainty of observations (OU)**

(after some tests...)  
(Feb 2021)

$$MQI_{forecast} = \sqrt{\frac{\frac{1}{N} \sum_{i=1}^N (M_i - O_i)^2}{\frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2}}$$

$$P_i = O_{i-1-forecast\ horizon} \pm OU(O_{i-1-forecast\ horizon})$$

$MQO_{forecast}$  is fulfilled if  $MQI_{forecast} \leq 1$ ,

# FORECAST Modelling Quality Objective: DELTA tool proposal

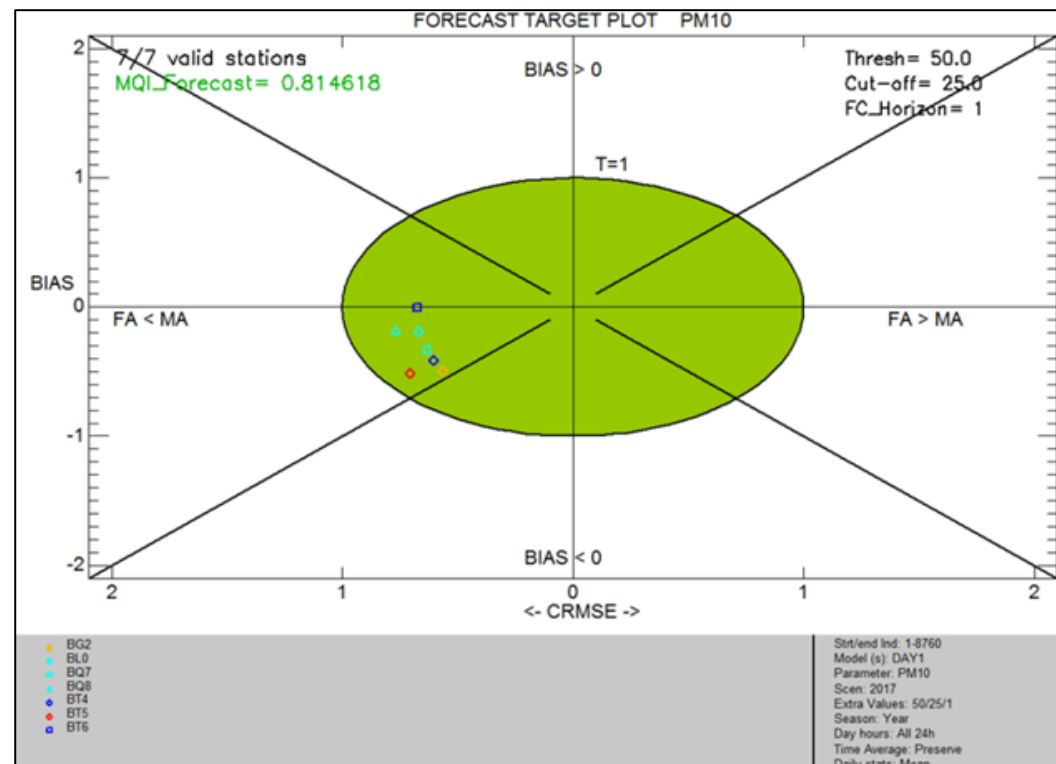
Detection of the start / end of an episode  
(sudden changes in the concentration)

comparison with the persistence model  
TARGET PLOT

- Target forecast: RMSE
- Y-axis: Bias
- X-axis: CRMSE
- Left/right asymmetry:

- **Cut-off:** remove lower concentration values to put more emphasis on high episodes

their arbitrary choice influence the outcomes  
(MQO fulfilment) -> consensus on taking this  
option ultimately off from the current version





# FORECAST Modelling Quality Objective: DELTA tool proposal

Detection of the start / end of an episode  
(sudden changes in the concentration)



comparison with the persistence model  
FORECAST MPI Plot

A **new evaluation diagram** was proposed and useful to support the interpretation of results

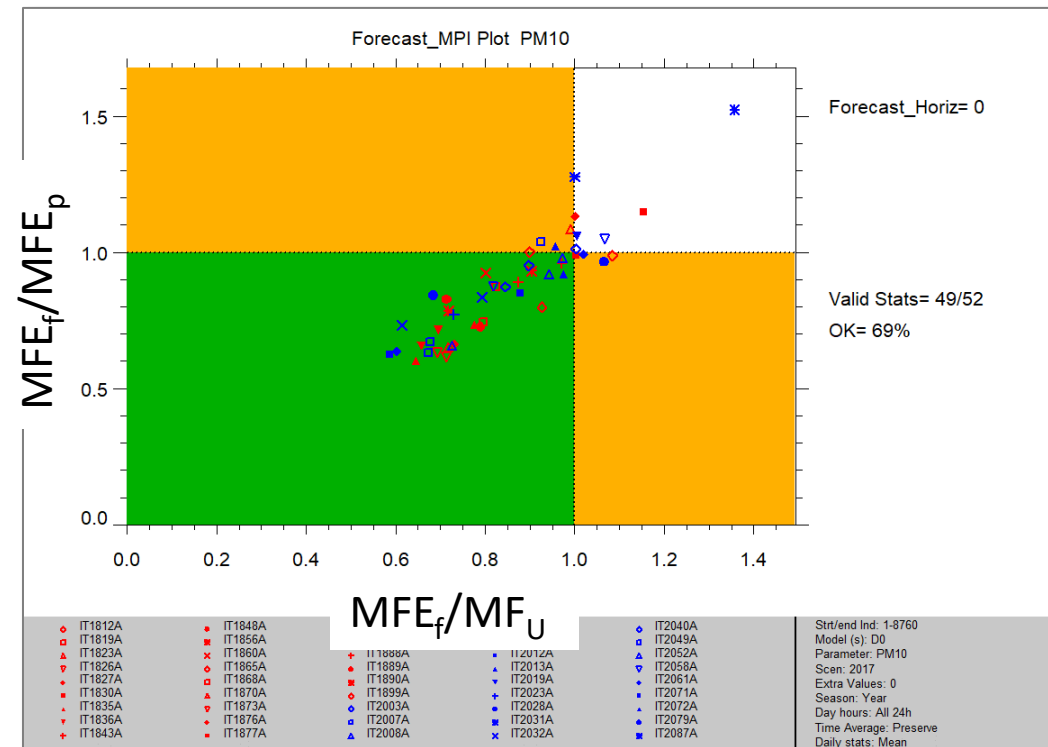
→ **the Forecast MPI Plot**, where MPIs are defined for the Mean Fractional Error (MFE) statistical indicator

$$MFE_f = \frac{2}{N} \sum_{i=1}^N \frac{|M_i - O_i|}{(M_i + O_i)}$$

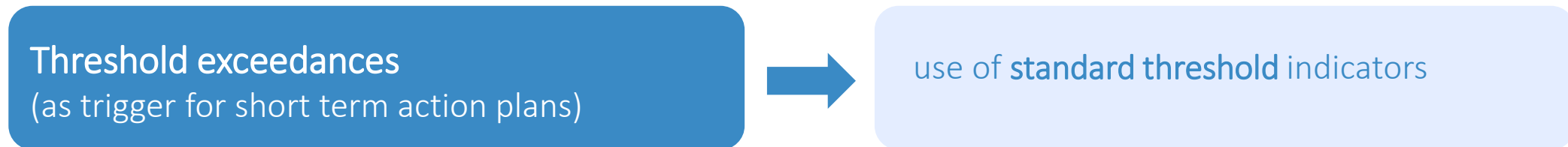
$$MFE_p = \frac{2}{N} \sum_{i=1}^N \frac{|P_i - O_i|}{(P_i + O_i)}$$

$$P_i = O_{i-1-\text{forecast horizon}} \pm U(O_{i-1-\text{forecast horizon}})$$

$$MF_U = \frac{1}{N} \sum_{i=1}^N \frac{2U(O_i)}{O_i}$$



## FORECAST Modelling Quality Objective: DELTA tool proposal



# FORECAST Modelling Quality Objective: DELTA tool proposal

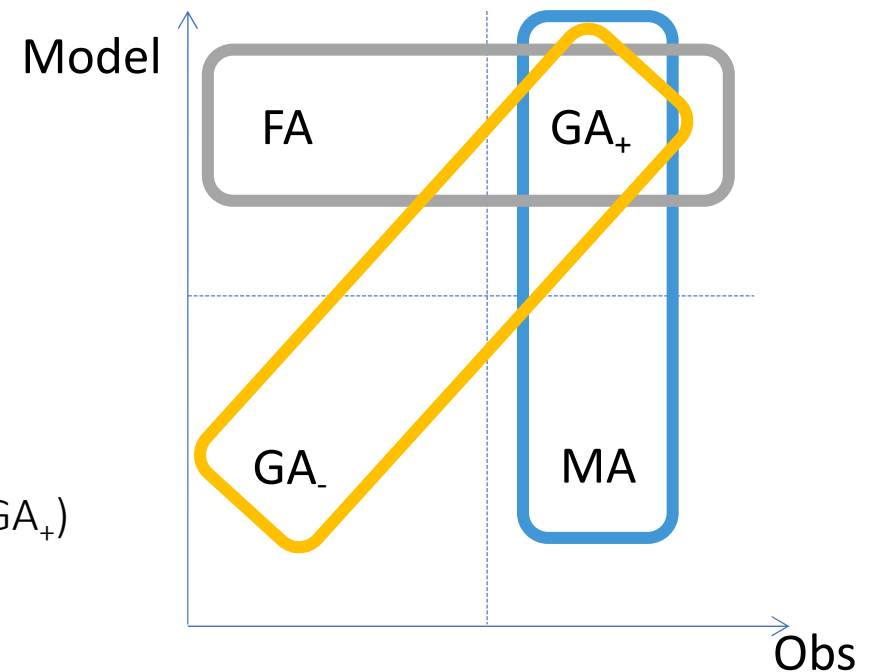
Threshold exceedances  
(as trigger for short term action plans)



use of standard threshold indicators

## Threshold exceedance indicators

- False Alarms (FA)
- Missed Alarms (MA)
- Good values below thr ( $GA_-$ )
- Good values above thr ( $GA_+$ )
  
- **Probability of detection**:  $PoD = GA_+ / (MA + GA_+)$
- **Success ratio**:  $SR = 1 - FAR = 1 - FA / (FA + GA_+) = GA_+ / (FA + GA_+)$
- **Accuracy**:  $ACC = (GA_+ + GA_-) / (GA_+ + GA_- + MA + FA)$



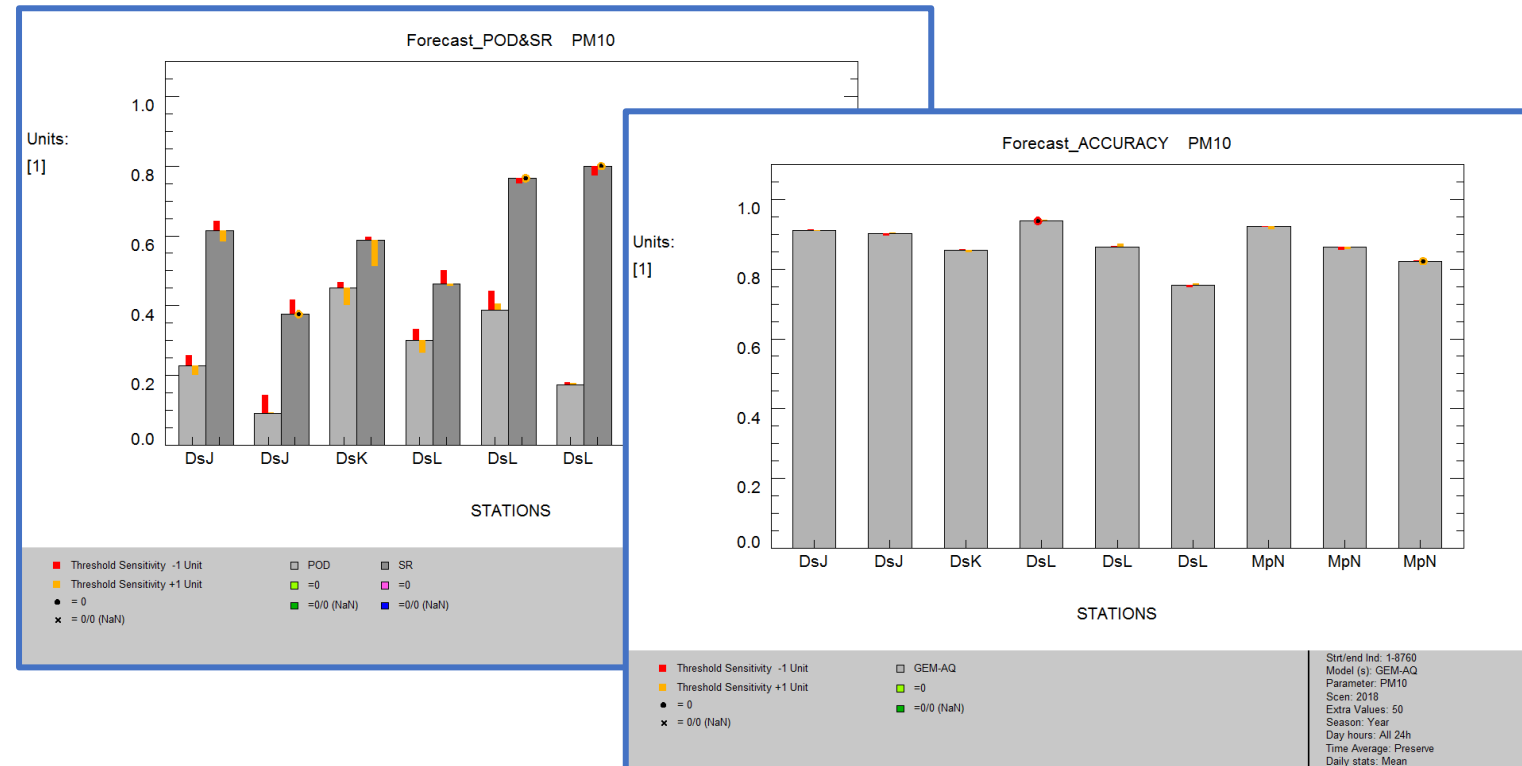
# FORECAST Modelling Quality Objective: DELTA tool proposal

Threshold exceedances  
(as trigger for short term action plans)



use of standard threshold indicators

- POD, SR & ACC = 1  
→ perfect model
- Sensitivity:
  - Red: threshold - 1unit
  - Yellow: threshold + 1unit

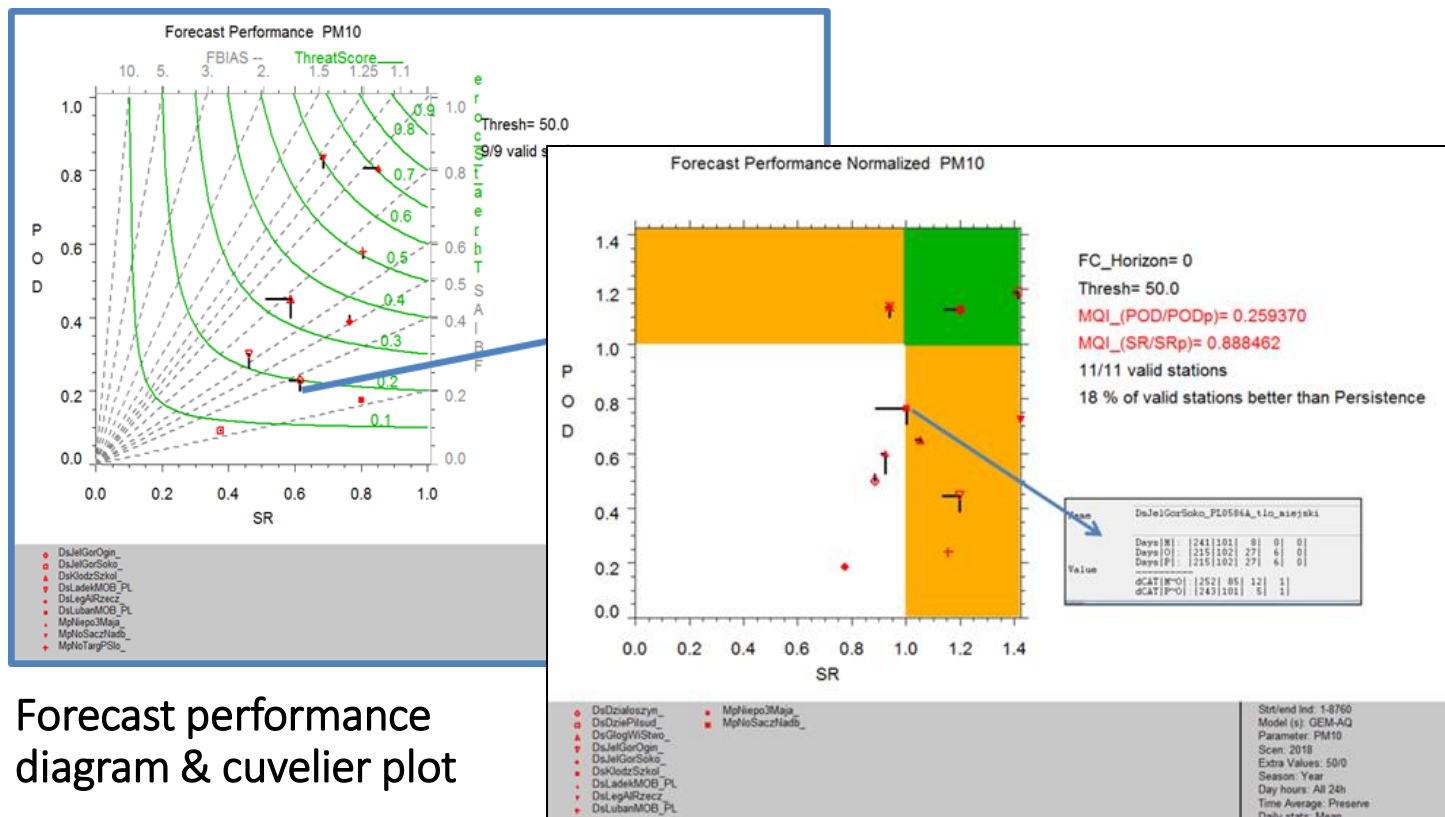


# FORECAST Modelling Quality Objective: DELTA tool proposal

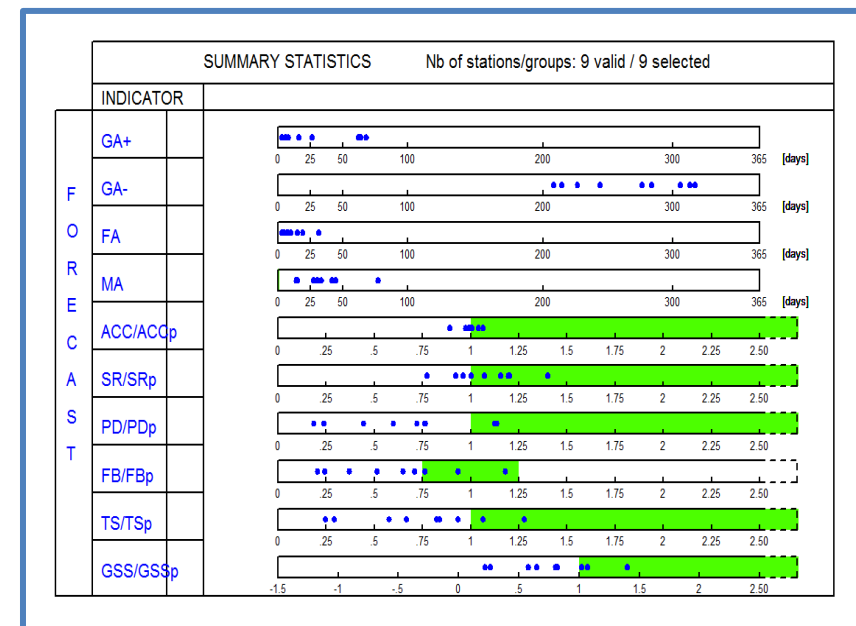
Threshold exceedances  
(as trigger for short term action plans)



use of standard threshold indicators



Forecast performance diagram & cuvelier plot



Summary report

# FORECAST Modelling Quality Objective: DELTA tool proposal

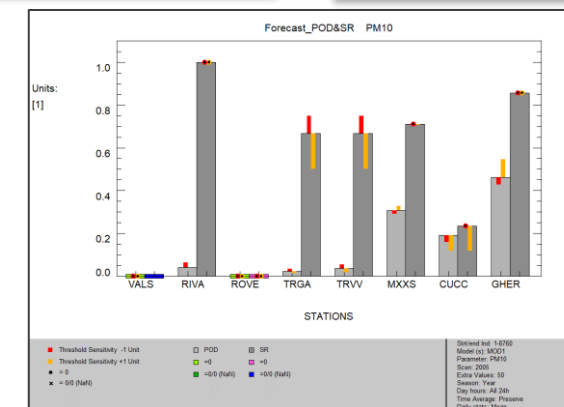
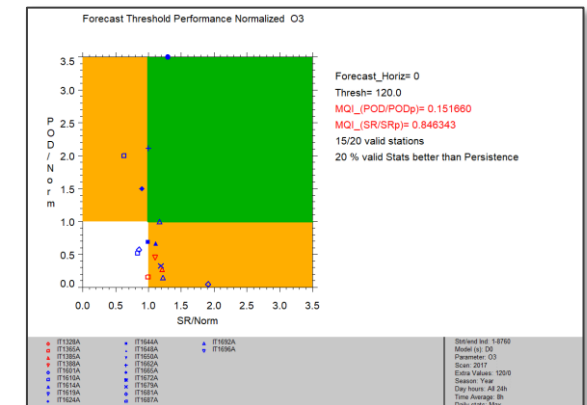
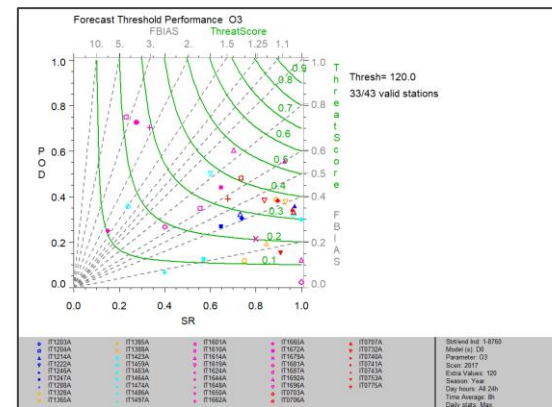
Threshold exceedances  
(as trigger for short term action plans)



use of standard threshold indicators

Comparison with Persistence Model:  
can be included by the user to better understand  
the results but it is not mandatory

It turned out that Persistence Model is very  
difficult to beat in predicting exceedances!



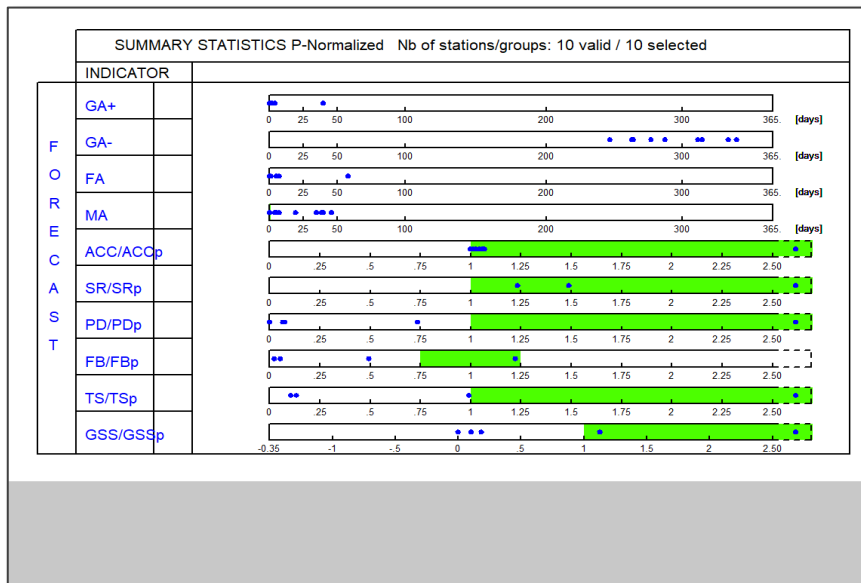
# FORECAST Modelling Quality Objective: DELTA tool proposal

Threshold exceedances  
(as trigger for short term action plans)

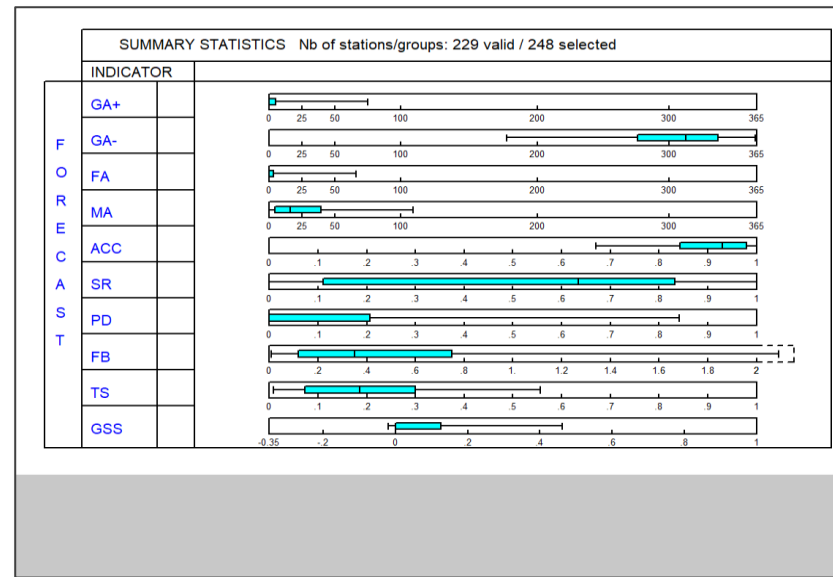


use of standard threshold indicators

Forecast Summary Report: a different graphical layout is applied depending on the n° stations

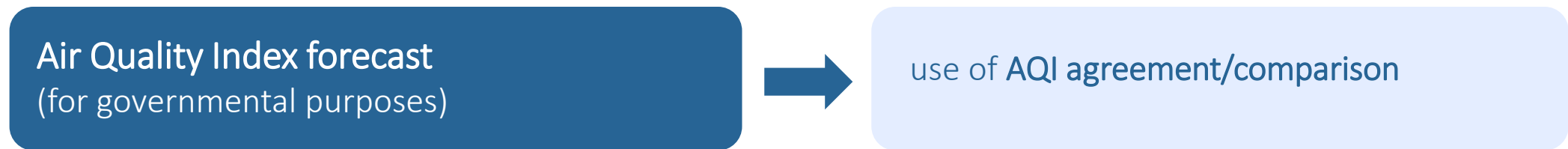


N° stations < 15 → “dots style”



N° stations ≥ 15 → “boxplots style”

# FORECAST Modelling Quality Objective: DELTA tool proposal



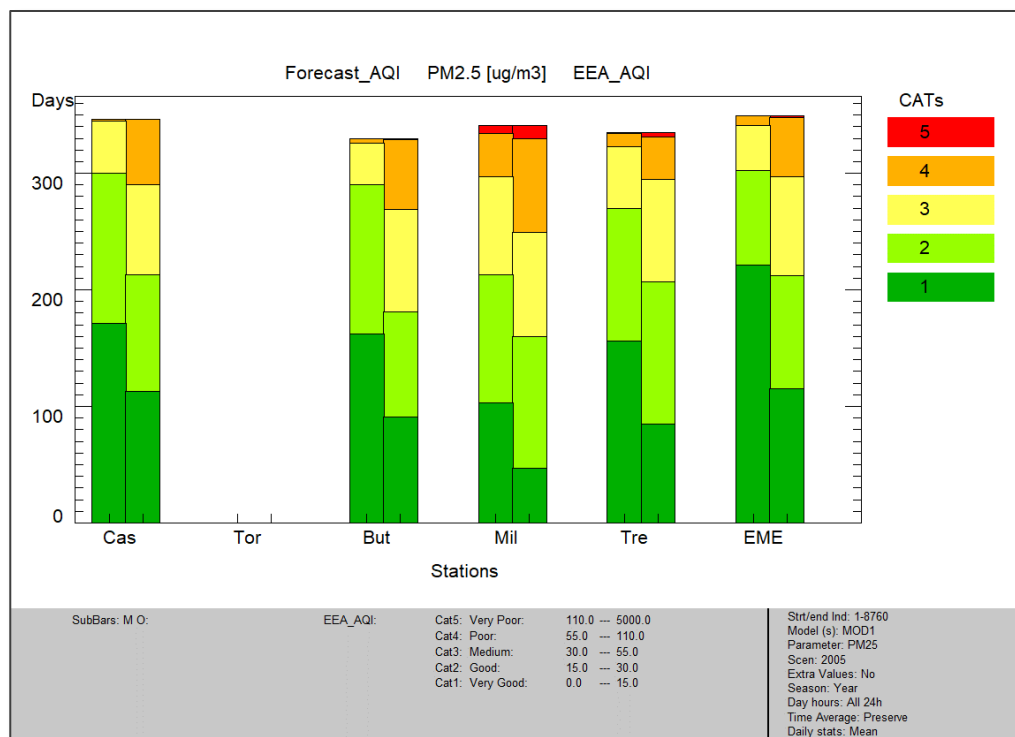


# FORECAST Modelling Quality Objective: DELTA tool proposal

Air Quality Index forecast  
(for governmental purposes)



use of AQI agreement/comparison



- In the current version the plot is produced **without the comparison with Persistence model (it was the same of observations column (shifted))**
- Information on total number of occurrences in each AQI class but **there is no information about the timing of the forecasted AQI levels**  
→ future improvement (Multi-category Contingency Table?)