FAIRMODE – Cross-cutting activities: Forecasting evaluations

F. Meleux
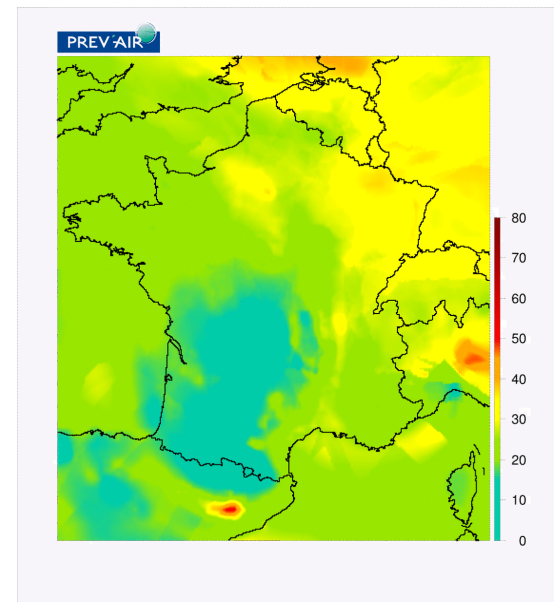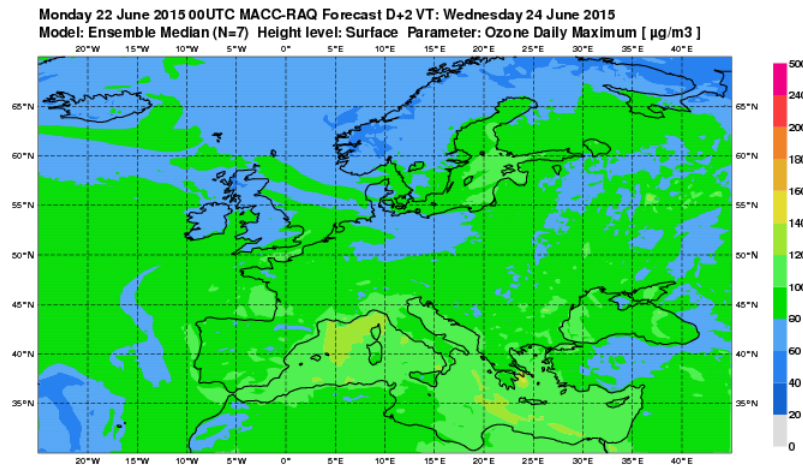
and VITO,BSC,CERC,JRC

INERIS

FAIRMODE
Forum for air quality modelling in Europe

INTRODUCTION

- Air quality platforms produce daily forecasts for the D+0, D+1 and D+…).
  - provide every day information related to the air quality levels
 - targetted pollutants: O3,NO2,PM10,PM2.5

OBJECTIVES

- What are the policy objectives for using AQ forecasts:
  - Predict & anticipate the development of a pollution episode
    - Inform and provide recommendations to the public
    - to identify the likely causes responsible for this episode to set-up efficient measures (short term action plans)

- What are the policy user needs to use AQ forecasts :
  - Get an assessment/knowledge about the capabilities of the forecasting system
    - How able are forecasts to detect / anticipate threshold exceedances ?
      - The triggering of measures is based on threshold values
    - How stable are the forecast scores from D+0 to D+n (usually n = 2 or 3)?
      - Policy measures are more efficient when they are taken earliest – so the goal is really to provide confident forecasts at least at D+1 (D+2 would be even better)

- ➢ Use the deltatool to provide evaluation of the capabilities of the Air quality forecasting system

FAIRMODE FORECAST TARGET

• To address these objectives, a specific diagram was developed in the delta-tool to focus on evaluation of forecasting system capabilities:

$$\text{Target}_{\text{forecast\_ji}} = \frac{\sqrt{\dfrac{1}{N}\sum_{i=1}^{N}\left(M_{ji}-O_i\right)^2}}{\sqrt{\dfrac{1}{N}\sum_{i=1}^{N}\left(O_{i-1}-O_i\right)^2}} \; with\_j = i-n,...,i$$
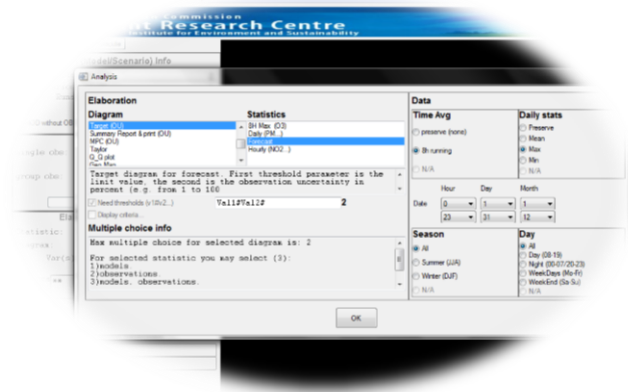
• Stating that the worst acceptable model is the persistent model, so at a given station the forecast (D+0) provides the observation of the eve (D-1).
• Assess the ability of the forecast to have a better daily variations than the persistent model even though the forecast is produced j days before
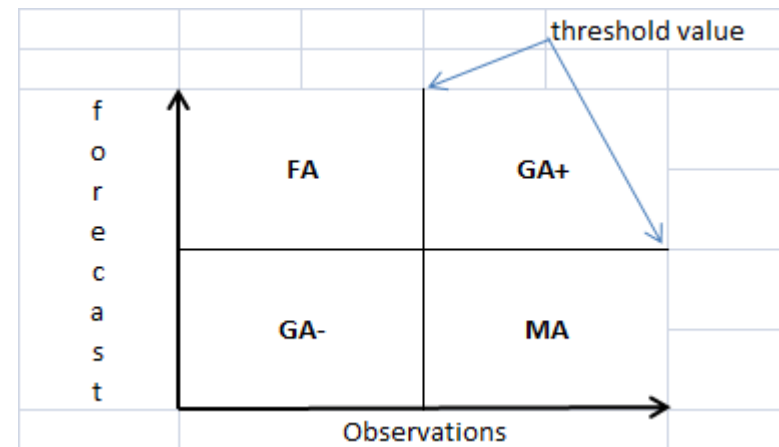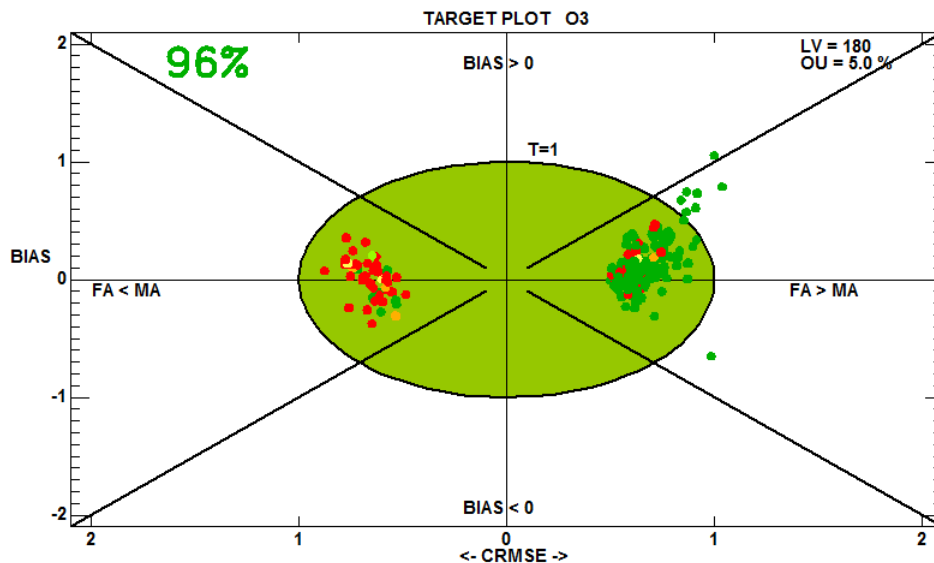
Forecast target diagram

- Keep part of indicators in the assessment diagram:
 - RMSE, BIAS, CRMSE
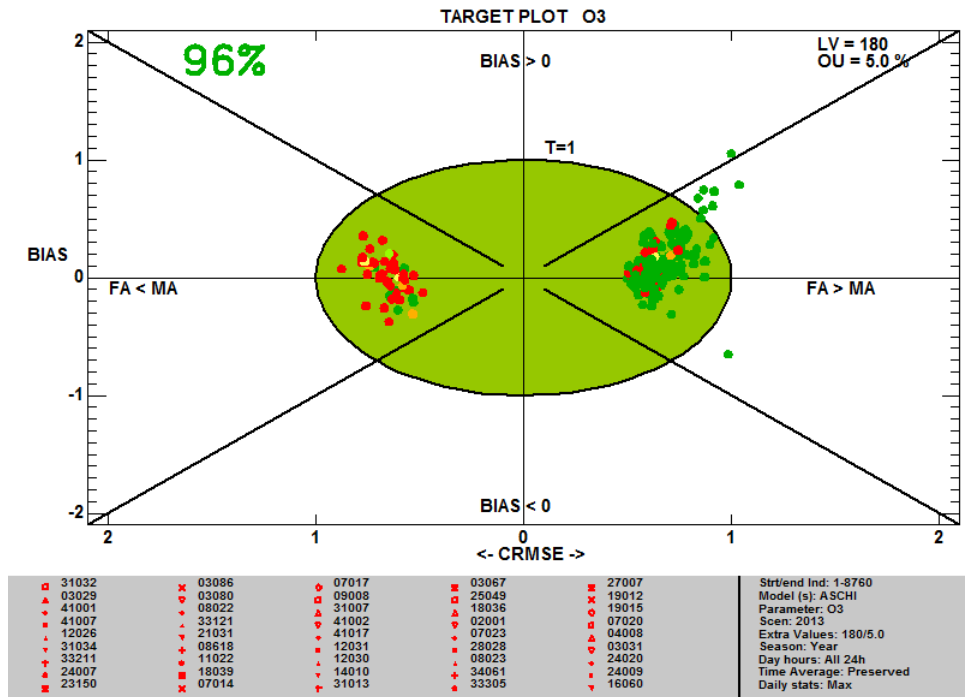
Forecast target diagram

- And to add new indicators more dedicated to the threshold exceedances:
  - FA (False alarm), MA (Missed alarm), GA+ and GA- (Good forecasts respectivily with and without threshold exceedances)

  - Ratio FA/MA   If $\dfrac{FA}{MA} \leq 1 \Rightarrow$ Left   If $\dfrac{FA}{MA} > 1 \Rightarrow$ Right

  - New suggestion to use the probability of good detection of threshold (FCF=GA$^+$/(GA$^+$+ MA)) and the rate of false alarm (FFA=FA/(GA$^+$+FA)). 2 formulations are proposed:
    - FCF-FFA (if > 0 means FA > MA and if <0 MA > FA).
    - FCF+FFA -1 (if > 0 means FA > MA and if <0 MA > FA).

Forecast target diagram

- And to add new indicators more dedicated to the threshold exceedances:
  - FA (False alarm), MA (Missed alarm), GA+ and GA- (Good forecasts respectivily with and without threshold exceedances)
    - $GA_+ / (FA + MA) \rightarrow GA_+ / (GA_+ + MA)$: probability of good forecast when exceedance occurs

TARGET PLOT O3
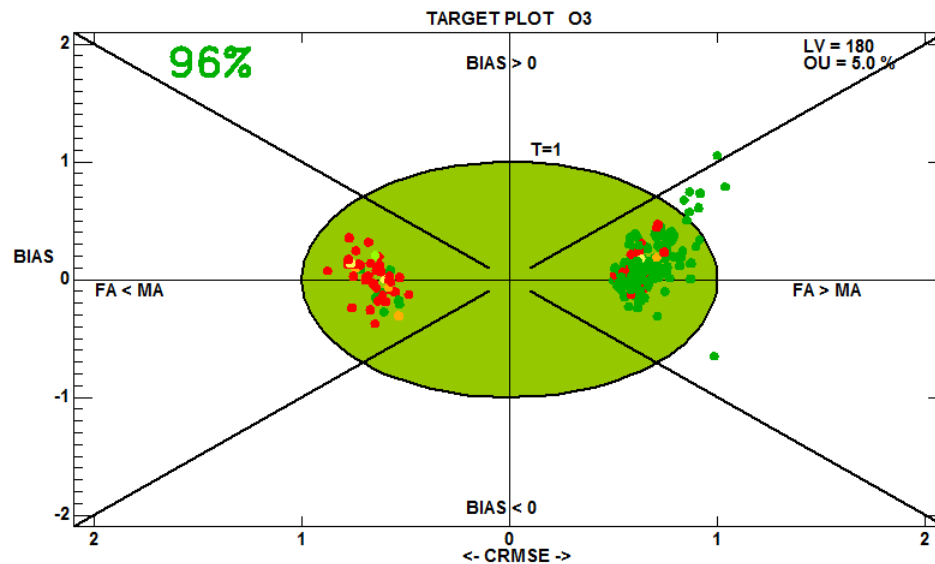
$$\frac{GA+}{FA+MA} < 0.2 \Rightarrow \text{Red}$$

$$0.2 \leq \frac{GA+}{FA+MA} < 0.4 \Rightarrow \text{Orange}$$

$$0.4 \leq \frac{GA+}{FA+MA} < 0.6 \Rightarrow \text{Yellow}$$

$$0.6 \leq \frac{GA+}{FA+MA} < 0.8 \Rightarrow \text{Light green}$$

$$0.8 \leq \frac{GA+}{FA+MA} \Rightarrow \text{Dark green}$$
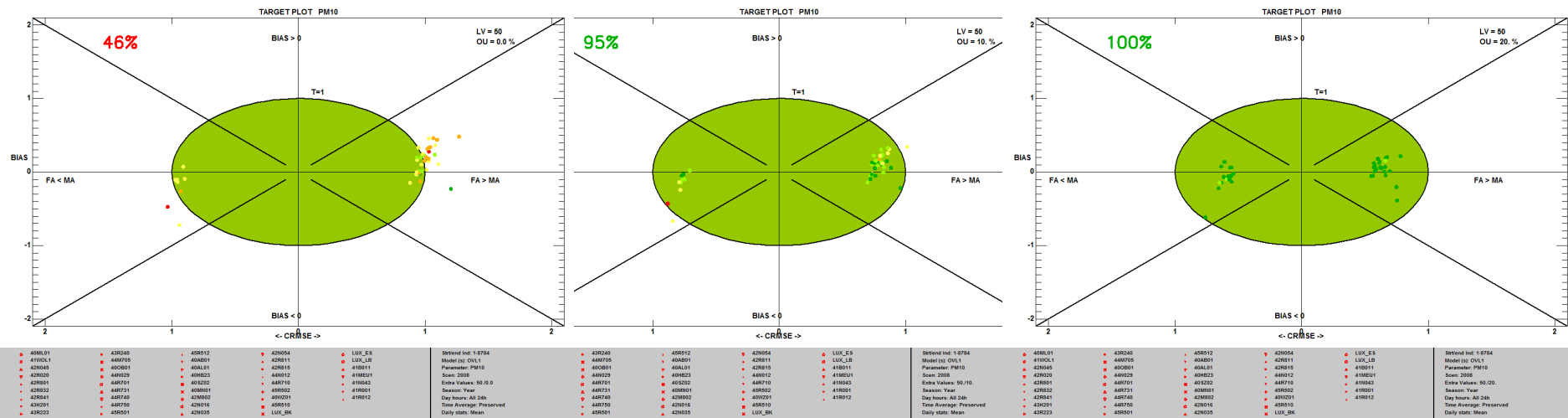
Forecast target diagram

- Introduction of an observation uncertainty (OU) attributed to the modeling outputs to take into account a margin of tolerance in the threshold exceedances:

$$\text{if } M_t < O_t \text{ then } M_t^* = \min(M_t * (1 + OU), O_t)$$

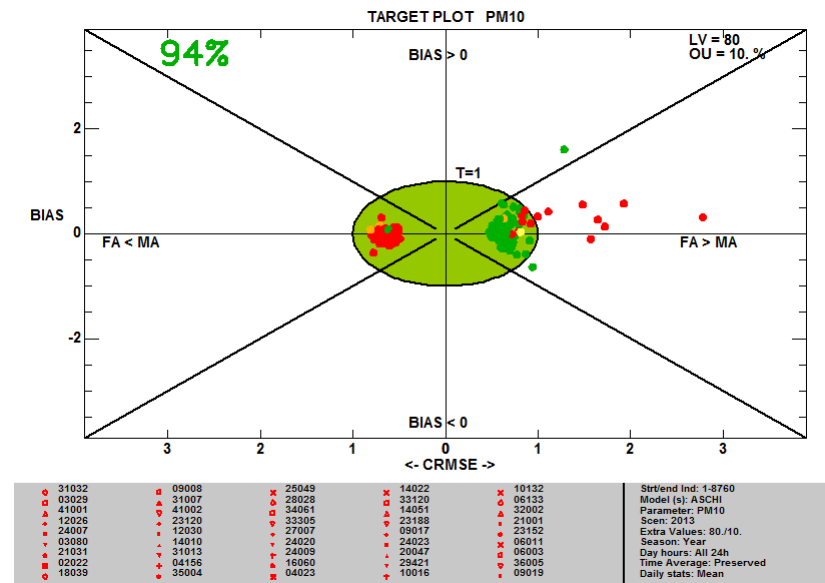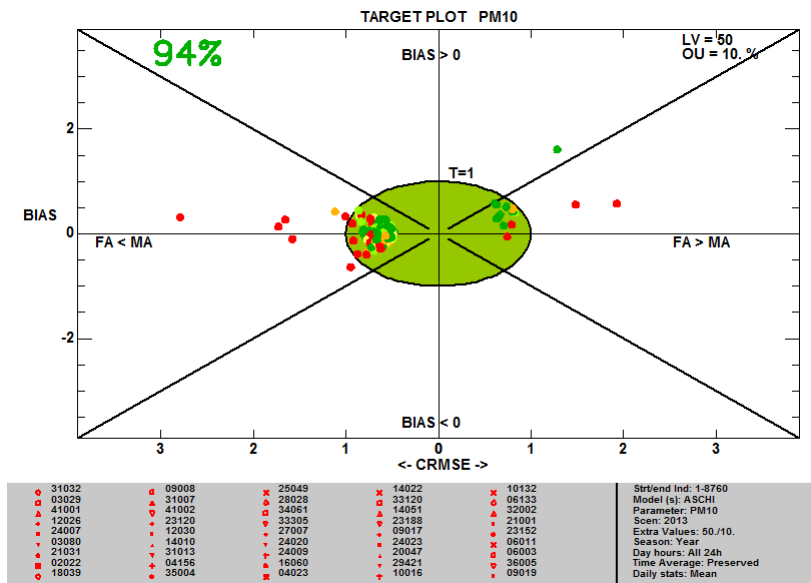$$\text{if } M_t \geq O_t \text{ then } M_t^* = \max(M_t * (1 - OU), O_t)$$

  - Model value becomes closer to the observations
  - OU= 5% for O3 , OU=10% for PM10

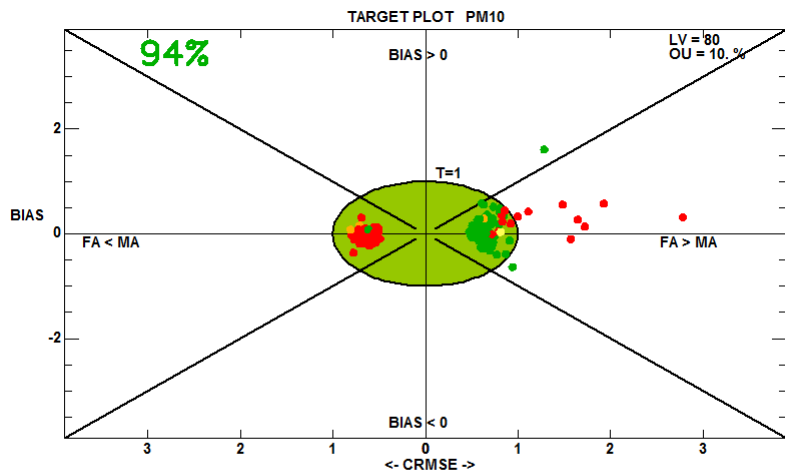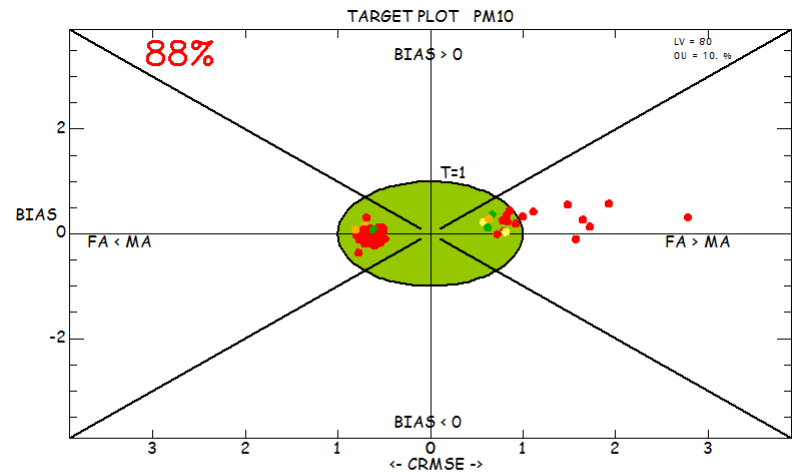- New suggestion(s) in the next presentations

Delta-tool developments for forecasting

- Data selection to compute scores when threshold exceedances occur
  - V4: all stations included
  - V5: Only the stations with at least one exceedances of the threshold either in the observation or in the forecast datasets
  - Suggestions for next version: all the station included but with white dots for stations without any exceedances

Delta-tool developments for forecasting

- Data selection to compute scores when threshold exceedances occur
  - V4: all stations included
  - V5: Only the stations with at least one exceedances of the threshold either in the observation or in the forecast datasets
  - Suggestions for next version: all the station included but with white dots for stations without any exceedances

Delta-tool developments for forecasting

- Extent the summary statistics report :
  - to include statistics about threshold exceedance detections – corresponding to the dot colors
  - to have a representation of the scores for multiple time-lags
    - Details about the consistency of the AQ forecast established for D+0, D+1, D+…
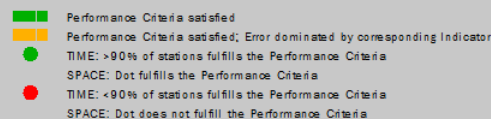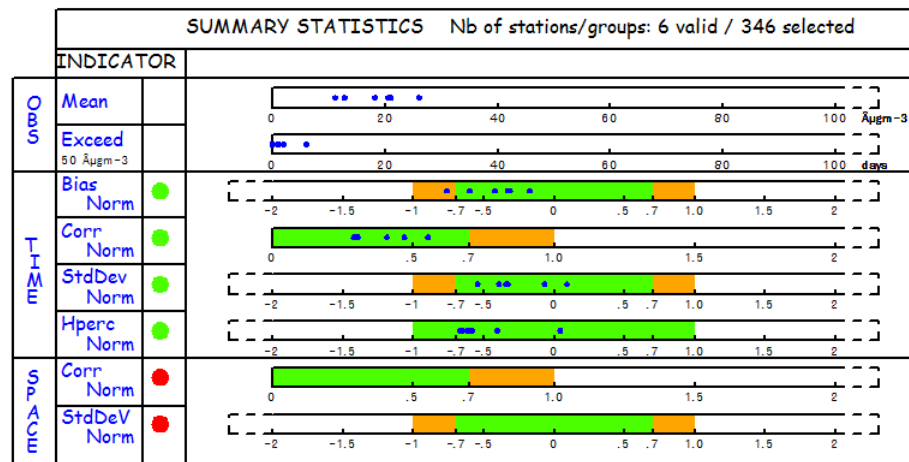
Evaluation of the PREV'AIR 2013 forecasts

- Modelling platform have two kinds of products:
  - Raw forecasts (directly produced by the model CHI)
  - Forecasts with corrective post-processing (ASCHI:statistical correction of CHI based on the past errors of the models)
- For O_3 : performances are in good agreement
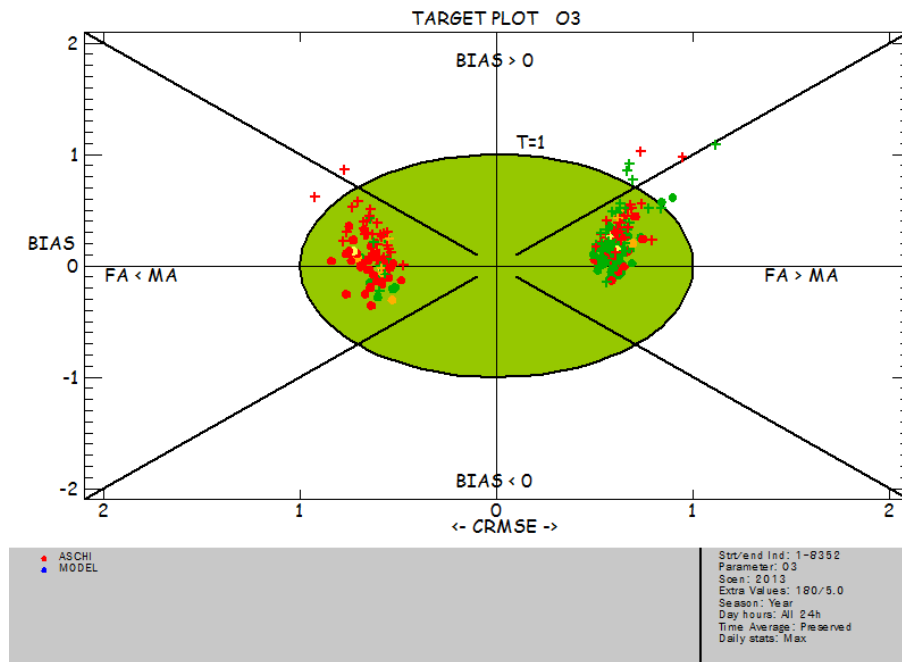
Evaluation of the PREV'AIR 2013 forecasts

- Modelling platform have two kinds of products:
  - Raw forecasts (directly produced by the model CHI)
  - Forecasts with corrective post-processing (ASCHI:statistical correction of CHI based on the past errors of the models)
- For PM$_{10}$: performances are quite different:
  - Good for ASCHI
  - Poor for CHI
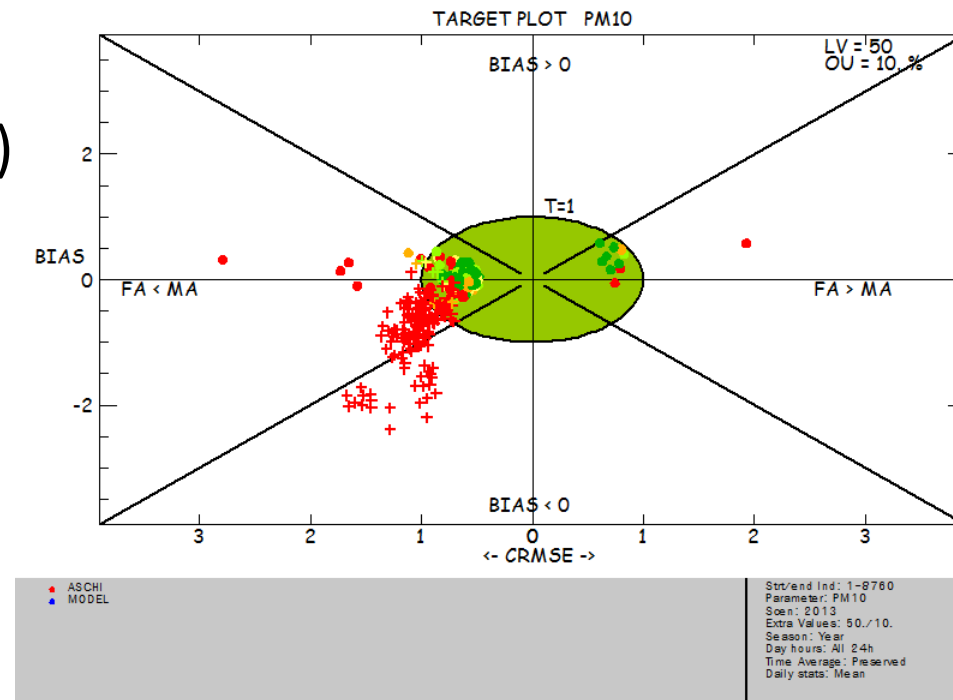
Conclusions

- First version of the target forecast diagram looks promising but some improvements still need to be done:
  - define our final choice about the indicators to represent in the deltatool
  - implement statistics about the threshold detection capabilities in the summary statistics report.
- Assessment of the consistency/accuracy of the whole forecasting system (global assessment for D+0,D+1, D+... and all thresholds)
- Continue to test
  - Additional tests are foreseen with the large European database from MACC projects (not yet done due to issues in the MACC NRT observation dataset for 2015)