



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



universidade de aveiro
theoria poiesis praxis



Overview of the available Diagrams for Forecast validation

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

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



1101 0110 1100
0101 0010 1101
0001 0110 1110
1101 0010 1101
1111 1010 0000



3 goals and 6 plots

Comparison with the Persistence Model	Capability in predicting Exceedances	Capability in predicting Air Quality Indices
Forecast Target Plot	Forecast Threshold Performance Plot (absolute or P-normalized)	Barplots of AQI
Forecast MPI Plot	Barplots of POD, SR, POD&SR, ACC	
	Forecast Summary Report (absolute or P-normalized)	

- 1 slide per type of diagram
- Reference → pages 28 and 74-81 of the Delta Tool User's Guide:
<https://aqm.jrc.ec.europa.eu/section/assessment/PU/pdf/DELTA-UserGuide-V7>

Before starting: GA+, GA-, FA, MA, ACC, SR, POD, FBIAS, TS, GSS

False Alarms (**FA**): Model values are above the threshold but not observations

- $M > \text{threshold}$ and $O \leq \text{threshold}$

Missed Alarms (**MA**): Model values are below the threshold but observed values are above it

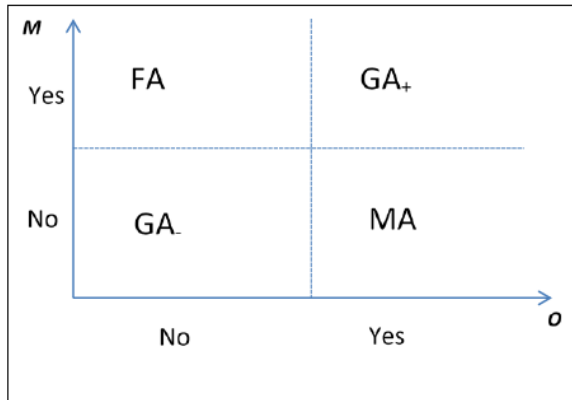
- $M \leq \text{threshold}$ and $O > \text{threshold}$

Good Alarms (low) below threshold (**GA-**): both model and observation are below or equal to the threshold.

- M and $O \leq \text{threshold}$

Good Alarms (high) above threshold (**GA+**): both model and observations are above the threshold.

- M and $O > \text{threshold}$



For a good forecast both FA and MA are small compared to GA+ and GA-. Based on these quantities the following indicators can be calculated:

- Probability of Detection: $POD = GA_+ / (MA + GA_+)$
- Success Ratio: $SR = 1 - \text{False Alarm Ratio} = 1 - FA / (FA + GA_+) = GA_+ / (FA + GA_+)$

The POD indicator is comparing the correct modelled alerts with the **observed** alerts whereas the SR indicator is comparing the correct modelled alerts with all alerts **issued** by the model.

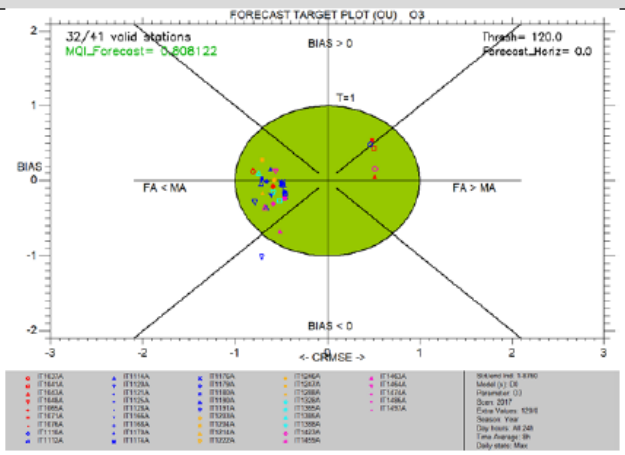
We also define four additional indicators as:

- FBias score: $FBIAS = (GA_+ + FA) / (MA + GA_+)$
- Accuracy: $ACC = (GA_+ + GA_-) / \text{Total}$
- Threat score: $TS = GA_+ / (MA + FA + GA_+) = GA_+ / (FA + CA)$
- Gilbert Skill score: $GSS = (GA_+ - H_{\text{random}}) / (MA + FA + GA_+ - H_{\text{random}})$
with $H_{\text{random}} = (GA_+ + MA)(GA_+ + FA) / \text{Total}$

1. versus Persistence: Forecast Target Plot – page 74

74

FORECAST TARGET PLOT



X axis	CRMSE / RMSE _p
Y axis	BIAS / RMSE _p
Parameters	NO2, O3, PM
Time Avg	See MQO below
Daily Stats	
Season	
Day	
Threshold	Used for exceedance calculation

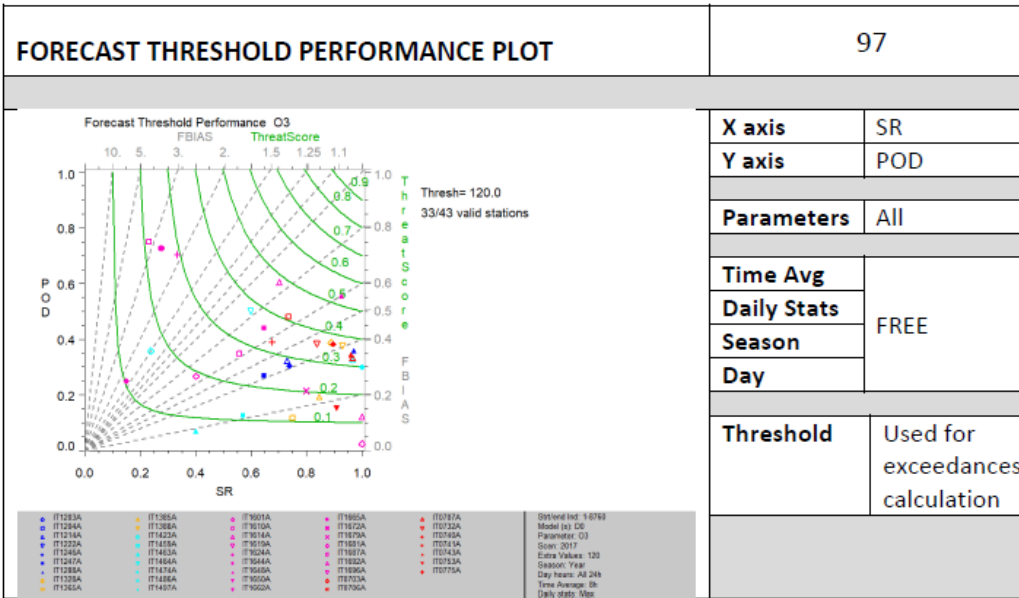
Description

In the Forecast target plot information is included on the following quantities (all normalized by the root mean squared error of the persistence model):

- RMSE: distance from the origin to the point
- BIAS: the bias can be either positive or negative and is represented along the vertical axis (Y)
- CRMSE: The CRMSE is always positive and given by the distance from the origin to the point along the X axis.
- False Alarm (FA) vs. Missed Alarm (MA): we use the FA/MA ratio to differentiate the right and left parts of the target diagram: $\frac{FA}{MA} < 1 \rightarrow \text{Left}$; $\frac{FA}{MA} \geq 1 \rightarrow \text{Right}$

- main diagram
- Values lower than one (points within the green circle) indicate better capabilities than the persistent model
- Values larger than one indicate poorer performances
- The MQI value corresponding to the 90th largest percentile is printed in the upper left corner and should be lower than 1

2. on Exceedances: Forecast Threshold Performance Plot – page 76

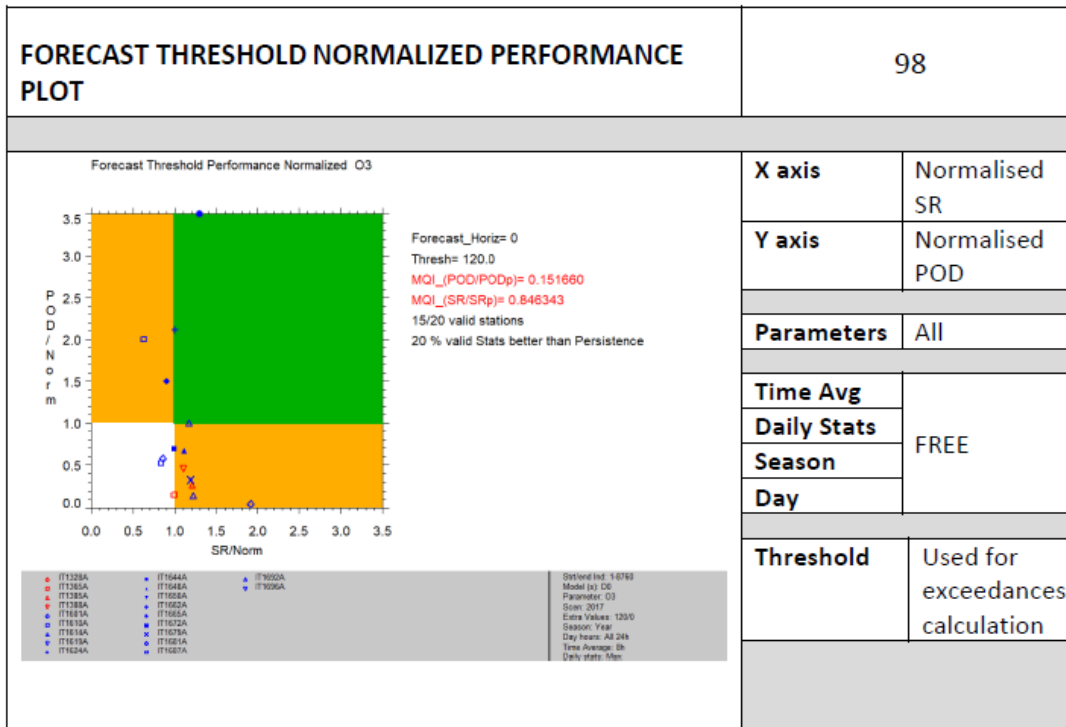


- Good forecasts with a high POD and SR are situated in the upper right corner

The Forecast Threshold Performance Plot shows the four forecast indicators POD (Probability of Detection), SR (Success Ratio), Fbias (FBias score) and TS (Threat score). It is based on the SR values on the X axis and POD values on the Y axis. Since Fbias and TS are indicators related to POD and SR, they are represented by additional isolines.

Good forecasts with a high POD and SR are situated in the upper right corner.

2. on Exceedances and vs Persistence: Forecast Threshold Normalized Performance Plot – page 77



- the green area represents forecasts with better POD and SR threshold indicators than the persistence model
- in the white zone, the model performs worse than the persistence model on both indicators
- In the orange zone, one of the two indicators is better than the benchmark.
- The normalized POD and SR values (i.e. POD/PODp and SR/SRp) are also given as indicative Modelling Performance Indicators. These MPis correspond to the 10th largest percentile value and should be larger than one for a “good enough” forecast.

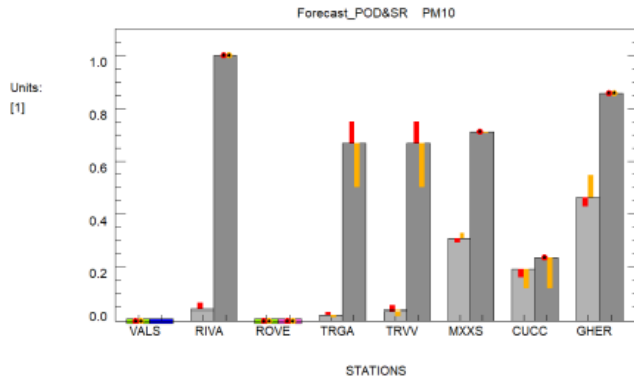
Description

In order to indicate whether a forecast model is “good enough”, in the Forecast Threshold Normalized Performance Plot the POD and SR values obtained with the persistence model are used as a benchmark (i.e. POD and SR of the forecast model are normalized with POD and SR of the persistence model).

2. on Exceedances: Barplots of POD, SR, POD&SR, ACC – page 78

BARPLOT FOR EXCEEDANCE INDICATORS (POD, SR, POD&SR, ACCURACY)

89, 90, 91, 92



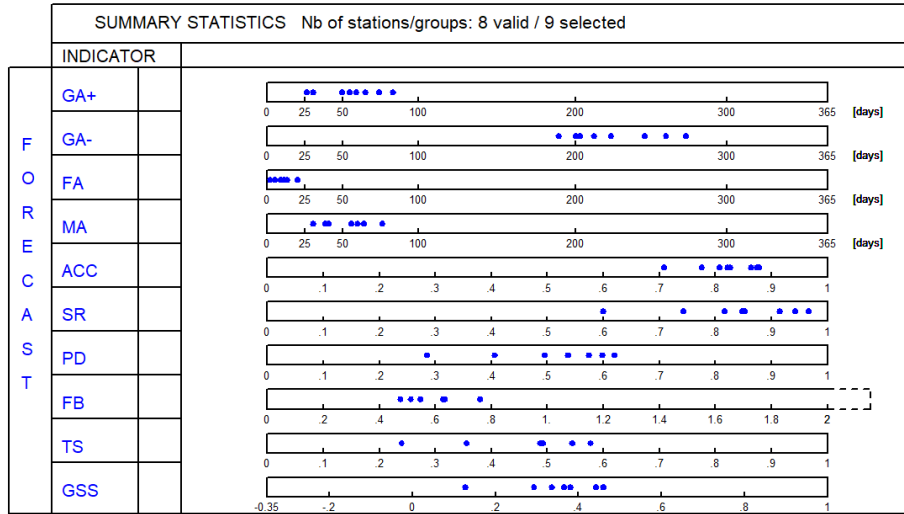
X axis	Stations
Y axis	POD, SR, POD&SR, ACCURACY
Parameters	All
Time Avg	FREE
Daily Stats	
Season	
Day	
Threshold	Used for exceedances calculation

- The red bar indicates the change in the indicator when the threshold is reduced by 1 unit
- the yellow bar the change when the threshold is increased by 1 unit

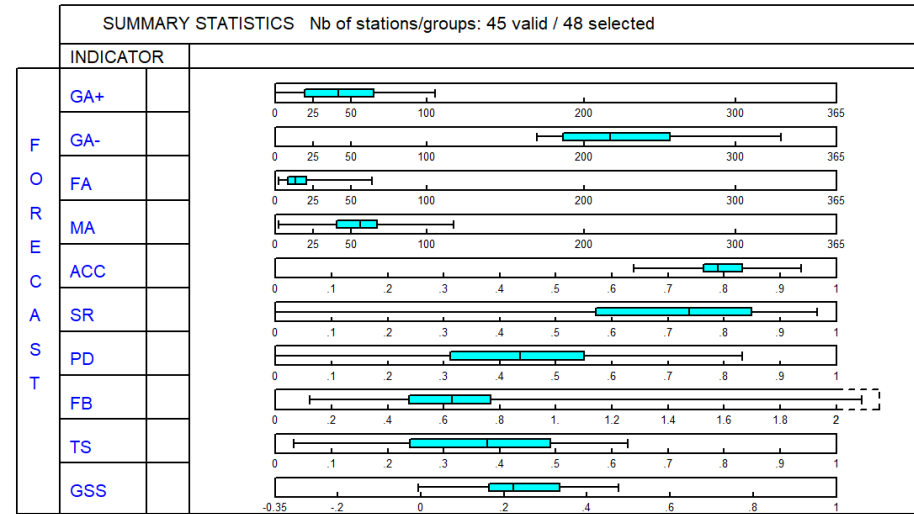
Description

Bar Plots show the values of the individual exceedances indicators (POD, SR, POD&SR, ACCURACY), together with their negative and positive sensitivities with respect to the threshold. The red bar indicates the change in the indicator when the threshold is reduced by 1 unit, the yellow bar the change when the threshold is increased by 1 unit.

2. on Exceedances: Forecast Summary Report – page 79

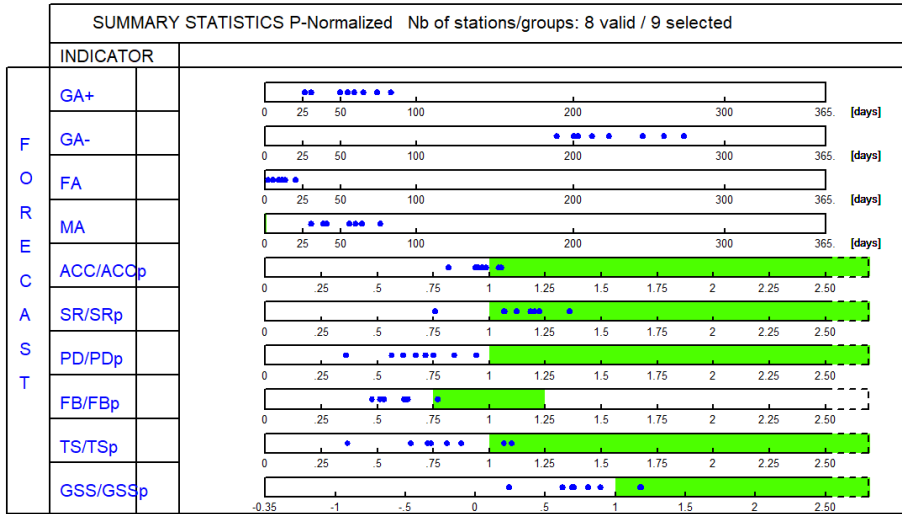


Points: < 15 stations

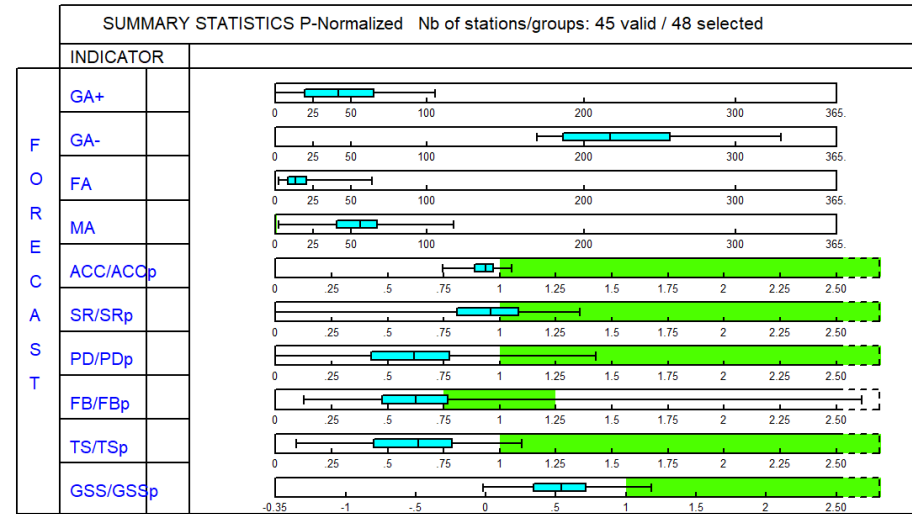


Bars: > 15 stations

2. on Exceedances and vs Persistence : Forecast Summary P-Normalized Report – page 80



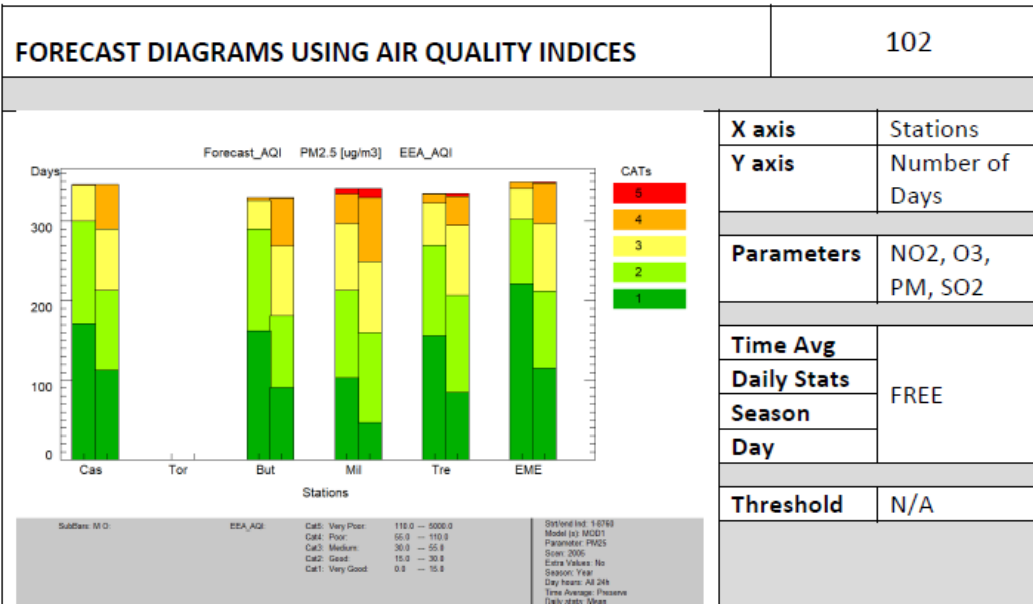
Points: < 15 stations



Bars: > 15 stations

the “good enough” zone is shaded in green and indicates that the model performs better than the persistence model for this particular indicator

3. on AQ Indices: Barplots of AQI – page 81



Description

This plot is based on multiple thresholds as they appear in the Air Quality Categories and their Indices, like EEA, UK or US EPA indicators. In this diagram we compare the number of days that the forecast model (M) and the Measurements (O) have in common in each of the Air Quality Categories.

The Index table itself is shown in the grey area below, the corresponding colours on the right-hand side of the graphic.

In the current version of the DELTA Tool the following AQI tables are available: EEA (5 indices), UK4 (4 indices), UK10 (10 indices), USEPA (7 indices), and can be selected in the aqibounds.dat file in the DELTA tool configuration folder.