European Commission

FAIRMODE
Forum for air quality modelling in Europe

WG3 (Forecast) update

WORK DONE & PLAN TILL END OF YEAR

Antonio Piersanti, Alexandra Monteiro, Lina Vitali

Prague (CZ), March 5-6, 2025

AGENDA

- **Status:** how the Working Group contributed to better modelling for assessment (and planning)

- **Future steps** - Plan till end of year & discussion on priorities for the next Roadmap

STATUS

Fairmode was invited to talk about forecast evaluation by CAMS Global colleagues

Performance indicators and headline scores for atmospheric composition and air quality forecasting

Online Workshop | Wednesday 19/02/2025 | 13:30-17:00

13:30 - 13:40 | Henk Eskes (KNMI)
Welcome and introduction to the workshop

13:40 - 13:50 | Johannes Flemming (ECWMF)
Developing best practices for the evaluation of Air Quality Forecasting System - a goal of the GAFIS initiative at GAW/WMO

13:50 - 14:10 | Athanasios Tsikerdekis, Henk Eskes (KNMI), Michael Schulz (MET-Norway)
Headline scoring to assess the performance of CAMS products

14:10 - 14:30 | Nick Schutgens (VU)
Evaluation of AEROCOM models with satellite data: lessons learned

14:30 - 15:00 | Dragana Kornic (ECCC)
AQPI and multi-model forecast evaluation for North America

Coffee/Tea break (15:00 - 15:30)

15:30 - 16:00 | Thomas Haiden (ECWMF)
Headline scores in NWP: benefits and challenges

16:00 - 16:30 | Antonio Piersanti (ENEA) and Alexandra Monteiro (UA)
Model quality indicators for air quality forecast: FAIRMODE approach

16:30 - 17:00 | Moderator: Johannes Flemming
Discussion

ENEA universidade de aveiro theoria poiesis praxis European Commission FAIRMODE Forum for air quality modelling in Europe

STATUS

Some open issues were presented and discussed at the last Plenary Meeting (Paris, February 26-27, 2024), according to the three main WG3 topics

- ✓ Concerning the Comparison with the Persistence Model: *how does adopting AQUILA-based parameters impact on Forecast MQI outcomes?*
- ✓ Concerning the Capability in predicting Exceedances: *What could be the target for predictability of episodes? How could objective criteria for threshold exceedance's indicators be defined?*
- ☐ Concerning the Capability in predicting Air Quality Indices: *How could the correct timing be included within AQI evaluation?*

✓ Work mainly done last year before last Technical Meeting (Dublin, October 2024)

✓ Ongoing work after last Technical Meeting

ENEA universidade de aveiro theoria poiesis praxis European Commission FAIRMODE Forum for air quality modelling in Europe

STATUS

1. COMPARISON WITH THE PERSISTENCE MODEL

MAIN ISSUE: *How does adopting AQUILA-based parameters impact on Forecast MQI outcomes?*

WORK DONE: An exercise was launched during 2024, aimed at testing the effect of adopting the new parameters proposed by AQUILA for measurement uncertainty estimates.
Results and feedback of the exercise were shared and discussed at a dedicated hackathon and during last Technical Meeting (Dublin, October 2024).

Contribution from *Paweł Durka, Aleksander Norowski IEP-NRI (Poland); Loris Colombo ARPA Lombardia (Italy); Alexandra Monteiro, Carla Gama UniAveiro (Portugal); Eivind G. Wærsted, Bruce R. Denby MET Norway (Norway)*

MAIN FINDINGS:

- criteria get more stringent (outcomes get worse) in most of the cases
- since the stringency factor (β value) is not included within Forecast MQI formulation there is no "control knob" to compensate the effect of changing uncertainty parameters
- **adopting AQUILA-based parameters impacts on Assessment MQI and Forecast MQI outcomes very differently**

STATUS

1. COMPARISON WITH THE PERSISTENCE MODEL

NEW OPEN ISSUES

(minor) → *Should we introduce β within Forecast MQI formulation? Ideas for setting β values?*

(major) → *Should we persist with the Persistence Model comparison?*
Should we look at the comparison with the Persistence Model in a more comprehensive way?

→ Feedback is needed from users on the actual use of Forecast MQI
(see next steps)

STATUS

2. CAPABILITY IN PREDICTING EXCEEDANCES

When a forecasting system is used for policy purposes, it is of ***utmost importance to verify its skill in predicting categorical answers (yes/no) in relation to exceedances of specific threshold levels***

Directive (Eu) 2024/2881 of the European Parliament and of the Council on Ambient Air Quality and Cleaner Air for Europe (adopted on 11 December 2024)

Article 15 - Exceedances of alert thresholds or information thresholds:

…Where any alert threshold laid down in Section 4, Point A, of Annex I is exceeded, or, where appropriate, if it is predicted to be exceeded based on modelling applications or other forecasting tools, Member States shall, where applicable, implement without undue delay the emergency measures indicated in the short-term action plans established pursuant to Article 20…

…Where any alert threshold or any information threshold laid down in Section 4 of Annex I is exceeded, or, where appropriate, if it is predicted to be exceeded based on modelling applications or other forecasting tools, Member States shall take the necessary steps to inform the public within the shortest possible timeframe and as far as possible within a few hours, in accordance with points 2 and 3 of Annex X, making use of different media and communication channels and ensuring broad public access.

STATUS

2. CAPABILITY IN PREDICTING EXCEEDANCES

MAIN ISSUES: *What could be the target for predictability of episodes?*
How could objective criteria for threshold exceedance's indicators be defined?

ONGOING WORK (in order to define objective criteria)

A. Based on the performance of a state-of-the-art air quality forecasting system, used at national level for policy purposes, can we fix “typical” (e.g. more frequent) values for exceedance's indicators outcomes?

B. Could CAMS Ensemble skills be a suitable “quality reference” for evaluating European forecast modeling systems applications?

STATUS

2. CAPABILITY IN PREDICTING EXCEEDANCES

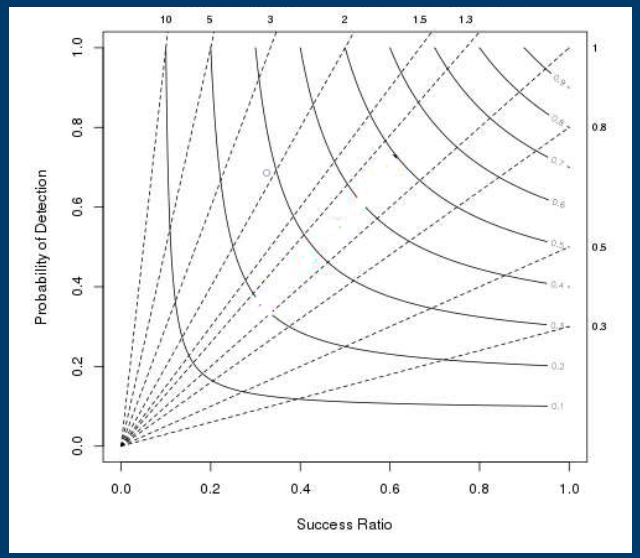
CONTINGENCY TABLE

Forecast events	Yes	FA	GA+
	No	GA−	MA
Contingency table		No	Yes
		Observed events	

CATEGORICAL METRICS

Probability of detection	$PD = \frac{GA_+}{GA_+ + MA}$
Success ratio	$SR = \frac{GA_+}{FA + GA_+}$
Accuracy	$ACC = \frac{GA_+ + GA_-}{FA + GA_+ + GA_- + MA}$

PERFORMANCE DIAGRAM



too few exceedances are available for NO_2, not enough for the statistical analysis

POLLUTANTS

~~NO_2~~ O_3

PM10 ~~PM2.5~~

no hourly or daily limit value or target value are set for PM2.5 within AAQD 2008

CATEGORICAL METRICS

POD	to evaluate the capability of detecting exceedances
SR	to evaluate the capability of avoiding false alarms
ACC	to evaluate the frequency of true answers (GA_- + GA_+)

STATUS

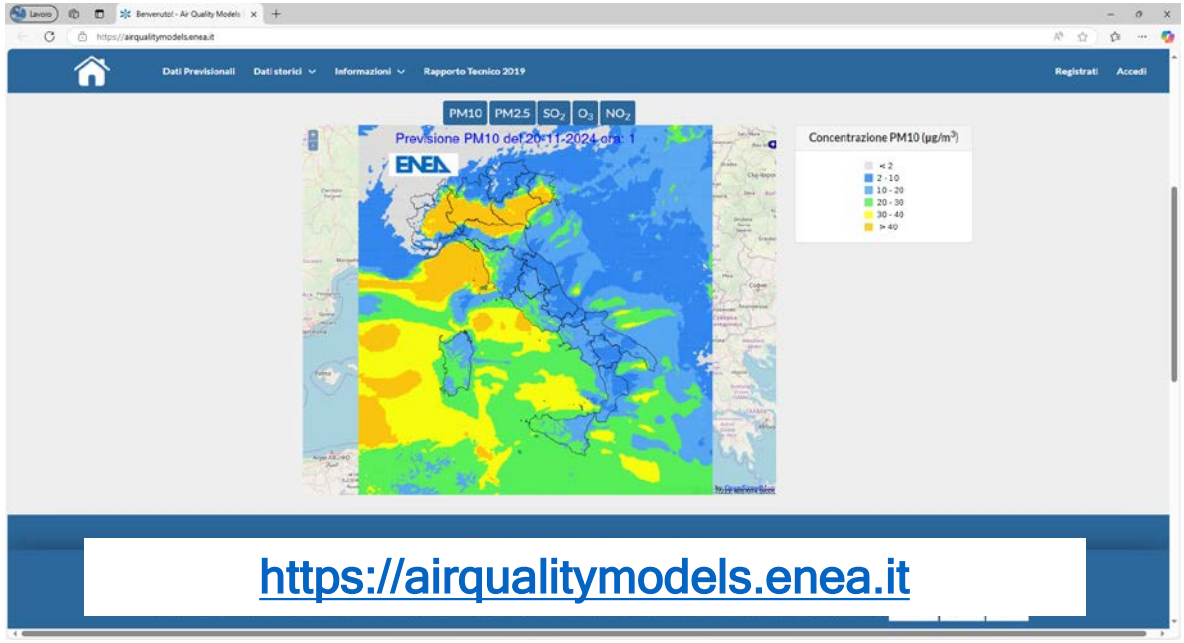
2. CAPABILITY IN PREDICTING EXCEEDANCES

DATA SET USED FOR THE EVALUATION

FORAIR-IT - National Italian Forecast System

hourly outcomes for year 2022

Forecast horizon: 3 days



https://airqualitymodels.enea.it

MODELLING SYSTEM OPERATIONAL MAINTENANCE

Massimo D'Isidoro (ENEA)

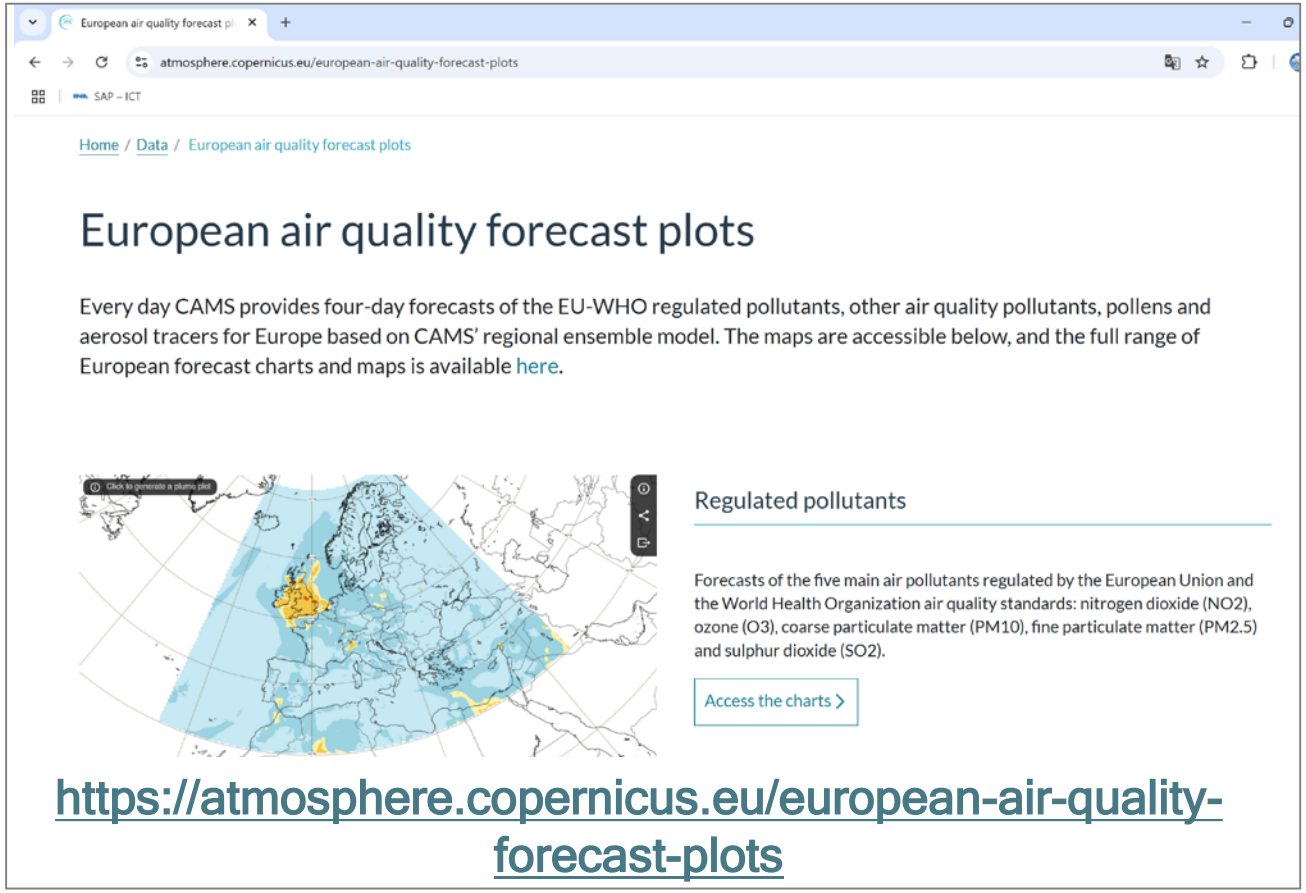
VALIDATION DATA BASE SETTING UP

Maria Gabriella Villani (ENEA)

CAMS Regional - European Air Quality Forecast

ENSEMBLE hourly outcomes for year 2021

Forecast horizon: 4 days



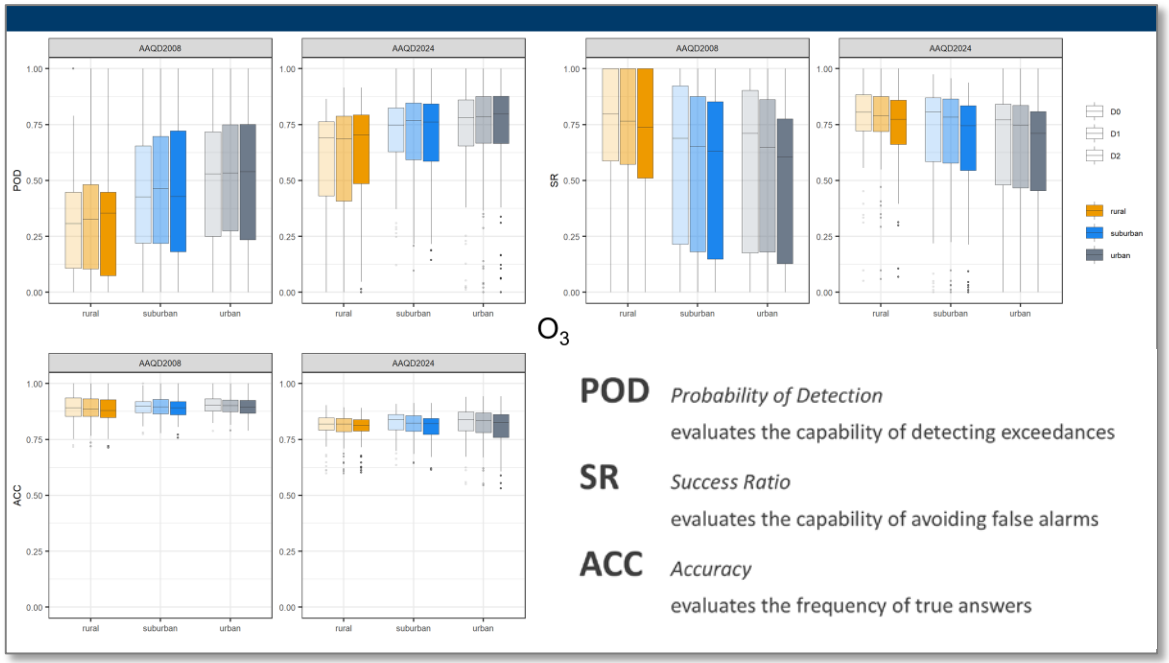
https://atmosphere.copernicus.eu/european-air-quality-forecast-plots

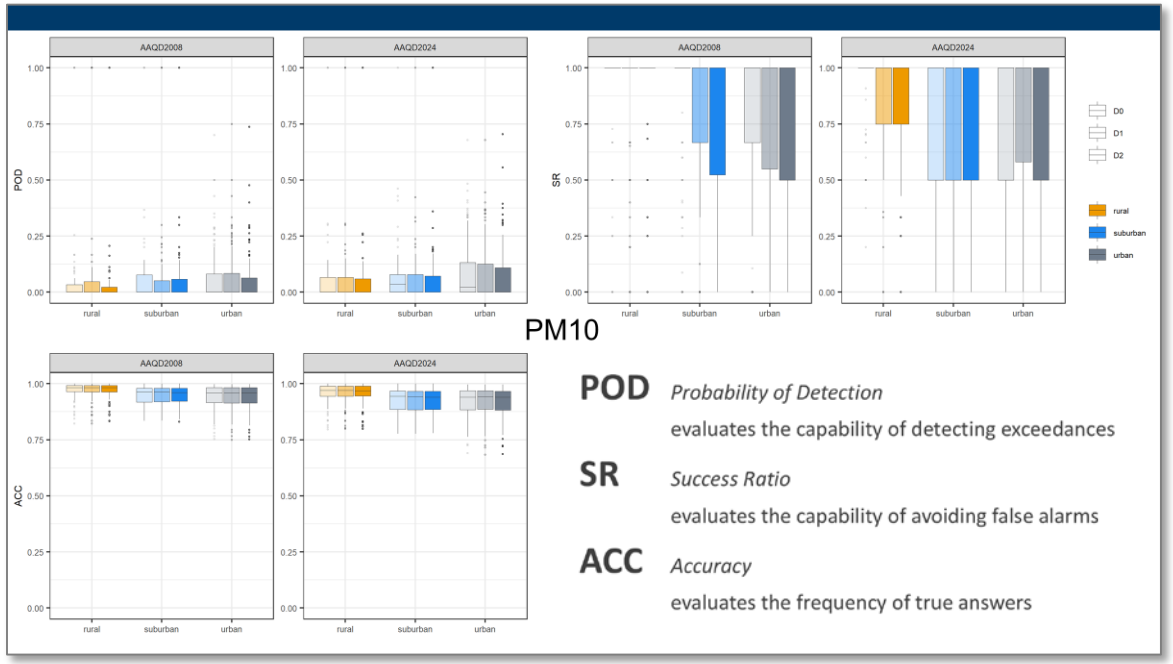
STATUS

2. CAPABILITY IN PREDICTING EXCEEDANCES

A. Based on the performance of **a state-of-the-art air quality forecasting system**, used at national level for policy purposes, can we fix “typical” (e.g. more frequent) values for exceedance's indicators outcomes?

FORAIR-IT





- ✓ SR distribution is very large ranging from 0 to 1, and concerning O_3 the same happens for POD
- ✓ Tighter distribution are shown for ACC

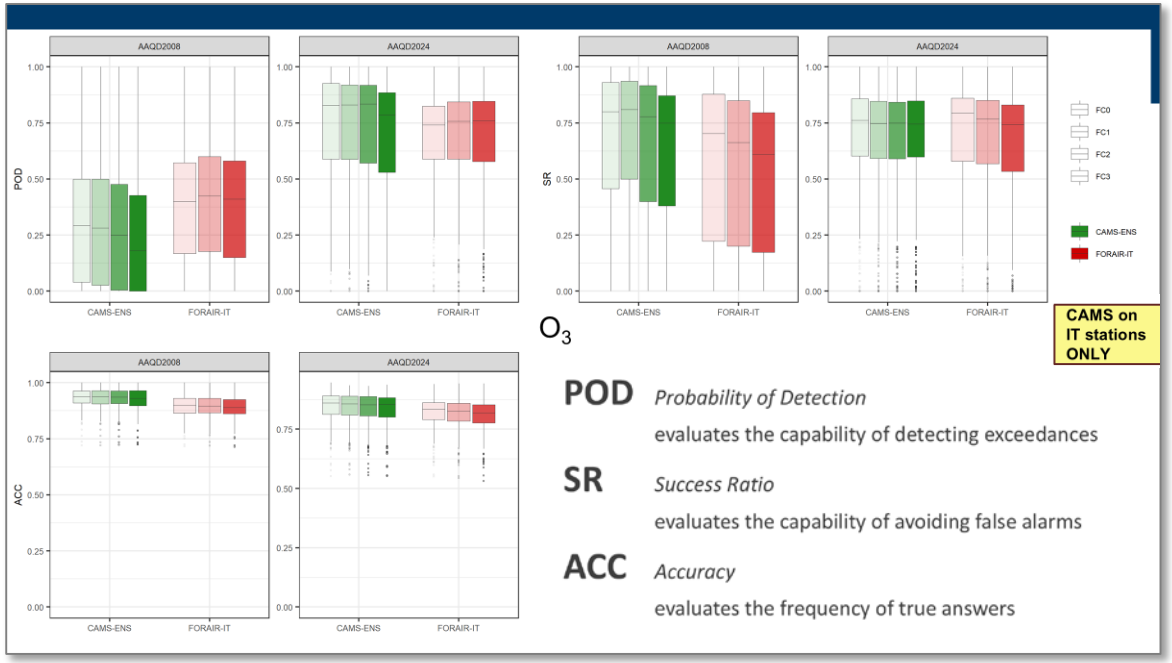
Fixing typical or more frequent values seems not straightforward

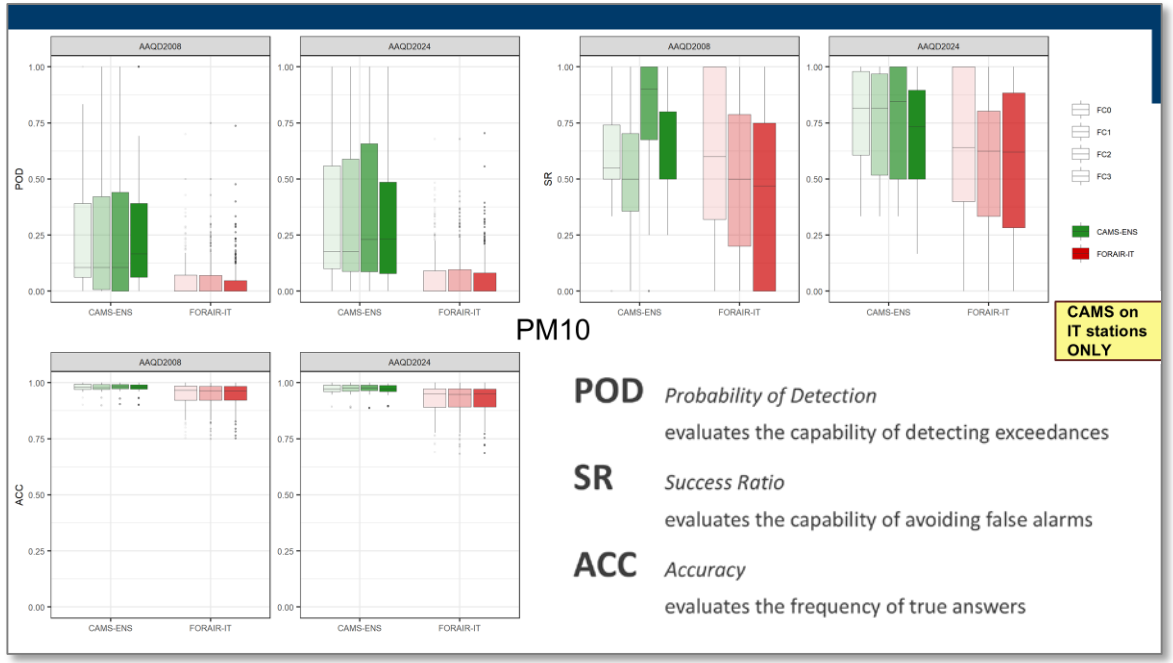
STATUS

2. CAPABILITY IN PREDICTING EXCEEDANCES

B. Could CAMS Ensemble skills be a suitable "quality reference" for evaluating European forecast modeling systems applications?

FORAIR-IT vs CAMS Ensemble





- ✓ O_3: FORAIR-IT performances are consistent with CAMS Ensemble ones
- ✓ PM10: FORAIR-IT performances look worse especially for POD

Extend the comparison to other national forecasting systems? Any volunteer to share their data?

FUTURE STEPS

2025

- survey on the actual use of Forecast MQI, March
- continuing searching criteria for exceedance indicators, comparing with CAMS
- correct timing of AQI forecast (multi-category contingency table and indicators) (Kees?)

After 2025

- continuing searching criteria for exceedance indicators, but changing focus, from exceedance of threshold values to the best possible forecast of episodes (ROC analysis, see literature)
- Communication of forecast exceedances to the citizens (AAQD art. 20 and Annex X)

→ **Online discussion session in April-May**

ENEA universidade de aveiro theoria poiesis praxis European Commission FAIRMODE Forum for air quality modelling in Europe