

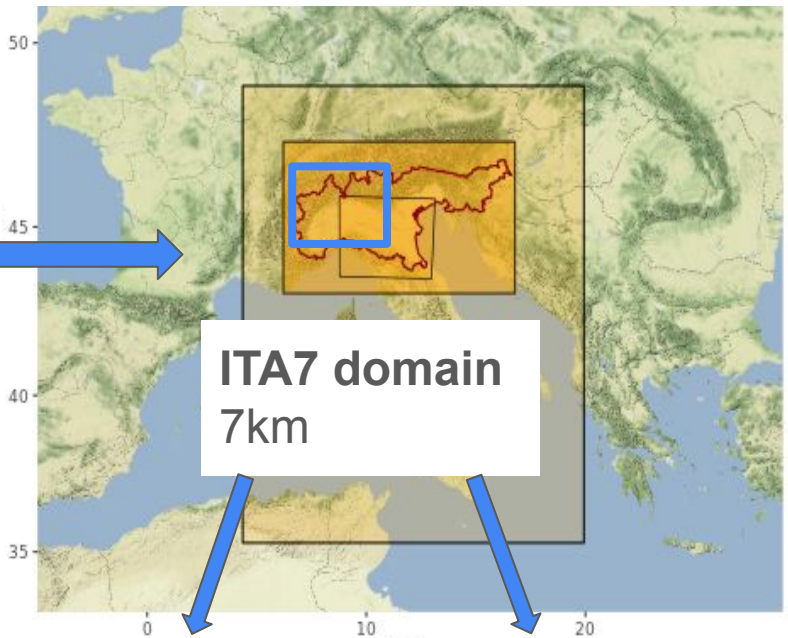
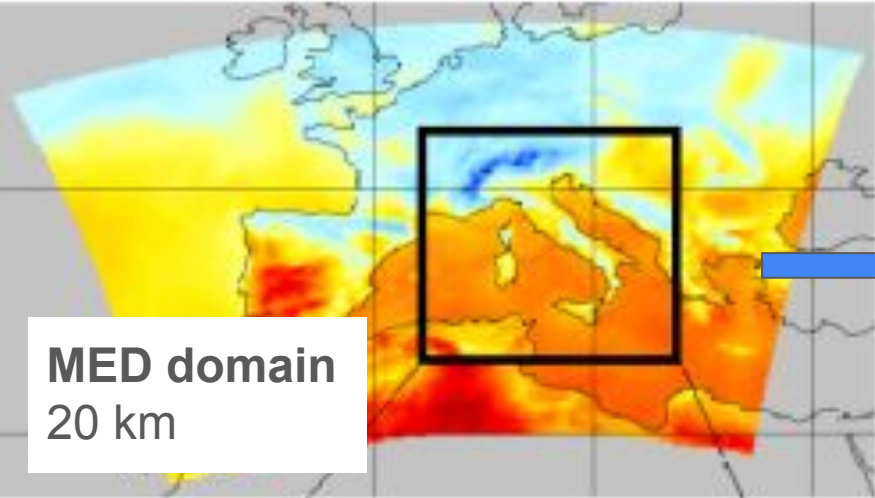
AQ Forecast in Emilia-Romagna, Italy

An attempt to verify forecast with
a probabilistic approach

Roberta Amorati & Michele Stortini

ramorati@arpae.it

Operational forecasting @ Arpae NINFA kAIROS



CTM

CHIMERE + COSMO/ICON

Analysis
 d_{-1}

Forecast
 $d_0 d_1 d_2$

hourly concentrations
 PM_{10} $PM_{2.5}$ O_3 NO_2

Northern Italy + Slovenia
5 km

LIFE 15 IPE IT 013

EMR3 domain
3km

Forecast used to activate extra actions to reduce emissions during episodes of high PM concentrations (concentration $> 50 \mu\text{g}/\text{m}^3$)

reliability is required

Post Processing: IBIS module

Adjusting quantitative CTM prediction with observed daily concentrations

Statistical bayesian model for real-time adjustment

Observed data model

$$Y_t(s) = Z_t(s) + \varepsilon_t(s) \quad \varepsilon_t(s) \sim N(0, \sigma_\varepsilon^2)$$

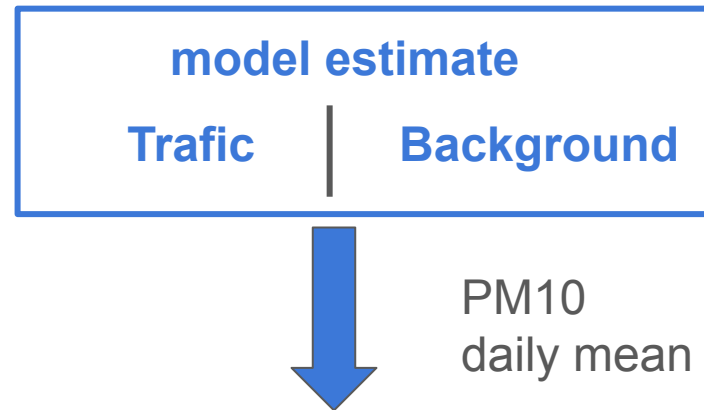
Latent process

$$Z_t(s) = \beta_0 + \beta_1 X_t(B) + \rho Z_{t-1}(s) + \eta_t(s) \quad s \in B$$

$$(\eta_{0,t}(s_1), \dots, \eta_{0,t}(s_n)) \sim N(0, \sigma_\eta^2 S_\eta(\phi, \nu)) \quad t = 1, \dots, 14$$

- ★ β_0 : global additive bias
- ★ $\beta_1 X_t(B_s)$: global multiplicative bias
- ★ $\rho Z_{t-1}(s)$: autoregressive term to model auto-correlation between PM_{10} measurements on successive days
- ★ $\eta_t(s)$: local additive bias (spatially varying intercept process)

autocorrelation
14 days



Probabilistic Prediction

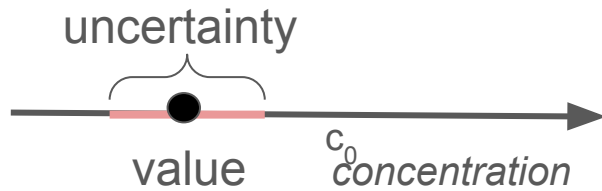
For every prediction location a distribution of possible values

S. K. Sahu, A. E. Gelfand, and D. M. Holland. High-resolution space-time ozone modeling for assessing trends. *Journal of the American Statistical Association*, 102:1221–1234, 2007.

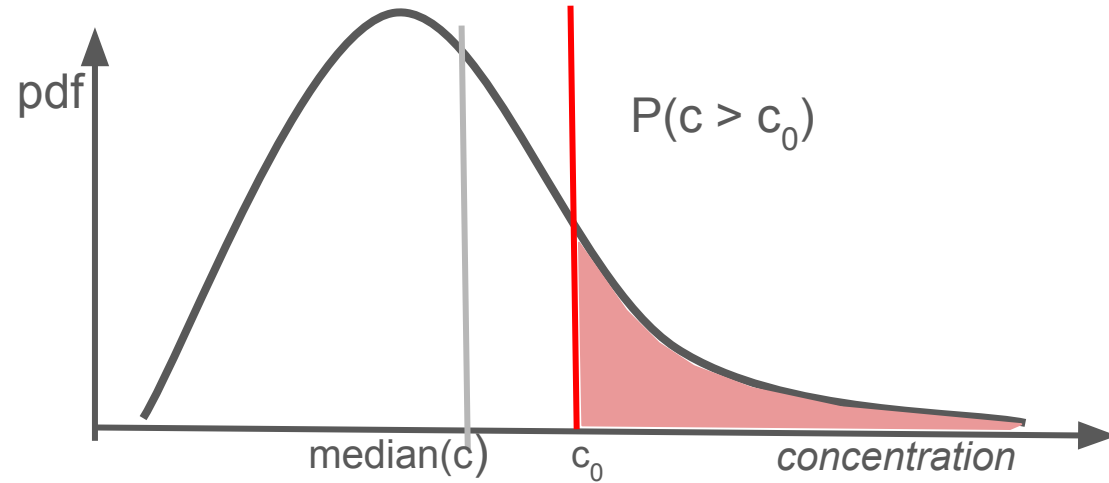
S. K. Sahu, S. Yip, and D. M. Holland. Improved space-time forecasting of next day ozone concentrations in the eastern us. *Atmospheric Environment*, 43:494–501, 2009.

K. S. Bakar and S. K. Sahu. spTimer: spatio-temporal Bayesian modeling using R. *Journal of Statistical Software*, 63(15):1–32, 2015.

Deterministic approach

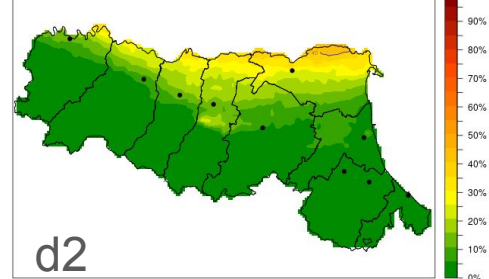
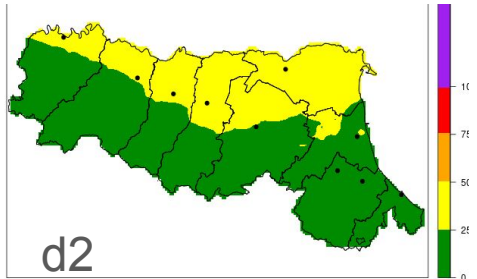
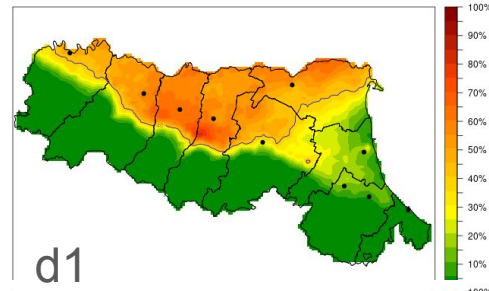
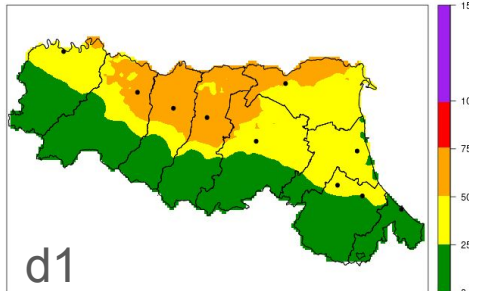
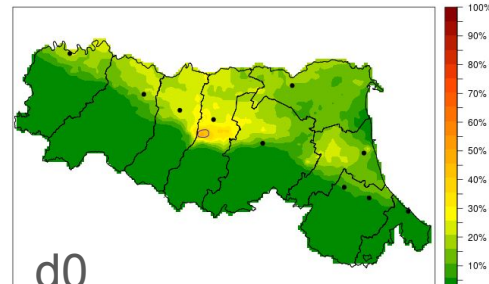
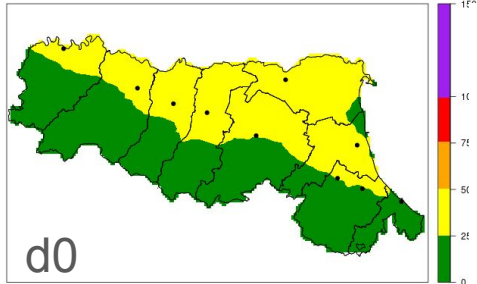


Probabilistic approach



concentration

exceedance probability



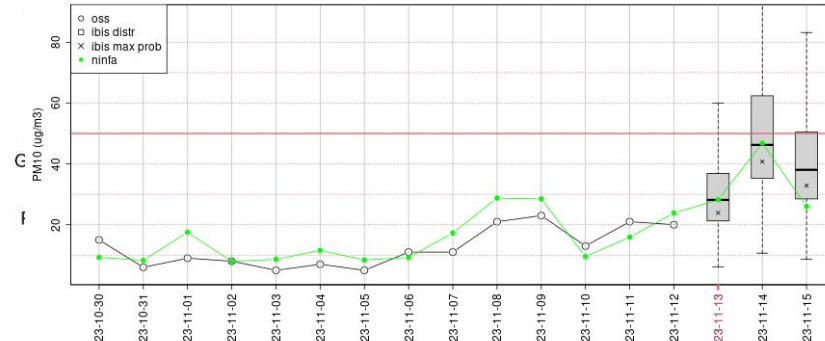
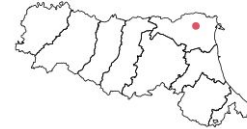
nome	prov	d0	d1	d2	P>50 d0	P>50 d1	P>50d2
CITTADELLA	PR	33	55	26	16	59	6
BADIA	PR	25	49	16	5	48	0
SARAGAT	PR	36	54	36	22	58	22
CASTELLARANO	RE	35	53	19	22	55	1
S. LAZZARO	RE	37	57	28	23	62	9
FEBBIO	RE	6	6	5	0	0	0
S. ROCCO	RE	35	53	36	19	56	22
REMESINA	MO	35	49	34	18	48	17
PARCO FERRARI	MO	38	53	30	26	56	11
GAVELLO	MO	31	52	42	14	52	33
PARCO EDILCARANI	MO	42	66	29	35	74	10
LUGAGNANO	PC	19	28	11	1	10	0
BESENZONE	PC	29	47	31	11	45	15
PARCO MONTECUCCO	PC	33	47	29	17	43	9
CORTE BRUGNATELLA	PC	6	5	5	0	0	0
PARCO RESISTENZA	FC	26	30	22	7	10	2
FRANCHINI-ANGELONI	FC	23	24	20	2	4	1
SAVIGNANO	FC	24	25	22	4	5	2
SAVIGNANO DI RIGO	FC	8	6	7	0	0	0

GI, SAT

Concentrazione di PM10 (ug/m3) osservata e prevista

GHERARDI (FE) - FondoRurale

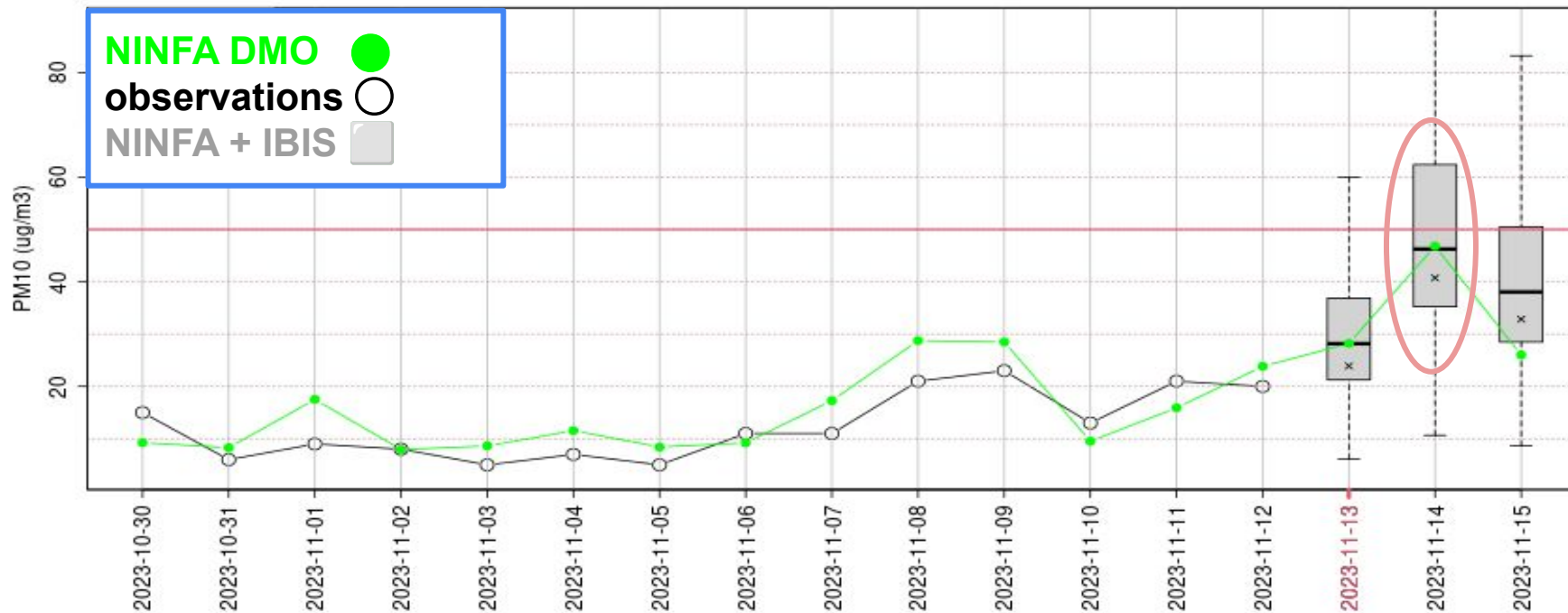
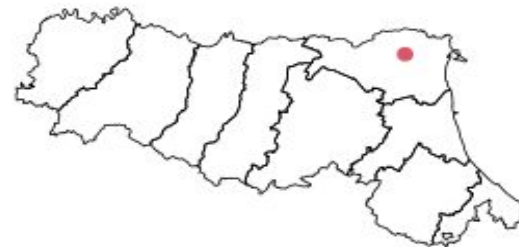
Emissione del 2023-11-13



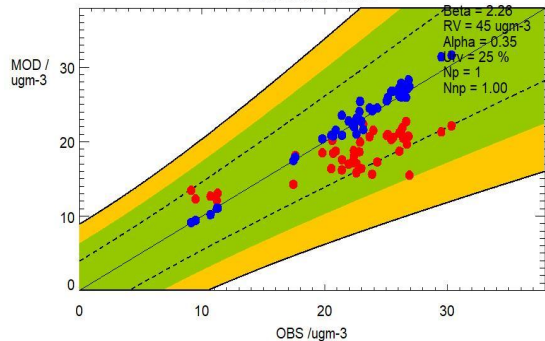
Concentrazione di PM10 (ug/m3) osservata e prevista

GHERARDI (FE) - FondoRurale

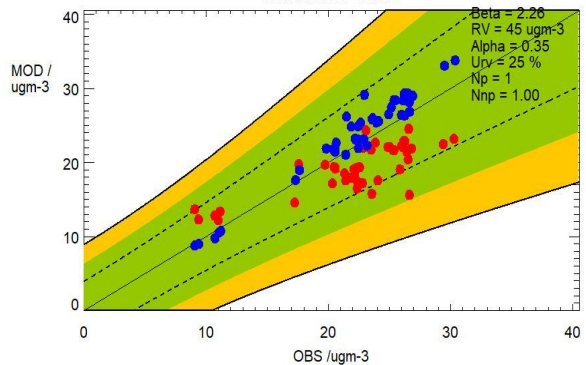
Emissione del 2023-11-13



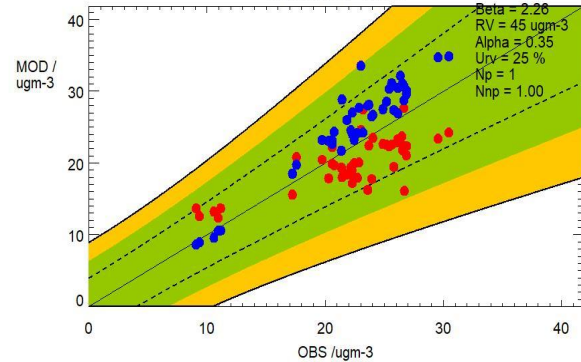
Scatter PLOT PM10



Scatter PLOT PM10



Scatter PLOT PM10

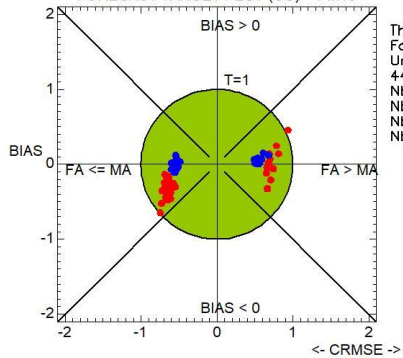


DMO

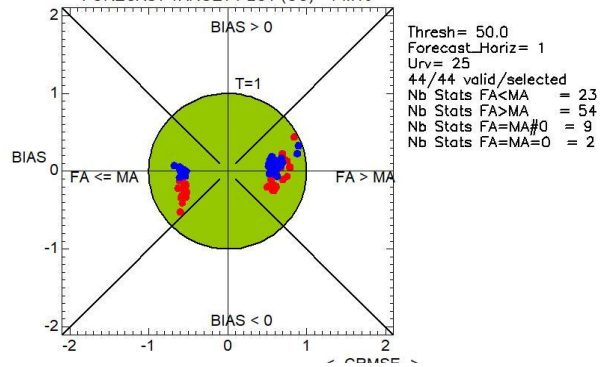
 IBIS
 NINFA DMO

2023

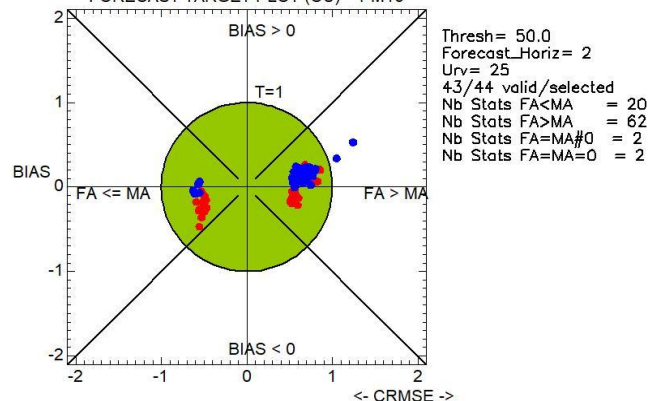
FORECAST TARGET PLOT (OU) PM10



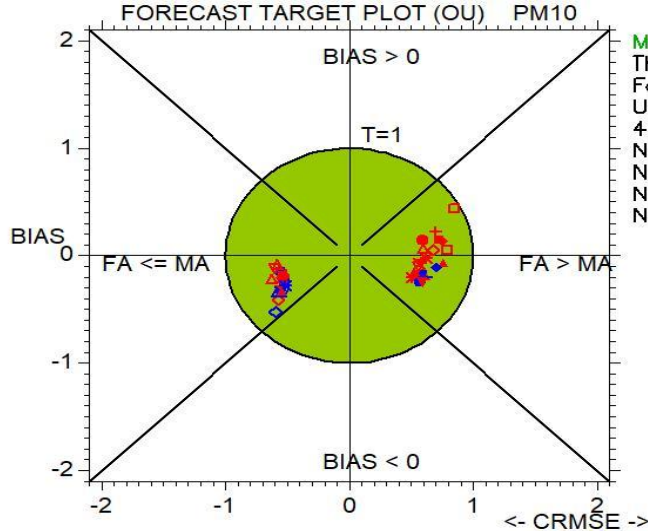
FORECAST TARGET PLOT (OU) PM10



FORECAST TARGET PLOT (OU) PM10

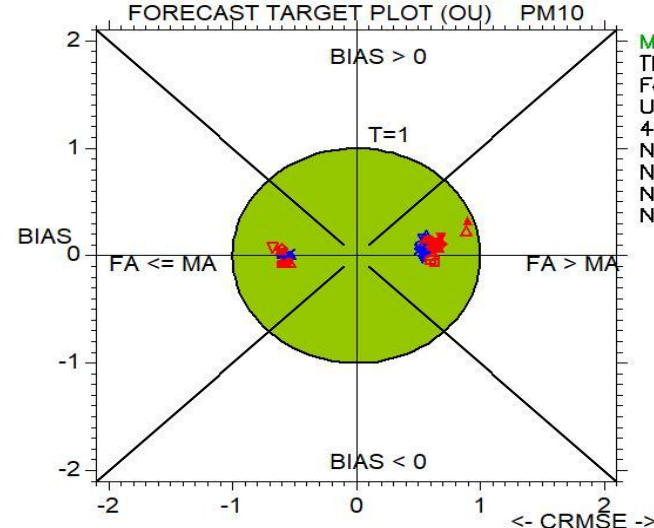


DMO



MQI = 0.750405
 Thresh= 50.0
 Forecast_Horiz= 1
 Urv= 25
 44/44 valid/selected
 Nb Stats FA<MA = 17
 Nb Stats FA>MA = 23
 Nb Stats FA=MA#0 = 4
 Nb Stats FA=MA=0 = 0

IBIS



MQI = 0.676998
 Thresh= 50.0
 Forecast_Horiz= 1
 Urv= 25
 44/44 valid/selected
 Nb Stats FA<MA = 6
 Nb Stats FA>MA = 31
 Nb Stats FA=MA#0 = 5
 Nb Stats FA=MA=0 = 2

◊ FLAMINIA	▼ SLAZZARO	◻ PARCOEDILCARA	▼ SAVIGNANO	◻ ISONZO
◊ MARECCHIA	▼ FEBBIO	◻ LUGAGNANO	▼ SAVIGNANODIRI	◻ GHERARDI
◊ VERUCCHIO	▼ SANROCCO	◻ GIORDANI-FARN	▼ DEAMICIS	◻ CENTO
◊ SANLEO	▼ TIMAVO	◻ BESENZONE	▼ GIARDINIMARGH	◻ VILLAFULVIA
◊ CITTADELLA	▼ GIARDINI	◻ PARCOMONTECUC	▼ PORTASANFELIC	◻ ZALAMELLA
◊ MONTEBELLO	▼ REMESINA	◻ CORTEBRUGNATE	▼ SANLAZZARO	◻ CAORLE
◊ BADIA	▼ PARCOFERRARI	◻ PARCOCRESISTEN	▼ SANPIETROCAPO	◻ DELTACERVIA
◊ SARAGAT	▼ SANFRANCESCO	◻ ROMA	▼ VIACHIARINI	◻ PARCOBERTOZZI
◊ CASTELLARANO	▼ GAVELLO	◻ FRANCHINI-ANG	▼ CASTELLUCCIO	

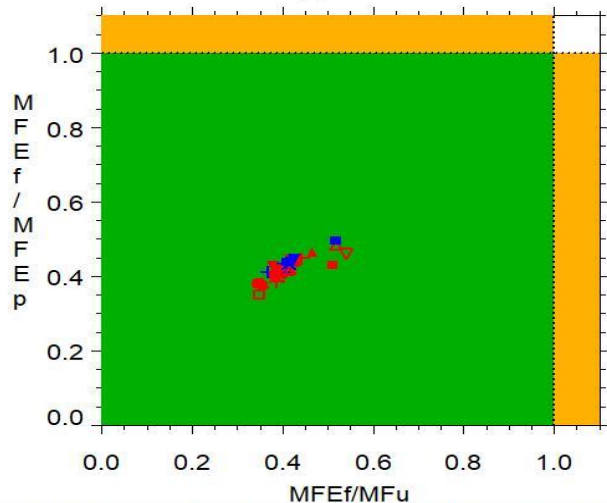
Strt/end Ind: 1-87
 Model (s): DMO
 Parameter: PM11
 Scen: 2023
 Extra Values: 50
 Season: Year
 Day hours: All 24
 Time Average: Pr
 Daily stats: Mean

◊ FLAMINIA	▼ SLAZZARO	◻ PARCOEDILCARA	▼ SAVIGNANO	◻ ISONZO
◊ MARECCHIA	▼ FEBBIO	◻ LUGAGNANO	▼ SAVIGNANODIRI	◻ GHERARDI
◊ VERUCCHIO	▼ SANROCCO	◻ GIORDANI-FARN	▼ DEAMICIS	◻ CENTO
◊ SANLEO	▼ TIMAVO	◻ BESENZONE	▼ GIARDINIMARGH	◻ VILLAFULVIA
◊ CITTADELLA	▼ GIARDINI	◻ PARCOMONTECUC	▼ PORTASANFELIC	◻ ZALAMELLA
◊ MONTEBELLO	▼ REMESINA	◻ CORTEBRUGNATE	▼ SANLAZZARO	◻ CAORLE
◊ BADIA	▼ PARCOFERRARI	◻ PARCOCRESISTEN	▼ SANPIETROCAPO	◻ DELTACERVIA
◊ SARAGAT	▼ SANFRANCESCO	◻ ROMA	▼ VIACHIARINI	◻ PARCOBERTOZZI
◊ CASTELLARANO	▼ GAVELLO	◻ FRANCHINI-ANG	▼ CASTELLUCCIO	

Strt/end Ind: 1-876
 Model (s): IBIS14C
 Parameter: PM10
 Scen: 2023
 Extra Values: 50/1
 Season: Year
 Day hours: All 24h
 Time Average: Pre
 Daily stats: Mean

TRAFIC
BACKGROUND

Forecast_MPI Plot PM10

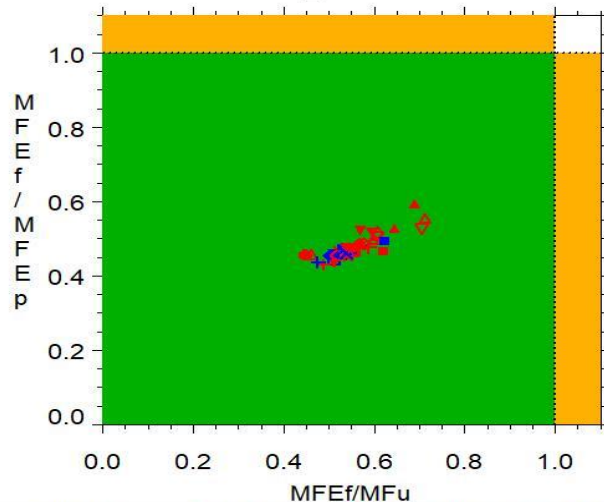


IBIS

Forecast_Horiz= 0

Valid Stats= 44/44
OK= 100%

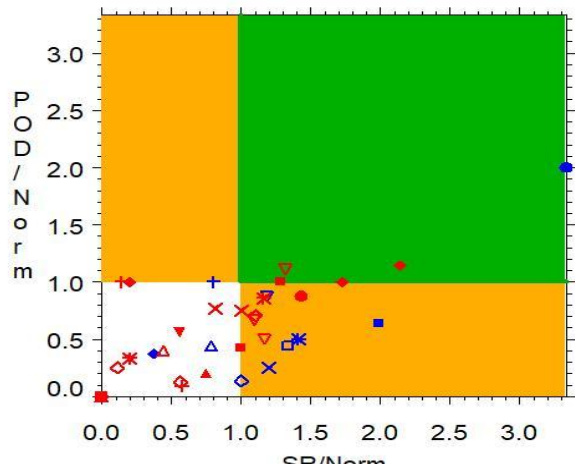
Forecast_MPI Plot PM10



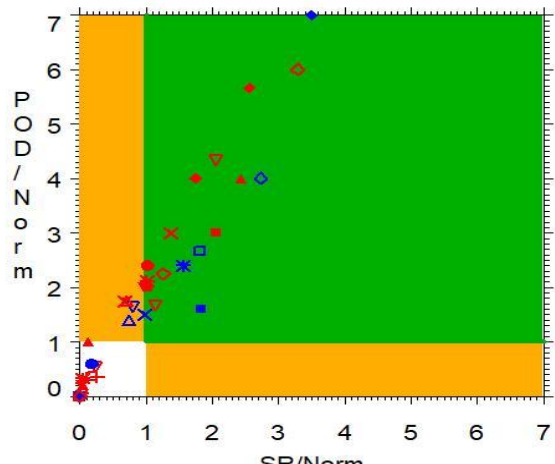
Forecast_Horiz= 2

Valid Stats= 43/44
OK= 100%

Forecast Threshold Performance Normalized PM10

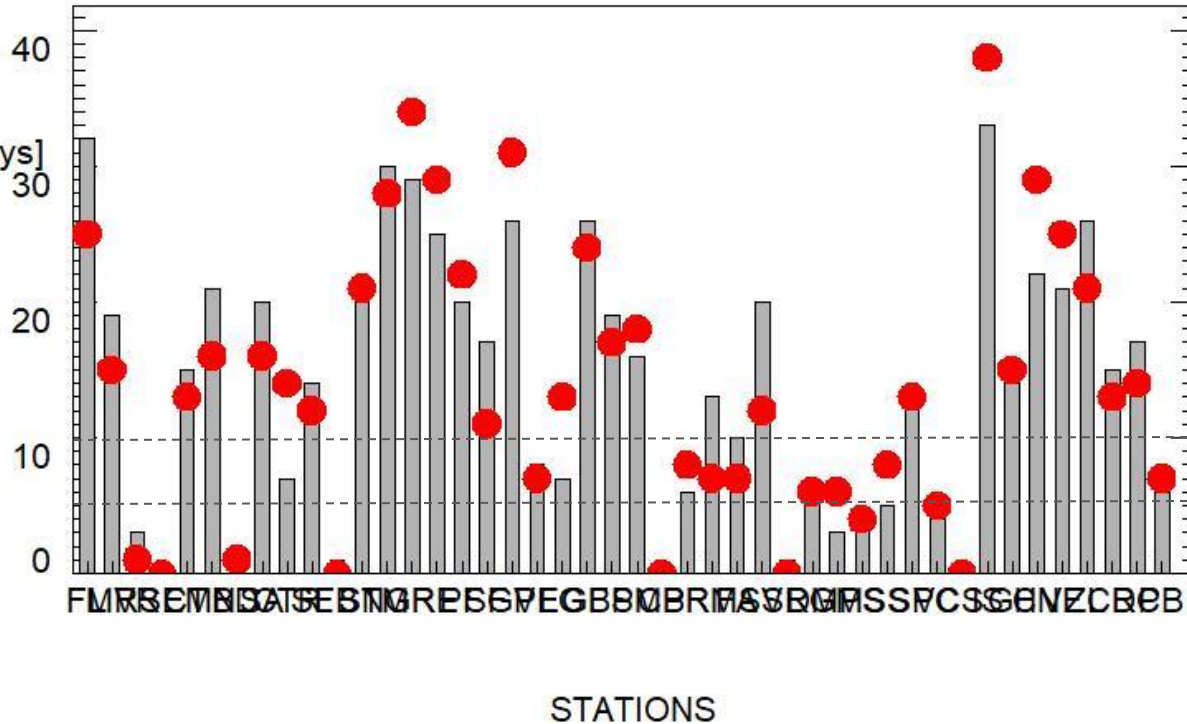
Forecast_Horiz= 0
Thresh= 50.0MPI_(POD/PODp)= 0.000
MPI_(SR/SRp)= 0.000000
41/44 valid stations
12 % valid Stats better th.

Forecast Threshold Performance Normalized PM10

Forecast_Horiz= 2
Thresh= 50.0MPI_(POD/PODp)= 0.000000
MPI_(SR/SRp)= 0.000000
39/44 valid stations
43 % valid Stats better than Per

ExcDays PM10

Units:
[Number of days]



for some
stations few
exceedances

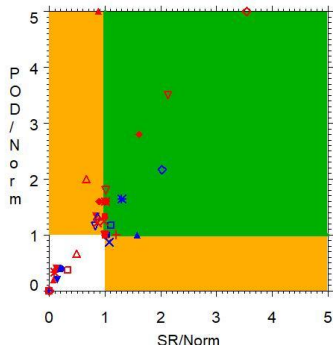
few
misclassification
produce big
errors

● PM10-IBIS14GD0-2023



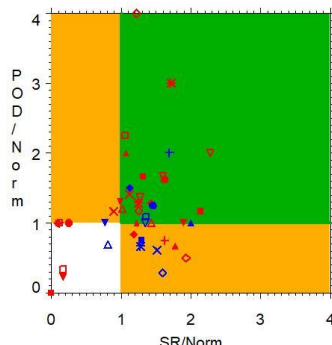
Strt/end Ind: 1-8760
Model (s): IBIS14GD0
Parameter: PM10
Scen: 2023
Extra Values: 50
Season: Year
Day hours: All 24h
Time Average: Preserve
Daily stats: Mean

Forecast Threshold Performance Normalized PM10



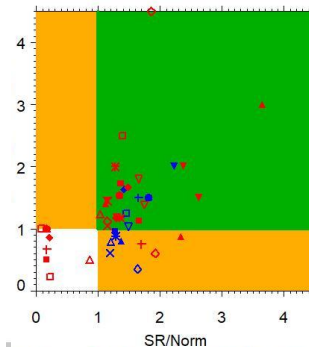
Forecast_Horiz= 1
 Thresh= 50.0
 MPI_(POD/PODp)= 0.160000
 MPI_(SR/SRp)= 0.0571428
 38/44 valid stations
 44 % valid Stats better than Per

Forecast Threshold Performance Normalized PM10



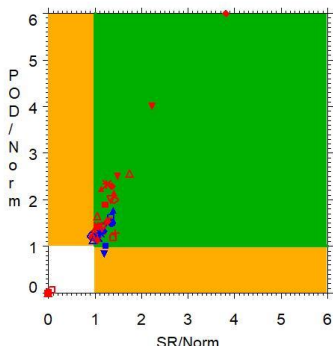
Forecast_Horiz= 1
 Thresh= 45.0
 MPI_(POD/PODp)= 0.534615
 MPI_(SR/SRp)= 0.166667
 43/44 valid stations
 55 % valid Stats better than Persistence

Forecast Threshold Performance Normalized PM10



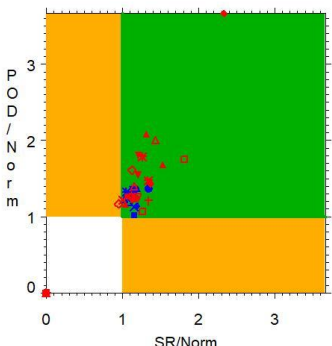
Forecast_Horiz= 1
 Thresh= 40.0
 MPI_(POD/PODp)= 0.603478
 MPI_(SR/SRp)= 0.185714
 44/44 valid stations
 61 % valid Stats better than Persistence

Forecast Threshold Performance Normalized PM10



Forecast_Horiz= 1
 Thresh= 35.0
 MPI_(POD/PODp)= 0.0266668
 MPI_(SR/SRp)= 0.0307693
 44/44 valid stations
 77 % valid Stats better than Pe

Forecast Threshold Performance Normalized PM10



Forecast_Horiz= 1
 Thresh= 30.0
 MPI_(POD/PODp)= 0.406898
 MPI_(SR/SRp)= 0.379168
 44/44 valid stations
 86 % valid Stats better than P

Threshold	% valid station better than persistence	
	IBIS	DMO
50	44	45
45	55	55
40	61	61
35	77	59
30	86	30

decreasing threshold increases performance with respect to persistence

What about probability?

**Can we say anything about model performance
in predicting probability of exceedance?**

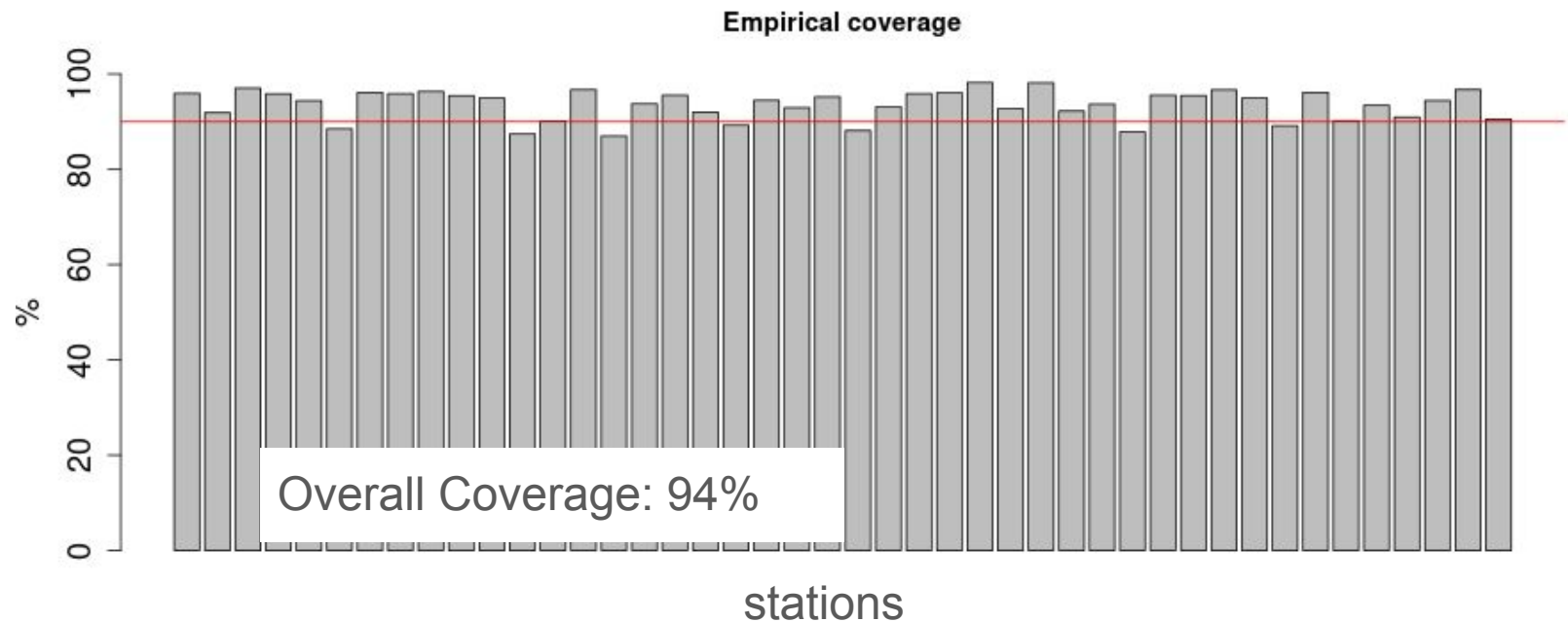
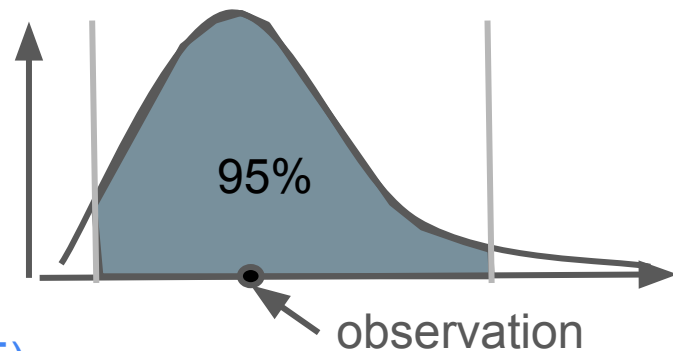
tentative probabilistic verification

Empirical coverage

How good is model in explaining variability?

how many times is the observation in the credible range of prediction, i.e. 95%

$\text{percentile}(0.025) < \text{observation} < \text{percentile}(0.975)$



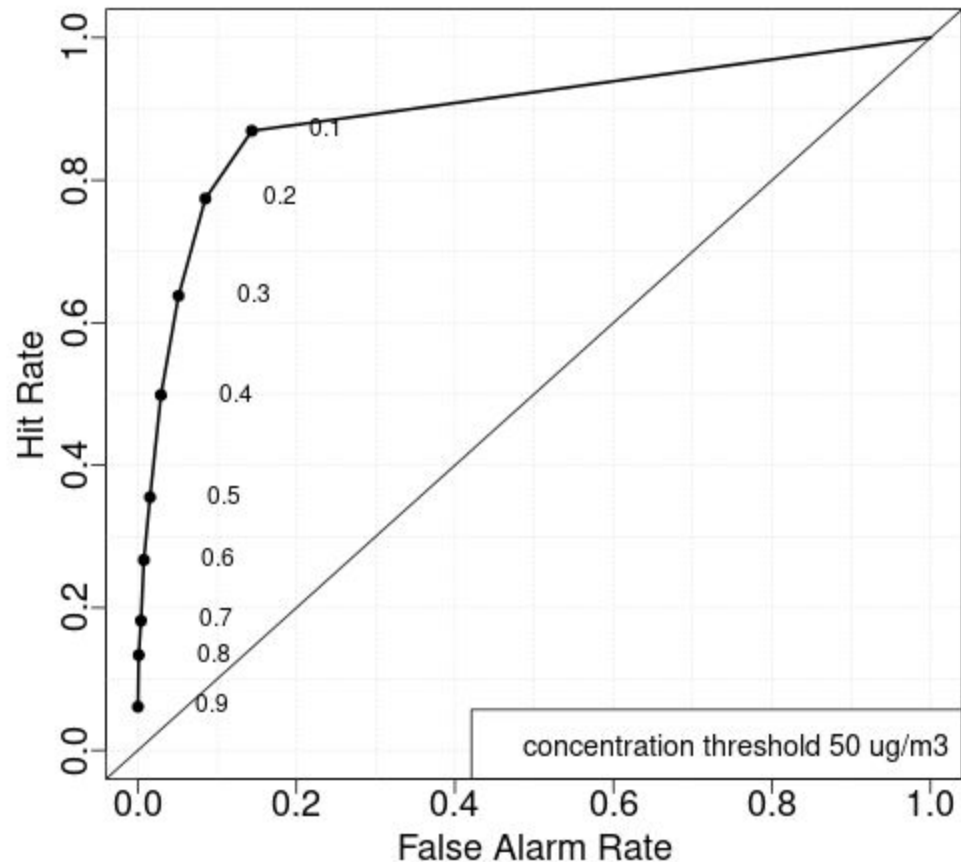
Relative Operating Characteristic

probabilistic prediction can be converted to deterministic for each probability threshold

if $p > p_{Th} \Rightarrow X=1$ yes

else $X=0$ no

Relative Operating Characteristics (ROC) curve - IBIS d0



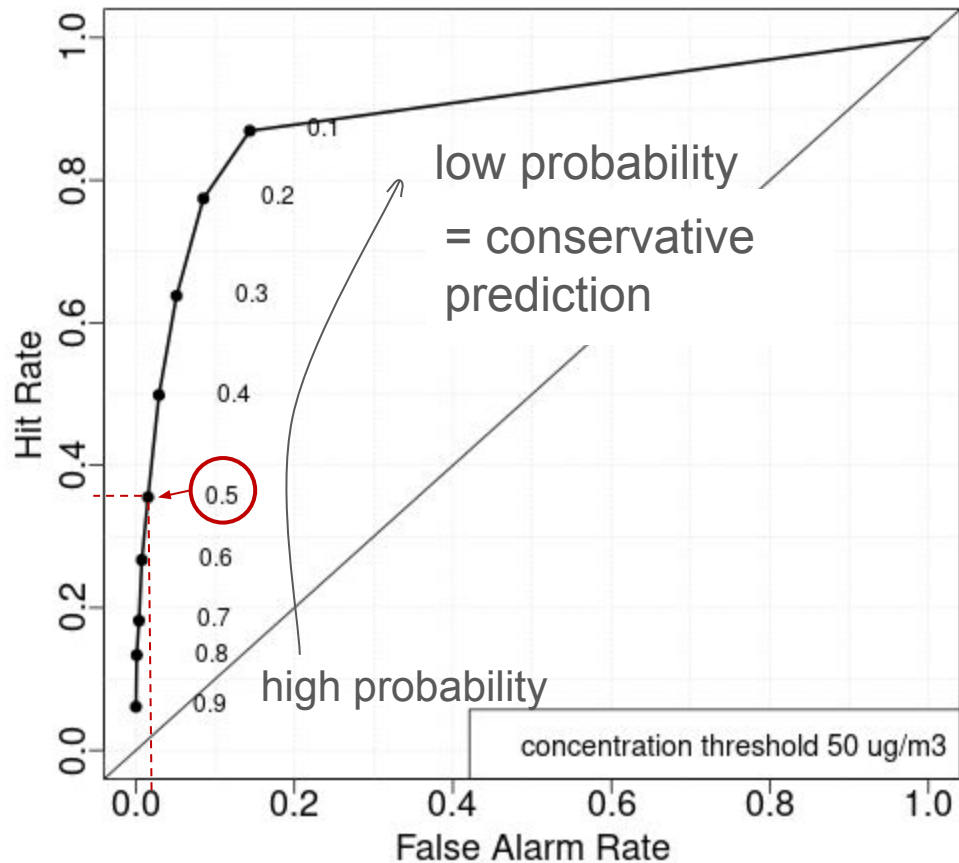
Relative Operating Characteristic

probabilistic prediction can be converted to deterministic prediction for each probability threshold

if $p > p_{Th}$ \Rightarrow $X=1$ yes
else $X=0$ no

event X: exceedance

Relative Operating Characteristics (ROC) curve - IBIS d0



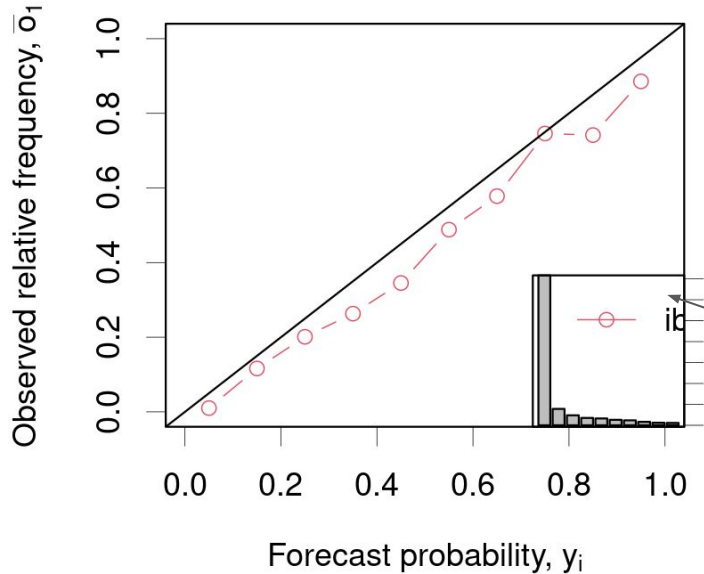
Reliability Diagram

shows relation between probability of event occurrence (predicted) and observed frequency of that event

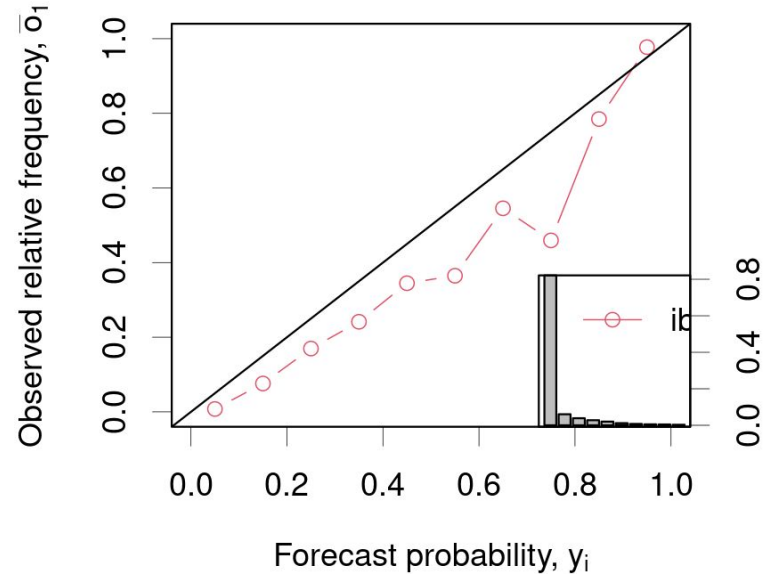
perfect prediction: points on the bisector

d0 40 ug/m3

d0 50 ug/m3



Sharpness
histogram: the frequency of forecasts in each probability bin (histogram) shows the sharpness of the forecast



Outcome

- The AQ forecast NINFA DMO satisfy the requirements
- The statistical post processing IBIS contributes to reduce quantitative bias
- The performance categorical prediction of exceedance depends on threshold
- A probabilistic approach is evaluated
- Open question: is it a probabilistic forecast worth for episode warning?

Thank you for your attention

ramorati@arpae.it