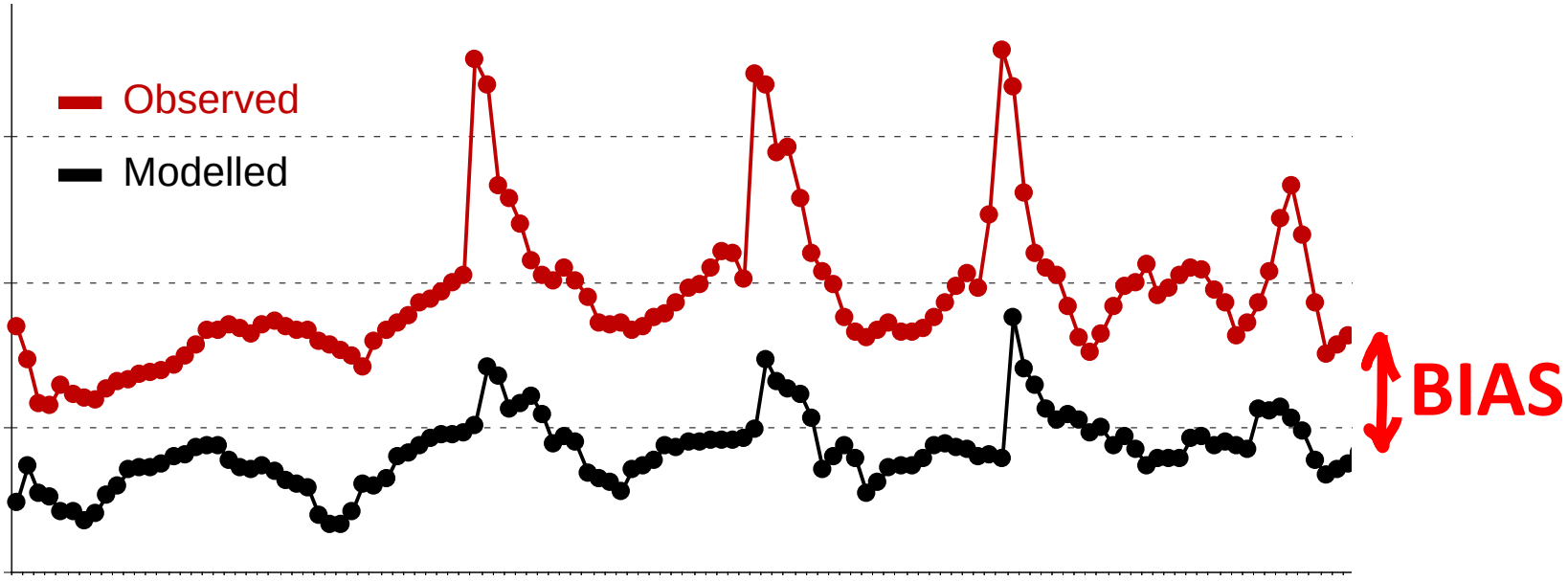
How different air quality forecasting systems (should) operate over Portugal?

Carlos Borrego | Alexandra Monteiro | Isabel Ribeiro | Ana Isabel Miranda
University of Aveiro | Portugal

M. T. Pay | S. Basart | J. M. Baldasano
BSC | Barcelona Supercomputing Center

HARMO14 conference
Kos, Greece | 2-6 October 2011

Motivation



Test, comprehend and evaluate different bias correction techniques

Very often, large errors and bias occur on forecasting modelling

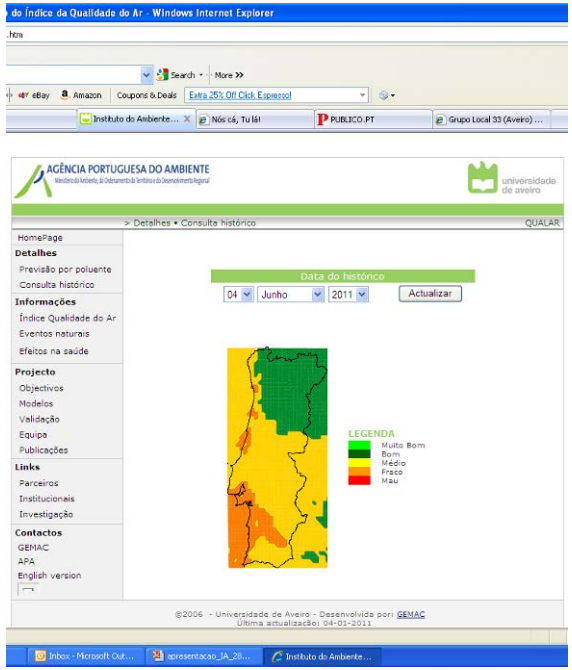
To improve the air quality forecast using bias-correction

Evaluate and test different bias correction techniques

How?

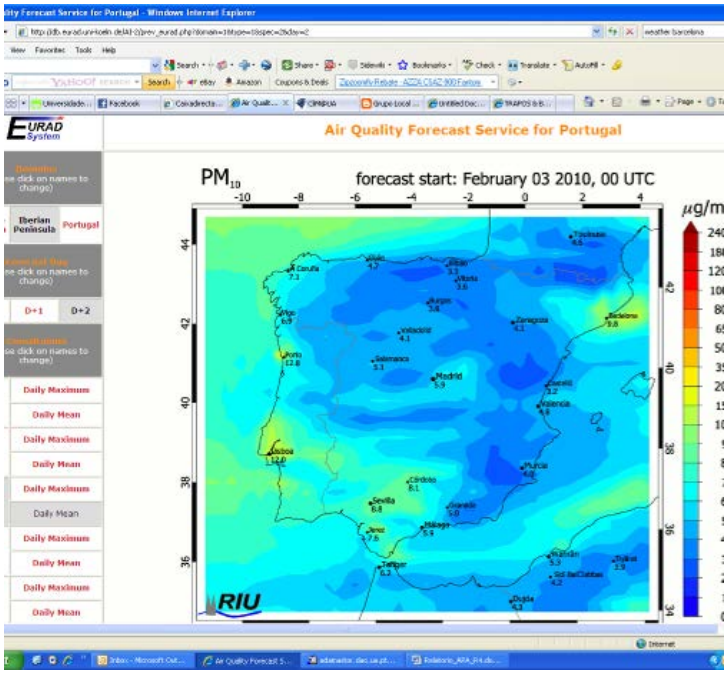
3 different models deliver daily operational forecasts over Portugal
Their performance with bias-correction was investigated for 2010 year

CHIMERE | UA

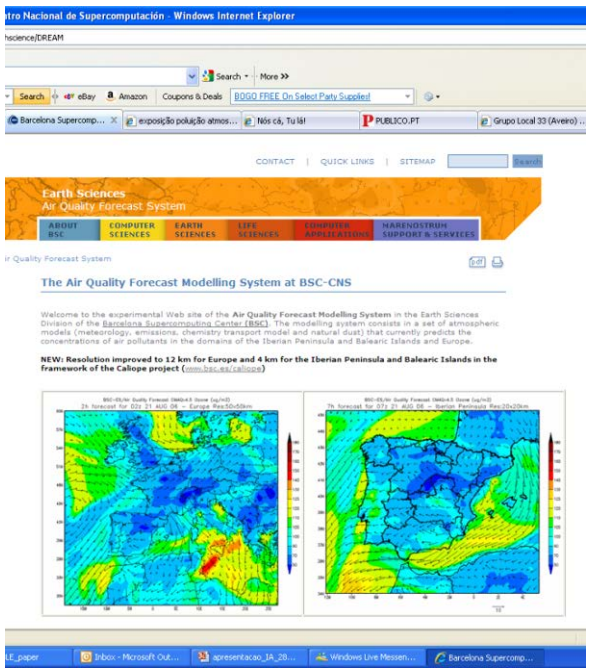


5 x 5 km

EURAD | RIU + UA



CALIOPE | BSC, Spain



4 x 4 km

Bias-correction techniques

We started with 2 pos-processing methods, comparing…

SUBST

an **additive correction** of the mean bias

$$C^{corrected}(h,day) = \frac{1}{ndays}\sum_{ndays}(C_h^{model} - C_h^{obs}) + C^{model}(h,day)$$

RAT

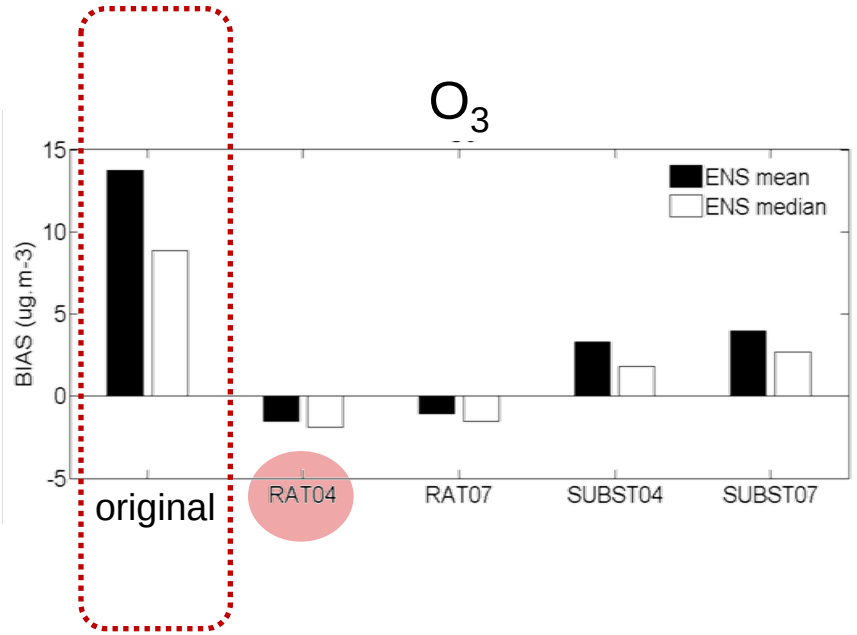
a **multiplicative ratio** correction

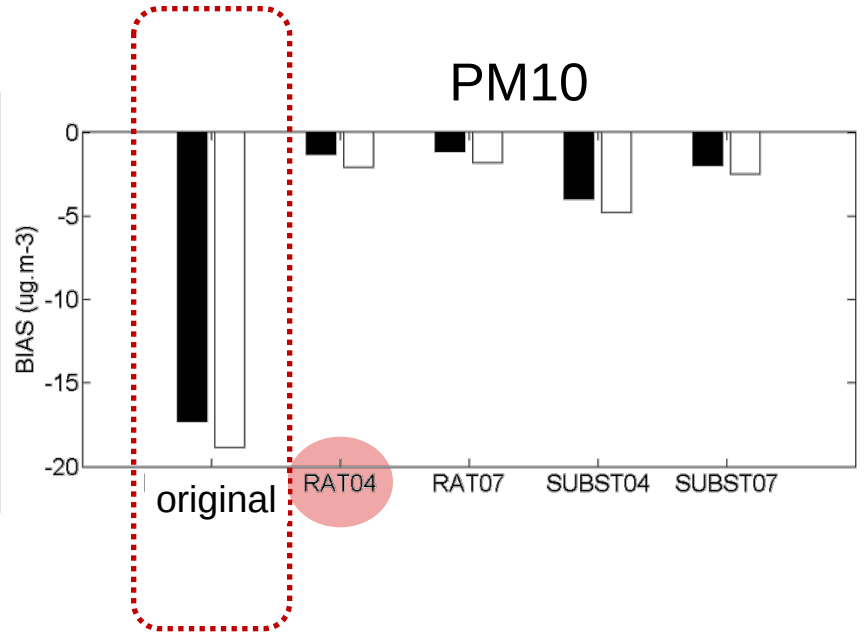
$$C^{corrected}(h,day) = \frac{\sum_{ndays} C^{obs}(h,day)}{\sum_{ndays} C^{model}(h,day)} \times C^{model}(h,day)$$

both tested using different period of previous days (3-4 and 7 days) to correct the forecast bias

Bias-correction techniques

RAT & SUBST





- after BIAS correction, model results have a decrease > 70% on average systematic error
- the multiplicative ratio: better correction technique
- synoptic conditions are characterized by a 3-4 day period.

Bias-correction techniques
Kalman Filter (KF)

- KF performance is sensitive to **error ratio ($\sigma^2\eta/\sigma^2\varepsilon$)**:

  $\sigma^2\eta$ forecast-error white-noise variance

  $\sigma^2\varepsilon$ true forecast-bias white-noise variance

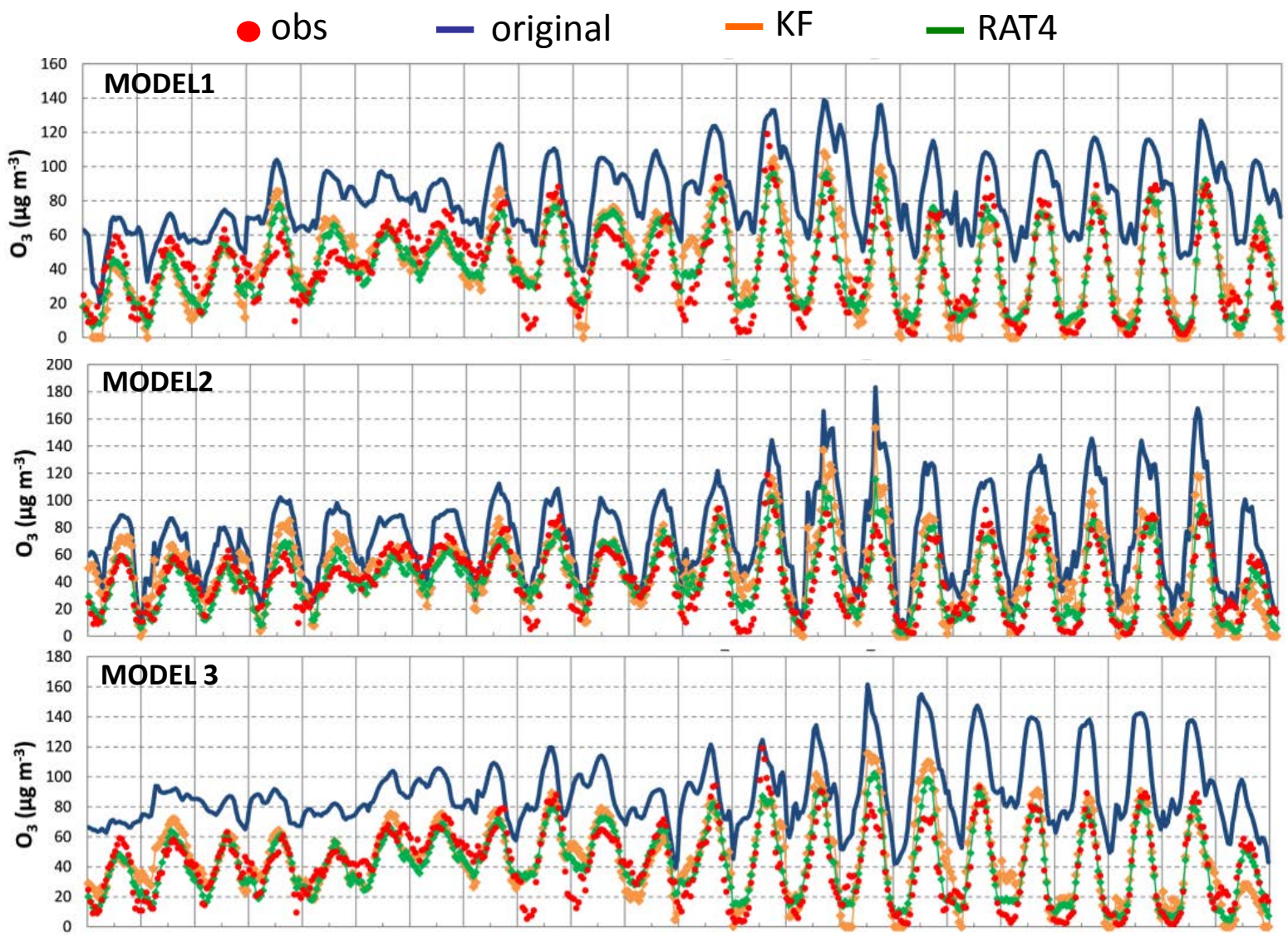
optimal error ratio to generate the best forecast

- an ***optimal error ratio*** was estimated minimizing **RMSE** and maximizing the **CORR** for all the stations

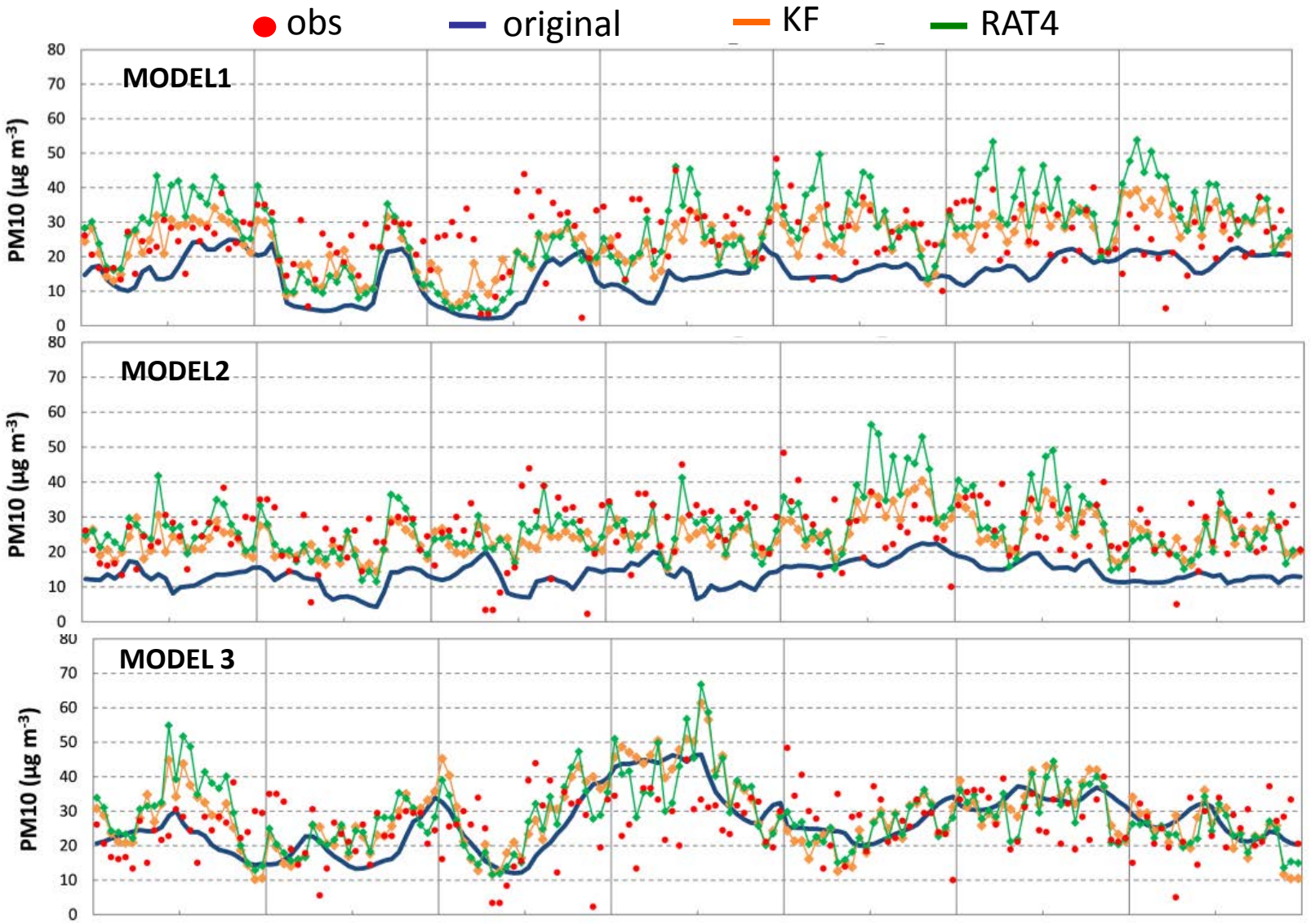
- Due to the relatively low extension of Portugal, it was assume **no spatial variability** of optimal error ratios over the country

How do bias-correction techniques perform in general?

O_3 time series

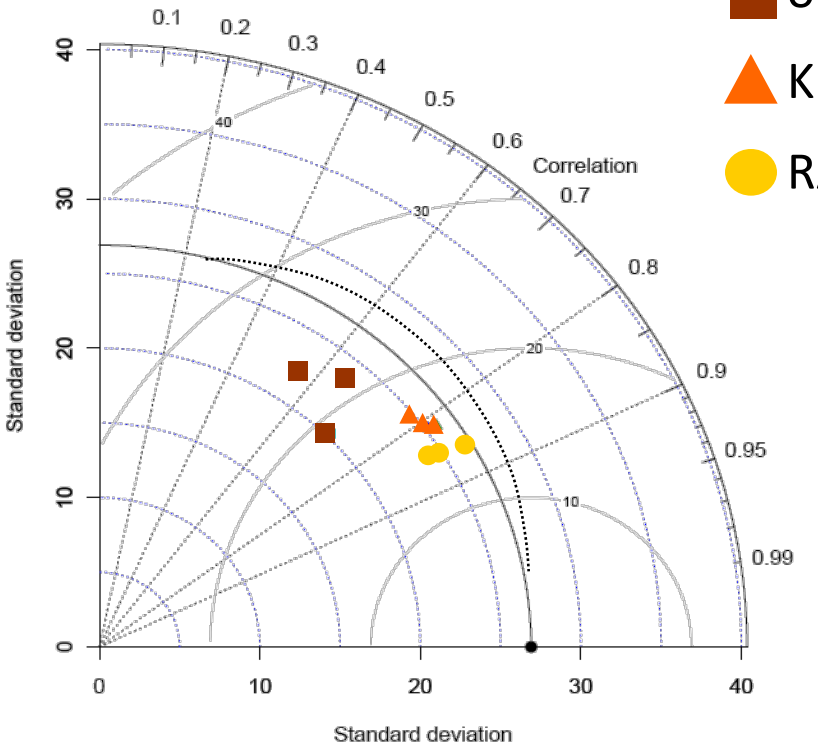


PM10 time series

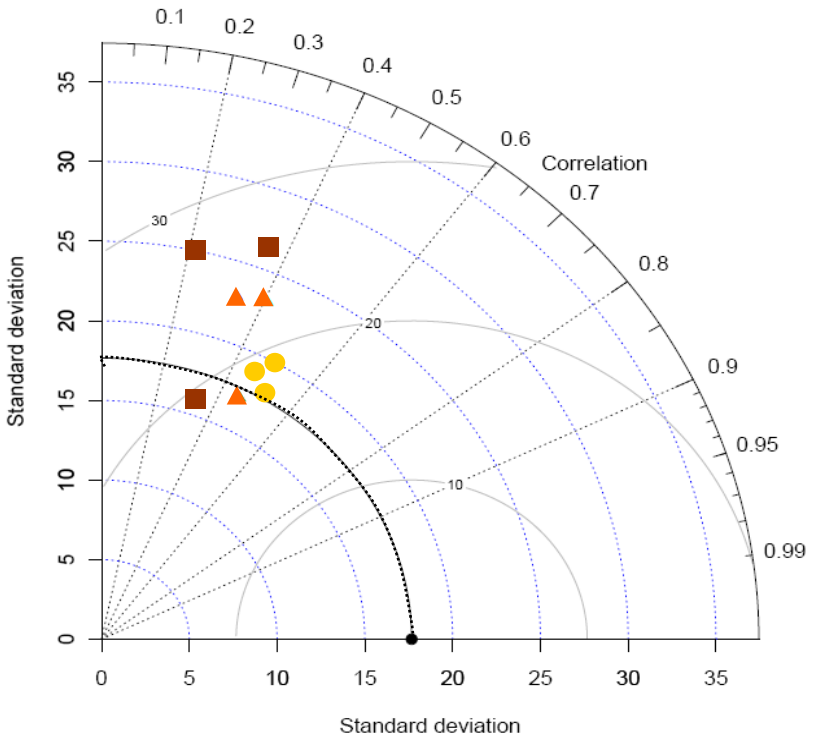


Taylor diagram results

O_3



PM10



Models skills improved significantly with bias-correction

No evident technique better than the other (mainly for O_3!!)

Did bias-correction techniques improve forecast performance?

PM10 episodes | daily limit value exceedances

FORECAST ORIGINAL MODEL

FORECAST AFTER BIAS CORRECTION

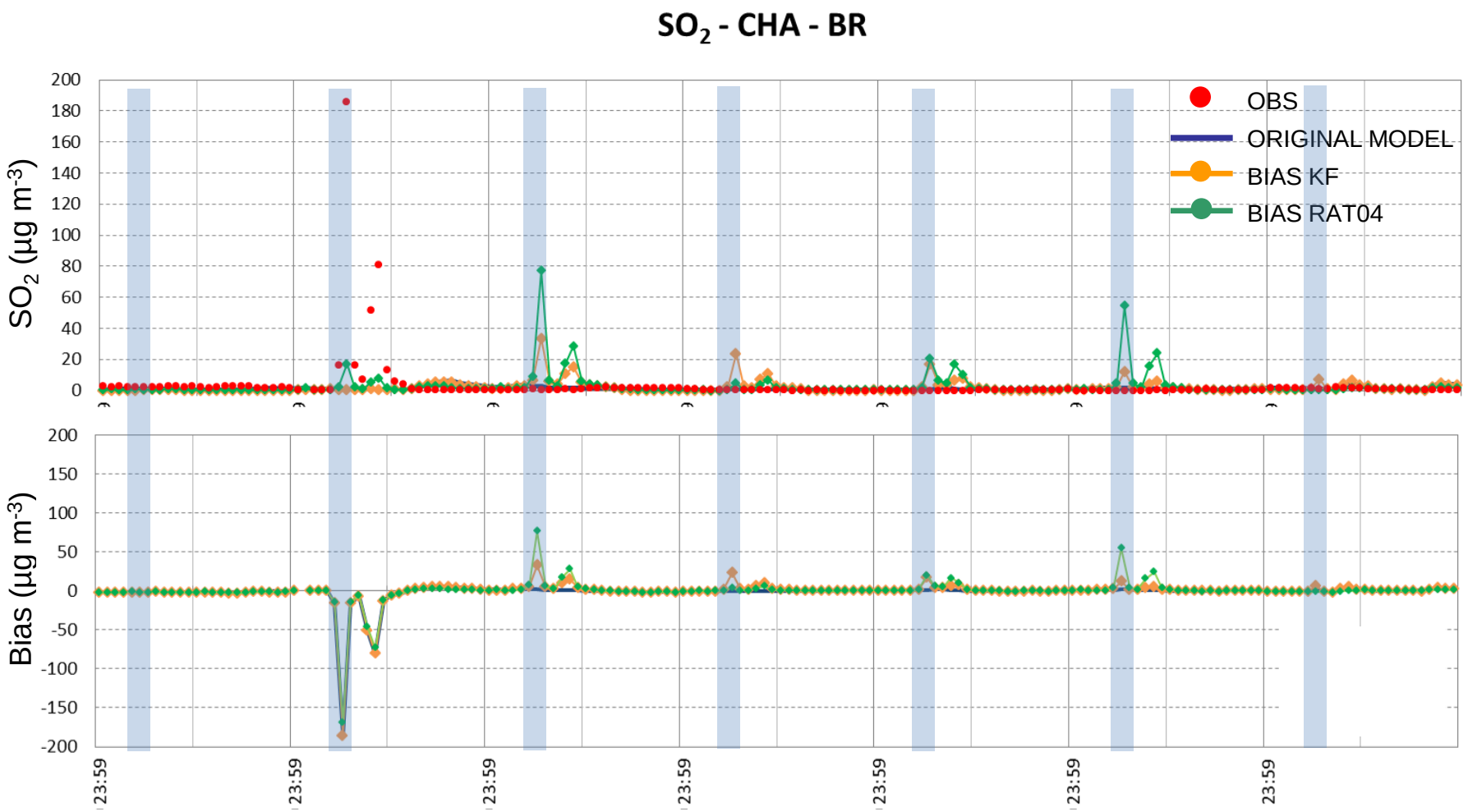
DATE	STATION	MODEL 1	MODEL 2	MODEL 3
04/07/2010	SOB	0	0	0
05/07/2010	SOB	0	1	0
06/07/2010	SOB	0	1	1
28/07/2010	SOB	1	1	1
29/07/2010	FUN	0	0	0
29/07/2010	SOB	1	1	1
04/08/2010	SOB	0	0	0
06/08/2010	SOB	0	1	0
07/08/2010	SOB	1	1	1
08/08/2010	FUN	1	1	1
09/08/2010	CAM	1	1	1
09/08/2010	ERV	0	1	1
09/08/2010	FUN	1	1	1
09/08/2010	OLI	0	1	1
10/08/2010	FUN	1	1	1
10/08/2010	MVE	1	1	1
11/08/2010	CAM	0	0	0
11/08/2010	FUN	0	0	0
11/08/2010	OLI	0	0	0
12/08/2010	FUN	0	0	0
30/08/2010	ERV	0	0	0
30/08/2010	SOB	1	1	0
31/08/2010	ERV	0	0	1
31/08/2010	SOB	1	1	0
20/10/2010	ERV	0	0	0
20/10/2010	OLI	0	1	0
21/10/2010	ERV	1	0	1
21/10/2010	OLI	1	1	1
22/10/2010	OLI	1	1	0
22/10/2010	SOB	1	1	1
23/10/2010	ERV	1	0	1
23/10/2010	SOB	1	1	1
27/10/2010	OLI	1	1	1
28/10/2010	OLI	1	1	1
28/10/2010	SOB	0	1	0
06/11/2010	OLI	0	1	1
11/11/2010	OLI	1	1	0
13/12/2010	SOB	0	0	0

Problems & advantages

of each technique

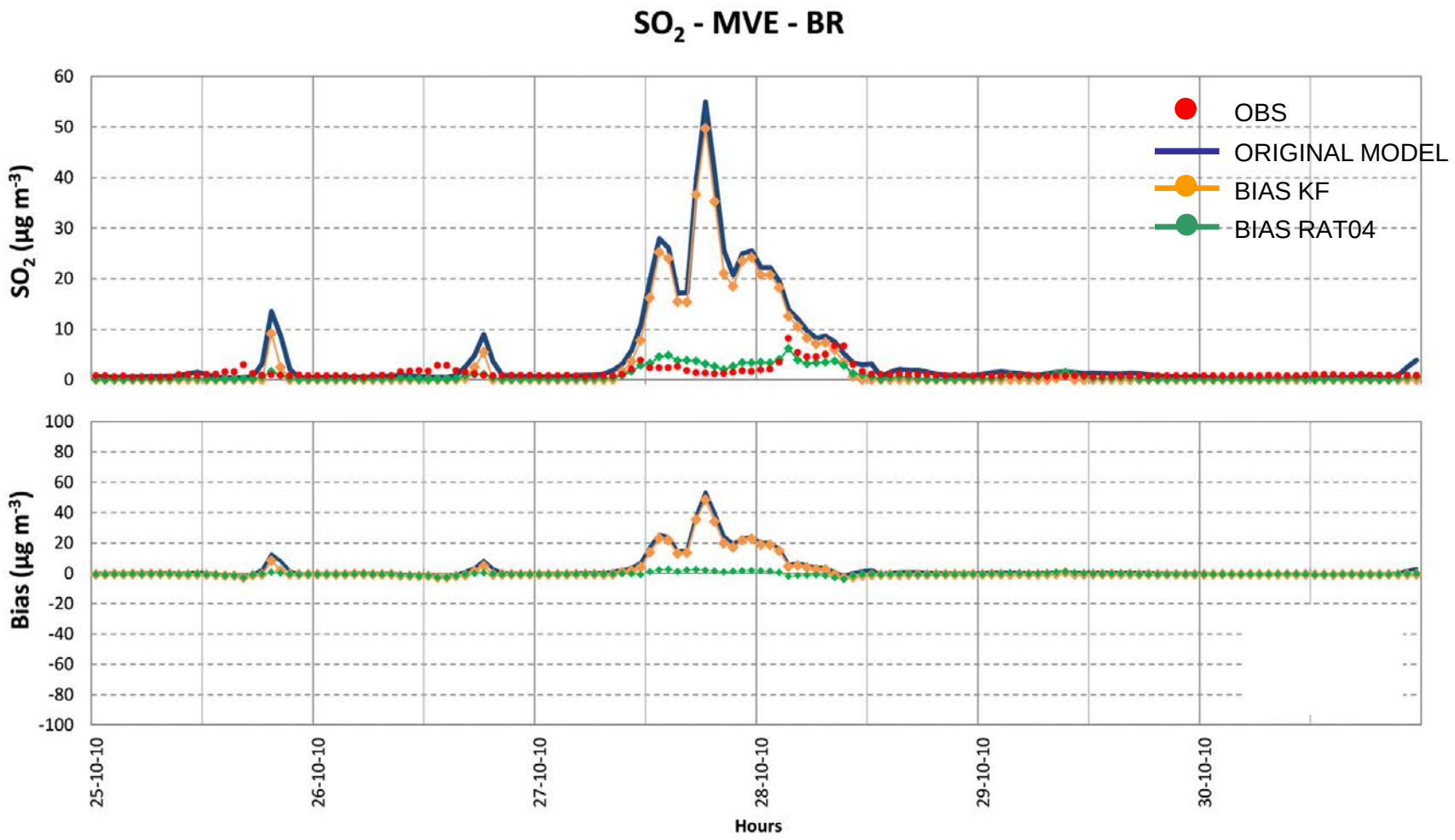
to understand the weekness and the streghth of each technique…in order to improve them and its application

where RAT04 fails…



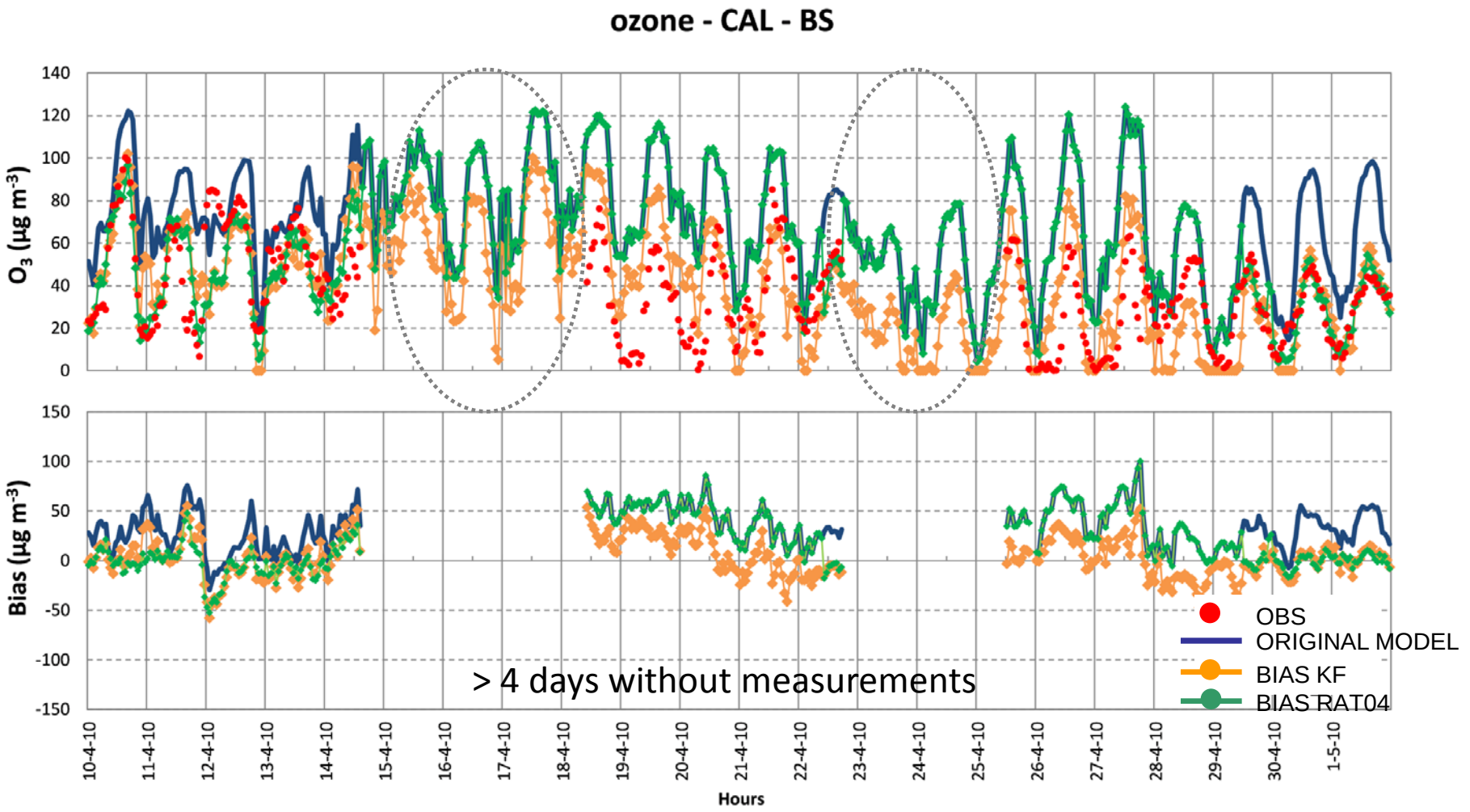
when an instantaneous peak happen in specific day/hours, this peak will be reproduced in the following days (4) by both techniques, but more strongly for RAT04

where Kalman filter fails…



When model predict an episode consistently, KF trust in model more than in observations…which do not happen with RATO4

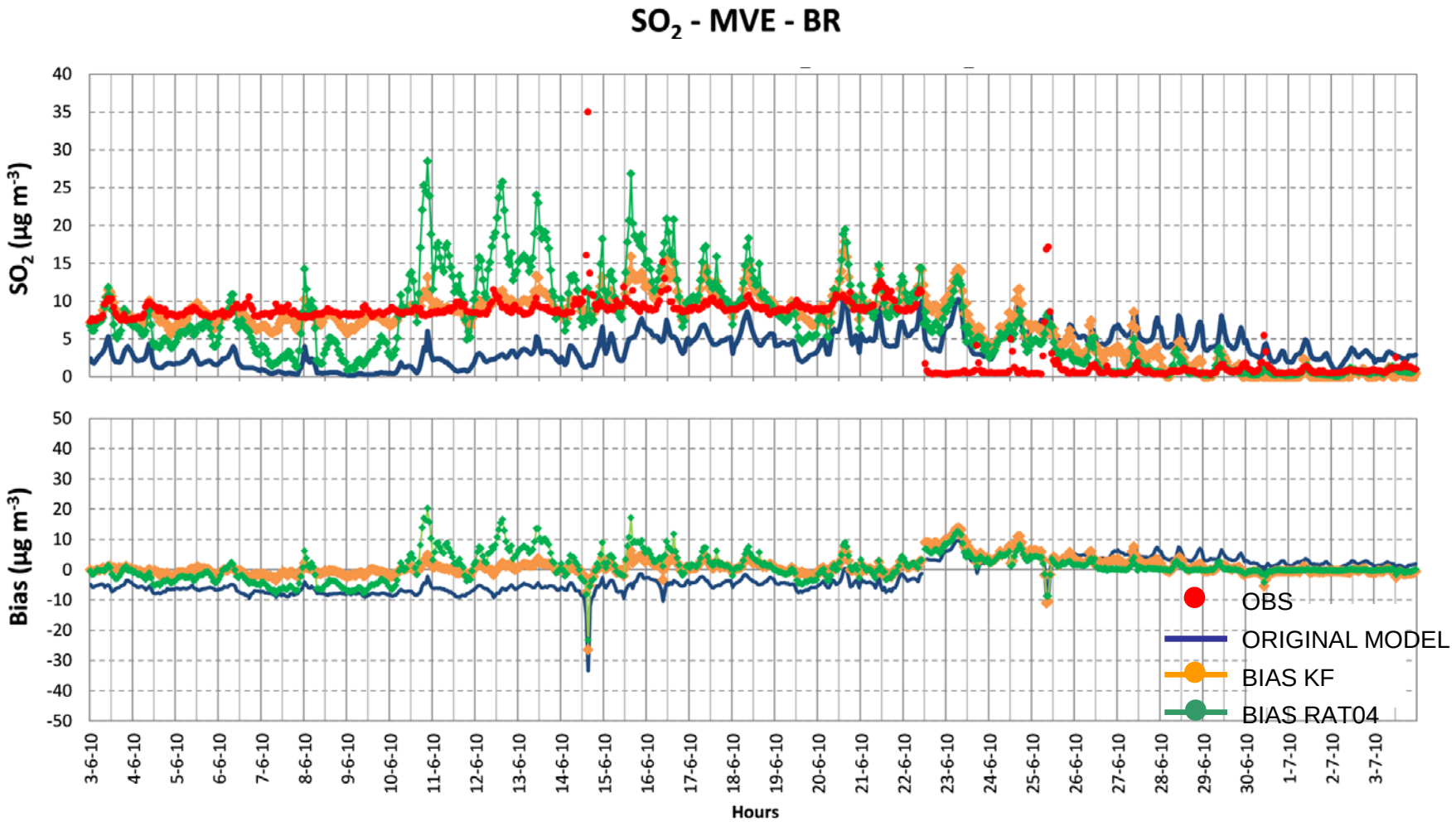
The “no data” problem



Both techniques don’t have values to be guided.

Values will not be corrected until new data appears.

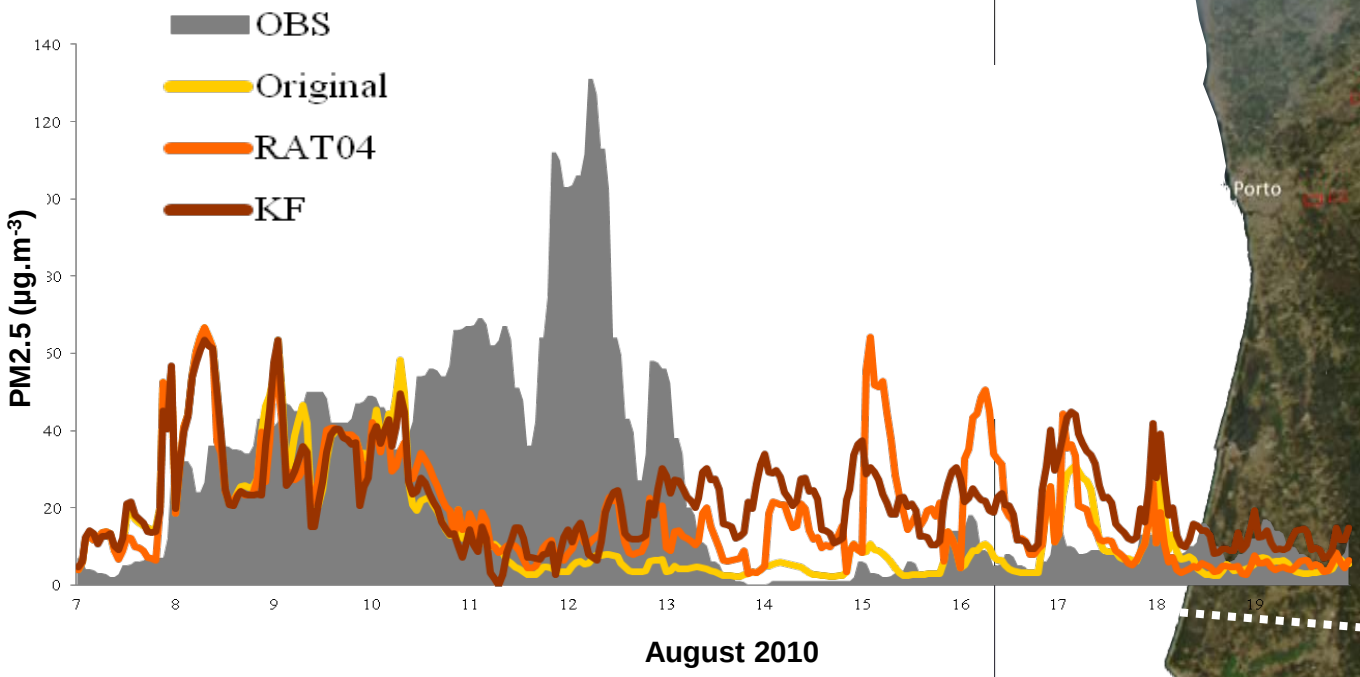
The “no validated data” problem



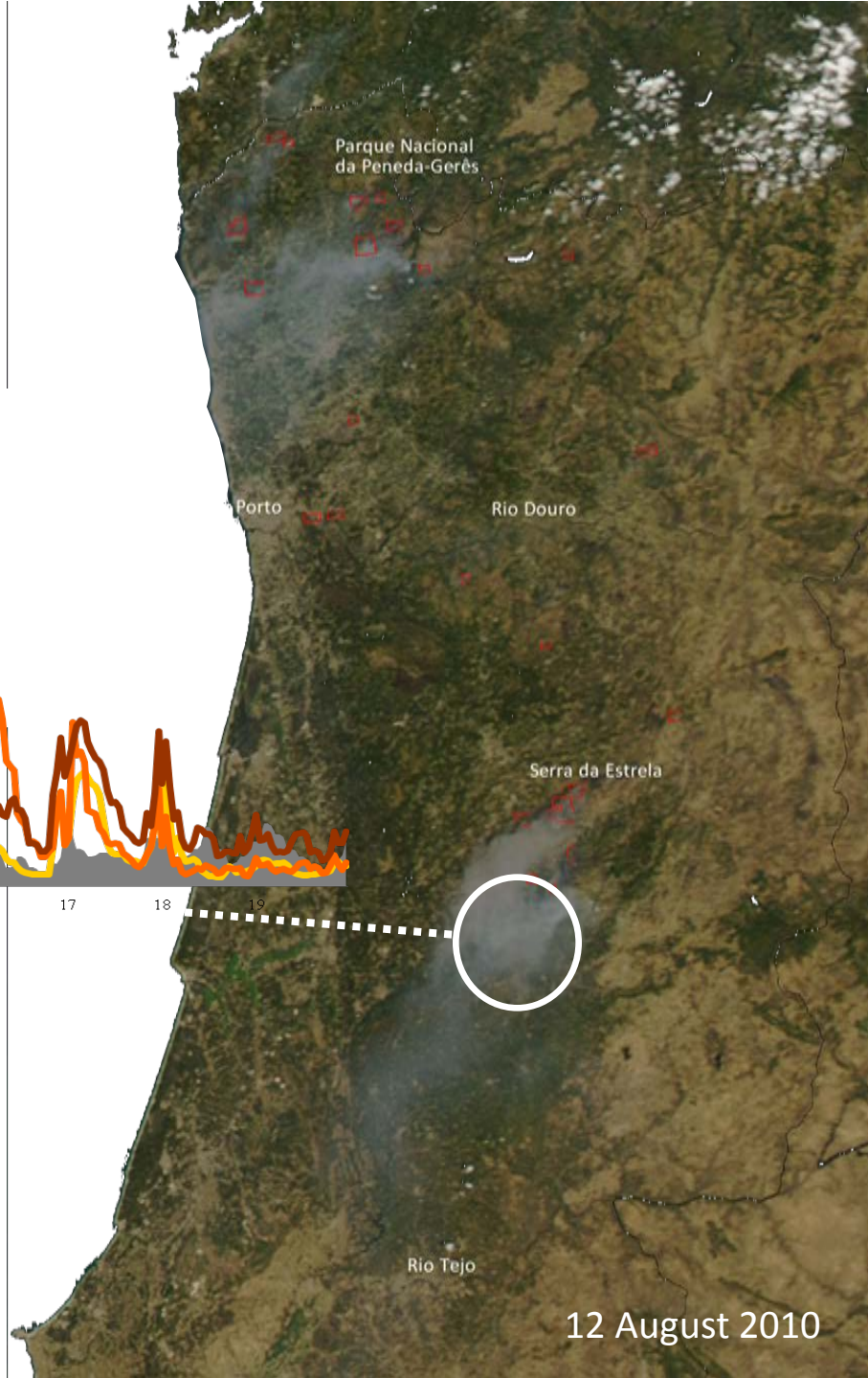
error on monitoring data (ex. calibration problems) can originate wrong bias-correction…

Were this bias-correction techniques able to improve the forecasting of all pollution episodes?

Forest fire episodes



Both techniques after the episode, worsted the forecast, because correction is based on previous 4 days



Comments

- Several **sources of problems**:
  no validated data;
  lack of measurements;
  measurement out of range, etc.

- **KF follows the model**: false modeled high peaks

- **RAT04 propagates error**: no modeled high peaks

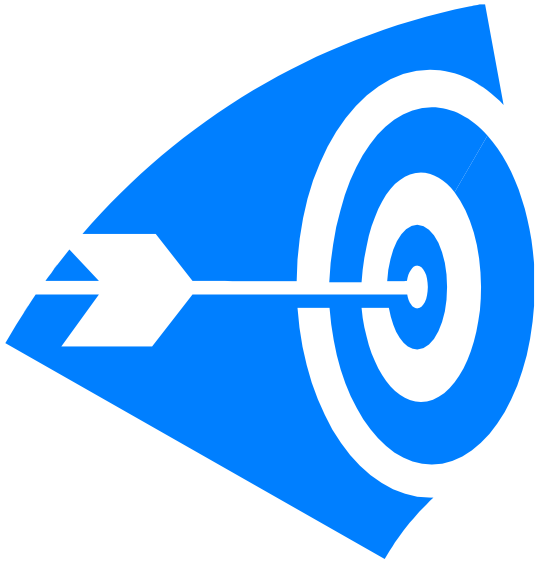
- KF is more robust than RAT04 when there is no data: bias is corrected faster in KF.

- Solutions can be found for each problem identified:

"No data": use short periods (< 4days) for bias-correction

"calibration problem": introduce criteria to identify this data problem…

Thank you for your attention!



universidade de aveiro

GEMAC

www.ua.pt/gemac