Ensemble techniques to improve air quality assessment

focus O_3 and PM over Portugal

A. Monteiro |I. Ribeiro| O. Tchepel|A. Carvalho
E. Sá|J. Ferreira | A.I. Miranda | C. Borrego
University of Aveiro, Portugal

S. Galmarini
JRC, Italy

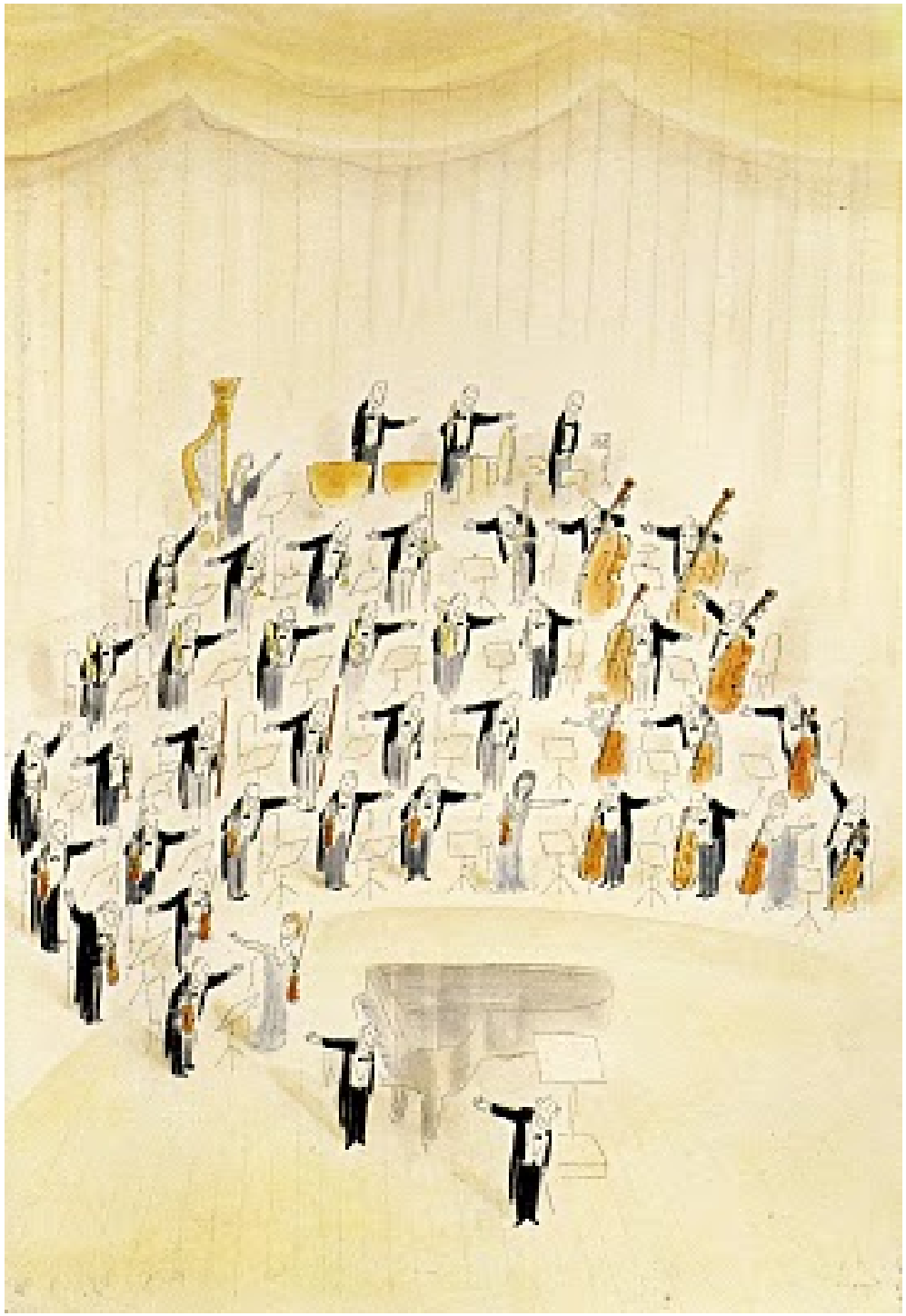
HARMO14 conference
Kos, Greece | October 2011

Models results sometimes are below our expectations…

How to improve model performance?
…there is hope!

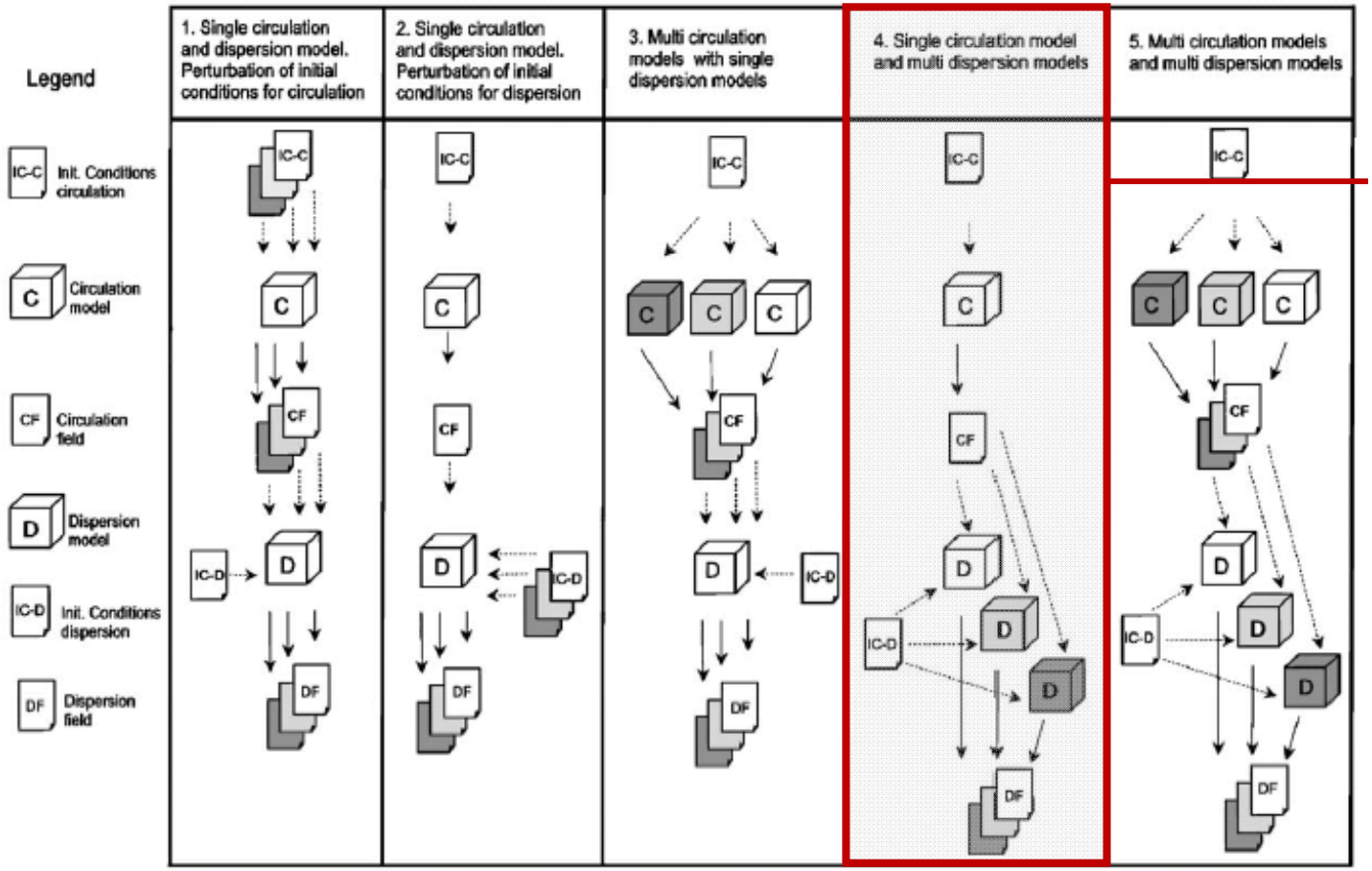
- using several **different models** ✓
- applying **bias correction** techniques ✓
- applying **ensemble** techniques ↙
- …

What type of ensemble?



ENSEMBLE approaches

- ensemble can be applied in different conceptual forms
- a single model and multiple inputs; or multimodel approach



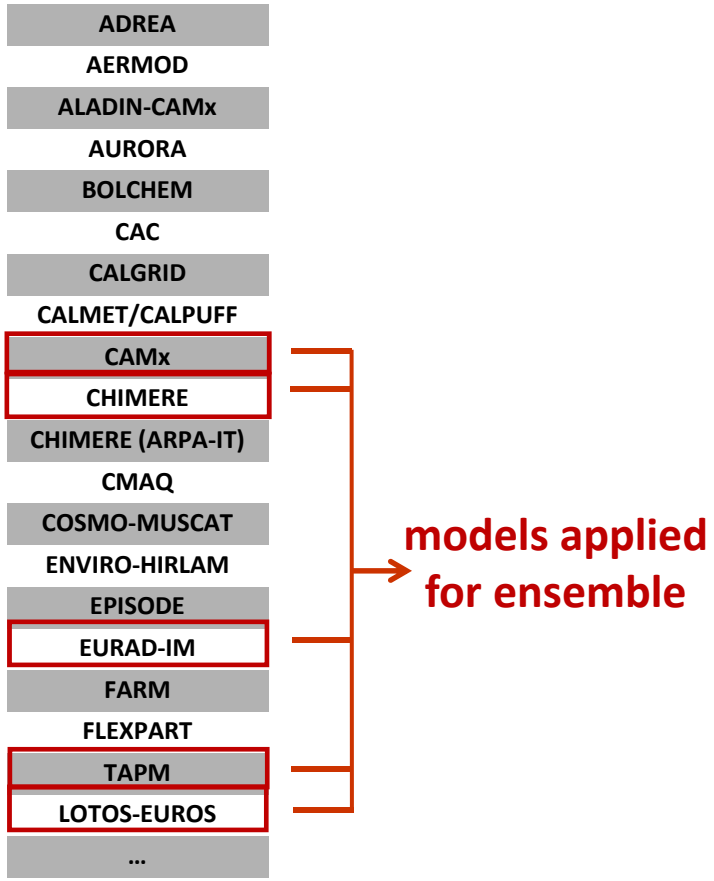
Same input data
Different CTM models

[Galmarini et al., 2004]

Air quality models

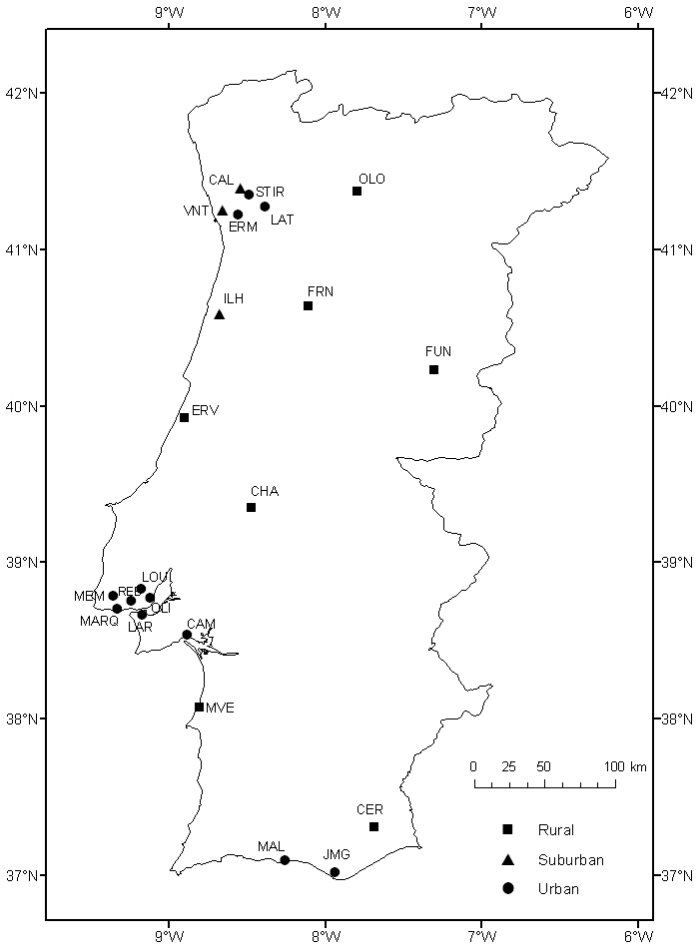
Several different models exist, with distinct:

- meteorological forcing
- emissions specification
- physical parameterizations
- chemical mechanisms
- aerosol formulation
- dry/wet deposition formulation
- etc...



Modelling setup

The different models were applied over Portugal, with high resolution
Data from 24 background stations were used for model validation



Domain
Portugal area

Resolution
5x5 km^2

Period
2006

Which ensemble techniques?

ENSEMBLE techniques

Median (MED)
Model weights are equal

Static Linear Regression (SLR)
Model weights are different but static in time

Dynamic Linear Regression (DLR)
Model weights are different and vary in time

Bayesian Model Averaging (BMA)
Model weights are different
Ensemble expressed as a probability density function (PDF)

Static Linear Regression (SLR)

Model weights are different but static in time

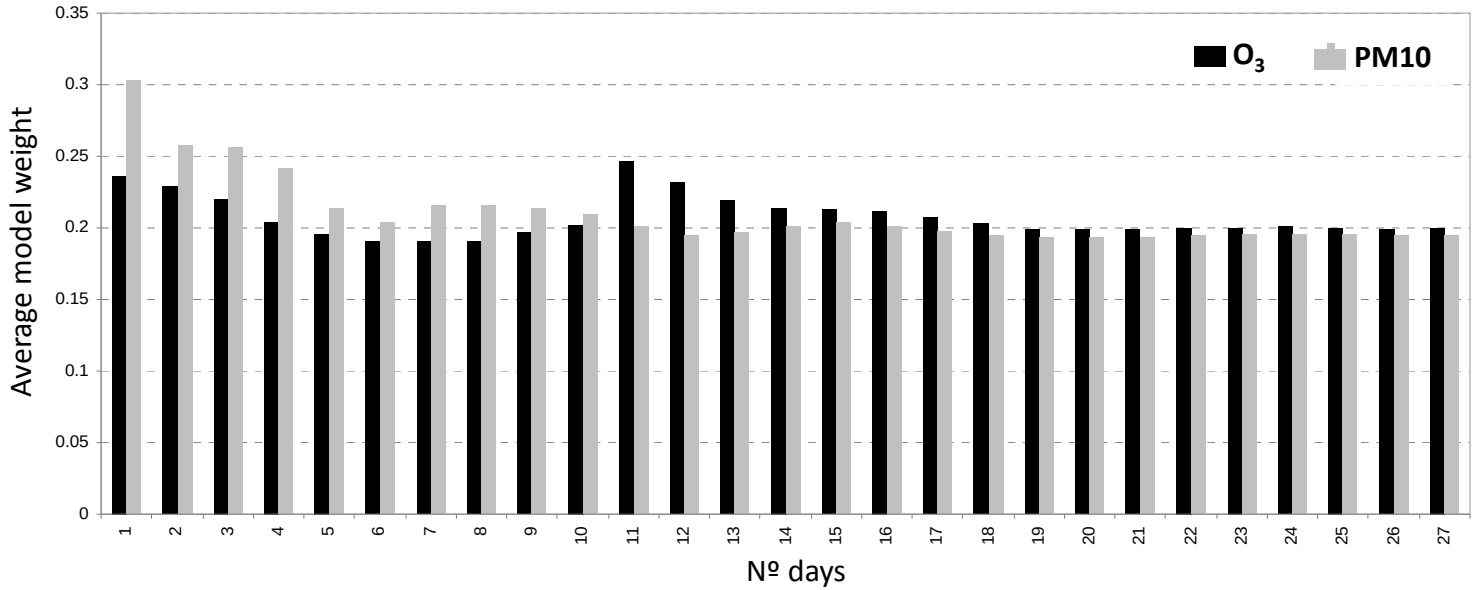
Weight (w_i) are found throughout linear combination…

$$\begin{bmatrix} m_{11} & m_{12} & \dots & m_{1J} \\ m_{21} & m_{22} & \dots & m_{2J} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ m_{I1} & m_{I2} & \dots & m_{IJ} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \\ \\ w_J \end{bmatrix} = \begin{bmatrix} o_1 \\ o_1 \\ \dots \\ \dots \\ o_I \end{bmatrix}$$

MODEL **OBS**

Static Linear Regression (SLR)

Influence of the training period on the model weights
(from 1 - 31 days)



Variability of weights decreases with training period length

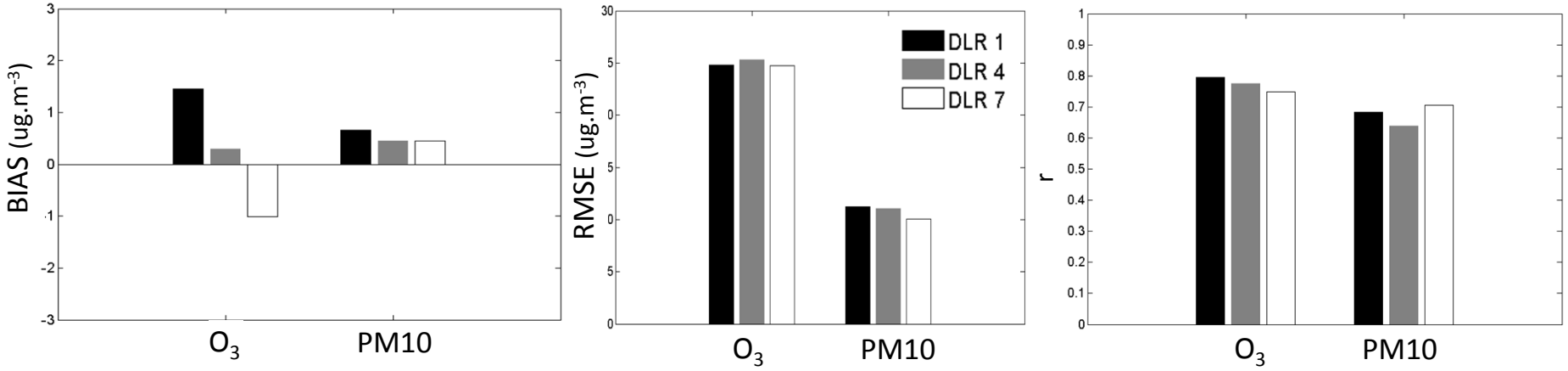
Weights vary significantly on first 10 days and little after 15 days of training

Dynamic Linear Regression (DLR)

Model weights are different and vary in time

Least square method

Different length of training periods was tested: 1, 4 and 7 previous days



No significant differences between the 3 training periods

Selection of DLR7 for ensemble comparisons

Bayesian Model Averaging (BMA)

BMA scheme describes the posterior probability density function (pdf) as a weighted average of probability distributions of individual models:

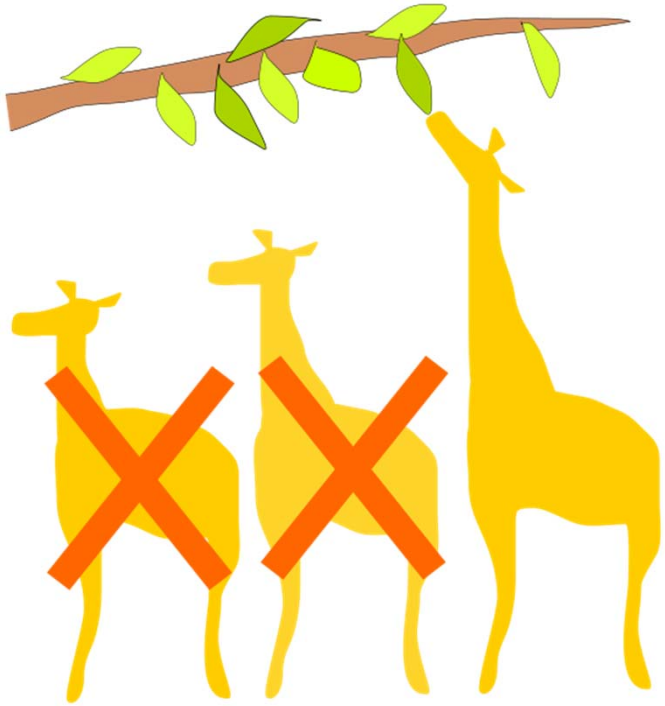
$$p(x \mid D) = \sum_{k=1}^{m} p_k(x \mid M_k, D)\, p(M_k \mid D)$$

w_k posterior probability of model
M_k best forecast in ensemble
p_k posterior probability that x occurs for model prediction M_k and observed O

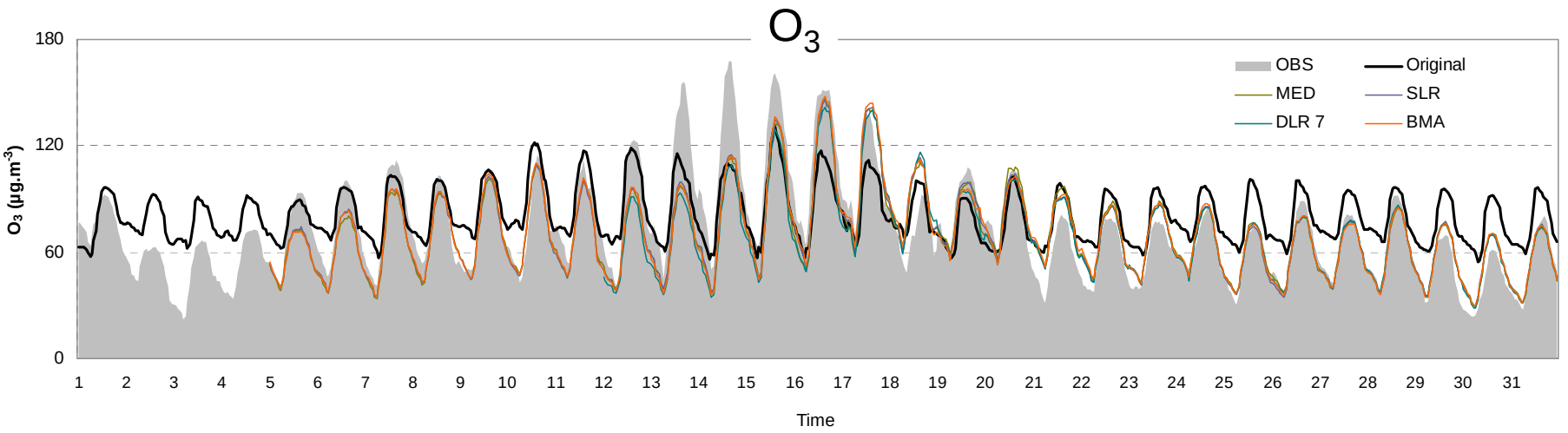
Comparison of model and observed pdf shows a good linearity for O_3

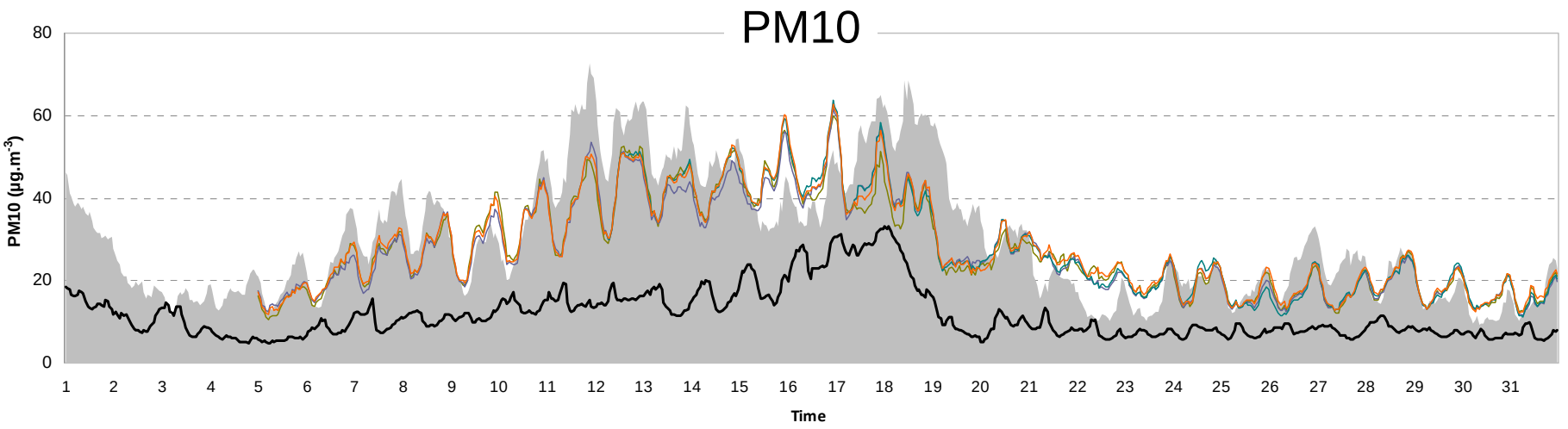
More complex behaviour is demonstrated for PM10

Which one is the
“best” technique?

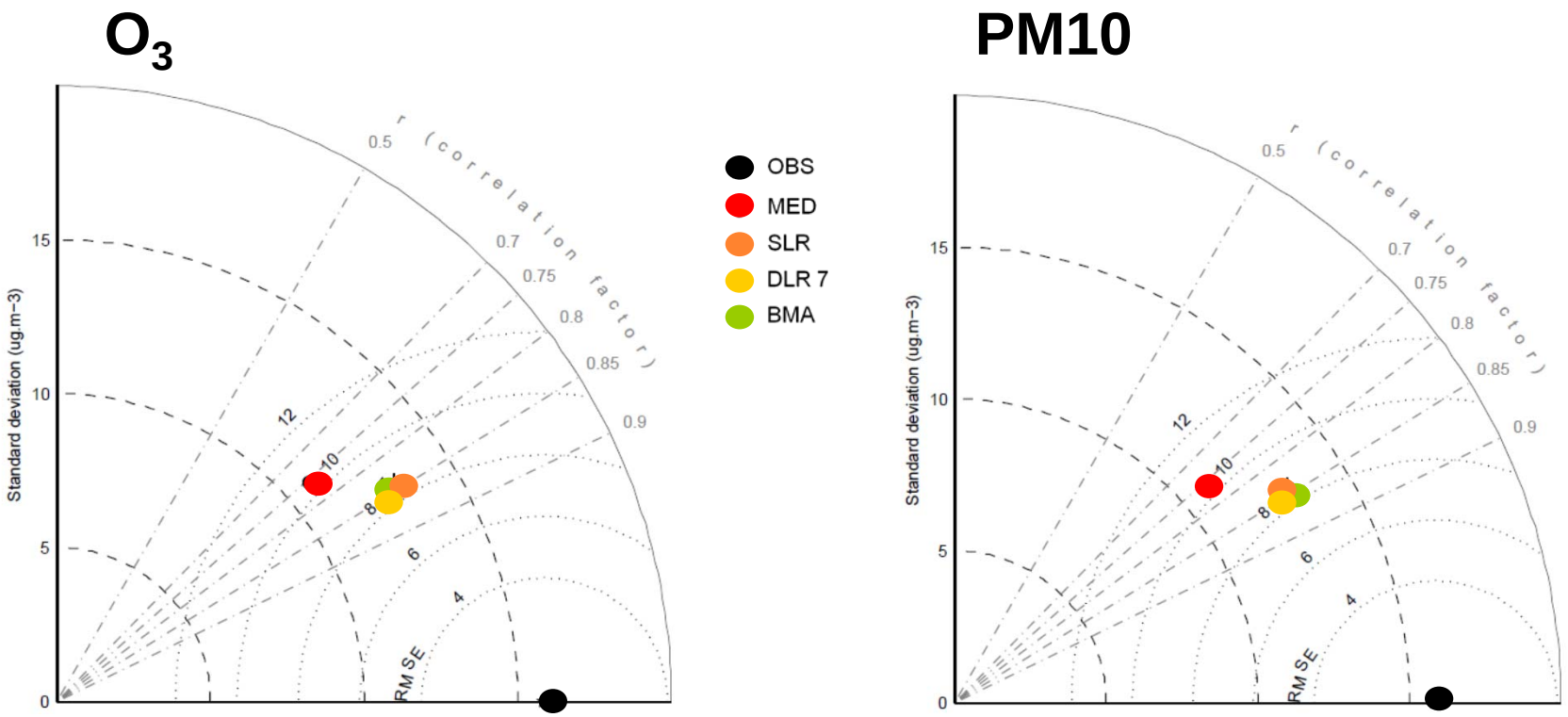


ENSEMBLE results | Time series

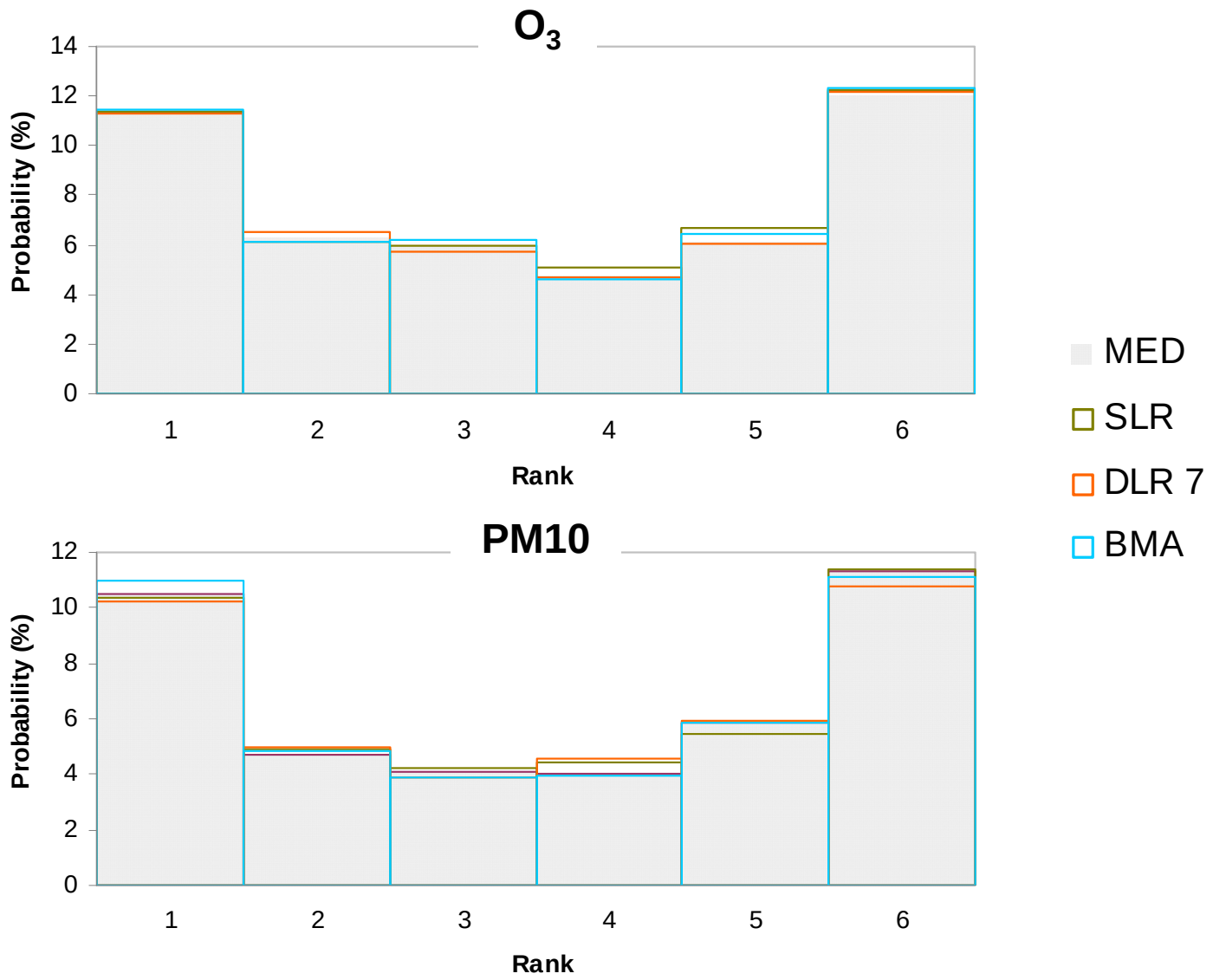




ENSEMBLE results | Taylor diagram



ENSEMBLE results | Rank histogram

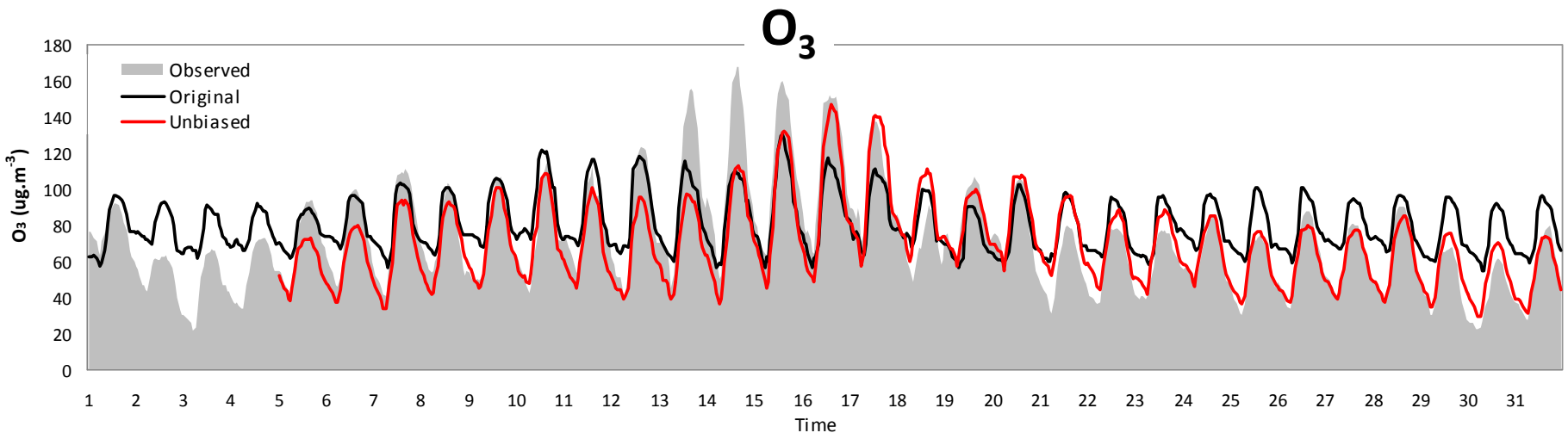


Is ensemble

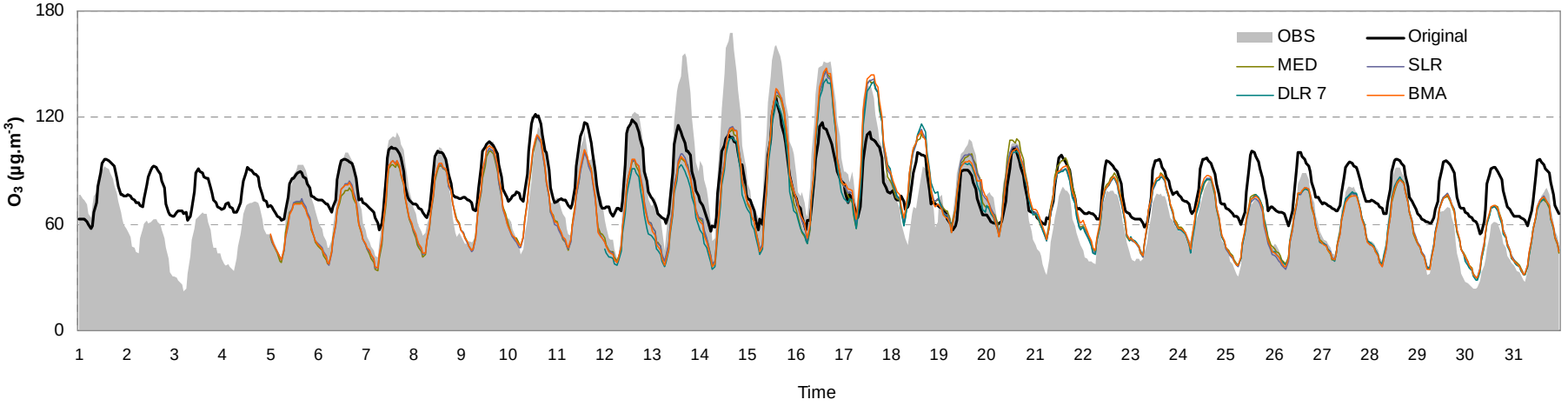
after bias correction

an added-value?

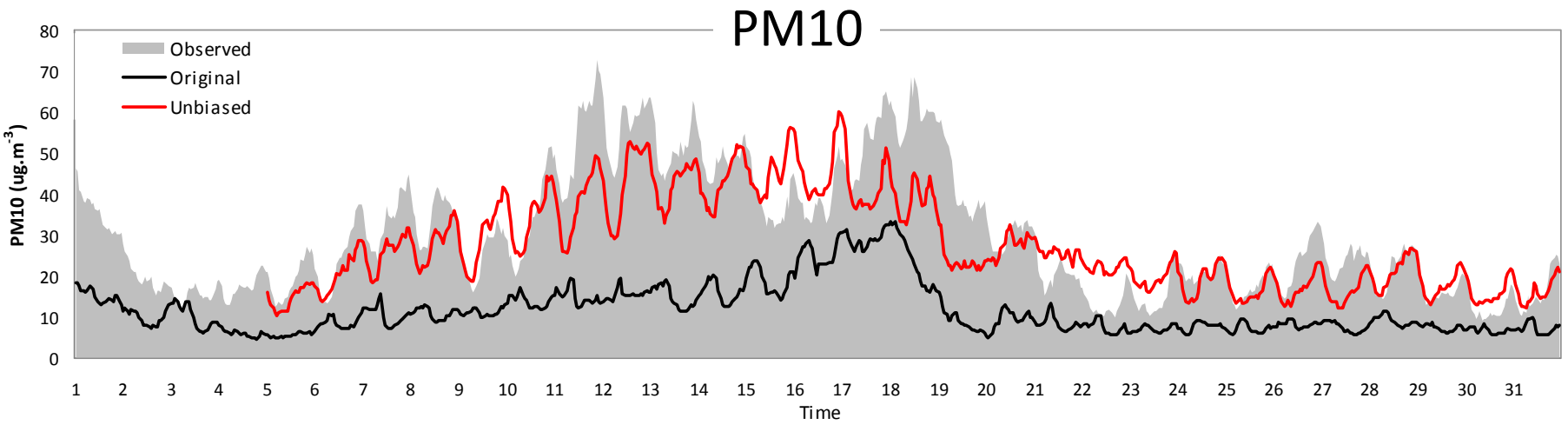
Bias correction



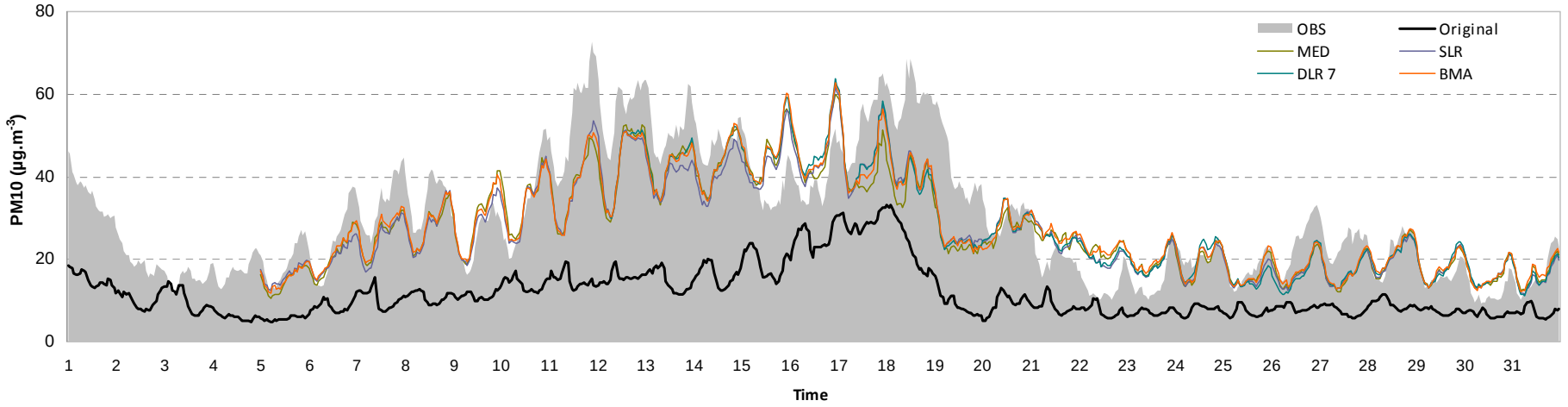
Bias correction + ensemble



Bias correction



Bias correction + ensemble



Final comments

- E**nsembles techniques** performed **similar** and better than single models
- Slight **improvement of weighted ensembles** compared to median
- Statistical analysis indicates **LR and BMA** ensembles are best “performers”
- The **SLR effortless implementation** can be an advantage
- **Ensemble** efforts are **not justified** for **bias-free** models


Thank you!