

Combination of model and observations for air quality assessment in Spain How can it be evaluated?

Fernando Martín

M.G. Vivanco, I. Palomino, J.L. Garrido

Atmospheric Pollution Division

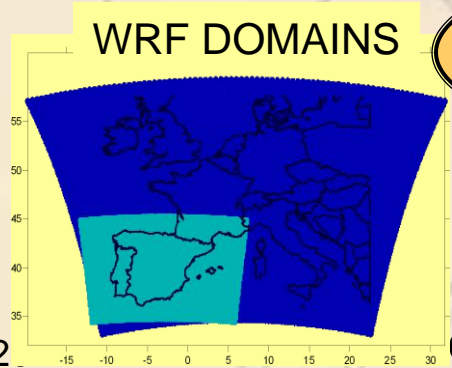
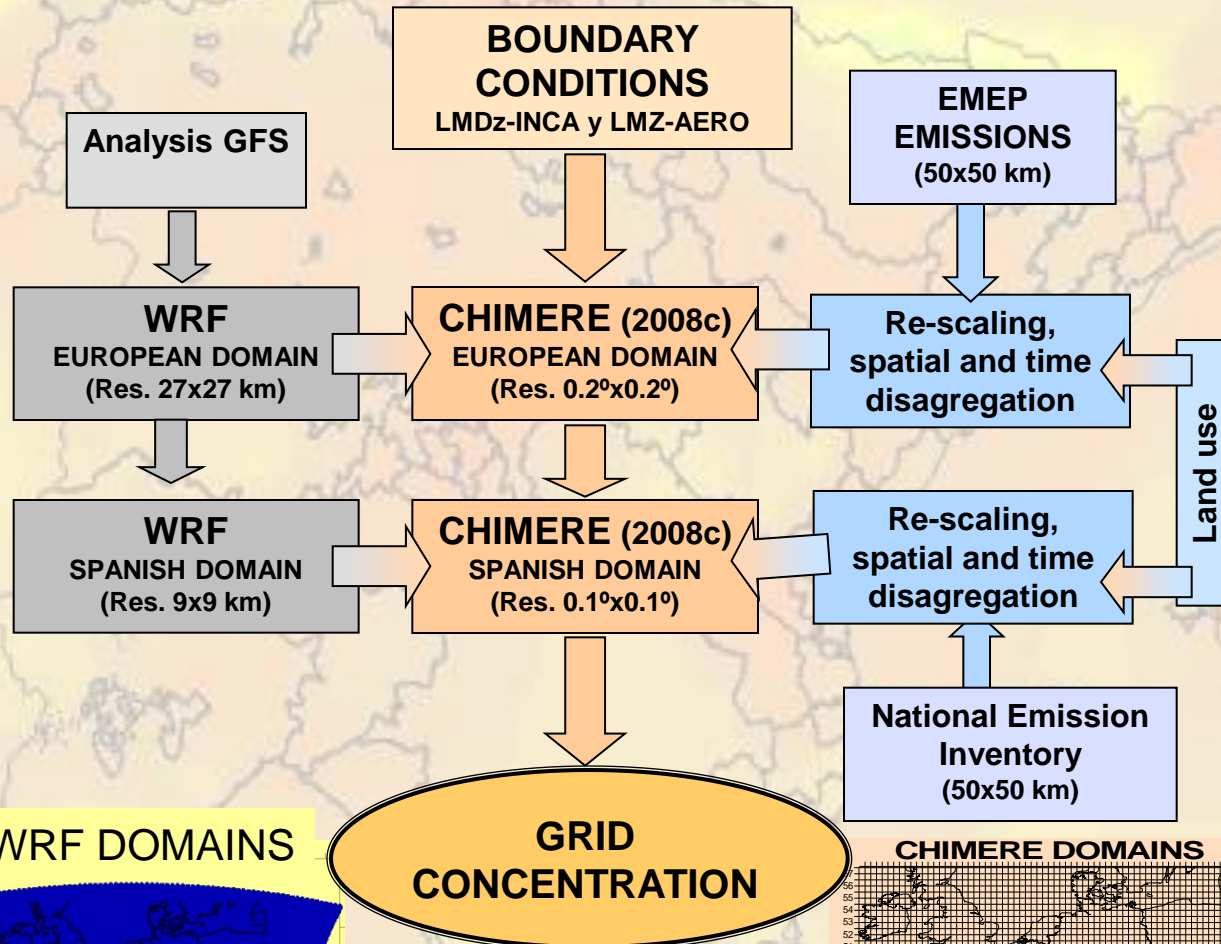
CIEMAT

Spain

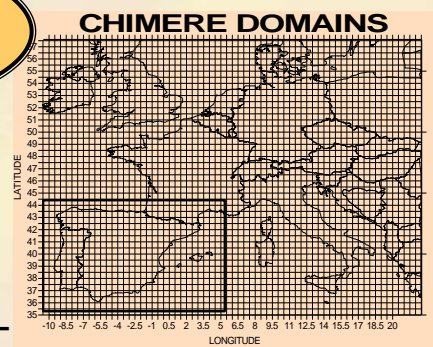
Introduction

- A methodology to combine measurements from air quality stations and estimates from the CHIMERE model for air quality assessment in Spain is described (*Martín et al, 2012, Int. J. Environment and Pollution, Vol. 49*)
- How can we measure the performance of the combination methodology?
- An simple exercise has been done for testing different methods.

Modeling scheme



**GRID
CONCENTRATION**



Combination of model and measured data

- How can we improve the air quality maps using the accuracy of the measurements and the good spatial coverage of the model outputs?

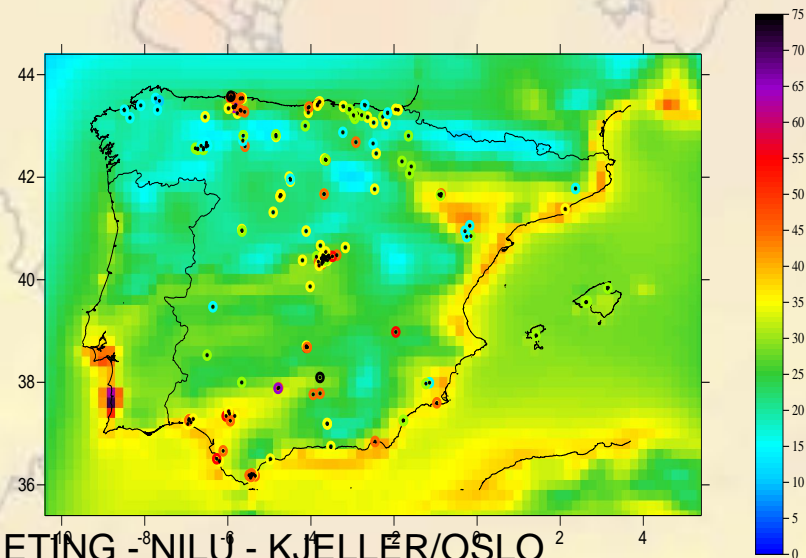
$$C_k = M_k + e_k + s_k$$

M_k = concentration estimate (i.e., by a dispersion model),

e_k = systematic error of the estimate (i.e., modelling error)

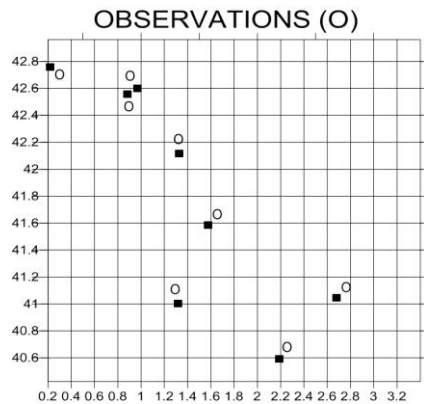
s_k = the inherent error or measurement error.

- ¿How to reduce e_k ?

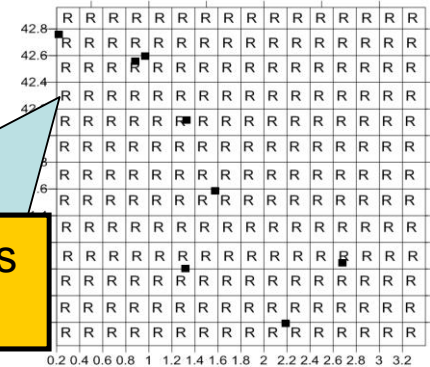


Combination of modelling and measured data.

Kriging

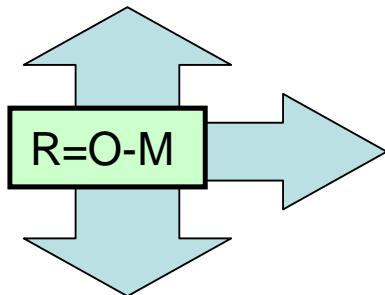
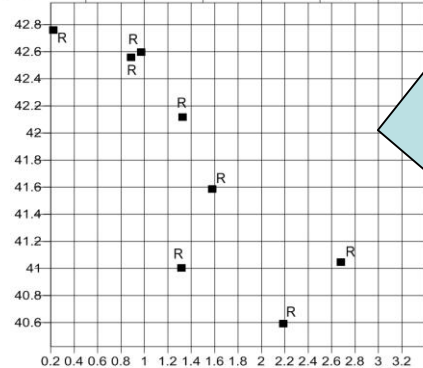


RESIDUALS IN CELLS (R=O-M)



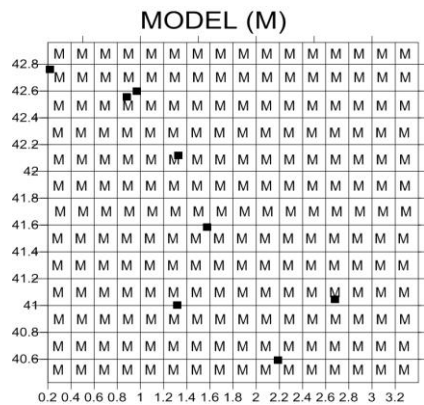
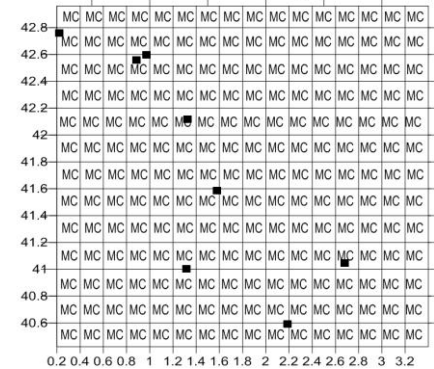
Interpolation of residuals using kriging

RESIDUALS IN STATIONS (R=O-M)

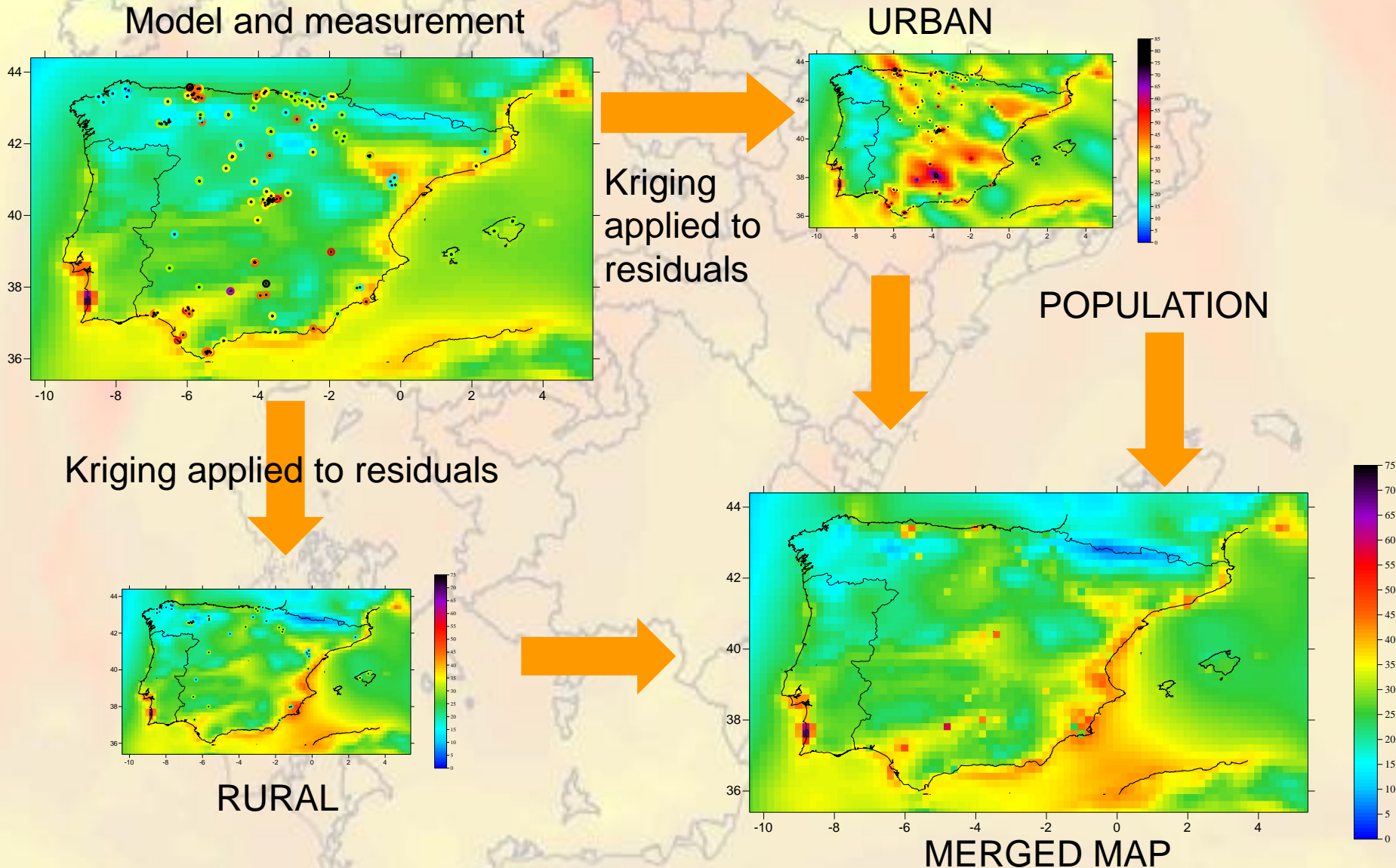


MC=M+R

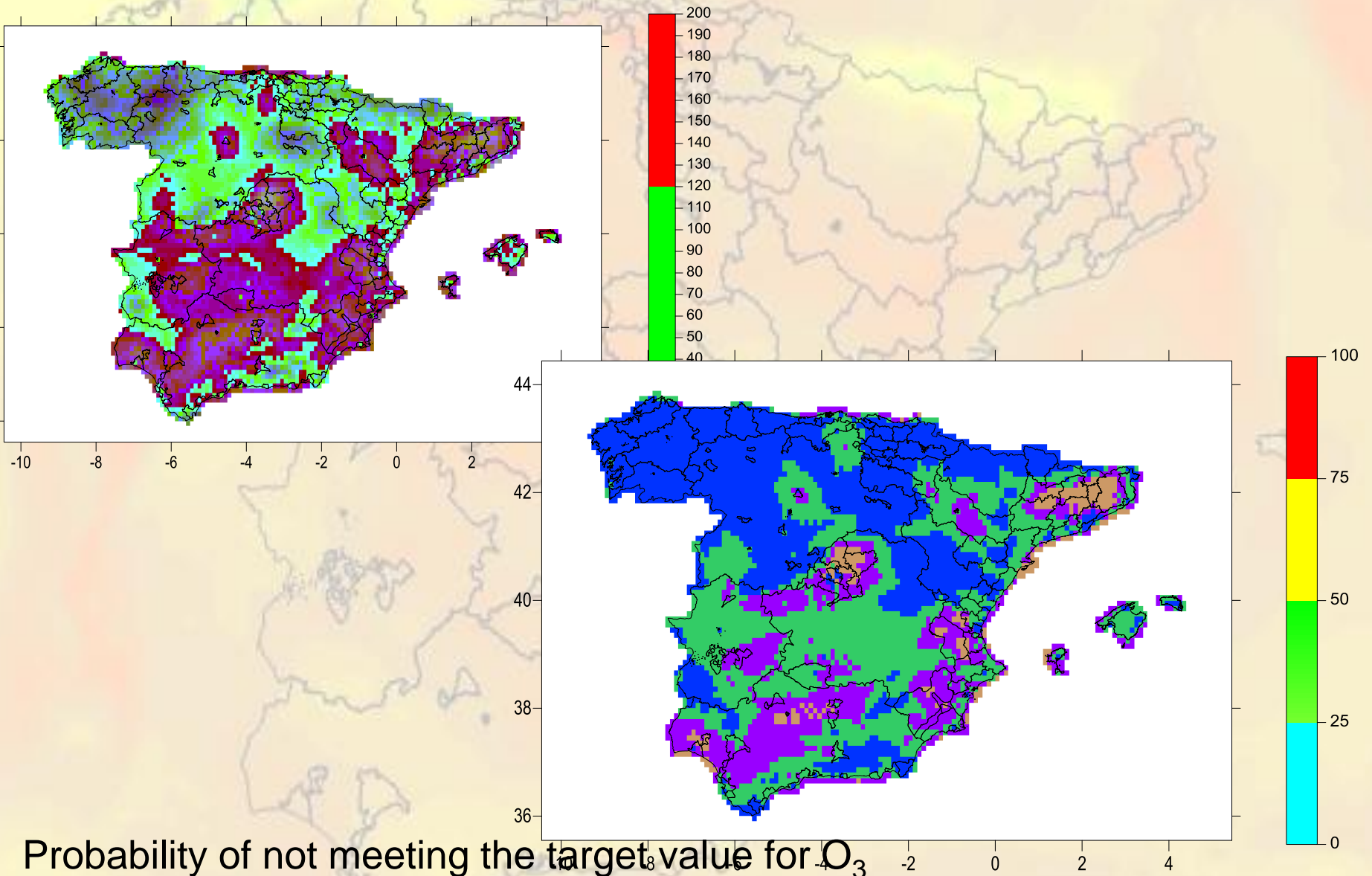
CORRECTED MODEL (MC=M+R)



Measurements and modeling combination



26th highest value of 8-hour O₃ concentration



Probability of not meeting the target value for O₃

Incertidumbre de la combinación de mediciones y modelos

- Directive Relative Error (RDE):

$$ERD = \frac{|O_{LV} - M_{LV}|}{LV}$$

- Máximum of RDE del Error Relativo de la Directiva (MERD).

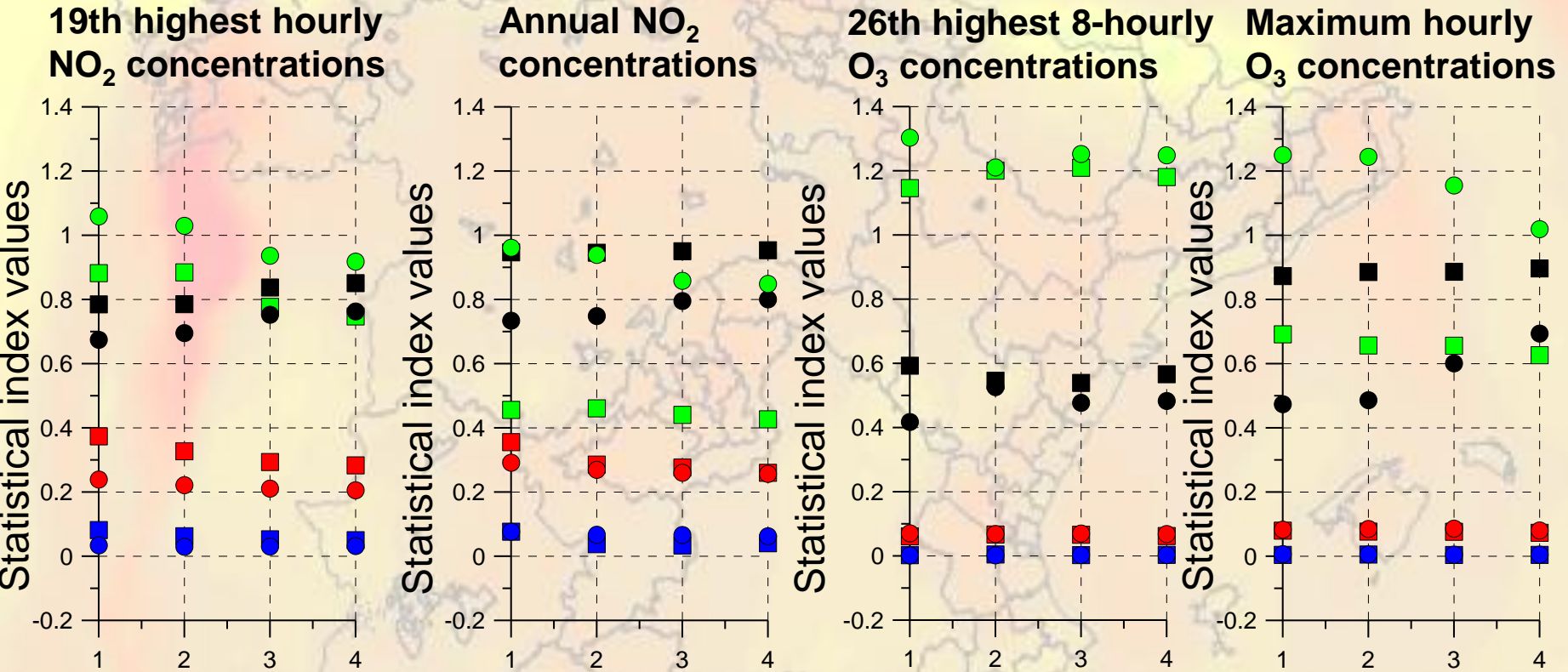
Reference value	MRDE Combination methodology	MRDE CHIMERE Model	Pollutant
Target value 120 $\mu\text{g m}^{-3}$ (eight-hour average)	0.1196	0.1570	O ₃
Information value 180 $\mu\text{g m}^{-3}$ (hourly average)	0.2056	0.2510	
Alert value 240 $\mu\text{g m}^{-3}$ (hourly average)	0.1542	0.2064	
Limit value 200 $\mu\text{g m}^{-3}$ (hourly average)	0.2315	0.3268	NO ₂
Limit value 40 $\mu\text{g m}^{-3}$ (annual average)	0.0549	0.3272	
Limit value 350 $\mu\text{g m}^{-3}$ (hourly average)	0.3288	0.5282	SO ₂
Limit value 125 $\mu\text{g m}^{-3}$ (daily average)	0.0804	0.2394	
Limit value 50 $\mu\text{g m}^{-3}$ (daily average)	0.2311	0.6217	PM10
Limit value 40 $\mu\text{g m}^{-3}$ (annual average)	0.1045	0.5224	

How do the statistics change when different stations are used for combination or for validation?

- Two methods for validation:
 - Leave-one-out
 - Selected set of data
- Several statistical index (R^2 , MFB, MFE, TARGET, etc).
- Several cases of data used for model-measurement combination and for validation
- Data of maps of air quality assessment for 2011 in Spain for O_3 and NO_2 .

STATISTICS vs DATA/METHOD FOR VALITATION

Leave-one-out method for validation



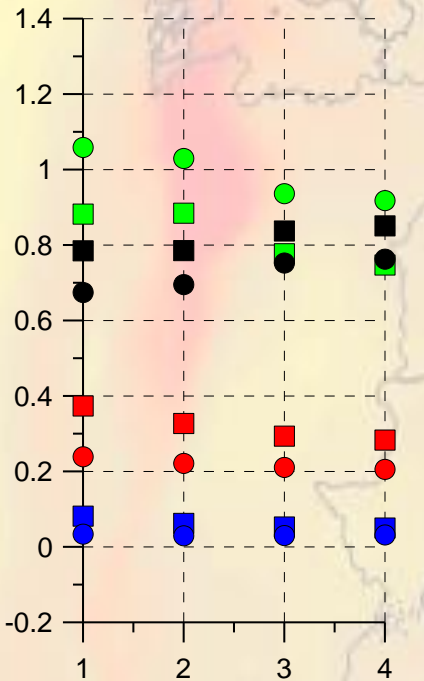
- ● R² (URBAN/SUBURBAN)
- ● MFB (URBAN/SUBURBAN)
- ● MFE (URBAN/SUBURBAN)
- ● TARGET (URBAN/SUBURBAN)
- ■ R² (RURAL)
- ■ MFB (RURAL)
- ■ MFE (RURAL)
- ■ TARGET (RURAL)

1 = 50% stations for combination and validation
 2 = 70% stations for combination and validation
 3 = 90% stations for combination and validation
 4 = 100% stations for combination and validation

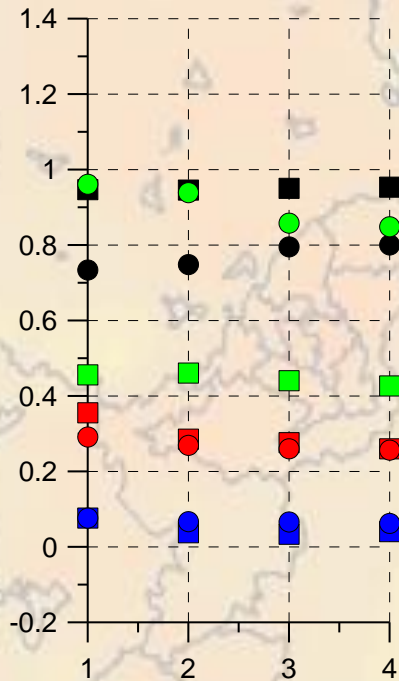
STATISTICS vs DATA/METHOD FOR VALITATION

Leave-one-out method for validation

19th highest hourly NO₂ concentrations



Annual NO₂ concentrations



26th highest 8-hourly Maximum hourly

Annual NO₂ :

- R for rural stations does not change, but for urban/suburban stations, more stations, better R.
- MFB for rural stations does not change, but for urban/suburban stations, slight improvement as amount of stations increases.
- MFE improves as increase amount of stations (clearer for rural stations).
- TARGET for urban/suburban stations does not change, but for rural stations, slight improvement as amount of stations increases.

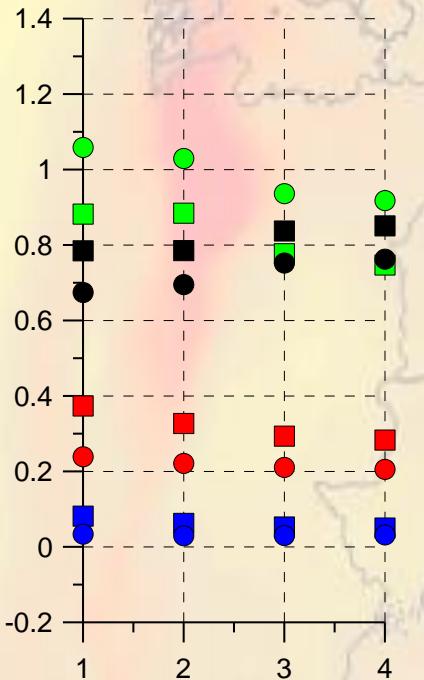
- R² (URBAN/SUBURBAN)
- MFB (URBAN/SUBURBAN)
- MFE (URBAN/SUBURBAN)
- TARGET (URBAN/SUBURBAN)
- R² (RURAL)
- MFB (RURAL)
- MFE (RURAL)
- TARGET (RURAL)

- 1 = 50% stations for combination and validation
- 2 = 70% stations for combination and validation
- 3 = 90% stations for combination and validation
- 4 = 100% stations for combination and validation

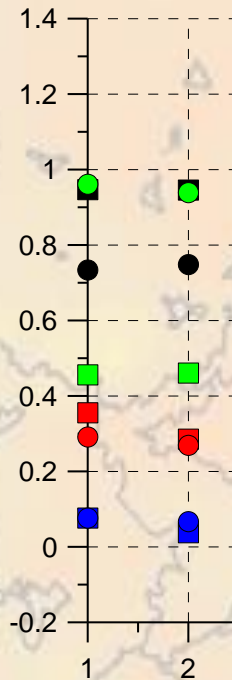
STATISTICS vs DATA/METHOD FOR VALITATION

Leave-one-out method for validation

19th highest hourly NO₂ concentrations



Annual NO₂ concentration



26th highest 8-hourly

Maximum hourly concentrations

19th highest hourly NO₂:

- R improves as increase amount of used stations.
- MFB for urban/suburban stations does not change, but for rural stations, slight improvement as amount of stations increases.
- MFE and TARGET improves as amount of stations increases (clearer for rural stations).

- ● ● R² (URBAN/SUBURBAN)
- ● ● MFB (URBAN/SUBURBAN)
- ● ● MFE (URBAN/SUBURBAN)
- ● ● TARGET (URBAN/SUBURBAN)
- ■ ■ R² (RURAL)
- ■ ■ MFB (RURAL)
- ■ ■ MFE (RURAL)
- ■ ■ TARGET (RURAL)

- 1 = 50% stations for combination and validation
- 2 = 70% stations for combination and validation
- 3 = 90% stations for combination and validation
- 4 = 100% stations for combination and validation

STATISTICS vs DATA/METHOD FOR VALIDATION

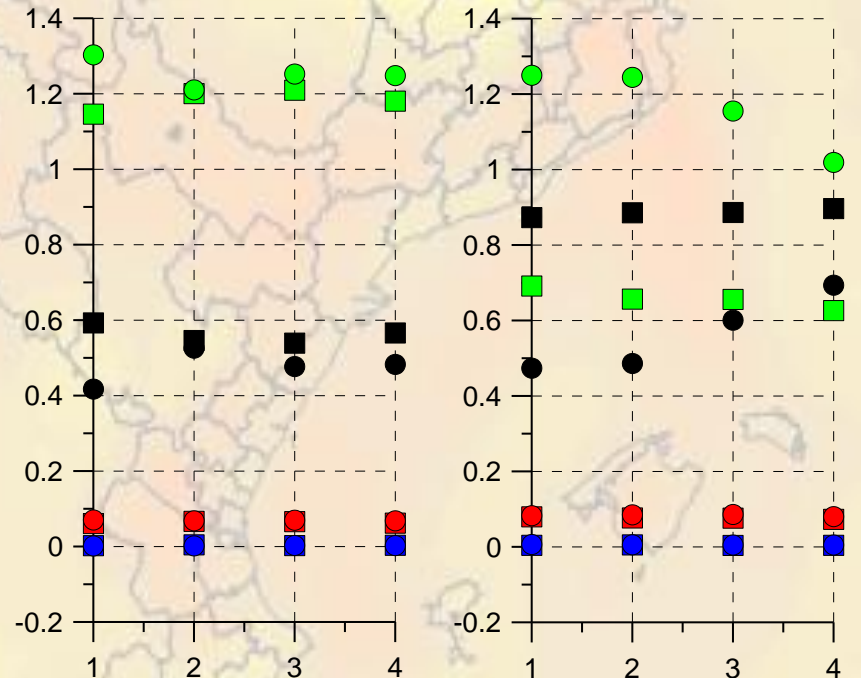
Leave-one-out method for validation

19th highest hourly Annual NO_x

26th highest 8-hourly O₃ concentrations Maximum hourly O₃ concentrations

26th highest 8-hourly O₃:

- R for rural stations, few differences, but best for 50% case, worst for 90% case.
- R for urban/suburban stations, worst for 50% case, best for 70% case.
- TARGET. For rural stations, best results for 50% case, few differences in other cases.
- TARGET. For urban/suburban stations, worst for 50% case, best for 70% case.
- TARGET. Few differences among 70%, 90% and 100% cases.



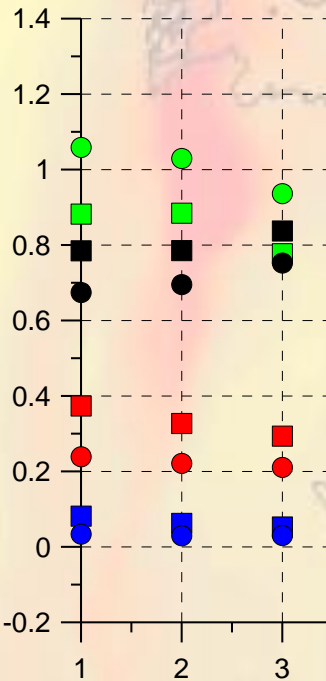
- ● ● TARGET (URBAN/SUBURBAN)
- ■ ■ R² (RURAL)
- ■ ■ MFB (RURAL)
- ■ ■ MFE (RURAL)
- ■ ■ TARGET (RURAL)

- 1 = 50% stations for combination and validation
- 2 = 70% stations for combination and validation
- 3 = 90% stations for combination and validation
- 4 = 100% stations for combination and validation

STATISTICS vs DATA/METHOD FOR VALITATION

Leave-one-out method for validation

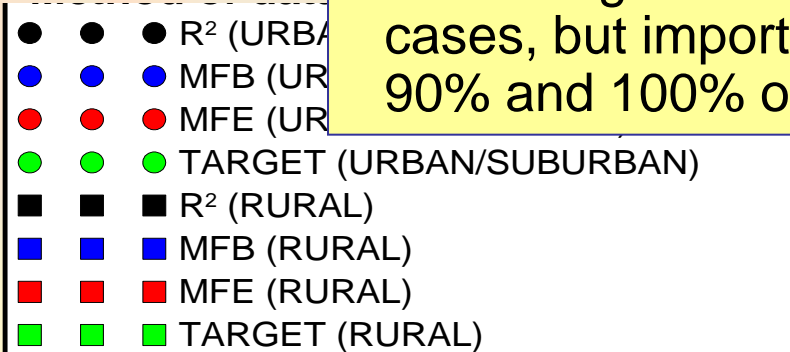
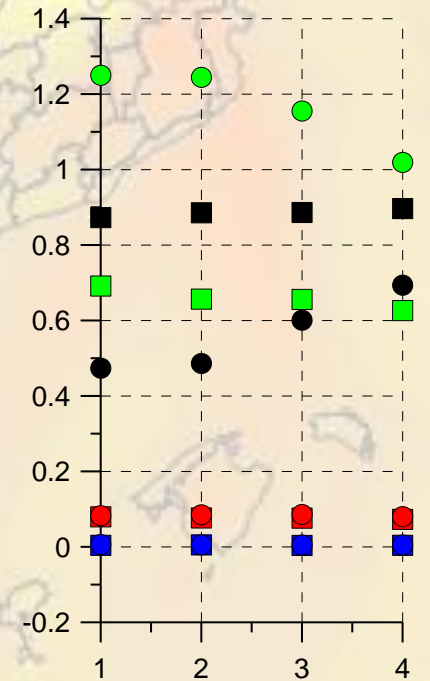
19th highest hourly NO₂ concentration



Maximum hourly O₃:

- R for rural stations, very slight improvement as amount of used stations increases.
- R for urban/suburban stations, significant improvement as amount of used stations increases.
- TARGET. For rural stations, slight improvement as amount of used stations increases.
- TARGET. For urban/suburban stations, few changes between 50% and 70% cases, but important improvement for 90% and 100% ones.

26th highest 8-hourly Maximum hourly O₃ concentrations



- 1 = 50% stations for combination and validation
- 2 = 70% stations for combination and validation
- 3 = 90% stations for combination and validation
- 4 = 100% stations for combination and validation

STATISTICS vs DATA/METHOD FOR VALIDATION

Leave-one-out method for validation

Comments:

- More stations used for combination, better validation statistics (generally).
- Better R and TARGET in rural stations
- For NO₂, statistics for cases of 90% and 100% stations are similar.
- For O₃, MFB and MFE does not change with the amount of used stations.

STATISTICS vs DATA/METHOD FOR VALITATION

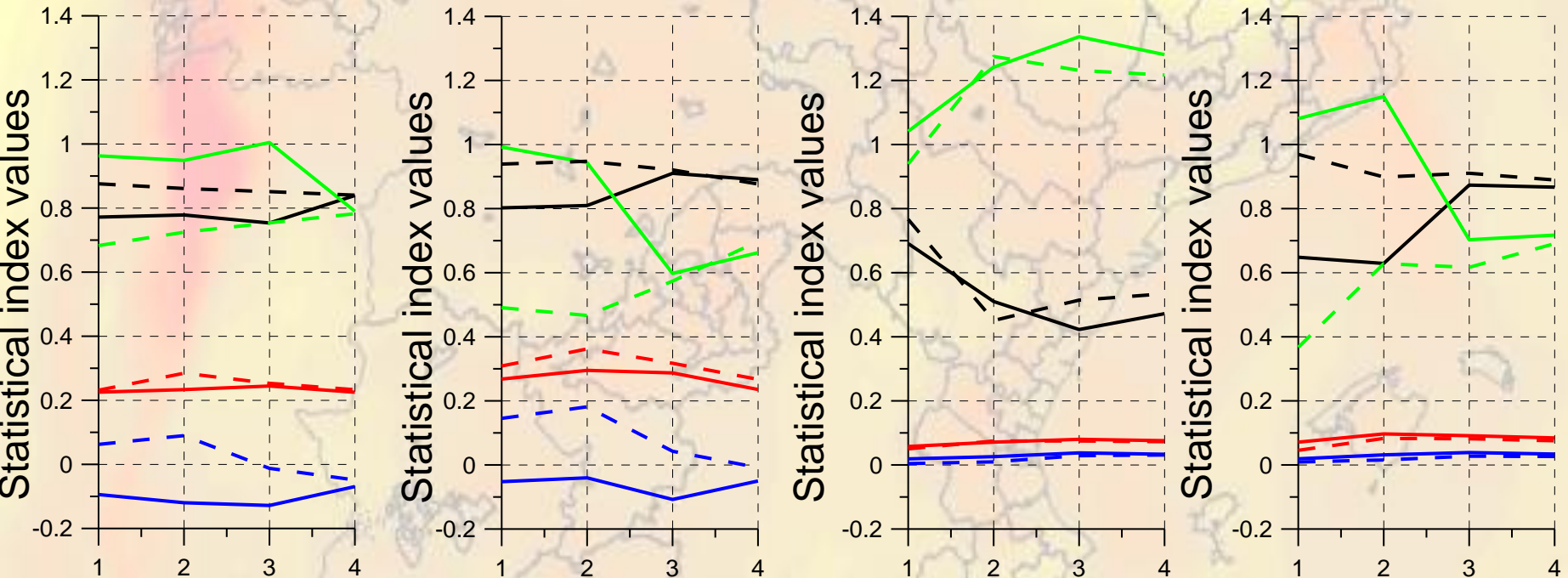
Set of selected data for validation

19th highest hourly NO₂ concentrations

Annual NO₂ concentrations

26th highest 8-hourly O₃ concentrations

Maximum hourly O₃ concentrations



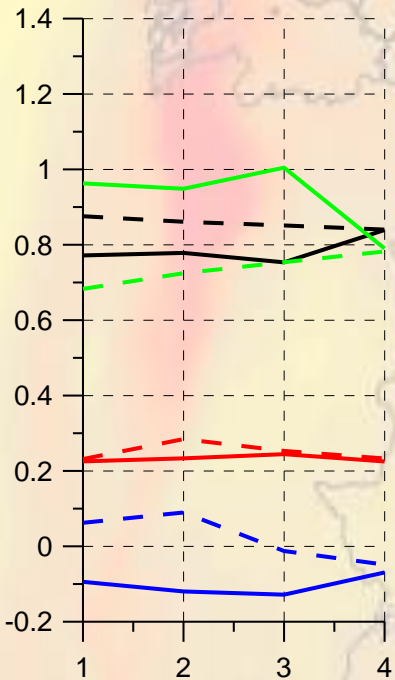
- R² (URBAN/SUBURBAN)
- MFB (URBAN/SUBURBAN)
- MFE (URBAN/SUBURBAN)
- TARGET (URBAN/SUBURBAN)
- - R² (RURAL)
- - MFB (RURAL)
- - MFE (RURAL)
- - TARGET (RURAL)

- Cases**
- 1 = 100% stations for combination and validation
 - 2 = 50% stations for combination and the other 50% for validation
 - 3 = 70% stations for combination and the other 30% for validation
 - 4 = 90% stations for combination and the other 10% for validation

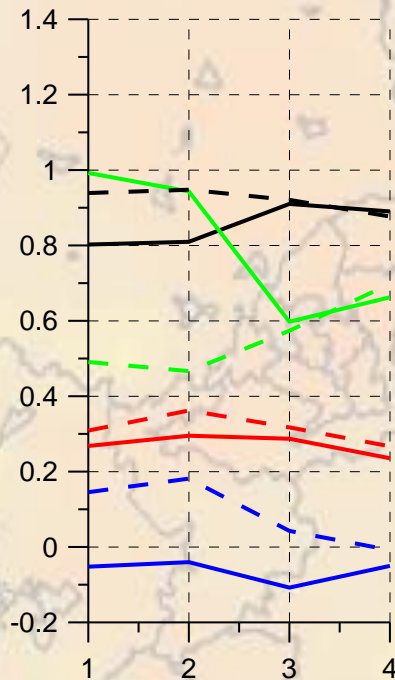
STATISTICS vs DATA/METHOD FOR VALITATION

Set of selected data for validation

19th highest hourly NO₂ concentrations



Annual NO₂ concentrations



Annual NO₂ :

- R for rural stations, slight decrease as used stations for combination increase and for validation decrease.
- R for urban/suburban stations, best results for cases 3 and 4.
- MFB for urban/suburban stations few changes
- MFB for rural stations, best results for cases 3 and 4.
- MFE. Less differences for urban/suburban stations than for rural ones. Worst results for case 2, best for 4.
- TARGET. Very different performance between rural (better) and urban/suburban stations for cases 1 and 2. Similar results for cases 3 and 4.

- R² (URBAN/SUBURBAN)
- MFB (URBAN/SUBURBAN)
- MFE (URBAN/SUBURBAN)
- TARGET (URBAN/SUBURBAN)
- R² (RURAL)
- MFB (RURAL)
- MFE (RURAL)
- TARGET (RURAL)

1 = 100%

2 = 50% stations for combination and the other 50% for validation

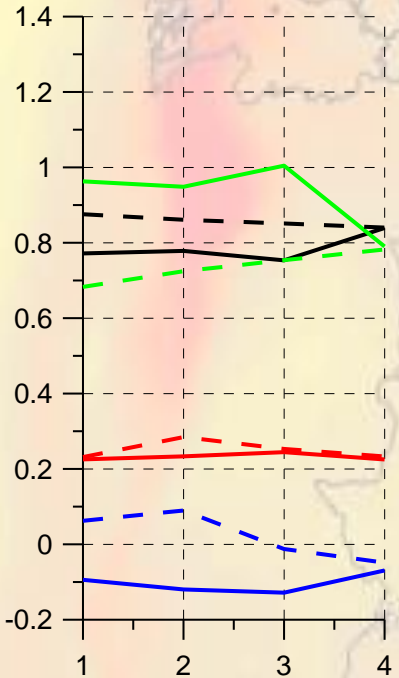
3 = 70% stations for combination and the other 30% for validation

4 = 90% stations for combination and the other 10% for validation

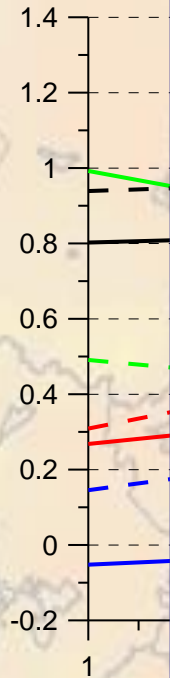
STATISTICS vs DATA/METHOD FOR VALITATION

Set of selected data for validation

19th highest hourly NO₂ concentrations



Ann
conc



19th highest hourly NO₂:

- R for rural stations almost does not change.
- R for urban/suburban stations does not change for cases 1-3. Better for case 4.
- MFB for rural stations, best for case 3. Case 4 underprediction.
- MFB for urban/suburban stations, few changes (underprediction). Best for case 4.
- MFE for urban/suburban stations, no changes, but for rural, some changes with worst result for case 2.
- TARGET for rural stations becomes worse from case 1 to case 4.
- TARGET for urban/suburban stations is much better for case 4.

- R² (URBAN/SUBURBAN)
- MFB (URBAN/SUBURBAN)
- MFE (URBAN/SUBURBAN)
- TARGET (URBAN/SUBURBAN)
- - R² (RURAL)
- - MFB (RURAL)
- - MFE (RURAL)
- - TARGET (RURAL)

- 1 = 100% stations for combination and validation
- 2 = 50% stations for combination and the other 50% for validation
- 3 = 70% stations for combination and the other 30% for validation
- 4 = 90% stations for combination and the other 10% for validation

STATISTICS vs DATA/METHOD FOR VALITATION

Set of selected data for validation

19th highest hourly

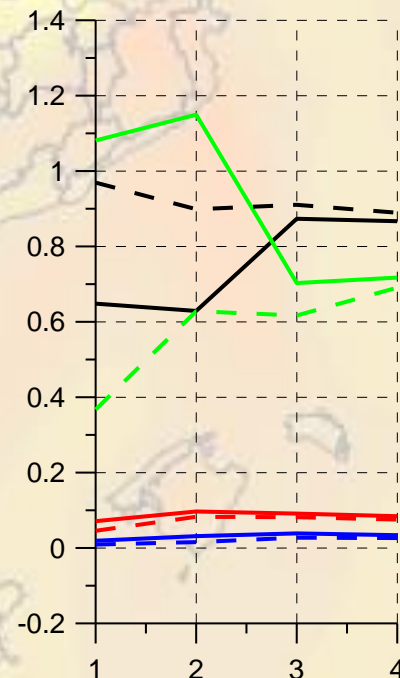
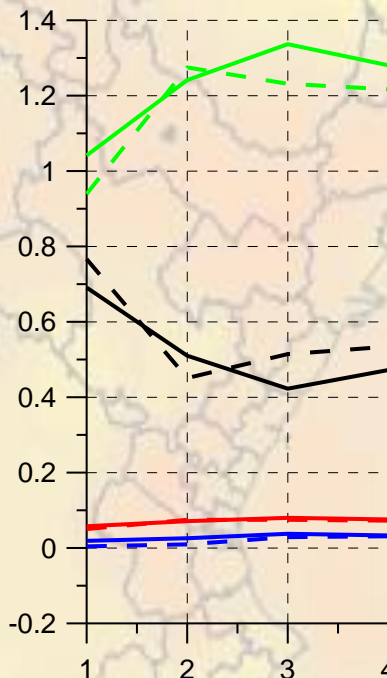
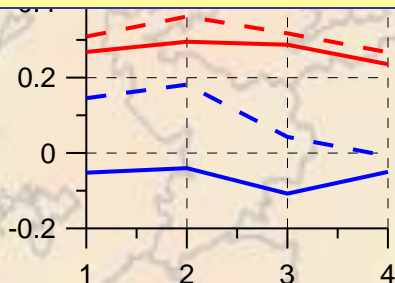
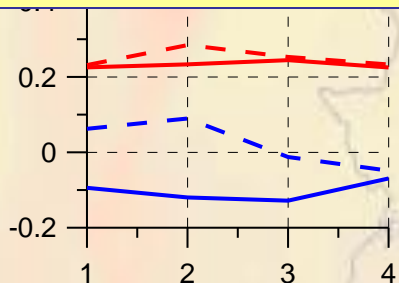
Annual NO_x

26th highest 8-hourly O₃ concentrations

Maximum hourly O₃ concentrations

26th highest 8-hourly O₃:

- R and TARGET for rural stations, improvement from case 2 to case 4.
- R and TARGET for urban/suburban stations, worse results for case 3.
- Few changes in MFB and MFE. Slightly better for rural stations.



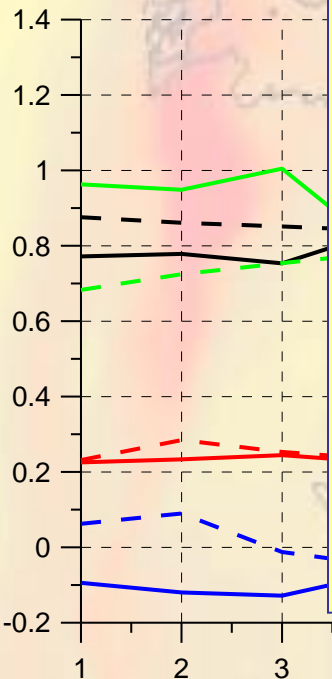
- R² (URBAN/SUBURBAN)
- MFB (URBAN/SUBURBAN)
- MFE (URBAN/SUBURBAN)
- TARGET (URBAN/SUBURBAN)
- - R² (RURAL)
- - MFB (RURAL)
- - MFE (RURAL)
- - TARGET (RURAL)

- 1 = 100% stations for combination and validation
- 2 = 50% stations for combination and the other 50% for validation
- 3 = 70% stations for combination and the other 30% for validation
- 4 = 90% stations for combination and the other 10% for validation

STATISTICS vs DATA/METHOD FOR VALITATION

Set of selected data for validation

19th highest hourly NO₂ concentration



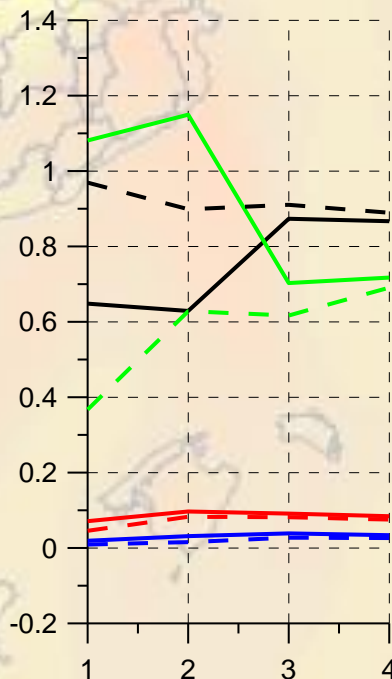
Annual NO₂

Maximum hourly O₃:

- R for rural stations, slight changes.
- R for urban/suburban stations, clearly best results for cases 3 and 4.
- TARGET. For rural stations, small changes from case 2 to case 4 (worst).
- TARGET. For urban/suburban stations, much better for case 3 and 4.
- Few changes in MFB and MFE. Slightly better for rural stations.

26th highest 8-hourly

Maximum hourly O₃ concentrations



- R² (URBAN/SUBURBAN)
- MFB (URBAN/SUBURBAN)
- MFE (URBAN/SUBURBAN)
- TARGET (URBAN/SUBURBAN)
- - R² (RURAL)
- - MFB (RURAL)
- - MFE (RURAL)
- - TARGET (RURAL)

- 1 = 100% stations for combination and validation
- 2 = 50% stations for combination and the other 50% for validation
- 3 = 70% stations for combination and the other 30% for validation
- 4 = 90% stations for combination and the other 10% for validation

STATISTICS vs DATA/METHOD FOR VALIDATION

Set of selected data for validation

Comments:

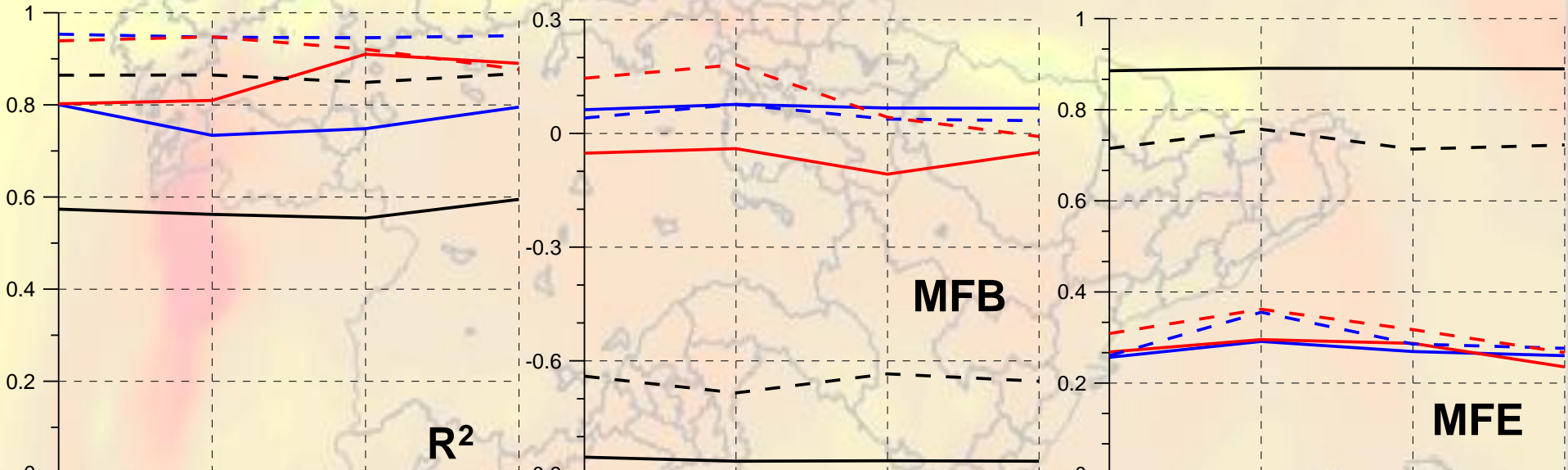
- More stations used for combination, better validation statistics (generally), but not in some cases (O_3 with R and TARGET). Of course, the size of the selected data for validation is an important factor!!
- Better R and TARGET in rural stations
- For O_3 , MFB and MFE does not change with the amount of used stations.
- For O_3 , MFB and MFE slightly better for rural stations



How different are the statistics from one method to other one?

STATISTICS vs DATA/METHOD FOR VALITATION

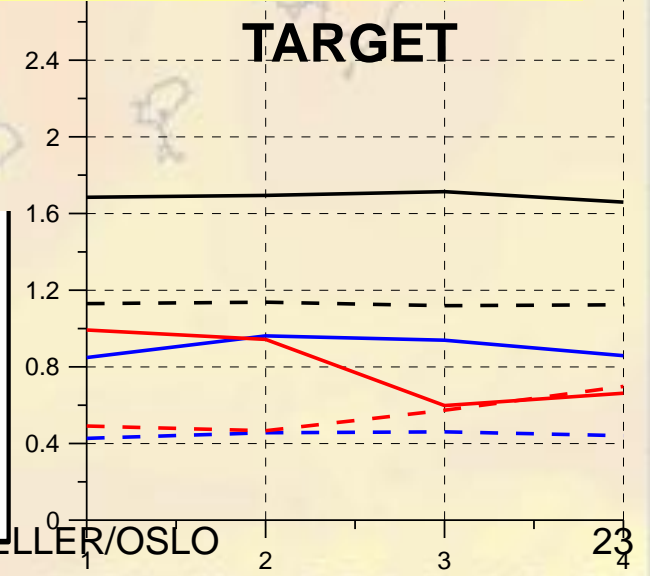
Annual NO₂ concentrations



MFE very similar for leave-one-out and for different selected data for validation

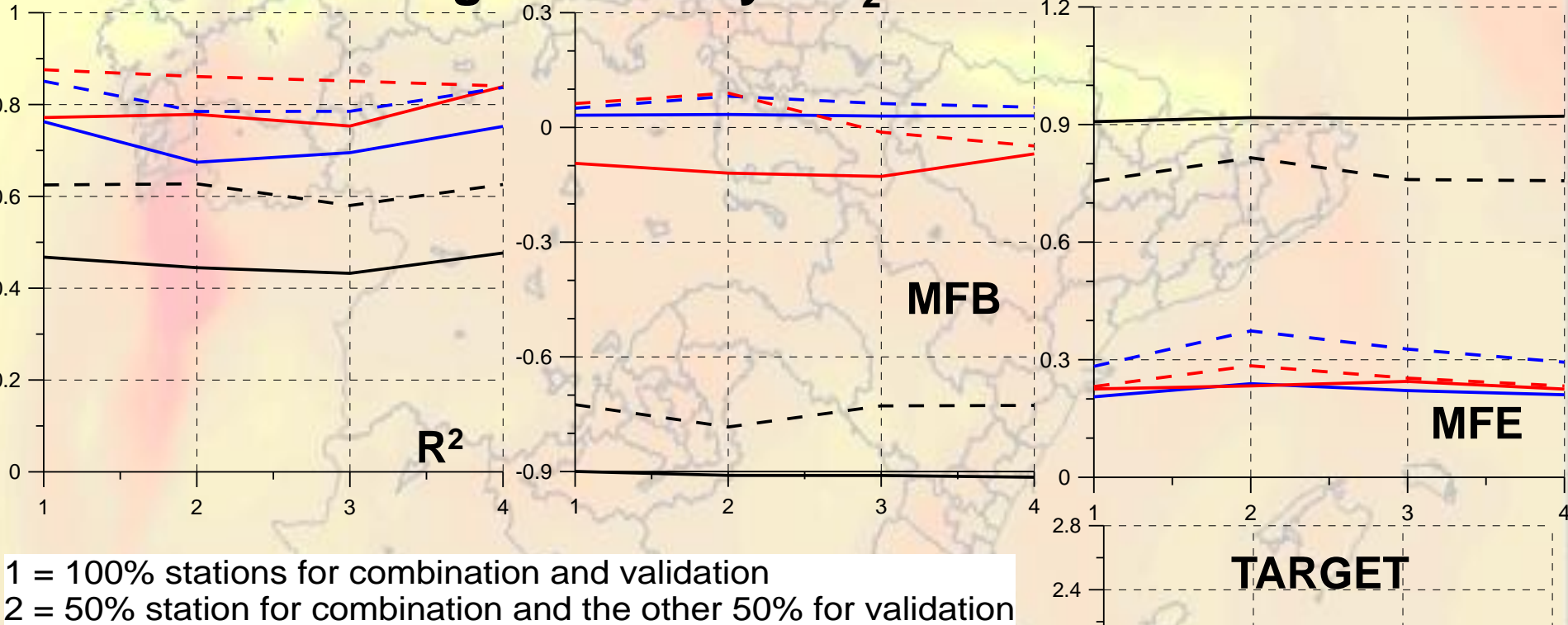
- 1 = 100% stations for combination and validation
- 2 = 50% station for combination and the other 50% for validation
- 3 = 70% station for combination and the other 30% for validation
- 4 = 90% station for combination and the other 10% for validation

- Model Urban-Suburban stations
- Combined model-observation/leave-one-out validation method/urban-suburban
- Combined model-observation/some stations for combination/urban-suburban
- - - Model Rural stations
- - - Combined model-observation/leave-one-out validation method/rural
- - - Combined model-observation/some stations for combination/rural



STATISTICS vs DATA/METHOD FOR VALITATION

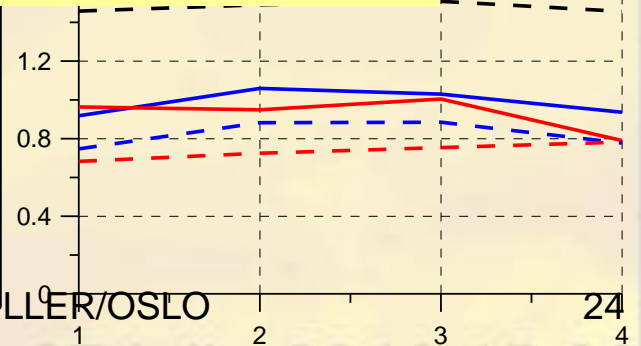
19th highest hourly NO₂ concentrations



1 = 100% stations for combination and validation
 2 = 50% station for combination and the other 50% for validation

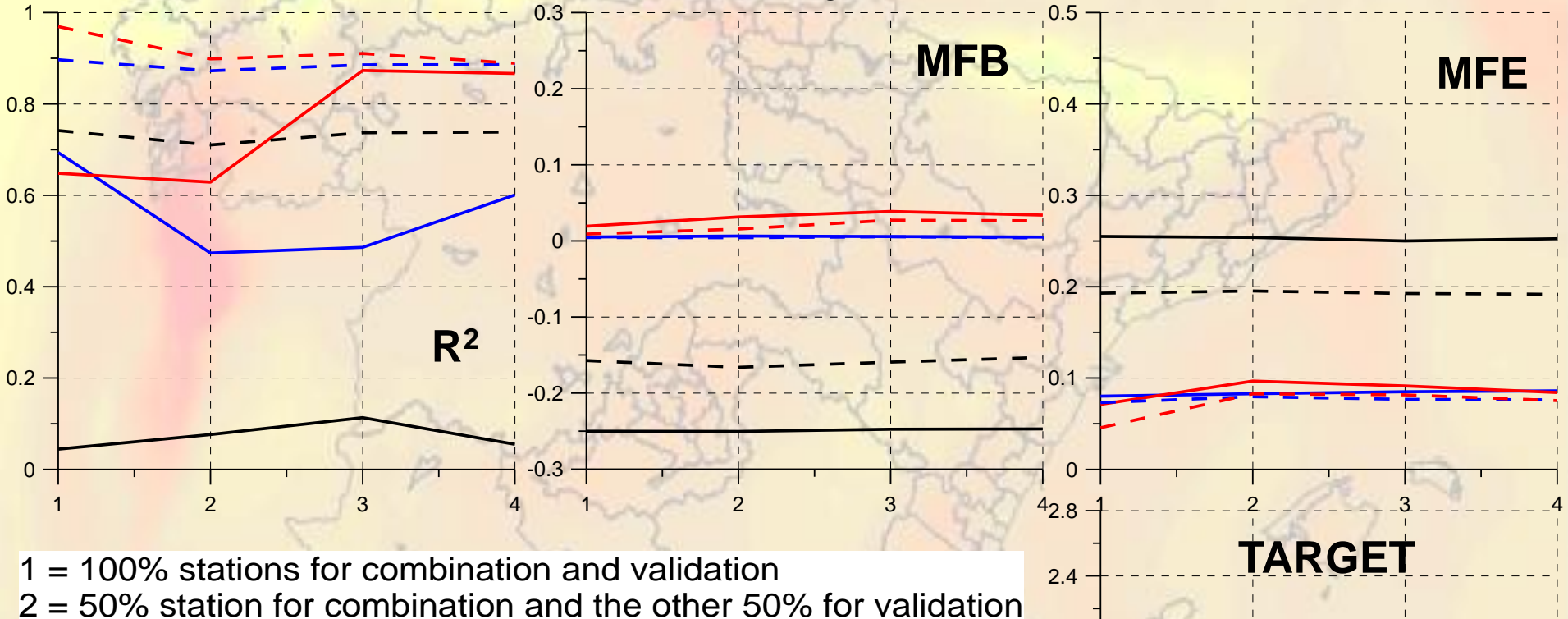
MFE and MFB similar for leave-one-out and for different selected data for validation for urban stations

- Combined model-observation/leave-one-out validation method/urban-suburban
- Combined model-observation/some stations for combination/urban-suburban
- - Model Rural stations
- - Combined model-observation/leave-one-out validation method/rural
- - Combined model-observation/some stations for combination/rural



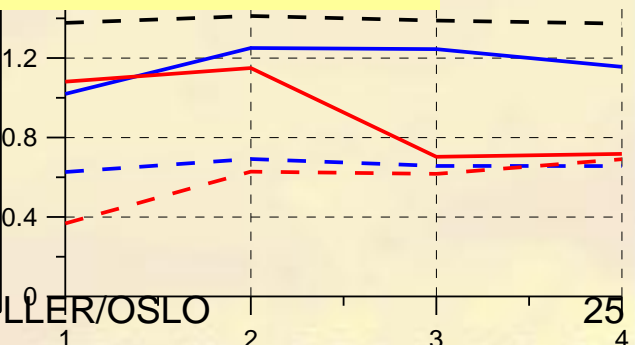
STATISTICS vs DATA/METHOD FOR VALITATION

Maximum hourly O₃ concentrations



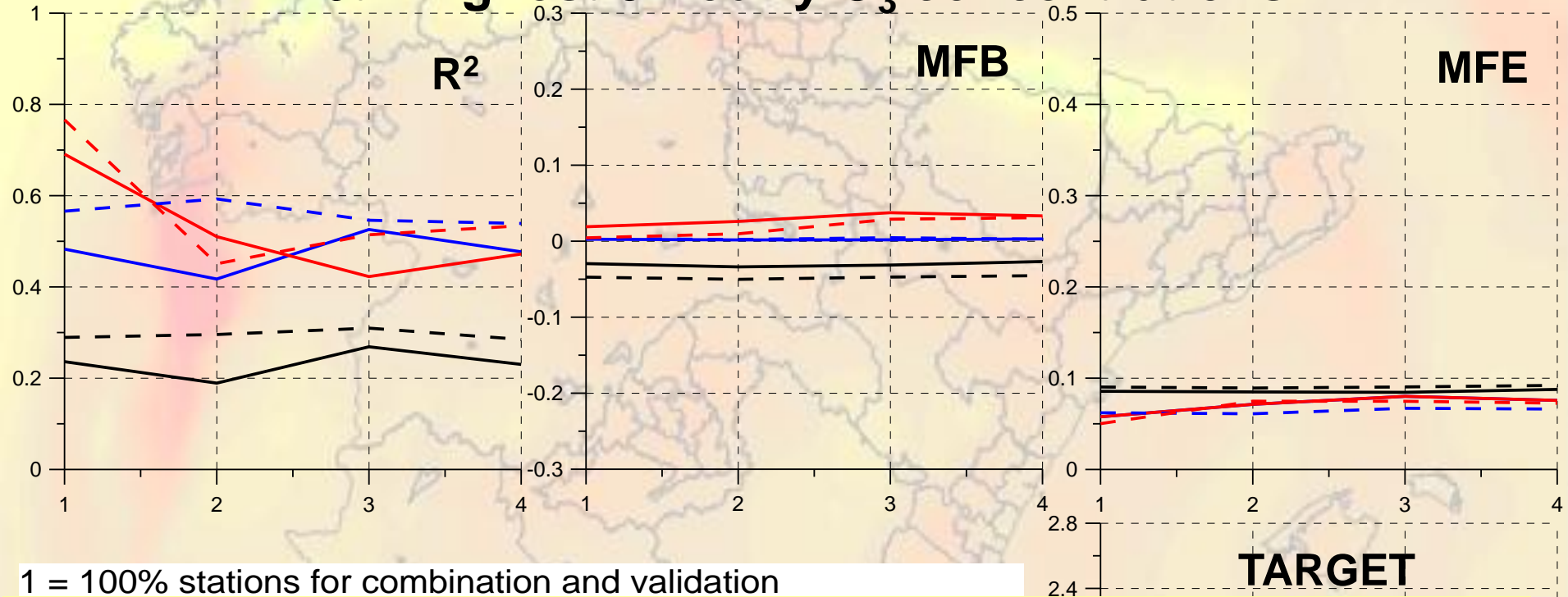
MFE and MFB similar for leave-one-out and for different selected data for Validation.

- Model Urban-Suburban stations
- Combined model-observation/leave-one-out validation method/urban-suburban
- Combined model-observation/some stations for combination/urban-suburban
- Model Rural stations
- Combined model-observation/leave-one-out validation method/rural
- Combined model-observation/some stations for combination/rural



STATISTICS vs DATA/METHOD FOR VALITATION

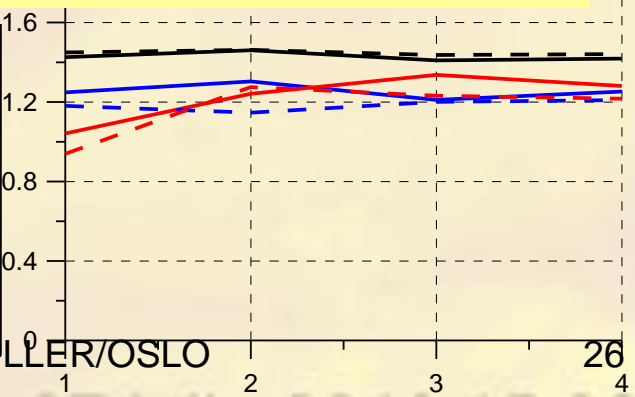
26th highest 8-hourly O₃ concentrations



TARGET

MFE, MFB and TARGET similar for leave-one-out and for different selected data for validation.

- Model Urban-Suburban stations
- Combined model-observation/leave-one-out validation method/urban-suburban
- Combined model-observation/some stations for combination/urban-suburban
- Model Rural stations
- Combined model-observation/leave-one-out validation method/rural
- Combined model-observation/some stations for combination/rural



Some conclusions and ideas

- In some cases statistical index are similar (especially for MFE), but in others no. Then values of statistical indexes are different depending on:
 - Method for validation
 - Data set
- It is not clear what method has to be used.
- It is not straightforward to get rules about how many stations has to be used for validation respect to the used ones for combination.
- Needs of more studies and tests.
 - Other data sets and cases.
 - Select more subsets of data for validation (ensemble).
 -

