# Prelimenary analysis of the redundancy in HTAP1 data -future directions for HTAP2

Stefano Galmarini, Efisio Solazzo and
Ioannis Kioutsioukis

Presented by Frank Dentener, San Francisco, December 2014

JRC/IES

AQMEII: 20 Models

What is the best and most beneficial way to build an ensemble of members? And how should the optimum size of the ensemble be determined in order to capture data variability as well as keeping the error low? These questions are addressed here by looking at optimal ensemble size and quality of the members.

**Background**:

➢ Multi model ensembles => ensemble of convenience
➢ Models are selected because they are available, but …
➢ Do they contribute additional information content to the ensemble result?
➢ Models only give full contribution to the ensemble result if they are 'truly' independent
➢ Our models could produce different results but not independent results!
➢ A screening of the model and ensemble performances is an important step often neglected- including HTAP Phase 1.

**Research questions**:

What is the level of independent information content produced by each model of the HTAP1 ensemble?
What is the level of redundancy of the results?
What is the effect of the redundancy on the estimate of ensemble error?

**Use of three techniques to analyse and estimate redundancy and its effects:**

1- Talagran Diagram (Rank Histogram)

2-  the calculation of the eigenvalues of the matrix of correlations among all models errors
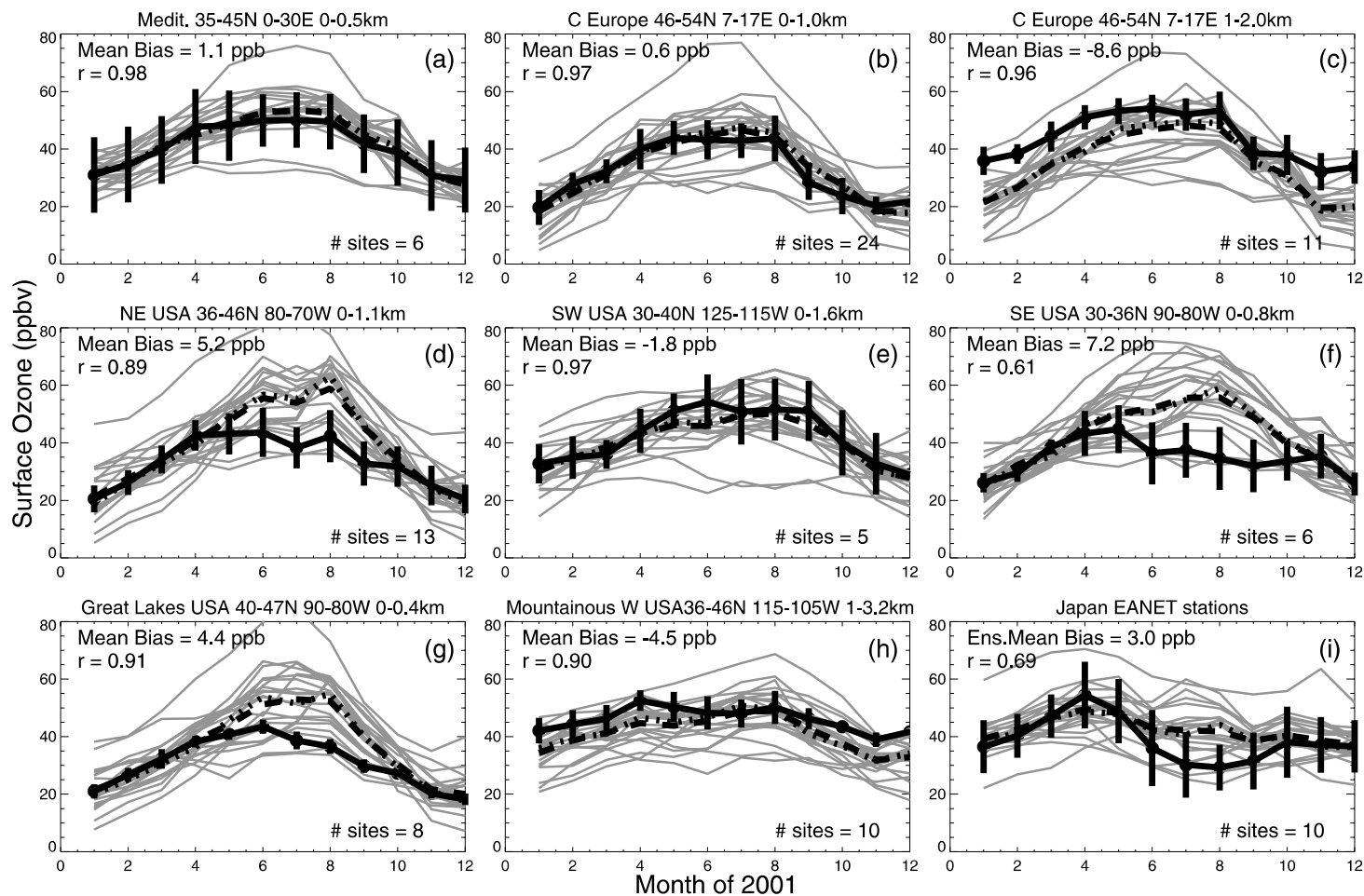
3- analysis of the minimum error

**Figure 2.** Monthly mean surface O$_3$ concentrations (ppb) for the year 2001. Observed values (black circles) represent the average of all sites falling within the given latitude, longitude, and altitude

Fiore et al. 2009; Figure 2

# 1- Ranked or Talagran histogram

- Each region 12 monthly values and 25 models realizations
- Min-Max=>20 bins=>the number of measurements in these bins
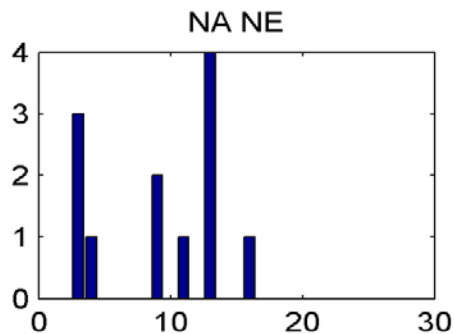- empty bins -> models predict a min-max not seen in the measurements
- Optimal diagram: all bins are equally filled.
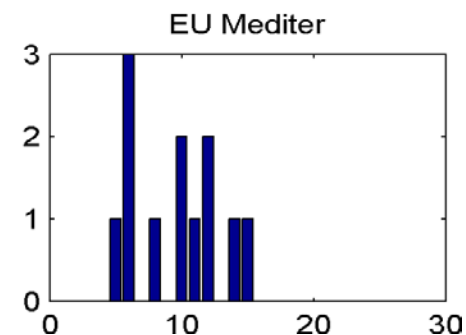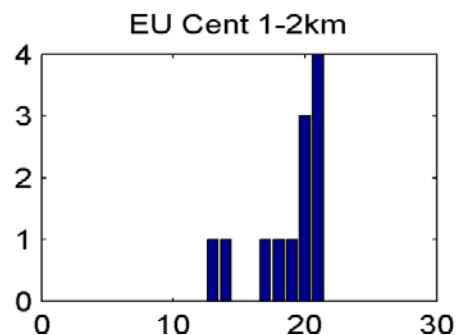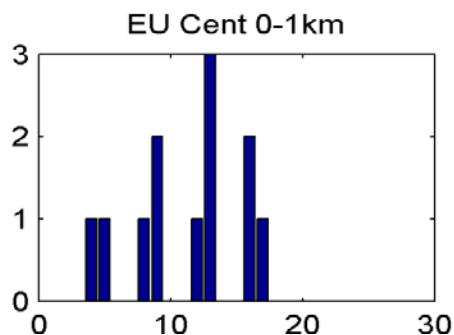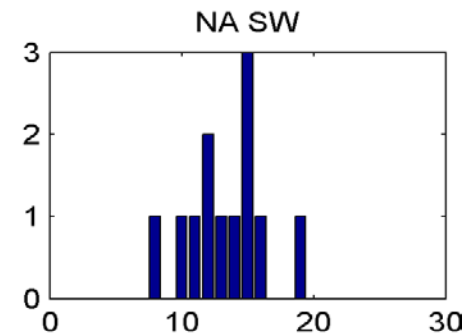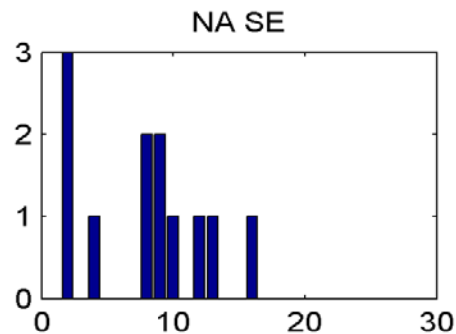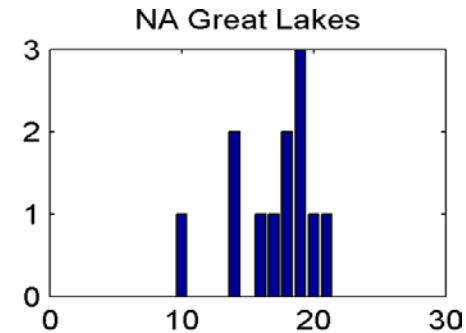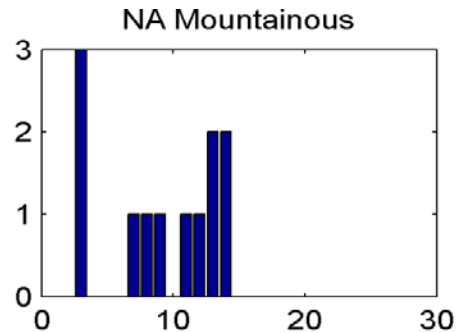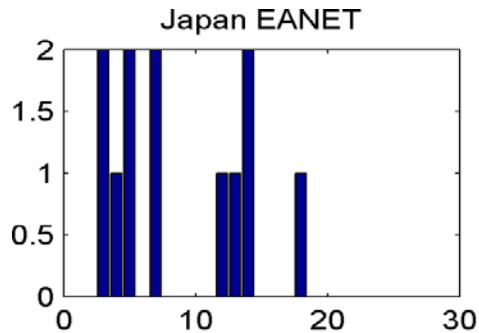
## 2- HTAP ozone monthly mean concentration at ground for the year 2001; Eigenvalue analysis of model error correlations

**21 models available**:

CAMCHEM-3311m13

ECHAM5-HAMMOZ-v21

EMEP-rv26

FRSGCUCI-v01

GEMAQ-EC

GEMAQ-v1p0

GEOSChem-v07

GEOSChem-v45

GISS-PUCCINI-modelA

GISS-PUCCINI-modelE

GMI-v02f

INCA-vSSz

LLNL-IMPACT-T5a

MOZARTGFDL-v2

MOZECH-v16

OsloCTM2

STOC-HadAM3-v01

STOCHEM-v02

TM5-JRC-cy2-ipcc-v1

ULAQ-v02

UM-CAM-v01

Measurements are made available in **nine** sub-regions. For each of those Fiore et al (2009) derived the ensemble mean.

Our redundancy test **based on eigenvalues analysis** shows that the number of effective models $M_{eff}$ is **between 2 and 4**.
Based on matrix of correlations of model errors corr($di,dj$)):
   with d: distance of model-observation

EU Mediterranean region (nrec=6) $M_{eff}$ = **4.0**
EU central region 0-1 km (nrec=24) $M_{eff}$ = **3.06**
EU central region 1-2 km (nrec=11) $M_{eff}$ = **3.5**
NE-USA (nrec=13) $M_{eff}$ = **1.86**
SW USA (nrec=5) $M_{eff}$ = **1.82**
SE USA (nrec=6) $M_{eff}$ = **1.89**
Great Lakes USA (nrec=8) $M_{eff}$ = **2.0**
Mountainous USA (nrec=10) $M_{eff}$ = **1.8**
Japan EANET (nrec=10) $M_{eff}$ = **2.6**

where nrec is the number of measurements in each sub-region.

# 3- Analysis of minimum error

RMSE is the root mean square error of modelled vs. measured ozone concentration [ppb]
Permutations of combinations of N (e.g. 21, or 7 or 3 models)=>errors
In this specific example: lowest error curve minimizes around 4-5 but not everywhere
The 'winning' combination is not a-priori known- need some redundancy
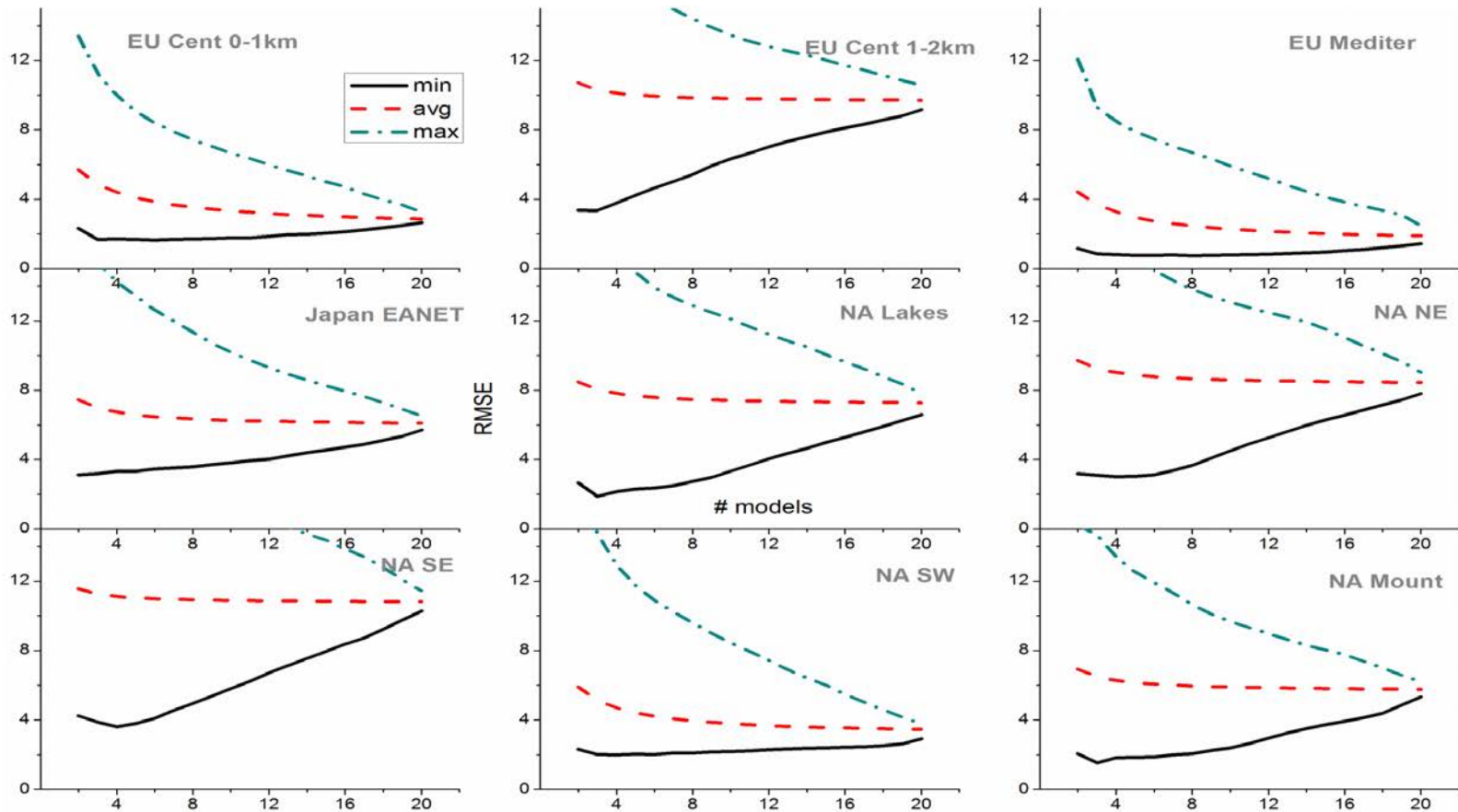
| Domain | score minRMSE |
|--------|---------------|
| EU central 0-1 km | RMSE=1.69 (2.65)  **- 36%**<br>PCC=0.98 (0.96)<br>$\sigma$=0.99 (1.10) |
| EU central 1-2 km | RMSE=3.35 (9.2)  **-63%**<br>PCC=0.98 (0.95)<br>$\sigma$=1.03 (1.25) |
| EU medit | RMSE=0.76 (1.44)  **- 47%**<br>PCC=0.99 (0.98)<br>$\sigma$=1.0 (1.13) |
| NA _SW | RMSE=2.0 (2.9)  **-31%**<br>PCC =0.95 (0.96)<br>$\sigma$=0.87 (0.86) |
| NA_SE | RMSE=3.61 (10.27)  **-65%**<br>PCC=0.77 (0.62)<br>$\sigma$=0.83 (1.81) |
| NA_NE | RMSE=3.01 (7.8)  **-61%**<br>PCC=0.93 (0.90)<br>$\sigma$=0.90 (1.56) |
| NA_Mountain | RMSE=1.53 (5.33)  **-71%**<br>PCC=0.93 (0.90)<br>$\sigma$=1.04 (1.44) |
| NA_LAkes | RMSE=1.89 (6.58)  **-71%**<br>PCC=0.97 (0.91)<br>$\sigma$=1.03  (1.45) |
| Japan | RMSE=3.11 (5.70)  **- 45%**<br>PCC=0.96 (0.79)<br>$\sigma$=0.66 (0.51) |

**Effect on error statistics when non redundant ensemble is selected**

Score of the min-RMSE combinations.
In parenthesis ensemble mean corresponding to Fiore et al.
Ensemble mean which was used to project future scenarios based  on emission perturbations

 - RMSE is the root mean square error
(the % reduction with respect to the full ensemble mean);

 - PCC is the linear correlation coefficient

- Sigma is the ratio between modeled and observed standard deviation (values closer to unity are desired)

In all ceases the reduced ensemble mean improves on accuracy and precision.

# Summary

- There is a necessity to diagnose multi model ensemble error prior to using it
- There are techniques to do this rather efficiently
- Each parameter may behave differently
- Probably different for long-range transport
     versus local effect (not tried yet).
- Proof of concept
- Future:
  - Ensemble reduction of HTAP2 dataset
  - Use of mixed global-regional ensemble datasets
   What is the redundancy in mixed global/regional results?