AEROCOM and GEMS-AER

lessons learned

Stefan Kinne

MPI-Meteorology
Hamburg, Germany

outline

- benefits from assimilation ?
- comparing AOD and Angstrom of
  - forecast (ECMWF)
  - GEMS assimilation
  - MACC assimilation
- strength and weaknesses
  - by ‘subjective’ comparison
  - by ‘objective’ scores
- exploring regional errors and differences
- finally something related to the title

the task

- to be evaluated monthly maps
  - ECMWF aerosol forecasts for 2003 and 2004
    - SO2/DMS, sulfate, o./bl. carbon (hydro. / nonhydro. each), seasalt (3*size), dust (3*size) : 12 tracers
  - GEMS assimilations for 2003 and 2004
    - MODIS (coll. 5) aerosol optical depth (no S removal)
  - MACC assimilations for 2003 and 2004 (GFED3)
    - MODIS (coll. 5) aerosol optical depth used
- **AOD (aerosol optical depth)** *(info on 'amount')*
- **Angstrom parameter** *(info on 'aerosol size')*

on terminology

- aerosol optical depth (AOD)
  - extinction along a (vertical) direction due to scattering and absorption by aerosol
    - here for the entire atmosphere
    - here for the mid-visible (0.55μm wavelength)

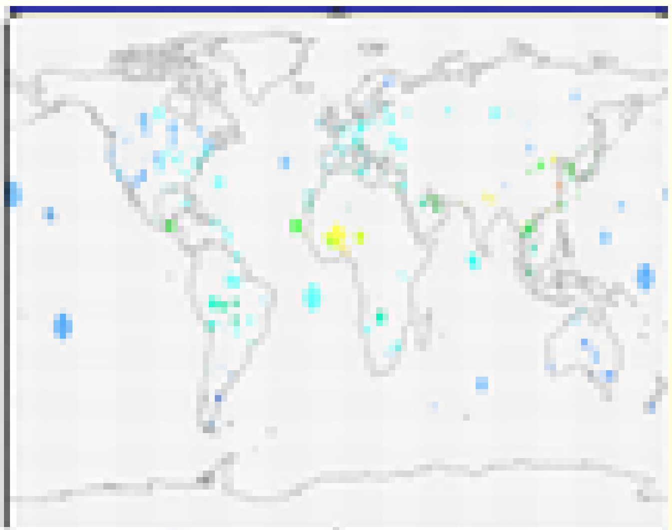
- Angstrom parameter (Ang)
  - spectral dependence of AOD *in the visible spectrum*
    - small dependence (Ang ~ 0) ⇨ aerosol > 1μm size
    - strong decrease (Ang > 1.2) ⇨ aerosol < 0.5μm size

the reference

- **sunphotometer AOD data**
  - **AERONET(~200) + GAW (15) + SKYNET (8)**
    - **monthly statistics combined on a 1X1 (lat/lon) grid**

annual AOD maps

2003 data

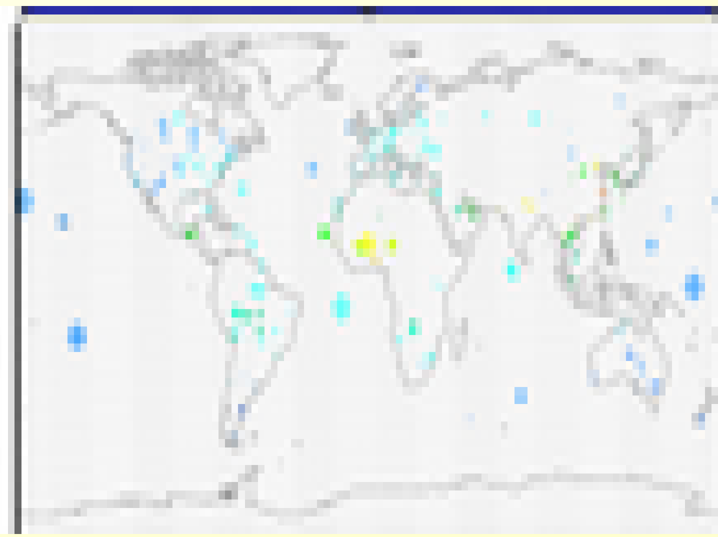


sun

0.188

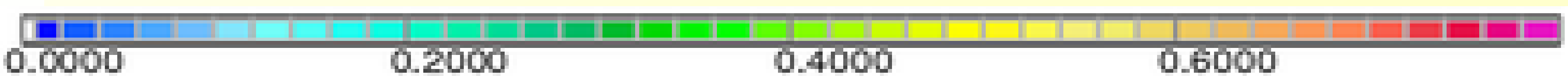
globally sparse

2004 data



sun

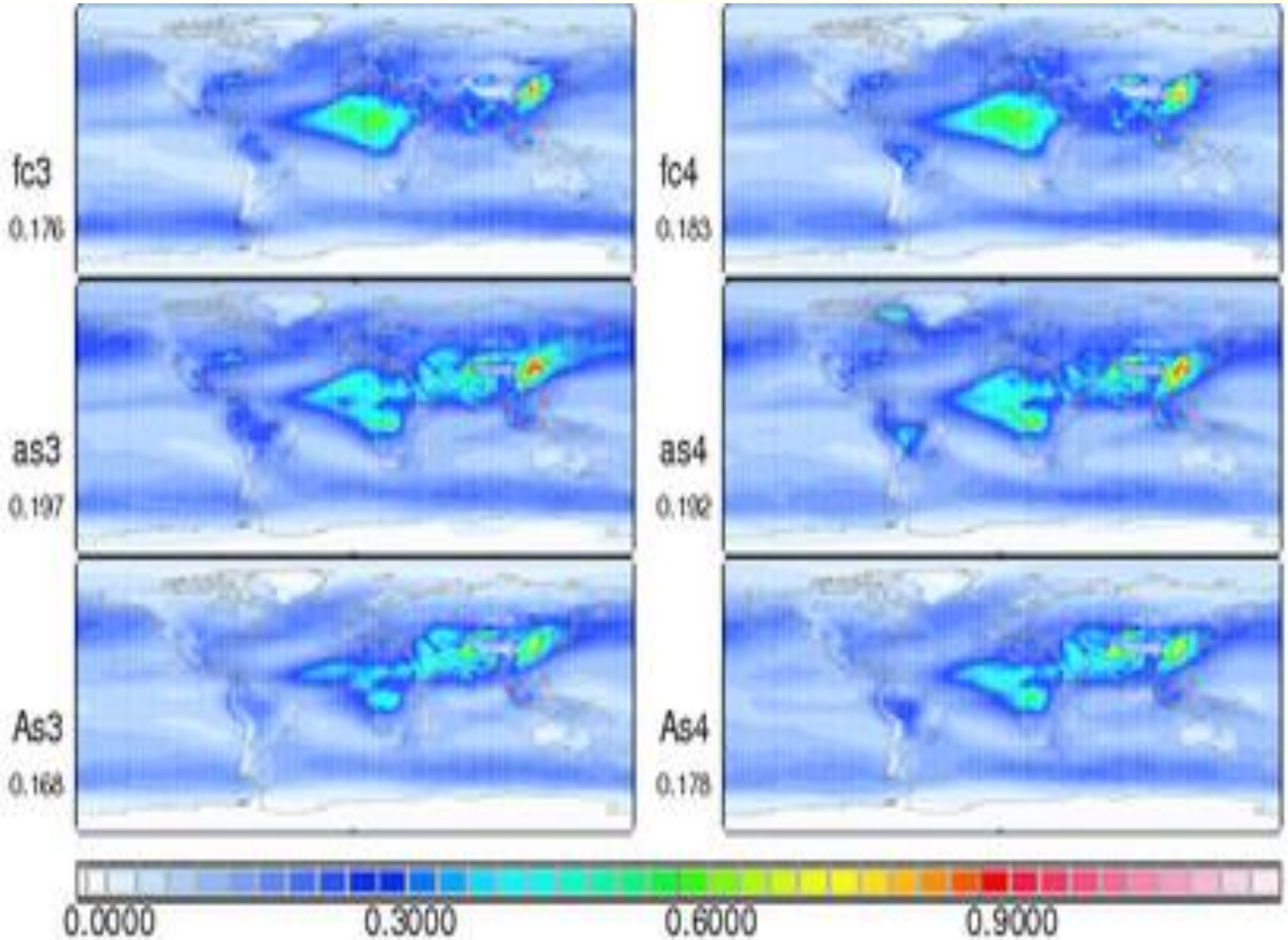
0.187



AOD simulations

year 2003 **year 2004**

annual maps

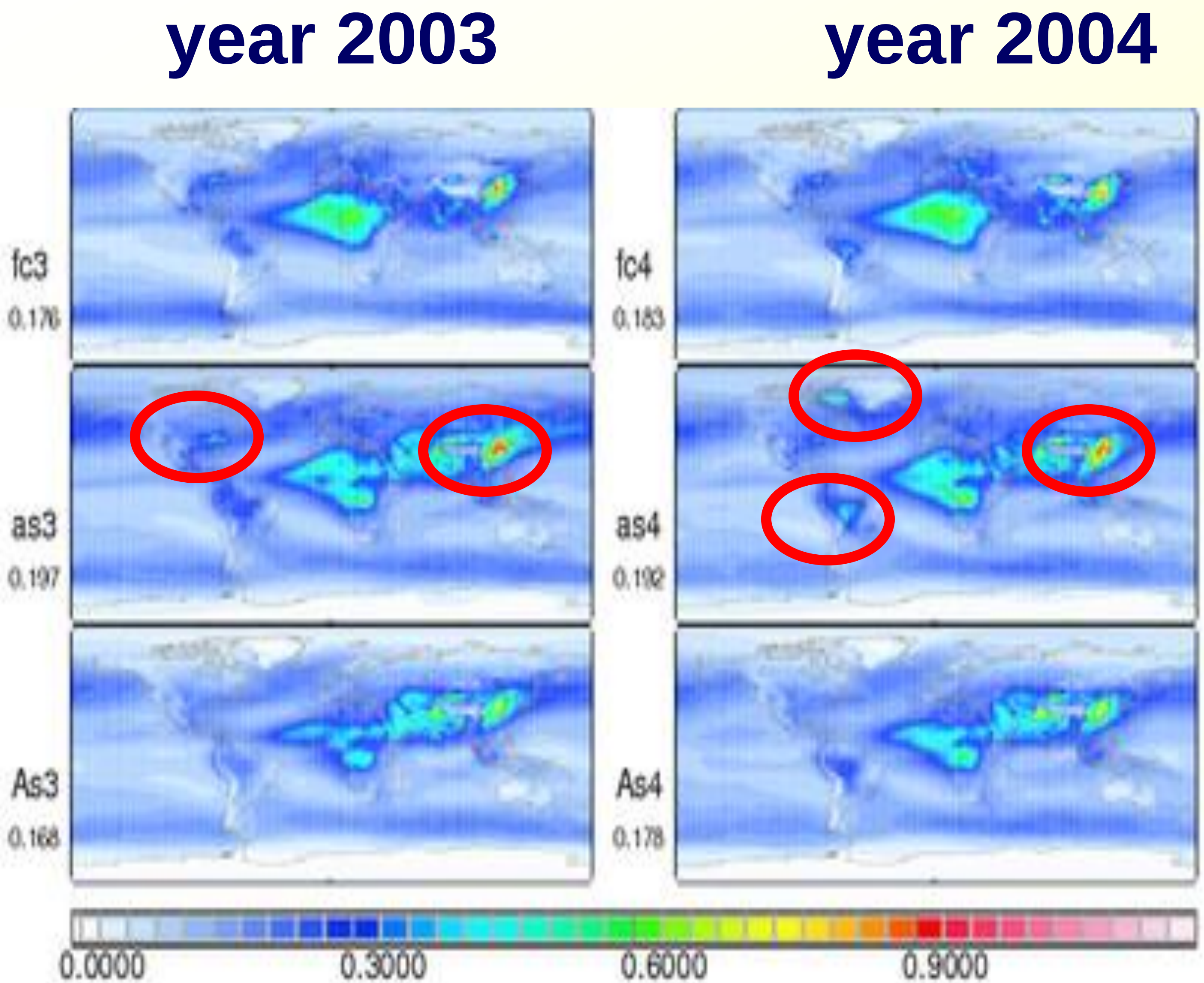


fc.. forecast
as.. GEMS
As.. MACC

AOD simulations

annual maps

GEMS deficiencies

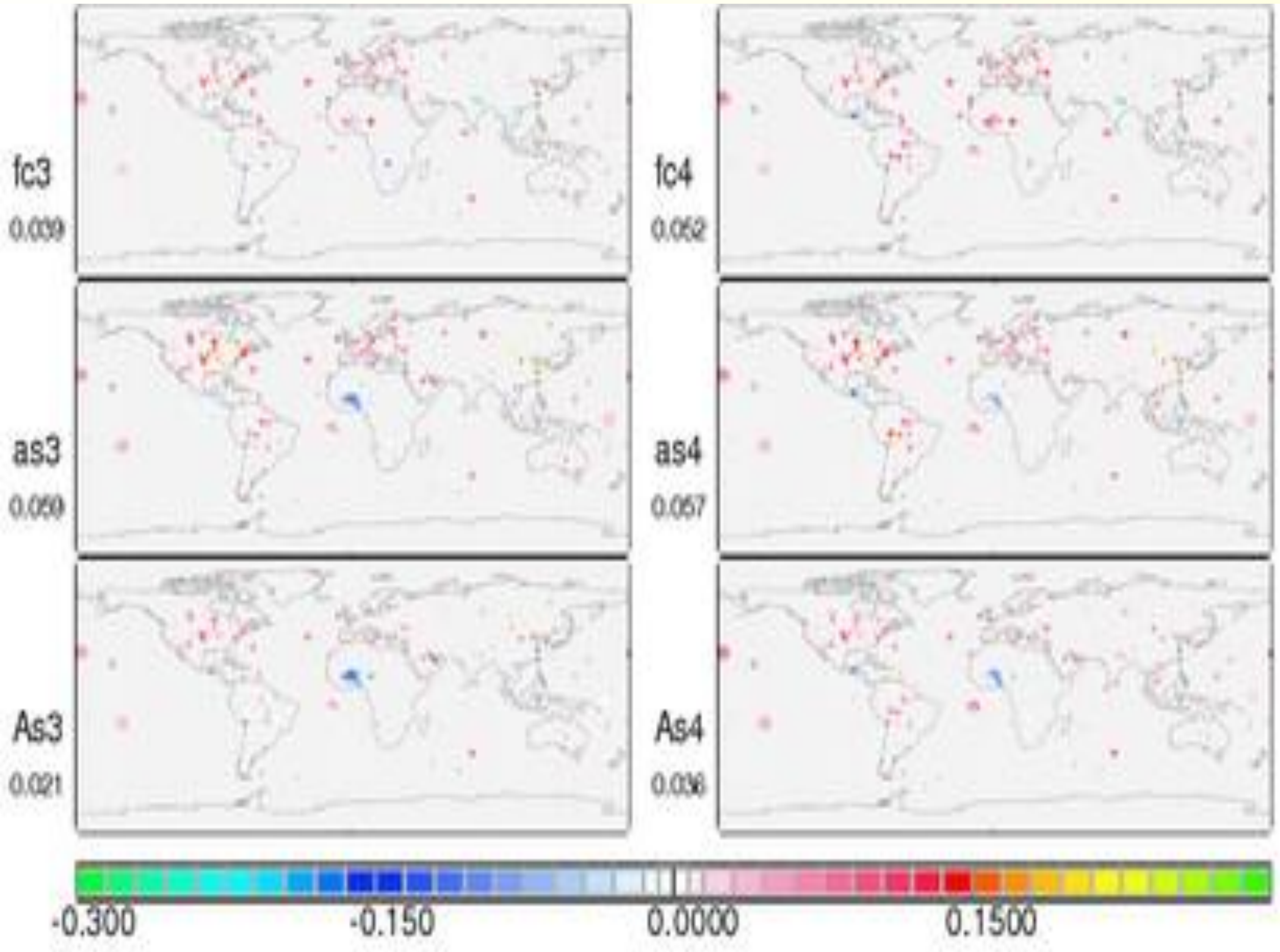


fc.. forecast
as.. GEMS
As.. MACC

AOD diff. to AERONET

year 2003 **year 2004**

annual maps

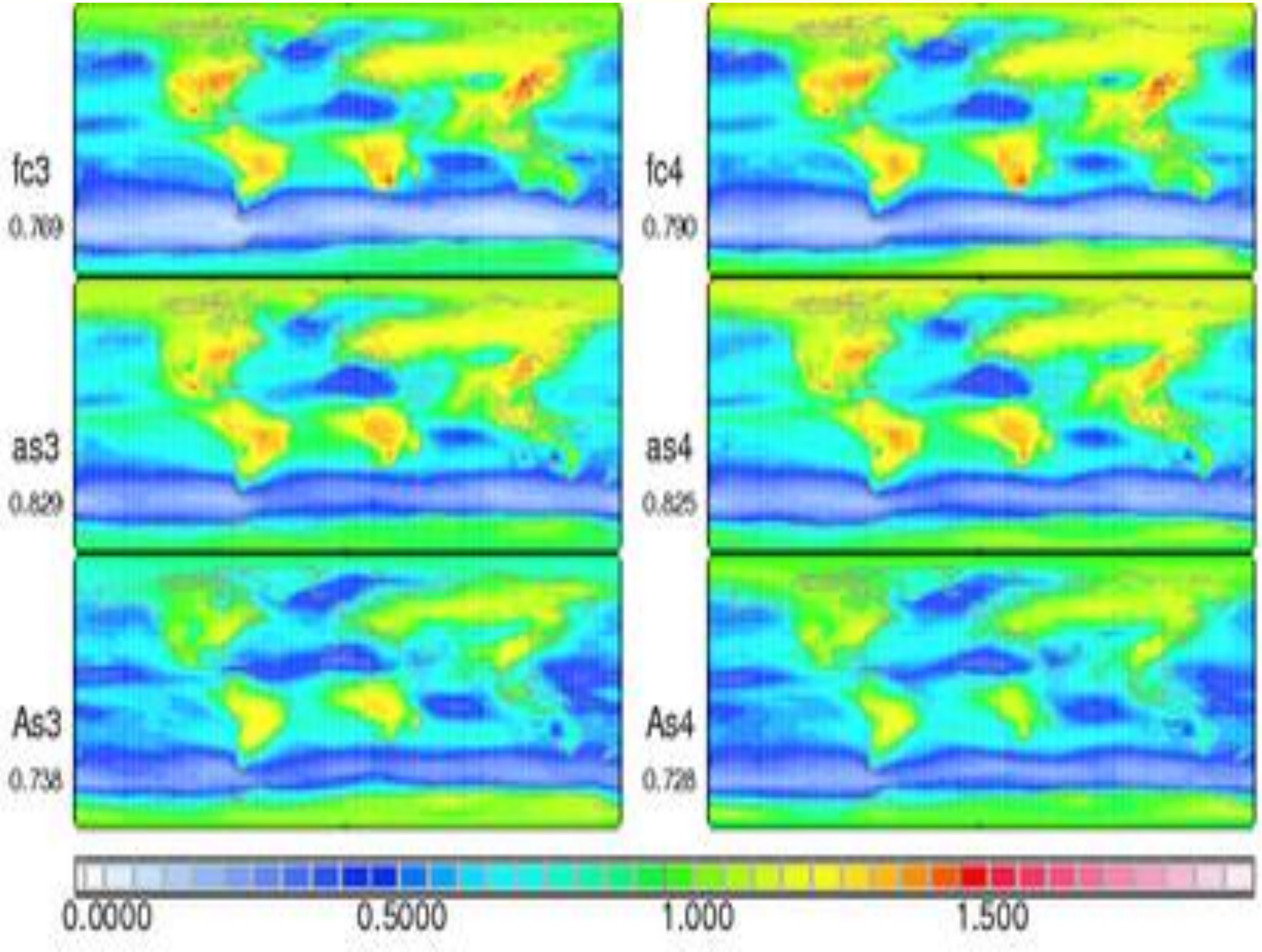


fc.. forecast
as.. GEMS
As.. MACC

Angstrom simulations

year 2003 **year 2004**

annual maps

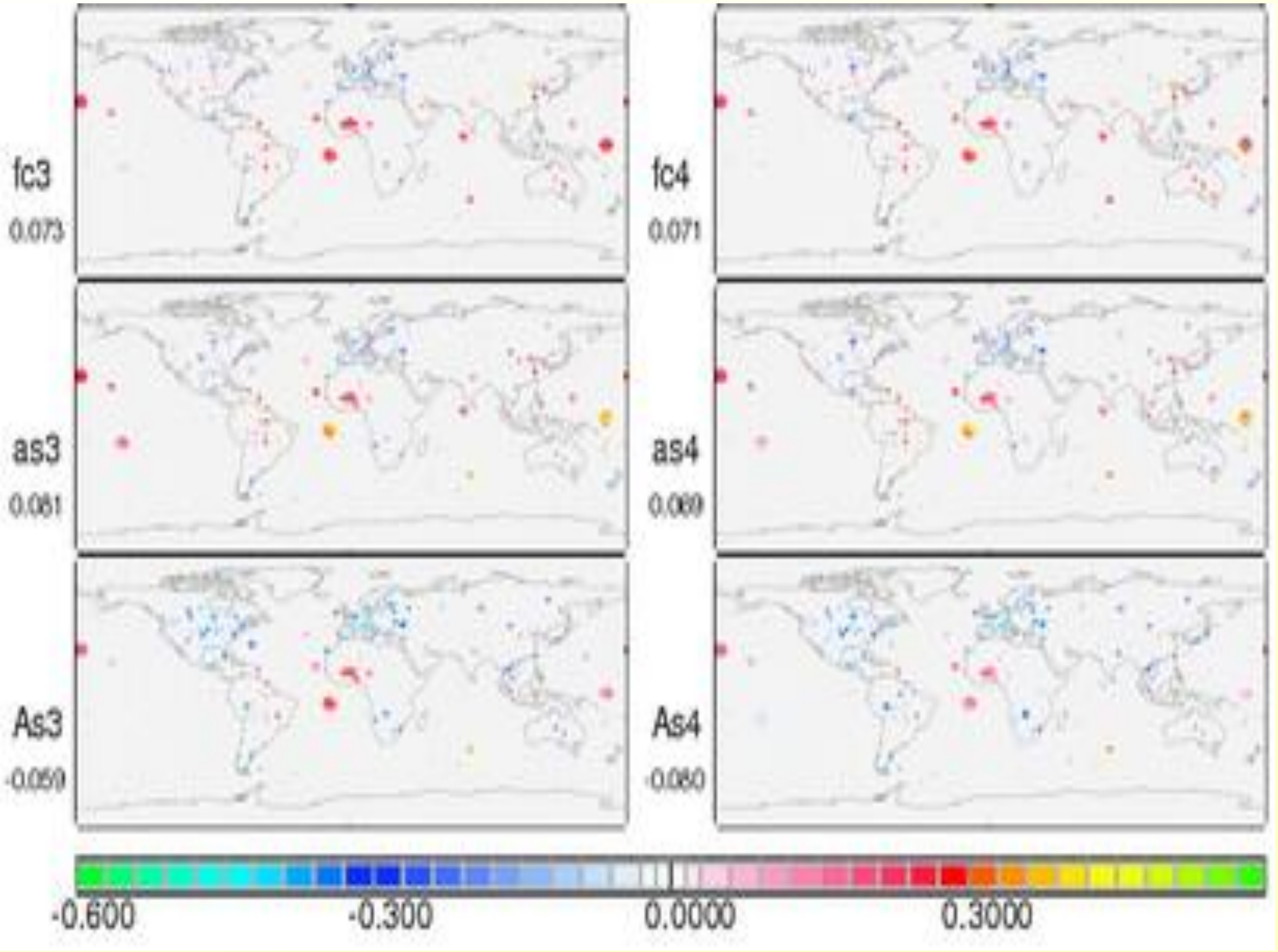


fc.. forecast
as.. GEMS
As.. MACC

Angstrom diff. to AERONET

year 2003 **year 2004**

annual maps



fc.. forecast
as.. GEMS
As.. MACC

first ‘subjective’ impressions

- MODIS data assimilation
  - increases high AOD bias over land
  - reduces Saharan dust
  - reduces Angstrom over NH continents
- MACC vs GEMS
  - reduced AOD for biomass over S.America
  - (further) reduced Angstrom parameters

… and now more ‘objective’

- quantify data performance by one number
- develop a score such that contributing errors to be traceable back to
  - bias
  - spatial correlation
  - temporal correlation
  - spatial sub-scale (e.g. region)
  - temporal sub-scale (e.g. month, day)
- make this score outlier resistant

- one possible scoring method ….

one number !

- 0.504

info on overall bias

sign of the bias

- 0.504

| 1 | is perfect …. 0 is poor

sign of
the bias

- 0.504

the closer to
absolute 1.0
… the better

product of sub-scores

sign of the bias

temporal correlation sub-score

bias sub-score

spatial correlation sub-score

- 0.504 = 0.9 * -0.7 * 0.8

the closer to absolute 1.0 … the better

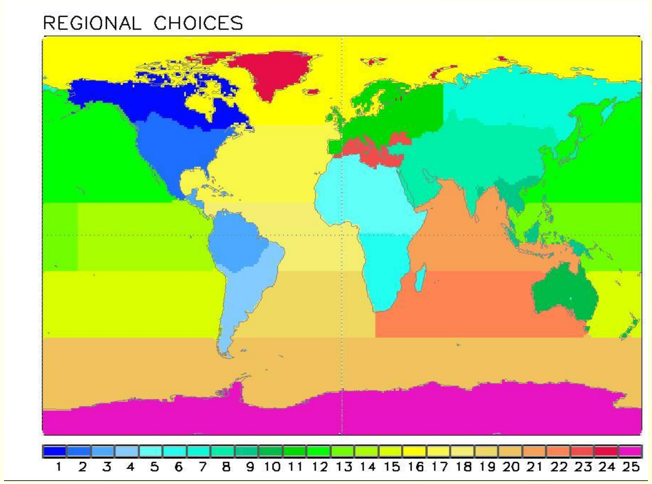
spatial stratification

sign of the bias	overall score	time score	bias score	spatial score

- 0.504 = 0.9 *- 0.7 * 0.8

regional surface area weights

spatial sub-scale scores

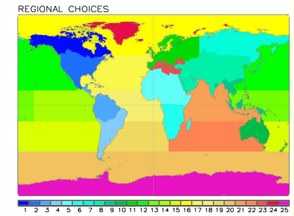


TRANSCOM regions

spatial stratification

sign of the bias | overall score | time score | bias score | spatial score

- 0.504 = 0.9 * - 0.7 * 0.8



spatial sub-scale scores

instantaneous median data

averaging in time

temporal sub-scale scores *(e.g. month or days)*

sub-score definition

- each sub-score S is defined
  - by an error e and
  - by an error weight w

time score | bias score | spatial score

0.9 * -0.7 * 0.8

$$S = 1 - w * e$$

spatial sub-scale scores

instantaneous median data

temporal sub-scale scores *(e.g. month or days)*

bias error

- **S = 1 – w * e**
- **all errors e are**
  rank-based

time score	bias score	spatial score
correlation	bias	correlation

- **e, bias = (rank-sum1 – rank-sum2) /(sum12)**
  (between e = -1 and e= +1)

how does the rank bias error work ?

- **set 1: 1 7 8** **value: 9 8 7 4 3 1** **rank-sum 1: 11**
- **set 2: 3 4 9** **rank: 1 2 3 4 5 6** **rank-sum 2: 10**

bias = (1-2)/(1+2) = (11-10)/21 ~zero ⇨ no clear bias

total score S

- $S = S_T * S_B * S_S$

  $= (1 - w * e_T)$

  $* (1 - w * e_B)$

  $* (1 - w * e_S)$

time score	bias score	spatial score
correlation	bias	correlation

- all errors e are "rank-based"
- weight $w = (75\%pdf - 25\%pdf - \Delta e) / 50\%pdf$

  … but not smaller than 0 & not larger than 1.0
- simply put …

  no variability (w = 0): errors do not matter/count

bias error e_B

- $S_B = 1 - w * e_B$

bias score

correlation | bias | correlation

- e_B = (rank-sum1 – rank-sum2) /(sum12)

(${e_B}$ between -1 and +1)
(sign of e_B indicates bias)

how does the rank bias error work ?

set 1: 1 7 8 value: 9 8 7 4 3 1 rank-sum 1: 11
set 2: 3 4 9 rank: 1 2 3 4 5 6 rank-sum 2: 10

bias = (rs1-rs2)/(rs1+rs2) = (11-10)/21 ~zero ⇨ no bias

spat.error e_S, temp error e_T

- $S = 1 - w * e_S$
- $S = 1 - w * e_T$

time score

spatial score

correlation | bias | correlation

- e_S = (1- rank_correlation coeff.) /2

  (correlated: e = 0, anti-correlated: e = 1)

  …using regional data at one time

- e_T = (1- rank_correlation coeff.) /2

  (correlated: e = 0, anti-correlated: e = 1)

  …using time series of regional median data

scoring approach

- one single score …
- … without sacrificing spatial and temporal detail !
- stratification into error contribution from
  - bias
  - spatial correlation
  - temporal correlation
- robustness against outliers
  - still … just one of many possible approaches
- now to some applications …

- now with real data

questions

- how did the different simulations score?
  - forecast
  - GEMS assimilation
  - MACC assimilation
- did assimilations improve the forecast?
  - overall ?
  - seasonality ?
  - spatial correlation ?
  - bias ?
  - in what regions ?
  - In what months ?

annual global scores

- *year 2003 - aod*

	TOTAL	seas	bias	corr
macc	.56	.90	.81	.77
gems	.56	.91	.77	.79
forec	.49	.81	.79	.77

- *year 2003 – Angstrom*

	TOTAL	seas	bias	corr
gems	.63	.87	.86	.85
forec	.63	.88	.85	.84
macc	-.59	.81	-.86	.85

- *year 2004 - aod*

	TOTAL	seas	bias	corr
macc	.55	.90	.80	.77
gems	.55	.90	.77	.79
forec	.50	.83	.79	.77

- *year 2004 - Angstrom*

	TOTAL	seas	bias	corr
mac	.67	.89	.87	.87
forec	.66	.89	.86	.86
gems	.65	.87	.87	.86

better overall AOD score

- *year 2003 - aod*

	TOTAL	seas	bias	corr
macc	.56	.90	.81	.77
gems	.56	.91	.77	.79
forec	.49	.81	.79	.77

- *year 2003 – Angstrom*

	TOTAL	seas	bias	corr
gems	.63	.87	.86	.85
forec	.63	.88	.85	.84
macc	-.59	.81	-.86	.85

- *year 2004 - aod*

	TOTAL	seas	bias	corr
macc	.55	.90	.80	.77
gems	.55	.90	.77	.79
forec	.50	.83	.79	.77

- *year 2004 - Angstrom*

	TOTAL	seas	bias	corr
mac	.67	.89	.87	.87
forec	.66	.89	.86	.86
gems	.65	.87	.87	.86

vs sun-photometry

better AOD seasonality

- *year 2003 - aod*

	TOTAL	seas	bias	corr
macc	.56	.90	.81	.77
gems	.56	.91	.77	.79
forec	.49	.81	.79	.77

- *year 2003 – Angstrom*

	TOTAL	seas	bias	corr
gems	.63	.87	.86	.85
forec	.63	.88	.85	.84
macc	-.59	.81	-.86	.85

- *year 2004 - aod*

	TOTAL	seas	bias	corr
macc	.55	.90	.80	.77
gems	.55	.90	.77	.79
forec	.50	.83	.79	.77

- *year 2004 - Angstrom*

	TOTAL	seas	bias	corr
mac	.67	.89	.87	.87
forec	.66	.89	.86	.86
gems	.65	.87	.87	.86

vs sun-photometry

pos. AOD bias – worse in GEMS

- *year 2003 - aod*

	TOTAL	seas	bias	corr
macc	.56	.90	.81	.77
gems	.56	.91	.77	.79
forec	.49	.81	.79	.77

- *year 2003 – Angstrom*

	TOTAL	seas	bias	corr
gems	.63	.87	.86	.85
forec	.63	.88	.85	.84
macc	-.59	.81	-.86	.85

- *year 2004 - aod*

	TOTAL	seas	bias	corr
macc	.55	.90	.80	.77
gems	.55	.90	.77	.79
forec	.50	.83	.79	.77

- *year 2004 - Angstrom*

	TOTAL	seas	bias	corr
mac	.67	.89	.87	.87
forec	.66	.89	.86	.86
gems	.65	.87	.87	.86

overall Ang - largely unchanged

- *year 2003 - aod*
  - TOTAL seas bias corr
- macc .56 .90 .81 .77
- gems .56 .91 .77 .79
- forec .49 .81 .79 .77

- *year 2003 – Angstrom*
  - TOTAL seas bias corr
- gems .63 .87 .86 .85
- forec .63 .88 .85 .84
- macc .59 .81 -.86 .85

- *year 2004 - aod*
  - TOTAL seas bias corr
- macc .55 .90 .80 .77
- gems .55 .90 .77 .79
- forec .50 .83 .79 .77

- *year 2004 - Angstrom*
  - TOTAL seas bias corr
- mac .67 .89 .87 .87
- forec .66 .89 .86 .86
- gems .65 .87 .87 .86

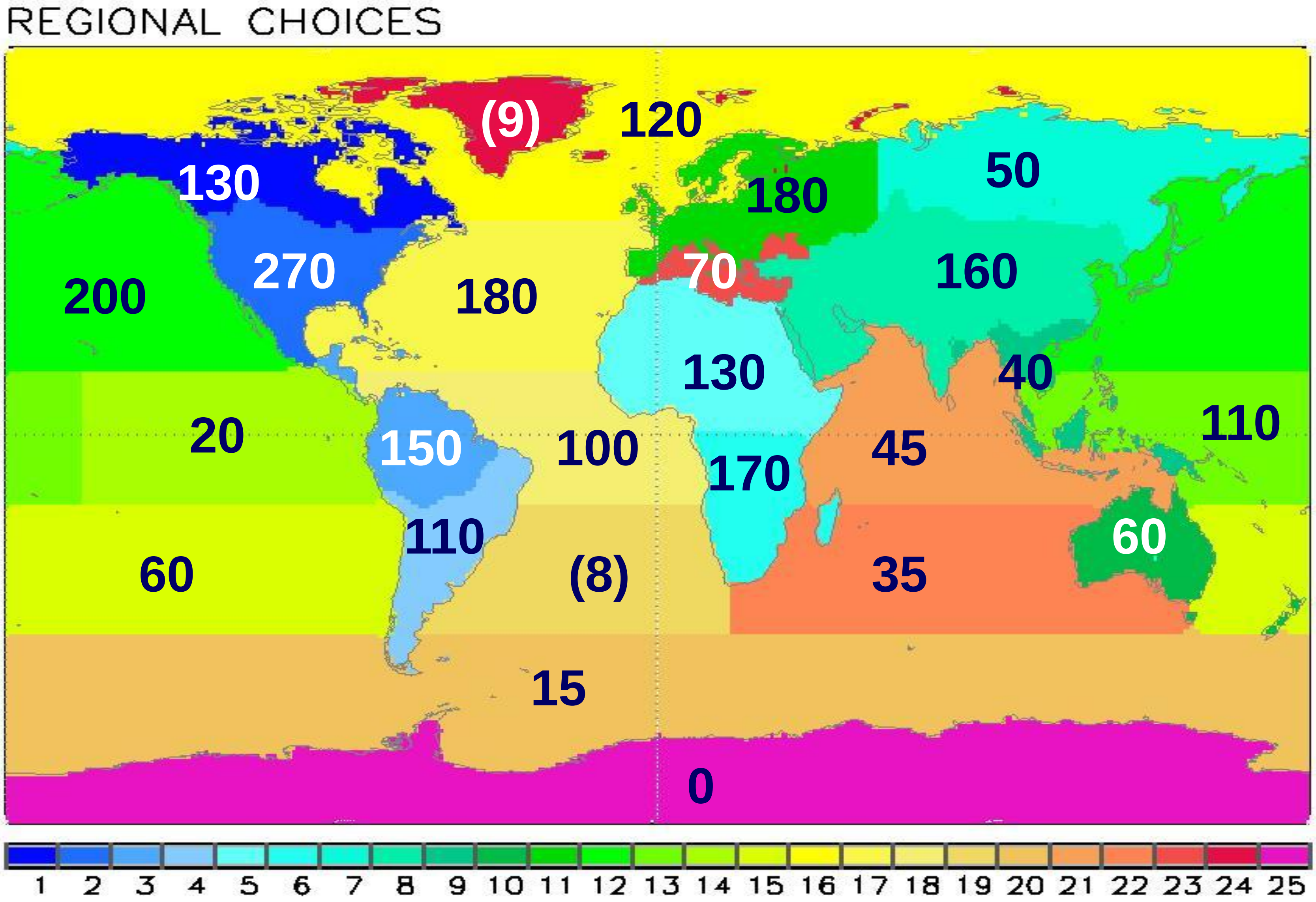
summary

- assimilations improved AOD score
- better AOD seasonality is the main reason
- positive AOD
  - stronger in GEMS than for the forecast
  - weaker than forecast in MACC
- ocean AOD more improved than land AOD
- Angstrom score largely unchanged
  - tendency to low bias only MACC

- …still the score is far from perfect

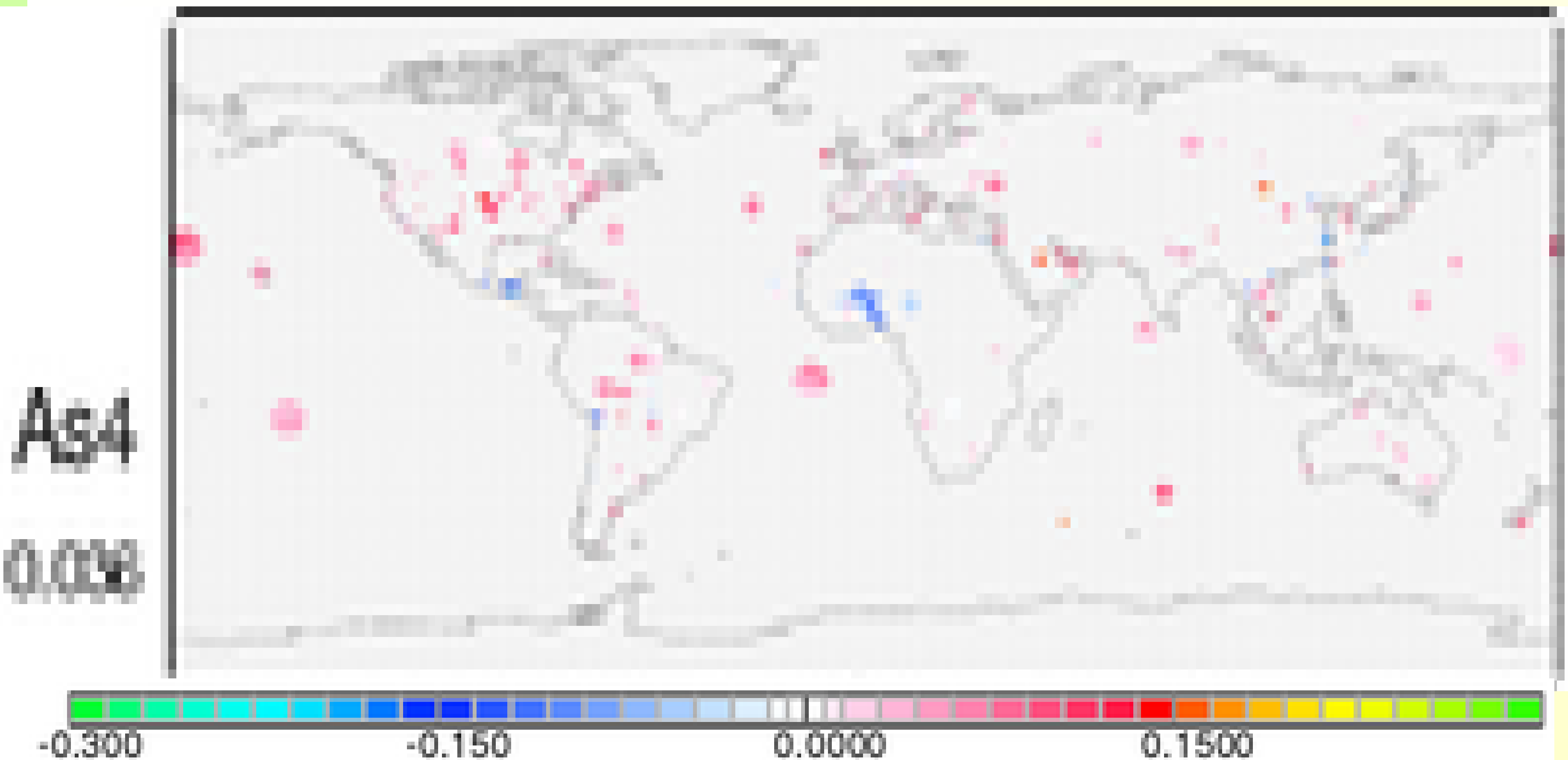
- **now look at regions**

regional stratifaction / data-pairs



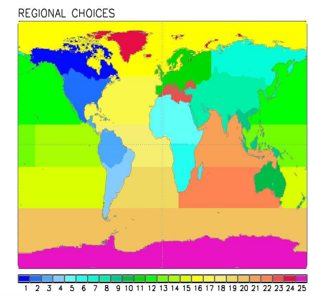
2004 AOD – how to quantify performance?

AOD difference to AERONET



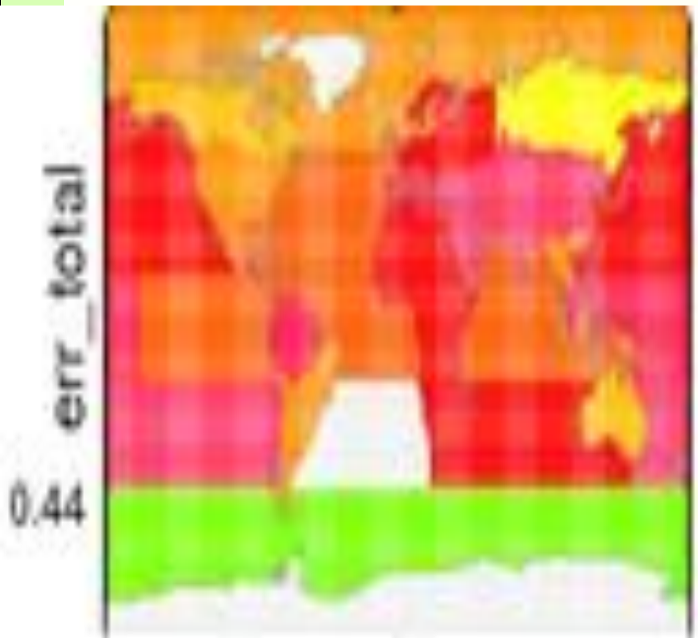
← underestimate overestimate →

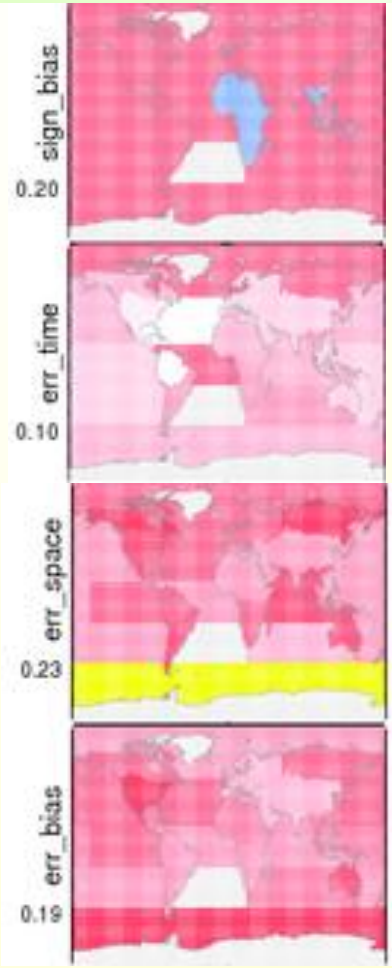
2004 AOD errors



error 0.44

(score = 0.56)





BIAS-sign

positive negative

temporal corr. error

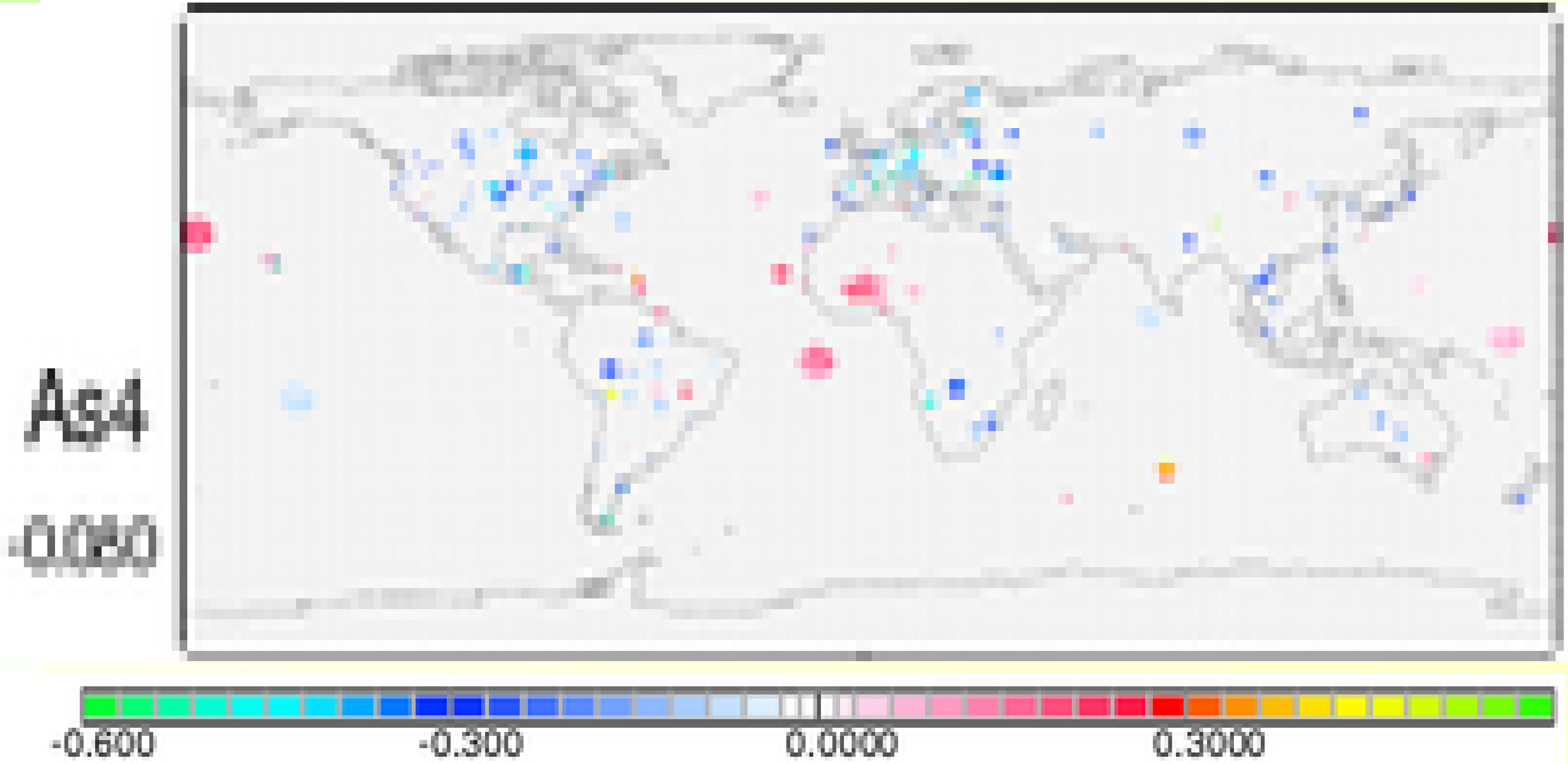
spatial correlat. error

bias strength error

increasing error →

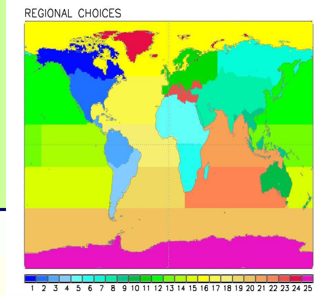
2004 Ang – how to quantify performance?

Angstrom difference to AERONET



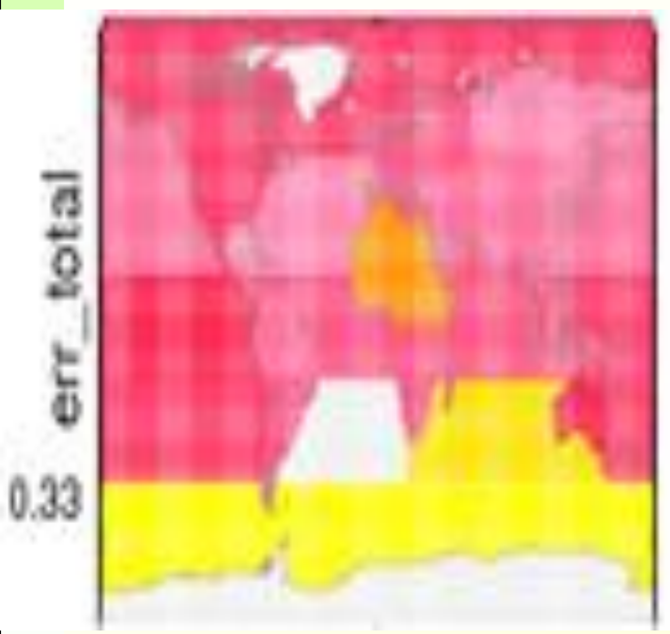
← underestimate overestimate →

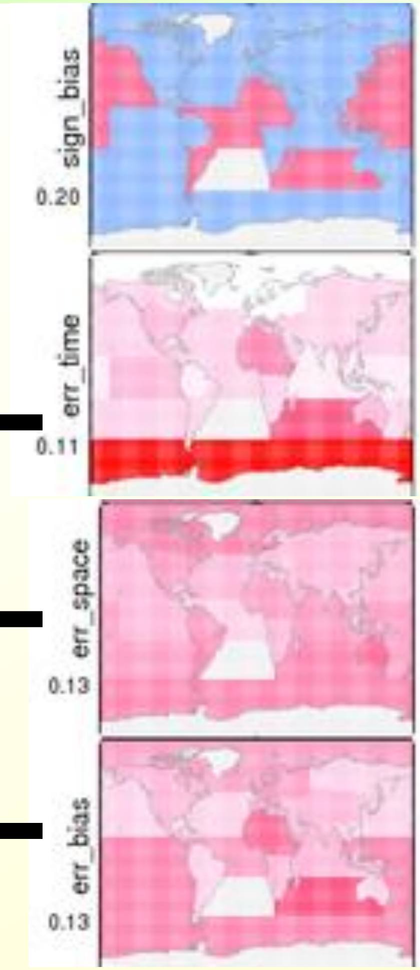
2004 Angstrom errors



error 0.33

(score = -0.67)





BIAS-sign

positive negative

temporal corr. error

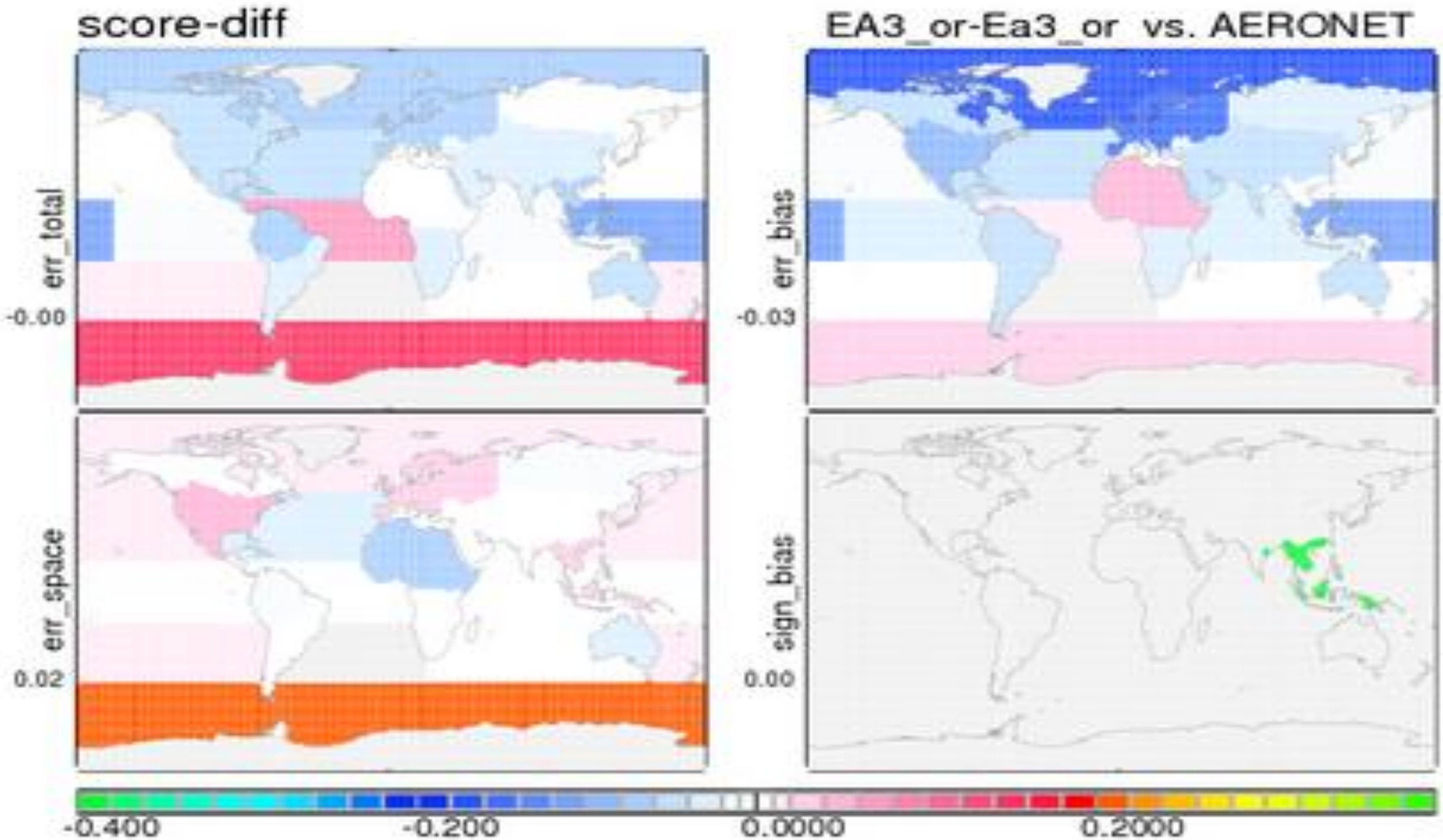
spatial correlat. error

bias strength error

increasing error →

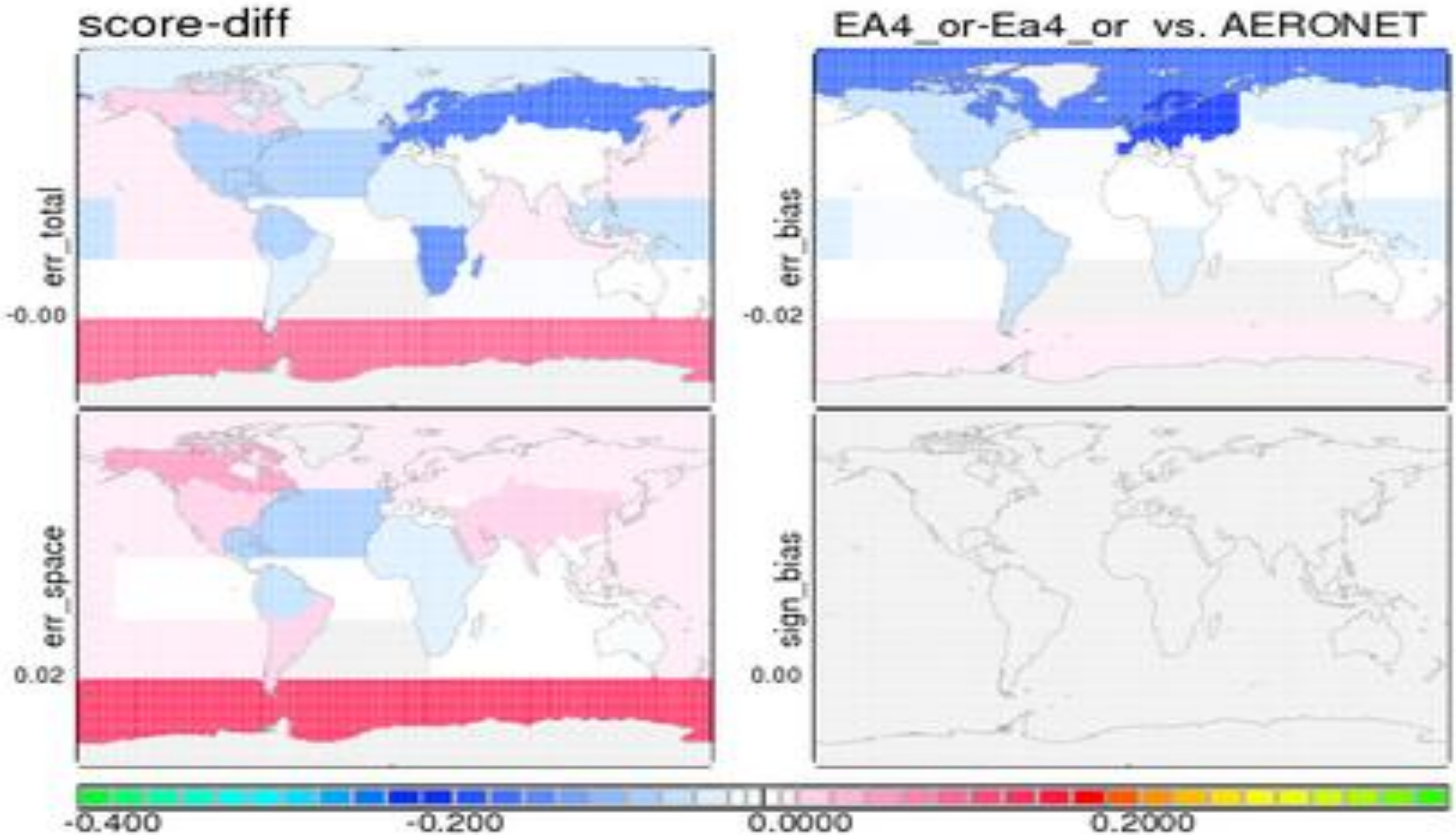
- comparing regional errors

AOD 2003 MACC vs GEMS



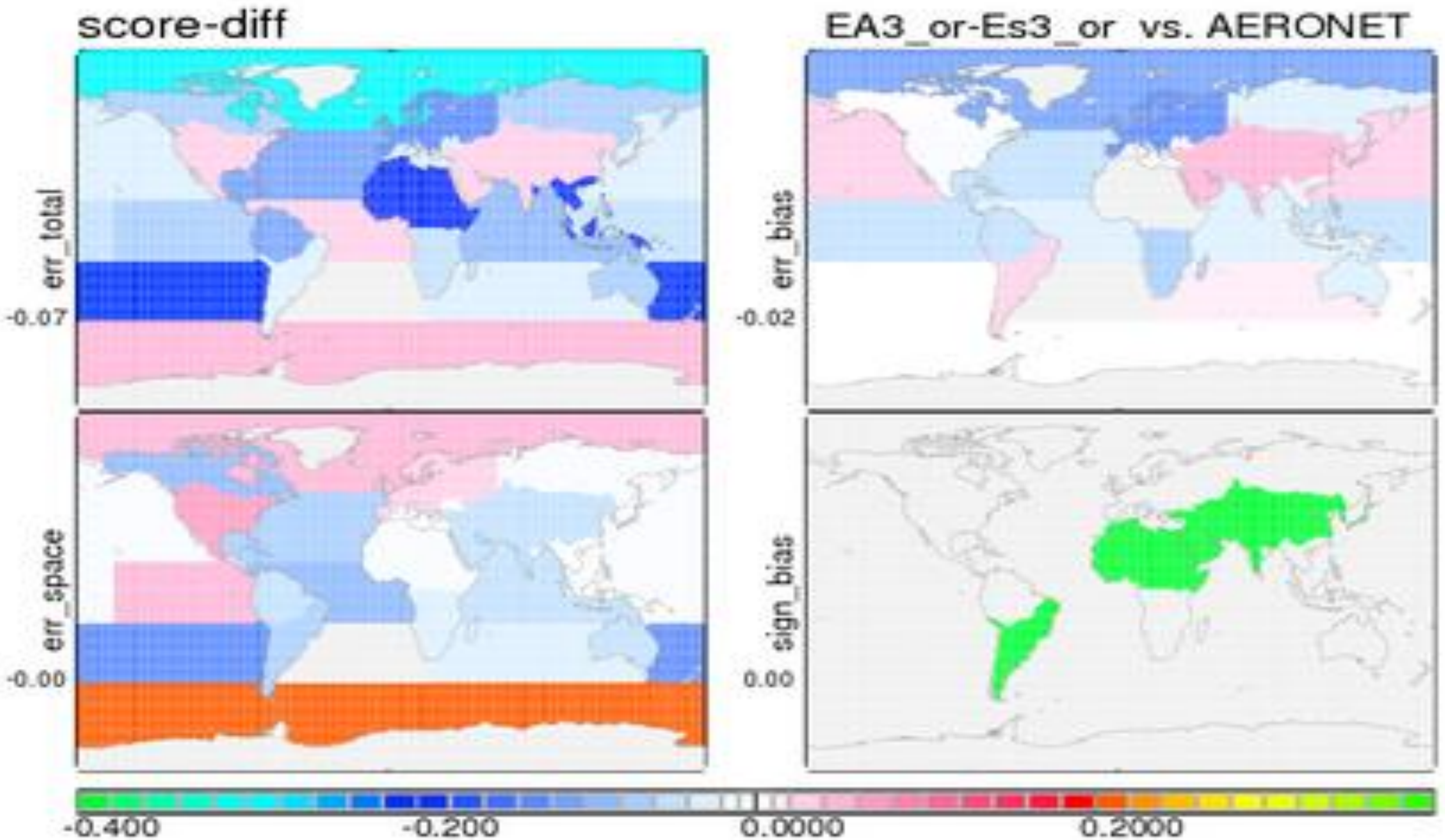
← **MACC better** **GEMS better** →

AOD 2004 MACC vs GEMS



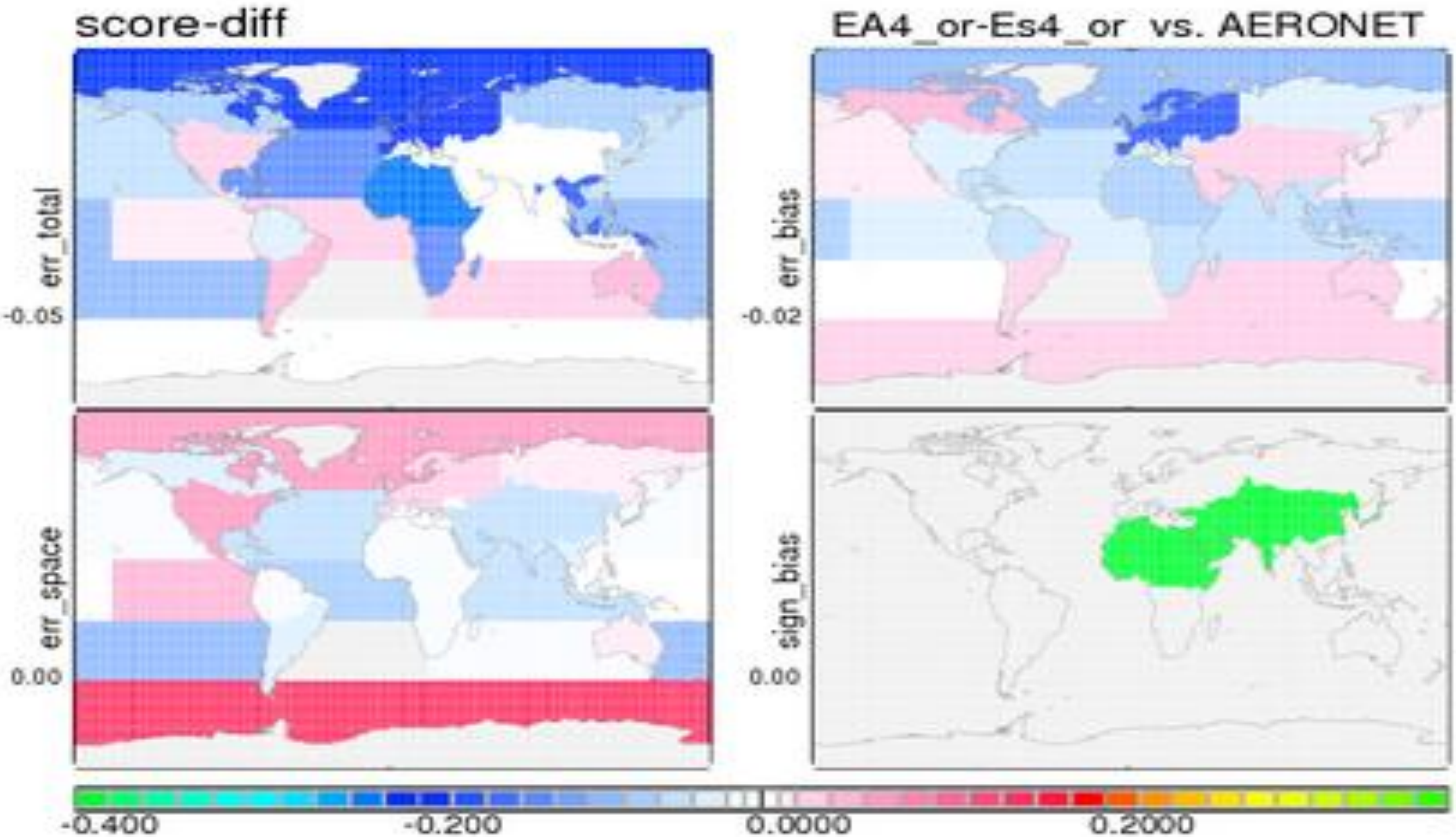
← MACC better GEMS better →

AOD 2003 MACC vs forecast



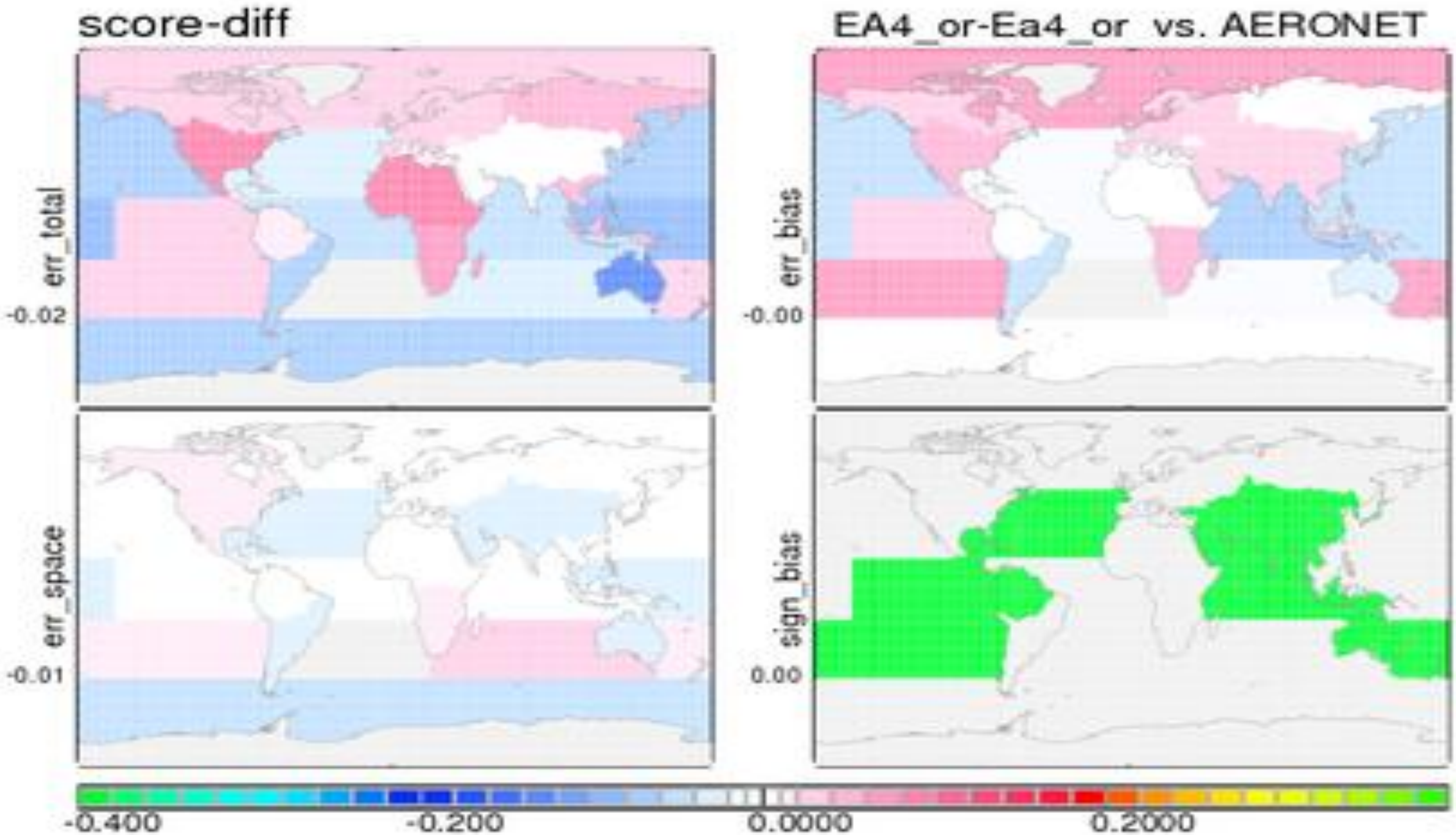
← MACC better forecast better →

AOD 2004 MACC vs forecast



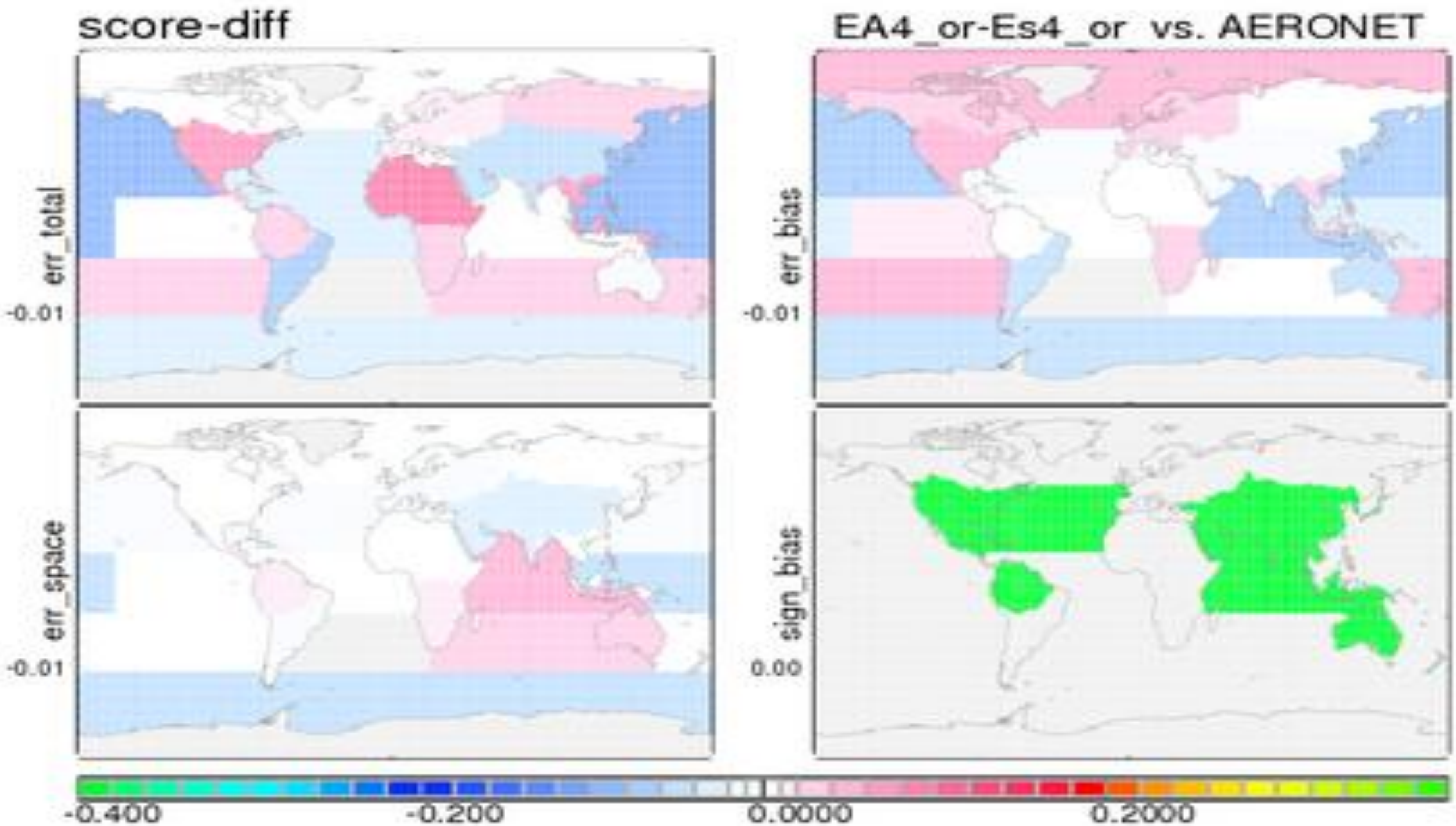
← MACC better forecast better →

Ang 2004 MACC vs GEMS



← MACC better GEMS better →

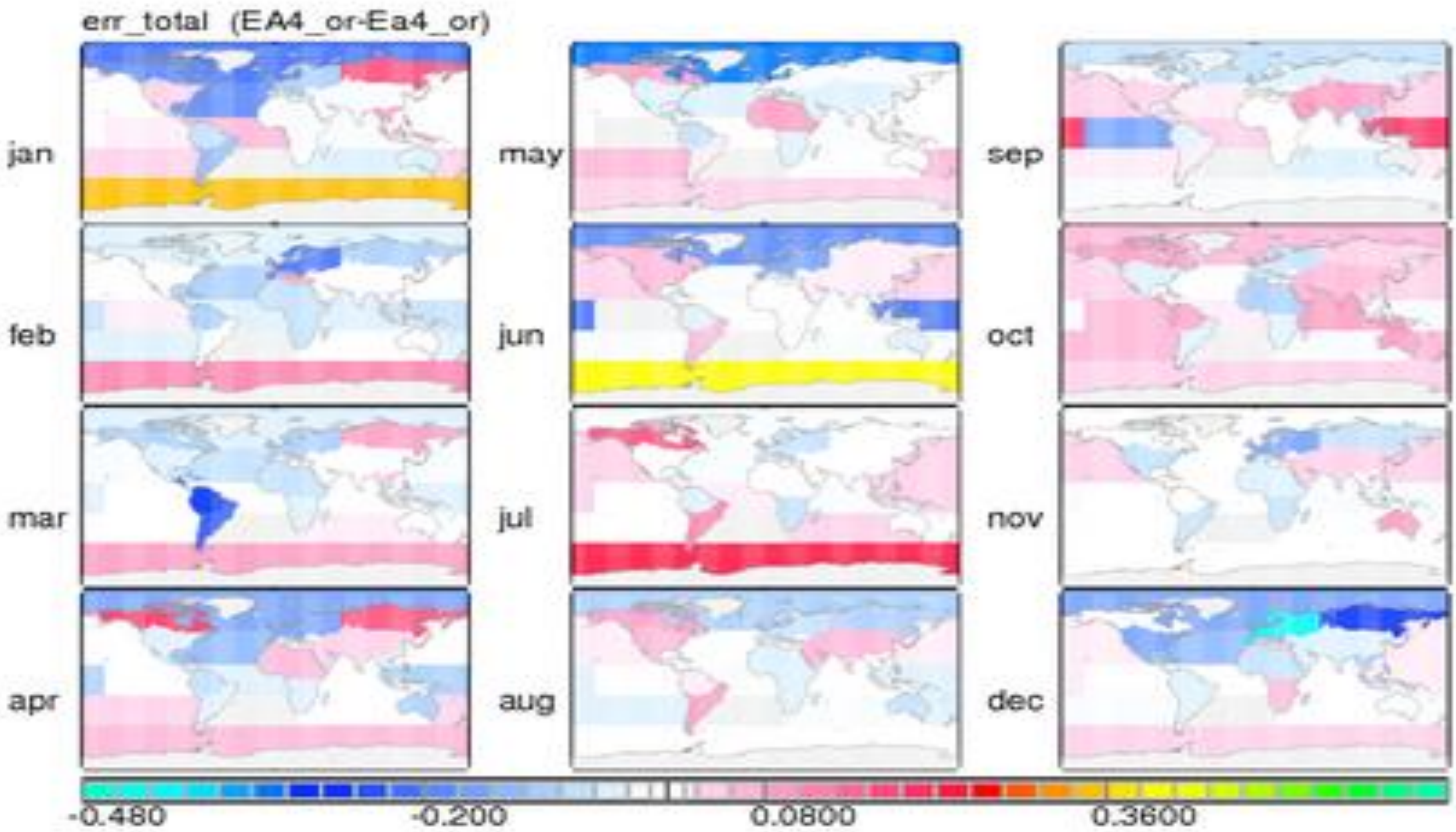
Ang 2004 MACC vs forecast



← MACC better forecast better →

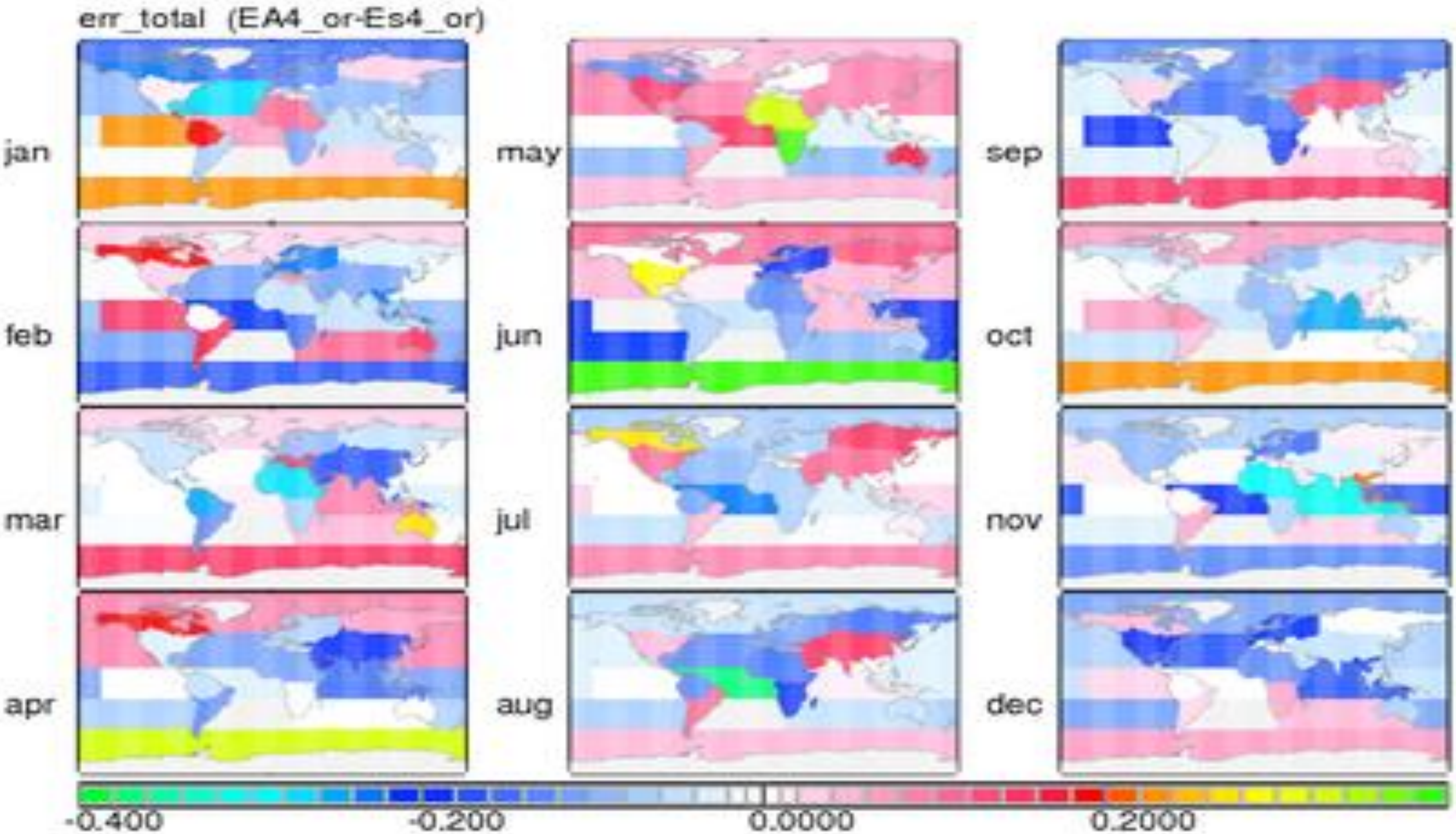
- now regionally and monthly errors

AOD 2004 MACC vs GEMS



← MACC better GEMS better →

AOD 2004 MACC vs forecast



← MACC better | forecast better →

- **and now an alternative method**

AeroCom

- open international science initiative for global aerosol modeling & comparisons to observations
- archive for aerosol global model data
  - http://dataipsl.ipsl.jussieu.fr/AEROCOM
  - IT infrastructure (idl, nco, perl, 10TB disk)
- annual workshops & sub-group activities
  - emission, u-physics, indirect effects, forcing ..
- steering group: M.Schulz, S.Kinne, M.Chin
  - funding (CNES, ESA, NASA, EU-projects)

AeroCom

- **and now some AeroCom web approaches**

  http://dataipsl.ipsl.jussieu.fr/AEROCOM

- **direct comparisons of simulations to**
  - **model simulation by many global model**
  - **comparison at AERONET sites**
  - **simple performance measures**

- **examples**
  - **latitidinal distribution at matches to AERONET**
  - **daily AOD correlations for the year 2003**

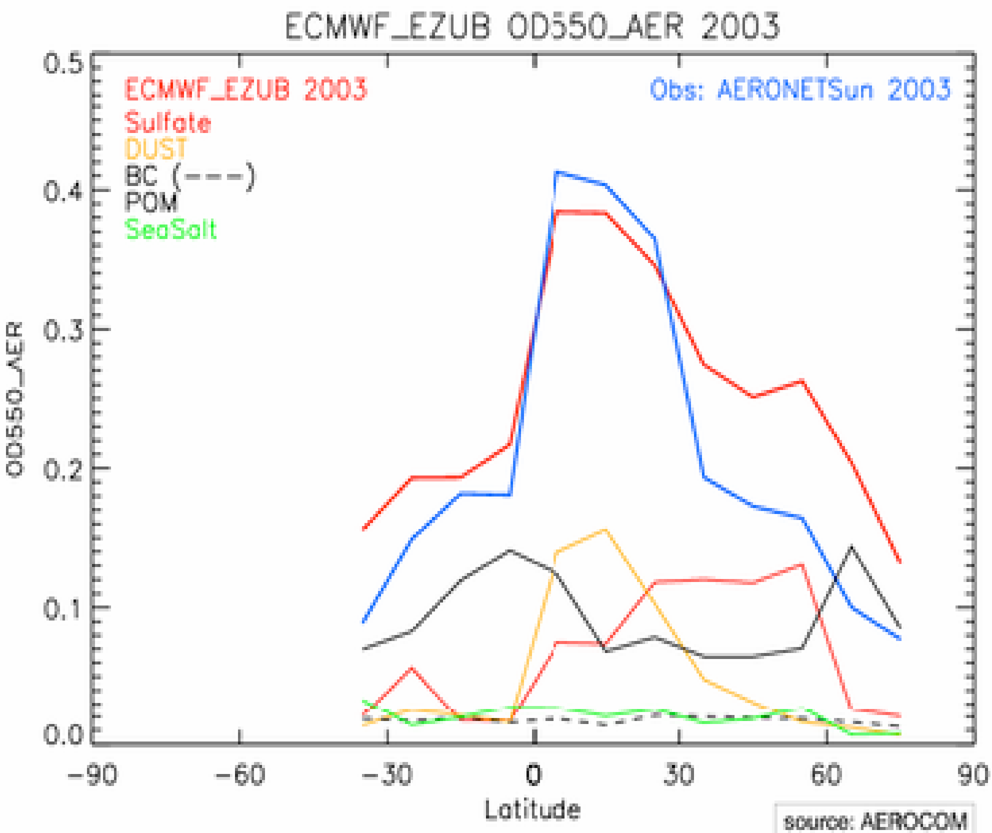
AOD match by latitude

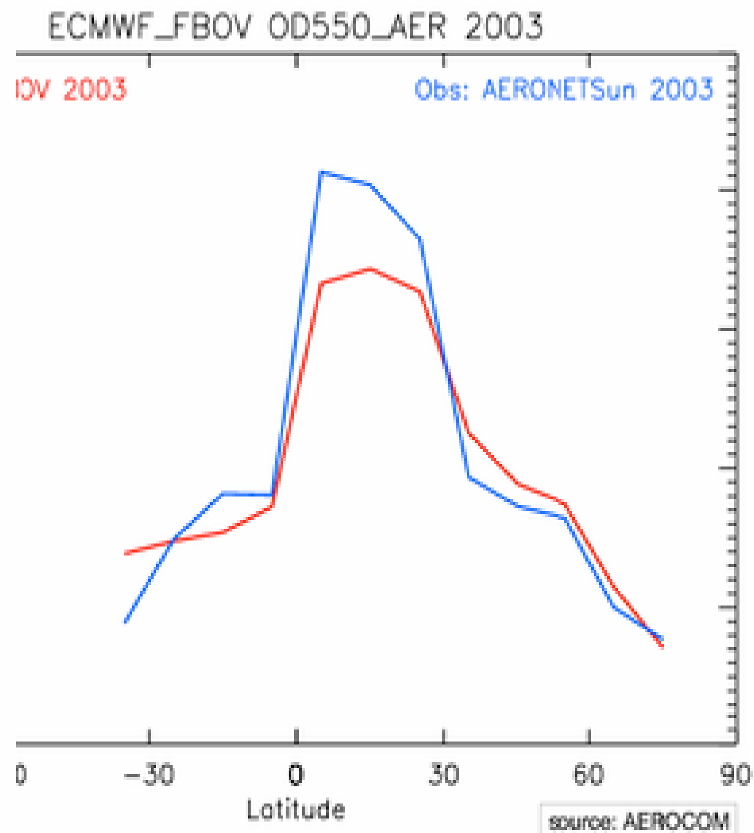
GEMS assimilation
with type stratification

AERONET

MACC assimilation

AERONET





statistics benchmarking

Progress of ECMWF-model assimilation MODIS-AOD into IFS

Aerocom global benchmarking against Aeronet+GAW+SKYNET – 1079 months / 2003 daily data / Stations below 1000m

	Correlation	RMS	Bias
1st forward model,SO4 error	0.70	0.13	+0.034
1st assimilation GEMS, SO4…	0.83	0.11	+0.057
2nd GEMS assimilation, SO4…	0.82	0.11	+0.047
1st MACC assimilation,	0.86	0.09	+0.005

lesson learned – AER-GEMS

- … it takes more time than you think
- assimilations of MODIS AOD improve the forecast
- improvements are weaker over land
  - ocean data are more accurate
- improvement is mainly of temporal nature
  - ‘events’ and overall seasonality
- improvement of AOD (amount) comes often at the expense of Angstrom (size)
  - additional assimilation of fine-mode-fraction ?

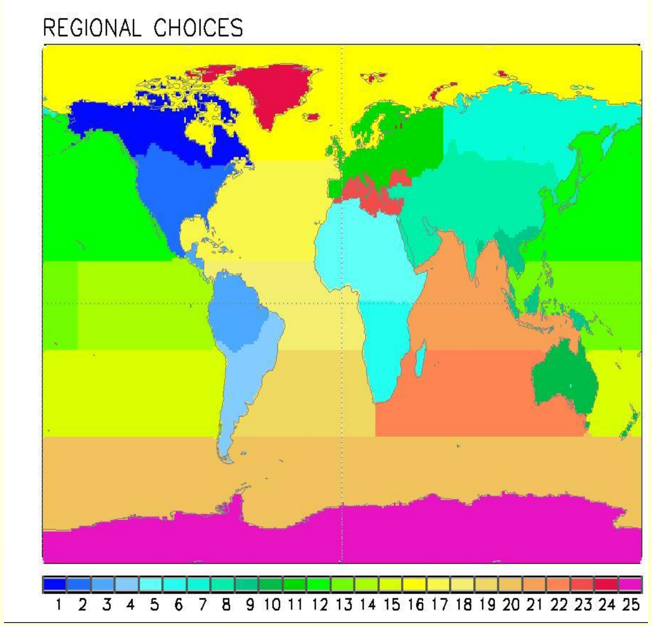
lesson learned – AeroCom

- common interests connect
  - the need to evaluate modeling
  - the need to connect model and data groups
    - to be exposed to all available data
    - to understand how data strength and limitation
    - to communicate data needs
  - a platform to interact
- major elements
  - data-sharing, web-tools, common papers
  - (annual) meetings (= reunions)
  - collaborative spirit (no money … no envy)

extras

the scoring

- sample at spatial and temporal sub-scales
  - month
  - regions ⇨



- spatial distribution score
  - spatial (rank-) correlation
- general bias score
  - compare ranks-sums
- seasonality score (only applied for annual scores)
  - temporal (rank-) correlation

rank based scoring - why ?

- **rank based scoring reduces the weight of outliers** ***(a better qualitative measure)***
- **rank correlation**
  - **are the ranks of data-pairs correlated ?**
- **rank bias**
  - **do the rank-sums compare ?** **… hereby data-set associated rank-sums are made up by the (value-) ranks of an array containing both data-sets**
  - ***example***
    - **set 1: 1 7 8** **value: 9 8 7 4 3 1** **rank-sum 1: 11**
    - **set 2: 3 4 9** **rank: 1 2 3 4 5 6** **rank-sum 2: 10**

bias = (1-2)/(1+2) = (11-10)/21 ~zero ⇨ no clear bias

scoring over scales

- individual scores for region and months
  - detail on local, seasonal performance and a tool for quantifying improvements
- combine monthly *(spatially correlation and bias scores)* to annual scores … and add a seasonality score (using monthly medians)
- combine regional annual scores into global scores (weigh by regional surface)
  - one-number summary
- combine scores of different properties

what is a good score ?

- each score S is defined via an error e
  - S = 1 – w*e, w is a weight factor based on the interquartile range – not to over-emphasize errors at low variability
  - correlation error = (1- correlation coeff.) /2.0
  - bias error = (sum1 – sum2) / (sum12)

    the sign of the bias matters and is carried on

- 1.0 is a perfect score … 0.0 is poorest
- total score = bias score
  * correlation score
  (* seasonality score)