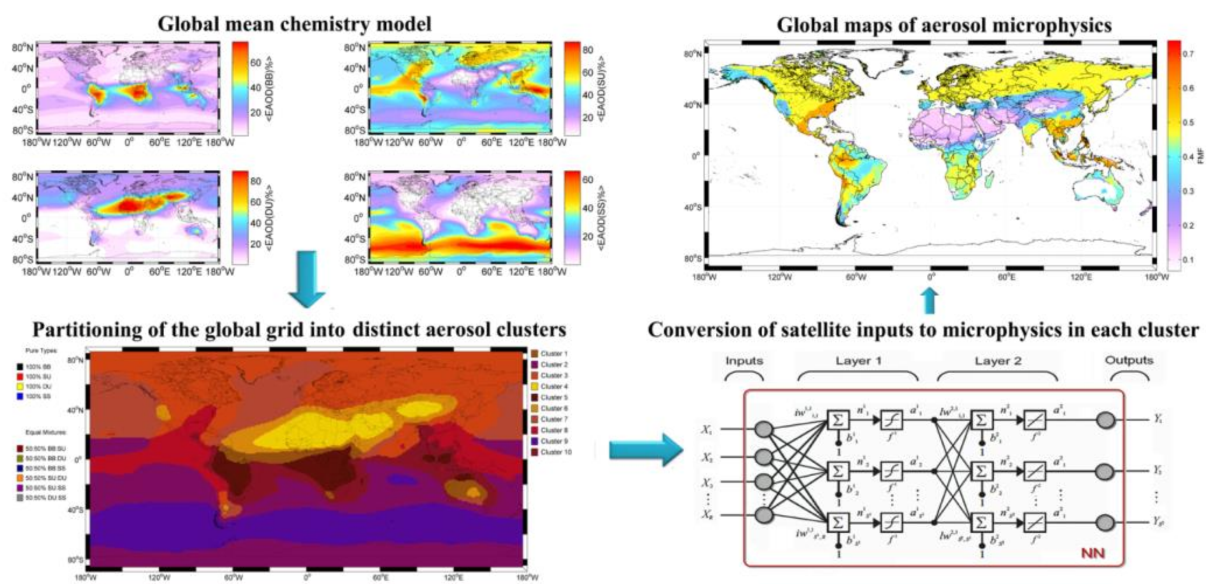
How a synergy of GOCART, MODIS and AERONET data can be used to train neural networks for producing global aerosol volume size distributions from space

Michael Taylor
National Observatory of Athens, Greece
mtaylor@noa.gr

2nd Gregory G. Leptoukh Online Giovanni Workshop:
November 13th, 2014 (Day 3): 10:00-11:00 (UTC/GMT-5)



OUTLINE:

- **Rationale** – to train neural networks (NNs) to invert satellite measurements to retrieve *daily* aerosol size distributions and associated microphysics globally at 1x1 degree
- **Prototype** – a NN model trained and tested for Saharan dust aerosol in Northern Africa
- **Global aerosol model** – a 3-step approach:
- STEP 1) application of ***cluster analysis*** to mean global GOCART aerosol optical depth data (per aerosol type) and identification of the composition & spatial distribution of aerosol mixtures at 1x1 degree
- STEP 2) train a NN to invert satellite measurements for each aerosol mixture/region to produce global maps of size distributions
- STEP 3) multimodal fitting & analysis of size distributions for determination of characteristic aerosol size & volume information
- **Case study** - quasi-realtime monitor (4-day average) of the Karthala volcanic eruption 12-23 January, 2007

Michael Taylor: mtaylor@noa.gr

DATA SOURCES: *A word of thanks*

A challenging experiment *conducted on a single 16Gb RAM, quad-core PC with MATLAB*:

From GIOVANNI 3 @NASA/GES-DISC: http://gdata1.sci.gsfc.nasa.gov

- Aqua/MODIS (Level 3 Collection 5 daily global 1x1 gridded values): 16 parameters x 360 degrees x 180 degrees x 3575 days [34.4Gb]
- Aura/OMI (Level 3 daily global 1x1 gridded values): 4 parameters x 360 degrees x 180 degrees x 3362 days [9.7Gb]

From GIOVANNI 4 @NASA/GES-DISC: http://giovanni.gsfc.nasa.gov/giovanni/

- GOCART (Version 4): 5 parameters x 2.5 x 2 gridded values (interpolated to 1x1) x 7 years x 12 months

From AERONET: http://aeronet.gsfc.nasa.gov

- 160 aerosol optical & microphysical parameters (Levels 1.5 and 2.0/Version 2 Inversion Products) x 715,288 records (969 sites) [7.2Gb]

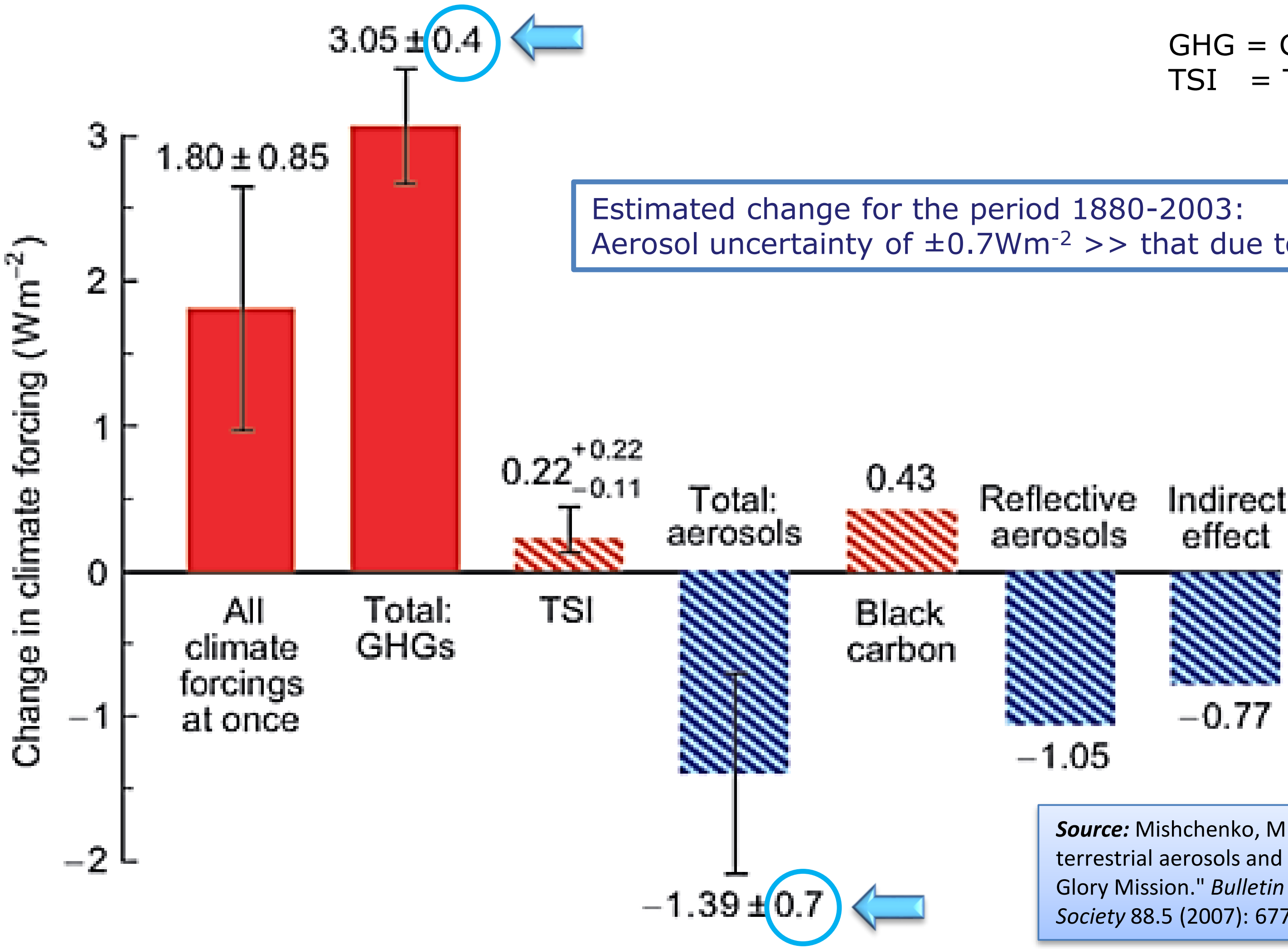
…which *would not have been possible* without the generous provision of open data by NASA/GES-DISC and the pioneering efforts of Gregory G. Leptoukh and co-workers to advance data interfacing, accessibility & visualization.

Michael Taylor: mtaylor@noa.gr

RATIONALE

CONTEXT: *There is a need to reduce uncertainty in aerosol climate forcing*

GHG = Green-House Gases
TSI = Total Solar Irradiance



Estimated change for the period 1880-2003:
Aerosol uncertainty of ±0.7Wm^{-2} >> that due to GHGs (0.4Wm^{-2})

Source: Mishchenko, M et al (2007) "Accurate monitoring of terrestrial aerosols and total solar irradiance: introducing the Glory Mission." *Bulletin of the American Meteorological Society* 88.5 (2007): 677-691.

Michael Taylor: mtaylor@noa.gr

CONTEXT: *Reducing uncertainty requires better aerosol characterization*

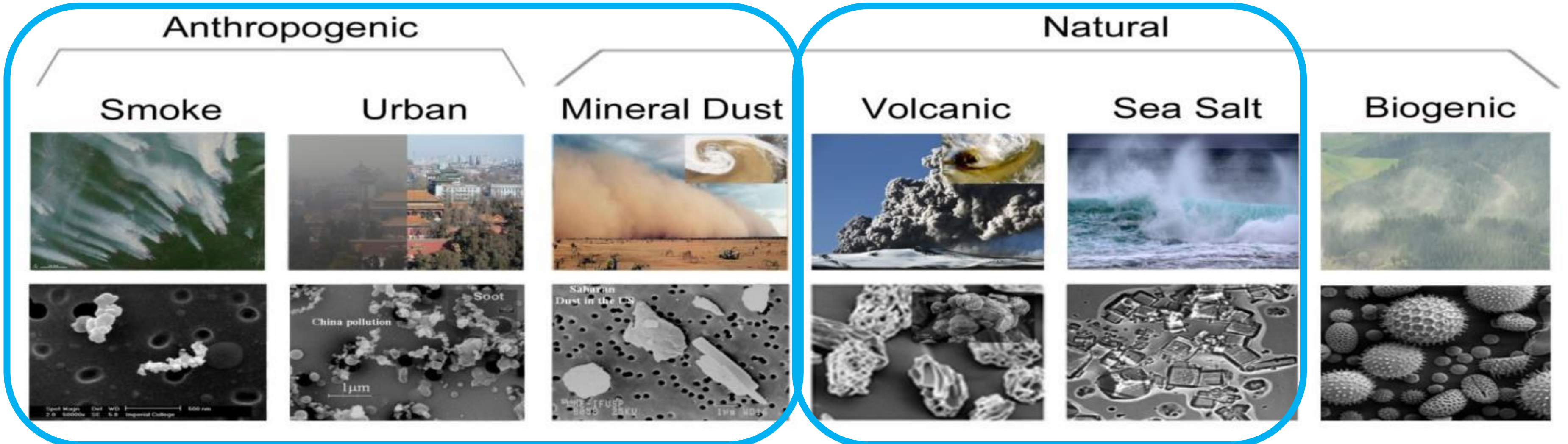
Year 2000 Emissions Tg yr⁻¹ or TgS yr⁻¹	Anthropogenic NMVOCs		Anthropogenic Black Carbon		Anthropogenic POA		Anthropogenic SO_2		Anthropogenic NH_3		Biomass Burning Aerosols	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Total	98.2	157.9	3.6	6.0	6.3	15.3	43.3	77.9	34.5	49.6	29.0	85.3

Source: IPCC/AR5 (2013)

Source	Natural Global	
	Min	Max
Sea spray	1400	6800
Mineral dust	1000	4000
Terrestrial PBAPs	50	1000
Dimethylsulphide (DMS)	10	40
Monoterpenes	30	120
Isoprene	410	600
SOA production from all BVOCs	20	380

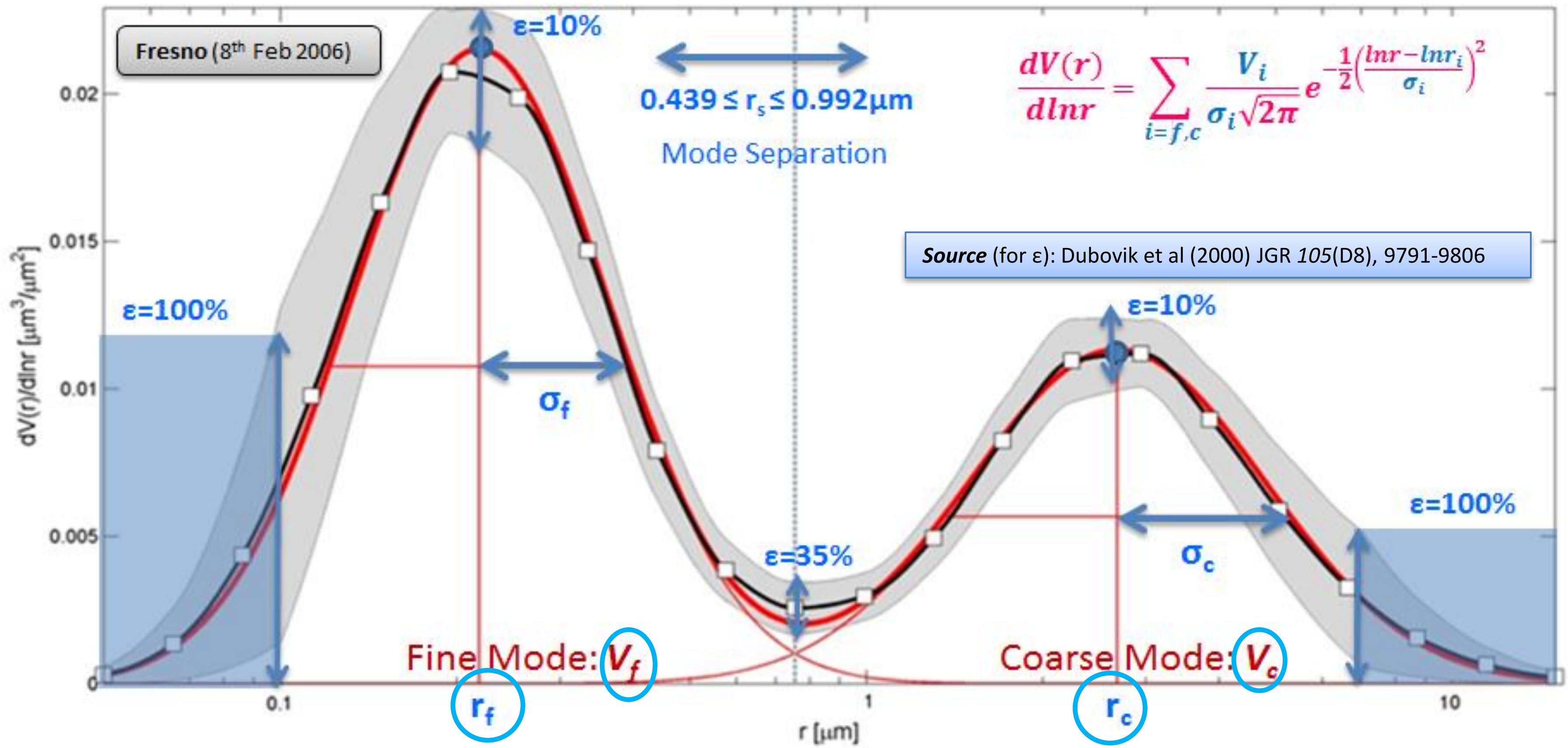
Dominant emissions are from deserts & oceans

Characterization = *f* (**size**, **volume**, shape, composition, source)



Michael Taylor: mtaylor@noa.gr

CONTEXT: *Aerosol volume size distribution (AVSD) – AERONET method*



$$V_i = \int_{r_1}^{r_2} \frac{dV(r)}{d\ln r}\, d\ln r$$

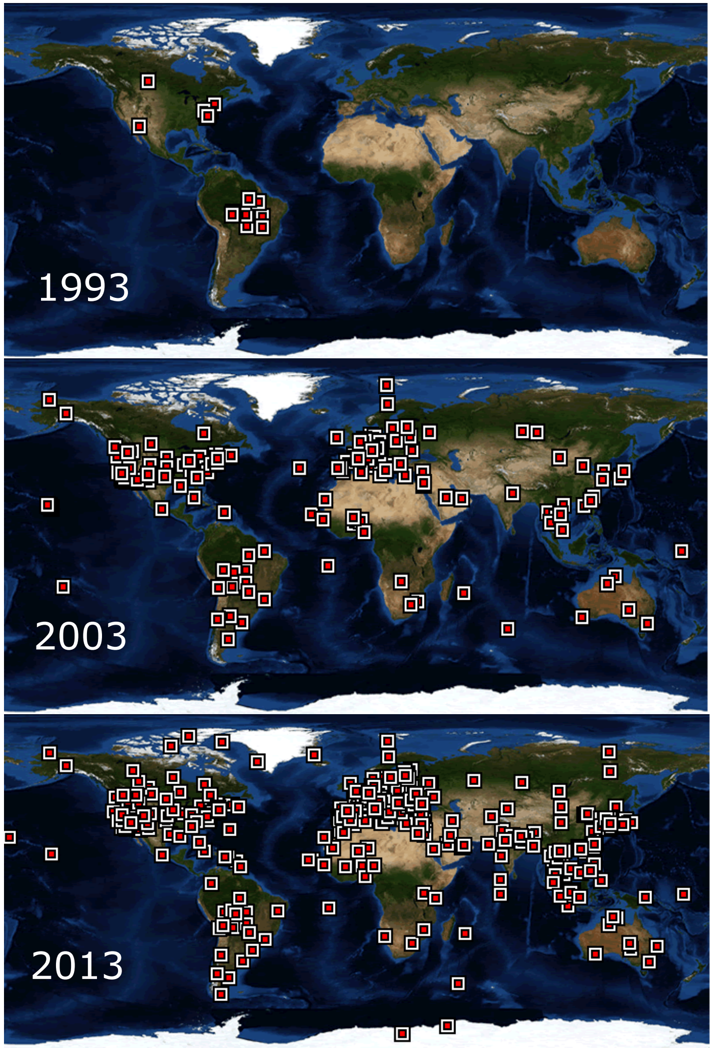
$$\ln r_i = \frac{\int_{r_1}^{r_2} \ln r \frac{dV(r)}{d\ln r}\, d\ln r}{\int_{r_1}^{r_2} \frac{dV(r)}{d\ln r}\, d\ln r}$$

$$\sigma_i = \sqrt{\frac{\int_{r_1}^{r_2} (\ln r - \ln r_i)^2 \frac{dV(r)}{d\ln r}\, d\ln r}{\int_{r_1}^{r_2} \frac{dV(r)}{d\ln r}\, d\ln r}}$$

Source: Taylor, Kazadzis & Gerasopoulos (2014) *Proceedings of the 12th International Conference on Meteorology, Climatology & Atmospheric Physics (COMECAP)*, Heraklion, Crete, 28-31 May, Vol. 3, 191-196.

Michael Taylor: mtaylor@noa.gr

CHALLENGE: *AERONET stations have inhomogeneous spatial coverage*



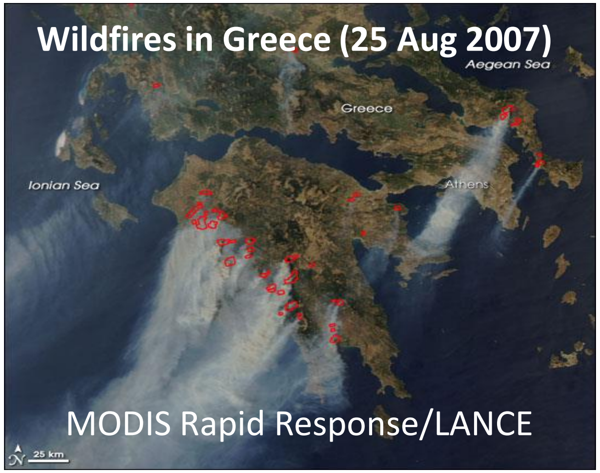
Aerosol inversion products date back to 27 May, 1993 @GSFC

Now >1000 sites…
… occupying only 343 pixels of the global grid (1x1 degree):

= 343/(360*180)

=0.54% membership of surface pixels

AERONET sites in Greece providing (Level 1.5) inversions during August 2007



Source: http://earthobservatory.nasa.gov

Source: http://aeronet.gsfc.nasa.gov

Source: http://aeronet.gsfc.nasa.gov

Michael Taylor: mtaylor@noa.gr

IDEA: *Retrieve AERONET parameters by inverting SATELLITE measurements*

Satellite inputs
(gridded 1x1 degree)

Function Approximator

"AERONET" outputs
(gridded 1x1 degree)

MODIS H2O (to cloud)
MODIS AOD (470nm)
MODIS AOD (550nm)
MODIS AOD (660nm)
OMI AOD (380nm)
OMI AOD (500nm)
OMI AAOD (500nm)

From: Giovanni 3

Neural Network (NN)

- H2O=Precipitble water vapour
- AOD=Aerosol optical depth
- AAOD=Absorption AOD
- AVSD=Aerosol volume size distribution
- CRI-R=Complex refractive index (real part)
- CRI-R=Complex refractive index (imaginary part)
- SSA=Single scattering albedo
- ASYM=asymmetry factor

AVSD (22 radial bins)
CRI-R (440, 675, 870, 1020nm)
CRI-I (440, 675, 870, 1020nm)
SSA (440, 675, 870, 1020nm)
ASYM (440, 675, 870, 1020nm)

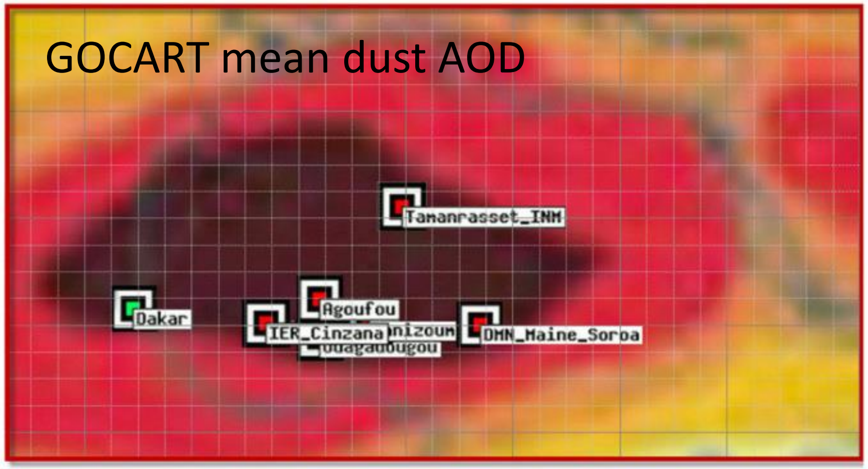
From: AERONET

Michael Taylor: mtaylor@noa.gr

PROTOTYPE:

A NN model for retrieval of Saharan Dust size distributions in Northern Africa

DATA: *GOCART-selected Northern African "dust" sites for the NN model*

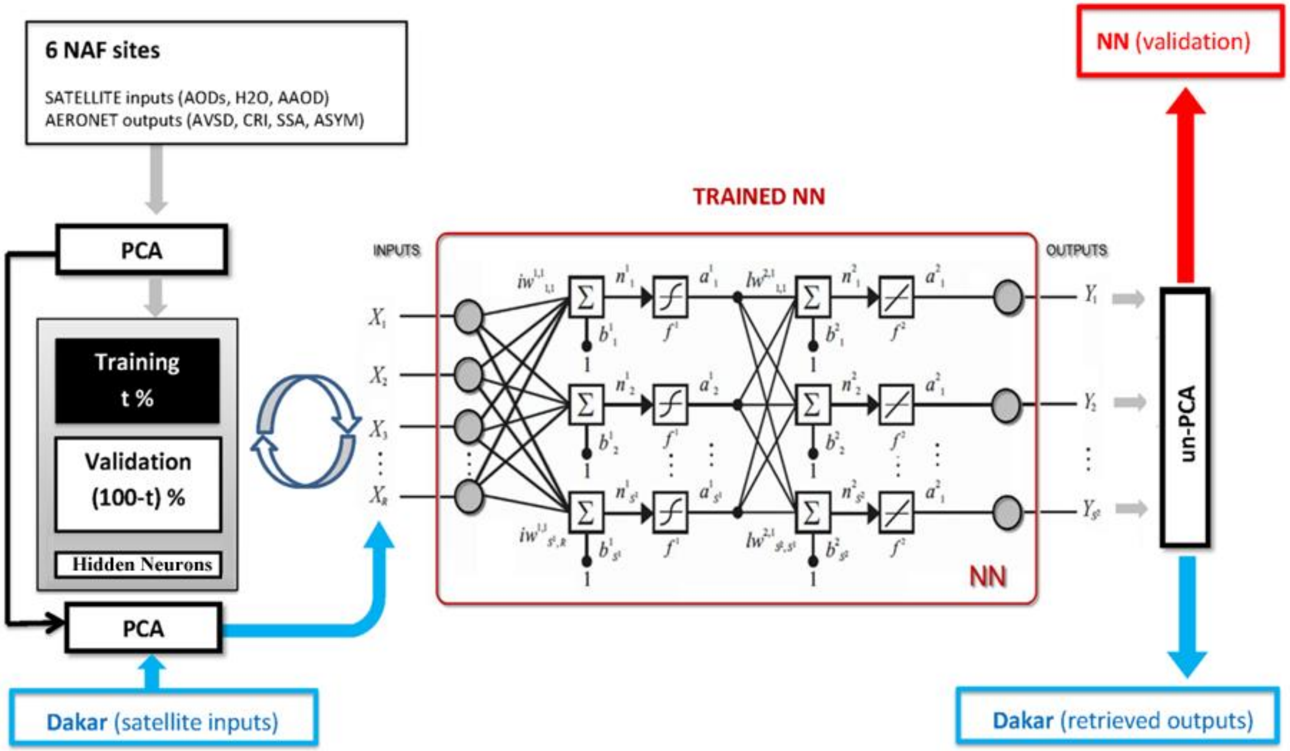


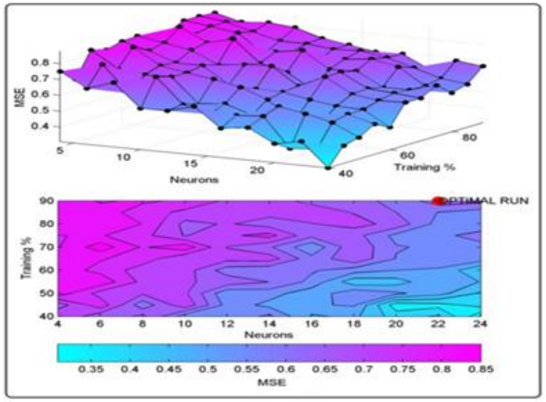
	SITE	*N*	GOCART Mean AOD & aerosol composition					
			< AOD >	% SO_2	% OC	% BC	% Sea Salt	% Dust
TRAINING	Tamanrasset INM	407	0.793	4.54 %	1.39 %	0.63 %	0.13 %	93.44 %
	Agoufou	1028	0.973	3.70 %	2.47 %	0.82 %	0.10 %	92.91 %
	Banizoumbou	2283	0.920	4.57 %	3.48 %	1.09 %	0.11 %	90.76 %
	DMN Maine Soroa	680	0.967	5.27 %	3.52 %	1.14 %	0.10 %	90.07 %
	IER Cinzana	1469	0.823	4.86 %	4.62 %	1.22 %	0.12 %	89.19 %
	Ouagadougou	966	0.776	6.06 %	7.47 %	1.93 %	0.13 %	84.41 %
SIMULATION	Dakar	1583	0.705	7.38 %	5.53 %	1.42 %	0.71 %	84.82 %

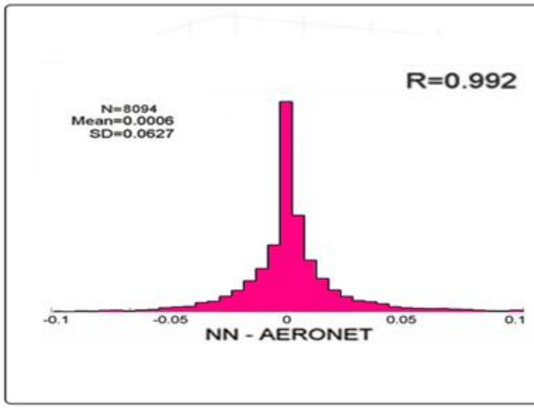
Source: Taylor, M., Kazadzis, S., Tsekeri, A., Gkikas, A., Amiridis, V. (2014) *Satellite retrieval of aerosol microphysical and optical parameters using neural networks: a new methodology applied to the Sahara desert dust peak*. Atmospheric Measurement Techniques 7, 3151-3175.

Michael Taylor: mtaylor@noa.gr

NN MODEL: *Scheme for objectivizing the NN architecture*







Source: Taylor et al (2014) *AMT* 7, 3151-3175

<u>A grid of 100 NNs were trained</u>:

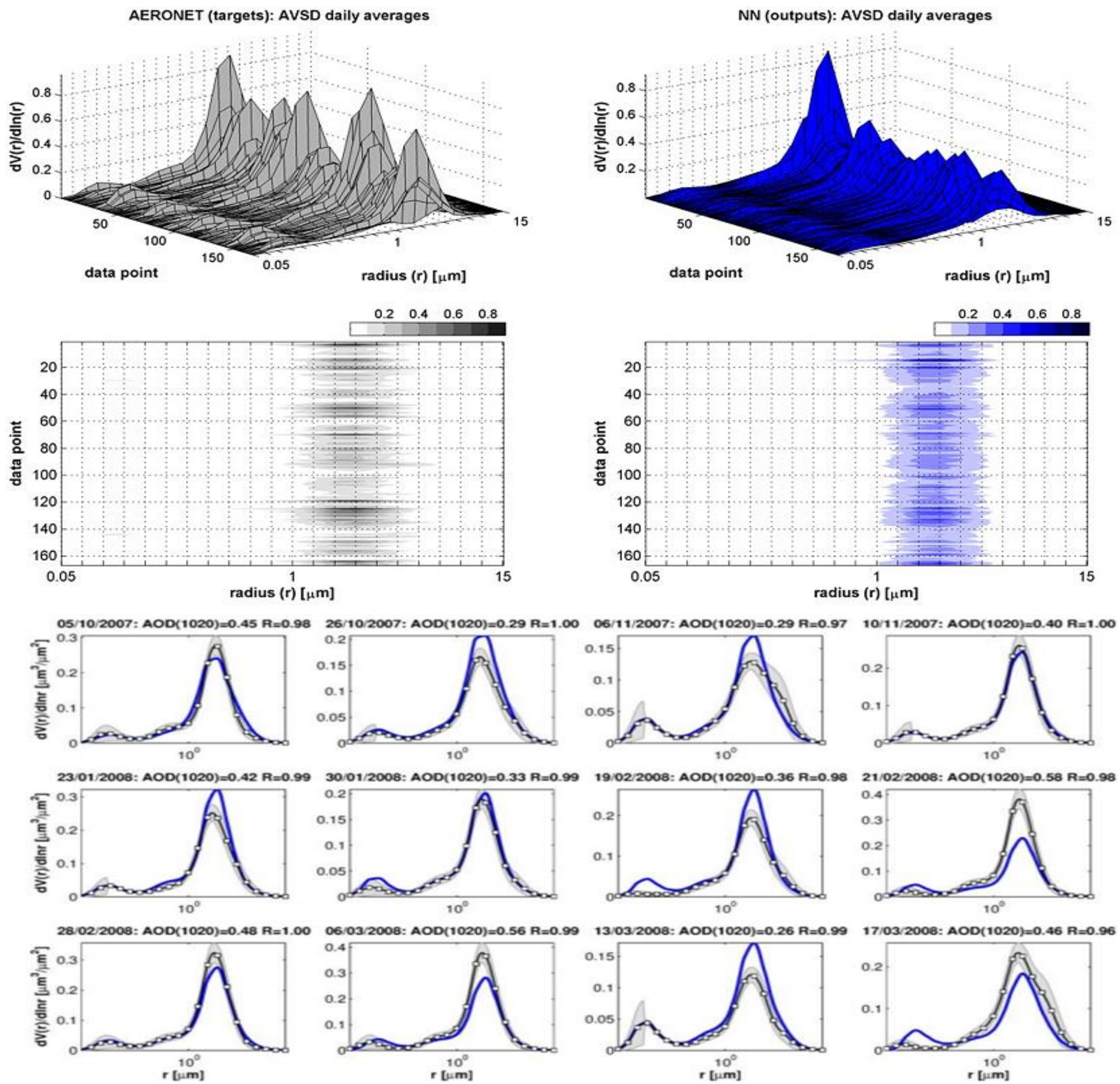
- Training fraction (t%) = [40%:5%:90%]
- *Tanh* hidden neurons = [4:2:24]

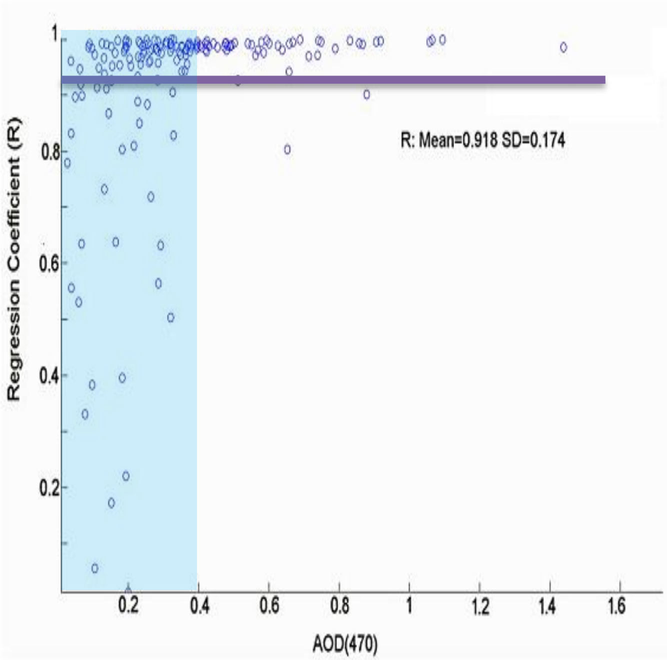
Minimum mean squared error → **Optimal NN**

PCA = Principal Components Analysis (98% variance)

Michael Taylor: mtaylor@noa.gr

NN SIMULATION (Dakar): *AVSD (181 co-located daily averages)*





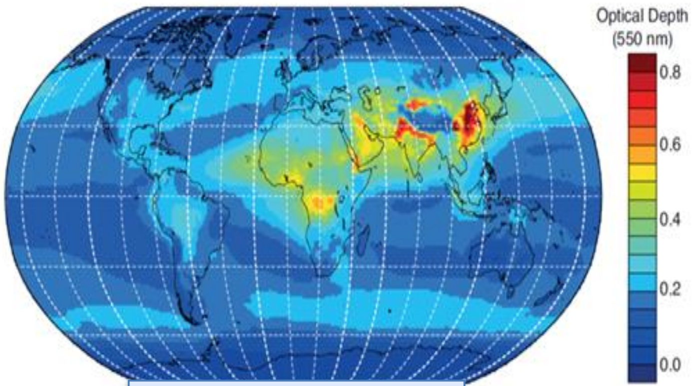
Good ***daily*** retrieval of the AVSD
for Saharan dust (<u>only</u>)

Source: Taylor, M et al (2014) *AMT* 7, 3151-3175

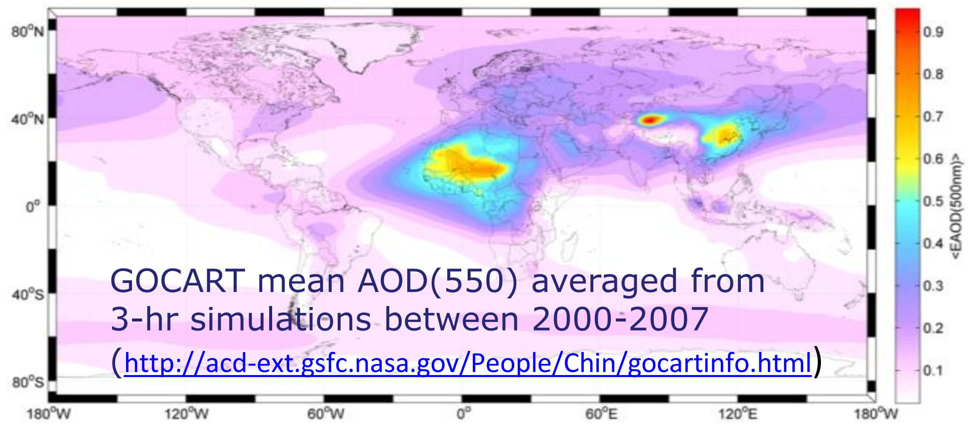
Michael Taylor: mtaylor@noa.gr

GLOBAL AEROSOL MODEL

STEP 1: *Use GOCART to partition by **mean** aerosol type **mixture***

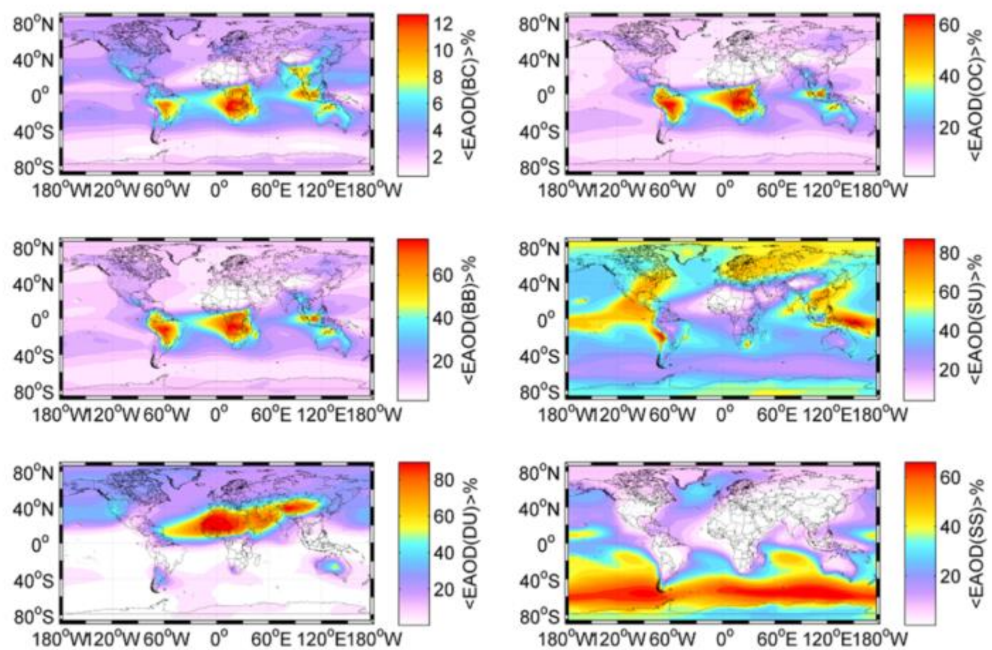


Source: IPCC/AR5 (2013)



GOCART mean AOD(550) averaged from 3-hr simulations between 2000-2007 (http://acd-ext.gsfc.nasa.gov/People/Chin/gocartinfo.html)

GOCART 2000-2007 global mean AOD per type : (BC, OC, SU, DU, SS)

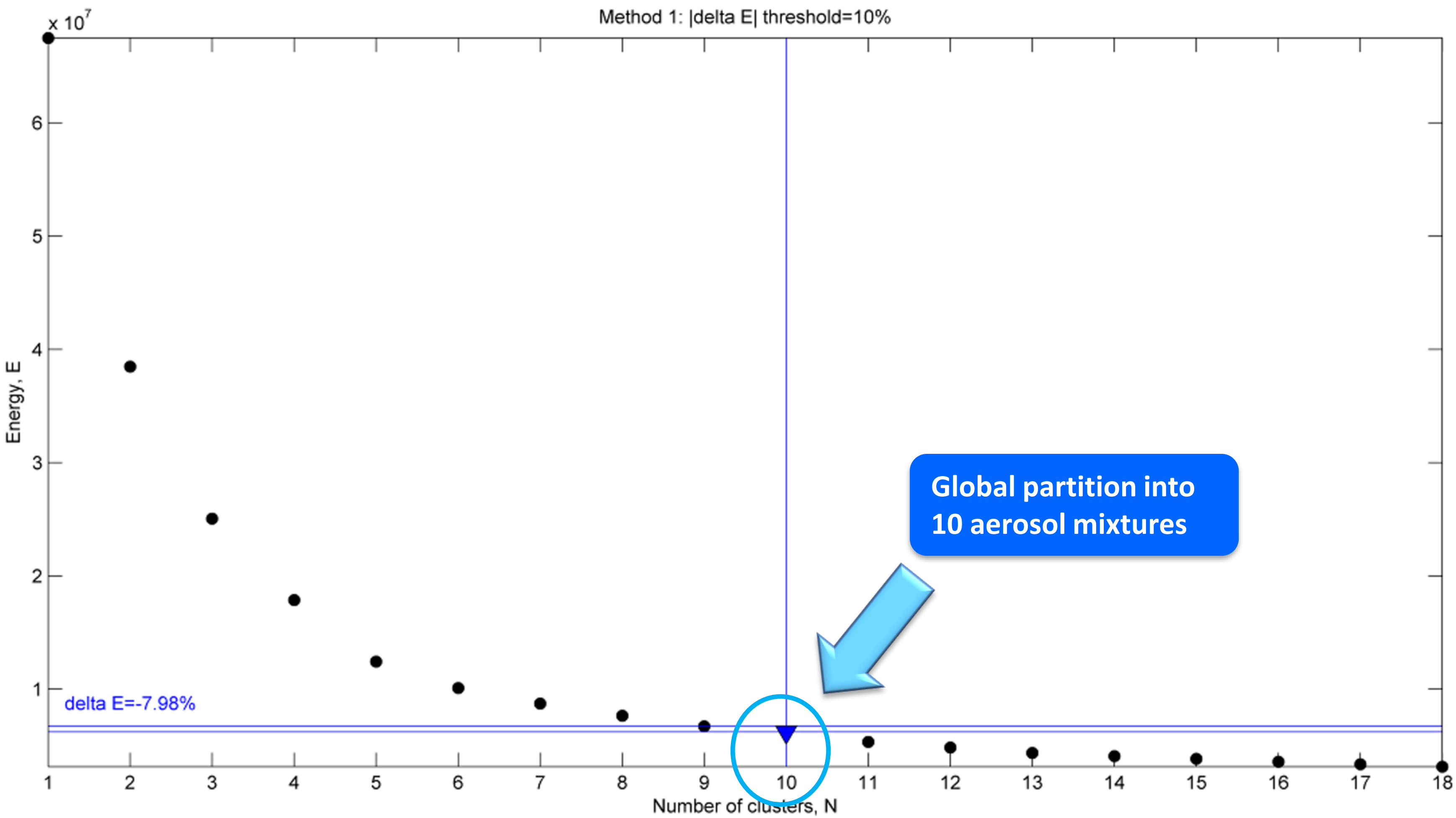


Source: MT

BC=Black Carbon
OC=Organic Carbon
SU=Sulfate
DU=Dust
SS=Sea Salt

Michael Taylor: mtaylor@noa.gr

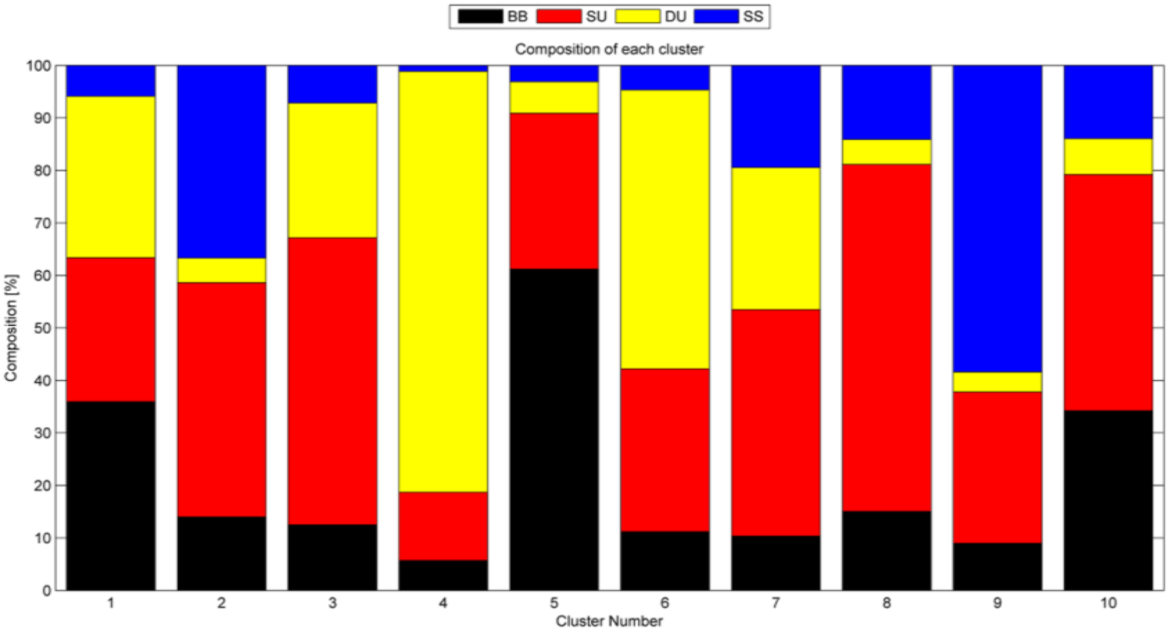
CLUSTER ANALYSIS: *of GOCART AOD (BB=BC+OC, SU, DU, SS)*



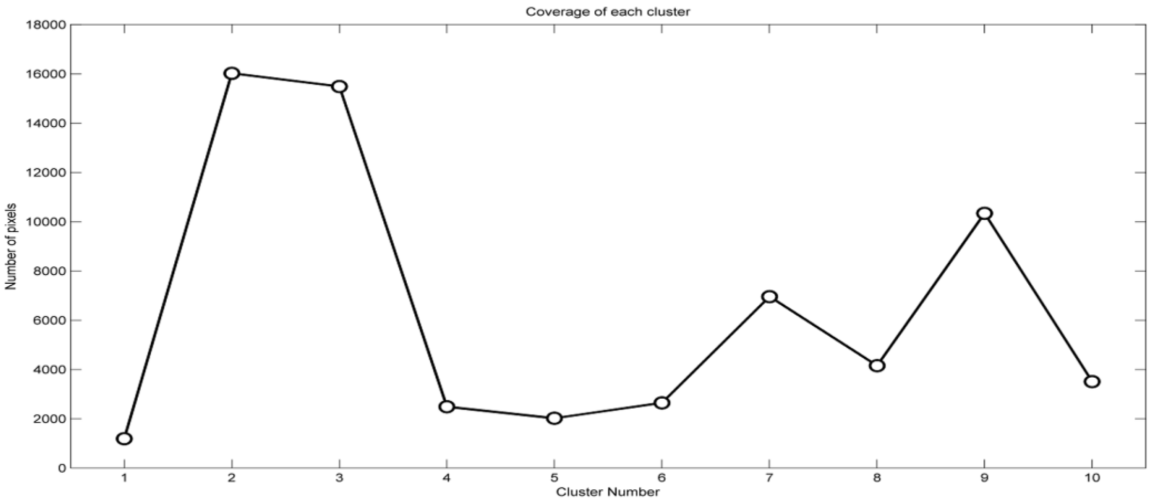
Add clusters (10 random seeds each time) until cluster centers do not change by > 10%

Michael Taylor: mtaylor@noa.gr

CLUSTER ANALYSIS: *Composition of 10 aerosol mixtures*



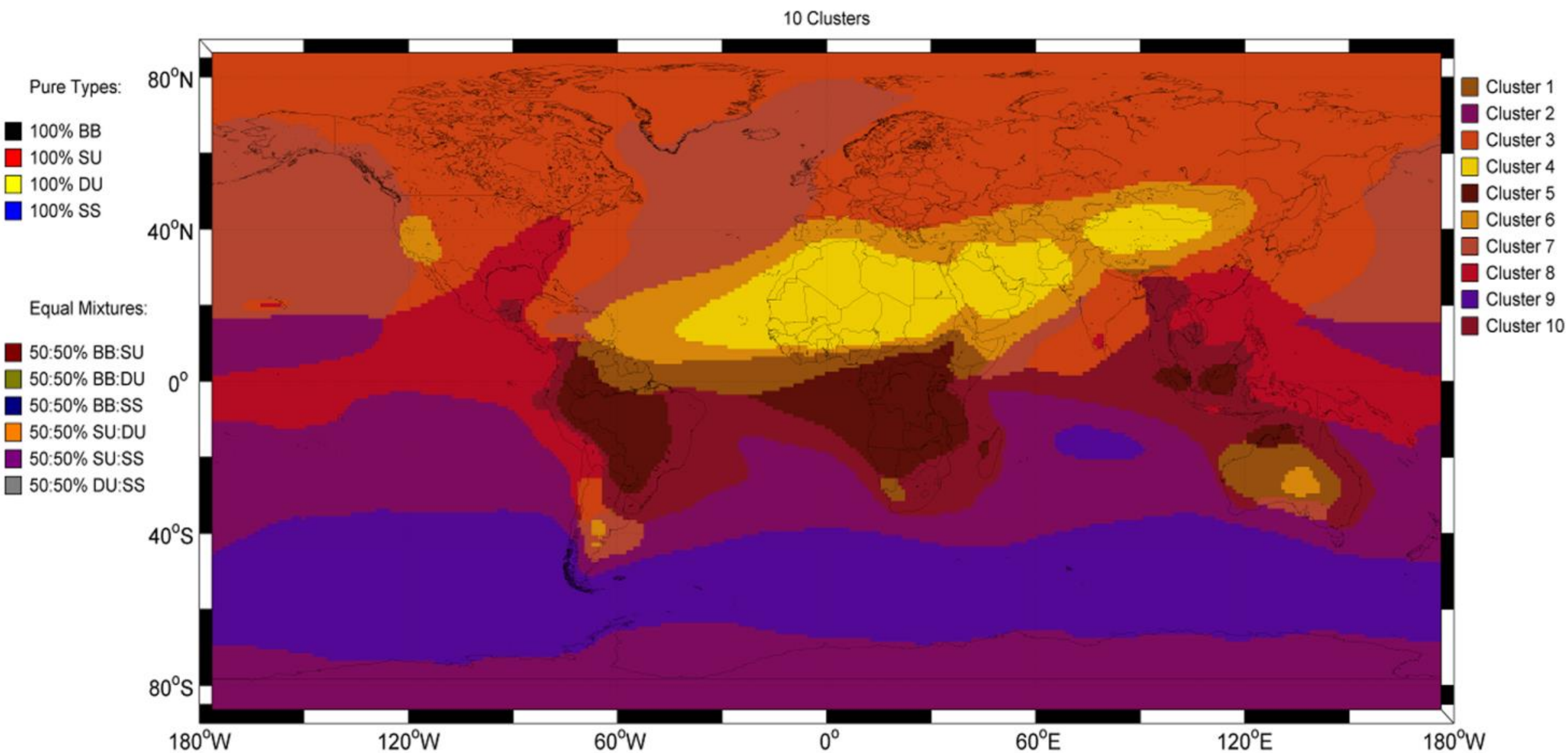
Cluster 1 = Dusty Sulfurous **SMOKE**
Cluster 2 = Marine **SULFATE**
Cluster 3 = Dusty **SULFATE**
Cluster 4 = **DUST**
Cluster 5 = Sulfurous **SMOKE**
Cluster 6 = Sulfurous **DUST**
Cluster 7 = Dusty Marine **SULFATE**
Cluster 8 = **SULFATE**
Cluster 9 = Sulfurous **MARINE**
Cluster 10 = Smokey **SULFATE**



Minimum cluster coverage > 1000 pixels (1x1 degree)

Michael Taylor: mtaylor@noa.gr

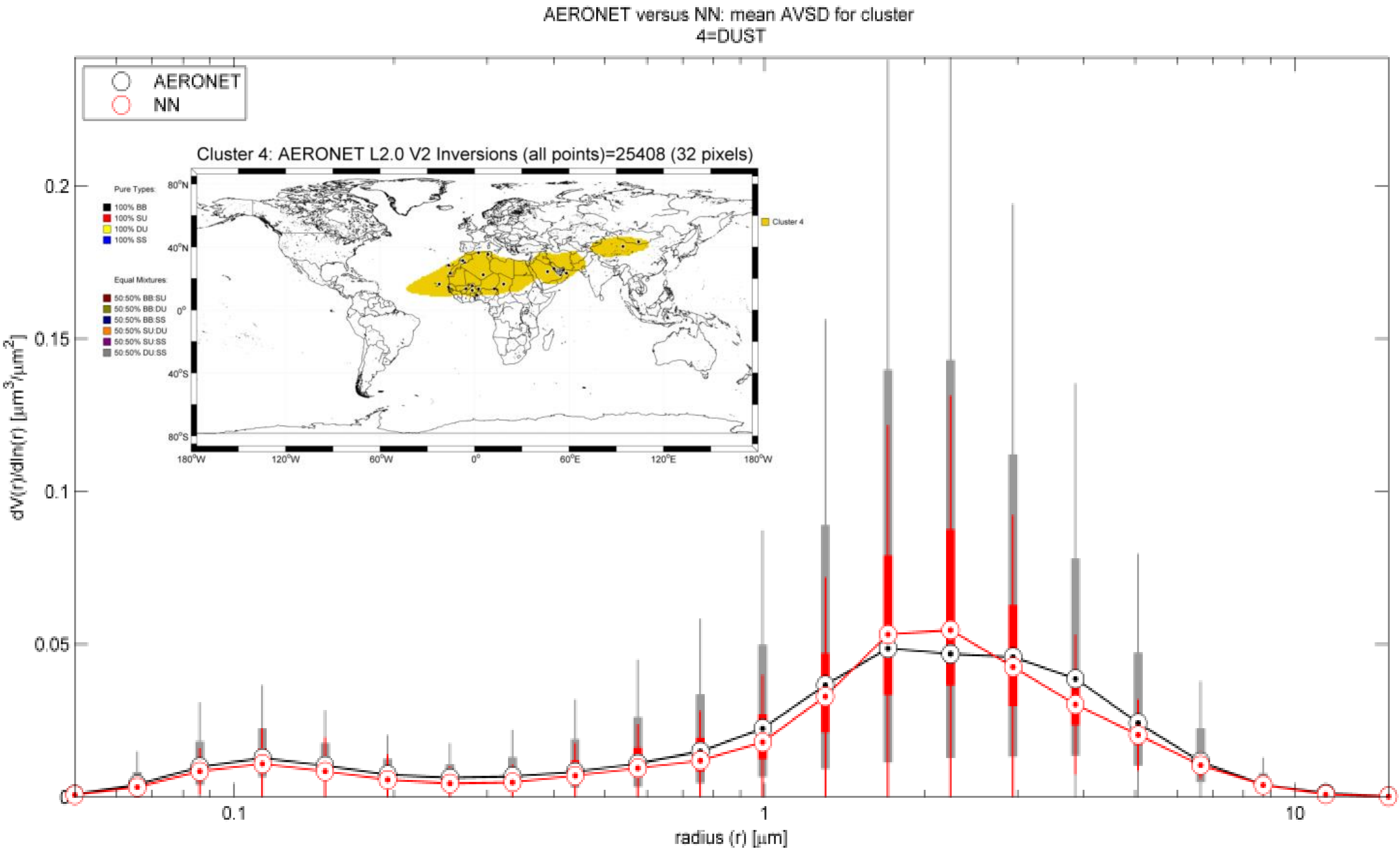
CLUSTER ANALYSIS: *Colour-mixed global mean aerosol partitions*



This colour scheme may help in visually interpreting aerosol mixtures

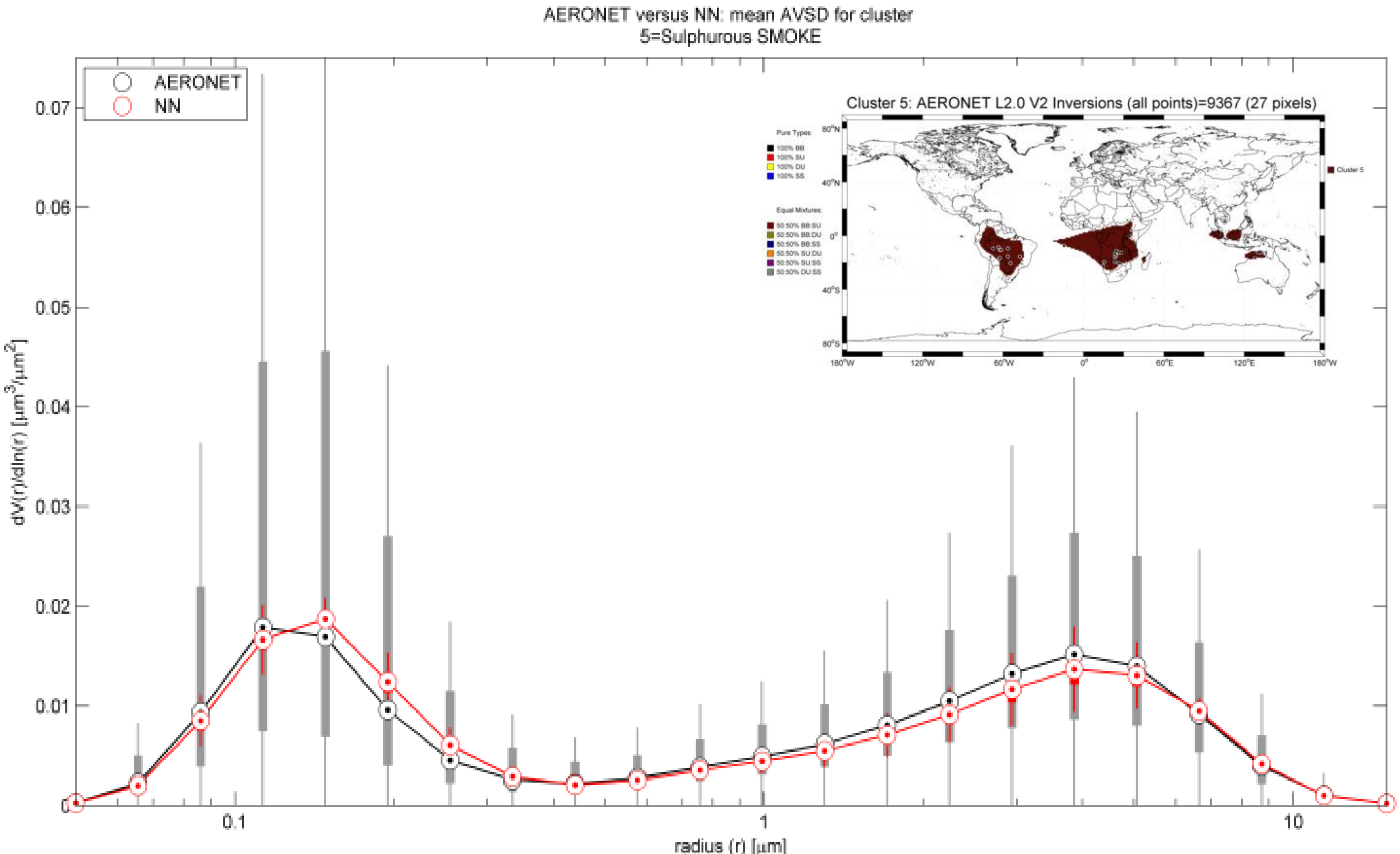
Michael Taylor: mtaylor@noa.gr

STEP 2: *Use co-located Satellite/AERONET data to train a NN for each cluster*



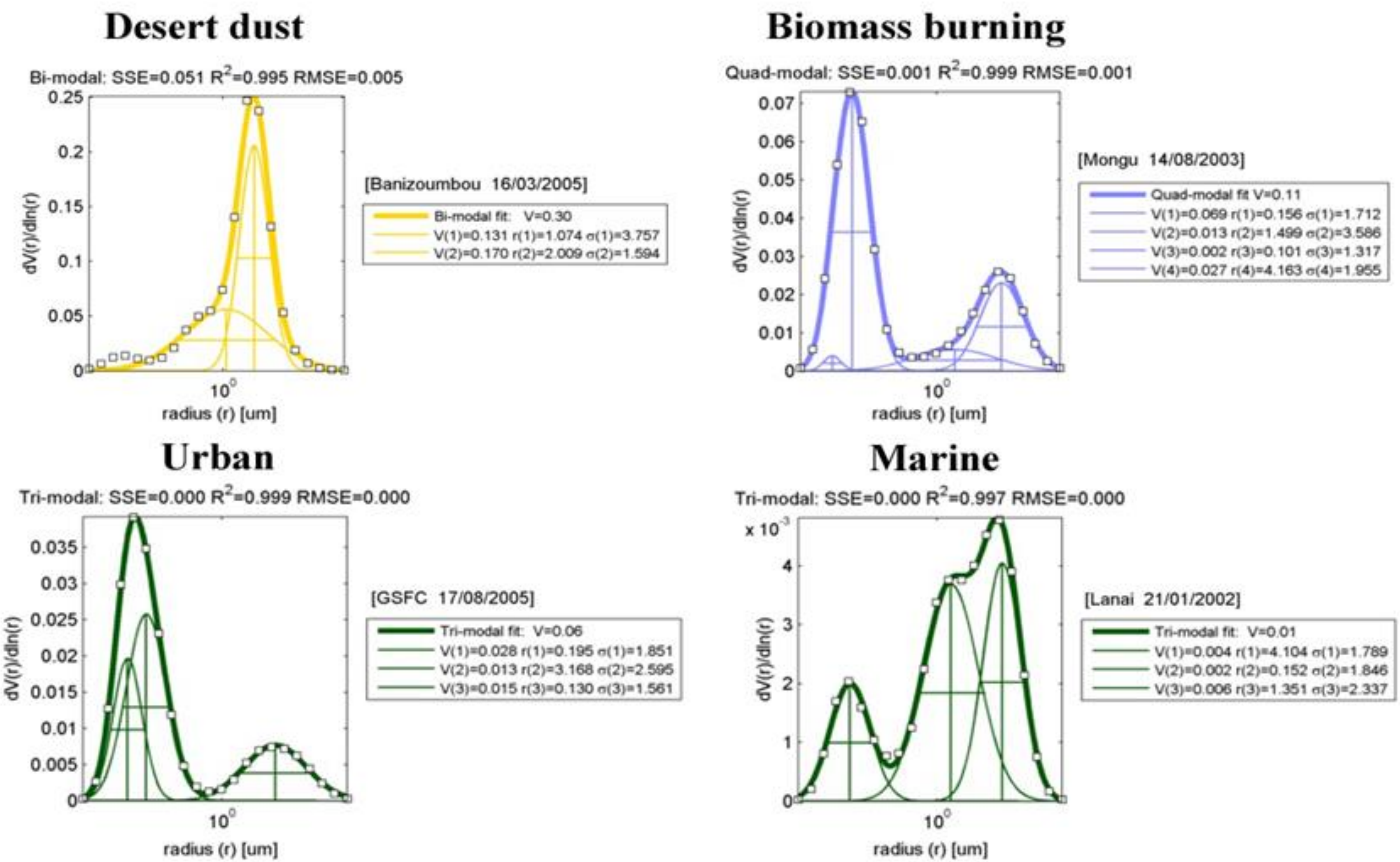
Michael Taylor: mtaylor@noa.gr

RESULTS: *The biomass burning cluster 5 ("Sulfurous SMOKE")*



Michael Taylor: mtaylor@noa.gr

STEP 3: *Multimodal fitting & analysis → additional size & volume info*



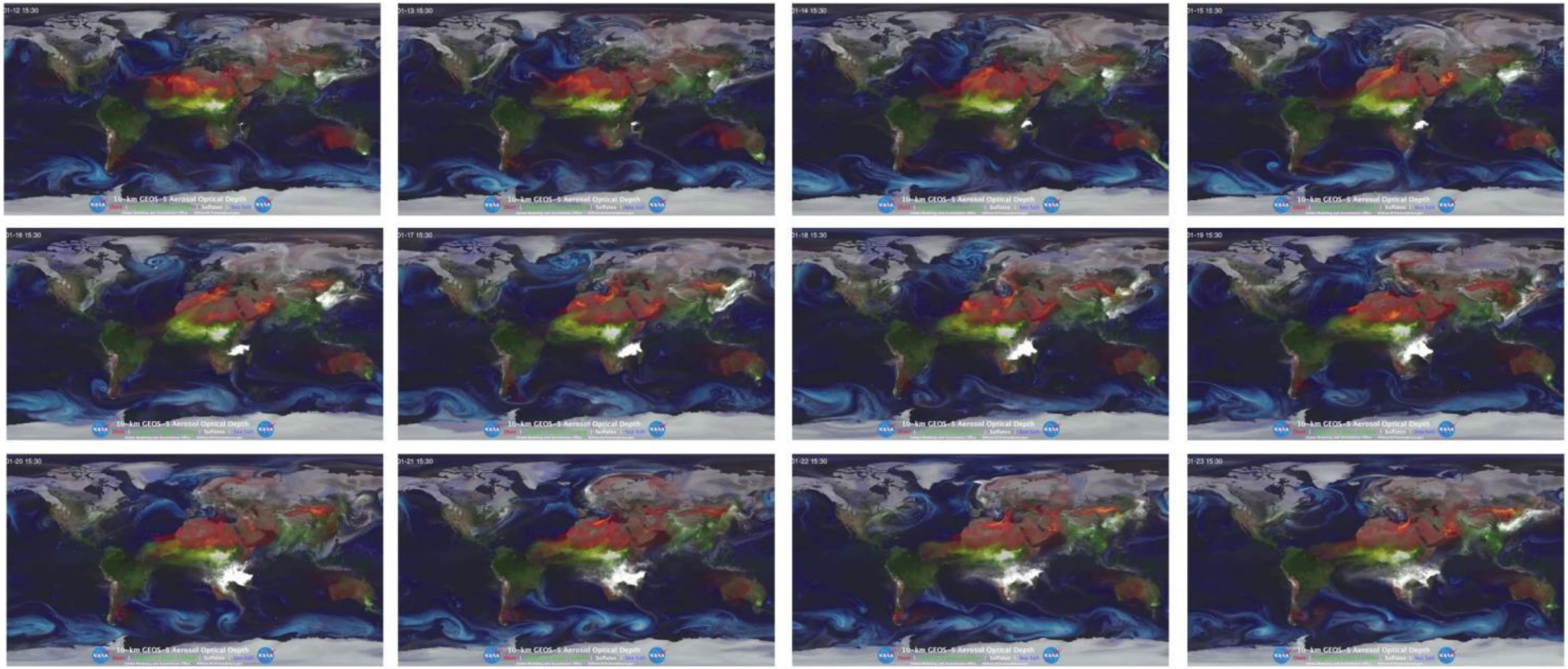
Source: Taylor, Kazadzis, Gerasopoulos (2014) *Multi-modal analysis of aerosol robotic network size distributions for remote sensing applications: dominant aerosol type cases*. **Atmospheric Measurement Techniques** 7, 839-858*AMT* 7, 839-858

Michael Taylor: mtaylor@noa.gr

CASE STUDY:

Quasi-real time monitoring of the Karthala Volcano Eruption
(12-23 January, 2007)

GOCART DATA: *Composite maps (BB, SU, DU and SS)*



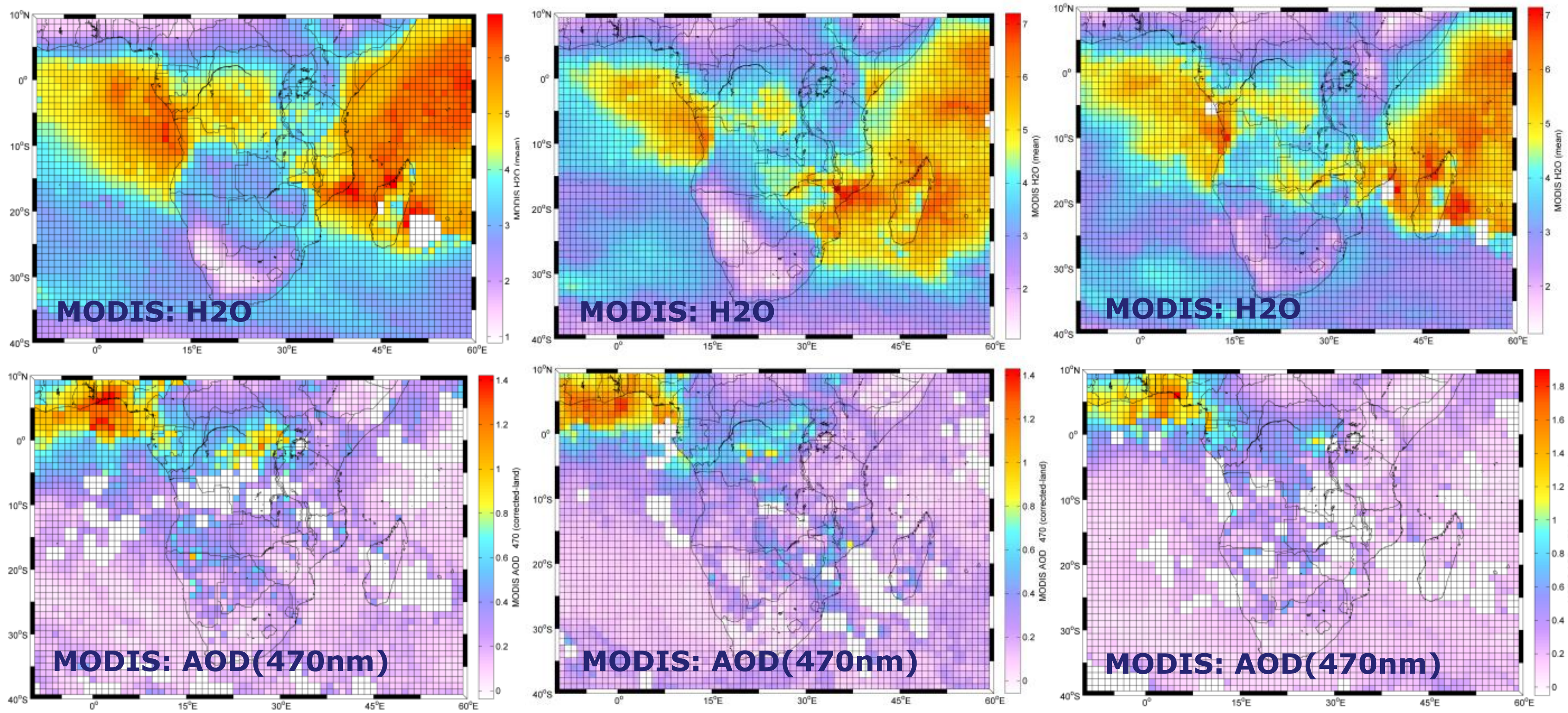
Karthala (Grande Comore Island)
Shield volcano eruption (VEI=2):
12-23 January, 2007

GOCART daily-averaged composite AOD
(BB=green, SU=white, DU=orange, SS=blue)

Michael Taylor: mtaylor@noa.gr

SATELLITE INPUTS: *4-day average (to smooth out "patchiness")*

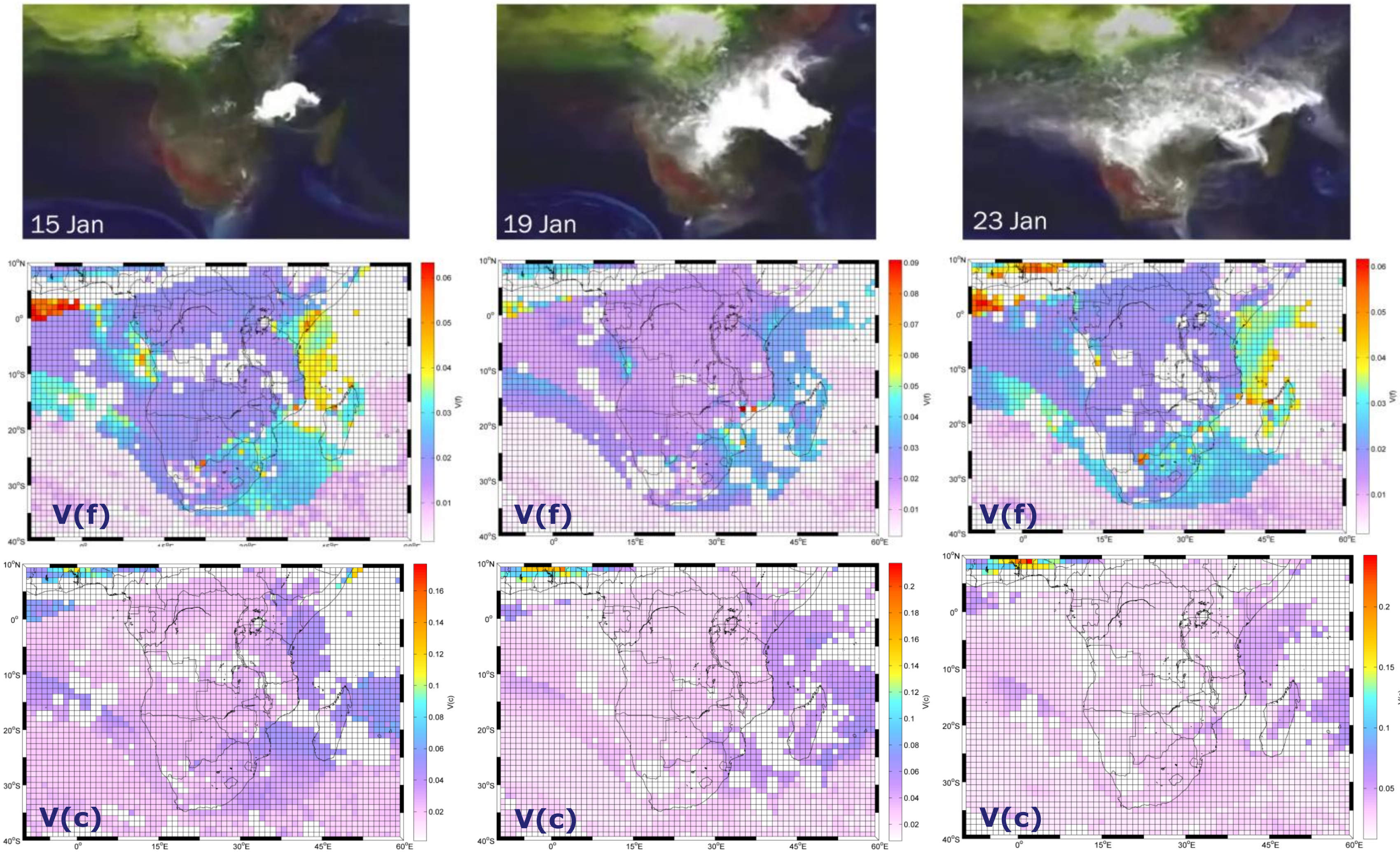
12-15 Jan 2007 | 16-19 Jan 2007 | 20-23 Jan 2007



Michael Taylor: mtaylor@noa.gr

4-DAY AVERAGE: *Quasi-realtime 1x1 degree grid of microphysics*

12-15 Jan 2007 | 16-19 Jan 2007 | 20-23 Jan 2007



Michael Taylor: mtaylor@noa.gr

SUMMARY:

Prototype model → it is possible to train NNs to invert satellite measurements to provide daily retrievals of the aerosol volume size distribution at 1x1 degree resolution

Global aerosol model → 1) Mean global GOCART AOD (per type) can be used to identify 10 clearly defined aerosol mixtures/partitions, 2) partitions allow extraction of co-located Satellite/AERONET data for training of aerosol mixture/region-dependent NNs, 3) Multimodal fitting & analysis of AVSDs can provide detailed size & volume information

Case study → quasi-realtime 1x1 degree maps of size distribution-derived microphysics can be produced at the 4-day timescale

Many thanks to:

@NOA: Stelios Kazadzis, Vangelis Gerasopoulos, Alexandra Tsekeri, Vasilis Amiridis and

@BSC/Earth Science: Antonis Gikas (for their co-authorship of articles)

@AERONET: PIs (for maintenance & provision of HQ open data inversion products)

@NASA/GES-DISC: Giovanni 3 & 4 PIs (for maintenance & provision of HQ open data from MODIS/Aqua, OMI/Aura and GOCART 4)

@ADNET SYSTEMS INC: Jennifer Brennan & Jim Acker for kindly organizing this workshop

Michael Taylor: mtaylor@noa.gr