Overview of the need and status of data assimilation and uncertainty estimates

Jeronimo Escribano (BSC), Enza Di Tomaso (BSC), Angela Benedetti (ECMWF)
with contribution from N. Schutgens, A. Da Silva, T. Tanaka and ICAP members

28/09/2019

AEROCOM / AEROSAT Barcelona 2019

Talk content

- Role of observations and their uncertainty in the solution of the data assimilation problem

- ICAP thoughts and operational practices

- A case study with a IASI dust product

The data assimilation problem

In the Kalman Filter, 4D-var, 3Dvar methods, the analysis is the solution of the minimization of the cost function J(x):

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x}))$$

x: state (or control) vector (model space; e.g., aerosol concentration)

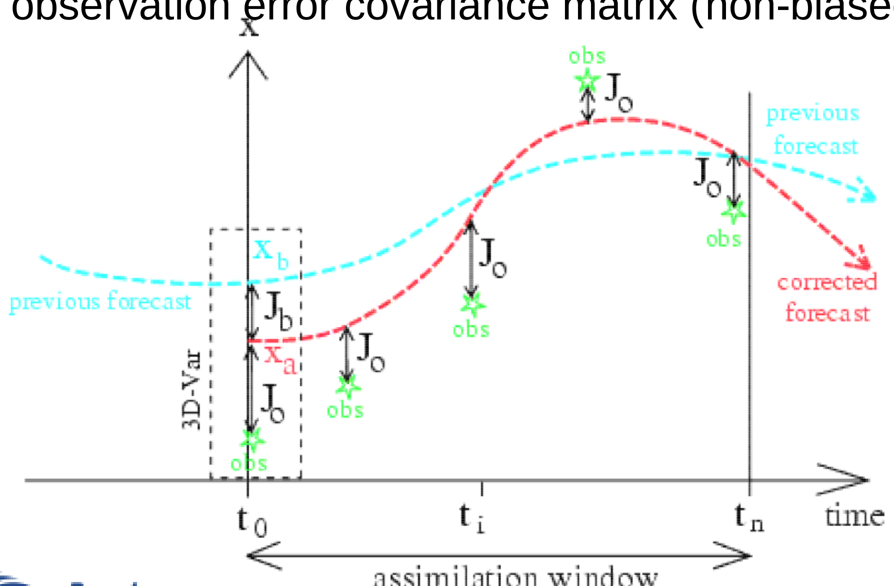
x^b : prior control vector (“first guess”)

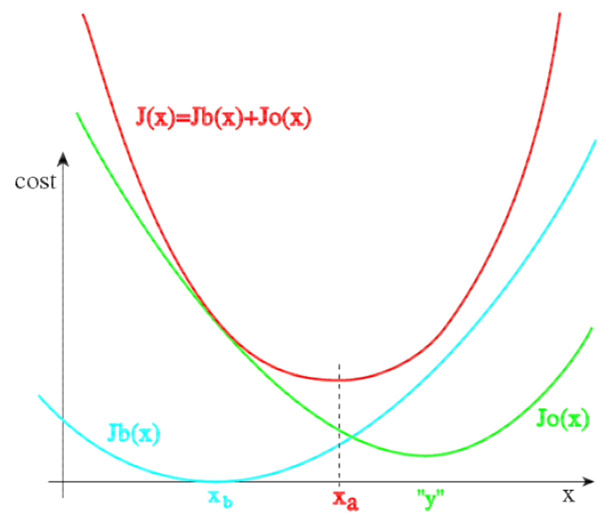
B: background or prior (forecast) error covariance matrix

y: observation vector (e.g., AOD 550nm)

H: observation operator (to calculate model-equivalents of the observations)

R: observation error covariance matrix (non-biased variable, random and gaussian assumptions)





Barcelona Supercomputing Center
Centro Nacional de Supercomputación

F. Bouttier and P. Courtier, ECMWF notes, 1999

Role of prior and observation uncertainty in data assimilation

Error statistics are used to represent mathematically observation and background uncertainty (both observations and background have errors).

From Niels Bormann's NWP training course, 2016:

- R and B together determine the weight of an observation in the assimilation.

- In the linear case, the minimum of the cost function can be found at x_a:

$$\underbrace{(\mathbf{x}_a - \mathbf{x}_b)}_{\text{Increment}} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}\underbrace{(\mathbf{y} - \mathbf{H}\mathbf{x}_b)}_{\text{Departure, innovation, "o-b"}}$$

- "Large" observation error → smaller increment, analysis draws less closely to the observations
- "Small" observation error → larger increment, analysis draws more closely to the observations

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

The different components of observational uncertaity

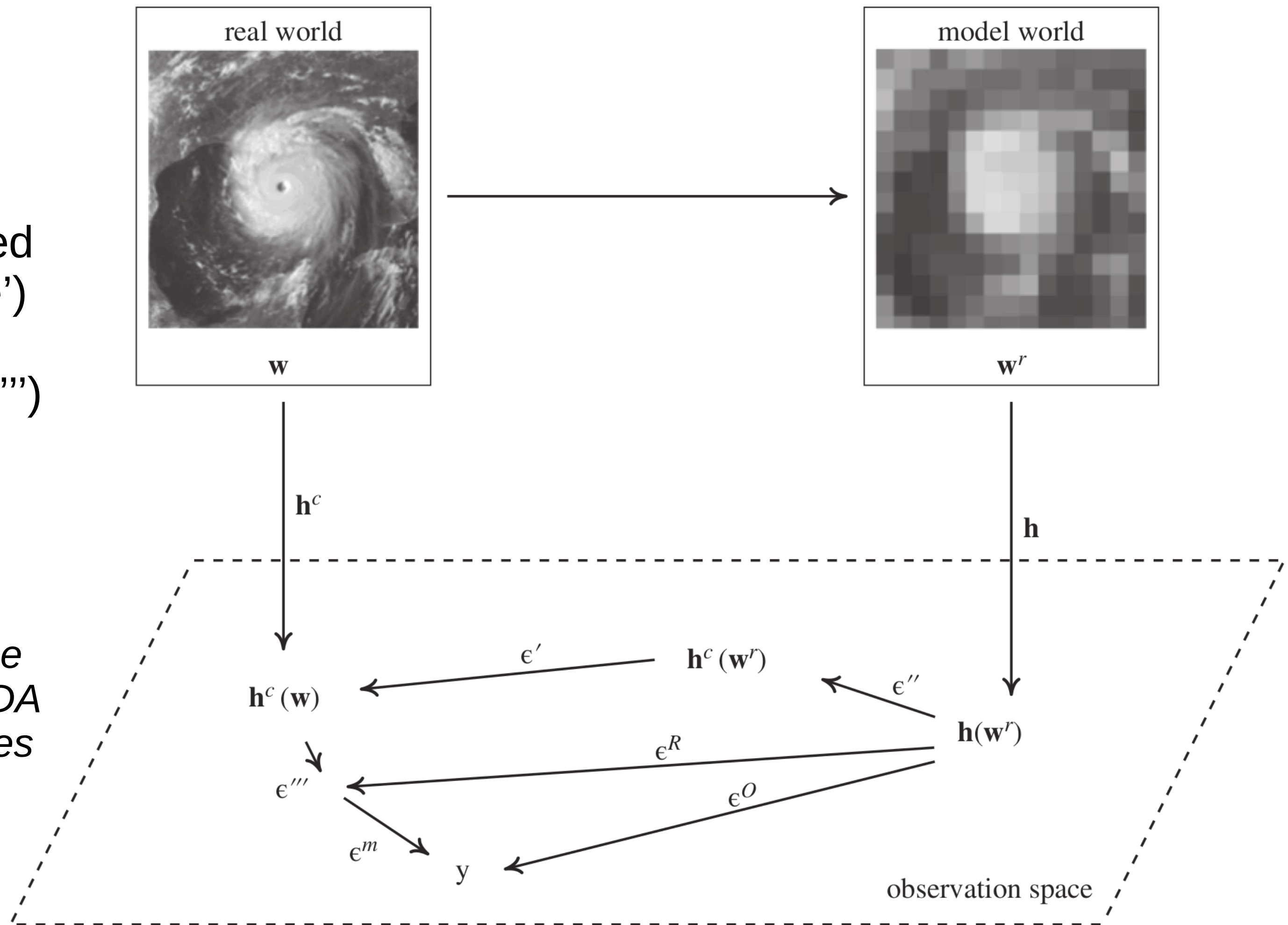
Observation error = measurement error (em) + representation error (er)

<u>Representation error</u> :

obs operator error (e'')

+ error due to unresolved scales and processes(e')

+ preprocessing error(e''')



While systematic or gross observation errors should be removed a-priori or during DA by bias correction techniques or quality control

Janjic et al., QJRMS, 2017

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Estimation of observation error statistics

Possible strategies to estimate observation error statistics for DA:

- Match-up with a reference observation / Error models
- Pixel-level uncertainties (physically-based) taking into account all possible source of error
- Fix values (can be situation-dependent, e.g., land/sea, satellite ...)
- Estimates based on DA innovations (observations minus background) statistics, with innovation statistics providing an upper bound

Some "tricks" commonly used:

- Neglect error correlations (diagonal covariance matrix) in space, time and between control vector variables
- Inflation of R elements, used to compensate for neglected errors (e.g. error correlations)
- Thinning by reducing obs density, used to avoid to have to estimate spatial error correlations (and to avoid over sampling)
- Globally constant covariance matrix
- Empirical adjustment of error estimates
- Balance between background and observation error matrices , eg., calculated through innovation statistics

$$E[\mathbf{d}_b^o\,(\mathbf{d}_b^o)^T] = \mathbf{H}\,\mathbf{B}\,\mathbf{H}^T + \mathbf{R}$$, where $\mathbf{d}_b^o = \mathbf{y} - H(\mathbf{x}^b)$

Thoughts from ICAP

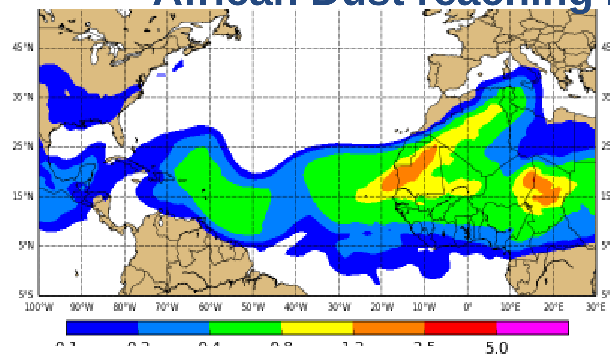
ICAP is the International Cooperative for Aerosol Prediction
(http://icap.atmos.und.edu/)

Participating members are: BSC, Copernicus/ECMWF, US Navy/FNMOC, NASA/GMAO, JMA, NCEP, UKMO, MeteoFrance, FMI

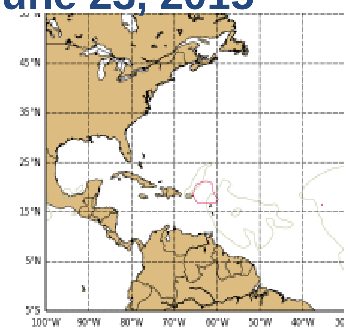
The ICAP participants provide global aerosol prediction operationally or quasi-operationally

They also coordinate the first global multi-model ensemble for aerosol forecasts (*Sessions et al 2015, ACP*)

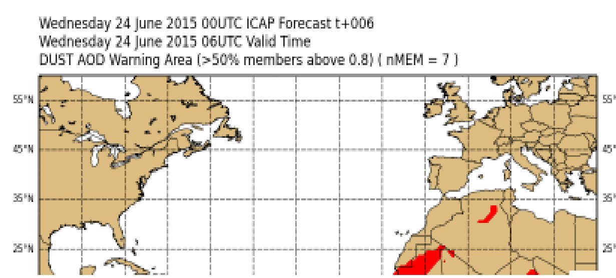
African Dust reaching DC, June 23, 2015



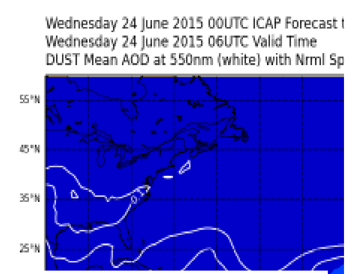
1. Ensemble Mean



2. AOT Contour (0.8)



3. Dust Warning Product



4. Normalized Ensemble Standard Deviation

Example of multi-model ensemble products for aerosol prediction

Thoughts from ICAP

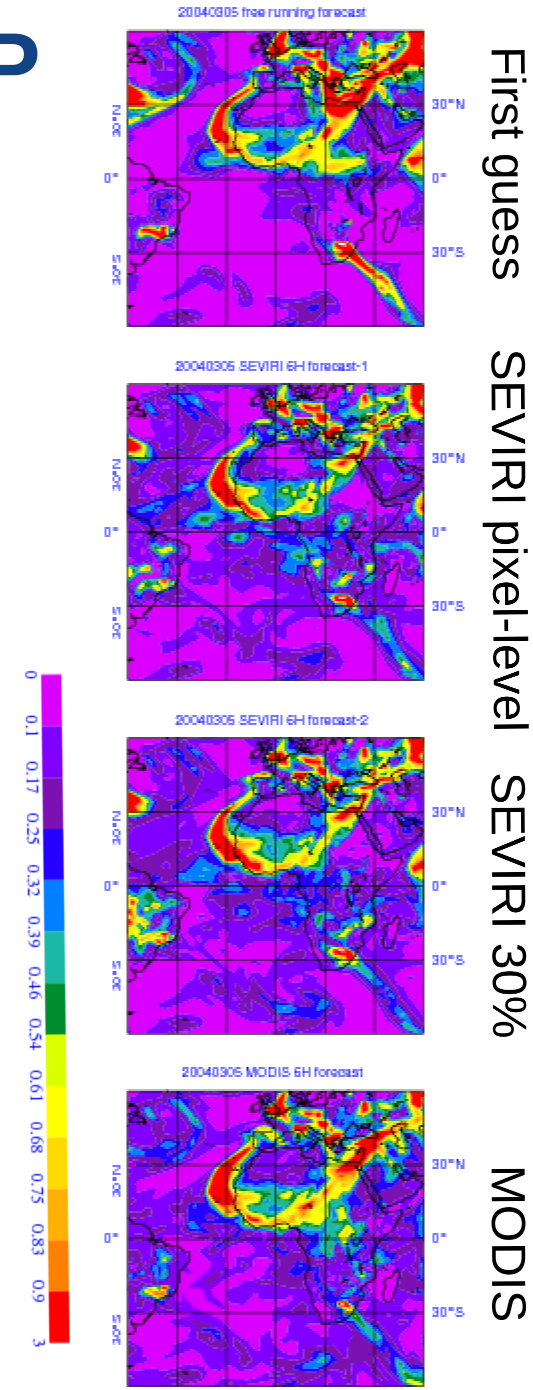
ECMWF/CAMS

- Fixed MODIS DT errors: 0.1 over land, 0.05 over ocean
- PMAp pixel-level uncertainties

GLOBAEROSOL project (ECMWF, 2009): showing that in the assimilation of SEVIRI AOD pixel-level uncertainties are beneficial (reported errors are larger for large and spurious AODs)

  - Using the retrieval errors has been found to provide more desirable first-guess and analysis departure distributions, with a smaller bias and a more Gaussian shape compared to using errors fixed at 30%

  - Although this latter choice allows assimilate twice as many data (observation errors are used in the quality control process, which means that the lower the errors, the more stringent the quality control is)



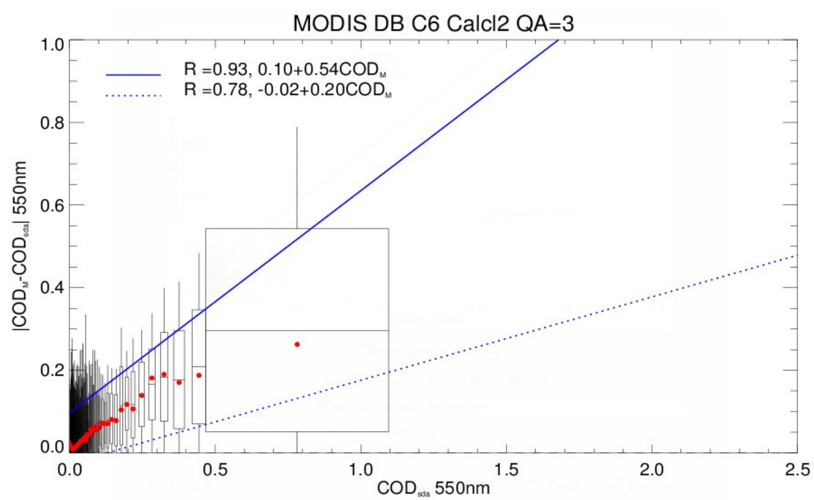
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Thoughts from ICAP

BSC

- NRL MODIS DT: pixel-level uncertainty for L3 (instrumental error variance and spatial representation error variance estimated by NRL); MODIS DB: error linear in AOD (based on *Sayer at al. 2014*) plus error representation component for L3. (*Di Tomaso et al., 2017*)

- Dust regional reanalysis: linear relation model for the observation error calculated using reference observations (regional AERONET SDA), plus balanced error based on innovation statistics.



(Paul Ginoux, GFDL)

A linear model of the error tends to give more weight to small AODs → it creates bias in the analysis (BSC + ECMWF's experience). Pragmatic solutions : fix error, independent of AOD, empirical adjustment of the linear coefficients

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Thoughts from ICAP

JMA

- Himawari-8 / AHI AOD (not operational) : pixel-level uncertainty from their retrieval and representativeness error estimated through the process of making super observation for the model grid

- JAXA aerosol product ver 2: observational uncertainty estimated by taking standard deviation from a match-up with MODIS AOD (provided observational uncertainty is too small).

NASA GSFC

- Observation error estimated from innovation statistics using a maximum-likelihood algorithm.

- DA obs error are dominated by forward model errors (due to uncertainties in optical properties) and representativeness (hence pixel-level uncertainties have to be inflated)

Case study: 25 April to 25 May 2017

First try!

Model:

NMMB - MONARCH multi-scale chemical weather prediction system model (Janjic et al., 2011, Perez et al., 2011)

- Aerosols: only dust configuration
- 0.66 degrees resolution, 40 vertical levels
- Dust emission schemes available

DA experiments:

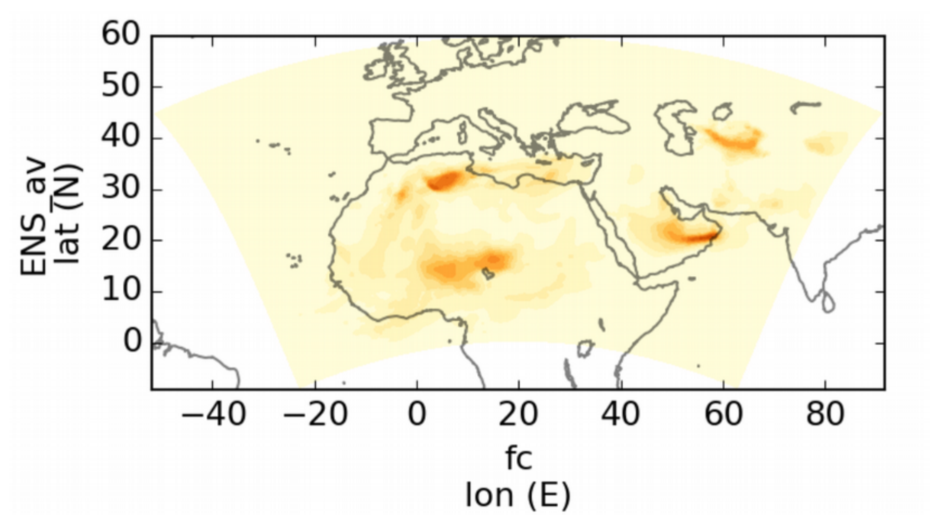
- Fix uncertainty
- Linear uncertainty
- Pixel-level unc + model uncertainty

Observations:

IASI AOD 10 um from ULB

(Clarisse et al., 2019)

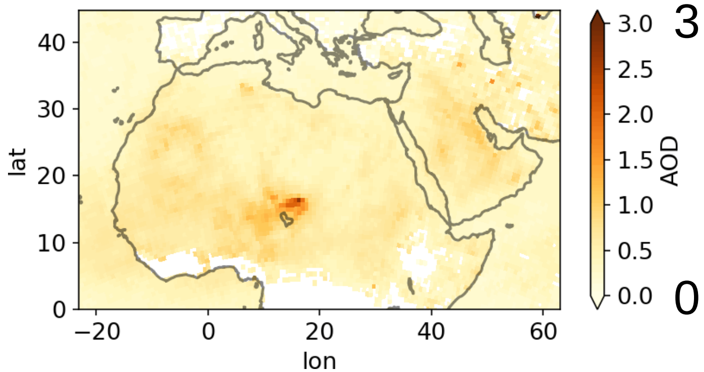
- MetopA (Ascending and Descending)
- Pixel-level uncertainties provided



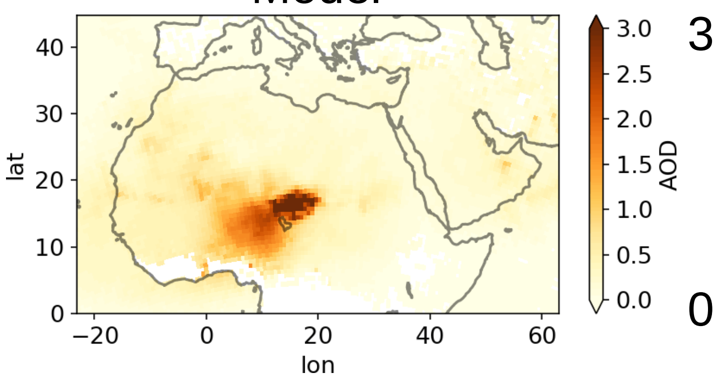
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Simulation without assimilation

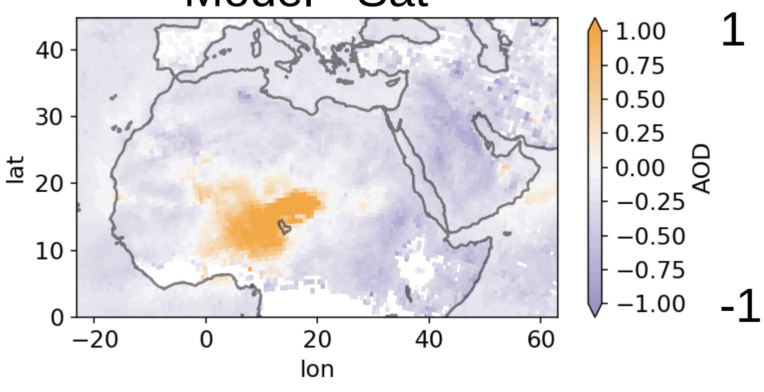
VIIRS-SNPP – DB / dust 550 nm



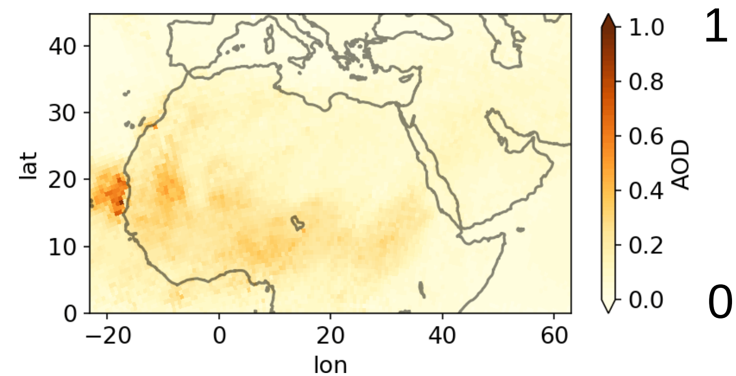
Model



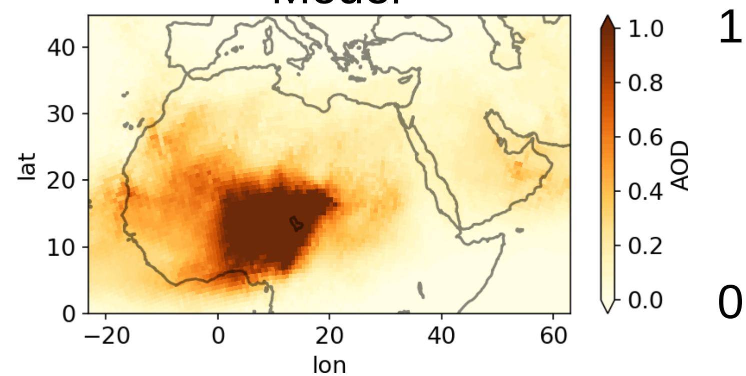
Model - Sat



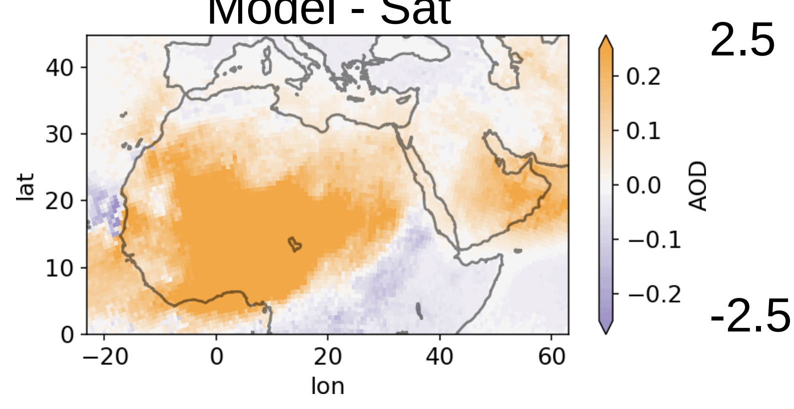
IASI – ULB 10 um



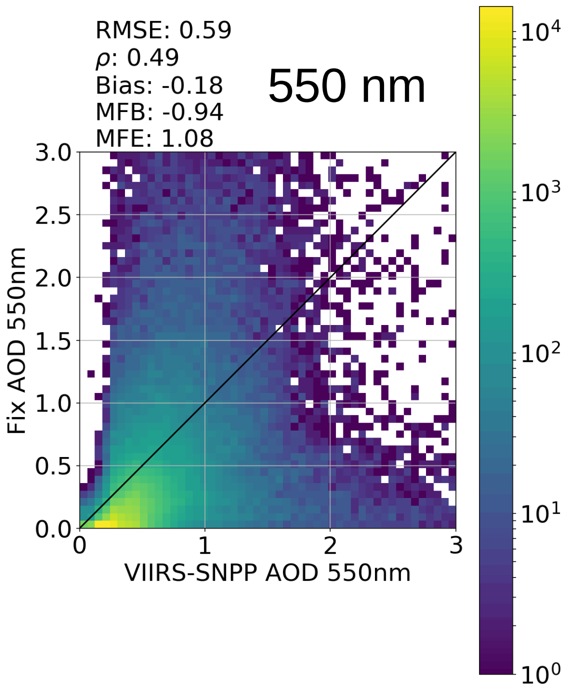
Model

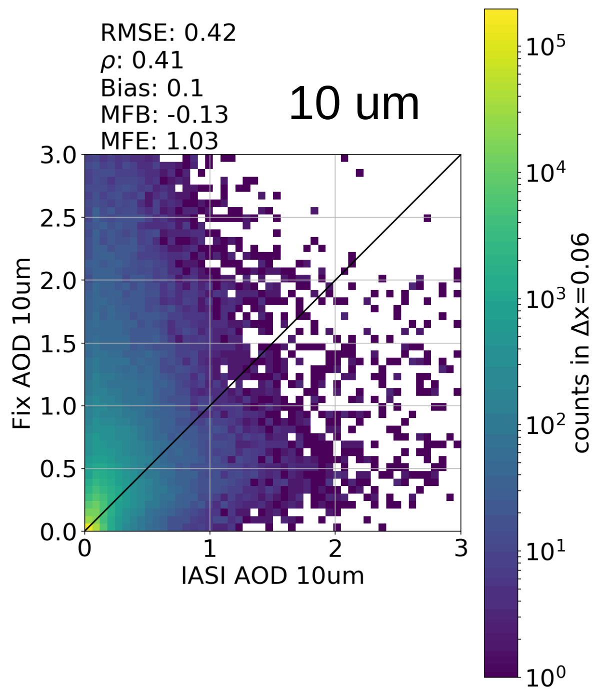


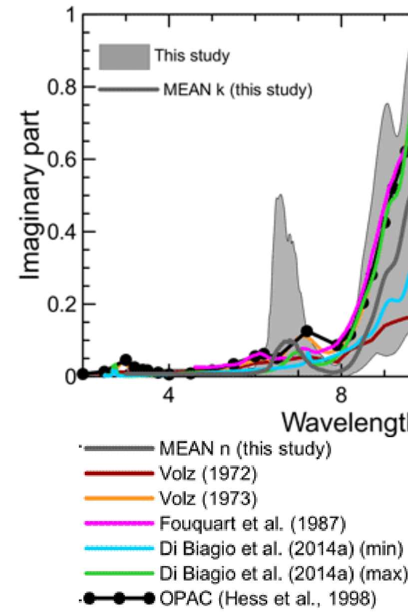
Model - Sat



(likely model bias in optical properties)







Di Biagio et al., ACP 2017

[DELETED FIGURE]

Gonçalves-Ageitos et al.(in preparation)

Mean bias (1 day forecast)

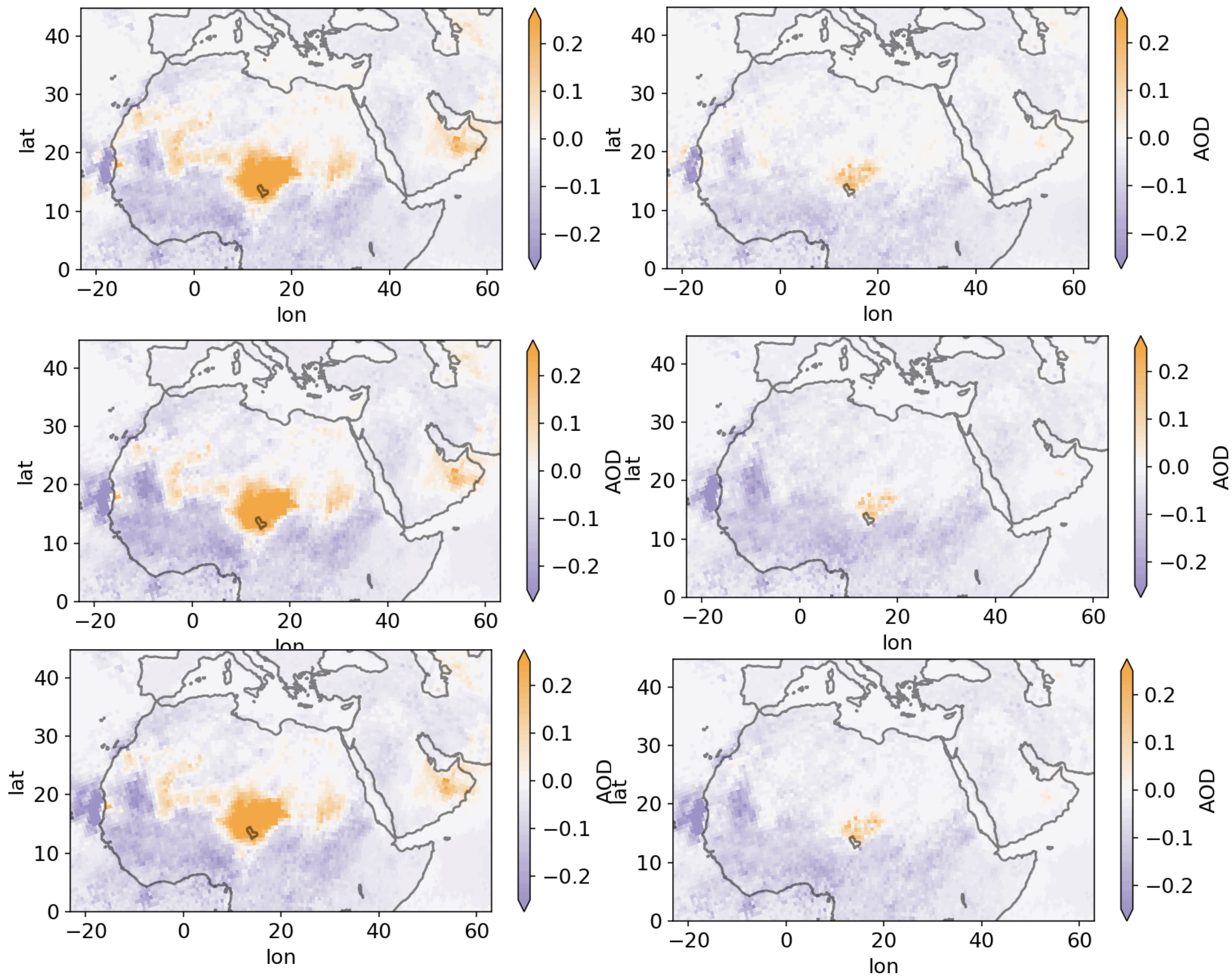
Forecast – IASI 10um

Analysis - IASI 10um

Fix error:
$e^o = 0.1$

Linear error:
$e^o = 0.015 + 0.1*AOD$

Pixel-level uncertainties and model error:
$(e^o)^2 = (e^m)^2 + 0.01^2$

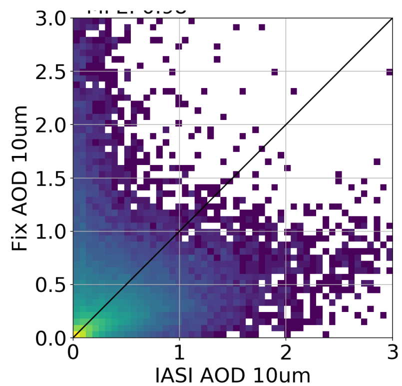


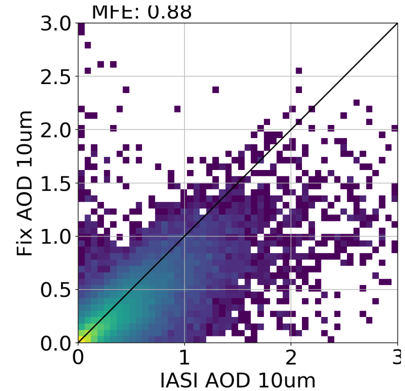
Mean bias (1 day forecast)

Forecast – IASI 10um | Analysis - IASI 10um

Fix error:
$e^o = 0.1$

RMSE: 0.22
ρ: 0.35
Bias: -0.02
MFB: -0.66
MFE: 0.98

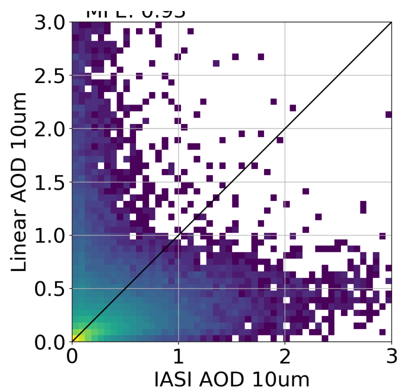


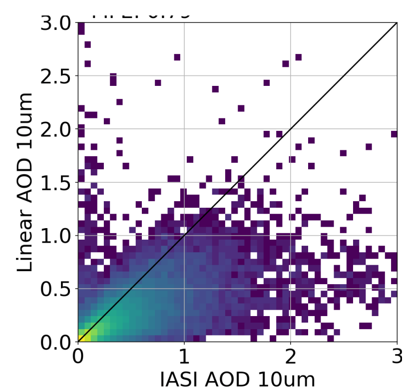


RMSE: 0.11
ρ: 0.77
Bias: -0.03
MFB: -0.64
MFE: 0.88

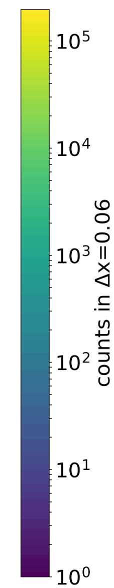
Linear error:
$e^o = 0.015 + 0.1 \cdot AOD$

RMSE: 0.23
ρ: 0.28
Bias: -0.03
MFB: -0.63
MFE: 0.93



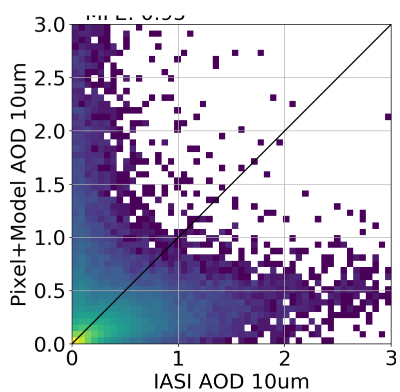


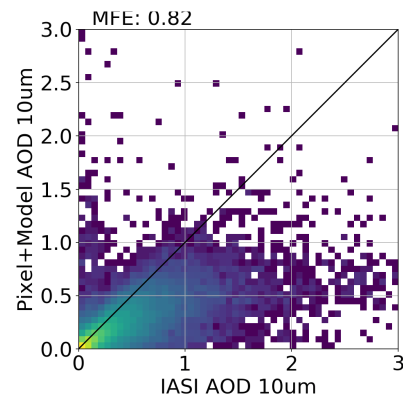
RMSE: 0.12
ρ: 0.71
Bias: -0.04
MFB: -0.6
MFE: 0.79



Pixel-level uncertainties and model error:
$(e^o)^2 = (e^m)^2 + 0.01^2$

RMSE: 0.23
ρ: 0.29
Bias: -0.03
MFB: -0.66
MFE: 0.95





RMSE: 0.12
ρ: 0.71
Bias: -0.04
MFB: -0.63
MFE: 0.82

Mean bias (1 day forecast) : 550 nm

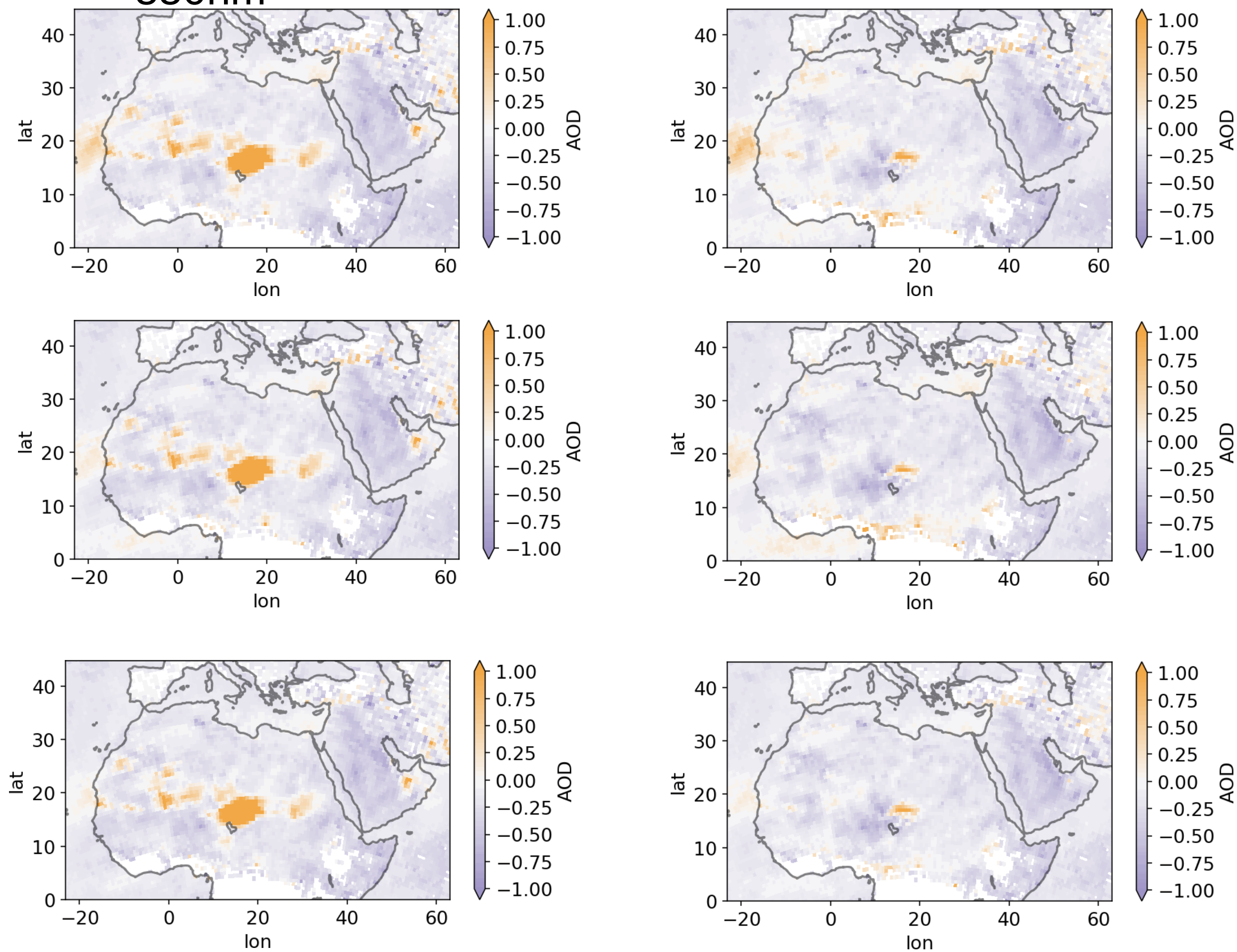
Forecast – VIIRS-DB 550nm

Analysis – VIIRS-DB 550nm

Fix error (at 10um):
$e^o = 0.1$

Linear error (at 10 um):
$e^o = 0.015 + 0.1*AOD$

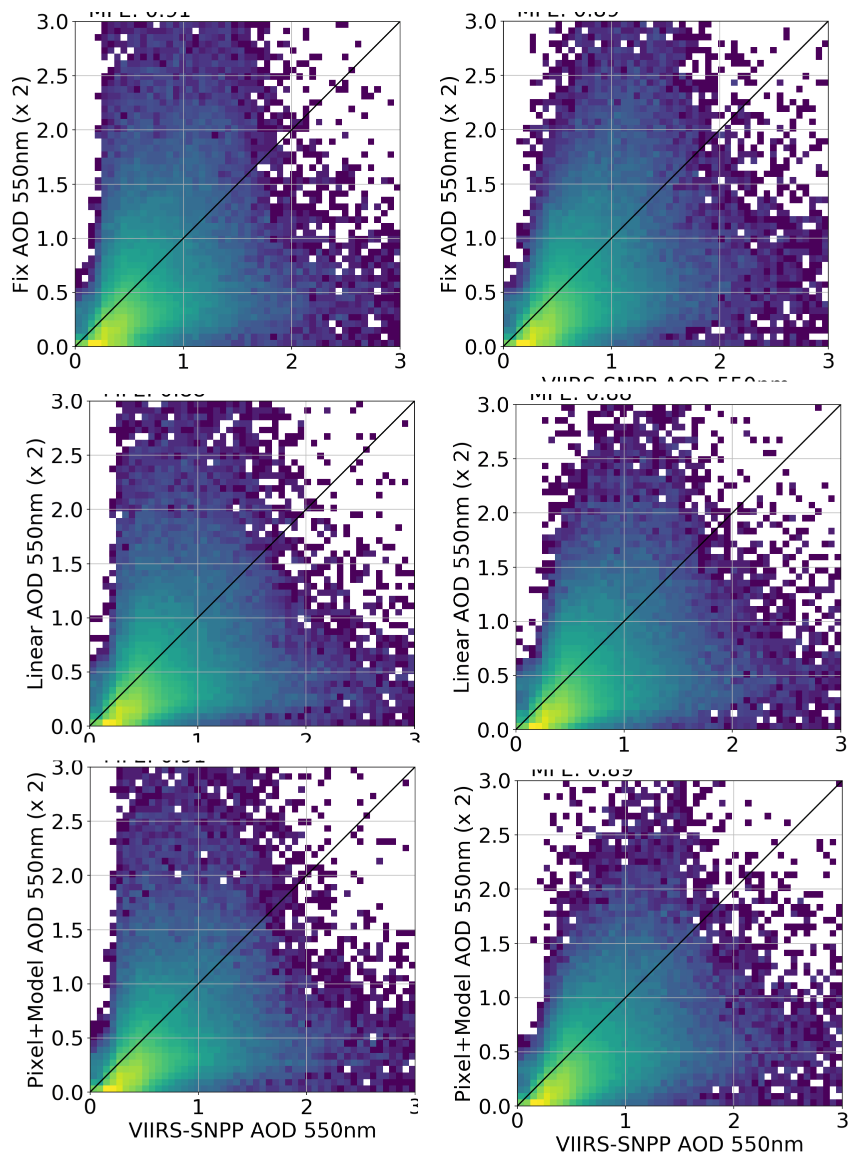
Pixel-level uncertainties and model error (at 10um):
$(e^o)^2 = (e^m)^2 + 0.01^2$



Mean bias (1 day forecast): 550 nm

Forecast – VIIRS-DB 550nm

Analysis – VIIRS-DB 550nm



Fix error:
$e^o = 0.1$

RMSE: 0.86
ρ: 0.41
Bias: -0.11
MFB: -0.74
MFE: 0.91

RMSE: 0.48
ρ: 0.53
Bias: -0.16
MFB: -0.76
MFE: 0.89

Linear error:
$e^o = 0.015 + 0.1*AOD$

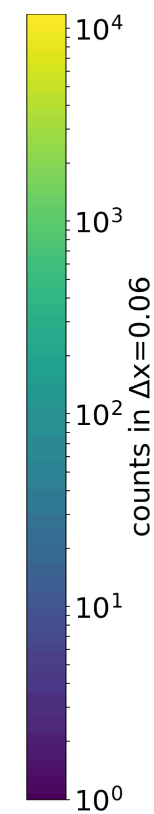
RMSE: 0.85
ρ: 0.38
Bias: -0.13
MFB: -0.73
MFE: 0.88

RMSE: 0.47
ρ: 0.48
Bias: -0.19
MFB: -0.76
MFE: 0.88

Pixel-level uncertainties and model error:
$(e^o)^2 = (e^m)^2 + 0.01^2$

RMSE: 0.85
ρ: 0.39
Bias: -0.14
MFB: -0.78
MFE: 0.91

RMSE: 0.46
ρ: 0.51
Bias: -0.19
MFB: -0.8
MFE: 0.89



Summary

Estimation of observation uncertainty in (aerosol) data assimilation:

- Measurement error is one of the ingredients of the observational error estimates used in data assimilation
- No correlation between errors in aerosol DA (unlike meteorological DA). Are they estimated?
- DA observational error have to be adjusted in each DA system (model, measurement and representativity errors) and a balance is needed ("diagnostics")

- Pixel uncertainties helps modellers to identify model errors in the DA context

Dust IASI data assimilation case study:

- Clear inconsistency in model optics: to be updated
- No large difference in experiments for IASI AOD: Error dominated by the model bias (overestimation of Bodélé emissions and underestimation elsewhere, more work is needed)
- More detailed analysis will follow: departure statistics, spread changes, obs influence, etc.
- Qualitative differences in the "corrected" AOD at 550

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Thank you!

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement H2020-MSCA-COFUND-2016-754433

STARS POST-DOCTORAL PROGRAMME

PRACE · GOBIERNO DE ESPAÑA · MINISTERIO PARA LA TRANSICIÓN ECOLÓGICA · AEMet Agencia Estatal de Meteorología · DustClim · ERA4CS European Research Area for Climate Services · esa

28/09/2019

AEROCOM / AEROSAT 2019

jeronimo.escribano@bsc.es