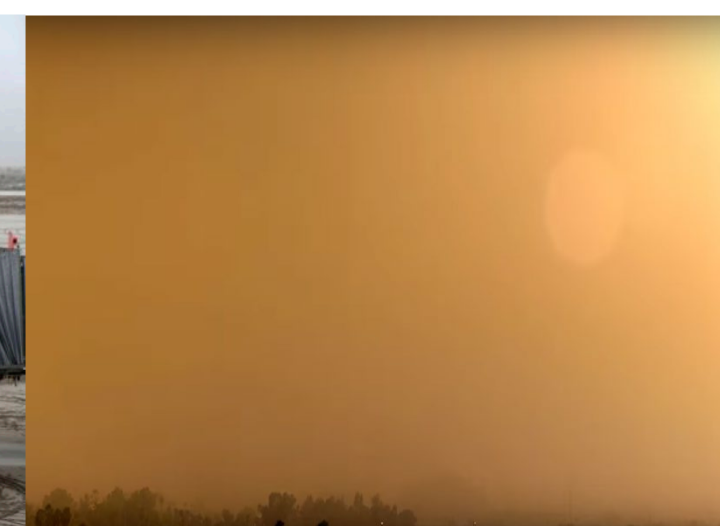
Data Assimilation and Machine Learning for Air Quality Forecasts - Emission Inversion

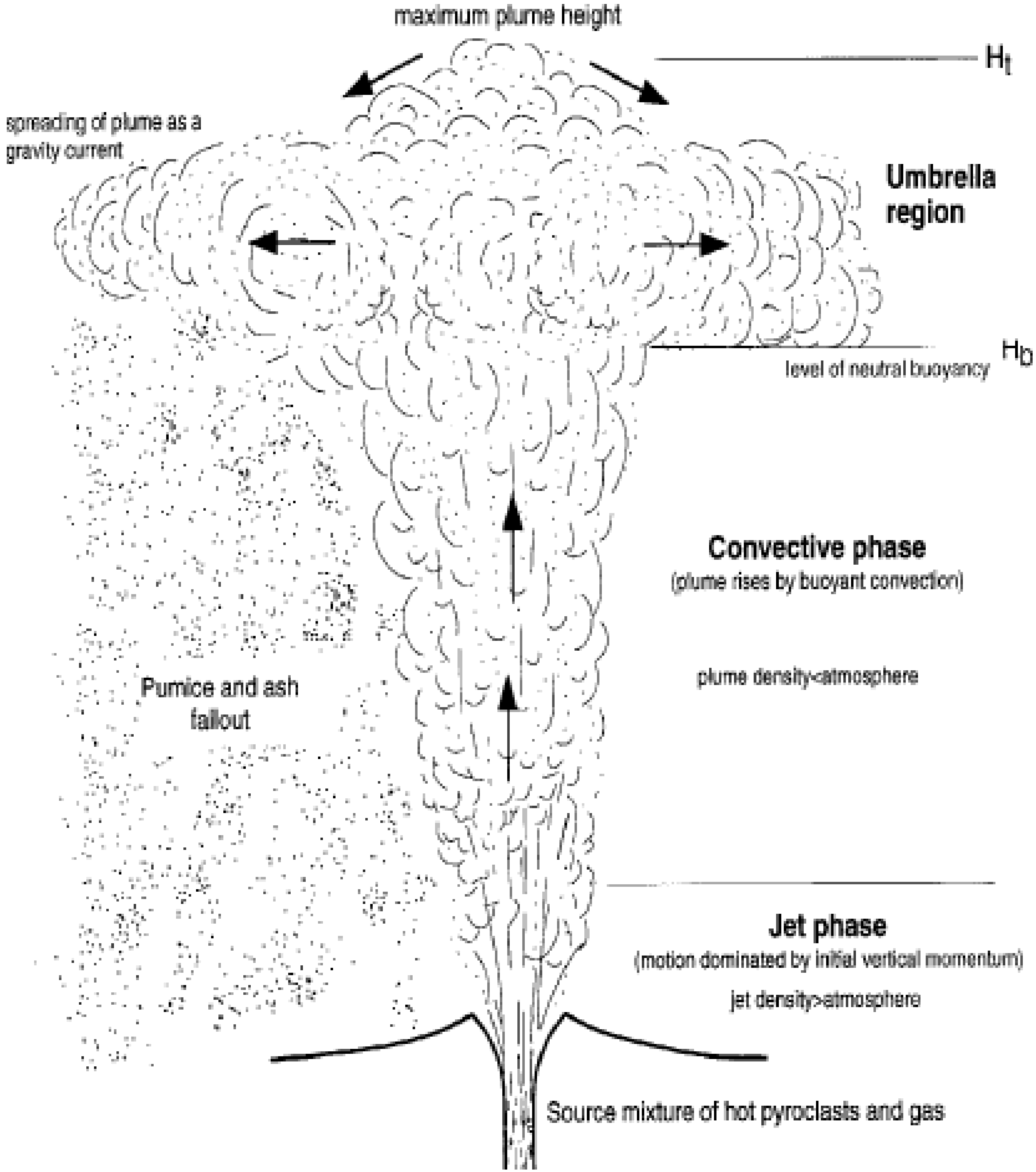
Hai Xiang Lin
Delft University of Technology &
Leiden University

In cooperation with
G. Fu, A.W. Heemink, J. Jin, S. Lu (TU Delft), A.J. Segers (TNO),
T. Palsson (Iceland), K. Weber (Germany), A.J. Prata (Norway)

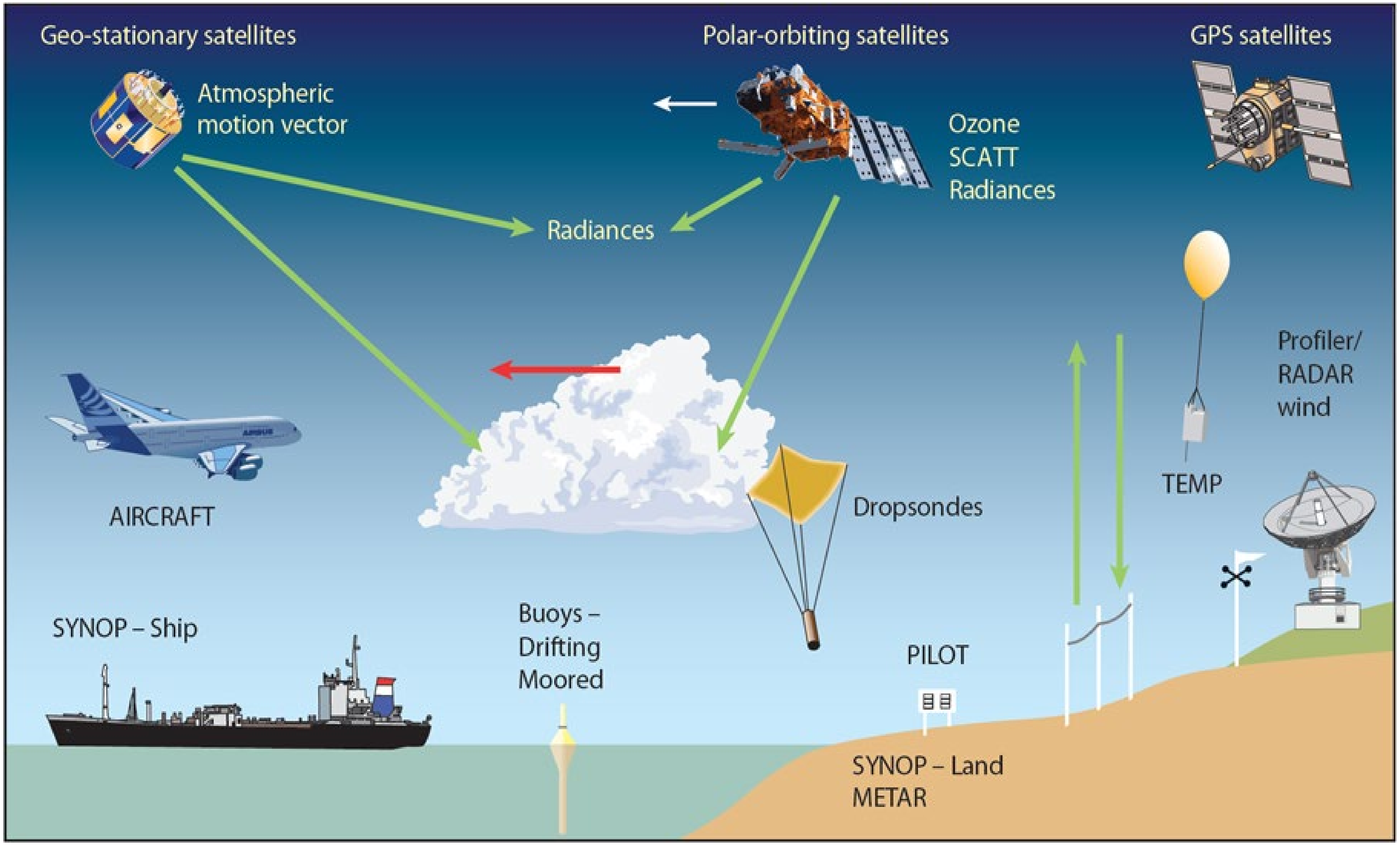


TUDelft

Hazard of volcanic ash:
An accurate forecast is important!

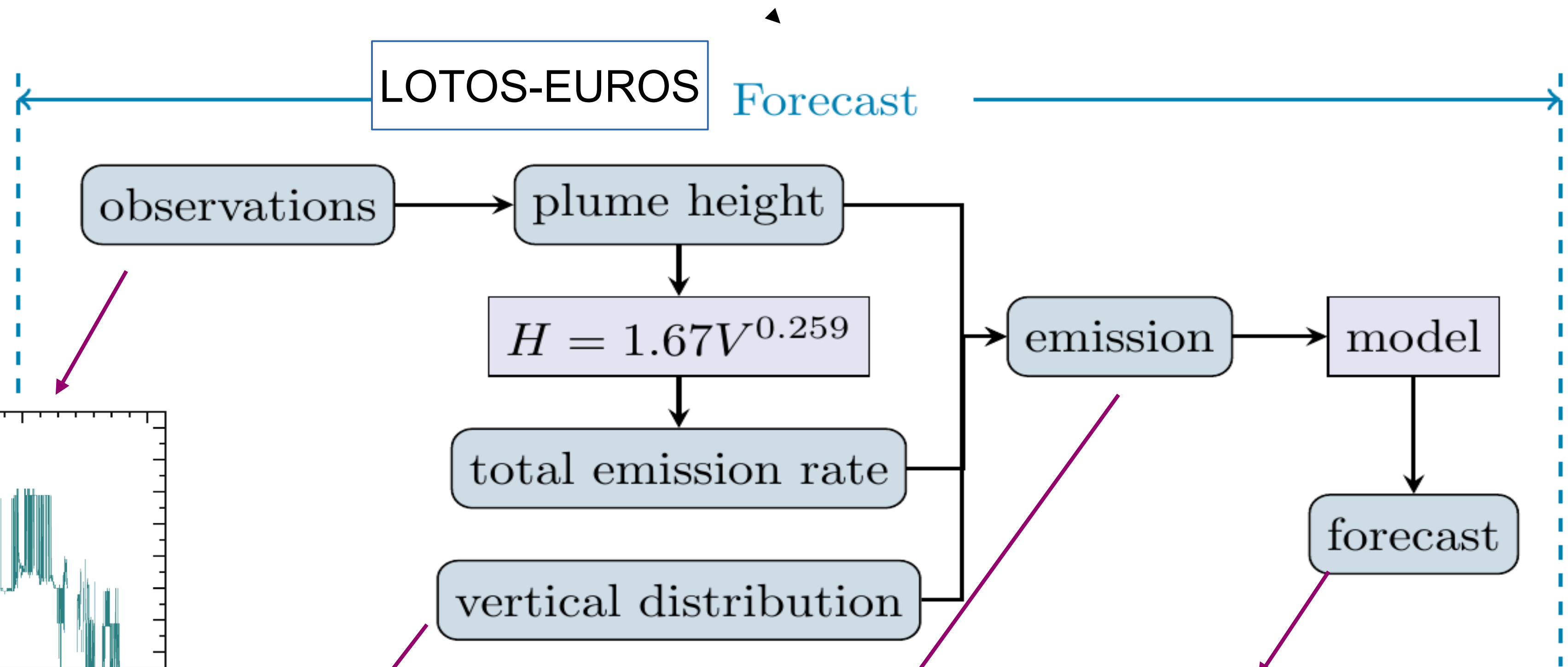

maximum plume height
H_t
spreading of plume as a gravity current
Umbrella region
H_b
level of neutral buoyancy
Convective phase
(plume rises by buoyant convection)
plume density<atmosphere
Pumice and ash fallout
Jet phase
(motion dominated by initial vertical momentum)
jet density>atmosphere
Source mixture of hot pyroclasts and gas

Observations of volcanic ash activity



TU Delft

Emission observation: plume height



Ground-based observations of plume height: Lidar, radar

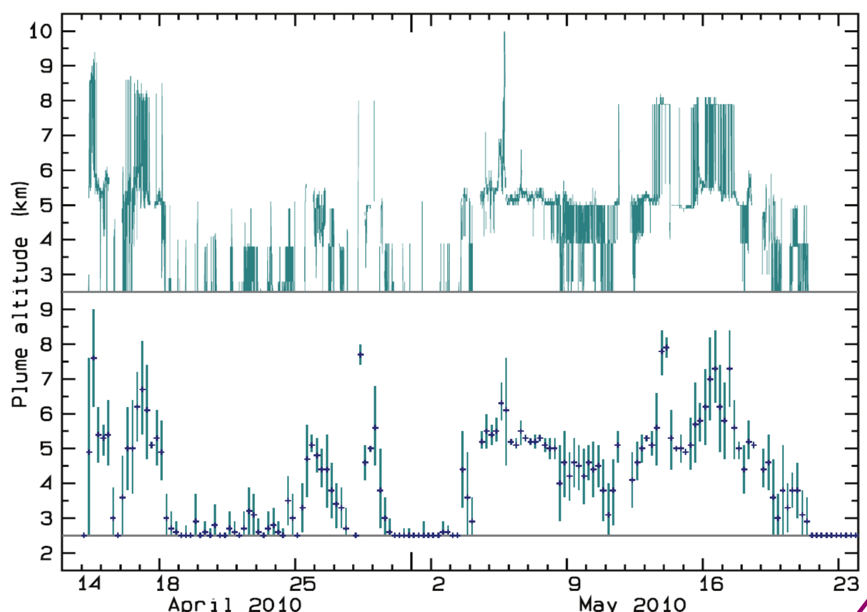


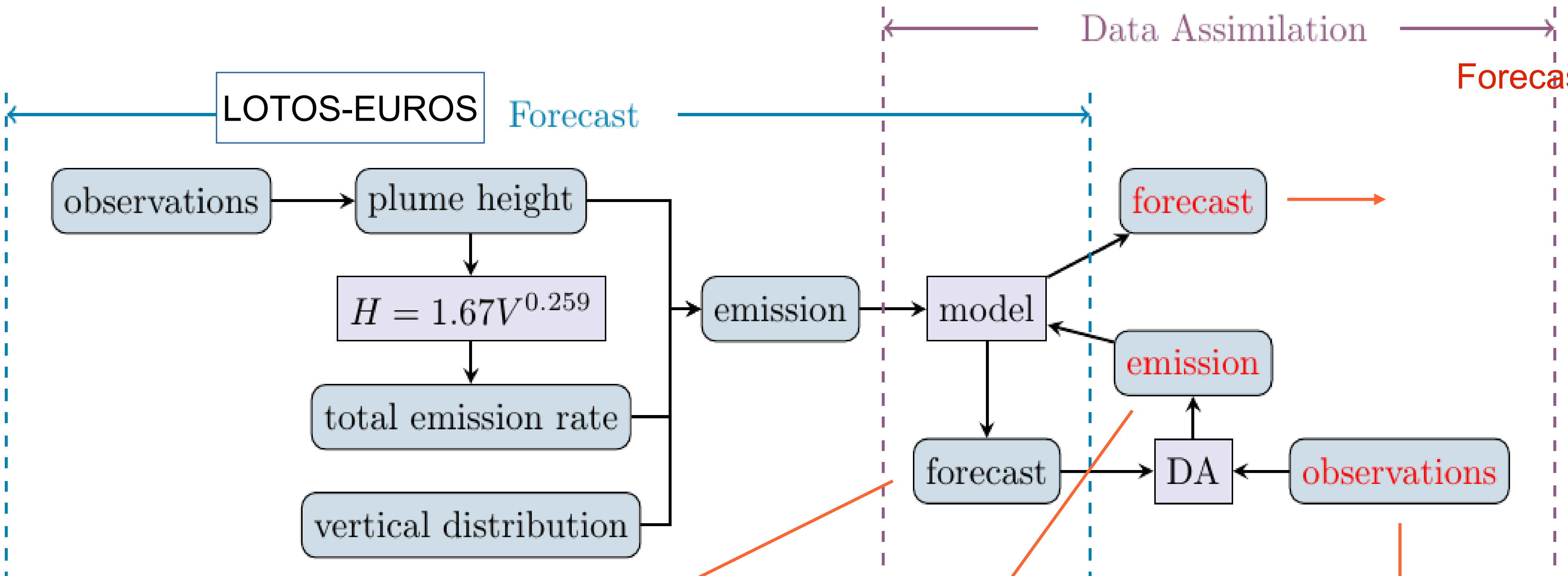
Illustration of the vertical profile of emission

Ash columns of forecast ash plume

Predefined vertical distribution:
Umbrella-shaped,
Exponentially-distributed,
Poisson-distributed

TUDelft

Forecast of volcanic ash cloud

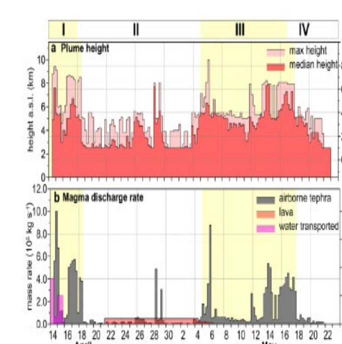


Forecast after DA

Forecast before DA

Illustration of emission after DA

Ash mass loading retrieved from satellite data;
Plume height from ground-based observations



TU Delft

Data Assimilation

Data assimilation combines information of observations and models and their errors to get a best estimate of atmospheric state (or other parameters)

N(0,B_k)

$$X_k^f = M_k(X_{k-1}^a, u) + w_{k-1}$$

$$y_k = H_k(X_k) + v_k$$

N(0,R_k)

The prior and likelihood are

$$f(X) \propto \exp(-\frac{1}{2}(X - X^f)^T B_k^{-1}(X - X^f)),$$ and

$$f(y \mid X) \propto \exp(-\frac{1}{2}(X - y)^T R_k^{-1}(X - y))$$

Posterior: $f(X \mid y) \propto f(X) f(y \mid X) \propto \exp(-\frac{1}{2} J(X))$

Minimize the cost function J:

$$J(X^a) = (X^a - X^f)^T B_k^{-1}(X^a - X^f) + (X^a - y)^T R_k^{-1}(X^a - y)$$

TU Delft

Chemical Transport Model – LOTUS-EUROS

$$\frac{\partial C}{\partial t} + U\frac{\partial C}{\partial x} + V\frac{\partial C}{\partial y} + W\frac{\partial C}{\partial z} = \frac{\partial}{\partial x}\left(K_h\frac{\partial C}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_h\frac{\partial C}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_z\frac{\partial C}{\partial z}\right)$$
$$+E+R+Q-D-H$$

Emissions

Transport
(advection, diffusion, radiation**)**

Model

sedimentation
(wet/dry deposition)

Chemical reactions

TU Delft

Using 4D-Var to estimate the emissions

Model representation:

$$\begin{aligned}\mathbf{x}_k &= M_k(\mathbf{x}_{k-1}, \mathbf{u}_k + \mathbf{w}_k),\\ \mathbf{y}_k &= H_k(\mathbf{x}_k) + \mathbf{v}_k.\end{aligned}$$

Model uncertainty lies in the emission 'u'

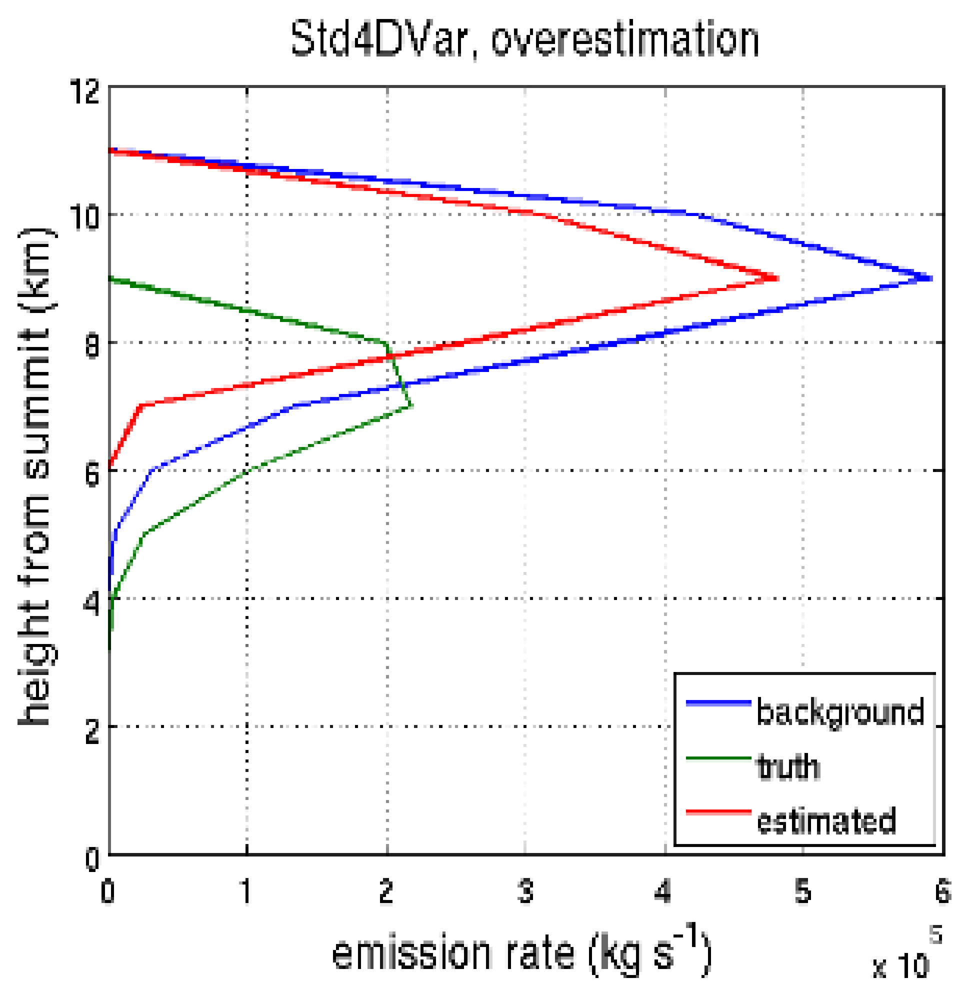
Measurement uncertainty (error) v_k

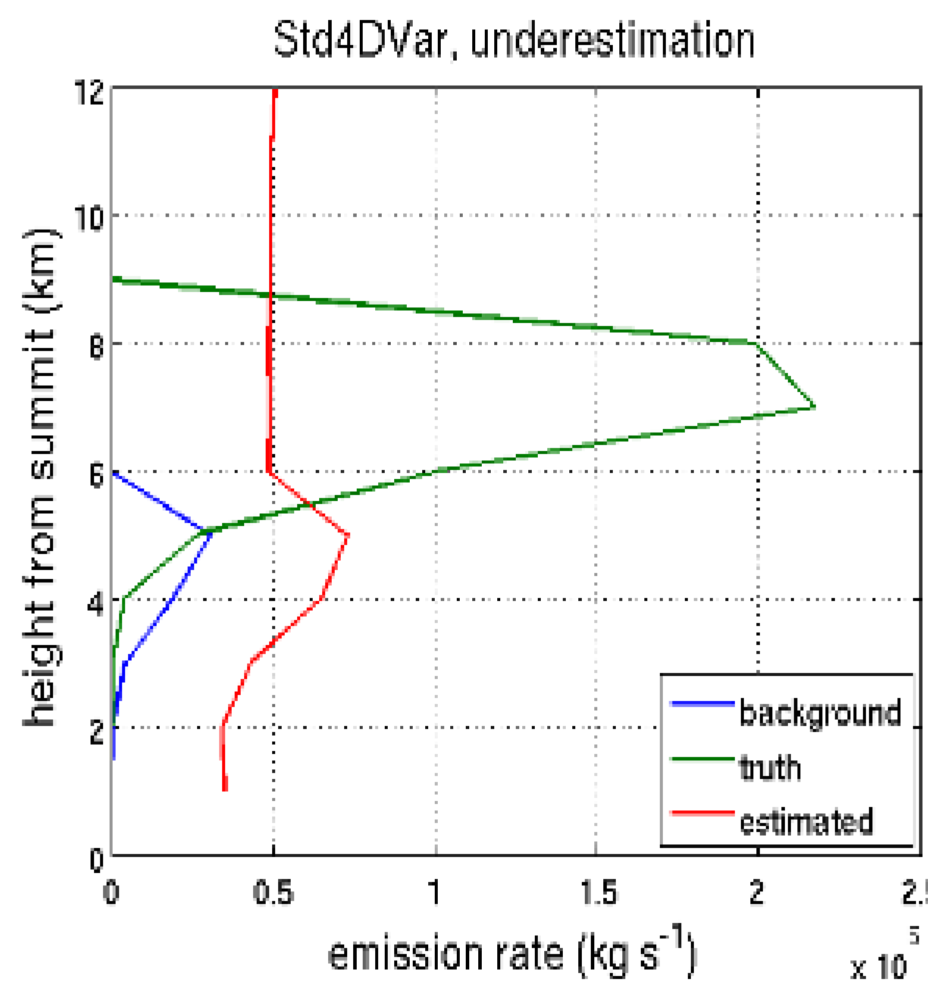
Typical cost function of standard 4DVar:

$$J(\mathbf{u}_k) = \frac{1}{2}\sum_{k=0}^{N}\left(\mathbf{u}_k - \mathbf{u}_k^b\right)^T \mathbf{B}_k^{-1}\left(\mathbf{u}_k - \mathbf{u}_k^b\right) + \frac{1}{2}\sum_{k=0}^{N}(\tilde{\mathbf{y}}_k - \mathbf{y}_k)^T \mathbf{R}_k^{-1}(\tilde{\mathbf{y}}_k - \mathbf{y}_k)$$

$$= J^b + J^o,$$

Ill-conditioned problem
due to ‘spurious relationship’



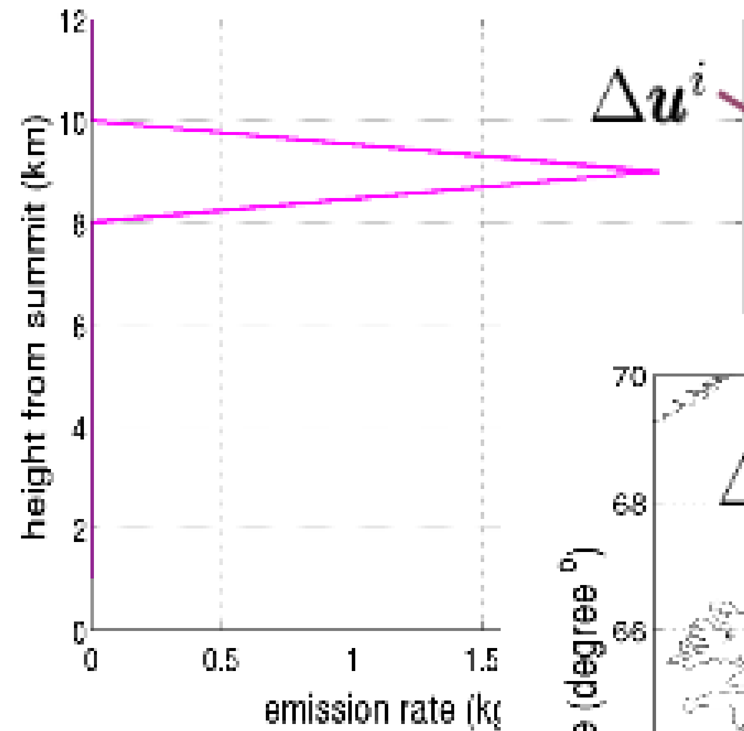


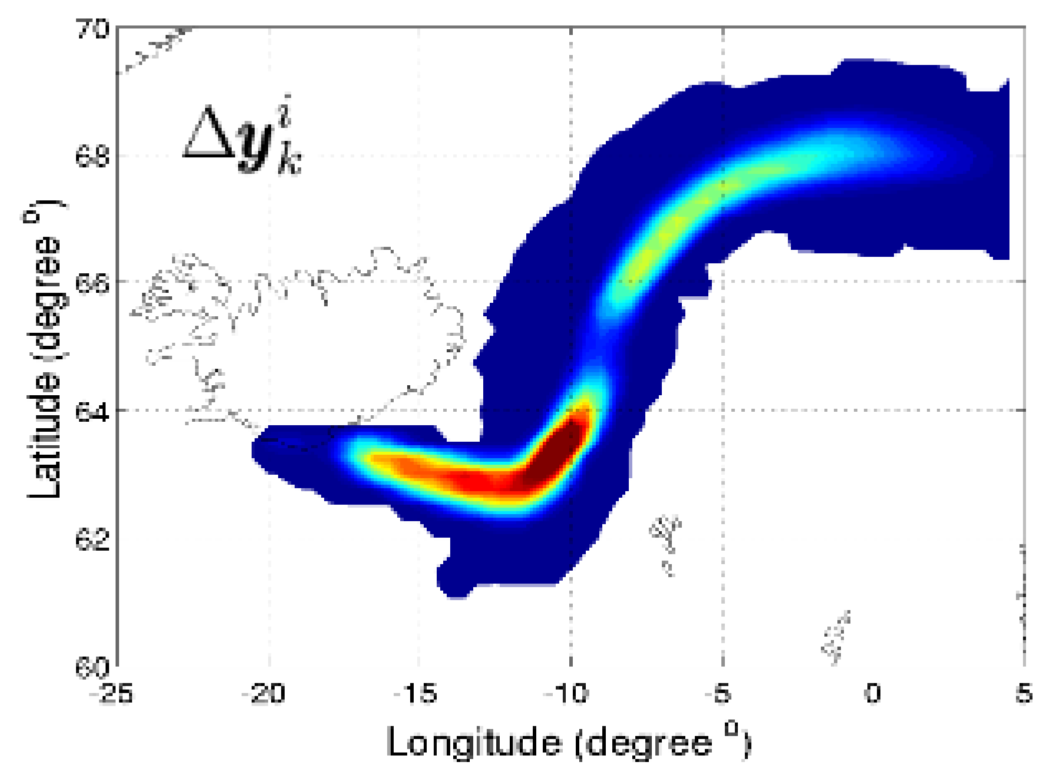
Trajectory-based 4D-Var

Assumption/observation:
Horizontal transport is stronger than vertical transport and diffusion.

The emission is assumed to be a linear combination of the perturbation sets:

$$\boldsymbol{u} = \boldsymbol{u}^b + \sum_{i=1}^{p} \beta^i \Delta \boldsymbol{u}^i,$$





Trajectories
:

$$\boldsymbol{y}_k^0 = \mathcal{H}_k(\mathcal{M}_k(\boldsymbol{x}_{k-1}, \boldsymbol{u}^b)).$$

$$\Delta \boldsymbol{y}_k^i = \mathcal{H}_k(\mathcal{M}_k(\boldsymbol{x}_{k-1}, \boldsymbol{u}^b + \Delta \boldsymbol{u}^i)) - \boldsymbol{y}_k^0.$$

The reformulated cost function of Trj4DVar:

$$\begin{aligned} J(\boldsymbol{\beta}) &= \frac{1}{2} \sum_{k=1}^{Nt} \left(\sum_{i=1}^{p} \beta^i \Delta \boldsymbol{y}_k^i + \boldsymbol{y}_k^0 - \boldsymbol{y}_k \right)^T [\mathbf{R}_k]^{-1} \left(\sum_{i=1}^{p} \beta^i \Delta \boldsymbol{y}_k^i + \boldsymbol{y}_k^0 - \boldsymbol{y}_k \right) \\ &+ \frac{1}{2} \left(\boldsymbol{u} - \boldsymbol{u}^b \right)^T [\mathbf{B}_k]^{-1} \left(\boldsymbol{u} - \boldsymbol{u}^b \right) \\ &= J^o + J^b, \end{aligned}$$

Given:

$$\mathbf{x}_k = M_k(\mathbf{x}_{k-1}, \mathbf{u}_k + \mathbf{w}_k), \tag{1}$$

$$\mathbf{y}_k = H_k(\mathbf{x}_k) + \mathbf{v}_k . \tag{2}$$

with $\mathbf{u} = \mathbf{u}^b + \sum_{i=1}^{p} \beta^i \Delta \mathbf{u}^i$ Eq.(2) can be rewritten as

$$\begin{aligned}
\mathbf{y}_k &\approx H_k[M_k(\mathbf{x}_{k-1}, \mathbf{u}^b)] + \sum_{i=1}^{p} \beta^i \mathbf{H}_k \mathbf{M}_k(\mathbf{x}_{k-1}, \mathbf{u}^b) \Delta \mathbf{u}^i + \mathbf{v}_k \\
&\approx \mathbf{y}_k^0 + \sum_{i=1}^{p} \beta^i \{ H_k[M_k(\mathbf{x}_{k-1}, \mathbf{u}^b + \Delta \mathbf{u}^i)] - \mathbf{y}_k^0 \} + \mathbf{v}_k \\
&= \mathbf{y}_k^0 + \sum_{i=1}^{p} \beta^i \Delta \mathbf{y}_k^i + \mathbf{v}_k , \qquad (5)
\end{aligned}$$

leading to the trajectory-based 4D-Var formulation:

$$\begin{aligned}
J(\beta) &= \frac{1}{2} \sum_{k=1}^{Nt} \left(\sum_{i=1}^{p} \beta^i \Delta \mathbf{y}_k^i + \mathbf{y}_k^0 - \mathbf{y}_k \right)^{\mathrm{T}} [\mathbf{R}_k]^{-1} \left(\sum_{i=1}^{p} \beta^i \Delta \mathbf{y}_k^i + \mathbf{y}_k^0 - \mathbf{y}_k \right) \\
&\quad + \frac{1}{2} \sum_{k=1}^{Nt} (\mathbf{u} - \mathbf{u}^b)^{\mathrm{T}} [\mathbf{B}_k]^{-1} (\mathbf{u} - \mathbf{u}^b) + \mu \|\nabla \mathbf{u}\|^2 \\
&= J^o + J^b + J^r ,
\end{aligned}$$

TUDe

Satellite
Data
assimilation

Estimates of emissions in twin experiments

E15 2010 eruption events as a case study: LOTOS-EUROS model, meteorological situation, synthetic observations

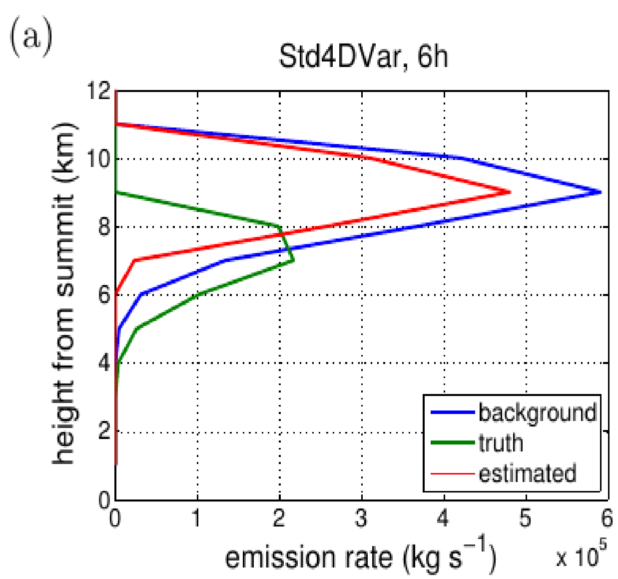
Testing the methodology:

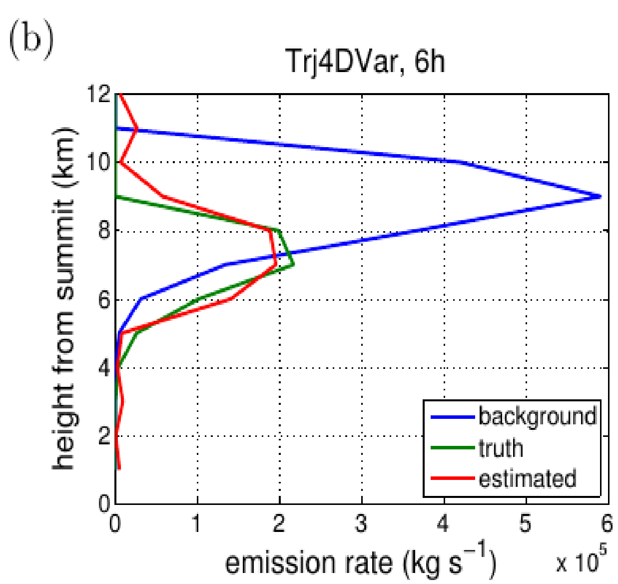
- Standard 4D-Var vs.
- Trajectory-based 4D-Var

- Deterministic model

- Observations: synthetic ash column data

- 6-hour assimilation window





Testing some of the choices:

- Trajectory-based 4D-Var

- Stochastic model

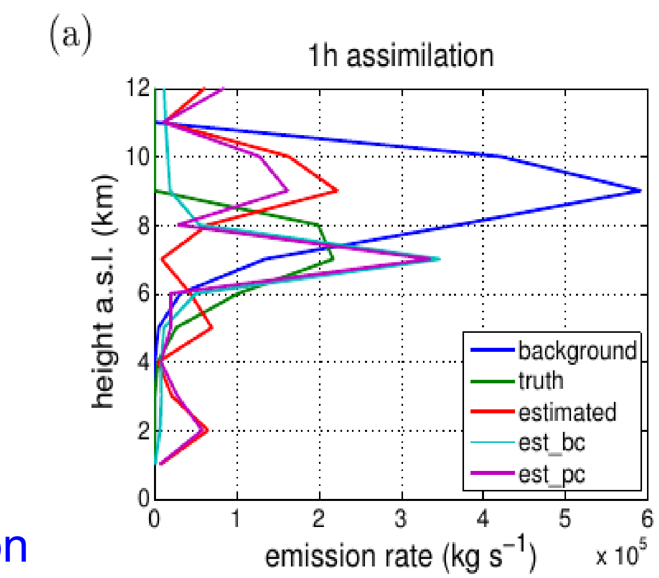
- Observations: synthetic ash columns with 50% uncertainty & plume height and mass eruption rate

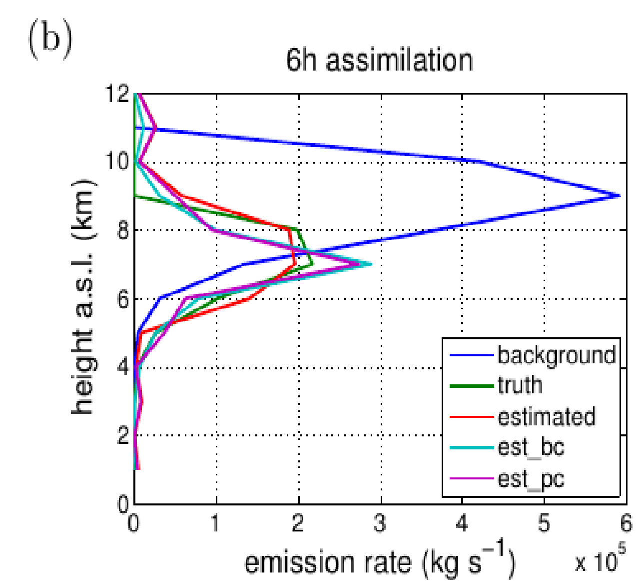
Settings and choices:

- Length of assimilation windows

- Uncertainties of Observations

- Schemes to integrate
- multi-observations





TU Delft

Lu, S., Lin, H.X., Heemink, A.W., Fu, G., and Segers, A.J. (2016). Estimation of Volcanic Ash Emissions Using Trajectory-Based 4D-Var Data Assimilation. ***Monthly Weather Review*** *144, 575-589.*

Dust storm emission inversion using multiple data sources

Dust storm models
(chemical transport model)

- **Emissions;**
- **Transport;**

  advection, diffusion, radiation
- **Sedimentations.**

  wet, dry deposition

$$\frac{\partial C}{\partial t}+U\frac{\partial C}{\partial x}+V\frac{\partial C}{\partial y}+W\frac{\partial C}{\partial z}$$

$$=\frac{\partial}{\partial x}\left(K_h\frac{\partial C}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_h\frac{\partial C}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z\frac{\partial C}{\partial z}\right)$$

$$+E+R+Q-D-H$$

Example: Lotos-Euros/Dust over East Asia

2.1. Data assimilation with Lotos-Euros: algorithm

Data assimilation: to find a solution that fits both the observation and priori.

Traditional 4 dimensional variational (4DVar) data assimilation:

$$J(\delta \boldsymbol{f}) = \frac{1}{2}\,\delta \boldsymbol{f}\;\mathbf{B}^{-1}\;\delta \boldsymbol{f} + \frac{1}{2}\sum_{i=1}^{k}(\mathbf{H}_i\mathbf{M}_i\delta \boldsymbol{f} + \boldsymbol{d}_i)^T\,\mathbf{O}_i^{-1}\,(\mathbf{H}_i\mathbf{M}_i\delta \boldsymbol{f} + \boldsymbol{d}_i)$$

$$\boldsymbol{d}_i = \mathcal{H}_i(\;\mathcal{M}_i(\boldsymbol{f})\;) - \boldsymbol{y}_i$$

$\mathbf{M}_i$ **full tangent linear model**
Order of **O(10^5)**

Reduced-tangent-linearization 4DVar

$$\mathbf{B} = \mathbf{U}\mathbf{U}^T \approx \tilde{\mathbf{U}}\tilde{\mathbf{U}}^T$$

$$\delta \boldsymbol{f} \approx \tilde{\mathbf{U}}\delta \boldsymbol{w}$$

$$J(\delta \boldsymbol{w}) = \frac{1}{2}\delta \boldsymbol{w}^T\delta \boldsymbol{w} + \frac{1}{2}\sum_{i=1}^{k}(\mathbf{H}_i\tilde{\mathbf{M}}_i\tilde{\mathbf{U}}\delta \boldsymbol{w} + d_i)^T\,\mathbf{O}_\mathbf{i}^{-1}\,(\mathbf{H}_i\tilde{\mathbf{M}}_i\tilde{\mathbf{U}}\delta \boldsymbol{w} + d_i)$$

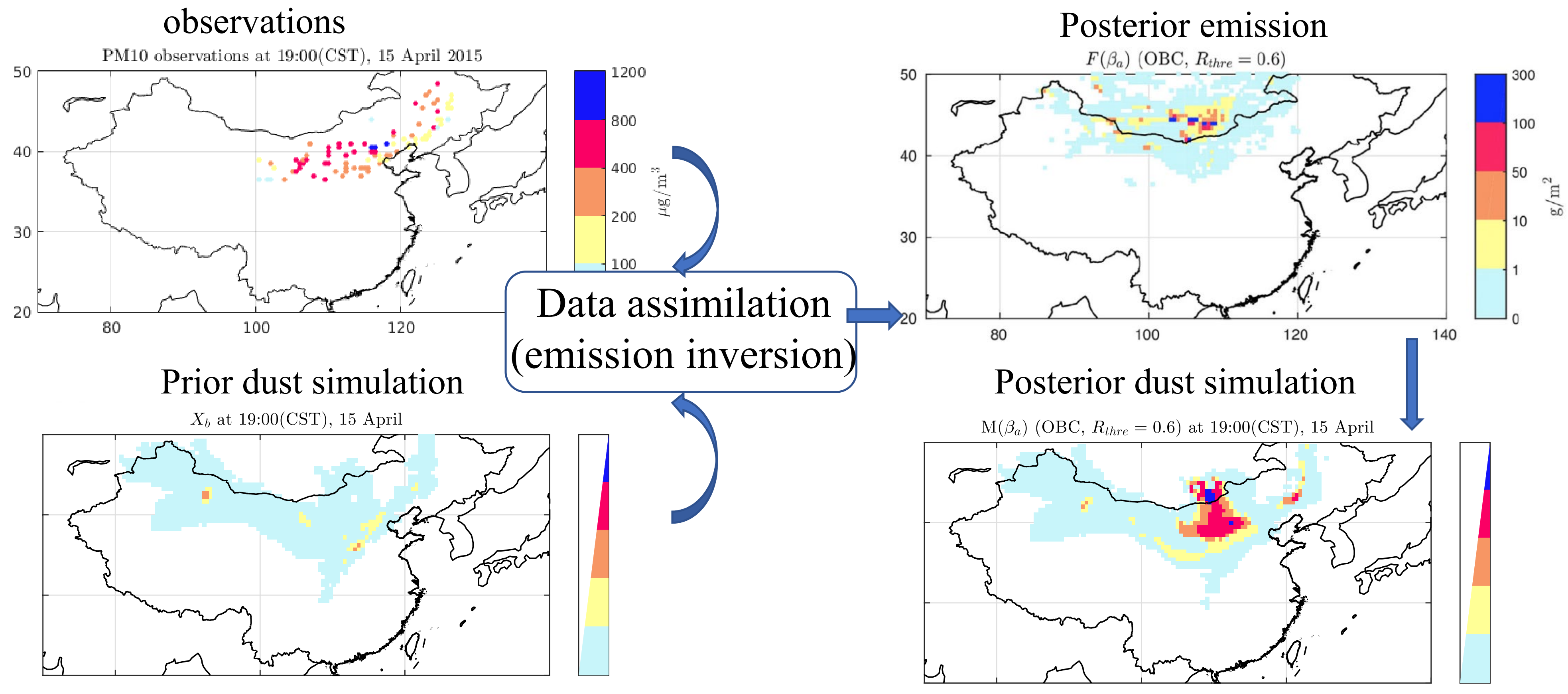
$\tilde{\mathbf{M}}_i$ **reduced tangent linear model**
Order of **O(10^2)**

Sensitivity-based parameter filters: To reduce the size of $\delta \boldsymbol{f}$, improve the computation efficiency

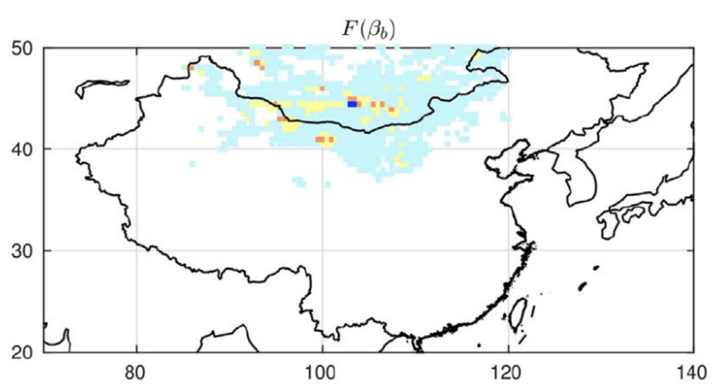
Dust storm emission inversion using multiple data sources

2.1 Data assimilation with Lotos-Euros: assimilating PM10

China MEP monitoring network:
- PM_{10}, $PM_{2.5}$, SO_2, NO_X, O_3
- wide coverage
- high accuracy
- hourly measured
- Over 1,500 sites



Prior emission

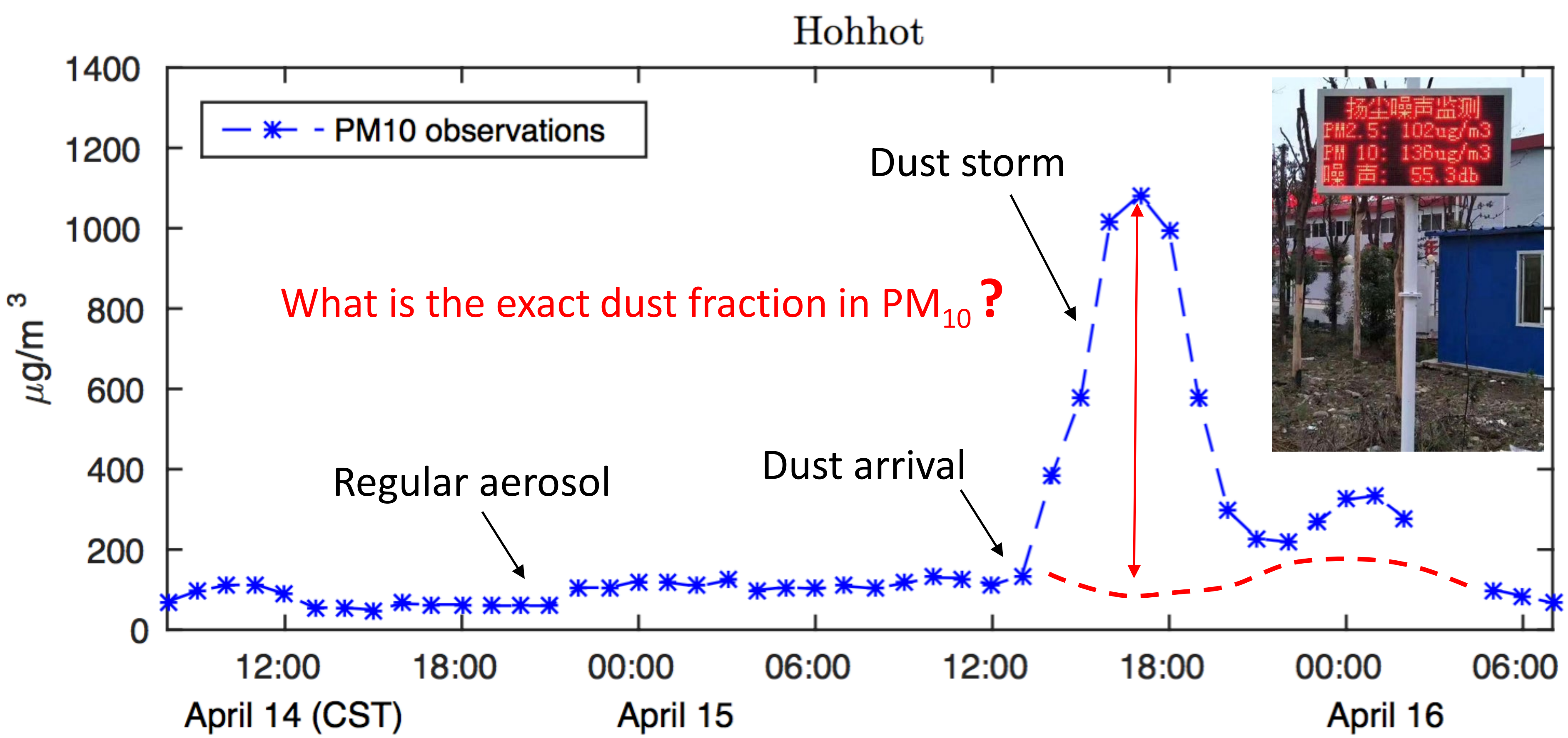


***J.Jin** et al., 2018, Spatially varying parameter estimation …, Atmospheric Environments*

Dust storm emission inversion using multiple data sources

2.2 Machine learning based observation bias correction: bias/baseline

Existence of bias in PM10 concentration for its use in data assimilation

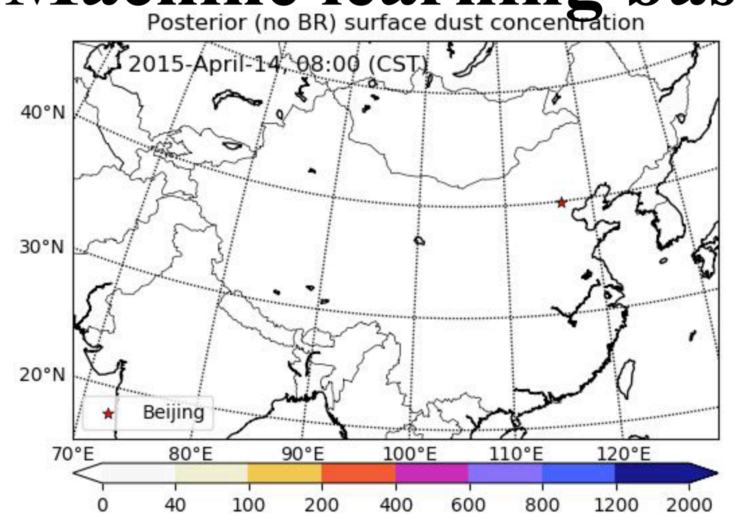


- **PM10 observation is a sum of non-dust and dust aerosols, thus includes a bias when representing the dust concentration.**
- **Issue:** the data assimilation algorithm cannot calculate whether the error is caused by the model deficiency or observation bias.
- **Challenge:** bias with strong spatial and temporal variability
- **Why not full aerosol model???**

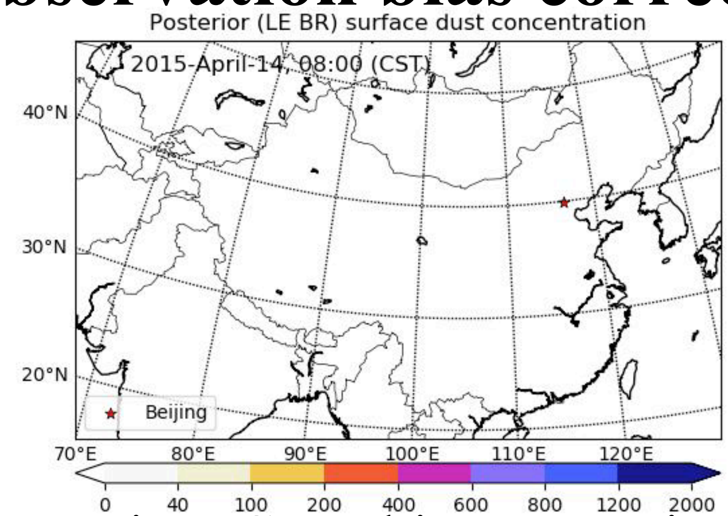
TU Delft Delft University of Technology
TNO innovation for life

Dust storm emission inversion using multiple data sources

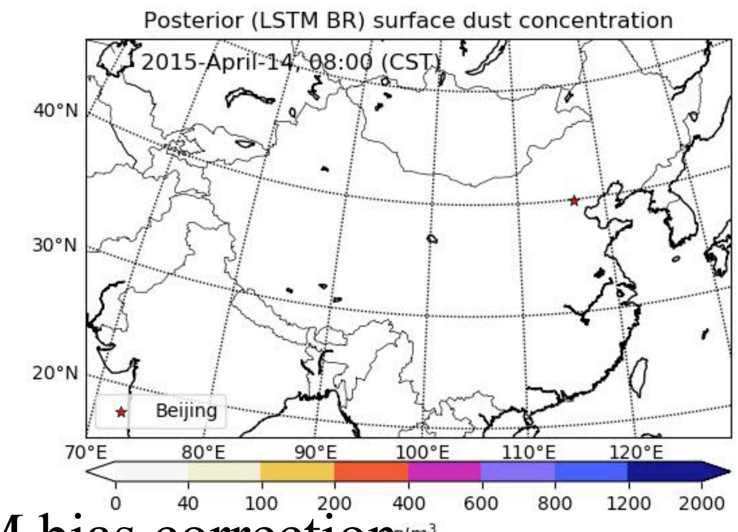
2.2 Machine learning based observation bias correction: assimilation evaluation



Posteriors no bias correction



Posteriors CTM bias correction



Posteriors LSTM bias correction

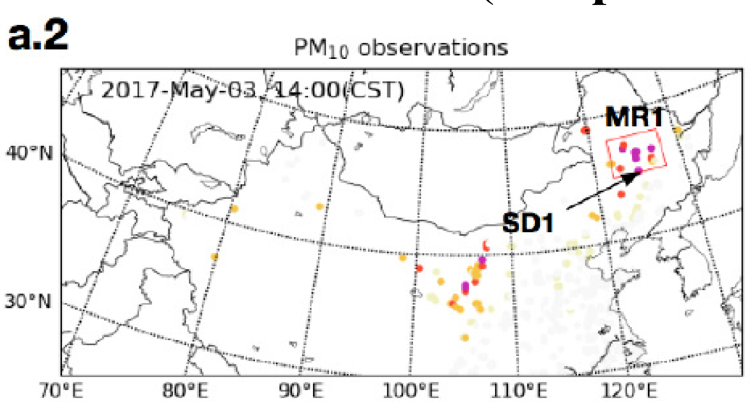
- Assimilation of machine learning bias corrected data gives the most accurate posterior;
- **Direct assimilation of PM_{10} causes overestimation of dust simulations.**

J.Jin *et al, Machine learning for observation bias correction with application to dust storm data assimilation. (ACP Discussion)*

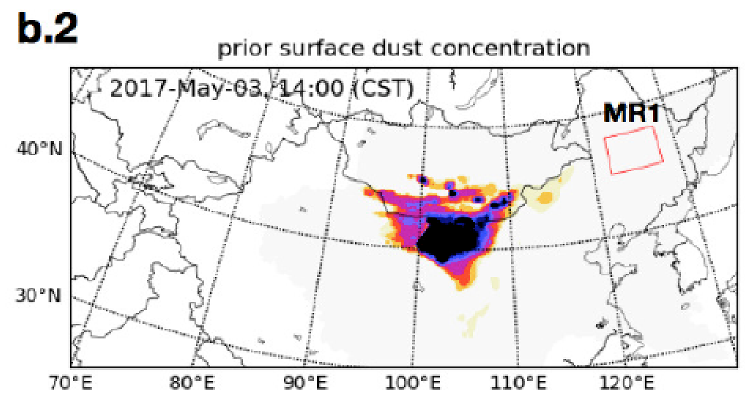
H.X.Lin, ***J.Jin*** *et al. air quality forecast through integrated data assimilation and machine learning. ICCART, Prague, 2019.*

2.3 Emission detection using adjoint: no dust simulated in northeast China

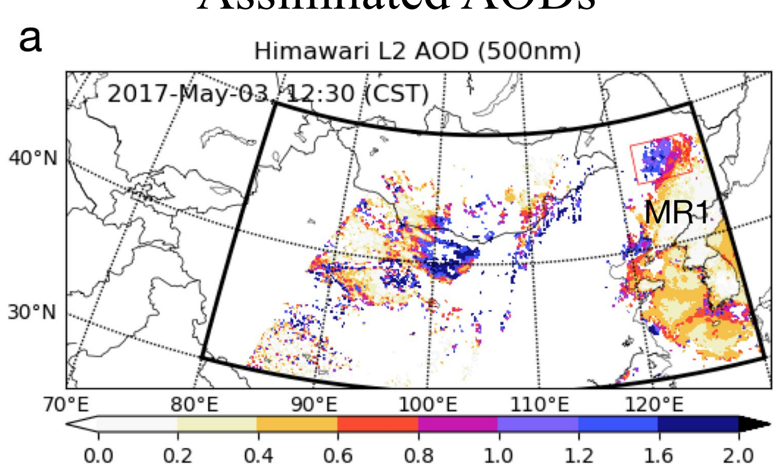
PM10 observations (**independent**)



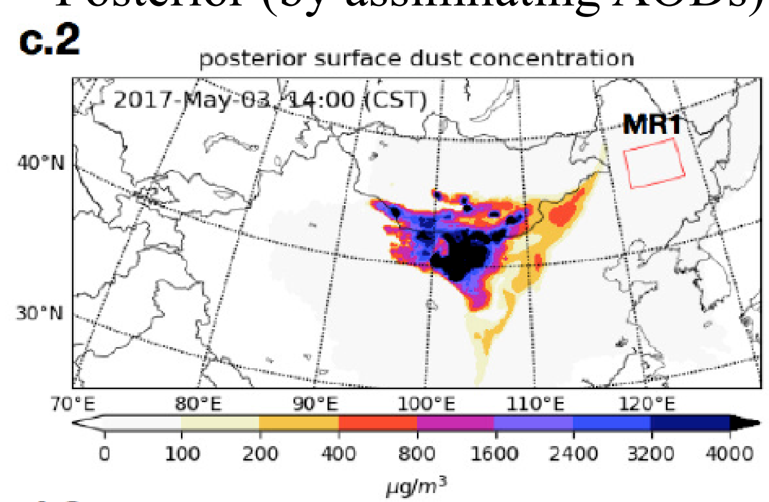
Prior



Assimilated AODs



Posterior (by assimilating AODs)



- No dusts are simulated in prior or posterior model;
- Other **two** dust outbreaks are also not reproduced.
- **Solution: to detect the (missing) sensitive emissions for the dust outbreak**

2.3 Emission detection using adjoint: theory

- CTM Model: $x^{i+1} = \mathcal{M}^i(x^i,\ f^i)$
- Sensitivity calculation
- Linearized model operator: $\frac{\partial x^{i+1}}{\partial f^i} = \frac{\partial \mathcal{M}^i}{\partial f^i} = \mathrm{M}^i_f$ and $\frac{\partial x^{i+1}}{\partial x^i} = \frac{\partial \mathcal{M}^i}{\partial x^i} = \mathrm{M}^i_x$

finite difference: **accurate but intensive**

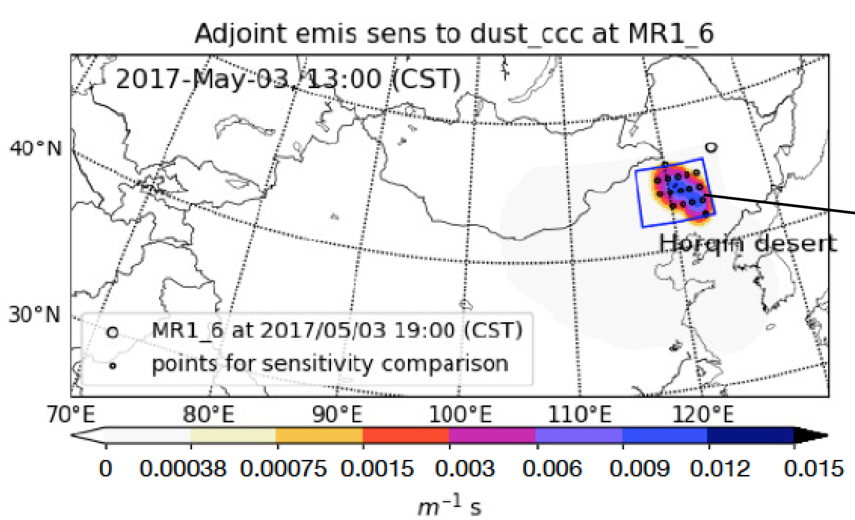
$\boxed{(x^i,\ f^i)}$ forward model $\boxed{x^{i+1}}$

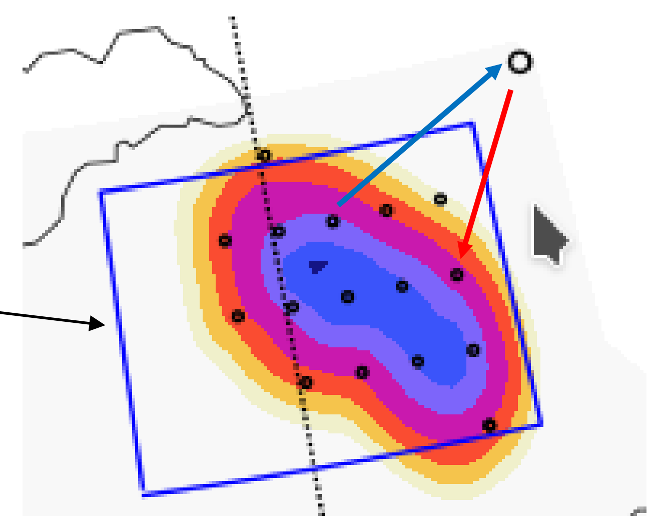
backward model

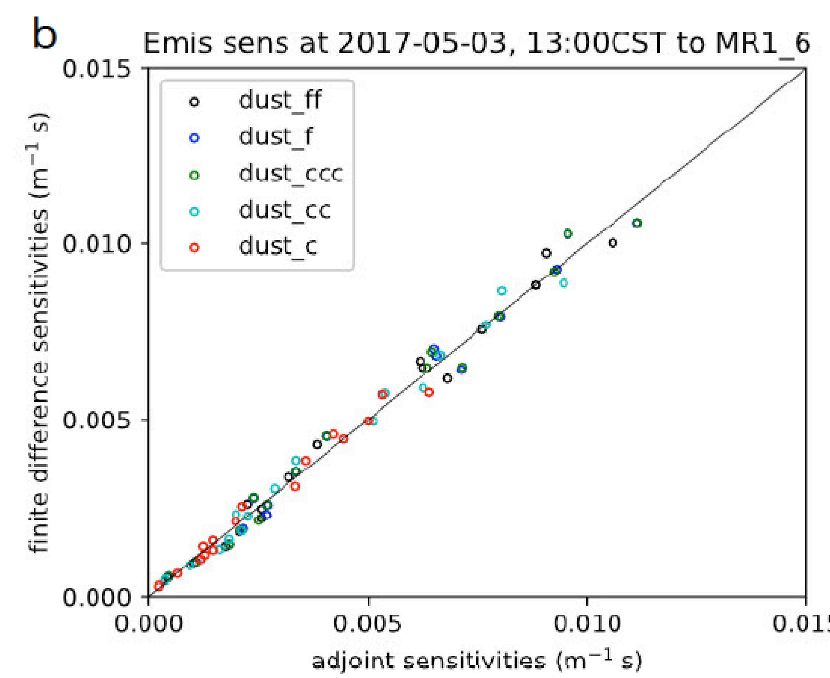
adjoint method: **efficient but ……**

Gradient of model response J to parameters: $\nabla_{f^j}\mathcal{J}(x^i) = (\mathrm{M}^j_f)^T \cdot (\mathrm{M}^{j+1}_x)^T \cdots (\mathrm{M}^{i-1}_x)^T \cdot \{\frac{\partial \mathcal{J}(x^i)}{\partial x^i}\}^T$

(adjoint vs. finite difference):





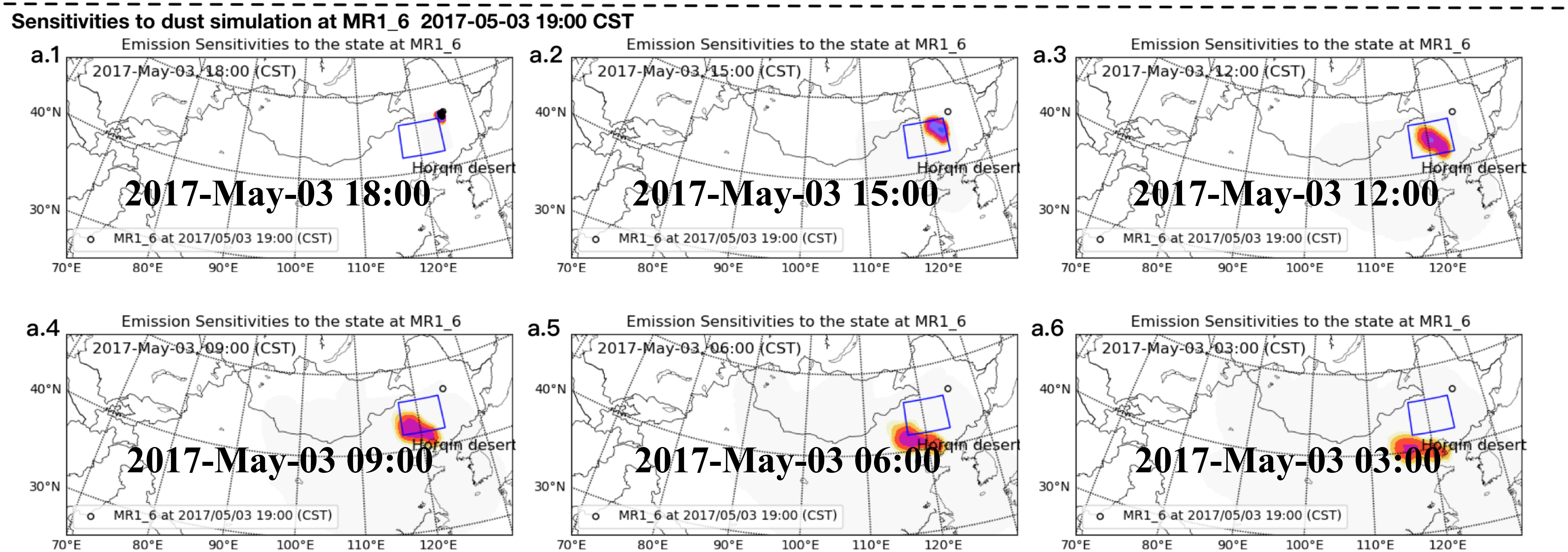


Delft University of Technology
TU Delft
TNO innovation for life

Dust storm emission inversion using multiple data sources

2.3 Emission detection using adjoint: emission backtracking

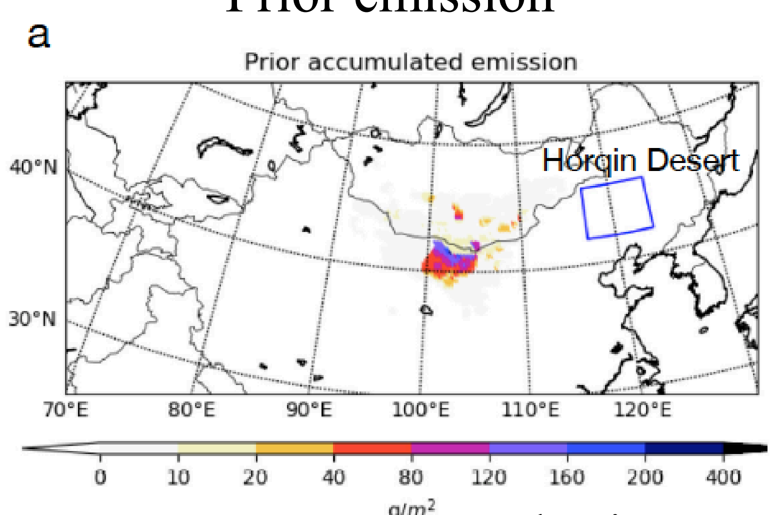
Time series of emission sensitivities to a state X at **2017-May-03 19:00**



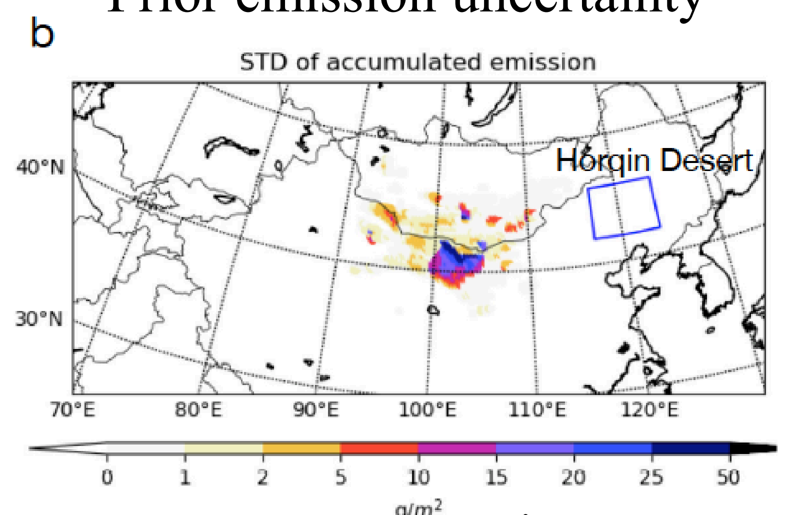
Dust storm emission inversion using multiple data sources

2.3 Emission detection using adjoint: guided emission reconstruction

Prior emission

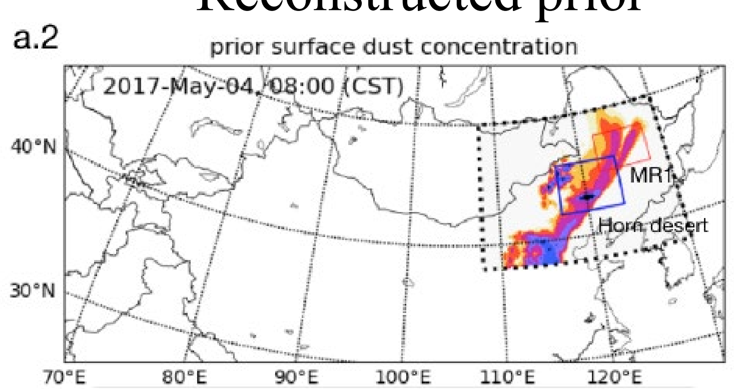


Prior emission uncertainty

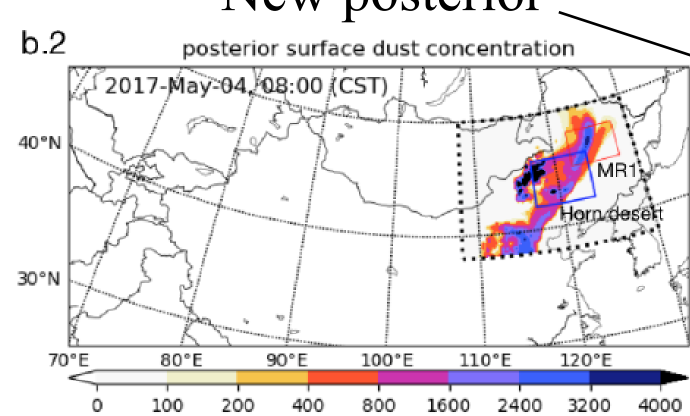


Landcover over **Horqin**

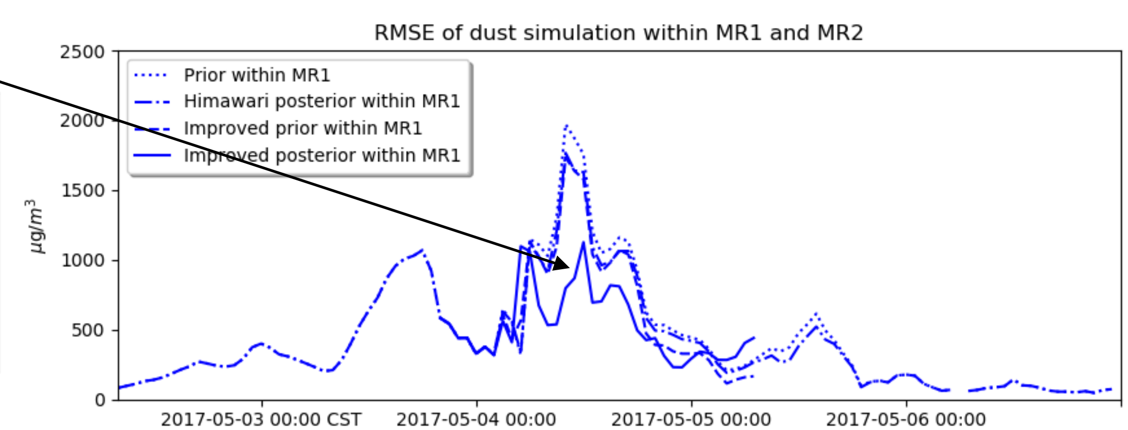
Reconstructed prior



New posterior

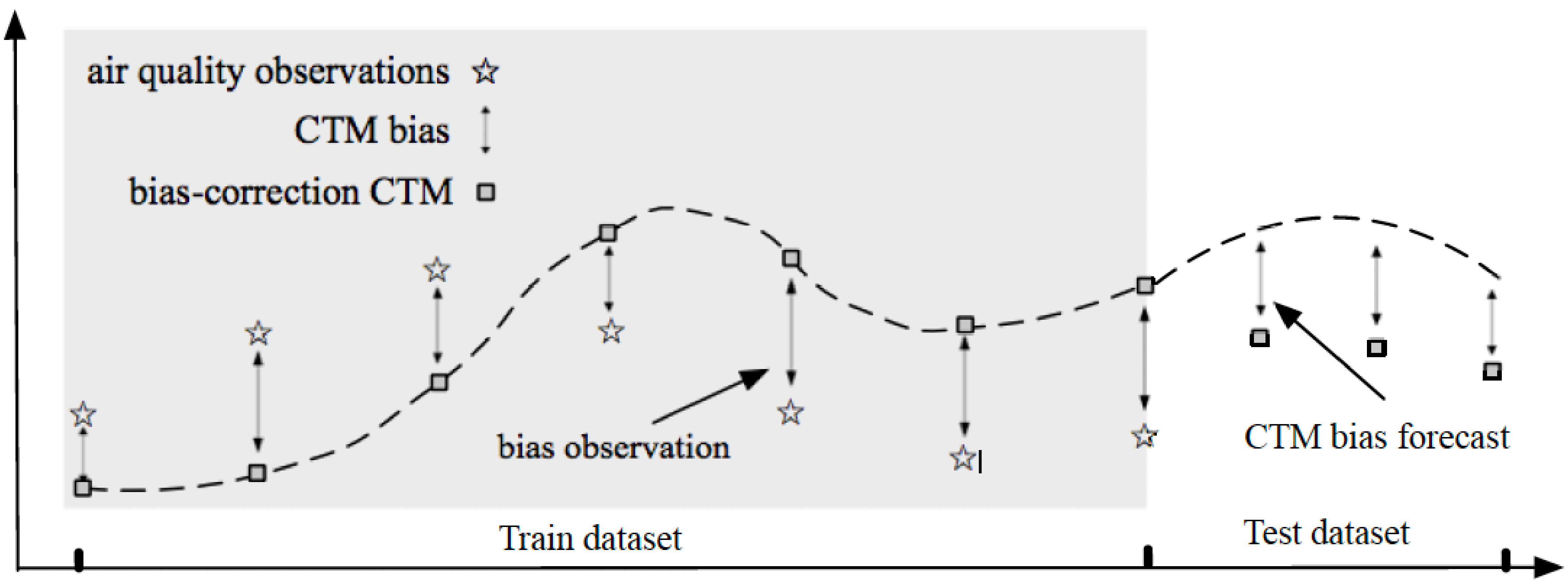


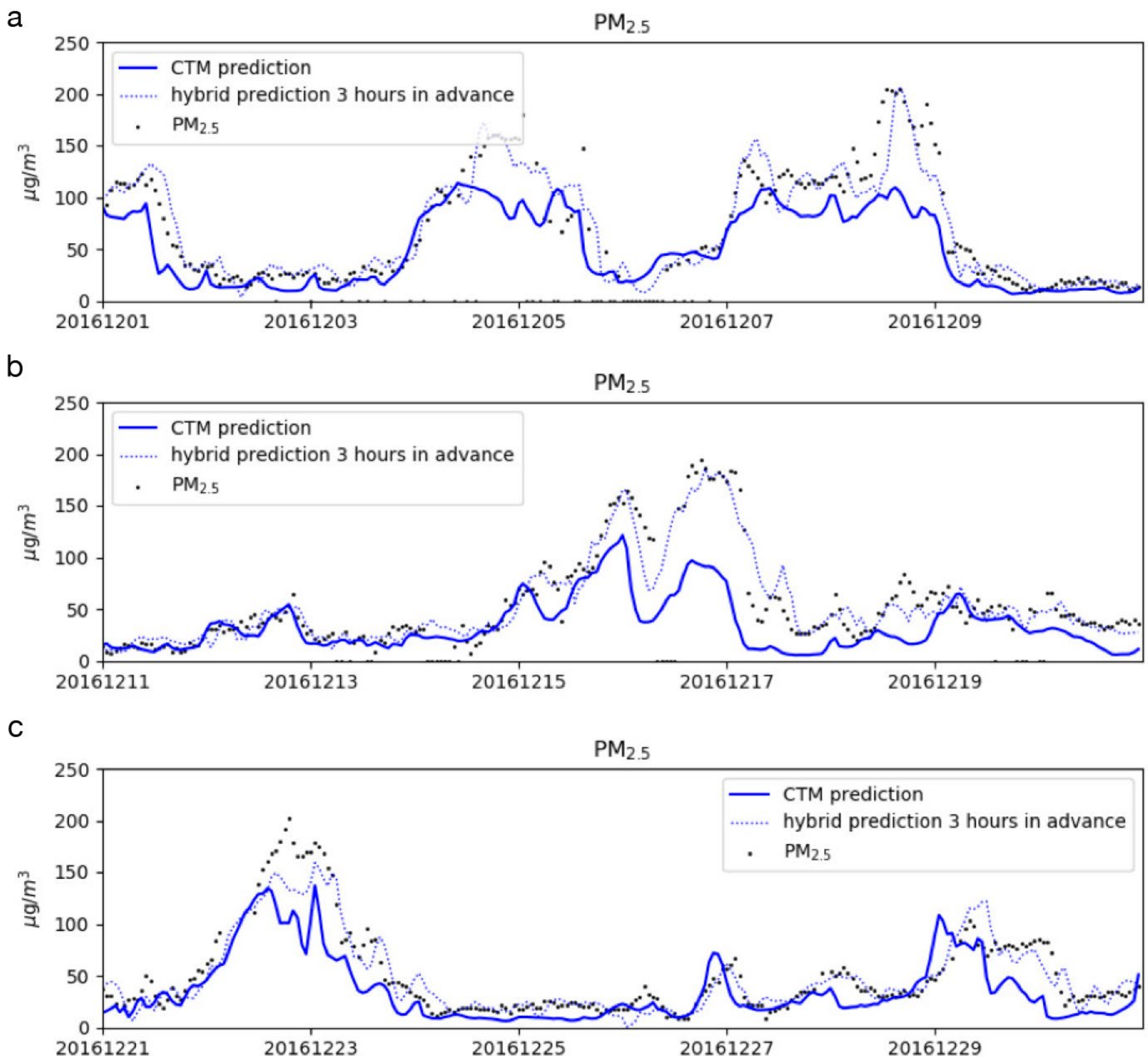
RMSE



J.Jin *et.al.*, Source backtracking for dust storm emission inversion using an adjoint method.
Atmospheric Chemistry and Physics, 2021

2.4 Machine Learning: error bias correction





Improved forecast after ML bias correction

Figure 4: Time series of the $PM_{2.5}$ observation, CTM prediction and hybrid forecasts with 3 hours in advance. (a): Dec 01-10; (b): Dec 11-20; (c): Dec 21-30.

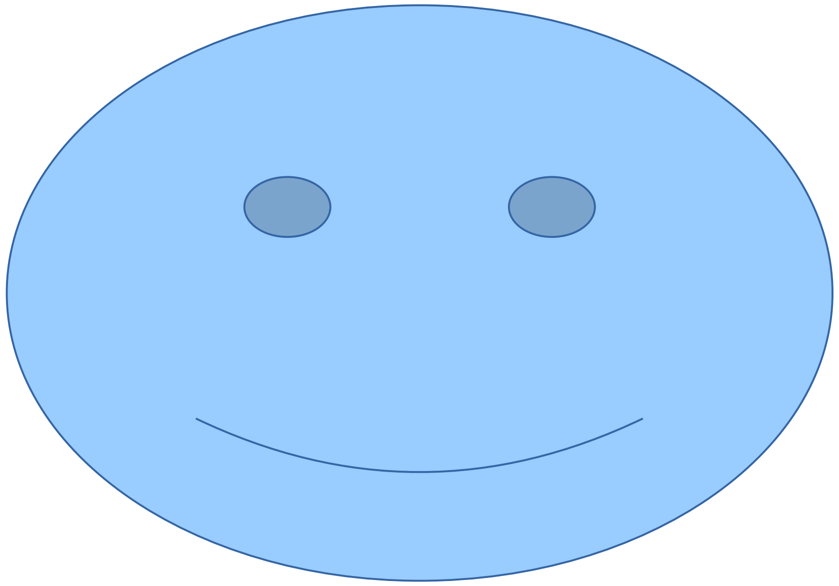
Delft University of Technology
Universiteit Leiden

Data assimilation and machine learning for air quality forecasts

- Trajectory-based 4DVar is effective in estimating volcano eruption plume shapes using satellite data
- The integration of machine learning and data assimilation results in more accurate air quality forecast during dust storms.
- Inclusion of physical model knowledge further enhances the learning process in (atmospheric) modeling.
- Integration ML and DA opens many new possibilities, such as filling unmodelled prrocess in CTM with ML, using an ML surrogate to replace computation intensive (sub)models, …

Challenges

- Explainable ML model
- Can DA sometimes converge to ‘truth’ (certainty)? Accurate Uncertainty quantification is the key.
- Computational cost
- …



Thank you!

References

1. J. Jin, A.J. Segers. H.X. Lin, B. Henzing, X. Wang, A.W. Heemink, H. Liao (2021). Position correction in dust storm forecast using LOTUS-EUROS v2.1: grid distorted data assimilation v1.0, Geoscientific Model Development.
2. M. Xu, J. Jin, G. Wang, A Segers, T. Deng, H.X. Lin (2021), Machine learning based bias correction for numerical chemical transport models, Atmospheric Environment, 118022
3. J. Jin, A. Segers, H. Liao, A.W. Heemink, R. Kranenburg, H.X. Lin (2020), Source backtracking for dust storm emission inversion using adjoint method: case study of northeast China, Atmospheric Chemistry and Physics.
4. C. Xiao, O. Leeuwenburgh, H.X. Lin, A.W. Heemink (2021), Conditioning of Deep-Learning Surrogate Models to Image Data with Application to Reservoir Characterization, Knowledge Based Systems.
5. C. Xiao, H.X. Lin, O. Leeuwenburgh, A.W. Heemink (2022), Surrogate-assisted inversion for large-scale history matching: comparative study between projection-based reduced-order modelling and deep neural network, Journal of Petroleum Science and Engineering.
6. H.X. Lin, J. Jin, H.J. van den Herik (2019), Air Quality Forecast through Integrated Data Assimilation and Machine Learning, in Proc. 11th International Conference on Agents and Artificial Intelligence (ICAART 2019).
7. J. Jin, H.X. Lin, A. Segers, Y. Xie, A.W. Heemink (2019), Machine learning for observation bias correction with application to dust storm data assimilation, Atmospheric Chemistry and Physics, Vol.19, pp. 10009-10026.
8. J. Jin, A.W. Heemink, A. Segers, M. Yoshida, W. Han, H.X. Lin (2019), Dust Emission Inversion Using Himawari-8 AODs Over East Asia: an Extreme Dust Event in May 2017, Journal of Advances of Modeling Earth Systems, pp. 446-467.
9. J. Jin, H.X. Lin, A.W. Heemink, A. Segers (2018), Spatially varying parameter estimation for dust emissions using reduced-tangent-linearization 4DVar, Atmospheric Environment, Vol.187, pp. 358-373.
10. S. Lu, A. Heemink, H.X. Lin, G. Fu, A Segers (2017), Evaluation criteria on the design for assimilating remote sensing data using variational approaches, *Monthly Weather Review* 145(6), pp.2165-2175
11. G. Fu, H.X. Lin, A.W. Heemink, S. Lu, A. Segers, N. van Velzen, T.C. Lu, and S.M. Xu (2017), Accelerating volcanic ash data assimilation using a mask-state algorithm based on an ensemble Kalman filter: a case study with the LOTOS-EUROS model (version 1.10), *Geoscientific Model Development*, 10, pp.1751–1766
12. G. Fu, F. Prata, H.X. Lin, A.W. Heemink, S. Lu, A.J. Segers (2017) Data assimilation for volcanic ash plumes using a Satellite Observational Operator: a case study on the 2010 Eyjafjallajokull volcanic eruption, *Atmospheric Chemistry and Physics*, Vol17(2), pp. 1187—1205, DOI: 10.5194/acp-17-1187-2017
13. S. Lu, H.X. Lin, A. Heemink, A Segers, G. Fu (2016) Estimation of volcanic ash emissions through assimilating satellite data and ground-based observations, *Journal of Geophysical Research: Atmospheres*, Vol. 121 (18), pp. 10971–10994.
14. G. Fu, A.W. Heemink, S. Lu, A.J. Segers, K. Weber, H.X. Lin (2016) Model-based aviation advice on distal volcanic ash clouds by assimilating aircraft in-situ measurements, *Atmospheric Chemical Physics,* doi:10.5194/acp-2016-166,
15. S. Lu, H.X. Lin, A.W. Heemink, G. Fu, A. Segers (2016), Estimation of volcanic ash emissions using trajectory-based 4D-Var data Assimilation, *Monthly Weather Review,* Vol. 143, pp.575-589.
16. G. Fu, H.X. Lin, A.W. Heemink, A.J. Segers, S. Lu, T. Palsson (2015) Assimilating aircraft-based measurements to improve Forecast Accuracy of Volcanic Ash Transport, *Atmospheric Environment,* Vol.115, pp. 170-184.