Inspire Create Transform

Vigilada Mineducación

TU Delft Delft University of Technology

UNIVERSIDAD EAFIT®

Covariance Localization and Parameter Estimation using Ensemble-Based Data Assimilation.

Santiago Lopez-Restrepo

Olga Lucía Quintero Montoya
Nicolás Pinel
Arnold Heemink

Mathematical Modelling Research Group

Doctoral Seminar 3

Medellín, Colombia.

2018

Inspire Create Transform

Outline

Motivation

EnKF

Covariance Localization

Parameter Estimation

Preliminary Results

Motivation

Ensemble-Based data assimilation

Uses an ensemble to model the statics of the first guess

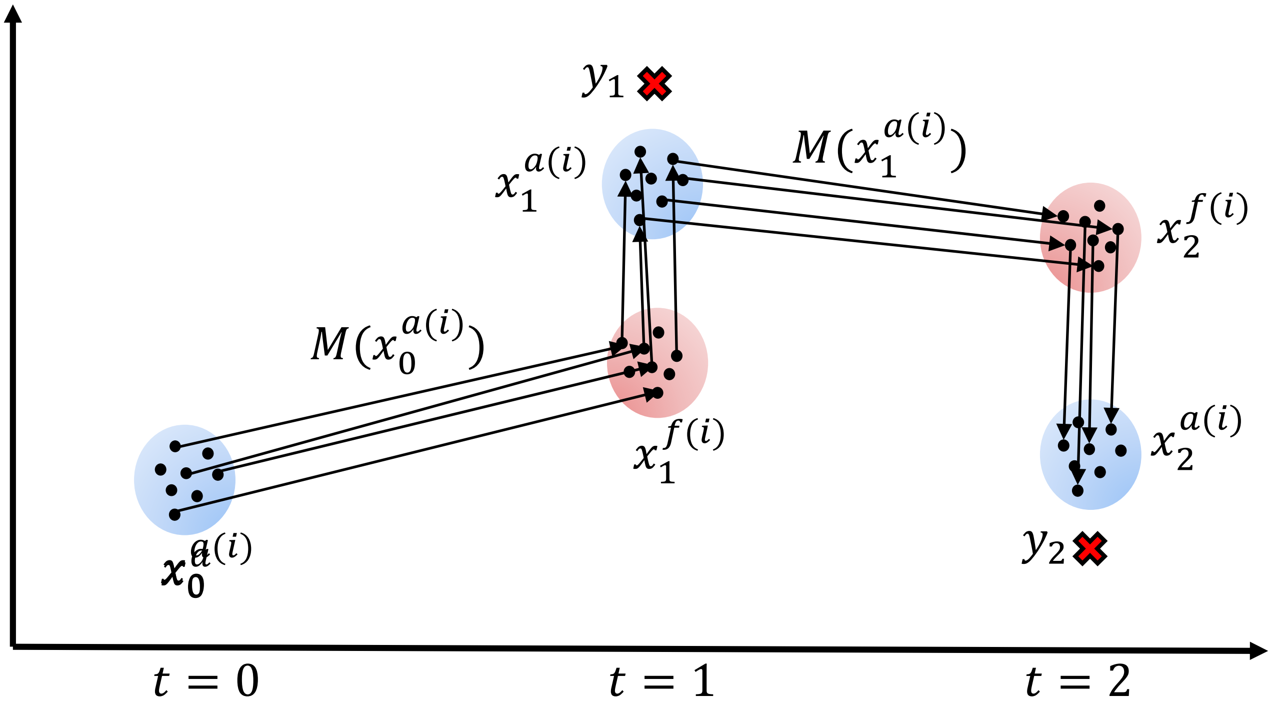
Updating

Using observation this forecast is modified at each step

Advantage

Easy implementations, low computational cost, general formulation

Ensemble Kalman Filter



Sequential estimation scheme for the Ensemble Kalman Filter. The x-axis denotes time; the y-axis is the estimated variable

Ensemble Kalman Filter

x_{t+1}^{f}

Forecast step

$$\boldsymbol{x}_t^{f(i)} = \boldsymbol{\mathcal{M}}_t(\boldsymbol{x}_{t-1}^{a(i)})$$

$$\boldsymbol{L} = [\left(\boldsymbol{x}_t^{f(1)} - \overline{\boldsymbol{x}}_t^{f}\right), \dots, (\boldsymbol{x}_t^{f(N)} - \overline{\boldsymbol{x}}_t^{f})]$$

$$\widehat{\boldsymbol{P}}^{f} = [\boldsymbol{L}\boldsymbol{L}^T]/(N-1)$$

Analysis step

$$\boldsymbol{x}_t^{a(i)} = \boldsymbol{x}_t^{f(i)} + \widehat{\boldsymbol{K}}_t\left(\boldsymbol{y}_{t-}\boldsymbol{H}_t\boldsymbol{x}_t^{f(i)}\right)$$

$$\widehat{\boldsymbol{K}}_t = \widehat{\boldsymbol{P}}_t^{f}\boldsymbol{H}_t'(\boldsymbol{H}_t\widehat{\boldsymbol{P}}_t^{f}\boldsymbol{H}_t' + \boldsymbol{R}_t)^{-1}$$

$$\widehat{\boldsymbol{P}}_t^{a} = (1 - \widehat{\boldsymbol{K}}_t\boldsymbol{H}_t)\widehat{\boldsymbol{P}}_t^{f}$$

x_t^{a}

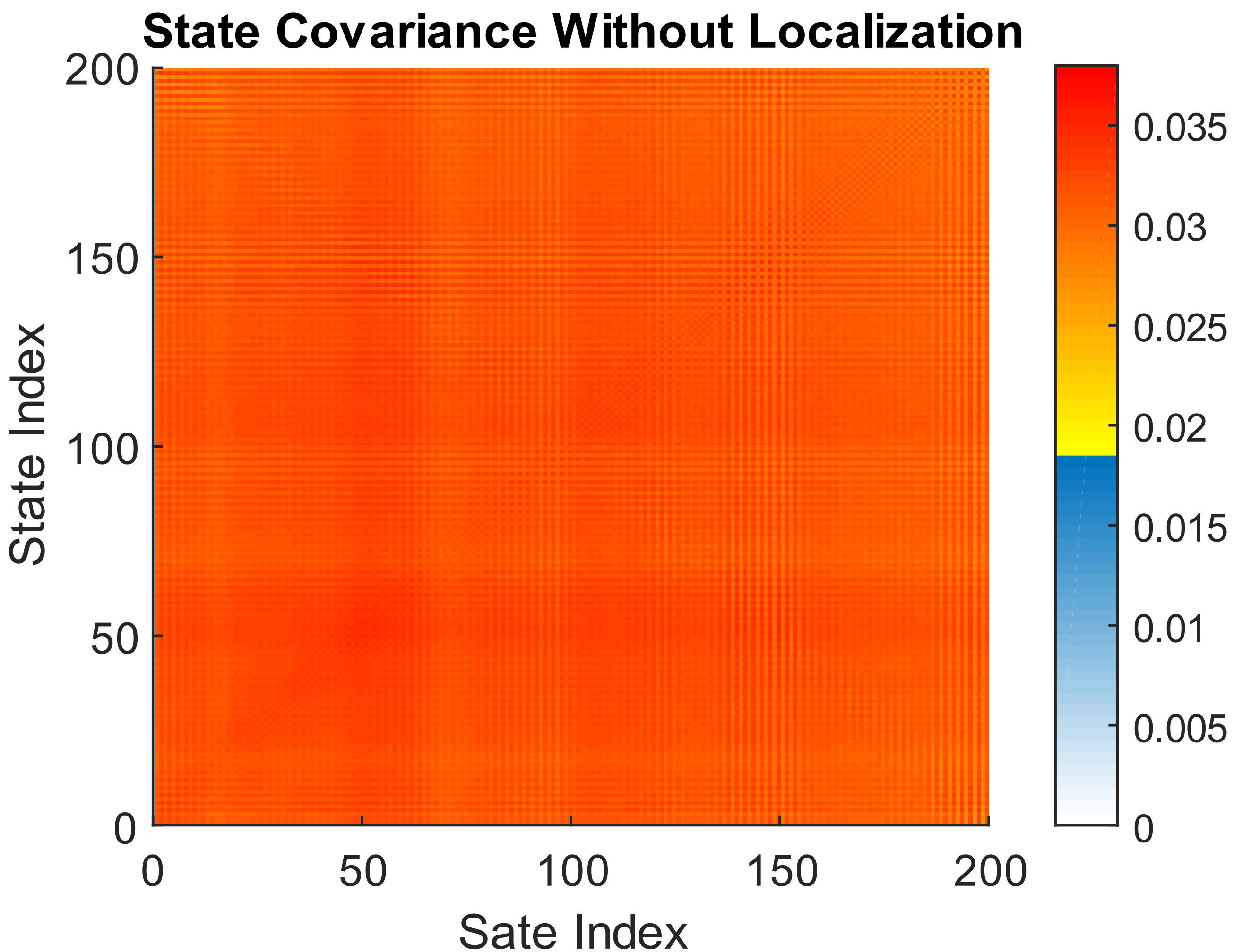
Ensemble Kalman Filter

De St Venant equations

$$\frac{\partial h}{\partial t} + D\frac{\partial u}{\partial x} = 0$$

$$\frac{\partial u}{\partial t} + g\frac{\partial h}{\partial x} + fu = 0$$

Ensemble Kalman Filter



Covariance Localization

Spurious correlations

Correlations between states not physically related.

Causes

Approximation of the forecast covariance by the ensemble covariance

Consequences

State variable may be incorrectly impacted by a remote observation

Covariance Localization

Covariance localization

Cutting off longer range correlations at a specified distance.

How

Applying a Schur product $(A \circ B) = A_{ij}B_{ij}$

Where

Between $\widehat{\boldsymbol{P}}_t^f$, and a correlation function with local support, ρ.

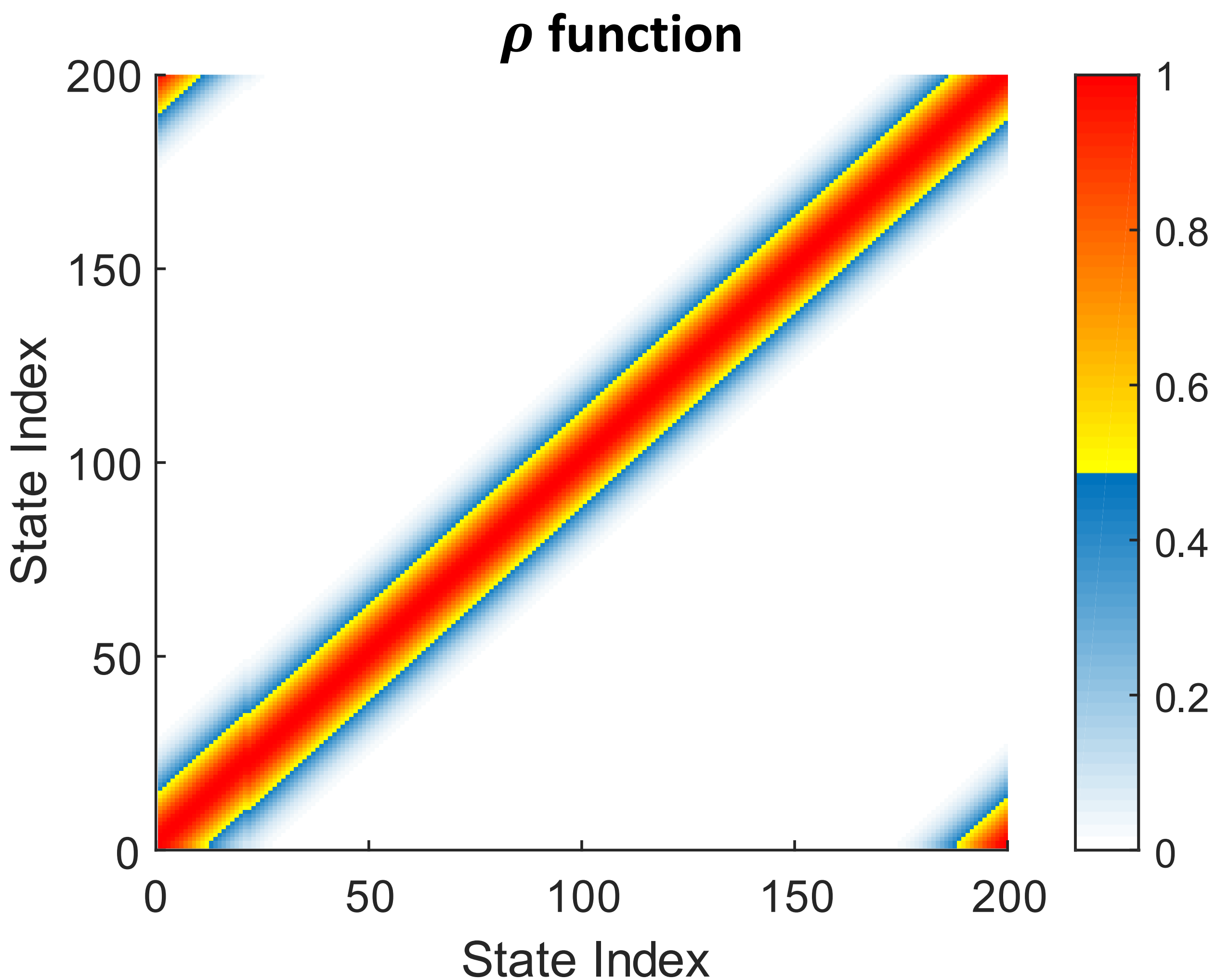
Covariance Localization

- The correlation function ρ is commonly taken as defined in Gaspari and Cohn (1999), such that

$$\rho = \begin{cases} -\frac{1}{4}\left(\frac{|z|}{c}\right)^5 + \frac{1}{2}\left(\frac{|z|}{c}\right)^4 + \frac{5}{8}\left(\frac{|z|}{c}\right)^3 - \frac{5}{3}\left(\frac{|z|}{c}\right)^2 + 1, & 0 \le |z| < c \\ \frac{1}{12}\left(\frac{|z|}{c}\right)^5 - \frac{1}{2}\left(\frac{|z|}{c}\right)^4 + \frac{5}{8}\left(\frac{|z|}{c}\right)^3 + \frac{5}{3}\left(\frac{|z|}{c}\right)^2 - 5\left(\frac{|z|}{c}\right) + 4 - \frac{2}{3}\left(\frac{c}{|z|}\right), & c \le |z| < 2c \\ 0, & 2c \le |z|. \end{cases}$$

Here z is the Euclidean distance between either of the grid points in physical space. A length scale c is defined such that beyond this the correlation reduces from 1 and at a distance of more than twice the correlation length scale the correlation reduces to zero.

Covariance Localization



Covariance Localization

In the EnKF the ensemble Kalman gain is given by

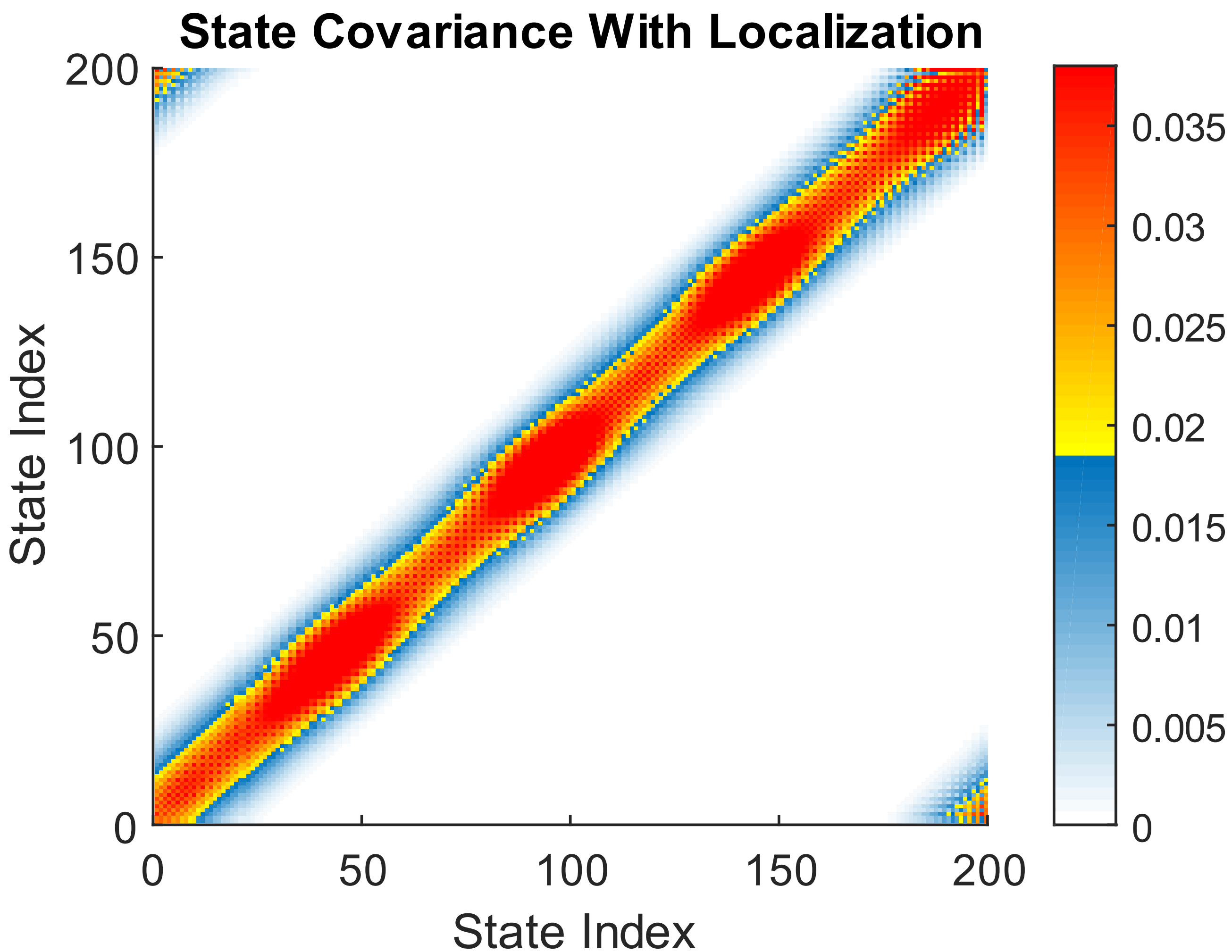
$$\widehat{\boldsymbol{K}}_t = \widehat{\boldsymbol{P}}_t^f \boldsymbol{H}_t'(\boldsymbol{H}_t \widehat{\boldsymbol{P}}_t^f \boldsymbol{H}_t' + \boldsymbol{R}_t)^{-1}$$

The forecast error covariance appears twice within this equation and strictly speaking the Schur product should be taken with each of these occurrences such that we have

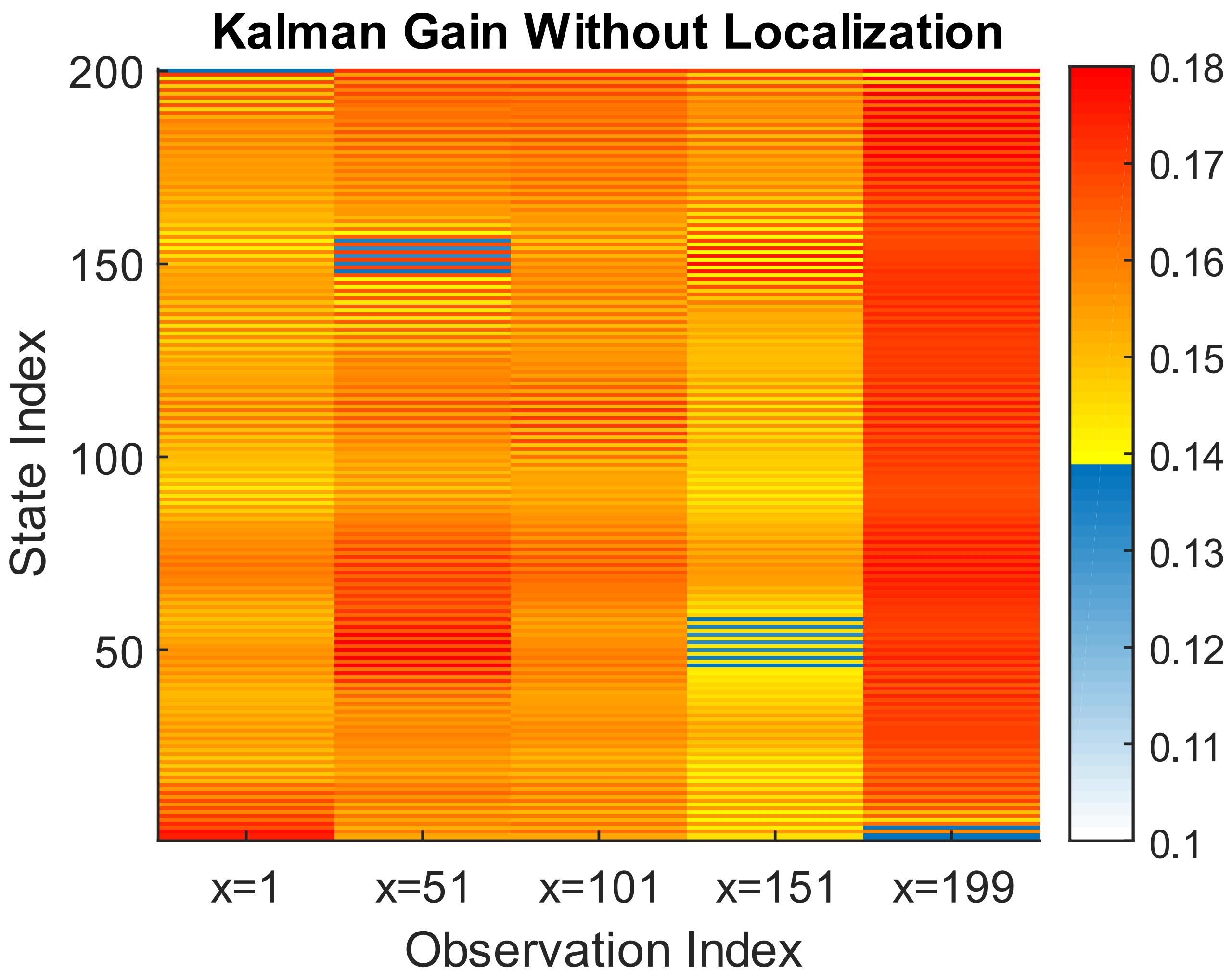
$$\widehat{\boldsymbol{K}}_t = (\rho \, o \, \widehat{\boldsymbol{P}}_t^f)\boldsymbol{H}_t'[\boldsymbol{H}_t(\rho \, o \, \widehat{\boldsymbol{P}}_t^f)\boldsymbol{H}_t' + \boldsymbol{R}_t]^{-1}$$

Since ρ is a covariance matrix and $\widehat{\boldsymbol{P}}_t^f$ is a covariance matrix then it can be proved that $(\rho \, o \, \widehat{\boldsymbol{P}}_t^f)$ is also a covariance matrix (Horn,1990).

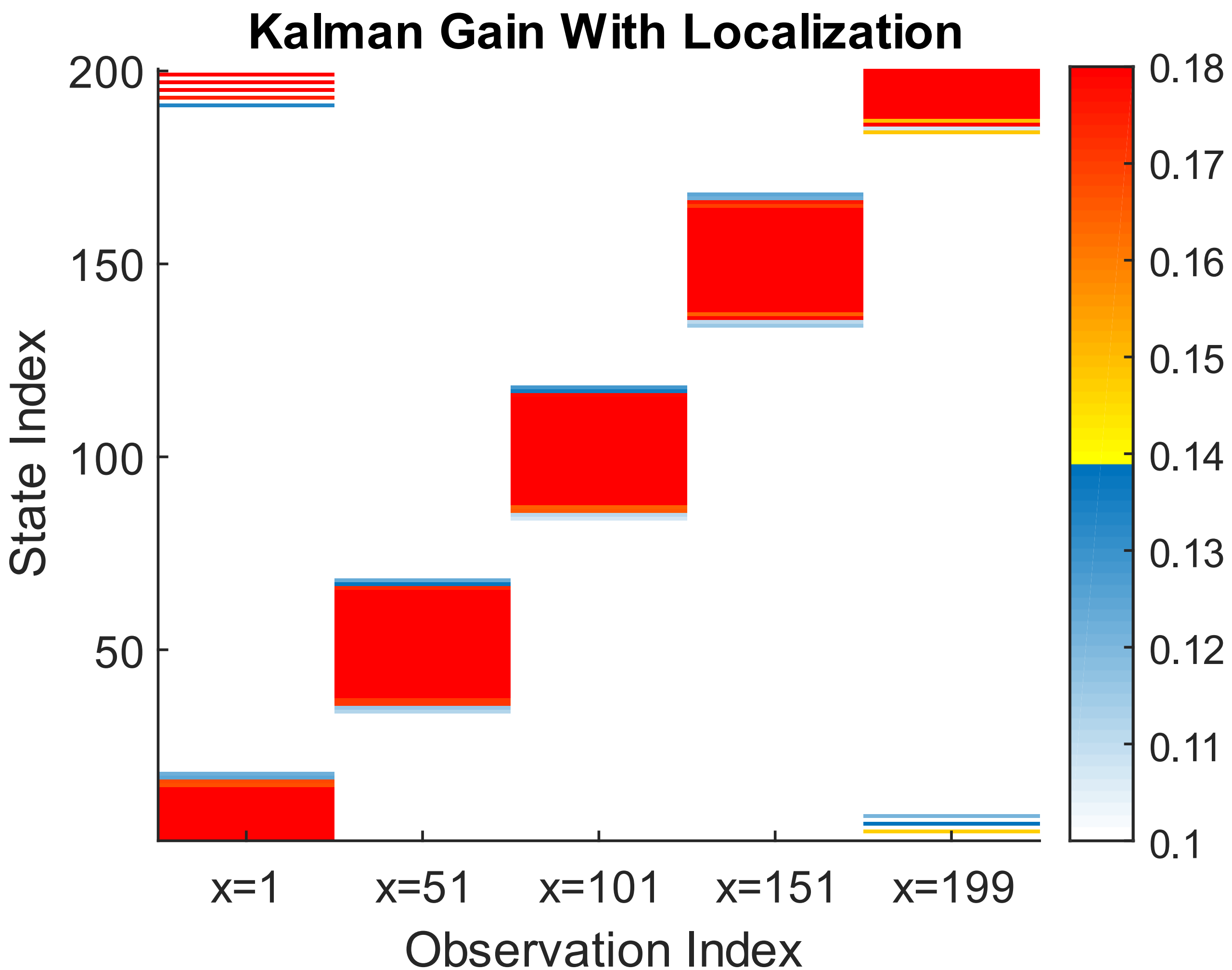
Covariance Localization



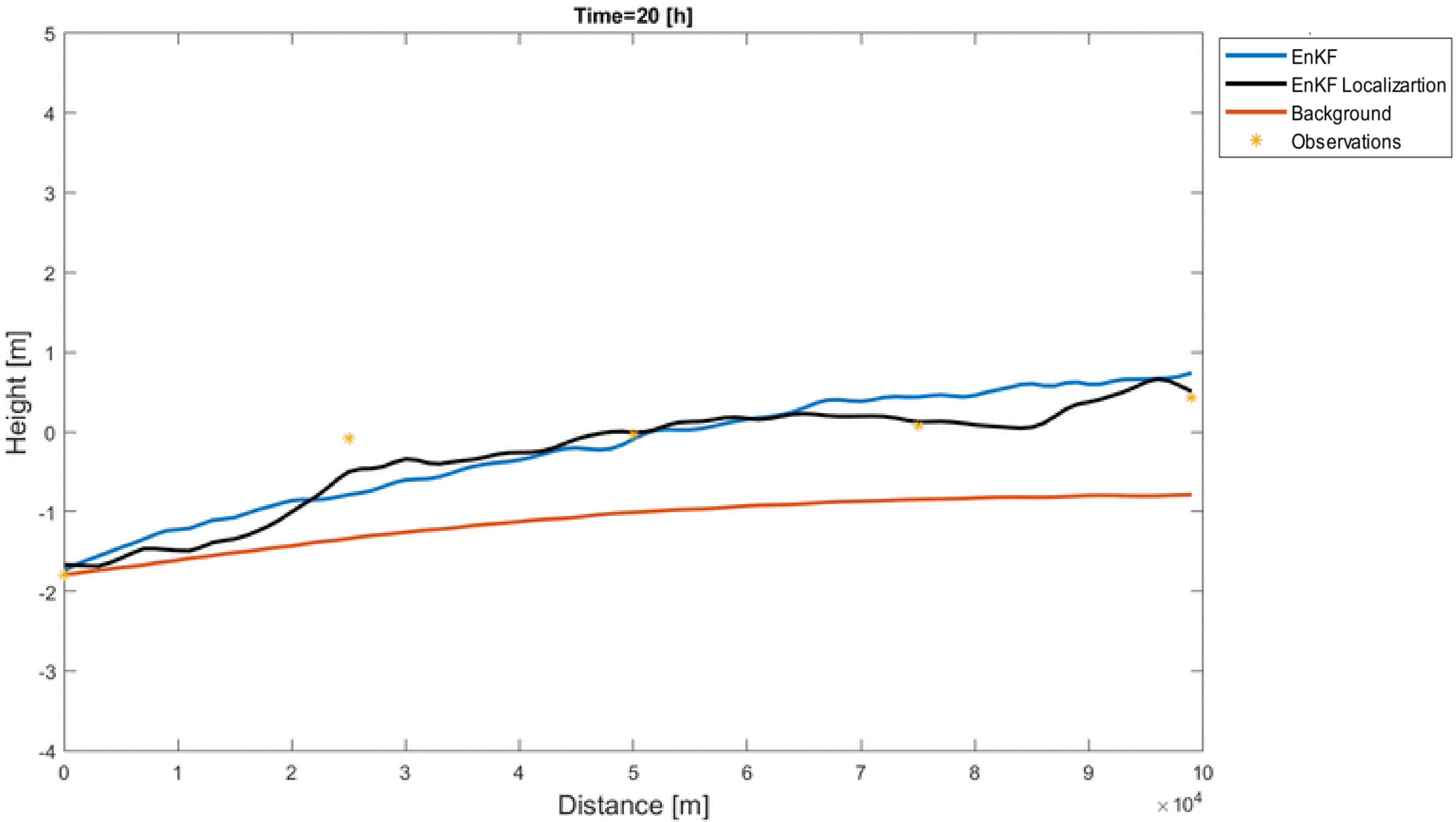
Covariance Localization



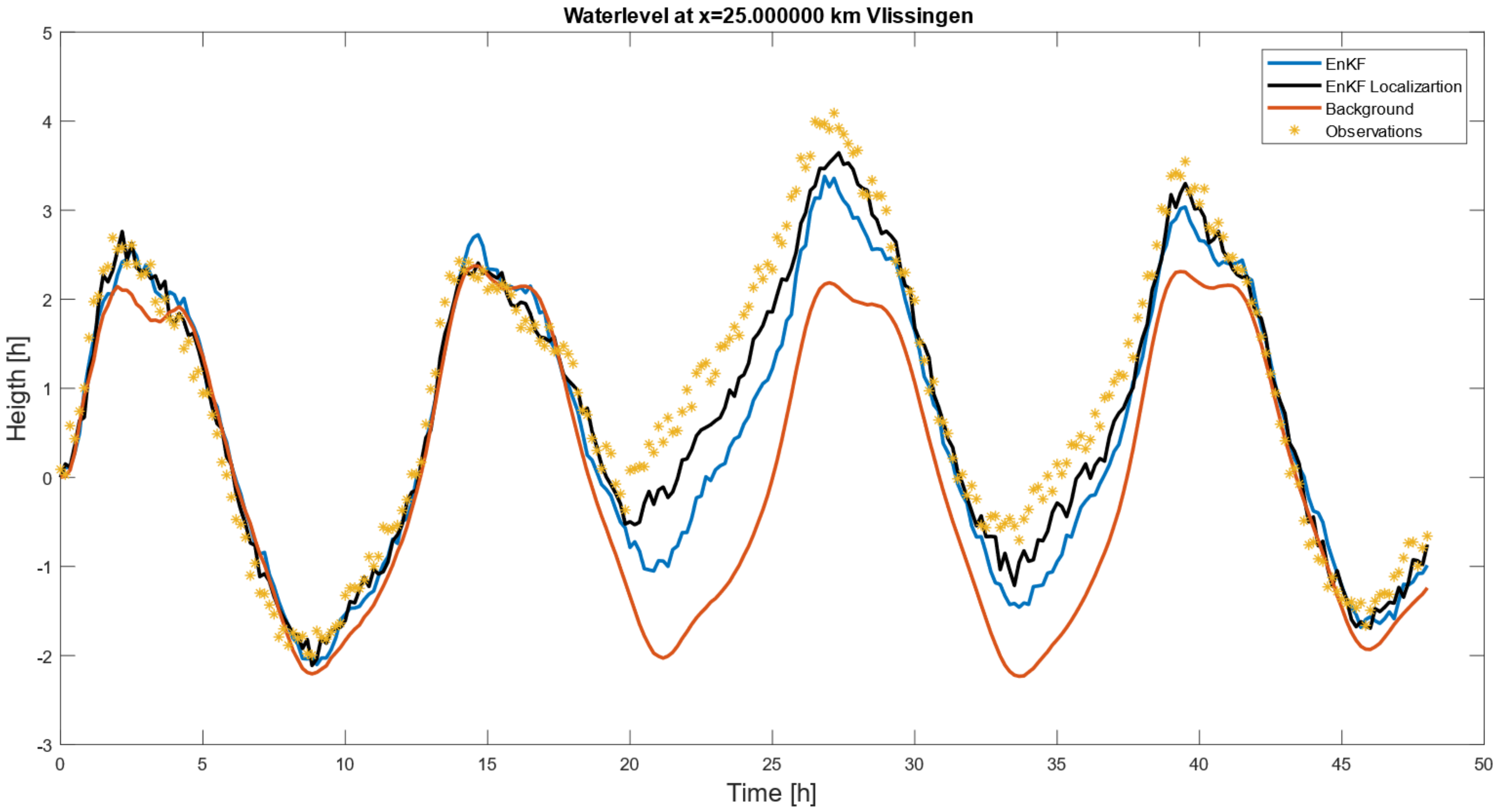
Covariance Localization



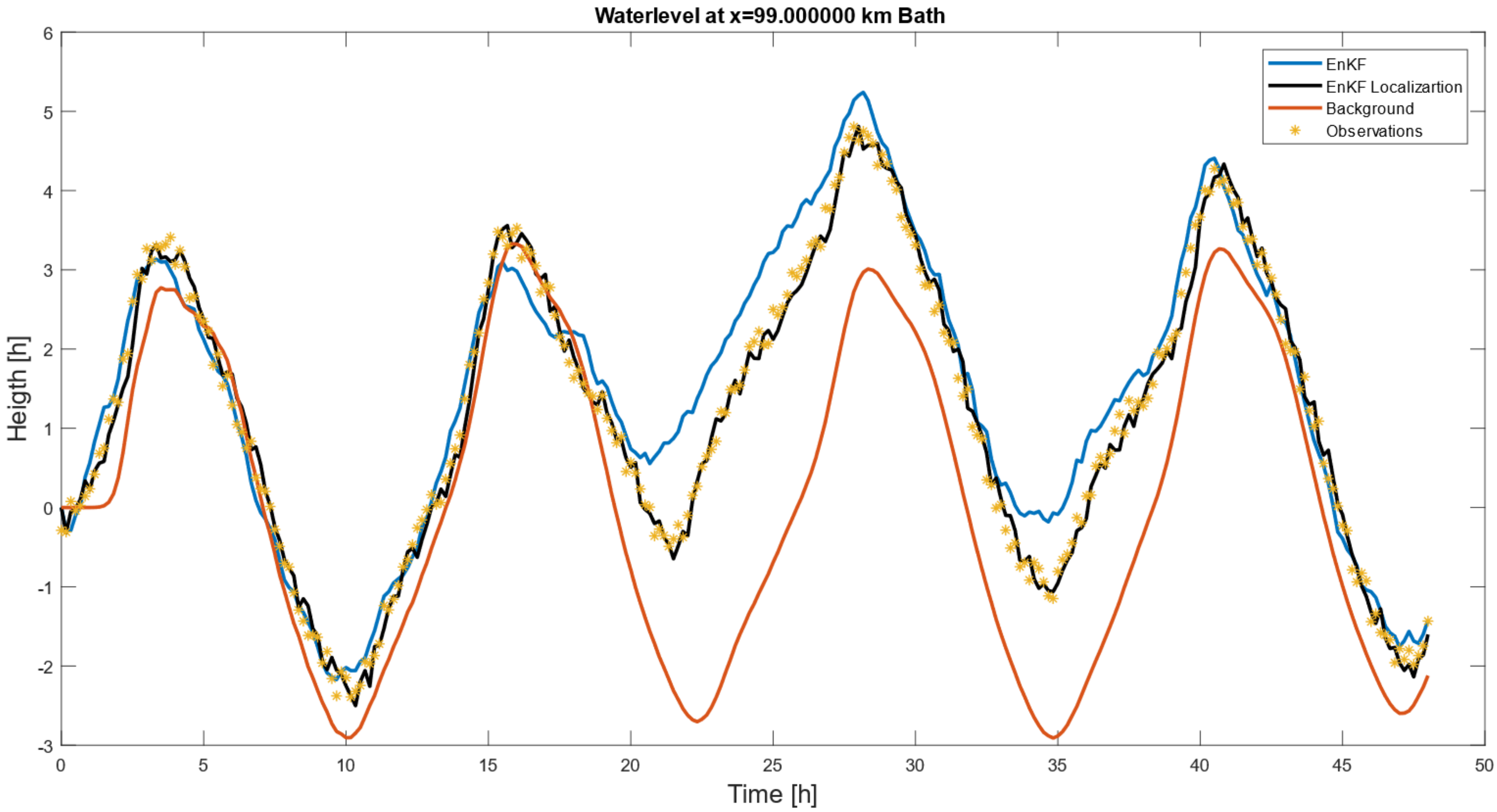
Covariance Localization



Covariance Localization



Covariance Localization



Parameter Estimation

Parameter Estimation

In applications, it is necessary estimate unknow parameters

In the EnKF

Parameters only appears implicit in the model operator

Include the parameters in the state vector

This new vector state is called augmented state.

Parameter Estimation

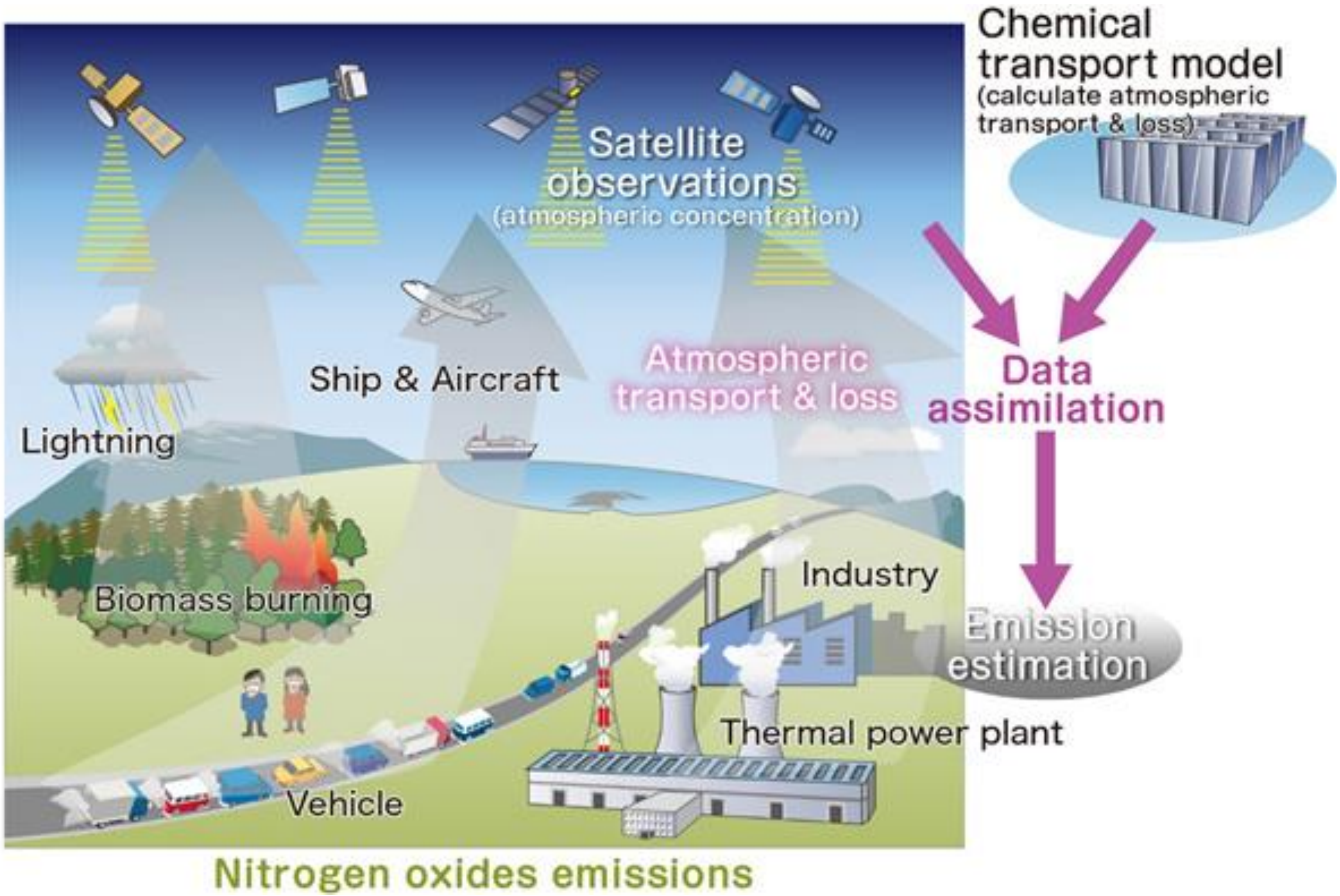
$$[\boldsymbol{x}_t] = [\boldsymbol{M}(\boldsymbol{x}_{t-1}, \boldsymbol{\theta})]$$

where $\boldsymbol{\theta}$ is the vector of unknown parameters. Since the parameters are a source of uncertainty, we can model them as a stochastic process $\Gamma(\boldsymbol{\theta})$.

$$\begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{\theta}_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{M}(\boldsymbol{x}_{t-1}, \boldsymbol{\theta}_{t-1}) \\ \Gamma(\boldsymbol{\theta}_{t-1}) \end{bmatrix}$$

In this way, the EnKF is capable of estimating the augmented state vector including the parameters.

Parameter Estimation



Parameter Estimation

The emission of PM2.5 is considered as follows:

$$\widehat{\boldsymbol{e}_t} = \boldsymbol{e}_t(1 + \boldsymbol{\delta e})$$

where $\boldsymbol{e}_t$ is the default emissions (emission inventory), $\boldsymbol{\delta e}$ is the correction factor and $\widehat{\boldsymbol{e}_t}$ is the estimated emission. With this we augmented the state vector as:

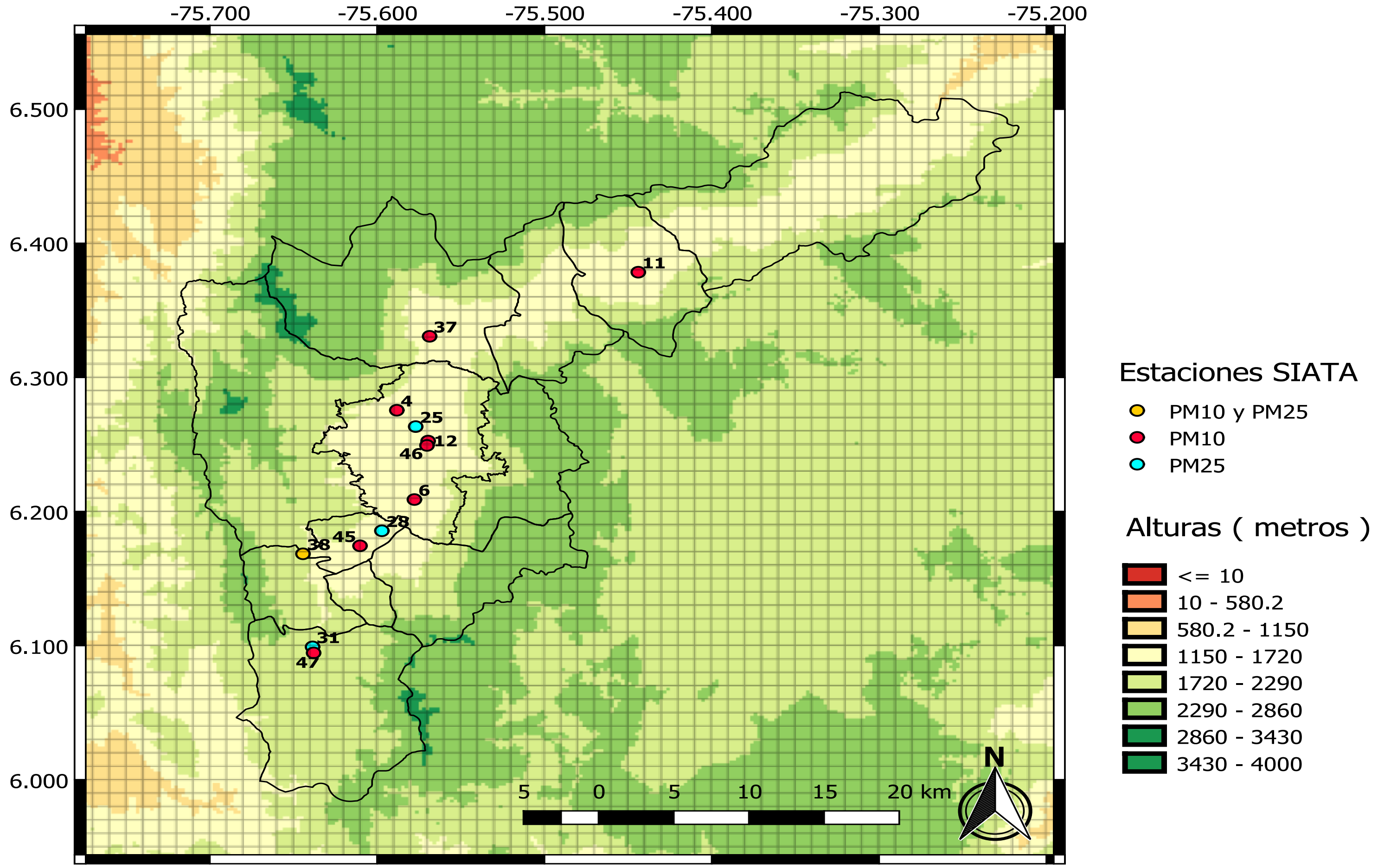
$$\begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{\delta e}_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{M}(\boldsymbol{x}_{t-1}, \boldsymbol{\theta}_{t-1}) \\ \alpha \boldsymbol{\delta e}_{t-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ \sqrt{1 - \alpha^2} \end{bmatrix} \boldsymbol{w}_k$$

The factor $\boldsymbol{\delta e}_t$ is modeled as a colored noise process, forced by a white noise $\boldsymbol{w}_k$ with a mean zero and a standard deviation of 30%. A time correlation of $\alpha = 0.95$ ensures that the samples are smoothed in time

Preliminary Results

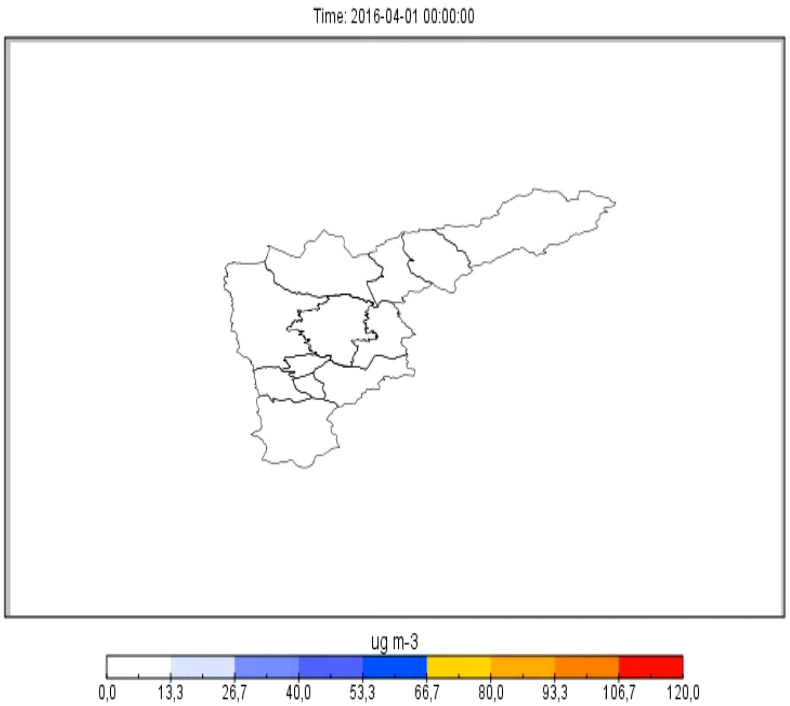
Resolution of LE	0.01°x 0.01° ≈ 1 km x 1km
Emission Inventory	EDGAR 4.0
EDGAR Inventory Resolution	0.1°x 0.1° ≈ 10 km x 10km
Species assimilated	PM10 and PM2.5
Period of simulation	April 1, 2016-April 12,2016

Preliminary Results

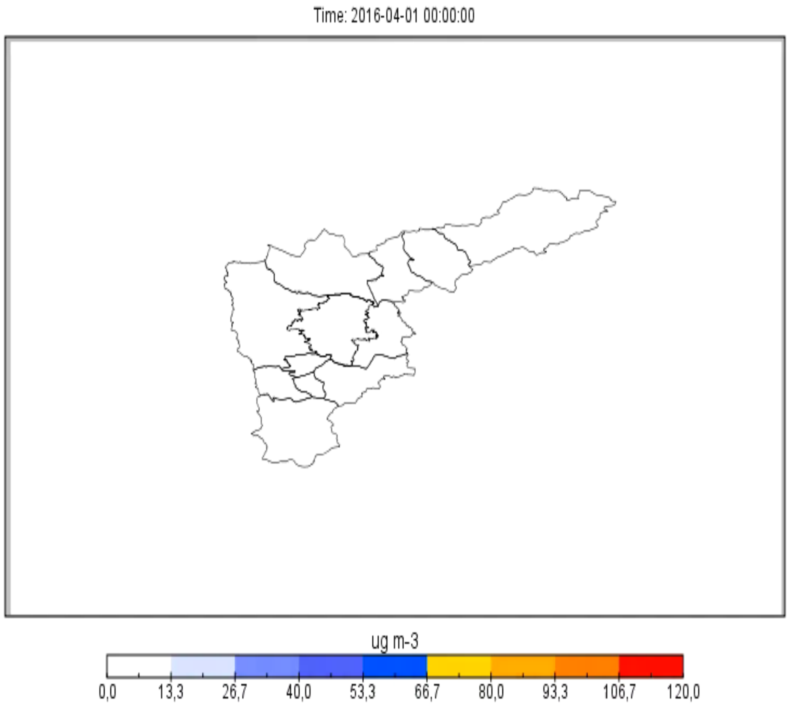


Preliminary Results

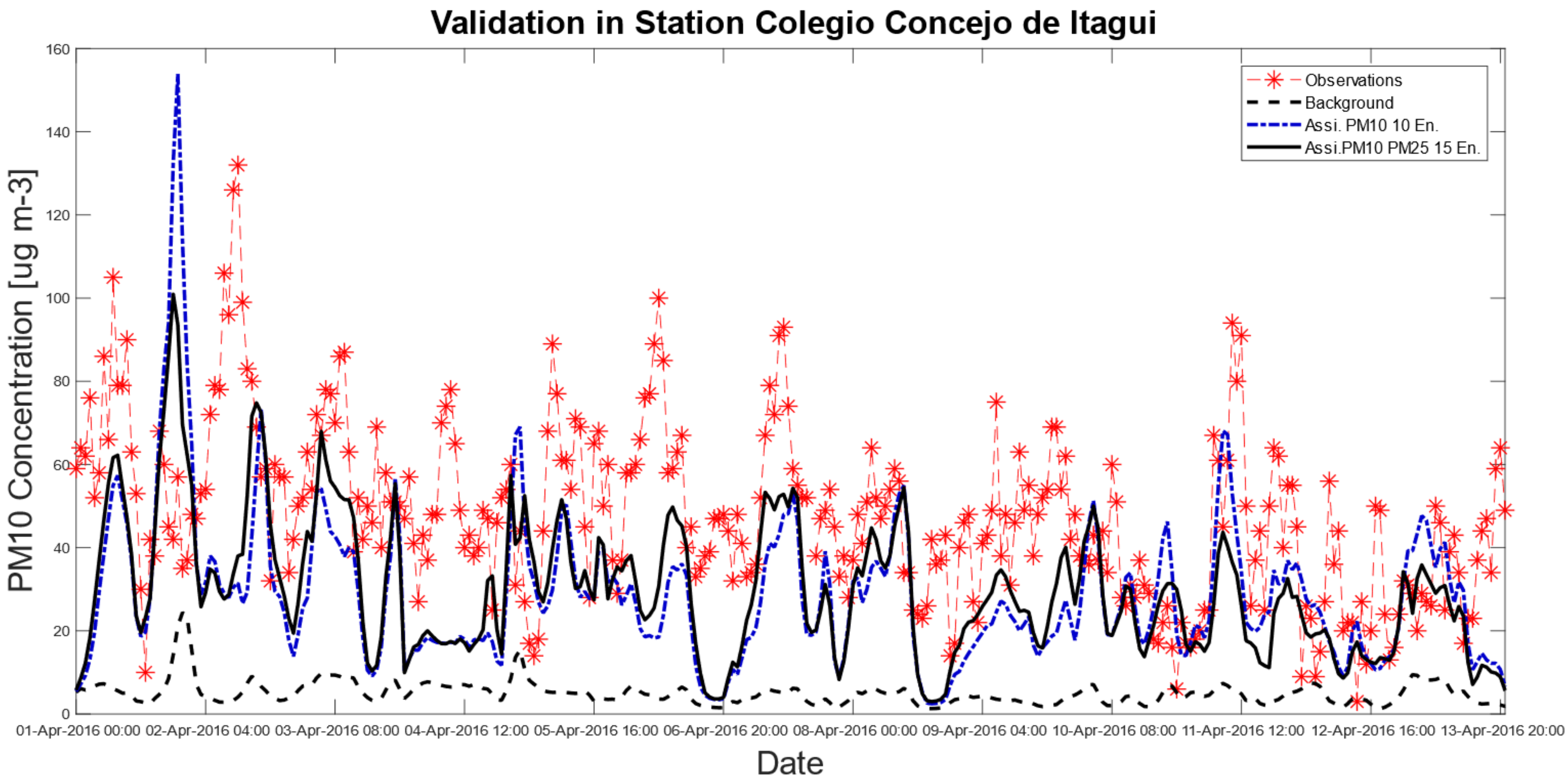
Background



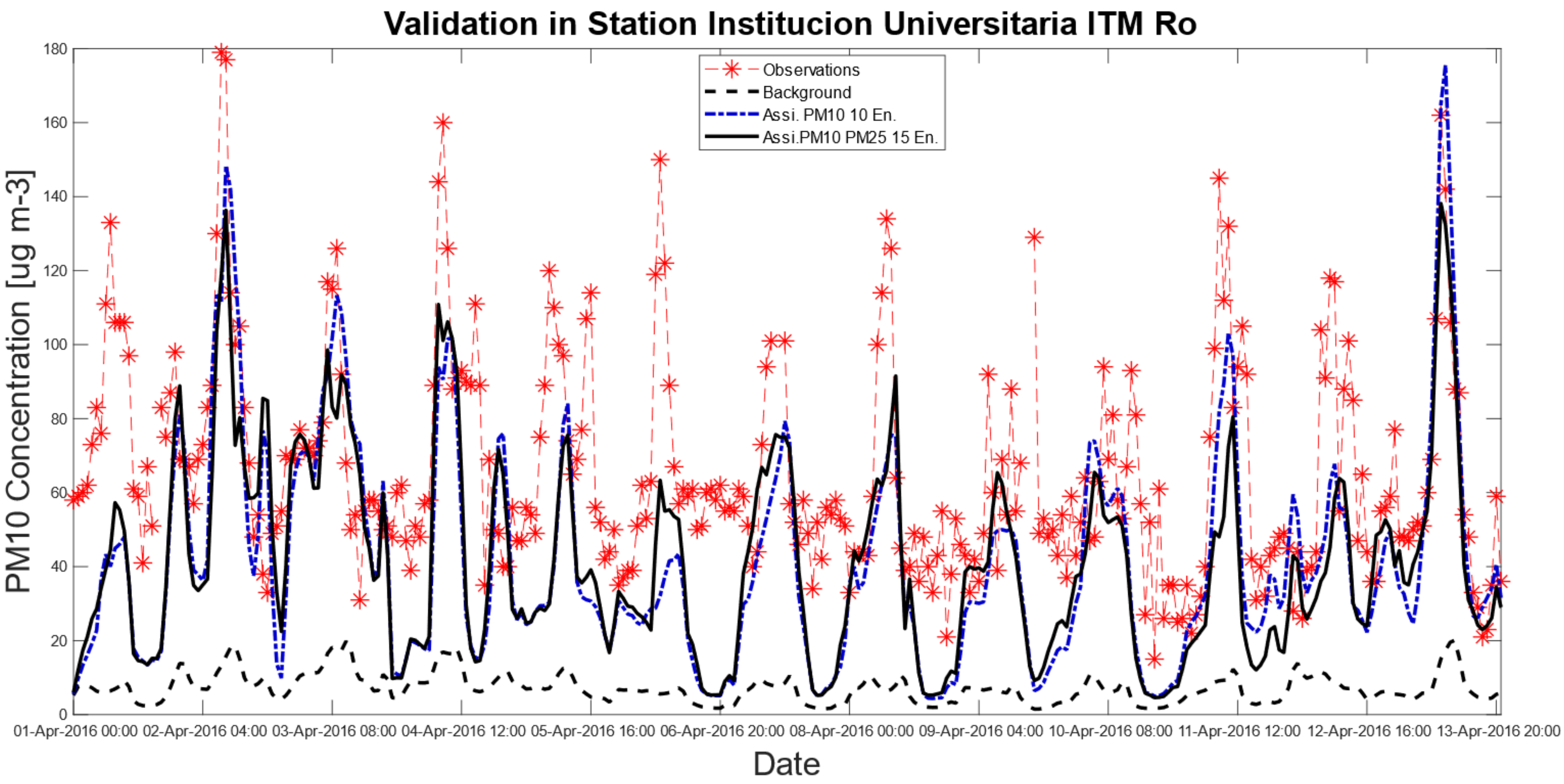
EnKF



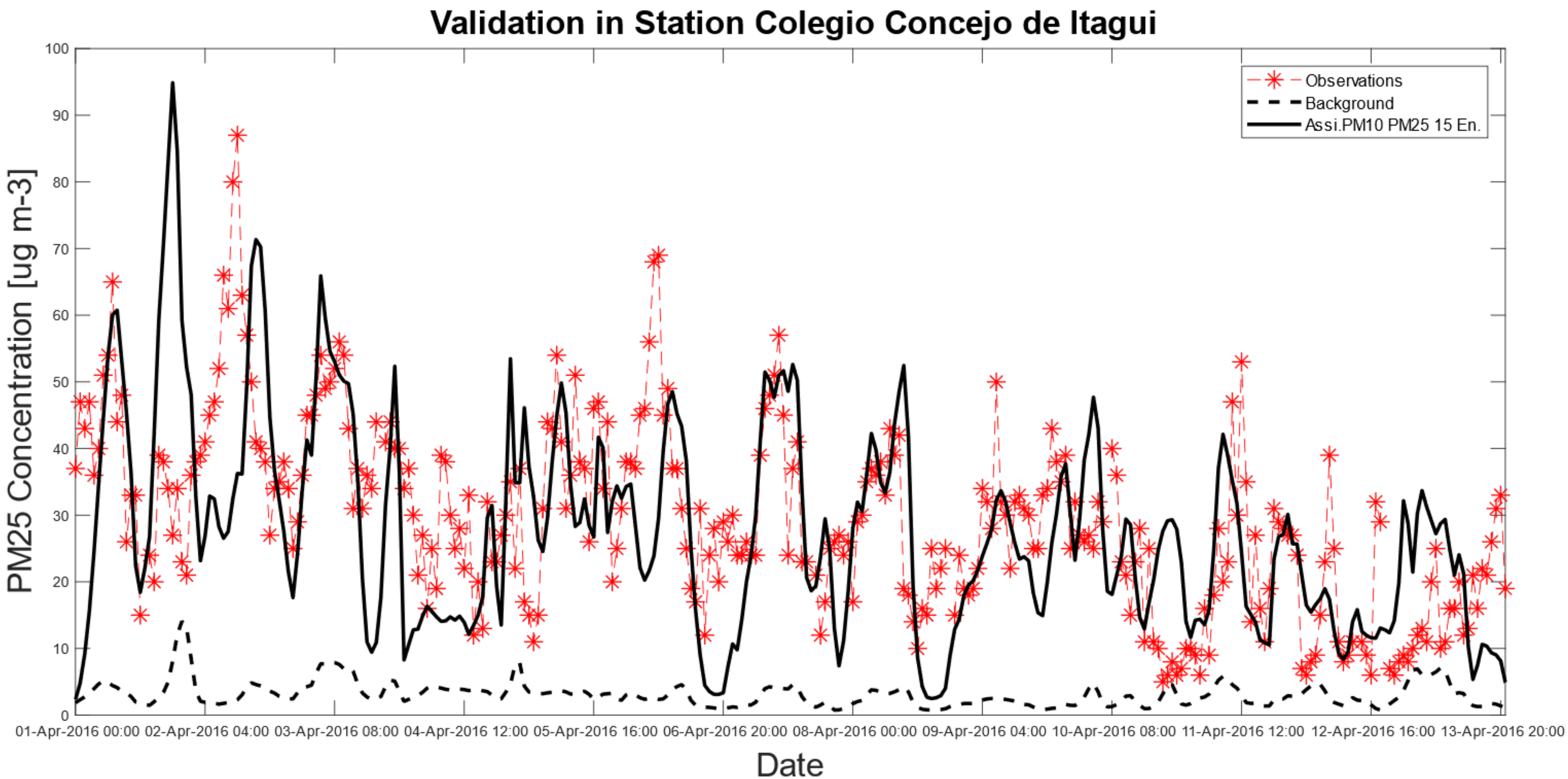
Preliminary Results



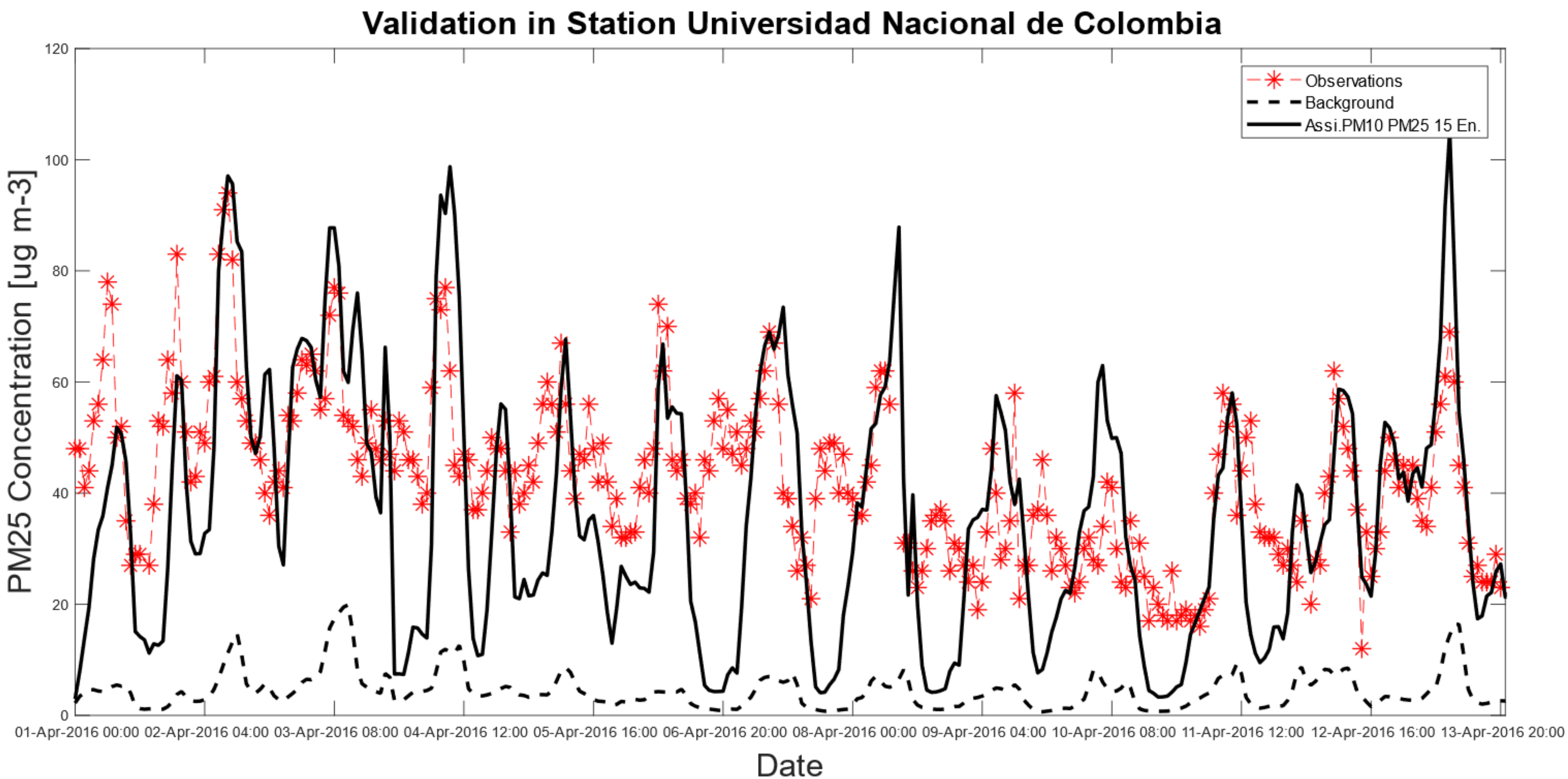
Preliminary Results



Preliminary Results

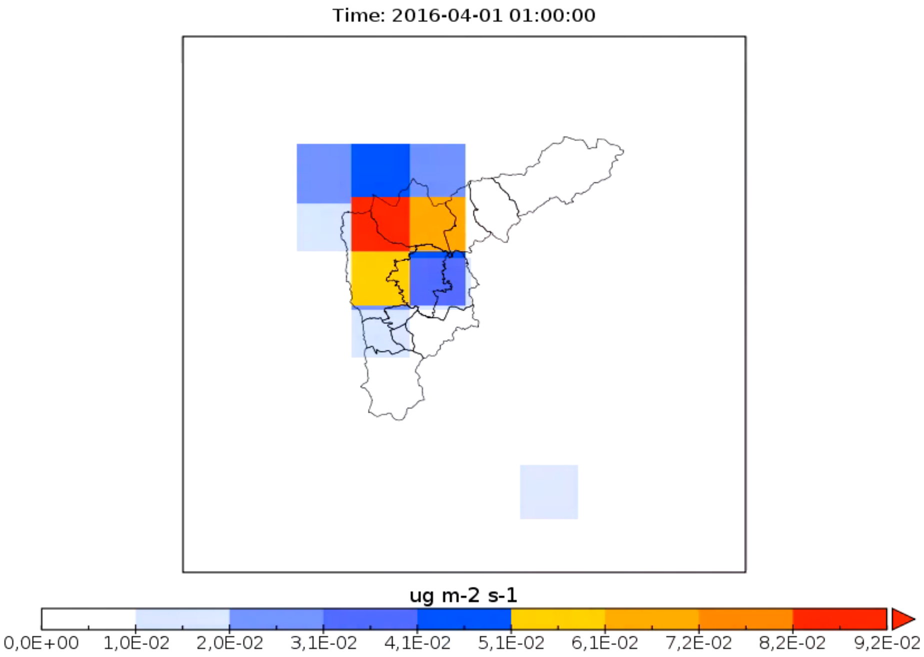


Preliminary Results

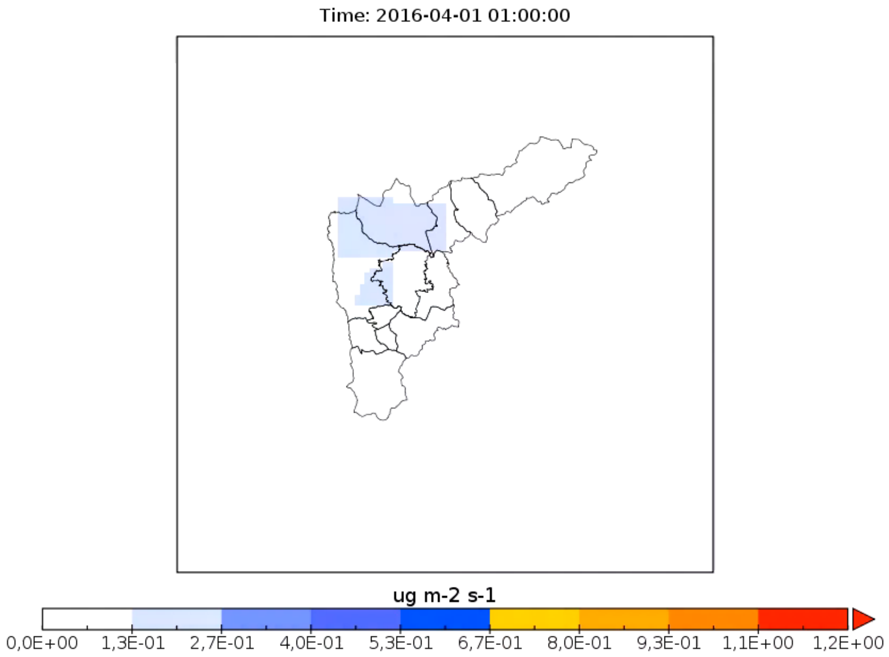


Preliminary Results

Background Emissions PM2.5



EnKF Emissions PM2.5



References

Barbu AL, Segers AJ, Schaap M, Heemink AW, Builtjes PJH. A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe. *Atmos Environ*. 2009; 43(9):1622–1631.

Fu, G., Prata, F., Lin, H. X., Heemink, A., Segers, A., and Lu, S. *Data assimilation for volcanic ash plumes using a satellite observational operator: a case study on the 2010 Eyjafjallajökull volcanic eruption* Atmos. Chem. Phys., 2017, 17, 1187-1205, doi:10.5194/acp-17-1187-2017,

Fu, G., Heemink, A., Lu, S., Segers, A., Weber, K., and Lin, H.-X. *Model-based aviation advice on distal volcanic ash clouds by assimilating aircraft in situ measurements* Atmos. Chem. Phys., 2016,16, 9189-9200, doi:10.5194/acp-16-9189-2016.

Green J, Sánchez S. Air Quality in Latin America: An Overview. Clean air Institute. Washington D.C., USA. 2012.

Hendriks C, Kranenburg R, Kuenen J, van Gijlswijk R, Kruit RW, Segers A, van der Gon HD, Schaap M. The origin of ambient particulate matter concentrations in the Netherlands. Atmos Environ. 2013; 69, 289–303.

References

Kumar A, Jiménez R, Belalcázar L, Rojas N. Application of WRF-Chem Model to Simulate PM10 Concentration over Bogota. *Aerosol Air Qual Res*. 2016; 16:1206–1221.

Lu, S., H. X. Lin, A. Heemink, A. Segers, and G. Fu. *Estimation of volcanic ash emissions through assimilating satellite data and ground-based observations* J. Geophys. Res. Atmos., 2016, 121, 10,971-10,994, doi:10.1002/2016JD025131

Lu, S., Lin, H. X., Heemink, A. W., Fu, G., and Segers, A. J. Estimation of volcanic ash emissions using trajectory-based 4D-Var data assimilation Monthly Weather Review, 2015, e-View, doi:10.1175/MWR-D-15-0194.1

Marécal V, Peuch VH, Andersson C, Andersson S, Arteta J, Beekmann M, *et al*. A regional air quality forecasting system over Europe: The MACC-II daily ensemble production. *Geosci Model Dev*. 2015; 8(9):2777–2813.

Rendón AM, Salazar JF, Palacio CA, Wirth V, Brötz B. Effects of urbanization on the temperature inversion breakup in a mountain valley with implications on air quality. *J App Meteorol Climatol*. 2014; 53:840–858.