EnKF-based estimation of Carbon monoxide surface fluxes using simulated observations during summer 2015

Vikram Khade[1], Saroja Polavarapu[1,2], Michael Neish[1], Dylan Jones[2], Pieter Houtekamer[1] and Seung-Jong Baek[1]

[1]Environment and Climate Change Canada
[2]Department of Physics, University of Toronto, Canada

08 June 2021
16th EnKF workshop (virtual)
Norway

UNIVERSITY OF TORONTO

Environment and Climate Change Canada | Environnement et Changement climatique Canada

CSA ASC

EC-CAS (Environment Canada Carbon Assimilation System)

Estimate meteorology, GHG gases and their fluxes with Ensemble Kalman Filter (EnKF).

Today I will present the results of estimation of Carbon monoxide (CO) and its fluxes.

- Forecast model = GEM-MACH-GHG (0.9 degrees, 81 levels, model top = 0.1 mb)
- Simulated meteorological observations assimilated include radiosonde, surface, aircraft, GPS-RO and scatterometer. These are generated at their real time and locations.
- Simulated in situ observations of CO are assimilated from a hypothetical network (HYPNET).

Why use simulated observations approach ?

- The truth is known perfectly in this approach. Errors can be computed exactly.
- A crucial advantage is that a known amount (and type) of error/uncertainty can be injected into the system.
- 64 ensemble members are used.
- 65^{th} ensemble member is defined to be the truth.
- Observations are drawn from the truth trajectory.
- These simulated observations (rather than the real observations are assimilated).

EC-CAS building blocks

- GEM-MACH *(Moran et. al. 2010)* at ECCC has been developed over many years.
- GEM-MACH-GHG has been recently developed and tested for GHG forecasting *(Polavarapu et. al. 2016)*.
- EnKF at ECCC has been developed over many years *(Houtekamer et. al. 2005)*.
- ECCC is the first to develop an operational ensemble based assimilation and forecasting system for meteorology.
- EC-CAS is being developed as the extension to estimate and forecast GHG and their fluxes.

What is flux estimation ?

- Flux estimation is basically parameter estimation. Typically parameters are not observed.
- Flux values are usually prescribed in the forecast model.
- However if one assumes flux values as uncertain then to estimate them the flux field is treated as a vector to be estimated.
- In this work, forecast model for flux is persistence. This is the simplest possible model.

Augmented analysis state

ECH_4 = Methane flux

ECO = Carbon Monoxide flux

ECO_2 = Carbon dioxide flux

CH_4 = Methane

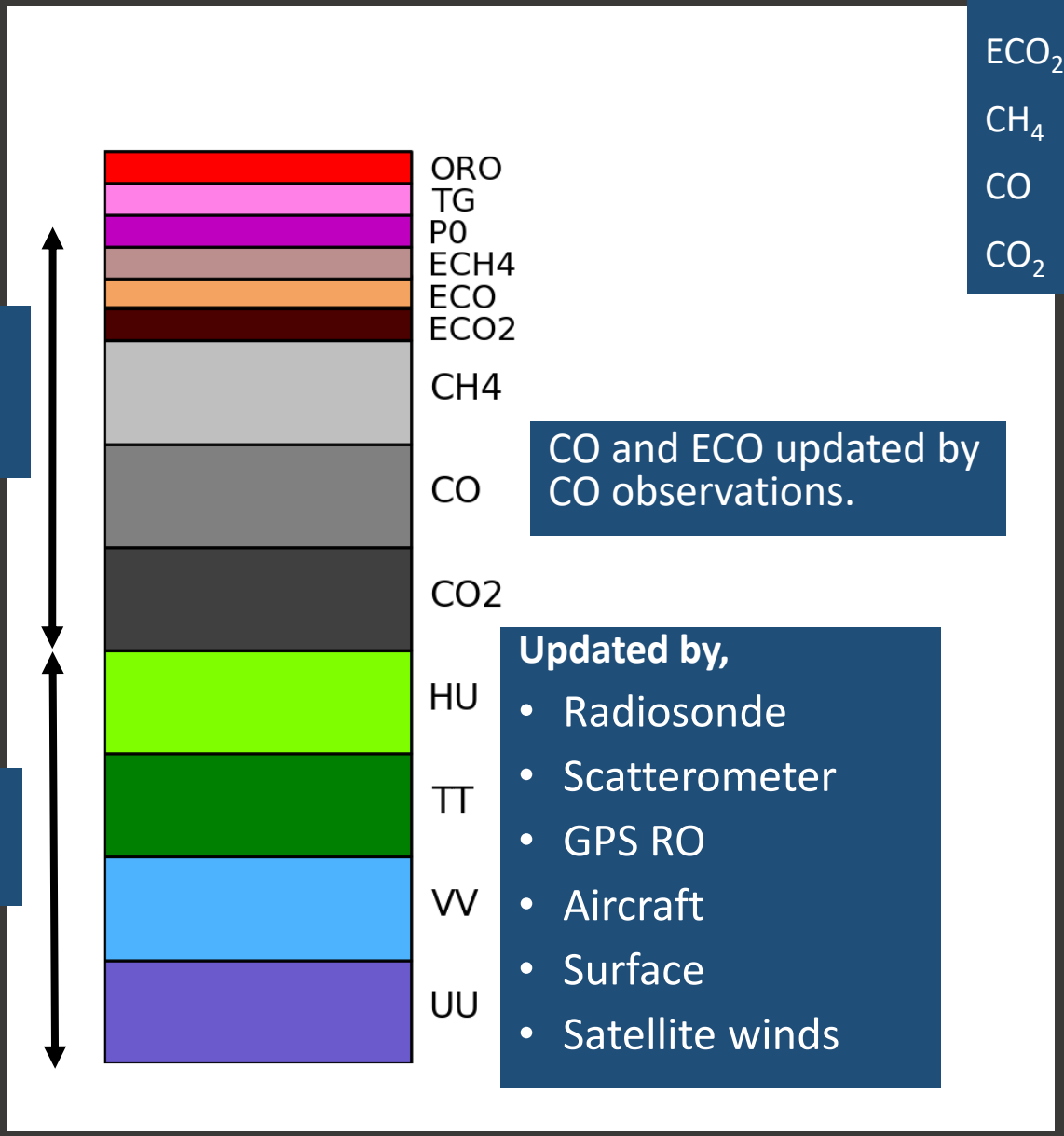
CO = Carbon Monoxide

CO_2 = Carbon dioxide

Augmented GHG state

Correlation is set to ZERO (Variable localization)

Meteorological state



CO and ECO updated by CO observations.

Updated by,

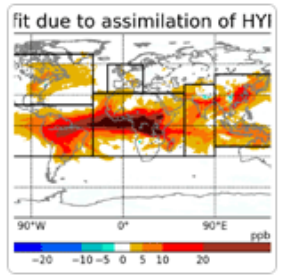
- Radiosonde
- Scatterometer
- GPS RO
- Aircraft
- Surface
- Satellite winds

EC-CAS 1.0

Geosci. Model Dev., 14, 2525–2544, 2021
https://doi.org/10.5194/gmd-14-2525-2021

Development and technical paper

06 May 2021

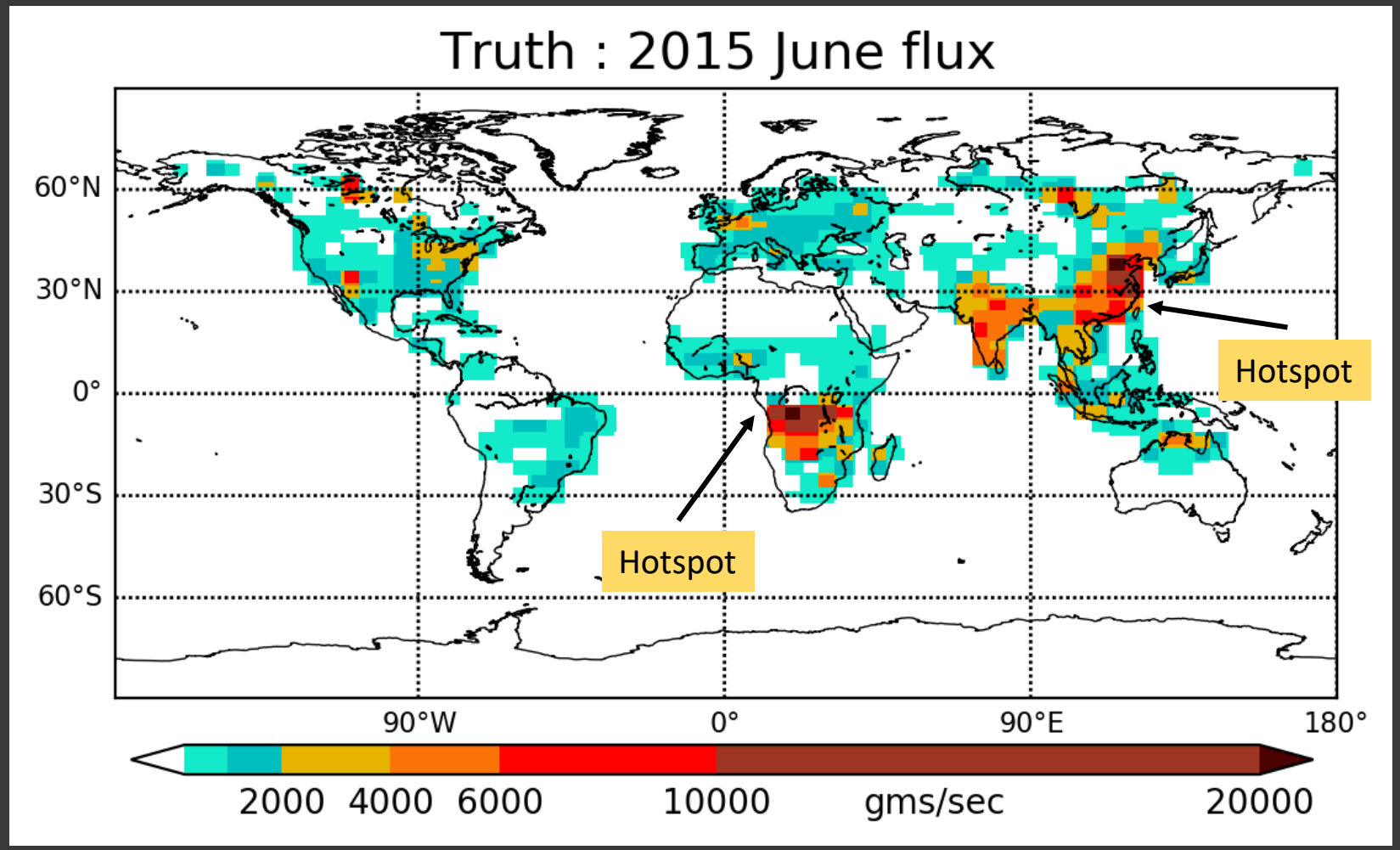


The Environment and Climate Change Canada Carbon Assimilation System (EC-CAS v1.0): demonstration with simulated CO observations

Vikram Khade[1,2], Saroja M. Polavarapu[1,2], Michael Neish[1], Pieter L. Houtekamer[3], Dylan B. A. Jones[2], Seung-Jong Baek[3], Tai-Long He[2], and Sylvie Gravel[3]

In this paper results of estimation of CO state and meteorology are presented. The flux was assumed to be uncertain but was not estimated.

Today I will present the extension of this work which includes estimation of the CO flux.



This 2d truth field is assumed to be constant with time.

In reality flux changes on timescales from hours to months.

Can the Ensemble Kalman Filter (EnKF) recover the true flux at each grid point at the surface starting from an incorrect estimate of truth by assimilating observations of Carbon Monoxide ?

(Flux is not observed.)

Ensemble Kalman Filter (EnKF)

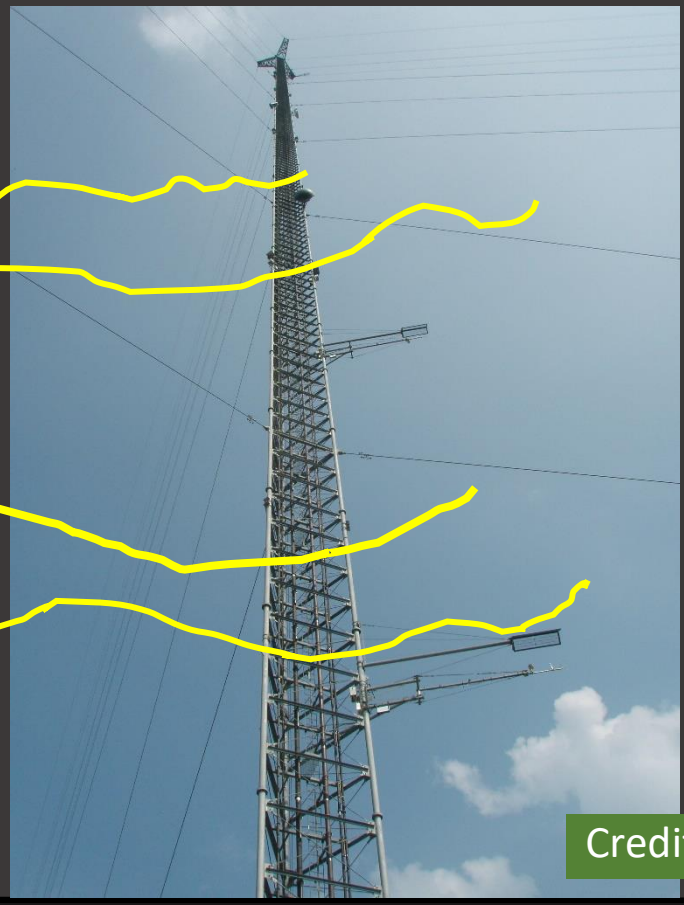
$$\mathbf{x}^a(t) = \mathbf{x}^f(t) + \mathbf{P}^f\mathbf{H}^T \frac{(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f)}{\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}}$$

- In an EnKF the covariance is calculated by using the ensemble members (sample).
- This estimate is state dependent.
- GEM-MACH-GHG state is ~ 10^8.
- The analysis uncertainty is quantified by the analysis covariance matrix.

Flux estimation

Credit :
E360.yale.edu
David Biello

Winds



Credit : *NOAA ESRL*

Source
Past time

Observing location
Current time (at DA)

Time required for signal to reach instrument =

Distance / (wind speed)

Flux estimation

Innovation is calculated at location of the instrument

$$\mathbf{x}^a(t) = \mathbf{x}^f(t) + \mathbf{P}^f\mathbf{H}^T\frac{(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f)}{\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}}$$

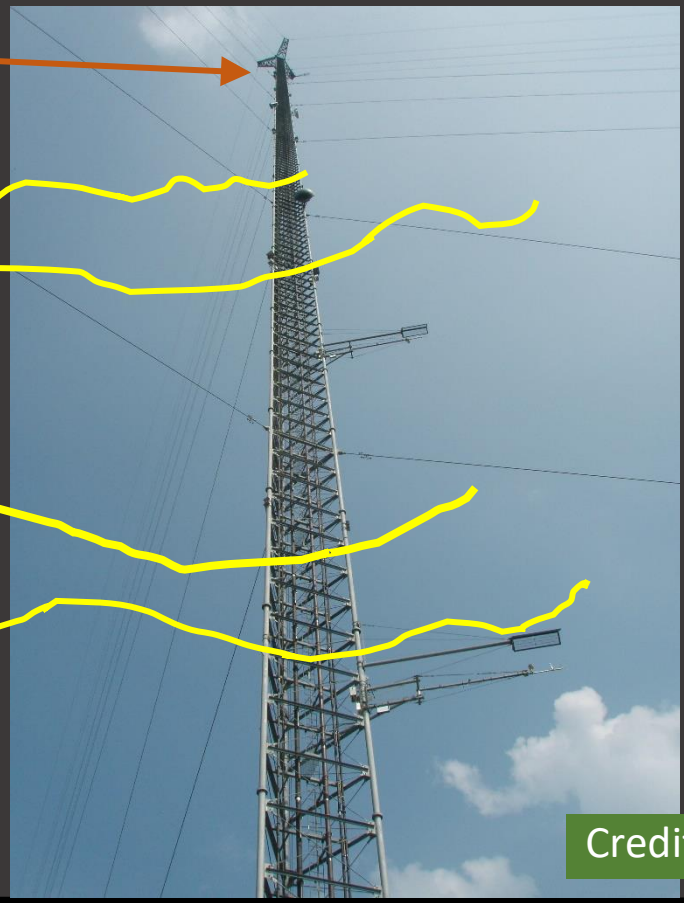
State (Flux, tracer)

is updated at the location of the source.

Correlation

Winds

Credit :
E360.yale.edu
David Biello



Credit : *NOAA ESRL*

Source
Past time

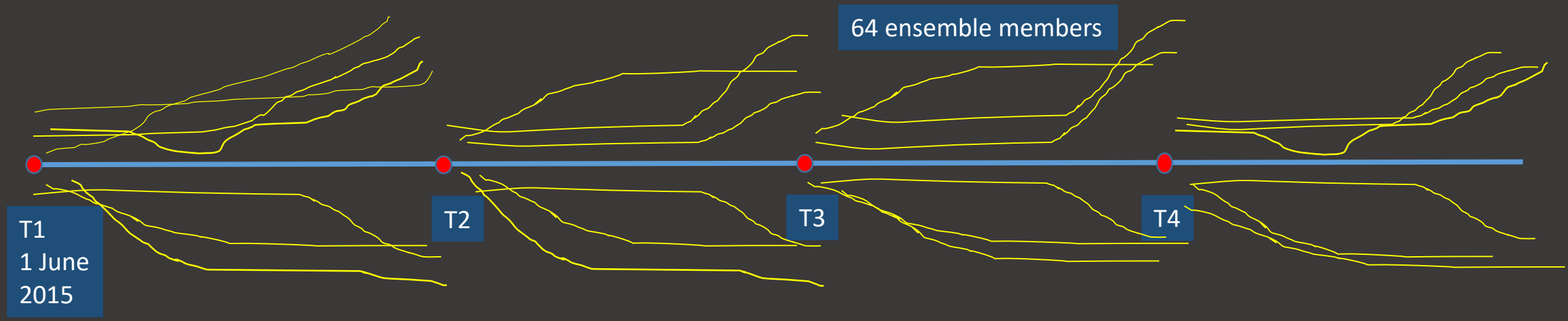
Observing location
Current time (at DA)

Time required for signal to reach instrument =

Distance / (wind speed)

Why estimate winds (and meteorology) simultaneously with CO and its flux ?

- Uncertainty in tracer (CO) depends both on error in winds and flux field.
- The CO ensemble accounts for both these uncertainties since a different realization of wind is associated with each ensemble member.
- State-dependent transport error is inherent in the ensemble !



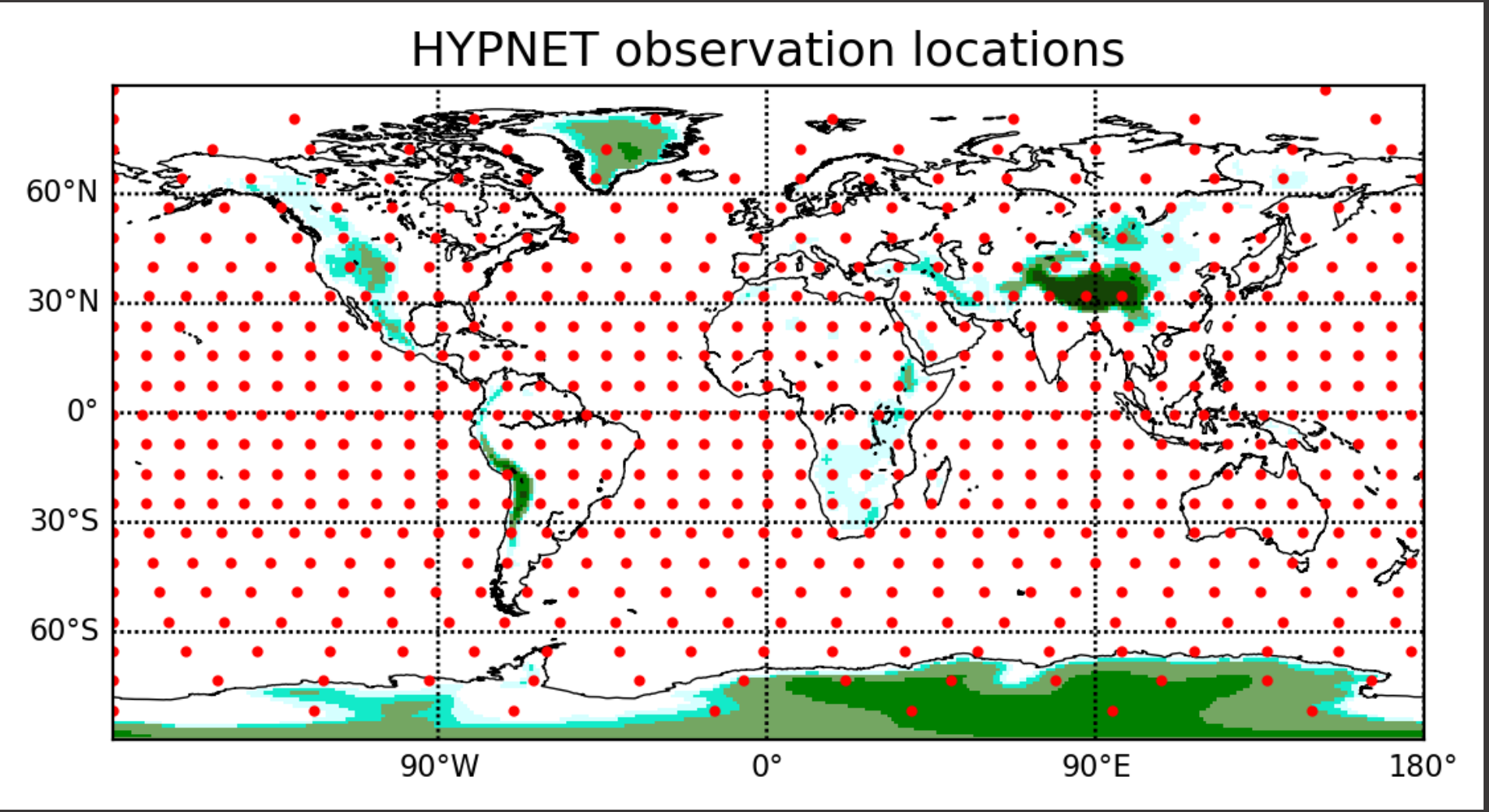
- Experiment runs from June 1 – 30, 2015.
- Spin up period is from June 1 - 10. Only meteorological observations are assimilated.
- After June 10 both meteorological and CO observations are assimilated.
- Simulated observations are assimilated every 6 hours. The DA scheme uses Incremental Analysis Update (IAU).
- Smoother approaches are more common in flux estimation. But we have used filtering.

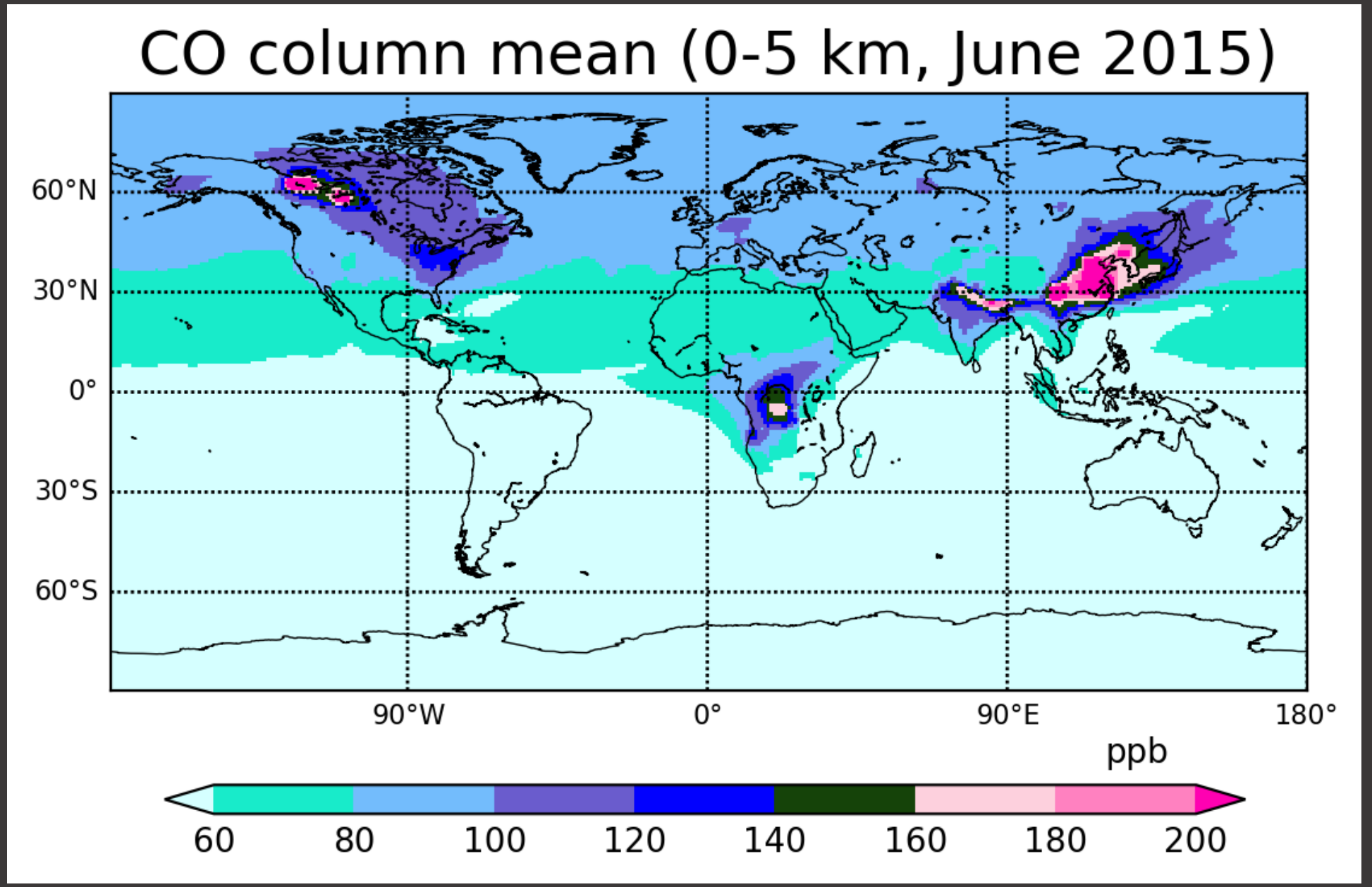
HYPNET :

CO is observed every **1000 km** in horizontal every 6 hours.

These observations exist at three planes - 400m, 2000m, 5000m

Observations error = 10%



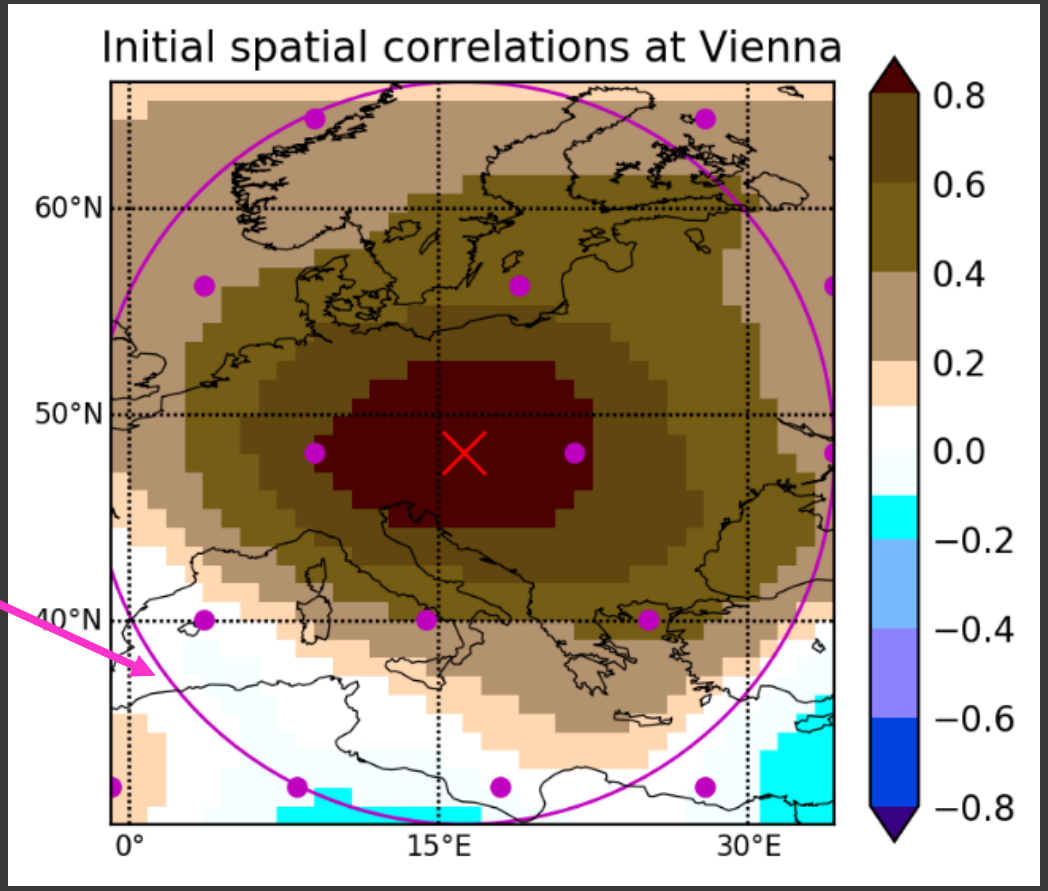


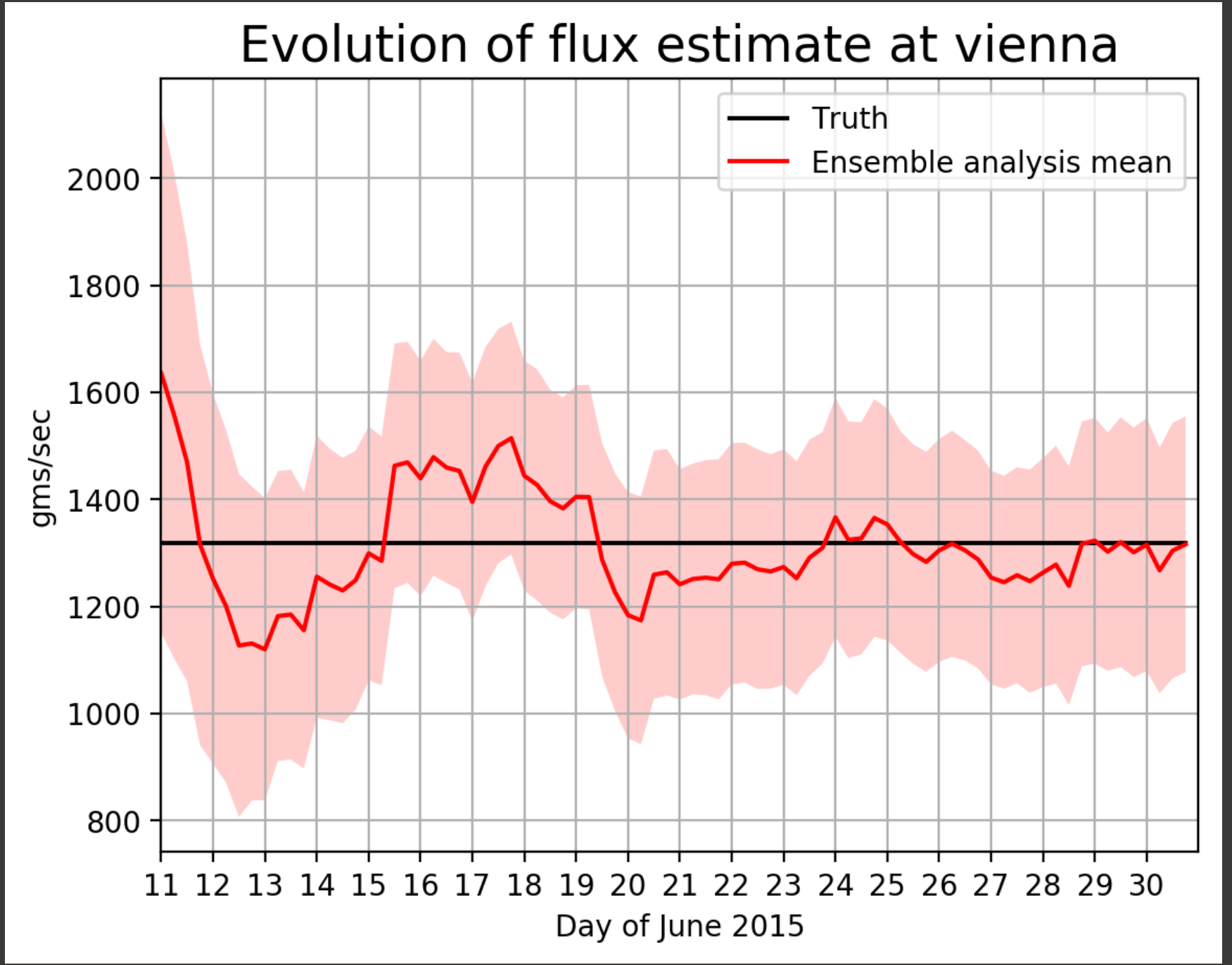
Ensemble mean averaged over column and time.

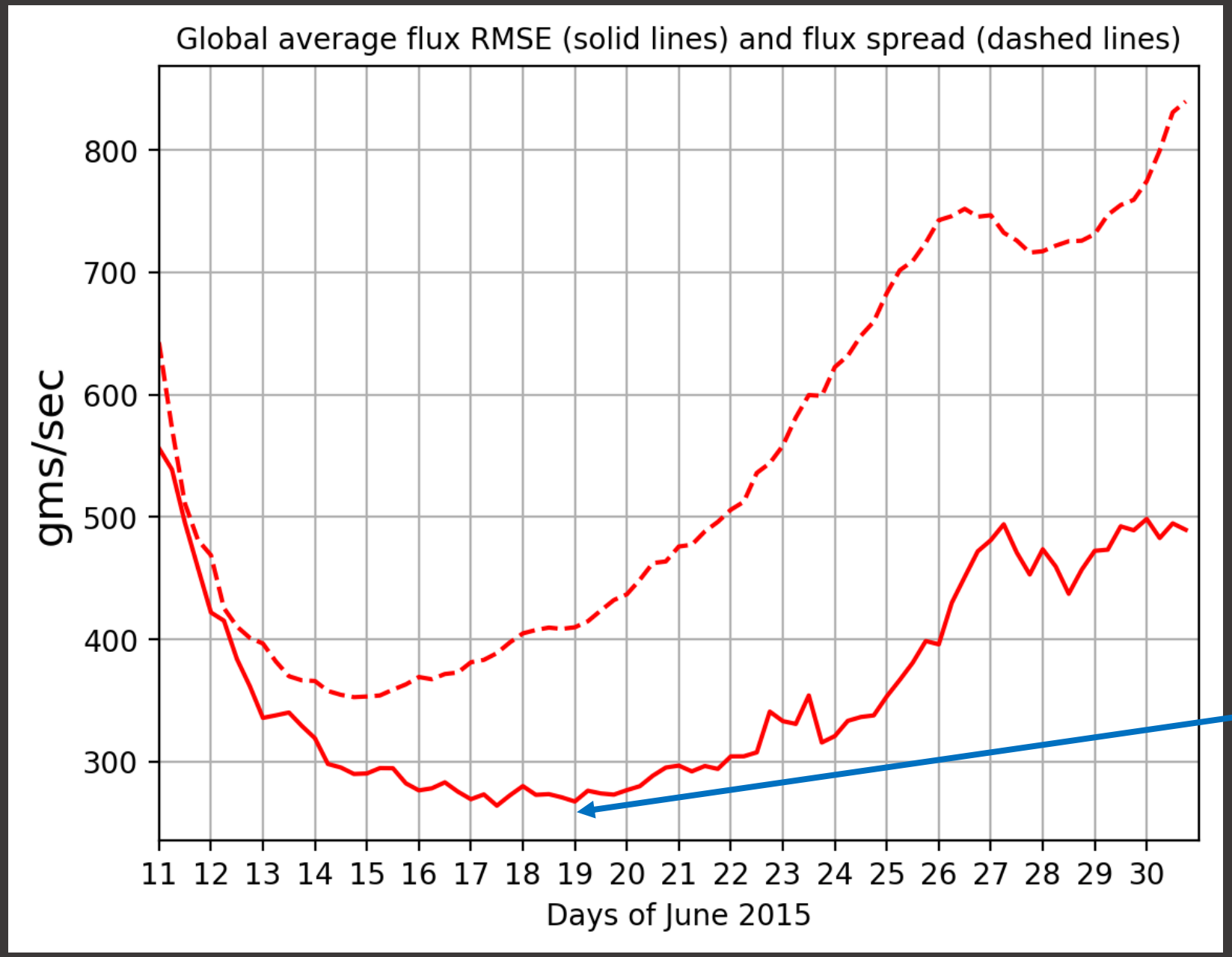
Flux ensemble generation

- Ensemble of prior flux field (on 1 June) are generated offline with an error of 30%.
- A spatial correlation of 500 km is imposed.
- This flux ensemble is updated every 6 hours starting on 10 June by observations of CO.

Localization (1000 km)

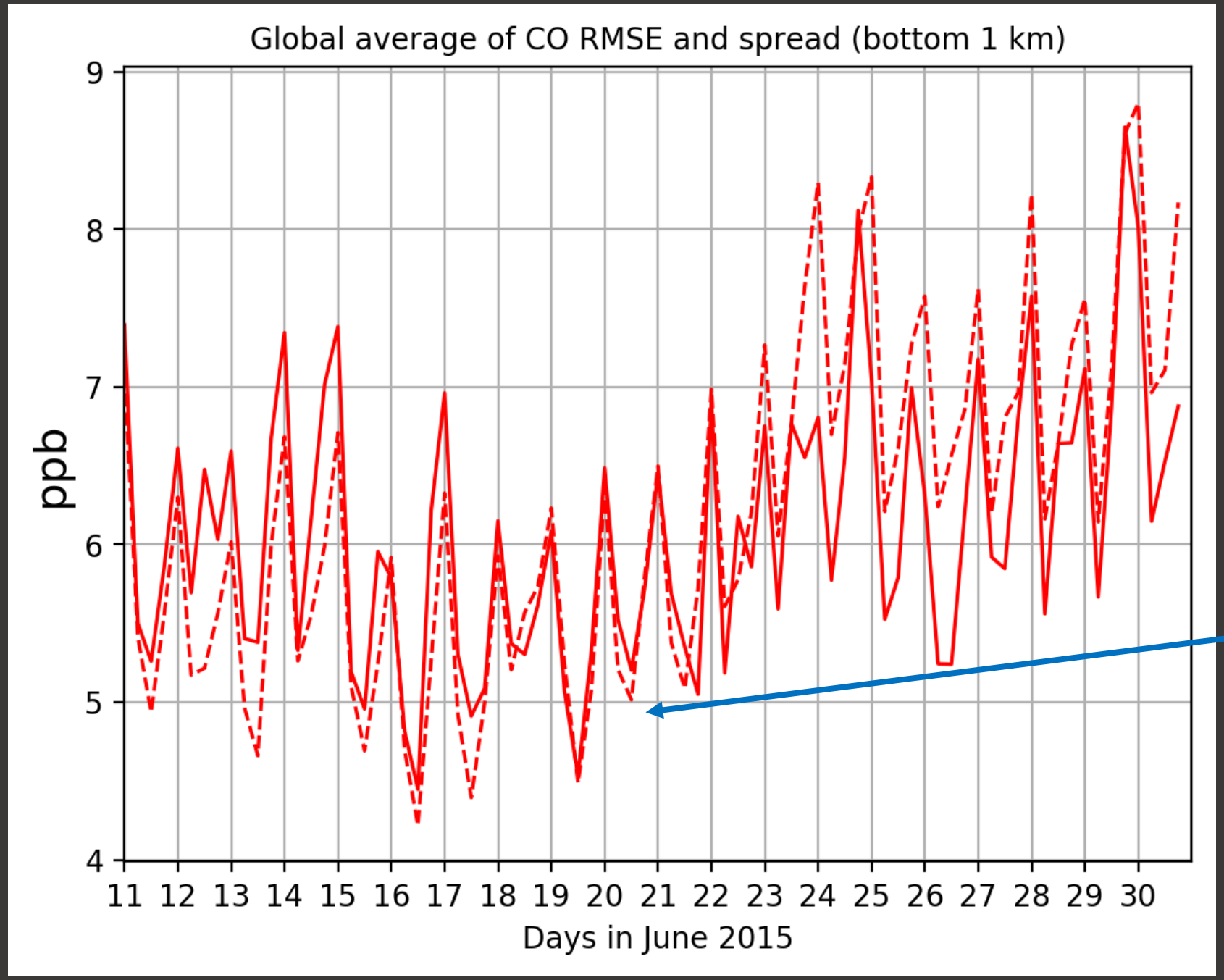



Evolution of flux estimate at vienna
Truth
Ensemble analysis mean
gms/sec
2000
1800
1600
1400
1200
1000
800
11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Day of June 2015



The RMSE hits the minimum value – but then it starts increasing !

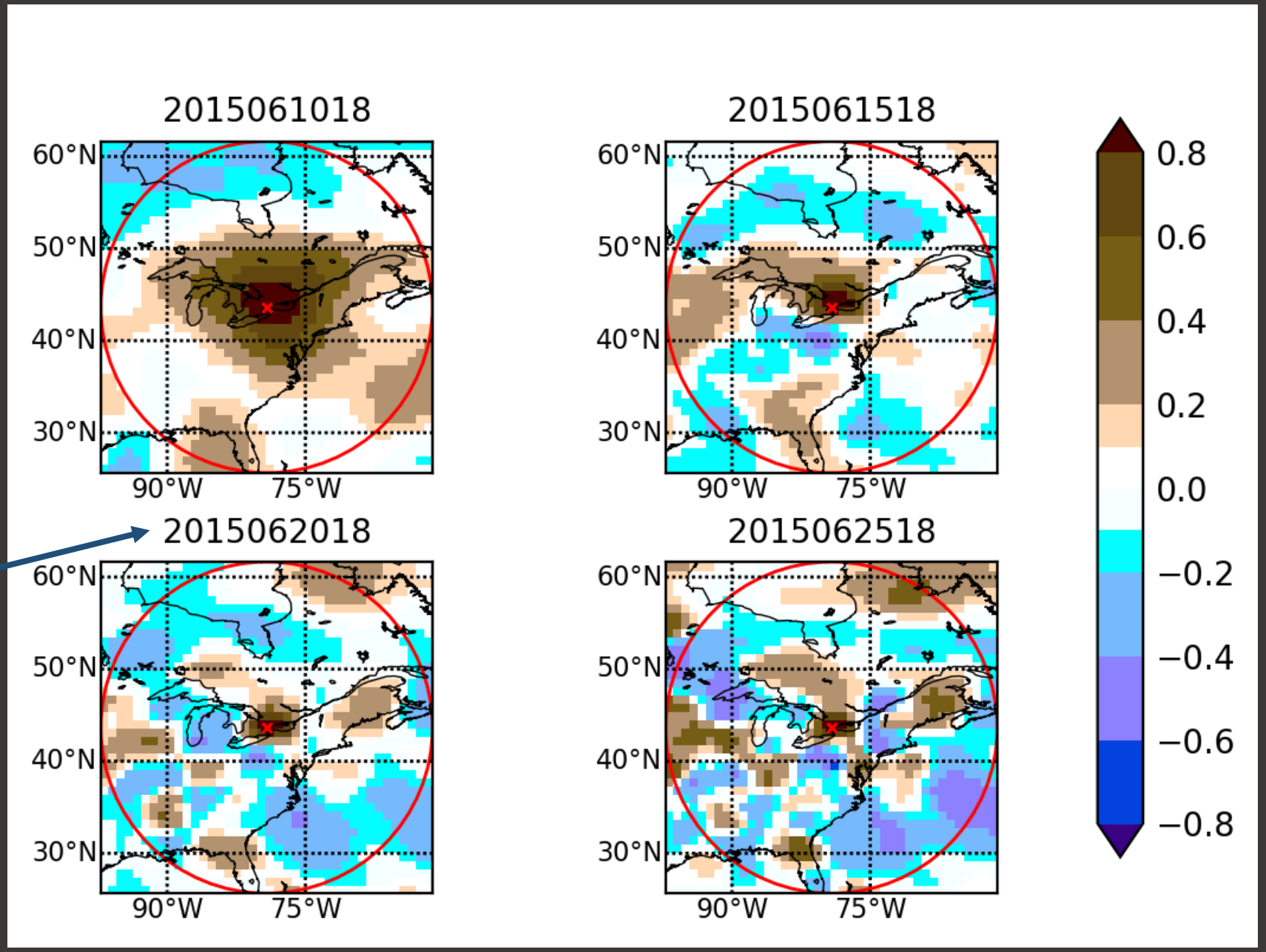
Spread is over dispersive !



The RMSE hits the minimum value – but then it starts increasing !

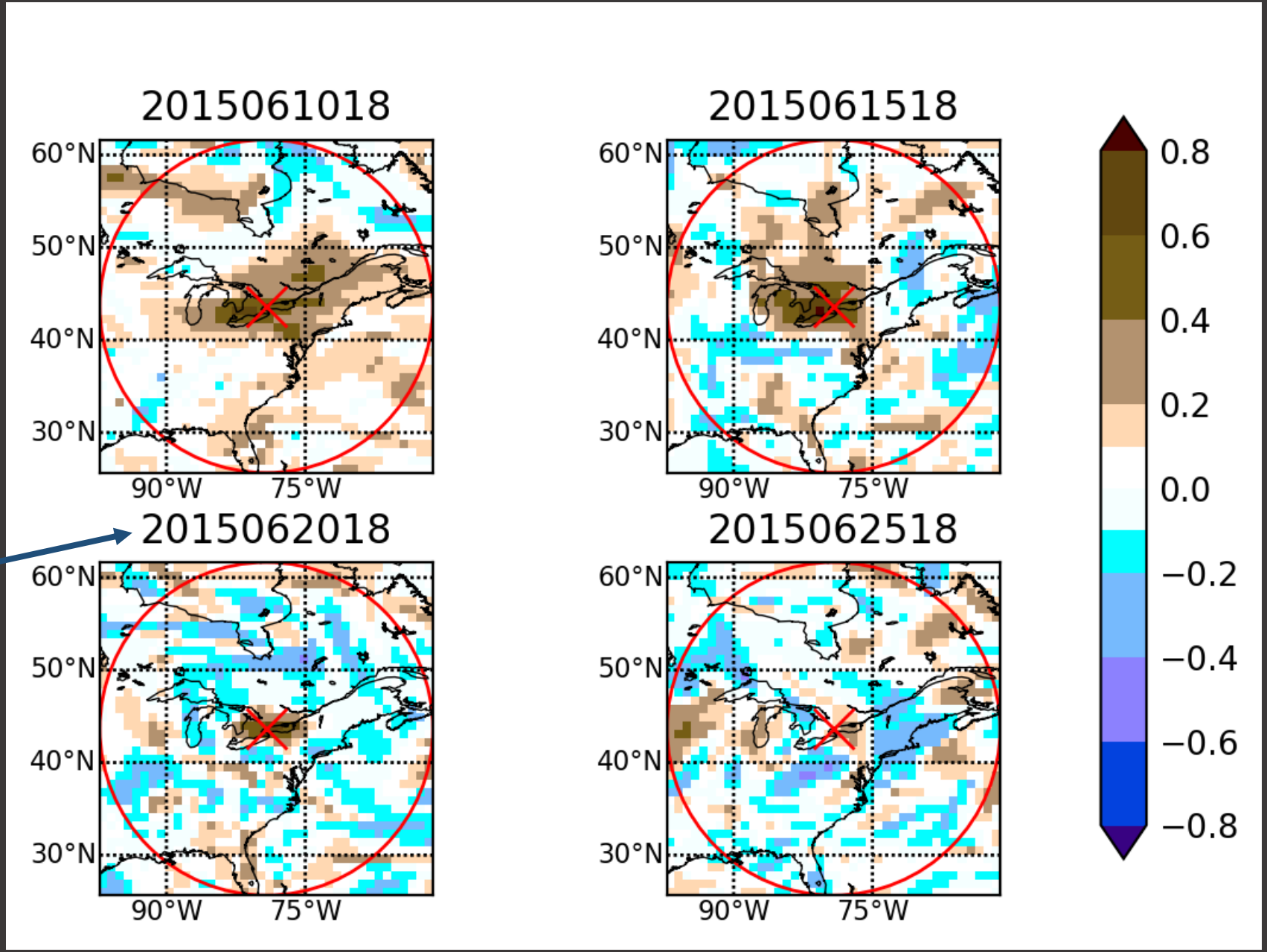
Correlation between flux at Toronto and other grid points.

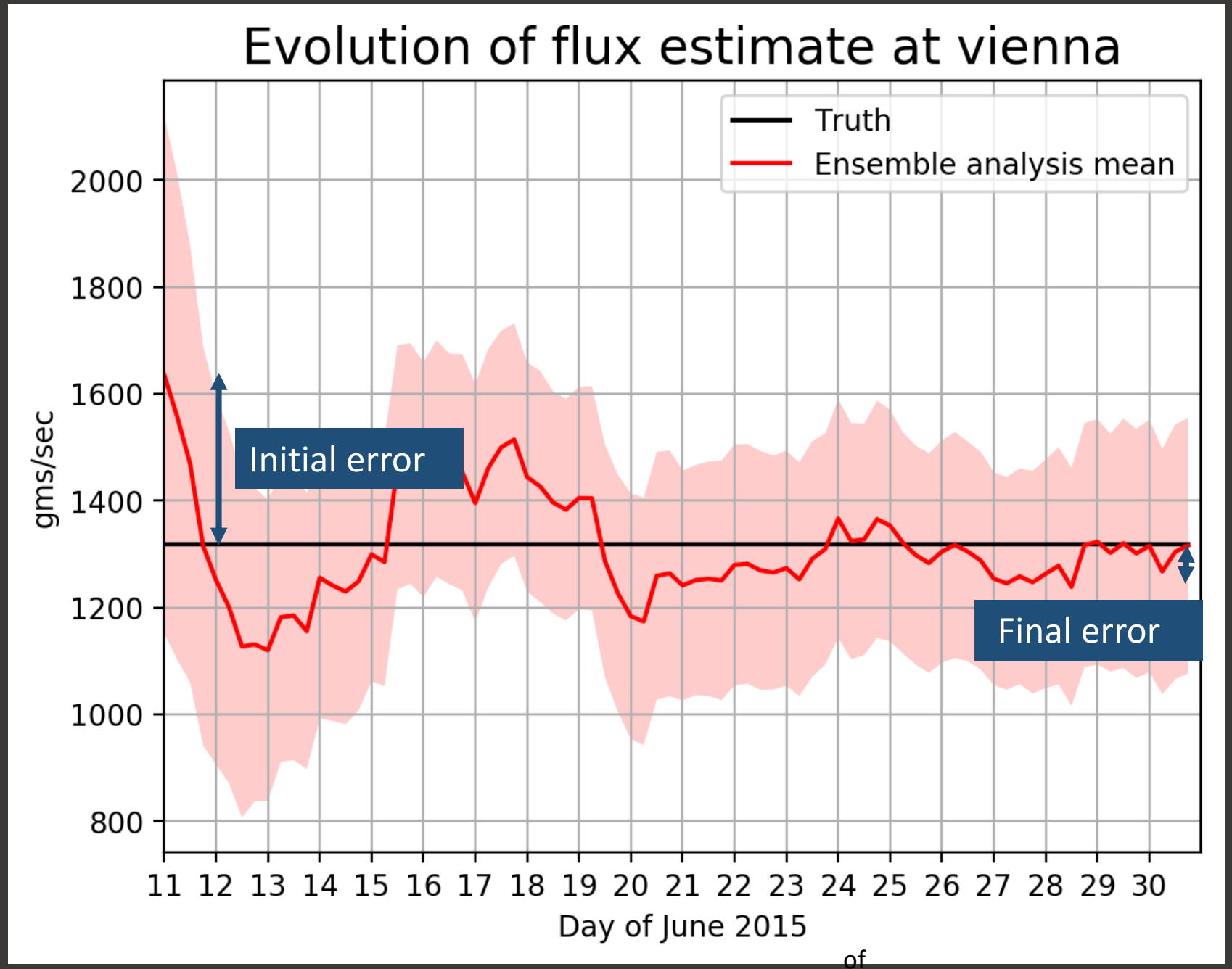
The spatial correlation in the flux field has weakened substantially by 20 June !



Correlation between flux at Toronto and CO at 1 km height.

The correlation between the flux and CO tracer has weakened substantially by 20 June !





Metric =

|Initial error| - |Final error|

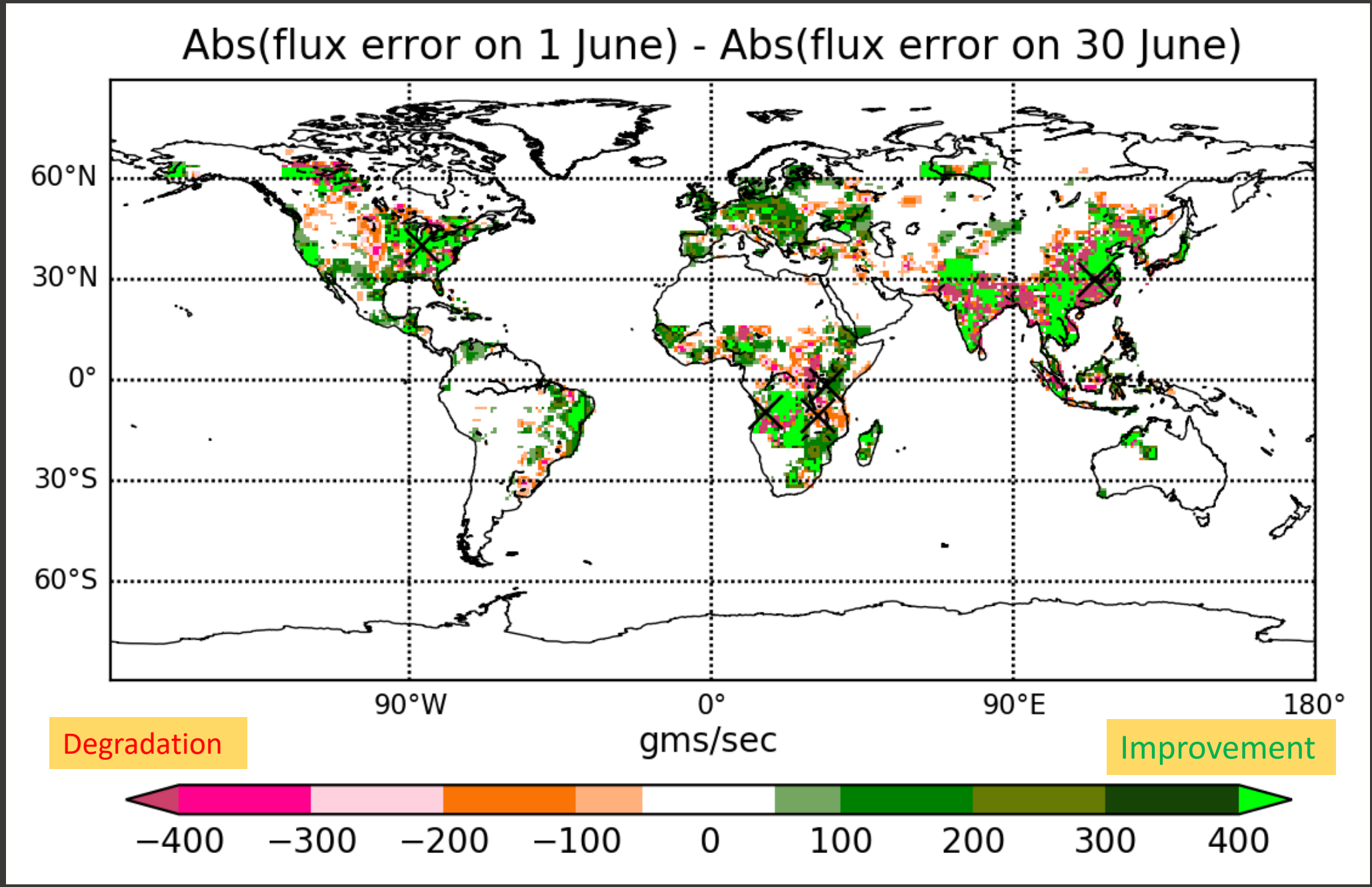
Metric > 0

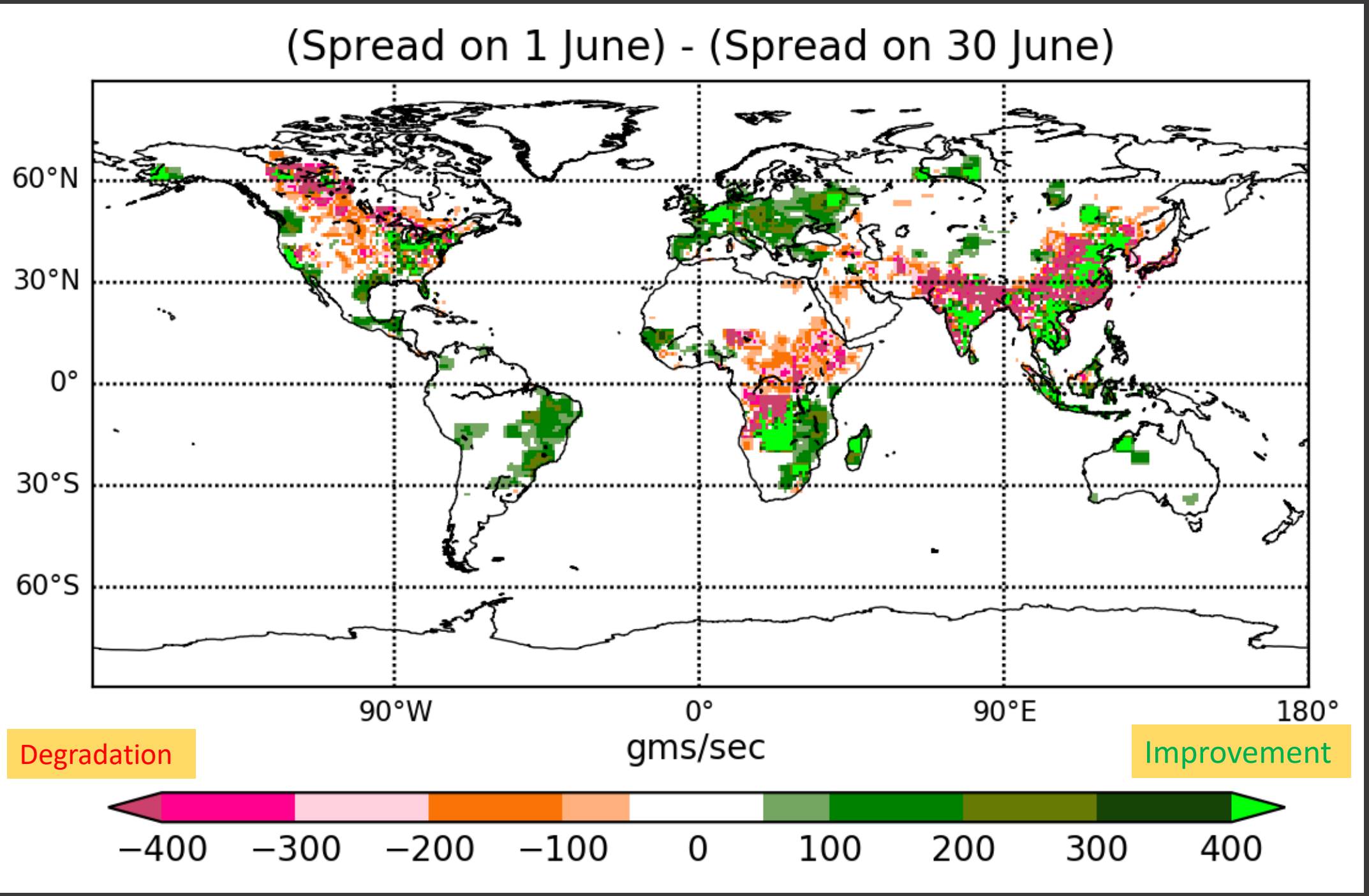
means EnKF improves the estimate.

Metric < 0

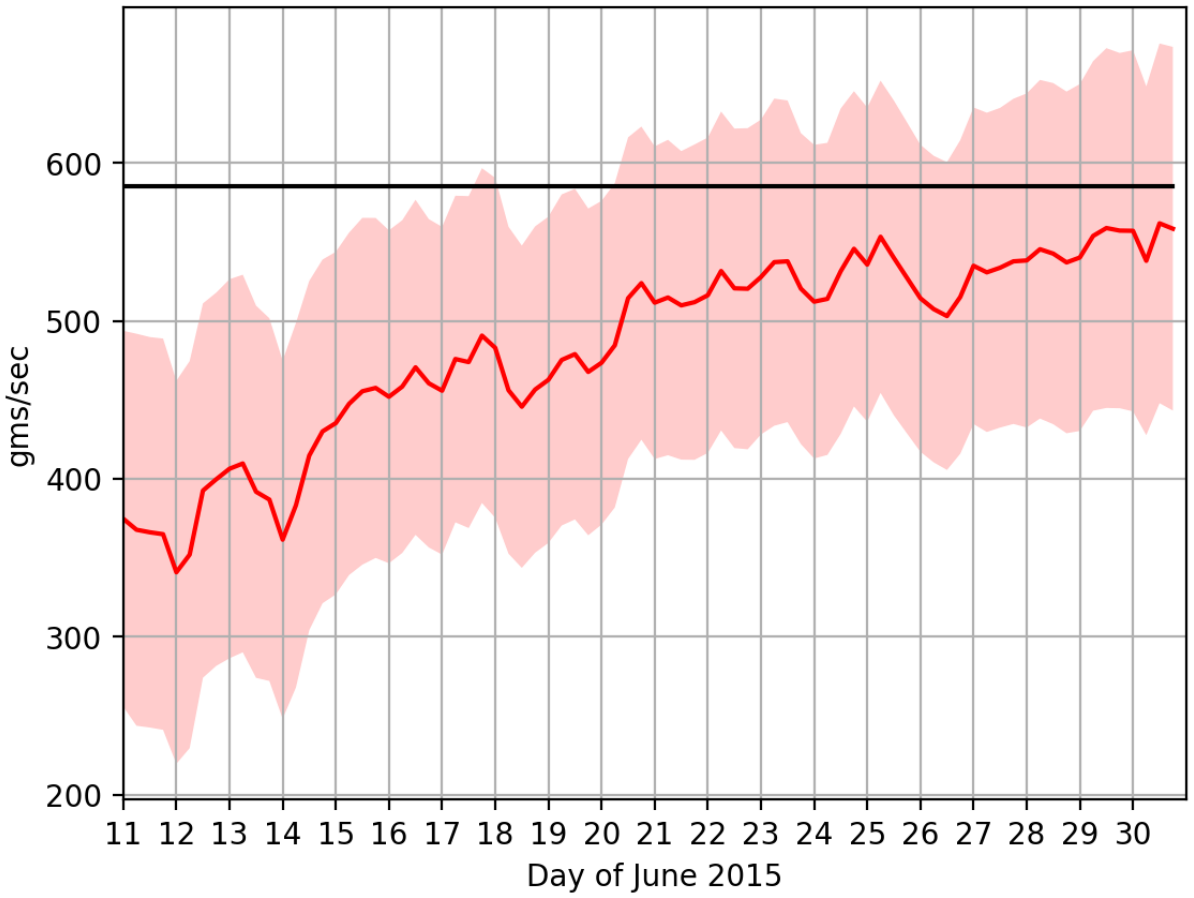
means EnKF degrades the estimate.

of

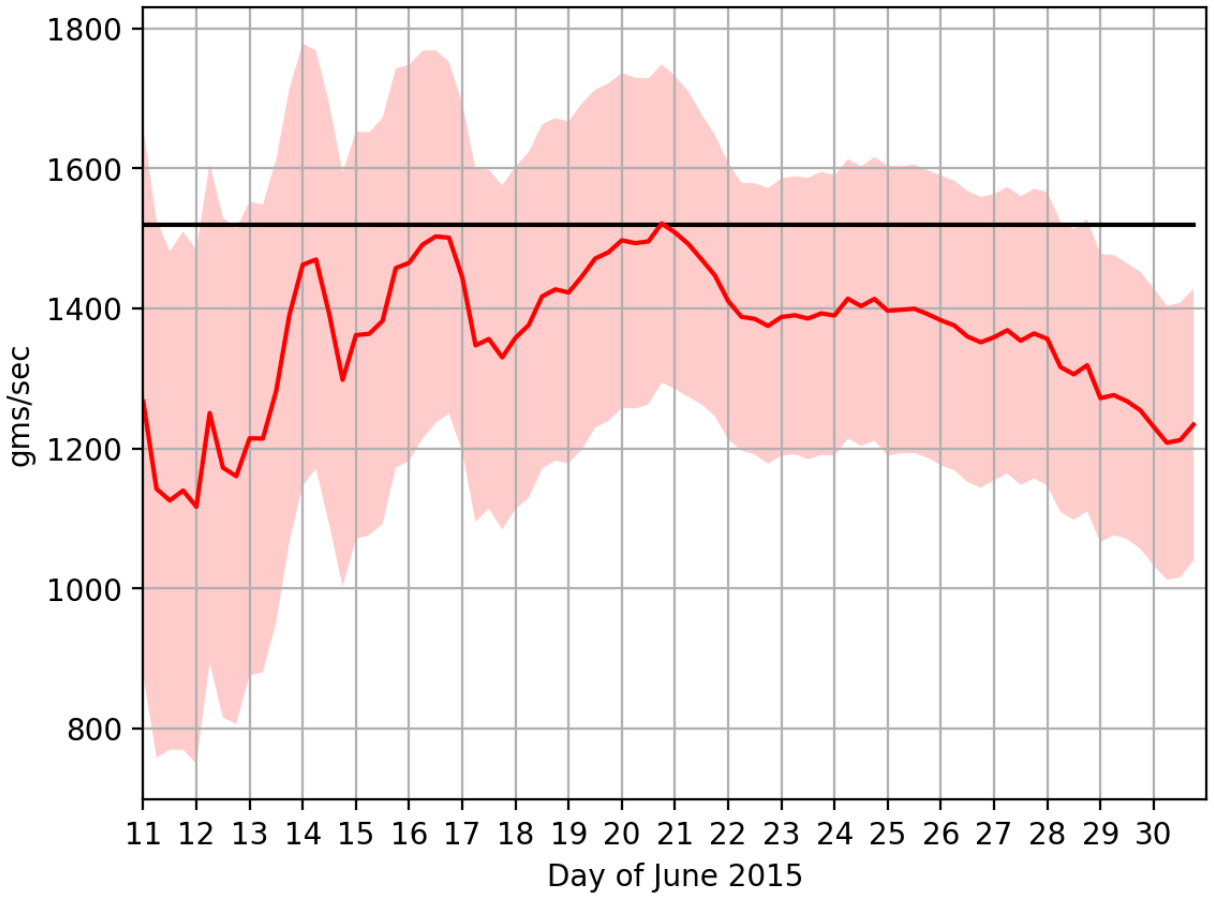

Abs(flux error on 1 June) - Abs(flux error on 30 June)
60°N
30°N
0°
30°S
60°S
90°W
0°
90°E
180°
Degradation
gms/sec
Improvement
−400
−300
−200
−100
0
100
200
300
400


(Spread on 1 June) - (Spread on 30 June)
60°N
30°N
0°
30°S
60°S
90°W
0°
90°E
180°
Degradation
gms/sec
Improvement
−400
−300
−200
−100
0
100
200
300
400

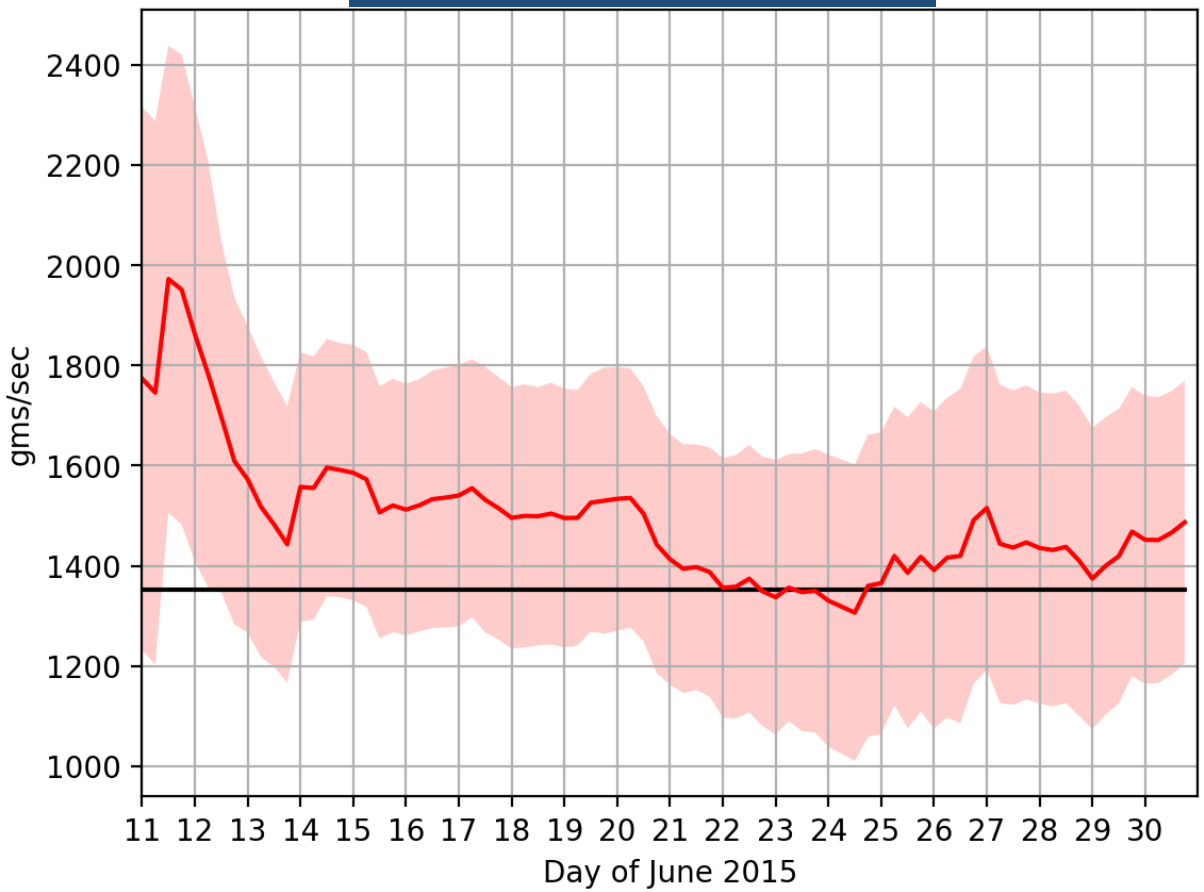
Musoma, Tanzania



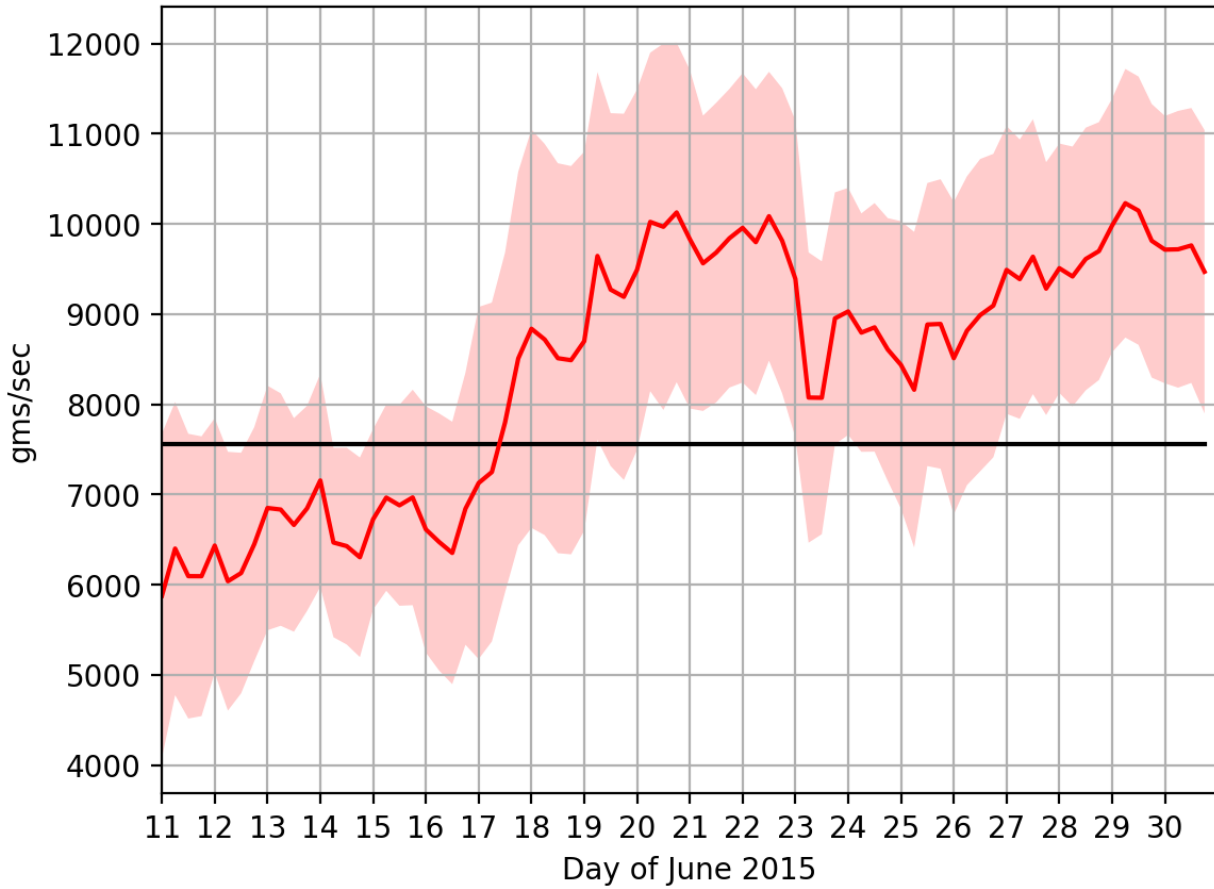
Kasama, Zambia



Maputo, Mozambique



Delhi, India

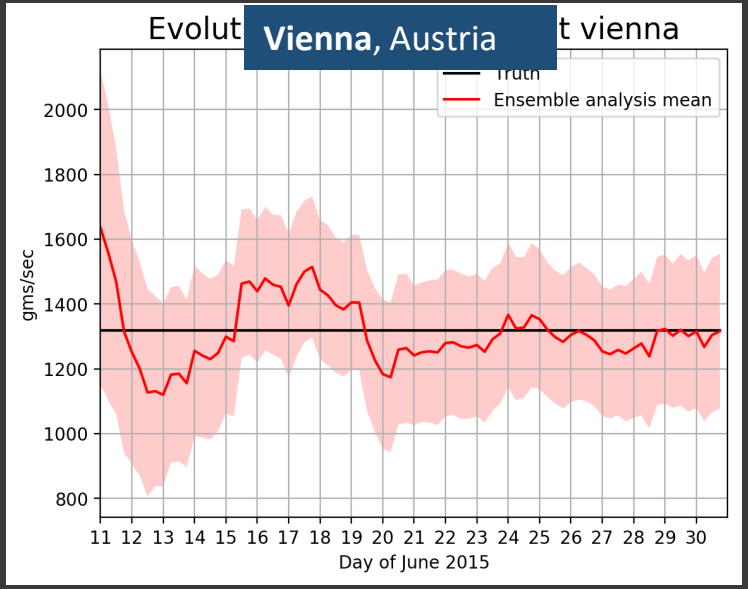


At some points such as Musoma, Maputo the EnKF performs quite well in estimating the truth.

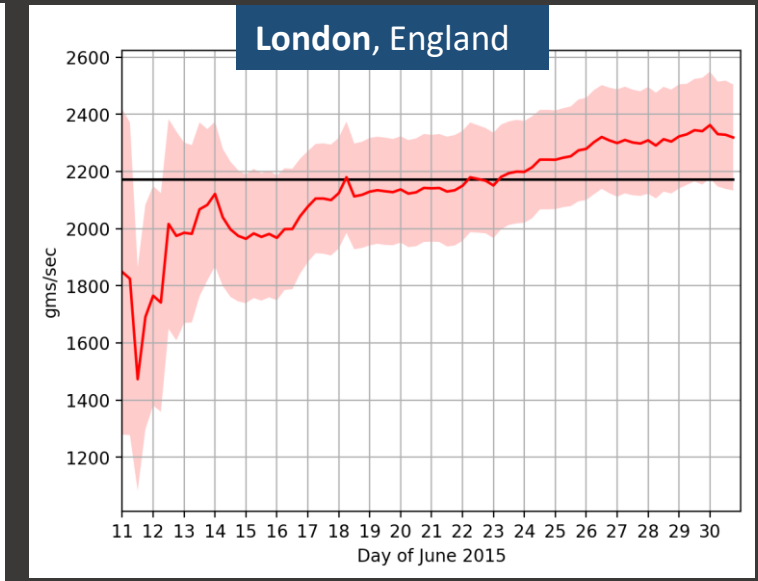
But at other points like Kasama, Delhi etc. the EnKF does not converge (Why ?).

EnKF performs quite well in Europe.

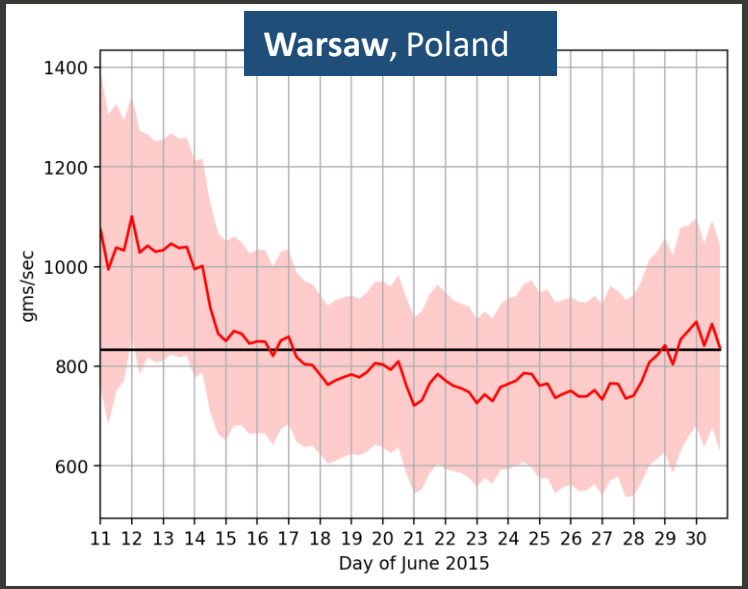
The analysis mean converges towards truth. However the spread appears over dispersive at some locations.



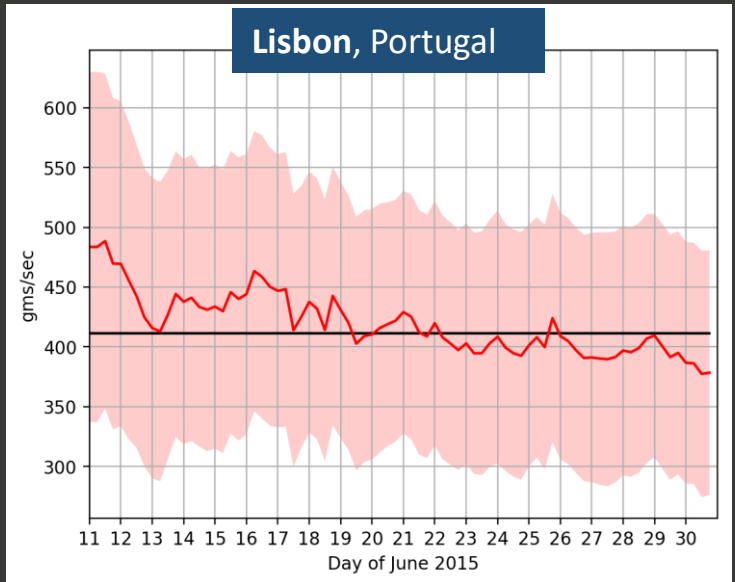
Vienna, Austria



London, England

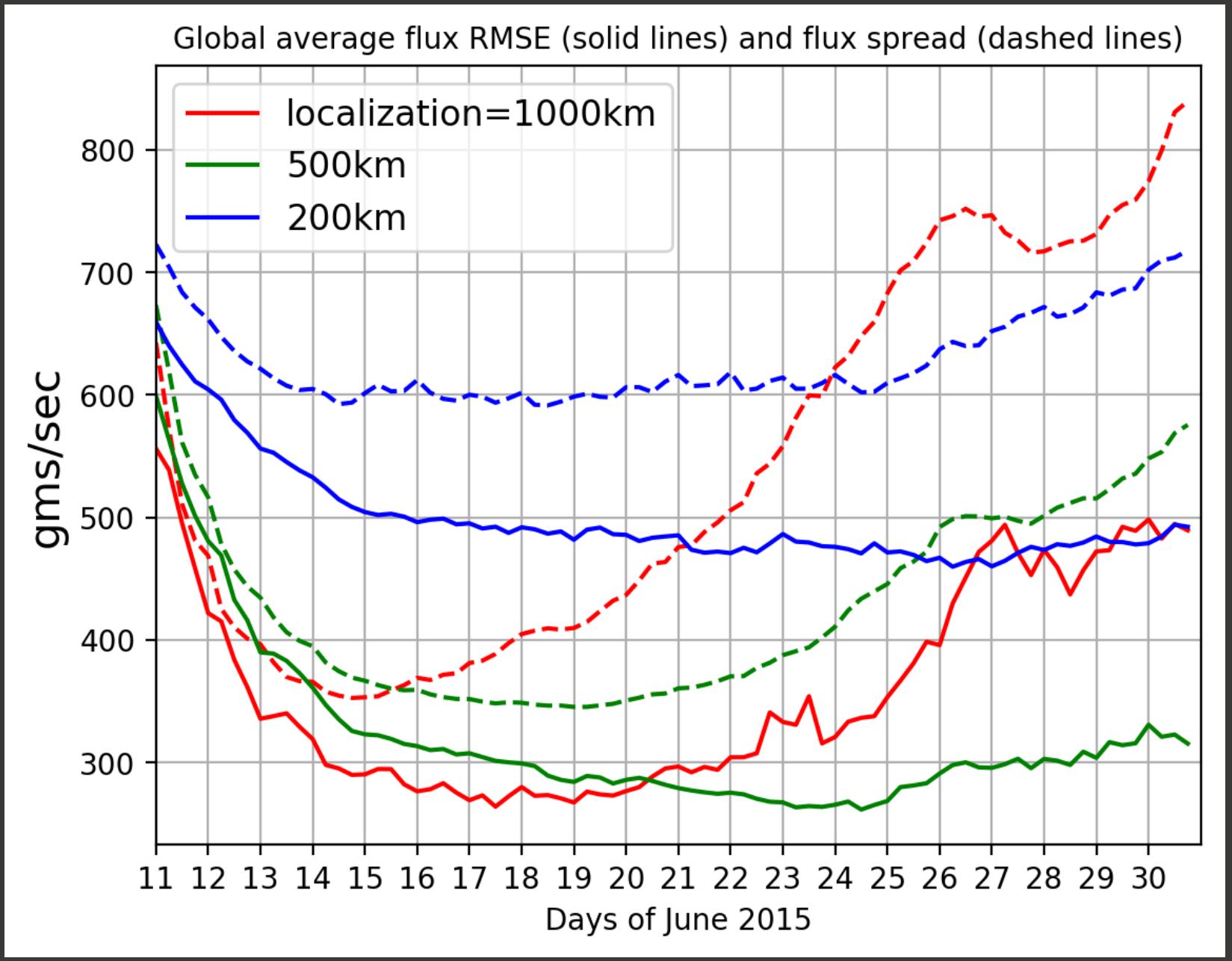


Warsaw, Poland



Lisbon, Portugal

Sensitivity to localization radius

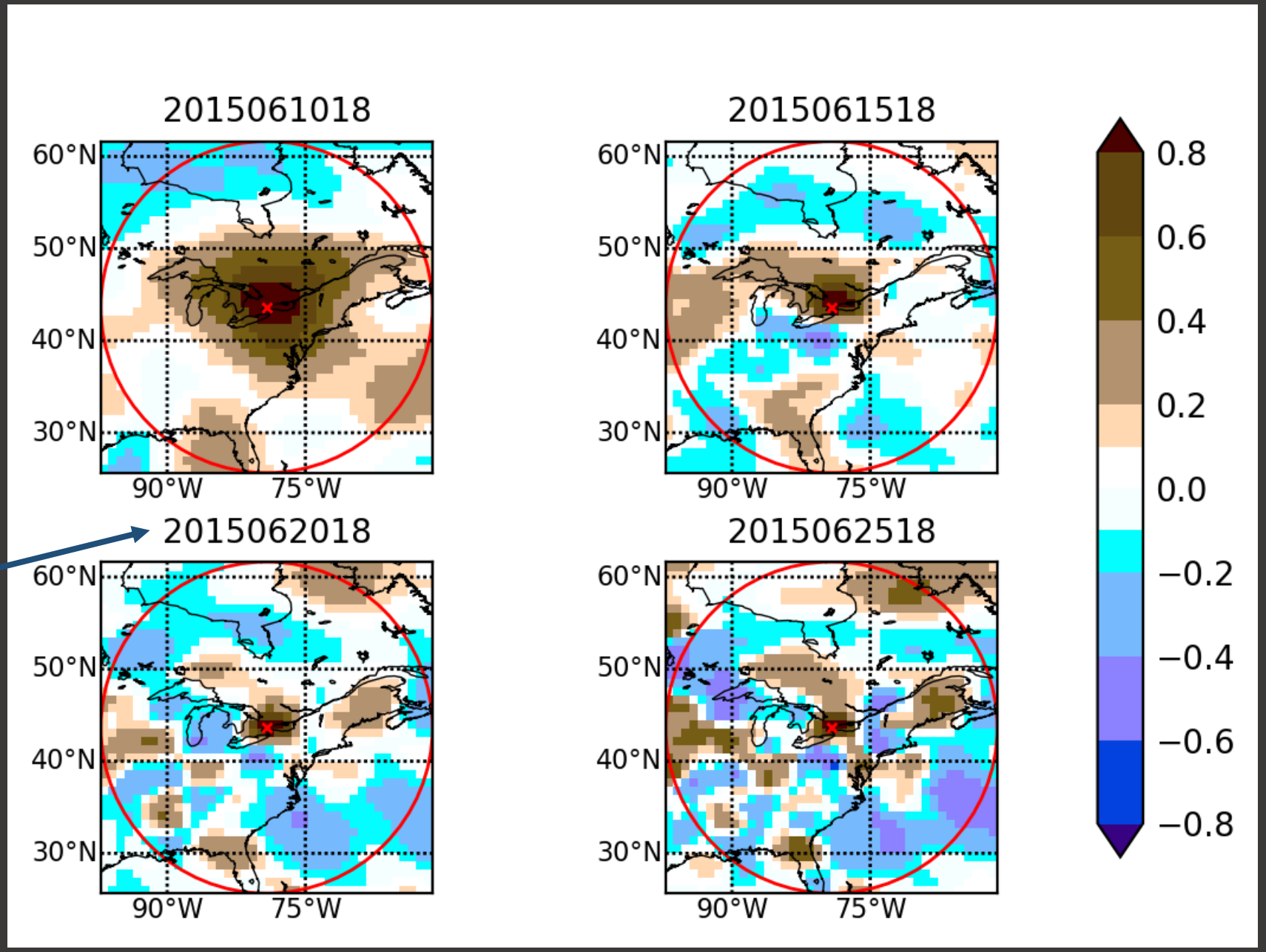


200 km is suboptimal.

500 km does not solve the problem of RMSE increasing.

Correlation between flux at Toronto and other grid points.

The spatial correlation in the flux field has weakened substantially by 20 June !



Conclusions and current status

- Preliminary results suggest that it is possible to constrain the flux field by assimilating CO observations.
- Currently investigating why EnKF is not able to perform well at many grid points.
- Hypothesis is that the spatial correlation of flux (and CO) is weakening as DA cycles proceed.
- Working on including an additive error scheme where the error is a correlated perturbation of flux and CO.

Inflated ensemble member
=
analysis mean + analysis pert + additive pert

These perturbations are from an independent ensemble run.
Perturbations are added to the CO state and flux.