The 6th EnKF workshop @ Buffalo, NY
May 18-22, 2014

KIAPS
www.kiaps.org

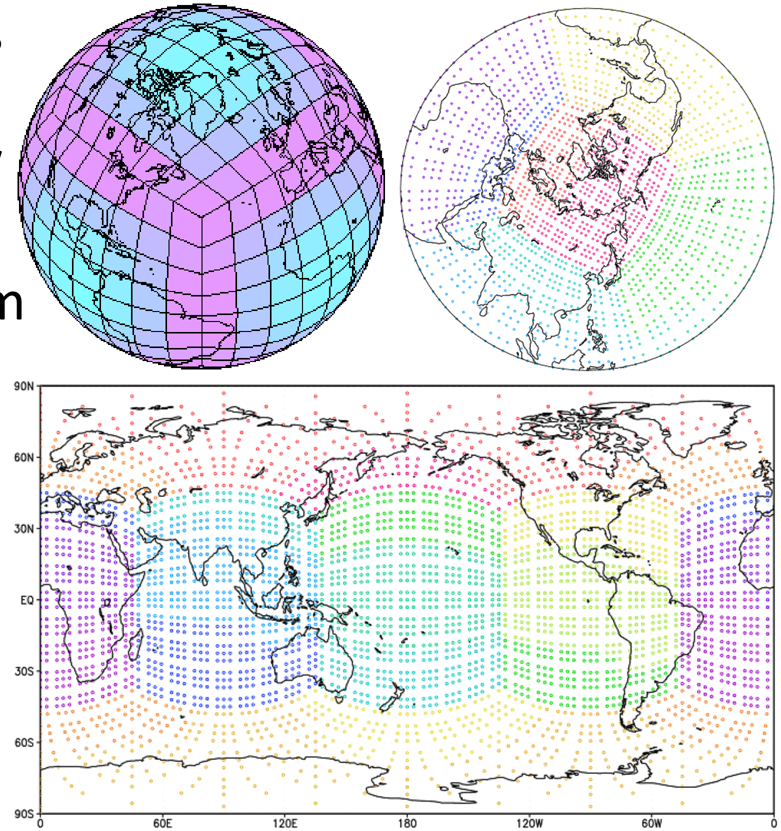
Estimation of Forecast Sensitivity to Observations within KIAPS-LETKF data assimilation system

Ji-Sun Kang*,
Byoung-Joo Jung*,
Youngsoon Jo*, and Yoichiro Ota#

*Korea Institute of Atmospheric Prediction Systems (KIAPS)
#Japanese Meteorological Agency (JMA)

KIAPS-LETKF Data Assimilation System

- **KIAPS-LETKF data assimilation system** has been developed **for cubed-sphere grid** model (e.g. NCAR CAM-SE, KIAPS-GM now being developed).
- We have examined the KIAPS-LETKF system in various simulation experiments, and have **successfully assimilated real data** (conventional data from NCEP prepbufr) into the system.
  - *Posters of Dr. Jung and Ms. Jo*
- Now, (AMSU-A and IASI) *radiance* data assimilation using RTTOV and *GPS RO* data assimilation using ROPP are in progress.
- We also would like to estimate observation impact using EFSO technique introduced by Kalnay et al. (2012) within KIAPS-LETKF.



Unstructured grid on the cubed sphere▲

Ensemble Forecast Sensitivity to Observations (EFSO)

- **Estimating forecast sensitivity to observations, *without TLM/ADJ of the model***
  - Kalnay et al. (2012) has introduced a simple formula to estimate forecast sensitivity to observations, within an ensemble data assimilation cycle.
  - **Ota et al. (2013)** has *applied* the method of Kalnay et al. (2012) to *NCEP GFS-EnSRF data assimilation system* and has shown **promising results.**

➔ *We decided to implement this method to KIAPS-LETKF system*

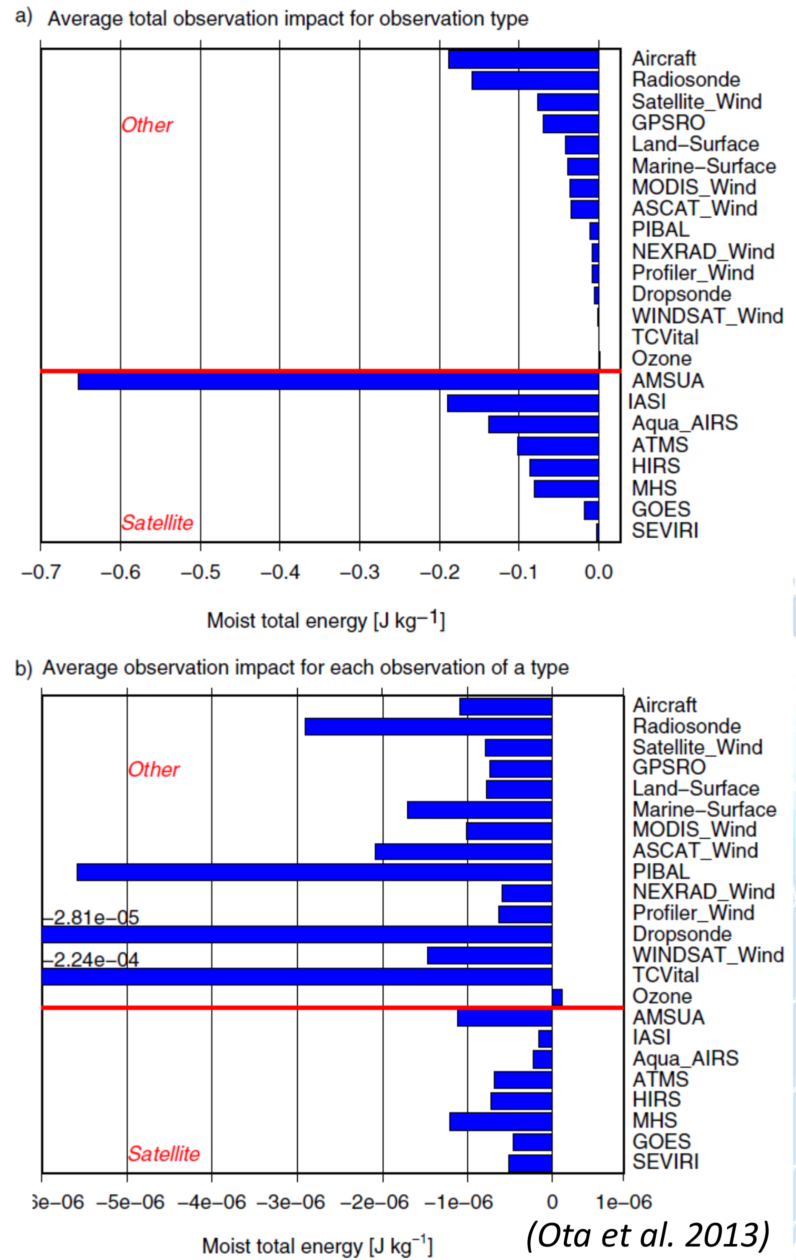
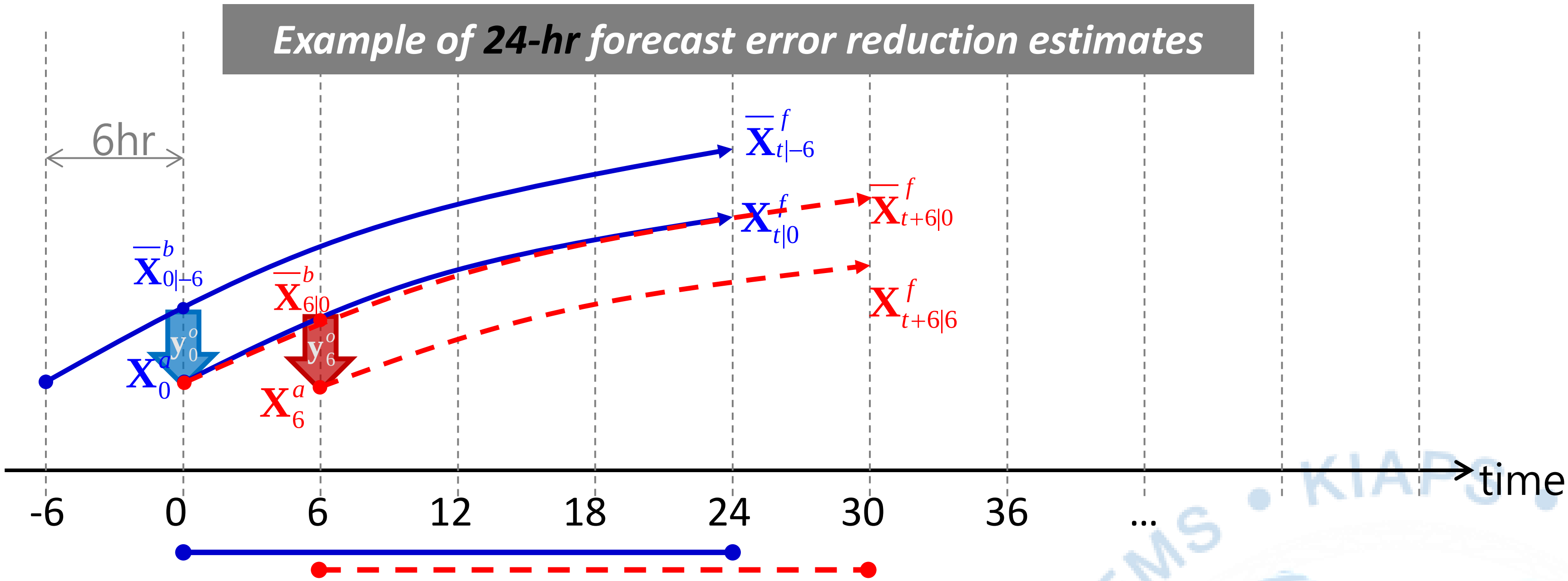


Fig. 2. Estimated average 24-hour forecast error reduction contributed from each observation types (moist total energy, J kg^{-1}). (a) represents the total error reduction and (b) represents error reduction per observation.

(Ota et al. 2013)

KIAPS

Ensemble Forecast Sensitivity to Observations (EFSO)

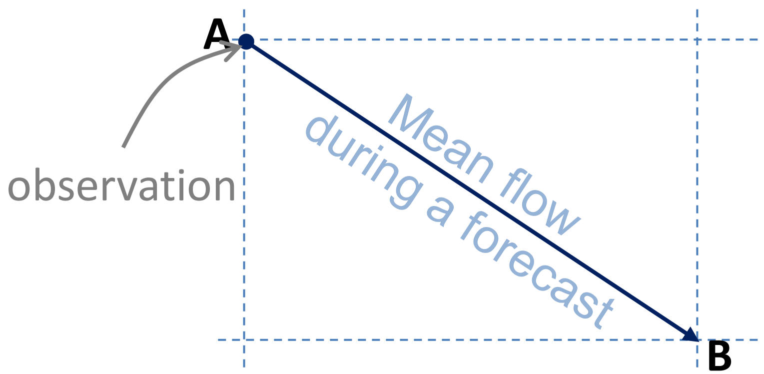


- **Forecast error reduction is defined as** $J = \mathbf{e}^T_{t|0}\mathbf{C}_{jj}\mathbf{e}_{t|0} - \mathbf{e}^T_{t|-6}\mathbf{C}_{jj}\mathbf{e}_{t|-6}$
  - Here, $\mathbf{e}_{t|0} = \overline{\mathbf{X}}^f_{t|0} - \mathbf{X}_t$, $\mathbf{e}_{t|-6} = \overline{\mathbf{X}}^f_{t|-6} - \mathbf{X}_t$ ($\mathbf{X}_t$ should be the best estimate of true atmosphere), and C_{jj} is a total moist energy norm.
- **EFSO formula** introduced by Kalnay et al. (2012)

$$\left(\Delta e^2\right)_{j,l} = \frac{1}{K-1}\left(\delta \mathbf{y}_0\right)_l \left[\rho_j \mathbf{R}^{-1}\mathbf{Y}^a_0\left(\mathbf{X}^{f\mathrm{T}}_{t|0}\right)_j C_{jj}\left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right)_j\right]_l ,$$

How to move a localization function (Kalnay et al., 2012)

- **In order to estimate observation impact at time t, forecast fields valid at t+τ are used.**
  - Observation impact tends to be advected and diffused as a result of atmospheric flow during the forecast hours



- **We need to localize information**
- ***Kalnay et al. (2012)* shows different results from several localization functions**
  - *As τ increases, no localization case gives better performance*

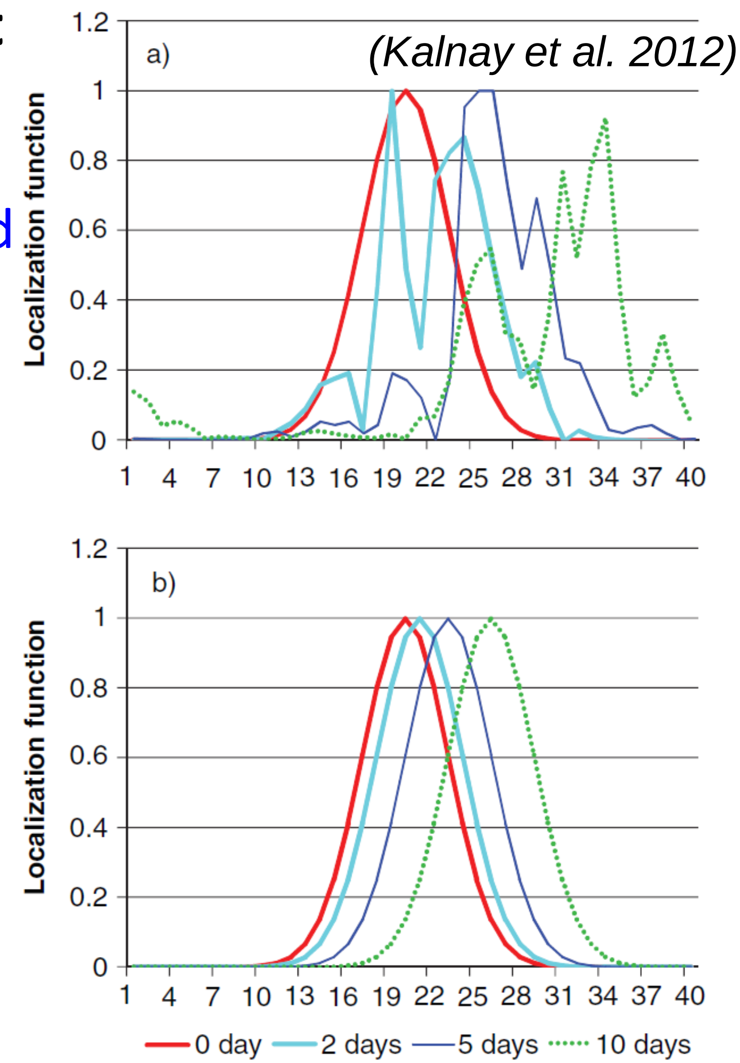


Fig. 5. Example of the evolution of the localization function with two methods. (a) Localization function with the non-linear incremental evolution (NL-loc); and (b) localization function moving with constant group velocity (CL-loc).

Localization Function in Kalnay et al. (2012)

Table 2. Same as Table 1 but with 10 ensemble members and localization when indicated

	6 h	12 h	1 d	2 d	3 d	5 d	7 d
Truth	−23 (0)	−32 (0)	−56 (0)	−121 (2)	−213 (6)	−582 (27)	−1144 (47)
Analysis	−25 (0)	−31 (0)	−51 (0)	−115 (2)	−206 (6)	−570 (27)	−1126 (46)
(4-ADJ)	−21 (0)	−25 (0)	−39 (0)	−77 (1)	−132 (4)	−344 (17)	−627 (52)
(5-OLD) Noloc	−12 (0)	−14 (0)	−24 (1)	−50 (3)	−69 (4)	136 (22)	1163 (65)
(6-NEW) Noloc	−32 (0)	−39 (0)	−62 (1)	−133 (4)	−227 (6)	−574 (41)	−1220 (80)
(5-OLD) Fxloc	−25 (0)	−30 (0)	−47 (0)	−89 (2)	−117 (4)	−9 (9)	486 (29)
(6-NEW) Fxloc	−28 (0)	−33 (0)	−50 (0)	−86 (2)	−106 (4)	−125 (11)	−162 (29)
(6-NEW) NL-loc	−24 (0)	−24 (0)	−47 (0)	−107 (2)	−178 (4)	−422 (27)	−788 (57)
(6-NEW) CL-loc	−28 (0)	−33 (0)	−54 (0)	−99 (2)	−151 (5)	−284 (18)	−462 (30)

▲Truth/Analysis: $J = \mathbf{e}_{t|0}^T \mathbf{C}_{jj} \mathbf{e}_{t|0} - \mathbf{e}_{t|-6}^T \mathbf{C}_{jj} \mathbf{e}_{t|-6}$ (6-NEW): $(\Delta e^2)_{j,l} = \frac{1}{K-1} (\delta \mathbf{y}_0)_l \left[\rho_j \mathbf{R}^{-1} \mathbf{Y}_0^a \left(\mathbf{X}_{t|0}^{f\mathrm{T}} \right)_j C_{jj} \left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6} \right)_j \right]_l$

- **NoLoc gives better results than other localization function cases (NL-loc and CL-loc) as the forecast lead time increases.**
- **NL-loc is not feasible for real case due to too much computational cost, but CL-loc has significant underestimation.**
- **Therefore, we would like to try modifying localization function *not only moving the center but also chaning the radius of the function*.**

Localization Function in Kalnay et al. (2012)

Table 2. Same as Table 1 but with 10 ensemble members and localization when indicated

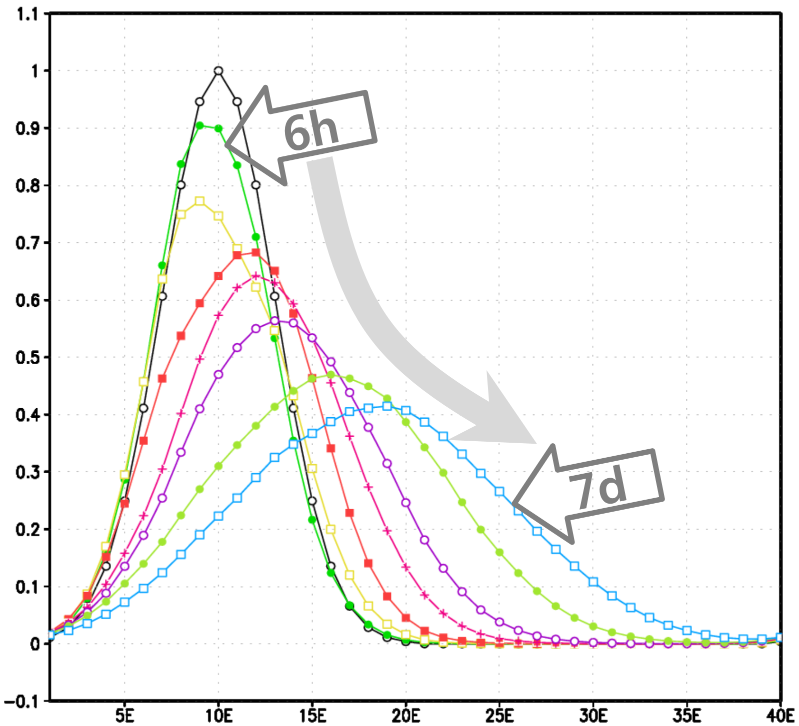
	6 h	12 h	1 d	2 d	3 d	5 d	7 d
Truth	−23 (0)	−32 (0)	−56 (0)	−121 (2)	−213 (6)	−582 (27)	−1144 (47)
Analysis	−25 (0)	−31 (0)	−51 (0)	−115 (2)	−206 (6)	−570 (27)	−1126 (46)
(4-ADJ)	−21 (0)	−25 (0)	−39 (0)	−77 (1)	−132 (4)	−344 (17)	−627 (52)
(5-OLD) Noloc	−12 (0)	−14 (0)	−24 (1)	−50 (3)	−69 (4)	136 (22)	1163 (65)
(6-NEW) Noloc	−32 (0)	−39 (0)	−62 (1)	−133 (4)	−227 (6)	−574 (41)	−1220 (80)
(5-OLD) Fxloc	−25 (0)	−30 (0)	−47 (0)	−89 (2)	−117 (4)	−9 (9)	486 (29)
(6-NEW) Fxloc	−28 (0)	−33 (0)	−50 (0)	−86 (2)	−106 (4)	−125 (11)	−162 (29)
(6-NEW) NL-loc	−24 (0)	−24 (0)	−47 (0)	−107 (2)	−178 (4)	−422 (27)	−788 (57)
(6-NEW) CL-loc	−28 (0)	−33 (0)	−54 (0)	−99 (2)	−151 (5)	−284 (18)	−462 (30)

▲Truth/Analysis: $J = \mathbf{e}_{t|0}^{T}\mathbf{C}_{jj}\mathbf{e}_{t|0} - \mathbf{e}_{t|-6}^{T}\mathbf{C}_{jj}\mathbf{e}_{t|-6}$ (6-NEW): $\left(\Delta e^{2}\right)_{j,l} = \frac{1}{K-1}\left(\delta\mathbf{y}_{0}\right)_{l}\left[\rho_{j}\mathbf{R}^{-1}\mathbf{Y}_{0}^{a}\left(\mathbf{X}_{t|0}^{f\mathbf{T}}\right)_{j}C_{jj}\left(\mathbf{e}_{t|0}+\mathbf{e}_{t|-6}\right)_{j}\right]_{l}$

- **NoLoc gives better results than other localization function cases (NL-loc and CL-loc) as the forecast lead time increases.**
- **NL-loc is not feasible for real case due to too much computational cost, but CL-loc has significant underestimation.**
- **Therefore, we would like to try modifying localization function *not only moving the center but also chaning the radius of the function*.**

How large the localization scale should be?

- **We have estimated a time mean of nonlinear localization function (Nl-Loc) of Kalnay et al. (2012).**



As forecast lead time gets longer, the center of the localization function has been shifted.
Also, *localization scale* gets remarkably larger!

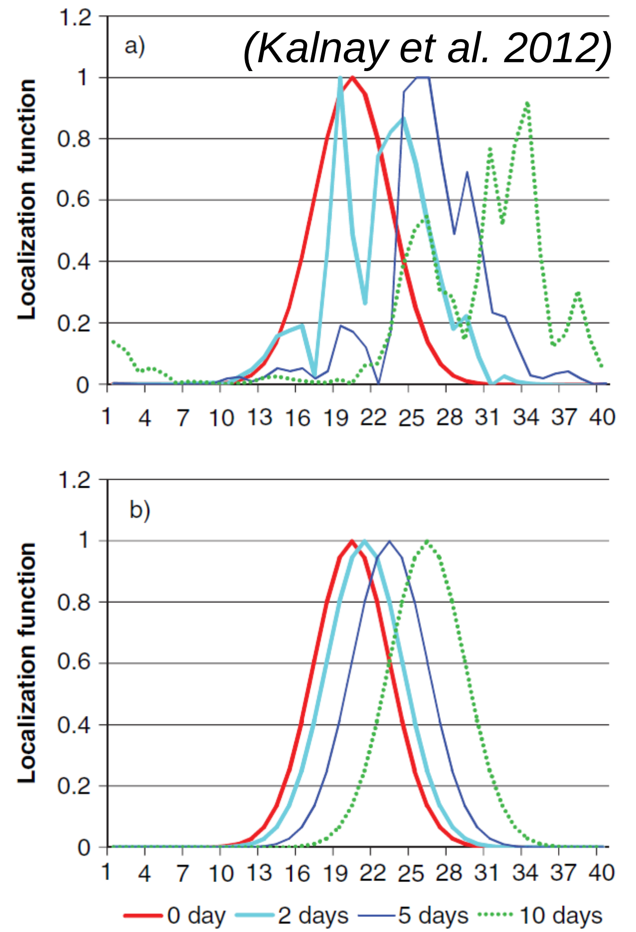
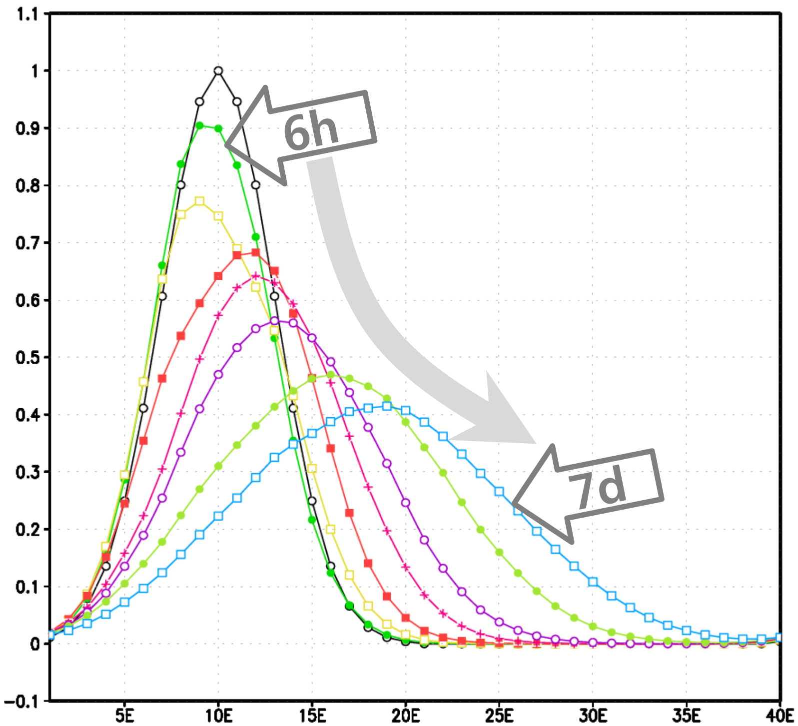
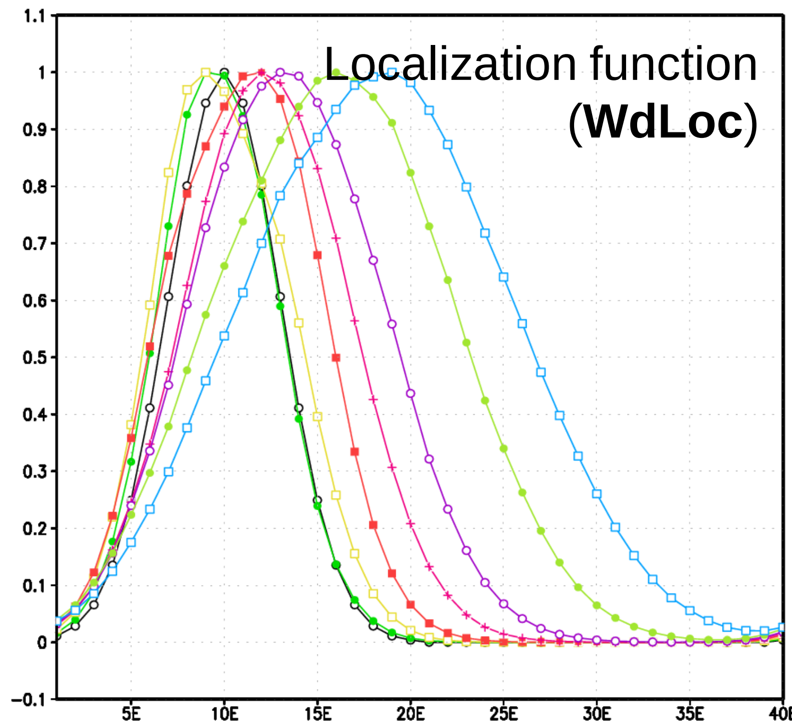


Fig. 5. Example of the evolution of the localization function with two methods. (a) Localization function with the non-linear incremental evolution (NL-loc); and (b) localization function moving with constant group velocity (CL-loc).

How large the localization scale should be?

- **We have estimated a time mean of nonlinear localization function (Nl-Loc) of Kalnay et al. (2012).**





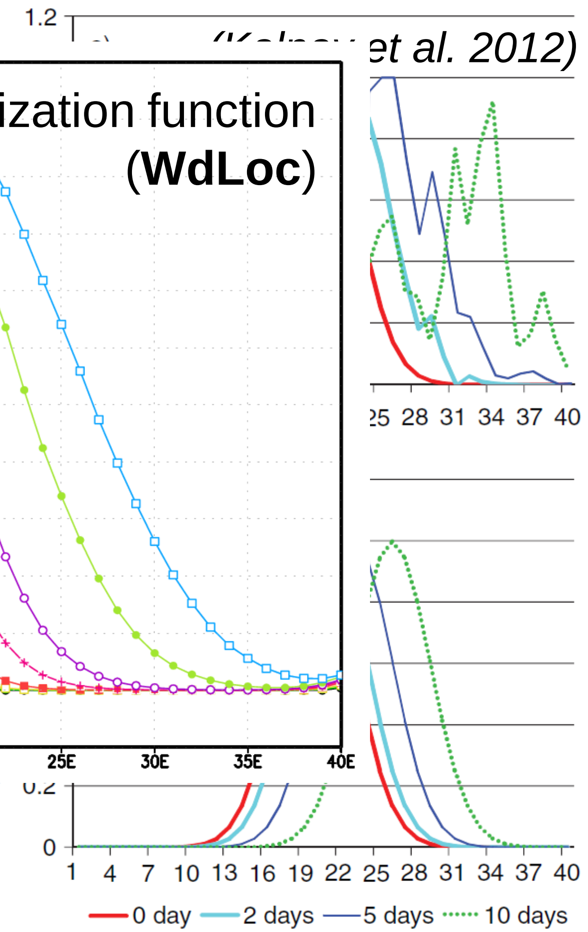


Fig. 5. Example of the evolution of the localization function with two methods. (a) Localization function with the non-linear incremental evolution (NL-loc); and (b) localization function moving with constant group velocity (CL-loc).

As forecast lead time gets longer, the center of the localization function has been shifted.
Also, *localization scale* gets remarkably larger!

Simple Experiment with Lorenz 40-var model

- We have reproduced the table as shown in the previous slide and compare our estimation with a different localization function.

	6 h	12 h	1 d	2 d	3 d	5 d	7 d
Truth	**-23.4**	**-32.4**	**-56.06**	**-121.76**	**-214.38**	**-584.14**	**-1139.08**
Analysis	-24.96	-30.74	-51.12	-115.06	-206.76	-572.46	-1121.2
FxLoc	-28.28	-33.54	-50.12	-86.46	-107.82	-124.52	-155.46
NoLoc	-31.84	-38.62	-62.48	-133.68	-227.8	-582.82	-1224.4
ClLoc	-28.28	-33.54	-53.86	-99.88	-151.88	-284.66	-451.14
NlLoc	-24.26	-24.4	-47.54	-107.26	-177.9	-416.3	-763.68
WdLoc	-28.38	-33.54	-57.5	-118.98	-201.1	-509.34	-1033.92

- Results show that **WdLoc consistently outperforms ClLoc and NlLoc when the forecast lead time is greater than 6 hour.**
  - NlLoc takes too long *(~ four times slower than others)*
- Thus, we conclude that **localization scale changes the results significantly. In order to improve the result, we need to consider more sophisticated localization function in terms of its *width*.**

Simple Experiment with Lorenz 40-var model

- We have reproduced the table as shown in the previous slide and compare our estimation with a different localization function.

	6 h	12 h	1 d	2 d	3 d	5 d	7 d
Truth	**-23.4**	**-32.4**	**-56.06**	**-121.76**	**-214.38**	**-584.14**	**-1139.08**
Analysis	-24.96	-30.74	-51.12	-115.06	-206.76	-572.46	-1121.2
FxLoc	-28.28	-33.54	-50.12	-86.46	-107.82	-124.52	-155.46
NoLoc	-31.84	-38.62	-62.48	-133.68	-227.8	-582.82	-1224.4
ClLoc	-28.28	-33.54	-53.86	-99.88	-151.88	-284.66	-451.14
NlLoc	-24.26	-24.4	-47.54	-107.26	-177.9	-416.3	-763.68
WdLoc	-28.38	-33.54	-57.5	-118.98	-201.1	-509.34	-1033.92

- Results show that **WdLoc consistently outperforms ClLoc and NlLoc when the forecast lead time is greater than 6 hour.**
  - NlLoc takes too long *(~ four times slower than others)*
- Thus, we conclude that **localization scale changes the results significantly. In order to improve the result, we need to consider more sophisticated localization function in terms of its *width*.**

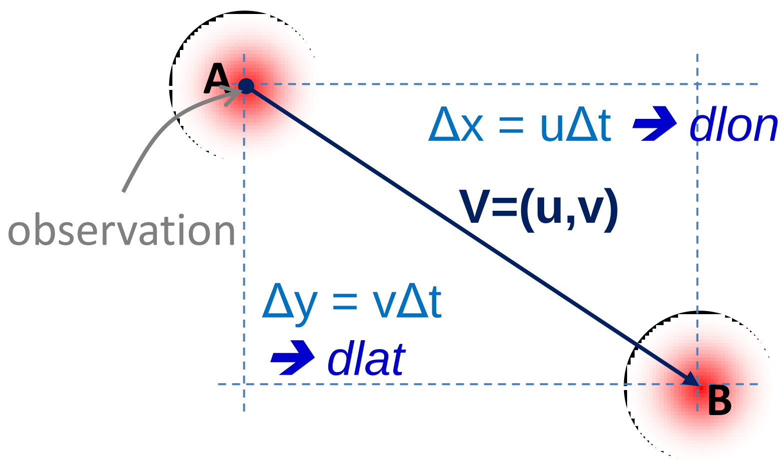
KIAPS

Findings from Lorenz model experiments

- **Tuning width of localization function with respect to forecast lead time changes results significantly.**
  - *Use of **greater localization scales for longer forecast lead time** improves the estimation of EFSO*
- **In more realistic system, it is not easy to define climatological localization scales as the case of Lorenz**
  - We have considered several ways to tune the localization scales adaptively.

Localization function of Ota et al. (2013) – (0)

- **Moving the center of localization functions**
  - *without changing a localization scale*
  - "The coefficient that multiplies the average horizontal wind is tuned"



Information at the forecast time, collected for estimating an impact of the observation at A

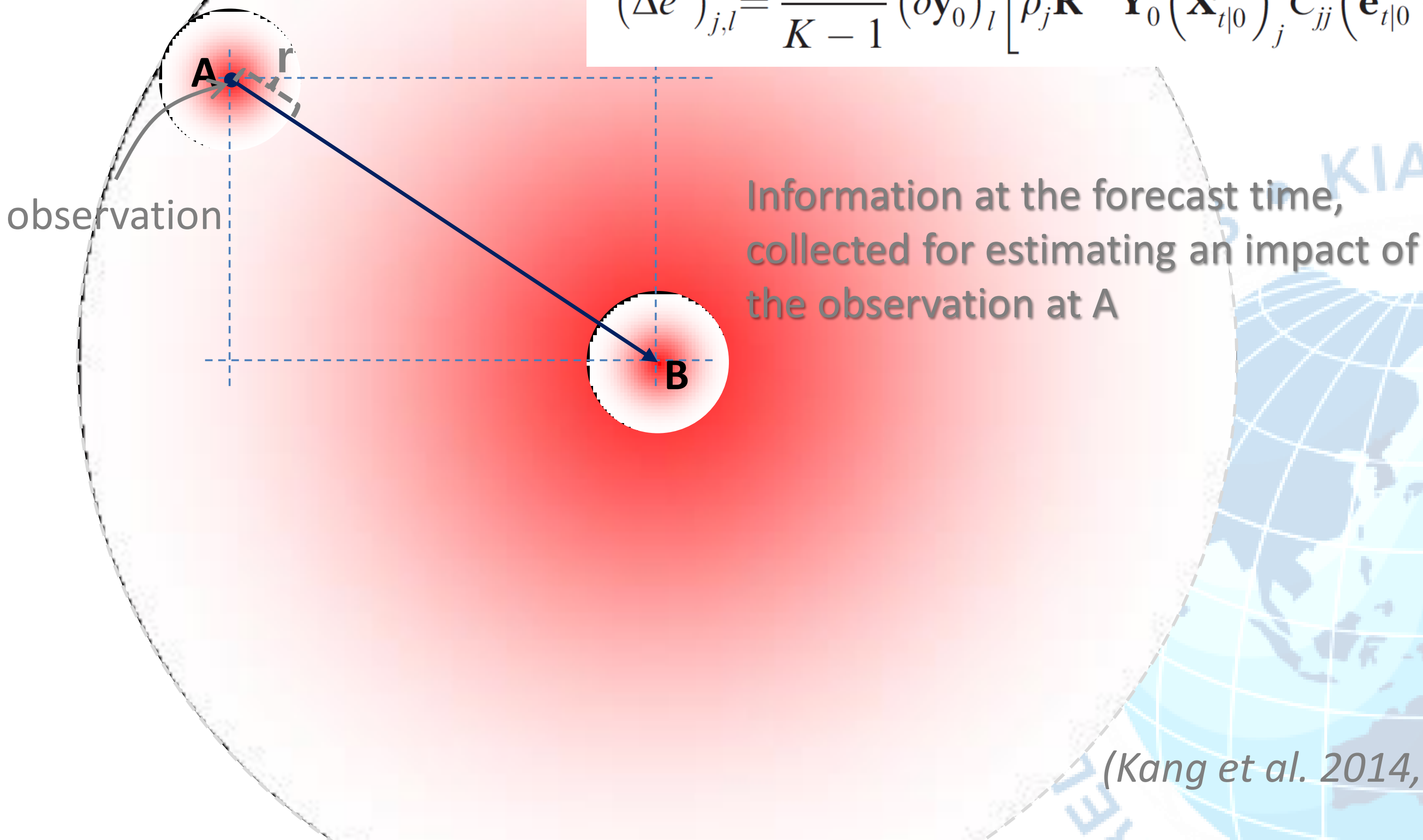
$$\left(\Delta e^2\right)_{j,l} = \frac{1}{K-1}\left(\delta \mathbf{y}_0\right)_l \left[\rho_j \mathbf{R}^{-1}\mathbf{Y}_0^a \left(\mathbf{X}_{t|0}^{f\mathrm{T}}\right)_j C_{jj}\left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right)_j\right]_l ,$$

- Since it already computes displacement in zonal and meridional directions, we may use those values for the width of localization function

Various localization functions for EFSO – (1)

- **Changing localization scale based on the moving distance**
  - Circular shape of localization function as a standard
  - Its radius is defined by r + sqrt($\Delta x^2 + \Delta y^2$)

$$\left(\Delta e^2\right)_{j,l} = \frac{1}{K-1}\left(\delta \mathbf{y}_0\right)_l \left[\rho_j \mathbf{R}^{-1}\mathbf{Y}_0^a\left(\mathbf{X}_{t|0}^{f\mathbf{T}}\right)_j C_{jj}\left(\mathbf{e}_{t|0}+\mathbf{e}_{t|-6}\right)_j\right]_l,$$

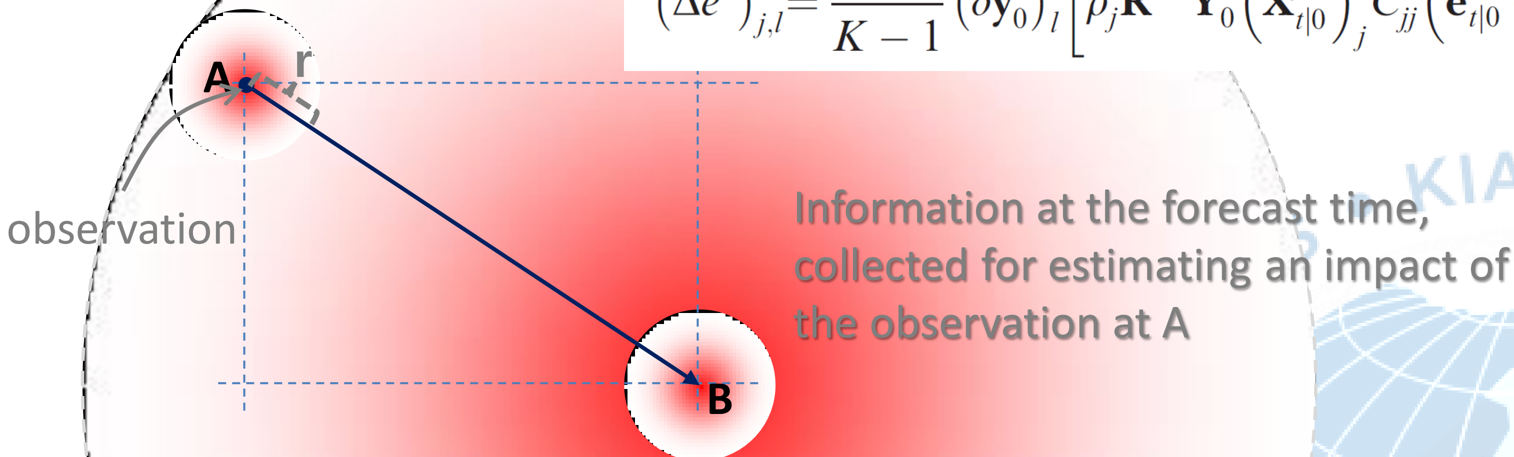


(Kang et al. 2014, in prep.)

Various localization functions for EFSO – (1)

- **Changing localization scale based on the moving distance**
  - Circular shape of localization function as a standard
  - Its radius is defined by r + sqrt($\Delta x^2 + \Delta y^2$)

$$\left(\Delta e^2\right)_{j,l} = \frac{1}{K-1}\left(\delta \mathbf{y}_0\right)_l \left[\rho_j \mathbf{R}^{-1}\mathbf{Y}_0^a \left(\mathbf{X}_{t|0}^{f\mathrm{T}}\right)_j C_{jj}\left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right)_j\right]_l,$$

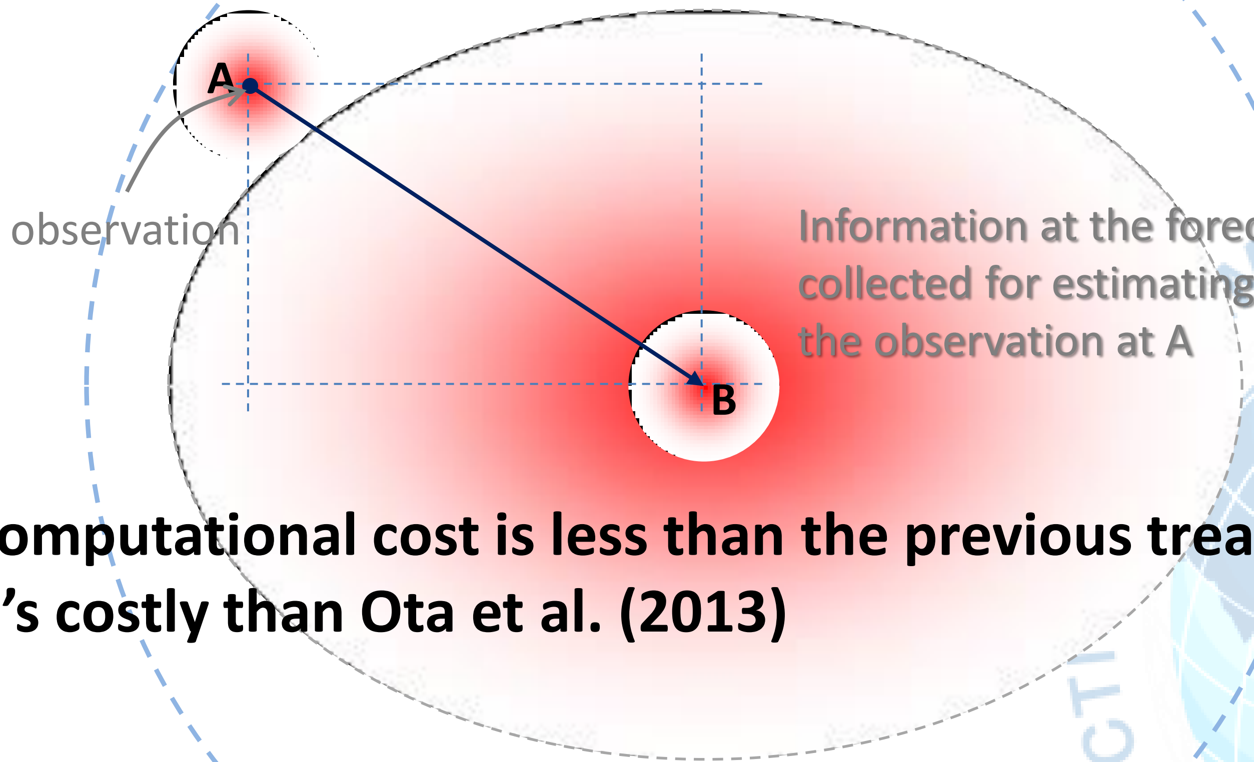


- **Computational burden is increased alot…☹**
- **Δx and Δy tend to be very much different**
  - Suppose a jet region: Δx >> Δy

(Kang et al. 2014, in prep.)

Various localization functions for EFSO – (2)

- **Changing localization scale based on the moving distance of x- and y-direction separately**
  - Elliptic shape of localization function based on ΔX (dlon) and ΔY (dlat)

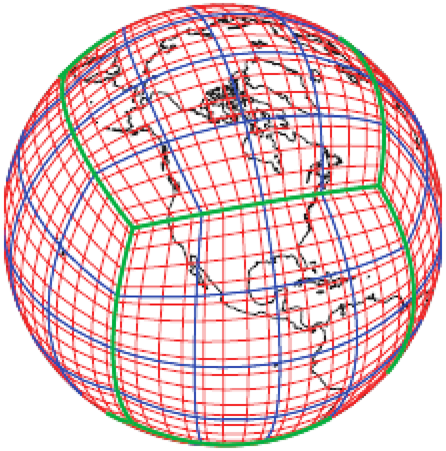


- **Computational cost is less than the previous treatment although it's costly than Ota et al. (2013)**

(Kang et al. 2014, in prep.)

Experimental Design

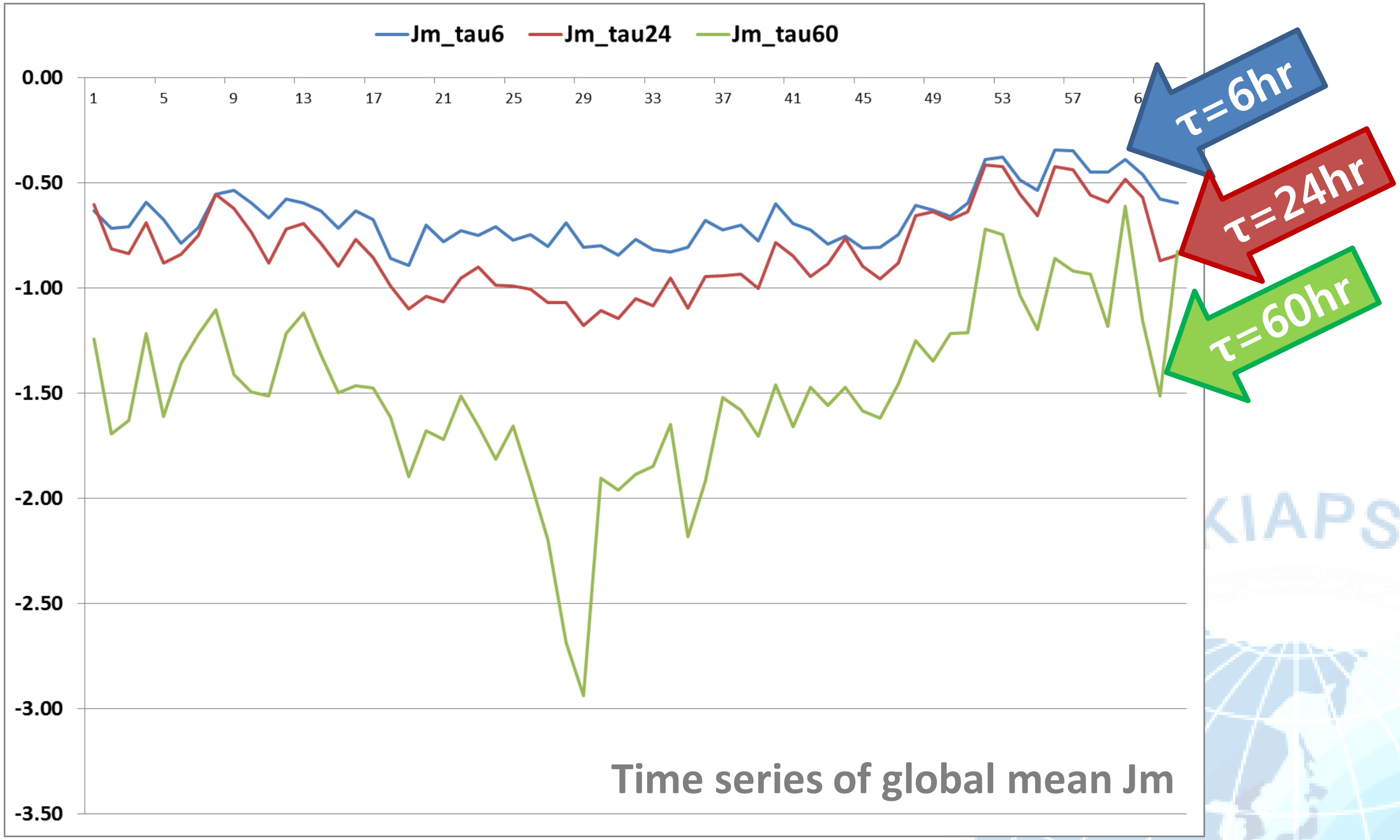
- **KIAPS-LETKF system implemented to NCAR CAM-SE model**
  - Resolution: ne16np4 (~2.5˚) with 30 levels (~2hPa top)
  - Observing System Simulation Experiments (OSSEs)
    - ➤ Sonde, surface Ps, and AIRS T and q retrieval data



- **Impact of localization strategies**
  - Compute $J = \mathbf{e}_{t|0}^{T}\mathbf{C}_{jj}\mathbf{e}_{t|0} - \mathbf{e}_{t|-6}^{T}\mathbf{C}_{jj}\mathbf{e}_{t|-6}$ at every model grid point, **Jm**
  - Compute the same *J, but at the observation space with different localization functions*, **Jo**
  - Compare global mean of **Jm** and **Jo**
    - ➤ ***NO EFSO formula incorporated!***

➔ Assuming perfect FSO formula (w.r.t. truth), we can see the ***impact of localization function only***.

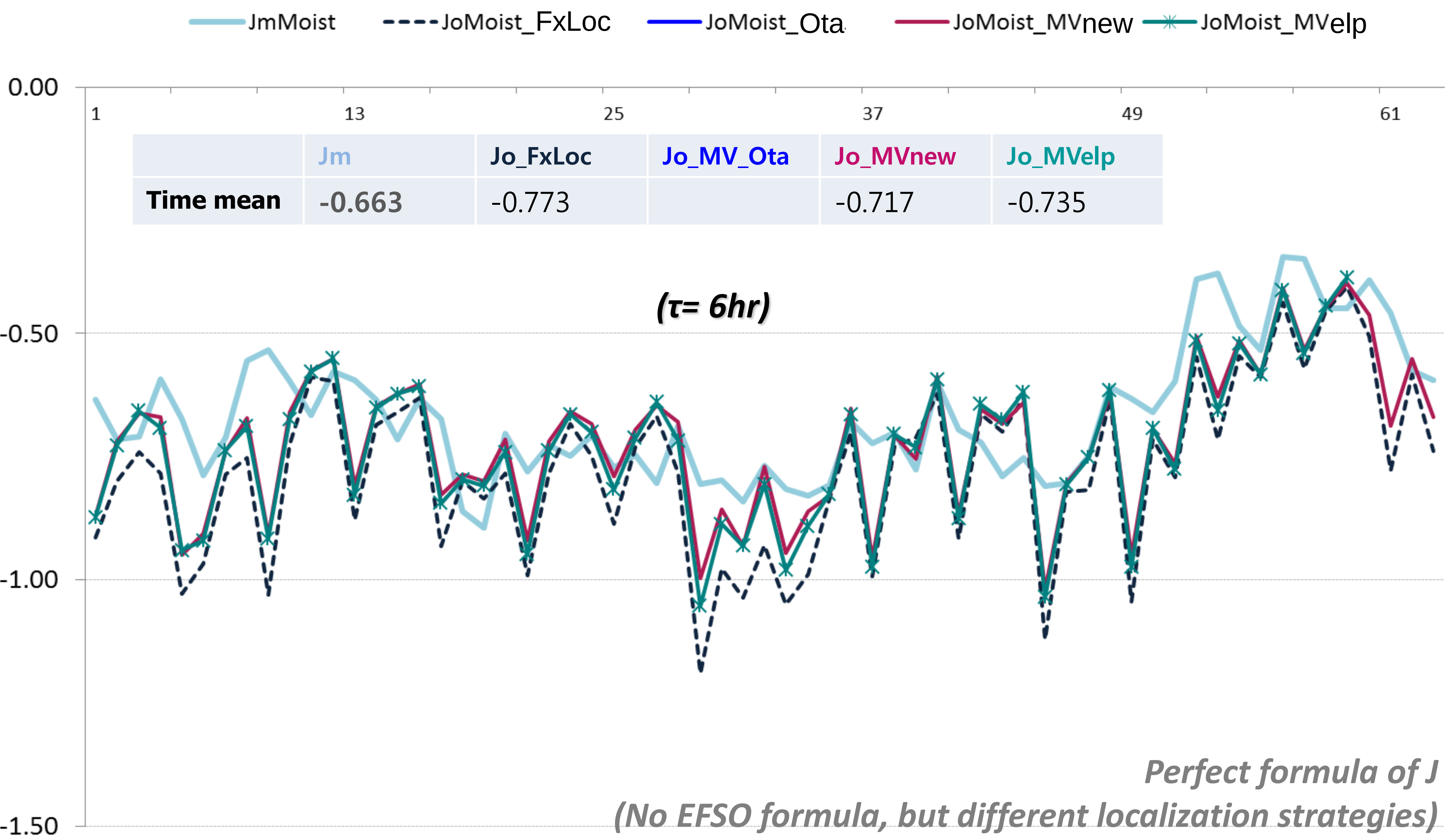
True forecast error reduction J_m



- As expected, forecast error reduction (Jm) becomes greater as the forecast lead time gets longer.
- We will see how the estimations obtained at the observation space (Jo) with different localization methods look.

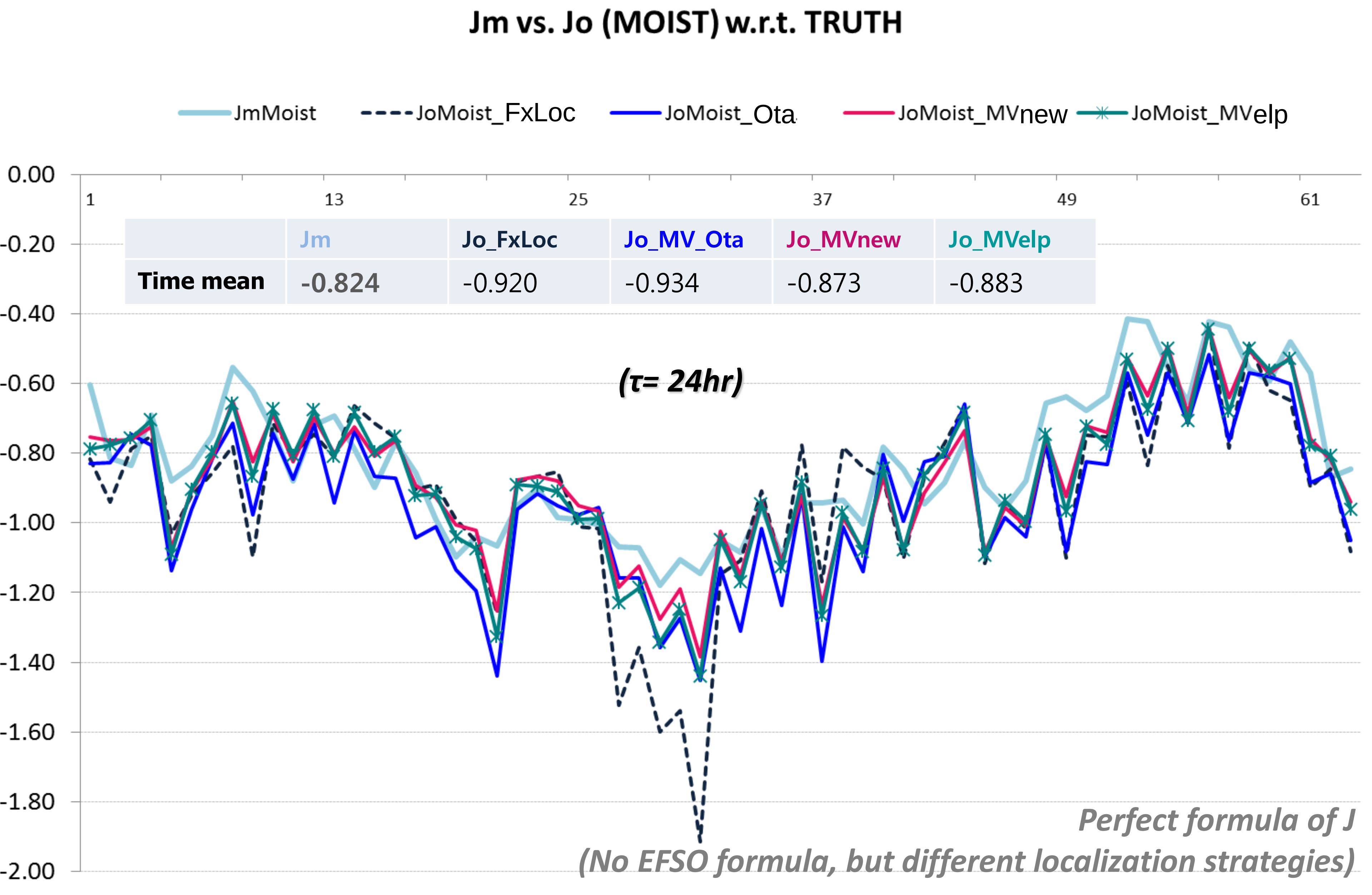
J_m vs. J_o with different localization functions (τ=6hr)

Jm vs. Jo (MOIST) w.r.t. TRUTH [tau=6hr]



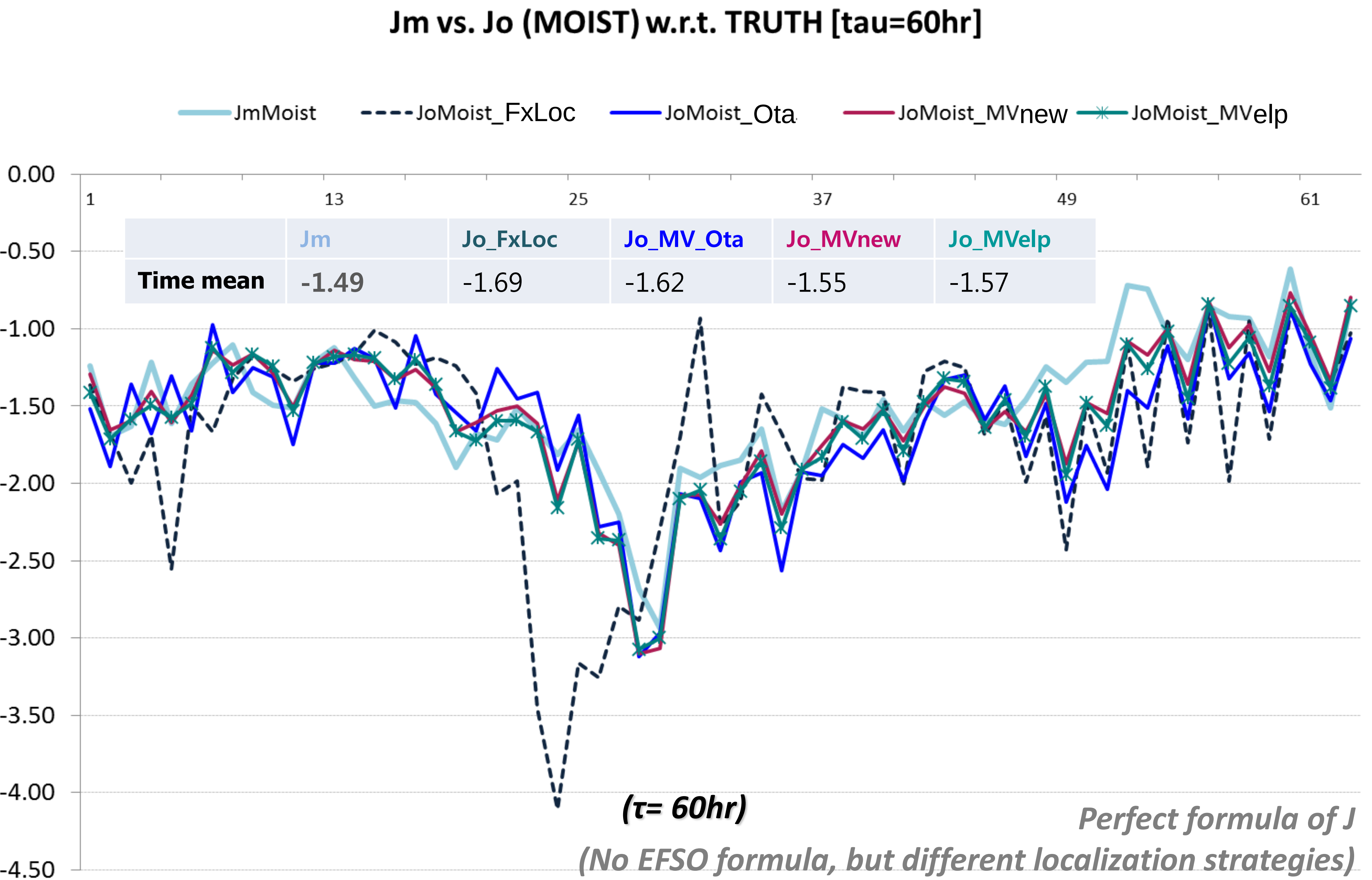
	Jm	Jo_FxLoc	Jo_MV_Ota	Jo_MVnew	Jo_MVelp
Time mean	-0.663	-0.773		-0.717	-0.735

J_m vs. J_o with different localization functions (τ=24hr)



	Jm	Jo_FxLoc	Jo_MV_Ota	Jo_MVnew	Jo_MVelp
Time mean	-0.824	-0.920	-0.934	-0.873	-0.883

J_m vs. J_o with different localization functions (τ=60hr)



	Jm	Jo_FxLoc	Jo_MV_Ota	Jo_MVnew	Jo_MVelp
Time mean	-1.49	-1.69	-1.62	-1.55	-1.57

Perfect formula of J

(No EFSO formula, but different localization strategies)

Summary and Plans

- **Different strategies of localization function for estimating EFSO have been investigated.**
  – We have shown that *tuning localization scales* gives positive impact on the result, using Lorenz model and OSSEs with KIAPS-LETKF.
  – OSSEs of KIAPS-LETKF system will also quantify how much EFSO estimates can be degraded by imperfect forecast error estimates (due to imperfect analysis), EFSO formula, etc.
  – We are generating EFSO estimates now and will analyze the results carefully.

- **We will test this technique in KIAPS-LETKF with real data** *(conventional data, AMSU-A and IASI radiance data, and GPS RO data),* **which are in progress.**