
Author(s)

Noel Chawang¹ and G Kutty¹

¹ Department of Earth and Space Science (ESS), Indian Institute of Space Science and Technology (IIST)

Presented By:

Noel M. Chawang
Email : chawang.noel@gmail.com
Observation impact

Data assimilation is how observations play a role in meteorological applications.

There are various ways of studying the observation's impact, the *Observing System Experiments (OSEs)* is a well-known traditional method.
The main problem that lies with OSEs are that with increasing observation system this method becomes computationally expensive.

Rani et al. (2016)
The EFSO method
New ensemble method for observation impact

With OSEs being so computationally expensive, it brought forth a new method to study the impact observations have on the forecast. Kalnay et al (2012) following the Lui and Kalnay (2008) brought about a manner of finding out the positive or negative impact the observation through a simple formulation for ensemble forecast sensitivity to observation (EFSO).

The ESFO method utilizes the ensemble perturbations as a means to estimate the observation's impact on the forecast.

\[
\Delta e^2 = \frac{1}{K - 1} \delta y^T_0 R^{-1} Y^a_0 X^f_{t|0} (e_{t|0} + e_{t|-6})
\]
EFSO Design

With the new EFSO method you require only 2 forecast runs, one with assimilating the observations and one without to study the impact of all the observations that are being assimilated as long as you have the innovation vector and observation error covariance matrix.
EFSO formulation

We define our errors as such: \( e_{t|0} = x^f_{t|0} - x^a_t \quad e_{t|-6} = x^f_{t|-6} - x^a_t \)

From Langland and Baker the cost function to find positive or negative impact goes as:

\[
J = \Delta e^2 = (e^2_{t|0} - e^2_{t|-6}) = [e^T_{t|0} e_{t|0} - e^T_{t|-6} e_{t|-6}]
\]

And with short range forecast, \((x^f_{t|0} - x^f_{t|-6}) \simeq M(x^a_{0} - x^b_{0|-6})\) and we know \( \delta x_0 = (x^a_0 - x^b_{0|-6}) = K \delta y_0 \)

Here \( \delta y_0 \) is the innovation vector. Placing them in the equations we get;

\[
\Delta e^2 = e^T_{t|0} e_{t|0} - e^T_{t|-6} e_{t|-6} \\
= (e_{t|0} - e_{t|-6})^T (e_{t|0} + e_{t|-6}) \\
\Delta e^2 = [M(x^a_0 - x^b_{0|-6})]^T (e_{t|0} + e_{t|-6}) \\
= [MK \delta y_0]^T (e_{t|0} + e_{t|-6}) \\
\Delta e^2 = \delta y_0^T K^T M^T (e_{t|0} + e_{t|-6})
\]

\[
K = \frac{1}{K-1} X^a_0 X^{aT} H^T R^{-1} \\
Y^a_0 = H X^a_0 \\
MK = \frac{1}{K-1} M X^a_0 X^{aT} H^T R^{-1} \approx \frac{1}{K-1} X^T_{t|0} Y^a_0 R^{-1}
\]

\[
\Delta e^2 \approx \frac{1}{K-1} \delta y_0^T R^{-1} Y^a_0 X^T_{t|0} (e_{t|0} + e_{t|-6})
\]
Forecast metric

For computing the observation impact with EFSO, a suitable forecast metric should be chosen, in this study the forecast metric is KE energy.

\[ KE = \frac{1}{2} \frac{1}{D} \int_D \int_{\sigma_{\text{bottom}}}^{\sigma_{\text{top}}} \left[ u^2 + v^2 \right] d\sigma \, dD \]

Here \( u \) and \( v \) are the zonal and meridional wind components computed over the horizontal domain \( D \) between two levels.

Interpreting cost function \( J \)

\[ \Delta e^2 = e^T_{t|0} e_{t|0} - e^T_{t|-6} e_{t|-6} \]

\[ = (e_{t|0} - e_{t|-6})^T (e_{t|0} + e_{t|-6}) \]

In the EFSO method, the idea is that if the observations have positive impact the forecast error that is computed in the run when the observations are assimilated will be lower than when observations are not assimilated. This means that the observations have a positive impact if the cost function \( J \) (or \( \Delta e^2 \)) is negative and when \( J \) is positive that would mean the observations have a negative impact.
Data Assimilation - Ensemble Kalman Filter

In this study, the Ensemble Adjustment Kalman filter (EAKF) available in DART is utilized. The ensemble kalman filter uses recursive filtering to find the optimal solution for the analysis by sampling the pdf of the forecast and analysis. It follows the updation of each ensemble member by computing the gain matrix $K$.

\[
X^a = X^b + K(y - H(x^b))
\]

$x^a$ : analysis state
$x^b$ : background state
$y$ : observations
$H(x^b)$ : $x^b$ on the observation space
$K$ : Kalman gain matrix

Aksoy 2003
Localization in EFSO method

Localization helps us in limiting the no. of observation taken into account and also helps in giving weights to how impactful the observation is on basis of it’s distance to the gridpoint.

In localization, with a gaussian-type graph we shall be placing weights on the various observations it has in about a certain distance. So, say if we have a limit of 1530km (which is also chosen for this study) which will affect the the grid point, all the observation within this limit will also have a weightage to how much it can affect the gridpoint. The localization function used is the Gaspari Cohn taper function to remain consistent with our data assimilation system.

In the $R$ matrix, we take off diagonal elements to be zero, or as to saw the observations have no correlation to each other, bringing about spurious correlations. Localization is implement to get rid of this spurious correlation that arise.

EFSO method will change and an introduction of a weight matrix $X_i$ with with an element by element multiplication is to take place

$$\Delta e^2 \approx \frac{1}{K-1} (\delta y_0^T \otimes \xi) R^{-1} Y_0^a X_{t|0}^f (e_{t|0} + e_{t|-6})$$
EFSO localization

- Gridpoint
- Observation not considered
- Observation considered

If obs present here, Weight depending on the distance from the gridpoint

Weight = 1 when observation is right over the gridpoint
Weight = 0 when observation is at distance d from the gridpoint

Distance d
Experimental design

Experiment Layout

A detailed study of the SATWIND (Atmospheric wind vectors from geostationary satellite) observations is conducted using the EFSO method.

The SATWIND impact study is conducted over the Bay of Bengal region during intensification period of two tropical cyclones, Hudhud (2014) and Phailin (2013).

For Hudhud this is from 1200 UTC 10 October 2014 to 0600 UTC 12 October 2014; and for TC Phailin (2013) this period starts from 1200 UTC 10 October 2013 to 0600 UTC 12 October 2013.
Observation spread of SATWIND that plays a role in 500hPa.
### Domain and parameters

<table>
<thead>
<tr>
<th>WRF model</th>
<th>3.8.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>(43.125E - 134.875E, 29.135S - 51.345N)</td>
</tr>
<tr>
<td>DA</td>
<td>Dart Lanai - Eakf</td>
</tr>
<tr>
<td>Grid size</td>
<td>27000m</td>
</tr>
<tr>
<td>Microphysics scheme</td>
<td>Option 4 : WRF Single–moment 3–class and 5–class Schemes</td>
</tr>
<tr>
<td>No. of ensembles</td>
<td>80</td>
</tr>
<tr>
<td>Observations assimilated</td>
<td>Radiosonde and dropsondes (ADPUPA), surface station (ADPSFC), aircraft (AIRCFT and ACAR), ships and buoys (SFCSHP) and satellite derived wind products (SATWIND).</td>
</tr>
<tr>
<td>Initial conditions</td>
<td>NCEP Global Data Assimilation System (GDAS) final analysis (FNL)</td>
</tr>
<tr>
<td>EFSO study</td>
<td>SATWIND observations</td>
</tr>
<tr>
<td>Forecast metric</td>
<td>KE energy</td>
</tr>
<tr>
<td>Localization function</td>
<td>Gaspari Cohn function</td>
</tr>
<tr>
<td>Localization limit</td>
<td>1530km horizontally and 1.2lnp vertically</td>
</tr>
<tr>
<td>Forecast run</td>
<td>6h and 12h DA runs, EFSO study</td>
</tr>
</tbody>
</table>

Experiment layout

Gen Ensembles 08 12h Oct → Run Ensembles to 09 00h Oct → DA performed from 0900h Oct → to 10 06h Oct (5 DA cycles)

This experiment layout is the same for TC Hudhud (2014) and TC Phailin (2013) with 5 da cycles performed before the intensification stage.

6h ensemble and mean forecast from the DA cycle and 12h mean forecast from the precious DA cycles
Results

No localization

Forecast metric: Meridional wind

Forecast metric: Zonal wind

500hpa
Localization performed

Slp overlaid with wind vectors and the spatial spread of observation impacts

Right is the observation impact in J/kg units

Bottom is the slp in hPa

TC hudhud 6h SATWIND impact

Total impact during TC Hudhud 2014

Normalised impact during TC Hudhud 2014

TC Hudhud 6h impact

1

2

3
TC Phailin 6h SATWIND impact

TC Phailin 6h impact
Fraction of 6h SATWIND observations showing positive impact
Spatial Impact plot over Max and min impact days

Spatial AMV impact distribution at time of max positive impact
Phailin (18h 10102013) hudhud (00h 12102014)

Spatial AMV impact distribution at time of min positive impact
Phailin (12h 11102013) hudhud (18h 10102014)
Increased forecast lead time SATWIND impact
Targeted domain

Spatial plot of observation impacts over 150 -50 Hpa

Spatial plot of observation impacts over 850 -750 Hpa

Faster computation
Conclusions

The detailed study of the SATWIND observations revealed a lot to us, contrary to our expectation, only a little over half the observations in fact have a positive impact.

During the intensification of the tropical cyclone, it is not only the intensity of the tropical cyclone that plays a factor on the impact of the observations, but mainly the intensification of the tropical cyclone that plays a major role on the observation impact. At points where there is rapid intensification, the already chaotic and dynamically unstable weather event of a TC, makes the observations show low positive or a negative impact altogether.

We also see how the forecast length plays a role in the impact of SATWIND observations, with the decrease of SATWIND impact with increase in forecast length.

Add to all of this, the targeted domain study shows that the position of the observation in a flow also plays a major factor in the impact of the observation. With observations present in the upstream region of a straight flow showing more positive impact and observations present at downstream regions of the flow showing low positive or negative impact altogether.
Thankyou