### Proactive Quality Control based on Ensemble Forecast Sensitivity to Observations

### Daisuke Hotta<sup>1,2</sup>

## Eugenia Kalnay<sup>1</sup>, Yoichiro Ota<sup>2</sup>

<sup>1</sup> University of Maryland <sup>2</sup> Japan Meteorological Agency

July 16<sup>th</sup>, 2014 EMC Seminar

# Outline

- 1. EFSO (Ensemble Forecast Sensitivity to Observations) and "Proactive QC"
- 2. EFSR (Ensemble Forecast Sensitivity to Observation Error Covariance matrix R) and tuning of R
- 3. Future Directions: Operational Applications

(Appendix: semi-implicit Lorenz N-cycle scheme)

## Part 1:

## **EFSO** (Ensemble Forecast Sensitivity to Observations) and "Proactive QC"

## Motivation: The NCEP "forecast skill dropout" problem



- NCEP's 5-day Forecast skill is generally very high (~ 0.9 level)
- However, it occasionally drops to a low level (= "dropout")
- In some cases, all NWP centers suffer.
- But in some cases, NCEP does suffer while ECMWF does not.

## Motivation: The NCEP "forecast skill dropout" problem



- "Culprit" is not the model but "bad observations" (or inability of DA system to properly assimilate them)
- → How can we detect those "flawed" observations?

#### **EFSO: Ensemble Forecast Sensitivity to Observations**



- Quantifies how much each observation improved/degraded the forecast
- First invented for a variational DAsystem using the *adjoint method* by Langland and Baker (2004)
- Liu and Kalnay (2008) adapted it to LETKF (*no adjoint*)
- Kalnay et. al (2012) gave an improved, simpler formulation
- The new formulation is
  - more accurate
  - simpler and easier to implement
  - applicable to any formulation of EnKF
- Ota et al. (2013) successfully implemented the new EFSO into the NCEP's operational GFS system

#### Ota et al. (2013): Identification of "flawed" observations by 24-hour EFSO

Table 3. List of local 24-hour forecast failure cases (initial time from 00 UTC, 8 January 2012, to 18 UTC 7 February 2012)

Initial	Area	Size	Rate	N	Denied observation (denied number/total number)	Change (estimate)
06 UTC JAN 12	50N-80N 145E-175W	1.99	1.36	5	AMSUA ch4, 5, 6 (2735/125063)	-8.7% (-19.5%)
00 UTC JAN 16	30N-60N 20W-0	2.71	1.35	6	GPSRO 600–950 hPa (50/4918)	-0.6% (-4.3%)
18 UTC JAN 27	30S-0 105E-120E	2.40	1.21	1	AIRS (19908/670041)	-0.2% (-6.0%)
00 UTC JAN 30	70S-40S 165E-165W	2.00	1.25	6	AMSUA ch1, 3, 4, 5, 15 (3822/164934)	-4.7% (-12.8%)
06 UTC FEB 2	50N-80N 150W-110W	3.01	1.22	5	GPSRO 250-400 hPa, 600-850 hPa (407/13 092)	-11.7% (-8.7%)
06 UTC FEB 4	30N-60N 150W-130W	1.81	1.26	3	IASI (57950/1177256), HIRS ch3, 4, 9, 11, 12,	-25.5% (-81.6%)
					14, 15 (785/73419), Aircraft 950 hPa ~, 125–600 hPa	
					(5794/100 896)	
18 UTC FEB 6	60N-90N 40E-100E	1.71	1.38	2	MODIS_Wind (10 970/43 452)	-28.4% (-17.7%)

- Identified 7 cases of potential "regional forecast skill dropouts"
- Rerun the analyses and forecasts without using "flawed" obs. identified by 24-hour EFSO
- The forecast errors were substantially reduced.

## "Proactive QC": Proposed Algorithm

Suppose we wish to identify and delete "flawed" obs. at 00h.

- 1 Run regular DA cycle from -06h to 00h.
- 2 Run regular DA cycle from 00h to 06h.
- 3 Detect "regional dropouts" using the information available from (1) and (2).
- 4 Perform 6-hour EFSO to identify "flawed" obs. at 00h.
- 5 If "flawed" obs. are identified, repeat 00h analysis without using the detected "flawed" obs.

# Key questions to be addressed in order for the Proactive QC to work

- Are 6 hours long enough for detecting "flawed observations"?
  - Forecast errors are computed as Forecast-minus-Analysis
  - When compared to the errors of very short-term forecast, analysis errors might not be small.
  - $\rightarrow$  Estimation of forecast errors becomes more difficult.
- What is the best criterion for rejection of observations? Rejecting too many observations might lead to forecast degradation, but too few would make little difference.

 $\rightarrow$  How to strike the best balance?

Does rejection of those observations really improve analysis and forecast?

## Experiments with quasi-operational NCEP's GFS/GSI system Experimental Set-up

(Implemented on top of GFS/hybrid GSI ported to JCSDA's S4 by Dr. Jim Jung)

- Forecast Model: NCEP's GFS model
- **Resolution**: *half* of the operational:
  - T254L64 (deterministic), T126L64 (ensemble)
- DA system: hybrid GSI (as in the operational), but EnKF part replaced by LETKF
- **Observations**: same as the NCEP operational system
- **Period**: 34 days (Jan Feb, 2012)
- LETKF: Covariance localization and inflation (same as the operational)
- EFSO:
  - Localization: same as LETKF + moving localization of Ota et al. (2013)
  - Error norm: Moist total energy norm

# EFSO's sensitivity to forecast lead time (1) Time average

#### Average net observation impact for each observation type



• EFSO results are not very sensitive to the choice of evaluation lead time.

#### EFSO's sensitivity to forecast lead time (2) Individual Cases

Example: MODIS wind near the North Pole on Feb 06 18UTC, 2012

Geographical distribution of EFSO from each. obs.

Red: negative impact ; Blue: positive impact



- Again, EFSO results are not very sensitive to the choice of evaluation lead time, *even for individual cases.*
- → 6 hours are long enough for detection of "flawed" observations.

# Key questions to be answered

Are 6 hours long enough for detecting "flawed observations"?

→ Yes. 6-hr EFSO is equally capable of detecting "flawed" obs. as 24-hr EFSO.

- What is the best criterion for rejection of observations?
- Does rejection of those observations really improve analysis and forecast?

#### Data Denial Experiments Case study (1): MODIS case 2012-Feb-06-18Z, [60°N—90°N] x [40°E—100°E]



- MODIS wind identified as "flawed" (i.e., with net negative impact).
- There are both **helpful** and **harmful** observations.
- How can we decide which / how many obs. should be denied?



Data Denial Experiments

How many obs. should we reject? Case study (1): MODIS case 2012-Feb-06-18Z, [60°N—90°N] x [40°E—100°E] Relative 24-hr fcst. improvement:= ( $e^{f}_{beforeQC} - e^{f}_{afterQC}$ )/  $e^{f}_{beforeQC}$  x 100 [%] Data selection based on 6-hour EFSO



- allobs: overall improvement, but with several areas with degradation
- **allneg**: enhanced improvement, reduced degradation
- one-sigma & netzero: less improvement, but with further reduced degradation

#### Data Denial Experiment Summary of the relative 24-hr forecast improvement for 20 cases (20x4x2=160 experiments)

Case			6-1	hour		24-hour			Case		6-hour			24-hour					
#		allobs	allneg	one- sigma	netzero	allobs	allneg	one- sigma	netzero	#		allobs	allneg	sigma	netzero	allobs	allneg	sigma	netzero
1	max.imp. max.deg. avg.imp.	12% -9% 0.0%	$\begin{array}{c} 11\% \\ -1\% \\ 0.2\% \end{array}$	4% -1% 0.1%	$5\% \\ -1\% \\ 0.1\%$	12% -9% 0.0%	20% -1% 0.3%	0% -1% -0.0%	$6\% \\ 0\% \\ 0.1\%$	11	max.imp. max.deg. avg.imp.	$11\% \\ -6\% \\ 0.5\%$	$9\% \\ -5\% \\ 0.3\%$	$2\% \\ -2\% \\ 0.1\%$	$3\% \\ 0\% \\ 0.1\%$	$\begin{array}{c} 22\% \\ -5\% \\ 0.9\% \end{array}$	$15\% \\ -6\% \\ 0.9\%$	$1\% \\ 0\% \\ 0.0\%$	$2\% \\ 0\% \\ 0.2\%$
2	max.imp. max.deg. avg.imp.	14% -5% -0.1%	$\begin{array}{c} 11\% \\ -4\% \\ 0.3\% \end{array}$	N/A	$4\% \\ 0\% \\ 0.2\%$	N/A	$10\% \\ -5\% \\ 0.1\%$	N/A	$2\% \\ 0\% \\ 0.1\%$	12	max.imp. max.deg. avg.imp.	$37\% \\ -14\% \\ 0.7\%$	$\begin{array}{c} 39\% \\ 12\% \\ 0.7\% \end{array}$	$19\% \\ -2\% \\ 0.5\%$	$\begin{array}{c} 19\% \\ -2\% \\ 0.5\% \end{array}$	$\begin{array}{c} 37\% \\ -14\% \\ 0.7\% \end{array}$	$38\% \\ -19\% \\ 0.4\%$	$1\% \\ 0\% \\ 0.0\%$	$12\% \\ -6\% \\ 0.2\%$
3	max.imp. max.deg. avg.imp.	$13\% \\ -15\% \\ 0.0\%$	$7\% \\ -5\% \\ 0.2\%$	$2\% \\ -1\% \\ 0.0\%$	$4\% \\ -2\% \\ 0.0\%$	7% -8% -0.1%	12% -7% 0.1%	$0\% \\ 0\% \\ 0.0\%$	2% -3% -0.1%	13	max.imp. max.deg. avg.imp.	24 % -9% 1.4%	30% 0% 0.8%	-10% 0.3%	$\begin{array}{c} 19\% \\ -12\% \\ 0.4\% \end{array}$	$\begin{array}{c} 24\% \\ -9\% \\ 1.3\% \end{array}$	26% -10% 1.1%	$0\% \\ 0\% \\ 0.0\%$	$8\% \\ -6\% \\ 0.1\%$
4	max.imp. max.deg. avg.imp.	$25\% \\ -5\% \\ 0.6\%$	$27\% \\ -5\% \\ 0.7\%$	$\begin{array}{c} 15\% \\ -2\% \\ 0.3\% \end{array}$	$13\% \\ -2\% \\ 0.2\%$	3% -6% -0.3%	$4\% \\ -3\% \\ 0.1\%$	1% -5% -0.3%	0% 0% -0.0%	14	max.imp. max.deg. avg.imp.	$5\% \\ 0\% \\ 0.3\%$	$ \begin{array}{c c} 3\% \\ 0\% \\ 0.1\% \end{array} $	$1\% \\ 0\% \\ 0.0\%$	$     \begin{array}{c}       1\% \\       0\% \\       0.1\%     \end{array} $	$5\% \\ 0\% \\ 0.3\%$	$3\% \\ -1\% \\ 0.1\%$	$0\% \\ 0\% \\ 0.0\%$	$0\% \\ 0\% \\ 0.0\%$
5	max.imp. max.deg. avg.imp.	15% -32% -0.2%	19% 81% -0.2%	23% -30% 2%	$22\% \\ -13\% \\ 0.3\%$	12% -78% -1.3%	10% -21% -0.4%	1% -1% -0.0%	1% -1% -0.0%	15	max.imp. max.deg. avg.imp.	$3\% \\ -2\% \\ 0.1\%$	$1\% \\ -1\% \\ 0.1\%$	1% -1% -0.0%	$1\% \\ -1\% \\ 0.0\%$	13% -16% -0.1%	$35\% \\ -18\% \\ 0.8\%$	$1\% \\ -1\% \\ 0.0\%$	$7\% \\ -10\% \\ 0.2\%$
6	max.imp. max.deg. avg.imp.	9% -9% 0.0%	$15\% \\ -6\% \\ 0.4\%$	$12\% \\ -3\% \\ 0.3\%$	$3\% \\ -1\% \\ 0.1\%$	$24\% \\ -38\% \\ 0.0\%$	9% -10% 0.1%	$2\% \\ -2\% \\ 0.0\%$	$3\% \\ -2\% \\ 0.0\%$	16	max.imp. max.deg. avg.imp.	27% -15% 1.9%	$\begin{array}{r} 30\% \\ 21\% \\ 1.8\% \end{array}$	23% -4% 1.3%	$\begin{array}{c} 16\% \\ -2\% \\ 0.7\% \end{array}$	$\begin{array}{c} 30\% \\ -20\% \\ 2.1\% \end{array}$	$33\% \\ -43\% \\ 1.2\%$	$1\% \\ -1\% \\ 0.0\%$	$7\% \\ -1\% \\ 0.3\%$
7	max.imp. max.deg. avg.imp.	17% -9% -0.0%	$13\% \\ -5\% \\ 0.4\%$	$2\% \\ -3\% \\ 0.0\%$	$0\% \\ 0\% \\ 0.0\%$	$\begin{array}{c} 19\% \\ -36\% \\ 0.3\% \end{array}$	$26\% \\ -28\% \\ 0.6\%$	$0\% \\ 0\% \\ 0.0\%$	$4\% \\ -1\% \\ 0.2\%$	17	max.imp. max.deg. avg.imp.	$39 \ 6 \ -15\% \ 0.8\%$	$\begin{array}{c c} 48\% \\ 4\% \\ 2.1\% \end{array}$	$26\% \\ -2\% \\ 1.2\%$	20% -2% 0.8%	$\begin{array}{c} 45\% \\ -15\% \\ 0.7\% \end{array}$	51% -8% 1.6%	0% -1% -0.0%	$15\% \\ -2\% \\ 0.5\%$
8	max.imp. max.deg. avg.imp.	41 % -18% 0.9%	41% -14% 1.1%	21% -5% 0.8%	$10\% \\ -2\% \\ 0.4\%$	41% -18% 0.9%	26% -10% 1.2%	0% 0% -0.0%	$4\% \\ -1\% \\ 0.2\%$	18	max.imp. max.deg. avg.imp.	$\begin{array}{c} 46\% \\ -9\% \\ 2.4\% \end{array}$	46% 2 2%	25% -3% 1.0%	$\begin{array}{c} 21\% \\ -2\% \\ 0.8\% \end{array}$	$36\% \\ -14\% \\ 1.6\%$	$47\% \\ -13\% \\ 2.1\%$	0% -1% -0.0%	$20\% \\ -4\% \\ 0.6\%$
9	max.imp. max.deg. avg.imp.	7% -21% -0.6%	8% -10% -0.4%	8% -3% 00%		3% -2% -0.1%	5% -1% 0.1%	3% -1% 0.0%	3% -1% 0.0%	19	max.imp. max.deg. avg.imp.	$\begin{array}{c} 44.5 \\ -24\% \\ 2.2\% \end{array}$	37% -10% 2.2%	17% -1% 1.0%	$     14\% \\     -1\% \\     1.0\%   $	6% -17% -0.2%	8% -7% 0.2%	$0\% \\ 0\% \\ 0.0\%$	$2\% \\ -1\% \\ 0.0\%$
10	max.imp. max.deg. avg.imp.	25% -6% 1.1%	$19\% \\ -6\% \\ 0.7\%$	N/A	$6\% \\ 0\% \\ 0.2\%$	N/A	$17\% \\ -12\% \\ 0.8\%$	N/A	$2\% \\ 0\% \\ 0.2\%$	20	max.imp. max.deg. avg.imp.	$\begin{array}{c} 12\% \\ -3\% \\ 0.2\% \end{array}$	$ \begin{array}{c c} 10\% \\ -1\% \\ 0.3\% \end{array} $	$5\% \\ -1\% \\ 0.2\%$	$3\% \\ -1\% \\ 0.0\%$	$\begin{array}{c} 12\% \\ -3\% \\ 0.2\% \end{array}$	N/A	1% -2% -0.0%	$9\% \\ -1\% \\ 0.2\%$

- With allneg:
  - Hemispheric-scale forecast error reduced in 18 out of 20 cases.
  - Local improvement over 30% in 7 cases

#### Data Denial Experiment Summary of the results for 20 cases

- Data selection based on 6-hour EFSO:
  - **allobs**: improvement mixed with degradation
  - allneg: enhanced improvement, reduced degradation
     Hemispheric-scale forecast error reduced in 18 out of 20 cases.
     Local improvement over 30% in 7 cases.
  - one-sigma & netzero: diminished improvement, but with further reduced degradation
  - For all of the 7 most successful cases, MODIS wind was identified as "flawed."
- Data selection based on 24-hour EFSO: Similar to 6-hour EFSO, but with less improvements.

# Key questions to be answered

- Are 6 hours long enough for detecting "flawed observations"?
  - → Yes. 6-hr EFSO is equally capable of detecting "flawed" obs. as 24-hr EFSO is.
- What is the best criterion for rejection of observations?
  - → A matter of trade-off: if some degradation is tolerable, "allneg" should be favorable; else "onesigma" or "netzero" should be used.
- Does rejection of those observations really improve analysis and forecast?
  - → Yes, with >30% local improvement in 7 out of 20 cases.

# Summary for **Proactive QC**

- **"Flawed" observations** that potentially lead to forecast skill dropouts **can be detected by EFSO** diagnostics **after only 6 hours** from the analysis.
- Proactive QC does improve forecast and analysis.
- Proactive QC is **innovative**:
  - The first fully flow-dependent QC
  - based on whether observations actually improve/ degrade forecast

## Part 2:

## **EFSR** (Ensemble Forecast Sensitivity to Observation Error Covariance matrix **R**) and **Tuning of R**

# Motivation

- Data Assimilation combines information from background and observations with an "optimal weight."
- The "optimal weight" is determined based on the background -and observation- error covariances
   B and R.
- In EnKF, B (=P<sup>b</sup>) is dynamically estimated, but R is still an external parameter.
  - Truth is unknown.  $\rightarrow$  True **R** is also unknown.
  - NWP centers specify it empirically and subjectively.
- $\rightarrow$  We need a systematic method for tuning **R**.

# **EFSR Formulation**

- Daescu and Langland (2013) proposed an *adjoint-based* formulation of forecast sensitivity to R matrix.
- We can formulate an *ensemble* version based on **EFSO** by Kalnay et al. (2012) :

$$\left[\frac{\partial e}{\partial \mathbf{R}}\right]_{ij} \approx \frac{\partial e}{\partial y_i} z_j \approx -\frac{1}{K-1} \left[ \mathbf{R}^{-1} \mathbf{Y_0^a} \mathbf{X_t^{fT}}_{\mathbf{t}|\mathbf{0}} \mathbf{C} \left( \mathbf{e_{t|0}} + \mathbf{e_{t|-6}} \right) \right]_i \left[ \mathbf{R}^{-1} \delta y^{oa} \right]_j$$

- We know whether fcst. will be improved or degraded by the increase or decrease of **R**.
- $\rightarrow$  We can optimize **R**.

# **EFSR: Experiments**

- Perfect-model experiment with Lorenz '96 system
  - Run two DA cycles, one with incorrect **R**, the other with correct **R**
  - Perform EFSR to the two experiments. Examine if EFSR can detect mis-specification of **R**.
- Real NWP system experiment & Tuning of **R** 
  - Diagnose forecast **R**-sensitivity for each observation type by EFSR.
  - Tune **R** based on EFSR and run the DA cycle again.
     Examine if the tuning improves the EFSO impacts of the tuned observation types.

# Perfect-model Experiment: Experimental Setup

- Model: Lorenz '96 model with N=40 and F=8.0  $\frac{\mathrm{d}x_j}{\mathrm{d}t} = x_j (x_{j+1} - x_{j-2}) - x_j + F_j$
- **DA method**: 40 member LETKF, no localization
- EFSR: no localization
- Observations: available at every grid point.
- Specification of R:

Name	True obs error variance	Prescribed error variance
SPIKE	$\sigma_j^{o,\text{true}^2} = \begin{cases} 0.8^2 & j = 11\\ 0.2^2 & j \neq 11 \end{cases}$	$\sigma_j^{o2} = 0.2^2$ everywhere
STAGGERED	$\sigma_j^{o,\text{true}^2} = \begin{cases} 0.1^2 & j: \text{ odd} \\ 0.3^2 & j: \text{ even} \end{cases}$	$\sigma_j^{o2} = 0.2^2$ everywhere
LAND-OCEAN	$\sigma_{j}^{o,\text{true}^{2}} = \begin{cases} 1 \le j \le 20\\ 0.3^{2} & (\text{``land''})\\ 21 \le j \le 40\\ 0.1^{2} & (\text{``ocean''}) \end{cases}$	$\sigma_j^{o2} = 0.2^2$ everywhere

- Erroneous obs. variance only at the 11-th grid pt.
- DA system assumes constant **R** for all grid pts.

Design is inspired by Liu and Kalnay (2008)

## Perfect-model Experiment: **Result (SPIKE experiment)**



• For "incorrect-R," EFSR detects the mis-specification of R at the 11<sup>th</sup> grid point.

 $\rightarrow$  We can detect mis-specified **R** 

• For "correct-R," EFSR diagnoses almost-zero sensitivity.  $\rightarrow$  No "false alerts" 26

### EFSR for GFS / GSI-LETKF hybrid



- Aircraft, Radiosonde and AMSU-A: large positive sensitivity
- **MODIS wind** : negative sensitivity
- → Tuning experiment:
  - Aircraft, Radiosonde and AMSU-A: reduce **R** by 0.9
  - MODIS wind: increase **R** by 1.1

## Tuning Experiment: Result EFSO **before/after** tuning of **R**



- Aircraft, Radiosonde, AMSU-A:
  - significant improvement of EFSO-impact
     (as expected)
- MODIS wind :
  - No improvement in EFSO (contrary to expectation)

#### Why no improvement in MODIS?

- MODIS had "flawed" obs. along with "helpful" obs.
- The "flawed" obs. might have resulted in incorrect estimation of EFSR.

# Excluding cases where MODIS wind had negative impact



- MODIS wind exhibited several negatively-impacting cases.
- Exclude negative cases
- → EFSR for MODIS becomes neutral
- → Consistent with the result of tuning experiment

Lesson:

Before performing
 EFSR, we should
 remove "bad" obs.

# Summary for EFSR

- EFSR gives information on whether we should increase/reduce prescribed observation error covariance **R**.
- Tuning of **R** based on this diagnostics improves the EFSO.

•  $\rightarrow$  EFSR can be used to *systematically* optimize **R**.

# **Future Directions:**

# Future plans

Immediate future:

- Implementation of Proactive QC into the real operational system
  - Can the operational system
    - 1. wait for 6 hours?
    - 2. afford to do analysis again?

Long-term Future Directions:

- Applications of EFSO and EFSR
  - Collaboration with instrument developers
  - Acceleration of development for assimilation of new observing systems

Implementation to the real operational system (1) Can we wait for 6 hours?

# Idea: Exploit the time lag between "early analysis" and "cycle (final) analysis"

(suggested by Dr. John Derber, 2013)



**cycle (final) analysis**: maintains analysis-forecast cycle Time **early analysis**: provides initial condition for extended forecast

Implementation to the real operational system (1) we don't need to wait 6 hours!

# Idea: Exploit the time lag between "early analysis" and "cycle (final) analysis"

(suggested by Dr. John Derber, 2013)



**cycle (final) analysis**: maintains analysis-forecast cycle Time **early analysis:** provides initial condition for extended forecast

Implementation to the real operational system (2) can we afford to do analysis twice?

Idea: Use **approximated analysis** rather than doing analysis again:

- Using the approximation to Kalman gain:

$$\mathbf{K} \approx \frac{1}{K-1} \mathbf{X}_0^a \mathbf{X}_0^{aT} \mathbf{H}^T \mathbf{R}^{-1} \approx \frac{1}{K-1} \mathbf{X}_0^a \mathbf{Y}_0^{aT} \mathbf{R}^{-1}$$

the change in analysis by the denial of observations can be approximated by:

$$\bar{\mathbf{x}}_{0}^{a,\text{deny}} - \bar{\mathbf{x}}_{0}^{a} \approx -\mathbf{K}\delta\bar{\mathbf{y}}_{0}^{ob,\text{deny}} \approx -\frac{1}{K-1}\mathbf{X}_{0}^{a}\mathbf{Y}_{0}^{aT}\mathbf{R}^{-1}\delta\bar{\mathbf{y}}_{0}^{ob,\text{deny}}$$

As inexpensive as EFSO.

#### $\rightarrow$ No need to repeat analysis

 $\rightarrow$  Can minimize the time delay

#### Application of EFSO: (1) Collaboration with instrument developers

- In our experiments, in several cases, MODIS wind showed large negative impacts that caused "regional dropouts."
- → EFSO can be used to build a database of "flawed" observations along with their relevant metadata.
  - Provide such database to instrument developers so that they can fix problems with their algorithms.
  - Collaborate with instrument developers to determine which metadata would be helpful to them.

#### Application of EFSO: (2) Acceleration of development for assimilation of new observing systems

- Traditional approach: compare
  - Test: with a new observing system
  - Control: without a new observing system
- Difficulty with this approach:
  - Signals from a new observing system is obscured by the many observations that are already assimilated in Control.
  - $\rightarrow$  Hard to obtain statistically significant results
- $\rightarrow$  EFSO-based data selection will enable efficient determination of an optimal way to assimilate new observing systems.
- An optimal specification of **R** is also a difficult issue for assimilation of new observing systems.
- $\rightarrow$  Our EFSR diagnostics should provide useful guidance.

# Conclusions

- 6-hour EFSO can successfully identify "flawed" observations.
- Rejecting them (Proactive QC) does improve the analysis and forecast.
- Database of "flawed" obs. can help instrument developers to improve their algorithms.
- Proactive QC can readily be implemented into the operational system.
- **EFSR** enables systematic tuning of **R** matrix.
- **EFSO and EFSR** together can accelerate development of the assimilation of new observing systems.

# Acknowledgements

- My advisor Prof. Eugenia Kalnay for advice and encouragement
- **Prof. Takemasa Miyoshi** of RIKEN/AICS and University of Maryland for guidance and help
- JCSDA and its Director Dr. Sid Boukabara for access to the S4 supercomputer
- **Dr. Jim Jung** of NCEP for kind guidance on S4
- JPSS program for support
- Mr. Yoichiro Ota for EFSO and GFS-LETKF code
- Government of Japan, in particular Japan Meteorological Agency (JMA) for funding and support
- **Dr. Jeff Whitaker, Dr. Sajal Kar and Dr. Jim Purser** for discussion and advice (on the Appendix)
- ... and many others and all of you.

