# Adjoint-based Forecast Sensitivity to Observations (FSO)

Wei Sun

July 22–July 24, 2019
National Center for Atmospheric Research
Boulder, CO

# FSO – Outline

- Introduction
- Implementation in WRF/WRFDA
- Applications
- Limitations
- Conclusions

# Introduction

- What
  - Diagnostics tool, quantitatively determine which observations improved or degraded the forecast
  - Using adjoint technique, trace back to observations used in the analysis
  - Impact of observations has traditionally been assessed with observing system experiments (OSEs), while FSO complements the OSEs
- Why
- Who
- How
- How much

# Introduction

- What
- Why
- Who
- How
- How much

Two ways of OSEs :

a) baseline with a minimum amount of observation

add at least one more observation type

b) baseline with full observations

removing one particular or various datasets

# Introduction

- What
- Why
- Who
- How
- How much

| Type of Data | Description |
|---|---|
| OZONE | Satellite ozone retrieval |
| GOES | Geostationary satellite infrared sounder radiances |
| Meteosat | Geostationary satellite infrared sounder radiances |
| AMSU-B | Satellite microwave sounder radiances related to $H$ |
| SSMI-TCWV | Satellite microwave imager radiances related to clouds and precipitation |
| SSMI | Satellite microwave imager radiances related to $H$ and surface wind speed |
| AIRS | Satellite infrared sounder radiances related to $H$ and $T$ |
| AMSU-A | Satellite microwave sounder radiances related to $T$ |
| HIRS | Satellite infrared radiances |
| ERS-QuikSCAT | Satellite microwave scatterometer |
| AMVs | Atmospheric Motion Vectors derived from satellite cloud imagery |
| GPS-RO | Satellite GPS radio occultation |
| PILOT | Sondes and American, European and Japanese wind profiler ($u$, $v$) |
| TEMP | Radiosondes from land and ship measuring $p$, $T$, $RH$, $u$ and $v$ |
| AIREP | Aircraft measurements of $T$, $u$ and $v$ |
| DRIBU | Drifting buoy measuring $p$, $T$, $RH$, $u$ and $v$ |
| SYNOP | Surface observations from land and ship stations: measuring $p$, $RH$, $u$ and $v$ |

Operational datasets in the OSE control experiment for summer 2006 and winter 2007 (Kelly and Thépaut, 2007).

- **require considerable computational resources**
- **DA and forecast system must be run independently to evaluate the impact of each observation**

# Introduction

Ronald Gelaro & Yanqiu Zhu (2009) Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models, Tellus A: Dynamic Meteorology and Oceanography, 61:2, 179-193, DOI: 10.1111/j.1600-0870.2008.00388.x

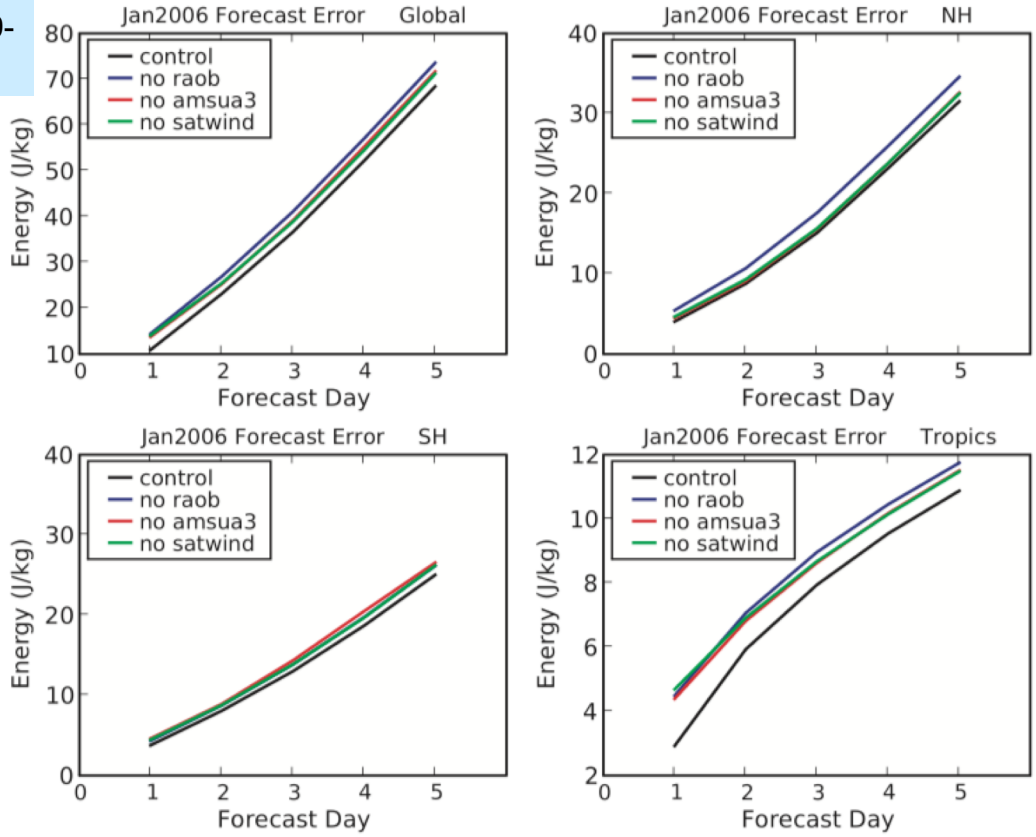



*Fig. 5.* Average values of $e$ over the globe (upper left-hand side), Northern Hemisphere (upper right-hand side), Southern Hemisphere (lower left-hand side) and tropics (lower right-hand side) as a function of forecast length for various OSEs during January 2006.

# Introduction

- What
- Why
- Who
- How
- How much

- Calculate the impact of each observation simultaneously, only need one run
- NWP centers use FSO routinely to monitor their Data Assimilation and Global Observing System
- Help access impact of specific sensors, tune quality control, bias correction

# Introduction

- What
- Why
- Who
- How
- How much

- Naval Research Laboratory (Monterey, CA)
- NASA/GMAO (Washington, DC)
- ECMWF (Reading, UK)
- Environment Canada (Montreal, Canada)
- Meteo-France (Toulouse, France)
- NCAR/MMM (Boulder, CO)

# Introduction

- What
- Why
- Who
- How
- How much

- Non-Linear (NL) forecast models can be linearized (with simplifications).
- The resulting **Tangent-Linear** (TL) represents the linear evolution of small **perturbations**.
- The mathematical transpose of the TL code is called the Adjoint (ADJ) and it transports **sensitivities** back in time.

- The ADJ of the Data Assimilation system is needed to compute the sensitivity to observations It can be computed with various methods:
  - Ensemble (ETKF, Bishop *et al.* 2001)
  - Dual approach (PSAS, Baker and Daley 2000, Pellerin *et al.* 2007)
  - Exact ADJ calculation (Zhu and Gelaro 2007)
  - Hessian approximation (Cardinali 2006)
  - Lanczos minimization (Fisher 1997, Tremolet 2008)

# Introduction

- What
- Why
- Who
- How
- How much

- 2 runs of non-linear forecast model
- 2 runs of adjoint model
- 1 run of adjoint of analysis
- The computer cost is estimated to 10-15 times the cost of the forecast model.

# FSO – Outline

- Introduction
- Implementation in WRF/WRFDA
- Applications
- Limitations
- Conclusions

# Implementation in WRF/WRFDA

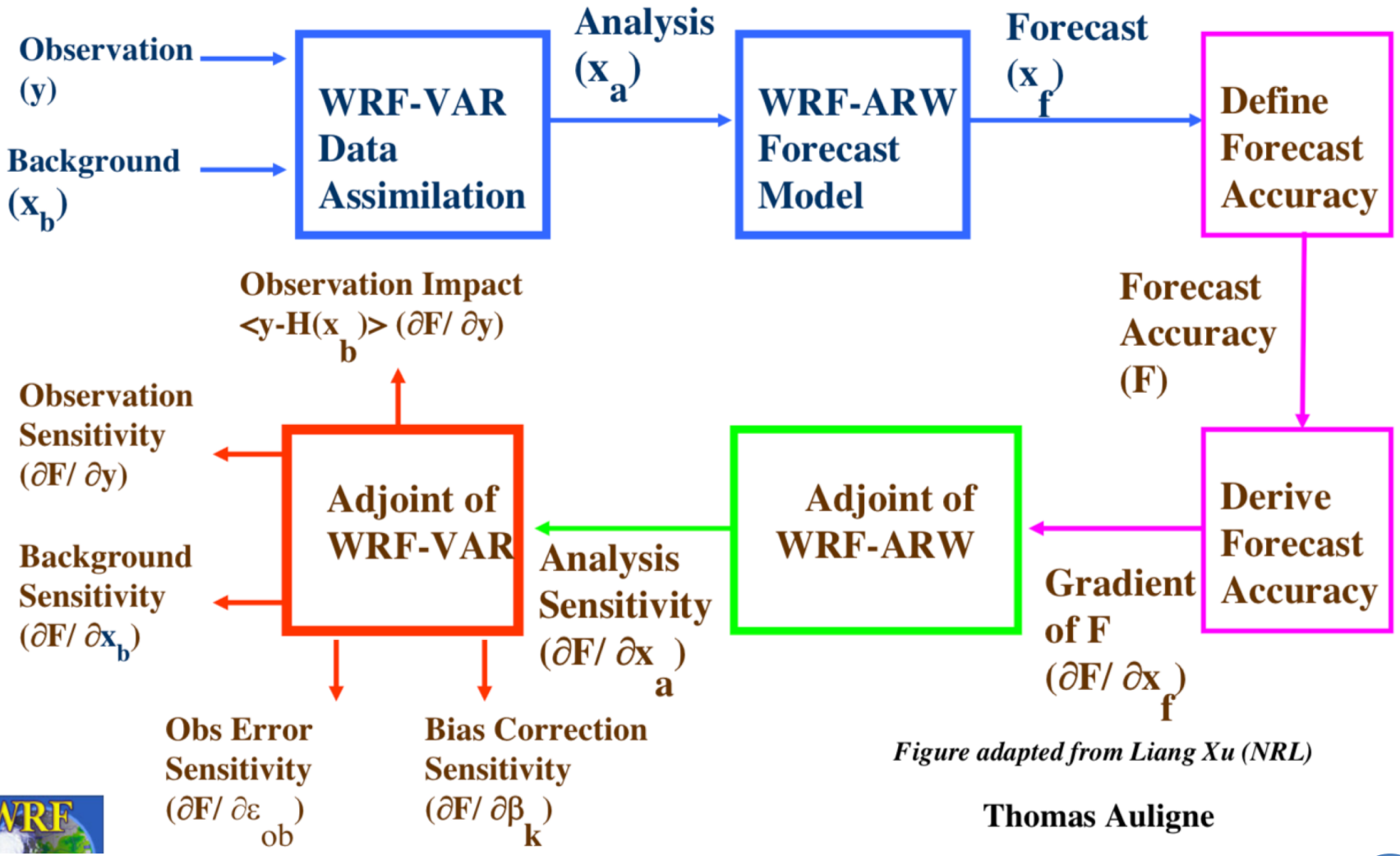
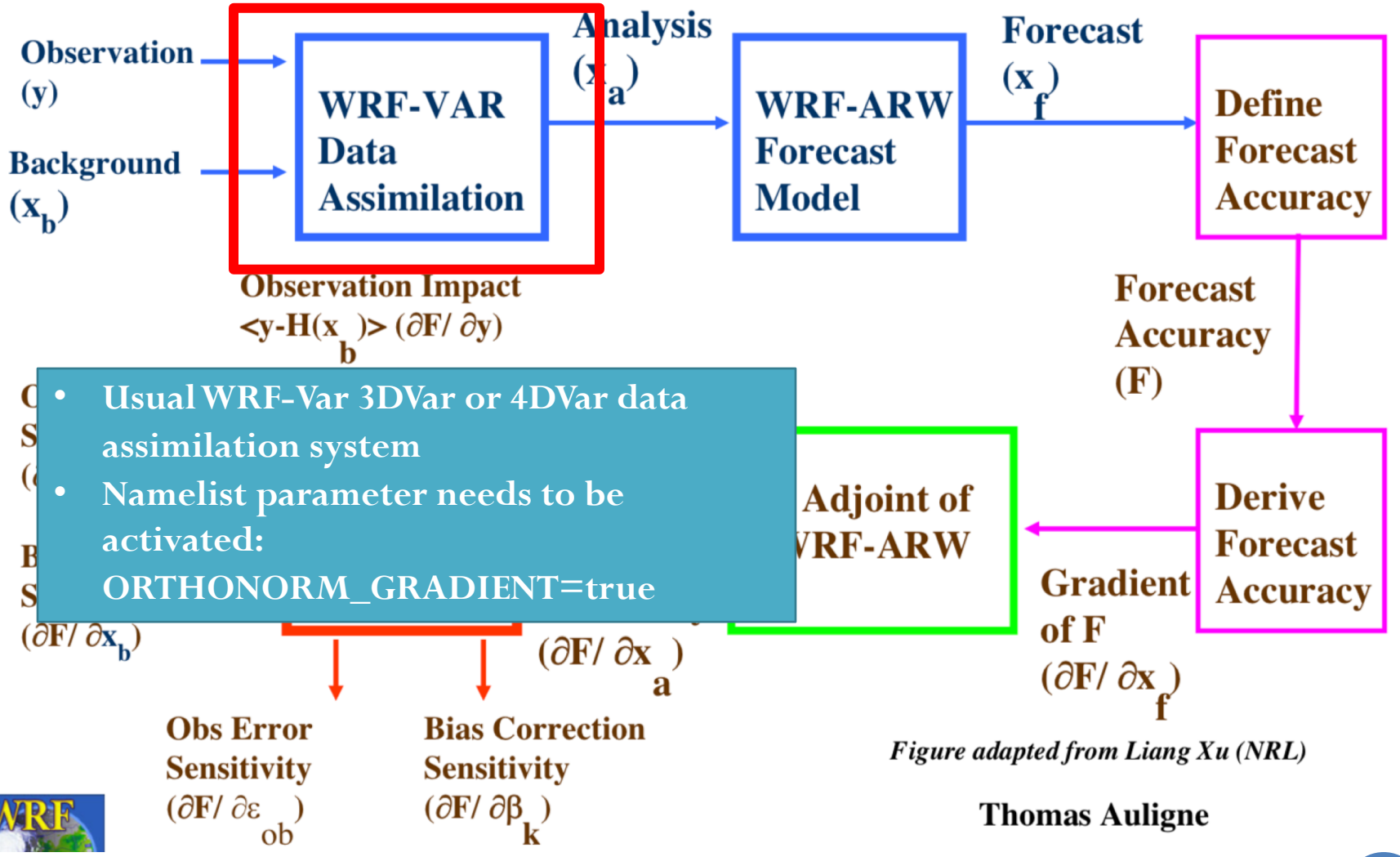



*Figure adapted from Liang Xu (NRL)*

Thomas Auligne

# Implementation in WRF/WRFDA

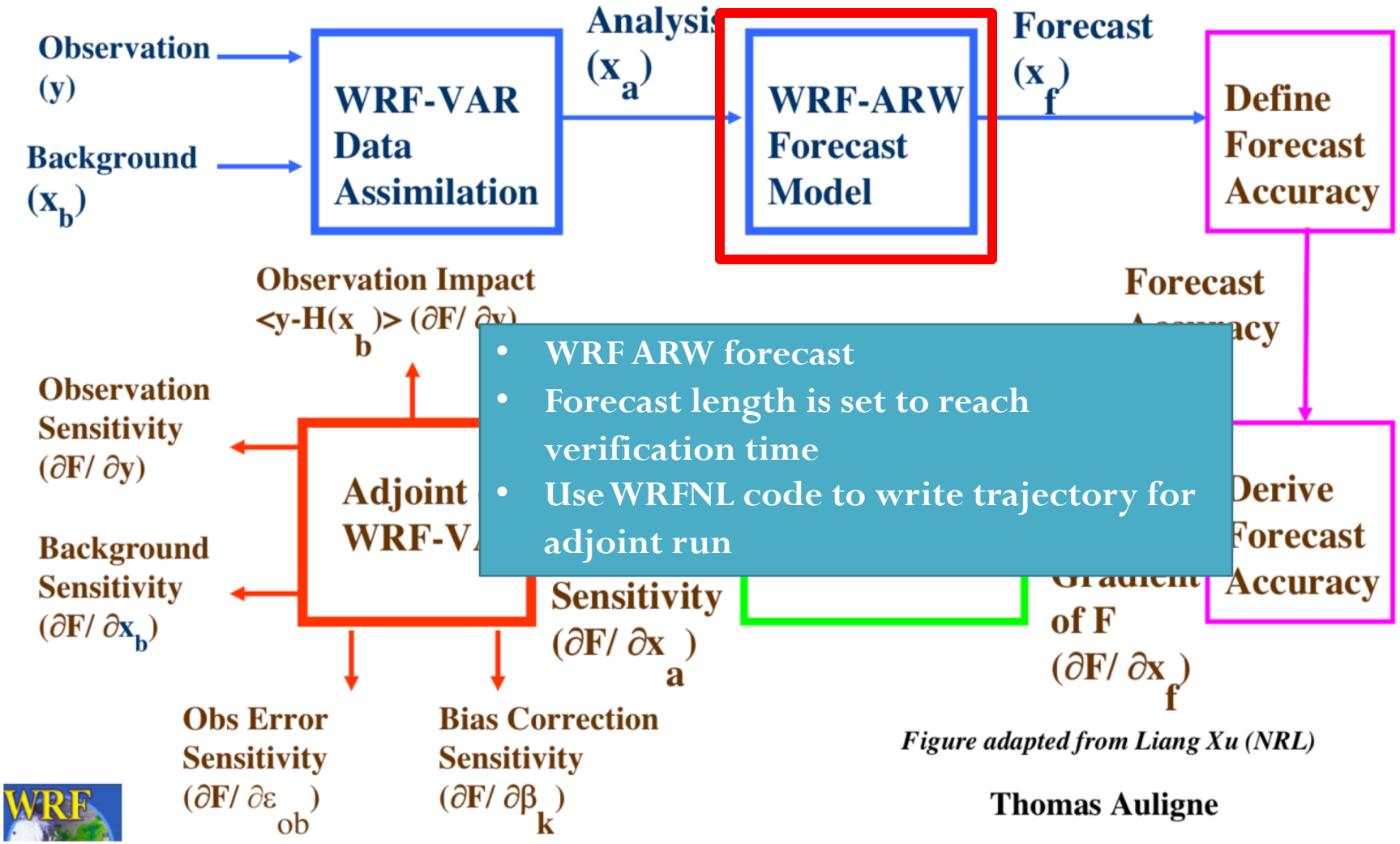



- Usual WRF-Var 3DVar or 4DVar data assimilation system
- Namelist parameter needs to be activated: ORTHONORM_GRADIENT=true

*Figure adapted from Liang Xu (NRL)*

Thomas Auligne

# Implementation in WRF/WRFDA




- WRF ARW forecast
- Forecast length is set to reach verification time
- Use WRFNL code to write trajectory for adjoint run

*Figure adapted from Liang Xu (NRL)*

Thomas Auligne

# Implementation in WRF/WRFDA

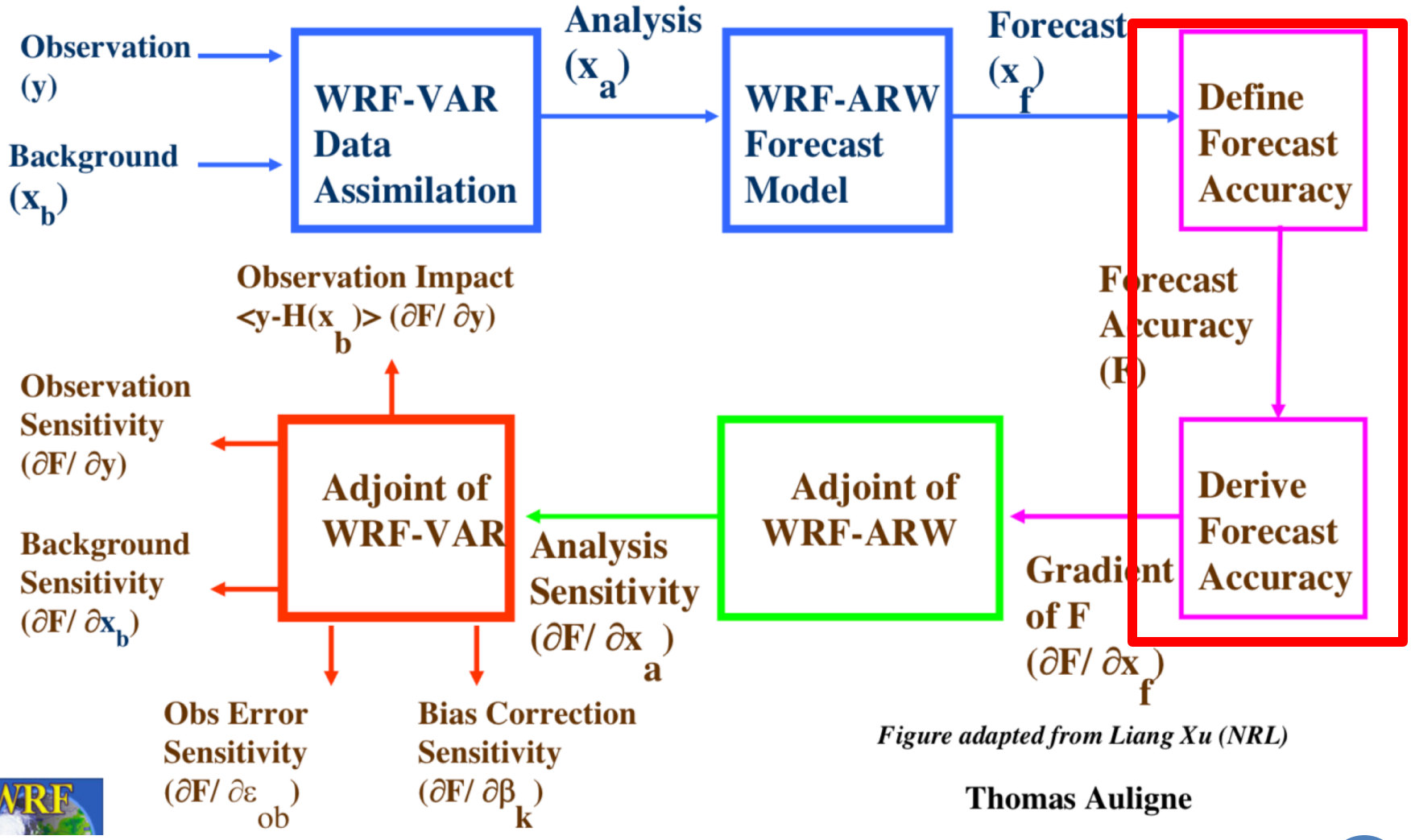
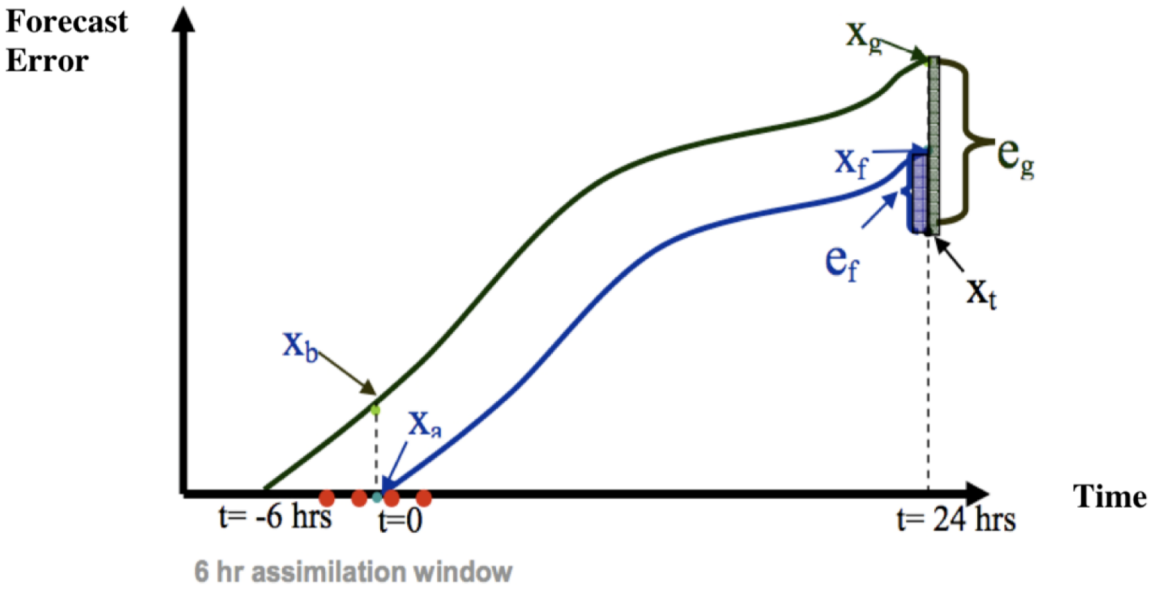



*Figure adapted from Liang Xu (NRL)*

Thomas Auligne

# Implementation in WRF/WRFDA

*From Langland and Baker (2004)*




*$x_t$ is the true state, estimated by the analysis at the time of the forecast*
*$x_f$ is the forecast from analysis xa*
*$x_g$ is the forecast from first-guess at the time of the analysis $x_a$*

**Impact of analysis: $F = De^{f,g} = e^f - e^g$**

# Implementation in WRF/WRFDA

- **Reference state**: Namelist `ADJ_REF` is defined as
  - 1: $X^t$ = Own (WRFVar) analysis
  - 2: $X^t$ = NCEP (global GSI) analysis
  - 3: $X^t$ = Observations

- **Forecast Aspect**: depends on reference state
  - 1 and 2: Total Dry Energy

$$< \mathbf{x}, \mathbf{x} >= \frac{1}{2} \iiint_{\Sigma} [u'^2 + v'^2 + \left( \frac{g}{\bar{N}\bar{\theta}} \right)^2 \theta'^2 + \left( \frac{1}{\bar{\rho} c_s} \right)^2 p'^2] d\Sigma$$

  - 3: WRFVar Observation Cost Function: $J_o$

- **Geo. projection**: Script option for box (default = whole domain)
  *ADJ_ISTART, ADJ_IEND, ADJ_JSTART, ADJ_END, ADJ_KSTART, ADJ_KEND*

- **Forecast Accuracy Norm**: $e = (x^f - x^t)^T C (x^f - x^t)$

# Implementation in WRF/WRFDA

- Use WRF+ code to compute WRF-ARW adjoint with namelist ADJ_ENS=true
- WRF+ is run for both trajectories from xb and xa
- both sensitivities are added together
- Script variable ADJ_MEASURE defined as:
  - 1: first order
  - 2: second order
  - 3: third order
  - 4: variant of third order

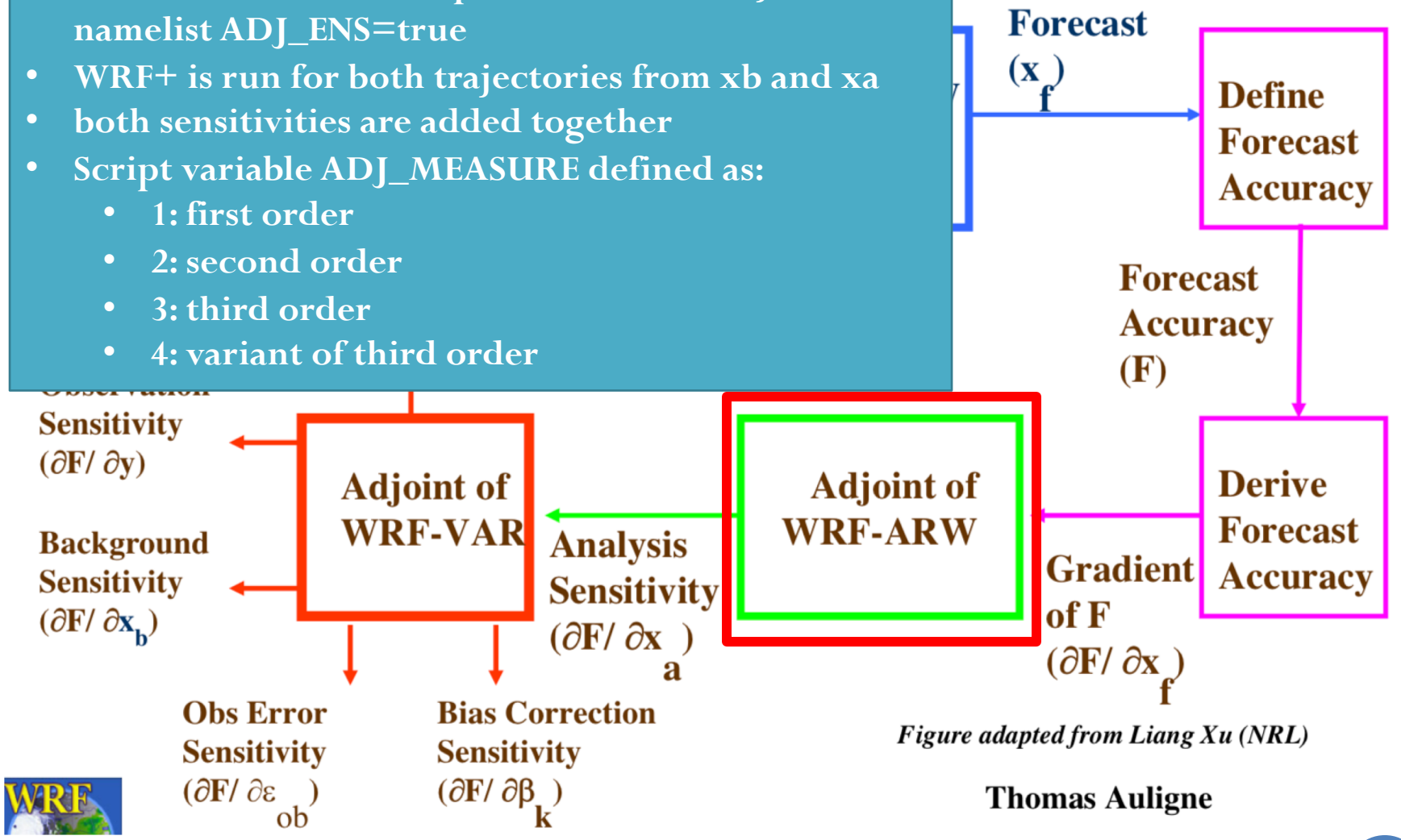



*Figure adapted from Liang Xu (NRL)*

Thomas Auligne

# Implementation in WRF/WRFDA

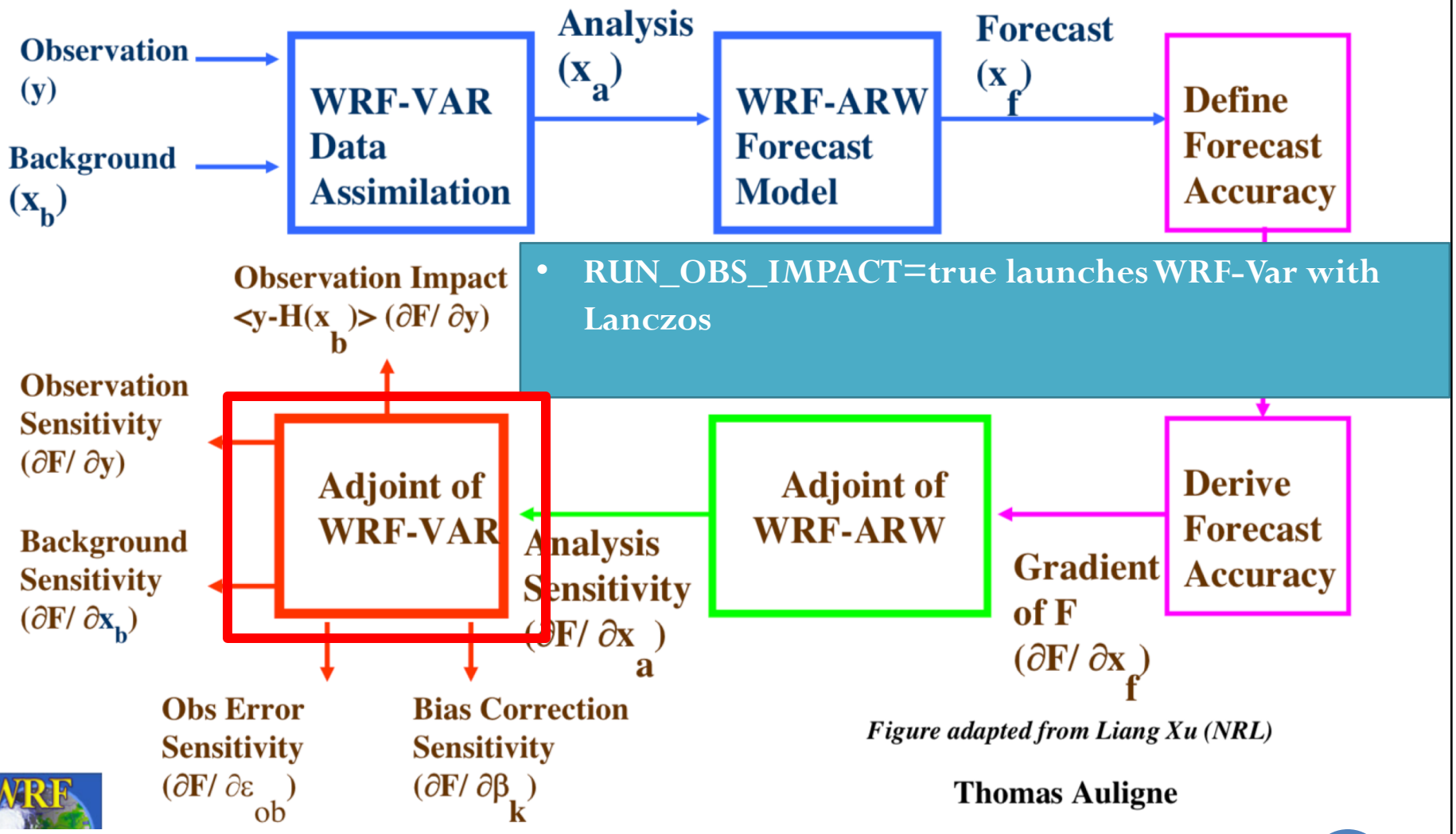



*Figure adapted from Liang Xu (NRL)*

Thomas Auligne

# FSO – Outline

- Introduction
- Implementation in WRF/WRFDA
- Applications
- Limitations
- Conclusions

# Applications

- Install WRF
- Install WRFDA
- Install WRFPLUS
- Download WRFDA_TOOLS

https://github.com/wrf-model/WRFDA_TOOLS

# Applications

- A wrapper script "wrapper_run_fso.ksh" is provided to call the necessary scripts in order for FSO.

```
# 1 Decide which stages to run (run if true):
#-----------------------------------------
# Step 1: run assimilation at initial time (parallel)
# run assimilation at final time if verifying against own analysis (parallel)
export RUN_WRFVAR=true
export RUN_UPDATE_BC=true

# Step 2: run WRF Non-Linear and Adjoint Model (parallel)
export RUN_ADJ_SENS=true

# Step 3: run adjoint of the Data Assimilation at initial time (parallel)
export RUN_OBS_IMPACT=true
export NL_USE_LANCZOS=false
export NL_WRITE_LANCZOS=true
export NL_EPS=1E-5
```

step 1 in the flowchart

step 2-4 in the flowchart

step 5 in the flowchart

Step 1. The log file will be located in "run/2007010100/wrfvar/".

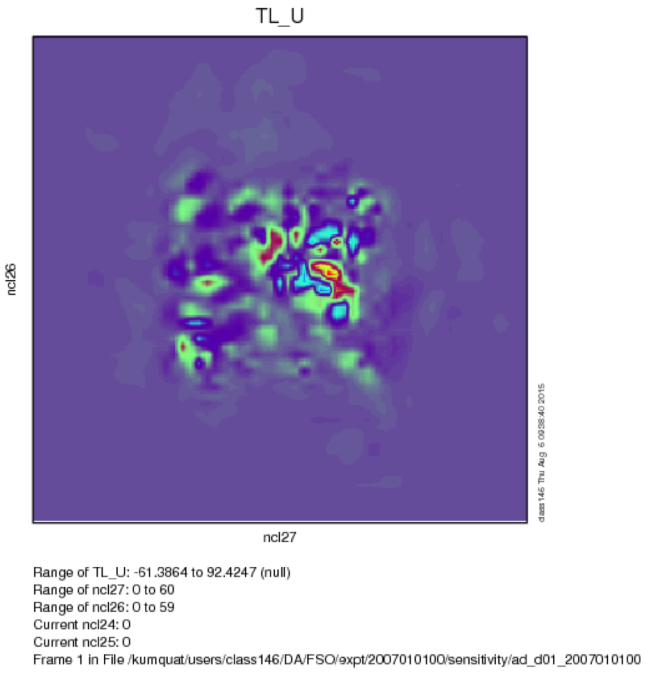
Step 2. The log file will be located in "run/2007010100/sensitivity/".

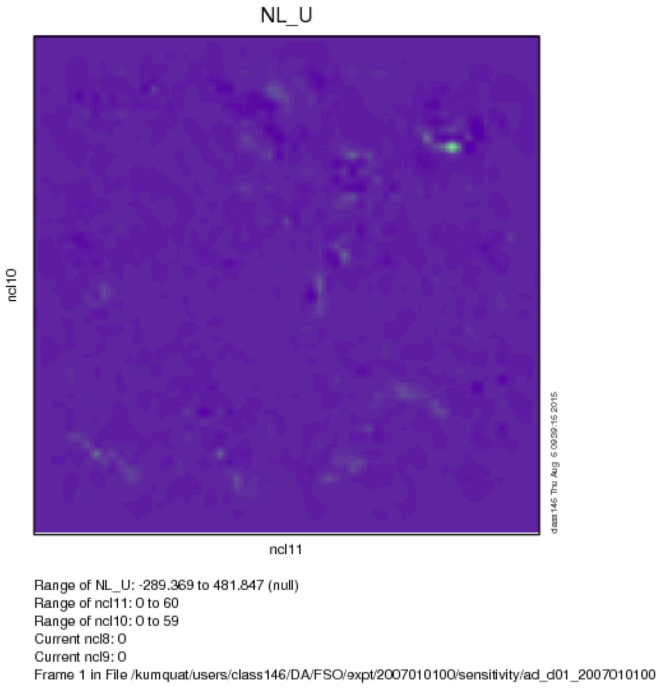
Step 3. The log file will be located in "run/2007010100/obsimpact/".

# Applications

## Check Output

- The error trapping in FSO is not perfect, so for certain problems the system may still report "SUCCESS". Make sure all the files below were created, or else you may have set some variables incorrectly.
- fso/run/2007010100/sensitivity/ad_d01_2007010100 is the **gradient of the forecast error norm to the initial conditions**.



# Applications

Check Output

- fso/run/2007010100/sensitivity/index.html contains the text output from running var/graphics/ncl/adj_diagnostics.ncl to get the **diagnostics of error reduction in the forecasts from Xb (background) to Xa (analysis)**.

- fso/run/2007010100/obsimpact/rsl.out.0000 contains the **cumulated impact of each observation type on the forecast errors**.

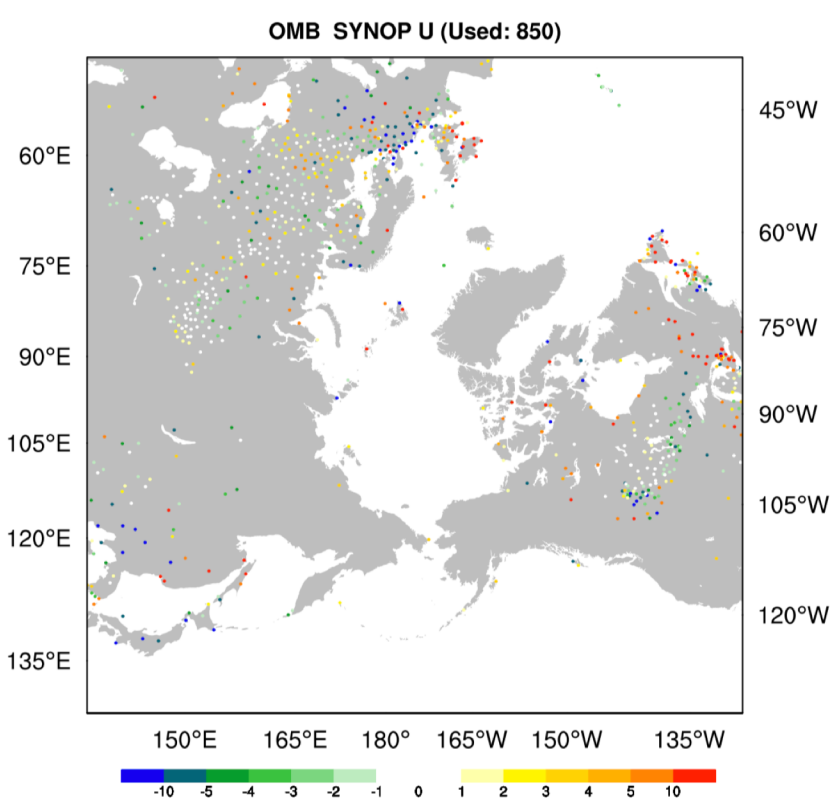
# Applications

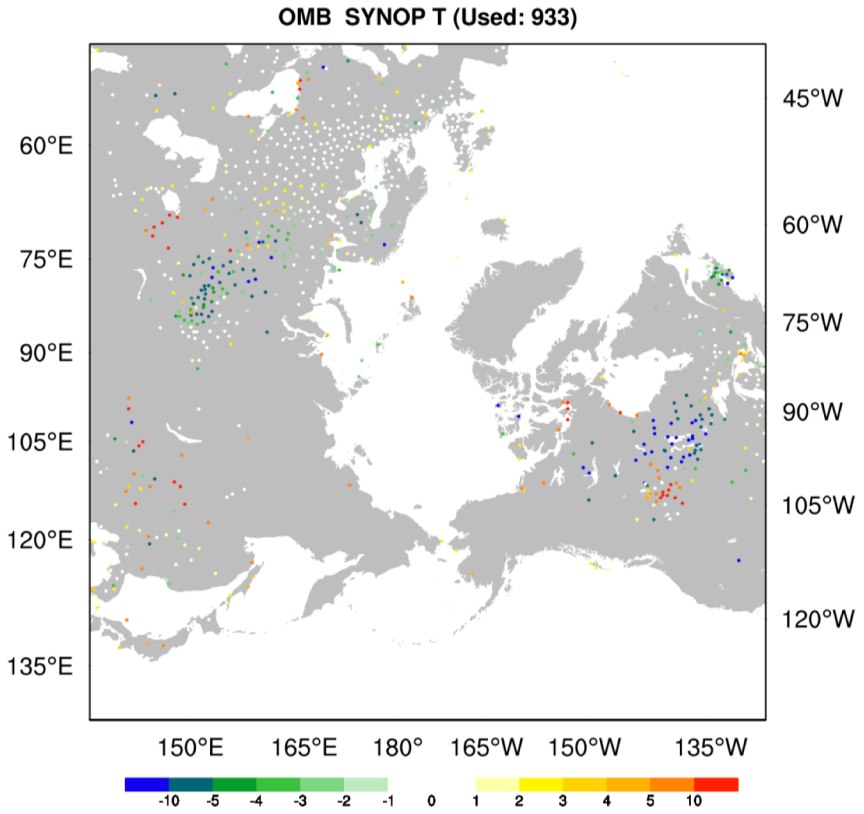
*fso/plot/plot_gts_omb_oma.ncl* will output plots of the observations used for the experiment by type and pressure level

```
;---------------------------------------------------------------------------
; plot options
;===========================================================================
;
; NOTE: actual file to read is gts_fullname
; gts_fullname = datdir1+date+datdir2+filename_gts will be defined later in the main script
; search for gts_fullname and modify it for your own directory structure
plotdir      = "/glade/work/weisun/2019/fso_aws/fso/plot/"
expt         = ""  ; for output naming purpose
datdir1      = "/glade/work/weisun/2019/fso_aws/fso/run/"        ; the path before DATE
datdir2      = "/obsimpact/working/"                      ; the path after  DATE
filename_gts   = "gts_omb_oma_01"
; NOTE: actual file to read is gts_fullname
; gts_fullname = datdir1+date+datdir2+filename_gts will be defined later in the main script
; search for gts_fullname and modify it for your own directory structure
fname = "/glade/work/weisun/2019/fso_aws/fso/run/2007010100/wrfvar/working/wrfinput_d01"  ; for retrieving mapping info
start_date   = "2007010100"
end_date     = "2007010100"
cycle_period = 12
plotscatter  = False   ;;    True
write_netcdf = False
input_source = "ascii"   ; reading data from gts_omb_oma_01
;input_source = "netcdf"  ; reading data from the output of write_netcdf
nc_datdir    = "./"
out_type     = "pdf"    ; or ncgm
expt         = ""  ; for output naming purpose
filename_gts   = "gts_omb_oma_01"
```

# Applications

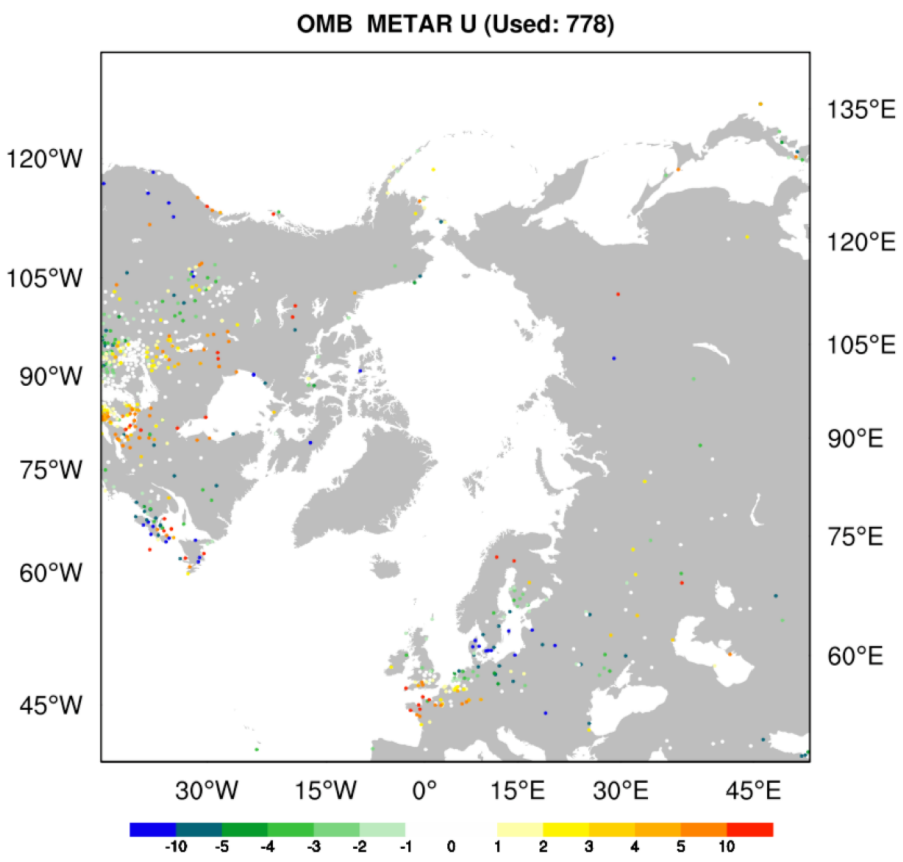
Test case in FSO  average between 2007010100 - 2007010112

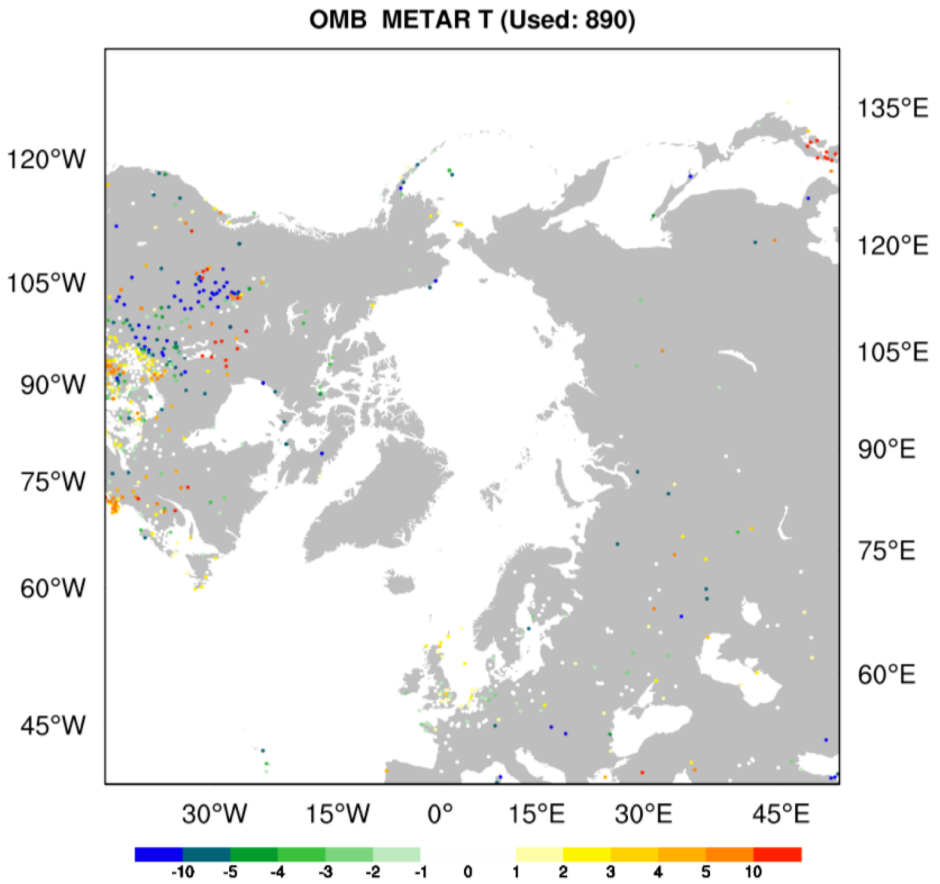


# Applications

Test case in FSO  average between 2007010100 - 2007010112



# Applications

*fso/plot/plot_levs_impact.ncl* will create the file "multi_levs_sound_2007010100.pdf", which will show the observation impact on the forecast by observation type at each vertical level.

```
;=================================================================
; Main code
;=================================================================

  start_date   = "2007010100"
  end_date     = "2007010112"
  cycle_period = 12
  date         = start_date
  datdir1      = "/glade/work/weisun/2019/fso_aws/fso/run/"        ; the path before DATE
  datdir2      = "/obsimpact/working/"                        ; the path after  DATE
  filename_gts   = "gts_omb_oma_01"
```

# Applications

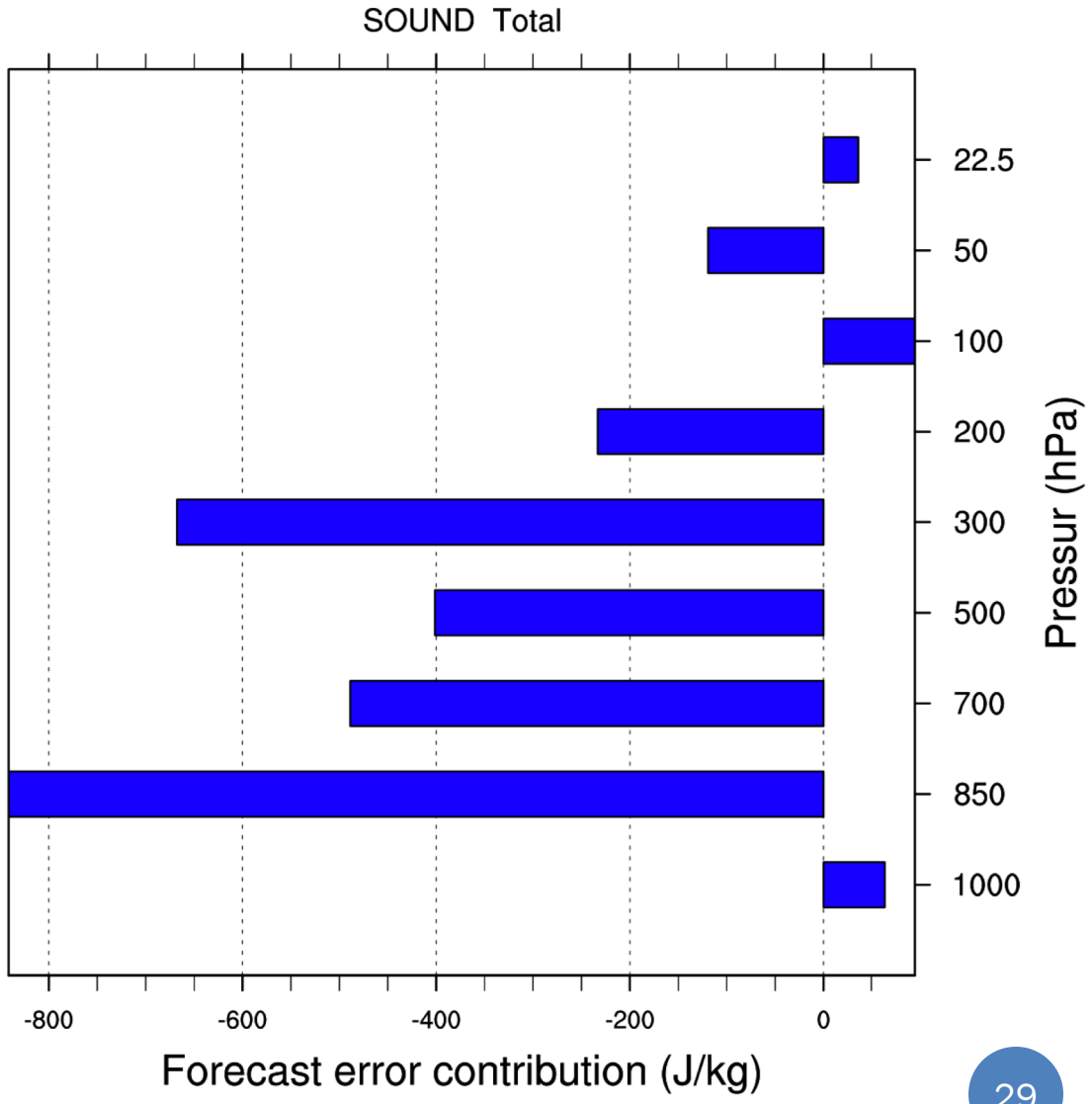
**negative value**

↓

**reduce forecast error**

↓

**positive forecast contribution**

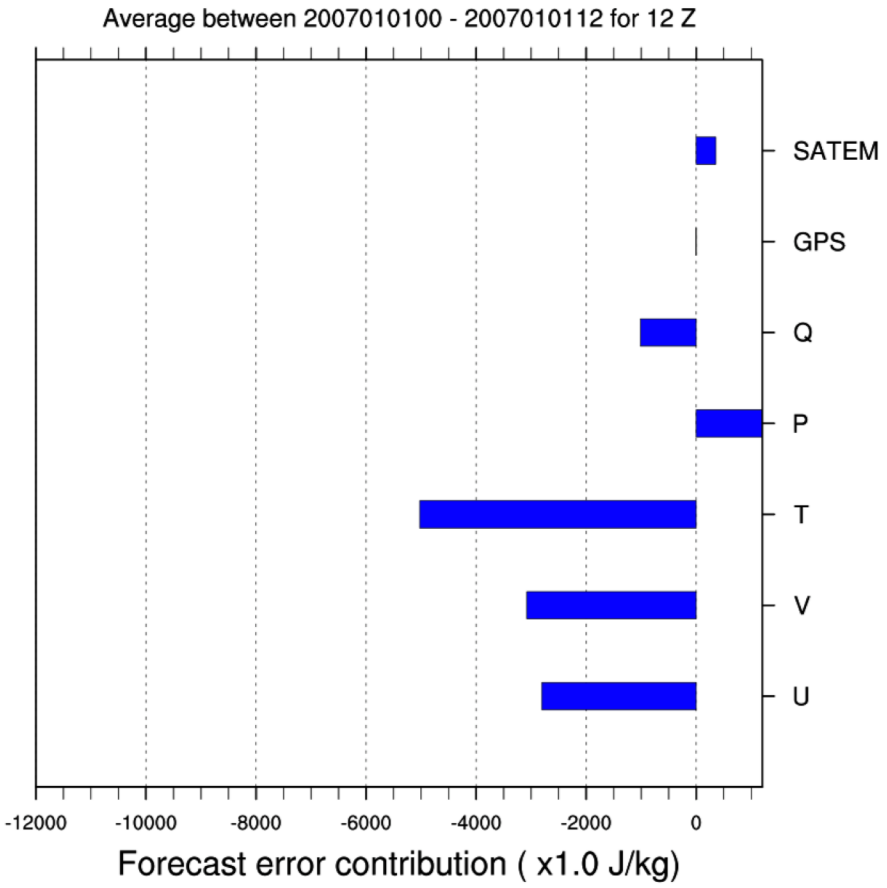


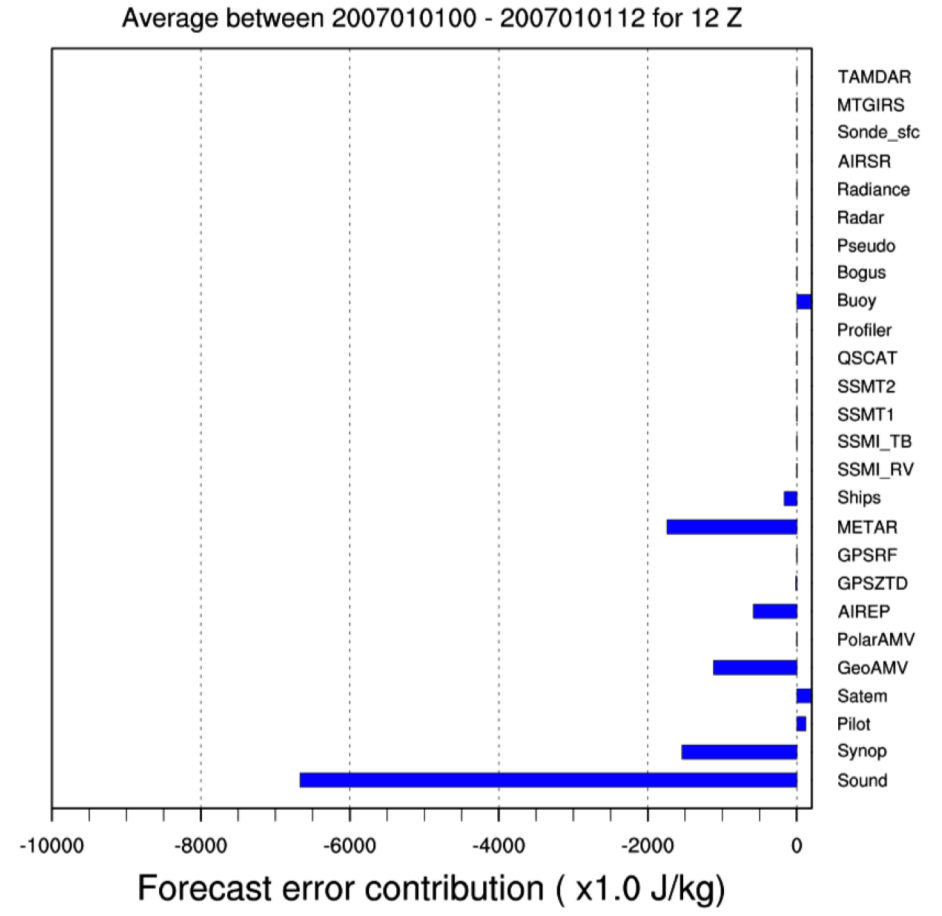
# Applications

*fso/plot/plot_avg_obs_impact.ncl* will create the file " avg_obs_impact_bar.pdf", which will show the observation impact on the forecast by observation type.

```
;================================================================
; Main code
;================================================================
  start_date   = "2007010100"
  end_date     = "2007010112"
  utc = "12 "
  cycle_period = 12
  expt_dir     = "/glade/work/weisun/2019/fso_aws/fso/run/"
  obsimpact_dir= "/obs_impact"
```

# Applications



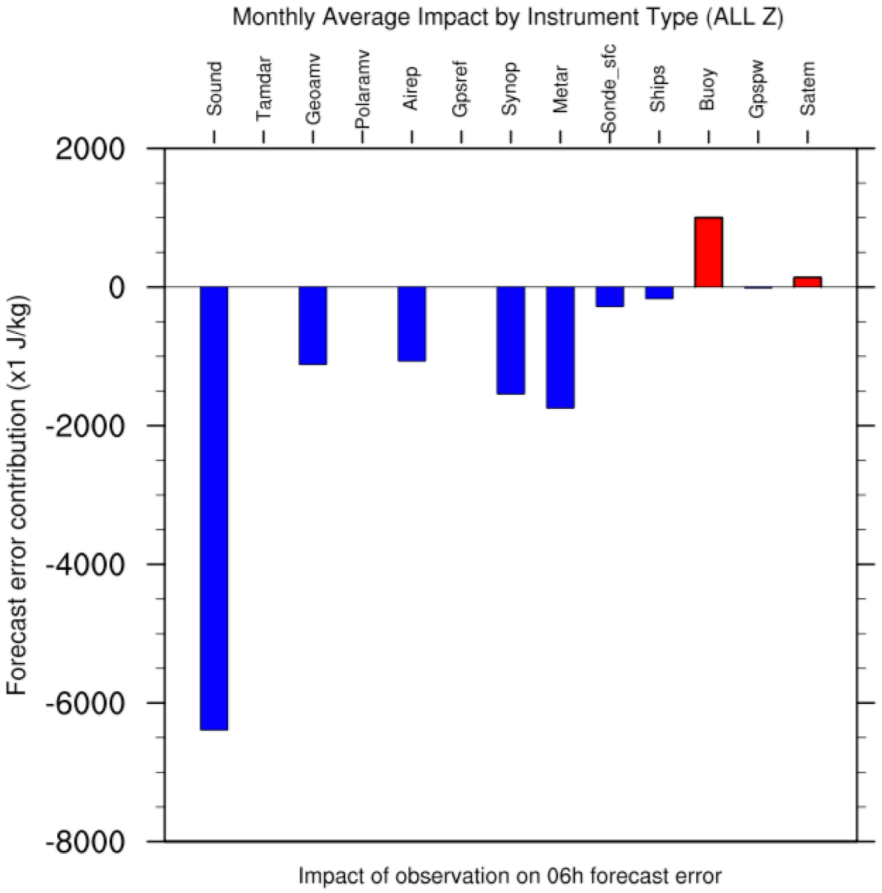



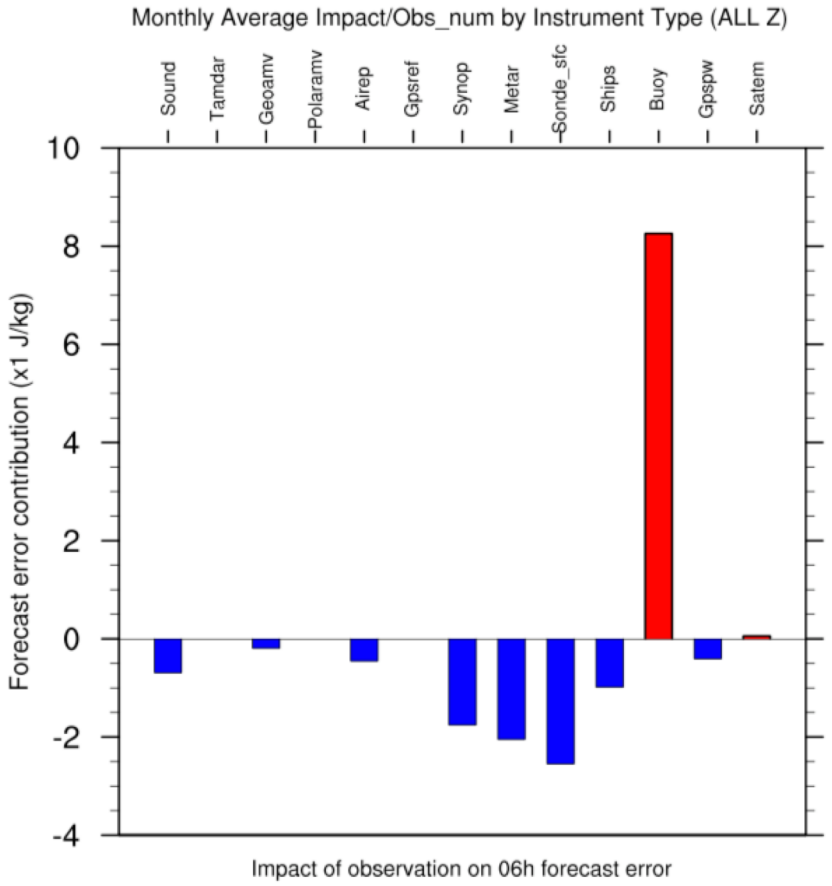
# Applications

*fso/plot/plot_impact_by_obsnum_time_average.ncl* will create the file "impact_by_obs2007010100-2007010112.pdf", which will show the observation impact on the forecast by observation type.

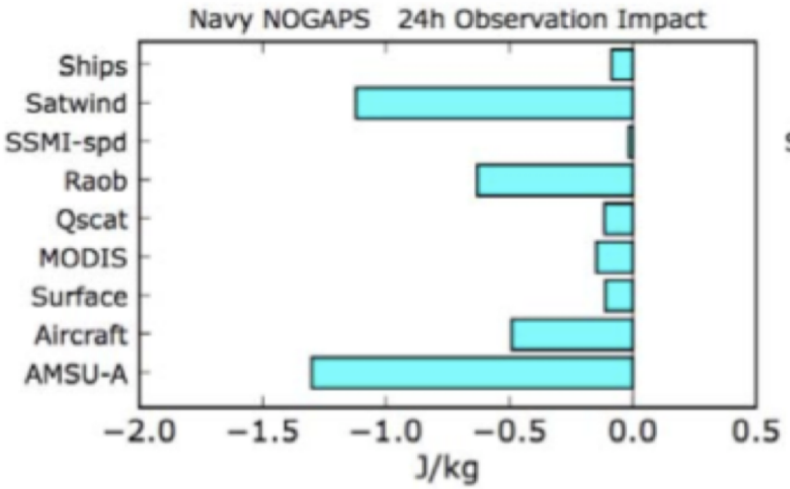
```
;=================================================================
; Main code
;=================================================================

  start_date   = "2007010100"
  end_date     = "2007010112"
  cycle_period = 12
  date         = start_date
  datdir1      = "/glade/work/weisun/2019/fso_aws/fso/run/"        ; the path before DATE
  filename_gts   = "/gts_omb_oma_01"
```

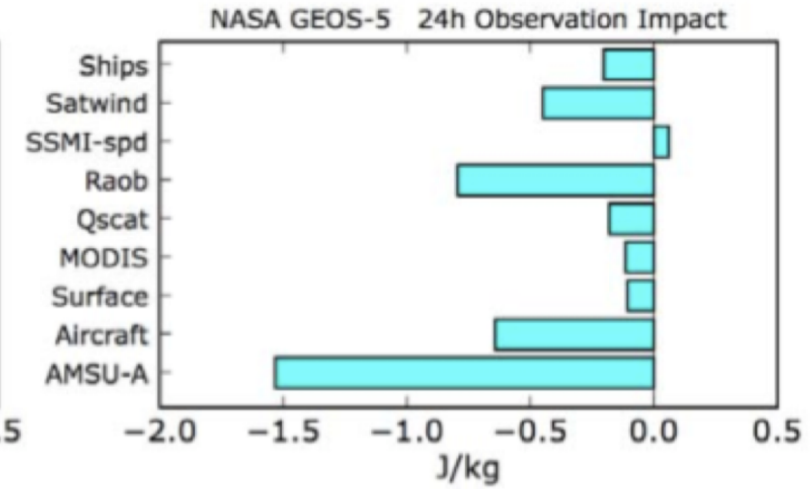
# Applications

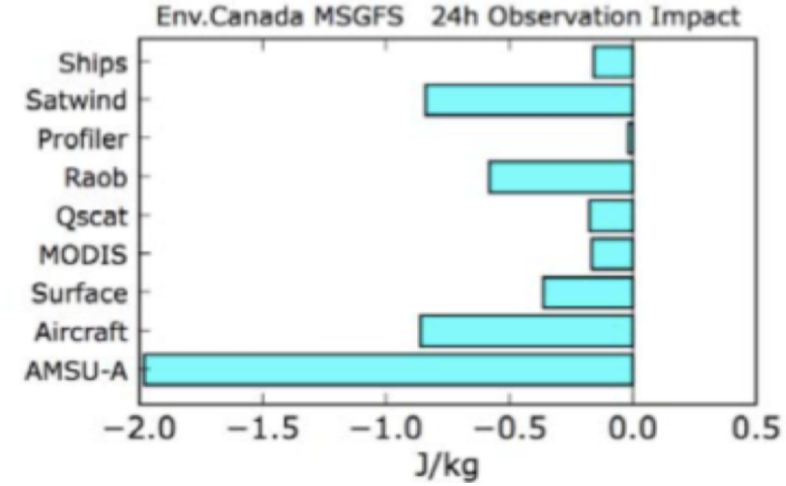


# Applications




***AMSU-A Observations Have the Greatest Benefit at all Three Centers.***

*from Gelaro et al. 2009*

# FSO – Outline

- Introduction
- Implementation in WRF/WRFDA
- Applications
- Limitations
- Conclusions

# Limitations

- Uncertainties are difficult to estimate
  - The reference for the calculation of the forecast accuracy is NOT perfect and often correlated with the initial analysis
  - The adjoint model is not an accurate representation of the NL model behavior (linearization, simplification, dry physics)
- Results are strongly dependent on the norm chosen to define the forecast accuracy
- The adjoint-based technique is restricted by the tangent linear assumption, valid from 1 to 3 days,

# Conclusions

- All code and scripts for FSO are available in current WRF public release
- Testing package & User's Guide available online http://www2.mmm.ucar.edu/wrf/users/wrfda/download/fso.html
- Have fun!