Ensemble Assimilation of Global Large-Scale Precipitation

Guo-Yuan Lien[1,2]

in collaboration with Eugenia Kalnay[2], Takemasa Miyoshi[1,2]

[1]RIKEN Advanced Institute for Computational Science
[2]University of Maryland, College Park

Special thanks to Daryl Kleist, George Huffman,
Yoichiro Ota, and Daisuke Hotta

RIKEN Advanced Institute for Computational Science
September 10, 2014

Outline

- Background
- Objectives
- Transformation of precipitation
- OSSEs with the SPEEDY model
- GFS-LETKF assimilation of the TMPA data
  - Part I: Statistics of the precipitation variable.
  - Part II: Real data assimilation experiments.
  - Part III: EFSO for precipitation.
- Summary

Background

- Precipitation has long been an important meteorological observations.
- Past studies of precipitation assimilation have been done mostly with the nudging or variational systems.
  - They are successful during the assimilation (e.g., NARR), but the model forgets about the changes soon after the assimilation stops.
  - The change in moisture is not an efficient way to update the potential vorticity field, which is the “master” dynamical variable that primarily determines the evolution of the forecast in NWP models.

Background

- Assimilation of precipitation is difficult because of:
  - the nonlinear precipitation process.
    (difficult to create the tangent linear model and adjoint model)
    → Solution: Ensemble Kalman filter (EnKF)
  - the non-Gaussianity of the precipitation variable.
    (violate the Gaussian error assumption in most of data assimilation schemes)
    → Solution: CDF-based Gaussian transformation

OSSE with a simplified GCM

  - the imperfect precipitation parameterization in numerical models.
    (model errors)
  - the unknown errors associated with the precipitation observations.
    (observation errors)

More challenges with real model and data

Objectives

- Proposed methods of precipitation assimilation:
    - Local ensemble transform Kalman filter (LETKF).
    - Cumulative distribution function (CDF)-based transformation of precipitation, instead of logarithm transformation.
- For proof of concept, we conducted idealized perfect-model experiment with the SPEEDY model.
- Application to the real data assimilation with a realistic model:
    - TRMM Multisatellite Precipitation Analysis (TMPA)
    - NCEP Global Forecasting System (GFS).

Transformation of precipitation

- Most of practical data assimilation schemes for large systems assume Gaussian error distributions for both observations and the model backgrounds
  - However, it is unavoidable that we need to use observations with a certain non-Gaussianity.
  - Variable transformation is a cheaper solution to alleviate the problem.
- Transformation methods used in our study:
  - No transformation (**NT**)
  - Logarithm transformation (**Log**)
    (Lopez 2011, 2013; others)
  - Gaussian transformation (**GT**)
    (Schöniger et al. 2012; Lien at al. 2013)

Transformation of precipitation

- Logarithm transformation:
  - $\tilde{y} = \ln(y + \alpha)$
- Gaussian transformation (Gaussian anamorphosis):

$F^G(\tilde{y}) = F(y)$ or $\tilde{y} = {F^G}^{-1}[F(y; \text{location, period of year})]$

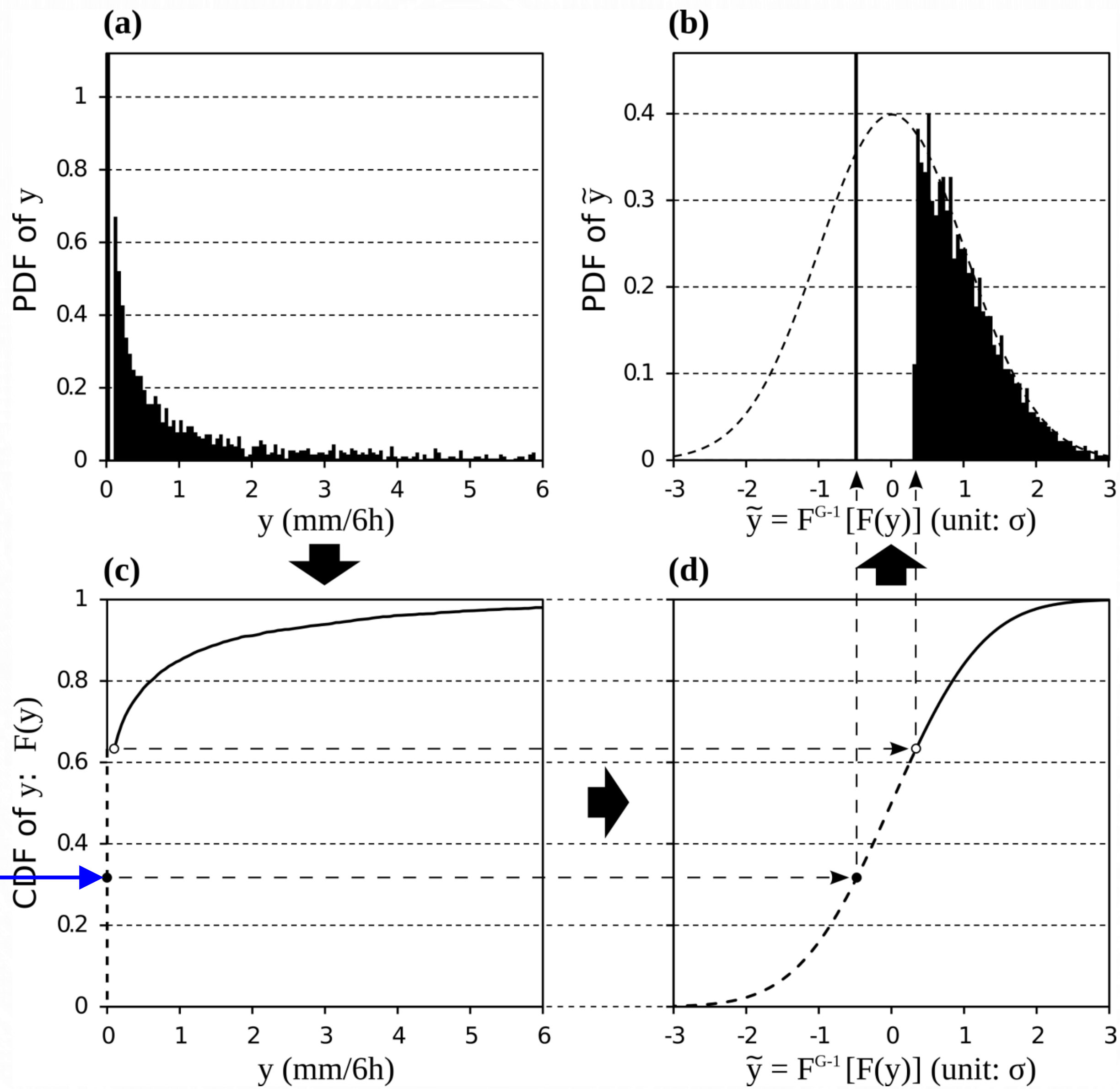
  - $F()$: The cumulative distribution function (CDF) of the original variable
  - $F^G()$: The CDF of a standard normal distribution. [${F^G}^{-1}()$: The inverse function of $F^G()$]
  - The CDF of precipitation variables is empirically determined based on the model/observation climatology at each grid point and period of year. It requires a long period of model/observation data.
  - Two methods for transform the zero precipitation: **GTcz** and **GTbz**. → Will be introduced later.
  - It results in a Gaussian climatology, but not necessarily the Gaussian error distribution required in the data assimilation. → Will be verified later.

Example of Gaussian precipitation transformation

PDF

CDF

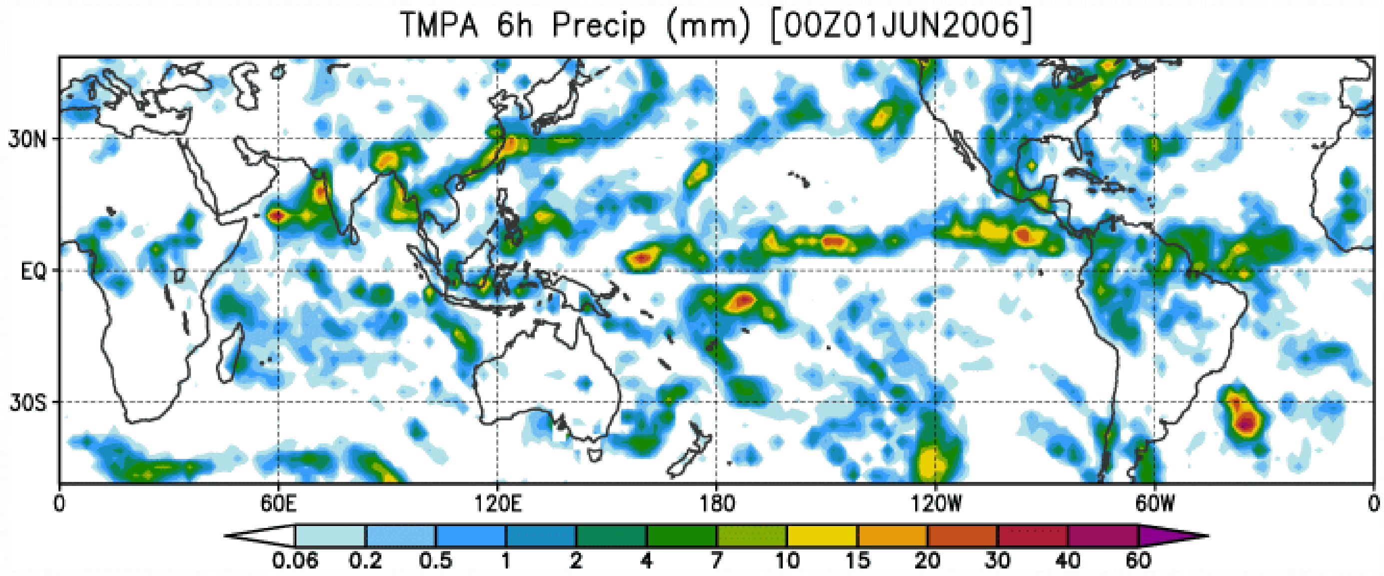
The climatological median of zero (GTcz) method



Original variable **Transformed variable**

Example of Gaussian precipitation transformation

Original variable



Transformed variable

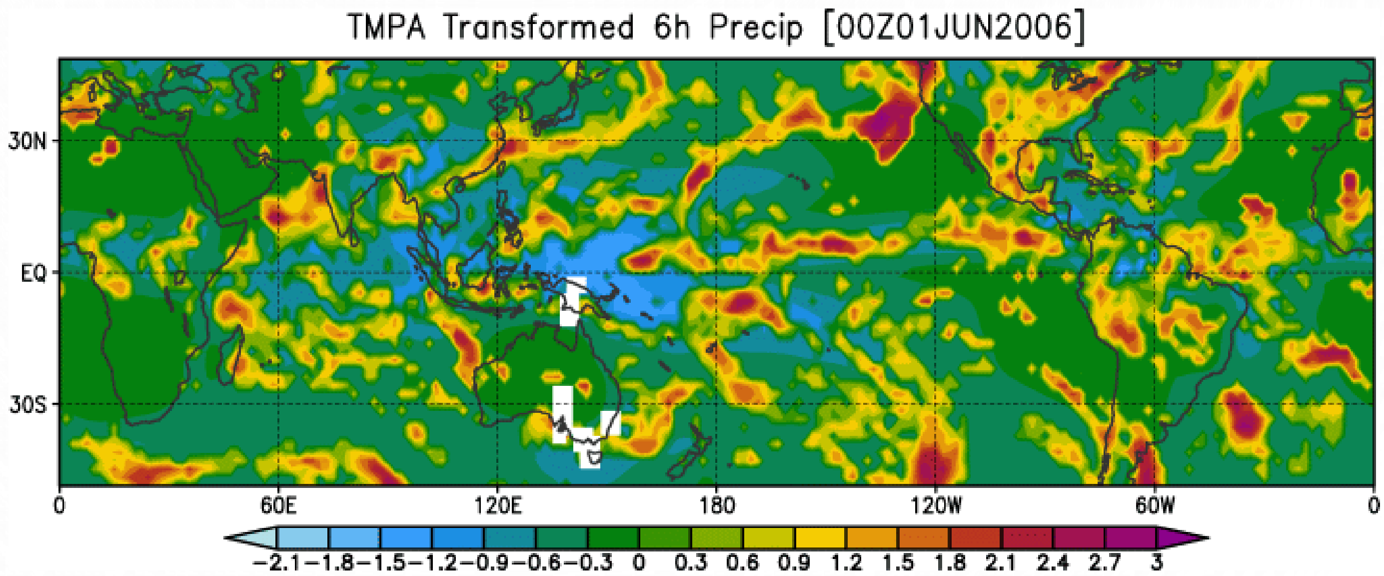


Illustration of error distribution and zero precipitation transform

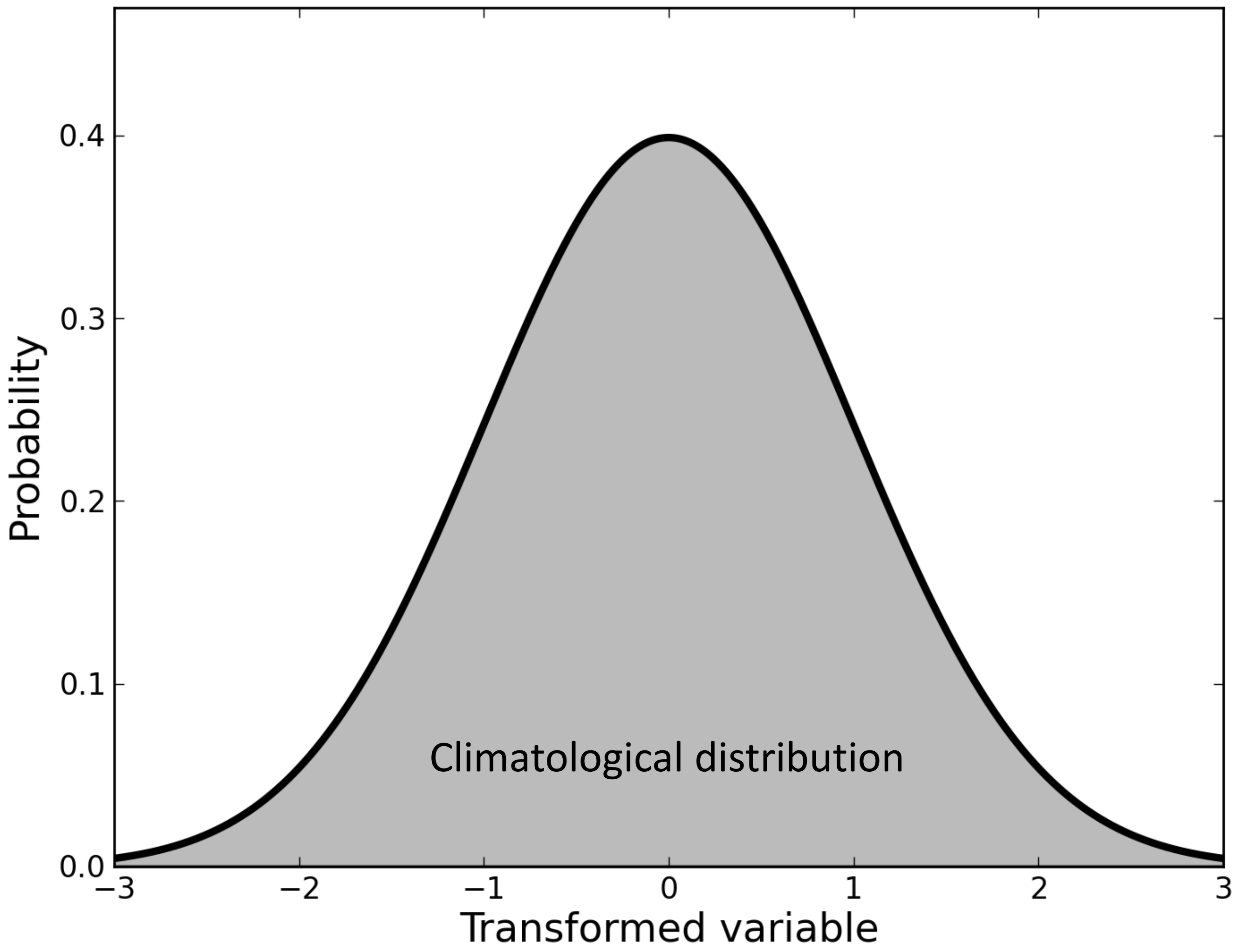


Illustration of error distribution and zero precipitation transform

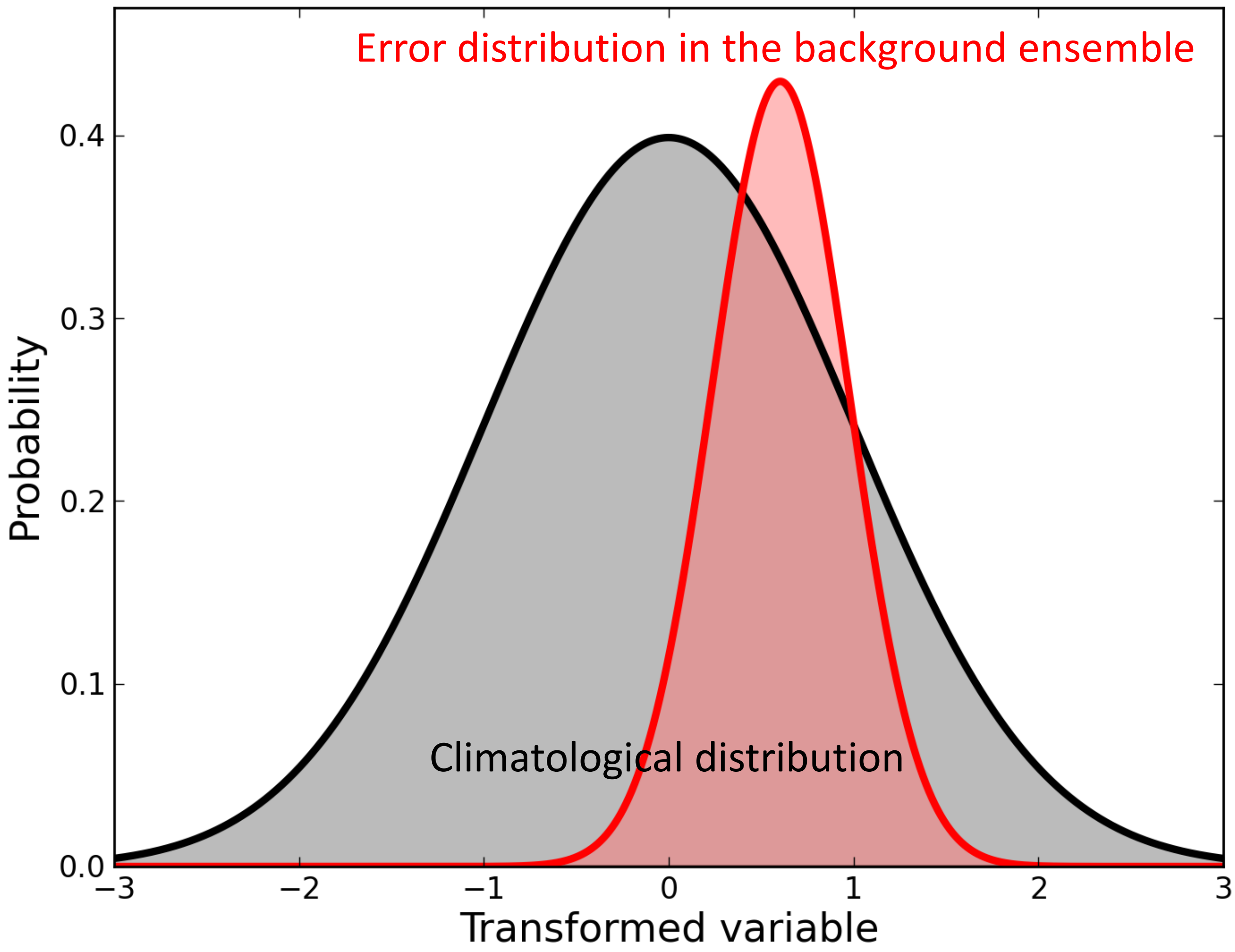
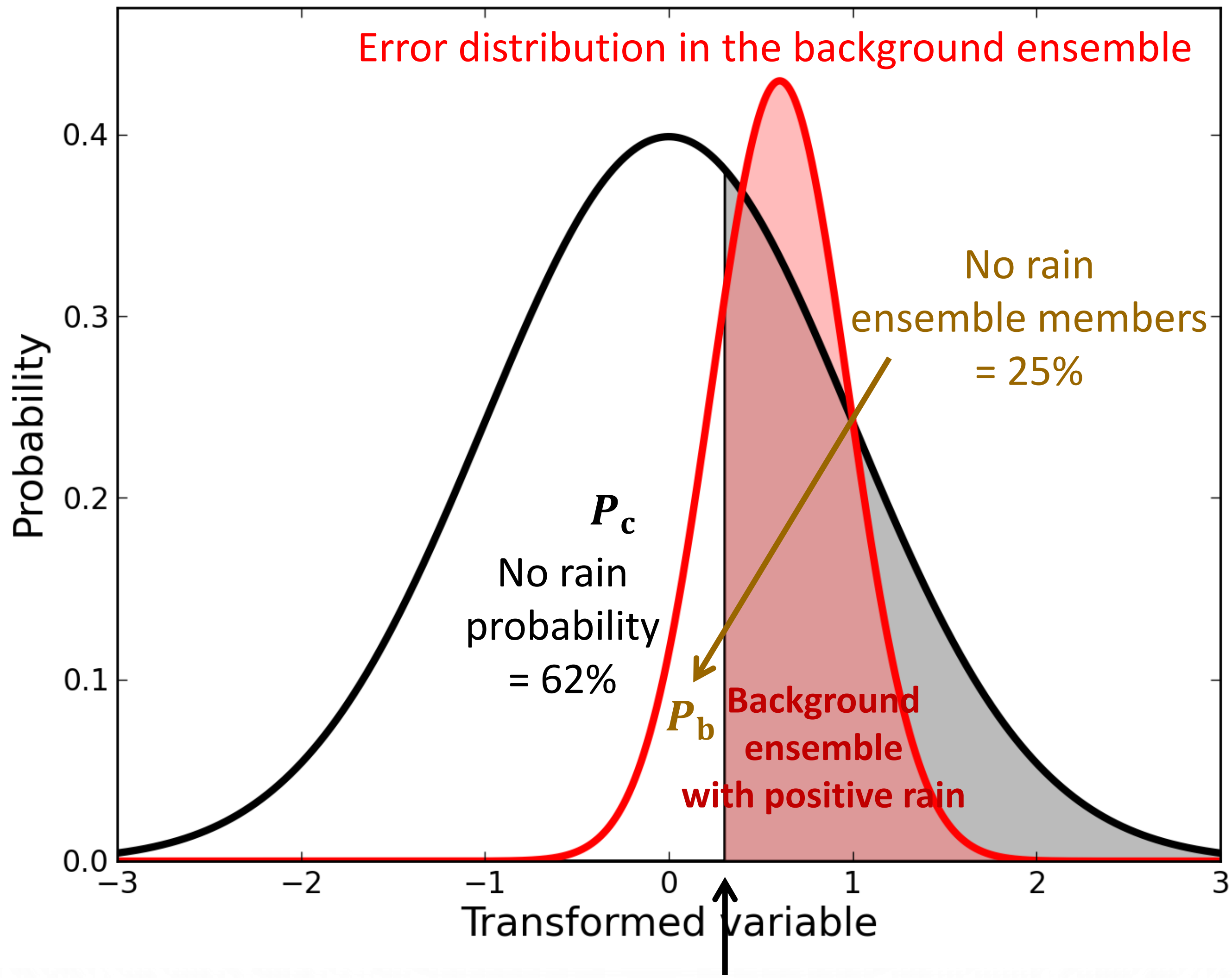


Illustration of error distribution and zero precipitation transform



$\tilde{y}_{\mathrm{trace}}$ **: Boundary of rain/no rain**

Illustration of error distribution and zero precipitation transform

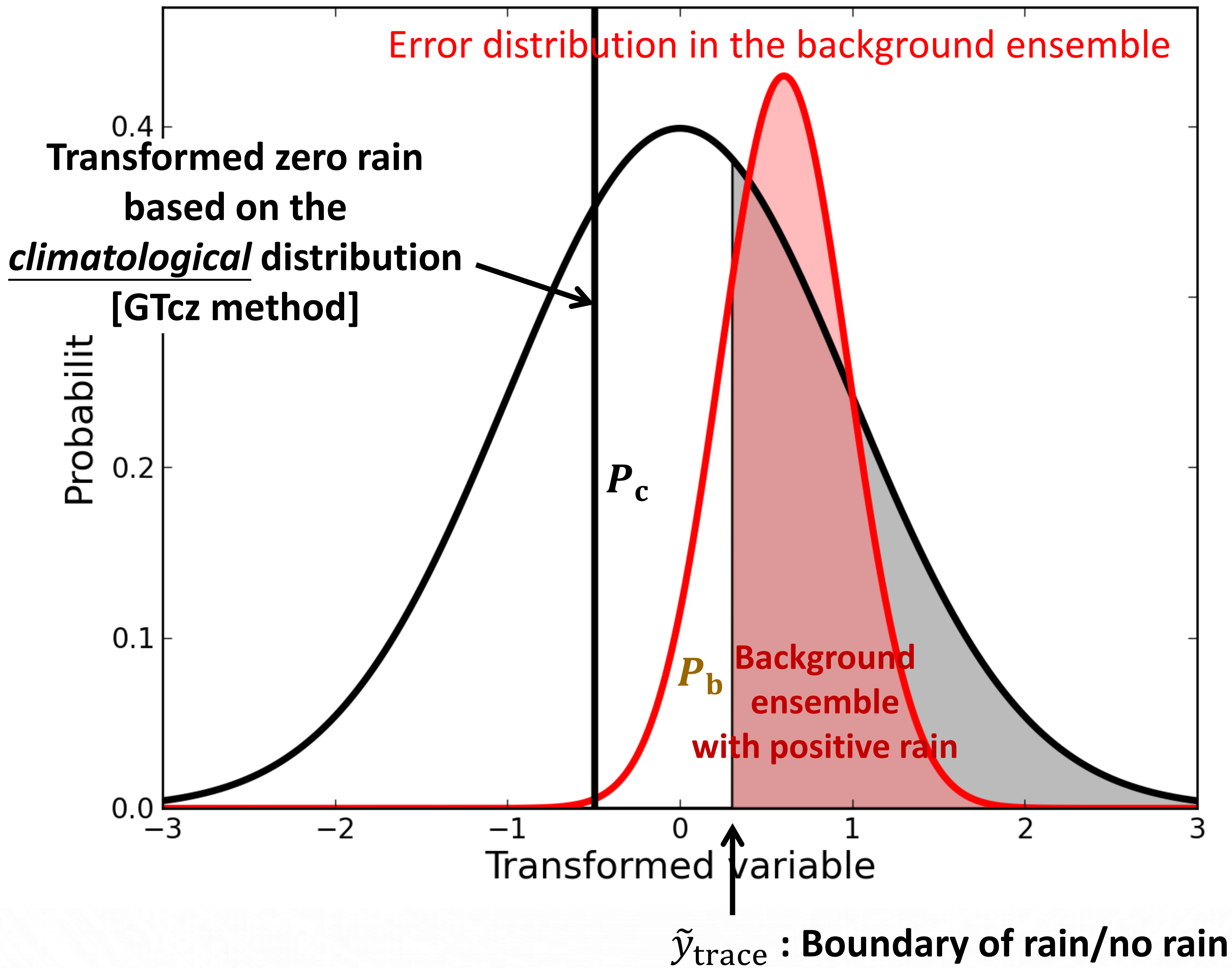
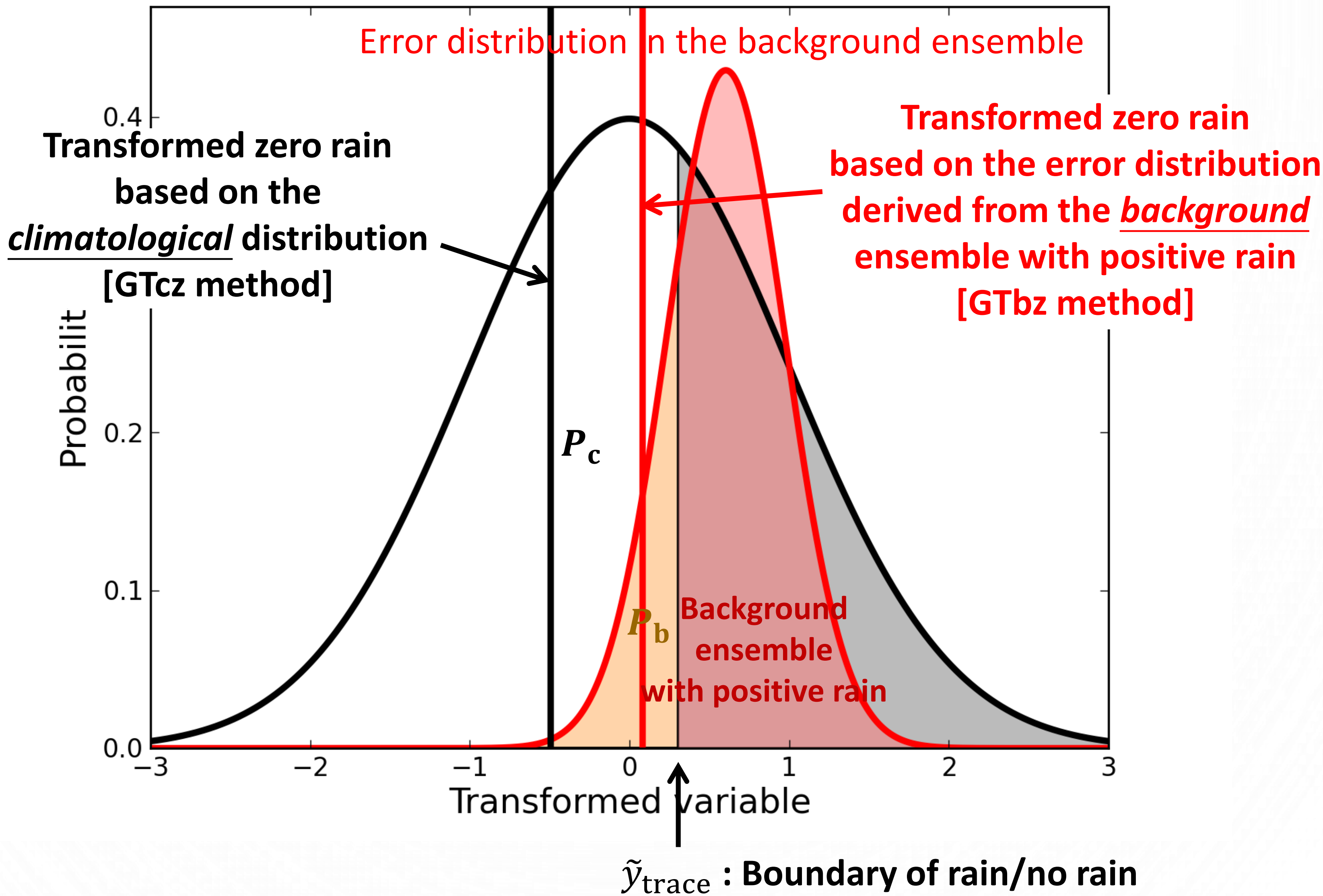


Illustration of error distribution and zero precipitation transform



Outline

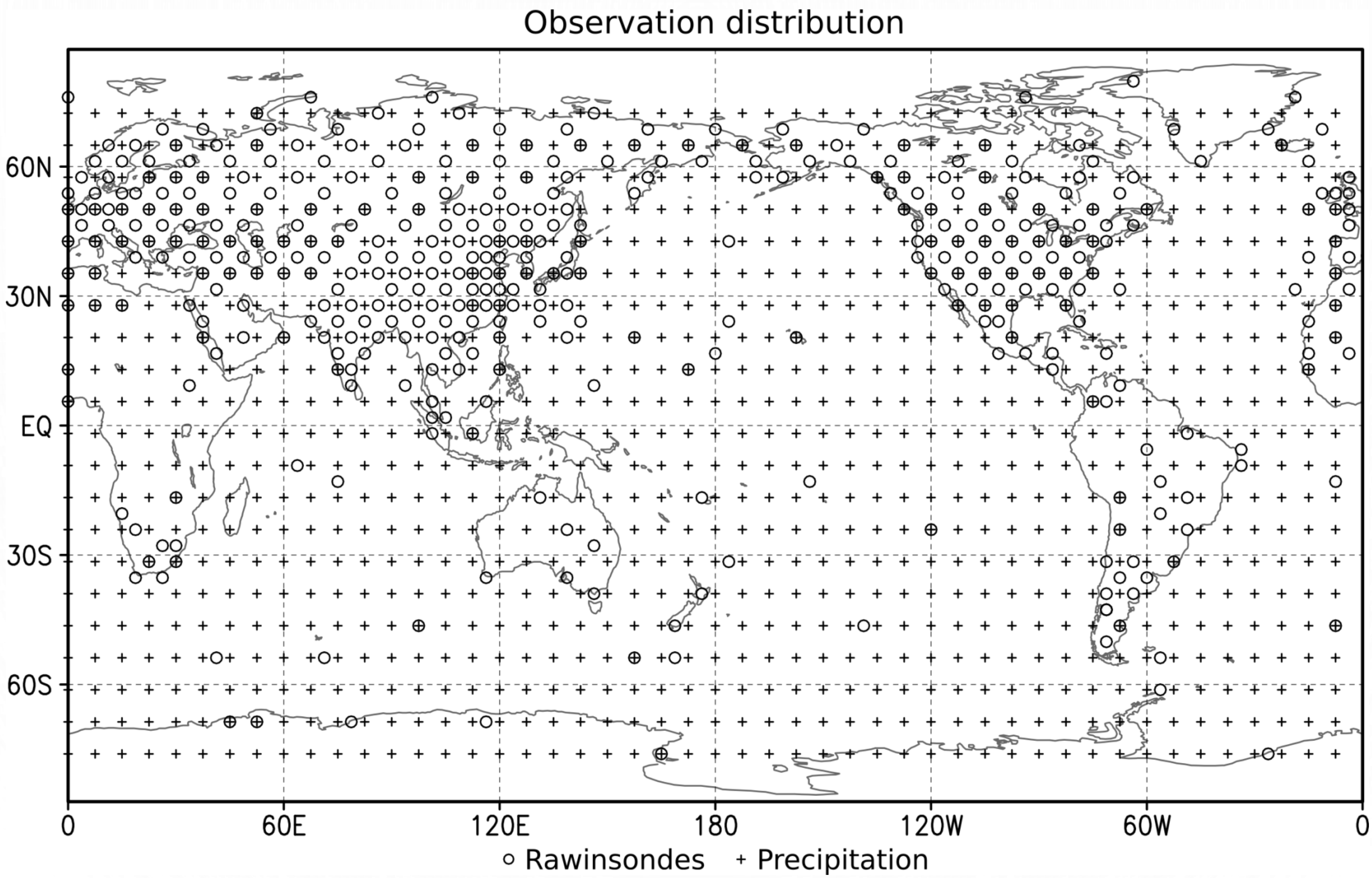
- Background
- Objectives
- Transformation of precipitation
- OSSEs with the SPEEDY model
- GFS-LETKF assimilation of the TMPA data
  - Part I: Statistics of the precipitation variable.
  - Part II: Real data assimilation experiments.
  - Part III: EFSO for precipitation.
- Summary

OSSEs with the SPEEDY model

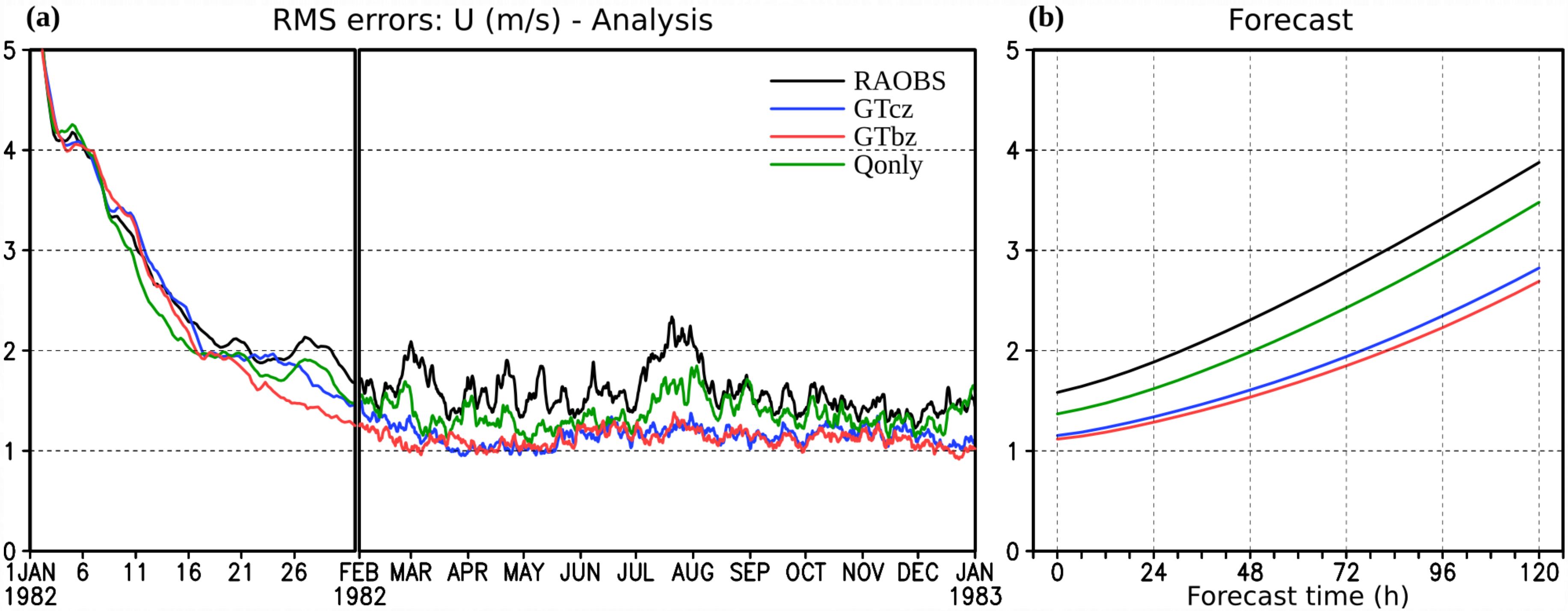
- 1-year identical-twin observing system simulation experiment (OSSE).
- Ensemble size = 20
- Adaptive inflation (Miyoshi 2011)
- Observation selection criteria for precipitation assimilation:
  - The traditional "***ObsR*** criterion": only assimilating precipitation at the location with observed positive precipitation (> 0.1 mm/6h).
  - The "***10mR*** criterion": only assimilating precipitation at the location where more than 10 background members have positive precipitation.

Experiment	Observations		Transf (precip)	Selection criteria (precip)	Obs error (precip)
	Raobs	Precip			
Raobs	v				
GTcz	v	v	GTcz	10mR	20%
GTbz	v	v	GTbz	10mR	20%
Qonly	v	v (only updating Q)	GTcz	10mR	20%
ObsR	v	v	GTcz	ObsR	20%

Observation distribution



Average analysis and forecast errors



(Spin-up) (After the spin-up) (11-month average after the spin-up period)

RAOBS: Assimilate rawinsonde observations

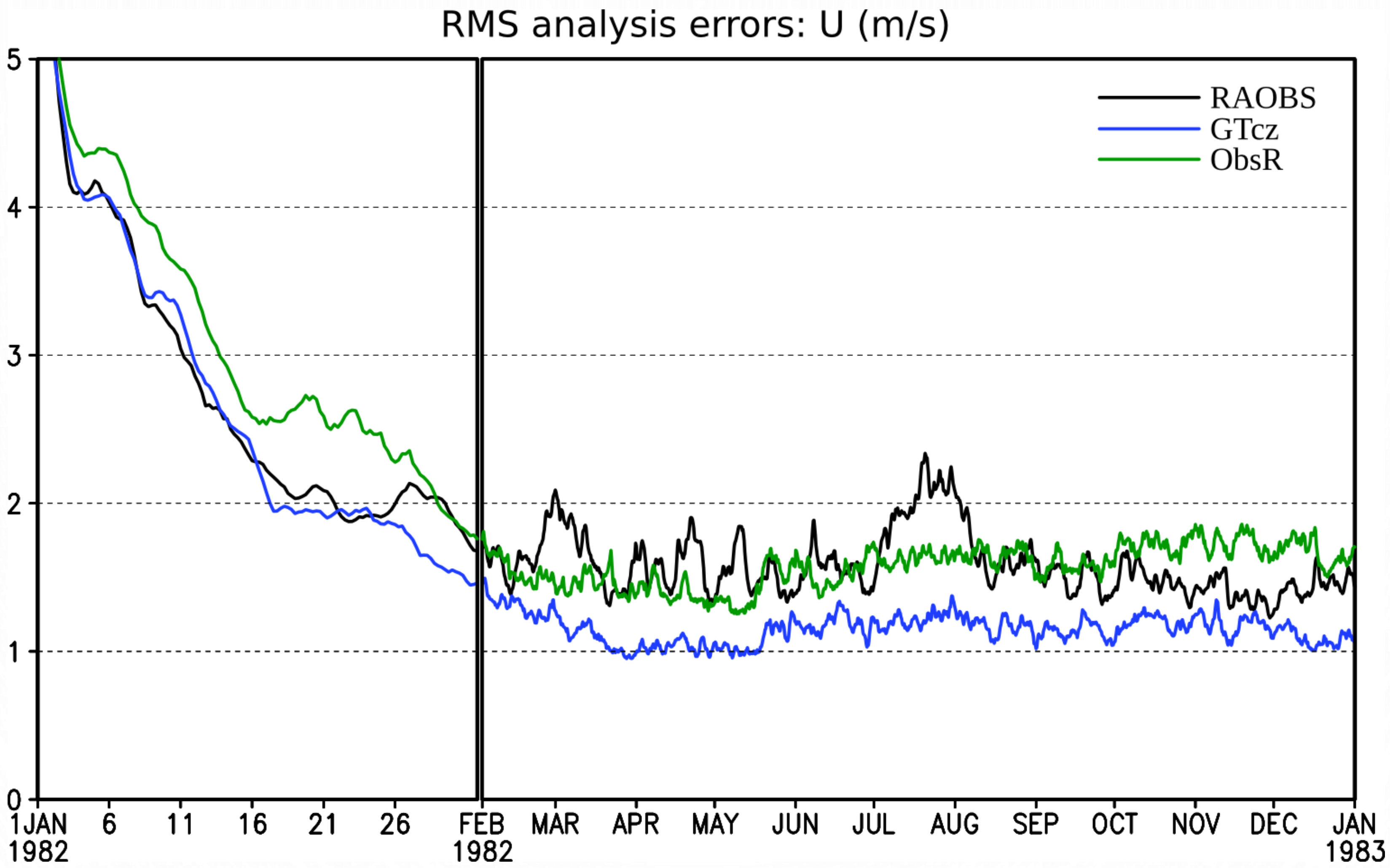
GTcz: Assimilate rawinsondes + uniformly distributed global precipitation using **GTcz**

GTbz: Assimilate rawinsondes + uniformly distributed global precipitation using **GTbz**

Qonly: Same as **GTcz**, but only update moisture field by precipitation assimilation

(Other variables show similar results)

Impact of observation selection criteria



Summary of the SPEEDY OSSEs

- Precipitation assimilation using an EnKF and with
  - Gaussian transformation
  - “10mR” criterion

  can significantly improve the analyses and medium range forecasts in the SPEEDY model.
- The EnKF can effectively update all model variables by the precipitation assimilation.

Outline

- Background
- Objectives
- Transformation of precipitation
- OSSEs with the SPEEDY model
- GFS-LETKF assimilation of the TMPA data
  - Part I: Statistics of the precipitation variable.
  - Part II: Real data assimilation experiments.
  - Part III: EFSO for precipitation.
- Summary

GFS-LETKF assimilation of TMPA data

- More realistic experimental design:

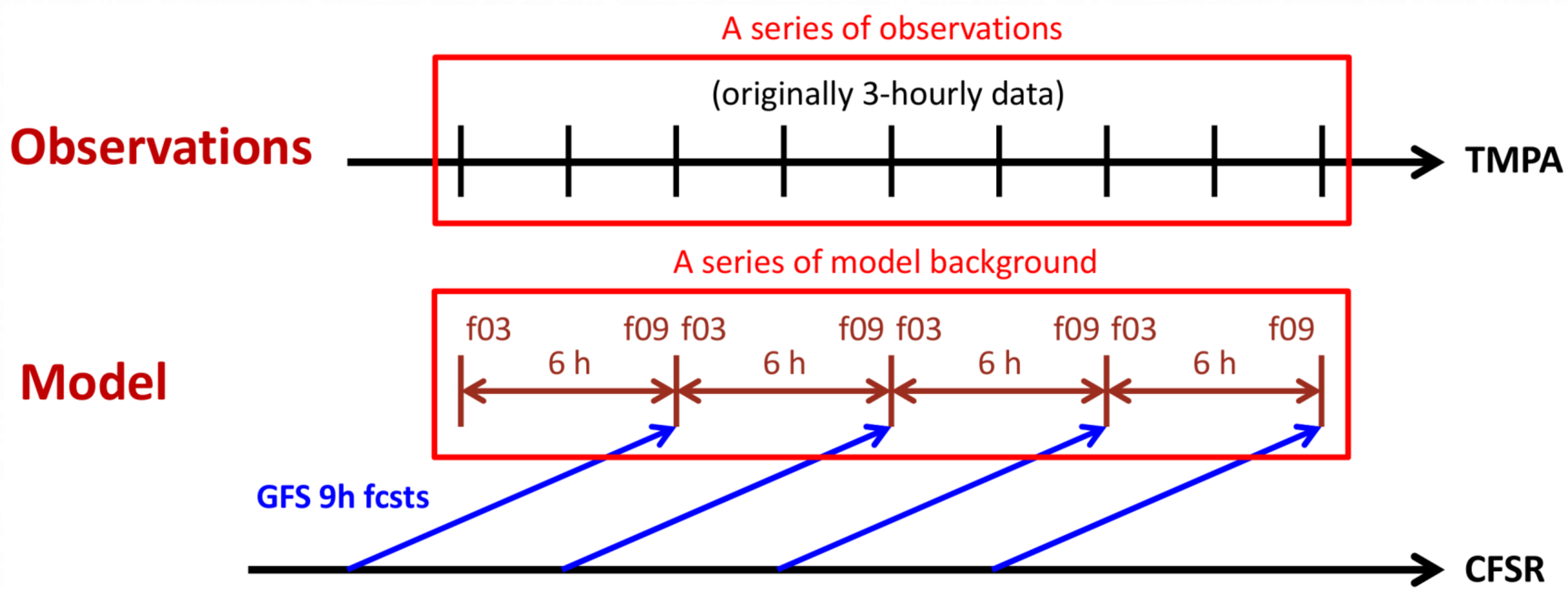
  Assimilation of the TRMM Multi-satellite Precipitation Analysis (TMPA) into a low-resolution NCEP GFS model.

Three parts of the study:

- **Part I**: Statistics of the precipitation variable in both the GFS model and the TMPA observations.
- **Part II**: Real data assimilation experiments.
- **Part III**: Ensemble forecast sensitivity to observations (**EFSO**) for precipitation.

Part I: Statistics of model/observed precipitation

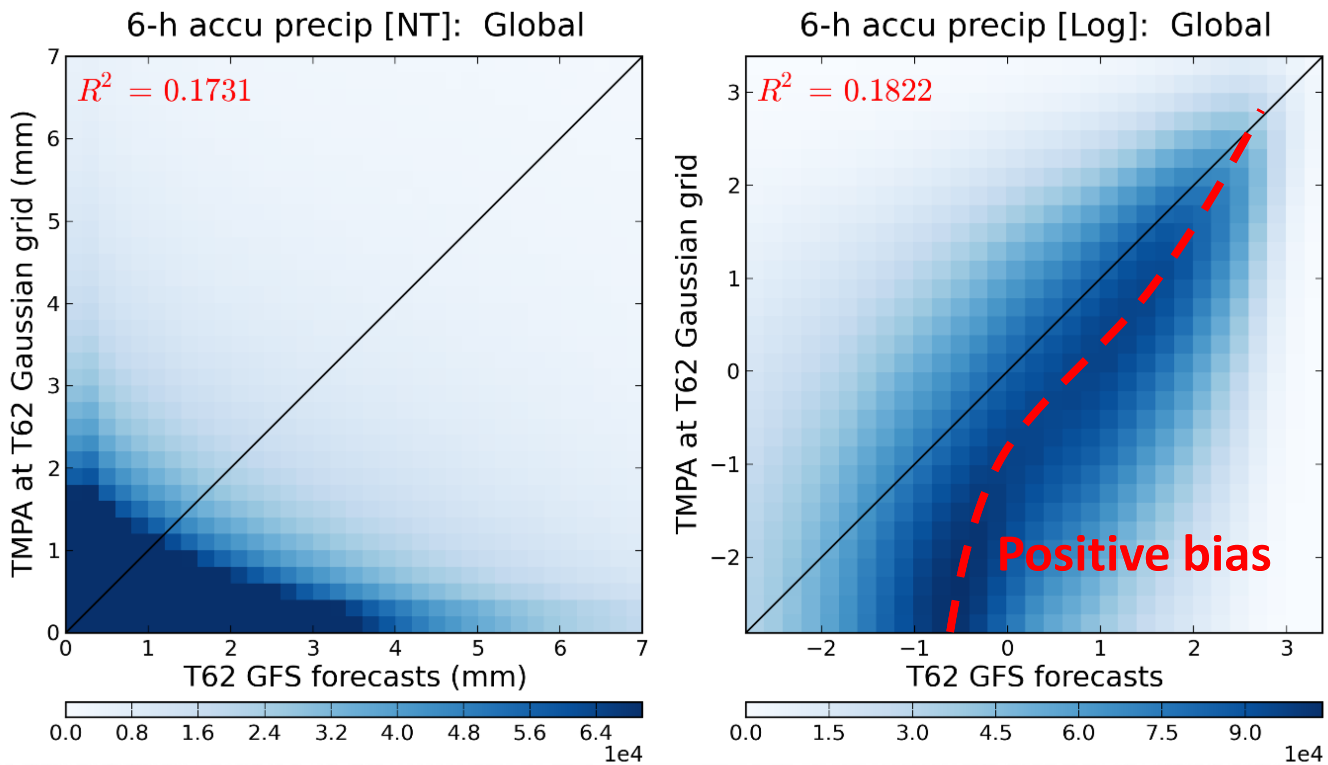
- **Model**: NCEP GFS model at a T62 resolution
- **Observations**: TMPA version 7 (3B42), upscaled to the Gaussian grid used by the T62 GFS model using an areal conservative remapping.
- **Variables**: precipitation rate or 6-h accumulated precipitation.
- **Sample**: 2001-2010 (10 year) period.



Joint probability distribution diagrams

No transformation **Log transformation**

TMPA at T62 grids

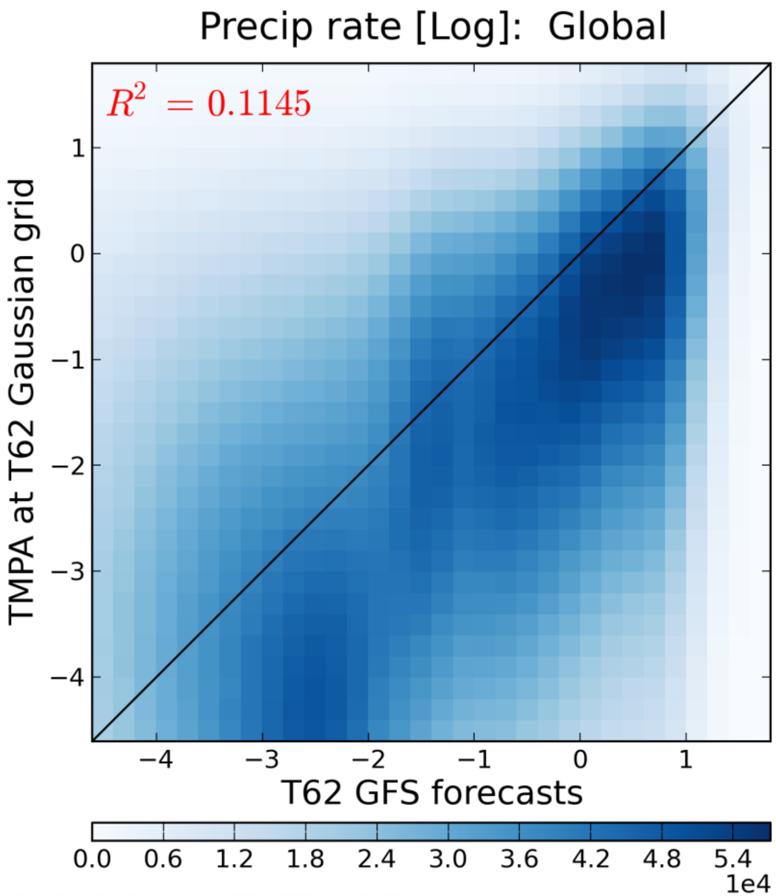


T62 GFS model background

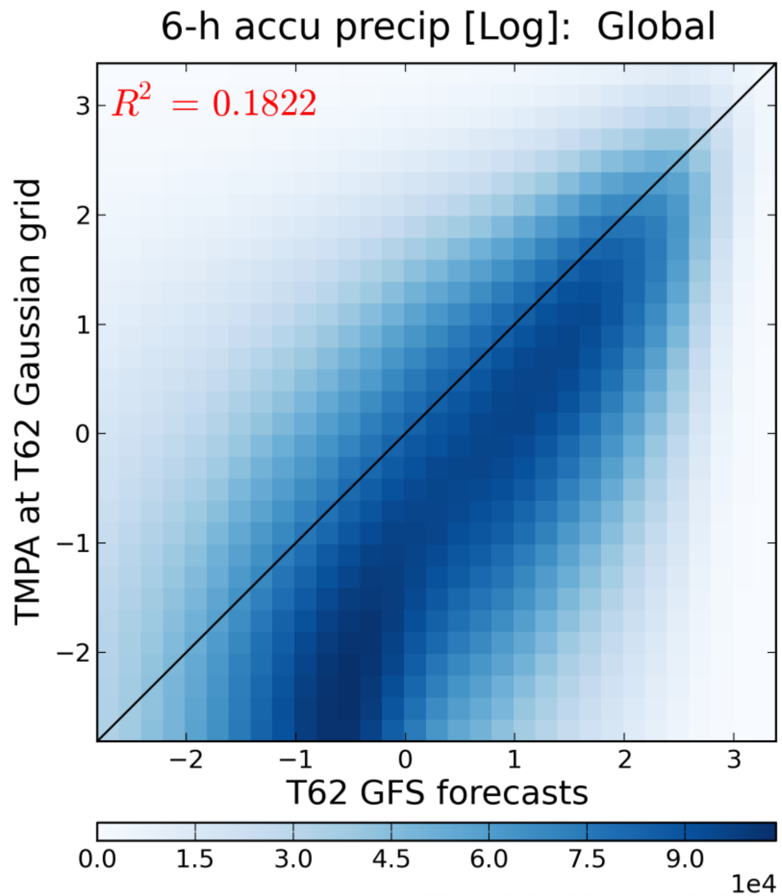
- The 6-h accumulated precipitation is used.
- Only positive precipitation is shown in all figures.

Instantaneous vs. accumulated precipitation

Instantaneous precip rate

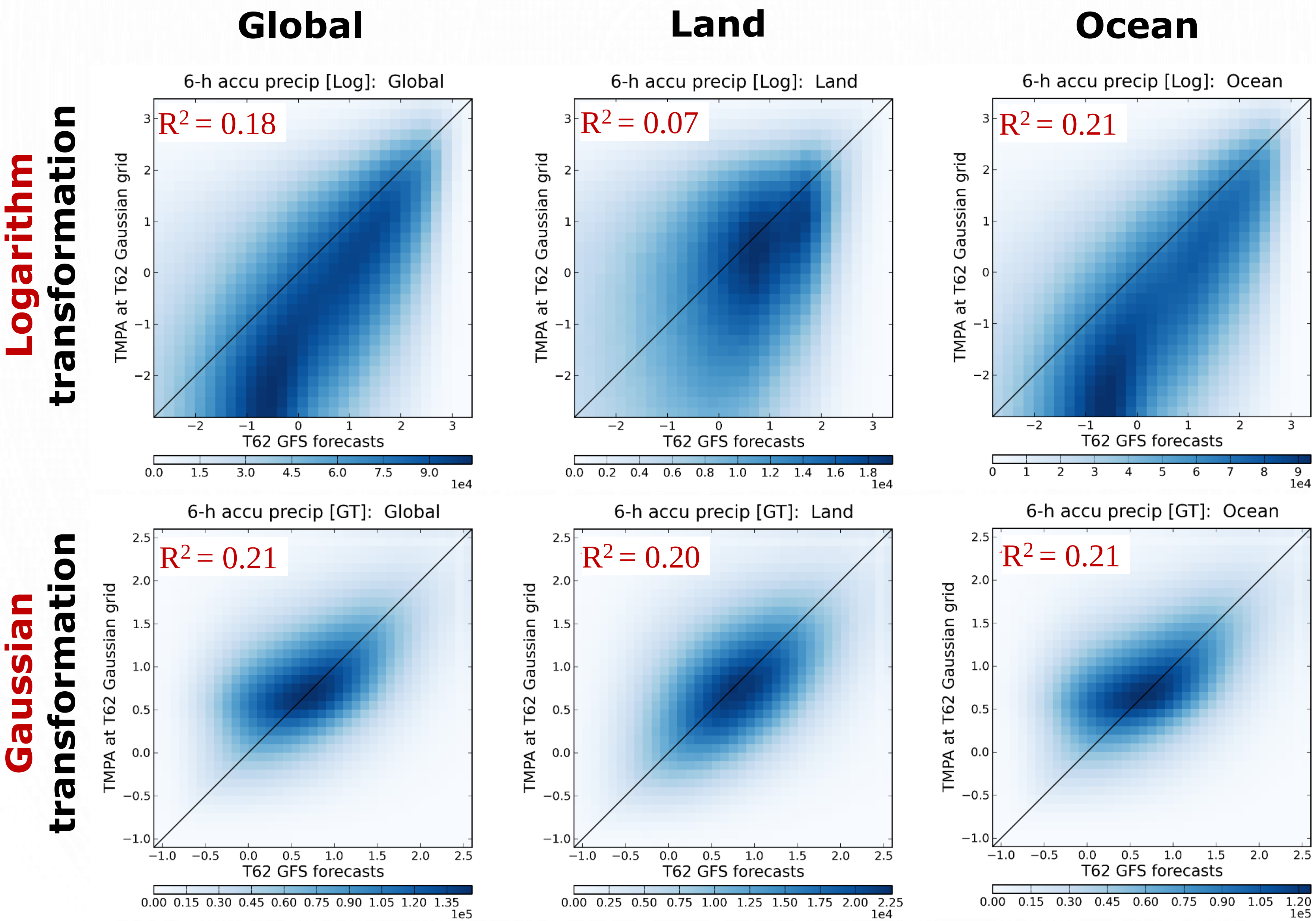


6-h accumulated precipitation



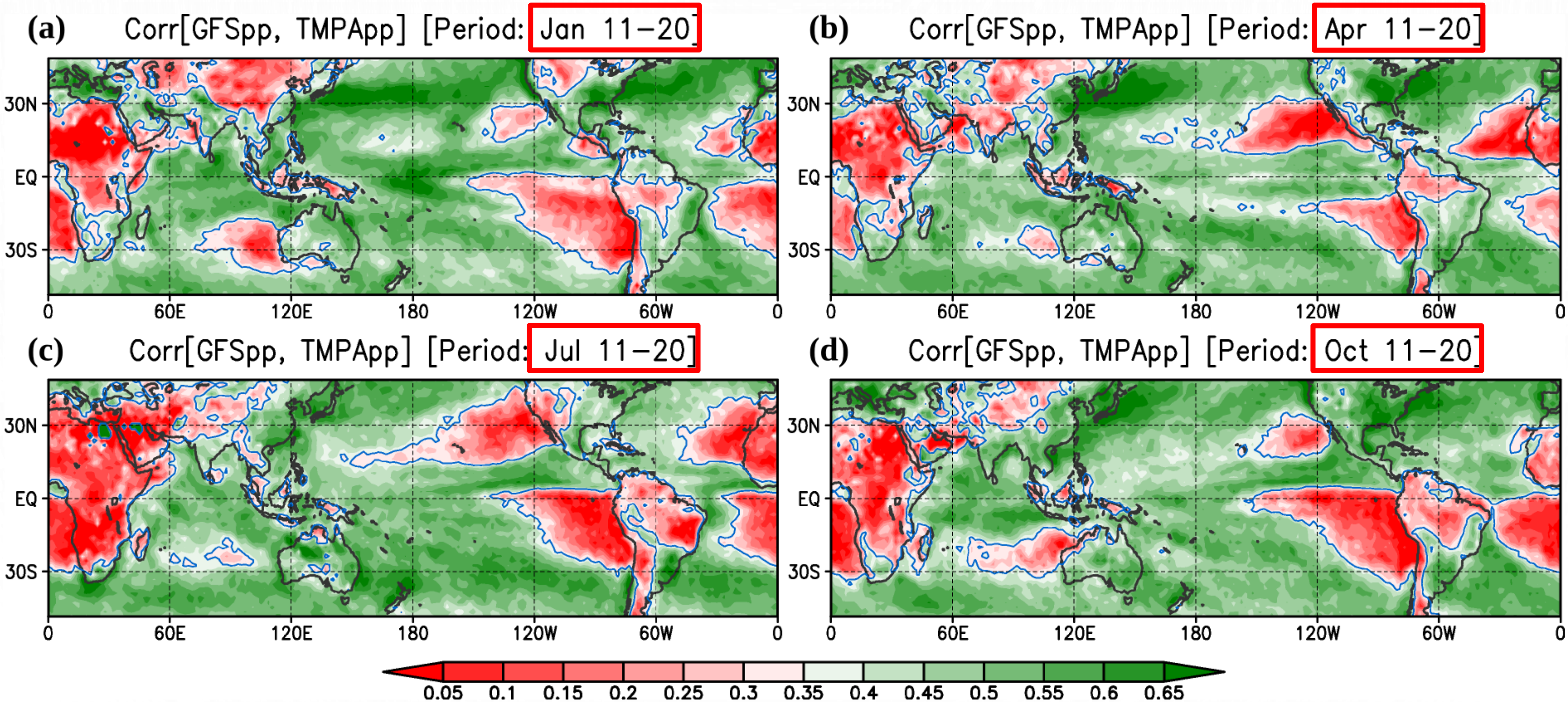
(Figures plotted with logarithm transformation)

Distribution after the Gaussian transformation



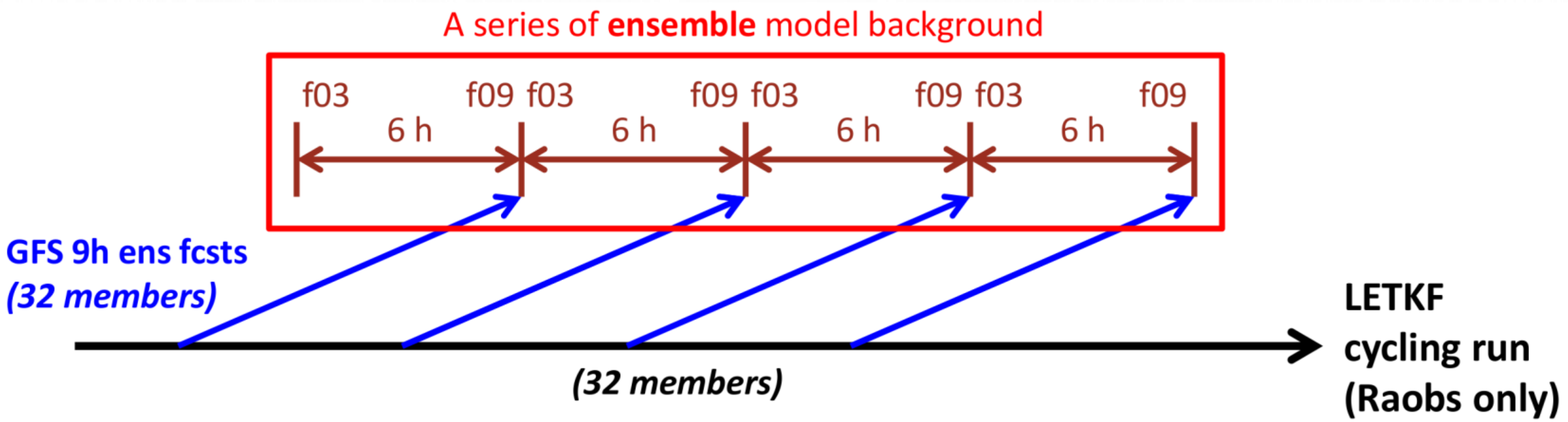
Correlation maps

- Correlation between the model backgrounds and the observations at each grid point.
- Blue contours: Corr = 0.35, **will be used to define the QC**.



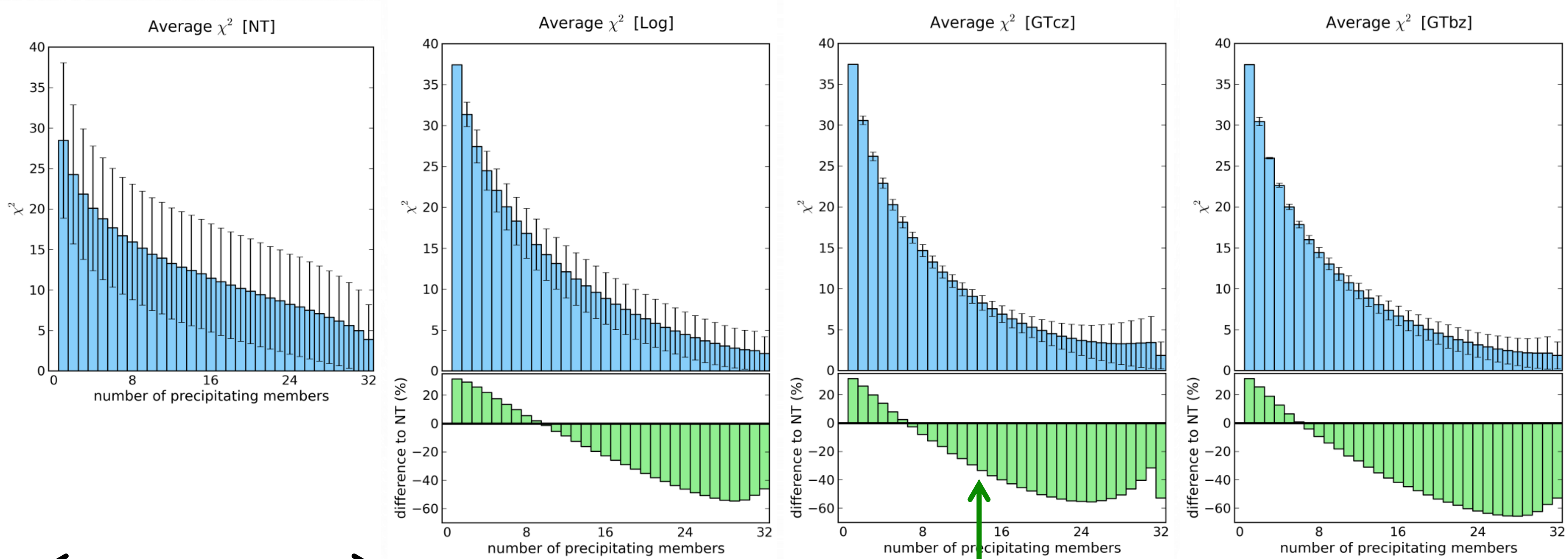
Gaussianity statistics

- Measure of (non-)Gaussianity:
  - $\chi^2 = \sum_{k=1}^{K} \frac{\left(y_k - y_k^{\text{expected}}\right)^2}{\sigma^2}$ (sorted y_k)
  - y_k : samples in the ensemble.
  - y_k^{expected} : samples taken from a Gaussian distribution with the mean and variance same as the ensemble.
- χ^2 values are computed for each precipitation observation, and then averaged.
- Sample: Year 2008 every 30 hours (skip every 4 of 5 cycles).



Average χ^2 wrt. precipitating members

NT **LOG** **GTcz** **GTbz**



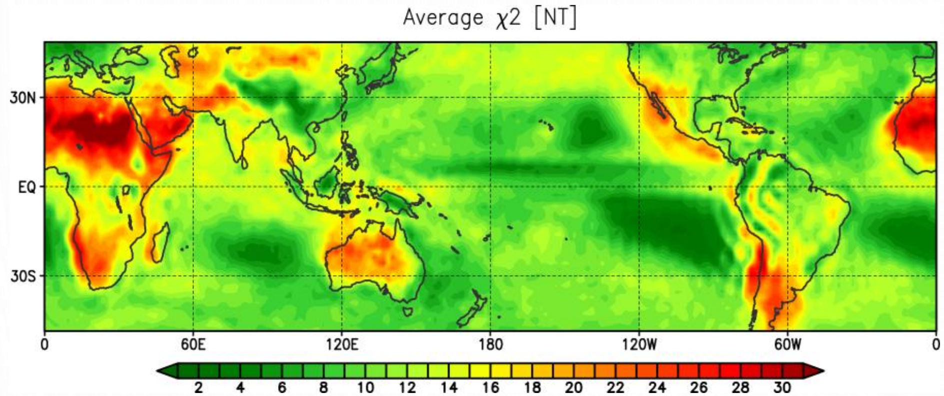
Relative improvement to NT

0 16 32 (all)

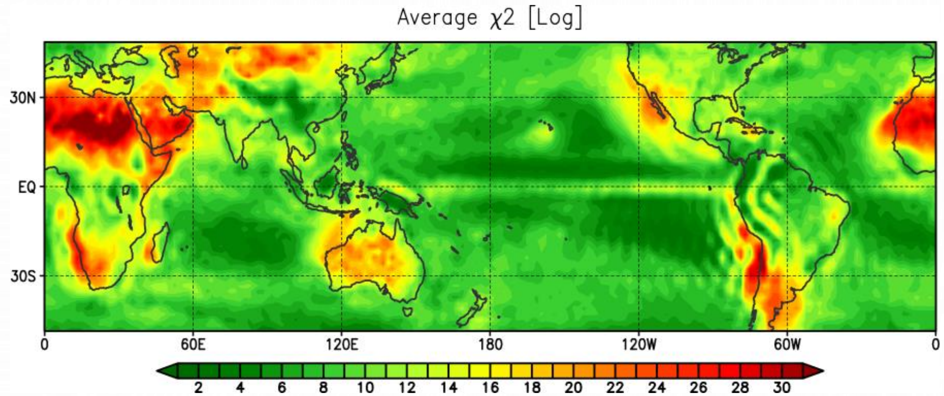
Precipitating members in the background

Average χ^2 maps

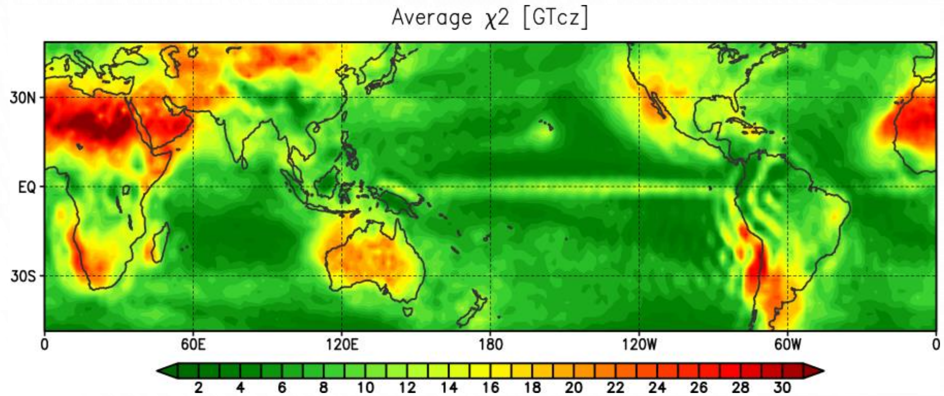
NT



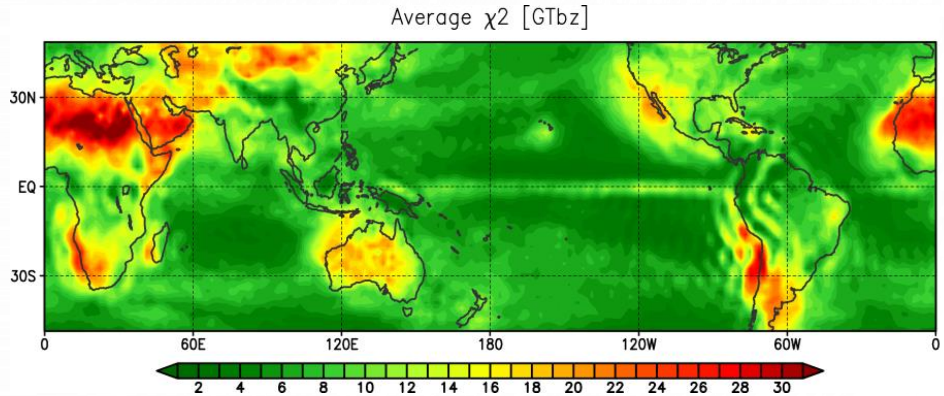
Log



GTcz



GTbz



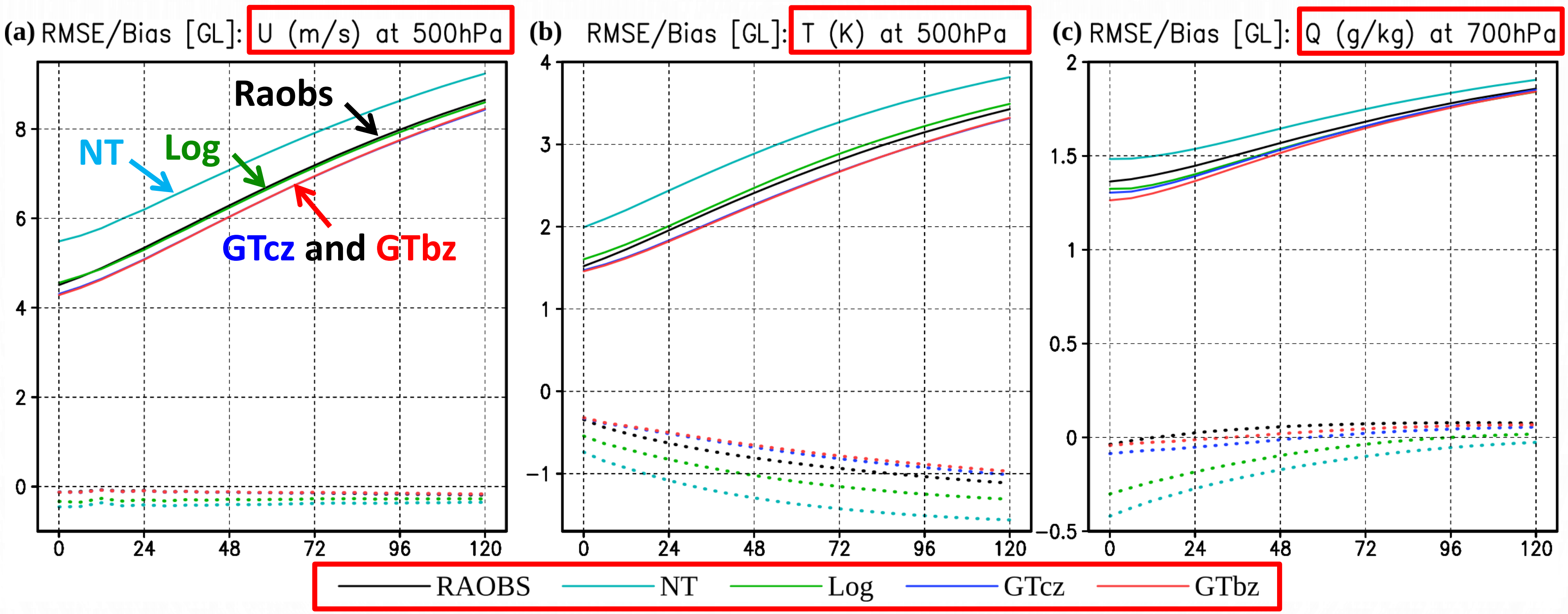
Problems and strategies in the real data case
(applicable to the large-scale, non-convective precipitation)

- Inconsistent probability distribution between model and observation climatology (amplitude-dependent biases).
  → Apply Gaussian transformation to model/observation precipitation separately to correct the biases.
- Timing errors in the forecast precipitation.
  → Use the 6-h accumulated amount.
- Wrong precipitation parameterization at some regions.
  → Don't use observations in those areas.

Part II: GFS-LETKF assimilation of TMPA data

- **RAOBS**: Conventional radiosonde only
- **NT**: Raobs + TMPA without transformation
- **Log**: Raobs + TMPA with the Log transformation
- **GTcz**: Raobs + TMPA with the GTcz transformation
- **GTbz**: Raobs + TMPA with the GTbz transformation
- General settings:
  - Horizontal localization: 500 km
  - Vertical localization: 0.4 ln(P)
- Settings for TMPA assimilation:
  - **24mR**: require >= 24 members (out of 32) are precipitating (> 0.06 mm/6h)
  - **Corr0.35**: Assimilated only at where Corr[GFSpp, TMPApp] > 0.35
  - No selection rule based on observed values
  - Horizontal localization: 350 km
  - Vertical localization: 0.4 ln(P), from 850 hPa
- Verification period: One-year (2008) cycling run after 1 month spin-up.
- Verified against the ECMWF ERA interim reanalysis.

Average RMSE/bias vs. forecast time

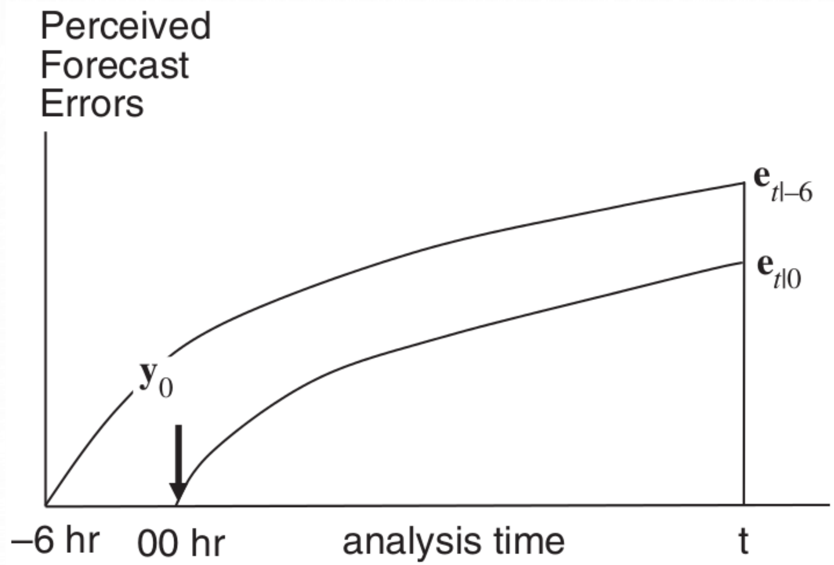


Global results Solid lines: RMS errors Dashed lines: Biases

- **NT** gives very bad results.
- **Log** transformation leads to marginal results.
  - Good for moisture, but bad for temperature.
- **GTcz** and **GTbz** are almost the same, both leading to clear positive impacts.

Part III: Ensemble forecast sensitivity of observations (EFSO)

- Estimate the forecast error reductions of any subset of observations.
- Economical alternative to OSEs.
- Thanks for the code and guidance from Daisuke Hotta and Yoichiro Ota.



$$\Delta e^2 = e_{t|0}^2 - e_{t|-6}^2 = \mathbf{e}_{t|0}^T C \mathbf{e}_{t|0} - \mathbf{e}_{t|-6}^T C \mathbf{e}_{t|-6}$$

$$\cong \frac{1}{K-1} \delta \mathbf{y}^T \, \mathbf{R}^{-1} \, \mathbf{Y}^a \, \mathbf{X}_{t|0}^{f\,T} \, C\left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right)$$

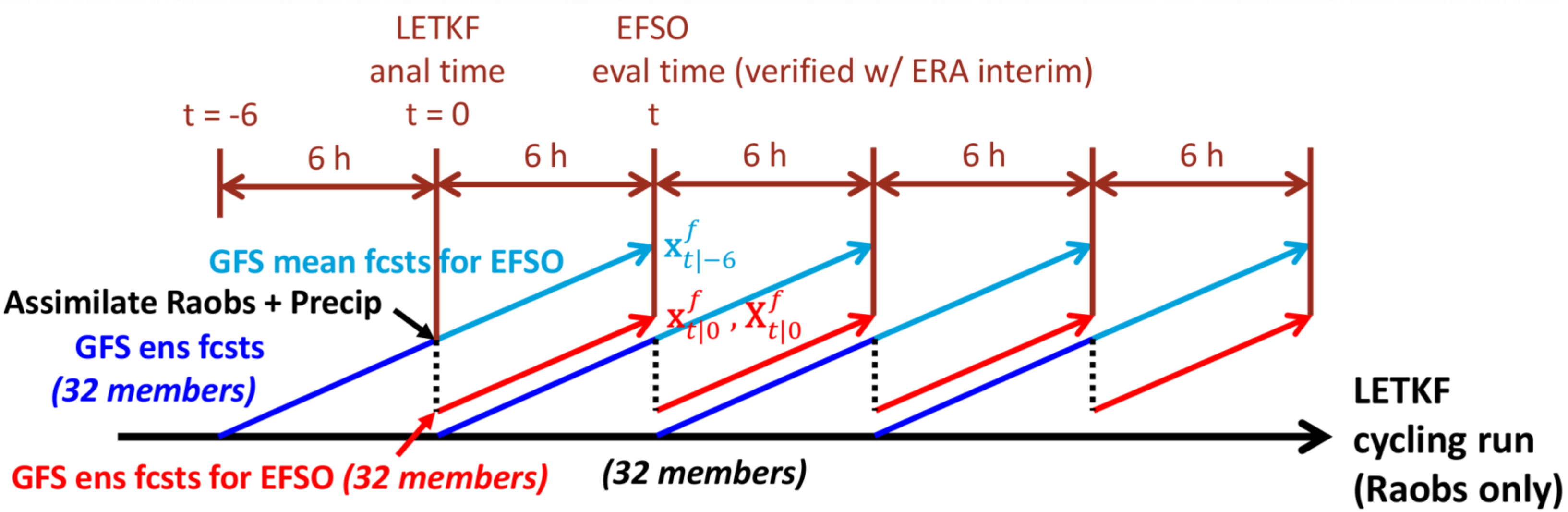
Moist total energy norm: $e^2 = \frac{1}{2}\frac{1}{S}\int_S \left[\int_0^1 \left(u'^2 + v'^2 + \frac{C_\mathrm{p}}{T_\mathrm{r}} T'^2 + \frac{L^2}{C_\mathrm{p} T_\mathrm{r}} q'^2\right) d\sigma + \frac{R_\mathrm{d} T_\mathrm{r}}{P_\mathrm{r}^2} P_\mathrm{s}'^2\right] dS$

Kinetic energy **Moist energy**

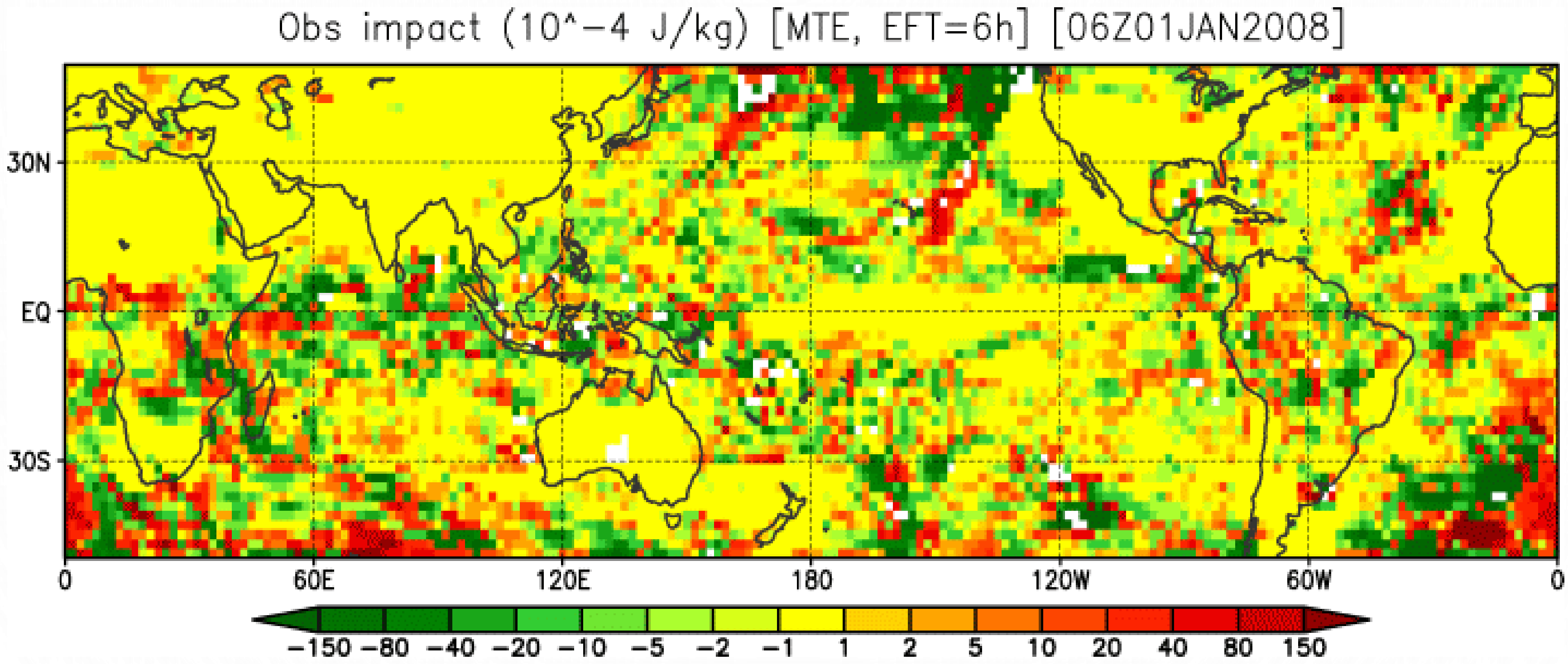
Potential energy

EFSO for precipitation

- Since all good and bad precipitation observations are assimilated, we intend not to carry out cycling assimilation in order to prevent the degrading trend of the analysis.
- Evaluation forecast time (EFT): 6 hours
- EFSO values are computed for each precipitation observations, then averaged in terms of various factors.
- Sample: Year 2008 every 30 hours (skip every 4 of 5 cycles).

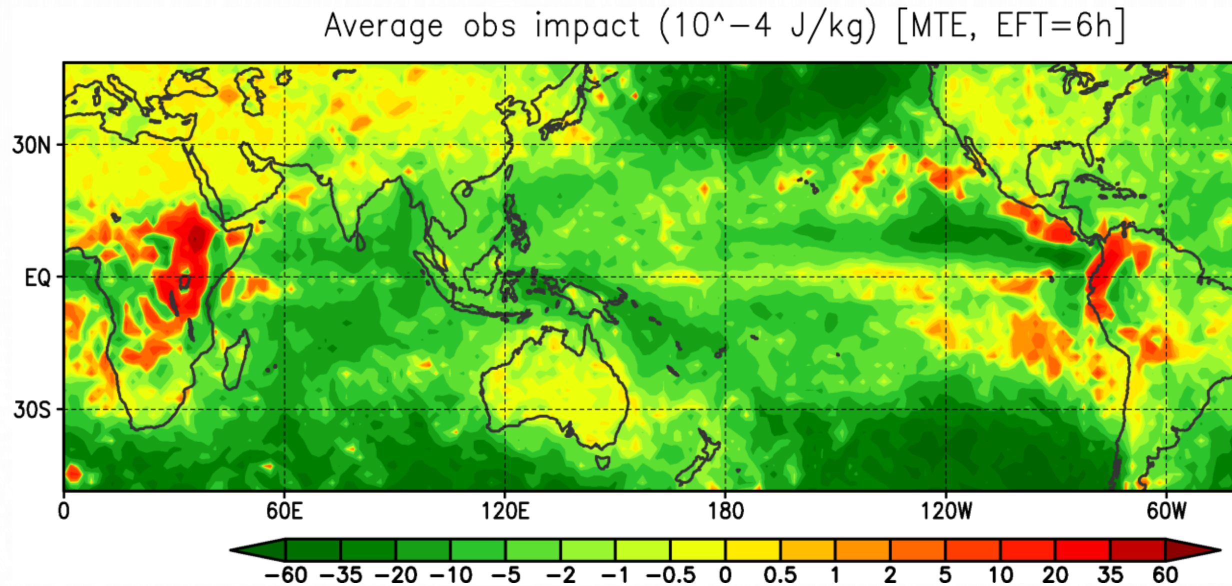


EFSO of TMPA precipitation

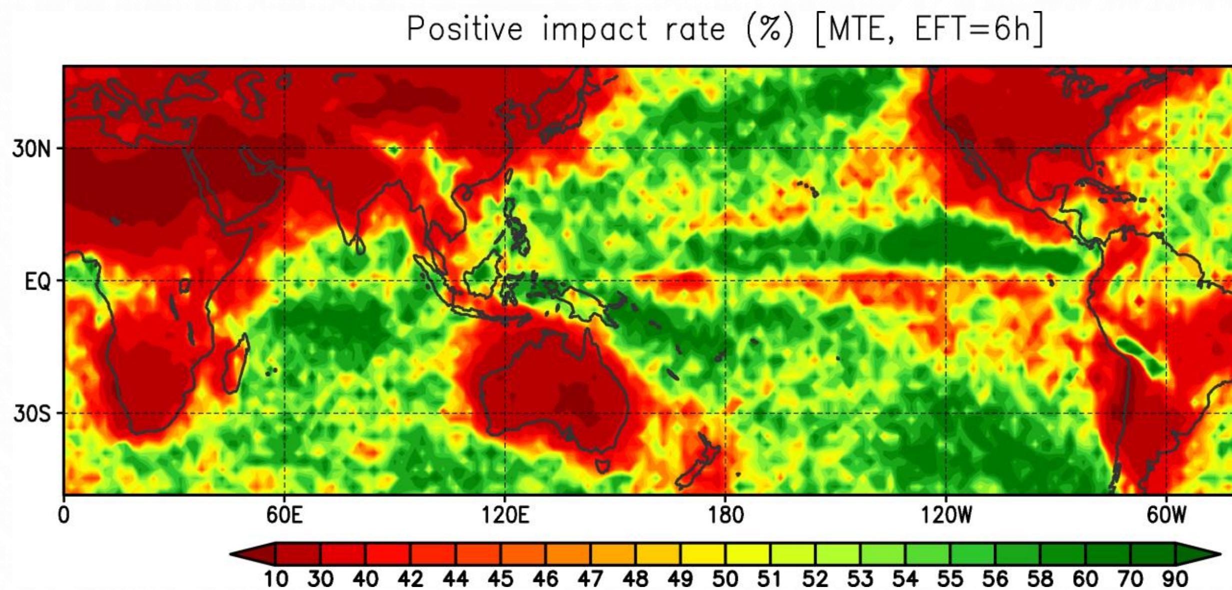


Average EFSO maps

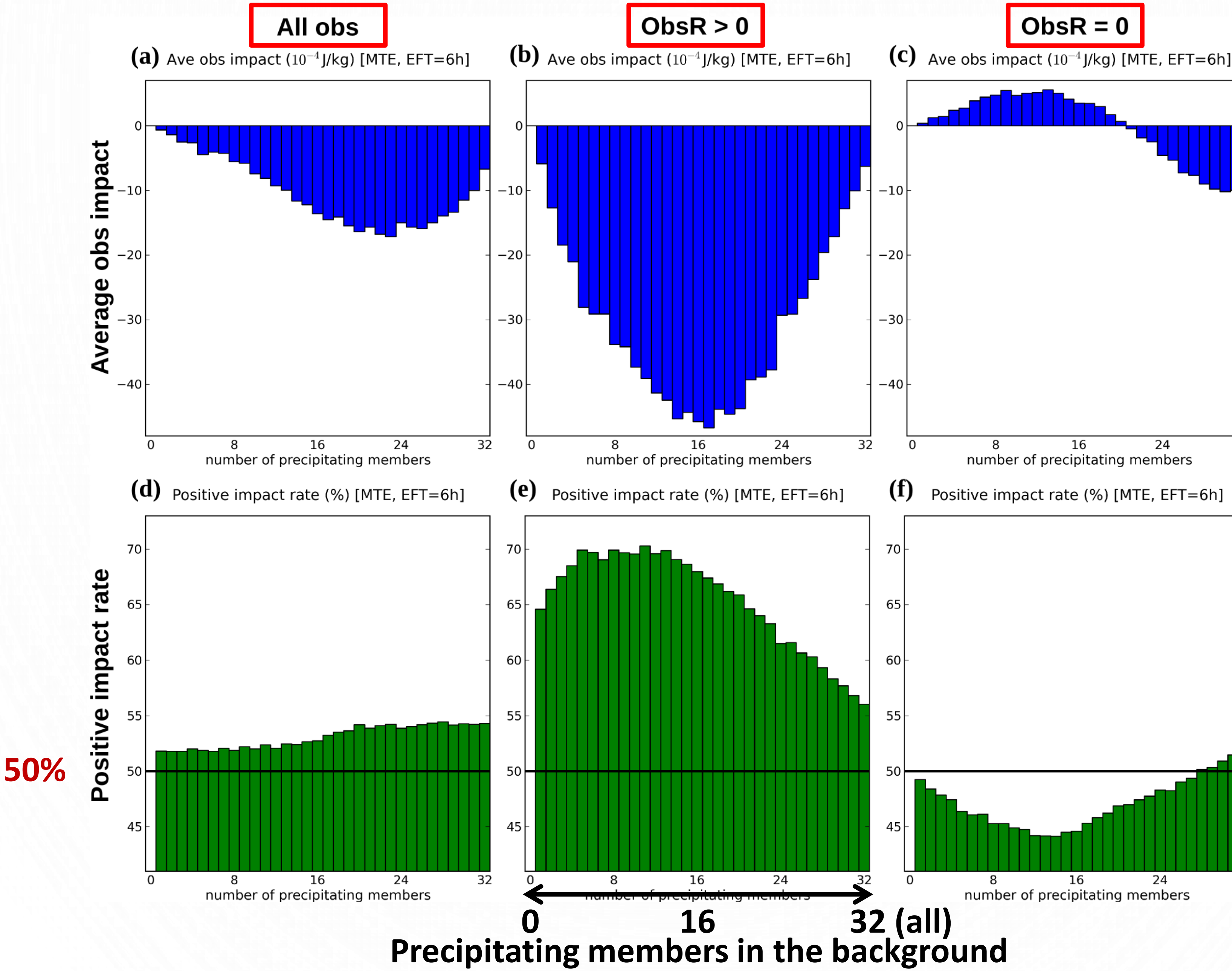
Average forecast error reduction



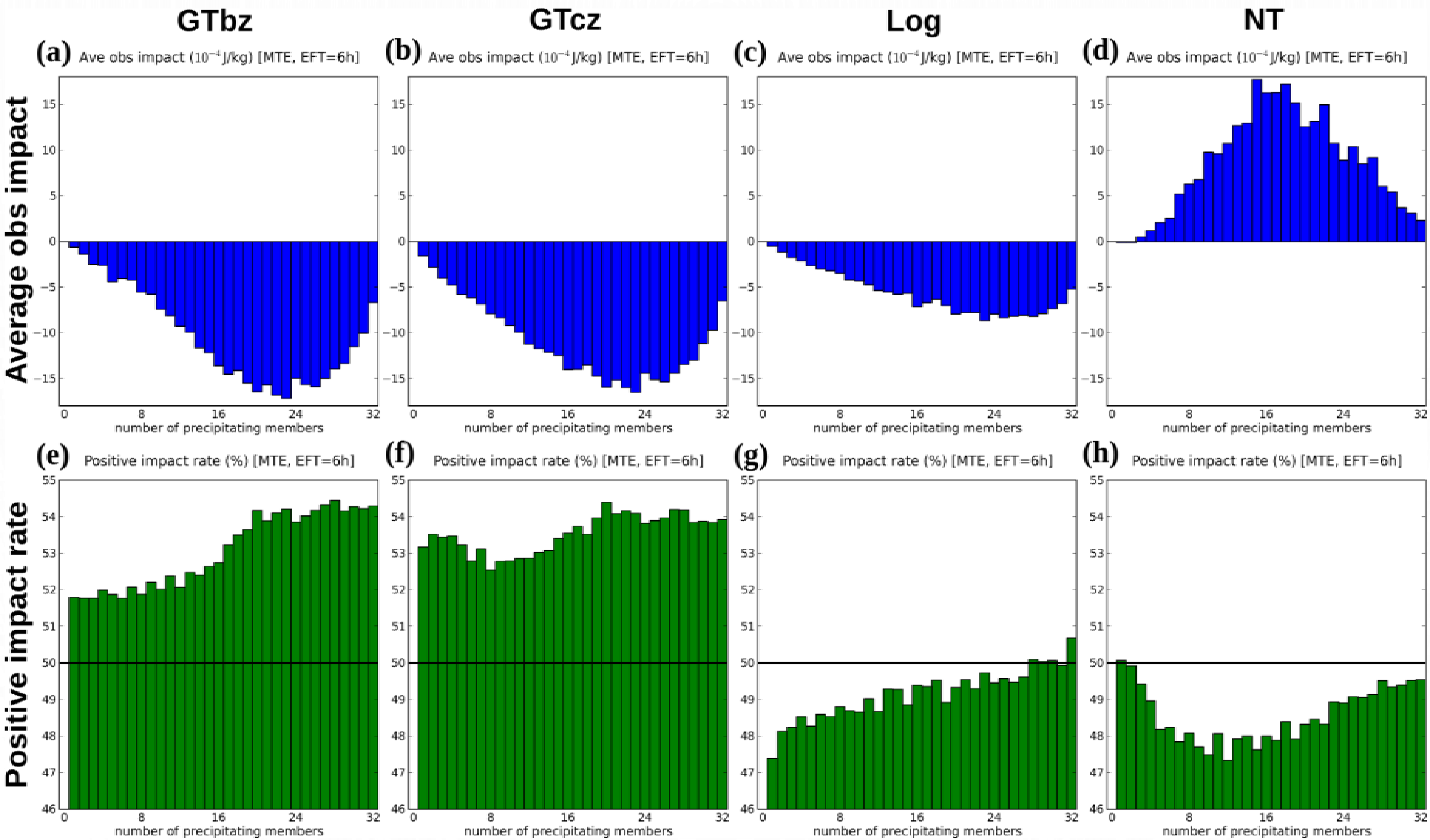
Positive impact rates



EFSO wrt. precipitating members

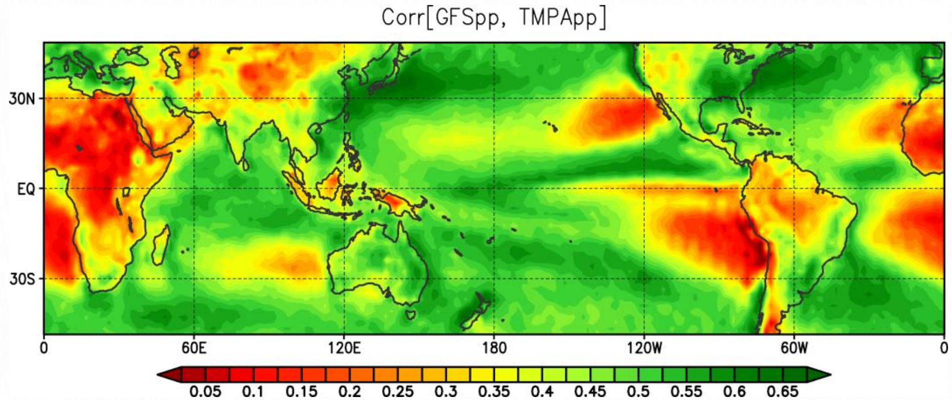


EFSO using different transformation methods

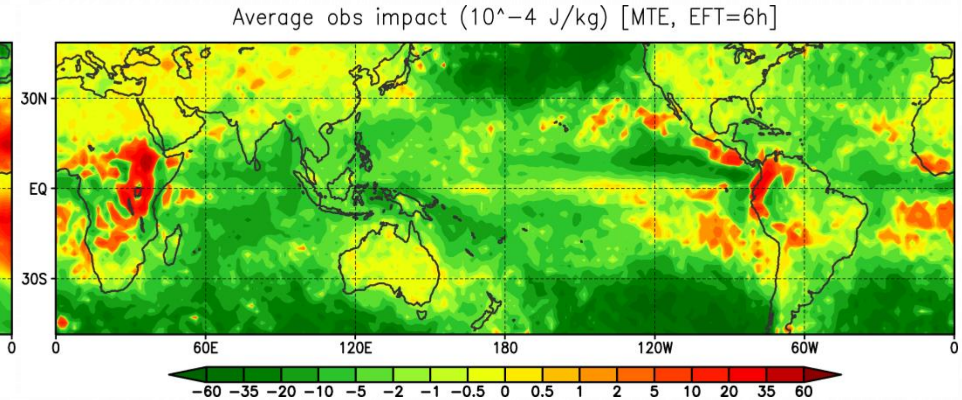


Reconsideration of the precipitation QC

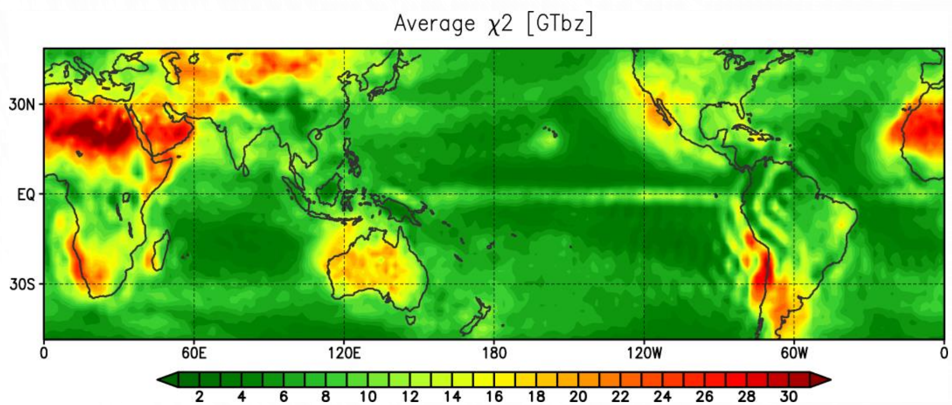
Correlation between the model and observations



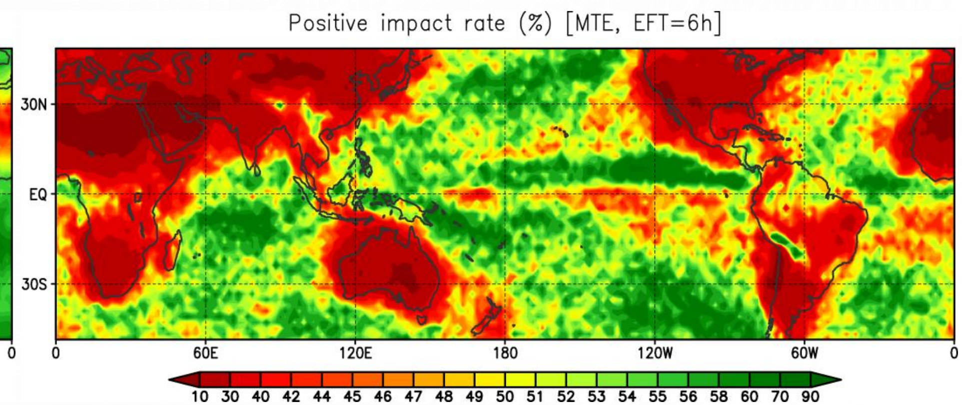
EFSO: Average observation impacts



Average (non-)Gaussianity (χ^2)



EFSO: Positive impact rates



Summary

- We successfully obtained positive impacts by assimilating precipitation, in both idealized OSSEs and a realistic model and observations, using the LETKF and the Gaussian transformation.
- The impacts are seen in both analyses and 5-day forecasts.
- Gaussian transformation is beneficial to the precipitation assimilation:
  - Applying Gaussian transformation to model/observation precipitation separately can correct the bias and increase the correlation between these two quantities.
  - Gaussian transformation based on the climatology does produce more Gaussian background error distribution of precipitation.
- Statistical characteristics of the precipitation variable can give us useful hints in the real precipitation data assimilation.
- The model error is a very important issue in the real precipitation data assimilation.
- We demonstrated how to use the EFSO to efficiently analyze the effectiveness of a new observing system.

Acknowledgements

- Advisors:
  - Prof. Eugenia Kalnay
  - Prof. Takemasa Miyoshi
- Guidance on TMPA:
  - Dr. George Huffman
- Guidance on GFS/GSI
  - Dr. Daryl Kleist
  - Dr. Henry Huang
  - Dr. Runhua Yang
  - Tetsuro Miyachi
- EFSO code:
  - Yoichiro Ota
  - Daisuke Hotta
- Other discussion:
  - Dr. Ross Hoffman
  - Prof. Kayo Ide
  - Prof. Brian Hunt
  - Prof. James Carton
  - Prof. Takeshi Enomoto
- Grants:
  - NASA (including the support of Discover cluster)
  - NOAA
  - National Oceanographic Partnership Program (NOPP)