Local Ensemble Transform Kalman Filter Assimilation of Precipitation with the NCEP Global Forecasting System

Guo-Yuan Lien[1], Eugenia Kalnay[1], Takemasa Miyoshi[1,2]

[1]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, USA

[2] RIKEN Advanced Institute for Computational Science, Kobe, Japan

National Oceanographic Partnership Program

Introduction

- Many in-situ and satellite based precipitation observations have been made available, but the assimilation of precipitation is still difficult because of:
  - the nonlinear observation operator.
  - the non-Gaussianity of the precipitation variable.
  - the imperfect precipitation parameterization in the numerical model.
  - the unknown errors associated with the precipitation observations.
- It is relatively easy to force the model precipitation to be close to the observed values; however, since this is not an efficient way to modify the potential vorticity field that the model would remember, model forecasts tend to lose their additional skill after few forecast hours.
- Proposed method of precipitation assimilation:
  - Local ensemble transform Kalman filter (LETKF).
  - Cumulative distribution function (CDF)-based transformation applied to the precipitation variable (instead of logarithm transformation).
  - Ensemble background-based observation selection criterion (instead of observation-based criterion).
- We first successfully tested our idea with a simplified but still realistic general circulation model (i.e., SPEEDY model) in a observing system simulation experiment (OSSE) framework. We are currently working on the assimilation of the TRMM Multisatellite Precipitation Analysis (TMPA) with the NCEP Global Forecasting System (GFS).

CDF-based Gaussian transformation

- The "Gaussian anamorphosis" (also used by Schöniger et al. 2012 in hydrology):

$$y_{trans} = G^{-1}[F(y)]$$

y : Original variable.

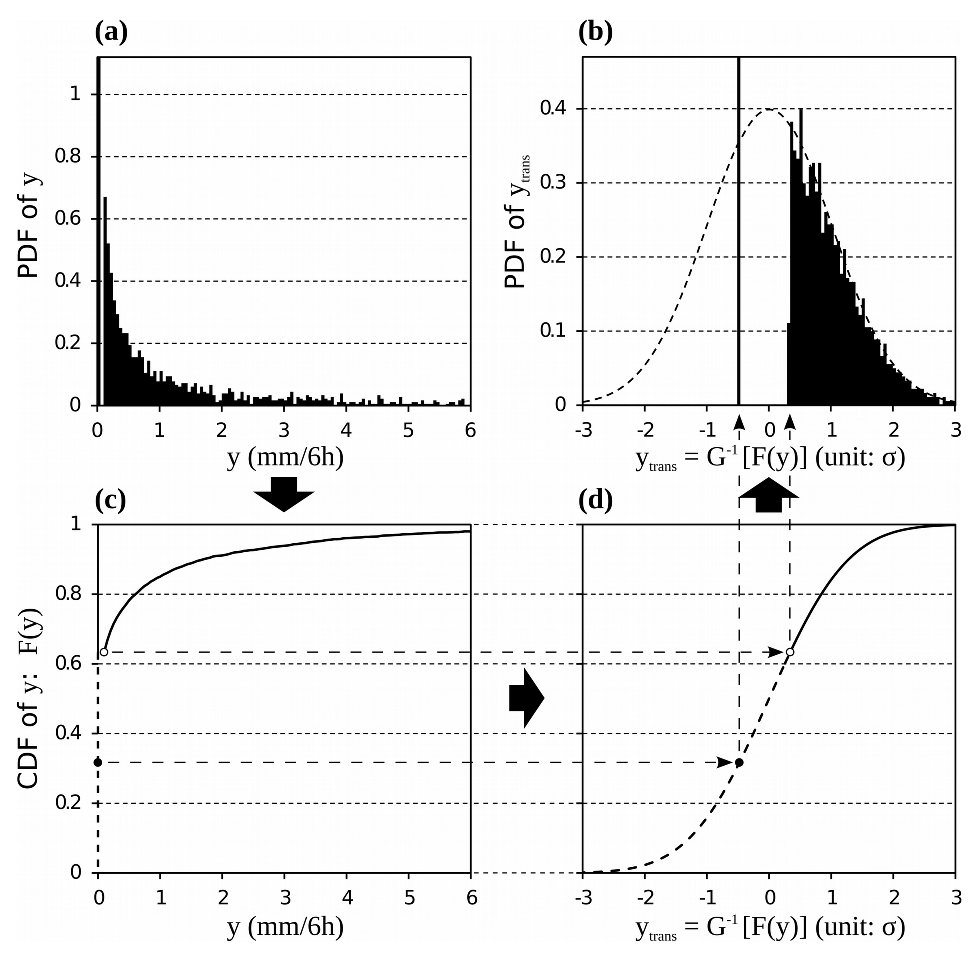
F : Empirical cumulative distribution function (CDF) of y based on a long period of observations or model climatology.

G^{-1} : Inverse CDF of normal distribution.

$$G^{-1}(x) = \sqrt{2}\,\mathrm{erf}^{-1}(2x-1)$$

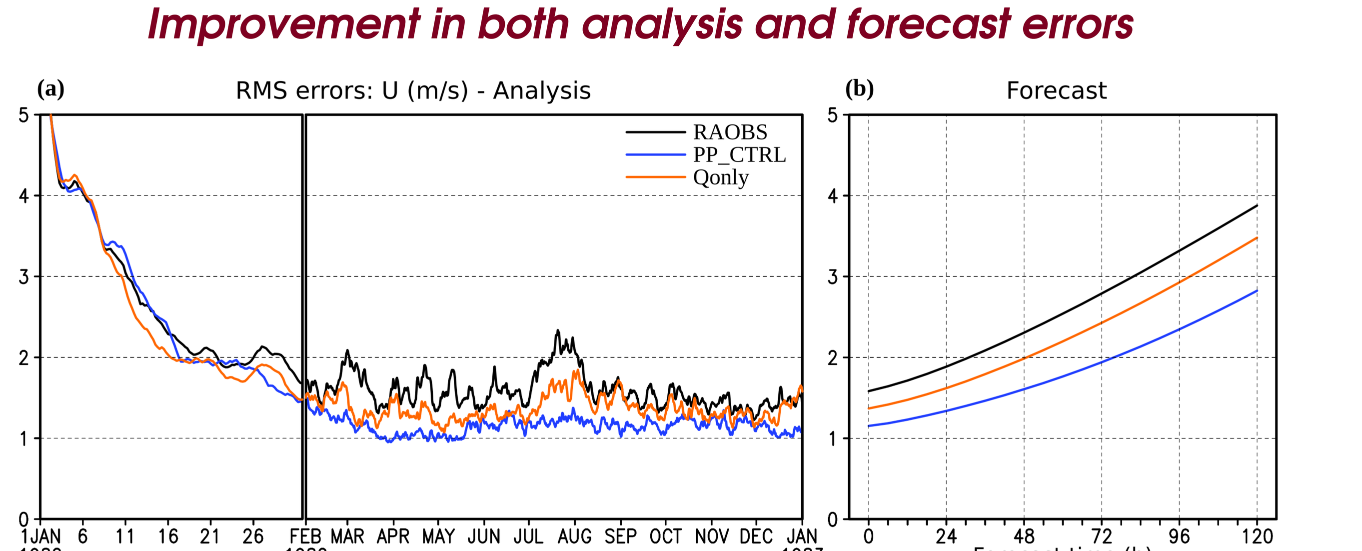
- Precipitation variables contain a large portion of zero values.
  - Zero precipitation values have to be considered in the transformation.
  - A natural choice: assigning the middle value of zero-precipitation cumulative probability to $F(0)$.
- LETKF assimilation of precipitation is performed on the transformed space.
- The transformation can apply to observed values and model background values separately. In this case, it not only transforms an arbitrary variable into a Gaussian variable, but also functions as a "CDF-based bias correction."

Example of the Gaussian transformation



Perfect model experiment with SPEEDY model

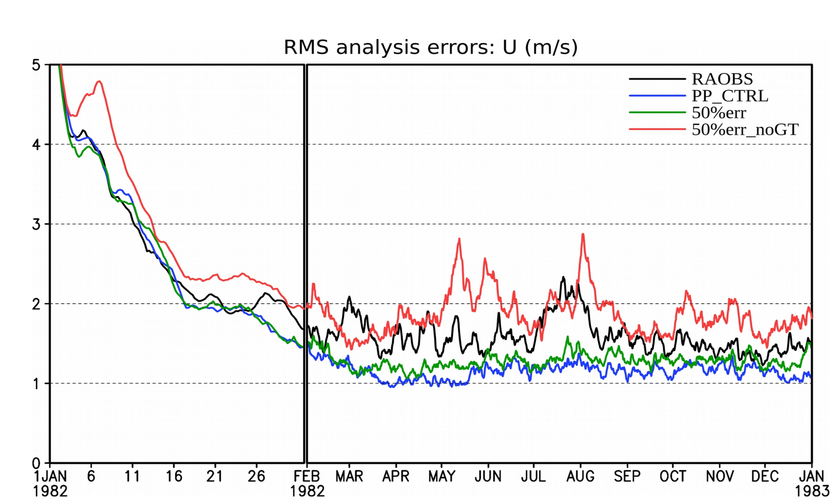
Improvement in both analysis and forecast errors



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

- **RAOBS**: Assimilate rawinsonde observations
- **PP_CTRL**: Assimilate rawinsonde observations + uniformly distributed global precipitation
- **Qonly**: Same as PP_CTRL, but only update moisture field by precipitation assimilation

Impact of Gaussian transformation

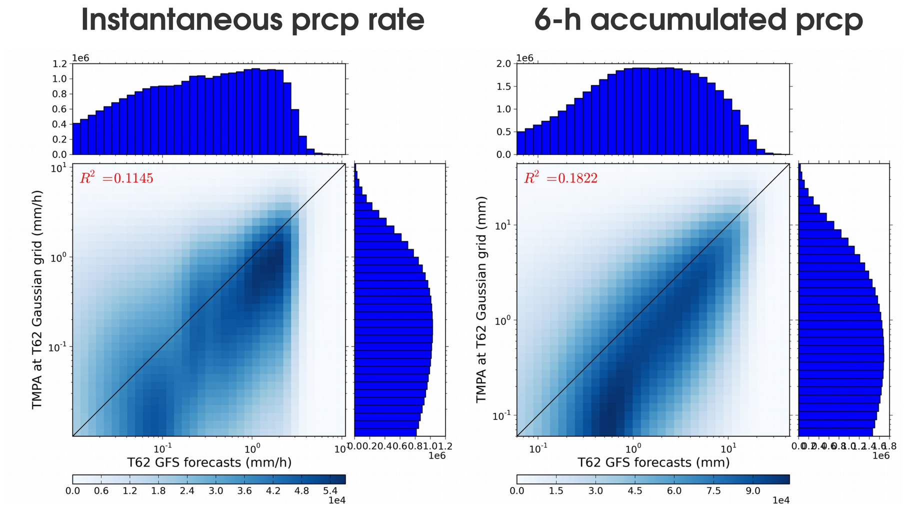


- **50%err**: Same as PP_CTRL, but increase the observation error of precipitation observations from 20% to 50%.
- **50%err_noGT**: Same as 50%err, but do not use the Gaussian transformation

Statistics with TMPA and GFS 3-9 hour forecasts

- We target to run assimilation experiments at a T62 resolution. TMPA data are upscaled to the Gaussian grid used by the T62 GFS model using an areal conservative remapping.
- 2001-2010 (10 year) period is chosen to compute all these statistics.
- 9-hour GFS model forecasts initialized from every 6-hour NCEP Climate Forecast System Reanalysis (CFSR) are conducted within the 10-year period, in which the 3-9 hour forecasts (i.e., assimilation window) are compared to the TMPA data. In the LETKF data assimilation, this period of model forecasts is used as background.

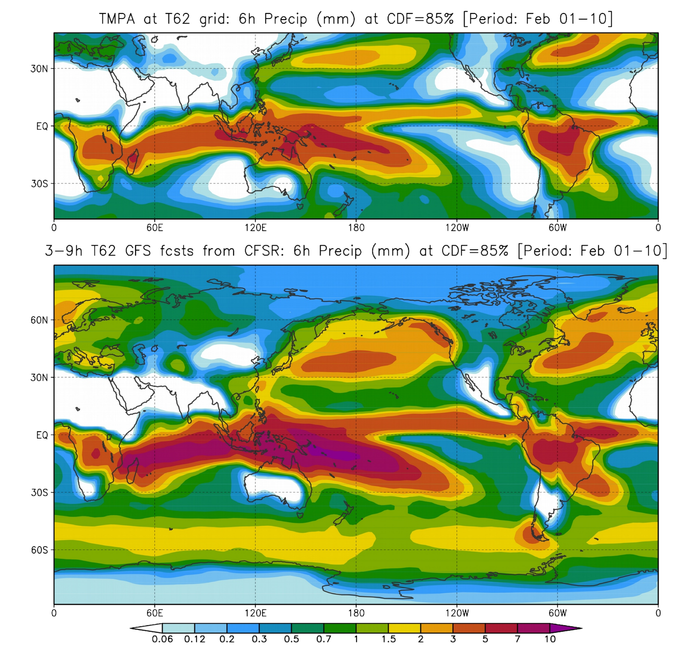
Instantaneous prcp rate v.s. 6-h accumulated prcp



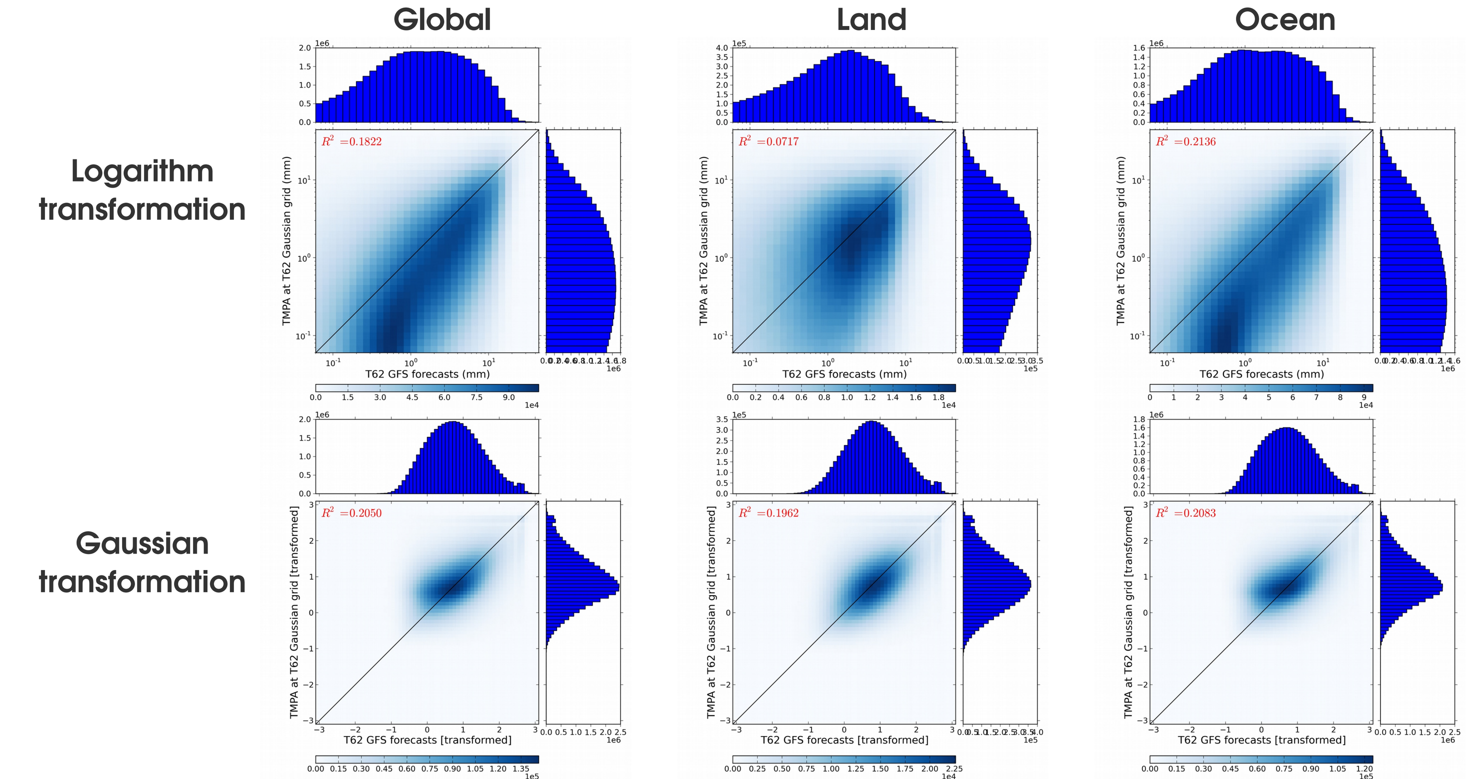
- Figures plotted with logarithm transformation.
- Only positive precipitation is shown in all distribution figures.

CDFs of both TMPA and GFS forecasts

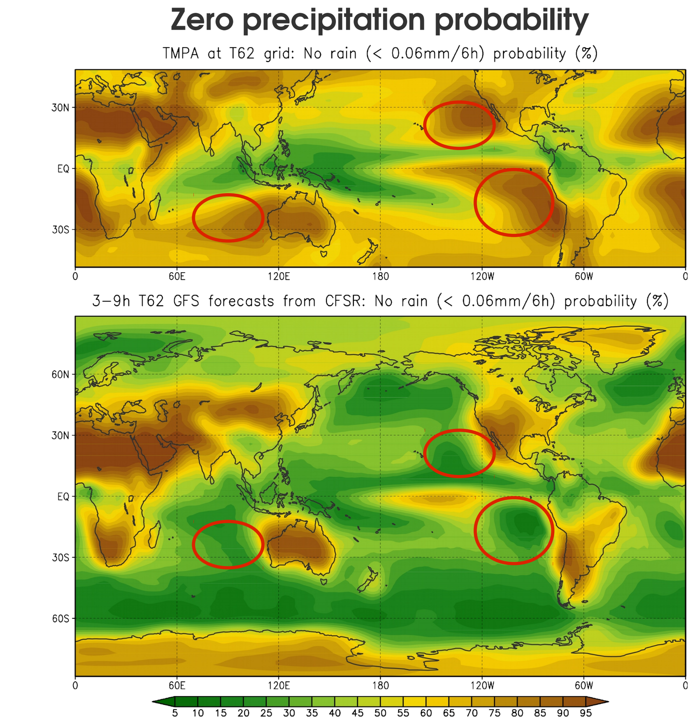
- Empirical CDFs as a function of geographic location and 10-day period of the year (i.e., 1-10 Jan., 11-20 Jan., … etc.) are constructed from the 10-year TMPA data and GFS forecasts, in order to define the Gaussian transformation of the precipitation variables for both observations and model background.
- When computing the CDF at each grid point and each 10-day period, all data within 500-km radius and +/- 20 day period are considered in order to obtain spatially and temporally smooth CDFs.



Logarithm transformation v.s. Gaussian transformation



Problem with the marine stratocumulus precipitation in the GFS



Conclusions

- In observing system simulation experiments (OSSEs) using a simplified general circulation model, we have obtained a very promising improvement by assimilating precipitation.
  - The Gaussian transformation of precipitation variables is useful in the case of 50% observation errors.
- Assimilation of the TMPA data with the NCEP GFS model is in progress. Several statistics have been performed to understand the sources of the difficulties of the precipitation assimilation with real data and a real model.
  - Accumulated precipitation is a better variable rather than the instantaneous precipitation rate for precipitation assimilation.
  - Empirical CDFs as a function of geographic location and 10-day of the year are constructed for both TMPA and GFS 3-9 h forecasts.
  - Applying Gaussian transformation to both precipitation observations and model background precipitation corrects the bias and increases the correlation between these two quantities.
  - The GFS model has a severe problem in parameterizing the marine stratocumulus precipitation at a T62 resolution.
- ***Future plans***: OSSE assimilation of precipitation is successful in the LETKF because of its multivariate character, which allows it to modify the potential vorticity field, rather than the column moisture or temperature as done in conventional approaches. With real data, if the model fails to represent the dynamics leading to precipitation in the real data, it is more difficult to successfully assimilate precipitation, so that we have not yet obtained clear improvements from assimilating TMPA observations. Based on the precipitation statistics we will focus initially on regions where precipitation is determined by potential vorticity dynamics rather than, e.g., marine stratocumulus or convection. This will require defining additional criteria to select the TMPA precipitation observations that should be assimilated in the LETKF.