Bologna 2023

Incorporating non-Gaussian observation errors into variational methods

Chih-Chi Hu

Princeton University, Colorado State University

Peter Jan van Leeuwen

Colorado State University

Alan Geer

European Centre for Medium-Range Weather Forecasts

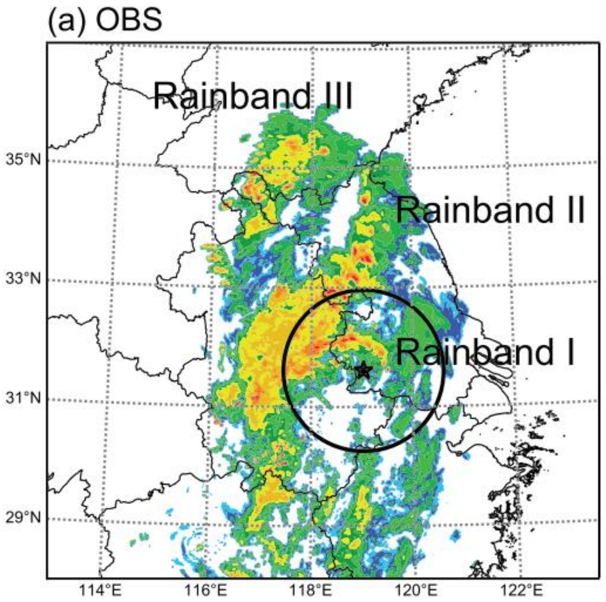
2023/10/19 (Thu)

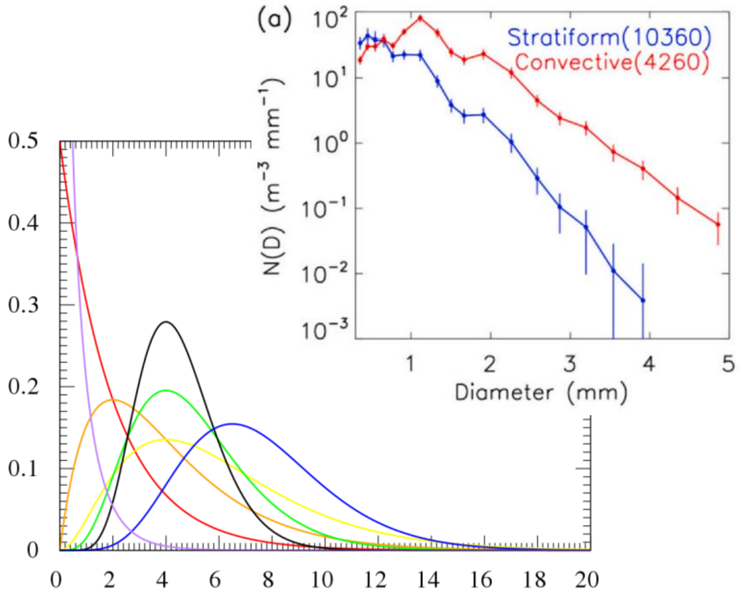
The observation errors can be quite non-Gaussian due to the errors in the observation operator

- Consider a perfect model state and a perfect observation:

$$y^{truth} \neq H\left(x^{truth}\right)$$

- The errors in the observation operator H come from
  - e.g., inappropriate assumptions in H





figures from Wang et al. (2020); Thomas et al. (2021)

<u>Example</u>: cloudy radiances

$$H = \left(RTM \circ M_{t_{initial} \rightarrow t_{obs}}\right)$$

The error can be quite non-Gaussian when predicting clouds (e.g., due to the displacement error)

How do we incorporate non-Gaussian errors into DA?

- Nonlinear DA methods, e.g., Particle Flow Filter *Daum and Huang (2011); Liu and Wang (2016); Pulido and van Leeuwen (2019); Hu and van Leeuwen (2021)*

- Variational methods
  - Derive the cost-function for the non-Gaussian pdf
  - Examples: varQC *Anderson and Järvinen (1999)*, log-normal *Fletcher and Zupanski (2006)*, Huber norm *Huber (1972); Tavolato and Isaksen (2015)*

- These methods require a prior assumption of the parametric form of the error distribution.

- **Motivation**: to propose a method that can incorporate non-Gaussian errors into the variational methods in a more general way

The "evolving-Gaussian" method for the variational methods

- Key idea: approximate the cost-function iteratively using Gaussian errors.

$J^{NG}(x)$ the cost-function from an arbitrary non-Gaussian pdf

$J^{NG}(x_0)$ —— find a Gaussian G_0 s.t. $\nabla J^{NG}(x_0) = \nabla J^{G_0}(x_0)$ ——→ $J^{G_0}(x_0)$

the first-guess (background) ↗ x_0

update the state ↓

$J^{NG}(x_1)$ —— find G_1 s.t. $\nabla J^{NG}(x_1) = \nabla J^{G_1}(x_1)$ ——→ $J^{G_1}(x_1)$

⋮

⋮

How do we find “the Gaussian(s)” in the evolving-Gaussian method?

- Assumptions:
  - Independent observation errors ⇒ the observation error pdf reduces to a 1-D pdf
  - We have access to the 1-D observation error pdf, denoted as $f_{\varepsilon^o}\big(y - H(x)\big)$

- Consider the cost-function for one observation (at the *i* th iteration):

$$J^{NG}(x_i) = J_b(x_i) + J_o^{NG}(x_i)$$

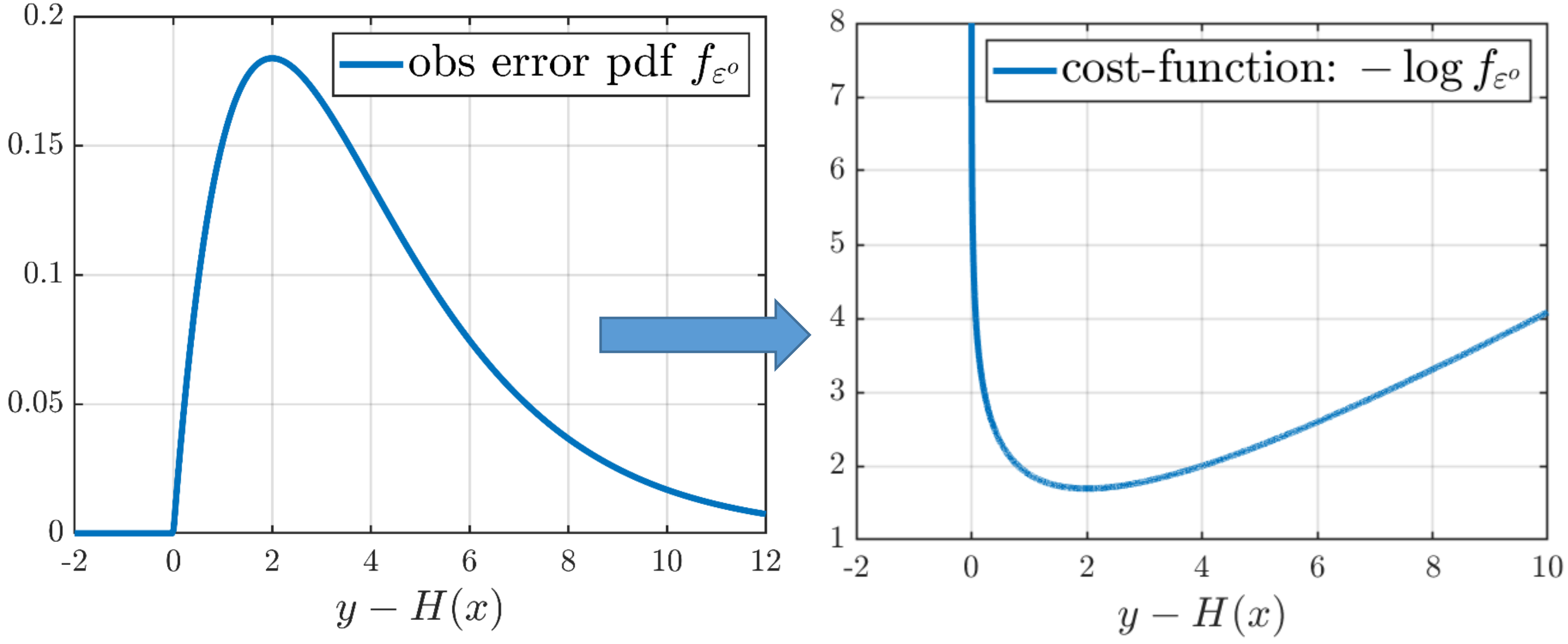
$$J_o^{NG}(x_i) = -\log f_{\varepsilon^o}\left(y - H(x_i)\right)$$

$$J_o^{G_i}(x_i) = \frac{1}{2}\left(\frac{H(x_i) - y + \mu_i}{\sigma_i}\right)^2$$

set μ_i = the mode of f_{ε^o}

and solve σ_i by $\nabla J_o^{NG}(x_i) = \nabla J_o^{G_i}(x_i)$

An example of finding “the Gaussian(s)” for a Gamma pdf



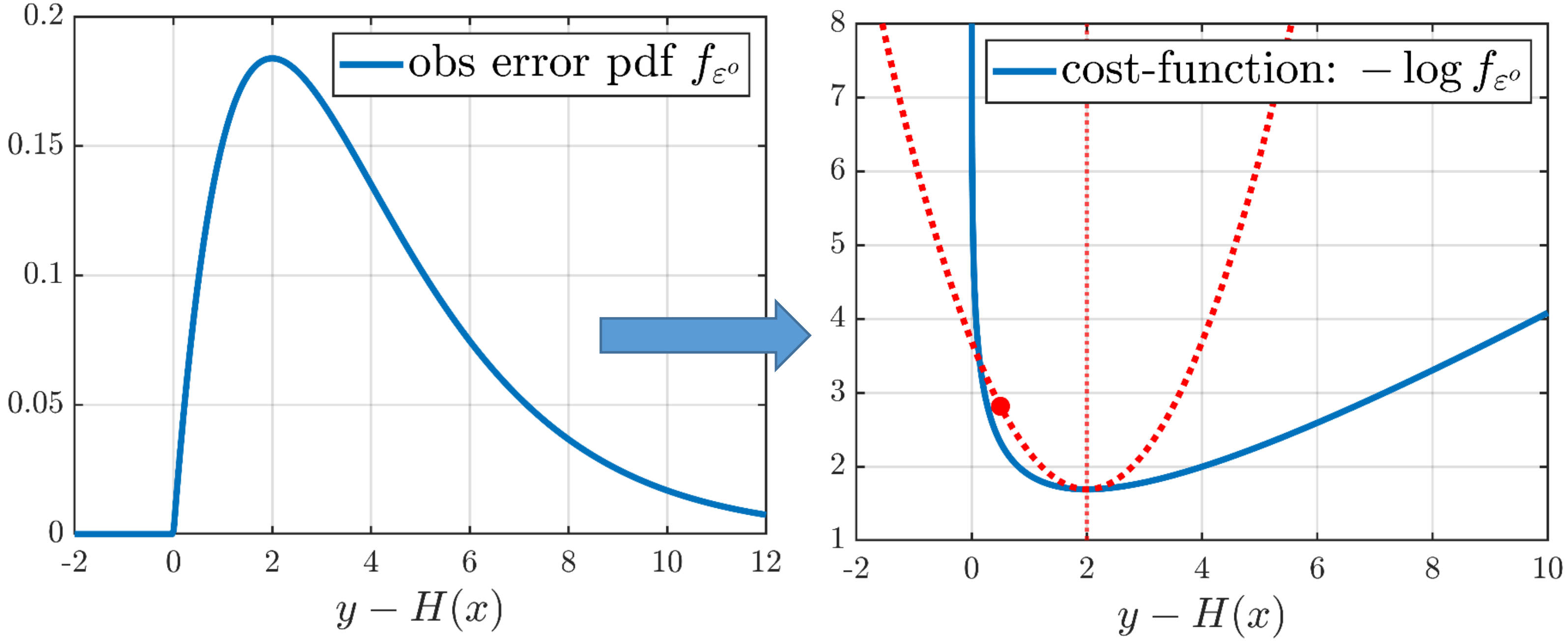
non-Gaussian obs error pdf

$$f_{\varepsilon^o}\big(y - H(x)\big)$$

non-Gaussian cost-function

$$J_o^{NG} = -\log f_{\varepsilon^o}$$

An example of finding “the Gaussian(s)” for a Gamma pdf

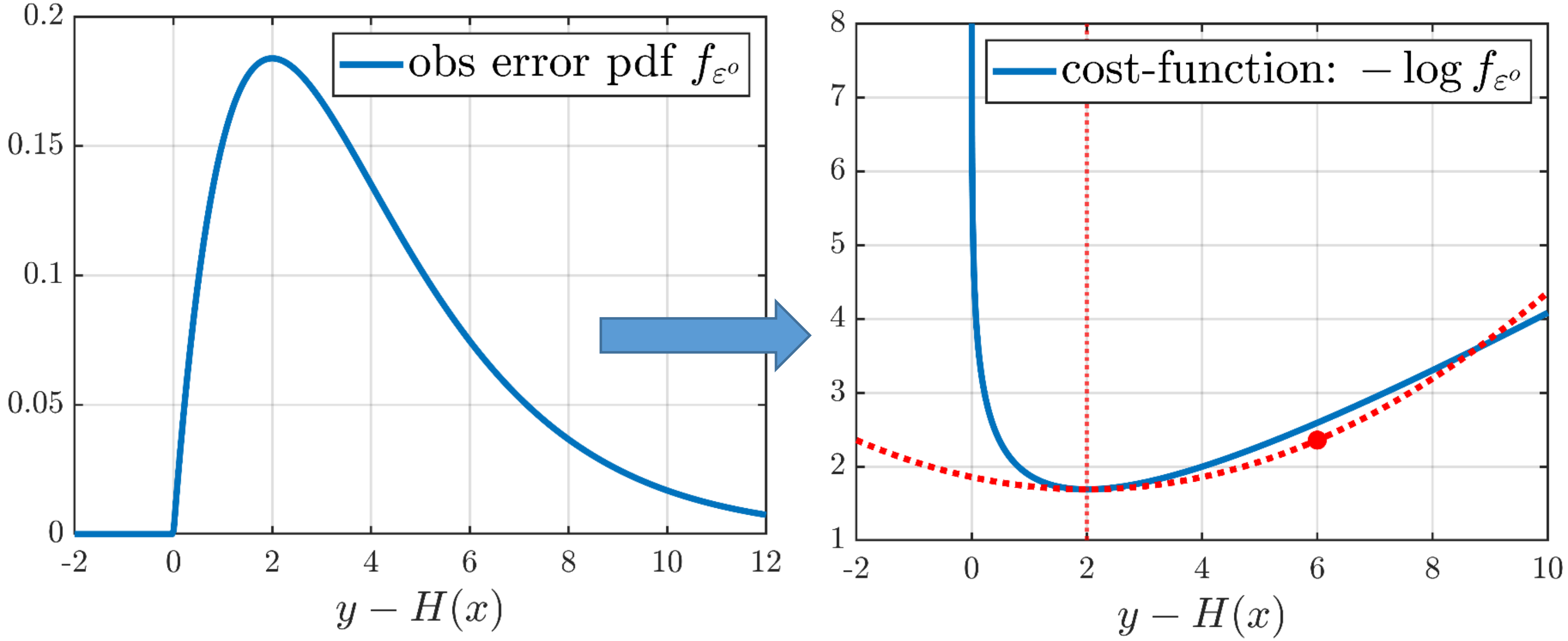


non-Gaussian obs error pdf $f_{\varepsilon^o}\big(y - H(x)\big)$

(blue) non-Gaussian cost-function $J_o^{NG} = -\log f_{\varepsilon^o}$

(red) Gaussian cost-function $J_o^{G_i}$ when $y - H(x) = 0.5$

An example of finding “the Gaussian(s)” for a Gamma pdf

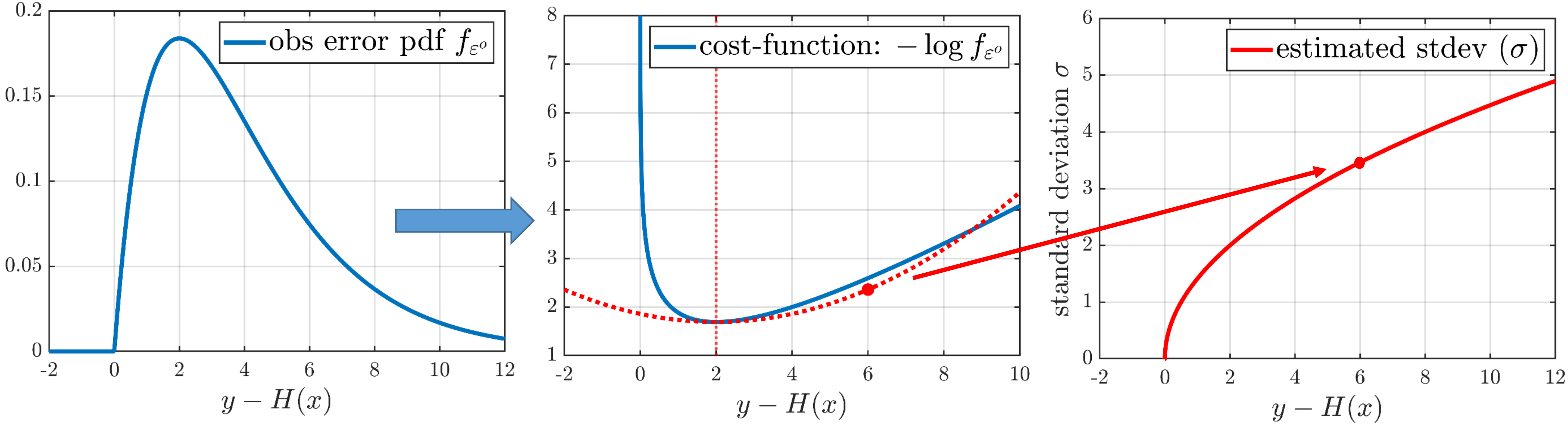


non-Gaussian obs error pdf
$f_{\varepsilon^o}\big(y - H(x)\big)$

(blue) non-Gaussian cost-function $J_o^{NG} = -\log f_{\varepsilon^o}$

(red) Gaussian cost-function $J_o^{G_i}$ when $y - H(x) = 6.0$

An example of finding “the Gaussian(s)” for a Gamma pdf



non-Gaussian obs error pdf

$f_{\varepsilon^o}\big(y - H(x)\big)$

non-Gaussian cost-function

$J_o^{NG} = -\log f_{\varepsilon^o}$

Summary of the Gaussian(s)

$\sigma\big(y - H(x)\big)$

Remarks of the evolving-Gaussian method

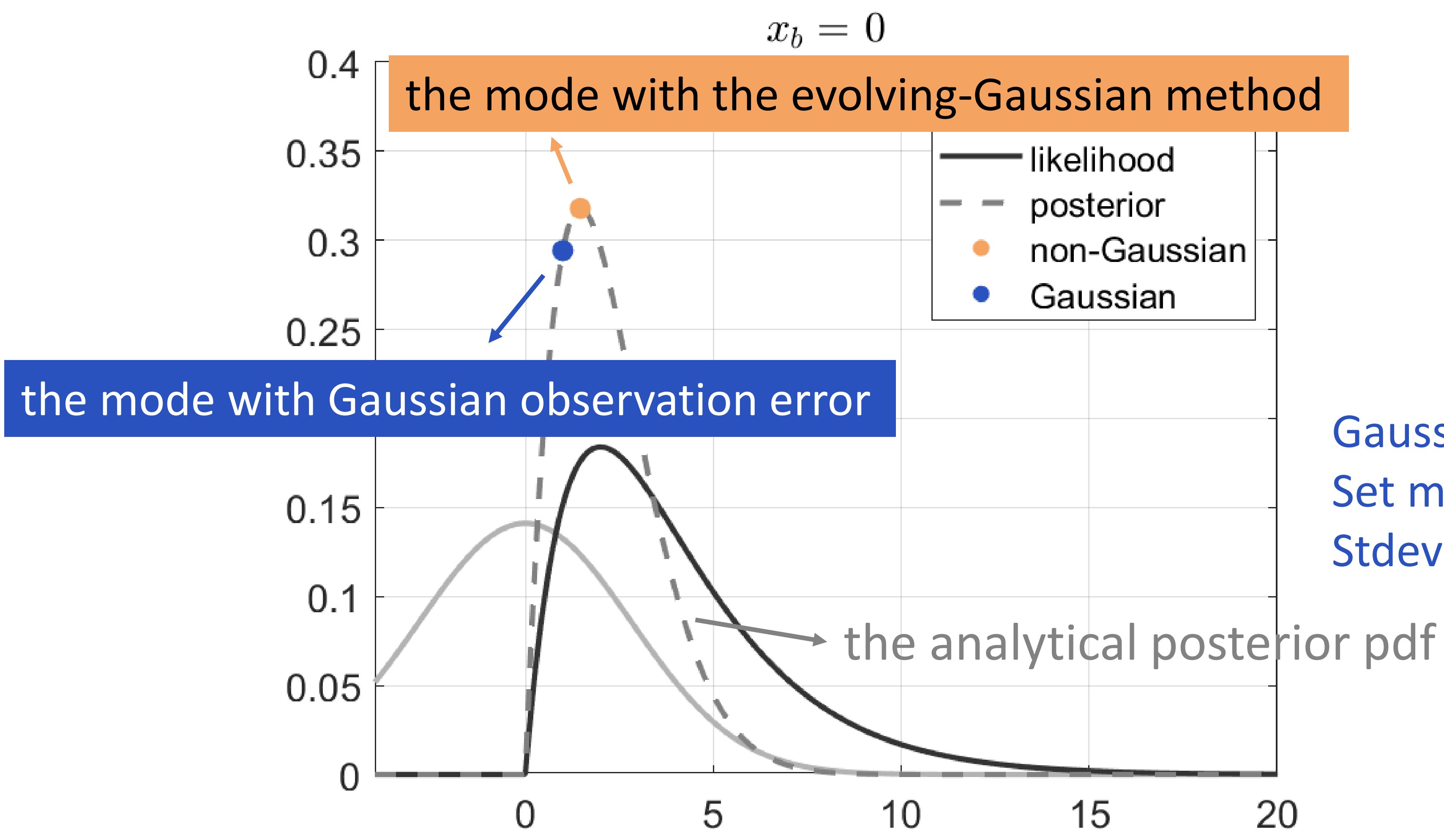
- Easy: the cost-function remains the same, no need to change the cost-function at all.

- Efficient: all the work can be done offline, only need a lookup table $\sigma\big(y - H(x)\big)$

- In theory, we need to evolve the Gaussian for every inner and outer loops.
  - need to evaluate H for each inner loop

- What if we **only** evolve the Gaussian in the outer loop?
  - A similar idea to linearize H in each outer loop in the incremental (3)4D-Var
  - The cost-function remains quadratic in the inner loops

An idealized test for the evolving-Gaussian method in the outer loop

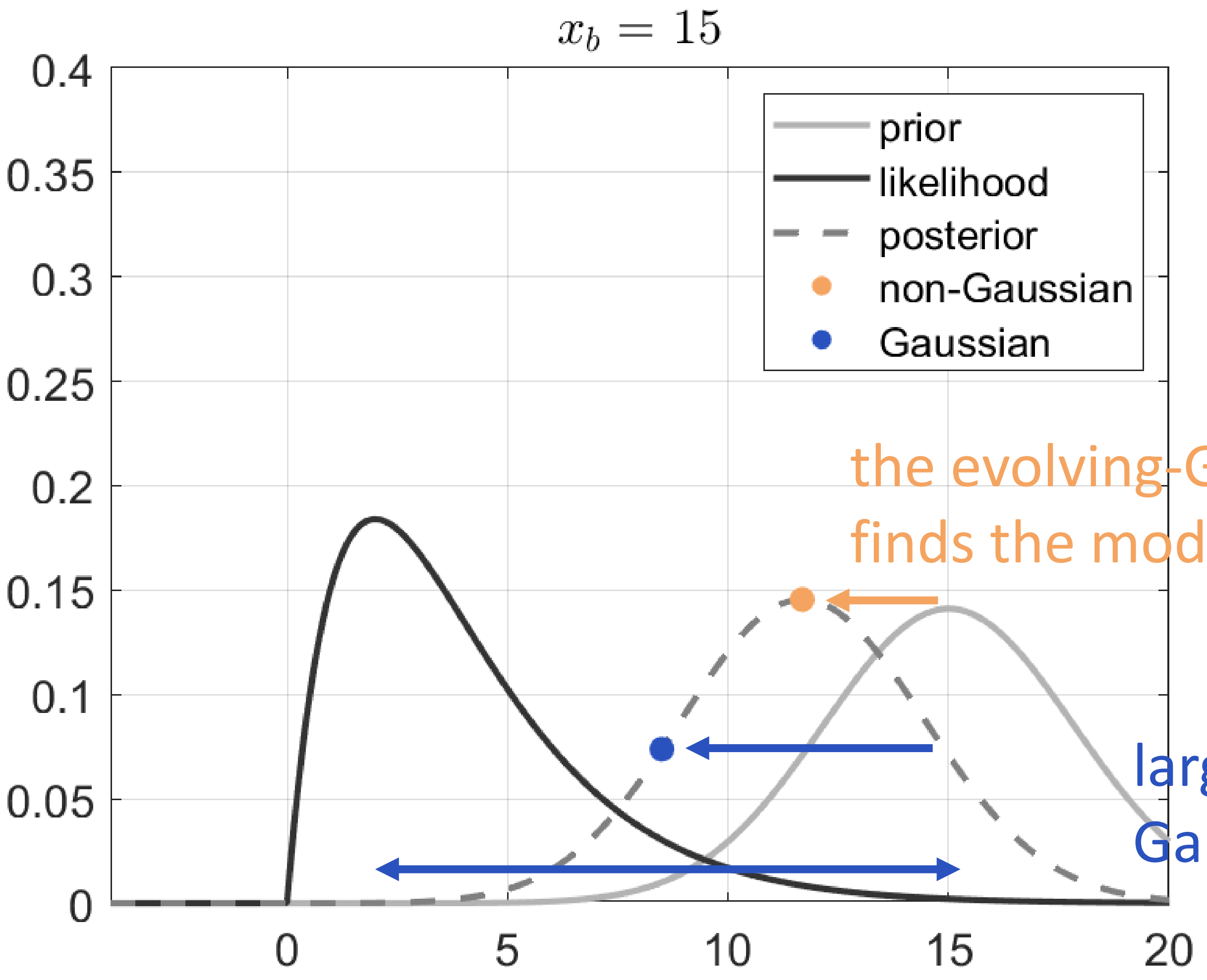


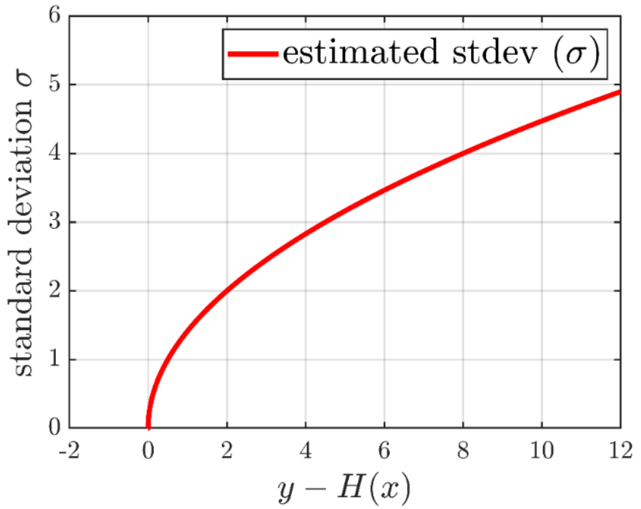
Gaussian experiment:

Set mean = mode of non-Gaussian pdf

Stdev = stdev of the non-Gaussian pdf

What can we learn from the idealized experiment?





the evolving-Gaussian method accurately finds the mode of the posterior pdf

large discrepancy between O and B

Gaussian solution: too large increment

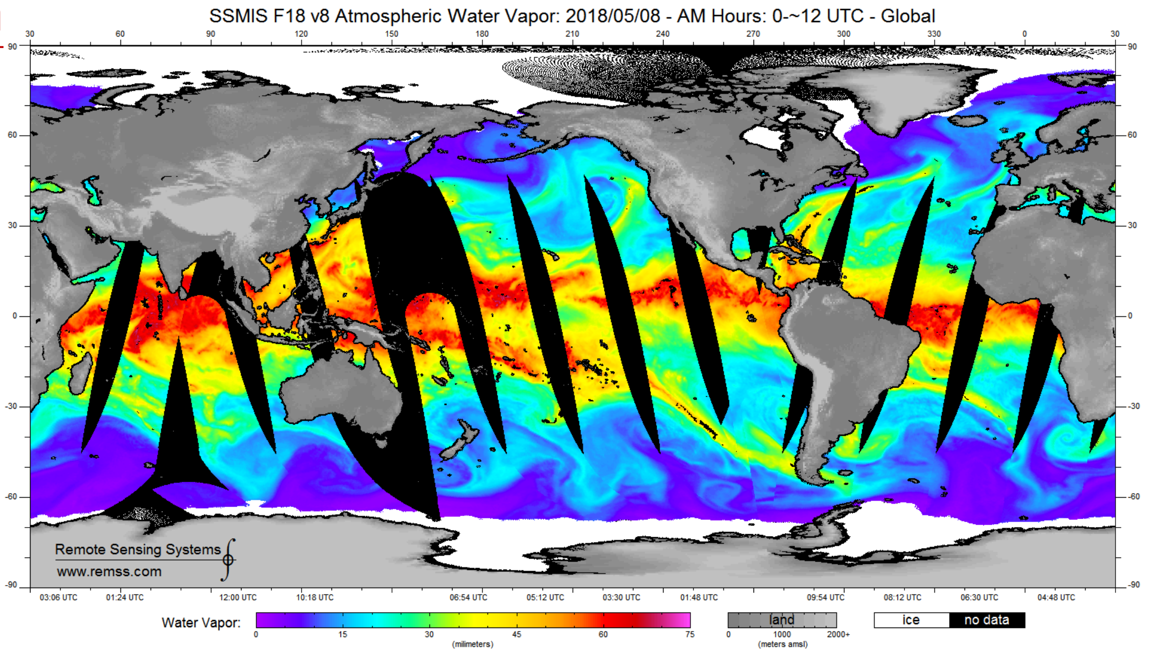
An implementation for the all-sky microwave radiance in IFS-ECMWF

- Selected SSMIS channels: ch12 (19H), ch13 (19V), ch14 (22V), ch16 (37V), ch17 (91V)
- The assumptions for the assimilation of the all-sky microwave radiances:
  - A ”symmetric cloud”-dependent **Gaussian** observation error model

non-Gaussian

$$\overline{C_{37}} = \frac{C^{o}_{37} + C^{b}_{37}}{2}$$

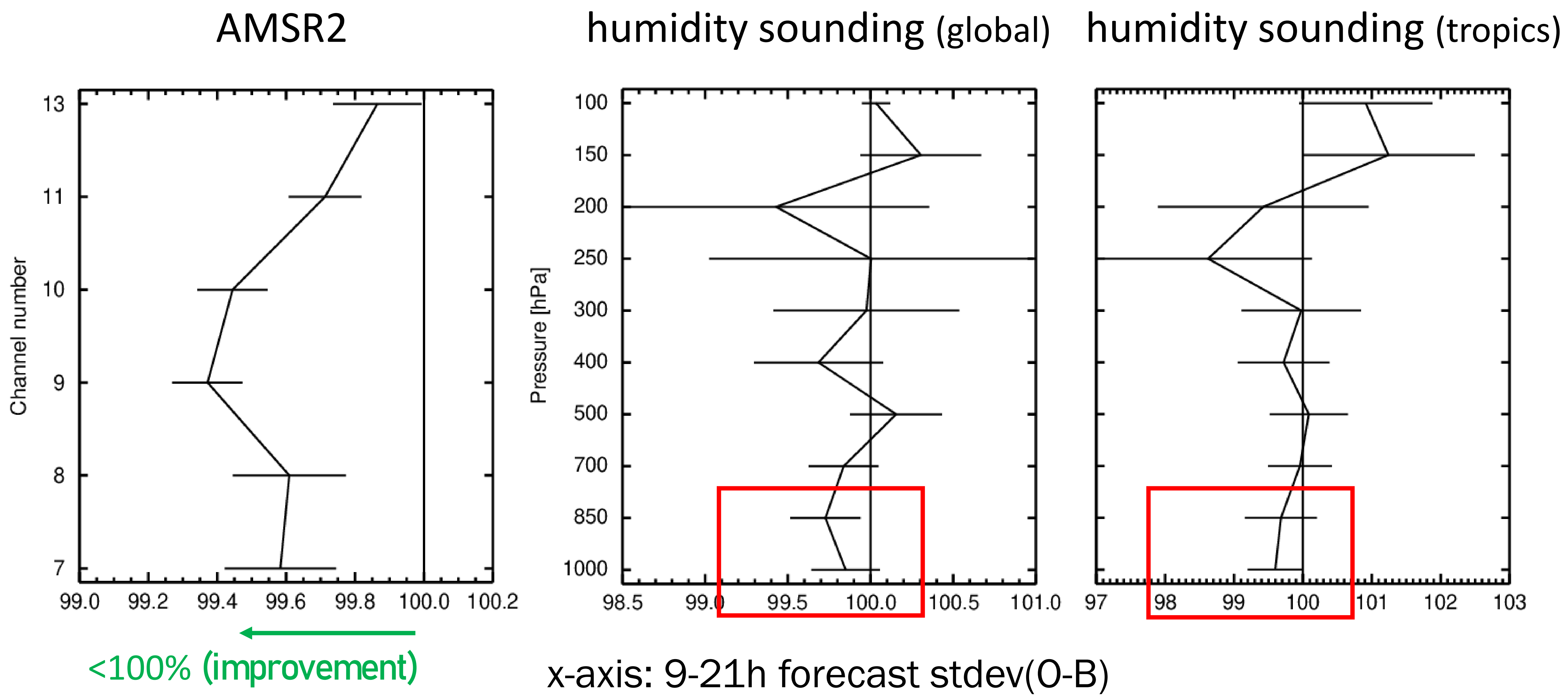
Geer and Bauer (2011)



  - The O-B (innovation) pdf is assumed to be the observation error pdf*

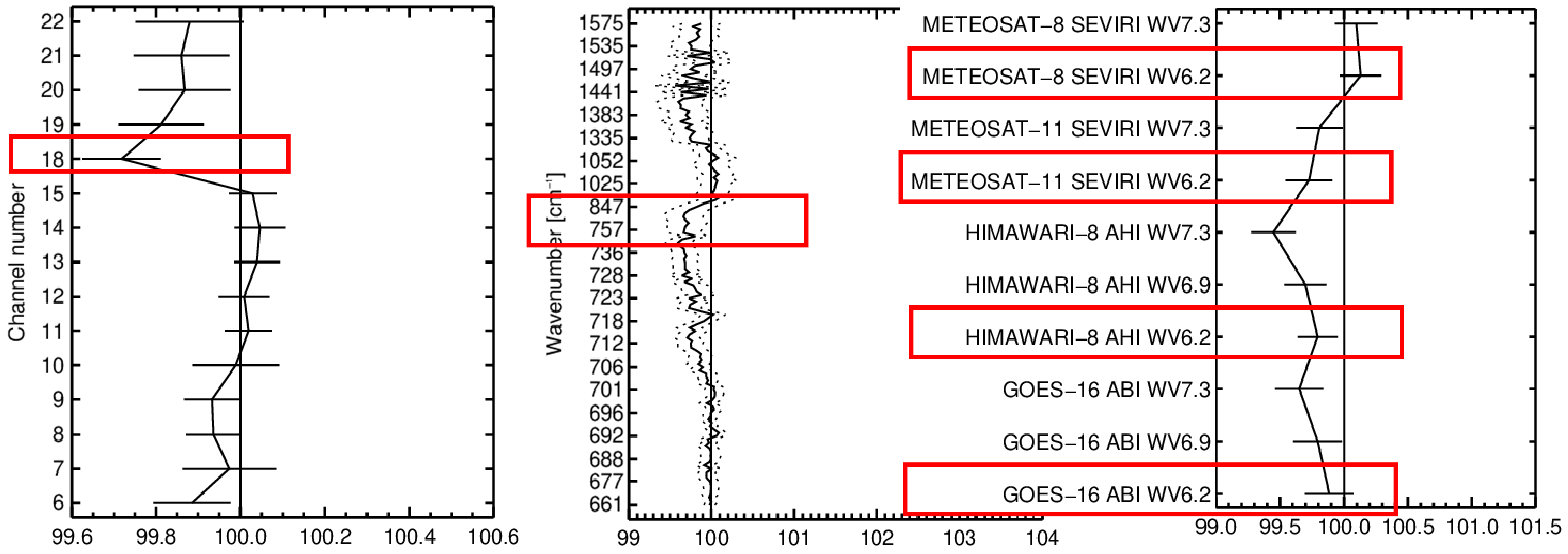
*(*subtracting the background error from O-B does not improve the results)*

Preliminary results suggest lower tropospheric water vapor, cloud and precipitation are improved, especially in the tropics.



Preliminary results suggest lower tropospheric water vapor, cloud and precipitation are improved, especially in the tropics.

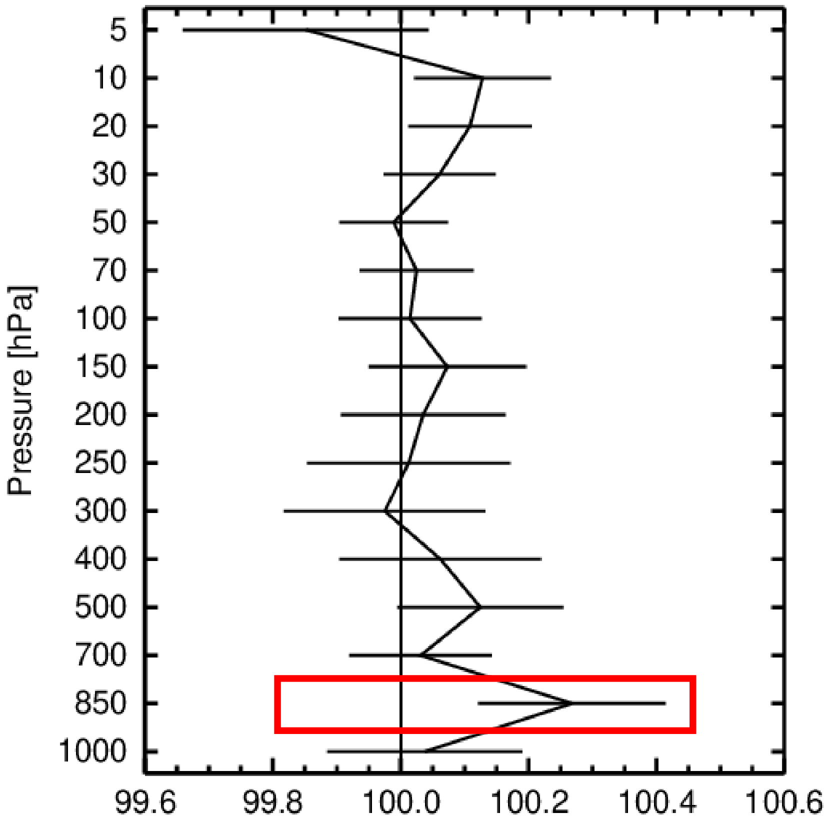
ATMS | CRIS | Geostationary satellites



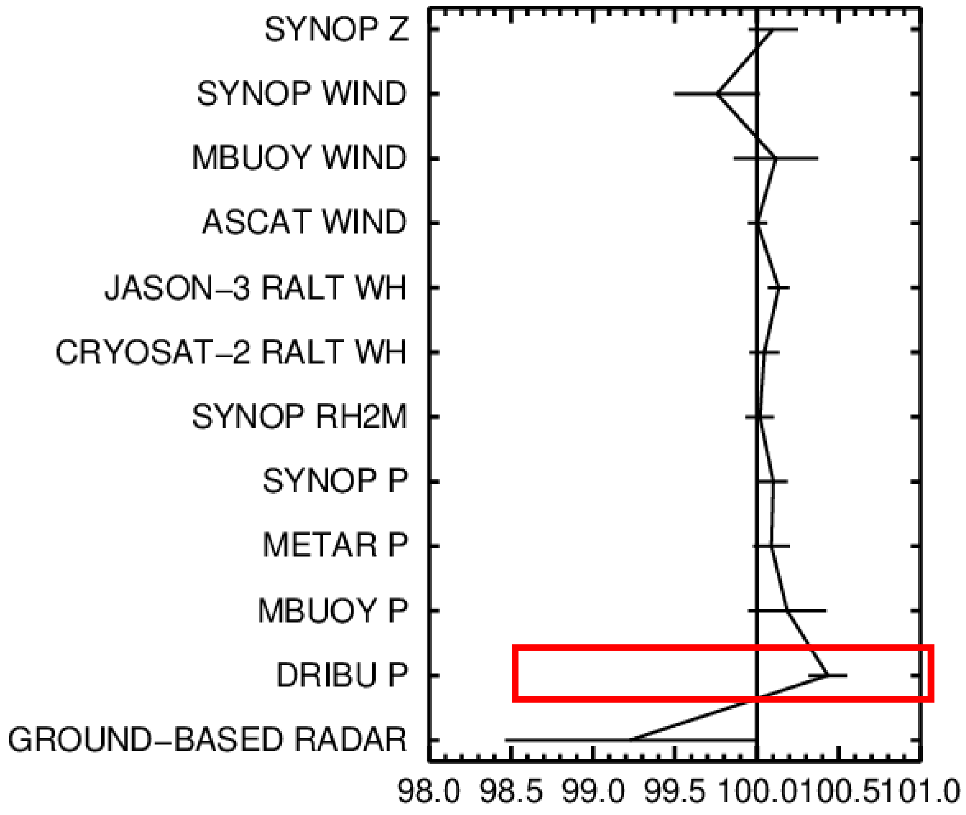
x-axis: 9-21h forecast stdev(O-B)

There are degradations in temperature and surface pressure

temperature sounding



surface observations



x-axis: 9-21h forecast stdev(O-B)

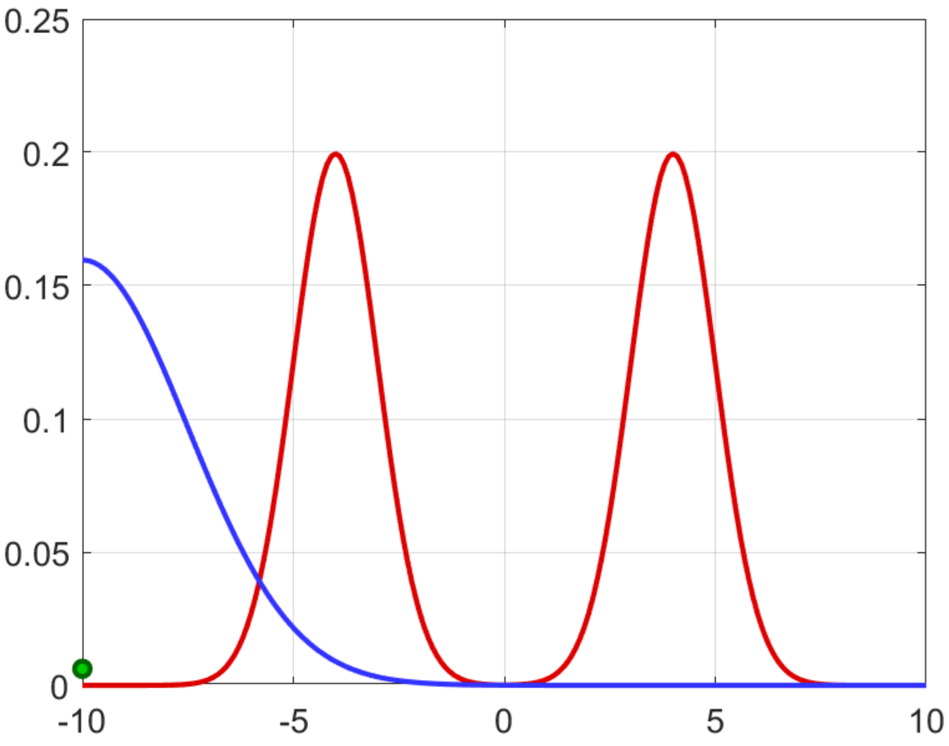
Deconvolution-based Observation Error Estimation (DOEE)

Hu, van Leeuwen and Geer (under review)

- Currently, the innovation pdf is assumed to be the observation error pdf (for the all-sky microwave radiances).

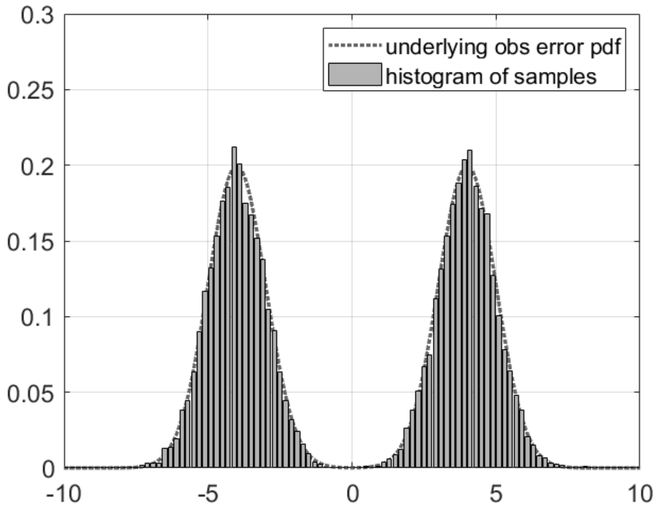
How do we identify the pdf of an arbitrary non-Gaussian observation error?

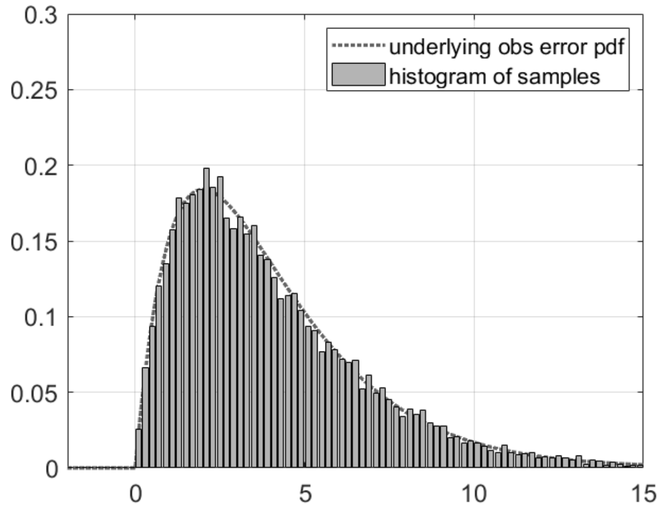
- DOEE can be used to find a *non-parametric* observation error pdf:



$$D = Y - H(X) = \varepsilon^o - \left[H(X) - H\left(X^{truth}\right)\right]$$

$$f_D = \left(f_{\varepsilon^o} * f_{H(X)-H\left(X^{truth}\right)}\right)$$





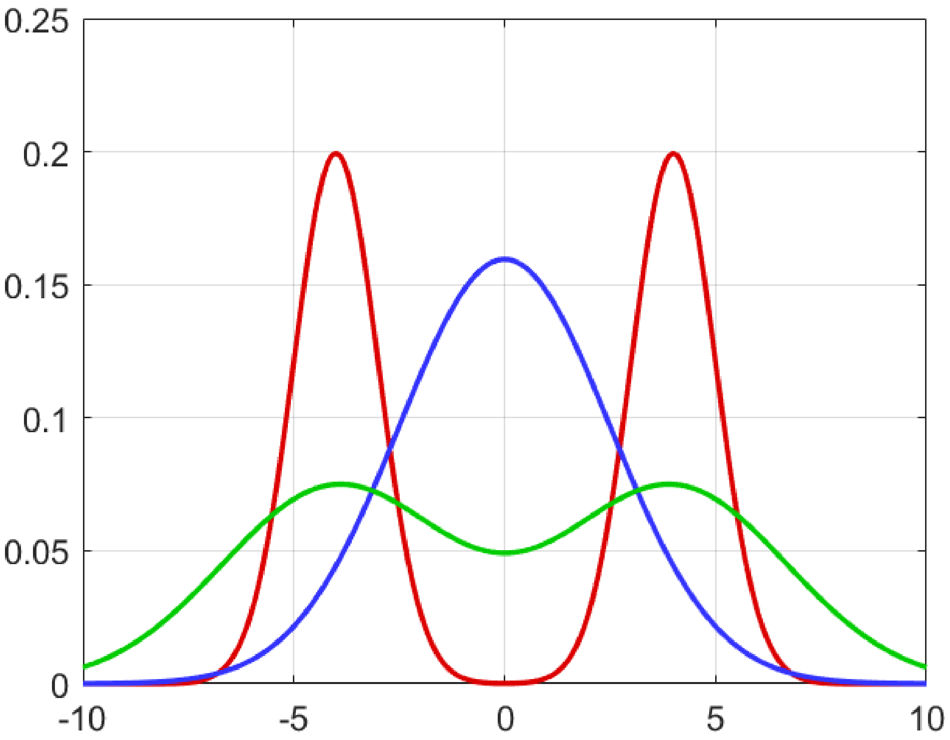
Deconvolution-based Observation Error Estimation (DOEE)

Hu, van Leeuwen and Geer (under review)

- Currently, the innovation pdf is assumed to be the observation error pdf (for the all-sky microwave radiances).

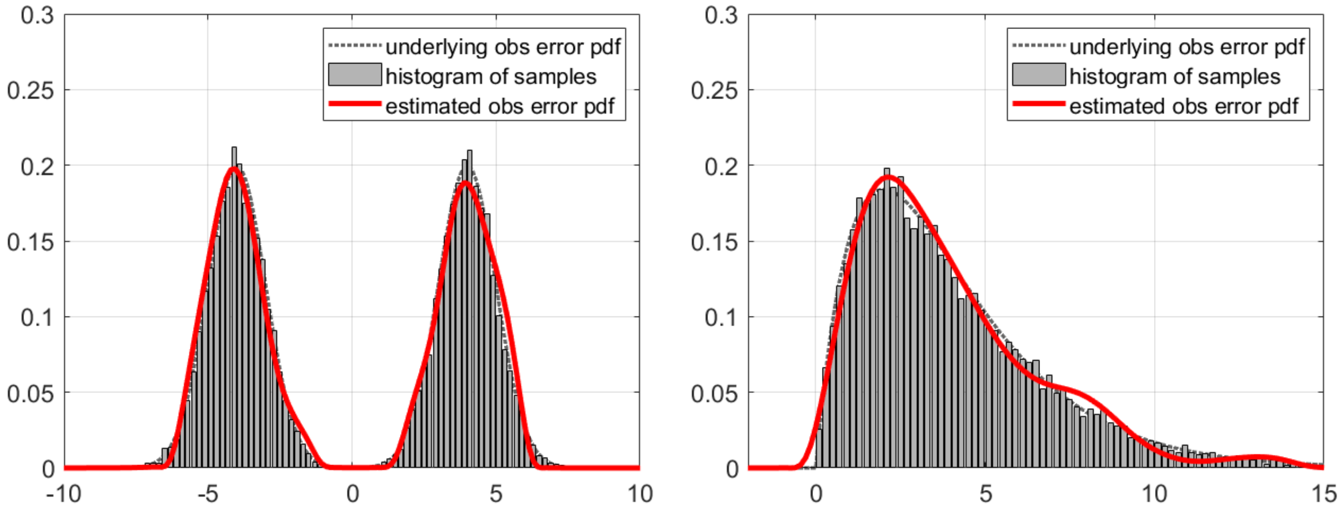
How do we identify the pdf of an arbitrary non-Gaussian observation error?

- DOEE can be used to find a *non-parametric* observation error pdf:



$$D = Y - H(X) = \varepsilon^o - \left[H(X) - H\left(X^{truth}\right)\right]$$

$$f_D = \left(f_{\varepsilon^o} * f_{H(X)-H\left(X^{truth}\right)}\right)$$



Future work and possible applications

- Incorporate non-Gaussian observation error for more observation types.
- The evolving-Gaussian method for the inner-loop using approximated $H(x)$
- **Possible applications:**
  - ML based observation operator ("surrogate observation operator")
  - The errors from the ML observation operator can be difficult to estimate
  - DOEE can be used to estimate the errors from the ML observation operator
  - DOEE + evolving-Gaussian method can be applied in 4D-Var.

Conclusions

- We propose a new way, the evolving-Gaussian method, to incorporate general non-parametric observation error pdfs into the variational methods.

- The evolving-Gaussian method does not require revising the cost-function, and therefore it is easy to implement in a full-scale weather prediction system without adding much computational cost.

- We have implemented the evolving-Gaussian method for the all-sky microwave radiances in IFS-ECMWF, resulting in improved short-term forecasts of the lower-tropospheric cloud, water vapor and precipitations, especially in the tropics.

- The evolving-Gaussian method, together with the non-Gaussian observation error estimation technique DOEE, can potentially address the issues regarding the complicated errors in the observation operators.

Contact Info

- Feel free to send me an email for any questions, comments or more info about DOEE/the evolving-Gaussian method!

- Emails

  - Chih-Chi.Hu@noaa.gov
  - ch0683@princeton.edu