NON-GAUSSIAN DATA ASSIMILATION METHODOLOGIES

STEVEN J. FLETCHER

With thanks to Milija and Dusanka Zupanski, Andy Jones, Laura Fowler, Manajit Sengupta and Thomas Vonder Haar

DoD Center for Geosciences/Atmospheric Research (CG/AR)R

COOPERATIVE INSTITUTE FOR RESEARCH IN THE ATMOSPHERE

COLORADO STATE UNIVERSITY

PRESENTATION OUTLINE

- ❑ MOTIVATION FOR THE WORK
- ❑ REAL LIFE EXAMPLES OF NON-GAUSSIAN VARIABLES
- ❑ MATHEMATICAL ILLUSTRATIONS OF THE DRAWBACKS OF CURRENT ATTROACHES
- ❑ 3-D HYBRID LOGNORMAL – NORMAL DATA ASSIMILATION APPLICATIONS
- ❑ 4-D HYBRID LOGNORMAL – NORMAL DATA ASSIMILATION

CG/AR

MOTIVATION

The main assumption made in variational and ensemble data assimilation is that the state variables and observations are Gaussian distributed

Note: The difference between two Gaussian variables is also a Gaussian variable.

Is this true for all state variables?

Is this true for all observations of the atmosphere?

State Variables

Miles et al 2000, lists cloud variables that are not Gaussian

Dee and Da Silva, 2003, humidity

Most positive definite variables!!

Observations

Stephens et al 2002, many of the CLOUDSAT observations, non-Gaussian:
i.e. Optical depth, Infra red flux differences

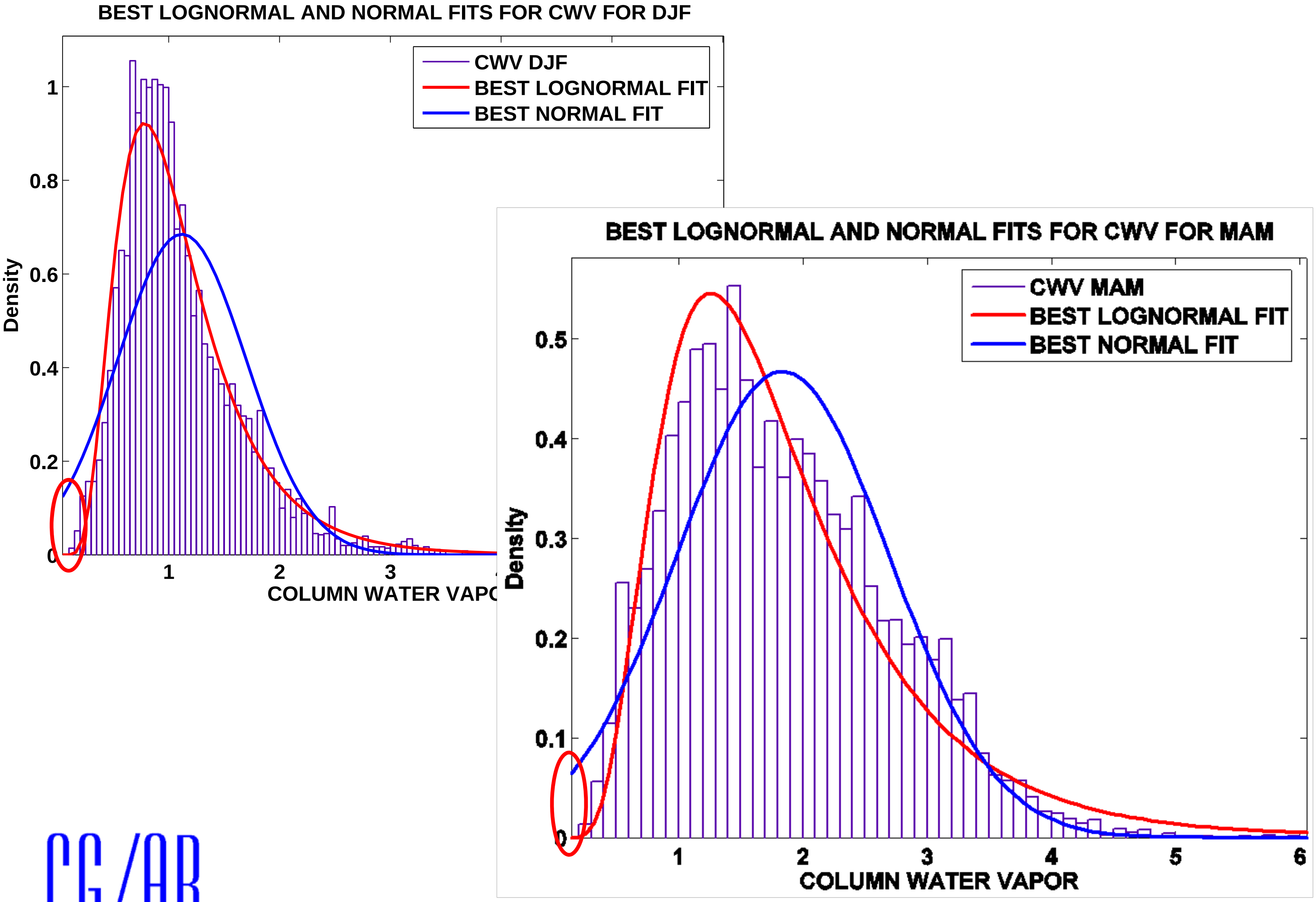
Sengupta et al 2004: Cloud base height, Liquid water path.

Deblonde and English 2003, Boukabara *et al.* 2007: Humidity retrievals.

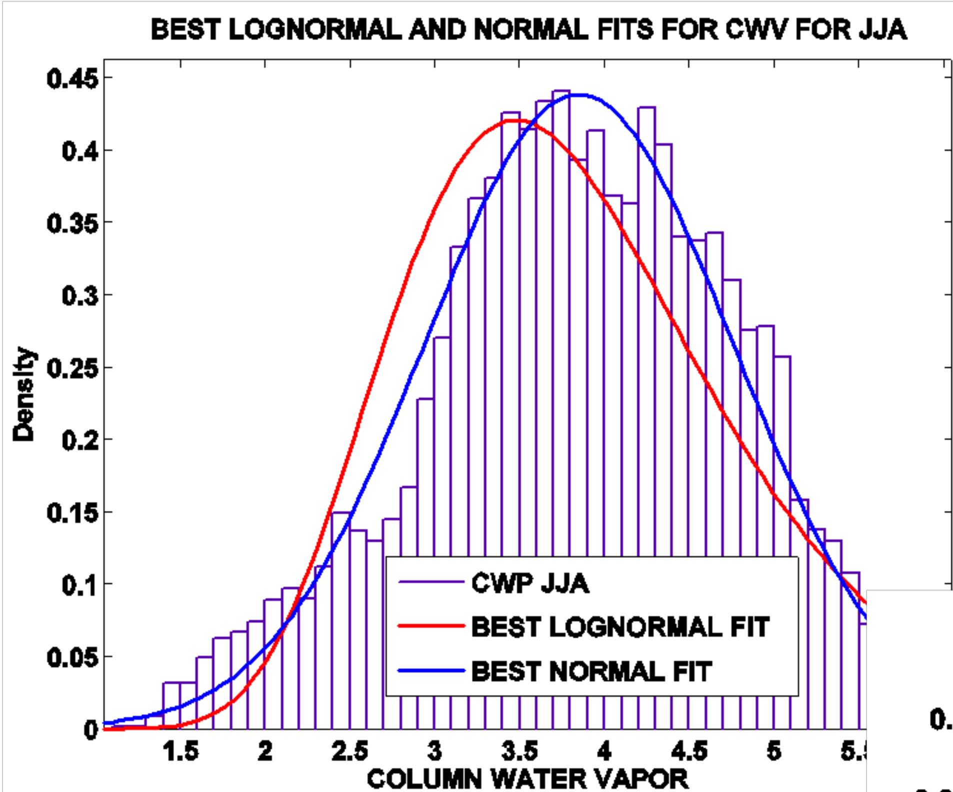
REAL LIFE EXAMPLES

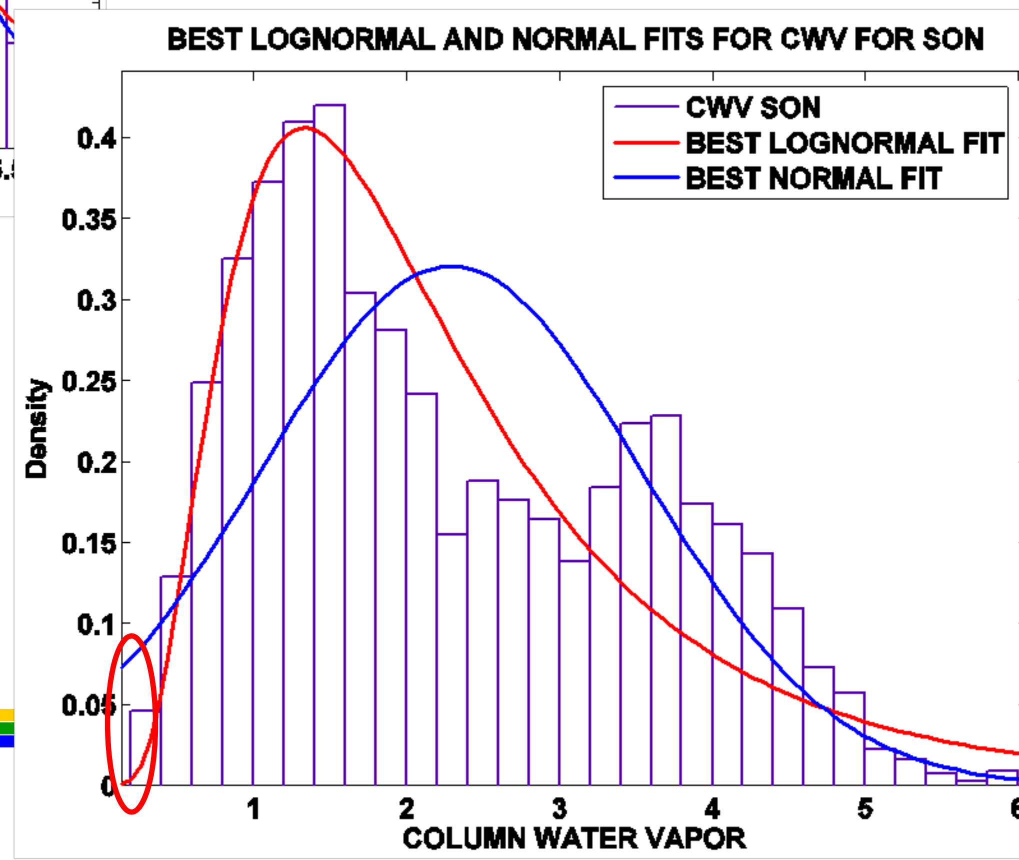
THIS DATA IS COLUMN WATER VAPOUR CLIMATOLOGIES FROM THE OKLAHOMA ARM SGP SITE FROM 1997 – 2000 WHERE THE DATA ARE OBSERVED FOR DAYS WITH BOUNDARY LAYER CLOUDS.

THE DATA HAS BEEN BROKEN DOWN BY SEASON AS WELL AS FOR THE WHOLE FOUR YEARS. THE DATA WAS COLLECTED FROM A MICROWAVE RADIOMETER

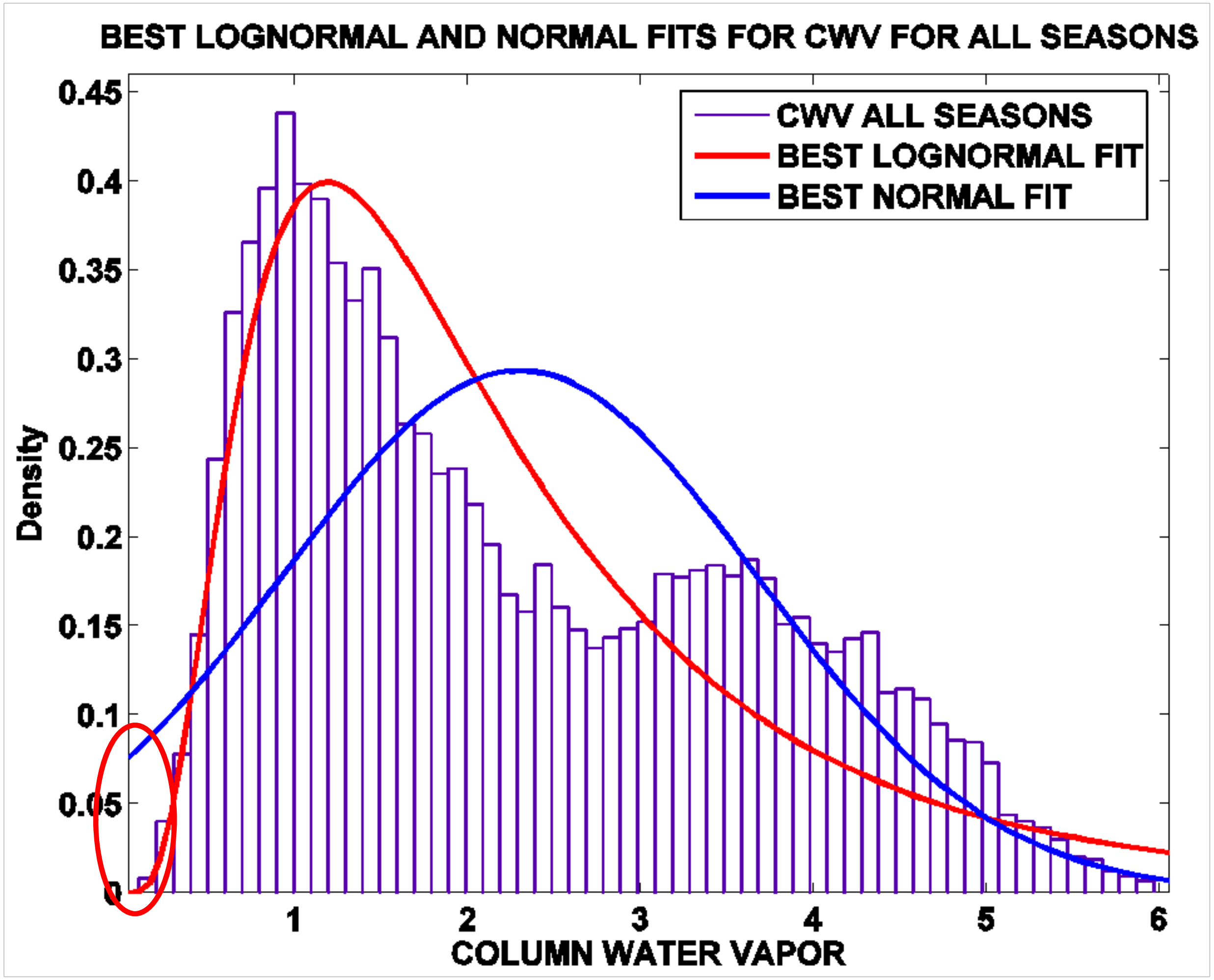


BEST LOGNORMAL AND NORMAL FITS FOR CWV FOR DJF
CWV DJF
BEST LOGNORMAL FIT
BEST NORMAL FIT
Density
COLUMN WATER VAPO
0
0.2
0.4
0.6
0.8
1
1
2
3
BEST LOGNORMAL AND NORMAL FITS FOR CWV FOR MAM
CWV MAM
BEST LOGNORMAL FIT
BEST NORMAL FIT
Density
COLUMN WATER VAPOR
0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
5
6



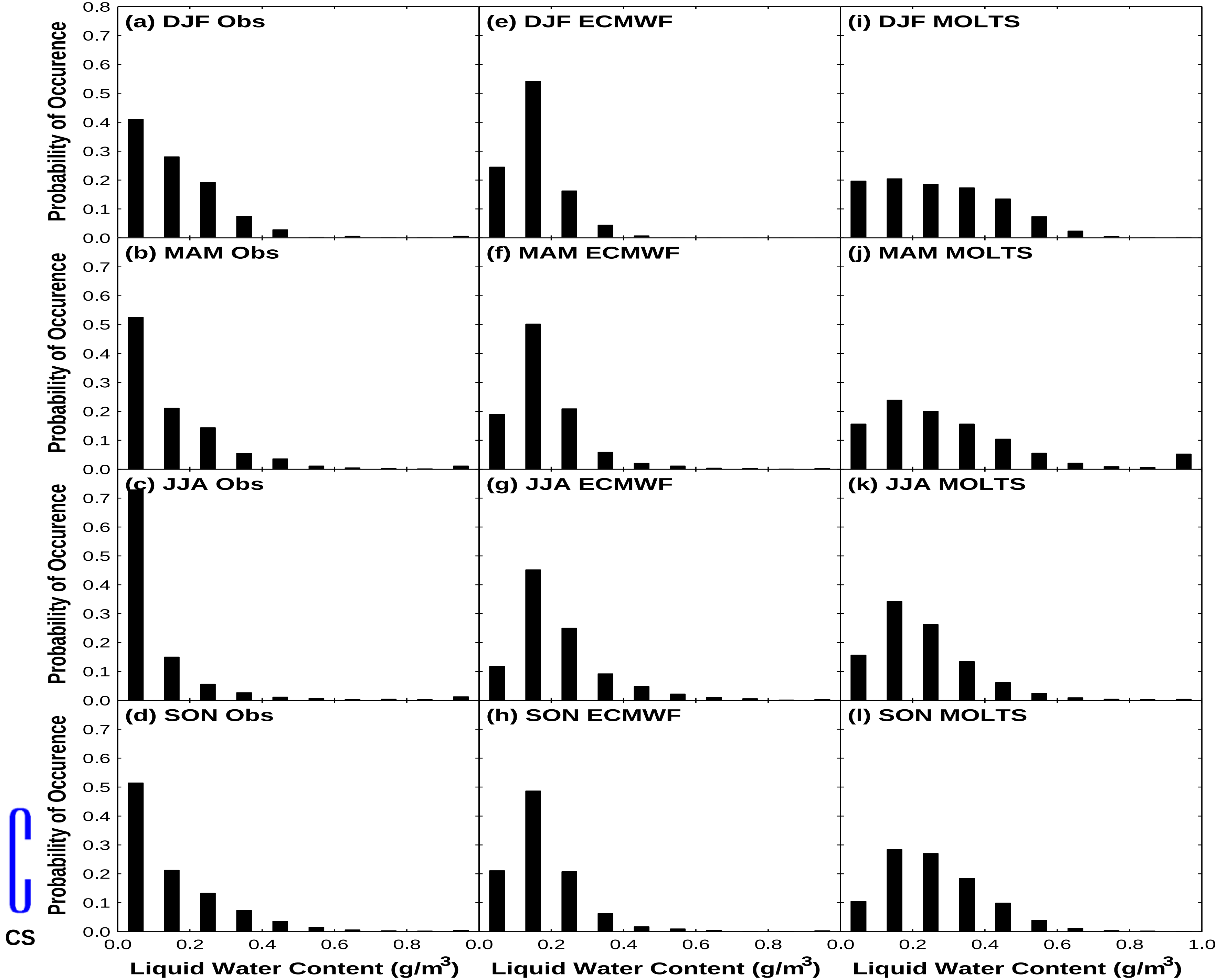


CSU/CIRA Dr. Steven J. Fletcher


BEST LOGNORMAL AND NORMAL FITS FOR CWV FOR ALL SEASONS
CWV ALL SEASONS
BEST LOGNORMAL FIT
BEST NORMAL FIT
Density
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1
2
3
4
5
6
COLUMN WATER VAPOR

<u>ANOTHER REAL LIFE EXAMPLE FROM SENGUPTA *ET AL.* (2004)</u>

Here we using liquid water path where there the plots consist of model outputs form ECMWF and the MOLTS models. This for the same data sets just shown.



(a) DJF Obs
(e) DJF ECMWF
(i) DJF MOLTS
(b) MAM Obs
(f) MAM ECMWF
(j) MAM MOLTS
(c) JJA Obs
(g) JJA ECMWF
(k) JJA MOLTS
(d) SON Obs
(h) SON ECMWF
(l) SON MOLTS
Probability of Occurence
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.0
0.2
0.4
0.6
0.8
1.0
Liquid Water Content (g/m^3)
CS

Current Techniques used with non-Gaussian Variables

1) Transform by taking the LOGARITHM of the original state variable. This then makes the new variable ALMOST GAUSSIAN. Minimize the cost function with respect to this variable, TRANSFORM BACK and initialize with this state. STATE FOUND IS A NON-UNIQUE MEDIAN OF THE ORIGINAL VARIABLE, Fletcher and Zupanski 2006a, 2007.

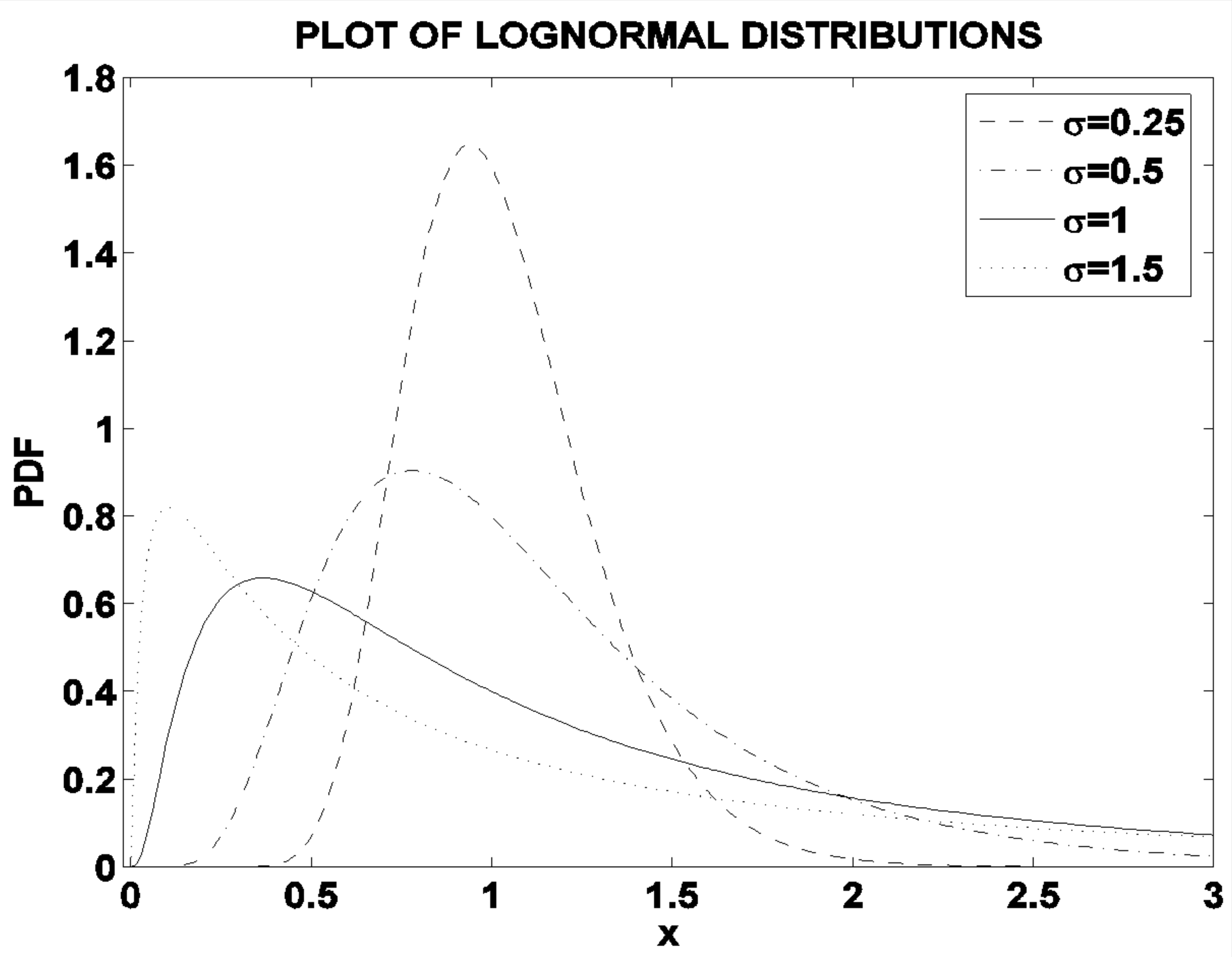
2) Assumed Gaussian assumption and BIAS CORRECT.

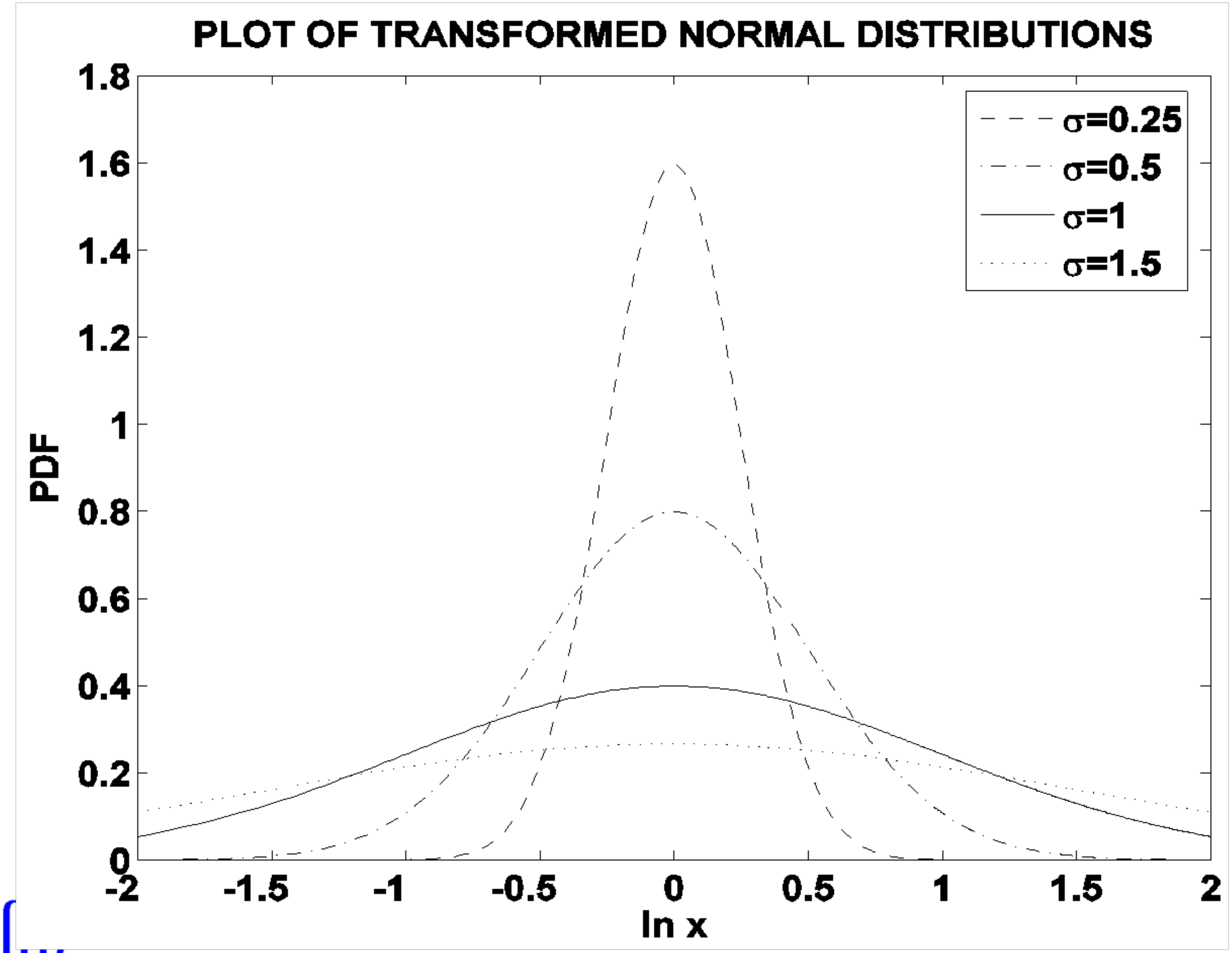
3) Another Technique, employed in other fields, is to use a GAUSSIAN SUM FILTER.

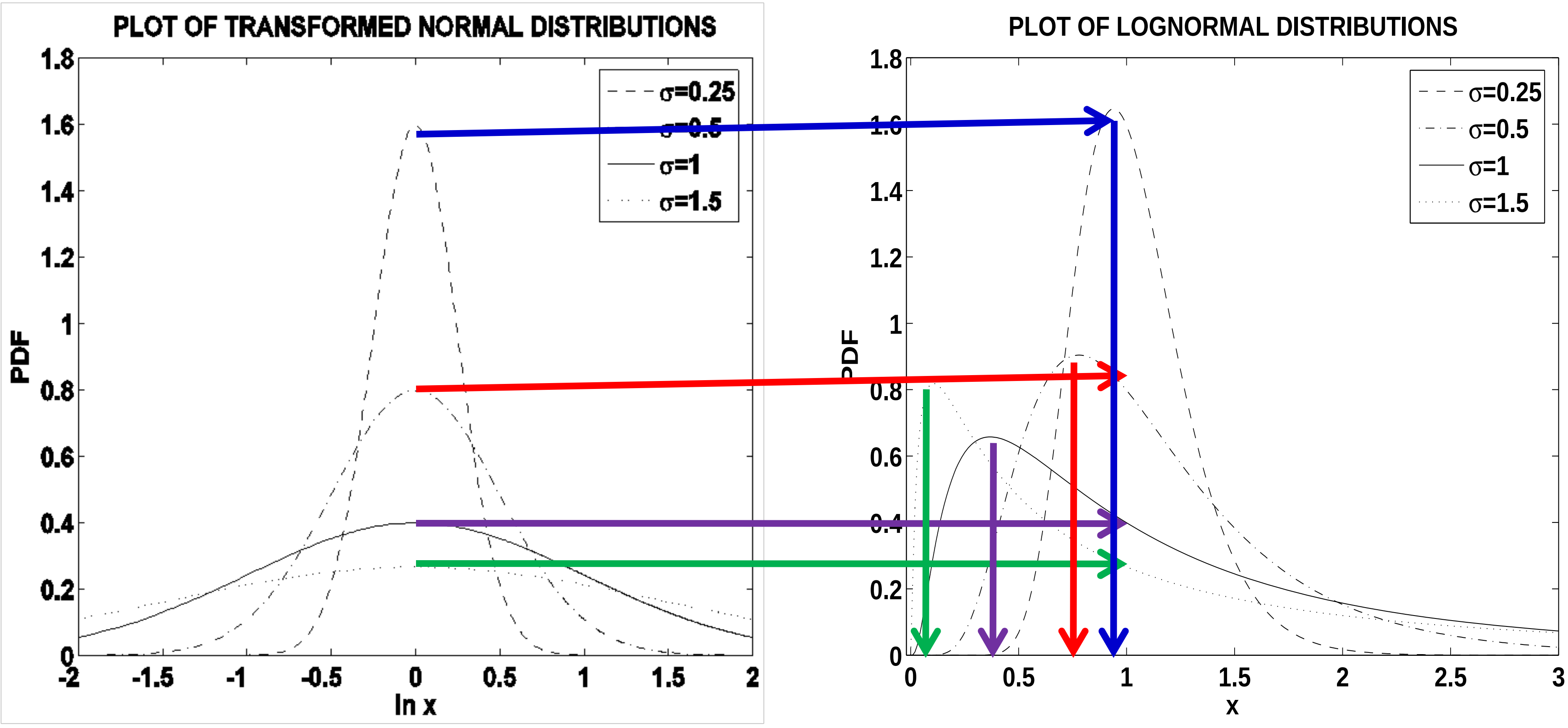
4) A recently suggested technique for meteorological applications is a MAXIMUM ENTROPHY FILTER.

5) Using a Markov-Chain Monte-Carlo approaches (Posselt *et al.* 2008)

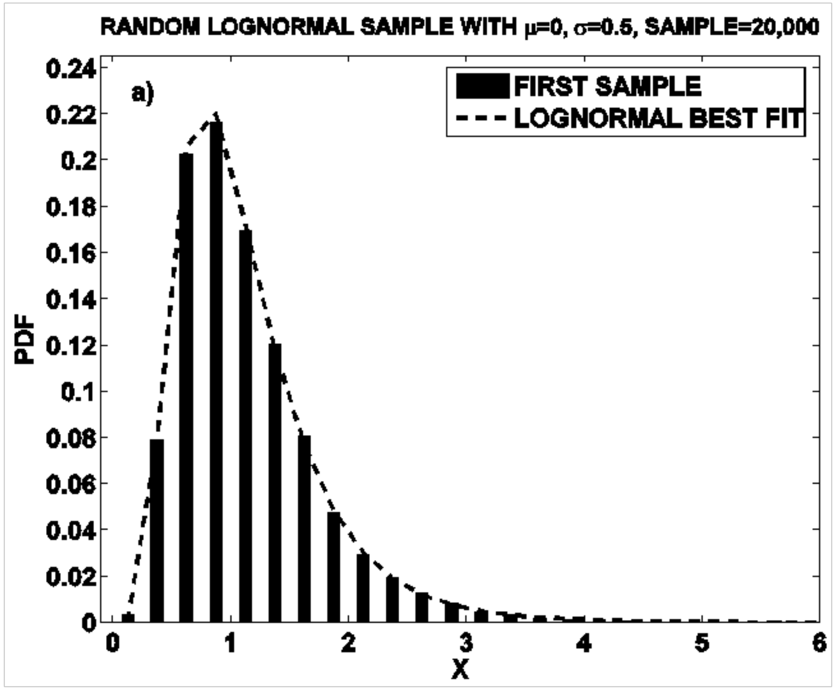
CG/AR

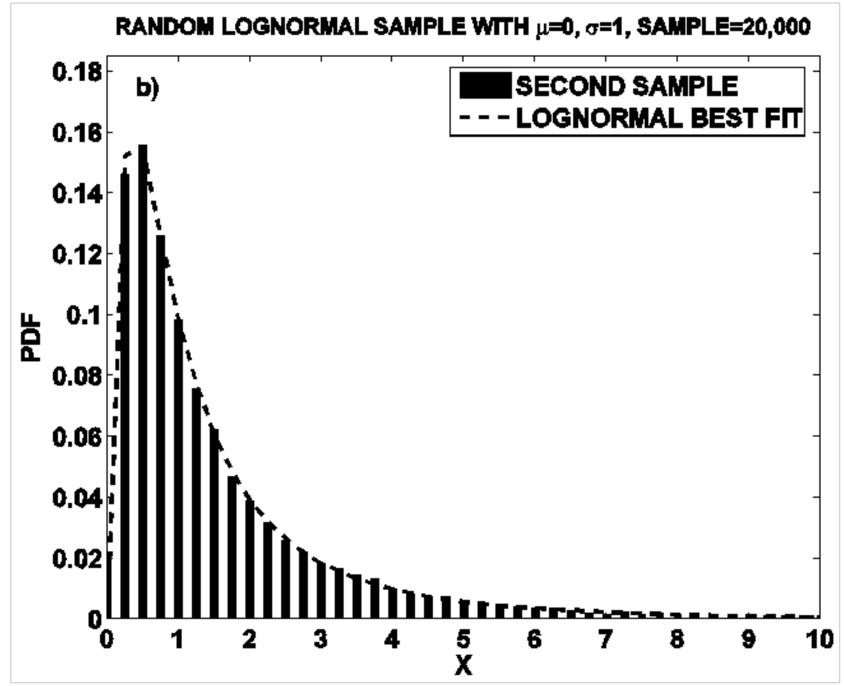

PLOT OF LOGNORMAL DISTRIBUTIONS
σ=0.25
σ=0.5
σ=1
σ=1.5
PDF
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
0
0.5
1
1.5
2
2.5
3


PLOT OF TRANSFORMED NORMAL DISTRIBUTIONS
σ=0.25
σ=0.5
σ=1
σ=1.5
PDF
ln x
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2

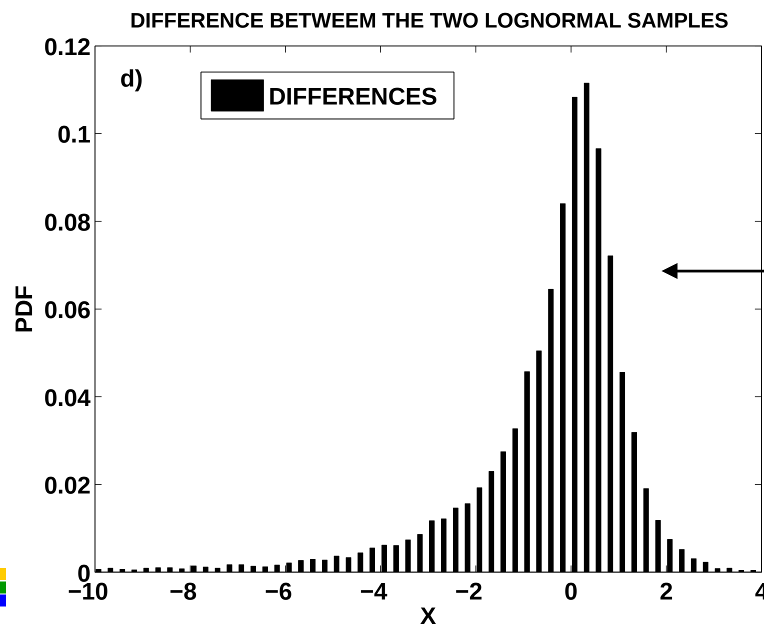


All skewness information is lost

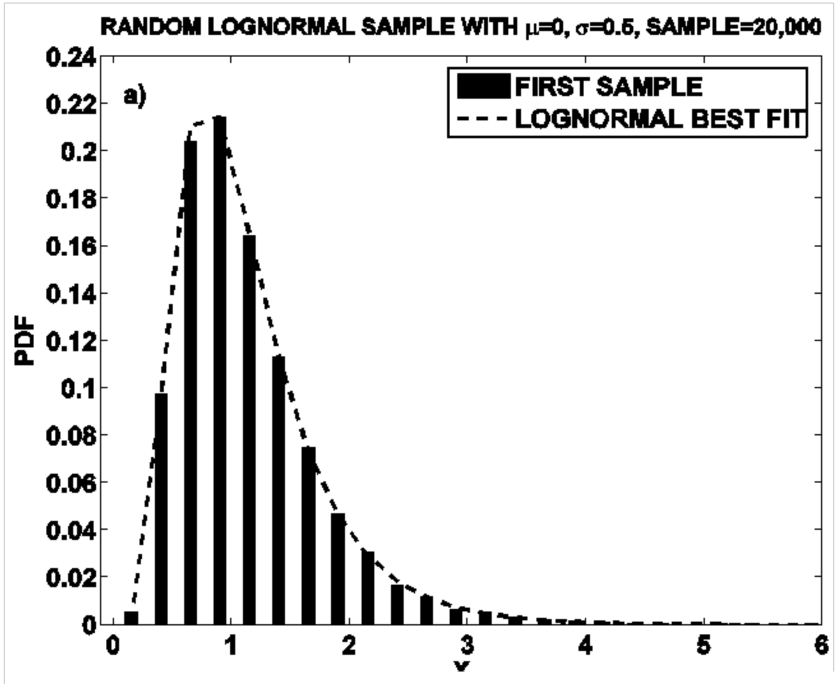


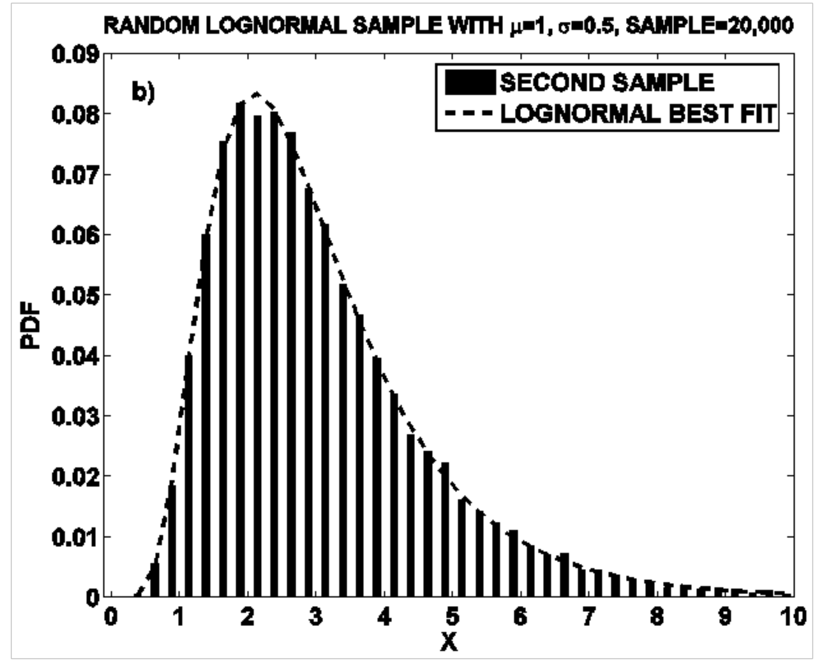


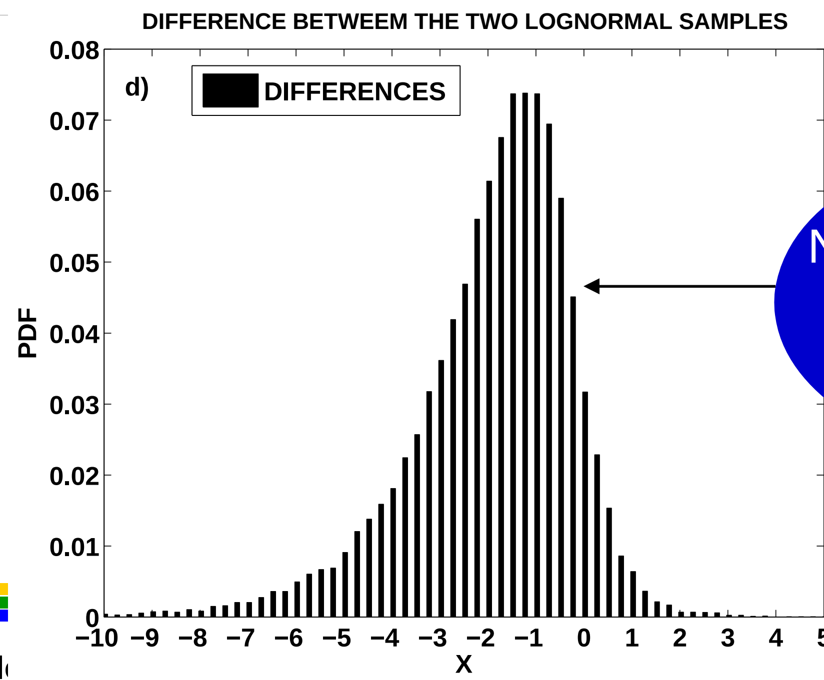
ASSUMED GAUSSIAN APPROACH



NOTE: DIFFERNCE IS NOT A GAUSSIAN







NOTE: DIFFERNCE IS NOT A GAUSSIAN

PROBLEMS ASSOCIATED WITH CURRENT TECHNIQUES

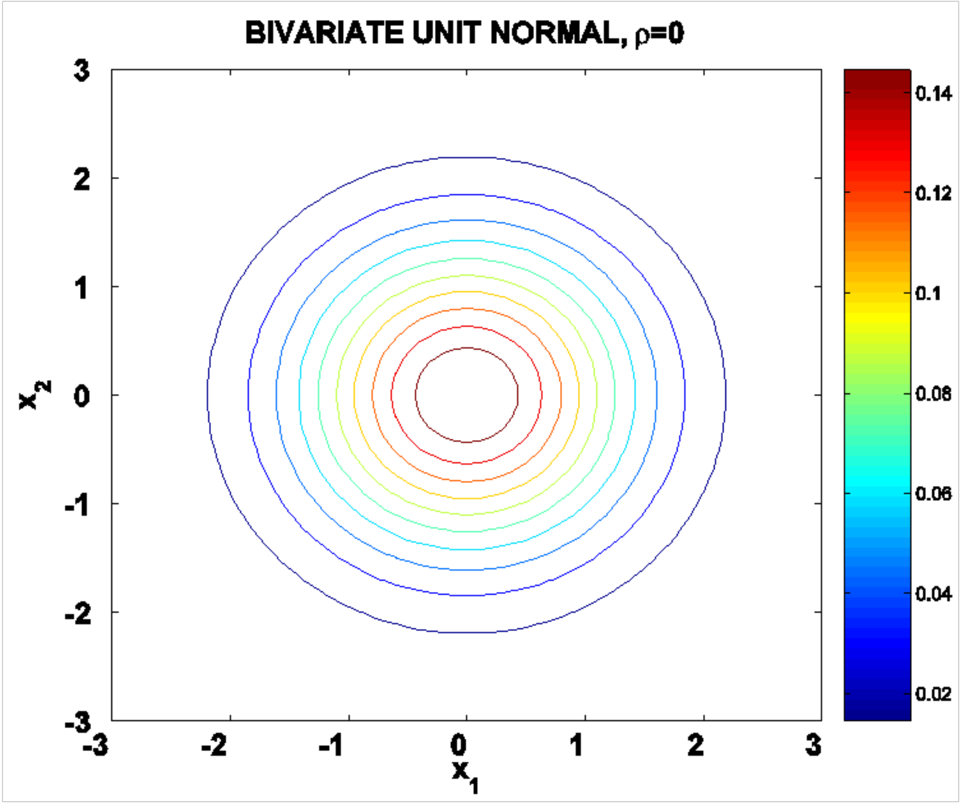
ASSUMED GAUSSIAN:

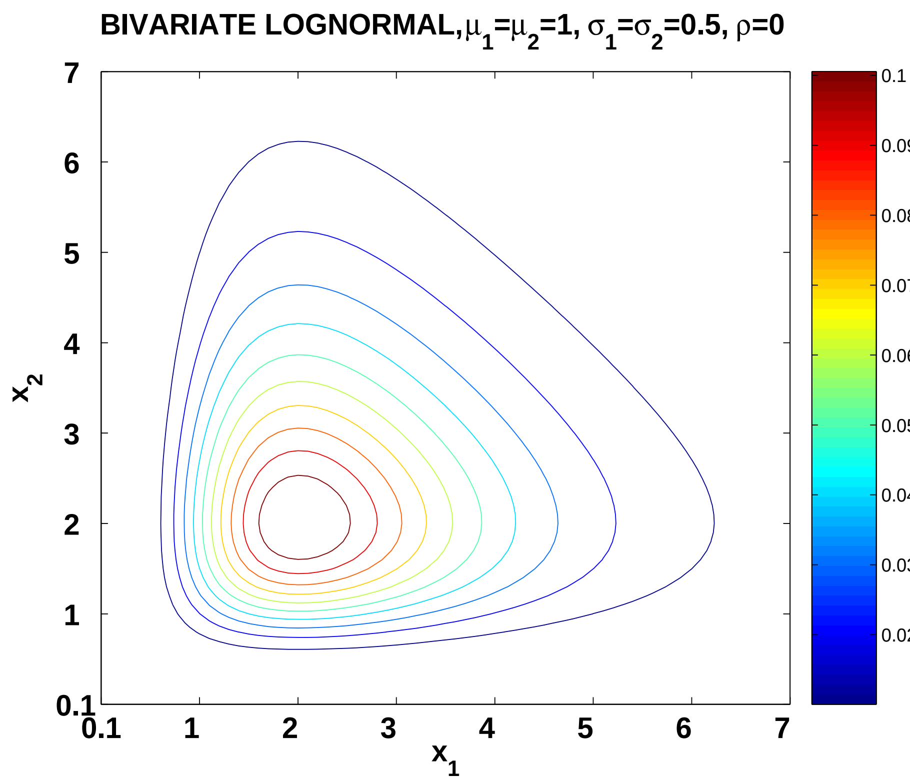
IMPACT 1: WRONG PROBABILITIES ASSIGNED TO THE OUTLIERS.

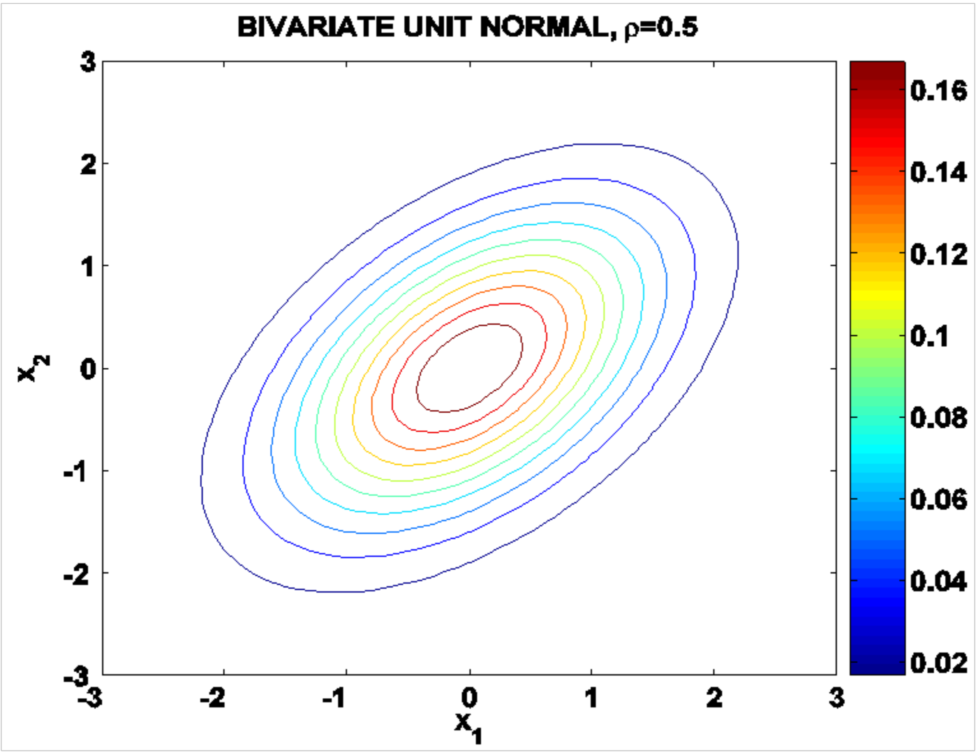
IMPACT 2: PROBABILITIES ASSIGNED TO UNPHYSICAL VALUES.

IMPACT 3: WRONG STATISTICS USED TO APPROXIMATE THE VARIABLE'S DISTRIBUTION.

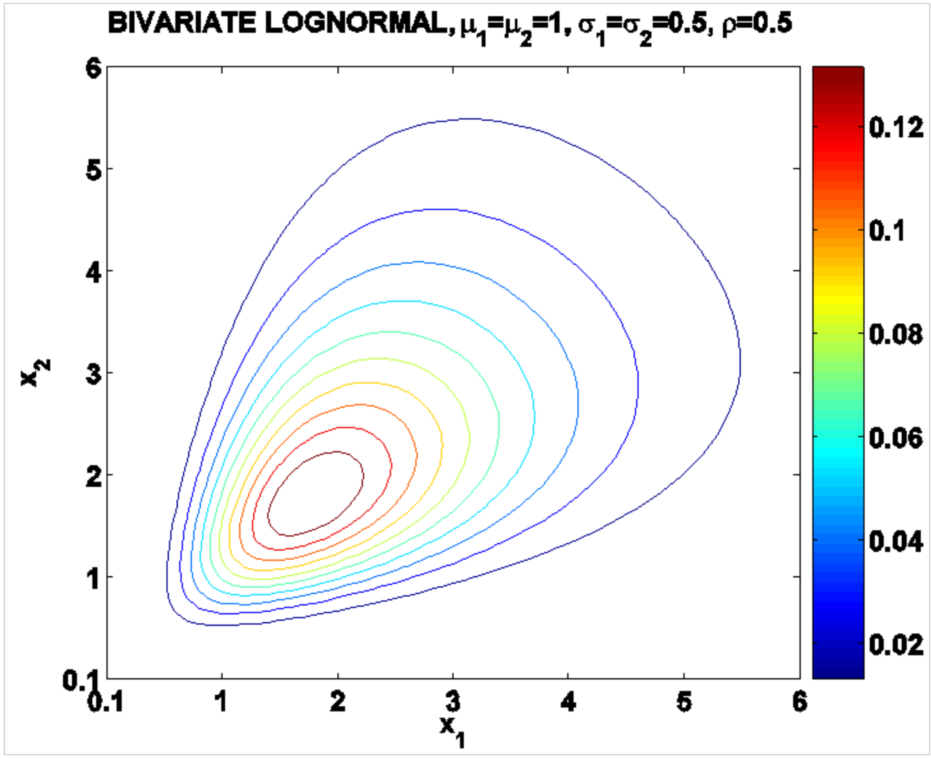
INEDPENDENT RANDOM VARIABLES







CORRELATED RANDOM VARIABLES



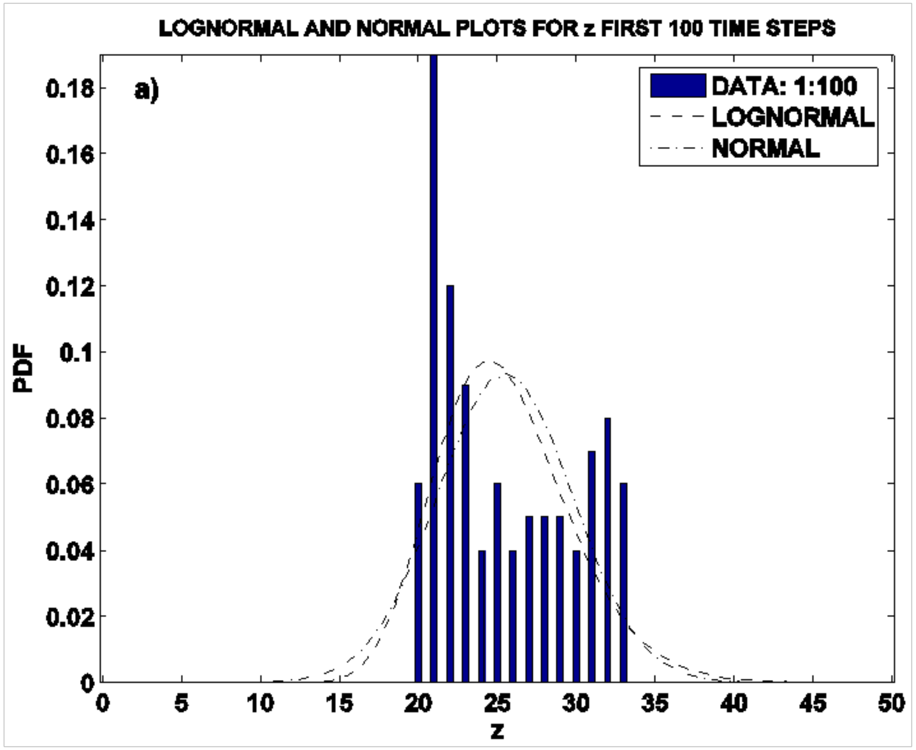
EXAMPLE WITH THE LORENZ'63 MODEL

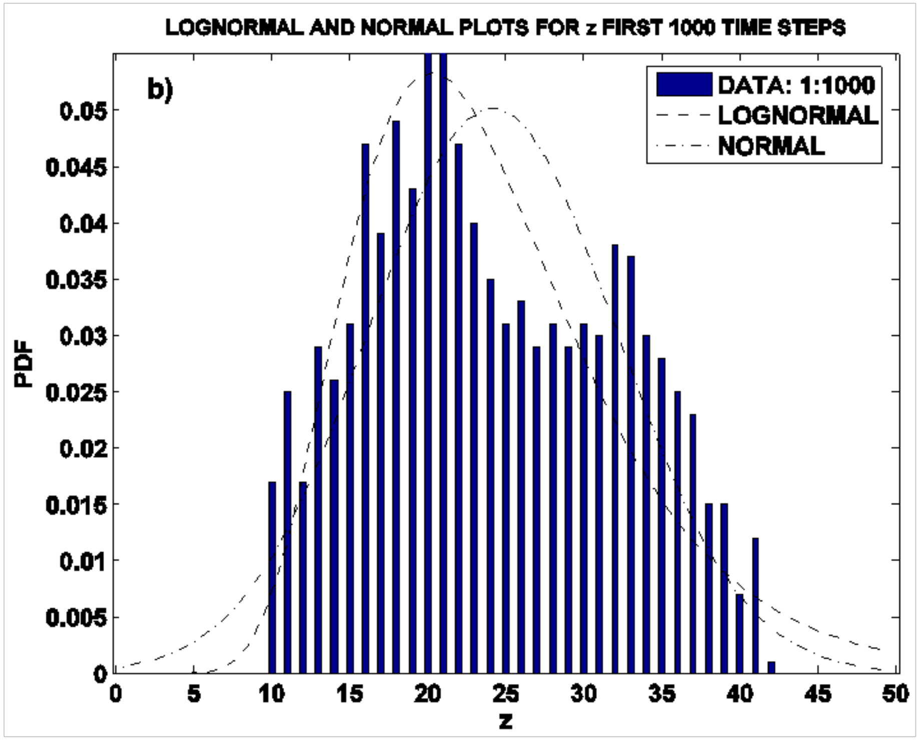
THE MODEL CONSIST OF THE COUPLED SYSTEM OF THREE NON-LINEAR PDES

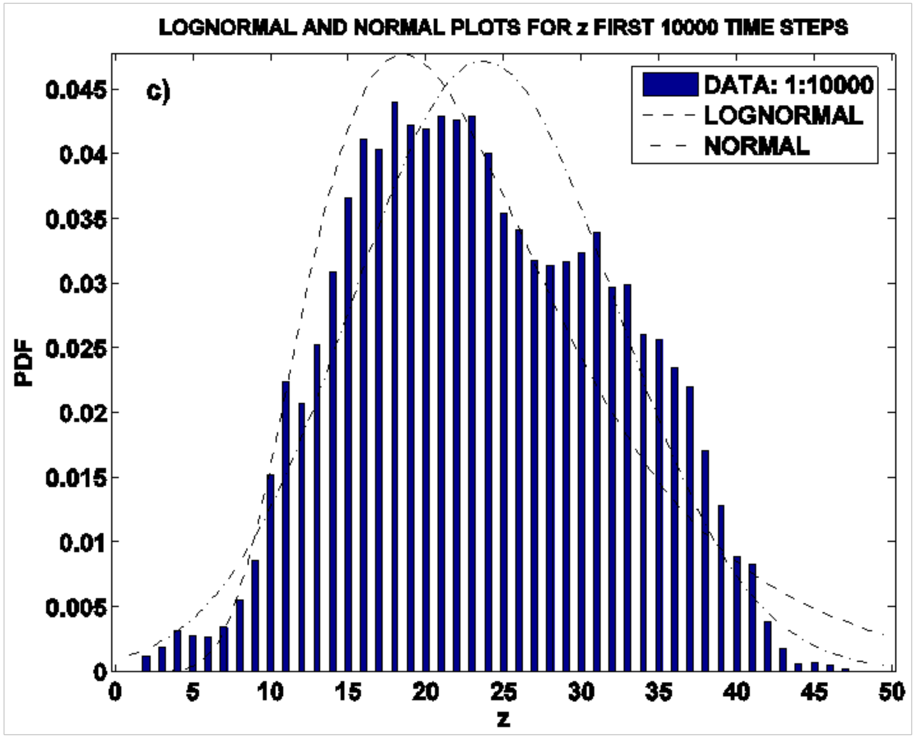
$$\dot{x} = -\sigma x + \sigma y$$
$$\dot{y} = -xz + \rho x - y$$
$$\dot{z} = xy - \beta z$$

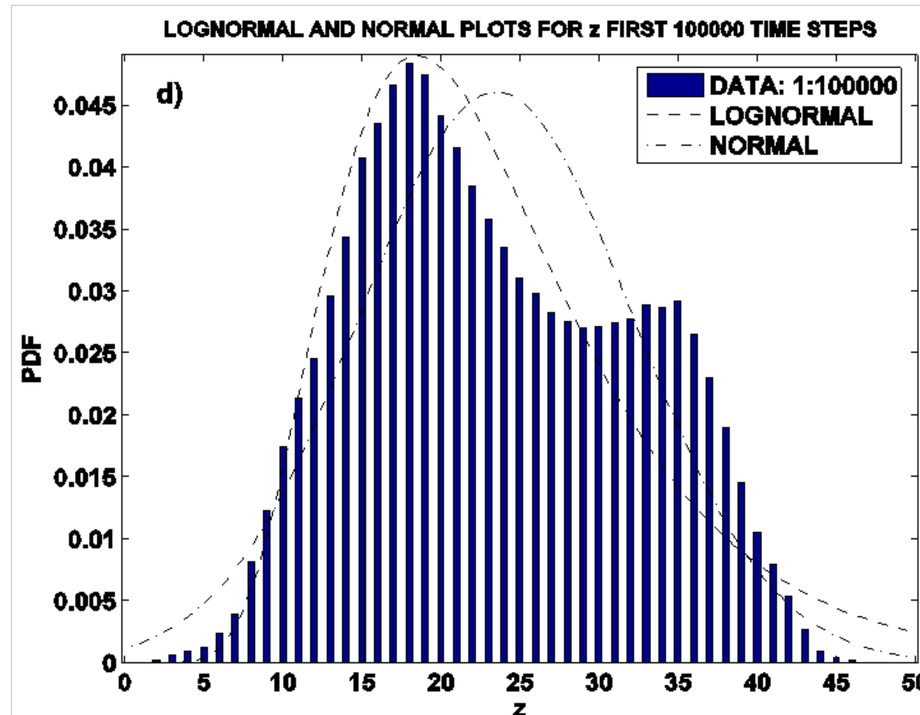
$$\beta = \frac{8}{3}, \quad \sigma = 10 \quad \text{AND} \quad \rho = 28$$

$$x_0 = -5.4458, \quad y_0 = -5.4841 \ \text{AND} \ z_0 = 22.5606$$


LOGNORMAL AND NORMAL PLOTS FOR z FIRST 100 TIME STEPS
a)
DATA: 1:100
LOGNORMAL
NORMAL
PDF
z


LOGNORMAL AND NORMAL PLOTS FOR z FIRST 1000 TIME STEPS
b)
DATA: 1:1000
LOGNORMAL
NORMAL
PDF
z


LOGNORMAL AND NORMAL PLOTS FOR z FIRST 10000 TIME STEPS
c)
DATA: 1:10000
LOGNORMAL
NORMAL
PDF
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0 5 10 15 20 25 30 35 40 45 50
z


LOGNORMAL AND NORMAL PLOTS FOR z FIRST 100000 TIME STEPS
d)
DATA: 1:100000
LOGNORMAL
NORMAL
PDF
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0 5 10 15 20 25 30 35 40 45 50
z

LOGNORMAL DATA ASSIMILATION

We start by consider which statistic to use to best represent the underlying analysis pdf.

The three descriptive statistics are the mode 'most likely state', median 'unbiased state' and the mean 'minimum variance'.

Unlike with the Gaussian distribution and other symmetric distributions these three statistics are not identical so which one to use?

PROPERTIES OF THE MULTIVARIATE LOGNORMAL DISTRIBUTION

PROPERTY 1: Median is non-unique.

PROPERTY 2: Moments do not determine the distribution uniquely.

PROPERTY 3: Mean is independent of covariances, and is unbounded with respect to the variances.

PROPERTY 4: Mode is bounded and finite with respect to the variances and covariances.

PROPERTY 5: MODE IS UNIQUE!!!

LOGNORMAL DATA ASSIMILATION

FLETCHER AND ZUPANSKI (2006a; 2007)

By following Lorenc (1986) and extending the error definition from Cohn (1997) we can define a cost function for lognormal background and observational errors as

$$J(\mathbf{x}) = \frac{1}{2}\left(\ln\mathbf{x} - \ln\mathbf{x}_\mathbf{b}\right)^{\mathrm{T}}\mathbf{B}^{-1}\left(\ln\mathbf{x} - \ln\mathbf{x}_\mathbf{b}\right) + \frac{1}{2}\left(\ln\mathbf{y} - \ln\mathbf{h}(\mathbf{x})\right)^{T}\mathbf{R}^{-1}\left(\ln\mathbf{y} - \ln\mathbf{h}(\mathbf{x})\right)$$
$$+ \sum_{i=1}^{N}\left(\ln\mathbf{x} - \ln\mathbf{x}_\mathbf{b}\right) + \sum_{j=1}^{N_o}\left(\ln y_j - \ln h_j(\mathbf{x})\right) \qquad (1)$$

Where the ERRORS are defined by

$$\boldsymbol{\varepsilon}_\mathbf{b} = \frac{\mathbf{x}}{\mathbf{x}_\mathbf{b}} \propto LN(\mathbf{0},\mathbf{B}), \quad \mathbf{AND} \quad \boldsymbol{\varepsilon}_\mathbf{o} = \frac{\mathbf{y}}{\mathbf{h}(\mathbf{x})} \propto LN(\mathbf{0},\mathbf{R}) \qquad (2)$$

MISCONCEPTIONS ABOUT LOGNORMAL DATA ASSIMILATION

1) The theory holds as the background solution is independent of the true solution, it is only an approximation and statistically has no information about the true solution.

2) The theory holds for the observational component as the observations are independent of the observations operator and vice-versa.

3) If two solutions have a relative error of 50% then we are still out by a factor of two in both cases no matter what order of magnitude.

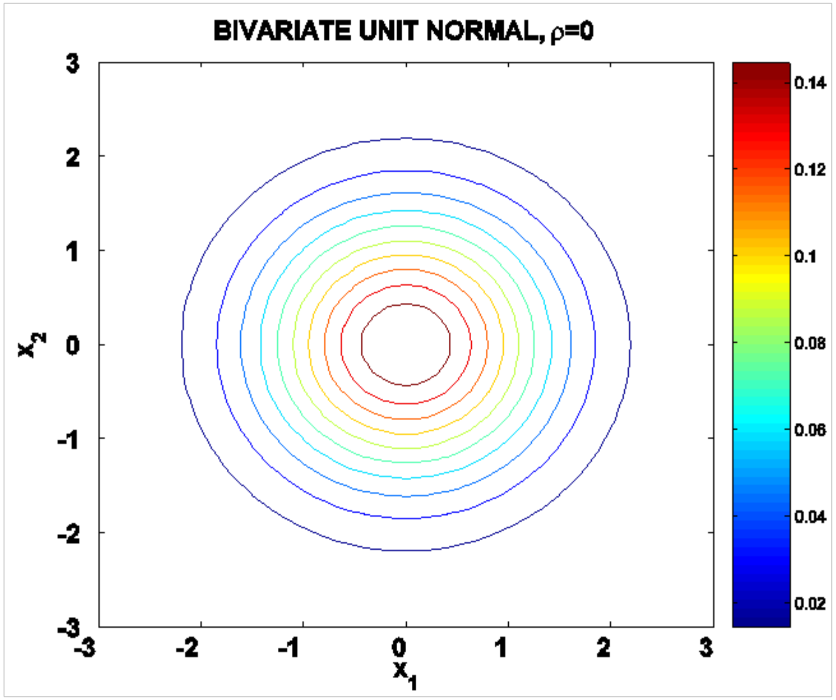
HYBRID DISTRIBUTION
FLETCHER AND ZUPANSKI (2006b)

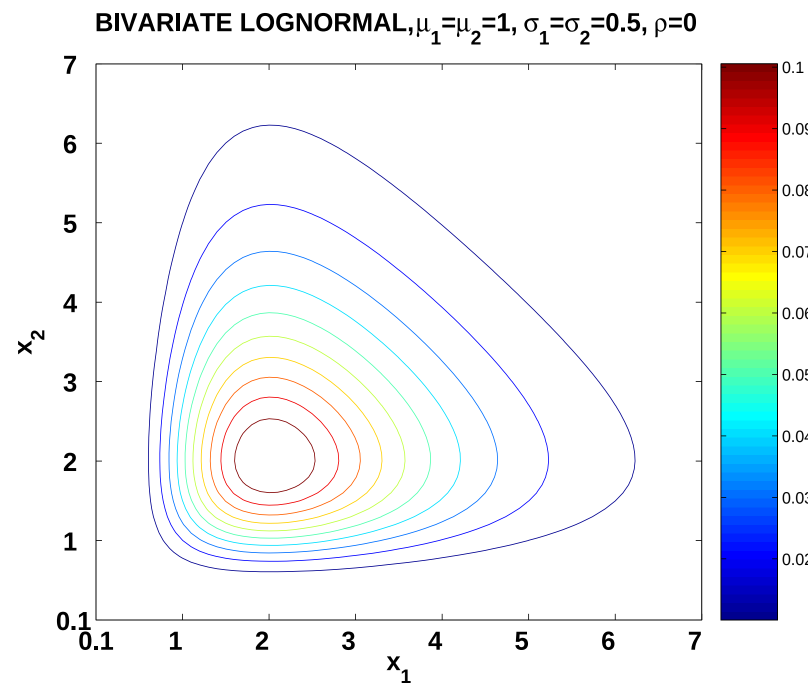
Can define a hybrid normal-lognormal multivariate probability density function of the form

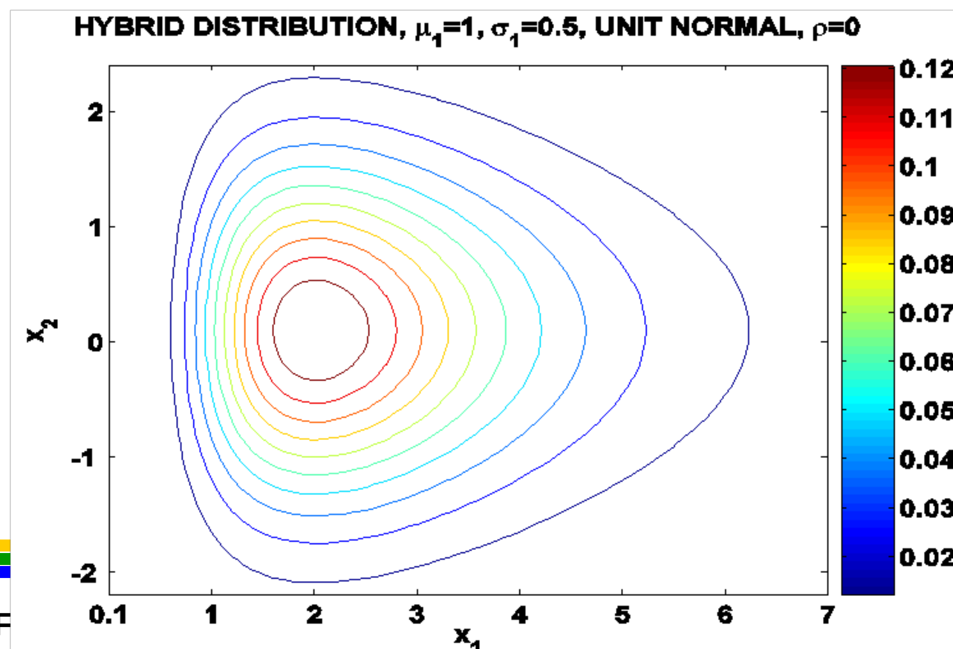
$$f_{p,q}(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{N}{2}} |\boldsymbol{R}|^{\frac{1}{2}}} \left(\prod_{i=p+1}^{N} \frac{1}{x_i} \right) \exp\left\{ -\frac{1}{2} (\hat{\boldsymbol{x}} - \boldsymbol{\mu})^T \boldsymbol{R}^{-1} (\hat{\boldsymbol{x}} - \boldsymbol{\mu}) \right\} \quad (3)$$

WHERE

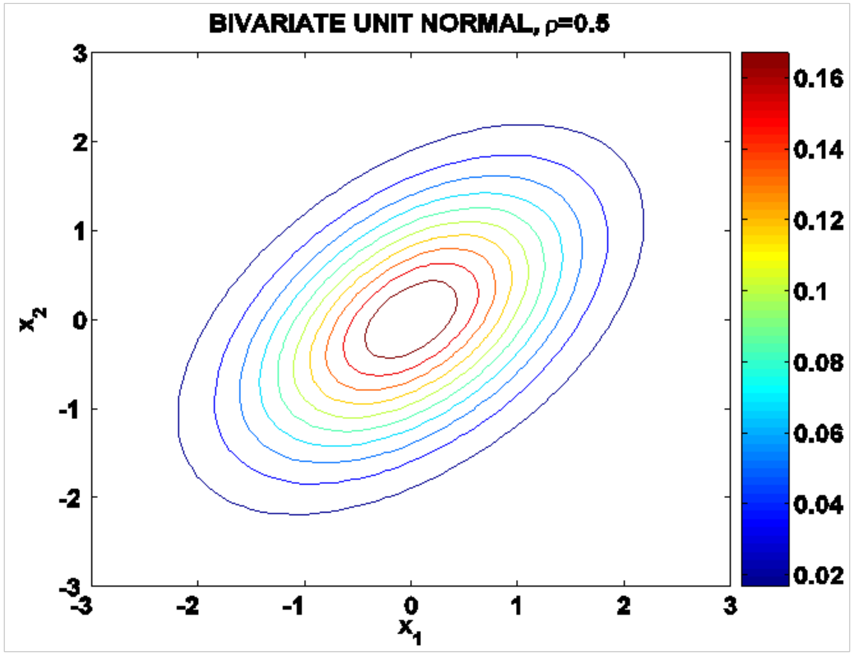
$$\hat{\boldsymbol{x}} = \begin{pmatrix} \boldsymbol{x}_p \\ \ln \boldsymbol{x}_q \end{pmatrix}$$

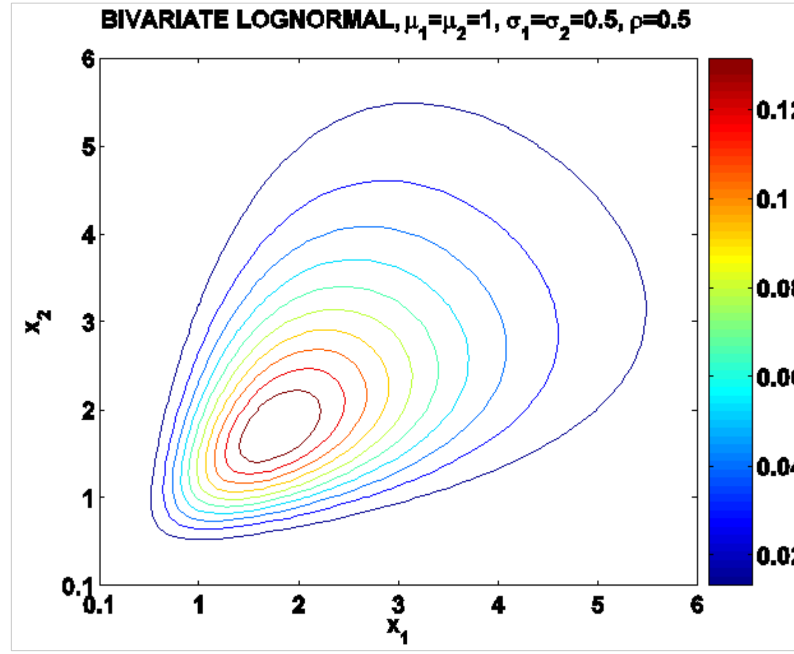


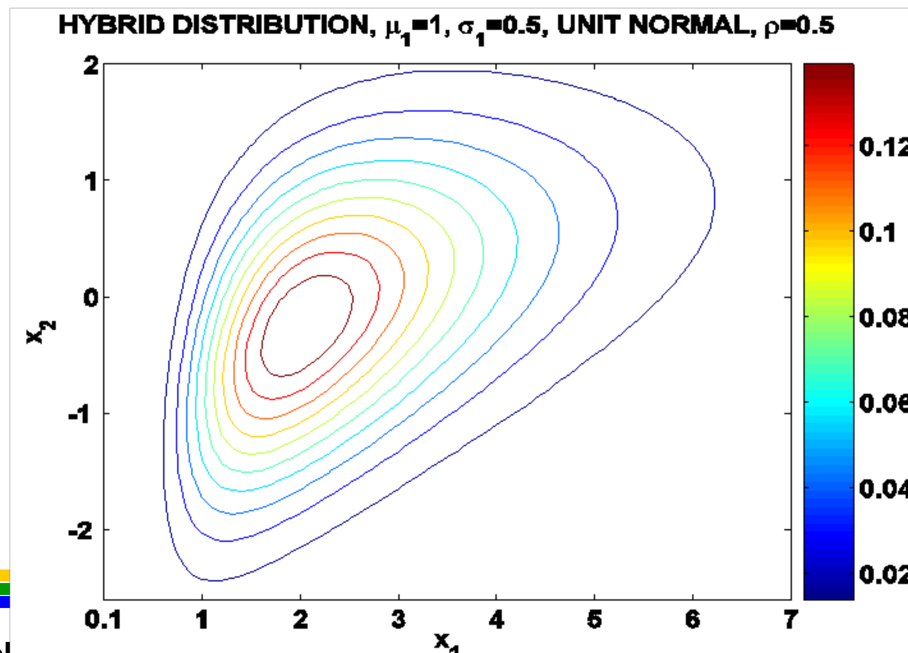




INDEPENDENT







CORRELATED

CG/AR

HYBRID ASSIMILATION

FLETCHER AND ZUPANSKI 2006b, 2007

From the distribution defined in (3) it is possible to defined a cost function following the maximum likelihood approach as set out in Lorenc (1986). Therefore the associated cost function for hybrid background and observational errors is

$$J(\boldsymbol{x}) = \frac{1}{2}\hat{\boldsymbol{\varepsilon}}_b^T \boldsymbol{B}^{-1}\hat{\boldsymbol{\varepsilon}}_b + \frac{1}{2}\hat{\boldsymbol{\varepsilon}}_o^T \boldsymbol{R}^{-1}\hat{\boldsymbol{\varepsilon}}_o + \sum_{i=p_1+1}^{N}\hat{\varepsilon}_{bi} + \sum_{j=p_2+1}^{N_o}\hat{\varepsilon}_{oj} \qquad (4)$$

Where

$$\hat{\boldsymbol{\varepsilon}}_b = \begin{pmatrix} \boldsymbol{x}_{p_1} - \boldsymbol{x}_{b,p_1} \\ \ln \boldsymbol{x}_{q_1} - \ln \boldsymbol{x}_{b,q_1} \end{pmatrix} \qquad \hat{\boldsymbol{\varepsilon}}_o = \begin{pmatrix} \boldsymbol{y}_{p_2} - \boldsymbol{h}_{p_2}(\boldsymbol{x}) \\ \ln \boldsymbol{y}_{q_2} - \ln \boldsymbol{h}_{q_2}(\boldsymbol{x}) \end{pmatrix} \qquad (5)$$

Example with the Lorenz 1963 model

The three non-linear differential equations are given by (Lorenz 1963)

$$\dot{x} = -\sigma x + \sigma y$$

$$\dot{y} = -xz + \rho x - y$$

$$\dot{z} = xy - \beta z$$

$$\beta = \frac{8}{3}, \quad \sigma = 10 \ \text{ AND } \ \rho = 28$$

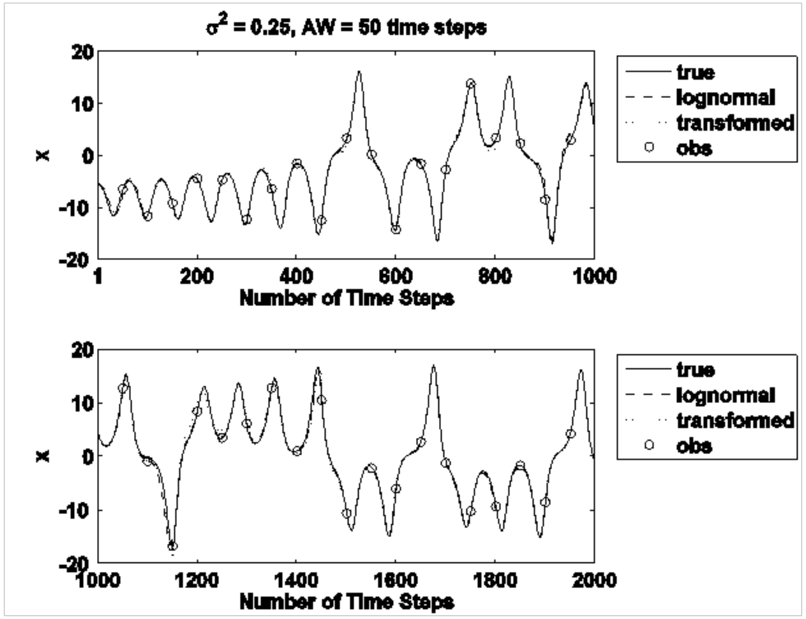
$$x_0 = -5.4458, \quad y_0 = -5.4841 \ \text{ AND } \ z_0 = 22.5606$$

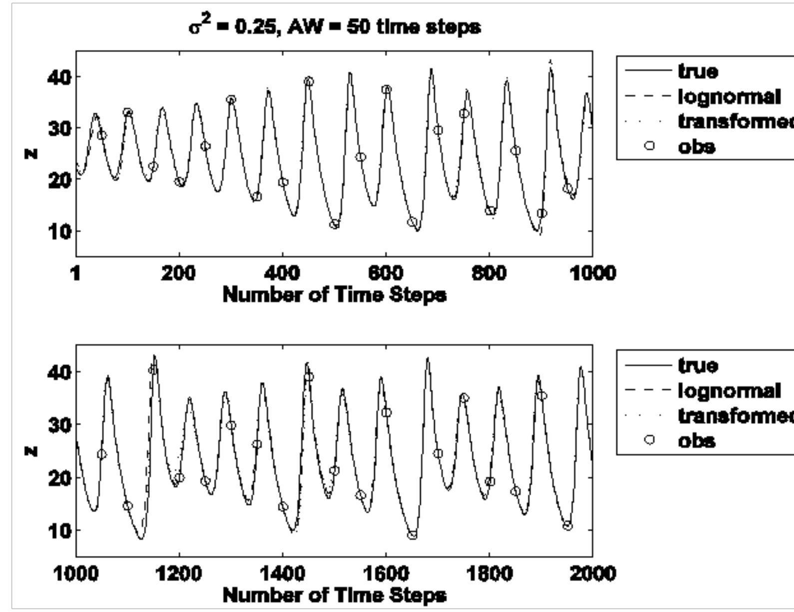
Going to assume x and y components and the associated obs are Gaussian, z is lognormal
(Fletcher and Zupanski 2007)

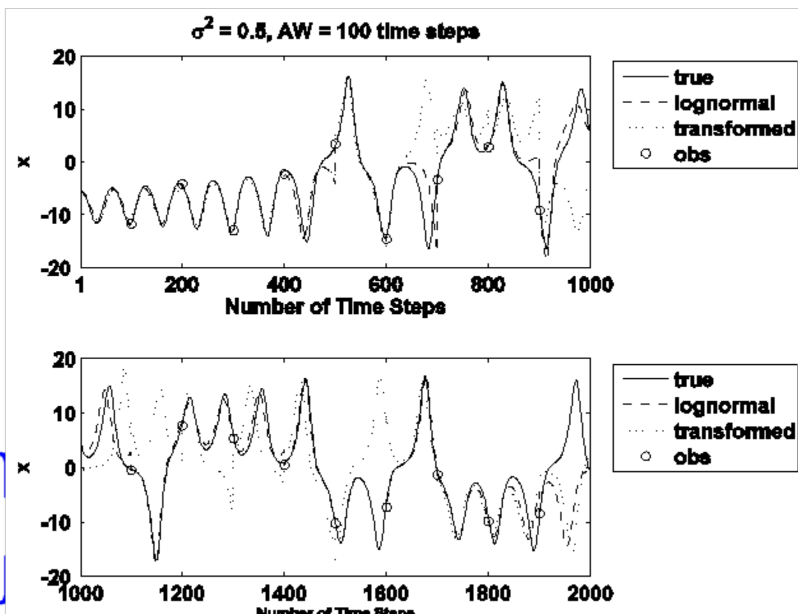
Experiments

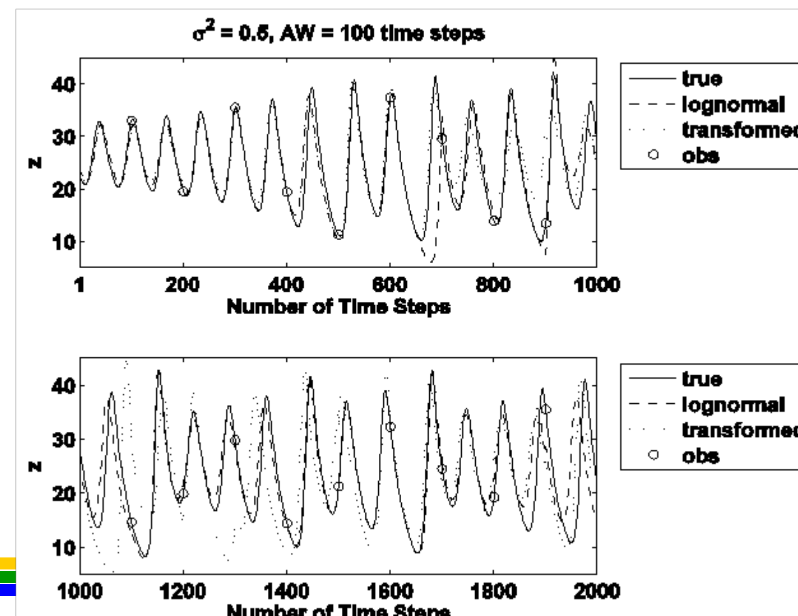
1) Different standard deviations: $\sigma^2 = 0.25, 1$

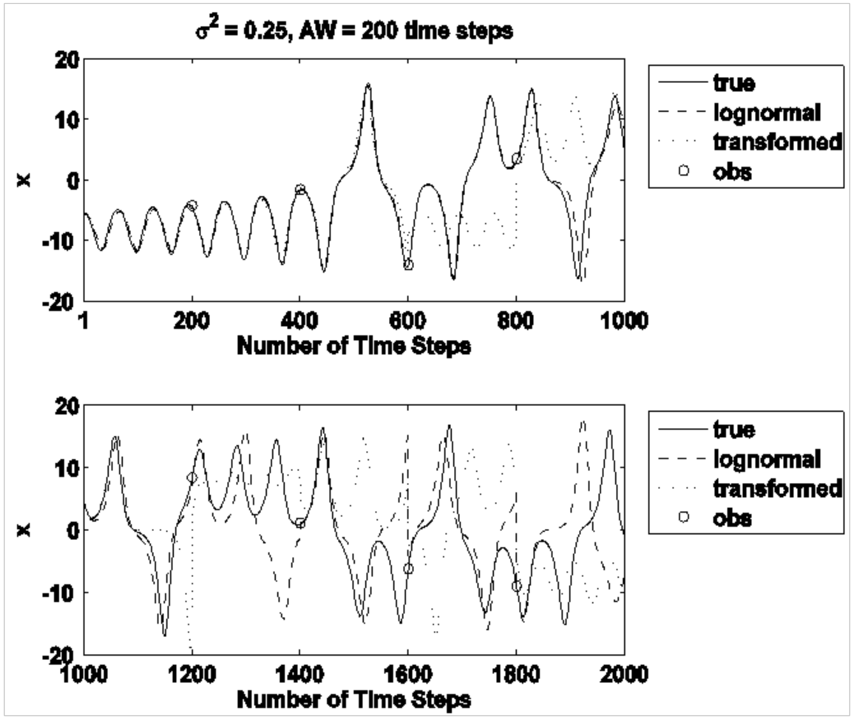
2) Different assimilation window lengths: 50, 100, 200 time steps.


σ² = 0.25, AW = 50 time steps
x
Number of Time Steps
true
lognormal
transformed
obs

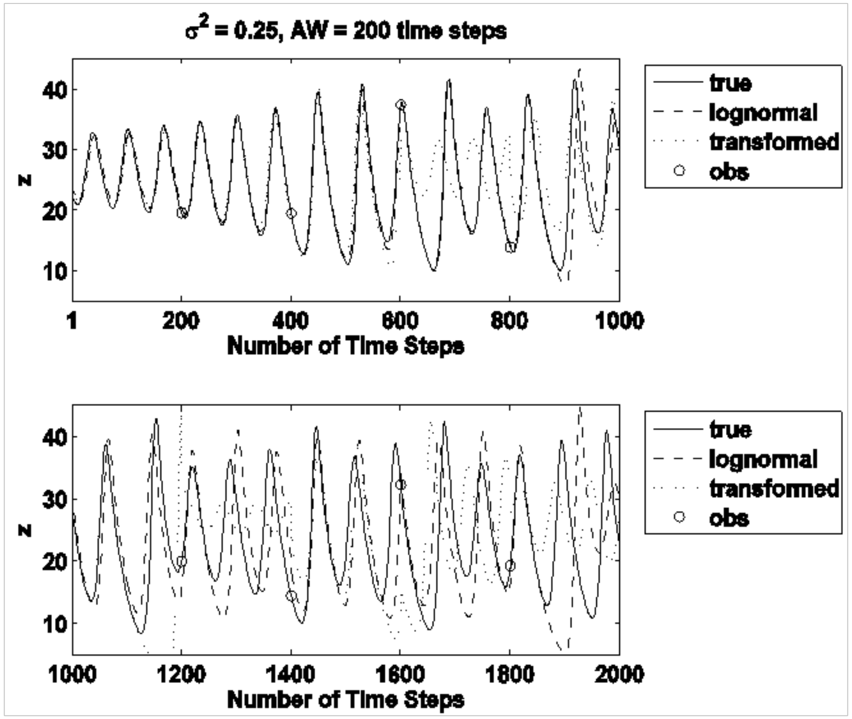

σ² = 0.25, AW = 50 time steps
z
Number of Time Steps
true
lognormal
transformed
obs


σ² = 0.5, AW = 100 time steps
x
Number of Time Steps
true
lognormal
transformed
obs

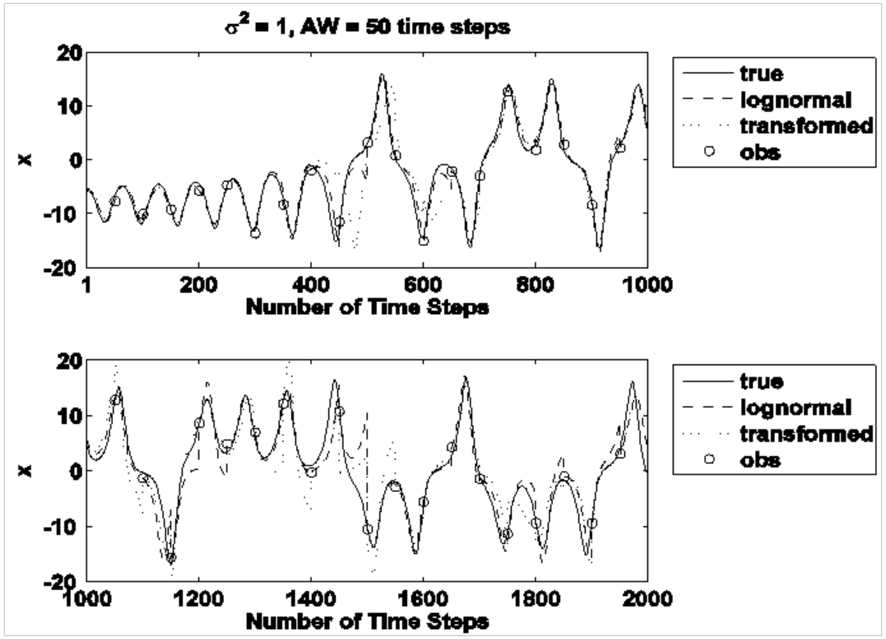

σ² = 0.5, AW = 100 time steps
z
Number of Time Steps
true
lognormal
transformed
obs

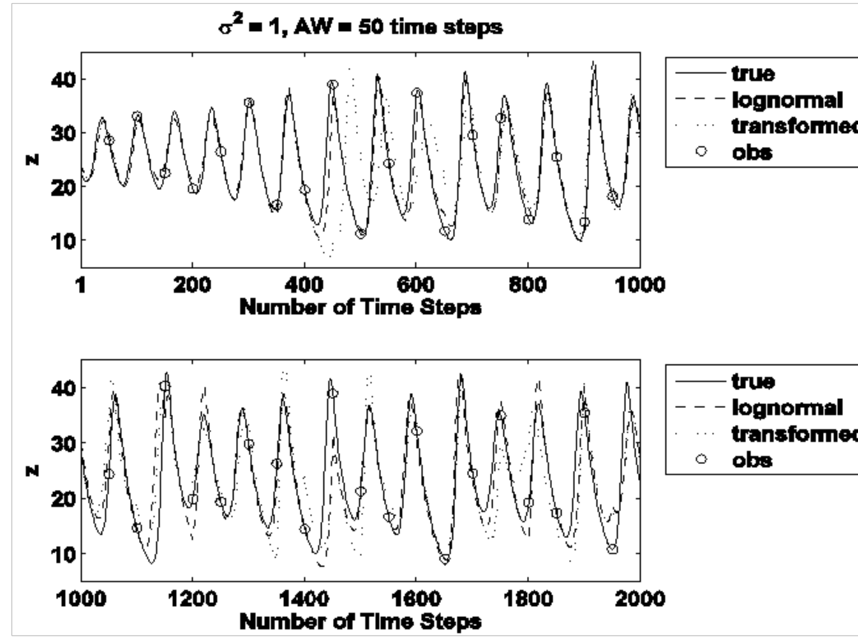


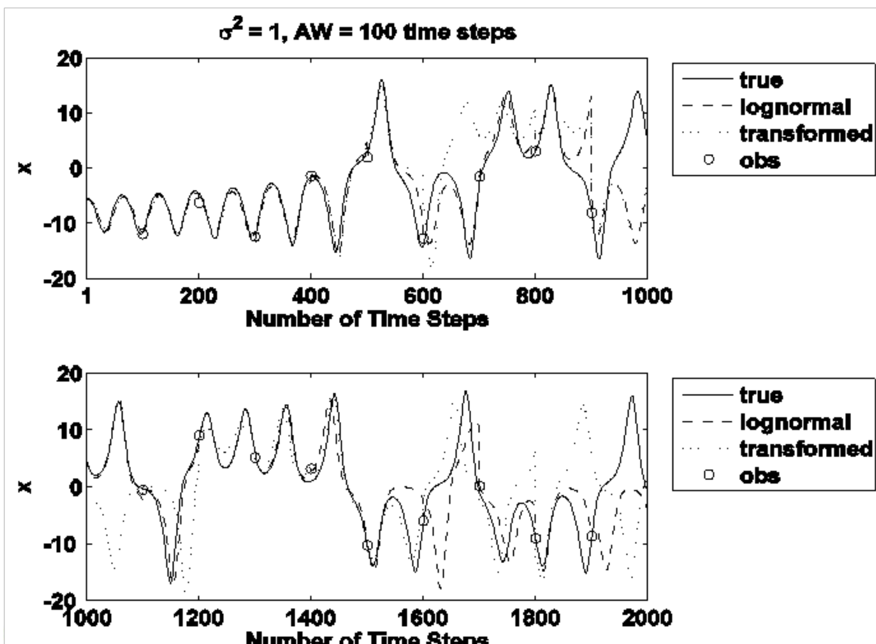
σ^2 = 0.25, AW = 200 time steps
x
Number of Time Steps
true
lognormal
transformed
obs

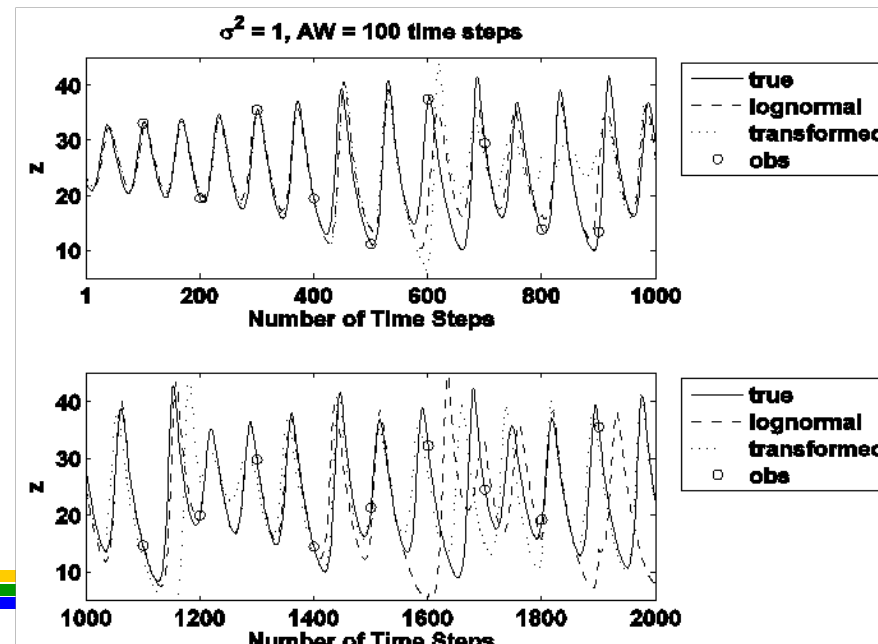


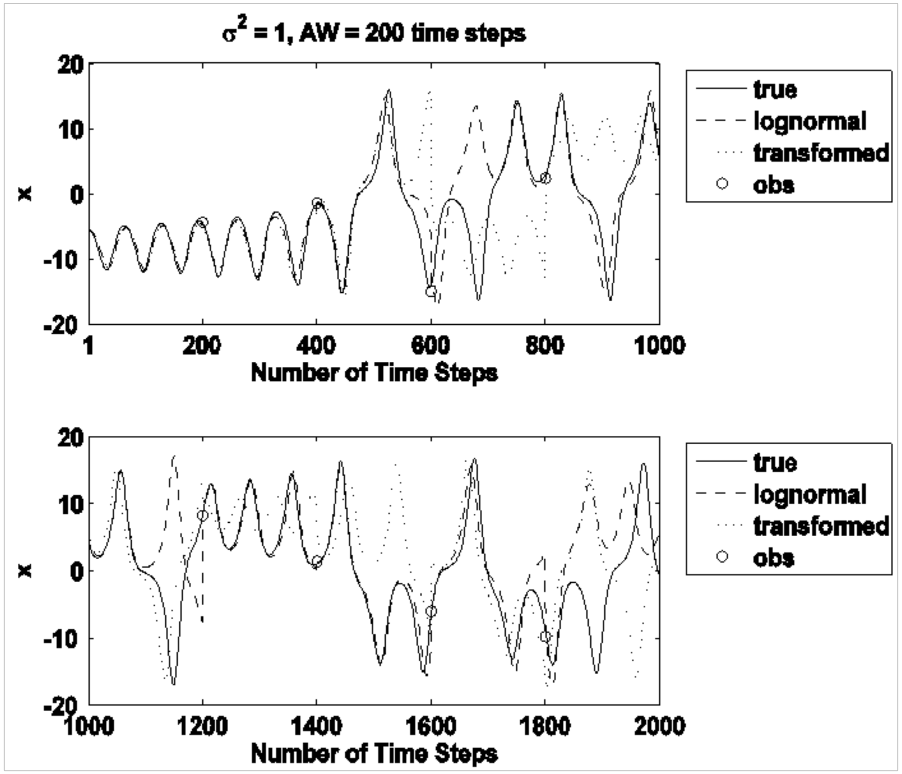
σ^2 = 0.25, AW = 200 time steps
z
Number of Time Steps
true
lognormal
transformed
obs

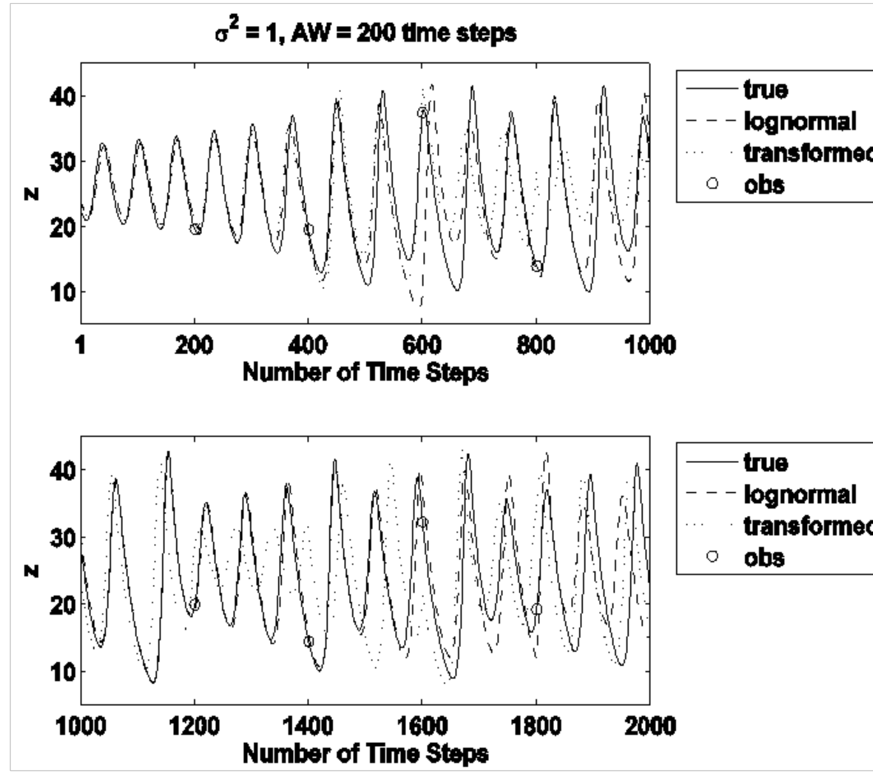

σ² = 1, AW = 50 time steps
x
Number of Time Steps
true
lognormal
transformed
obs


σ² = 1, AW = 50 time steps
z
Number of Time Steps
true
lognormal
transformed
obs


σ² = 1, AW = 100 time steps
x
Number of Time Steps
true
lognormal
transformed
obs


σ² = 1, AW = 100 time steps
z
Number of Time Steps
true
lognormal
transformed
obs


σ² = 1, AW = 200 time steps
x
Number of Time Steps
true
lognormal
transformed
obs


σ² = 1, AW = 200 time steps
z
Number of Time Steps
true
lognormal
transformed
obs

SUMMARY OF 3D VAR RESULTS

- For many observations with small variance then the transform approach is no different than the hybrid approach.
- For larger time between observations then the hybrid approach is more reliable.
- When the time between observations is long and the observations are less accurate then the hybrid approach out performed the transform approach.
- Some observations were ignored by the transform approach even though they are on the correct attractor.
- NOTE: The *z* component of the Lorenz'63 model is neither Gaussian nor lognormal but that the lognormal distribution does capture the first mode.

4D LOGNORMAL DATA ASSIMILATION

Unlike with the three dimensional version of variational data assimilation, the four dimensional version is defined as a weighted least squares problem.

The Gaussian weighted least squares approach to 4D VAR is defined through a calculus of variation problem with initial conditions found through the adjoint.

This weighted least squares approach can be defined for a lognormal framework, which is defined by the following inner product

$$g_1(\boldsymbol{x_0}) = \iiint_A \sum_{i=1}^{N_o} \frac{1}{2} \left\langle \ln \boldsymbol{y}_i - \ln \boldsymbol{h}_i\left(\boldsymbol{M}_i\left(\boldsymbol{x}_0\right)\right), \boldsymbol{R}_i^{-1}\left(\ln \boldsymbol{y}_i - \ln \boldsymbol{h}_i\left(\boldsymbol{M}_i\left(\boldsymbol{x}_0\right)\right)\right)\right\rangle$$

As with the Gaussian case we know that the first variation of the functional defined on the previous is equivalent to

$$\delta g_1(\boldsymbol{x}_0) = \sum_{i=1}^{N_0} \left\langle \ln \boldsymbol{y}_i - \ln \boldsymbol{h}_i\left(\boldsymbol{M}_i(\boldsymbol{x}_0)\right), \boldsymbol{W}_{o,i}\boldsymbol{H}_i\mathbf{M}_i\boldsymbol{R}_i^{-1}\delta\boldsymbol{x}_0 \right\rangle$$
$$= \left\langle \nabla g_1(\boldsymbol{x}_0), \delta\boldsymbol{x}_0 \right\rangle$$

Through using the properties of inner products we get that the gradient is

$$\nabla g_1(\boldsymbol{x}_0) = \sum_{i=1}^{N_o} \left(\boldsymbol{W}_{o,i}\boldsymbol{H}_i\mathbf{M}_i\boldsymbol{R}_i^{-1}\right)^T \left(\ln \boldsymbol{y}_i - \ln \boldsymbol{h}_i\left(\boldsymbol{M}_i(\boldsymbol{x}_0)\right)\right)$$

CG/AR

The solution is a median and not the mode and hence is independent of the variance.

We need to define the functional as

$$g_2\left(\boldsymbol{x_0}\right)=$$

$$\iiint\limits_A \sum_{i=1}^{N_o} \frac{1}{2} \left\langle \ln \boldsymbol{y}_i - \ln \boldsymbol{h}_i\left(\boldsymbol{M}_i\left(\boldsymbol{x}_0\right)\right) + \boldsymbol{R}^T \boldsymbol{1}, \boldsymbol{R}_i^{-1}\left(\ln \boldsymbol{y}_i - \ln \boldsymbol{h}_i\left(\boldsymbol{M}_i\left(\boldsymbol{x}_0\right)\right)\right)\right\rangle$$

Which then has a gradient of

$$\nabla g_2\left(\boldsymbol{x_0}\right) = \sum_{i=1}^{N_o} \left(\boldsymbol{W}_{o,i} \boldsymbol{H}_i \mathbf{M}_i \boldsymbol{R}_i^{-1}\right)^T \left(\ln \boldsymbol{y}_i - \ln \boldsymbol{h}_i\left(\boldsymbol{M}_i\left(\boldsymbol{x}_0\right)\right) + \boldsymbol{R}_i^T \boldsymbol{1}\right)$$

CG/AR

Current Gaussian approach

$$J_G(\mathbf{x}_0)=\sum_{i=1}^{N_0}\left\langle \mathbf{R}_{G,i}^{-1}\left(\mathbf{y}_i-\mathbf{h}_i\left(M_{0,i}\left(\mathbf{x}_0\right)\right)\right),\left(\mathbf{y}_i-\mathbf{h}_i\left(M_{0,i}\left(\mathbf{x}_0\right)\right)\right)\right\rangle$$

Improved Transform technique

$$J_{TR}(\mathbf{x}_0)=\sum_{i=1}^{N_0}\left\langle \mathbf{R}_{L,i}^{-1}\left(\ln\mathbf{y}_i-\ln\left(\mathbf{h}_i\left(M_{0,i}\left(\mathbf{x}_0\right)\right)\right)\right),\left(\ln\mathbf{y}_i-\ln\left(\mathbf{h}_i\left(M_{0,i}\left(\mathbf{x}_0\right)\right)\right)\right)\right\rangle$$

New Lognormal 4D VAR approach:

$$J_{LN}(\mathbf{x}_0)=\sum_{i=1}^{N_0}\left\langle \mathbf{R}_{L,i}^{-1}\left(\ln\mathbf{y}_i-\ln\left(\mathbf{h}_i\left(M_{0,i}\left(\mathbf{x}_0\right)\right)\right)+\mathbf{R}_{L,i}\mathbf{1}_{N_{0,1}}\right),\left(\ln\mathbf{y}_i-\ln\left(\mathbf{h}_i\left(M_{0,i}\left(\mathbf{x}_0\right)\right)\right)\right)\right\rangle$$

Term that gives the mode

PROBABILITY APPROACH

$$P\left(x_0, x_1, x_2, \ldots, x_{N_o} \middle| y_1, y_2, y_3, \ldots, y_{N_0}\right) =$$
$$P(x_0)P(x_1|x_0)P(y_1|x_1, x_0)P(x_2|y_1, x_1, x_0)$$
$$\ldots P\left(y_{N_O} \middle| x_N, y_{N_o - 1}, x_{N_o}, \ldots, y_1, x_1, x_0\right)$$

Bayesian networks allow us to remove terms that are not conditioned on other random variables.

$$P\left(\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_o} \middle| \mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_{N_0}\right) = P(\mathbf{x}_0) \prod_{i=1}^{N_o} P(\mathbf{y}_i | \mathbf{x}_0)$$

Taking the negative logarithm of the circled pdf in the previous slide results in

$$\min\left\{J\left(\mathbf{x}_0\right)=-\ln P\left(\mathbf{x}_0\right)-\sum_{i=1}^{N_0}\ln P\left(\mathbf{y}_i\middle|\mathbf{x}_0\right)\right\}$$

This can now be used to derive a 4D VAR system for any distributed random variable

$$P\left(\boldsymbol{x_0},\boldsymbol{x_1},\boldsymbol{x_2},\ldots,\boldsymbol{x}_{N_o}\,\middle|\,\boldsymbol{y_1},\boldsymbol{y_2},\boldsymbol{y_3},\ldots,\boldsymbol{y}_{N_0}\right)=$$

$$P(\boldsymbol{x_0})\prod_{i=1}^{N_o}P\left(\boldsymbol{y}_i\,\middle|\,\boldsymbol{x_0}\right)$$

For the multivariate Gaussian case we have

$$P(\boldsymbol{x_0})\propto\exp\left\{-\frac{1}{2}\left(\boldsymbol{x_0}-\boldsymbol{x_{b,0}}\right)^T\boldsymbol{B}_0^{-1}\left(\boldsymbol{x_0}-\boldsymbol{x_{b,0}}\right)\right\}$$

$$P\left(\boldsymbol{y}_i\,\middle|\,\boldsymbol{x_0}\right)\propto\exp\left\{-\frac{1}{2}\left(\boldsymbol{y_i}-\boldsymbol{h}_i\left(\boldsymbol{M}_i\left(\boldsymbol{x_0}\right)\right)\right)^T\boldsymbol{R}_i^{-1}\left(\boldsymbol{y_i}-\boldsymbol{h}_i\left(\boldsymbol{M}_i\left(\boldsymbol{x_0}\right)\right)\right)\right\}$$

For the multivariate lognormal case we have

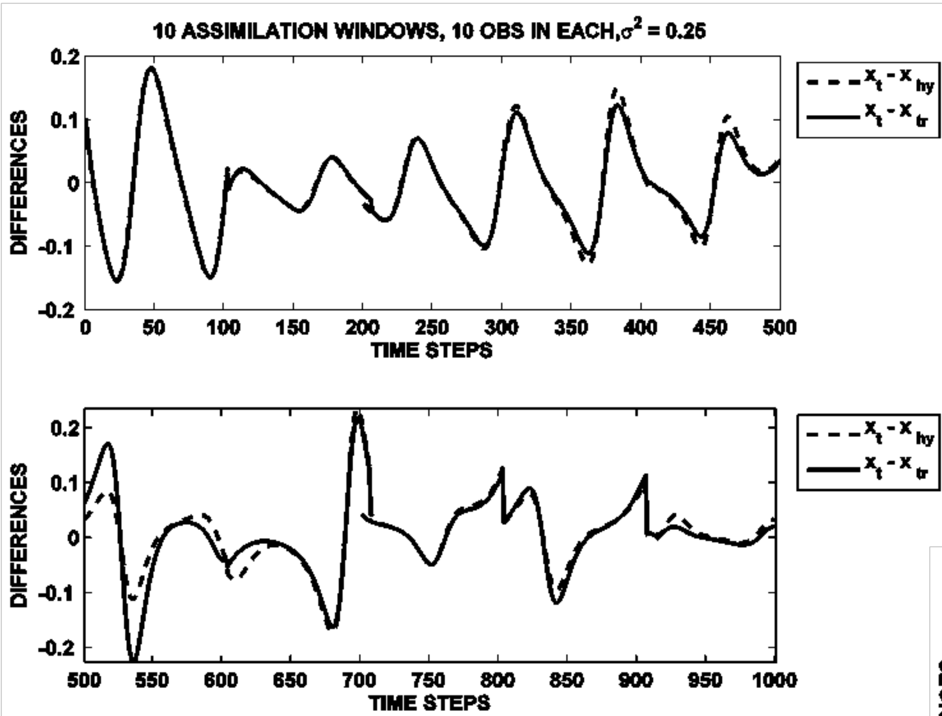
$$P(\boldsymbol{x_0}) \propto \left(\prod_{j=1}^{N} \frac{\boldsymbol{x_{0,i}}}{\boldsymbol{x_{b,0,j}}} \right) \times$$

$$\exp\left\{ -\frac{1}{2} \left(\ln \boldsymbol{x_0} - \ln \boldsymbol{x_{b,0}} \right)^T \boldsymbol{B}_0^{-1} \left(\ln \boldsymbol{x_0} - \ln \boldsymbol{x_{b,0}} \right) \right\}$$

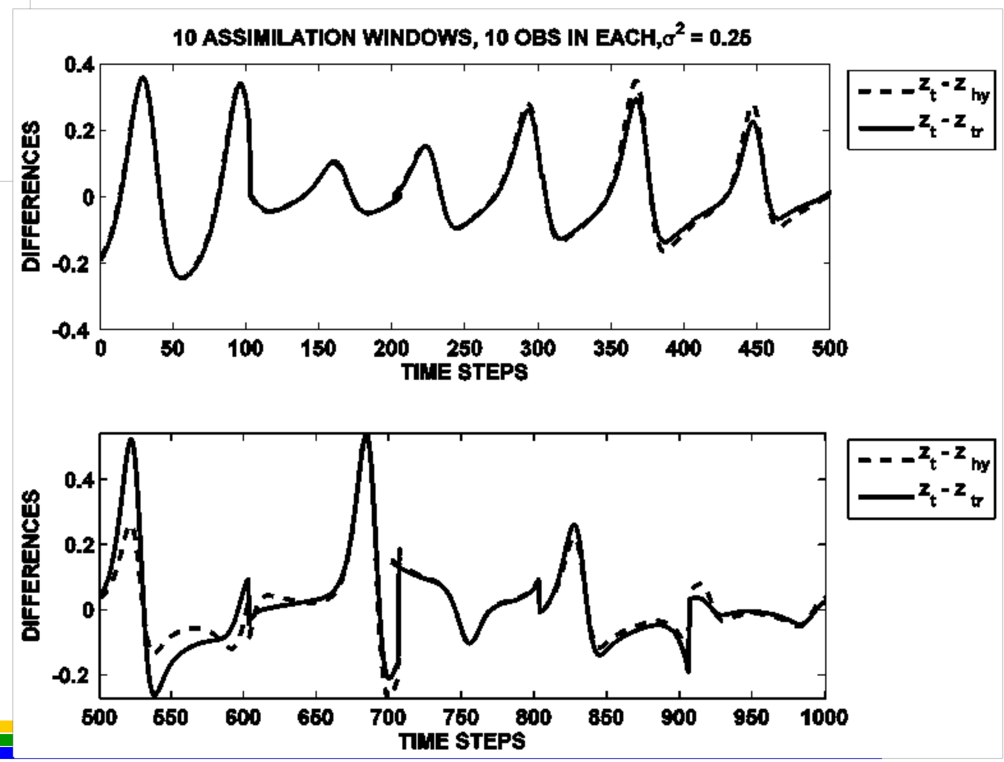
$$P(\boldsymbol{y}_i | \boldsymbol{x_0}) \propto \left(\prod_{k=1}^{N_{o,i}} \frac{h_{i,k}\left(\boldsymbol{M}\left(\boldsymbol{x}_0\right)\right)}{y_{i,k}} \right) \times$$

$$\exp\left\{ -\frac{1}{2} \left(\boldsymbol{y_i} - \boldsymbol{h}_i\left(\boldsymbol{M}_i\left(\boldsymbol{x_0}\right)\right) \right)^T \boldsymbol{R}_i^{-1} \left(\boldsymbol{y_i} - \boldsymbol{h}_i\left(\boldsymbol{M}_i\left(\boldsymbol{x_0}\right)\right) \right) \right\}$$

Results with the Lorenz 1963 model

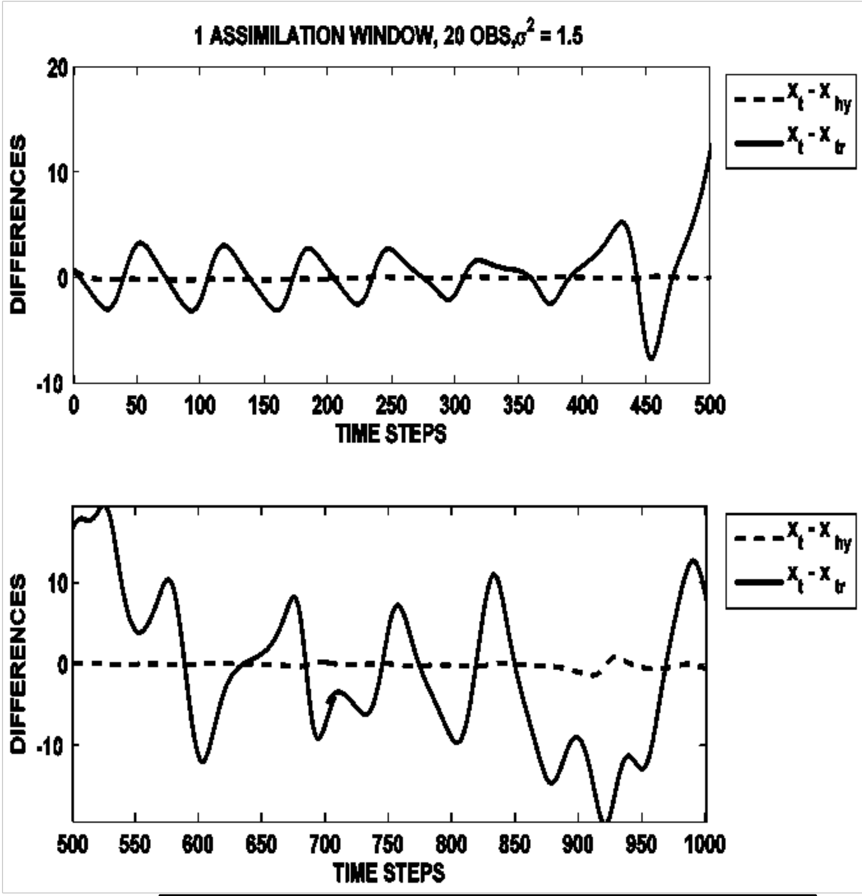


Plots of the differences in the trajectories with many accurate obs with short assimilation windows

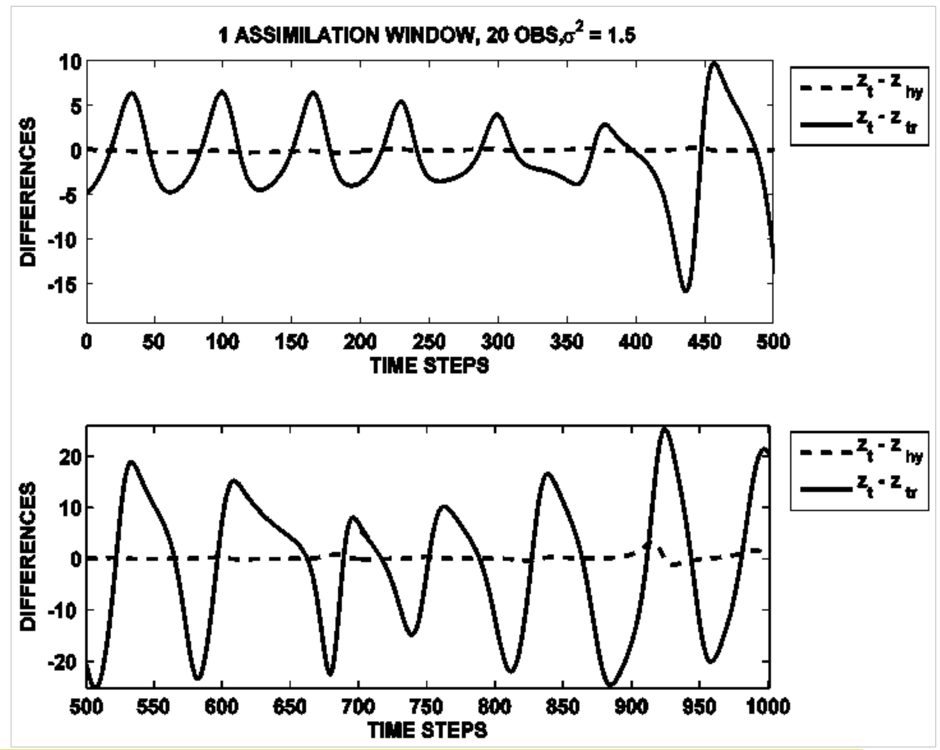


Note: transform and hybrid approach quite similar

BEST CASE SCENERIO



When assimilation window is large (1000ts) with fewer and less accurate obs, hybrid approach is more accurate



Transform approach converges quickly to the wrong solution

Conclusions, Implications and Further Work

- Careful which statistic to use to analyses
- Mode is closer to the true trajectory in the Lorenz 63 model
- Possible to assimilate variables of mixed types simultaneously
- Incremental version???
- Combine other distributions?? i.e. Gamma, Normal, Lognormal
- Faster method for finding positive definite variables
- No need to change background error covariance matrix
- Improved moisture fields
- More reliable forecasts, less likely to issue false warnings
- Better prediction of clouds and dust storms as the moisture field is more accurate
- Better prediction of dust storms, hurricane intensity, super cells.

Conclusions, Implications and Further Work

- **Implement the hybrid cost function into the MLEF at CIRA**

- **Implement the hybrid cost function into the Weather, Research and Forecasting (WRF) 3D VAR.**

- **Develop a new version of humidity/temperature retrievals from brightness temperature with the hybrid method**

- **Develop an incremental version similar to the operational centres**

- **Derive new variational schemes for other distributions.**

REFERENCES

COHN, S.E., 1997: AN INTRODUCTION TO ESTIMATION THEORY. *J. Met. Soc. Japan,* 75, 420-436.

DEBLONDE, G. AND S. ENGLISH, 2003: ONE-DIMENSIONAL VARIATIONAL RETRIEVALS FROM SSMISS-SIMULATED OBSERVATIONS. *J. Appl. Met.* 42, 1406-1420

DEE, D.P. AND A.M. DA SILVA, 2003: THE CHOICE OF VARIABLE FOR ATMOSPHERIC MOISTURE ANALYSIS. *Mon. Wea. Rev.* 131, 155-171

DERBER, J.C. AND W.-S. WU, 1998: THE USE OF TVOS CLOUD-CLEARED RADIANCES IN THE NCEP SSI ANALYSIS SYSTEM. *Mon. Wea. Rev.* 126, 2287-2299

FLETCHER, S.J., 2009: A BAYESIAN NETOWRK APPROACH FOR GAUSSIAN AND NON-GAUSSIAN FOUR DIMENSIONAL DATA ASSIMILATION. To be submitted to *Q. J. Roy. Meteor. Soc.*

FLETCHER, S.J. AND M. ZUPANSKI, 2006a: A DATA ASSIMILATION METHOD FOR LOGNORMAL DISTRIBUTED OBSERVATIONAL ERRORS. *Q. J. Roy. Meteor. Soc*. 132, 2505-2520

FLETCHER, S.J. AND M. ZUPANSKI, 2006b: A HYBRID MULTIVARIATE NORMAL AND LOGNORMAL DISTRIBUTION FOR DATA ASSIMILATION. *Atmos. Sci. Letters.* 7, 43-46

FLETCHER, S.J. AND M. ZUPANSKI, 2007: IMPLICATIONS AND IMPACTS OF TRANSFORMING LOGNORMAL VARIABLES INTO NORMAL VARIABLES IN VAR, IN PRINT: *Meteorologische Zeitschrift.*

LORENC, A.C., 1986: ANALYSIS METHODS FOR NUMERICAL WEATHER PREDICTION. *Q. J. Roy. Meteor. Soc.* 112, 1177-1194

LORENZ, E.N., 1963: DETERMINSTIC NONPERIODIC FLOW. *J. Atmos. Sci.* 20, 130-141

MILES, N.L., J. VERLINDE AND E.E. CLOTHIAUX, 2000: CLOUD DROPLET SIZE DISTRIBUTION IN LOW-LEVEL STRATIFORM CLOUDS. *J. Atmos. Sci.* 57, 295-311

SENGUPTA, M., E.E. CLOTHIAUX AND T.P. ACKERMAN, 2004: CLIMATOLOGY OF WARM BOUNDARY LAYER CLOUDS AT THE ARM SGP SITE AND THEIR COMPARISIONS TO MODELS. *J. Clim.* 17, 4760-4782

STEPHENS, G.L. AND COAUTHORS, 2002: THE CLOUDSAT MISSION AND THE A-TRAIN. *Bull. Amer. Meteor. Soc.* 83, 1771-1190