Utilizing Machine Learning to Optimize the Choice of Error Distribution in Data Assimilation

Steven J. Fletcher Senne Van Loon Md. Jakir Hossen
Micheal R. Goodliff Anton J. Kliewer

Cooperative Institute for Research in the Atmosphere
Colorado State University

BIRS,
22nd March 2023

COLORADO STATE UNIVERSITY

Outline

Outline

Lognormal Observational Errors

In Steve Cohn's seminal paper in 1997, Cohn [1997], there is a definition for the lognormally distributed errors associated with direct observations of a lognormally distributed control variable, which is in terms of the ratio of the observed and the model equivalent. In Fletcher and Zupanski [2006a] the definition for the lognormally distributed observational error, $\boldsymbol{\varepsilon_o}$, was extended to the case of non-direct observations as

$$\boldsymbol{\varepsilon_{o,i}} \equiv \frac{\boldsymbol{y_j}}{\boldsymbol{h_j}\left(\boldsymbol{x}\right)}, \qquad j = 1, 2, \ldots, N_o, \tag{1}$$

where $\boldsymbol{y}$ is the vector of observations, $\boldsymbol{h}$ is the nonlinear observation operator, $\boldsymbol{x}$ is the model state at the time of the observation, and N_o is the total number of observations.

The reason for using the ratio instead of the difference is because it enables us to use the property that the ratio of two independent lognormally distributed random variables is also a lognormally distributed random variable.

Lognormal Background Errors

Given the definition for the lognormal observational errors, the next step is to introduce the lognormal equivalent for the background errors, $\varepsilon_{\boldsymbol{b}}$, which comes from Fletcher and Zupanski [2007], again defined as a ratio, given by

$$\varepsilon_{\boldsymbol{b},\boldsymbol{i}} \equiv \frac{\boldsymbol{x}_{\boldsymbol{i}}^{\boldsymbol{t}}}{\boldsymbol{x}_{\boldsymbol{i}}^{\boldsymbol{b}}}, \qquad i = 1, 2, \ldots, N, \tag{2}$$

where $\boldsymbol{x}^{\boldsymbol{t}}$ is the true state, $\boldsymbol{x}^{\boldsymbol{b}}$ is the background state, and N is the total number of state variables.

Outline

Bayes Theorem

Given the definition of the errors, the next step in deriving the variational form of data assimilation is the consider Bayes theorem, which is given by

$$P(A \mid B) \propto P(B \mid A) P(A), \tag{3}$$

where $P(A \mid B)$ is referred to as the posterior distribution, $P(B \mid A)$ is the likelihood distribution, and $P(A)$ is the apriori distribution.

From Lorenc [1986] the events in (3) are defined as $A : \boldsymbol{x} = \boldsymbol{x}^{\boldsymbol{t}}$ and $B : \boldsymbol{y} = \boldsymbol{y}^{\boldsymbol{o}}$. Thus we are either seeking the state that minimises the variance of the posterior distribution (mean), the unbiased state (median) or the state that has the highest probability of occurring, referred to as the maximum likelihood state (mode).

For variational data assimilation we seek the mode and the approach that is applied here is to find the minimum of the negative logarithm of (3).

Lognormal Distribution

First we need the definition of the multivariate lognormal distribution, which is given by

$$LN(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \equiv \prod_{i=1}^{n} \left(\frac{1}{x_i}\right) \frac{1}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} (\ln \boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\ln \boldsymbol{x} - \boldsymbol{\mu})\right\}, \quad (4)$$

where $\boldsymbol{\mu} \equiv \mathbb{E}[\ln \boldsymbol{x}]$ is the mean of $\ln \boldsymbol{x}$, not $\boldsymbol{x}$, with $\mathbb{E}$ representing the expectation operator, and $\boldsymbol{\Sigma}$ is the covariance matrix of $\ln \boldsymbol{x}$.

An important property of the lognormal distribution to note here is that the **logarithm of a lognormally distributed random variable is a Gaussian random variable**, whilst the **exponential of a Gaussian random variable is a lognormal random variable**

Lognormal 3DVAR

Thus, given the definitions for the lognormally distributed background and observational errors, and the definition for the multivariate distribution, it is possible to derive the associated 3DVAR cost function as

$$\begin{aligned} J\left(\boldsymbol{x}^{\boldsymbol{t}}\right) &= \frac{1}{2}\left(\ln \boldsymbol{x}^{\boldsymbol{t}} - \ln \boldsymbol{x}^{\boldsymbol{b}}\right)^{T} \boldsymbol{B}_{\boldsymbol{L}}^{-1}\left(\ln \boldsymbol{x}^{\boldsymbol{t}} - \ln \boldsymbol{x}^{\boldsymbol{b}}\right) \\ &+ \left\langle\left(\ln \boldsymbol{x}^{\boldsymbol{t}} - \ln \boldsymbol{x}^{\boldsymbol{b}}\right), \mathbf{1}_{\boldsymbol{N}}\right\rangle \\ &+ \frac{1}{2}\left(\ln \boldsymbol{y} - \ln \boldsymbol{h}\left(\boldsymbol{x}\right)\right)^{T} \boldsymbol{R}_{\boldsymbol{L}}^{-1}\left(\ln \boldsymbol{y} - \ln \boldsymbol{h}\left(\boldsymbol{x}\right)\right) \\ &+ \left\langle\left(\ln \boldsymbol{y} - \ln \boldsymbol{h}\left(\boldsymbol{x}\right)\right), \mathbf{1}_{\boldsymbol{N_o}}\right\rangle, \end{aligned} \tag{5}$$

where $\boldsymbol{B_L}$ and $\boldsymbol{R_L}$ are the lognormal background and observational error covariance matrices respectively, and $\mathbf{1}$ is a column vector of 1s of dimension N or N_o.

Why the mode?

As part of the original work in Fletcher and Zupanski [2006a] a reviewer suggested that we picked the mode because it was the easier one to find. This is not true! The reason why we picked the mode is because it is the only one of the descriptive statistics that is degenerate with respect to the variance, but is also unique. For left (positive) skewed distributions you have the property that

$$\boldsymbol{mode} \leq \boldsymbol{median} \leq \boldsymbol{mean}$$

where for the lognormal distribution the three descriptive statistics can be shown to be

$$\exp\left\{\boldsymbol{\mu} - \langle \boldsymbol{\Sigma}, \mathbf{1}_{\boldsymbol{N}} \rangle\right\} < \exp\left\{\boldsymbol{\mu}\right\} < \exp\left\{\boldsymbol{\mu} + \frac{diag\left(\boldsymbol{\Sigma}\right)}{2}\right\}. \tag{6}$$

It is clear from the inequality above that as the variances increase that the mode is tending towards 1, whereas for the same situation the mean is increasing, whilst the median is invariant.

Outline

Fletcher-Zupanski Distribution

In Fletcher and Zupanski [2006b] we derived a new probability density function (PDF), where the starting point is to assume that there are p Gaussian random variables and q lognormal random variables, such that $N = p + q$. Thus the associated PDF is given by

$$\begin{aligned} MX\left(\boldsymbol{\mu_{mx}}, \boldsymbol{\Sigma_{mx}}\right) &\equiv \prod_{i=p+1}^{N} \left(\frac{1}{x_i}\right) \frac{1}{(2\pi)^{\frac{N}{2}} \left|\boldsymbol{\Sigma_{mx}}\right|^{\frac{1}{2}}} \\ &\times \exp\left\{ \begin{pmatrix} \boldsymbol{x_p} - \boldsymbol{\mu_p} \\ \ln \boldsymbol{x_q} - \boldsymbol{\mu_q} \end{pmatrix}^T \boldsymbol{\Sigma}_{\boldsymbol{mx}}^{-1} \begin{pmatrix} \boldsymbol{x_p} - \boldsymbol{\mu_p} \\ \ln \boldsymbol{x_q} - \boldsymbol{\mu_q} \end{pmatrix} \right\}, \end{aligned} \tag{7}$$

where $\boldsymbol{\Sigma_{mx}}$ is the covariance matrix between the Gaussian and lognormal random variables. Of note here is the associated mode:

$$\boldsymbol{x_{mode}} = \begin{pmatrix} \boldsymbol{\mu_p} - \langle \boldsymbol{\Sigma_{pq}}, \mathbf{1_q} \rangle \\ \exp\left\{ \boldsymbol{\mu_q} - \langle \boldsymbol{\Sigma_{qq}}, \mathbf{1_q} \rangle \right\} \end{pmatrix} \tag{8}$$

Fletcher-Zupanski Distribution

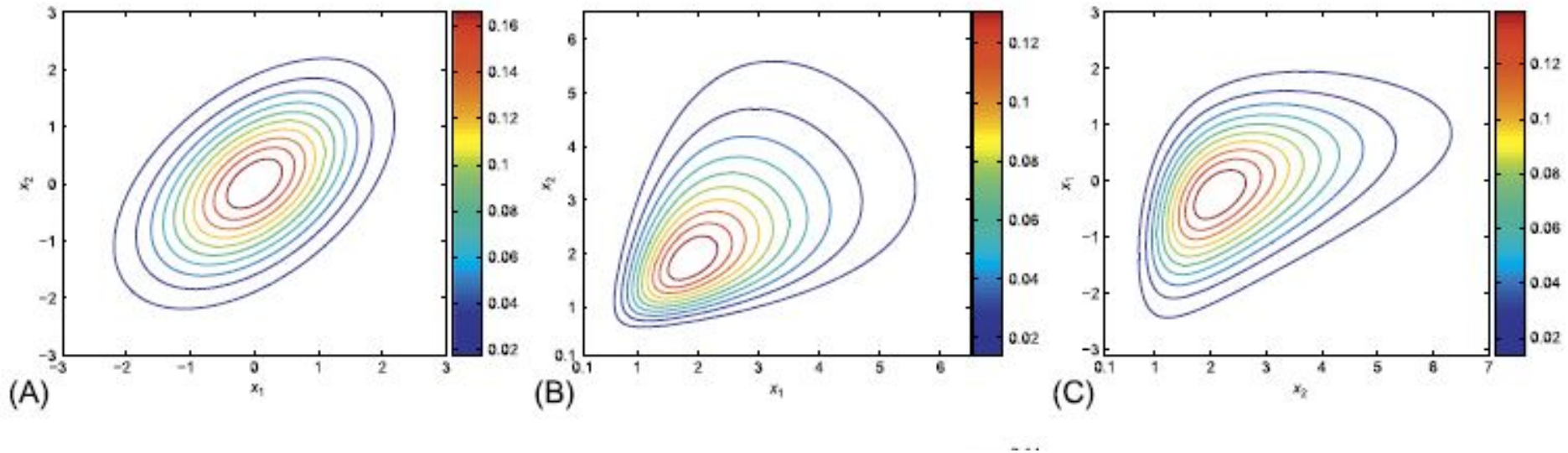


Figure 1: Plot of bivariate Gaussian, lognormal, and Fletcher-Zupanski distributions with $\rho = 0.5$.

Mixed Gaussian-lognormal 3DVAR

To be able to derive the associated 3DVAR cost function for the mixed distribution we need the definition for the mixed distributed errors, which are given by

$$\boldsymbol{\varepsilon}^{\boldsymbol{b}}_{\boldsymbol{mx}} \equiv \begin{pmatrix} \boldsymbol{x}^{\boldsymbol{t}}_{\boldsymbol{p}_1} - \boldsymbol{x}^{\boldsymbol{b}}_{\boldsymbol{p}_1} \\ \ln \boldsymbol{x}^{\boldsymbol{t}}_{\boldsymbol{q}_1} - \ln \boldsymbol{x}^{\boldsymbol{b}}_{\boldsymbol{q}_1} \end{pmatrix}, \quad \boldsymbol{\varepsilon}^{\boldsymbol{o}}_{\boldsymbol{mx}} \equiv \begin{pmatrix} \boldsymbol{y}_{\boldsymbol{p}_2} - \boldsymbol{h}_{\boldsymbol{p}_2}(\boldsymbol{x}) \\ \ln \boldsymbol{y}_{\boldsymbol{q}_2} - \ln \boldsymbol{h}_{\boldsymbol{q}_2}(\boldsymbol{x}) \end{pmatrix} \tag{9}$$

where there are p_1 Gaussian distributed background errors, p_2 distributed Gaussian observational errors, q_1 lognormally distributed background errors, and q_2 lognormally distributed observational errors, with $N = p_1 + q_1$, $N_o = p_2 + q_2$, and it maybe the case that $p_1 \neq p_2$ and $q_1 \neq q_2$.

Mixed Gaussian-lognormal 3DVAR

Thus given the definitions for the mixed distribution errors and the multivariate PDF, through following the maximum likelihood approach the associated cost function is

$$
\begin{aligned}
J_{mx}(\boldsymbol{x}) &= \frac{1}{2}\begin{pmatrix} \boldsymbol{x}^{t}_{p_1} - \boldsymbol{x}^{b}_{p_1} \\ \ln \boldsymbol{x}^{t}_{q_1} - \ln \boldsymbol{x}^{b}_{q_1} \end{pmatrix}^{T} \boldsymbol{B}^{-1}_{mx} \begin{pmatrix} \boldsymbol{x}^{t}_{p_1} - \boldsymbol{x}^{b}_{p_1} \\ \ln \boldsymbol{x}^{t}_{q_1} - \ln \boldsymbol{x}^{b}_{q_1} \end{pmatrix} \\
&+ \left\langle \begin{pmatrix} \boldsymbol{x}^{t}_{p_1} - \boldsymbol{x}^{b}_{p_1} \\ \ln \boldsymbol{x}^{t}_{q_1} - \ln \boldsymbol{x}^{b}_{q_1} \end{pmatrix}, \begin{pmatrix} \boldsymbol{0}_{p_1} \\ \boldsymbol{1}_{q_1} \end{pmatrix} \right\rangle \\
&+ \frac{1}{2}\begin{pmatrix} \boldsymbol{y}_{p_2} - \boldsymbol{h}_{p_2}(\boldsymbol{x}) \\ \ln \boldsymbol{y}_{q_2} - \ln \boldsymbol{h}_{q_2}(\boldsymbol{x}) \end{pmatrix}^{T} \boldsymbol{R}^{-1}_{mx} \begin{pmatrix} \boldsymbol{y}_{p_2} - \boldsymbol{h}_{p_2}(\boldsymbol{x}) \\ \ln \boldsymbol{y}_{q_2} - \ln \boldsymbol{h}_{q_2}(\boldsymbol{x}) \end{pmatrix} \\
&+ \left\langle \begin{pmatrix} \boldsymbol{y}_{p_2} - \boldsymbol{h}_{p_2}(\boldsymbol{x}) \\ \ln \boldsymbol{y}_{q_2} - \ln \boldsymbol{h}_{q_2}(\boldsymbol{x}) \end{pmatrix}, \begin{pmatrix} \boldsymbol{0}_{p_2} \\ \boldsymbol{1}_{q_2} \end{pmatrix} \right\rangle . \qquad (10)
\end{aligned}
$$

Application of mixed Gaussian-lognormal Distribution

In Kliewer et al. [2016] the mixed distribution was introduced in to the CIRA 1-Dimensional Optimal Estimator (C1DOE) by Dr. Anton Kliewer to asses the impact on retrieving temperature and mixing-ratio values from microwave brightness temperatures. It was assumed that the observational errors were Gaussian distributed, whilst for the background errors it was assumed that errors for temperature were Gaussian distributed, and those for the mixing-ratio were lognormally distributed.

An important finding from Kliewer et al. [2016] was that by using a lognormal model for the mixing-ratio enables us to fit better to the temperature channels, so called O-A statistics.

Results from Kliewer et al (2016)

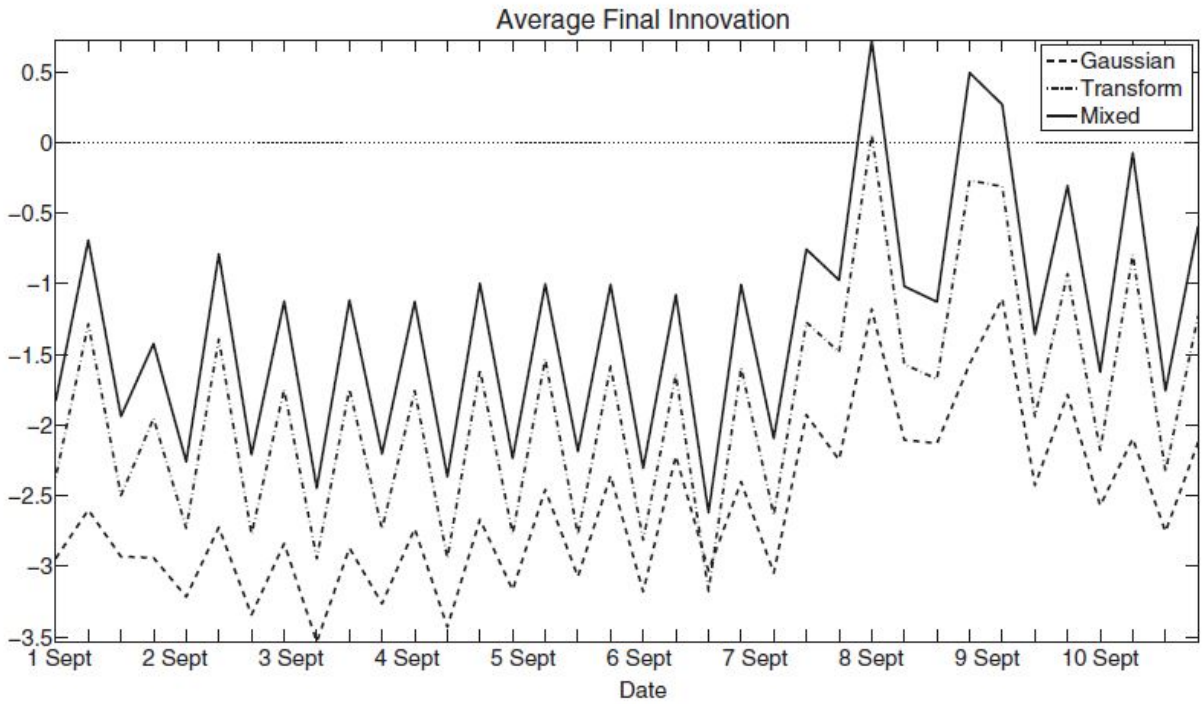


Figure 10. Similar to Figure 9: the average final innovation for temperature in the troposphere (AMSU-A channel 6, 54.4 GHz) for each retrieval. This is the average of all innovations for all points in the last Newton–Raphson iteration. The mixed distribution clearly has the smallest differences between the retrieved state and the observation. The noted oscillation is due to limb darkening via varying zenith angles.

Figure 2: O-A statistics from Kliewer et al. [2016].

CIRA Data Assimilation Testbed: CDAT

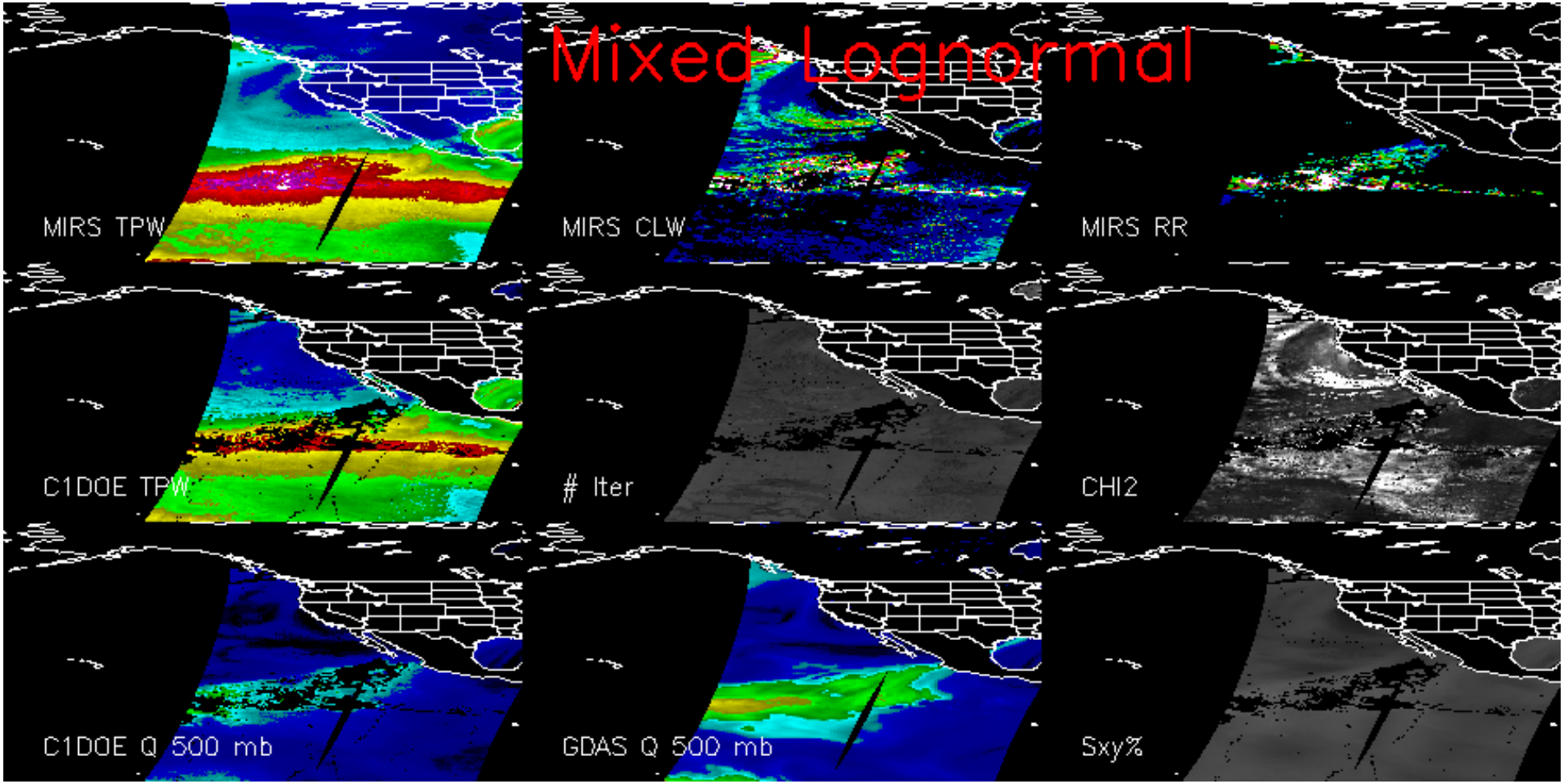


Figure 3: Sample of the output from CDAT. https://cdat.cira.colostate.edu

Is the mode the best statistic?

Whilst undertaking the research for Kliewer et al. [2016] we used a test case where we created a fake brightness temperature from a specific first guess and added some small and larger errors to check if the retrieval worked as we thought, but it did not and when we gave it the correct first guess the mode was worse the the logarithmic transform (median). But why?

Upon coding up an one variable equivalent of C1DOE it became clear that if the aprori state is within a bound of the true state, then the median was the best minimiser. However, if this was not the case then there were regions where either the mode or the mean could minimise the errors.

Fletcher et al. [2019]

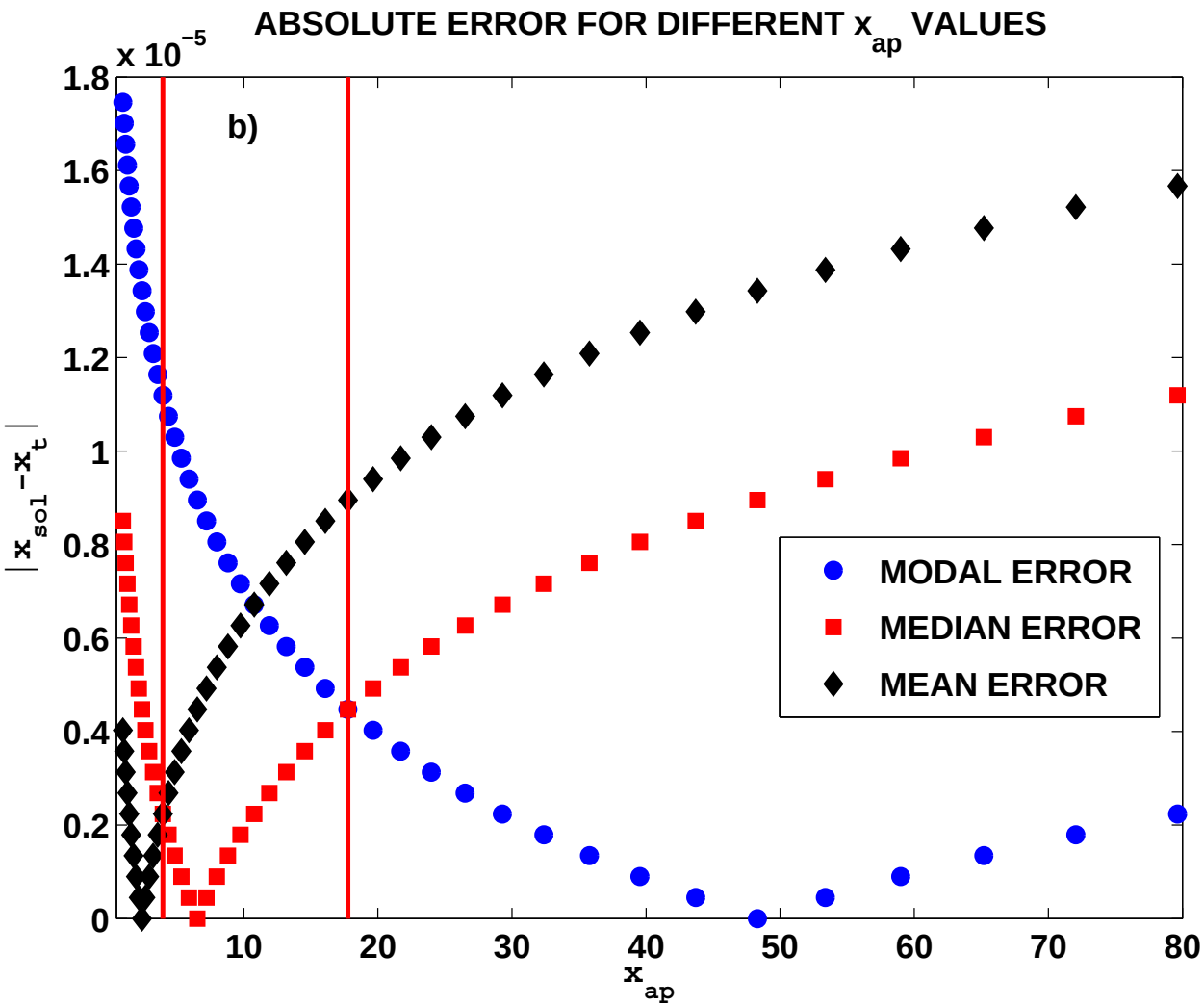


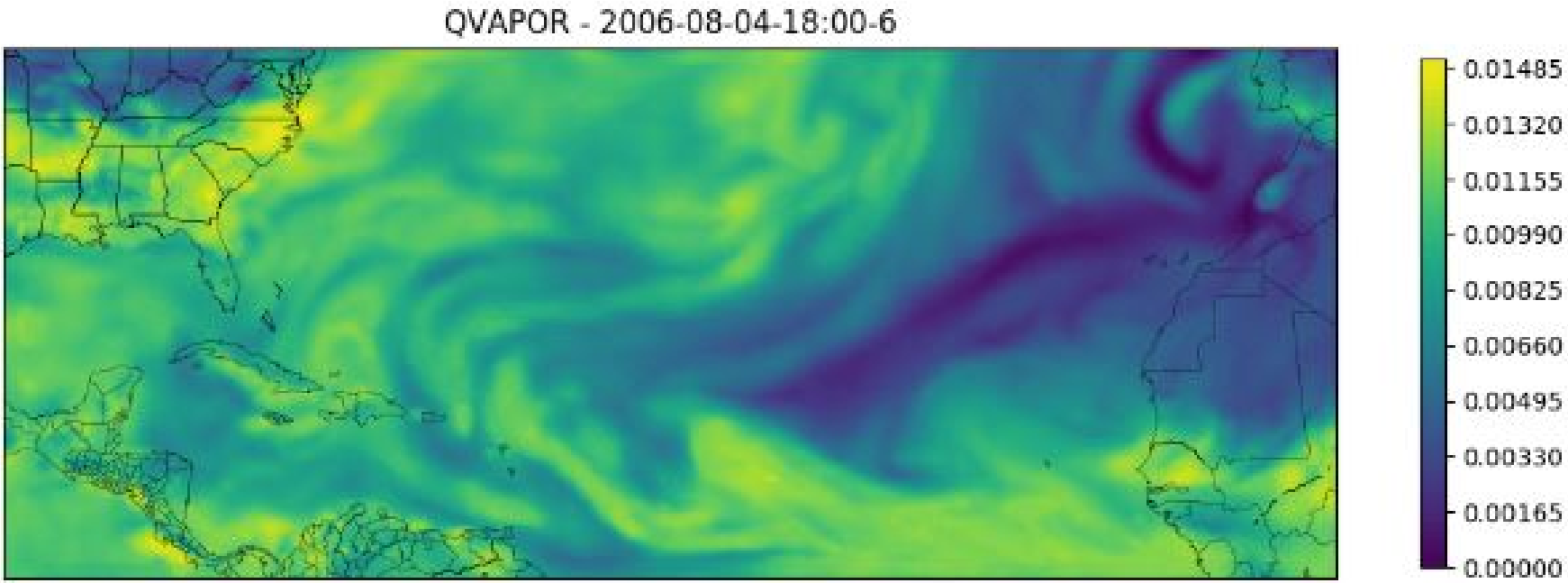
Figure 4: Results from Fletcher et al. [2019].

Outline

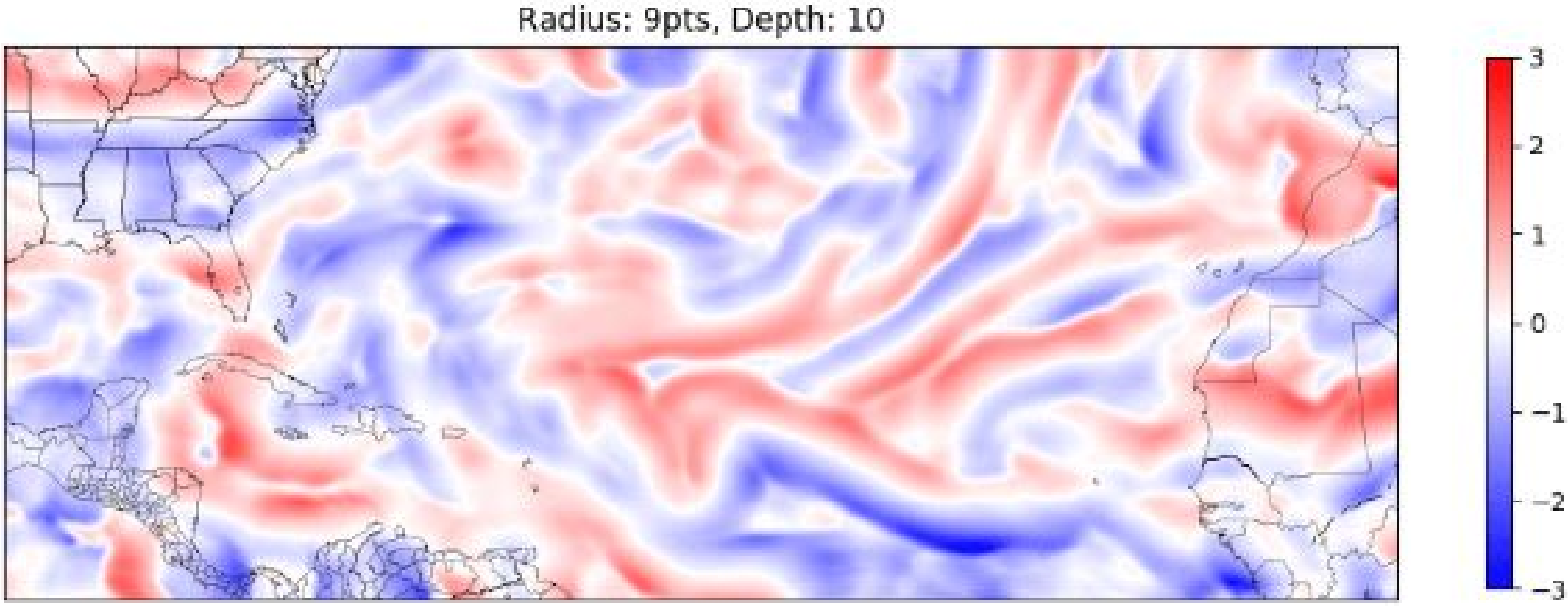
Summary of non-Gaussian developments at CIRA

- Full Field mixed 4DVAR, Fletcher [2010].
- Incremental mixed VAR, Fletcher and Jones [2014].
- Mixed Gaussian-lognormal Kalman filter, Fletcher et al. [2023b]
- Lognormal and mixed Gaussian-lognormal based Buddy check system for observational quality control, Fletcher et al. [2023a]

Is lognormal the only other distribution for moisture?



Is lognormal the only other distribution for moisture?



The answer is no, and the plots above show that there is a negative-skewed distribution present here. We have determined this to be a reverse lognormal distribution.

Summary of Reverse Lognormal DA

The reverse lognormal distribution is part of the 3-parameter family of lognormal distributions, Foster et al. [2006]. Its definition is similar to that of the lognormal distribution, but now has an upper bound ξ. Thus the reverse lognormal distribution is defined on $(-\infty, \xi)$ and is given by

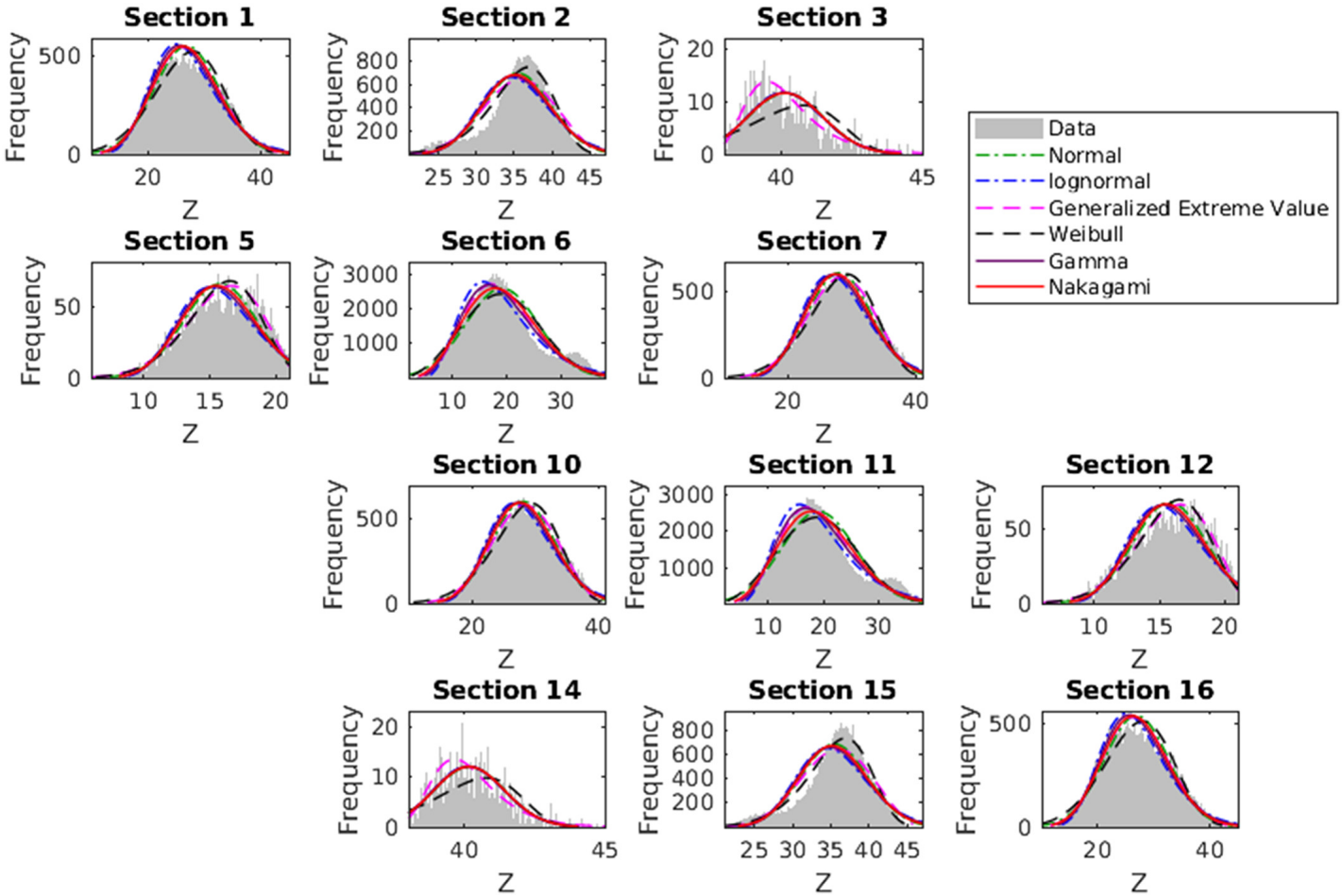
$$P(x) \equiv \frac{1}{(\xi - x)\,\sigma\sqrt{2\pi}} \exp\left\{ -\frac{(\ln(\xi - x) - \mu)^2}{\sigma^2} \right\}. \tag{11}$$

It is possible to define a mixed distribution that contains Gaussian, lognormal, and reverse lognormal random variables, Fletcher [2022] and then derive a 3DVAR cost function, Goodliff et al. [2023], as well as a Kalman filter like assimilation scheme, Van Loon and Fletcher [2023].

Outline

Which Distribution to Use?

In Goodliff et al. [2020] we use a support vector machine (SVM) and a neural network (NN) approach to determine whether or not it is possible to detect and predict a change in the distribution in the Lorenz 1963 model, Lorenz [1963]. Below is a copy of figure 5 from Goodliff et al. [2020].



When to use which distribution?

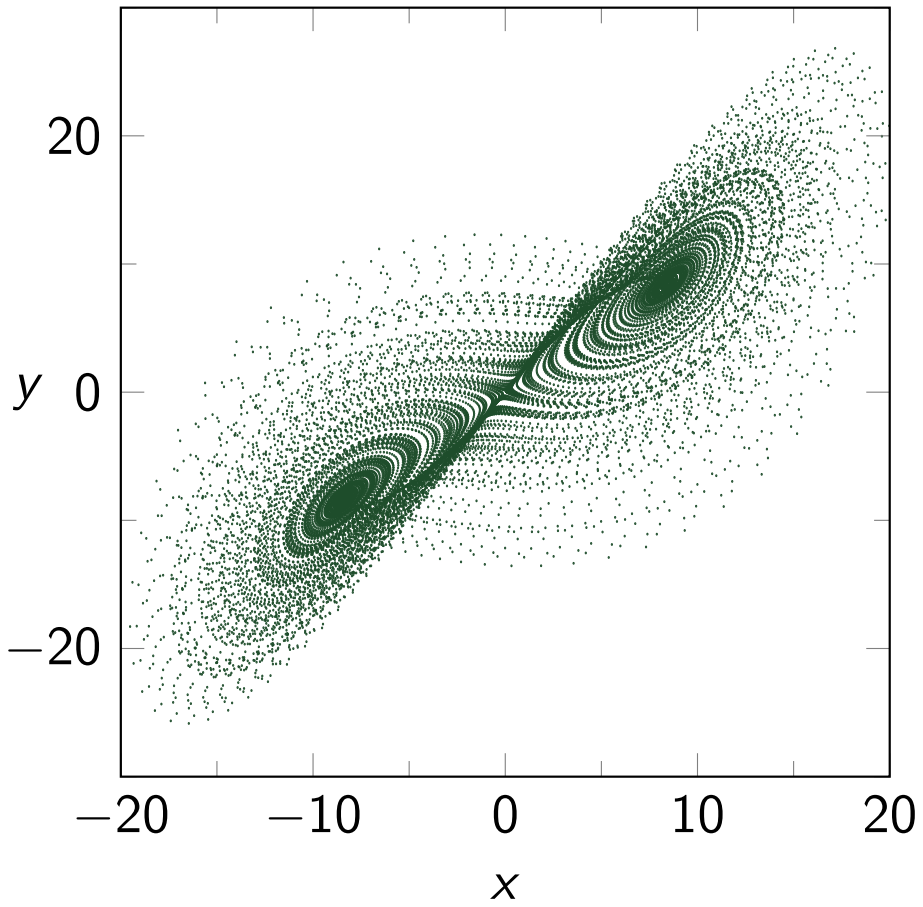
There are two approaches that were considered to detect the change of distribution in Goodliff et al. [2020]: 1) Detect that the sample mean and mode are not equal, 2) Detect a positive of negative skewness in the sample. It is the latter approach that has proven the most successful.

The SWM approach was implemented for the Gaussian and lognormal approaches for 3DVAR with the L63 model in Goodliff et al. [2022] with a study on performance relative to different times between observations and sample size to determine the detection of skewness.

Work undertaken by Dr. Jakir Hossen indicated that a third machine learning technique, K-Nearest Neighbors (KNN) performed better at predicting the change in distribution in the L63 model than the other two techniques.

K-Nearest Neighbours

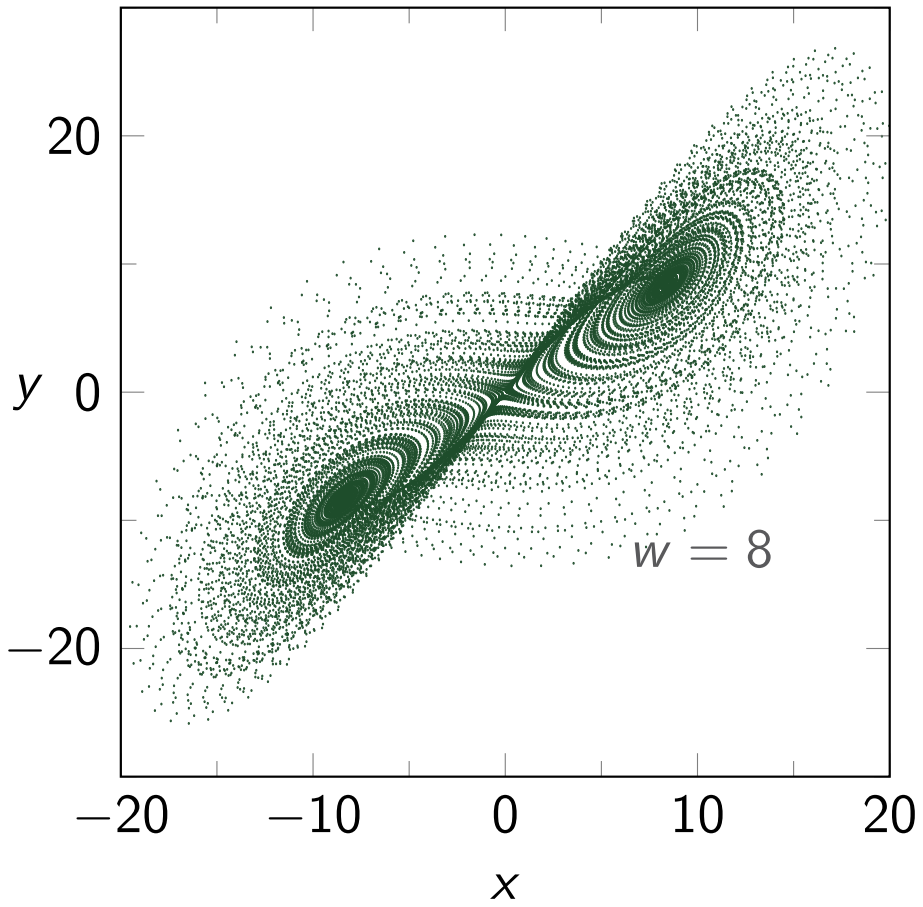
- Generate long run of Lorenz-63 model



K-Nearest Neighbours

- Generate long run of Lorenz-63 model
- Test skewness of z over some time period

$$s(t_k) = \text{skewtest}\{z(t_{k-w}), \ldots, z(t_{k+w})\}$$



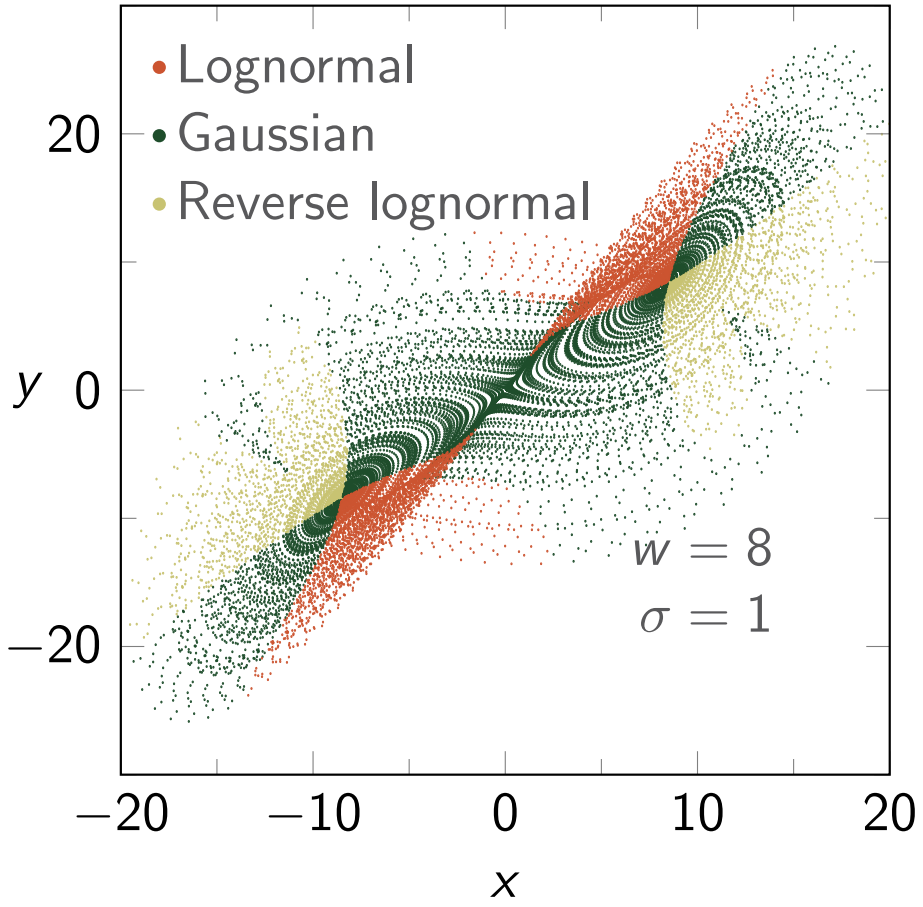
K-Nearest Neighbours

- Generate long run of Lorenz-63 model
- Test skewness of z over some time period

$$s(t_k) = \text{skewtest}\{z(t_{k-w}), \ldots, z(t_{k+w})\}$$

- Divide into three bins

$$\begin{aligned} s &\geq \sigma && \text{Lognormal,} \\ -\sigma < s &< \sigma && \text{Gaussian,} \\ s &\leq -\sigma && \text{Reverse lognormal.} \end{aligned}$$



K-Nearest Neighbours

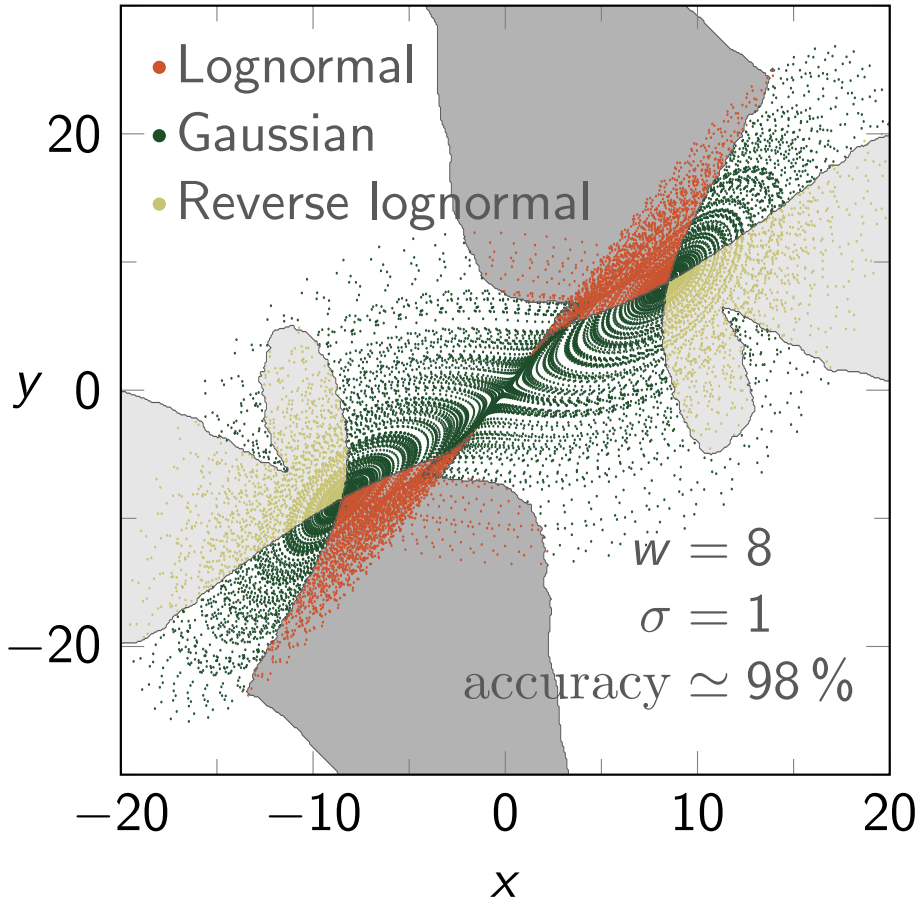
- Generate long run of Lorenz-63 model
- Test skewness of z over some time period

$$s(t_k) = \text{skewtest}\{z(t_{k-w}), \ldots, z(t_{k+w})\}$$

- Divide into three bins

$$\begin{aligned} s &\geq \sigma && \text{Lognormal,} \\ -\sigma < s &< \sigma && \text{Gaussian,} \\ s &\leq -\sigma && \text{Reverse lognormal.} \end{aligned}$$

- Train k-nearest neighbor classifier on $\{x(t_k), y(t_k)\}$ as input



Optimal Training Data

Table 1: Performance of ML techniques with different length of data (two-third used for training), from Goodliff et al. [2023].

Methods	10000	20000	30000	36000	50000	100000
			Radius 4			
KNN	89.9%	92.7%	92.2%	94.0%	94.5%	94.6%
NN	85.5%	89.4%	87.9%	92.4%	91.4%	90.5%
SVM	85.4%	84.0%	84.6%	86.3%	84.6%	85.1%
			Radius 8			
KNN	89.9%	92.2%	93.3%	95.2%	95.2%	96%
NN	88.2%	92.7%	90.1%	92.8%	93.9%	93.0%
SVM	85.7%	85.0%	86.2%	88.3%	86.8%	87.8%
			Radius 12			
KNN	92.0%	92.9%	94.9%	96.1%	96.0%	97.1%
NN	90.9%	91.4%	92.0%	93.8%	92.8%	93.4%
SVM	85.5%	86.3%	87.7%	88.6%	89.1%	89.9%

Optimal Training Data, Goodliff et al. [2023]

- We consider three different radii: 4, 8 & 12 time steps i.e 9, 17 & 25 points around the current point of which skew value is calculated.
- After training the model with the training data set, we use spatial coordinates of testing data set to predict the skewvalue s_p corresponding to the z-component in the testing data set s_t.
- Using s_t and s_p, we calculate the accuracy (%).
- We found that the accuracy does not change significantly after 36000 time steps for all the three radii. Using a higher number of time steps (e.g. 50000 and 100000), the accuracy rather decreases for the case of NN and SVM.
- Using a training trajectory of 36000 time steps, the KNN method provides accuracy of 94%, 95% and 96% for 4, 8 & 12 time step radius, respectively.
- Similarly, the NN method provides the accuracy of 92%, 93% and 94% and SVM provides 86%, 88% & 88%.

Outline

Conclusions and Further Work

- Have developed many forms of lognormal and reverse-lognormal based variational and Kalman filter based data assimilation systems.
- Comparing different ML techniques to determine which distribution the z component of the Lorenz 63 model is following, we have seen that the K-Nearest Neighbour approach appears to be optimal at predicting which version of the cost function, or Kalman filter to use to minimise the analysis error.
- Have shown that 36000 time steps of the Lorenz 63 model appears to be the optimal time needed to train the KNN for this model and that it is a waste of resources to go beyond this.
- We have been developing a lognormal and a reverse lognormal version of the MLEF.
- Have extended the buddy check observational quality control to lognormal and mixed Gaussian-lognormal formulations, currently working on the reverse lognormal version.

References I

S. E. Cohn. An introduction to estimation error theory. *Journal of the Meteorological Society of Japan*, 75:257–288, 1997.

S. J. Fletcher. Mixed lognormal-Gaussian four-dimensional data assimilation. *Tellus*, 62A:266–287, 2010.

S. J. Fletcher. *Data Assimilation for the Geosciences: From Theory to Applications, 2nd Edition*. Amsterdam, Elsevier, 2022.

S. J. Fletcher and A. S. Jones. Multiplicative and additive incremental variational data assimilation for mixed lognormal-Gaussian errors. *Mon. Wea. Rev.*, 142: 2521–2544, 2014.

S. J. Fletcher and M. Zupanski. A data assimilation method for log-normally distributed observational errors. *Quart. J. Roy. Meteor. Soc.*, 132:2505–2519, 2006a.

S. J. Fletcher and M. Zupanski. A hybrid normal and lognormal distribution for data assimilation. *Atmospheric Science Letters*, 7:43–46, 2006b.

S. J. Fletcher and M. Zupanski. Implications and impacts of transforming lognormal variables into normal variables in VAR. *Meteor. Z.*, 16:755–765, 2007.

References II

S. J. Fletcher, A. J. Kliewer, and A. S. Jones. Quantification of optimal values for the parameters in lognormal variational data assimilation and their chaotic effects. *Mathematical Geosciences*, 51:187–207, 2019.

S. J. Fletcher, S. Van Loon, M. R. Goodliff, A. J. Kliewer, A. S. Jones, and J. M. Forsythe. Lognormal and mixed gaussian-lognormal observational quality control measures for data assimilation methods. *To be submitted to Q. J. R. Meteor. Soc.*, 2023a.

S. J. Fletcher, M. Zupanski, M. R. Goodliff, A. J. Kliewer, A. S. Jones, J. M. Forsythe, T.-C. Wu, M. J. Hossen, and S. Van Loon. Lognorml and mixed Gaussian-lognormal Kalman filters. *Mon. Wea. Rev:*, 151:761–774, 2023b.

J. Foster, M. Bevis, and W. Raymond. Precipitable water and the lognormal distribution. *J. G. R. Atmospheres*, 111:D15102, 2006.

M. R. Goodliff, S. J. Fletcher, A. J. Kliewer, J. M. Forsythe, and A. S. Jones. Detection of non-Gaussian behavior using machine learning techniques: A case study on the Lorenz 63 model. *J. Geophys. Res.: Atmospheres*, 125(2): e2019JD031551, 2020.

References III

M. R. Goodliff, S. J. Fletcher, A. J. Kliewer, A. S. Jones, and J. M. Forsythe. Non-gaussian detection using machine learning with data assimilation applications. *Earth and Space Sciences*, 9:e2021EA001908, 2022.

M.R. Goodliff, M. J. Hossen, S. Van Loon, and S. J. Fletcher. Non-gaussian data assimilation: Reverse lognormal. *Submitted to Q. J. R. Meteor. Soc.*, 2023.

A. J. Kliewer, S. J. Fletcher, A. S. Jones, and J. M. Forsthye. Comparison of Gaussian, logarithmic transform and mixed distribution Gaussian-lognormal distribution based 1DVAR microwave temperature-water vapour mixing ratio retrievals. *Quart. J. Roy. Meteor. Soc.*, 142:274–286, 2016.

A. C Lorenc. Analysis methods for numerical weather prediction. *Q. J. R. Meteor. Soc.*, 112:1177–1194, 1986.

E. N. Lorenz. Deterministic Nonperiodic Flow. *J. Atmos. Sci.*, 20(2):130–141, 1963.

S. Van Loon and S.J. Fletcher. Dynamic Gaussian, lognormal, and reverse lognormal kalman filter. *Submitted to Q. J. R. Meteor. Soc.*, 2023.