# Experiments with variational DA in L63

NCEO Intensive Course on Data Assimilation

University of Reading, 2022

The objective of this practical is to perform state estimation using 3DVar and 4DVar in the Lorenz 1963 system.

# 1 Description of the code

## 1.1 Files

These are the python files used in this part of the activity:

- *ControlL63Var.py*. This is the control file - it is the one that you will be running and modifying.
- *L63model.py*. This file contains the instructions for running the L63 model.
- *L63misc.py*. This file generates different observation operators, creates the observations, and generates a simple background error covariance matrix.
- *L63var.py*. This file contains the routines to perform 3D-Var and strong constraint 4D-Var. This includes computing the tangent linear model and transition matrices.
- *L63plots.py*. This file contains the code for producing different output plots.

## 1.2 Instructions

You will run different sections of the file *ControlL63Var.py*. These are numbered as comments of the file (recall that in python # is used for comments). To run **only** a subsection of a file you can highlight the desired lines of code with the mouse, and then press F9.

- The first lines of the file import the different packages that the file uses: numpy, matplotlib, and the functions we have created for this activity.
- ***Section 1: The Nature Run***. This section generates the **nature** run for the experiment, i.e. what we consider to be the true system trajectory. You can change the initial conditions for the nature run, the final time (the model time step is fixed at 0.01 time units), and the initial background guess from which the assimilation will start. You can also play with the 3 parameters of the model to see how the behaviour of the system changes for different combinations of values. However, for the final experiment you should leave their values at $\boldsymbol{\theta} = (10, 8/3, 28)$. Running this section should also plot the 3D phase space (time is implicit in this figure), and time evolution plots for each of the 3 model variables.

- ***Section 2: The observations***. This section generates the observations, the observation error covariance matrix and the observation operator. You can select to observe different combinations of variables by changing the value of *'obsgrid'*: setting $obsgrid = 'xyz'$ will observe all variables, or you could choose a subset e.g. $'xz'$ or $'y'$. Different choices will create the appropriate observation operator. The **R** matrix is designed to be diagonal (common assumption), but you can choose the observation error variance. You can also choose the frequency of the observations (in number of model steps). As a rule of thumb, observations every 8 steps yield a quasi-linear problem, whereas observations every 25 steps yield a fully non-linear problem.

- ***Section 3: Data assimilation***. This short section creates the climatological **B** matrix for this model. There is a scaling tuning parameter that can be varied depending on the observation frequency.

- ***Section 3a: 3DVar***. This section runs the 3D-Var assimilation. It also computes the background and analysis RMSE (root mean squared error) with respect to the truth. The trajectories and the RMSE's are plotted.

- ***Section 3b: 4DVar***. This section runs the 4D-Var assimilation. You can select the length of the assimilation window, which is expressed in terms of the number of observation times per window. It also computes the background and analysis RMSE with respect to the truth. The trajectories and the RMSE's are plotted.

# 2 Exercises

1. Run the Lorenz 1963 model to a maximum time $t = 10$ and plot the trajectory. This nature run will be the basis of our experiments. Generate a set of synthetic observations with the settings described next.

2. 3DVar. In this case the following parameters can vary: the frequency of observations in time (in model steps), the density of observations in space, the observational error covariance (here we consider it to be diagonal with the same variance for the all observed variables), the background error covariance (for now we let it be fixed as obtained by the code). Perform experiments with the combinations indicated in the following table:

| Obs frequency | Obs density | Obs std |
|---|---|---|
| 10 | xyz | 1 |
| ”” | ”” | $\sqrt{2}$ |
| ”” | xz | 1 |
| ”” | x | 1 |
| 20 | xyz | 1 |
| ”” | ”” | $\sqrt{2}$ |
| ”” | ”” | $\sqrt{2}$ |
| 20 | z | 0.2 |

How does the observational frequency influence the estimation? What about the observational density? Are all variables of this model equally 'informative'?

3. 4DVar. In this case there is another parameter to vary: the length of the assimilation window. In this exercise it expressed as the number of observational times per window. Try the following numbers: $\{1, 2, 4, 6\}$ for some of the configurations of the previous exercise. What is the difference between 3DVar and 4DVar with only 1 observational time? How does the performance change as you include more and more observational times? Is this a steady behaviour?