

Using the EnKF in L63

NCEO Intensive Course on Data Assimilation

University of Reading, 2022

1 Summary

The objective of the practical is to perform state estimation and state/parameter estimation in the Lorenz 1963 system. In this small (3-variable) model, localisation is not required.

2 Description of the code

2.1 Files

These are the python files used in this activity.

- *ControlL63Enkf.py*. This is the control file. You will run and modify this file.
- *L63model.py*. This file contains all the instructions to run the L63 model.
- *L63misc.py*. This file generates different observation operators, creates the observations, and generates a simple background error covariance matrix.
- *L63kfs.py*. This file contains the routines to perform SEnKF and ETKF.
- *L63plots.py*. This file has instructions for different plots.

2.2 Instructions

You will run different sections of the file *ControlL63EnKF.py*. These are enumerated as comments of the file (recall that in python `#` is used for comments). To run **only** a section of a file you can highlight the desired instructions 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*. This section generates the **nature** run of the experiment, i.e. what we consider to be the true system. You can change the initial conditions, the final time (consider that the model time step is 0.01 time units), and the initial guess from which the assimilation will start. You can also play with the parameters of the model to see how the behaviour of the system changes. However, for the final experiment you should leave this at $\theta = (10, 8/3, 28)$. Running this section should also plot the 3D phase space (time is implicit in this figure), and time evolution plots for each one of the variables.

- *Section 2.* This section is related to the observations. You can select to observe different variables: e.g. all of them: 'xyz', or subsets: 'xz' or 'y'. Different choices will create the observation matrix. The \mathbf{R} matrix is designed to be diagonal (common assumption), but you can choose the observational variance. You can also choose the observational period (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 full non-linear problem.
- *Section 3.* In this section you perform the DA experiments using the Kalman filets. You can vary the number of ensemble members M , the inflation ρ , and the method: either 'SEnKF' or 'ETKF'.
- *Section 4.* In this section you can perform parameter estimation. You can vary the initial 'guess' value of the parameters (parambad), as well as the variance of the random evolution of the parameters in the forecast (α).

Note that in this exercise a new metric is computed along the RMSE. This is the spread of the analysis ensemble. In a healthy ensemble system these two should coincide (in average).

3 Experiments

3.1 State estimation

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 synthetic set of observations with some of the options described next.
2. State estimation 1. In the Lorenz 1963 system we only have 3 variables, so we can easily afford to run more ensemble members than variables (a luxury we seldom have in real life!). There is not a well-defined 'physical distance' between these variables, so the concept of localisation does not apply. For this first experiment let us fix the observational standard deviation $\sqrt{2}$, and vary the following: the observational frequency, the observation density, the ensemble size. Select SEnKF as method.

Obs frequency	Obs density	Ensemble size
10	xyz	100
""	""	10
""	yz	100
""	""	10
20	xyz	100
""	""	10
""	y	20
""	z	100

How does the observational frequency influence the estimation? What about the observational density? What about the ensemble size? Repeat the experiments (or some of them) with SEnKF and compare the results.

3. State estimation 2. Now let us try a more challenging setting by limiting our ensemble size to $N_e = 3$, using ETKF first and then SEnKF. In this case it is possible you will need inflation to make the filter work. Suggestion: try values of $\mathcal{O}(10^{-2})$ for frequent observations, and $\mathcal{O}(10^{-1})$ for infrequent observations

Obs frequency	Obs density
10	xyz
""	y
""	z
20	xyz
""	y
""	xz

How are the results in this case? What happens as you increase the inflation values? Can you always find an inflation value that ensures a good result?

3.2 Joint state/parameter estimation

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 observations with covariance $\mathbf{R} = \mathbf{I}$, observing only y and z .
2. State and parameter estimation 1. In this exercise you will estimate both the 3 variables of the L63 system, as well as the 3 parameters of the system. Let us fix some parameters: the sample size to $M = 10$ and the method to ETKF. Perform the following experiments.

Initial param values	ρ	α
[6.0,3.0,25.0]	0.01	0
["", "", ""]	0.1	0.1
["", "", ""]	0.01	0
["", "", ""]	0.1	0.1
[0.3,1.0,-1.0]	0.01	0
["", "", ""]	0.1	0.1
["", "", ""]	0.01	0
["", "", ""]	0.1	0.1

What can you conclude about this experiments? Parameter estimation is a powerful too, but how much does the starting initial condition matter? What is the importance of α ?