

DA ingredients practical

Instructions

All code and links can be found on the Github repository.

In part 1 of the practical, we will use streamlit to look at the role of the error covariances in a simple 1D system.

In part 2 we will use Jupyter notebooks to look at the DA ingredients needed to assimilate data into the toy Lorenz 96 model.

streamlitApp

Controls

Background function shape

Flat

First observation location

0 4

Second observation location

5 10

Observation error standard deviation

0.00 1.00

Background error standard deviation

0.00 1.00

Background error correlation length scale

0.00 2.00

Observation error correlation length scale

0.00

Optimal interpolation demonstration

Introduction

In this interactive demo, we perform data assimilation in order to estimate the values of a function over the interval $[0,10]$. A simple data assimilation scheme, called *optimal interpolation*, is used. This computes an updated estimated of the state given a first guess (the background), \mathbf{x}_b , observations, \mathbf{y} , and the mapping from the state variables to those observed, $h(\cdot)$.

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - h(\mathbf{x}_b)).$$

\mathbf{K} controls the weighting given to the observations versus the background and is a function of the error covariance matrices for the background, \mathbf{B} , and the observations, \mathbf{R} , as well as the Jacobian of the observation operator, \mathbf{H} .

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}.$$

The exercises explore two ideas:

1. the effect of the uncertainty in the observations and background on the analysis.
2. how the information from observations is spread to the unobserved variables.

Instructions

Use the selection boxes and sliders in the left-hand control panel to adjust parameter values. The plot will automatically update, although you might have to wait a few seconds after changing a parameter value.

Share ☆ ↻ ⋮

jupyter Control_DA_components Last Checkpoint: 18 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

Introduction

Over the next few days we will be exploring the use of different DA algorithms for estimating and forecasting the state of a system. To enable this, in today's practical we will begin by setting up an idealised system in which the true underlying dynamics are described by the Lorenz 96 model. From the true dynamics we will then generate a background and observations consistent with their assumed error characteristics. This approach is known as identical twin experiments and allows us to fully evaluate the success of different DA algorithms.

First of all we need to import the functions needed.

```
In [2]: import numpy as np
from tools.L96_model import lorenz96
from tools.obs import createH, gen_obs
from tools.cov import getBcanadian
from tools.plots import plotL96, plotL96obs, plotL96_Linerr, plotH, tileplotB
from tools.diag import compute_lin_error
```

The Lorenz 96 model

The Lorenz 96 model describes the simplified evolution of a univariate large-scale atmospheric system on a one dimensional circle of latitude; simulating external forcing, internal dissipation and advection. It is given by the following equations

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F \quad \text{for } i = 1, 2, \dots, n$$

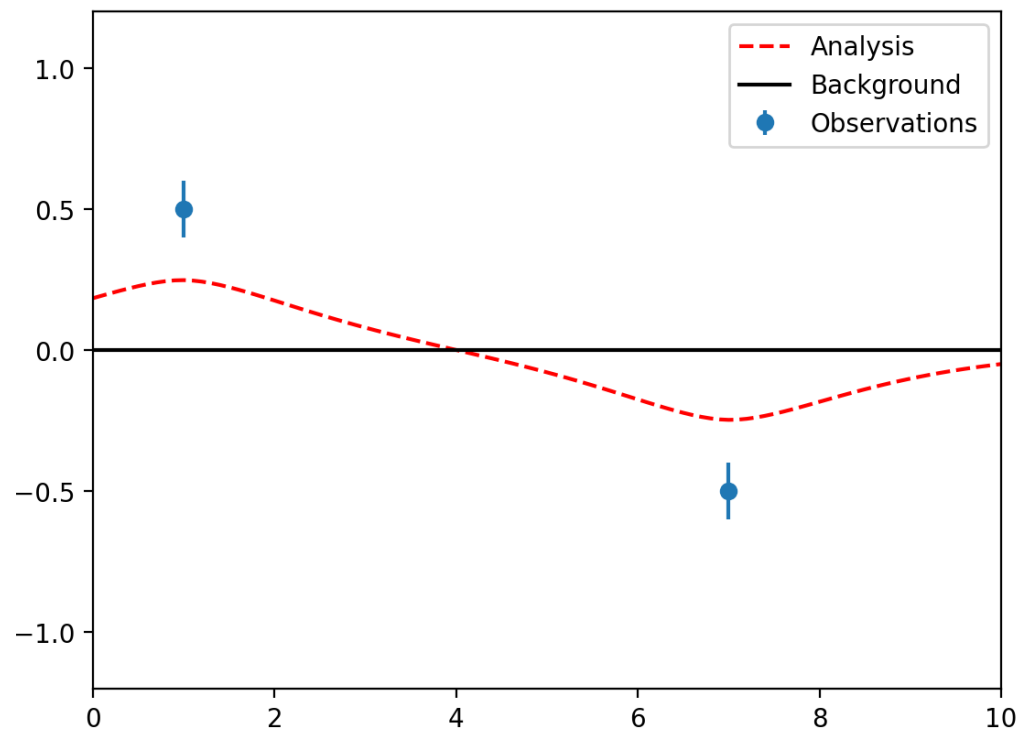
where x_i is the state variable at position i , and F is a forcing term.

Running the next section of code will produce a Hovmoller plot of the Lorenz 96 model when $n = 12$ and $F = 8$. What structures are identifiable within this plot?

```
In [4]: n = 12 # No. of state variables (No. of nodes in Lorenz-96)
F = 8.0 # Forcing F>5.0 guarantees chaos
deltat = 0.025 # Model timestep, 0.025 equivalent to approximately 3 hours
tf = 5.0 # The final time of the simulation
```

Streamlit: Optimal interpolation demonstration

The effect of the uncertainty in the observations and background on the analysis



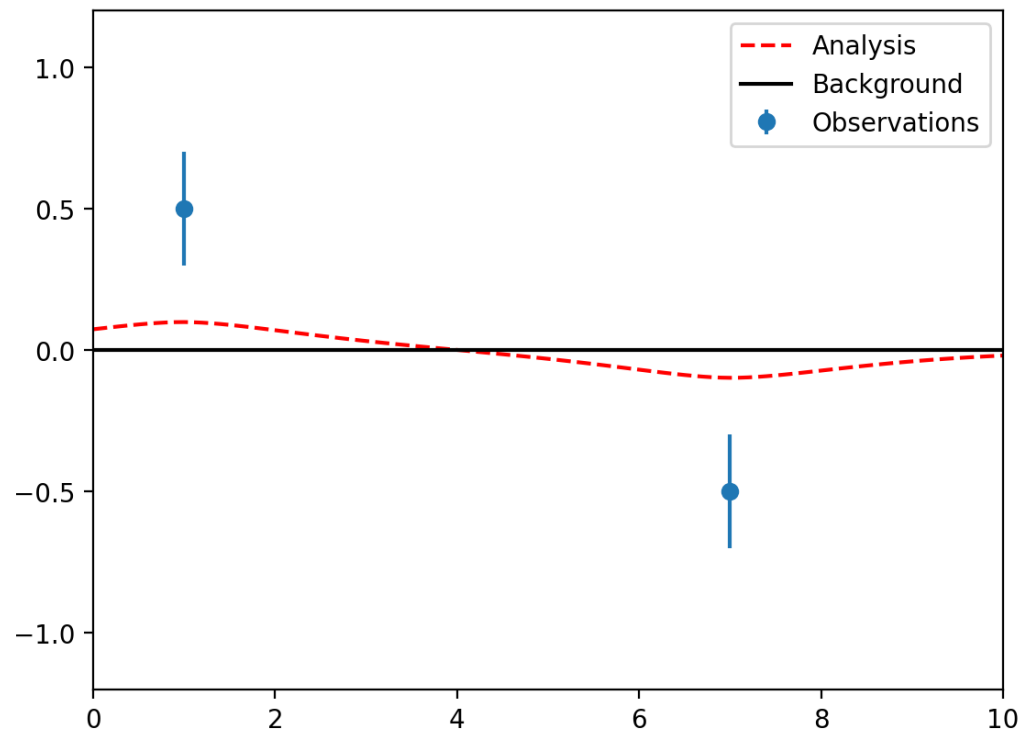
Parameters:

Observation error standard deviation 0.1

Background error standard deviation 0.1

Streamlit: Optimal interpolation demonstration

The effect of the uncertainty in the observations and background on the analysis



Parameters:

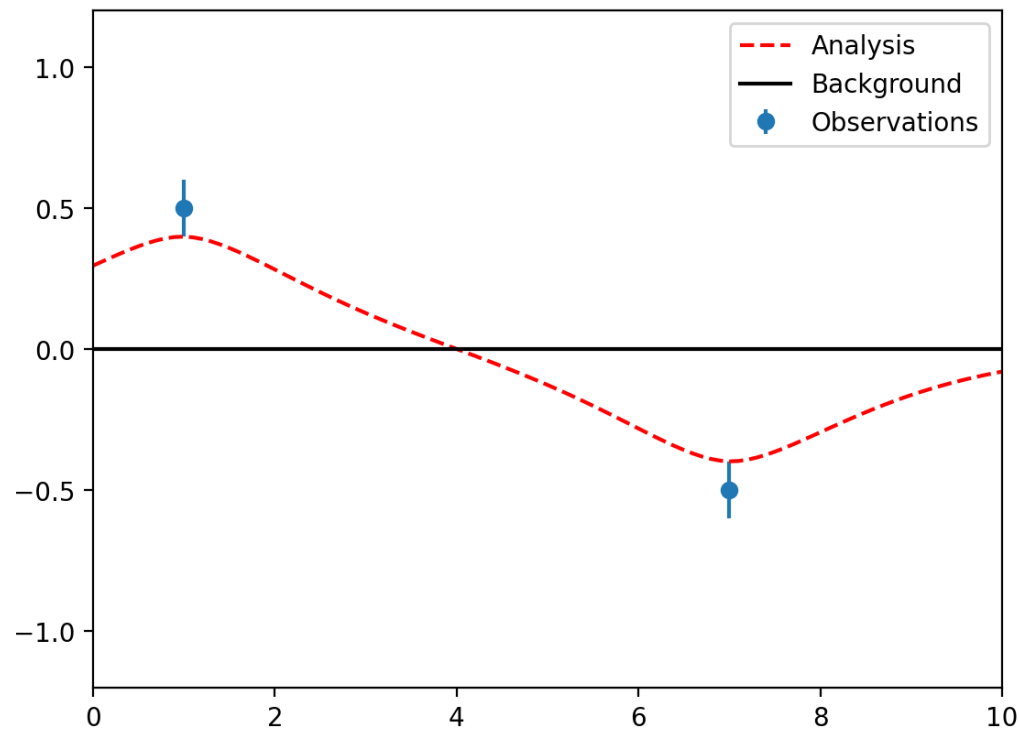
Observation error standard deviation 0.2

Background error standard deviation 0.1

Increasing the observation error standard deviation reduces the analysis fit to the observations

Streamlit: Optimal interpolation demonstration

The effect of the uncertainty in the observations and background on the analysis



Parameters:

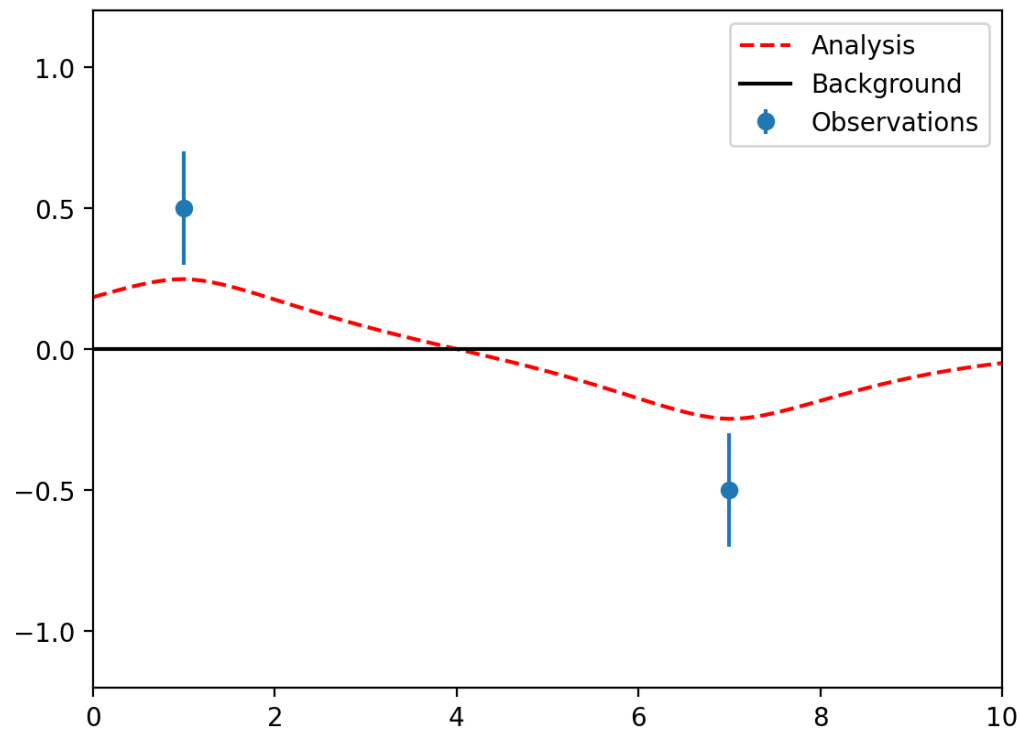
Observation error standard deviation 0.1

Background error standard deviation 0.2

Increasing the background error standard deviation increases the analysis fit to the observations

Streamlit: Optimal interpolation demonstration

The effect of the uncertainty in the observations and background on the analysis



Parameters:

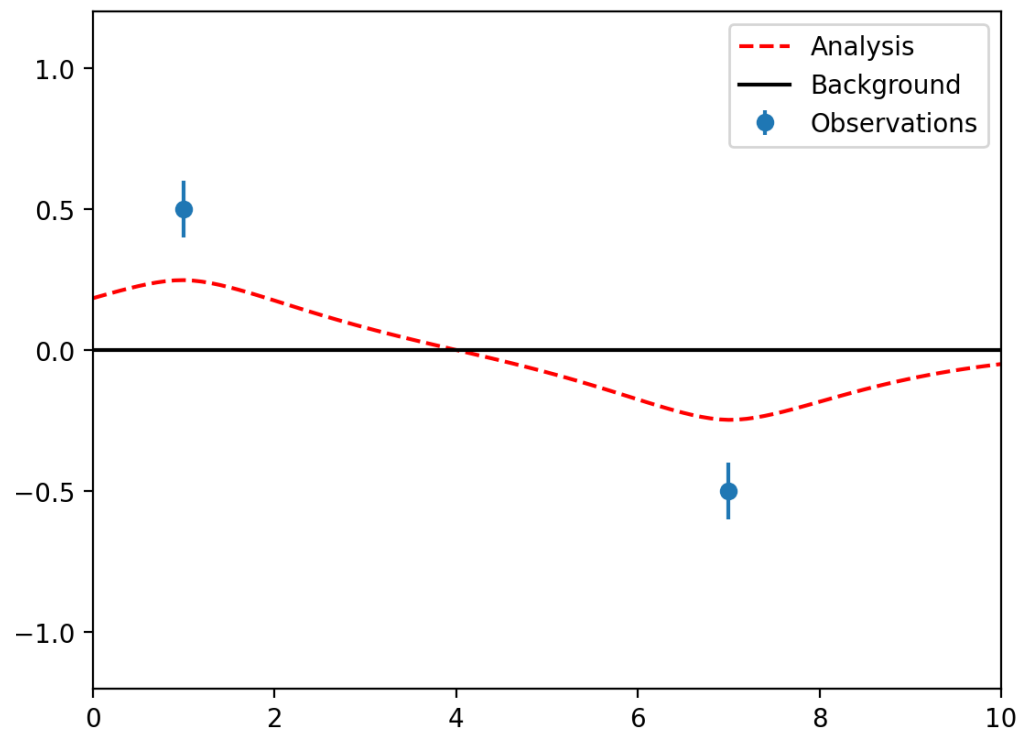
Observation error standard deviation 0.2

Background error standard deviation 0.2

It is the ratio of the observation to background error standard deviation that is important

Streamlit: Optimal interpolation demonstration

The effect of the uncertainty in the observations and background on the analysis



Parameters:

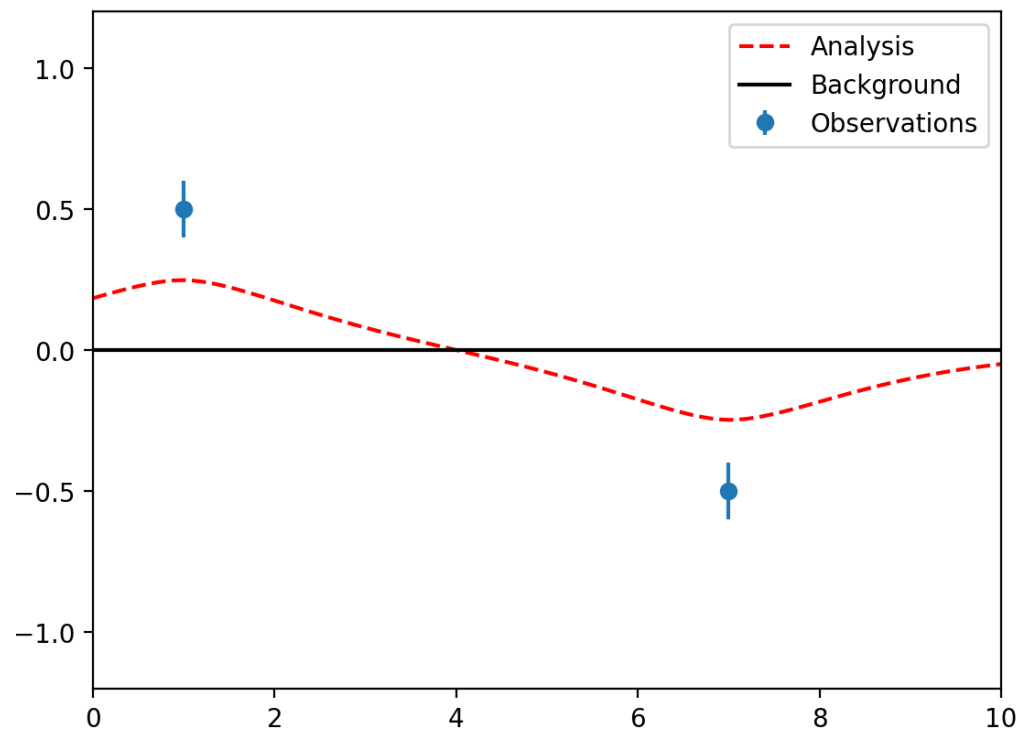
Observation error standard deviation 0.1

Background error standard deviation 0.1

It is the ratio of the observation to background error standard deviation that is important

Streamlit: Optimal interpolation demonstration

How the information from observations is spread to the unobserved variables: the role of the background error correlations

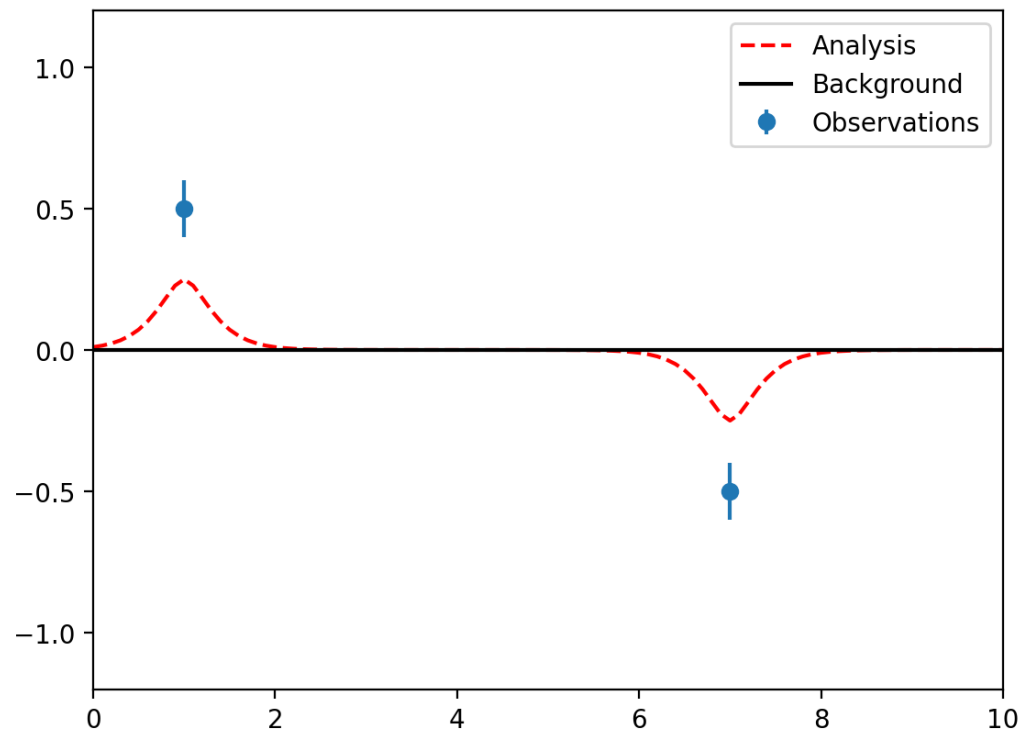


Parameters:

Observation error standard deviation 0.1
Background error standard deviation 0.1
Background error correlation length scale 1
Observation error correlation length scale 0

Streamlit: Optimal interpolation demonstration

How the information from observations is spread to the unobserved variables: the role of the background error correlations



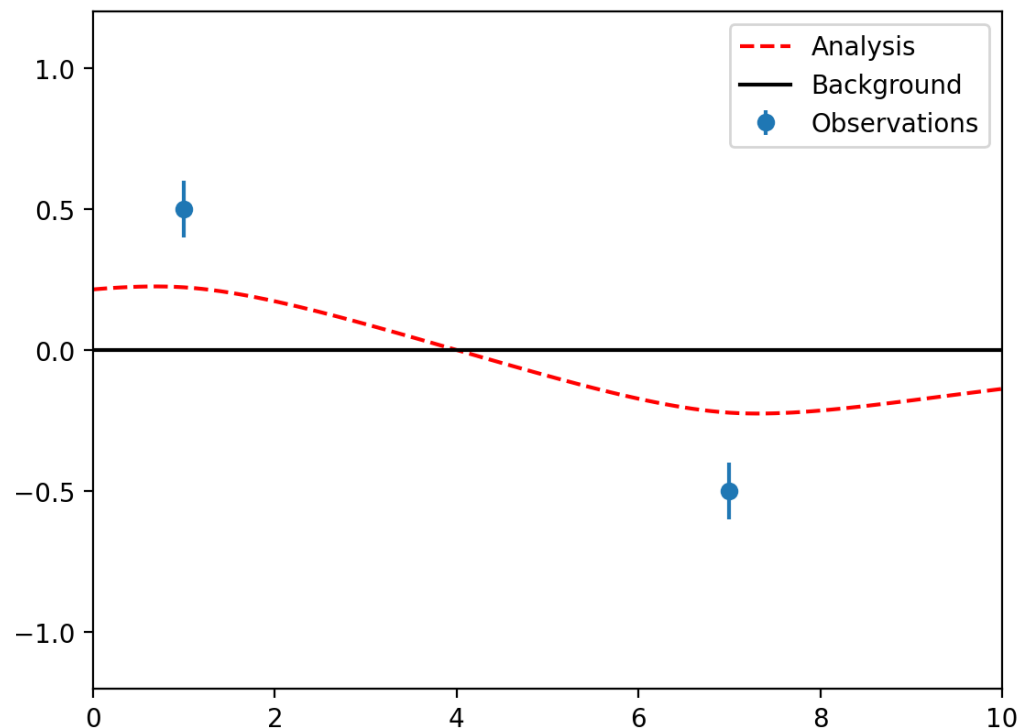
Parameters:

Observation error standard deviation 0.1
Background error standard deviation 0.1
Background error correlation length scale 0.2
Observation error correlation length scale 0

decreasing the background error correlation length scale reduces the spread of information from the observations.

Streamlit: Optimal interpolation demonstration

How the information from observations is spread to the unobserved variables: the role of the background error correlations



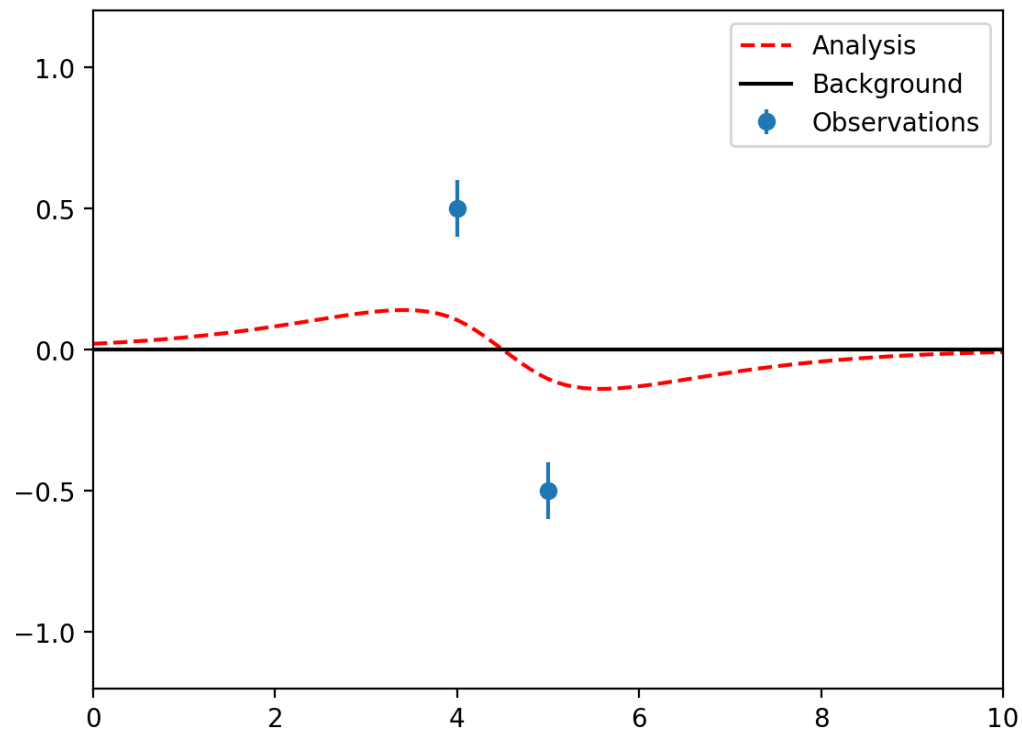
Parameters:

Observation error standard deviation 0.1
Background error standard deviation 0.1
Background error correlation length scale 2
Observation error correlation length scale 0

Increasing the background error correlation length scale increases the spread of information from the observations.

Streamlit: Optimal interpolation demonstration

How the information from observations is spread to the unobserved variables: the role of the observation error correlations

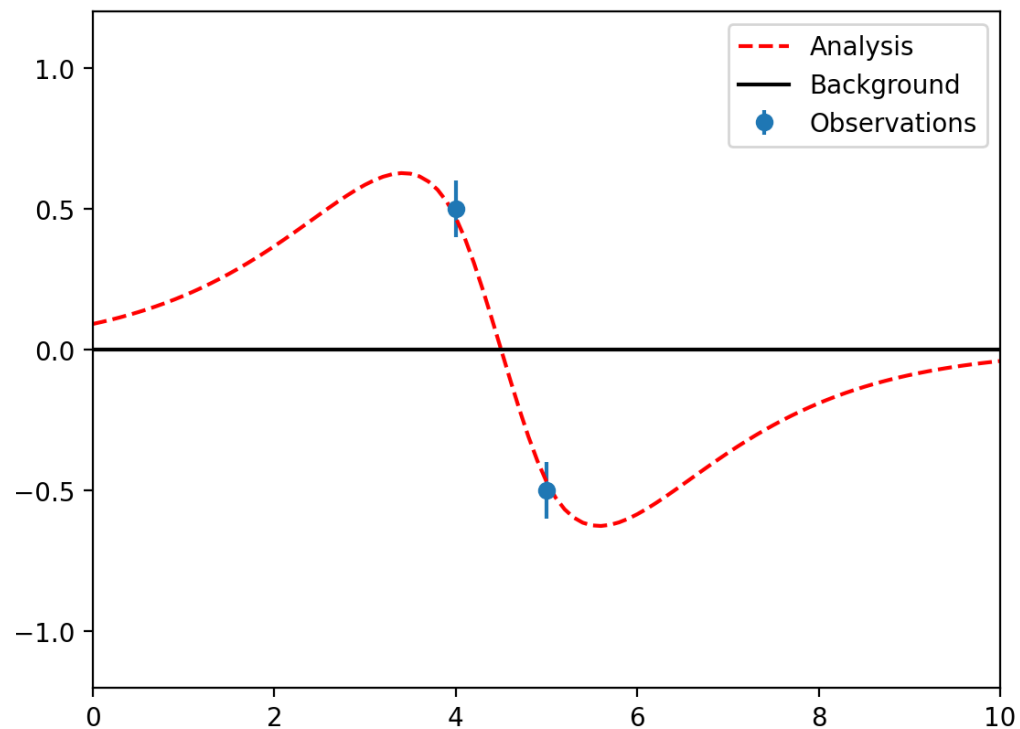


Parameters:

Observation error standard deviation 0.1
Background error standard deviation 0.1
Background error correlation length scale 1
Observation error correlation length scale 0

Streamlit: Optimal interpolation demonstration

How the information from observations is spread to the unobserved variables: the role of the observation error correlations

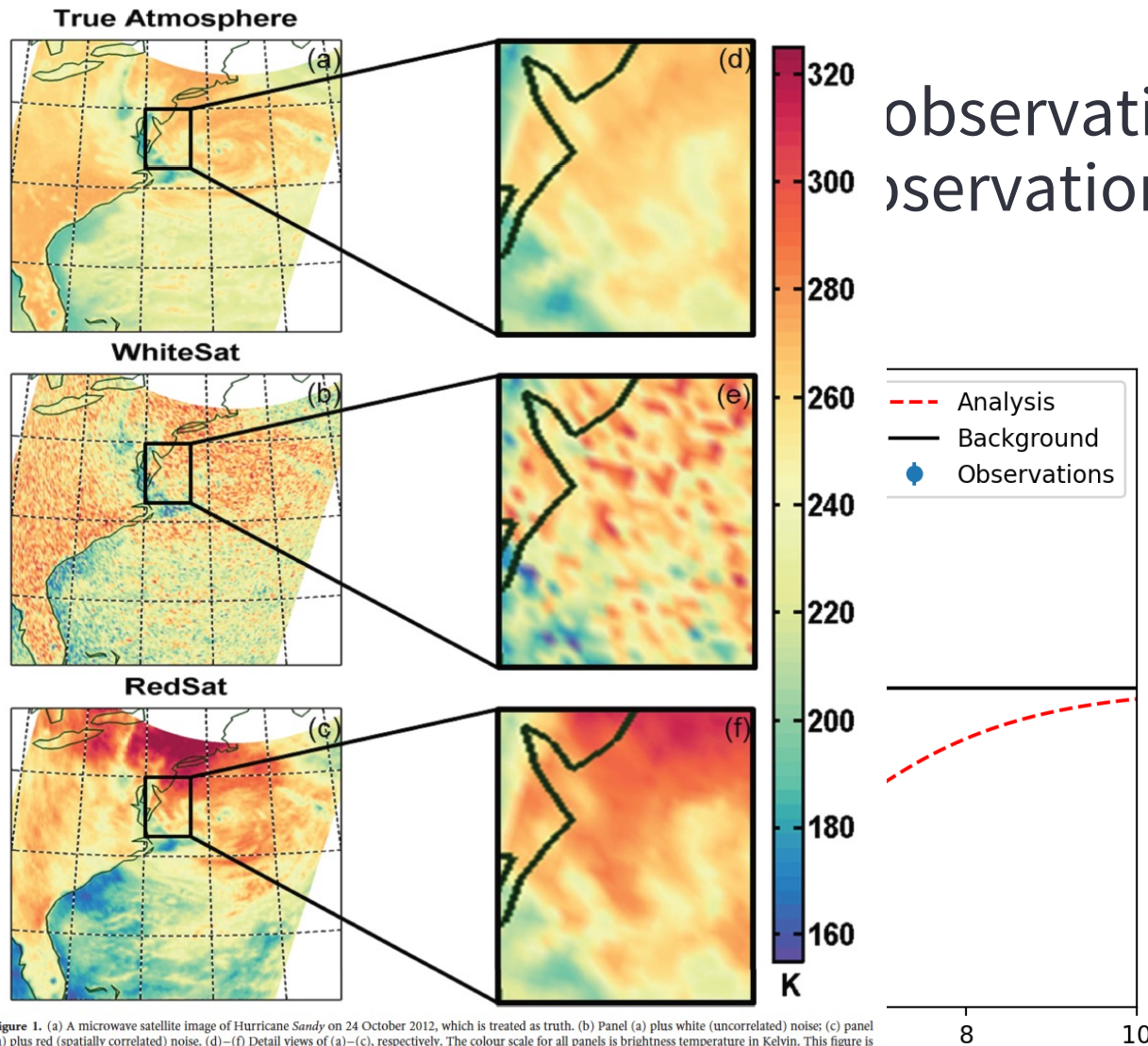


Parameters:

Observation error standard deviation 0.1
Background error standard deviation 0.1
Background error correlation length scale 1
Observation error correlation length scale 5

Increasing the observation error correlation length scale allows the observations to provide more information about gradients.

Streamlit: Optimal interpolation demonstration



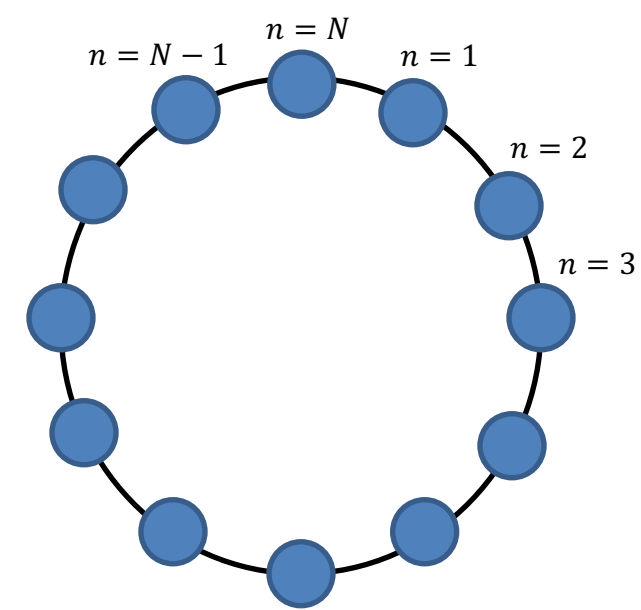
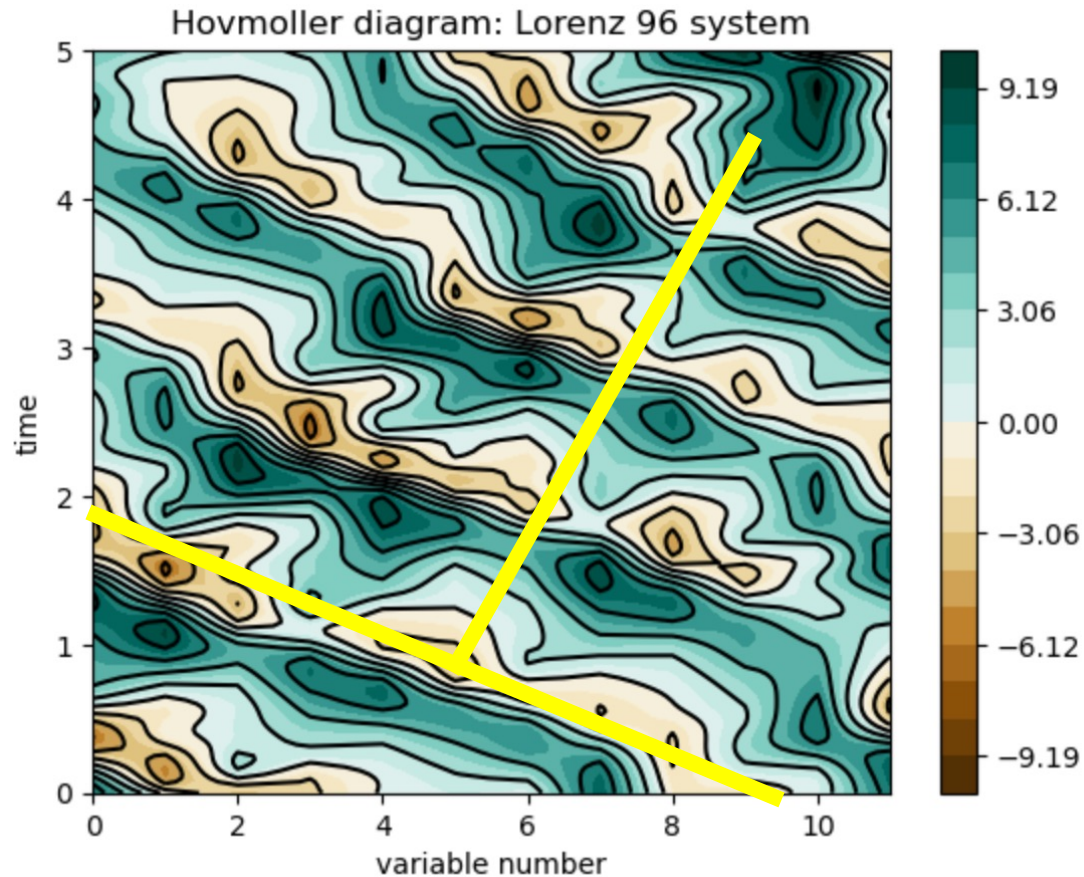
observations is spread to the unobserved
 observation error correlations

Parameters:

- Observation error standard deviation 0.1
- Background error standard deviation 0.1
- Background error correlation length scale 1
- Observation error correlation length scale 5

Increasing the observation error correlation length scale allows the observations to provide more information about gradients.

Lorenz 96 model



$$\frac{dx_i}{dt} = \underbrace{(x_{i+1} - x_{i-2})x_{i-1}}_{\text{advection}} - \underbrace{x_i}_{\text{dispersion}} + \underbrace{F}_{\text{External forcing}} \quad \text{for } i = 1, 2, \dots, n$$

$F=8$ is the chaotic regime.

Hovmoller plot demonstrates the presence of 'waves' with positive phase speed and negative group speed. Spatially coherent waves move left at a speed of 4 grid points per time unit

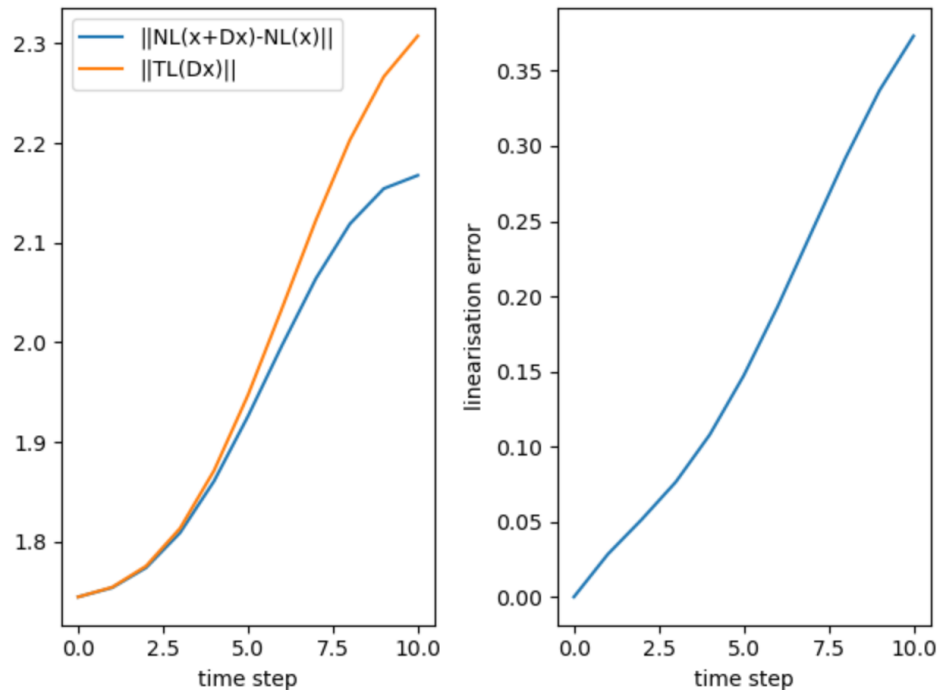
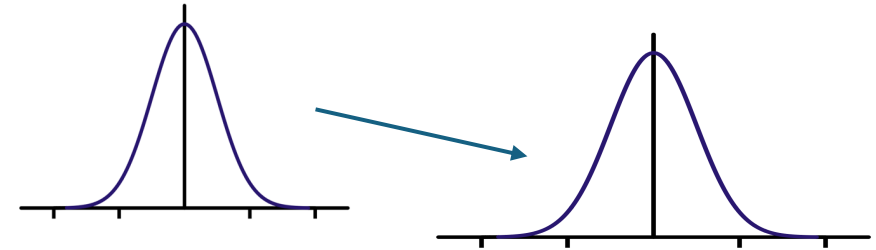
Lorenz 96 model – error in linear assumption

Many DA algorithms assume that the propagating model is near linear.

This is related to the Gaussian assumption.

The linearization error can be quantified as

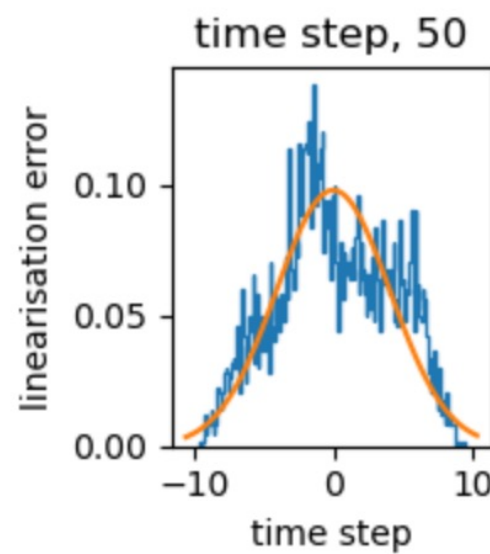
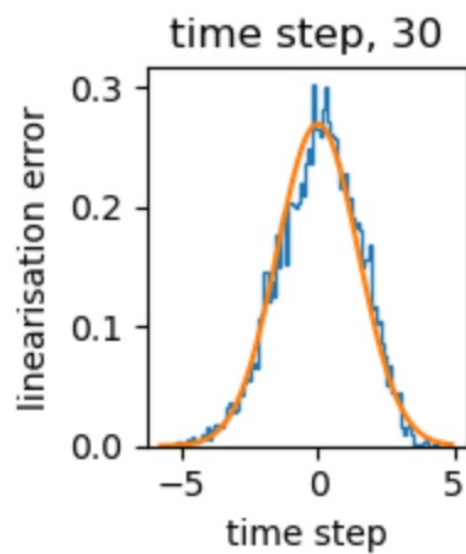
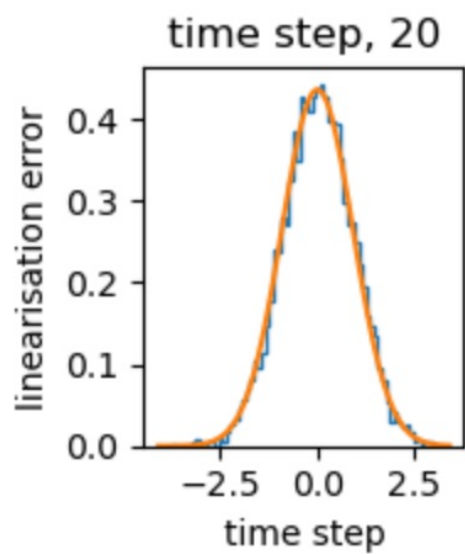
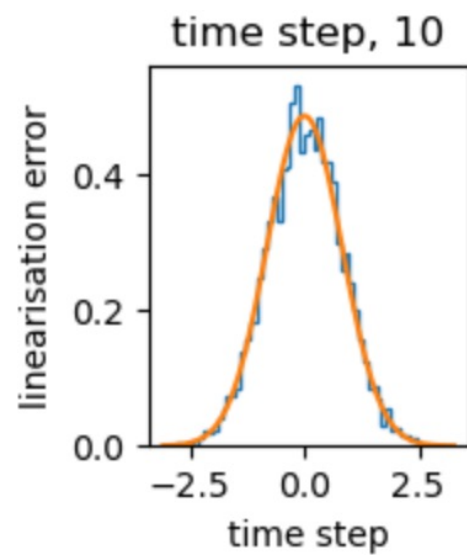
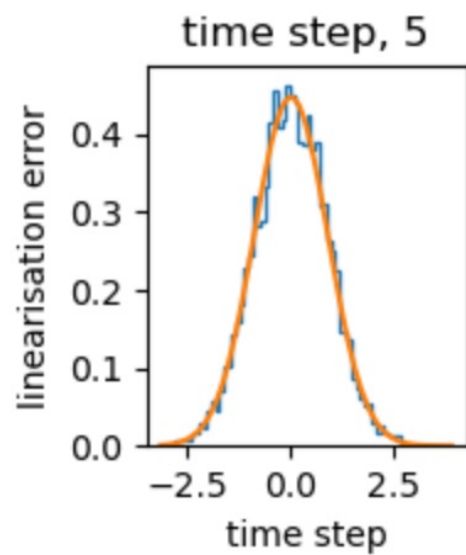
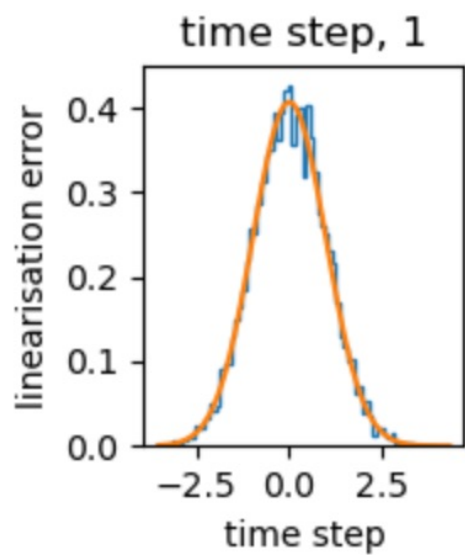
$$\text{linerror}(t) = ||NL(\mathbf{x} + \Delta\mathbf{x}, t) - NL(\mathbf{x}, t) - TL\Delta\mathbf{x}, t||$$



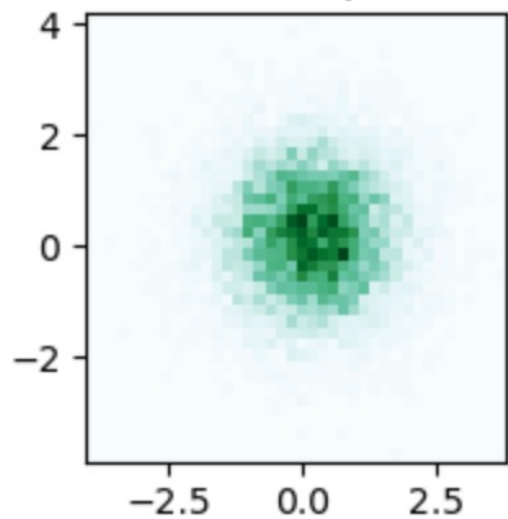
The linearization depends on the magnitude of the perturbation, $\Delta\mathbf{x}$, and the length of the simulation.

The size of the perturbation is related to the size of the corrections we want to make during the DA process.

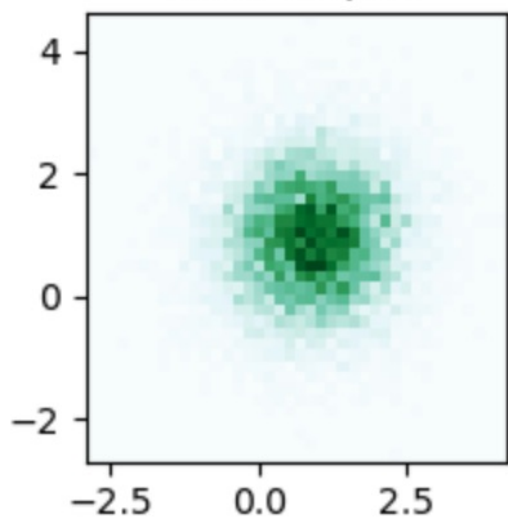
The linearization error should be used to guide how frequently the assimilation should be performed.



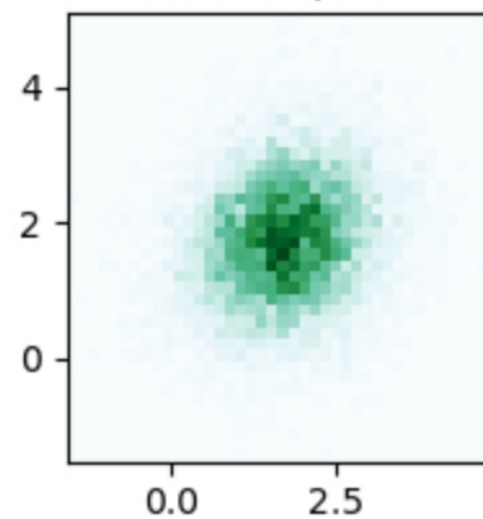
time step, 1



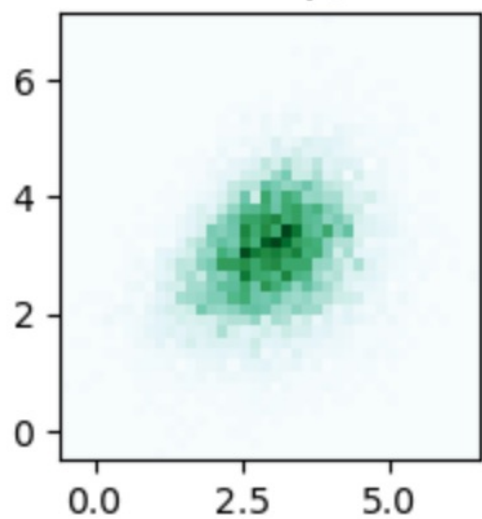
time step, 5



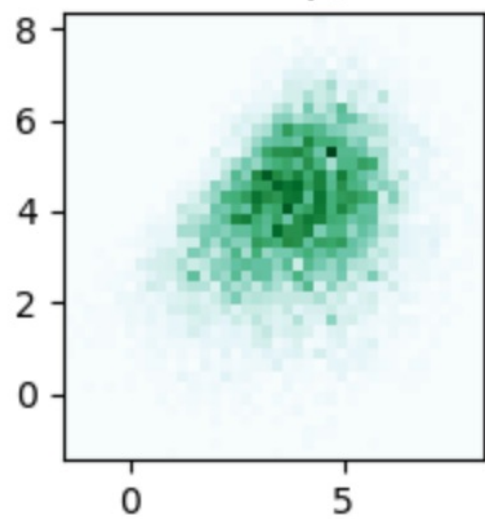
time step, 10



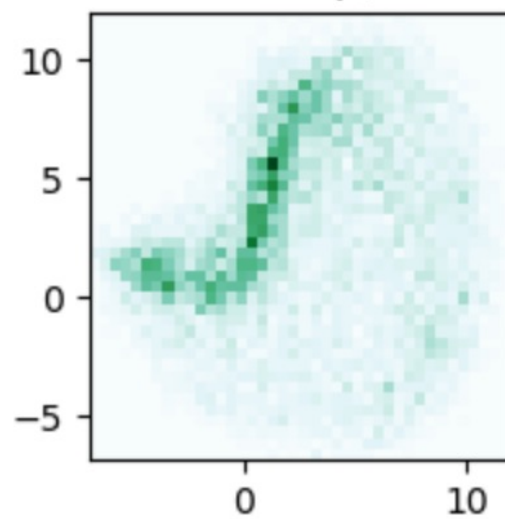
time step, 20



time step, 30



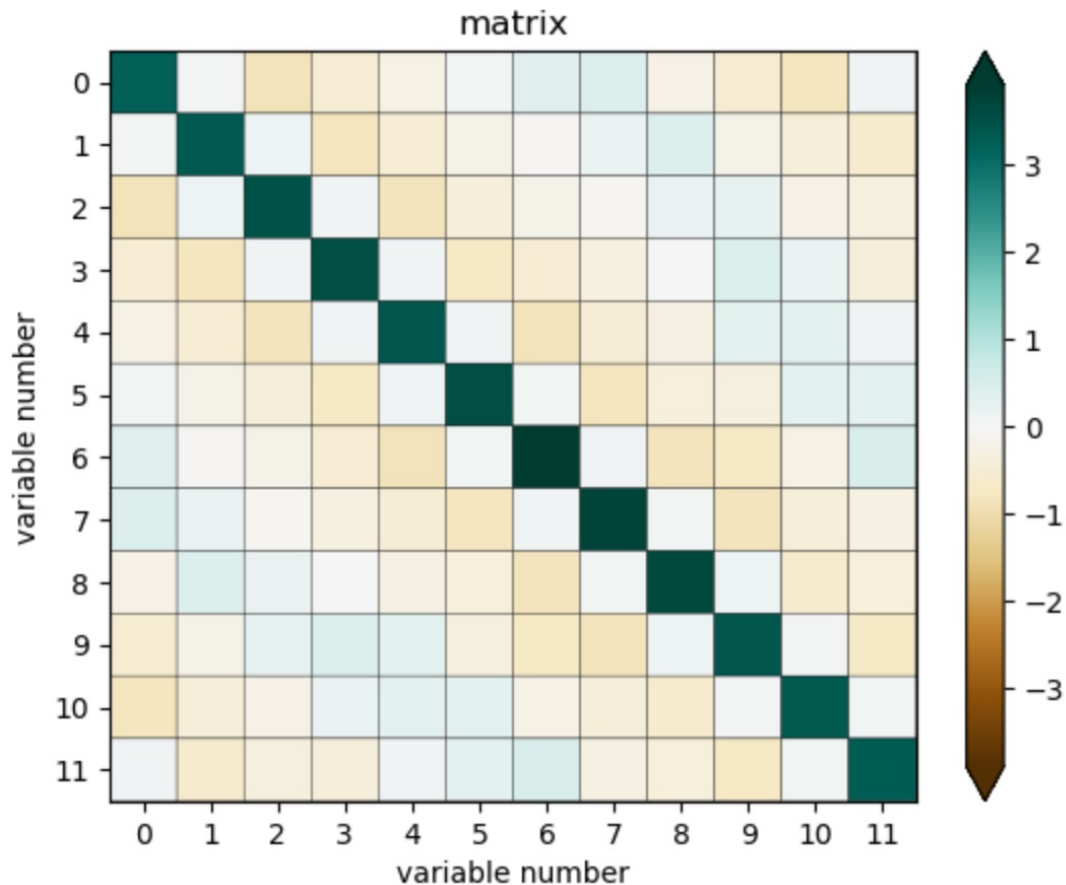
time step, 50



Lorenz 96 – B matrix

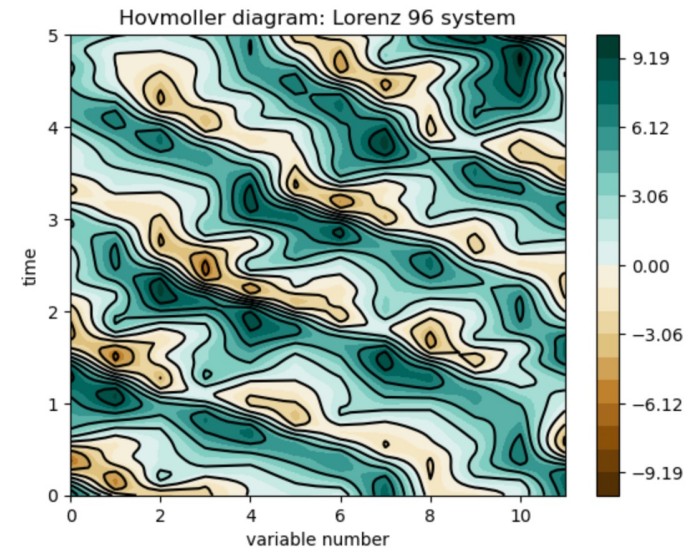
We generate an approximate B matrix using the Canadian Quick method

$$\epsilon = (\mathbf{x}(t + T) - \mathbf{x}(t)) / \sqrt{(2)}$$



Sample size = 10000

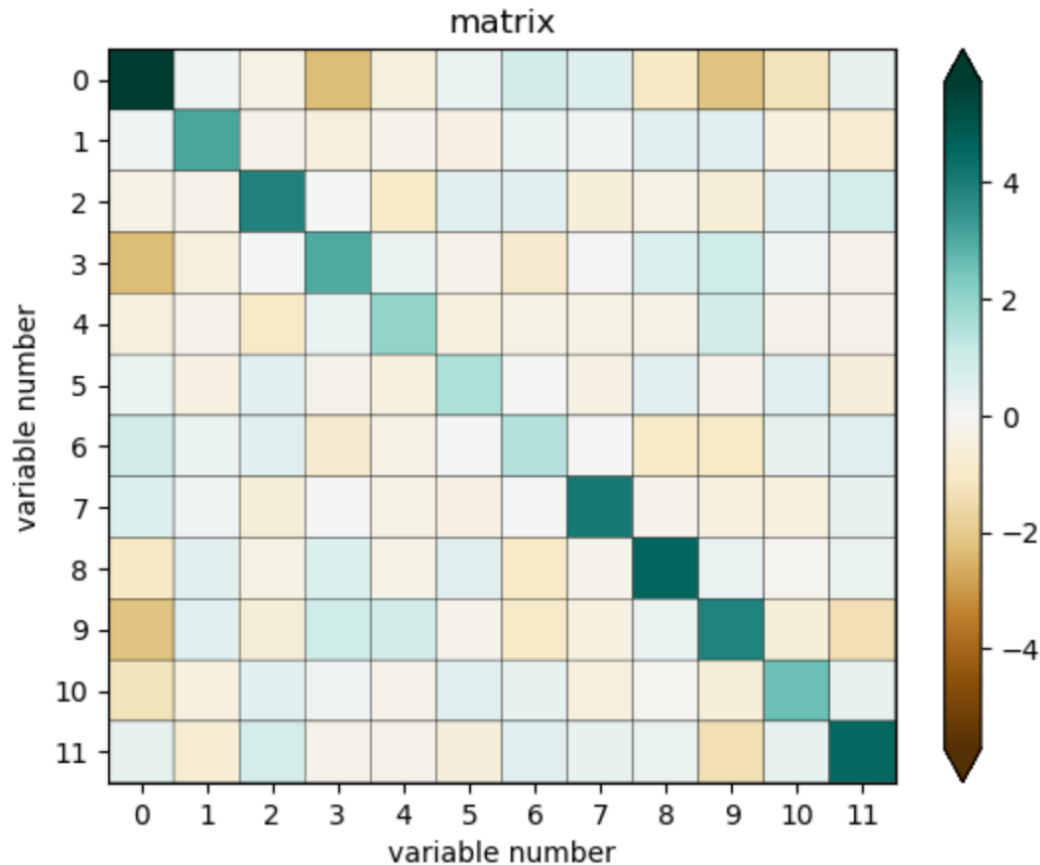
The structure in the B matrix can be compared back to the Hovmoller plot.



Lorenz 96 – B matrix

We generate an approximate B matrix using the Canadian Quick method

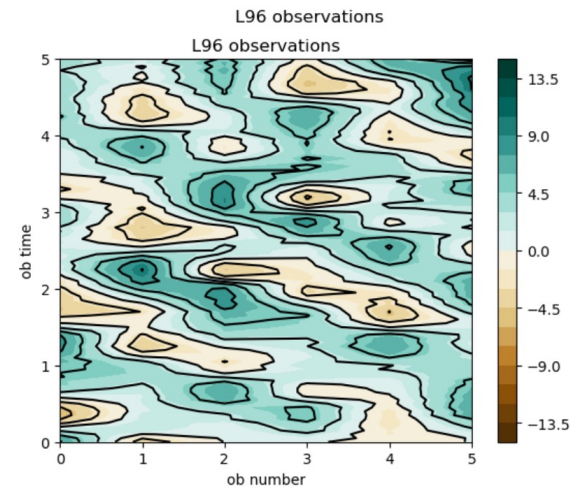
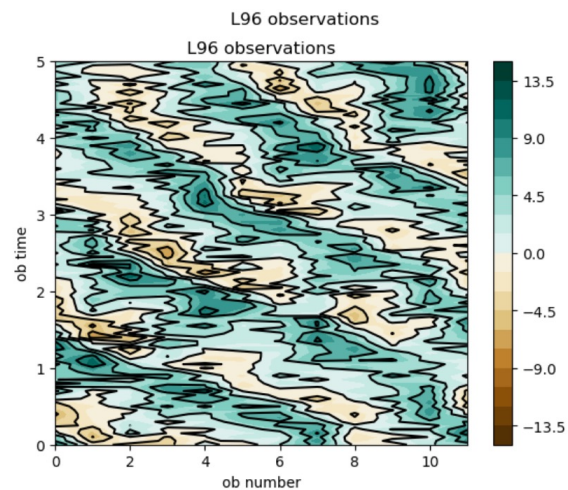
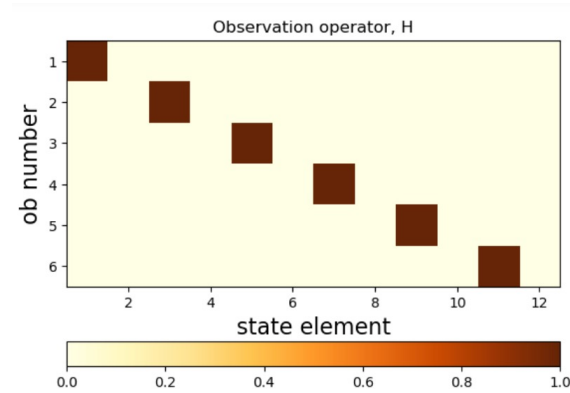
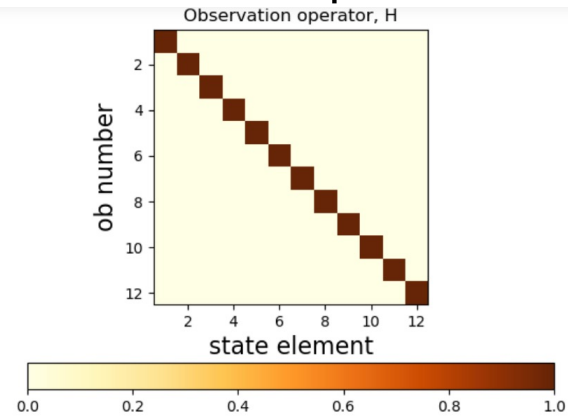
$$\epsilon = (\mathbf{x}(t + T) - \mathbf{x}(t)) / \sqrt{2}$$



Reducing the sample size to 200 results in noisy unphysical correlations that will affect how the information in the observations is used.

Lorenz 96 – generating observations

We generate observations consistent with the observation error standard deviations and observation operator



Which is better observing every grid point with a large error standard deviation or observing every other grid point with a small error standard deviation?

We can explore this question in the practicals other the rest of the week.