

## Data assimilation and inverse problems – Homework Set 3

### EnKFs for L'96

Build EnKFs (perturbed obs. and square root) for the L'96 model. You can follow the steps below to build these systems.

1. Write code to numerically solve the L'96 equations (use  $F = 8$ ) using a fourth-order Runge-Kutta method. Keep the number of variables as a parameter in your code. Use  $n = 40$  as the default dimension for this assignment.
2. To find an initial condition on (or close to) the L'96 attractor do the following. Solve the equations starting with an arbitrary initial condition for “a long time”, e.g., start with a uniformly distributed initial condition and simulate for 100 dimensionless time units. This should bring you close to the attractor. Use the final state of this simulation as your initial condition for the simulations below.
3. Create a synthetic data set by simulating the L96 model starting from an initial condition on the attractor (see above) for 10 dimensionless time units. Observe every other state variable every  $\Delta t = 0.1$  dimensionless time units, perturbed by Gaussian noise with mean zero and covariance matrix  $R = I$  (here  $I$  is the  $20 \times 20$  identity matrix).
4. Assimilate the synthetic data by an EnKF in perturbed obs. implementation. For your initial ensemble, use randomly chosen states from a “long” simulation of L96 (1000 or more dimensionless time units). Do *not* use localization or inflation. Plot RMSE and spread to check if the algorithm works, i.e.,  $\text{RMSE} \approx \text{spread}$ . Disregard the first half of your simulation/assimilation as “spin-up”. Increase the ensemble size until the algorithm works. How large must your ensemble be?
5. Implement localization and inflation for your perturbed obs. EnKF. Fix the ensemble size  $N_e = 20$ . Tune localization and inflation by performing data assimilation over a grid of localization/inflation parameters. For each set of parameters, compute the average RMSE and average spread (over time). Don't forget to disregard the spin-up time. Use UA's HPC system to perform this “optimization”. Use the parameters that yield the smallest RMSE and plot RMSE and spread for the optimal parameters.
6. Repeat steps 4 and 5, but replace “perturbed obs. EnKF” by “square root EnKF”.
7. Increase dimension to  $n = 400$ . Create synthetic data (using steps 2 and 3 above). Run your tuned perturbed obs. EnKF and square root EnKFs with ensemble size  $N_e = 20$  to assimilate these data. Plot RMSE and spread. Do these filters “work”? Why?
8. Assimilate the data of a 40-dimensional and 400-dimensional model that are posted on the website. These are observations of every other variable of a L96 model (starting with  $x_1$ ) that I obtained every 0.1 dimensionless time units.