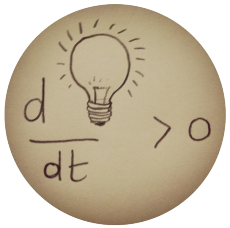
Data Assimilation and Genetic Algorithms for the Parameter Estimation Problem in Simple Climate Models

Morgan R. Frank & Andrew Reagan

Department of Mathematics & Statistics
Vermont Complex Systems Center
Computational Story Lab
Vermont Advanced Computing Center
University of Vermont



The Parameter Estimation Problem

The Parameter Estimation Problem

- Start with some observation data for some phenomenon

The Parameter Estimation Problem

- Start with some observation data for some phenomenon
- Start with a well-principled model for that phenomenon

The Parameter Estimation Problem

- Start with some observation data for some phenomenon
- Start with a well-principled model for that phenomenon
- Can we find model parameters to reproduce the observation data?

Systems of Interest

Lorenz '63

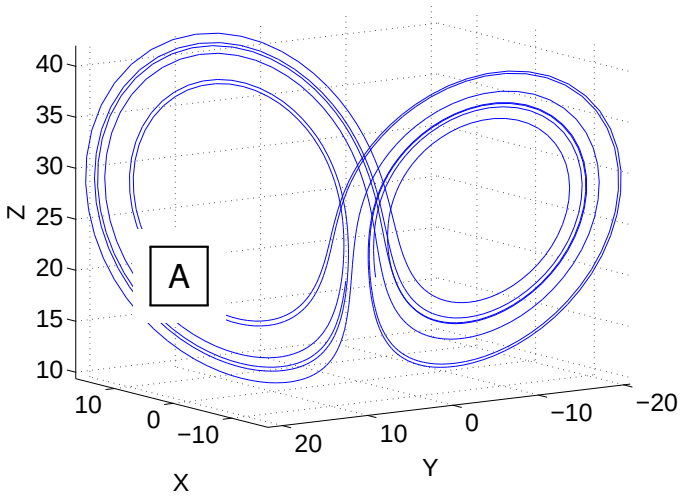
$$\begin{aligned}\frac{dx}{dt} &= \sigma(y - x)\\ \frac{dy}{dt} &= \rho x - y - xz\\ \frac{dz}{dt} &= xy - \beta z.\end{aligned}$$

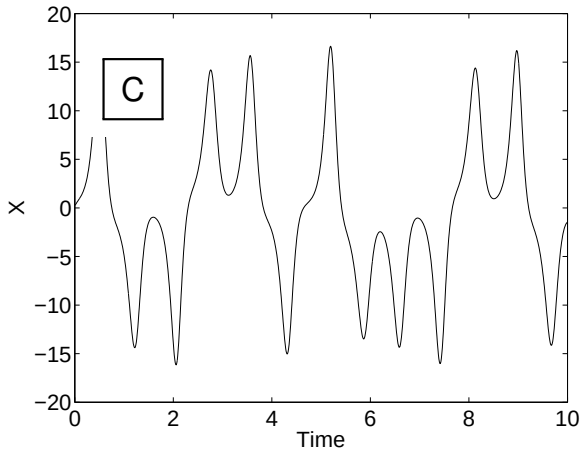
Lorenz '96

$$\begin{aligned}\frac{dx_i}{dt} &= x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F - \frac{hc}{b}\sum_{j=1}^{J} y_{(j,i)}\\ \frac{dy_{(j,i)}}{dt} &= cby_{(j+1,i)}(y_{(j-1,i)} - y_{(j+2,i)}) - cy_{(j,i)} + \frac{hc}{b}x_i\end{aligned}$$

Systems of Interest

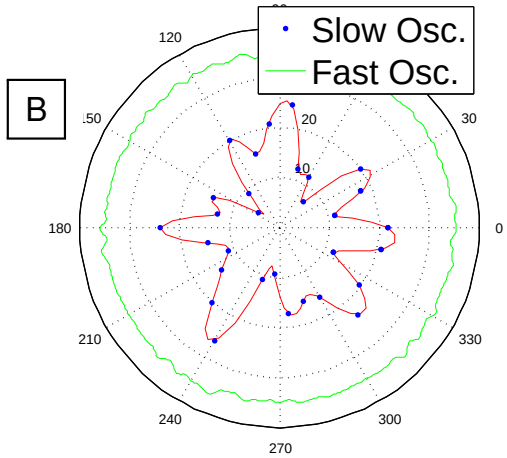
Lorenz '63 (L63)

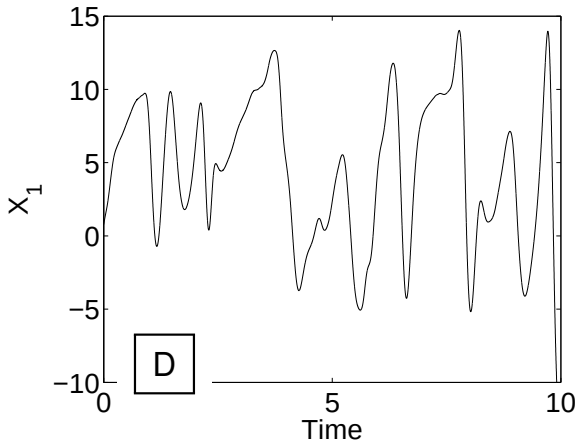




Lorenz '96 (L96)

I=30, J=5, F=14





Genetic Algorithms (GA)

- *Evolution is a pretty sweet problem solver. . . just need to fit the paradigm*

Genetic Algorithms (GA)

- *Evolution is a pretty sweet problem solver. . . just need to fit the paradigm*
- **Population:** real-valued vectors (called **individuals** or **genes**), each index represents a parameter choice (or **alleles**)

Genetic Algorithms (GA)

- *Evolution is a pretty sweet problem solver. . . just need to fit the paradigm*
- **Population:** real-valued vectors (called **individuals** or **genes**), each index represents a parameter choice (or **alleles**)
- **Genetic Mutation:** with some probability, add or subtract a small amount from an index in a vector

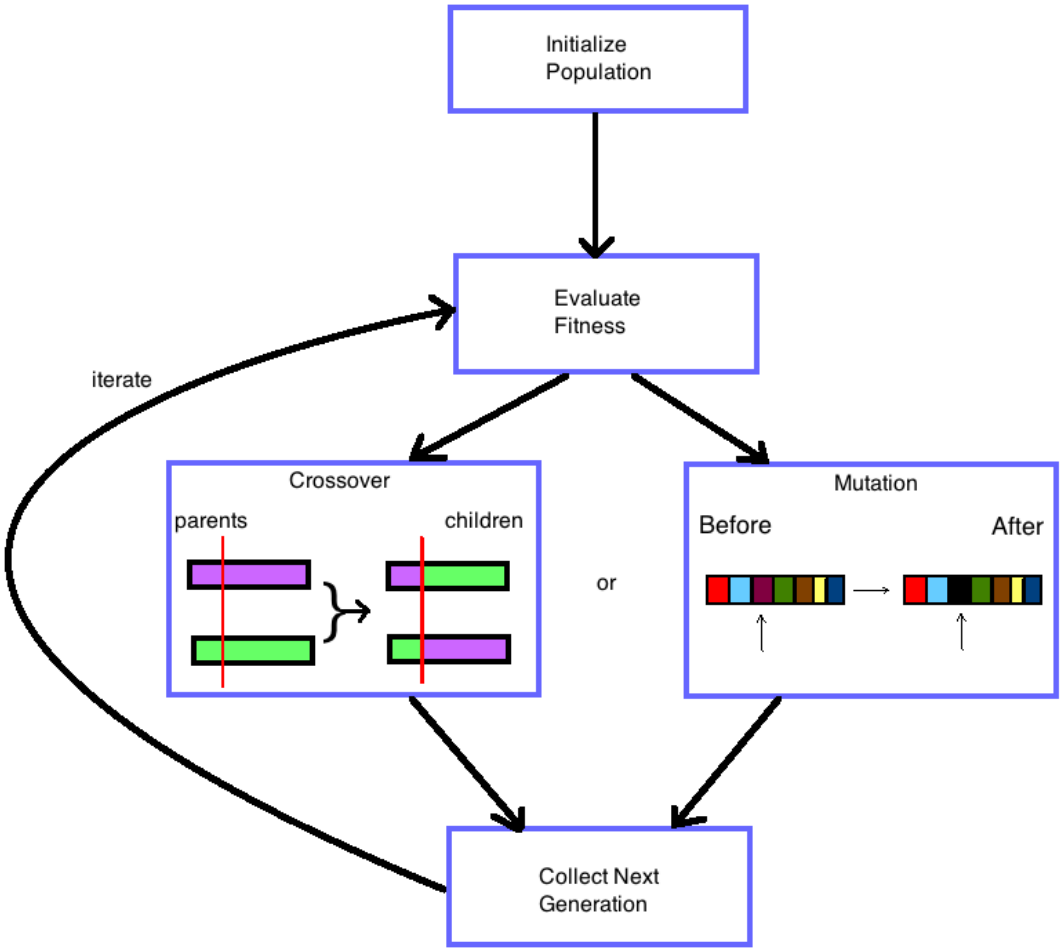
Genetic Algorithms (GA)

- *Evolution is a pretty sweet problem solver. . . just need to fit the paradigm*
- **Population:** real-valued vectors (called **individuals** or **genes**), each index represents a parameter choice (or **alleles**)
- **Genetic Mutation:** with some probability, add or subtract a small amount from an index in a vector
- **Selection Pressure:** integrate model using parameters from an individual in the population. The point-wise root mean square error of this time series compared to observation data (or **fitness**)

Genetic Algorithms (GA)

- *Evolution is a pretty sweet problem solver. . . just need to fit the paradigm*
- **Population:** real-valued vectors (called **individuals** or **genes**), each index represents a parameter choice (or **alleles**)
- **Genetic Mutation:** with some probability, add or subtract a small amount from an index in a vector
- **Selection Pressure:** integrate model using parameters from an individual in the population. The point-wise root mean square error of this time series compared to observation data (or **fitness**)
- **Reproduction/Crossover:** select two individuals from the population, called **parents**, and select a random index. Create to **children** vectors. The first child copies the entries from Parent 1 up to the selected index, and fills the remaining entries from Parent 2. Swap the roles of the parents to create the second child. The children replace the parents in the

Genetic Algorithms



The GA Experiments

- For the L63: $\sigma = 10$ & $\beta = 8/3$

The GA Experiments

- For the L63: $\sigma = 10$ & $\beta = 8/3$
- Test the effects of system dynamics on GA by tuning $\rho \in [22, 28, 35]$

The GA Experiments

- For the L63: $\sigma = 10$ & $\beta = 8/3$
- Test the effects of system dynamics on GA by tuning $\rho \in [22, 28, 35]$
- For the L96: $h = 1$, $b = c = 10$, $F = 14$ & $J = 4$

The GA Experiments

- For the L63: $\sigma = 10$ & $\beta = 8/3$
- Test the effects of system dynamics on GA by tuning $\rho \in [22, 28, 35]$
- For the L96: $h = 1$, $b = c = 10$, $F = 14$ & $J = 4$
- Test the effects of system dimensionality on GA by tuning $I \in [4, 8, 10, 15]$

The GA Experiments

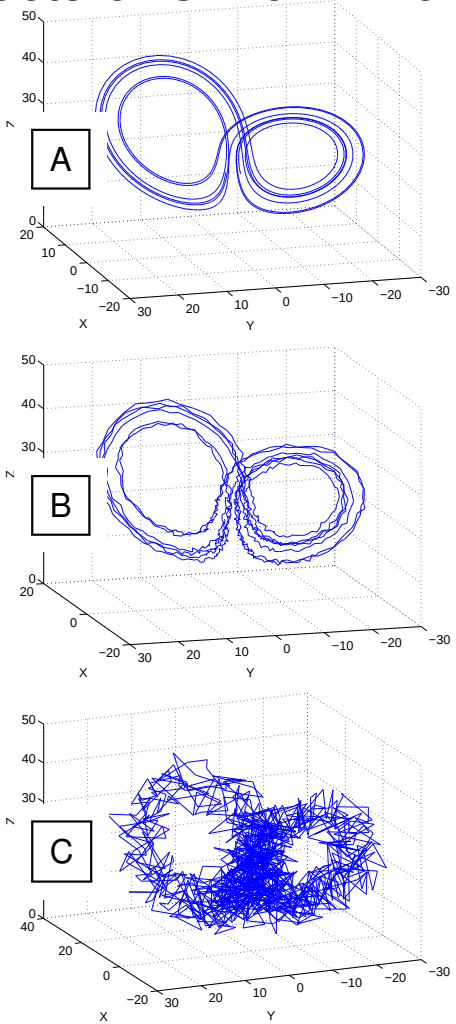
- For the L63: $\sigma = 10$ & $\beta = 8/3$
- Test the effects of system dynamics on GA by tuning $\rho \in [22, 28, 35]$
- For the L96: $h = 1$, $b = c = 10$, $F = 14$ & $J = 4$
- Test the effects of system dimensionality on GA by tuning $I \in [4, 8, 10, 15]$
- Test the effects of subsampling

The GA Experiments

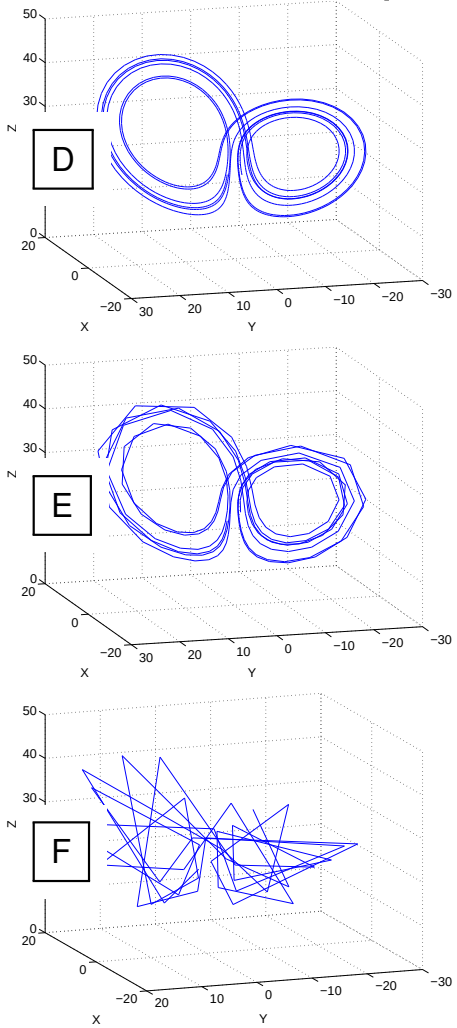
- For the L63: $\sigma = 10$ & $\beta = 8/3$
- Test the effects of system dynamics on GA by tuning $\rho \in [22, 28, 35]$
- For the L96: $h = 1$, $b = c = 10$, $F = 14$ & $J = 4$
- Test the effects of system dimensionality on GA by tuning $I \in [4, 8, 10, 15]$
- Test the effects of subsampling
- Test the effects of normal and uniformly distributed noise in the observation data

The Effects of Subsampling and Artificial Noise in Observation Data

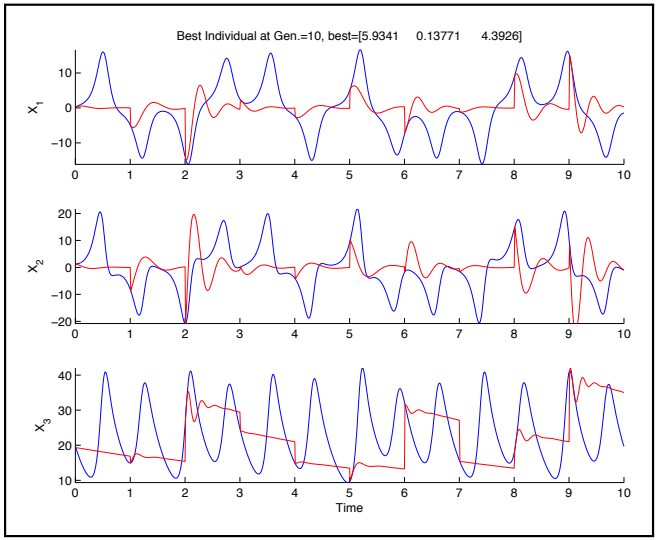
Effects of Uniform Noise:

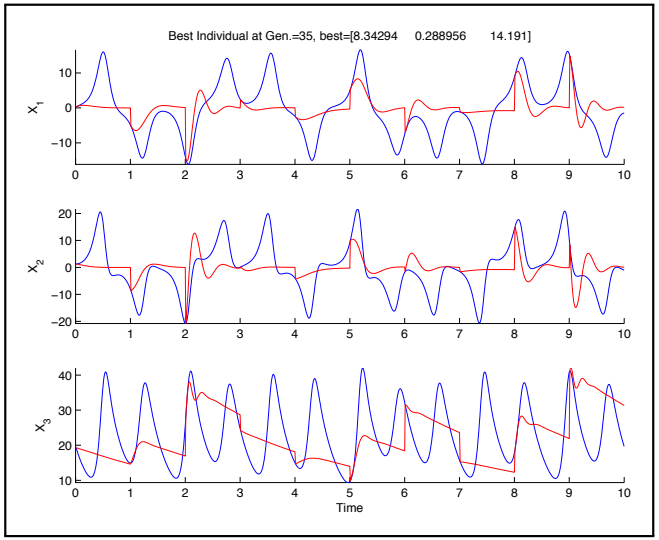


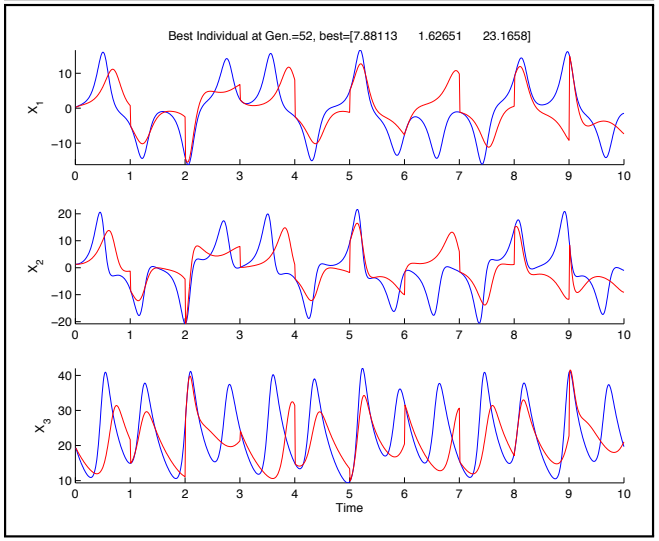
Effects of Subsampling:

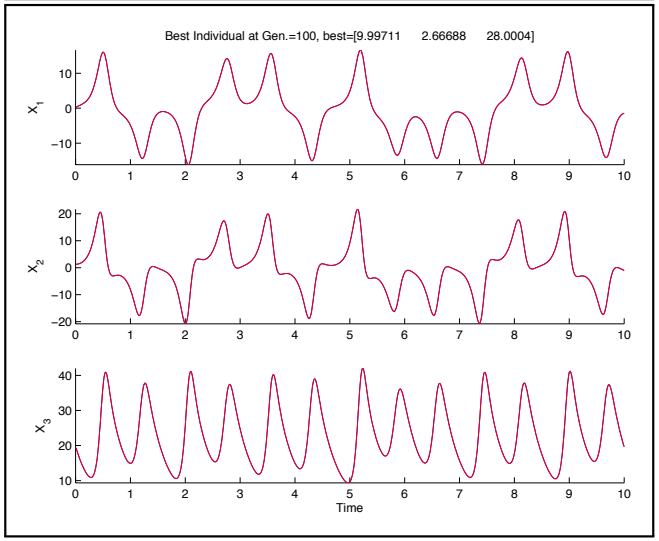


Running the GA

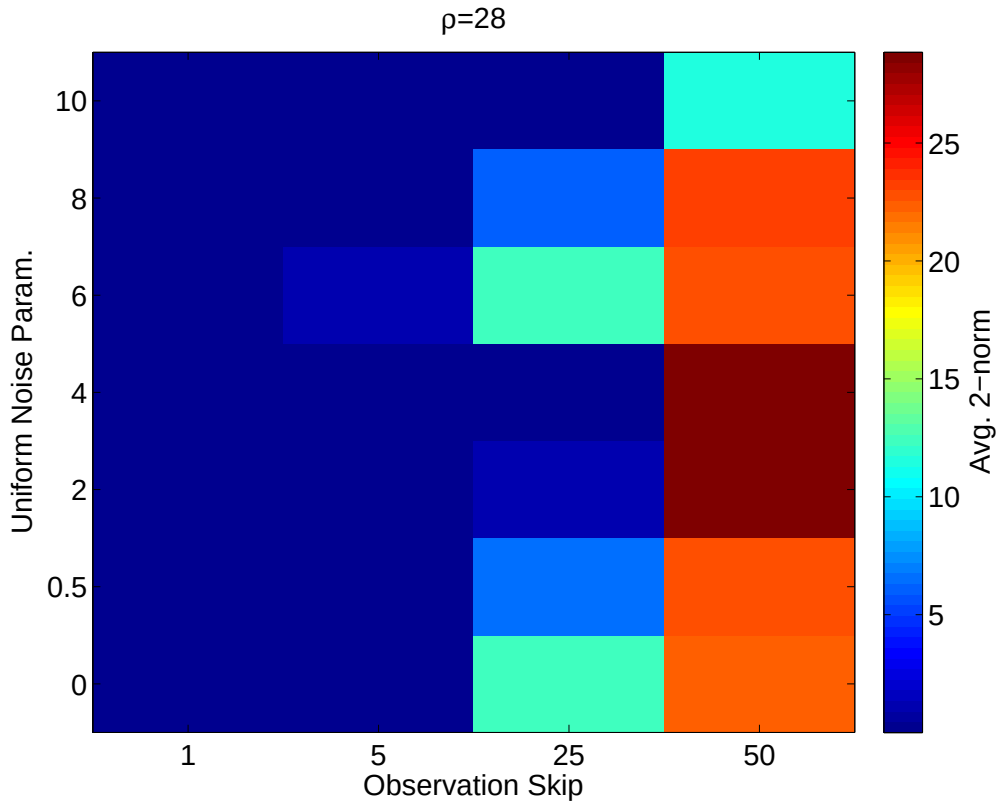




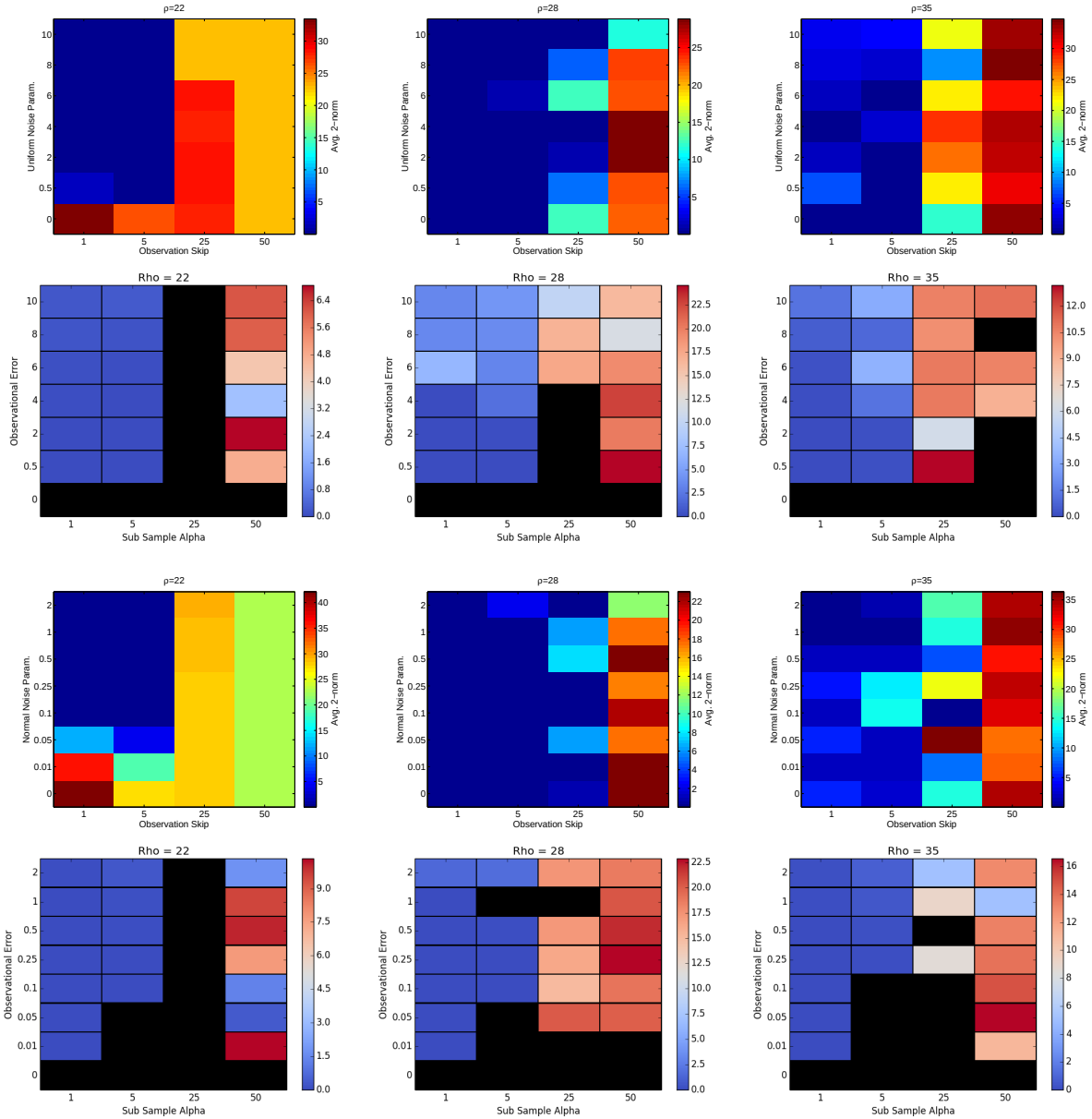




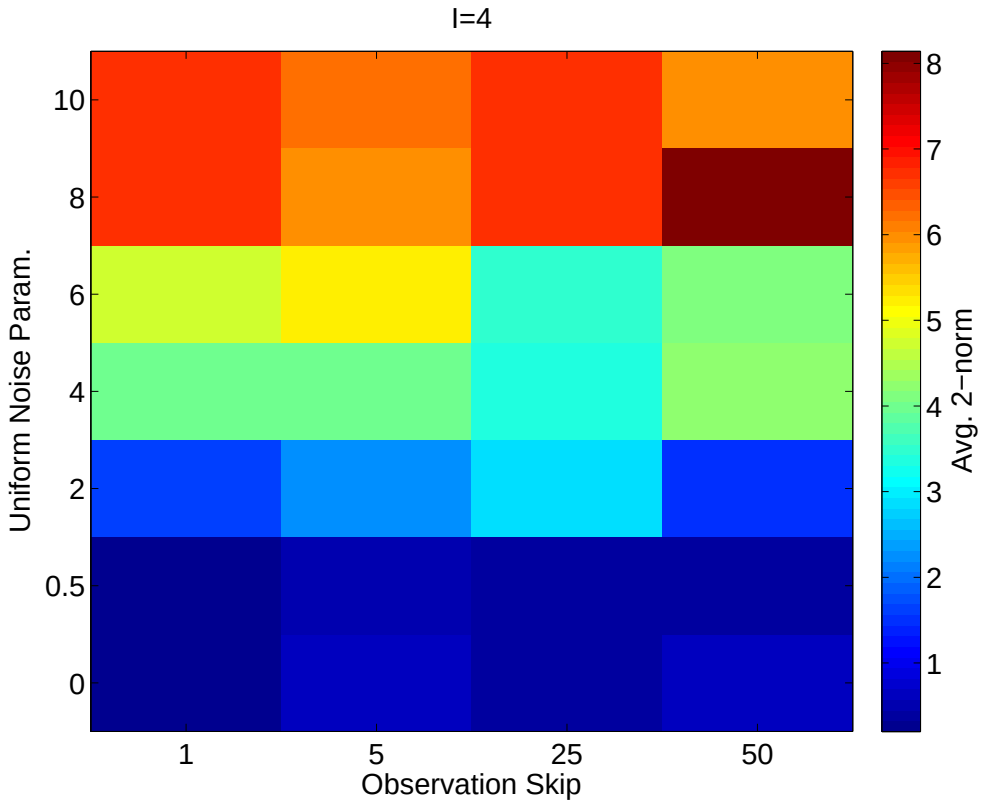
System Dynamics in L63



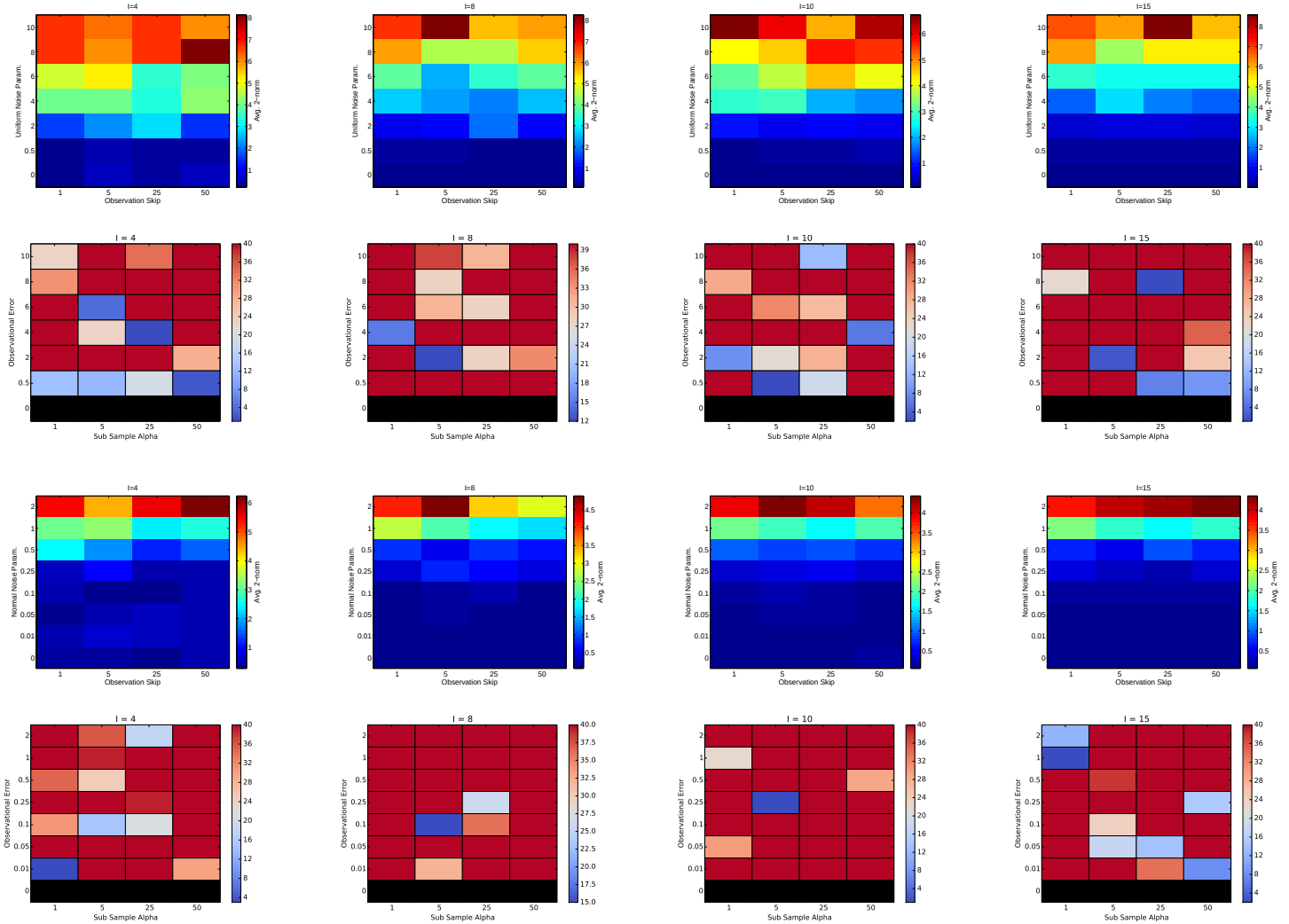
System Dynamics in L63



Dimensionality in L96



Dimensionality in L96



Discussion & Future Work

Discussion & Future Work

- GA performs comparably to Data Assimilation (DA) for Parameter Estimation as “out-of-the-box" tools

Discussion & Future Work

- GA performs comparably to Data Assimilation (DA) for Parameter Estimation as “out-of-the-box" tools
- DA diverged for some experiments. Steps can be taken to prevent this. GA performs without special considerations.

Discussion & Future Work

- GA performs comparably to Data Assimilation (DA) for Parameter Estimation as “out-of-the-box" tools
- DA diverged for some experiments. Steps can be taken to prevent this. GA performs without special considerations.
- Initial results show GA **FAILS** the one-variable problem, while DA got it right for many cases.

Discussion & Future Work

- GA performs comparably to Data Assimilation (DA) for Parameter Estimation as “out-of-the-box" tools
- DA diverged for some experiments. Steps can be taken to prevent this. GA performs without special considerations.
- Initial results show GA **FAILS** the one-variable problem, while DA got it right for many cases.
- need to try with other simple models. Need to see how far up the model hierarchy before GA is impractical or ineffective.

Discussion & Future Work

- GA performs comparably to Data Assimilation (DA) for Parameter Estimation as “out-of-the-box" tools
- DA diverged for some experiments. Steps can be taken to prevent this. GA performs without special considerations.
- Initial results show GA **FAILS** the one-variable problem, while DA got it right for many cases.
- need to try with other simple models. Need to see how far up the model hierarchy before GA is impractical or ineffective.
- Attempt a real-world problem