Please Stop and Smell the Tracers:

Predicting Tracer Concentration Behavior in Low-Order Models with Data Assimilation

FAIRUZ ISHRAQUE
Colgate University

July 2021

Outline

1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

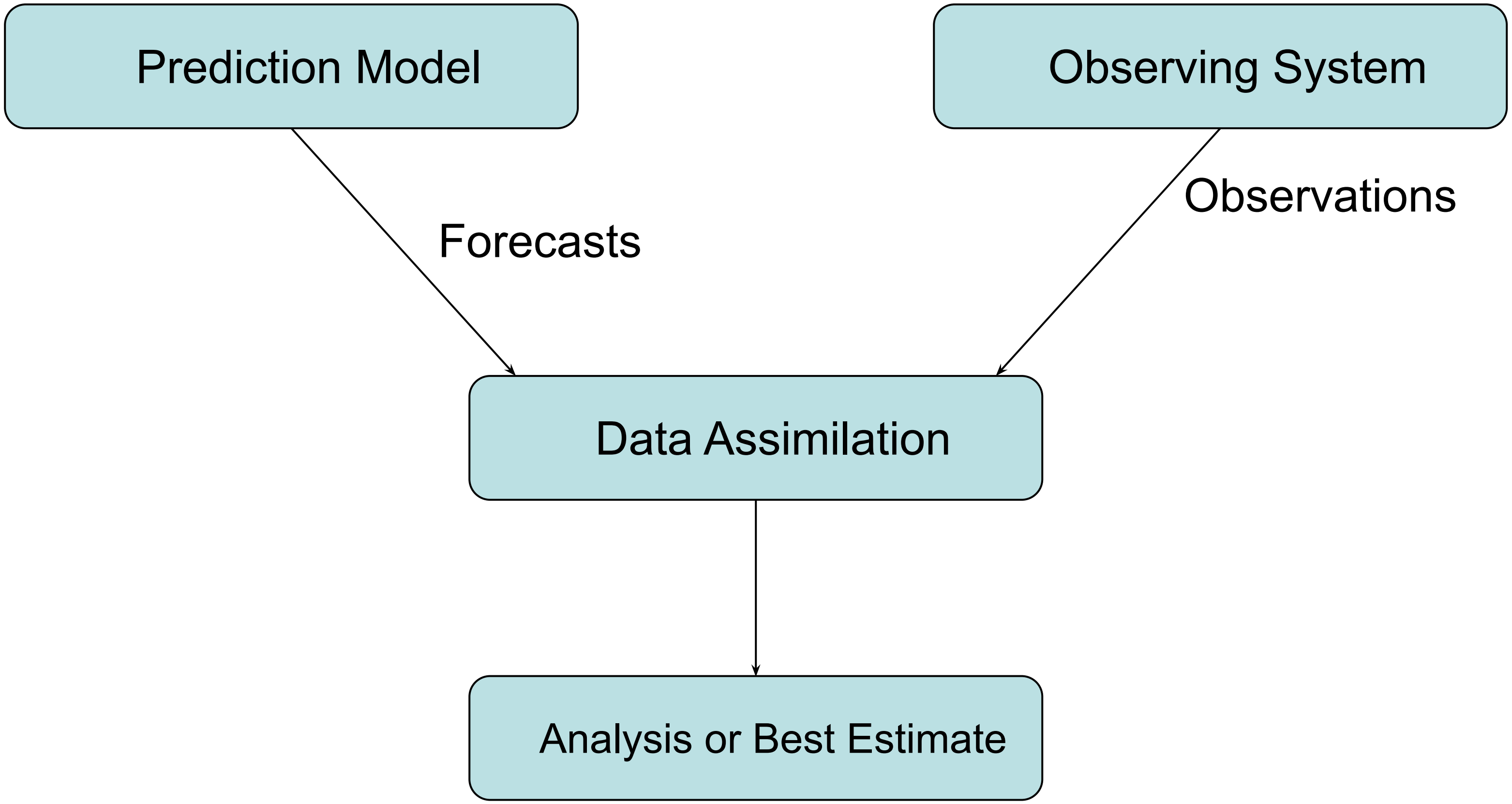
Motivation

Determining the harm posed by these tracer contaminants/pollutants requires better understanding of tracer concentration behaviors in the atmosphere as well as source characterization

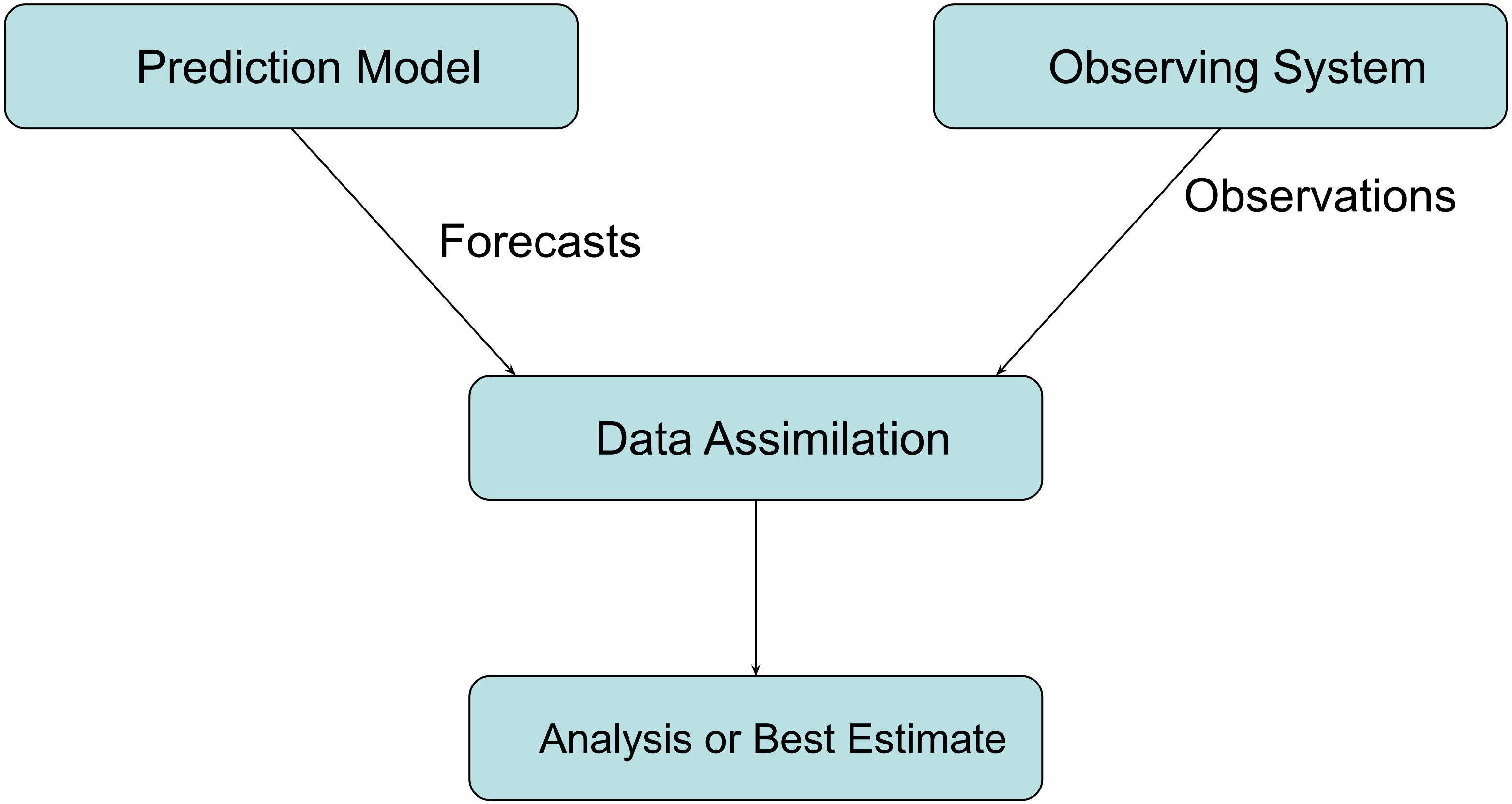
Outline

1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz 96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

Data Assimilation: A General Description



Data Assimilation: A General Description



Data assimilation combines model forecasts with observations to produce better predictions

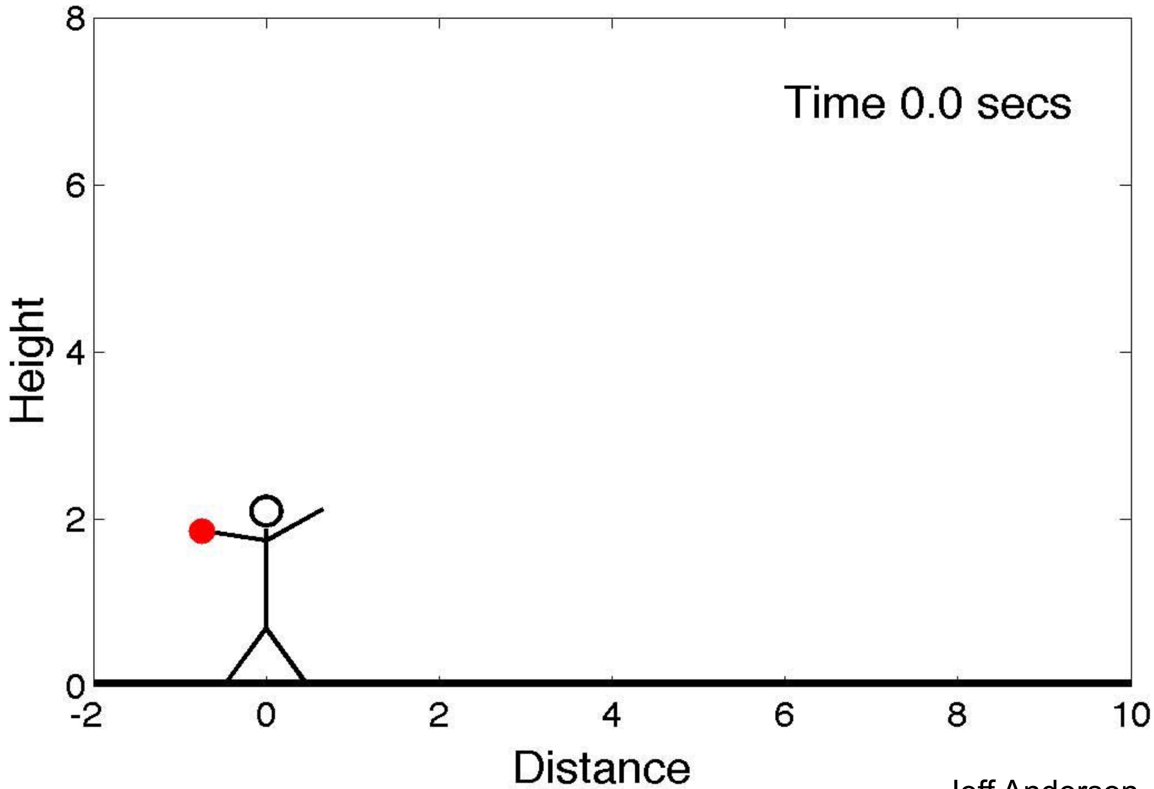
- Often used in numerical weather prediction
- Very good at predicting behaviors of systems sensitive to initial conditions

Throwing a ball

The prediction model for this system is quite simple:

$$x = x_{initial} + u_{initial} t$$

$$y = y_{initial} + v_{initial} t - 1/2\, g t^2$$



Jeff Anderson

Data Assimilation: A General Description

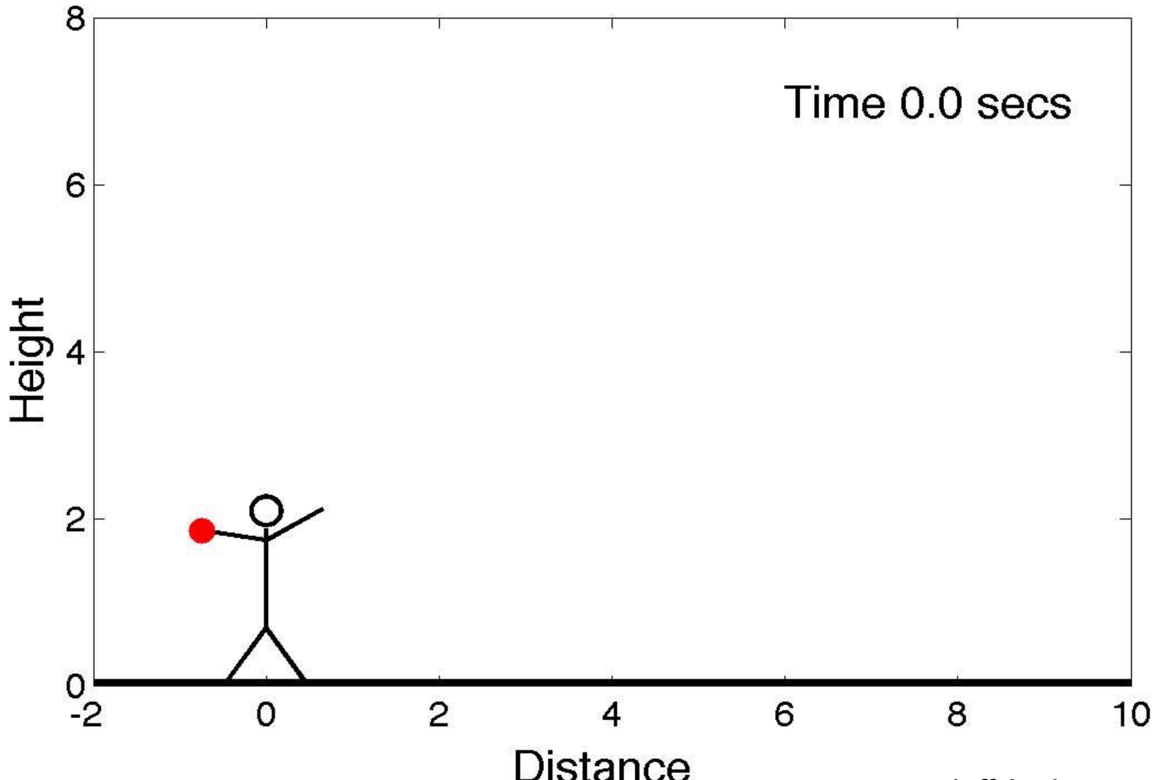
Throwing a ball

The prediction model for this system is quite simple:

$$x = x_{initial} + u_{initial} t$$

$$y = y_{initial} + v_{initial} t - 1/2\, g t^2$$

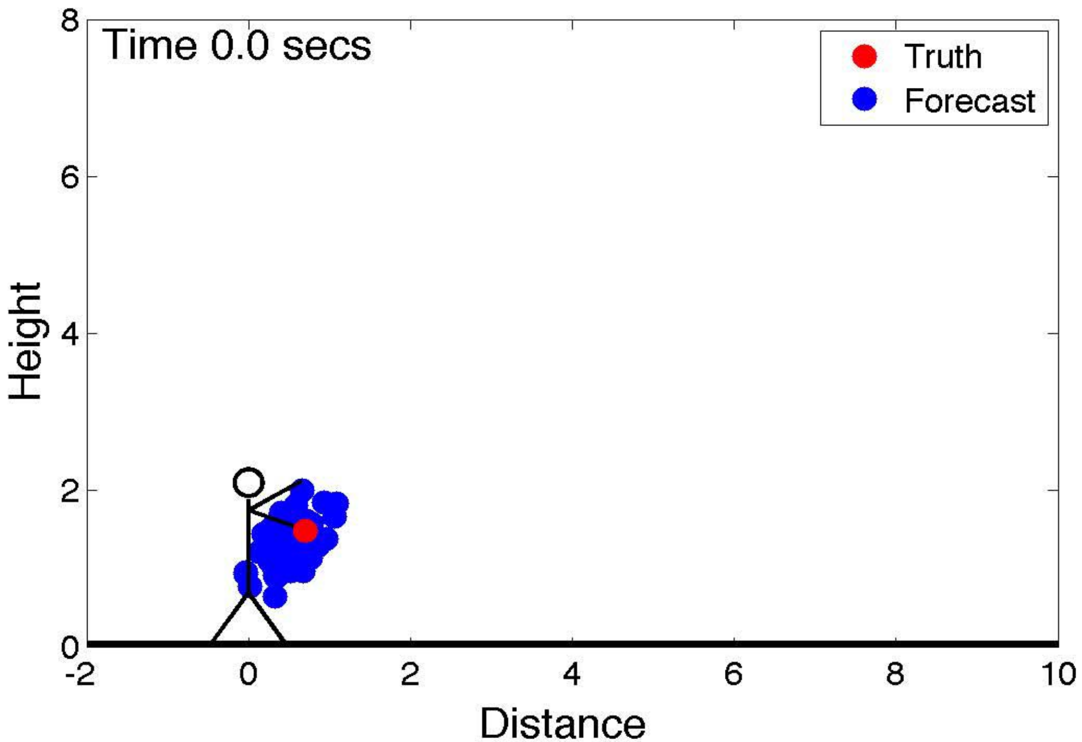
Realistically, the initial conditions are uncertain



Jeff Anderson

Data Assimilation: A General Description

We run the model with an **ensemble** of different initial conditions

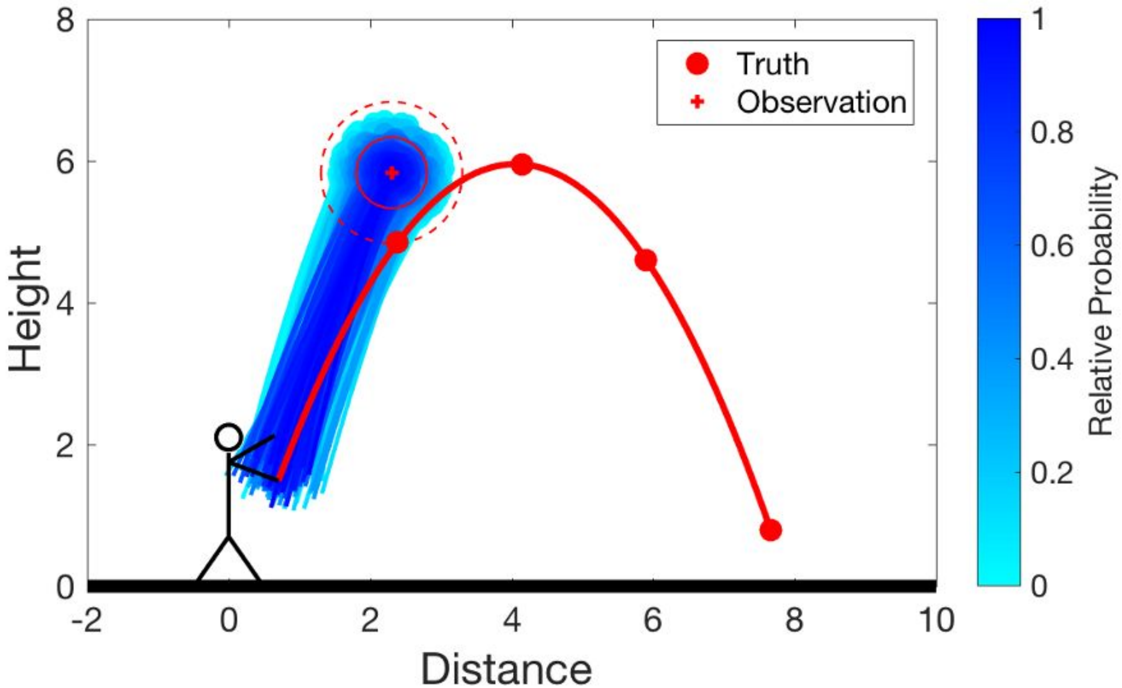


Jeff Anderson

Data assimilation narrows down the range of results the ensemble returns by combining observations with ensemble results

Data Assimilation: A General Description

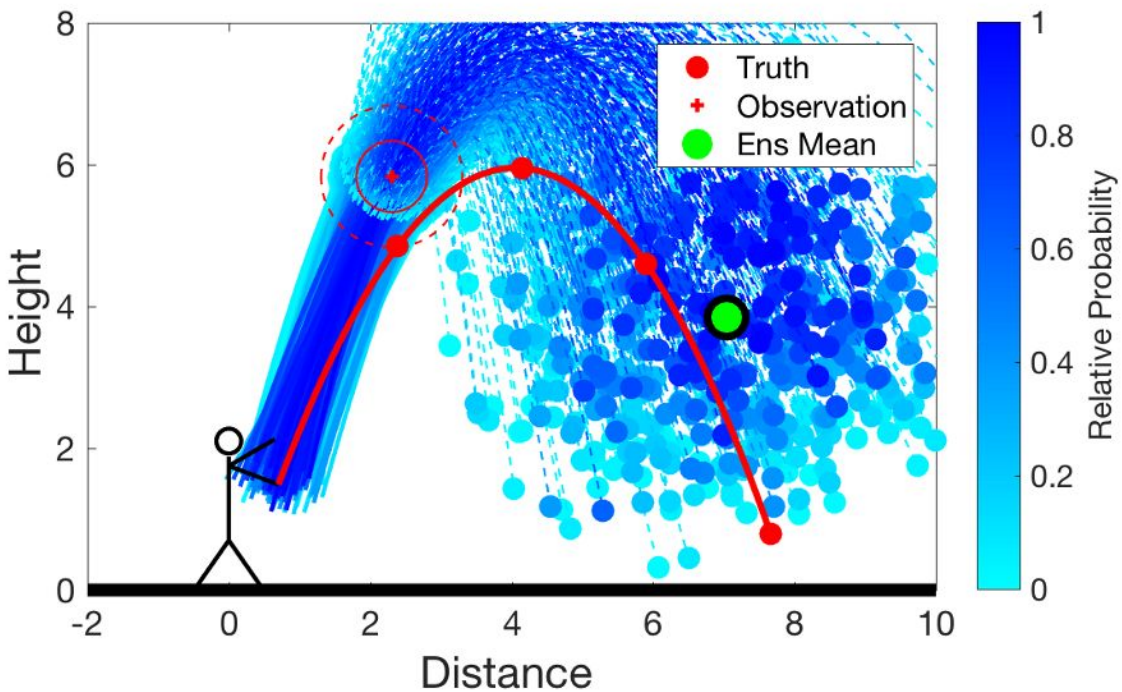
Now we combine **ensemble** results with **observations** of the position of the ball (every 2 seconds) to improve estimates



Jeff Anderson

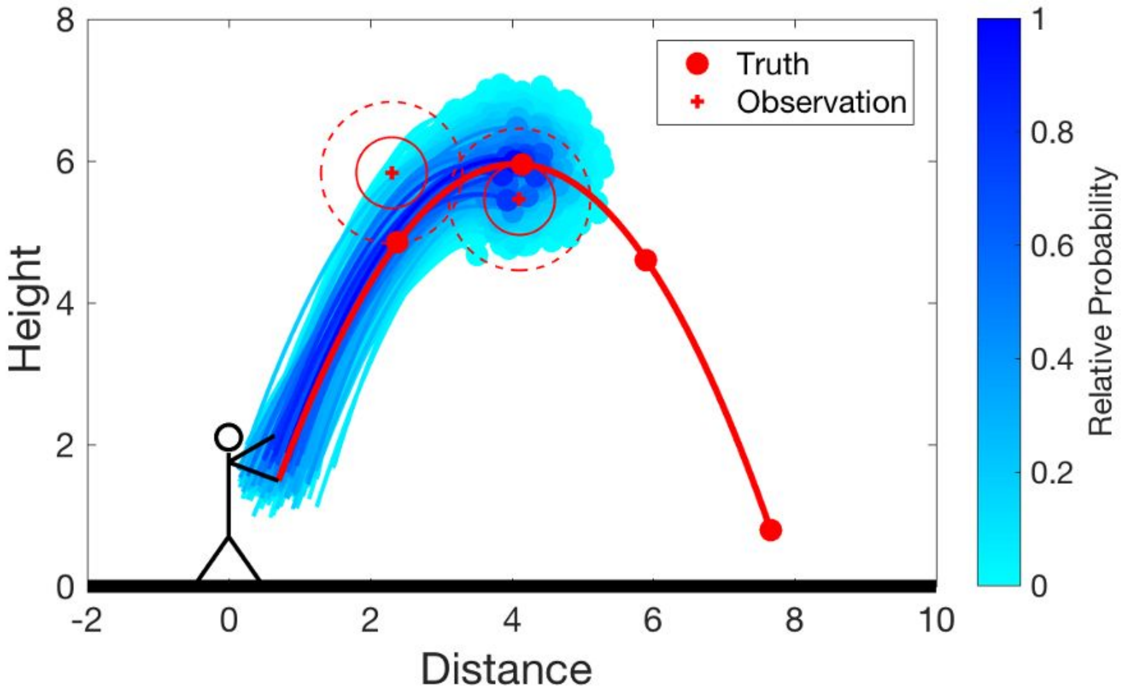
Data Assimilation: A General Description

Now we combine **ensemble** results with **observations** of the position of the ball (every 2 seconds) to improve estimates



Jeff Anderson

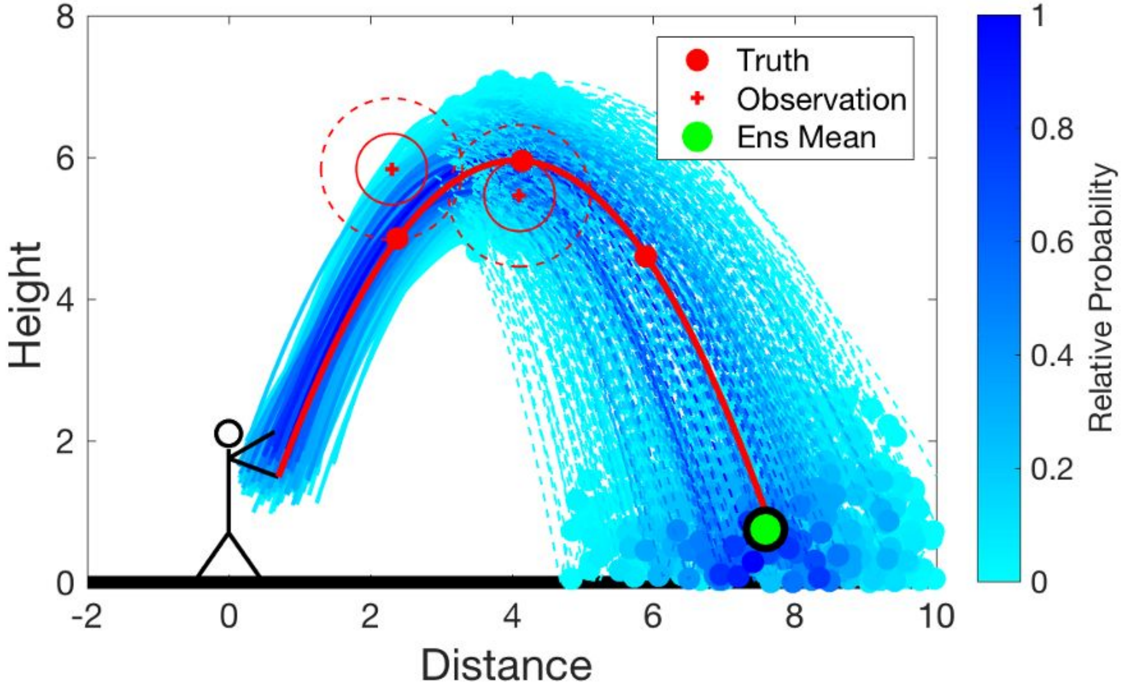
Now we combine **ensemble** results with **observations** of the position of the ball (every 2 seconds) to improve estimates



Jeff Anderson

Data Assimilation: A General Description

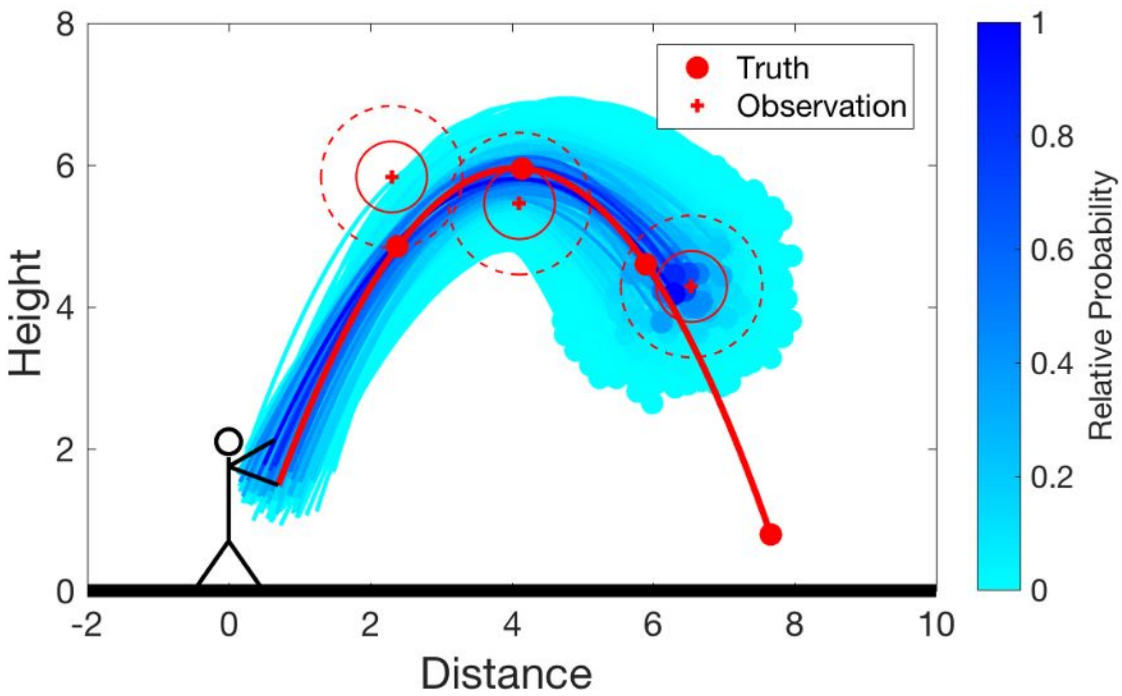
Now we combine **ensemble** results with **observations** of the position of the ball (every 2 seconds) to improve estimates



Jeff Anderson

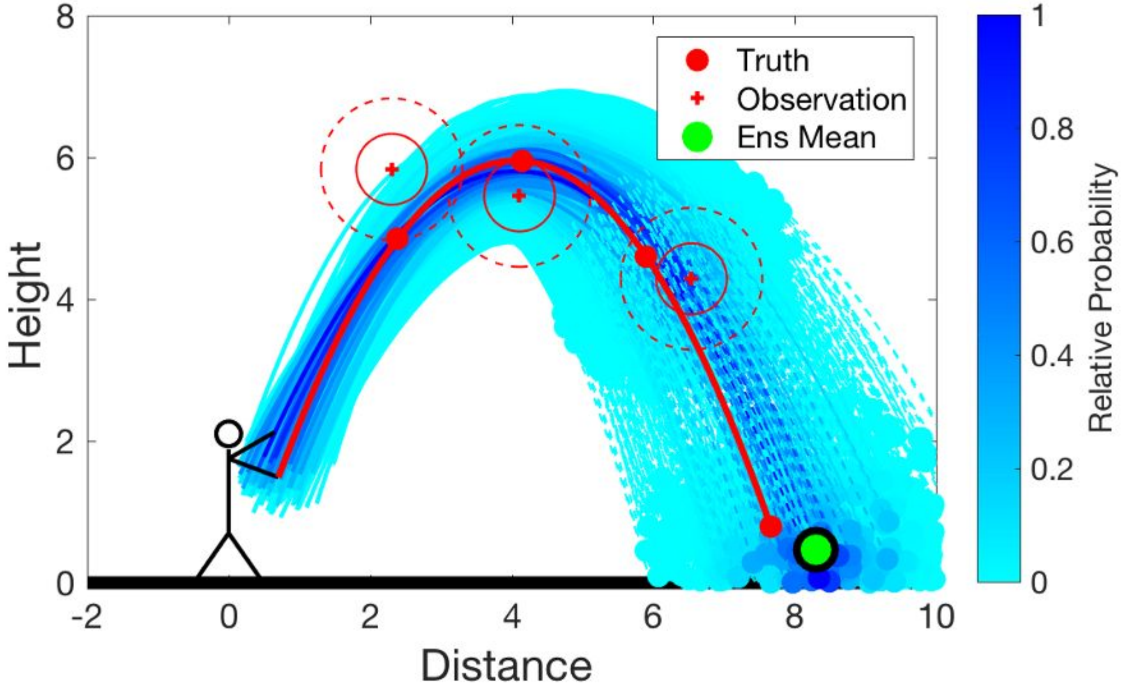
Data Assimilation: A General Description

Now we combine **ensemble** results with **observations** of the position of the ball (every 2 seconds) to improve estimates



Data Assimilation: A General Description

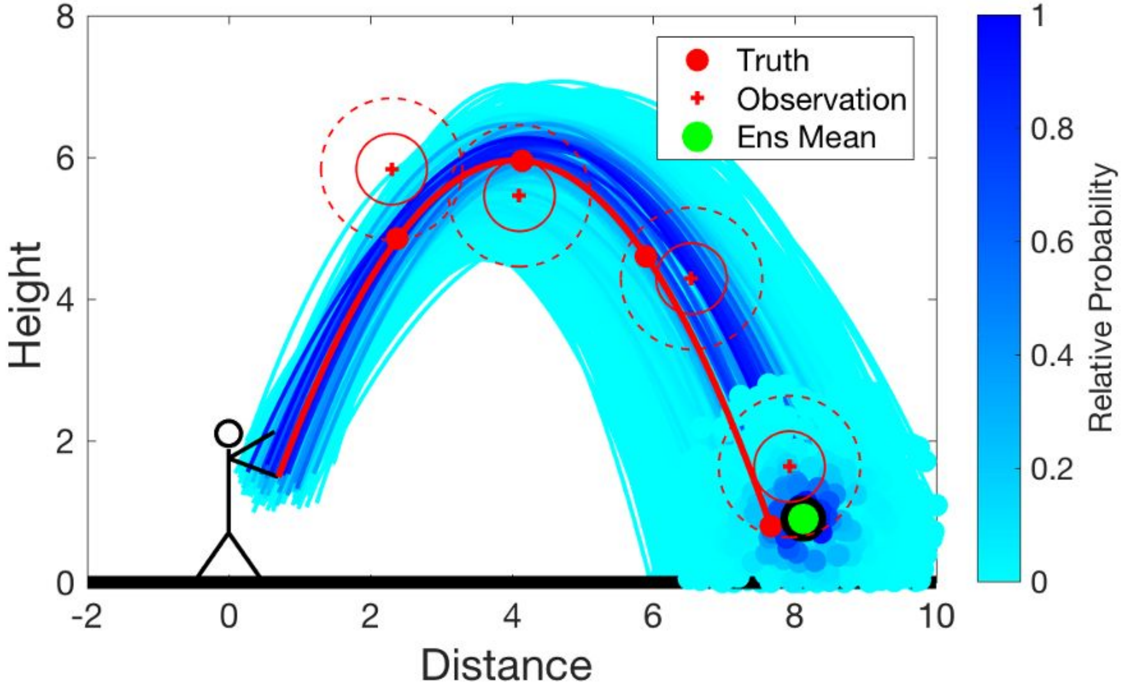
Now we combine **ensemble** results with **observations** of the position of the ball (every 2 seconds) to improve estimates



Jeff Anderson

Data Assimilation: A General Description

Now we combine **ensemble** results with **observations** of the position of the ball (every 2 seconds) to improve estimates

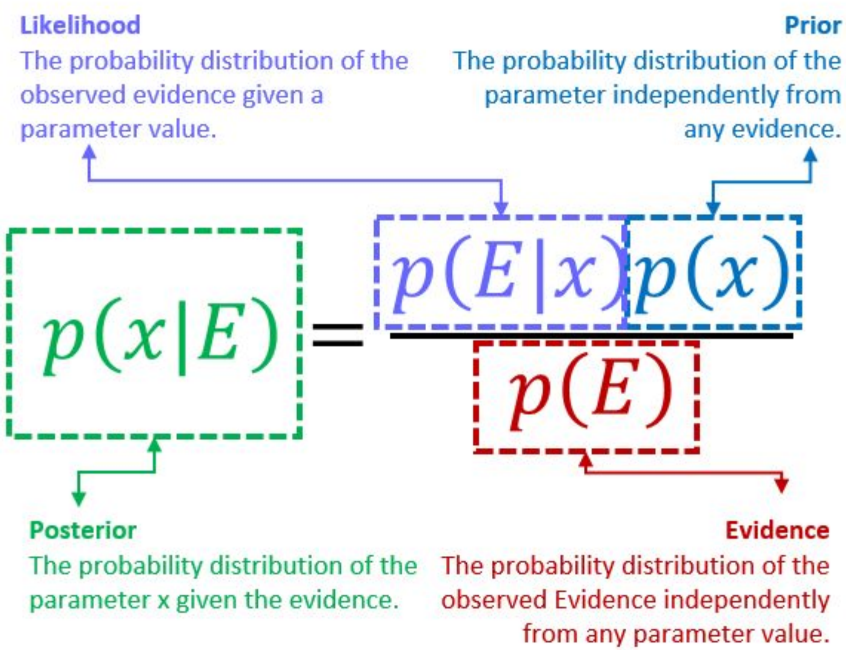


Jeff Anderson

Outline

1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz 96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

Bayes’ Theorem can be thought of a way of updating our **beliefs** on a **hypothesis** in the light of **new information**



Prior: The model forecast

Likelihood: The observations

Evidence: Normalization constant

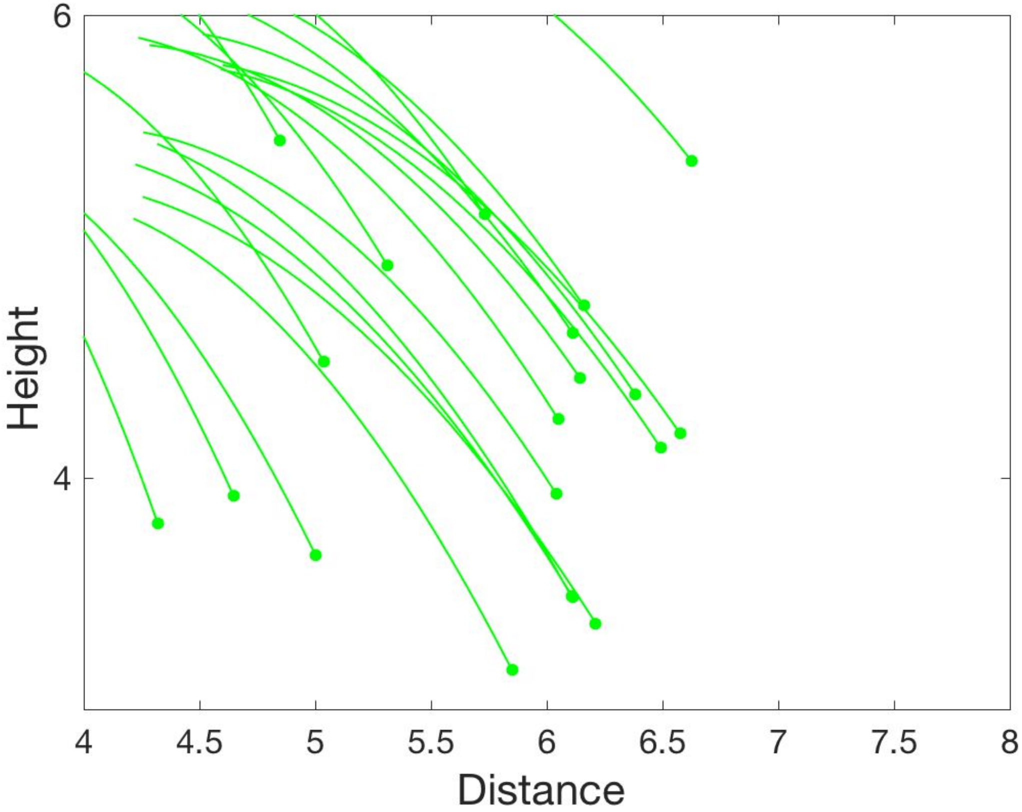
Posterior: Updated estimate

Outline

1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz 96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

Ensemble Adjustment Kalman Filter

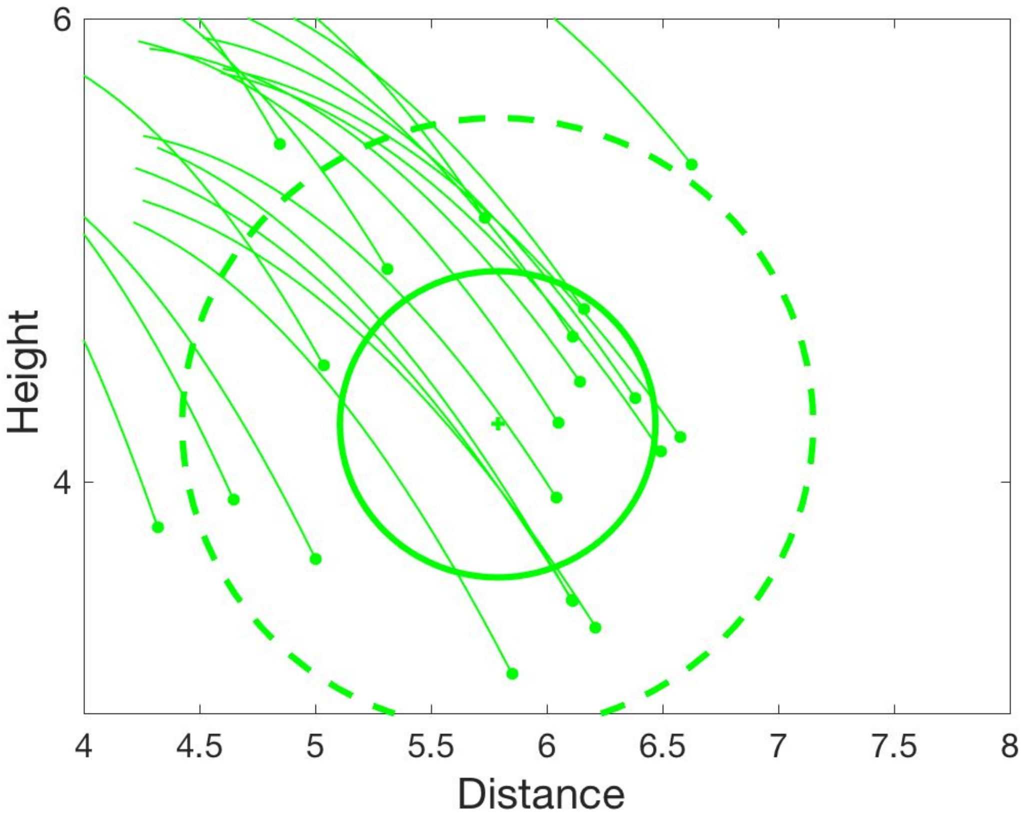
1. Prior ensemble:



Courtesy of Jeff Anderson

Ensemble Adjustment Kalman Filter

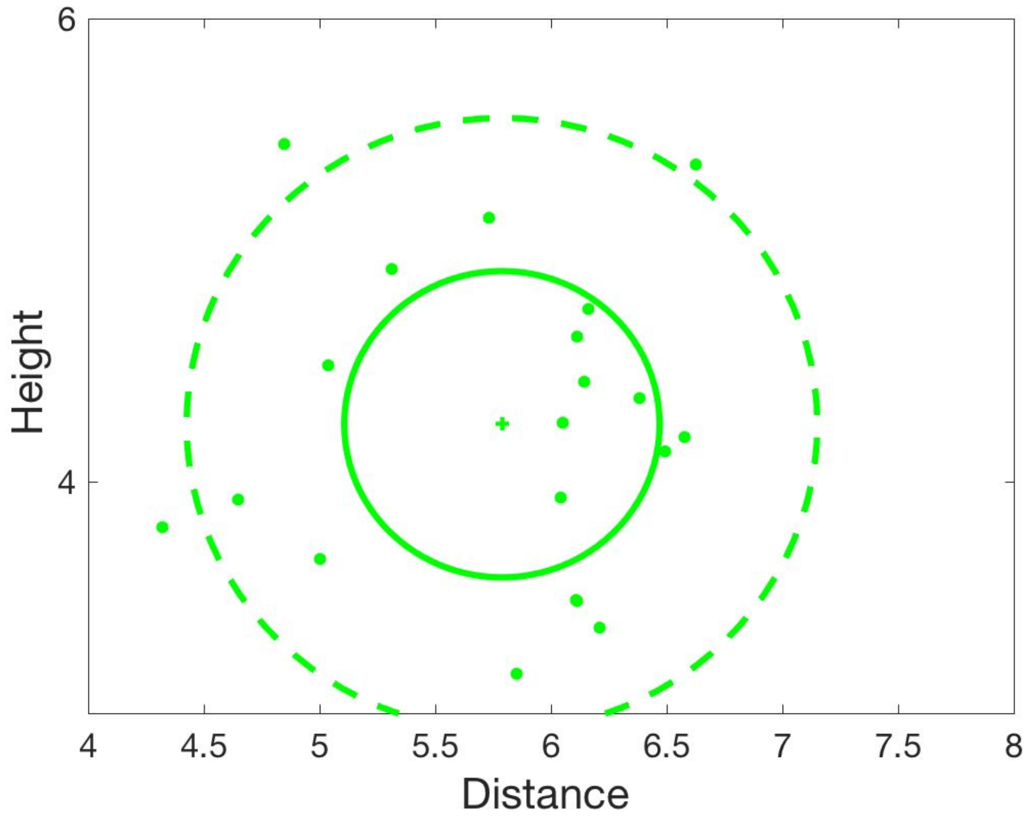
1. Prior ensemble: fit a gaussian, sample mean and covariance (pdf is green contours).



Courtesy of Jeff Anderson

Ensemble Adjustment Kalman Filter

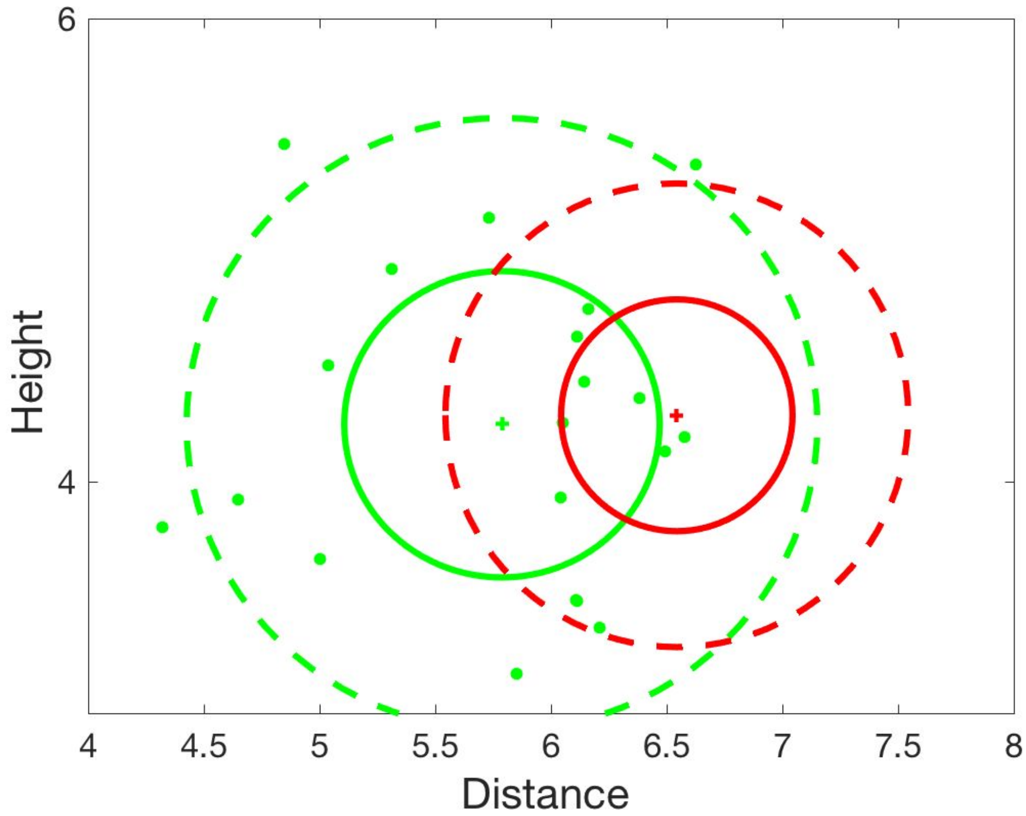
1. Prior ensemble: fit a gaussian, sample mean and covariance (pdf is green contours).



Courtesy of Jeff Anderson

Ensemble Adjustment Kalman Filter

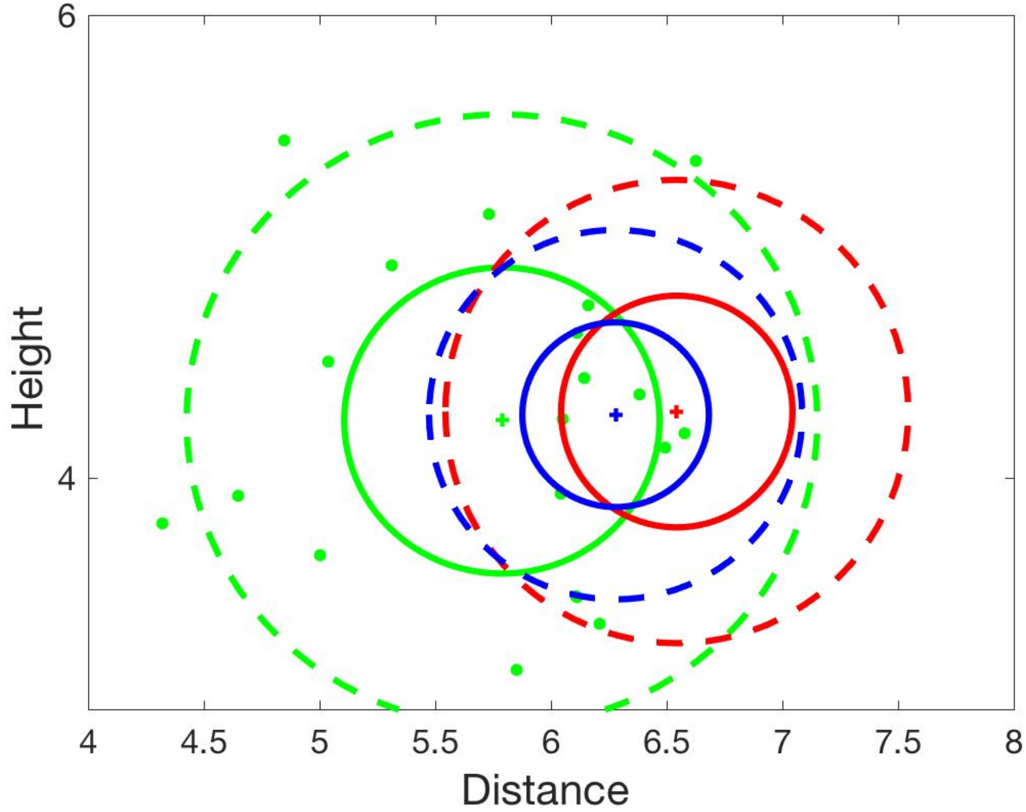
1. Prior ensemble: fit a gaussian, sample mean and covariance (pdf is green contours).
2. Kalman filter product (Bayes' Theorem) of observation likelihood (red) with prior (green)



Courtesy of Jeff Anderson

Ensemble Adjustment Kalman Filter

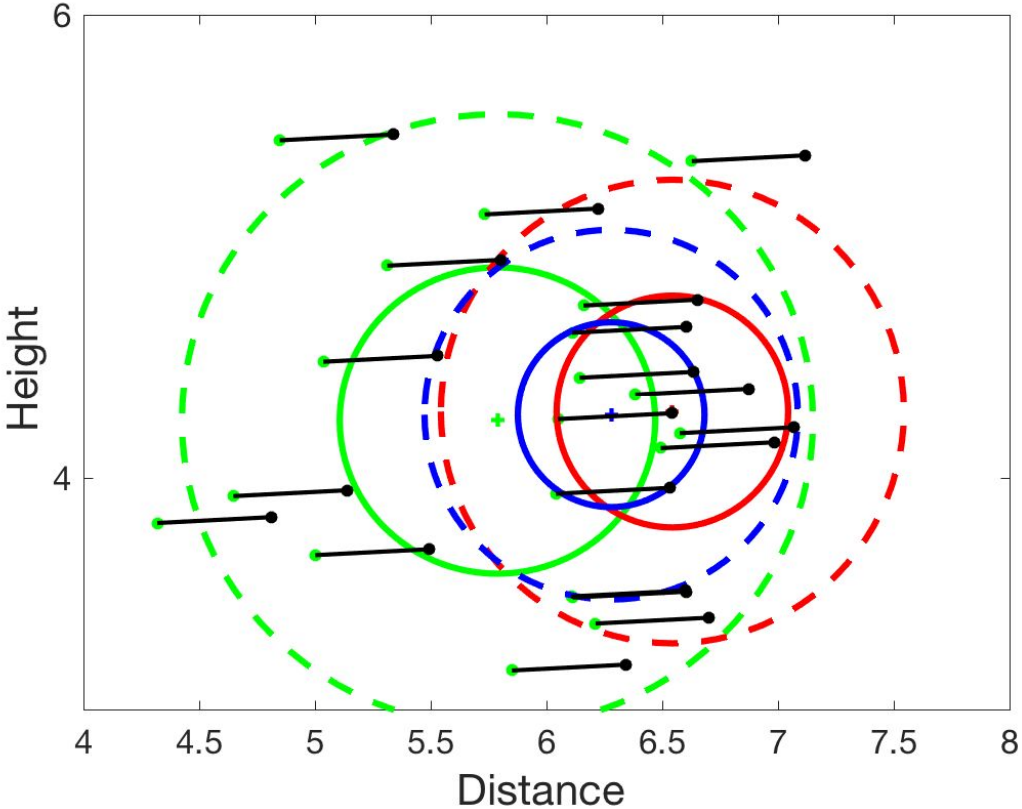
1. Prior ensemble: fit a gaussian, sample mean and covariance (pdf is green contours).
2. Kalman filter product of observation likelihood (red) with prior (green) to get gaussian posterior (blue).



Courtesy of Jeff Anderson

Ensemble Adjustment Kalman Filter

1. Prior ensemble: fit a gaussian, sample mean and covariance (pdf is green contours).
2. Kalman filter product of observations (red) with prior (green) to get gaussian posterior (blue).

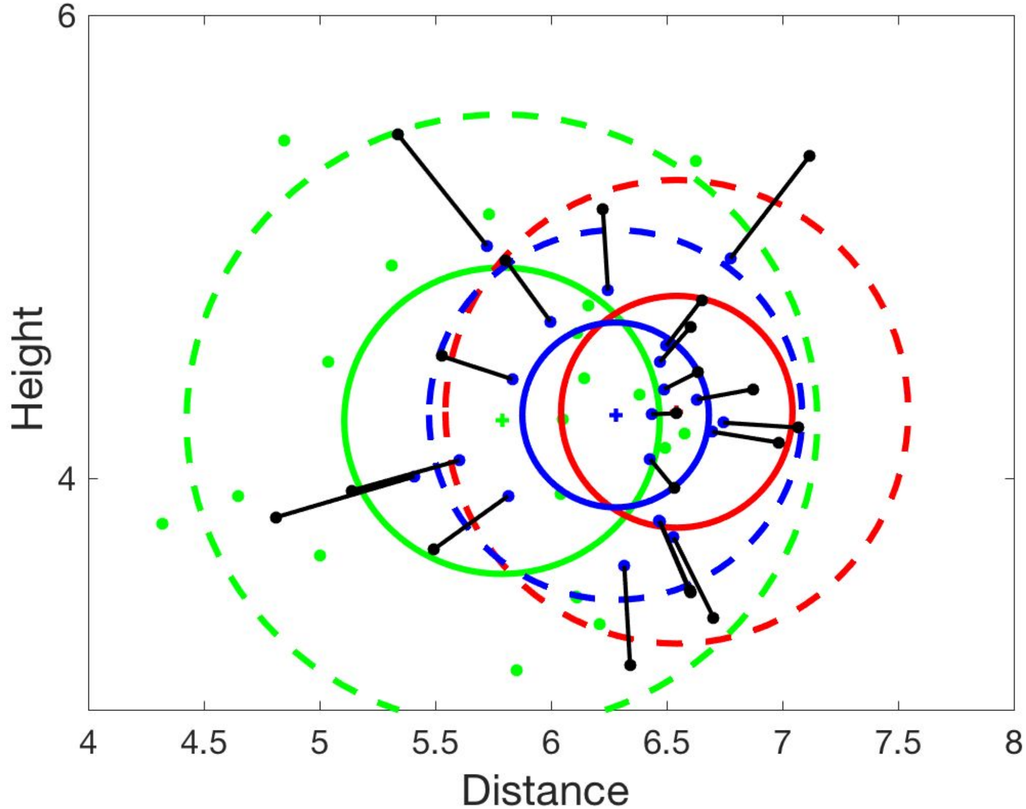


3. Shift ensemble members to have posterior mean.

Courtesy of Jeff Anderson

Ensemble Adjustment Kalman Filter

1. Prior ensemble: fit a gaussian, sample mean and covariance (pdf is green contours).
2. Kalman filter product of observations (red) with prior (green) to get gaussian posterior (blue).



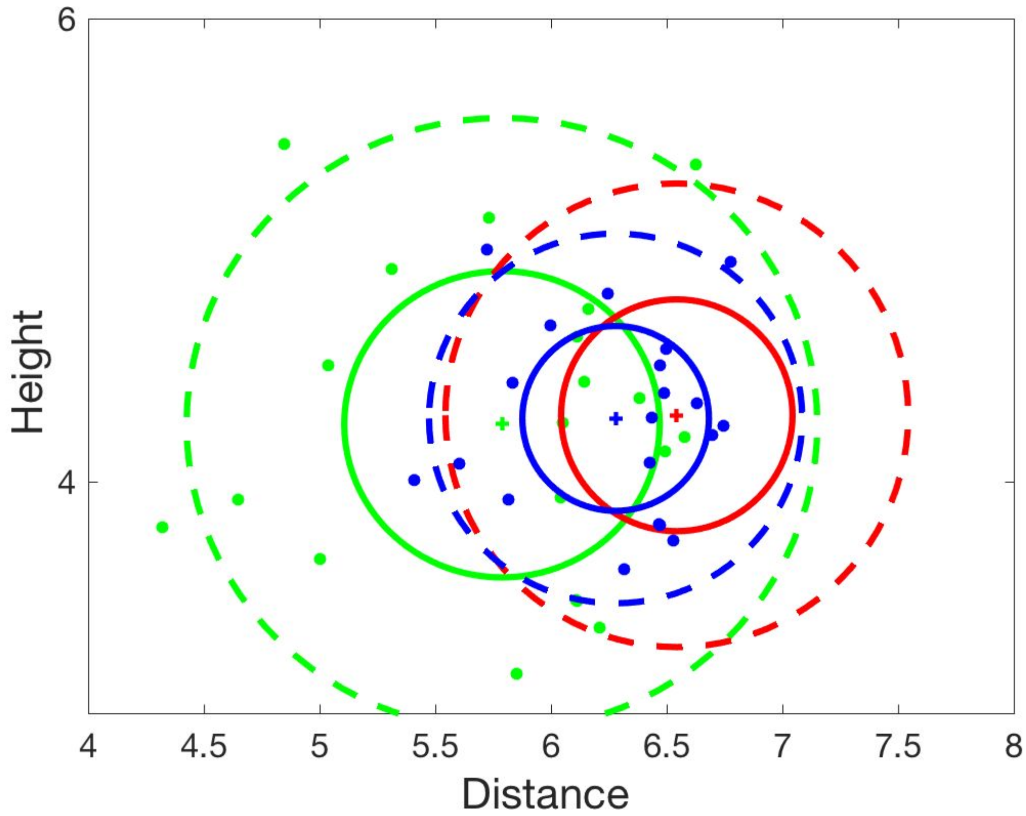
3. Shift ensemble members to have posterior mean.

4. Compact ensemble to have posterior covariance.

Courtesy of Jeff Anderson

Ensemble Adjustment Kalman Filter

1. Prior ensemble: fit a Gaussian, sample mean and covariance (pdf is green contours).
2. Kalman filter product of observations (red) with prior (green) to get Gaussian posterior (blue).



3. Shift ensemble members to have posterior mean.

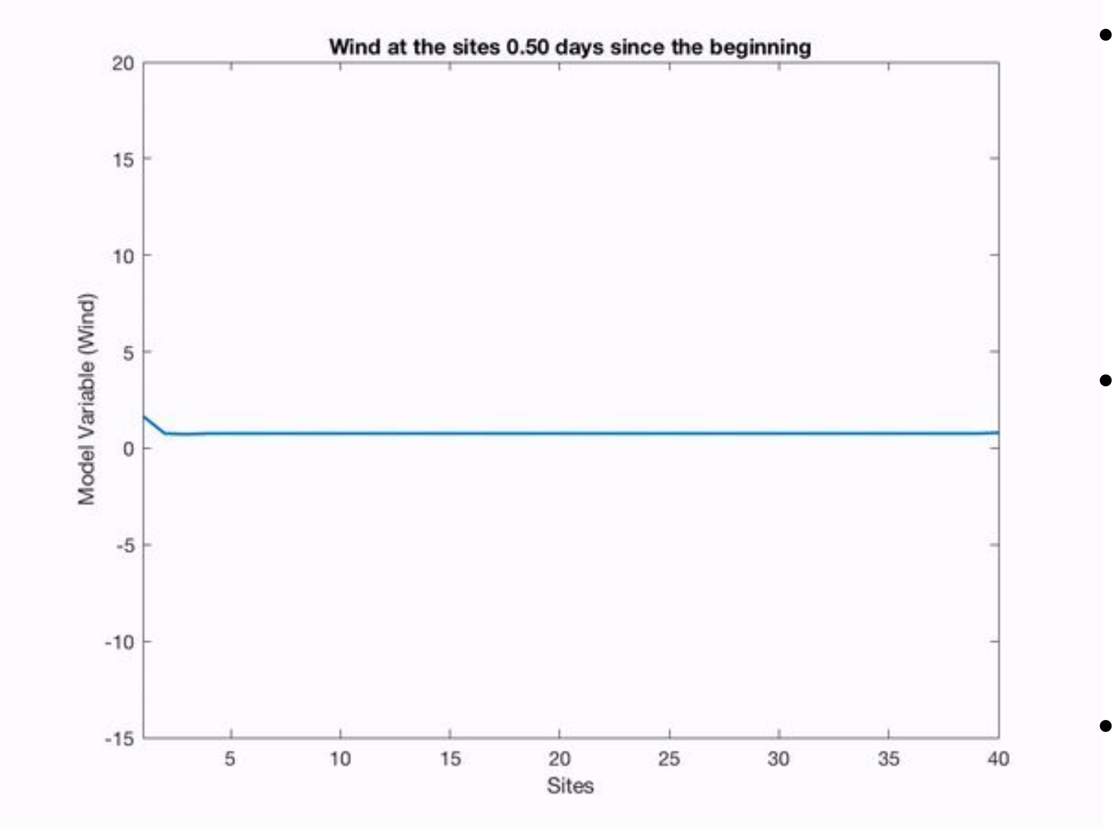
4. Compact ensemble to have posterior covariance.

Courtesy of Jeff Anderson

Outline

1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz 96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

The Model – The Lorenz 96



- A dynamical system constructed by Edward Lorenz as a problem for numerical weather prediction
- Describes a single scalar quantity as it evolves (through forcing, dissipation, and advection) on a circular array of sites
- Commonly used as model problem in data assimilation

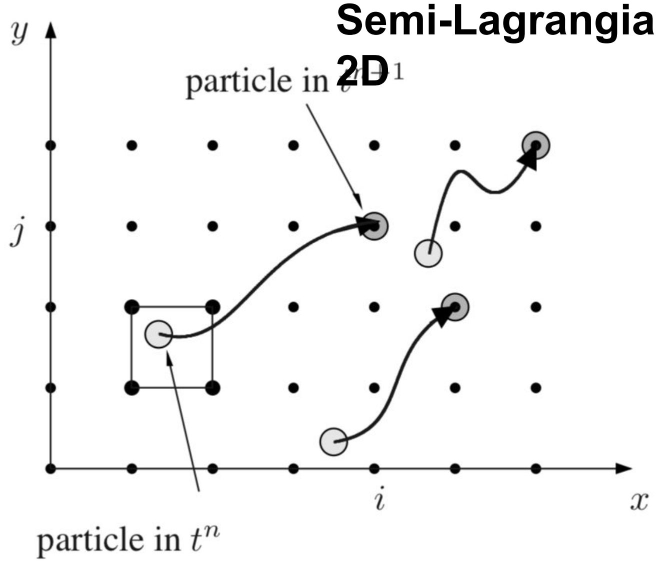
Outline

1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz 96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

The Model – Semi-Lagrangian Advection

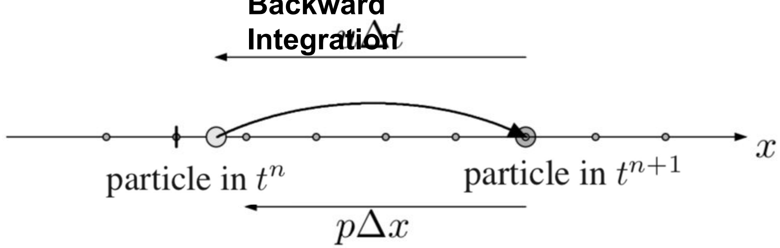
The **Semi-Lagrangian** scheme is used to model how tracer particles get distributed upstream across the grids by the Lorenz 96 winds

Semi-Lagrangian in 2D



Semi-Lagrangian in 1D

Backward Integration



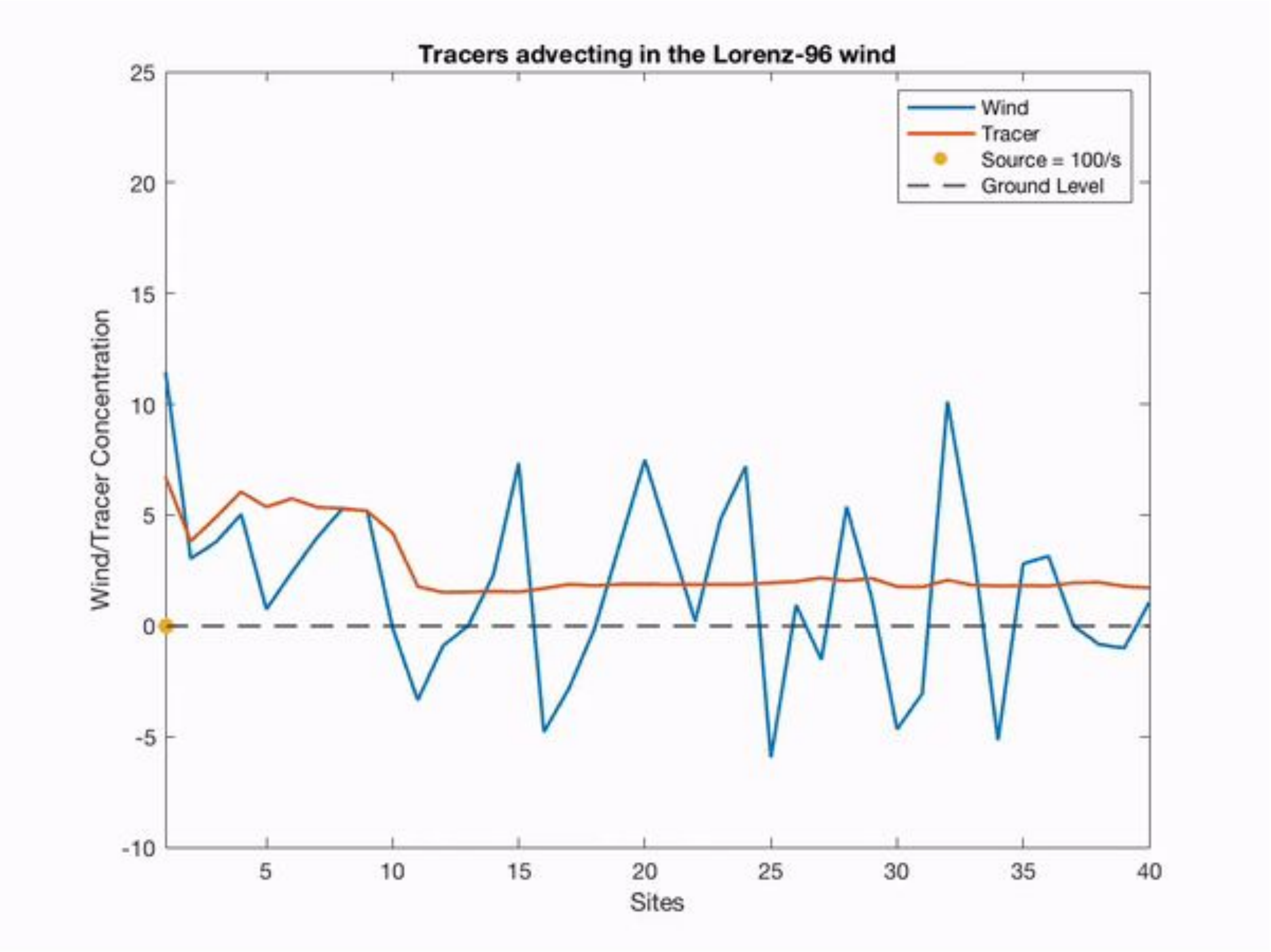
- Tracer particle land on predefined gridpoint at t^{n+1}
- Tracer particle trajectory backward integrated by one time step to time t^n , often landing between gridpoints
- The concentration of tracer at t^{n+1} is determined by linearly interpolating tracer concentration at the position during time t^n

Cushman-Roisin & Beckers

The Model – Semi-Lagrangian Advection

The **Semi-Lagrangian** scheme is used to model how tracer particles get distributed upstream across the grids by the Lorenz 96 winds

- Source at site 1 with rate of 100/s
- Exponential sinks at every site



Outline

1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz 96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

Results – Assimilated Timeseries

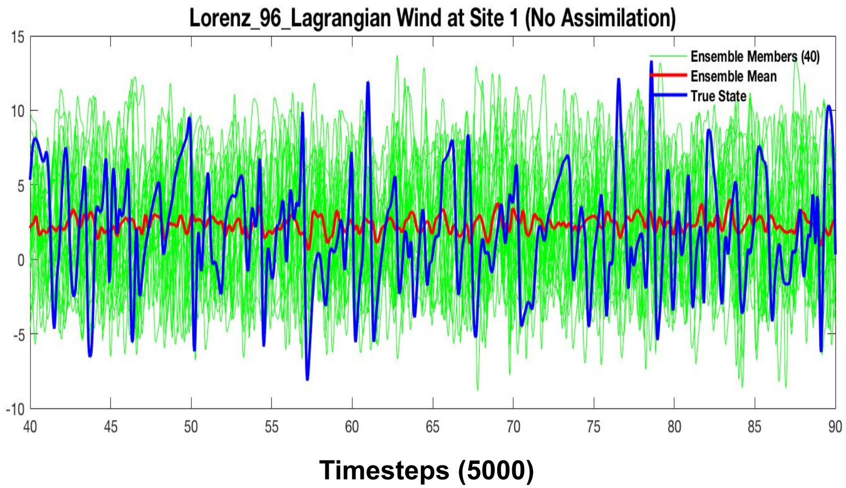
Nature of the Observations Assimilated

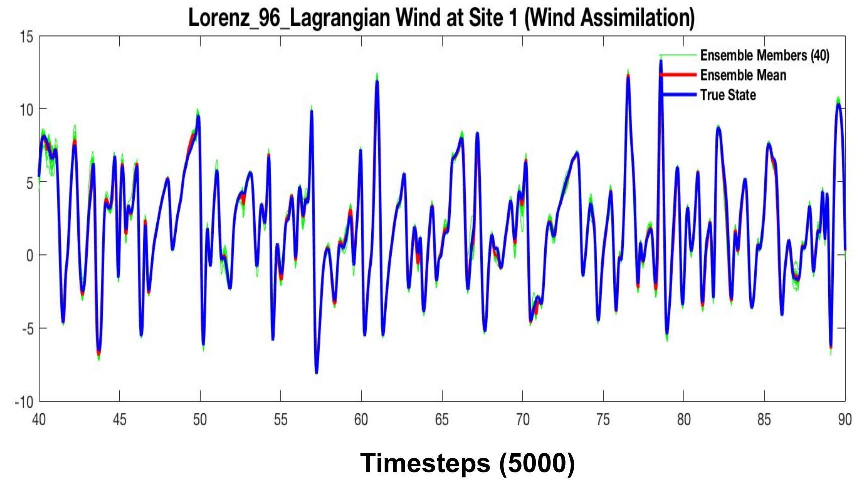
- All observations were synthetic
- 40 observations for wind at each timestep
- 40 observations for tracer concentration at each time step
- Observations were randomly distributed among the 40 Sites

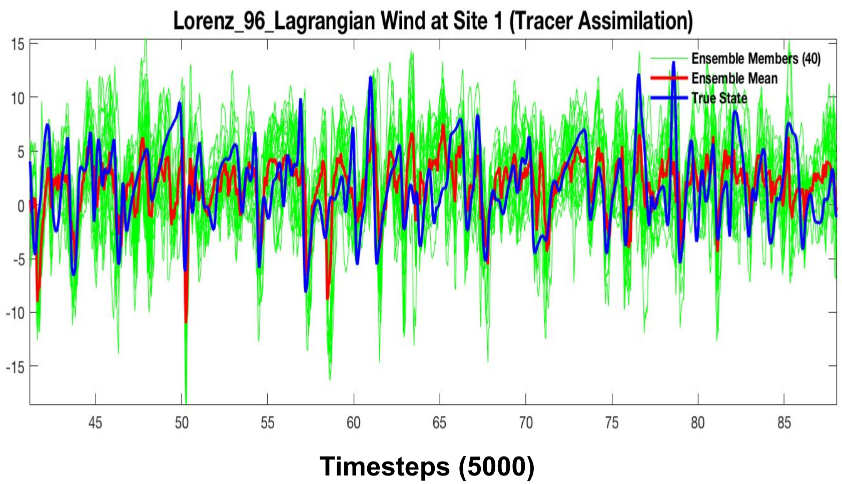
Assimilation Settings

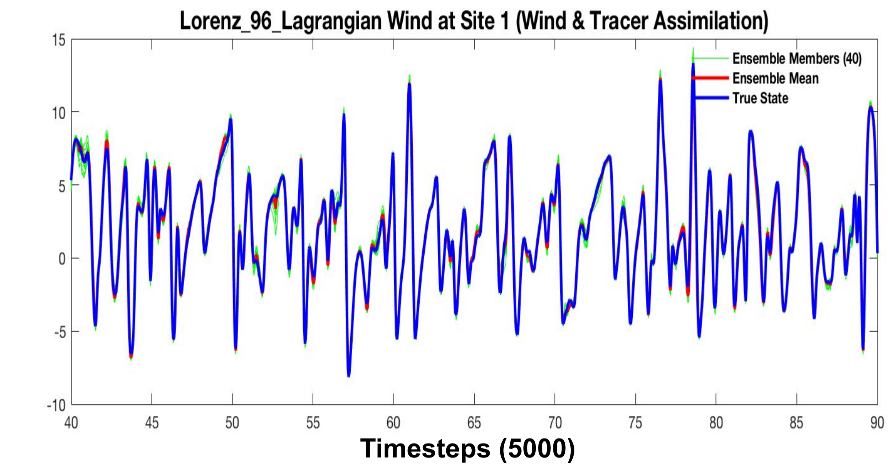
- 40 ensemble members for each assimilation run
- Maximum allowed localization – 1.2
- Localization cutoff – 0.2
- Timestep - 0.05
- Forcing - 8

Assimilation Results of Wind Values at Site 1 for 50 Timesteps



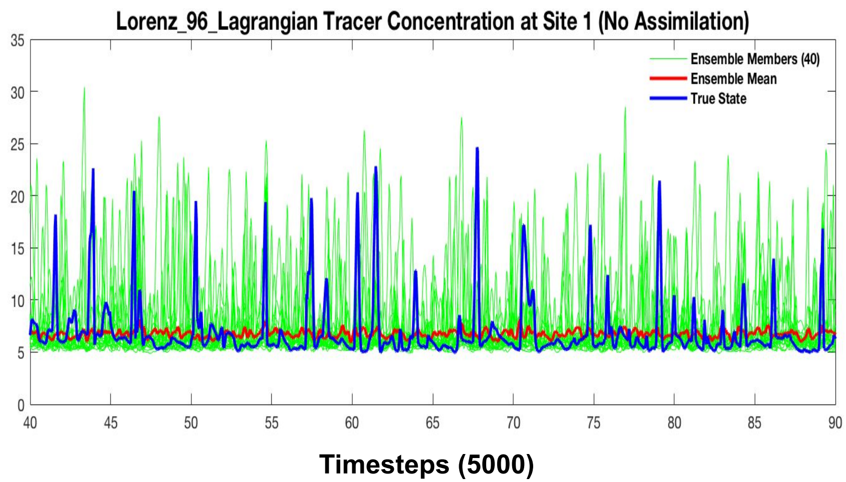


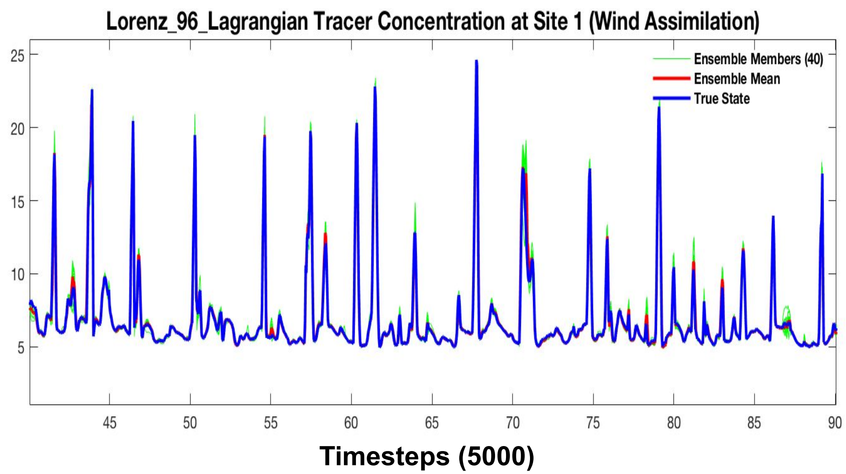


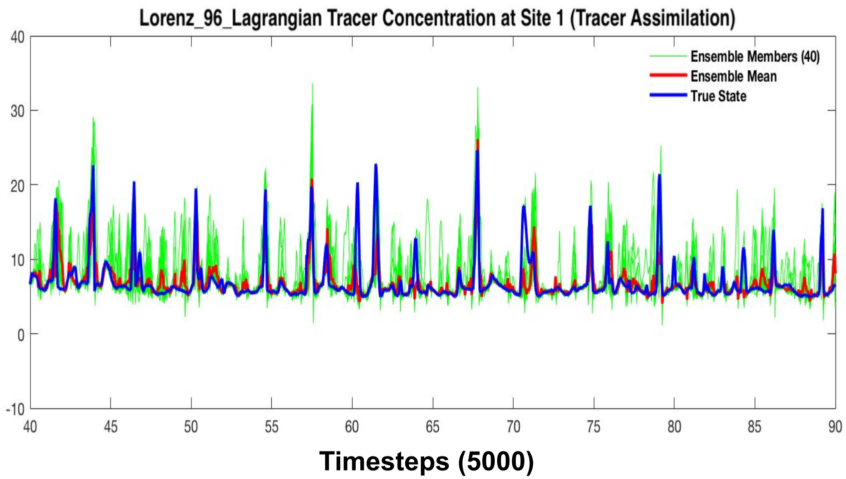


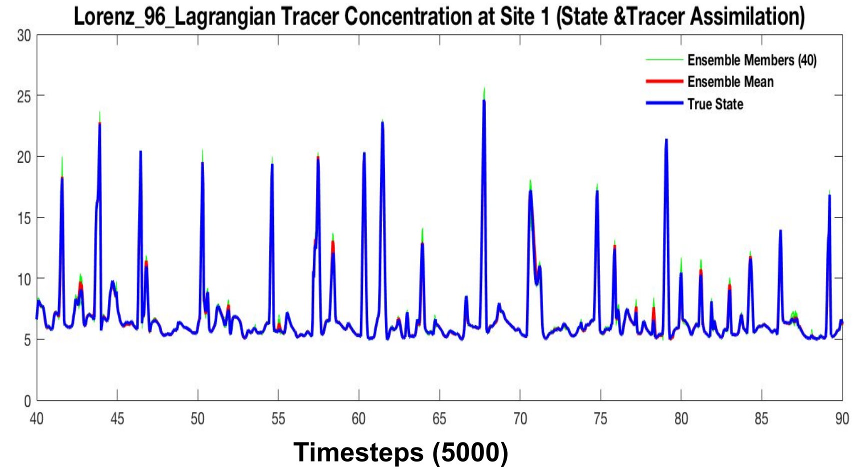
Results – Assimilated Timeseries

Assimilation Results of Tracer Concentration Values at Site 1 for 50 Timesteps









Outline

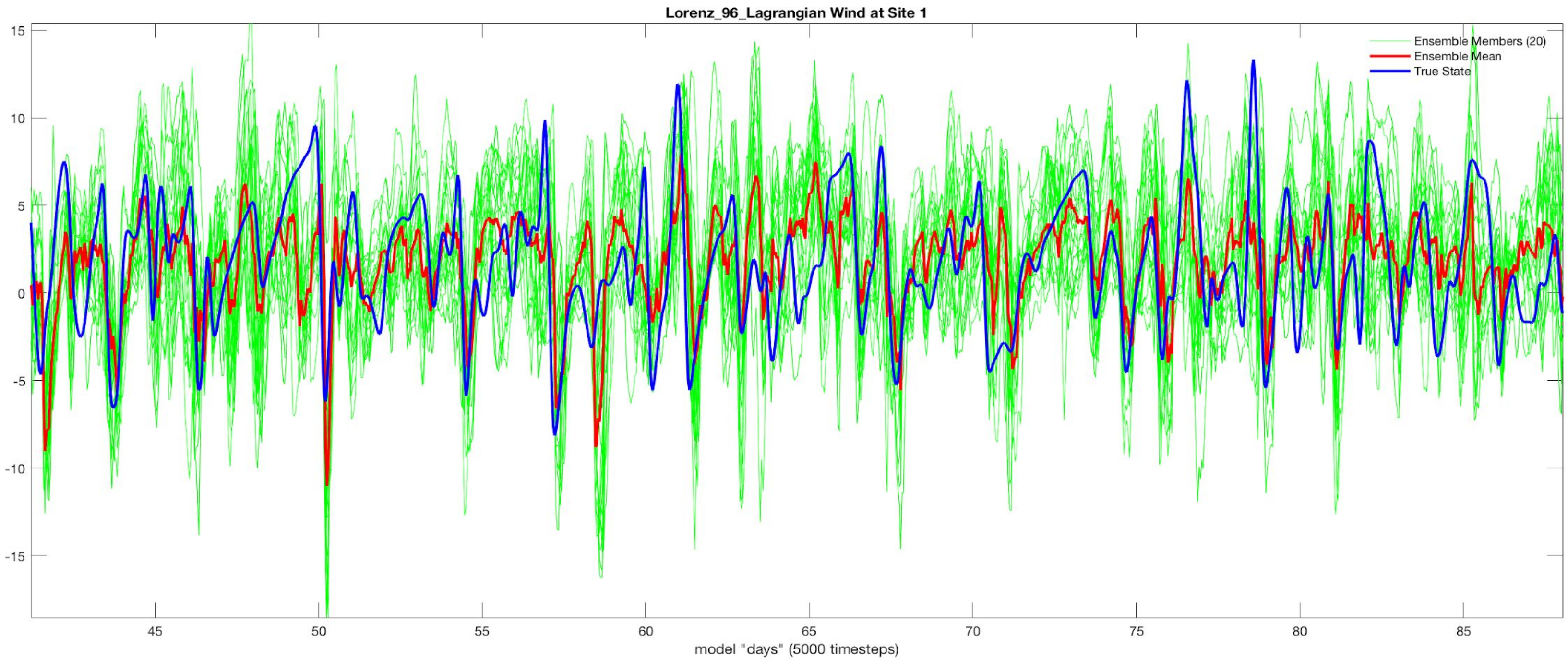
1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz 96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

Results – Assimilation Error

Run Type	State Ensemble Mean Error	State Ensemble Mean Spread	Tracer Ensemble Mean Error	Tracer Ensemble Mean Spread
No Assimilation	23.1688	22.9384	7.9196	9.1916
Assimilate State Obs	1.7656	1.9191	1.043	1.1736
Assimilate Tracer Obs	22.8191	22.4523	3.7832	3.7914
Assimilate State and Tracer Obs	1.6284	1.7289	0.71986	0.7956

Results – Assimilation Error

Only Assimilating Tracer Observations (Zoomed in at Wind



Ensemble Mean Error - 22.8191

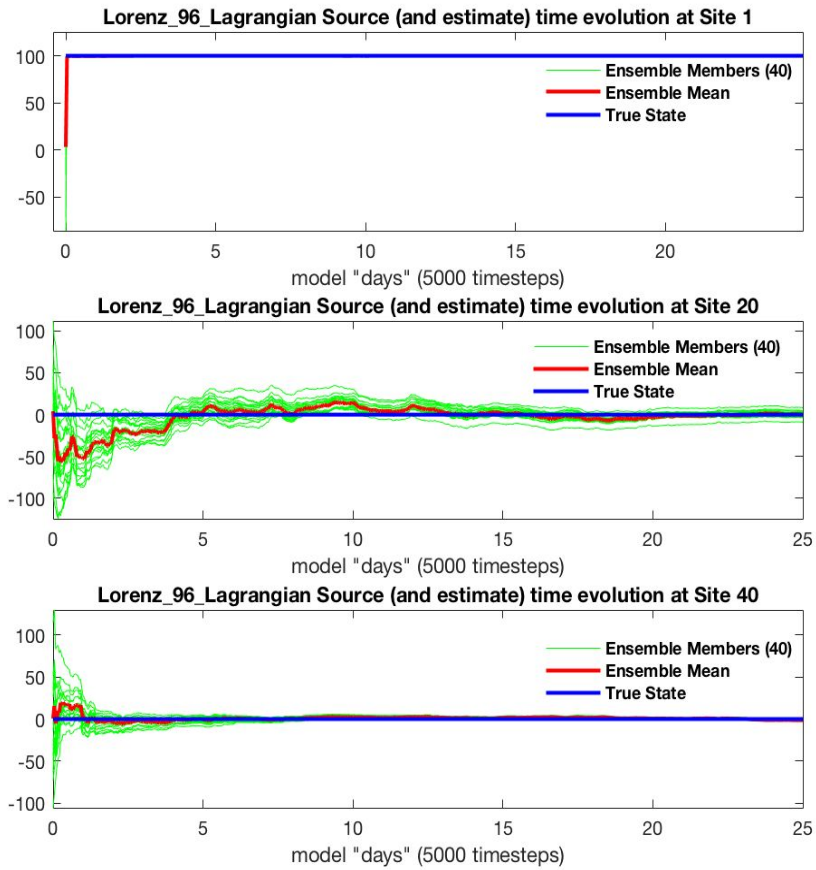
Outline

1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz 96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

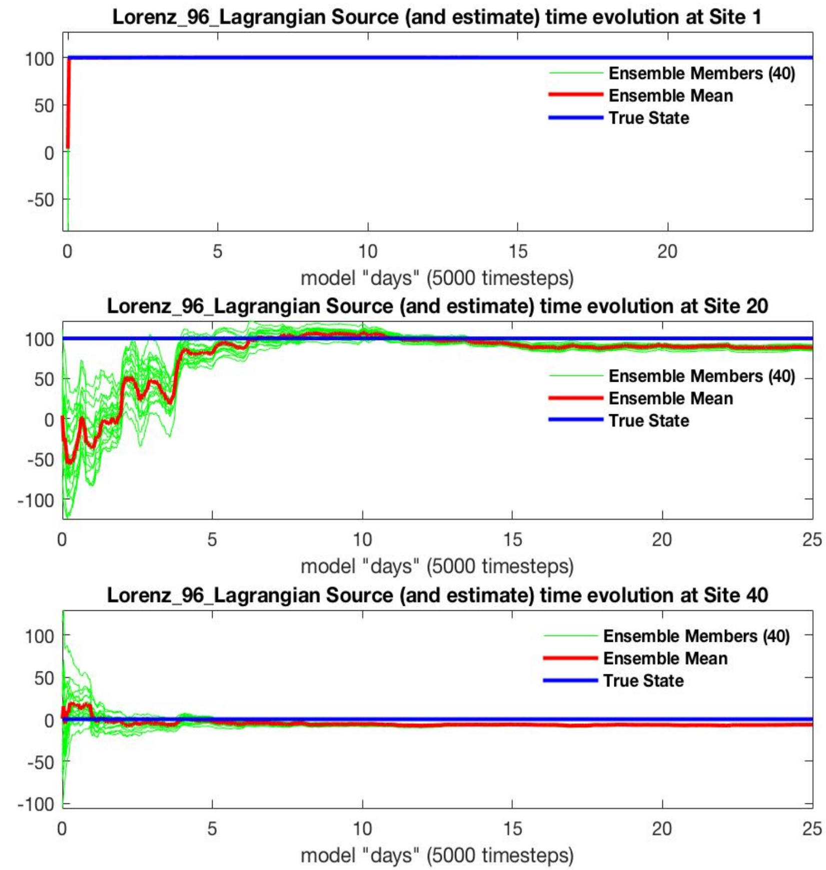
Results – Source Characterization

Characterizing source location and rate (of 100) by assimilating wind and tracer observations

Source at Site 1



Source at Site 1 and 20



Outline

1. Motivation
2. Data Assimilation: A general description
3. Bayes' Theorem
4. Ensemble Adjustment Kalman Filter
5. The Model – The Lorenz 96
6. The Model – Semi-Lagrangian Advection
7. Results – Assimilated Timeseries
8. Results – Assimilation Error
9. Results – Source Characterization
10. Further Steps

Further Steps

- Exploring source characterization capabilities of data assimilation in Lorenz 96 with lower quality observations
- Implementing tracer advection in higher level circulation models
- Exploring novel assimilation techniques designed specifically for tracers

Acknowledgements

- Mentors: Jeff Anderson, Helen Kershaw
- Data Assimilation Research Section
- CODE Team
- Fellow SIParCS Interns

Any Questions?

You can reach me at:

- Email: ishraque@colgate.edu
- LinkedIn: Fairuz Ishraque