# Inflation in Ensemble Filters: Why, How and When?

## Real High-Dimensional Atmospheric and Hydrologic Applications

**Moha Gharamti**

https://dart.ucar.edu/

gharamti@ucar.edu

Date: Mar. 5, 2020

National Center for Atmospheric Research

Data Assimilation Research Section (DAReS) - TDD - CISL

CISL Computational & Information Systems Lab

siam Society for Industrial and Applied Mathematics

NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

# The questions: Why, How, & When?

# 1. Ensemble Covariance Inflation: Why?

- Simply, because we're not in optimal EnKF settings:
    1. Highly nonlinear models
    2. Massive model dimensions forces us to use small ensemble sizes; can never really satisfy this:

$$\boxed{\lim_{N\to\infty} \widehat{\mathbf{P}} = \mathbf{B}}$$

    3. Deal with many non-Gaussian phenomena (e.g., precipitation)
    4. Unavoidable model errors

# 2. Ensemble Covariance Inflation: How?

**Spatially and Temporally Varying Adaptive Covariance Inflation:**

$$p(\lambda|d) \propto p(\lambda) \cdot p(d|\lambda) \tag{1}$$

- **Prior** $p(\lambda)$; assumed Inverse Gamma
- **Likelihood** $p(d|\lambda)$; a Gaussian density where
  - $d = |y^o - \overline{x}_b|$ is the innovation
  - formulated using innovation statistics [Derosiers et al. 2005]
    $\mathbb{E}(d) = 0; \quad \mathbb{E}(d^2) = \sigma_o^2 + \lambda \widehat{\sigma_b}^2$
- **Posterior** $p(\lambda|d)$

# 2. Ensemble Covariance Inflation: How?

**Spatially and Temporally Varying Adaptive Covariance Inflation:**

$$p(\lambda|d) \propto p(\lambda) \cdot p(d|\lambda) \tag{1}$$

- **Prior** $p(\lambda)$; assumed Inverse Gamma
- **Likelihood** $p(d|\lambda)$; a Gaussian density where
  - $d = |y^o - \overline{x}_b|$ is the innovation
  - formulated using innovation statistics [Derosiers et al. 2005]
    $\mathbb{E}(d) = 0; \quad \mathbb{E}(d^2) = \sigma_o^2 + \lambda \widehat{\sigma_b}^2$
- **Posterior** $p(\lambda|d)$

## Characteristics

- Adaptive in time; posterior becomes prior the next DA cycle
- Varies in space (affects the rank of the covariance)
- Variance increase is proportional to the size of the innovation

# 3. Ensemble Covariance Inflation: When?

The algorithm can be used to inflate the prior covariance [Anderson 2009; El Gharamti 2018], the posterior covariance [e.g., El Gharamti et al. 2019], or both actually. So, what to do?

# 3. Ensemble Covariance Inflation: When?

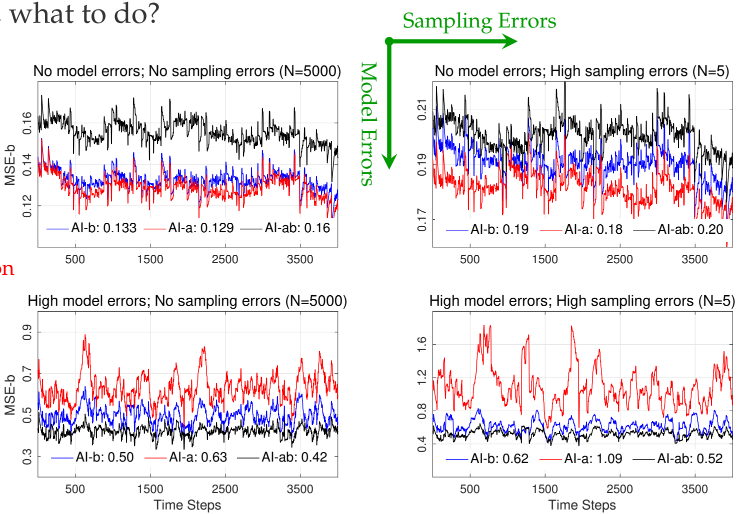
The algorithm can be used to inflate the prior covariance [Anderson 2009; El Gharamti 2018], the posterior covariance [e.g., El Gharamti et al. 2019], or both actually. So, what to do?

Answer: It depends!

The algorithm can be used to inflate the prior covariance [Anderson 2009; El Gharamti 2018], the posterior covariance [e.g., El Gharamti et al. 2019], or both actually. So, what to do?

Lorenz-63 System

◇ prior inflation
◇ posterior inflation
◇ both



# Application I: Atmospheric DA

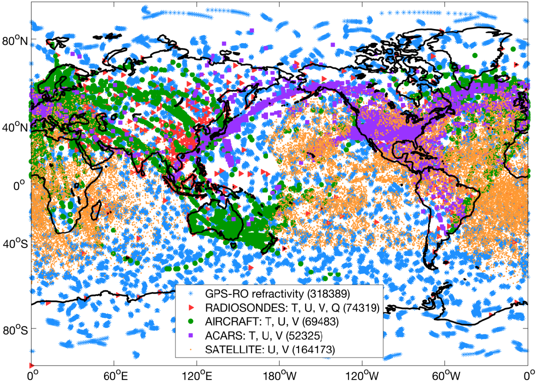
# 4.1 Atmospheric DA: Configuration

- The Community Atmosphere Model (CAM; Neal et al. 2013)
- The Data Assimilation Research Testbed (DART; Anderson et al. 2003)

# 4.1 Atmospheric DA: Configuration

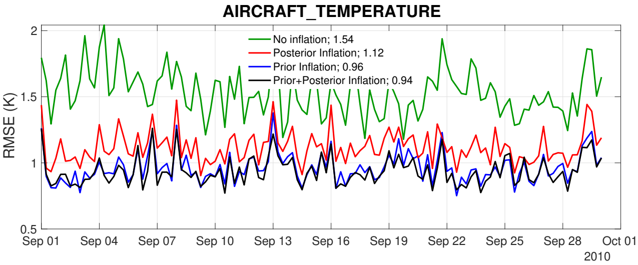
- The Community Atmosphere Model (CAM; Neal et al. 2013)
- The Data Assimilation Research Testbed (DART; Anderson et al. 2003)

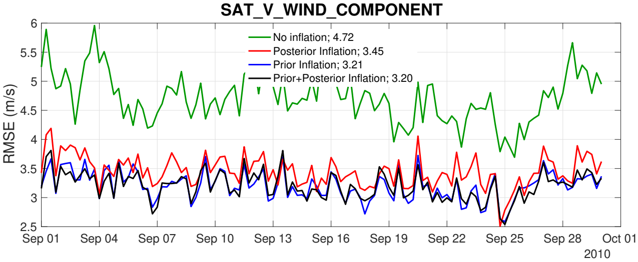
- $2°$ model + 26 levels
- 80 members; 6 weeks
- Localization: GC ~ 960 km
- Variables: `PS`, `T`, `U`, `V`, `Q`, ..




Typical observations assimilated every 6 hours over CONUS

# 4.2 Atmospheric DA: Obs-space diagnostics

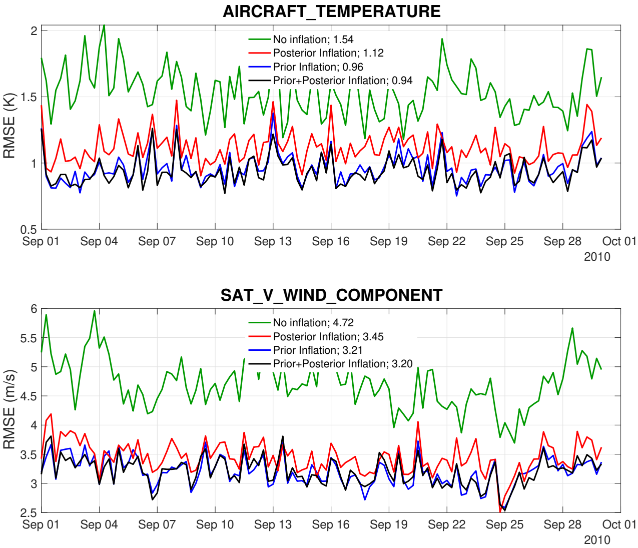



Time-series of T and V over the tropics (i.e., −20:20° N) and averaged over all layers

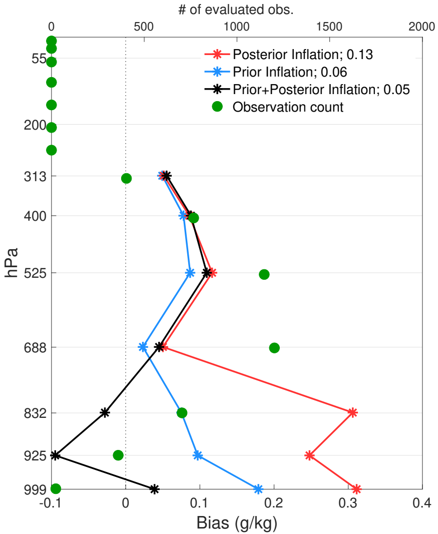
# 4.2 Atmospheric DA: Obs-space diagnostics

- Failing to use inflation yields low-quality estimates
- Posterior inflation performs fairly well
- Prior inflation outperforms posterior inflation
- Best accuracy is obtained after combining both inflation schemes




Time-series of T and V over the tropics (i.e., −20:20° N) and averaged over all layers

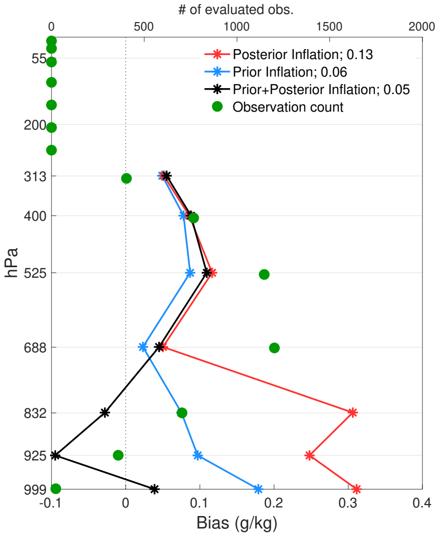
# 4.2 Atmospheric DA: Obs-space diagnostics




Bias profile of Q over the Northern Hemisphere (i.e., 20:90° N)

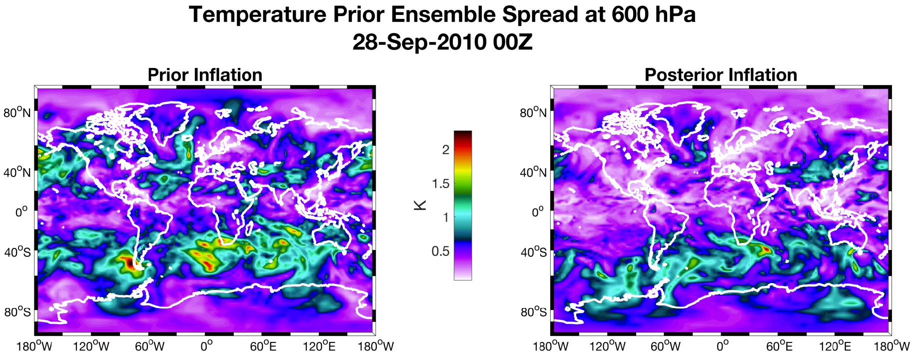
# 4.2 Atmospheric DA: Obs-space diagnostics

- Radiosonde specific humidity (Q) was not assimilated. It was kept aside for verification purposes only
- Largest biases are observed near the surface
- Prior inflation is more effective than posterior inflation at mitigating the bias



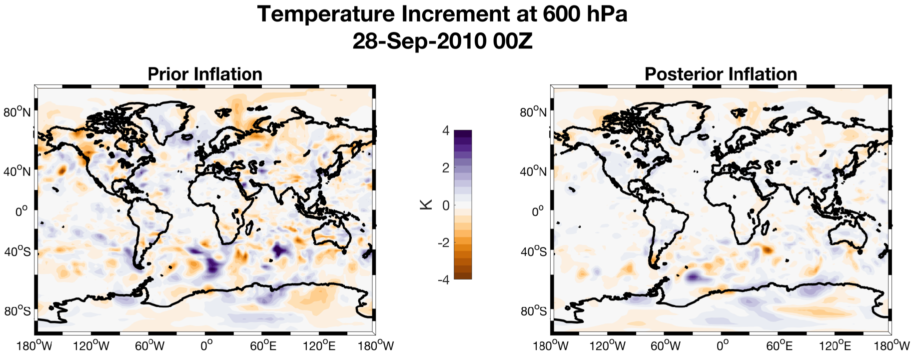

Bias profile of Q over the Northern Hemisphere (i.e., 20:90° N)

# 4.3 Atmospheric DA: State-space diagnostics

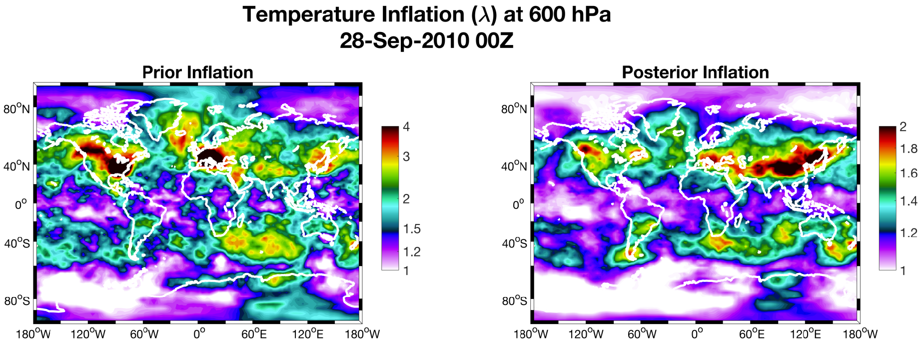



- Largest uncertainties are present in the Southern Ocean (sparsely observed)
- Prior inflation yields larger ensemble spread than posterior inflation at this elevation

# 4.3 Atmospheric DA: State-space diagnostics




- Largest increments where the ensemble spread is high
- Larger DA increments suggested by prior inflation

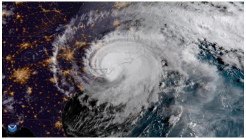



- Prior inflation is the largest in the vicinity of the observations (e.g., CONUS, Europe)
- Posterior inflation could point to locations where sampling error is the largest?

# Application II: Flood Prediction

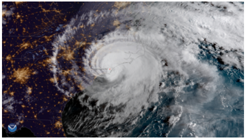
# 5.1 Hurricane Florence Flooding

- **Category 4 hurricane:** Carolinas on Sep. 14, 2018
- Precipitation exceeded 35″ in certain areas
- Caused major flooding and catastrophic damages



# 5.1 Hurricane Florence Flooding

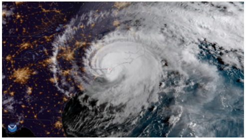
- **Category 4 hurricane:** Carolinas on Sep. 14, 2018
- Precipitation exceeded 35″ in certain areas
- Caused major flooding and catastrophic damages



∼ 50 casualties and $25B losses in coastal communities

# 5.1 Hurricane Florence Flooding

- **Category 4 hurricane:** Carolinas on Sep. 14, 2018
- Precipitation exceeded 35″ in certain areas
- Caused major flooding and catastrophic damages

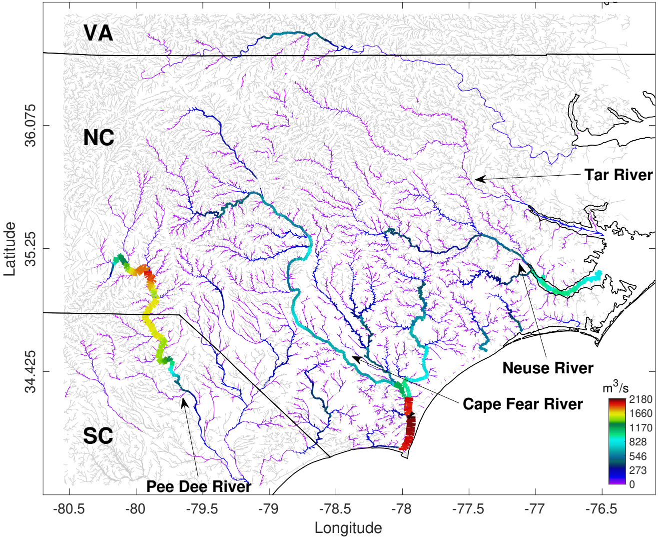


Can we enhance flood prediction using DA and available streamflow models?

$\sim$ 50 casualties and \$25B losses in coastal communities

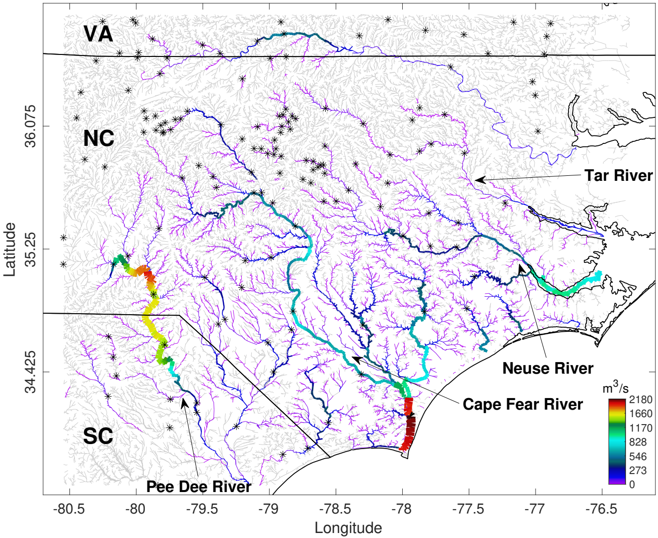
# 5.2 WRF-Hydro and DA Configuration

- Interface DART to WRF-Hydro (NOAA's NWM; Gochis, 2020)



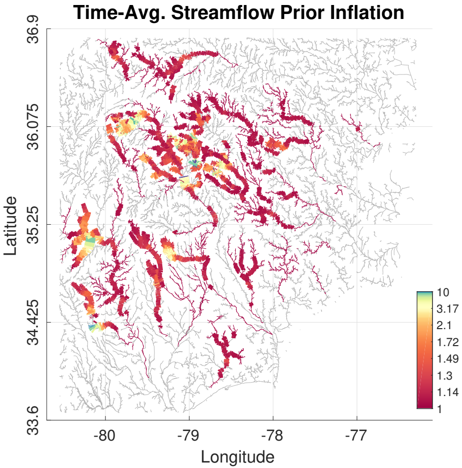

- Total of ~ 70K reaches (streams)

- Interface DART to WRF-Hydro (NOAA's NWM; Gochis, 2020)

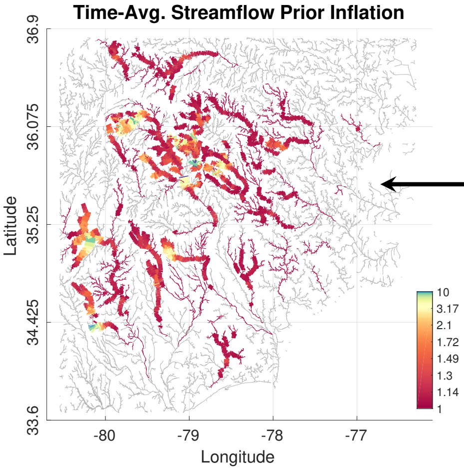


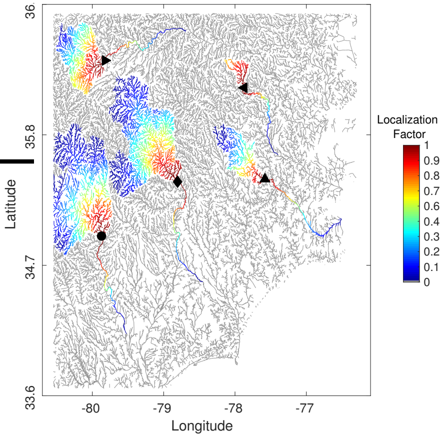

- Total of ~ 70K reaches (streams)
- Hourly streamflow assimilation
- 107 USGS gauges

# 5.3 Streamflow Inflation in Space



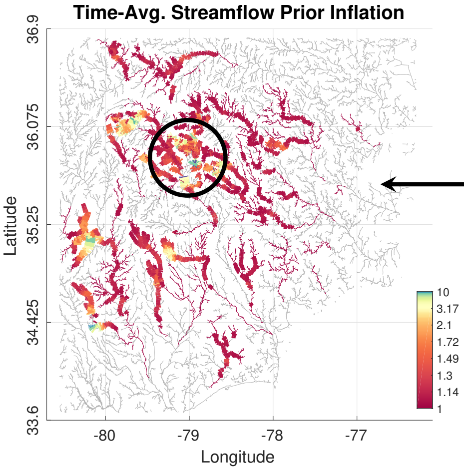
# 5.3 Streamflow Inflation in Space

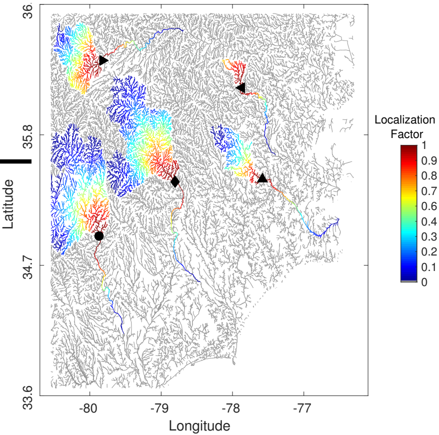



- Inflation confined in space to the stream network thanks to Along-The-Stream Localization (El Gharamti et al. 2021)

# 5.3 Streamflow Inflation in Space

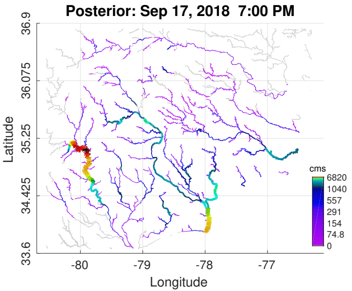


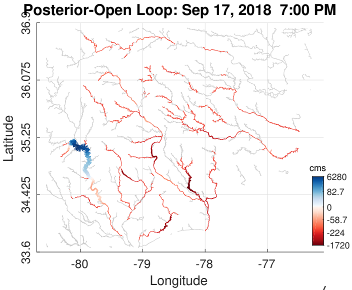




- Inflation confined in space to the stream network thanks to Along-The-Stream Localization (El Gharamti et al. 2021)
- Larger inflation in densely observed watersheds

# 5.4 Streamflow Bias Mitigation

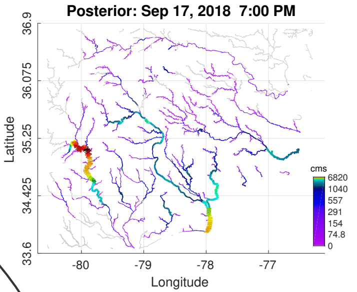
After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina

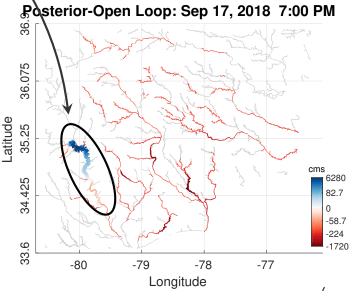


# 5.4 Streamflow Bias Mitigation

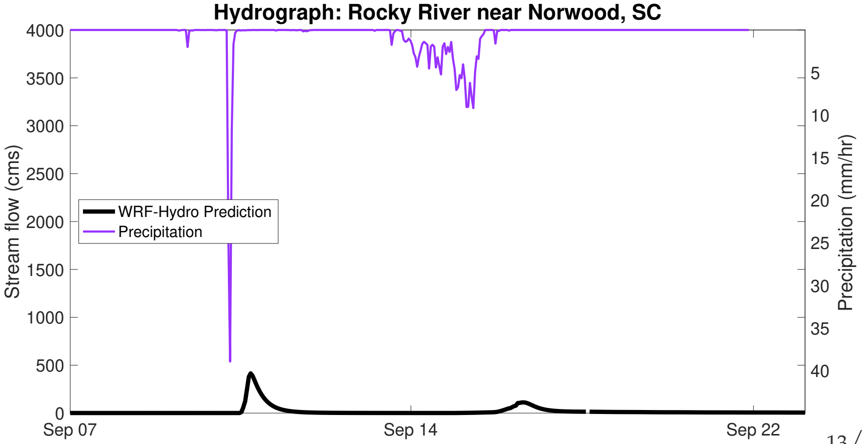
After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina



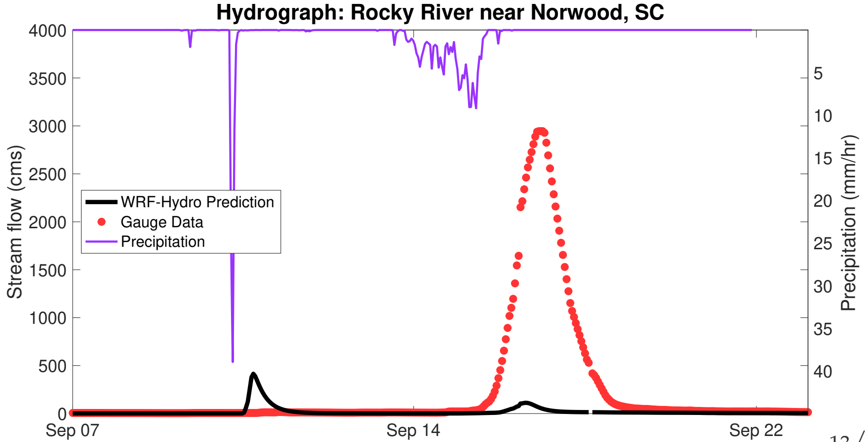
# 5.4 Streamflow Bias Mitigation

After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina



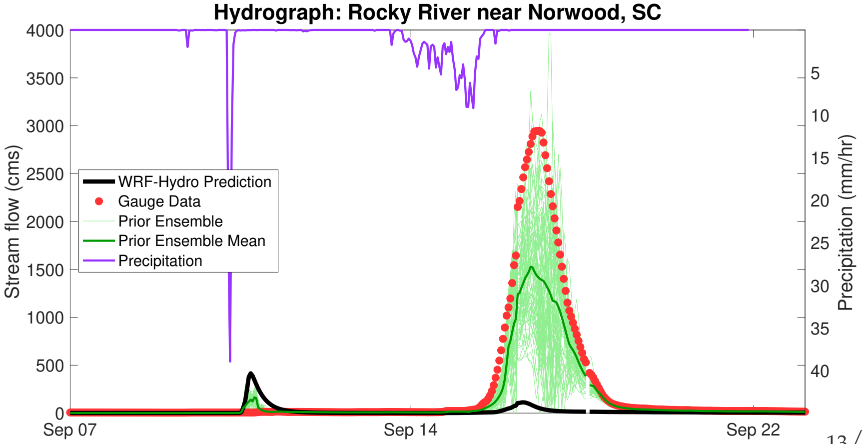
# 5.4 Streamflow Bias Mitigation

After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina



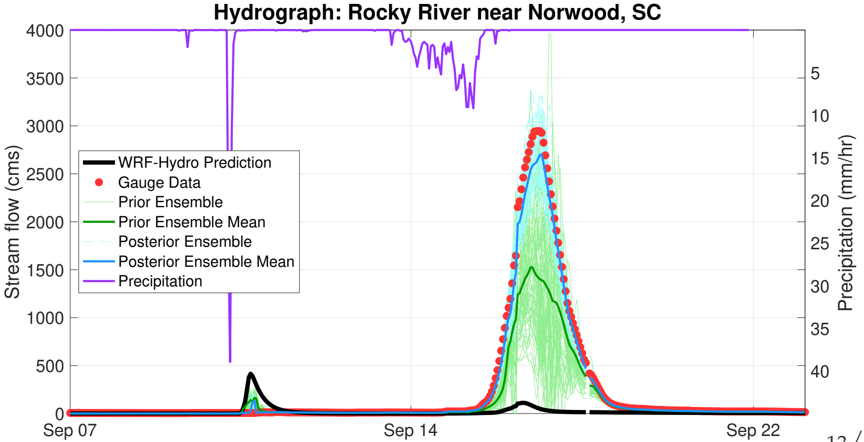
# 5.4 Streamflow Bias Mitigation

After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina



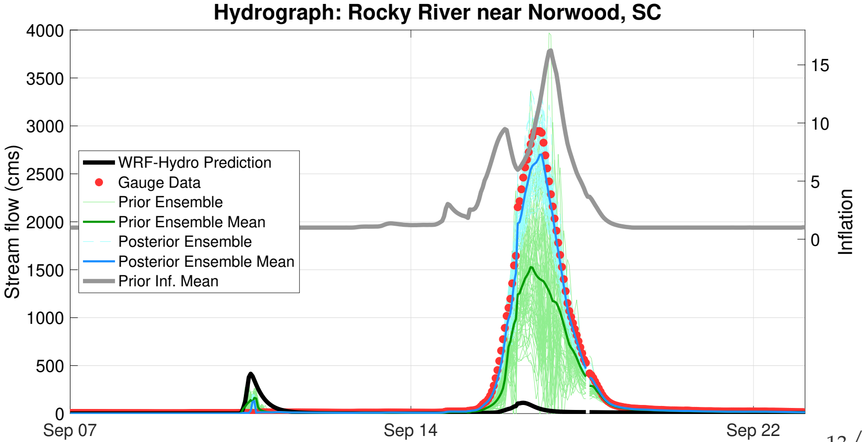
# 5.4 Streamflow Bias Mitigation

After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina



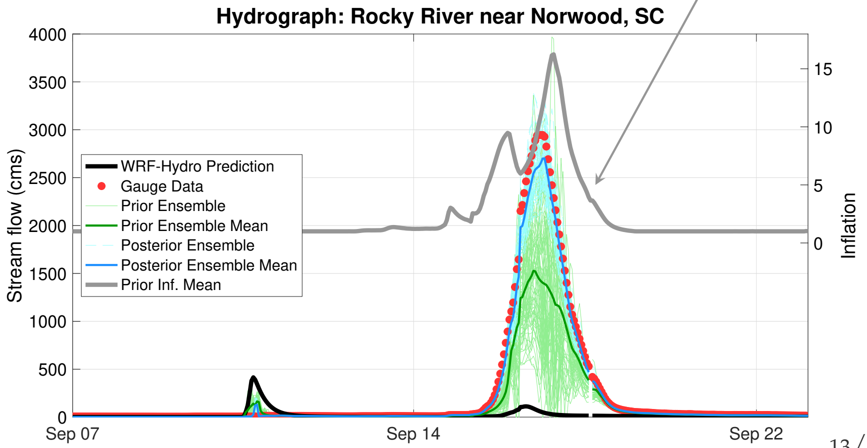
# 5.4 Streamflow Bias Mitigation

After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina



After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina

A sizable increase in prior inflation to counter the bias in the modeled streamflow!



# 6. Summary

- Spatially and temporally varying adaptive inflation algorithm

## 6. Summary

- Spatially and temporally varying adaptive inflation algorithm
- Indispensable in NWP systems: Massive number of available data cause huge reduction in ensemble spread that need to be restored

# 6. Summary

- Spatially and temporally varying adaptive inflation algorithm
- Indispensable in NWP systems: Massive number of available data cause huge reduction in ensemble spread that need to be restored
- Combing prior and posterior inflation can tackle different issues in the ensemble simultaneously

# 6. Summary

- Spatially and temporally varying adaptive inflation algorithm
- Indispensable in NWP systems: Massive number of available data cause huge reduction in ensemble spread that need to be restored
- Combing prior and posterior inflation can tackle different issues in the ensemble simultaneously
- For flood prediction, adaptive inflation is found useful and can serve as a rigorous bias correction scheme when the model's prediction is highly uncertain

# 6. Summary

- Spatially and temporally varying adaptive inflation algorithm
- Indispensable in NWP systems: Massive number of available data cause huge reduction in ensemble spread that need to be restored
- Combing prior and posterior inflation can tackle different issues in the ensemble simultaneously
- For flood prediction, adaptive inflation is found useful and can serve as a rigorous bias correction scheme when the model's prediction is highly uncertain

https://dart.ucar.edu/


Thank You!