Enhanced Streamflow Forecasting using Ensemble DA

Hurricane Florence Flooding 2018

Moha Gharamti

`https://dart.ucar.edu/`

`gharamti@ucar.edu`

Special thanks to: James McCreight, Tim Hoar, Seong Jin Noh, Arezoo RafieeiNasab, Ben Johnson, Nancy Collins

AIMES Virtual Workshop 06/15/2021

National Center for Atmospheric Research
Data Assimilation Research Section (DAReS) - TDD - CISL

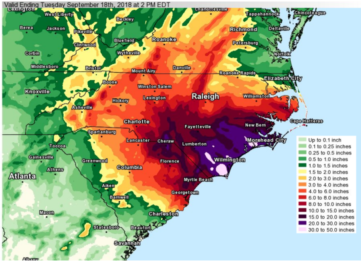
CISL Computational & Information Systems Lab

NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

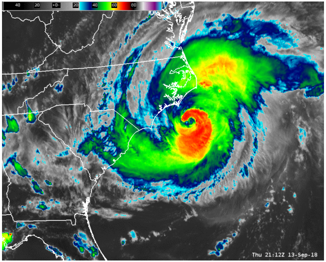
1. Why Streamflow Forecasting?

Hurricane Florence (2018):

- Tropical wave ⇝ tropical storm ⇝ **Category 4 Hurricane**
- Landfall on Sep. 14 (Carolinas) with winds up to 150 mph
- Catastrophic damages to coastal communities [$25 billion]
- Flooding magnitude greatly exceeded the levels observed due to Hurricane Matthew (2016) and Floyd (1999) **combined!**



Rainfall estimates from Hurricane Florence (*Source: NWS*)



Hurricane Florence eye during landfall (*Source: NWS*)

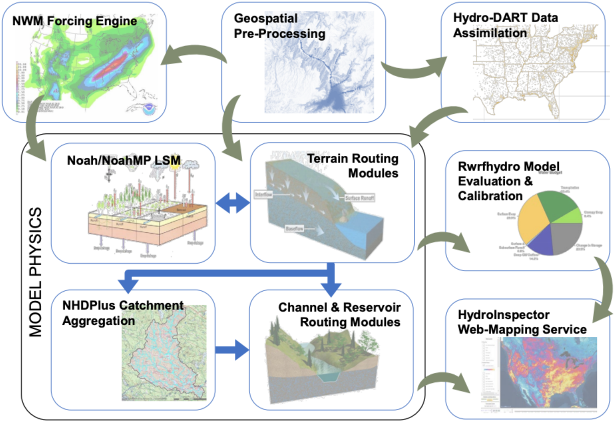
1. Why Streamflow Forecasting?

- Predicting major floods during extreme rainfall events is crucial
  1. Save lives (~ 50 people died due to Florence Flooding)
  2. Limit damages (via advance warnings)
  3. Protect infrastructure, socio-economic impacts, ...

Flooded city of New Bern, NC

2.1 The Coupled Modeling-DA Framework, HydroDART

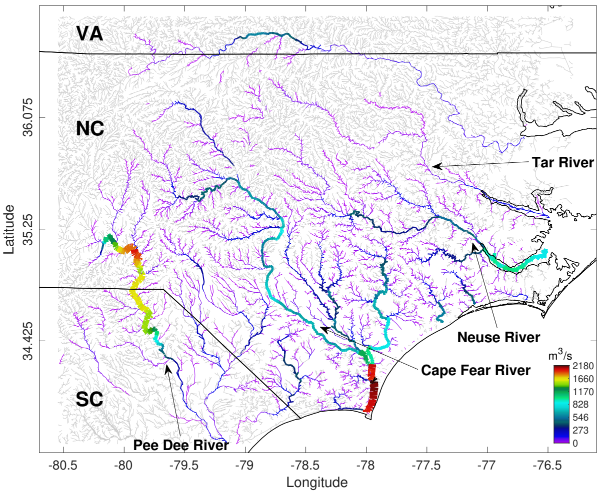
- Interface the Data Assimilation Research Testbed [DART: Anderson et al., 2008; BAMS] with WRF-Hydro [Gochis et al., 2020]



WRF-Hydro (NOAA's NWM) modeling framework
(*Source: https://ral.ucar.edu/projects/wrf_hydro*)

2.1 The Coupled Modeling-DA Framework, HydroDART

- Interface the Data Assimilation Research Testbed [DART: Anderson et al., 2008; BAMS] with WRF-Hydro [Gochis et al., 2020]

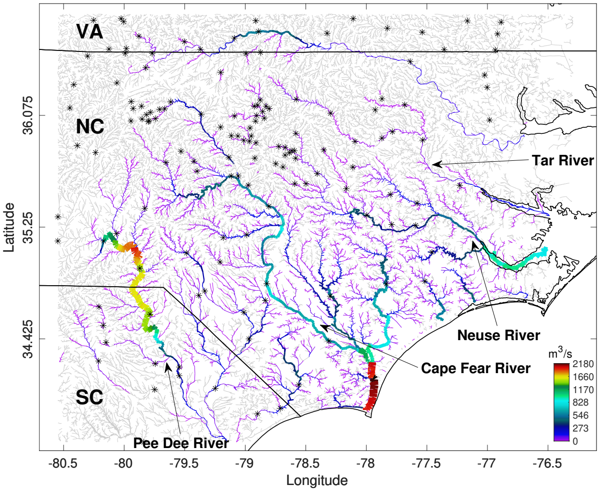


WRF-Hydro®

- Regional subdomain of the NWM CONUS
- NWM channel network based on NHDPlus V.2
- ~ 70K reaches

2.1 The Coupled Modeling-DA Framework, HydroDART

- Interface the Data Assimilation Research Testbed [DART: Anderson et al., 2008; BAMS] with WRF-Hydro [Gochis et al., 2020]



WRF-Hydro®

- Regional subdomain of the NWM CONUS
- NWM channel network based on NHDPlus V.2
- ~ 70K reaches
- Hourly streamflow assimilation
- 107 USGS gauges
- EnKF: 80 members

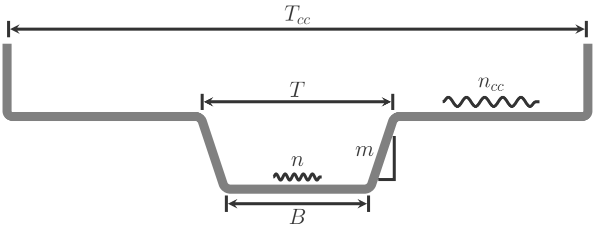
https://dart.ucar.edu/

2.2 Enhancements to the DA System

1. *Forcing and Ensemble Uncertainty:*

2.2 Enhancements to the DA System

1. *Forcing and Ensemble Uncertainty:*
   - Perturb fluxes to the channel and groundwater bucket
   - Multi-configuration ensemble; perturb 6 channel parameters

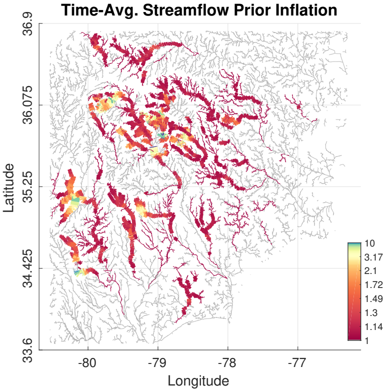


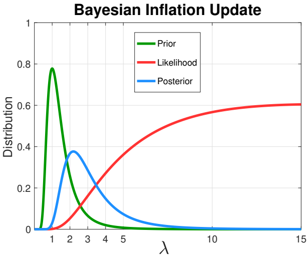
2.2 Enhancements to the DA System

1. *Forcing and Ensemble Uncertainty:*
2. *Adaptive Inflation:*

2.2 Enhancements to the DA System

1. *Forcing and Ensemble Uncertainty:*
2. *Adaptive Inflation:*
   - Tackle variance underestimation due to sampling errors and model biases
   - Spatially and Temporally varying algorithm [El Gharamti 2018; MWR]
   - Inflation assumed a random variable; updated using the data

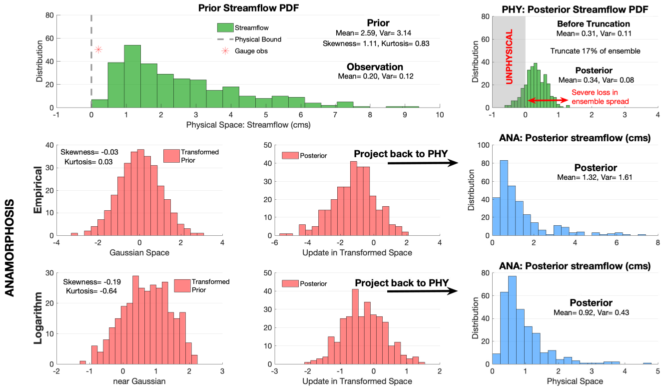




2.2 Enhancements to the DA System

1. *Forcing and Ensemble Uncertainty:*
2. *Adaptive Inflation:*
3. *Gaussian Anamorphosis:*

1. *Forcing and Ensemble Uncertainty:*
2. *Adaptive Inflation:*
3. *Gaussian Anamorphosis:*
   - Streamflow is a positive quantity ⇒ non-Gaussian
   - A variable transform during the update



2.2 Enhancements to the DA System

1. *Forcing and Ensemble Uncertainty:*
2. *Adaptive Inflation:*
3. *Gaussian Anamorphosis:*
4. *Along-The-Stream (ATS) Localization:*

2.2 Enhancements to the DA System

1. *Forcing and Ensemble Uncertainty:*
2. *Adaptive Inflation:*
3. *Gaussian Anamorphosis:*
4. *Along-The-Stream (ATS) Localization:*
    - Small ensemble sizes produce imperfect sample covariances
    - Taper spurious correlations
    - Channel routing model: unstructured grid (stream network)

2.3.1 Along-The-Stream (ATS) Localization

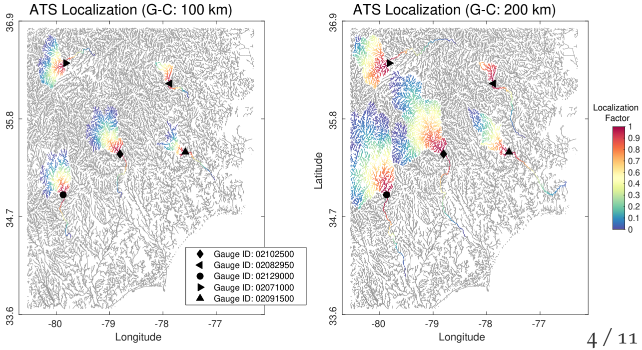
$$\mathbf{x}_{j,k}^{a(i)} = \mathbf{x}_{j,k}^{f(i)} + \alpha \Delta x_j^{(i)} \qquad 0 < \alpha < 1 \qquad \text{(Localization Factor)}$$

- ATS localization [El Gharamti et al., 2020; HESS] aims to mitigate not only spurious correlations but also **physically incorrect correlations between unconnected state variables in the river network**
- 2 reaches could be physically close but unrelated (particularly through error correlations) if they belong to different catchments/basins

2.3.1 Along-The-Stream (ATS) Localization

$$\mathbf{x}_{j,k}^{a(i)} = \mathbf{x}_{j,k}^{f(i)} + \alpha \Delta x_j^{(i)} \qquad 0 < \alpha < 1 \qquad (\text{Localization Factor})$$

- ATS localization [El Gharamti et al., 2020; HESS] aims to mitigate not only spurious correlations but also **physically incorrect correlations between unconnected state variables in the river network**
- 2 reaches could be physically close but unrelated (particularly through error correlations) if they belong to different catchments/basins

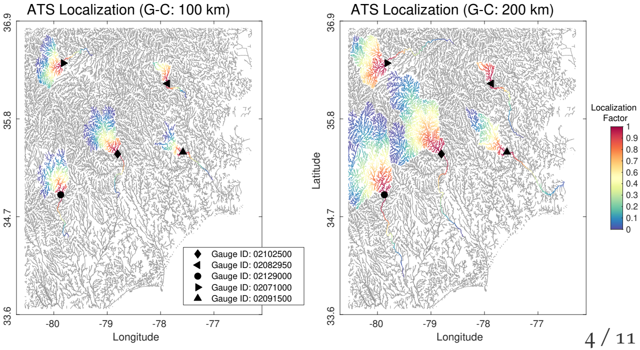


2.3.1 Along-The-Stream (ATS) Localization

$$\mathbf{x}_{j,k}^{a(i)} = \mathbf{x}_{j,k}^{f(i)} + \alpha \Delta x_j^{(i)} \qquad 0 < \alpha < 1 \qquad (\text{Localization Factor})$$

Functionality/Characteristics:

- Reaches **only upstream and downstream from a gauge** are impacted

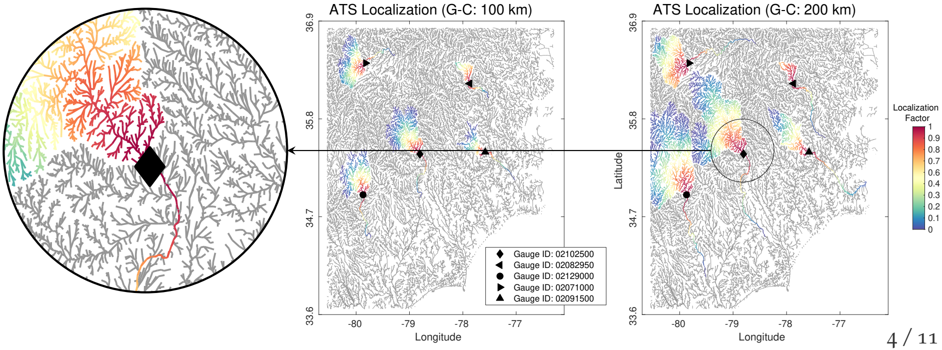


2.3.1 Along-The-Stream (ATS) Localization

$$\mathbf{x}_{j,k}^{a(i)} = \mathbf{x}_{j,k}^{f(i)} + \alpha \Delta x_j^{(i)} \qquad 0 < \alpha < 1 \qquad \text{(Localization Factor)}$$

Functionality/Characteristics:

0. Reaches **only upstream and downstream from a gauge** are impacted
1. Downstream from a gauge, information flows only downstream (tree-like shapes)

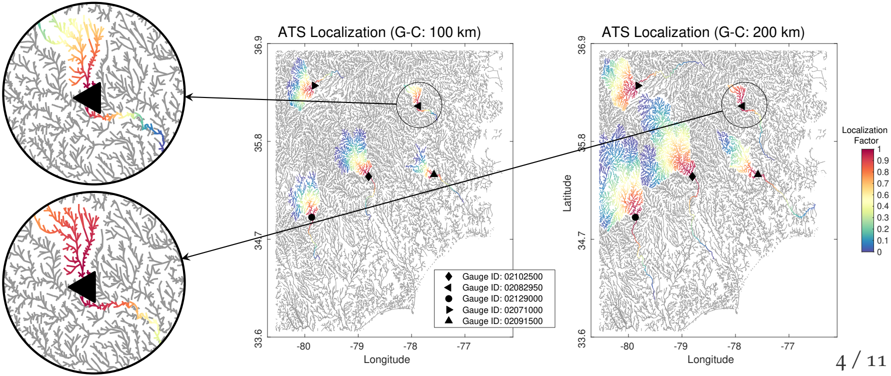


2.3.1 Along-The-Stream (ATS) Localization

$$\mathbf{x}_{j,k}^{a(i)} = \mathbf{x}_{j,k}^{f(i)} + \alpha \Delta x_j^{(i)} \qquad 0 < \alpha < 1 \qquad (\text{Localization Factor})$$

Functionality/Characteristics:

0. Reaches **only upstream and downstream from a gauge** are impacted
1. Downstream from a gauge, information flows only downstream (tree-like shapes)
2. Total number of close reaches depend on the size of the basin

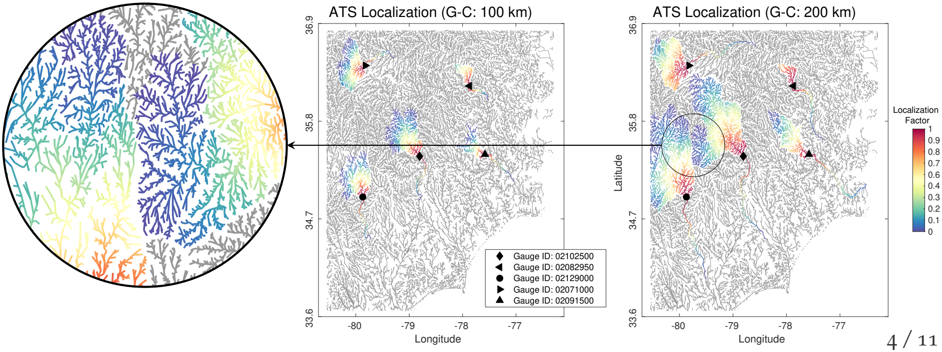


2.3.1 Along-The-Stream (ATS) Localization

$$\mathbf{x}_{j,k}^{a(i)} = \mathbf{x}_{j,k}^{f(i)} + \alpha \Delta x_j^{(i)} \qquad 0 < \alpha < 1 \qquad \text{(Localization Factor)}$$

Functionality/Characteristics:

0. Reaches **only upstream and downstream from a gauge** are impacted
1. Downstream from a gauge, information flows only downstream (tree-like shapes)
2. Total number of close reaches depend on the size of the basin
3. Observations in different catchments do not have common close reaches

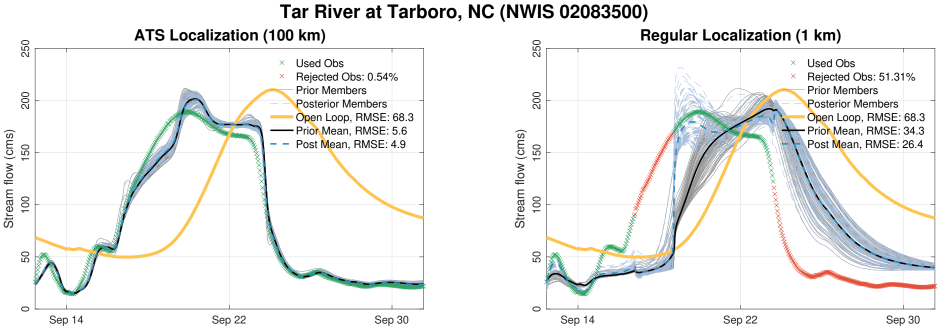


2.3.2 Does regular localization even work?

Tar River at Tarboro (NWIS 02083500)		ATS	Reg 20	Reg 10	Reg 5	Reg 2	Reg 1
	Prior RMSE	5.58	18.54	8.86	33.46	41.61	34.32
	Posterior RMSE	4.93	17.82	6.75	25.11	33.66	26.41
	Prior Bias	-1.13	-11.65	-1.71	-20.24	-18.09	-11.07
	Posterior Bias	-0.85	-11.41	-0.74	-20.37	-17.16	-10.01
	Prior Spread	1.20	3.29	2.80	10.90	10.84	9.54
	Posterior Spread	1.55	3.00	2.27	6.28	6.43	5.17

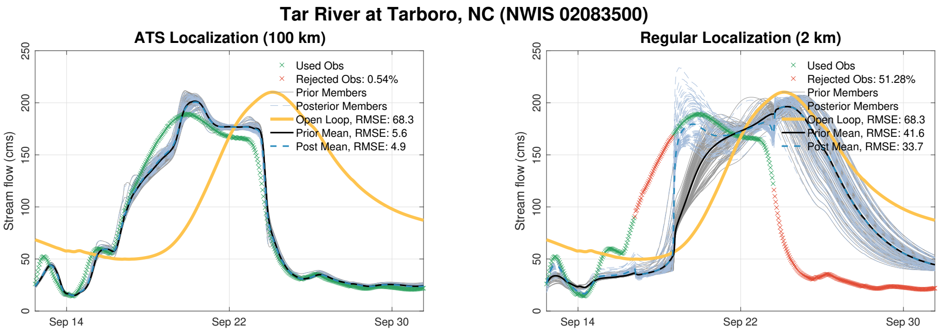
- Performance using ATS localization is significantly better ($\sim$ 40%)
- Using ATS, one can increase the effective localization radius
- Regular localization with large radii fails (correlating physically unrelated variables)

2.3.2 Does regular localization even work?



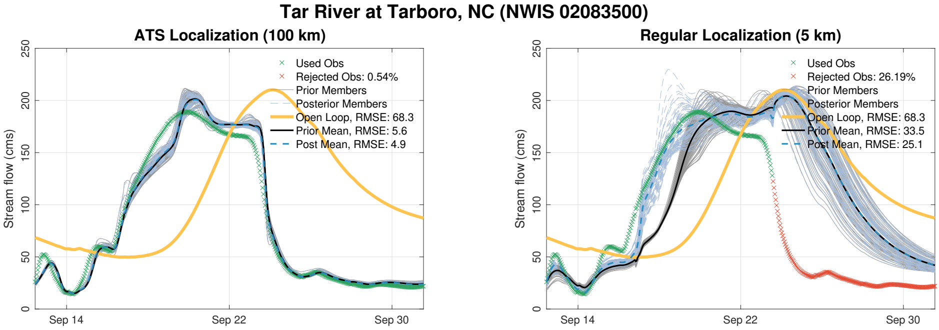
- Performance using ATS localization is significantly better (~ 40%)
- Using ATS, one can increase the effective localization radius
- Regular localization with large radii fails (correlating physically unrelated variables)

2.3.2 Does regular localization even work?



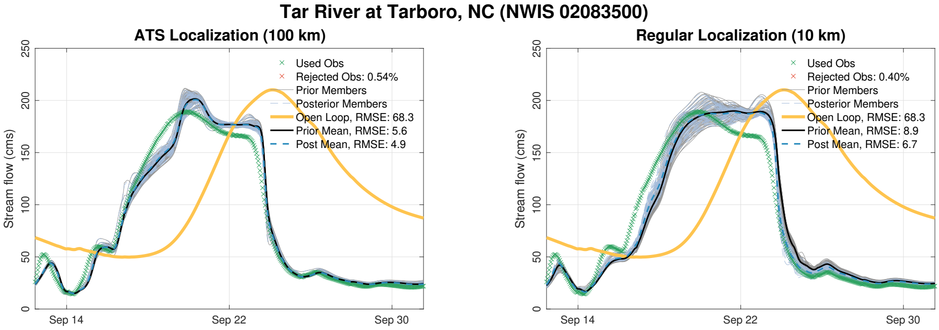
- Performance using ATS localization is significantly better (~ 40%)
- Using ATS, one can increase the effective localization radius
- Regular localization with large radii fails (correlating physically unrelated variables)

2.3.2 Does regular localization even work?



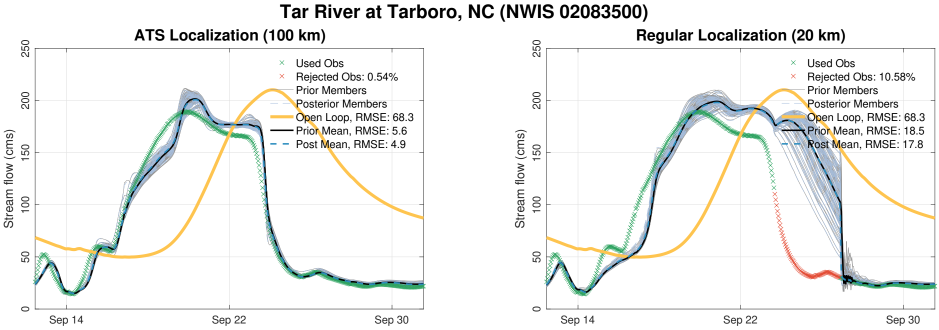
- Performance using ATS localization is significantly better (~ 40%)
- Using ATS, one can increase the effective localization radius
- Regular localization with large radii fails (correlating physically unrelated variables)

2.3.2 Does regular localization even work?



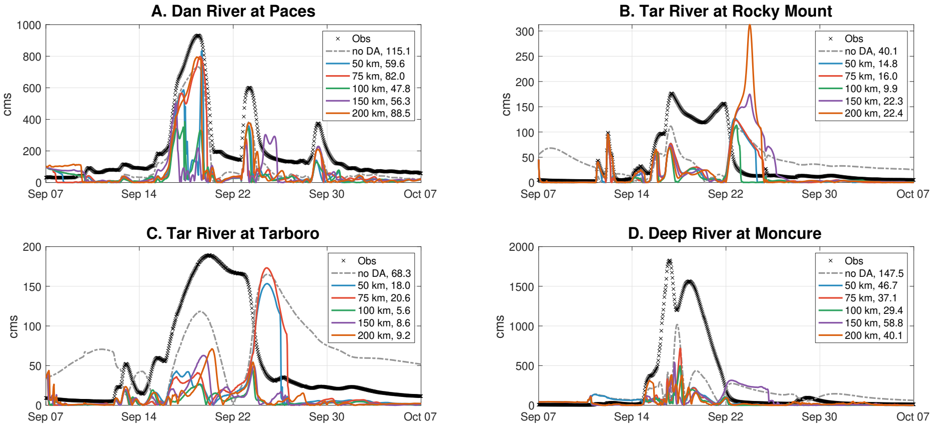
- Performance using ATS localization is significantly better (~ 40%)
- Using ATS, one can increase the effective localization radius
- Regular localization with large radii fails (correlating physically unrelated variables)

2.3.2 Does regular localization even work?



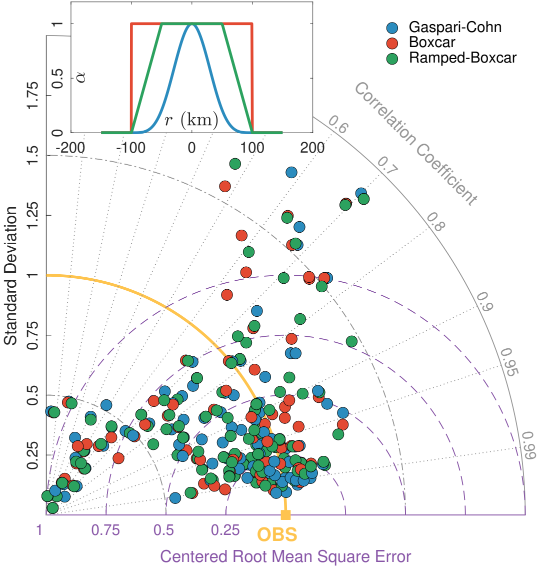
- Performance using ATS localization is significantly better (~ 40%)
- Using ATS, one can increase the effective localization radius
- Regular localization with large radii fails (correlating physically unrelated variables)

2.3.3 Tuning ATS Localization; [i] Radius



- Test with different localization radii: 50, 75, 100, 150, 200 km
- Larger radii degrade the accuracy (giving rise to spurious correlations)
- Smaller radii limit the amount of useful information
- Best performance with 100 km

2.3.4 Tuning ATS Localization; [ii] Correlation Function



- Averaging over all gauges, the correlation coefficient was: Gaspari-Cohn (**0.83**), Boxcar (**0.77**) and Ramped-Boxcar (**0.79**)
- Gaspari-Cohn outperforms other functions

3.1 Summary

- **HydroDART** is a state-of-the-art streamflow prediction system that couples WRF-Hydro and DART
  1. Provides hourly skillful streamflow estimates
  2. Enhanced ensemble uncertainty assessment
  3. Introduces Along-The-Stream localization
  4. Supports a variety of DA algorithms: e.g., Adaptive Inflation
  5. Supports parameter *(model + hyper)* estimation

- **ATS Localization:**
  1. Topological localization scheme that adheres to the river network
  2. Improves information propagation (eliminates error covariances between unconnected streams)
  3. Produces significantly better results than regular Euclidean distance-based approach

3.1 Summary

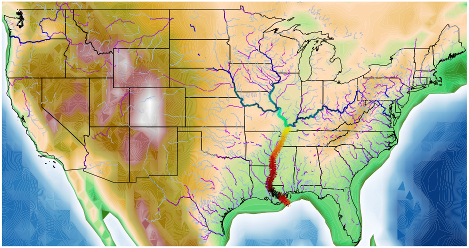
- **HydroDART** is a state-of-the-art streamflow prediction system that couples WRF-Hydro and DART
  1. Provides hourly skillful streamflow estimates
  2. Enhanced ensemble uncertainty assessment
  3. Introduces Along-The-Stream localization
  4. Supports a variety of DA algorithms: e.g., Adaptive Inflation
  5. Supports parameter *(model + hyper)* estimation

https://github.com/NCAR/wrf_hydro_nwm_public
https://github.com/NCAR/wrf_hydro_dart
https://github.com/NCAR/DART

- **ATS Localization:**
  1. Topological localization scheme that adheres to the river network
  2. Improves information propagation (eliminates error covariances between unconnected streams)
  3. Produces significantly better results than regular Euclidean distance-based approach

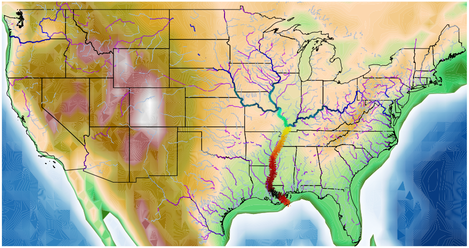
3.2 Future Research Directions

- Full CONUS streamflow reanalysis for the past 30 years:
  $\longrightarrow$ Explore hybrid EnKF-OI approaches:
  Adaptive: [El Gharamti 2021; MWR]
  Analogs: [Grooms 2021; QJRMS]



3.2 Future Research Directions

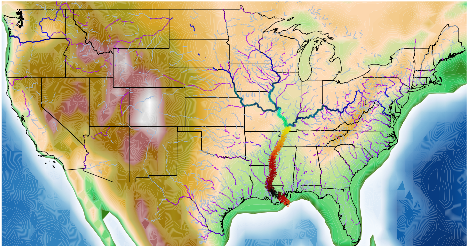
- Full CONUS streamflow reanalysis for the past 30 years:
  → Explore hybrid EnKF-OI approaches:
  Adaptive: [El Gharamti 2021; MWR]
  Analogs: [Grooms 2021; QJRMS]



- A collaborative project with USGS; 2 main goals:
  1. Assimilate gauge temperature data (investigate effects on streamflow)
  2. Placement of gauges (OSSE studies)

3.2 Future Research Directions

- Full CONUS streamflow reanalysis for the past 30 years:
  → Explore hybrid EnKF-OI approaches:
  Adaptive: [El Gharamti 2021; MWR]
  Analogs: [Grooms 2021; QJRMS]



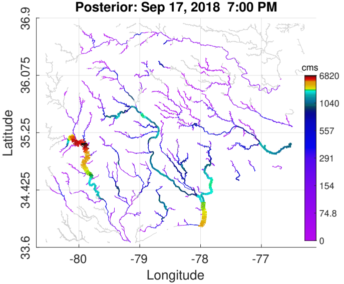
- A collaborative project with USGS; 2 main goals:
  1. Assimilate gauge temperature data (investigate effects on streamflow)
  2. Placement of gauges (OSSE studies)

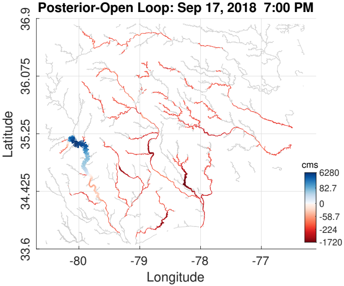
- Coupling the LSM with WRF-Hydro:
  1. Assimilate soil moisture & streamflow; weak vs strong coupling
  2. Assimilate snow data (thickness, SWE, ...)

Backup Slides

Bias Mitigation

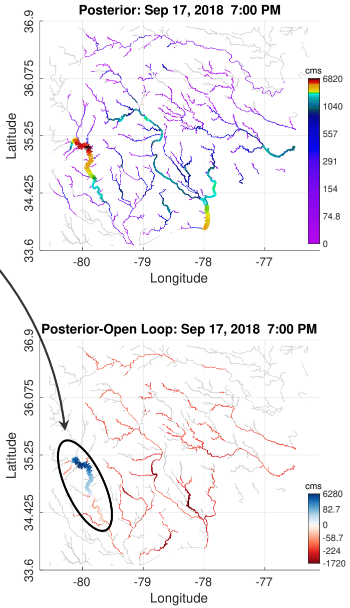
After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina





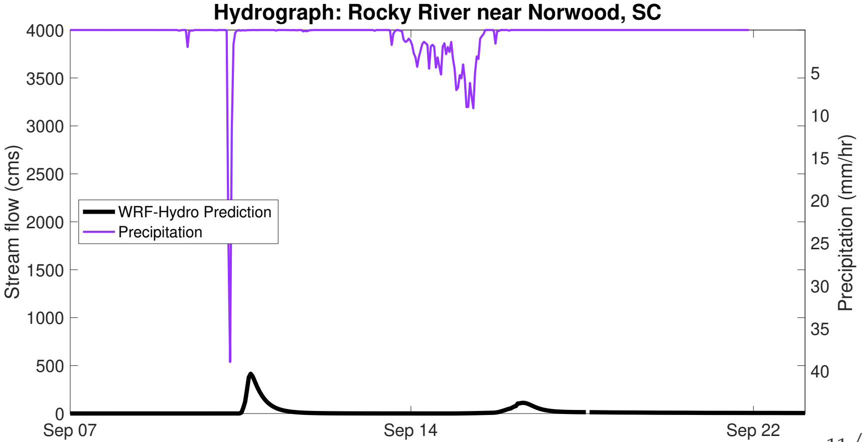
Bias Mitigation

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



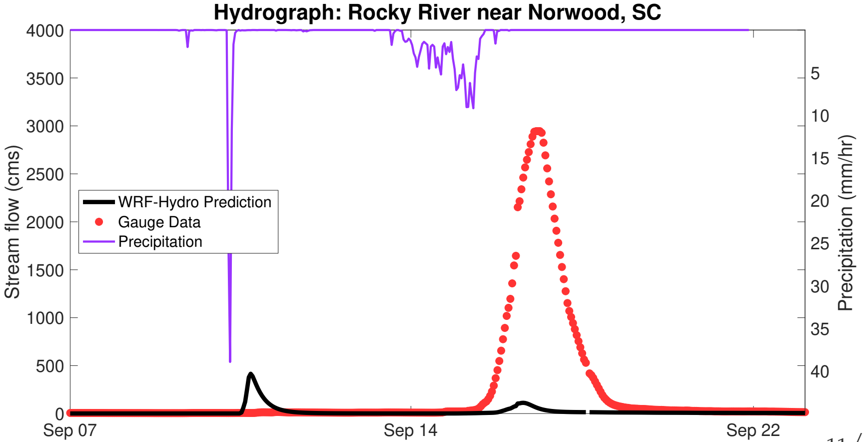
Bias Mitigation

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



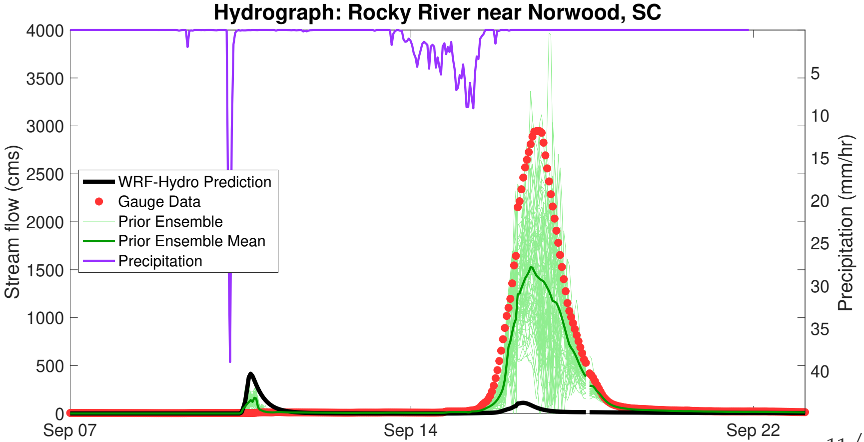
Bias Mitigation

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



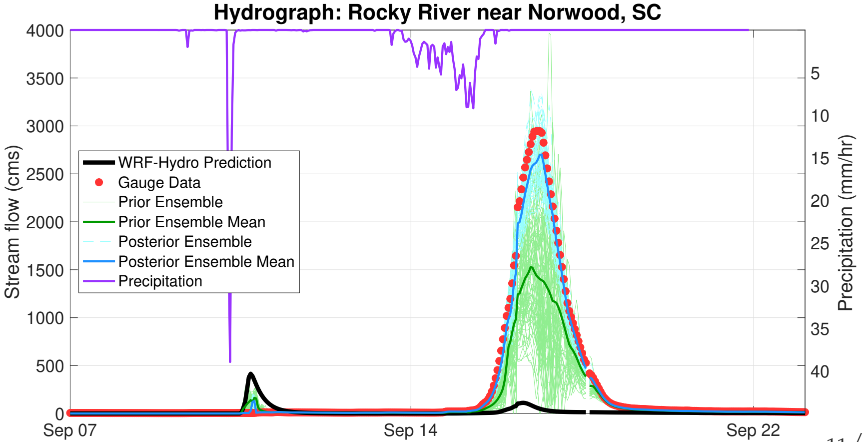
Bias Mitigation

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



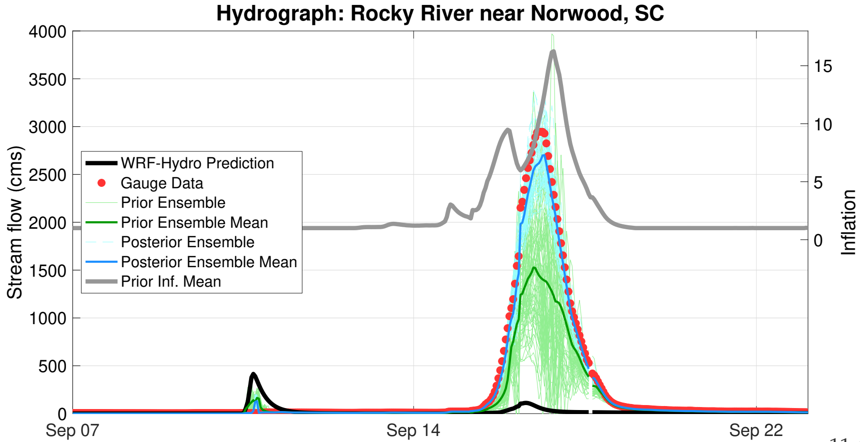
Bias Mitigation

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



Bias Mitigation

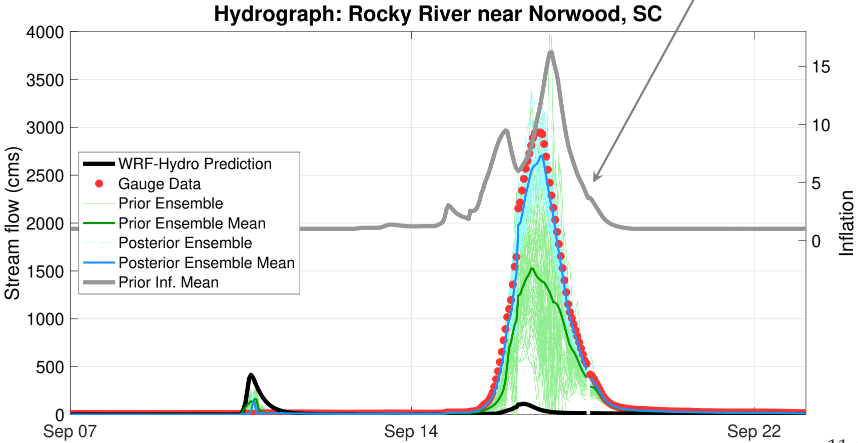
After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina



Bias Mitigation

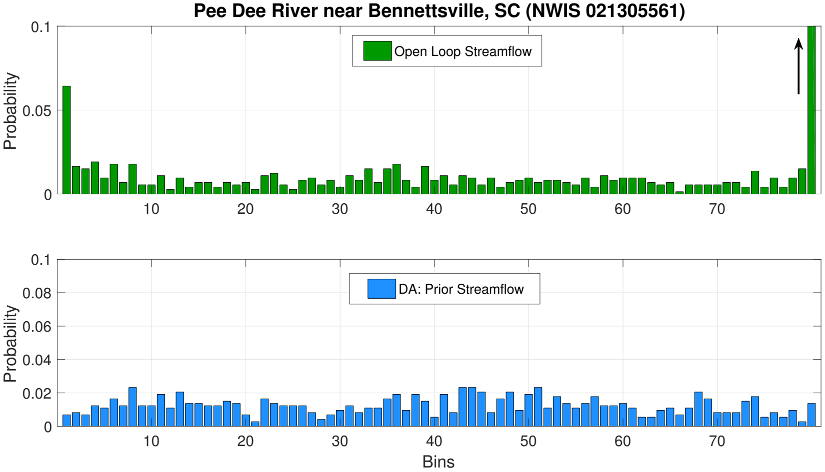
After landfall, the model's streamflow prediction (Open Loop) is significantly smaller than the posterior along Pee-Dee River in South Carolina

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



Bias Mitigation

The rank histogram for the open loop is heavily skewed to the right indicating that the gauge data is larger than the ensemble



Bias Mitigation

The rank histogram for the open loop is heavily skewed to the right indicating that the gauge data is larger than the ensemble