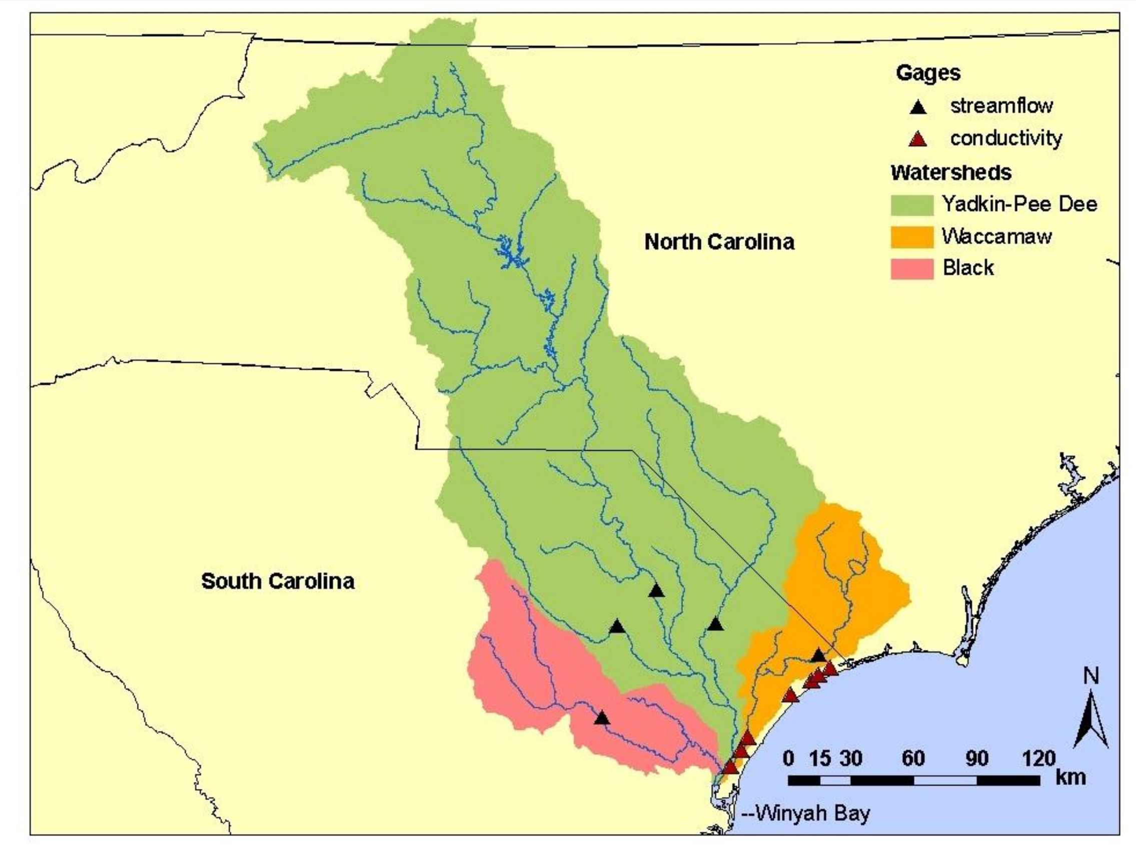
Assessing methods to disaggregate daily precipitation for hydrological simulation

Peng Gao, Gregory Carbone, Daniel Tufford, Aashka Patel, and Lauren Rouen

University of South Carolina

Background

- **CISA (Carolinas Integrated Sciences and Assessments)**



- **Hydrological modeling: how climate affects water supply and quality in major watersheds in Carolinas**

Background

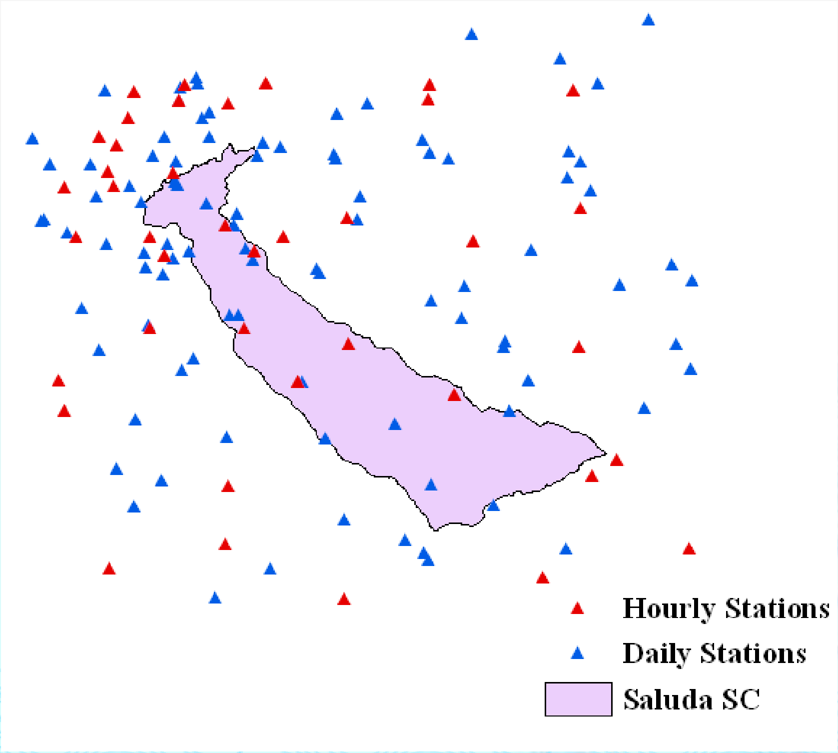
Continuous simulation modeling

(e.g. Hydrologic Simulation Program-Fortran (HSPF))

- **a principle tool to investigate the impacts of climate change on water resources**
- **high spatial and temporal resolution (e.g. hourly or subdaily) rainfall data**

Challenges - the constraint of data availability

- **Precipitation data are often available only at coarser levels (i.e., daily) (*25,000* daily recording stations, *8,000* hourly stations in U.S.)** **(Booner, 1998)**



Challenges - the constraint of data availability

- **Meteorological variables from the GCMs (General Circulation Models) needed for hydrological simulation - typically at monthly or daily scales**

Geo Data Portal

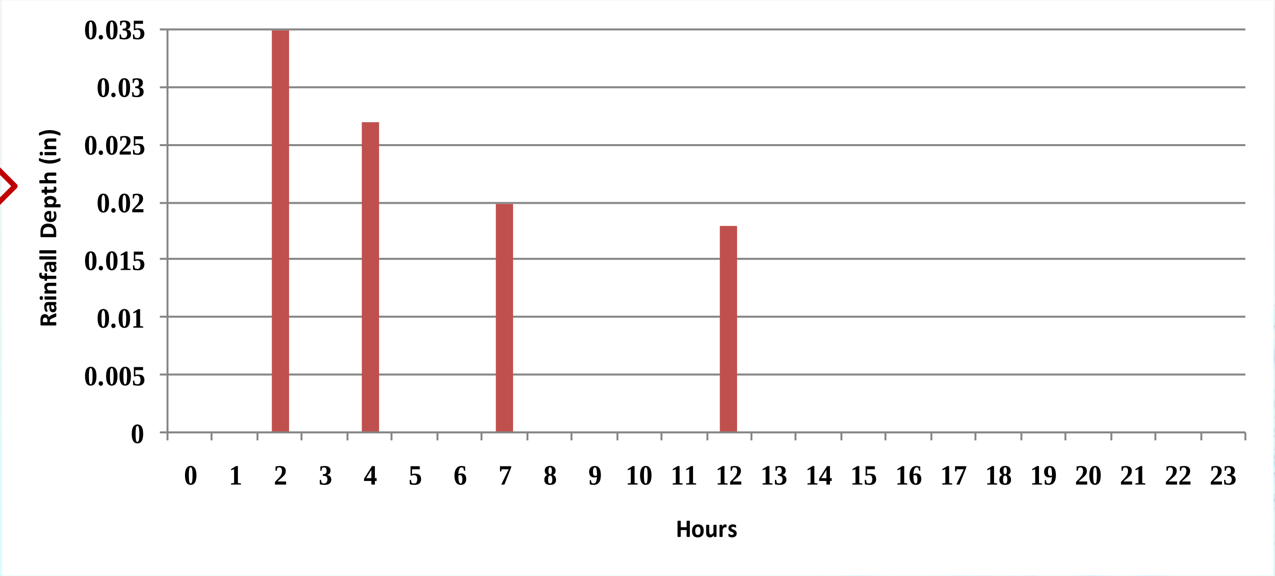
Downscaled Climate Projections by Katharine Hayhoe

Precipitation. The downscaling model for precipitation is similar to that for temperature in many aspects, but with some key differences. First, for practical reasons an AOGCM predictor had to be chosen that was commonly archived at the daily scale. Although upper-level humidity and geopotential height have shown promise in downscaling precipitation, few AOGCMs have preserved daily outputs. Thus, 24h cumulative precipitation was selected as the predictor for precipitation, with the additional refinement of incorporating convective and large-scale precipitation if both predictors were available. For models with these variables, the downscaling approach selects from three possible predictors the one best suited to each month: convective, large-scale, or total. This refinement significantly improved the method's ability to simulate precipitation over arid and semi-tropical regions. Second, EOF filtering of the GCM output is not performed since we found that to degrade the results along with introducing negative values for precipitation. Finally, the logarithm of precipitation values is used instead of raw precipitation amount. This was found to decrease the residuals of the regression.

Solutions

- **Disaggregate the daily rainfall to hourly time series**

Date	Prec. (inch)
Jun 16	0
Jun 17	0.1
Jun 18	0
et al.	



Background

- **Many disaggregation methods**

- **Few tests to assess the performance of these methods on hydrological simulations**

Overview of the Study

- Examine three different disaggregation methods to construct hourly precipitation time series from daily precipitation

- Use those time series as input and compare simulated flows against observed flows

Three Disaggregation Methods

Method1 – Triangular by HSPF

Daily rainfall needs to be disaggregated: 0.10

daily total		0	0.01	0.02	0.04	0.08	0.16	0.32	…
ratio for each hour	10	0	0	0	0	0	0.01	0.01	…
	11	0	0	0	0.01	0.01	0.04	0.05	…
	12	0	0.01	0.01	0.02	0.03	0.06	0.1	…
	13	0	0	0.01	0.01	0.03	0.04	0.1	…
	14	0	0	0	0	0.01	0.01	0.05	…
	15	0	0	0	0	0	0	0.01	…

- **Find the daily total closest to but larger than the daily rainfall that needs to be disaggregated**
- **Distribute the daily rainfall proportionally to ratio for each hour**

Method 2 and 3

DISAGGREGATION OF DAILY RAINFALL FOR CONTINUOUS WATERSHED MODELING

By Scott Socolofsky,[1] E. Eric Adams,[2] Members, ASCE, and Dara Entekhabi[3]

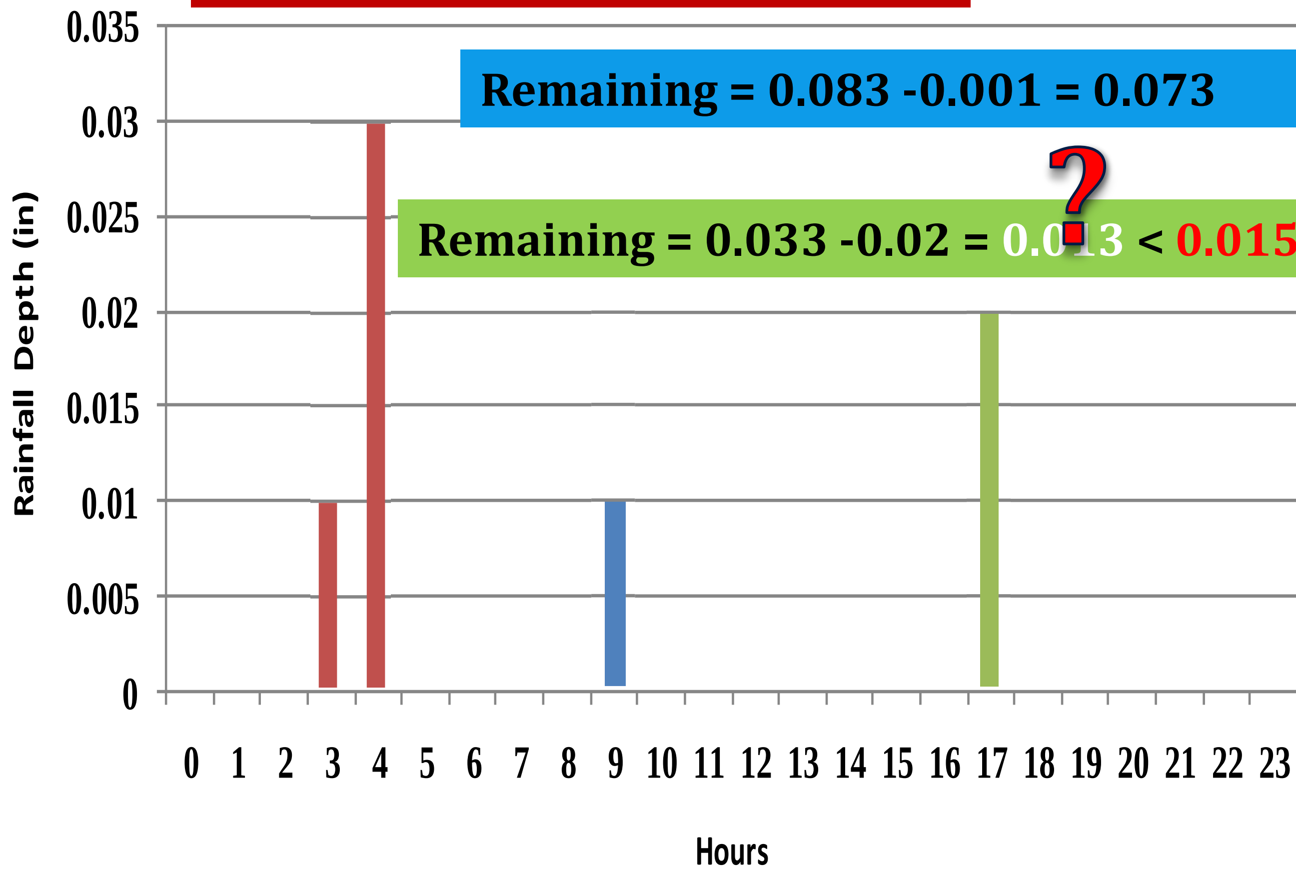
- It iteratively searches the rainfall events from the existing rainfall event database until the remaining amount is lower than an assigned minimum threshold
- The disaggregated hourly rainfalls reserve the probability distribution of the existing rainfall event database

To disaggregate a 0.083 inch daily total with the assigned minimum threshold: 0.015 inch

Remaining = 0.073 -0.04 = 0.033

Remaining = 0.083 -0.001 = 0.073

Remaining = 0.033 -0.02 = 0.013 < 0.015 ?



Rainfall events

Total: 0.04 inch
(0.01 inch at 3 o clock and 0.03 inch at 4 0 clock)

Total: 0.02 inch
(0.02 inch at 8 o clock)

Total: 0.01 inch
(0.01 inch at 9 o clock)

Total: 0.01 inch
(0.01 inch at 6 o clock)

Total: 0.01 inch
(0.01 inch at 17 o clock)

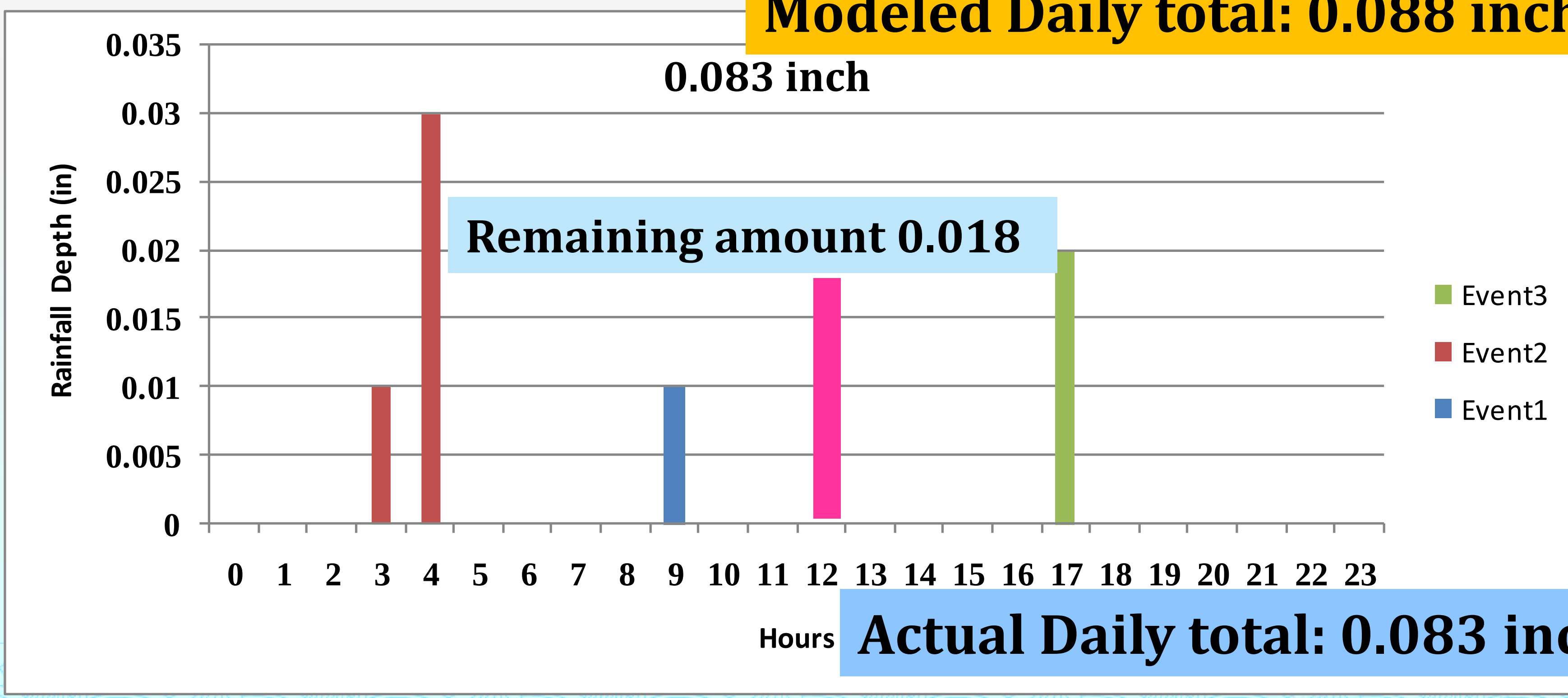
et al.

How to deal with the remaining amount of 0.013 inch?

Method2: Socolofsky method

- The remaining amount follows exponential distribution (e.g. Modeled remaining amount = -Actual remaining amount (*i.e., -0.013*)***log (random seed)** at random hour from (0 to23))

Modeled Daily total: 0.088 inch

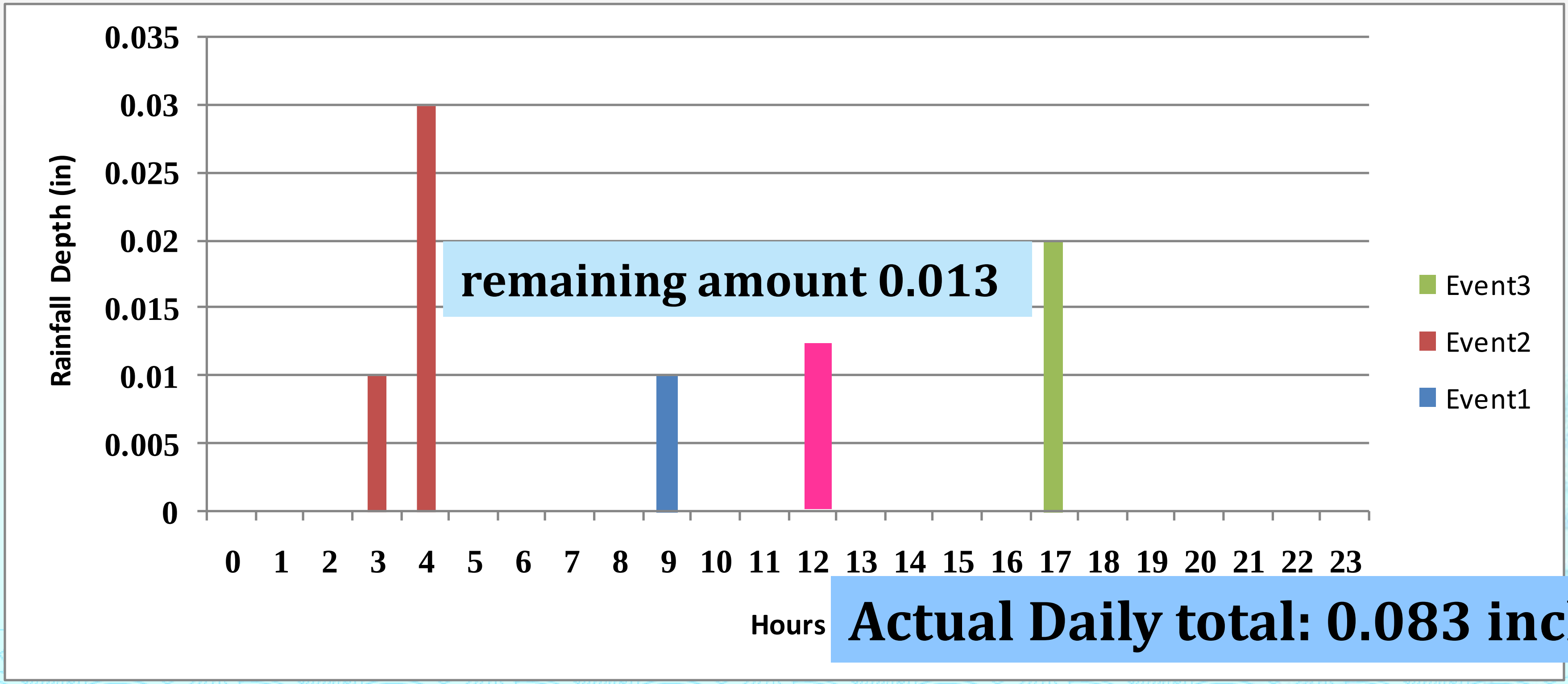


Actual Daily total: 0.083 inch

Method3: Adjusted Socolofsky method

- The remaining amount is directly placed into a one-hour storm event

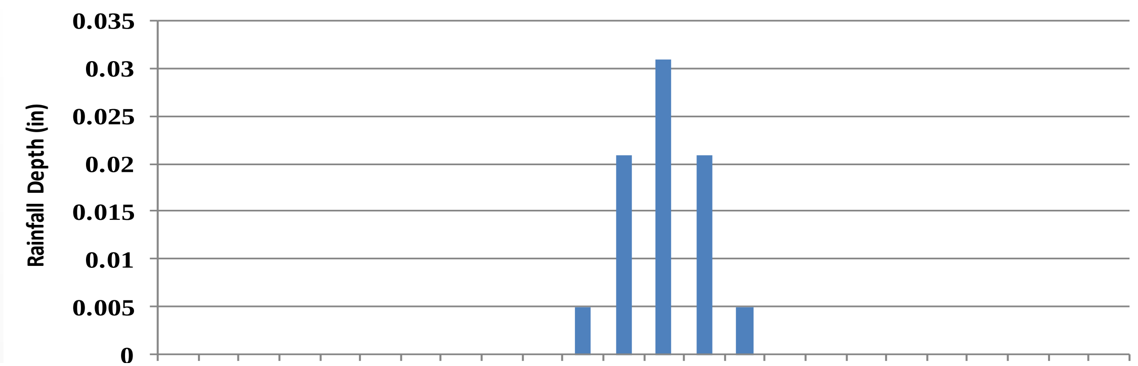
Modeled Daily total: 0.083 inch



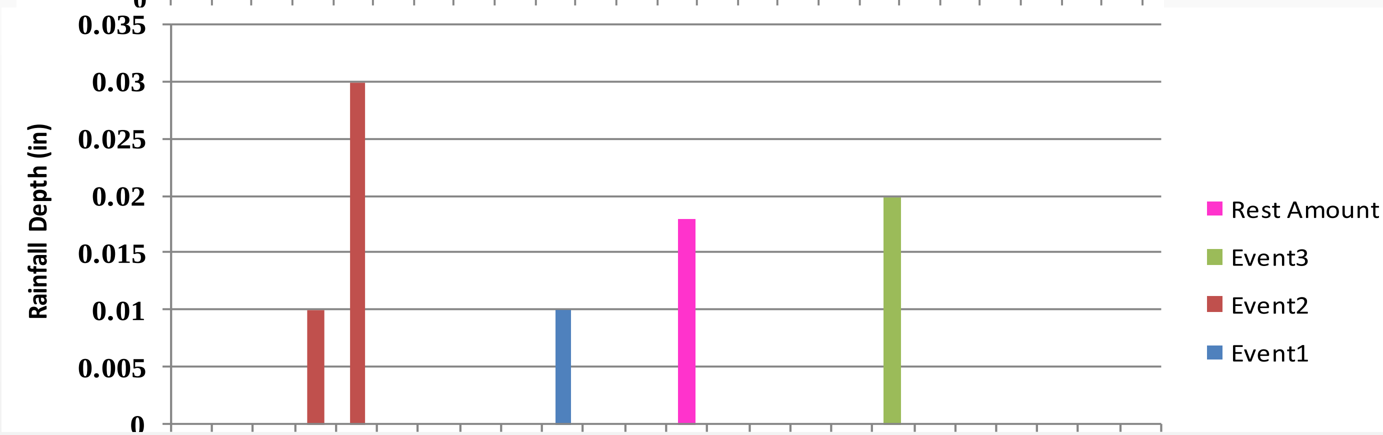
Actual Daily total: 0.083 inch

Daily total: 0.083 inch

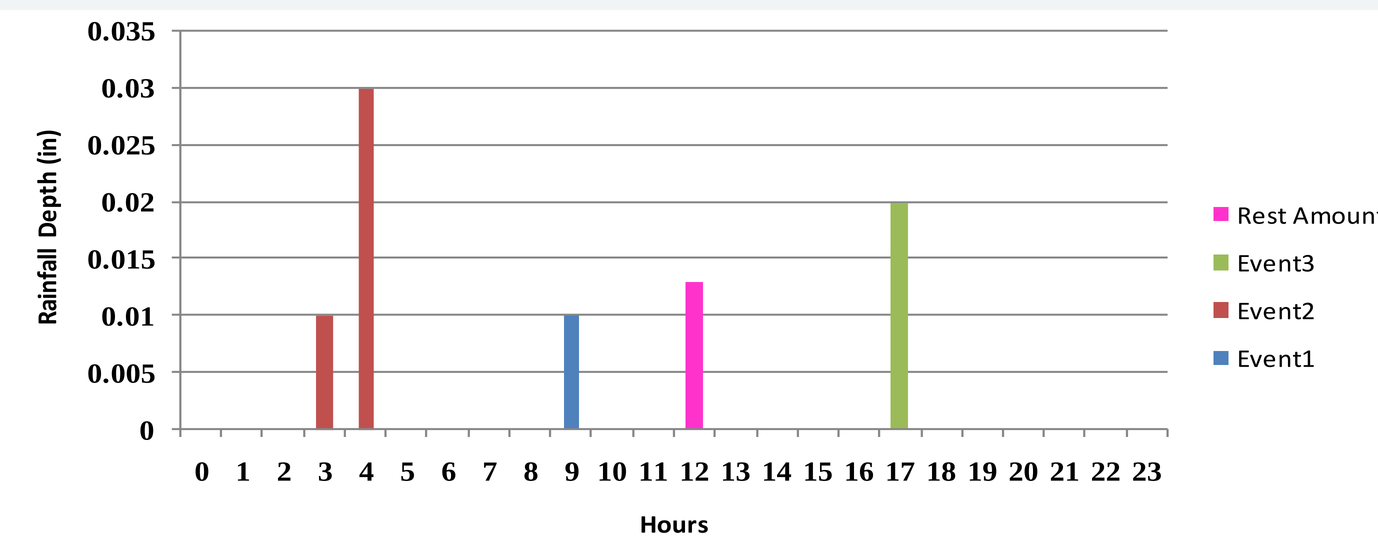
Triangular Distribution



Socolofsky method



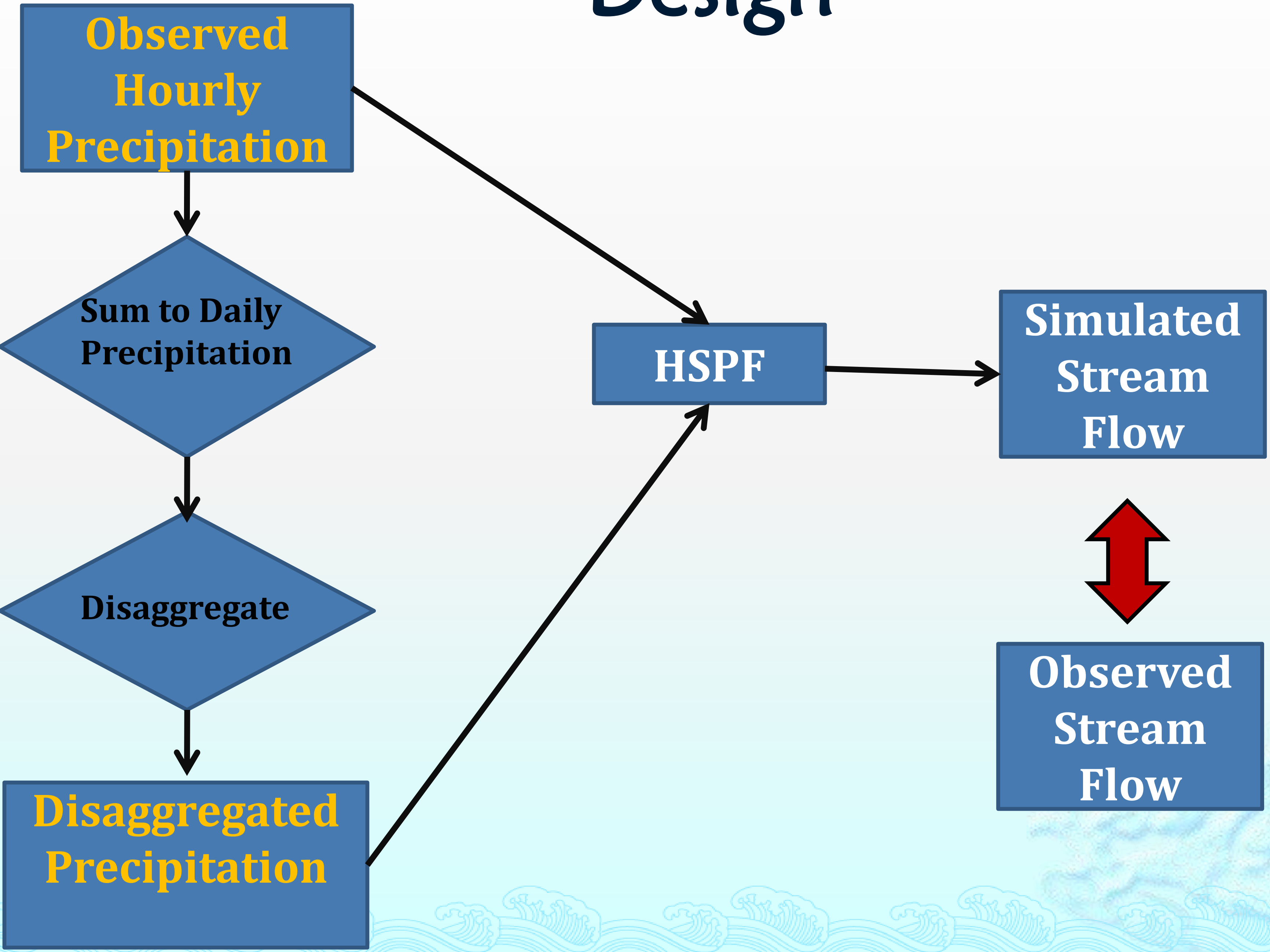
Adjusted Socolofsky method



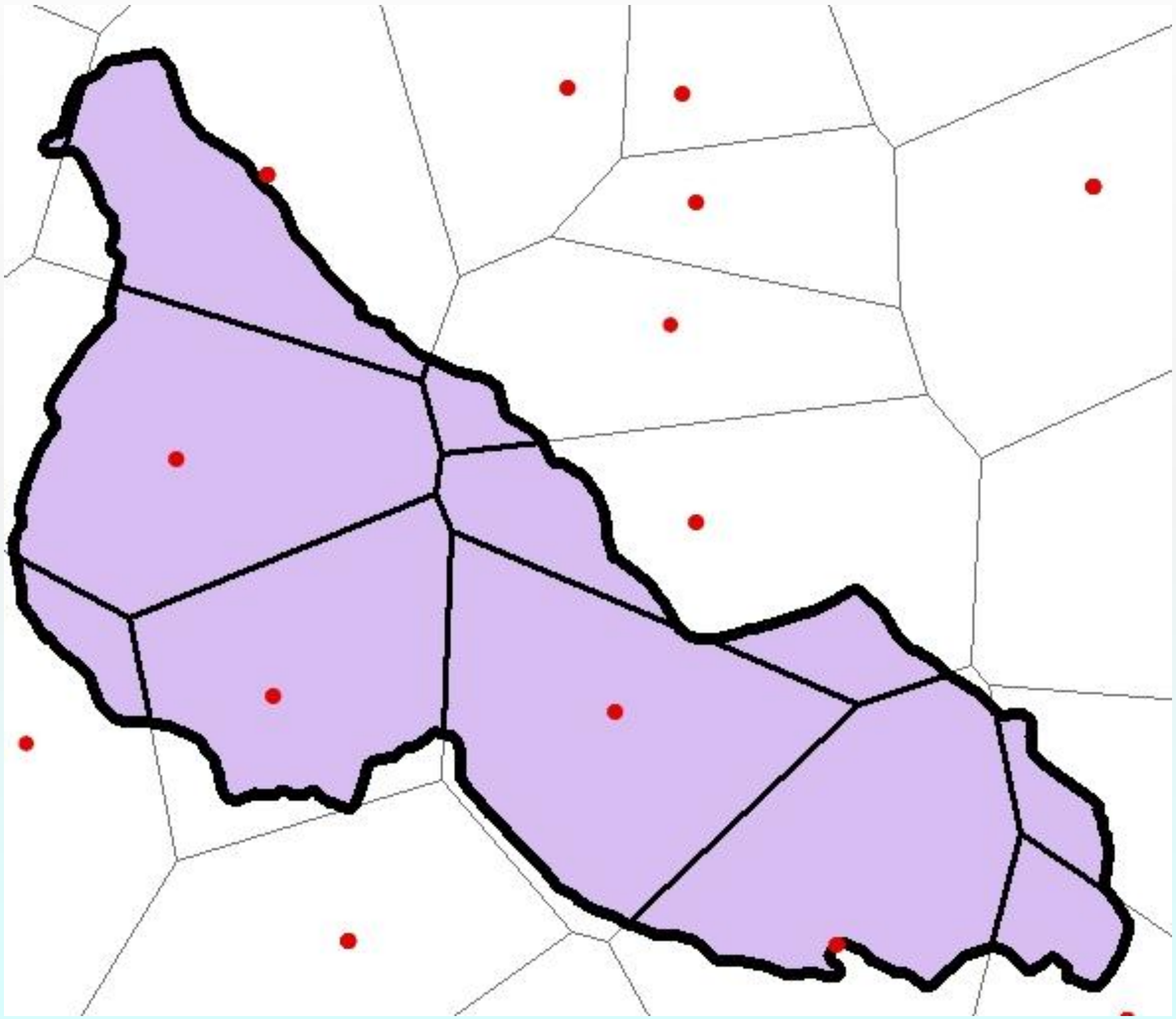
Characteristics Comparison

	Triangular Distribution	Socolofsky Method	Adjusted Socolofsky Method
Stochastic	N	Y	Y
Using existing hourly data	N	Y	Y
Conserve daily total	Y	N	Y
Conserve PDF	N	Y	Y
Multi events	N	Possible	Possible

Design


Observed Hourly Precipitation
Sum to Daily Precipitation
Disaggregate
Disaggregated Precipitation
HSPF
Simulated Stream Flow
Observed Stream Flow

Observed Hourly Precipitation

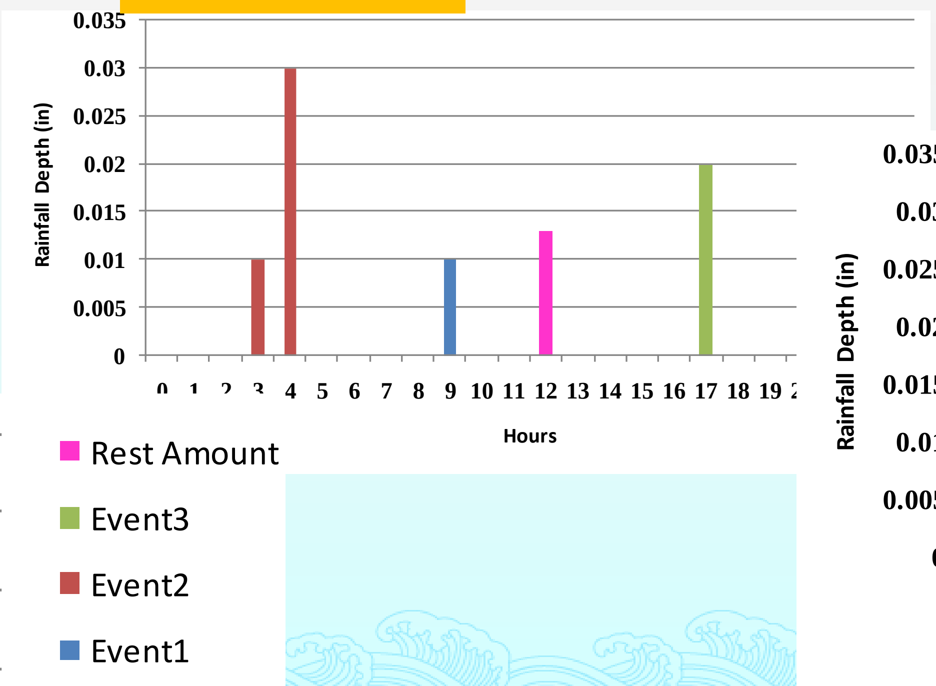


- Virtual station: area weighted mean of stations whose Thiessen polygons fall into the watershed

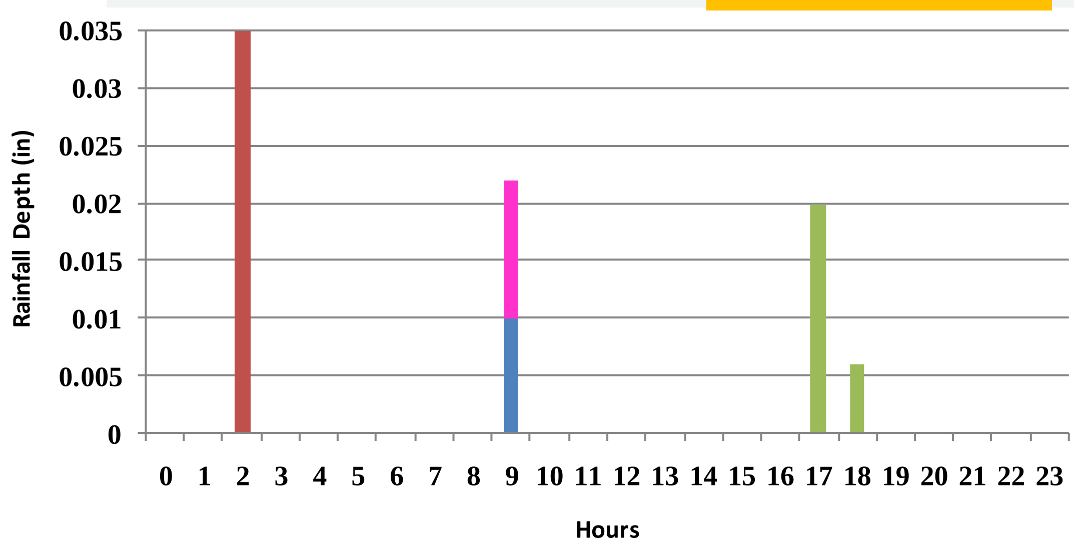
Disaggregated Precipitation

- **Triangular Distribution**
- **Ten simulations for each Socolofsky method and Adjusted Socolofsky method *using precipitation from the virtual station as the existing rainfall event database***

Simulation 1



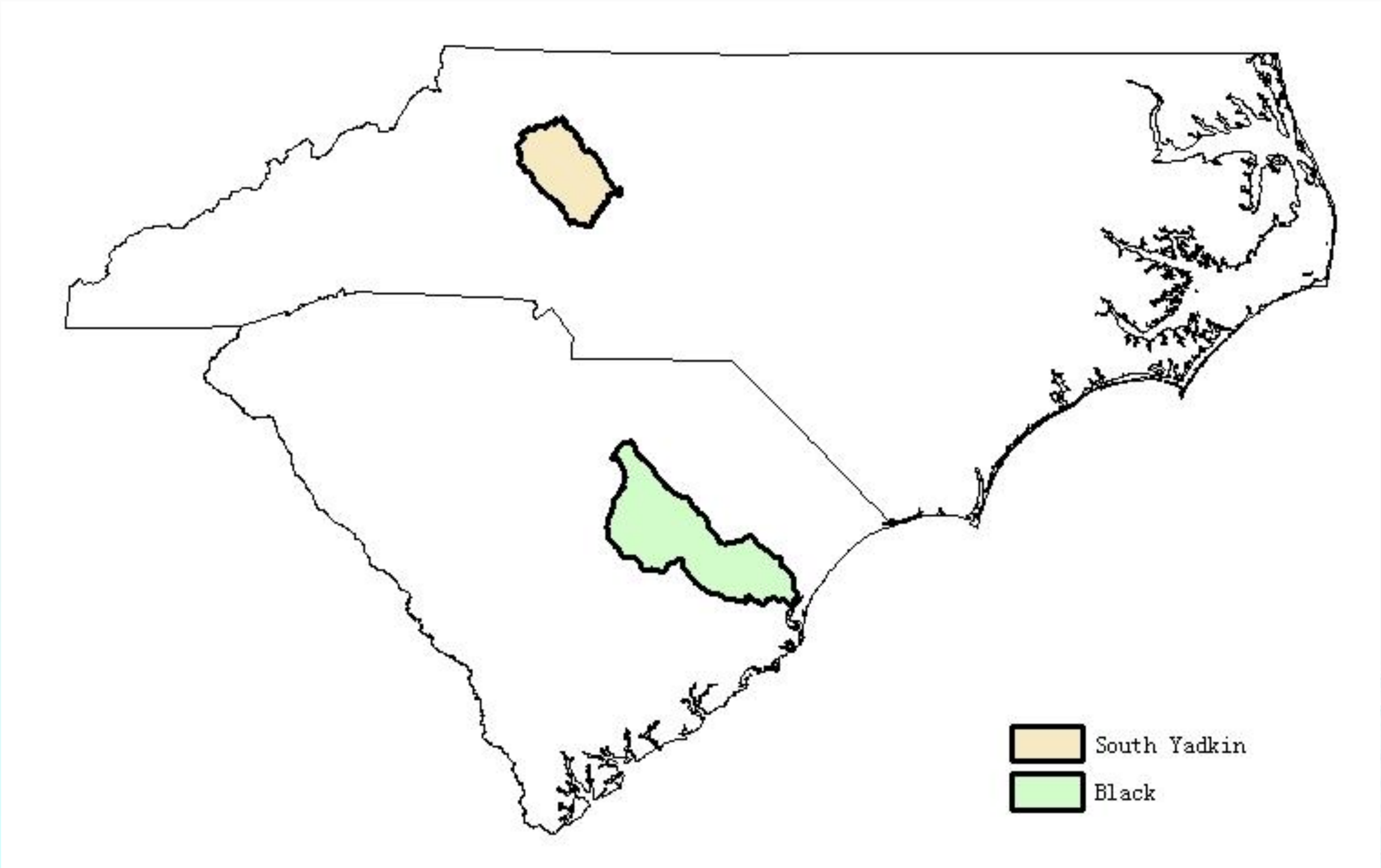
Simulation 2



Comparisons

- **Four types of hourly precipitation time series**
  - ***Precipitation from the virtual station (a combination of observed hourly precipitation)***
  - **Triangular distribution**
  - **Socolofsky Method (ten simulations)**
  - **Adjusted Socolofsky Method (ten simulations)**
- **The simulated stream flows VS. the observed stream flows in the verification time period**

Study Area



Model performance evaluation

- Indices
  - index of agreement (d) (higher, better)
  - mean absolute error (MAE) (lower, better)
  - Nash-Sutcliff efficiency (NS) (closer to 1, better)
  - percent bias (p-bias) (lower, better)
  - root mean squared error (RMSE) (lower, better)
  - Willmott Index (dr) (Willmott et al. 2011) (closer to 1, better)

Nash-Sutcliff efficiency (NS)

- range from $-\infty$ to 1
- NS= 1: a perfect match
- NS = 0 :the model predictions are as accurate as the mean of the observed data
- $NS < 0$: the observed mean is a better predictor than the model

$$\mathrm{NS} = 1 - \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{O_i} - \mathrm{P_i})^2}{\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{O_i} - \bar{\mathrm{O}}_{\mathrm{i}})^2}$$

Nash and Sutcliffe (1970)

where: P_i is predicted value, O_i is observed value, $\bar{O}$ is the observed mean

Willmott Index

- Measures ratio of mean absolute error of modeled vs. observations to mean absolute error of observations relative to observed mean
- d_r=1: perfect agreement
- d_r=0.5: sum of error magnitudes is half perfect model deviation and observed deviation magnitudes
- d_r=-0.5: sum of error magnitudes is twice perfect model deviation and observed deviation magnitudes
- d_r=-1: model-estimated deviations about observed mean are poor estimates of observed deviation **<u>OR</u>** means little observed variability

$$d_r = \begin{cases} 1 - \dfrac{\sum_{i=1}^{n} |P_i - O_i|}{c \sum_{i=1}^{n} |O_i - \bar{O}|}, \textit{when} \\ \displaystyle\sum_{i=1}^{n} |P_i - O_i| \leq c \sum_{i=1}^{n} |O_i - \bar{O}| \\ \dfrac{c \sum_{i=1}^{n} |O_i - \bar{O}|}{\sum_{i=1}^{n} |P_i - O_i|} - 1, \textit{when} \\ \displaystyle\sum_{i=1}^{n} |P_i - O_i| > c \sum_{i=1}^{n} |O_i - \bar{O}| \end{cases}$$

Willmott et al. (2011)

where: P_i is predicted value, O_i is observed value, $\bar{O}$ is the observed mean, and c=2

Results

- **Socolofsky (S) and Adjusted Socolofsky (AS) VS. Virtual station (V)**

Statistic of daily stream flow in watershed South Yadkin, NC

Results

	d	MAE	NS	p-bias	RMSE	dr
V	**0.89**	**83.10**	**0.64**	**-15.47**	**186.8**	**0.67**
S1	0.86	87.60	0.57	-15.48	204.7	0.66
S2	0.79	93.66	0.40	-18.10	241.3	0.63
S3	0.86	86.43	0.58	-10.27	202.2	0.66
S4	0.84	83.09	0.51	-8.91	218.3	0.67
S5	0.81	82.70	0.48	-11.40	225.4	0.68
S6	0.78	105.15	0.30	-26.64	260.2	0.59
S7	0.76	92.13	0.32	-10.23	256.5	0.64
S8	0.78	100.31	0.35	-25.47	251.9	0.61
S9	0.81	93.72	0.42	-15.36	236.6	0.63
S10	0.86	87.46	0.58	-17.88	201.8	0.66

	d	MAE	NS	p-bias	RMSE	dr
V	**0.89**	**83.10**	**0.64**	**-15.47**	**186.8**	**0.67**
AS1	0.84	87.68	0.51	-14.22	217.8	0.66
AS2	0.84	86.53	0.54	-13.47	210.2	0.66
AS3	0.82	87.20	0.48	-13.74	224.1	0.66
AS4	0.82	91.19	0.47	-14.45	227.3	0.64
AS5	0.86	83.31	0.59	-12.86	199.3	0.67
AS6	0.85	84.10	0.56	-12.73	206.5	0.67
AS7	0.86	84.76	0.59	-13.32	200.0	0.67
AS8	0.82	88.52	0.47	-13.76	227.0	0.65
AS9	0.83	90.31	0.49	-14.35	222.5	0.65
AS10	0.83	85.77	0.51	-13.18	217.4	0.66

Socolofsky (S) and Adjusted Socolofsky (AS) produced similar statistics to precipitation from the virtual station (V)

Statistic of daily stream flow in watershed South Yadkin, NC

Results

	d	MAE	NS	p-bias	RMSE	dr
V	**0.89**	**83.10**	**0.64**	**-15.47**	**186.8**	**0.67**
S1	0.86	87.60	0.57	-15.48	204.7	0.66
S2	0.79	93.66	0.40	-18.10	241.3	0.63
S3	0.86	86.43	0.58	-10.27	202.2	0.66
S4	0.84	83.09	0.51	-8.91	218.3	0.67
S5	0.81	82.70	0.48	-11.40	225.4	0.68
S6	0.78	105.15	0.30	-26.64	260.2	0.59
S7	0.76	92.13	0.32	-10.23	256.5	0.64
S8	0.78	100.31	0.35	-25.47	251.9	0.61
S9	0.81	93.72	0.42	-15.36	236.6	0.63
S10	0.86	87.46	0.58	-17.88	201.8	0.66

	d	MAE	NS	p-bias	RMSE	dr
V	**0.89**	**83.10**	**0.64**	**-15.47**	**186.8**	**0.67**
AS1	0.84	87.68	0.51	-14.22	217.8	0.66
AS2	0.84	86.53	0.54	-13.47	210.2	0.66
AS3	0.82	87.20	0.48	-13.74	224.1	0.66
AS4	0.82	91.19	0.47	-14.45	227.3	0.64
AS5	0.86	83.31	0.59	-12.86	199.3	0.67
AS6	0.85	84.10	0.56	-12.73	206.5	0.67
AS7	0.86	84.76	0.59	-13.32	200.0	0.67
AS8	0.82	88.52	0.47	-13.76	227.0	0.65
AS9	0.83	90.31	0.49	-14.35	222.5	0.65
AS10	0.83	85.77	0.51	-13.18	217.4	0.66

Socolofsky (S) varied more than Adjusted Socolofsky (AS) because it does not conserve the depth of daily rainfall

Results

- **Triangular (T) VS. Virtual station (V)**

South Yadkin NC （V is better）

	d	MAE	NS	p-bias	RMSE	dr
V	0.89	83.10	0.64	-15.47	186.82	0.67
T	0.86	92.99	0.51	-15.66	219.06	0.64

Black basin SC (T is better)

	d	MAE	NS	p-bias	RMSE	dr
V	0.89	387.22	0.69	-18.73	656.1	0.75
T	0.91	332.88	0.75	-9.14	586.3	0.79

Conclusion

- **The adjusted Socolofsky method**
  - **most robust in terms of performance when compared to the model verification run using the observed hourly precipitation as input**
  - **a useful means of disaggregating the daily precipitation from GCMs under different scenarios**
- **comparisons in more watersheds to test the robustness and consistency of these methods**