# Flood Forecasting Using Satellite Precipitation with Statistical Models


**Krish Chhabra[1]*, Vinit Sehgal[2], Binayak P. Mohanty[3]**

[1]Computer Science and Engineering, [2]Water Management and Hydrological Science, [3]Biological and Agricultural Engineering, Texas A&M University


Presenter
Krish Chhabra

## Why is flood forecasting important?

Climate change has resulted in more **unpredictable and extreme precipitation patterns**. An improved ability to predict flooding can both **save lives and reduce costly damages.**

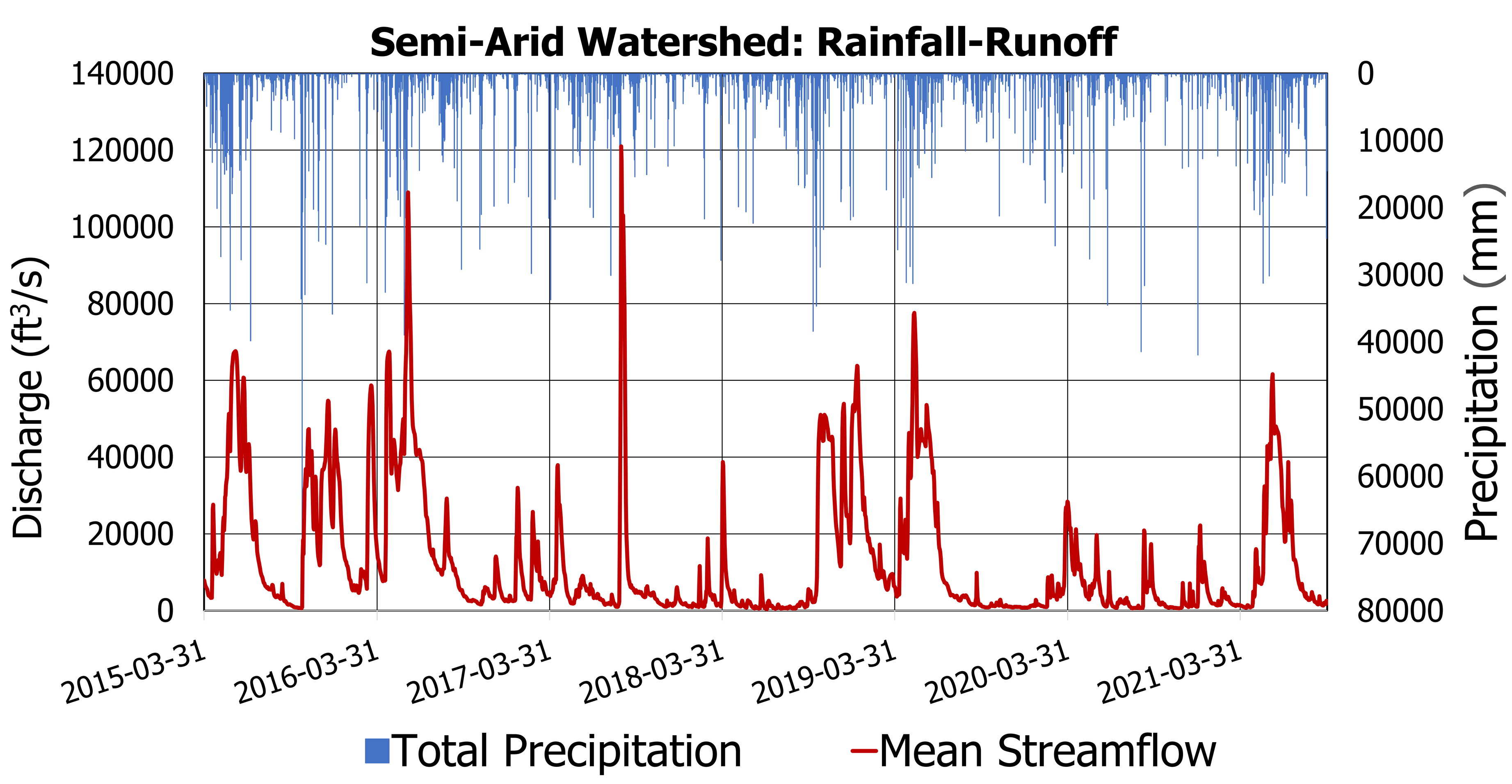
## Satellite Precipitation Data

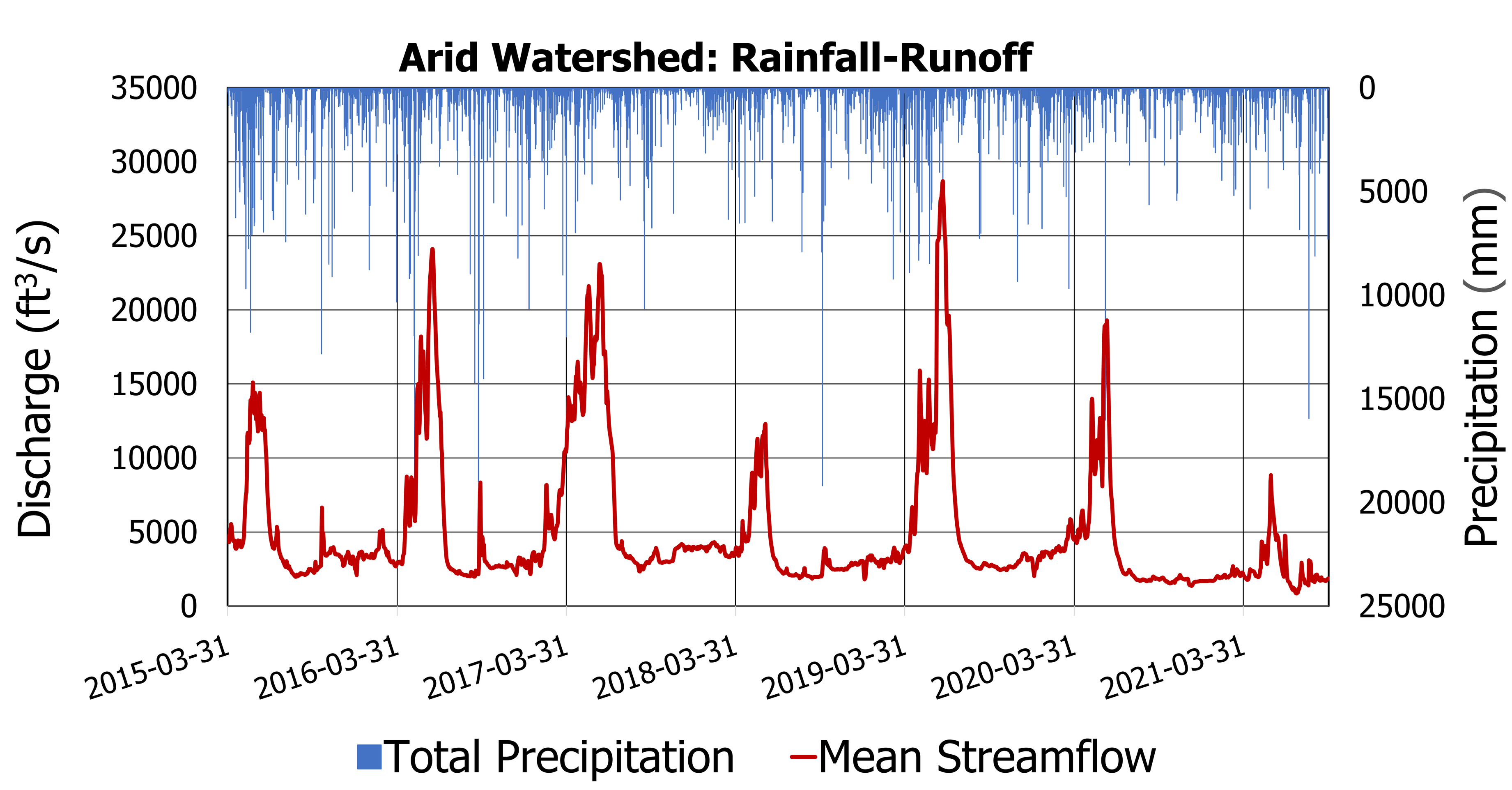
Satellite precipitation provides a **labor- and cost-effective alternative** to gage-based measurements, while better capturing the **spatial variability** of precipitation.

This study incorporates satellite-based precipitation data from **NASA's Global Precipitation Measurement (GPM) Constellation** to forecast daily-streamflow.

## The Precipitation-Runoff Relationship

This models in this study seek to represent and predict the **physical relationship** between precipitation and runoff. The graphs below **visualize this relationship** for each watershed of interest from 3/31/2015 to 9/30/2021.

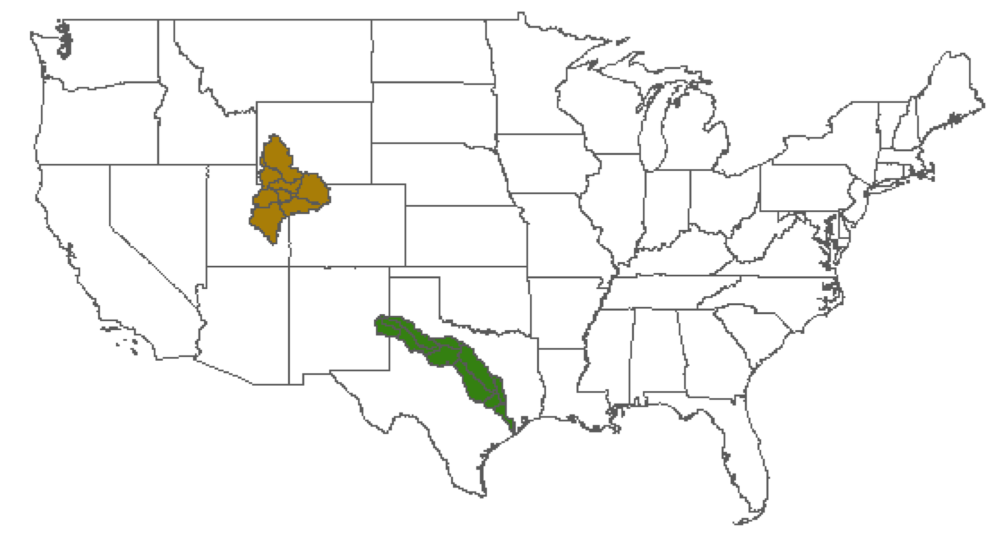


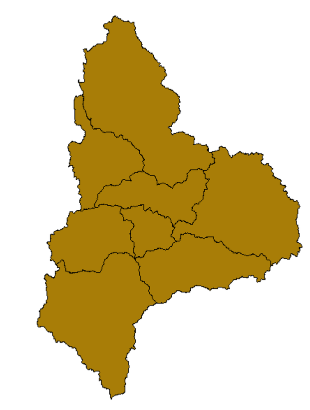




Improving flood forecasting by incorporating satellite data of precipitation and soil moisture in statistical and machine learning models.

## What are Watersheds?

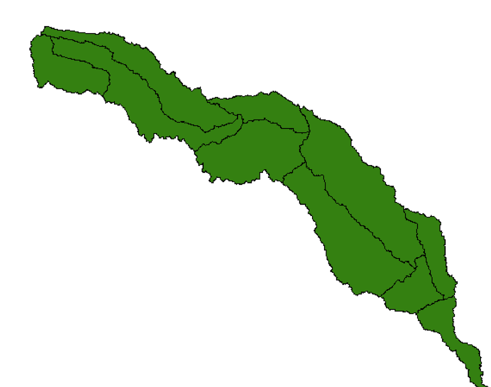
A watershed is a cohesive unit of land which entirely drains into **one exit stream** or body of water. This study examines **two watersheds** within the United States with **contrasting hydroclimates**.





**Upper Colorado River Basin**

- **Arid** climate
- Streamflow from United States Geological Survey (USGS) stream station 09328920



**Brazos Valley River Basin**

- **Semi-arid** climate
- Streamflow from USGS stream station 08116650

## Stochastic Statistical Models

### Autoregressive (AR) Model

Predicts current streamflow, $Q(t)$, from past $p$ days of streamflow.

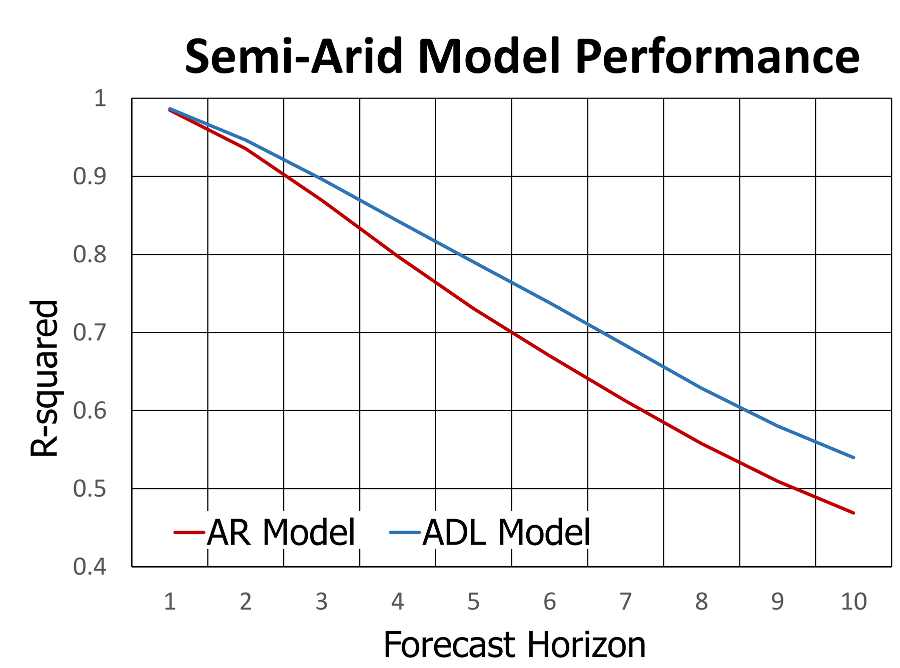
$$AR(p) := Q(t) = \phi_0 + \phi_1 Q_{t-1} + \phi_1 Q_{t-1} + \phi_2 Q_{t-2} + \ldots + \phi_p Q_{t-p} + \varepsilon_t$$

### Autoregressive Distributed Lag (ADL) Model

Predicts current streamflow from past $p$ days of streamflow and past $q$ days of precipitation, $P(t)$.

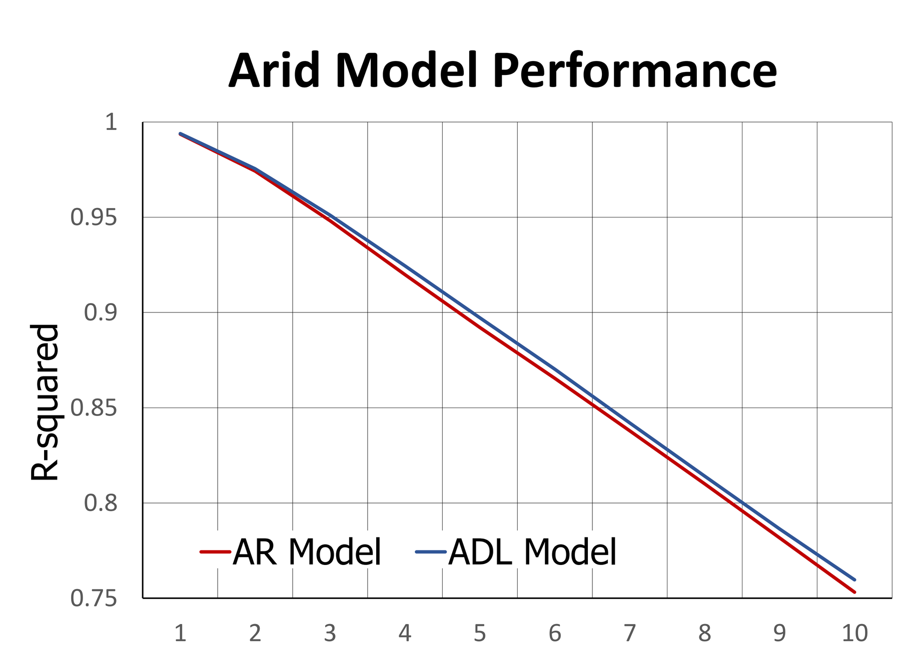
$$ADL(p, q) := Q(t) = \beta_0 + \beta_1 Q_{t-1} + \ldots + \beta_p Q_{t-p} + \delta_0 P_t + \ldots + \delta_q P_{t-q} + \varepsilon_t$$

## Results of the Study




### Semi-Arid Watershed Analysis

- **Significant improvement** from the inclusion of precipitation data, especially at **longer lead times**
- Reliable forecasting up to **2-3 days lead time** with ADL model
- **Greater variability** in streamflow reduces accuracy of models compared to arid region at longer lead times




### Arid Watershed Analysis

- Minimal improvement from inclusion of precipitation data because of sparse precipitation in the region
- Reliable forecasting ($R^2 > 0.9$) up to **4-5 days lead time** with ADL model
- Relatively **steady streamflow** allows for more accurate forecasting

## What is Next?

- Incorporating **soil moisture stress (SMS)** data from **NASA's Soil Moisture Active Passive (SMAP) satellite.**
- Additional models trained using **machine learning algorithms** to better capture the true, **nonlinear relationship**