# A probabilistic approach for estimating monthly catchment water balances from satellite and ground data

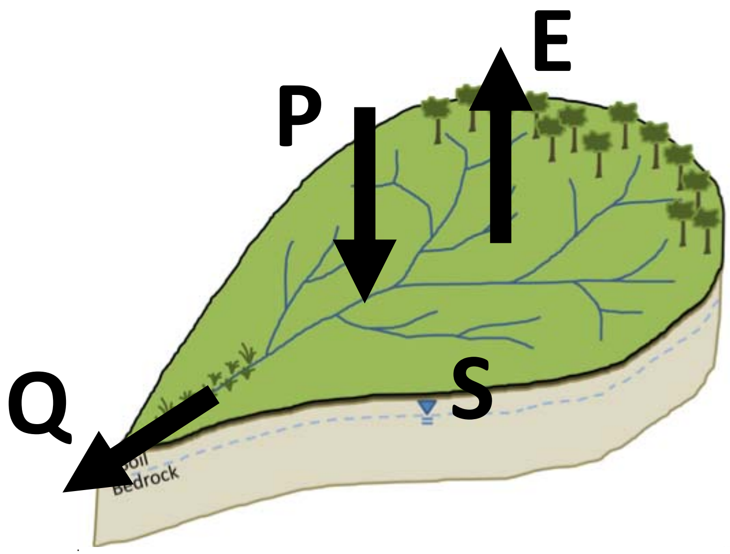
Gerrit Schoups
Delft University of Technology

# Goal

- Estimate monthly basin water balance terms from data and water balance constraints

$$S_t = S_{t-1} + P_t - E_t - Q_t$$

- $S$ is storage
- $P$ is precipitation
- $E$ is evaporation
- $Q$ is river discharge



# Overview

**Data**

- Off-the-shelf monthly data for each water balance term
- Period: 2006-2015; 346 MOPEX basins

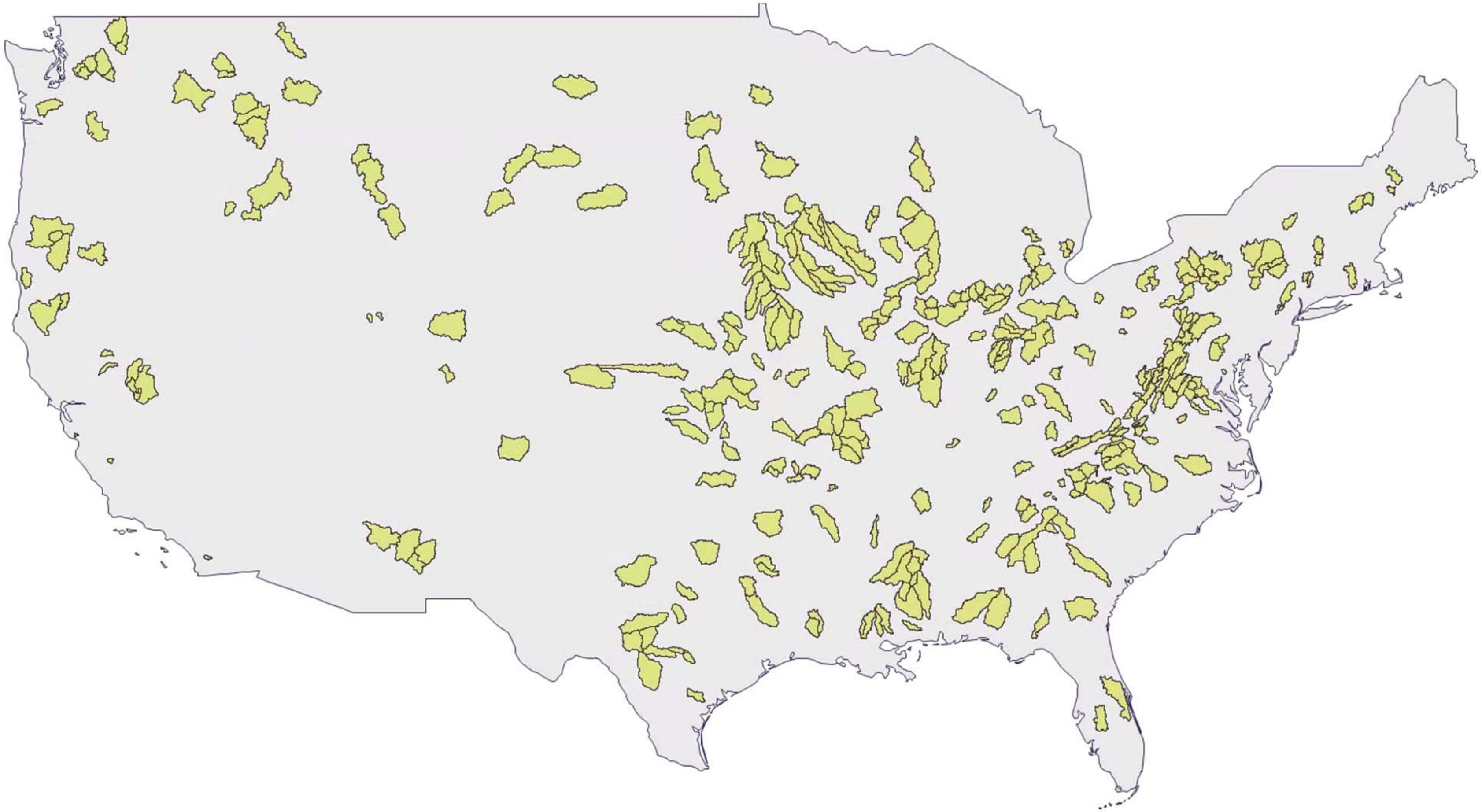
**Model**

- Monthly water balance constraints
- Parameterized systematic and random data errors
- Joint Bayesian estimation of water balance terms and error parameters

**Results**

- How accurate are estimates of monthly basin-scale water balance terms?
- How large are systematic and random data errors?
- Any relation to basin characteristics?

# MOPEX basins



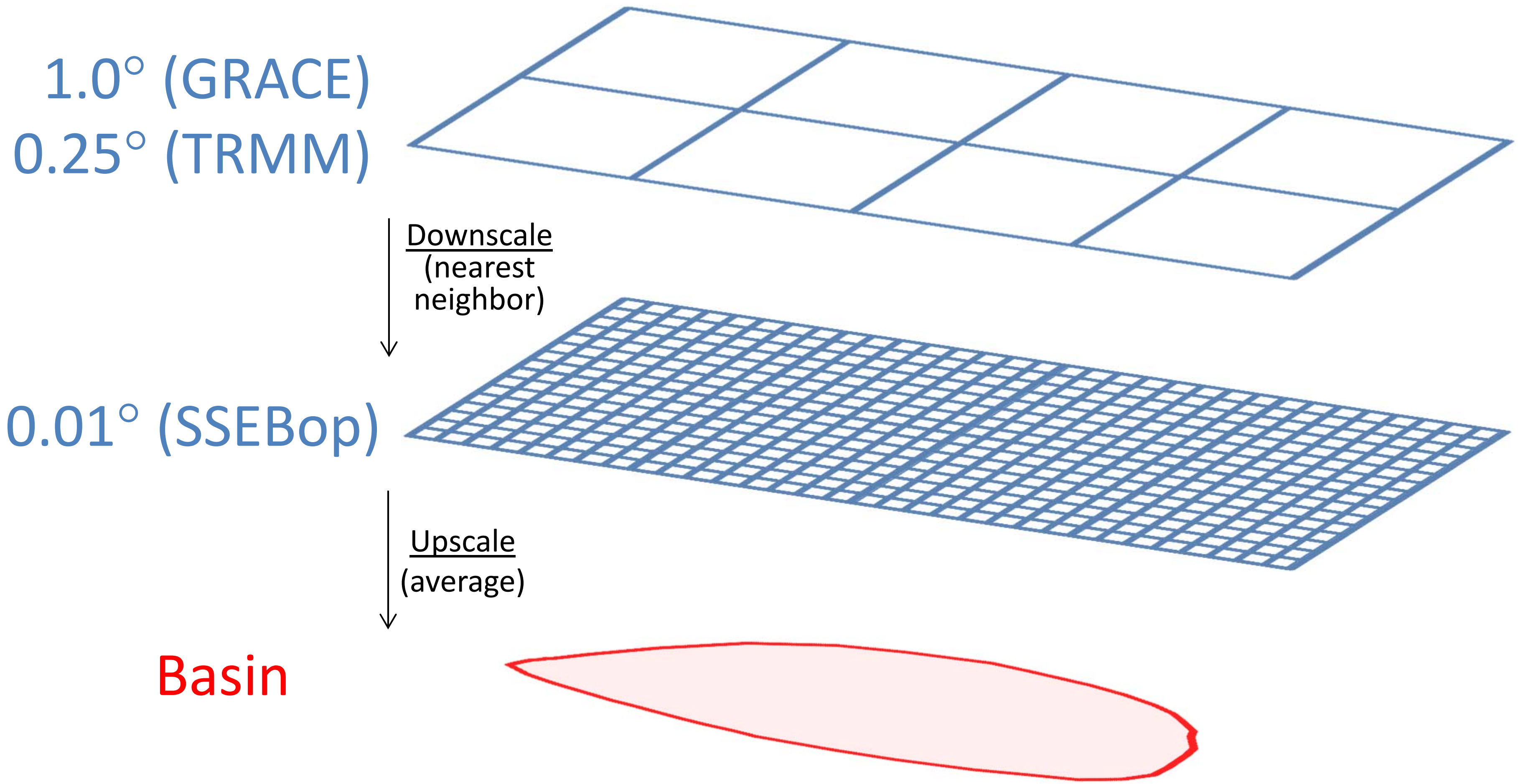
Basin sizes: 60 – 10,000 km$^2$

# Monthly data

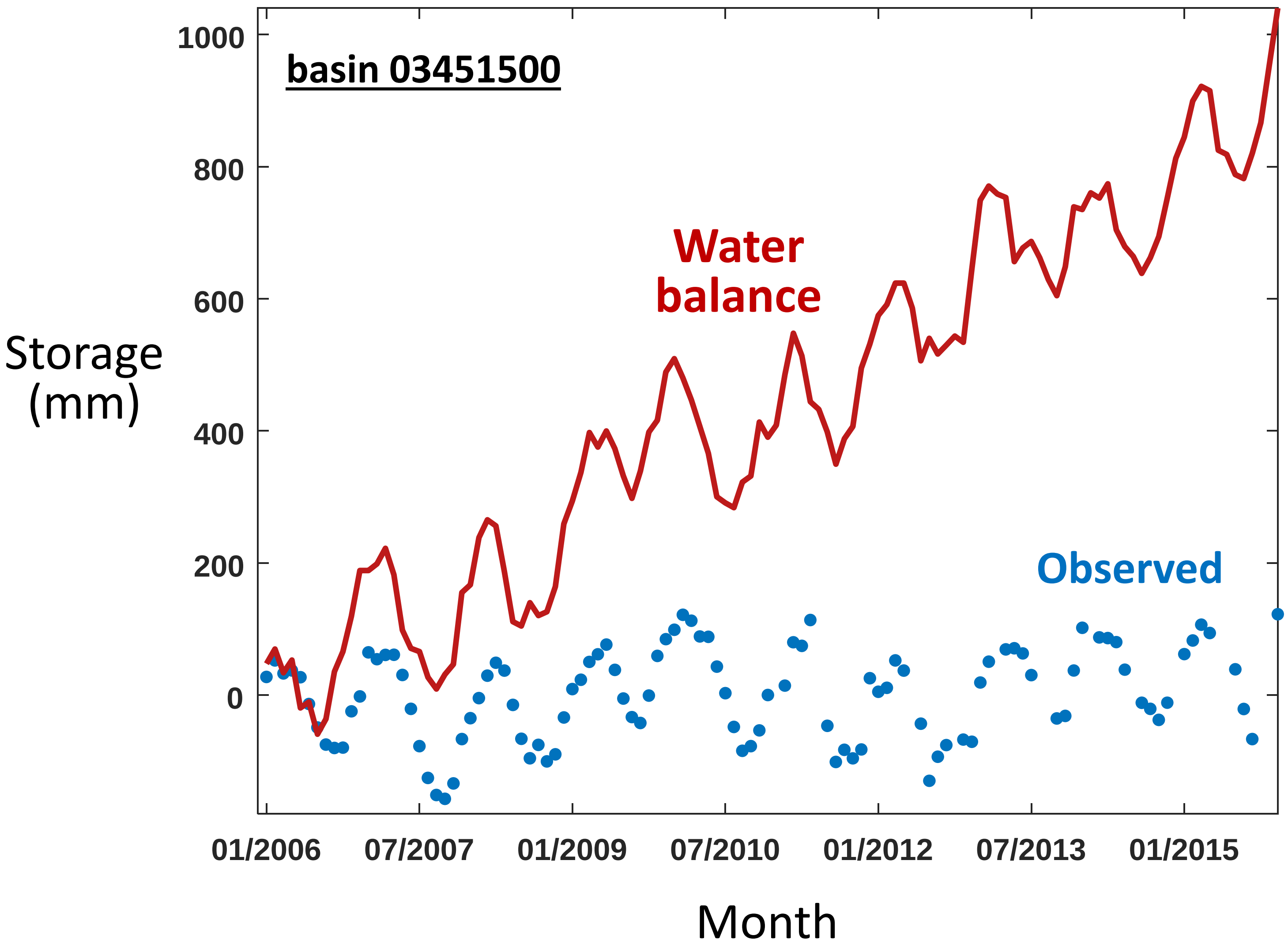
| | Source | Type | Spatial resolution |
|---|---|---|---|
| Precipitation | TRMM-B43 | Satellite + ground | 0.25° |
| Evaporation | SSEBop | Satellite | 0.01° |
| River discharge | USGS stream gauges | Ground | Basin |
| Storage | GRACE | Satellite | 1.0° |

Time period: 2006-2015

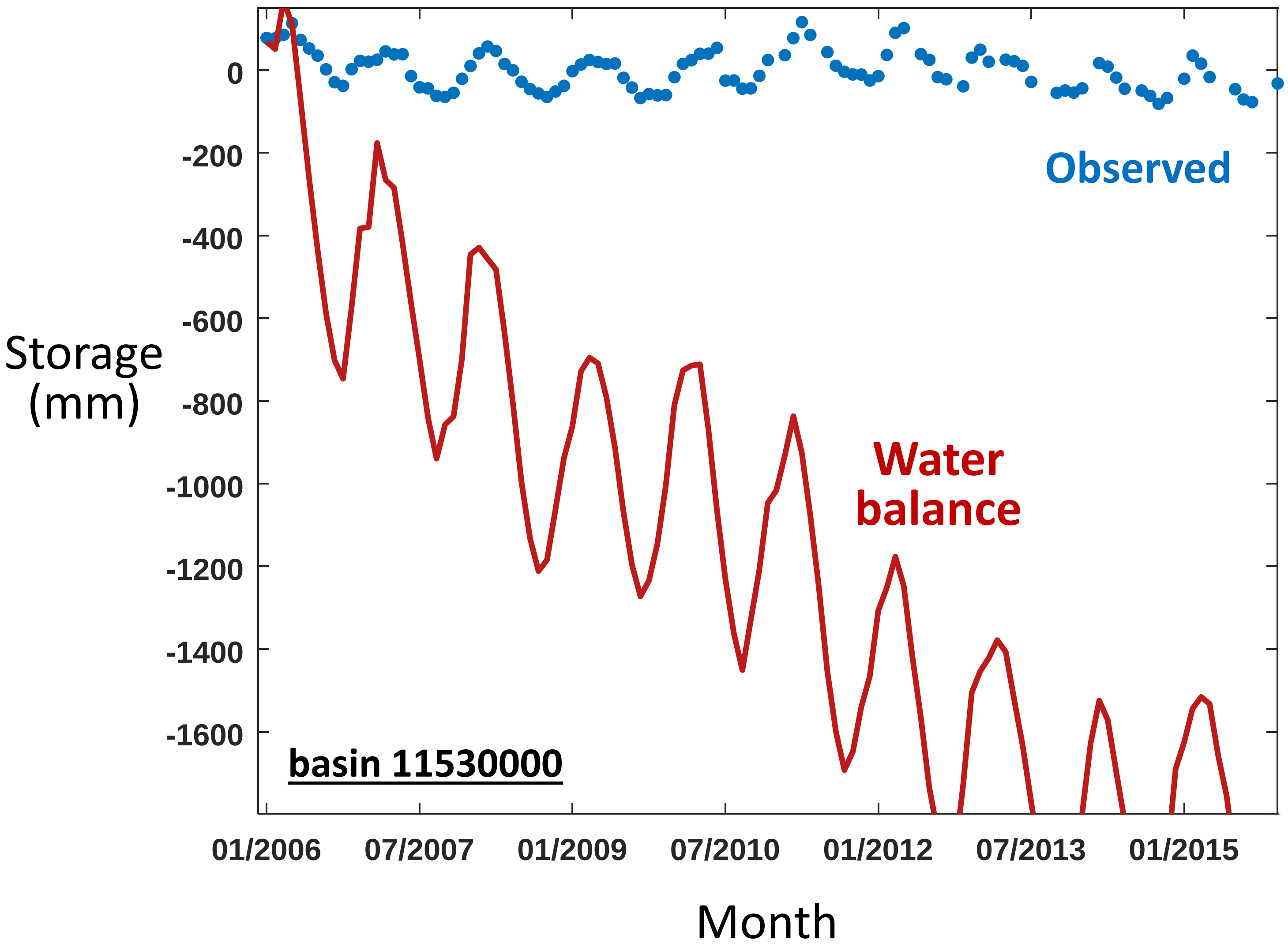
# Spatial scaling



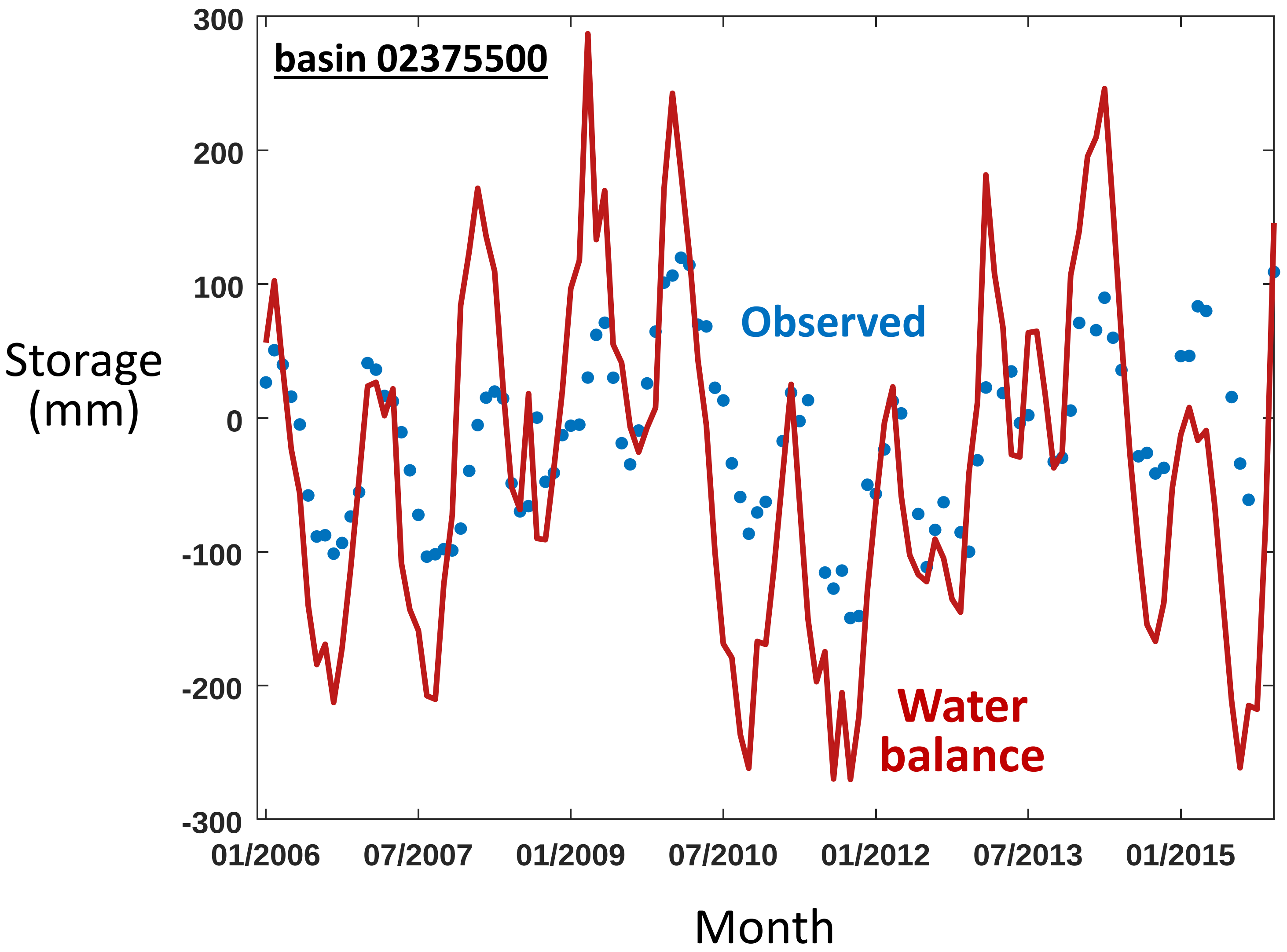
# How well do data fit water balance?



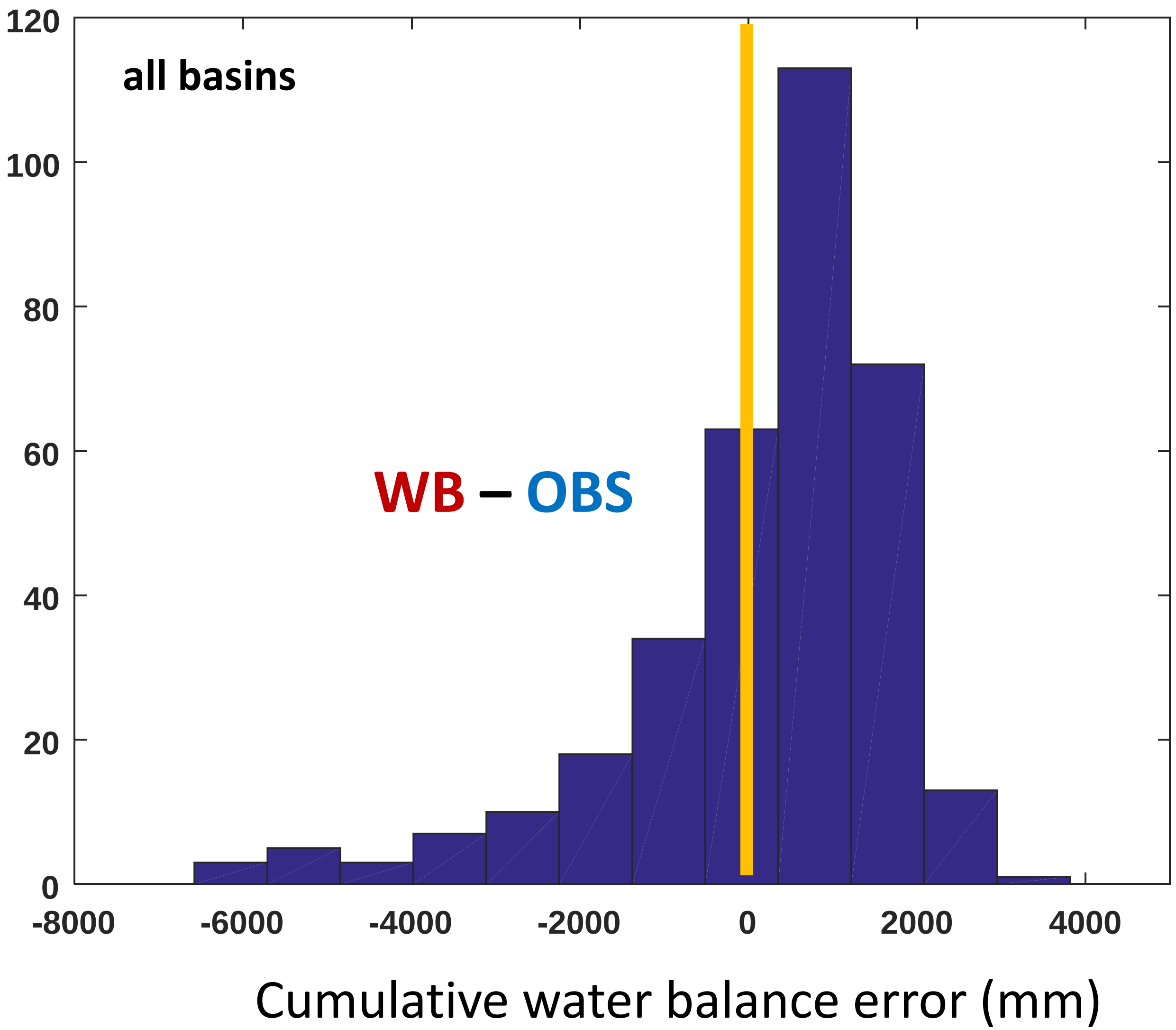
# How well do data fit water balance?



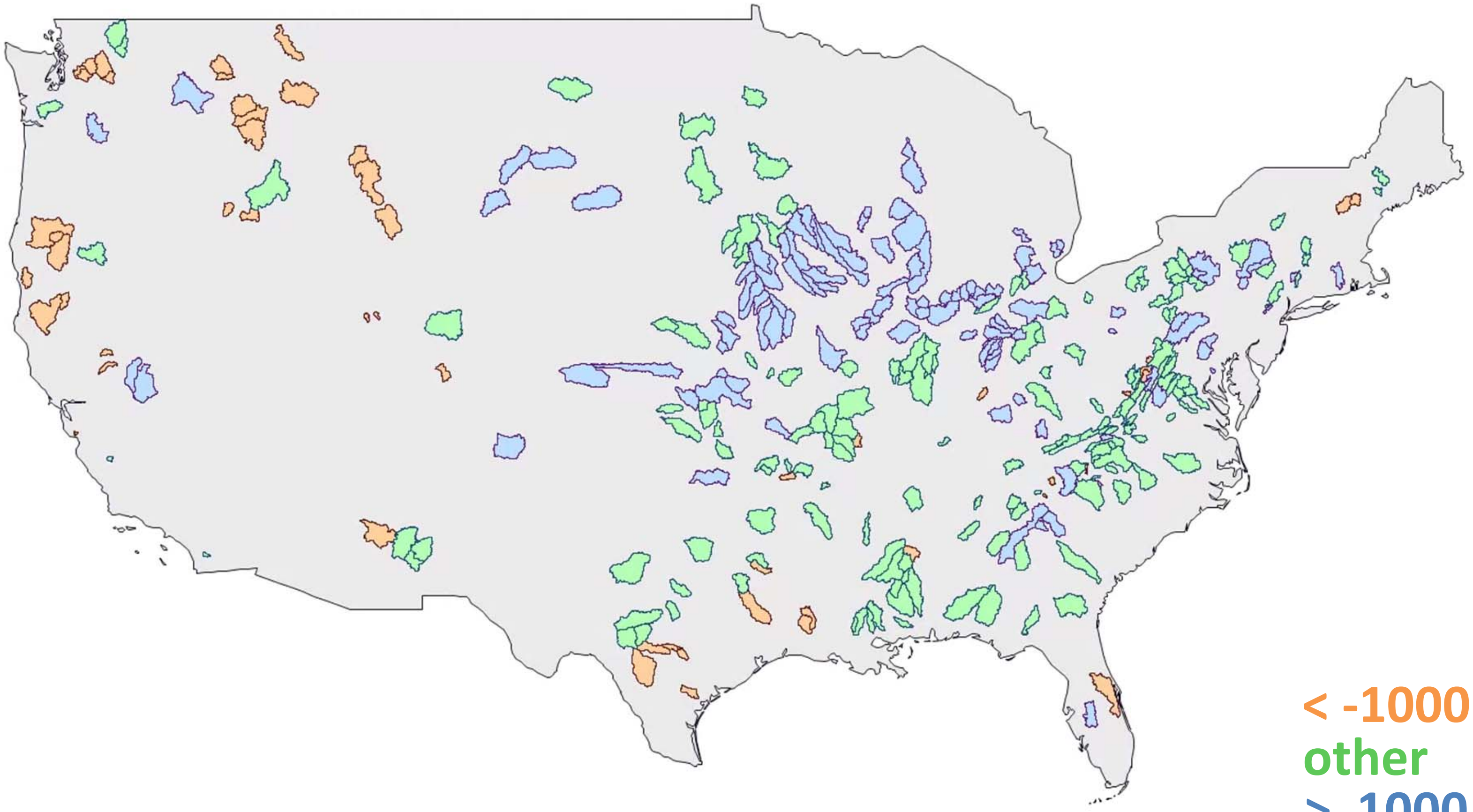
# How well do data fit water balance?



# How well do data fit water balance?



# How well do data fit water balance?




Cumulative water balance error (mm)

# Probabilistic model

- Systematic and random deviations between true and observed water balance terms

$X^{true} \sim N(fX^{obs}, \sigma^2)$ constant noise

$X^{true} \sim N(fX^{obs}, (fX^{obs}CV)^2)$ proportional noise

- $f$ is scaling factor (multiplicative bias)
- $\sigma$ is standard deviation
- $CV$ is coefficient of variation

# Probabilistic model: parameters

| | Noise parameter | Prior (mode, CV) | Bias parameter | Prior (mode, CV) |
|---|---|---|---|---|
| Evaporation | $CV_E$ | $Gam(0.3, 0.9)$ | $f_E$ | $Gam(1, 0.9)$ |
| Storage | $\sigma_S$ | $Gam(30, 0.9)$ | $f_S$ | $Gam(1, 0.9)$ |
| River discharge | $CV_Q$ | $Gam(0.1, 0.4)$ | $f_Q$ | $\delta(1.0)$ |
| Precipitation | $\sigma_P = (1-w)\sigma_{trmm}^{nc} + w\sigma_{trmm}^{pc}$ <br> $w \sim U(0,1)$ | | $f_P$ | $\delta(1.0)$ |

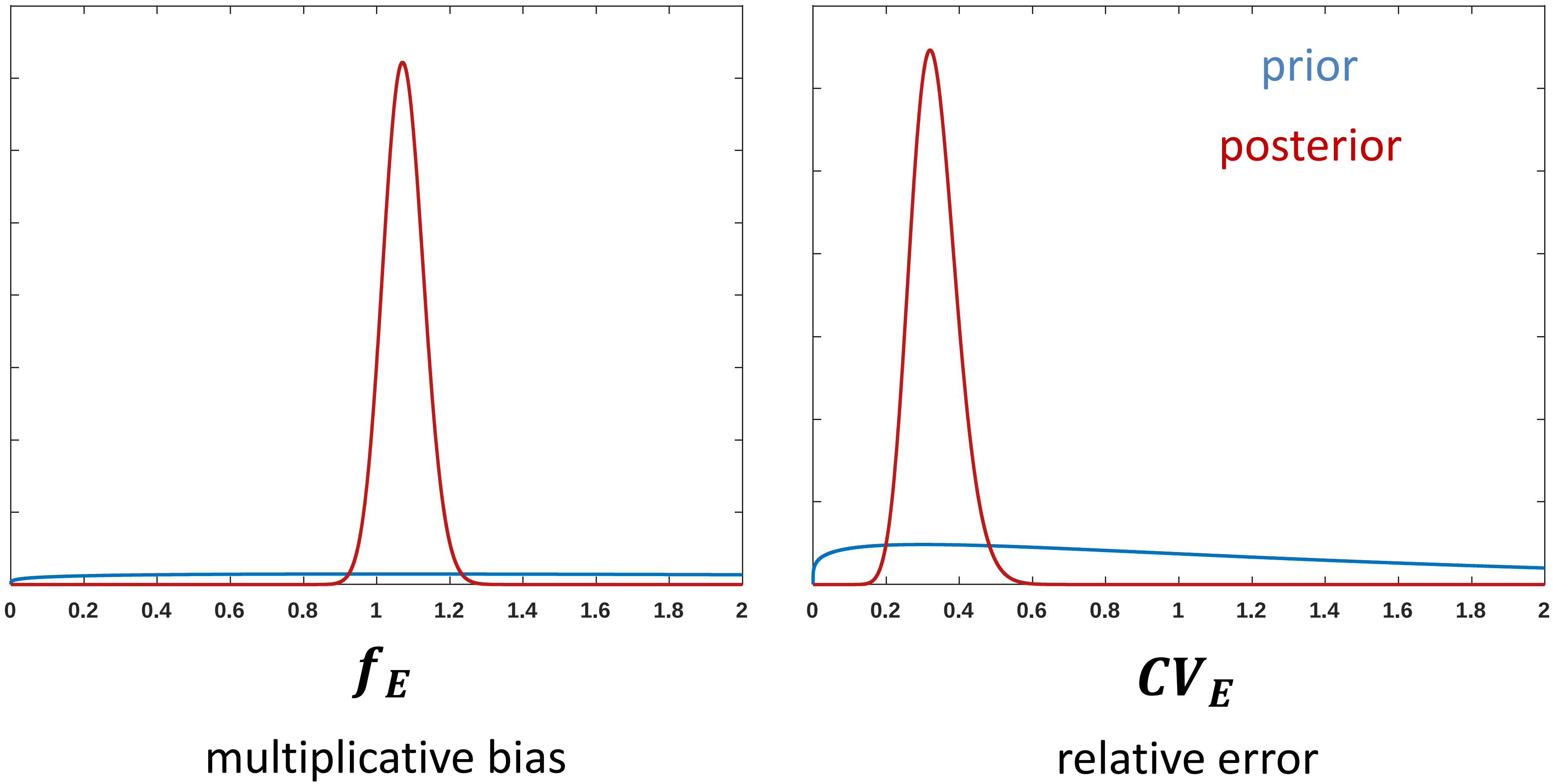
- TRMM and stream gauge data: assume unbiased
- Six parameters to be estimated

# Solving the model

- Compute posterior distributions of
  - parameters: how large are systematic and random data errors?
  - water balance terms: how accurately can each be estimated?
- Methods
  - Sampling (MCMC) for parameters
  - Kalman smoothing for water balance terms

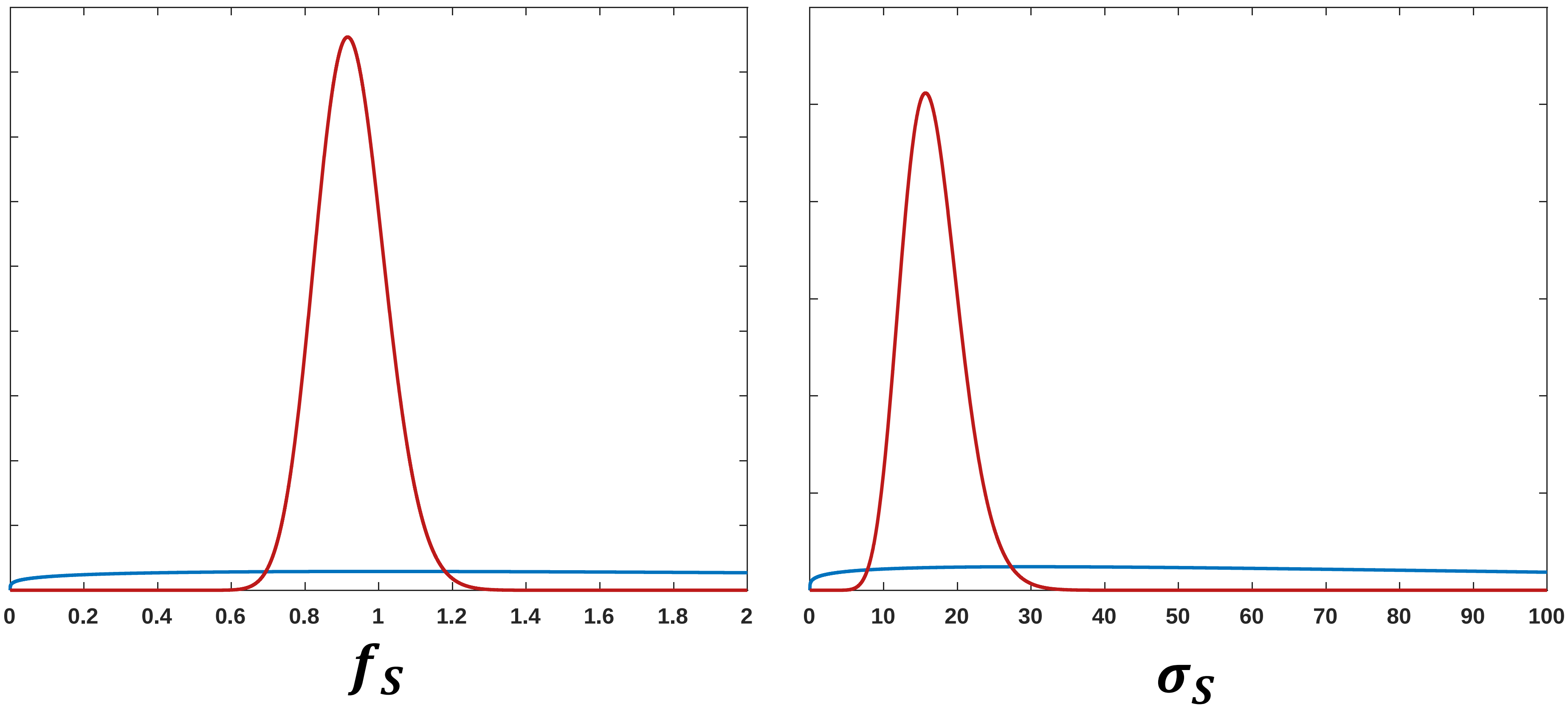
# How large are **SSEBop** data errors?

**parameter posteriors – <u>basin 03451500</u>**



# How large are **GRACE** data errors?

**parameter posteriors – <u>basin 03451500</u>**

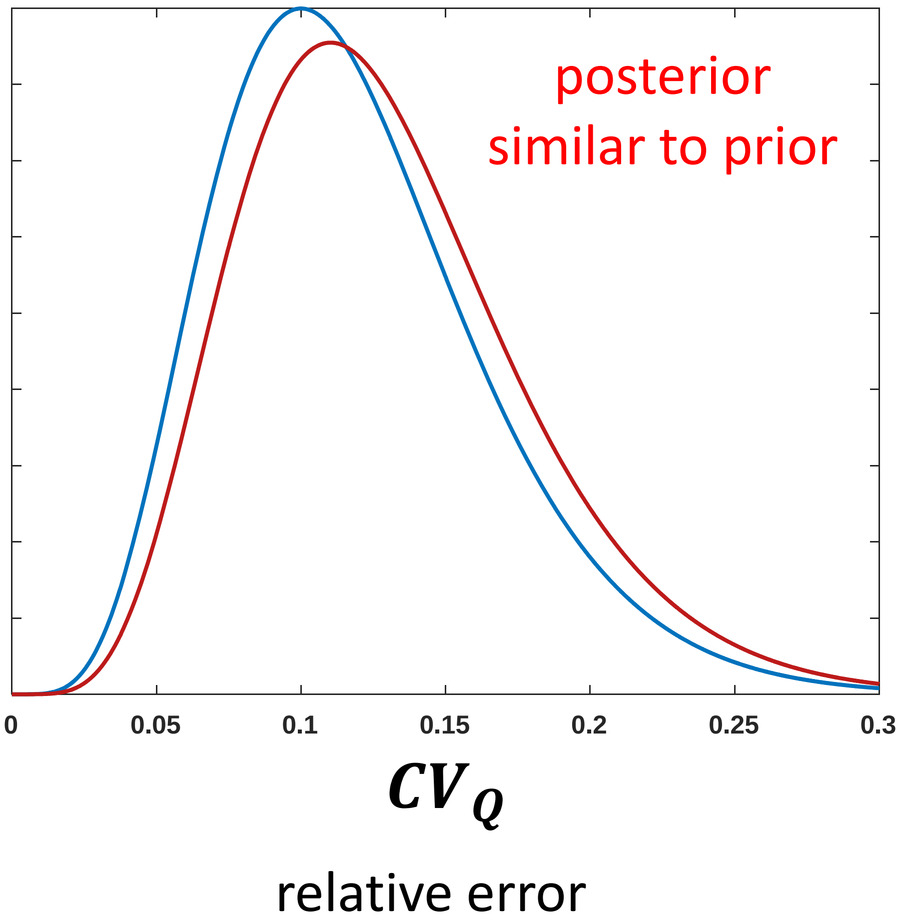


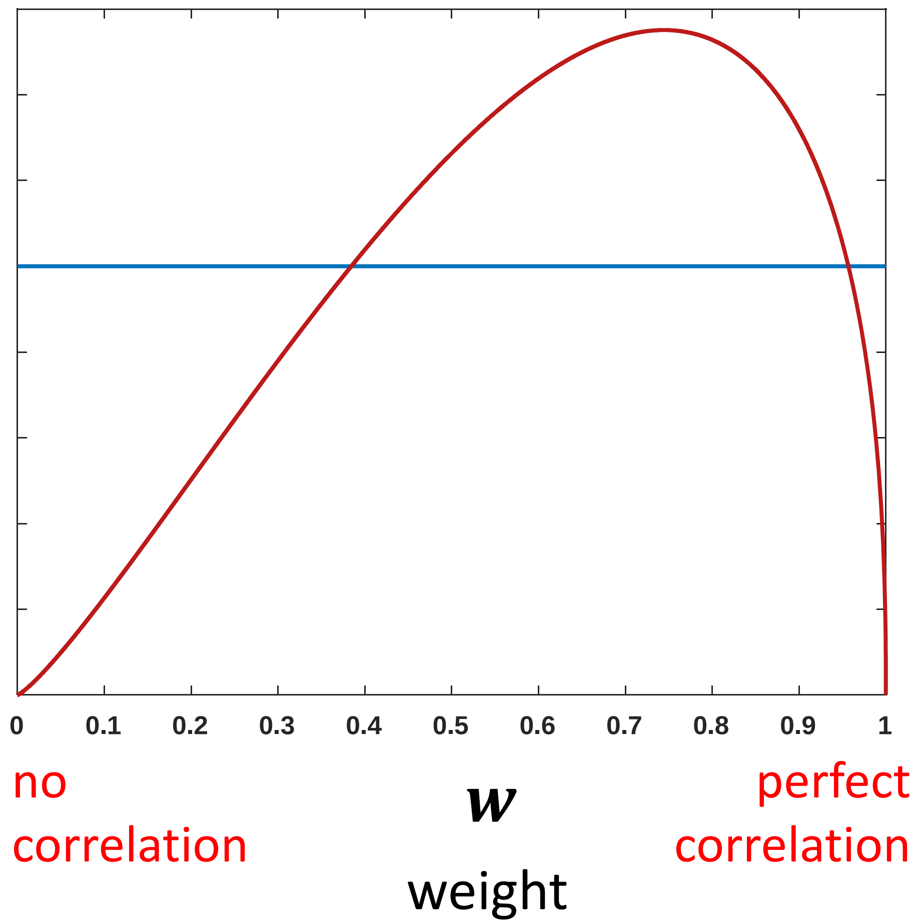

multiplicative bias

absolute error (mm)

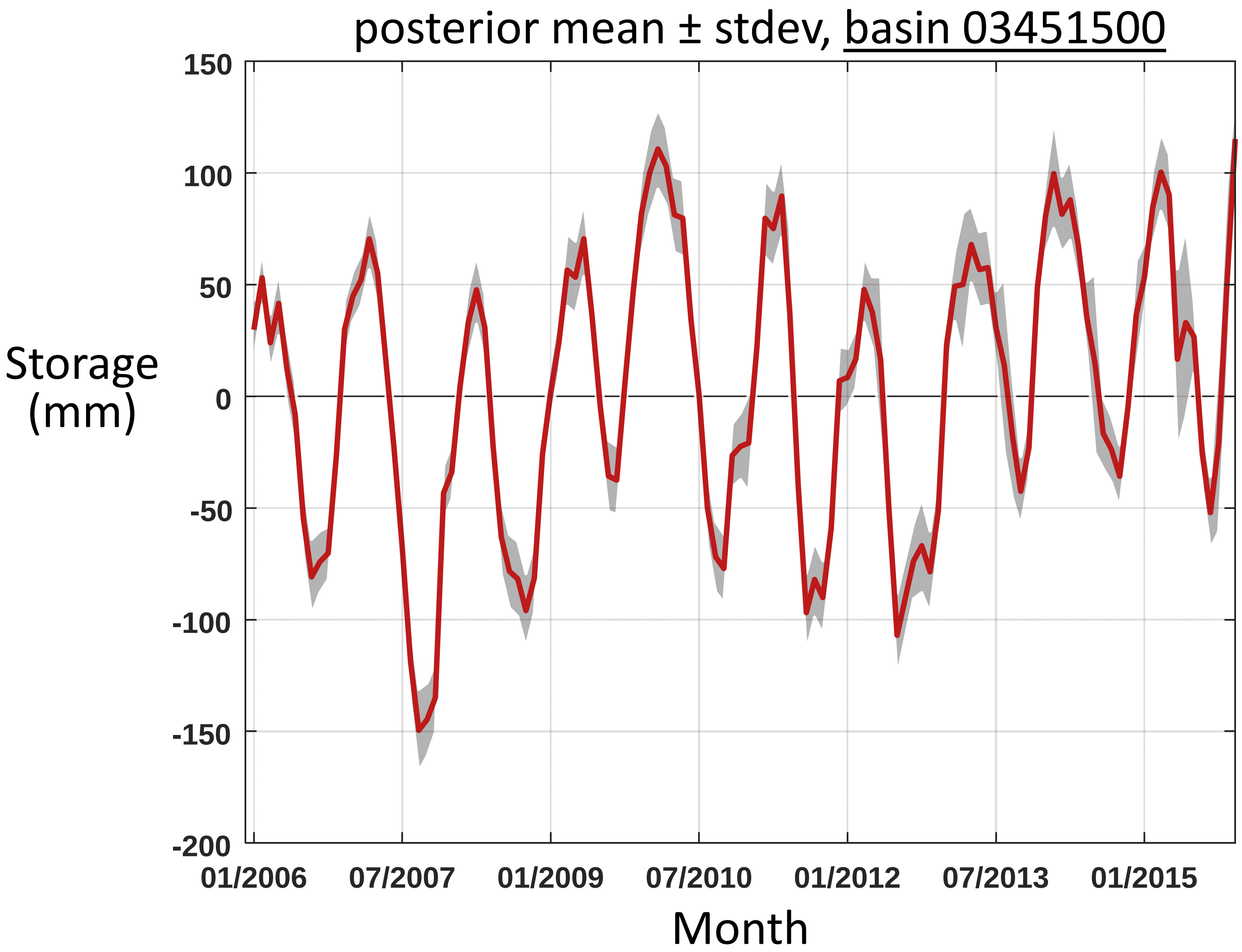
# How large are **Q** & **P** data errors?

**parameter posteriors – <u>basin 03451500</u>**

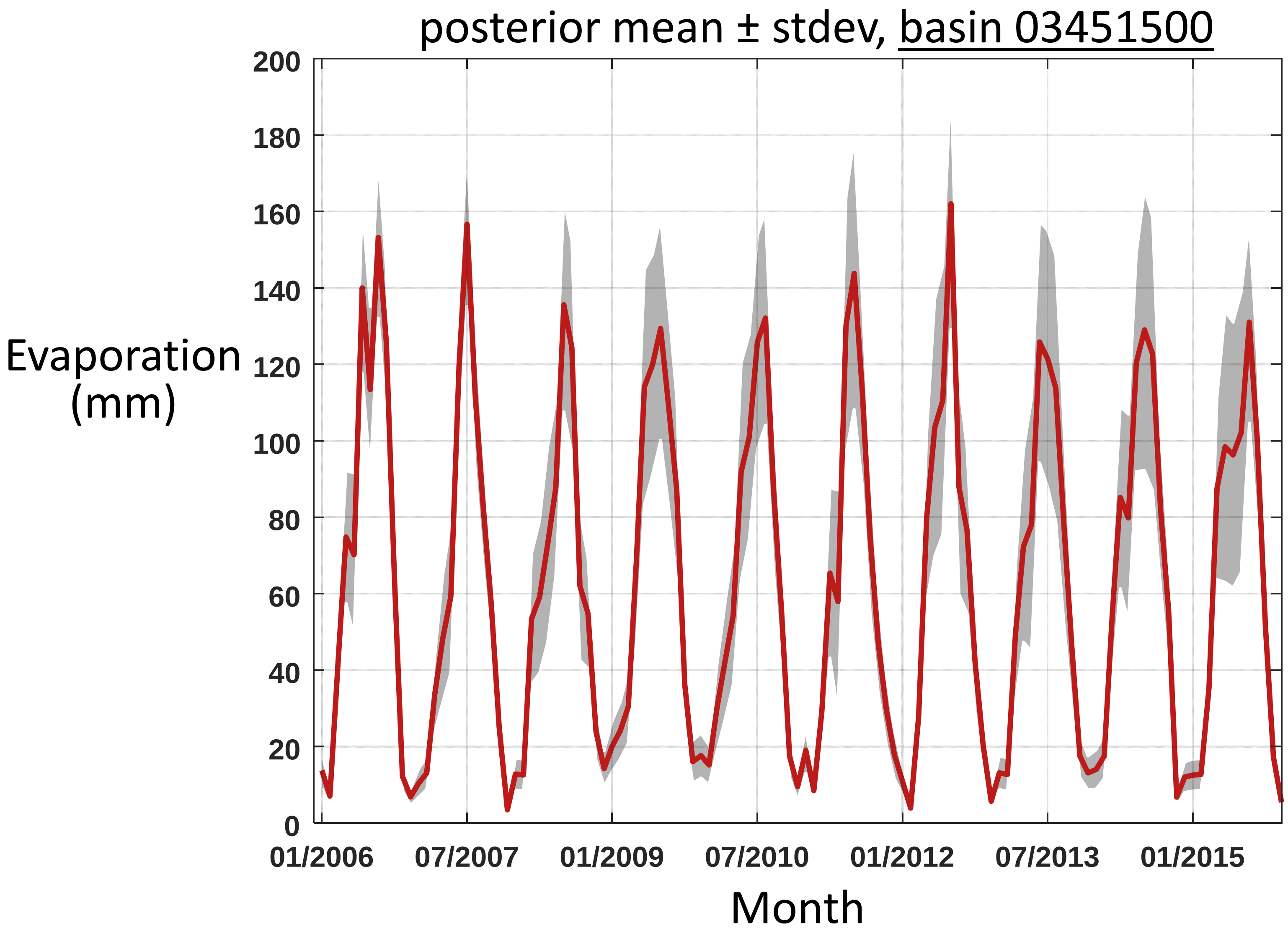


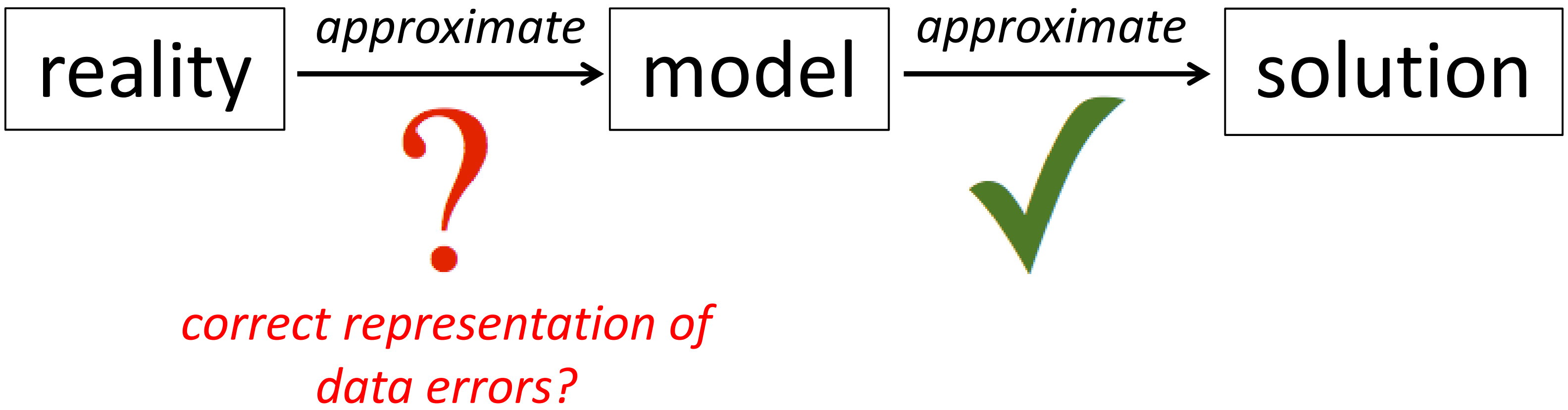
# How accurately can $\boldsymbol{S}$ be estimated?



# How accurately can $\boldsymbol{E}$ be estimated?

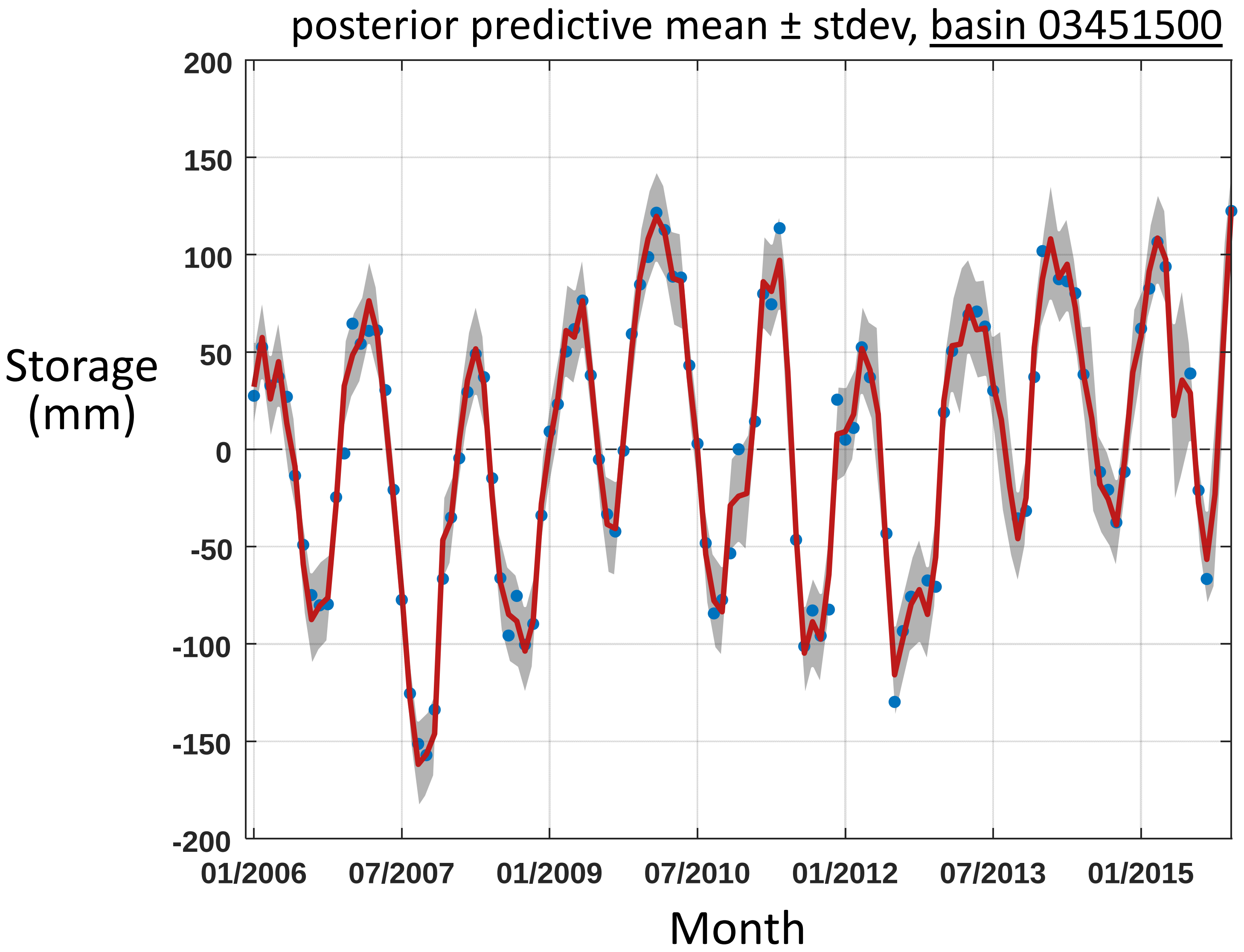


# How reliable are these results?

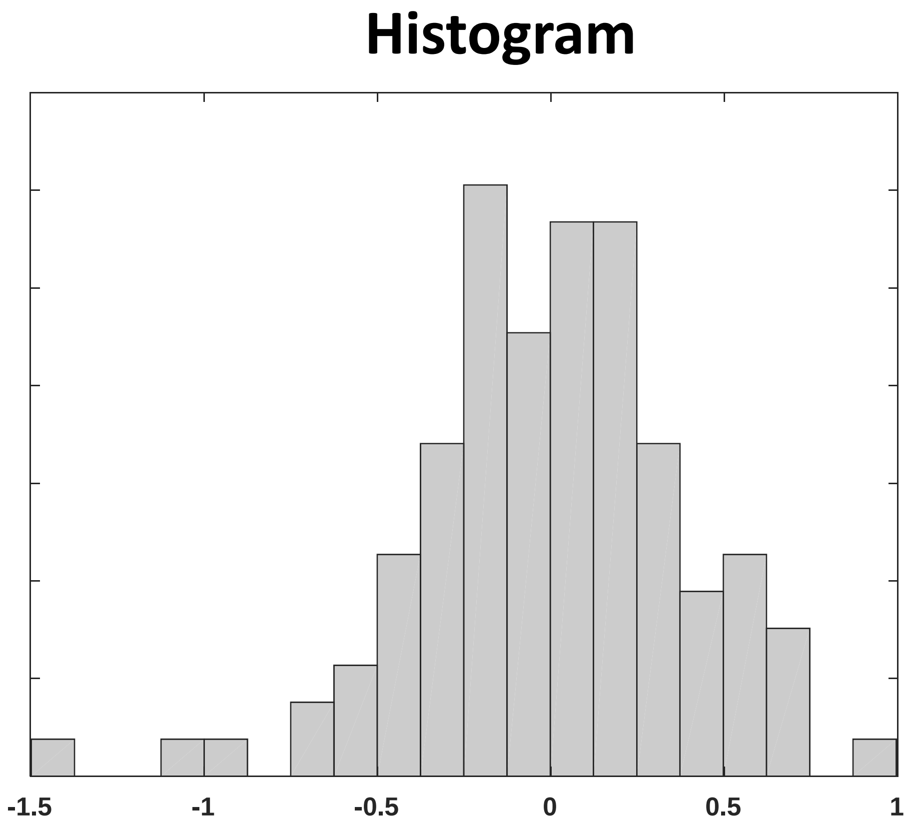



- Model checking
  - compare model-predicted data to actual data
  - residual plots

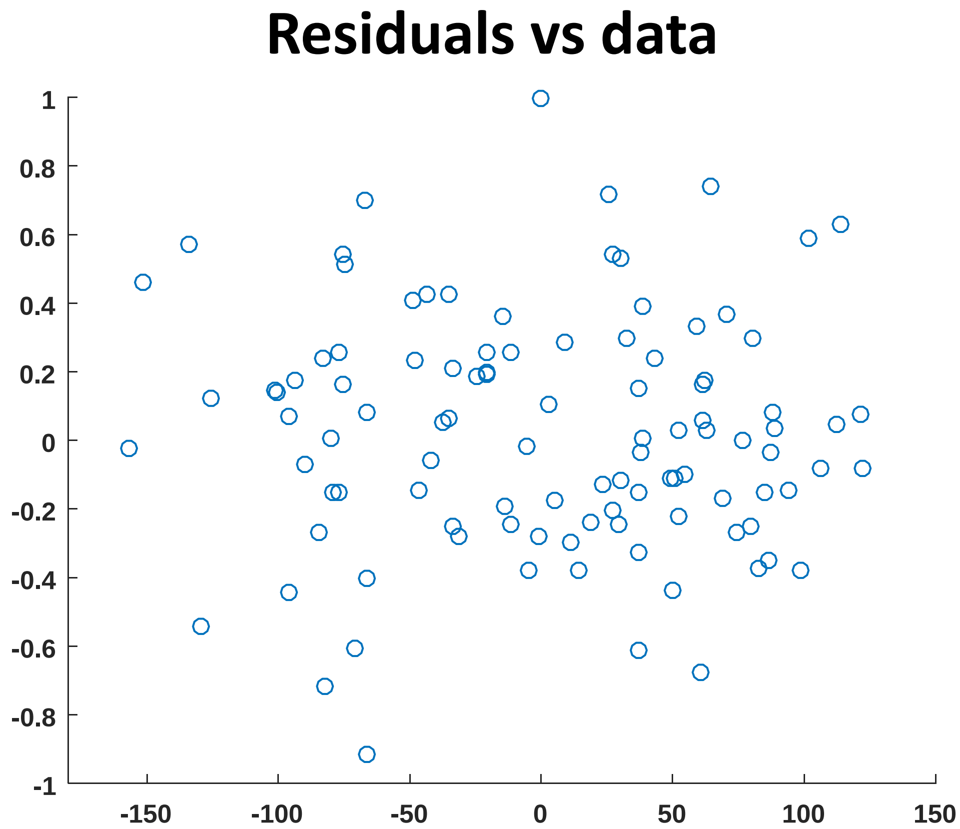
# Model-predicted **GRACE** data



# Residual plots for $\boldsymbol{S}$

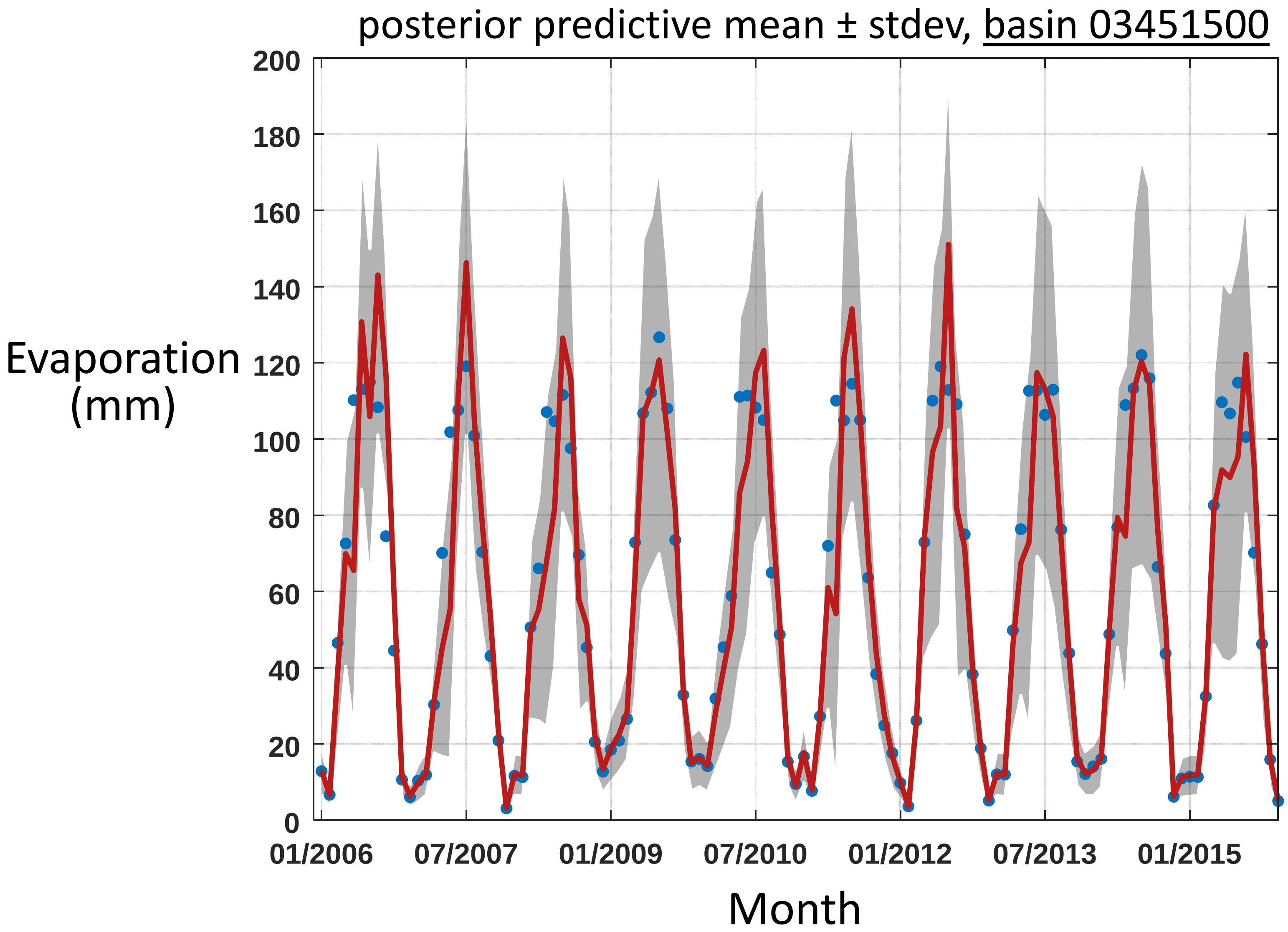



should look Gaussian



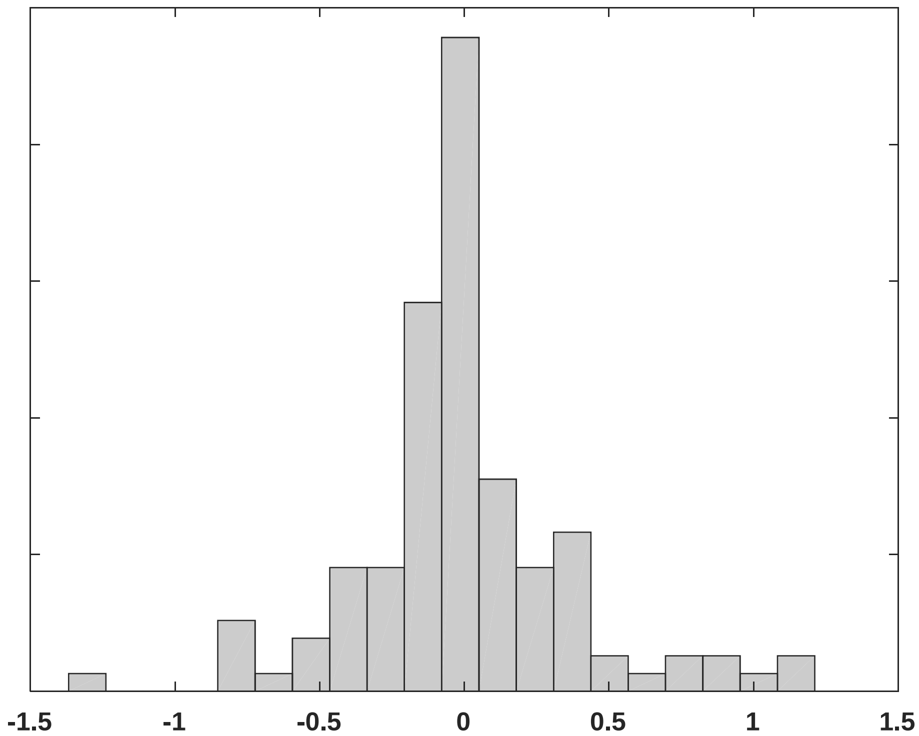

should look random

# Model-predicted **SSEBop** data

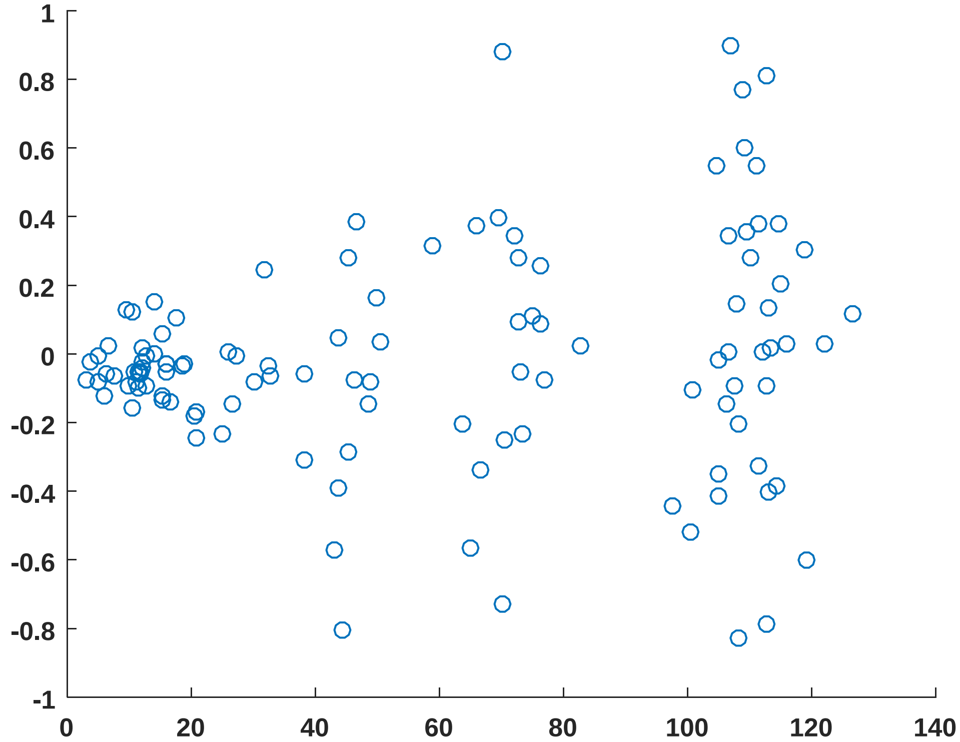


# Residual plots for $\boldsymbol{E}$

**Histogram**




**Residuals vs data**

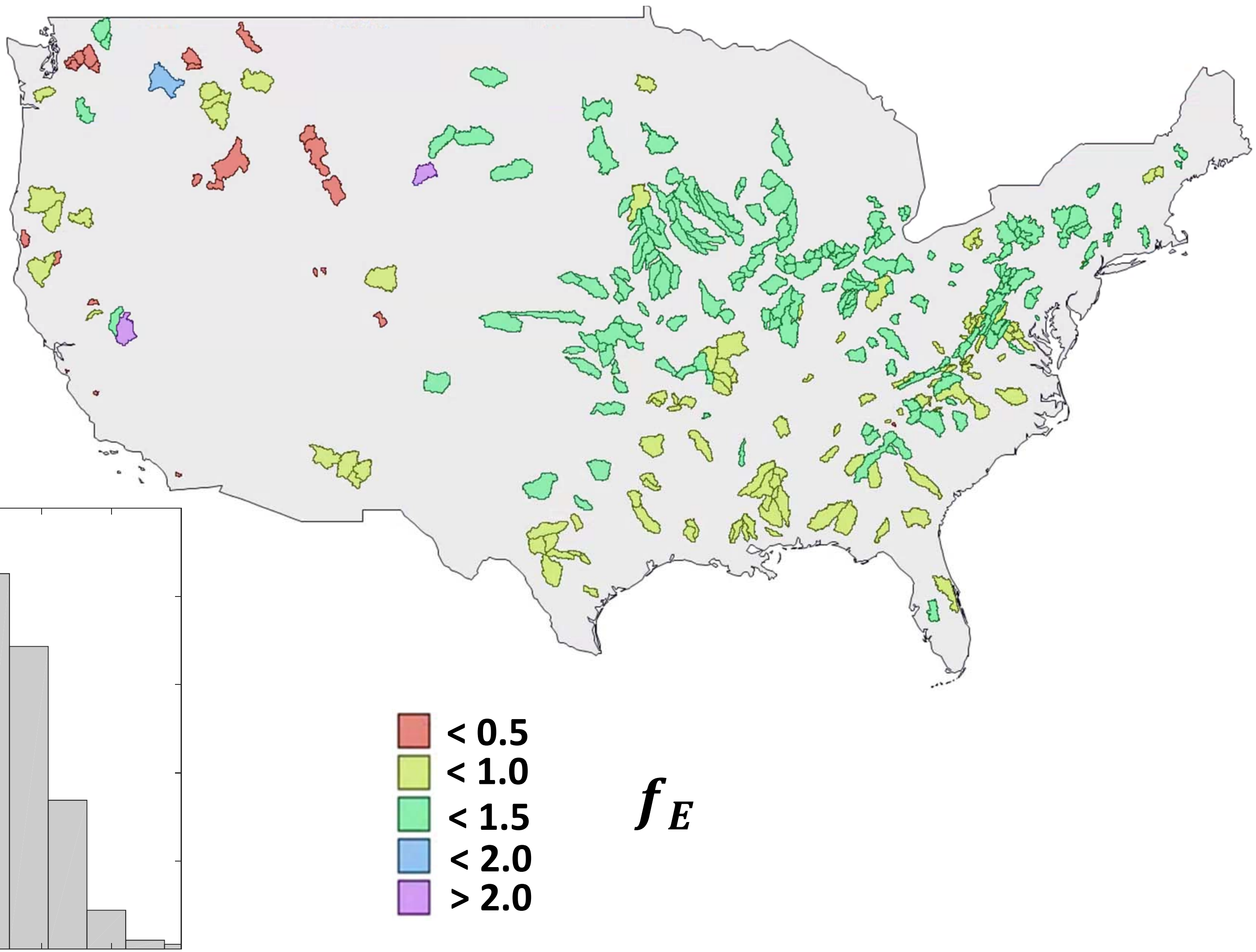



should look Gaussian

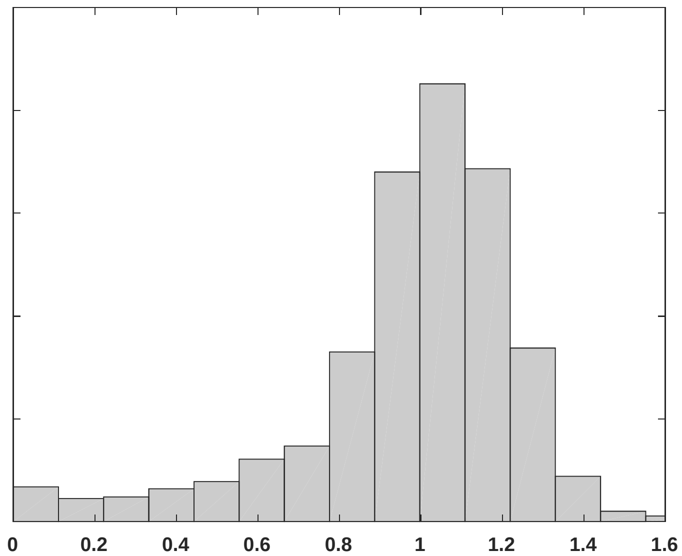
should look random

# **SSEBop** multiplicative bias: all basins

**posterior means**

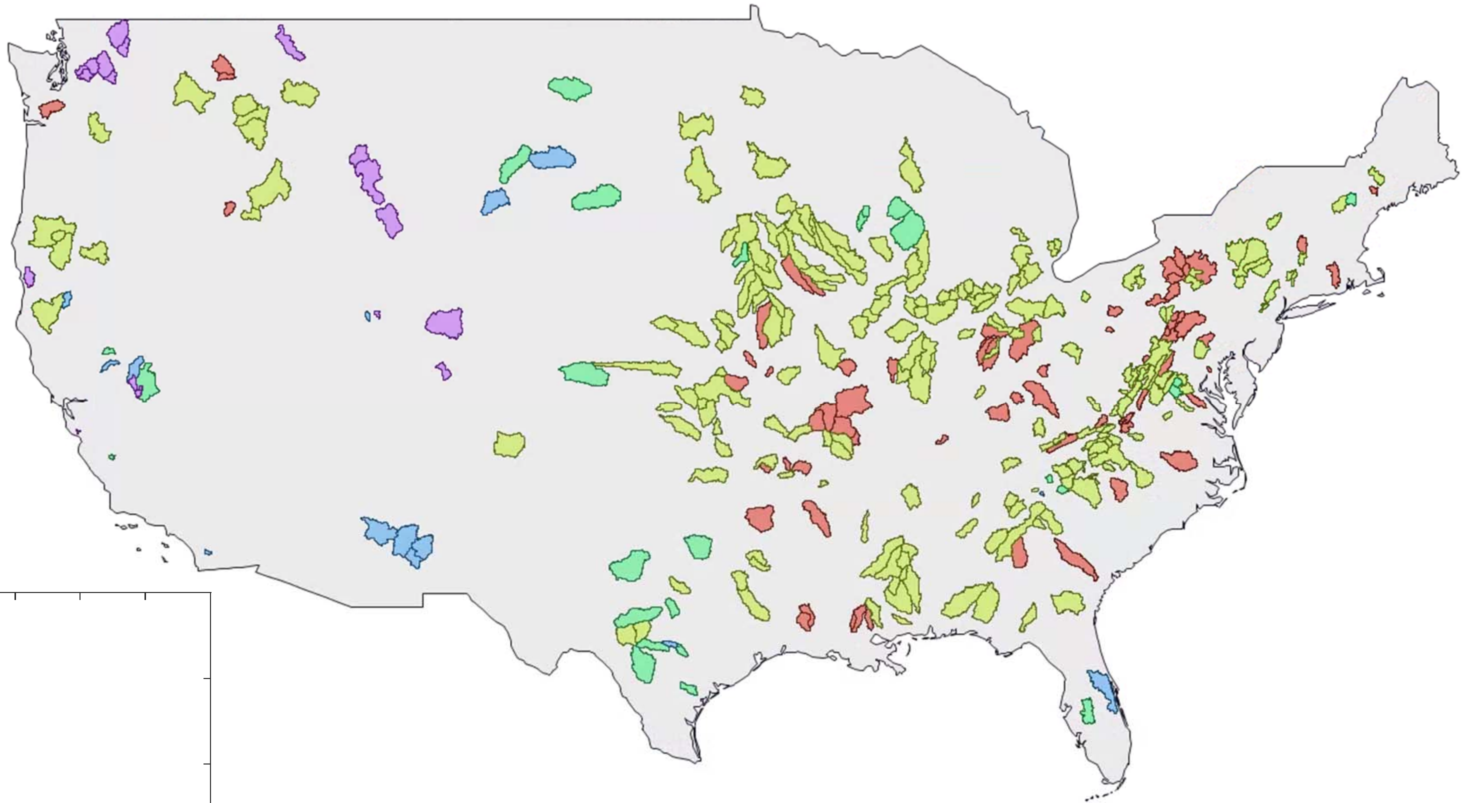



**all posteriors**



# **SSEBop** relative error: all basins

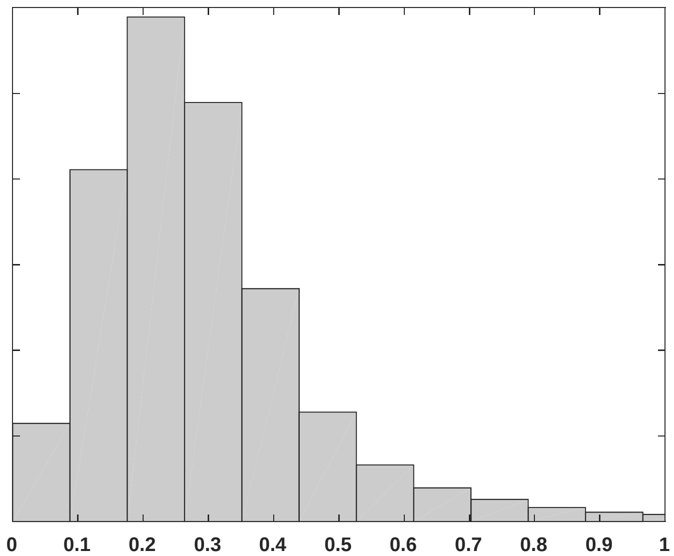
**posterior means**



< 0.2
< 0.4
< 0.6
< 1.0
> 1.0

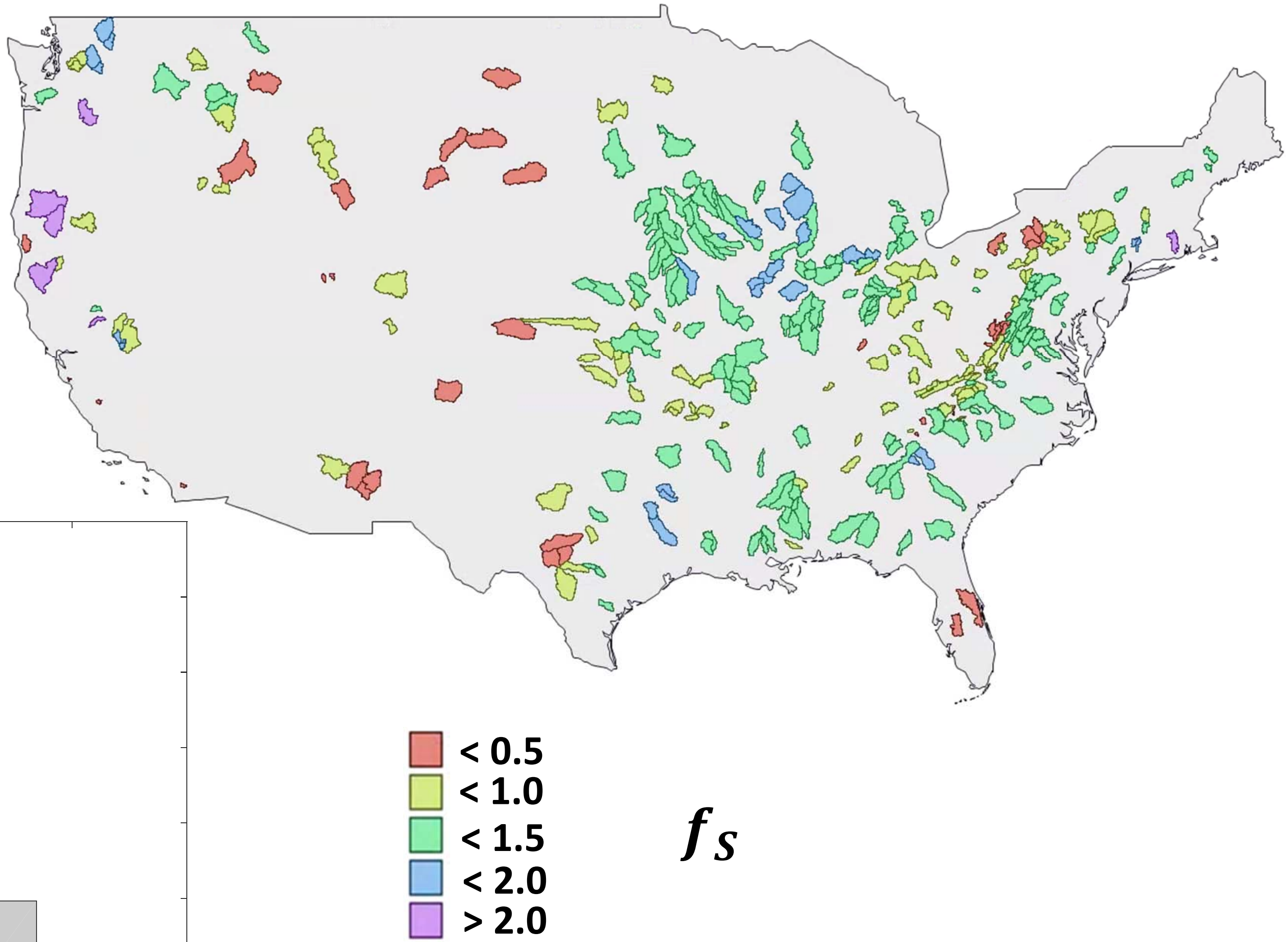
$CV_E$

**all posteriors**



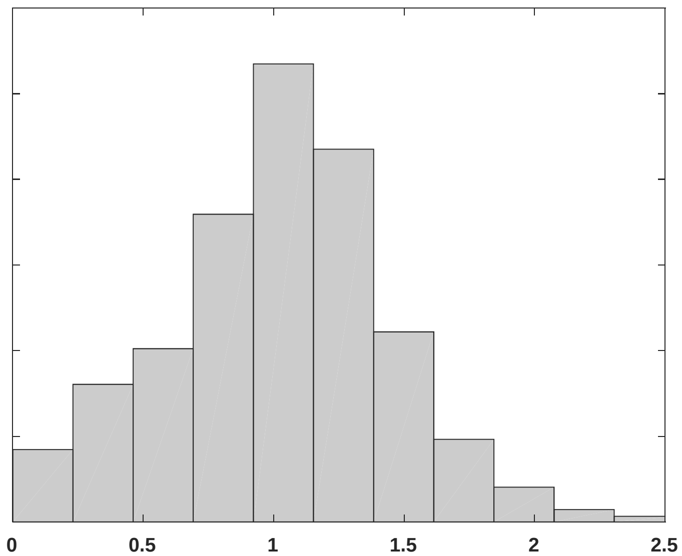
# **GRACE** multiplicative bias: all basins

**posterior means**



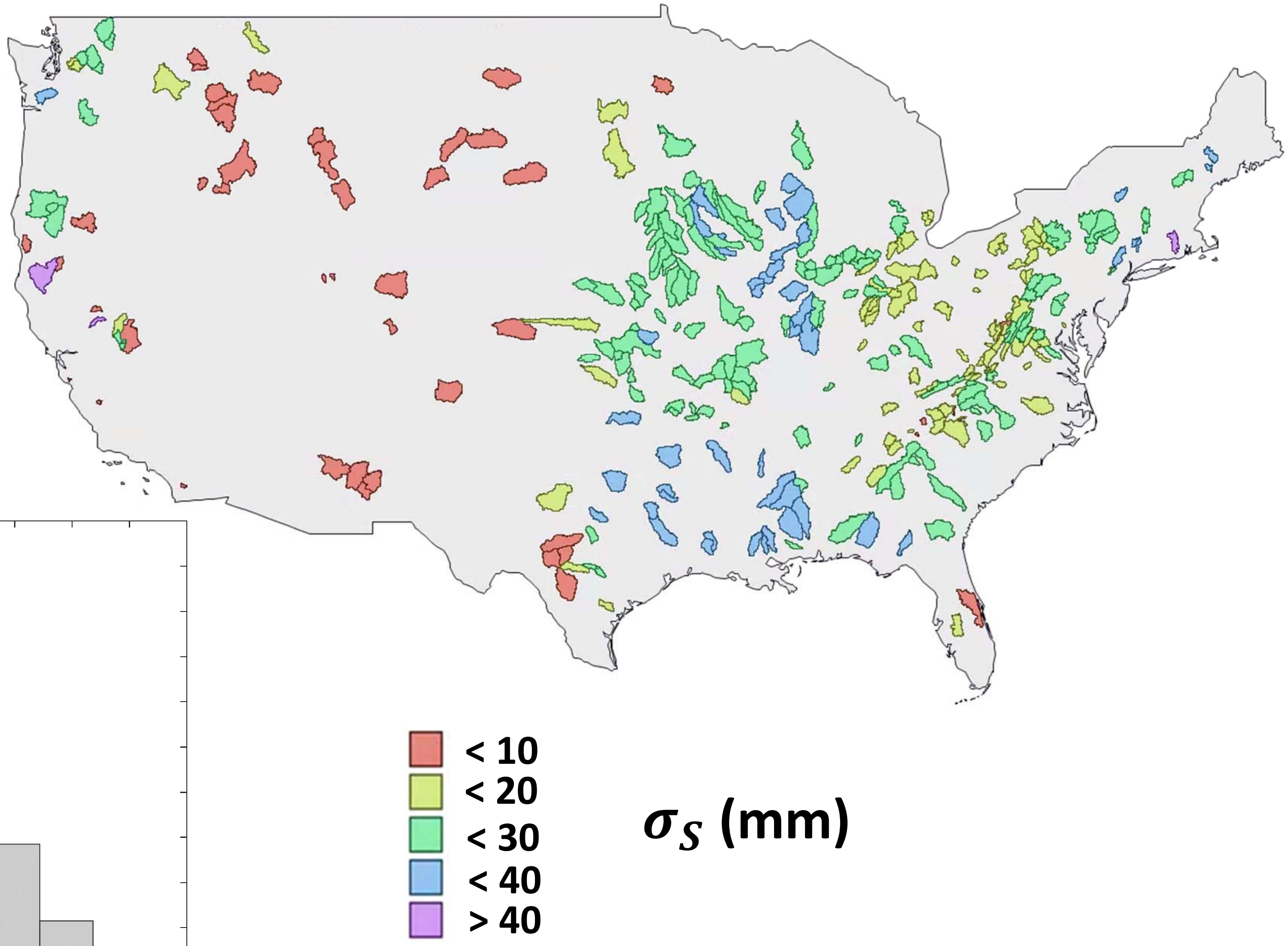

$f_S$

**all posteriors**

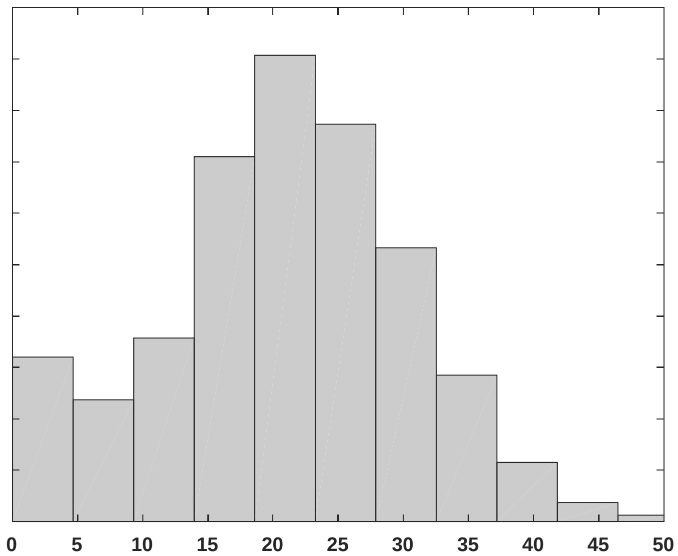


# **GRACE** absolute error: all basins

**posterior means**




**all posteriors**



# To be continued...

- Figure out spatial patterns
- More / better data
  - P, E: compare/combine multiple data products
  - S: use higher resolution GRACE data
- Better error models
  - Q: rating curve error analysis
  - E: non-Gaussian, seasonal bias/noise
  - temporal correlation?