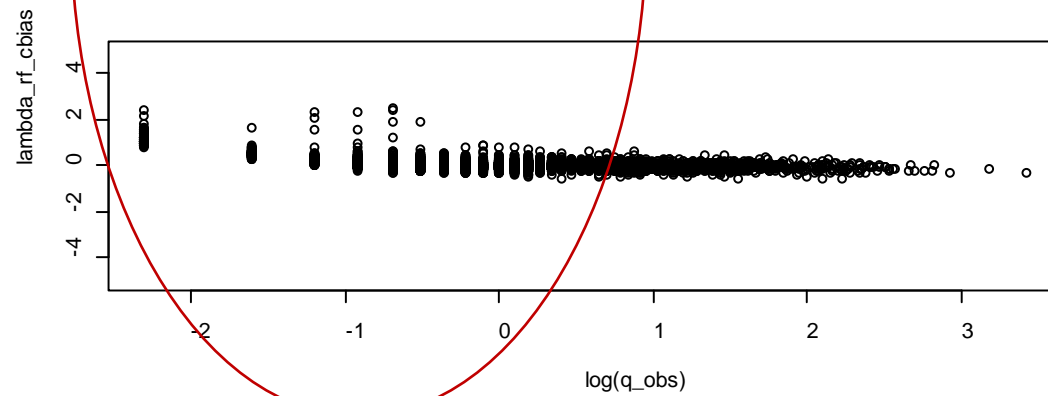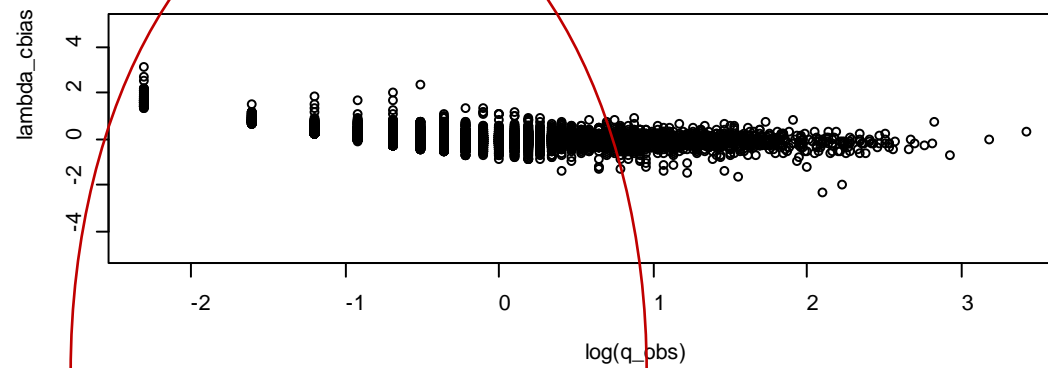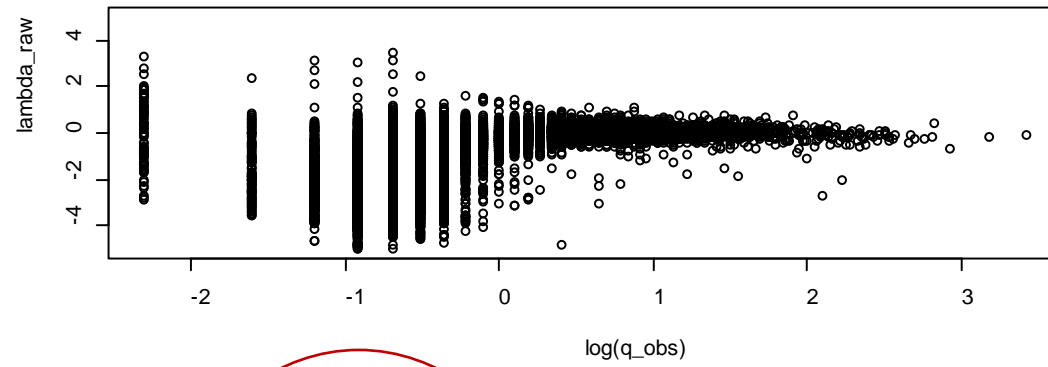# Conditonal bias estimation and logration SWM

Data/Methods

- New Shasta (SHA) from Sungwook in mm units (did not convert to cfs or kcfs)

- Tried out monthly fitted loess model ('cbias') and a random forest model ('rf_cbias') using Qsim and lag/lead Qsim covariates

- All conditional bias model attempt to predict a debiased estimate of Qobs -> P(Qobs|Qsim)

Both forms of conditional debiasing dampen bias and high variability of lambdas at low flow values

Histogram of lambda_raw
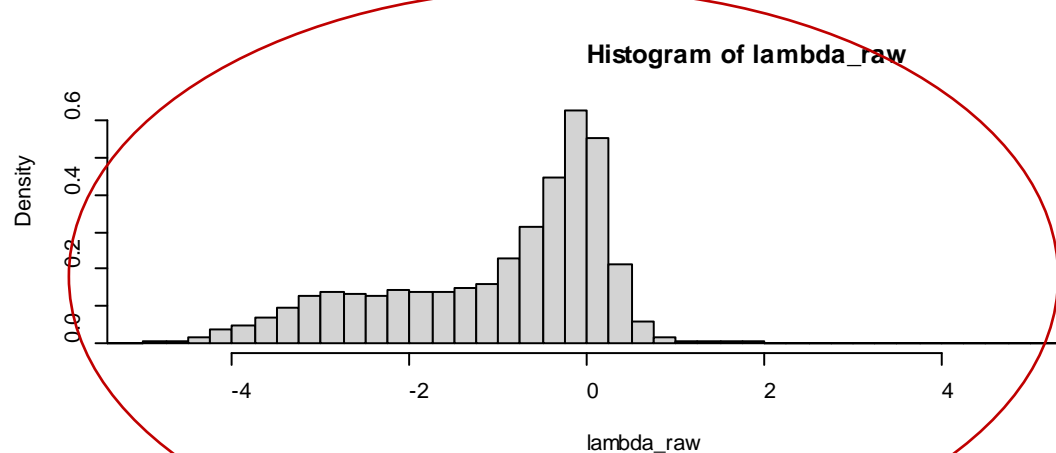
Raw lambdas very skewed; conditional debiasing seems to mostly correct this and substantially reduce variance
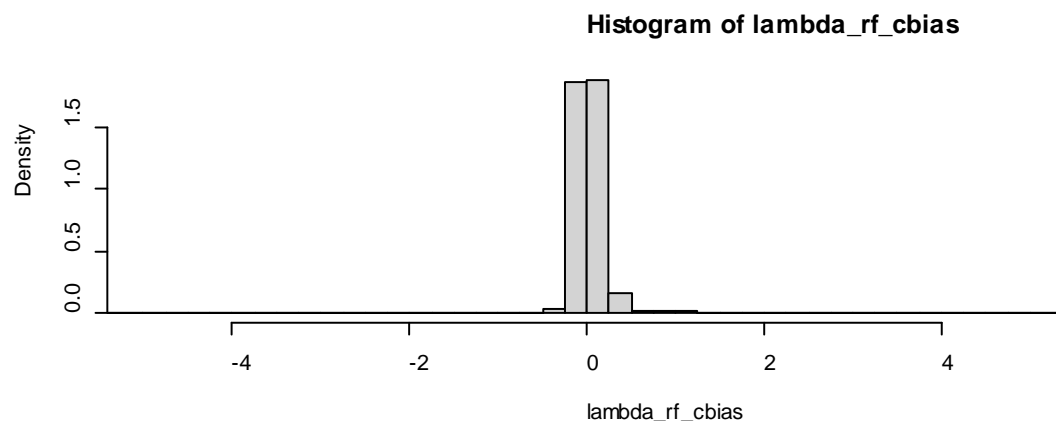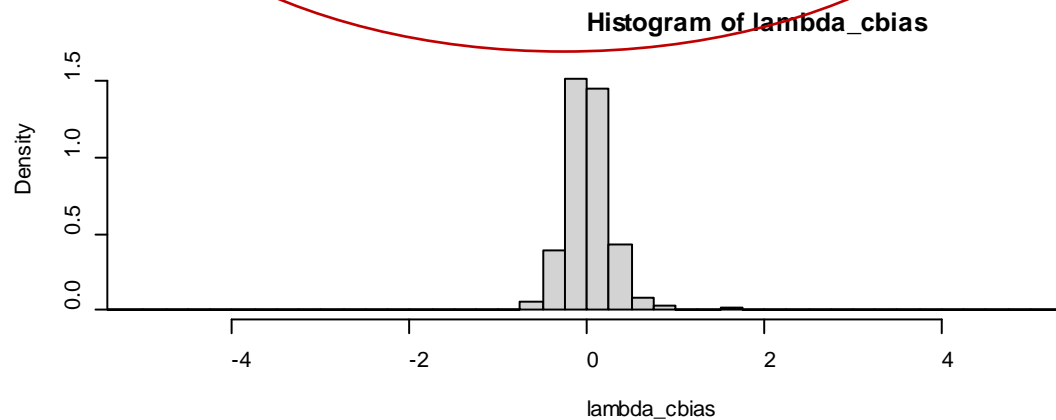
Closer inspection of two conditional debias methods shows relatively well behaved Gaussian distributions, though some positive skew remains

Histogram of lambda_cbias

Histogram of lambda_rf_cbias
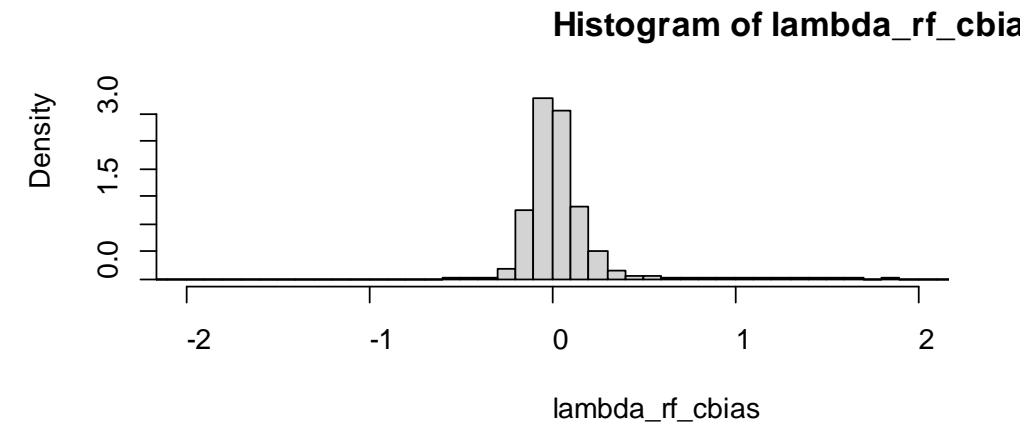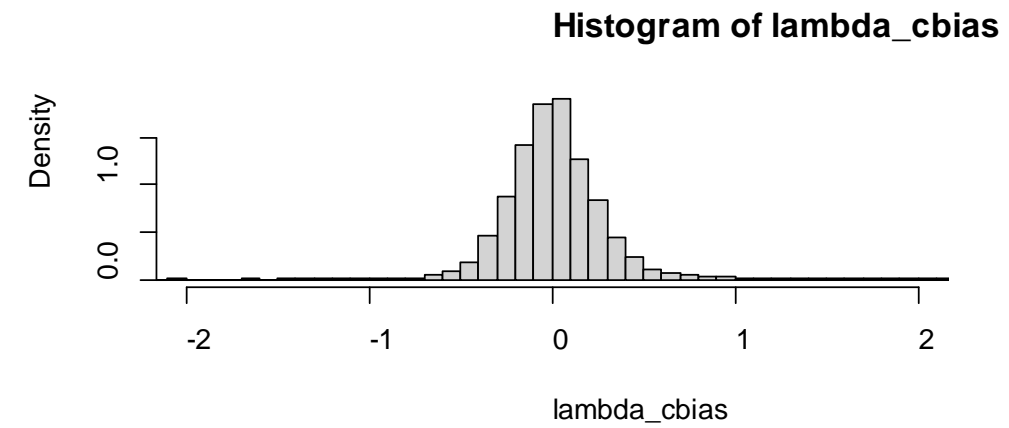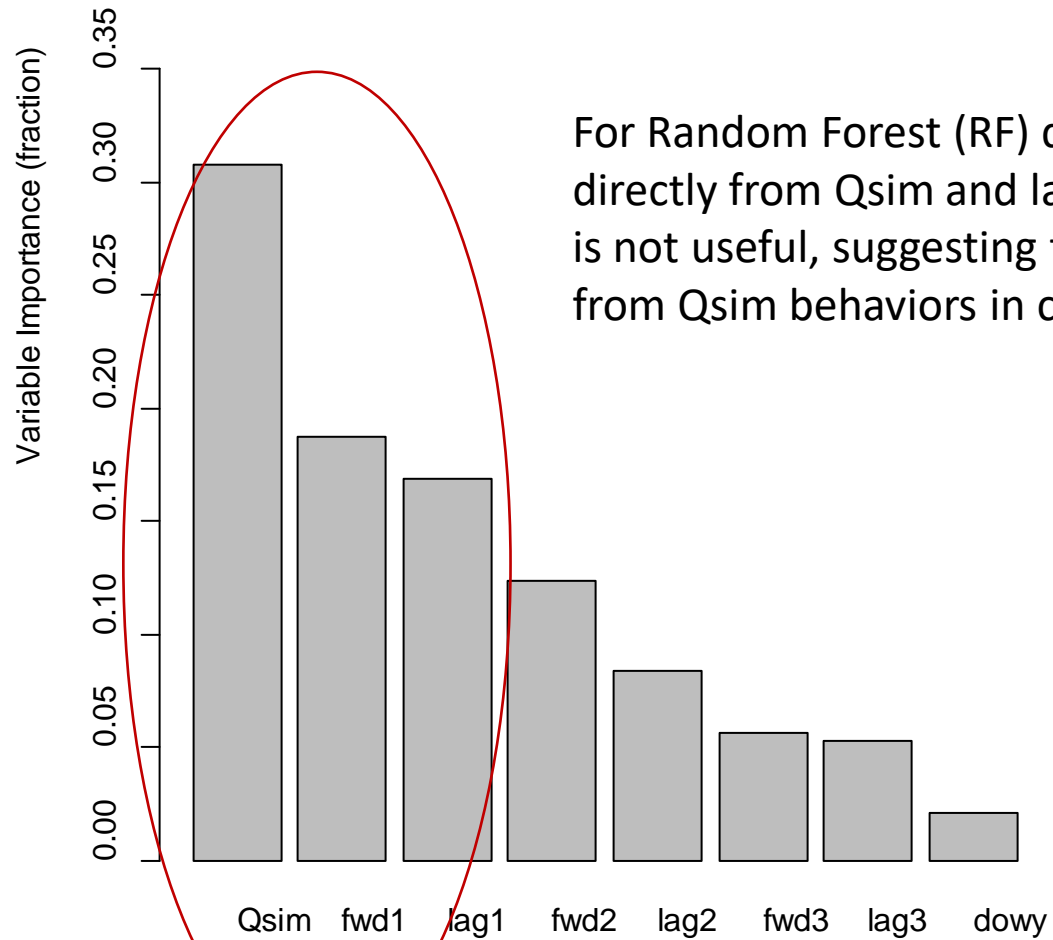
For Random Forest (RF) debiasing method, most information comes directly from Qsim and lag -1/+1 covariates. Day-of-water-year (dowy) is not useful, suggesting that the model gets most of what it needs from Qsim behaviors in close temporal proximity to predictand (Qobs)

Brief SWM output analysis

- SWM simulations using raw (uncorrected) log-ratio method can produce unrealistically large simulation values (maximum flows more than 10X what has been observed). This appears to be largely fixed by applying the conditional bias correction, which also helps the central tendency estimation

- Both the LOESS conditionally debiasing and the RF model may be overfit to calibration data with no cal-val-test subsetting. This would be an important consideration for truly understanding out-of-sample uncertainty quantification

```
> #compare maximum values
> max(q_obs)
[1] 30.6
> max(q_swm_raw) #very large, can be > 10X q_obs
[1] 985.0616
> max(q_swm_cbias)
[1] 34.37096
> max(q_swm_rf_cbias)
[1] 22.16145
`
```

```
> #mean values
> mean(q_obs)
[1] 1.081639
> mean(q_swm_raw) #well above mean of q_obs, driven high by high maximum values
[1] 6.34467
> mean(q_swm_cbias) #pretty close
[1] 1.062816
> mean(q_swm_rf_cbias) #pretty close but low
[1] 1.066988
`
```

```
> #median values
> median(q_obs)
[1] 0.7
> median(q_swm_raw) #median is not bad
[1] 0.8205531
> median(q_swm_cbias) #pretty close
[1] 0.6597482
> median(q_swm_rf_cbias) #pretty close but lower
[1] 0.6513245
```