QUANTILE MAPPING

ERT 474/574

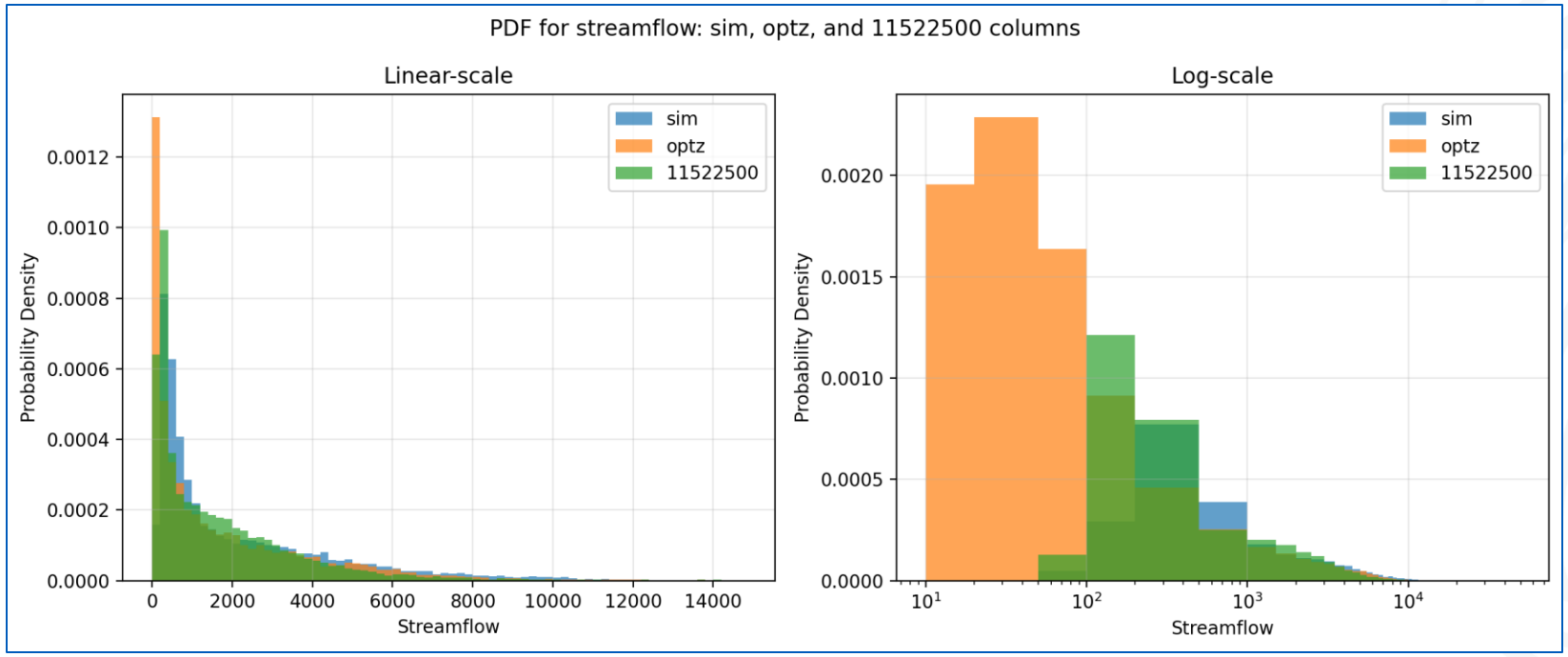
Open-Source Hydro Data Analytics

Oct 1st 2025

University at Buffalo The State University of New York

Announcements

- **Homework #3**
  - Due date: Wednesday (Oct 1)



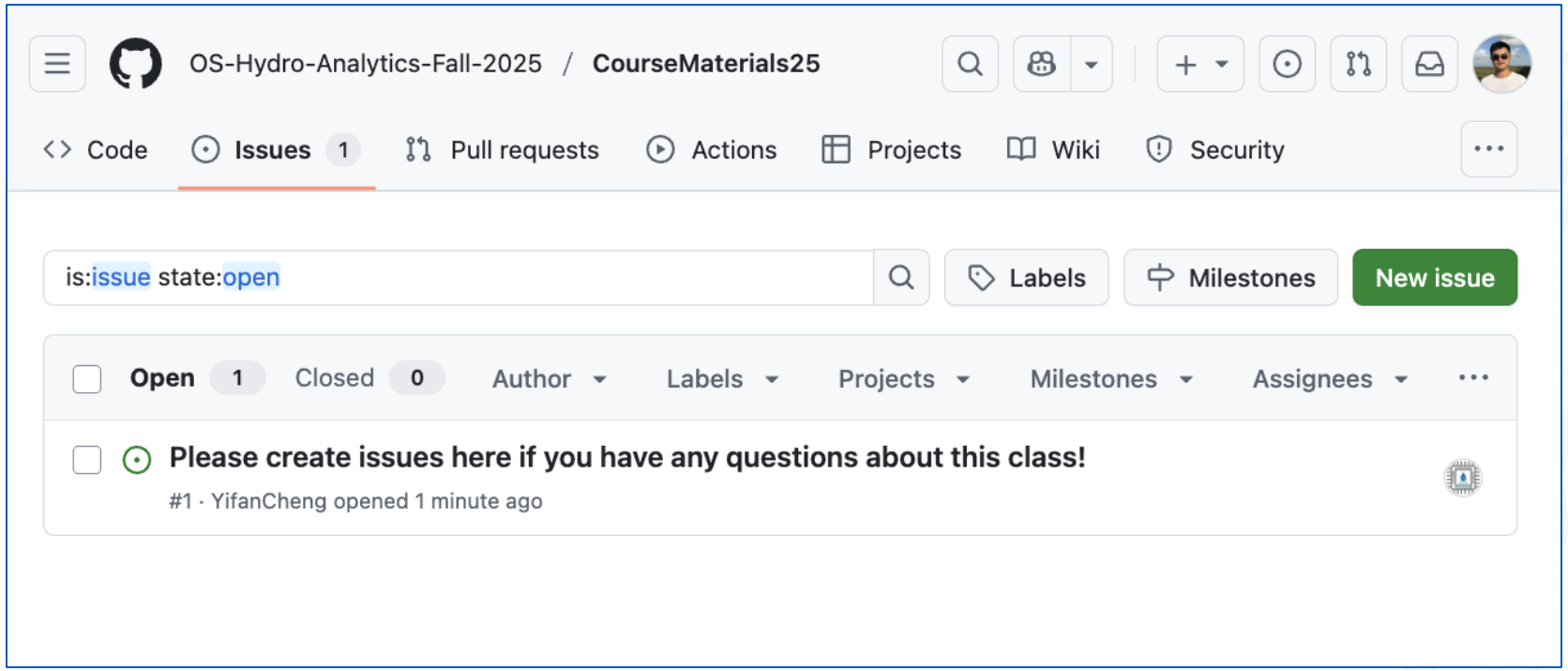
Someone asked me:

How can we visualize the three histograms in the same figure?

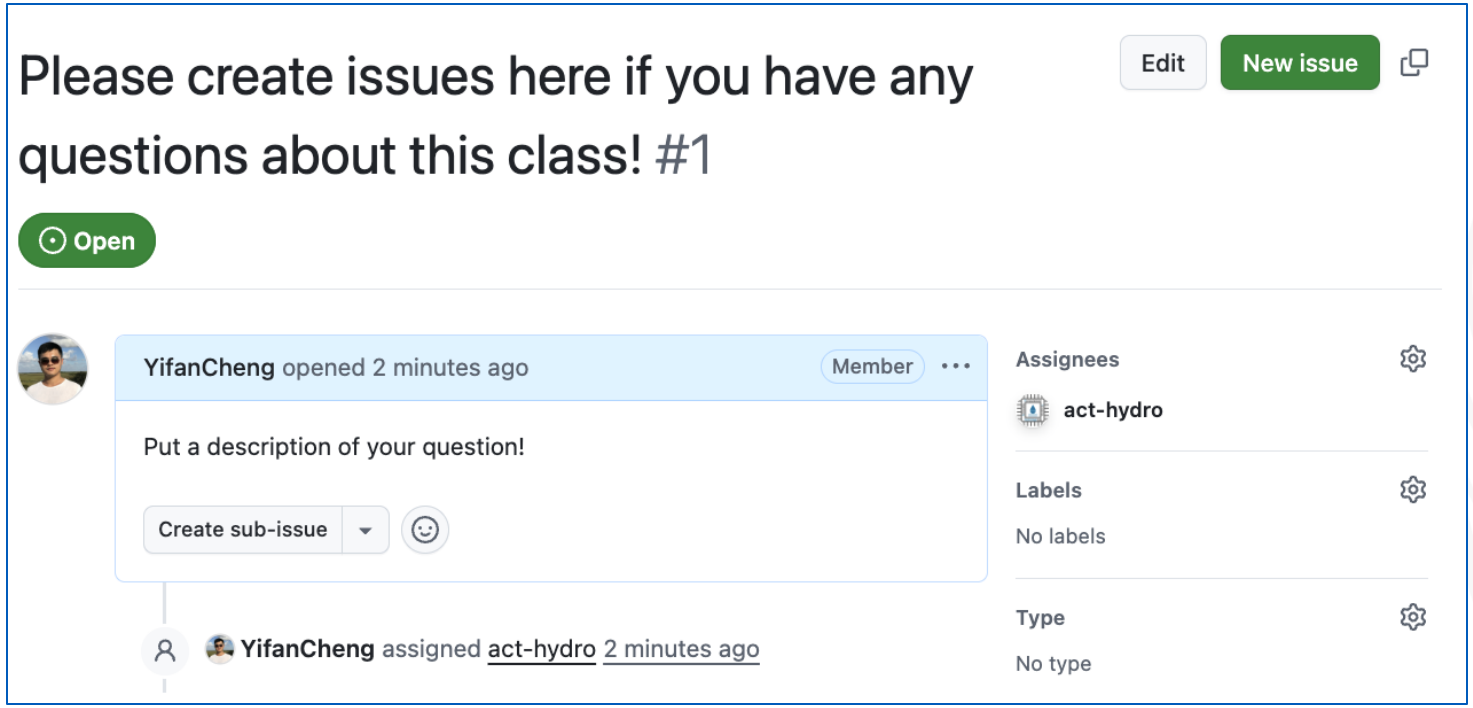
The solution is to add the transparency.

```
plt.hist(…., alpha=0.5)
```

Post questions on Issue page



Please assign this issue to **act-hydro**

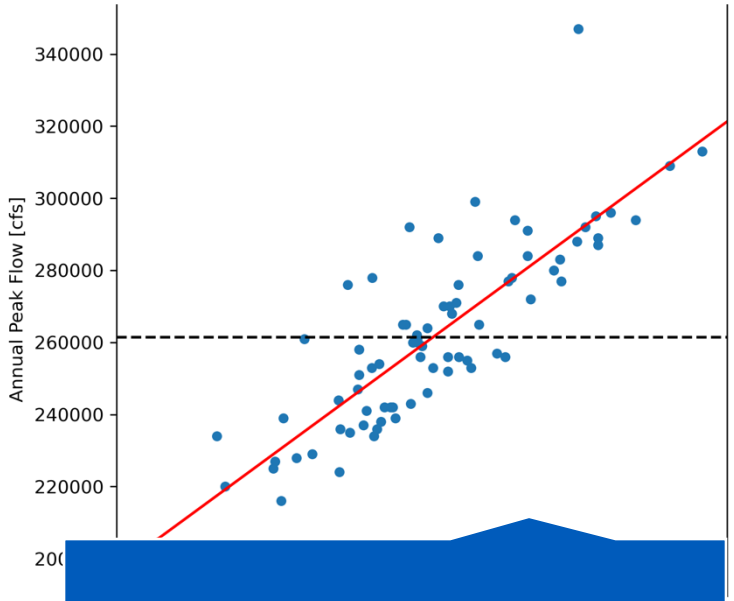


Then I will receive a notification and will get back to your issue ASAP!

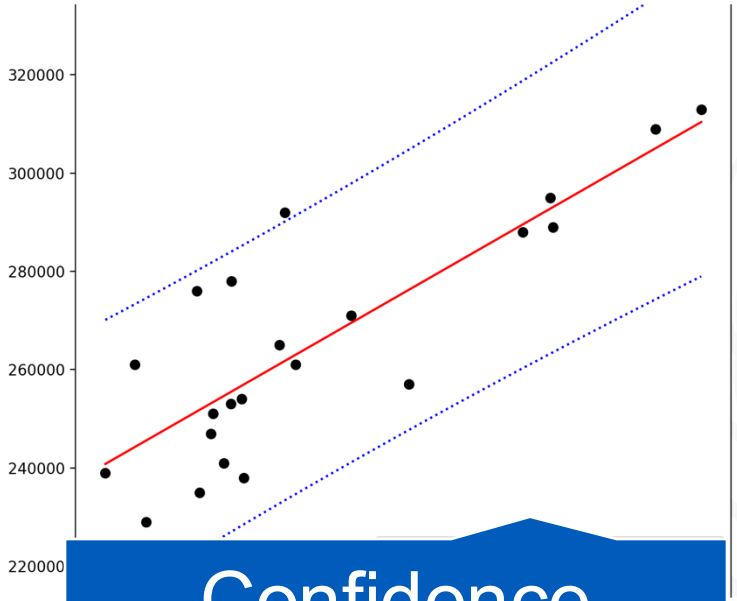
This way, others who might share similar questions can see our discussion!

Please don't hold questions to yourself and guess what I mean. If you cannot figure something out, please feel free to reach out to me!

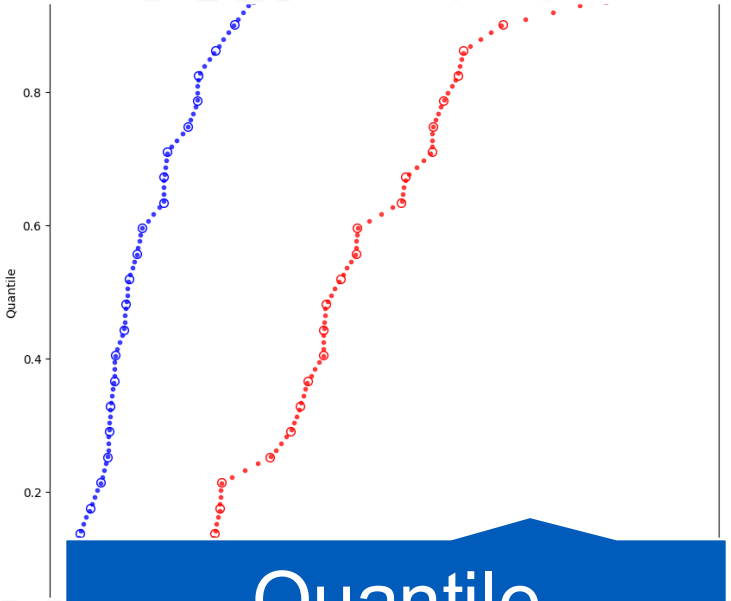
Trend analysis



Linear regression



Confidence intervals



Quantile regression

Quantile Regression

Quantile Regression

Advantage of Quantile regression

- Do not require that the underlying probability distributions are known or have any particular form.
- A linear relationship between the two variables is not required.
- The time series of the data need not be the same (or even from the same times) in the explanatory and dependent variables. That is, paired data is not required (although in many cases it is desirable).

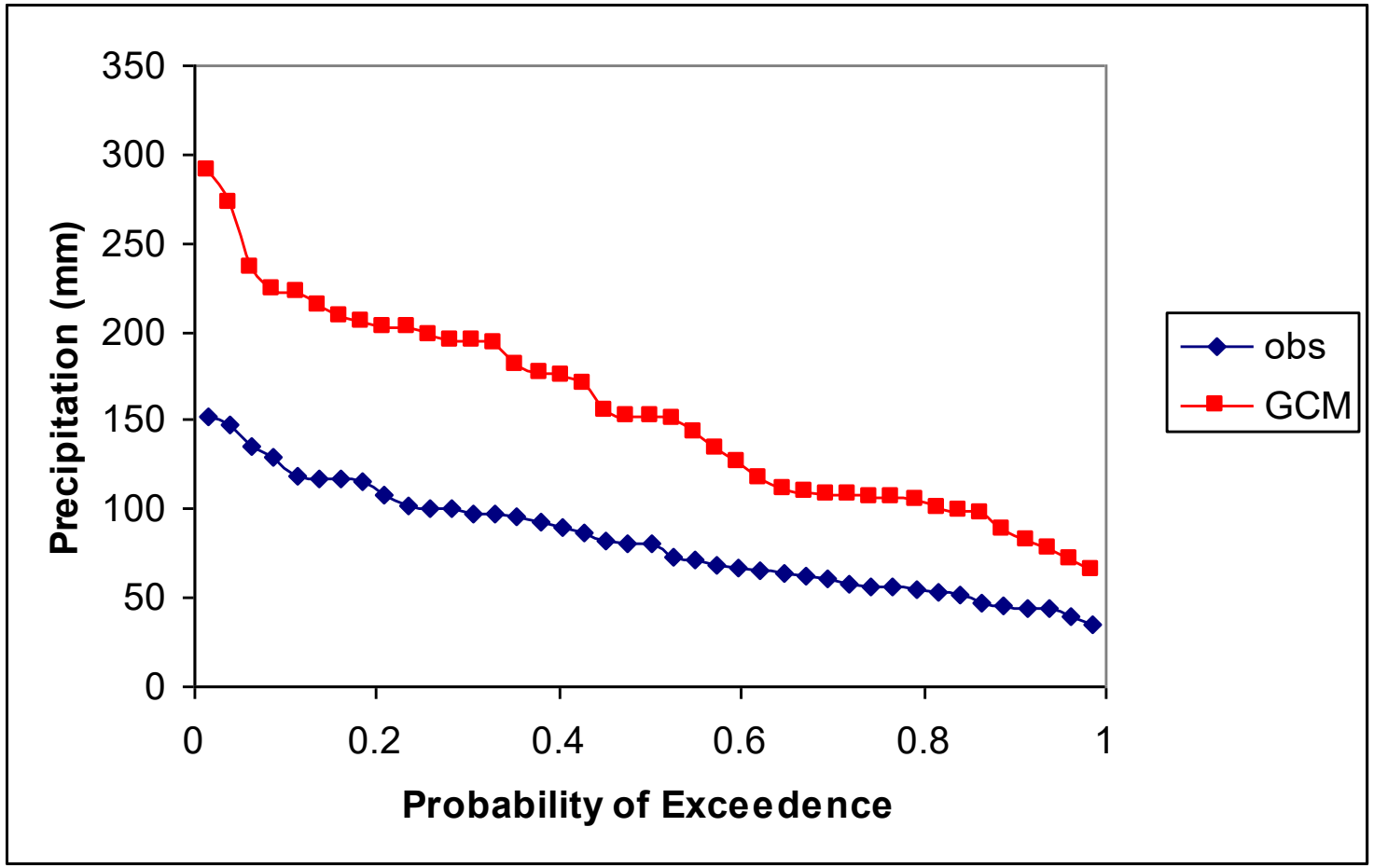
Example

Bias-correct simulated precipitation from Global Climate Model (GCM) using local observations

How does Quantile Regression work?

Step 1: For each of your two datasets, create an empirical CDF

We presume that relative ranking and frequencies of events are correct, even if actual values don't match up in a linear way.



Example

GCM Input = 190

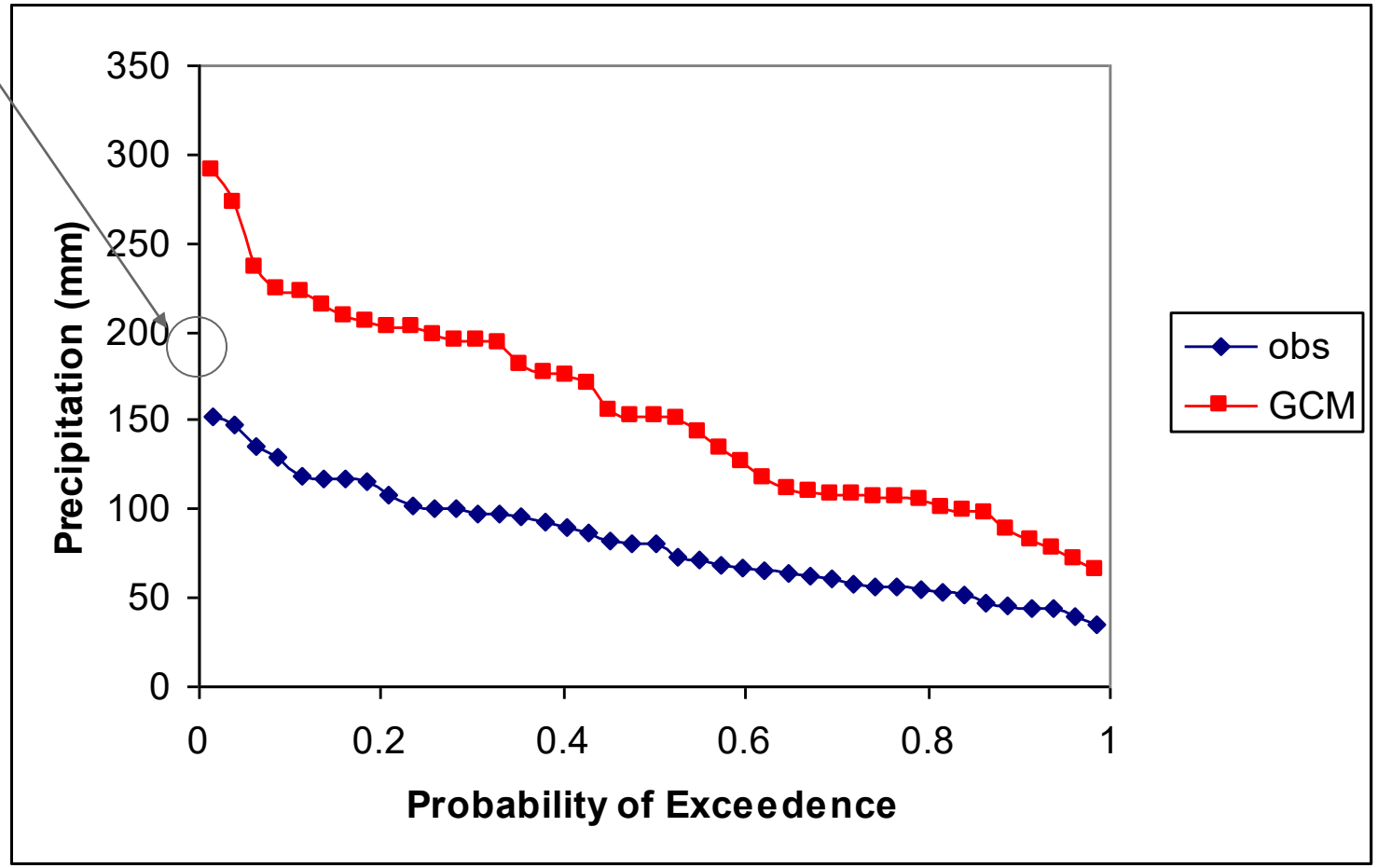
We presume that relative ranking and frequencies of events are correct, even if actual values don't match up in a linear way.

Bias-correct simulated precipitation from Global Climate Model (GCM) using local observations

How does Quantile Regression work?

Step 1: For each of your two datasets, create an empirical CDF

Step 2: Use the two empirical CDFs as a way of looking-up (or mapping) values from the predictor to the predictand, by matching which physical value corresponds to the same quantile



Example

GCM Input = 190

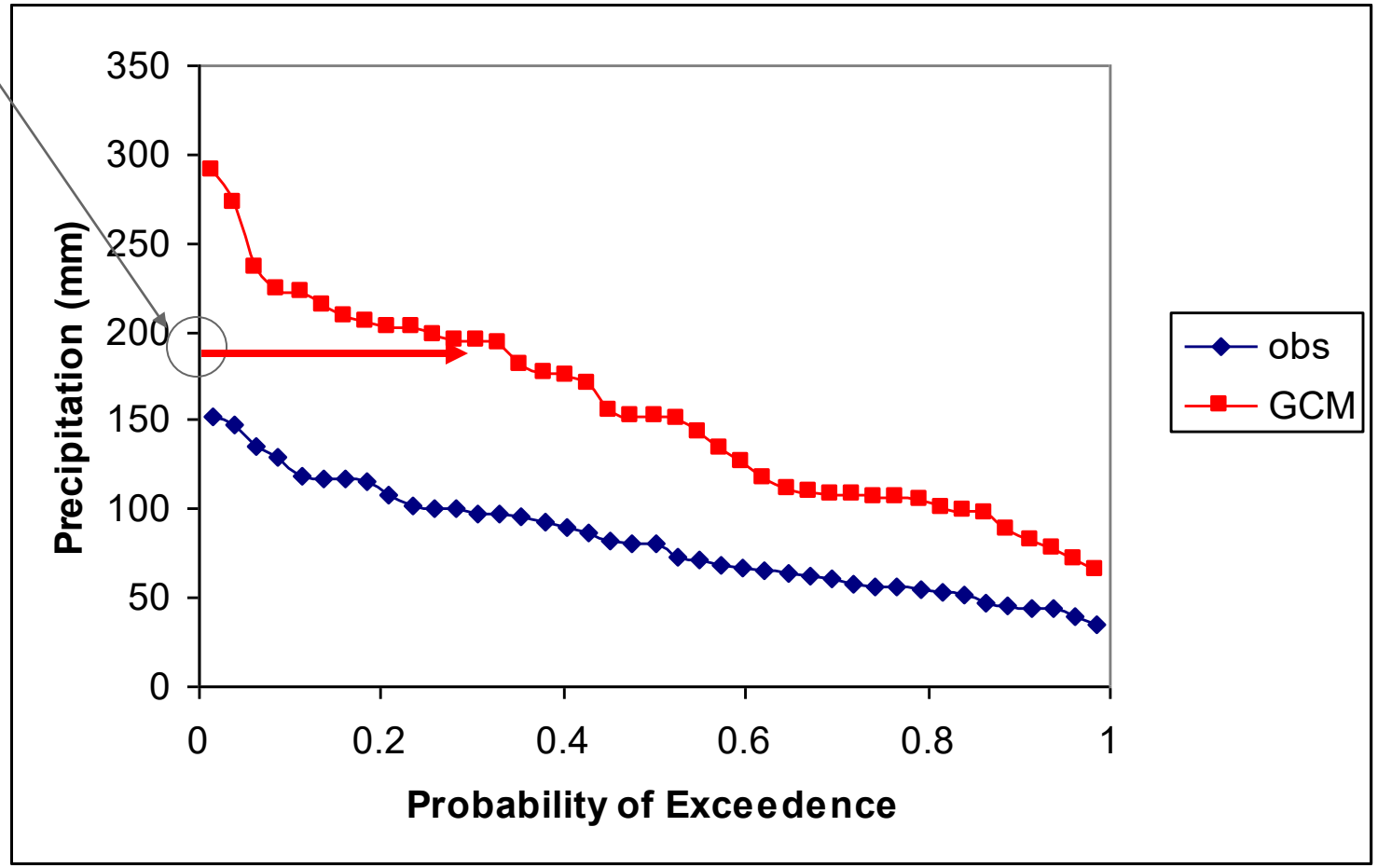
We presume that relative ranking and frequencies of events are correct, even if actual values don't match up in a linear way.

Bias-correct simulated precipitation from Global Climate Model (GCM) using local observations

How does Quantile Regression work?

Step 1: For each of your two datasets, create an empirical CDF

Step 2: Use the two empirical CDFs as a way of looking-up (or mapping) values from the predictor to the predictand, by matching which physical value corresponds to the same quantile



Example

GCM Input = 190

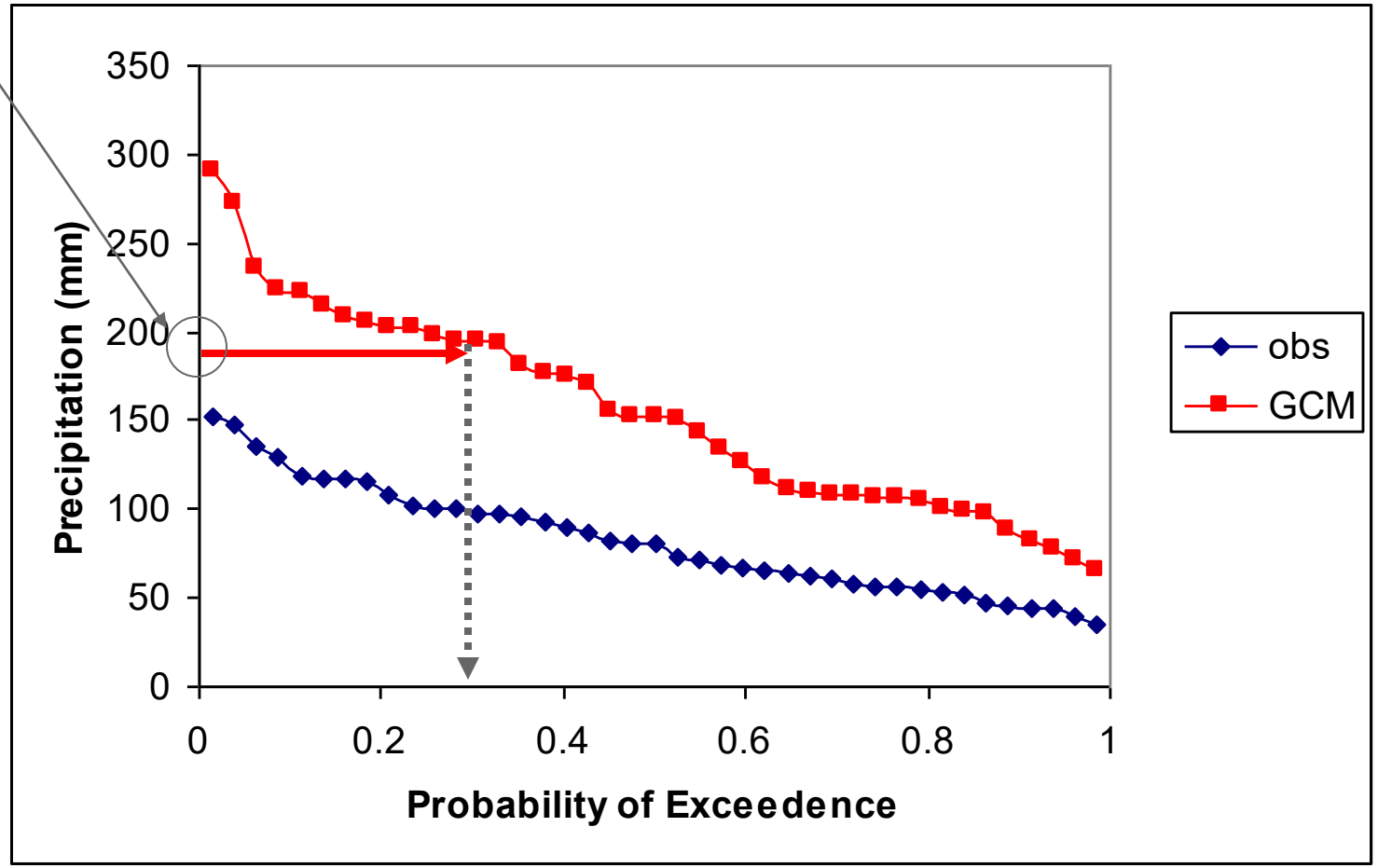
We presume that relative ranking and frequencies of events are correct, even if actual values don't match up in a linear way.

Bias-correct simulated precipitation from Global Climate Model (GCM) using local observations

How does Quantile Regression work?

Step 1: For each of your two datasets, create an empirical CDF

Step 2: Use the two empirical CDFs as a way of looking-up (or mapping) values from the predictor to the predictand, by matching which physical value corresponds to the same quantile



Example

GCM Input = 190

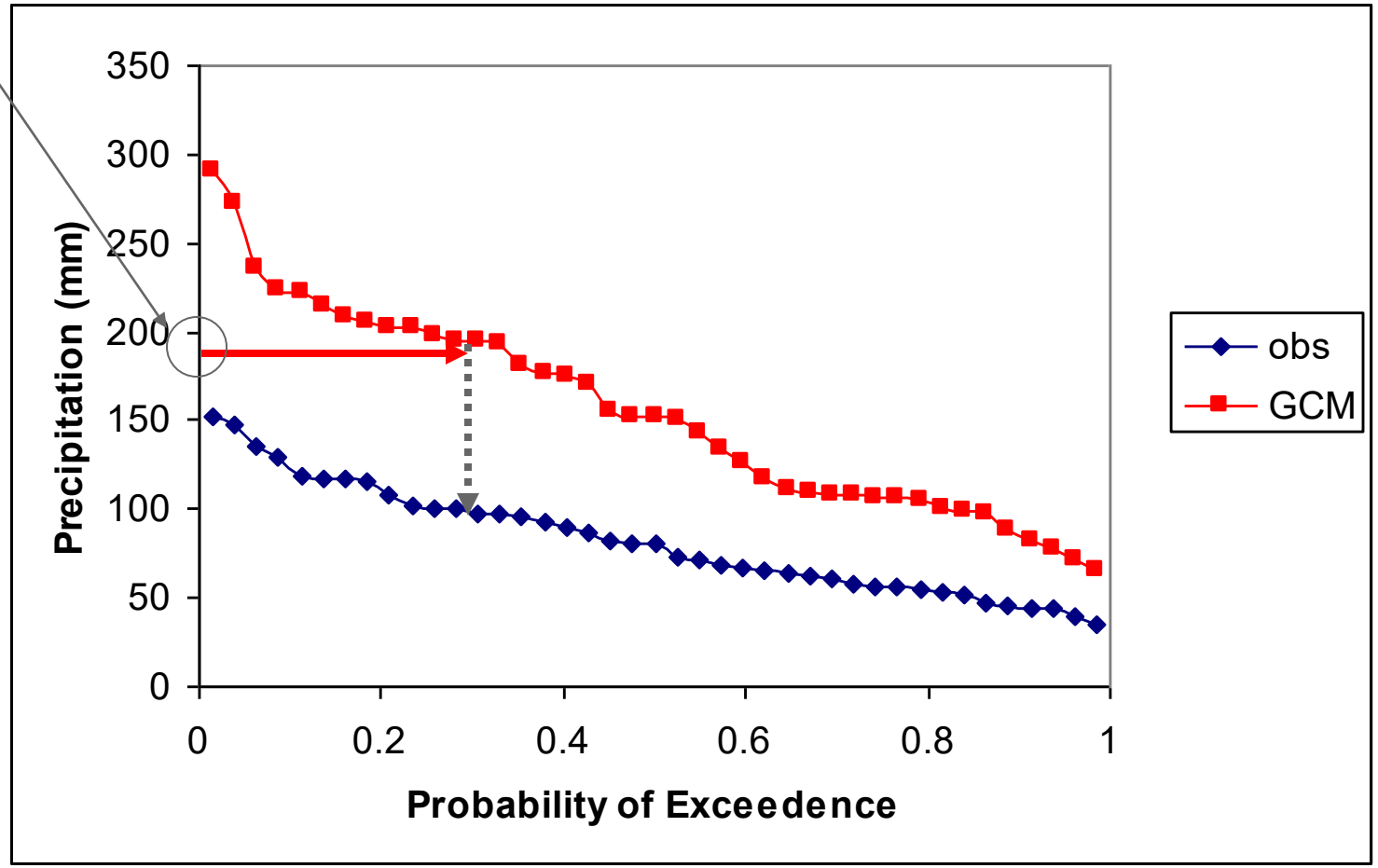
We presume that relative ranking and frequencies of events are correct, even if actual values don't match up in a linear way.

Bias-correct simulated precipitation from Global Climate Model (GCM) using local observations

How does Quantile Regression work?

Step 1: For each of your two datasets, create an empirical CDF

Step 2: Use the two empirical CDFs as a way of looking-up (or mapping) values from the predictor to the predictand, by matching which physical value corresponds to the same quantile



Example

GCM Input = 190

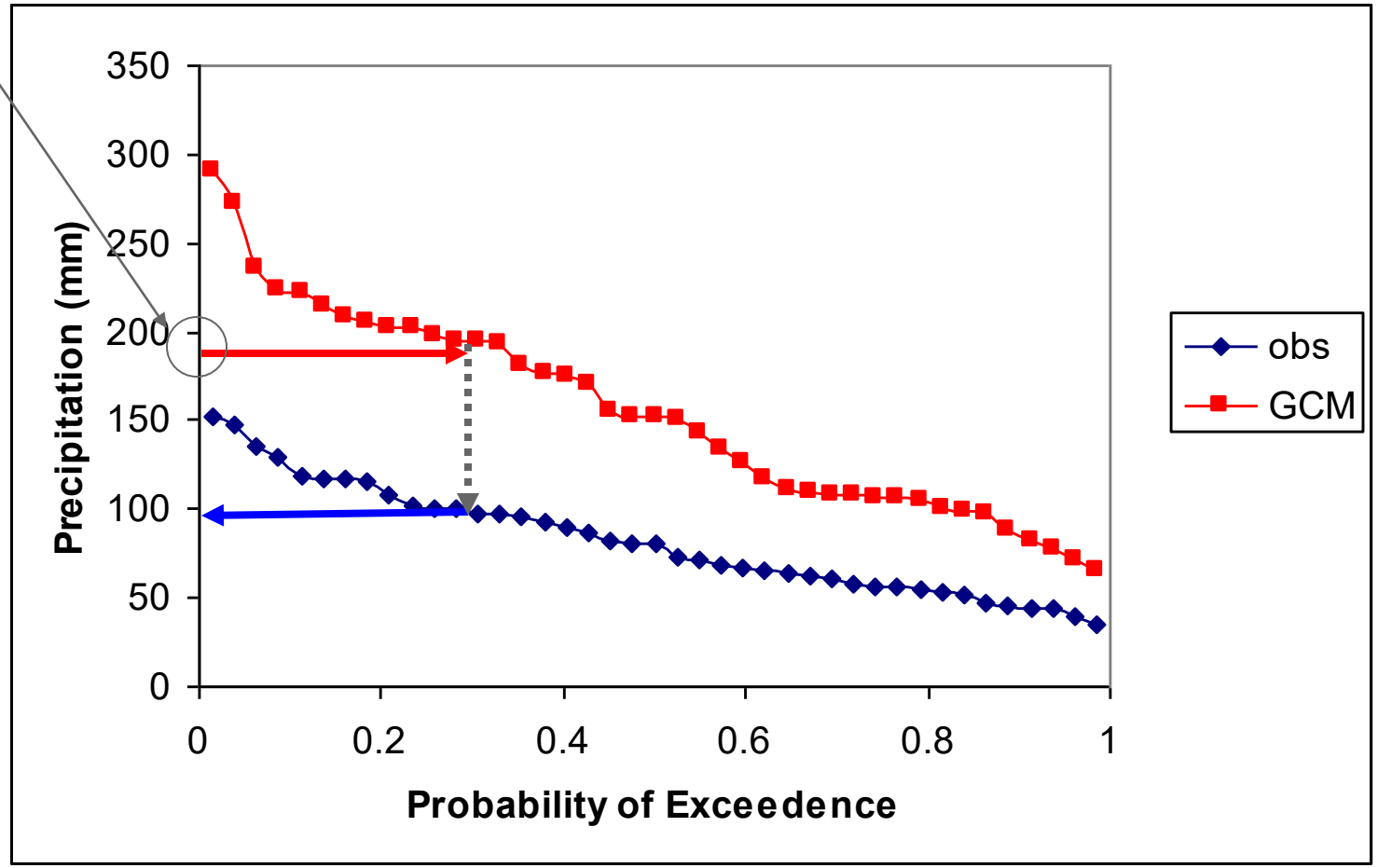
We presume that relative ranking and frequencies of events are correct, even if actual values don't match up in a linear way.

Bias-correct simulated precipitation from Global Climate Model (GCM) using local observations

How does Quantile Regression work?

Step 1: For each of your two datasets, create an empirical CDF

Step 2: Use the two empirical CDFs as a way of looking-up (or mapping) values from the predictor to the predictand, by matching which physical value corresponds to the same quantile



Example

GCM Input = 190

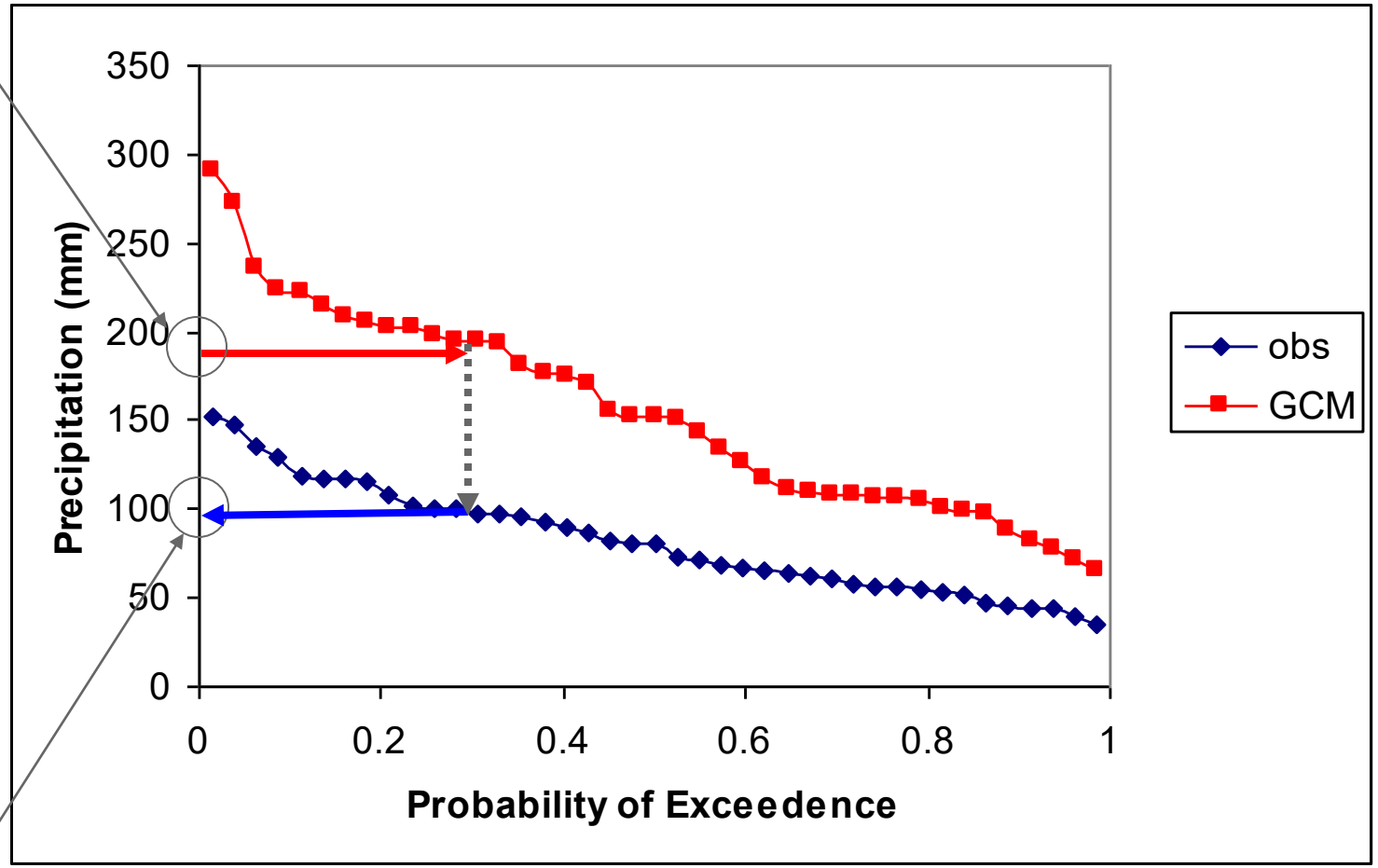
We presume that relative ranking and frequencies of events are correct, even if actual values don't match up in a linear way.

Bias-correct simulated precipitation from Global Climate Model (GCM) using local observations

How does Quantile Regression work?

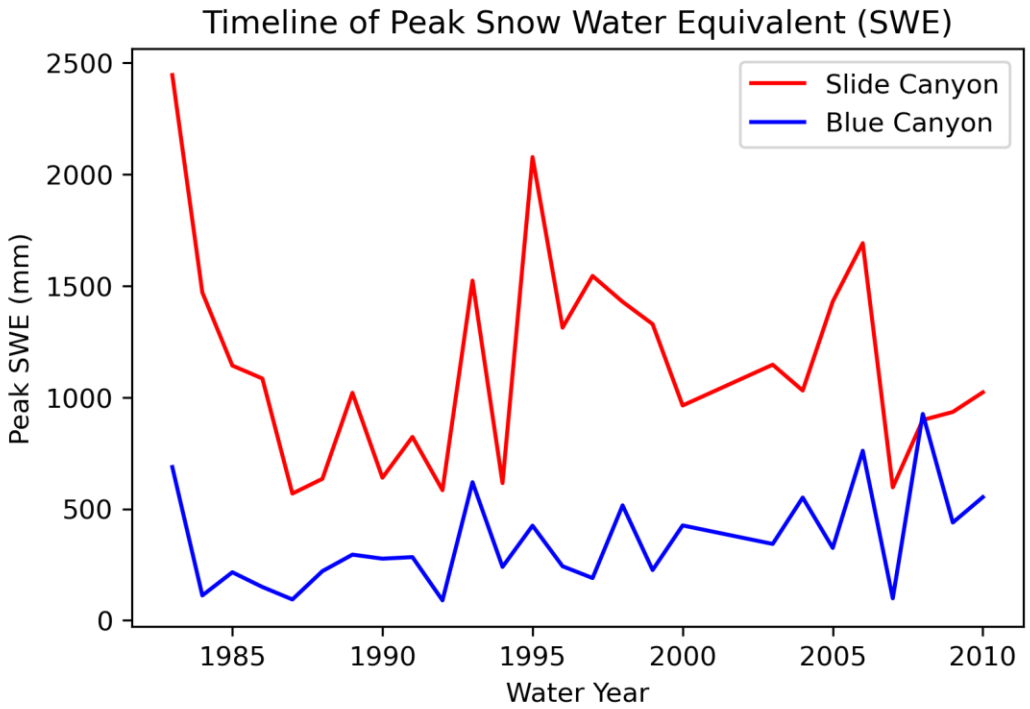
Step 1: For each of your two datasets, create an empirical CDF

Step 2: Use the two empirical CDFs as a way of looking-up (or mapping) values from the predictor to the predictand, by matching which physical value corresponds to the same quantile

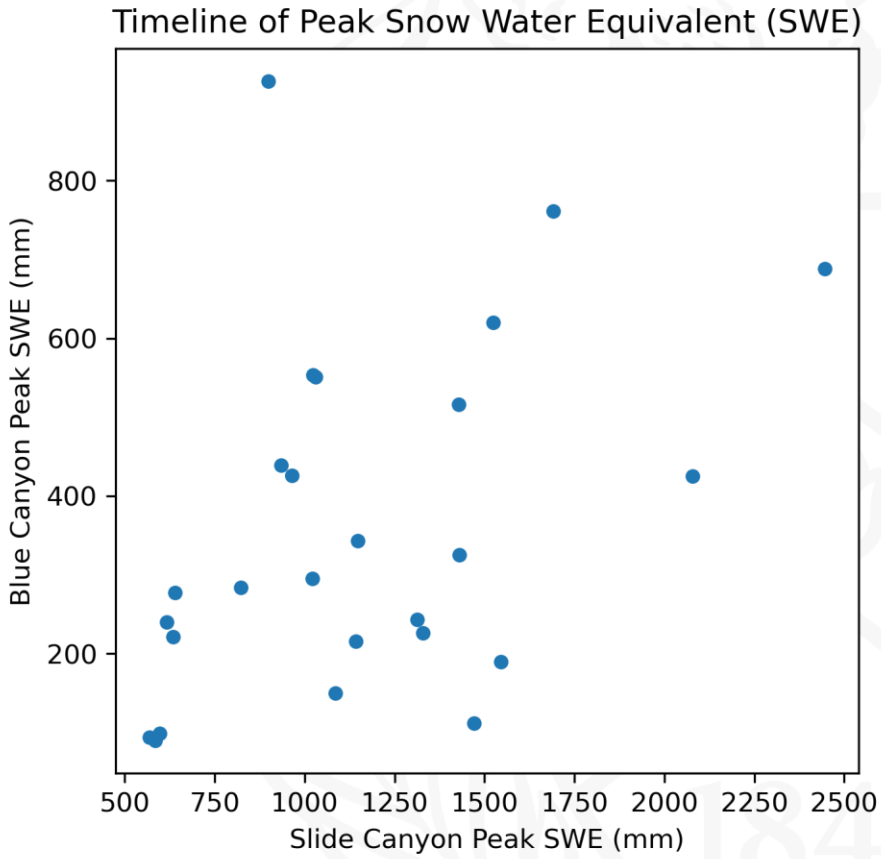


Bias Corrected Output = 100

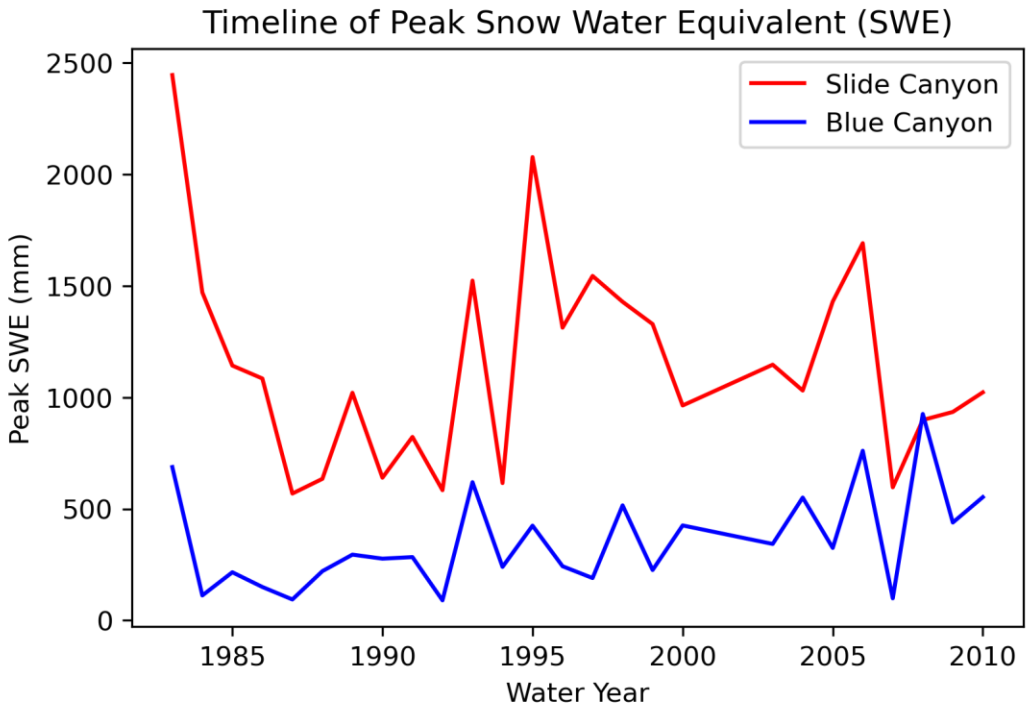
Example: Use SWE in Slide Canyon to predict SWE in Blue Canyon



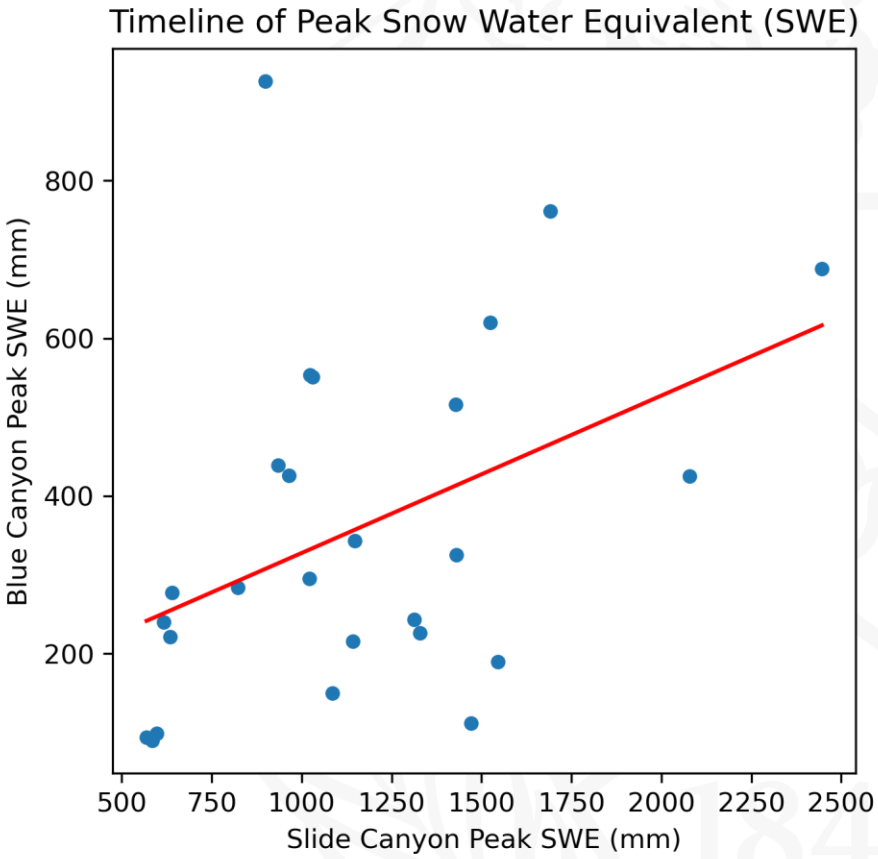
We can use the Linear Regression Model!



Example: Use SWE in Slide Canyon to predict SWE in Blue Canyon



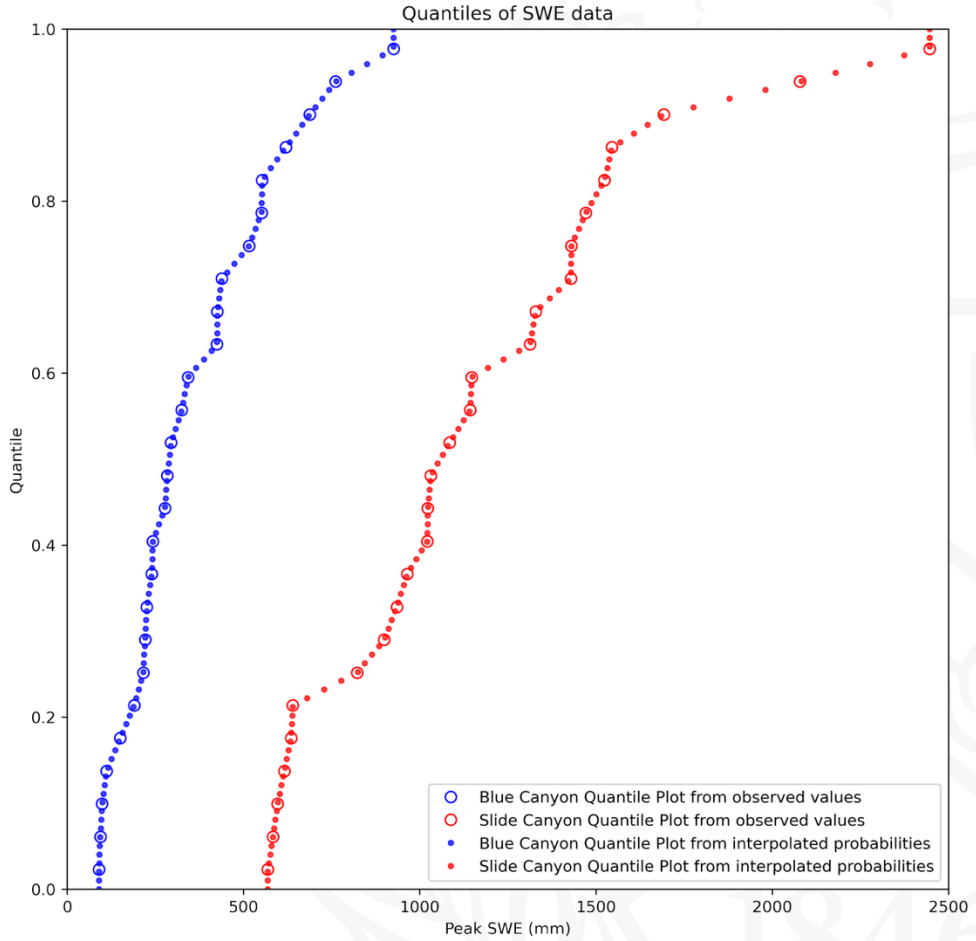
We can use the Linear Regression Model!



Example: Use SWE in Slide Canyon to predict SWE in Blue Canyon

How does Quantile Regression work?

Step 1: For each of your two datasets, create an empirical CDF

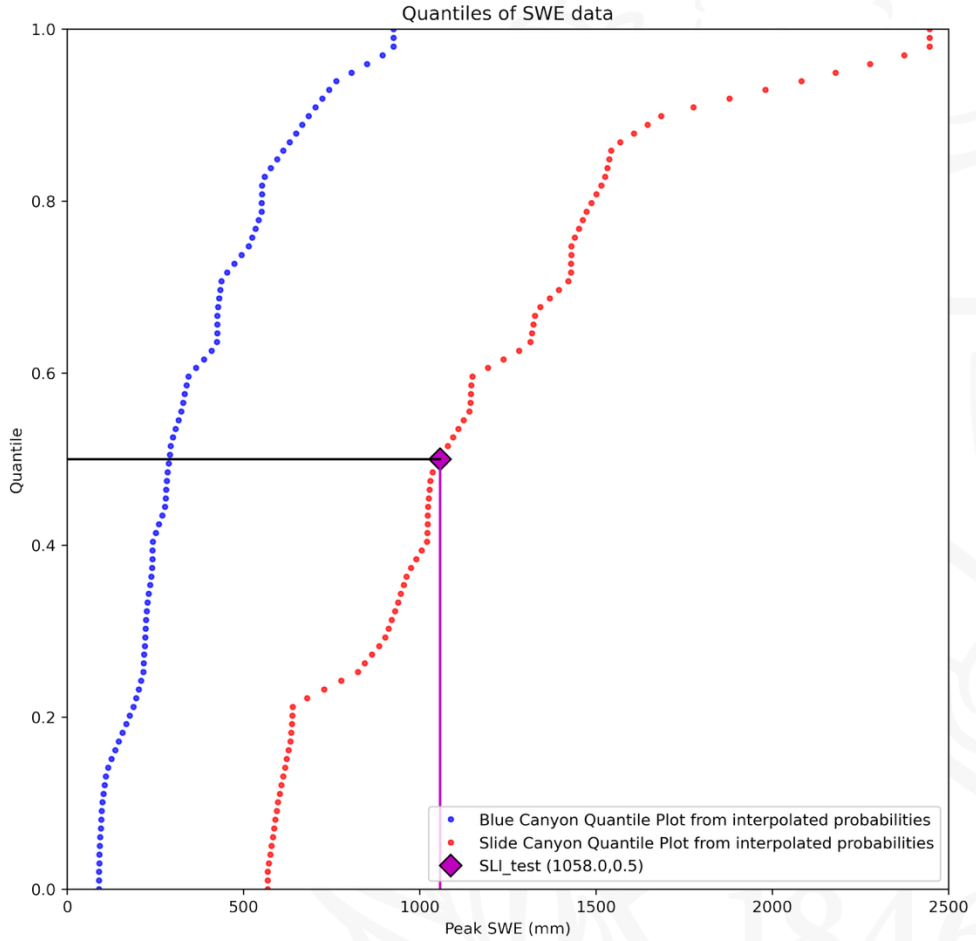


Example: Use SWE in Slide Canyon to predict SWE in Blue Canyon

How does Quantile Regression work?

Step 1: For each of your two datasets, create an empirical CDF

Step 2: Use the two empirical CDFs as a way of looking-up (or mapping) values from the predictor to the predictand, by matching which physical value corresponds to the same quantile

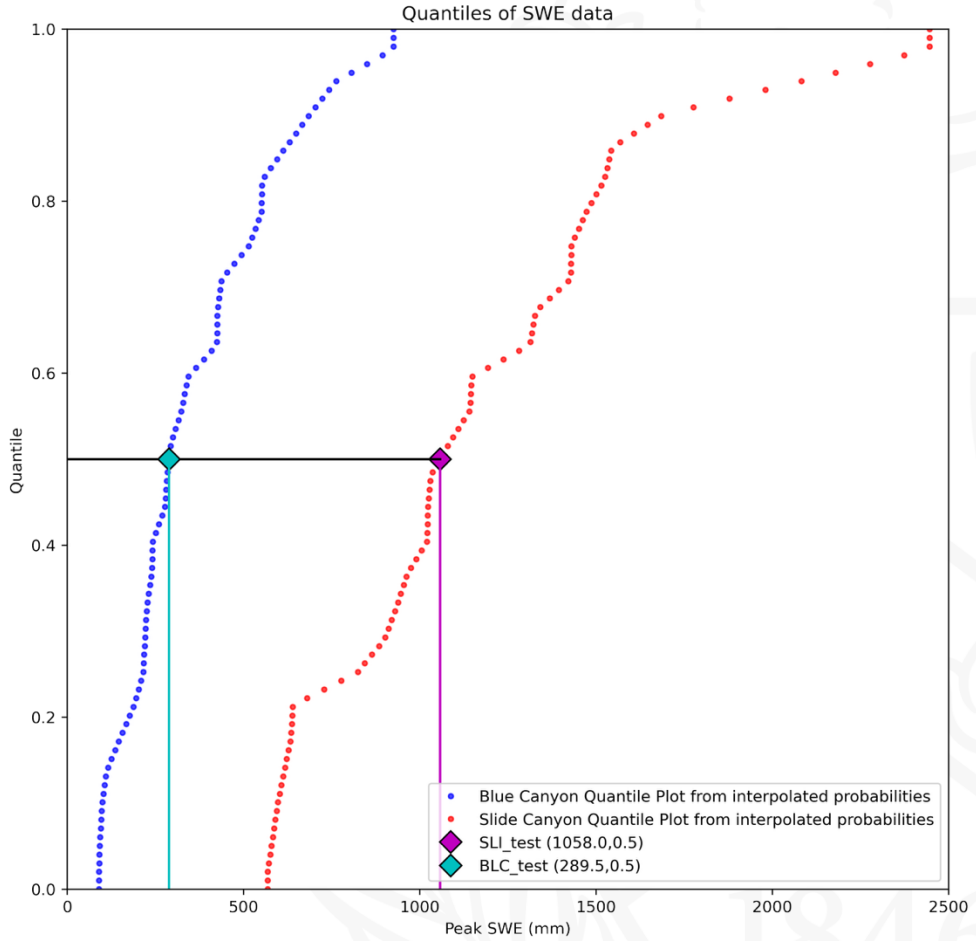


Example: Use SWE in Slide Canyon to predict SWE in Blue Canyon

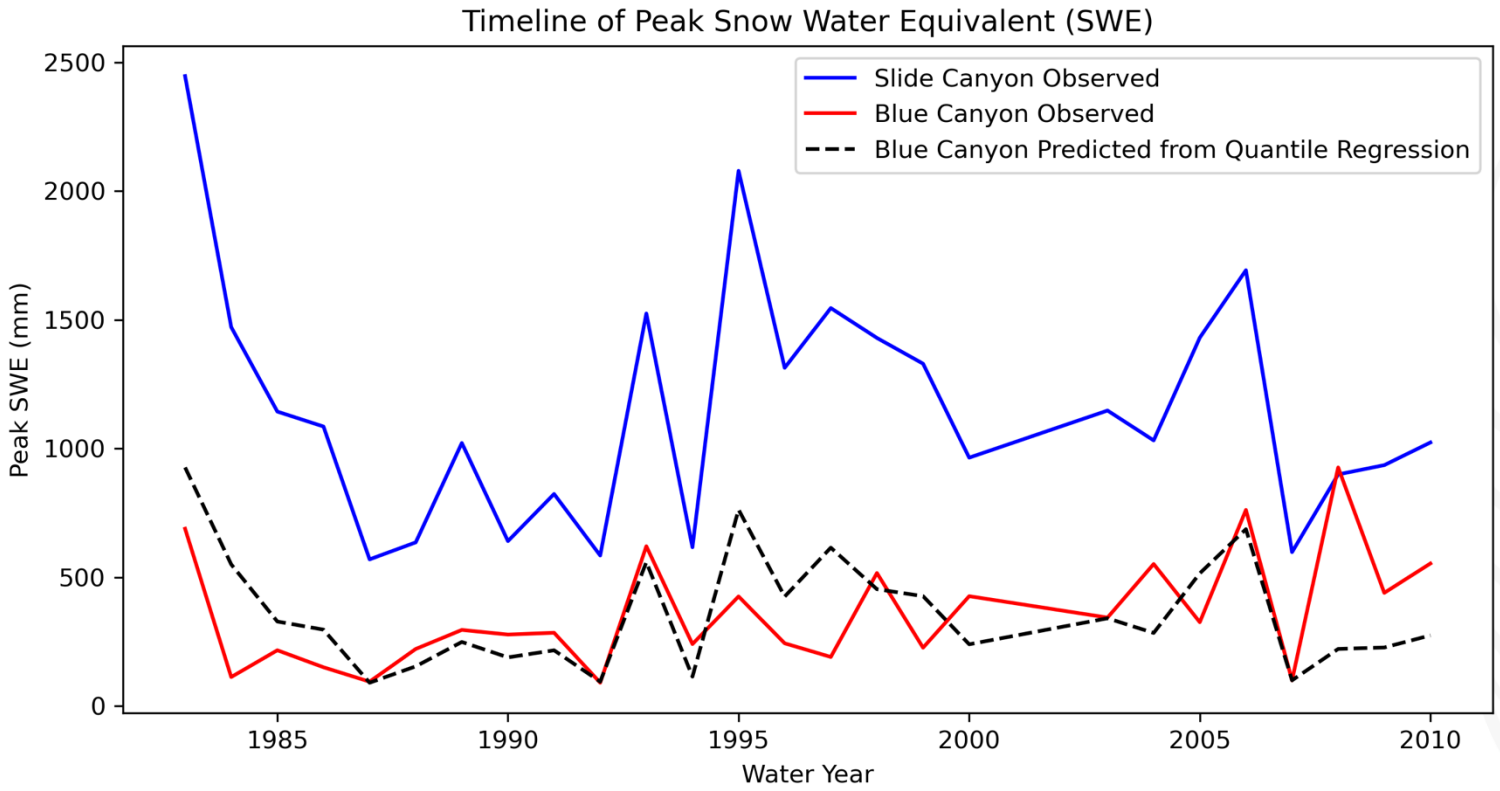
How does Quantile Regression work?

Step 1: For each of your two datasets, create an empirical CDF

Step 2: Use the two empirical CDFs as a way of looking-up (or mapping) values from the predictor to the predictand, by matching which physical value corresponds to the same quantile



Example: Use SWE in Slide Canyon to predict SWE in Blue Canyon



Minute paper time

- Today's class marks the end of basic statistics.
- Over the past month, we learned
  - Intro into Python (Python literacy, Numpy & Matplotlib)
  - Basic Statistics (Pandas)
  - Extreme events (100-year flood, 7Q10 low flow events)
  - Hypothesis testing (Significant changes, Z-test, Student test)
  - Model evaluation (Nash-Sutcliffe Efficiency, Kling-Gupta Efficiency)
  - Trend Analysis (Linear regression, confidence intervals, quantile mapping)

Please consider:

1) What topics are you mostly interested in? Maybe most useful to your research?
2) In which topic do you want to learn a bit more in-depth?
3) Currently, our classes involve lectures, labs, and homework. For each learning method, what do you like about it? Do you see any opportunities for improvement?

I will try my best to incorporate your comments towards the latter part of this class.