ROC Station

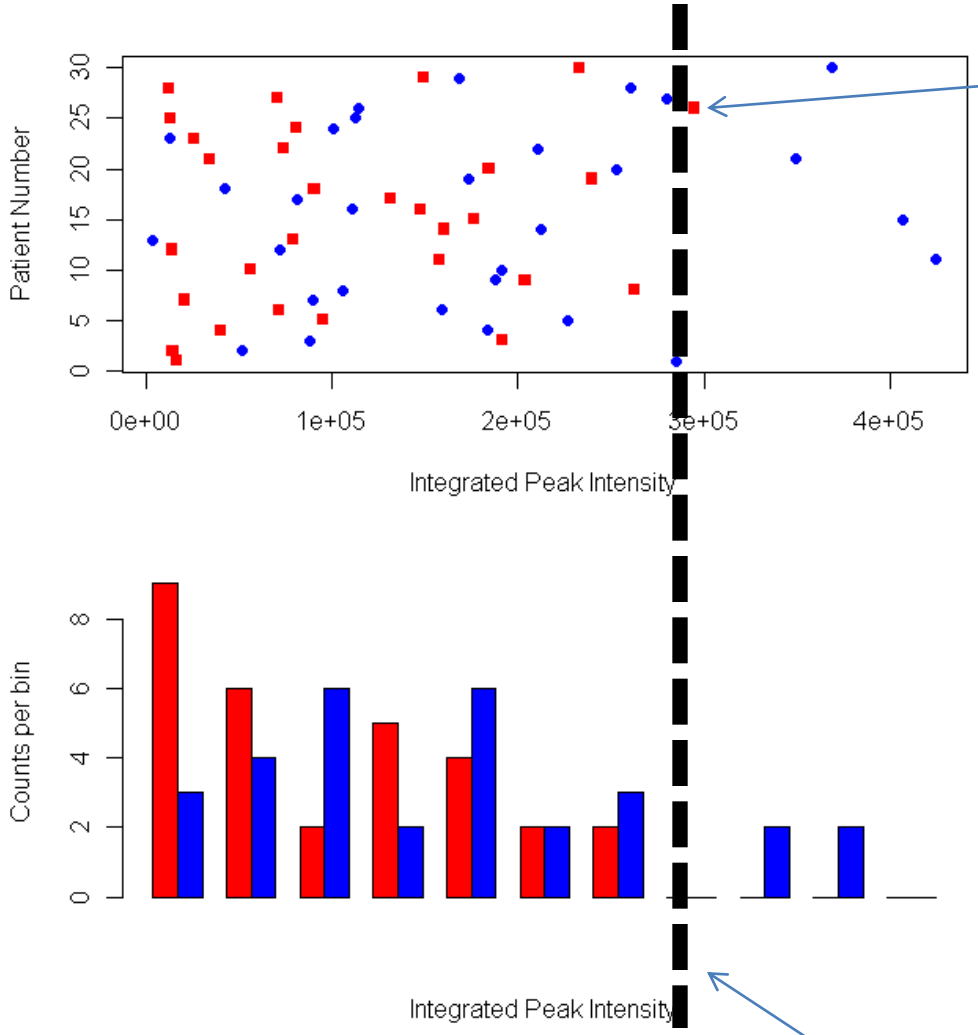
athanas@vastscientific.com

110110

Motivation

What is the significance of a measured ROC AUC (area under the curve)?

Current Method of Calculating ROC



One treatment (red) measurement above threshold

Four control measurements (blue) above threshold

Sensitivity:
True Positives = $\frac{4}{30}$

1-Specificity:
False Positives = $\frac{1}{30}$

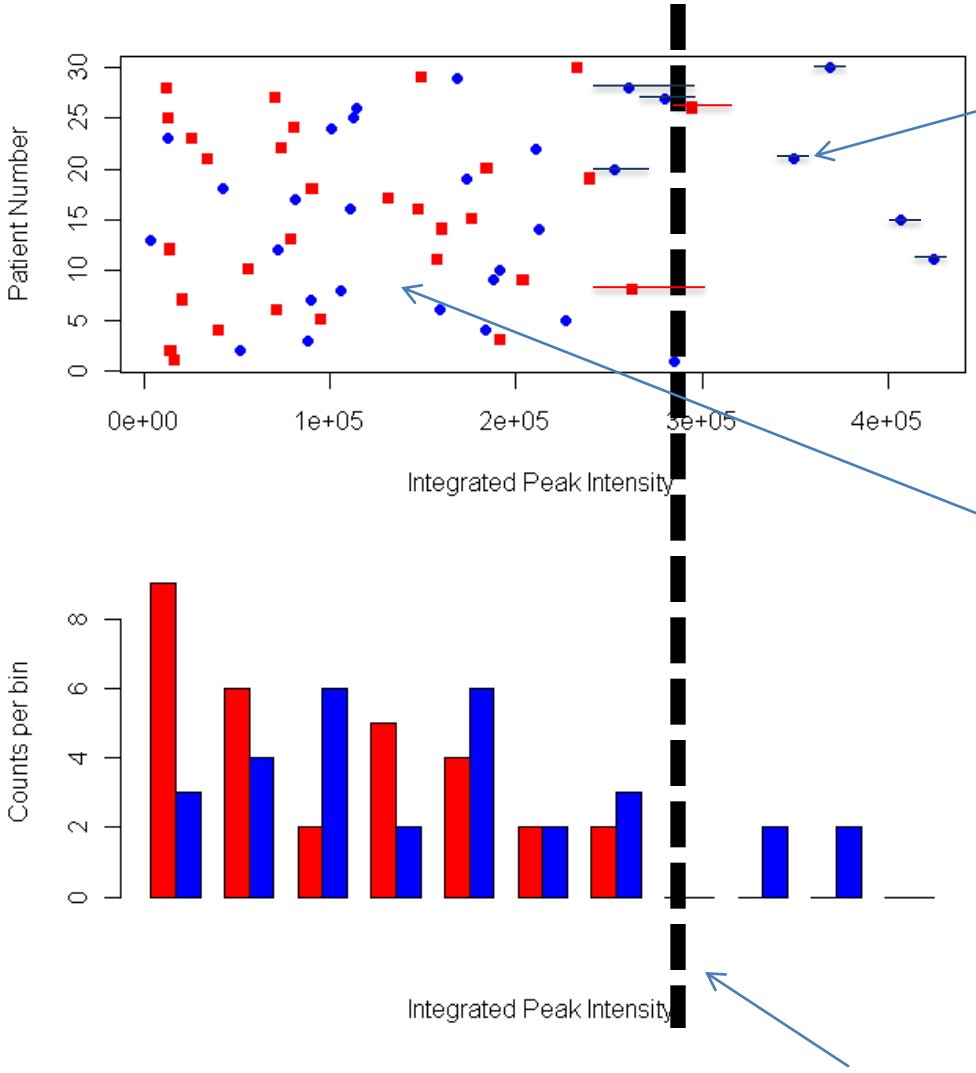
Pick a threshold value (arbitrary)

Problem with this approach

- What is the significance of AUC?

- Needs to be done:
  - Proper error propagation analysis
  - ROC Station

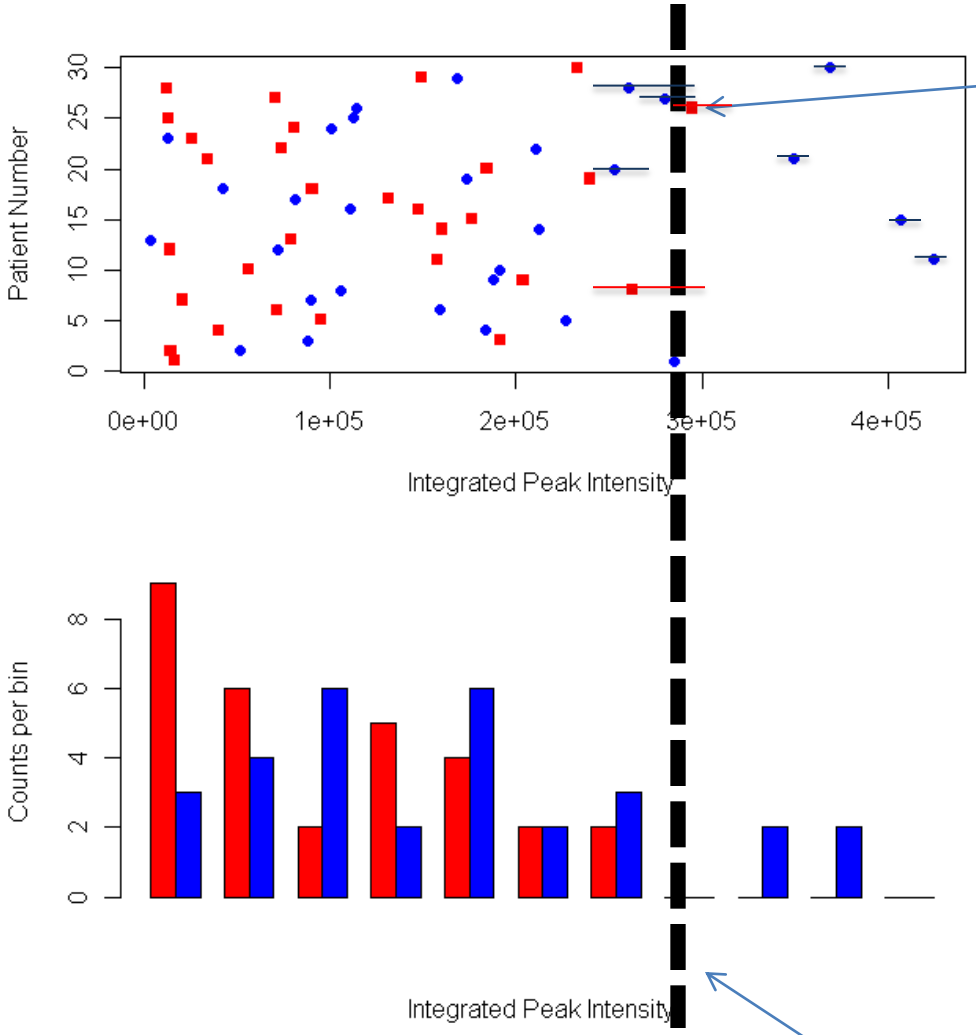
Measurement Uncertainty



Uncertainty is inherit with any kind of measurement.

(I am too lazy to draw the bars on all these dots.)

Measurement Uncertainty



One treatment (red) measurement above threshold

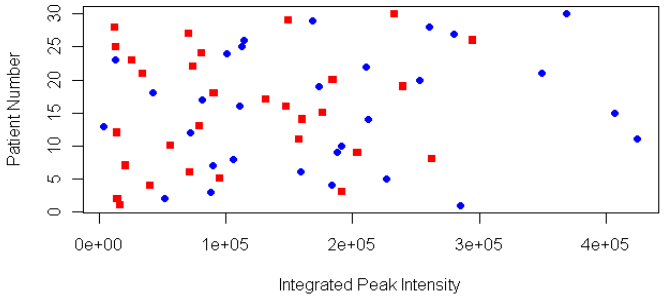
True positives: 4/30 →6/30

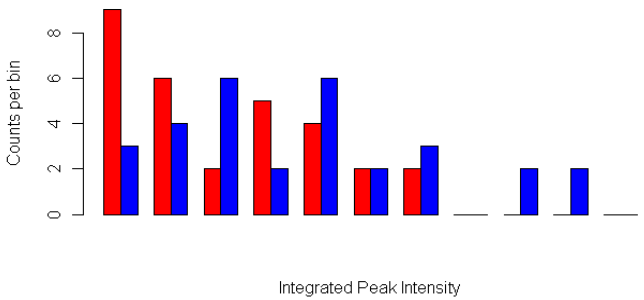
False positives: 0/30→ 2/30

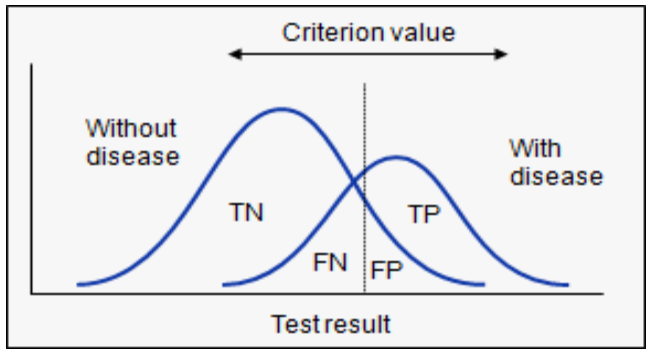
Pick a threshold value (arbitrary)

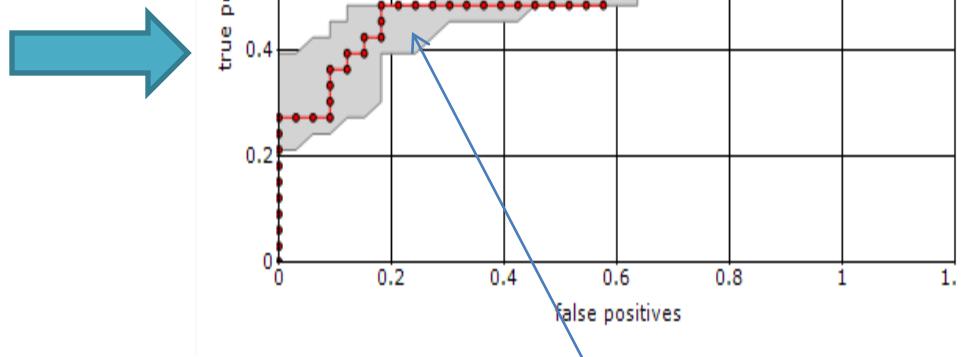
ROC Plot

Plot *Sensitivity* versus (*1-Specificity*) for all possible thresholds







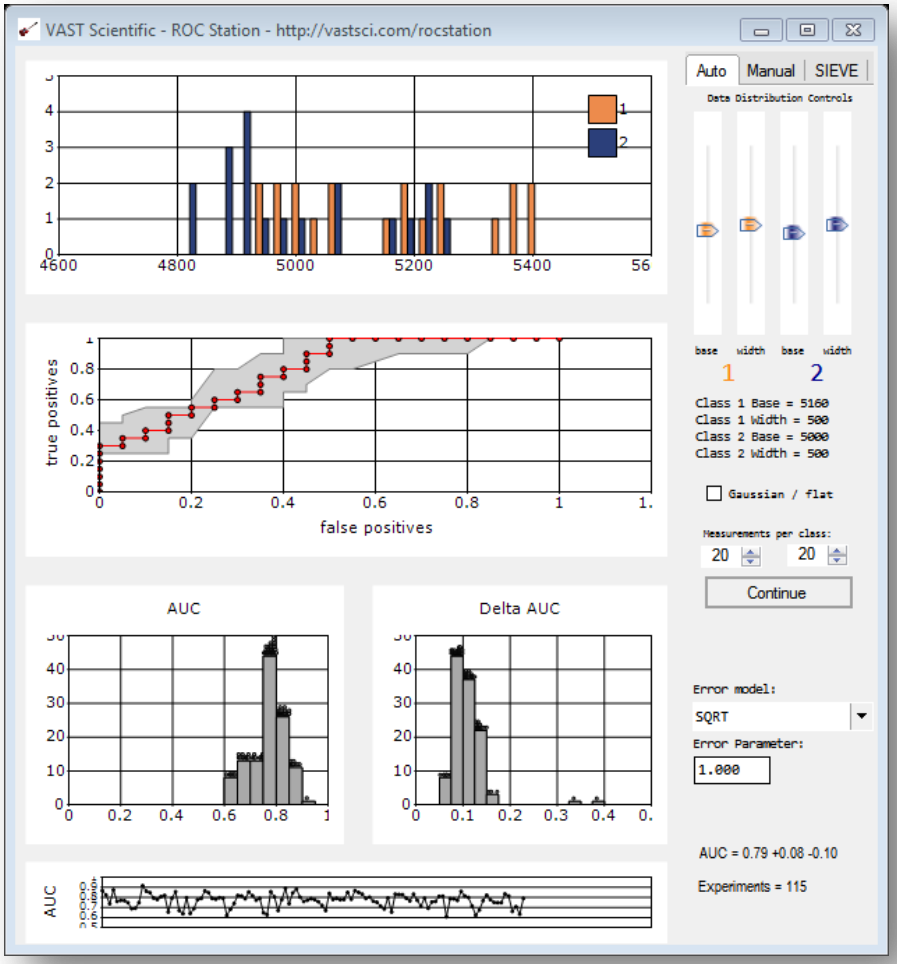


Band represents error bounds

http://vastsci.com/rocstation

ROC STATION

ROC Station

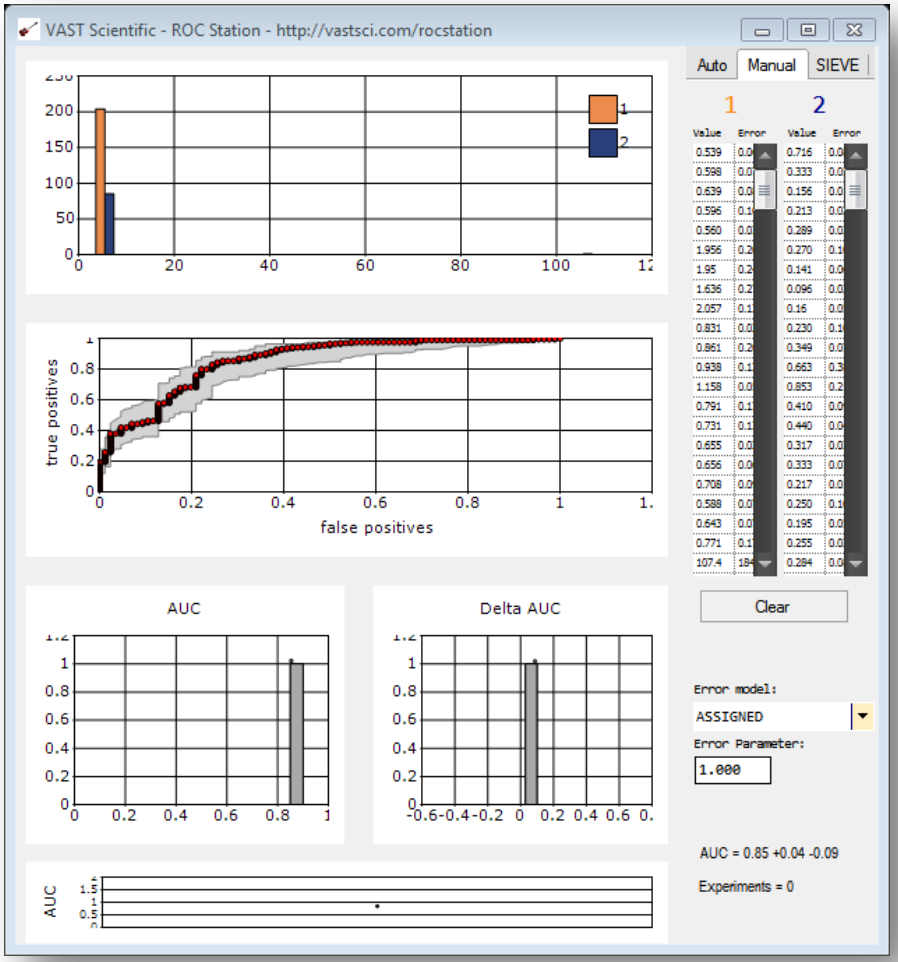


How to ROC – a learning center

Generate synthetic data to understand the effects of:

- Number of measurements
- Measurement data distributions
- Measurement error models
- Accumulated effects over repeated measurements

ROC Station

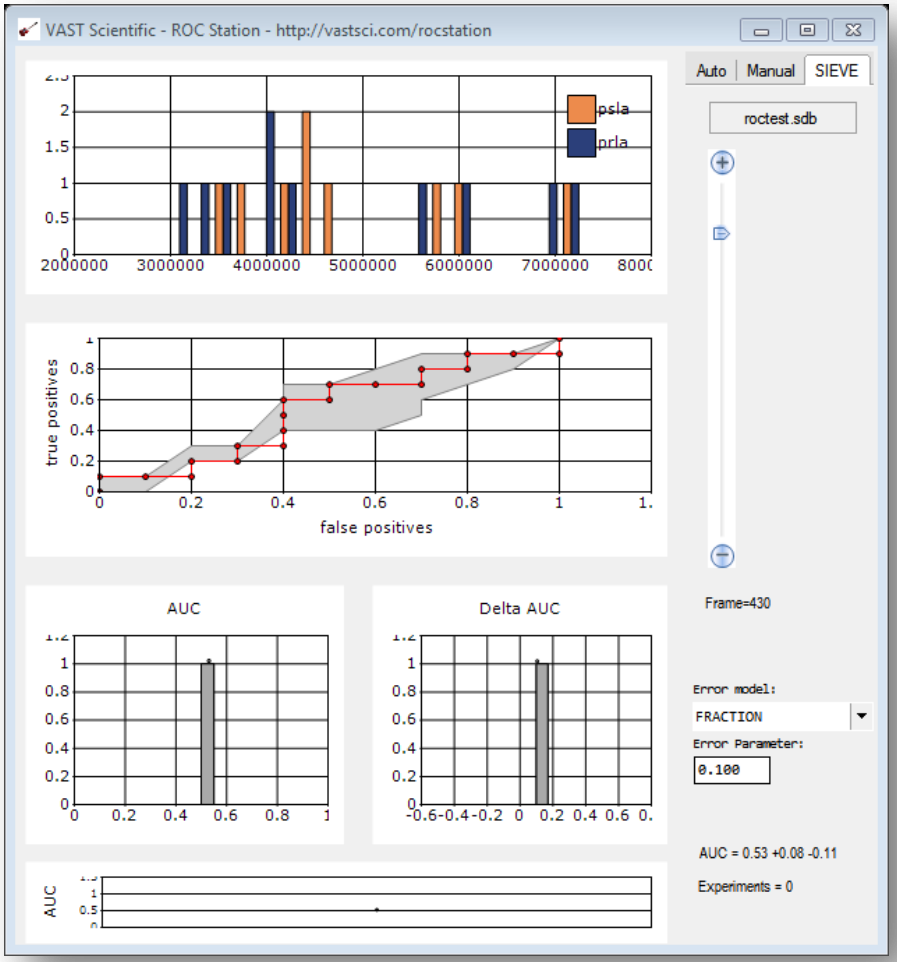


Measure real data

Paste in real measurements

Measurements can included assigned errors

ROC Station



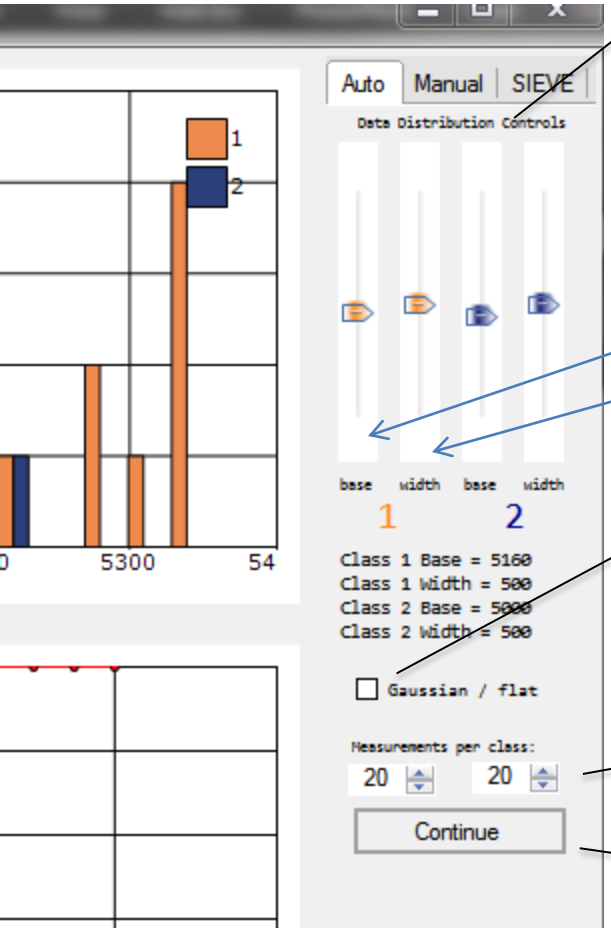
Explore SIEVE Results

Open a SIEVE result file

Explore the effects of AUC values with different error models.

DATA GENERATION

Auto Data Generator



Controls for generation of data for two classes of data

Data = Base + RandomError

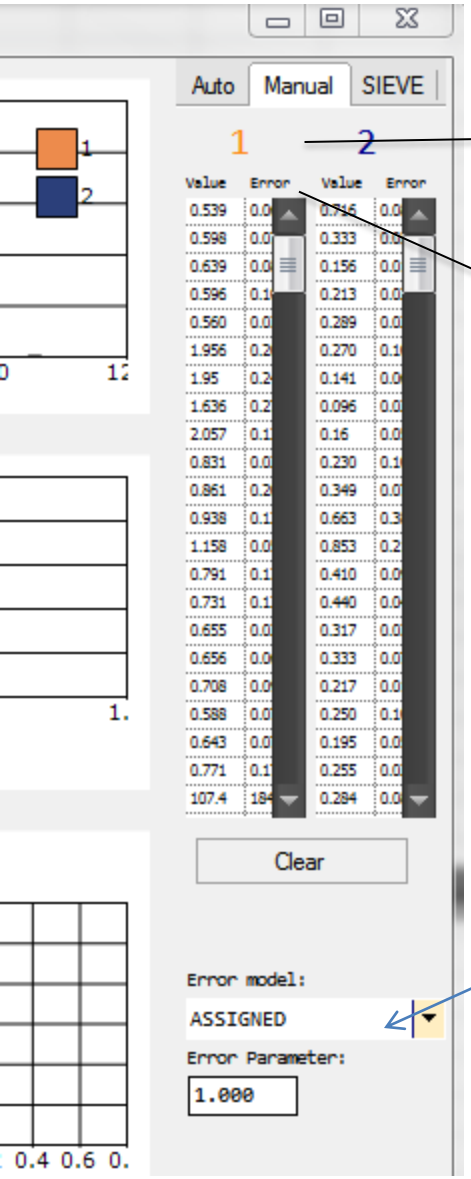
Two types of RandomError:

1. Flat distribution with defined width
2. Gaussian distribution with defined width

Number of measurements in each class

Pause live data generation

Manual Data Entry

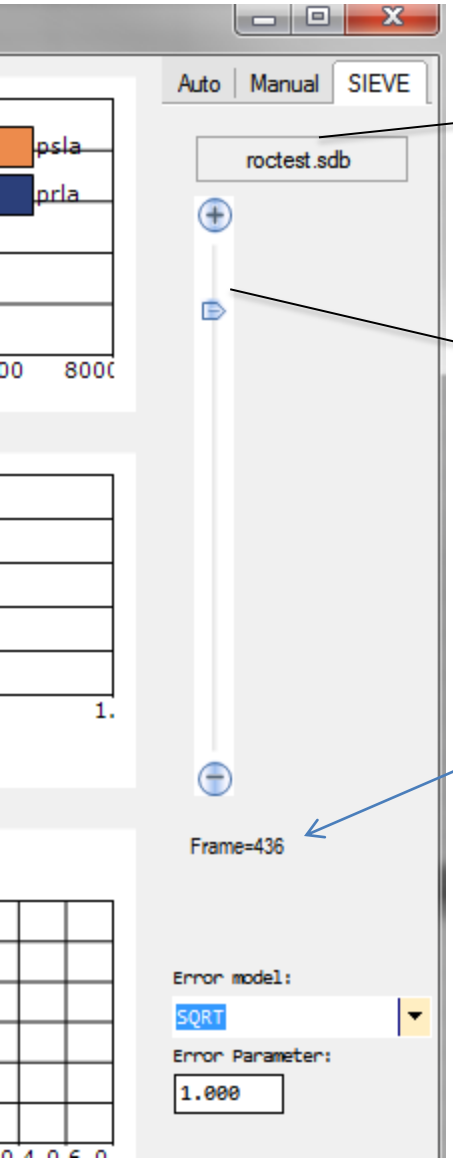


Enter data for two classes of measurements

Optionally, add error values for each measurement (ASSIGNED)

Paste values into the table from Excel

SIEVE Data Entry



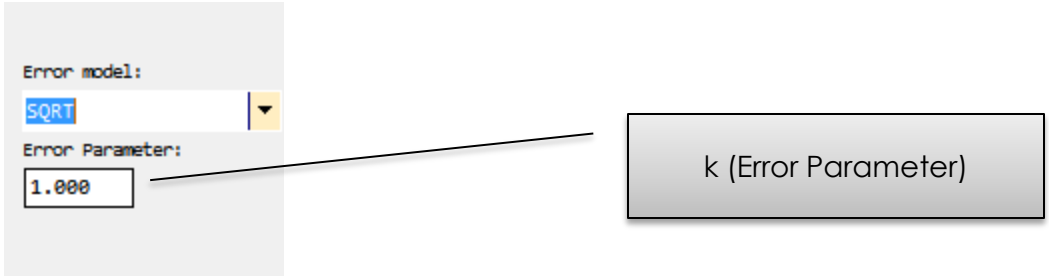
Open a SIEVE results file

Browse frames by moving the slider

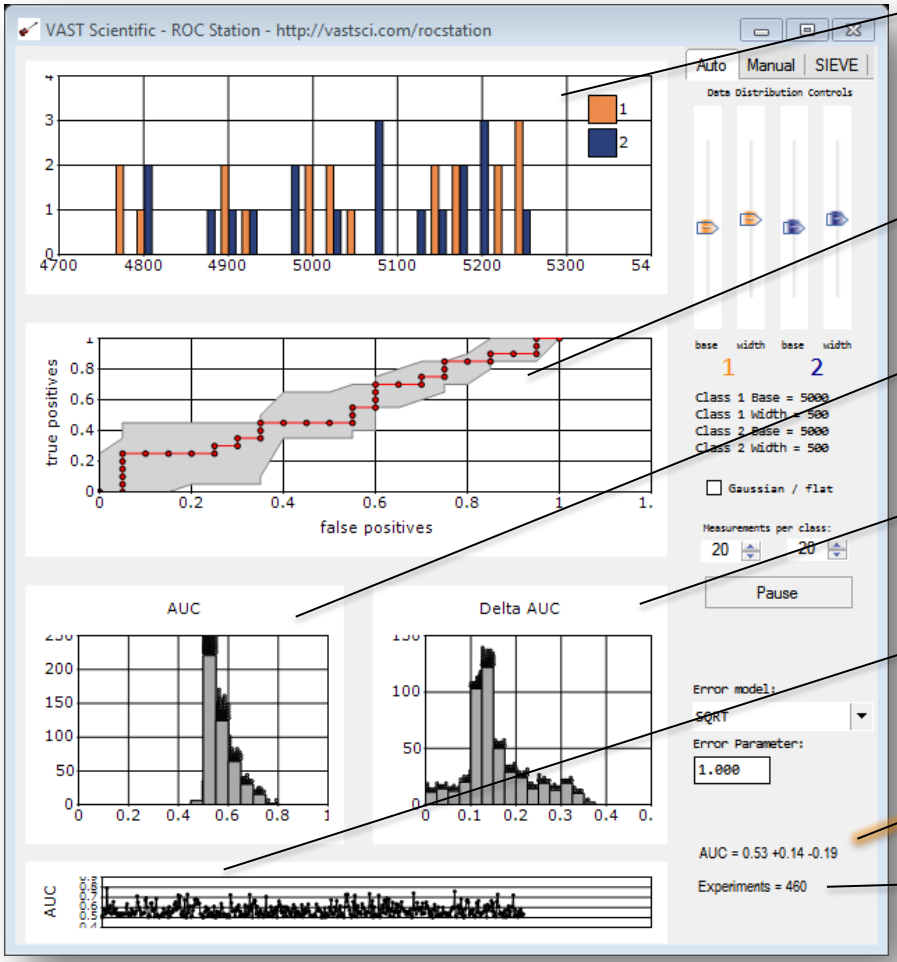
ERROR MODELS

5 Kinds of Error Models

- NONE
- SQRT: Error = k*SQRT(value)
- CONSTANT: Error = k
- FRACTION: Error = k*(value)
- ASSIGNED: for Manual data entry



Results View



Histogram of data distribution

ROC Plot

Distribution of AUC for entire experiment

StdDev(AUC)

Strip chart of AUC measurements for experiment

AUC for current measurement with error quotation

Number of measurements in the experiment

Questions, comments, complaints

athanas@vastscientific.com