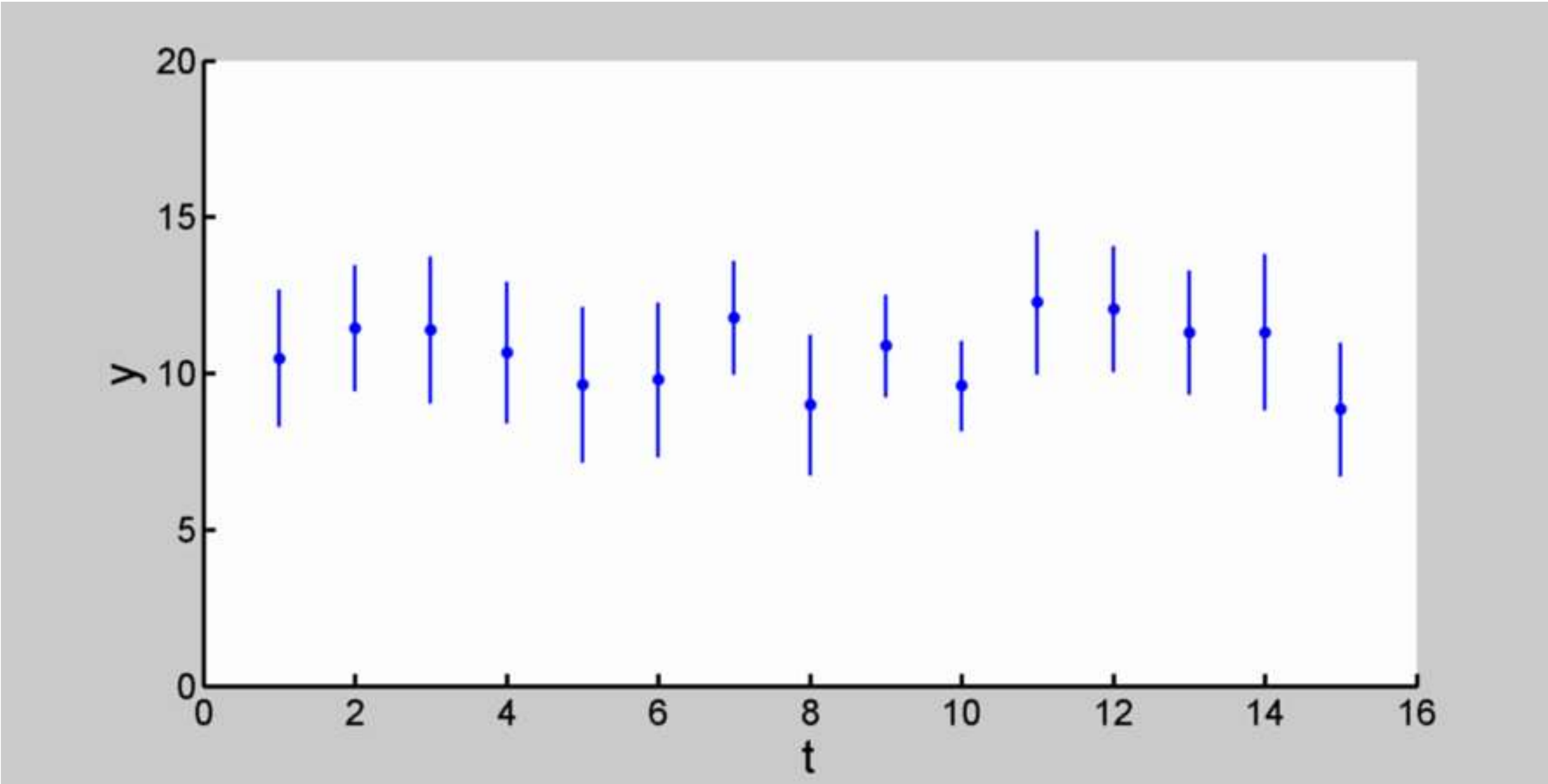
Analyzing choices in data reduction: Model Selection

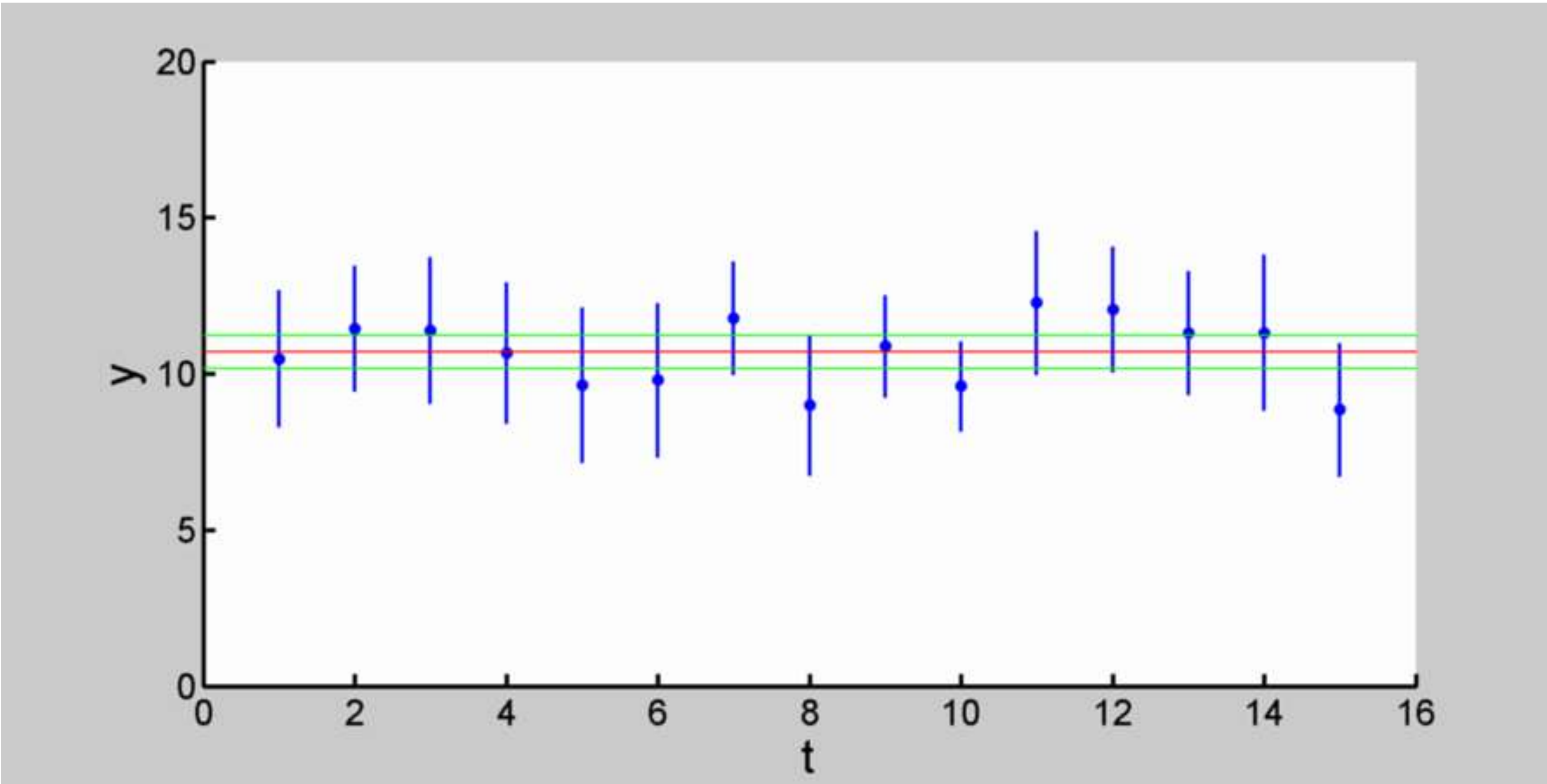
Noah McLean

Model Selection



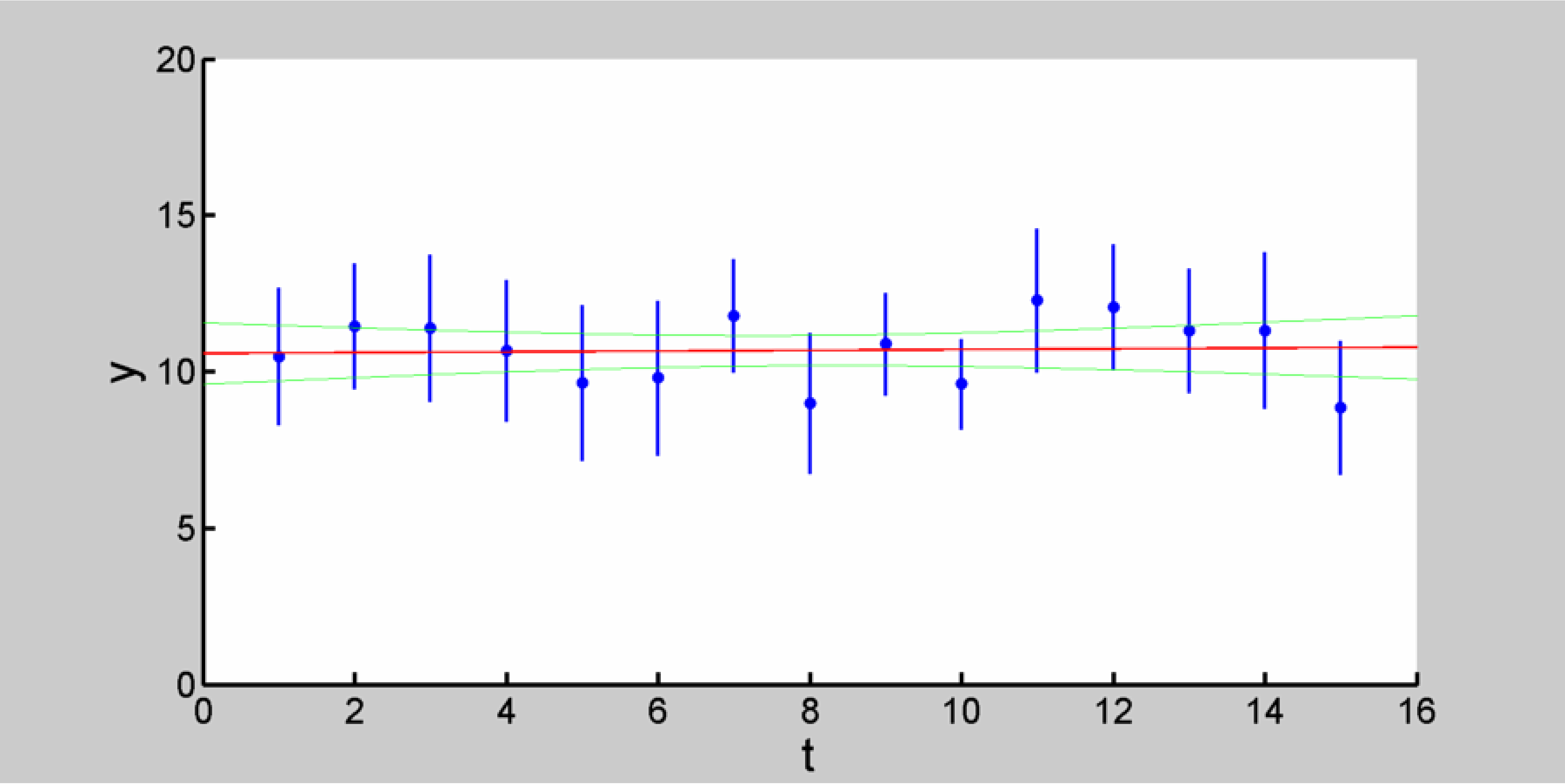
Do these measured data have a trend? Should I use a mean or a linear fit to describe the data?

Model Selection



Weighted mean: 10.69 ± 0.53 (2σ), MSWD = 1.38

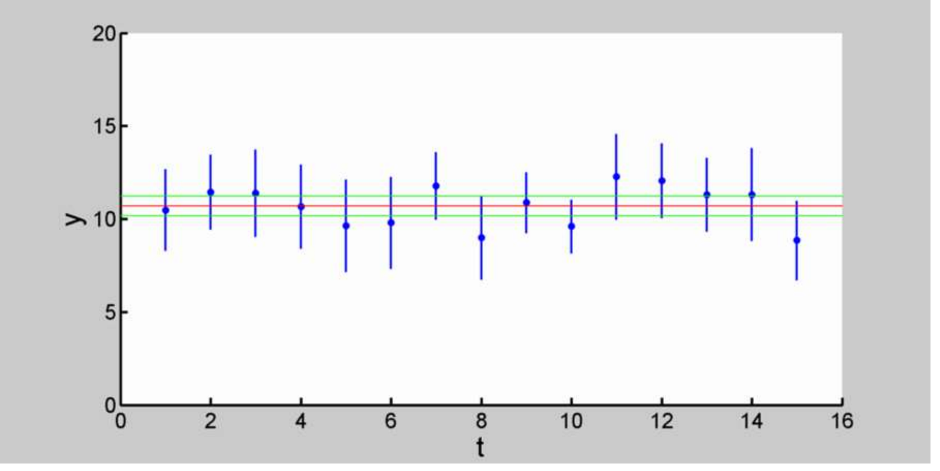
Model Selection



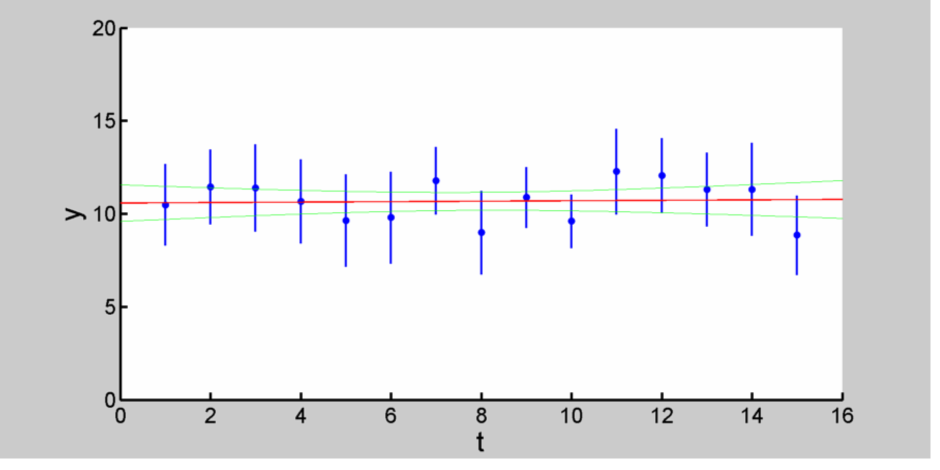
Line fit: $(0.012 \pm 0.11)t + (10.58 \pm 0.98)$, MSWD = 1.13

$\boldsymbol{\rho_{ab} = -0.871}$

Model Selection



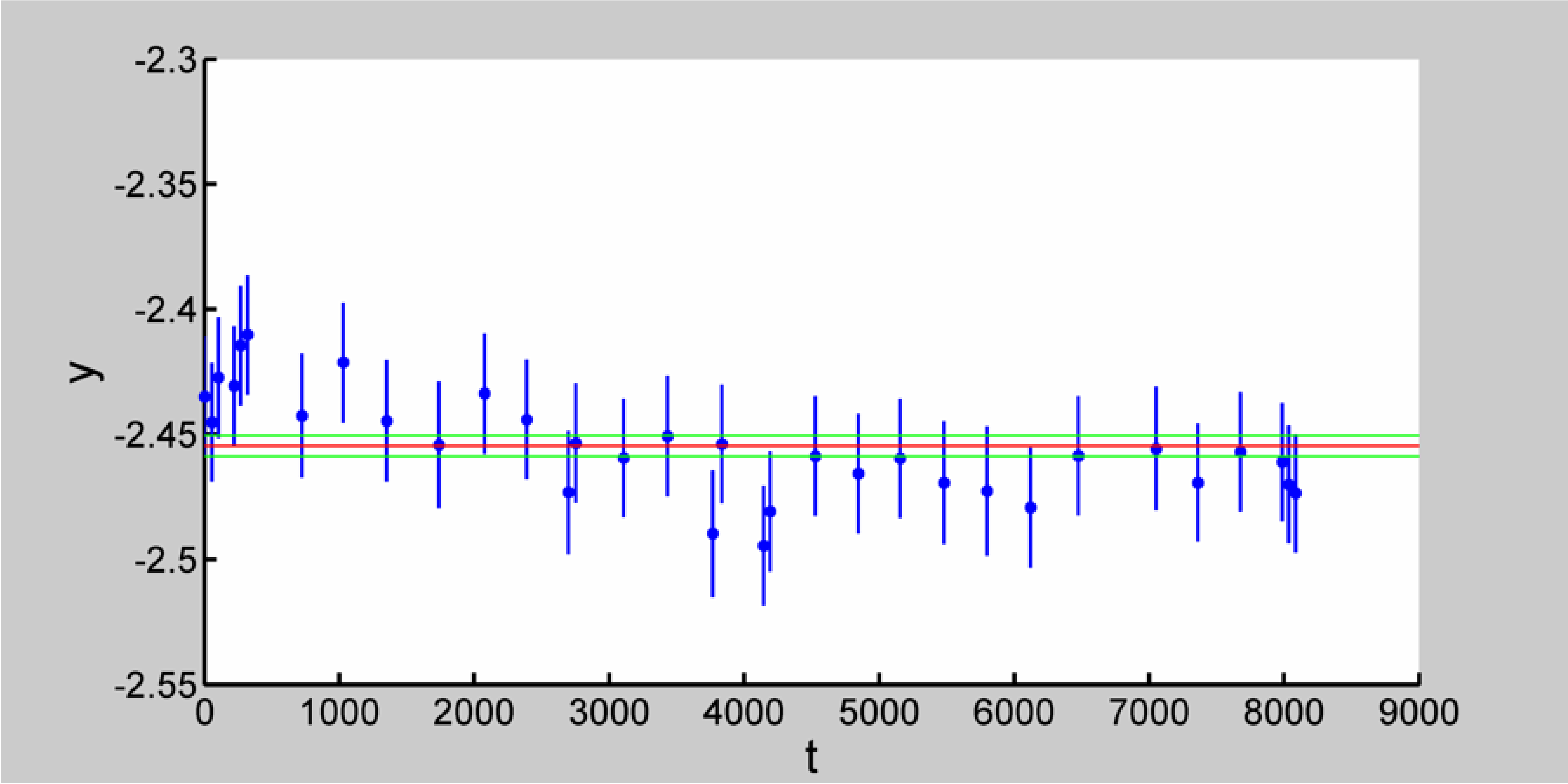
<u>Weighted mean</u>:
10.69 ± 0.53 (2σ),
MSWD = 1.38, **BIC = 20.7**



<u>Line fit</u>: (0.012 ± 0.11)*t* +
(10.58 ± 0.98),
MSWD = 1.13, **BIC = 19.9**

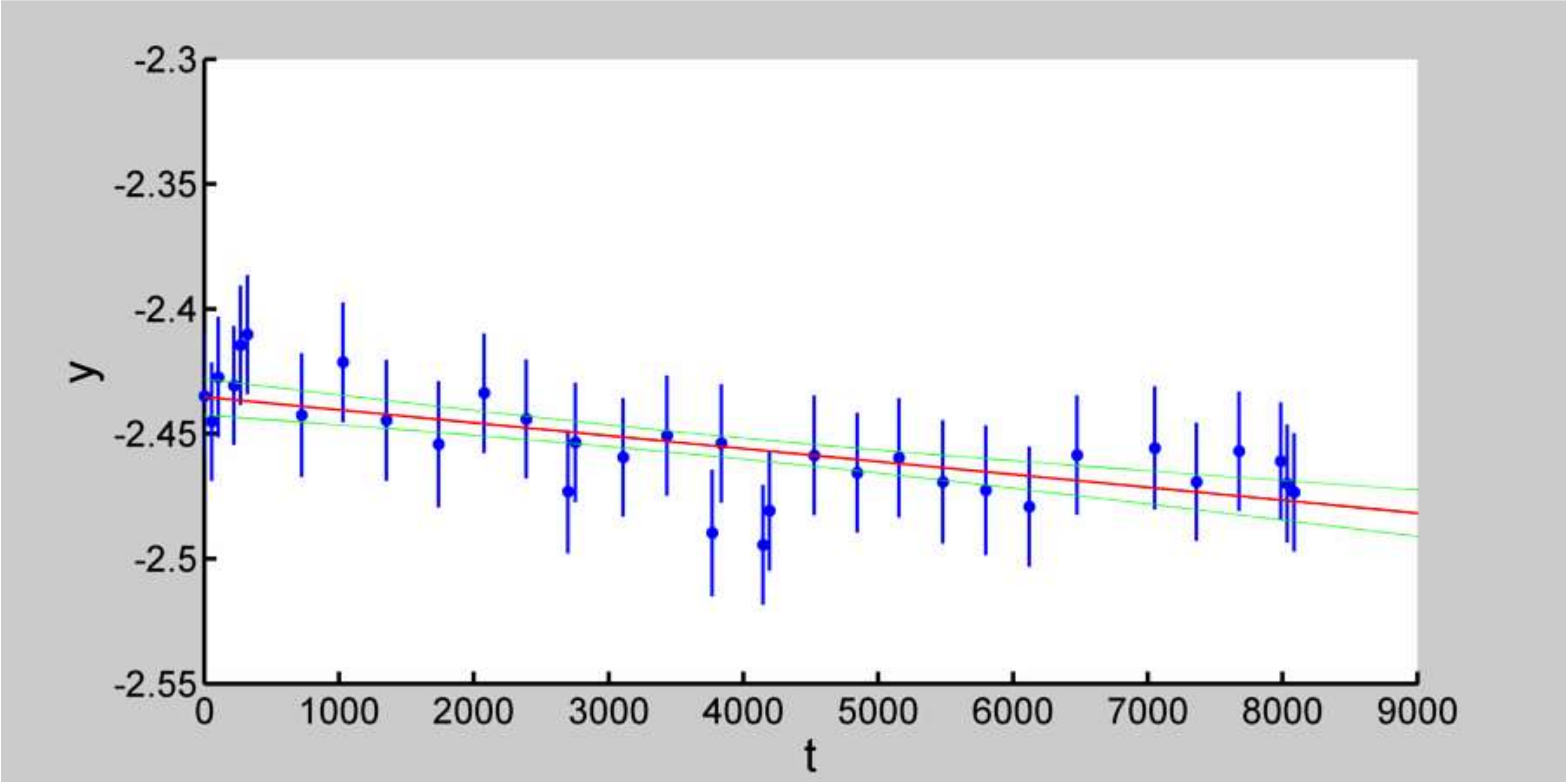
Lowest BIC wins: the **line** is most likely the best fit.

Session fits:



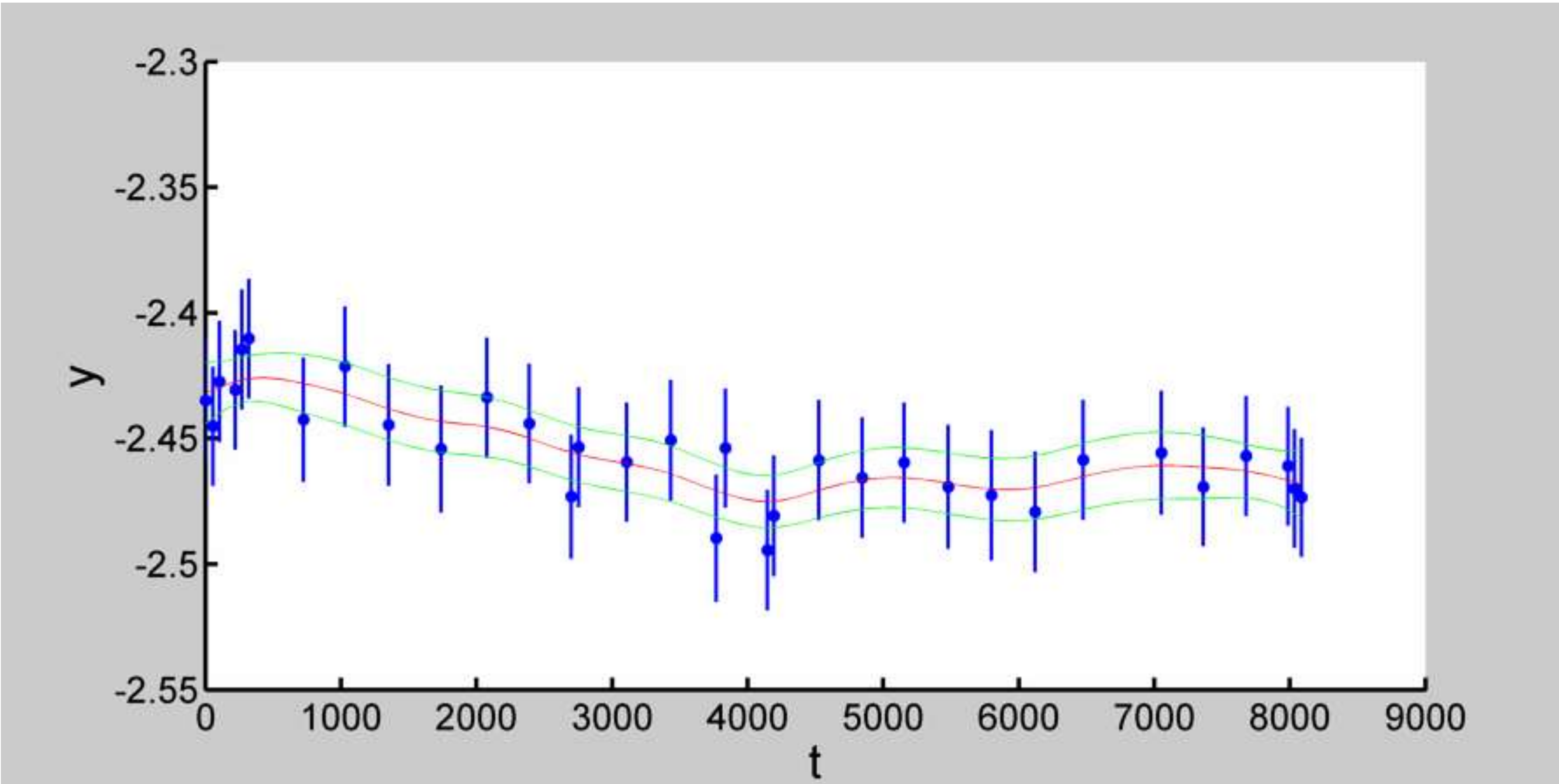
Weighted mean: -2.4545 ± 0.0042 (2σ), MSWD = 2.88

Session fits:

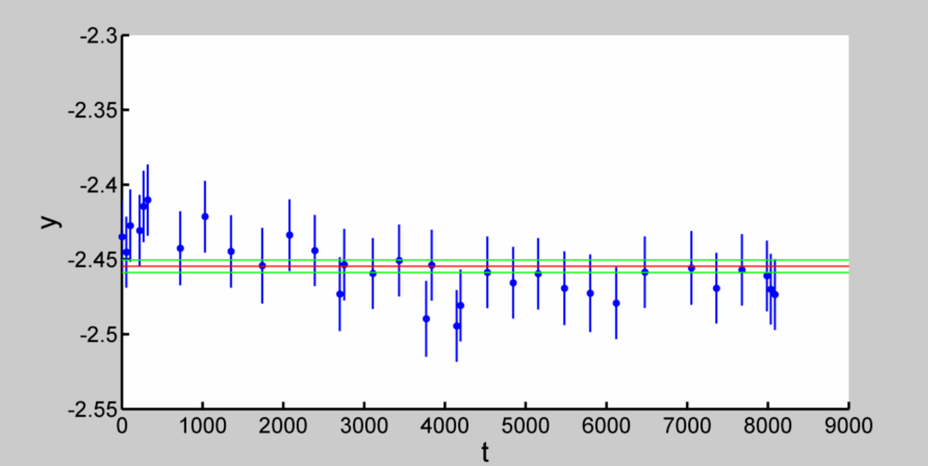


Line fit: MSWD = 1.54

Session fits:



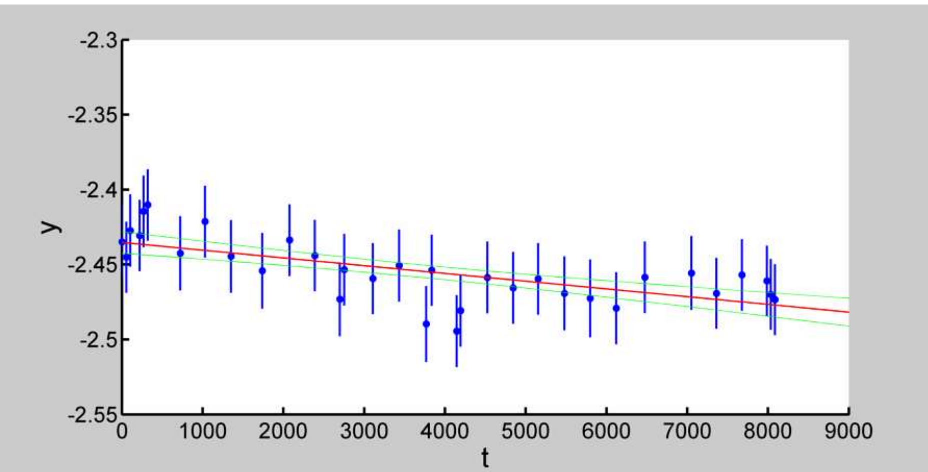
Spline fit: MSWD = 1



Weighted mean:

MSWD = 2.88,

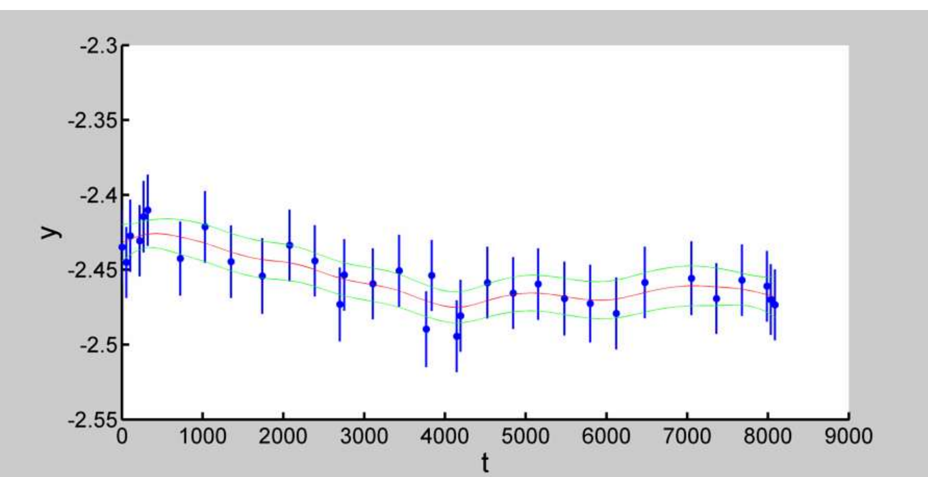
BIC = -196



Line:

MSWD = 1.54,

BIC = -235.6



Spline:

MSWD = 1.00,

BIC = -236.1

Calculation of the BIC

$$BIC = -2\log(L) + m\log(n)$$

m	number of parameters used in the model
n	number of data points measured
n-m	degrees of freedom
L	likelihood function

Calculation of the BIC

$$BIC = -2\log(L) + m\log(n)$$

m	number of parameters used in the model
n	number of data points measured
n-m	degrees of freedom
L	likelihood function

$$\log L = -\frac{1}{2}\left[\ (x-\hat{x})^{\mathrm{T}}\,\Sigma^{-1}(x-\hat{x}) + \log|\Sigma|\ \right]$$

Calculation of the BIC

m	number of parameters used in the model
n	number of data points measured
n-m	degrees of freedom
L	likelihood function

$$BIC = (x - \hat{x})^{\mathrm{T}} \Sigma^{-1} (x - \hat{x}) \ + \ \log |\Sigma| \ + \ m \log(n)$$

Calculating the effective degrees of freedom:

1. Use the trace of the hat matrix:

$$\hat{y} = Hy$$

Calculating the effective degrees of freedom:

1. Use the trace of the hat matrix:

$$\hat{y} = Hy$$

2. Shatterthwaite approximation:

$$edf = \mathrm{trace}(H^{\mathrm{T}}H)^2 / \mathrm{trace}(H^{\mathrm{T}}HH^{\mathrm{T}}H)$$

Failure of BIC:

- For fits that include an overdispersion (excess variance) calculation: if you count the overdispersion itself as an extra parameter, the mean always wins.

$$\log L = -\frac{1}{2}\sum_{i=1}^{n}\left[\ (x-\hat{x})^{\mathrm{T}}\,\Sigma^{-1}(x-\hat{x}) + \log|\Sigma|\ \right]$$

Other Model Selection parameters:

- *BIC: Bayesian Information Criterion*

$$BIC = -2\log(L) + m\log(n)$$

- AIC: Akaike Information Criterion

$$AIC = -2\log(L) + 2m$$

$$AICc = AIC + \frac{2k(k+1)}{n-k-1}$$

- MDL: Minimum Description Length