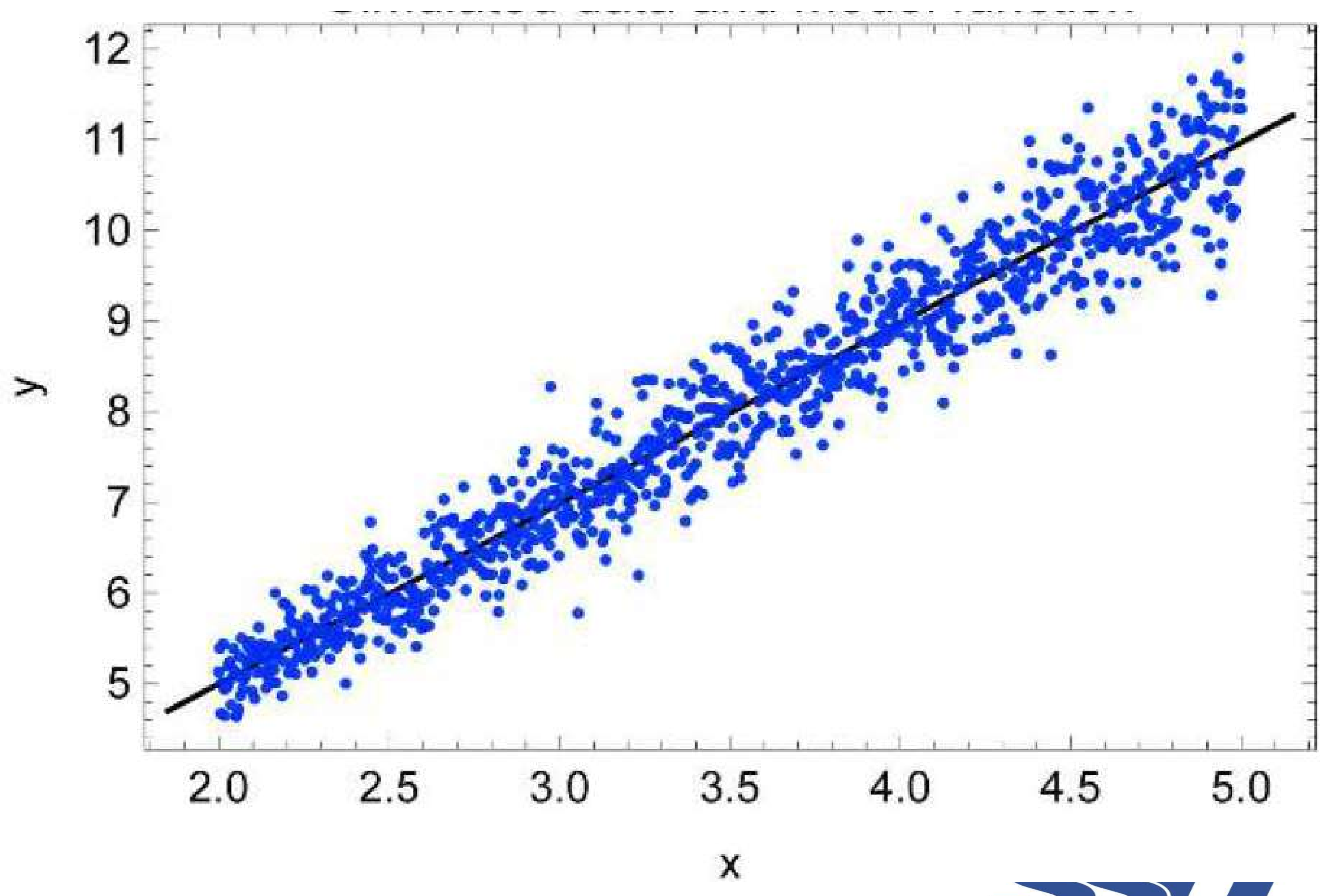
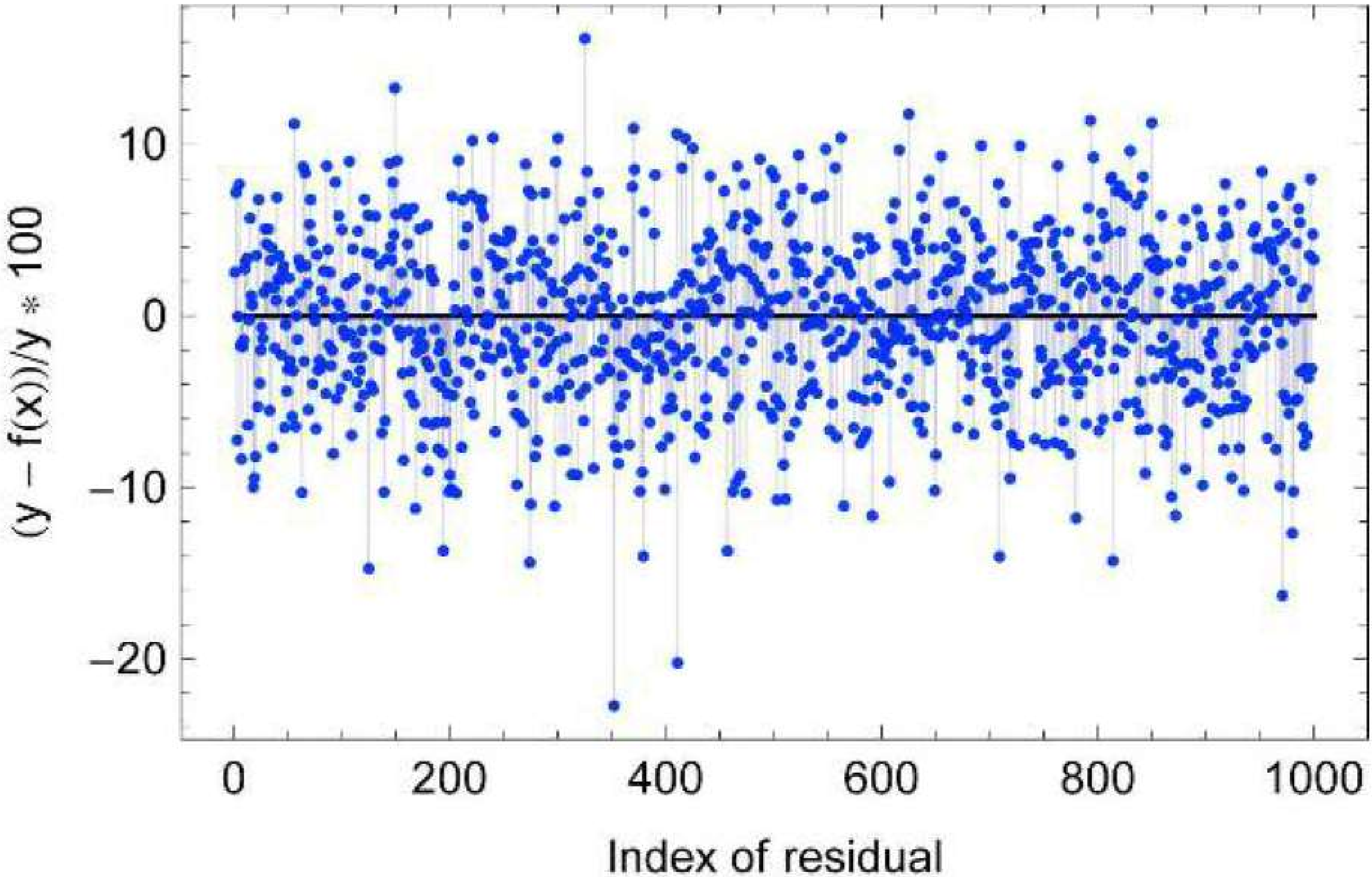
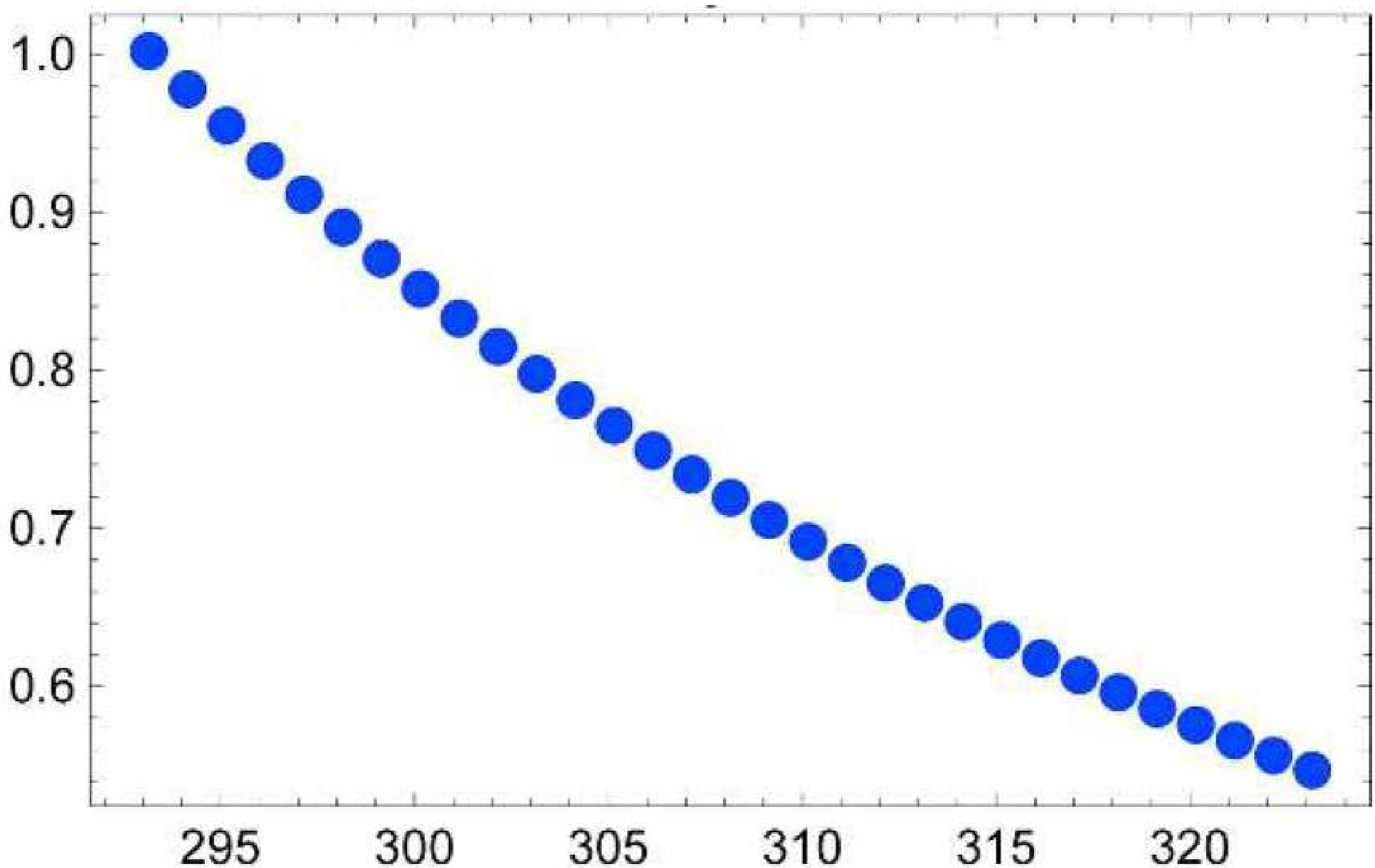
Least Mean Square Error

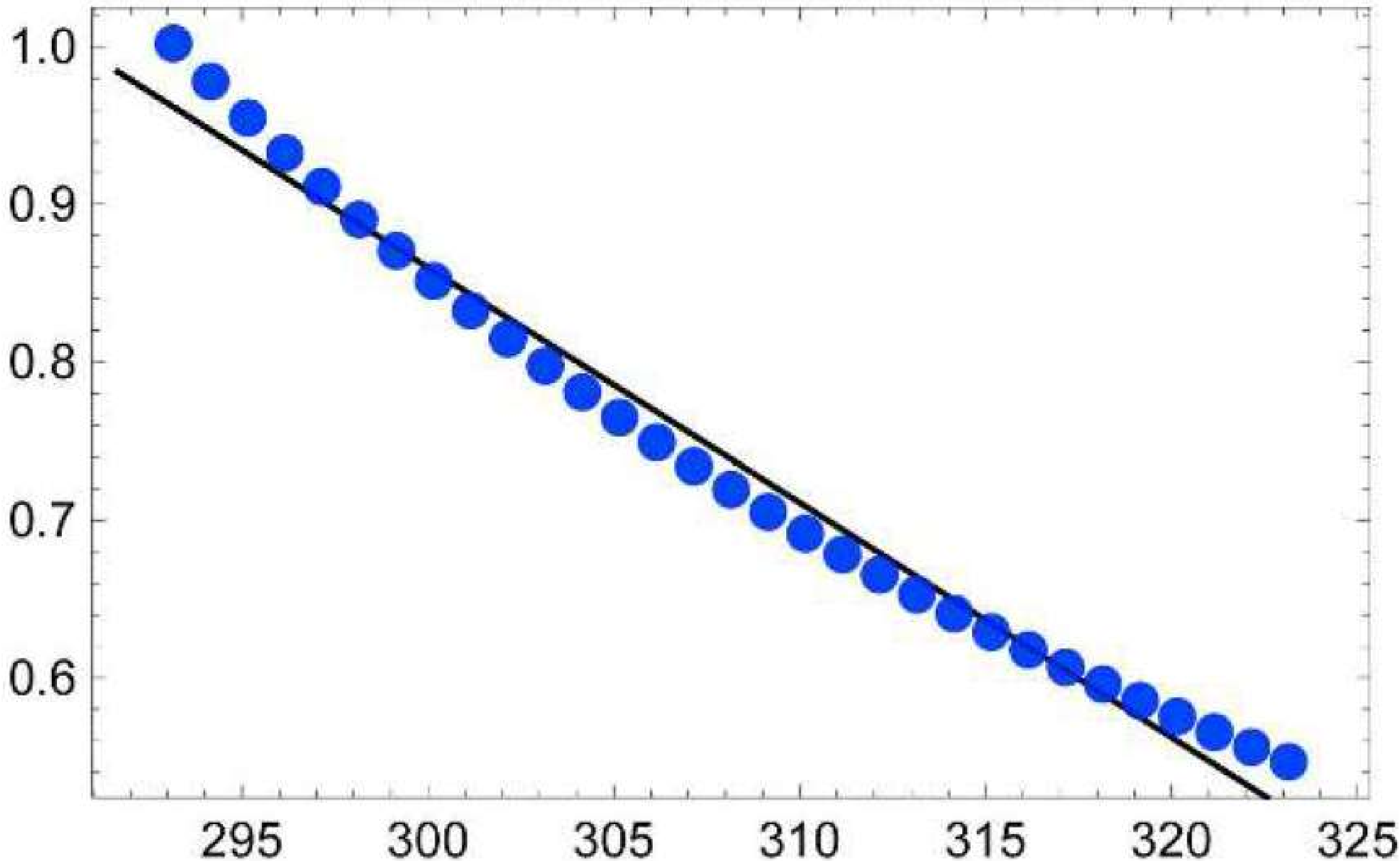

12
11
10
9
8
7
6
5
y
2.0
2.5
3.0
3.5
4.0
4.5
5.0
x


(y − f(x))/y * 100
10
0
−10
−20
0
200
400
600
800
1000
Index of residual

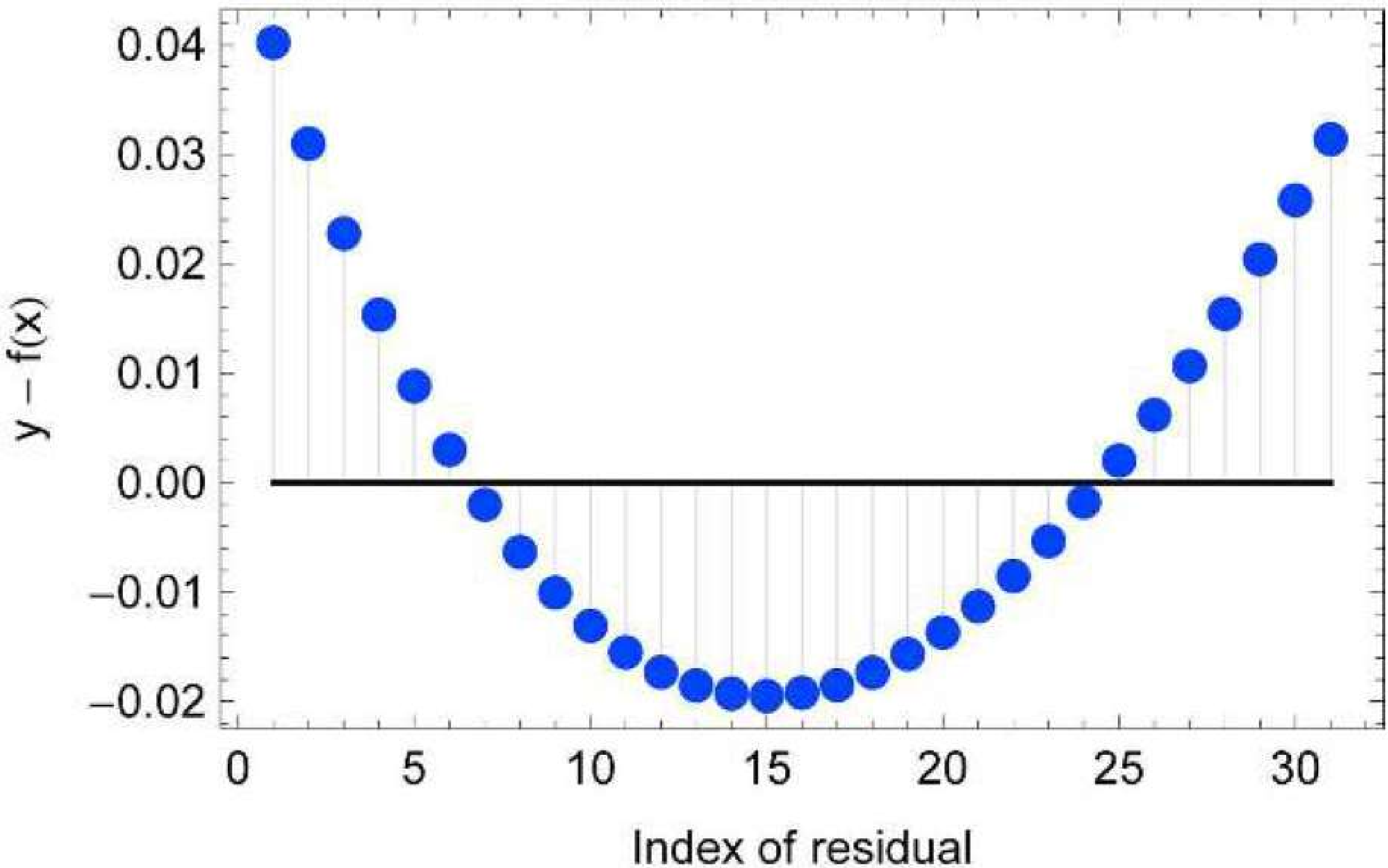
Given data



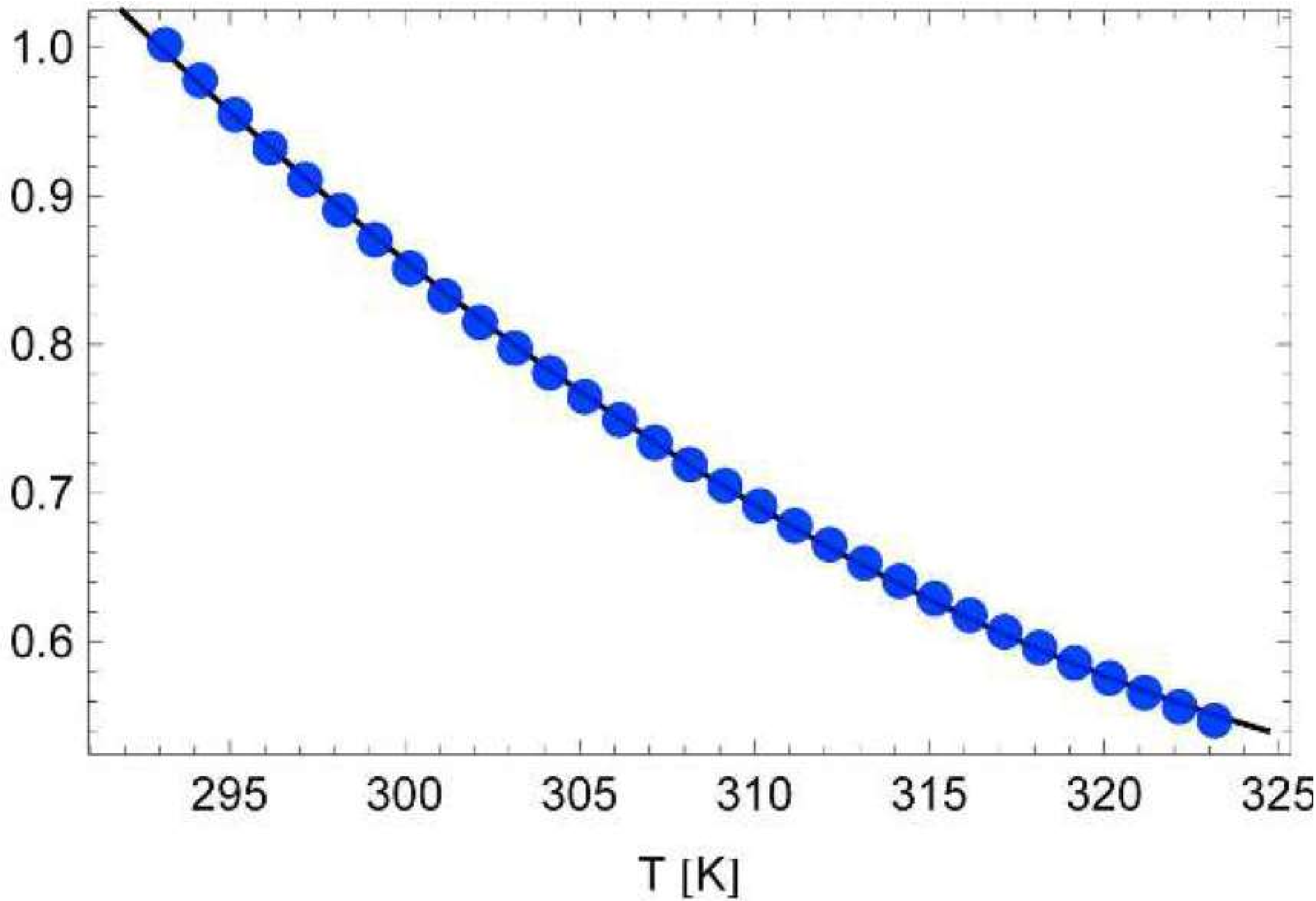
Assume a function to fit $y = a_1 + a_2x$



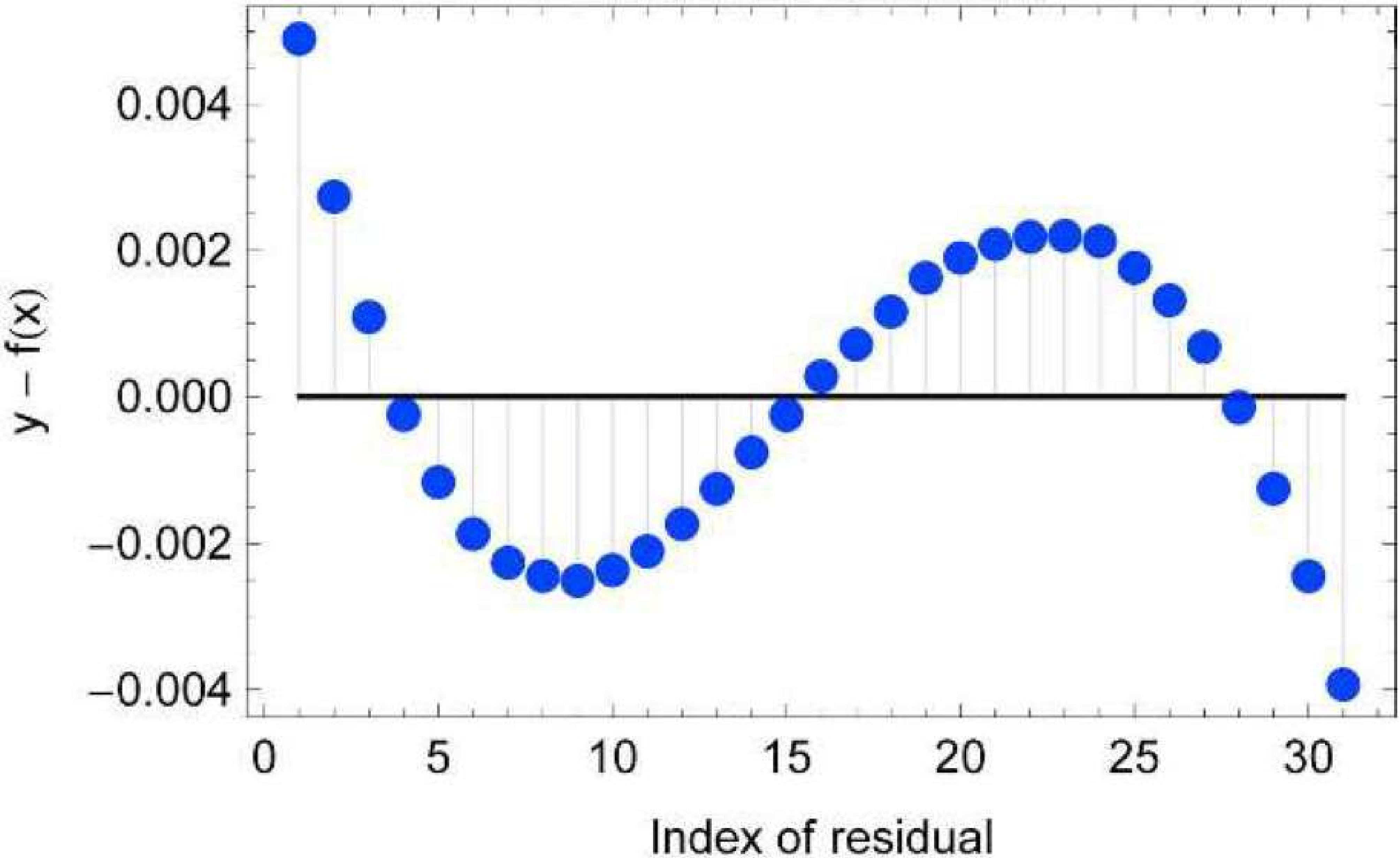
Residual or errors @ various points



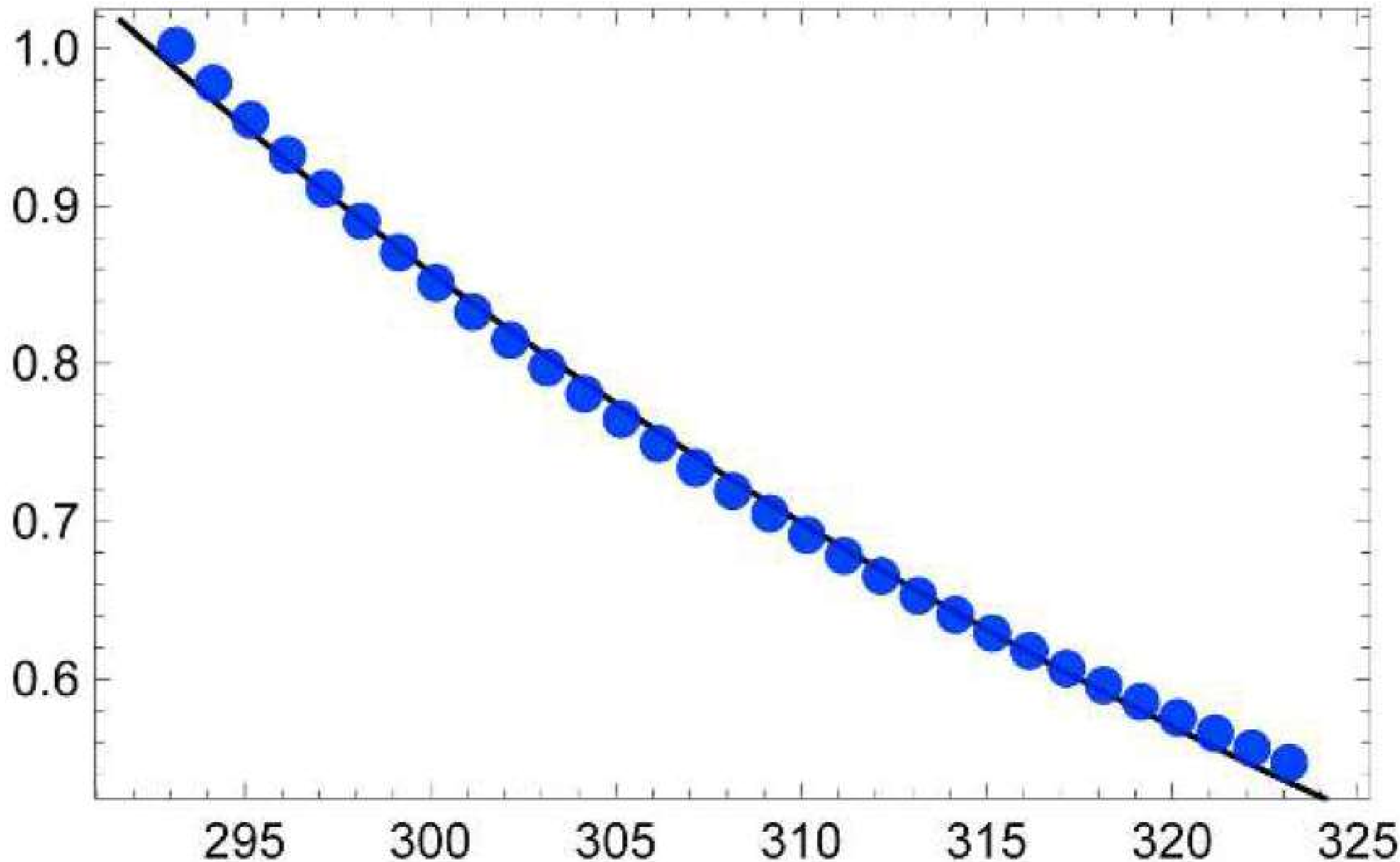
Assume function to fit $y = a_1 + a_2x + a_3x^2$



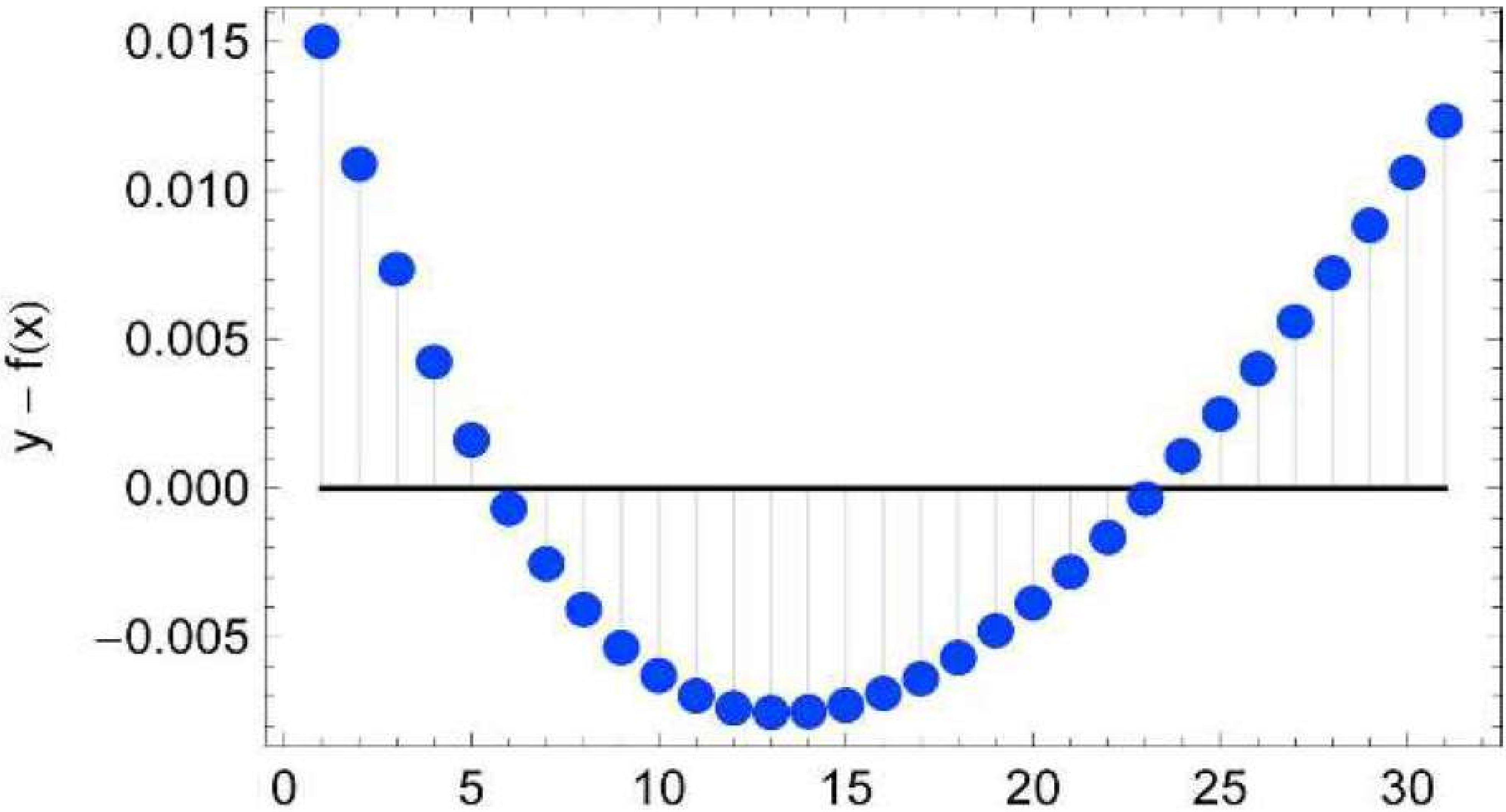
Residual or errors @ various points



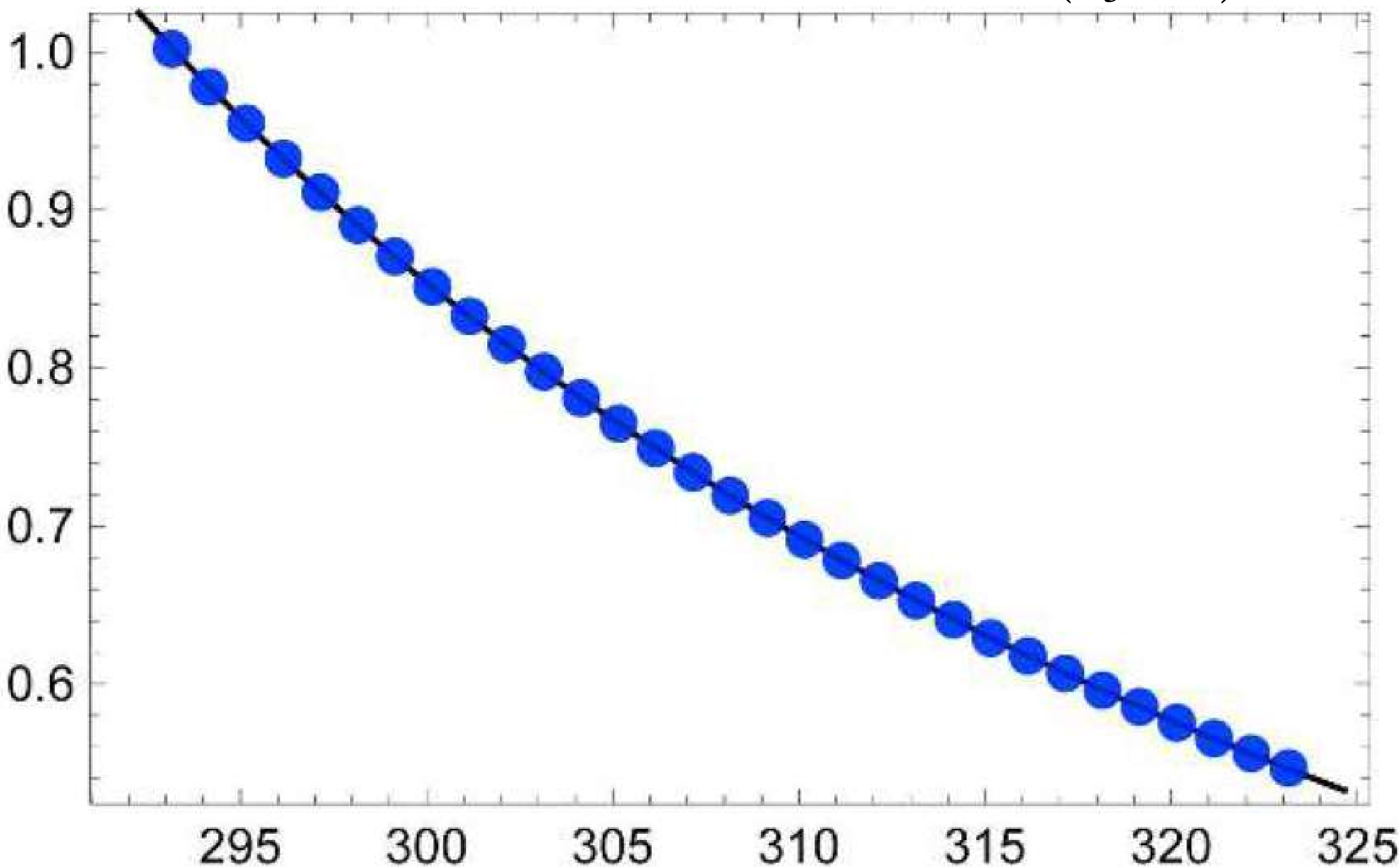
Assume function, $y = a_1 . e^{a_2 x}$



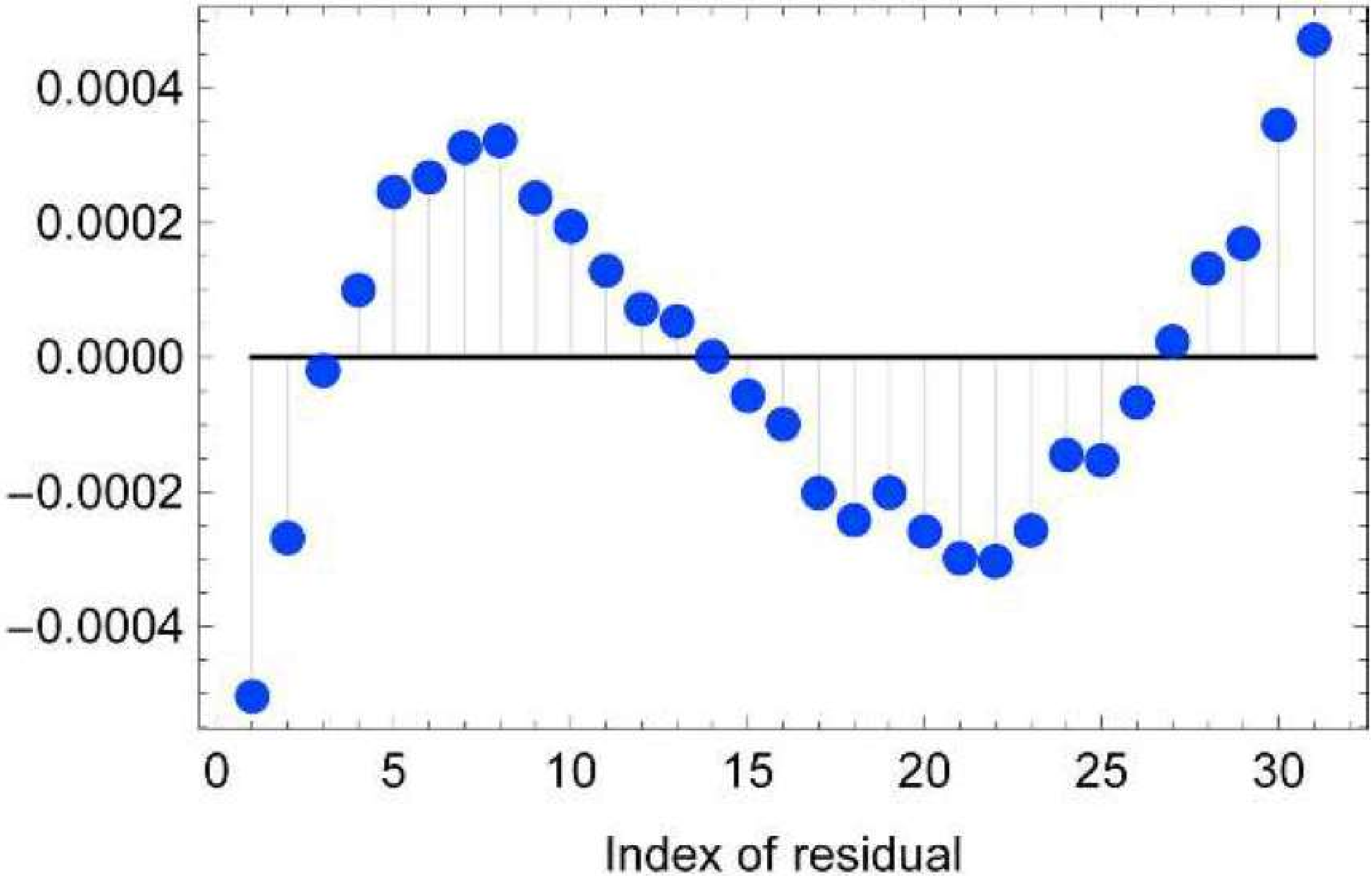
Residual or errors @ various points



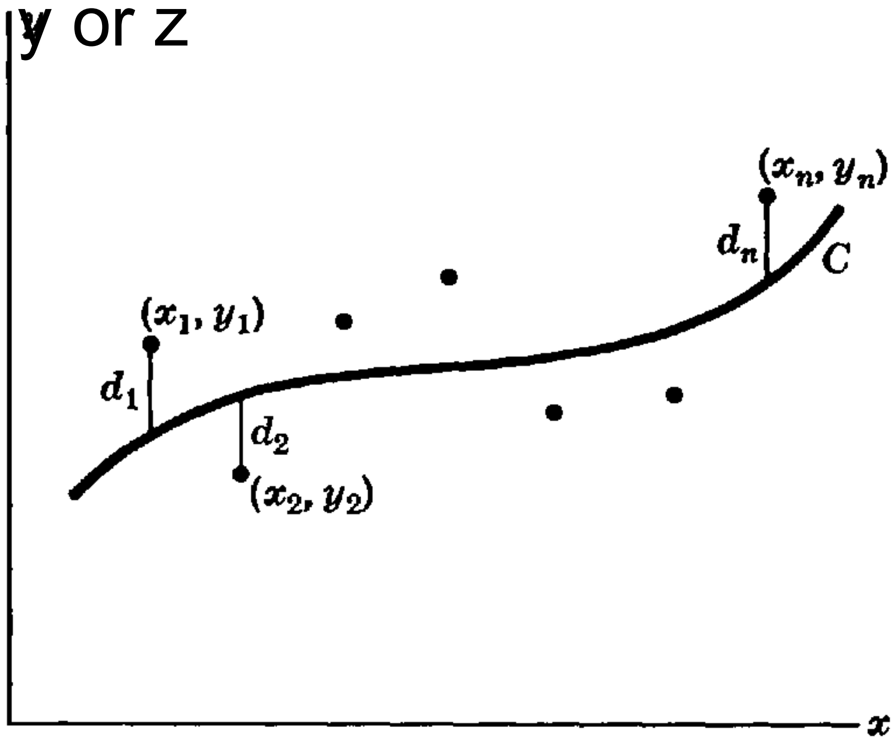
Assume function to fit , $y = a_1 + \frac{a_2}{(a_3 - x)}$



Residual or errors @ various points



Deviation - error



- Given, x_i and y_i
- K number of data points
- Error = $y_i - f(x_i)$
- How many errors?
- K errors

Sum of Squared errors

$$d_1^2 + d_2^2 + \cdots + d_k^2 = \text{ a minimum}$$

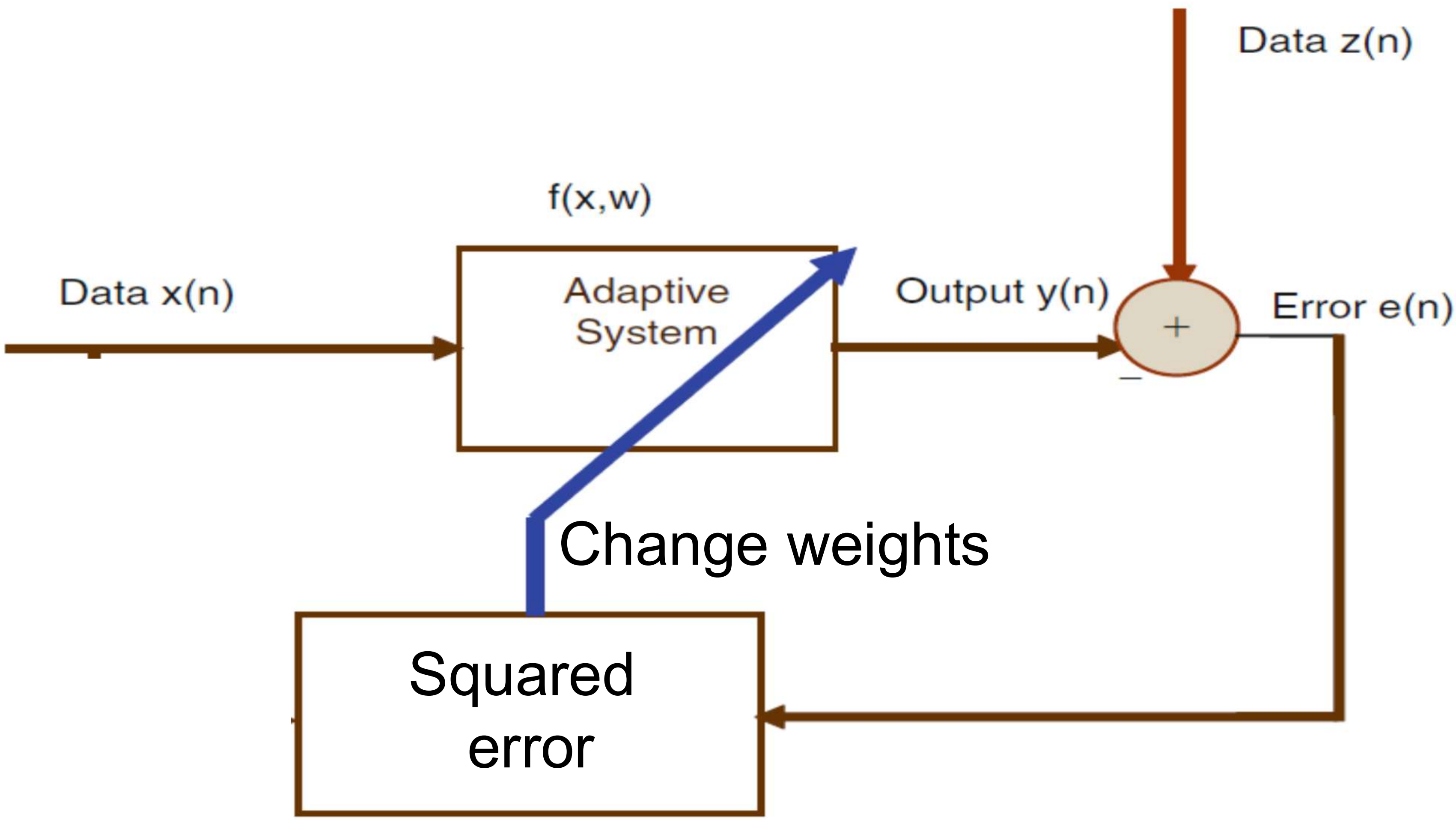
$$\sum_{i=1}^{K} (y_i - f(x_i))^2 \longrightarrow \text{minimize}$$

Root mean square error

$$\mathrm{RMSE} = \sqrt{\frac{1}{K}\sum_{i=1}^{K}\left(y_i - f\left(x_i\right)\right)^2}$$

Least MSE - Schematic

Squared error = $\left[\left(z(n) - f(\mathbf{w}, x(n))\right)^2\right]$

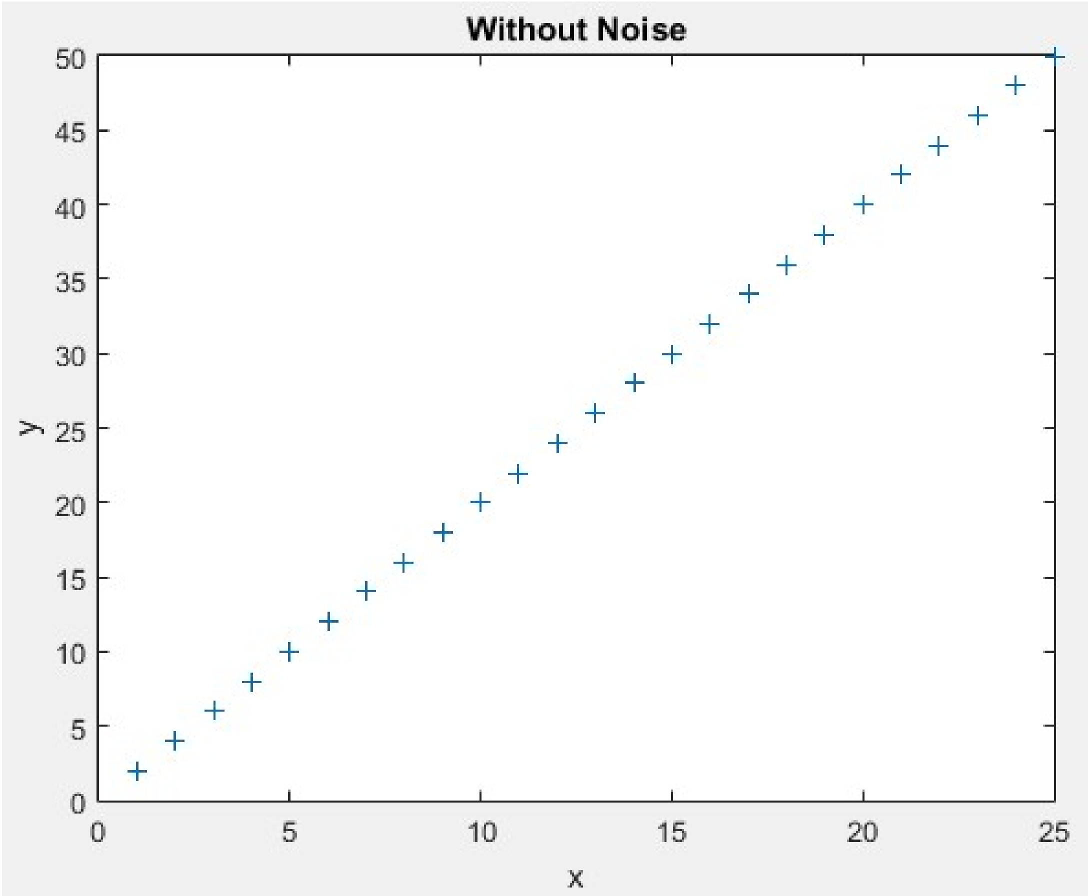


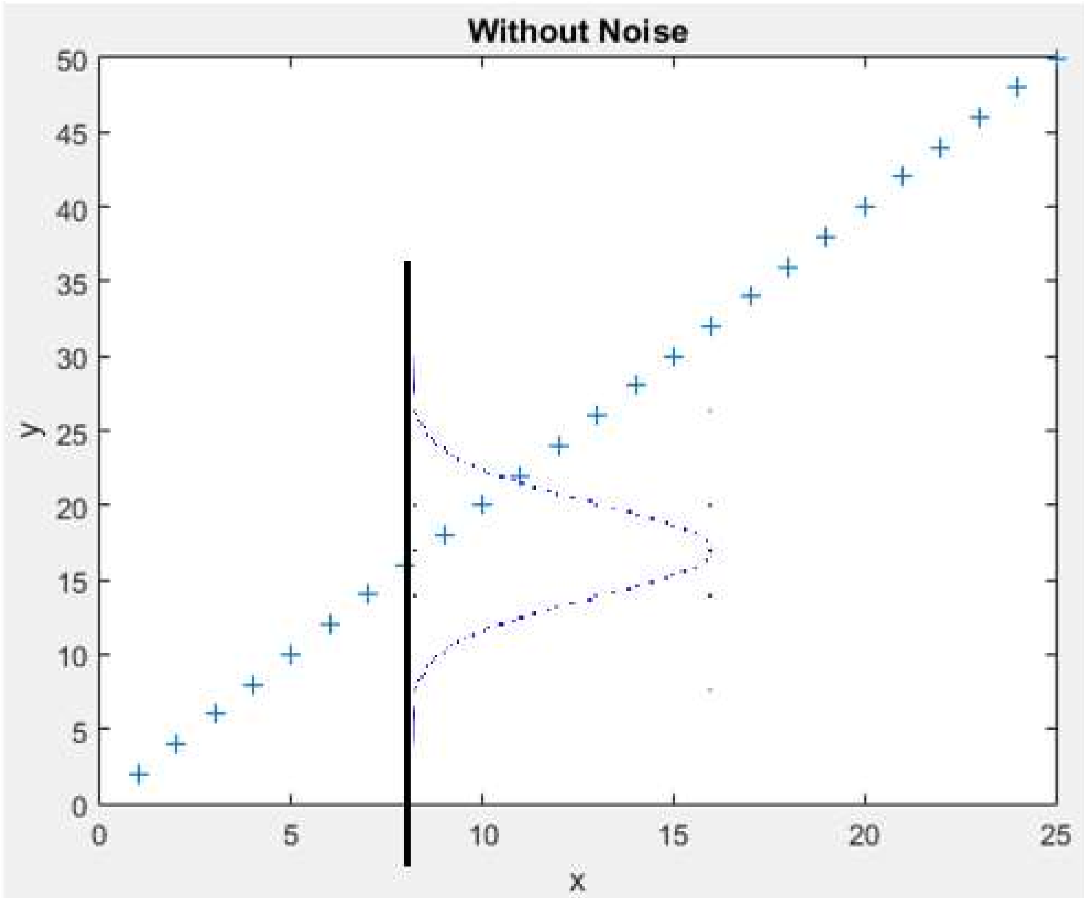
Data points generated: t=2x

X	y
1	2
2	4
3	6
4	8
5	10
6	12
7	14
8	16
9	18
10	20
11	22
12	24
13	26
14	28

x	y
15	30
16	32
17	34
18	36
19	38
20	40
21	42
22	44
23	46
24	48
25	50

t=2x points plotted

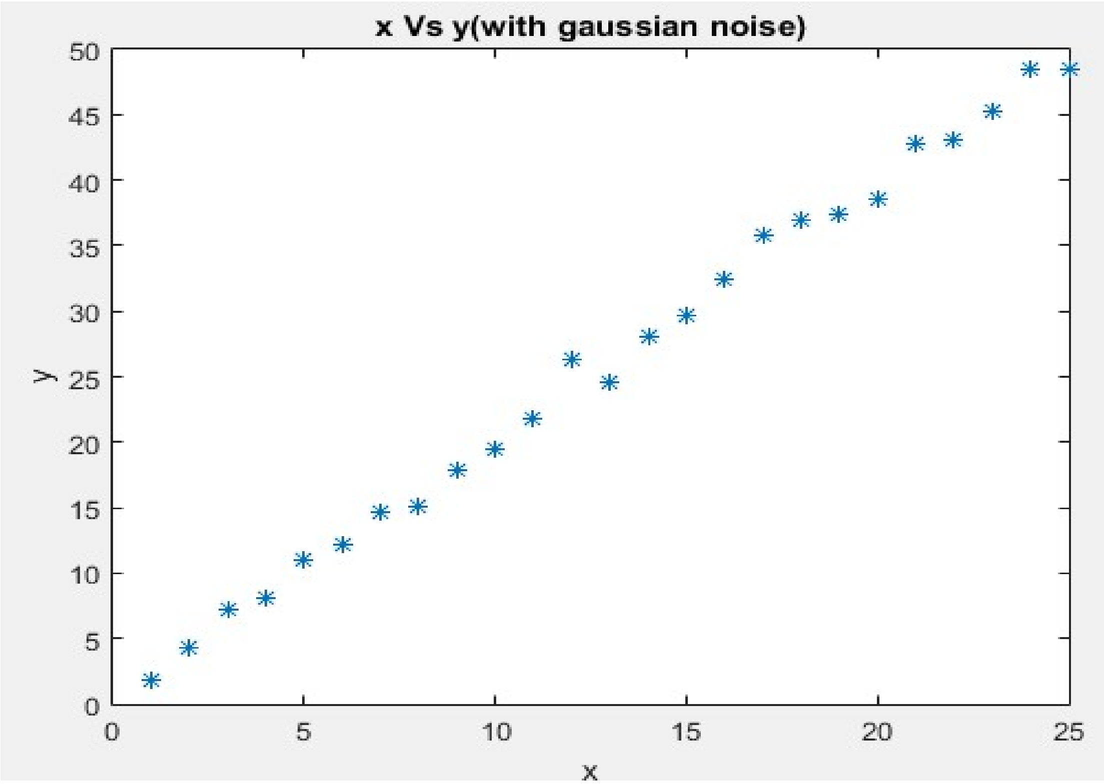




Without Noise
y
x
0
5
10
15
20
25
30
35
40
45
50
0
5
10
15
20
25

t=2x + Gaussian noise (σ=2)

All the points are disturbed



Algorithm

- We have decided to fit using y=m.x
- x=[---, ---, ---, ---, …., ---]
- t=[---, ---, ---, ---, …., ---]
- Choose m

Predict y=[---, ---, ---, ---, …., ---]

<u>Find out error</u>

e=[---, ---, ---, ---, …., ---]

<u>Generate squared error by squaring the elements of e</u>

se=[---, ---, ---, ---, …., ---]

<u>Compute the mean of se (MSE)</u>

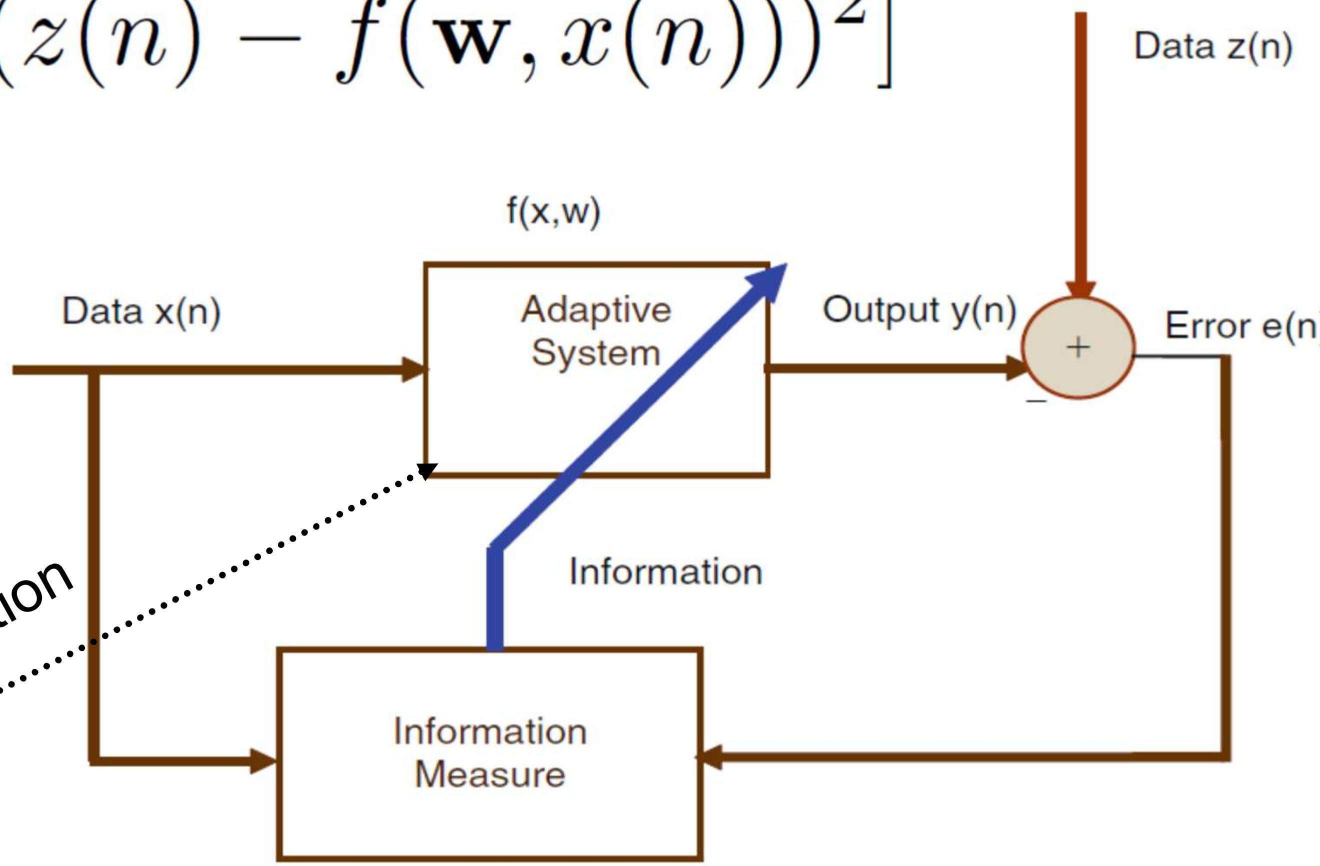
MSE = $\sum$ ***se***

- Change m; repeat 1 to 6
- Plot MSE versus m; choose the m corresponding to minimum MSE

MSE for a linear system

$$J_w(e(n)) = E[(z(n) - f(\mathbf{w}, x(n)))^2]$$

E refers to mean



Assume a linear function

$$y(n) = w.x(n)$$

m is replaced by w

Cost function: J

Cost function

Observation 1

1. Cost function is a function of error

Observation 2

- Error is [z- f(w,x)]

1. Error is function of expected output (z)
2. Error is function of input (x)
3. Error is function of weights (w)

Inference

- Cost function is a function of z, x and w

Make an adaptive system

- Cost function = function of z,x, &w
- x & z not in our control

Make expected output = predicted output

↓

Minimize error

↓

Minimize cost function

↓

Minimize cost function w.r.t weight

$$J_w(e(n)) = E[(z(n) - f(x(n), \mathbf{w}))^2]$$

$$\frac{\partial J(e(n))}{\partial \mathbf{w}} = 0$$

$$\boxed{\frac{\partial J(e(n))}{\partial \mathbf{w}} = \frac{\partial J(e(n))}{\partial e(n)} \frac{\partial e(n)}{\partial \mathbf{w}} = 0}$$

How error function changes w.r.t to error?

How error changes w.r.t to w?

$$\frac{\partial J(e(n))}{\partial \mathbf{w}} = E\left[\frac{\partial e^2(n)}{\partial e(n)} \frac{\partial e(n)}{\partial \mathbf{w}}\right] = -2E[e(n)\mathbf{x}(n)] = 0$$

MSE condition – linear system

x	y (generated)	z (desired or expected)	e = z-y	e.X
X1	Y1	Z1	e1	X1.e1
X2	Y2	Z2	e2	X2.e2
X3	Y3	Z3	e3	X3.e3
…	…	…	…	…
…	…	…	…	…
…	…	…	…	…
Xn	Yn	Zn	en	Xn.en

Expectation of last column should go to zero