

#"k" is considered as the time slots for the Mapper and Reducer Max value update
 $k = \{0, 40, 80, 120, 160, 200, 240\}$

#delta k is the different in the time frames between two readings

delta k = 40

#u(k) considered as percentage of job completed by Mapper divided by 10
 $uWave(k) = \{0.8, 2.2, 4.2, 6.5, 8.7, 10.0\}$

#p(k) considered as Job Execution Rate from the values

$pWave(k) = (uWave(k)/(\text{delta } k)) * 6$

$pWave(k) = \{0.1200, 0.3300, 0.6300, 0.9750, 1.3050, 1.5000\}$

#Linear Function: $p(k+1) = a * p(k) + b * u(k)$

Mean input and output:

uDash = .54

pDash = .81

Values of Si Estimates for the Data

$S = \{1.0112, 2.1443, 11.2305, 0.8111, 2.5900\}$

Parameter a and b calculation:

$a = (S(3)*S(4)-S(2)*S(5))/(S(1)*S(3)-(S(2))^2);$

$b = (S(1)*S(5)-S(2)*S(4))/(S(1)*S(3)-(S(2))^2);$

a = 0.5261

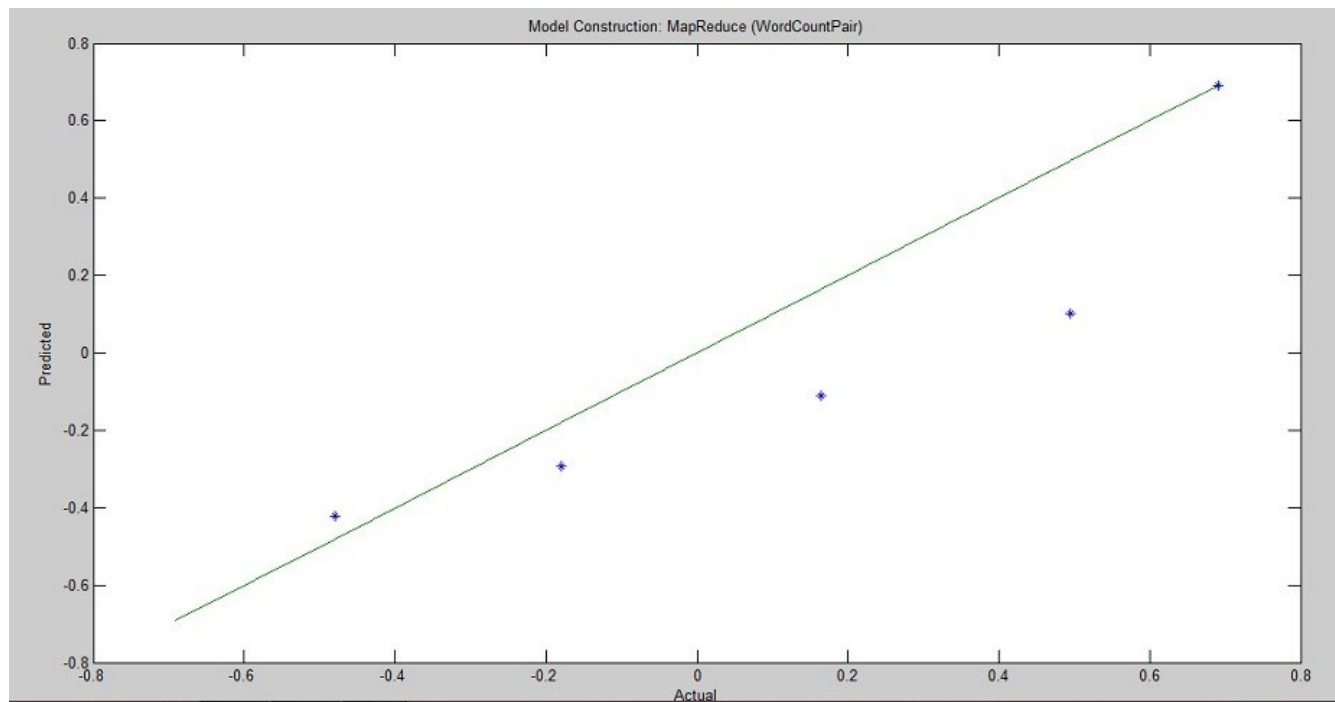
b = 0.1302

pHat =
-0.4229 -0.2942 -0.1103 0.1011 0.6900

Final Data Table:

k	uWave(k)	pWave(k)	u(k)	p(k)
40	.08	.12	-0.46	-0.69
80	2.2	.33	-0.32	-0.48
120	4.2	.63	-0.12	-0.18
160	6.5	.975	0.11	0.165
200	8.7	1.305	3.3	0.495
240	10.0	1.5	4.6	0.69

Final Graph:



RMSE, R², Correlation Coefficient:

- **RMSE:**

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k+1) - \hat{y}(k+1)]^2}$$

$$= \left\{ [(-0.0571)^2 + (0.1142)^2 + (0.0547)^2 + (0.3939)^2 + (0)^2] / 5 \right\}^{1/2}$$

$$= \left\{ [0.1745] / 5 \right\}^{1/2}$$

$$= \left\{ 0.0349 \right\}^{1/2}$$

$$= 0.1868$$

- **R²:**

$$R^2 = 1 - \frac{\text{var}(y - \hat{y})}{\text{var}(y)}$$

$$= 1 - (0.036839)/(0.29745)$$

$$= 1 - 0.123849386$$

$$= 0.876150614$$

- **Correlation Coefficient**

$$\text{CC} = \frac{\sum_k e(k)u(k)}{\sqrt{\text{var}(e(k))\text{var}(u(k))}}$$

$$= (0.018272 - 0.013704 + 0.006017 + 1.29987 + 0) / (0.036839 * 5.1795)^{1/2}$$

$$= (1.3106) / (0.1906)^{1/2}$$

$$= (1.3106 / 0.4366)$$

$$= 3.0018$$