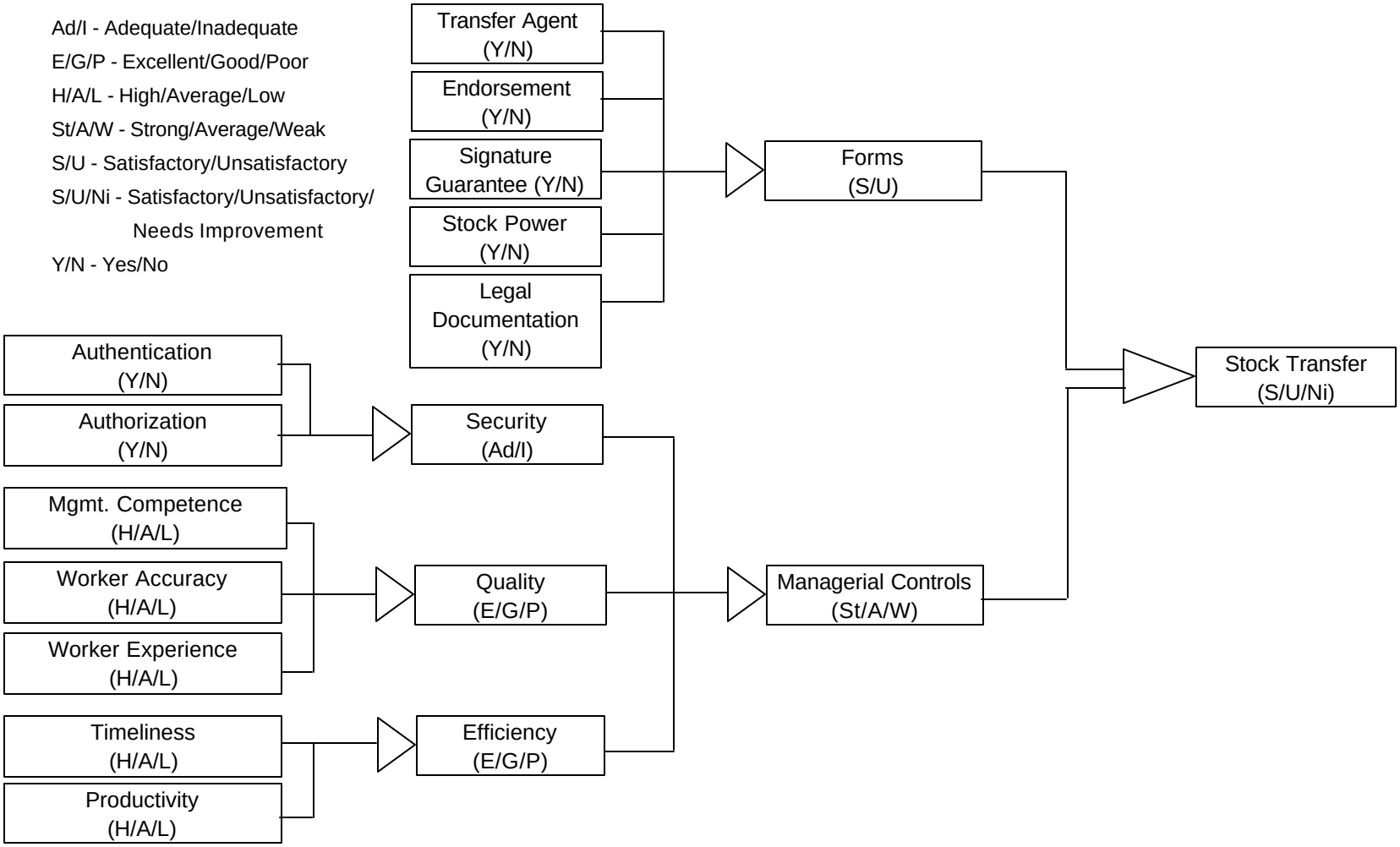
DISS 710 - Decision Support Systems: Assignment Three

by
Ronald G. Wolak
wolakron@nova.edu
810-944-4775

Problem 1 - Decision Tables and Dependency Diagrams

1a. Construct a dependency diagram.



1b. Construct a decision table to represent how Quality is determined.

Rules	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Mgmt. Competence	L	-	-	H	H	H	A	H	A	A	A
Worker Accuracy	-	L	-	H	H	A	H	A	H	A	A
Worker Experience	-	-	L	H	A	H	H	A	A	H	A
Quality	*P*	*P*	*P*	*E*	*E*	*E*	*E*	*G*	*G*	*G*	*G*

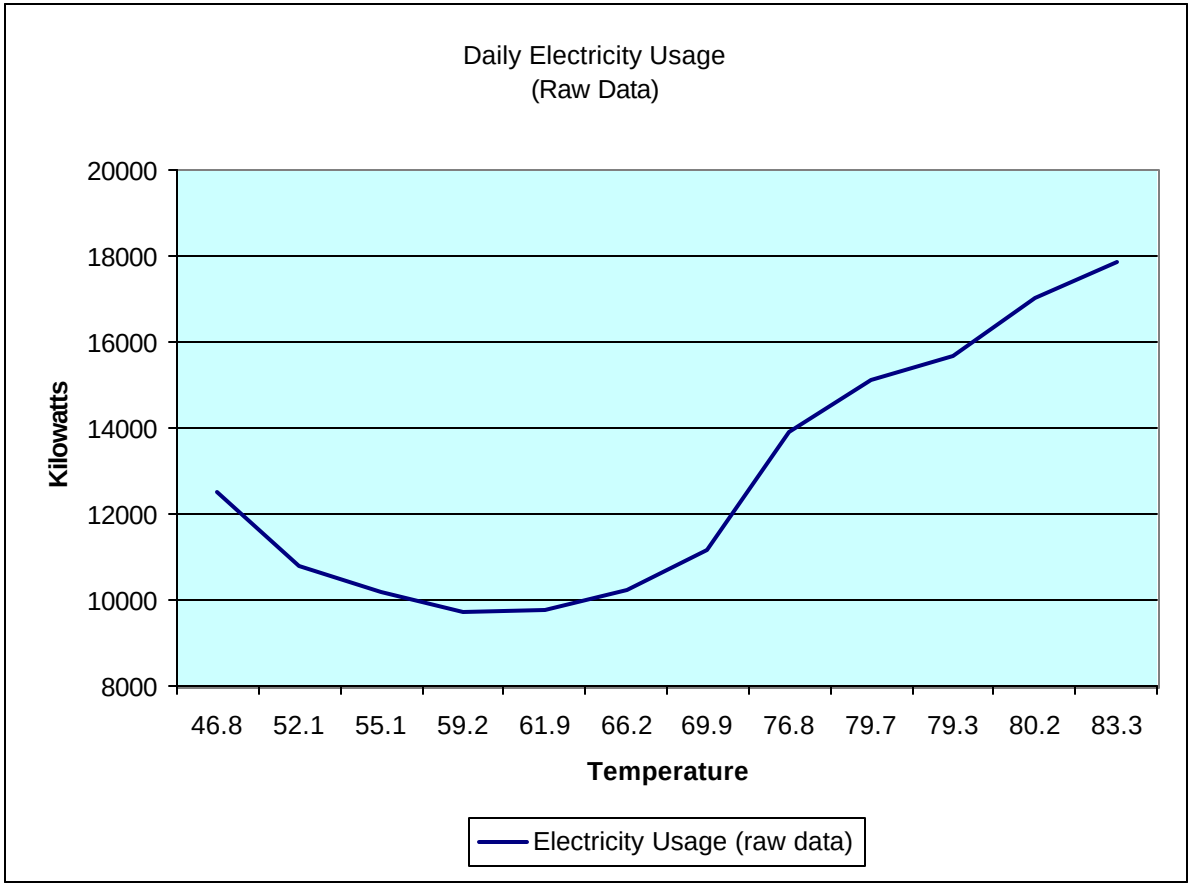
1c. Represent the decision table as IF … THEN rules.

R1.	IF Mgmt. Competence = L	THEN Quality = P
R2.	IF Worker Accuracy = L	THEN Quality = P
R3.	IF Worker Experience = L	THEN Quality = P
R4.	IF Mgmt. Competence = H AND Worker Accuracy = H AND Worker Experience = H	THEN Quality = E
R5.	IF Mgmt. Competence = H AND Worker Accuracy = H AND Worker Experience = A	THEN Quality = E
R6.	IF Mgmt. Competence = H AND Worker Accuracy = A AND Worker Experience = H	THEN Quality = E
R7.	IF Mgmt. Competence = A AND Worker Accuracy = H AND Worker Experience = H	THEN Quality = E
R8.	IF Mgmt. Competence = H AND Worker Accuracy = A AND Worker Experience = A	THEN Quality = G
R9.	IF Mgmt. Competence = A AND Worker Accuracy = H AND Worker Experience = A	THEN Quality = G
R10.	IF Mgmt. Competence = A AND Worker Accuracy = A AND Worker Experience = H	THEN Quality = G
R11.	IF Mgmt. Competence = A AND Worker Accuracy = A AND Worker Experience = A	THEN Quality = G

Problem 2 - Neural Networks

2a. Plot the raw data

Temperature (X)	Kilowatts (Y)
46.8	12530
52.1	10800
55.1	10180
59.2	9730
61.9	9750
66.2	10230
69.9	11160
76.8	13910
79.7	15110
79.3	15690
80.2	17020
83.3	17880



What pattern do you see?

As the temperature goes above and below 59 degrees the power usage increases.

What do you think is really affecting electricity usage?

As the temperature drops below 59 degrees electricity usage rises due to a greater demand for heating.
As the temperature rises above 59 degrees electricity usage rises due to a greater demand for air conditioning

2b. Solve this problem with a linear regression (Excel Data Analysis - Regression)

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.776552467
R Square	**0.603033734**
Adjusted R Squa	0.563337107
Standard Error	1967.153189
Observations	12

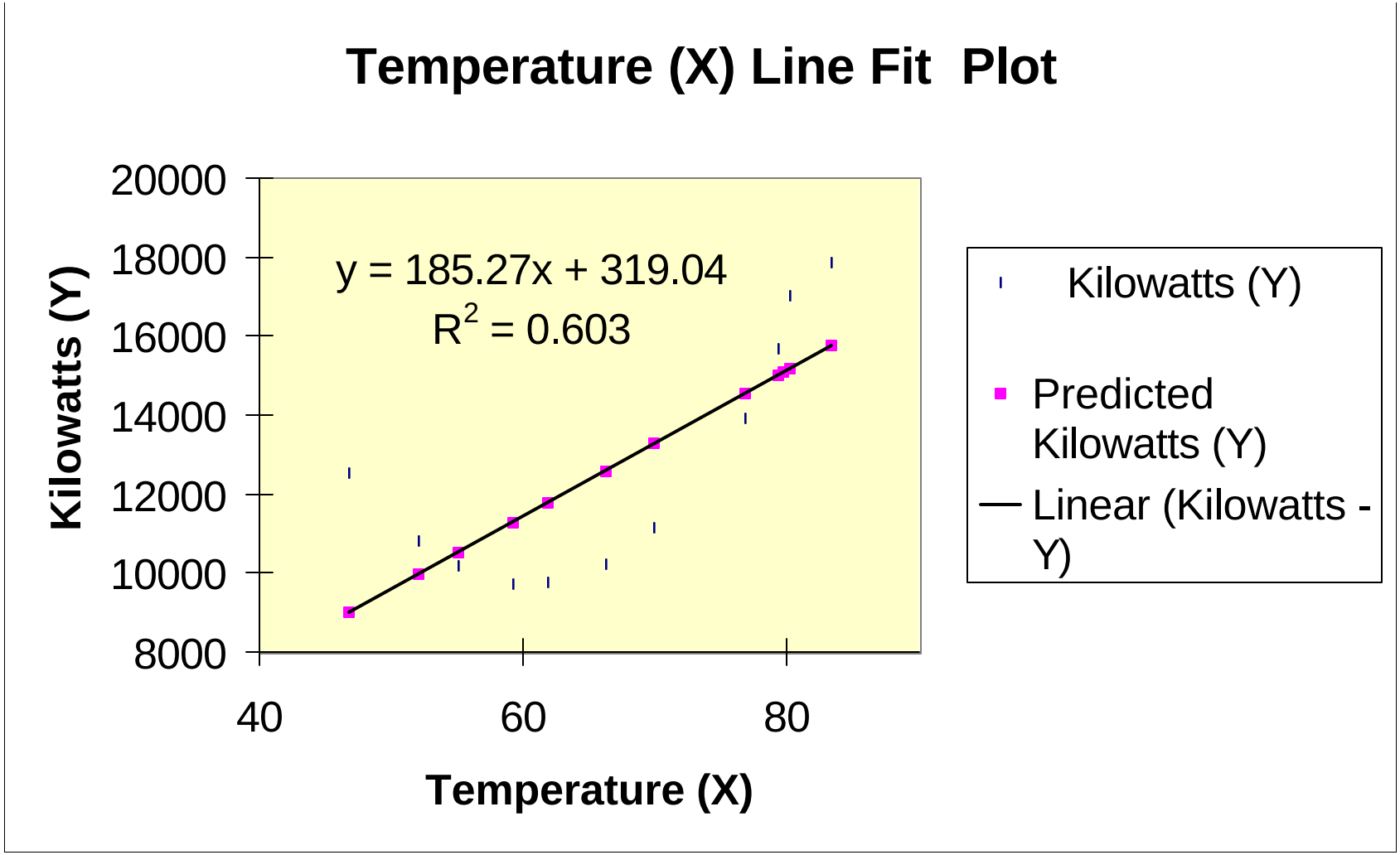
ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	58784708	58784708.3	15.19105741	0.0029727
Residual	10	**38696917**	3869691.67		
Total	11	97481625			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 98.0%*	*Upper 98.0%*
Intercept	319.0414124	3260.4128	0.09785307	0.923982528	-6945.612	7583.69513	-8691.9979	9330.0807
Temperature (X)	185.2702073	47.534791	3.89757071	0.002972726	79.356075	291.184339	53.89486	316.64555

RESIDUAL OUTPUT

Observation	*Predicted Kilowatts (Y)*	*Residuals*
1	8989.687116	3540.3129
2	9971.619215	828.38079
3	10527.42984	-347.42984
4	11287.03769	-1557.0377
5	11787.26725	-2037.2672
6	12583.92914	-2353.9291
7	13269.42891	-2109.4289
8	14547.79334	-637.79334
9	15085.07694	24.923062
10	15010.96885	679.03115
11	15177.71204	1842.288
12	15752.04968	2127.9503



What is wrong? Since R Square = .60 the linear regression is a poor fit to the raw data.
The sum of the squares error = Residual SS value = 38696917

2c. Solve this problem with nonlinear regression (quadratic function).

Temperature (X)	X Squared	Kilowatts (Y)
46.8	2190.2	12530

SUMMARY OUTPUT

RESIDUAL OUTPUT

52.1	2714.4	10800
55.1	3036.0	10180
59.2	3504.6	9730
61.9	3831.6	9750
66.2	4382.4	10230
69.9	4886.0	11160
76.8	5898.2	13910
79.7	6352.1	15110
79.3	6288.5	15690
80.2	6432.0	17020
83.3	6938.9	17880

Regression Statistics	
Multiple R	0.9923059
R Square	**0.984671**
Adjusted R Squa	0.9812646
Standard Error	407.47115
Observations	12

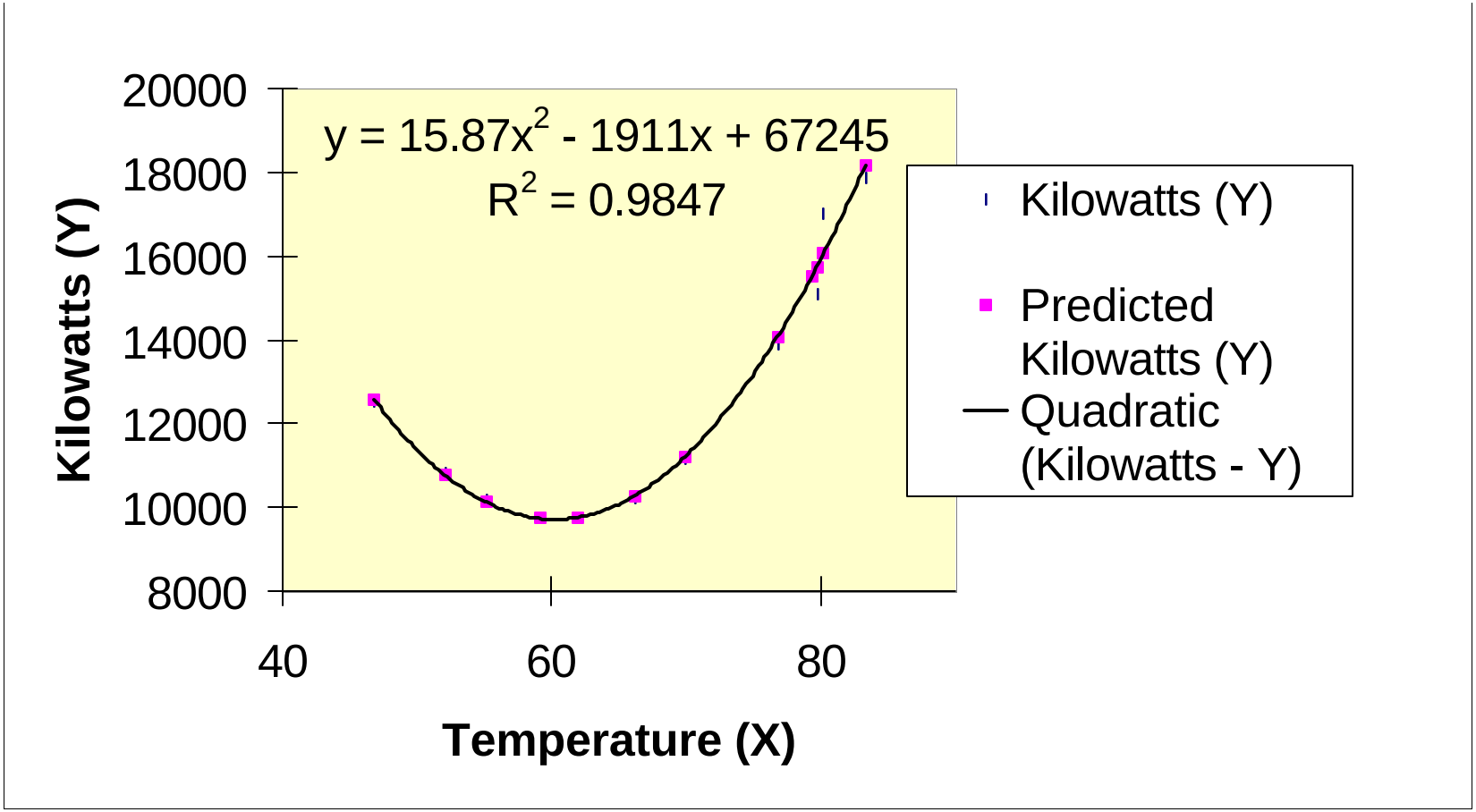
Observation	*Predicted Kilowatts (Y)*	*Residuals*
1	12567.751	-37.751223
2	10757.845	42.154797
3	10128.534	51.465861
4	9730.4526	-0.4526034
5	9759.6753	-9.6753168
6	10283.903	-53.903454
7	11204.737	-44.736717
8	14082.701	-172.70118
9	15743.307	-633.30703
10	15498.388	191.61209
11	16056.597	963.40254
12	18176.108	-296.10777

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	95987330	47993665.2	289.061457	6.836E-09
Residual	9	**1494294.6**	166032.738		
Total	11	97481625			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 98.0%*	*Upper 98.0%*
Intercept	67245.16853	4521.7304	14.8715562	1.21556E-07	57016.296	77474.0412	54487.402	80002.935
Temperature (X)	-1911.038841	140.39	-13.612358	2.61324E-07	-2228.623	-1593.4544	-2307.14	-1514.938
X Squared	15.87004186	1.0602009	14.9689008	1.14862E-07	13.471699	18.2683847	12.878755	18.861329

Temperature (X) Line Fit Plot



Did the method work well? Yes, because the relationship between the dependent and independent variables is quadratic rather than linear. The R Square is .98 as compared to .60
The sum of the squares error = Residual SS value = 1494295

2e. Build a neural network to solve the problem.

- The neural network software used in this problem was NNModel32 version 1.0.2.0
- The network used the back error propagation algorithm and consisted of an input layer, hidden layer, and output layer.
- 100,000 iterations (determined by the max count parameter) were used to train the model and the time taken was 9 seconds.

Parameters used:

Max Hidden Neurons	4	Learning Rate	0.75
Eon	100	Hlearning Rate	1.5
Max Train Count	100000	Tlearning Rate	0.75
Tolerance	0.05	IO Learning Rate	0.1
Error Tolerance	0.001	Alpha	0.8
Good RSQ	0.9	Theta	0.5
Signif Increase	0.05	Random Fact	0.5
No Signif Increase	0.005	InRandom Fact	0
Hidden Freeze	0.75	Seed	15
Best Eon	0	CG Max Iterations	1

Raw Data

(used for training)

Temp (X)	Kilowatts (Y)
46.8	12530
52.1	10800
55.1	10180

Test Values

(generated from quadratic equation above)

Temp (X)	Kilowatts (Y)
40	16197
41	15571.5
42	14977.7

Temp (X)	Kilowatts (Y)
66	10248.7
67	10448.4
68	10679.9

59.2	9730
61.9	9750
66.2	10230
69.9	11160
76.8	13910
79.7	15110
79.3	15690
80.2	17020
83.3	17880

43	14415.6
44	13885.3
45	13386.8
46	12919.9
47	12484.8
48	12081.5
49	11709.9
50	11370
51	11061.9
52	10785.5
53	10540.8
54	10327.9
55	10146.8
56	9997.32
57	9879.63
58	9793.68
59	9739.47
60	9717
61	9726.27
62	9767.28
63	9840.03
64	9944.52
65	10080.8

69	10943.1
70	11238
71	11564.7
72	11923.1
73	12313.2
74	12735.1
75	13188.8
76	13674.1
77	14191.2
78	14740.1
79	15320.7
80	15933
81	16577.1
82	17252.9
83	17960.4
84	18699.7
85	19470.8
86	20273.5
87	21108
88	21974.3
89	22872.3
90	23802

Back-Error Propagation Model raw_test.bep
Inputs / Hidden / Outputs: 1 / 4 / 1
Input Labels:
V1
Output Labels:
V2
Parent is: RAW_TEST
Training Matrix: 12

Test Matrix: 51
Max Hidden: 4
Trained to: 100000
Max Count: 100000
Sum Sq Error: 0.000211
Training Time 0 00:00:09

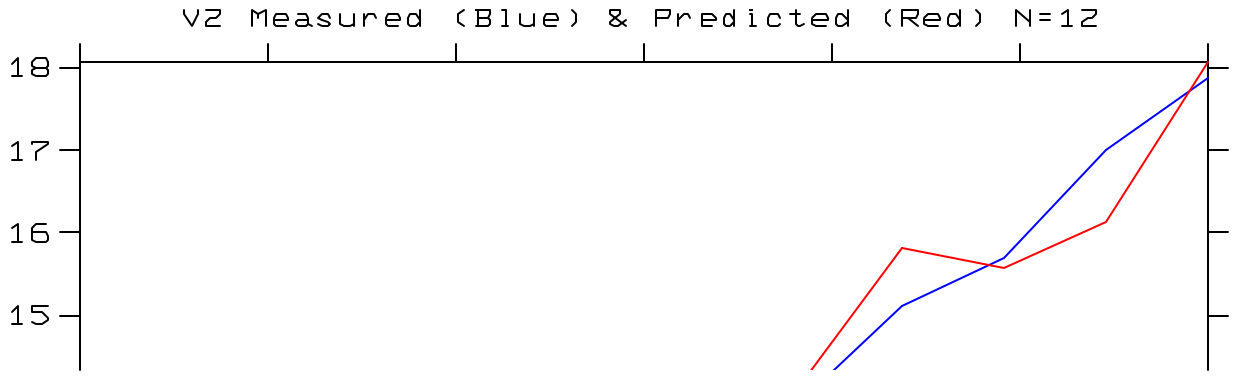
Input to hidden layer weights

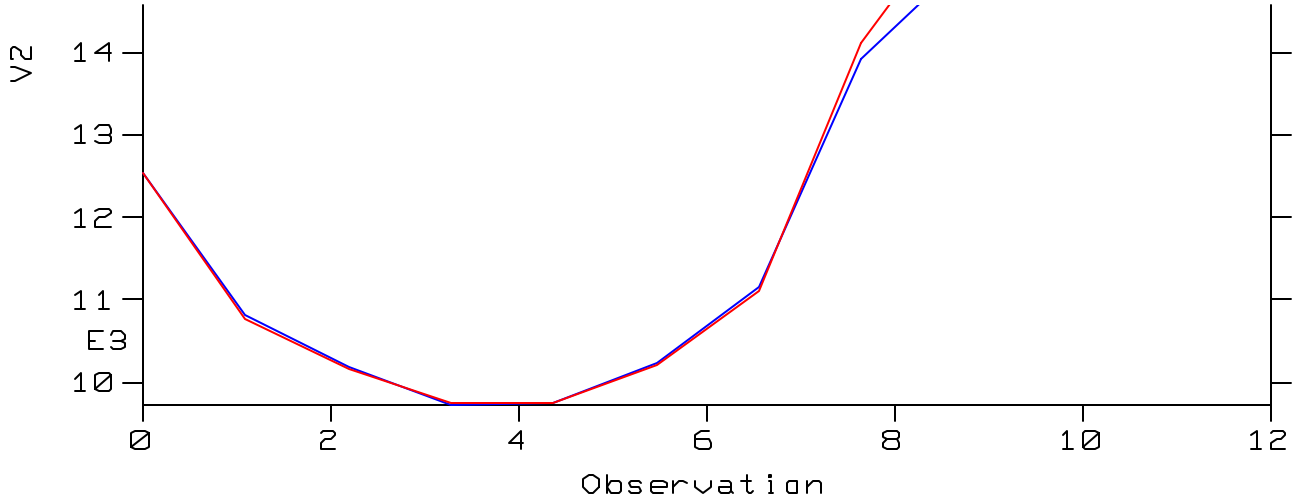
	HID001	HID002	HID003	HID004				
V1	-5.84E+00	-6.72E+00	-7.40E-01	1.06E+00				
Theta	3.71E+00	1.98E+00	2.12E-01	-3.42E-01	TLearn	1.17E-02	4.69E-02	1.88E-01
CircW	-3.26E-01	-8.29E-02	-3.66E-01	-1.10E-01				

Hidden to output layer weights

	V2				
HID001	-8.25E+00				
HID002	6.85E+00				
HID003	3.70E-01				
HID004	1.93E-01				
Theta	2.71E+00	TLearn	7.50E-01	Learn	7.50E-01

Results Using Training (Raw) Data



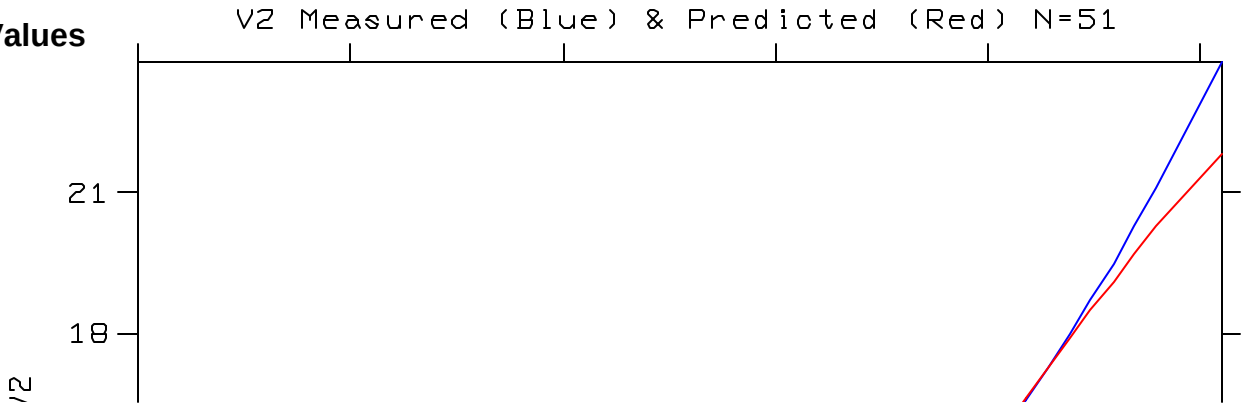


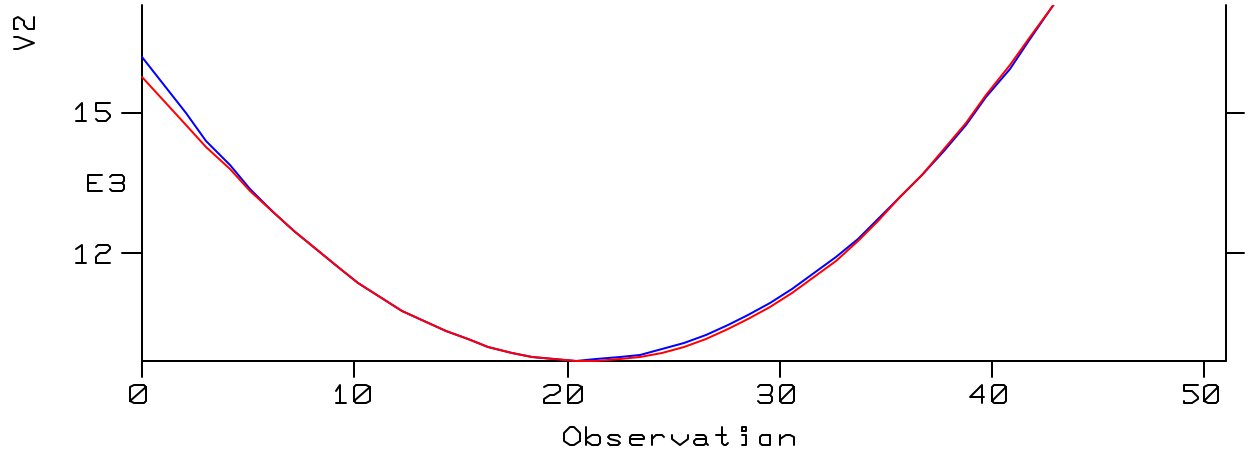
MODEL raw_test.bep
Model Performance Statistics
(Training Matrix) # of Points = 12
11-Jun-00

Variable	Mean	Std Dev	Minimum	Maximum	Sum Sq
V1	67.541667	12.477577	46.799999	83.300003	1712.5893
V2					
Measured	12832.5	2976.9054	9730	17880	97481625
Predicted	12831.939	2971.4665	9738.8183	18077.251	97125747
Residual	0.560384	356.31492	-714.5693	886.18847	**1396563.6**
R Square	**0.985674**				

95.0% Confidence Intervals	Tail	Cover	Intervals
	0.338681	0.340998	-714.569336,886.188477

Results of Trained NM using Test Values





MODEL raw_test.bep
Model Performance Statistics
(Test Matrix) # of Points = 51
11-Jun-00

Variable	Mean	Std Dev	Minimum	Maximum	Sum Sq
V1	65	14.866069	40	90	11050
V2					
Measured	13519.25	3840.4908	9717	23802	737468515
Predicted	13350.235	3549.3019	9716.6562	21817.238	629877211
Residual	169.01485	396.67391	-76.7705	1984.7617	**7867509.7**
R Square	**0.989332**				

95.0% Confidence Intervals	Tail	Cover	Intervals
	0.089672	0.969077	-76.770508,1984.761719

How well did this work?

Very well, however when the max count was set to only 100 the results were disappointing.
The R Square for the raw data was only .68 & the test data was worse at .45 .
When the max count was increased to 100,000 the R Square increased to .99 for both.

2f. Which method worked the best?

Comparing the R Square values from each of the methods:

Linear Regression 0.6

Nonlinear Regression	0.98
Neural Network	0.99

Neural Network worked slightly better than Nonlinear (quadratic) regression.

Other Neural Network Programs Investigated were:

AiNet
ANN-PRO
Back Propagation Neural Network
EasyNN
Hybid Neural Networks
NeuroOffice
NeuralSolutions
Pythia
Qnet 2000
QwikNet
WANND

I attempted to use EasyNN, NeuralSolutions and Back Propagation Neural Network but found these programs less intuitive than NNModel32. The above program links are also available at:

http://www.scisstudyguides.addr.com/decision_support_study_guide.htm