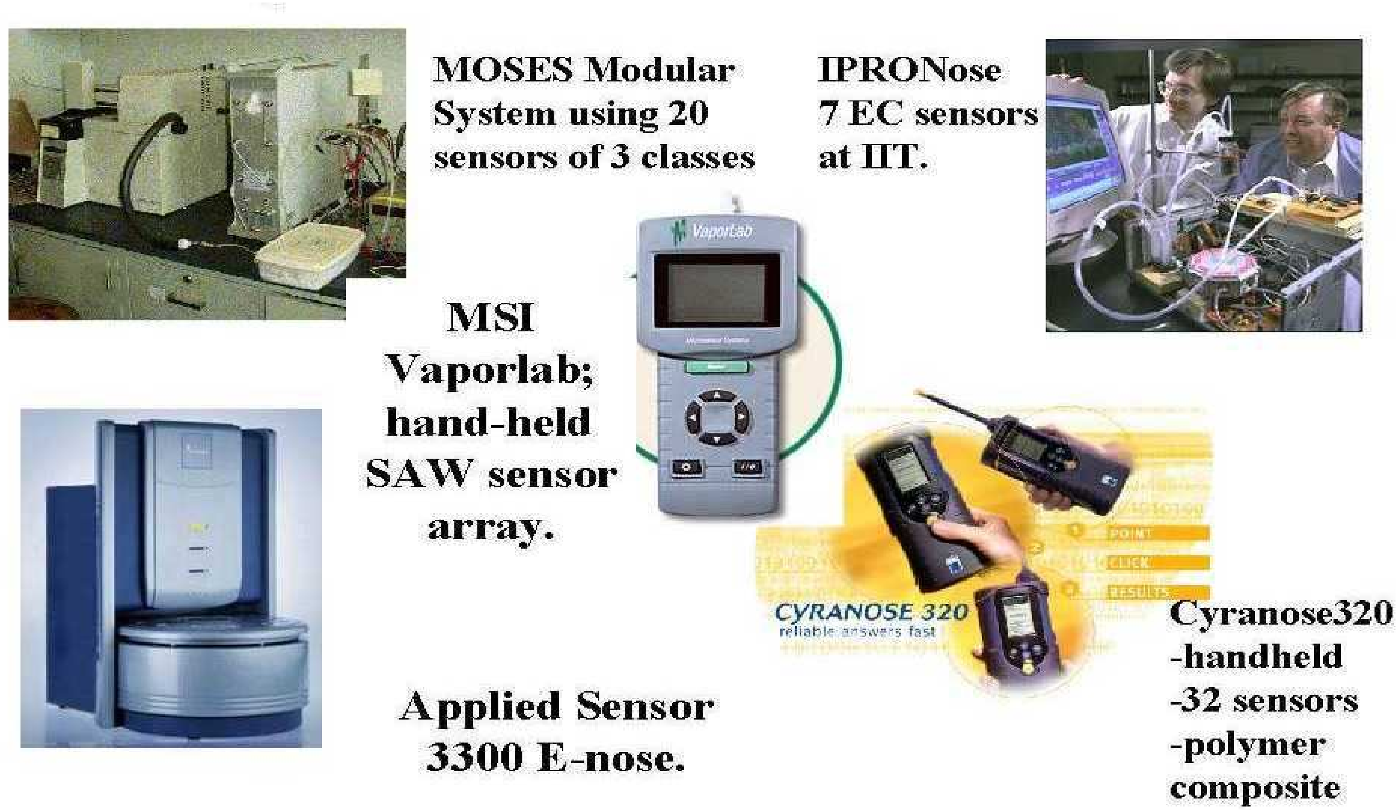
Artificial Nose Technology: The WI -Nose

A Profitability and Market Analysis for the Development of Artificial Nose Technology to Monitor the Fermentation Process in Wine

Shawna M. Linehan
Sarosh N. Nizami

What is an E-Nose?

- An artificial smelling device that identifies the specific components of an odor and analyzes its chemical makeup to identify it

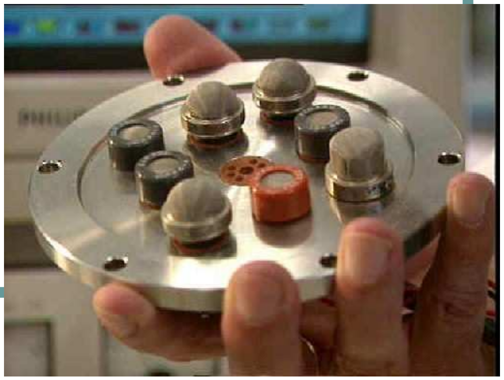


What Is It Made Of?

- Electronic Olfactory System: looks nothing like an actual nose but works similar to one
- Two main components
  - Chemical Sensing System
    1. Acts like receptors in our nasal passages
    2. Odor-reactive sensor array
  - Automated Pattern Recognition System
    1. Acts like our brain
    2. Artificial Neural Networks (ANN)

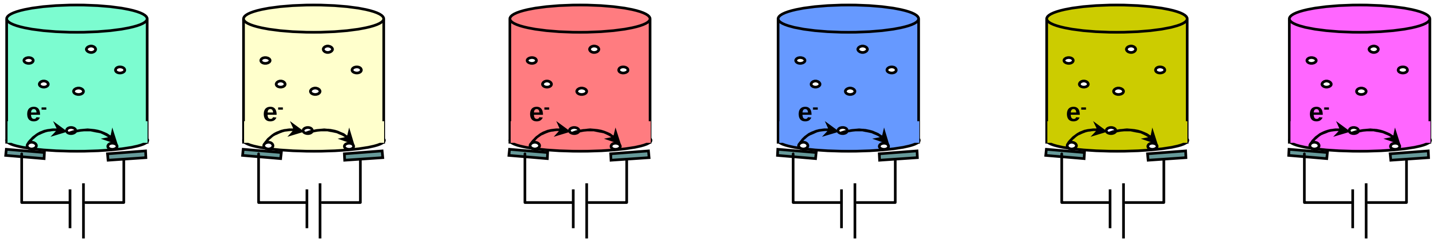
How Does An E- Nose Work?

- The sensor array generally consists of different polymer films, which are specially designed to conduct electricity.
- When a substance is absorbed into these films, the films expand slightly, and that changes how much electricity they conduct.
- Each electrode reacts to particular substances by changing its electrical resistance in a characteristic way.



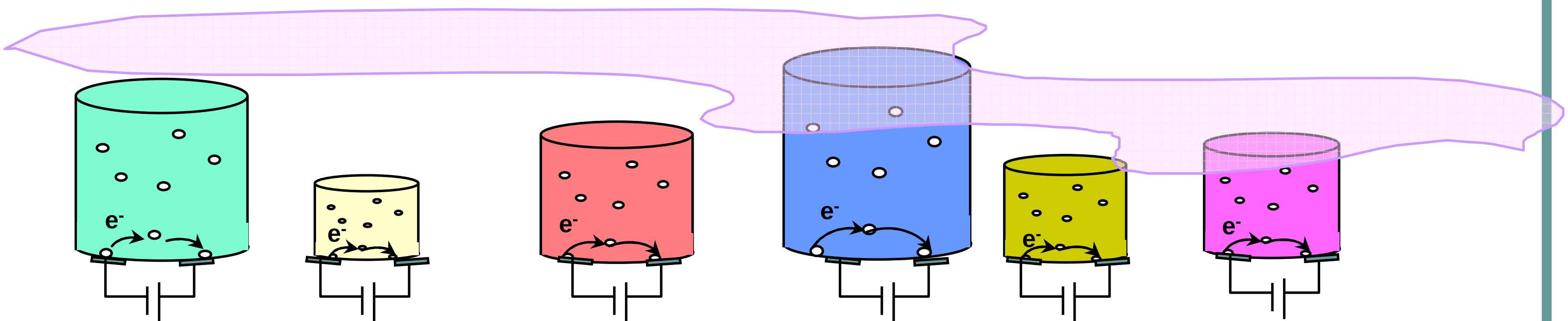
Baseline Resistance

All of the polymer films on a set of electrodes (sensors) start out at a measured resistance, their *baseline resistance*. If there has been no change in the composition of the air, the films stay at the baseline resistance and the percent change is zero.

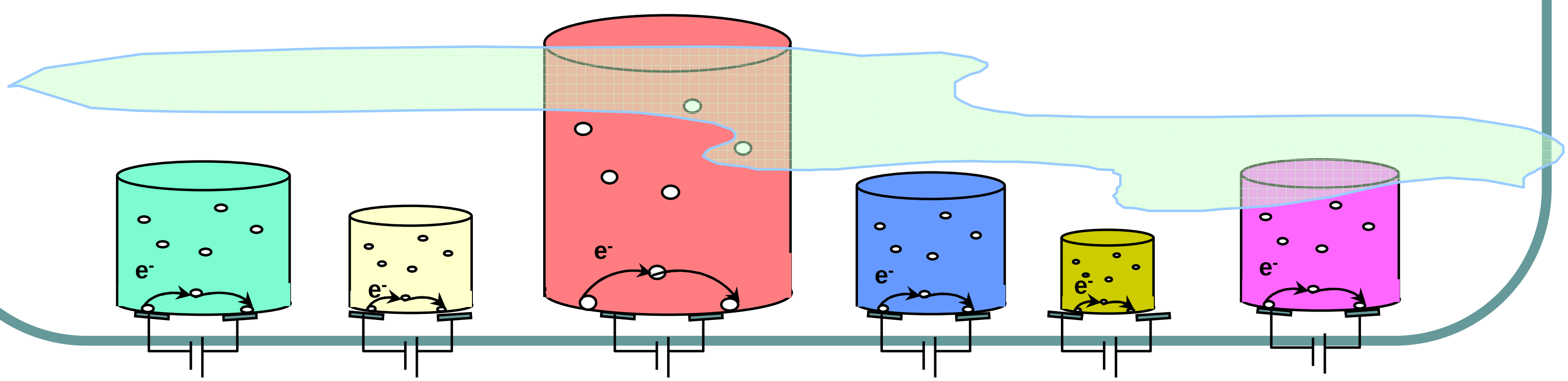


The E-Nose Smells Something

Each polymer changes its size, and therefore its resistance, by a different amount, making a pattern of the change



If a different compound had caused the air to change, the pattern of the polymer films' change would have been different:

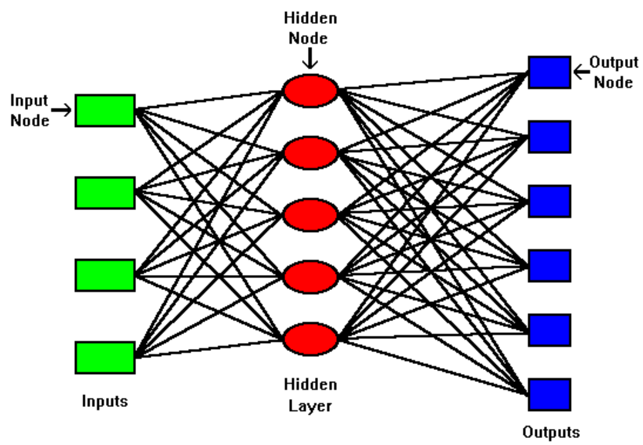


“Smell-Prints”

- Each chemical vapor presented to a sensor array produces a pattern characteristic of the vapor.
- By presenting many different chemicals to the sensor array, a database of signatures is built up which is then used to train the pattern recognition system.
- Combining the signals from all the electrodes gives a "smell-print" of the chemicals in the mixture that neural network software can learn to recognize.

Artificial Neural Networks (ANN)

- An information processing system
- Collections of mathematical models
- Learning typically occurs by example – through exposure to a set of input-output data



Why use an ANN?

- Well suited to pattern recognition and forecasting.
  - Like people, learn by example.
  - Can configure, through a learning process, for specific applications, such as identifying a chemical vapor.
- Capability not affected by subjective factors such as working conditions and emotional state.

Global Markets

- Companies have taken the E-Nose technology and expanded to various markets:
  - Cyrano Sciences (Pasadena, California)
  - Neotronics (Essex, England)
  - Alpha MOS (Toulouse, France)
  - Bloodhound Sensors (Leeds, England)
  - Aroma Scan (Manchester, England)
  - Illumina (Cambridge, Massachusetts)
  - Smart Nose (Zurich, Switzerland)

Applications: NASA

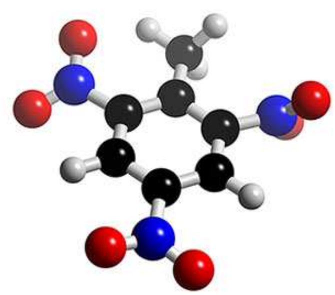
- NASA started the E-Nose Project to detect leaked ammonia onboard space station.
- Ammonia is just one of about 40 - 50 compounds necessary on the space station which humans can't sense until concentrations become dangerously high.

Current Applications: Environmental Monitoring

- Environmental applications include:
  - analysis of fuel mixtures
  - detection of oil leaks
  - testing ground water for odors
  - identification of household odors
  - identification of toxic wastes
  - air quality monitoring
  - monitoring factory emissions
  - check for gas buildups in offshore oil rigs
  - check if poisonous gases have collected down in sewers

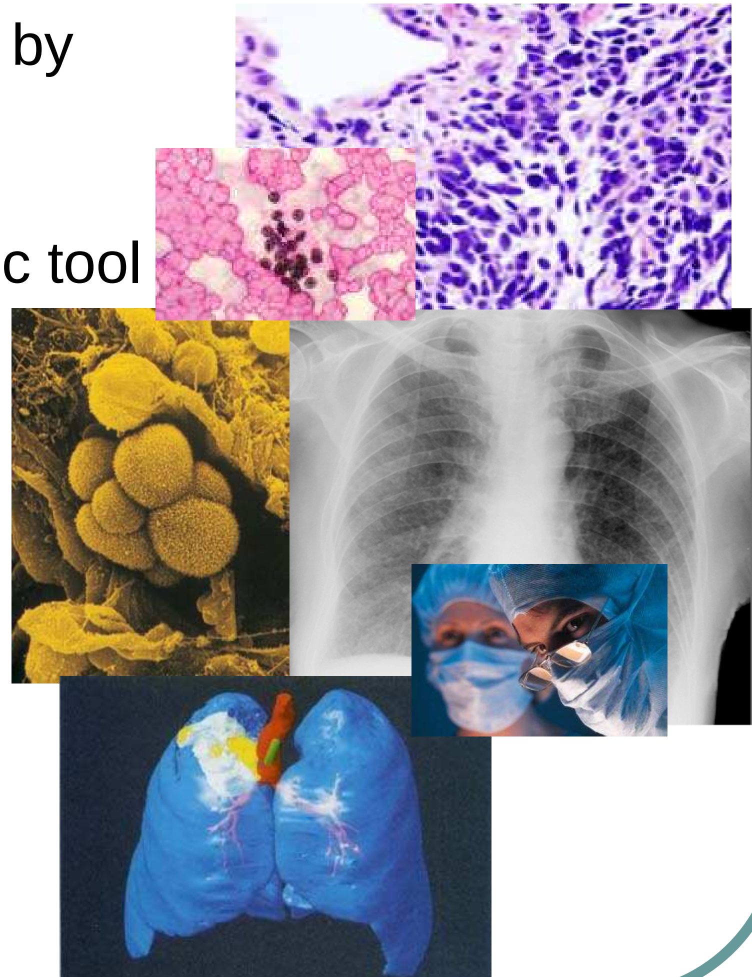
Current Applications: Explosives Detection

- Detection of bombs, landmines, TNT, and other explosive devices.
- Specific Applications:
  - Homeland Security
  - Airport security
  - Military
  - Battlefields



Current Applications: Medical Diagnostics

- Detecting diseases and disorders by odor
- Relatively new technology
- Provides a non-invasive diagnostic tool
- Potential applications include:
  - Detecting bacterial infections as well as type and severity of cancer, specifically lung cancer
  - Diagnosing gastrointestinal disorders, diabetes, liver problems, and diseases such as Tuberculosis.



Current Applications:
The Food Industry

- Assessment in food production
- Inspection of food quality
- Control of food cooking processes

- Specific applications include:
  - Inspection of seafood products
  - Grading whiskey
  - Wine testing
  - Inspection of cheese composition
  - Monitoring fermentation process

Fermentation In Wine

- Fermentation in wine is the process where yeast convert sugar into carbon dioxide and ethyl alcohol.
  - $C_6H_{12}O_6$ ---> $2CO_2 + 2C_2H_5OH$

- Three Stages of Wine Fermentation
  - Primary or Aerobic Fermentation
  - Secondary or Anaerobic Fermentation
  - Malo-Lactic Fermentation (possible 3rd stage)

Primary or Aerobic Fermentation

- Typically lasts for the first 4-7 days
- On average, 70% of fermentation activity will occur during these first few days.
- Carbon dioxide, produced by yeast, leaves the solution in the gaseous form, while the alcohol is retained in mix.
- Critical stage for yeast reproduction

Secondary or Anaerobic Fermentation

- Remaining 30% of fermentation activity will occur
- Usually lasts anywhere from 2-3 weeks to a few months, depending on available nutrients and sugars.
- Should take place in a fermentation vessel fitted with an airlock to protect the wine from oxidation

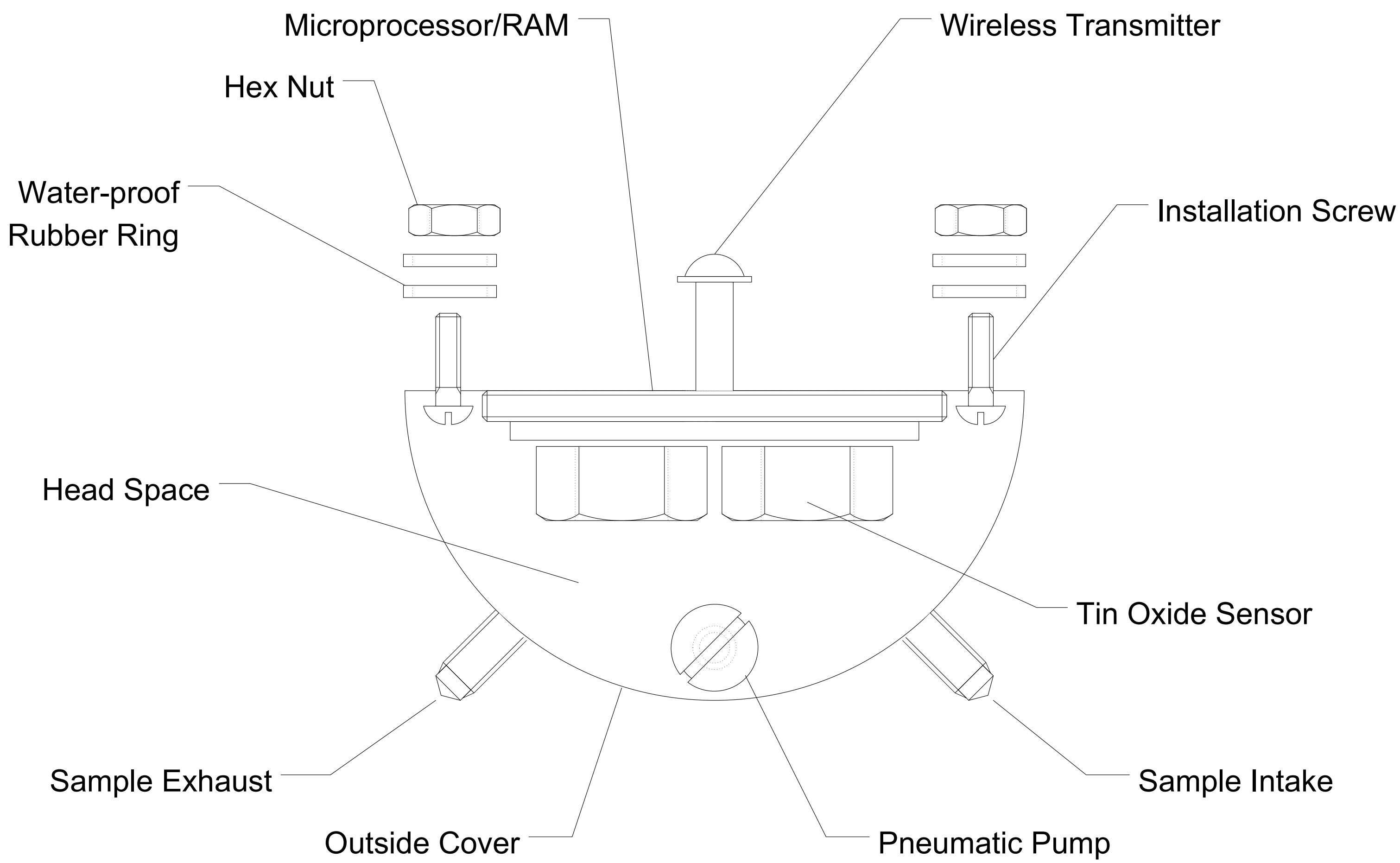
Malo-lactic Fermentation (Possible 3rd stage)

- A continuation of fermentation in the bottle is to be avoided
  - Can result in a buildup of carbon dioxide which can cause bottles to burst
  - Often results in a semi carbonated wine that does not taste good.
- If initiated pre-bottling, results in a softer tasting product
  - Is often induced after secondary fermentation by inoculating with lactobacilli to convert malic acid to lactic acid
  - Lactic acid has approximately half the acidity of malic acid, resulting in a less acidic wine with a much cleaner, fresher flavor.

Why Is It Important to Monitor the Fermentation Process in Wine?

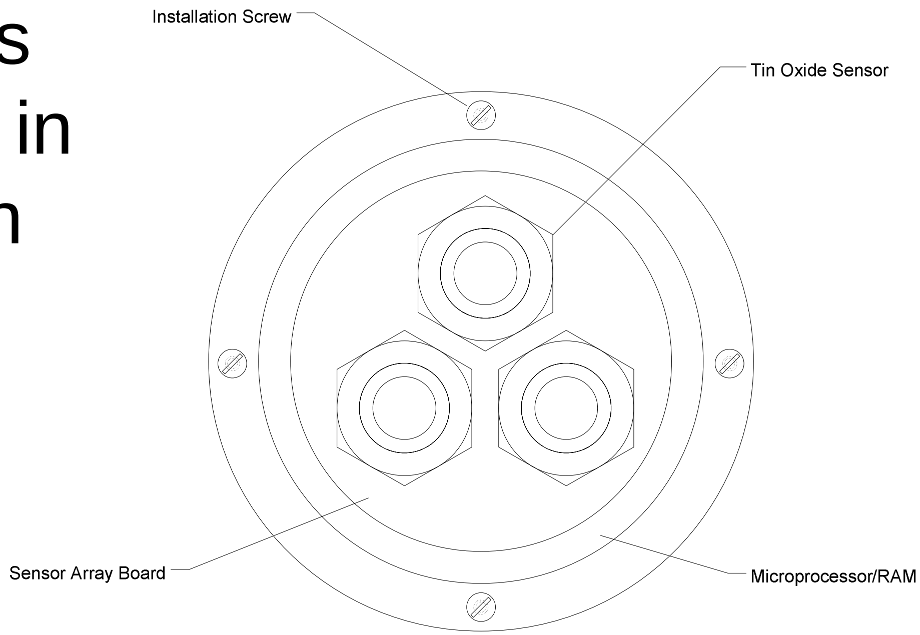
- The wine industry needs to know the stage of their products in order to:
  - Precisely induce Malo-lactic fermentation
  - Add rock sugar and additional yeast needed to produce champagne and sparkling wines
  - Bottle batches of champagne and sparkling wine
  - Add additional nutrients and/or yeast enabling products
  - Add acidity to the wine

Design: Wi-Nose (Cross-section)



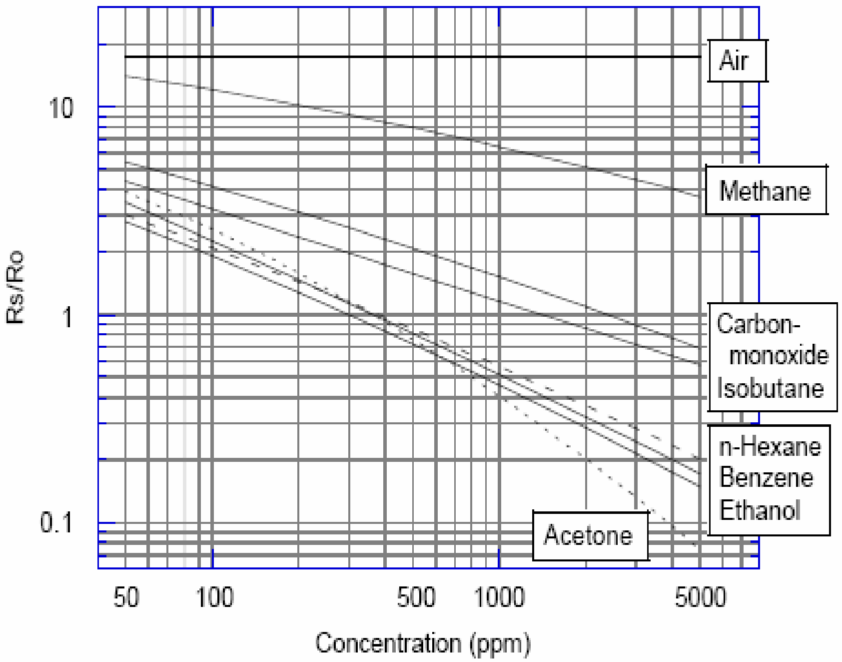
Design: Wi-Nose (Top View)

- Most of these units are to be installed in metal fermentation vats
  - Reduce Rusting
    - Rubber O-Rings
  - Avoid Moister Contact
    - Unique hemisphere design

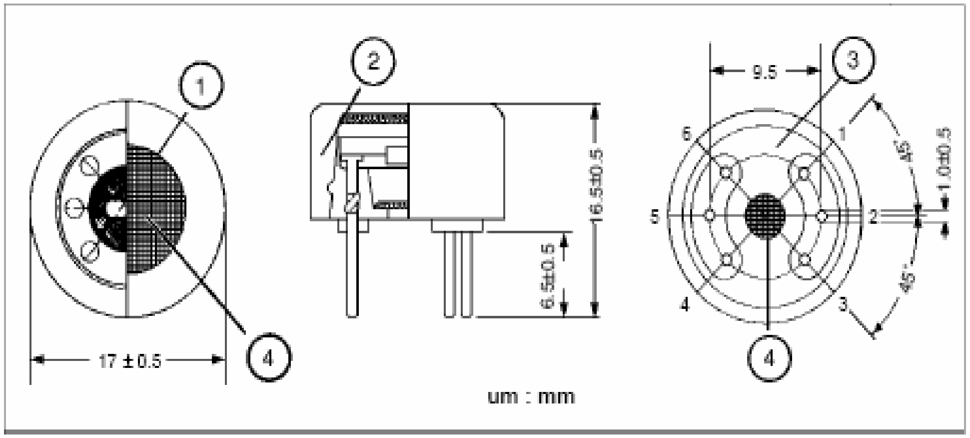


Choice of Sensors: TGS 822

- High sensitivity to organic solvent vapors such as ethanol
- Is not responsive to carbon dioxide
- High stability and reliability over a long period (lifetime ≥ 5 years, up to 200 ºC)
- Long life and low cost



Choice of Sensors: TGS 822



(1) Sensing Element:
SnO_2 is sintered to form a thick film on the surface of an alumina ceramic tube which contains an internal heater.

(2) Cap:
Nylon 66

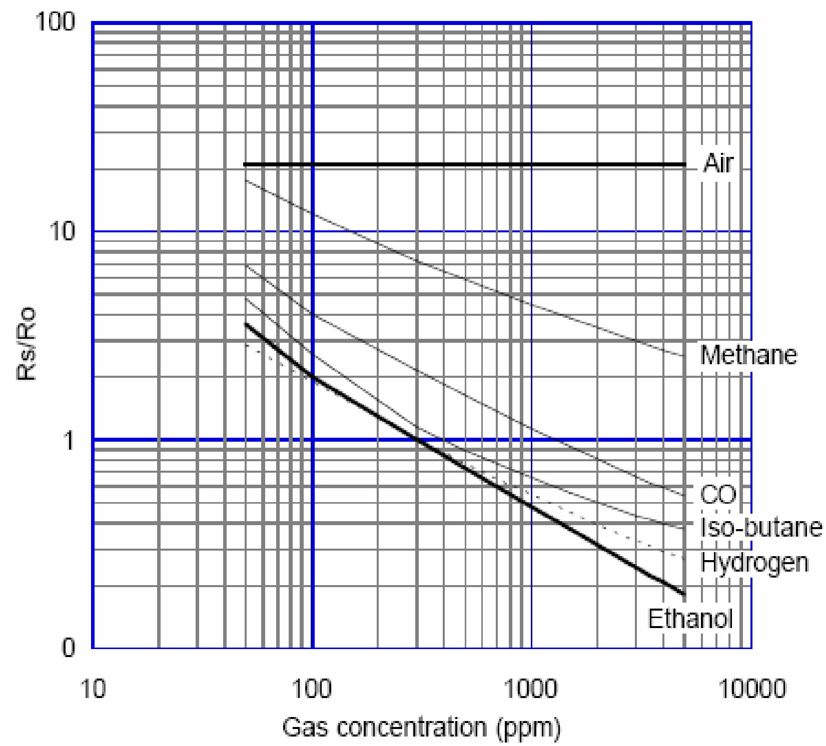
(3) Sensor Base:
Nylon 66

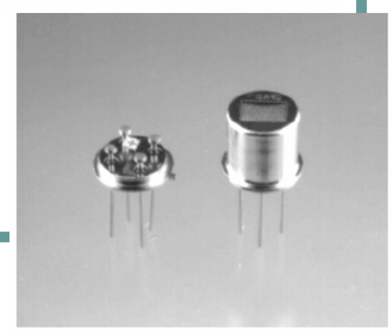
(4) Flame Arrestor:
100 mesh SUS 316 double gauze

- Uses simple electrical current to produce a resistance output in response to a detectable gas's concentration (ppm)

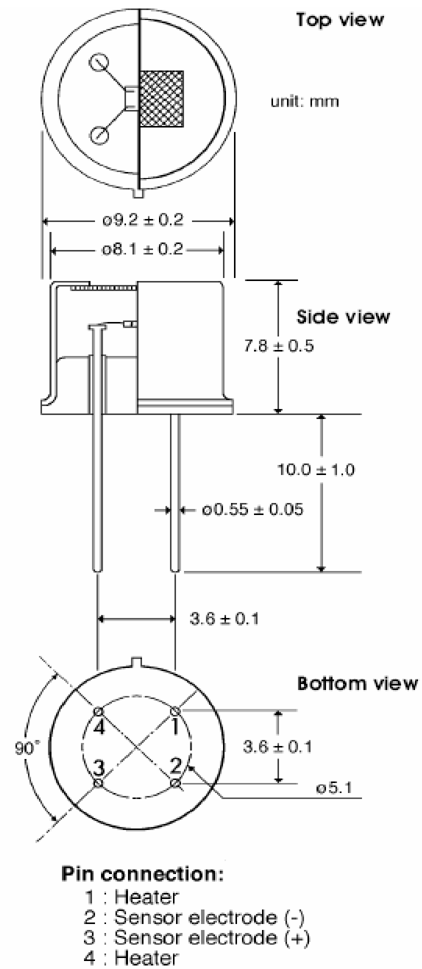
Choice of Sensors: TGS 2620

- Low power consumption
- High sensitivity to alcohol and organic solvent vapors
- Not responsive to carbon dioxide
- Long life and low cost
- Uses simple electrical circuit





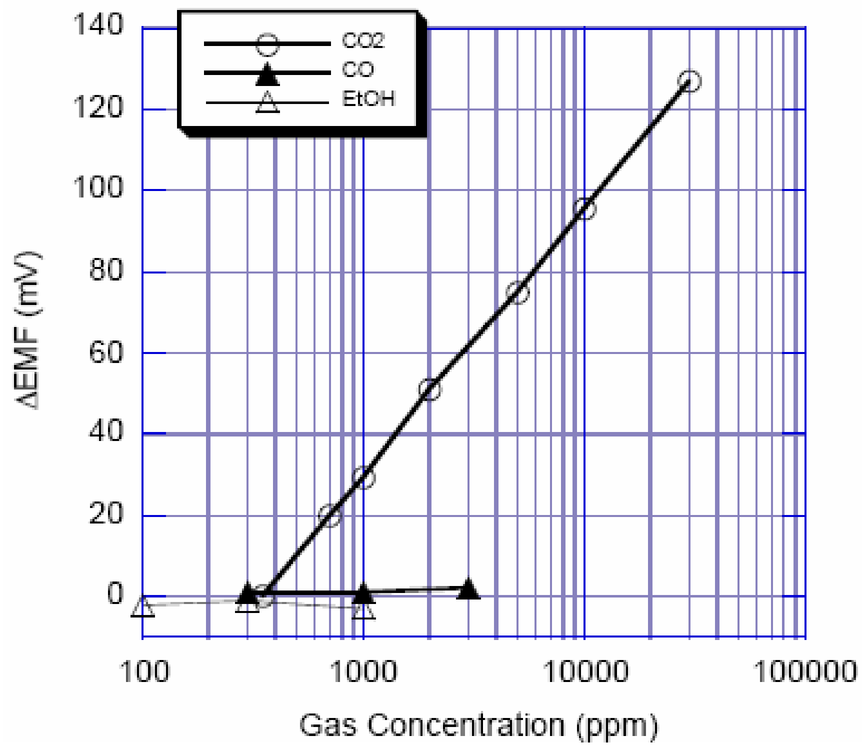
Choice of Sensors: TGS 2620

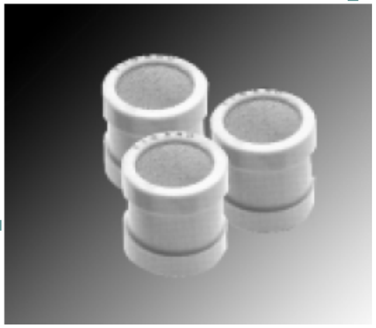


- Comprised of a metal oxide semiconductor layer formed on alumina substrate
- Simple electrical circuit provides an output signal based on changes in conductivity that corresponds with gas concentration

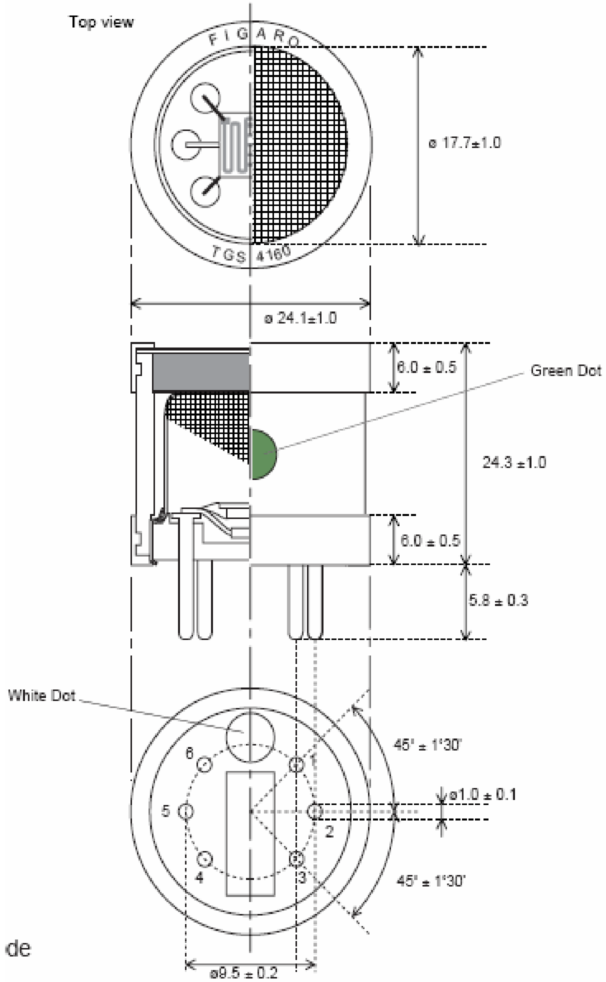
Choice of Sensors: TGS 4160

- High selectivity for carbon dioxide
- Unresponsive to ethanol
- Compact size
- Long life
- Electomotive force is used to create a signal output that corresponds to a detectible gas's concentration

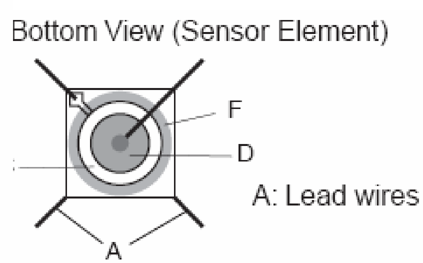


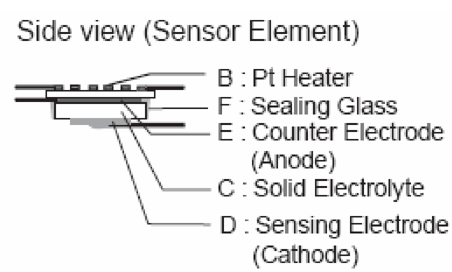


Choice of Sensors: TGS 4160



- Ethanol exposure tests confirm that the sensors response is not affected by the presence of ethanol
- The zeolite filter is installed in the sensor cap and eliminates the influence of interference gases



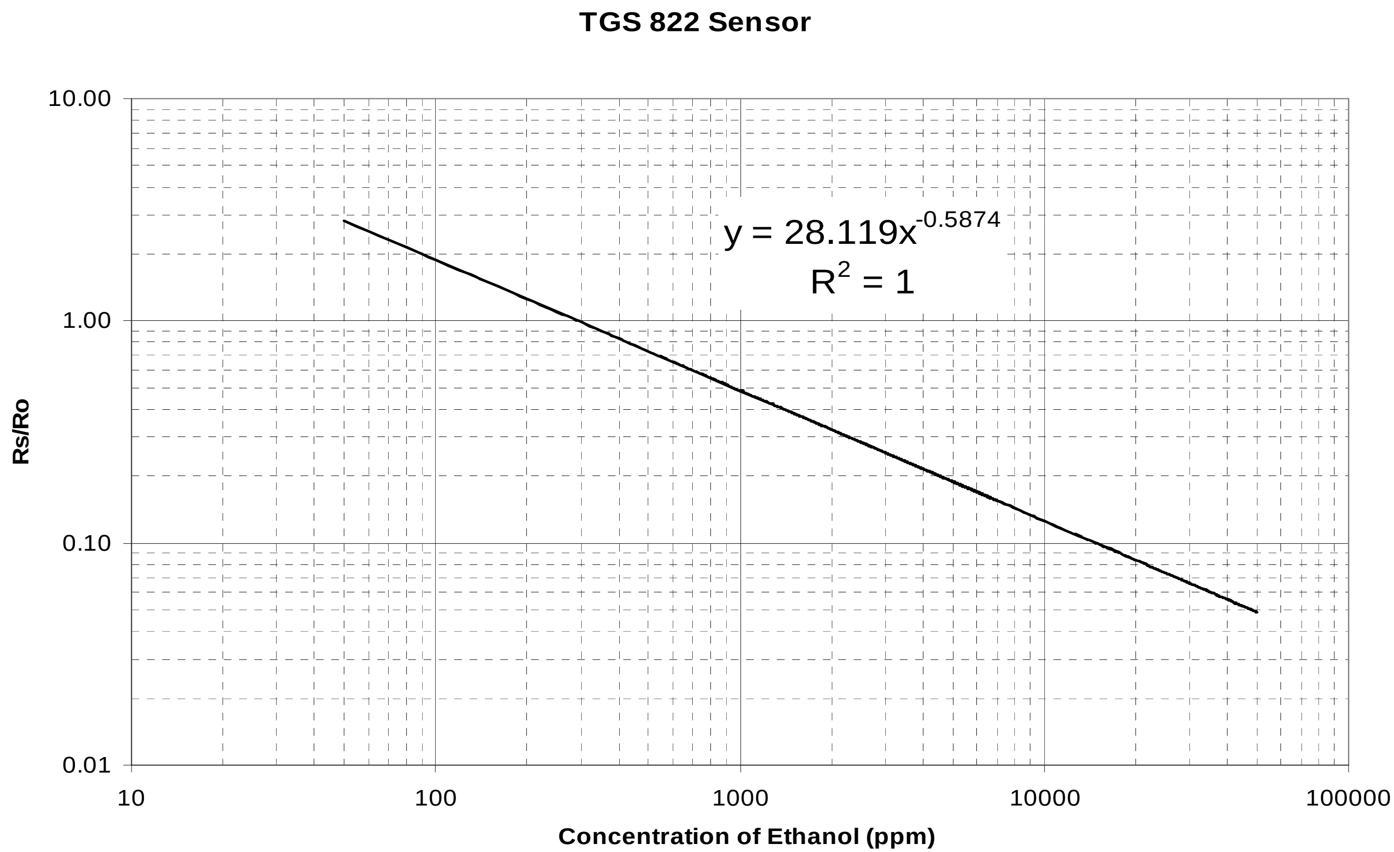


Sensor Data

- Each sensor has a different output signal versus concentration relationship.
  - Log-Log or Semi Log plots
- Graphs were reproduced in Microsoft excel by using the following methodology:
  - Output = m*(Concentration)$^{n-1}$
  - M and n were allowed to vary while the sum of the square of the difference of output and calculated output was minimized in the Excel Solver add in.

Sensor Data

- A typical reproduced output vs. concentration plot



Sensor Data

- These plots were then used to develop an Excel spreadsheet with data representing the output signal as a function of concentration.

- Based on a known experimental process (*Camen Pinheiro, Carla M. Rodrigues, Thomas Schafer, Joao G. Crespo*) the vaporized concentration limits for 1st, 2nd, and 3rd stage of fermentation were calculated.

- The data was then classified using these limits

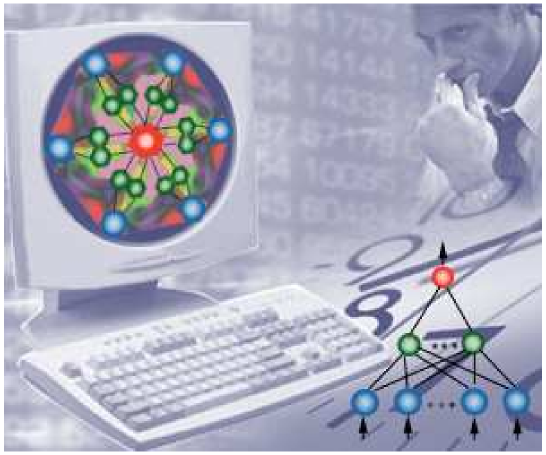
Sensor Data

- A sample of the original Microsoft Excel spread sheet

TGS 822 Rs/Ro	TGS 822 Concentration (ppm)	TGS 2620 Rs/Ro	TGS 2620 Concentration (ppm)	TGS 4160 EMF (mV)	TGS 4160 Concentration ppm	FERMENTATION STAGE
0.16585	6244.00000	0.155831	6244	212.4276762	63531.37931	FIRST
0.16583	6245.00000	0.155815	6245	212.4114168	63567.81609	FIRST
0.16582	6246.00000	0.1558	6246	212.3951668	63604.25287	FIRST
0.1658	6247.00000	0.155784	6247	212.3789261	63640.68966	FIRST
0.16578	6248.00000	0.155769	6248	212.3626947	63677.12644	FIRST
0.16577	6249.00000	0.155754	6249	212.3464725	63713.56322	FIRST
0.16575	6250.00000	0.155738	6250	212.3302597	63750	FIRST
0.16574	6251.00000	0.155723	6251	212.0926114	64286.48844	SECOND
0.16572	6252.00000	0.155707	6252	211.8569381	64822.97688	SECOND
0.16571	6253.00000	0.155692	6253	211.6232072	65359.46532	SECOND
0.16569	6254.00000	0.155676	6254	211.3913871	65895.95376	SECOND
0.16568	6255.00000	0.155661	6255	211.1614467	66432.4422	SECOND
0.16566	6256.00000	0.155645	6256	210.9333558	66968.93064	SECOND
0.16564	6257.00000	0.15563	6257	210.7070848	67505.41908	SECOND

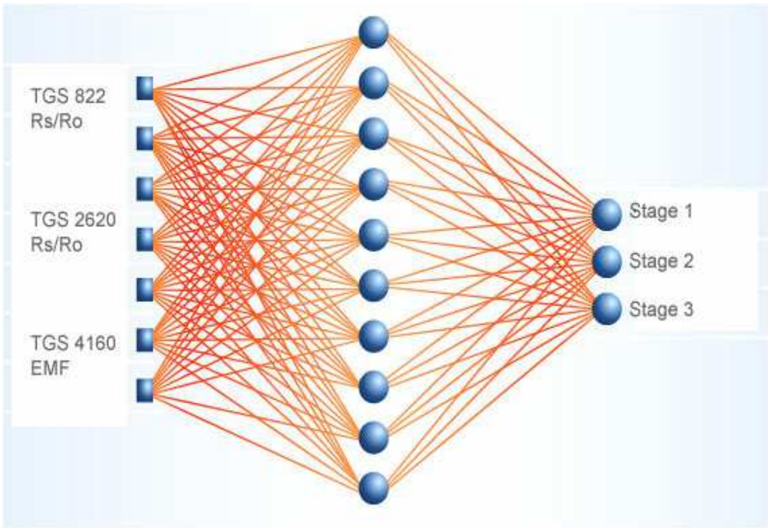
NeuroSolutions for Excel 5

- NeuroSolutions 5 creates the most powerful and easy to use neural network simulation environment on the market today.
- Allows for the use of a neural network while working within a familiar spreadsheet environment



NeuroSolutions Problem Definition

- Trained a neural network to classify stages of fermentation
  - 1st, 2nd, or 3rd.
- Data collected from 2458 samples of data:
  - 1741 1st Stage data
  - 692 2nd Stage data
  - 25 3rd Stage data
- Preprocess Data
  - Randomize Row Function to randomize samples
- Tagged data columns as Input, Output, and rows as Training, Cross Validation, Testing



NeuroSolutions Problem Definition

- Excel sheet sample with input and output tags

Sample	R1	C1	R2	C2	R3	C3	Fermentation	Stage 1	Stage 2	Stage 3		R1 / C1	R2 / C2	R3 / C3
2318	0.157375	6827	0.147437	6827	162.2097	373303.8	SECOND	0	1	0		TGS 822	TGS 2620	TGS 4160
239	0.285261	2480	0.276306	2480	267.7789	9021.954	FIRST	1	0	0				
1278	0.173418	5787	0.163353	5787	221.047	46879.77	FIRST	1	0	0				
221	0.29817	2300	0.289527	2300	269.9192	8366.092	FIRST	1	0	0				
1737	0.165815	6246	0.1558	6246	212.3952	63604.25	FIRST	1	0	0				
2290	0.157755	6799	0.147814	6799	163.3744	358282.2	SECOND	0	1	0				
2445	0.058322	37000	0.051682	37000	157.4556	441437.9	THIRD	0	0	1				
1970	0.162286	6479	0.1523	6479	181.8729	186605.9	SECOND	0	1	0				
2013	0.161657	6522	0.151677	6522	178.5675	209674.9	SECOND	0	1	0				
1958	0.162463	6467	0.152476	6467	182.8686	180168	SECOND	0	1	0				
1006	0.178392	5515	0.168305	5515	227.7823	36968.97	FIRST	1	0	0				
1622	0.167635	6131	0.157606	6131	214.3278	59414.02	FIRST	1	0	0				
2146	0.159752	6655	0.149789	6655	170.2616	281027.8	SECOND	0	1	0				
1187	0.17504	5696	0.164967	5696	223.1272	43564.02	FIRST	1	0	0				
134	0.394175	1430	0.388784	1430	283.4255	5196.092	FIRST	1	0	0				
2120	0.160119	6629	0.150154	6629	171.7053	267079.1	SECOND	0	1	0				
1522	0.169262	6031	0.159222	6031	216.1225	55770.34	FIRST	1	0	0				
211	0.306057	2200	0.297621	2200	271.182	8001.724	FIRST	1	0	0				
1770	0.165303	6279	0.155291	6279	206.1377	79308.16	SECOND	0	1	0				
958	0.17931	5467	0.16922	5467	229.1566	35220	FIRST	1	0	0				
1818	0.164565	6327	0.15456	6327	198.1638	105059.6	SECOND	0	1	0				
1012	0.178278	5521	0.168192	5521	227.6151	37187.59	FIRST	1	0	0				
1201	0.174788	5710	0.164716	5710	222.7971	44074.14	FIRST	1	0	0				
1257	0.173789	5766	0.163722	5766	221.5137	46114.6	FIRST	1	0	0				
98	0.467378	1070	0.465408	1070	291.6762	3884.368	FIRST	1	0	0				
1291	0.17319	5800	0.163126	5800	220.7619	47353.45	FIRST	1	0	0				
609	0.186394	5118	0.176288	5118	241.8593	22503.56	FIRST	1	0	0				
909	0.180261	5418	0.170168	5418	230.6319	33434.6	FIRST	1	0	0				
1811	0.164672	6320	0.154666	6320	199.196	101304.2	SECOND	0	1	0				
2324	0.157294	6833	0.147357	6833	161.9662	376522.8	SECOND	0	1	0				

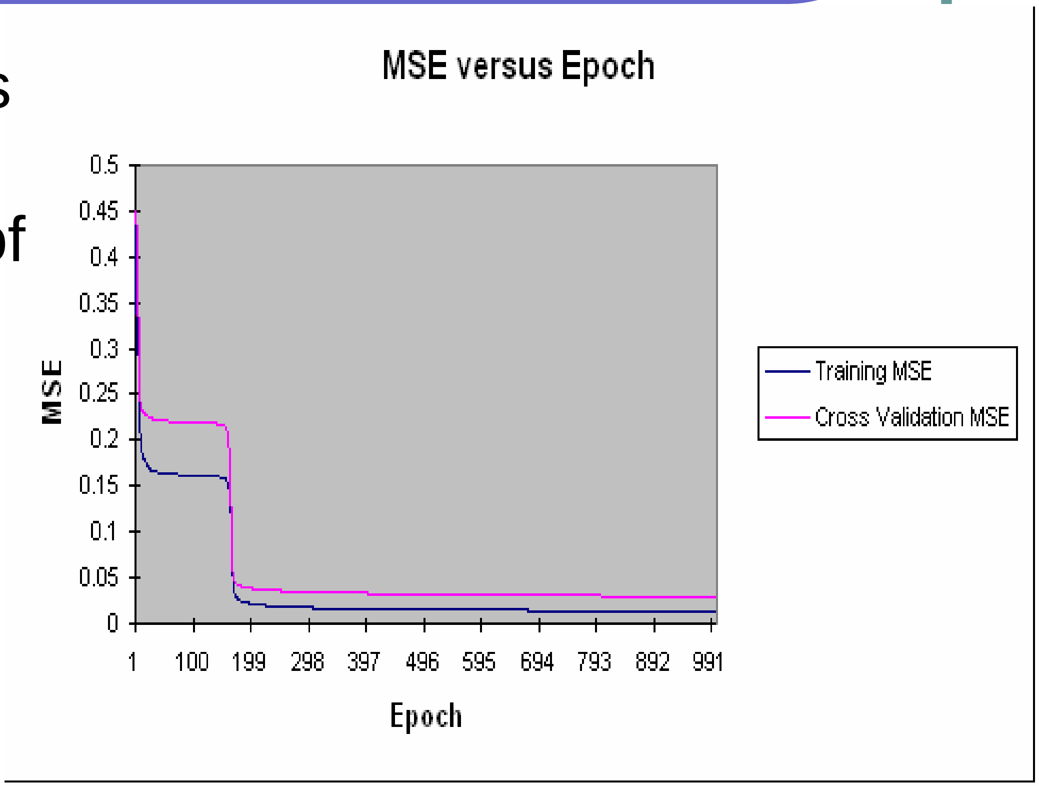
Neural Network Training Results

- Trained network using 1000 epochs
- Generated report summarizing training results:
  - Plot showing learning curve of training and cross validation data
  - Table with minimum and final mean-squared errors

Training MSE	Cross Validation MSE
0.43410369	0.44981748
0.340294	0.44223499
0.33280918	0.42945878
0.32267846	0.40656058
0.30733597	0.36021234
0.2787291	0.30898905
0.25573399	0.27647618
0.2279212	0.26232052
0.21351986	0.24691495
0.20126831	0.23870294
0.19461823	0.23526743
0.18960026	0.23345988
0.18549146	0.23249844
0.18321123	0.23088157
0.18100021	0.22958992
0.17932922	0.22886859
0.17807501	0.22835228
0.17679957	0.22799805
0.17564384	0.22763973
0.17452552	0.22717646
0.1733015	0.22671047
0.17203555	0.22629662
0.17081494	0.22591041
0.16972628	0.2255373
0.1688617	0.22516113
0.16821932	0.2247676
0.16772743	0.22437058
0.16733089	0.22399581
0.16700198	0.2236585
0.16672238	0.22336425

Neural Network Training Results

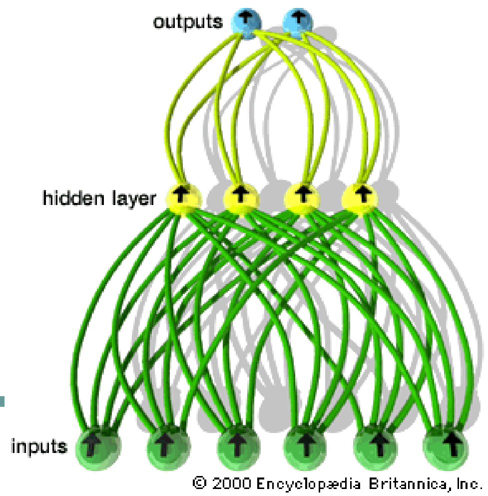
- Examine learning curves to see if trained neural network did a good job of learning the data
- To verify conclusion, need to run a testing set through the trained neural network model



Best Networks	Training	Cross Validation
Epoch #	1000	1000
Minimum MSE	0.013864389	0.029511521
Final MSE	0.013864389	0.029511521

Neural Network Testing Classifiers

- Determine the classification performance of the “Training” data set
- Test classification performance of data that network has never seen
  - This will tell us whether the neural network simply memorized the training data or truly learned the underlying relationship.



Training Data Classification Results

- Classification report generated
- Confusion matrix summarizes classification results in an easy to interpret format.
- Table lists various performance measures.
  - Percentage of samples classified correctly for each class

R1	R2	R3	Stage 1	Stage 2	Stage 3	Stage 1 Output	Stage 2 Output	Stage 3 Output	Output (Symbolic)
0.17444713	0.16437732	222.355111	1	0	0	1.01618618	-0.0219309	-0.0004503	Stage 1
0.16400239	0.15400168	193.256052	0	1	0	0.00263022	0.99366977	-0.0047552	Stage 2
0.45030135	0.44746936	289.872722	1	0	0	1.05540843	-0.0555258	-0.0209847	Stage 1
0.16617488	0.15615682	212.771294	1	0	0	0.81697629	0.18717517	-0.0057397	Stage 1
0.3477586	0.34060275	277.36157	1	0	0	1.05538075	-0.055508	-0.0133749	Stage 1
0.1641843	0.15418207	194.756638	0	1	0	0.02093752	0.97624006		
0.18539579	0.17529099	239.779437	1	0	0	1.05361962	-0.0542112		
0.16583079	0.15581537	212.411417	1	0	0	0.80224309	0.20344161		
0.17405485	0.163987	221.851812	1	0	0	1.01213263	-0.0181152		
0.17520269	0.16512926	223.341489	1	0	0	1.02304447	-0.0282908		
0.20129142	0.19119972	250.925538	1	0	0	1.05501702	-0.0552216		
0.18006565	0.16997323	230.324516	1	0	0	1.04670594	-0.0489073		
0.16771535	0.15768595	214.414855	1	0	0	0.87620098	0.12235204		
0.16033242	0.15036465	172.572865	0	1	0	-0.0530539	1.05050438		
0.1576465	0.14770579	163.036695	0	1	0	-0.0546353	1.05280952		

Output / Desired	*Stage 1*	*Stage 2*	*Stage 3*
Stage 1	134	1	0
Stage 2	0	44	1
Stage 3	0	0	0

Performance	*Stage 1*	*Stage 2*	*Stage 3*
MSE	0.007348178	0.013371824	0.004951796
NMSE	0.038624428	0.071316397	0.89630269
MAE	0.055570194	0.06102064	0.01696378
Min Abs Error	0.000225056	0.000279677	2.0753E-05
Max Abs Error	0.73125206	1.054121039	0.897267385
r	0.982389226	0.967126613	0.323188199
Percent Correct	100	97.77777778	0

Testing Data Classification Results

- True test of a network is how well it can classify samples that it has not seen before
- Another classification report generated with confusion matrix and table
  - See if you have developed a good model for the data

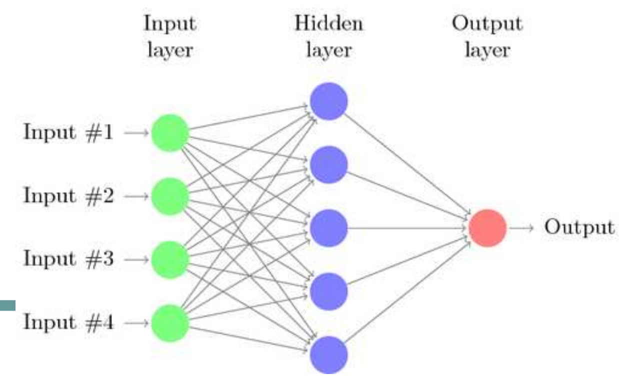
R1	R2	R3	Stage 1	Stage 2	Stage 3	Stage 1 Output	Stage 2 Output	Stage 3 Output	Output (Symbolic)
0.17444713	0.16437732	222.355111	1	0	0	1.01618618	-0.0219309	-0.0004503	Stage 1
0.16400239	0.15400168	193.256052	0	1	0	0.00263022	0.99366977	-0.0047552	Stage 2
0.45030135	0.44746936	289.872722	1	0	0	1.05540843	-0.0555258	-0.0209847	Stage 1
0.16617488	0.15615682	212.771294	1	0	0	0.81697629	0.18717517	-0.0057397	Stage 1
0.3477586	0.34060275	277.36157	1	0	0	1.05538075	-0.055508		
0.1641843	0.15418207	194.756638	0	1	0	0.02093752	0.97624006		
0.18539579	0.17529099	239.779437	1	0	0	1.05361962	-0.0542112		
0.16583079	0.15581537	212.411417	1	0	0	0.80224309	0.20344161		
0.17405485	0.163987	221.851812	1	0	0	1.01213263	-0.0181152		
0.17520269	0.16512926	223.341489	1	0	0	1.02304447	-0.0282908		
0.20129142	0.19119972	250.925538	1	0	0	1.05501702	-0.0552216		
0.18006565	0.16997323	230.324516	1	0	0	1.04670594	-0.0489073		
0.16771535	0.15768595	214.414855	1	0	0	0.87620098	0.12235204		

Output / Desired	Stage 1	Stage 2	Stage 3
Stage 1	50	2	0
Stage 2	0	22	0
Stage 3	0	0	0

Performance	Stage 1	Stage 2	Stage 3
MSE	0.015448211	0.015027416	0.000767293
NMSE	0.070495337	0.06857511	#DIV/0!
MAE	0.064506955	0.064156293	0.015647248
Min Abs Error	0.003877229	0.00154224	0.000332898
Max Abs Error	0.695174671	0.685019732	0.097046681
r	0.96786176	0.969005846	#DIV/0!
Percent Correct	100	91.66666667	#N/A

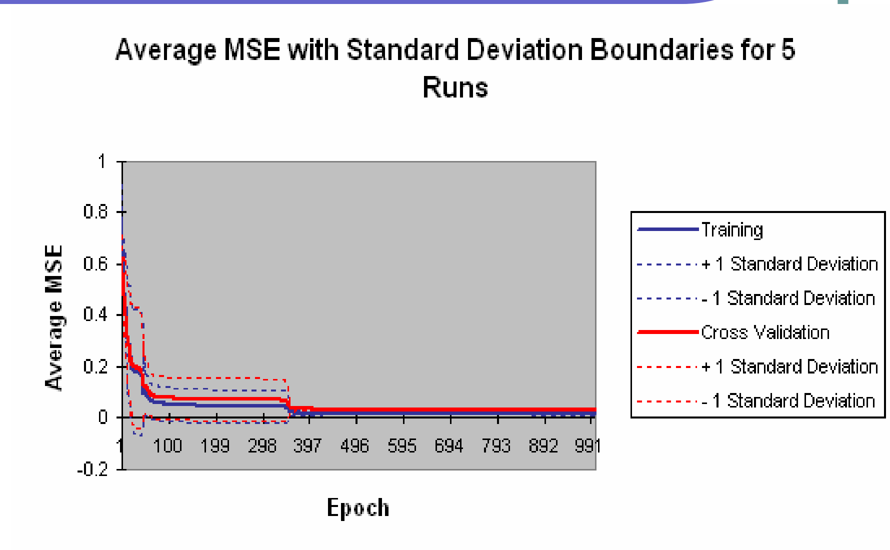
Neural Network Multiple Training

- Unlike a linear system, a neural network is not guaranteed to find the global minimum.
- A neural network can actually arrive at different solutions for the same data given different values of the initial network weights.
- Thus, in order to develop a statistically sound neural network model, the network must be trained multiple times.
- Networks were trained 3,4, and 5 times.
- 1000 epochs for each training run



Neural Network Multiple Training Result

- Graph gives average of multiple training runs along with standard deviation boundaries.
- 2 tables also generated
  - Average of Minimum MSE's & Average of Final MSE's
  - Information about best network over all of the runs

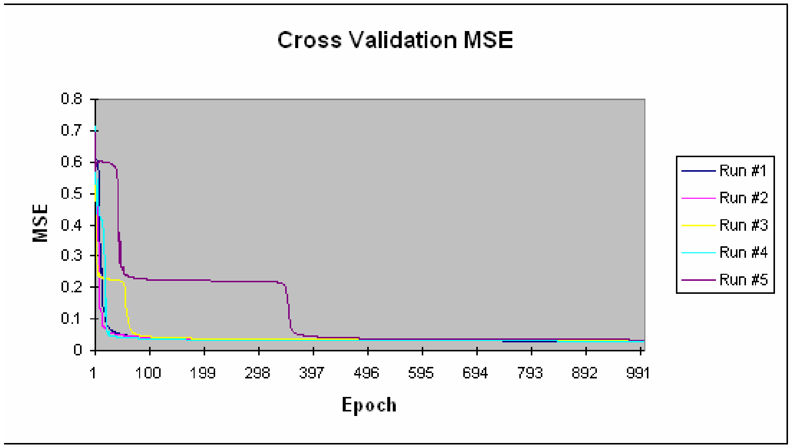


All Runs	*Training Minimum*	*Training Standard Deviation*	*Cross Validation Minimum*	*Cross Validation Standard Deviation*
Average of Minimum MSEs	0.014580204	0.000759845	0.030901978	0.002563347
Average of Final MSEs	0.014580204	0.000759845	0.030901978	0.002563347

Best Networks	*Training*	*Cross Validation*
Run #	4	1
Epoch #	1000	1000
Minimum MSE	0.013835231	0.028684369
Final MSE	0.013835231	0.028684369

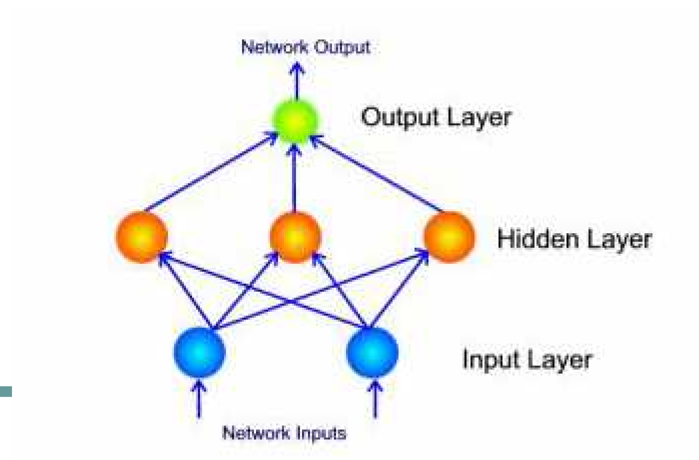
Neural Network Multiple Training Result

- Graph is a plot of learning curves for each of the runs
- Goal is to try and find a neural network model for which multiple trainings approach the same final MSE



Varying Network Parameters

- Developed a training process to train a neural network multiple times while varying:
  - Hidden layer processing elements
  - Step size
  - Momentum rate
- Develop an optimized neural network solution by varying any one of the network parameters to see which gives the best results

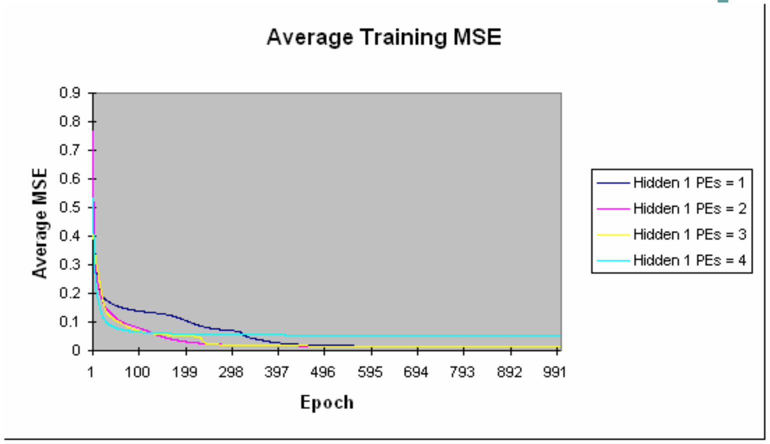


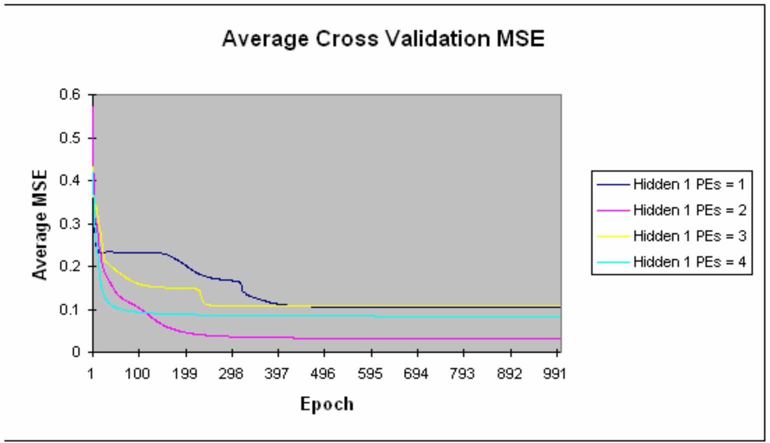
Varying Hidden Elements

- “Parameter Variation” training process to determine the optimum number of hidden processing elements for learning sensor data
- Number of hidden processing elements varied from 1 to 4.
- Each run for 1000 epochs and network run ‘n’ times for each parameter value
  - ‘n’ = optimal training number previously found

Varying Hidden Elements Results

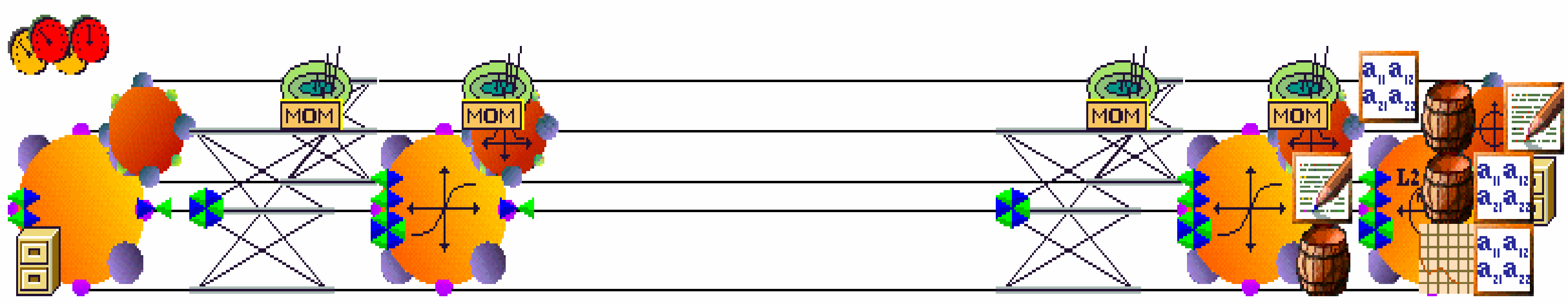
- Networks do not generally fully learn the problem with only 1 processing element in the hidden layer.
- Increasing the number of hidden processing units to 2 results in significant improvement in minimum MSE.
- Further increasing the number of processing elements eventually results in final MSE converging to same general value.
- Usually the network with more processing elements tends to learn faster.





Testing the Optimal Network

- Use data set tagged as “Testing” to test performance of best network found
- Testing report and confusion matrix should have improved results in learning to classify fermentation stages.



Testing the Optimal Network Results

R1	R2	R3	Stage 1	Stage 2	Stage 3	Stage 1 Output	Stage 2 Output	Stage 3 Output	Output (Symbolic)
0.15945615	0.14949694	169.146946	0	1	0	-0.0477093	1.03787958	0.0418876	Stage 2
0.22480933	0.21486511	256.265662	1	0	0	1.05312184	-0.0532125	-0.002688	Stage 1
0.17072968	0.16068037	217.801637	1	0	0	0.94758824	0.04419747	-0.0008302	Stage 1
0.16086159	0.15088878	174.833769	0	1	0	-0.0439408	1.03350547	0.0297924	Stage 2
0.22552228	0.21558478	256.418698	1	0	0	1.05313476	-0.0532279	-0.0026791	Stage 1
0.15624873	0.14632311	158.97768	0	1	0	-0.0508077	1.04254958	0.06819225	Stage 2
0.17368263	0.16361667	221.379582	1	0	0	0.99530406	-0.0022267	-0.0005765	Stage 1
0.15935803	0.14939979	168.784921	0	1	0	-0.0478795	1.03809848	0.04272011	Stage 2
0.20050563	0.19041158	250.736529	1	0	0	1.05252714	-0.0525384	-0.0028445	Stage 1
0.17401931	0.16395163	221.806497	1	0	0	0.99926153	-0.0059199	-0.0005575	Stage 1
0.18680154	0.1766949	242.745424	1	0	0	1.05066757	-0.0508165	-0.0022984	Stage 1
0.54644789	0.54893286	299.25453	1	0	0	1.05432032	-0.055003	0.005855	Stage 1
0.15889765	0.14894402	167.13785	0	1	0	-0.0485792	1.0390298	0.04660268	Stage 2
0.17648185	0.16640271	225.065752	1	0	0	1.02140057	-0.0259995	-0.0005079	Stage 1
0.18682306	0.1767164	242.792837	1	0	0	1.05068489	-0.0508313	-0.0023039	Stage 1
0.15783725	0.14789454	163.63031	0	1	0	-0.0497417	1.04072617	0.05539535	Stage 2
0.16969211	0.15964931	216.608336	1	0	0	0.9241341			
0.1727701	0.16270899	220.243017	1	0	0	0.983201			
0.18057418	0.17048021	231.130719	1	0	0	1.0404644			
0.41876567	0.41444211	286.355834	1	0	0	1.0541986			
0.16807013	0.15803823	214.801257	1	0	0	0.8791531			
0.16750665	0.15747874	214.188978	1	0	0	0.8610140			
0.1723706	0.16231169	219.754478	1	0	0	0.9772152			
0.25761638	0.24810795	262.850262	1	0	0	1.0535621			
0.15666186	0.14673171	160.127996	0	1	0	-0.050582			
0.18346267	0.17336135	236.065415	1	0	0	1.0468111			
0.17769251	0.16760843	226.76653	1	0	0	1.028821			
0.16975854	0.15971531	216.683826	1	0	0	0.9257551			

Output / Desired	*Stage 1*	*Stage 2*	*Stage 3*
Stage 1	50	2	0
Stage 2	0	22	0
Stage 3	0	0	0

Performance	*Stage 1*	*Stage 2*	*Stage 3*
MSE	0.015628101	0.015000623	0.000585812
NMSE	0.071316233	0.068452844	#DIV/0!
MAE	0.062833451	0.060256236	0.012355805
Min Abs Error	0.000738466	0.002226664	0.000184181
Max Abs Error	0.702854507	0.693608444	0.070025657
r	0.966287274	0.967849116	#DIV/0!
Percent Correct	100	91.66666667	#N/A

Run Parameter Results

Run #	Training	Cross-Validation	Testing	Optimal Runs	Optimal Processing Elements	Stage 1 Accuracy (%)	Stage 2 Accuracy (%)	Stage 3 Accuracy (%)
1	0.6	0.15	0.25	5	3	100	91.67	N/A
2	**0.7**	**0.1**	**0.2**	**3**	**4**	**100**	**100**	**N/A**
3	0.5	0.1	0.4	4	3	100	97.4	0
4	0.6	0.1	0.3	4	4	100	92	0
5	0.8	0.1	0.1	5	3	100	100	N/A

- Table illustrating different runs/best run
- 5 different runs with varying training, cross-validation, and testing percentages
- Best Run - #2

NeuroSolutions Evaluation Mode Limitations

- Maximum of 300 exemplars
  - Thus, we could use only 12% of all the data collected
- For more accurate results, require Full Version, so we can train, cross-validate, and test all samples
- Towards the end of the project, the full version without exemplar limitations was available.
  - Utilized ASCII text files instead of Excel
  - The inputs and desired variables were the same.

Full Version Results

- 80% training (10% cross-validation) & 20% testing,
  - entire data set used to train and test neural network model
- Results for stage 1 and 2 were quite accurate
  - 100% classification - stage 1
  - 99% classification - stage 2
- However, the original problem still remained
  - all stage 3 data was classified as stage 2

Full Version Results

- As a final try, an “optimized” data set was used
  - All stage 3 data and portions of stage 1 and 2 that were at the stage boundaries
- This ended up giving the best results overall, with a 100% classification rate for all 3 stages.
- The optimal neural network model had been found!

Justification of Neural Network

- Because gaseous carbon dioxide is produce in much greater quantities than gaseous ethanol
  - Neural network allows for each to have a different weight in determining the classification.
- Neural network allows for the addition of more sensors, including sensors that can detect more than one gas
- Future work on this project will include
  - Varying the number/type of sensors
  - Weighting the concentration measurements of ethanol more that the concentration measurements of carbon dioxide

Customer Satisfaction – Model Development

- Consumer satisfaction is based not only on demand but on the quality of the product.
- Consumer satisfaction, *S*, can be represented as follows:

$$S = \left(d_1^{\rho} + d_2^{\rho}\right)^{\frac{1}{\rho}}$$

  - Where d_1= demand for the WI – Nose
    d_2= demand for the competitor's product
    ρ = pre-determined factor = .76

Customer Satisfaction – Model Development

- The maximum consumer satisfaction solution can therefore be defined as follows:

$$p_1 d_1^{1-\rho} = p_2 d_2^{1-\rho}$$

  - Where p_1 = price of the WI-Nose
    p_2 = price of the competitor's device
- Suggests when prices of products are equal, demands will also be equal (not realistic)
- Therefore, model must be further developed to take into consideration the effect of product quality on demand

Customer Satisfaction – Model Development

- The following relationship is generated introducing two variables to account for this effect.

$$p_1 d_1^{1-\rho} = \left(\frac{\alpha}{\beta}\right) p_2 d_2^{1-\rho}$$

- The parameters α and β represent the inferiority function and the superiority function
  - Inferiority function = consumer's knowledge for the product of interest.
  - Superiority function = consumer's preference for the product of interest in comparison to the competitor's product

Customer Satisfaction – Model Development

- The parameter *Y* represents the consumer's budget and can be represented as follows:

$$Y \leq p_1 d_1 + p_2 d_2$$

- Consumer satisfaction should be maximized while still satisfying the consumer's budget

$$p_1 d_1 = p_2 \left(Y - p_1 d_1 \right)^{1-\rho} d_1^{\rho}$$

Customer Satisfaction – Model Development

- By satisfying these conditions, the following solution to the consumer satisfaction maximization can be derived as an implicit equation for d_1

$$\Phi(d_1) = p_1 d_1 - \left(\frac{\alpha}{\beta}\right)^{\rho} p_2 \left[\frac{Y - p_1 d_1}{p_2}\right]^{1-\rho} d_1^{\rho} = 0$$

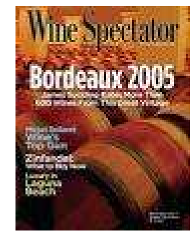
  - Where $\beta = H_2/H_1$
    - H_1= consumer's preference for the WI-Nose and
    - H_2 = consumer's preference for the competition's product
- These can be calculated as follows: $H_i = \sum w_i y_i$
- Where the w_i's are the weights associated with respective y_i's, or happiness functions

Market Evaluation - Proposal

- Number of wineries in the U.S. = 4740
- Proposed Market: California
  - Accounts for 90% of American wine production
- Relatively small number of wineries in California implies that information about Wi – Nose can and will be spread quickly.
- This implies that an *α* value of 1 will be reached within the first year.

Market Evaluation - Advertising

- We plan on accomplishing this by advertising
  - www.WineBusiness.com
    - Most highly trafficked website for the wine industry
  - WineBusiness Monthly
    - Industry's Leading Publication for Wineries and Growers
    - latest developments and trends in the global business of making wine, emphasis on new products
  - Unified Wine and Grape Symposium (UWGS)
    - Has become the largest wine and grape show in the nation



Consumer Satisfaction Model

- To calculate β, we need to calculate H_1 and H_2, the consumer preference for the WI-Nose and the competition's device.
- Three device design characteristics were allowed to vary
  - Accuracy
  - Size
  - Weight

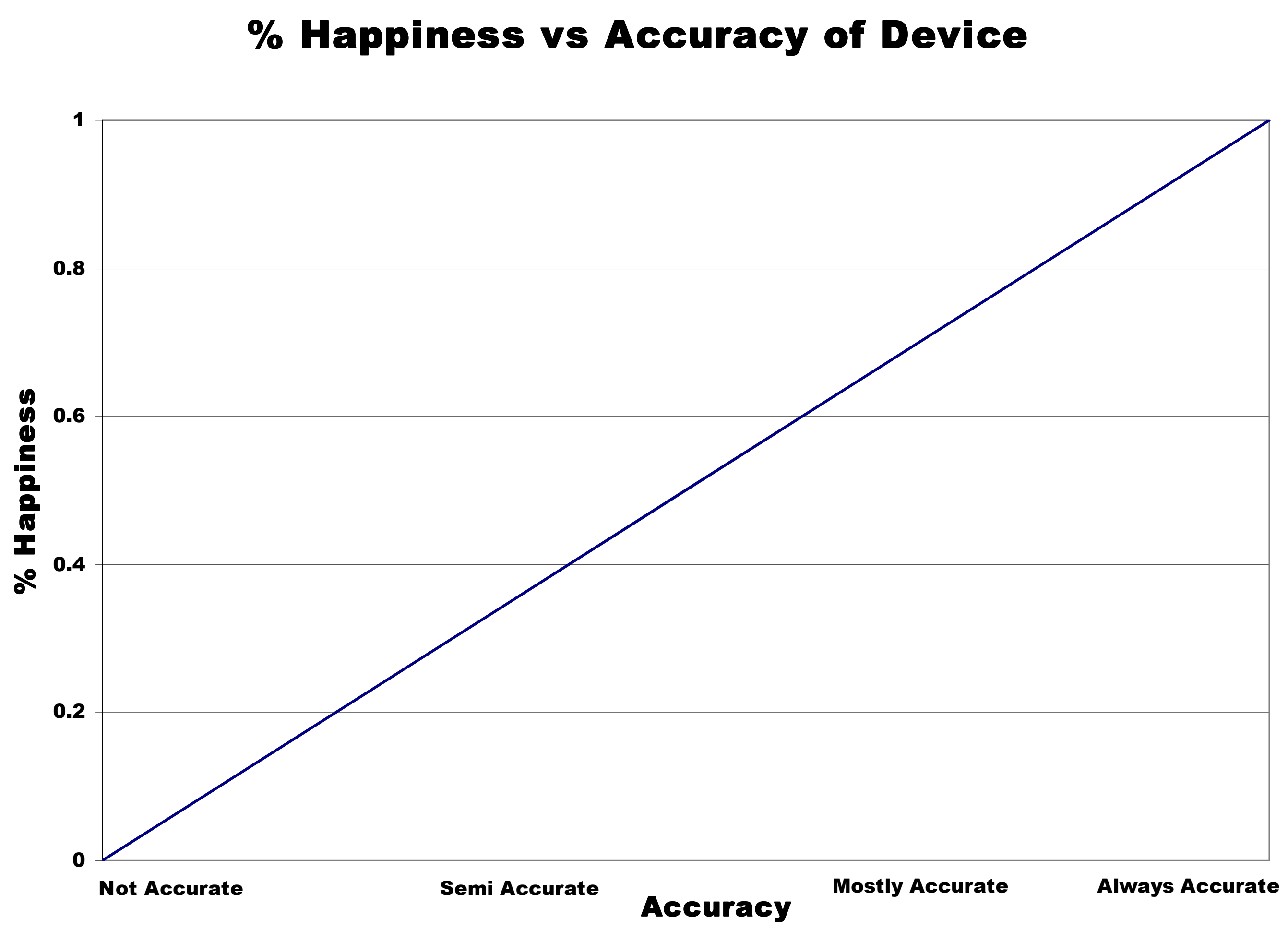
Consumer Satisfaction Model

- An informal survey was performed to determine optimal consumer satisfaction based on these three device characteristics
- This resulted in the following weights:

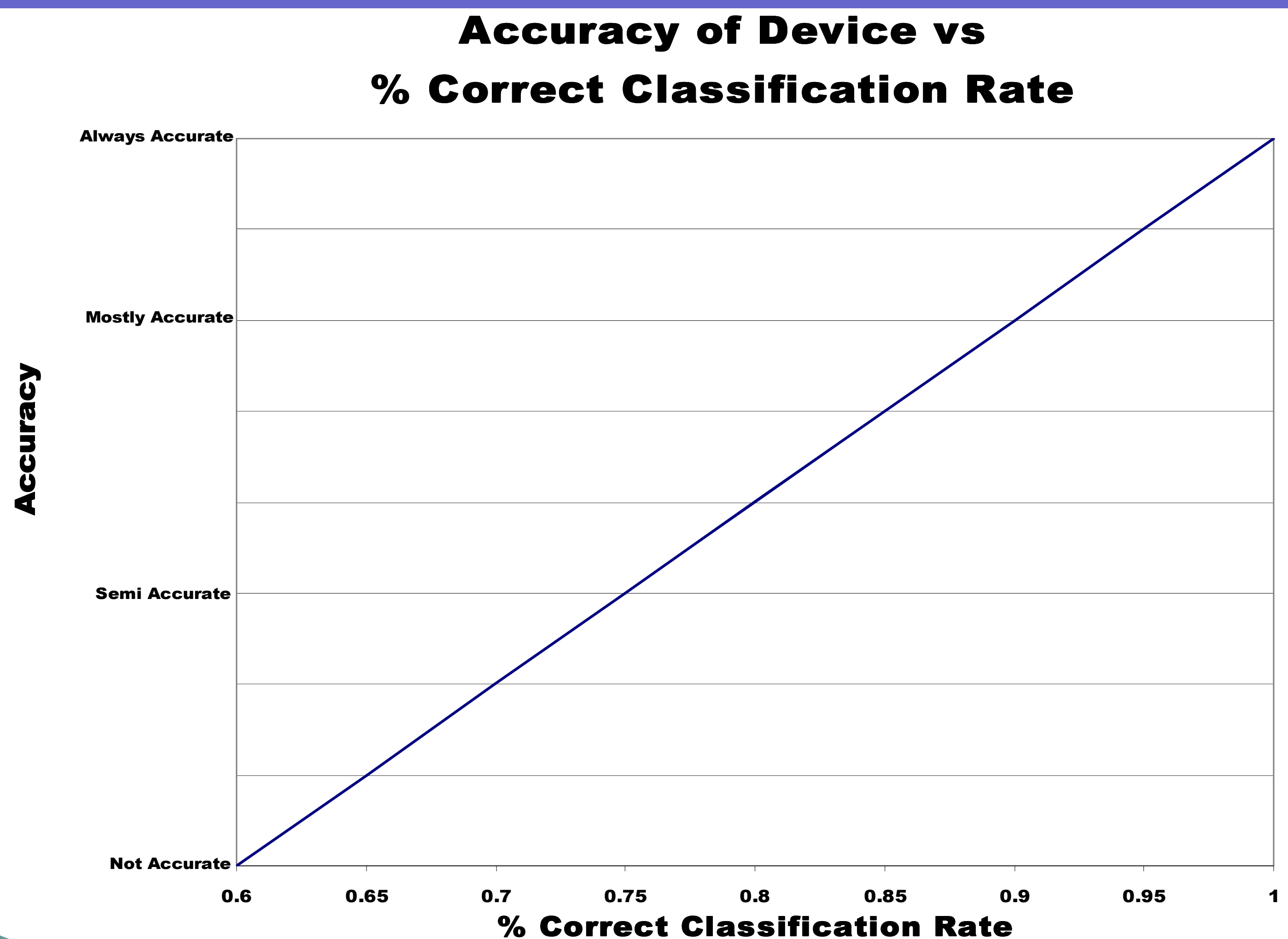
Design Characteristic	Weight
Accuracy	0.43
Size (cc)	0.23
Weight (pounds)	0.34

- Now, the happiness functions y_i's for the three design characteristics must be determined.

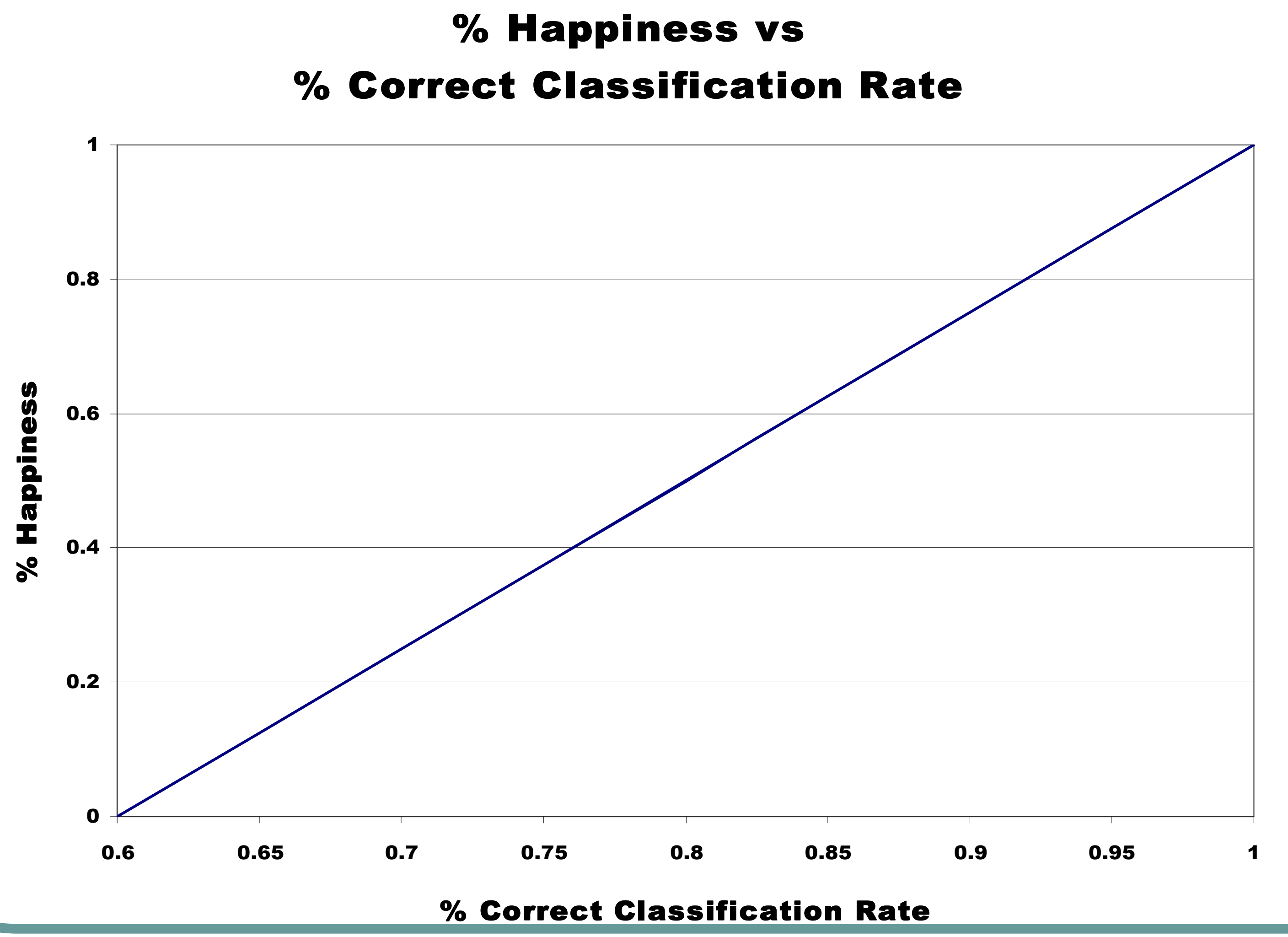
Consumer Satisfaction Model



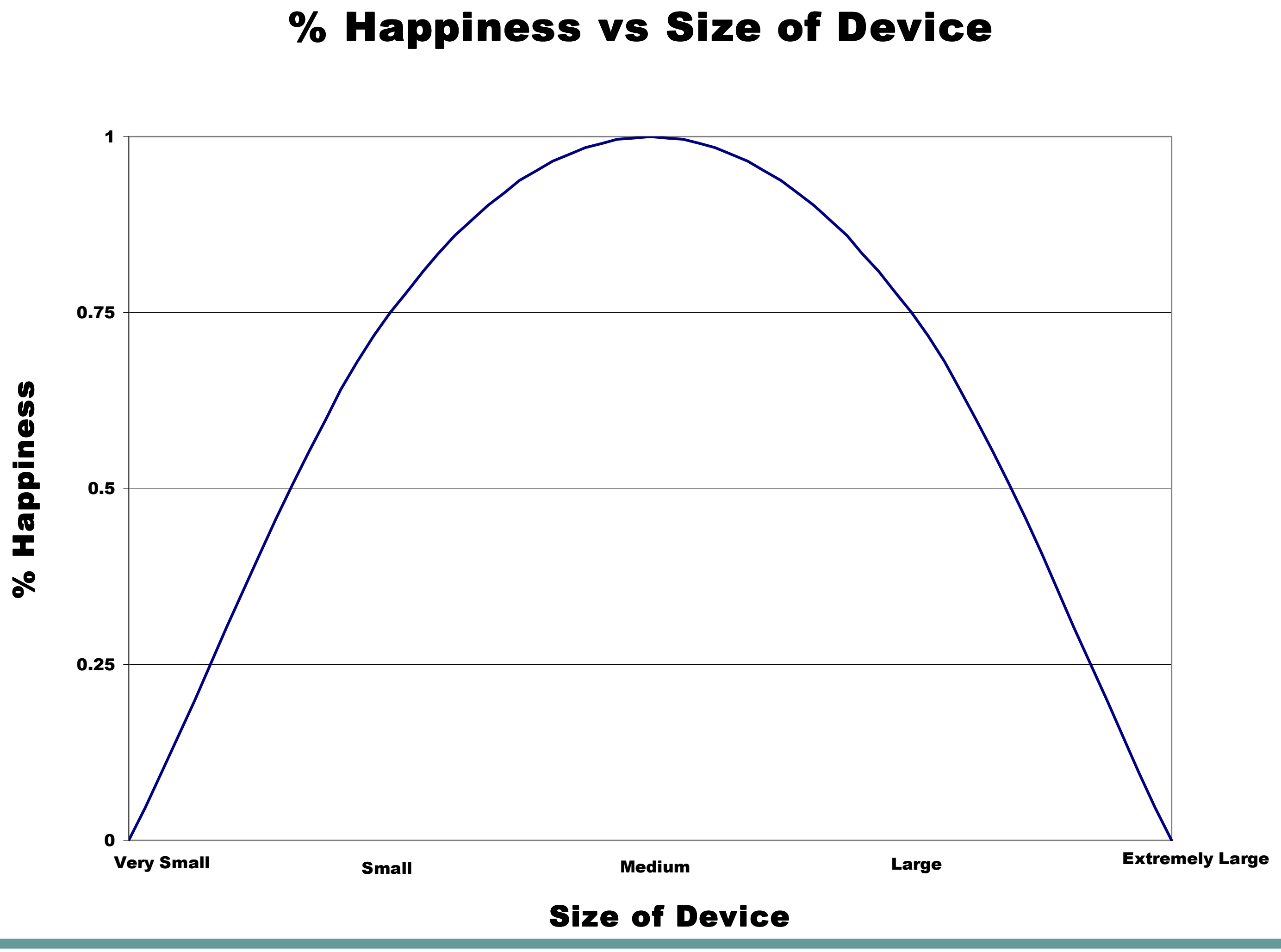
Consumer Satisfaction Model



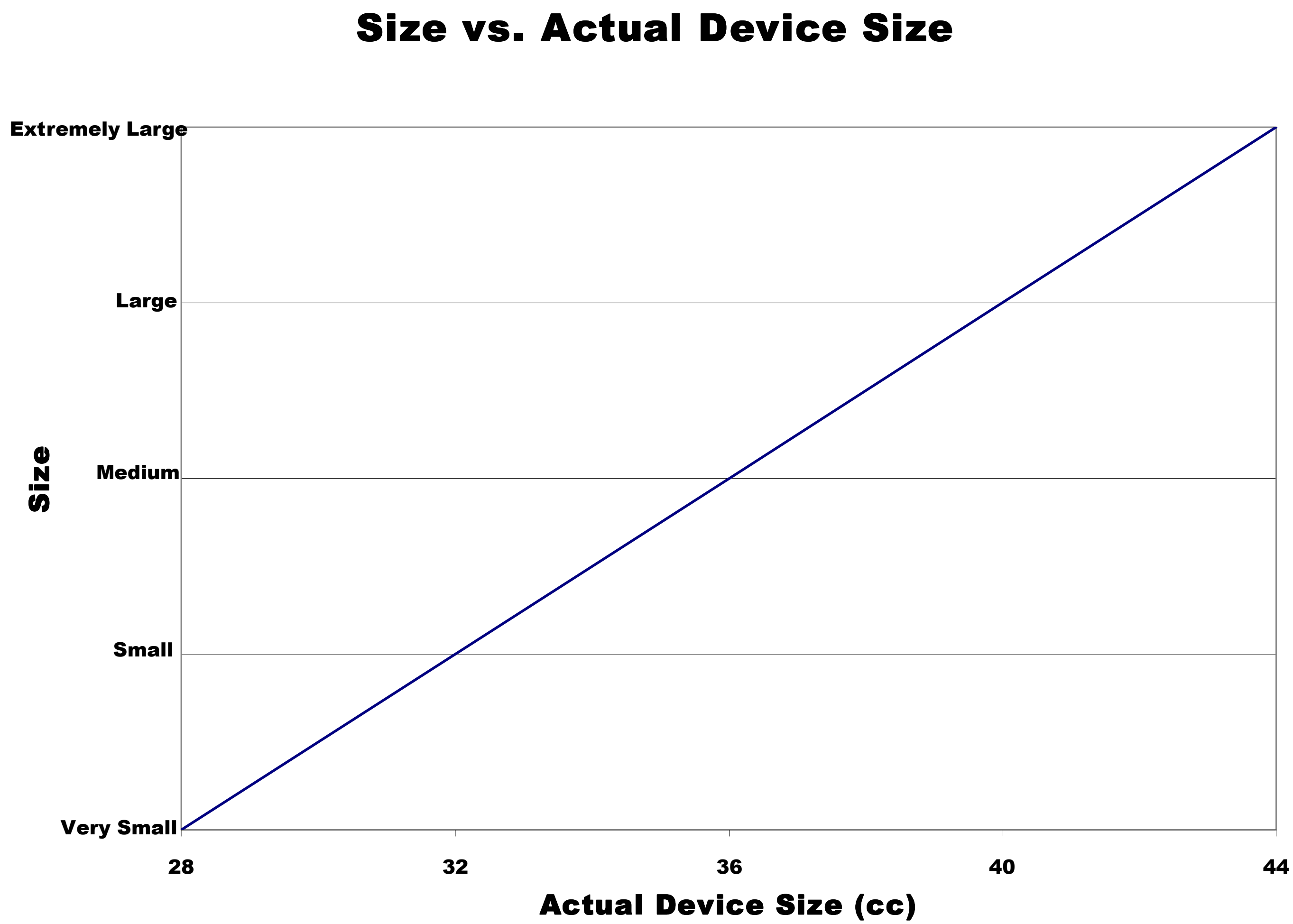
Consumer Satisfaction Model



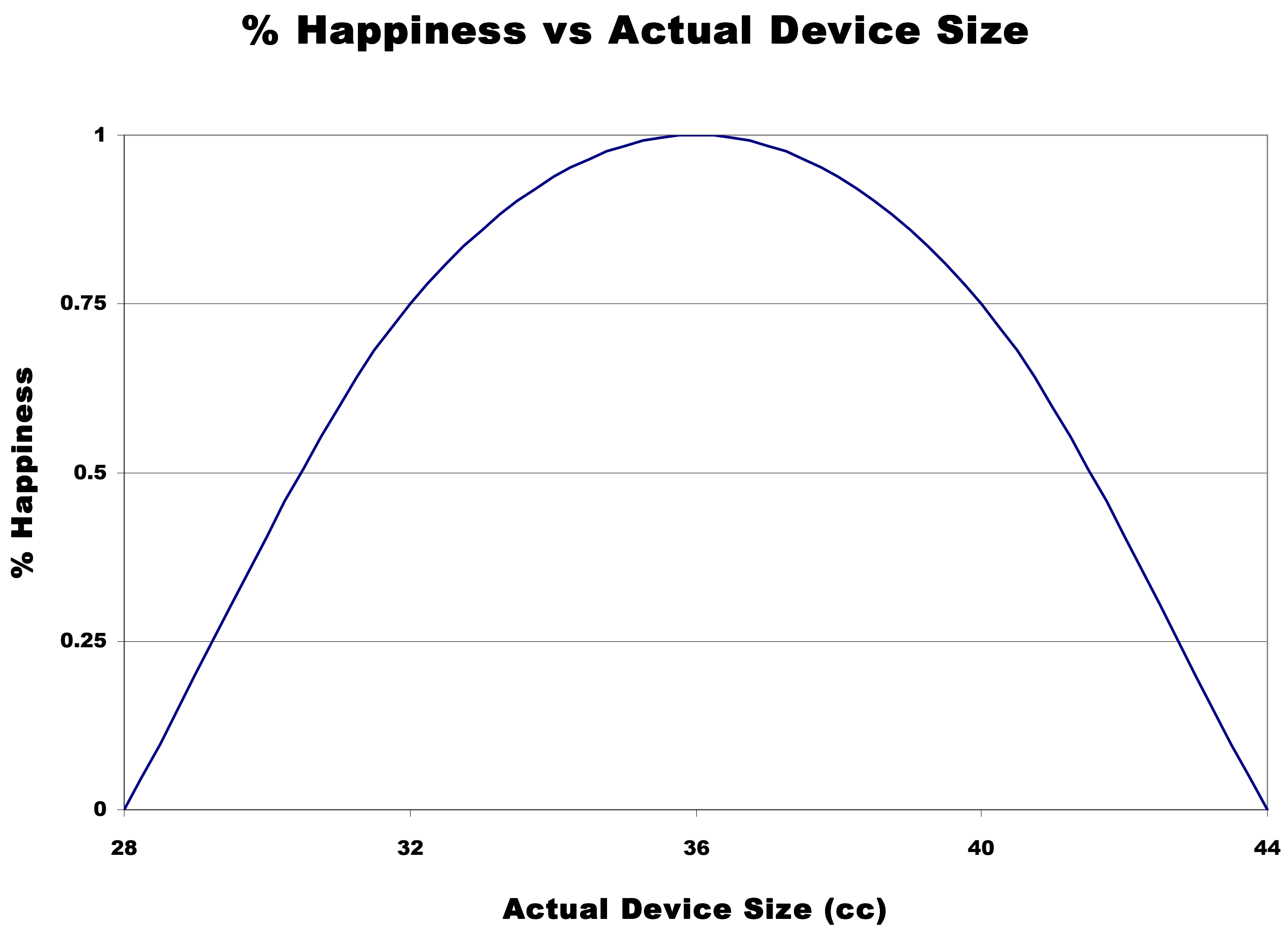
Consumer Satisfaction Model



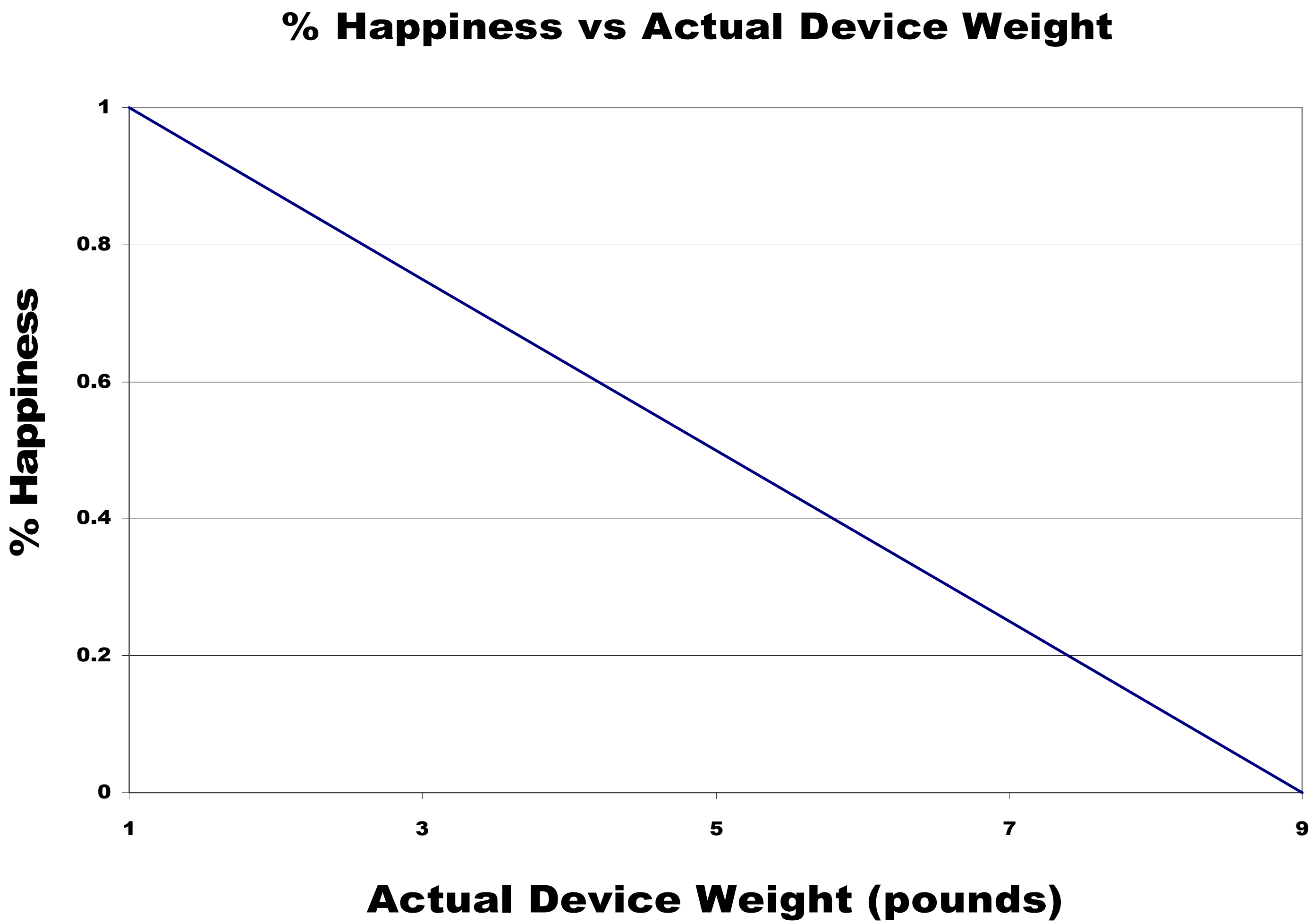
Consumer Satisfaction Model



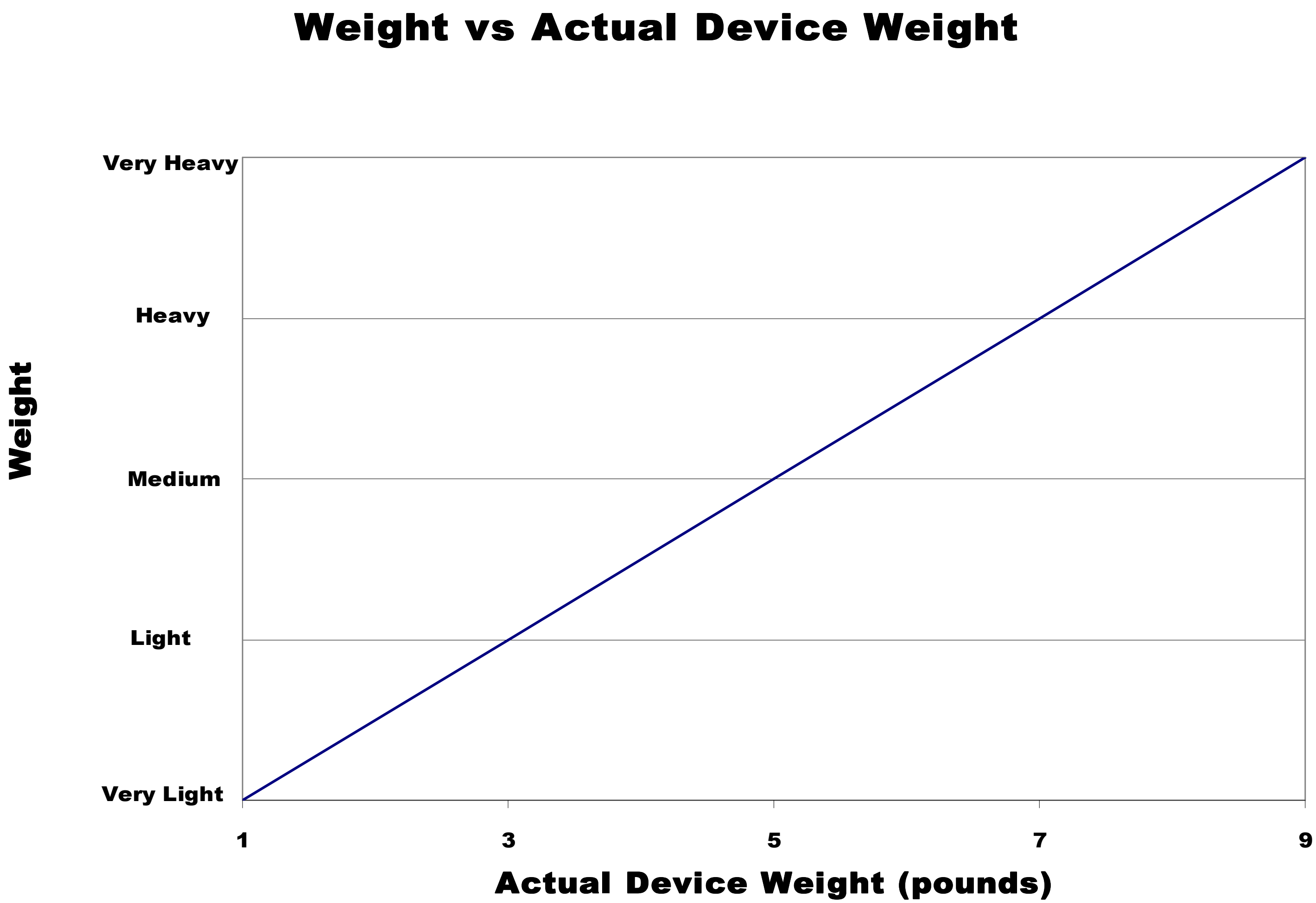
Consumer Satisfaction Model



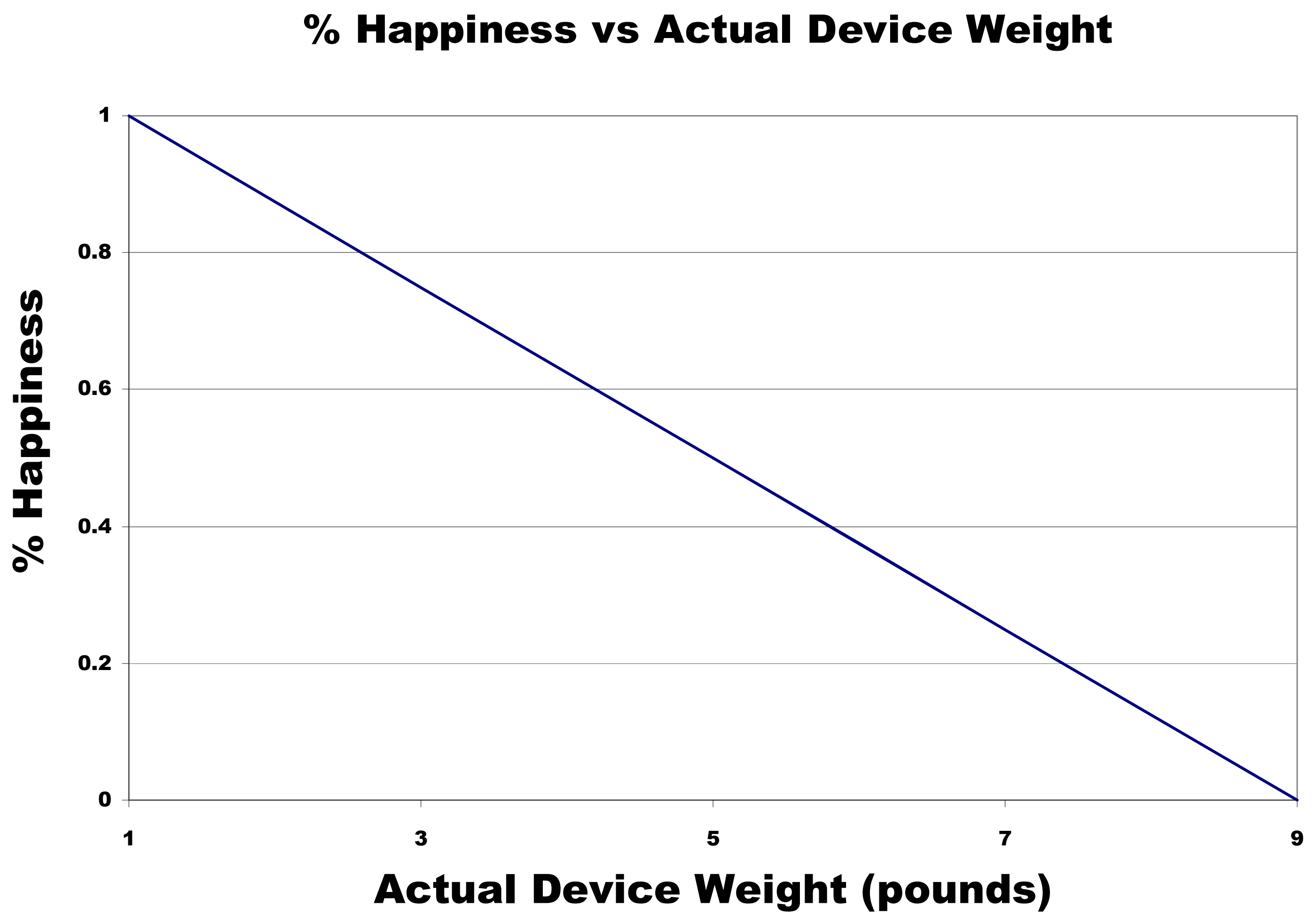
Consumer Satisfaction Model



Consumer Satisfaction Model



Consumer Satisfaction Model



Consumer Satisfaction Model Competition – Cyranose 320

- Weight: ~ 2.5 lbs
- Dimensions: ~ 100 cc
- Currently used in diverse industries including petrochemical, chemical, food, packaging, plastics, pet food and many more.
  - Accuracy for wine fermentation stage classification: semi-accurate (75%)
- Cost: ~ $10,000

Consumer Satisfaction Model

Device Characteristic	Our Device	y_i Our Device	Weights	H1	H2	Beta
Accuracy	1	1	0.43	0.77	0.6	0.779221
Size (cc)	28	0	0.23			
Weight (lbs)	1	1	0.34			

- The happiness functions were then combined with the appropriate weights to calculate H_1.
- H_2 was calculated using the given characteristics for the Cyranose 320.
- The β value was then calculated.

Consumer Satisfaction Model

- This Beta value was used to determine demand for various product prices.
- This methodology was repeated for various values of the design characteristics to attain many different demands.

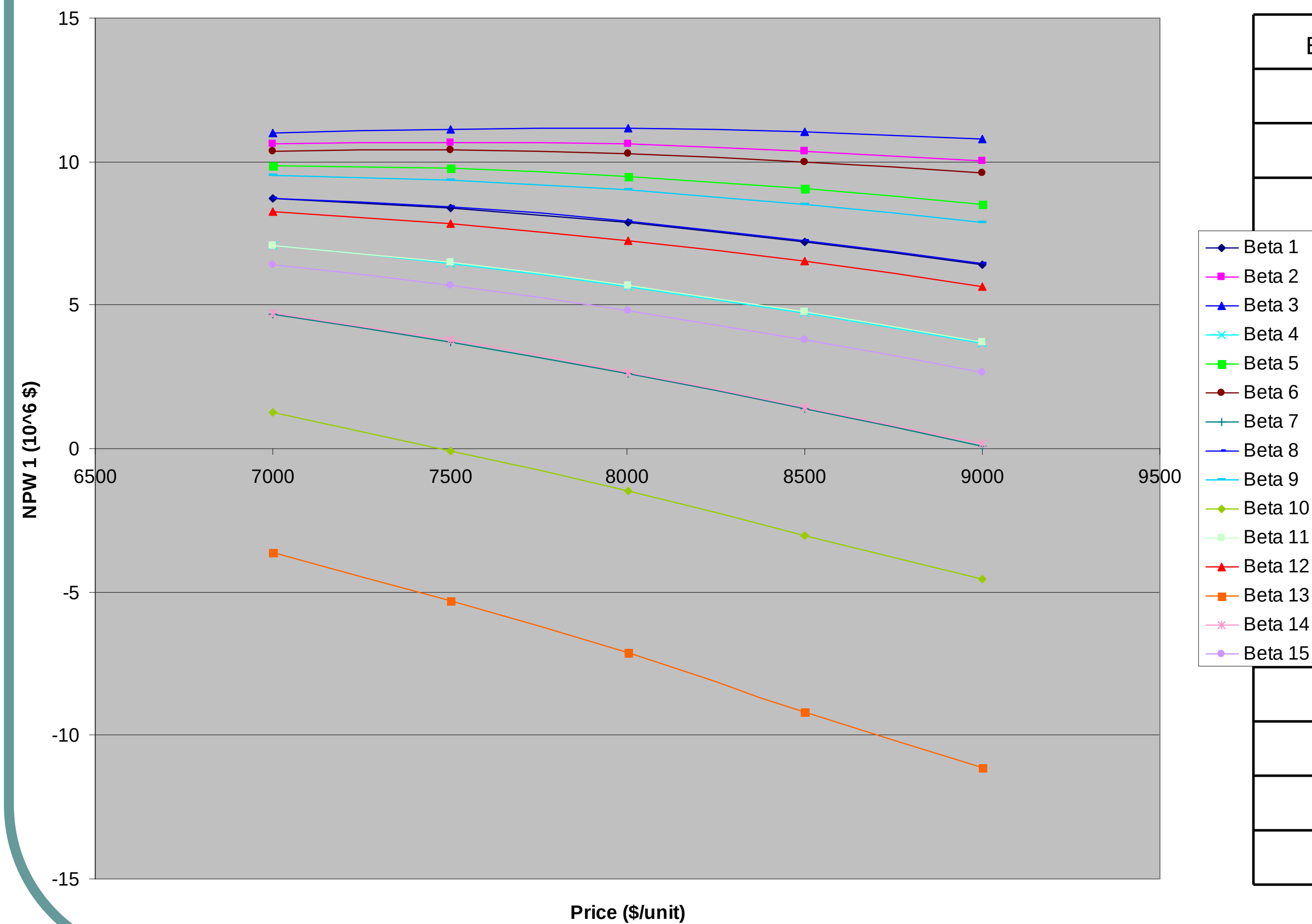
Price 1 ($/unit)	Price 2 ($/unit)	D1 B1-28,1
9000	10000	1215.83
8500		1341.86
8000		1480.93
7500		1634.94
7000		1806.44

H2	Alpha	Rho	Y
0.6	1	0.76	14500000

Consumer Satisfaction Integration

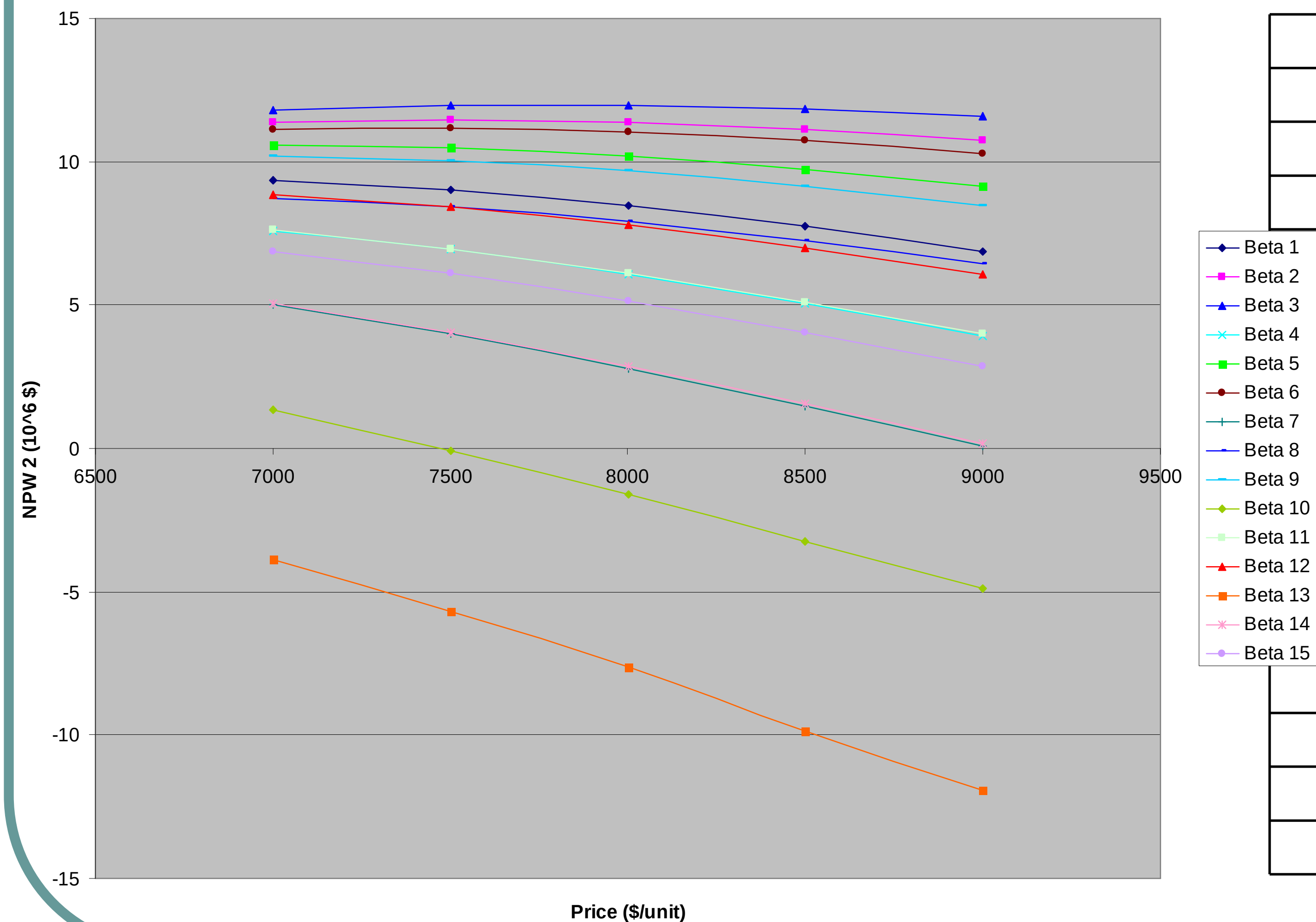
- Using these price-demand combinations, net present worth's were attained.
  - NPW 1 using Annual End-of-Year Cash Flows and Discounting
  - NPW 2 with Continuous Cash Flows and Discounting
- Accounted for the size and weight that contributed to specific *β*'s by adjusting raw materials costs
- Ultimately graphed NPW vs. product price for each of the β's.

Consumer Satisfaction Integration



Beta	Value
1	0.779220779
2	0.636604775
3	0.6
4	0.875912409
5	0.699708455
6	0.655737705
7	1
8	0.776699029
9	0.722891566
10	1.165048544
11	0.872727273
12	0.805369128
13	1.395348837
14	0.995850622
15	0.909090909

Consumer Satisfaction Integration



Beta	Value
1	0.779220779
2	0.636604775
3	0.6
4	0.875912409
5	0.699708455
6	0.655737705
7	1
8	0.776699029
9	0.722891566
10	1.165048544
11	0.872727273
12	0.805369128
13	1.395348837
14	0.995850622
15	0.909090909

The "Best" Product Design

- Beta 3 proved to be the most profitable
  - Accuracy = 100%
  - Size = 36 cc
  - Weight = 1 pound
- Price = $8,000
- Demand = 1651 units
- Total Capital Investment (TCI) = $6.514 million
- Total Annual Value of Products = $13.21 million
- Total Annual Cost of Raw Materials = $2.07 million
- Return on Investment (ROI) = 49.2%
- Payback Period = 1.5 years
- Net Return = $2.22 million
- NPW 1 = $11.15 million
- NPW 2 = $11.97 million

The “Best” Product Design

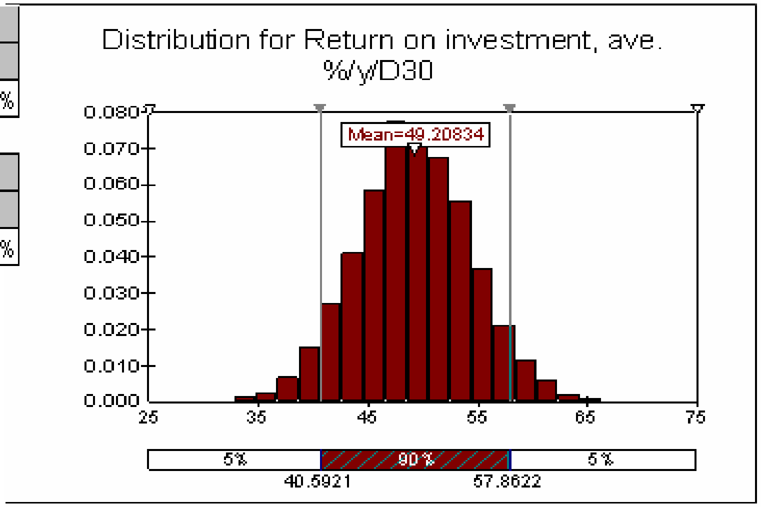
- Generally, the optimal happiness product is not the most profitable due to costs associated with its desired characteristics.
- With our device the optimal happiness product is also the most profitable.
  - The characteristics that were varied (size & weight) have very little costs associated with them (cover-$2/unit, board-$1.50/unit, wiring- $2/unit).
  - This is unlike other cases in which the product’s characteristics have much more significant costs associated with them.

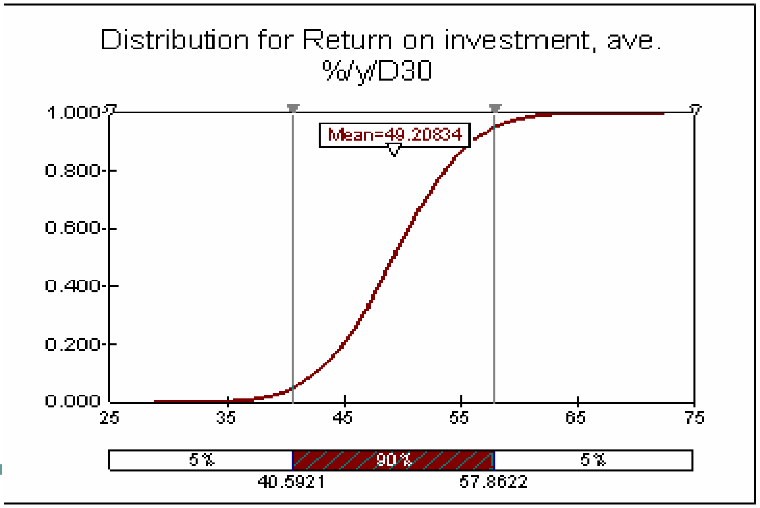
Risk Analysis

Output		Statistics						
Name	Cell	Minimum	Mean	Maximum	x1	p1	x2	p2
Return on investment, ave. %/	Evaluation!D30	28.9	49.2	72.2	40.6	5%	57.9	95%

Input		Statistics						
Name	Cell	Minimum	Mean	Maximum	x1	p1	x2	p2
Annual_Raw_Materials_Cost	Materials&Labo	0.20228821	2.074001661	3.697429657	1.370520473	5%	2.767649889	95%

- 20% variability in raw material costs for device
- Normal Distribution
- 10,000 iterations
- Monte Carlo Sampling Type
- Desired Output - ROI





Future Considerations/Work

- Get more 3rd stage data
- Vary number/type of sensors to get different values of accuracy
- See if a device can be designed that will give higher NPW but is not the “perfect” product
  - Sensor and software costs more significant than size and weight costs

Questions, Comments, Concerns, Suggestions?