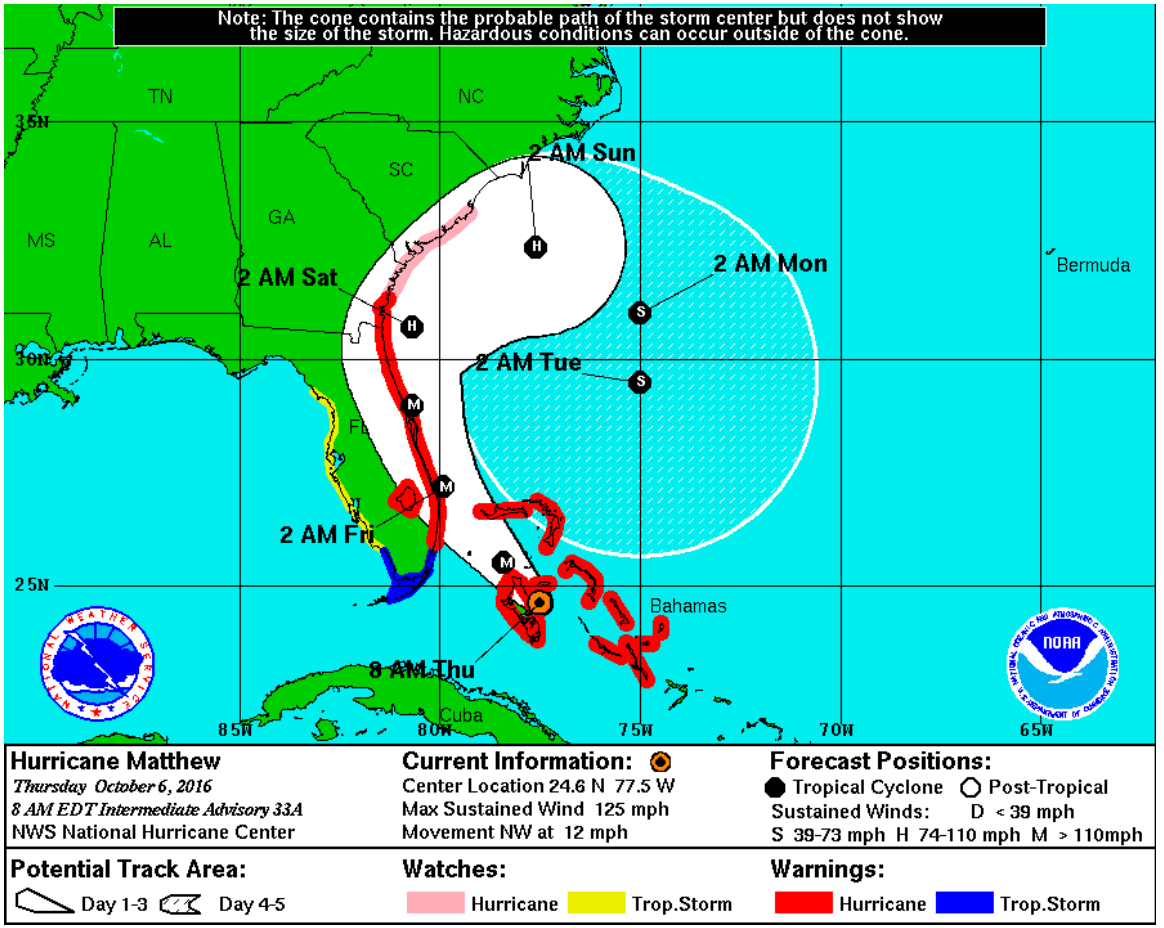
Decision Making in the Face of Uncertainty

Ronald Michaels, PhD
Technical Director
Phenotype Screening Corporation
November 3, 2016

Decision Making in the Face of Uncertainty

- Statistical Methods Interpret the Past
- All Decisions made on Incomplete Knowledge
- Decision Making is about the Future

Predicting the Future is Not Easy



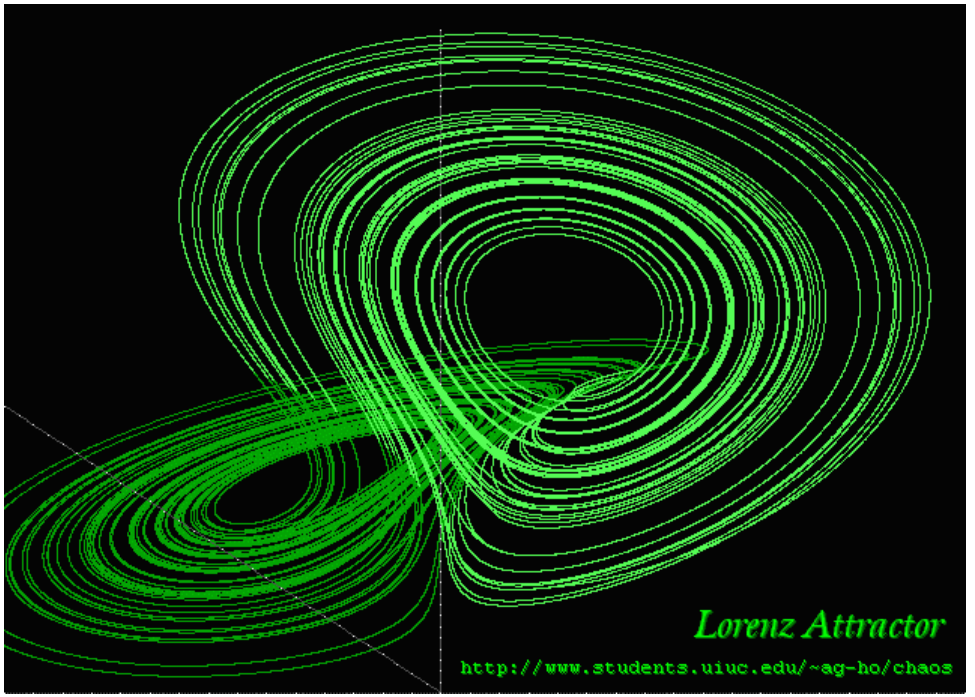
The further out into the future one attempts to predict, the less precise the prediction will be.

Complexity

- Complexity is not Randomness
- Requires an Immense Number of Variables to describe a model
- Many Variables are not Known or Measured
- Many Causal Relationships are Hidden
- Predictive Models are not Predictive Outside of the Range of Training Data
- Good News: Big Data

Chaotic Processes

- Purely Deterministic in Concept
- Extremely Sensitive to Initial Conditions
- Chaotic Attractor – no Equilibrium State
- Limited Predictability
- Lorenz Attractor:
- Good News: Dimensionality of Attractor can be known (complexity of system)



Nonlinearity

- Nonlinearity: 2x fertilizer does not produce 2x yield
- Complex Processes are Inherently Nonlinear
- Nonlinear relationships are mathematically and computationally difficult
- Good News: Computers can handle nonlinear problems – example: finite element methods
- More Good News: the ascendancy of Artificial Intelligence – an inherently nonlinear technology

Quantum Processes

- Example: Photosynthesis
- Heisenberg Uncertainty Principle
- Quantum Processes are Probabilistic
- Good News: At the Classical Newtonian level, the effect of Quantum Randomness is small
- Bad News: Biology works in the transition zone from Quantum to Newtonian

The Likelihood of Improbable Events

- The Future includes an incredible number of potential events; each one highly Improbable
- It is a Statistical Certainty that an Improbable Event will occur because there are so many possibilities
- Example: The Birthday Problem – in a room of 23 people the chance of two people having the same birth date is more than 50%

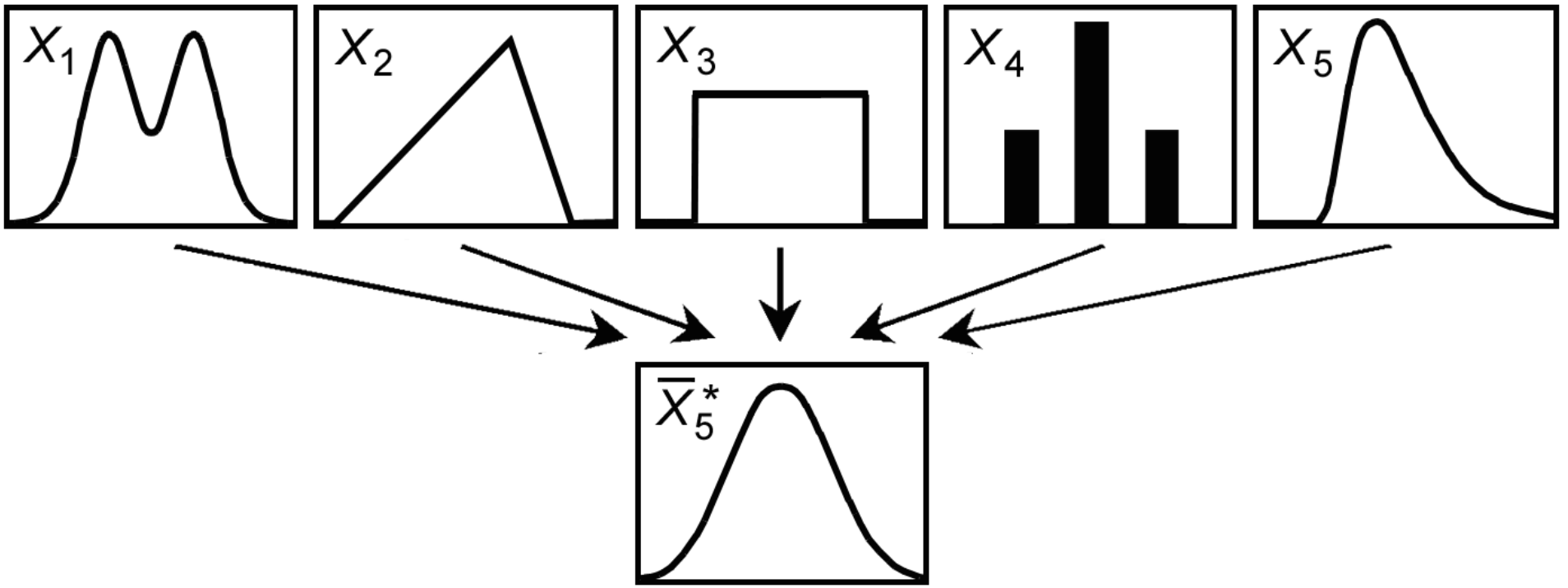
The High Cost of Knowledge

- Use Large Data Sets to “Average Out” unknown or unmeasured variables
- Need for more Field Trial Instrumentation
  - Within Field Variations – fertility, pests, etc.
  - Weather
- Ratings and Scores are Inexpensive, but Imprecise, and Ephemeral
- Need for PSC Growth Trials (small size, simplified growth platform)

Examples: Additional Knowledge

- Central Limit Theorem
- Pattern Classification
- Effects of Surfactants

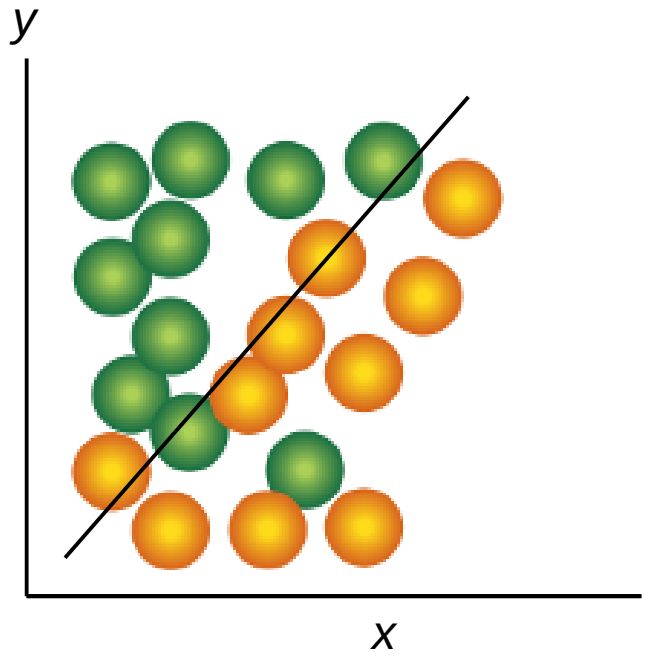
Central Limit Theorem



- Combination of several distributions is Gaussian
- A non-Gaussian Experimental Distribution Suggests that a very few Processes Dominate the Experimental Result
- Provides insight into the Complexity of a Process

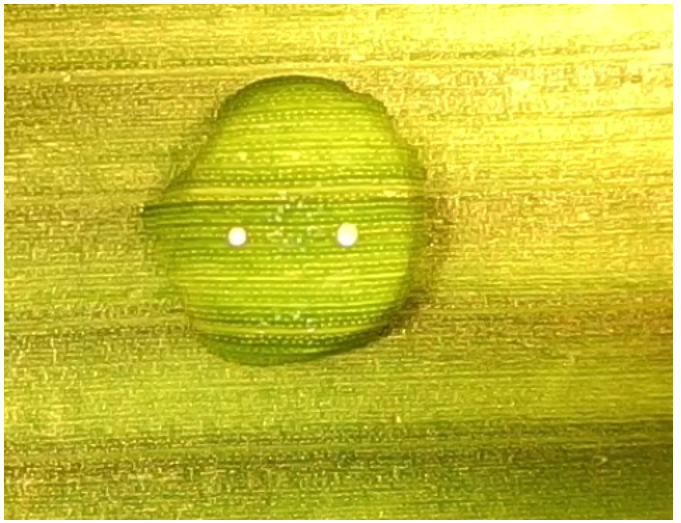
Pattern Classification

- Classes overlap along *x* and *y* axes
- A diagonal line that takes both *x* and *y* information into account can separate the classes
- Entire body of experimental results can reveal relationships not apparent in any one variable

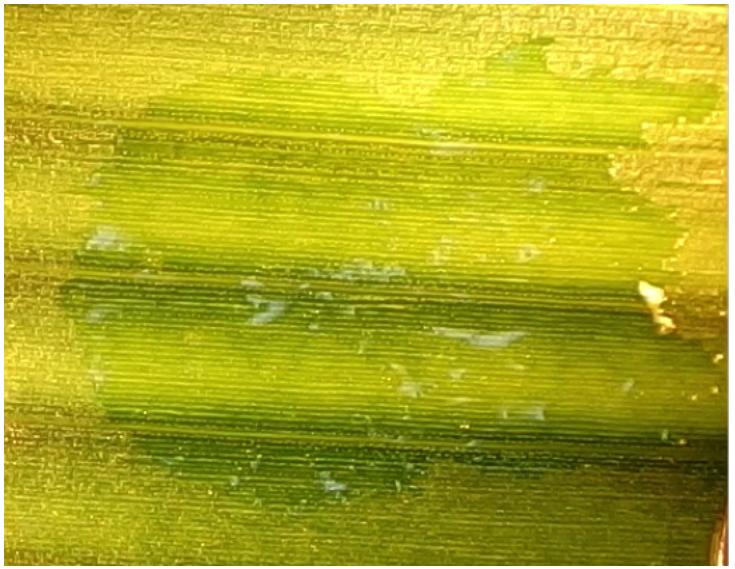


Effect of Surfactant on Foliar Spray

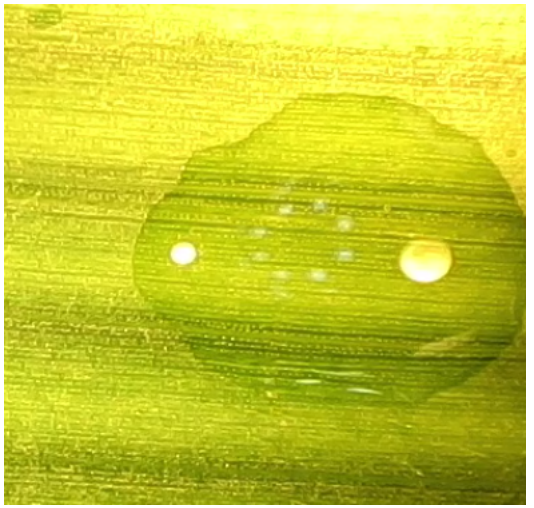
Interactions between active ingredient, surfactant, and leaf affect spreading of foliar spray



treatment only on Maize leaf

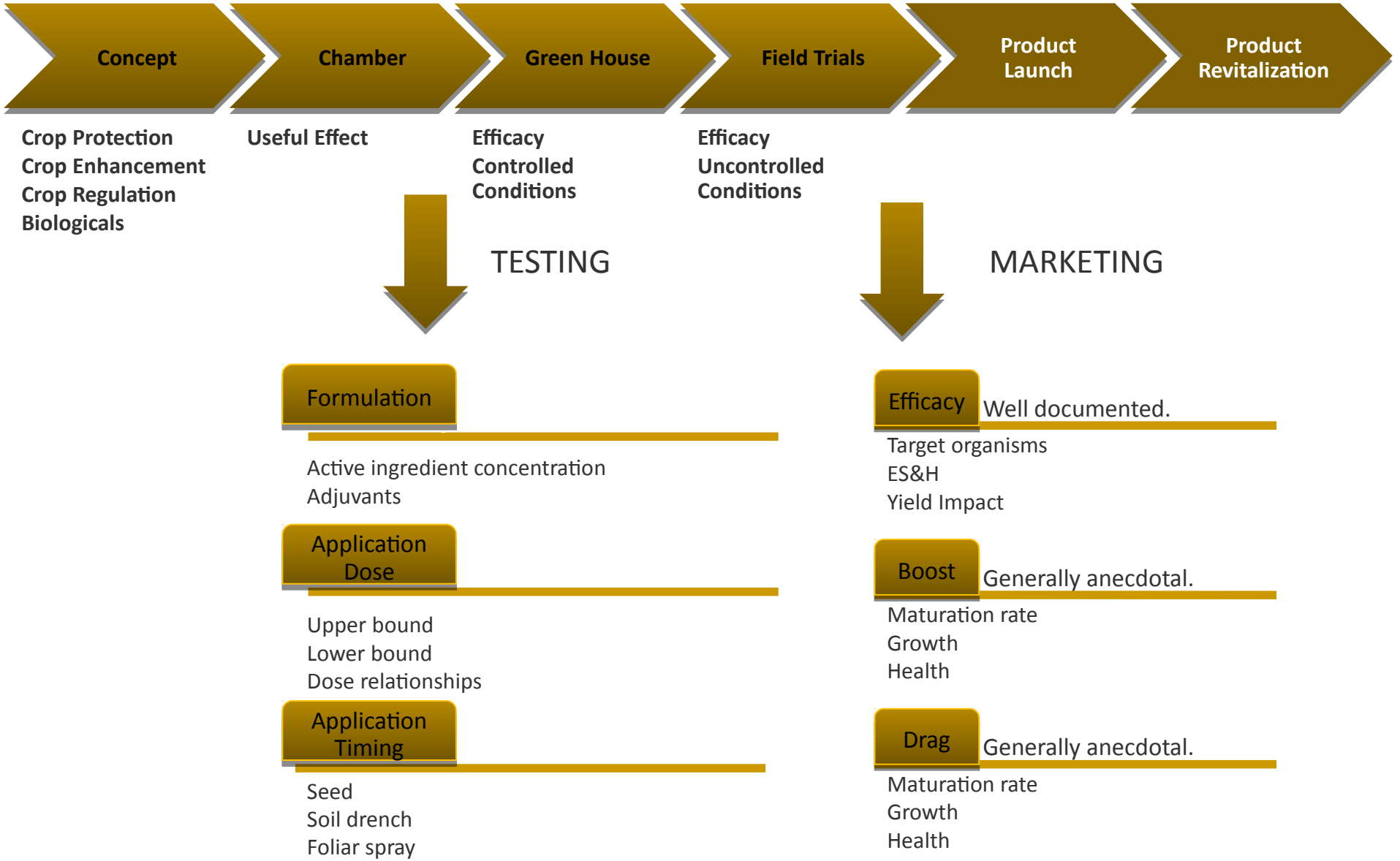


with surfactant A



with surfactant B

Phenotype Screening Corporation
Product Development Services



Integration of Knowledge in order to Better Predict the Future

- Multi-dimensional View of all the Data
- Decisions take into account All Available Knowledge
  - General Knowledge
  - Past Experience
  - System Properties
- Frequentist Statistics – Best Practices
- Bayesian Method