

Risk-Based Sampling: the PPQ Experience



Catherine S. Katsar¹, ByeongJoon Kim²
Robert Griffin¹, and Steve Hong²

¹USDA-APHIS-PPQ, 1730 Varsity Drive, Suite 400, Raleigh, NC, 27606 USA.

²Center for Integrated Pest Management, North Carolina State University, Raleigh, NC 27695, USA

Outline





United States Department of Agriculture

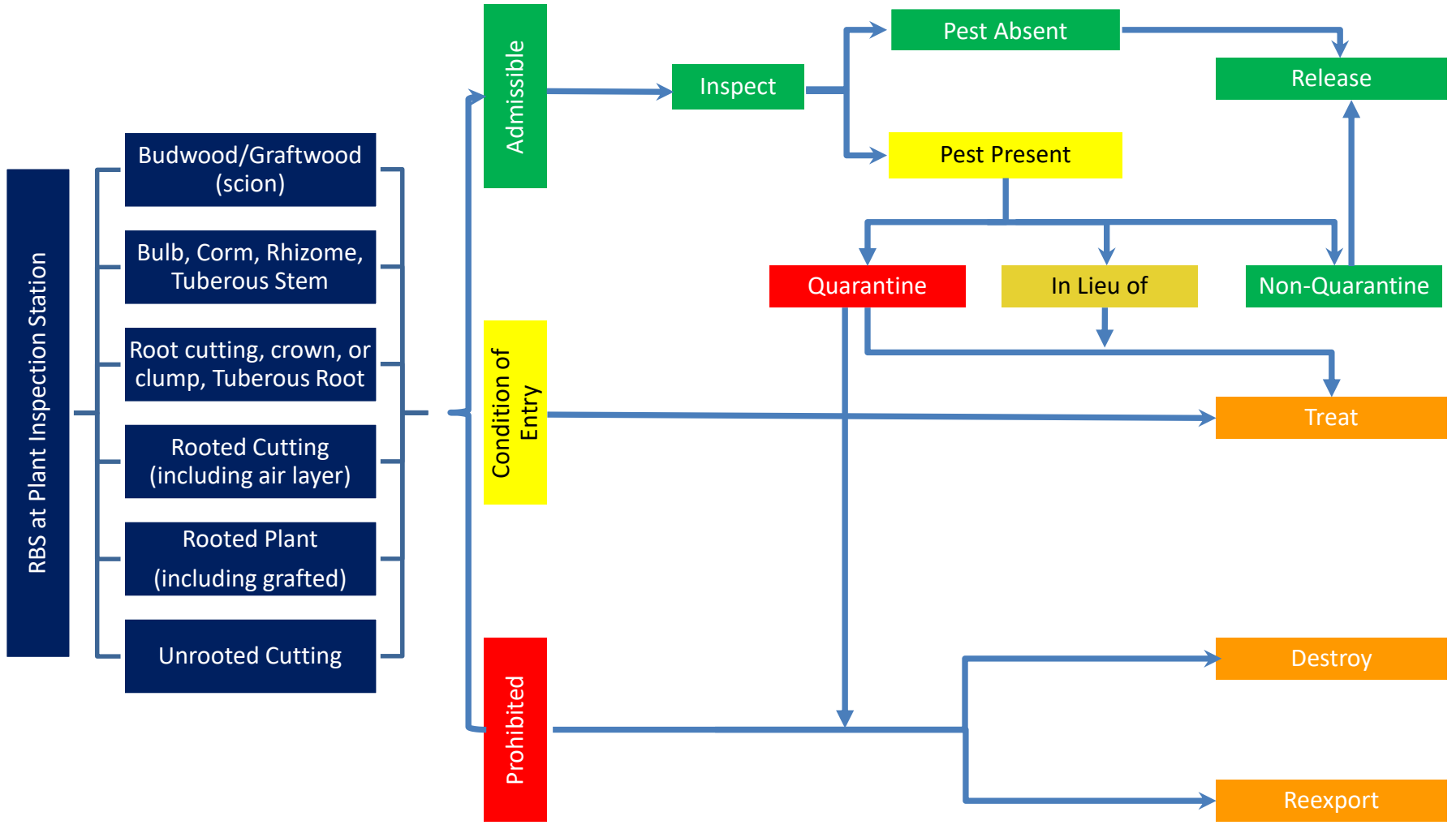
Data Availability & Quality



Safeguarding America's Agricultural and Natural Resources

United States Department of Agriculture | Animal and Plant Health Inspection Service | **Plant Protection and Quarantine**

Where does the data come from?



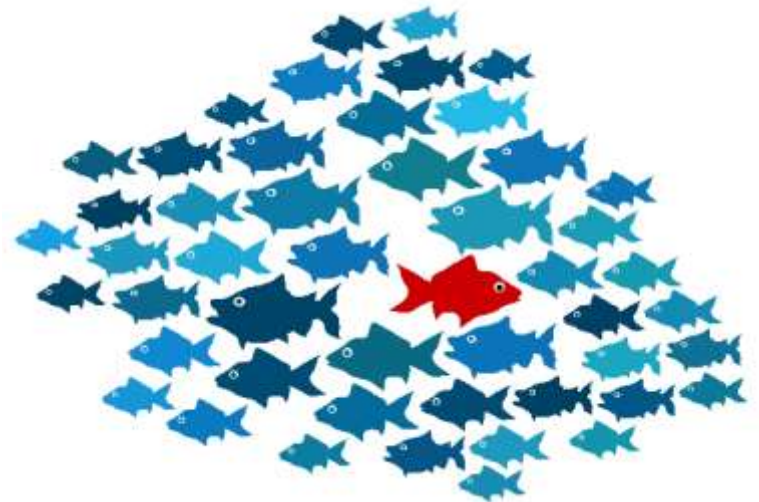
Requested RBS Data Elements



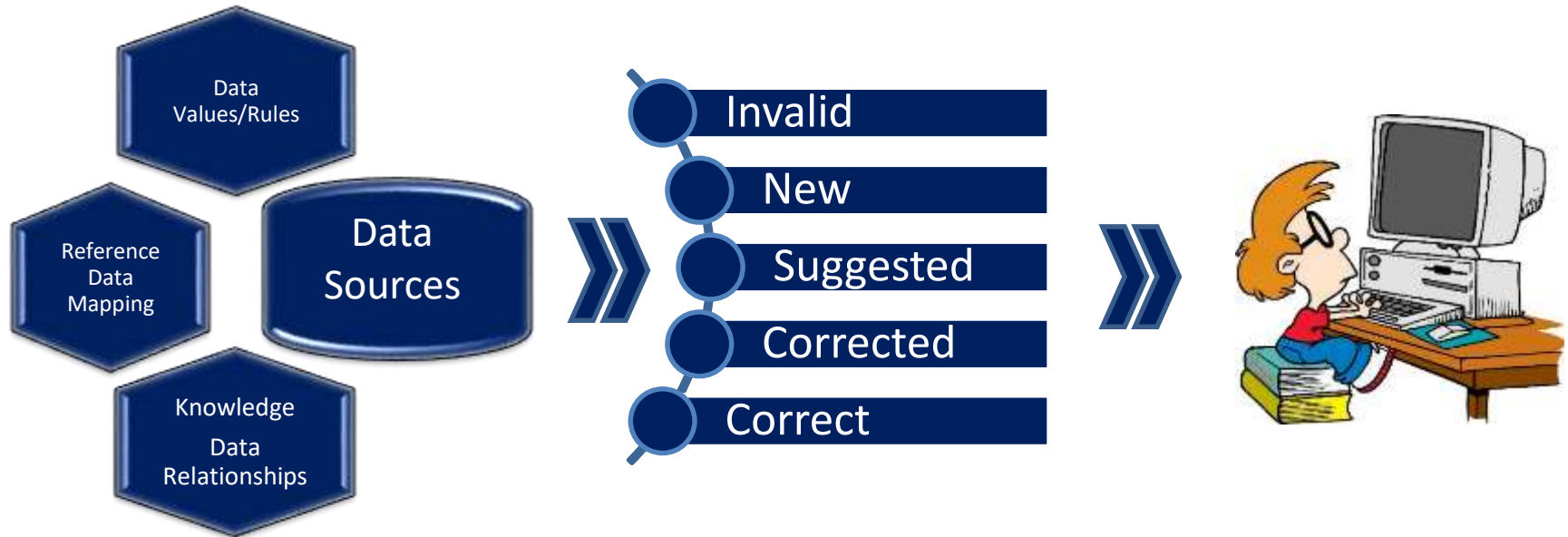
Data Elements	
Certified Facility Number	RBS Inspection?
Commercial or Non-Commercial	Regulatory Type
Farm Manufacturer Identification Number (MID)	Sample Unit - Amount
Harmonized Tariff Code	Sample Unit - Type
Inspectional and Sampling Units Comingled?	Size/Stage of Growth
Inspectional Unit - Amount	Species and/or Variety Detail
Inspectional Unit - Type	Time Inspection Completed
Propagative Material Type	Time Inspection Started

Common Sources of Data Inconsistency

- Quarantine Policy Effects
- Disposition codes
- Pathway Determination
- Host Plant
- Quantity
- Misalignment of Identification Codes among datasets



Data Quality Control and Assurance



COMMODITY	ORIGIN	PRE-QC ACTION RATE	POST-QC ACTION RATE	PERCENT CHANGE
Egeria	Indonesia	0.18	0.27	9.09%
Gardenia	Guatemala	0.10	0.17	6.90%
Aglaonema	Dominican Republic	0.02	0.04	1.75%
Liriope	Guatemala	0.03	0.04	1.22%
Renanthera	Taiwan	0.09	0.00	-9.09%
Calathea	Malaysia	0.08	0.00	-8.33%
Epidendrum	Taiwan	0.08	0.00	-8.33%
Plumeria	Mexico	0.08	0.00	-7.69%



United States Department of Agriculture

Sampling & Sample Size Calculator



Safeguarding America's Agricultural and Natural Resources

United States Department of Agriculture | Animal and Plant Health Inspection Service | **Plant Protection and Quarantine**

Comparing 2% and Hypergeometric

- Inconsistent level of detection
- Action rates, approach rates, and infestation rates are not true and cannot be used to demonstrate effectiveness
- No consistent level of risk management
- Limited analytical value
- Frequently misunderstood and misused

2%

Inspect-O-Matic® Summary Report: Podunk/July 22, 2014

Shipment No.	Confidence	Lot size	Sample Rate (%)	Sample size	Detection level	Inspection time (hrs)
001	.95	500 boxes	2	10	>.20	<1
002	.95	1,000 boxes	2	20	.14	<1
003	.95	33,000 plants	2	660	.004	2
004	.95	100 boxes	2	2	>.20	<1
005	.95	5,000 boxes	2	100	.03	<1
006	.95	200 boxes	2	4	>.20	<1
007	.95	20,000 boxes	2	400	.008	1
008	.95	70,000 flasks	2	1400	.002	4.5
009	.95	10 boxes	2	1	>.20	<1
010	.95	600 boxes	2	12	>.20	<1

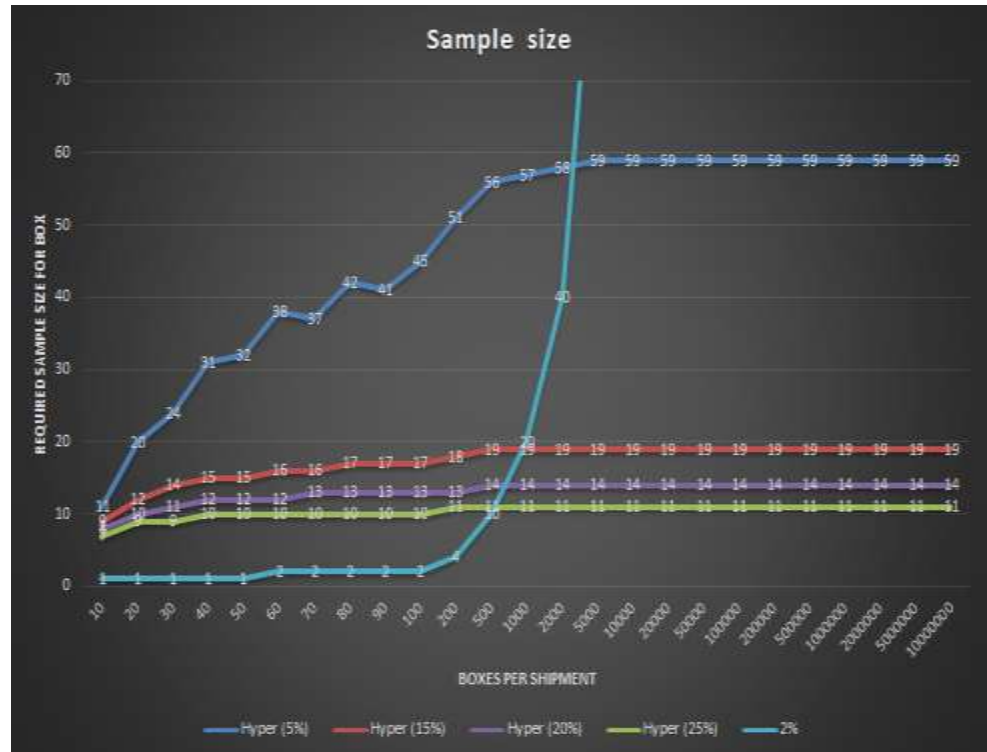
Hypergeometric

Inspect-O-Matic® Summary Report: Podunk/July 23, 2014

Shipment No.	Confidence	Lot size	Sample Rate (%)	Sample size	Detection level	Inspection time (hrs)
001	.95	500 boxes	11	56	.05	<1
002	.95	1,000 boxes	6	57	.05	<1
003	.95	33,000 plants	0.1	59	.05	<1
004	.95	100 boxes	47	45	.05	<1
005	.95	5,000 boxes	1	59	.05	<1
006	.95	200 boxes	25	51	.05	<1
007	.95	20,000 boxes	0.2	59	.05	<1
008	.95	70,000 flasks	0.08	59	.05	<1
009	.95	10 boxes	100	10	.05	<1
010	.95	600 boxes	1	56	.05	<1

Sample Size

- Account for inspection intensity / frequency
- Statistically Robust (e.g., randomness)
- Operationally feasible
- Calculates sample size and identifies systematic-random sample units.



Potential Parameters

- Infestation level
- Confidence level



United States Department of Agriculture

Statistical Analysis & Risk Ranking



Safeguarding America's Agricultural and Natural Resources

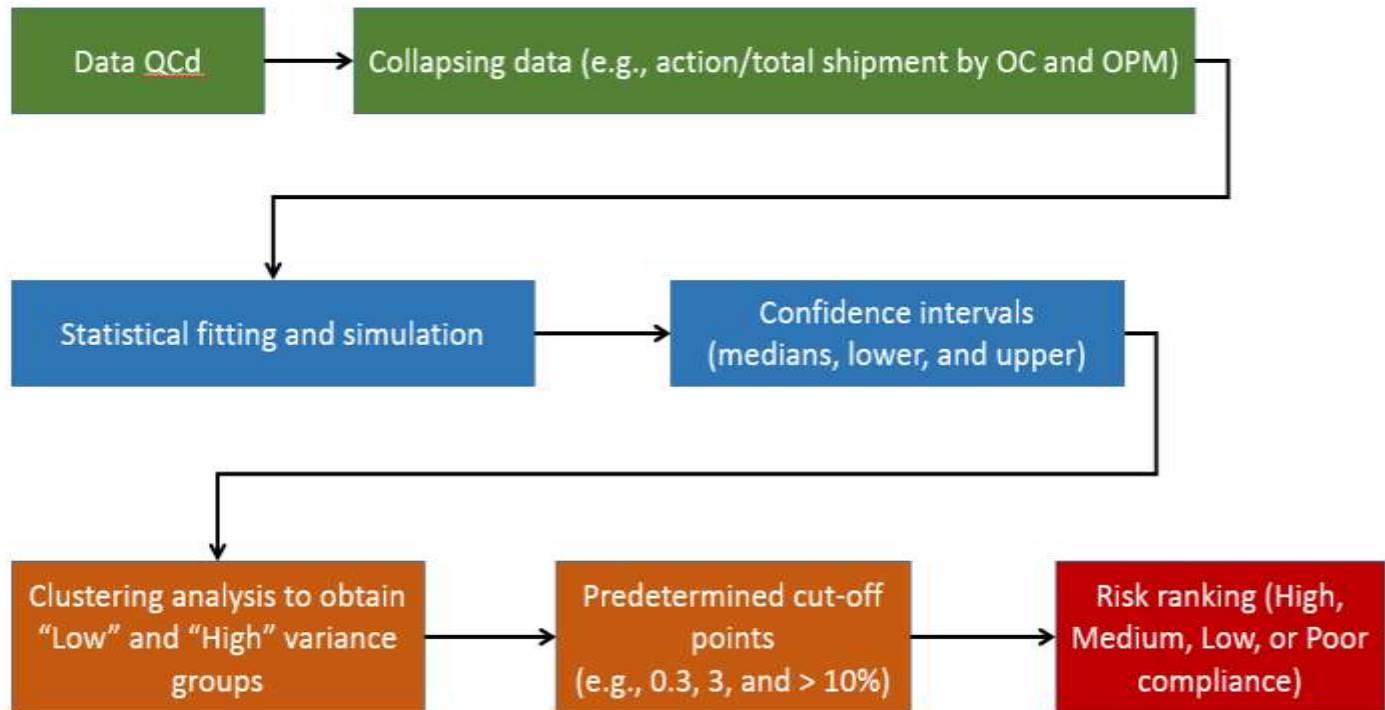
United States Department of Agriculture | Animal and Plant Health Inspection Service | **Plant Protection and Quarantine**

Sample Size Calculator

- Sample size can be estimated based on risk ranking
 - Hypergeometric distribution (95% confidence level)
 - Intensity
 - More boxes for low compliance shipment
 - Less boxes for high compliance shipment

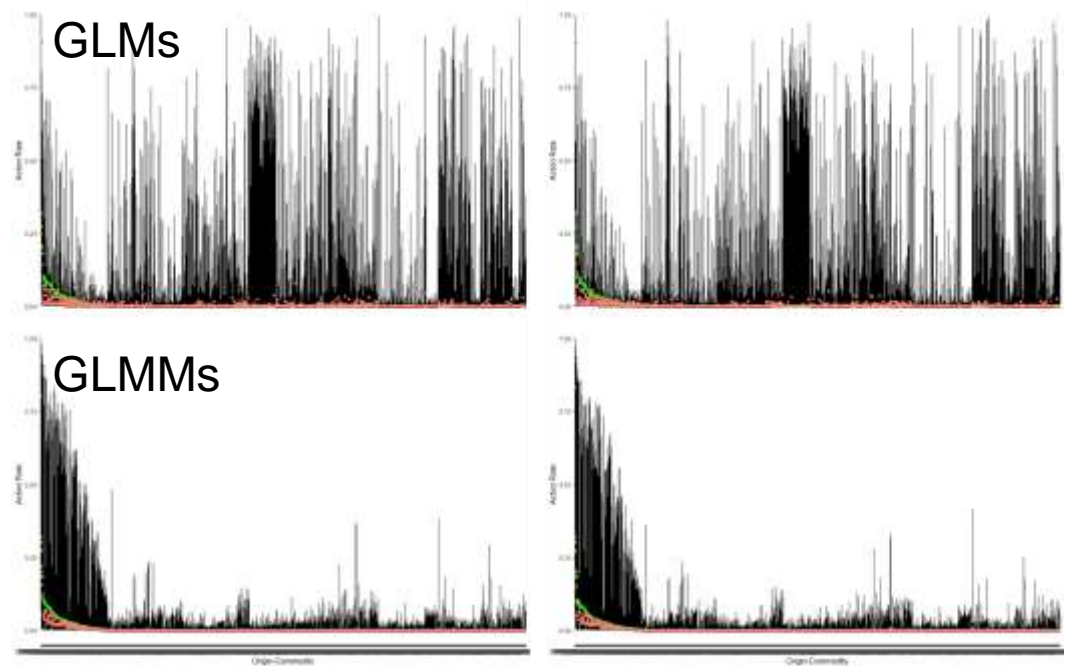
Hypergeometric Sampling Tool	
Total boxes or bag	150
Total plant quantity	500000
Confidence level for box	95%
Origin	CHN
PM type	Rooted Cutting (including air layer)
Analysis Output	
OPM	CHN_Rooted Cutting (including air layer)
PCL	High compliance
Infestation Rate	25%
Required box number for inspection	10

Statistical Analysis Procedures



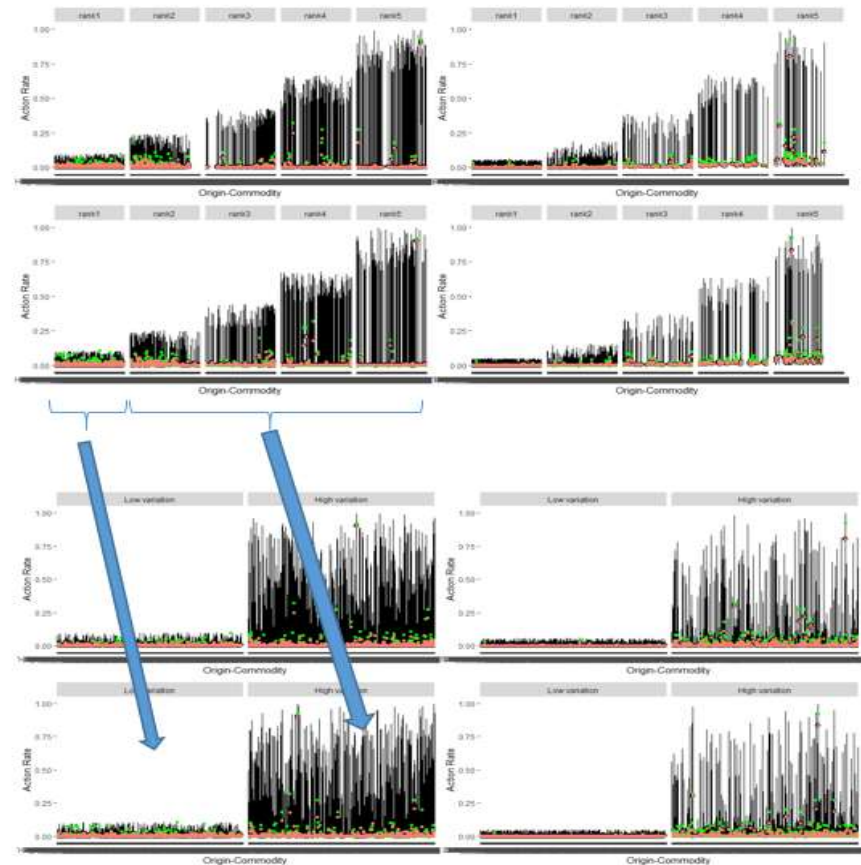
Data Analysis: Predicted and Observed Action Rate

- Origin-Commodity combos
 - Likelihood of carrying quarantine pests
 - Action rates
- Different types of statistical models were tested to predict action rate.
 - Bayesian generalized linear regression (GLM)
 - Generalized linear mixed-effects models (GLMM)



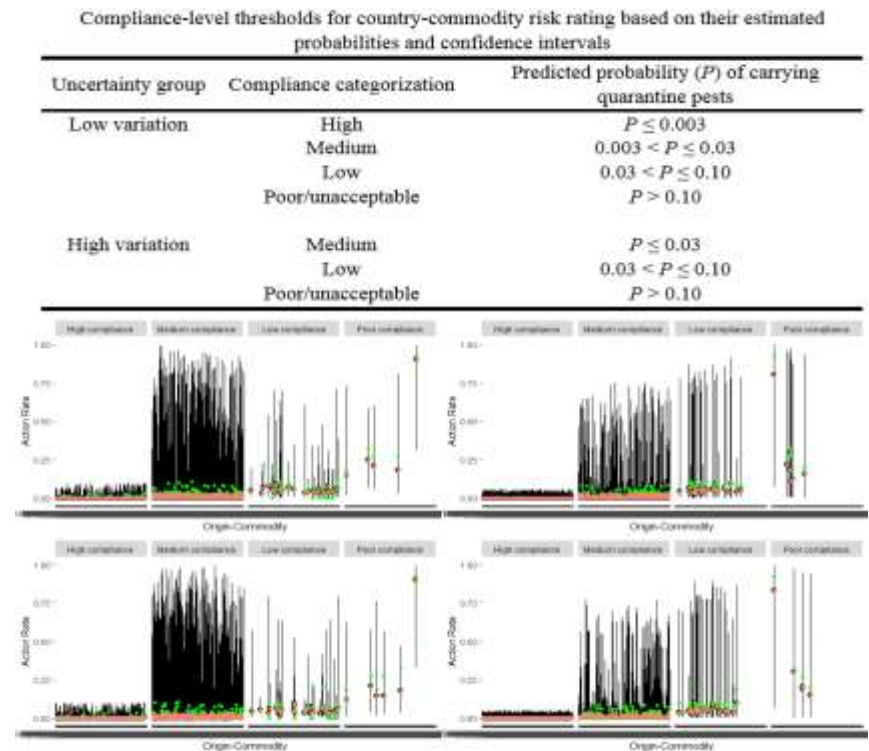
Data Analysis: Unsupervised Learning - Clustering

- Using confidence interval range, origin-commodity combos were categorized to groups (4 plots in top right)
- Low and high variability
 - Narrower range: low variability in predicted action rate
 - Wider range: high variability in predicted action rate

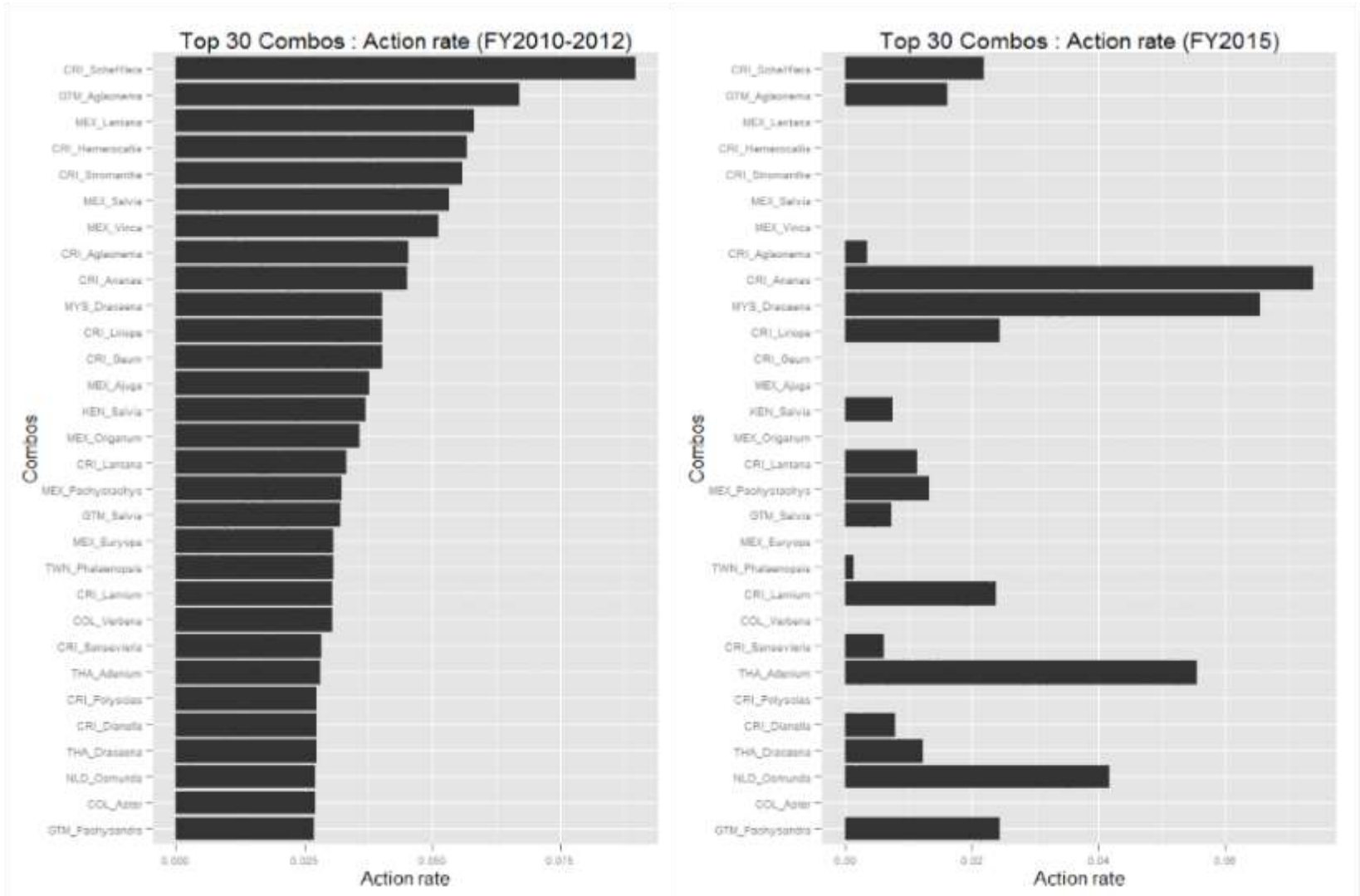


Risk Ranking

- Grouping combos with action rates and their uncertainties
- High compliance combos only in low variation group
- How to determine them is the most challenging part (needed?)
- Need to consider impact of slippage given thresholds on environmental and/or economic damage (on-going issues)
- Predetermined cut-off points (0.3, 3, and 10%); can be adjustable to achieve operational feasibility



Comparing Inspections





United States Department of Agriculture

Operational Challenges & Enforcement Framework



Safeguarding America's Agricultural and Natural Resources

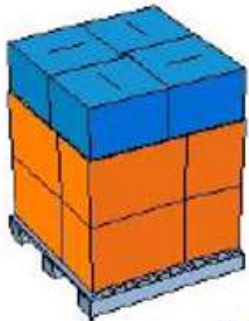
United States Department of Agriculture | Animal and Plant Health Inspection Service | **Plant Protection and Quarantine**

Plant Material Type



Commingling of Plant Genera Occurs at Multiple Levels

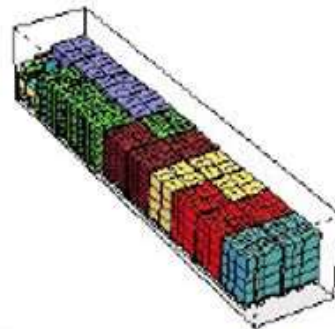
- **Shipment Level** – Imported plant genera mixed within a container or on a Bill
- **Sampling Unit Level** – Imported plant genera mixed within a carton/box/baggie or other sampling unit of measure



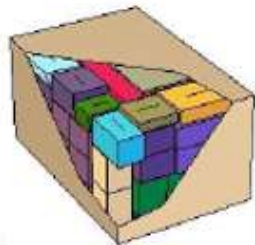
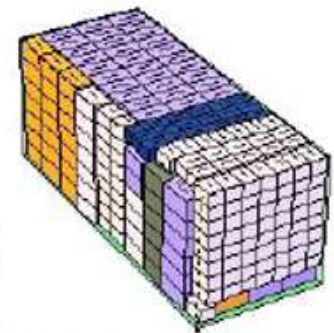
Pallet



ULD (Air pallet)



Container/Truck



Carton



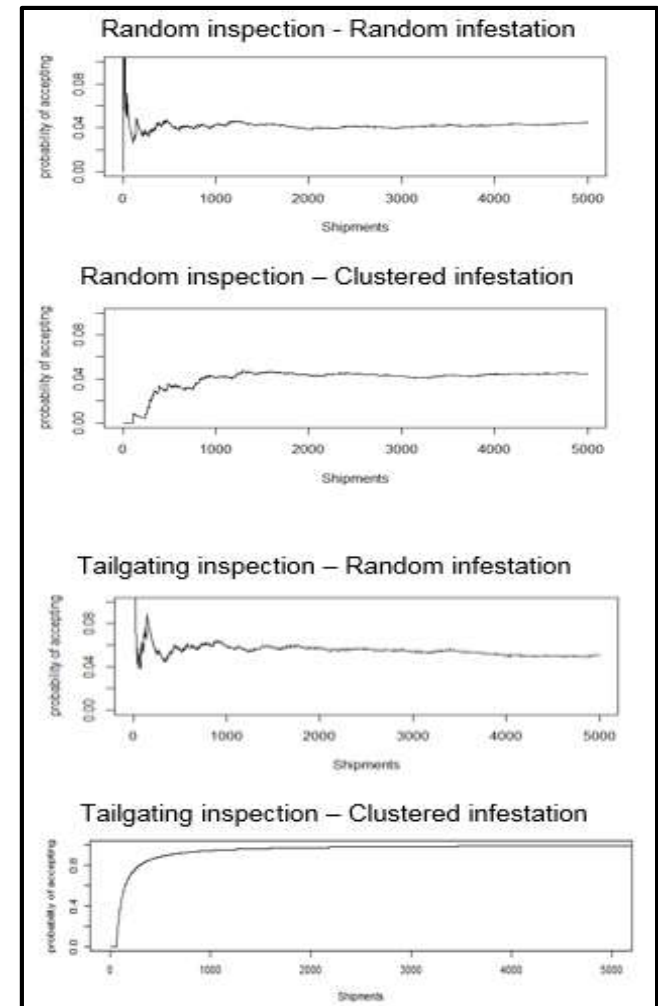
Importance of Randomness: Random vs Tailgate Inspection

- Random inspection

- The computer picked boxes randomly for inspection.
- Probability of accepting randomly infested shipment is converged at 5%.
- Similarly, probability of accepting clustered-infested shipment is converged at 5%.

- Tailgating inspection

- The computer picked boxes at 'tail' part of a shipment for inspection.
- Probability of accepting randomly infested shipment is converged at 5%.
- **However**, probability of accepting clustered-infested shipment is converged at 99%.

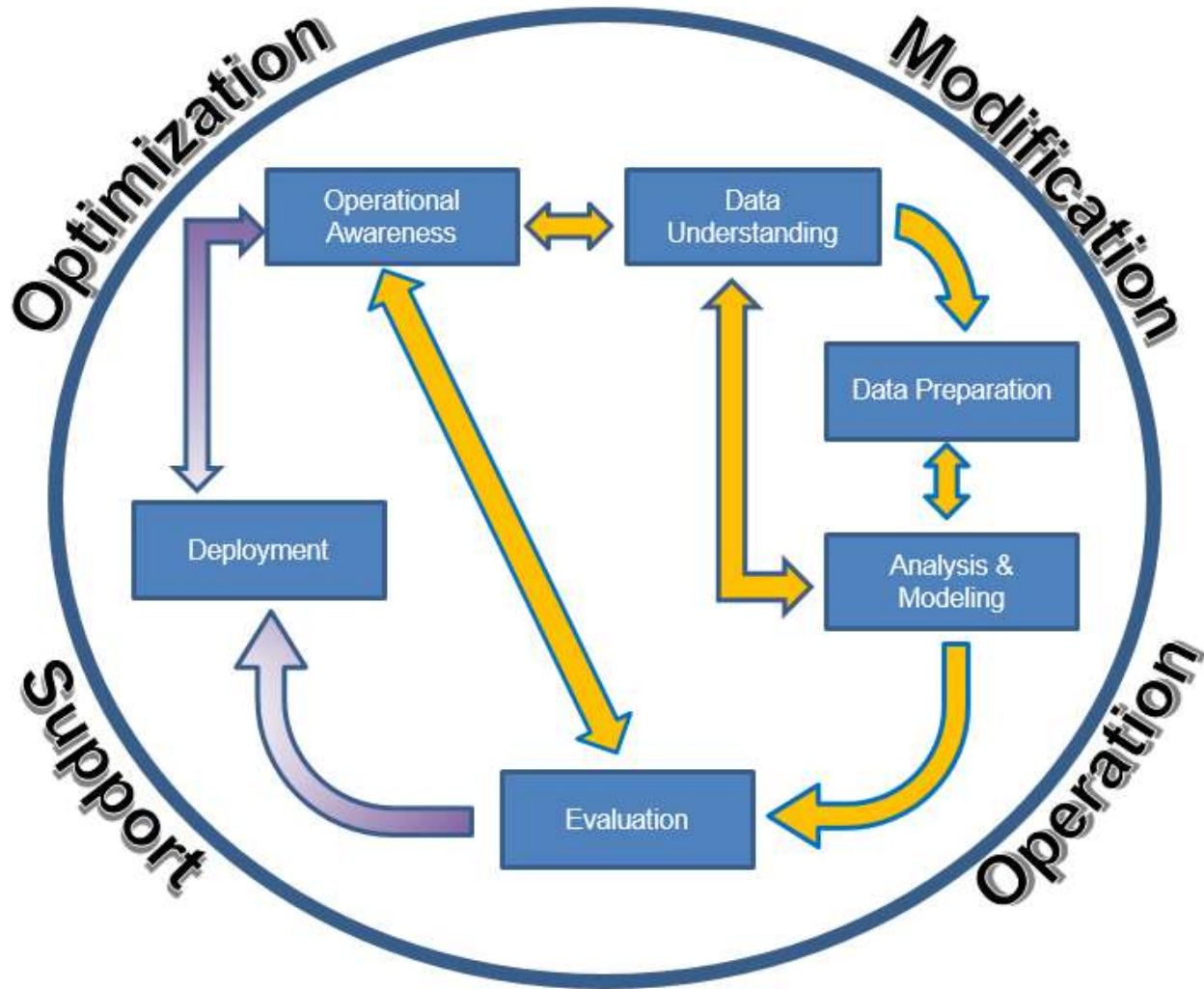


Deployment & Enforcement Framework

- Compliance level
- Confirm essential data fields
- Adjust sampling tool and analytical process
- Set standard operational rules and procedures
- Identify reward(s) and distribution
- Establish structure, processes and standard operating procedures for:
 - Operational guidelines
 - Analytical support
 - Overall system management



Summing up the RBS System



Acknowledgements

- Special thank you to Marla Cazier Mosley and Dave Farmer, USDA-APHIS-PPQ Field Operations for their continued guidance and support.
- Valuable assistance provided by Barney Caton (PERAL) and Karl Suiter (CIPM).
- Thank you to all the PPQ Plant Inspection Stations for their cooperation and support.

