LT2 Round 1 Monitoring DCTS Data and Calculated Bin Results

Alexa Obolensky, O.C.
Steve Via, AWWA
Mike Hotaling, Newport News Waterworks

EPA Public Meeting November 15, 2012

Long Term 2 Enhanced Surface Water Treatment Rule

"Monitoring Data Analysis, Occurrence Forecasts, Binning, and the Microbial Toolbox"

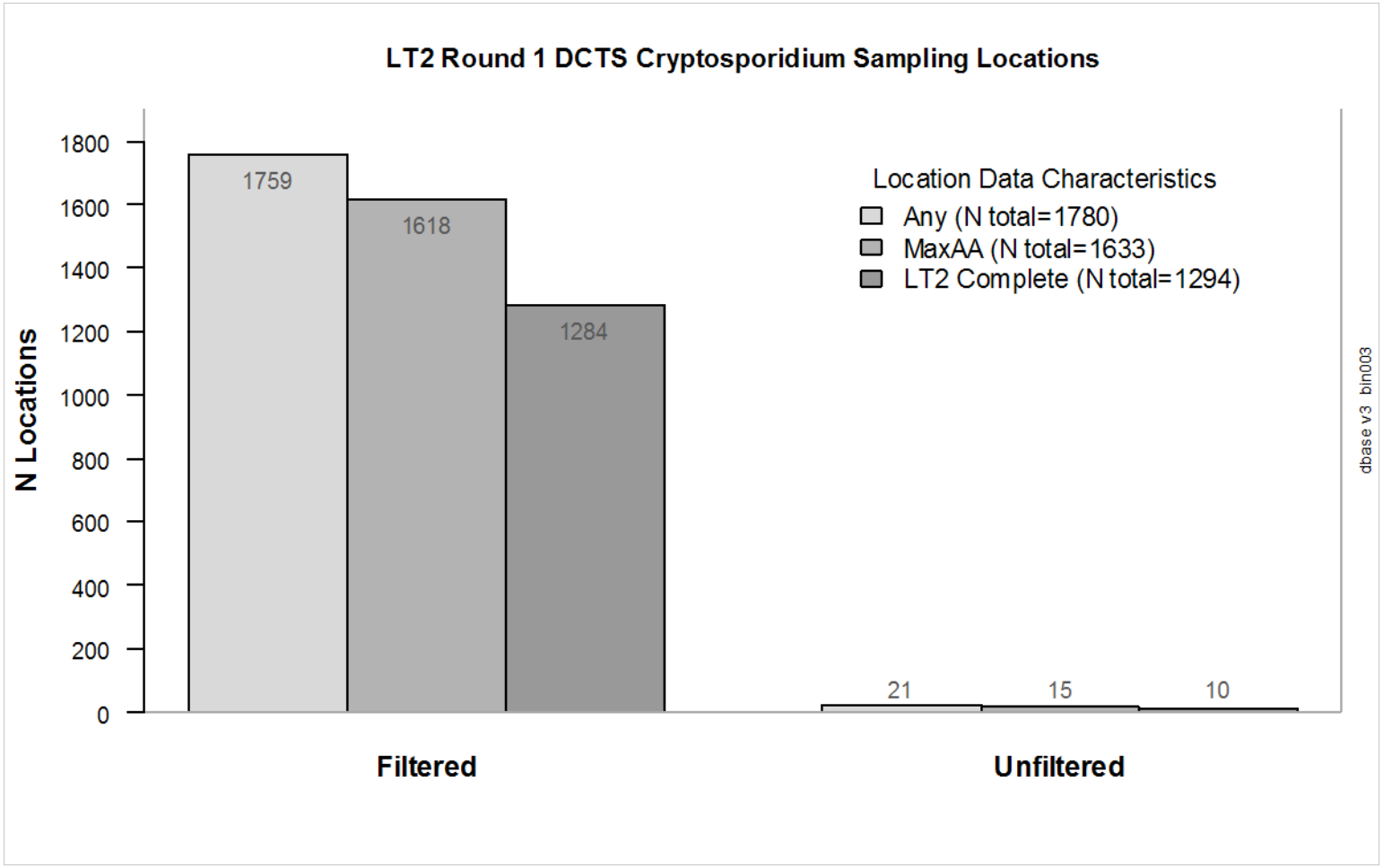
Presentation Outline

- Overview of DCTS Data
- Sampling and Analytical Results
- Calculated Bins
- *Cryptosporidium* Recovery
- *E*. coli /*Cryptosporidium* Relationship

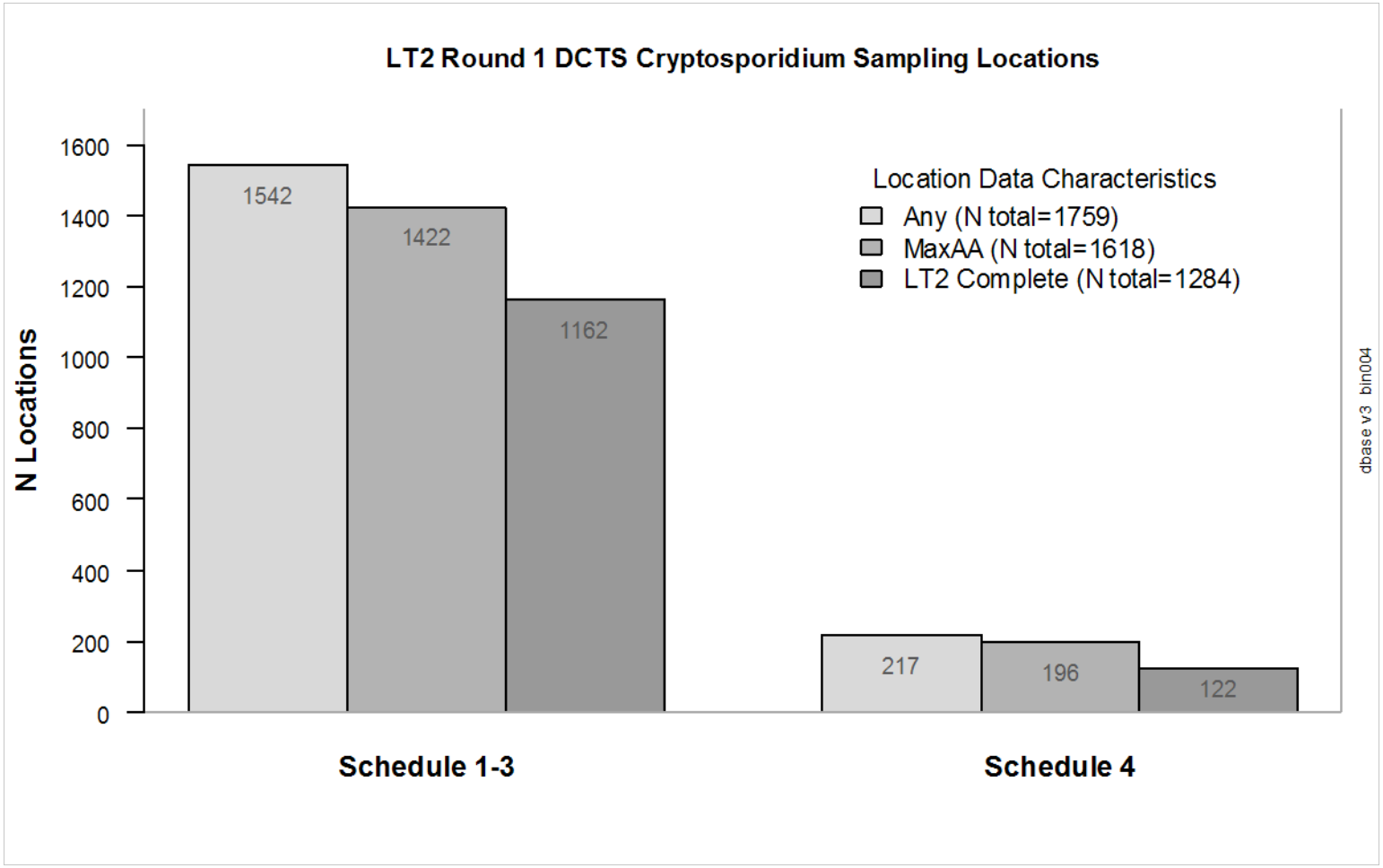
DCTS Data

- **Data Collection and Tracking System**
- 2280 sampling locations (1780 Crypto, 2146 E. coli)
- electronic submission required for large systems (schedule 1-3)
- sparse representation of small systems (schedule 4)
- grandfathered data not included (~ 900 locations)
- *Cryptosporidium* and E. coli public data posted June 2012
- examination showed need for data cleaning and updating
  - sampling point attributes differ between sampling events
  - redundant sample entries
  - conflicting or missing source water category information
  - filtration status misclassification
  - E. coli calculated results missing
  - miscellaneous erroneous or suspect data entries
- “cleaned data” posted October 2012 with documentation

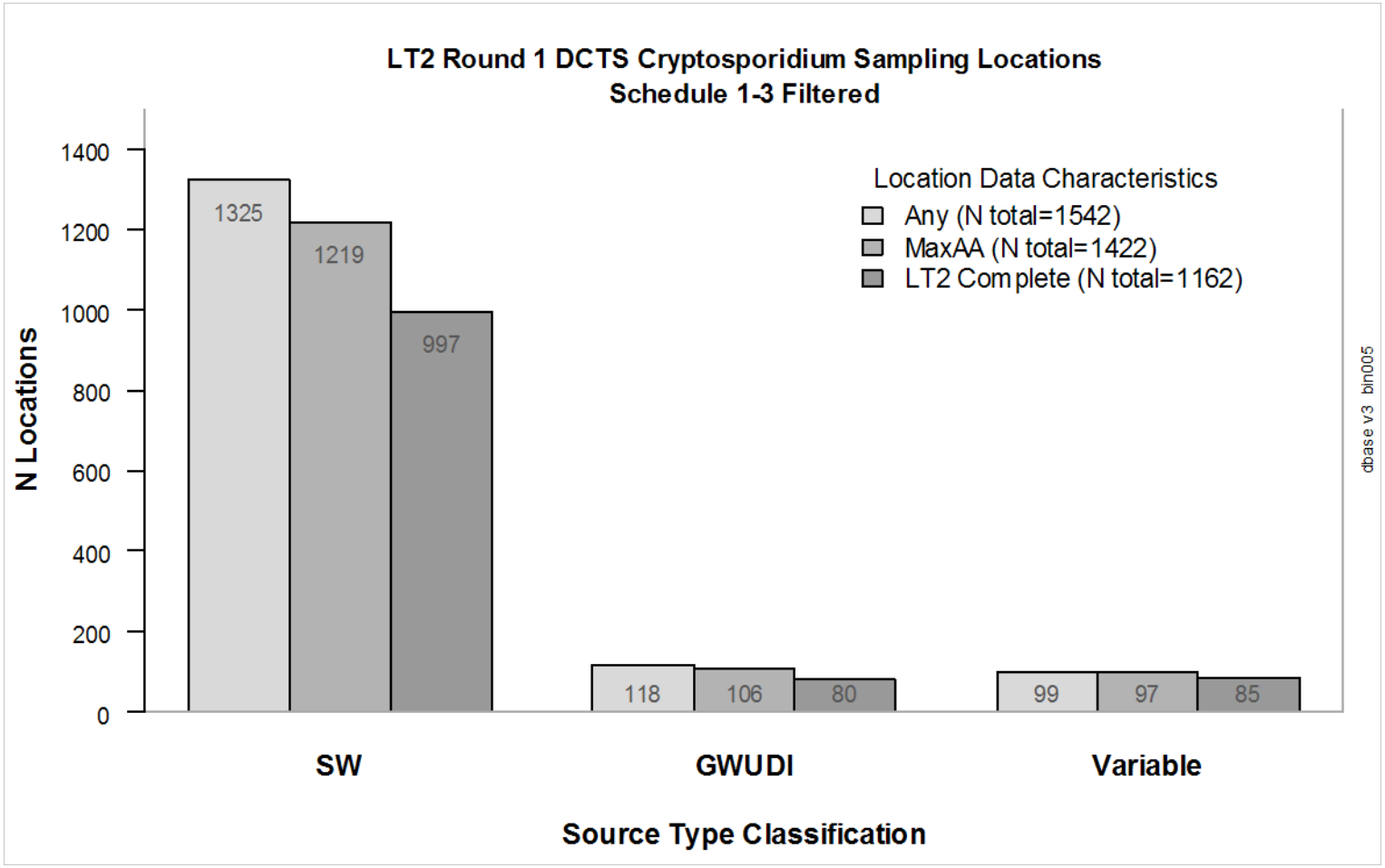
Cryptosporidium monitoring locations in DCTS data



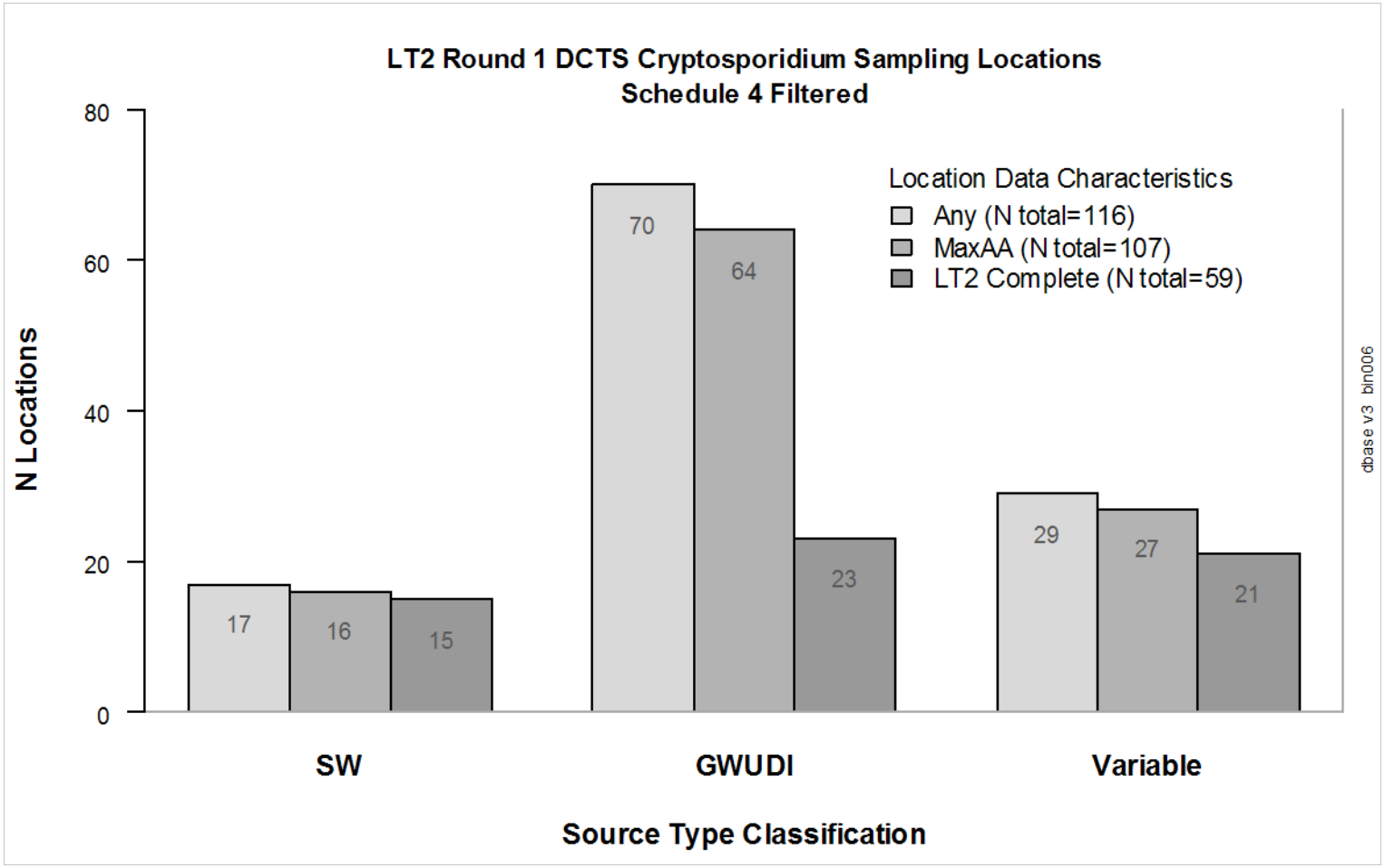
Cryptosporidium monitoring locations at Filtered Facilities



Cryptosporidium monitoring locations at Filtered Facilities Schedule 1-3 (large systems)

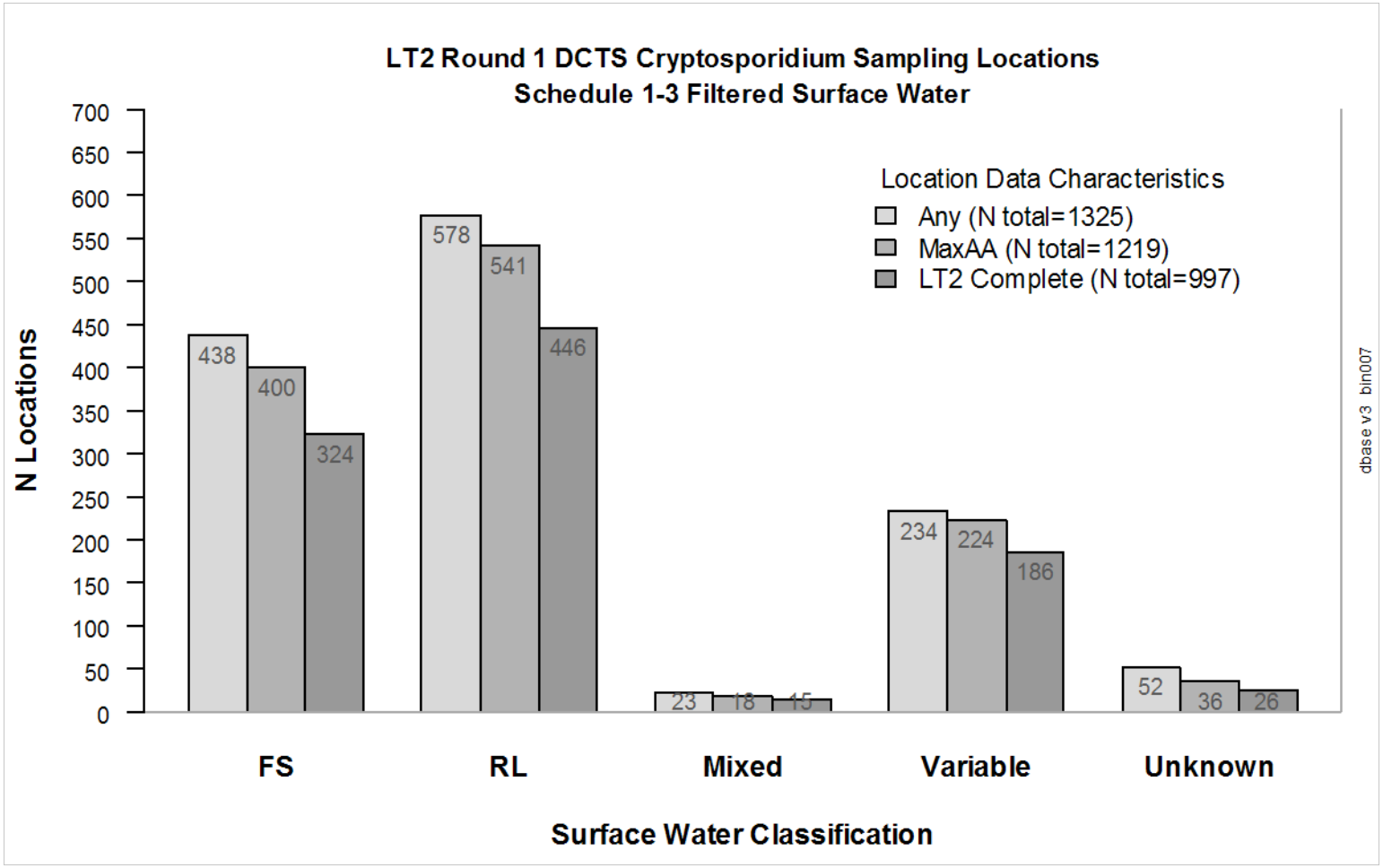


Cryptosporidium monitoring locations at Filtered Facilities Schedule 4 (small systems)



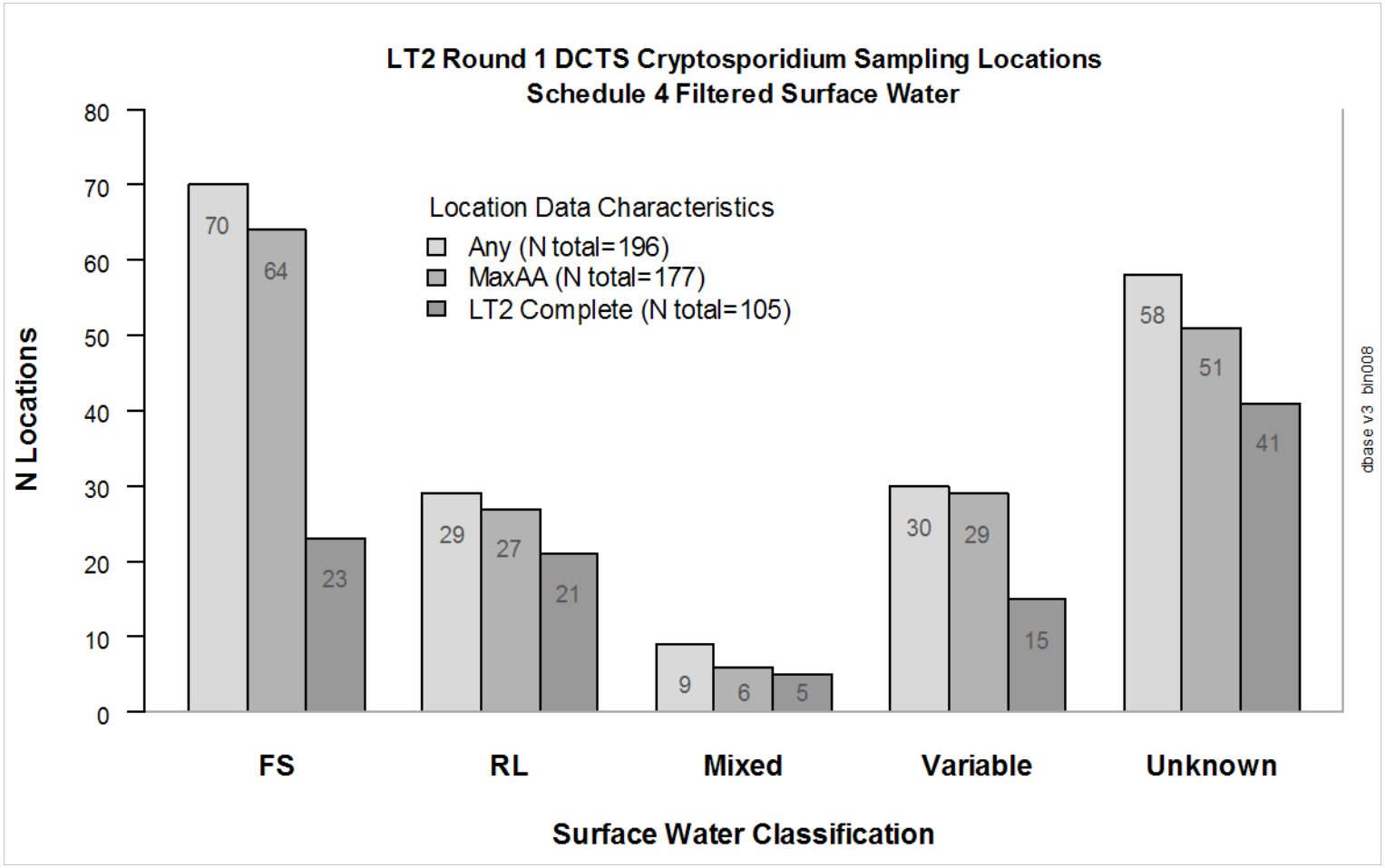
Cryptosporidium Surface Water monitoring locations

Large System (Schedule 1-3) Filtered Facilities

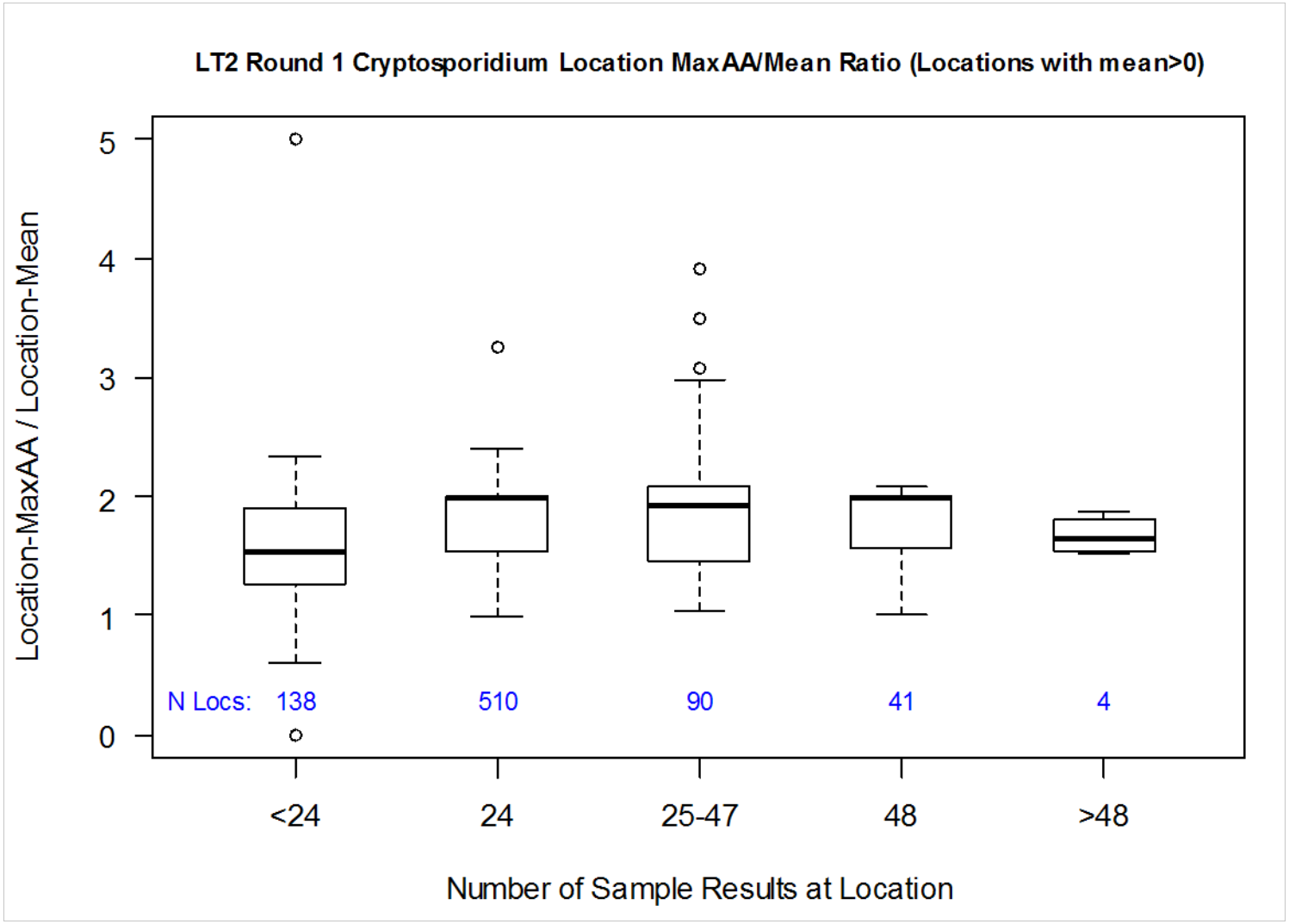


Cryptosporidium Surface Water monitoring locations

Small System (Schedule 4) Filtered Facilities

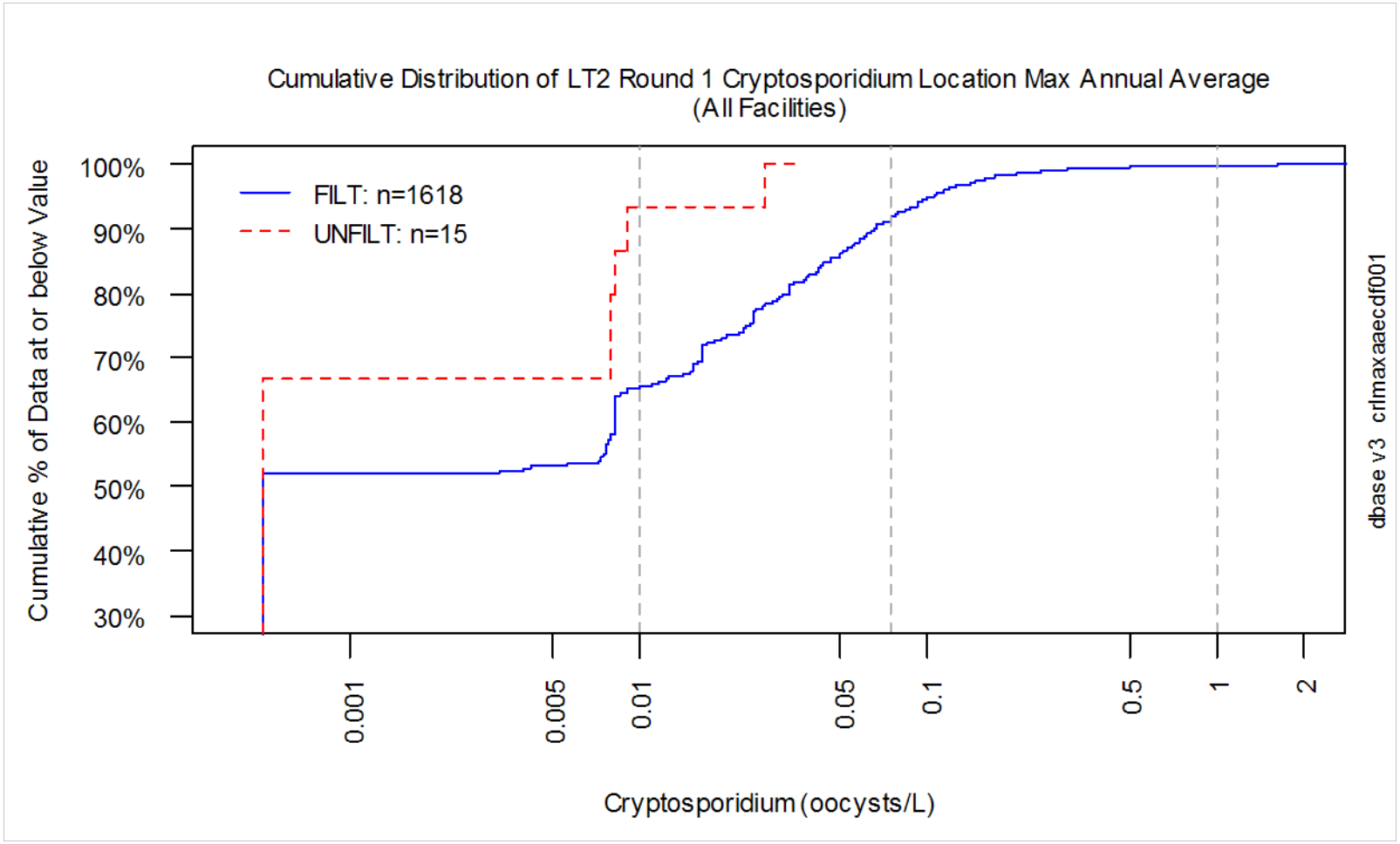


Location MaxAA or Location Mean?
Ratio median near 2
Regulatory Impact Will Differ Substantially



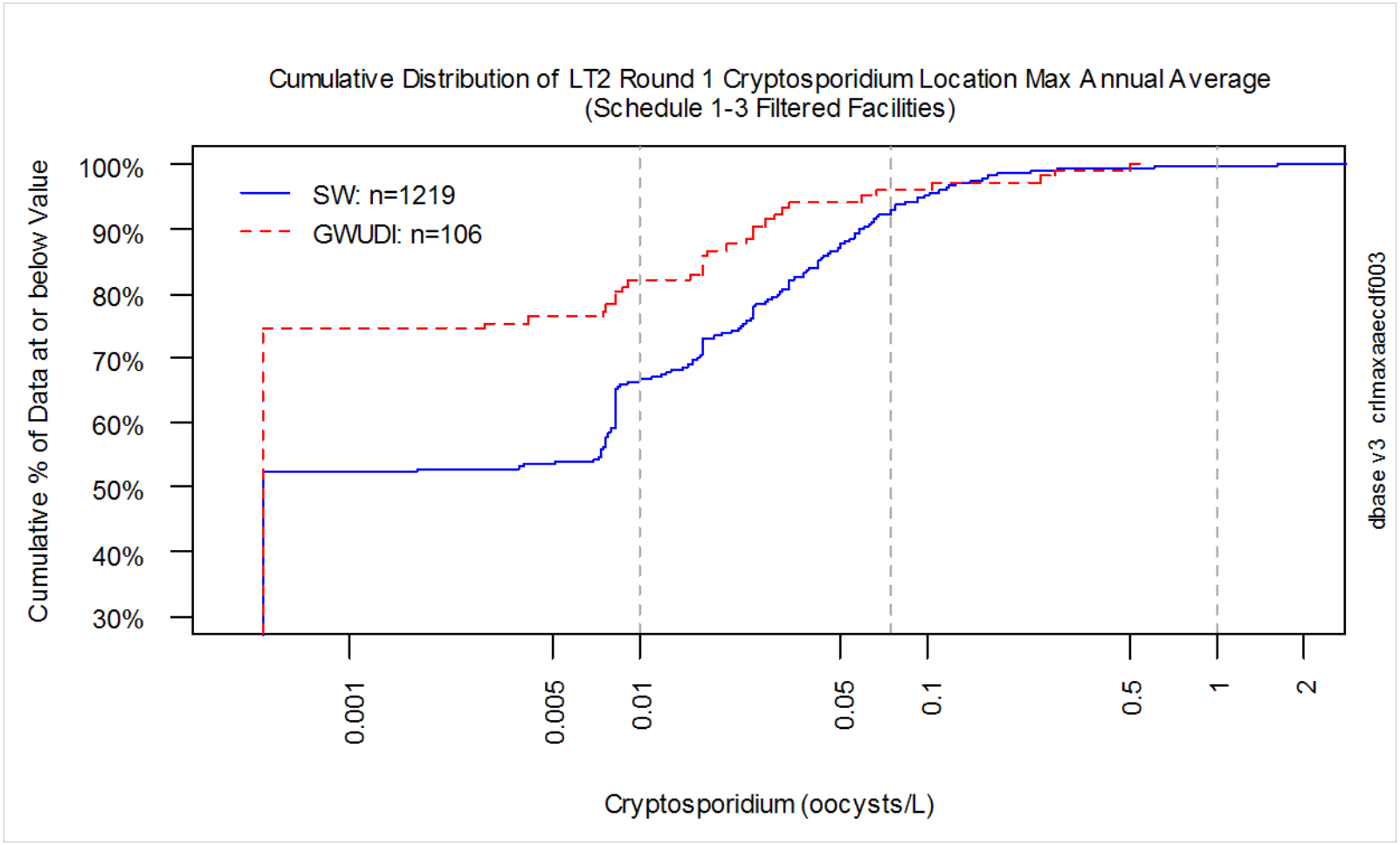
Filtered vs. Unfiltered
All Locations

Cumulative Distribution of Location MaxAA *Cryptosporidium*



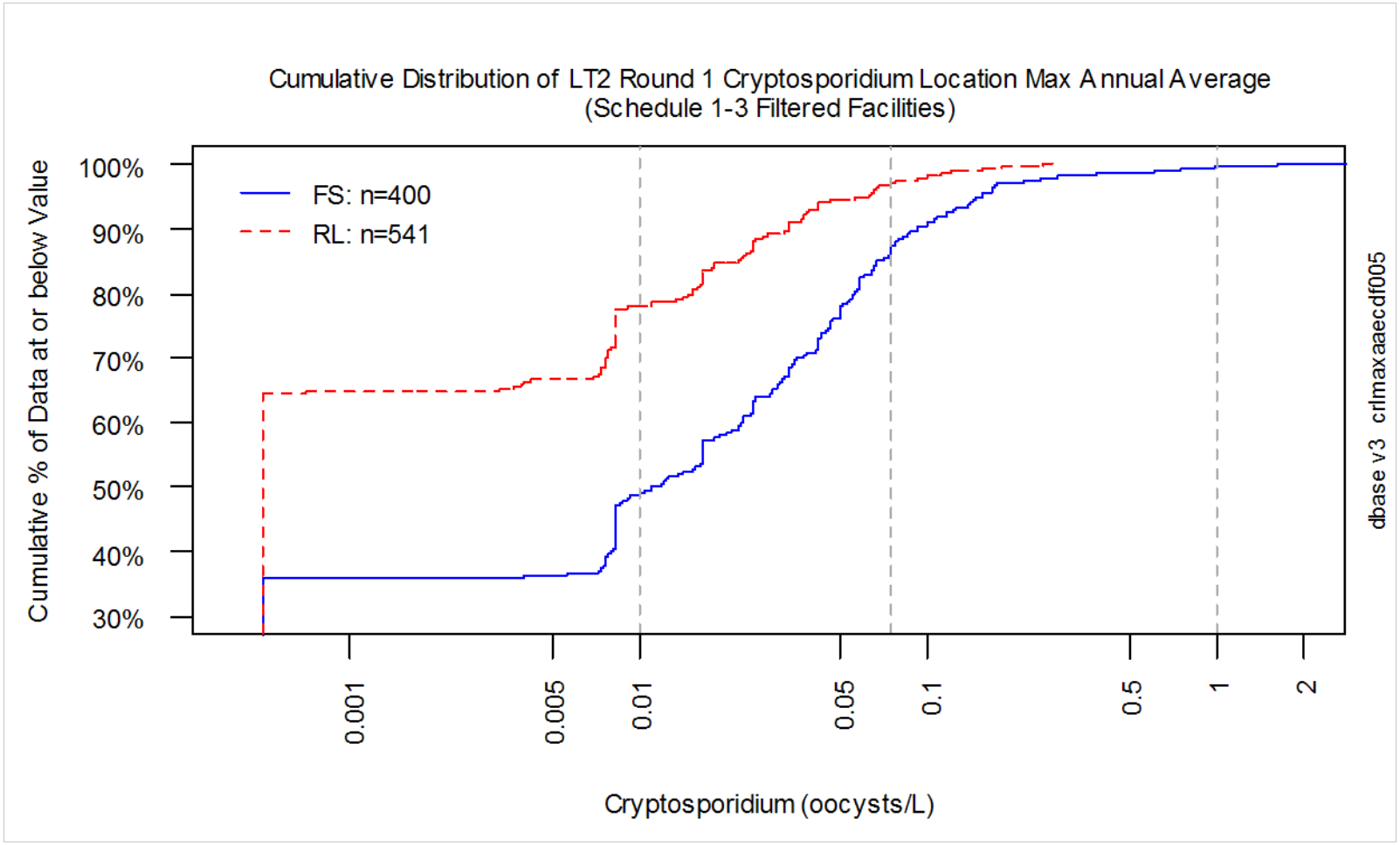
SW vs. GWUDI
Large System Filtered

Cumulative Distribution of Location MaxAA *Cryptosporidium*



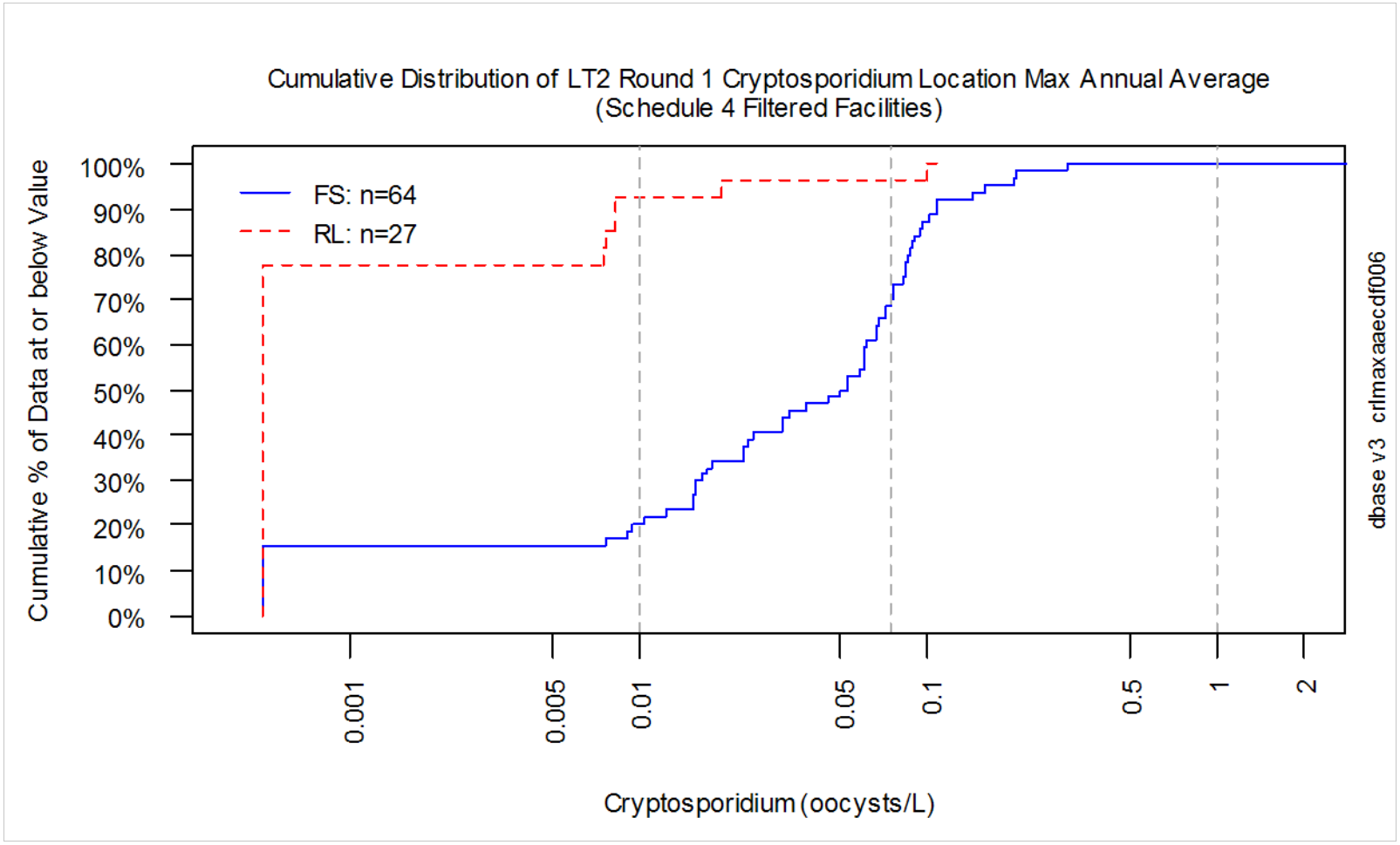
Flowing Stream vs. Reservoir/Lake
Large System Filtered Surface Water

Cumulative Distribution of Location MaxAA *Cryptosporidium*



Flowing Stream vs. Reservoir/Lake
Small System Filtered Surface Water

Cumulative Distribution of Location MaxAA *Cryptosporidium*



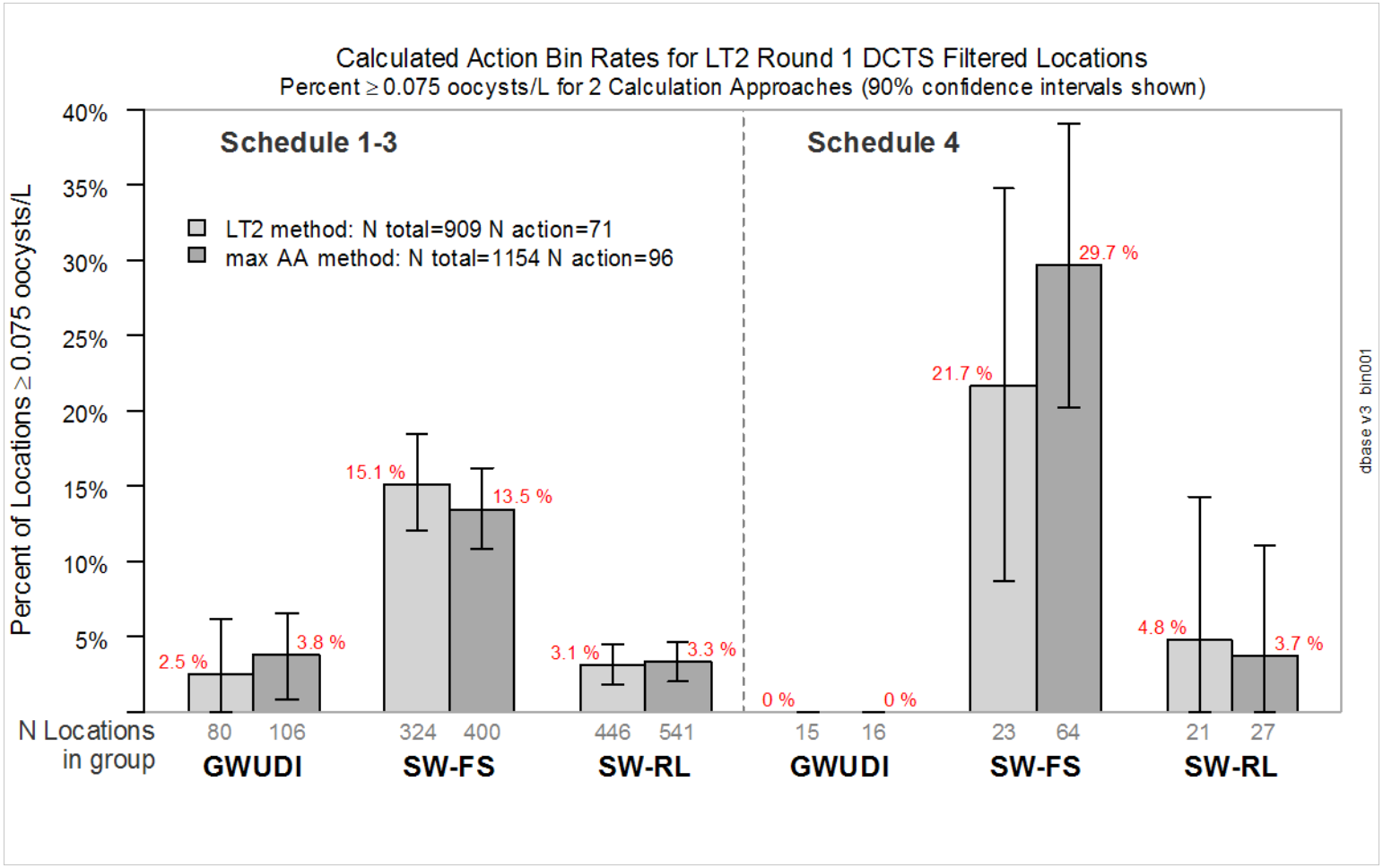
Bin Calculation from DCTS Data

Schedule 1-3 (large system) Filtered Facilities

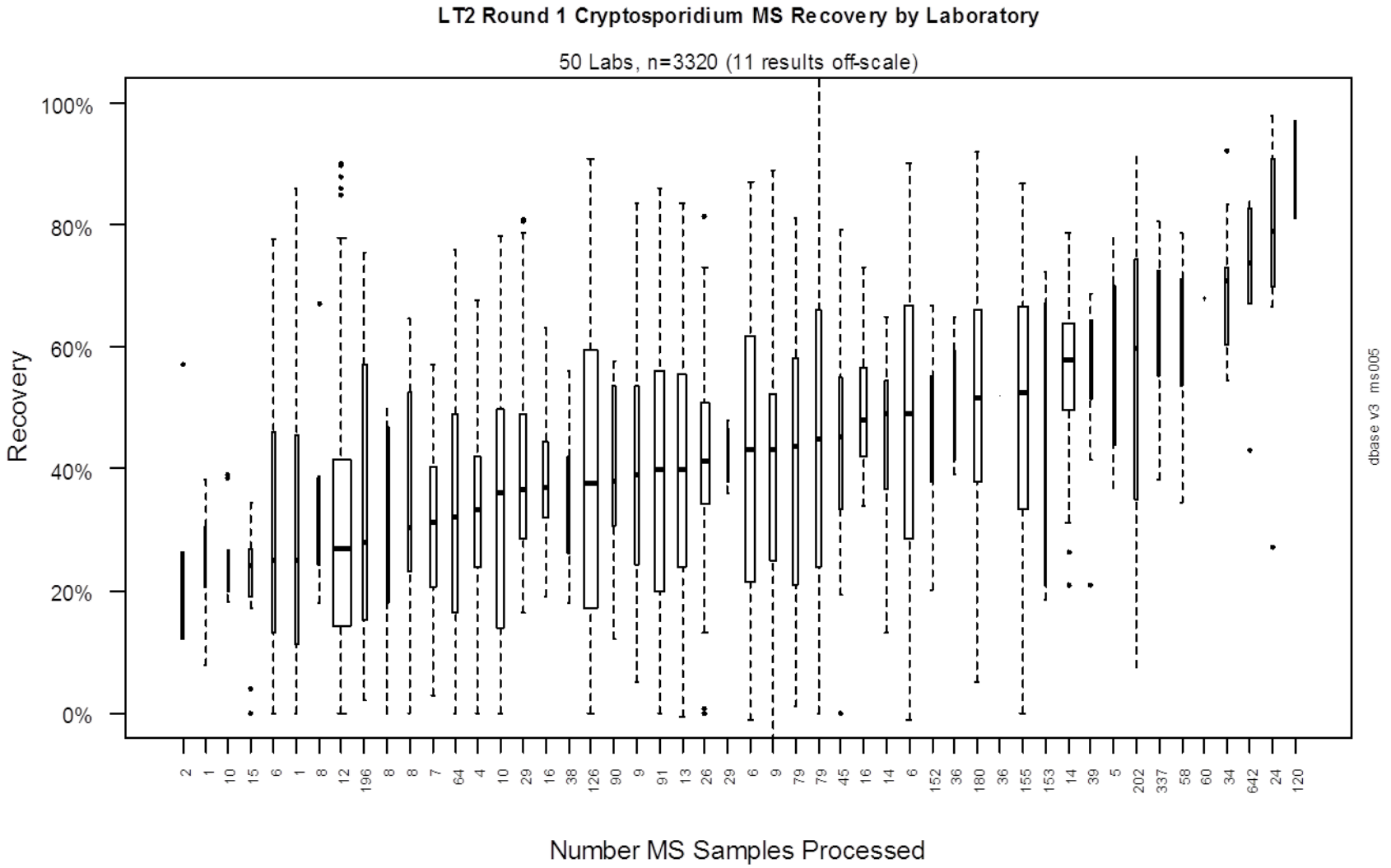
Location Type	N Total	N Binned	N (%) with any Detected Crypto	N (%) Bin 1	N (%) Bin 2	N (%) Bin 3
All	1542	1162	576 (50%)	1079 (93%)	83 (7%)	0
GWUDI	118	80	19 (24%)	78 (98%)	2 (2%)	0
SW	1325	997	505 (51%)	921(92%)	76 (8%)	0
Source Type Variable	99	85	52 (60%)	80 (94%)	5 (6%)	0
SW-FS	438	324	224 (69%)	275 (85%)	49 (15%)	0
SW-RL	578	446	168 (38%)	432 (97%)	14 (3%)	0
SW-uncertain	309	227	113 (50%)	214 (94%)	13 (6%)	0

Alternative Bin Calculations for Incomplete Data

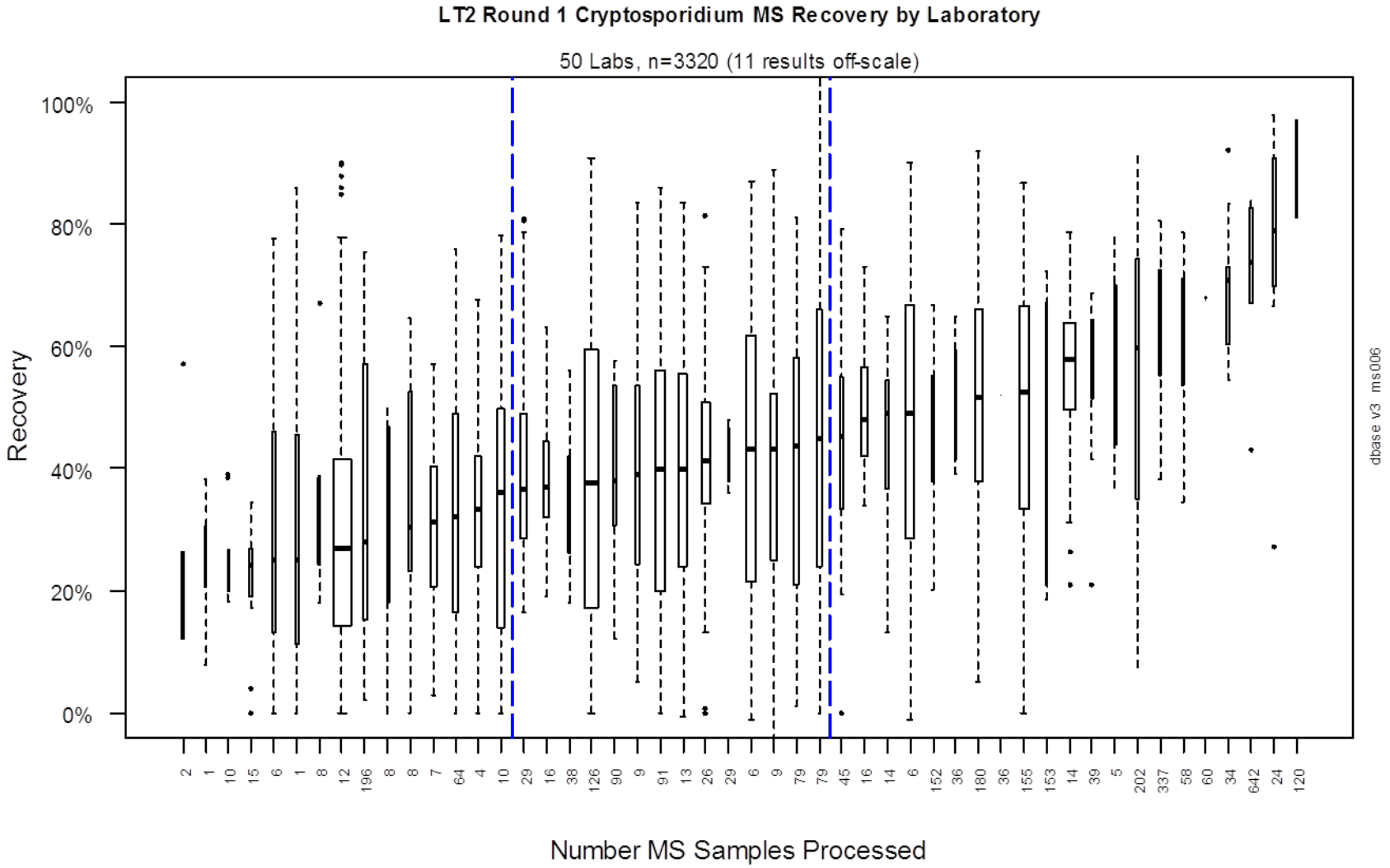
Filtered Locations



Distributions of *Cryptosporidium* Recovery by Laboratory

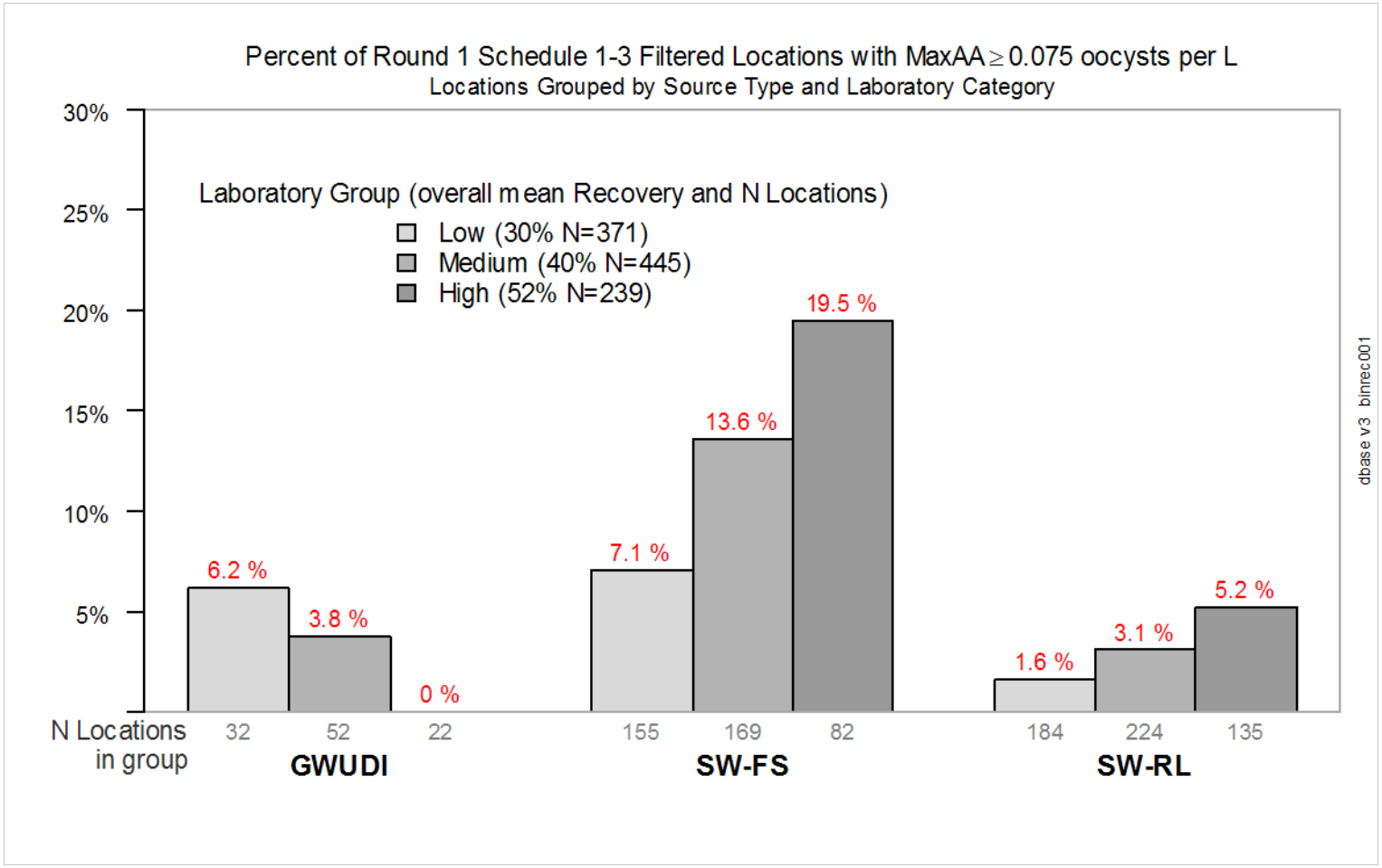


Laboratories Grouped into Low/ Medium/ High Recovery



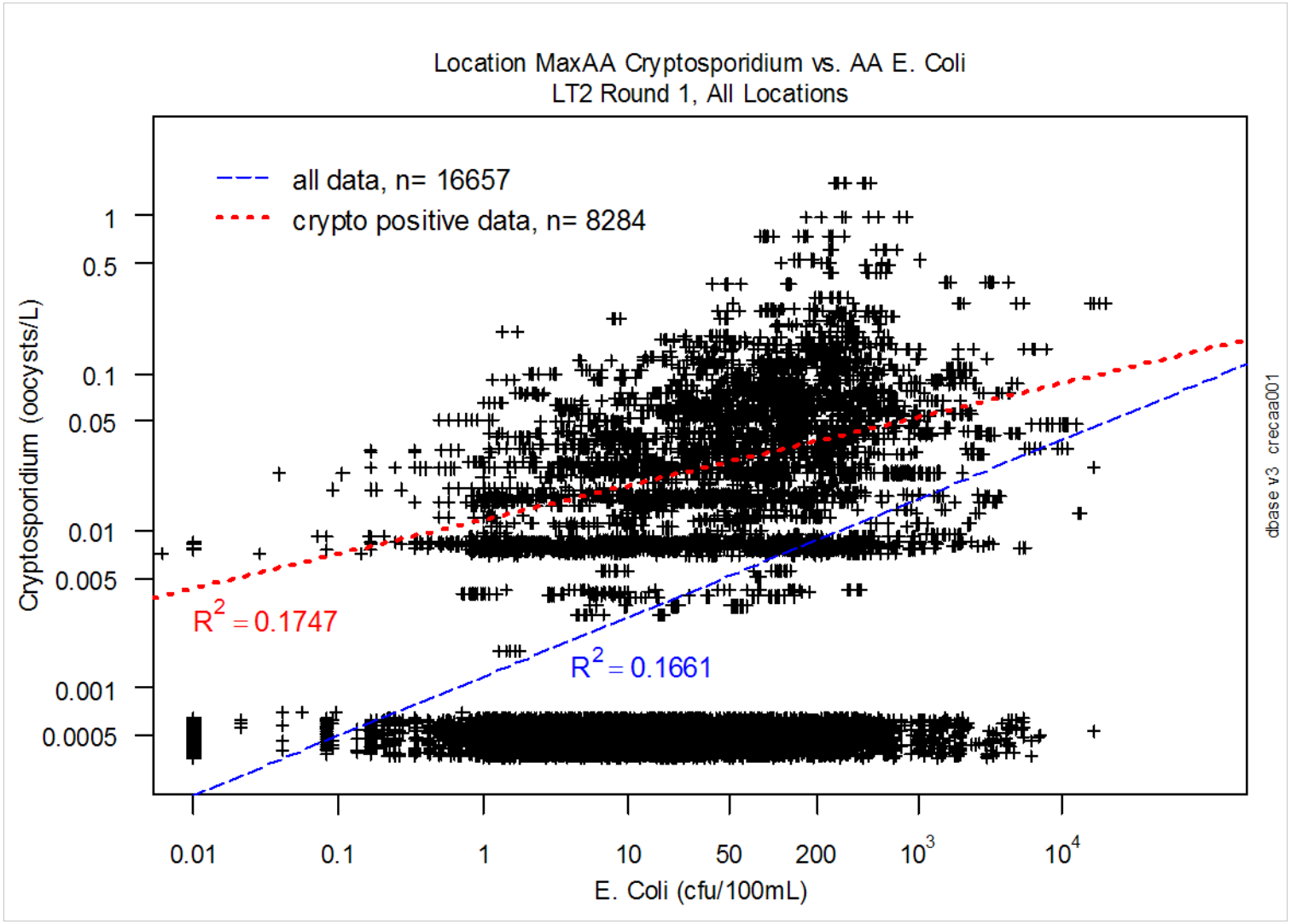
Higher *Cryptosporidium* Results at High Recovery Labs

locations grouped by L/M/H laboratory

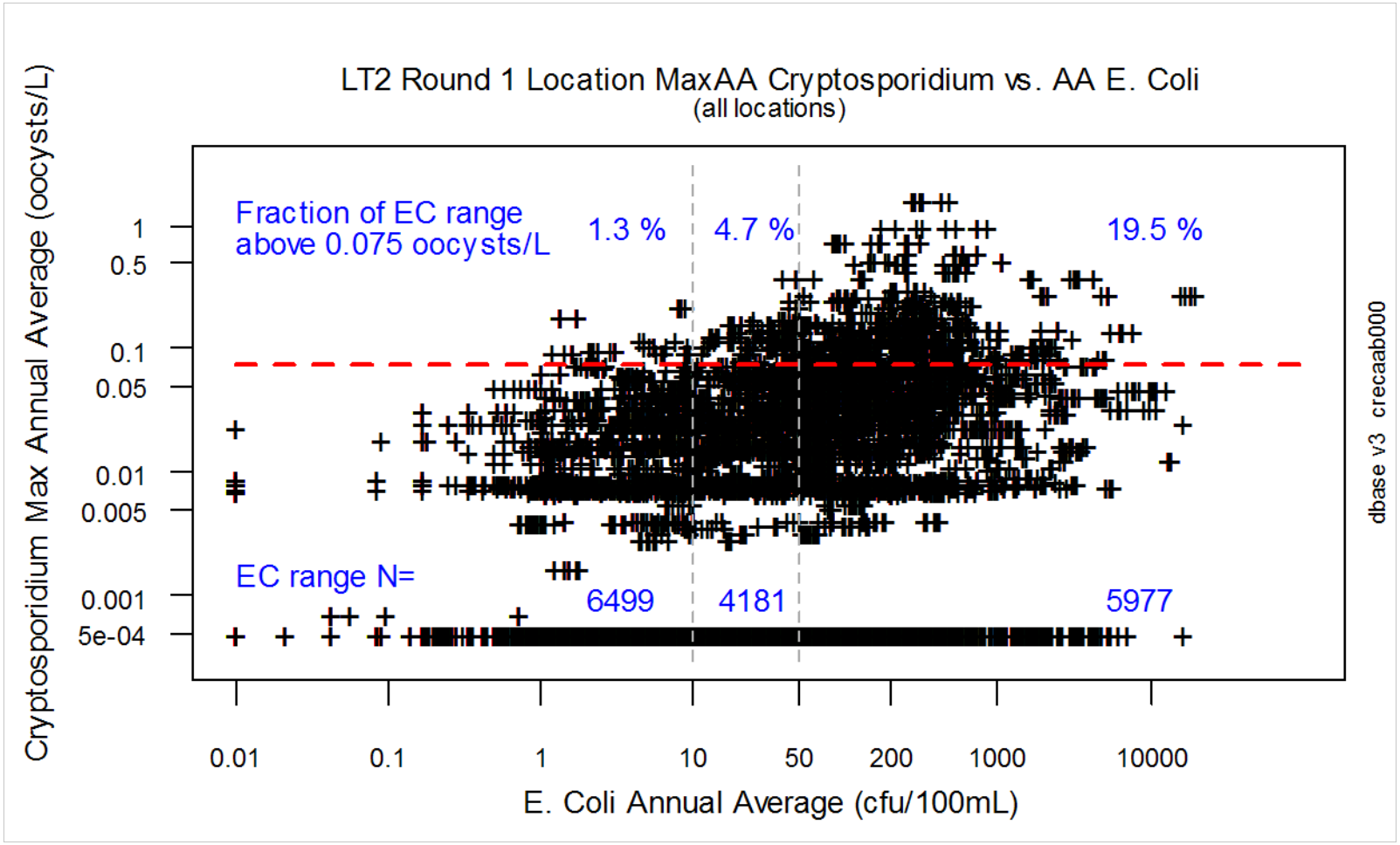


Location MaxAA *Cryptosporidium* vs. Location AA E. coli

1633 locations



Location MaxAA *Cryptosporidium* vs. Location AA E. coli 1633 locations



Collection and Use of Compliance Data

- Water utilities expend considerable effort and cost for data collection
- Do we intend to leverage compliance data for analyses to inform sound science for regulations?
- If so, data collection and handling processes should support that goal to best extent

Data Analysis Transparency

- Round 1 data analysis pictures may be confusing
  - datasets have overlapping coverage and information
  - data handling can make comparison on equal footing difficult
  - sub-population patterns differ widely, dataset compositions must be well-understood for appropriate comparisons
- Organize analyses consistent with rule structure
  - maxAA vs. mean
  - individual intakes vs. utilities
- Projections based on random draws from population won't inform how recovery affects individual intakes
  - more information needed to support understanding

Recommendations for Future Data Collection

- better data integrity rules
  - location descriptor consistency
  - classification variables
- capture data from all population categories
- capture grandfathered data
- supporting information for *Cryptosporidium* recovery analysis
  - UV254
  - alkalinity

Thank You

Questions for clarification?