

# **Methodologies in the Wine Interlaboratory Testing**

**Examination of  
methodologies and  
results in the ASEV-CTS  
Interlaboratory Program**



## An aside about consensus statistics

- Large Group Statistics are preferred
  - More robust
  - Small groups must have protocols to reduce influence of possible outliers
  - Generally, not advisable to subdivide except to mitigate an identifiable bias
- Interlaboratory testing (statistical model used) is concerned with measurement agreement, not determination of wine properties.

# Statistical stuff - Please stay awake.

- The Grand Mean is an estimate of the true value, e.g. amount of volatile acidity in a batch.
- Bias
  - A **biased** estimator is one that over or underestimates the quantity that is being estimated
- Residuals (Noise)
  - A **residual** is an observable *estimate* of the unobservable error
- Noticed how often I had to say *estimate*

# The Most Simple Statistical Consideration

- Interlaboratory Testing = Measurement Comparison
- Subdivide to eliminate meaningful bias
- Should these measurements be compared?

## Meanwhile, back in the real world...

- Even if we determine meaningful bias, is it a good idea to subdivide?
- Industry must ponder non-technical considerations.
  - Communication, regulation, confusion and expense

## And finally onto Interlaboratory Testing

- Troublesome properties in the ASEV-CTS Interlaboratory Program

Sugar (The Past) 

SO<sub>2</sub> (The Present)

Alcohol (The future?)

# Sugar – The Split

- History

The LPTGC, supported by ample statistical evidence, endorsed subdividing the sugar analysis in Spring 2005 (Cycle 019)

- Terminology

Old Analysis 908 Residual Sugar

New Analysis 908 Reducing Sugar

New Analysis 910 Glucose + Fructose

# Successful Sugar Split?

- Very pronounced effect in dry reds
- Significant in dry whites
- Negligible in sweeter blush

Please note that I am not proposing a general rule. These conclusions simply reflect the wines we have examined in the program.

- **Readily understood by large majority of participating laboratories.**

# Split Sugar Grand Means

Wines	Reducing Mean	Gluc + Fruc Mean	Difference	
Dry Red	6.96	5.11	1.85	36%
	6.55	4.73	1.82	38%
Dry Red	7.17	5.52	1.64	30%
	7.05	5.53	1.52	28%
Dry White	6.68	6.17	0.50	8%
	6.91	6.40	0.51	8%
Sweet Blush	33.57	33.23	0.34	1%
	21.14	20.81	0.32	2%

## Successful Sugar Split? #2

- So we have two sets of measurements with tighter variances, right?
- Problem of smaller groups
  - between 44 and 50 participants in Glucose + Fructose
  - between 18 and 20 participants in Cu Reduction
- Variability in Glucose + Fructose is slightly better for dry reds and the same for all others.
- Variability in Reducing Sugar is slightly worse in all cases.

# SO<sub>2</sub> – Ripper, AO and Other

- History

CTS and LPTGC have been examining the Ripper v AO question since the second testing round  
Interlaboratory data has been used in multiple research papers and publications

- Difficult to judge

Popularity of methods now tied to wine type.  
In the past few labs reported Total SO<sub>2</sub> by AO  
(<10 between reports 3 and 10).  
Relatively large variation in Free SO<sub>2</sub>

# Total SO<sub>2</sub> Ripper v AO

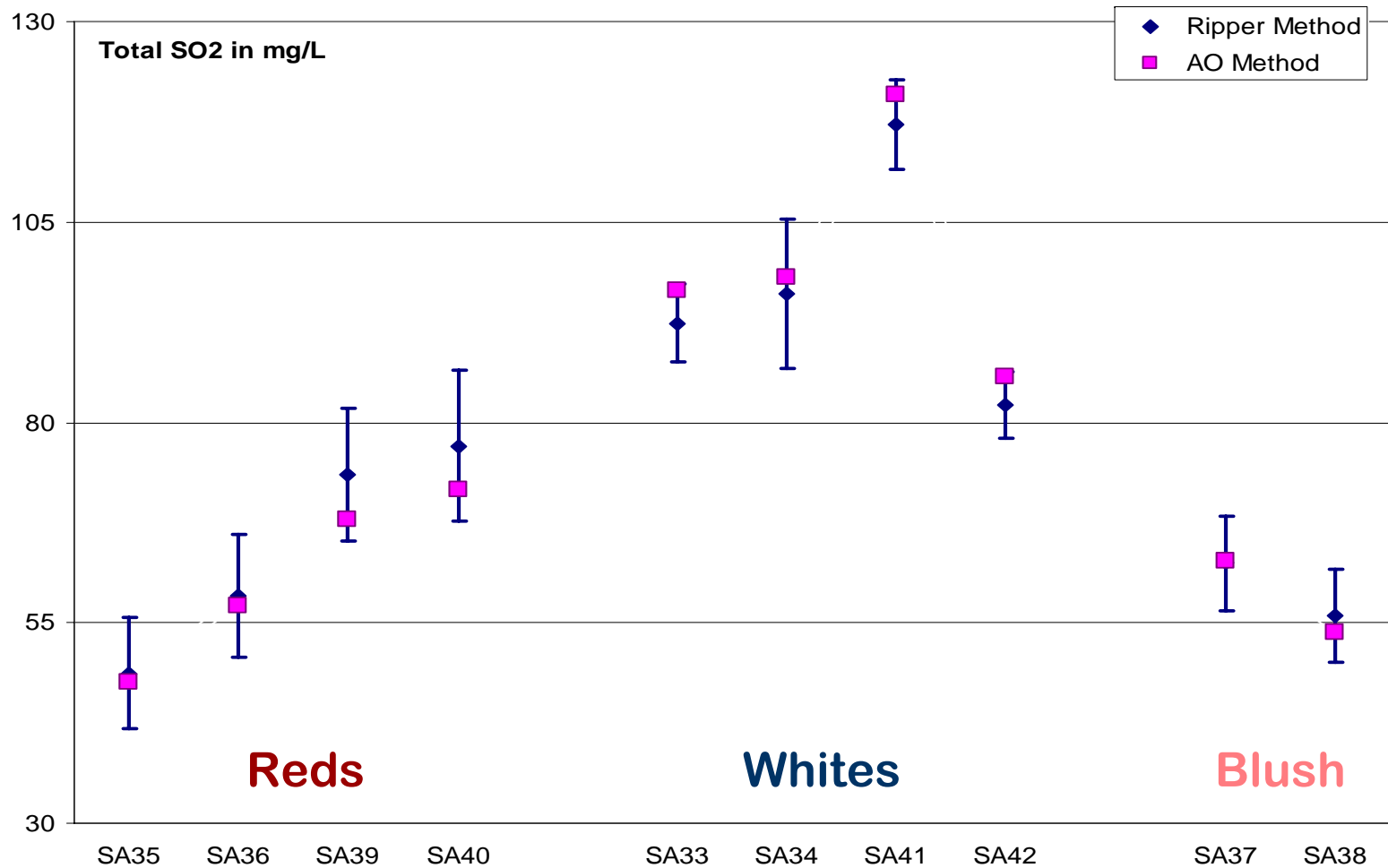
<i>Total SO2 (Analysis 902)</i>			Ripper Method				AO Method			
Wine Type	Difference Ripper v AO		Mean	SSTD	%CV	Count	Mean	SSTD	%CV	Count
Dry Red	1.2	2.5%	48.7	6.9	14.2%	32	47.5	5.1	10.7%	19
Dry Red	1.3	2.3%	58.4	7.6	13.0%	32	57.1	6.3	11.0%	19
Dry Red	5.6	7.9%	73.5	8.3	11.3%	30	67.9	3.4	5.0%	16
Dry Red	5.5	7.3%	77.1	9.4	12.2%	30	71.6	4.3	6.0%	16
Dry White	-4.0	-4.2%	92.4	4.9	5.3%	30	96.4	8.9	9.2%	14
Dry White	-2.0	-2.1%	96.1	9.3	9.7%	30	98.1	9.3	9.5%	14
Dry White	-3.7	-3.1%	117.2	5.6	4.8%	29	120.9	5.9	4.9%	13
Dry White	-3.4	-4.0%	82.2	4.2	5.1%	29	85.6	7.0	8.2%	13
Sweet Blush	-0.2	-0.3%	62.5	5.9	9.4%	33	62.7	5.7	9.1%	15
Sweet Blush	1.9	3.4%	55.8	5.8	10.4%	33	53.9	7.2	13.4%	15

# Free SO<sub>2</sub> Ripper v AO

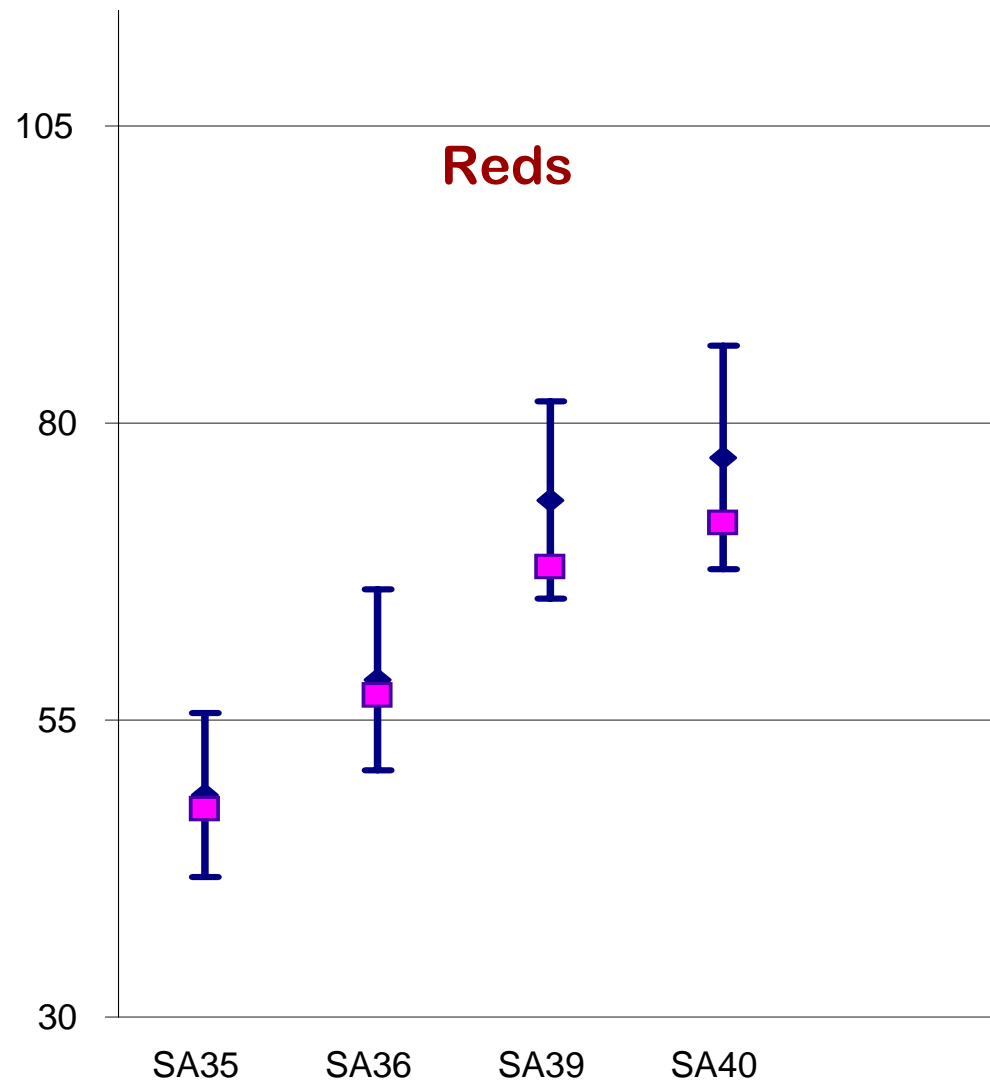
<i>Free SO<sub>2</sub> (Analysis 903)</i>			Ripper Method				AO Method			
Wine Type	Difference Ripper v AO		Mean	SSTD	%CV	Count	Mean	SSTD	%CV	Count
Dry Red	3.6	16.8%	24	2.5	10.4%	15	20.4	3.6	17.6%	36
Dry Red	3.0	14.9%	22.2	2.4	10.8%	15	19.2	2.7	14.1%	36
Dry Red	4.0	15.1%	29.4	2	6.8%	16	25.4	4.3	16.9%	31
Dry Red	4.6	16.7%	30.9	2.1	6.8%	16	26.3	3.5	13.3%	31
Dry White	2.2	12.3%	18.9	2.8	14.8%	20	16.7	2.7	16.2%	24
Dry White	2.3	11.1%	22.0	3.3	15.0%	20	19.7	3.4	17.3%	24
Dry White	3.3	12.8%	27.5	3.4	12.4%	20	24.2	3.1	12.8%	26
Dry White	3.1	12.1%	27.4	2.7	9.9%	20	24.3	3.4	14.0%	26
Sweet Blush	2.4	24.5%	11.5	1.7	14.8%	18	9.1	1.9	20.9%	32
Sweet Blush	3.6	40.9%	11.4	2.1	18.4%	18	7.8	1.8	23.1%	32

# Total SO<sub>2</sub> Ripper v AO

Total SO<sub>2</sub> Ripper / AO Comparison

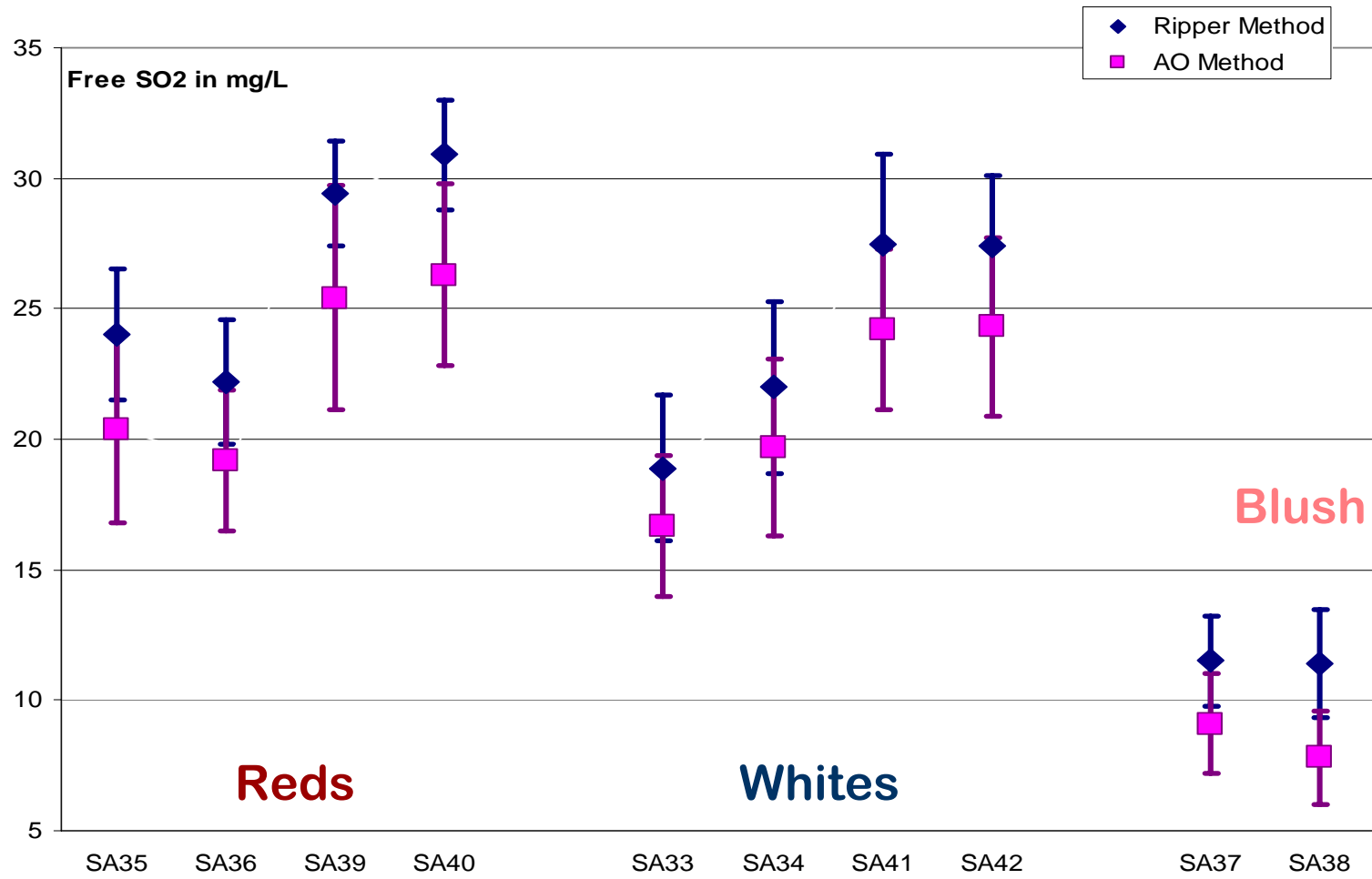


# Total SO<sub>2</sub> Ripper v AO

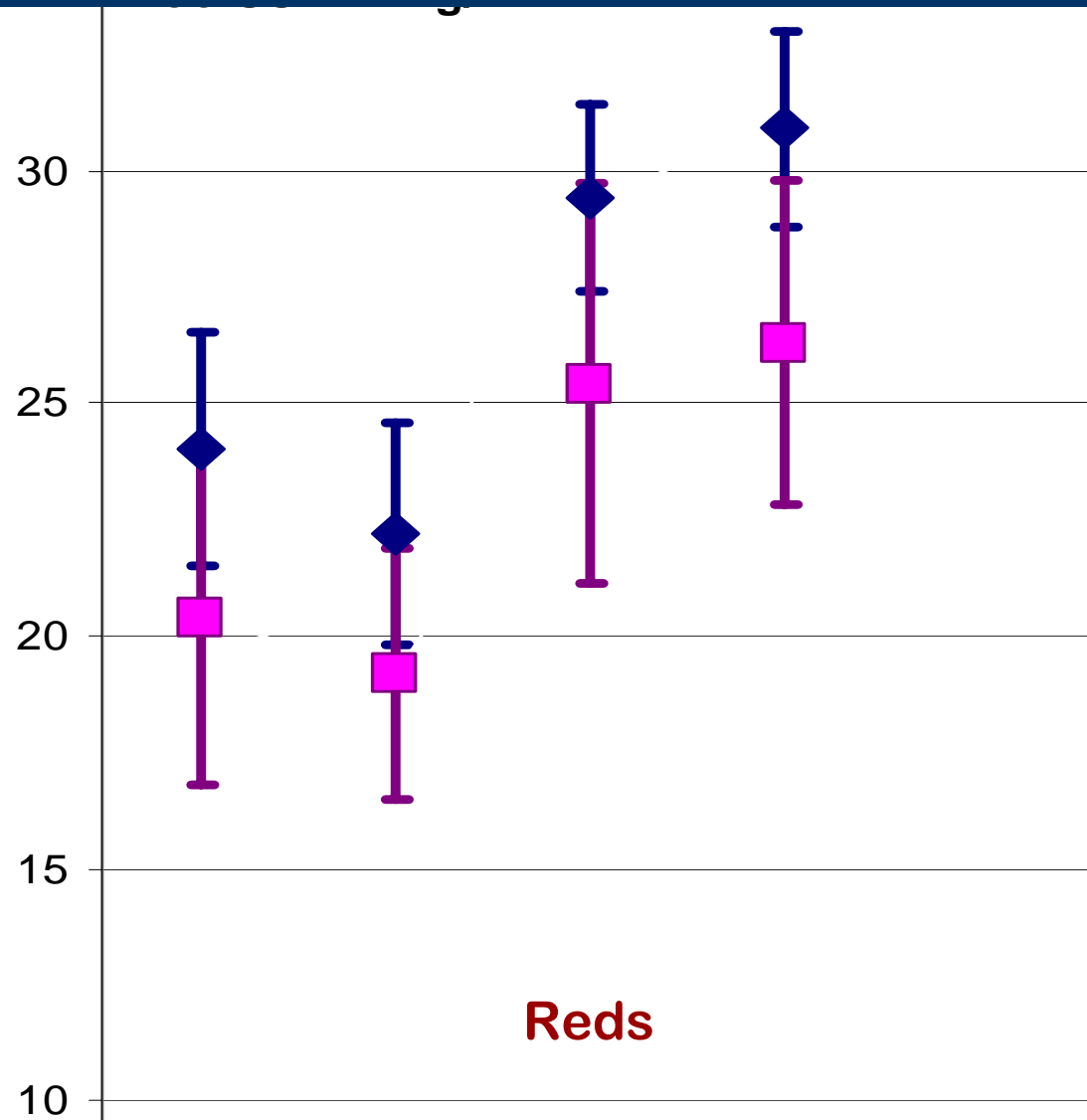


# Free SO<sub>2</sub> Ripper v AO

Free SO<sub>2</sub> Ripper / AO Comparison



# Free SO<sub>2</sub> Ripper v AO



# SO<sub>2</sub> – Ripper v AO

- General conclusions – Methodology Bias

Total SO<sub>2</sub>: Continually just *inside* statistically significant

Free SO<sub>2</sub>: Significant for reds and blush wines and *may* be significant for white wines.

# SO<sub>2</sub> – Ripper v AO

- General conclusions – Methodology Usage
  - Total SO<sub>2</sub>: Ripper used more (2:1)
  - Free SO<sub>2</sub>: AO is more common and (2:1) for red wines
- “Other Methods” slowly increasing in usage. Too little data for conclusions

# SO<sub>2</sub> – Roadmap

- Ample statistical support for split in Free SO<sub>2</sub>
- Split in Total SO<sub>2</sub> is question
- How to handle increasing participation in other methods

# %Ethanol – Problem?

- Skinny ellipses

% ethanol is one of the properties that routinely exhibits highly correlated behavior and results in long thin ellipse.

Between lab variation is proportionally larger than (standard deviations between 0.08 and 0.13) between sample variations (correlation coefficient of best fit line on two sample plot is generally approaches 0.8).

- So, is there a problem with biases between the multiple methodologies ?

# % Ethanol – Method Averages

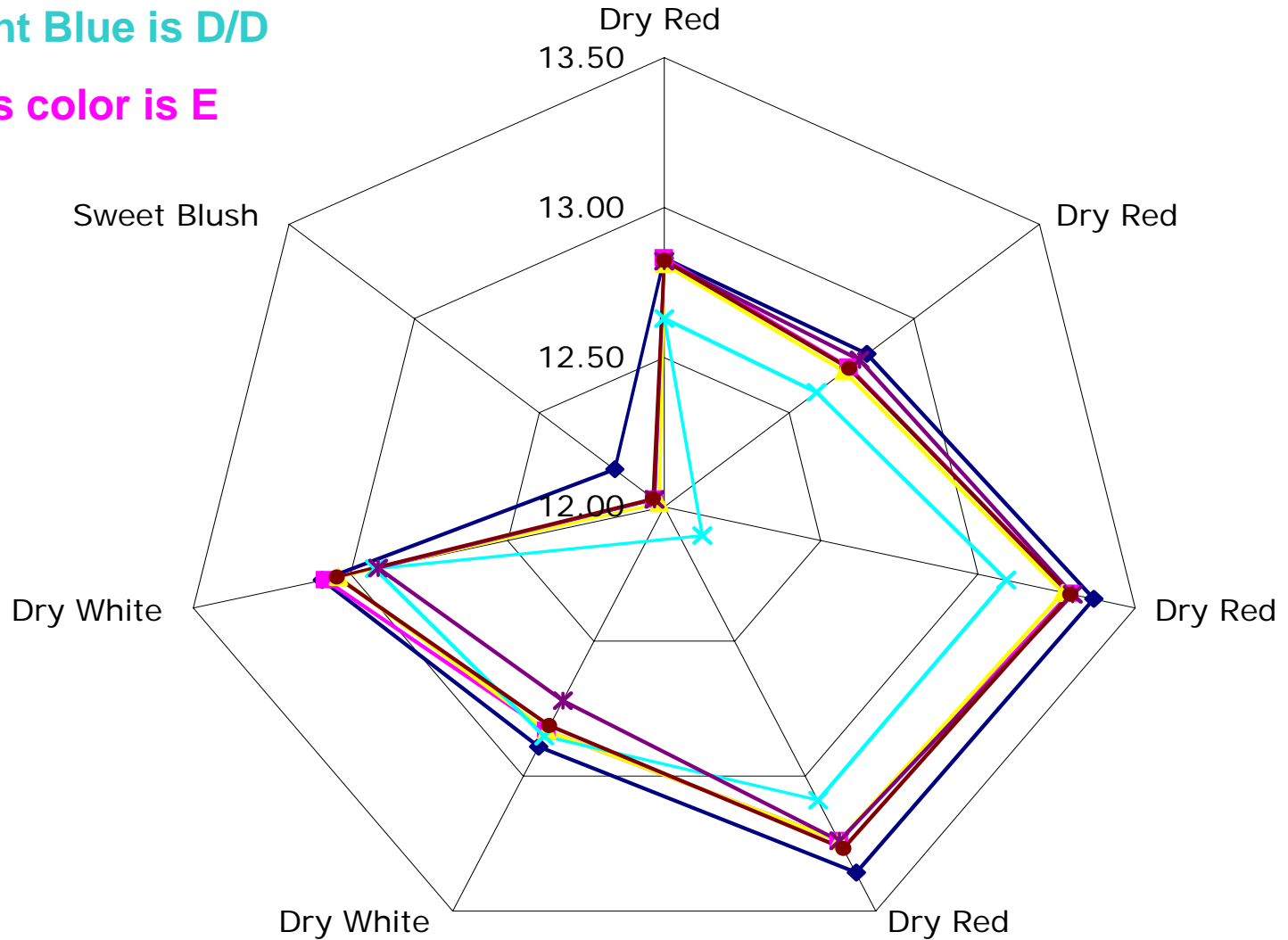
Cycle	Wine	Ebulliometer	GC	NIR	D/D	FTIR
022	Dry White	12.89	12.83	12.83	12.85	12.72
022	Dry White	13.09	13.08	13.04	12.92	12.91
021	<b>Dry Red</b>	13.37	13.29	13.27	13.09	13.30
021	<b>Dry Red</b>	13.36	13.24	13.24	13.09	13.24
020	<b>Sweet Blush</b>	10.51	10.37	10.25	10.16	10.71
020	<b>Sweet Blush</b>	12.20	12.04	12.02	11.85	12.04

\* This data reflects no exclusions.

# % Ethanol – Method Averages

Light Blue is D/D

This color is E



## %Ethanol – Problem?

- Means of different methodology groups **generally agree**. Differences between methodologies largely less than 0.1%, even without exclusions.

The lab-to-lab variation within each methodology is comparable to the lab-to-lab variation of the entire analysis.

## %Ethanol – Problem?

- No problem apparent in comparing methods
- Bias predominately between *labs* not *methods*
- But...

# % Ethanol – Ebulliometer Problem

Cycle	Wines	Ebulliometer Reporting	Ebulliometer Excluded	NIR Reporting	NIR Excluded
022	Dry White	10	5	25	2
021	Dry Red	15	5	21	1
020	Sweet Blush	14	6	21	0
019	Dry Red	16	5	20	0
018	Dry White	17	4	19	1



Most, but certainly not all, are excluded for consistency (14 of 25).

7 labs have both results too high; 4 have both too low.

# **Methodologies in the Wine Interlaboratory Testing**

**Examination of  
methodologies and  
results in the ASEV-CTS  
Interlaboratory Program**

---

**The End**