

**C13 Surveillance Panel**  
OF  
ASTM D02.B0.02  
June 18, 2007  
Loews Miami Beach Hotel, Miami Beach, FL

THIS DOCUMENT IS NOT AN ASTM STANDARD: IT IS UNDER CONSIDERATION WITHIN AN ASTM TECHNICAL COMMITTEE BUT HAS NOT RECEIVED ALL APPROVALS REQUIRED TO BECOME AN ASTM STANDARD. IT SHALL NOT BE REPRODUCED OR CIRCULATED OR QUOTED, IN WHOLE OR IN PART, OUTSIDE OF ASTM COMMITTEE ACTIVITIES EXCEPT WITH THE APPROVAL OF THE CHAIRMAN OF THE COMMITTEE HAVING JURISDICTION AND THE PRESIDENT OF THE SOCIETY. *COPYRIGHT ASTM, 100 BARR HARBOR DRIVE, WEST CONSHOHOCKEN, PA 19428-2959.*

---

**ACTION ITEMS**

1. **CAT discuss completing the Research Report**
  2. **Surveillance Panel conduct conference call to discuss a reference oil target update.**
- 

**MINUTES**

- 1.0 Call to order
  - 1.1 The C13 Surveillance Panel was called to order by Chairman Jim Gutzwiller at 10:05 a.m. on Monday, June 18, 2007, in the Poinciana 2 Room of the Loews Miami Beach Hotel, Miami Beach, FL.
  - 1.2 There were 10 members present and 16 guests present. The attendance list is shown as Attachment 2.
- 2.0 Agenda
  - 2.1 The agenda is included as Attachment 1.
- 3.0 Minutes
  - 3.1 The minutes from the January 24, 2007 meeting and the February 9, 2007 and the May 11, 2007 conference calls were approved as written.
- 4.0 Membership
  - 4.1 There were no membership changes. Phil Spengler represented CAT.
- 5.0 Scope and Objectives
  - 5.1 Phil Spengler will take a message to Hind to discuss completing the ASTM research report.
- 6.0 Ring Rating Workshop Report
  - 6.1 Elisa Santos completed an analysis of the 2007 rating workshop but could not attend the meeting. Jim Gutzwiller presented her analysis. See Attachment 3. Workshops from 2006 and 2007 were analyzed. The deposits on the rings change slightly over long periods of time, so rings from the first workshop were not used for the second workshop.
  - 6.2 Conclusions suggest that variability increases with demerits magnitude. It is unlikely that variability has not changed from one year to the other. Jim Rutherford asked about the variability in the same performance region. Further analysis suggests that variability is reduced in 2007 compared to 2006 for deposits in the same region. Rater precision seemed to improve during the 2007 workshop. Two ring sets were rated by everyone the

first day and results shown to all. Raters discussed among themselves the various rating levels, then the same rings were coded differently and rated again. The precision improved as a result of discussion and agreement between raters. The improvement may not be statistically significant. The data suggest that there is a rater effect. The likely variation between raters is +/- 2.4 demerits. Some questions have not been addressed yet.

- 6.3 Jim McCord asked whether rubbed carbon and polished carbon are rated the same as heavy carbon or not and if that accounts for the increased variability at higher deposit levels. The CRC manual defines heavy carbon as carbon making contact whether it is rubbed or polished.
- 7.0 Parts Supplier Report
  - 7.1 No parts issues at this time. There is a one and a half to two year supply of C13 liners at this time.
- 8.0 TMC Report
  - 8.1 Jeff Clark presented the TMC report. See Attachment 4. C13 reference test activity is low; 3 calibrated labs with 5 calibrated stands. Reference oil 831 (PC-10B) has been exhausted. The TMC is waiting on a re-blend. 15 tests exist in the database on this oil, but some of those are from a lab that never calibrated and some other runs are "large impact tests". Jeff suggests a more thorough Surveillance Panel review before updating targets. Values were shown for all 15 tests, 14 tests from calibrated labs, and 13 tests removing one test that was very severe for R2TC. The delivery of the re-blend is unknown at this time. A re-blend is supposed to be delivered.
  - 8.2 Chairman Gutzwiller suggests that further analysis be performed and a conference call held to decide on a reference oil target update.
- 9.0 Oil Consumption Delta Calculation
  - 9.1 A question was raised on a previous conference call about what to do with a negative oil consumption value. The result is a square root transformation for the application of any potential industry correction factors. Several suggested techniques were proposed. Mark Sutherland **moved** that for negative OC deltas, the transformed values be blanked out and the same negative value be shown as the final value. Jim McCord seconded. The **motion carried** by voice vote.
- 10.0 Inlet Air Pressure
  - 10.1 Some labs on low barometer days run below the 93 kPa setpoint. When that was set-up, it was agreed to review that setpoint. The intent is to run at 93, but the panel recognizes that on low barometer days, the inlet air pressure may drop below 93. On high barometer days, the labs should trouble shoot low inlet air pressure conditions.
- 11.0 Other Business
  - 11.1 Labs have had trouble getting fuel. One lab that uses rail cars has had much trouble getting fuel; having to wait months. The perception is that fuel delivery is getting worse. The situation exists for both PC-9 and PC-10 fuel. The HDEOCP will be notified. Labs have had to delay starts to know that fuel was on the way.
- 12.0 Next meetings
  - 12.1 A conference call will be set up to further discuss ring rating workshop analyses and reference oil target updates.
- 13.0 The meeting was adjourned at 11:20 am.

## **C13 Surveillance Panel Meeting**

### **Meeting Agenda June 18, 2007 10:00 am – noon Loews Miami Beach Hotel Poinciana 2**

Membership	Stacy Bond
Approve Minutes January 24, 2007 Meeting February 9, 2007 Conf. Call May 11, 2007 Conf. Call	Stacy Bond
Review Scope and Objectives	Jim Gutzwiller
Ring Rating Workshop Report	Jim Gutzwiller
Parts Supplier Report	
TMC Report	Jeff Clark
Oil Consumption Delta Calculation	Jeff Clark
Inlet Air Pressure Range or Spec.	Group
New Business / A.O.B.	Group
Next Meeting	Jim Gutzwiller

6/18/07

# CIS SURVEILLANCE PANEL LOEWS MIAMI BEACH

JIM MORITZ	INTERTEK
Jim GUTZWILLER	INFINEUM
Jim Rutherford	ORONITE
JEFF CLARK	TMC
Jim McCord	SWRI
MARK SUTHERLAND	ORONITE
Bob Campbell	AFTON
Frank Fernandez	ORONITE
GARY PARSONS	ORONITE
Doc KERSEY	Valvoline
Tom Buch	TEI
DAVID GLAENZER	AFTON
Steven Kennedy for Riccardo Conti	ExxonMobil
PAT FETTERMAN	INFINEUM
Phil Spangler	Caterpillar Inc.
PAUL L. STRIGNER (CONSULTANT-FACILITATOR) <a href="mailto:stingbeans@ca.inta.net">stingbeans@ca.inta.net</a>	
TELEPHONE: 613-746-0647	



# **Analysis of the C13 2ND RING TOP CARBON Workshop data: 2006 & 2007**

Elisa Santos

June 15<sup>th</sup>, 2007

## Questions:

1. “Have the raters improved their precision rating the second ring deposits: Jan2006 vs. the latest workshop Feb2007?”
2. “Did the precision of the raters improve during the 2007 workshop?”
3. Is there a rater effect?

## Overall structure of the work:

- Data Source
- Plots of the data
- Analysis of the data
- Conclusions

# Data Source

- Workshop 2006 & workshop 2007
- 192 test results
  - 2006: 80
  - 2007: 112
- Ring Sets
  - 2006: Prelim, G, A, B, C, D, "E, Gfinal
  - 2007: A through L (A & B have repeats for Day 1 & Day 2)
  - Ring sets common to both workshops: **NONE**
- Raters:
  - There are 11 raters in both workshops
  - 7 are common to both workshops

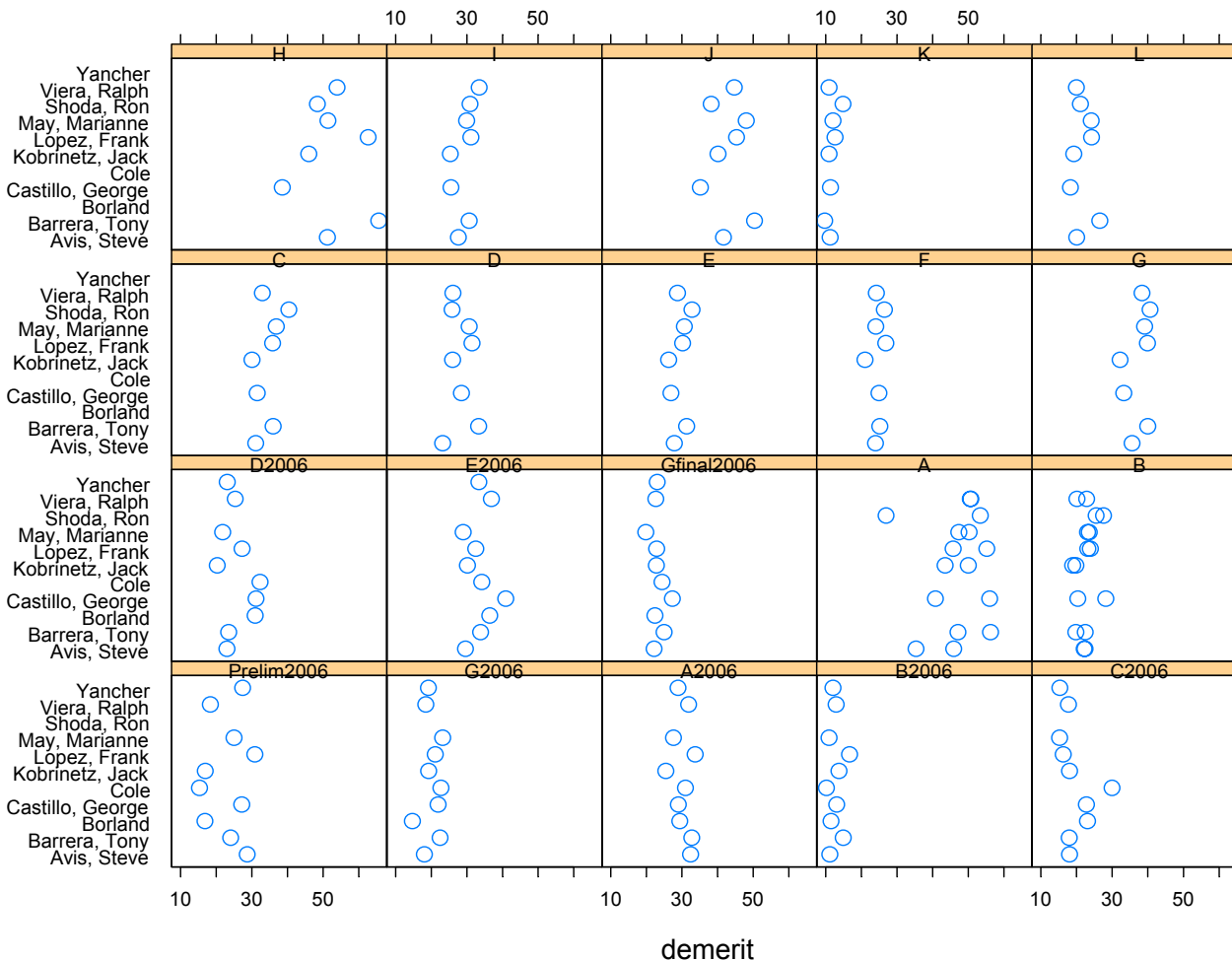
# Question #1:

- “Have the raters improved their precision rating the second ring deposits: Jan2006 vs. the latest workshop Feb2007?”



# Plots (1)

## Demerits by Rater and Ring Set for 2006 and 2007



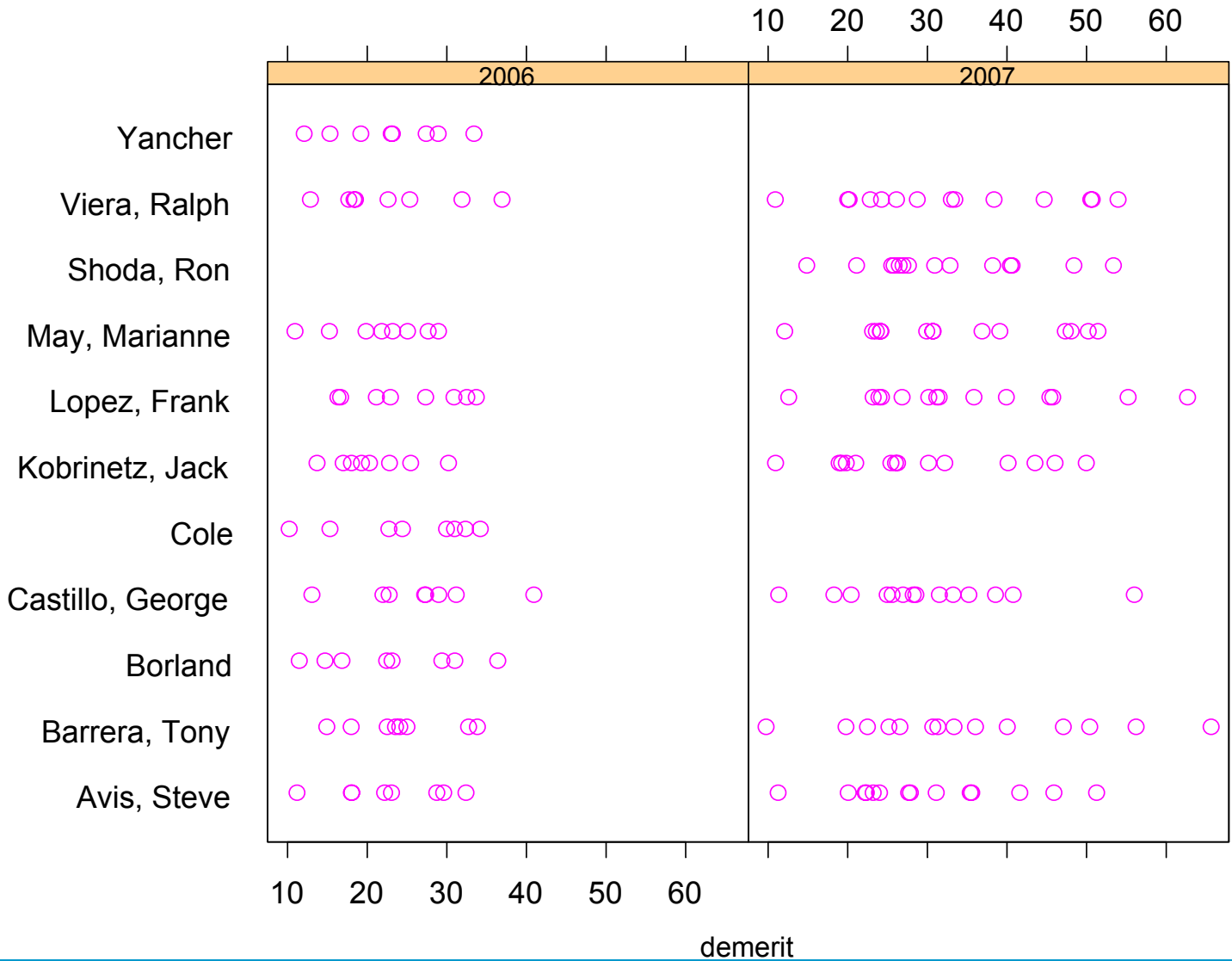
- Each panel corresponds to a Ring Set. The straighter the column of points, the better the agreement among raters

- There are 8 sets for 2006 and 12 sets for 2007. Seven raters are common to both workshops. **See plot in the next slide.**

- Note that only Ring sets A and B (for 2007) have repeats corresponding to Day 1 and Day 2 of the workshop.

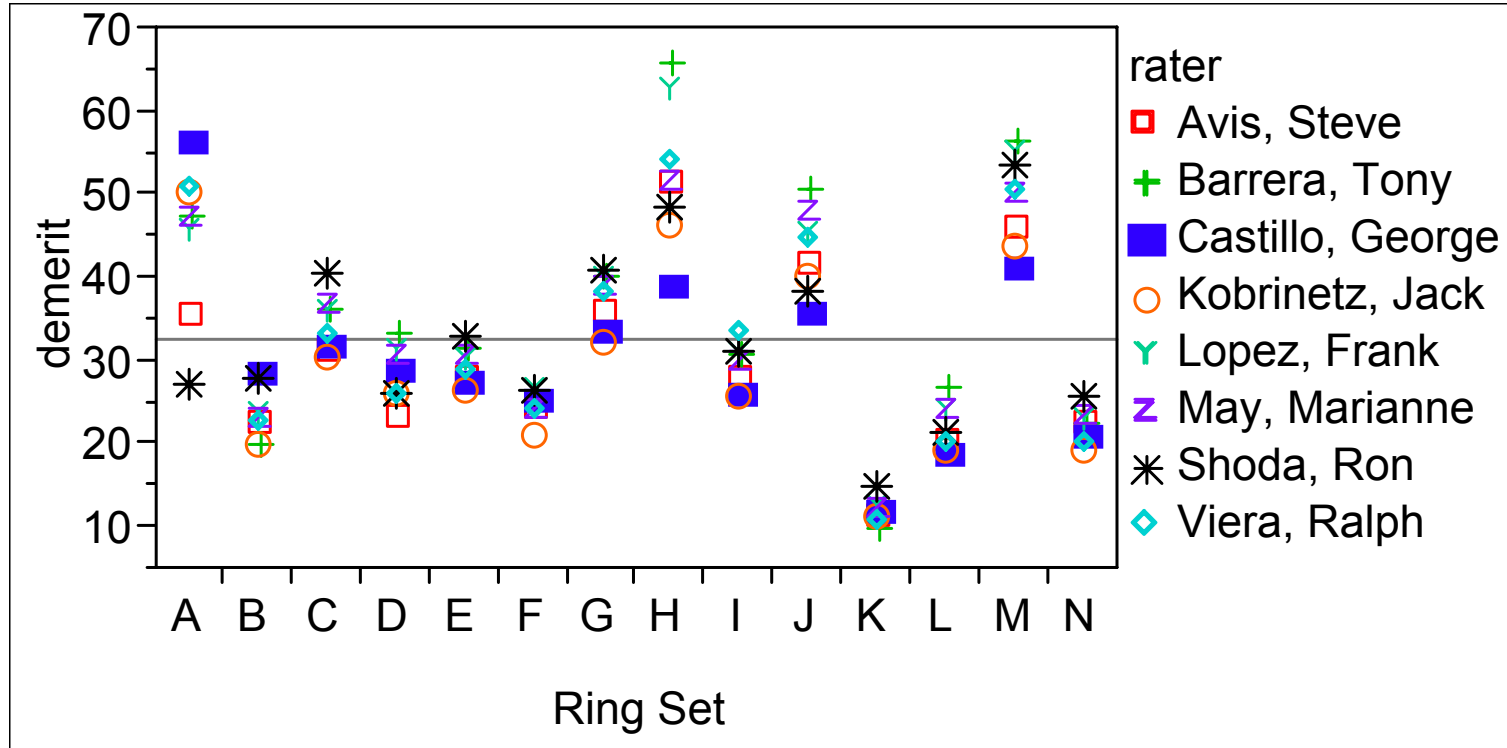
# Plots (2)

## Demerits by Rater and Year (ignoring Ring Set)



Note that the range of demerits is larger for 2007 than for 2006.

## Demerits by Rater and Ring Set for 2007



Ring set M is Ring set A on Day 2  
Ring set N is Ring set B on Day 2

- Models that describe demerits were selected by Elisa Santos.
- The statistical analysis based on the models was shared in advance with Jim Rutherford, Phil Scinto and Todd Dvorak.
- Jim had questions that were tentatively answered in Appendices 3 and 4.
- Additional questions and comments are welcome!
  - Elisa can be reached at [Elisa.Santos@Infineum.com](mailto:Elisa.Santos@Infineum.com)
- The details about the statistical analysis are omitted here.

# Conclusions for question #1

- Variability increases with demerits magnitude.
  - The model selection suggests that the demerits can be described by a model that allows for the variability to change as a power of the fitted value within both years.
  
- It is likely that the underlying variability of the measurements has not changed from one year to the other
  
- It is likely that the variability changes with the fitted value for both years, but that 2006 is observed in a narrower range than 2007.

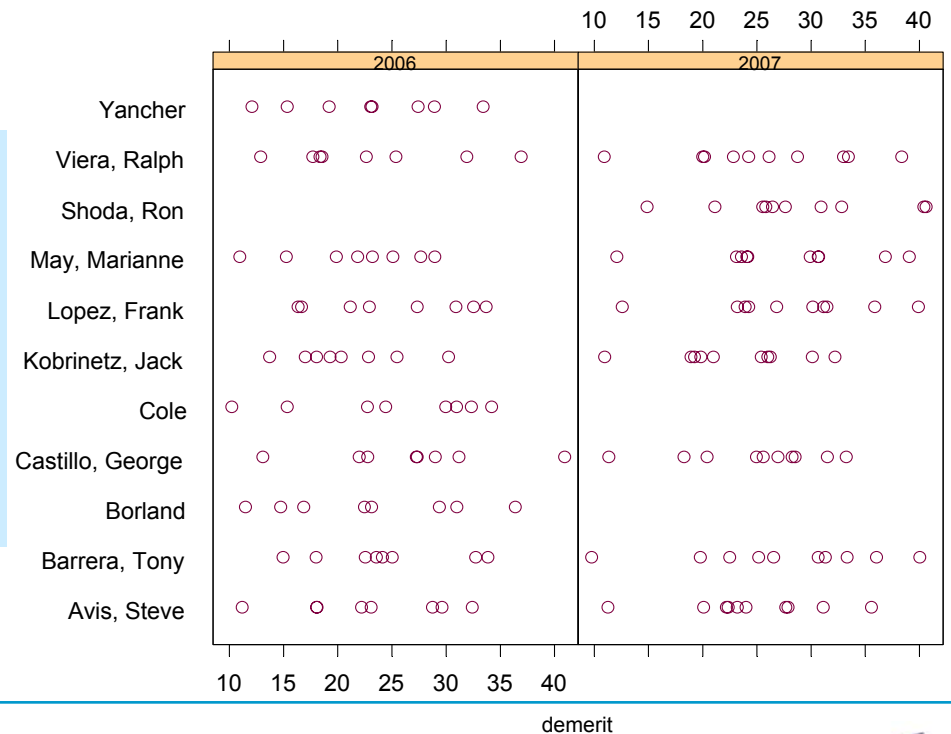
# Another way of answering Question #1

## Jim Rutherford's question (a)

- Jim said: "I agree with your conclusion that residual error looks like function of deposit. How about comparing years using only ring sets that are in same performance region?"
- Elisa said: "Jim, I created a subset of Ring set, eliminating Ring sets J, A and H from the 2007 workshop. The plot of the data is below. Then, I analyzed the subset. "

### Main conclusion:

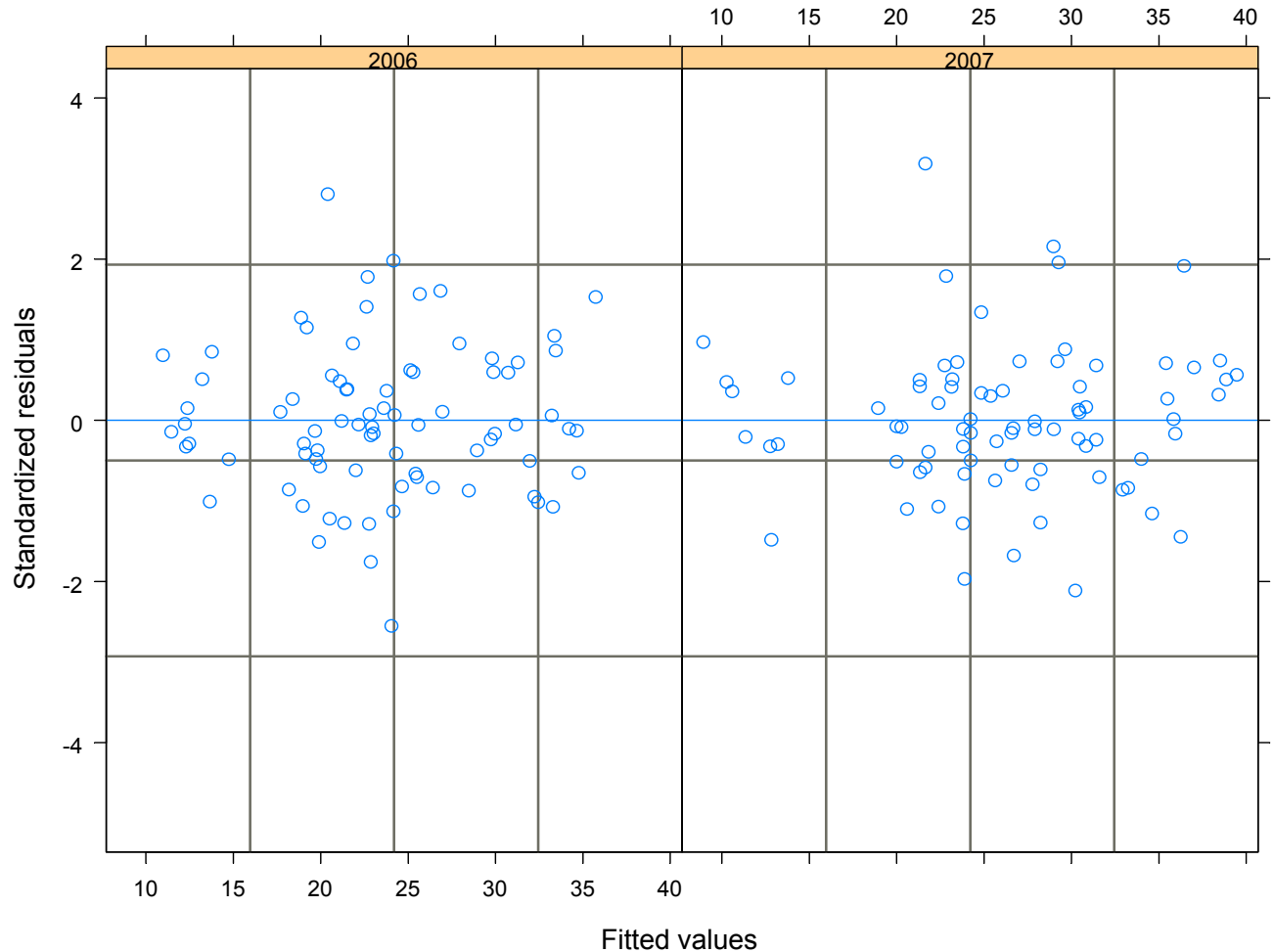
The statistical analysis of the subset indicates that the variability associated to 2007, **in the selected range of demerits**, has been reduced when compared to 2006.



# Standardized Residuals versus fitted values for model lme1c

Option 1: Allowing for the variability to be constant within each year, but different across years (model is called lme1a on slide 20)

Attachment 3; Page 11 of 16



The residual plot shows no pattern left in the data, indicating the adequacy of the model used.

# Question #2:

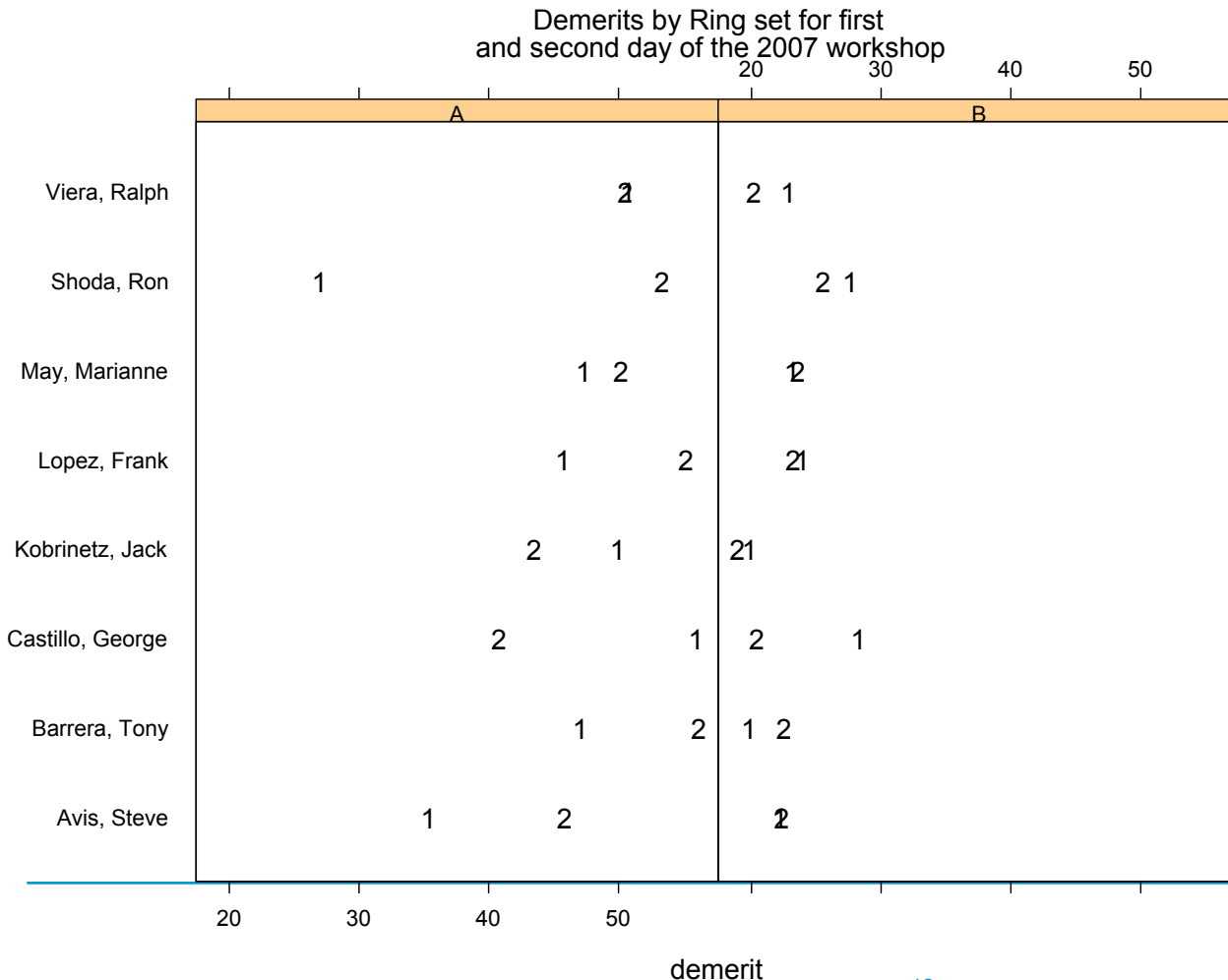
○ “Did the precision of the raters improve during the 2007 workshop?”

- Each panel corresponds to a Ring Set: A or B

- For each Ring set there are two measurements, one for Day 1 and one for Day 2

- Note that the variability associated to Ring Set B seems to be smaller than the variability associated to Ring A.

- Note also that the demerits for Ring A are larger than the demerits for Ring B.

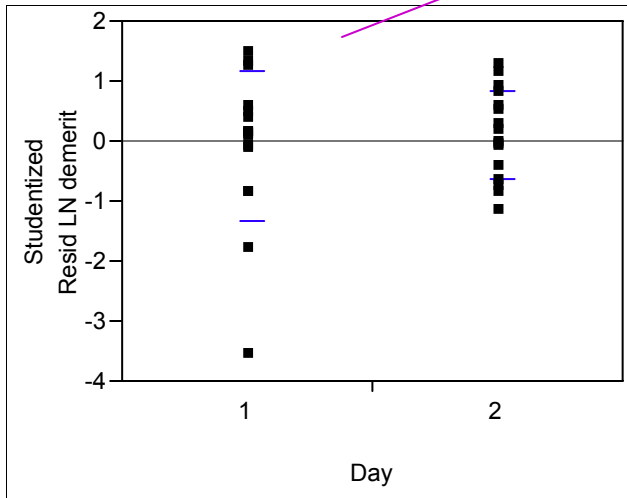




# Testing for differences between variability of Day 1 and Day 2 after removing the effect of Ring Set

Attachment 3; Page 13 of 16

Residual plot: Visualizing what is left after removing the effect of Ring Set

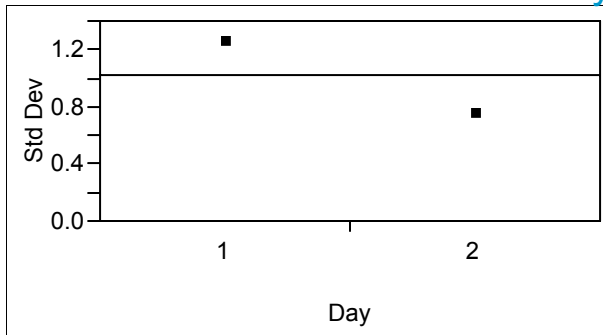


Level	Count	Std Dev	MeanAbsD if to Mean	MeanAbsDif to Median
1	16	1.252421	0.8471722	0.8050389
2	16	0.738785	0.6056735	0.6056735

Test	F Ratio	DFNum	DFDen	p-Value
O'Brien[.5]	1.4803	1	30	0.2332
Brown-Forsythe	0.5917	1	30	0.4478
<b>Levene</b>	<b>0.9745</b>	<b>1</b>	<b>30</b>	<b>0.3314</b>
Bartlett	3.8693	1	.	0.0492
F Test 2-sided	2.8738	15	15	0.0491

Standard Deviation versus Day



Using Levene's test (robust to non normality), there is no statistical evidence to reject the hypothesis that the Days have the same variability.

## Conclusion for question #2

- There is no statistical evidence that the variability associated to Day 1 is different from the variability associated to Day 2.

## Question #3:

### Jim Rutherford's question (b) - using all data

Is there a rater effect?

Yes, there is a rater effect. There is a statistically significant difference between raters.

The likely variation between individual raters assessment is plus or minus 2.4 demerits

Is it a practical difference?

○ Jim said: “Isn't part of the 2006 versus 2007 question whether raters agree? Seems like you have focused on within rater variability. Could you address among rater variability?”

○ Elisa said:

“Jim,

The final model selected to describe demerits (slide 11) has a random effect for raters. This means that the variability among raters is statistically significant: **there is a rater effect**. Below, I present the comparison of the models with a without the random effect for raters.

Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
lme1e	1	23 1028.839	1101.231	-491.4193			
glsb	2	22 1037.341	1106.585	-496.6703	1 vs 2	10.50195	0.0012

○ The 95% interval for the random effect is given by

Level: rater

	lower	est.	upper
sd((Intercept))	0.6745845	1.202507	2.143577

# Test Monitoring Center Report to the C13 Surveillance Panel

**June 18, 2007**  
**Miami, FL**



# Reference Oil Testing Summary

- C13 Test Activity
  - Low
    - Labs: 3 active & calibrated
    - Stands: 5 active & calibrated

# C13 Reference Oil Update

- Reference Oil 831 (PC-10B)
  - TMC supply has been exhausted
  - Waiting on reblend from supplier
  - Some samples still remain at labs

# C13 Reference Oil Update

- Test Targets for 831 (PC-10B)
  - Total of 15 tests
  - Issues with data
    - uncalibrated test lab
    - ‘large impact’ tests
      - R2TC, OC
    - oil supply
  - Suggest fuller SP review



# R2TC Results – Oil 831

<u>Dataset</u>	<u>N</u>	<u>Mean</u> (ln)	<u>Std. Dev.</u> (ln)
Targets	8	2.8954	0.2055
All Data	15	2.8785	0.2800
Cal. Labs	14	2.8828	0.2900
'Screened'	13	2.8337	<b>0.2336</b>

# OC Results – Oil 831

<u>Dataset</u>	<u>N</u>	<u>Mean</u> (sqrt)	<u>Std. Dev.</u> (sqrt)
Targets	8	5.7336	<b>0.7280</b>
All Data	15	5.6378	0.8502
Cal. Labs	14	5.5089	<b>0.7141</b>
'Screened'	13	5.5654	<b>0.7100</b>

# TGC Results – Oil 831

<u>Dataset</u>	<u>N</u>	<u>Mean</u>	<u>Std. Dev.</u>
Targets	8	45.18	<b>7.42</b>
All Data	15	46.60	6.11
Cal. Labs	14	46.02	5.90
'Screened'	13	46.32	6.03

# TLC Results – Oil 831

<u>Dataset</u>	<u>N</u>	<u>Mean</u>	<u>Std. Dev.</u>
Targets	8	<b>24.99</b>	7.59
All Data	15	22.04	7.63
Cal. Labs	14	21.87	7.89
'Screened'	13	22.78	7.41