VIF Precision Matrix Analysis

Statistics Group

Date: December 15, 2016

Statistics Group

- Arthur Andrews, ExxonMobil
- Doyle Boese, Infineum
- Jo Martinez, Chevron Oronite
- Kevin O'Malley, Lubrizol
- Martin Chadwick, Intertek
- Richard Grundza, TMC
- Lisa Dingwell, Afton
- Todd Dvorak, Afton
- Travis Kostan, SwRI

Summary

- Analyses include the results of 18 valid precision matrix tests which reflects the surveillance panel's decisions
- Simulations suggest a change in baseline weighting could improve test precision (estimated standard deviation decreases as much as 0.02% (8% reduction) for FEI1; 0.01% (11% reduction) for FEI2)
- Analyses indicate that engines may not differentiate oils similarly
- These data suggest that second run tests may be the highest. In particular, higher than first run tests. This could have implications on the engine hours corrections, engine calibration, and/or severity adjustments
- It is not clear, based on the data obtained, whether a nonlinear type of engine hours correction or lack of consistency in oil discrimination across the engines and engine life or combination of these effects exists

Input is needed from the surveillance panel for analysis to proceed – some options are provided

PM Data for Analysis

- Precision Matrix (PM):
  - On 11-7-16 the surveillance panel passed a motion to include 18 tests in the statistical analysis.

Run Order	EOT Engine Hours	SwRI #1		SwRI #2		IAR #1		IAR #2	
1	350	Stage 1 Sense Check	543 112952-VIF	Stage 2 Sense Check	1011 112953-VIF	Stage 1 Sense Check	542-2 112957-VIF	Stage 2 Sense Check	1011 112955-VIF Baseline Shift
2	550		542-2 112951-VIF		542-2 116037-VIF		543 112958-VIF		543 113824-VIF
3	750		542-2 113818-VIF		1011 112954-VIF		543 113823-VIF		1011 112956-VIF
4	950		543 113819-VIF		543 113820-VIF		542-2 113822-VIF EBP Calibration Shift 542-2 113231-VIF		542-2 116030-VIF
5	1150	1011 117508-VIF		543 113821-VIF Worn Throttle Controller 543 [illegible]-VIF		1011 116832-VIF		542-2 116031-VIF Baseline Shift	
6	1350	543 117626-VIF		1011 117509-VIF		543 113825-VIF		1011 117495-VIF	
7	1550	542-2 116038-VIF		542-2 117511-VIF		1011 117496-VIF		543 117494-VIF	
8	1750	1011 117510-VIF				542-2 117493-VIF			

Excluded From Analysis

LZ	
Additional Testing	1011 118268-VIF
	543 118267-VIF
	542-2 119631-VIF
	1011 119628-VIF

Test Reported | Under Review | Invalid

- Table is from Frank Faber's 6-21-16 matrix update plus 4 additional tests

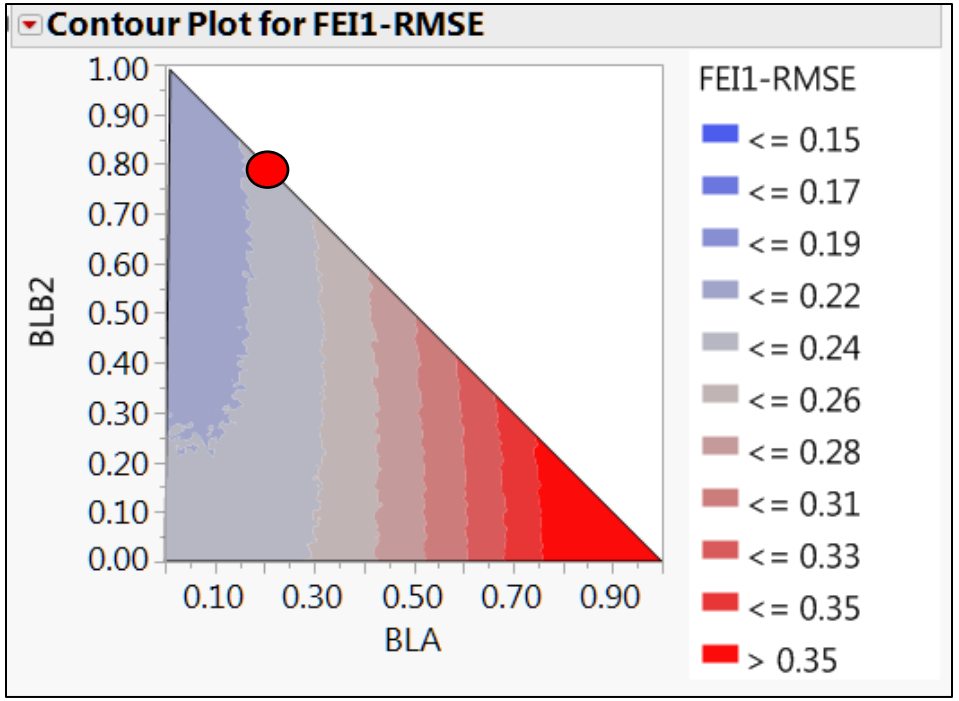
Evaluating Baseline Weight Scenarios

- Excel Program developed to evaluate 10,000 different weight combinations of BLB1, BLB2, and BLA
- Excel based prediction model for precision (RMSE) included Lab, Eng(Lab), Oil, and EngHr factors
- All BL weight combinations summed to a value of 1.0
- For those runs that included a BLB3, BL weights were applied to BLB2 & BLB3 in lieu of BLB1 & BLB2
- Results are shown on the following slides

Evaluating Baseline Weight Scenarios

- Plot of RMSE vs. baseline (BL) weight combinations for FEI1 shown below:
  - RMSE of weights can be interpreted from plot- if BL weights sum to 1.0
  - VID & VIE FEI1 Baseline weights are 80% & 20% (shown in red circle)
  - VIF test precision can be improved with weight factor of 1.0 for BLB2

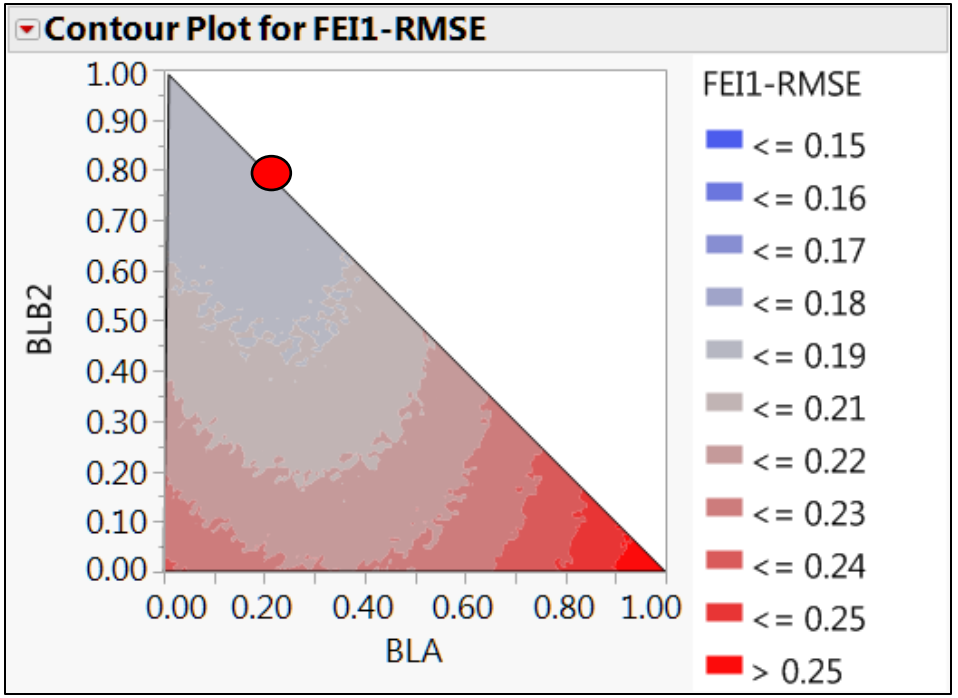
VIF Precision - BL Weights			
BLB1	**BLB2**	**BLA**	**FEI1-RMSE**
0	0.8	0.2	0.2225
0	1	0	0.2050



Evaluating Baseline Weight Scenarios

- Plot of RMSE vs. BL weight combinations for FEI1-with 1st run data deleted is shown below *(n = 14)*
  - VID & VIE FEI1 Baseline weights are 80% & 20% (shown in red circle)
  - Traditional BL weights appear to be better suited for this reduced data set
  - BL shifts tend to be higher during first run tests & may affect the BL weights and RMSE

VIF Precision - BL Weights			
BLB1	**BLB2**	**BLA**	**FEI1-RMSE**
0	0.8	0.2	0.1896
0	1	0	0.1912

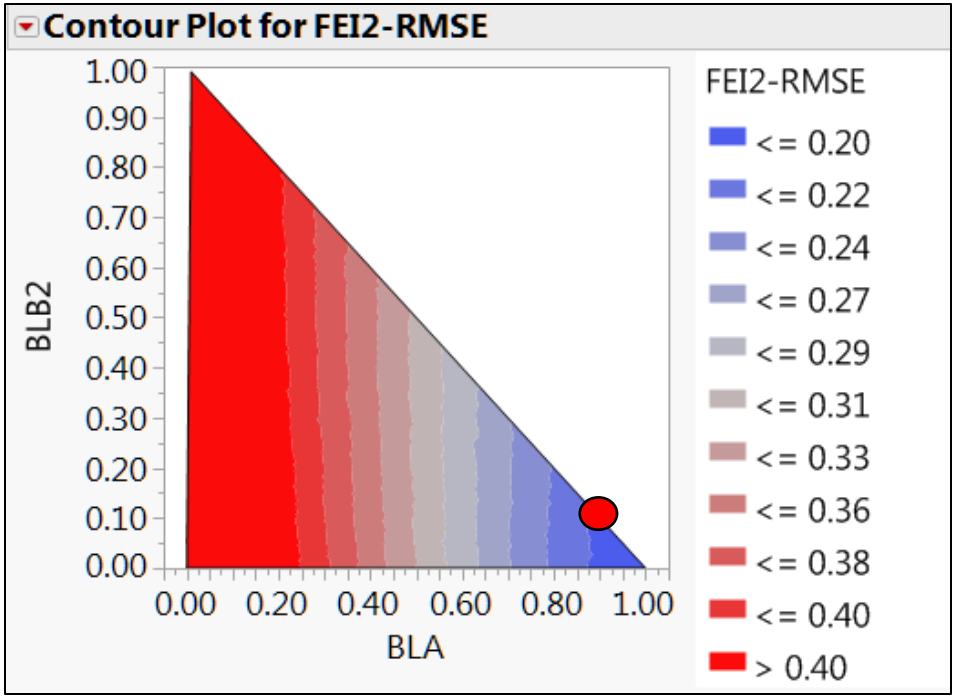


Evaluating Baseline Weight Scenarios

- Plot of RMSE vs. baseline weight combinations for FEI2 shown below
  - RMSE of weights can be interpreted from plot- if BL weights sum to 1.0
  - VID & VIE FEI2 Baseline weights are 10% & 90% (shown in red circle)
  - Test precision can be decreased with other BL weighting combinations

VIF Precision - BL Weights				
BLB1	**BLB2**	**BLA**	**FEI2-RMSE**	**EngHr Factor**
0	0.1	0.9	0.1971	Yes
0	0	1	0.1753	Yes
0	0	1	0.1775	No[1]

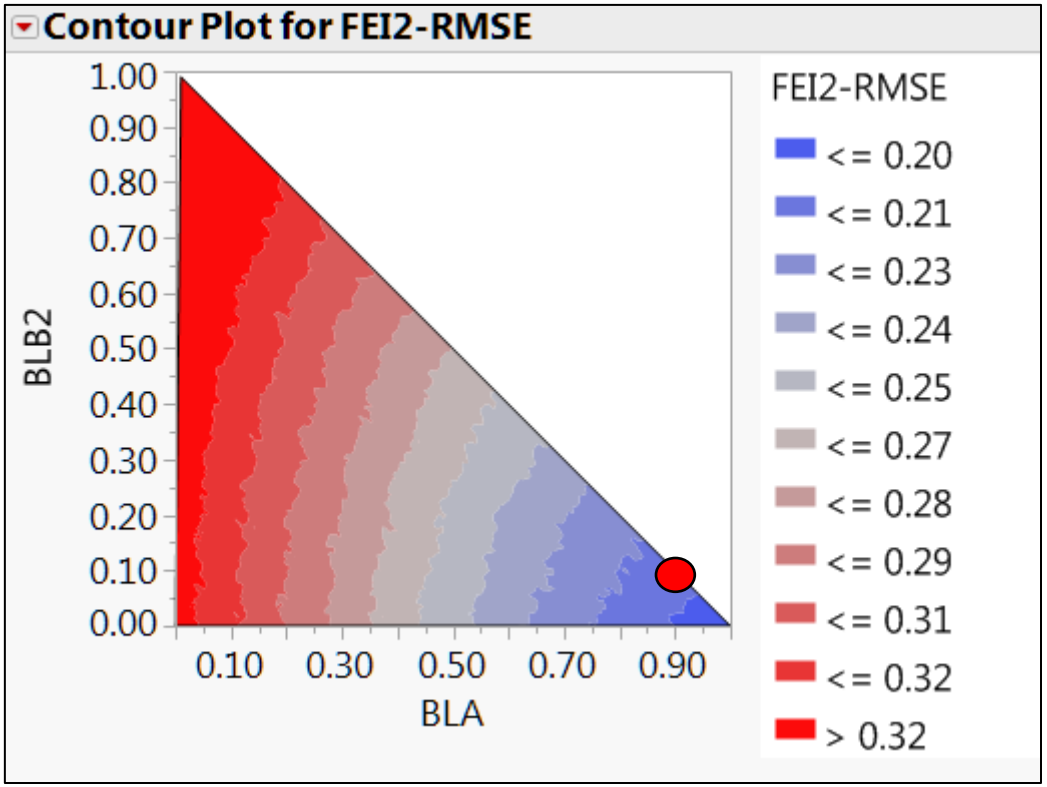
Note 1: Plot shown at right includes engine hour factor



Evaluating Baseline Weight Scenarios

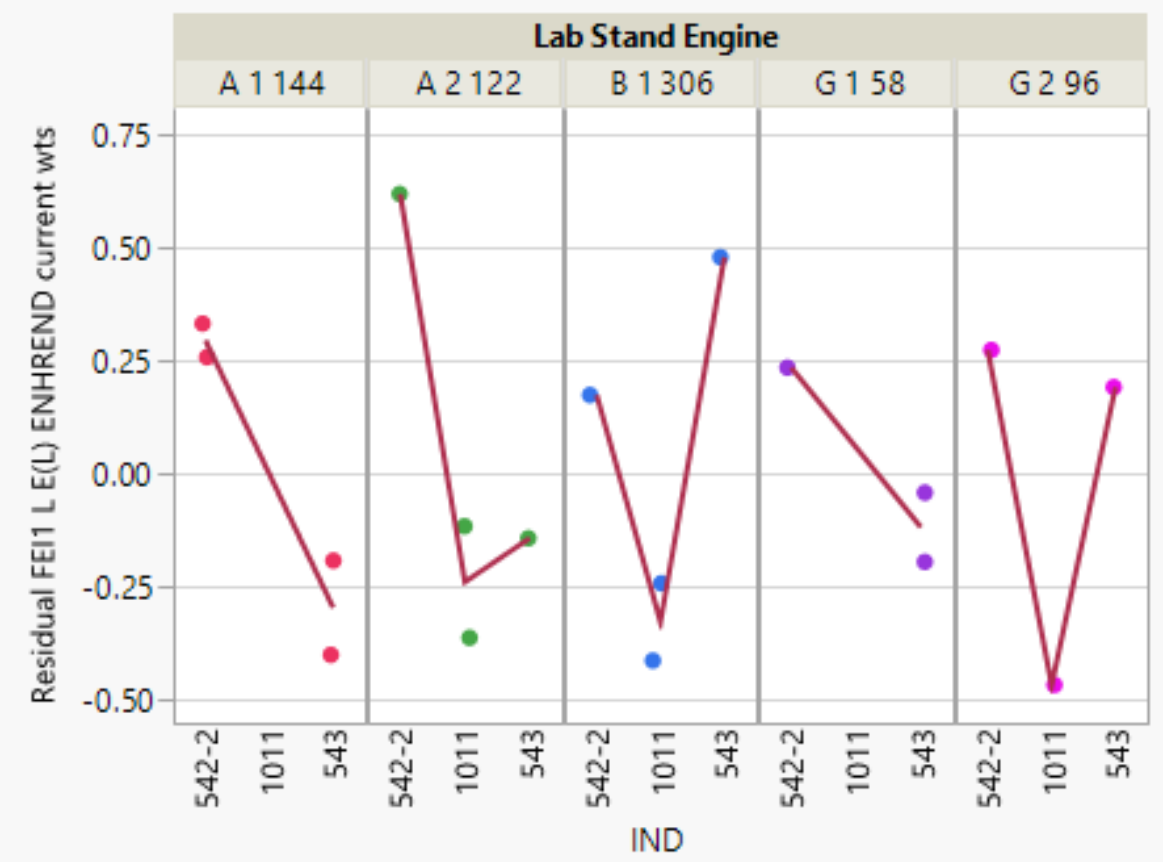
- Plot of RMSE vs. BL weight combinations for FEI2-with 1st run data deleted is shown below *(n = 14)*
  - VID & VIE FEI1 Baseline weights are 10% & 90% (shown in red circle)
  - Precision can be slightly improved with revised BL weights

VIF Precision - BL Weights			
BLB1	**BLB2**	**BLA**	**FEI2-RMSE**
0	0.1	0.9	0.2059
0	0.0	1.0	0.1910



Oil Discrimination Consistency – FEI1

- Engines do not appear to separate oils the same way, but caution should be used when basing conclusions on limited data.
- Similar differences are observed when baseline weights are used which improve test precision as shown in previous slides (100% BLB2 chosen as a representative)

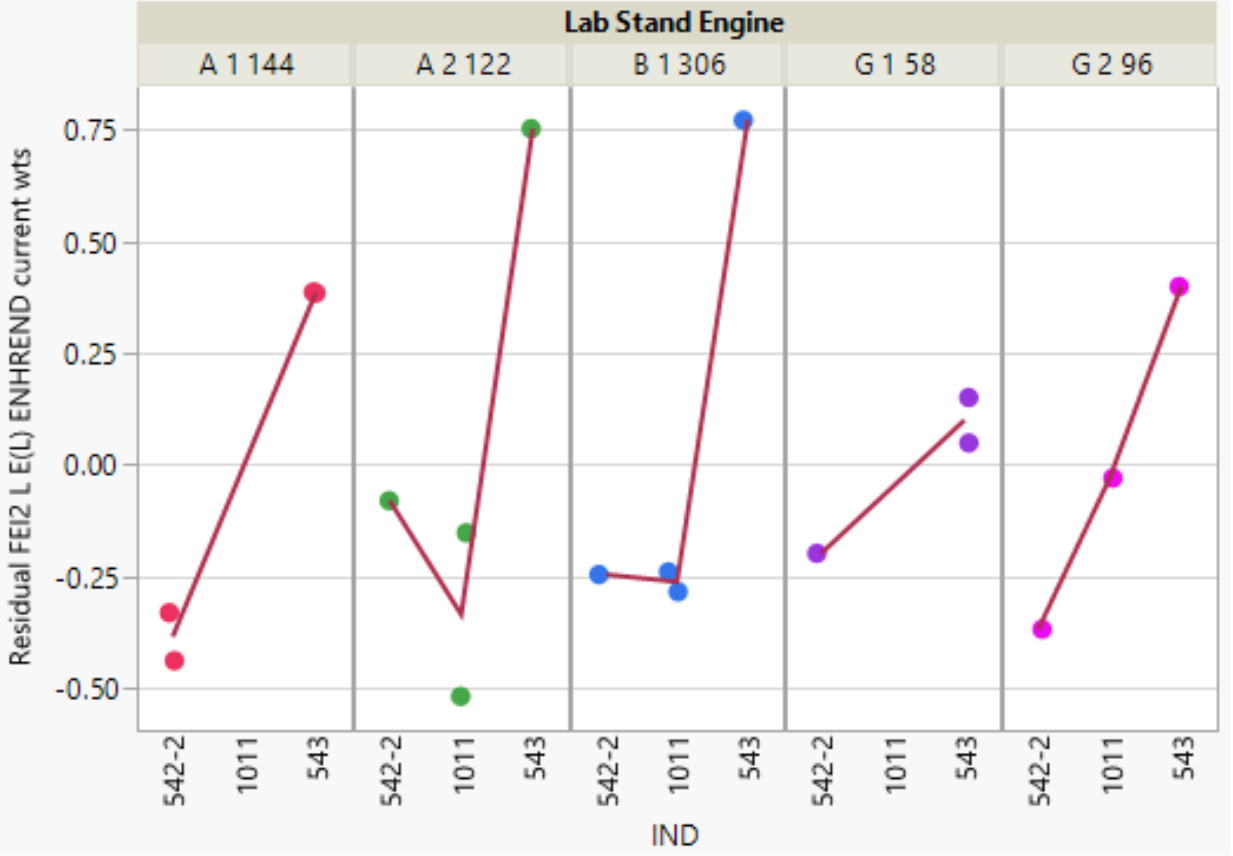


Plot assumes current/historical baseline weights: 80%BLB2 and 20% BLA

Residuals are based on models with LTMSLAB, ENGNO(LTMSLAB), and ENHREND effects

Oil Discrimination Consistency – FEI2

- Engines do not appear to separate oils the same way, but caution should be used when basing conclusions on limited data.
- Similar differences are observed when baseline weights are used which improve test precision as shown in previous slides (100% BLA chosen as a representative)

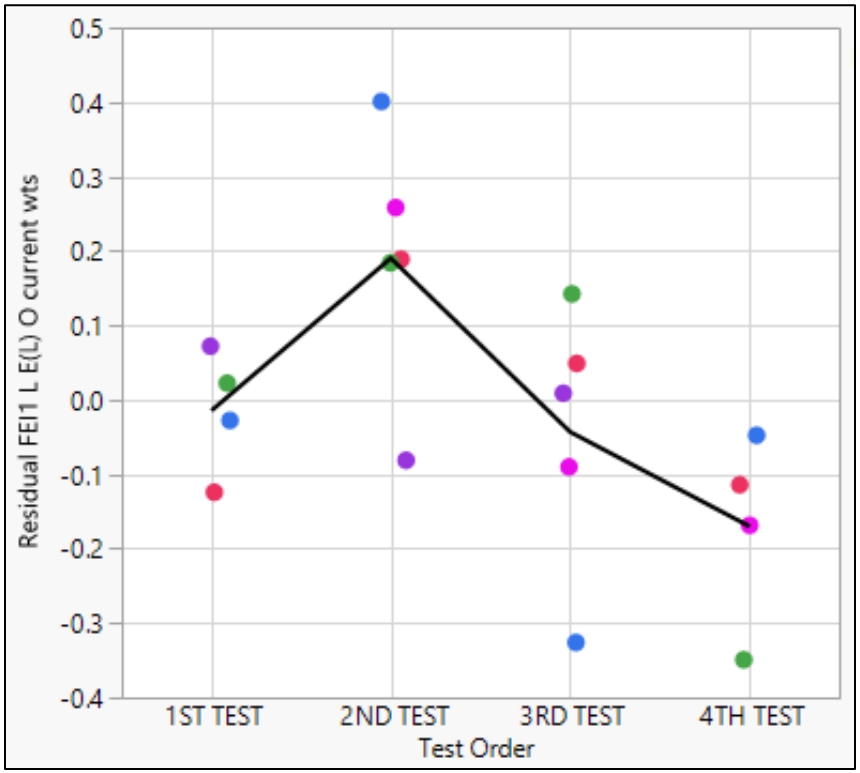


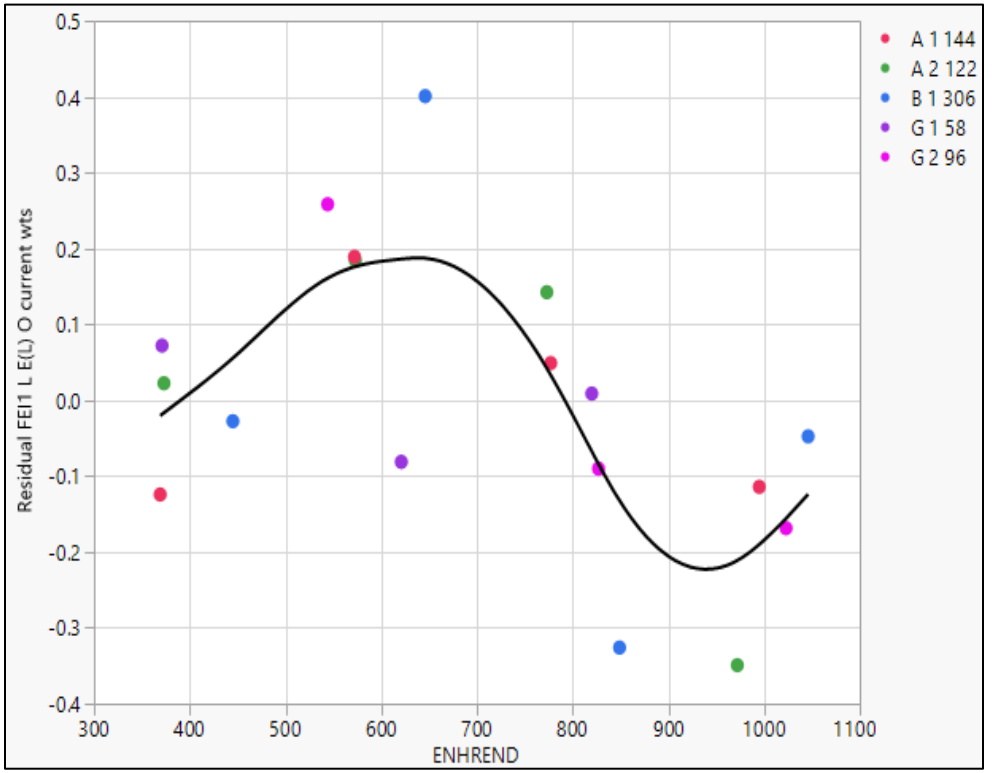
Plot assumes current/historical baseline weights: 10%BLB2 and 90% BLA

Residuals are based on models with LTMSLAB, ENGNO(LTMSLAB), and ENHREND effects

Engine Hours Effect– FEI1

- The second tests run within engines are generally the highest (in particular, higher than the first test). This could have implications on the engine hours correction used and/or engine calibration/severity adjustments.
  - Engine hour corrections in this situation are viable – See Appendix for one possibility
- Similar effect is observed when baseline weights are used which improve test precision as shown in previous slides (100% BLB2 chosen as a representative)



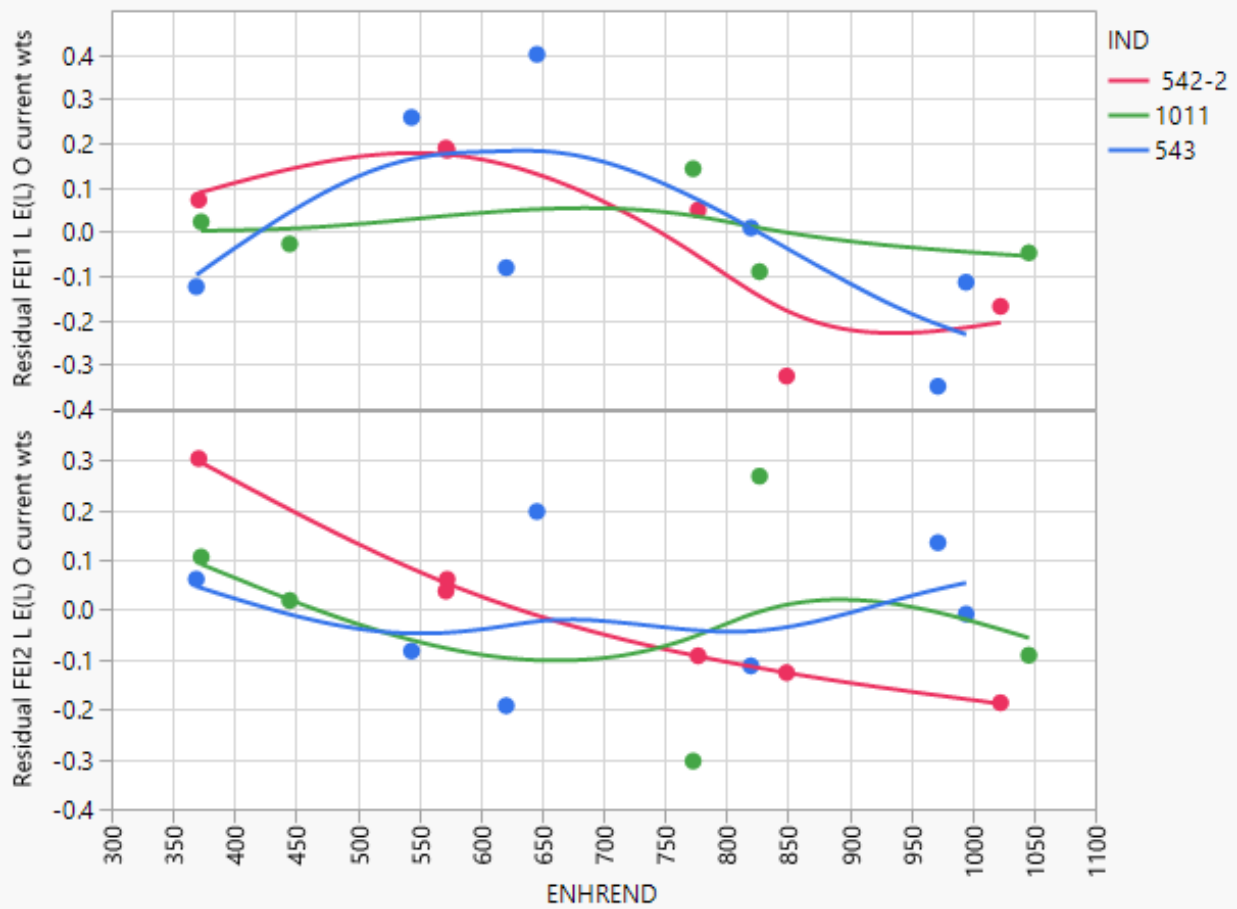


Plots assume current/historical baseline weights: 80%BLB2 and 20% BLA

Residuals are based on models with LTMSLAB, ENGNO(LTMSLAB), and Oil effects

FEI2 - Engine Hours Effect

- For FEI2, 542-2 tends to have a different engine hours effect compared to the other oils
- Although the engine hour effects for oils in FEI1 don't significantly differ, it should be pointed out that the results of the second tests within engines have an influence on the observed engine hours trend. In particular, there is lack of 1011 data in this range of engine hours.



Similar FEI2 effect is observed when baseline weights are used which improve test precision as shown in previous slides (100% BLB2 chosen as a representative)

Plots assume current/historical baseline weights: 80%BLB2 and 20% BLA

Residuals are based on models with LTMSLAB, ENGNO(LTMSLAB), and Oil effects

Evaluating Different FEI1 Modeling Scenarios

- Different FEI1 models[1] were evaluated by changing the Base Line Weights, Engine Hour effect coding, and elimination of 1st run test data.
  - For the full data set (*n=18*), the minimum RMSE corresponds to BL weights of 1.0 and 0.0 for BLB2 and BLA, respectively.
  - For the reduced data set (no first run data n=12), the minimum RMSE corresponds to the traditional BL weights of 0.8 and 0.2 for BLB2 and BLA, respectively.
  - A table of the various scenarios that were evaluated is provided below.

	Model	N Size	BLB2 Weight	BLA Weight	Model RMSE	Piece-Wise EngHr	EngHr *p* value	LSMeans			*Contrast Significant (p ≤ 0.05)*		
								RO_1011 (A)	RO_542-2 (B)	RO_543 (C)	A - B	A - C	B - C
	FEI1	18	0.8	0.2	0.2225	No	0.132	1.45	2.23	1.88	Yes	Yes	Yes
	FEI1	18	0.8	0.2	0.1965	Yes (Hrs=646)	0.031	1.47	2.22	1.87	Yes	Yes	Yes
No First Run	FEI1	14	0.8	0.2	0.1896	No	0.018	1.55	2.15	1.90	Yes	No	No
	FEI1	18	1.0	0.0	0.2050	No	0.001	1.47	2.22	1.89	Yes	Yes	Yes
	FEI1	18	1.0	0.0	0.1866	Yes (Hrs=646)	0.003	1.51	2.21	1.87	Yes	Yes	Yes
No First Run	FEI1	14	1.0	0.0	0.1912	No	0.009	1.52	2.08	1.87	Yes	No	No

[1] *Models are based on LTMSLAB, ENGNO(LTMSLAB), and Engine Hour factors (Regular vs. Piecewise)*

Evaluating Different FEI2 Modeling Scenarios

- Different FEI2 models[1] were evaluated by changing the Base Line Weights, Engine Hour effect coding, and elimination of 1st run test data.
  - For the full data set (*n=18*), the minimum RMSE corresponds to BL weights of 0.0 and 1.0 for BLB2 and BLA, respectively.
  - For the reduced data set (no first run data n=12), the minimum RMSE corresponds to the traditional BL weights of 0.1 and 0.9 for BLB2 and BLA, respectively.
  - A table of the various scenarios that were evaluated is provided below.

	Model	N Size	BLB2 Weight	BLA Weight	Model RMSE	Piece-Wise EngHr	EngHr *p* value	LSMeans RO_1011 (A)	LSMeans RO_542-2 (B)	LSMeans RO_543 (C)	*Contrast Significant (p ≤ 0.05)* A - B	A - C	B - C
	FEI2	18	0.1	0.9	0.1971	No	0.208	1.41	1.52	2.25	No	Yes	Yes
	FEI2	18	0.1	0.9	0.2057	Yes (Hrs=646)	0.380	1.42	1.52	2.24	No	Yes	Yes
No EngHrs	FEI2	18	0.1	0.9	0.1941	No Hr Factor	N/A	1.37	1.42	2.26	No	Yes	Yes
No First Run	FEI2	14	0.1	0.9	0.2059	No	0.658	1.36	1.42	2.26	No	Yes	Yes
	FEI2	18	0.0	1.0	0.1753	No	0.569	1.40	1.52	2.24	No	Yes	Yes
	FEI2	18	0.0	1.0	0.1771	Yes (Hrs=646)	0.720	1.40	1.52	2.39	No	Yes	Yes
No EngHrs	FEI2	18	0.0	1.0	0.1775	No Hr Factor	N/A	1.37	1.45	2.27	No	Yes	Yes
No First Run	FEI2	14	0.0	1.0	0.1910	No	0.837	1.45	1.45	2.27	No	Yes	Yes

[1] *Models are based on LTMSLAB, ENGNO(LTMSLAB), and Engine Hour factors (Regular, Piecewise, and none)*

Questions for the Surveillance Panel

- Should we treat 1st run results differently than the remaining tests?
- Should we change the baseline weights?
- Should we pursue a non-linear engine correction factor?
- Should we consider tests beyond the first 4?
- Should we consider FEI2, exclusively?
- Should additional testing be pursued to understand which effect(s) are "real" (oil discrimination consistency across engines, oil discrimination across engine hours, and test order)?

Some Options:

1. In the opinion of the SP the VIF data indicates performance that was not taken into account during the matrix design. Additional test development or additional test data designed to better quantify these differences is necessary. The industry will consider redevelopment or the stats group will provide additional matrix runs in an attempt to help clarify the current concerns.

   - Absolute Minimum (Engine 1: 542-2, 1011 and Engine 2: 543, 1011); 3 or 4 runs per engine is better

   - Preferred (3 to 5 engines; 6 to 8 runs per engine; revised break-in?)

2. In the opinion of the SP the VIF may perform in a fundamentally different manner from the VIE. The analysis should take this into account and minimize the variability of the available VIF data set by considering different BL weights, engine hour correction calculation methods, run limitations, etc. with the understanding that individual data points will carry significant weight in determining these changes due to the small data set available.

3. In the opinion of the SP the VIF should be similar to the VIE and any disagreement between the VIE methods of analyzing the results with the VIF matrix data is caused by the small data set available for analysis. The VIF analysis should proceed using the same BL weights, engine hour correction calculation methods, run limitations, etc. as the VIE used.

   - Engine referencing should include two tests

   - Gather 5th run (6th if we allow 3 candidates) data similar to the VIE

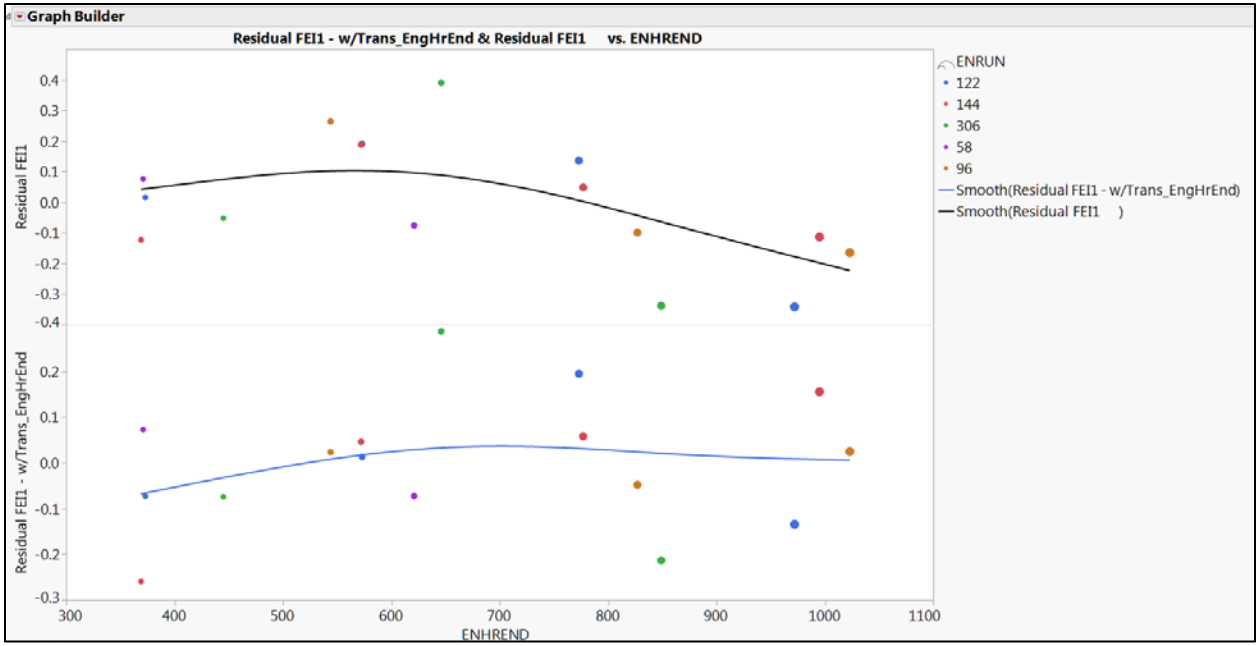
   - Revisit assumptions with more data

APPENDIX

Engine Hours Effect – FEI1

- Based on a [1]residual analysis, piecewise engine hour adjustment may be a viable alternative for FEI1.
  - If EngHrEnd > 646 then *Trans_EngHrEnd = (EngHrEnd – 646)*
  - If EngHrEnd ≤ 646 then *Trans_EngHrEnd = 0*

		VIF Precision - BL Weights			
Trans_EngHrEnd at 646	**EngHrEnd *p* value**	**BLB1**	**BLB2**	**BLA**	**FEI1-RMSE**
No	0.132	0	0.8	0.2	0.2225
No	0.001	0	1	0	0.2050
Yes	0.031	0	0.8	0.2	0.1965
Yes	0.003	0	1	0	0.1866



[1]Residuals are based on models with LTMSLAB, ENGNO(LTMSLAB), and Oil effects