Statistical Method Part A – Analysis of proposed options

[Draft]

Analysis 1

Statistical Method Part A – Analysis of proposed Options
Key concerns on Option B

1. COP based Option B is not in line with the current ISC test procedure (e.g. regarding the max. sample size); in addition, Option B is following a COP based evaluation concept which is different to the current practice in ISC
   - ISC concept is a single vehicle evaluation: pass/fail based on “inside tolerance yes/no”
   - COP concept is evaluation of the deviation between the measurements (standard deviation of the sample)
2. Option B not feasible as this procedure is evaluating the standard deviation:
   - Procedure justified for COP as vehicles in COP are expected to be statistically close to each other
   - Procedure not justified for Part A as with agreed tolerances a higher standard deviation can be expected
3. Although all single vehicles of a sample are within specified tolerances, additional measurements are required (see examples below)

		From Part A Test	From Indikator			
Cumulative sample size		SOCE_measured	SOCE_read			
1	1	85	86	Option A1	PASS	
2	1	70	73	Option A2	PASS	
3	1	90	95	Option B (Proposal EU-COM with **1.05**)		ONE MORE VEHICLE
4	0			Option B (Proposal JPN with **5%** delta)		ONE MORE VEHICLE
		From Part A Test	From Indikator			
Cumulative sample size		SOCE_measured	SOCE_read			
1	1	80	80	Option A1	PASS	
2	1	82	80	Option A2	PASS	
3	1	76	80	Option B (Proposal EU-COM with **1.05**)		ONE MORE VEHICLE
4	0			Option B (Proposal JPN with **5%** delta)		ONE MORE VEHICLE

➔ Consequence: unnecessary additional costs, efforts in vehicle acquisition and test capacity without additional value

Statistical Method Part A – Analysis of proposed Options

Further example supporting concerns on Option B

- Option B with ratio 1.05 is not working for the following example

		From Part A Test	From Indikator				
Cumulative sample size		SOCE_measured	SOCE_read				
1	1	**95**	**100**	Option A1	PASS		
2	1	**85**	**90**	Option A2	PASS		
3	1	**75**	**80**	Option B (Proposal EU-COM with **1.05**)		FAIL	
4	0			Option B (Proposal JPN with **5%** delta)	PASS		

- Option B with ration 1.05 and delta of 5% are requiring “unnecessary” additional measurements

		From Part A Test	From Indikator				
Cumulative sample size		SOCE_measured	SOCE_read				
1	1	**95**	**90**	Option A1	PASS		
2	1	**85**	**90**	Option A2	PASS		
3	1	**75**	**75**	Option B (Proposal EU-COM with **1.05**)			ONE MORE VEHICLE
4	0			Option B (Proposal JPN with **5%** delta)			ONE MORE VEHICLE

		From Part A Test	From Indikator				
Cumulative sample size		SOCE_measured	SOCE_read				
1	1	**95**	**100**	Option A1	PASS		
2	1	**85**	**85**	Option A2	PASS		
3	1	**75**	**75**	Option B (Proposal EU-COM with **1.05**)			ONE MORE VEHICLE
4	0			Option B (Proposal JPN with **5%** delta)			ONE MORE VEHICLE

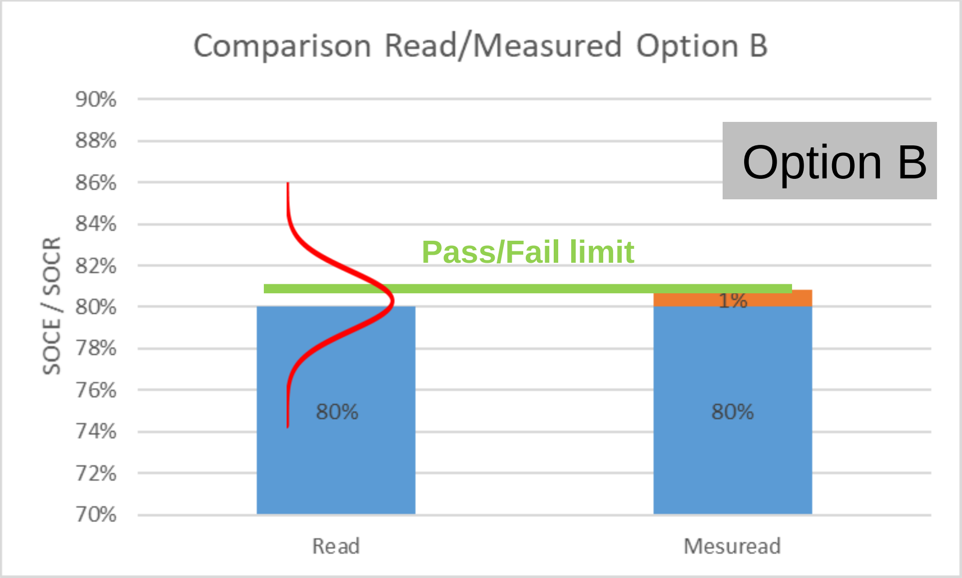
➔ First example: At least 7 vehicles are required to get a pass (although first three vehicles are within the tolerances)

➔ Second example: At least 5 vehicles are required to get a pass (although first three vehicles are within the tolerances)

Conclusion: Option B as now is resulting in unnecessary additional costs, efforts in vehicle acquisition and test capacity.
This additional efforts should be avoided as Part A vehicles are customer vehicles and as these add no additional value.

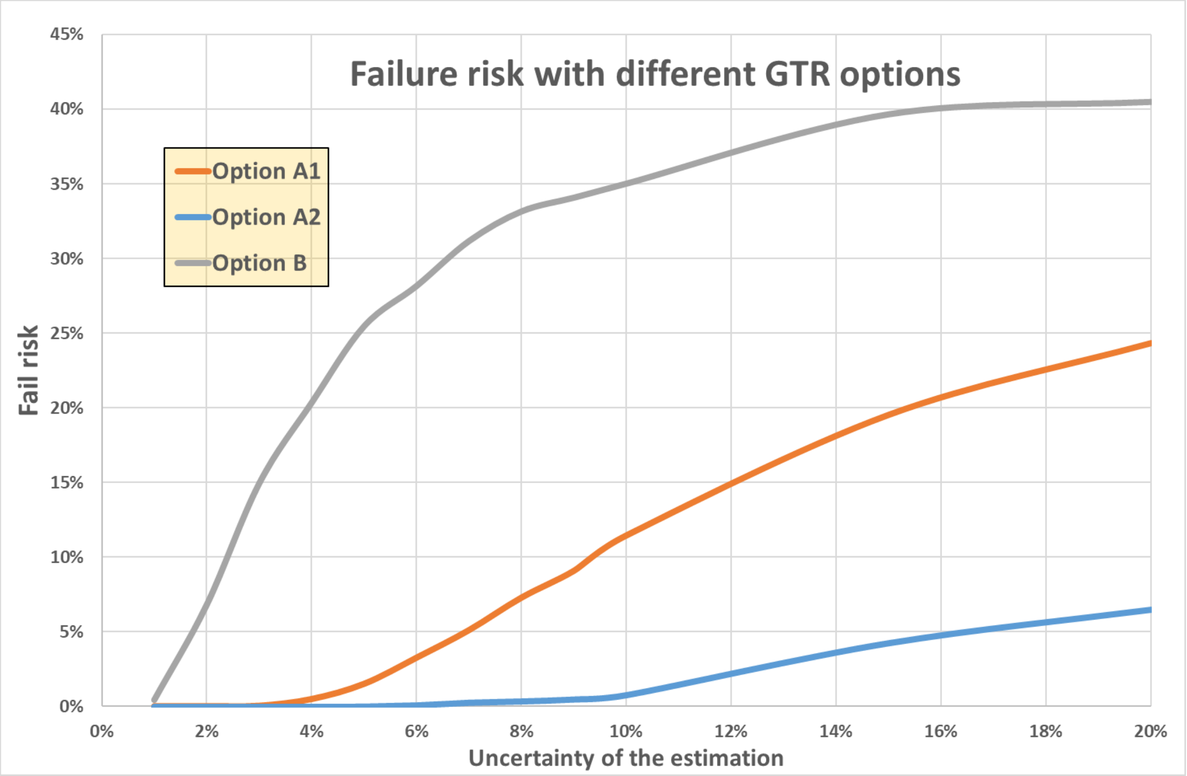
Statistical Method Part A – Analysis of proposed Options
Further explanations supporting concerns on Option B

- **Option B : SOCE/SOCR read are compared to SOCE/SOCR measured**
- **Due to the factor A [1.01] , the pass/fail limit is less than 1% over the measured value.**
- **With the uncertainty of the estimation, a non negligible part of the population is over the pass/fail limit**



Making simulation with different uncertainties to calculate the risk of failure of a sample lot :

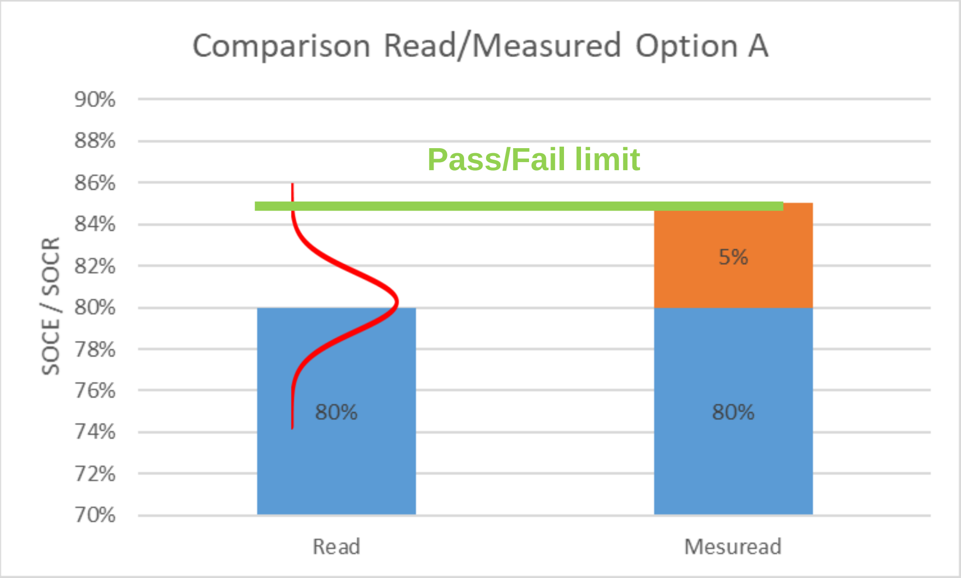
- **the more stringent Option is B**



Statistical Method Part A – Analysis of proposed Options

Reflections on Option A

- **Option A : SOCE/SOCR read are compared to SOCE/SOCR measured + 5% points**
- **With the uncertainty of the estimation, a large part of the population is under the pass/fail limit**



Option A1

WLTP	10								
	9								
	8								
	7								
	6								
	5								
	4								
	3								
	2								
	1								
	0								
seuil accept		0	1	1	1	2	2	3	5
seuil refus		3	3	4	5	5	6	6	6

Option A2

R83	20																		
	19																		
	18																		
	17																		
	16																		
	15																		
	14																		
	13																		
	12																		
	11																		
	10																		
	9																		
	8																		
	7																		
	6																		
	5																		
	4																		
	3																		
	2																		
	1																		
	0																		
seuil accept		0	1	1	2	2	3	4	4	5	5	6	6	7	8	8	9	9	11
seuil refus		-	-	5	6	6	7	8	8	9	9	10	11	11	12	12	13	13	12

Statistical Method Part A – Analysis of proposed Options

Option A as fallback solution if Option B concerns cannot be eliminated

If Option B concerns cannot be resolved and eliminated, the only way forward is Option A (A1 or A2).

Reflections on Option A1:

- Option A1 in line with new ISC test procedure in WLTP
- Weights the single vehicle more and therefore statistics will come faster to a decision

Reflections on Option A2:

- Option A2 in line with old ISC test procedure and preferable from a statistical point of view

Analysis 2

Statistical Method Part A – Analysis of proposed Options
Background and scope of the Analysis

<u>Compare Part A methods for verification of in-use monitors (SOCE/SOCR)*</u>

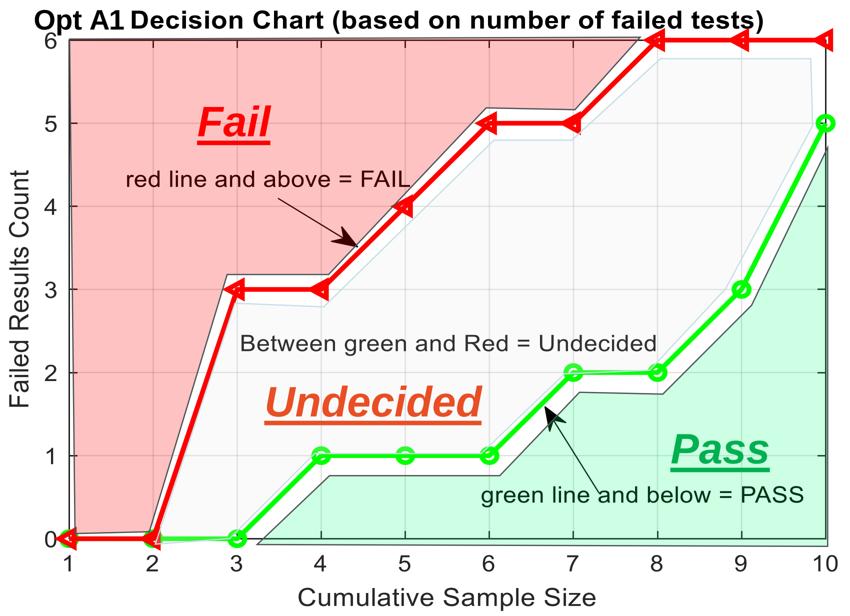
- Three statistical methods have been proposed in the GTR draft (Opt A1, Opt A2, and Opt B) to make a PASS or FAIL decision for the PART A monitor family.
- A Monte Carlo simulation with virtual test data was used to evaluate the various methods.
- Two variables were used in this analysis → $SOCE_{read}$ and $SOCE_{measured}$.
- Both quantities are expressed as a percentage of the declared value and have the units of %.
- The PASS rate for the three methods were compared.

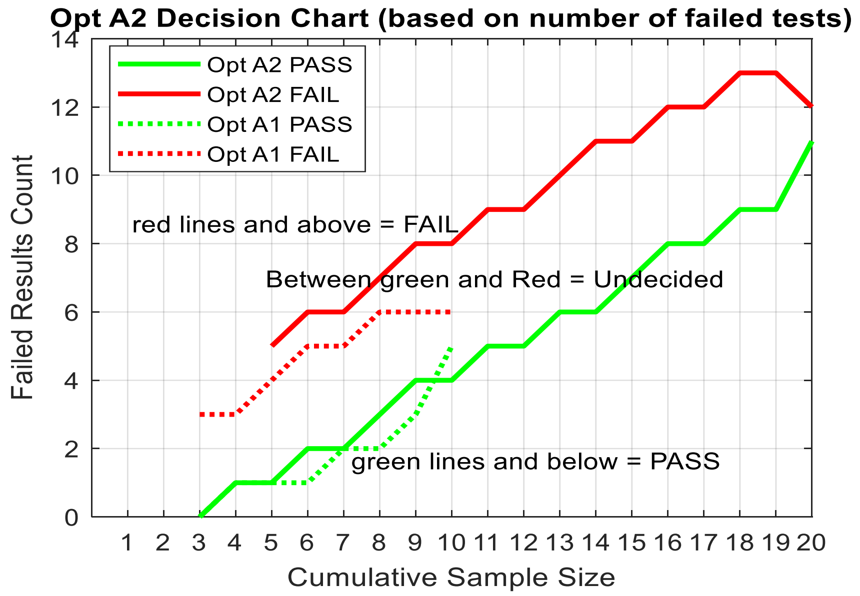
*See 6.3. Part A: Verification of SOCR/SOCE monitors in *“Proposal for a new UN GTR on In-vehicle Battery Durability for Electrified Vehicles”*

Statistical Method Part A – Analysis of proposed Options

Method Option A1 and Option A2 to verify a Monitor Family

- Both methods use multiple tests (or samples) for the verification.
- Decision of PASS/FAIL of a monitor family is based on the number of failed individual tests used for the verification.
- Individual test Pass/Fail ➔ "A vehicle test shall be considered a pass (p) when the estimated SOCR and estimated SOCE read from the vehicle are both not more than 5 percentage points greater than the respective measured value."
  - Pass ➔ if $SOCE_{read} - SOCE_{measured} <= 5$
  - Fail ➔ if $SOCE_{read} - SOCE_{measured} > 5$
- Refer to charts on right for a PASS/FAIL decision for the monitor family.





Observations

- Opt A1 uses max of 10 tests, while Opt A2 uses 20 tests.
- Within use of 10 tests, Opt A1 is easier to FAIL than Opt A2.
- Within 10 tests, Opt A1 is more difficult or equal to PASS than opt A2, except when using exactly 10 tests.

Statistical Method Part A – Analysis of proposed Options
Method Option B to verify a Monitor Family

X_{test} = mean for x_i *s=standard deviation for x_i*   ***A=1.01***

$$x_i = \frac{SOCE_{read,i}}{SOCE_{measured,i}}$$

For each number of tests $3 \leq N \leq 16$, one of the three following decisions can be reached, where the factor A shall be set at [1.01]:

(i) Pass the family if $X_{tests} \leq A - \left(t_{P1,N} + t_{P2,N}\right) \cdot s$

(ii) Fail the family if $X_{tests} > A + \left(t_{F1,N} - t_{F2}\right) \cdot s$

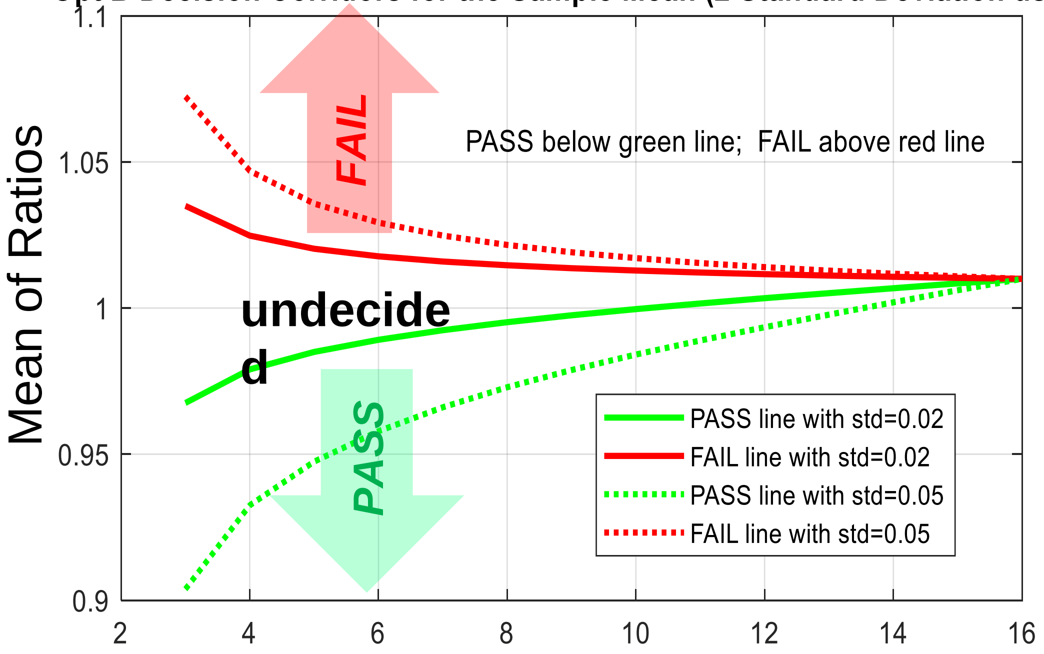
(iii) (UND) Take another measurement if:

$$A - \left(t_{P1,N} + t_{P2,N}\right) \cdot s < X_{tests} \leq A + \left(t_{F1,N} - t_{F2}\right) \cdot s$$

where the parameters $t_{P1,N}$, $t_{P2,N}$, $t_{F1,N}$ and t_{F2} are taken from Table 4.

	PASS		FAIL	
Tests (N)	$t_{P1,N}$	$t_{P2,N}$	$t_{F1,N}$	t_{F2}
3	1.686	0.438	1.686	0.438
4	1.125	0.425	1.177	0.438
5	0.850	0.401	0.953	0.438
6	0.673	0.370	0.823	0.438
7	0.544	0.335	0.734	0.438
8	0.443	0.299	0.670	0.438
9	0.361	0.263	0.620	0.438
10	0.292	0.226	0.580	0.438
11	0.232	0.190	0.546	0.438
12	0.178	0.153	0.518	0.438
13	0.129	0.116	0.494	0.438
14	0.083	0.078	0.473	0.438
15	0.040	0.038	0.455	0.438
16	0.000	0.000	0.438	0.438

Opt B Decision Corridors for the Sample Mean (2 Standard Deviation used)



Observations

- Read / Measured ratio < 1 is required to get a high PASS rate for a monitor family.
- Mean and standard deviation for the read / measured ratios are the determining factors for this method.

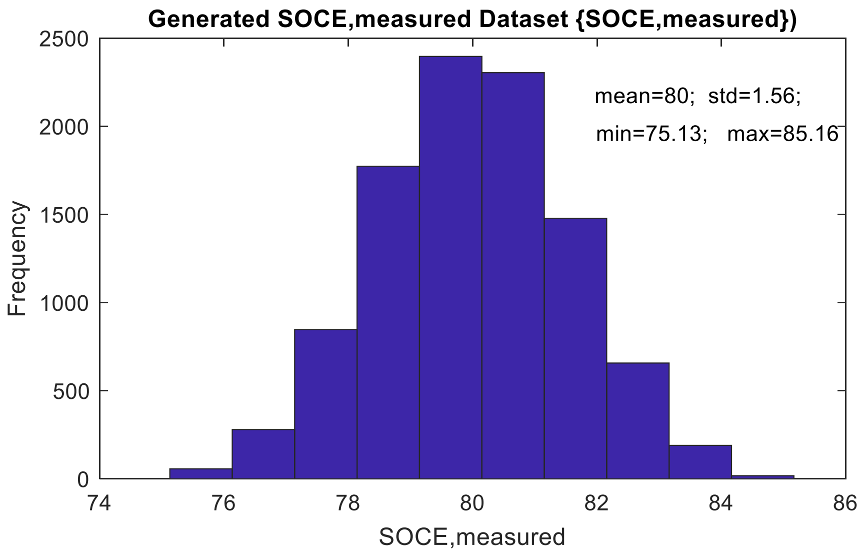
Statistical Method Part A – Analysis of proposed Options
General observation from methods Option A1, A2 and Option B

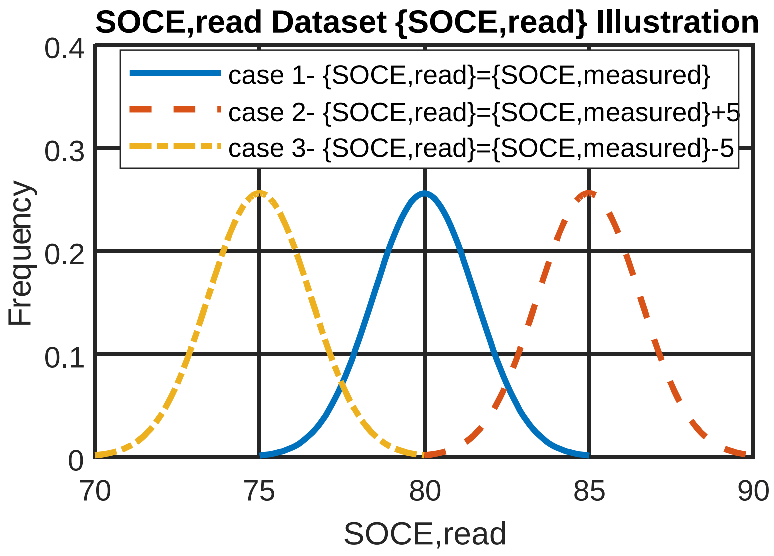
- To reach a final PASS/FAIL decision for a monitor family, the number of tests can be up to 10 (Opt A1), 20 (Opt A2), and 16 (Opt B).
- Based upon the basic definition of the three methods, the expected results due to range/relationship of $SOCE_{read}$ to $SOCE_{measured}$ are below:

Relationship between $SOCE_{read}$ & $SOCE_{measured}$	Opt A1&Opt A2	Opt B
$SOCE_{read} > SOCE_{measured} + 5$	FAIL	FAIL
$SOCE_{measured} < SOCE_{read} <= SOCE_{measured} + 5$	PASS (partial)	FAIL (likely)
$SOCE_{read} <= SOCE_{measured}$	PASS	PASS (likely)

Statistical Method Part A – Analysis of proposed Options
Simulation and Data Set Information

- Virtual test data was generated for $SOCE_{measured}$
- Assumptions for $SOCE_{measured}$ dataset
  - Ranges - 75~85
  - Data distribution = normal with mean = 80 and StdDev between 1~2
- $SOCE_{read}$ dataset scenarios evaluated
  - Case 1: $\{SOCE_{read}\} = \{SOCE_{measured}\}$
  - Case 2: $\{SOCE_{read}\} = \{SOCE_{measured}\} + 5$
  - Case 3: $\{SOCE_{read}\} = \{SOCE_{measured}\} – 5$
- Simulation – An example for one run
  1) *Randomly* select a pair of samples from $\{SOCE_{read}\}$ & $\{SOCE_{measured}\}$
  2) For method A1 and A2, perform individual pass/fail test (fail if sample $SOCE_{read}$ - sample $SOCE_{measured} > 5$)
  3) Count the # of fails (initially start with 3 pairs)
  4) Apply A1 & A2 decision charts for the PASS/FAIL of the family. If the decision is undecided, add a pair of samples and repeat 1~4.
  5) For Opt B, calculate mean and StdDev and apply decision corridors. Add a pair of samples in the case of an undecided result. Repeat until PASS/FAIL.





Statistical Method Part A – Analysis of proposed Options
Summary

Observations:

- Case 3 results in a similar PASS rate for all three methods (using any number of tests).
- Cases 1 & 2 results in a significantly lower PASS rate for Opt B
- Opt A2 has slightly higher PASS rate than Opt A1 in Case 2 when using any number of tests.

Recommendations

- **Opt B is not recommended:**
  - Inherently forces a stricter Pass/Fail criteria than Opt A1 or A2
  - The $SOCE_{read}$ must be less than $SOCE_{measured}$ for it to be a logical usage
  - Does not reflect the 5% allowance for the individual pass/fail test criteria in the GTR text
- Opt A1 & A2 are recommended with a preference for A2 having a slightly higher passing rate for the cases evaluated

Summary Table 1 PASS Rate (Approximation) from Three Methods Using Number of Tests Matching Opt A1			
	Opt A1	Opt A2	Opt B
Case 1 ({SOCEread} = {SOCEmeasured})	100%	100%	8%
Case 1 # of tests used	3~7	3~7	3~7
Case 2 ({SOCEread} = {SOCEmeasured}+5)	55%	45%	0%
Case 2 # of tests used	3~10	3~10	3~10
Case 3 ({SOCEread} = {SOCEmeasured}-5)	100%	100%	80%
Case 3 # of tests used	3	3	3

Summary Table 2 PASS Rate (Approximation) from Three Methods Using *ANY* Number of Tests			
	Opt A1	Opt A2	Opt B
Case 1 ({SOCEread} = {SOCEmeasured})	100%	100%	83%
Case 1 # of tests used	3~7	3~7	3~16
Case 2 ({SOCEread} = {SOCEmeasured}+5)	55%	80%	0%
Case 2 # of tests used	3~10	3~20	3~10
Case 3 ({SOCEread} = {SOCEmeasured}-5)	100%	100%	100%
Case 3 # of tests used	3	3	3~7

Conclusion

Statistical Method Part A – Analysis of proposed Options
Conclusion

Considering Option A

- If the « uncertainty » of the estimation (Standard deviation/average) is under 5% ,
  Option A1 or Option A2 are acceptable with a failure risk under 1%.
- Option A2 is always more acceptable than Option A1

Considering Option B

- With an uncertainty of 2%, there is a failure risk of 5% for a sample lot
- To reach a failure risk of 2% (or less) with Option B, this would be only possible an underestimation of the SOCE/SOCR

Challenges coming along with range test on customer vehicles

- Customer acceptance for a range test will be questionable (as significantly more mileage) → challenge for the manufacturer

→ That is why manufacturer’s interest will be to keep the number of measured vehicles in a sample at the minimum

→ Preference would be to have no range test on dyno but understood that this is required for determination of UBE