An Example

With augmented plans we cannot use ANOVA methods to derive parameter estimates. Instead, we recommend maximum likelihood estimation with standard errors calculated from the Fisher information matrix. We describe an artificial example to demonstrate the analysis when the total number of measurements is $N = 60$ and the number of operators is m = 3. We use the recommended plan A(6,2,24). The Stage 1 and Stage 2 data are given below. In Stage 1, six randomly selected parts were measured twice by each of the three operators. In Stage 2, another 24 parts were randomly selected and eight were measured once by each of the three operators.

Stage 1

Part	Operator	Replicate	Response	Part	Operator	Replicate	Response
1	1	1	10.4	4	1	1	22.5
1	1	2	9.6	4	1	2	22.3
1	2	1	10.8	4	2	1	24.4
1	2	2	10.8	4	2	2	24.1
1	3	1	10.6	4	3	1	23.3
1	3	2	10.4	4	3	2	23.3
2	1	1	23.8	5	1	1	22.9
2	1	2	22.5	5	1	2	22.6
2	2	1	23.3	5	2	1	22.3
2	2	2	21.7	5	2	2	22.2
2	3	1	23.1	5	3	1	22.5
2	3	2	23.3	5	3	2	23.1
3	1	1	25	6	1	1	18.1
3	1	2	25.3	6	1	2	17.5
3	2	1	26	6	2	1	19.8
3	2	2	25.4	6	2	2	20.6
3	3	1	27.6	6	3	1	17.7
3	3	2	28	6	3	2	19.6

Stage 2

Operator 1	Operator 2	Operator 3
18.8	12	31.6
16.3	18.9	18.2
5.1	18	11.2
23.8	18.5	15
28.3	26.4	28.6
11.9	25.6	22.3
17.2	27.1	27.4
23.4	24.6	24.5

With this data we called the Matlab code MLEresults – see example call inside the code.
From the Matlab code we find that the estimate of γ is $\hat{\gamma} = 0.15$ with a standard error of 0.03.
The estimates of the operator means are $\hat{\mu}_1 = 20.07, \hat{\mu}_2 = 20.87, \hat{\mu}_1 = 20.97$ so that $\hat{\sigma}_0 = 0.402$ with standard error 0.18.
The estimates of σ_d and σ_m are 0.65 and 0.55 with corresponding standard errors 0.18 and 0.09 respectively.

For one to analyse a data set such as this using the Matlab code, one must enter the data into the command in the following manner. Suppose first that the analysis consists of N total measurements where k parts are sampled in Stage 1, and measured n times by each of m operators, and where k_{T} parts (T represents augmented plan type; A or B) are sampled for Stage 2. The way in which the k_{T} parts are measured in Stage 2 depends on whether the measurement study used Augmented Plan A or B. Furthermore, the analysis depends on whether a part-by-operator interaction is assumed to exist.

For the analysis to work properly, the kmn pieces of Stage 1 data $\{y_1, y_2, ..., y_{kmn}\}$ must be ordered by part, and then by operator. Also, the k_{T} pieces of Stage 2 data $\{z_1, z_2, ..., z_{k_T}\}$ must be ordered by operator. The Matlab call would look as follows:

[res]=MLEresults(k, m, n, k_{T}, T, I ,[$y_1, y_2, ..., y_{kmn}$],[$z_1, z_2, ..., z_{k_T}$]);

Note that T can only take on values 1 and 2. Here, 1 represents Augmented Plan A, and 2 represents Augmented Plan B. Also, I can only take on values 0 and 1, where 1 is entered if a part-by-operator interaction is to be included, and 0 is entered if it is not.

For the example described above, the Matlab input and output are given below. The output gives the maximum likelihood estimates and standard errors of all of the associated parameters; namely, the operator means (μ_j 's), the components of variation ($\sigma_p, \sigma_o, \sigma_d, \sigma_m$), and the primary metric of interest, γ .

```
[res]=MLEresults(6,3,2,24,1,1,[10.4; 9.6; 10.8; 10.8; 10.6; 10.4; 23.8; 22.5; 23.3;
21.7; 23.1; 23.3; 25.0; 25.3; 26.0; 25.4; 27.6; 28.0; 22.5; 22.3; 24.4; 24.1; 23.3;
23.3; 22.9; 22.6; 22.3; 22.2; 22.5; 23.1; 18.1; 17.5; 19.8; 20.6; 17.7; 19.6],[18.8;
12.0; 31.6; 16.3; 18.9; 18.2; 5.1; 18.0; 11.2; 23.8; 18.5; 15.0; 28.3; 26.4; 28.6;
11.9; 25.6; 22.3; 17.2; 27.1; 27.4; 23.4; 24.6; 24.5]);

Augmented Plan type A with 6 parts 3 operators and 2 repeated measurements, plus 24
additional parts in the augmented plan

total number of measurements = 60

            operator means
MLEs    = 20.0735      20.8661      20.9752
MLE SEs = 1.1472       1.1472       1.1472
             sp           so           sd           sm         gamma
MLEs     = 6.0812     0.40179      0.65348      0.54735      0.15314
MLE SEs = 0.79838     0.17722      0.18418     0.091224     0.030936
```

Now that the estimates have been found, it is important to assess whether the model being used is adequate. The model makes a number of assumptions, some of which can be checked using simple graphical tools. Since Stage 1 uses a standard plan, we can follow the AIAG recommendations and plot these data in a number of ways. A QQ plot of the Stage 2 data, after subtracting the corresponding operator average can be used to check both the overall normality and the assumption that the measurement variation is the same for each operator.

Figure 1 below displays various graphical outputs resulting from a GR&R ANOVA analysis on the Stage 1 data from MINITAB. The first plot (top left-hand corner) depicts the extent to which the various variance components contribute to the overall variation. We can see clearly that part-to-part variation accounts for almost all of the total variation, further supporting the MLE estimate of gamma, $\hat{\gamma} = 0.15$, which says that the measurement system is adequate. The next two plots of interest are the top two on the right-hand side of Figure 1. Respectively, they show the measurement data stratified by part and by operator. In the uppermost plot, we see that relative to the part-to-part variation, there is little measurement variation. The lower plot suggests the operator averages are similar. The plot located in the bottom right-hand corner of Figure 1 has been replicated on a larger scale due to it's importance. See Figure 2.

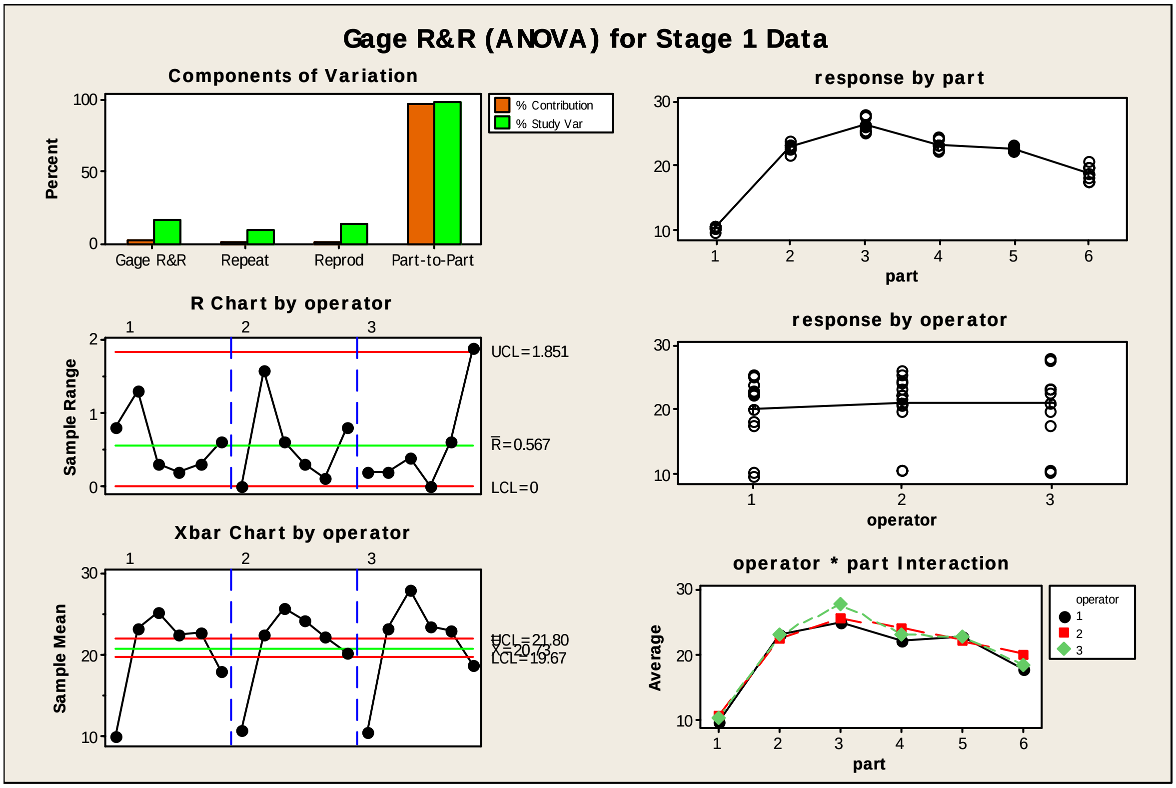

Figure 1: MINITAB Gauge R&R Plots

Figure 2 suggests a part-by-operator interaction. Particularly on Part #3, and to an extent on Parts #4 and 6, we notice a divergence from the more tightly clustered sets of points that we see for Parts #1, 2, and 5. This tells us that operator relative bias is not the same across all parts, and instead varies part-by-part.

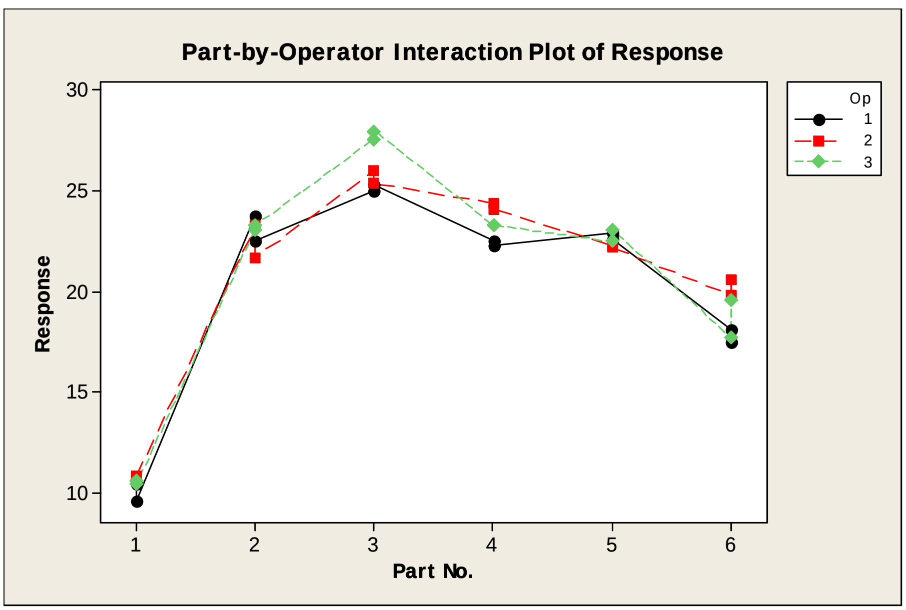

Figure 2: Part by Operator Interaction Plot

Figure 3 plots all the measurement data stratified by both part and operator. As in Figure 2 it suggests a part-by-operator interaction. For example, we see more variation in the operators' measurements on Parts #3, 4, and 6, than we do on Parts #1, 2, and 5. In particular, Operator 3's repeated measurements on Part #3 are much higher than those taken by Operator 1 and 2. Similarly, we see that Operator 1's measurements on Part #4 are considerably lower than those taken by Operator 2 and 3. The repeated measurements on Part #6 seem to be quite different for each Operator. Again, this exemplifies the fact that the operator relative bias differs on a part-by-part basis.

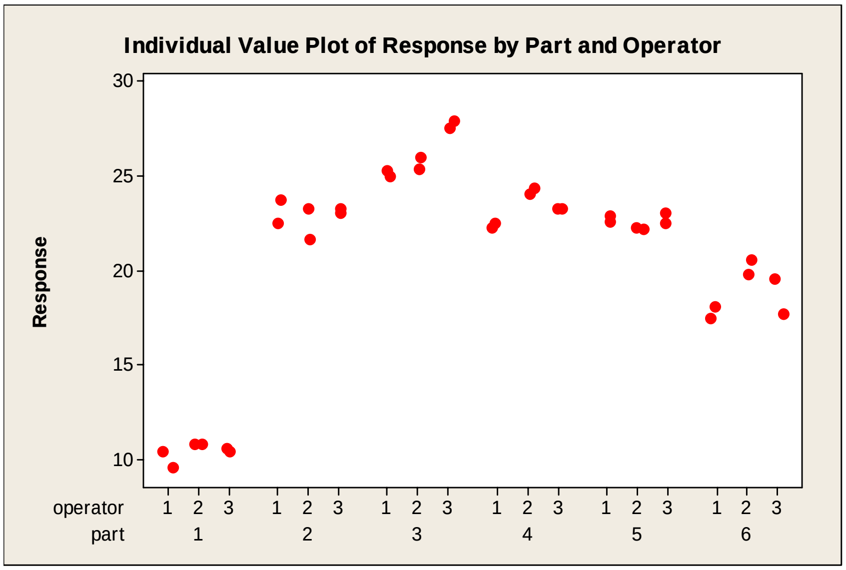

Figure 3: Measurement Data Stratified by Part and Operator

The ANOVA table associated with this GR&R study is given below. The Sum of Squares breakdown confirms what has been depicted pictorially: part-by-part variation accounts for almost all of the overall variation, and although somewhat overshadowed, the part-by-operator interaction accounts for a significant proportion of the variation (P-value = 0.002)

Source	DF	SS	MS	F	P
part	5	927.907	185.581	134.560	0.000
operator	2	5.012	2.506	1.817	0.212
part * operator	10	13.792	1.379	4.606	0.002
Repeatability	18	5.390	0.299		
Total	35	952.100			

We now consider the Stage 2 data. First of all we plot the 24 measurement by operator in Figure 4. Due to the large part to part variation we are not surprised to see that the results for each operator seem to have the roughly the same mean and variation.

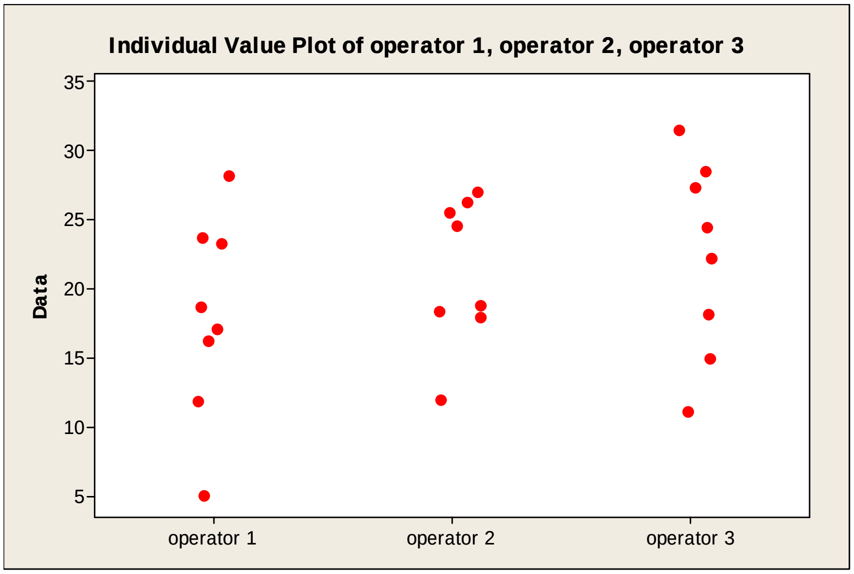

Figure 4: Stage Two Data Stratified by Operator

Figure 5 gives a graphical analysis of the distributional assumptions associated with the Stage 2 augmented data. In this analysis we assume that the data are normally distributed (with mean 0 and constant standard deviation). To check these assumptions, we subtract the respective operator mean off of each of their measurements in Stage 2, and construct a QQ-plot of this adjusted data. The QQ-plot measures quantile-by-quantile the probability that the empirical data could have come from the theoretical distribution. A relatively straight line on a 45 degree angle would suggest that our assumptions are appropriate. Although we do not see a perfectly straight line, we do however notice that all of the points fall within the two outer bands, which represent a 95% confidence interval.

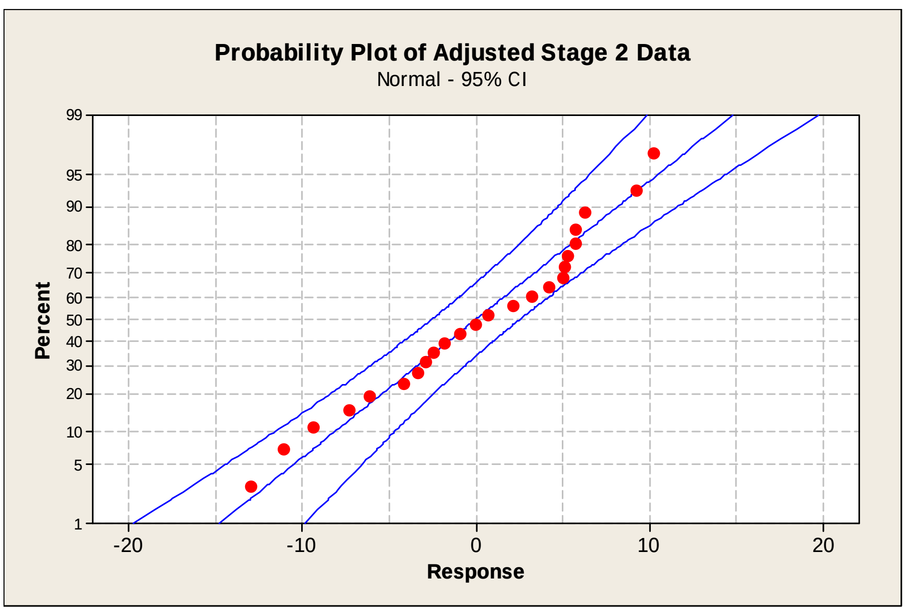

Figure 5: Normal QQ-Plot of Adjusted Stage 2 Data

All in all, it would appear as though the assumptions associated with the model are realistic, and hence the model is adequate. In turn, this suggests that the maximum likelihood estimates of the parameters are reliable. In particular, the estimate $\hat{\gamma} = 0.15$ is reliable. Note that a measurement system is deemed to be acceptable if γ is less than 0.1, unacceptable if γ is greater than 0.3, and is in need of improvement if $0.1 < \gamma < 0.3$. Thus, because $\hat{\gamma}$ is closer to 0.1 than to 0.3 ($\hat{\gamma} \approx 0.1$), we can conclude that the measurement system used in this example is acceptable.