

# Good Data, Bad Data, and Process Behavior Charts

Donald J. Wheeler

Over the years many different superstitions have grown up regarding the importance of having high quality measurements before you start using them on a process behavior chart. This paper outlines a simple experiment you can perform for yourself that should clarify this issue.

## Our Model

---

In order to understand how different amounts of measurement error will affect the ability of a process behavior chart to detect process changes we shall construct some artificial data with a known structure and with known changes and see how the charts react.

Let  $Y$  denote the product values. Let  $E$  denote the measurement errors. And let  $X$  denote the product measurements. Thus our model for a given measurement is:

$$X = Y + E$$

We shall assume that the  $Y$  values are normally distributed with a mean of zero and a standard deviation of  $\sigma_y$ . We shall also assume that the  $E$  values are normally distributed with a mean of zero and a standard deviation of  $\sigma_e$ . Then, as a matter of course,  $X$  will also be normally distributed with a mean of zero and a variance of:

$$\sigma_x^2 = \sigma_y^2 + \sigma_e^2$$

Clearly, when  $\sigma_e$  is small relative to  $\sigma_y$ , the values for  $E$  will be small,  $X$  will be quite similar to  $Y$ , and the measurements will be said to be good. As  $\sigma_e$  increases in size relative to  $\sigma_y$ , measurement error will become a more prominent part of each number, and the measurements will contain more and more noise. The traditional way of quantifying this relationship was introduced by Sir Ronald Fisher in 1921 and is defined to be the ratio of the product variation  $\sigma_y^2$  to the total variation  $\sigma_x^2$ :

$$\text{Intraclass Correlation Coefficient} = \rho_1 = \frac{\sigma_y^2}{\sigma_x^2}$$

This ratio correctly shows that proportion of the total variation in  $X$  that can be attributed to the product values  $Y$ . The complement to this ratio:

$$1 - \rho_1 = \frac{\sigma_x^2 - \sigma_y^2}{\sigma_x^2} = \frac{\sigma_e^2}{\sigma_x^2}$$

will characterize that proportion of the total variation in  $X$  that can be attributed to the measurement error.

The common estimate of the Intraclass Correlation Coefficient is a biased estimate which will be called the Intraclass Correlation Statistic:

$$r_I = \frac{\hat{\sigma}_y^2}{\hat{\sigma}_x^2} = \frac{\hat{\sigma}_x^2 - \hat{\sigma}_e^2}{\hat{\sigma}_x^2} = 1 - \frac{\hat{\sigma}_e^2}{\hat{\sigma}_x^2}$$

We use the last form for this statistic because, in practice, we cannot estimate  $\sigma_y$  directly. While the Intraclass Correlation Coefficient is bounded by 0 and 1, the statistic above can occasionally turn out to be less than zero. When this happens it is simply an indication that the product variation is so small that it has been overwhelmed by the uncertainty in the estimates of  $\sigma_x$  and  $\sigma_e$ .

Rather than using this traditional statistic, the industrial technique known as a “gauge R&R study” uses the ratio of  $\sigma_e$  to  $\sigma_x$  to characterize measurement systems into three categories. Since this gauge R&R ratio is completely different from the traditional Intraclass Correlation Coefficient, Figure 1 is provided to show how they are related. The upper scale shows values of the gauge R&R ratio and the three categories into which it sorts all measurement systems. The lower scale shows the values for the Intraclass Correlation Statistic.

The left side of Figure 1 represents a measurement system with no measurement error—the values are determined by the measured item alone. Using our model, the left side represents the limiting condition  $X = Y$ . The right side of Figure 1 represents a random number generator which produces values that are 100 percent measurement error—pure noise containing no information about the measured item. In terms of our model, the right side represents the limiting condition  $X = E$ . Between these two extremes, the nonlinear relationship between the gauge R&R ratio and the Intraclass Correlation Coefficient is shown by the diagonal lines which connect corresponding values on these two scales.

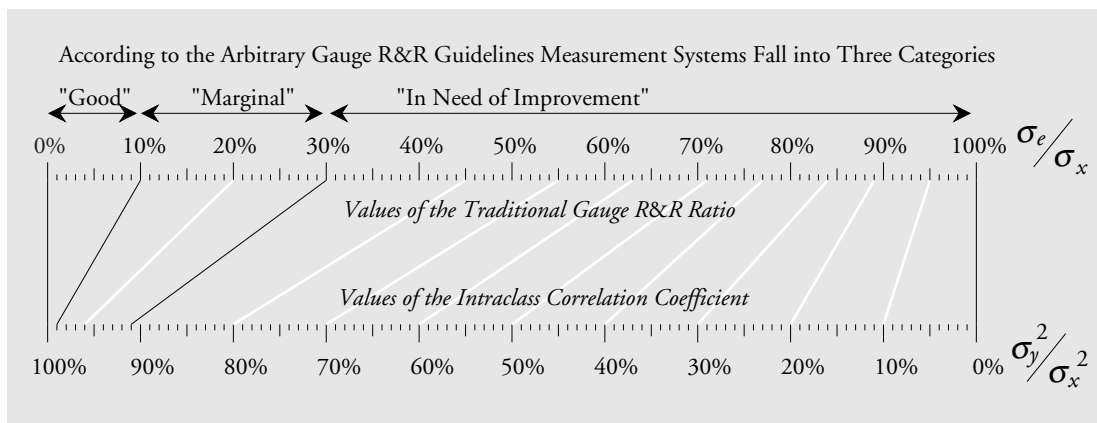


Figure 1: The Intraclass Correlation Coefficient and the Gauge R&R Ratio

In the gauge R&R nomenclature, “good” systems have a  $\sigma_e$  that “consumes” less than 10% of the total variation. (These will have an Intraclass Correlation better than 99%.) “Marginal” systems have a  $\sigma_e$  that “consumes” between 10% and 30% of the total variation. (These will have an Intraclass Correlations between 99% and 91%.) And “bad” systems are those for which  $\sigma_e$  that “consumes” more than 30% of the total variation. (Everything with an Intraclass Correlation below 91%!)

As we shall see, these gauge R&R guidelines for characterizing measurement systems as good, marginal, or bad are excessively conservative. They have no basis in fact , they have no basis in practice, and they have no relevance to the ability to use the data on a process behavior chart.

## A Simple Experiment

We can examine the effect of measurement error upon the performance of a process behavior chart by means of a simple experiment that will generate values to use in our structural model,  $X = Y + E$ . We begin by using a random number generator from a spreadsheet program to generate 100 observations from a normal distribution having a mean of 0.00 and a standard deviation of 1.00. One such set of values is shown in Figure 2.

We will consider these values to represent the actual values for some product characteristic (the  $Y$  values) and will use the first 50 values to compute the limits for our chart. For Figure 2 the first 50 values have an Average of  $-0.14$  and an Average Moving Range of  $1.17$  which yields a  $\text{Sigma}(Y)$  value of  $1.04$ . These values are reasonably close to the theoretical mean and standard deviation. Since we are going to use these data as a blank slate, it is preferable that they contain no false alarms or spurious signals, either in the 50 point baseline or in the remaining 50 values.

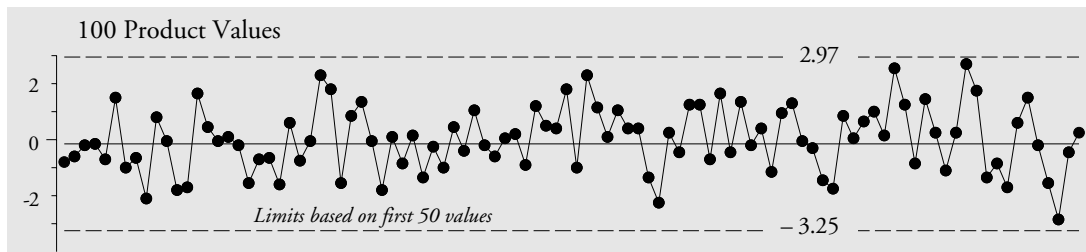


Figure 2: Individuals Chart for One Hundred Product Values  $Y$

The chart in Figure 2 is typical of what we would find with a predictable process and a perfect measuring system. With no measurement error we would have an Intraclass Correlation Statistic of 1.00, and the limits would represent nothing but the routine product variation.

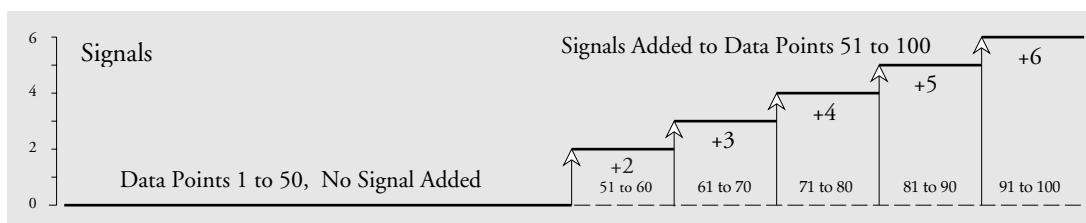


Figure 3: Signals Used

Next we are going to add the signals of Figure 3 to the product values in Figure 2. By leaving the first 50 values unchanged we obtain a baseline to use in obtaining limits. By adding 2.0 units to values 51 to 60 we simulate a two-sigma shift in the production process. Likewise, the other signals represent three-, four-, five-, and six-sigma shifts in the production process. Since experience has shown that three-sigma to six-sigma shifts in location are not uncommon in practice, the signals in Figure 3 provide a useful model for those process changes that we want to detect.

When we combine Figure 2 and Figure 3, and use the first 50 values as our baseline, we get the Individuals Chart in Figure 4. Each of the five different signals are clearly detected by the chart. But then we

should have expected this, since theory tells us that even a two-sigma shift has better than an 80 percent chance of being detected within ten points following the shift.

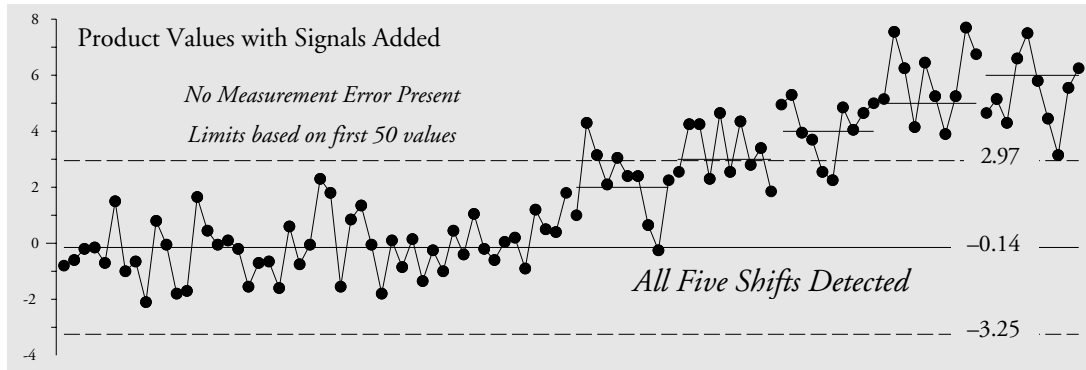


Figure 4: Individuals Chart for Product Values Plus Signals when Intraclass Correlation is 1.00

Figure 4 represents our product values,  $Y$ , after they have been subjected to a known set of process changes. We will now use the sequence of 100 values shown in Figure 4 to examine how different levels of measurement error,  $E$ , will affect our ability to detect these known shifts in the production process.

When our measurement system is less than perfect it will contribute some measurement error to the picture in Figure 4. This measurement error can only widen the limits and dilute the strength of the signals seen there. To represent the effects of measurement error we will use our random number generator to obtain another 100 values from a normal distribution. If we again use a normal distribution with a mean of 0.00 and standard deviation of 1.00, we will get values like those shown in Figure 5. The baseline for the limits shown is all 100 values. The average is  $-0.11$ , which is reasonably close to the theoretical mean of 0.00. The average moving range is 1.27 giving a  $Sigma(E)$  of 1.125, which is reasonably close to the theoretical standard deviation of 1.00. Once again, for the purposes of our simple experiment, it is best to have a set of measurement errors that display no false alarms.

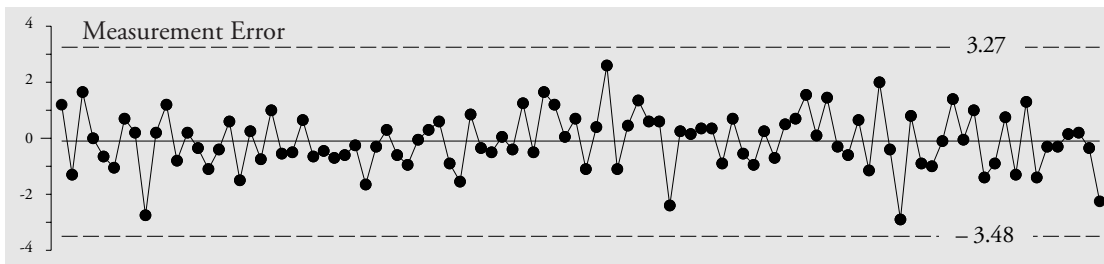


Figure 5: Individuals Chart for Measurement Error Set Number One

When we add the Product Values,  $Y$ , of Figure 2 to the Measurement Errors,  $E$ , of Figure 5 we get a set of Product Measurements,  $X$ . When we add the Signals of Figure 3 to this set of Product Measurements we get the data shown in Figure 6.

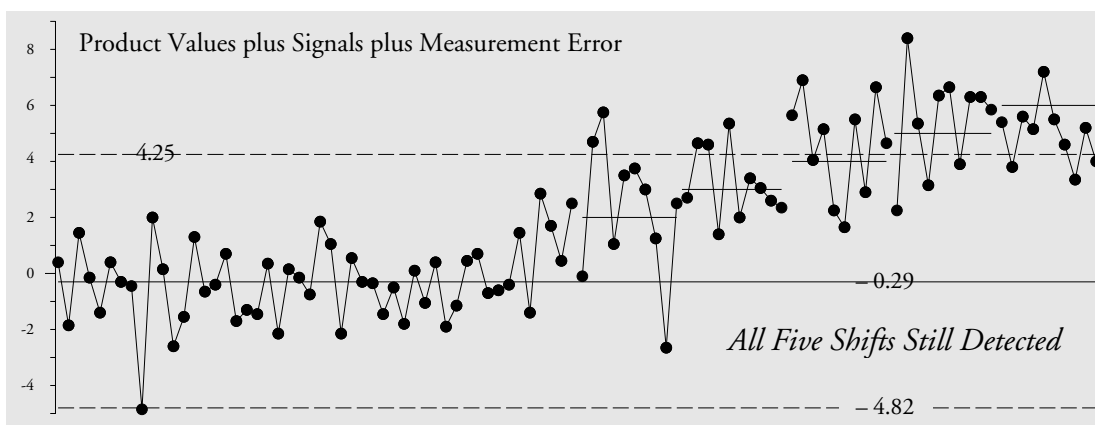


Figure 6: X-Chart for Product Measurements Plus Signals when Intraclass Correlation is 0.50

The limits shown in Figure 6 are based on the first 50 values. Since the first 50 values contain no added signals, they represent the Product Measurements alone. Based on the fact that both Figure 2 and Figure 5 were obtained from  $N(0,1)$  distributions, the Intraclass Correlation Coefficient for these Product Measurements is:

$$\rho_1 = \frac{\sigma_y^2}{\sigma_x^2} = \frac{1.00}{1.00 + 1.00} = 0.50$$

However, since the Measurement Errors had  $\text{Sigma}(E) = 1.125$ , while the Product Values had  $\text{Sigma}(Y) = 1.04$ , the Intraclass Correlation Statistic is:

$$r_1 = 1 - \frac{\hat{\sigma}_e^2}{\hat{\sigma}_x^2} = 1 - \frac{1.26}{2.35} = 0.46$$

Thus, in Figure 6 we have a fairly severe situation—measurement error,  $E$ , accounts for approximately 50 percent of the variation present in these Product Measurements,  $X$ . Only 46% of the variation present can be attributed to the Product Values,  $Y$ . (The gauge R&R ratio is up in the vicinity of 75%—a very bad number indeed.)

As expected, the limits in Figure 6 are considerably wider than those in Figure 4. These wider limits now encompass the four-sigma shift in the process. *Yet, in spite of these wider limits, the process behavior chart still detects each one of the five shifts in these data!*

From a theoretical perspective, within  $k = 10$  points following a shift, the two-sigma shift had a 45% chance of being detected by a point outside the limits, and a 95% chance of being detected using all four rules of the Western Electric Zone Tests.\* The three-sigma shift had an 87% chance of being detected by a point outside the limits, and a 100% chance of being detected using the Western Electric Zone Tests. And the four-sigma shift had a 98% chance of being detected by a point outside the limits.

Think about this. In spite of the fact that the noise of the measurement system dominates the product variation here, the chart for individual values is still able to detect two-sigma and larger shifts in the process. An Intraclass Correlation of 50% is not one to inspire confidence in the product measurements, and yet these data are still useful in detecting process changes.

\* Detection Rule One: A single point outside the three-sigma limits is a signal of a process change.  
 Detection Rule Two: Two out of three points beyond one of the two-sigma lines is a signal.  
 Detection Rule Three: Four out of five points beyond one of the one-sigma lines is a signal.  
 Detection Rule Four: Eight successive points on the same side of the central line is a signal.

When you think that the Gauge R&R guidelines will condemn a measurement system whenever its Intraclass Correlation falls below 91%, you can begin to see how completely meaningless those guidelines are in determining when you can use your measurements on a process behavior chart.

## How Much Measurement Error is Too Much?

To push our simple experiment to an extreme, consider doubling the magnitude of each of the values in Figure 5. Call this Measurement Error Set Number Two. (An alternative would be to generate another 100 observations from a normal distribution with a mean of 0.0 and a standard deviation of 2.0.) The doubled values shown in Figure 7 have an average of  $-0.21$  and a  $\text{Sigma}(E)$  value of 2.25.

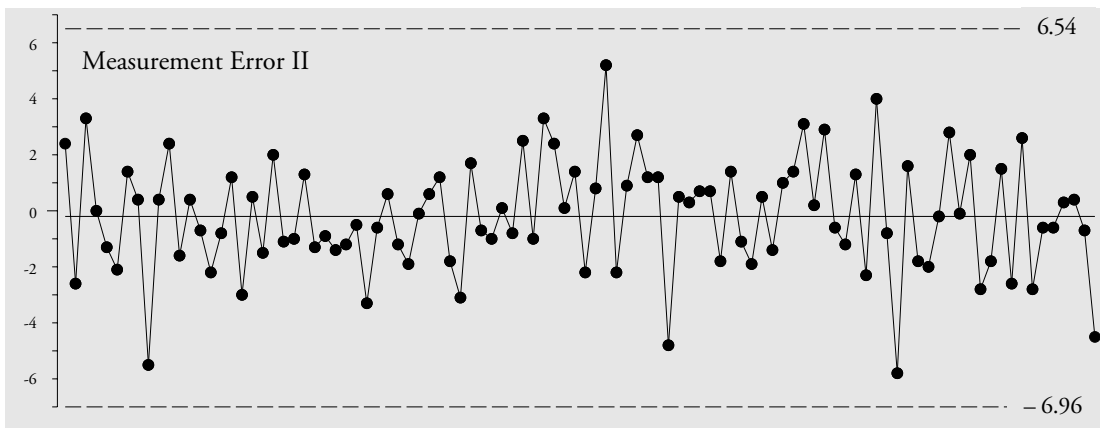


Figure 7: Measurement Error Set Number Two

When we add the Product Values,  $Y$ , of Figure 2 to the Measurement Errors,  $E$ , of Figure 7 we get our second set of Product Measurements,  $X$ . When we add the Signals of Figure 3 to this second set of Product Measurements we get the data shown in Figure 8.

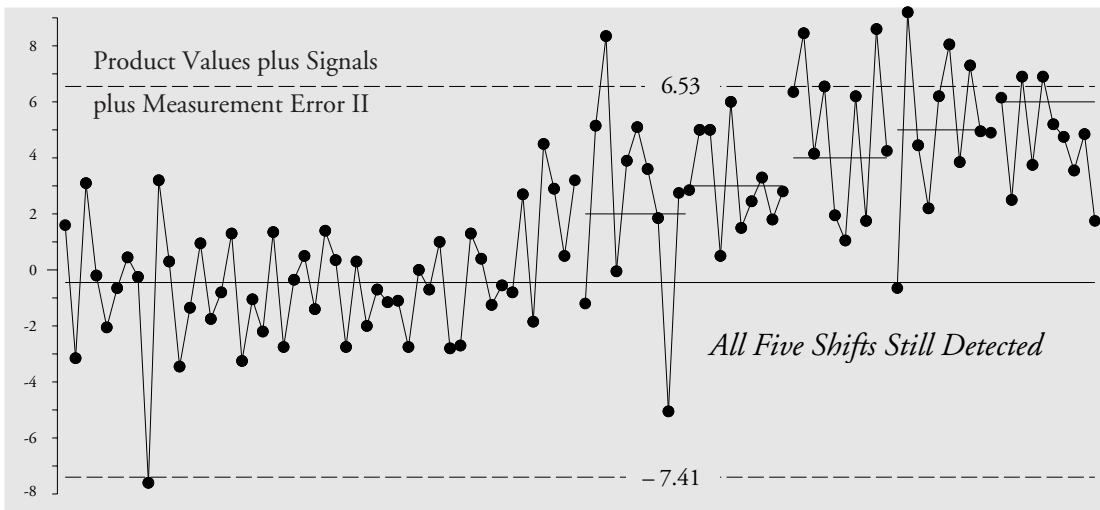


Figure 8:  $X$ -Chart for Product Measurements Plus Signals when Intraclass Correlation is 0.20

The limits shown in Figure 8 are based on the first 50 values. Since the first 50 values contain no added signals, they represent the second set of Product Measurements alone. Based on the fact that Figure 2 was obtained from a  $N(0,1)$  distribution and Figure 7 was effectively obtained from a  $N(0,4)$  distribution, the Intraclass Correlation Coefficient for this second set of Product Measurements is:

$$\rho_1 = \frac{\sigma_y^2}{\sigma_x^2} = \frac{1.00}{1.00 + 4.00} = 0.20$$

However, since the Measurement Errors in Figure 7 had  $\text{Sigma}(E) = 2.25$ , while the Product Values had  $\text{Sigma}(Y) = 1.04$ , the Intraclass Correlation Statistic is:

$$r_1 = 1 - \frac{\hat{\sigma}_e^2}{\hat{\sigma}_x^2} = 1 - \frac{5.06}{6.14} = 0.18$$

Therefore, Figure 8 shows what happens when over 80% of the variation is noise.

The limits in Figure 8 are over twice as wide as those in Figure 4. Yet either by runs or by points outside the limits, all five shifts are still detected within  $k = 10$  points of when they occurred.

From a theoretical perspective, within  $k = 10$  points following a shift, the two-sigma shift had a 16% chance of being detected by a point outside the limits, and a 60% chance of being detected using the Western Electric Zone Tests. The three-sigma shift had a 40% chance of being detected by a point outside the limits, and a 91% chance of being detected using the Western Electric Zone Tests. And the four-sigma shift had a 70% chance of being detected by a point outside the limits, and a 98% chance of being detected using the Western Electric Zone Tests.

While your own running of this experiment will turn out differently, I think that you will be surprised by the sensitivity of the process behavior chart in the face of overwhelming measurement error. As long as you are detecting signals your measurements are good enough. You have something to investigate, and you do not need to worry about the quality of your measurements. The next two sections will show why this is the case.

## So What Do the Gauge R&R Ratios Characterize?

In order to obtain theoretical support for the results illustrated in the preceding sections we will need to return to the ratio commonly used in gauge R&R studies—the ratio of  $\sigma_e$  to  $\sigma_x$ . Although gauge R&R studies interpret this ratio as the amount of the total variation that is consumed by measurement error, this interpretation is incorrect.

To see what the ratio of  $\sigma_e$  to  $\sigma_x$  actually represents we begin with our structural model:

$$X = Y + E$$

Assume that the *measurement system* experiences a shift in location, and that this shift is equal to  $3\sigma_e$  in magnitude. Then our measurements could be expressed as:

$$X = Y + E + 3\sigma_e$$

And if  $\sigma_e / \sigma_x$  is 0.10, which is taken as the borderline between good and marginal, this will become:

$$X = Y + E + 3\sigma_e = Y + E + 0.3\sigma_x$$

Thus, in this case, a three standard deviation shift in the measurement system will look like a 0.3 standard deviation shift to the *product measurements*. Here the signal of a change in the measurement system was

attenuated by 90%. When subtracted from 1.00, the ratio of  $\sigma_e$  to  $\sigma_x$  defines the amount by which signals of change in the *measurement system* are attenuated before they show up in the *product measurements*.

A similar argument will show that, when subtracted from 1.00, the ratio of  $\sigma_y$  to  $\sigma_x$  defines the amount by which signals of change in the *production process* are attenuated before they show up in the *product measurements*. When these two signal attenuation curves are plotted against the Intraclass Correlation Coefficient we get Figure 9. For reference, Table 1 contains some of the values plotted in Figure 9.

Table 1: The Signal Attenuation Curves

Intraclass Correlation Coefficient	Production Process Signal Strengths	Measurement System Signal Strengths	Intraclass Correlation Coefficient	Production Process Signal Strengths	Measurement System Signal Strengths
0.99	.9950	.1000	0.40	.6325	.7746
0.91	.9539	.3000	0.30	.5477	.8367
0.80	.8944	.4472	0.20	.4472	.8944
0.70	.8367	.5477	0.09	.3000	.9539
0.60	.7746	.6325	0.01	.1000	.9950
0.50	.7071	.7071			

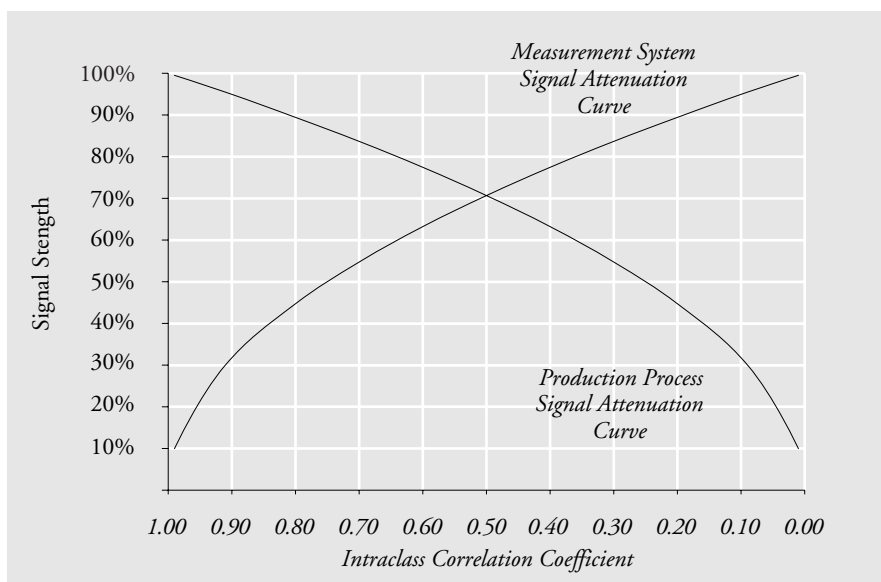


Figure 9: The Signal Attenuation Curves

Clearly, as the Intraclass Correlation goes from 1.00 to 0.50, the strength of the signals from the production process does not decline as rapidly as the strength of the signals from the measurement system increases.

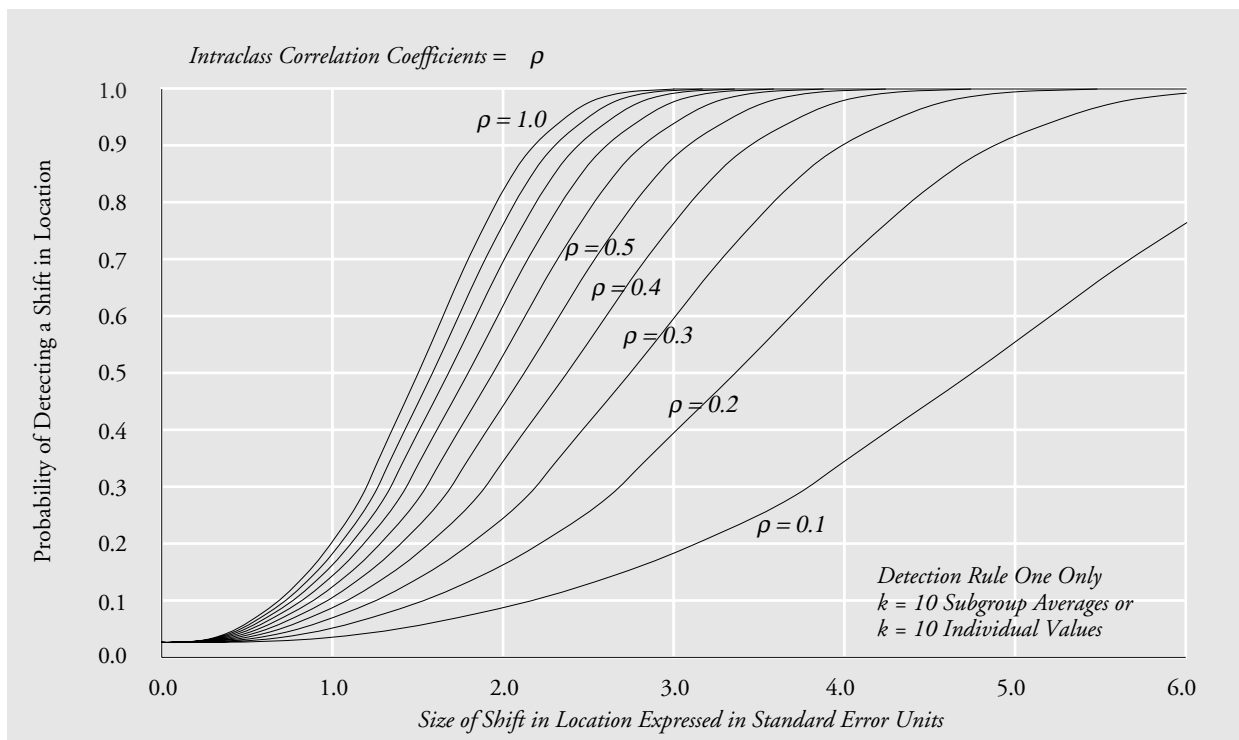
When the Intraclass Correlation is 91% a 1.0 unit change in location in the production process will show up in the product measurements as a 0.954 unit change. On the other hand, a 1.0 unit change in location in the measurement system will show up as a 0.3 unit change in the product measurements. Yet this is the situation where the gauge R&R ratio is on the border between marginal and bad!

By using these signal attenuation curves with published tables of the power function for Average Charts we can determine the probabilities of detecting a shift in location as a function of the Intraclass Correlation. This will be done in the next section.

## The Probability of Detecting a Shift

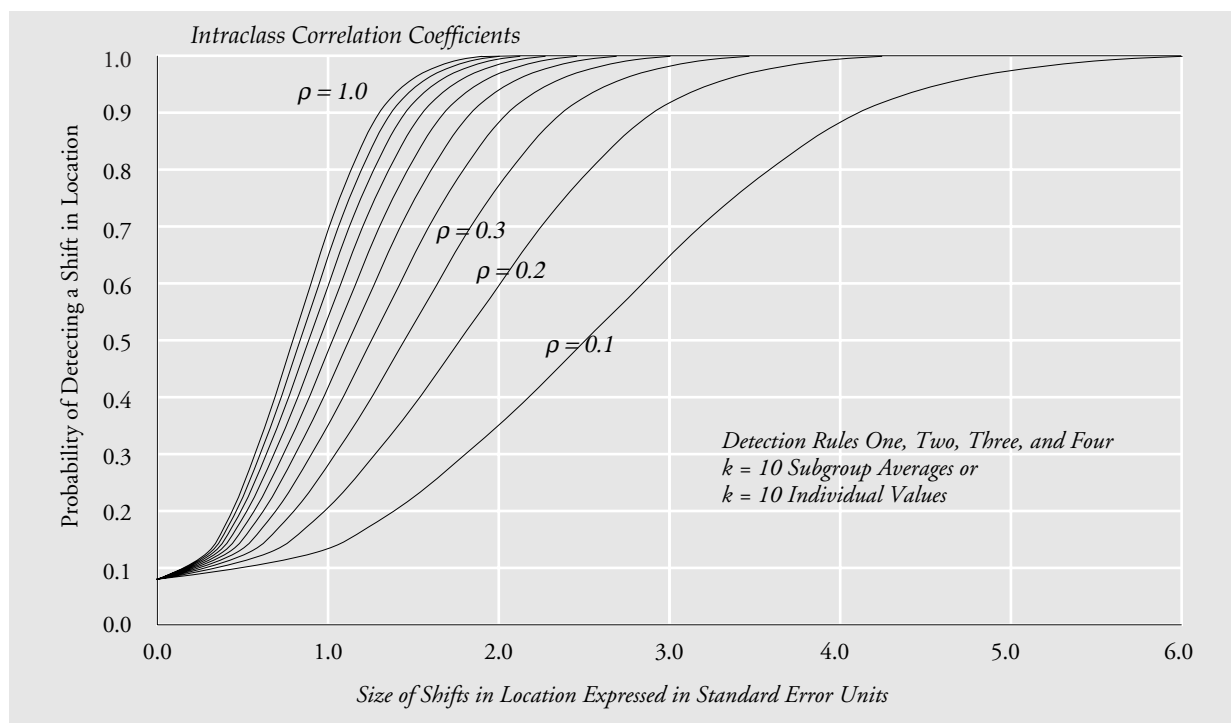
Since empirical studies like those in the earlier sections do not always turn out the same way, we resort to theory as a guide for practice. Figure 10 shows the Power Function for charts for location based on Detection Rule One alone. The curves shown represent the probability of detecting a shift within 10 points following that shift. The vertical scale shows the probabilities and the horizontal scale shows the size of the shifts. When used with a Chart for Individual Values the horizontal scale represents the size of the shift in location as a multiple of  $\sigma_x$ . When used with an Average Chart the horizontal scale represents the size of the shift in location as a multiple of  $\sigma_x/\sqrt{n}$ .

The curve on the left for  $\rho = 1.0$  is based on published tables of the power function [Wheeler, 1983]. The curves for other values of the Intraclass Correlation Coefficient were obtained from the first curve by using the signal strengths of Figure 9. For example, when the Intraclass Correlation Coefficient is 0.80, a production process signal is attenuated down to a signal strength of 0.8944. Therefore, a 2.00 standard error shift when  $\rho = 0.80$  would have the same probability of detection (0.70) as a 1.79 standard error shift when  $\rho = 1.00$ .



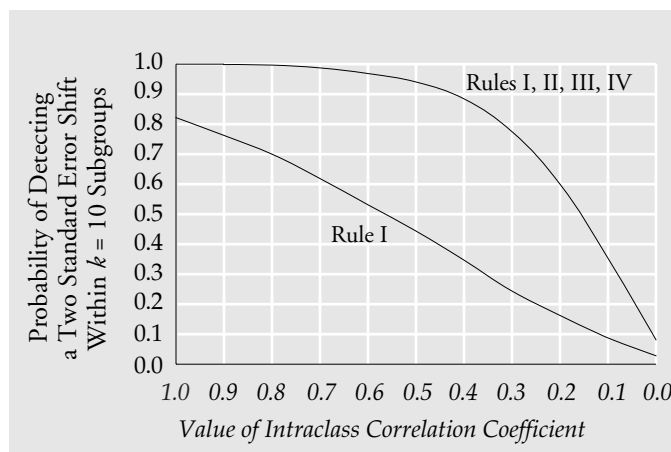
**Figure 10: The Effect of the Intraclass Correlation Coefficient Upon Power Functions for Charts for Location for Detection Rule One alone when  $k = 10$**

Figure 11 shows the Power Function for charts for location when using all four detection rules in the Western Electric Zone Tests. As before, the left curve for  $\rho = 1.0$  comes from published tables of the power function and the remaining curves were found by taking into account the signal attenuations.



**Figure 11: Power Functions for Charts for Location for Detection Rules One, Two, Three, and Four, with  $k = 10$ , and for Different Values of the Intra-class Correlation Coefficient**

By looking at a fixed size of shift on the horizontal axis in Figures 10 and 11, and by plotting the probabilities as a function of the Intra-class Correlation Coefficient, we can obtain figures which show how the probability of detecting a shift of a given size changes with the Intra-class Correlation Coefficient.



**Figure 12: Probabilities of Detecting Two-Sigma Shifts**

Figure 12 shows the probabilities of detecting a two standard error shift in location within ten subgroups on an Average Chart or within ten values on a Chart for Individual Values. As expected, the probabilities of detection drop off as the Intra-class Correlation moves from 1.0 down to 0.0. Also, as ex-

pected, increasing measurement error will dramatically diminish the probability of a two-sigma shift being detected by a point outside the limits. But with the use of the run-tests (Detection Rules Two, Three, and Four) we can keep the probability of detecting a two-sigma shift above 90% all the way down to an Intraclass Correlation of 0.43.

Figure 13 shows the probabilities of detecting a three standard error shift in location within ten subgroups. Contrary to our expectations, but consistent with the results of the simple investigations reported earlier, these probabilities of detecting a shift do not drop off very rapidly. A three-sigma shift has at least a 90 percent chance of being detected within 10 subgroups as long as the Intraclass Correlation exceeds 54%. The Intraclass Correlation has to drop to less than 25% before we have a 50-50 chance of detecting a three-sigma shift with Detection Rule One. With the run-tests we have better than a 90% chance of detecting a three-sigma shift all the way down to an Intraclass Correlation of 0.20.

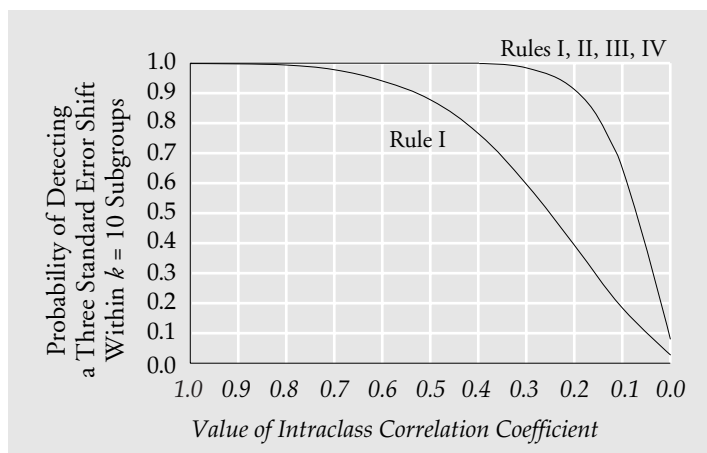


Figure 13: Probabilities of Detecting Three-Sigma Shifts

Figure 14 shows the probabilities of detecting a four standard error shift in location within ten subgroups. Here we have better than a 90% chance of detection with Rule One alone all the way down to an Intraclass Correlation of 0.30.

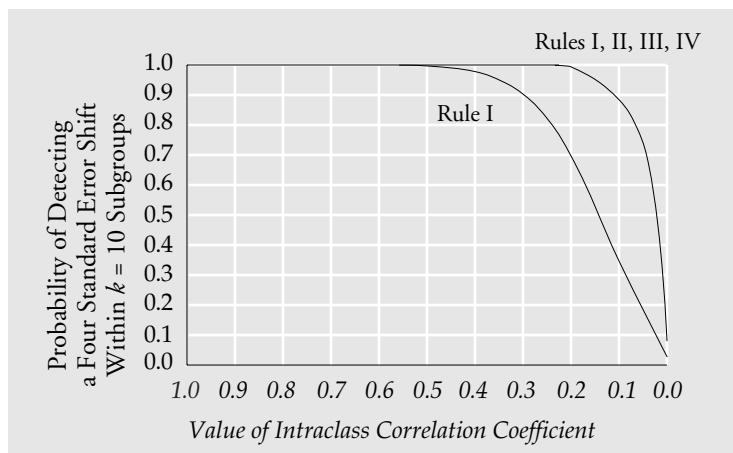


Figure 14: Probabilities of Detecting Four-Sigma Shifts

Figure 15 shows the probabilities of detecting a five standard error shift in location within ten subgroups. Here we have better than a 90% chance of detection with Rule One alone all the way down to an Intraclass Correlation of 0.20.

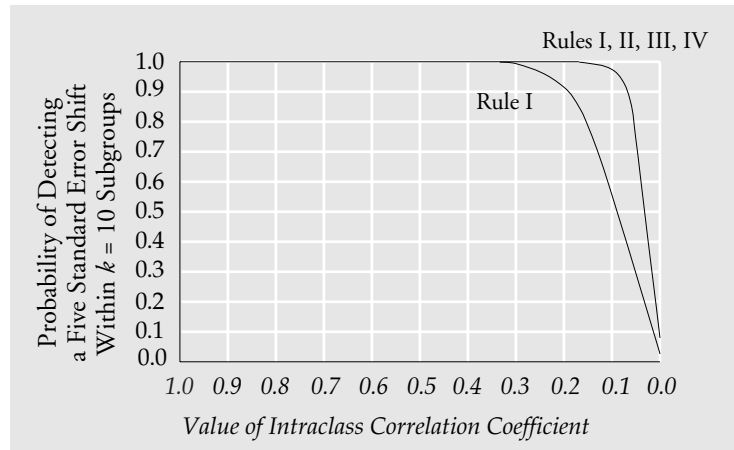


Figure 15: Probabilities of Detecting Five-Sigma Shifts

Six-sigma shifts in location have better than a 90 percent chance of being detected within 10 subgroups with Rule One alone all the way down to an Intraclass Correlation of 0.15.

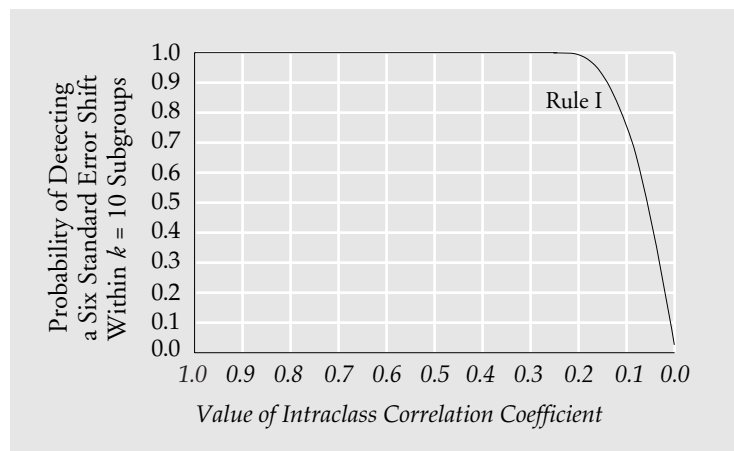


Figure 16: Probabilities of Detecting Six-Sigma Shifts

Thus, even very weak and soft data can be successfully used on process behavior charts to detect those changes in the production process that are large enough to have an economic impact. You do not have to wait until you have “good” measurements. You do not have to upgrade your measurement process whenever it has a gauge R&R ratio above 30%. As long as you are detecting signals you have opportunities to improve your process.

## Four Categories of Data for Process Monitoring

Recall that Shewhart designed process behavior charts to detect changes that were large enough to have an economic impact upon the process. In this paper we have seen both theoretical and empirical justifications of the fact that we can detect these changes even when the measurement system contributes more variation to the product measurements than does the production process itself.

So what does this mean about the effects of measurement error upon the performance of a process behavior chart? Just how good do the data have to be before we can use them on a process behavior chart? Since there is no sharp cut-off, but rather a gradual degradation, it is helpful to have different categories to use in describing how to use measurements on process behavior charts.

### First Class Monitors

Figure 17 reproduces Figure 13 and a portion of Figure 9. There we see that whenever the Intraclass Correlation Coefficient exceeds 80% we will have better than a 99% chance of detecting a Three Standard Error Shift within ten subgroups of the time that it occurred using Detection Rule One alone. So while we may use all four detection rules, we will rarely need more than Detection Rule One when our measurement system has an Intraclass Correlation Coefficient that exceeds 80%.

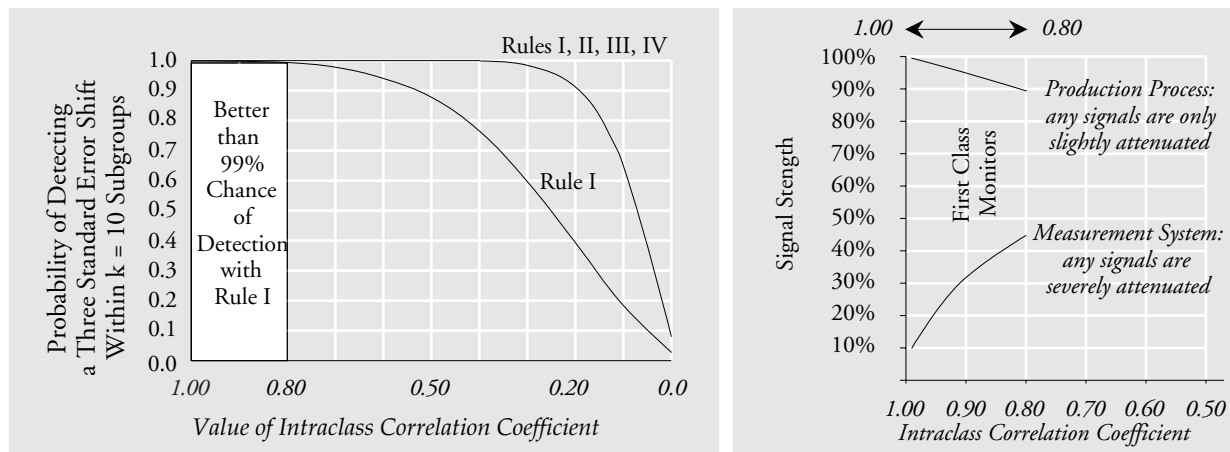


Figure 17: Characteristics of First Class Monitors

We also see that whenever the Intraclass Correlation Coefficient exceeds 80% any signals from the production process are attenuated by less than 11%, while any signals of changes in the measurement system are attenuated by more than 55%. (Attenuation is the complement of signal strength.)

Therefore, whenever the Intraclass Correlation Statistic exceeds 80% the measurements can be considered to provide First Class Monitors of the production process and any signals on the process behavior chart should be interpreted as coming from the production process.

- Better than 99% chance of detecting a three std. error shift using Detection Rule One alone.
- May also use Detection Rules Two, Three, and Four.
- Signals from Production Process are only slightly attenuated (0% to 11%).
- Signals from Measurement System are severely attenuated (90% to 55%).

## Second Class Monitors

Figure 18 reproduces Figure 13 and a portion of Figure 9. There we see that whenever the Intraclass Correlation Coefficient is between 80% and 50% we will have better than an 88% chance of detecting a Three Standard Error Shift within ten subgroups of the time that it occurred using Detection Rule One alone. Using all four detection rules will boost this up to a virtual certainty. So even when measurement error is a substantial portion of the overall variation, we still have very high probabilities of detecting changes when they occur.

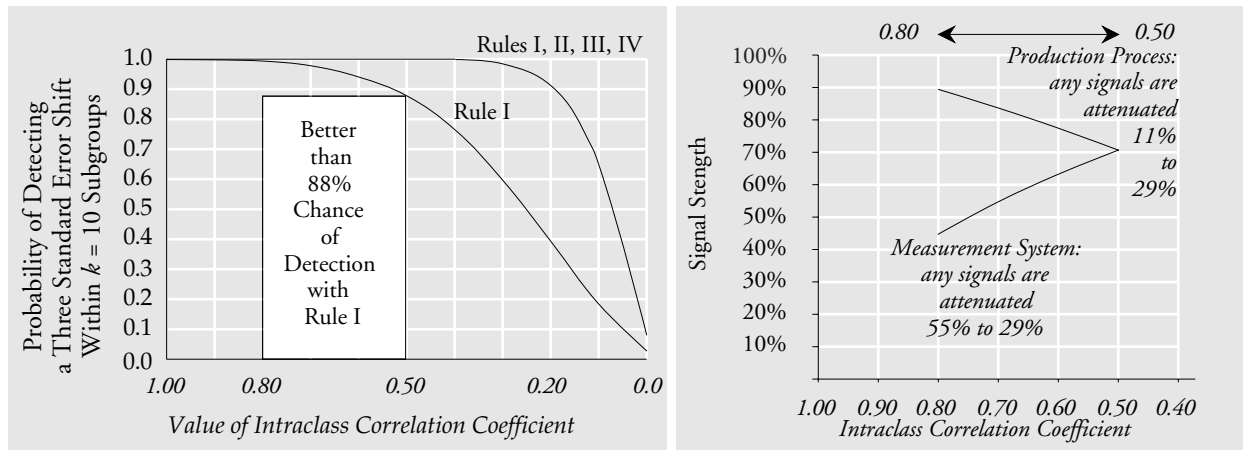


Figure 18: Characteristics of Second Class Monitors

Whenever the Intraclass Correlation Coefficient is between 80% and 50% any signals from the production process will be attenuated from 11% to 29% while any signals that come from the measurement system will be attenuated by 55% to 29%. This means that the process behavior chart for the production process will still be more sensitive to changes in the production process than to changes in the measurement system.

Therefore, whenever the Intraclass Correlation Statistic is between 80% and 50% the measurements can be considered to provide Second Class Monitors of the production process. Second Class Monitors still provide a very high likelihood of detecting changes in your production process. Moreover, since production process signals are less attenuated than signals coming from the measurement system, it is more likely that any signals on the process behavior chart come from the production process. However, with Second Class Monitors we cannot rule out the possibility of signals from the measurement system showing up on the process behavior chart. If your measurement system is reasonably predictable, then the process behavior chart for the production process may be all that you need to use. But if you have doubts about the predictability of your measurement system you may wish to maintain a process behavior chart for the measurement system in addition to the process behavior chart for the production process.

- Better than 88% chance of detecting a three std. error shift using Detection Rule One alone.
- Virtual certainty of detecting this size shift using Detection Rules One, Two, Three, and Four.
- Signals from Production Process are attenuated (11% to 29%).
- Signals from Measurement System are more attenuated (55% to 29%).
- May choose to track measurement system on process behavior chart.

### Third Class Monitors

Figure 19 reproduces Figure 13 and a portion of Figure 9. There we see that whenever the Intraclass Correlation Coefficient is between 50% and 20% we will have better than a 91% chance of detecting a Three Standard Error Shift within ten subgroups of the time that it occurred using all four detection rules in the Western Electric Zone Tests. So even when measurement error is the dominant portion of the overall variation, we still have high probabilities of detecting changes when they occur.

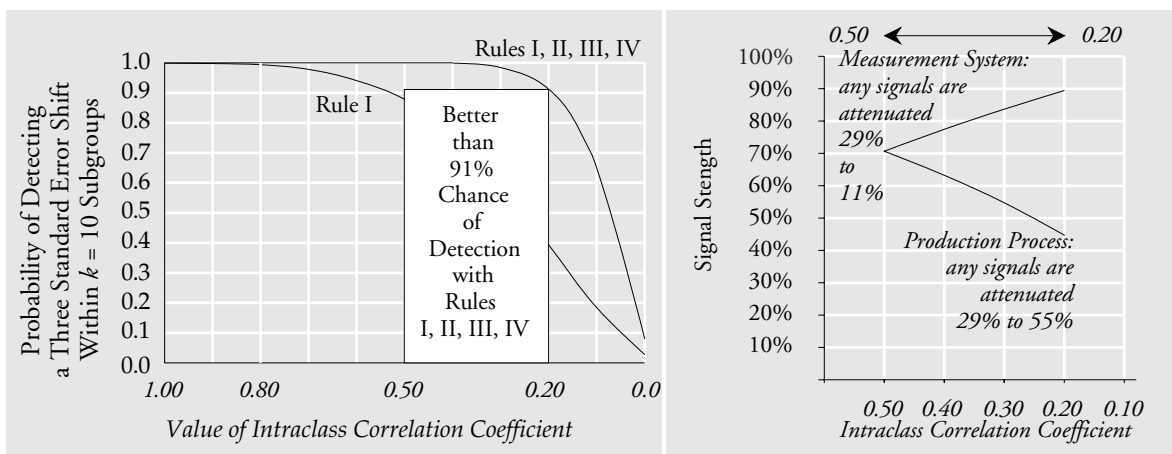


Figure 19: Characteristics of Third Class Monitors

Whenever the Intraclass Correlation Coefficient is between 50% and 20% any signals from the production process will be attenuated from 29% to 55% while any signals that come from the measurement system will be attenuated by 29% to 11%. This means that the process behavior chart maintained for the production process will be more sensitive to a shift in the measurement system than to changes in the production process.

Therefore, whenever the Intraclass Correlation Statistic is between 50% and 20% the measurements can be considered to provide Third Class Monitors of the production process. Third Class Monitors can still detect changes in your production process, but only when you can rule out the possibility that a change has occurred in your measurement system. This means that in order to use Third Class Monitors to track your production process you will have to maintain, on a concurrent basis, a process behavior chart that monitors your measurement system. Signals on both charts are interpreted as a change in the measurement system. Signals on the production process chart alone are interpreted as coming from the production process.

- Better than 39% chance of detecting a three std. error shift using Detection Rule One alone.
- Better than 91% chance of detecting this size shift using Detection Rules I, II, III, and IV.
- Signals from Production Process are attenuated (29% to 55%).
- Signals from Measurement System are less attenuated (29% to 11%).
- Must also track measurement system on process behavior chart.

### Fourth Class Monitors

Figure 20 reproduces Figure 13 and a portion of Figure 9. There we see that whenever the Intraclass Correlation Coefficient is below 20% the chances of detecting a three standard error shift rapidly vanish. While very large changes in the production process may be detected with these measurements, the attenuation is so great that you are likely to be made aware of the process changes by other means. Here the use of a process behavior chart can only be an act of desperation. Obviously, such a chart would have to be accompanied by a parallel chart that tracked the measurement system at the same time.

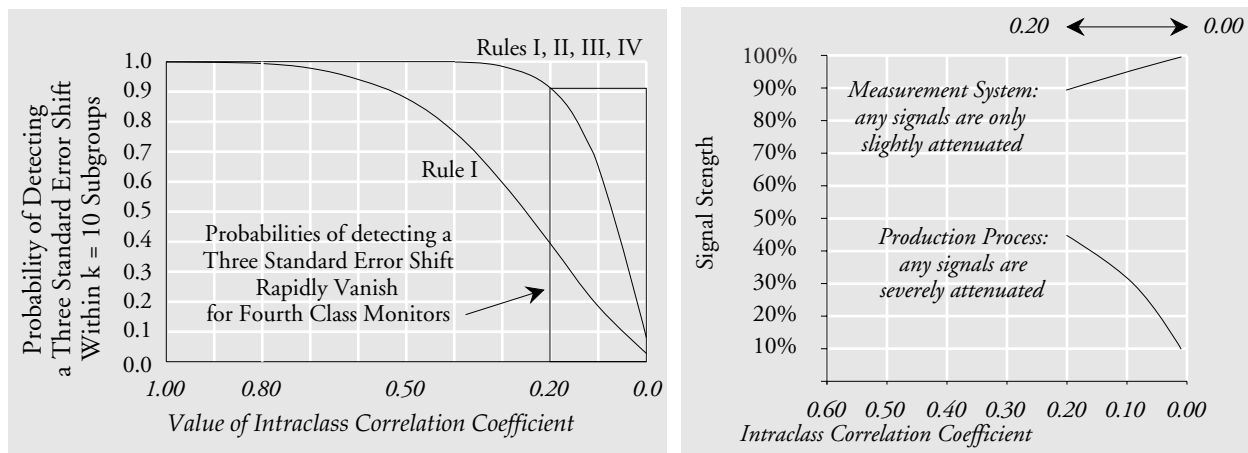


Figure 20: Characteristics of Fourth Class Monitors

Therefore, whenever the Intraclass Correlation Statistic is below 20% the measurements can be considered to provide Fourth Class Monitors of the production process. Such data should only be used when no other measures are available. Fourth Class Monitors contain very little useful information about the product itself.

### Summary

We have seen that it is the ratio of the Variance of  $Y$  to the Variance of  $X$  that describes the proportion of the routine variation in the product measurements that can actually be attributed to the production process. In a similar vein, it is the ratio of the Variance of  $E$  to the Variance of  $X$  that describes the proportion of the routine variation in the product measurements that must be attributed to the measurement system. Thus, in characterizing routine variation, it is the Intraclass Correlation, and its complement, that are the proper measures to use.

On the other hand, we have seen that when the ratio of  $\sigma_e$  to  $\sigma_x$  is, say, 10% this does not mean that measurement error consumes 10% of the total variation, but rather that any signals of unpredictable behavior that might originate in the measurement system will be attenuated by 90% before they show up on a process behavior chart. Gauge R&R studies have completely misinterpreted this ratio for over 40 years.

We have obtained results that completely contradict the guidelines of virtually all gauge R&R studies. While these guidelines have been carefully handed down from one generation of engineers to another for over 40 years, in all this time there has not been any rigorous examination of the basis for these gauge R&R guidelines. In fact, over the years, the same set of guidelines has been applied to different ratios, in spite of the fact that these different ratios were, in no sense, equivalent. In short, the gauge R&R guide-

lines are merely a superstitious tradition.

The results obtained here are theoretically sound, rigorously developed, and empirically verifiable. They reflect the reality that many practitioners have observed. And they do not needlessly condemn the measurement system. Instead these results provide guidelines on how to get useful information even when measurement error is the dominant component of your measurements. If you still feel compelled to use the arbitrary gauge R&R scale, you will need to use cut-offs that are substantially different from the traditional ones. But the better alternative is to use the Intraclass Correlation Coefficient which correctly compares the different sources of variation.

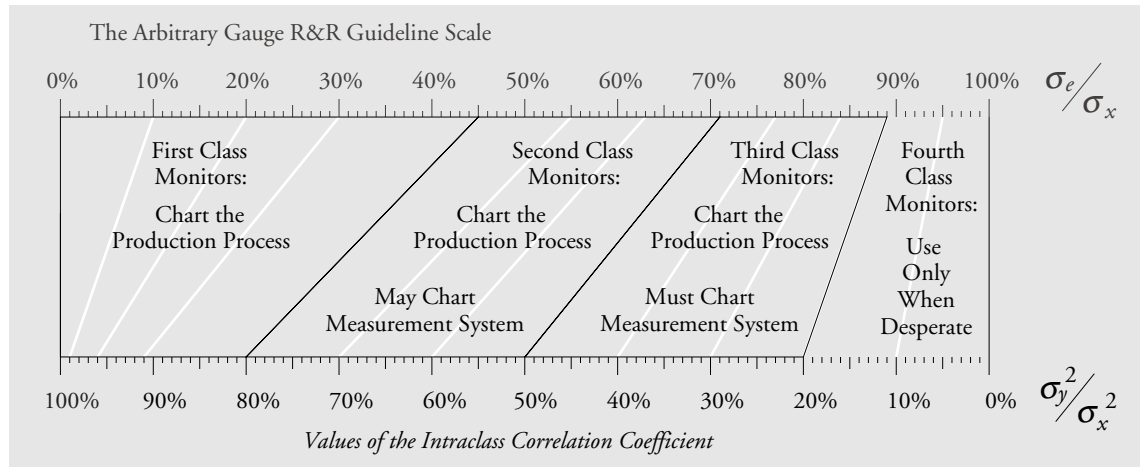


Figure 21: The Truth About Measurement Error and Process Behavior Charts

There is no difference in the way that you construct, interpret, or maintain a process behavior chart for First or Second Class Monitors. Even with Third Class Monitors, the construction, maintenance, and interpretation of the chart for the production process will be done in the same way as for the first two classes. In most instances, when you have Third or Fourth Class Monitors you will know that your measurements are soft even without having to do a measurement error study. Therefore, it is not necessary to carry out a measurement system study prior to placing your data on a process behavior chart. Just place the product measurements on a process behavior chart. If you find signals, assume that you probably have First or Second Class Monitors, and look for an explanation in the production process. If you do not detect any signals, then consider if that might be because you are using Third or Fourth Class Monitors, and do your measurement system study. This is why measurement error studies have never really been a prerequisite for placing data on a process behavior chart.

### References

- Fisher, Sir Ronald, "On the Probable Error of a Coefficient of Correlation Deduced from a Small Sample," *Metron*, v.1, No.4, pp. 3-32, 1921.
- Wheeler, Donald J., "Detecting a Shift in Process Average, Tables of the Power Function for X-bar Charts," *Journal of Quality Technology*, v.15, No.4, pp.155-170, 1983.

Copyright © 2003 SPC Press

Contact Dr. Wheeler at [www.spcpress.com](http://www.spcpress.com).