100% Inspection and Measurement Error

How to handle the problem of misclassified product

Donald J. Wheeler

One hundred percent inspection is commonly used to avoid shipping nonconforming product. Each of the items produced is measured and judged to be either conforming or nonconforming. The conforming items get shipped and the nonconforming items get set aside for use as factory authorized replacement parts or some other equally ignominious fate. It all seems quite simple until you do the math. The purpose of this article is to show what happens when a product stream is subjected to 100% screening inspection with less than perfect measurement systems.

To simplify the mathematics I will assume that the production process is being operated predictably, and that the stream of product values can be represented by our generic, worst case probability model, the normal distribution. In our model we shall denote these product values by Y. In Figure 1 we see the distribution of product values and the specifications for this product characteristic. This distribution represents the product stream, and when we measure an item, it is the value for Y that we are trying to determine.

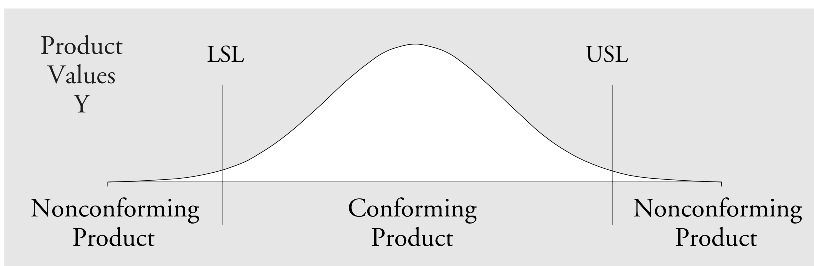

Figure 1: The Distribution of the Product Values Y

When we measure an item having the value Y we get an observed value, X. This observed value will differ from Y, and this difference may be denoted as E, the amount by which our measurement has erred. Since the traditional model for measurement error is also a normal distribution, we end up with a normal distribution for the observed values, X, as shown in Figure 2. It is customary to assume that the distribution for X has the same mean as Y, but that the variance for X will be equal to the variance for Y plus the variance for the measurement errors, E.

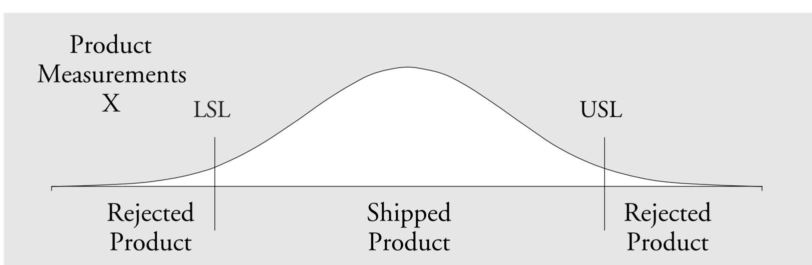

Figure 2: The Distribution of the Observed Values X

These two probability models are often shown when discussing the problem of inspection, but these two one-dimensional probability models do not tell the whole story. In particular, they do not show how the values for X are related to the values for Y, and this relationship can only be shown by using a bivariate probability model. Here we shall use a bivariate normal distribution. The means and variances for X and Y will be the same as the means and variances shown in Figures 1 and 2. In addition to the means and variances of the two marginal distributions, the bivariate normal distribution requires a correlation coefficient to be fully specified. In this case, the correlation between X and Y will be the square root of the Intraclass Correlation Coefficient discussed in my December 2010 column. This bivariate normal probability model may be shown in the $X\,Y$ coordinate plane by a series of ellipses where each ellipse is a contour of equal height for the joint probability density function. The eccentricity of the ellipses is determined by the correlation between X and Y. Figure 3 shows the one-, two-, and three-standard deviation contours for two bivariate normal distributions where the Intraclass Correlation Coefficients are 0.95 and 0.80 respectively.

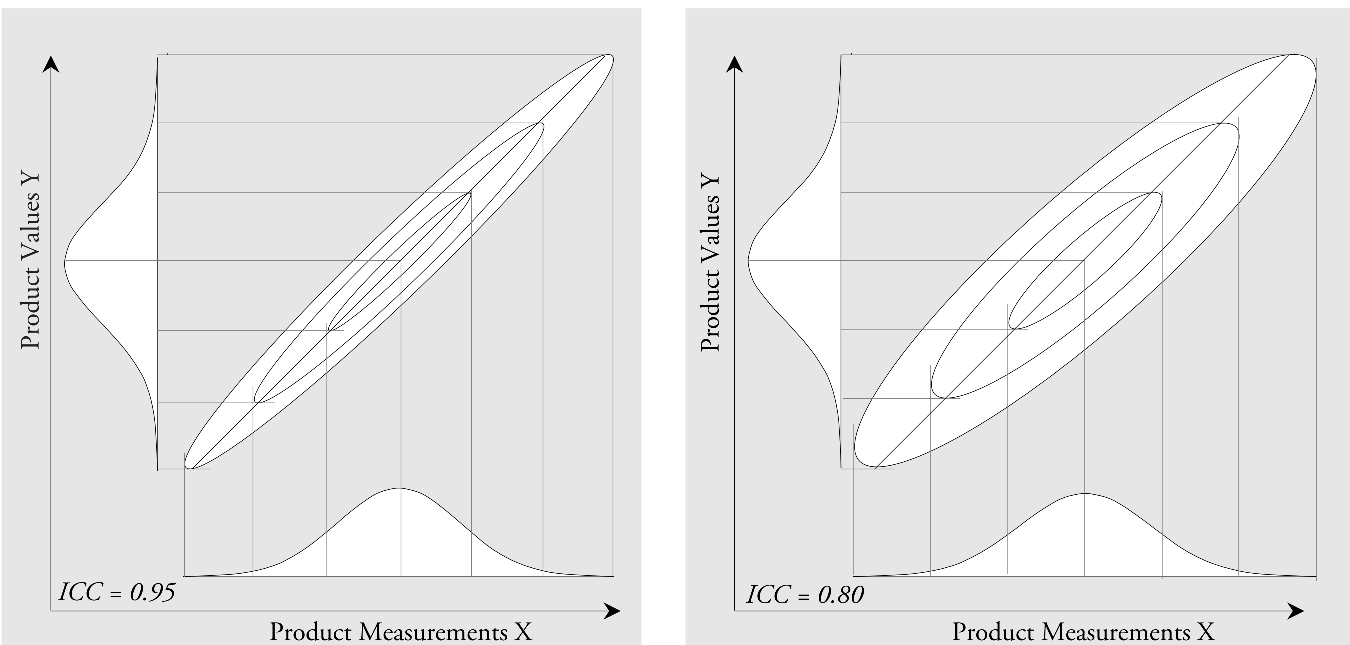

Figure 3: Contours of Two Bivariate Normal Distributions

From Figure 3 we see that as the measurement error increases the ellipses for the bivariate normal distribution get fatter, the major axis skews around, and the dispersion for the marginal distribution of X increases. While we could return to Figure 1 to compute the proportion of conforming product in the product stream, we need to use graphs like those in Figure 3 to understand how the set of product measurements X vary with the product values Y. Specifically, we are interested in those values for the product measurements X that correspond to the conforming items, Y. This is shown in Figure 4 where the region in white corresponds to those (X,Y) coordinates that coincide with the conforming items.

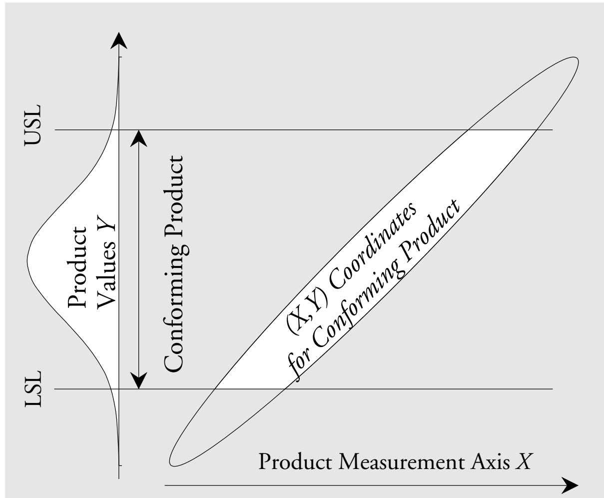

Figure 4: The (*X,Y*) Coordinates for Conforming Product

In contrast to Figure 4, Figure 5 shows the (*X,Y*) coordinates that correspond to the product that gets shipped following the 100% inspection.

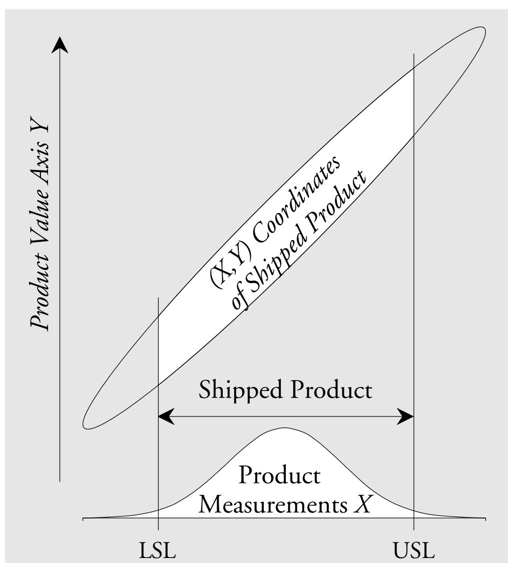

Figure 5: The (*X,Y*) Coordinates for Product That Gets Shipped

Thus, the region of conforming product is not the same as the region of product that gets shipped. The difference between these two regions is the essence of the problem of inspection. Using "good" as shorthand for conforming and "bad" as shorthand for nonconforming, we end up with four distinct outcomes:

- Good product that gets shipped (*GS*);
- good product that gets rejected (*GR*);
- bad product that gets shipped (*BS*); and

- bad product that gets rejected (*BR*).

Of course, good product that gets shipped (*GS*) and bad product that gets rejected (*BR*) are the appropriate outcomes for the inspection procedure. The good product that is rejected (*GR*) is a problem for the producer and the bad product that gets shipped (*BS*) is a problem for the customer. Figure 6 shows four regions of misclassification and three regions of correct classifications. There are two regions where good product will be rejected (GR1 and GR2), and two regions where bad product will be shipped (BS1 and BS2). In addition there are two regions where bad product will be rejected (BR1 and BR2), and the one central region where good product is shipped (GS).

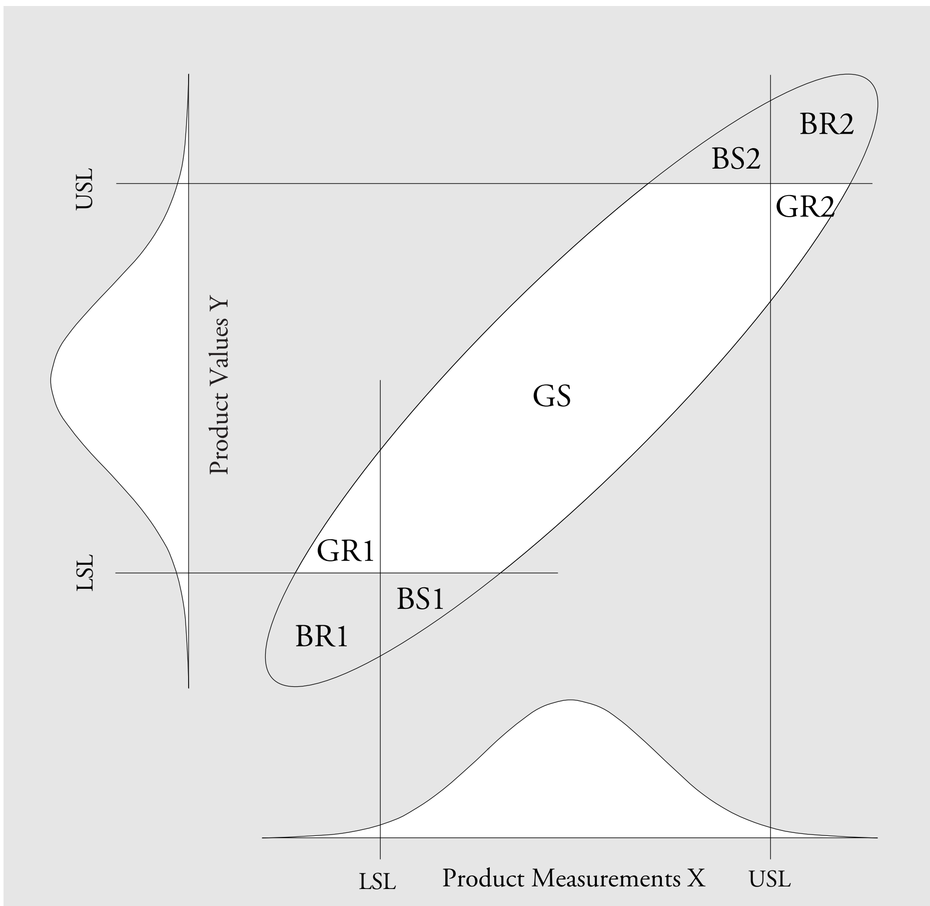

Figure 6: The Four Outcomes of 100% Inspection

The traditional way of characterizing an inspection operation like the one shown in Figure 6 is to look at two proportions. The first of these is the proportion of good product that gets shipped:

$$\textit{Proportion of Good Product that Is Shipped} = PGS = \frac{GS}{GS + GR1 + GR2}$$

The producer will want *PGS* to be as large as possible since the complement of *PGS* will be the proportion of good product that is rejected.

The other proportion is the proportion of bad product that gets rejected:

$$\textit{Proportion of Bad Product that Is Rejected} = PBR = \frac{BR1 + BR2}{BR1 + BR2 + BS1 + BS2}$$

The customer will want *PBR* to be as large as possible since the complement of *PBR* will be the proportion of bad product that is shipped.

In order to study the behavior of these two proportions we will need to determine two things. The first will be the overall fraction nonconforming in the product stream. This will depend upon the relationship between the distribution for *Y* and the specification limits. Assuming that the production process is centered within the specifications, as shown in Figure 6, we will consider six different levels of nonconforming product. These six levels are shown in Table 1. The second thing that affects the two proportions given above is the width of the ellipse in Figure 6. As we saw in Figure 3, this width will be inversely proportional to the value of the Intraclass Correlation Coefficient. As the *ICC* drops the ellipse will get wider, and both *PGS* and *PBR* will tend to get smaller.

Table 1 begins with a virtually perfect measurement system having an *ICC* of 0.995 and then considers some smaller *ICC* values to show the effects of increasing amounts of measurement error in the product measurements, *X*.

Table 1: Proportion Good Product Shipped (*PGS*) and Proportion Bad Product Rejected (*PBR*) Based on 100% Inspection

		Percentage of Nonconforming Product in Product Stream					
ICC		*50%*	*40%*	*30%*	*20%*	*10%*	*1%*
0.995	*PGS*	.963	.973	.980	.987	.993	.999
	PBR	.965	.962	.958	.954	.946	.925
0.99	*PGS*	.947	.961	.972	.981	.990	.999
	PBR	.951	.947	.942	.935	.926	.900
0.95	*PGS*	.874	.905	.930	.952	.973	.996
	PBR	.896	.888	.879	.865	.850	.814
0.90	*PGS*	.811	.857	.893	.925	.958	.993
	PBR	.857	.847	.836	.821	.801	.757
0.85	*PGS*	.759	.815	.861	.901	.942	.990
	PBR	.828	.817	.806	.789	.770	.724
0.80	*PGS*	.711	.776	.829	.877	.926	.986
	PBR	.804	.793	.781	.764	.745	.700

The first pair of entries in Table 1 tell us that when the product stream consists of 50% bad stuff and our measurements are very good we will manage to ship about 96 percent of the good stuff and to also filter out about 96 percent of the bad stuff. As we move to the right we see how these two proportions change as the fraction nonconforming declines. When we have only one percent bad stuff in the product stream, our very good measurements will ship virtually all of the good stuff and will reject about 92 percent of the bad stuff.

As we move down Table 1 we see the effects of increasing amounts of measurement error. When considered as a whole, Table 1 reveals two inherent characteristics of 100% inspection. As the incoming quality level gets better, the inspection operation will do a better job of shipping the good stuff, but it will also do a poorer job of finding the bad stuff. We can see both of these characteristics by looking at how the *PGS* values get larger in each row while the *PBR* values get smaller. In addition, both sets of proportions tend to get smaller as the measurement error increases.

To illustrate this point, consider inspecting 10,000 items with a measurement system that has an *ICC* value of 0.99. Let us assume that there are 100 nonconforming items in these 10,000 items (one percent). Table 1 tells us that 100% inspection will ship 0.999 of the 9,900 good items. This works out to be 9,890, and we would expect about 10 good items to be rejected. At the same time, Table 1 shows that 0.900 of the 100 bad items will be rejected, leaving about 10 bad items that will be shipped. Thus, this inspection plan found 99.9% of the good items, but only 90% of the bad items. The better the initial quality level, the lower the *PBR* value will become. This means that the strategy of 100% inspection will always have a diminishing return, and at some point its expense will outweigh its benefits.

AVOIDING MISCLASSIFICATIONS

Since it is the misclassifications that are the fly in the ointment of 100% inspection, it is logical to consider how to avoid them.

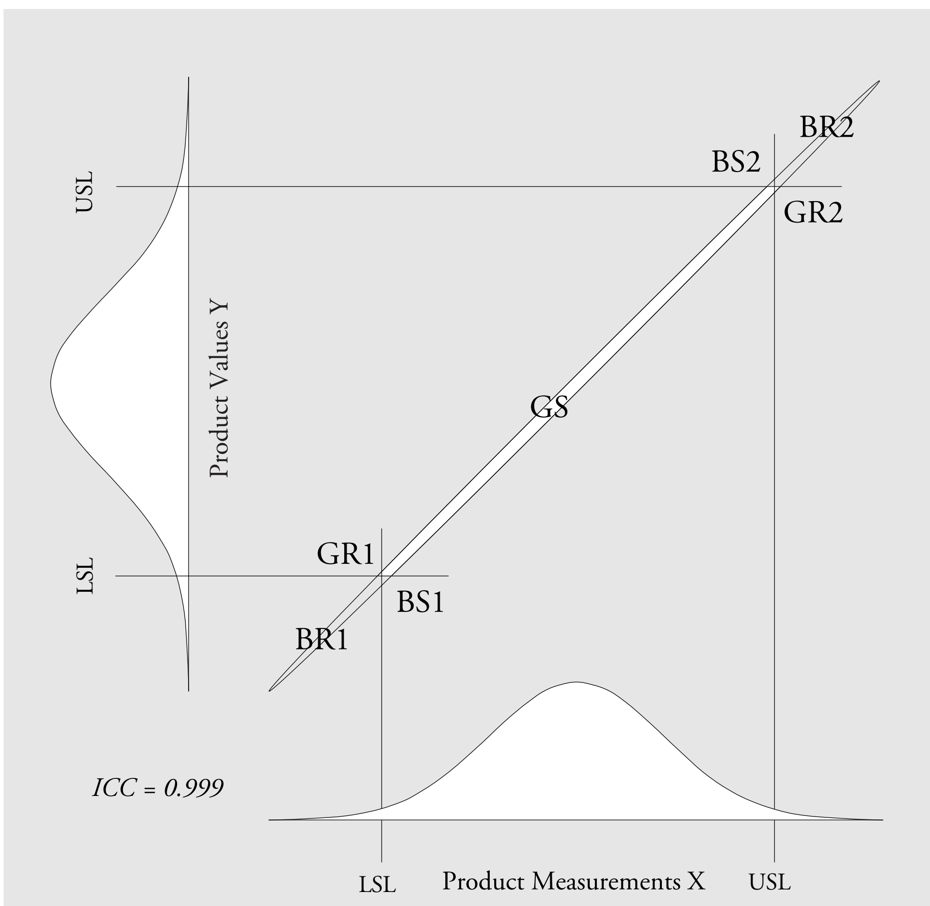

Figure 7: A Virtually Perfect Measurement System

Both Figure 7 and Table 1 suggest that one way to avoid the problems of misclassification is to have a perfect measurement system. When the *ICC* reaches 1.00 the measurement error will disappear, the ellipse will become a straight line, X will always equal Y, and both *PGS* and *PBR* will become 1.000.

Let me be clear on this point. It is perfection that is needed here. Figure 7 shows that as long as there is any measurement error there will be some misclassifications. Returning to Table 1 we see from the first two rows that when measurement error amounts to only one part in 200, you will suffer up to 3.5 misclassifications per hundred good parts inspected and from 3.5 to 7.5 misclassifications per hundred bad parts inspected.

Let measurement error climb to one part per hundred and you will suffer up to 5 misclassifications per hundred good parts inspected and from 5 to 10 misclassifications per hundred bad parts inspected.

And should measurement error amount to one part per 20, you will suffer up to 12 misclassifications per hundred good parts inspected and from 10 to 19 misclassifications per hundred bad parts inspected.

So, if you are wedded to using inspection at the end of the line to separate the good stuff from the bad stuff, then you will need to be committed to using only perfect measurement systems. Of course, when virtually perfect measurement systems exist, they will tend to be both expensive to own and expensive to operate. Since money spent to purchase and operate measurement systems is 100% overhead, these costs directly reduce productivity and profitability. Thus, 100% inspection will increase the costs of production while it does a poorer job of removing bad product as the quality level of the production process increases. This is why 100% inspection provides a less than sterling solution to the problem of how to ship good product.

A seminar with a company that made brake parts began with the students asking, “How can we be sure that we are not shipping any nonconforming parts?” They asked this on Monday. They asked this on Tuesday. They asked this on Wednesday. So I finally got the message that until I answered this question I would not be allowed to leave town. In light of what we have just seen, and given the fact that virtually all measurement systems will be less than perfect, there is only one way left to be sure that you are not shipping some nonconforming product, and that is to avoid making any bad stuff in the first place.

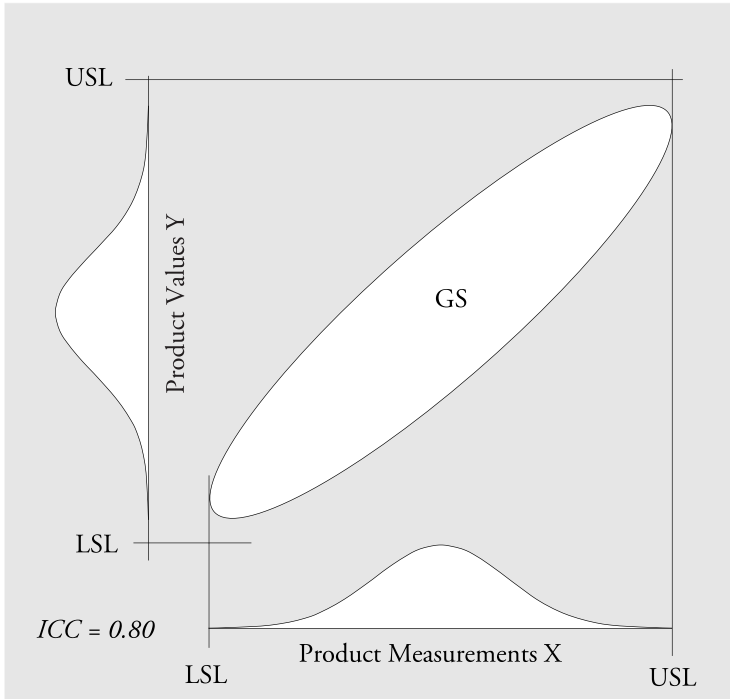

Figure 8: The Only Sure Way to Avoid Shipping Nonconforming Product

Figure 8 shows a second way to avoid the problems of misclassification. In order for the three-sigma ellipse to fit entirely within the *GS* region we will need to operate our process predictably, on-target, and with a capability ratio in excess of 1.00. Here we will find 100% inspection to be unnecessary and there will be no problems of misclassification. This will not only save the cost of the perfect measurement system, but it will also eliminate the costs of inspection and thereby increase productivity. But how can you do this? A proven way of reducing both process variation and production costs is through the effective use of process behavior charts for continual improvement. Moreover, as we saw last month, this can be done using measurements that contain substantial amounts of measurement error. For an explanation of how this works see my column "Two Routes to Process Improvement, Parts I and II" of May 5 and 6, 2010.

USING IMPERFECT MEASUREMENTS PART ONE

So, what happens when we have less than perfect measurements? Obviously games are played. One of my clients had a process where only 50% of the product met the blueprint specifications. However, due to the pressure of production schedules, the customer had agreed to accept product that met "deviated specifications" that were 133% larger than the blueprint specifications. This deviation allowed about 85% of the product to be shipped, but the producer still had the inspection costs associated with the deviated specifications. As the producer learned how to operate this process predictably and on-target they got to the point where they were shipping better than 99.5% based on the deviated specifications. At that point the customer confessed that their actual specifications were 150% of the blueprint specs and agreed to change the blueprints accordingly.

This is fairly typical. Customers determine what they want and then divide it in half, and then the producers complain until they get a deviation. In other words, specifications are more often the result of negotiations than of a careful consideration of what will work in practice. In this environment you will use whatever measurement system you have with the deviated specifications, and both the customer and the producer will suffer their respective costs of misclassification as the price of doing business.

In the presence of measurement error, 100% inspection has to trade off one type of misclassification against the other. Table 1 showed what happens when the inspection operation uses the customer specifications. In this section and the next we will look at two alternatives. The first of these is widened specifications.

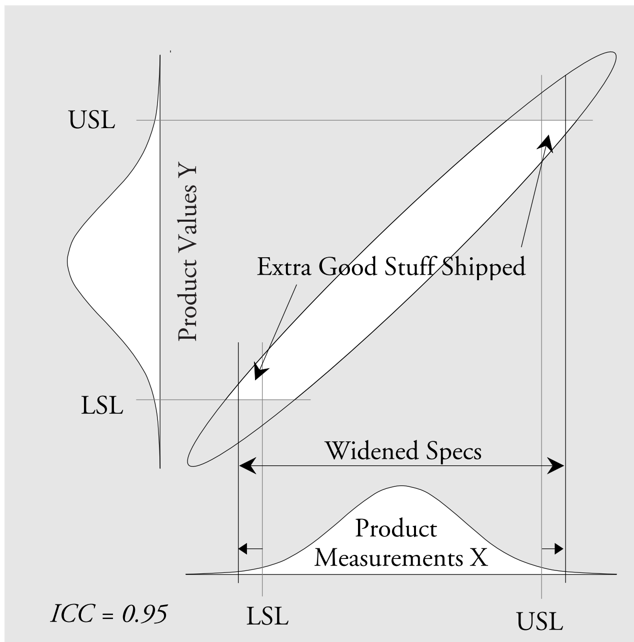

Figure 9: Specifications Widened by Two Probable Errors

Since the producer is usually the one performing the inspection, let us see how the producer might reduce the amount of good product that gets rejected. Since the Probable Error defines the effective resolution of a measurement system, we will consider what would happen if the producer widens the specifications on the X axis by two Probable Errors. (The Probable Error was explained in my April Column as 0.675 times the standard deviation of measurement error.) When the producer widens the specifications by two Probable Errors a higher percentage of the good stuff will get shipped, as shown in Figure 9. Table 2 lists the *PGS* and *PBR* values for 100% inspection using widened limits.

Table 2: Proportion Good Product Shipped (*PGS*) and Proportion Bad Product Rejected (*PBR*) Based on 100% Inspection Using Specifications Widened by 2 Probable Errors

		Percentage of Nonconforming Product in Product Stream					
ICC		*50%*	*40%*	*30%*	*20%*	*10%*	*1%*
0.995	*PGS*	.996	.997	.998	.999	.999	1.000
	PBR	.881	.870	.858	.841	.818	.759
0.99	*PGS*	.995	.996	.997	.998	.999	1.000
	PBR	.835	.821	.806	.784	.756	.681
0.95	*PGS*	.987	.990	.993	.995	.998	1.000
	PBR	.662	.640	.615	.585	.546	.451
0.90	*PGS*	.981	.986	.990	.993	.996	.999
	PBR	.553	.529	.503	.474	.434	.352
0.85	*PGS*	.976	.982	.987	.991	.995	.999
	PBR	.480	.457	.432	.405	.369	.300
0.80	*PGS*	.971	.978	.984	.989	.993	.999
	PBR	.425	.403	.380	.356	.323	.265

This widening of the specifications on the *X* axis has the effect of making virtually all of the *PGS* values reasonably close to 1.00. This means that almost all of the good stuff will get shipped. Unfortunately, the *PBR* values go from bad to worse. Since *PBR* tracks the proportion of bad stuff that gets rejected, this approach will end up shipping a lot of the bad stuff along with the good stuff. In fact, if shipping all of the good stuff is the goal, then why not skip the charade of 100% inspection and simply ship everything? The customer will scarcely know the difference between the shipments inspected by the plans above and the uninspected shipments.

USING IMPERFECT MEASUREMENTS PART TWO

However, in the situation where the specifications might actually be tied to some underlying reality, and where the producer has to *certify* that the product shipped does, indeed, meet the specifications, then it is essential to know how to minimize the *BS* misclassification. This was the topic of my columns, "Is the part in spec?," and "Where do manufacturing specifications come from?" in June and July of 2010. There the recommended approach was to *tighten* the specifications on the *X* axis by two Probable Errors. The results of this approach are summarized in Figure 10 and Table 3.

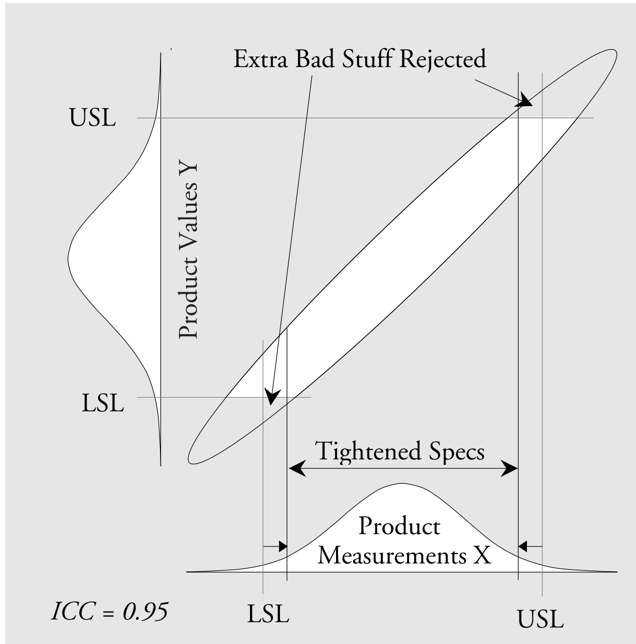

Figure 10: Specifications Tightened by Two Probable Errors

Table 3: Proportion Good Product Shipped (*PGS*) and
Proportion Bad Product Rejected (*PBR*)
Based on 100% Inspection Using Specifications Tightened by 2 Probable Errors

		Percentage of Nonconforming Product in Product Stream					
ICC		***50%***	***40%***	***30%***	***20%***	***10%***	***1%***
0.995	*PGS*	.868	.903	.930	.953	.975	.997
	PBR	.997	.996	.996	.995	.994	.992
0.99	*PGS*	.810	.860	.897	.930	.962	.995
	PBR	.995	.994	.994	.993	.992	.989
0.95	*PGS*	.544	.652	.739	.817	.895	.982
	PBR	.989	.988	.987	.985	.984	.979
0.90	*PGS*	.321	.473	.595	.707	.823	.966
	PBR	.985	.983	.982	.980	.978	.971
0.85	*PGS*	.143	.321	.470	.606	.752	.946
	PBR	.987	.980	.978	.976	.973	.967
0.80	*PGS*	—	.185	.351	.509	.679	.921
	PBR	—	.981	.975	.972	.970	.963

As we might have expected these tightened limits result in high proportions of bad product being rejected (*PBR*). While there are small differences in the *PBR* values across each row and as we move down Table 3, all of the inspection plans shown will remove almost all of the nonconforming product. Unfortunately, except in the last column and in the first row, these plans do not ship high proportions of the good product (*PGS*). In the last column we see reasonable values for both *PGS* and *PBR*, suggesting that this approach will work reasonably well when the overall quality level is good. So, while these plans make sense when the fraction nonconforming is small, and while these plans allow the producer to certify that the shipments will consist almost entirely of conforming product, they tend to punish the producer for not learning how to operate his process predictably and on-target with a capability of 0.80 or greater.

If we are going to take inspection seriously, we have to seek to avoid shipping bad product, even at the expense of rejecting good product. Manufacturing specifications allow you to do this while choosing your own degree of certainty. Any other approach either ignores measurement error or else treats specifications as a game to be played where the rules are subject to negotiation.

SUMMARY

A plant manager had a problem. Several of his measurement devices had been classified as "not acceptable" according to the AIAG guidelines. When he looked at replacing these devices he found that the cost would be about 3 million dollars. Before he spent this kind of money he wanted a second opinion, so he called Terry, the company statistician. Terry observed that the measurement systems involved were being used for prediction and representation, rather than for 100% inspection, and found that they all had Intraclass Correlations between 0.85 and 0.90. He told the plant manager that the classification of these measurement devices as "not acceptable" had been a false alarm. These devices were all capable of tracking the production process and quantifying process improvements. Needless to say, the manager chose to follow Terry's recommendation and avoid the unnecessary 3 million dollar expense.

As we saw above, if you are using a less than perfect measurement system for 100% inspection, you can use tightened manufacturing specifications to certify the conformity of the product you ship (when such certification is necessary). However, since this will usually be less than satisfactory from the producer's perspective, you will need to either work to improve the measurement system, or work to improve the production process.

Upgrading a measurement system will increase overhead expenses. It will not make the product any better. It will not eliminate the need for inspection. It will simply lower your productivity by increasing the costs associated with finding the burnt toast.

Improving the production process will reduce both the excess costs of production and the excess costs of use. Moreover, as has been proven time and again, these process improvements can usually be accomplished without capital expenditures and without having to improve the imperfect measurement system. It is always better to learn how to quit burning the toast than to be a world-class toast scraper!

This is why an excessive interest in the quality of the measurements is counterproductive. Imperfect measurement systems may be used to substantially improve the quality and consistency of a production process. These improvements can eliminate the need for 100% inspection while they reduce other excess costs for both the producer and the customer.

However, as we have seen, imperfect measurements will not do a good job of sorting product. If we focus on improving the measurement system so that we can sort the product to meet the specifications, our effort, our time, and our resources will be spent on increasing our overhead rather than on reducing our costs.