Title of Publication *Edited by*
TMS (The Minerals, Metals & Materials Society), Year

IMPROVING CARBON PLANT OPERATIONS THROUGH THE BETTER USE OF DATA.

Keith A. Sinclair[1], Barry A. Sadler[2]
[1]Sinclair Associates, Inc., 12620 Ridgepath Lane, Knoxville, TN 37922 USA.
keith@sinclairassociates.com
[2] Net Carbon Consulting Pty Ltd, P.O. Box 286, Kangaroo Ground, Victoria 3097 Australia.
barry.sadler@bigpond.com.au

Keywords: data traps, specifications, process control, sampling, averages, analysis, improvement

Abstract

In most Carbon Plants an enormous amount of data is collected and used for short and long term control, improvement of plant operations, reporting, and to monitor anode quality for the customer. In the experience of the Authors, however, the way that data are used can be improved considerably - full value is rarely extracted from the information and several “data traps” are common:

- ❑ Control decisions and actions are based only on comparisons to arbitrary specifications
- ❑ Averages are used inappropriately, e.g. with a lack of attention to variability in the data and data is over-aggregated
- ❑ Insufficient attention is given to the impact of sampling methods on the usefulness of data

This paper will demonstrate these data traps, describe their implications, and suggest ways of avoiding these pitfalls using plant examples including information from the sampling and testing of anode cores.

Introduction

A quick review of the weekly or monthly reports generated in a typical Carbon Plant will show that a very large quantity of numbers are collected, utilized in some way, and reported as a part of the running of these facilities. All too often these reports consist of vast tables of numbers (often averages of some kind) with an occasional graph thrown in to emphasize what is considered to be a critical parameter. Only rarely is it evident that full value is being extracted from the data collated in these reports; unfortunately it is much more common to find errors in the way the data is used and interpreted. Of these “data traps”, the Authors have encountered the following most frequently:

- Decisions on whether to take control actions (or not) are based solely on comparisons to “arbitrary” specification limits; these limits having little or no technical validity, and are commonly justified with “that is what we have always done”. They rarely have a clear connection with the needs of the Customer (i.e. the next stage in the process) but are used as the “magical” trigger point at which action or explanation is suddenly required.
- The over-aggregation of data into averages and use of these averages without appropriate attention to the variation within the data are rampant in Carbon Plants (And the other parts of Aluminium Smelters and many other Process Industries). Averages used in this way obscure variation in the process, thereby hiding legitimate signals that action may be required.
- Inferences regarding process behavior or product quality are made without sufficient attention given to the sampling method and the resulting usefulness of the data. Sampling processes must be designed based on the purpose of collecting the data. It is a common trap to use data specifically sampled for one purpose to make other, unrelated decisions.

In previous publications the Authors have looked at the need for “Statistical Thinking” in Carbon Plant operations [1] and presented the potential opportunities to improve processes by integrating Lean Manufacturing and Six-Sigma methodologies [2]. What follows in this paper is an expansion on the implications of the data traps that have been outlined above, and suggestions on how to avoid the waste associated with these traps using Carbon Plant data as examples.

Control by conformance to arbitrary specifications

This approach to process management is based on the view that as long as a parameter is within set limits (specifications), no action is required. This good value (inside the limits) versus bad value (outside the limits) form of “binary” control model results in action only when the parameter moves outside the limits. There are a number of problems/issues with the binary control model:

I. Specification limits are often set in a very arbitrary way and do not have a valid basis stemming from the impact on the Customer or the process. They are commonly derived from:
   a. History – “always been that way”
   b. Experience – often outdated, and rarely challenged
   c. “Experts” applying general industry standards (also often outdated) without considering the needs of specific Customers or the impact of local conditions.

   With such a model, all results within the limits are all good, are assumed to not cause any loss to the process, and therefore, require no action. (See Figure 1). As will be seen later, this is rarely, if ever, the case in reality. An implication of “within the limits means zero loss” is that a parameter can continually be measured to be just inside a specification limit and this will be considered acceptable with no action triggered. A small drift in the parameter could then move it outside of the limit. Suddenly the result is unacceptable as loss is now deemed to be occurring and action is called for – at the very least an explanation will be required. The reality is that the drift that resulted in the parameter moving beyond the limit could have only a negligibly small impact on the process, and does not warrant special attention.

II. Process variation can be "Common" cause - random variation which is part of the process and does not normally have an assignable cause; hence it should not trigger a response, and "Special" cause – which is not expected as a part of the process and normally will have an identifiable cause that should trigger a response [1]. The binary control model can result in two primary forms of loss to the business. First, it can cause over control or "tweaking" associated with reacting to common cause variation as though it were a special cause. This increases variation in the process and product. Second, it can result in under control, or the lack of action on a special cause change in the process (still within specification limits), leading to problems not being identified and actioned until the damage associated with the process disturbance escalates.

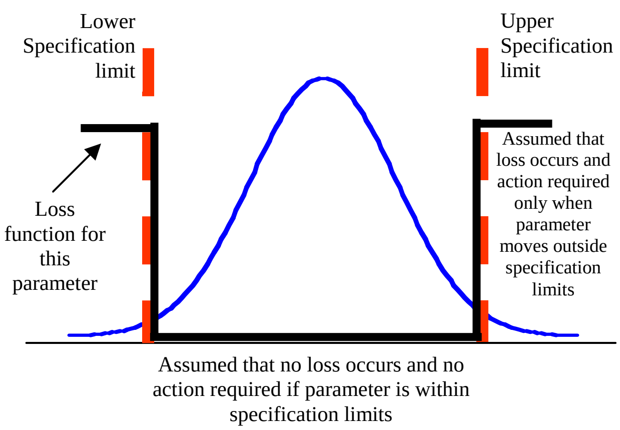

Figure 1. "Binary" control model assumes that as long as a parameter remains within arbitrary specifications no loss is incurred by the process. (See loss function curve) resulting in increased process variation and cost.

The issues associated with binary control can be explored further by using anode core Air Permeability as an example.

<u>"Binary" control of anode core Air Permeability.</u>

Binary control of Air Permeability (AP) assumes that variation in AP has no impact on anode performance in the cell (i.e. on the Customer's process) until it reaches an upper limit, at which point anode performance suddenly deteriorates unacceptably, impacting on the cell performance – i.e. loss occurs (See Figure 2.).

Ideally, the upper specification limit for AP would have been established through rigorous scientific research into the in-cell performance of anodes with different AP values, covering a wide range of specific conditions such as: cell design and operating parameters, anode raw materials, and anode production conditions. The results of this work would then be used to set the specification limits. Unfortunately this is not what actually happens.

In reality, the AP specification limits are set based on standard industry values, and these are normally attributable to historical standards or expert opinion. It is rare for plant data to be used to establish the point at which loss in the cell becomes "economically" unacceptable, i.e. according to binary control, the point at which cell performance is suddenly impacted.

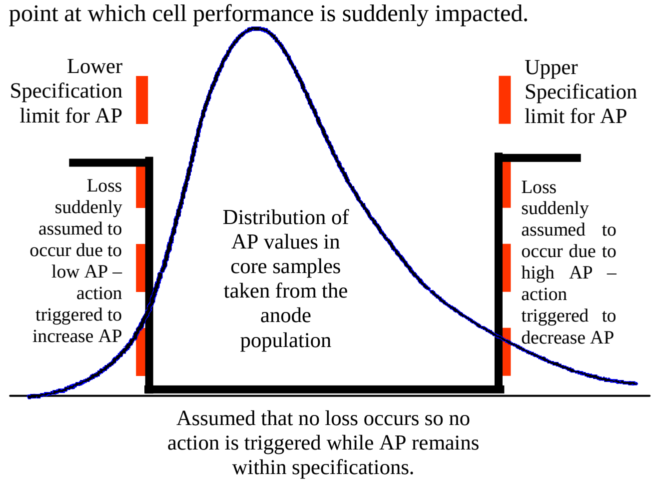

Figure 2. "Binary" control of anode core Air Permeability.

The result of the binary control approach is that AP can be anywhere within the arbitrary specification limits and no action will be taken, even if the measured AP varies wildly or if it tracks very close to one of the specification limits. Should the AP value move outside the specification limits, however, suddenly it is assumed that the AP value is no good, that loss occurs, and that action is required. All of this because the measured AP moved outside an arbitrary limit that has little technical validity or direct correlation with how the anode will perform in the customer's cells. The actions taken to "recover" the situation and get AP back into specification are likely to cause unnecessary variation in other anode properties. There is also the distinct possibility that if no action was taken, the AP values would come back within the specification limit of their own accord - if the move outside the limits was due to common cause variation.

There is an appropriate role for valid specifications; however, the "binary" method of control by comparison to specification should be avoided when making decisions about when action is required on the process. There is a better alternative:

<u>"On target with minimum variation" control of anode core permeability.</u>

There is likely to be an optimum AP value that results in the least loss to anode production and cell operations. AP values less than this optimum may result in loss due to difficulty in baking the anodes. The very low permeability hinders the release of pitch volatiles during baking and can result in internal cracking of anodes if "normal" heat-up rates are exceeded during the anode baking process. This results in loss through poor anode performance in the cells due to internal cracking, and/or through a reduction in baking furnace capacity due to the need to use longer fire cycles to reduce the anode heat-up rate to avoid internal cracking. With significant test work it might be possible to determine the lowest limit for AP based on the longest fire cycle that can be used and still meet production requirements and at what point internal cracking starts to have an unacceptable impact on the cells. The likelihood of this work being done is not high.

Air Permeability values greater than the optimum value cause loss by increasing the rate at which anode carbon is consumed in a non-productive way by gaseous attack in the cell, primarily due to reaction with Carbon Dioxide. A crude estimate of the magnitude of this loss can be made with Fischer's equation for calculating net anode consumption [3]. In this equation, the AP term has a coefficient of 9.3, suggesting that for each 1 nPm increase in average AP, an increase in anode consumption rate of 9.3 kg Carbon/Tonne of Aluminium can be expected. The upper specification limit could be set at the point where the Customer determines that the economic impact on the process from the additional carbon loss due to increasing AP is not acceptable.

A better alternative to "control to specification" is demonstrated in Figure 3, and can be described as "on target with minimum variation". Control decisions and actions are based on a view that economic loss occurs as the process varies from an optimum point regardless of whether it is in or out of specifications. Operational definitions of control points (action limits) are determined from a statistical analysis of process performance.

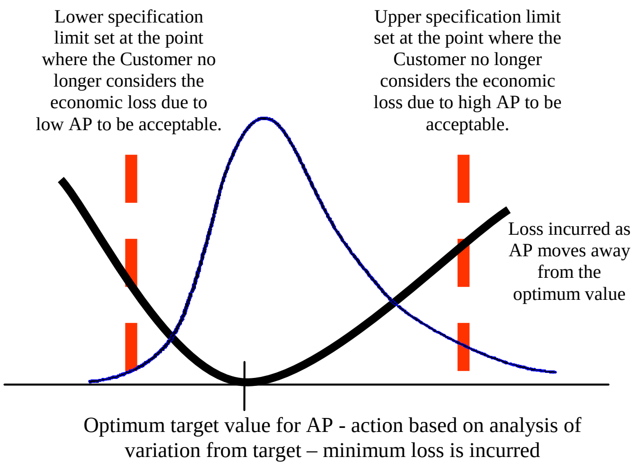

Figure 3. "On target with minimum variation" approach to control of anode core Air Permeability reduces variation and loss (new loss function)

In relation to Figure 3, action based on "comparison of current performance to specification" is not relevant. Appropriate process decisions and actions to achieve on target with minimum variation at the lowest total cost are:

1. Monitor the process through time, using statistically derived action limits based on actual performance behavior.
2. If special causes indicating a process change occur, to either side of target – take action to investigate, understand, and respond to these unique or assignable cause events.
3. If the process is stable (no special cause signals), take no action on specific outcomes, as action will likely result in increased variation.
4. If the process is stable, and DOES NOT meet specification limits, take action to improve the system through process redesign, method changes, engineering actions and so on.
5. If the process is stable and DOES meet the specification limits, determine if further process improvements are economically justified. If so, plan further improvements to reduce variation around an optimum target or to shift the target to a new low cost optimum.

In the "on target with minimum variation" approach, process decisions and actions for stability only require knowledge of the statistically derived limits and a running record of process performance. Specification limits are needed when comparing a stable system to minimum acceptable limits of capability and determining what, if any, further action is required to improve the process.

To properly apply the "on target with minimum variation" approach to process management requires the application of Shewhart's "Principles for Economic Control of Manufacturing Processes" by using control charts [4] to provide the following:

a. Control actions are based on cost effective responses to variation. Actions are triggered when control limits (calculated from process data) are exceeded, indicating that the variation is special cause and investigation is justified to find and remove (or incorporate) the root cause.
b. If the process is not capable of meeting valid specification limits, the focus is on improving the "process design" to reduce variation so that the specifications are achieved. (This contrasts with the "control to specification' approach of reacting to just the points that are outside the specification limits.) Until the process is capable of meeting Customer requirements (i.e. conforming to the specification limits), it may be necessary to reject non-conforming product; this is normally expensive and provides a further financial imperative to improve the process (i.e. reduce variation).
c. The never-ending cycle of analysis, action and improvement inherent in the "on target with minimum variation" approach is directly opposed to the "maintaining the status quo" outcome of controlling to specification. It provides the means to achieve levels of process performance previously unattained, yet necessary in this environment of intense competitive pressure for Carbon Operations.

"On target with minimum variation" - summary

The "on target with minimum variation" approach can be summarized as:

- Aim to achieve a process target value with minimum variation around this target so economic loss to the business is minimized and specification limits are not exceeded.
- Use thorough statistical analysis of the process to determine the control limits that trigger action and the impact of variation on the Customer.
- Take action when control limits calculated from the data are exceeded (or other statistical signals are present).
- If the process does not always result in values that meet the Customer's specification, the focus should be on reducing common cause variation in the process until capability to specifications is achieved.
- The objective is to maintain the process at minimum total cost resulting in maximum business value.

Inappropriate use of averages

A common practice in Carbon Plants is to aggregate data into averages for analysis, decision making – including about when action is required, and for reporting process status. Some of the problems associated with this data trap were covered in a previous publication [1], however the frequency at which the Authors encounter this trap suggests that it warrants further attention.

Reducing variation is a key to improving processes. Averaging data is the best way to hide variation in the data. The issue is immediately obvious – averaging data inhibits process improvement. It is a surprisingly simple fact, but one very often overlooked, that averages tell you nothing about the amount of variation within the data set that makes up the average. A number of issues/pitfalls are associated with inappropriate use of averages:

- What is reported is not what is sent to the Customer. This is due to the practice of using averages to report conformance to specifications or other quality criteria.
- "Phantom" averages – where unusual distributions in the data mean that there is very little, or in some cases no, "average" product in the batch.
- Patterns in data, crucial to understanding variation and enabling process improvement, are hidden by averages.
- Averages alone do not tell you anything about the historical context for the data – providing the historical context of data enables predictions to be made, the essence of effective process management and improvement.

Each of these pitfalls will be covered in more detail using examples drawn from plant operating data.

<u>What is reported is not what was sent to the customer.</u>

It is common for management to report performance to targets, specifications, etc by using shiftly, daily, weekly, or monthly averages of plant data. This practice is a good way to unintentionally or deliberately hide variation. Averaging (and the distortion of results that goes with it) appears to know no bounds. Often daily averages are the average of shiftly averages; weekly averages are the average of daily averages and so on. Taking the average of averages is an even better way of hiding variation.

As an example of "what is reported is not what is sent to the Customer", Figure 4 shows the distribution of paste elongation values (A measure related to paste viscosity at a fixed temperature) reported to the Potrooms customer on a monthly basis in a Soderberg smelter. The comments provided with this data were along the lines of "paste quality has met the required standards." The response from the Potlines was similar to "Thanks for the good paste quality – we can see that paste quality does not appear to explain why we are seeing some dry anode tops".

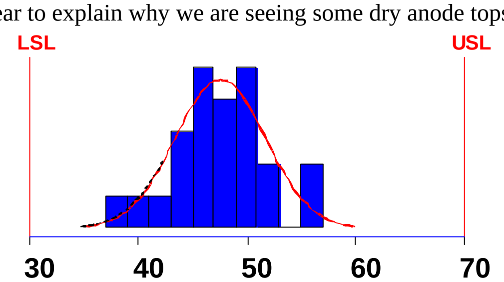

Figure 4. Histogram - paste elongation data as reported to Potrooms.

Analysis of Figure 4 suggests that the Carbon plant comments may appear appropriate – paste elongation (as reported) for the month was well within the Customer's specification limits. The problem is that the data presented to the Customer was a daily average of the elongation values taken every two hours. A very different picture emerges when the individual elongation values from the same monthly data set are plotted on a histogram (See Figure 5). The data plotted in Figure 5 shows that in reality there was quite poor conformance to the Customer's specifications during the month. The affect of averaging reduced the perceived variation to about a quarter of the level seen by the Customer. If this had been reported to Potlines it is not difficult to imagine that the paste quality discussions would have had a different tone and degree of urgency. The reporting of this data would also have suggested that paste quality be pursued as a possible cause of the dry anode tops.

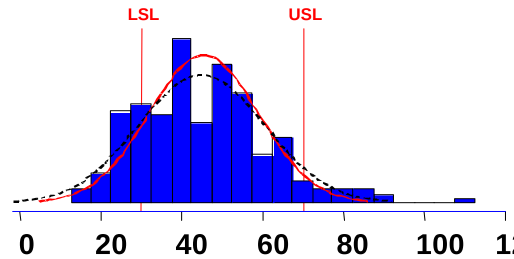

Figure 5. Histogram of individual paste elongation values. These values produced the daily averages plotted in Figure 4.

This problem of averages obscuring the real product quality sent to the Customer is also commonly encountered when examining how anode core property test results are reported to Potrooms.

<u>"Phantom" averages.</u>

When we see an average, a common assumption is that the average represents the most typical value of the parameter being measured, i.e. the average is the most representative value of the items measured. For example, if the Volatile Matter (VM) in green petroleum coke from a particular supplier was measured on each railcar delivered, and the individual railcar VM values were averaged for a week, then it would not be surprising to see this average VM value used to set up the kiln parameters for a calcining run for this green coke. In this example, only the average VM for the weekly green coke batch was reported from the Laboratory to the Kiln Superintendent.

There was one coke source that always caused problems for the Kiln Superintendent – with very unstable kiln operations and high variation in calcined coke Real Density. When investigating the problem, the individual railcar VM analyses were plotted as a histogram (Figure 6.).

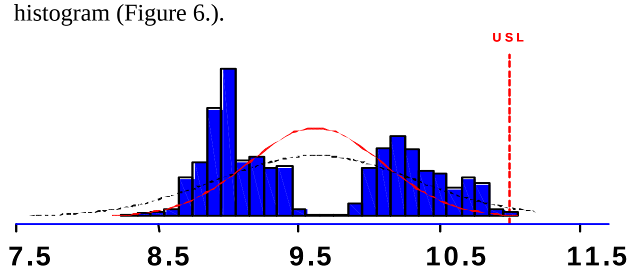

Figure 6. Histogram of individual railcar analyses of green coke VM delivered to a calciner. The average VM reported for this batch of green coke was 9.6%.

The average VM reported to the Kiln Superintendent was 9.6%. This was then used to set up the kiln for the coke batch from this supplier. As can be seen from Figure 6, there was not a single railcar load of green coke from the supplier that had a measured VM of 9.6%. The Superintendent was setting the kiln up for a coke that did not exist. As the slugs of higher and lower VM coke came through to the kiln for calcination, the kiln went out of control due to the variation in VM.

What had happened in this case was that the supplier was using two banks of delayed cokers to produce the green coke. Initially the banks had produced similar green coke qualities, but over time the qualities had drifted resulting in the different VM's evident in Figure 6. As the batch average VM value was all that was reported to the Kiln Superintendent, the divergence in quality was not picked up. The Laboratory had noticed something strange in the results, but had not commented as they had been told once that their job is to just generate the analytical data they are asked to, not to comment on the results. (This is an all too common perception by Laboratory staff and is very wasteful. Perhaps, they should be the first group to raise a warning flag that there might be something unusual in the results they generate.)

Unusual data patterns hidden by averaging.

Unfortunately not all of the data collected and used for decision making and action is reliable. When people are involved in data handling, there can be subtle pressure to not report values outside the limits that trigger action. Averaging the data usually removes any indication that the data is not accurate. The result is that decisions and actions are taken on the averaged (poor) data resulting in increased variation and waste, with opportunities to learn about how to improve the process missed. Plotting the data instead of using averages can, however highlight such data inconsistencies which can then be addressed. Examples of unusual trends in manually reported data are given in Figures 7 & 8.

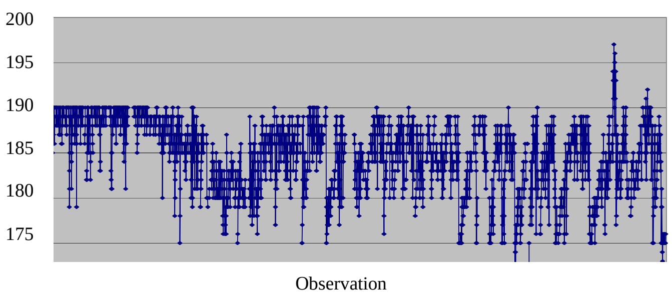

Figure 7. Manually collected mixing temperature (^{o}C) data. Action is triggered if the reported temperature exceeds 190^{o}C.

In both of these examples "unusual" cut-offs in the data can be seen at the point beyond which action is required. It would appear that the data has been manipulated to avoid additional work or other repercussions from exceeding the limits. The nature of these cut-offs in the data and their likely cause is not evident when just averages of the parameters are reported. Using poor data for decision making and control actions is as bad, or even worse, than having no data at all.

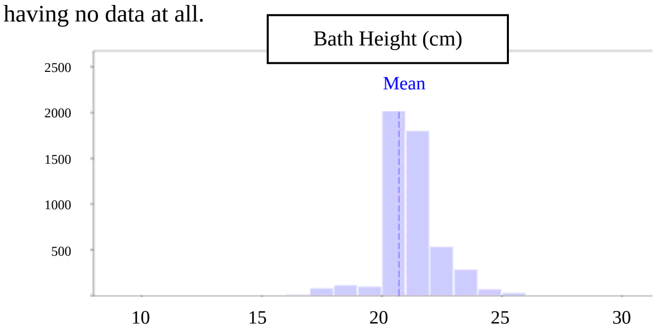

Figure 8. Manually collected bath height (cm) data. Action is required if the bath height is less than 20cm

Using averages without a historical context.

To extract full value from the most recent data, it must be analysed within its historical context. Just looking at the most recent averaged data gives no insight from its historical context. A Green Carbon plant has two Ball Mills for fine coke production. Blaine Index (BI) of the Ball Mill product is used to monitor and control Mill performance. The fine coke quality from the Mills is reported on a weekly basis as shown in Table I:

Table I. Weekly average and standard deviation values for fine coke products from Ball Mills 1 & 2.

Ball Mill 1- Blaine Index		Ball Mill 2 – Blaine Index	
Average	Standard Deviation	Average	Standard Deviation
3095	116	3102	117

The data provided in Table I suggest that the performance of the Mills is very similar, both in terms of the average BI and variation in BI over the week as measured by the standard deviation of the individual values. The conclusions are a lot different, however if the BI data are looked at in their historical context.

An Individual and Moving Range (IMR) control chart of the BI values from Mill 1 is shown in Figure 9. This shows the data to be randomly spread around the average and that under normal conditions we can expect BI values to be within the range 2746 to 3444. This is important as it enables us to predict the future performance of Mill 1 – unless a change is made to the operation of the Mill, we can expect that it will continue to produce fine coke with these properties. It is a relatively stable process with nearly all common cause variation.

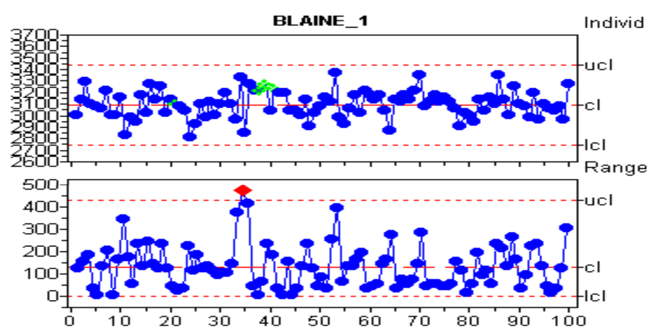

Figure 9. IMR Control chart of Blaine Index values of the fine product from Ball Mill 1 over a week.

If we now look at the performance of Mill 2 within its historical context, we can see different patterns in the Mill product data (See Figure 10.). In this case the BI values are not randomly spread around the average – as a result of instabilities in the Mill operation, there is a distinct pattern to the data with steps in BI. These steps will result in jumps up and down in aggregate pitch demand and variation in green anode quality. As expected with this, the anodes produced with fine coke from Mill 2 would be more variable than those produced from Mill 1 fine coke. This would not be suspected by just looking at the BI averages and standard deviations of the fine products from the two Ball Mills shown in Table II. Predicting the future performance of Mill 2 is more difficult than Mill 1 as the process not stable.

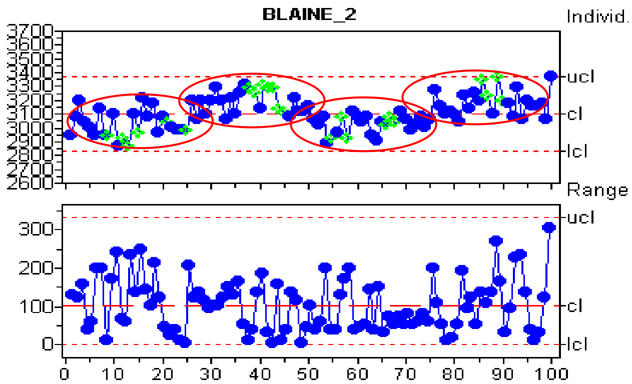

Figure 10. IMR Control chart of BI values of the fine product from Ball Mill 2 over a week. An unstable pattern of steps is evident in the data.

Avoiding the inappropriate use of averages – summary

- Using averages & standard deviations to describe process or product performance is useful however; it requires that the individual values that make up those statistics arise from a single, homogenous population (a stable process), which can only be determined through analyzing data in its historical context.
- Evaluate histograms and control charts plotted from the original data. If the volume of data involved prevents this (e.g. Individual green anode data) use appropriate averages with a measure of variation (standard deviation) and use a histogram to look "inside" the averages.
- Use the appropriate control chart such as average & standard deviation or average & range. Always determine and report conformance to specification limits and other criteria through proper statistical methods for capability analysis and using individual data points, never with averages.

Impact of sampling methods on the usefulness of data.

This data trap is commonly encountered when evaluating anode core property data, and this example will be expanded on. (The whole topic of the validity of anode core sampling strategies is worthy of a more detailed discussion than can be undertaken in this paper. For this reason it will only be covered briefly here, but will be the subject of a future publication.)

Data sampling strategies are determined by the purpose of the data being collected, i.e. what is the question you want to answer? This question is the starting point for designing a sampling strategy – as opposed to simply collecting and analyzing data and hoping it will answer your questions. You should not use data to answer questions for which the sampling strategy was not designed.

In anode core sampling, a question that is rarely answered well is "What is the purpose of sampling and testing cores taken from baked anodes?" There are at least two different answers:

1. To monitor and characterize the quality of anodes sent to Potrooms. In reality, the ability to provide useful data for this purpose is very limited – the turnaround time for core drilling, sample preparation, and testing is too slow compared with the in-cell life of an anode and the frequency of sampling is normally much lower than required to detect an important change in anode properties in a useful timeframe. Anode core properties sampled using typical strategies do not give a good representation of the overall anode quality sent to the Customer. At best these data may help with long-term trend analysis.
2. To monitor on-going process behavior in the Green Carbon plant – to study the impact of "upstream" process changes (such as raw material quality, green anode production conditions) on anode core properties.

To answer question 1, i.e. "What is the quality of the anodes sent to Potrooms?" the sampling strategy must capture all of the variation in green anode properties and all the variation due to the baking process (e.g. section, pit, layer, location in layer, pit condition, "straight" sections, cross-over sections, etc) .

To answer question 2, the sampling plan in the baking furnaces should be restricted to taking core samples from one location within the same pit in selected sections. This helps to "block" out the impact of some of the variation in the baking process.

If the sampling strategy for question 2 is adopted and you try to use the resulting data to give answer 1, the amount of variation in the anode core property results will be understated, masking potential sources of problems.

Alternatively, if the question 1 sampling strategy is adopted, and the results are used to try to give answer 2, the amount of variation in anode core data will be larger, making it difficult to detect changes in anode quality. The effect of upstream process changes are masked by, or confounded with, the variation coming from the baking process.

In either case, using data to answer questions for which the sampling strategy was not designed will result in a poor basis for decision-making and misleading results. This issue is largely ignored when anode core property data are used to try to address the myriad of questions that can be asked about it.

Conclusions

To avoid the data traps outlined in this paper and improve Carbon Plant operations through better use of data:

1. Adopt "on target with minimum variation" for process management instead of the binary "control to specification".
2. Avoid using averages for process analysis, decision making and reporting. Wherever possible use a graphical presentation of the individual data points instead.
3. Only use data to answer the questions that the sampling strategy was designed to answer.

References:

1. K.A. Sinclair, and B.A. Sadler, "Going Beyond SPC – why we need statistical thinking in operations such as Carbon Plants", *Light Metals 2004, TMS (The Minerals, Metals & Materials Society)*, 597-572.
2. K.A. Sinclair, R. Phelps, and B.A. Sadler, "The Integration of Lean and Six Sigma – A powerful improvement strategy for Carbon Plants," *Light Metals 2005, TMS (The Minerals, Metals & Materials Society)*, 641-646.
3. R&D Carbon Ltd, *Anodes for the Aluminium Industry*, (R&D Carbon Ltd, Sierre, Switzerland, 1st Ed. 1995), 342.
4. D.J. Wheeler, *Understanding Variation – the key to Managing Chaos.* (SPC Press, Knoxville, Tenn., 1993)