UNSUPERVISED CHANGE DETECTION FOR WITHIN-RUN MANUFACTURING DATA

Alan H. Brothers, Ph.D.
Benjamin II Carlos

Manufacturing Engineering Business Intelligence and Data Analytics
Western Digital Corporation, 109 Technology Ave., SEPZ, Laguna Technopark,Binan, Laguna, Philippines
alan.harold.brothers@wdc.com
benjamin.carlos@wdc.com

ABSTRACT

Optimal control over a manufacturing process requires monitoring of multiple critical process parameters (e.g. temperatures, pressures, flow rates, voltages, etc.) to ensure: (1) consistency between tools; (2) consistency over time within each tool; and (3) compliance with the prescribed recipe. While modern technology makes collection and storage of high-resolution within-process data feasible, the quantity of such data makes it nearly impossible for engineers to effectively monitor, especially in high-volume manufacturing where controlling a single process may require monitoring of many parameters, for each of many different products/recipes, and for each of dozens or even hundreds of machines.

This problem may sometimes be addressed by reducing the full within-run data to a single parameter such as a mean, maximum, slope, etc., which can be easily visualized to monitor compliance within each tool, and across tools. However, parameterization of this sort is not always feasible, for example because the appropriate parameterization is not obvious due to the complex shape of the data, or varies widely across the different parameters being monitored.

The objective of this work was to develop a more flexible platform for process control via monitoring of process tool parameters, which could be broadly applied to parameter data irrespective of its shape. The resulting platform achieves this through the use of unsupervised pattern matching algorithms such as dynamic time warping (DTW), allowing engineers to monitor and investigate hundreds of pieces of data using a simple web-based platform with a few simple visualizations.

1. 0 INTRODUCTION

The need for a new tool parameter monitoring approach became clear when engineers noticed that drift in performance was related to changes in the shape of chamber pressure vs. time within a deposition tool. The engineers could use the performance data to identify the 'ideal' shape of the curve, and could restore performance by restoring this shape via tool maintenance; but it was not obvious how to parameterize the complex curve for each run, such that deviations from the ideal curve could be monitored over time within each tool, across dozens of tools, in order to drive more effective maintenance. And even if a method were found for this parameter, it was likewise not obvious how the approach could be generalized and applied to all the other parameters measured in the deposition tools. They therefore requested help to develop a general platform for monitoring deviation of within-run parameter curves.

Put another way, the problem was to define an algorithm to capture the deviation between an actual and an ideal curve in a single 'score', for example the deviation between the actual pressure vs. time in a particular process run, vs. the ideal pressure vs. time as determined by analysis of downstream outcomes. Given such a score, tool maintenance can be improved by creating a platform for continuously computing the scores of actual process runs, and allowing engineers to visually monitor the scores of each tool over time, and the differences in scores between tools at any time. To further reduce the workload, engineers could be given automated alerts based on conventional statistical process control logic applied to the scores, such that visual monitoring would only be needed when the time for maintenance was near.

2. 0 REVIEW OF RELATED WORK

Not Applicable.

3.0 METHODOLOGY

Although many approaches could be used, it is natural to derive the above 'score' from some measurement of distance between the actual and ideal curves. For example, one could compute the score as the mean absolute difference in pressure between the actual and ideal curves at each time step, often referred to as the 'Euclidean' distance. However, this intuitive approach suffers from several well-known limitations when applied to real process data. For example, suppose a deposition process involves applying a pulse of electrical current, starting whenever a stable chamber pressure is reached. If the duration and magnitude of the pulse are important but its exact timing is not, Euclidean distance would be inappropriate, since shifts in the pulse timing would yield large Euclidean distances, and thus false alarms. A number of similar scenarios can be envisioned, which show that naïve Euclidean distance can cause undesirably high 'scores' under real process conditions[1].

A well-documented alternative is the DTW distance algorithm, which was developed in the 1970's and has been applied in many domains that are similar to within-run process data, for example analysis of electrocardiograms[1]. The principle of the DTW algorithm is to compute distance by evaluating many potential mappings between the points in the actual curve and those of the ideal curve, and then choosing the mapping which gives the lowest distance / best alignment between the two. This approach creates some flexibility in scoring real data, reducing over-estimation of deviation by allowing for some translation and distortion.

Various constraints are also commonly applied to the DTW mapping process to prevent excessive / unrealistic distortion of the actual data, which would lead to under-estimation of the real deviation. In addition, variants of the algorithm have been created to make it more robust with real data, for example the psi-DTW variant, in which some leading and/or trailing data points may be ignored during fitting, to accommodate issues such as poor 'cropping' of the data[2].

This approach is used by default for analyzing process data in the platform described here, although technically the platform can accommodate any algorithm yielding a positive number as a 'score'. Specifically, the platform is based on a set of defined 'signals', each of which is defined by a process tool, recipe name, recipe step, and parameter. Each signal is associated with one or more 'triggers', consisting of a scoring algorithm such as psi-DTW, and a set of parameters to be passed for that algorithm. Scoring is performed relative to a 'golden pattern', usually fitted on a specific time period, or specific set of runs, defined by the users based on strong downstream performance. Each trigger is also associated with a set of trigger conditions, for example hard and soft upper limits, which are used to generate automated alerts to the engineers when the scores computed using the golden pattern meet certain conditions.

Data for each signal are monitored continuously, and new runs are scored according to the defined triggers, with alerts being generated whenever the trigger conditions are met. Users are given ownership of signals and triggers, allowing them to manage the algorithms, algorithm parameters, and alert criteria via a web interface. Users are also able to view the score data at any time using commercial business intelligence software such as Tableau, Power BI, or Spotfire. In addition, users are given a set of web-based utilities for building new golden patterns based on historical data, allowing them to update the 'best-known' parametric curves whenever needed.

It is important to note that this approach is best described as unsupervised, in the sense that the score only quantifies deviation of any kind from the golden pattern. As there are an infinite number of possible ways a real curve may deviate, including many deviations which will map to the same score, no attempt is made to assess whether a specific score/deviation is harmful or helpful, i.e. to predict downstream performance. The platform exists to let engineers see when significant deviations occur, and to let them quickly visualize the deviant data, and take action without waiting on downstream outcomes. They employ their domain expertise in the assessment of what action to take, if any.

4.0 RESULTS AND DISCUSSION

The platform has been applied successfully in production for monitoring at two factories, where it helps the engineering teams to review both tool stability and tool-to-tool variation at a glance.

Figure *1* shows the score vs. time for one example signal, with each point representing a process run. At a glance, the engineer can observe that: (a) there is significant deviation from the ideal pattern after each preventive maintenance cycle (see for example, 10 Oct), which regresses towards ideal behavior over time; (b) the rate of recovery after preventive maintenance has been slowing (see for example, the more shallow slope of the curve after the 18 Jan cycle compared to the 10 Oct cycle); (c) in addition to the cyclic behavior around preventive maintenance, there is also a long-term regression taking place over several months; and (d) a major step-change took place around 12 Feb. Such insights would be tedious or even impossible to gather by reviewing individual run data manually, but are easily acquired from the score data, leading to deeper understanding of the process and the ability to better engineer maintenance procedures.

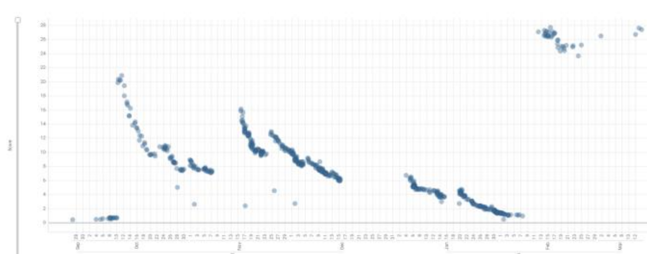

Figure 1: Example score vs. time data, showing clear structure that would be difficult to see when reviewing the raw data manually.

Viewing scores from the same measurement in more than one tool easily reveals tool to tool variation as well, as shown in Figure *2*, where both tools show regression within the maintenance cycle, but one tool (in red) shows consistently lower scores / closer-to-optimal within-run behavior when compared to the other (in blue).

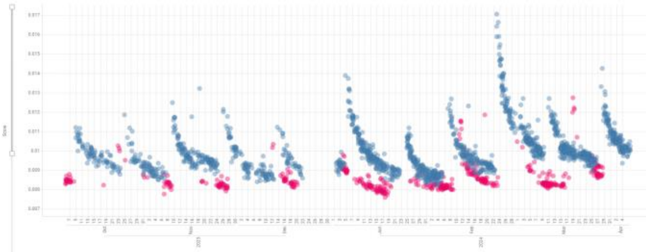

Figure 2: Score data like the previous figure, except comparing two process tools. The tool in red has consistently lower scores than the tool in blue, indicating a closer fit to the golden pattern.

The ease with which the data can be reviewed also promote best practices in the form of more frequent and regular monitoring, and several instances have been captured as a result in which critical tool sensors fell out of calibration. Proactive monitoring allows these cases to be captured, and recalibration to be triggered automatically.

Figure 3 shows one such example. In this case, the signal consists of the ratio of two sensor measurements, the physical relationship of which is known, and captured in the golden pattern. Deviation of this composite signal over a period of months can be seen, and used directly as a means of triggering calibration of the underlying sensors.

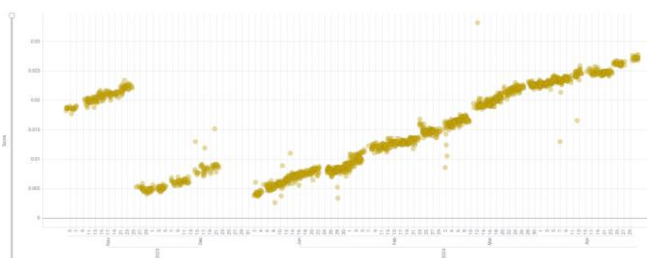

Figure 3: Score data from a composite signal reflecting a fixed, known physical relationship. Scores from signals configured in this way are a direct reflection of sensor miscalibration, and as such can be used to trigger calibration.

Finally, in at least one case, score data revealed the presence of a cooling water leak well in advance of the leak being physically discovered. This example is shown below in Figure *4*, which shows scores for water residual gas analyzer signals at various locations in a process. While physical evidence of a water leak in the tool was only found in mid November (right edge of the figure), the score data clearly reveal that the origin of the leak was in mid October (center left of the figure). With hindsight, it would certainly have been possible to also identify the origin of the leak by looking at the raw data in mid October, without the platform. However, this is just one of hundreds of signals across several machines running constantly in production, so the level of surveillance needed for this is not sustainable; a suitably defined automatic trigger inside the platform guarantees detection with no manual surveillance.

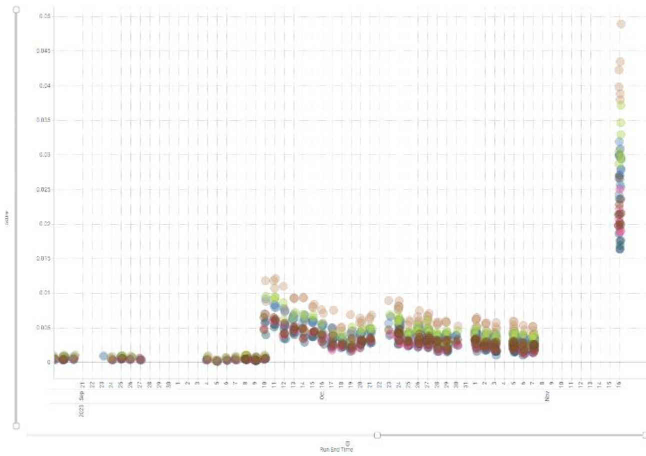

Figure 4: Score data from a residual gas analyzer, revealing the onset of a water leak (center) nearly one month before it worsened to the point of being physically visible (right edge).

Finally, it is important to remember that reducing a within-run parametric time series of potentially hundreds or thousands of measurements down to a single score represents heavy compression, which will not allow engineers to diagnose the nature of the deviation or assess its risk. Therefore, the platform must always provide a quick way to pivot from high-level scores back to the original within-run time series data. In the current platform, this is done simply by selecting points from the plot of scores vs. time, and then clicking to open the corresponding raw data. An example of this is shown below, in which some anomalous runs (dashed circle, top center of the upper panel) were selected for detailed raw data review (lower panel). Raw data from these anomalous runs (at bottom) deviated significantly from the golden pattern (in gold), as expected.

Figure 5: Example of pivoting from high-level observations based on scores, back to examination of specific within-run data of interest.

5.0 CONCLUSION

While modern technology makes it theoretically possible for engineers to monitor nearly every critical parameter in every process tool, in practice such surveillance is not possible without platforms designed to analyze the data, and point the engineers towards only those tools, parameters, and runs that they need to attend to. While supervised methods based on downstream outcomes may offer a solution in some such cases, the resources needed to train and maintain these models quickly becomes prohibitive as the numbers of parameters, tools, and outcomes increase.

As an alternative, this paper described a platform for within-run monitoring based on flexible similarity metrics such as the DTW distance, which provides more scalability than supervised methods, but more utility and flexibility than naïve methods like tracking aggregate statistics. This platform promotes best practices by making it easy for engineers to quickly review within-tool and tool-to-tool variation, and to pivot back to the raw within-run data to investigate any changes of interest. In this capacity it has helped the engineering team to both expose patterns of machine behavior not previously known, and to develop better, more intelligent processes for tool and sensor maintenance and calibration.

6.0 RECOMMENDATIONS

It has already been noted that the platform is not limited to the DTW algorithm. However, even when DTW is used, it is important to point out while DTW is most intuitively employed for time series data, e.g. pressure vs. time within a process run, it may be employed for any ordered series. For example, the algorithm could be applied to analysis of spatial rather than temporal variation. Doing so would require some effort to generalize the implementation in code but would significantly expand the value of the platform in manufacturing, as spatial variations across products, tooling and fixturing, or even facility spaces are also important to understand and monitor.

7.0 ACKNOWLEDGMENT

The authors gratefully acknowledge the support of the Western Digital IT, analytics, and engineering teams, without whom such a complex platform would not be possible. In particular, the authors would like to thank Yimin Zeng, Abhishek Kumar Gaur, Alan Huang, Pakawat Suaysuwan, Phuwanai Bunnak, Budi Suswadi, Chee Kheng Lim, and Luis Geoffrey Macapayag. They are also extremely grateful for the efforts of Tze Ping Kee and Zhenni Li of the Western Digital Media Operations team, which inspired this work.

8.0 REFERENCES

1. T. Giorgino, **J. Stat. Software**, 31 (7), 2009, 1-24.
2. Silva, D. F., Batista G. E. A. P. A. and Eamonn Keogh, ***IEEE 16th International Conference on Data Mining (ICDM)***, 2016, 209-1214.

9.0 ABOUT THE AUTHORS

Alan H. Brothers is a Data Scientist at Western Digital Corporation, focusing on analytics within the slider fabrication process. He has more than 15 years of experience as a researcher, developer, process engineer, and data scientist in the data storage, electronics manufacturing, and defense industries.

Benjamin G. Carlos II is Program Manager for Business Intelligence and Analytics for Manufacturing Engineering at Western Digital Corporation, with more than 15 years of experience in Test and HDD Recording Sub Systems Integration.