INCREASING EFFICIENCY BY ELIMINATING END OF LINE TESTING
EXECUTIVE SUMMARY

End-of-Line (EOL) testing is a crucial part of the automotive manufacturing process, acting as a final check on complex assemblies to minimize the risks of shipping defective units to customers. However, despite the obvious benefits, there are significant limitations to end of line testing, and the need to conduct end-of-line tests on every finished unit reduces throughput and diverts resources from more valuable activities.

This white paper lays out the case for and against comprehensive end of line testing and offers an alternative quality methodology. By integrating advanced analytics and machine learning into production lines, automakers can mitigate the need to subject every unit to end-of-line testing while simultaneously reducing warranty claims, lowering costs, and improving throughput.
What is EoL Testing?
EOL testing is the process of testing the overall functionality of a manufactured product. In the automotive industry, creating a method of evaluating a complex assembly's function is a crucial step in building a profitable assembly line that manufacturers reliable products.

For example, for automakers producing geared assemblies, the noise those assemblies generate can be an early indicator of manufacturing defects. These defects can lead to both in-passerenger comfort issues or even failures that could risk passenger safety. Using noise, vibration, and harshness (NVH) testing at the end of the assembly line can identify the early warning signals indicative of such problems.

This is why end-of-line testing is so important, since it is the last checkpoint before a product leaves the factory. Millions of auto parts are manufactured every day, representing hundreds of millions of dollars of potential revenue. EOL testing ensures those parts have been manufactured to the appropriate specifications and gives engineers an indication of when to start troubleshooting to reduce failure rates.

Benefits of EoL
EOL testing can drastically reduce the total failure rate of an assembly line. End-of-line testing also benefits companies as a whole, since fewer defects mean fewer warranty claims or recalls in the future. It's especially important for high-value products, including car seats, engines, transmissions, electrical systems, and other complex assemblies.
LIMITATIONS OF EOL

Despite the obvious benefits of end-of-line testing, like any quality process, it’s also limited in several important ways. Under the wrong circumstances, these limitations can negatively impact an automaker’s bottom line and undermine the very reasons for having EOL testing in the first place. The three major issues with end-of-line testing are: cost, accrued inefficiency, and lack of impact. Let’s look at each one in turn.

Cost
While there are always costs associated with manufacturing quality, the costs of end of line testing in automotive manufacturing tend to be particularly high.

Consider hot testing, in which a completed engine is run on a test bench with the aim of checking all the engine’s operating parameters just as they would function in an actual vehicle. Obviously, conducting such tests requires dedicated equipment and between rigging, derigging and running the test itself, the entire process can take anywhere from 18-45 minutes. Eliminating the need to run even a fraction of these tests would have a significant impact on the bottom line.

Accrued Inefficiency
The obvious response to the preceding issue is that even if end-of-line testing isn’t perfect, no quality process is and it’s better than nothing. If the point of EOL testing is to minimize the risks of warranty issues, it only needs to cost less than those issues to be worthwhile. The problem with this response is that it doesn’t consider all the potential outcomes of end-of-line testing. When an EOL test correctly identifies a defective unit, what then?

The manufacturer is faced with a choice between re-working the part, scrapping it, or in some cases just testing it again and hoping for better results. Though not uncommon, the last option should give any quality engineer pause, since it inherently introduces more uncertainty into their measurements. For the first two options, the result is decreased efficiency, since the end of production is the costliest point at which a manufacturer can scrap or rework an assembly.

Simply put: the sooner one catches a problem, the less costly it is to fix it.

Lack of Impact
Even when taken together, one might think that the cost and the potential for accrued inefficiencies are outweighed by the value of having a final check at the end of the line which minimizes the risk of shipping defective products. However, end-of-line testing is not the only way to minimize such risk, and its third major limitation can be reason enough on its own to seek alternative methods.

In AIAG & Deloitte’s Quality 2020 Report, automotive OEMs and suppliers both ranked Problem Solving as one of the most critical issues impacting quality. In addition, both cited Lack of Root Cause Analysis as one of the main reasons for this issue. While end-of-line testing can point to the existence of a problem, it cannot explain why the problem is occurring or how to fix it. To put it another way: end-of-line testing may catch defective parts, but it cannot improve first time yield on its own.

Reworking a single gear pulled off the production line is not too costly, pulling a gear out of an assembled transmission is expensive, but pulling a transmission from a completed vehicle to fix that gear is prohibitively expensive and unacceptable – that vehicle actually reaching a customer, worse yet.

- Brüel & Kjær – “End-of-line testing – All geared up and ready to go”
Although there are significant limitations to end-of-line testing, it would be presumptuous to eliminate the process entirely, especially in automotive. Fixing problems at the end of the line is indeed costly and time consuming, but it’s well worth it compared to the costs of dealing with those problems if they reach the customer.

In fact, a recent review of the costs of quality for the automotive industry found that external failures can cost a company 10-15% of sales revenue, versus 4-6% of revenue for measuring, evaluating or auditing products, including end-of-line testing.

However, the same review also found that preventative approaches only cost 1% of sales revenue, and so if end-of-line testing can be even partially eliminated in favour of a more proactive, preventative methodology, there are significant efficiencies to be gained and costs to be saved.

**What’s the Alternative**

Ask any quality professional for their advice on reducing scrap and re-work (the most likely result of end-of-line testing), and the first suggestion is almost always to collect more data. Whether working under lean principles or other quality management techniques, manufacturers working to improve their defect rates and the associated costs will tend to focus on traceability. By collecting data from every stage of production, defects encountered during end-of-line testing will be traceable to specific parts or processes.

*Effective identification of root causes was consistently cited by OEMs and suppliers. Leverage advanced predictive analytics capabilities to sift through big data and improve root cause analysis capabilities.*

- **Key Takeaways - AIAG & Deloitte Quality 2020 Report**
This approach also points to an alternative to subjecting every unit to end-of-line testing. With enough traceability data and the right set of analytics tools, automotive manufacturers can predict the likelihood that a given unit will pass an end of line test. Using these predictions, automakers can thereby eliminate the need for end-of-line testing for many or even most of their finished units.

Type I Error: False Positive, e.g., classifying a non-defective unit as defective
Type II Error: False Negative, e.g., classifying a defective unit as non-defective

The key is using machine learning for classification. The performance of a machine learning classifier is assessed via confusion matrices, which compare the actual condition of a unit (e.g., defective or not) with the classifier’s predicted condition. Accuracy is defined in terms of true positives and true negatives—a positive in this context is a defective unit. However, because defective units represent a minority class in manufacturing, accuracy alone is insufficient as a measure of classifier performance.

<table>
<thead>
<tr>
<th></th>
<th>True Defect</th>
<th>True No Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Defect</td>
<td>7%</td>
<td>5% (Type I Error)</td>
</tr>
<tr>
<td>Predicted No Defect</td>
<td>8% (Type II Error)</td>
<td>80%</td>
</tr>
</tbody>
</table>

False Positive Rate = 5/(80+5) = 6%
True Positive Rate = 1 - 8/(7+8) = 47%

As the confusion matrix above illustrates, a machine learning model could operate at 87% accuracy but still miss nearly half of all defective units. In order to get a more complete picture of machine learning model performance, data scientists judge classifiers in terms of their true positive rate and their false negative rate.
Each line represents the performance of a different machine learning model as a result of adjusting the threshold at which it classifies a unit as defective (positive) or non-defective (negative). These curves appear more or less smooth depending on the number of units involved, with fewer units resulting in the appearance of a step function as adjustments are made to the threshold for classification.

Exactly how the threshold is adjusted will depend on the needs and priorities of each manufacturer. Many automakers may wish to maximize sensitivity (i.e., the true positive rate), in order to capture as many defective units as possible.

On the other hand, some manufacturers may be more interested in maximizing specificity (i.e., the false positive rate) so that all non-defective units are identified as such. The former case is focused on identifying units that are highly likely to fail an end-of-line test, while the latter focuses on identifying those which are highly likely to pass.

In either case, machine learning reduces the need to actually put units through the end-of-line test by predicting the outcome before it’s conducted, eliminating unnecessary tests.
To the extent that a manufacturer can eliminate end-of-line testing, they can also avoid the limitations noted in the previous section. By reducing their reliance on a single point of insight at the end of the line in favour of advanced analytics combined with traceability throughout the manufacturing process, automakers can see fewer warranty issues, lower costs and greater throughput.

**Fewer Warranty Issues**
Many automotive manufacturers are already collecting data from every stage of their production process, but few are actually leveraging that data to its full potential. In fact, it’s been estimated that as much as 70% of the data within an enterprise goes unused.

However, with a fully-integrated analytics platform ingesting data from every stage of the production process, the odds of a defective unit escaping the factory drop considerably. In this context, the idea is not to replace end-of-line testing with machine learning, but rather to use the latter to augment the former.

Machine learning models are capable of detecting issues that other quality processes miss, such as identifying variations which are predictive of failures, even when those variations fall within normally acceptable tolerances.

By supplementing end-of-line testing with advanced analytics, automotive manufacturers can significantly reduce their defect rate, and hence the number of potential warranty claims.

**Improved Throughput**
In addition to increasing product quality and lowering operating costs, using advanced analytics to eliminate the need to subject every unit to end-of-line testing can also improve overall throughput. By removing end-of-line testing as a bottleneck in production and replacing it with an analytics platform that constantly monitors the data coming off the line, automotive manufacturers can produce more units without sacrificing quality.

Indeed, they will most likely see an increase in quality, further contributing to their first time through yield. Fewer units failing their end of line tests—whether because they are bypassing the test based on machine learning model predictions or because the models are catching issues earlier in production that could result in failure—means fewer units that need to be scrapped, reworked or re-tested.

**Lower Cost**
Given the supposed necessity of end-of-line testing, manufacturers will design their production schedules to account for lost time and money due to testing, scrap, rework, and re-testing. If every assembly is being put through end-of-line testing at least once, that means the requisite test equipment must be operated and maintained by one or more dedicated personnel. Moreover, any disruption to the testing station, such as unexpected calibrations or maintenance overhauls, will inevitably slow down production, require additional expenditures, or both.

By eliminating their overreliance on end-of-line testing by adopting advanced analytics, automotive manufacturers can free up human and other resources which can then be reassigned to other, more profitable areas of the business.
USE CASE

Consider a typical Tier 1 automotive manufacturer making complex mechanical assemblies, such as axles or transmissions. Many such companies use noise, vibration and harshness (NVH) testing as a surrogate for other data points at the end of the line, such as gear contact patterns. Based on the way an assembly sounds, the company’s quality engineers can determine its gear pattern and whether it needs to be adjusted.

This is a time-intensive process, adding minutes to the total manufacturing time for every unit.

However, the manufacturer can save considerable time by using a machine learning model that compiles data from the entire assembly process and uses it to predict what the result of that NVH test will be.

Even if the model does not predict with 100% probability that a unit will pass or fail the end-of-line test, it can mitigate the need to conduct NVH testing on every finished assembly. How much mitigation is possible depends on the classification model and the manufacturer’s specific needs.

For example, if the threshold at which a unit will be classified as a fail is set very high (e.g., 90% probability or more), the model will generate very few false positives (i.e., units that are predicted to fail the test but which actually pass) but many false negatives (units that are predicted to pass the test but which actually fail).

Alternatively, if the threshold is set significantly lower (e.g., 30%), the model will generate many false positives, but few false negatives. The receiver operating characteristic (ROC) curve shows how this discrimination threshold operates in more detail.
REFERENCES


• “Metric of the Month: Reduce Scrap and Rework Costs” Supply and Demand Chain Executive, July 2, 2018 https://www.sdcexec.com/warehousing/blog/21004027/apqc-reduce-scrap-and-rework-costs

• “Hadoop is Data’s Darling for a Reason” Forrester Blog, Mike Gaultieri, January 21, 2016 https://go.forrester.com/blogs/hadoop-is-dataas-darling-for-a-reason/


