IMPROVEMENTS IN GAKUSHIN WEAR TESTING THROUGH LABORATORY AUTOMATION

by

B. HAGEMAN,* J. HOEFLER, J. GREGORY AND W. STAPELKAMP
The Dow Chemical Company, Spring House Technical Center
727 NORRISTOWN ROAD, SPRING HOUSE, PA 19477, USA

ABSTRACT

Coatings designed for leather finishes are subjected to ever increasing performance specifications, requiring accurate assessments to properly judge performance. Unfortunately, the results of many tests utilized to determine leather finish performance are plagued by high levels of variability. One such test is the Gakushin wear resistance test, which is required by a number of Japanese auto manufacturers, but can be used more generally to measure wear performance important to most automotive OEMs. Gakushin wear testing also requires intensive operator monitoring over lengthy test cycles and does not present an easily determined endpoint. It is therefore often run unmonitored as a pass/fail test, forfeiting the potential to detect performance gradients among samples.

In order to address these challenges, a new system has been developed that preserves the critical components of the wearing system, but adds automation and imaging systems that allow it to run completely unattended. A digital control circuit that periodically interrupts the wearing cycle initiates the capture and storage of high resolution images of the leather samples under test. After the test is completed, the operator is able to view the sequential images, selecting the exact point of coating failure. One benefit of this system is higher productivity, since the test can be run unattended. Also, since the operator is able to evaluate images of the coating both before and after the selected endpoint, the error associated with endpoint detection is minimized and the test yields more precise data than when performed in pass/fail mode. This combination of higher productivity and more precise endpoint detection facilitates the use of statistical data analysis on the results. In addition, the images showing coating wear following the endpoint allow an analysis of failure mode. These factors combine to make the improved Gakushin instrument a potent product development tool for wear resistant topcoats.

RESUMEN

Los acabados diseñados para las terminaciones del cuero son sometidos cada vez a mayores especificaciones de performance, lo que requiere de una evaluación exacta a fin de juzgar adecuadamente el desempeño. Desafortunadamente, los resultados de varios ensayos utilizados para determinar la performance del acabado del cuero se ven afectadas por altos niveles de variabilidad. Uno de estos ensayos es la prueba de resistencia al desgaste Gakushin, el cuál es requerido por una serie de fabricantes japoneses de automóviles, pero puede ser utilizado generalmente para medir la performance de desgaste, importante para los fabricantes de autopartes. La prueba de resistencia al desgaste Gakushin también requiere un seguimiento intensivo del operador en largos ciclos de prueba y no presenta un punto final fácilmente determinable. Por lo tanto, a menudo se ejecuta sin monitoreo como un test pasa/no pasa, perdiendo el potencial de detectar gradientes de performance entre las muestras.

Para hacer frente a estos desafíos, un nuevo sistema ha sido desarrollado protegiendo los componentes críticos del sistema de desgaste, y añadiendo los sistemas de automatización y de imagen que permiten que se ejecute completamente sin monitoreo. Un circuito de control digital periódicamente interrumpe el ciclo iniciando la captura y el almacenamiento de imágenes de alta resolución de las muestras de cuero a ensayar. Después que la prueba se ha completado, el operador es capaz de ver las imágenes secuenciales, seleccionando el punto exacto de la falla del acabado. Una ventaja de este sistema es una mayor productividad, ya que la prueba se puede ejecutar sin vigilancia. También, puesto que el operador es capaz de evaluar las imágenes del acabado, tanto antes como después de que el punto final seleccionado, el error asociado con la detección del punto final se minimiza y el ensayo proporciona datos más precisos que cuando se realiza en modo pasa/no pasa. Esta combinación de una mayor

---

*Corresponding author e-mail: bradleykhageman@dow.com

A **Technical Note** presented at the 108th Annual Meeting of the American Leather Chemists Association, June 7-10, 2012, Grandover Resort and Conference Center, Greensboro, NC
Manuscript received March 5, 2013, accepted for publication March 6, 2013

productividad y mayor precisión en la detección de punto final facilita el uso de análisis de datos estadísticos sobre los resultados. Además, las imágenes que muestran el desgaste del acabado tras el punto final permiten un análisis de modo de falla. Estos factores se combinan para hacer del instrumento Gakushin mejorado una herramienta potente para el desarrollo de productos acabados resistentes al desgaste.

INTRODUCTION

General

In many automobile product lines, leather upholstery is offered as an optional upgrade at additional cost. Leather upholstery is generally viewed as a higher quality, luxury alternative to cloth upholstery with a better aesthetic presentation, including appearance and comfort. Positioned at a higher price point than cloth, leather interiors must also maintain their performance and quality for the life of the vehicle. In order to meet these requirements, automotive OEMs require leather finishes designed for automotive interior upholstery to meet stringent performance requirements. One performance area of particular interest, especially for Asian auto manufacturers, is resistance to abrasion. A number of test methods exist for evaluating wear, including Taber wear (ASTM D3884), Chrysler Wyzenbeek (Chrysler 463KB-06-01), and Ford Pilling (FLTM BN108-14)

Description of Gakushin Wear Test

Many Asian auto OEMs specify use of the Gakushin instrument to perform test method JIS L 0849, which measures the colorfastness of a textile subjected to rubbing. They have also made modifications to this test to enable the wear resistance of leather upholstery finishes to be evaluated with the Gakushin instrument (hereafter referred to as the Gakushin test and shown in Figure 1). In this variation, leather articles to be tested are cut into a rectangular shape measuring 10X50mm using a hydraulic press and cutting die. The leather strips are then fitted to the weighted test heads of the instrument with double-sided tape. The abrading medium consists of pieces of #6 cotton duck cut with the press and die into strips measuring 30X250mm. The cotton fabric is mounted to the movable curved platen of the instrument and held in place with the attached spring-loaded clamps. Five hundred gram weights are added to each test head, which brings the total weight of each head to 1000g. The movable platen travels 100 mm forward and then returns to the starting position to complete a cycle, operating at a rate of 30 cycles per minute. Periodically during the run, the test is interrupted to determine if the endpoint has been reached. To make this determination, the operator places the heads in the raised position used for sample loading and examines the appearance of the finish using either the naked eye or a handheld magnifying glass. At the first instance of the russet becoming visible under the abraded finish, the operator notes the number of completed cycles as the test endpoint. The most rigorous specifications require 20,000 cycles to be achieved in order to pass the test. At the specified rate of 30 cycles per minute, total elapsed time for the test, not including interruptions for sample evaluation, is slightly over 11 hours.

Figure 1. Gakushin test instrument.

Advantages and Disadvantages of the Gakushin Test

Although the Gakushin test is reported to provide high quality data when compared to other abrasive wear tests for upholstery finishes,[1] it is not without disadvantages. If the instrument is simply programmed to complete a specified number of cycles and the samples are then evaluated on a pass/fail basis, minimal operator input is required. At the same time, however, this approach provides only the coarsest level of detail and does not allow a relative ranking of performance among samples. On the other hand, if the operator monitors the test and determines the number of cycles-to-failure for each sample, a much richer data set is obtained. This type of procedure captures enough experimental detail to allow a ranking of every sample against its neighbors. This level of detail comes at a price, however, as the experimentalist must closely monitor the entire 11 hour test, ensuring that the moment of breakthrough to russet is observed in order to record the correct number of cycles. When testing is conducted by a trained observer, a moderate portion of the test only requires cursory monitoring, since a finish that is near the failure point exhibits dull, pitted grooves. In the absence of this appearance, the test may be allowed to run largely unattended. However, as soon as the grooves start to develop, careful attention must be given to the sample in question in order to detect the exact endpoint. Since the failures of all six samples loaded onto the instrument may occur sequentially, this can add up to a significant amount of time. Near an endpoint, the test must be interrupted every 50-200 cycles to examine the condition of the finish. The distractions of the typical workday make it difficult to devote undivided attention during this period, often resulting in missed endpoints and a record of endpoints across a six sample series which all represent different degrees of failure.

Gakushin Test Variability

The standard deviation of a set of measured values expresses how widely the points vary from the mean. In an ideal test, the individual data points would be clustered closely around the mean resulting in the ability to accurately predict the mean with a small number of measurements.

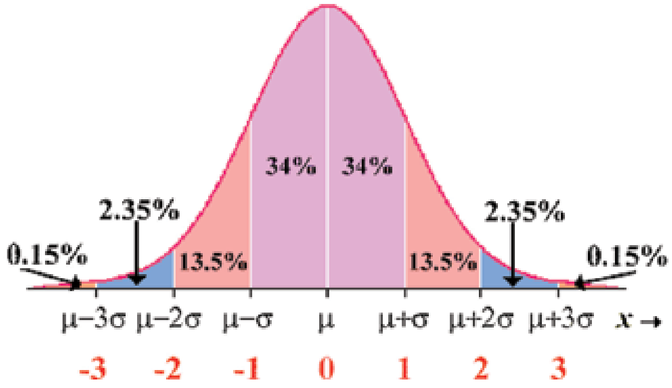


Figure 2. Standard deviation of a normal distribution about the mean[2].

Typically, this is not the case with data produced by the Gakushin test. For example, in a recent set of experiments involving six replicate measurements, the mean was 2824 cycles, while the standard deviation was 1156 cycles. Since 99.7% of measured results will be within plus or minus three standard deviations of the mean when the points are normally distributed (as shown in Figure 2), this determination yields a range of expected results from 0 to 6292. As a consequence of this large window of expected results, many replicates must be tested before subtle differences in performance may be determined. If an insufficient number of samples are tested, one could easily draw the wrong conclusions about sample performance, including ranking samples in the wrong order.

Figure 3 shows the results of a typical pair of samples, including the mean (vertical columns), three standard deviations around the mean (error bars), and the six individual data points that were measured for each sample (small boxes). As may be observed from the graph, the data points show quite a bit of scatter.

Instrument Automation

The literature contains a number of references to conventional material property and performance tests that have been improved through laboratory automation.[3,4] US patent #20050050942, in particular, describes an automated system that improves on the conventional scrub test, a wet abrasive resistance test for wall paint. The scrubbing cycles are controlled and counted by a microprocessor while a digital camera captures images of the samples during the test. An image processing algorithm uses the pictures to determine the extent of wear of each sample and determine the endpoint. This system improves the accuracy of the results by standardizing determination of the endpoint and frees the operator to perform other tasks during the test.

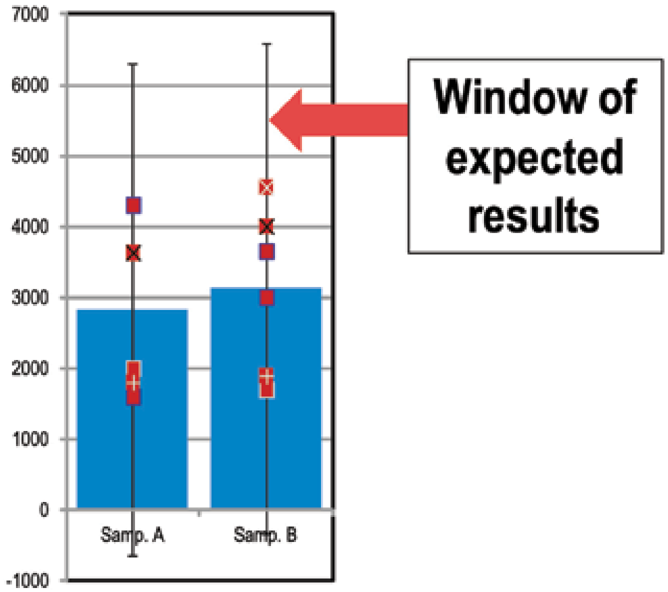


Figure 3. Gakushin cycle counts.

EXPERIMENTAL

Automated Instrument

By incorporating the concepts used to automate the scrub machine, a new Gakushin instrument was developed that was similarly automated. One of the important considerations was that the physical aspects of the abrasion phase of the new test be as similar as possible to that of the conventional test, but with improvements to the endpoint detection phase of the experiment. Maintaining the wearing characteristics of the existing test was important in order to maintain a high correlation between the existing and new tests. Ideally, the new test would yield the same average result as the old test regardless of whether the sample required a low, moderate, or high number of cycles to reach failure. Establishing this type of correlation is important so that results on the new instrument can easily be translated to performance in a test recognized by the industry.

The fully functional automated Gakushin instrument, shown in Figure 4, has been named VERALAB (Vision Enabled Robotic Abrasion LABoratory). As described above, and shown below in Figure 5, the form and function of the conventional wearing mechanism has been preserved as much as possible, including the shape and arrangement of the curved platen, weighted sample heads, and arms.

Most notable in Figure 6 is the imaging system revealed underneath the retracted platen. A motorized drive controls the position of an attached digital camera, which, after the platen is withdrawn, moves across the samples and takes close-up pictures of each leather swatch. While the camera is activated, an eight inch square light bar is energized, providing virtually glare-free light to properly illuminate the samples.

Behind the door on the left side of the cabinet is the programmable logic controller, which coordinates both the wearing and imaging functions of the instrument. An on-board Ethernet connection transmits the digital images to the image processing program running on a nearby computer.

VERALAB User Interface

The user interface for the system is a custom software package that controls the automation process from the computer, including equipment control and data/image acquisition. All of the required experimental parameters are entered on this interface, such as total number of cycles desired and number of cycles between image captures. Image analysis is also controlled by the user interface program, which uses a thresholding technique to detect the difference between the color of the intact finish versus exposed areas of russet (See Figure 7). Image pixels which are flagged for color change by the thresholding step are counted as worn through. Pixels which are not flagged are assumed to represent intact finish. By dividing the number of flagged pixels by the number of unflagged ones, a wear ratio is calculated which is used by the software to calculate the endpoint. The true endpoint does not occur with the first flagged pixel, since dust and other artifacts result in pixels detected by the thresholding step. The software-determined endpoint can later be adjusted by the user based on an examination of the stack of saved images.

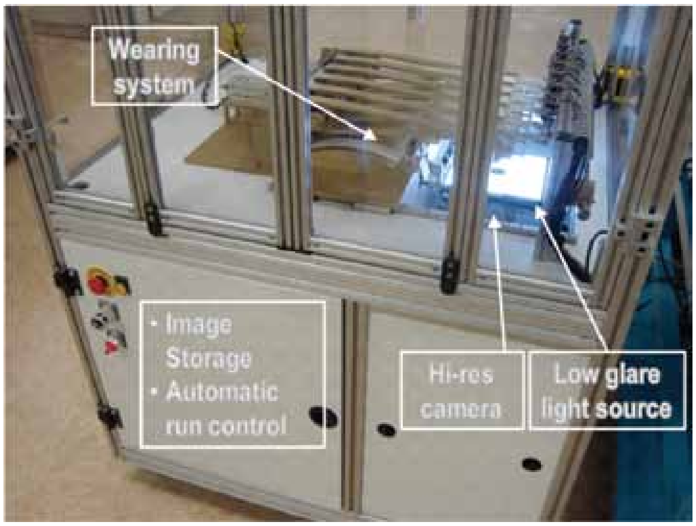


Figure 4. Automated Gakushin.

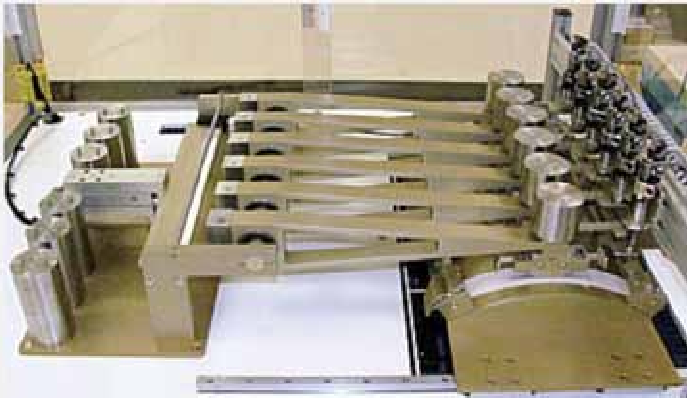


Figure 5. Wearing hardware.

VERALAB Image Player

After a run is completed, the operator is able to view all of the images captured for each sample in chronological order (See example images in Figure 8). By advancing or rewinding the sequence of images, the exact moment of failure can be determined reliably, without fear of missing the endpoint or selecting it prematurely. Since the pictures showing the point of coating failure can be compared across the samples in the set, a high level of consistency in selecting the endpoint can be ensured. Additionally, the image player displays a graph showing the wear ratio as a function of elapsed cycles, the value used by the software to determine the endpoint (See Figure 9). The images captured from the endpoint to the end of the test and the plot of the wear ratio offer the possibility of

Figure 6. Imaging hardware.

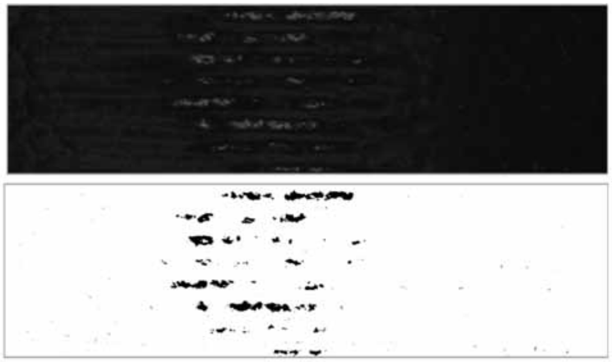


Figure 7. Picture of worn sample before and after thresholding.

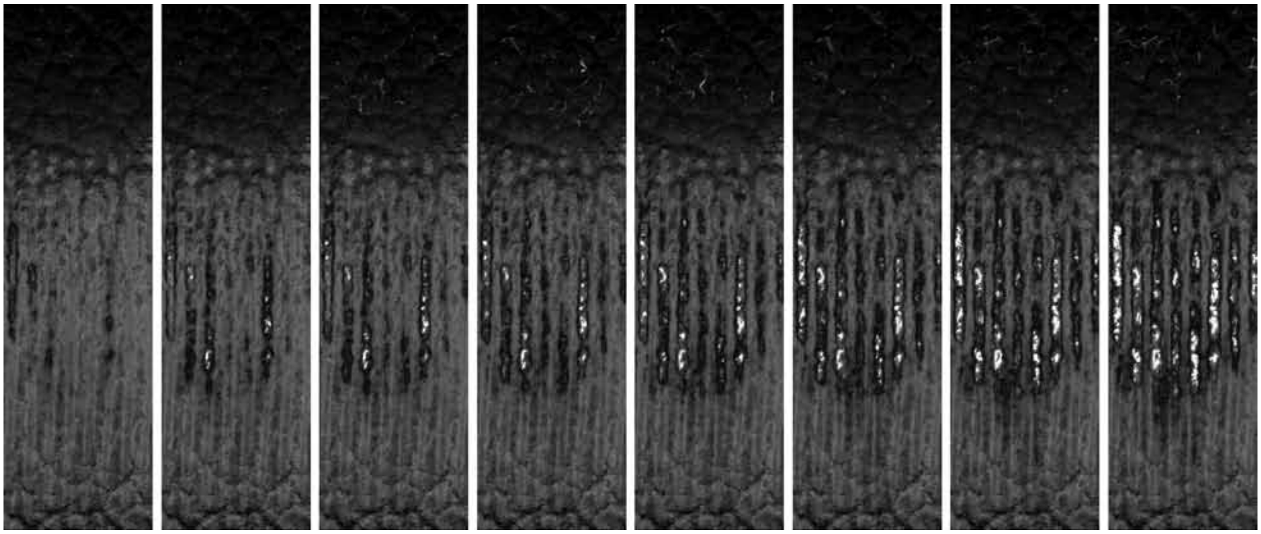


Figure 8. Example of progression of wear with elapsed cycles.

Figure 9. Image player graph displaying wear ratio.

determining the failure mode (delamination, abrasion, intercoat adhesion) of each sample. Knowledge of the failure mode offers an opportunity to more closely tie test results to long-term product performance and develop improved products more quickly and easily.

Leather Finishes

In order to test the performance of the new system, a set of finished leathers were prepared. Different levels of wear resistance were conferred to the samples by the selection of three different binders and two silicone additives incorporated into the topcoats. The six topcoats were applied to sections of a single hide. Six replicate Gakushin samples representing each topcoat were tested on the conventional and automated unit for a total of 72 measurements. The purpose of the experiment was to compare the range of results across the experiment from one instrument to the other. The hypothesis upon which the experiment was based was that the two instruments were performing an identical test and would find the same performance differences from sample to sample. The details related to this experiment follow.

Crust Preparation

For the topcoat testing of leather samples detailed in this article, black colored full grain russet was used after base- and color-coat application and embossing. The base- and color-coats used and embossing procedures employed were typical of those used for automotive upholstery.

Topcoat

The topcoats were prepared using two binder systems, A and B. A third combination of A + B was also employed. 1st and 2nd generation silicone additives were added to the topcoats containing the three binder variations to give a total of six topcoats. The formulations are shown in Table I.

RESULTS AND DISCUSSION

Gakushin Testing

Approximately one week after application, the samples were tested in parallel on the two machines. The results, including average cycles and standard deviation, are presented below in Table II and Table III. Testing of this magnitude, even if conducted on only a single machine, is a significant investment of time since the test must be closely monitored to detect the first instance of coating failure in each sample. At the same time, due to the experimental variability demonstrated by the large standard deviations, many replicate measurements were required to detect real differences in performance. Presented in table form, it is difficult to visualize the effect of experimental variables on performance and the significance of the standard deviation relative to the number of cycles.

Figure 10 and Figure 11 show the results in graphical form and present a more intuitive way to understand the differences in performance. The graphs show that the 2nd generation silicone additive consistently performs better than the 1st

TABLE I
Topcoat formulations.

	Top 1	Top 2	Top 3	Top 4	Top 5	Top 6
Water	150	150	150	150	150	150
Binder System A	-	180	-	-	180	-
Binder System B	200	-	230	200	-	230
Bound Duller A	400	400	-	400	400	-
Bound Duller B	-	-	410	-	-	410
Generation 1 Silicone Additive	70	70	70	-	-	-
Generation 2 Silicone Additive	-	-	-	70	70	70
Flow Agent	10	10	10	10	10	10
Black Colorant	5	5	5	5	5	5
Thickener	x	x	x	x	x	X
50% Isocyanate	150	150	150	150	150	150

TABLE II
Conventional Gakushin results by binder and silicone.

	Binder A/B		Binder A		Binder B	
	Gen. 1 Silicone	Gen. 2 Silicone	Gen. 1 Silicone	Gen. 2 Silicone	Gen. 1 Silicone	Gen. 2 Silicone
Run #	Top 1	Top 4	Top 2	Top 5	Top 3	Top 6
1	4303	3000	900	4303	3000	1696
2	3614	4552	2660	3940	2660	2660
3	3629	4000	3629	4000	2000	2643
4	2000	1697	1697	2000	1697	2499
5	1600	3651	2400	4800	1200	2400
6	1800	1892	3000	3200	1800	1307
Mean	2824	3132	2381	3707	2060	2201
S.D.	1156	1154	967	986	661	564

TABLE III
VERALAB results by binder and silicon.

	Binder A/B		Binder A		Binder B	
	Gen. 1 Silicone	Gen. 2 Silicone	Gen. 1 Silicone	Gen. 2 Silicone	Gen. 1 Silicone	Gen. 2 Silicone
Run #	**Top 1**	**Top 4**	**Top 2**	**Top 5**	**Top 3**	**Top 6**
1	1300	1300	2800	1800	900	1400
2	2100	1600	3000	2800	1500	1300
3	2300	2000	2500	2500	1200	2200
4	1200	4100	1200	2800	1300	1000
5	900	2800	1600	1700	800	1300
6	1400	3100	1800	5600	900	1100
Mean	1533	2483	2150	2867	1100	1383
S.D.	547	1050	720	1422	276	426

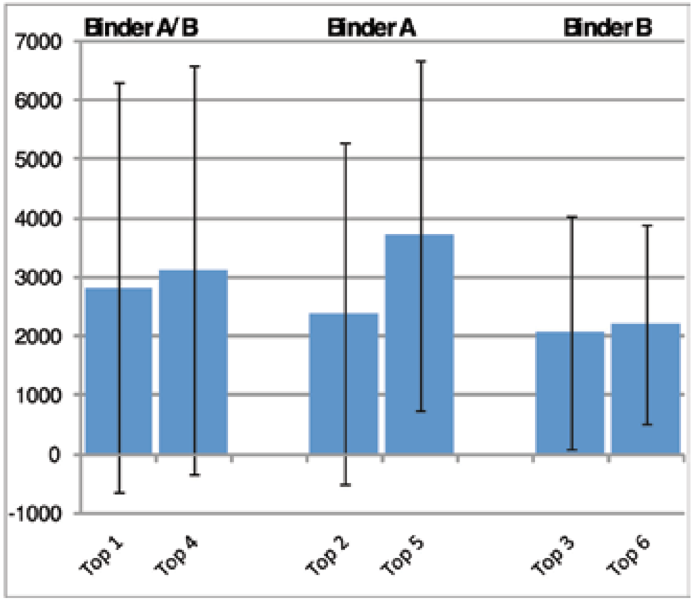


Figure 10. Conventional results by binder and silicone.

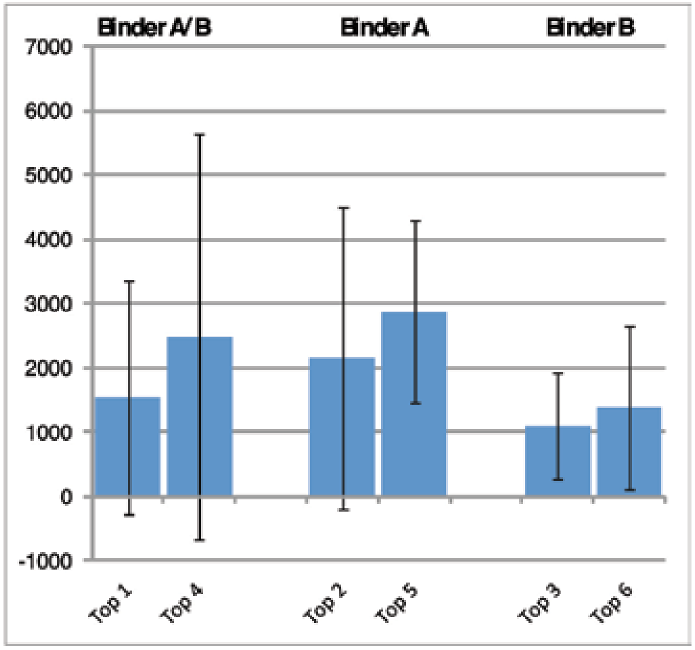


Figure 11. VERALAB results by binder and silicone.

generation silicone, and Binder B underperforms the other binders. Both instruments show the same trends, although the VERALAB tests were completed at fewer elapsed cycles. The error bars represent +/- 3 standard deviations and make it apparent that there is a large amount of overlap in the performance of the samples. The considerable variability calls into question the significance of the observed performance differences and necessitates a more sophisticated data analysis using JMP statistical software.

JMP Analysis

The data displayed above were analyzed with JMP Pro statistical software, version 9.0.3. This was done to determine the statistical validity of the observed trends in performance described above in light of the large standard deviations associated with the data. The specific factor to be studied was the change in wear performance based on choice of binder or silicone. A second factor of interest was the correlation of results between the conventional and automated test methods.

Validation of Model

Using the JMP software, a two-way linear model was constructed studying the factors of binder and silicone selection which generated an F-ratio of 5.76 and a P-value of 0.0002 (Figure 12). This P-value indicates that there is less that a 0.02% chance that such a large F-ratio would be obtained by chance, rather than by actual variations from manipulating the experimental factors. The F-ratio indicates that the model contains almost six times more explained variance than unexplained variance. These measurements provide a high confidence that the model is accurately representing the data.

Validation of Effects

Having validated the overall model, the individual factors were also studied to determine their statistical significance. The model effects of binder and silicone were shown to have very high significance, with large F-ratios and P-values of less than 1%, as shown in Figure 13. These statistics confer a high level of confidence to the model's predictions with reference to the effects of binder and silicone type.

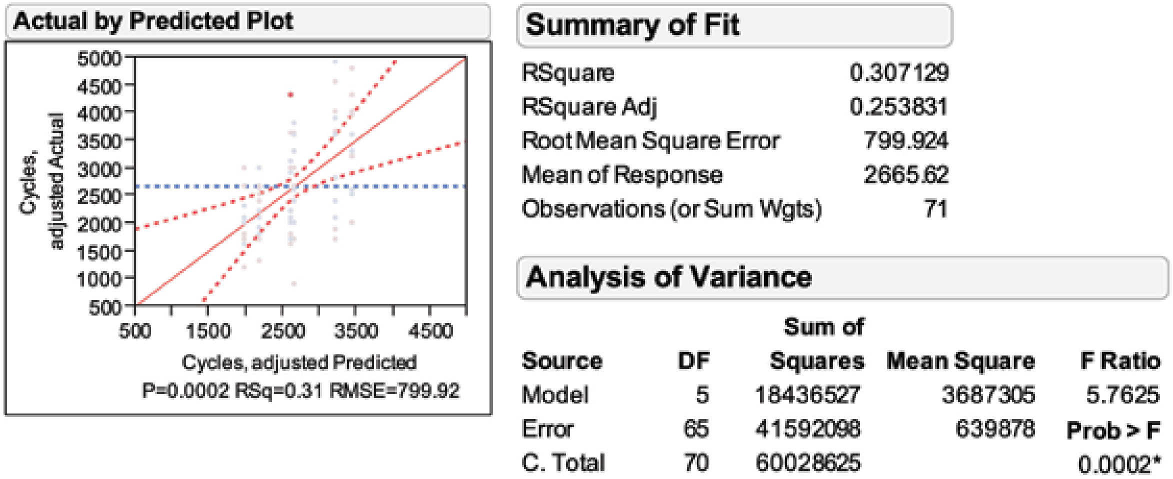
Thus, the model had little difficulty in distinguishing performance differences as the formulation was changed, despite the high variability observed in the data points.

Coating Performance

The least squares means plots highlight the lower performance of Binder B and the higher performance of the 2nd generation silicone. In these formulations, the 2nd generation silicone has an average wear performance advantage of 27%. These data are shown in Figure 14.

Comparison of Abrasion Method

Having confirmed the validity of the model and effects, the final objective of this study was to determine if the new automated method yields the same results as the traditional method. An unexplained observation was that the automated test consistently reached the endpoint in approximately 30% fewer cycles that the conventional instrument. After that difference is added back (as a scaled offset) to the VERALAB results, the endpoints agree quite well across the different variables tested, as shown in Figure 15, below. An additional observation is the reduced variability in the VERALAB data,



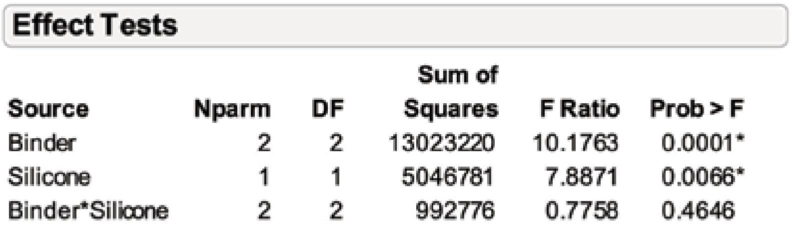
Summary of Fit

RSquare	0.307129
RSquare Adj	0.253831
Root Mean Square Error	799.924
Mean of Response	2665.62
Observations (or Sum Wgts)	71

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	18436527	3687305	5.7625
Error	65	41592098	639878	Prob > F
C. Total	70	60028625		0.0002*

Figure 12. Statistical analysis of JMP experimental model.



Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Binder	2	2	13023220	10.1763	0.0001*
Silicone	1	1	5046781	7.8871	0.0066*
Binder*Silicone	2	2	992776	0.7758	0.4646

Figure 13. Statistical analysis of individual JMP model effects.

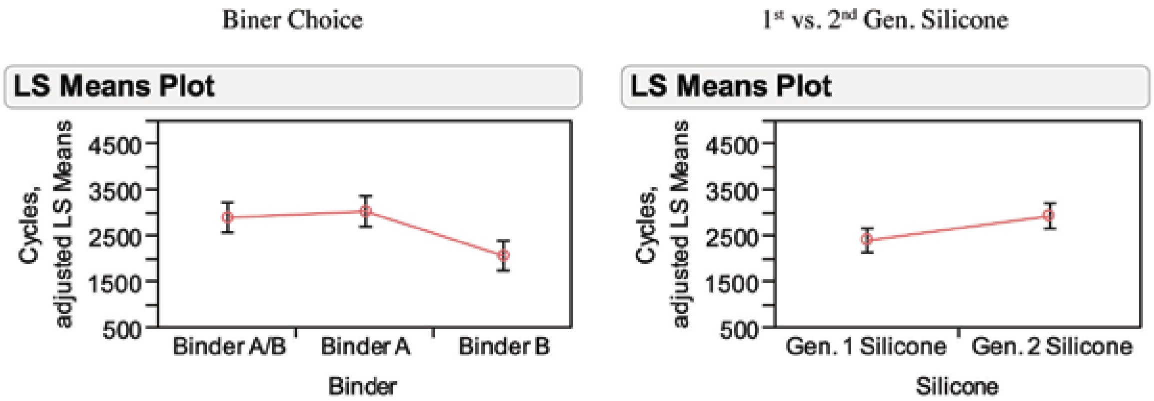


Figure 14. LS means plots showing effect of variables.

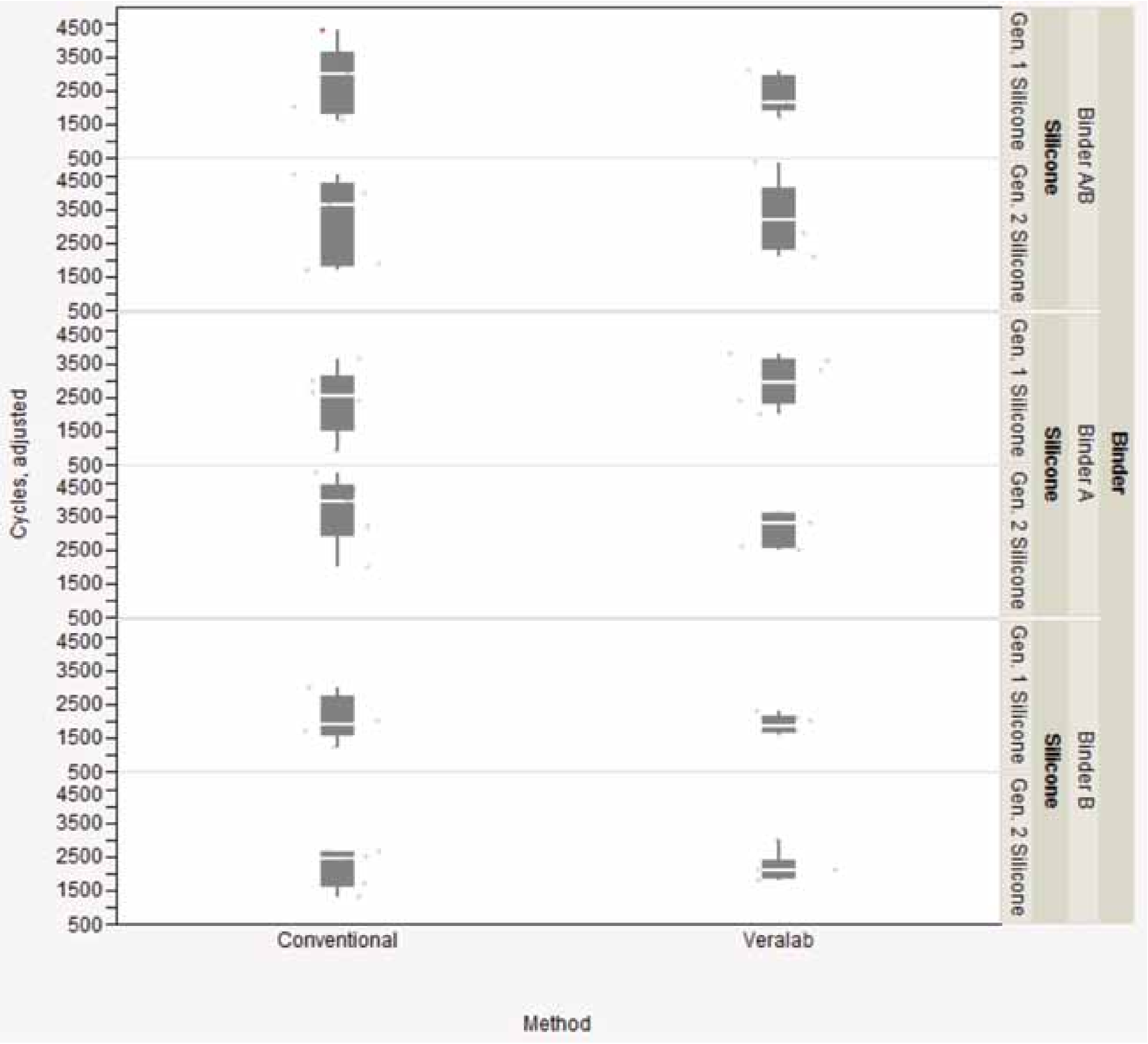
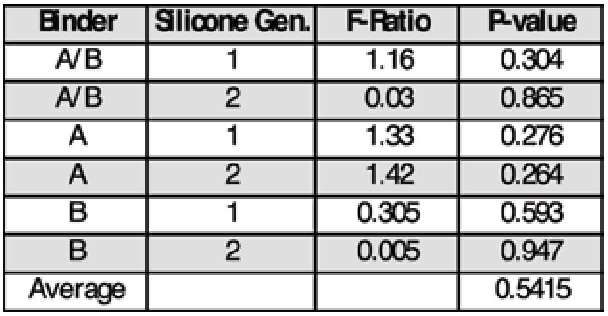


Figure 15. Scaled average cycles as a function of test method.

likely due to more consistency in endpoint detection. This is illustrated in Figure 15 by the smaller vertical height of the boxes in the right hand (VERALAB) column. The boxes signify the spread between the first and third data quartiles.

A one-way analysis of variance was performed for each combination of binder and silicone to look for statistically significant differences between the conventional and automated tests, where a P-value of less than 0.05 indicates a difference. The average P-value over the six variations was 0.54 and the smallest P-value obtained was 0.27, indicating that for this experiment, there is no significant difference between the two instruments for any of the variables studied. The numerical results of these analyses are shown in Figure 16.



Binder	Silicone Gen.	F-Ratio	P-value
A/B	1	1.16	0.304
A/B	2	0.03	0.865
A	1	1.33	0.276
A	2	1.42	0.264
B	1	0.305	0.593
B	2	0.005	0.947
Average			0.5415

Figure 16. Method-to-method comparisons: conventional test vs. VERALAB.

CONCLUSIONS

Recognizing both the advantages and limitations of conventional Gakushin abrasion testing, the authors undertook the development an automated system that, relative to the existing test, would provide:

1. Average performance data that is statistically indistinguishable from conventionally obtained data

2. Less variability in wear results

3. Greatly reduced operator involvement

4. Greater confidence in test endpoints.

The authors have demonstrated, through statistical analysis, an automated system that delivers the sought-after characteristics. This greatly improved test instrument represents a powerful tool in the hands of the leather technologist seeking to develop upholstery topcoats with improved abrasive wear.

ACKNOWLEDGEMENTS

The authors would like to extend their thanks to Mr. Richard Cummins for contributing his statistical expertise. We also acknowledge the contributions of Dr. Rebecca Smith, Dr. Andrew Zalusky, Mr. Michael Franken, and Mr. Hans-Peter Bergelt to this project.

REFERENCES

1. Casey, P., et al.; Comparative Study of Test Methods for Measuring the Wear and Tear (or Wear) Resistance of Leathers for the Automotive industry. Leather Technology Magazine **73**, 8-13, 2010.
2. www.oswego.edu/~srp/stats/6895997.htm. Accessed on 5/6/2012
3. US patent no. 7997118
4. US patent no. 6736017