R_f Method of Reciprocity Performance Testing – A Statistical Approach to Reducing the Time Needed to Reach Accurate Conclusions – Part I

Adam R. Bush, Brian Lindstrom, Joseph LaBarca, Douglas Bugner, Jon Kapecki; Eastman Kodak Company; Rochester, NY/USA

Abstract

Modern image stability testing typically involves the isolation of each environmental variable believed to impact image permanence and the measurement and quantification of the effects produced by exposure to this variable over time. While natural aging at "real-world" levels of these variables is considered the only certain test for image permanence, the high stability of modern photographic products makes testing under ambient conditions too lengthy for new product development cycles. Thus, a widely used alternative to natural aging is accelerated aging, where the impacting variables are held at levels considerably greater than ambient, forcing the image to a failure point in a far shorter length of time. This concept relies heavily upon the law of reciprocity, originally proposed in 1862 by Bunsen and Roscoe. Unfortunately, as is often the case with traditional and digital media, this law only applies over a limited exposure range and can fail as the difference between high and low intensity becomes increasingly large. Provided that the extent of this failure can be quantified and accounted for, meaningful conclusions of product permanence under ambient conditions can be drawn from accelerated testing in a relatively short period of time. This paper will explore a new method of testing for reciprocity performance, with the intent of minimizing the time required for test completion, while improving the accuracy of the results and predictions.

Introduction

The long-term stability of photographs has long been of interest in the field of imaging and photography. The stability of imaging products can be affected in many different ways including, but not limited to, discoloring and yellowing of substrates, brittle and cracked supports, loss of sharpness, and dye fade. With all of these effects ultimately leading to product performance falling short of customer expectations, the need for efficient testing, evaluation, and a means for potential improvement, is evident.

There are four main environmental variables known to impact image permanence: light, heat, moisture, and air pollutants, such as ozone. Modern image stability testing typically involves the isolation of each variable, measuring and quantifying the effects produced by each variable as a function of exposure intensity and time. While the only error-proof method for testing image permanence is natural aging under "real-world" levels of these variables, the high stability of modern photographic products makes testing under ambient conditions too lengthy a process for the majority of practical uses. Thus, a widely used alternative to natural aging is accelerated aging, where the impacting variables are held at levels considerably greater than ambient, forcing the image to a failure point in a far shorter length of time. In the case of accelerated light fade testing, this approach relies heavily on the assumption that product behavior, when exposed to highly elevated levels of a given variable for short times, is equal to product behavior when exposed to an ambient or slightly accelerated level of that variable for longer times, a relationship defined by Bunsen and Roscoe in 1862, known as the Reciprocity Law.

This paper will explore a new method of testing for reciprocity law adherence or failure, with the intent of minimizing the time required for test completion, while improving the accuracy of the results and predictions. By combining proper application of the reciprocity law with appropriate statistical methodologies, meaningful conclusions of product longevity under ambient conditions can be drawn from accelerated testing in a relatively short period of time.

Background

The reciprocity law, originally proposed to describe light-induced chemical reactions, states that the product of a photochemical reaction is determined simply by the total exposure, that is, by the product of irradiance and time, and is independent of the two factors separately [1]. In the context of light-induced fade of photographs, the reciprocity law can be described such that the change in density (ΔD), resulting from high-intensity illumination (HI) for a short period of time, is equal to the change in density that results from low-intensity illumination (LI) for a long period of time, given that the two conditions yield an equal amount of cumulative exposure, as defined by the product of intensity and time. This relationship can be expressed mathematically as:

$$\Delta D_{LI} = \Delta D_{HI} \tag{1}$$

(for a given cumulative exposure)

In the definition above, the more general term "intensity" can be substituted for "irradiance," which allows for a more broad application of this law to reactions resulting from exposure to other degradation factors as well [2].

Unfortunately, as is often the case with traditional and digital media, this law applies only over a limited exposure range, and can fail as the difference between HI and LI becomes increasingly large. This is known as reciprocity failure. Although the existence of reciprocity failure in an accelerated test initially compromises the assumption stated above, it does not necessarily mean that this particular test has lost merit, provided that the extent of the failure can be understood, quantified, and accounted for. What does becomes obvious, however, is that high-intensity, highly accelerated testing, alone, is not sufficient in understanding product performance, and some form of low-intensity check for reciprocity failure must accompany it.

Completion of a LI reciprocity test to an endpoint, for example, a 30% loss of density in a primary colorant, can be very time consuming and, in order to reach a conclusion in a shorter length of time, an experimenter might be tempted to either skip the LI test or to extrapolate a result from an abbreviated set of data. It is currently recommended that if reciprocity failure exists, the LI data should be used for making product lifetime predictions because it is closer to ambient conditions and considered more representative of the "real world" environment [3]. In this scenario, it is easy to speculate that product longevity claims and comparisons will often need to be made using values that have been extrapolated from LI data.

It is important to note that while extrapolation is a useful technique in many areas of applied regression analysis, the level of statistical confidence and certainty often resulting from its use can prove to be unacceptably low and provides a final conclusion with little value. A complete analysis of variance (ANOVA) must always be carefully considered. In general, the greater the level of extrapolation of a set of data, the lower the level of statistical confidence in a value predicted from that extrapolation. This relationship between confidence and extrapolation is precisely defined in any textbook on applied regression analysis [4]. As an example, a reciprocity test that is run only to a 15% density change, with results calculated from an extrapolation of these data to a 30% density change, may yield a longevity prediction of 75 years. However, after a complete analysis of variance, it may be found that there is significant lack of fit in the model and, at 95% confidence, the only certainty is that the true longevity of this product falls somewhere between 40 and 110 years. Obviously, a prediction with a confidence interval of this magnitude has questionable value for making useful product longevity claims and comparisons.

Methodology

Reciprocity Factor (R_f)

Let us consider, in contrast, a different technique. As stated above, the existence of reciprocity failure in an accelerated test does not mean that the test has lost merit. We simply must quantify the extent of the failure and adjust the HI data to account for it. Examining Equation 1 above, one could easily assert that if reciprocity failure exists between an LI and HI condition, the extent of that failure for any given cumulative exposure is equal to the constant that relates the two. Since image change resulting from treatment under LI test conditions is considered more representative of change occurring under ambient conditions (less acceleration), reciprocity failure is attributed to the HI condition, and can be expressed as follows:

$$\Delta D_{LI} = (R_f)\Delta D_{HI} \tag{2}$$

(for a given cumulative exposure)

Here, R_f represents the constant factor by which the ΔD_{LI} and the ΔD_{HI} data are related -- the reciprocity factor.

Rearranging this equation to solve for R_f, we get:

$$R_f = \frac{\Delta D_{LI}}{\Delta D_{HI}} \tag{3}$$

(for a given cumulative exposure)

Observe here that R_f will assume a value of 1.0, if the LI and HI test conditions yield an equal change in density. Alternatively, R_f will assume a value greater than 1.0, if the LI condition yields more density change than the HI condition. Finally, R_f will assume a value less than 1.0, if the HI condition yields more density change than the LI condition. Each scenario has been observed in image permanence testing, and each scenario should be considered a possible outcome in testing [5].

Now that R_f is defined for a given cumulative exposure, we examine how it behaves throughout the entire practical range of exposure. Note that at zero cumulative exposure there will be no change in density in either the LI or HI conditions, and reciprocity failure cannot exist. Inevitably, some finite exposure will yield complete density loss in both the LI and HI conditions, and R_f will assume a value of one. Neither case is useful to test, however, because there are no alternative outcomes. A reasonable range to explore might be from the portion of the data that begins with the initial sample exposure to the portion of the data that yields a ΔD equal to the endpoint at which longevity predictions are to be made.

In the early stages of accelerated testing, small changes in density are not detectable from within normal test system variability, and any calculation of R_f would be significantly compromised by noise. However, once a signal above noise is detected in both the LI and HI conditions, R_f should quickly assume its true value; and in a "well-behaved" test, it would remain constant throughout the entire practical range of exposure. A well-behaved test implies a number of constraints but, most importantly, it requires that the test variable under experimentation be truly isolated. With this requirement satisfied, there will be only one mechanism of degradation acting upon the image, at least from a practical point of view, and the rate at which each condition progresses relative to one another should remain constant.

To further develop this point, failure to completely isolate the specific environmental factor being tested may result in apparent reciprocity failure, most notably in the LI test, where ambient laboratory levels of some controlled or uncontrolled variable contribute a long-term degradative effect on a given test product. In this scenario, extended exposure to this additional secondary factor would give the appearance of reciprocity failure when, in fact, it is not occurring. For example, it has been shown that for light-induced fade of inkjet prints, many early reports of gross reciprocity failure were, in fact, caused by tests that were confounded by the presence of ambient ozone in the test chamber, which resulted in proportionately more loss of density being observed during the longer tests run at the lower intensity condition [6].

In cases like this, where the factors cannot be completely separated, a constant R_f likely would not be achieved. However, the R_f method described in this paper could still provide a more accurate result than current LI methodologies. Current methodologies require that the LI reciprocity test be run fully to an endpoint. Here, the secondary factor would likely have a longer term and a potentially much larger cumulative effect on density change. In contrast, as a result of the relatively brief test period that the R_f method encompasses, this false reciprocity failure signal could be largely avoided. The key to success would be in identifying R_f in as short a timeframe as possible. This minimizes the amount of time that a confounding factor, if present, can contribute to undesirable image degradation, which allows for a more accurate assessment of reciprocity law failure or adherence [6].

Any test anomalies may void the ability of this method to give an accurate measure of reciprocity; therefore, it is important to recognize other characteristics of a well-behaved test. Beyond the isolated variable requirement just stated, the high-intensity fade curve must take some expected mathematical form, usually either linear or logarithmic, and have a sufficiently high measure of fit to that function. R^2 is usually a good initial indicator of this fit, with a value of 0.95 or higher expected. Also, both high- and low-intensity conditions must produce a detectable signal from total system noise, otherwise a calculation of R_f will have no meaning. Typical test system noise with modern image permanence equipment is roughly 3–5% of any measured value [7], therefore a reliable calculation of R_f might not be expected until at least approximately a 10% change in density.

Data and Discussion

We begin with two sets of sample data that were taken directly from testing done at Kodak's Image Stability Technical Center over a seven-month period in 2004 and 2005. One set represents the ΔD in a test sample after 80 klux HI fluorescent light treatment over 112 days, and the other represents ΔD in an identical sample after 5.4 klux LI fluorescent light treatment over 224 days. See Figure 1.

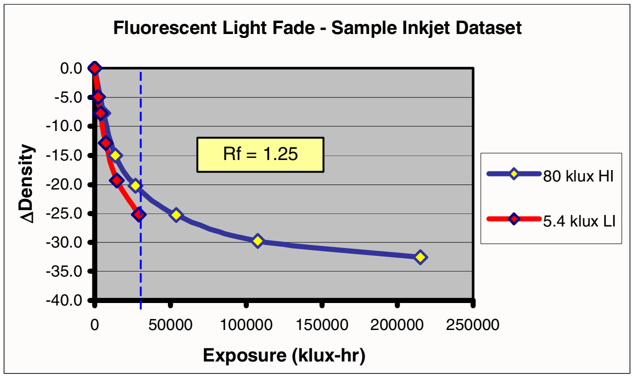

Figure 1. *HI and LI fluorescent light-fade data. R_f equals 1.25, calculated at 29030 klux-hr cumulative exposure.*

To calculate R_f, we pick a cumulative exposure common to both LI and HI test conditions (represented by the vertical line drawn in Figure 1) and calculate the ratio of the change in density at each condition, following the example of Equation 3.

$$R_f = \frac{\Delta D_{LI}}{\Delta D_{HI}} = \frac{-25.2}{-20.2} = 1.25 \qquad (4)$$

(at 29,030 klux - hr cumulative exposure)

It is important to use experimental data to perform this calculation as opposed to values interpolated from a fitted model. Model fitting introduces an additional source of error and uncertainty into the analysis that, at this point, will only confound the precise calculation of R_f. The smoothing that occurs when fitting a curve gives the false appearance that R_f is still changing, when in fact, it has truly leveled. Using true experimental data, we are subject only to test system noise and, with proper sampling and replication, can quickly and accurately identify R_f.

Using the data from Figure 1, we can plot a trend line of R_f as a function of cumulative exposure to explore its behavior throughout the entire practical range of exposure. This can be seen in Figure 2.

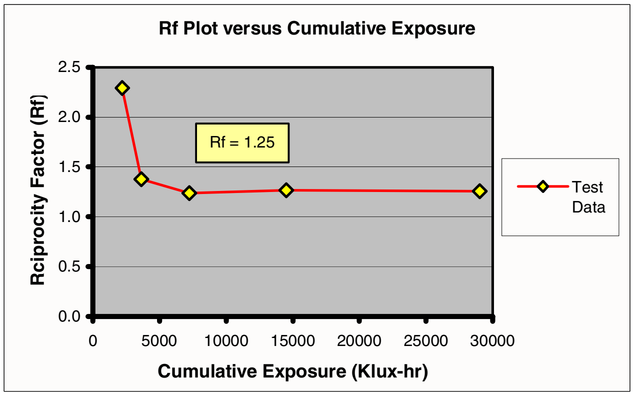

Figure 2. *R_f is plotted against cumulative exposure demonstrating the assertion that over a practical range of exposure the rates of change between two reciprocity conditions are constant.*

This plot offers visual confirmation that once a ΔD signal can be detected above the system noise, R_f will quickly assume its true value and remain constant throughout the entire practical range of exposure. Once a steady R_f has been demonstrated, that value can be used to "correct" the HI data, accounting for the exact amount of reciprocity failure exhibited in the test. A model can now be fit to the corrected data, and an interpolation made at some chosen endpoint. This method of correction and interpolation is shown in Figure 3.

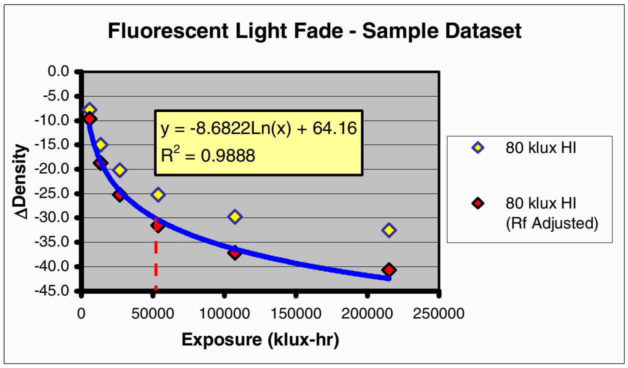

Figure 3. *Here a 1.25 reciprocity factor (R_f) is applied to each value in the 80 klux HI data set, correcting each value for the observed reciprocity failure. A 30% ΔD endpoint is reached at approximately 50000 klux-hr of cumulative exposure.*

Statistical Error Discussion

It is important to note that the two sets of data used in the example above were not specifically designed to test the premise of the R_f method that is outlined in this paper, and hence, may not best illustrate the power this method offers. Considering Figure 2, it is easy to see that an increased level of sampling and further replication in the critical exposure range, i.e., where signals begin to resolve from noise, may have allowed a determination of R_f in even less time. Five exposure samplings were measured and evaluated for R_f in this example. However, the two final readings, representing more than half of the measured exposure range, provided little additional information. If those two measurements had been made prior to 10000 klux-hr, more pertinent information could have been included in the analysis. It is obvious that the full 224 days of LI testing were not necessary to reach a conclusion here, and the test could have been stopped after only a fraction of that time. It appears that R_f may actually reach constant value at approximately 5000 klux-hr (39 days), and a more appropriate data sampling frequency could resolve this abbreviated time span. It still must be demonstrated that R_f positively has settled prior to ending the LI test, but in the example provided, we can estimate that 8000 klux-hr (60 days) would be adequate.

Further, since the accuracy in calculating R_f is based solely on the level of test system noise present in generating the respective LI and HI data points, it is important to note that sample replication will be very useful in maximizing the accuracy of the result. In general, replication will drive down system noise by a factor of the square root of the number of replicates [8]. For example, we shall assume that the variability of a given test system is 5% of any measured value. If quadruplicate samples were tested at each condition, variability (noise) would be reduced in half, following the equation:

$$TestSystemNoise = \frac{0.05}{\sqrt{4}} = 0.025. \tag{5}$$

This would make a substantial difference in both the confidence and accuracy of the R_f prediction.

Conclusions

In the example given in Figure 2, the first indication of the true R_f can be calculated at approximately 7250 klux-hr of exposure and 10% dye fade. At 5.4 klux intensity, 7250 klux-hr represents 56 days of testing. Compared with current reciprocity testing, where a test may take from six months to a year or longer to complete, often relying on extrapolation to reach a conclusion, 56 days is a vast improvement. By eliminating the extended length of time currently required for LI testing where apparent reciprocity failure can confound results, this method may, indeed, be more accurate.

In light of the observations made in completing the analysis for this paper, the authors offer this discussion as a Part I in a continuing study to more fully investigate the implications and potential of this method. Further experiments will explore the behavior of R_f throughout the full length of a test and at that critical exposure range where signals begin to rise above system noise. While the time required to complete a test of this nature made it impractical to have the analysis completed for this preliminary publication, an analysis of existing data provided excellent agreement with the premises proposed for this method. It is worth stating that, although only one example system was shown here, many systems were analyzed, covering most major printing technologies, all showing similar results, given the constraints of a well-behaved test.

References

[1] The Theory of the Photographic Process, 4th Edition, T.H. James, Ed. (MacMillan, NY, 1977) pg.133.

[2] D.Bugner, R. Van Hanehem, M. Oakland, P. Artz, D. Zaccour, R. Levesque, "Ozone Concentration Effects on the Dark Fade of Ink Jet Photographic Prints," J. Imag. Technol., 49: 317–325 (2005).

[3] S. Guo, N. Miller, Estimating Light-fastness of Inkjet Images: Accounting for Reciprocity Failure, Proc. NIP17, pp. 186–191 (2001).

[4] N.R. Draper and H. Smith, Applied Regression Analysis, 3rd Edition (Wiley, NY, 1998), pg. 83.

[5] J. Smitley, B. Thomas, T. Silliman, and T. Izyk, Sensitivities of Kodak's Thermal Dye Transfer Prints to Environmental Factors, Proc. NIP21, pp. 339–243 (2005).

[6] D.E. Bugner, D. Kopperl, and P. Artz, Proc.12th Int. Symp. on Photofinishing Technology, pp. 54–57 (2002).

[7] D. F. Kopperl, R.J. Anderson, R. Codori, and N. Rollo, "Quality Improvements and Control Procedures at the Image Stability Technical Center," J. Imag. Technol., 16, 5 (1990), pp. 198–202.

[8] T.B. Barker, Quality by Experimental Design (Marcel Dekker, NY, 1985) pg. 139.

Author Biography

Adam Bush received his BS in Chemistry/Biology from the Rochester Institute of Technology. After working as an analytical chemist and product development engineer at Xerox Corp. in Webster, NY, Adam joined Eastman Kodak Company in 2000 as a photoscience engineer. He currently supports the Image Stability Technical Center in Rochester, NY.