

L•PRIZE®

Lumen Maintenance Testing of the Philips 60-Watt Replacement Lamp L Prize Entry

Updated September 2016

Updated: July 2015

Updated: July 2013

Original publication: July 2012

Prepared for:

Solid-State Lighting Program

Building Technologies Office
Office of Energy Efficiency and
Renewable Energy
U.S. Department of Energy

Prepared by:

Pacific Northwest National
Laboratory

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY

operated by

BATTELLE

for the

UNITED STATES DEPARTMENT OF ENERGY

under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,

P.O. Box 62, Oak Ridge, TN 37831-0062;

ph: (865) 576-8401

fax: (865) 576-5728

email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service

5301 Shawnee Rd., Alexandria, VA 22312

ph: (800) 553-NTIS (6847)

email: orders@ntis.gov <<http://www.ntis.gov/about/form.aspx>>

Online ordering: <http://www.ntis.gov>



This document was printed on recycled paper.

(8/2010)

Lumen Maintenance Testing of the Philips 60-Watt Replacement Lamp L Prize Entry

KL Gordon
RP Hafen
JE Hathaway
JJ McCullough

Original publication: July 2012
Updated: July 2013
Updated: July 2015
Updated: September 2016

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

Executive Summary

The Pacific Northwest National Laboratory, operated by Battelle for the U.S. Department of Energy, developed a procedure and facility for testing the lumen maintenance of light-emitting diode (LED) replacement lamps submitted to the L Prize[®] competition in the 60-watt replacement lamp category. The Accelerated Long-term Test Apparatus (ALTA)^{*}, in which lamps are operated continuously in a 45°C environment, measures the relative lumen and chromaticity maintenance of up to 200 screw-based lamps. Two hundred (200) samples of the 60-watt LED replacement product submitted by Philips Lighting were first measured over a period of 7,000 hours to determine whether the lamps met the minimum requirements of the L Prize competition. Measurements have continued thereafter for the purpose of long-term data collection. Spectroradiometric measurements performed by a mobile integrating sphere took place every 100 hours for the first 5,000 hours of operation and weekly (every 168 hours) thereafter.

The full sample of 200 Philips L Prize lamps were tested for a total of 25,000 hours, a period reached at the end of April 2013. Average lumen maintenance for the lamps at that time was still greater than 100% of average initial output of the lamp samples. The L Prize competition required at least 70% lumen maintenance at 25,000 hours. Based on the data available after the first 7,000 hours of operation, lumen maintenance at 25,000 hours was projected to be 97.1%, and in fact, at the 25,000-hour mark, was 100%. Further, the L Prize competition requirements allowed for a change in chromaticity of no more than .004 as measured on the International Commission on Illumination (CIE) 1976 (u',v') color diagram. The average change in chromaticity of the 200 Philips lamps after the first 7,000 hours of operation was 0.0006, and at 25,000 hours it remained less than .002. Partly on the basis of the measured performance at 7,000 hours, the Philips Lighting product was declared the winner of the L Prize in the 60-watt replacement lamp category on August 3, 2011.

After reaching the milestone of 25,000 hours, the number of L Prize lamps continuing under test was reduced to 31, a quantity sufficient to maintain statistical significance for the product sample. Updated results for 40,000 and 50,000 hours are provided in this document in sections 7.0 and 8.0, respectively. As of August 2016, the 31 lamps passed 50,000 hours of operation, with average lumen maintenance of 93% and average chromaticity change of .002. The results reported here are for the specific product tested as part of the L Prize competition and should not be generalized to other LED lighting products. However, as the L Prize test results represent the first publicly-available data confirming actual performance of a high-quality LED product at 50,000 hours, they can serve as an indicator for the long-term potential for a well-designed and constructed product, and validate the methods being used for extrapolating and predicting long-term performance.

^{*} Formerly called the lumen maintenance test apparatus (LMTA). References to the test apparatus appear as LMTA in Sections 1.0-6.0 of this report and as ALTA in subsequent sections.

Acronyms and Abbreviations

CIE	International Commission on Illumination
DOE	U.S. Department of Energy
EISA	Energy Independence and Security Act of 2007
IES	Illuminating Engineering Society
L Prize	Bright Tomorrow Lighting Prize
LED	light-emitting diode
PLC	programmable logic controller
PNNL	Pacific Northwest National Laboratory
SSL	solid-state lighting

Contents

Executive Summary	iii
Acronyms and Abbreviations	iv
1.0 Introduction	1.1
1.1 Estimating LED Lifetime	1.2
1.2 L Prize Lifetime Metric	1.2
2.0 L Prize Requirements	2.1
3.0 Lumen Maintenance Testing	3.1
3.1 Test Facility	3.1
3.2 Test Apparatus Development	3.1
3.2.1 Systems Integration	3.5
3.2.2 Data Collection	3.6
4.0 Interpreting Results: Statistical Method	4.1
4.1 Non-Parametric Statistical Approach	4.1
4.2 Determining Point of Maximum Flux	4.3
5.0 Test Results for First L Prize Winner	5.1
5.1 Lumen Maintenance	5.1
5.2 Chromaticity Maintenance	5.2
5.3 Updated Interim Results Since L Prize Award	5.3
6.0 Update: Results After 25,000 Hours	6.1
6.1 Future Testing and Analysis	6.3
7.0 Update: Results After 40,000 Hours	7.1
8.0 Update: Results After 50,000 Hours	8.1
9.0 References	9.1

Figures

Figure 1.1. DOE SSL Program Approach.....	1.1
Figure 3.1. LMTA Exterior.....	3.2
Figure 3.2. LMTA Interior.....	3.2
Figure 3.3. LMTA with Original Plexiglas Walls.....	3.3
Figure 3.4. Close-up of U-Channels	3.3
Figure 3.5. Integrating Sphere and Computer-Numerical-Control Track System	3.4
Figure 3.6. Diagram of Systems Integration	3.6
Figure 3.7. Data Collection Showing Luminous Flux	3.7
Figure 4.1. Illustration of Non-Parametric Statistical Approach	4.1
Figure 4.2. Measurement Data for All Lamps at 100-Hour Intervals	4.3
Figure 4.3. Residuals from the Loess Fit	4.4
Figure 4.4. Loess Fit with 99% Upper Bound for Mean Trend Function	4.5
Figure 4.5. Frequency Distribution by Run Hours.....	4.5
Figure 4.6. Distribution of Four Possible Maximum Output Periods	4.6
Figure 5.1. Measured Lumen Maintenance	5.1
Figure 5.2. Forecast for 13 th -Worst-Performing Lamp	5.2
Figure 5.3. Change in Chromaticity over Test Period	5.3
Figure 5.4. Average Lumen Maintenance and Sphere Temperature (°C) to 14,000 Hours.....	5.4
Figure 5.5. Range of 25,000-Hour Lumen Maintenance Forecasts	5.4
Figure 5.6. Predictions Between 4,500 and 14,000 Hours.....	5.5
Figure 5.7. Forecasts of 7,000-Hour Data Similar to 14,000-Hour Data.....	5.5
Figure 5.8. Change in Chromaticity up to 14,000 Hours	5.6
Figure 5.9. Average Lumen Maintenance After 18,000 Hours of Operation	5.7
Figure 5.10. Lumen Maintenance After 18,000 Hours of Operation.....	5.8
Figure 5.11. Chromaticity Maintenance After 18,000 Hours of Operation	5.8
Figure 6.1. L Prize Average Lumen Maintenance After 25,000 Hours	6.1
Figure 6.2. Comparison of Best, Worst, 13 th -Worst, and Average L Prize Lumen Maintenance After 25,000 Hours.....	6.2
Figure 6.3. L Prize Average and 13 th -Worst Chromaticity Maintenance Relative to Required Limit.....	6.2
Figure 7.2. Comparison of Best, Worst, and Average L Prize Lumen Maintenance After 40,000 Hours.....	7.2
Figure 7.3. L Prize Best, Worst, and Average Chromaticity Maintenance Relative to Required Limit.....	7.3
Figure 8.1. L Prize Lumen Maintenance After 50,000 Hours	8.1
Figure 8.2. Comparison of Best, Worst, and Average L Prize Lumen Maintenance After 50,000 Hours.....	8.2

Figure 8.3. L Prize Best, Worst, and Average Chromaticity Maintenance Relative to Required Limit After 50,000 Hours.....	8.3
---	-----

1.0 Introduction

The Bright Tomorrow Lighting Prize (L Prize®) was established by Congress in the Energy Independence and Security Act of 2007¹ (EISA). The purpose of the L Prize is to encourage the lighting industry to develop energy-efficient, solid-state lighting (SSL) replacements for common light sources, including the 60-watt incandescent light bulb and the PAR38 halogen incandescent lamp (floodlight). The EISA legislation directed the U.S. Department of Energy (DOE) to hold the L Prize competition. DOE publicly announced the competition in May 2008² and in September 2009 the first entry was received—a sample of 2,000 60-watt replacement lamps submitted by Philips Lighting North America.

These lamps were subjected to a variety of tests over the next 18 months, and the Philips lamp was awarded the first L Prize in August 2011. This paper describes testing conducted to evaluate the product's ability to meet the lifetime/lumen maintenance requirement of the competition, which was: “having 70 percent of the lumen value under subparagraph (A) [producing a luminous flux greater than 900 lumens] exceeding 25,000 hours under typical conditions expected in residential use.” A custom test apparatus was designed and constructed for this testing and a statistical approach was developed for use in evaluating the test results.

The L Prize competition is carried out by DOE as part of a comprehensive SSL research and development and commercialization support program (see Figure 1.1 below). The purpose of the program is to advance the energy efficiency and performance of SSL technology and support successful application of SSL products in the marketplace to realize national energy savings, environmental, and economic goals.



Figure 1.1. DOE SSL Program Approach

¹ PUBLIC LAW 110–140—DEC. 19, 2007. ENERGY INDEPENDENCE AND SECURITY ACT OF 2007. Section 655. Bright Tomorrow Lighting Prizes.

² See <http://www.lightingprize.org/lightfair08.stm> for information about the L Prize launch event at LIGHTFAIR® 2008.

1.1 Estimating LED Lifetime

Determining the lifetime of light-emitting diode (LED) lighting products is not straightforward. LEDs can continue to produce light over periods of time much longer than most other light sources. Barring failure of other components (such as the driver, electrolytic capacitor, or other parts), LED light output typically declines slowly over time. Useful life is usually defined in terms of the number of operating hours until the light output has decreased to a given percentage of initial output, often 70% (denoted as L_{70}), but it may differ depending on the needs of the lighting application.

The challenge is how to verify LED lumen maintenance *without* undertaking full life testing; 25,000 hours of continuous operation takes nearly three years. An industry standard test procedure, Illuminating Engineering Society (IES) LM-80-08, “Approved Method for Measuring Lumen Maintenance of LED Light Sources,” provides a method for LED packages to be tested for at least 6,000 hours, with light output and chromaticity measurements made at least every 1,000 hours. An accompanying document, Technical Memorandum (TM)-21-11, “Projecting Long Term Lumen Maintenance of LED Light Sources,” provides a method for estimating lumen maintenance for a period up to six times the LM-80-08 test period.

At the time the L Prize requirements were developed, LM-80-08 was only recently approved and TM-21-11 did not exist. The scope of LM-80-08 is lumen and chromaticity maintenance of LED packages operated under specific temperature conditions, and testing is not conducted on end-use products such as integral replacement lamps or complete luminaires. Therefore, a methodology for evaluating full-lamp lumen and chromaticity maintenance over time was developed by DOE specifically for the L Prize contest.

1.2 L Prize Lifetime Metric

The incandescent light sources targeted for replacement by the L Prize competition have a much shorter lifetime than LED sources. Standard incandescent 60-watt bulbs, for example, are typically rated for 1,000 hours of use. That rating means that, of a statistically significant sample of lamps, 50% will have failed (burned out) by 1,000 hours of operation. There is no significant lumen depreciation over the lifetime of incandescent lamps, as they fail catastrophically before degradation of the light output becomes noticeable.

Other conventional light sources have known lumen depreciation behavior over their lifetimes. Lamp manufacturers publish initial lumens and “design lumens” for linear fluorescent, compact fluorescent, and high-intensity discharge lamps. Design lumens typically refer to output at 40% of the lamp’s rated average life.

In the L Prize competition, DOE adopted a more rigorous metric. Whereas typical industry standard lifetime ratings as described above for incandescent lamps assume a 50% failure rate (denoted B_{50}), DOE applied a 10% (B_{10}) failure rating, meaning output dropping below 70% of initial output. The sample size for long-term testing was set such that DOE could state with 95% confidence that 90% of the lamps would perform at least as well as the 13th-worst-performing lamp in the sample. This metric is explained in greater detail in the following section.

2.0 L Prize Requirements

The L Prize Competition Requirements document, dated June 26, 2009, listed the following requirement for Product Lifetime – Lumen Maintenance:

Products shall have a lifetime exceeding 25,000 hours defined as $L_{70}/C_{95}/B_{10}$, where: L_{70} refers to 70% lumen maintenance, C_{95} refers to 95% statistical confidence, and B_{10} refers to no more than 10% of the products dropping below 70% of their initial (defined as the maximum) lumen value after 25,000 hours.

The document further specified the following evaluation method:

It is anticipated that lumen and color maintenance testing will be conducted on a sample of 200 lamps, oriented base up in a 45°C environment, randomly drawn from a population of at least 2,000 lamps. Using a non-parametric methodology for establishing confidence intervals, the lower 95%/10% (C_{95}/B_{10}) tolerance limit is the 13th worst performing product, when results are ranked in order of lumen maintenance performance, from lowest to highest. L_{70} is estimated using a best-fit exponential decay function extrapolation, based on data collected from the point in time at which the product reaches its maximum flux (t_{max}) to a point in time 5,000 hours later ($t_{max} + 5,000$).

Compliance with the 25,000-hour color maintenance requirement will be tested in a similar manner. When rank ordered from most color shift to least, the 13th worst performing product must not exhibit color shift of more than .004 (u',v') over the time from t_{max} to $t_{max} + 5,000$.

A statistical approach yielding a high-confidence interval was adopted for several reasons. First, long-term testing of an integral LED lamp replacement product had not been done before by an independent (non-manufacturer) entity. DOE wanted to provide a high level of rigor and confidence in the results of this testing. Second, the L Prize partners who contributed to development of the L Prize program rules encouraged DOE to conduct long-term testing and do so in a highly rigorous, statistically defensible manner. The extra level of confidence in long-term product performance was seen by energy efficiency program sponsors as critical to the successful uptake of a high-performance product likely to carry a premium price upon market introduction.

To enter the competition, manufacturers were required to submit 2,000 lamps. From these, a sample of 200 was randomly selected for short-term photometric testing according to industry standard test procedure IES LM-79-08, “Approved Method: Electrical and Photometric Measurements of Solid-State Lighting Products.” Testing was conducted by two independent test laboratories. The same 200 lamps were then evaluated for longer-term lumen maintenance using the equipment, procedures, and statistical method described in this paper.

3.0 Lumen Maintenance Testing

3.1 Test Facility

DOE required a facility for testing SSL replacement lamps entered in the L Prize competition. The facility had to maintain specific operating conditions and take accurate, reliable, and repeatable measurements over the intended test period. The requirements identified for the test facility were as follows:

1. Accommodate at least 200 samples operating simultaneously.
2. Accommodate two types of lamps: A19 or PAR38, both types with standard Edison medium screw (E26) base.
3. Operate lamps in base-up position.
4. Maintain 45°C ambient temperature with a temperature tolerance of $\pm 2^\circ\text{C}$ or less.
5. Operate lamps continuously.
6. Monitor power conditions to allow detection of failures and anomalies within 24 hours or less.
7. Perform relative photometric measurements of each sample at defined intervals, with measurement equipment placed in exactly the same location relative to the sample each time.
8. Provide lamp spacing to allow for measurement equipment to be positioned beneath each sample and to capture necessary data without interference from adjacent samples.
9. Perform relative measurements of each sample at predetermined time intervals for the following purposes:
 - a. To determine when the lamp sample has reached its maximum light output.
 - b. To measure changes in spectral output (lumen depreciation and chromaticity) during an additional 5,000-hour period of operation, beginning at the point of maximum light output.
10. Record and store specified data automatically throughout the test period.

3.2 Test Apparatus Development

The Lumen Maintenance Test Apparatus (LMTA) was designed and constructed in Richland, Washington, for the purpose of evaluating long-term lumen maintenance of LED lamps submitted under the L Prize competition. The apparatus has space for testing 202 lamps (including one calibration lamp and one rest-space position) and was designed to maintain a constant ambient temperature of 45°C at a plane coincident with the electronic components of the devices under test.

Figure 3.1 displays the exterior of the LMTA, showing the exhaust fan (top), integrating sphere and track mounting (through open door), and power supply and data collection equipment (at right). Figure 3.2 shows the interior with lamps mounted base-up, integrating sphere, and track mounting. Lamps had not yet been powered on in this photo.

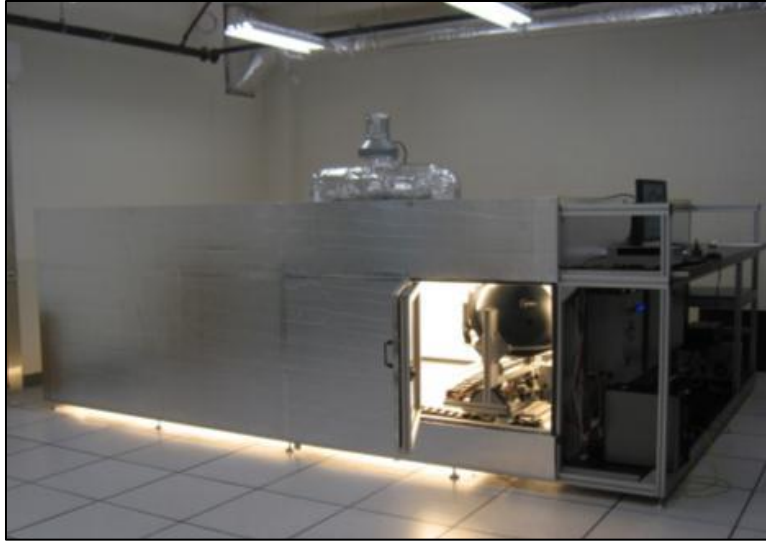


Figure 3.1. LMTA Exterior *Source: PNNL*



Figure 3.2. LMTA Interior *Source: PNNL*

The apparatus was constructed of Bosch aluminum tubing enveloped by rigid polyisocyanurate insulation (1.5 inches thick on the top and 1.0 inches thick on the sides of the apparatus). Original plans called for Plexiglas walls to enable viewing of the testing operation (see Figure 3.3); however, the elevated temperature caused the panels to bow, thus allowing free convection. As a result, the apparatus was unable to reach 45°C passively and the panels were removed and replaced with additional rigid insulation.



Figure 3.3. LMTA with Original Plexiglas Walls *Source: PNNL*

The apparatus was designed to hold the lamps in a rectangular grid configuration, in base-up position, and spaced exactly 8 inches on center. The lamps are mounted within 7-inch aluminum U-channels (see Figure 3.4) measuring approximately 11 feet long, 17 inches total, each supporting 12 lamps. The channels are mounted in parallel with a 1-inch space between. The 1-inch space is maintained between all channels and the perimeter containing walls such that, when negative pressure is applied above, the resulting airflow will be approximately uniform.



Figure 3.4. Close-up of U-Channels *Source: PNNL*

The lamps are mounted 3 feet above the finished floor. Below the lamps, a custom-engineered, 20-inch integrating sphere is mounted to a platform affixed to a computer-numerical-controlled track system, allowing it to move in the x and y directions. The z (vertical) direction is accomplished by a pneumatic

cylinder. A pneumatic system was selected for two reasons: (1) it provides compliance in the system for channel deflection and (2) it provides the ability to adjust the force applied to the channel for sealing purposes.



Figure 3.5. Integrating Sphere and Computer-Numerical-Control Track System *Source: PNNL*

The integrating sphere, supplied by Orb Optronix, Inc., incorporated several custom elements. The most notable difference from a traditional sphere is that the meeting axis for the two lobes is oriented 45° from vertical. This allows for a seamless fixed aperture for lamp insertion. Further, the sphere is outfitted with additional thermocouples for temperature measurement.

The type of emission and lamp type (omnidirectional vs. directional) drives the method of photometry for sphere measurement. Measurement of omnidirectional lamps employs four pi (4π) geometry, in which the lamp is fully inserted into and rests at the center of the sphere during measurement. In contrast, 2π geometry is used for directional lamps. In this geometry the plane of the emitting surface of the lamp is oriented at the sphere aperture such that all emission is captured within the sphere.

The 200 sample lamps sent to independent testing laboratories were measured using 4π geometry in large (2-meter-diameter) integrating spheres. Those lamps were then sent to PNNL for the elevated temperature testing. Only the relative change in flux/chromaticity needed to be monitored to meet L Prize and statistical method requirements. While the LMTA system is capable of great accuracy, testing with full 4π geometry measurement would require a considerably larger apparatus, as the lamps would need to be spaced much farther apart to allow for clearance when the sphere comes up to meet and envelop the lamp.

The ultimate measurement geometry used in the LMTA was neither 4π nor 2π but rather an “abridged 2π .” The desire was to capture all the omnidirectional emission without requiring the lamp to be centered within the integrating sphere. In this abridged 2π geometry, approximately two-thirds of the lamp’s maximum overall length is within the sphere during measurement, capturing all the light emission.

3.2.1 Systems Integration

The need for a system that simultaneously maintains a precise thermal environment, samples and records spectral data, and provides a means for repeatability required the integration of three discrete systems into a singular automation and control scheme. To meet the challenge, engineers selected off-the-shelf systems with robust features and flexible configurations and contracted with Orb Optronix to customize their SpectralSuite™ software to act as the unifying element and provide the graphical user interface for spectral data collection.

SpectralSuite captures and logs spectral data readings, makes requests of the programmable logic controllers (PLCs) to move the integrating sphere, and conducts a series of checks to ensure the system remains within tolerance. The software provides the user with real-time reports of radiometric and photometric measurements as well as the temperatures within the sphere at the time of the reading. The user is also able to view all prior data collection sequences to visually compare changes over the duration of the testing.

The PLCs that control sphere movement receive motion requests from SpectralSuite, providing the electrical signals to operate the xy motors and z pneumatic cylinder, and to monitor the various limit and interlock switches. The PLCs also monitor the status of error modes provided by the Campbell Scientific, Inc., measurement and control system and, in the event of a fault (temperature out of tolerance, power anomalies, etc.), will automatically shut the system down and autodial the operator.

The Campbell Scientific system monitors and logs temperatures inside the apparatus via an array of thermocouples, the current at each of the 17 rows (each row containing 12 lamps, except the first row which has 10), and the output voltage at the main power supply. In addition, the system controls the exhaust fan atop the apparatus, cycling ambient room air through the lamps to maintain $45^{\circ}\text{C} \pm 1^{\circ}\text{C}$.

Figure 3.6 depicts the system integration. The blue arrows indicate directional or bi-directional monitoring or communication between systems, red arrows indicate data collection (logging), and green arrows indicate control.

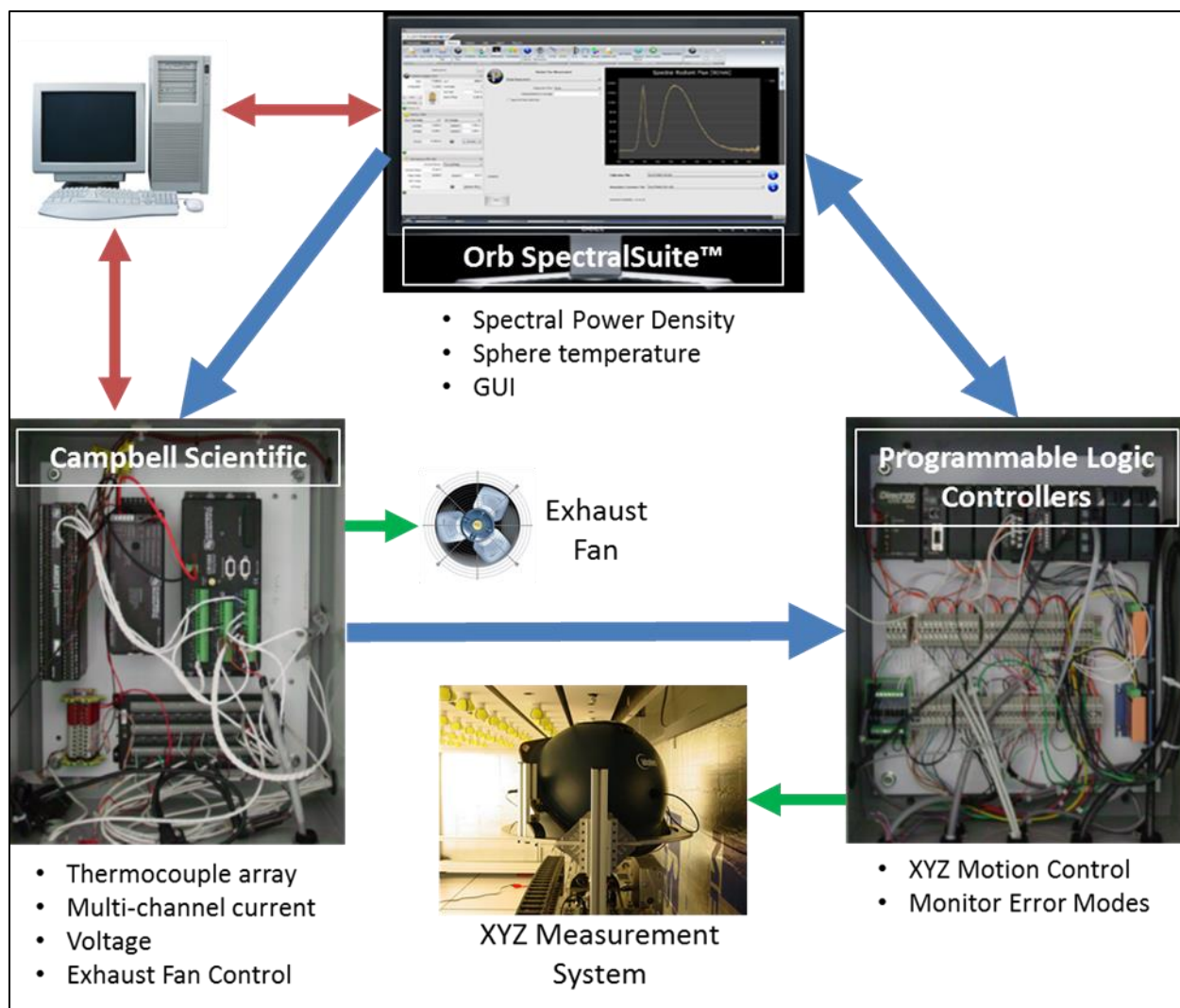


Figure 3.6. Diagram of Systems Integration

3.2.2 Data Collection

The LMTA continuously monitors and logs the temperature within the apparatus, the current to each lamp row, and the output voltage of the main power supply. In the event of a catastrophic lamp failure, the system, by virtue of its normal parameter recording, will know when a failure occurred by the change in current to that channel. The exact failed lamp can only be identified by operator visual inspection or the next photometric data collection sequence.

Measurements of each lamp are made weekly (every 168 hours) using the integrating sphere. (NOTE: For the first approximately 2,000 hours of LMTA operation, measurements were made more frequently—every 100 hours—to provide greater detail for identification of the maximum average output point for the lamps under test.) The data collection sequence takes just over one hour for all 202 lamps. Each time the integrating sphere runs a measurement sequence, the following data are recorded:

1. Full spectral power distributions from which all radiometric, photometric, and colorimetric properties are calculated

2. The temperature within the sphere at the time of measurement.

Figure 3.7 shows a screen capture of data collection in progress. The system has completed measurements on all lamps represented by green cells in the grid and is currently taking measurements on the lamp in cell K7. The display is set to show luminous flux but can be changed, under Property to Display in the lower middle of the screen, to show temperature, radiant flux, or saturation (related to chromaticity).

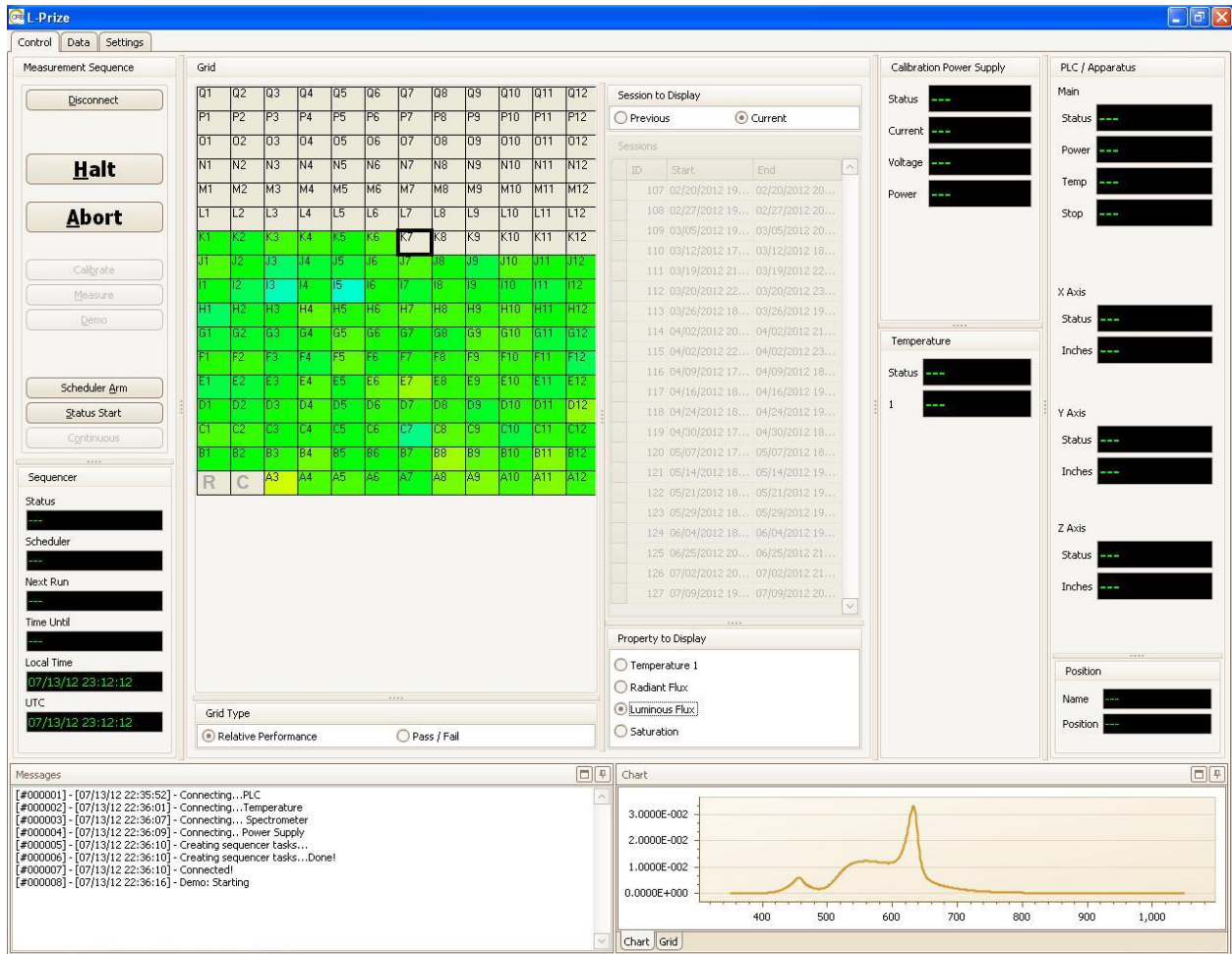


Figure 3.7. Data Collection Showing Luminous Flux

4.0 Interpreting Results: Statistical Method

The software and test equipment capture and store significant and highly accurate data regarding the performance of the lamps under test. Each weekly measurement yields 143,000 separate data points (the number is high because the measurements include full spectral power distributions each time, for each lamp). Several analytical techniques were used to evaluate these data and verify the level of performance of the lamps relative to the L Prize requirements.

4.1 Non-Parametric Statistical Approach

Non-parametric confidence interval estimation was used to drive the sample size and analyze the resulting data because, when the L Prize competition rules were written, DOE did not know which companies would enter the competition, and therefore did not know in advance what kind of lumen maintenance behavior (performance distribution) to expect from an entry. This statistical approach makes no assumptions about the shape of the products' lumen maintenance distribution. The reason this is necessary is illustrated in the figure below.

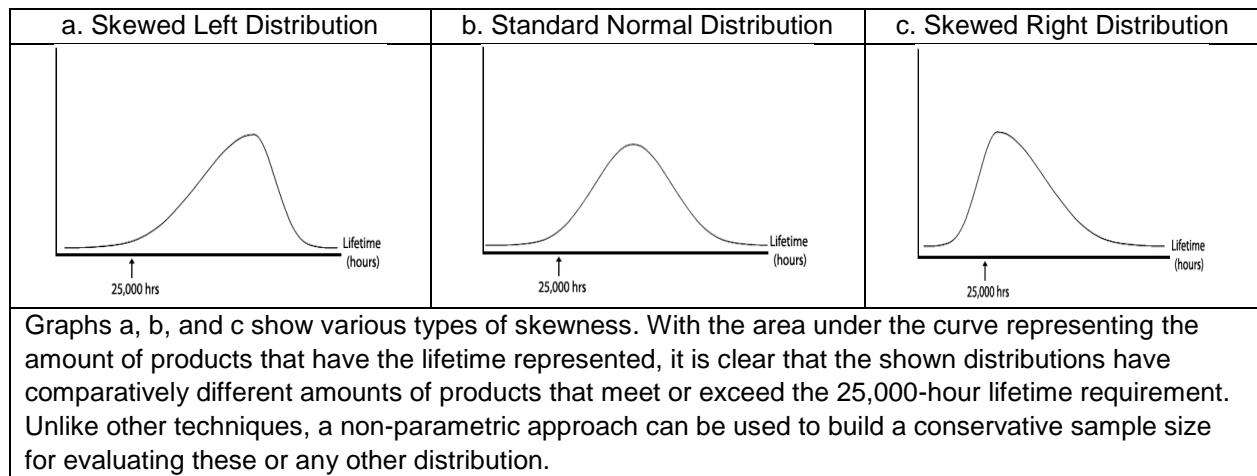


Figure 4.1. Illustration of Non-Parametric Statistical Approach

The statistical approach for evaluating an L Prize entrant's ability to meet the long-term performance requirements of the competition was based on the following parameters:

1. Lumen maintenance at least 70%: Entrant maintains a forecasted lamp output of at least 70% of its maximum (initial) lumen rating.
 - a. The EISA legislation defined lifetime for the L Prize product as follows: The product must maintain at least 70% of the defined light output (greater than 900 lumens for the 60-watt replacement category) for at least 25,000 hours of operation.
2. Failure rate 10%: No more than 10% of products are expected to fail (fall below L_{70}) within 25,000 hours of operation.
 - a. In accordance with goals for large-scale manufacturing, requiring a failure rate of no greater than B_{10} puts measurable bounds on quality assurance. When production goals are in the millions of

fixtures, having a failure rate of greater than 1 in 10 is expensive and inefficient; entrant's products must be both capable and reliable.

3. Confidence level 95%: 95% statistical confidence in the L Prize tolerance limit.
 - a. 95/90 tolerance limit: Lumen value, based on a sample of tested fixtures, assures with 95% confidence level that 90% of the underlying fixture lumen values will exceed.
 - b. Because of uncertainty in the production process, testing samples only give an estimation of how the entire population of devices will perform. As the confidence level, commonly 90%, 95%, or 99%, (C_{90} , C_{95} , or C_{99}), increases, the tolerance limit will decrease, thereby increasing the possibility of rejecting a product even though it performs satisfactorily. Conversely, lowering the confidence level increases the tolerance limit, generally making it easier for a product to qualify. The specific confidence level used will influence the likelihood of making such incorrect decisions.
 - c. C_{95} was selected for this application. Increasing the confidence level protects consumers from receiving a faulty product when the manufacturer meets the statistical criteria, while a lower confidence level makes it easier and less expensive for manufacturers to qualify but may increase the risk of consumers receiving faulty lamps.
4. Testing will be performed on a random sample of 200 from the 2,000 quality control lamps provided.
5. A non-parametric approach will be used to determine confidence intervals.
6. According to this non-parametric approach, when test results are ranked from worst (x_1) to best (x_{200}), the 13th-worst-performing unit (x_{13}) is the lower tolerance limit, which asserts that, subject to C_{95}/B_{10} criteria, we have 95% statistical confidence that 90% of the lamps produced will perform at least as well as the 13th-worst unit.

For the C_{95} & B_{10} standards, 10% or 200 of the 2,000 samples submitted is a sufficiently large portion to provide accurate results while not being encumberingly large. A larger sample size would increase the confidence limit, but would also increase testing costs.

According to this non-parametric approach, when test results are ranked from worst (x_1) to best (x_{200}), the 13th-worst-performing unit (x_{13}) provides a 95/90 tolerance limit. This approach does not depend on the underlying shape of the lumen distribution, but rather only on the $n = 200$ tested units and the 10% (B_{10}) target. When appropriate distributional knowledge is available for the lamps being evaluated, satisfying the B_{10} criteria for a sample of 200 would result in a tighter lower bound, or a smaller sample could be used to attain a comparable lower bound.

To protect against faulty assumptions about the underlying distribution, this non-parametric approach was adopted and will widen the lower bound as compared to a parametric approach due to the increase in the assumed uncertainty or, in this case, increase sample size requirements to attain the same lower bound.

While this approach may potentially discount lighting products that would otherwise have met the L Prize criteria, it ensures that the samples that pass are highly qualified.

4.2 Determining Point of Maximum Flux

The lumen maintenance testing and evaluation approach defined by DOE required identification of the point of maximum output of the lamps under test, i.e., the time in run hours at which the normalized lumen maintenance trend reached its maximum. To identify this point required an accurate estimate of the average behavior of these lamps over time.

The data showed that the behavior of normalized lumen maintenance was variable from one lamp to another. Characterization of the mean normalized lumen maintenance across run hours was developed based on the raw measurements for all 202 lamps at each run hour.

The goal was to obtain a correct estimate of the mean across run hours. Figure 4.2 is a scatterplot of the data with a local quadratic loess (local regression) fit with span = 0.5 superimposed. The “X” on the plot corresponds to the maximum value of the mean function (2,694 hours). The solid black dots correspond to the means at each measurement hour.

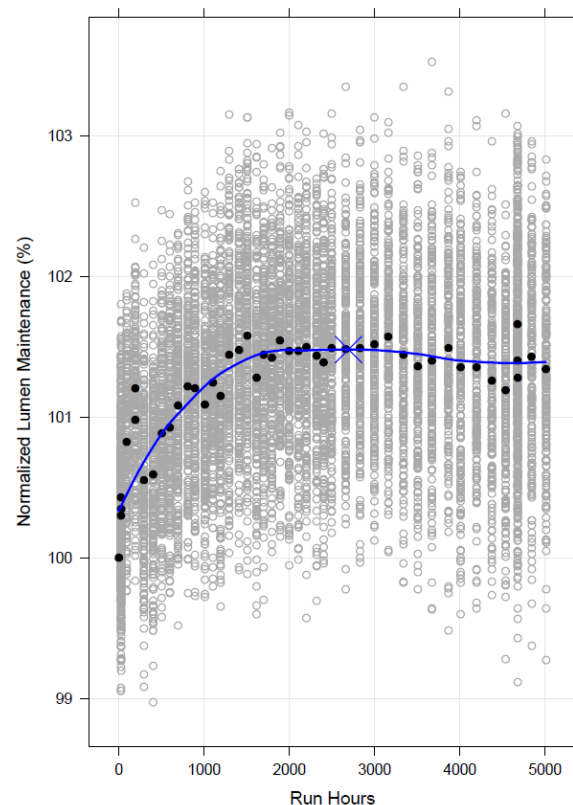


Figure 4.2. Measurement Data for All Lamps at 100-Hour Intervals

While there might be a parametric model that could reasonably fit these data, we were mainly interested in obtaining the best fit to the mean trend, for which non-parametric local methods such as loess are suited. For this loess modeling, local quadratic fitting with a span parameter of 0.5 was used to handle the curvature in the lumen maintenance trend. The span parameter controls how smooth or wiggly the fit to the data is. Figure 4.3 displays the residuals from the loess fit to demonstrate the reasonability of the fitting decisions and to check for deviations from zero.

In this plot, the solid line is a loess fit to the residuals; this time it is simply to highlight deviations from zero. In the first 500 hours, the loess model is not fitting the data well due to the sudden peak in the data during that time. However, it is not important to have reliable mean estimates within this region of time. After the first 500 hours, it appears that the mean of the residuals is appropriately varying around 0 (representing an unbiased estimating procedure).

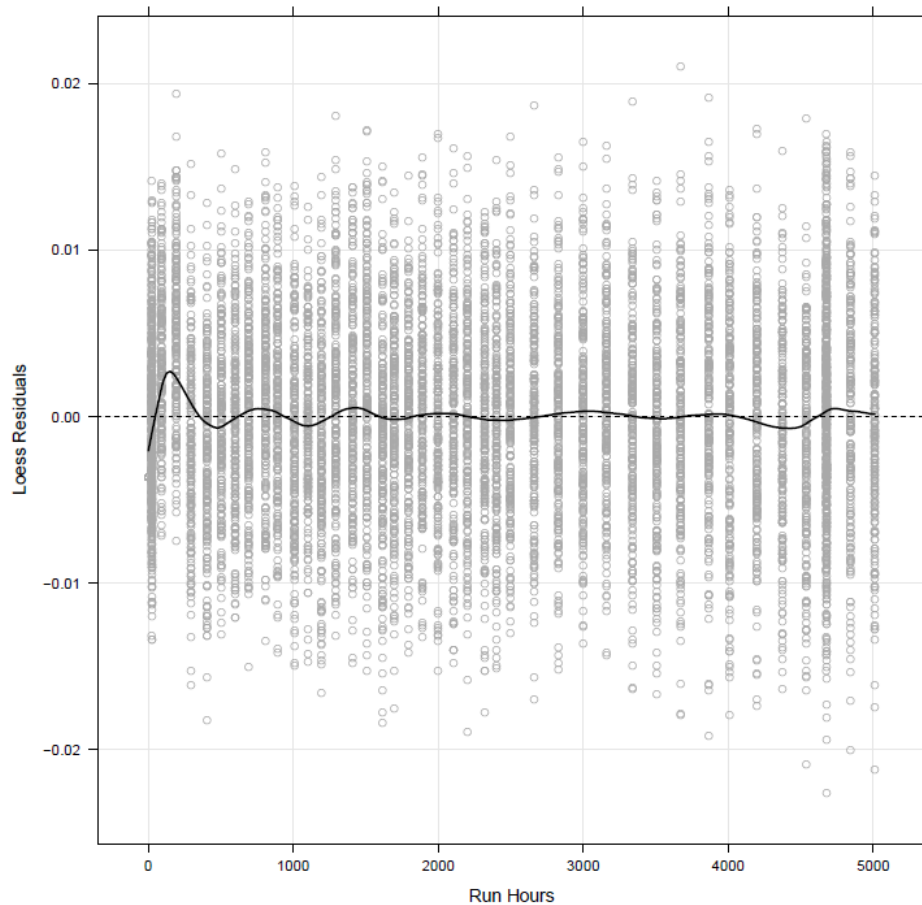


Figure 4.3. Residuals from the Loess Fit

The mean estimate described above is just that, an estimate. As with all estimates, there is a measure of uncertainty associated with it; thus there is also uncertainty about the maximum value obtained from the mean estimate. We can calculate the uncertainty about the mean estimate provided that the residuals are normally distributed (which they are) and use this to construct confidence intervals about the mean.

Figure 4.4 shows the scatterplot of the data with the loess fit superposed as before, but also adds a 99% upper bound for the mean trend function. This bound means that we are 99% confident that the mean value lies below this line. Note that this plot is focused within a smaller y-axis ranged for a better view of the curve. Following the assumption that the data follow the pattern of increasing and then reaching a plateau before descending again, it is possible that this plateau is reached before the observed maximum—highlighted by the orange box superposed in the plot. The lower bound of the orange box is at 1,636 hours. Choosing this as the point at which the data reach the plateau does not seem reasonable, but there is a local maximum inside the region of the box that is almost as high as the true maximum.

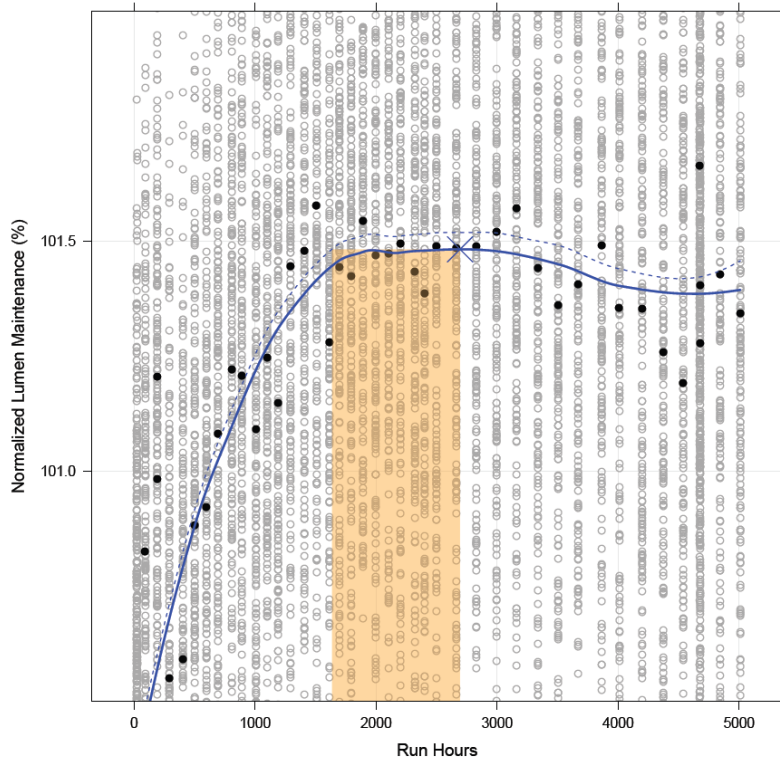


Figure 4.4. Loess Fit with 99% Upper Bound for Mean Trend Function

To better determine where the mean trend is expected to reach its plateau, a statistical technique called the “bootstrap” was used. The bootstrap process requires the data to be resampled, with replacement, many times and for each sample. The loess mean fit is calculated and the maximum value for that fit is recorded. The bootstrap process was done 1,000 times to get the results shown in Figure 4.5.

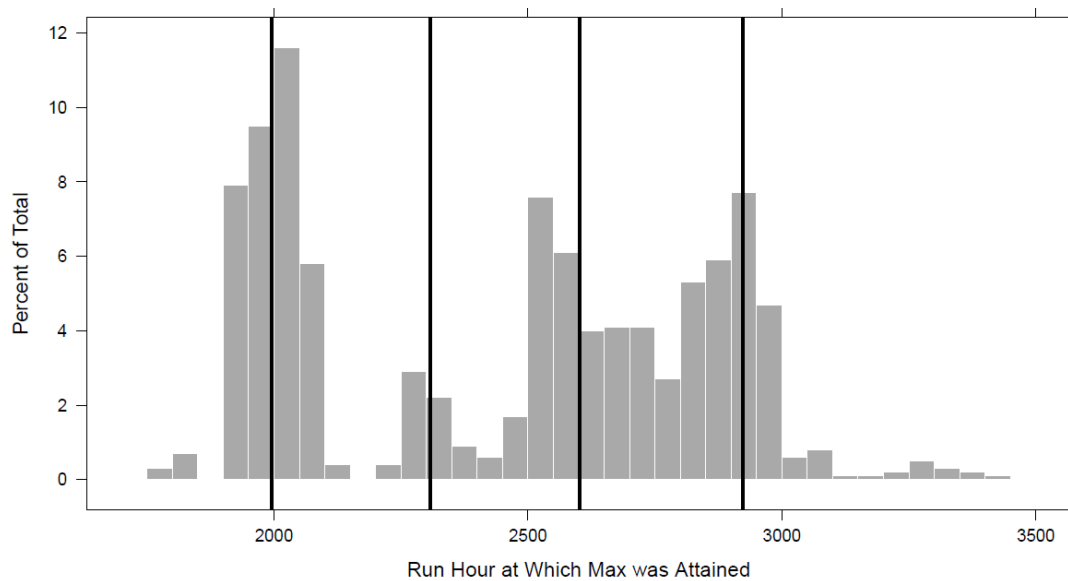


Figure 4.5. Frequency Distribution by Run Hours

Figure 4.5 shows the frequency of different run hour values at which the maximum occurred. It appears that there are four clusters, meaning there are four different places competing to be the maximum. Figure 4.6 shows that, regardless of which one is chosen, the distribution of the resulting maximum is essentially the same. This is good evidence that the plateau is reached at the first run hour cluster (having a mean of just under 2,000 run hours).

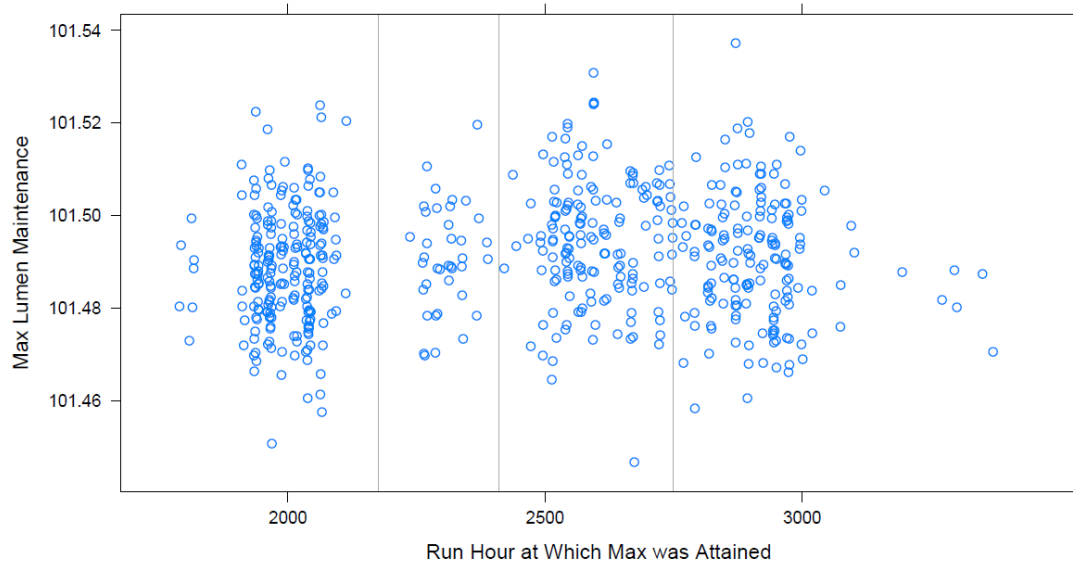


Figure 4.6. Distribution of Four Possible Maximum Output Periods

5.0 Test Results for First L Prize Winner

5.1 Lumen Maintenance

The maximum average light output for the 200 lamp samples under test (L_{\max}) was statistically estimated to have occurred at 1,996 hours. L Prize rules required lumen maintenance to be based on $L_{\max} + 5,000$ hours, = 6,996 hours. This point was reached on April 10, 2011. At the 14,000 hour and 25,000 mark, performance remained consistent.

Figure 5.1 shows the relative lumen maintenance data from 2,000 to 7,000 hours for the 13th-worst-performing sample, which represents the 95% confidence limit. These data were used to estimate future lumen maintenance of the 200 samples. As explained below, the prediction method proved to be quite accurate.

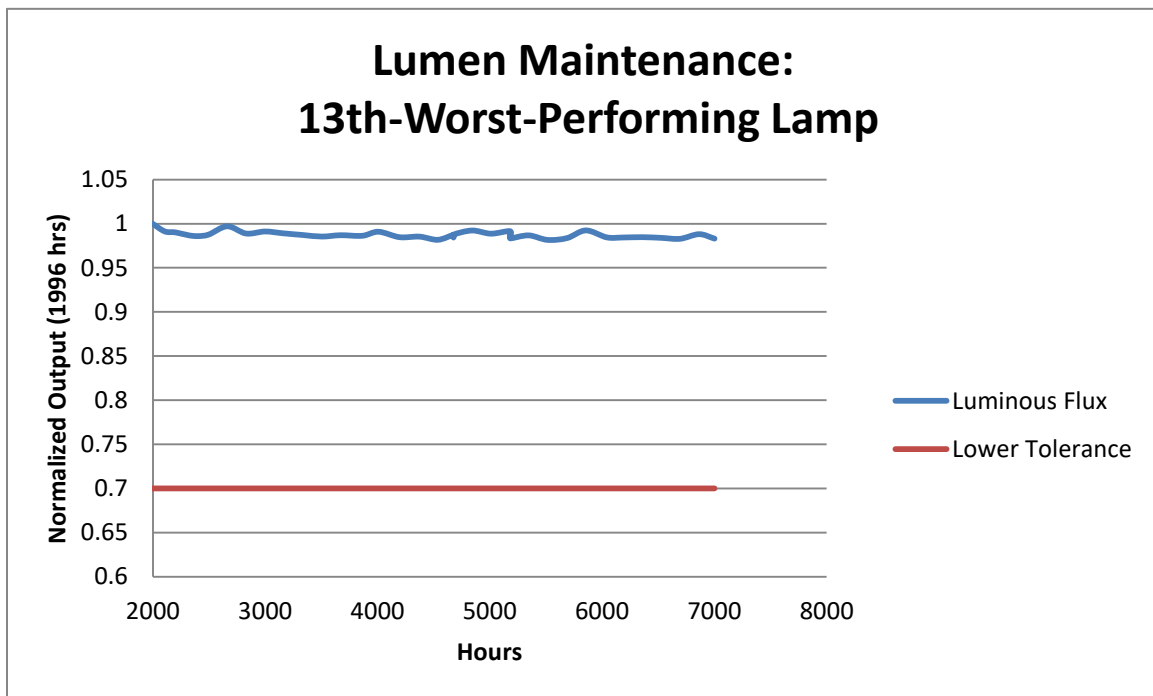


Figure 5.1. Measured Lumen Maintenance

The normalized lumen data were used to forecast the lumen output at 25,000 hours for each individual lamp (see Figure 5.2). The box plot at 25,000 hours shows the range of predictions from each of the 200 lamps. The red dot identifies the 13th-worst-performing item. The non-parametric statistical methodology says the $L_{70}/B_{10}/C_{95}$ estimate is based on performance of this 13th-worst-performing lamp in the 200-lamp sample. When an exponential decay function is fit to the data for the 13th-worst-performing lamp, lumen maintenance must be $\geq 70\%$ at 25,000 hours. As of April 10, 2011, at test hour 7,001, the 13th-worst-performing lamp of the 200 samples under test had projected lumen maintenance of 97.08% at 25,000 hours.³ As of late April, 2013, at the 25,000-hour mark, the actual average lumen maintenance of the

³ Projected lumen maintenance was originally reported as 99.3%. An error in the projection calculations was subsequently discovered and corrected. The corrected value is reported here.

samples remained 100% of the maximum average output, and the projected 70% lumen maintenance point would be well beyond 150,000 hours, the maximum that could be stated using the six-times guideline in TM-21 (i.e., projected L_{70} hours should not exceed six times the actual test period).

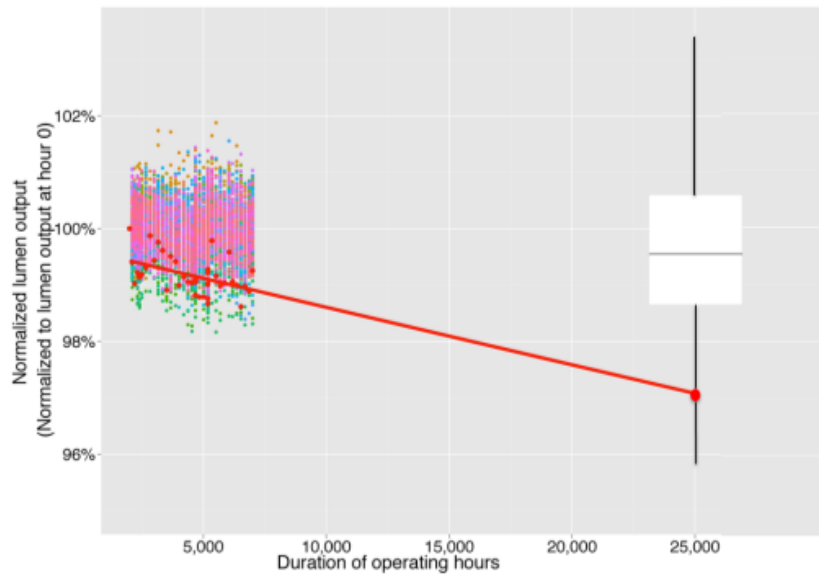


Figure 5.2. Forecast for 13th-Worst-Performing Lamp

5.2 Chromaticity Maintenance

Chromaticity data were collected at the same time as lumen maintenance data in the same apparatus. Since there is no consensus methodology for extrapolating chromaticity maintenance data, L Prize rules required lamps to show no more than $0.004 \Delta u', v'$ at $L_{\max} + 5,000$ hours. Using the same non-parametric methodology, we selected the 13th-worst-performing lamp to determine compliance with this requirement. At test hour 7,001, the 13th-worst-performing lamp had a measured chromaticity maintenance value of $0.0006 \Delta u', v'$. Figure 5.3 shows the chromaticity performance of the 200 lamps over the first 7,000 hours. The black dots and line show the trend at each measurement time. The horizontal blue line marks the L Prize established criterion below which each entrant's measurements must be measured at 7,000 hours. All 200 products were well within the requirement. At 14,000 hours and again at 25,000 hours, chromaticity remained at less than .002 on the CIE 1976 (u', v') diagram. This means color is holding steady, within the acceptable range. See sections 5.3 and 6.0 below for more detail on the results and analysis after 14,000, 18,000, and finally after the full 25,000 hours of operation.

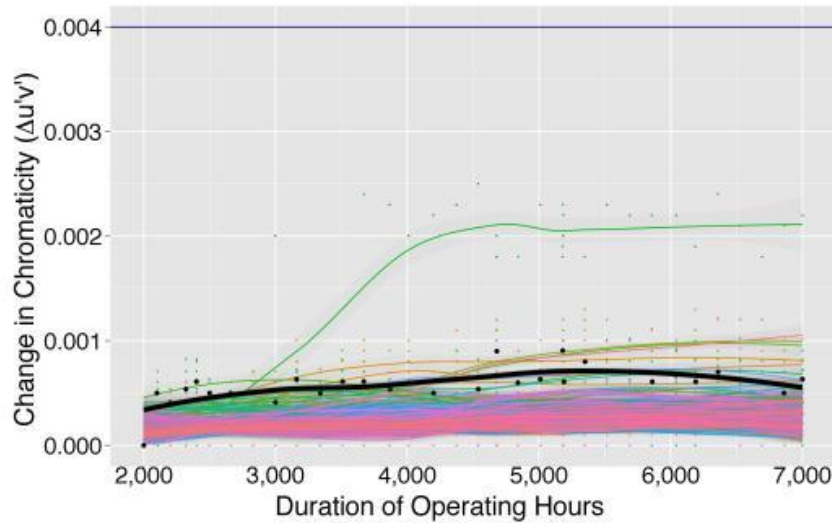


Figure 5.3. Change in Chromaticity over Test Period

5.3 Updated Interim Results Since L Prize Award

Although the minimum test period defined for the competition ($t_{\max} + 5,000$ hours) was reached in April 2011, DOE continues to operate the Philips lamps in the LMTA. Automatic measurements take place weekly. The reasons for longer-term operation and measurement of the lamps, even after the L Prize evaluation was completed and the award made, are the following:

1. To collect independent, high-quality, long-term, statistically valid lumen and chromaticity maintenance data for the L Prize-winning LED replacement lamp product and make it publicly available.
2. To compare long-term data with lumen maintenance predictions based on shorter-term (6,000-hour) lamp aging data and LM-80-08 data for the LED packages used in the integral lamp.

In January 2012, the accumulated operating hours for these lamp samples had passed 14,000 run hours; some additional statistical analysis was done at that time, including a comparison of the accuracy of the prediction method based on the 7,000-hour data compared to the 14,000-hour data. Figure 5.4 depicts the average lumen maintenance (blue line) for the 200 lamps installed in the apparatus, normalized to 100% at hour 1,996. Note the scale of the vertical axis is “zoomed in” to show just the region between 99.5% and 103.0% normalized lumen maintenance and 38° to 44°C average sphere temperature. The red line represents the average temperature inside the integrating sphere. The temperature drop that occurred between 10,000 and 11,000 hours and the corresponding uptick in average light output was anomalous and due to a door being left open temporarily on the test apparatus.

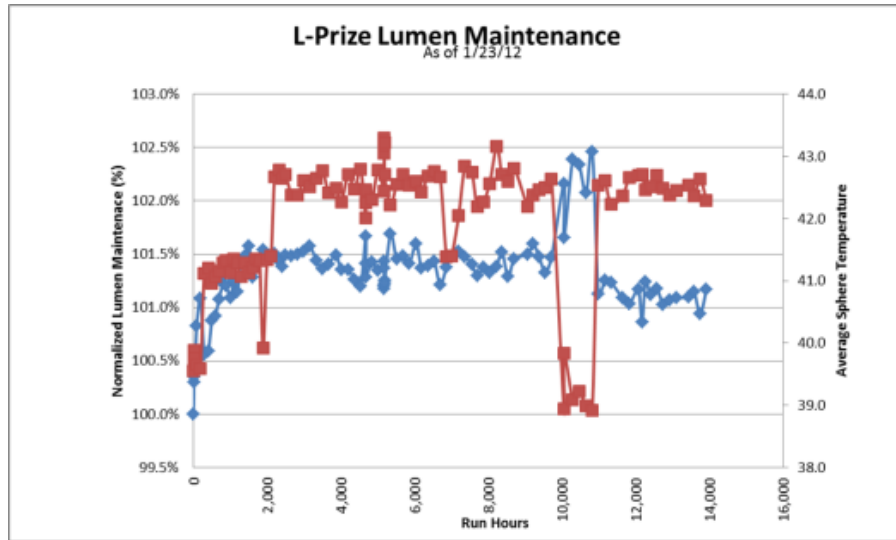


Figure 5.4. Average Lumen Maintenance and Sphere Temperature (°C) to 14,000 Hours

The following graphs summarize the results for each of the 200 lamps. Each box plot represents the range of 25,000-hour predictions of the individual lamps, based on the data collected to date by that time. The first observation was made after 2,000 hours.

Figure 5.5 shows how the ability to project 25,000-hour lumen maintenance improves with longer data collection.

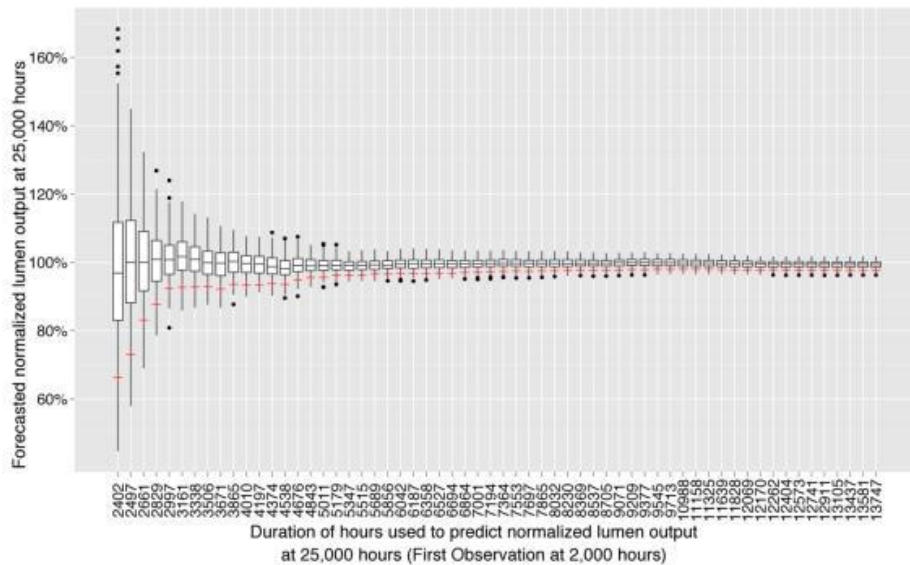


Figure 5.5. Range of 25,000-Hour Lumen Maintenance Forecasts

Figure 5.6 zooms in on the predictions between 4,500 and 14,000 hours, showing that the range of predictions started to moderate between 6,000 to 7,000 hours.

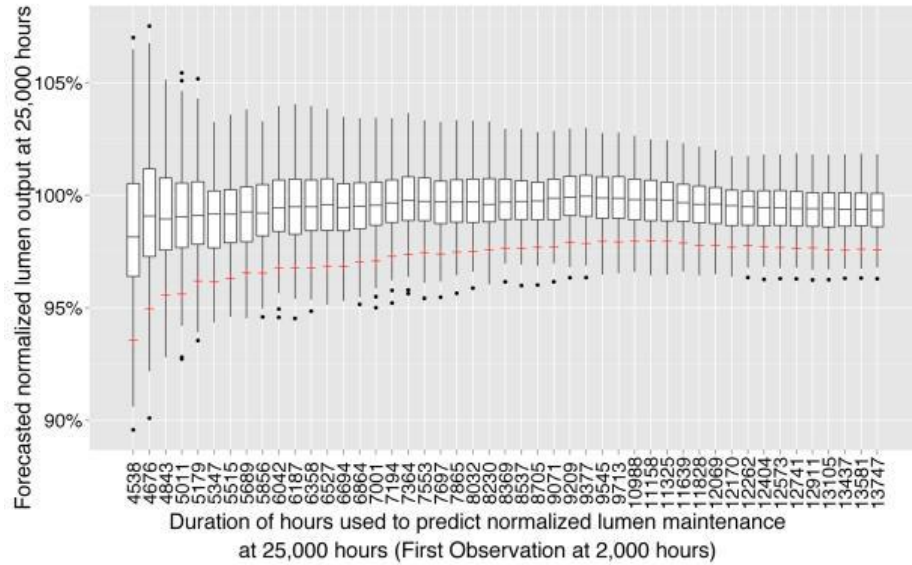


Figure 5.6. Predictions Between 4,500 and 14,000 Hours

In fact, the forecasts using 14,000 hours of data were very similar to the forecasts using 7,000 hours as shown in Figure 5.7.

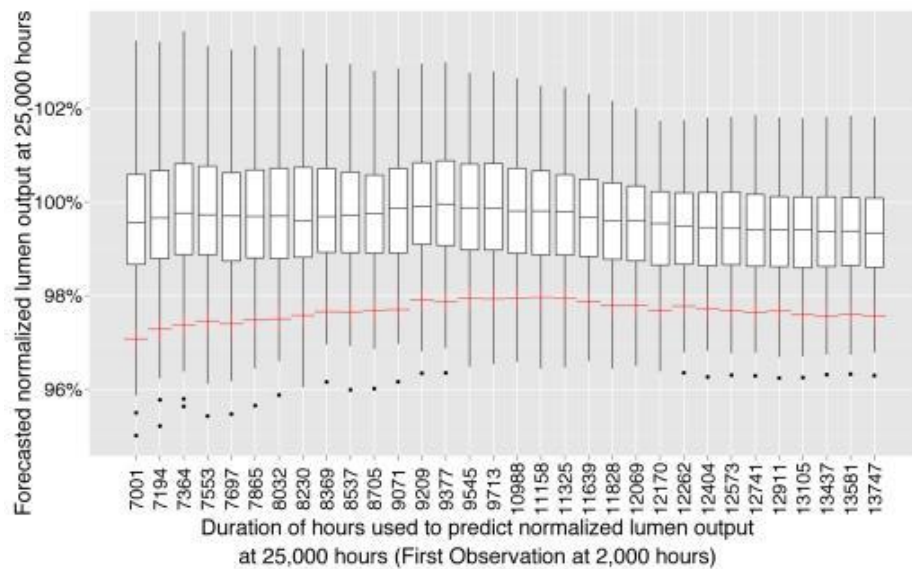


Figure 5.7. Forecasts of 7,000-Hour Data Similar to 14,000-Hour Data

The chromaticity maintenance of the L Prize-winning lamps after 14,000 hours of operation remained well within the parameters required by the competition for the $L_{\max} + 5,000$ hours minimum test period, as shown in Figure 5.8.

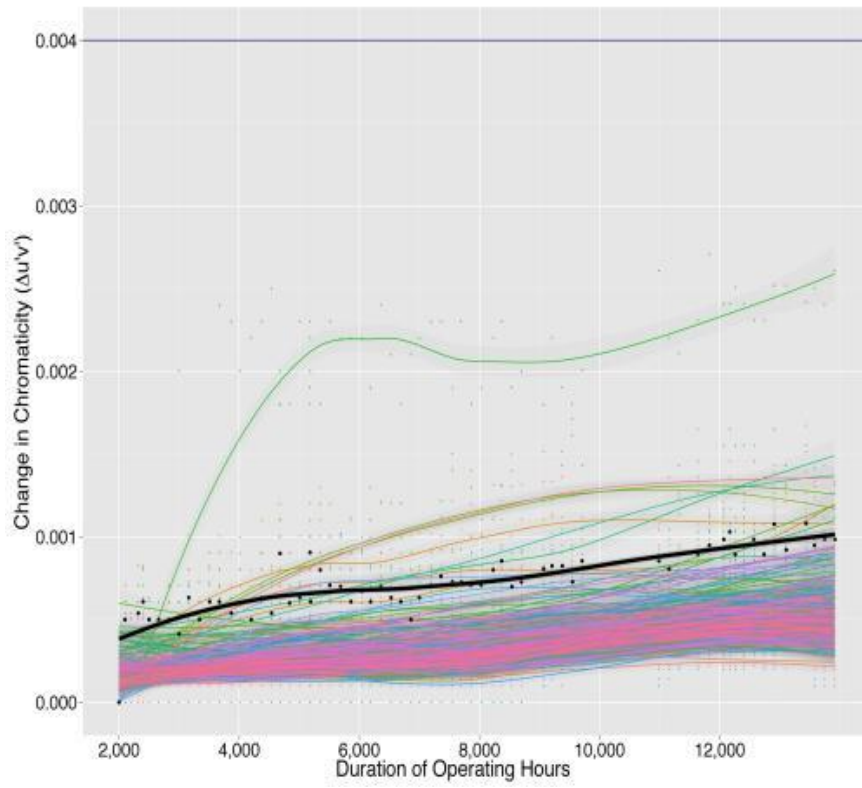


Figure 5.8. Change in Chromaticity up to 14,000 Hours

In July 2012, accumulated hours of operation passed 18,000 hours. Figure 5.9 shows the average lumen maintenance for the entire sample of lamps under test to that point.

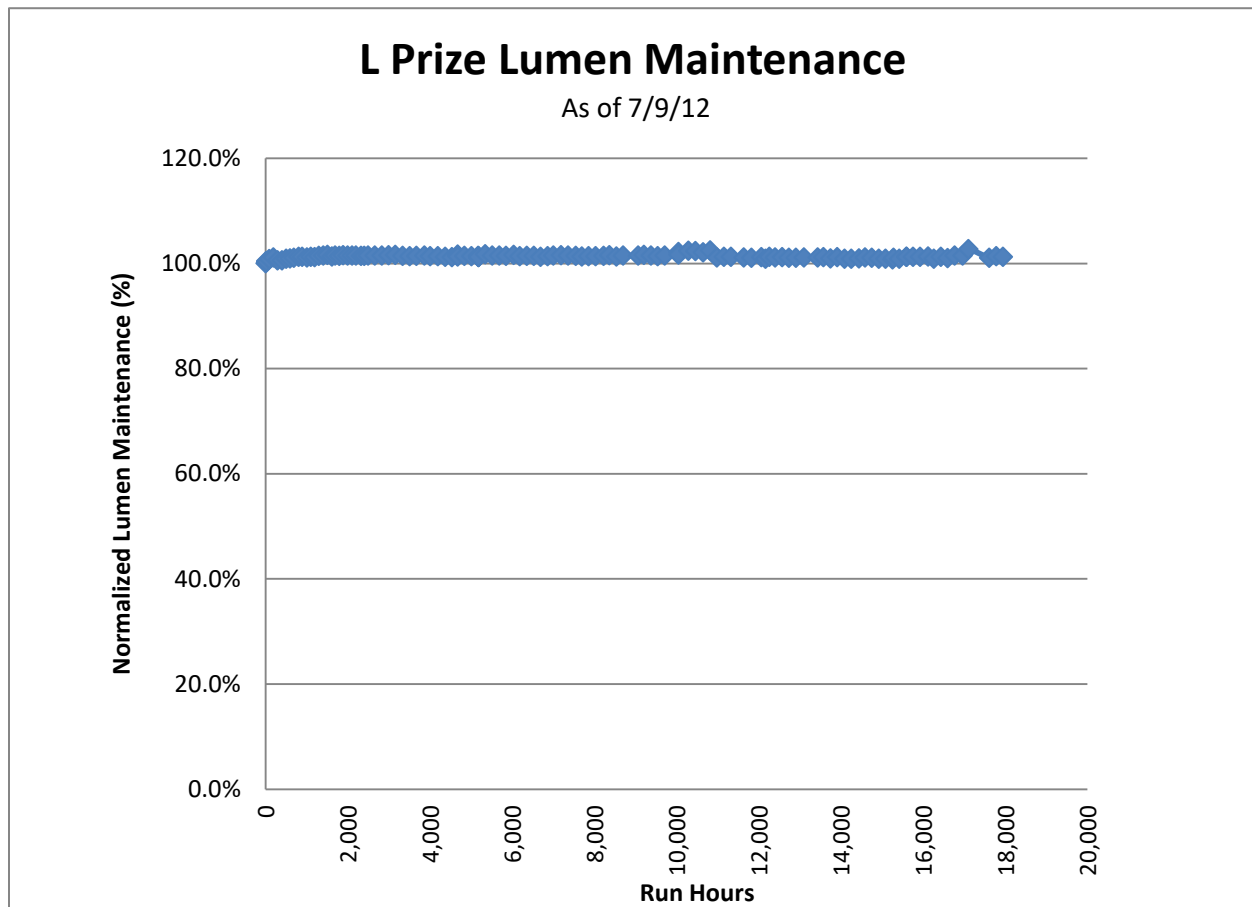


Figure 5.9. Average Lumen Maintenance After 18,000 Hours of Operation

Figure 5.10 and Figure 5.11 show updated lumen maintenance and chromaticity data, respectively, for the 13th-worst-performing lamp, representing the 95% confidence limit. Based on these data, projected lumen maintenance at 25,000 hours for the 13th-worst-performing lamp was 97.8%.

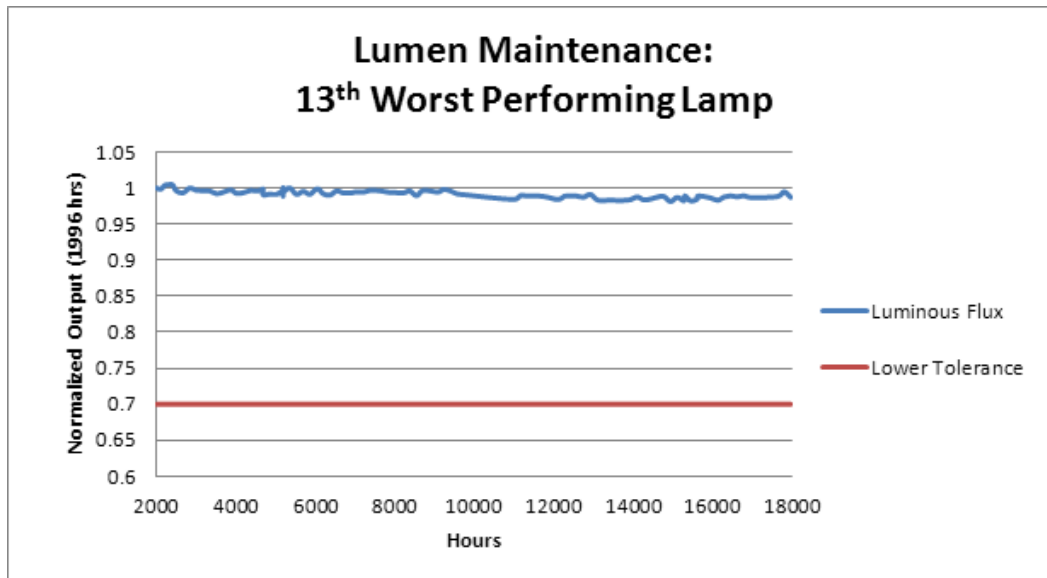


Figure 5.10. Lumen Maintenance After 18,000 Hours of Operation

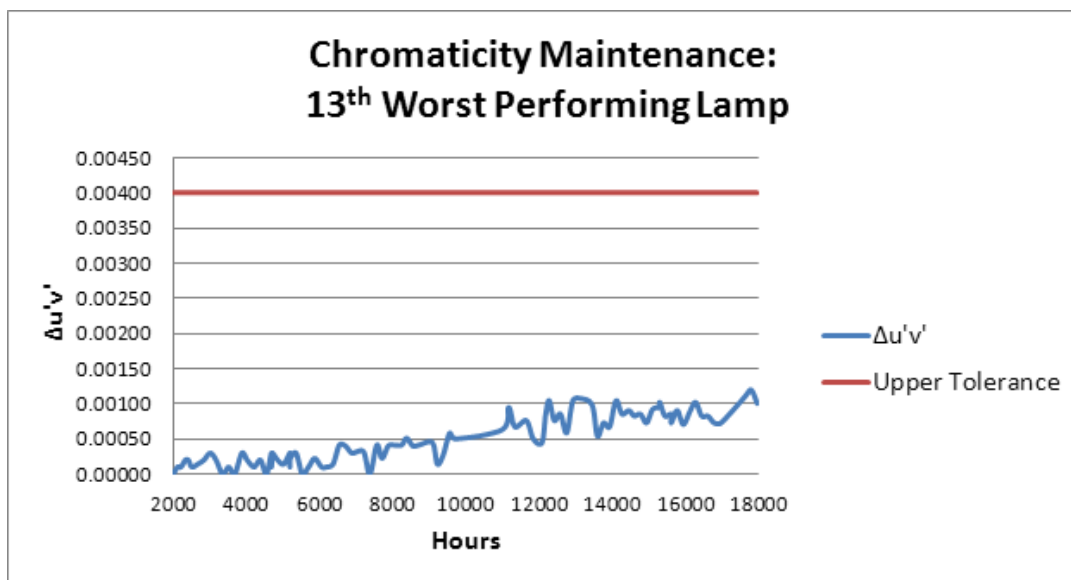


Figure 5.11. Chromaticity Maintenance After 18,000 Hours of Operation

6.0 Update: Results After 25,000 Hours

As of April 29, 2013, the 200 Philips L Prize lamps completed 25,000 hours of continuous operation in the lumen maintenance test facility. None of the samples has failed to date, and average lumen maintenance of the 200 lamps remains above 100% of initial light output. The projected time to 70% lumen maintenance would be well in excess of 150,000 hours (the six-times guideline in TM-21 limits forward projections to no more than six times the actual test period). Projected lumen maintenance should not be interpreted as projected lifetime, as it assumes all other components in the lamp continue to operate normally. Lamp failure could result from electronic or other component failure before or after LED output has reached 70%.

Figure 6.1 summarizes the average lumen maintenance of the 200 lamps under test, as of 25,000 hours of operation. Lumen maintenance is relative to the maximum output of this lamp sample, which occurred at 1,996 hours, as described above in section 4.2.

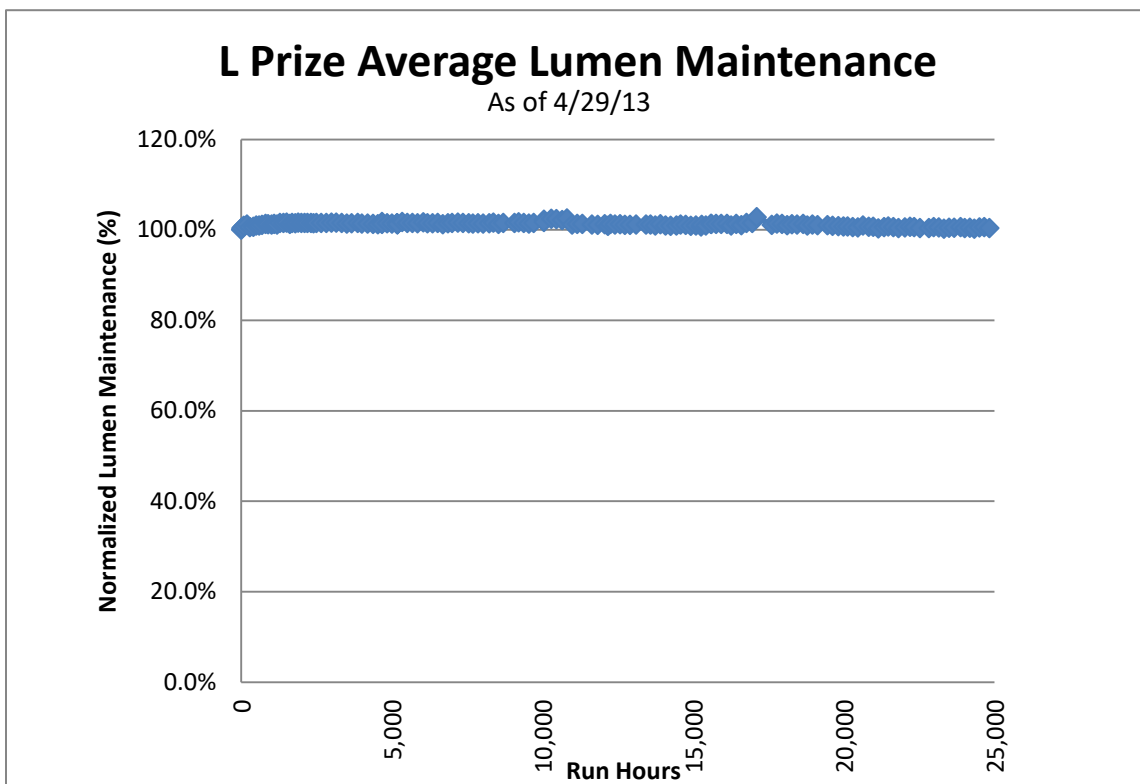


Figure 6.1. L Prize Average Lumen Maintenance After 25,000 Hours

Figure 6.2 provides additional information on lumen maintenance, separating out the best-performing and worst-performing lamps in the sample, as well as the 13th-worst-performing lamp, which, according to the non-parametric statistical method (for details see section 4.1 above) used in this analysis, represents the 95% confidence limit.

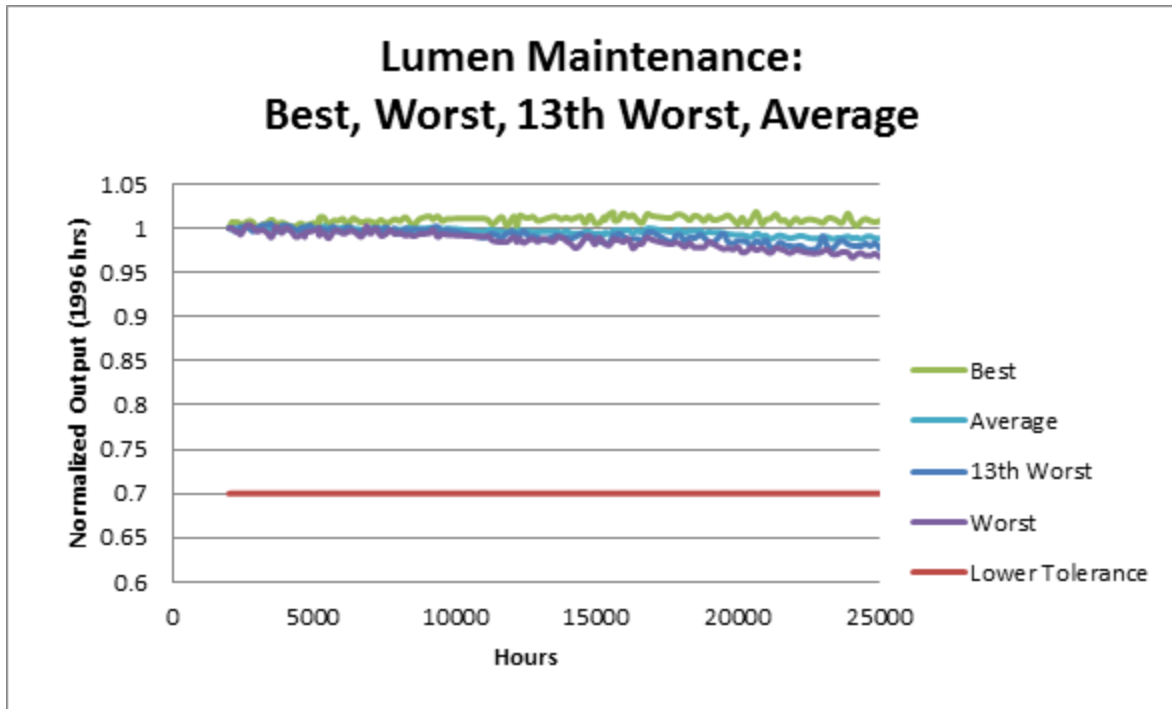


Figure 6.2. Comparison of Best, Worst, 13th-Worst, and Average L Prize Lumen Maintenance After 25,000 Hours

The average chromaticity change remains well below the L Prize requirement for the original evaluation period (time to maximum output plus 5,000 hours). After 25,000 hours, the chromaticity change measured in the lamps under test remains less than .002 on the CIE 1976 (u' , v') diagram.

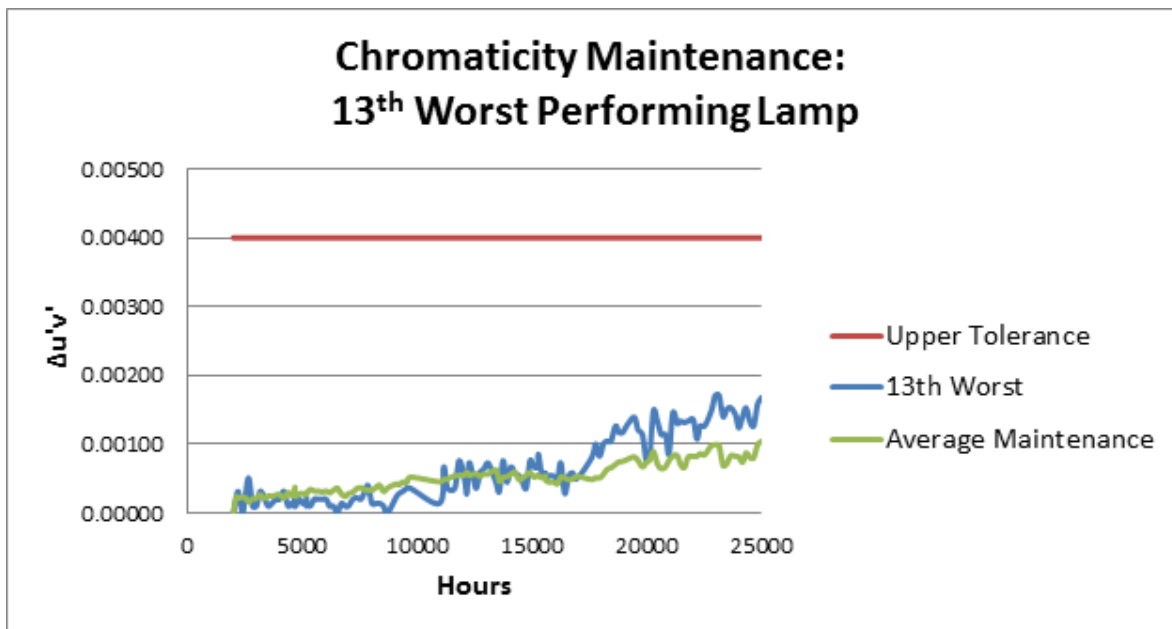


Figure 6.3. L Prize Average and 13th-Worst Chromaticity Maintenance Relative to Required Limit

The results of 25,000 hours of L Prize testing are significant in several respects. First, this is the largest publicly available data set representing long-term performance of a statistically significant sample of LED integral lamps tested at elevated temperature, under tightly controlled thermal and electrical conditions, and with automated data collection to minimize measurement variability. A number of organizations have conducted long-term “rack” testing of LED lamps, but open-air conditions are subject to much more variation in temperature and air flow, so measurement variability may be significant. Measurements are typically done manually, which introduces additional potential for variability and error.

Second, 25,000 hours is a commonly estimated, targeted, or required minimum L70 lifetime for LED integral lamps, including the ENERGY STAR[®] program.⁴ Energy efficiency programs and users would ideally like to know actual performance of LED lamps at 25,000 hours, but full testing for that long is impractical, requiring nearly three years of continuous operation. The L Prize test results indicate that the lumen maintenance projection method based on IES TM-21-11 is reasonable. During the first one to two thousand hours, LED lamp output typically continues to increase. In testing the L Prize lamps, great pains were taken to identify the maximum average output of the lamps in the test sample, which turned out to be 1,996 hours for this particular product. (It wasn’t easy to determine the point of maximum output because luminous flux remained relatively flat after the first few hundred hours.) This kind of data analysis is not normally practical in product testing, so TM-21 stipulates exclusion of the first 1,000 hours of data. In the L Prize case, the variability of 25,000-hour predictions using the data to 7,000 hours was found to differ little compared to the 14,000 hour data. Note that TM-21-11 technically applies to use of LM-80 data, which is lumen maintenance of the LED package (as opposed to the LED integral lamp) over time, under specific ambient temperature and drive current conditions. In the absence of an industry standard method for projecting long-term integral lamp performance, DOE applied the TM-21 method to full-lamp data.

Third, the L Prize results, while they pertain to one arguably highly specialized product, demonstrate an important point: well-designed LED integral lamps can operate very reliably, with excellent lumen and chromaticity maintenance, over long periods of time. The L Prize lamp was designed to go above and beyond the very demanding requirements of the competition. Producing that lamp in 2009 was very expensive, but in the five years since then, LED chip efficacy, materials, thermal management, circuit design, and other aspects of LED product design have continued to advance. The L Prize shows that excellent long-term performance by LED integral lamps is achievable, and continued advances allow manufacturers to balance efficacy, output, color quality, and longevity to meet the needs of users.

6.1 Future Testing and Analysis

DOE plans to continue to operate a subset of the L Prize-winning lamps in the LMTA. This will be the only publicly available, third-party data set of long-term LED product operation.

Depending on how the SSL market evolves, how the technology changes over time, and whether buyers and producers coalesce around well-designed products with strong lumen and chromaticity maintenance performance, these test results indicate it might be possible to pay less attention to lumen and chromaticity testing in the future, saving producers and buyers significant money, and allowing them to shift their attention to other luminaire performance attributes.

⁴ ENERGY STAR[®] Program Requirements for Integral LED Lamps, version 1.4.

7.0 Update: Results After 40,000 Hours

A sample of 31 L Prize-winning lamps produced by Philips remains under test in the Accelerated Long-Term Test Apparatus (ALTA) at PNNL. These lamps have now passed 40,000 hours of operation under elevated ambient temperature (45° C) conditions. Measurements taken as of 40,890 hours indicate average lumen maintenance for the 31 lamps under test was 95.6%. At that rate, the average operating hours to reach 70% lumen maintenance would be 194,765 hours (using TM-21 forecasting methodology). Projected lumen maintenance should not be interpreted as projected lifetime, as it assumes that all other components in the lamp continue to operate normally. Lamp failure could result from electronic or other component failure before or after LED output has reached 70%.

Figure 7.1 below summarizes the lumen maintenance of the 31 lamps under test, as of 40,890 hours of operation. Lumen maintenance is relative to the maximum output of the original lamp sample, which occurred at 1,996 hours, as described above in section 4.2.

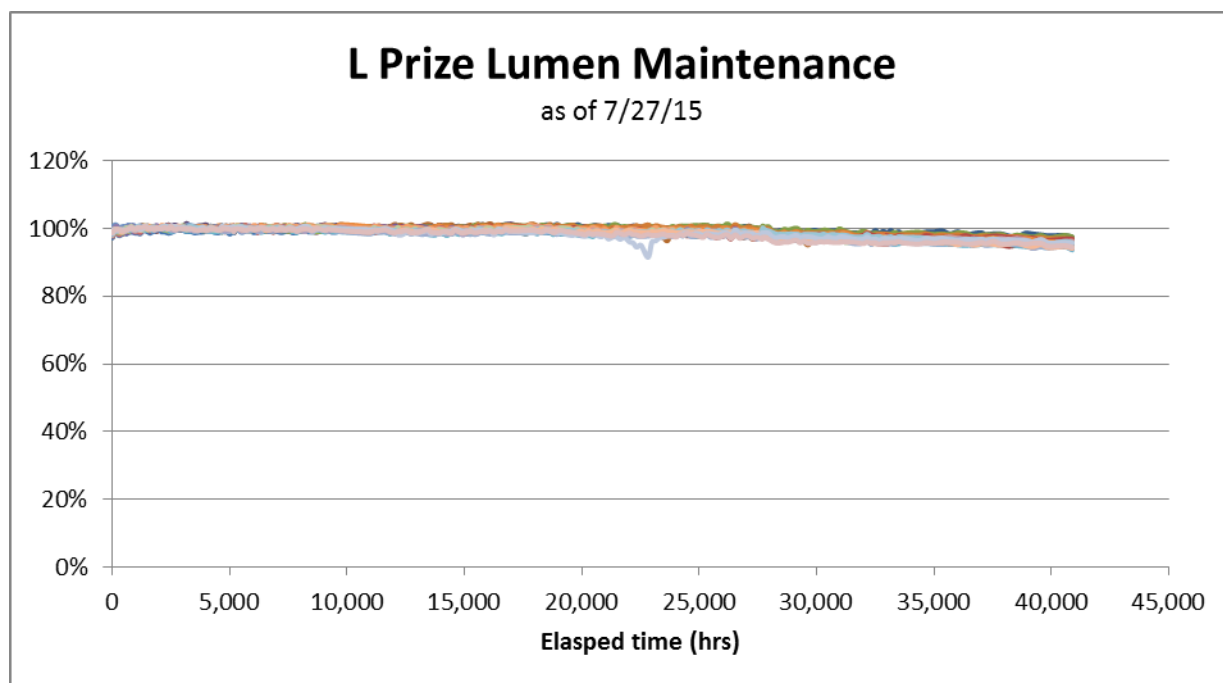


Figure 7.1. L Prize Lumen Maintenance After 40,000 Hours

Figure 7.2 below provides additional information on the lumen maintenance of the best, worst, and average of the 31-lamp sample.

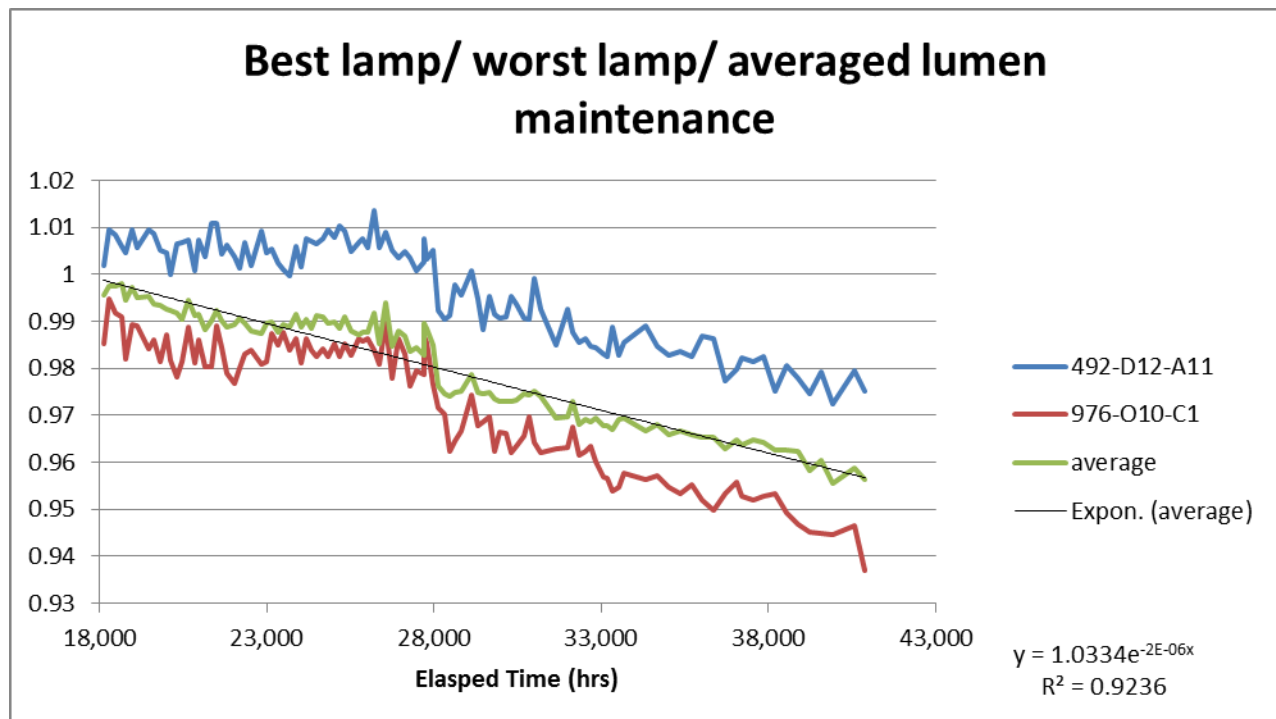


Figure 7.2. Comparison of Best, Worst, and Average L Prize Lumen Maintenance After 40,000 Hours

The best-performing lamp in the sample of 31 (blue line, 492-D12-A11) had 97.5% lumen maintenance, while the worst (red line, 976-O10-C1) was still at 93.7%.

The average chromaticity change ($\Delta u'v'$) for the lamps under test as of 40,890 hours was .00093, well below the original L Prize requirement (0.007 at 25,000 hours). Figure 7.3 shows that the worst-performing lamp (green line, 1897-Q7-C9) had a net change of .0030. This sample had exhibited erratic behavior earlier, spiking to more than .004 at around 23,000 hours, then returning to a very low level of change (less than .001) relative to initial values, then increasing again from .002 to nearly .004 over the period 28,000 to 37,000 hours before turning downward again. Meanwhile the best sample (purple line, 1403-B7-A6) showed a change of only .0002 at 40,890 hours, relative to initial value.

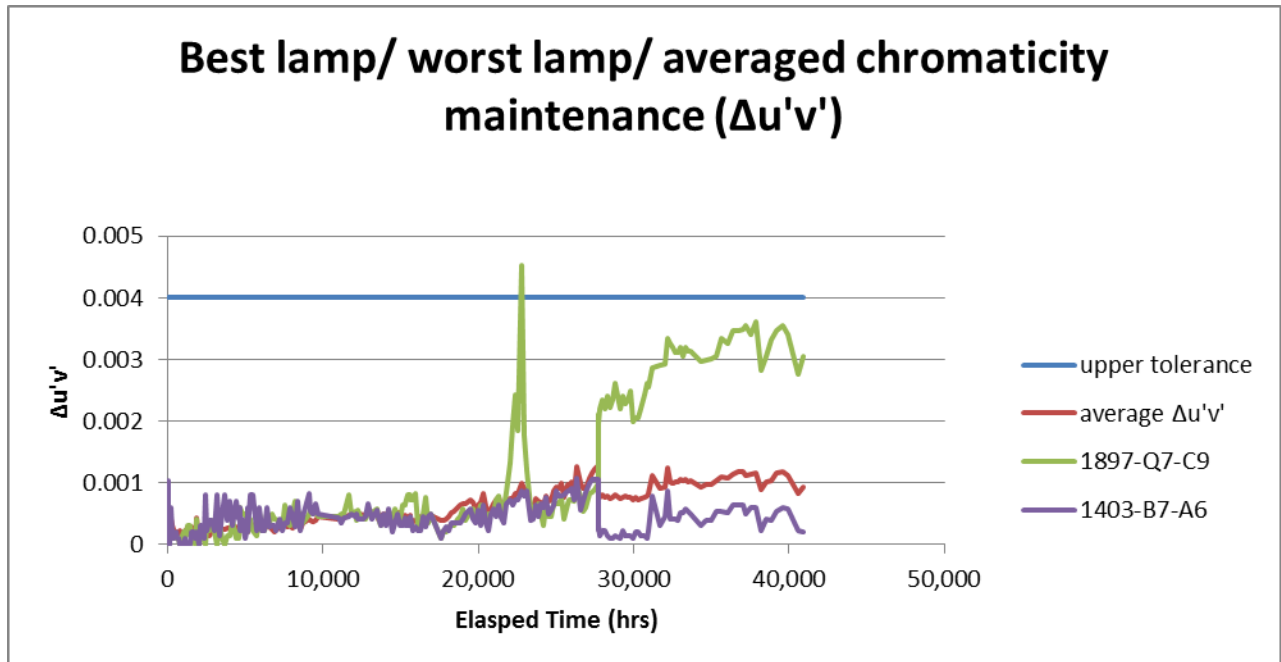


Figure 7.3. L Prize Best, Worst, and Average Chromaticity Maintenance Relative to Required Limit

On average, the L Prize-winning lamps continue to demonstrate very stable behavior in terms of both light output and chromaticity. Even the worst-performing sample in terms of lumen maintenance remains well above the parametric failure level (70%). No catastrophic failures have occurred to date.

8.0 Update: Results After 50,000 Hours

A sample of 31 L Prize winning lamps produced by Philips remain under test in the Accelerated Long-Term Test Apparatus (ALTA) at PNNL. These lamps have now passed 50,000 hours of operation under elevated ambient temperature (45°C) conditions. Measurements taken as of 50,182 hours on August 17, 2016, indicate average lumen maintenance for the 31 lamps under test was 93.0%. At that rate the average operating hours to reach 70% lumen maintenance would be 192,340 hours (using TM-21 forecasting methodology). Projected lumen maintenance should not be interpreted as projected lifetime, as it assumes all other components in the lamp continue to operate normally. Lamp failure is likely from electronic or other component failure before or after LED output has reached 70%.

Figure 8.1 summarizes the lumen maintenance of the 31 lamps under test, as of 50,182 hours of operation. Lumen maintenance is relative to the maximum output of the original lamp sample, which occurred at 1,996 hours, as described above in section 4.2.

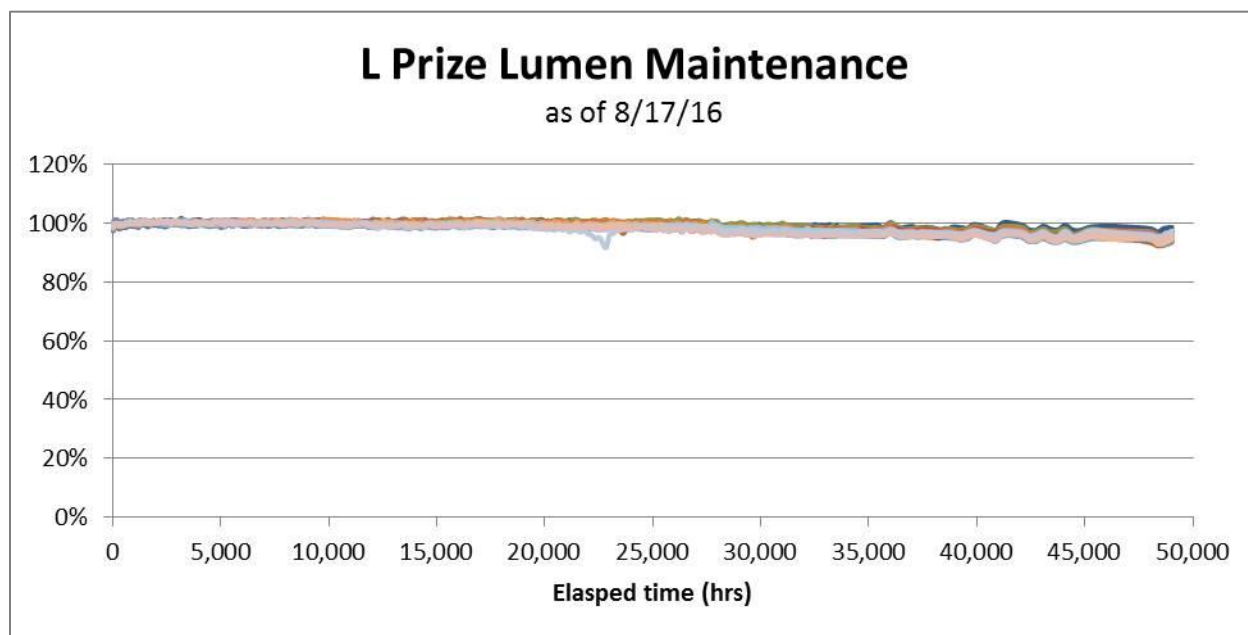


Figure 8.1. L Prize Lumen Maintenance After 50,000 Hours

Figure 8.2 below provides additional information on the lumen maintenance of the best, worst, and average of the 31-lamp sample.

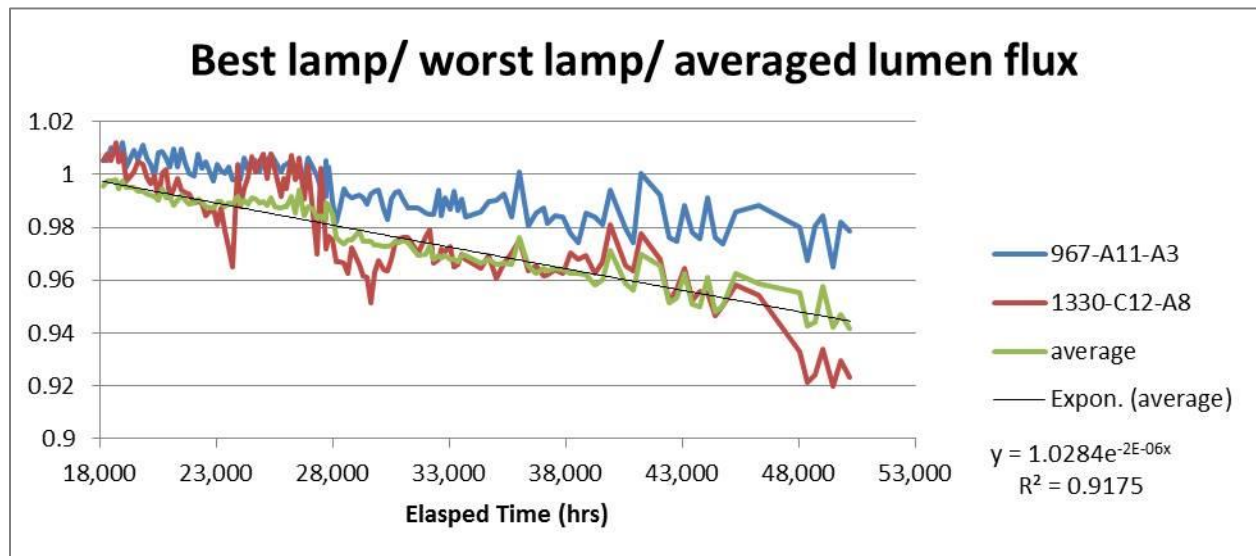


Figure 8.2. Comparison of Best, Worst, and Average L Prize Lumen Maintenance After 50,000 Hours

The best-performing lamp in the sample of 31 (blue line, 967-A11-A3) had 97.8% lumen maintenance, while the worst (red line, 1330-C12-A8) was still at 92.3%.

The average chromaticity change ($\Delta u'v'$) for the lamps under test as of 50,182 hours (red line) was .002, still below the original L Prize requirement (0.007 at 25,000 hours). Figure 8.3 below shows the worst-performing lamp (purple line, sample number 1897-Q7-C9) had a net change of .005. This sample had exhibited erratic behavior earlier, spiking to more than .004 at around 23k hours, then returning to a very low level of change (less than .001) relative to initial values, then displaying still erratic but overall increasing color shift between 30,000 and 50,000 hours. Meanwhile the best sample (green line, 456-A3-A4) showed a net change of only .001 at 50,182 hours, relative to initial value, and little variation over the test period.

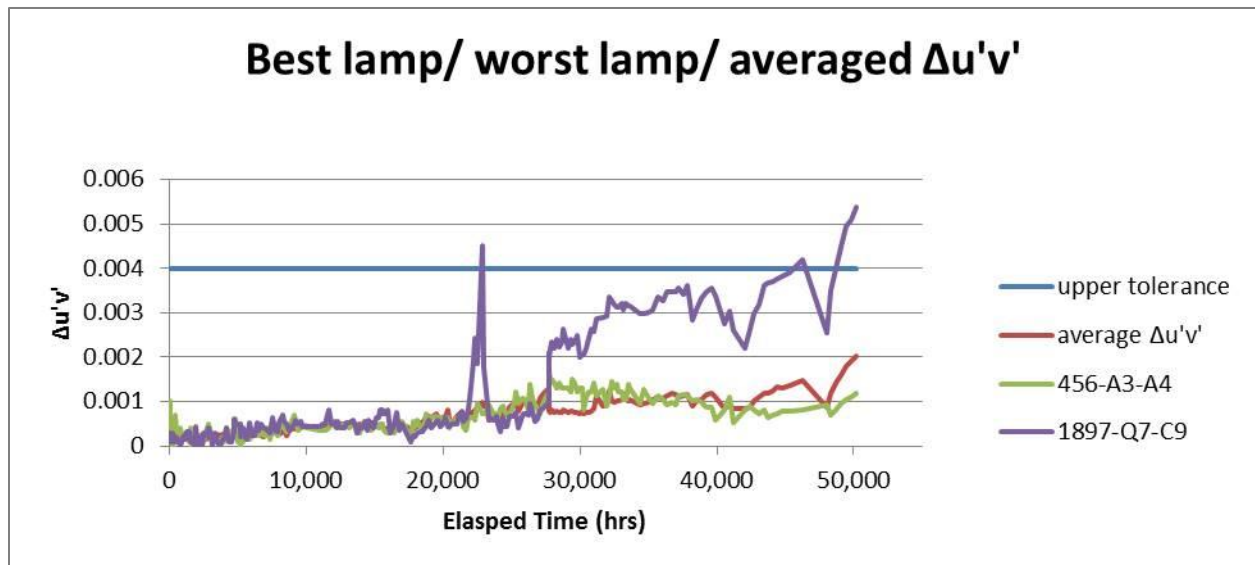


Figure 8.3. L Prize Best, Worst, and Average Chromaticity Maintenance Relative to Required Limit After 50,000 Hours

On average, the L Prize winning lamps continue to demonstrate very stable behavior in terms of both light output and chromaticity. Even the worst-performing sample in terms of lumen maintenance remains well above the parametric failure level (70%). No catastrophic failures have occurred to date.

9.0 References

1. Energy Independence and Security Act of 2007, Pub. L. no. 110–140 (2007). Print.
2. DOE – U.S. Department of Energy. Revision 1, June 26, 2009. *Bright Tomorrow Lighting Competition (L Prize)*. 08NT006643.
3. IES – Illuminating Engineering Society. 2008. “Approved Method: Electrical and Photometric Measurements of Solid-State Lighting Products.” IES LM-79-08.
4. IES – Illuminating Engineering Society. 2008. “Approved Method for Measuring Lumen Maintenance of LED Light Sources.” IES LM-80-08.
5. IES – Illuminating Engineering Society. 2011. “Projecting Long Term Lumen Maintenance of LED Light Sources.” IES TM-21-11.