Evaluating Color Rendition Using IES TM-30-15

IES TM-30-15 is a new system of several related measures and graphics that can be used together to effectively evaluate and communicate a light source's color rendering properties. The development of the method involved synthesizing multiple related research efforts and combining ideas into a single, cohesive system of objective information that can be used to aid decision-making processes, such as finding the preferred light source for a given application or evaluating the tradeoffs between efficacy and color rendering.

Developing new measures for evaluating color rendering has been a goal of lighting researchers—and the broader lighting industry—for decades. The utility of the Commission Internationale de l'Eclairage's (CIE's) general color rendering index (R_a, commonly called CRI)[1] has been demonstrated by its continued use for 50 years, but its limitations and deficiencies are becoming ever more prominent with the advent of solid-state lighting, which provides greater opportunity for spectral engineering than other source types. Within this context, the Illuminating Engineering Society (IES) formed the color metrics task group, operating under the IES Color Committee, which produced the technical memorandum *IES TM-30-15: IES Method for Evaluating Light Source Color Rendition*.[2]

Alone, TM-30 simply describes a voluntary calculation procedure. Achieving maximum benefit will require lighting manufacturers, specifiers, researchers, and regulators to examine its features, understand its unique ability to characterize multiple aspects of color rendition, and put it to use. Fortunately, TM-30 requires no additional measurement, so the burden to begin reporting and specifying the included measures is minimal. The greater challenge will be developing new sources that fulfill the potential of a more advanced characterization scheme, rather than just trying to maximize similarity to reference light sources.

This fact sheet describes the basic features of the TM-30 method and how they apply to various situations. More information can also be found at:

- DOE/IES Introduction Webinar: *http://energy.gov/eere/ssl/webinar-understanding-and-applying-tm-30-15*
- DOE/IES Technical Webinar: *http://energy.gov/eere/ssl/downloads/webinar-technical-discussion-tm-30-15*.
- Frequently asked questions page: *http://energy.gov/eere/ssl/tm-30-frequently-asked-questions*

TM-30 Calculation Engine

Like the CIE Test-Colour Method (i.e., CRI), TM-30 is a reference-based method, where a test source is compared to a reference source at the same correlated color temperature (CCT). This approach is widely understood and compatible with the lighting design process where color temperature

1 For more information, see CIE 13.3-1995 or Tutorial: Color Rendering and Its Applications in Lighting (Houser KW, Mossman M, Smet K, Whitehead L. 2015. Leukos. DOI:10.1080/15502724.2014.989802.)

2 IES TM-30-15 is available at https://www.ies.org/store/product/ies-method-for-evaluating-light-source-color-rendition-3368.cfm. Excel tools for performing calculations are supplied with the document.

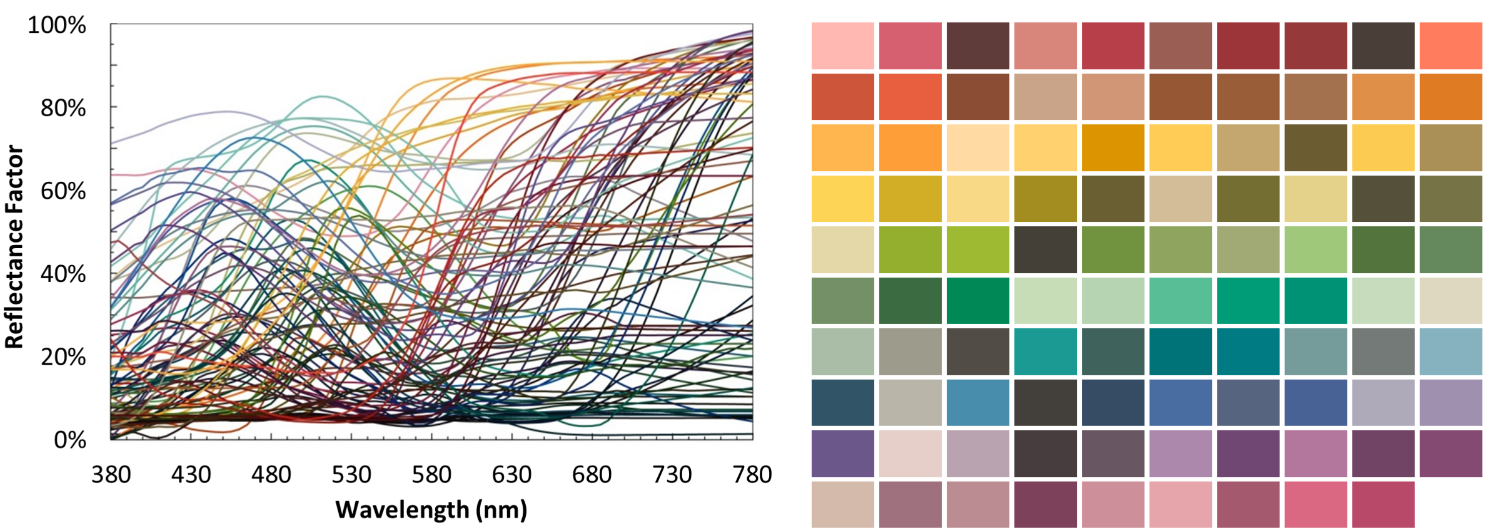

Figure 1. At the core of TM-30-15 is a set of 99 color samples that are used to characterize the difference between the test source and reference illuminant. In the calculation procedure, the samples are represented by their reflectance function (left), which is the percent of radiant flux reflected at each wavelength. The color swatches on the right illustrate the approximate appearance of the samples under the 5000 K reference illuminant.

(or chromaticity[3]) is considered first, separately from color rendering. That is, chromaticity and color rendering can be considered independently. The reference illuminants used for the TM-30 method are the same as used in the CIE CRI system—blackbody radiation and the CIE D Series models of daylight[4]—but the two are blended proportionally between 4500 K and 5500 K to avoid the abrupt changeover at 5000 K that exists for CIE CRI. This has a minimal affect on any score compared to the previous method, but is more appropriate given the increasing prevalence of color-tunable sources.

Although it uses a familiar reference-based approach, there are many underlying technical differences between TM-30 and CIE CRI, along with a more comprehensive suite of numerical and graphical outputs. Table 1 documents the major differences between the CIE CRI system and the TM-30 system.

At the core of the TM-30 method is a new set of color evaluation samples (CES). The 99 CES, illustrated in Figure 1, were statistically selected from a library of approximately 105,000 spectral reflectance function measurements for real objects, which included paints, textiles, natural objects, skin tones, inks, and more. The sampling methodology dictated even coverage of color space by dividing the color volume into cubes and selecting one sample from each, with the additional stipulation that the combined set minimized the difference in sensitivity for individual wavelengths.[5] This minimization of sensitivity is important because non-uniformity can lead to false optimizations based on features of the evaluation samples that are not present in typical applications. Compared to the eight samples used to calculate CIE's general color rendering index R_a,[6] the 99 CES mitigate selective optimization, so the output values are a better prediction of real-world performance.

The other major technical difference between CIE CRI and TM-30 is the use of the CAM02-UCS (uniform color space), which offers substantially better uniformity than the CIE 1964 U*V*W* color space. The improved uniformity allows for more accurate calculations of color difference, which in turn means more accurate results, as well as a reduction in undesirable CCT-related effects.[7] The CAM02-UCS also includes an improved chromatic adaptation transformation,[8] which means more effective calculations when the test source is off the blackbody locus (i.e., has a greater absolute value for D_{uv}).

3 Color rendering and chromaticity (or the appearance of the light itself) are distinct characteristics, although the two may be considered together in one color quality specification.

4 Blackbody radiation and the CIE D Series illuminants are both mathematically derived spectral power distributions that represent a heated mass at increasing temperatures and different phases of daylight, respectively. They can be calculated at any correlated color temperature (CCT), and have long served as references in the lighting industry. Although they are natural sources, they are not necessarily ideal sources. TM-30 details the deviation from these references, providing more technical information for comparison when choosing a light source.

5 For more on the sample selection procedure, see: David et al. 2015. Development of the IES Method for Evaluating Light Source Color Rendition. Opt. Express 23, 15888-15906. DOI: 10.1364/OE.23.015888

6 R_a is commonly referred to simply as CRI, but it is technically just one component of the CIE Test-Colour Method, which includes six special color rendering indices and one general color rendering index. There are a total of 14 color samples in the system, but only eight are used for calculating R_a.

7 Smet et al. 2015. Why Color Space and Spectral Uniformity are Essential for Color Rendering Measures. Leukos 12(1). DOI:10.1080/15502724.2015.1091356

8 Chromatic adaptation transformations are employed in color rendering metrics and color appearance models to account for differences caused by the test and reference conditions eliciting different chromatic adaptation in the viewer, rather than differences caused by the distortion of the actual colors.

Table 1. Comparison of the CIE Test-Colour Method (commonly known as CRI) and IES TM-30-15.

	CIE 13.3-1995 (CRI)	IES TM-30-15
Year of Issuance	1965, 1974 (Revision), 1995	2015
Color Space	CIE 1964 U*V*W*	CAM02-UCS (CIECAM02)
Number of Color Samples	8 general (for R_a) plus 6 special (for R_is)	99
Color Volume Coverage	Limited	Full and equal
Saturated Samples	No	Yes
Sample Types	Munsell samples only (limited pigments)	Variety of real objects
Sample Spectral Uniformity	No	Yes
Reference Illuminants	Blackbody radiation, CIE D series	Blackbody radiation, CIE D series
Reference Transition	Sharp at 5000 K	Blended between 4500 K and 5500 K
Output Measures	General index, R_a (fidelity) 6 special indices, R_i (fidelity)	Fidelity Index, R_f Gamut Index, R_g Color Vector/Saturation Graphics 16 hue-based fidelity indices 16 hue-based chroma indices 1 skin-specific fidelity index 99 individual fidelity values
Score Ranges	Max 100 with no lower limit, variable scaling	0 to 100, consistent scaling

Evaluation Measures

Although the technical differences between the IES TM-30-15 and the CIE CRI system are noteworthy, they may be entirely invisible to many users. Of greater practical relevance is the new information that TM-30 provides to supplement the improved average value for color fidelity.

TM-30 delivers a hierarchy of information, ranging from high-level averages to highly detailed characterizations for specific color samples. Each output of TM-30 is described in Table 2, with additional discussion in the subsequent sections. It is important to understand the meaning and limitations of each of the TM-30 outputs.

Average Values

Color rendering is undoubtedly a complex subject, and any color rendering measure must make compromises in distilling complex information to a digestible number or graphic. For most users, the most familiar TM-30 value will be the fidelity index, R_f, which characterizes the average difference in color for the 99 CES under the test and reference conditions. It is analogous to CIE R_a, although the underlying

Table 2. Description and use of the measures described in IES TM-30-15.

Measure	Abbreviation	Description
Fidelity Index	R_f	Analogous to CIE R_a (CRI). Characterizes the average color shift of the 99 CES to characterize the overall level of similarity between the test source and reference illuminant. Values range from 0 to 100.
Gamut Index	R_g	Compares the area enclosed by the average chromaticity coordinates in each of 16 hue bins to characterize the average saturation level of the test source compared to the reference illuminant. A neutral score is 100, with values greater than 100 indicating an increase in saturation and values less than 100 indicating a decrease in saturation. The range in values grows as fidelity decreases.
Color Vector Graphic		Provides a visual representation of hue and saturation changes based on the average rendering in each hue bin, relative to the reference. The graphic provides a quick understanding of how different hues are rendered in different ways.
Color Saturation Graphic		Provides a simplified visual representation of only saturation changes based on the average performance in each hue bin.
Hue Fidelity Indices	$R_{f,hj}$ (j = 1 to 16)	Provides a numerical characterization of color fidelity in each of 16 hue bins (j), which can be used to evaluate how similarly the test source renders reds, yellows, greens, blues, or in-between hues compared to the reference. Values range from 0 to 100. Specific values may be used to supplement average values if one hue type is of particular concern. Specifying limits for all values is also possible. These scores are analogous to the special indices of the CRI system (e.g., R_9), but are more robust because they combine several samples with different spectral features.
Chroma Change by Hue Indices	$R_{g,hj}$ (j = 1 to 16)	Provides numerical values for relative chroma change in each of 16 hue bins (j), which can be used to evaluate saturation (positive values) or desaturation (negative values) of reds, yellows, greens, blues, and in-between hues compared to the reference. Supplementary criteria could be set for all values, or just specific values, such as red (bin 1).
Skin Fidelity Index	$R_{f,skin}$	Characterizes the similarity of skin tones (CES15 and CES18) as rendered by the test source compared to the reference source. Values range from 0 to 100. $R_{f,skin}$ can be used to supplement other values when skin is an important consideration.
Sample Fidelity Indices	$R_{f,CESi}$ (i = 1 to 99)	Characterizes the similarity of each CES (i) as rendered by the test source compared to the reference source. Values range from 0 to 100. Individual values may have little predictive power for other objects, but examining the scores in a combined chart can indicate the source's object-to-object consistency.

math and new sample set make it a tougher test that prevents selective optimization (commonly called "gaming").[9] R_f has a range of 0 to 100, with 100 indicating an exact match with the reference. As an average value, R_f does not provide an indication of what types of colors are most distorted, or if the distortions are increases in saturation, decreases in saturation, or hue shifts. This is where the gamut index, R_g, begins to fill the void of fidelity-only evaluation.

R_g documents the overall increase or decrease in saturation relative to the same-CCT reference illuminant, with a score of 100 indicating the same average gamut area. Scores above 100 indicate an average increase in saturation, whereas scores below 100 indicate an average decrease in saturation. R_g can help capture the effect of the neodymium incandescent lamp, for example, which trades fidelity for increased saturation. Because they are based on the same reference illuminant and underlying calculations, R_f and R_g form a comprehensible two-axis system (Figure 2) where the relationship between fidelity and saturation can be evaluated and compared across multiple sources. In this way, the exact reference becomes less critical than in a system where fidelity is the only aspect evaluated; because one can understand *how* a source differs from the reference, rather than just the magnitude of the difference, there is a way to judge the suitability of a source for a given application.

9 Although CIE R_a and TM-30 R_f have similar scales, many sources that were previously optimized for CIE R_a are no longer optimized when considering R_f, and thus have lower scores. This is particularly true of sources with narrow peaks, such as triphosphor fluorescent lamps, which were specifically designed to balance efficacy and CIE R_a score.

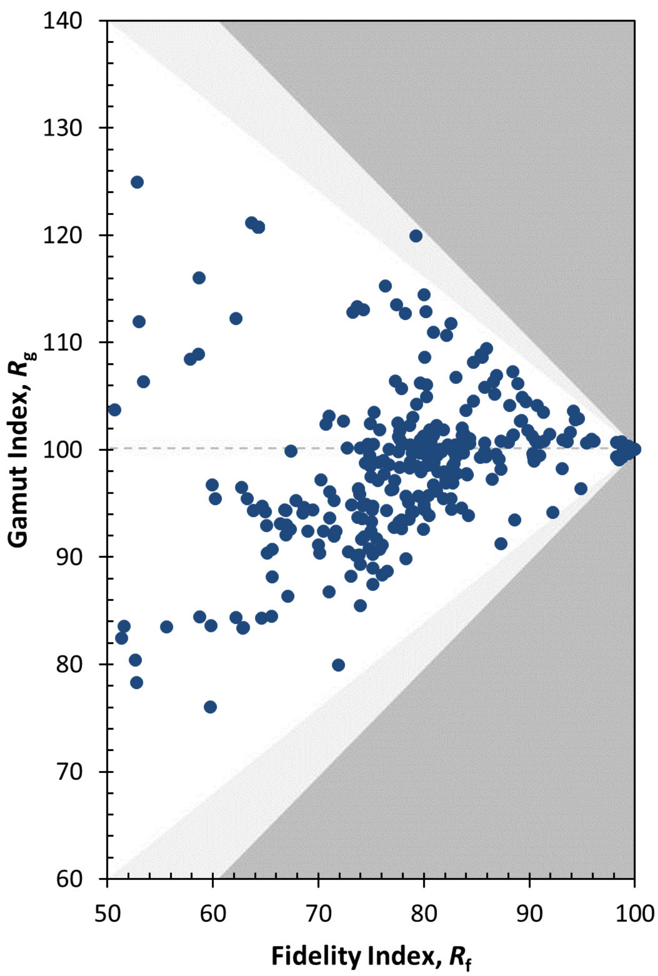

Figure 2. R_f and R_g form a complimentary two-measure system, which can be plotted to visually illustrate the tradeoff between fidelity and saturation. The blue points represent real sources, with the shaded areas approximating combinations that are not possible for sources on the blackbody locus (light gray) or not classified as white light (dark gray).

The suitability of a source for a given application—or color preference—is a key consideration in color rendering metrics. TM-30 does not include a single-value measure that characterizes preference, because the most suitable source is not universal to all applications or populations. Some applications, like a hospital, require high color fidelity. Other applications, such as retail, might be better served with a boost in saturation that makes the colors of the goods more vibrant and eye-catching. With TM-30, such decisions can be made in a purposeful and educated manner.

Graphical Representations and Detailed Values

In many applications, the average values R_f and R_g may be sufficient for specifying the most suitable source, and will allow manufacturers to design and market novel light sources more effectively. However, as with R_9 in the CIE CRI system, supplemental information about specific hues may be important in color-critical applications. Hue-specific information can be readily gleaned from the color vector graphic (Figure 3), which shows a combination of saturation changes and hue changes of the test source (red line) versus the reference (black circle). The color saturation icon shows similar information in a simplified form, which may be easier to convey but does not show hue shifts. For example, the graphics can easily be used to distinguish sources that saturate reds from those that saturate greens, two features that may be related to preference and efficacy.

In some cases, it may be necessary to specify numerical values for specific hue regions, such as red. The measures $R_{f,hj}$ and $R_{g,hj}$ (where j indicates the number of the hue bin) correspond to the fidelity and chroma change for the average chromaticity in each of 16 hue bins. The bins are numbered in a counterclockwise order beginning with the positive a' axis of the CAM02-UCS, as shown in Figure 4. For example, $R_{f,h1}$ corresponds to the fidelity in hue bin 1, which is red. $R_{f,h1}$ is correlated with the CIE R_9 value, but does not share the same scale.[10] All fidelity values in TM-30 are scaled from 0 to 100, with 100 indicating an exact match with the reference.

If still further detailed information is needed, one can examine the fidelity value for any of the 99 CES, or a plot of all the values (Figure 5). Examining all the values

*10 R_9 has an unusual scale due to the non-uniformity of the CIE 1964 U*V*W* color space, in which it is calculated.*

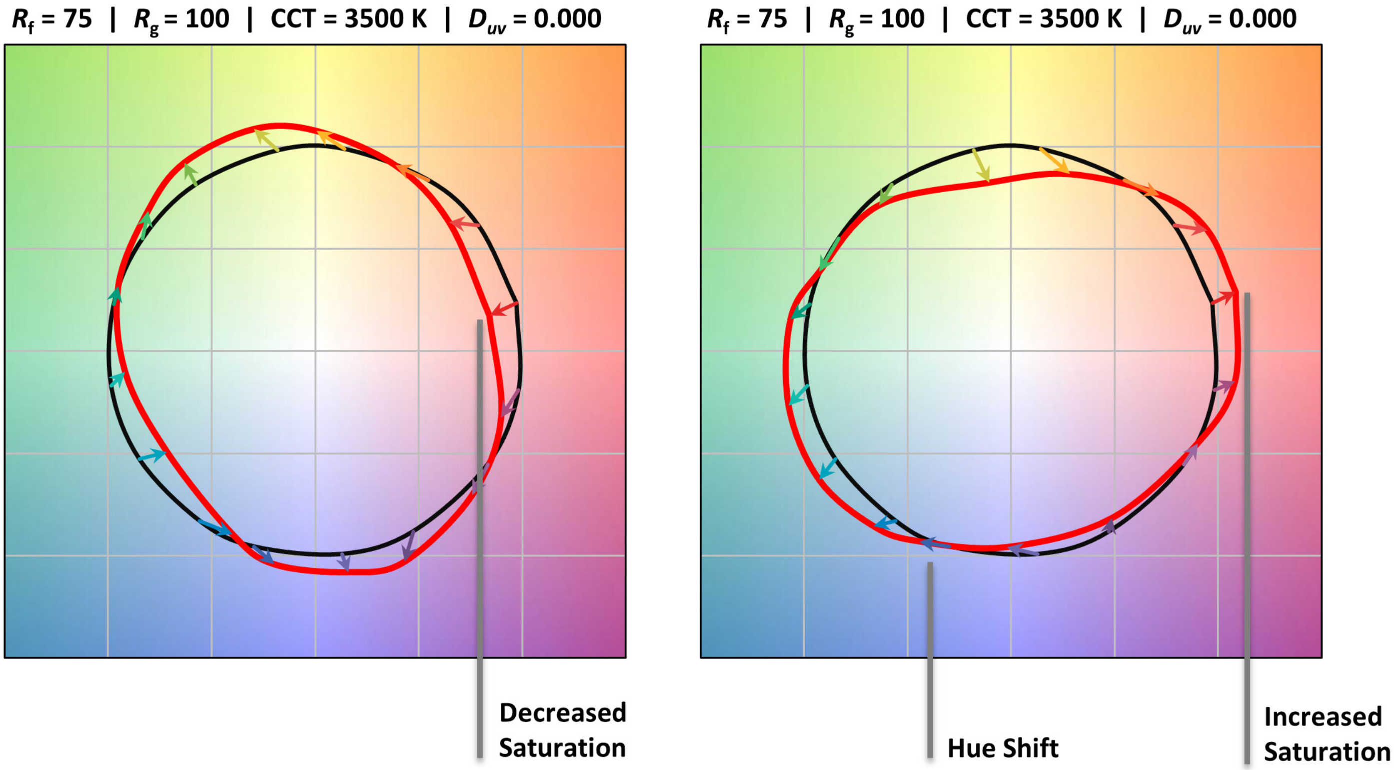

Figure 3. The color vector graphic illustrates average hue and saturation changes in each of 16 hue bins. The reference source is normalized to a black circle, whereas the test source is represented by the red line. Arrows show specific distortions: arrows pointing in from the black circle indicate a region of desaturation, arrows pointing out from the black circle indicate a region of increased saturation, and arrows tangential to the black circle indicate hue shift only.

simultaneously can help one understand the consistency of color rendering. If a source has drastically different values for adjacent CES, it indicates that it may render similar colors very differently, relative to the reference. Other sources, usually those with a more continuous spectral power distribution, may have more consistent performance from sample to sample.

One other measure, $R_{f,skin}$, is available to document a source's performance in rendering human skin tones. $R_{f,skin}$ is the average of CES15 and CES18, which are two skin reflectance measurements that provide the highest correlation in fidelity score compared to a large library of human skin tones.

Design Criteria and Guidance

IES TM-30-15 documents a calculation procedure; it does not specify criteria (i.e., minimum values) for any measure or stipulate how the measures should be applied. Guidance and criteria can be developed by appropriate groups using past experiences and industry feedback. Guidance and criteria for familiar characterizations, such as fidelity, may come easy, whereas guidance for new characterizations such as gamut area or gamut shape may take more time as the industry adapts to the new degrees of freedom.

Although many in the lighting industry are eager to see new resources for evaluating color rendering, the transition period does not have a defined timeline. If users find TM-30 suitable, it will be adopted rapidly. Some manufacturers are already supplying TM-30 information, and specifiers can calculate TM-30 measures from the spectral power distribution (SPD) of any given source.

Conclusion

TM-30 synthesizes decades of research on color rendering to provide a suite of objective characterizations that can be used to aid in choosing the best source for an application and in engineering novel light sources. It remedies flaws and limitations of the widely used CIE CRI method, while providing complementary and more detailed information.

To obtain maximum benefit from TM-30, one must move beyond a higher-number-is-better mentality; the best source depends on the needs of the application, and isn't always the reference illuminant. In this way, TM-30 follows a now common approach of solving complex problems (i.e., characterizing color rendering), with a data-driven approach. It is a powerful new tool, but also respects and supports the role of design and engineering.

The increased availability of data to characterize color rendering can benefit specifiers, manufacturers, and researchers alike, although transitioning from the familiarity and simplicity of CIE CRI will take considerable effort, with an undefined timetable. Pending needs include establishing

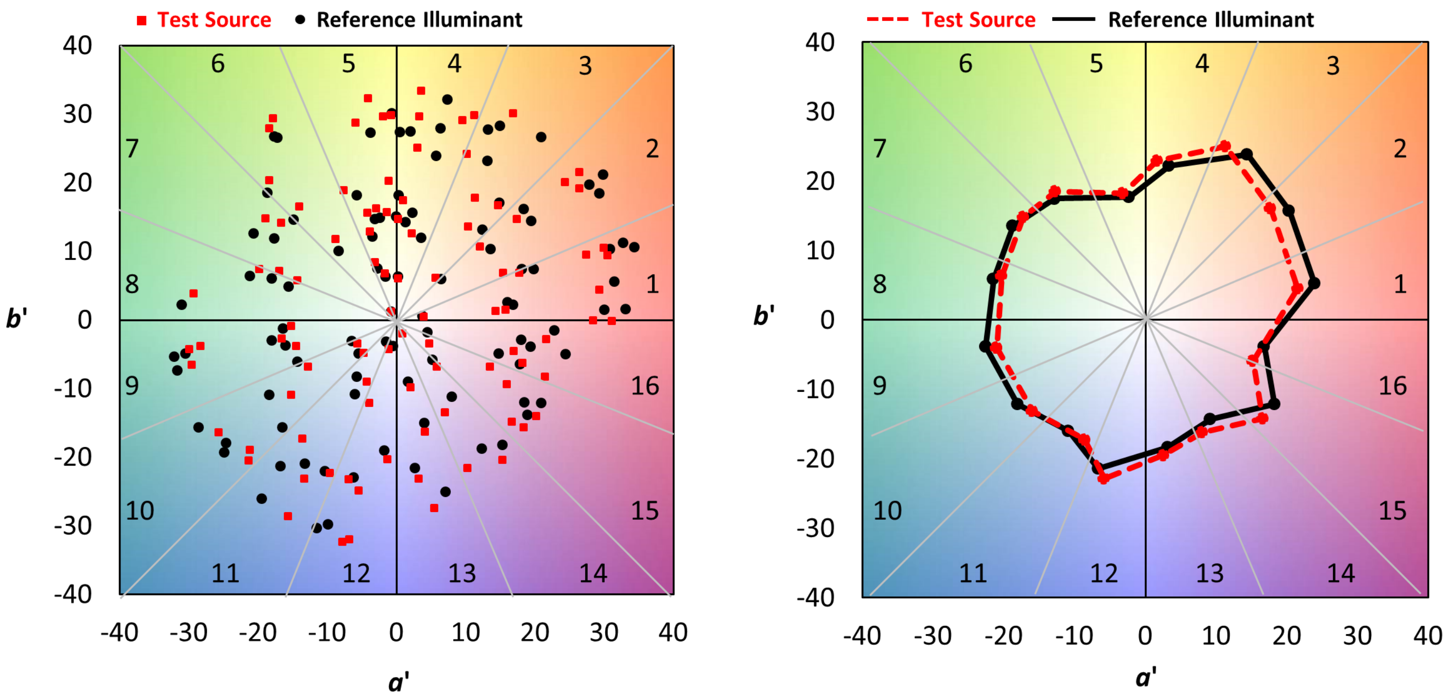

Figure 4. The 99 samples can be binned into 16 groups based on their (a', b') coordinates (left). The equally-sized bins are numbered sequential beginning with the positive a' axis. Within each bin average coordinates can be calculated based on the test source and reference illuminant (right). R_g is a comparison of the area encompassed by the two resulting polygons.

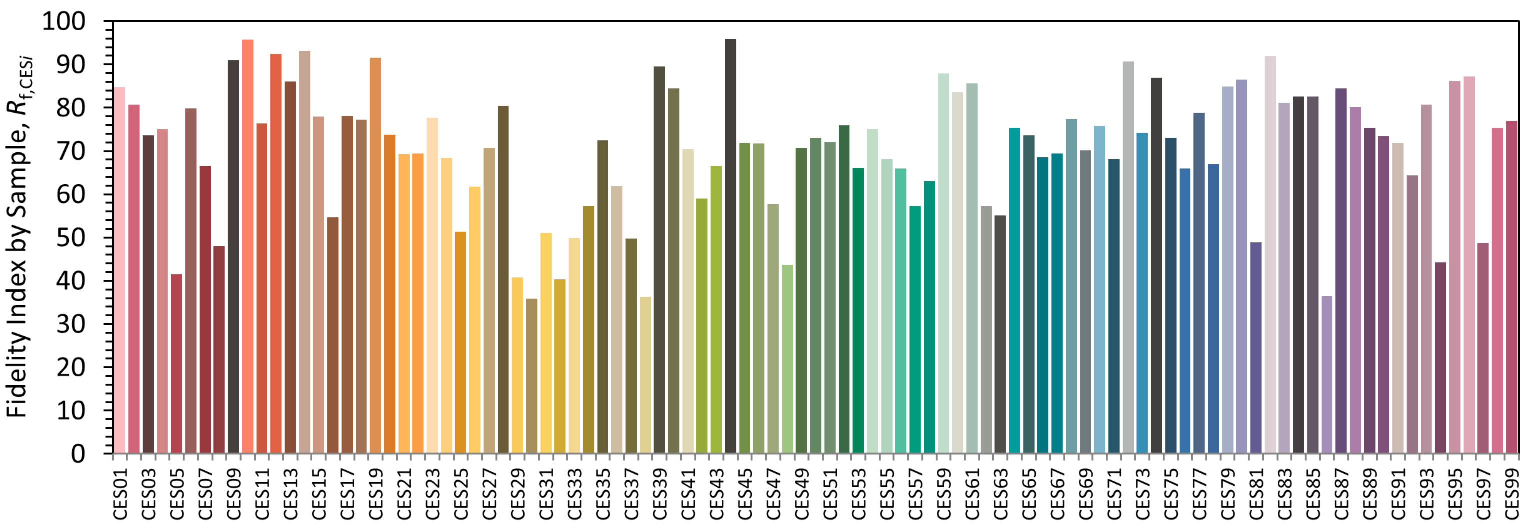

Figure 5. The individual sample fidelity scores ($R_{f,CESi}$) can be plotted together, in hue order, to create a “signature” for the source. This signature indicates the consistency of color rendition. That is, are color evaluation samples with similar hues rendered with the same amount of shift relative to the reference illuminant?

specification criteria, design guidance, and personal experience with the meaning of the measures and how they can be used to match a light source with an application. Additionally, an infrastructure of data exchange needs to grow through the cooperation of manufacturers, specifiers, researchers, and regulators.

U.S. DEPARTMENT OF
ENERGY
Energy Efficiency &
Renewable Energy

For more information, visit: ssl.energy.gov
Direct fact sheet feedback to: SSL.Fact.Sheets@pnnl.gov

PNNL-SA-114005 • October 2015

Printed with a renewable-source ink on paper containing at least 50% wastepaper, including 10% post consumer waste.