Spectral resolution in colour rendering assessments and LED lighting

Luke L. A. Price

Public Health England, Chilton, DIDCOT, OX11 0RQ, UK
Email: luke.price@phe.gov.uk

The colour rendering properties of LED domestic lighting solutions were assessed. The majority of the sample were white LEDs emitting partially phosphor-converted blue LED light. Compact fluorescent (CFL), incandescent, and tungsten-halogen (TH) lamps were included for comparison. In addition to familiar fidelity and gamut metrics, an often neglected third class of metric which evaluates spectral richness for colour rendering was examined, and used as a tool to consider the importance of spectral resolution in colour rendering.

Received 06 September 2017; revised 01 November 2017; accepted 01 December 2017
Published online: 26 March 2018

Introduction

In 2015, the Chartered Institution of Building Services Engineers (CIBSE) and the Society of Light and Lighting (SLL) commissioned a study of human responses to LED lighting [1] to address the areas where data were lacking and apply the latest knowledge of human responses to light and international exposure guidelines. Measurements were to be taken of a range of LED lighting products on sale to the public, businesses and commissioners of outdoor lighting projects in the UK.

The primary aim of this paper concerns developments into the next generation of colour rendering metrics, with a secondary aim to present the colour rendering results for LEDs available to consumers in the UK and elsewhere. Whilst it is not intended to give a full introduction to the subject area, it is necessary to introduce three areas of background: the recent Illumination Engineering Society (IES) metrics in report TM-30-15 and other widely used colour rendering metrics; spectral dispersion, spectral resolution and metrics for spectral richness; and a metric adapted from another field, the Gini coefficient, which is used to measure wealth inequalities. Other publications can provide excellent introductions to colour science [2-4] and the history of colour and colour rendering metrics [3-5].

Colour rendering metrics in common use

At about the same time as this project was underway, TM-30-15 was put forward with significant backing within the lighting community, as a replacement for the Color Rendering Index (CRI) and Color Quality Scale (CQS) systems of metrics, and the CRI-related Gamut Area Index (GAI) [6-8]. TM-30-15 uses 99 spectral reflectances [9-10], with uniform coverage in both the colour and wavelength spaces, and provides fidelity index, R_f and relative-gamut index, R_g.

CQS itself was intended to be a replacement for the CRI, and provides a number of metrics, including a widely quoted fidelity index, Q_a, based on a combined score for reproduction of 15 high-chroma spectral reflectances. The similar CRI metric R_a is based on just 8 spectral reflectances, which are pastel-like in appearance.

A fidelity index quantifies the absolute distances in a colour space, for a range of spectral reflectances, between the test lamp and a suitable reference lamp. The solid lines ending in diamonds in Figure 1 show these distances for the eight reflectances in the CRI; as this is known as uv-space, the distances are called *duv*1 to *duv*8. The choice of reference lamp typically depends on the correlated colour temperature, T_{CP}, of the test lamp. This avoids variations in T_{CP} having a significant impact on the colour rendering score, as T_{CP} is a good explanatory variable for the main colour differences between the range of naturally lit conditions to which a scene lit by a given lamp might be faithful (hence "fidelity").

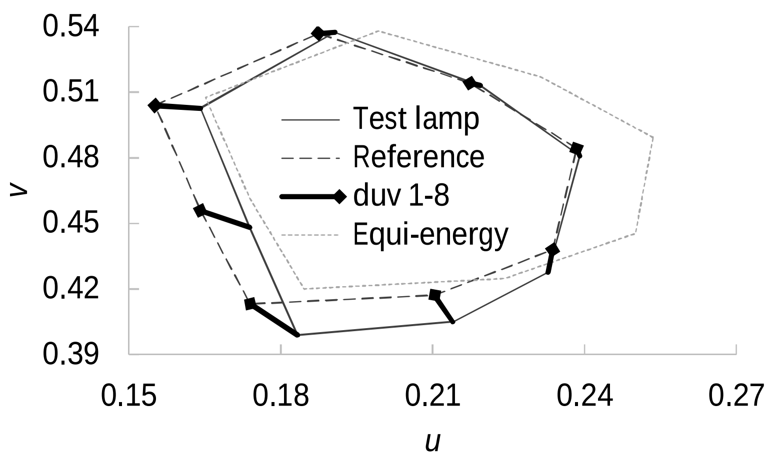

Figure 1: The representation of the distances in the calculation of CRI R_a and the areas in the GAI calculation. (T_{CP} = 6515 K, R_a = 84.8%, GAI = 95.3%).

The eight spectral reflectance of the CRI are also used to calculate the gamut area index (GAI), which corresponds to the area inside the dotted line in Figure 1 for the test lamp divided by the same area for an equi-energy spectrum [11], rather than the reference spectrum. The gamut area index represents the vividness with which the light source renders object colours [11].

Despite the publication of TM-30-15 and the adoption of a closely related system by the CIE [12], at present many practitioners continue to use the outdated CRI system. The method of investigation here relates to colour fidelity metrics; it helps to validate many of the latest metrics, provides a useful explanation of a shortcoming in older metrics such as CRI and CQS, and identifies a necessary characteristic in a successful metric. It does not obviate the need for gamut area assessments.

Spectral richness

The International Commission on Illumination (CIE) has separated into two technical committees the work of developing an improved colour fidelity index to replace CRI Ra and the work of investigating the scope for other metrics, particularly related to colour preference [13]. This note only addresses the former area, and in particular the standardisation of the subjective experience of illumination is considered to be scientifically inappropriate. A subjective matter may be investigated to uncover non-subjective elements masked within the whole, but purely subjective elements of a response to lighting are not reproducible.

The reasons for wishing to replace CRI are closely linked with the development of LED lighting and other modern lighting technologies, and the need for fair rankings between different types of light source with different spectral characteristics. The new metrics use modern colour spaces, in which the mathematical distances between points more closely reflects their perceptual difference, but what is most notable is the trend to use increasing numbers of spectral reflectances in an attempt to solve this problem of fairness.

Before the emergence of CRI [e.g. 5], colour rendering assessments often considered non-colorimetric properties of the relative spectral power distribution (RSPD) directly, i.e. ‘spectral band methods’. The role of spectral band methods will be considered further in the discussion. However, note that the main method in use before CRI employed brightness-weightings and only eight spectral bands [e.g. 8] with a low wavelength resolution.

The question of required wavelength resolution is closely linked with object reflectances. To consistently reproduce colour information from objects with steep gradients or narrowband features in their reflectances, a finer measurement resolution is needed than for smooth and broad reflectances. For a given spectral range, the number of bins is inversely proportional to the wavelength resolution, see Figure 2.

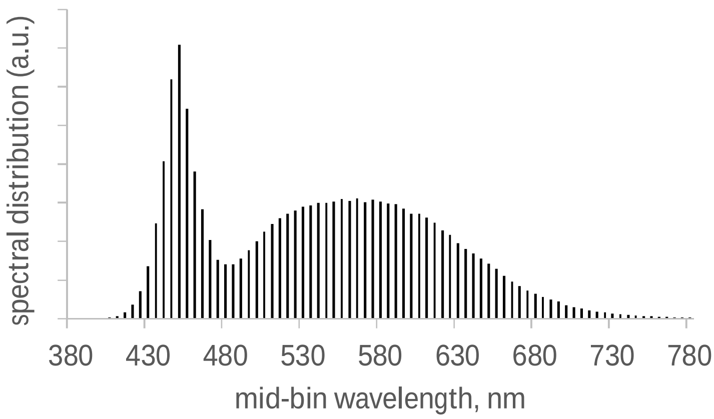

Figure 2: The spectral energy distribution of an LED. The LED spectrum is divided here into 81 spectral bins of width 5 nm.

As the bin count reduces, information is thrown away and no metric can be devised that would reintroduce information from coarsely resolved spectra. For example, the CRI R_a uses eight reflectances, which reduces the information content of the RSPD to the equivalent of, or coarser than, just eight 50 nm bins, based on (780-380 nm)/8 and TM-30-15 uses 99 samples, approximately equivalent to, or coarser than, 4 nm bins, based on (780-380 nm)/99. Bouma's spectral band method [14] was too coarse for present-day lighting technologies, with around 50 nm per bin (like CRI).

Figure 2 shows the division of the spectral energy distribution from 380 nm to 780 nm into 81 spectral bins of 5 nm each (note the actual wavelength range is thus 377.5 nm to 782.5 nm). This subdivision can be used to calculate any of the metrics discussed, although a 1 nm resolution (379.5-780.5 nm) is supported for TM-30-15 inputs. Subject only to measurement constraints, an arbitrarily fine resolution is theoretically possible for the spectral richness metrics discussed.

Two more modern attempts have sought to draw attention to unweighted spectral richness, the full spectrum color index (FSCI) and spectral entropy metrics [8, 15]. Both have the advantages of objectivity and a robustness against "gaming", where the spectral distribution is tweaked to produce better scores without necessarily better colour rendering. They are correlated, although not exactly, as the FSCI method rewards certain bin-orders, and actually penalises smoothness (presumably unintentionally). Calculation time is another limitation for FSCI, whereas spectral entropy scores need several decimal places to reveal the spectral richness clearly.

Before continuing, it is important to clarify why spectral richness can be used to evaluated colour fidelity. Colour fidelity metrics evaluate relative spectral richness compared to a reference illuminant. This is achieved by adopting colorimetric weighting systems that evaluate colour differences (or colour dispersion) between the test illuminant and a low-dispersion (or colour faithful) reference illuminant. Additionally, it has previously been shown [15] that spectral richness measured using spectral entropy is positively correlated to the colour fidelity metrics in the CRI and CQS systems as well as FSCI.

The metrics of money

Spectral richness metrics are analogous to econometrics for wealth inequalities. The aim of both is to assess the dispersion in a set of values; for lamps the concern is whether colour is distorted by the spectrum, for economists the concern is whether wealth is adequately and appropriately distributed according to some political philosophy. Statistics such as the richest 5% of the population own 95% of the wealth, might be quoted for any given country; but these statistics may be chosen with some subjectivity. Why choose 5% and/or 95% as the important measures? To avoid this, economists have developed robust metrics capable of characterising the dispersion of wealth within the whole economy. Perhaps the most widely quoted of these is the Gini coefficient [16-17].

Due to the mathematical similarities, the Gini coefficient can be used directly as a spectral richness metric. The similarities are so close that Spectral Entropy, H_{Sp}, [15] also already exists independently in wealth dispersion economics within the generalised entropy index [18]. The existing inequality metrics of economics are all bin-order independent, and due to their abstract nature can be used to examine the relevance of bin-size, *i.e.* wavelength resolution, in colour rendering metrics, as has been done here.

Using the Gini coefficient as inspiration, the spectral richness of a lamp, or other light source, can be calculated using the Lamp Gini index. The bins from an RSPD such as in Figure 2 are arranged in order of increasing size, and accumulated to form a Lorentz curve. This is shown in Figure 3, where the dashed line is the Lorentz curve for the spectral distribution shown on the same chart as a dotted line. The same process for the equi-energy spectrum produces the solid line. The Lamp Gini index, G_L, is given by the ratio of areas below these two Lorentz curves:

$$G_L = A / (A + B) \quad (1)$$

where A refers to the area above the dotted line, but below the solid line, and B to the area below the dotted line (in Figure 3). It is easy to show that G_L is equal to $2A/n$, where n is the number of wavelength bins. To avoid having to reorder and accumulate the values, there is a convenient mathematical result for quick calculations:

$$G_L = (2 \sum_{i=1}^{n} p(\lambda_i) r(\lambda_i) - 1 - n) / n \quad (2)$$

where $r(\lambda_i)$ is the unique rank of the ith value of the RSPD $p(\lambda_i)$, which can also be thought of as the probability distribution of emitted spectral power for n consecutive wavelength bins denoted by λ_i. i = 1, 2, 3, ..., n relate to λ_i = 380 nm, 385 nm, 390 nm, ..., 780 nm.

The advantage of equation (2) is that it takes less processing time, and considerably less than the metrics of CRI, GAI, FSCI, CQS and TM-30-15. With this metric it is important to realise that a low score is desirable, just as more equal wealth distributions are often held to be superior.

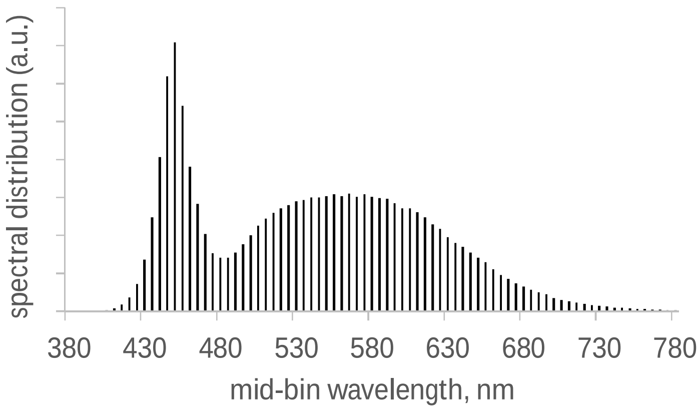

Figure 3: Visualizing the Lamp Gini index for an LED RSPD (dotted, not to scale). Sorting wavelength bins into increasing values and accumulating produces the Lorentz curve (dashed). This curve divides the area below the Lorentz curve for perfect spectral uniformity (solid) into areas A and B. The Lamp Gini index equals 2A/n or A/(A+B). Here A+B = 0.5 n, and there are n = 81 times 5 nm bins from 380 nm to 780 nm inclusive.

Method

Domestic LED lamps were purchased for spectral measurements, purposively sampled from different suppliers and distribution channels, aimed at consumers in the UK, with a range of colour temperatures, and including both dimmable and non-dimmable lamps. All 23 lamp models were based on partially phosphor-converted light of blue LED chips. The main report also considers LED-based office panel lighting and street lighting [1].

The RSPD data $p(\lambda_i)$ were calculated, and colorimetric calculations were carried out using a specially built and verified spreadsheet. The published TM-30-15 spreadsheet used later provided a second verification of these results. In a further tool, the Lamp Gini index and Spectral Entropy calculations were configured to allow for changes in wavelength resolution to be studied.

As mentioned earlier, high values of L_G are associated with poorer spectral richness (or with inequality in economics), and poorer rendering. The concepts related to spectral richness can be defined as follows: Spectral richness is a description of a radiometric spectral distribution with a relatively low statistical dispersion between the radiometric quantity (*e.g.* spectral irradiance) in its spectral bins. A source having a high statistical dispersion can be referred to as spectrally selective, whereas low dispersion indicates a spectrally-rich source. Note that the concept of spectral richness is based on the RSPD, and that when a monochromatic source is added to a spectrally rich source it tends to reduce the spectral richness in the combined distribution. The less formal term 'spiky' is commonly used to describe a spectrally-selective source dominated by two or more distinct lines, such as mercury emission lines.

Starting with 81 bins of width 5 nm (378.5 nm to 782.5 nm), increasingly reduced resolution representations can be derived directly with 27 bins of width 15 nm, 9 of width 45 nm and 3 of width 135 nm. Based on 1 nm resolution data, similar wavelength ranges can be derived with a range of integer bin counts (*i.e.* with no partial bins). Mainly to supplement the limited number of coarser resolutions, the following bin counts n were included: 400, 20, 12, 10, 6, 5, 4, 2 bins with 400/n nm resolution.

Results

Figure 4 shows the distribution, for the domestic LED lamps only, of traditional fidelity metrics CRI R_a vs CQS Q_a, the two new fidelity and gamut metrics TM-30-15 R_f vs R_g and the spectral richness metrics Lamp Gini index G_L vs Spectral Entropy H_{Sp}. Under these measures, two LED lamps with poorer colour rendering properties and one with superior scores, stand out. Due to the purposive sampling, care should be taken in the interpretation of the R^2 values, which may not necessarily indicate general associations for domestic LED lighting in general.

Figure 5 illustrates the impact of resolution on these two spectral richness metrics for typical domestic LED, TH and CFL lamps. When based on coarse resolutions, fidelity and spectral richness metrics are inefficient and the spectral richness metrics are unpredictable. All the quality and fidelity metrics in current use are essentially spectral richness metrics incorporating reflectance and colorimetric weightings. In theory, the weightings may add robustness to coarser resolutions, but in practice, where reflectance functions have been selected without independence (*e.g.* CRI), useful spectral richness information is lost together with the unnecessary information.

Metrics are often employed to rank lamps. Figures 6a and 6b show that the resolution threshold for lamp ranking in a spectral richness metric becomes stable for resolutions below around 5 nm to 20 nm. These points can be also be illustrated using the correlations in Figure 6c. The two spectral richness metrics are highly negatively correlated, with similar, but numerically reversed, rankings at the finest resolutions. At 5 nm, the best and worst performing of the LED sample tested can be easily traced back from Figures 6a and 6b to their positions in Figure 4c.

Discussion

As these results also relate to various types of acceptable lighting, they give some hope that, for general use, a typical spectral resolution of target reflectance features can be derived, perhaps in the range 5 nm to 20 nm. Figures 5 and 6 help to show how spectral richness varies over this range of wavelength resolutions, based on the Lamp Gini index and Spectral Entropy. Note that for the coarser resolutions, colour rendering information is lost or diluted, particularly when using a metric with no dependence on bin-order.

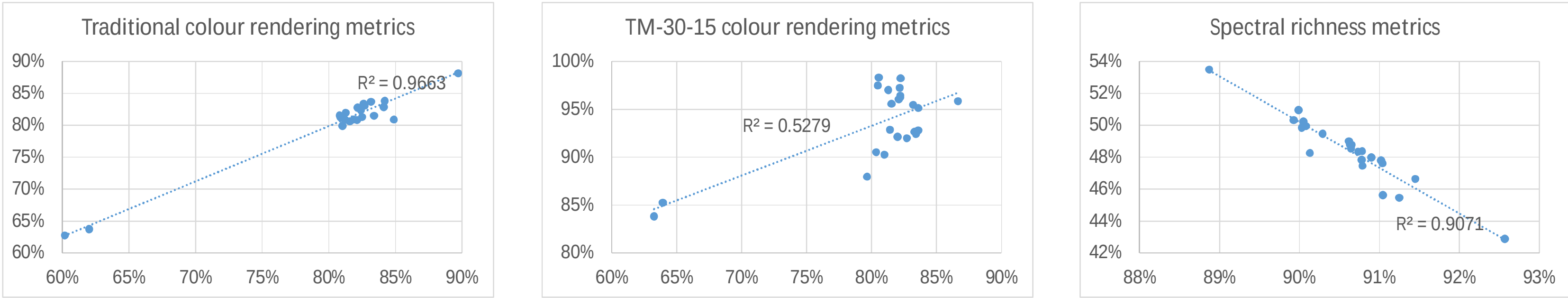

Figure 4: Colour rendering metric performance for domestic LED lighting sample. Vertical axis vs horizontal. From the left a. R_a vs Q_a b. R_f vs R_g c. G_L vs H_{Sp}.

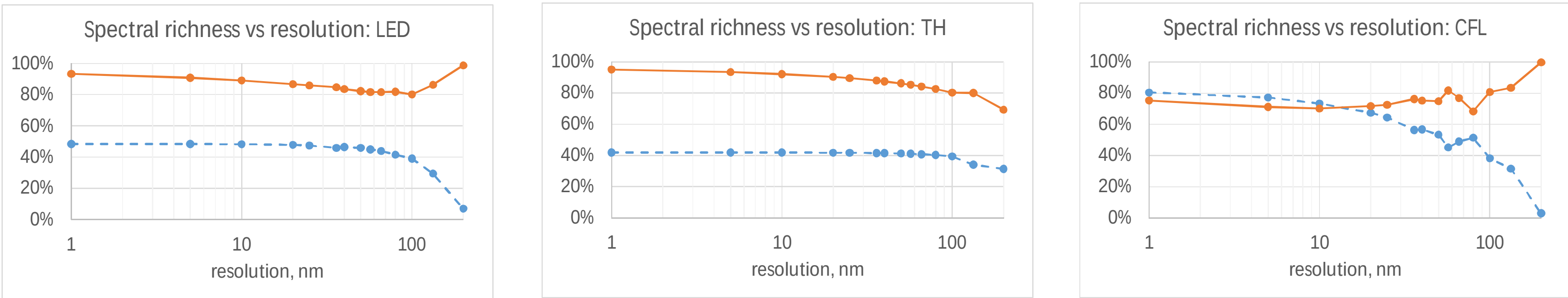

Figure 5: Spectral richness of domestic lighting for metrics calculated at different wavelength resolutions (H_{Sp} solid, G_L dashed). From the left a. LED b. TH c. CFL.

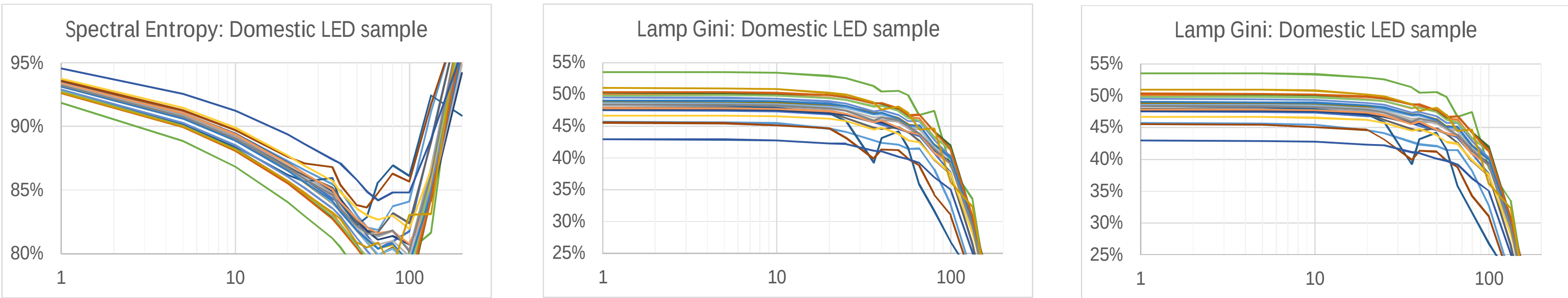

Figure 6: a. H_{Sp} and b. G_L both as Fig 5.a. c. Correlations to 1 nm metrics (H_{Sp} solid, G_L dashed) and negative correlations between H_{Sp} and G_L (dark): domestic LED sample.

Note that the second worst LED tested does not stand out clearly in Figure 4.c based on the values of the spectral richness metrics. In the context of this sample, this lamp spectrum might be investigated to demonstrate whether unweighted spectral richness metrics, at any resolution, are sufficient to predict colour rendering for standard human vision.

Spectral band methods were largely abandoned with the emergence of the CRI, and one criticism of Bouma's method is notable [19]. The argument against spectral band methods relied on the ill-defined distinction between a “reduced” spectral band representation of a light source, and “the complete spectral energy distribution”. This could refer to (at least) two concepts, either the loss of information due to the spectral resolution or, alternatively, the difference between mathematical representations with an arbitrarily low (*i.e.* detailed) spectral resolution and psychophysical tests. As such visual tests are only possible with the actual source, and therefore apply to any metric using spectral input, this distinction is not considered to be a fair criticism. In the former case, the clear implication is that with a sufficiently fine spectral resolution, spectral band methods would remain valid. Whilst other arguments are put forward, they consistently fall into this pattern – either they apply equally to other calculation methods, or it is simply the low spectral resolution in Bouma's and other spectral band methods of the time that were found wanting.

Note that before calculating CRI, CQS, GAI or the TM-30-15 metrics it is necessary to represent the spectrum as an RSPD in 5 nm bins (or at some other resolution), so in practice these various indices are spectral band methods in disguise. This is the final argument that dispels the myth about the inappropriateness of spectral band methods for colour rendering. To replace the ill-defined distinction between complete and incomplete, there are two clearer distinctions between metrics which are convergent and preserve rankings as the spectral resolution decreases and those which do not.

It is beyond the scope of this paper to demonstrate beyond all doubt that Spectral Entropy and the Lamp Gini are convergent rank-preserving metrics. What is clear, from Figures 5 and 6, is that these behave like convergent rank-preserving metrics to well below the spectral resolution threshold for present day colour rendering applications.

Conclusions

The sample of domestic LED lights based on delivering partially phosphor-converted blue LED light provided some reassuringly high colour rendering metric scores, which compared favourably against fluorescent lamps, as previously reported [1]. Blackbody-like incandescent and tungsten-halogen lamps have a natural advantage in spectral richness, smoothness and naturalness, which is also built into reference lamps used in many fidelity metrics.

However, there were two clear cases of inadequate colour quality in the LED sample, which perhaps shows that cheap or old designs can still be problematic. These came from a well-known general purpose online supplier, and were not branded; there were other safety concerns with these two models, and they were not representative of the bulk of the domestic LED lighting market.

To select or derive a colour rendering metric for quality, fidelity or spectral richness without considering wavelength resolution is not ideal, to say the least. It is fair to consider such an approach as underspecifying the purpose or purposes to which the metric is to be put. Regrettably, many studies into colour quality, fidelity and spectral richness will be found lacking in this respect. However, the effects may be less serious or non-existent where the resolution is finer than a required threshold.

TM-30-15 has undoubtedly improved many aspects of the colour rendering toolbox, but it will be interesting to see what effect overlooking the importance of wavelength resolution has had on the fidelity index, and whether it will be the last word in addressing the shortcomings of the CRI system. 99 reflectance samples, with uniform coverage over the wavelength space, may or may not be sufficiently powerful. The sample tested here is not sufficiently diverse in spectral distribution to address this question.

Resolution may not create quite such a great problem for gamut area metrics and an important advantage of the TM-30-15 gamut area index over its predecessors is that it is relatively independent of correlated colour temperature [1].

Acknowledgements

I would like to thank Marina Khazova of Public Health England for supplying data from careful measurements, to CIBSE and SLL for initiating and funding the wider project, and to Wendy Davis of the University of Sydney for assistance with the early history of colour rendering.

References

1. Chartered Institution of Building Services Engineers (2016), CRCE-RDD 01-2016. Human Responses to Lighting Based on LED Lighting Solutions (in association with Public Health England, Didcot, UK and Society of Light and Lighting, Balham, UK), CIBSE, Balham, UK.
2. Wyszecki G and Stiles WS (1982), *Color Science*, 2nd Edition, New York: Wiley.
3. SPIE Vol. MS 77 (1993), Selected papers on *Colorimetry–Fundamentals*, SPIE Milestones Series.
4. Boyce PR (2014), *Human Factors in Lighting*, 3rd Edition, Boca Raton, FL: CRC Press.
5. Davis W (2015), History of color metrics, in *Handbook of Advanced Lighting Technology*, Karlicek R, Sun C-C, Zissis G, Ma R (eds.), Switzerland: Springer.
6. CIE 13.3-1995 (1995), *Method of Measuring and Specifying Colour Rendering of Light Sources*, 3rd Edition, CIE, Vienna.
7. Davis W and Ohno Y (2010), Color quality scale, *Optical Engineering*, **49** (3), 033602-033602.
8. Rea MS and Freyssinier-Nova JP (2008), Color rendering: A tale of two metrics, *Color Research and Application*, **33** (3), 192-202.
9. Illuminating Engineering Society (2015), *TM-30-15 IES Method for Evaluating Light Source Color Rendition*, New York: Illuminating Engineering Society.
10. Ashdown I, Avilés G, Bennett LC, Burkett R, Choi A, Conway K, Deroos M, Druzik J, Gregory P, Herst D, Houser K, Innes M, Israel C, Luedtke W, Oberkircher F, Paolini S, Rosen S, Rosenfeld S, Royer M, Sanders M, Siminovitch M, Smet K, Stone C, Sundin J, Tonello G, Van Der Burgt P, Van Kemenade J, Veitch JA, Wei M, Whitehead L and Wood M (2015), Correspondence: In support of the IES method of evaluating light source colour rendition, *Lighting Research and Technology*, **47** (8), 1029-1034.
11. Rea MS (2013), *Value Metrics for Better Lighting*, Bellingham, WA: SPIE Press.
12. Yaguchi H (2017), CIE 2017 colour fidelity index, *Proceedings of the CIE 2017 Midterm Meeting*, CIE, Vienna.
13. CIE Position Statement (2015), *CIE Position Statement on CRI and Colour Quality Metrics*, CIE, Vienna.
14. Bouma PJ (1937), *Colour reproduction in the use of different sources of 'white' light*, Philips Technical Review, **2**,1-7.
15. Price LLA (2012), *Entropy, color, and color rendering*, *Journal of the Optical Society A*, **29** (12), 2557-2565.
16. Gini C (1912), *Variabilità e mutabilità*, reprinted in *Memorie di metodologica statistica*, Pizetti E and Salvemini T (eds.), Rome: Libreria Eredi Virgilio Veschi 1.
17. Gini C (1921), Measurement of inequality of incomes, *The Economic Journal*, **31** (121), 124-126.
18. Shorrocks AF (1980), The class of additively decomposable inequality measures, *Econometrica: Journal of the Econometric Society*, 613-625.
19. Henderson ST and Waigh DT (1953), *The colour-rendering properties of fluorescent lamps and a proposed new method of specification*, *Transactions of the Illuminating Engineering Society*, **18** (4), 113-121.