To Average, or not to Average ...

Robert Clear and Samuel Berman
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

Introduction

A common lighting design problem is evaluating distributions of power levels, or illuminances, luminances, or some other lighting parameter. It is a problem for IES committees, as well as the practioners. We discuss the advantages and disadvantages of three techniques that are being used.

Background

One approach to evaluating a distribution is to use a generalized measure of the overall location of its values. The location of a distribution is usually given in terms of some measure of its "center". The most common of these measures is the average. In this energy conscious era every designer is probably familiar with computing average watts/ft^2.

Unfortunately evaluations in terms of averages do not always make sense. Laurence Maloney in his article on sports lighting had the following pithy comment on the calculation and evaluation of averages: "From a calculation standpoint, one takes total lumens delivered to an area and divides by the area. This is a little like sitting on a stove with one's feet in a bucket of ice - on the average, you are comfortable.".[1]

Instead of averaging Maloney suggests the use of target percentile values. A sample probability distribution for illumination over the space is generated from measurements at a number of points throughout the area. This is the type of procedure that was used for ESI evaluations.[2] He defines the target percentile value as the value that a fixed percentile of the points meets or exceeds, and suggests using the 70th percentile as the target percentile. However, this does not actually solve the dilemma that Maloney has raised. For instance, if we take Maloney's example and modify it by adding a few more ice and stove surfaces for our subject to touch, we can get examples where the subject is miserable even though the percentile value would insist the opposite.

If we used the 50th percentile as the target value we would actually be taking a median, which is just another of the measures of the center of a distribution. Choosing the 70th percentile mixes information about the "location" of the values with information about the width of the distribution. A third method of evaluating a distribution is to specify both its center and width explicitly.

For lighting design purposes the width of the distribution seems to be normally given in terms of a "uniformity ratio", which is defined as the average/minimum or possibly as the maximum/minimum. One example of this approach can be seen in the Roadway lighting recommendations, which are given in terms of a minimum value for the average, and a maximum value for the uniformity ratio. This type of procedure is capable of giving reasonable predictions of the comfort of Mr. Maloney's poor subject. A limitation of this procedure is that it ranks designs into only two classes: pass and not pass. In fact, a two parameter description, such as is given here, cannot be used to unambiguously rank designs without some rule to convert the two parameters to one parameter or attribute. Without this rule there is no way to rank system "A" against system "B" when "A" is better in one attribute and worse in the other.

Discussion

Ideally we would like a single number value that, as Maloney puts it, "does everything". Let us look Maloney's comment on averages again. What Maloney meant, is that in his test the average temperature is comfortable, but the individual is not. The situation most likely to produce problems with averages is where we are interested in one type of quantity, and are measuring and averaging a different, related quantity, such as comfort and temperature, respectively, in our example. The average comfort level correctly ranks Maloney's example as being uncomfortable, and will usually give a fairly good one-parameter value for evaluating comfort in a space. Unfortunately, in lighting it is generally difficult to measure the quantity of most direct interest.

Let us presume that the functional relationship, C=f(T), between a particular temperature, T, and comfort, C, the probability of temperatures in a room, p(T), and the resultant comfort distribution, pc(C), calculated from f(T) and p(T), are as shown in figure 1. The average comfort level, $<C> = <f(T) * p(T)>$, where <> denotes averaging (in this case an average over temperature) is not equal to the comfort level at the average temperature, $C = f(<T>)$, or the comfort level at the 70th percentile temperature, C=f(T70). In general, for any dependent-independent variable pair, y = g(x), you can only be assured that $<y> = g(<x>)$ when the function g(x) is linear in x (g(x)=a + bx).

This is the root of our problems in using averages to evaluate lighting designs. Distributions of illuminances, luminances, Visual Comfort Probabilities (VCPs), Equivalent Spherical Illuminations (ESIs), and so on, are what we measure or calculate, but what we are actually interested in are the distributions of comfort, visual performance, and productivity. The relationships between what we measure and what we are interested in are not linear, so averaging over the measured values is often not appropriate.

It is instructive to review the uniformity constraint, and percentile procedures, in the above context. As we show below, under the proper conditions, both approximate an averaging over some quantity of interest. Note that almost any function is approximately linear when the range of the independent variable is small enough. Thus if the uniformity ratio for the independent variable, x, is small enough, then the average of the dependent variable is approximately given by $<y>$ *l* $f(<x>)$. In words, if the uniformity constraint is tight enough, the average value is useful.

The acceptability of the percentile procedure presupposes some explict knowledge about the shape of f(x). The 70th percentile can be expressed as the median of the bottom 60% of the distribution. The values of the upper 40% do not count. This is what you want to do if f(x) saturates for the top 40% of the values of x, so that the actual x values are irrelevant. Figure 2 gives a rough idea of how performance is related to light level. It is a very compressive function, with performance improving rapidly at low light levels, and then tending to saturate, or at least level off, at higher light levels.

When an arbitrary function, y = f(x), is monotonic, the percentiles in x are directly related to the percentiles in y, thus y70 = f(x70). For a symmetric distribution the median and the average are same. The very compressive and monotonic nature of the performance-light level relationship will, for most practical light level distributions, x, tend to make the performance distribution, y, very skewed. Normally one would not characterize a distribution with a 70th percentile, instead of a median or average, but for a very skewed distribution a percentile value such as y70 may be closer to the average, <y>, than either the median or f(<x>). Thus, the success of the percentile approach is basically due to the fact, that under the proper conditions, it approximates the average of the quantity we are ultimately interested in.

A problem with the percentile approach is that it is not clear which percentile level should be chosen as the target level. Maloney's comment on this point sums the situation up: "Why 70? It seems reasonable from a practical standpoint.". The choice of target level mixes information about both the functional dependance of y on x, and the form of the distribution of x values. Thus there is no one target percentile level that will give the best estimate of y for all x distributions. Ideally what we want is a measure whose validity is independent of the form of the x distribution.

The lighting community's experience with ESI provides an example of the difficulties that arise from the non-linear nature of the illuminance - performance relationship. ESI was developed to go a step beyond the use of illuminance by including the effect of contrast in the evaluation of visibility. The result was expressed in terms of equivalent footcandles at a fixed contrast level. At normal office light and contrast levels, visibility is far more sensitive to contrast than it is to light level, although it is not particularly sensitive to either. Small variations in contrast give large variations in ESI, with the result that ESI distributions tend to be many times broader than the corresponding illuminance distributions. Since the performance - illuminance function is usually non-linear over the range of the distribution, the average ESI is not a good estimator of performance, and may even be worse than average illuminance. Percentile calculations can help, but as we noted above, it is a question of luck as to whether a particular calculation accurately reflects the true performance potential of the space.

The most important point of our discussion is that one needs to understand the application, or context, of an evaluation. The application determines the quantity of interest, y. The average, <y>, is, in a certain sense, the "best" estimate of the value for any sample taken from the distribution. This is why averaging can be so useful. However it should be clear that taking independent random samples from a distribution is not an appropriate model for all applications. For example, if the application is to build chains to support weights, then the user will be interested in the minimum strength of the links, and by inference, the chains. The average link or chain strength is not interesting for this application. Note that when averaging is not appropriate, the percentile and average plus uniformity constraint procedures are usually not appropriate either.

If there is more than one application for the data, or the applications are not known in advance, or the transform, y = f(x), is not known, the distribution cannot be simply evaluated. The problem faced is instead one of efficiently describing the distribution. The most efficient procedure is to give the parameters of a fit to the distribution so that the user can perform accurate evaluations as needed. In this interim stage of description, the average may be one of a number of parameters that are useful, but by itself, it is rarely a sufficient description.

Conclusion

The quantitative evaluation of a lighting design is a multi-step process. Important steps are the determination of an explicit measure of what we are actually interested in, the dependent variable y, the determination of the relationship between y and x, $y=f(x)$, where x is a measureable quantity such as illuminance, or luminance, and finally, the determination of the distribution of x, $p(x)$. Given $f(x)$ and $p(x)$ we can calculate the distribution of y, $g(y)$. For most lighting applications, given $g(y)$, $\langle y \rangle$ is the single number that is all that is needed to be known about the distribution. Thus, for example, what we might really like to know about highway lighting is the expected average number of vehicular accidents, ($\langle y \rangle$), as a function of the distribution, $p(x)$, of measureable lighting quantities, such as luminance, glare, and contrast.

In practice, there is a lot we still do not know about the relationships between lighting designs and lighting needs, and the functions that have been developed are approximations to the real, imperfectly known relationships.[3,4,5,6,7] Nonetheless, using these approximations is the best we can do, and averaging over the y values allows us to include as much as we know about the situation in question, while avoiding the limitations of the two other approaches that we have discussed. Perhaps the only disadvantage of averaging in the above manner, relative to the other methods discussed, is that it takes more work.

Averaging is often misused. It almost sounds paradoxical that at the same time it is underused.

Acknowledgement

This work was supported by the Assitant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technologies, Building Equipment Division of the U.S. Department of Energy under Contract No. DE-ACO3-76SFOOO98.

References

1) Laurence J. Maloney. 1986. Sports Lighting - Some new ideas on quantity and quality. LD&A, 16 (no. 4): 37-41.
2) Design Practice Committee of the IES. 1977. Recommended practice for the specification of an ESI rating in interior spaces when specific task locations are unknown. JIES, 6 (2): 111-123.
3) Publication Commission Internationale de L'Eclairage, 19/2. 1981. An Analytic Model for Describing the Influence of Lighting Parameters upon Visual Performance. Paris: International Commission on Illumination.
4) M. Rea. 1986. Toward a Model of Visual Performance: Foundations and Data. JIES 15 (no. 2): 41-58
5) R. Clear and S. Berman. 1981. Cost-effective visibility-based design procedures for general office lighting. JIES 10 (no. 4): 228-236.
6) V. P. Gallagher. 1976. A Visibility Metric for Safe Lighting of City Streets. JIES 5 (no. 2): 85-91.
7) R. Clear and S. Berman. 1985. Target Size, Visibility, and Roadway Performance. JIES 15 (no. 1): 167-180.