NRC Publications Archive
Archives des publications du CNRC

Discussion of Mardaljevic's paper: Verification of program accuracy for illuminance modelling: assumptions, methodology and an examination of conflicting findings

Reinhart, C. F.

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.
For the publisher's version, please access the DOI link below./ Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:
https://doi.org/10.1191/1477153504li120oa
Lighting Research & Technology, 36, September 3, pp. 239-240, 2004-09-01

NRC Publications Record / Notice d'Archives des publications de CNRC:
https://nrc-publications.canada.ca/eng/view/object/?id=58177590-c8c2-4b6b-a6b5-9b1bee4132ce
https://publications-cnrc.canada.ca/fra/voir/objet/?id=58177590-c8c2-4b6b-a6b5-9b1bee4132ce

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at
https://nrc-publications.canada.ca/eng/copyright
READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site
https://publications-cnrc.canada.ca/fra/droits
LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at
PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

National Research Council Canada
Conseil national de recherches Canada

National Research Council Canada
Conseil national de recherches Canada

NRC · CNRC

Discussion of Mardaljevic's paper: Verification of program accuracy for illuminance modelling: assumptions, methodology and an examination of conflicting findings

Reinhart, C.F.

NRCC-47043

A version of this document is published in / Une version de ce document se trouve dans :
Lighting Research & Technology, v. 36, no. 3, Sept. 2004, pp. 239-240
Doi: 10.1191/1477153504li120oa

http://irc.nrc-cnrc.gc.ca/ircpubs

construction
Institute for Research in Construction
Institut de recherche en construction
IRC

Lighting Res. Technol. 36,3 (2004) pp. 217–242

Verification of program accuracy for illuminance modelling: assumptions, methodology and an examination of conflicting findings

J Mardaljevic BSc MPhil PhD
Institute of Energy and Sustainable Development (IESD), De Montfort University, Leicester, UK

Received 30 December 2003; revised 10 April 2004; accepted 13 April 2004

This paper examines the role of assumptions commonly made in validation studies for lighting simulation programs and quantifies the sensitivity of results to uncertainties in key model parameters such as sky conditions and surface reflectivity. Actually occurring overcast skies are often taken to approximate the CIE standard overcast sky for the purpose of comparing predictions against measurements in real buildings. The validity of this assumption is tested against measurements of the sky luminance distribution for real skies. Illuminance predictions are particularly sensitive to the assigned reflectance of surfaces when the direct component of illumination is small. A number of confounding factors that can lead to imprecise estimates of surface reflectivity for building facades are identified and a methodology to minimize their effect is proposed. This study reveals that commonly made assumptions with respect to sky conditions and moderate imprecision in model parameters can lead to erroneous assessments of program accuracy. The degree to which existing validation findings can be extrapolated to very different application scenarios is discussed in the context of reported conflicting assessments of program accuracy.

1. Introduction

Consulting engineers routinely use computer simulation to predict the internal environment and energy consumption of buildings at the design stage. The results from simulation studies may form the basis for significant features of the building design, with ensuing cost implications for time and materials. As a consequence, practitioners choose to employ those simulation programs that have undergone validation tests and proven to be fit for the task to which they will be applied. For any given simulation program applied to a particular design, it is highly unlikely that the validation studies for that program were carried out under conditions identical to the building design under evaluation. For example, the key validation studies for dynamic thermal simulation programs were carried out using instrumented test-cell enclosures. If the fundamental processes (e.g., conduction, convection etc.) can be shown to have been accurately modelled by the program, then this instils confidence that buildings more complex than the test-cell can also be simulated by the program with comparable accuracy.

Validation of predictions from a lighting simulation program, is in principle, a straightforward task. Predictions are compared with measurements and an assessment of the program's accuracy is made. In practice however,

Address for correspondence: J Mardaljevic, Institute of Energy and Sustainable Development (IESD), De Montfort University, The Gateway, Leicester LE1 9BH, UK. E-mail: jm@dmu.ac.uk

© The Chartered Institution of Building Services Engineers 2004 10.1191/1477153504li120oa

it is often difficult to specify with certainty the actually occurring conditions in the simulation model because knowledge of reality is usually both incomplete and imprecise. If the degree by which model and reality diverge is not known, then it becomes difficult to attribute the cause of any disagreement between predictions and measurements: was it due to the underlying algorithms or was the validation scenario wrongly described in the simulation? This paper examines the role of assumptions commonly made in the validation of lighting simulation programs and quantifies the sensitivity of results to uncertainties in key model parameters. Actually occurring overcast skies are often taken to match the CIE standard overcast sky for the purpose of comparing predictions against measurements in real buildings. The validity of this assumption is tested using luminance-mapped measurements of real skies. Illuminance predictions are very sensitive to surface reflectivity in situations where light is received after several reflections between surfaces. The sensitivity of predictions to small changes in surface reflectance is investigated for a generic dense urban setting. As noted, the building model and setting used for a validation study will almost always differ from actual application scenarios. It is instructive therefore to attempt to gauge the degree to which the findings from existing validation studies can be extrapolated to other scenarios. This is considered in Section 5.

2. The validation benchmark study

The *Radiance* lighting simulation system[1] has been the subject of a number of validation studies, more perhaps than any other comparable system.[2–8] In particular, for daylight modelling, *Radiance* has been validated using the BRE-IDMP dataset.[9] The BRE-IDMP dataset consists of simultaneous measurements of the sky luminance distribution, the direct normal illuminance and the internal illuminance in a full-size mock office with south-facing glazing (together with other measurements). This unique dataset, permits, for the first time, a near-complete specification of the sky luminance pattern at the time of measurement. Thus it was possible to make a true assessment of the intrinsic accuracy of the *Radiance* predictions for internal illuminance under a wide range real sky/sun conditions, including heavily overcast. That study demonstrated that the accuracy of the *Radiance* predictions was very high: 66% of predictions were within ±10% of the measured values, and 95% were within ±25%. The accuracy of the *Radiance* predictions was, in fact, shown to be comparable with that of the measuring instruments themselves and much higher than that demonstrated for scale models.[10]

For the BRE-IDMP validation, the building model (3D geometry, surface reflectivities, etc.) and the natural luminous environment (sky luminance distribution etc.) were measured to a high degree of precision and were described in the simulation accordingly. In validation studies where these data are known with less precision, or are not available at all, it is possible to arrive at erroneous conclusions for the intrinsic accuracy of lighting simulation programs. The testing of *Radiation* predictions using the BRE-IDMP dataset is arguably the most rigorous validation study of daylight illuminance modelling to date. Hence it is referred to here as the benchmark validation test against which the less exacting validation scenarios are compared.

3. Overcast sky conditions: reality and the CIE standard

A number of validation studies have tested computer-simulated illuminance predictions against measurements under real skies.[11–13] Where the sky luminance distribution has not been measured it has to be derived from basic quantities using a sky model. For example, the CIE standard overcast model is normally used to represent actually occurring overcast conditions. When this is done for the purpose of testing the accuracy of a computer-simulated predictions of illuminance, it is important that the real sky luminance pattern is identical—or at least very similar—to that used for the simulation. Otherwise it becomes impossible to determine the cause of any divergence between measured and modelled illuminance: it could equally result from poor performance of the simulation program or because the modelled sky was different from the real sky at the time of measurement. To illustrate this, the luminance pattern for a real sky (taken from the BRE-IDMP dataset) is shown alongside that for a sky generated using the CIE Overcast Standard formula, Figure 1. The skies were normalized to produce the same diffuse horizontal illuminance. The real sky was measured during stable conditions and the global horizontal illuminance was $\sim$10 klux. The direct normal illuminance for this sky was effectively zero, thus the global and diffuse horizontal illuminances were equal.

There are a number of criteria that have been employed to infer that a real sky is heavily overcast, and therefore one which can be described using the CIE standard overcast pattern. These criteria include the following:

- Visual observation—the sky appears heavily overcast with no discernable circumsolar region and a 'smooth' brightness distribution.
- Low global horizontal illuminance—usually in conjunction with visual observations.
- When global horizontal illuminance is equal to diffuse horizontal illuminance—in other words, confirmation by measurements that the contribution of direct solar horizontal illuminance is negligible.
- Small fluctuations in short time-step sequential measurements of global horizontal illuminance for seemingly overcast skies.

The aim here is to test the validity of a number of objective criteria that could be used to infer that CIE standard overcast sky conditions occurred at the time of measurement. Visual criteria are not examined because of their subjective nature. In any case, the logarithmic response function of the eye to changes in luminance means that visual perception cannot reliably be used to assess

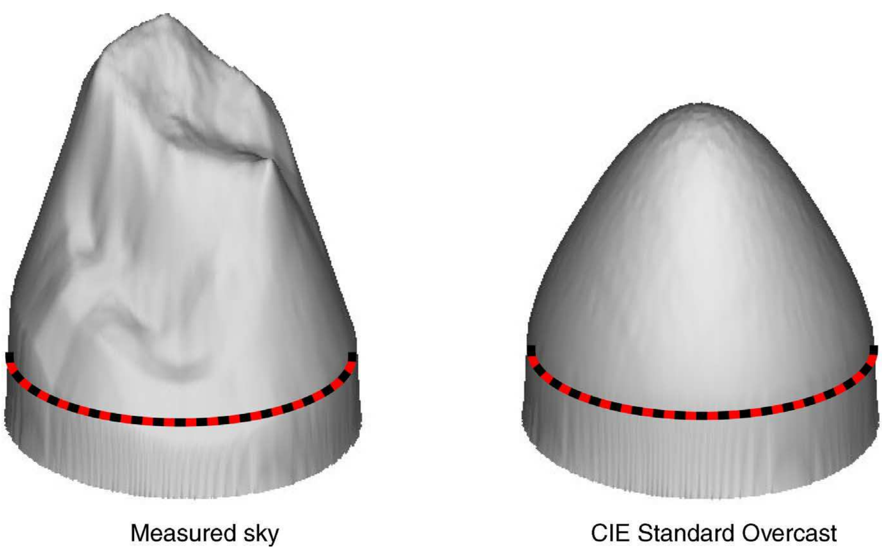

Figure 1 Luminance surface plots for measured sky (266_92_13h00) and CIE Standard Overcast

the regularity or otherwise of sky luminance patterns. The criteria used in studies to infer CIE standard overcast sky conditions are tested here by comparing daylight factors derived directly from measurements in a full-size office space with precise daylight factor values predicted using exact CIE standard overcast sky conditions. The measured data are taken from the BRE-IDMP validation dataset collected by the Building Research Establishment as part of the International Daylight Measurement Programme. The BRE-IDMP dataset consists of simultaneous measurements of a wide range of quantities including:

- the sky luminance distribution measured at 145 points on the sky vault
- global horizontal illuminance and diffuse horizontal illuminance
- the direct normal illuminance
- vertical north, east, south, and west illuminance
- the internal illuminance at six points in a full-size mock office.

As noted, this dataset was used by the author to validate illuminance predictions from the *Radiance* lighting simulation system under real sky conditions.[9] The aforementioned validation proved that the *Radiance* system predictions were capable of high accuracy (typically $\pm 10\%$ of measured values) if the simulation is 'driven' correctly. On the basis of that validation, it is reasonable to assume that the daylight factors predicted for the BRE office are of a similar high accuracy i.e., within $\pm 10\%$ of the true value.

3.1 Filtering criteria

The various physical quantities in the BRE-IDMP dataset permits the formulation of a number of filtering criteria to test the assumption that an actual overcast sky conforms to the CIE standard. In all, nine filtering criteria (labelled cases A to I) were devised and tested. These criteria are divided into three classes—'basic', 'intermediate' and 'refined'— depending on the type of the measured quantities used for the filtering. The 'basic' class uses only measurements of global horizontal illuminance to infer sky conditions. At the next level, the 'intermediate' class uses measurements of both global horizontal illuminance and diffuse horizontal illuminance to infer sky conditions. Finally, for the class designated 'refined', measurements for the four vertical illuminances—north, east, south and west—together with global horizontal illuminance are used to formulate the filtering criteria.

The nine filtering criteria are summarized in Table 1. The first two, A and B, are in the class designated 'basic '. Here the condition in each case that the global horizontal illuminance is within a specified range. For case A, daylight factors were derived for all the skies in the validation dataset where the global horizontal illuminance was greater than 1000 lux

Table 1 Filtering criteria for the different cases

Case	Class	Filtering criteria					Number of skies [754 max]
A	Basic	Gh > 1 klux	and	Gh < 25 klux			480
B		Gh > 1 klux	and	Gh < 5 klux			123
C	Inter.	Gh > 1 klux	and	Gh < 25 klux	and	(Gh-Dh)*100/Dh < 10%	395
D		Gh > 1 klux	and	Gh < 25 klux	and	(Gh-Dh)*100/Dh < 1%	332
E		Gh > 1 klux	and	Gh < 25 klux	and	(Gh-Dh)*100/Dh < 1%	95
F	Ref.	CoV < 25%					334
G		CoV < 10%					188
H		CoV < 5%					81
I		CoV < 5%	and	Gh > 1 klux	and	Gh < 5 klux	34

and less than 25 000 lux. For case B, the range was narrowed to only those skies with global horizontal illuminances greater than 1000 lux and less than 5000 lux. The second class of filtering criteria, designated 'intermediate' make use of an additional quantity: diffuse horizontal illuminance. The diffuse horizontal illuminance together with global horizontal illuminance can identify those instances where the illuminance contribution from the sun is small, and therefore indicative—or rather, suggestive—of overcast conditions. Note that the diffuse horizontal illuminance used here was in fact derived from subtracting the horizontal component of measured direct normal illuminance from global horizontal illuminance. This is considered a more reliable value for diffuse horizontal illuminance than that measured directly using an illuminance meter with shadow band. The percentage relative difference between global horizontal illuminance and diffuse horizontal illuminance is used as one of the filtering criteria for these 'intermediate' cases (labelled C, D, and F). Lastly, for the class designated 'refined' (cases F, G, H and I), the filtering criteria are based on the coefficient on variation between the four measurements of vertical N, E, S and W illuminance. The coefficient of variation, here expressed as a percentage, is a measure of the overall difference between the four vertical illuminances.

The BRE-IDMP dataset consists of 754 simultaneous measurements of internal and external parameters taken from 27 days of monitoring at the BRE during 1992. The 27 days were pseudo randomly selected by the BRE to cover the full range of naturally occurring sky conditions from heavily overcast, through intermediate to clear. Because the filtering criteria were designed to pick out only those skies likely to be overcast, each one results in the selection of a subset of the total number of skies (maximum 754) depending on how many are included by the conditions. The number of skies selected by the filtering criteria ranged from 480 (case A) down to just 34 (case I)—see last column in Table 1.

The BRE-IDMP dataset includes measurements of the sky luminance distribution. This data could be used to make an absolute comparison of the measured sky luminance patterns against the CIE standard overcast pattern.[15] It is almost invariably the case that real sky luminance patterns will diverage from model ideals to a greater or lesser degree. However, quantification of the divergence between the two patterns—real and model—does not reveal how the skies will perform for the purpose of predicting internal illuminance, or as is the case here: daylight factors. Moreover, the aim here is to test the credibility of assumptions that are used in validation exercises, specifically, those pertaining to sky conditions. Thus, the measured luminance patterns were not examined for an *a priori* conformance to the CIE overcast standard, but rather on the basis of the applied filtering criteria.

3.2 Daylight factors based on measurements

Daylight factors for the BRE office were calculated from measurements of internal illuminance and global horizontal illuminance for all of the 754 skies in the BRE-IDMP dataset. The nine filtering criteria (cases A to I) were applied to the predictions and, for each case, the mean measured daylight factor (MMDF) at the six photocell locations was computed. Scatter plots showing the measured daylight factor (x symbol) for the skies selected by each of the filtering criteria are given in Figure 2. Note that a logarithmic scale is used on the abscissa and that many of the points plotted overlap with each other. The mean measured daylight factor at each photocell is marked with a ◆ symbol and the 'standard' daylight factor predicted under exact CIE overcast skies is marked with a ▮ symbol. Evident in the plots is the large range in measured daylight factors at each photocell for cases A to F. The 'standard' daylight factor and the MMDF for each of the cases are given in Table 2. The percentage relative difference

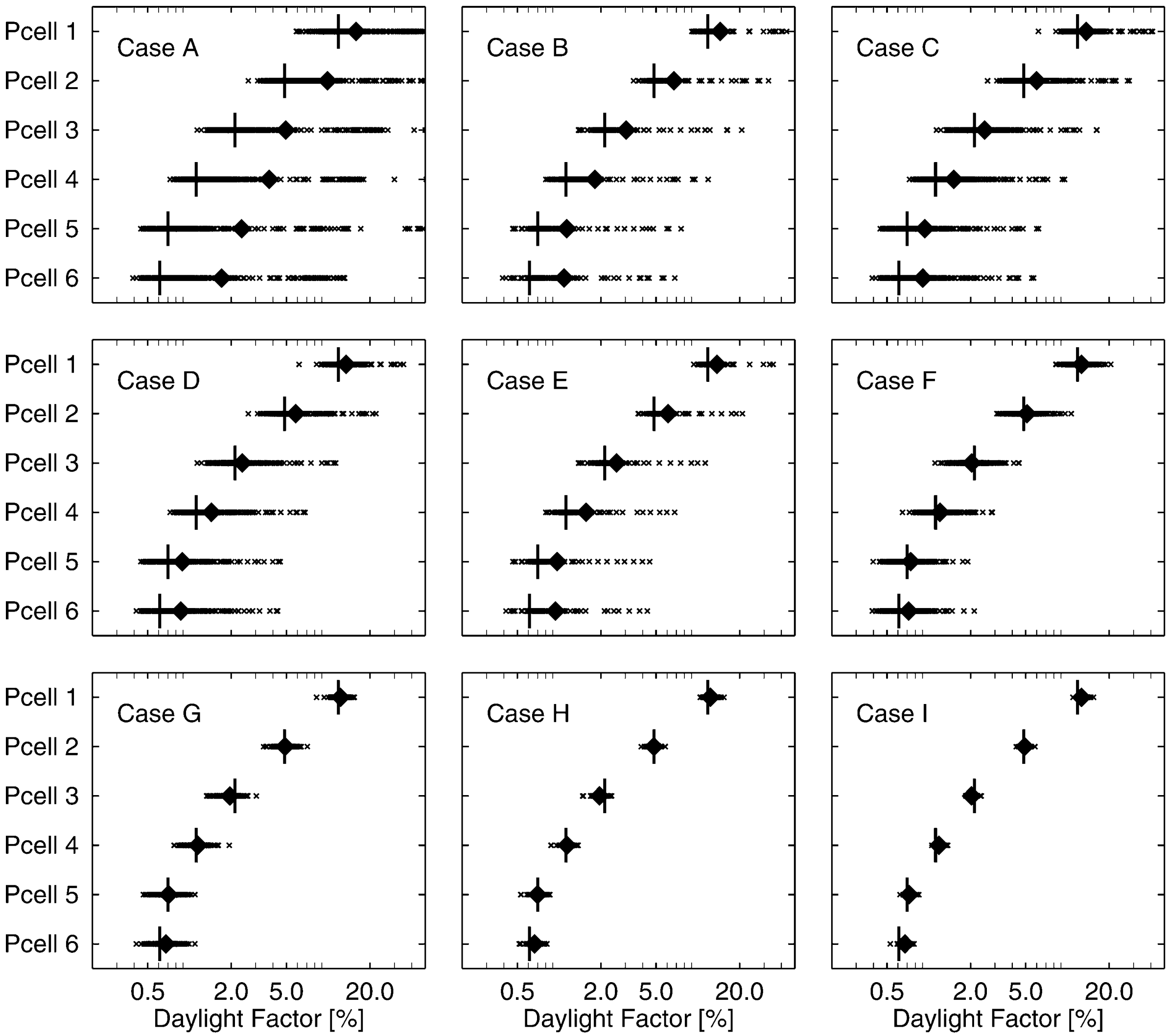

Figure 2 Plots showing the scatter in measured DF, the mean measured DF (◆) and the standard DF (▮)

between the mean measured daylight factor and the true daylight factor value is also presented. Taking first class A and B ('basic' filtering), the mean measured daylight factor is consistently greater than the true value at all six photocell locations. Case B results offer an improvement over case A due to the smaller range in global horizontal illuminances that are considered. But even here the MMDF is greater than the true value by $\sim$ 22% nearest the window and $\sim$ 77% at the back of the office. The 'intermediate' level filtering for the next three cases (C, D and E) improves matters, but the relative differences between the MMDF and the true value at the back of the office is never better than (+)41%. Last are the 'refined' level filtering cases (F, G, H and I) which are based on the coefficient of variation (CoV) between the measurements of the four vertical illuminances. In all four cases, the MMDF shows much better agreement with the true value

Table 2 Mean measured daylight factors for the different cases

Location	Standard DF	Mean Measured Daylight Factor [Relative differences from standard value]								
Pcell 1	11.81	15.86 [34]	14.49 [22]	13.67 [15]	13.44 [13]	13.73 [16]	12.61 [6]	12.27 [3]	12.32 [4]	12.62 [6]
Pcell 2	4.84	9.88 [104]	6.74 [39]	6.00 [23]	5.82 [20]	6.11 [26]	5.11 [5]	4.88 [0]	4.83 [0]	4.91 [1]
Pcell 3	2.13	4.95 [132]	3.04 [42]	2.53 [18]	2.40 [12]	2.59 [21]	2.04 [−4]	1.95 [−8]	1.95 [−8]	2.03 [−4]
Pcell 4	1.12	3.76 [235]	1.82 [62]	1.51 [34]	1.44 [28]	1.57 [39]	1.20 [7]	1.15 [2]	1.14 [1]	1.18 [5]
Pcell 5	0.70	2.37 [239]	1.13 [61]	0.94 [33]	0.89 [26]	0.97 [38]	0.74 [5]	0.71 [0]	0.70 [0]	0.72 [3]
Pcell 6	0.61	1.70 [179]	1.08 [77]	0.91 [48]	0.86 [41]	0.94 [54]	0.72 [17]	0.68 [11]	0.66 [8]	0.67 [10]
Case		A	B	C	D	E	F	G	H	I
Class		Basic		Intermediate			Refined			
Summary of filtering criteria		Limit range in global horizontal illuminance		Limit range in global horizontal illuminance and percentage difference between global and diffuse horizontal illuminance			Limit range in coefficient of variation between vertical N, E, S and W illuminances (and limit range in global horizontal illuminance for Case I)			

at all six photocell locations. Between these four 'refined' level filtering cases, there is little in the way of significant differences. In fact, there is only marginal improvement in the overall agreement between the MMDF and the true value from case F (CoV $< 25\%$) to case G (CoV $< 10\%$), and then only for the photocell at the back of the office. On the basis of correspondence between the MMDF and the true value, differences between cases G, H and I are marginal. The scatter plots however reveal that there is a narrower distribution in measured daylight factors for case H than for case G, Figure 2.

3.3 Real and assumed overcast sky luminance patterns

A comprehensive evaluation of real and model skies is beyond the scope of the work reported here. For the purpose of this study however, a visual examination of measured sky luminance patterns against the CIE standard overcast pattern will be sufficient to reveal the effectiveness or otherwise of the various filtering criteria. The luminance of the measured skies along a semi-circular arc (south-horizon to zenith to north-horizon) was determined for each of the 754 skies in the validation dataset. The luminance of a CIE standard overcast sky along the same arc was also generated. For the purpose of comparison, the measured skies were normalized to give the same horizontal illuminance as the CIE standard overcast sky. In this case, 10 000 lux, though the value is arbitrary. The various filtering criteria were applied, and all the luminance-arcs for the skies in each case were plotted together with the luminance arc for the CIE standard overcast. For brevity, only two plots are given in this paper, Figures 3 and 4. The plots show that only 'refined' level filtering (case H, Figure 4) was effective is selecting skies that largely matched the CIE standard overcast pattern. The degree of variance for the luminance-arcs

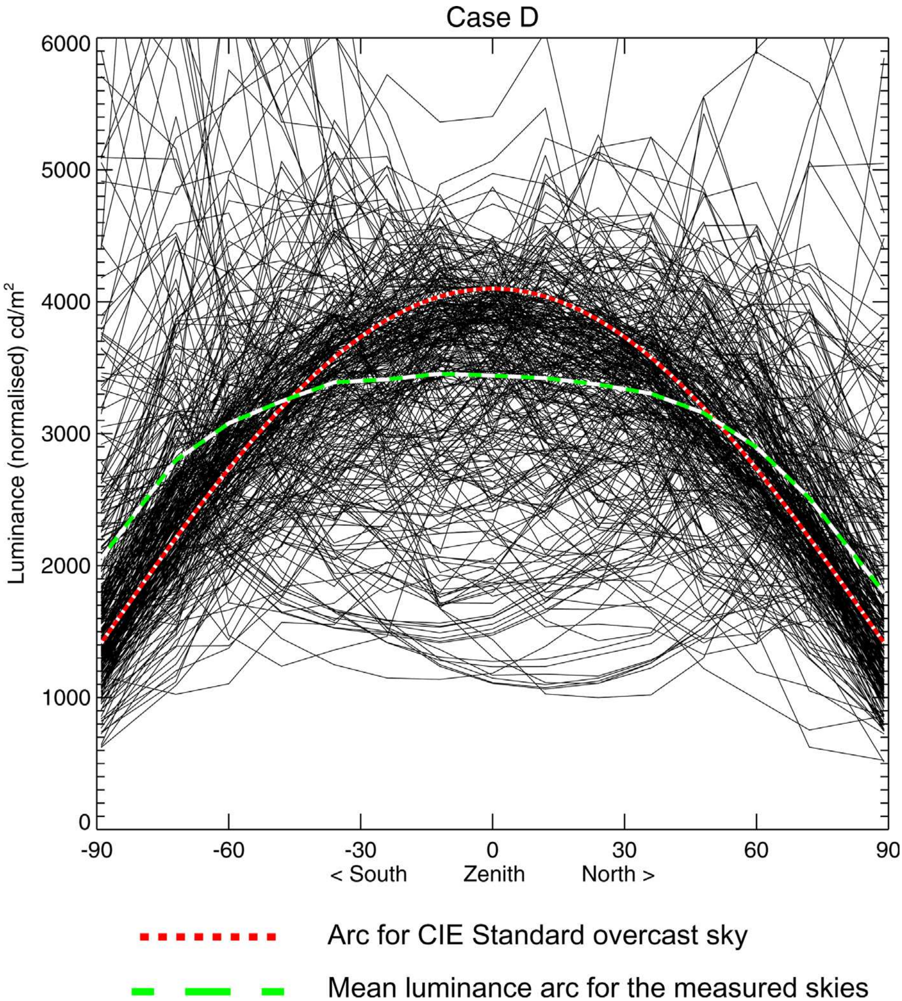

Figure 3 Normalized sky luminance arcs (case D)

selected using 'intermediate' filtering (case D, Figure 3) is perhaps surprising. Recall that for case D, global horizontal illuminance was never more than 1% greater than diffuse horizontal illuminance. Nevertheless, there were many skies where the luminance at low altitude was markedly greater than the zenith luminance, and this was so for both the south and north directions.

An arc showing the mean of all the luminance-arcs for the measured skies is also shown. For case H, the mean-arc is very close to that for the CIE standard overcast (Figure 4). Whereas, for case D the mean-arc shows marked divergence from the arc for the CIE sky (Figure 3). In particular, the mean-arc has a lower luminance at the zenith and a higher luminance towards the horizon. This explains the systematic trend whereby measured daylight factors for the BRE office tended to be in excess of the correct value because the luminance of the sky 'seen' through the glazing was generally higher than that of the CIE standard overcast. This was so for all skies selected using 'basic' and 'intermediate' level filtering. For buildings

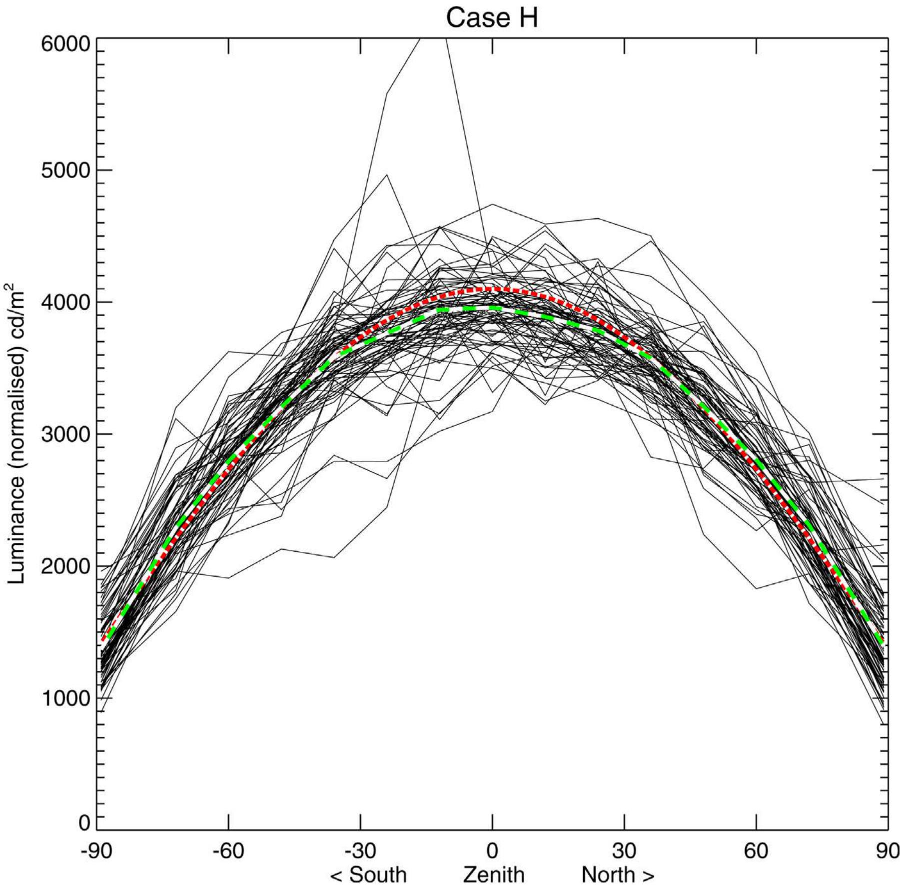

Figure 4 Normalized sky luminance arcs (case H)

with roof-rights, or heavily obstructed sites where the illumination is predominantly from the zenith, the situation is reversed and the tendency would be for measured daylight factors to be less than the correct value.

3.4 Can sky stability be used to infer standard overcast conditions?

Stable sky conditions have been used as one of the criteria to infer that real overcast skies conform the CIE Standard Overcast model.[12,13] For one of these studies, a time-series of global horizontal illuminance was examined for 'fluctuating' sky conditions. A difference of less than 10% in sequential measurements of global horizontal illuminance (i.e., 'stable' skies) was taken to suggest the presence of CIE standard overcast conditions. This assumption is tested here using the coefficient of variation (CoV) in the four vertical illuminances. In the previous section it was established that, for the criteria examined, only the CoV could be used as a reliable discriminator for CIE standard overcast conditions. Consequently, low fluctuations in global horizontal illuminance should be associated with low CoV if 'stable' skies are indeed an indicator of CIE standard overcast conditions. This is tested as follows. The percentage difference in sequential measurements of global horizontal illuminance, called here the stability index, was determined together with the CoV for all 4515 skies in the (5 minute time-step) BRE-IDMP database. Note that the database with the sky scanner measurements was at 15-minute time-step and for fewer days than the 5-minute time-step database. Of course, many of these measurements were for plainly non-overcast skies. To make fair comparison with the method used in Ng,[12] filtering criteria D (Table 1) was applied to select candidate overcast skies for the test. This resulted in the selection of 1854 candidate overcast skies. A scatter plot of stability index versus CoV is given in Figure 5 with dashed lined marking the 10% values for both axes. Evident is the lack of any relation between the stability index and the CoV. The Pearson correlation coefficient for the sample was 0.17. The breakdown in the distribution of the number of skies with respect to the 10% lines (dashed) for stability index (SI) and CoV is given in Table 3. Approximately half of the skies selected by filtering criteria D have a CoV $< 10\%$; these skies can be considered to be close matches to the CIE standard overcast sky (Table 1). A similar fraction had an SI $< 10\%$. However, the condition that SI $< 10\%$ selected only half of those skies which can be reliably assumed to be CIE standard overcast (i.e., CoV $< 10\%$). In other words, a low stability index cannot be used as an indicator of CIE standard overcast conditions, and indeed it would appear to

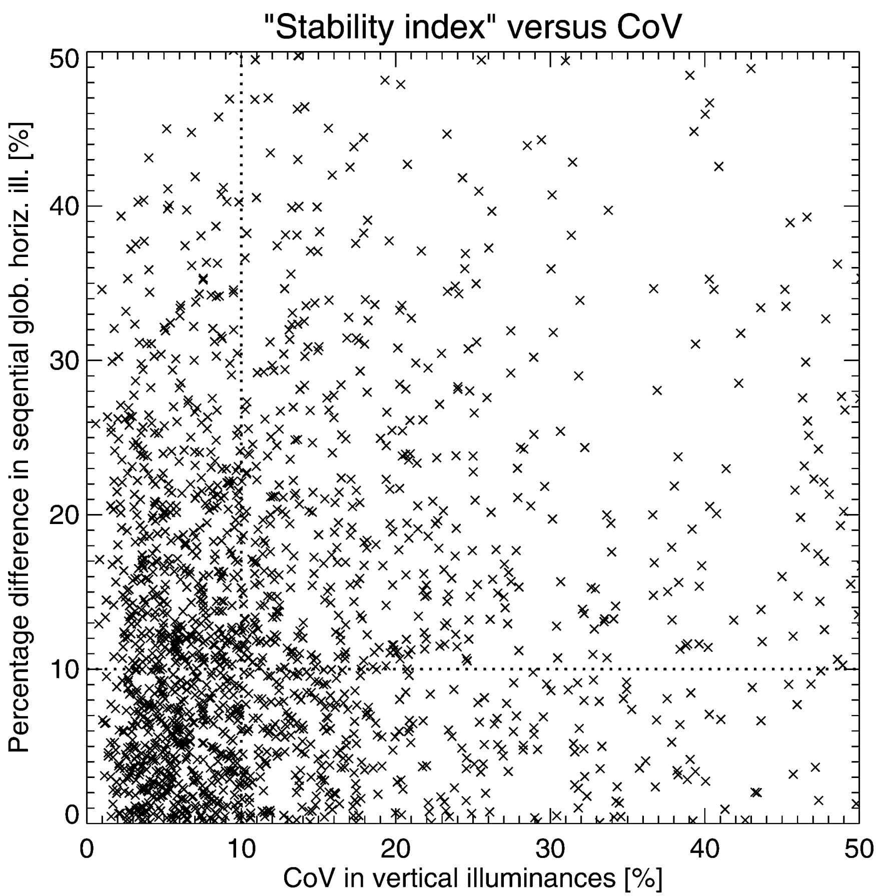

Figure 5 Stability index versus CoV in vertical illuminances

Table 3 Breakdown in numbers for stability index and CoV with respect to the 10% lines

Case (Filtering Criteria D)	Number (Total = 1854)	Fraction of total
StLn < 10%	747	0.40
StLn > 10%	1107	0.60
CoV < 10%	897	0.48
CoV > 10%	957	0.52
StLn < 10% and CoV < 10%	423	0.23

be no better than random sampling. Note that this is for skies that have already been filtered as likely to be overcast (criteria D). Real overcast skies that closely match the CIE standard pattern do not show any tendency to fluctuate less than (real) overcast skies that diverge markedly from the CIE standard pattern. This finding was unexpected as a correspondence, it seems, had long been assumed.

3.5 Summary

On the basis of the above findings, the following conclusions are drawn:

- Filtering criteria founded only on the magnitude of global horizontal illuminance (e.g., 'basic', cases A and B), are unlikely to be a robust indicator that actual sky luminance distributions conform to the CIE standard overcast pattern.
- Filtering which is founded on reliable values for diffuse horizontal illuminance in addition to global horizontal illuminance (e.g., 'intermediate', cases C, D and E) offers an improvement over 'basic' filtering. However, divergence between MMDFs and the true values for the rear half of the office is still significant (+26 to +54%) indicating that many of the actual sky luminance patterns were markedly different to the CIE standard.
- Filtering based on limiting the range in the coefficient of variation between the four vertical illuminances to CoV < 10% resulted in MMDFs that were very close to the true values, i.e., within ±10%. This suggested that CoV < 10% is sufficient to ensure that actual sky conditions were a very close match to the CIE overcast standard. Examination of the luminance-arcs confirmed that this was indeed the case.
- A low stability index (i.e., small fluctuations in global horizontal illuminance) is no indicator that an actual sky conforms to the CIE standard overcast pattern.

If the methodology described was reversed, and the mean measured daylight factors were used to validate the computer predictions of illuminance, it is probable that erroneous conclusions would be drawn depending on the criteria used to filter the measured skies. For example, if the measured DF at the back of the office was determined to be, say 0.91% (case C, Table 2), an accurate computer prediction of 0.61% would be considered to have fallen short of 'true' value by approximately one third. Indeed, it can be seen from Table 2 that taking measured DFs as 'true' would lead to the erroneous conclusion that (accurate) simulated DFs were systematically less than the 'true' value. As noted earlier, when the illumination is predominantly from the zenith (e.g., from rooflights or in a heavily obstructed urban setting), the situation will be reversed and accurate-simulated DFs will be erroneously concluded to be in excess of the 'true' measured value.

4. Sensitivity of results to building model parameters

4.1 Measuring surface reflectivity 'in the field'

Diffuse reflectivities are usually determined from measurements of the surface luminance of the unknown material against that for a known standard. The reflectivity ρ_t of the test material is:

$$\rho_t = \frac{L_t \rho_r}{L_r} \tag{1}$$

where ρ_r is the reflectivity of the reference standard; L_t and L_r are the measured luminance for the test and reference materials respectively. For small planar, homogeneous samples (e.g., painted surfaces) measured under carefully controlled conditions, a precision of ± 0.02 in the determination of reflectivity may be achievable (private communication: P Littlefair, BRE). The measurement of surface reflectivity 'in the field', say, for building facades, presents a number of practical difficulties that are likely to result in significantly lower precision than that stated earlier. For example, the test material and the reference standard must be subject to exactly the same illumination conditions during measurement. To determine the reflectivity of building facades, the measurements will invariably be carried out under daylight conditions. Identical conditions will not be maintained if either the daylight illumination changes between measurements or if the two samples are not subject to exactly the same daylight exposure. When this occurs, it will introduce an error into the calculation of reflectance in the order of E_t/E_r where E_t and E_r are the daylight illuminances incident on the test and reference materials, respectively. In principle, these illuminances could be measured and corrected for:

$$\rho_t = \left(\frac{E_r}{E_t}\right)\left(\frac{L_t \rho_r}{L_r}\right) \tag{2}$$

This method (with or without correction) assumes that the materials are ideal Lambertian (i.e., diffusing) so that the surface luminance is proportional to the incident illuminance, i.e., $L \propto E$.

The measurement of surfaces that are not immediately to hand presents a number of practical difficulties. The usual method for determining the reflectance of, say, an obstructing building facade is to aim a spot-luminance meter at one or more parts (i.e., targets) of the facade and take measurements.[12] At least one of the measurements must be of the reference standard. Ideally, the reference standard and the target samples should be co-located to ensure that they are both subject to the same daylight exposure, i.e., that $E_t = E_r$. A further requirement of course is that the sky conditions remain steady during measurement. Co-location of measurement samples is not always possible e.g., for 'sealed' facades where it is not practicable to place (or drape) a reference standard. Where co-location is not possible, incident illuminances should be measured and any difference corrected for using Equation (2). However, if co-location of test materials is not practicable, then the placement of illuminance meters will most likely not be practicable either.

4.2 Effects of surface 'texture' on reflectivity

For any non-planar building facade, the degree of articulation or surface texture can influence the calculation of surface reflectivity from measurements. The effect of surface texture can operate at many scales, large and small. For a simple overhang with a horizontal projection of 1 m, the shading effect is apparent up to 3 m below it, Figure 6a. Note that this effect is greater for CIE overcast sky conditions than diffuse uniform illumination because the zenith is the brightest part of the CIE sky, and so shadow cast by the overhang is more pronounced than would be the case for uniform sky conditions. For a standard measured at location X, the absolute reflectivity of a target sample at Y will be underestimated, Figure 6a. If the places were reversed, then the reflectivity of a target sample will be overestimated. Where the acceptance angle of spot-luminance meter includes features of the building facade that are articulated, the effect of surface textures is to lower

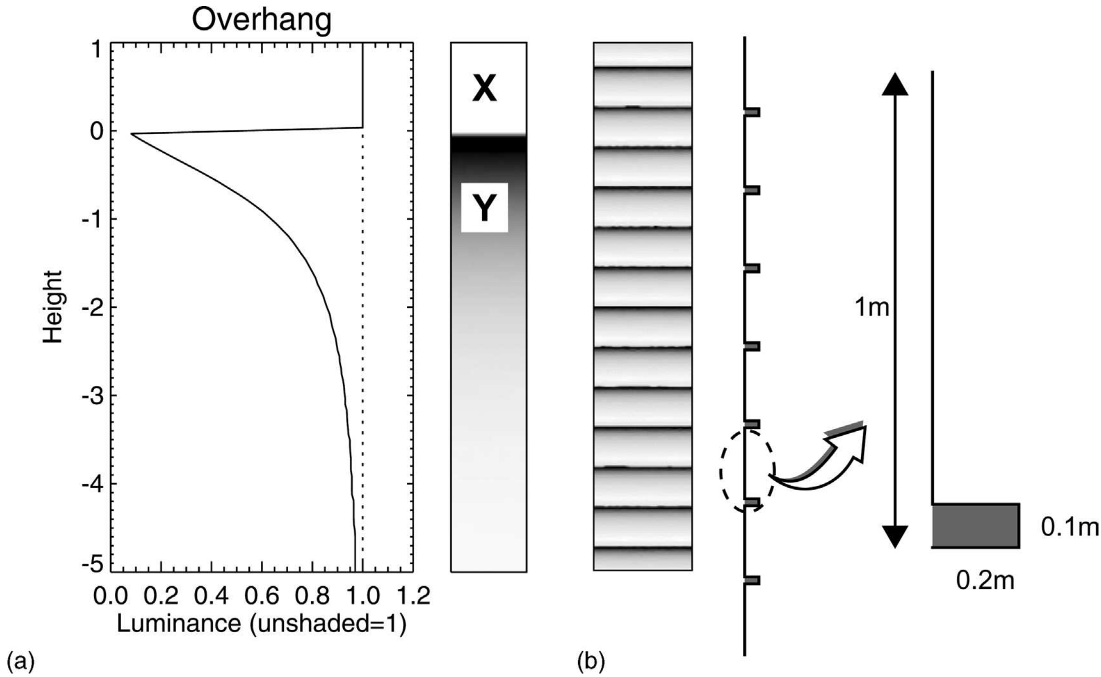

Figure 6 Facade luminance reduced under overhang (a) and effect of surface texture (b)

the apparent reflectivity of the material. In other words, a homogeneous textured surface will always have an overall reflectivity which is less than the absolute reflectivity of the planar surface finish. The likely magnitude of the effect is illustrated in Figure 6b for a facade with just a small degree of articulation. The equivalent reflectivity (under CIE overcast sky illumination) is 0.16 compared to 0.20 for a planar surface.

For the purpose of using measured reflectivities of building in a lighting simulation, the following measurement strategies are suggested depending on the nature of the building facade, Figure 7. For a planar facade (Method A), it is a relatively straightforward matter to determine an area-weighted mean reflectivity from measurements of the component materials (Figure 7a). If the facade has any marked articulation e.g., overhangs, recesses, etc., then these could be modeled explicitly in the simulation to account for the effective reduction in overall reflectivity due to the surface texture, Method B (Figure 7b). Thus, for a *Radiance* simulation, rays would be traced in and out of the recesses etc. A high degree of explicit detail in the 3D model of a building facade introduces overheads both for the geometrical modeling and the simulation of light transfer. These could be avoided, and the surface texture effect accounted for, by the using the luminance meter to determine a bulk reflectivity for a representative section of the facade, Method C (Figure 7c).

For the first two of the methods described above (A and B), controlled illumination conditions must be observed (see Section 4.1). For Method C, the daylight illumination should not change appreciably over the scale of the section of facade that is within the acceptance angle of the spot luminance meter. Also, the standard material needs to be positioned to receive the same illumination as the facade section, but it should not be shaded by any part of the facade. Ideally, it should be draped just proud of the articulated surface. And of

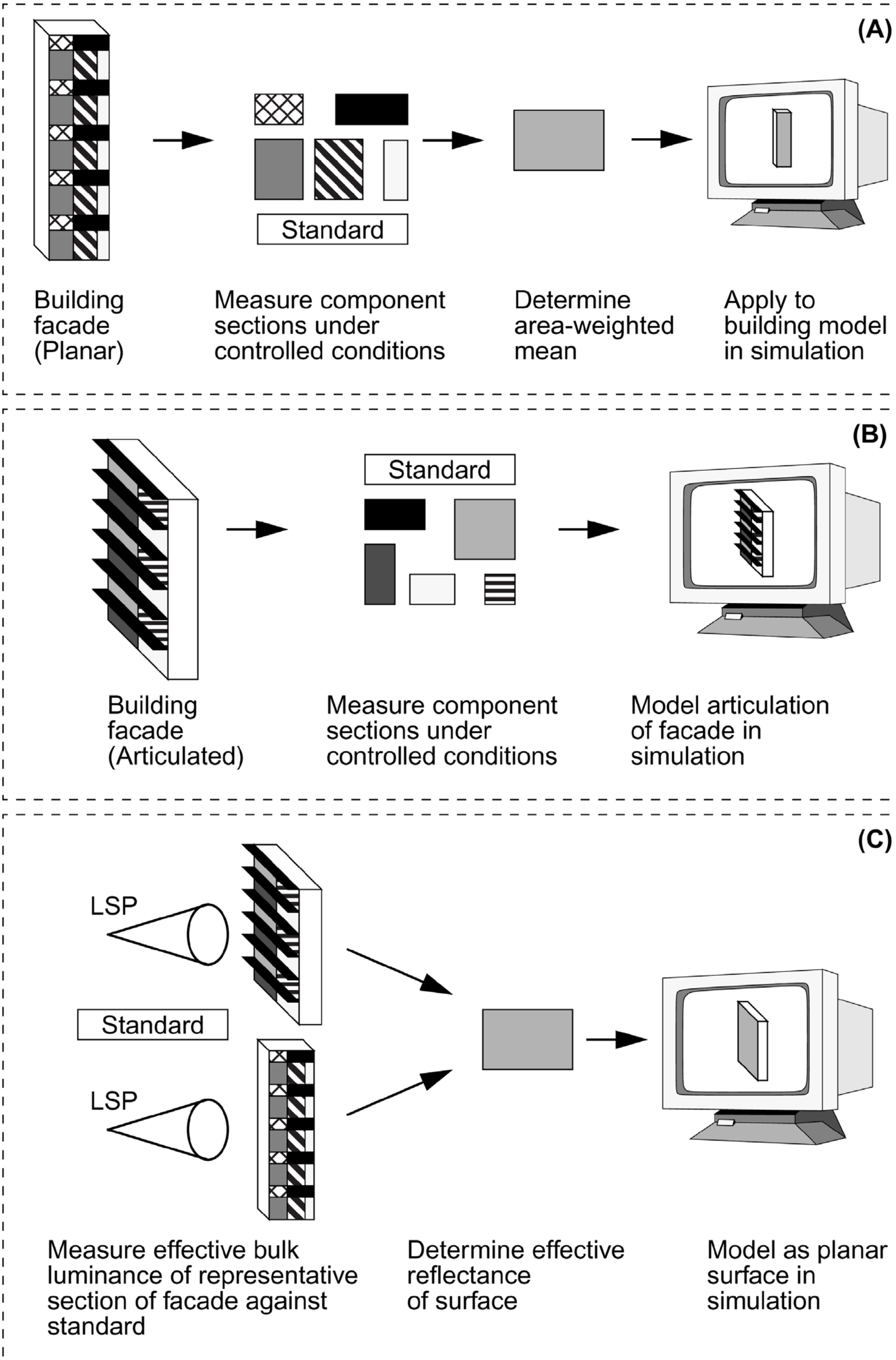

Figure 7 Measurement strategies to determine facade reflectances for simulation

course, illumination conditions should not change appreciably between measurements. Note that the (planar) reflectivity for the articulated facade determined using Method C is dependent on the illumination conditions at the time of measurement. Thus, to be used correctly for modeling purposes, the sky conditions specified for the simulation must be very similar to the (real) sky conditions that occurred at the time of measurement.

The measurement strategies described are, strictly speaking, applicable to facades comprised of diffuse-reflecting materials only. Most real building facades will, of course, contain glazing for which the reflection properties are strongly view dependent i.e., specular. The degree to which the specular properties of glazing may effect modeling studies can be gauged from the curve given in Figure 8. This shows a simulation of the luminance of a vertical glass facade as a function of the angle of incidence with respect to the glazing normal. Under the glass facade was an opaque material of zero reflectance and the scene was illuminated by a CIE standard overcast sky. The increase in luminance with angle of incidence results from two effects combined. For larger angles, the specular reflection is both stronger and directed towards higher altitude sky where the luminance is greatest (for the CIE standard overcast sky). Viewed nearly normally, the equivalent diffuse reflectivity (EDR) for the glazing is approximately 0.08 (Figure 8). This results from weak specular reflections 'seeing' the low altitude sky behind the 'observer'. In a city setting, the EDR would be lower still because the 'observer' is likely to 'see' a lower luminance building in the reflection rather than sky. At 50° incidence, the EDR is $\sim$0.2 and at 65° it rises to $\sim$0.5. The significance of this effect for modeling studies will depend on the scenario. Evidently, the larger the proportion of glazing on the facade, the greater the magnitude of the effect. For obstructing buildings with highly glazed facades, the specular properties of the glass may need to be modeled in the simulation to adequately represent reality in the same simulation, especially for validation studies where high accuracy results are needed.

4.3 Sensitivity of predictions to variations in surface reflectivity

The dimensions and surface reflectivities of the building model for the BRE-IDMP validation were measured to a high degree of precision. For example, the overall office dimensions were measured to an accuracy of $\sim$1 cm and the glazing dimensions to an accuracy of $\sim$0.2 cm. Surface reflectances were measured by the originators of the dataset under controlled conditions.[16] There were no external obstructions of any significance, and so none were included in the 3D model. Where external obstructions are present, the daylight factors will be sensitive to the magnitude of the reflectivities assigned to the obstructing facades. As discussed in the previous section, the measurement and characterization of building facades for simulation studies is subject to a number of confounding factors. As a consequence, it is highly unlikely that actual building facades could be measured and modeled in a simulation with comparable precision to that attained for the benchmark BRE-IDMP validation.

The consequences for simulation studies of imprecision in the measured diffuse reflectivity is investigated in this section. An idealized obstructed (i.e., urban) setting was devised to test the sensitivity of illuminance predictions to variations in diffuse surface reflectivity. For this, a 'tower block' building was positioned centrally in an open top 'box' of equal height to represent a heavily obstructed urban environment, Figure 9. The absolute dimensions are unimportant; the purpose here is to reveal the magnitude of the likely effects. Vertical daylight factors were predicted at six points evenly spaced along a line on the central tower with the reflectivity of all the surfaces set to 0.10. This was repeated for

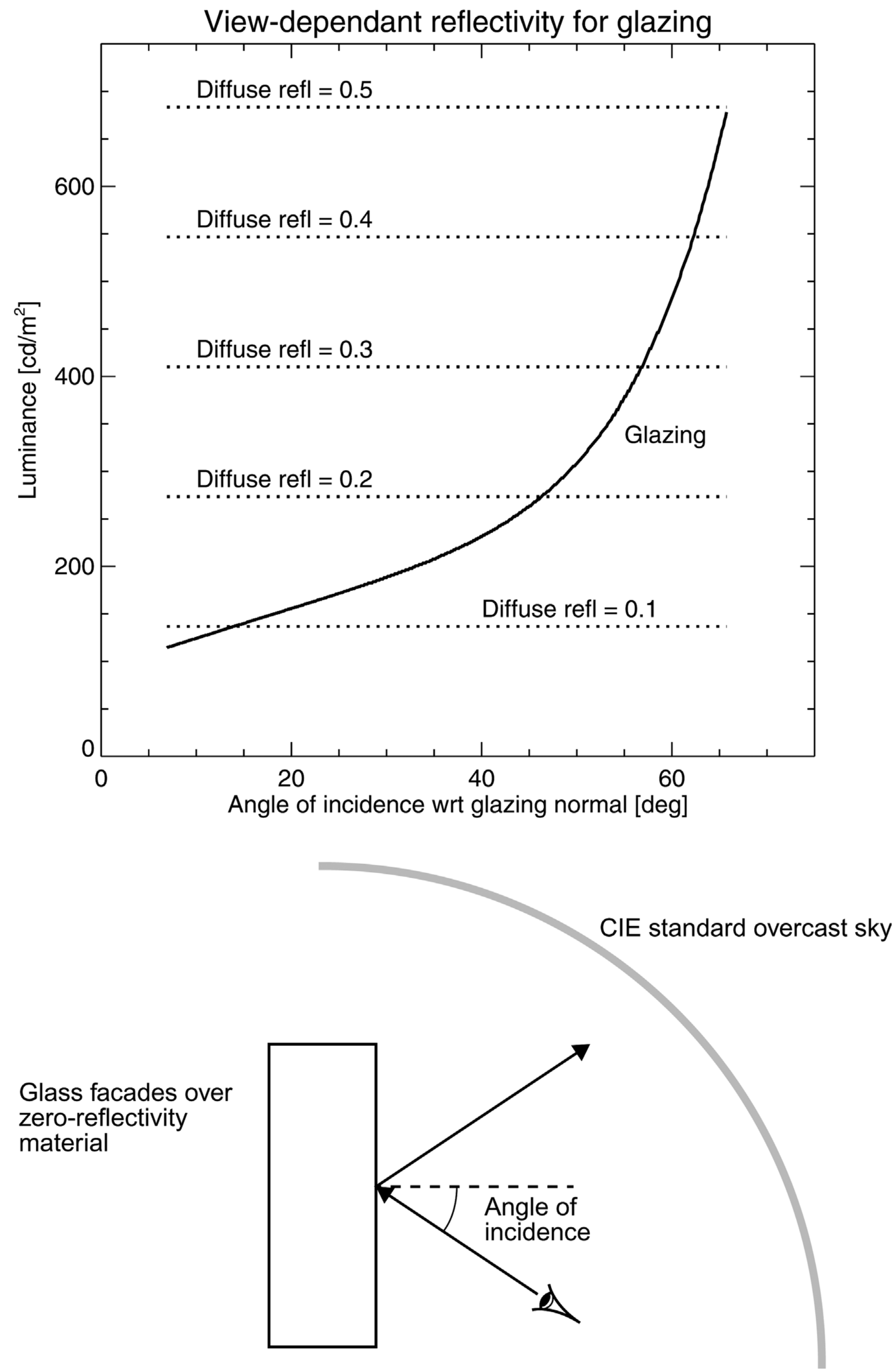

Figure 8 View-dependant reflectivity

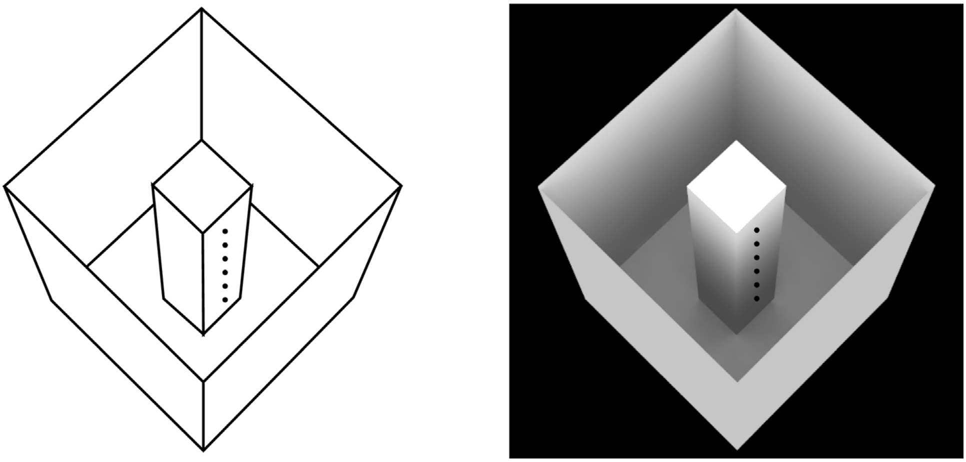

Figure 9 Simplified obstructed site model

reflectivities up to 0.60 in steps of 0.10. The vertical daylight factors predicted for each of the reflectivities are given in Table 4. Daylight factors at the bottom of the tower show the greatest sensitivity to surface reflectivity values. For example, at location 6 the daylight factors for surfaces reflectivities of 0.10 and 0.20 were 1.0% and 1.8% respectively, giving a relative difference of approximately 100% depending on which value is taken as the datum, see Table 4. As noted in Section 4.1, an uncertainty in surface reflectivity of 0.05 may in fact be small compared to the actual uncertainty that can result when real measurements are taken under outdoor 'field' conditions. This simple exercise demonstrates that very large errors in illuminance predictions for obstructed settings can result where there is relatively small imprecision in the value of the surface reflectivity.

5. An examination of conflicting findings

As noted in the Introduction, the application scenarios for simulation programs invariably differ from the validation scenarios under which they were tested. It is important therefore to consider the degree to which the results from validation studies might be indicative of program performance under 'real-world' conditions. In other words, is it possible to estimate a 'domain of validity' for the simulation program, and are likely applica-

Table 4 Vertical daylight factor as a function of reflectance

		Vertical daylight factor [%]					
Location	1 (top)	28.1	30.3	32.8	35.7	39.1	43.1
	2	10.1	12.2	14.7	17.6	21.1	25.6
	3	4.4	5.8	7.6	9.9	12.9	16.7
	4	2.3	3.3	4.6	6.3	8.6	11.9
	5	1.3	2.1	3.1	4.5	6.4	9.2
	6 (bottom)	1.0	1.8	2.7	4.1	5.9	8.5
Reflectance		$r = 0.1$	$r = 0.2$	$r = 0.3$	$r = 0.4$	$r = 0.5$	$r = 0.6$

tion scenarios expected to fall within or outside of it? The 'domain of validity' contains those scenarios where the (lighting simulation) program is expected to perform with comparable accuracy to that shown in the validation, proving that:

a) the simulation is 'driven' correctly; and,
b) the scenario (buildings, sky, etc.) is accurately described in the simulation.

Necessarily, this consideration must be based in large part on reasoned extrapolations from one scenario to another. If it is suspected, say, from some test measurements, that the accuracy demonstrated in the validation cannot be repeated for the application scenario, then it is vital to determine the cause.

Rather than construct arbitrary application scenarios, an example is drawn from a recently published study[12] on the accuracy of illuminance predictions from *Radiance* using a scenario which appears, at least at first, to be very different from that used for the benchmark validation. The recent study was carried out in a heavily obstructed urban setting and it concluded that *Radiance* overestimated the vertical illuminance by $\sim$50% for high obstruction angles (i.e., at the low heights on the building facade). This was noted to be 'a very serious error indeed'.[12] Given that *Radiance* is one of the most widely used lighting simulation programs, and that many commercial and research applications depend on its reliability and perceived high-accuracy, the noted poor performance deserves attention. Was the simulation program used outside of its 'domain of validity', or was there some other cause for the reported failing of the program?

In order to answer this, both validation tests should, ideally, be repeated to ensure reproducibility of results. If the conflict in the assessment of accuracy persists, a hypothetical next-step could determine the program's accuracy for a scenario that is judged to be some intermediate of the BRE-IDMP office and the heavily obstructed setting. As neither of these options are immediately available, it is necessary to employ other methods. Two are suggested. The first is a comparison of the two scenarios: is there some fundamental, qualitative difference between the settings that might be the cause of the conflicting findings? Here it will be instructive to extrapolate by increments from one scenario to another. In the course of the extrapolation, it may be possible to identify a point at which there might be a compelling reason to believe that the validation results from one setting no longer apply to the other. The second is a critical comparison of the validation methodologies. The analysis of prediction errors resulting from imprecision in the description of the scenario (Section 3 and Section 4) will inform this examination.

5.1 Comparison of the BRE-IDMP and Hong Kong validation scenarios

For the heavily obstructed scenario in Hong Kong, measurements of the vertical daylight factor (that is, vertical illuminance) at three heights on a building facade were compared with predictions from the *Radiance* program. Measurements were taken during 'stable' sky conditions that had low ($<$25000 lux) global horizontal illuminance and the sky luminance pattern was assumed to be CIE Standard Overcast.[12] At first sight, the heavily obstructed scenario appears markedly different to the BRE-IDMP office, Figure 10. However, two simple geometric transforms applied to the BRE office scenario produce a scene that can be considered similar to a heavily obstructed urban site, at least for the purpose of predicting vertical illuminance. The two transforms are a 13 times magnification and a 90° rotation applied in any order, Figure 10(a). The BRE offices scene can now, in terms of the 'inside', be thought of as comparable to four tower blocks, each 117 m tall, placed so close together that the corners join. There is an overhang at the top (the sill in the office

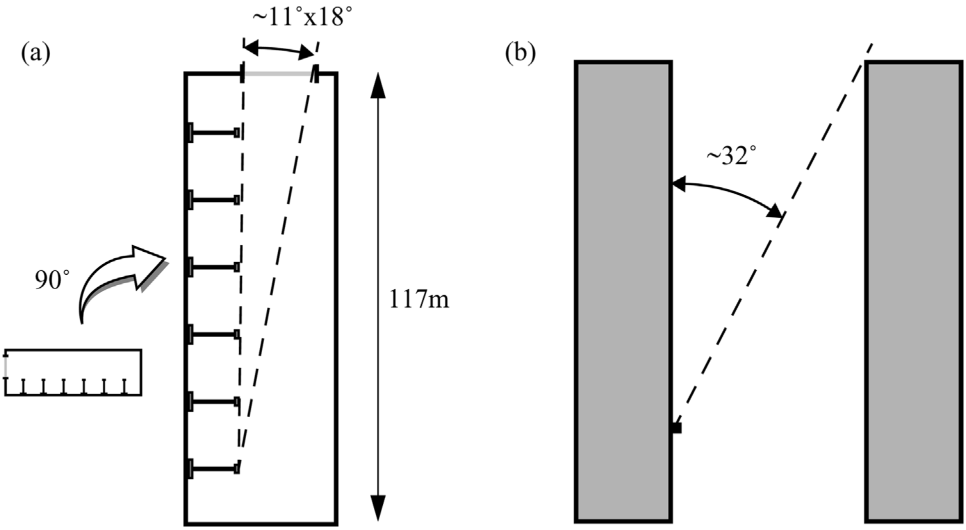

Figure 10 Rotational and Scaling transform of the BRE office (a) and schematic of tower block setting (b) with maximum obstruction angles indicated

model) that further obstructs the sky. For the photocell now at the base of the transformed office, the window subtends a solid angle approximately 11° by 18°. Viewed in this way, it is evident that the modeling of internal illuminance for the office scene is largely equivalent to modeling vertical illuminance in heavily obstructed space between four tall towers. Furthermore, the photocell at the back of the office was, in actuality, more heavily obstructed than any of the points on the tower block in Dr Ng's study, Figure 10(b). The six calculation points for the BRE office were in a horizontal row, whereas three points in the vertical plane were considered for the obscured building. For the BRE office, *Radiance* had to model light reflections off and between the vertical (walls) and horizontal (floor and ceiling) surfaces of the office space in order to accurately predict illuminance at all six photocell locations. It is highly unlikely therefore that the orientation of calculation points—top to bottom as opposed to across—can affect the accuracy of *Radiance*. It is important to note that the propagation of daylight—and the algorithms used in *Radiance* to model it—are not affected in any way by these geometrical transforms.[14] In other words, variations in scale and orientation alone are not sufficient to suggest an essential difference between the two scenarios.

Could sky conditions be a factor? For the Hong Kong study, the CIE standard overcast sky luminance distribution was used in the simulations. Overcast skies accounted for approximately 300 of the 754 unique skies in the BRE-IDMP dataset. Thus it cannot be said that the conditions (assumed) for the Hong Kong study were significantly different from a large number of skies in the BRE-IDMP validation. Moreover, the ratio between the maximum and minimum illuminances measured in the Hong Kong study was ~ 7 (as indicated by the vertical daylight factor) which falls within the range of ~ 5 to ~ 25 for the 754 skies in the BRE-IDMP dataset. The actual vertical illuminances that were recorded for the Hong Kong were not given but can be estimated from the plot of global horizontal illuminance (Figure 9 in Ng[12]). Values in the range $\sim 5-10$ klux are identified as candidate data that might be suitable indicators of overcast sky conditions. Based on this and the three vertical daylight factor values, vertical illuminances in the range ~300 lux to ~5000 lux were probably recorded. By comparison, illuminances in the range 7 lux to 12 666 lux were accurately predicted by *Radiance* in the BRE-IDMP study (for photocells not illuminated by direct sunlight). It can be reasonably claimed therefore that skies in the BRE-IDMP dataset encompass a very wide range of sky conditions—both in terms of sky type and the internal illuminance levels they produce. Thus in terms of sky type, illuminance ratios and absolute illuminance values, the conditions monitored in Hong Kong cannot be claimed to be extremes that were outside of the range found in the BRE-IDMP dataset.

The only significant difference between the two scenarios is quantitative rather than qualitative: the heavily obstructed scenario is geometrically more complex than the BRE-IDMP office scenario. This fact alone should not prove to be a fundamental limiting factor on the program's accuracy. Indeed, because the facades of the tower blocks were modeled as planar polygons, the difference in complexity of the *modeled* scenarios is much less than it would appear from photographs of the actual setting (Figures 8 and 10 in Ng[12]). Unless evidence can be found to the contrary, there is no compelling reason to suppose that the geometrical complexity of the obstructed scenario (as modeled) should be the cause of larger errors in predicted illuminance. It is likely to be the case however that the *Radiance* simulation parameters would need to be adjusted in response to the complexity of the scene to accurately model light transfer and achieve a converged result.[17]

5.2 Differences in applied methodologies

It was concluded in the previous section that differences between the two scenarios do not constitute sufficient reason to presume that the setting for the Hong Kong study lies outside of the 'domain of validity' established by the BRE-IDMP validation. The possibility that differences in the applied methodologies were the cause of the discrepancy are discussed here.

As noted, the BRE-IDMP validation was carried out under conditions that were carefully monitored and specified to a high degree of precision in the simulation. Perhaps the most notable difference between the BRE-IDMP validation and the Hong Kong study is that the former used measured sky luminance patterns whilst sky conditions were assumed to confirm to the CIE standard overcast for the latter. The following procedures were employed for the Hong Kong study to obtain 'useful data' (i.e., overcast conditions) from the 10 weeks of monitored data sampled at 5-minute intervals:[12]

- Weather reports from the Hong Kong Observatory were used to pin-point the cloudy days.
- Global horizontal illuminance was examined to identify periods when sky was fluctuating. A difference in excess of 10% between one data point to another was eliminated.
- Periods where the global horizontal illuminance exceeded 25 klux were eliminated.

The analysis given in Section 3 has shown that these procedures alone are not sufficient to ensure that actual sky conditions are a close match to the CIE standard overcast pattern. A low coefficient of variation in the four measurements of vertical illuminance (cases F, G, H and I) is needed to guarantee that actual luminance patterns closely match the CIE standard overcast pattern (Table 2). This was the case for skies in the UK. There is no compelling reason to presume that real overcast skies in Hong Kong diverge from the CIE overcast standard pattern in a fundamentally different way to that from UK skies. As noted in Section 5.1, the BRE-IDMP scenario shares many similarities with that for Hong Kong. Consequently, assessments of program accuracy based on a comparison of measurements under real skies against predictions using the CIE standard overcast may be similarly prone to erroneous conclusions (Section 3.5).

Illumination at the back of the BRE office or, equally, at high obstruction angles for the Hong Kong study, is dominated by inter-reflected light. Thus the predictions will be very sensitive to the reflectance assigned to participating surfaces. It is almost certainly the case that the BRE office reflectivities were measured to a higher degree of accuracy than those for the Hong Kong study. The BRE office consisted of a small number of homogenous planes e.g., walls, floor and ceiling. It was relatively easy therefore to measure their dimensions and reflectivities and describe the entire scene in the simulation at high precision. In the Hong Kong study, the actual building facades exhibit a marked degree of articulation (Figure 10 in Ng[12]). The analysis given in Section 4 shows that the measurement of facade reflectivity is subject to a number of confounding factors e.g., facade surface texture, stability of sky conditions, presence of specular (i.e., window) surfaces, differing daylight exposure for the test samples and reference standard, etc. Each of these are potential sources of significant imprecision in the determination of facade reflectivity. Furthermore, the demonstration examples of a heavily obstructed setting used in Section 4, showed that even moderate imprecision in reflectivity can lead to large uncertainties in vertical illuminance at high obstruction angles (i.e., at the base of the tower).

5.3 Summary

On the basis of the preceding analysis, it is highly probable that the poor accuracy in *Radiance* predictions for the Hong Kong setting resulted from imprecision in the model specification rather than the intrinsic accuracy of the simulation program. The two most important factors are likely to have been the assumption of CIE overcast sky conditions (Section 3) and the uncertainty in the effective reflectivity of textured building surfaces (Section 3) (Section 4). It is proposed therefore that the *Radiance* program is indeed capable of high accuracy predictions for heavily obscured urban settings, including Hong Kong. However, modeling such environments presents considerable challenges, not least in the specification of the buildings (geometry and reflectance) and the sky conditions.

6. Conclusion

Current and future use of lighting simulation programs depends in large part on their accuracy as demonstrated in published validation studies. For practitioners (e.g., consulting engineers), the results from simulation studies may play a major part in the environmental design and/or operation of a building. No less important is the role of simulation in research where a wide range of lighting-related phenomena are investigated using lighting simulation programs. Validation studies underpin the work of both practitioners and researchers using the simulation programs.

The study reported here has shown that validation of lighting simulation predictions in and around buildings under real sky conditions is subject to a number of confounding factors that can lead to erroneous findings. The assumption of CIE overcast sky conditions on the basis only of limited ranges and/or low fluctuations in global horizontal illuminance has been shown to be unreliable. In the absence of measured sky luminance patterns, only low variance in the four measurements (N, E, S and W) of vertical illuminance seems to guarantee close approximation of actually occurring sky conditions to the CIE overcast standard. Moderate imprecision in surface reflectivity can result in large uncertainties in illuminance and daylight factor predictions, particularly for heavily obstructed settings. The characterization and representation of urban buildings for lighting simulation presents many problems. Surface articulation and the presence of glazing can affect measurements of reflectivity, which may have view dependant (i.e., specular) properties that are problematic to account for accurately in the simulation.

It is prudent to both repeat established validation tests to ensure consistency of findings and to devise new trial scenarios to test predictions under different conditions. To this end, new validation studies are to be welcomed, especially those in urban settings. However, given the confounding factors noted earlier, and the practicalities of measurement 'in the field', it seems unlikely that urban settings could be represented in a lighting simulation with a precision equal to that used for the benchmark BRE-IDMP validation study. The BRE-IDMP scenario was a controlled, carefully monitored setting. It was also a real setting: a full-size office space under real sky/sun conditions. In addition, the BRE-IDMP setting has sufficient similarity to an urban setting (e.g., degree of obstruction) to reasonably assume that equivalently accurate predictions are possible, at least in principle if the sky conditions and the urban buildings could be represented with adequate precision in the simulation model. As a consequence, caution is advised in applying judgements on the intrinsic accuracy of a simulation program where any of the uncertainties in the model representation noted in this paper are present.

Acknowledgements

The study reported here was carried out in support of the activities of IEA Task 31 (Subtask C) and CIE TC 3.33. Sky luminance etc. data were supplied by Paul Littlefair and Maurice Aizlewood of the UK Building Research Establishment. The referees' comments on this paper are acknowledged.

7. References

1 Ward Larson G, Shakespeare R. *Rendering with radiance: the art and science of lighting visualization*. San Francisco: Morgan Kaufmann, 1998.
2 Grynberg A. *Validation of radiance*, LBID 1575, LBL Technical Information Department, Lawrence Berkeley Laboratory, Berkeley, California, July 1989.
3 Mardaljevic J. Validation of a lighting simulation program under real sky conditions. *Lighting Res. Technol*. 1995; 27: 181–88.
4 Rushmeier, HE Ward G, Piatko C, Sanders P, Rust B. *Comparing real and synthetic images: some ideas about metrics: proceedings of the Sixth Eurographics Rendering Workshop*, revised version in Hanrahan and Purgathofer, editors. Rendering Techniques 1995, Springer 1995.
5 Khodulev AB, Kopylov EA. *Physically accurate lighting simulation in computer graphics software: Proceedings GraphiCon 96: The 6th International Conference on Computer graphics and Visualization*. St. Petersburg, Russia, 1–5 July, 1996: Vol. 2, 111–19.
6 Mardaljevic J. *Validation of lighting simulation program: a study using measured sky brightness distributions*. Lux Europa 97 proc. 555–569, Amsterdam, 1997.
7 McNamara A, Chalmers A, Troscianko T, Gilchrist I. Comparing real and synthetic scenes using human judgements of lightness. In Piroche B, Rusheier H, editors. *Proceedings of the eurographics workshop in Brno*. Eurographics, June 2000. Comparing Real and Synthetic Scenes using Human Judgements of Lightness, EGWR 2000, Brno.
8 Mardaljevic J. *Daylight simulation: validation, sky models and daylight coefficients* PhD thesis. Leicester: De Montfort University 2000.
9 Mardaljevic J. The BRE-IDMP dataset: a new benchmark for the validation of illuminance prediction techniques. *Lighting Res. Technol*. 2001; 33.
10 Cannon-Brooks SWA. Simple scal models for daylighting design: analysis of sources of error in illuminance predicting. *Lighting Res. Technol*. 1997; 29: 135–42.
11 Bellia L, Cesarano A, Sibilio S. Daylight contribution of interior lighting: experimental

verification of software simulation results. *Lighting Res. Technol*. 1994; 26: 99–105.
12 Ng E. *A study on the accuracy of daylight simulation of heavily obstructed buildings in Hong Kong*. International Building Performance Simulation Association Conference, Brazil, 2001.
13 Galasiu A, Atif MR. Applicability of daylight computer, modeling in real case studies: comparison between measured and simulated daylight availability and lighting consumption. *Building and Environment*. 2002; 37: 363–77.
14 Hopkinson RG, Petherbridge P, Longmore J. *Daylight*. Londari Heinneman, 1966.
15 Enarun D, Littlefair P. Luminance models for overcast skies: assessment using measured data. *Lighting Res. Technol.* 1995; 27: 53–58.
16 Aizlewood ME. Innovative daylight systems: An experimental evaluation. *Lighting Res. Technol*. 1993; 24: 141–52.
17 Mardaljevic J. Daylight simulation. In Ward Larson G, Shakespeare R, editors. *Rendering with radiance: the art and science of lighting visualization*. San Francisco: Morgan-Kaufmann, 1998: Chapter 6, 341–90.

Discussion

Comment 1 on 'Verification of program accuracy for illuminance modelling: assumptions, methodology and an examination of conflicting findings' by J Mardaljevic

CF Reinhart (National Research Council, Institute for Research in Construction, Ottawa)

This excellent paper is required reading for anybody involved in validation studies of daylight simulation programs. Dr Mardaljevic rightfully questions some of the assumptions that are commonly made during the validation of lighting simulation programs. As a remedy, he proposes practical methods on how to filter out measured sky conditions whose luminance distributions resemble the overcast CIE sky and how to estimate the diffuse reflectances of neighboring facades.

Dr Mardaljevic's findings, that measured daylight factors do vary considerably under seemingly similar overcast skies, is in agreement with earlier work by Tregenza,[1] who concluded that measured daylight factors under overcast sky conditions can vary by over 50% in either direction, and Littlefair,[2] who observed that because the CIE overcast sky has a relatively dark horizon, it tends to underestimate internal illuminances in sidelit rooms. Based on Dr Mardaljevic's analysis it seems likely that the poor agreement of RADIANCE simulations of external vertical illuminances in Hong Kong with measurements can indeed be attributed to the use of the CIE overcast sky model and the way surrounding facades were modeled.

I was surprised that the author never suggested the use of a dynamic sky model such as Perez[3] (1993) instead of the CIE overcast sky model. The Perez model considers direct and diffuse illuminances or irradiances and has been demonstrated to adequately model the sky luminous distribution under a larger number of different sky conditions. I would also like to learn the author's opinion on the notion that—from a practitioner's point of view—a relevant parameter to characterize the accuracy of a daylight simulation program is the compounded error of the sky model and the lighting simulation program used together. While the error introduced by a sky model can in principle be side-stepped by using sky luminous distributions, this input is not widely available so that practitioners mostly have to use direct and diffuse irradiances or illuminances as simulation inputs.

References

1 Tregenza PR. The daylight factor and actual illuminance ratios. *Lighting Res. Technol.* 1980; 12: 64–68.
2 Littlefair PJ. Modeling daylight illuminance in building environmental performance analysis. *J. Illumin. Eng. Soc.* 1992; Summer; 25–34.
3 Perez R, Seals R, Michalsky J. All-weather model for sky luminance distribution—

preliminary configuration and validation. *Solar Energy* 1993; 50: 235–45.

Comment 2 on 'Verification of program accuracy for illuminance modelling: assumptions, methodology and an examination of conflicting findings' by J Mardaljevic

P J Greenup (Specialist Engineer, Arup Australasia, Sydney, NSW, Australia)

Dr Mardaljevic has provided yet another important and useful paper regarding the application of lighting simulation tools. This paper, like his several before it (references 3, 6, 8, 9, 17), covers important topics, and presents his work as rating amongst the best in the world in this field.

This paper effectively demonstrates the difficulties involved in performing validation of computer simulation of real world problems. The importance of this issue for validation of lighting simulation programs is demonstrated by the activities of Subtask C of the IEA Task 31—Daylighting Buildings in the 21st Century, and CIE Technical Committee (TC) 3.33—Test Cases for Assessment of Accuracy of Lighting Computer Programs. The objective of the latter group is to help lighting program users and developers to assess the accuracy of lighting computer programs and to identify areas of weakness. The work presented in the paper by Dr Mardaljevic will be highly valuable to the work undertaken by these two groups.

Dr Mardaljevic revealed that the assumption of CIE standard overcast sky conditions could be the cause of significant error (discrepancy between simulation and reality). Acknowledgement of this information has led to development of several alternative sky models, including those related by Igawa *et al*,[1] Kittler, *et al*,[2] and Perez *et al*.[3] The list of 15 standard sky types proposed by Kittler *et al* has recently become the new CIE Standard General Sky. It is suggested that future modelling and validation efforts should be based upon use of this new sky standard. A Radiance algorithm for application of this model is now available from the Radiance Online web site.[4]

Dr Mardaljevic's suggestion to model opposing facades as homogeneous surfaces with area-averaged reflectances is particularly appropriate for comparative studies. The primary purpose of most forms of building lighting simulation is to assist in the design process. As such, the general approach is to simulate several different design options (e.g., facade or glazing options) and compare results achieved between designs. In this situation, absolute results are not as important as comparative results between design options. Reasonable assumptions concerning interior furnishings and surfaces can be applied, and opposing facades can be modelled as homogeneous surfaces with area-averaged surface reflectances.

To ensure accuracy in comparative studies, software validations should assess the accuracy of comparative results obtained under 'standard' internal arrangements. The work of the CIE TC 3.33 is addressing this issue by individually validating numerous components of simulation models (e.g., directional transmittance of clear glass), rather than validating the whole package against reality. This approach to validation will assist in avoiding inaccurate assessments of the intrinsic accuracy of a simulation program, of the type described in section 5 of Dr Mardaljevic's paper.

The work presented in this paper displays Dr Mardaljevic's continued important contribution to the field of computer lighting simulation. The work performed by Dr Mardaljevic, the IEA Task 31 and CIE TC 3.33 will go a long way to providing practitioners with the confidence required in the simulation tools that they use, whilst also providing useful guidance on best practice in lighting simulation.

References

1 Igawa N, Nakamura H. All sky model as a standard sky for the simulation of daylit environment. *Building and Environment*, 2001; 36: 763–70.
2 Kittler R, Perez R, Darula S. A new generation of sky standards. *Conf. Lux Europa.* pp. 359–73, 1997.
3 Perez, R, Seals, R, Michalsky, J. All-weather model for sky luminance distribution—preliminary configuration and validation, *Solar Energy* 1993; 50: 235–45.
4 Greenup P. *SSLD Sky Model*, www.radiance-online.org/patches/ssld/SSLD algorithm with PDF.zip (2003).

Author's response to CF Reinhart and P Greenup

J Mardeljevic

The author welcomes the comments and issues raised by Dr Reinhart and Dr Greenup. Both rightly observe that general purpose sky models (Perez All Weather or the new CIE Standard General Sky) *may* provide more faithful luminance distributions than the CIE Standard Overcast Sky when using measured data. What the model skies cannot do, of course, is reproduce the unique, discontinuous luminance patterns that real skies exhibit (Figure 3). The issue of practical application is distinct from that of validation or the creation of benchmark datasets. Reliable validation studies require high-precision data collected under conditions that can be adequately described in the simulation. The goal is to determine the absolute accuracy of the prediction technique or simulation program with the minimum of uncertainty brought about by confounding factors. For any study, research or practical application, founded on performance under the CIE Standard Overcast Sky, the sky luminance pattern is predetermined and can be described (in simulations at least) with absolute precision. Thus the sky operates as a 'controlled' environment in which predictions are made rather than providing some description of 'reality'. This is not the case however where absolute values of time-varying illuminance are predicted. Here, the goal is to determine actual illuminance levels under realistic skies (with sun) based on 'local' meteorological data.[1,2] Accordingly, the analyses are usually founded on hourly Test Reference Year (TRY) data for a full year—any shorter period would be unrepresentative. The model(s) used for the hourly-varying sky luminance patterns is now a matter of choice for the operator. As indeed is the luminous efficacy model if irradiances need to be converted to illuminances. Notwithstanding these factors, the annual simulation offers a realistic measure of daylighting performance because it accounts for the full range of naturally occurring sky conditions and, importantly, illumination from the sun. In comparison, the daylight factor offers a caricature, a partial one at that, of daylight in and around buildings. The gulf in performance information between an annual time-series of illuminance values and a (static) daylight factor value is so great that it should inform our concerns regarding the compounded error and its significance. The compounded error itself has yet to be precisely defined. I would favour a definition that was based on the long-term (preferably annual) divergence between measured illuminances and those predicted using skies generated from basic quantities (i.e., comparable to those in TRY data). It is, of course, still important to identify the source of any significant divergence between measurement and prediction so that we know where to direct future effort—do we need better simulation 'engines' or sky model formulations, or both? What should not be in doubt is the need to advance beyond the daylight factor towards realistic measures of daylight illumination.

References

1 Mardaljevic J. The simulation of annual daylighting profiles for internal illuminance. *Lighting Res. Technol.* (2000); 32(3).

2 Reinhart C F Herkel S. The simulation of annual illuminance distributions — a state-of-the-art comparison of six RADIANCE-based methods. *Energy and Buildings*, 2000; 32:167–87.