Calibration and validation of climate-based daylighting models based on one-time field measurements: Office buildings in the tropics*

Geraldine Quek[1] **and J. Alstan Jakubiec**[2]

[1]Laboratory of Integrated Performance in Design (LIPID), School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[2]John H. Daniels Faculty of Architecture, Landscape and Design & School of the Environment University of Toronto, Toronto, Canada

ABSTRACT

Calibrated climate-based lighting simulation models of buildings have the capacity to perform an essential role in post-occupancy evaluations, such as annual frequency assessments of daylighting quality and visual discomfort. However, in most post-occupancy case studies the role of lighting analysis is temporally limited by instantaneous measurements or limited in scale by requiring constant monitoring with expensive sensors. It is challenging to build calibrated models based on point-in-time measurements due to the presence of electric lighting, transient use of dynamic shades, limited information on the material specifications, and short durations of accessibility to the spaces being studied. The authors propose and present a calibration process for annual daylighting and electric lighting simulation models based on one-time field measurements of large daylit and electrically-lit spaces exemplified through a data set of 540 individual office desks across 10 office spaces. The calibration process includes measuring lighting, physical and material data during a one-time visit that are used to calibrate high dynamic range images and lighting simulation models using actual weather data. The calibration accuracy is validated based on measured and simulated luminance and illuminance data. Comparing measured and simulated illuminance, relative RMSE values were 25.8% and 45.5% for horizontal and vertical measurements respectively. When tracking errors using log10(illuminance), approximating human perceptual differences, errors of 4.3% and 6.8% were achieved. Vertical illuminance was found to vary more with measured data due to the uncertainty of monitor screen luminances. The authors aim to achieve calibrated lighting models reliable enough to be used in assessing the relationship of annualized lighting metrics to participants long-term perceptions of lighting quality, thereby enabling simulation models to be used in the post-occupancy evaluation process of building lighting. This paper demonstrates that measured data through one-time visits can be utilized to build reliable calibrated lighting simulation models to integrate long-term annual lighting results in post-occupancy evaluations.

A version of this paper is published in:
LEUKOS: The Journal of the Illuminating Engineering Society of North America

Please cite this paper as:
Quek, G., Jakubiec, J. A. (2019). Calibration and validation of climate-based daylighting models based on one-time field measurements: Office buildings in the tropics. LEUKOS.
https://doi.org/10.1080/15502724.2019.1570852

*Accepted by LEUKOS on 14 January 2019

1 Introduction and Background

Simulating daylighting and electric lighting in digital architecture models in various stages of design has become quintessential for informing design decisions to meet quantitative and qualitative lighting goals by predicting building performance prior to construction. Post-occupancy evaluations (POE's) of completed buildings and spaces close the loop by providing feedback on the end-quality of a design to the building and construction industry who can assess if the project brief was met (Oseland, 2007) as well as serve an important role by allowing researchers to generate knowledge of occupant well-being based on actual user experiences. This can well lead to new recommendations to inform future designs of buildings. Most POE's focus on instantaneous measures of illuminance or visual comfort using High Dynamic Range (HDR) photography techniques within a monitoring time period. However, architects, energy consultants, engineers, lighting designers and researchers designing for daylight today use computer daylighting simulations and annualized lighting measures (Reinhart and Fitz, 2006) and recently including climate-based daylighting metrics (CBDMs) and annual visual comfort analysis to assess potential designs instead of static daylighting metrics (Reinhart et al., 2006). To truly 'close the loop', these annual daylighting metrics need to be evaluated based on a comparison with overall user perceptions of a space and the annual lighting they experience. To this end, this manuscript describes a process of calibrating daylight simulation models during POE field studies which can be used to calculate annual CBDMs as a component of building POE's based on short-term visits instead of long-term monitoring. Having said that, the proposed methodology does not aim to replace current POE methods but rather to extend and refine the capabilities of POE's by correlating actual annual daylighting metrics to the overall queried experiences of occupants. Models of 10 office buildings with 540 calibration points—occupant seating locations of participants—are calibrated and validated in this manner as a proof of concept. This study is the first large-scale calibration effort that aims to utilize long-term lighting simulation data for post-occupancy studies, enabling the comparison of CBDMs to occupant impressions in daylit spaces over a long period of time. Although the calibration and validation study is situated in the tropical climate in Singapore, the proposed work flow is applicable to any climate.

Many measurement-driven POE studies in the research literature are based on single instantaneous measurements and do not utilize long-term monitoring nor annual predictive data. Bear and Bell measured illuminance, source luminances, surface reflectance, geometric factors and subjective information for 471 participants as early as 1992 (Bear and Bell, 1992). Osterhaus surveyed 86 participants in 9 office spaces (Osterhaus, 2001). Parpairi, et al. measured luminance values manually and illuminance at limited points in time in 3 libraries from 26 typical users of the spaces (Parpairi et al., 2002). Dahlen, et al. measured discomfort glare metrics using High Dynamic Range (HDR) photography along with illuminance and subjective information from 298 female college students over a two month period, categorizing overcast and clear sky conditions separately (Dahlan et al., 2009). Choi, et al. measured instantaneous luminance using HDR photography and illuminance, among other IEQ factors, at 402 occupant desks using a mobile sensor cart, although they spread their measurements out over a period of time in order to collect some level of seasonal information (Choi et al., 2012). Hirning, et al. measured discomfort glare using HDR photography for 419 occupants not experiencing discomfort due to reflections from their monitor screen and photographs were taken from the occupant's point of view close in time to when participants filled out a short subjective survey (Hirning et al., 2013, 2014). Mangkuto, et al. used HDR photography to measure luminance and glare metrics in an Indonesian library space over a 6-day period, collecting 86 measures paired with subjective information (Mangkuto et al., 2017). Hirning, et al. repeated their study in Malaysia with 341 participants in 6 separate office spaces (Hirning et al., 2017). In summary, for direct

measurement POE studies, it is feasible to quickly collect a large number of data from a diverse range of participants. The mean number of subjects or data points in these studies is 292. However, limitations do emerge—the vast majority of studies collect limited seasonal or temporal data, localizing the results at the specific point-in-time the measurements are taken.

Another approach employed by researchers has been to utilize long-term data monitoring in order to gather a more holistic representation of POE participant's lighting experiences. Fan, Painter, and Mardaljevic recorded HDR photographs and frequent subjective information using a computer application for 5 occupant workstations. Cameras were mounted as close to the head position of the occupant as feasible, which resulted in errors generally below 25% (Fan et al., 2009). Their study was later carried on for a period of one year (Painter et al., 2010). Konis recorded ambient environmental conditions, HDR photographs, and subjective ratings on a continuous scale for 14 participants over a 2-week period collecting 523 data points overall. A custom sensor and polling device was created and placed upon occupant's desks; it prompted participants periodically for subjective feedback using a physical slider at which time a HDR photograph would be taken (Konis, 2013, 2014). Drosou, et al. installed 2 high-quality HDR camera capture setups in classrooms, monitoring their lighting performance every 10-minutes for an entire year. Limitations included a period of data loss due to camera shutter failure and a single, fixed viewing location not from an occupant point of view (Drosou et al., 2016). Bellia, et al. continuously measured occupants in 3 private offices during a 3-day period while measuring luminance and illuminance (Bellia et al., 2017). Extrapolating from the above studies, continuous measurement studies are limited by the cost and maintenance of equipment, thereby enhancing the quality of data per participant but limiting the number of participants for which data can be reasonably and affordably collected.

A third, but less utilized, approach to gather experiential daylighting data inside of spaces has been to employ daylighting simulations, often using a validated Radiance-based engine (Ward, 1994) such as Daysim (Reinhart and Walkenhorst, 2001). Reinhart, et al. used quasi-calibrated CBDM results paired with subjective data to identify annual lighting performance levels that correlate with perceptions of 'daylit.' (Reinhart et al., 2014) Jakubiec and Reinhart used a daylighting model based on measured material properties, exacting geometric reconstructions, and specific weather data to assess annual lighting and glare metrics for 123 participants in a POE study at a 6-minute time interval (Jakubiec and Reinhart, 2016), but the model's calibration was not checked with measurements, and contributions from electric lighting were ignored. Bellia expressed the opinion that this approach was not feasible due to the complexity and time commitments of modeling data. (Bellia et al., 2017) Mardaljevic, et al. noted that CBDMs are difficult to validate in practice due to obstructions on the workplane where sensors would ordinarily be placed for long-term monitoring and that illuminance data is not ordinarily a part of building management systems. (Mardaljevic et al., 2016) They propose to use a continuous luminance camera to derive illuminance on vertical surfaces in order to validate CBDM illuminance calculations. Other researchers have noted the importance of appropriate material properties in simulations to achieve accurate results. (Jakubiec, 2016; Brembilla et al., 2015)

While not commonly utilized in POE lighting studies, calibrating CBDM daylighting models allows researchers to refine the POE process and extend it to design processes by analyzing the same annualized lighting metrics. Instead of a long-term monitoring approach, physical and lighting measurements are collected during a short visit, and calibrated daylighting models are then built and validated before being used to simulate annual CBDM results. Time-series data from the occupants' point of view can be easily obtained for an unlimited number of participants, not limited by sensor availability or manpower. However,

limitations are given by the accuracy of the models and a difficulty in evaluating their accuracy before being used to simulate annual daylighting conditions. Therefore this paper takes an approach to validate the performance of lighting simulation models for utilization in POE lighting studies based on limited measurements. Specifically, the problem of simultaneous presence of electric lighting and daylighting at the same time is addressed. First, a methodology is presented by which detailed field measurements are taken in currently occupied buildings and used to calibrate a combined daylighting and electric lighting simulation model. Next, this method is applied to the calibration of 10 models of office spaces, and model accuracy is validated using standard root mean squared error and mean bias error in addition to perceived errors modified for logarithmic human vision. Finally, the discussion focuses on what the methodology proposed in this manuscript means for future work in POE lighting analysis and research.

2 Methodology

Measurements at office desks of 540 occupants in 10 offices were recorded during the period between October 2016 and August 2017. There were no specific criteria applied when choosing the office buildings —just to be an office building located in Singapore. The offices the authors measured were a mix of public and private sector office buildings. Some characteristics of the 10 offices are detailed in Table 1, including the typology, types of roller blinds, glazing transmittance (Tvis), window to wall area ratio (WWR) and if any Green Mark (GM) award was attained. Green Mark is a green building certification system in Singapore which mirrors that of others known certification systems such as LEED, CASBEE, Green Star, etc. Calibrated climate-based daylighting models were then built according to field measurements: HDR images, illuminance and luminance measurements, material reflectance measures, space measurements, and external weather data. This section describes measurement techniques and calibration steps used to create and validate calibrated models for further climate-based daylighting analysis. Figure 1 explains the general overall workflow. Computational scripts were written to automate each process due to the large amount of data.

Office Building	Typology	Rollerblinds*	Tvis	WWR	Award
1	12-Storey Office Tower	O,T	25%	62%	GM Platinum
2	17-Storey Office Tower	O,T	40%	44%	GM Platinum
3	20-Storey Office Tower	T	35%	64%	GM Platinum
4	27-Storey Office Tower	O,T	10%	60%	GM Gold
5	Office space in 5-Storey Shopping Mall	T	15%	25%	-
6	Office space in 10-Storey Industrial Building	T	55%	22%	-
7	3-Storey Retrofitted Building	T	50%,56%	60%	-
8	27-Storey Office Tower	T	37%	62%	GM Platinum
9	17-Storey Office Tower	O,T	32%	60%	GM Platinum
10	27-Storey Office Tower	O,T	37%	62%	GM Platinum

**O: Opaque, T: Translucent*

Table 1. Characteristics of the ten measured office spaces

2.1 Field Measurements

Of the 10 offices measured, some occupy multiple floors (Offices 1, 8 and 10 in Table 1), and 2 offices are situated in the same building separated by 12 floors(Offices 8 and 10 in Table 1). Beyond formal and layout differences, each office has different material finishes and luminaire selection. High Dynamic Range (HDR) photographs, workplane and vertical illuminance, and luminance measurements from a neutral grey card were captured and recorded at each of the 540 occupants' desks. Each set of measurements took approximately 3 to 5 minutes to complete. Occupants were asked immediately preceding the measurements

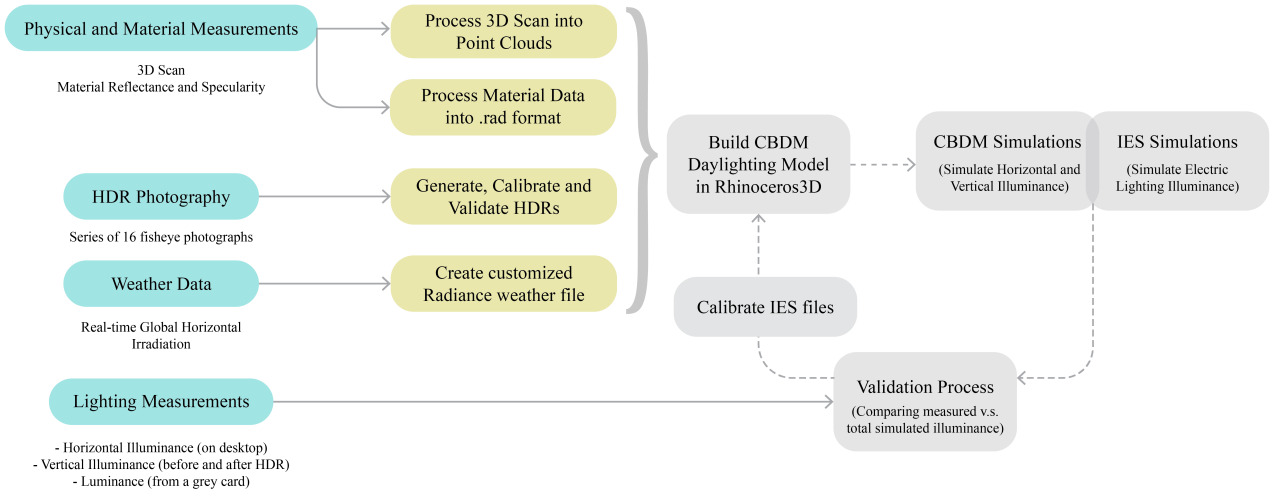


Figure 1. Overall workflow for calibrating climate-based daylighting models from single point-in-time measurements

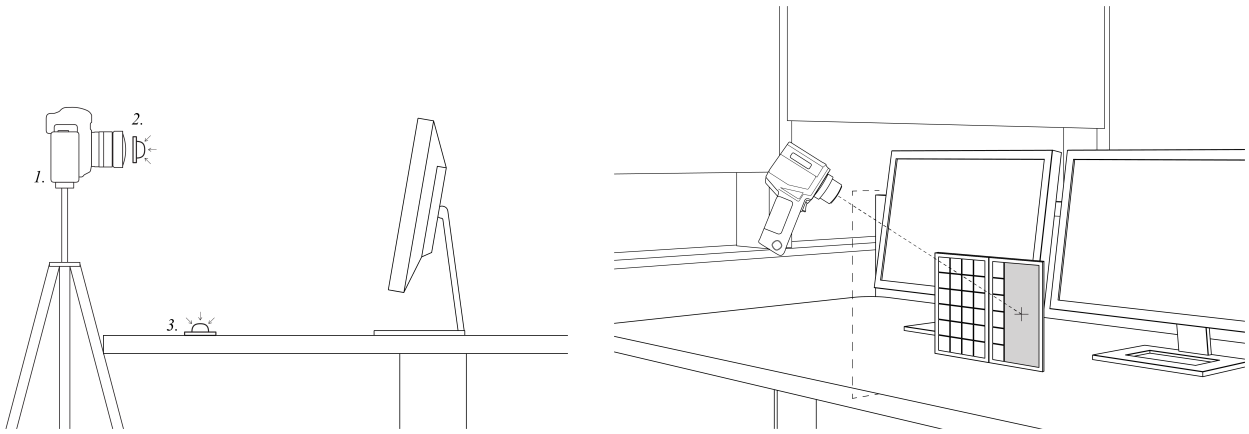
to fill out a short subjective survey on their perceptions of instantaneous and long-term lighting quality; however, those subjective components are not the focus of this paper.

2.1.1 Illuminance and Luminance Measurements

Before measurements are taken, occupants are asked to leave their desk such that the camera and lighting sensors can be positioned freely at the eye position and on the workplane. Horizontal illuminance (lux) on the desktop surface is measured once, after the HDR photographs are taken. Figure 2 illustrates the positioning of the illuminance and luminance meters for the respective measurements. Luminance measurements (cd/m^2) from a grey card are recorded with a luminance meter (Konica Minolta LS-100), and vertical illuminance (lux) is recorded with an illuminance meter (Konica Minolta Illuminance Meter T-10A or Illuminance Spectrophotometer CL-500A) before the first exposure and after the last exposure of the HDR photograph are captured, in front of the fisheye lens from the occupant's viewpoint. The average value of the before and after luminance measurements are used in calibrating the images, to account for minor changes in lighting levels during the HDR capture. If illuminance or luminance values differ significantly, the image is discarded.

2.1.2 HDR Photography

HDR fisheye photographs were taken at each occupant's desk to capture luminance values from their point of view shortly after filling out the subjective survey. The methods used are according to the recommendations proposed by Inanici and Jakubiec, et al. (Inanici, 2006; Jakubiec et al., 2016). The occupant is asked to vacate their desk before the measurements start—see Figure 2a. A full frame DSLR camera (Canon EOS 5D Mark III) with a fish-eye lens (Canon EF 8-15mm f/4L Fisheye USM or Sigma 8mm f/3.5 EX DG Circular Fisheye Lens) was used with a stable tripod. A series of 16 photographs with exposure times at an interval of 1 stop from 8 sec. to 1/4000 sec. were taken. A f-stop of f/11 and ISO speed of 100 were used throughout. The white balance was set to the daylight setting. Examples of HDR luminance maps of randomly selected participant desks from the 10 office spaces are shown in Figure 3. The individual exposures were converted to the Radiance angular fisheye (-vta) image projection (Ward, 1994) based on measurements of the lens projections taken with a panoramic tripod head. The



(a) 1. Camera setup for HDR photography, 2. Vertical illuminance measurement location, 3. Horizontal illuminance measurement location

(b) Luminance measurement setup

Figure 2. Measurements and HDR photography setup

images are then cropped to a 180 degree opening angle with a circular mask and resized to 800x800 pixels. The series of photographs were then converted into HDR images using Photosphere (Ward, 2005). Approximately half of the HDR images captured in this study did not have a corresponding luminance measurement. These images were calibrated based on vertical illuminance measurements taken at the camera lens. Images with a corresponding luminance measurement were then calibrated in Photosphere manually using the luminance data measured on the grey card. Finally, vignetting correction (Jakubiec et al., 2016) was applied to the HDR photos to correct for light attenuation at the edge of the fisheye photograph.

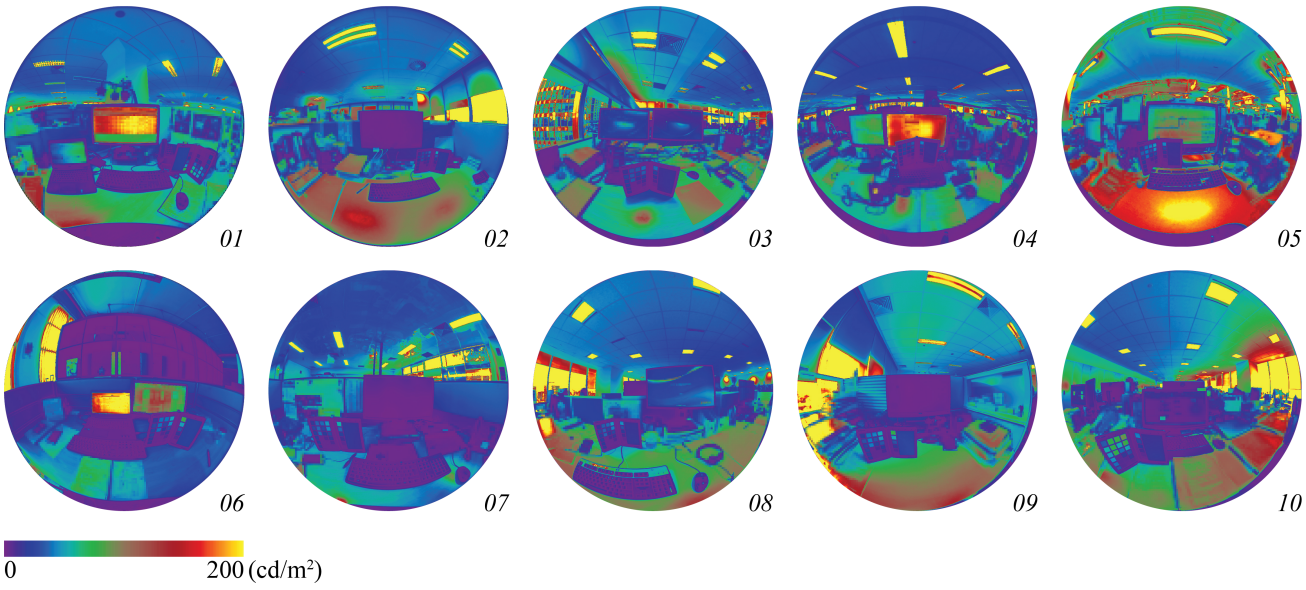


Figure 3. An example of HDR image collected from each of the 10 office spaces measured

2.1.3 Physical Measurements

A laser scanner (FARO FOCUS 3D X330) was used to scan the interiors and exteriors of the office spaces. This acts as a geometric reference for the precise construction of a 3D model. The 3D scan data, which is in the form of a point cloud, approximates the interior layout including dimensions and positions of furniture and electric lighting—see Figure 4 for an example. The scan files were processed and exported and imported into Rhinoceros 3D (McNeel and Associates, 2017) and used to build a detailed surface model. Due to the large file size of point clouds, screen captures of the relevant clipped views and measurements in 3D space using a point-cloud analysis software were used to calibrate the geometry of the model, instead of the actual point cloud. However, laser scanning is an optional step and may be replaced by hand measurements, or using as-built BIM-models after verifying their measurements. The surface geometry created in this step is converted to polygonal meshes and exported to the Radiance .rad format (Ward, 1994) using the DIVA plugin (Jakubiec and Reinhart, 2011).

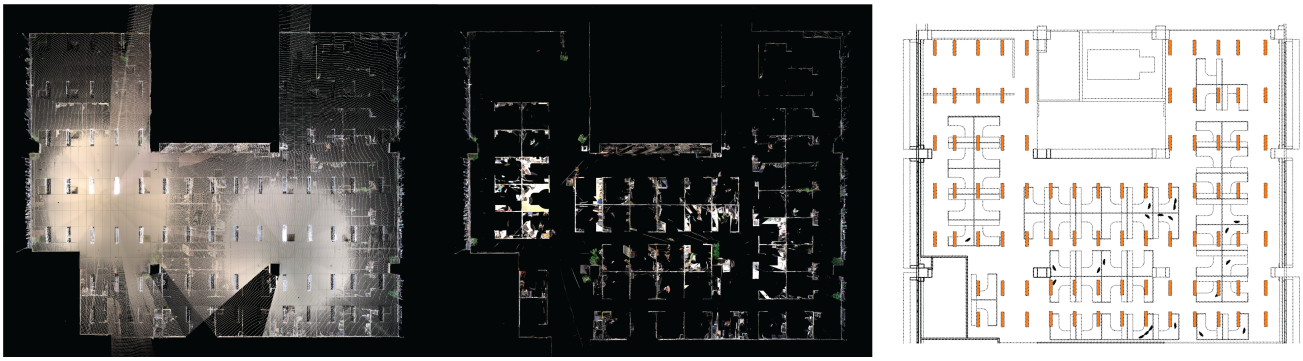


Figure 4. Example of 3D scan point cloud and screen captures of furniture plan and luminaires (orange rectangles) from Office 1

2.1.4 Measurement of Material Reflectance Data & Translation to Radiance Material Definitions

Material reflectance data of surfaces were measured using a spectrophotometer (Konica Minolta CM-2600d Spectrophotometer) based on an average value of three-point measurements across each surface material type. Each material finish had two types of reflectance recorded, Specular Component Included (SCI) and Specular Component Excluded (SCE) respectively which are both full-spectrum colorimetric measurements. The difference between SCI and SCE is used to derive the specular versus diffuse components of reflectance. A photo is also taken for reference, and roughness values are estimated based on the visual appearance of the images (Jones and Reinhart, 2017). The material data was converted to Radiance material definitions as described by Jakubiec (Jakubiec, 2016). Where the glazing transmittance information is not obtainable, it was either measured if the windows are operable or estimated from initial visualization simulations. Where the glazing transmittance was known, it was converted to transmissivity by multiplying by 1.09 (Jacobs, 2012). A generic monitor model was used for all the desks, modeled according to Jones (Jones and Reinhart, 2017) but with modifications to match the average observed screen luminances in Singapore with a high-state pixel luminance of 125 cd/m^2 and low-state pixel luminance of 33 cd/m^2. A generic roller blind with a direct normal transmittance of 0.01, front reflectance of 0.37, back reflectance of 0.56 and diffuse transmittance of 0.2, and transmittance that drops to zero at 84 degrees, was used. It can be accordingly calibrated for specific roller shade types in different offices. A typical set of Radiance material definitions measured and produced for one of the office spaces is shown in Table 2. Alternatively, cheaper and substantially accurate methods to measure material reflectance are also available, such as the illuminance method, using the CIBSE color chart or RAL color fan, according to

Gradillas' comparisons. (Gradillas, 2015) Another possible technique is one developed by Mardeljevic et al., using HDR imaging to derive surface reflectance maps through a interpolated illuminance map and a HDR luminance map. (Mardaljevic et al., 2015)

	Material	Radiance Definition
Interior	Wood Laminate Table Top	void plastic TableTop 0 0 5 0.6112 0.4779 0.3081 0.0141 0.15
	Beige Partition Fabric	void plastic PartitionFabric 0 0 5 0.6249 0.5803 0.4762 0.0097 0.10
	Grey Carpet	void plastic GreyCarpet 5 0.0731 0.0708 0.0654 0.0000 0.40
	Grey Mullions	void plastic GreyMullions 5 0.4618 0.4716 0.4765 0.0374 0.05
	White Wall	void plastic WhiteWall 5 0.8911 0.8933 0.8495 0.0107 0.30
	White Column	void plastic WhiteColumn 5 0.8884 0.8896 0.8423 0.0113 0.30
	Acoustic Ceiling Panels	void plastic AcousticCeilingPanels 5 0.8752 0.8717 0.8471 0.0079 0.40
	Light Shelf (Bottom)	void plastic LightShelfBottom 5 0.4851 0.4963 0.4958 0.0552 0.05
	Light Shelf (Top)	void plastic LightShelfTop 5 0.4929 0.5007 0.4951 0.0515 0.05
	Opaque Roller Shade	void plastic OpaqueShade 5 0.5812 0.5560 0.4676 0.0059 0.20
	Glazing	void glass Glazing 3 0.3815 0.3815 0.3815
	Roller Blinds	void BRTDfunc RollerBlinds 10 0 0 0 tspec tspec tspec 0 0 0 mechoshade.cal 0 9 0.37 0.37 0.37 0.56 0.56 0.56 0.2 0.2 0.2
Exterior	White Painted Wall	void plastic WhiteExteriorWall 0 0 5 0.4067 0.4968 0.5054 0.0050 0.3
	External Mullions	void plastic ExternalMullions 0 0 5 0.4732 0.5545 0.5009 0.0252 0.15
	Decorative Floor Tiles	void plastic DecorativeExternalFloorTiles 0 0 5 0.1545 0.1628 0.1390 0.0089 0.2
	Specular Steel Handrail	void plastic SpecularSteelHandrail 0 0 5 0.2247 0.2896 0.2991 0.3287 0.05
	Asphalt	void plastic Asphalt 0 0 5 0.1086 0.0998 0.0850 0.0004 0.4
	Wood Plank Walkway	void plastic WoodPlankWalkway 0 0 5 0.1162 0.0958 0.0841 0.0011 0.4
Monitor	Screen	void trans MonitorScreen 0 0 7 0.575 0.575 0.575 0.033 0.01 0.88 1
	High-State Pixel	void glow MonitorHigh 0 0 4 1.396 1.396 1.396 0
	Low-State Pixel	void glow MonitorLow 0 0 4 0.3352 0.3352 0.3352 0
	Dark Plastic	void plastic MonitorPlasticBlack 0 0 5 0.054 0.054 0.062 0.013 0.1
	Light Plastic	void plastic MonitorPlasticSilver 5 0.464 0.470 0.452 0.078 0.1

Table 2. An example list of Radiance material definitions from Office 5)

2.2 External Measurements

Real-time weather data was collected from an existing weather station located at the rooftop of the authors' university campus at 36 m above ground level with no urban obstructions. Global horizontal solar irradiation (W/m^2) is measured from a silicon pyranometer and recorded every 5 minutes by a data logger. Global horizontal solar irradiation was split using the Reindl (Reindl et al., 1990) method into direct-horizontal and diffuse-horizontal irradiance. A customized Radiance .wea weather file (Reinhart and Walkenhorst, 2001) was then created for each office space back-dated one year from the date of visit. For example, if the date of visit was 16 May 2017, a custom weather file was created between 16 May 2016 and 17 May 2017.

2.3 Validation of HDR Images

To validate the accuracy of the HDR images, which will be used for glare analysis and point-in-time calibration of the daylighting model, measured vertical illuminance, E_{mea}, was compared to the total pixel illuminance contribution from the equi-angular HDR images E_v, as calculated in equation 1. As the HDR images with illuminance measurements were calibrated to the measured illuminances, only the HDR images that were calibrated through the luminance measurements are included in this validation, shown in Figure 5. The grey line illustrates the ideal calibration of HDR images, and it is notable that the total pixel illuminance contribution fell slightly below than sensor measured vertical illuminance for luminance-calibrated images.

$$E_v = \sum_{\theta_p < 90^\circ} L_p \, \omega_p \cos \theta_p \tag{1}$$

where L_p is the luminance (cd/m^2) of the individual pixel, p, ω_P is the solid angle (str) of that pixel, and θ_p is the incident angle from the pixel to the center of the photograph. The authors then calculate the root mean squared error (RMSE) between the HDR illuminance and sensor-measured illuminance with equation 2 as follows:

$$RMSE = \sqrt{\frac{\sum(E_v - E_{mea})^2}{n}} \tag{2}$$

where E_v is the total pixel illumination contribution derived from n number of HDR images, E_{mea} is the measured vertical illuminance. The relative RMSE, $RMSE_{rel}$, refers to the percentage deviation from the mean while relative mean bias error MBE_{rel} is the Mean Bias Error relative to the mean. Both were calculated according to Equations 3 to 5:

$$RMSE_{rel} = \frac{RMSE}{\bar{E}_{mea}} \tag{3}$$

$$MBE = \frac{\sum(E_v - E_{mea})}{n} \tag{4}$$

$$MBE_{rel} = \frac{MBE}{\bar{E}_{mea}} \tag{5}$$

where $\bar{E}_{mea}$ is the mean of the measured vertical illuminances. RMSE on the logarithmic scale is also calculated, as the photopic sensitivity response of the human eye to lighting intensity is on a logarithmic scale rather than on a linear basis.(Reinhart and Andersen, 2006) Hence $logRMSE_{rel}$ is also calculated as per Equations 6 to 7:

$$logRMSE = \sqrt{\frac{\sum(log_{10}E_v - log_{10}E_{mea})^2}{n}} \tag{6}$$

$$logRMSE_{rel} = \frac{logRMSE}{log_{10}\bar{E}_{mea}} \tag{7}$$

The results showed a $RMSE_{rel}$ of 23.24% in linear space, $logRMSE_{rel}$ of 9.60% in logarthmic visual space, a slight negative bias of -9.91% MBE_{rel}, where MBE in linear space is -21.9lux, as illustrated in Figure 5.

2.4 Calibration of Electric Lighting

As the exact model and lamp selection of the luminaires were not available, an appropriate IES file (Committee et al., 1991) was selected to be used in the simulation models, based on luminaire dimensions and photometric distributions observed in the captured HDR images. After loading the IES file into the lighting simulation model, the luminaires were located as per the 3D scan data. An occupant's viewpoint was simulated using high-quality ambient Radiance parameters (see Table 3), and the global horizontal

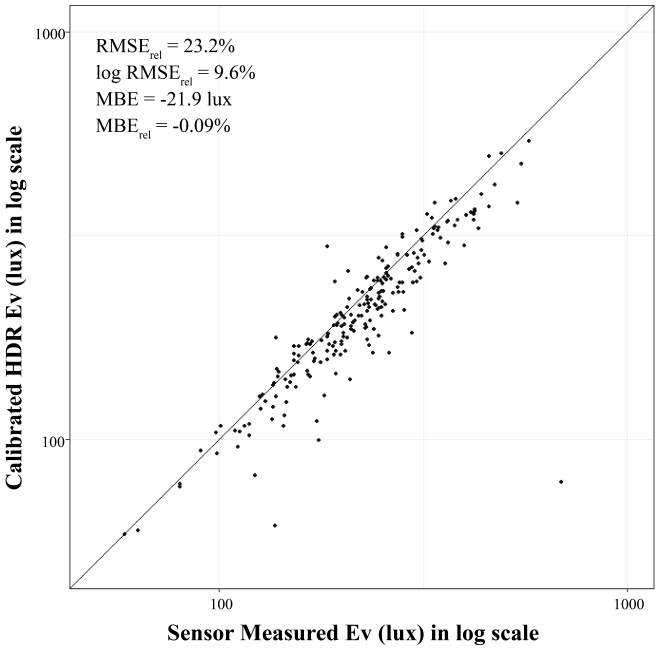


Figure 5. Total pixel illuminance contribution v.s. Measured Vertical Illuminance for luminance-calibrated HDR images

irradiance value that is recorded by the weather station nearest to the time of field measurement was used as the input for the Perez all-weather sky model (Perez et al., 1993) during a calculation. To keep daylight contribution to a minimum, a viewpoint further away from the facade was chosen when calibrating the photometric distribution and intensity of electric lighting data. A scale factor is approximated and *ies*2*rad* (Ward, 1994) was ran to multiply the brightness output of the luminaires. This process was repeated until a appropriate scaling factor was reached as shown in Figure 6. Although a desk further away from the facade is selected to reduce errors due to the high variability of daylight which may skew the calibration process of electric lighting, daylight is still included as ambient lighting in the scene.

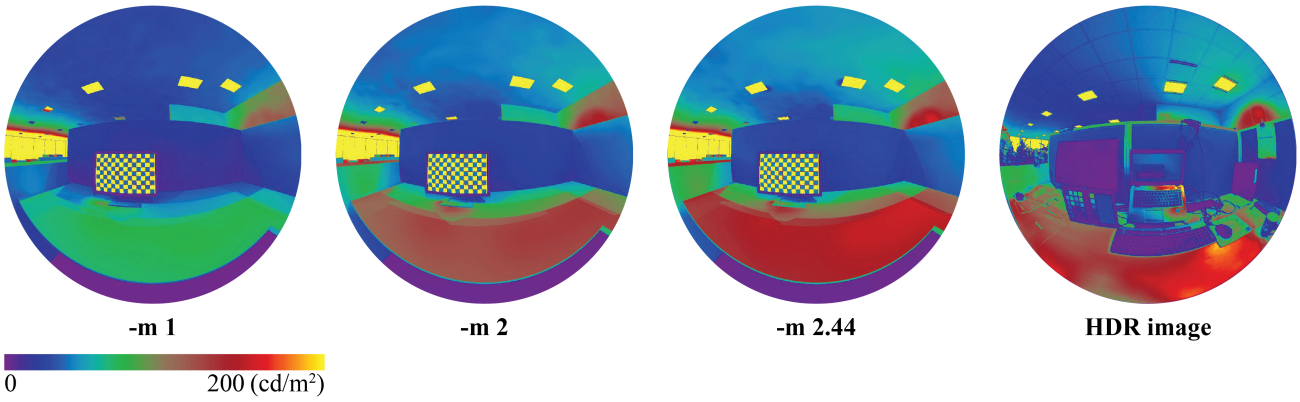


Figure 6. Calibration of electric lighting

2.5 CBDM Simulation Data and IES Simulation Data

Grid-based annual lighting simulations without roller blinds and electric lighting were ran for all 10 office spaces indicating the various daylighting conditions due to the diverse building typologies, materials, floor plans and building depths. Figure 7 illustrates the variation of daylight experienced by the participants in the 10 surveyed office spaces represented by Annual Mean Illuminance. A red line illustrates a Daylight Autonomy value of [75%] at an illuminance threshold of 300 lux. For validation purposes, climate-based daylighting simulations were ran using Daysim (Reinhart and Walkenhorst, 2001) for each office space, with roller blinds, per 5-minute time step using the custom weather data file generated for each office space. The Radiance parameters used for the climate-based simulations are as shown in Table 3 with 6 ambient bounces. Since light is additive, the simulated daylight and electric light illuminances were simulated separately and are summed up to calculate total simulated horizontal or vertical illuminance from electric lighting and daylighting. Electric lighting simulations includes only light from luminaires and monitor screens and are simulated at 4 ambient bounces instead of 6. These simulated values at the nearest simulated time step to the field HDR and illuminance measurements are validated against the measured lighting data which are detailed further in the next section.

Radiance Parameters	Visualizations	Climate-based Simulations
Ambient bounces (ab)	4	6
Ambient divisions (ad)	1000	2000
Ambient super-samples (as)	500	1000
Ambient resolution (ar)	500	1000
Ambient accuracy (aa)	0.1	0.1

Table 3. Typical Radiance parameters used for visualizations and validation simulations

3 Results

After the iterative process of calibrating the electric and daylighting portions of the lighting simulation models (see Figure 1), the authors compared the accuracy of the results with measurements by extracting the point-in-time simulation data nearest to the time of measurement for each of the 540 occupant desks measured during the 13 month period of this study. $RMSE_{rel}$ values of 25.8% and 45.5% were achieved between sensor measured illuminance and simulated point-in-time illuminances horizontally and vertically as illustrated in Figure 8. The grey line is a $y = x$ identity line representing the ideal calibration situation. Vertical illuminance deviates from the measured illuminance more than that of horizontal illuminance. Vertical $RMSE_{rel}$ values are exasperated by the variety of monitor types and configurations measured at various workstations where on/off status, the brightness setting, monitor type, monitor size, and monitor number vary wildly between participants; all of these will have an impact on measured vertical illuminance. Logarithmic differences are known to express perceived lighting differences by the human eye better than actual absolute differences in different lighting situations. (Reinhart and Andersen, 2006) Hence, the authors decided to use $logRMSE_{rel}$ as a paired validation measure of the CBDM daylighting models indicating how close calculations come to human perceptual differences in addition to standard linear lighting unit errors. In the base-10 logarithmic visual space, the models had 4.3% and 6.8% $logRMSE_{rel}$ for horizontal and vertical illuminances respectively. The overall MBE was 0.3 lux (0.0006%) and 61.8 lux (23.3%) for horizontal and vertical simulated illuminance respectively.

Figure 10 compares two selected calibrated HDR images to simulated visualizations using falsecolor luminance images scaled between 0 cd/m^2 and 300 cd/m^2. Overall these images are representative of

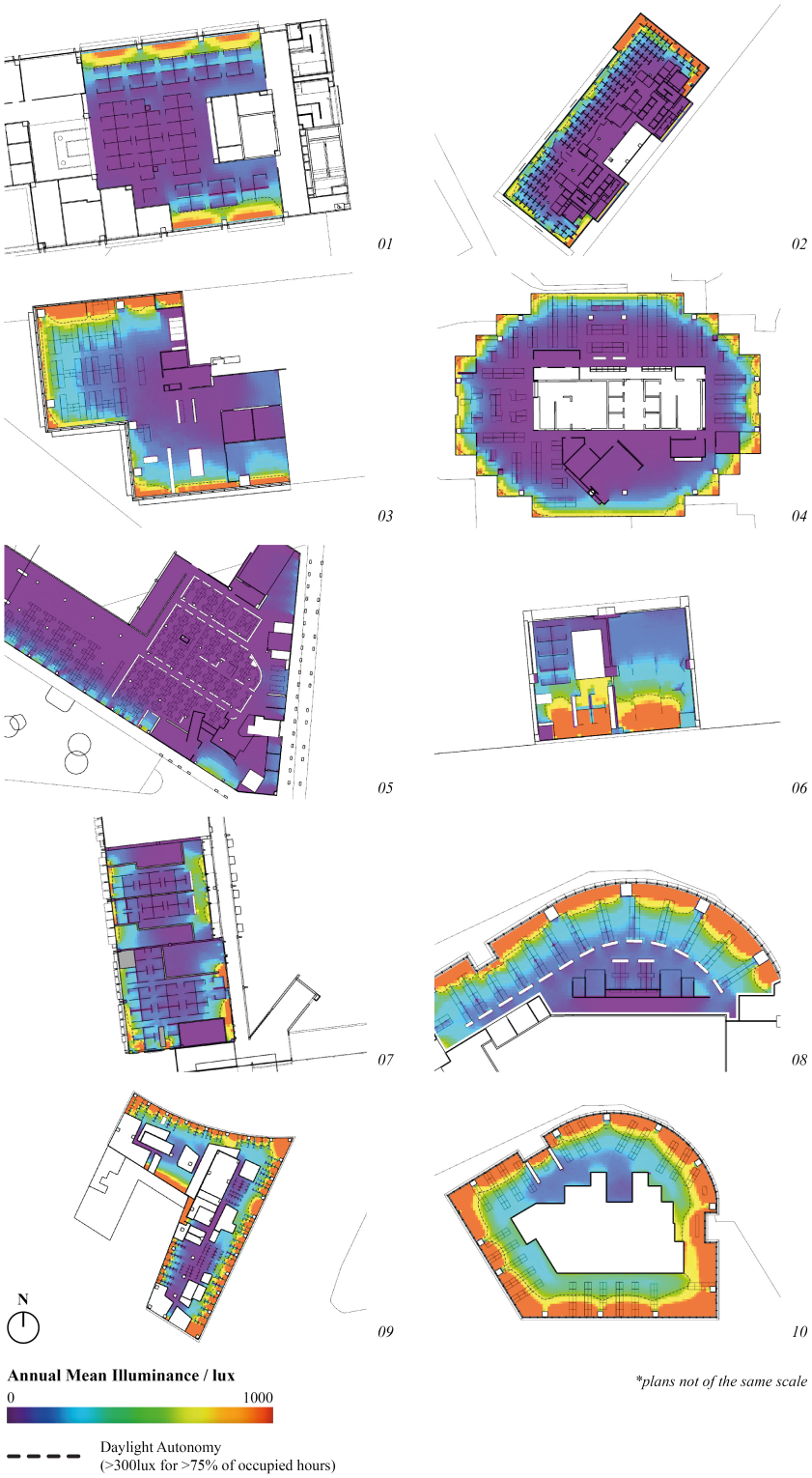


Figure 7. Annual Mean Illuminance and Daylight Autonomy for all 10 offices simulated (without roller blinds and electric lighting)

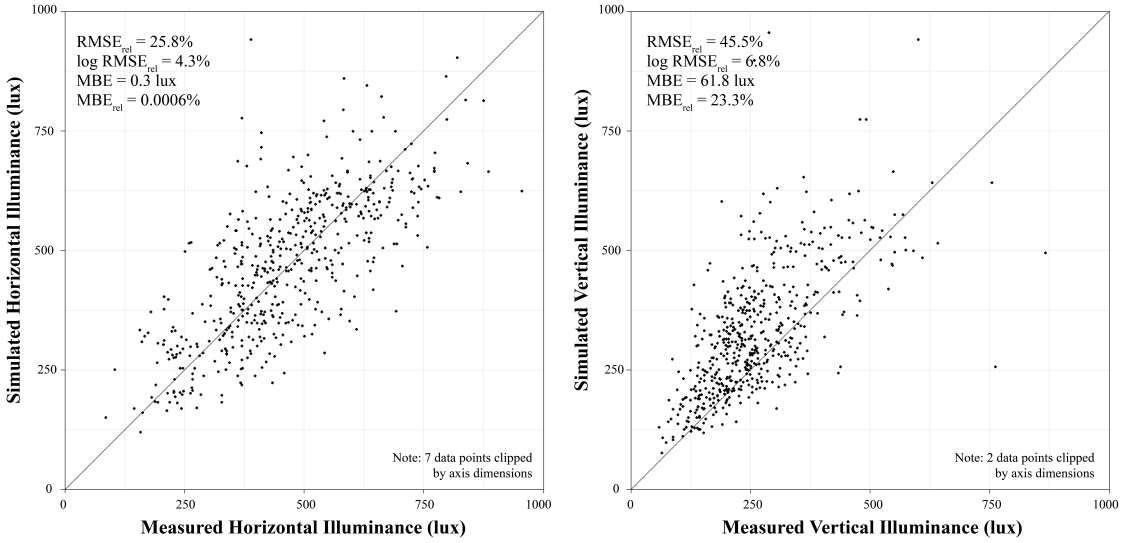


Figure 8. Scatter plots of measured vs simulated horizontal and vertical illuminance

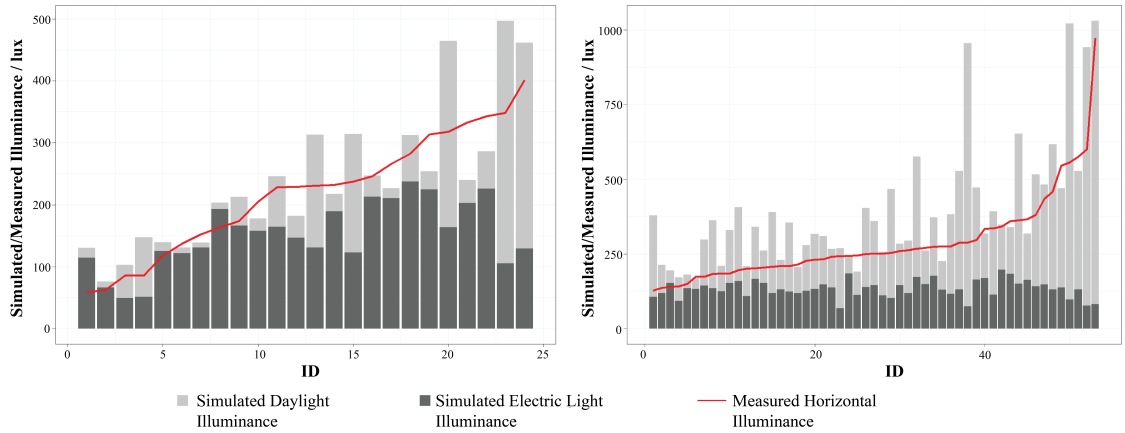


Figure 9. Simulated daylight and electric light illuminances of Office 7 and 10 as stacked column plots, and field-measured horizontal illuminance as a red line

those simulated during the calibration process. Differences specifically in the monitor brightnesses can be observed based on discrete user factors as mentioned as a cause of high vertical errors in the previous paragraph. Figure 11 compares the accuracy of the simulation models per office space (as shown in Figure 7) by their horizontal (yellow) and vertical (blue) $RMSE_{rel}$ values. Office 4 is an example of a well-calibrated model with both horizontal and vertical $RMSE_{rel}$ values of approximately 25%. Figure 9 depicts individual horizontal illuminance simulation results (stacked bars) compared to measured values (red lines) for buildings 7 and 10. Results are split between the electric lighting and daylighting simulation contributions. Simulated illuminances follow the overall trend of field-measured horizontal illuminance, and daylighting components of the calculated illuminance tend to be more volatile than electric lighting calculations as might be expected considering desired even workplane illuminances from electric lighting in most office spaces.

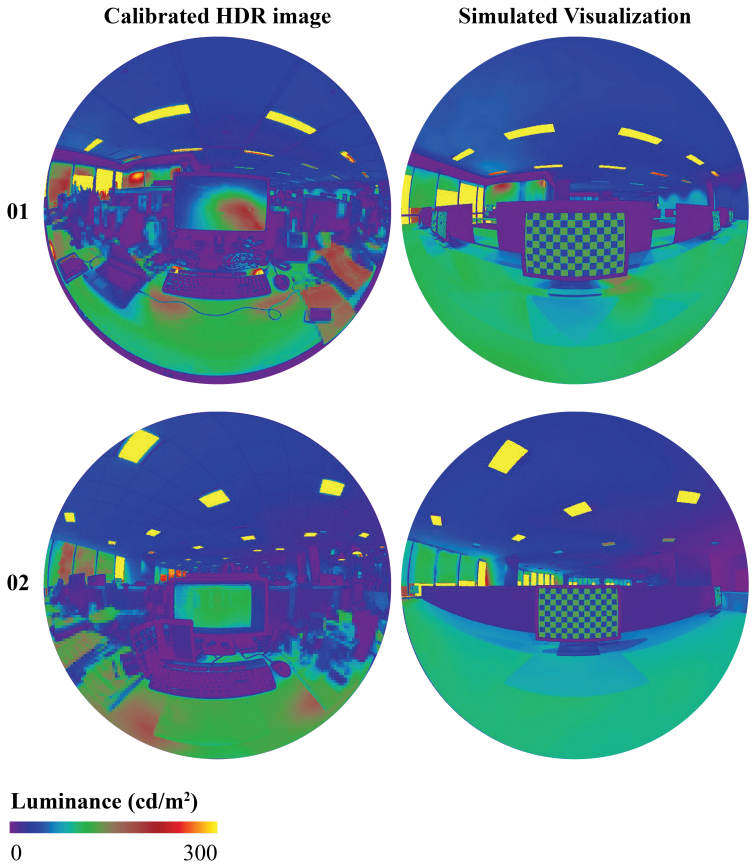


Figure 10. Comparison of calibrated HDR images and simulated visualizations from the same viewpoint

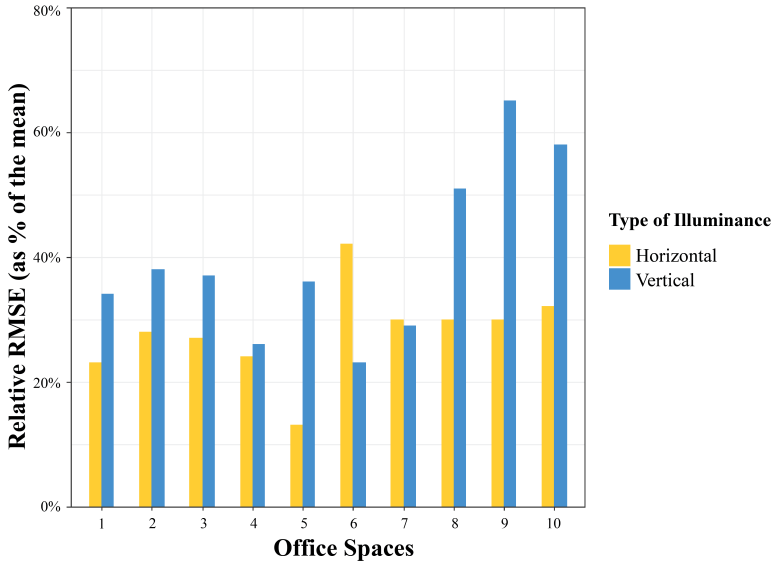


Figure 11. Horizontal and Vertical $RMSE_{rel}$ values between measured and simulated point-in-time values

4 Discussion

The authors believe that sharing their experience with the field measurement and calibration process of the 10 models in this paper would be beneficial to the reader. To achieve more accurate simulation results from the lighting models, the authors suggest to identify problematic data points—illuminance calculations at participant desks with large discrete errors—and resolving modeling issues for those desks by comparing their HDR images to simulated model views using measured solar data and a Perez sky model (Perez et al., 1993). Large discrete errors found in this study during the early calibration process were mostly attributed to modeling errors that influenced daylighting and electric lighting issues, or both. For electric lighting discrepancies, typical corrections include identifying missing luminaire fixtures, making sure they were not coplanar with ceiling surfaces, and adjusting their luminous power (see Figure 6). Calibration of electric lighting may prove difficult but is important to capture in simulations to describe the occupants' experience. The proposed method can also be used by measuring the building with electric lights turned off; however, that was not an option in this study. On the other hand, for daylighting discrepancies, adjustments of roller blind position and missing roller blinds were common modelling errors. In some cases, the transmission and angular properties of roller blinds had to be adjusted based on measured values. Roller blind statuses may be uncertain throughout the year, but the authors' focus is validating the instantaneous condition by simulating the horizontal and vertical illuminance values, at the same time-step as measurement, to be validated against actual measurement values. In addition, reproducing actual user behavior is impossible without long-term monitoring, so a first possibility is to apply the methodology to a design-type model without active shades. Secondly, some researchers have also suggested using blind control algorithms, and this may also be applied to the CBDM daylighting models with ease, if desired. For example, Gunay et al. have suggested outputting ranges of results based on different plausible models (Gunay et al., 2017) while others are mostly prescriptive. (Nezamdoost et al., 2014; Reinhart, 2004). Another third method is to monitor the blind usage with a drawstring potentiometer to capture actual blind behavior over a period of time (McNeil, 2018). Future work on field measurement of roller blinds would be beneficial, as diffuse and specular transmission data was rarely available in the authors experience and is difficult to measure in the field. There were often difficulties in discerning the sources of predictive errors between simulation and reality due to the presence of both daylight and electric lighting in the HDR images used to compare with the simulated visualizations. In most cases, building-level biases (increased horizontal illuminance levels at each desk) were able to identify when the electric lighting was the cause of a significant error.

In addition, field calibrations of lighting models have many difficulties in practice that have not been found in other calibration studies which are often based upon unoccupied spaces. The issues arising from monitor variety, screen image, user-specific brightness settings, and the choice to turn a monitor off before a measurement have already been described. Initially, the authors noticed a higher $MBE_{overall}$ of around 100 lux for vertical illuminance ultimately caused by monitor screens. Based on the HDR measurements, the authors attributed the main cause to the effective peak luminance of monitor screens being too high at $250 cd/m^2$, and the peak luminance was reduced to $125 cd/m^2$ which reduced the relative and bias errors. Each individual participant also has a variety of desk layout differences and personal or work items or clutter present that will influence the lighting results. Observation of Figure 3 will illustrate placement differences in keyboards and monitors as well as a wealth of objects which are impractical to model individually and likely change from day-to-day. These are real limitations that must be accepted as error sources when field-calibrating lighting models. Other error sources which are difficult to surmount are the difficulty in selecting IES datasets with accurate photometric distributions, unknown maintenance and

cleaning cycles, a variety of lamps with varied lumen outputs installed throughout buildings, and proxy geometry of the simulated electric lights not reflecting the direct view of the light source from the actual luminaire geometry in reality even though the overall photometric distribution is accurate.

The proposed field calibration method presented in this paper has several limitations. It is sometimes difficult to obtain information of actual glazing transmittances for fixed glazing, since they cannot be measured easily in the field. Modeling directional and diffusing properties of complex glazing and roller blinds also presents challenges where no detailed goniophotometer measurements exist. Where possible, the authors have tried to obtain accurate information from the individual building developers to increase the accuracy of the models. In addition, the location of the weather station may also cause discrepancies in the custom sky models used to simulate the instantaneous lighting data used in calibrations due to global horizontal irradiance being recorded from a different location on the island up to 25 km away in the case of this study. In the future, portable irradiance data loggers should be utilized at least during the field measurement period. Although the authors used the Reindl, et al. splitting model (Reindl et al., 1990) to derive direct and diffuse components from the global horizontal irradiance, there exists the possibility to derive diffuse fractions from global measures, time, and location. Based on previous research (Chandrasekaran and Kumar, 1994), researchers compared measured diffuse fractions and a computed model for a tropical location to the one preferred by Reindl, et al. It was found that clear sky conditions under humid tropical conditions had a higher diffuse fraction than the model of Reindl, et al., but that the relative standard deviation only improves from 32.2% to 29.4% by using a tropically-derived splitting model. Although the International Weather for Energy Calculation (IWEC) weather data is also mostly derived from global horizontal irradiance methods using splitting algorithms and sometimes observed cloud cover, although further improvements can be achieved with seasonally-specific splitting models.

Six participant data points were removed during data analysis due to horizontal and vertical illuminance errors of more than 1,000 lx, which may be attributed to weather data inaccuracies likely due to disparate cloud cover pattens at specific times. These discrepancies are typical and expected of field measurements, unlike precise laboratory measurements that can be controlled to a higher level of accuracy. Dynamic shading devices are included in the calibration models, but their operation will obviously vary across the year. The authors suggest simulating a calibrated model without dynamic shades represents the base design case, which is useful to compare with subjective participant opinions. Behavioral and automated dynamic shading models may also be applied to resultant annual simulations.

5 Conclusion

The proposed process to build calibrated climate-based models for POE's based on one-time field measurements at each building is validated through comparing measured and simulated illuminance data, and the results are sufficiently positive with $logRMSE_{rel}$ values of 4.3% and 6.8% and $RMSE_{rel}$ values of 25.8% and 45.5% for horizontal and vertical illuminances. While 20% is often seen as a best-case validation result in the lighting community, the authors suggest that the additional horizontal error observed in this study is acceptable given the complex field conditions and the fact that the errors only make up a small *logarithmic* perceived difference in lighting. Vertically with regards to occupant monitors, errors begin to grow unacceptably large, and it may be beneficial to specifically measure monitor screen brightnesses or to turn them off in future field measurement and calibration studies. As occupied spaces are usually of limited access to researchers, the non-invasive method presented here can allow reliable annual lighting information to be used for POE's through short visits instead of relying on long-term and invasive direct

monitoring data using illuminance sensors or HDR photographs. As the proposed workflow does not require constant monitoring and is non-invasive, this could also increase willingness of participants to engage in lighting POE studies.

The authors find that the proposed workflow of building calibrated models of existing spaces opens up possibilities of assessing long-term quantitative lighting results after a short visit for measurements which can be analyzed in correlation to subjective occupant responses that are collected in the same visit. This provides a reasonably reliable set of instantaneous and long-term lighting results and subjective responses from occupants for in-depth analysis, of which initial results have been reported elsewhere (Jakubiec et al., 2018). The authors acknowledge that the proposed methodology may be more complex than a simpler point-in-time measurement (which is also included in the proposed process as well), but building calibrated CBDM daylighting models enables researchers to simulate annual daylighting metrics based on measured annual weather data. At the same time it is significantly cheaper than long-term monitoring techniques due to less demands of manpower and equipment as measurements can be collected from multiple users using a single set of instruments during a short visit. Predicted lighting results from the design or construction phase of the building may also be analyzed with the post-occupancy results later on, for an additional feedback loop to architects and lighting designers and policy makers for lighting requirements in sustainable building assessment criteria.

Although the accuracy of field-calibrated models is not expected to surpass that of those in controlled laboratory setups, it is still a sufficiently reliable method to determine and detect lighting variations in a POE study. Problematic data points with large discrete errors have to be identified and rectified while calibrating the simulation models, as well as correcting for global errors due to electric lighting. In addition, some limitations in practice such as individual user differences in monitor screens, items in the workspace, and shade use must be accepted. The authors believe that the validation process presented in this manuscript is important to accurately drive POE results based on calibrated annual lighting simulations by nearly eliminating bias errors and minimizing relative errors.

6 Funding and Acknowledgements

This research is supported by the National Research Foundation, Prime Minister's Office of Singapore under its Research, Innovation & Enterprise (RIE) Funding's Green Buildings Innovation Cluster (GBIC) program through grant number GBIC-R&D / DCP 05. The authors further acknowledge support for measurement equipment from the SUTD-MIT International Design Centre (IDC). Any opinions expressed herein are those of the authors and do not reflect the views of the IDC. The authors are extremely grateful to Denise Kwok from the BCA's Green Mark Department for facilitating access to many of the buildings studied herein. The authors additionally thank the 10 building or office managers who granted us access to their buildings. Finally, the authors appreciate the comments of the two anonymous reviewers in strengthening the content and presentation of the paper.

7 Disclosure Statement

The authors have no financial interests to declare.

References

Bear, A. R. and R. I. Bell
1992. The csp index: A practical measure of office lighting quality as perceived by the office worker. *Lighting Research & Technology*, 24(4):215–225.

Bellia, L., F. Fragliasso, and E. Stefanizzi
2017. Daylit offices: A comparison between measured parameters assessing light quality and users' opinions. *Building and Environment*, 113:92–106.

Brembilla, E., J. Mardaljevic, and C. J. Hopfe
2015. Sensitivity analysis studying the impact of reflectance values assigned in climate-based daylight modelling. In *Proceedings of Building Simulation*.

Chandrasekaran, J. and S. Kumar
1994. Hourly diffuse fraction correlation at a tropical location. *Solar Energy*, 53(6):505–510.

Choi, J. H., V. Loftness, and A. Aziz
2012. Post-occupancy evaluation of 20 office buildings as basis for future ieq standards and guidelines. *Energy and Buildings*, 46:167–175.

Committee, I. C. et al.
1991. Ies standard file format for electronic transfer of photometric data and related information. Technical report, Technical report IES LM-63-1991, New York.

Dahlan, N. D., P. J. Jones, D. K. Alexander, E. Salleh, and J. Alias
2009. Daylight ratio, luminance, and visual comfort assessments in typical malaysian hostels. *Indoor and Built Environment*, 18(4):319–335.

Drosou, N., E. Brembilla, J. Mardaljevic, and V. Haines
2016. Reality bites: measuring actual daylighting performance in classrooms. In *Proceedings of PLEA*.

Fan, D., B. Painter, and J. Mardaljevic
2009. A data collection method for long-term field studies of visual comfort in real-world daylit office environments. In *Proceedings of PLEA*, Pp. 251–256.

Gradillas, M.
2015. Material characterization, presentation, diva day 2015.

Gunay, H. B., W. O'Brien, I. Beausoleil-Morrison, and S. Gilani
2017. Development and implementation of an adaptive lighting and blinds control algorithm. *Building and Environment*, 113:185–199.

Hirning, M. B., G. L. Isoardi, and I. Cowling
2014. Discomfort glare in open plan green buildings. *Energy and Buildings*, 70:427–440.

Hirning, M. B., G. L. Isoardi, S. Coyne, V. R. Garcia-Hansen, and I. Cowling
2013. Post occupancy evaluations relating to discomfort glare: A study of green buildings in brisbane. *Building and Environment*, 59:349–357.

Hirning, M. B., G. L. Isoardi, and V. R. Garcia-Hansen
2017. Prediction of discomfort glare from windows under tropical skies. *Building and Environment*, 113:107–120.

Inanici, M. N.
2006. Evaluation of high dynamic range photography as a luminance data acquisition system. *Lighting Research & Technology*, 38(2):123–134.

Jacobs, A.
2012. Radiance cookbook.

Jakubiec, J. A.

2016. Building a database of opaque materials for lighting simulation. In *Proceedings of PLEA*.

Jakubiec, J. A., G. Quek, and T. Srisamranrungruang

2018. Towards subjectivity in annual climate-based daylight metrics. In *Proceedings of Building Simulation*.

Jakubiec, J. A. and C. F. Reinhart

2011. Diva 2.0: Integrating daylight and thermal simulations using rhinoceros 3d, daysim and energyplus. In *Proceedings of Building Simulation*.

Jakubiec, J. A. and C. F. Reinhart

2016. A concept for predicting occupants' long-term visual comfort within daylit spaces. *Leukos*, 12(4):185–202.

Jakubiec, J. A., K. G. Van Den Wymelenberg, M. N. Inanici, and A. Mahic

2016. Accurate measurement of daylit interior scenes using high dynamic range photography. In *Proceedings of CIE Lighting Quality and Energy Efficiency*.

Jones, N. L. and C. F. Reinhart

2017. Experimental validation of ray tracing as a means of image-based visual discomfort prediction. *Building and Environment*, 113:131–150.

Konis, K.

2013. Evaluating daylighting effectiveness and occupant visual comfort in a side-lit open-plan office building in san francisco, california. *Building and Environment*, 59:662–677.

Konis, K.

2014. Predicting visual comfort in side-lit open-plan core zones: results of a field study pairing high dynamic range images with subjective responses. *Energy and Buildings*, 77:67–79.

Mangkuto, R. A., K. A. Kurnia, D. N. Azizah, R. T. Atmodipoero, and F. N. Soelami

2017. Determination of discomfort glare criteria for daylit space in indonesia. *Solar Energy*, 149:151–163.

Mardaljevic, J., E. Brembilla, and N. Drosou

2015. Illuminance-proxy high dynamic range imaging: a simple method to measure surface reflectance. In *Proceedings of the 28th Session of the International Commission on Illumination*.

Mardaljevic, J., E. Brembilla, and N. Drosou

2016. Real-world validation of climate-based daylight metrics: mission impossible? In *Proceedings of CIBSE Technical Symposium 2016*.

McNeel, R. and Associates

2017. Rhinoceros 3d version 5.0.

McNeil, A.

2018. Presentation.

Nezamdoost, A., A. Mahic, and K. Van Den Wymelenberg

2014. Annual energy and daylight impacts of three manual blind control algorithms. In *2014 IES Annual Conference Proceedings. Pittsburg, PA*.

Oseland, N.

2007. *British Council for Offices guide to post-occupancy evaluation*. British Council for Offices.

Osterhaus, W.

2001. Discomfort glare from daylight in computer offices: how much do we really know. *Proceedings of LUX Europa*, Pp. 448–456.

Painter, B., J. Mardaljevic, and D. Fan

2010. Monitoring daylight provision and glare perception in office environments. In *Proc CIB World Congress*.

Parpairi, K., N. V. Baker, K. A. Steemers, and R. Compagnon
2002. The luminance differences index: a new indicator of user preferences in daylit spaces. *Lighting Research & Technology*, 34(1):53–66.
Perez, R., R. Seals, and J. Michalsky
1993. All-weather model for sky luminance distribution—preliminary configuration and validation. *Solar energy*, 50(3):235–245.
Reindl, D. T., W. A. Beckman, and J. A. Duffie
1990. Diffuse fraction correlations. *Solar energy*, 45(1):1–7.
Reinhart, C. and A. Fitz
2006. Findings from a survey on the current use of daylight simulations in building design. *Energy and Buildings*, 38(7):824–835.
Reinhart, C. F.
2004. Lightswitch-2002: a model for manual and automated control of electric lighting and blinds. *Solar energy*, 77(1):15–28.
Reinhart, C. F. and M. Andersen
2006. Development and validation of a radiance model for a translucent panel. *Energy and Buildings*, 38(7):890 – 904.
Reinhart, C. F., J. Mardaljevic, and Z. Rogers
2006. Dynamic daylight performance metrics for sustainable building design. *Leukos*, 3(1):7–31.
Reinhart, C. F., T. Rakha, and D. Weissman
2014. Predicting the daylit area—a comparison of students assessments and simulations at eleven schools of architecture. *Leukos*, 10(4):193–206.
Reinhart, C. F. and O. Walkenhorst
2001. Validation of dynamic radiance-based daylight simulations for a test office with external blinds. *Energy and Buildings*, 33(7):683–697.
Ward, G.
1994. The radiance lighting simulation and rendering system. In *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*, Pp. 459–472.
Ward, G.
2005. Photosphere, anyhere software.
Wienold, J.
2004. Evalglare–a new radiance-based tool to evaluate daylight glare in office spaces. In *3rd International Radiance Workshop*.