Computing Long-term Daylighting Simulations from High Dynamic Range Imagery Using Deep Neural Networks

Yue Liu

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Mehlika Inanici, Chair

Alex Colburn

Robert Mugerauer

Program Authorized to Offer Degree:
College of Built Environments

University of Washington

Abstract

Computing Long-term Daylighting Simulations from High Dynamic Range Imagery Using Deep Neural Networks

Yue Liu

Chair of the Supervisory Committee:

Associate Professor, Dr. Mehlika Inanici
Department of Architecture

Practitioners and researchers utilize long-term daylighting analysis to evaluate the luminous environment under a wide range of naturally occurring sky and sun conditions. Annual simulation tools and metrics are commonly based on illuminance calculations, which focus on daylight availability and potential energy savings. Design and research practices are shifting from illuminance-based metrics towards luminance-based (human-centric) metrics and evaluations. However, hourly annual luminance maps require either labor-intensive and computationally expensive simulation processes, or impracticable long-term in-situ measurements. Simulations are performed by rendering the environment using physically-based rendering techniques. Annual luminance-maps are generated by solving the radiance regression

function, which is a non-linear mapping of pixel-scale luminance values from local and contextual attributes of surface points for every hour of daylight (approximately 4,000 simulations). This process can take weeks to compute with complex building geometries.

This dissertation presents a novel data driven machine learning approach to make annual luminance-based evaluations more efficient and accessible. The methodology is based on predicting annual luminance maps from a limited number of point-in-time high dynamic range (HDR) imagery by utilizing a deep neural network (DNN).

An initial study used a small dataset to train a DNN model that accurately predicts annual luminance maps. The training dataset utilizes 200 hourly samples from the annual luminance map dataset, approximately 5% of the entire dataset. A second study refines the methodology of the first study and applies it to panoramic luminance maps (with 360° horizontal and 180° vertical field of view). Panoramic views are more data intensive compared to perspective views that have a much narrower field of view, but they provide better information about an occupant's visual experience. A thorough sensitivity analysis is performed to develop guidelines for determining the minimum and optimum data collection periods, and the minimum number of imagery needed for accurately generating the annual luminance maps. The analyses suggest that the proposed DNN model can faithfully predict high-quality panoramic luminance maps from one-month hourly imagery generated or collected during daylight hours around the equinoxes. Alternatively, 9 days of hourly data collected around the spring equinox, summer solstice, and winter solstice would suffice (2.5% of total annual renderings) to predict the point-in-time luminance maps for the rest of the year. Data generation is greatly accelerated with this method, and only requires 30 minutes of training time on a NVIDIA GeForce ^{1080Ti} GPU and a few

seconds of computing time. Therefore, this approach alleviates the problem of extensive calculation time that typically hampers the utilization of annual luminance-based daylighting simulations and metrics.

With the developed workflow, practitioners and researchers can efficiently incorporate long-term luminance-based metrics into the design and research process. This study also lays the groundwork for generating annual luminance maps utilizing short-term HDR photographs of existing environments, which will enable quantitative analysis of daylit environments without requiring a time-consuming modeling process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor, Professor Mehlika Inanici, for her constant trust, patient guidance, and enthusiastic encouragement of this research work. During these six years, she never hesitated to share her knowledge and wisdom in research, teaching, and writing. She is the one who encouraged me to pursue the research direction of long-term daylighting accelerations. Although some parts of the research have gone beyond the scope of traditional architectural techniques, and beyond her expertise, she still provides a lot of support and understanding and has been a source of inspiration and challenges through our discussions.

I would also like to extend my gratitude to my thesis committee. Dr. Alex Colburn has contributed a great deal to my dissertation research, much more than what can be expected from a traditional committee member, especially on the machine learning issues. He has helped and challenged me to push my research forward. His invaluable advice and guidance were indispensable to this dissertation. I am also grateful to Professor Brian Curless for opening my eyes to connecting daylighting and computer graphics, for his technical insight and guidance, and for helping to connect me professionally. Professor Marc Schiler has been a trusted mentor since 2011 when I began pursuing my Master's degree. Marc is the one who inspired my interest in architectural lighting and encouraged me to continue my path to research in building technology and daylighting. I thank him for his endless trust and support for my success. I would like to thank Professor Robert Mugerauer, who has led the Ph.D. in Built Environment program

for many years. He never stopped from caring and supporting every one of us in our Ph.D. program.

In addition to those already mentioned, I am grateful to all my previous mentors during studies and internships, Professor Karen Kensek, Professor Douglas Noble, Professor Kyle Konis, and Professor Joon-ho Choi from USC; Seth Ely from Stantec; and Eleanor Lee from LBNL. They have demonstrated what brilliant and hard-working researchers and professionals can accomplish. I would like to thank the broader daylighting research community for many inspiring discussions over the years, especially Greg Ward, for his incredible willingness to answer questions, and every member of the RADIANCE Online Group.

I would also like to express my gratitude to UW College of Built Environments for providing me a year-long fellowship and the Illuminating Engineering Society of North America for awarding me the Richard Kelly Grant and the Robert E. Thunen Memorial Scholarship. Part of the computing resources used in this study is through Microsoft Azure Research Award.

During the long journey of Ph.D. study, I have had the fortune to make many good friends. I thank my UW Ph.D. peers, Huazeng, Jiawen, Mia, Naeun, Xuan, and Yanjun who shared both challenges and joys of the doctoral studies, cheered me on during difficult times, and celebrated each accomplishment. Thanks also go to Nat, Ruigang, Yao, and Zhang for their encouragement and constructive suggestions.

Last but not least, I am deeply grateful to my family for their infinite support and love. I am grateful to my parents, Hongcheng Liu and Xingmei Huang, for giving me the opportunities and experiences that have made me who I am. They selflessly encouraged me to seek my own destiny and have always been generous with their encouragement despite the long distance

between us. My love Lei has been my best friend and great companion, supported, encouraged, entertained, and helped me get through the long journey of Ph.D. in the most positive way.

DEDICATION

To my love, Lei Fu

Chapter 1. INTRODUCTION

1.1 BACKGROUND

Strategic use of daylight in buildings can reduce the energy demands and improve occupants' comfort and satisfaction. Daylight is the most sustainable source of illumination for building interiors. According to the most recent Commercial Buildings Energy Consumption Survey, electrical lighting takes the most significant portion (17%) of electricity consumption in commercial buildings (U.S. Energy Information Administration 2012). Use of daylight in buildings can significantly reduce the electrical energy consumption. Adequate daylight controls have the potential to reduce cooling loads by limiting the solar radiation penetrating into building interiors, and decrease the peak energy demand. Therefore, daylight utilization is an important part of green and sustainable building design, and it is an essential step towards net-zero energy buildings. Along with energy saving potentials, the presence of daylight improves occupants' health, alertness, productivity, and feelings of well-being (Boyce 2014). Light is the primary time cue for synchronizing our internal circadian body (~24h) clock to local time. Several optical radiation characteristics, such as quantity, spectrum, timing, duration, pattern, and prior optical radiation exposure, affect circadian rhythms (IES 2008). Proper daylighting designs can have a positive impact on occupant's circadian rhythms, visual comfort, and productivity.

However, there is no guarantee that daylight in buildings will always save energy and improve occupants' comfort. Uncontrollably maximizing daylight penetration into buildings can lead to undesirable luminous environments that reduce the visibility of the visual task, or create visual discomfort. These conditions have the potential to hamper visual productivity and cause undesirable physical symptoms, such as eye fatigue and headaches (IES Handbook 2011).

Nowadays, people spend 90% of their time indoors (Klepeis et al. 2001). The luminous environment should be carefully designed and evaluated to utilize and control daylight under dynamic conditions.

Meaningful daylighting evaluations require long-term analysis to have a comprehensive understanding of the luminous environment under a wide range of naturally occurring sky and sun conditions. Ideally, daylighting evaluation period should be annual to capture the seasonal variations of daylight, and the time step should be small enough (e.g., 1 hour) to capture the observed short-term daylight variations. Design and research practices of daylighting are shifting from illuminance-based metrics towards human-centric luminance-based evaluations. Luminance maps capture the subject's field of view and provide high resolution data that can be used to evaluate the luminance variations, contrast, positions and intensity of light sources, and adaptation luminances. Luminance distribution patterns determine the visual comfort (e.g., discomfort glare), visual appearance and task visibility in a given scene (Wienold and Christoffersen 2006, Jakubiec and Reinhart 2012, Suk et al. 2013, Konis 2014, van den Wymelenberg and Inanici 2015, Rockcastle et al. 2016). However, annual luminance-based simulations require generating a luminance map at each time step of the entire year (approximately 3500-4000 daylight hours), resulting in significant simulation time. Illuminance-based annual simulations are accessible to lighting professionals through a number of programs and metrics. Luminance-based annual simulations are computed through multi-phased daylight coefficient methodologies (Ward et al. 2011, McNeil 2013a, McNeil 2013b), which have steep learning curves and long simulation time.

From a practical perspective, existing luminous environments can be easily captured using HDR photographic techniques without modeling the light source, material, and geometry information of the scene (Debevec 1997). Studies have shown that HDR photographic techniques

are able to capture the luminous environment with reasonable precision and repeatability (Inanici 2006). Compared with traditional simulation methods, HDR photographic techniques open up new possibilities in lighting research and design by recording the scene luminance distribution with low-cost and high efficiency. However, HDR photographs only capture the moment. A single HDR photograph can only provide lighting information about the luminous environment for a specific sun position and sky condition. The limited lighting information is not informative about the dynamic nature of the luminous environment over the entire year. For example, an HDR capture of the scene under the overcast sky cannot represent the luminous environment of the same scene under sunny skies. Long-term daylighting evaluations using HDR photographs require long-term field studies, which is not feasible due to time concerns and other restrictions.

1.2 RESEARCH OBJECTIVES

Recognizing the need to improve the availability and accessibility of annual luminance-based lighting evaluations, this study develops a research framework to accelerate annual luminance-based simulations by generating annual perspective and panoramic luminance maps of indoor spaces from a small subset of HDR images. The research is based on deep learning techniques, which have been used widely in various research and industrial fields, but are rarely used in architectural lighting studies. The deep learning methodology is utilized to compute the radiance regression function, which is a non-linear mapping from local and contextual attributes of surface points (such as position, viewing direction, and lighting condition) to their illumination values. Architects, lighting designers, and researchers can efficiently use the workflow developed in this study to perform annual luminance-based evaluations over multiple view directions. This study also lays the foundation for generating annual luminance maps utilizing HDR captures of existing environments. The specific objectives of the study are:

- To create a deep learning framework to conduct architectural luminance predictions;
- To develop a robust workflow that accurately and efficiently predicts annual panoramic luminance maps from a small number of HDR images;
- To analyze the sensitivity of prediction accuracy to sample data size, which informs both the optimum and minimum data collection period and the minimum sample size; and
- To study how the weather and seasonal changes affect the accuracy of the outcome given a fixed data collection period.

The goal is to enhance the integration of annual luminance-based evaluations into the architectural design process by making the simulations more efficient and more accessible, without compromising accuracy. The availability of feasible annual luminance data will enable the design of better luminous environments that support human-centric metrics such as occupant visual comfort, productivity, and well-being.

1.3 Overview of the Dissertation

Chapter 2 discusses per-pixel lighting simulation (physically-based rendering) and measurement (HDR photography) techniques along with illuminance- and luminance-based evaluation metrics within the domain of architectural daylighting analysis. Best-practices in accelerating daylighting simulations are described; and the need for an efficient and accessible annual luminance-based daylighting evaluation method is demonstrated.

Chapter 3 describes the context of the machine learning techniques employed in the dissertation, and reviews the lighting-related machine learning studies.

Chapter 4 introduces the methodology for predicting annual luminance maps of perspective views from a subset of HDR images. The machine learning workflow is introduced in detail.

The methodology developed in Chapter 4 is refined in Chapter 5 to increase the model robustness and accuracy. The improved method extends the image format to panoramic luminance maps, which provide the ability to analyze head and gaze movements from a fixed point.

Chapter 6 incorporates a sensitivity analysis performed to improve the efficiency and applicability of the method. The minimum and optimum data collection period for generating annual luminance maps are explored. The impact of weather and other seasonal changes on the accuracy of the results are investigated. The analysis results provide guidelines for selecting training samples and data collection periods.

Finally, Chapter 7 summarizes the major contributions of the dissertation, and concludes with a discussion on future work.

Chapter 2. ARCHITECTURAL DAYLIGHTING SIMULATION AND MEASUREMENT

Chapter 2 provides the context for the state-of-the-art research and practice in daylight simulations, measurements, and analyses. Sections 2.1 and 2.2 focus on physically-based daylighting simulations and HDR photography technique as the two methods that can be used to generate and measure the luminance variations in daylit environments. Section 2.3 discusses the daylighting evaluation metrics used for daylight performance assessment. A better understanding of these metrics highlights the importance of luminance-based evaluations and demonstrates the need for accelerating the computation of long-term luminance-based simulations. Section 2.4 reviews the current best practices in accelerating annual daylighting evaluations.

2.1 PHYSICALLY-BASED DAYLIGHTING SIMULATIONS

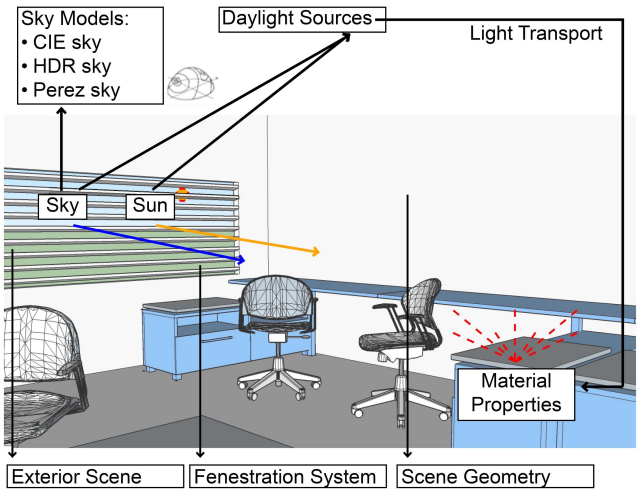


Figure 1. Physically-based daylighting simulations calculate the light transport of a given scene, based on geometry, material properties and light sources.

Physically-based daylighting simulation is a specific subset of typical architectural renderings. While most renderings aim to express the design concept with varying degrees of photorealism or abstraction, physically-based daylighting simulations provide quantifiable representations of luminous environments with reasonable accuracy and precision; and they are utilized by design and research teams to evaluate and compare different daylighting schemes. In physically-based daylight simulations, daylight levels are predicted by calculating the light transport in a scene, based on measurements and assumptions about the physical world (Figure 1). To perform such simulations, designers and researchers need to establish a scene model with physically-accurate geometry, light sources and material definitions as input, and calculate the light transport in the scene through validated global illumination algorithms. The daylighting parameters and global illumination algorithms will be further discussed in the following sections.

2.1.1 *Daylight Sources: Sun and Sky*

Although all daylight originates from the sun, the sun and sky are considered as two distinct sources in physically-based daylighting simulations. The sun is modeled as a directional light source with 0.5° steradian solid angle. There is a clear pattern of the daily and seasonal movements of the sun at a particular geographical location. Therefore, site longitude and latitude, date and time, can determine where the sun is located for a given site at a particular date and time. In simulations, the position of the sun is defined with its solar altitude and azimuth angles. The intensity of the solar disc is defined with its luminance (RGB) values, which corresponds to the direct normal irradiances divided by the disc's solid angle.

Modeling the sky is more complicated than simulating the sun. The diffuse portion of the sky exhibits distinct characteristics due to weather, atmospheric humidity, and turbidity (IES Lighting Handbook: Reference and Application 2011). In physically-based lighting simulations, the

geometry of the sky is defined as an invisible hemisphere. The sky luminance distribution is modeled through theoretical sky models, which are generalized mathematical functions derived from long-term sky measurements. These sky models describe the amount of light coming from each part of the celestial hemisphere at a particular time. The oldest sky model is the uniform sky model, which assumes that there is no direct sunlight and the sky has the uniform luminance. This model is not representative of any naturally occurring sky conditions, therefore, it is no longer used. Later, two commonly available types of sky models were suggested by CIE, which are the CIE standard Overcast Sky and the CIE standard Clear Sky (Figure 2). The CIE standard Overcast Sky (CIE 1970) is defined with a 3:1 gradation of sky luminance from the zenith to the horizon. The CIE standard Clear sky (CIE 1973), originally known as Kittler sky (Kittler 1967), describes the sky luminance distribution that varies with the solar position and across the sky dome, but only applies to the skies without any cloud cover. The CIE Intermediate Sky (Nakamura et al. 1985) fills in the blank between CIE Clear and CIE Overcast skies, and defines the luminance distribution of skies with thin or moderate cloud cover (Figure 2).

To cover the whole range of sky conditions from clear to overcast, Perez All Weather sky (Perez et al. 1993) as well as CIE Standard General sky (Kittler et al. 1998) models were developed. The CIE Standard General Sky defines fifteen discrete sky models with five overcast, five clear and five transitional skies. Perez All Weather Sky defines a continuous function to describe sky luminous distribution based on direct normal and diffuse horizontal radiation (W/m^2). Compared with previous sky models, Perez Sky has many advantages: 1) It uses a continuous function, and there is no need to select from discrete standard skies, which is required for the use of CIE General Skies. 2) It is capable of describing any sky condition, where the weather data is available. The EnergyPlus site provides the measured radiation data in hourly weather files for a

large number of locations around the world (US Department of Energy 2019). The radiation data in weather files could either be in the format of direct normal and diffuse horizontal radiation or as the global horizontal radiation, and the latter can be split into the former two components using various models such as the Reindl et al. model (1990). Perez Sky is widely used to simulate dynamic lighting performance. Perez skies take local climate into consideration, they can be generated from the weather files that are utilized in thermal simulations (Reinhart 2012), and they provide adequate accuracies (Inanici and Liu 2016, Inanici and Hashemloo 2017).

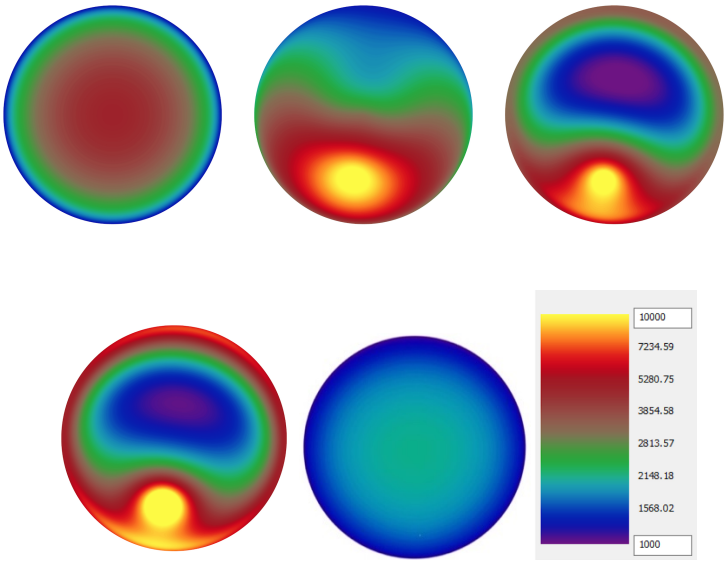


Figure 2. Three CIE Standard Sky types simulated In Radiance (Ward 1994) using Gensky on October 2nd at 11 am: (Upper Left) CIE Overcast Sky, (Upper Middle) CIE Intermediate Sky, and (Upper Right) CIE Clear Sky; and Two Perez Skies simulated in Radiance using Gendaylit: (Lower Left) Most sunny day October 2nd at noon with direct and diffuse irradiances of 922 and 59 W/m^2 and (Lower Right) Most cloudy day January 8th at noon with direct and diffuse irradiances of 2 and 46 W/m^2 in Seattle.

However, the actual sky luminous distributions are far more complex than theoretical sky models, and none of these theoretical sky models can represent the complex cloud compositions. Further accuracy can be achieved by image-based sky models. Image-based sky models, which are

captured HDR fisheye images of the skydome (Stumpfel et al. 2004), can be applied in simulation, replacing the original sun and sky model and lit the scene. These HDR sky models can capture instantaneous cloud conditions and surrounding environment (e.g., high-rise buildings and trees), but require a significant amount of labor to obtain and calibrate (Debevec 2002, 2005, Inanici 2010, Inanici and Liu 2016, Inanici and Hashemloo 2017).

2.1.2 *Material Modeling and Measurement*

Setting realistic optical material properties is an essential step towards building a physically-accurate daylight model. The material properties describe how light interacts with surfaces of the scene geometries. When light is incident on a surface, it can be reflected, transmitted, scattered, and absorbed. Among all the interactions, the reflection and transmission have a dominant impact on the luminous environment. In physically-based daylighting simulations, the reflection and transmission optical properties are described using mathematical models. For example, in Radiance (Ward 1994) simulation program, the "plastic" mathematical modifier is built to describe diffuse reflectance properties of opaque surfaces with RGB values ranging between zero and one. The specular reflectance properties are modeled with two additional parameters, specularity and roughness. Other material modifiers, such as "metal," "trans," and "glow" are also built with mathematical definitions, and complex materials are modeled by mixing material definitions in a modular format (Ward 1994).

In lighting simulations, accurate measurement and modeling of material optical properties is a challenge. Some measurements can be conducted in the field, and some require laboratory-based devices.

Luminance meters and Macbeth charts are simplified instruments used in the field to estimate the optical properties of diffuse materials. By calculating the ratio between the luminance of the

reference material with known reflectance and the luminance of the target material under identical lighting conditions, luminance meters are capable of estimating the percentage of reflectance of diffuse surfaces. Macbeth charts are usually used in a similar way. The photographs of a Macbeth chart and the target material are taken under constant lighting conditions. The diffuse color of the target surfaces is computationally estimated based on the difference of Macbeth chart under a reference light source and in the field. Hand-held spectro-reflectometers provide a higher level of precision by measuring light as a function of wavelengths and can be used to measure the diffuse reflection and a specular reflection at a given angle. To accurately measure the complicated spread reflectance, laboratory-based devices such as a goniophotometer is required. The goniophotometer captures the material's bidirectional scattering distribution function (BSDF), which is a four-dimensional function, and can be divided into two main components as bidirectional reflectance distribution (BRDF) and bidirectional transmittance distribution function (BTDF). The goniophotometer moves a light source across all incident illumination directions of the full hemisphere dome, and the radiance for each incident direction is captured.

2.1.3 *Light Transport*

Global illumination algorithms compute the complex interactions between geometry, materials, and light sources using the Kajiya rendering equation (Kajiya 1986). Radiosity (Goral et al. 1984) and ray-tracing (Whitted 1980) are two generic algorithms employed in architectural physically-based renderings to solve the rendering equation. Radiosity was originally developed to compute radiative heat transfer between surfaces based on form factors. It defines the ratio of radiative energy leaving a given surface to the energy directly arriving at a second surface. The principal idea of radiosity, which is luminous equilibrium, suggests that the direct and inter-reflected light reach a stable state within an enclosed space (Cohen and Wallace 1993). However, luminous

equilibrium is based on the assumption that each surface is perfectly diffuse, and does not support specular surface properties, which are indispensable in daylighting simulations. Due to the theoretical and practical limitations of the classic radiosity approach, backward ray-tracing is used to simulate these complex surface properties. In backward ray-tracing, rays are emitted from the point of interest and traced backward until they either hit a light source or another object. When a ray intersects with a surface other than a light source, new rays are spawned toward the light source direction, and toward the reflection or refraction direction. The process is recursively repeated until the maximum number of reflections has been reached. Backward ray-tracing is more efficient than forward ray-tracing because most of the rays emitted from the light source within a scene never reach the particular point of interest.

Monte Carlo backward raytracing is employed in simulation engines like Radiance (Ward 1994). Forward ray-tracing is an add-on to Radiance to simulate effects such as caustics. Radiance has been validated against physical measurements (Mardaljevic 2000a) and is widely used in physically-based lighting simulations and evaluations. One key simulation parameter of Radiance is the ambient bounces (ab), which defines the maximum number of diffuse recasts before a ray is discarded. Figure 3 shows a series of Radiance rendered luminance maps of an interior office scene under the clear sky with different ab values to illustrate the effect of the ab parameters. The luminance value of an exemplary pixel of the scene image is analyzed. When ab = 0, only the impact of direct sunlight is considered, and the sky can be viewed, but it has no impact on the luminous environment. The sun patches and sharp shadows are presented in the resulting luminance map, while the rest of the scene has 0 luminance. The calculated luminance of the exemplary pixel located at the position with no direct view of the sun is 0. An ab = 1 adds diffuse daylight (i.e. sky) into calculation along with the interreflections from direct sunlight. An ab = 2

includes the interreflections from scene surfaces. The reference pixel has notable increases in luminance values until ab values reach to 3 in this scene. While higher ab values are associated with more accurate simulation results, increasing the ab values also leads to noticeable increases in calculation time, as more rays are involved. The luminance values reach a plateau after 6-8 bounces. The ab value should be selected to efficiently approximate the luminous environment of the scene without largely compromising the accuracy. The complexity of the scene has an impact on the appropriate ab value.

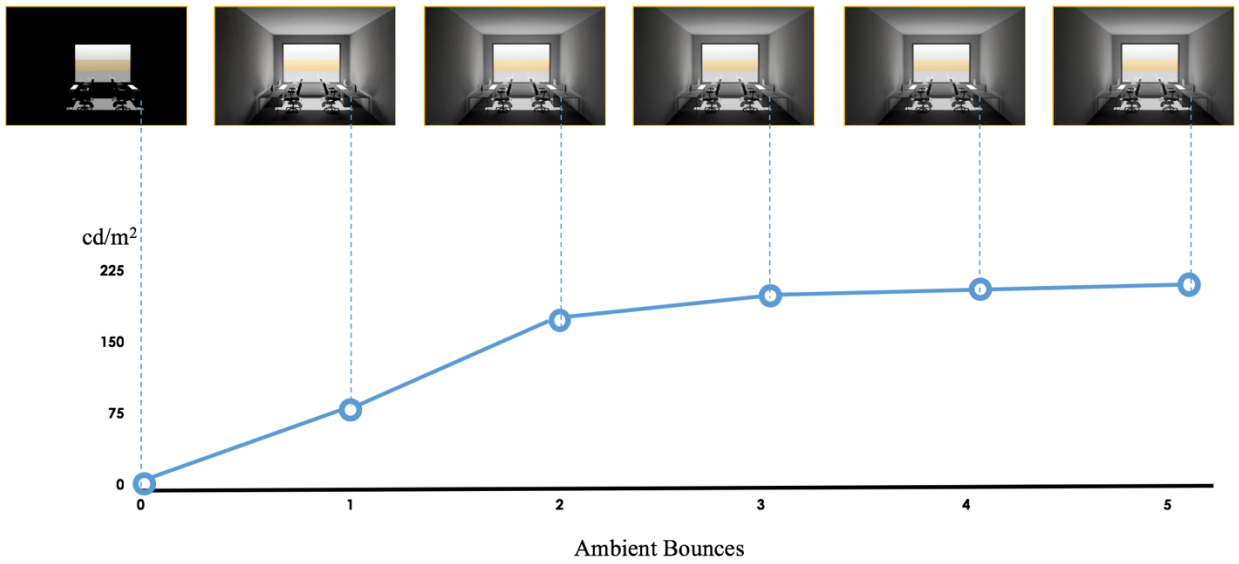


Figure 3. Ambient bounce parameter study of an exemplary pixel luminance value of an interior office scene.

2.2 HDR PHOTOGRAPHY FOR LUMINANCE-BASED MEASUREMENTS

Daylit scenes produce a much broader dynamic luminance range than a conventional digital camera device can capture with one exposure. The range from starlight to sunlight is 10^{-6} to 10^{9} cd/m^2. HDR photography technique (Debevec & Malik 1997) is developed to capture the wide luminance range through multiple Low Dynamic Range (LDR) photographs taken at different exposures. LDR photographs are computationally merged to a single HDR image. After post-

processing for lens and camera aberrations, each pixel value can correspond to the physical quantity of luminance. HDR photography technique is useful to efficiently collect high-resolution lighting data (i.e. luminance maps) from existing architectural spaces. Luminance maps provide much more information about the quantitative and qualitative dynamic of lighting across the scene than a limited number of luminance and illuminance measurements. The accuracy of the luminance maps has been validated against luminance measurements with less than 10% error, using appropriate calibrations and vignetting corrections (Inanici 2006). The resulting luminance maps can be further analyzed using the evaluation metrics described in section 2.3, for assessment of visual perception, preference, and comfort in daylit spaces (Wienold and Christoffersen 2006, Jakubiec and Reinhart 2012, Suk et al. 2013, Konis 2014, Van Den Wymelenberg and Inanici 2015, Bergen and Young 2018).

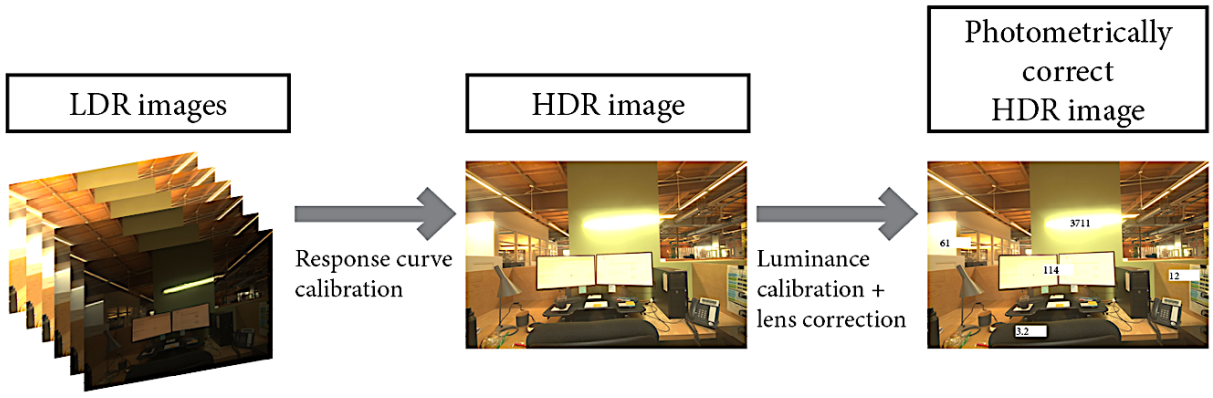


Figure 4. Creating photometrically correct HDR images from LDR images

Using HDR photography techniques for luminance measurements has three steps (Figure 4):

- Capturing a sequence of LDR photographs of the static scene with different exposures, where each exposure records a different luminance range;

- Computing a unique camera response function from the LDR images of the scene, and use the function to fuse the LDR images into one single HDR image with a specialized computer program such as Photosphere (Ward 2005); and
- Calibrating the HDR image with field luminance measurements and post-process the HDR image to correct for the vignetting effect (i.e., reduction of luminance at an image's periphery compared to the center) and other optical aberrations.

As a result, each pixel of an HDR image is associated with an absolute luminance (RGB) value in cd/m^2.

To get the finest results, users should follow the best practices for HDR image capturing (Inanici 2006, Jakubiec et al. 2016a, 2016b) to ensure accurate and consistent data:

- Set the white balance to daylight and the ISO sensor sensitivity to 100;
- Keep the aperture size constant to fix the depth of field and eliminate vignetting variance, only vary the shutter speed to get different exposure values (EVs). The aperture size is suggested to be f/11 (Jakubiec et al. 2016a), changed from f/4 (Inanici 2006), to capture a larger luminance range (~850, 000 cd/m^2).
- Measure the luminance of a neutral grey target using a luminance meter and measure the vertical illuminance at the front of the fisheye lens using an illuminance meter, every time before and after the HDR capture process. These readings serve multiple purposes including calibrating and fine-tuning the HDR images, detecting the changing light levels during the capturing process, and identifying and correcting the luminous overflow (i.e., high luminance values that are not captured around the solar corona).

The use of HDR photography to acquire luminance maps bypasses the modeling and material measurements process of physically-based simulations. However, one can only capture one HDR image at the moment, and long-term daylighting evaluations require a long-term capturing process. Therefore, HDR photography is mainly used as a technique to collect short-term luminance data for validating the physically-based daylighting simulations and analyzing human visual experience/preference/comfort but is rarely used for annual daylighting assessment without the supplement of simulations.

2.3 DAYLIGHTING EVALUATION METRICS

The discussions in Sections 2.1 and 2.2 are focused on the methodologies for generating or collecting daylight data. Daylighting evaluation metrics are utilized to analyze a given luminous environment. An understanding of these metrics is important, as they will directly affect the process of design decisions. The existing daylighting evaluation metrics can be divided into two categories: illuminance-based metrics and luminance-based metrics.

Illuminance is defined as the density of luminous flux incident on a surface and is measured in lumens per unit area (IES Handbook 2011), and its unit is lux or foot candle. Luminance is defined as the luminous intensity leaving, passing through, or arriving at a surface per unit area in a given direction (IES Handbook 2011), and the unit is candelas per unit area. Human visual system receives luminance reaching the eye, so luminance and luminance-based metrics correlate better with human visual performance and visual comfort, compared to illuminance-based counterparts.

There are many historical reasons for the prevalence of the illuminance-based metrics: 1) The goal for daylighting utilization has long been focused mainly on energy-savings. Hence, daylight availability has been the dominant evaluation criteria. 2) Both the measurements and calculations of illuminance quantities are less expensive compared to luminance values due to lower resolutions

of calculation points and lack of dependency on view directions. Therefore, they have been widely used by lighting professionals for evaluating daylight availability and distribution within a scene. Illuminance-based metrics have been first integrated into dynamic daylight simulations to assess the space based on the full range of daylighting conditions throughout a year. These annual simulations produce a substantial amount of information, thousands of calculation time steps (e.g., each daylight hour) multiplied by the number of sensors in the calculation grid, resulting in millions of illuminance data. Yet, the resolution of sensors in a calculation grid is miniscule compared to the resolution of pixels in a luminance map. Many climate-based metrics have been gradually developed and refined to and interpret and reduce the calculated illuminance data into a single (or a few) value.

In 1989, the classic Daylight Autonomy (DA) is proposed as 'the percentage of the year that a given sensor point (normally at the work plane level) in a scene is above a specified illuminance level' (Association Suisse des Electriciens). A later version refines the definition to take building type and occupant schedule into consideration, defined as 'the percentage of the occupied time of the year when the minimum illuminance requirement for a specific building type at a given sensor is met by daylight alone' (Reinhart and Walkenhorst 2001). In 2006, Continuous Daylight Autonomy (CDA) is proposed to acknowledge the partial contribution of low-level daylight, as it is still beneficial to illuminate the space (Roger 2006). CDA gives credit to time steps when illuminance levels are lower than the DA requirement, but higher than zero. However, neither DA or CDA take the negative impacts of excessive daylight into consideration.

Useful Daylight Illuminance (UDI) (Nabil and Mardaljevic 2005) was proposed to evaluate not only the sufficiency of daylight, but also the penalty of excessive daylight. UDI is defined as the percentage of annual daytime hours that a given sensor point is between 100 lux (too dark to

light the space) and 3000 lux (potential leading to glare) (Mardaljevic et al. 2012). The latest version of UDI separates the light into four bands: the underlit (<100 lux), the supplemental (>100 and <300 lux), the useful (>200 and <3000 lux) and the overlit range (>3000 lux).

Annual Sunlight Exposure (ASE) is proposed by IES LM-83 to determine the extent of direct sun penetrations and the potential for visual discomfort (IESNA 2012). $ASE_{1000lx,250h}$ is defined as 'the percentage of an analysis area that exceeds 1000 lux due to direct sunlight for more than 250 hours per year'. ASE is found to have the best correlation with the subjective response to glare (Heschong Mahone Group 2012). It is a simplified metric and it does not consider the contributions of diffuse light or inter-reflections. It is typically used to supplement the Spatial Daylight Autonomy (sDA). sDA takes the spatial distribution of illuminance into consideration to determine the daylight illuminance sufficiency for a given area" (IESNA 2012). $sDA_{300,\ 50\%}$ reports the percentage of space (sensor points) that exceeds the illuminance target, 300 lux, for at least 50% of occupied hours.

All of the above metrics are based on horizontal illuminances (derived from a grid of upward facing illuminance sensors) (Figure 5). Vertical illuminances are promoted by many human factor studies. For example, vertical illuminances are more effective than horizontal illuminances for evaluating the visual discomfort in a scene (Wienold 2009, Suk et al. 2013, Konis 2014, van den Wymelenberg and Inanici 2014, 2016) due to their alignment with human view direction and visual field (Figure 5). However, vertical illuminance is only one value and still cannot provide enough information to describe the luminance distribution over the entire field of view. For instance, the same vertical illuminance could result from a uniformly bright scene or a relatively small light source with a high luminance. Our perception of the scene will be different, yet the vertical illuminance (the total amount of light reaching the eye) will be similar. The vertical illuminance

has been found to have less accuracy for predicting visual discomfort caused by high contrast in comparison to high luminance (Kleindienst and Andersen 2011, Jakubiec and Reinhart 2012).

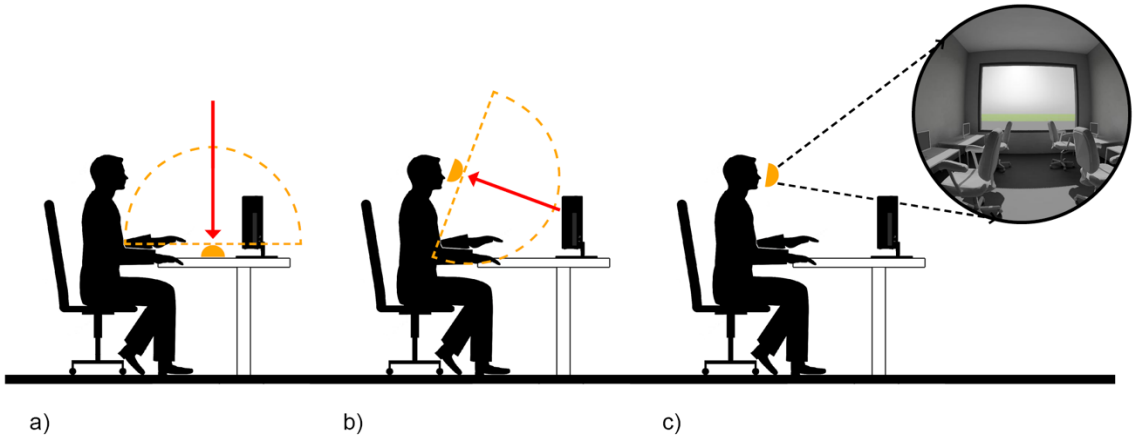


Figure 5. Diagrams show a) horizontal illuminances measured at the desk, b) vertical illuminances measured at the eye, and c) luminance maps that capture the subject's field of view.

Recently, design and research practices of daylighting are shifting from illuminance-based metrics towards human-centric luminance-based evaluations. Luminance maps capture the subject's field of view (Figure 5) and provide millions of values that can be used to evaluate the variabilities in the luminous environment, such as the luminance variations, contrast, positions and intensity of light sources, and background luminance levels. Luminance distribution patterns determine the visual comfort (e.g., discomfort glare), visual appearance, and task visibility in a given scene (Wienold and Christoffersen 2006, Jakubiec and Reinhart 2012, Suk et al. 2013, Konis 2014, van den Wymelenberg and Inanici 2015, Rockcastle et al. 2016). Commonly used luminance-based metrics including falsecolor images, glare-related visual comfort indices, and contrast based evaluations.

Falsecolor images are visual presentations of luminance maps. The HDR luminance maps cannot be presented in absolute values on conventional display devices. In falsecolor images, a

limited number of colors are assigned to the luminance values that can span several orders of magnitude. The color of each image pixel corresponds to a luminance range as indicated by the color bar scale. Therefore, falsecolor images illustrate the spatial distribution of high and low luminance ranges with a scene. Logarithmic scales are utilized to correlate with the human visual system (Fechner 1966). Figure 6 shows a falsecolor image based on the HDR rendering of an interior scene. Areas with high luminance (>3000cd/m^2) are shown in yellow, while greenish and blueish color represent pixels with lower luminance.

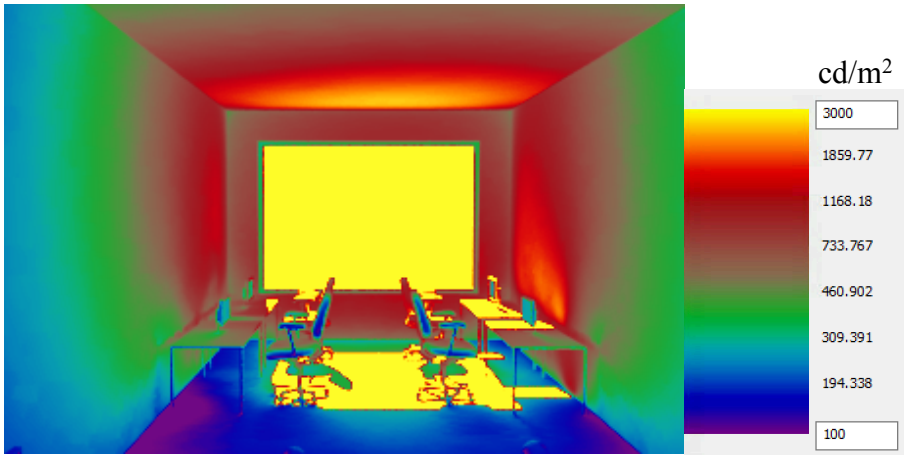


Figure 6. An example falsecolor image of an interior scene

Glare indices evaluate scenes for their potential to cause visual discomfort. They are defined as equations that describe the subjective magnitude of glare discomfort with high values illustrating higher possibility or intensity of glare perception. Four variables are known to participate in the perception of discomfort glare (*IES Lighting Handbook: Reference and Application* 2011): adaptation (background) luminance, the luminance of the glare source, size of the glare source, and location of the glare source in the field of view. Their relationship can be expressed using a simplified equation (Equation 1)

$$Discomfort\ Glare = \sum_{i=1}^{n} \frac{L_{s,i}^{exp} \omega_{s,i}}{L_b^{exp} P_i^{exp}}$$

Equation 1. Major components of a discomfort glare equation (R. G. Hopkinson and Collins 1970)

In this equation, the exponent is a weighting component applied to each variable. It can be observed that the intensity and possibility of discomfort glare increases while the luminance ($L_{s,i}$) and the solid angle size ($\omega_{s,i}$) of a glare source increase. On the contrary, the probability of discomfort glare perception will reduce if a glare source is further away from the center of the visual field. Originally developed by Guth (1949), the position index P_i is a convenient way to weigh a glare source based on its position in the field of view. P_i becomes larger when a glare source moves further from the center of the view, and it has a value of one if the source is located exactly at the center (Figure 7). Another important variable is the luminance of the background (L_b), which increases while the probability of discomfort glare decreases (R. G. Hopkinson and Collins 1970).

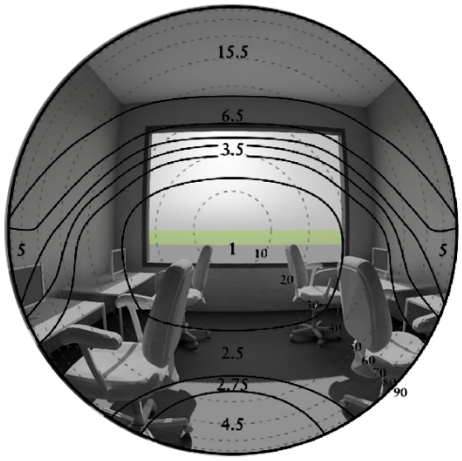


Figure 7. A diagram shows selected values of the Guth position index (Luckiesh and Guth 1949) plotted on a top view of 180° fisheye projection. The solid lines separate different index values while the dash lines indicate the angles between the view center and the objects.

Different datasets gathered through different human subject studies have led to different glare equations. Among all existing glare indices, only two glare indices are specifically developed and intended for use in daylit environments: the Daylight Glare Index (DGI) (Hopkinson 1972) and the Daylight Glare Probability (DGP) (Wienold and Christoffersen 2006). DGP is the first glare index based on human subjects study in a real daylit setting and is suggested to have the best performance among other existing glare indices for most daylit environments (J. Jakubiec and Reinhart 2012). The output of the DGP model is given on a four-point scale as shown in Table 1.

$$DGP = 5.87 \times 10^{-5} E_v + 9.18 \times 10^{-2} \log\left(1 + \sum_{i=1}^{n} \frac{L_{s,i}^2 \omega_{s,i}}{E_v^{1.87} P_i^2}\right) + 0.16$$

where,

E_v is the vertical illuminance at the eye (lux),

$L_{s,i}$ is the luminance of one glare source (cd/m²),

$\omega_{s,i}$ is the solid angle of one glare source (sr), and

P_i is the position index of the element of the glare source

Equation 2. Components of DGP calculation (Wienold and Christoffersen 2006)

Table 1. Output of the DGP model

Degrees of Perceived Glare	**DGP Value**
Intolerable glare (estimated 45% or more of people would feel discomfort)	DGP > 0.45
Disturbing (estimated 40% to 45% of people would feel discomfort)	0.40< DGP <=0.45
Perceptible (estimated 35% to 40% of people would feel discomfort)	0.35< DGP <=0.40
Imperceptible (estimated less than 34% of people would feel discomfort)	DGP < =0.35

DGP (and other glare indices such as DGI) can be automatically calculated by Evalglare (Wienold 2010) through analyzing the luminance maps of the scene. In Evalglare, glare sources are identified (Figure 8) based on a threshold value that can be: 1) a fixed value defined manually by users; 2) a value that is certain times higher than the average luminance of a task area; 3) a value that is certain times higher than the average luminance of the entire image (the adaptation level) (Wienold and Christoffersen 2006). It is recommended to calculate DGP using fisheye maps to cover the entire field of views of subjects.

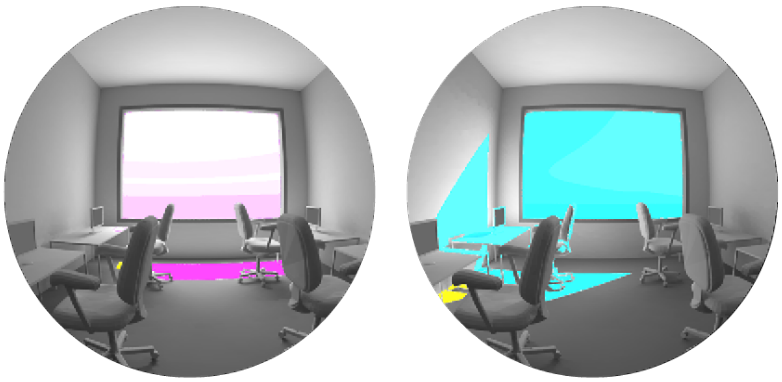


Figure 8. Examples of luminance maps evaluated by Evalglare with colors highlight the detected glare sources

As opposed to the glare-related visual comfort metrics, which focus on evaluating the negative aspect (i.e. undesirable distribution or range) of daylight, perceptual performance criteria are developed to evaluate the positive experiences (i.e. the contrast-based perceptual characteristics of daylight). The contrasting patterns of sunlight and shadows have long been employed by architects and designers. The benefits of these patterns and their impact on spatial interest and aesthetics of daylight scenes cannot be directly measured, but instead requires human subjects' assessment. Researchers showed that lighting preferences of human subjects can be evaluated using physically-based renderings, instead of real-world scenes (Newsham et al. 2005). In 2016, Rockcastle et al.

explored the correlation between the contrast measures of images and human subjects' assessment of Radiance rendered luminance maps of daylit scenes under various sky conditions (Rockcastle et al. 2016). Their research showed that spatial contrast as measured in RAMMG (Rizzi et al. 2004) and the modified RAMMG have the highest correlation score with human subject assessment of excitement levels (calm versus exciting), among all other tested contrast-based metrics such as Michelson (Michelson 1927), Difference of Gaussian (DOG) (Tadmor and Tolhurst 2000), Root Mean Square (RMS) (Pavel et al. 1987), and Retinal-like Subsampling Contrast (RSC) (Simone 2009). RAMMG (Equation 3) performs a pyramid subsampling of the image to various levels and combines the average contrast calculated at each level. At each level, the image resolution is reduced to half of the previous level, and the contrast is calculated as the sum of the neighborhood contrast for every pixel. This approach is based upon the principle that perception mechanisms can be explained with multi-channel analysis (McCann 1999). Other than RAMMG calculated with a series levels of image resolutions, modified RAMMG calculated with a specific (single) level, referred to as Modified Spatial Contrast, were also suggested to be used as perceptual predictors for contrast-based visual effects (Rockcastle et al. 2016).

$$RAMMG = \frac{1}{N_l}\sum_{l}^{N_l} \bar{C}_l$$

$$\bar{C}_l = \frac{1}{8m_l}\sum_{i=0}^{m_l}\sum_{j=0}^{8} M_{RAMMG}\left|L_i - L_j\right|$$

$$M_{RAMMG} = \frac{1}{4+2\sqrt{2}}\begin{bmatrix} \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} \\ 1 & 0 & 1 \\ \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} \end{bmatrix}$$

where,

$RAMMG$ *is the combined contrast for an image,*

N_l *is the number of levels,*

l *is the level number,*

$\bar{C}_l$ *is the mean contrast in level* l*,*

m_l *is the total number of pixels in level* l*,*

L_i *is the luminance value for one pixel* i *,*

L_j *is the luminance value of one surrounding pixel (out of eight neighbors) of pixel* i*, and*

M_{RAMMG} *is the weight matrix for the neighbor pixels.*

Equation 3. RAMMG contrast equation components (Rizzi et al. 2004)

There is an increasing interest in integrating two more dimensions into luminance-based evaluations: temporal and spatial variance (Jakubiec and Reinhart 2012, Jones and Reinhart 2017, Rockcastle et al. 2017). All existing luminance-based analyses, falsecolor imagery, DGP, and RAMMG, aim to evaluate the luminance map of a single moment. As discussed several times in the previous sections, dynamic daylighting conditions need to be considered in order to have a comprehensive evaluation of the luminous environment over the course of the entire year.

However, different from illuminance-based metrics, there are not any established annual luminance-based metrics. The main reason is that acquiring annual luminance maps is significantly labor-intensive and time-consuming, and as a result, practitioners are less motivated to integrate annual luminance-based metrics into the design evaluation process. Another fact that will restrain designers and consultants to consider such metrics is the overwhelming amount of luminance values or DGPs/RAMMGs resulting from annual calculations. Metrics that interpret and condense the information from long-term luminance-based calculations into practical evaluations are needed. In addition to the need for developing long-term evaluation metrics, researchers are investigating luminance-based evaluations that integrate subject's visual experience over multiple view directions (Jakubiec and Reinhart 2012, Rockcastle et al. 2017). We experience the luminous environment in 360° instead of a fixed view direction. A slight adjustment in the view direction can result in significant glare reduction or increase. This is especially common for open plan offices where a subject may change head and monitor orientation to avoid glare. Generating images of multiple view directions are more time-consuming than a single fixed-view image. Making annual luminance-based simulations and measurements more accessible will enable further development of luminance-based metrics for daylighting studies and foster wider usage of human-centric evaluations -of long-term and over multiple view directions - in the design processes.

2.4 Accelerating Annual Daylighting Calculations

As discussed in the previous sections, daylighting evaluation period should be a full year to capture seasonal variations of daylight, and the time step should be small enough (e.g., 1 hour) to capture the observed short-term daylight variations. Annual luminance-based evaluations require acquiring a luminance map at each time step of the entire year, resulting in labor-intensive and time-consuming simulations or impracticable long-term field measurements.

Accelerating annual daylight evaluations is an active area of research. *Daylight Coefficient* (DC) approach has been developed as a numerical methodology to accelerate annual daylighting simulations by using an efficient approximation (Tregenza and Waters 1983). The classic DC theoretically divides the celestial hemisphere into discrete (e.g.,145) sky segments and calculates the contribution of each segment to the illuminance (or luminance) level at various sensor points (Figure 9). Each illuminance value, $E_\alpha(x)$, at a sensor point and orientation, x, for a sky segment, α, is normalized with the luminance of the sky segment, L_α, and angular size, ΔS_α, and results in a daylight coefficient $DC_\alpha(x)$ (Equation 4). Daylight coefficient depends on the physical characteristics of the room and the external environment, such as the building geometry, surface material properties, and outside context, and is independent of the luminance of each sky patch. After gathering all daylight coefficient values, the illuminance of each sensor point (or the luminance at a pixel point) at each constant time, $E(x)$, is calculated by linear superposition of each daylight coefficient multiplied by its sky segment time-varying luminance (Equation 4). The luminance data of the sky segment can be calculated using meteorological data and sky models (e.g., Perez Sky Models). The most time-consuming step in DC approach is calculating the daylight coefficients, as ray-tracing is involved. But this step only consists of a limited number of simulations, depending on the sky division scheme. The rest of the calculations are mathematical and can be quickly performed by current computers.

$$DC_\alpha(x) = \frac{E_\alpha(x)}{L_\alpha \Delta S_\alpha}$$

$$E(x) = \sum_{\alpha=1}^{N} DC_\alpha(x) L_\alpha \Delta S_\alpha$$

Equation 4. Classic Daylight Coefficient approach

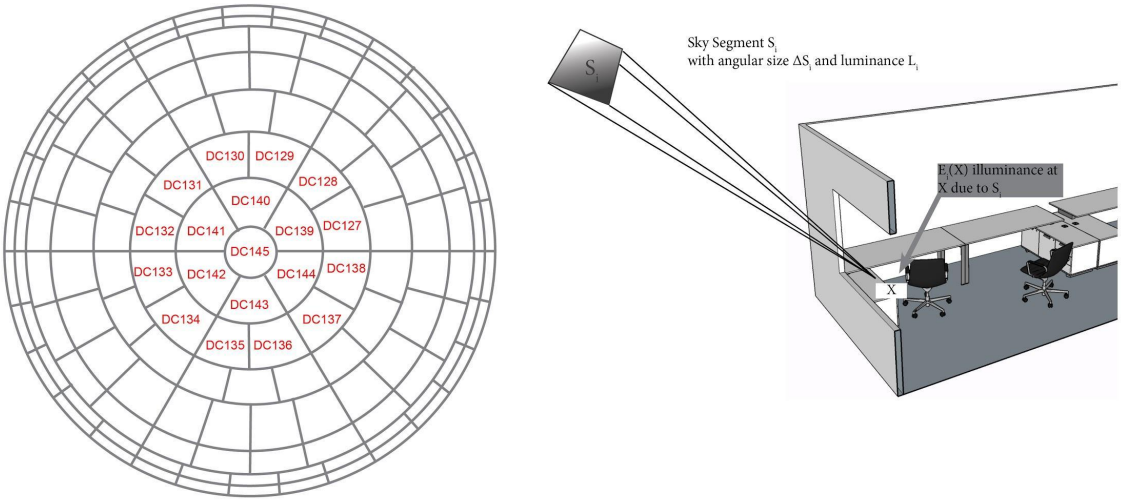


Figure 9. In classic DC method, (left) the celestial hemisphere is divided into 145 discrete segments, and (right) the contribution of each segment to the illuminance level at various sensor points is calculated.

Errors in classic DC method occur due to the difference between actual sun positions and sample sun positions (center of sky patches). This difference is referred to as the Sun Displacement Angle (SDA) (Mardaljevic 2000a). The direct component of illumination from the sun can be a significant source of errors. These errors could be reduced by using a larger number of sun positions. Direct light source calculation also requires less computational resources compared to the indirect light source calculation. Therefore, further development of DC methods involves differentiating direct and indirect contribution from the sun and the sky, resulting in four components of illumination (Figure 10). Researchers have explored using 5010 (Mardaljevic 2000b), 65 (Reinhart and Walkenhorst 2001), and 2305 (Bourgeois et al. 2008) evenly distributed direct solar positions, where the direct sun contribution is calculated by identifying the nearest sun positions. It has been proven that higher solar discretization resolution results in better captures of complex sunlight patterns, but also, increased calculation time (Bourgeois et al. 2008). Several studies suggest that dynamic lighting simulation, using the DC approach, Perez Sky models, and

backward ray-tracing, is able to reliably calculate time series of illuminance and luminance in buildings (Mardaljevic 2000b, Reinhart and Walkenhorst 2001).

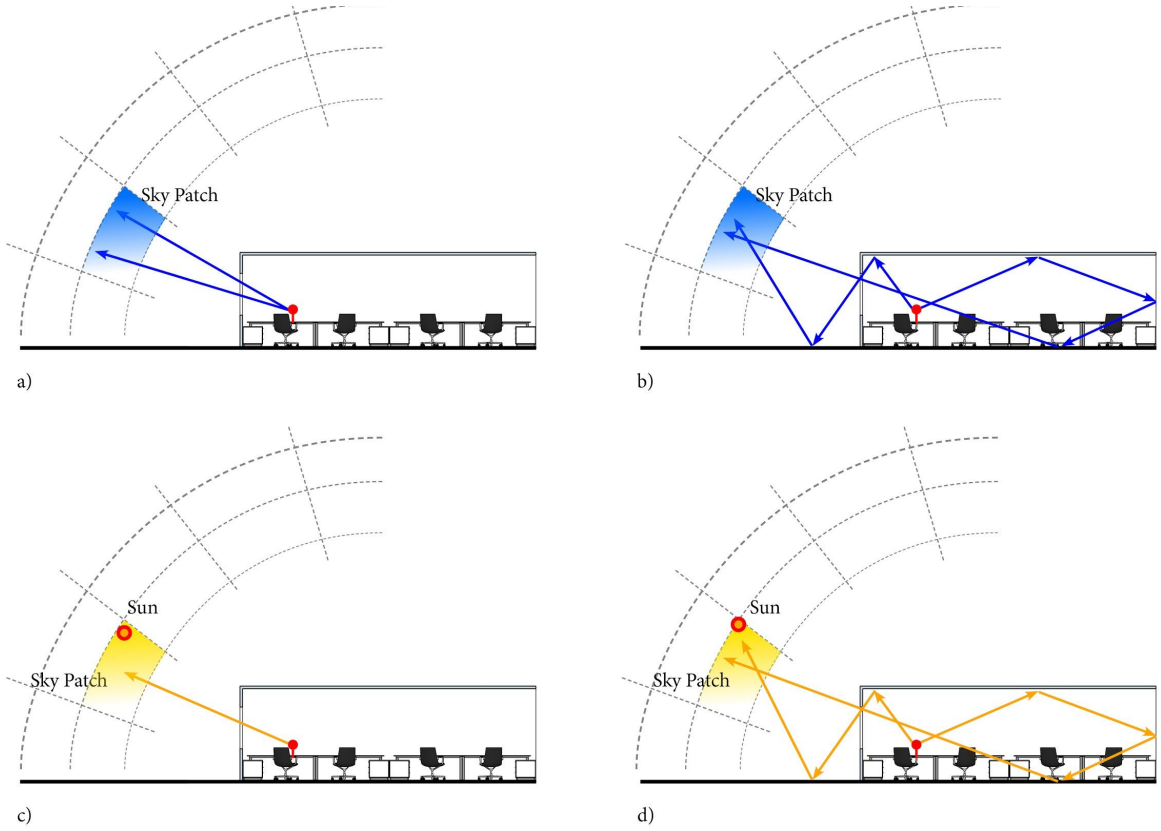


Figure 10. The four components of illumination in DC calculations: a) Direct sky component, b) Indirect sky component, c) Direct sun component, and d) Indirect sun component

DC technique does not adequately model complex fenestration systems (CFS), such as venetian blinds. The regular backward ray-tracing simulation process may not effectively calculate the inter-reflections and transmissions that occur between the layers of the fenestration system before a ray is passed to the interior space. Further developments of DC methods divide the light flux transfer process into three phases, named the Three-phase method, to better model CFS (Ward et al. 2011, McNeil 2013a). The single daylight coefficient in the DC method is expanded into the view matrix, the transmission matrix and the daylight matrix (Figure 11). BSDF data is

incorporated into the process to describe the complex light properties of CFS. Similar to the classic DC approach, the errors of the Three-phase method occur due to inaccurate calculations of direct solar contributions. The energy of the sun is spread to the closest three sky patches, which have a much larger area than sun disc. As a result, the finer details of sun penetrations in the scene may be missing. Five-phase method (McNeil 2013b) improves the result by replacing the direct solar contributions of the original Three-phase method with increased resolution ones (Figure 12). However, five phase method requires days to weeks of computation time.

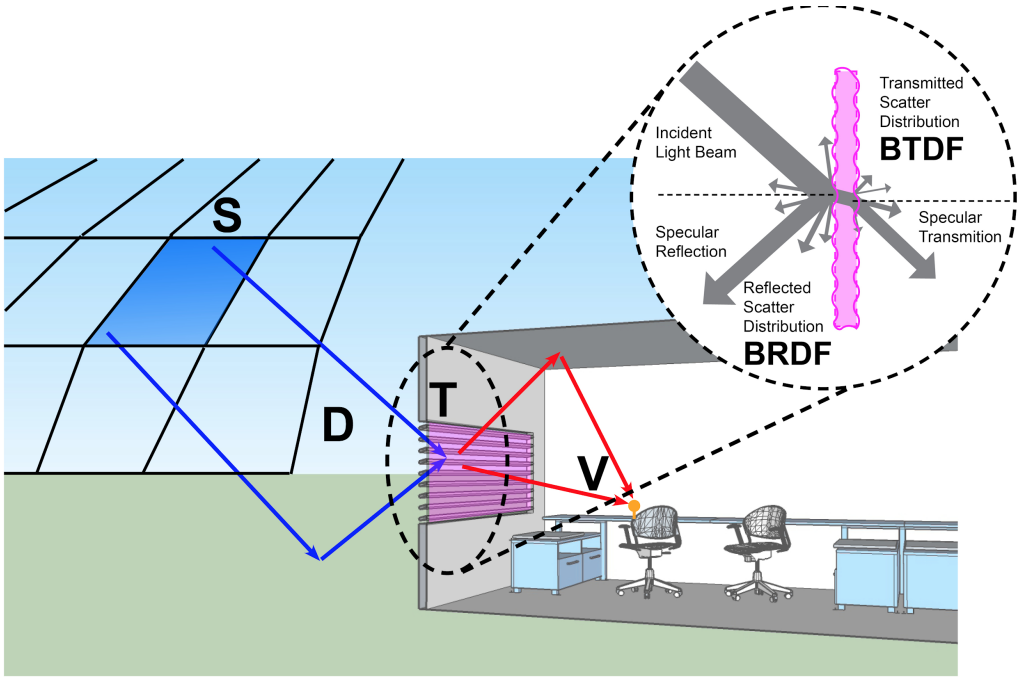


Figure 11. The Three-phase method splits up the daylight flux transfer from the sky patch (S) to an interior sensor point to the daylight matrix (D), the transmission matrix (T) and the view matrix (V). BSDF data is incorporated into the process to describe the complex light properties of CFS.

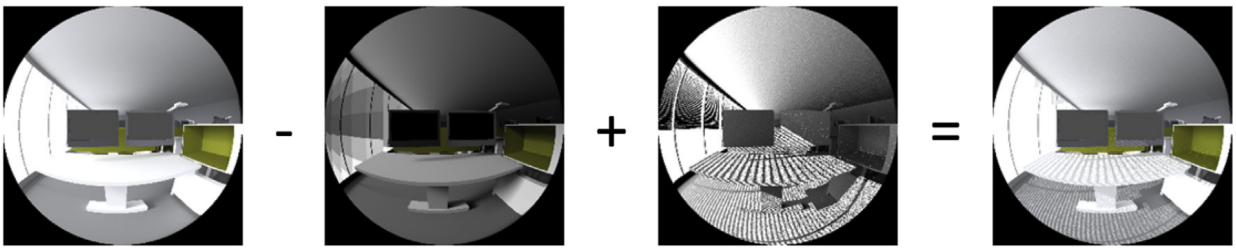


Figure 12. The transition between the Three-phase and Five-phase method techniques (Inanici and Hashemloo 2017).

Recent accelerating methodologies of long-term daylighting evaluations leverage the developments in modern computing technologies. One solution is to utilize parallel computing to trace multiple primary rays simultaneously on a graphics process unit (GPU) (Jones and Reinhart 2014, Zuo et al. 2014). Due to its highly parallel structure, a modern GPU can efficiently process a large block of data, which is not achievable by a general-purpose central processing unit.

Another approach is to utilize machine learning to create predictive models from observed data, bypassing the expensive computations involved with physically-based modeling systems. The methodology employed in this dissertation explores the use of machine learning, particularly DNNs, to accelerate long-term per-pixel luminance predictions of indoor scenes. Statistical methods have been previously investigated for per-pixel luminance predictions (Inanici 2013). They succeed in predicting images under overcast skies, but require further improvements for sunny sky conditions. The machine learning techniques and related context will be reviewed in detail in Chapter 3, and the methodology of the dissertation will be introduced in Chapter 4 and 5.

Chapter 3. MACHINE LEARNING

This chapter provides the context of the machine learning techniques employed in the dissertation. Section 3.1 introduces the basics of machine learning. Sections 3.2 and 3.3 describe two specific machine learning techniques which have been widely used for image-related tasks: conventional feedforward neural networks and convolutional neural networks. Section 3.4 outlines the techniques for improving learning quality. Finally, Section 3.5 reviews the studies that utilize machine learning for lighting-related tasks.

3.1 FUNDAMENTAL CONCEPTS

Machine learning, named by Arthur Samuel (1959), provides computer systems the capability of learning from data, and making data-driven decisions or predictions without being explicitly programmed. In many cases, it is easier to make a system learn from examples of desired input and output behavior, than manually specifying the exact logic of how to get from the input to the output of the task. Making the system self-learning is the primary goal of machine learning, which differentiates it from traditional statistics algorithms.

Supervised and unsupervised learning are two essential categories of machine learning (Figure 13). Supervised learning is the most widely used machine-learning method (Hastie et al. 2011), performed by using a dataset which contains input (features) and output (corresponding labels) pairs. The task of supervised learning is to predict the labels of the objects given the sets of features. The algorithm learns by comparing the predicted output with correct outputs to find errors, which are measured using a loss function (i.e., a function calculates the difference between predicted output and the ground truth labels). It then modifies the model accordingly. For example, in learning to detect architectural components (Figure 14), the tasks are 1) to detect regions that

may contain the architectural components in images of architectural interiors, and 2) to assign a classification label of "door", "window", or "furniture" to these regions. The second task of assigning classification labels provides a good example of supervised learning algorithm. The algorithm is provided with an image collection of interior scenes with various objects, each labeled with the correct category. The algorithm learns the patterns that relate the objects to their labels through improving the performance metrics (e.g., the accuracy of this object classifier). On the contrary, unsupervised learning attempts to identify hidden patterns in unlabeled input data. The algorithms learn and organize information without an error signal to evaluate the potential solution. Unsupervised learning methods are applied to tasks, where the labels are unobservable, infeasible, or even impossible to obtain, and can be advantageous in certain situations, since it lets the algorithm to look for patterns that have not been previously considered. For example, a subgroup of unsupervised learning is clustering, which aims at finding partitions of the data without labels that specify the desired partition. Figure 15 shows the calculation steps of one clustering algorithm, K-means (proposed by Stuart Lloyd in 1957 and published in 1982), with k=2, over an illustrative dataset. The user provides the algorithm with a set of data points $x_1 \ldots x_n$ and a parameter k (i.e., the number of resulting clusters). The algorithm initializes the clusters by placing k centroids $c_1 \ldots c_k$ at random locations. The clusters are updated with a two-step process: 1) each data point x_i, is assigned to the cluster with the nearest centroid c_j, 2) for each cluster j = 1…k, the algorithm updates the location of the centroid c_j to the center (i.e., the mean) of the points x_i assigned to that cluster. These two steps repeat until none of the cluster assignments change in the process. In the end, the K-means clustering algorithm partitions the data into k relatively equal-size groups. Note that the algorithm applies to high-dimensional data as well as low-dimensional data.

Supervised learning can be divided into two subgroups: regression and classification. Regression is the problem of estimating or predicting a continuous quantity. For instance, in the case of house price estimation, the task is to predict the house value based on various meaningful attributes of the house as its features. The historical data of house features and the price is provided to a regression estimator, which builds a continuous functional map between the features and the price values. Classification is the method that assigns observations into discrete categories. The previous example of object detection is a classification problem.

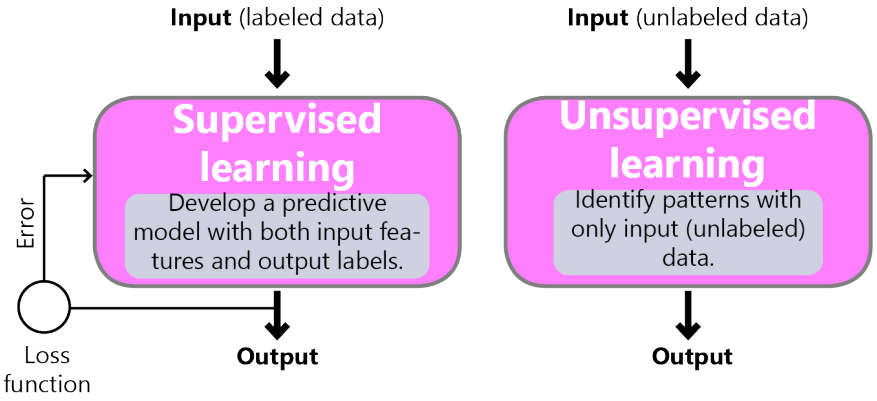


Figure 13. Diagrams show the concepts of (left) supervised learning and (right) unsupervised learning

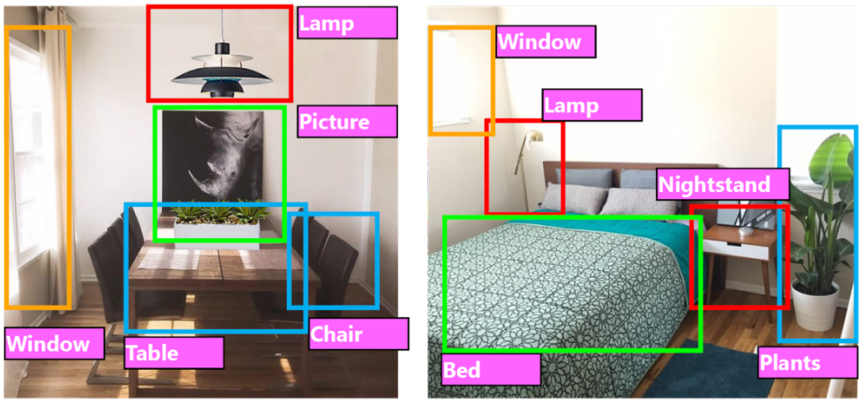


Figure 14. An example of a supervised task to detect architectural objects: the system learns through examples of input (images contains various objects) and corresponding output (objects labeled with the correct classifications), and then predicts the classification of new objects in images it has never seen (reproduced from Izadinia 2017).

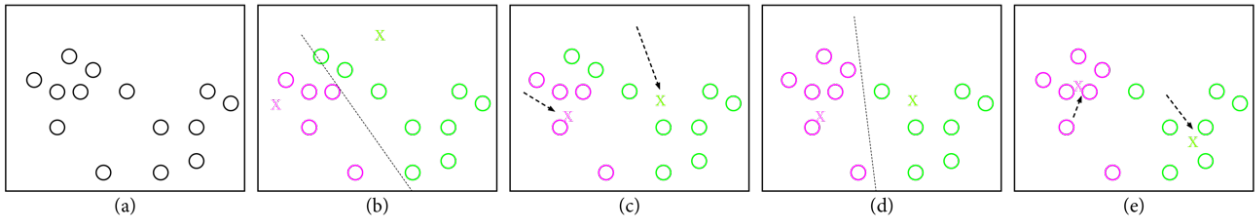


Figure 15. An example illustrates the calculation steps of K-means clustering algorithm with K = 2: (a) the original data (marked as circles) to be partitioned; (b) 2 (k = 2) centroids (marked as X) are placed at random locations, and each data point is assigned to the cluster with the nearest centroid, resulting in 2 clusters marked in red and green; (c) each centroid is updated and moved to the center (the x & y locations of the centroid = the mean x & y locations of all points) of the cluster; (d) each data point is assigned to the cluster with the smallest distance to the updated centroid; (e) repeat step (c). The step (c) and (d) will be repeated until none of the cluster assignments change.

3.2 Artificial Neural Networks

The concept of deep learning stems from the research on artificial neural networks (ANN) (Hinton and Salakhutdinov 2006). ANNs were originally proposed in 1943 (McCulloch and Pitts 1943). They are inspired by biological neurons of animal brains and consist of connected computing units, also called neurons. These neurons are organized in a hierarchical fashion of input layers, output layers, and hidden layers in between. The connection between two neurons has a weight which defines how much the previous neuron influences the next one. Each neuron has a bias parameter which defines how much the output of the transformation at the neuron is biased in the absence of any input. Every neuron is fully connected to all neurons in the previous layer. Figure 16(a) shows the basic schematic of a conventional feedforward ANN. The circles represent the neurons of the ANN, and the arrows represent the weights. The conventional feedforward ANN is the

quintessential deep learning model and forms the basis for many other important applications (Goodfellow et al. 2016).

ANNs condense or aggregate information into informative features and find relationships between inputs and outputs through a learning process. During the learning process, information is transferred from layer to layer and learned knowledge is stored in weights and bias. There are three stages of the network training: a forward stage, an error estimation stage, and a backward stage (Figure 17).

- In the forward stage, inputs are transferred to the output layer, with input features represented with the current weights in each layer. These weights are initialized with random values. Figure 16(b) shows how information is transferred through a typical neuron. The neuron receives a weighted sum of the inputs from the previous layer of neurons with added bias. The weighted sum then runs through an activation function, whose output serves as the input of the neurons in the next layer.
- In the error estimation stage, loss cost is calculated using a loss function.
- In the backpropagation (Linnainmaa 1970) stage, the goal is to increase the accuracy of the network predictions by minimizing the loss cost. This is most commonly done using a Gradient Descent optimization, which updates the model's parameters in the opposite direction of the gradient of the loss function (i.e., the partial derivative of the loss cost with respect to a learnable weight). In other words, the optimization algorithm follows the slope of the surface created by the loss function downhill until it reaches a valley, which is the (local) minimum. The learning rate determines the size of the steps the algorithm takes to reach a minimum. Further improvements of Gradient Descent such as Adagrad (Duchi et al. 2011) and Adam (Kingma and Ba

2015) aim at increasing the speed and reliability of the optimization. In this way, the loss cost is propagated backward from the output layer to the input layer, and all the weights and biases are updated and prepared for the next forward computation.

- The learning can be stopped when the loss is minimized after sufficient iterations of the forward and backpropagation stages.

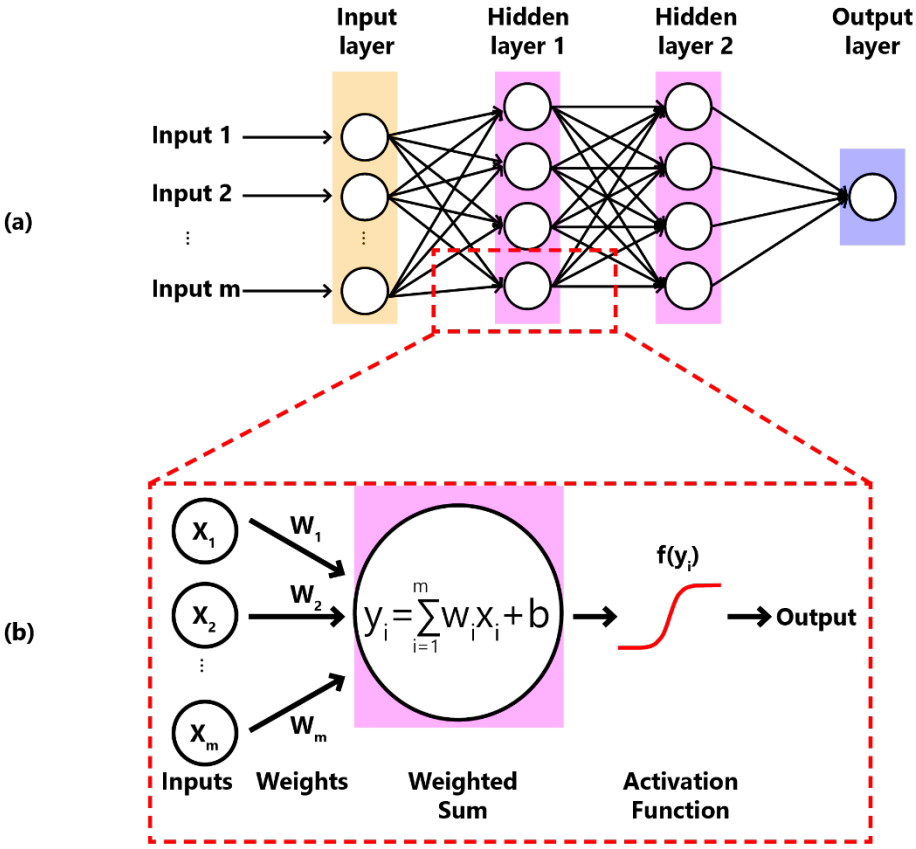


Figure 16. Diagrams show (a) a conventional feedforward neural network consists of fully connected neurons, information is transferred from input neurons through neurons in hidden layers and reach the output layer, and specifically (b) the input to each neuron is computed as the weighted sum of the neurons from the previous layer, followed by a non-linear activation function, whose output becomes the input for neurons in the next layer (Image adapted from CS231n 2019).

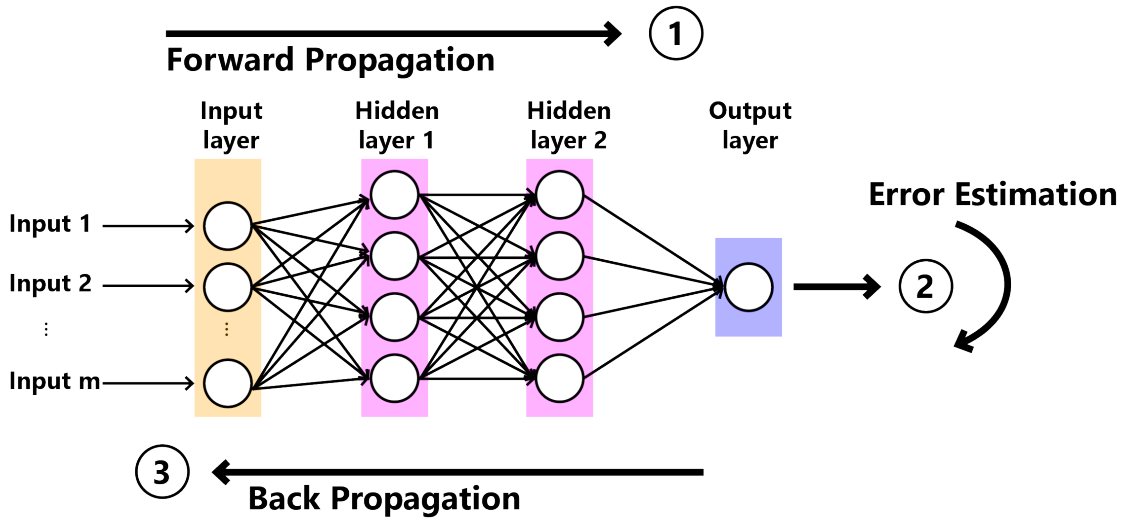


Figure 17. After computations flow forward from the input to output, at the output layer the error derivatives are calculated and backpropagate towards the input layer, so that weights can be updated to optimize the loss functions.

It is worth mentioning the importance of activation functions. The nonlinear activation functions, which are the most used activation functions, add the nonlinear transformation at the end of the neuron's work and prevent all the linear steps from combining. Without the activation functions, a neural network is a linear function. The nonlinearity makes the neural network able to learn and solve complex problems, such as image classification and language translations. Figure 18 shows four common non-linear activation functions used in neural networks. Among them, sigmoid (logistic function) (Figure 18(a)) and hyperbolic tangent (tanh) (Figure 18(b)) are smooth functions. They both transform the entire input range (from negative to positive infinity) into a small range of values: [0,1] for the sigmoid functions and [−1,1] for tanh functions. However, when input values become very positive or negative, the function will not discriminate between similar large values, and the training comes to a standstill. The Rectified Linear Units (ReLU) function (Hahnloser et al. 2000 first published, Nair and Hinton 2010 popularized) has been shown to have good results for most of the feedforward neural networks (Goodfellow 2016). It is a

function over 2 different function domains (Figure 18(c)). If the input is equal to or less than zero, the output is zero. Otherwise, the output is the same as the input. ReLU benefits from being nearly linear, which makes it easy to optimize with gradient-based methods and can generalize well. However, the ReLU suffers the same problem as sigmoid and tanh functions because it always returns zero for zero and negative values. Maas et al. (2013) proposed a generalization of ReLU - Leaky ReLU (Figure 18(d)), which instead of outputting a zero for any non-positive value, outputting the input scaled down by a factor of a small constant value (e.g., 0.01).

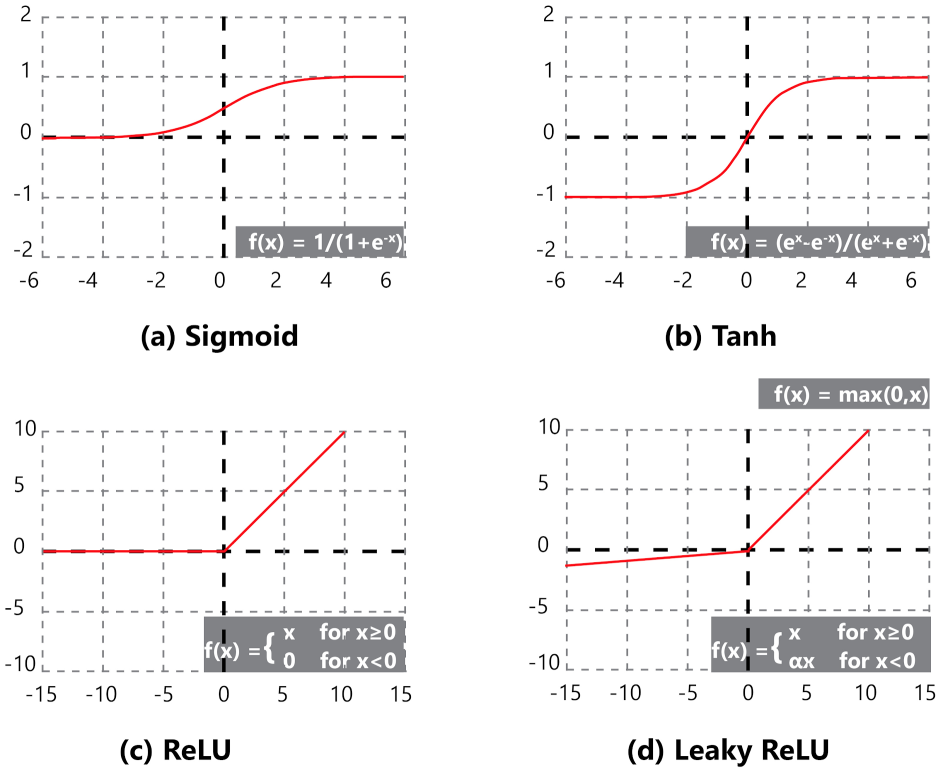


Figure 18. Diagrams of four commonly used activation functions in ANN models.

One drawback of conventional ANNs is that they are not designed for images as the inputs. For example, a 300×300×3 pixel image requires 270,000 neurons in the network's input layer. Each neuron corresponds to one pixel in the image. The full connectivity between different layers

of an ANN leads to a large number of parameters and consequently a large computational effort for training them. One way to reduce the number of parameters is via Convolutional Neural Networks.

3.3 Convolutional Neural Networks

The Convolutional Neural Network (CNN) is a deep learning approach, where multiple layers are trained in a robust manner (LeCun et al.1998). CNNs are similar to conventional ANNs in many ways but are designed for cases when the inputs are images, thus are commonly used in various computer vision research and applications. Compared with conventional ANNs, CNNs make the training process easier through some internal mechanisms which reduce the number of trainable parameters.

Unlike a conventional ANN, the neurons in layers of a CNN are arranged in three dimensions, as visualized in Figure 19. Every layer transforms an input 3D volume to an output 3D volume through convolution operations. An input image can be seen as an input volume with width, height, and depth of 3 (RGB channels), which can be directly imported into CNNs.

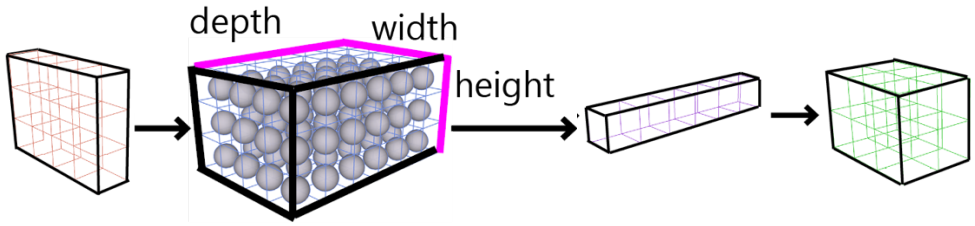


Figure 19. A CNN is made up of layers, in which neurons (represented by grey spheres) are arranged in three dimensions: depth, width, and height.

CNNs usually consist of three types of layers, which are convolutional layers, pooling layers, and fully connected layers. Different layers play different roles.

The core building blocks of a CNN are convolution layers. Convolution layers are inspired by the animal visual cortex organization: the complex arrangements of cells were constrained in the animal visual cortex in charge of light detection in overlapping and small sub-regions (receptive fields) of the visual field (Hubel and Wiesel 1968). Similarly, the weights in the convolution layer consist of a set of learnable filters (i.e., kernels). Each filter is applied to a small local region at different spatial locations in the input volume. These convolution operations are followed by an activation function, whose output is a two-dimensional feature map, as shown in Figure 20. Applying filters to images is performed by calculating the weighted sum of pixel values, followed by a non-linear activation function (Figure 21). The output of each convolution layer, the feature maps generated by the set of filters, become the input of the next layer. Compared with conventional ANNs, there are two obvious advantages of the convolution operations: 1) The convolution operations are only applied to small local regions rather than the whole input volume. Therefore, each neuron in the convolution layer is connected to a limited number of neurons in the previous layer. This characteristic of convolution layer is called local connectivity, which enables the algorithm to learn correlations among neighboring pixels. 2) The neurons on the same feature map are from the same filter. Therefore, the neurons share the weights of the filter. This is called weights sharing and causes CNNs to have a significantly reduced number of parameters when processing images, compared with conventional ANNs (Figure 22). These two mechanisms also make CNNs perform better for recognizing objects in images. Because the same filter (weights) is being moved over the image, the location of the object will not affect the patterns that the networks learn, known as translation invariance.

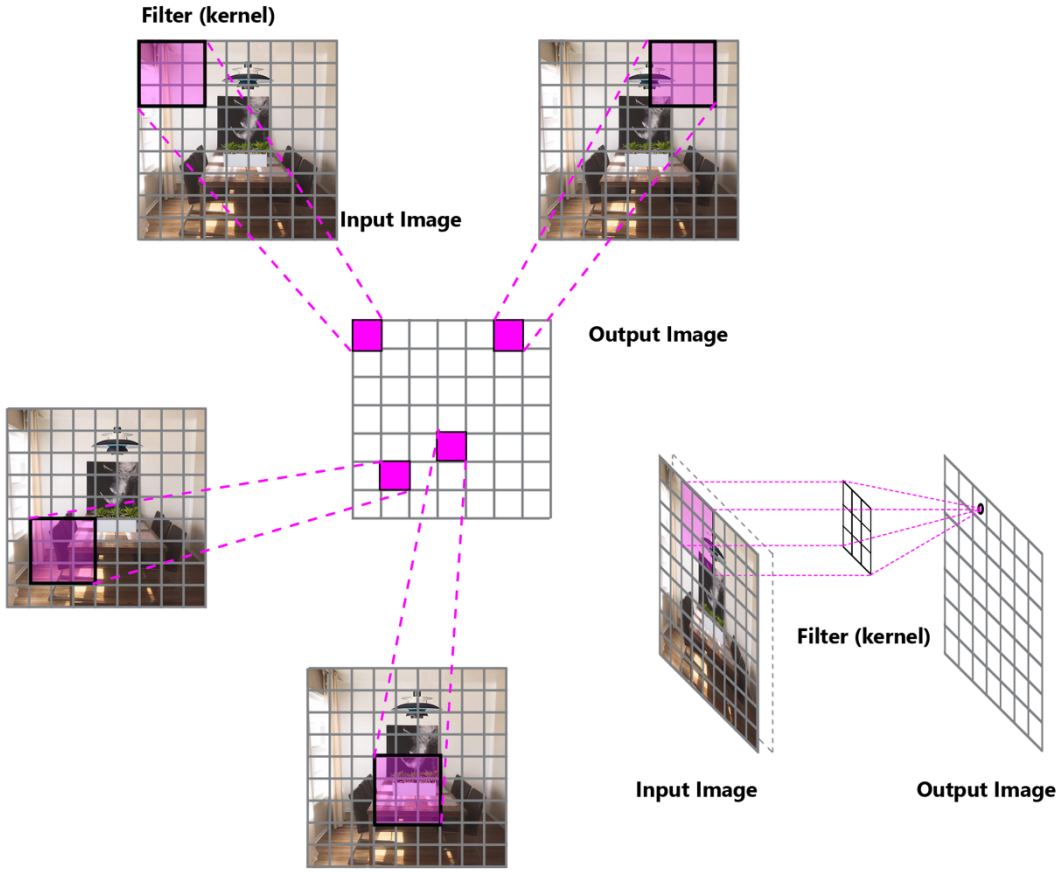


Figure 20. During the image convolution, the kernel filter is moved and applied across the image. The resulting value (a dot product between the kernel filter weights and the region they are connected to in the input image) of a filter operation then becomes the value for the corresponding pixel in the result. The diagram shows a subset of positions of the kernel filter in the input image, and the positions where their computed values go into the output (feature map).

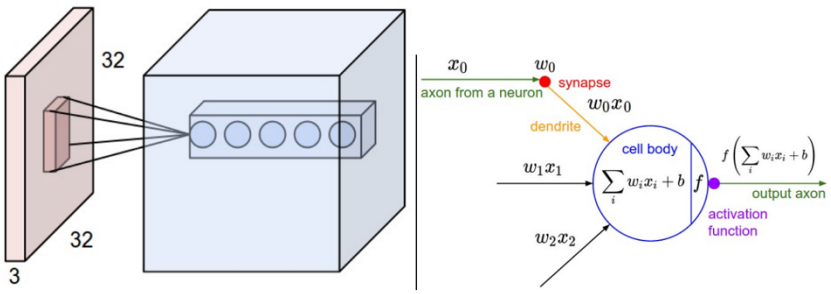


Figure 21. The diagram shows an exemplary input volume in red, and an example volume of neurons in the convolutional layer in blue. Each neuron in the convolutional layer is connected to a small local region in the input volume. Similar to conventional feedforward neural networks, a dot product of neuron weights with the input is computed, followed by a non-linear activation function. The architecture of the convolutional layer restricts the connectivity between neurons to be local spatially. (Image Source: CS231n 2019).

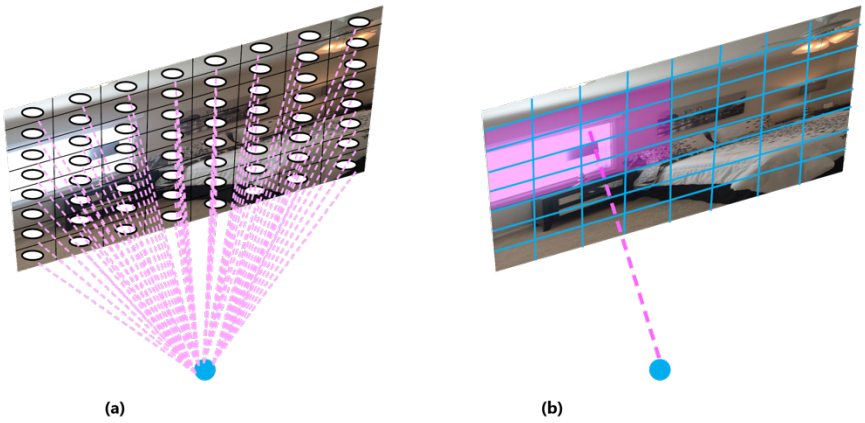


Figure 22. Comparison between conventional fully-connected ANNs and CNNs when dealing with high dimensional inputs such as images. (a) In conventional fully-connected ANNs, each neuron is connected to all neurons (each corresponds to one pixel in the image) in the input layer, resulting in a considerable number of learnable parameters. (b) In CNNs, each neuron will only be connected to a small region of the layer before it, resulting in a reduced number of parameters.

- Pooling layers, generally inserted in-between successive convolutional layers, are used to reduce the dimensions of the representations and network parameters. Max pooling and average pooling are two commonly used strategies. The max pooling is the operation of taking the maximum value of neurons in a small spatial region on the feature maps, while the average pooling takes the mean value. After applying the pooling operation to different regions, a new output feature map with smaller spatial size is generated. The pool operation is purely functional and there are no parameters learned through pooling layers. Similar to convolutional layers, pooling layers are also translation invariant.
- Fully-connected layers perform like conventional ANNs, in which each neuron has full connections to all neurons in the previous layer.

3.4 TRAINING QUALITY

The performance of a neural network depends on how the training is performed. A typical machine learning algorithm contains a dataset, a loss function, an optimization procedure, and a model. The dataset can be divided into non-overlapping training, validation and testing subsets. The validation and testing subsets are used to evaluate the successfulness of the training process by checking the model's performance with unseen data. A loss function measures the model's performance and is applied to all the training, validation and testing sets, resulting in training error, validation error, and testing error. The goal of the training is to minimize the training error and the gap between the training error and the validation/testing error. A model is underfitted if the model is not able to obtain a sufficiently low training error. It means the system does not learn well enough from the training data, and also does poorly when presented with new data. Every model is underfitted when training initially starts. A model is overfitted if the gap between training error and

validation/testing error is large. This means the system learned from the training data well but did not learn to generalize to new data. Whether a model is likely to overfit or underfit can be controlled by altering its capacity, which defines to what degree the model can fit a wide variety of functions. Overfitting can be a problem for models with the high capacity, the model tends to memorize the properties of the training data instead of learning generalized patterns from it (Figure 23). A neural network generally performs best when its capacity is appropriate for the size of the training data and real complexity of the task.

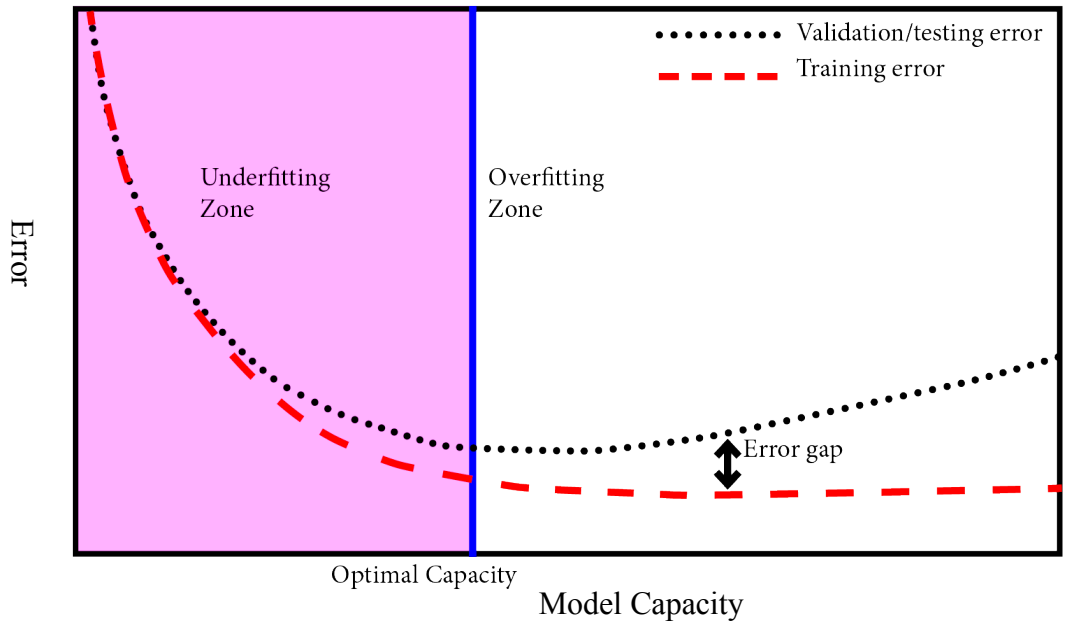


Figure 23. A diagram shows the typical relationship between a model's capacity and error. Initially, the training error and validation/testing error are both high (left end), and when the model capacity is increased, both the training error and validation/testing error decrease, which means the model is in the underfitting zone. Later, when the model capacity is increased to a certain degree, the training error and validation/testing error start to perform differently: while the training error still decreases, the validation/testing error starts to increase. As a result, the gap between training error and validation/testing error expands. Finally, the size of the error gap outweighs the decreases of the training error, which means the model is overfitted. As shown in the diagram, the overfitting and underfitting zones are divided by the optimal capacity of the model.

There are many ways to change a model's capacity and prevent overfitting. The primary strategy is regularization, which is a modification to the learning algorithm that aims to reduce its validation/testing error but not its training error (Goodfellow 2016). The common forms of regularization include early stopping, L^1/L^2 parameter regularization, data augmentation, bagging, dropout, and batch normalization.

- Early stopping is a method that saves the model's parameters with the lowest validation errors. Specifically, early stopping will record a copy of the model's parameters every time when the validation error starts to increase and stop the training algorithm after a certain number (defined by the user) of steps when validation error stops decreasing; even the training error is still decreasing. When the training algorithm stops, the recorded parameters will be returned instead of the latest ones. A rise in validation error (usually marks a rise in testing error as well) means the model is tuned too much to the training data, and the generalization ability is decreased. Therefore, early stopping method balances what can be learned from the training data and the generalization ability of the model. As an effective regularization, the cost of early stopping is not significant. The training time will only be slightly elevated for evaluating the validation error periodically. The cost can be further reduced by using a smaller validation set or by checking the validation error less frequently.
- L^1 and L^2 parameter regularizations are among the simplest and the most common forms of regularization for machine learning algorithms. By using a penalty term, they encourage the sum of the absolute values (L^1) or squares values (L^2) of the weights to be smaller across the entire model (Ng 2004). By controlling the weights of the neural

networks, these regularizations raise the chance that a simpler and hence a more generalizable solution will be selected, and still retaining a low training error.

- Data augmentation is a straightforward regularization strategy involves creating artificial data and adding it to the training set. The best way to make a machine learning algorithm more generalizable is to train it with more data. However, in practice, it is common that only a limited amount of data is available. The strategy of creating artificial data in addition to the original training data, if used appropriately, will make the algorithm generalize better, without introducing additional collecting and labeling costs. Commonly used data augmentation methods include image cropping, scaling, rotation, and gamma corrections. However, these transformations must be applied only in the case that they will not change the correct labels of the data set. For example, when trying to predict luminance maps (labels) from input features, it is not appropriate to augment the data by applying different gamma corrections to the original luminance maps because it will change the correct luminance values.
- Bagging (Breiman 1994), short for bootstrap aggregating, is a technique that reduces validation/testing error by combining several models. Specifically, bagging involves training multiple models separately and then combing their results by weighting their votes on the test dataset. The idea stems from a general machine learning strategy called model averaging, or ensemble methods, which is based on the fact that different model will usually not have the same errors on the test set. Although effective, bagging is an expensive technique, especially when each model is a large neural network. The reduced generalization error is at the cost of increased computation and memory.

- Dropout (Srivastava et al. 2014) is a computationally inexpensive approximation of bagging regularization. During each forward pass of a training cycle, the dropout technique will randomly drop a subset of neurons from the preceding layer and remove the weights associated with them. These selected neurons will not participate in either the forward or backward pass of this cycle. When a cycle is finished, and the weights are updated, the dropped neurons will be restored. Dropout is similar to bagging in a way that the network with partial neurons in each training cycle can be seen as a subnetwork. Without dropout, there is a chance that some neurons in the network are tuned to capture one specific type of information while ignoring other important data. With dropout, neurons are being randomly omitted, and therefore each subnetwork is forced to learn generalized patterns. Dropout is less effective when the dataset is very large or when only a few training samples are available (Goodfellow 2016).
- Batch normalization (Ioffe and Szegedy 2015) is a method of adaptive reparameterization, usually utilized for training very deep models. By standardizing the mean and variance of each computing layer of the network, batch normalization stabilizes the learning process and makes the learning significantly easier and more efficient.

3.5 MACHINE LEARNING RELATED LIGHTING STUDIES

Machine learning has been recently used a few times within the architectural lighting field. Researchers have investigated machine learning, and specifically neural networks, for indoor illuminance predictions, with wide applications in electric lighting energy estimations, evaluations of the effectiveness of the artificial lighting system, and optimizations of the shading systems. Conventional feedforward neural networks are utilized to estimate illuminance levels under

daylight (Kazanasmaz et al. 2009, Ahmad et al. 2017) and illuminance distributions under artificial lighting (Sahin et al. 2016) from input parameters related to light sources and room settings. Though inspiring, illuminance predictions in general are less computationally intensive compared with luminance predictions due to lower resolutions of calculation points, and lack of dependency on view directions. Annual illuminance predictions can be achieved through standard simulation practices without significant simulation time. To the best of the authors' knowledge, there is no previous architectural lighting study that utilizes machine learning to predict long-term luminance distributions of interiors.

While only few previous studies existed in architectural lighting field, which utilizes machine learning for lighting predictions, the technique has been increasingly employed in closely related problems, such as rendering and appearance synthesis, in computer vision and graphics. Machine learning has been investigated for image de-noising (Bako et al. 2017, Chaitanya et al. 2017). By creating high-quality images from noisy renderings with reduced sample rate, the method has great potentials in accelerating physically-based rendering processes. More related works involve rendering a scene under novel lighting conditions, which has been a long-lasting research question in computer vision and graphics field with applications in virtual reality, augmented reality, and visual effects. Recently, there is a trend of using learning-based models to tackle this long-studied problem.

The methodology employed in this dissertation research is inspired by Ren et al. (2015). In their study, a neural network model is used to learn the non-linear mapping of pixel-scale RGB intensities from local and contextual attributes of surface points. The model is trained with a small subset (hundreds) of images of the scene to create images under new lighting conditions. The developed method has been tested with HDR captures of the indoor scenes lit by an artificial light

source with impressive results. However, their method cannot be directly applied into architectural daylighting predictions because: 1) The illumination conditions caused by the natural varying sun and sky are more complicated than by a single point light source moved inside a 3D volume used in their study. The possible luminance range of a daylit scene will be orders of magnitude higher than of the scene lit by the artificial lighting. 2) Their method has a high computing requirement. They use an ensemble model created via a hierarchical clustering mechanism, which needs to be trained on a CPU cluster. More recently, Gardner et al. (2017) use two individual deep neural networks to predict the positions and then the intensities of light sources of indoor scenes, from a single LDR photograph. The estimated light sources are used to relight the virtual objects into the scene and composited into photographs. However, with the limited input information, this study aims at producing results with plausible visual appearance instead of physical accuracy. The light sources are blurred, and their intensities are predicted less accurately than their locations. In 2018, Xu et al. develop a neural network model to synthesize scene appearance under novel illuminations from five images. The model is trained on a large synthetic dataset consist of objects with various shapes and reflectance to learn a complex, non-linear function across scenes. By leveraging a single deep neural network model than several shallow neural networks, the computing requirement is reduced. This method is limited to relighting the objects rather than scenes and does not model the interreflections and cast shadows present in real images. Their study has a limited applicable luminance range, less than the range required to cover daylit scenes.

Machine learning has also been used to approximate high luminance range sky models from images. Satilmis et al. (2016) developed a neural network model to predict HDR skies from a sparse set of HDR captures of various sun positions and trained the network with illumination/image pairs. After training, given the parameters of the sun position, view direction,

and the sky turbidity level, the model can successfully predict an arbitrary HDR sky (exclude sun pixels). Hold-Geoffroy et al. (2017) proposed a deep learning-based method to predict outdoor illumination represented by sky models. In their study, a CNN model is trained to predict the three parameters for reconstructing the Hošek-Wilkie sky model (2012) from a single LDR image. Though all of these studies are encouraging for estimating the illumination of high luminance range, they only focus on the predictions of outdoor skies instead of the entire luminance environment of the daylit scenes. Skies can be well approximated with the low-dimensional analytical model. Indoor scenes involve light interreflections between surface which are more complex to approximate.

Chapter 4. PREDICTING PERSPECTIVE LUMINANCE MAPS

In this dissertation research, a novel end-to-end method is provided to generate annual luminance maps of indoor space from a subset of HDR images using DNNs. This dissertation documents the development of the methodology in three phases, where each phase builds upon the outcome of the previous one to refine the methodology. These phases are discussed in Chapters 4, 5, and 6. Chapter 4 introduces the study performed to predict annual luminance maps of single perspective views. This initial phase is used to build the research framework and to test the workflow for its applicability and accuracy.

Figure 24 gives an overview of the research framework with details to be introduced in this chapter. The sections of the chapter are structured as follows: Section 4.1 focuses on light transport model, which is the core concept utilized in developing the methodology. Section 4.2 and 4.3 describe the input and output of the developed DNN model, and how the input and output data is generated and processed. The architecture of the DNN model is discussed in Section 4.4, and the training strategy is discussed in Section 4.5. Section 4.6 documents the results and analysis of the study. The major findings of the study are outlined in Section 4.7, followed with a critical analysis of the methodology.

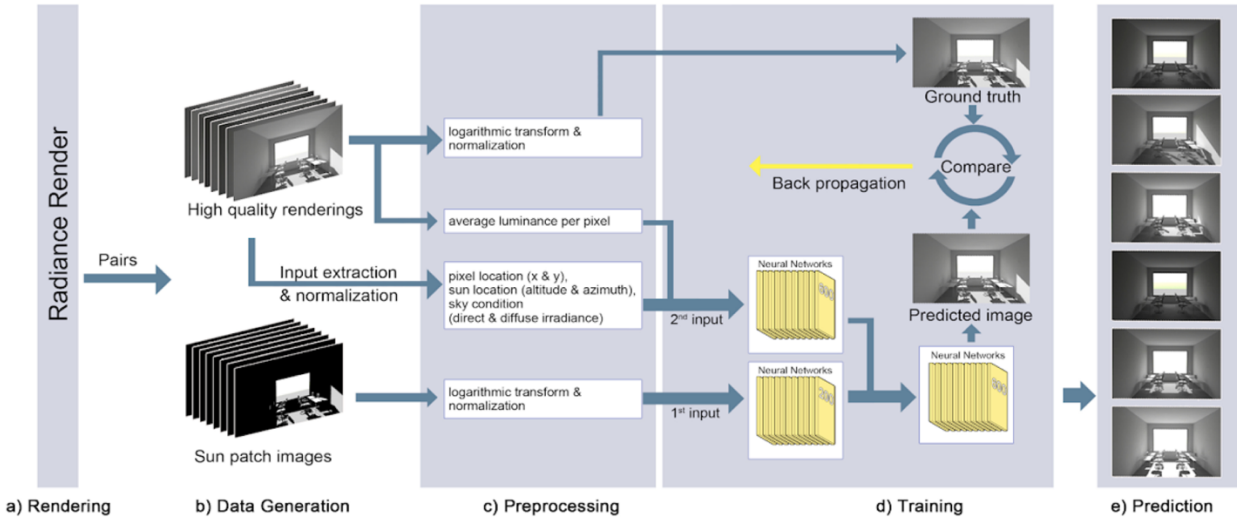


Figure 24. Overview of the research framework: a) Radiance simulation engine renders a small number of HDR images, which are used to train the DNN model. b) Pairs of high-quality interior maps and sun patches are generated. The interior maps are used as ground truth for training the DNN, while the sun patches are used as input to the DNN. c) The sun patches and simulation parameters are combined to create input for the DNN. d) The network is trained so that it accurately predicts the interior maps from the input. e) Once trained, the network can be used to generate interior maps from novel sun patches and simulation parameters.

4.1 Light Transport Model

The light transport model describes how light travels in a scene. In traditional physically-based renderings, a light transport model is computed using the complete scene information including the light sources, geometries, and materials. In this study, the light transport model is inferred from a series of luminance maps generated under various illumination conditions (where,).

$$\boldsymbol{M} = \boldsymbol{L} \times \boldsymbol{I}$$

where,

$\boldsymbol{M}$ is the light transport model;

$\boldsymbol{L}$ is the luminance of image pixels; and

$\boldsymbol{I}$ is a feature vector which describes the illumination.

Equation 5. Light transport model

A DNN model $\boldsymbol{M_{NN}}$ is used to estimate the light transport model ($\boldsymbol{M}$) through learning the non-linear relationship between illumination conditions (input $\boldsymbol{I}$) and the corresponding pixel luminance values (output $\boldsymbol{L}$). The estimated light transport model ($\boldsymbol{M}$) allows the scene to be rendered under new lighting conditions (Figure 25). Using this method, generating annual luminance maps becomes a three-step process that involves: 1) acquiring sparse samples of luminance maps; 2) estimating the light transport model from the sparse samples using the DNN, and; 3) predicting annual luminance maps using the estimated light transport model. The total calculation/acquisition time of annual luminance maps is significantly reduced compared to the traditional annual simulation/measurement method. Instead of rendering/capturing luminance maps for the entire year, we only need to render/capture a small set of them.

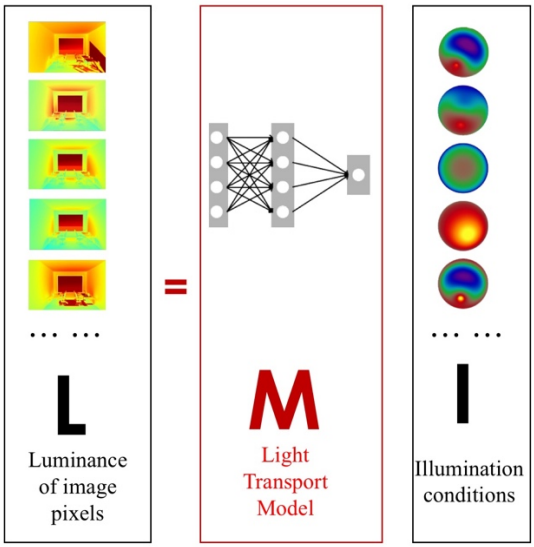


Figure 25. After the light transport model M is estimated, it can be used to generate luminance maps under novel lighting conditions

4.2 Input and Output of the DNN Model

The first task in designing the DNN model M_{NN} is to designate the input and output parameters. The illumination condition (input I) is designed to reflect the sun and sky conditions. Specifically, the illumination condition is parameterized with the following features: 1) the sun location defined by sun altitude and azimuth angles (*al, az*); 2) the sky condition defined by direct and diffuse irradiances (*dir, dif*); and 3) the pixel location defined by the x and y coordinates (*px, py*). These six parameters describe the luminous environment within an indoor scene at a given date and time. Direct and diffuse irradiances are the essential parameters required by Perez all-weather sky models (Perez et al. 1993). The input parameters are augmented with the average luminance of a pixel (*avg*) to improve convergence (Ren et al. 2015). Initial experiments show that adding the pixel luminance of the sun patches (*sunpatch*) aid the neural network in reconstructing sharp shadows and sun penetrations. More details about the advantages of adding sun patches as input are explained in Section 4.4. The final input feature vector includes eight parameters I = (*al, az,*

dir, dif, px, py, avg, sunpatch). The output vector *L* is the pixel luminance under the given lighting condition.

4.3 Data Generation and Processing

High-quality HDR renderings are utilized to train and evaluate the neural network to develop the methodology under controlled settings. The test room for the simulation is set to be an open plan office with basic furniture arrangement and a south facing side window, located in Seattle (47.6°N, 122.3°W) (Figure 26). The dimension of the room is 6m (width) by 14m (depth) and 4.5m (height). The camera is set to be in the front center of the room and 2.6m from the floor. The materials for walls, ceiling, floors, and outdoor ground have reflectance values of 50%, 80%, 20%, and 20% respectively.

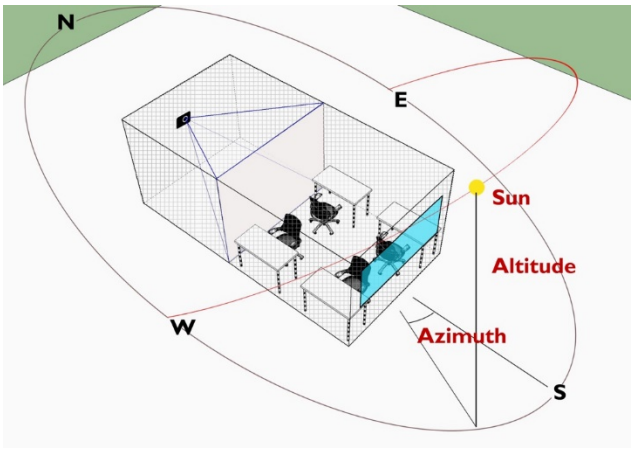


Figure 26. Room model used in Radiance simulations

The position of the sun is defined with solar altitude and azimuth angle. Perez all-weather sky model is used to generate a description of the sky with the direct and diffuse irradiances, extracted from Seattle EnergyPlus weather file. Two sets of images are generated using the Radiance RPICT method: the interior maps with four ambient bounces (-ab 4) and sun patches with zero ambient

bounce (-ab 0). These two sets of images share the other rendering parameters (-ps 2 -pt .05 -pj .9 -dj .7 -ds .15 -dt .05 -dc .75 -dr 3 -st .15 -aa .1 -ar 512 -ad 2048 -as 1024 -lr 8 -lw .005). 4379 pairs of high-quality interior maps and sun patches of the scene are generated at one-hour intervals from sunrise to sunset for the entire year. All images are rendered as perspectives from a single viewpoint with a pixel resolution of 1392×920 and downsized to 348×230 using Radiance pfilt for anti-aliasing. Generating such a large dataset is necessary for developing and testing the machine learning model. When applying the workflow into practice, the user only needs to provide a small amount of HDR images.

Eight images are constructed with the same resolution as the input interior map to provide px, py, al, az, dir, dif, avg, and $sunpatch$ as an input in each channel respectively. Stacking the N samples leads to an 8N-channel input. The generated interior maps are paired with the input to create N input-output-pairs; then pre-processed for the neural network.

Data preprocessing is an important step. Initially, all features are normalized to the range of [0, 1] using the scaling function: x = $\frac{x - x_{max}}{x_{\max} - x_{min}}$, where x_{max} and x_{min} represent the highest and lowest value of feature x. Feature scaling tends to improve the training process by ensuring that all features are in the same scale which: 1) makes each feature to contribute approximately proportionally to the learning process, and 2) makes the computation (i.e., time for gradient descent optimization to converge) faster (LeCun et al. 1998). However, due to the nature of high dynamic range renderings, the luminance values of the scene can span several orders of magnitude (e.g., the brightest part of the solar corona compared to the shadows in the room) and are not evenly distributed throughout the range. After the feature scaling, the majority of these values are very low (Figure 27). To solve this problem, a gamma-correction of 2.2 is employed to spread the range

of luminance more evenly over the range of [0,1] (Figure 27). When the training is completed, a reverse transform is applied to reconstruct the high dynamic range luminance maps.

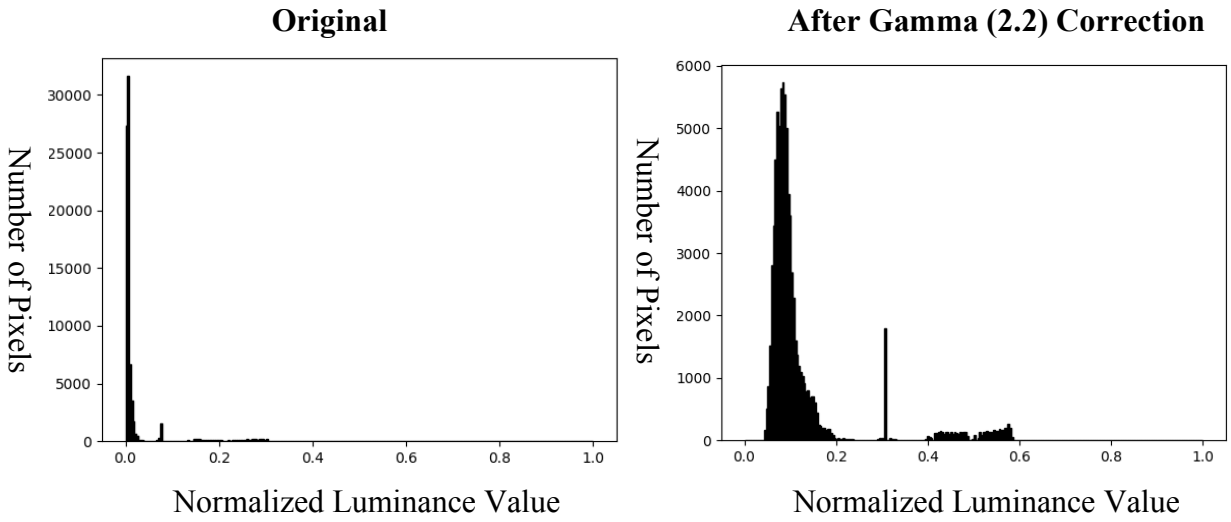


Figure 27. Diagrams show the distribution of the normalized luminance values of (Left) the original interior maps and (Right) the gamma corrected (gamma = 2.2) interior maps over the range of [0, 1]. The images contrast is reduced with the gamma-correction resulting in more even distribution.

4.4 DNN ARCHITECTURE

When designing the DNN architecture, one consideration is choosing the number of trainable parameters in the network. A DNN model with more trainable parameters may provide more computational power during the training process, but it requires more input data to avoid overfitting (Turmon and Fine 1994). The proposed model is designed with a comparable number of learnable parameters as the model described in Ren et al. (2015). The model differs from Ren's in that it utilizes a single DNN rather than an ensemble model created via a hierarchical clustering mechanism. Initial tests show that the model achieves the same accuracy as Ren et al., and can be trained in hours on a single machine rather than CPU cluster nodes. The DNN model is created with five hidden layers, each comprising of 600 nodes. In each layer, the network applies a linear

convolution to the output of the previous layer, and then applies a nonlinear activation function. ReLU activation function is selected, as it performs better than other activation functions (e.g., sigmoid and tanh) in the tests.

The model was initially designed with the feature vector composed of 7 inputs (*px, py, al, az, dir, dif, avg*). The sun patches (*sunpatch*) were not included. However, experiments show that this model can capture the smoothly varying illumination, but it fails to adequately capture the sharp shadows and sun penetrations caused by the direct sunlight (Figure 28 (b)(d)). Although this model can be improved by using more high-quality rendering samples that cover more sun positions, this defeats the purpose of training with a limited number of high-quality renderings. Therefore, a refined method is proposed which differentiates the direct and diffuse contribution from the sun and sky to improve the result. The pixel luminance of sun patches (*sunpatch*) is added as a second feature vector. Sun patches with 0 ambient bounce are rendered with contribution only from the direct sunlight; the contribution from the sky or the interreflections in the scene are not included. Compared with the high-quality renderings, quick sun patch renderings have two advantages: 1) shorter simulation time, they can be instantly calculated, and 2) they can be easily generated using a crude model constructed from either HDR photographs or basic CAD models. This information can aid the neural network in modeling the light transport of scenes with sharp shadows and sun penetrations, as shown in Figure 28 (c)(e).

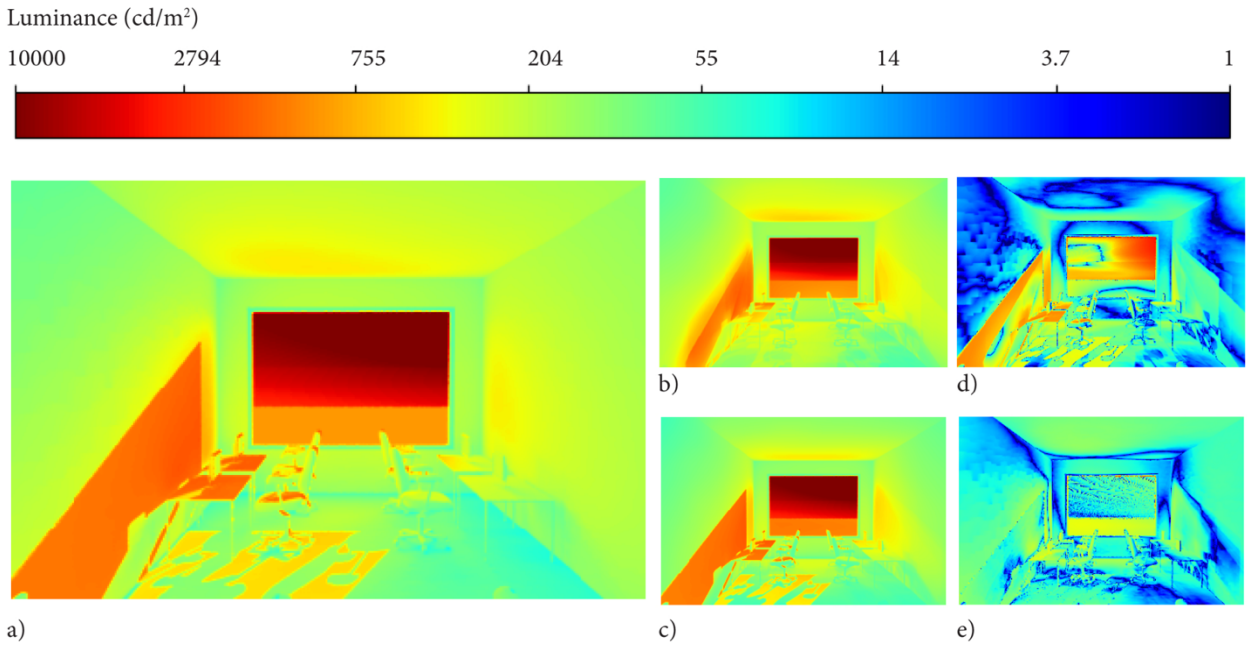


Figure 28. An example illustrates the comparison of the predictions made by the initial DNN model and the refined DNN model. The ground truth luminance map (a) is poorly predicted as in (b) with blurry shadows and sun penetrations by the initial DNN model. With the addition of quick sun patch renderings, the refined model generates more accurate results as in (c). (d) and (e) are errors maps for (b) and (c).

The refined network architecture (Figure 29) divides the input features into two separate vectors to a two-input CNN: The first input branch, which encodes *px, py, al, az, dir, dif,* and *avg*, consists of four convolutional layers, each with 600 filters. The second input branch, which encodes *sunpatch*, consists of a single 200 filters convolutional layer. The first input branch encodes parameters describing global illumination, while the second input branch encodes direct sun illumination. The two branches are concatenated to form an 800 filters layer, which is followed by a convolutional layer of 600 filters. The output layer is a convolutional layer with one filter that predicts luminance values. The ReLU activation functions are applied to all layers. All convolutional layers use the same kernel size of 1 and stride of 1. The kernel size was increased to

3×3, but resulted in lower accuracy. This model yields satisfactory results in this study, but will be further investigated in Chapter 5.

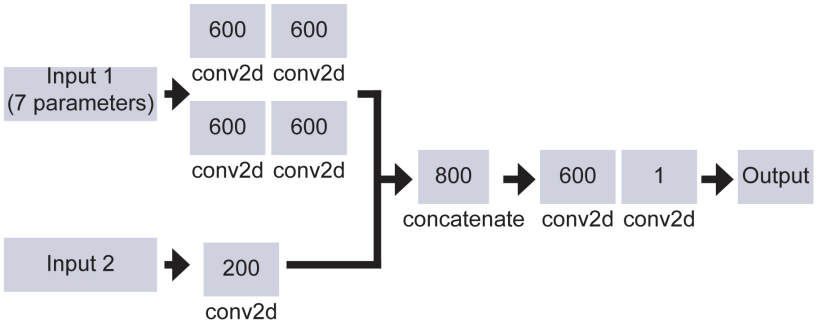


Figure 29. The refined DNN architecture. After a series of 2-dimensional convolutional layers, a concatenate layer merges two heads of input: one for regressing the sun patches (Input2) and the other for the location, sun and sky parameters (Input1). The number indicates the amount of filters in each layer (e.g., 600 conv2d).

4.5 TRAINING AND TESTING

The generated images in the database are divided into three groups: training, validation and test groups. The images in the training group are the input to feed into the neural network model. During the training process, only the model parameters with the improved performances on validation sets are saved. The test group images are used to evaluate the performance of the method. There is no overlap between the test group and the training/validation group. Hence, all tests are performed for lighting conditions that have not been seen by the network before. A proper method for selecting the training samples is important and will be discussed at length later in Section 4.6.

The next step is to train the developed DNN model with the processed training data. The main idea of training the neural network is to minimize the loss calculated by the loss function based on the difference between the predicted image and the ground truth image (high-quality interior

maps). The backpropagation algorithm is used to train the network via examples by repeatedly updating the neural network until the loss is minimized.

Choice of loss functions greatly impacts DNNs and their learning dynamics. Several loss functions are tested and a combination of mean square error $\mathcal{L}_{MSE}(y,t)$ (Equation 6) and relative error rate $\mathcal{L}_{RER}(y,t)$ (Equation 7) is selected: $\mathcal{L}\text{oss}(y,t) = \mathcal{L}_{MSE}(y,t) + 0.01 \times \mathcal{L}_{RER}(y,t)$, where y is the network prediction and t is the ground truth.

$$\mathcal{L}_{MSE}(y,t) = \frac{1}{n}\sum_{i=1}^{n}(y_i - t_i)^2$$

Equation 6. Mean Square Error (MSE) loss function

$$\mathcal{L}_{RER}(y,t) = \sqrt{\frac{\sum_i (y_i - t_i)^2}{\sum_i {t_i}^2}}$$

Equation 7. Relative Error Rate (RER) loss function

where,

y is the network prediction;

t is the ground truth;

n is the total amount of calculated samples; and

i is each sample in the dataset.

The neural network is trained with the ADAM optimizer (Kingma and Ba, 2014) with an initial learning rate of 10^{-3}, reducing the learning rate by a factor of 2 once learning stagnates, and an initial batch size of 4, reducing the batch size to 2 after 25 epochs, or if learning rate is reduced to a minimum threshold of 10^{-10}. The data sequence is shuffled once after each epoch to prevent the model from being influenced by the sequence of the data. The training stops when the

validation loss reaches a minimum threshold (10^{-10}) or the batch size is reduced to 1. During the training process, an early stopping regularization technique is applied to only save the model with a reduced error on the validation set. The training process of one data subset takes about 2 hours on NVIDIA Tesla K80 GPU.

Once the neural network is trained using a limited number of samples from the training group, an approximation of the light transport model $\boldsymbol{M_{NN}}$ is calculated. The model can then be tested on images under new illumination conditions, which have not been seen by the network before. The predicted interior maps are compared with the interior maps rendered by Radiance using per-pixel errors measured by $\mathcal{L}_{MSE}$ (Equation 6) and $\mathcal{L}_{RER}$ (Equation 7) used in the loss function, and evaluated using falsecolor analysis with image subtraction operations.

The visual comfort indicator DGP is added to evaluate how the results can influence lighting design decisions within the context of discomfort glare assessment. In DGP analysis, luminance maps are processed to determine the potential glare sources, and subsequently, to determine the percentage of the population who would find the scene as glary. DGP is calculated from the ground truth and the predicted images. Although these images are not 180° fisheye images that provide the best practices for glare assessment, both images have the same limited field of view. Therefore, DGP can be reliably used to compare one image to the other.

4.6 Results and Analysis

A sensitivity analysis is performed to understand how the number of training samples influences the prediction accuracy, and to find the minimum sample size. It uses the training sets that contain 50, 100, 200, 500, and 1000 images (out of 4379 images) and the test set with 500 images. The validation set contains 10% randomly selected images from the training set.

A K-means clustering algorithm (i.e., partitioning the data into k relatively equally distributed groups) is utilized to select the training samples that best represent the entire dataset and cover various sun positions and sky conditions (Figure 30). Random data sampling from the entire dataset biases the samples towards the most common sky conditions and sun positions. Sampling from the K-means clusters more evenly distributes the training data over the entire domain. The test sets contain 500 samples randomly selected from the test group.

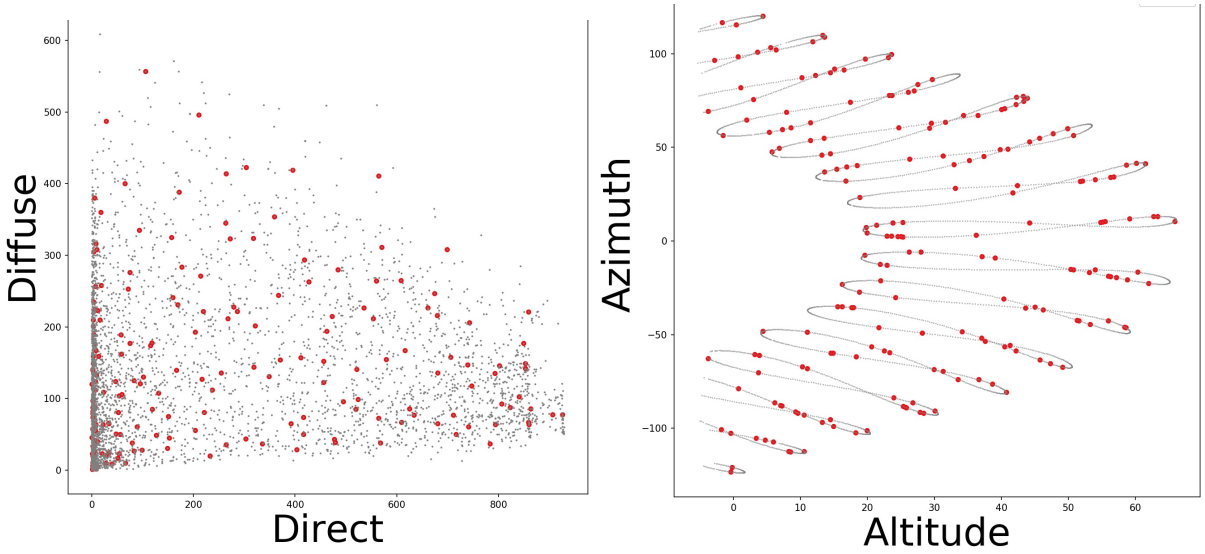


Figure 30. The K-means method is utilized to make the selected training samples well-distributed over the 4-dimensional light domain visualized in two 2-dimensional plots: (Left) distribution plot over sky condition parameters (direct and diffuse irradiances), and (Right) over sun position parameters (azimuth and altitude angle). Red dots represent the selected training samples while grey dots represent all data points. In this example, k= 200.

Figure 31 shows the result error curves with respect to the number of training samples. The predicted image error, measured by $\mathcal{L}_{MSE}$ and $\mathcal{L}_{RER}$ with normalized luminance values, decreases with the number of training samples. However, the error curves decrease slowly after the number of images reach 200. Therefore, 200 images are selected as the optimum sample size with

predictions $\mathcal{L}_{MSE}$ of 1e-05 and $\mathcal{L}_{RER}$ of 0.063 using a three-fold cross-validation. This demonstrates that using 5% (200 out of 4379) of high-quality renderings, the model can predict a luminance map of any time of the year with an average 93% accuracy.

A reverse transformation of preprocessing is then applied to both the predicted and the ground truth images to calculate DGPs. The results, as presented in Figure 32, show that DGPs evaluated using predicted images closely match those using ground truth images. The average absolute error of DGPs is 2.6e-08. Note that the maximum range of DGPs is below 0.4. Further investigations found that DGPs did not exceed 0.4 over the entire database. This is due to the high viewing position (2.6m from the floor) when generating the image data, resulting in a relatively low luminance range (0-3×10^4) of the dataset.

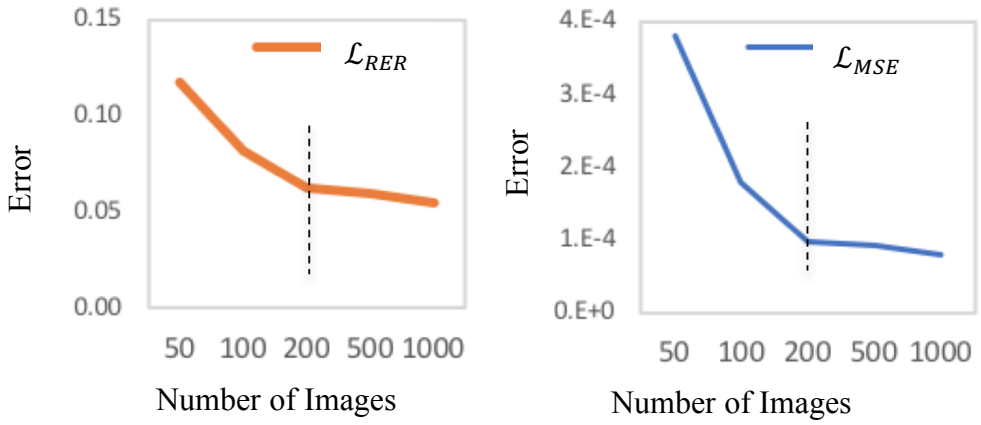


Figure 31. Sensitivity tests show the relationship between the number of input images and the testing errors evaluated by and $\mathcal{L}_{RER}$ (Left) and $\mathcal{L}_{MSE}$ (Right). The errors are all measured with normalized luminance values. The dashed lines indicate an optimal sample size of 200.

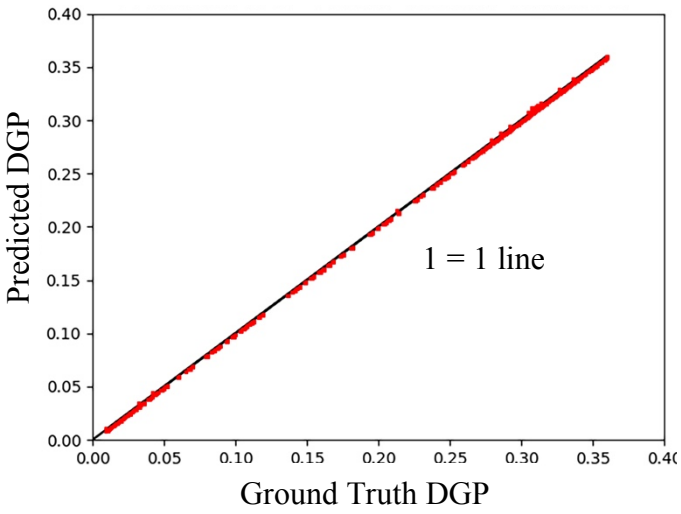


Figure 32. Scatterplot comparison of DGP values calculated using ground truth and predicted images. Red dots represent DGP pairs.

To better illustrate the results, Figure 33 shows several examples of different sky conditions (clear sky, intermediate sky, and cloudy sky) from the test set of this study. Predicted images are compared to ground truth images, with error maps illustrating the absolute differences. All images are shown in falsecolor with a logarithmic scale, and each image is labeled with MSE, RER, and DGP values. The results show that: 1) The predicted images are imperceptible from ground truth images in falsecolor without the help of error maps. This indicates that the proposed model is able to make predictions that will lead to the same design decisions of those real high-quality renderings, even with a small number of input samples (50 images out of 4379 images). 2) Similarly, DGPs of predicted images closely match those of ground truth images. This leads to a similar conclusion that the same design decision can be made with predictions based on small training samples. 3) As clearly shown in error maps, the differences decrease, when the number of images increases. This matches the conclusion of the previous sensitivity analysis. 4) Among three sky conditions, predicted images of a clear sky with high direct irradiance have the highest errors. 5) While the estimated result looks almost indistinguishable from the ground truth, the hardest part

of the prediction is the outdoor ground plane seen through the window. This planar geometry is not the Radiance generated ground hemisphere, which can be more easily predicted.

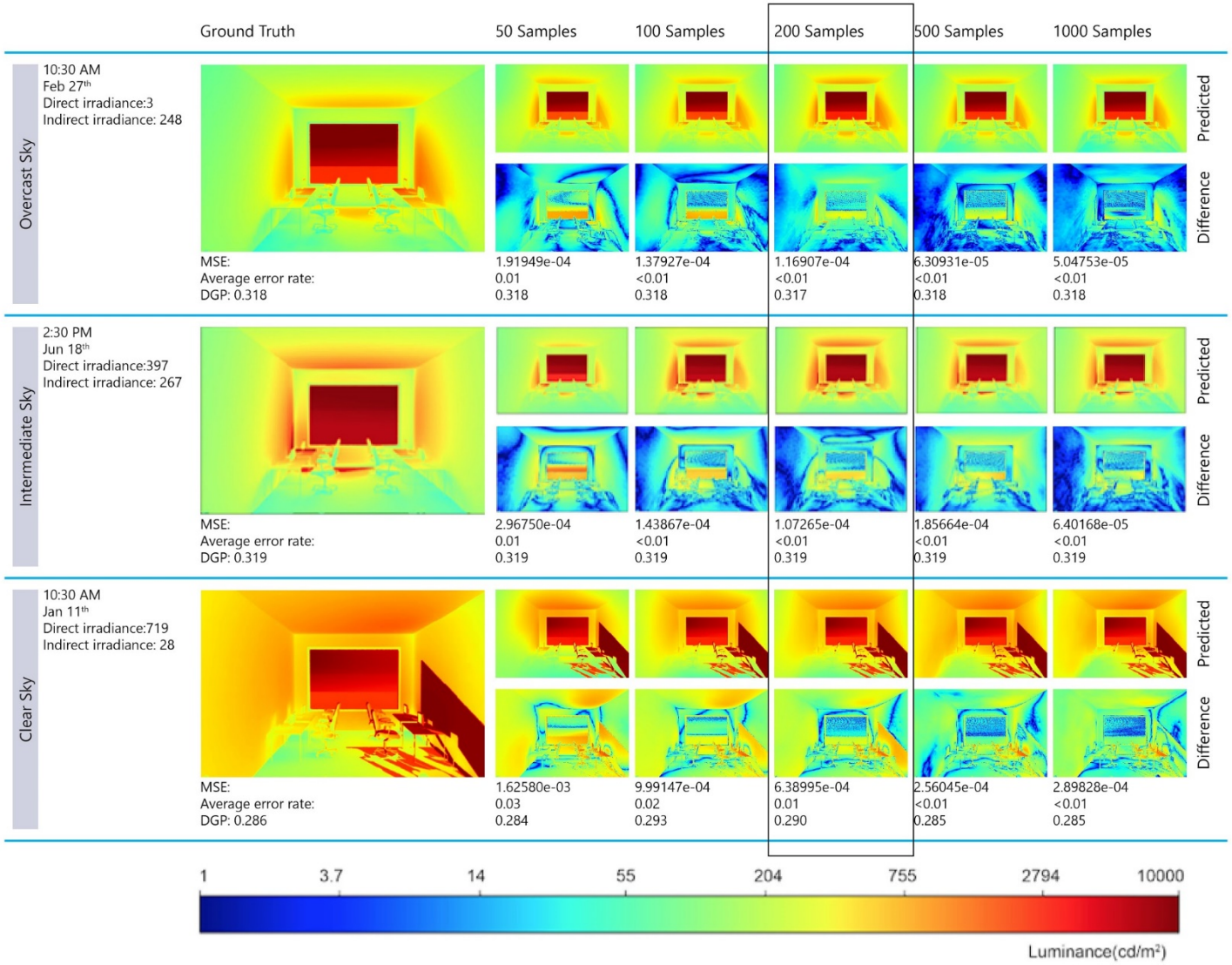


Figure 33. Sensitivity experiment results. Seven exemplary test cases with different sky types are displayed in falsecolor with logarithmic scale. The bounding box highlights an optimal sample size of 200.

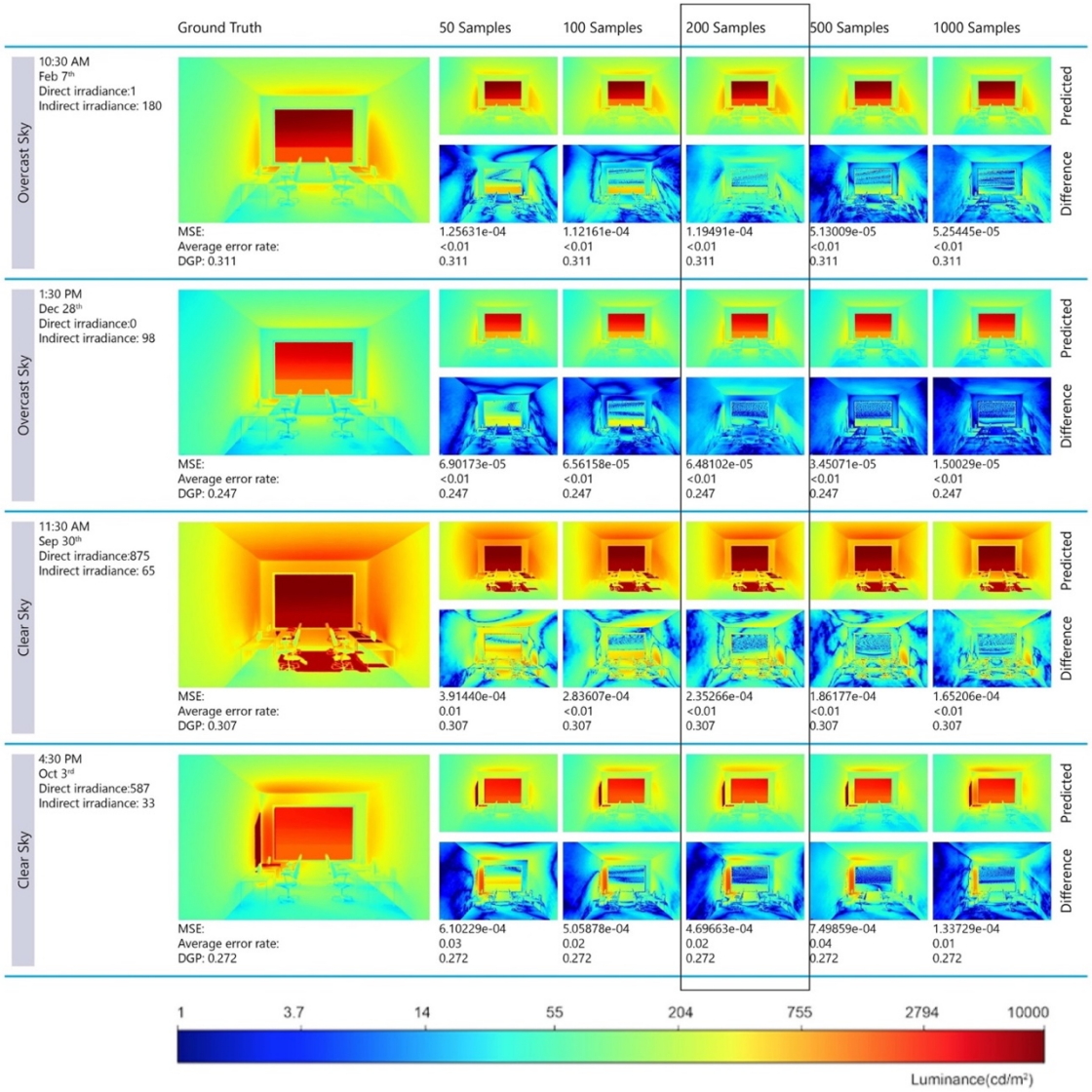


Figure 33. Sensitivity experiment results. Seven exemplary test cases with different sky types are displayed in falsecolor with logarithmic scale. The bounding box highlights an optimal sample size of 200.

4.7 Discussion and Next Steps

This study demonstrates that DNNs can be applied to solve complicated physically-based rendering problems. Specifically, it shows that a DNN approach is able to learn and approximate the fundamental, underlying non-linear relationships between the image pixel luminance and daylighting conditions without overfitting, and generate high-quality renderings. The study shows that by using a DNN model, it is possible to predict annual perspective luminance maps by using 5% high-quality renderings evenly sampled over the year as input and reach a comparable accuracy in the test set. This method can be used as an alternative to accelerate annual luminance-based simulations.

However, when applying the method to HDR captures, the collection period of 1 year is not practical, and the performed tests show that the current DNN model cannot make acceptable predictions with short-term data (e.g., one month). Further developments are required to shorten the minimum data collection period. Based on the lessons learned, several strategies are identified. Providing the DNN model with more information about the scene is identified as a viable strategy. As the current input images have a limited field of view, they may not include all the light sources (e.g., windows) and furniture of the scene. Therefore, panoramas (with a 360° horizontal field of view and 180° vertical field of view) will be tested as inputs (Figure 34) to provide more scene information during the training process. The generated panoramic luminance maps also have the advantages of offering more information about an occupant's visual experience for further analysis. The sky condition significantly affects the prediction accuracy (Figure 33). Augmenting the input data with sky luminance maps would provide the DNN model with more spatial information about the sky luminance distributions.

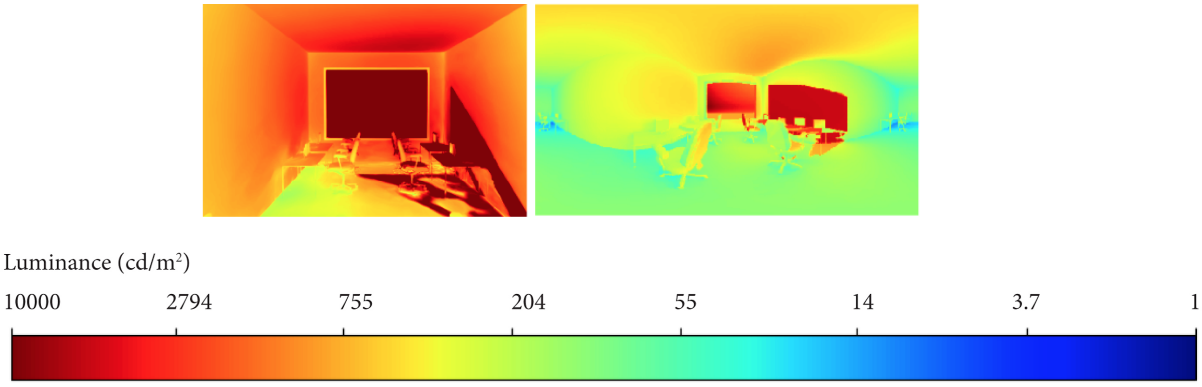


Figure 34. A sample input image (left) used in this study with a limited field of view, and (right) the same scene with a 360° horizontal and 180° vertical field of view (spherical panoramas).

Another lesson learned is that the image data should be generated (e.g., through adjusting the viewing position and direction) to have higher luminance range in order to understand the performance of the method on predicting glary scenes. It will also allow for more meaningful DGP comparisons. However, it is observed from the study that the higher luminance variance of the images leads to higher errors. Even after the normalization and gamma correction to spread the range of luminance more evenly over [0, 1], the pixels from the input images still show great variances in luminance (e.g., the images are generally dark with a few bright pixels around the window in the scene). The high variations of the input samples will lower the convergences speed of the gradient descent algorithm (LeCun et al. 1998). Therefore, the distribution of the input data luminance values will be further analyzed in order to formulate a more appropriate loss function, along with a new preprocessing method.

The study also serves to provide a better understanding of the calculation time when applying the methodology. The method is scene specific, which means the machine learning process needs to be performed whenever the method is applied to a new scene. Therefore, training time has a significant impact on the feasibility of the method. In this study, the average calculation time is a

few hours. It is desirable to further reduce the training time. DNN hyper-parameters, such as the parameters of the neural network architecture, have a great impact on the prediction efficiency and accuracy. They are not automatically optimized by the training algorithm, they will be fine-tuned in the next step.

The following two chapters are based on the lessons learned from this initial study. In Chapter 5, the prediction of annual panoramic luminance maps is demonstrated with a refined methodology. To understand the minimum data (captures / renderings) requirements for generating annual luminance maps, a thorough sensitivity analysis is performed and presented in Chapter 6.

Chapter 5. PREDICTING PANORAMIC LUMINANCE MAPS

Chapter 5 introduces a methodology to predict annual panoramic luminance maps (with 360° horizontal and 180° vertical field of view) from a small subset of HDR imagery. Unlike the perspective or fisheye projections that are commonly used in daylighting evaluations, panoramas allow full degree-of-freedom in camera roll, pitch, and yaw that can account for an occupant's changing view and gaze direction in a given environment. There is an increasing demand for panoramic luminance maps, as recent studies in luminance-based evaluations integrate subject's visual experience over multiple viewing directions (Jakubiec and Reinhart 2012, Rockcastle et al. 2017). However, generating panoramic luminance maps is more time-consuming than single fixed-view perspective maps.

The machine learning methodology is built upon the algorithm developed in Chapter 4, it is improved through multiple refinements. Figure 35 gives an overview of the revised methodology. The workflow starts with the sample data generation process using the Radiance rendering engine (Ward, 1994). For each daylighting condition, the proposed system uses Radiance to generate a set of three HDR panoramic images consisting of 1) high-quality interior maps, 2) sky maps, and 3) low-quality sun patches. The interior maps are used as ground truth, and the sky maps and sun patches are input to the neural network training. After the training is finished, the light transport model is estimated and can be used to predict luminance maps under new lighting conditions.

Each section of this chapter introduces and describes the methodology in detail. The design of the network's input and output parameters is discussed in Section 5.1. Sections 5.2 and 5.3 show the data generation and processing strategies. Section 5.4 and 5.5 present the design of the network architecture and the nuts and bolts of the training process. After the training process, the

predicted panoramas are validated against Radiance RPICT renderings using a series of quantitative and qualitative metrics, and the results are presented in Section 5.6. Finally, the lessons learned from the study are introduced in Section 5.7, followed with discussions about the next steps.

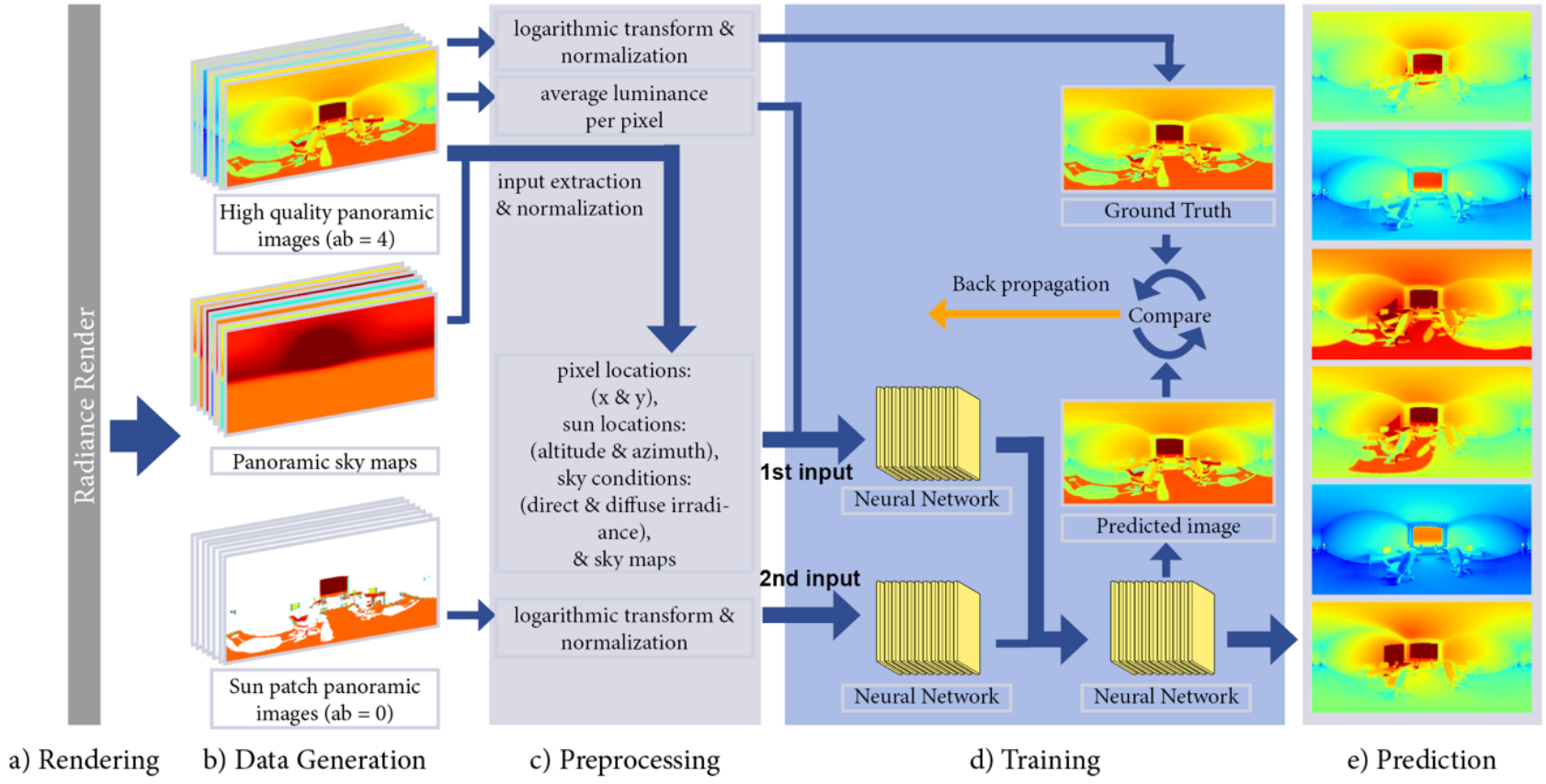


Figure 35. Overview of the workflow: a) A sparse data set is rendered using the Radiance simulation engine. b) Each data point includes three panoramas consisting of interior maps, sky maps, and sun patches. The interior maps are used as ground truth for the DNN training process, while the sky maps and the sun patches are used as the input to the DNN. c) All images are pre-processed and normalized. The sun patches and simulation parameters are combined to create the input for the DNN. d) The network is trained to predict the interior maps. e) The trained network can then be used to generate interior maps from novel sky maps, sun patches, and simulation parameters. Further long-term luminance-based analysis (e.g., glare and spatial contrast analysis) can be performed using these predicted interior panoramas.

5.1 INPUT AND OUTPUT

A DNN model is utilized to approximate the light transport model as a relationship between illumination conditions and the corresponding pixel luminance values. The input vector of the DNN model is constructed from several parameters describing the illumination conditions, and the output vector contains a sole value which is the pixel luminance in the given luminous environment. The input parameters include: 1) the location of the sun, defined by sun altitude and azimuth (*al, az*); 2) the luminance distribution of the sky, defined by direct irradiance and diffuse irradiance (*dir, dif*); 3) the location of the pixel defined by x and y coordinates (*px, py*); 4) the average luminance value for an image pixel (*avg*); and 5) the pixel luminance of the sun patches (*sunpatch*), to aid the neural network in reconstructing sharp shadows and sun penetrations. These eight input parameters are consistent with the previous study. Experiments show that augmenting the input by adding Perez sky luminance maps (*skymap*) improves the accuracy of the predicted luminance maps (Table 2). In comparison to direct and diffuse irradiance parameters, panoramic Perez sky maps provide spatial information of the sky luminance distributions (Figure 36). Both sun patches and sky maps can be simulated quickly, thus have little impact on the simulation efficiency. Sky maps in a given location do not change from project to project, as they are derived from weather data. The final input feature vector includes 9 parameters: *px, py, al, az, dir, dif, avg, sunpatch, skymap.*

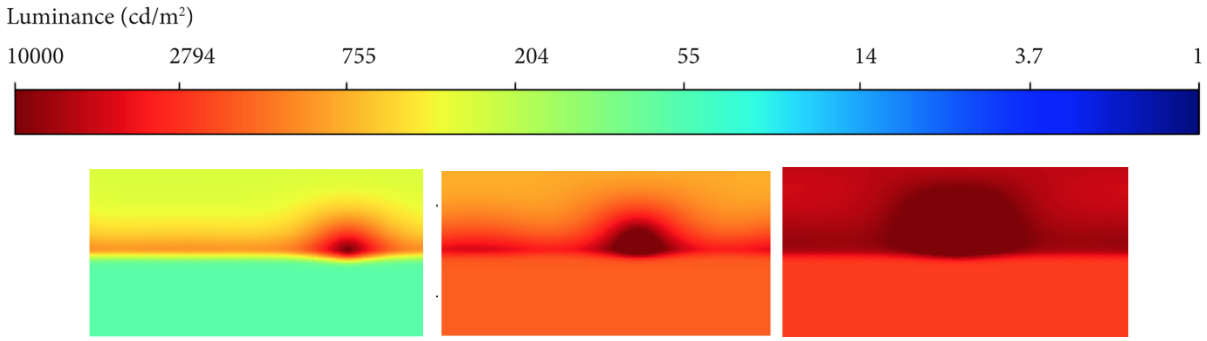


Figure 36. Samples of panoramic sky renderings generated using Seattle weather file (left) at 6 pm on September 6th, (middle) 2 pm on December 9th, and (right) 12 pm on January 25th; rendered with Radiance RPICT and showed in falsecolor with a logarithmic scale.

5.2 Data Generation

The neural network model is trained and evaluated using high-quality HDR renderings in order to develop the methodology under controlled settings. The test room model (Figure 37) is identical with the one utilized in Chapter 4. Different from the previous study, the camera is placed at the front center of the room at eye level of 1.6m (previously 2.6m) from the floor to correspond to a standing human's perspective. With the adjusted camera height and view, the rendered images in the database have a much broader luminance range ($0\text{-}10^{+8.3}$cd/m^2) as the direct sun may appear in the field of view, compared with the luminance range ($0\text{-}3\times10^4$) of the dataset in the previous study.

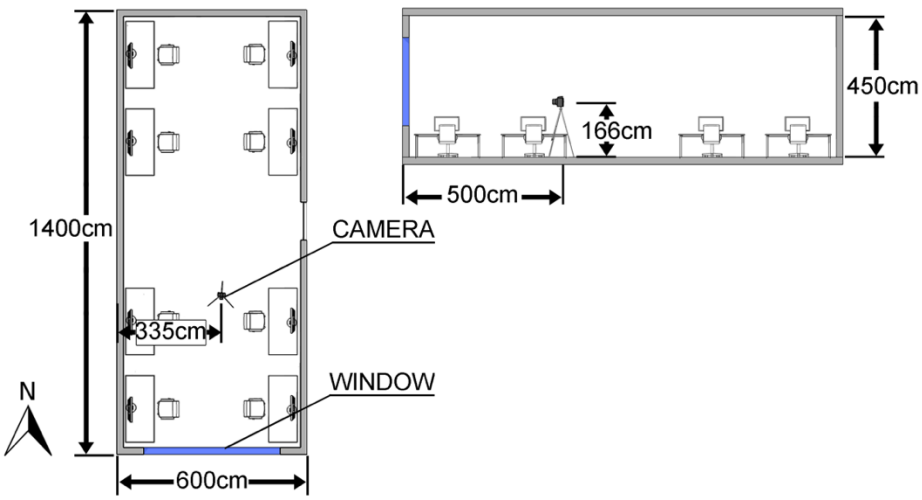


Figure 37. Plan and section of the test scene used in Radiance simulations.

Three sets of images are rendered using the Radiance RPICT method. The interior maps are rendered with 4 ambience bounces (-ab 4) and sun patches are rendered with 0 ambient bounce (-ab 0); they share the other RPICT rendering parameters (-ps 2 -pt .05 -pj .9 -dj .7 -ds .15 - dt .05 -dc .75 -dr 3 -st .15 -aa .1 -ar 512 -ad 2048 -as 1024 -lr 8 -lw .005). Sky maps are generated with the same rendering parameters as the interior maps. All data is generated with Perez all-weather sky model with inputs of direct and diffuse irradiances extracted from the EnergyPlus weather file. All images are rendered as equi-rectangular (spherical) panoramas from a single viewpoint (Figure 38). The panoramas are rendered with a pixel resolution of 1840×920 and downsized to 460×230 using Radiance pfilt for anti-aliasing. All sets of images are generated in 1-hour intervals from sunrise to sunset for the entire year, resulting in 4379 data samples.

The generated images are divided into training, validation and test sets. The model learns to approximate the light transport model from the data in the training set and evaluates its performance with the validation set during training. After the model is trained, its performance is evaluated using the test set. K-means clustering is utilized to select the training samples evenly

distributed over the light domains (**Error! Reference source not found.**Figure 30 in Chapter 4) i n order to cover various sun positions and sky conditions.

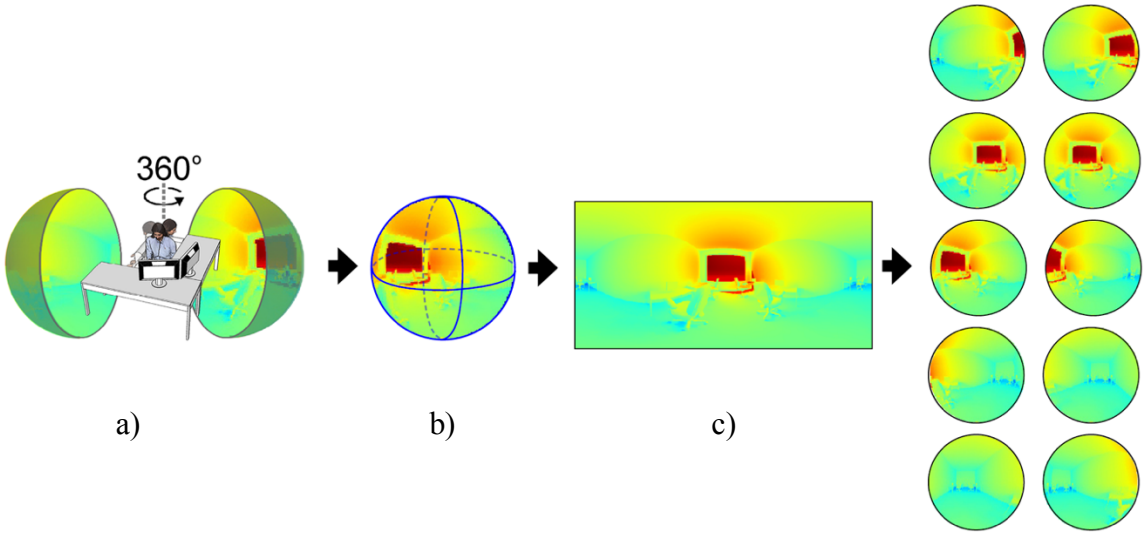


Figure 38. a and b) Full 360° field of view models an occupant's visual experience and provides more complete evaluations than a single fixed perspective view; Radiance generated equi-rectangular panoramas are used as (c) inputs to the DNN and then converted to fisheye images (d) for visual comfort analysis.

5.3 Data Pre-processing

After the data generation step, the input parameters and the output luminance is preprocessed and normalized to the range of [0,1], with a normalization function $x = \frac{x - x_{max}}{x_{\max} - x_{min}}$. Normally, x_{max} and x_{min} should be selected as the highest and lowest value of the parameter x in the training set and applied to the validation and test set when evaluating the model. In this study, the training set only contains a limited number of images which will not cover the entire luminance range in the database. Therefore, luminance values from the test set can be higher (or lower) than the maximum (or minimum) value of those from the training set. Setting the x_{min} and x_{max} of the luminance

parameters to be known luminance range of daylit scenes as 0 and $1.6 \times 10^{+9}$ cd/m^2 as the luminance of solar disc (Karandikar, 1955) provides a generic range for any given scene.

Normalization is done to reduce the variation in the parameters so that parameters of different scales and units contribute proportionately to the final results. However, normalizing the data in this study without other processing steps can be problematic as the majority of the luminance values are low after scaled by the luminance of the solar disc. Figure 39(a) shows the luminance distribution of images from the training set after the normalization without any further processing. It is clear that most of these normalized luminance intensities are very low. This is due to the high luminance range of the daylit scene, which spans several orders of magnitude. In this study, the maximum ground truth luminance is $10^{+8.3}$cd/m^2. To make the luminance values more evenly distributed over the range of [0, 1] after normalization, several additional preprocessing schemes are tested. Initially, a scheme similar to the study of Chapter 4 is performed with a gamma correction of 2.2 (Figure 39 (b)). However, this does not provide satisfactory results, because the luminance range of the dataset in this study is roughly 10^4 broader compared with the previous study. A better distribution is observed when further increasing the gamma to 10 (Figure 39(c)). This scheme fails to highlight important image features, such as the edges and corners (Figure 40(b)). The optimum scheme combines the following processing steps on the luminance intensities to distribute the values throughout the range more evenly: First, a $\log_{10}$ transform is applied to spread the range of luminance more evenly (Figure 39(d)), and the maps are normalized to the range of [0, 1]. A gamma correction of 1.5 is performed to shift the intensity distribution towards a mean value of ½ (Figure 39 (e)). Note the gamma value is customized to this room, but an algorithm that automatically chooses the optimum gamma can be developed. Our experiments show that the model performance is significantly improved after the pre-processing (Table 2) with

finer details (Figure 40(c)). After the training process, the inverse of the preprocessing transform is applied to reconstruct the final luminance maps.

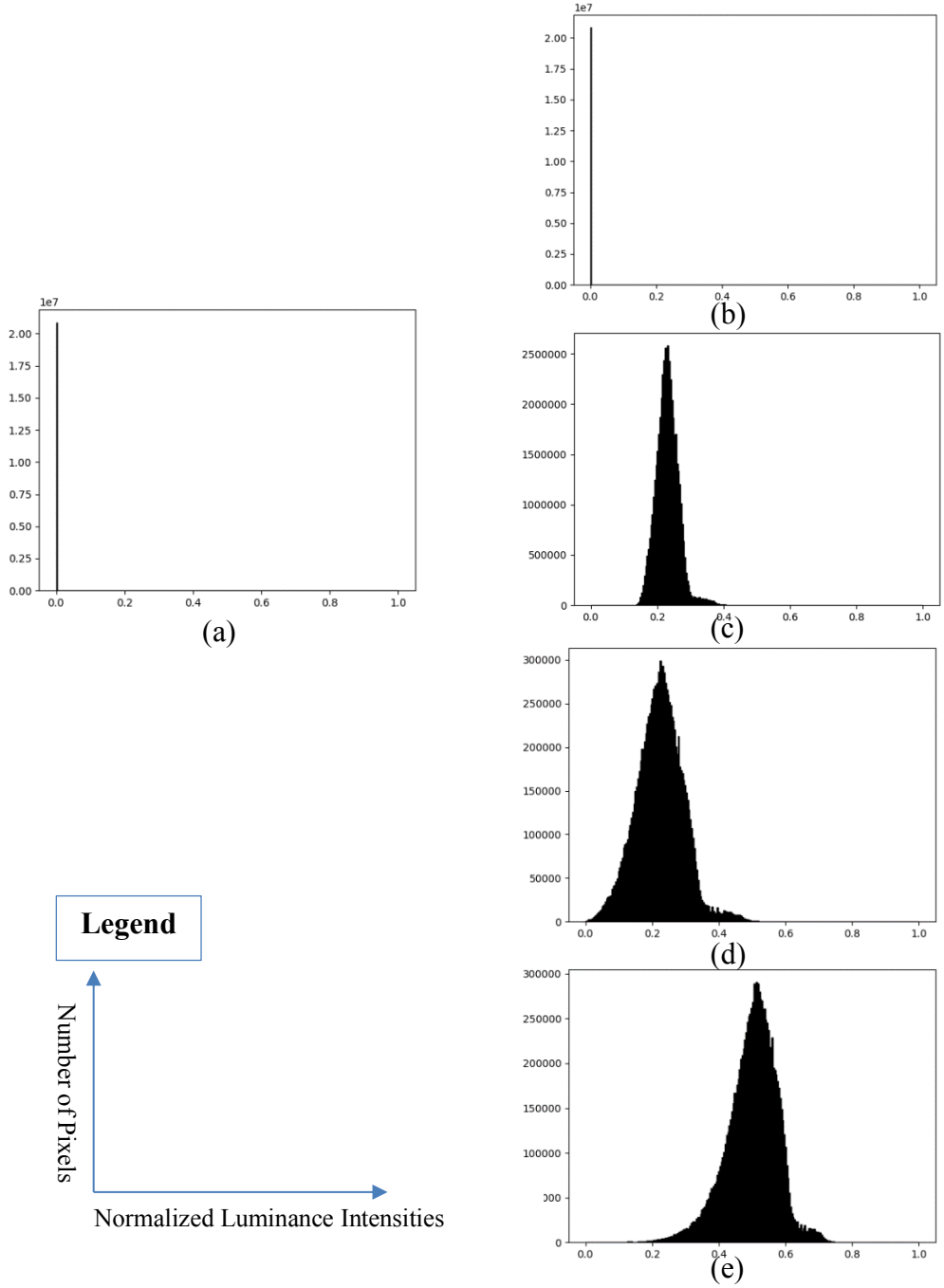


Figure 39. Diagrams show the distribution of the normalized luminance values with different preprocessing schemes: (a) the original, (b) gamma correction (gamma = 2.2), (c) gamma correction (gamma = 10), (d) $\log_{10}$ transform, and (e) $\log_{10}$ transform + gamma correction (gamma = 1.5); over the range of [0, 1]. Scheme (e) is selected as it spreads the range of luminance more evenly and shifts the intensity distribution towards a mean value of ½.

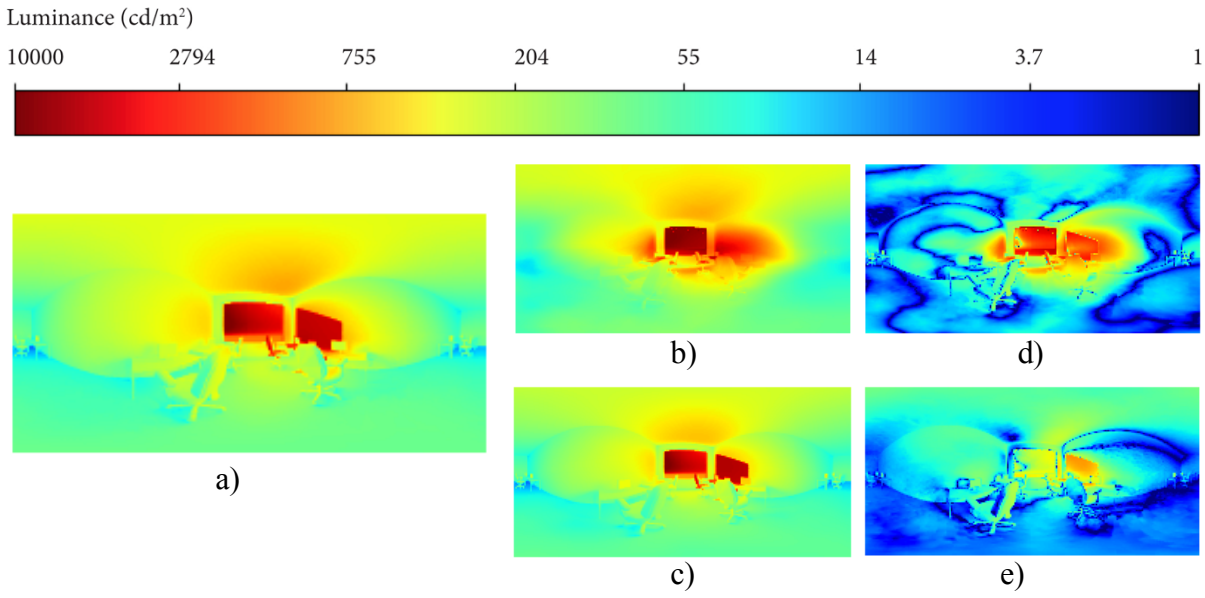


Figure 40. An exemplary test case (at 8am on October 25th with direct and diffuse irradiances of 476 and 77 W/m^2) demonstrates the impact of data preprocessing schemes on prediction results. The ground truth luminance map as shown in (a). The predicted luminance maps of two processing schemes are compared: (b) gamma correction (gamma = 2.2) and (c) a combination of $\log_{10}$ transform and gamma correction (gamma = 1.5). (d) and (e) are errors maps for (b) and (c).

5.4 Network Architecture

The network model in this study is built upon the model developed in Chapter 4. In that study, a CNN model is used to regress the light transport model from the training input. However, the CNN model has difficulties in learning fine details under different lighting conditions. Since training the CNN requires the complete panorama input (460 × 230 × 9), small batch sizes are required during the training due to GPU memory limitations. In other words, only a few lighting conditions are presented to the network in each batch step during training. The limited amount of data in the small batches does not fully represent the overall data distribution of the entire model. Training does not converge as well as it could if each batch better represented the range of the data distribution.

To overcome this limitation, the algorithm uses a multi-input deep dense network model architecture. The input data at each training step is sampled from various pixels and lighting conditions in the entire training set. Experiments show that the updated model and batch sampling strategy greatly reduces the prediction error (Table 2).

The next key consideration of the model architecture design is choosing the depth of the network and the width of each layer. The final network architecture is found via experimentation guided by monitoring the validation set error (Figure 41). Monitoring the training error is less meaningful as the error will always decrease with more trainable parameters, and testing error should not be used for fine tuning the parameters. The final developed network architecture divides the 9 input features into two separate vectors as input to a two-input deep network consisting of fully connected layers with ReLU activations. The first input branch, which encodes *px, py, al, az, dir, dif, avg*, and *skymap,* consists of four fully connected layers, each with 600 nodes (Figure 41). The second input branch, which encodes *sunpatch,* consists of a single 400 nodes fully connected layer. The first input branch encodes parameters describing global illumination, while the second input branch encodes direct sun illumination. The two branches are concatenated to form a 1000 node layer, which is followed by a fully connected layer of 600 nodes. The output layer is a single fully connected node that predicts luminance values. During training, the differences between predicted and ground truth luminance values are calculated using a loss function, and the model weights are updated accordingly.

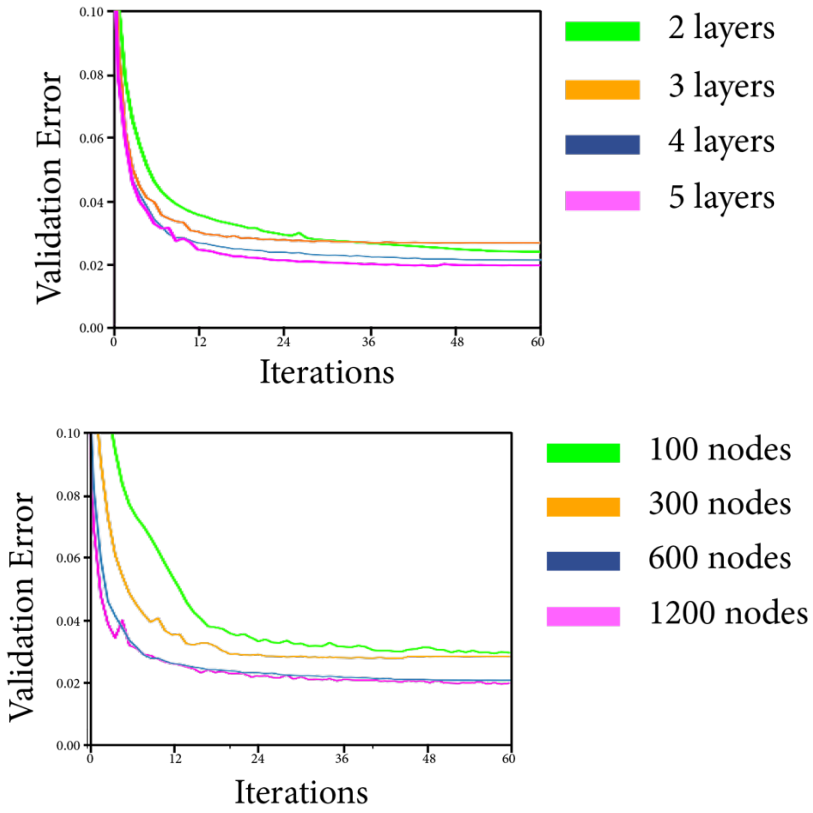


Figure 41. Diagrams show the comparison between the developed network model with different (Top) depths and (Bottom) width of each layer. The average validation error is measured after each iteration (i.e., a complete pass through the training set) as training proceeds. Convergence is attained after 28 to 30 passes. (Top) The model with a depth of 4 and 5 layers has the lowest validation error. 5-layers model has a slightly lower error, but considering the computational cost, 4-layers architecture is selected. (Bottom) The model with a width of 600 nodes in each layer has similarly low validation error as the 1200 nodes one, but with less computational cost.

5.5 Training Details

The loss function in this study is composed of two functions: $\mathcal{L}oss(y,t) = \mathcal{L}_{MSE}(y,t) + \lambda \times \mathcal{L}_{RER}(y,t)$, where y is the network prediction, t is the ground truth, and λ controls the relative weights between two functions. The first function calculates the L2 distance (mean square error) between the predicted and ground truth luminance $\mathcal{L}_{MSE}(y,t) = \frac{1}{n}\sum_{i=1}^{n} \omega_i (y_i - t_i)^2$, where ω_i is

the solid angle for pixel i. Solid angle weight ω_i is added so that each pixel is appropriately weighted for spherical panoramic solid angle projection (Figure 42). A second loss of $\mathcal{L}_{RER}(y,t) = \omega_i \sqrt{\frac{\sum_i (y_i - t_i)^2}{\sum_i t_i^2}}$ is added to improve the model performance. The weight λ is set to 10 as the result of a parameter search.

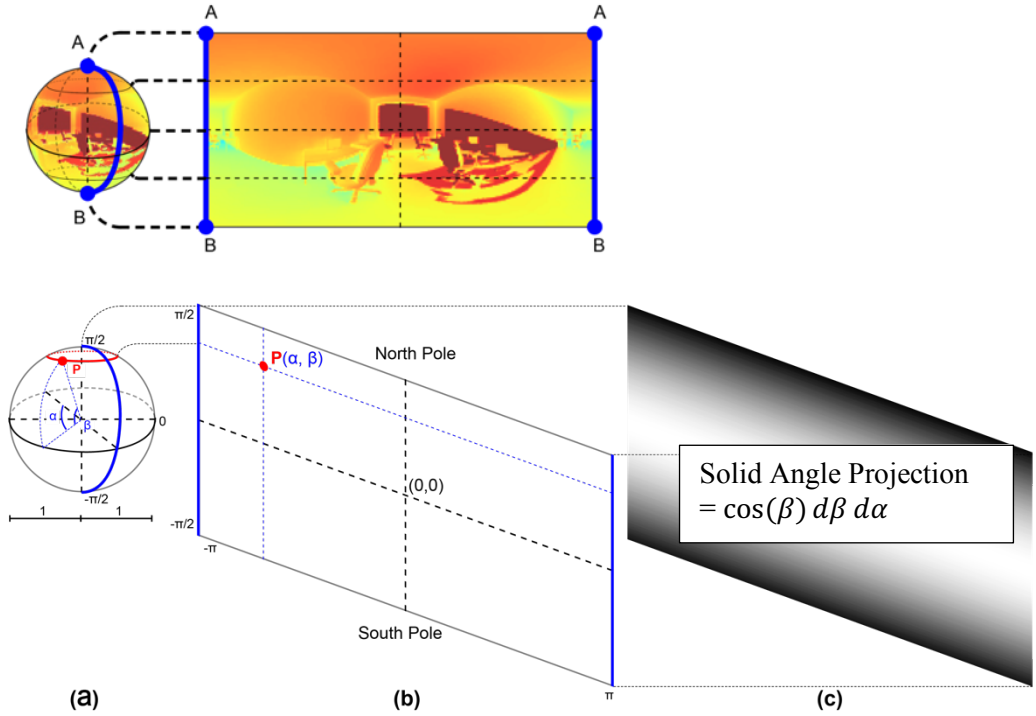


Figure 42. (Top) The equi-rectangular (spherical) panorama is mapped from a unit sphere. The north pole (A) and south pole (B) on the sphere become the top (AA) and bottom (BB) edges of the corresponding panorama. (Bottom) The image shows the solid-angle projection of each pixel in the panorama.

The model is trained with the ADAM optimizer with an initial learning rate of 10^{-3}, reducing the learning rate by a factor of 2 once learning stagnates. A mini-batch learning is utilized, it starts with an initial minibatch size of 6 × N, where N = width × height × 8 is the total number of input

parameters in one image, reducing the batch size by a factor of 2 after 30 epochs or if the learning rate is reduced to a threshold of 10^{-10}. The minibatch size of 6 × N is selected, so it fully utilizes the 11GB memory of the GPU. After a complete pass through the training set, the data is shuffled. The model is trained until the batch size is reduced to a minimum threshold of 1× N or the validation loss reaches 10^{-10}. During the training process, only the model with a reduced error on the validation set is saved. Training takes roughly 30 minutes on a NVIDIA GeForce ^{1080Ti} GPU. At test time, predicting a high-quality panoramic luminance map takes 1/10 second.

5.6 Results and Analysis

The neural network model is trained with 200 samples (5% of 4379 images) as suggested in Chapter 4. The performance of the trained model is then evaluated on the test set of 500 randomly selected lighting conditions, which have not been seen by the network before. The predicted panoramas are compared with the panoramas rendered by Radiance using per-pixel errors, falsecolor analysis with image subtraction operations, and human-centric metrics (visual comfort and perceptual metrics).

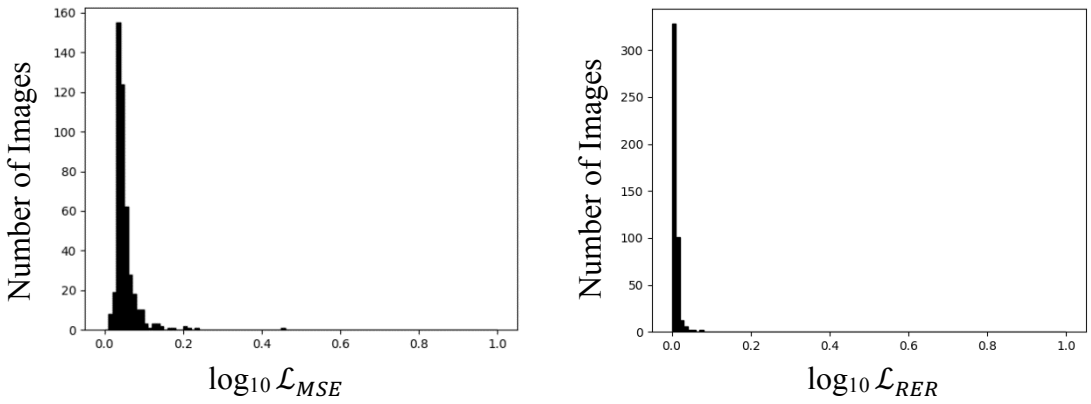


Figure 43. Diagrams show the error distribution maps over the test set of 500 images.

Similar to the previous study, the first set of metrics are per-pixel errors measured by $\mathcal{L}_{MSE}$ and $\mathcal{L}_{RER}$ used in the loss function. The errors are solid projected angle weighted and calculated with logarithmic ($\log_{10}$) luminance values to correlate with the human visual system (Fechner, 1966). Note the range of ground truth intensities is [0, 8.3]. The developed network succeeded in achieving a $\log_{10}\mathcal{L}_{MSE}$ of 9.54e-03 and a $\log_{10}\mathcal{L}_{RER}$ of 4.85e-02 on the test set, with the error distribution map presented in Figure 43.

Per-pixel errors do not necessarily prompt significant differences in architectural lighting design evaluations. For example, if the sun is estimated to be located a pixel off its ground truth location, a significant $\mathcal{L}_{MSE}$ and $\mathcal{L}_{RER}$ error will incur but will hardly lead to a different design decision. Here, the per-pixel errors are provided as benchmarks for future studies. Similar to the previous study, metrics that are commonly used by design professionals, such as falsecolor images and daylight glare probability (**DGP**) (Wienold and Christoffersen, 2006), are utilized to evaluate the results. Another human-centric metric RAMMG (Rizzi et al., 2004) is added in this study to compare the results.

5.6.1 *Falsecolor Image Comparison*

Six different test cases with different sky types (clear, intermediate, and overcast) are displayed in falsecolor in Figure 44. Error maps illustrate the absolute difference between ground truth and predicted luminance maps. The results show that: 1) The predicted panoramic luminance maps (generated from 5% sparse samples) are visually imperceptible from ground truth results in falsecolor images in a 1-10,000 cd/m^2 range without the utilization of error maps. This implies that the proposed method will lead to the same design decisions as conventional rendering methods, with orders of magnitude less calculation time. The error maps reveal that higher errors in each case occur in the window area. 2) Among all sky conditions, images of the clear sky with high

direct irradiance values are most challenging to predict with the highest errors. This is mostly due to the higher luminance variations in these images than those rendered under overcast sky conditions. Overall, the study obtains high-quality results illustrating the ability of the developed network to predict panoramic luminance maps under novel lighting conditions. Figure 45 shows the comparison between the model developed in Chapters 4 and 5. With several algorithmic refinements, the new model can handle higher luminance range and generate images with clearer boundaries and finer details. The applicable luminance range is widened from $0\text{-}3\times10^{4}$ to $0\text{-}10^{8.3}$. The impact of different refinements is illustrated in Table 2.

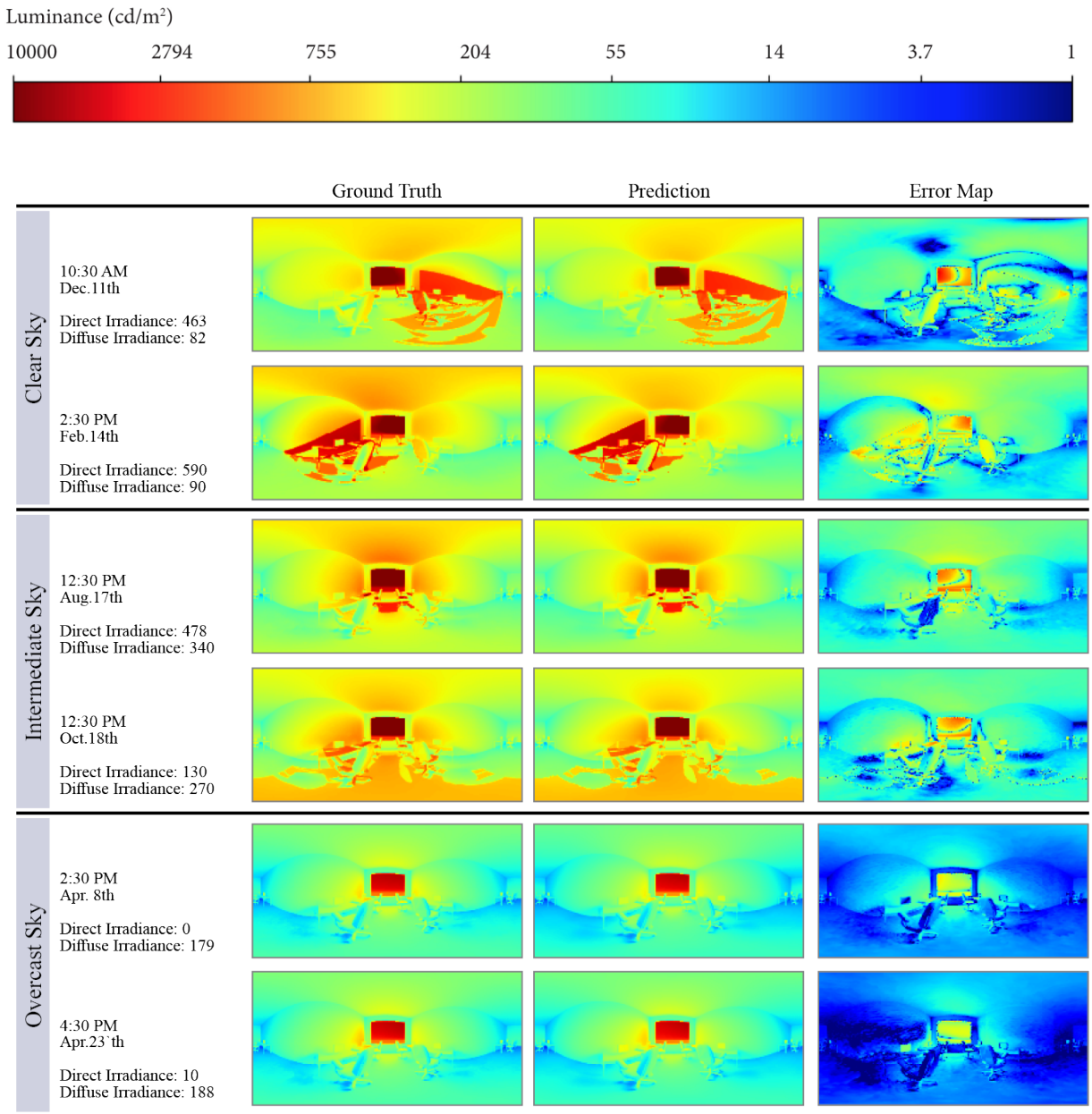


Figure 44. Six exemplary test cases with different sky types are displayed in falsecolor with a logarithmic scale.

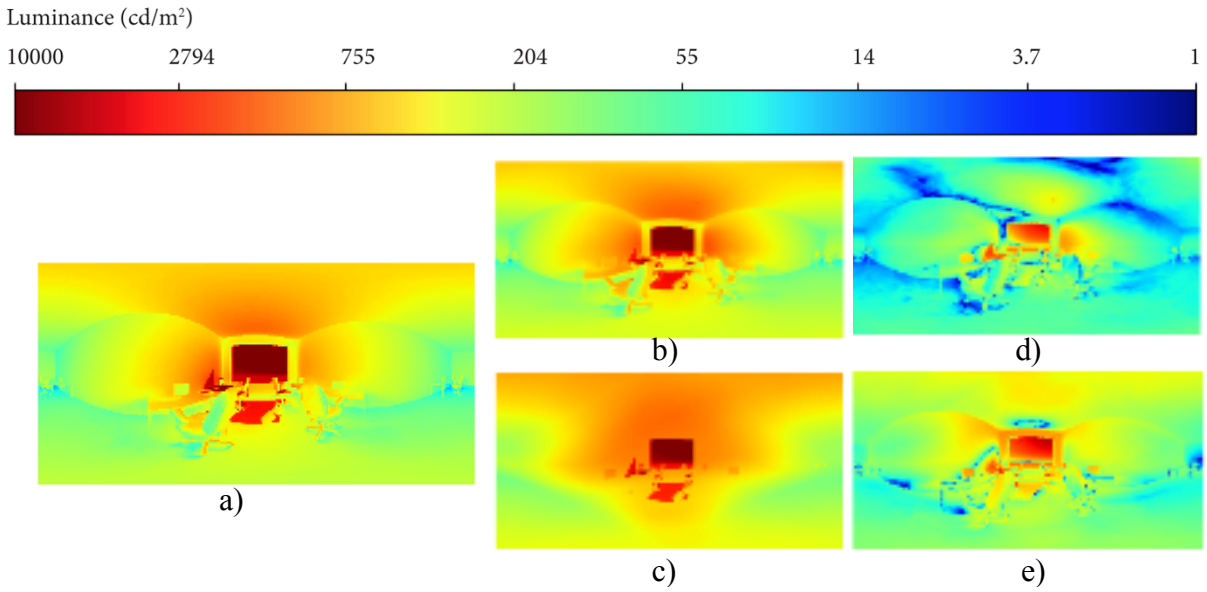


Figure 45. (a) the ground truth luminance map, (b) luminance map predictions of the refined model is compared to (c) the model developed in chapter 4. (b) has clearer boundaries and finer details, indicating higher accuracy on the lower range values. (d) and (e) are errors maps for (b) and (c).

Table 2. The final developed scheme is compared with other schemes including 1) without sky panoramas as input; 2) utilized a CNN model instead of a dense network, and 3) with a different data preprocessing transform. The final scheme shows better performance on all metrics.

Scheme	***Preprocessing***	***Network Architecture***	***Sky Panoramas***	***Log10 MSE***	***Log10 RER***	***DGP MSE***	***RAMMG MSE***
Final	$\log_{10}$ + gamma 1.5	Dense	Yes	**9.54e-03**	**4.85e-02**	**1.7e-04**	**29**
1	$\log_{10}$ + gamma 1.5	Dense	***No***	9.78e-03	4.91e-02	2.0e-04	35
2	$\log_{10}$ + gamma 1.5	***CNN***	Yes	1.78e-02	5.61e-02	2.78e-04	41
3	***gamma 10***	Dense	Yes	2.81e-02	8.31e-02	2.1e-03	151

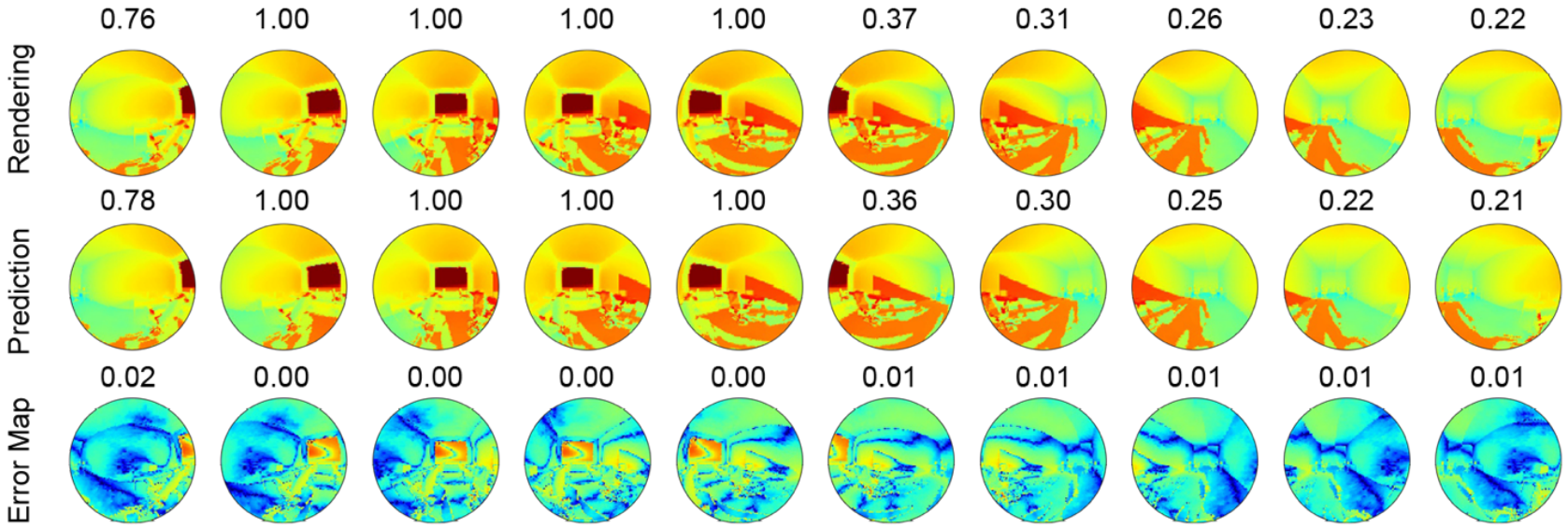


Figure 46. Comparison of Radiance rendered and DNN predicted fisheye images over a 360° view in 36° increments. The ground truth and predicted fisheye images are labeled with the DGP value. The error maps are labeled with absolute difference in DGP values. All images are shown in falsecolor with a logarithmic scale.

5.6.2 *DGP*

For DGP analysis, each panoramic luminance map in the test set is converted to 10 equidistant fisheye images over 360° of rotation in the y-axis using 36° increments (Figure 38). DGP is then applied to the predicted and ground truth fisheye images. Figure 46 shows the per-view-angle DGP comparison of an exemplary test lighting condition (January 21st 11:30 am, direct and diffuse irradiances of 553 and 87 W/m^2), with error maps illustrating the absolute differences. The result illustrates that: 1) a subject's visual comfort level can change vastly across various view directions of the same scene depending on the shift of the sun location in and out of the field of view and the relative position of the sun in the field of view, 2) the neural network model accurately predicts the overall luminance environment, with the predicted DGP values from fisheye images closely matching those of ground truth ones in all directions, and 3) the more significant errors occur towards the window directions.

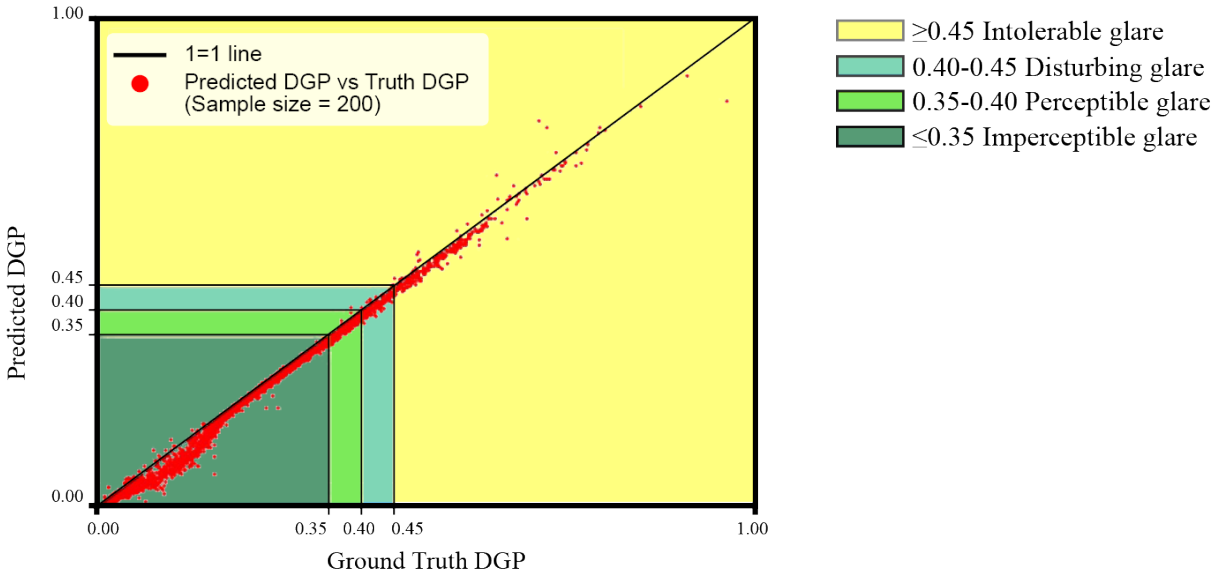


Figure 47. Scatterplot comparison of DGP values calculated using ground truth and predicted images

Figure 47 shows the comparison scatter plot of the DGP values for all of the test lighting conditions. The analysis shows a relatively strong agreement overall (with an r^2 of 0.99 and MSE of 1.7e-4) but exhibits some noise at higher ranges. Further investigation shows the high errors occur around winter noon time, when sun direct appears in the field of view through the south-facing window (Figure 48). The model has difficulties in estimating the sun luminance ($1.6 \times 10^{+9}$ cd/m^2) which is multiple orders of magnitude higher than the rest part in the scene. The failure cases are negligible because: 1) these cases only account for about 1% of the total tests cases; 2) the predictions still fall in the same zone (intolerable) as the ground truth DGPs, and 3) the situation is avoidable in the real world when people have an ability to change their view directions and positions to avoid a direct view of the sun. For scenes with actual DGP in the imperceptible, perceptible and disturbing zone, our method produced very little variation.

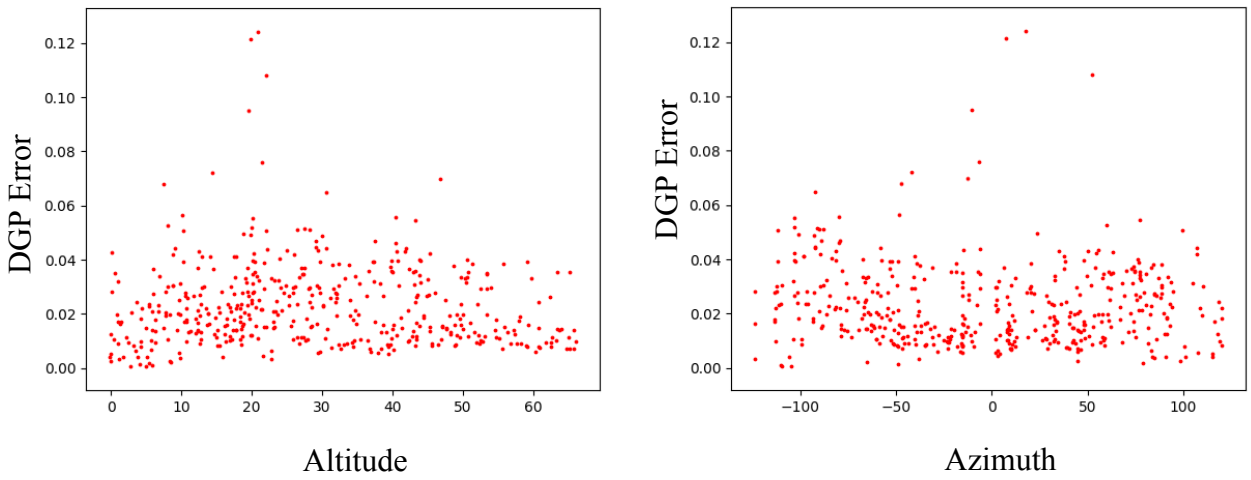


Figure 48. Scatter plots of DGP prediction errors (differences between predictions and ground truth) and the sun altitude (left) and azimuth (right). The max error occurs at sun altitude and azimuth of roughly 20° and 0°, respectively, when sun directly appears in the field of view through the south window in the Winter.

5.6.3 *RAMMG*

RAMMG is a contrast measure of images, and is utilized as a perceptual predictor of contrast-based visual characteristics of daylight (Rockcastle et al., 2016). RAMMG creates an image pyramid by sub-sampling an image to n-levels. At each level the resolution is reduced to half of the previous level. The RAMMG measure is the mean contrast calculated at each pyramid level. RAMMG is applied to both the predicted panoramas and Radiance rendered panoramas in the test set. Subsampling of five levels is used taking into consideration the image resolution (460 by 230) (Figure 49). Figure 50 illustrates the correlation between the predicted and ground truth RAMMGs; it shows a strong agreement (r^2 of 0.98) with a few errors at higher ranges, when the sun directly appears through the window (r^2 of 0.99 if 4 outlier cases are excluded).

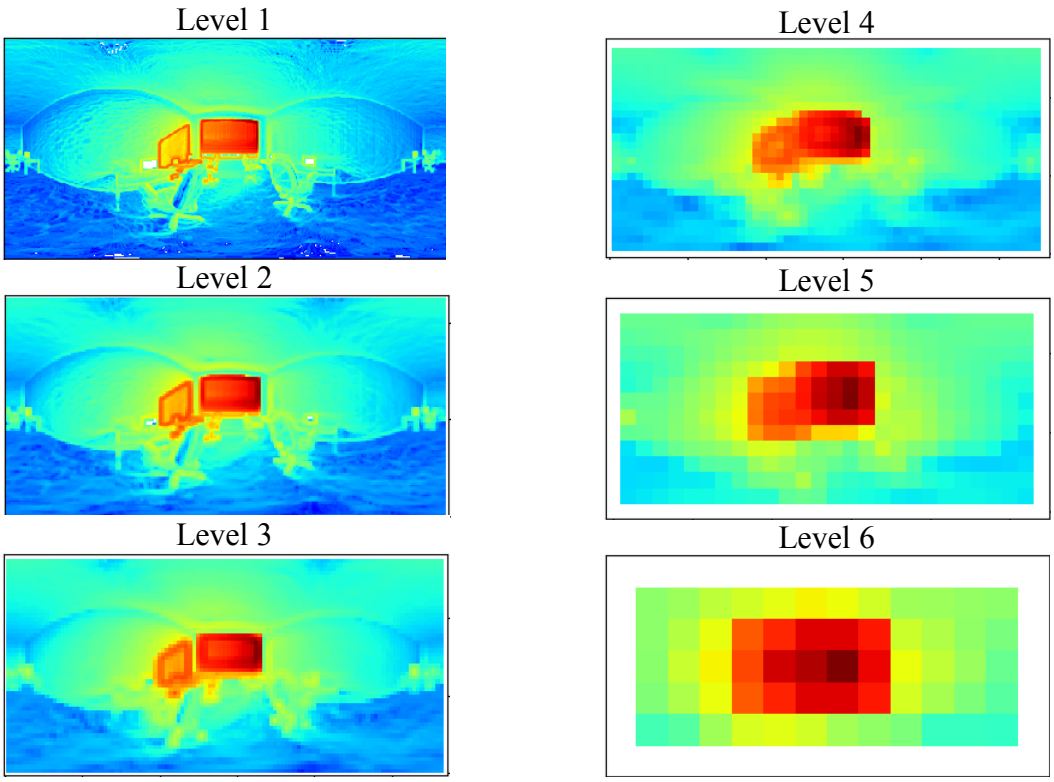


Figure 49. Each image represents a contrast map (each pixel represents the sum of neighborhood contrast of that pixel) of a sample panorama, calculated at every subsampling level starting from level 1. The image resolution at each level is reduced to half of the previous level. The contrast map starts providing less meaningful information from level 6.

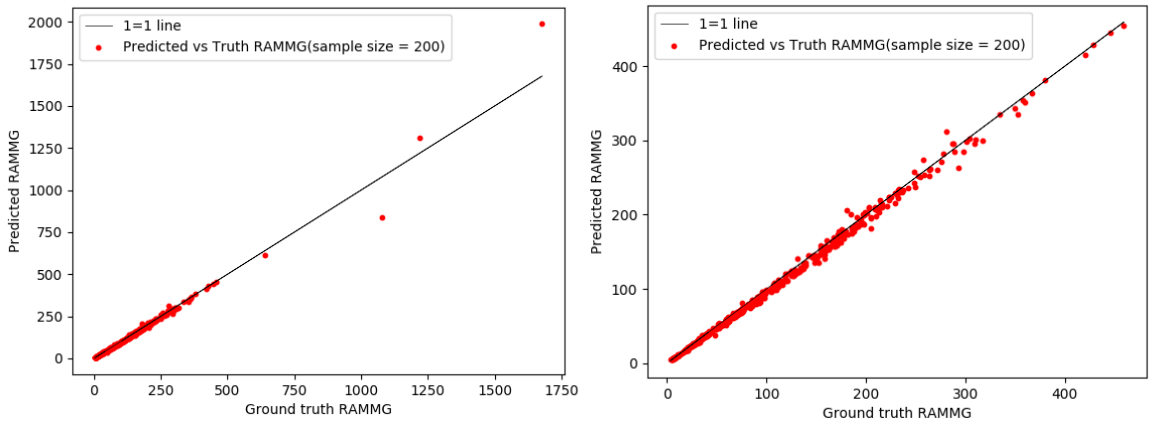


Figure 50. Comparison of RAMMG values calculated using ground truth and predicted images: (Left) original data, and (Right) zoomed in data.

5.7 Discussion and Next Steps

This study provides a robust workflow that regresses annual HDR panoramas from sparse samples (5%) using a DNN framework. The DNN approach accelerates the generation of panoramic luminance maps, and only requires a few seconds of computing time during evaluation. The predicted panoramic luminance maps can be transformed into fisheye projections for any view direction. The accuracy, applicability, and usefulness of the proposed DNN model is examined through comparisons with Radiance RPICT renderings using per-pixel error metrics, visual comfort, and contrast metrics. Architects, lighting designers, and researchers can effectively use the workflow to perform multi-view annual lighting analyses.

Compared with the previous study, the model has increased prediction accuracies under various sky conditions. The panoramic input format provides the ability to analyze head and gaze movements from a fixed point.

In Chapter 6, the focus shifts to improve the efficiency and applicability of the method through better selection of training samples. A sensitivity analysis is performed to inspect the relationship between the collection period of the training data and the accuracy of the predictions with the refined model developed in Chapter 5. Multiple data collection periods are tested to understand the minimum and optimum data collection periods for generating annual luminance maps. The difference between the predictions and ground truth renderings are measured by metrics utilized in this study.

Chapter 6. OPTIMIZING THE DATA COLLECTION PERIOD

Chapter 5 demonstrates that the developed machine learning model predicts annual panoramic luminance maps with high accuracy from a small subset of HDR images, and it accelerates annual luminance-based simulations. Although the training set is comprised of only 5% of the possible daylight hours, the data are evenly distributed over the entire year. This chapter presents results from two investigations where we show the impact of training samples on prediction results. Specifically, the objective is to minimize the duration of continuous data collection, while retaining high accuracy predictions, and study the impact of sun positions and sky conditions on the accuracy of the predictions. The results can demonstrate the applicability and feasibility of performing annual luminance-based evaluations from in-situ HDR photographs. Another investigation aims to further minimize the total data collection time and effort by utilizing a few well-timed short data collection intervals. Multiple training sets that contain images from short data collection intervals (such as 12 days, 9 days and 4 days selected during equinoxes and solstices) are tested. All analyses are performed by comparing the predicted HDR luminance maps to Radiance rendered luminance maps. Per-pixel errors, daylight glare probability, and perceptual performance metrics are utilized for comparisons.

6.1 THE MINIMUM AND OPTIMUM CONTINUOUS DATA COLLECTION PERIOD

In Chapters 4 and 5, the training samples are evenly distributed over the entire light domain of 4 dimensions (i.e., sun altitude and azimuth angles, sky direct and diffuse irradiances) using K-means clustering method. Such a training set can be generated using simulation methods, but it is not feasible with in-situ data collection. Therefore, it is necessary to explore training with samples of a single continuous collection period. A sensitivity analysis is performed to find the minimum and optimum collection period. The effect of different collection periods on the prediction accuracy of the annual luminance maps is tested. Longer collection periods, which

cover more sun positions and different sky conditions are expected to produce more accurate results, but the added burden of elongated data collection periods may have diminishing returns. The objective is to find the shortest collection period that produces similar accuracy to full year sampling. One-month data is selected (training set 1-1) and the collection period is further reduced to two weeks (training set 1-2) and one week (training set 1-3). The dataset utilized in Chapter 5 (training set 0) is used as the benchmark for comparisons (Table 3). The same machine learning workflow is performed with different training datasets to predict the annual luminance maps under various sky conditions. The validation set contains 20% of data randomly selected from the training set. The test set includes 500 randomly selected samples over the entire year.

Table 3. Proposed continuous data collection periods

Training Set Number	Total Duration	Number of Intervals	Continuous Data Collection Periods	Total Number of Image Samples
0	1 year	/	1 year (samples evenly distributed over the entire year)	200
1-1	1 month	1	January	268
			February	279
			March	356
			April	407
			May	453
			June	480
			July	478
			August	439
			September	368
			October	327
			November	301
			December	247
1-2	2 weeks	1	2 weeks in March during Equinox (March 14th-28th)	180
1-3	1 week	1	1 week in March during Equinox (March 17th -24th)	96

6.1.1 *Lighting Characteristics of One-month Training Sets*

Daylighting parameters for the one-month samples (training set 1-1) are illustrated in Figure 51-Figure 54. Figure 51 present

the distribution plots of the training samples over the sun position parameters (altitude and azimuth angles), with red, green, and blue dots represent training, validation, and test samples, respectively. Figure 52 shows the percentage of the area that the sun spans within the sunpath diagram in each month (in yellow). Compared to the evenly distributed samples (Figure 30), one-month samples cover a small portion of the daylighting variability. As a result, the model is trained with limited amount of variability. Figure 51 illustrates how altitude and azimuth angles of the sun change with the seasons. During equinoxes (March and September), the sun date lines spread to larger portions of the sky over the month (higher percentages of area in Figure 52), which means the data from these months can better represent the variance of the sun position parameters. Conversely, during the winter and summer solstices (December and June), the date lines remain closer to each other (lower percentage of area in Figure 52), which means the data from these months encapsulate limited variance of the annual sun path diagram.

Figure 53 shows the distribution plots of selected samples over the sky condition parameters (direct and diffuse irradiances). The red, green, and blue dots represent samples from training, validation, and test sets. The red dots cover more sky conditions in summer compared to winter. There are not any red dots in the upper right zones in winter months, which indicates that the training/validation sets from these months do not contain samples with skies of high direct and diffuse irradiances. Figure 54 presents the joint probability distributions of training samples over the sky condition parameters. All diagrams have been processed as four by four blocks and normalized by number of samples in the month, where the color of each block represents the

percentage of samples in that range. The diagram shows that among all months, the block with the lowest diffuse and direct irradiances (bottom left corner) has the highest probability density (i.e., overcast skies are the dominate sky conditions in Seattle). Summer months (July, August, and September) have darker blue blocks with high direct and low diffuse irradiances. Sunny skies are more common in these months. The blocks cover a larger portion of the sky irradiances over summer months, which means the data from these months can better represent the variance of the sky condition parameters.

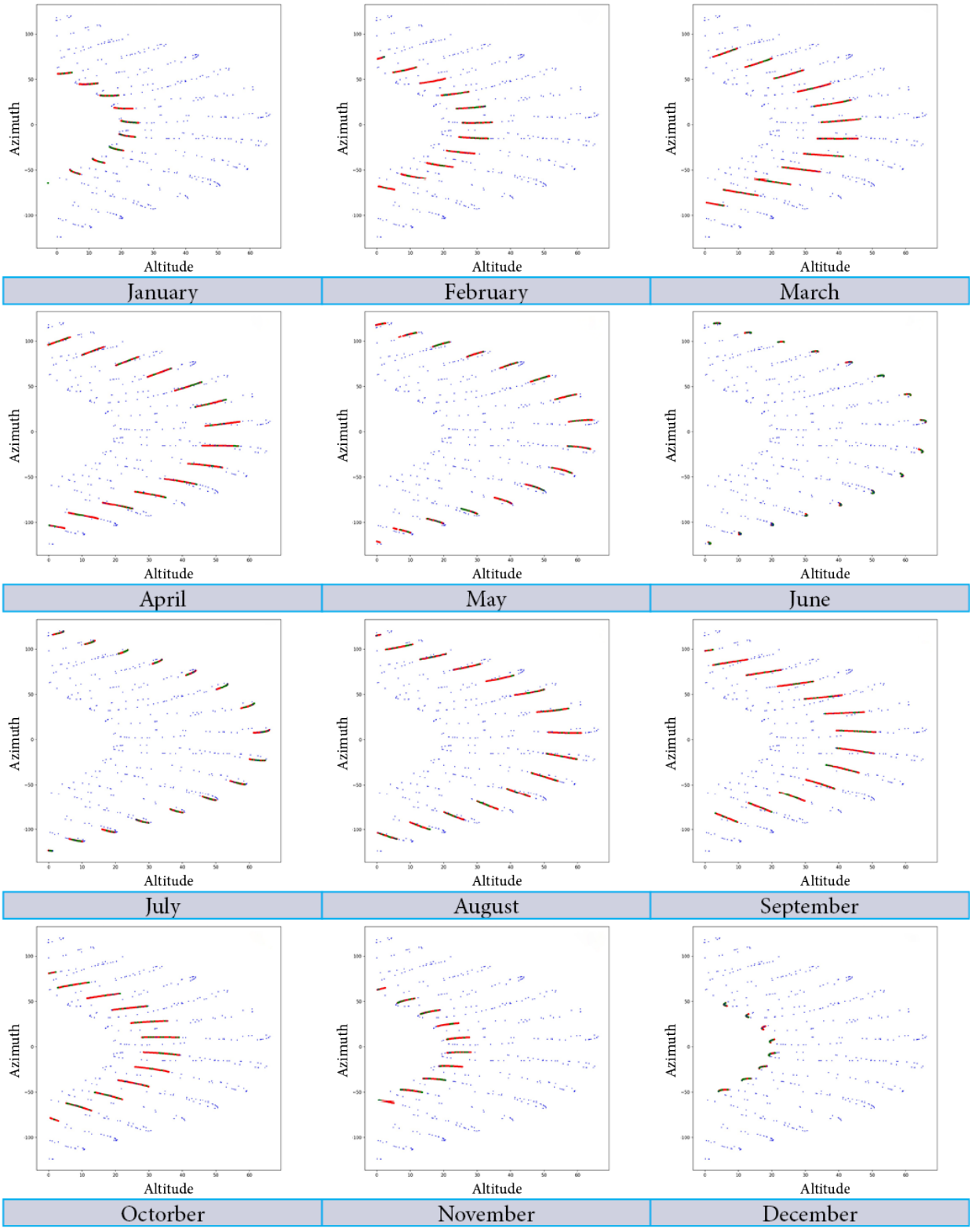


Figure 51. Distribution plots of selected samples over the sun position parameters (altitude and azimuth angles).

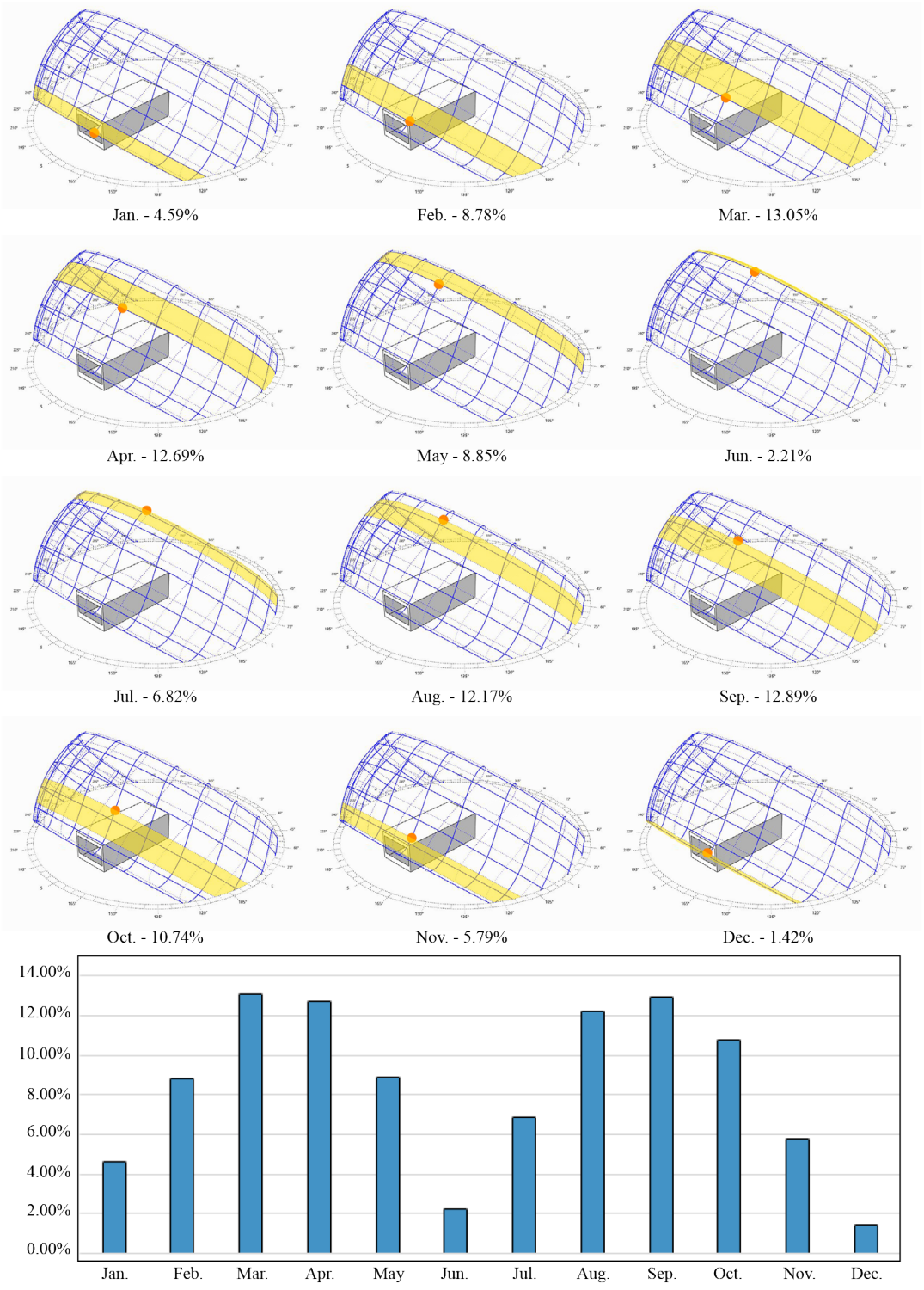


Figure 52. (Top) Sun path diagrams of Seattle with possible sun positions of each month shown in yellow. (Bottom) The percentage of the sun area that each month covers (yellow area) is shown in the chart.

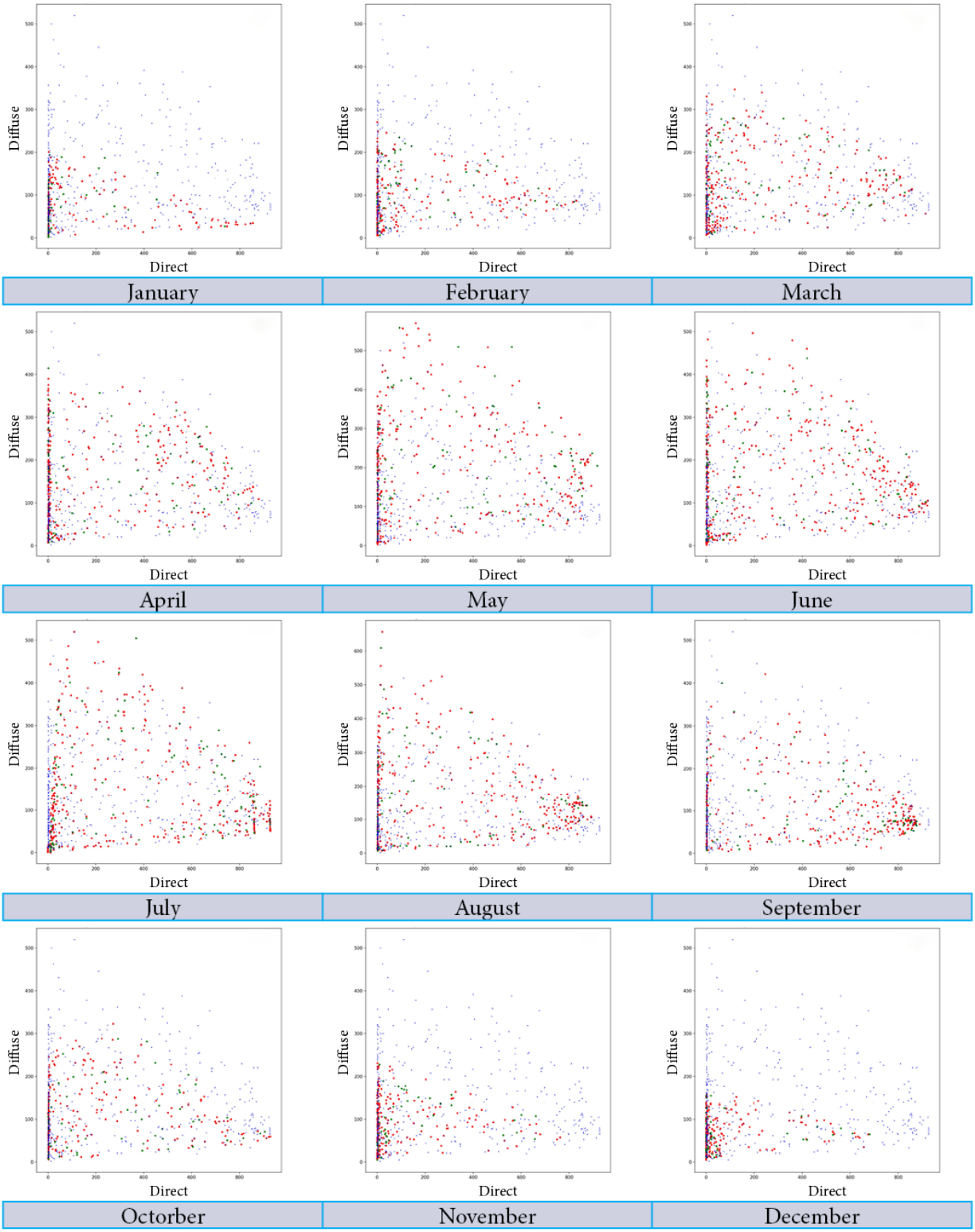


Figure 53. Distribution plots of selected samples over the sky condition parameters (direct and diffuse irradiances).

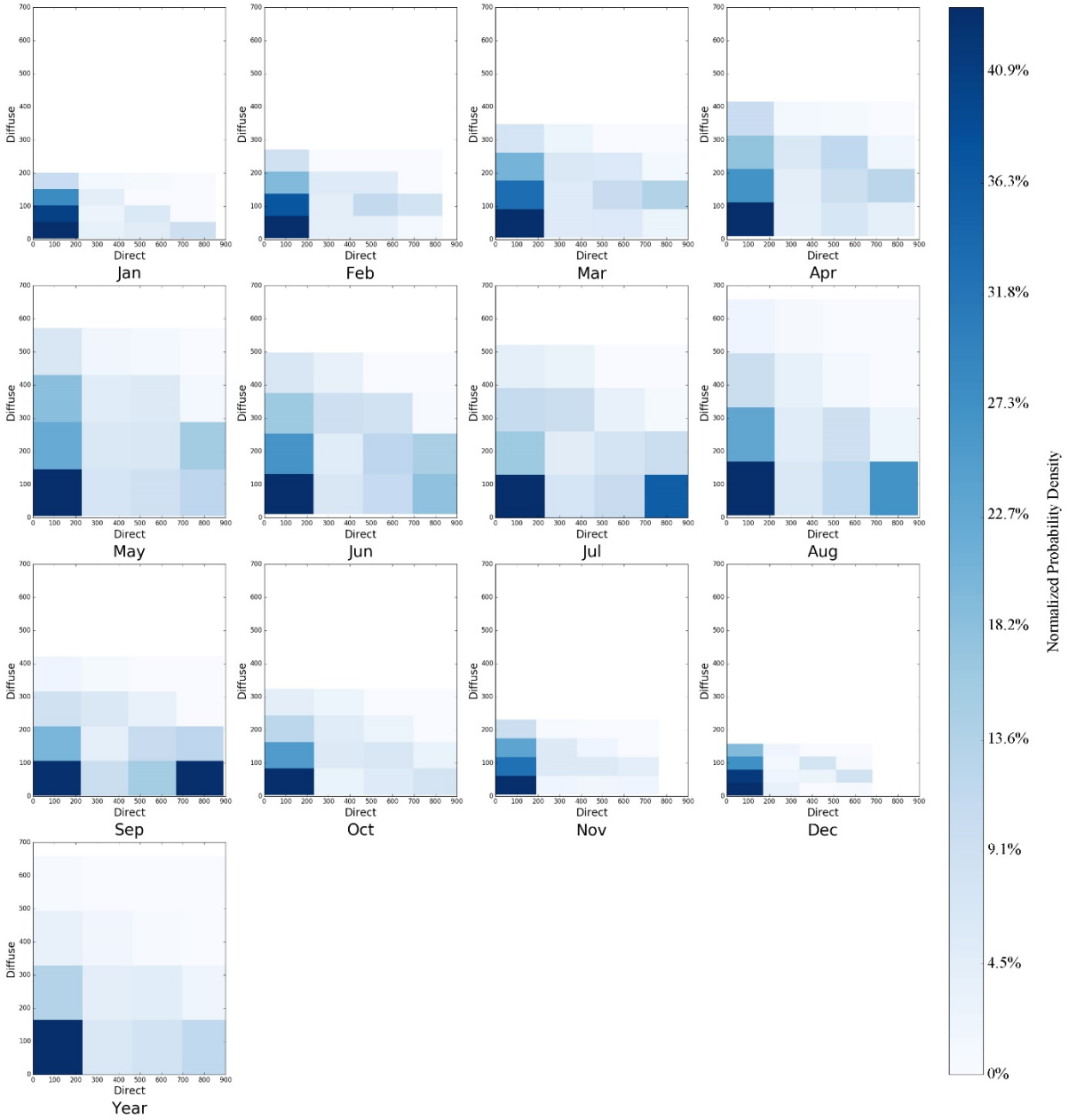


Figure 54. Joint probability distribution diagram of selected samples over the sky condition parameters (direct and diffuse irradiances).

6.1.2 *Comparisons of Predictions Generated with One-month Training Sets*

The predictions are generated with one-month training samples (training sets 1-1), and evaluated by per-pixel errors, visual discomfort metric DGP, and perceptual performance metric RAMMG. Figure 55 illustrates the comparisons of the predicted luminance maps when trained with one of the twelve month training sets and the training set 0. The differences are measured by Log_{10} MSE (Top) and Log_{10} RER (Bottom) error. The results show: 1) Significant differences are found among predictions generated with different training sets. For example, the Log_{10} MSE error generated with January training data is 10 times higher than the one generated with April training data. 2) There exists a pattern that lower prediction errors occur around the equinoxes, while highest errors occur around solstices, which implies that the most influential aspects on the prediction accuracies are the parameters of the sun positions, not the sky conditions. 3) Predictions generated with training samples from the summer months perform better than those from the winter months, indicating that the sky parameters also play a role.

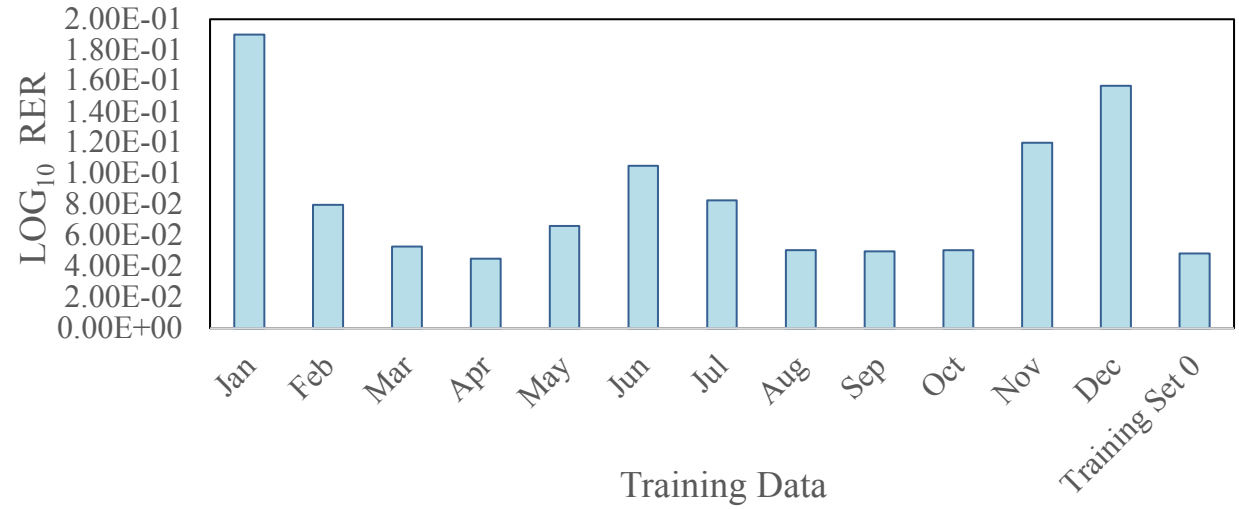
Figure 56 shows the DGP comparisons among the predictions generated with one-month training sets and training set 0. Figure 57 shows the DGP mean square errors (MSE) comparisons. The figures illustrate that: 1) similar to the per-pixel errors, DGP errors are lower, when trained with data selected around equinoxes, and higher with data selected around solstices; and 2) with training samples selected around equinoxes (e.g., March, April, September, and October), the predicted DGPs closely match to ground truth values.

RAMMG is applied to both the predicted panoramas generated with one-month training sets and Radiance rendered panoramas in the test sets to compare luminance contrasts. The preliminary study shows that RAMMG exhibits great variance at higher ranges in each calculation. This problem, as introduced in Chapter 4, is due to the difficulties in estimating the sun luminance,

which is orders of magnitude higher than the rest part in the scene. As a result, high errors occur when sun appears in the direct field of view (which accounts for less than 1% of the total tests cases). The errors bring significant noise to RAMMG analyses, but do not necessarily prompt differences in architectural design evaluations. Test cases with the direct viewable sun (1% of total cases) are excluded in the following RAMMG analyses. Figure 58 presents the comparison between training sets evaluated by RAMMG MSE. The similar trend is observed in this figure as in Figure 54 and Figure 55.

These analyses show that it is possible to accurately predict the annual luminance maps, with one-month data. The best months to capture or generate training data in Seattle are around equinoxes. The results also imply that a better strategy than evenly selecting samples over the light domain is adding more weights to the sun position parameters, as the sun positions have a more significant impact on the prediction accuracies than the sky conditions.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Training Set 0
LOG_{10} MSE	1.47 E-01	2.59 E-02	1.14 E-02	8.24 E-03	1.77 E-02	4.50 E-02	2.78 E-02	1.04 E-02	1.01 E-02	1.04 E-02	5.86 E-02	1.00 E-01	9.54E-03



Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Training Set 0
LOG_{10} RER	1.90 E-01	7.98 E-02	5.29 E-02	4.50 E-02	6.60 E-02	1.05 E-01	8.27 E-02	5.06 E-02	4.98 E-02	5.05 E-02	1.20 E-01	1.57 E-01	4.85E-02

Figure 55. Log_{10} MSE errors (Top) and Log_{10} RER errors (Bottom) between predicted panoramas generated with monthly training samples and the ground truth ones.

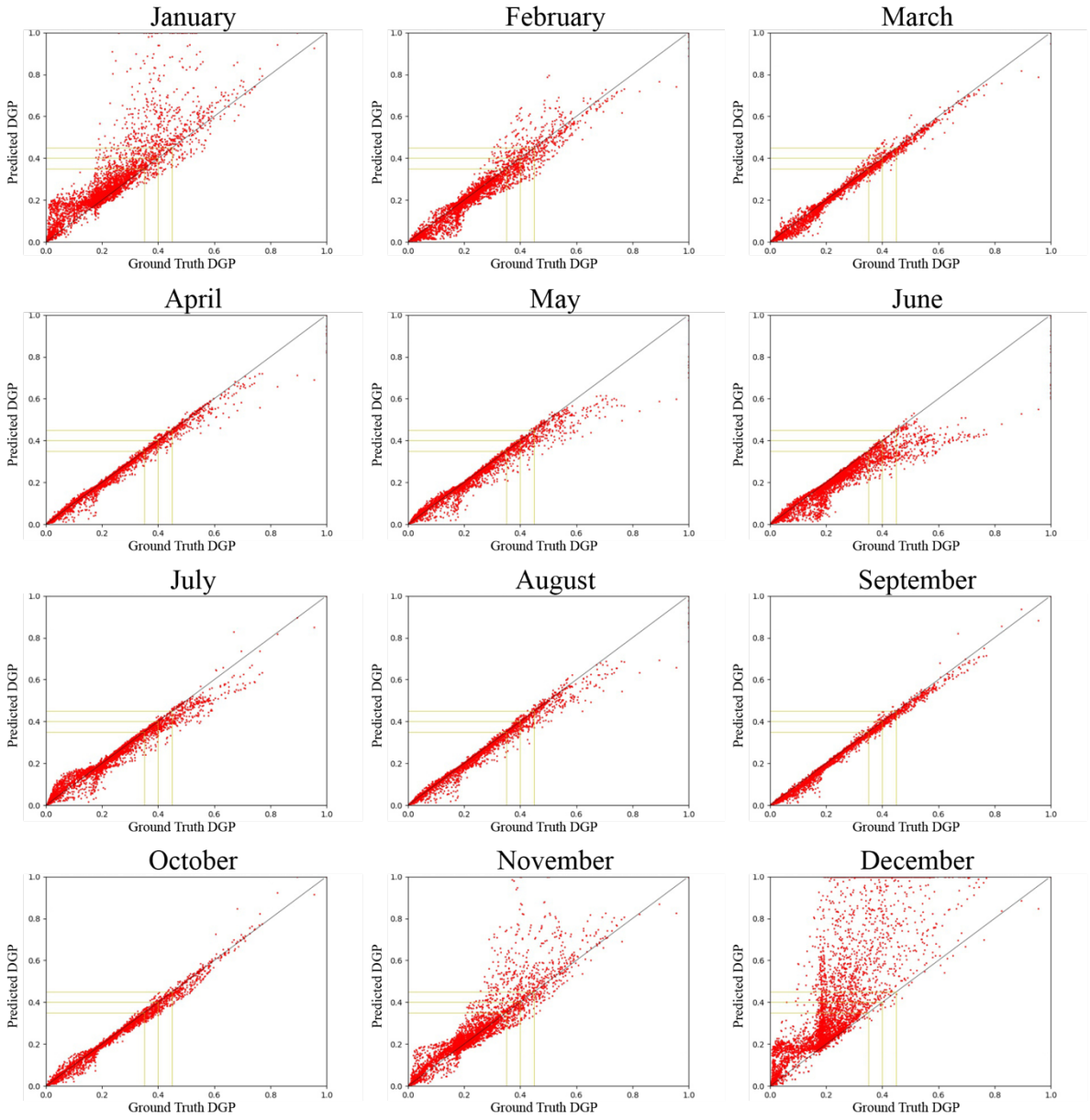
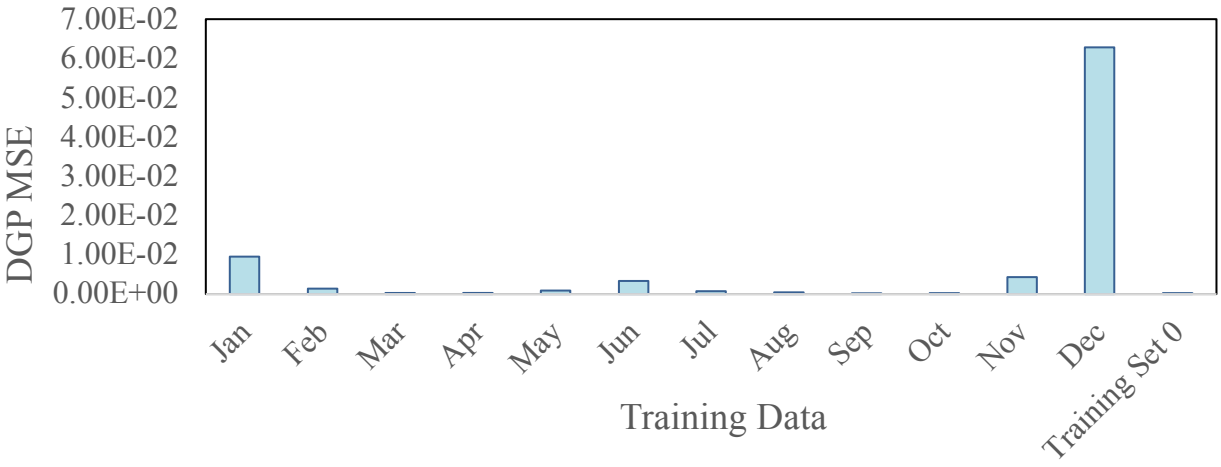
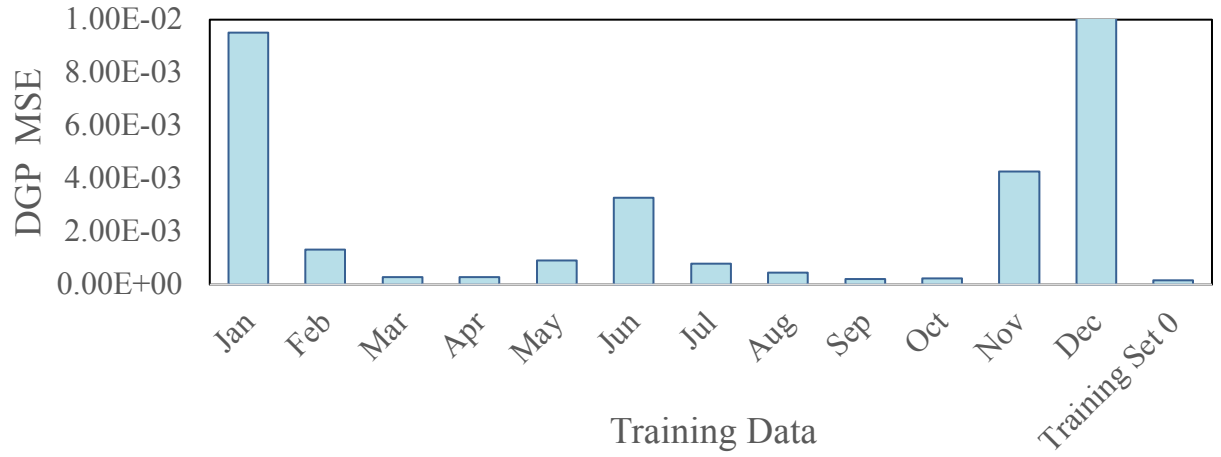


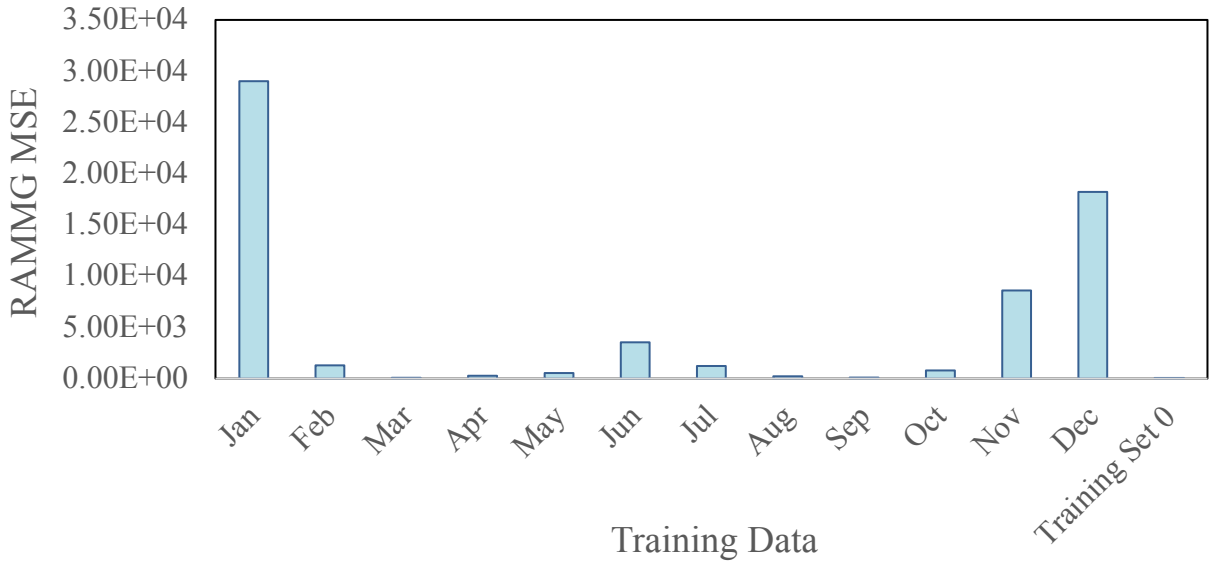
Figure 56. Scatter plots compare DGP values calculated using ground truth fisheye images and predicted fisheye images which are generated with monthly training samples.





Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Training Set 0
DGP MSE	9.51 E-03	1.31 E-03	2.81 E-04	2.89 E-04	9.12 E-04	3.28 E-03	8.01 E-04	4.41 E-04	2.00 E-04	2.40 E-04	4.27 E-03	6.28 E-02	1.74E-04

Figure 57. DGP MSEs calculated with predicted fisheye images generated by one-month training samples and the ground truth ones: (Top) Original; (Bottom) the December column was truncated in order to make the whole distribution display better.



Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Training Set 0
RAMMG MSE	2.90 E+04	1.27 E+03	8.00 E+01	2.78 E+02	5.37 E+02	3.54 E+03	1.21 E+03	2.25 E+02	9.94 E+01	7.96 E+02	8.62 E+03	1.82 E+04	2.92E +01

Figure 58. RAMMG MSEs calculated using predicted luminance maps generated with one-month training sets and the ground truth ones.

6.1.3 *Prediction Results Generated from an Exemplary One-month Training Set*

Section 6.1.2 shows that the machine learning model works better with training sets of months around equinoxes for predicting the annual luminance maps. This section further demonstrates the prediction accuracy generated with the spring equinox (March), evaluated by DGP and falsecolor image techniques (Figure 59-Figure 61).

Figure 59 illustrates the correlation between DGP values calculated from the predicted fisheye images (generated using the March training set), and those calculated from the ground truth fisheye images. Four colored zones represent four DGP scales, ranging from imperceptible (DGP <= 0.35),

perceptible (DGP >0.35 and <= 0.40), disturbing (DGP > 0.40 and < 0.45), to intolerable (DGP >= 0.45) glare. The diagram shows a strong agreement overall (r^2 of 0.99 and MSE of 2.81e-04). The majority of the predictions fall in the same zones as the ground truth DGPs.

Figure 60 shows the per-view-angle DGP comparison of an exemplary worse-case test lighting condition (January 21th 11:30 am, direct and diffuse irradiances of 553 and 87 W/m^2), with error maps illustrating the absolute differences. The result supports that the neural network model is able to predict the overall luminance environment with one-month data. In all directions, the predicted fisheye images closely match those of the ground truth ones.

Figure 61 shows the comparison between predicted panoramic luminance maps and Radiance generated ones in falsecolor with a logarithmic scale, with error maps illustrating the absolute differences. Six representative samples under different sky conditions (clear, intermediate, and overcast) are selected from the test set. The results show that the predicted panoramic luminance maps (generated from the March training set) are visually imperceptible from ground truth ones in falsecolor images without the help of error maps. This implies that the proposed method will lead to the same design decisions as conventional rendering method, but only requires 1/12 of the generated/captured data.

Overall, satisficing results are obtained in the study, which confirms the ability of the developed machine learning method to accurately predict panoramic luminance maps under novel lighting conditions, with one-month data.

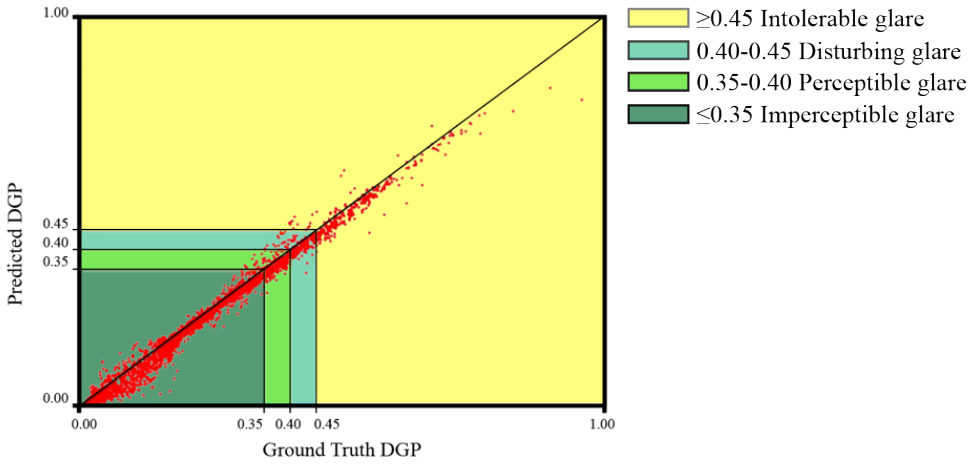


Figure 59. Correlation of DGP values calculated using predicted fisheye images, generated with one-month (March) training samples, and the ground truth ones.

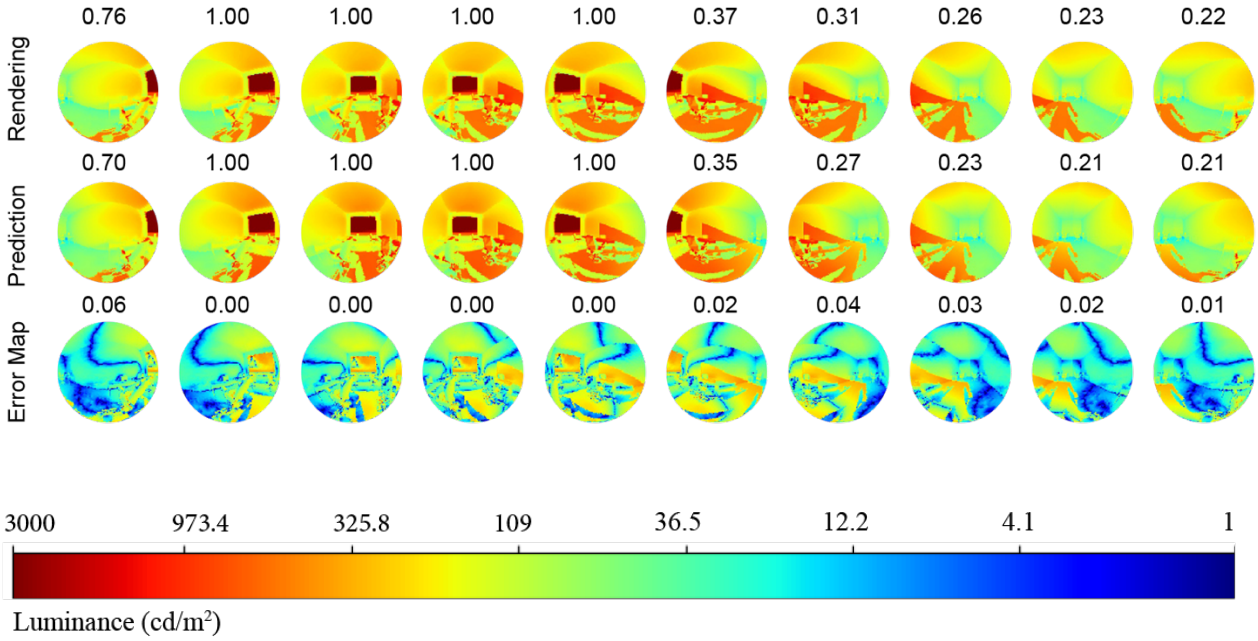


Figure 60. Comparison of Radiance rendered and DNN predicted fisheye images over a 360° view in 36° increments for one of the highest error cases (January 21th 11:30 am, direct and diffuse irradiances of 553 and 87 W/m$^{2)}$. The ground truth and predicted fisheye images are labeled with the DGP values. The error maps are labeled with the absolute difference in DGP values. All images are shown in falsecolor with a logarithmic scale.

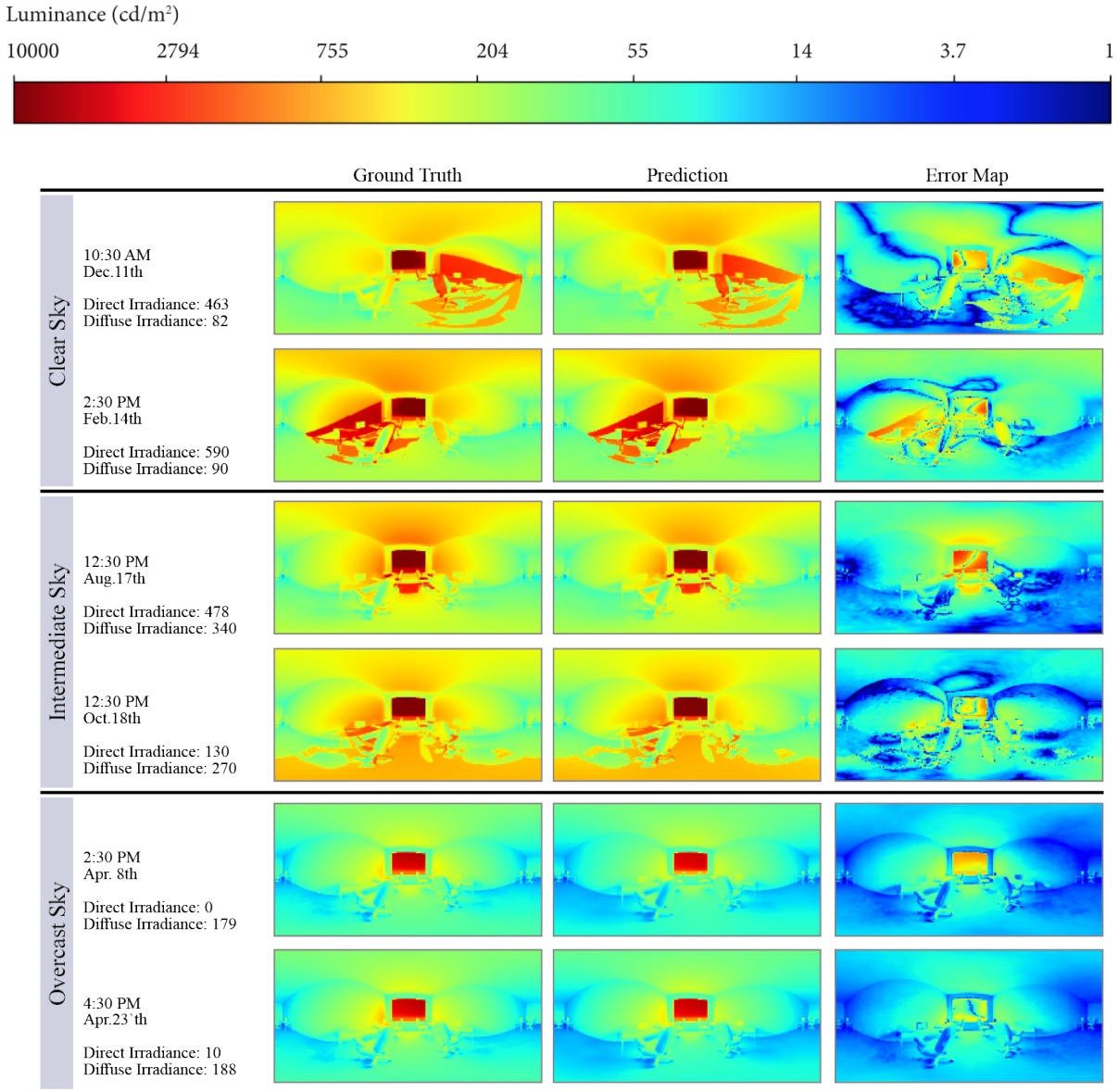


Figure 61. Six exemplary test cases with different sky types are displayed in falsecolor with a logarithmic scale.

6.1.4 *Further Shortening the Data Collection Period*

In this section, datasets of shorter collection periods, training set 1-2 (2 weeks) and training set 1-3 (1 week), are compared against the training set 1-1 (1 month) and training set 0. Both training

sets (1-2 and 1-3) are selected from March to keep consistent with the previous evaluations. Similar to the previous analyses, per-pixel errors, visual discomfort metrics, perceptual performance metrics, and falsecolor analysis with image subtraction operations are used to evaluate the differences between the predicted and ground truth luminance maps.

Figure 62 shows the comparison between predictions generated with different training sets and the ground truth, measured by Log_{10} MSE and Log_{10} RER error. The results show that using training sets of shorter collection periods will greatly reduce the prediction accuracies.

Figure 63 to Figure 65 demonstrate the differences in prediction accuracies with human-centric metrics, DGP and RAMMG. Figure 63 shows the relationship between the training sets and the DGP MSE and RAMMG MSE errors. The results align with the findings of Figure 62, and show that the predictions generated with training sets of shorter collection periods have increased errors. Figure 64 shows the correlations between the predicted and ground truth DGPs. The results confirm that DGP predictions generated with the training sets of shorter collection periods have weaker agreements with the ground truth, with r^2 of 0.770, 0.945 and 0.990 for training set 1-3, 1-2 and 1-1, respectively. DGP predictions generated with training set 1-2 and 1-3 have significant errors occur at high ranges, indicating that the model has diminished capability in accurately predicting the high luminance or high contrast scenes. Figure 65 shows a comparison between the predicted and the ground truth panoramas in falsecolor with a logarithmic scale. This particular case (January 20th 1:30 pm, direct and diffuse irradiances of 270 and 106 W/m^2) is selected as it has the lowest accuracy among all test cases. The result shows that: 1) the difference between the predicted luminance maps generated with training set 1-1 is imperceptible from the ground truth ones, without the help of the error maps; 2) discrepancies exist between the predictions generated

with training set 1-2 and 1-3 and the ground truth, with poorly predicted shadows, sun penetrations, and pixel luminances around the ceiling.

Figure 66 shows the per-view-angle comparisons under the same test lighting condition (the worst case). Each fisheye image is labeled with the calculated DGP value. All images are shown in falsecolor with a logarithmic scale. The DGP values of each view direction are compared in the bottom diagram, with colored zones represent four DGP scales. The results show that: 1) the predicted fisheye images generated with the training set 1-1, closely matches the ground truth images in all directions; 2) the differences between the predicted fisheye images, generated with the training set 1-2 and 1-3, and the ground truth images can be clearly noticed; 3) the DGP predictions generated with training set 1-1 have a strong agreement with the ground truth values, while the agreement is weaker between predictions, generated with training set 1-2 and 1-3, and the ground truth; 4) predicted DGPs all fall in the same zone as the ground truth DGPs when training with training sets 1-1 and 1-2, not with training set 1-3; and 5) the differences between the predicted and the ground truth DGPs are more significant in high contrast scenes (view direction 72° – 144°).

Overall, the results show that the minimum continuous data collection period for accurate predictions is one month. Further reducing the data collection period to two weeks or one week results in significant errors and may be misleading in a design process.

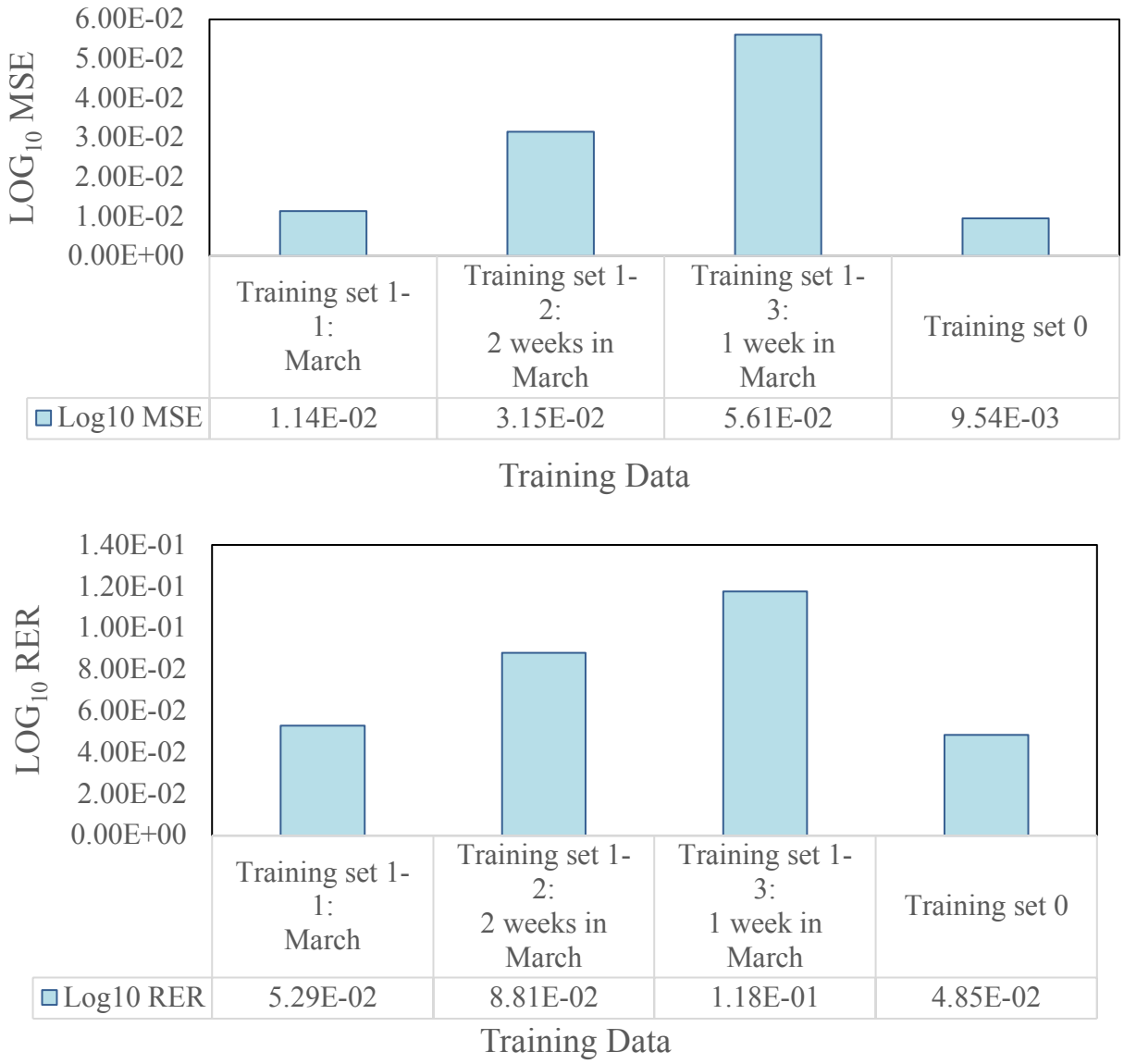


Figure 62. Prediction errors generated with different training sets, measured by (Top) Log_{10} MSE and (Bottom) Log_{10} RER error.

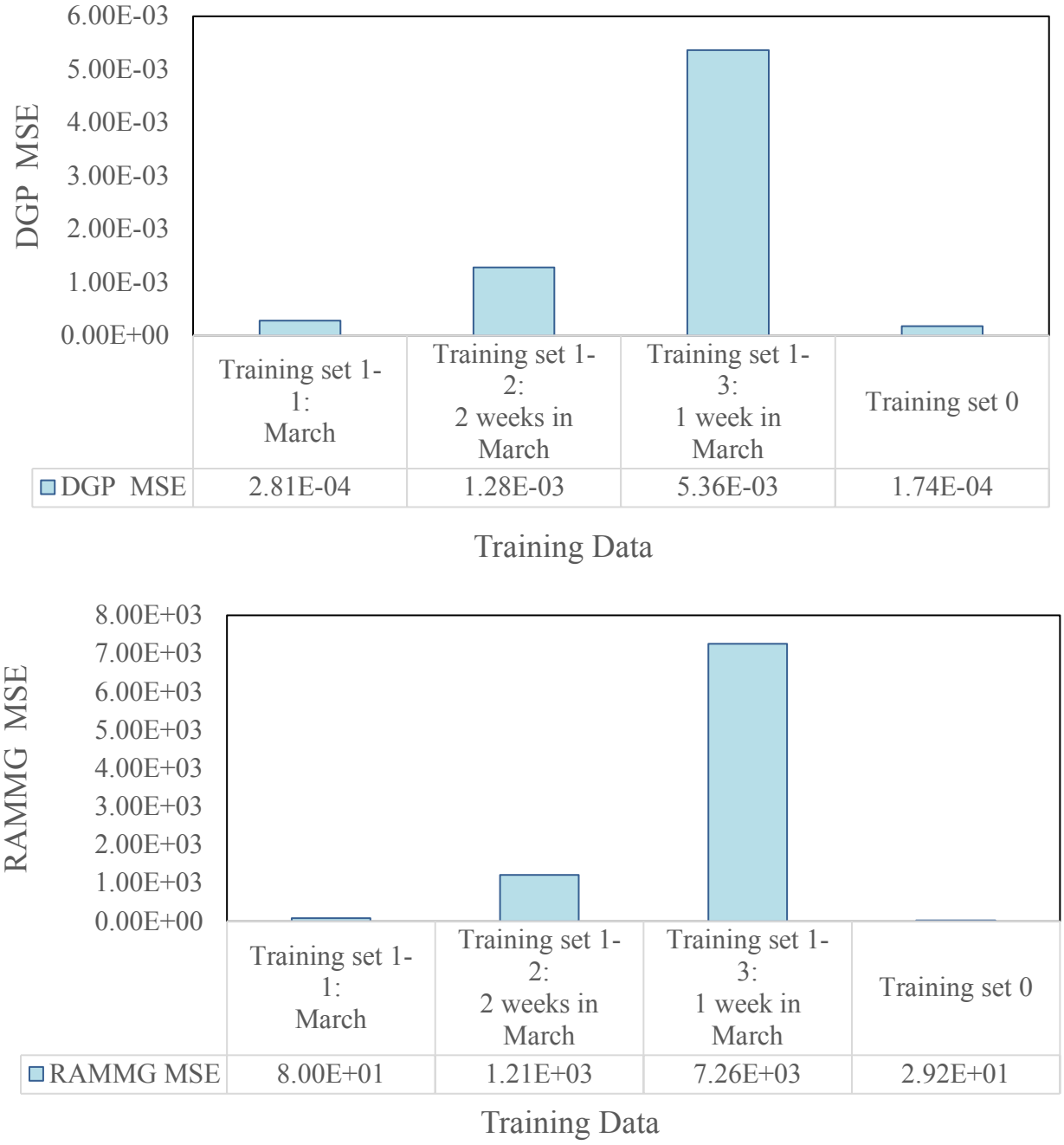
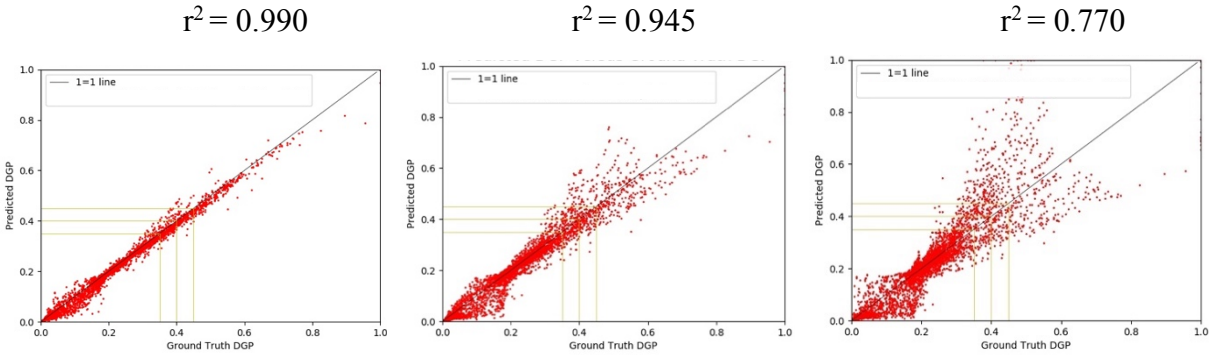


Figure 63. (Top) DGP MSE and (Bottom) RAMMG MSE between predictions generated with each training set and the ground truth.



Training set 1-1: March | **Training set 1-2: 2 weeks in March** | **Training set 1-3: 1 week in March**

Figure 64. Scatter plots comparison between DGP values calculated using the predicted fisheye images generated with three training sets and the ground truth ones.

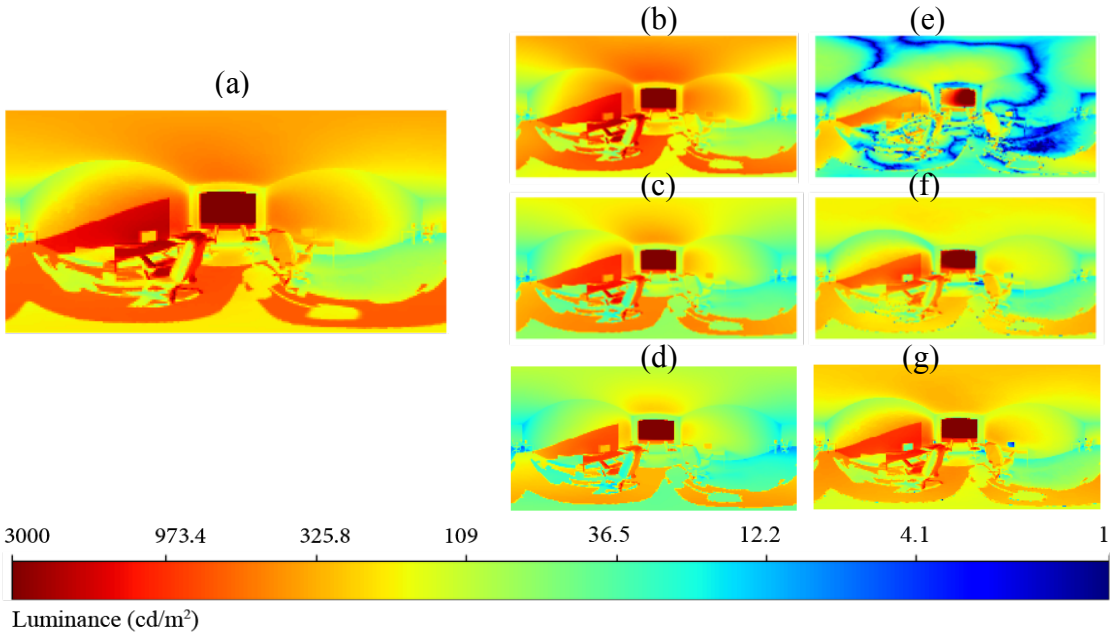


Figure 65. Worst-case (January 20th 1:30 pm, direct and diffuse irradiances of 270 and 106 W/m^2) falsecolor image comparison: (a) ground truth, (b) training set 1-1: March, (c) training set 1-2: 2 weeks in March, (d) training set 1-3: 1 week in March. (e), (f), and (g) are error maps of (b), (c), and (d).

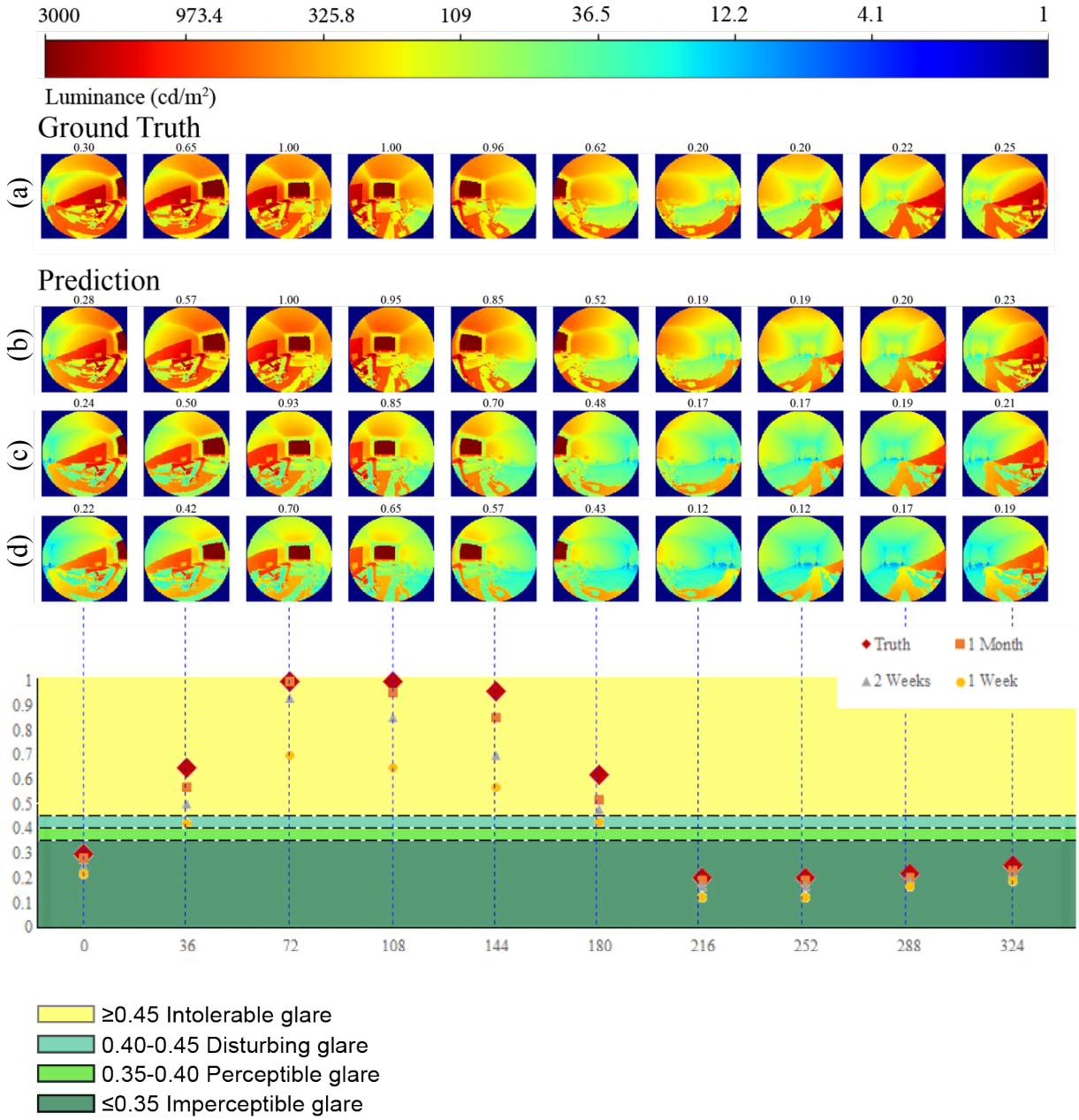


Figure 66. Worst-case (January 20th 1:30 pm) falsecolor fisheye images and DGPs comparison: (a) ground truth, (b) training set 1-1: March, (c) training set 1-2: 2 weeks in March, (d) training set 1-3: 1 week in March.

6.2 TRAINING WITH DATA COLLECTED FROM MULTIPLE SHORT PERIODS

Training with data collected from multiple short periods is expected to have increased prediction accuracy compared with the training with data collected from a single continuous period, due to the expanded distribution of the sun locations and sky conditions. Table 4 shows the proposed training sets and sample selection strategies in this study. Similar to the previous study in Section 6.1, the training set with 200 evenly distributed samples (training set 0), is utilized as the benchmark, against which all others are compared to in the subsequent analyses. The rest training sets contain samples collected from multiple short periods around the equinoxes and solstices. Equinoxes and solstices are selected for two reasons. Firstly, they cover the outlines of a sun path diagram (Figure 67), sun position variance is found to have the most significant impact on the prediction accuracies in previous analyses. Secondly, equinoxes and solstices are key time points which architects and lighting professionals typically generate luminance maps, when approximating a scene's annual lighting performances. Therefore, using these samples to generate year-round luminance maps is more acceptable. After a preliminary study, the investigation starts with the training set 2-1 containing samples collected from 4 short periods, each period consists of 3 days around the spring equinox (March 20th), summer solstice (June 21st), fall equinox (September 23rd) and winter solstice (December 21st). The second training set (2-2) excludes the samples of the fall equinox but keeps those of the spring equinox as they have similar sun positions. The duration of each period selected from equinoxes and solstices in the third training set (2-3) is reduced to 1 day. The total amount of samples in the training sets 2-1, 2-2 and 2-3 are 144, 108 and 48, respectively, all less than 200 samples in the benchmark training set.

To illustrate the lighting characteristics of all the training sets, Figure 67 shows the distribution plots of selected samples over the four-dimensional lighting domain (sky direct and

diffuse irradiances and sun azimuth and altitude angles). Red dots represent samples of each training set (i.e., training set 2-1,2-2, and 2-3) while green dots represent samples of the benchmark dataset (i.e., training set 0). The figure illustrates that compared to the benchmark dataset, training sets 2-1, 2-2, and 2-3 cover fewer sky conditions, especially skies with high diffuse irradiances; and have highly concentrated locations of the sun, and thus, cover less sun altitude and azimuth angles.

Table 4. Proposed training sets and sample selection strategies.

Training Set Number	Total Duration	Number of Periods	Training Samples	Total Number of Samples
0	1 year	/	Evenly distributed over the entire year selected using K-means clustering algorithm	200
2-1	12 days	4	**Around Equinoxes:** March 19th, 20th, 21st, September 22nd, 23rd, 24th **Around Solstices:** June 20th, 21st, 22nd, December 20th, 21st, 22nd	144
2-2	9 days	3	**Around Equinoxes:** March 19th, 20th, 21st **Around Solstices:** June 20th, 21st, 22nd, December 20th, 21st, 22nd	108
2-3	4 days	4	**Around Equinoxes:** March 20th, September 23rd **Around Solstices:** June 21st, December 21st	48

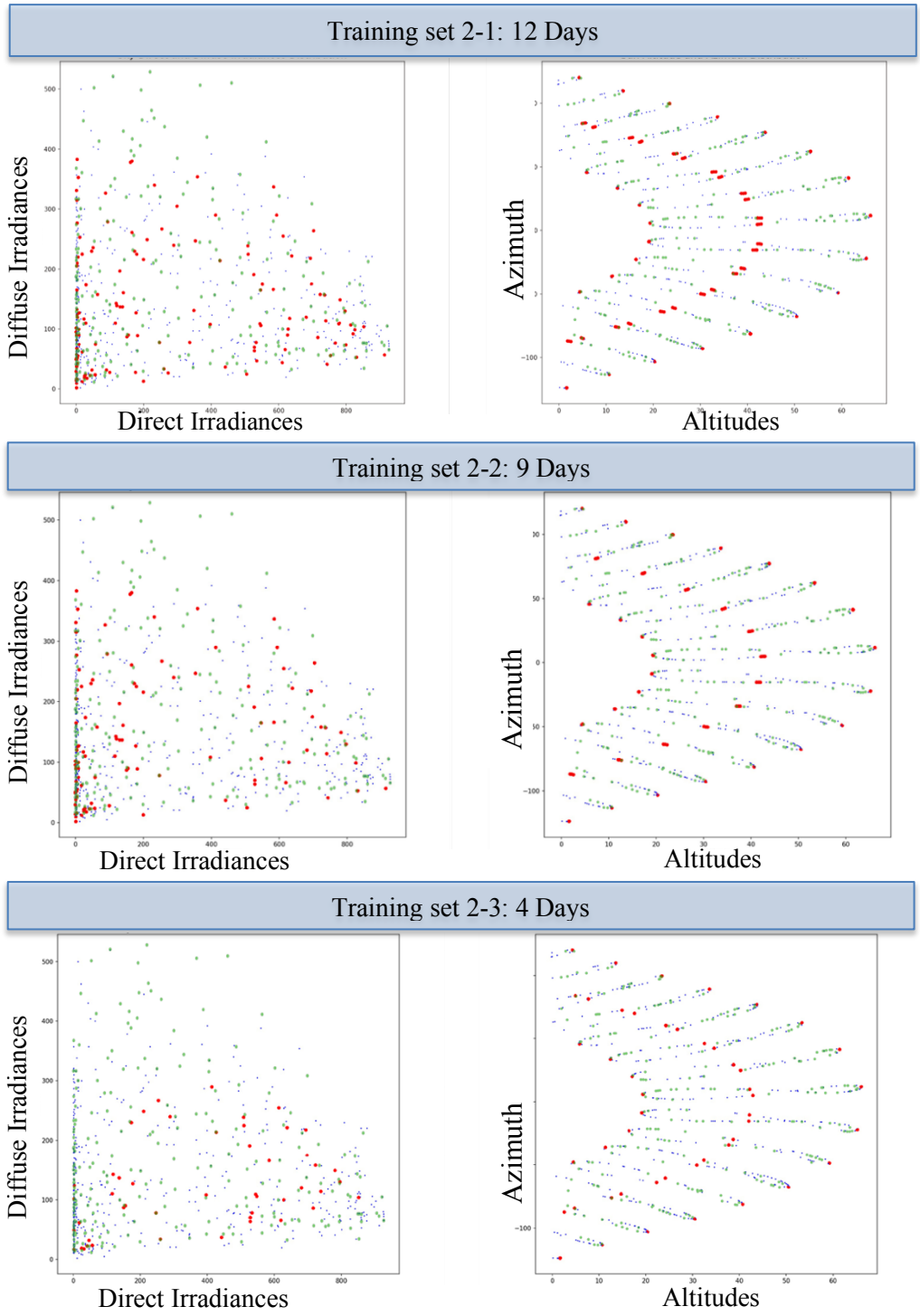


Figure 67. Distribution plots of selected samples over the four-dimensional lighting domain: (Left column) distribution plots over sky condition parameters (direct and diffuse irradiances), and (Right column) over sun position parameters (azimuth and altitude angles). Red dots represent samples in each training set while green dots represent samples of the benchmark dataset (training set 0).

6.2.1 *Evaluation of Prediction Results*

The results of the predictions generated with different training sets are evaluated by per-pixel errors (Log_{10} MSE, Log_{10} RER), human-centric metrics (DGP MSE, RAMMG MSE), and falsecolor images.

Figure 68 presents the comparisons between predictions generated with different training sets. In all four comparisons, predictions generated with the training set 2-1 containing 12-days samples have the lowest errors while those with the training set 2-3 containing 4-days samples have the highest errors. This indicates that the prediction errors increase with less training samples (or shorter data collection periods). The most important finding is that the training sets 2-1 and 2-2 have similar or even less per-pixel errors as achieved by the training set 0. This suggests that the method can have improved prediction accuracies with less training samples through better sample selection strategies. For DGP and RAMMG predictions, only the training set 2-1 achieves the same accuracy as the training set 0. Prediction errors of the training set 2-3 are higher than the benchmark errors measured by all four metrics.

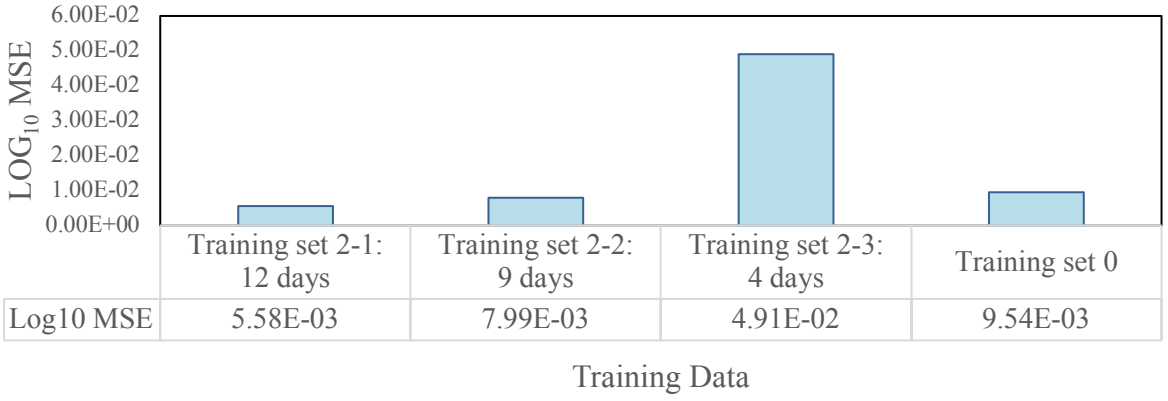
Figure 69 illustrates the correlations between DGP values calculated using predicted and ground truth luminance maps. The results show strong (r^2 = 0.99 and 0.98) agreements between predicted DGPs, generated with the training sets 2-1 and 2-2, and the ground truth DGPs, with a few errors (1% of the total test cases) at higher ranges, when the sun direct appears in the field of view through the south-facing window. The DGP predictions using the training set 2-3 have a relatively weak (r^2 = 0.89) agreement with the ground truth ones, and exhibit moderate noise at higher ranges.

Figure 70 shows the worst-case comparison of predicted luminance maps against ground truth ones, in falsecolor with a logarithmic scale. The test case occurs at the winter noon (January 20th

1:30 pm, direct and diffuse irradiances of 270 and 106 W/m^2) and has the lowest accuracy among all test cases with the sun appearing through the window. The figure illustrates that: 1) it is hard to differentiate between predicted luminance maps generated with the training set 2-1 and the ground truth images without the help of the error maps; 2) minor differences can be found between the predicted luminance maps generated of the training set 2-2 and ground truth ones, 3) the differences between the predicted luminance maps generated with the training set 2-3 and the ground truth one is perceptible, and may potentially lead to different design decisions (the scene is too bright overall); 4) the error maps reveal that in all cases the highest errors occur around the window and in the areas exposed to direct sunlight. Figure 71 shows the per-view-angle DGP comparison for the same test lighting condition (worst-case). The result illustrates that: 1) subjects can experience various visual comfort levels across different view directions of the same scene; 2) with dataset 2-1 and 2-2, the neural network accurately predicts the overall luminance environment, with the predicted DGP values closely matching those of the ground truth ones in all directions; and 3) the more significant errors occur towards the window directions (view directions of 72°-180°); 4) in these view directions (view directions of 72°-180°), the model with the training sets 2-1 and 2-2 slightly under-predicts the DGP values, due to underestimation of the luminance of the sun through the window; the predicted DGPs still fall in the same zones as the ground truth DGPs; 5) the most significant errors are with the training set 2-3, and the predicted DGPs fall into the different zones as the ground truth DGPs; 6) all training sets generate accurate predictions, when there is no direct sunlight in the field of view (view directions of 0°, 216°, 252°, 288°, 324°).

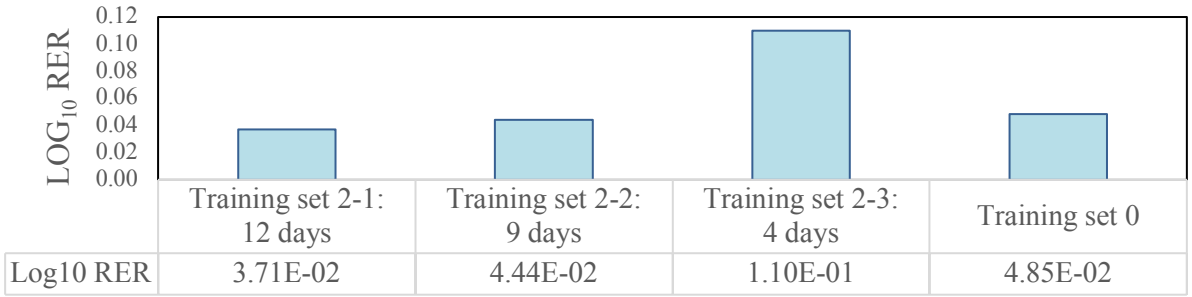
Overall, the training sets 2-1 with 12-days samples and 2-2 with 9-days samples succeed in generating luminance maps that are very close to ground truth along all view directions. Prediction

results of the training set 2-3 with 4-days samples yield less satisfactory results, and are thus, excluded in the subsequent analyses.



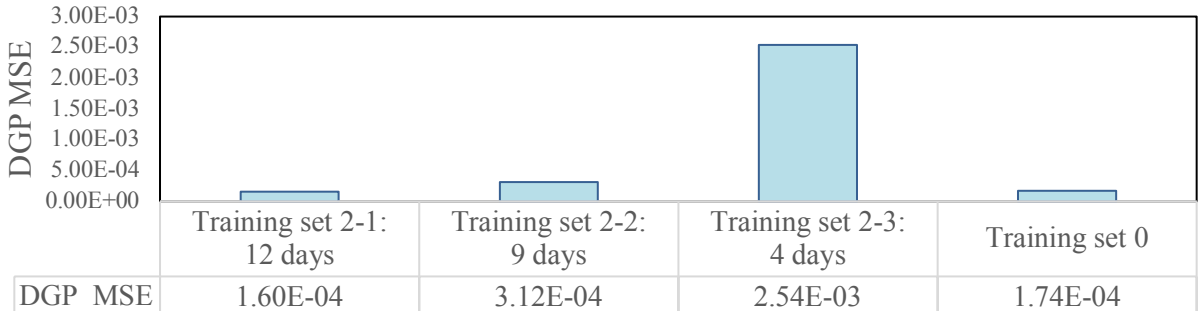
	Training set 2-1: 12 days	Training set 2-2: 9 days	Training set 2-3: 4 days	Training set 0
Log10 MSE	5.58E-03	7.99E-03	4.91E-02	9.54E-03

Training Data



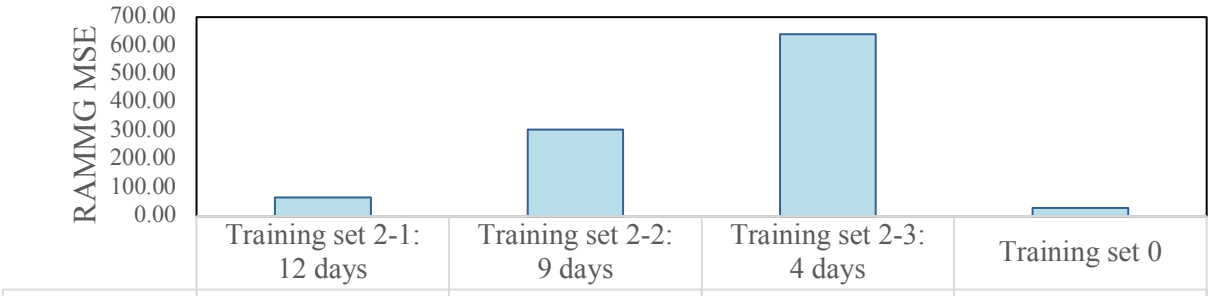
	Training set 2-1: 12 days	Training set 2-2: 9 days	Training set 2-3: 4 days	Training set 0
Log10 RER	3.71E-02	4.44E-02	1.10E-01	4.85E-02

Training Data



	Training set 2-1: 12 days	Training set 2-2: 9 days	Training set 2-3: 4 days	Training set 0
DGP MSE	1.60E-04	3.12E-04	2.54E-03	1.74E-04

Training Data



	Training set 2-1: 12 days	Training set 2-2: 9 days	Training set 2-3: 4 days	Training set 0
RAMMG MSE	6.60E+01	3.04E+02	6.40E+02	2.92E+01

Training Data

Figure 68. Comparison of predictions generated with different training sets, evaluated by (Top) Log_{10} MSE, (Upper Middle) Log_{10} RER, (Lower Middle) DGP MSE, and (Bottom) RAMMG MSE.

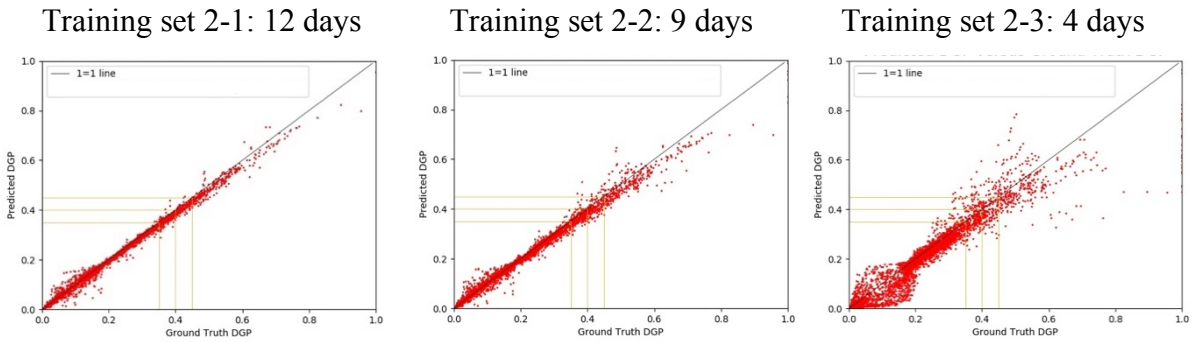


Figure 69. Scatterplots show the correlation of DGP values calculated using ground truth and predicted fisheye images. Red dots represent DGP pairs.

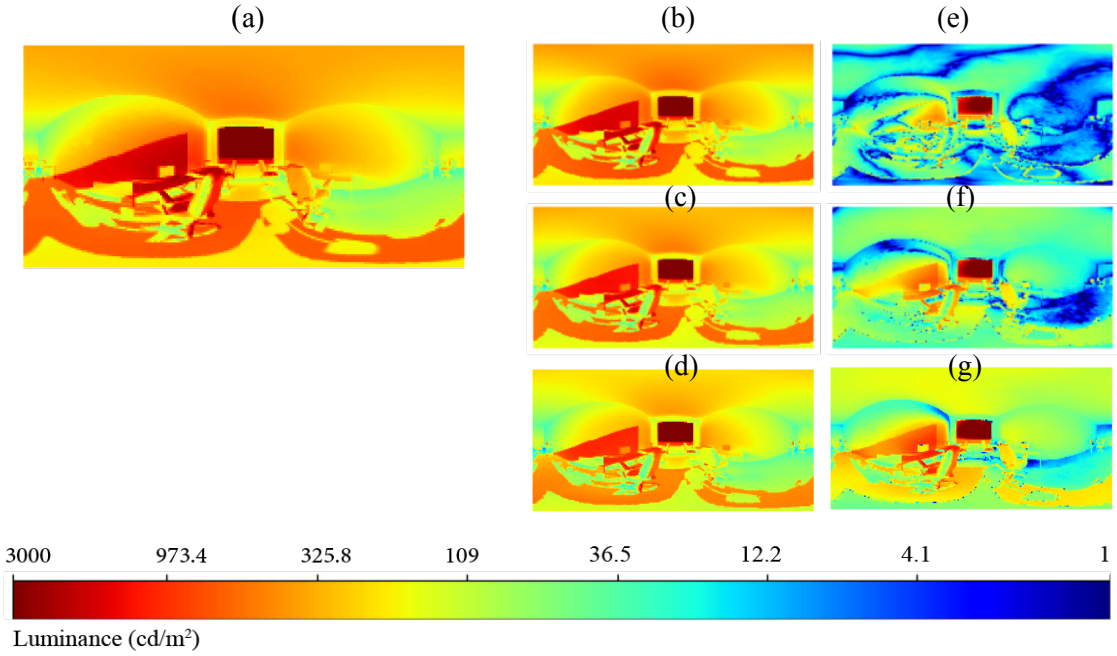


Figure 70. Predicted luminance maps are compared against ground truth ones under the worst-case test lighting condition (January 20th 1:30 pm, direct and diffuse irradiances of 270 and 106 W/m^2), in falsecolor with a logarithmic scale: (a) ground truth, (b) training set 2-1: 12 days, (c) training set 2-2: 9 days, (d) training set 2-3: 4 days. (e), (f), and (g) are error maps of (b), (c), and (d).

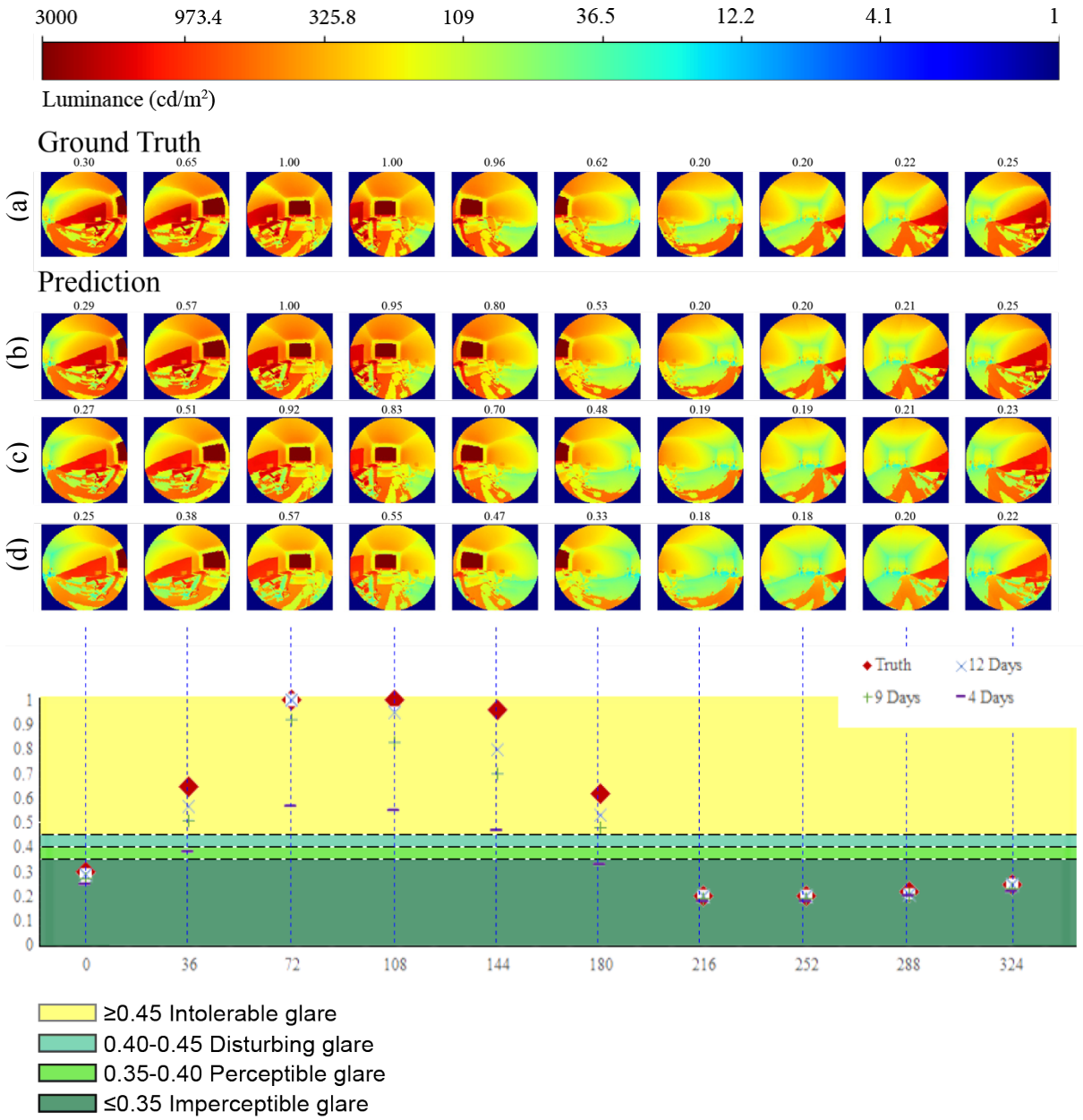


Figure 71. Per-view-angle DGP comparison under the worst-case test lighting condition (January 20th 1:30 pm, direct and diffuse irradiances of 270 and 106 W/m^2): a) ground truth, b) training set 2-1: 12-days samples, c) training set 2-2: 9-days samples, and d) training set 2-3: 4-days samples.

6.2.2 *Evaluations with Falsecolor Analysis*

This section evaluates the abilities of the training sets 2-1 and 2-2 in predicting accurate luminance maps under various lighting conditions. Six representative samples under clear, intermediate and overcast sky conditions are selected from the test case, with predicted luminance maps compared against Radiance renderings (Figure 72 and Figure 73). All images are shown in falsecolor with a logarithmic scale, and the error maps demonstrated the absolute differences between the predicted and ground truth luminance maps. The results show that the proposed method faithfully produces luminance maps under various novel lighting conditions. Without the help of error maps, the predicted panoramic luminance maps in both Figure 72 and Figure 73 are visually imperceptible from ground truth falsecolor images with a 1-10,000 cd/m^2 range. This implies that the same design decisions can be made with the proposed method and the conventional rendering methods, but the former requires orders of magnitude less of calculation time. The error maps reveal that: 1) higher errors in each case occur in the window area, and 2) the training set 2-2 with 12-days sample generated fewer errors than training set 2-3 with 9-days samples.

Overall, the study obtains high-quality results illustrating the ability of the proposed method to accurately predict panoramic luminance maps under various lighting conditions with a smaller set of samples (12 days and 9 days).

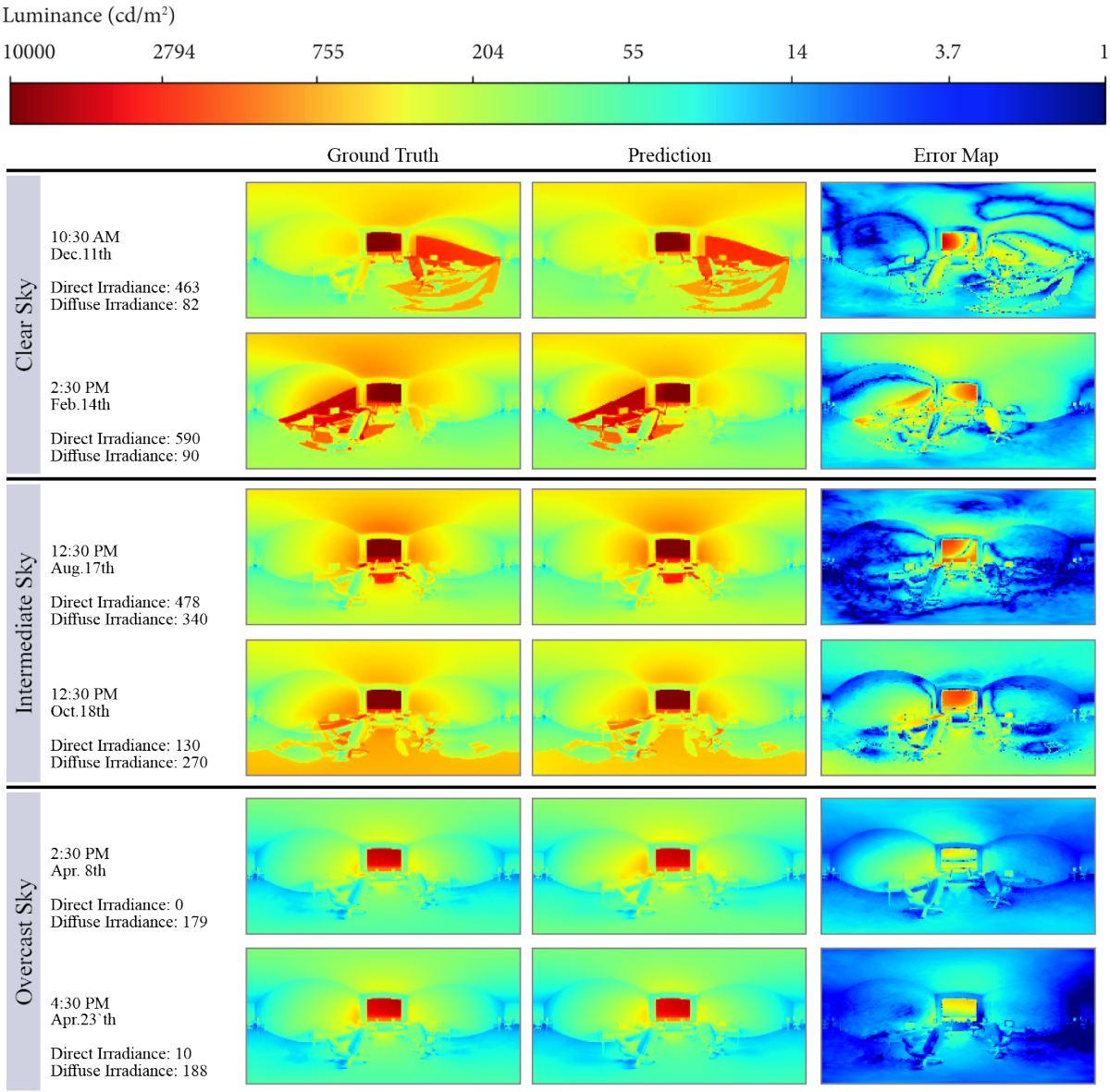


Figure 72. Six representative cases under different sky conditions, generated with training set 2-1 containing 12-days samples, are displayed in falsecolor with a logarithmic scale.

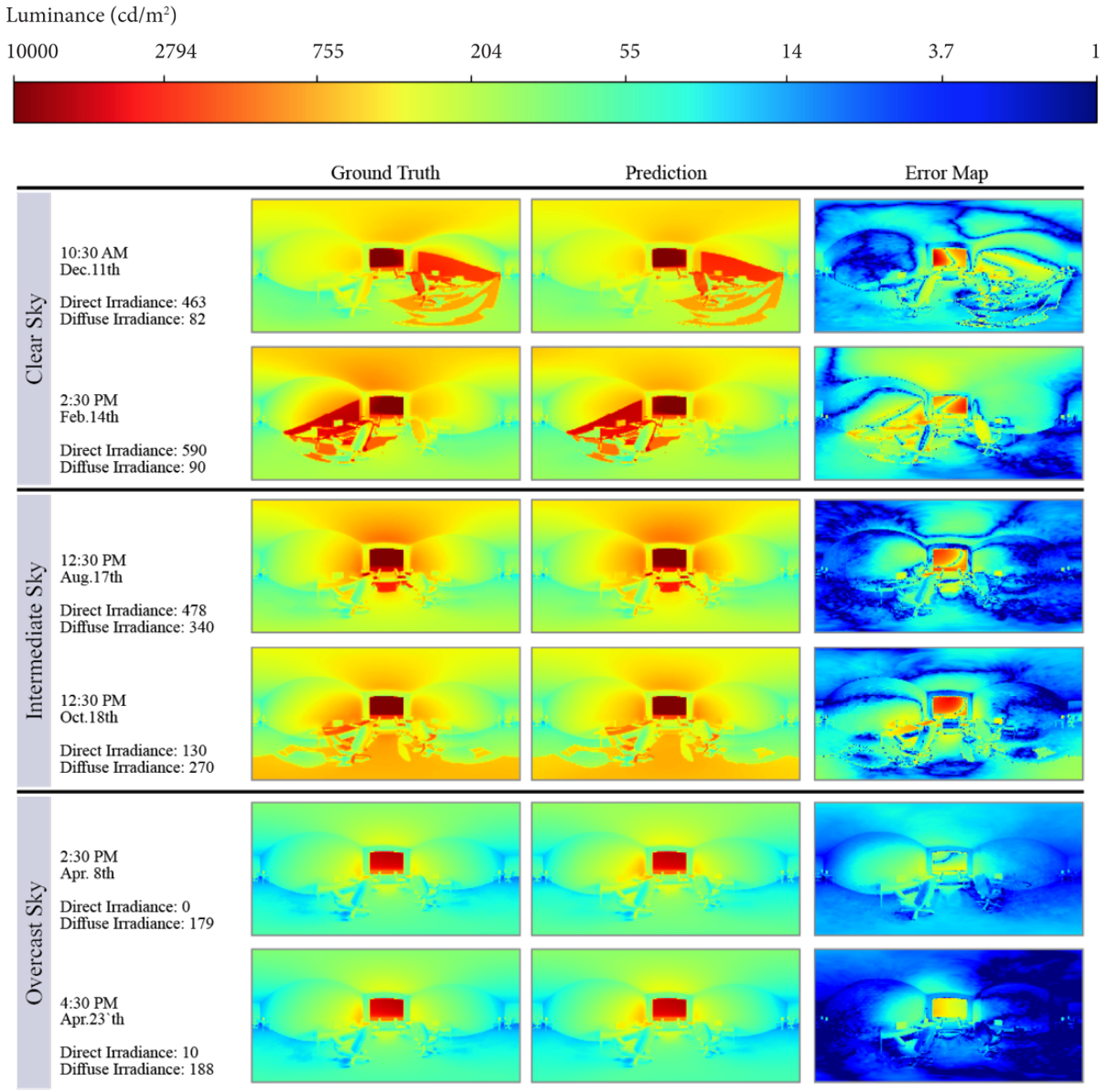


Figure 73. Six representative cases under different sky conditions, generated with the training set 2-2 containing 9-days samples, are displayed in falsecolor with a logarithmic scale.

6.3 DISCUSSION

In this chapter, a sensitivity analysis is performed to inspect the relationship between the collection periods of the training data and the accuracy of the predictions, in order to understand the minimum and optimum data collection periods for generating or capturing annual luminance maps. Multiple training data sets including samples from one continuous period (one month, two weeks, and one week) and multiple short periods (twelve days, nine days and four days selected around equinoxes and solstices) are tested. The predicted luminance maps are compared against Radiance renderings with a series of quantitative and qualitative metrics including per-pixel error, falsecolor image analysis, visual comfort, and perceptual performance metrics.

The results show that the method achieves improved efficiency and applicability through a better selection of training samples. Compared to the previous study, which utilizes evenly distributed training samples over the entire year, this study achieves comparable accuracy with shorter collection periods and smaller datasets. The analyses suggest that it is possible to shorten the data collection period to one month. This makes annual luminance-based evaluations using HDR captures more feasible. When applying the method to real HDR measurements, the user can capture one-month luminance maps and predict the rest, instead of capturing for the entire year. Another important finding is that the model can faithfully predict annual panoramic luminance maps under various sky conditions with only 2.5 % of 4379 samples from three short periods each contains three days selected during the spring equinox, summer solstice, and winter solstice (9 days in total). This helps further accelerate the renderings of annual luminance maps. The findings also reveal that compared to sky conditions, sun position parameters play a more important role in generating accurate predictions. The optimum time to collect the data in Seattle is around

equinoxes because data from these months can better represent variance of sun positions. This is useful information in refining the sampling strategies for future explorations.

Chapter 7. CONCLUSION AND FUTURE WORK

7.1 DISSERTATION SUMMARY

This dissertation provides a framework for predicting annual panoramic luminance maps of indoor spaces under varying sun and sky conditions, from a limited number of HDR images, using deep neural networks. The DNN approach accelerates the generation of panoramic luminance maps, and only requires a few seconds of computing time during evaluation. The accuracy, applicability, and usefulness of the proposed DNN model is examined through comparisons with Radiance RPICT renderings using per-pixel error metrics, visual comfort, and contrast metrics. Architects, lighting designers, and researchers can effectively use the workflow to perform multi-view annual lighting analyses.

Chapter 2 reviews the state-of-art research and practice in daylight simulations, measurements, and evaluations, and explains the need for an efficient and accessible annual luminance-based daylighting evaluation method. In this chapter, physically-based daylighting simulations and HDR photography techniques are discussed as the two methods for generating and measuring the luminous environments. Evaluation metrics are discussed for assessing the daylighting performances. A better understanding of these methods and metrics highlights the importance of luminance-based evaluations, demonstrates the need for accelerating the computation of annual luminance evaluations, and provides key information for designing the research methodology of the dissertation. The current best practices in accelerating annual daylighting evaluations are reviewed and our solution of utilizing machine learning to create predictive models from a subset of observed data during daylight hours is introduced in concept.

Chapter 3 describes the context of the machine learning techniques employed in the dissertation. It introduces the conventional feedforward neural networks and convolutional neural

networks as the two specific machine learning techniques for image-related tasks, and examines regularization techniques for improving training qualities. Understanding the underlying theory, assumptions, best practices, and limitations of these machine learning techniques is necessary to formulate the research framework. The most-current studies that utilize machine learning for lighting-related tasks are reviewed, and the challenges of simulating HDR images of daylit indoor scenes with physical accuracy and feasible computing times are discussed.

Chapter 4 to Chapter 6 documents the three phases of developing the methodology of the dissertation, where each one builds upon the results of the previous chapter. Chapter 4 introduces the initial study performed to build the research framework and test the applicability and accuracy of the workflow. The results show by comparison that the proposed machine learning model can predict annual luminance maps of single perspective views from 5% of annual HDR images, as accurate as Radiance does, for scenes within a moderate luminance range ($0\text{-}3\times10^{4}$). Chapter 5 increases the model's prediction accuracies and dynamic luminance range ($0\text{-}10^{8.3}$) under various sky conditions through several algorithmic refinements. It extends the image format to panoramic luminance maps (with 360° horizontal and 180° vertical field of view), which provide the ability to evaluate subject's visual experience over multiple viewing directions. In Chapter 6, the focus shifts to improve the efficiency and applicability of the method through a better selection of training samples. This chapter presents a sensitivity analysis performed to inspect the relationship between the collection period of the training data and the accuracy of the predictions. The results reveal that the minimum continuous data collection period for generating accurate annual luminance maps is one month hourly data during daylight hours. The months with greater variance in sun positions have better predictability power. The analyses show that solar position plays a more important role compared to sky parameters on prediction accuracy.

The minimum sample size of generated/captured data is reduced to 2.5% of annual images through selecting samples from three short periods each contains three days around the spring equinox, summer solstice, and winter solstice (9 days in total).

In this final chapter, the benefits of a fast and accurate annual luminance-based evaluation tool are discussed.

7.2 CONTRIBUTIONS

The dissertation addressed four primary objectives outlined in Chapter 1. Table 5 lists the major contributions. The accuracy and training time reported in Table 5 depend on many factors such as the speed of GPU, the room model complexity, the simulation parameters, and the machine learning parameters. Although these numbers may be different for other scenes and computer settings, they show the speed, accuracy, and the capabilities of the developed machine learning methodology.

Table 5. Main Contributions

Contribution		*Created a deep learning framework to conduct architectural luminance predictions (Objective 1).*	*Developed a robust workflow that accurately and efficiently predicts annual panoramic luminance maps from a small number of HDR images (Objective 2).*	*Found and explained the minimum and optimum data collection period and the minimum sample size (Objective 3). Studied the impact of the weather and seasonal changes on the prediction accuracies (Objective 4).*	
Training Sample Size		5% of annual renderings evenly distributed over the light domain	5% of annual renderings evenly distributed over the light domain	1 month	2.5% of annual renderings (9 days samples during Spring Equinox, Summer Solstice, and Winter Solstice)
Training Time (hours)		2	0.5	0.5	0.4
Maximum Luminance Range (cd/m²)		$3\times10^{+4}$	$10^{+8.3}$	$10^{+8.3}$	$10^{+8.3}$
Accuracy	*Log10 MSE*	1e-05	9.54e-03	1.14e-02	7.99e-03
	Log10 RER	0.063	4.85e-02	5.29e-02	4.44e-02
	DGP r²	0.99	0.99	0.99	0.98
	RAMMG r²	/	0.99	0.99	0.96
	Falsecolor	Predictions are visually imperceptible from ground truth results in falsecolor in a 1-10,000 cd/m² range without the utilization of error maps			

The contributions are beneficial to the architectural lighting community both on the practical and theoretical level.

On the practical level, the proposed workflow for generating annual luminance maps of indoor scenes from a limited set can be immediately useful to architects and lighting designers. Annual panoramic luminance maps can be faithfully predicted with 2.5% of annual renderings from three short periods (9 days in total), and only requires 30 minutes of computing time during training, and a few seconds during evaluation. Annual luminance-based simulation using this method is magnitudes faster (~40 times speedup) than conventional simulation methods with comparable accuracy examined by per-pixel errors, falsecolor image techniques, and visual comfort and perceptual metrics.

Making annual luminance-based simulations more efficient and accessible will foster more extensive usage of luminance-based evaluations considering temporal and spatial variance into the design processes. Compared with lighting values of point-in-time simulations (e.g., illuminance, luminance) which are most commonly utilized in current lighting analyses, annual panoramic luminance maps provide a comprehensive evaluation of the luminous environment over the course of the entire year under various sky conditions. The predicted panoramic luminance maps can be transformed into fisheye projections for any view direction. The occupants' visual experience in 360º can be evaluated instead of a commonly used fixed-view direction. With the workflow demonstrated here, practitioners and researchers can incorporate human-centric evaluations - of long-term and over multiple view directions - in the design processes without the lengthy computing processes.

The technique opens new avenues for many innovative applications. Architects and lighting designers can utilize the method to create annual animations of daylight scenes. This will help

designers to recognize and eliminate potential visual discomfort sources, understand the perceptual experience throughout the dynamic lighting conditions, and present and convey design ideas to other building professionals and clients. The generated annual panoramas can be easily imported to other visualization media (such as immersive media and animations) for both educational and commercial purposes. As a result, design of better luminous environments that support occupant visual comfort, productivity, and well-being will be promoted.

On the theoretical level, the proposed research demonstrates a novel multidisciplinary framework to apply deep learning to architectural daylighting studies. To the best of the author's knowledge, there are no previous studies that utilize deep learning to conduct architectural luminance predictions. There is great potential to apply emerging data science technologies to lighting research studies. Other major contributions of the dissertation on a theoretical level are summarized as follows:

- The study developed a DNN model that separates the direct and indirect contributions from the sun to reconstruct sharp shadows and accurate sun penetrations. The design of the model can be applied to many other studies with high dynamic scenes. The findings reveal that compared to sky conditions, sun position parameters play a more important role in generating accurate predictions. This is useful information in refining the sampling strategies for future explorations. The performance of the model is quantitatively investigated to meet the scientific accuracy requirement for applications in architectural lighting through a series of quantitative and qualitative metrics. This is useful for fields requiring physically-accurate lighting information (beyond plausible visual appearances).

- Different from illuminance-based metrics, there are only a few annual luminance-based metrics, because researchers and practitioners are less motivated to integrate annual luminance-based metrics into the design evaluation process due to the time-consuming rendering processes. Making annual luminance-based simulations more accessible will enable further development of luminance-based metrics over multiple view directions for daylighting studies.
- The results demonstrate that the minimum data collection is one-month hourly generated data and the optimum time for collection is during Equinox because data from these months can better represent the variance of sun positions.
- This method takes the first step towards generating annual luminance data from a limited number of HDR photographs taken in-situ. The user can utilize one-month hourly generated renderings during Equinoxes to predict the annual luminance maps. It is possible to apply this to HDR field measurements to shorten the capturing period, but further research is needed to address the complications of data collection in the field.
- The dataset of high-quality annual HDR panoramic luminance maps and the predictive model is available to third parties (https://github.com/yueAUW/neural-daylighting) to enable reproducibility and future explorations.

7.3 FUTURE WORK

Beyond the workflow and analysis techniques presented in the dissertation, this research opens up several research directions:

- In this study, we deliberately chose to focus on synthetic data for algorithm development and validation. The next step is to apply the method to real-world data

capture with HDR photographs. The required data collection period is determined as one month (around the equinoxes) to facilitate the feasible gathering of HDR photography-based training sets. It will be valuable to perform a comparison between captured annual luminance maps and the predicted ones generated from one-month HDR photographs of an existing scene. It is clear that photography based approach will require several post-processing and calibration phases in comparison to a synthetic image data. Such an approach enables quantitative analysis of daylit environments without the labor-intensive scene modeling and simulation processes, or the long-term measurement protocols. It can be used in post-occupancy evaluation studies to understand whether or not the luminous environment of an existing building support the original design intent and long-term occupant comfort and satisfaction. The evaluation results can be used as benchmarks for retrofits. As the evaluation process requires less labor and eventually less cost, it can be requested more often by consultants, building owners, employers of the occupants, and occupants themselves.

- The model performance and accuracy can be compared to multi-phased daylight coefficient methods. The multi-phased daylight coefficient methodology is the prevalent workflow to create long term luminance simulations. It can take a few days to generate annual data. The proposed methodology in this study can provide results within hours. However, it is important to test the capability of the develop DNN model for modeling complex fenestration systems (e.g., venetian blinds).
- The developed DNN approach could be tested with a database of rooms with different locations, room geometry, window size, and orientations. This would test the robustness of the model for discrete cases.

- A general DNN model can emerge if the algorithm could be extended to learn from multiple cases. This may further reduce the sample size required for generating annual luminance maps.
- A Graphic User Interface (GUI) will be a useful addition to the current workflow, and foster a wider usage of the proposed method. A successful GUI will allow the user to select the target 3D model and the location of the room, and generate annual panoramic luminance maps automatically.
- The study lays the foundations for many potential machine learning studies for architectural luminance predictions. The proposed method can be combined with other rendering accelerating techniques. For example, as introduced in Chapter 3, it is possible to create high-quality images from noisy renderings with reduced sample rate using a machine learning model. The annual luminance-based simulations can be further accelerated by generating high-quality annual luminance maps from a subset of quick noisy renderings. Other potential projects include predicting indoor luminous environment from outdoor HDR skies and predicting luminance maps of multi-view positions/directions with a limited number of captures using light-field cameras.

Chapter 8. BIBLIOGRAPHY

Ahmad, M.W., Hippolyte, J., Mourshed, M., and Rezgui, Y. 2017. Random Forests and Artificial Neural Network for Predicting Daylight Illuminance and Energy Consumption. Proceedings of IBPSA, San Francisco, August 7-9, 2017.

Samuel, A.L. 1959. Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development, 3: pp.210–229.

Bergen, T. and Young, R. 2018. Fifty Years of Development of Light Measurement Instrumentation. Lighting Research & Technology, 50(1): pp.141–153.

Bourgeois, D., Reinhart, C. F., and Ward, G. 2008. Standard Daylight Coefficient Model for Dynamic Daylighting Simulations. Building Research & Information, 36 (1): pp.68–82.

Bako, S., Vogels, T., McWilliams, B., Meyer, M., Novák, J., Harvill, A., Sen, P., DeRose, T., and Rousselle, F. 2017. Kernel-Predicting Convolutional Networks for Denoising Monte Carlo Renderings. ACM Transactions on Graphics (Proc. SIGGRAPH), 36(4): 97, 2017.

Boyce, P. 2014. Human Factors in Lighting. Second Edition. CRC Press.

Breiman, L. 1994. Bagging Predictors. Machine Learning, 24(2): pp.123–140.

Chaitanya, C.R.A., Kaplanyan, A., Schied, C., Salvi, M., Lefohn, A., Nowrouzezahrai, D., and Aila, T. 2017. Interactive Reconstruction of Noisy Monte Carlo Image Sequences using a Recurrent Autoencoder. ACM Transactions on Graphics (Proc. SIGGRAPH), 36(4):98, 2017.

CIE. 1970. International Recommendations for the Calculation of Natural Daylight. CIE No. 16.

CIE. 1973. Standardization of Luminance Distribution of Clear Skies. CIE No.22.

Cohen, M.F. and John R.W. 1993. Radiosity and Realistic Image Synthesis. San Francisco: Morgan Kaufmann.

CS231n, Convolutional Neural Networks for Visual Recognition, Retrieved from http://cs231n.github.io/convolutional-networks/ in public domain (accessed: May 2019).

Debevec, P. and Malik, J. 1997. Recovering High Dynamic Range Radiance Maps from Photographs. ACM Transactions on Graphics (Proc. SIGGRAPH), pp.369–378, 1997.

Debevec, P. 2002. Image-based Lighting. IEEE Computer Graphics and Applications, 22(2): pp.26–34.

Debevec, P. 2005. Making the Parthenon, 6th International Symposium on Virtual Reality, Archaeology, and Cultural Heritage, Pisa, Italy.

Duchi, J., Hazan, E., and Singer, Y. 2011. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. Journal of Machine Learning Research, 12: pp.2121–2159.

Fechner, G.T. 1966 (First published in 1860). Howes, D.H., and Boring, E.G. Elements of Psychophysics (Elemente der Psychophysik). volume 1. Translated by Adler, H E. United States of America: Holt, Rinehart and Winston.

Gardner, M.A., Sunkavalli, K., Yumer, E., Shen, X., Gambaretto, E., Gagne, C., and Lalonde, J.F. 2017. Learning to Predict Indoor Illumination from a Single Image. ACM Transactions on Graphics (Proc. SIGGRAPH), 36(6): 176, 2017.

Goodfellow, I., Bengio, Y., and Courville, A. 2016. Deep Learning. The MIT Press.

Goral, C.M., Torrance, K.E., Greenberg, D.P., and Bennett, B. 1984. Modeling the Interaction of Light between Diffuse Surfaces. ACM Transactions on Graphics (Proc. of SIGGRAPH), 18(3): pp.212–222, 1984.

Hahnloser, R.H.R., Sarpeshkar, R., Mahowald, M. A., Douglas, R.J., and Seung, H. S. 2000. Digital Selection and Analogue Amplification Coexist in a Cortex-inspired Silicon Circuit. Nature, 405: pp.947–951.

Hastie, T., Tibshirani, R., and Friedman, J. 2011. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd Edition, Springer.

Heschong Mahone Group. 2012. Daylight Metrics – PIER Daylighting Plus Research Program. California Energy Commission. Retrieved from https://www.energy.ca.gov/2012publications/CEC-500-2012-053/CEC-500-2012-053.pdf (accessed: May 2019).

Hinton, G. and Salakhutdinov, R. 2006. Reducing the Dimensionality of Data with Neural Networks. Science, 313 (5786): pp.504–507.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., et al. 2012. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. IEEE Signal Processing Magazine, 29(6): pp.82–97, 2012.

Hold-Geoffroy, Y., Sunkavalli, K., Hadap, S., Gambaretto, E., and Lalonde J.F. 2017. Deep Outdoor Illumination Estimation. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, July 21-26, 2017.

Hopkinson, R.G. 1972. Glare from Daylighting in Buildings. Applied Ergonomics, 3(4).

Hošek, L. and Wilkie, A. 2012. An Analytic Model for Full Spectral Sky-dome Radiance. ACM Transactions on Graphics (Proc. SIGGRAPH), 31(4): pp.1–9, 2012.

Hubel, D.H. and Wiesel, T.N. 1968. Receptive Fields and Functional Architecture of Monkey Striate Cortex. The Journal of Physiology, 195 (1): pp.215–243.

Illuminating Engineering Society. 2011. IES Lighting Handbook: Reference and Application.

Illuminating Engineering Society. 2012. IES LM-83-12-IES Spatial Daylight Autonomy (SDA) and Annual Sun-light Exposure (ASE), New York, NY, USA.

Inanici, M. 2006. Evaluation of High Dynamic Range Photography as a Luminance Data Acquisition System. Lighting Research and Technology, 38(2): pp.123–136.

Inanici, M. 2009. Applications of Image Based Rendering in Lighting Simulation: Development and Evaluation of Image based Sky. Proceedings of IBPSA, Glasgow, Scotland, July 27-30, 2009.

Inanici, M. 2010. Evaluation of High Dynamic Range Image Based Sky Models in Lighting Simulation. Leukos, Journal of the IESNA, 7(2): pp.69–84.

Inanici, M. 2013. Dynamic Daylighting Simulations from Static High Dynamic Range Imagery Using Extrapolation and Daylight Coefficient Methodologies. Proceedings of IBPSA Chambéry, France, August 25-28, 2013.

Inanici, M. and Hashemloo, A. 2017. An Investigation of the Daylighting Simulation Techniques and Sky Modeling Practices for Occupant Centric Evaluations. Building and Environment, 113: pp.220–231.

Izadinia, H., Shan, Q., and Seitz, S.M. 2017. IM2CAD. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, July 21-26, 2017.

Jakubiec, J.A. and Reinhart, C.F. 2012. The 'Adaptive Zone' - A Concept for Assessing Discomfort Glare throughout Daylit Spaces. Lighting Research and Technology 44(2): pp.149–170.

Jakubiec, J.A., Van Den Wymelenberg, K., Inanici, M., and Mahic, A. 2016a. Accurate Measurement of Daylit Interior Scenes Using High Dynamic Range Photography. CIE Conference Melbourne, Australia, March 3-5, 2016.

Jakubiec, J.A., Inanici, M., Van Den Wymelenberg, K., and Mahic, A. 2016b. Improving the Accuracy of Measurements in Daylit Interior Scenes Using High Dynamic Range Photography. PLEA 2016, July 11-13, Los Angeles.

Jones, N. and Reinhart, C. 2014. Irradiance Caching for Global Illumination Calculation on Graphics Hardware. ASHRAE/IBPSA-USA Building Simulation Conference, Atlanta, September 10-12, 2014.

Jones, N. and Reinhart, C. 2017. Experimental Validation of Ray Tracing as a Means of Image-based Visual Discomfort Prediction. Building and Environment (113): pp.131-150.

Kajiya, J.T. 1986. The Rendering Equation. ACM Transactions on Graphics (Proc. SIGGRAPH), 20(4): pp.143-150, 1986.

Karandikar, R.V. 1955. Luminance of the Sun. Journal of the Optical Society of America, 45: pp.483-488.

Kazanasmaz, T., Gunaydin, M., and Binol, S. 2009. Artificial Neural Networks to Predict Daylight Illuminance in Office Buildings. Building and Environment, 44: pp.1751–1757.

Kingma, D.P. and Ba, J.L. 2015. Adam: a Method for Stochastic Optimization. International Conference on Learning Representations, pp.1–13.

Kittler, R. 1967. Standardization of Outdoor Conditions for Daylight Factors for Clear Skies. CIE No. 22.

Kittler, R., Richard, P., and Stanislav, D. 1998. A Set of Standard Skies Characterizing Daylight Conditions for Computer and Energy Conscious Design, 54.

Kleindienst, S.A. and Andersen, M. 2011. The Adaptation of Daylight Glare Probability to Dynamic Metrics in a Computational Setting. Proceedings of the Lux Europa, Istanbul.

Klepeis, N.E., Nelson, W.C., Ott, W.R., Robinson, J.P., Tsang, A.M., Switzer, P., Behar, J.V., Hern, S.C., and Engelmann, W.H. 2001. The National Human Activity Pattern Survey (NHAPS): A Resource for Assessing Exposure to Environmental Pollutants. Journal of Exposure Analysis and Environmental Epidemiology, 11 (3): pp.231–252.

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: pp.67–79.

Krizhevsky, A., Sutskever, I., and Hinton, G. 2012. ImageNet Classification with Deep Convolutional Neural Networks. Proceedings of the International Conference on Neural Information Processing Systems. pp.1097–1105, 2012.

Laouadi, A., Reinhart, C.F., and Bourgeois, D. 2008. Efficient Calculation of Daylight Coefficients for Rooms with Dissimilar Complex Fenestration Systems. Journal of Building Performance Simulation, 1: pp.3–15.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. 1998. Gradient-based Learning Applied to Document Recognition. Proceedings of the IEEE, 86(11): pp.2278–2324.

LeCun, Y., Bottou, L., Orr, G., and Muller, K. 1998. Efficient BackProp. Neural Networks: Tricks of the Trade, Springer.

Linnainmaa, S. 1970. The Representation of the Cumulative Rounding Error of an Algorithm as a Taylor Expansion of the Local Rounding Errors. Master's Thesis (in Finnish), University of Helsinki, pp.6-7.

Lloyd, S.P. 1957. Least Square Quantization in PCM. Bell Telephone Laboratories Paper. Published in journal much later: Lloyd., S.P. 1982. Least Squares Quantization in PCM. IEEE Transactions on Information Theory, 28 (2): pp.129–137.

Luckiesh, M. and Guth, S.K. 1949. Brightnesses in Visual Field at Borderline Between Comfort and Discomfort (BCD), National Technical Conference of the Illuminating Engineering Society, French Lick, USA, pp.650-670.

Maas, A.L., Hannun, A.Y., and Ng, A.Y. 2013. Rectifier Nonlinearities Improve Neural Network Acoustic Models. International Conference on Machine Learning, Atlanta, June 16-21, 2013.

Mardaljevic, J. 2000a. Daylight Simulation: Validation, Sky Models and Daylight Coefficients. De Monfort University.

Mardaljevic, J. 2000b. Simulation of Annual Daylighting Profiles for Internal Illuminance. Lighting Research and Technology, 32 (3): pp.111–118.

Mardaljevic, J., Andersen, M., Roy, N., and Christoffersen, J. 2012. Daylighting Metrics: Is There a Relation Between Useful Daylight Illuminance and Daylight Glare Probability? Proceedings of Building Simulation and Optimization Conference (BSO), Loughborough, UK, September 10-11, 2012.

McCann, J.J. 1999. Spatial Contrast and Scatter: Apposing Partners in Sensations, Electronic Imaging. Conference on Human Vision and Electronic Imaging IV: pp.97-104, San Josè, USA.

McCulloch, W.S. and Pitts, W. 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity. The Bulletin of Mathematical Biophysics, 5 (4): pp.115–133.

McNeil, A. and Lee E.S. 2013. A Validation of the Radiance Three-phase Simulation Method for Modelling Annual Daylight Performance of Optically Complex Fenestration Systems. Journal of Building Performance Simulation, 6 (1): pp.24-37.

McNeil, A. 2013a. The Three-phase Method for Simulating Complex Fenestration with Radiance. Lawrence Berkeley National Laboratory, Berkeley, CA.

McNeil, A. 2013b. The Five-phase Method for Simulating Complex Fenestration with Radiance. Lawrence Berkeley National Laboratory, Berkeley, CA.

Michelson, A. 1927. Studies in Optics. Chicago, IL: University of Chicago Press.

Nabil, A. and Mardaljevic, J. 2005. Useful Daylight Illuminance: A New Paradigm for Assessing Daylight in Buildings. Lighting Research and Technology, 37: pp.41–59.

Nair, V. and Hinton, G. E. 2010. Rectified Linear Units Improve Restricted Boltzmann Machines. International Conference on Machine Learning, Haifa, Israel, June 21-24, 2010.

Nakamura, H., Oki, M., and Hayashi, Y. 1985. Luminance Distribution of Intermediate Sky. Journal of Light and Visual Environment, 9 (1): pp.6–13.

Navada, S.G., Adiga, C.S., and Kini, S.G. 2016. Prediction of Daylight Availability for Visual Comfort. International Journal of Applied Engineering Research, 11 (7): pp.4711–4717.

Newsham, R., Richardson, C., Blanchet, C., and Veitch, J. 2005. Lighting Quality Research Using Rendered Images of Offices. Lighting Research and Technology, 37(2): pp.93–115, 2005.

Ng, A.Y. 2004. Feature Selection, L1 vs. L2 Regularization, and Rotational Invariance, In International Conference on Machine Learning, Banff, Alberta, Canada, 2004.

Pavel, M., Sperling, G., Riedl, T., and Vanderbeek, A. 1987. Limits of Visual Communication: the Effect of Signal-to-noise Ratio on the Intelligibility of American Sign Language. Journal of the Optical Society of America, 4(12): pp.2355–2365.

Perez, R., Seals, R., and Michalsky, J. 1993. All-Weather Model for Sky Luminance Distribution—Preliminary Configuration and Validation. Solar Energy, 51: pp.423.

Reindl, D.T., Beckman, W.A., and Duffie, J.A. 1990. Diffuse Fraction Correlations. Solar Energy, 45 (1): pp.1-7.

Reinhart, C.F., and Walkenhorst, O. 2001. Validation of Dynamic RADIANCE-Based Daylight Simulations for a Test of Office with External Blinds. Energy and Buildings, 33(7): pp.683-697.

Reinhart, C.F. 2012. Simulation-Based Daylight Performance. In Building Performance Simulation for Design and Operation, edited by Hensen, J. and Lamberts, R. pp.235–276. Taylor & Francis.

Ren, P., Dong, Y., Lin, S., Tong, X., and Guo, B. 2015. Image Based Relighting Using Neural Networks. ACM Transactions on Graphics (Proc. SIGGRAPH), 34 (4): pp.111:1–111:12, 2015.

Rizzi, A., Algeri, T., Medeghini, G., and Marini, D. 2004. A Proposal for Contrast Measure in Digital Images. Proceedings of Second European Conference on Color in Graphics, Imaging and Vision, pp.187-192. Aachen, Germany.

Rockcastle, S. and Andersen, M. 2017. Contrast Measures for Predicting Perceptual Effects of Daylight in Architectural Renderings. Lighting Research & Technology, 49(7): pp.882–903.

Rogers, Z. 2006. Daylighting Metric Development Using Daylight Autonomy Calculations in the Sensor Placement Optimization Tool. Boulder, Colorado, USA: Architectural Energy Corporation.

Sahin, M., Oguz, Y., and Büyüktümtürk, F. 2016. ANN-based Estimation of Time-dependent Energy Loss in Lighting Systems, Energy Build, 116 (2016): pp.455–467.

Satilmis, P., Bashford-Rogers, T., Chalmers, A., and Debattista, K. 2016. A Machine Learning Driven Sky Model. IEEE Computer Graphics and Applications, 37 (1): pp.80-91.

Simone, G., Pedersen, M., Hardeberg, J.Y., and Rizzi, A. 2009. Measuring Perceptual Contrast in a Multi-level Framework. Proceedings of Electronic Imaging, San Jose, CA.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. 2014. Dropout: a Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research, 15: pp.1929–1958.

Stumpfel, J., Jones, A., Wenger, A., Tchou, C., Hawkins, T., and Debevec. P. 2004. Direct HDR Capture of the Sun and Sky. Proceedings of AFRIGRAPH, Stellenbosch, South Africa, 2004.

Suk, J.Y., Schiler, M., and Kensek, K. 2013. Development of New Daylight Glare Analysis Methodology Using Absolute Glare Factor and Relative Glare Factor. Energy and Buildings, (64): pp.113–122.

Tadmor, Y. and Tolhurst, D., 2000. Calculating the Contrasts That Retinal Ganglion Cells and LGN Neurones Encounter in Natural Scenes. Vision Research, 40: pp.3145–3157.

Turmon, M.J. and Fine T.L. 1994. Sample Size Requirements for Feedforward Neural Networks. Proceedings of International Conference on Neural Information Processing Systems 2 (1): pp.327– 334, 1994.

Tregenza, P.R. and Waters, I.M. 1983. Daylight Coefficients. Lighting Research and Technology, 15(2): pp.65–71.

U.S. Energy Information Administration. 2012. Commercial Buildings Energy Consumption Survey.

U.S. Department of Energy. 2019. Energy+ Weather Data, https://energyplus.net/weather (accessed: May 2019).

Van Den Wymelenberg, K. and Inanici, M. 2014. A Critical Investigation of Common Lighting Design Metrics for Predicting Human Visual Comfort in Offices with Daylight. Leukos, Journal of the IESNA, 10(3): pp.145–164.

Van Den Wymelenberg, K. and Inanici, M. 2016. Evaluating a New Suite of Luminance-Based Design Metrics for Predicting Human Visual Comfort in Offices with Daylight. Leukos, Journal of the IESNA, 12(3): pp.1-26.

Ward, G., Mistrick, R., Lee, E.S., Mcneil, A., and Jonsson, J. 2011. Simulating the Daylight Performance of Complex Fenestration Systems Using Bidirectional Scattering Distribution Functions within Radiance. Leukos, Journal of the IESNA, 7(4): pp.241-261.

Ward, G. 1994. The RADIANCE Lighting Simulation and Rendering System, ACM Transactions on Graphics (Proc. SIGGRAPH), pp.459-472, 1994.

Ward, G. 2005. Photosphere, 2005. http://www.anyhere.com/ (accessed: May 2019).

Whitted, T. 1980. An Improved Illumination Model for Shaded Display. Communications of ACM. 23(6): pp.343-349.

Wienold, J. and Christoffersen, J. 2006. Evaluation Methods and Development of a New Glare Prediction Model for Daylight Environments with the Use of CCD Cameras. Energy and Buildings, 38(7): pp.743-757.

Wienold, J. Evalglare version 0.9f. March 2010. Fraunhofer Institute for Solar Energy Systems, Freiburg, Germany. Retrieved from www.ise.fraunhofer.de/radiance (accessed: May 2019).

Xu, X., Sunkavalli, K., Hadap, S., and Ramamoorthi, R. 2018. Deep Image-based Relighting from Optimal Sparse Samples. ACM Transactions on Graphics (Proc. SIGGRAPH), 37(4):126, 2018.

Zuo, W., McNeil, A., Wetter, M., and Lee. E.S. 2011. Acceleration of Radiance for Lighting Simulation by Using Parallel Computing with OpenCL. 12th Conference of International Building Performance Simulation Association, pp.110–117.