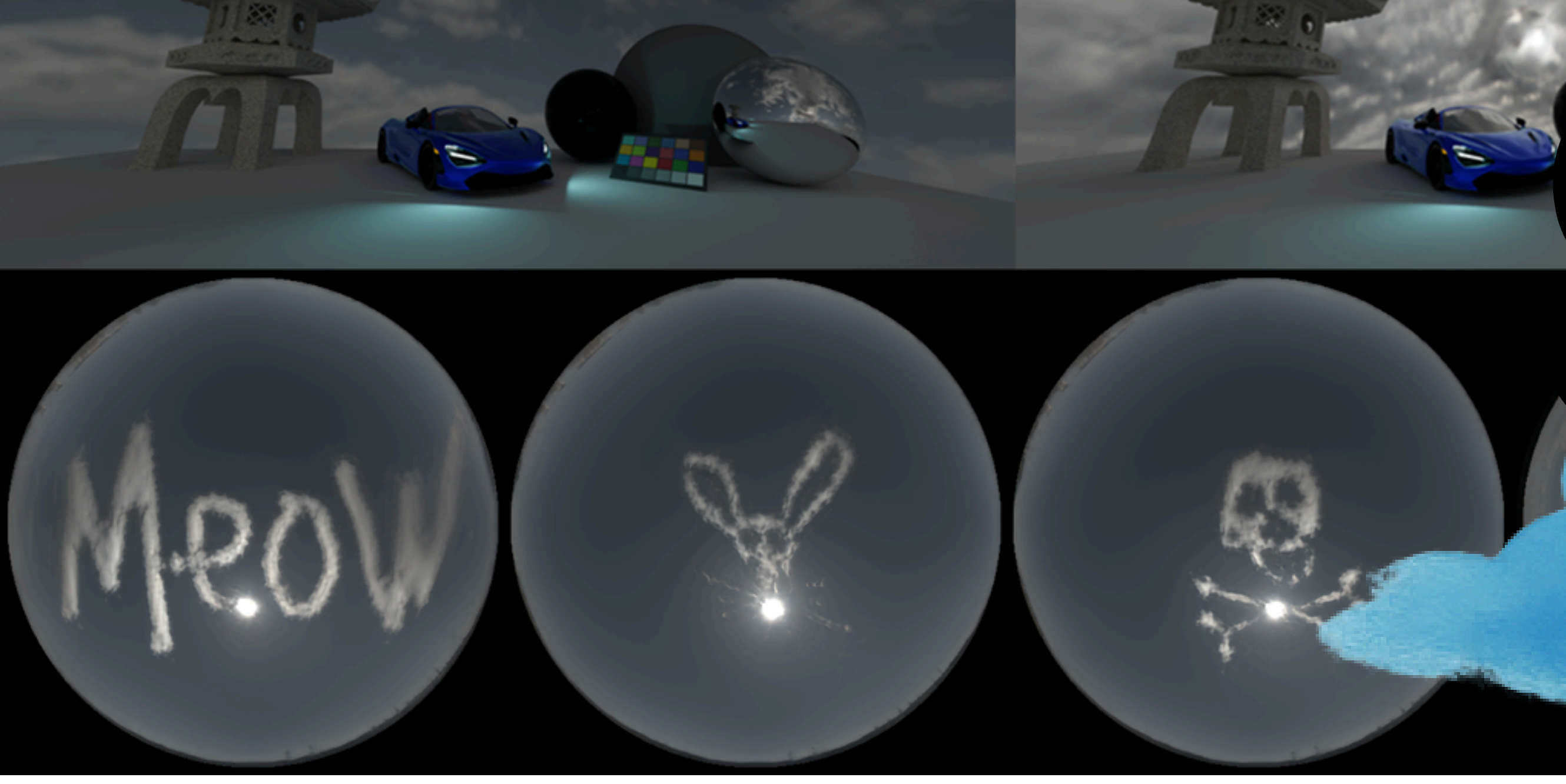
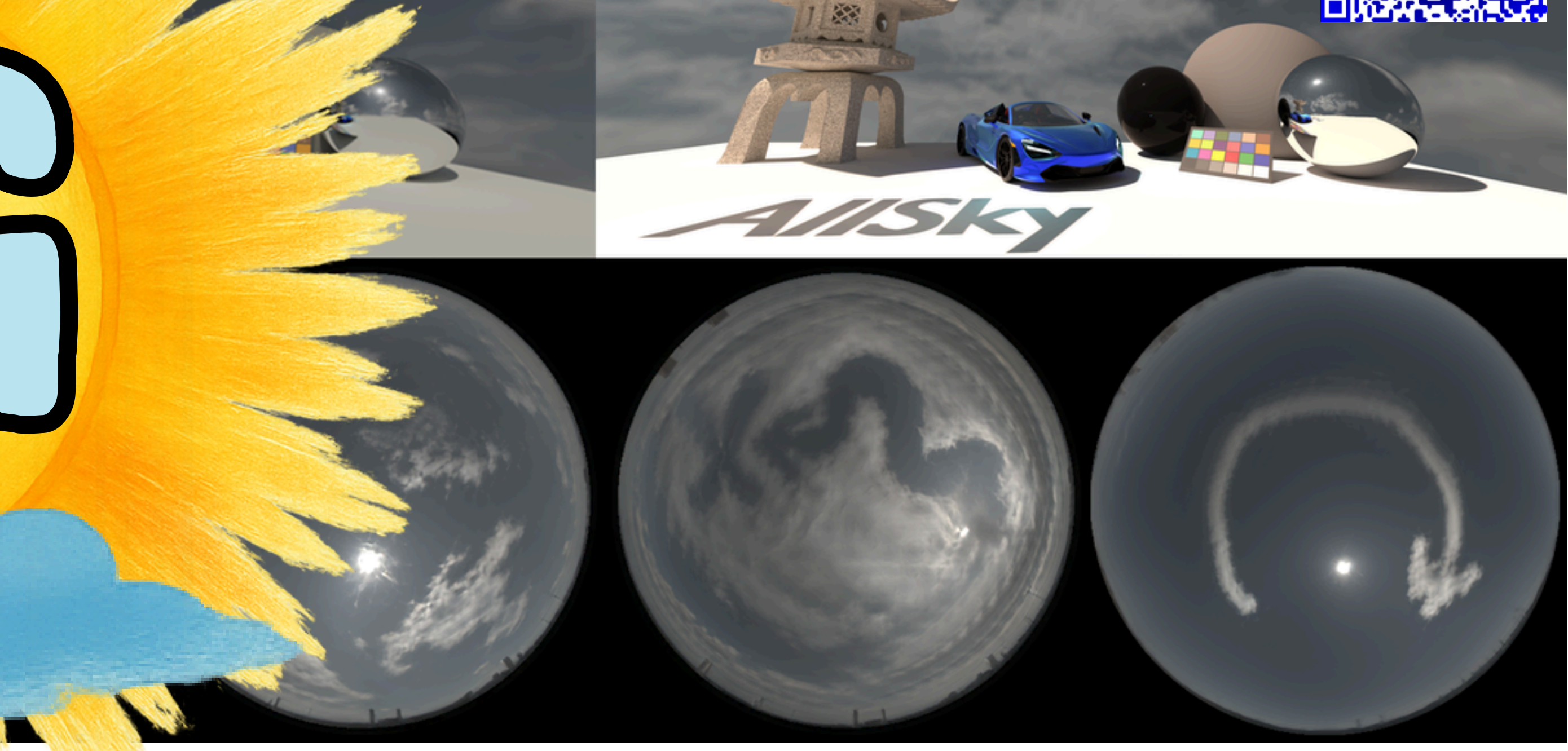
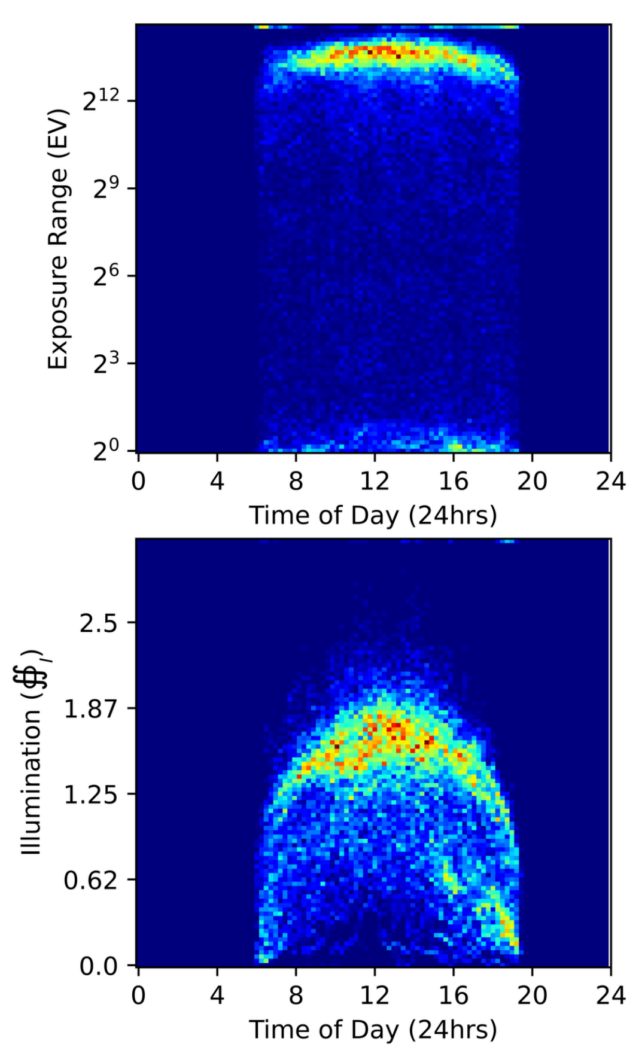


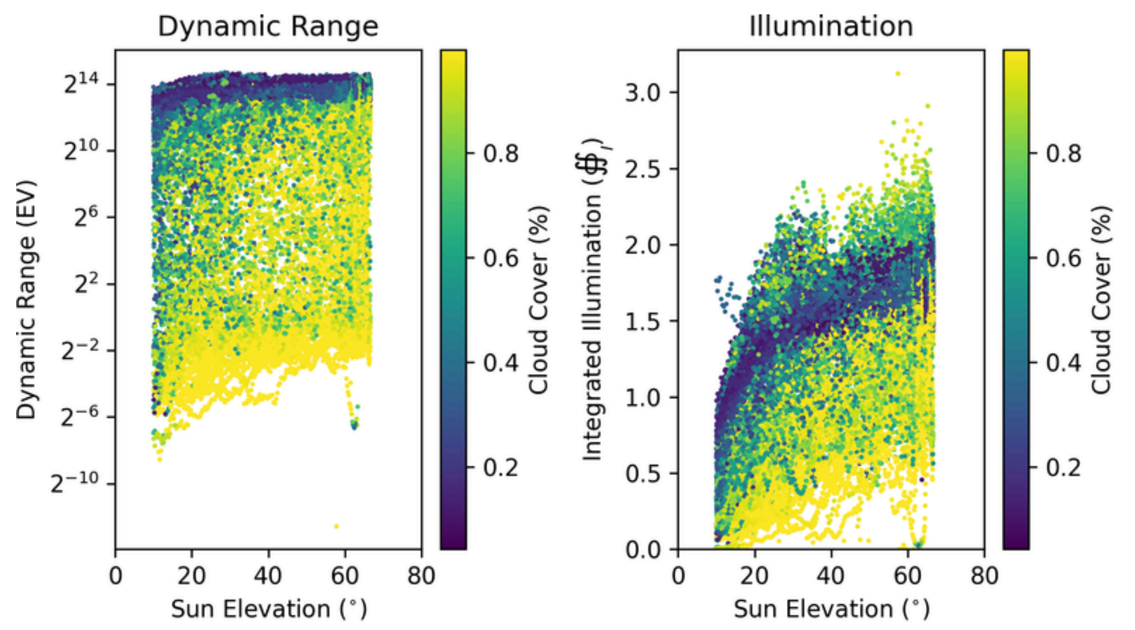
TOWARDS PHYSICALLY-BASED
SKY-MODELLING
FOR IMAGE BASED LIGHTING

Ian J. Maquignaz, Université Laval

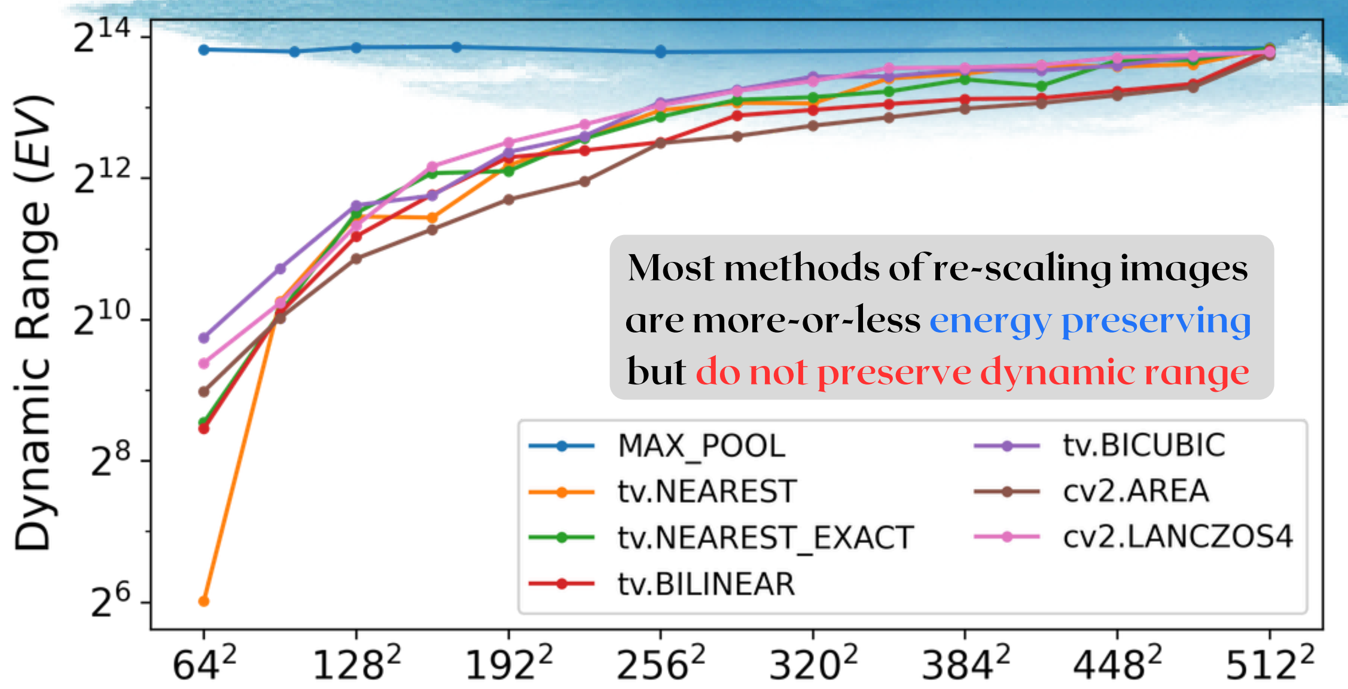


Skydome Dataset



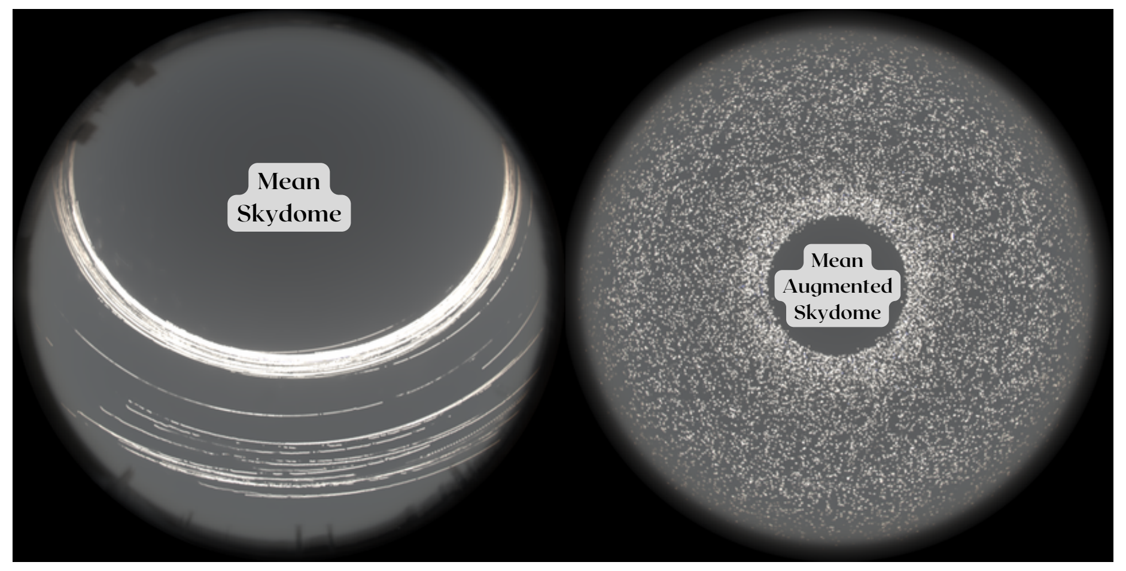


Dynamic range and illumination fluctuate depending on temperature, pressure, aerosols, airborne particles and other factors.



Most methods of re-scaling images are more-or-less energy preserving but do not preserve dynamic range

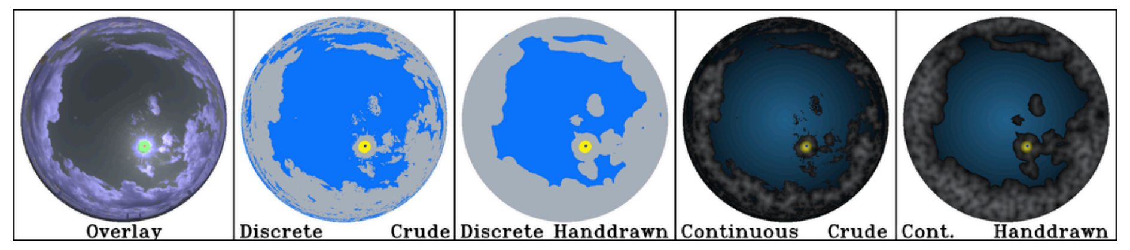
Augmenting the dataset through rotations and flips around the zenith allows for greater coverage.



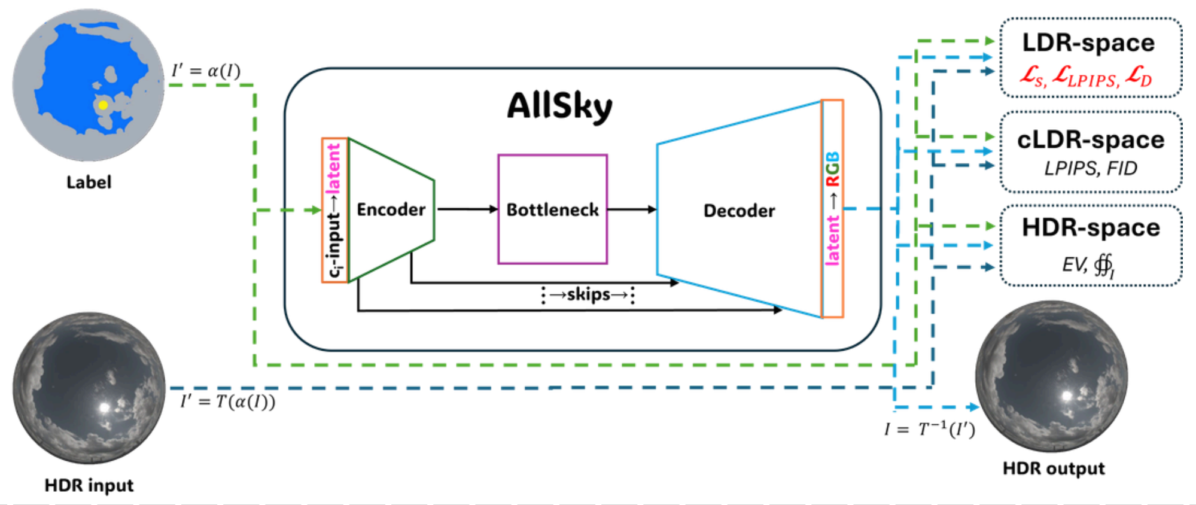
Input Modalities

Method #1: "Hand-Drawn" Segmentation Masks

<u>Clouds Segmentation</u> can be achieved by thresholding the ratio Y=Red/Blue or Y=(Blue-Red)/(Blue+Red). This method is crude, but simple and effective.

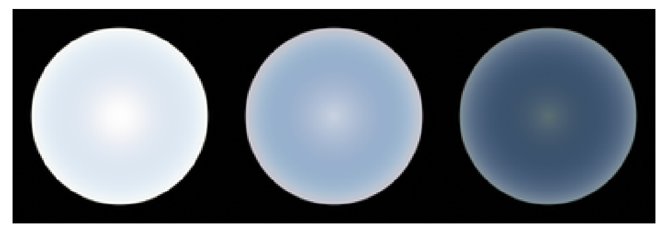


<u>Sun positioning</u> can be inferred as the brightest spot or calculated via ephemeris tables. The extraterrestrial solar solar disk is 0.5 angular degrees but, with atmospheric attenuation, is visually closer to 5 angular degrees.



Method #2:
Clear Sky Augmentation

Many recent models such as DeepClouds, SkyGAN, and LM-GAN elect to use parametric sky-models as priory. This enables users to control the generation of a clear-sky (sometimes inclusive of a solar disk) via minimal set of mostly-intuitive parameters, which can then be augmented by a DNN to include atmospheric formations.



Pretham, Nishita and Hošek-Wilkie
Clear Sky Models

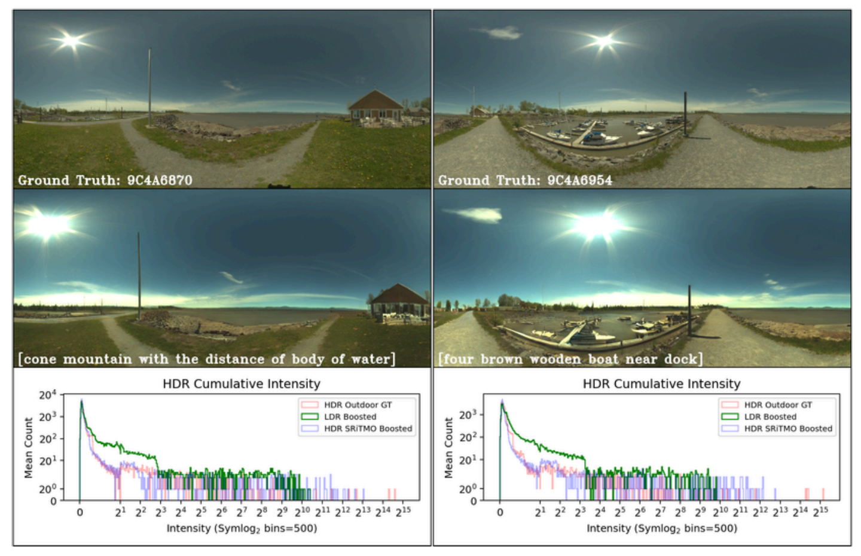
Though effective, the viability of these methods is questionable. Retention of user-configured clear-sky and, if applicable, solar disk, has not been demonstrated.

Method #3:
Inference from Textual Prompts

With the emergence of CLIP, models such as Text2Light championed textual prompts for IBL. As shown, this input modality is vague and imprecise. The language required to accurately describe a skydome is non-colloquial.

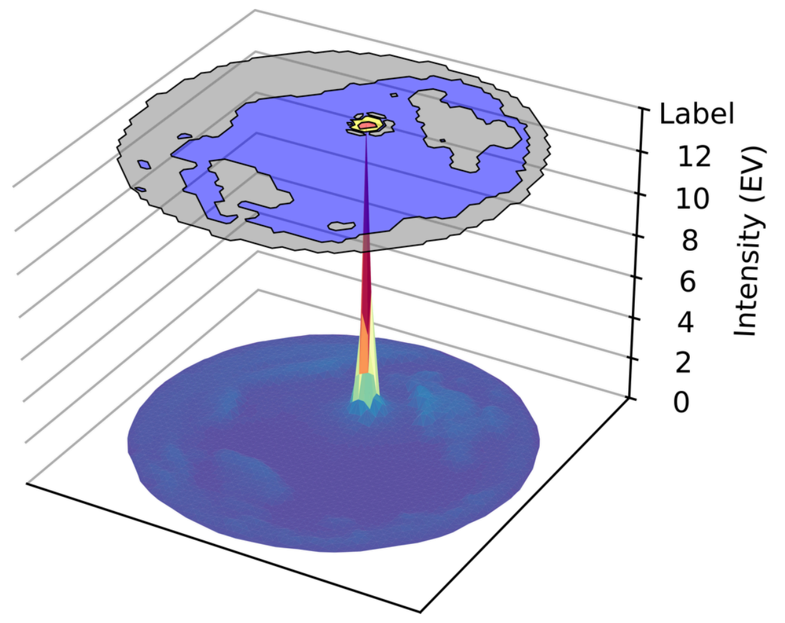
Boosting: $\rho = 4, \gamma = 1$ and $\beta = 0.7$

$$I'_b = I_b + (I_b M \rho); I'_b = e^{(I'_b - \overline{I'_b})\gamma - \beta}$$

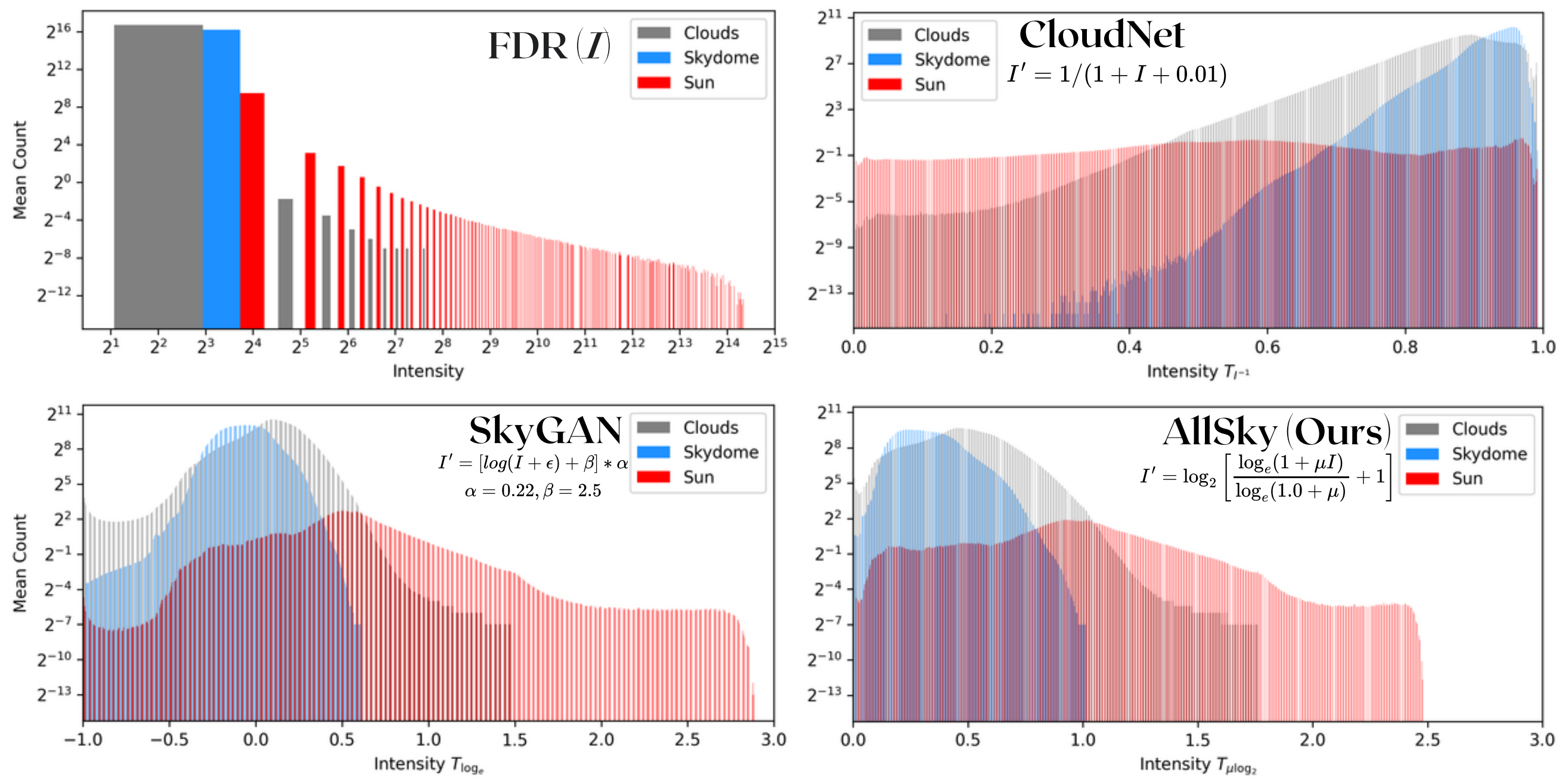


The importance of FDR

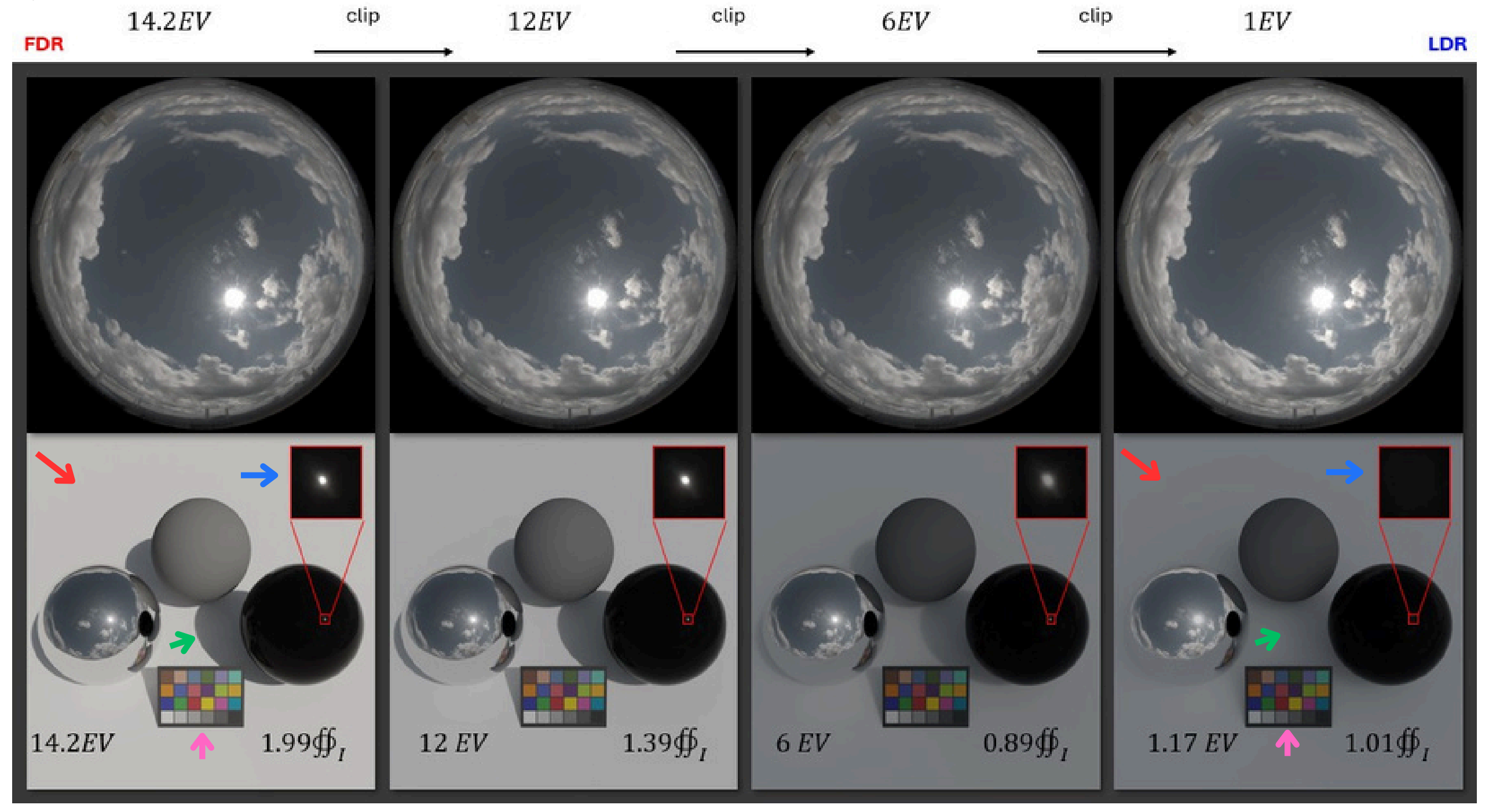
Illumination coherence can be key to distinguishing real and virtual-scene outdoor scenes. To accurately represent outdoor illumination, Full Dynamic Range (FDR) Imagery is necessary to <u>fully-capture</u> the dynamic range of an average real-world outdoor scene without saturation of the exposure range.



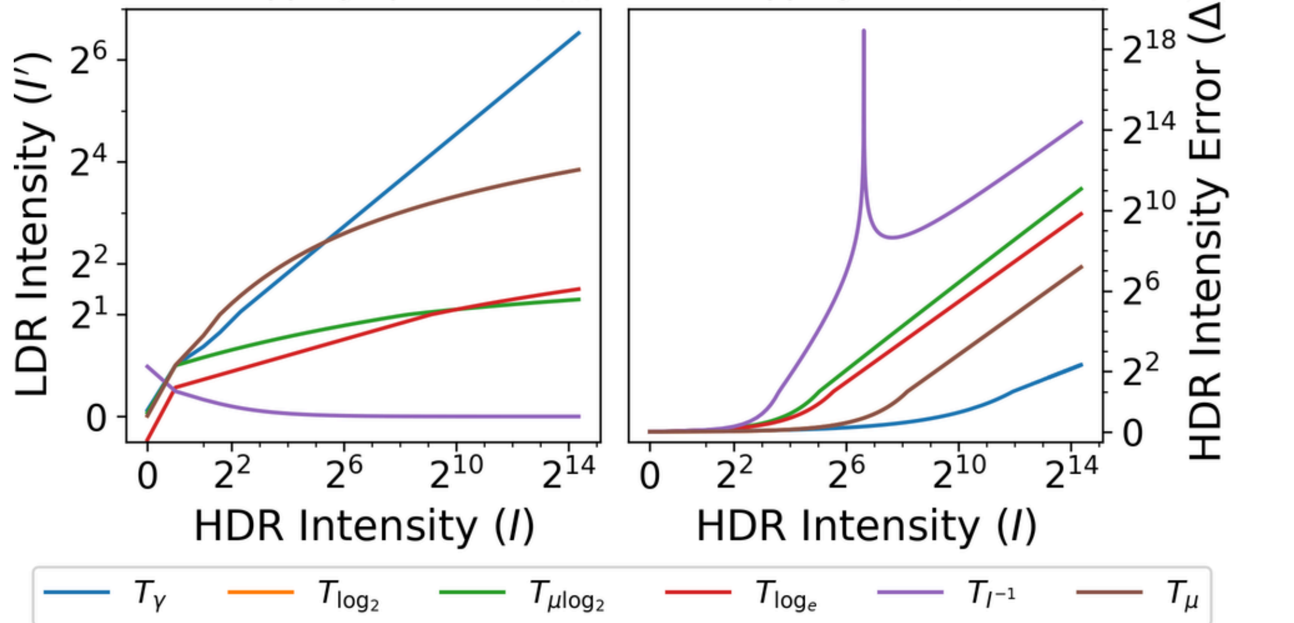
Full Dynamic Range is primarily required for sun pixels, which are less than 5 percent of skydome pixels but represent 60 percent of skydome illumination.



Skydomes which fail to preserve Dynamic Range are **indistinguishable** when viewed as **Low Dynamic Range** (LDR) images. **When rendered,** flaws emerge as changing illumination, surface albedos, shadows, and light transmission by materials.



We measure HDR as the Exposure Value (EV) of HDRI given $EV = log_2(|I|_{max} - |I|_{min} + 1)$, where $|I|$ is grayscale intensity. Tone-mapping operators (T_m) compress HDRI to a visible, displayable, or latent colour-space more favourable for DNN training given $I' = T_m(I)$. We investigate a range of operators, including: Power-Law (T_γ), logarithmic (T_{log_n}), μ-law (T_μ), and variations thereof. With the growing popularity of mixed-tone-mappers (combinations of tone-mappers) such as Hybrid Log-Gamma for display of 8K broadcasts, we propose μ-lawLog$_2$ ($T_{\mu\log_2}$) to better preserve texture-rich cloud and skydome components (generally $EV \lessapprox 1$) and aggressively compress texture-poor high-exposure regions which are generally saturated in LDR imagery.

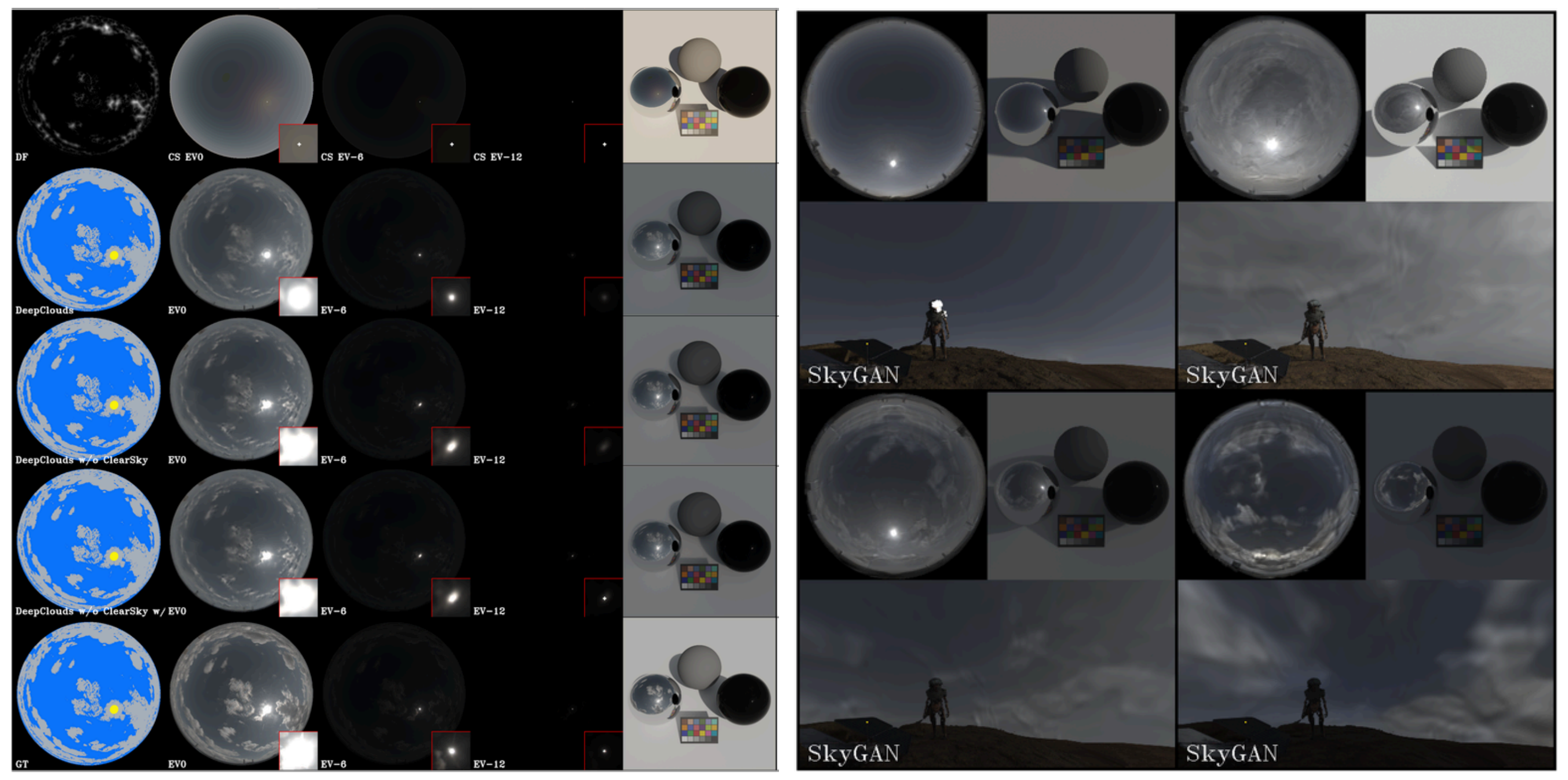


Tonemappers compress dynamic range to be suitable for DNN models. Tonemappers are bijections which introduce non-linear error between LDR- and HDR-space.

ALL-SKY Results

Table 1. Comparative results for Text2Light [10], CloudNet [50], SkyGAN [37], and AllSky (ours). AllSky offers unsurpassed visual quality (LPIPS, FID) for both crude (C) and hand-drawn (H) labels, and accurate FDR illumination (EV, $\oiint_I$) superseded only by SkyGAN. The results further demonstrate that L_1, L_2 and PNSR offer little correlation to visual quality or illumination. <u>Underlined</u> values indicate category best. Exposure matching is denoted by (*).

	LDR	T_γ-cLDR		HDR			
	PSNR$_{\log_2}$ ↑	LPIPS ↓	FID ↓	L_1 ↓	L_2 ↓	EV ←	$\oiint_I$ ←
Ground Truth 512^2	-	-	-	-	-	12.85	3.54
Text2Light SRiTMO Boosted	46.5	0.51	-	0.27	816	8.4	0.16
Text2Light SRiTMO Boosted*	46.5	0.42	-	0.32	1380	12.92	3.73
Text2Light LDR+Boost*	47.3	0.42	-	0.25	827	7.96	3.69
Ground Truth 512^2	-	-	-	-	-	12.7	1.05
CloudNet	99.4	0.17	35.2	0.07	414	6.3	0.53
CloudNet w/ Sun	99.4	0.17	35.1	0.17	1907	14.4	2.07
CloudNet w/o Clear Sky	92.1	0.18	41.5	0.08	418	6.7	0.64
CloudNet w/o Clear Sky w/Sun	92.1	0.18	41.5	0.18	1909	14.4	2.18
Ground Truth 256^2	-	-	-	-	-	11.12	1.15
SkyGAN	-	-	55	-	-	10.7	0.83
SkyGAN w/o Clear Sky	-	-	70	-	-	11.3	1.01
StyleGAN $T_{\mu\log_2}$	-	-	49	-	-	7.7	1.04
Ground Truth 256^2	-	-	-	-	-	8.09	1.13
AllSky $\mathbb{Z}^+$ $T_{\mu\log_2}$ C	104.3	0.13	14.9	0.07	965	7.24	0.85
AllSky $\mathbb{Z}^+$ $T_{\mu\log_2}$ H	98.1	0.16	22.6	0.43	344704	9.58	4.49
Ground Truth 512^2	-	-	-	-	-	9.11	1.14
AllSky $\mathbb{Z}^+$ $T_{\mu\log_2}$ C	102.8	0.14	17.1	0.06	467	8.23	0.82
AllSky $\mathbb{Z}^+$ $T_{\mu\log_2}$ H	93.3	0.18	24.8	0.11	9919	9.95	1.52



<u>Text2Light</u>

Skydomes generated via user-provided textual prompts

'Sunny Afternoon'

'A storm is approaching on the horizon'