Joint MAM MANER APPMAT Workshop on Material Appearance Modeling (2026)
D. Muselet and R. Pacanowski and H. Rushmeier (Editors)

A benchmark for BRDF and BTDF models

H. Mateo[1], J. Marco[1], P. Mézières[2], P. Santafe[1], A. Ferrero[3], R. Pacanowski[2]

[1]Universidad de Zaragoza–I3A, Spain
[2]Inria, France
[3]Instituto de Óptica, Consejo Superior de Investigaciones Científicas

Abstract

Fitting reflectance and transmittance models to measured data is a central challenge in material appearance modeling. While the graphics community has produced dense BRDF datasets, these measurements typically lack uncertainty estimates, unlike the sparse but metrologically rigorous data provided by the National Metrology Institutes (NMIs). In this paper, we present a benchmark evaluating several BRDF and BTDF models against real measurements with associated radiometric uncertainties. For BRDF, we assess models including Oren-Nayar, GGX, VMF and Micrograins. For BTDF, we evaluate different microfacet-based models (GGX, Beckmann,...) on a diverse set of transmissive materials. We study the influence of uncertainty-aware weighting schemes on fitting quality and compare the effect of different loss formulations. This benchmark provides a foundation for principled evaluation of appearance models under metrologically grounded conditions.

1. Motivation

Since the work by Matusik et al. [MPBM03], the graphics community (e.g., [FVH14]) has continued to develop imaging acquisition systems to measure BRDF or SVBRDF as densely as possible. However, none of the measurements made publicly available include uncertainty values. In contrast, the optics community—and national metrology institutes (NMIs) in particular—provide scattered measurements for which uncertainties are available. In this paper, we focus on evaluating 5 BRDF models, as well as, 4 BTDF models in terms of radiometric uncertainties and leave directional uncertainties for future work. More precisely, we test 4 different fitting metrics (i.e., losses) for BRDF and BTDF, which are modified to include the uncertainties provided by the National Institutes of Metrology and analyze and discuss our results.

2. Methodology

The spectral BRDF measurements from metrology institutes are converted to RGB color by selecting 3 representative wavelengths.

2.1. BRDF optimization

To optimize BRDF models to our real BRDF samples, we leverage the existing BBM library [Pee23], which implements a broad collection of BRDF models and metrics, and optimizes BRDF models using compass search. To mitigate falling into local minima, we run several iterations of the compass search with random parameter initialization for each BRDF model and keep the best result.

2.1.1. BRDF samples

Our dataset contains spectral angular measurements of several materials, most of them displaying diffuse-like reflectance, and a few of them with varying levels of glossiness. The vast majority of our samples correspond to in-plane measurements where the outgoing direction $\{\theta_o, \phi_o\}$ lies on the plane determined by the incoming direction and the surface normal. Our data has a coarse angular sampling rate of 10-degree steps in θ_i and θ_o In a few captured materials, we included out-of-plane measurements and finer sampling of near-specular directions. To keep the analysis tractable, we optimize model parameters in RGB space by selecting three representative wavelengths of each dataset.

2.1.2. BRDF models

Since our datasets are predominantly diffuse reflectance, we analyze a subset of diffuse-oriented models, adding support for highlights with Fresnel aggregate models and GGX microfacets. We incorporate energy-preserving Oren Nayar (EON) [dW24]; von Mises-Fisher (VMF) NDFs for rough diffuse surfaces [dW24]; and recent micrograin (Mg) BRDF models [LRPB23], which allow combining rough and diffuse micrograin scattering with a configurable base model (e.g. Lambertian, Oren-Nayar, GGX, etc). In total, we run optimizations on the following models: Lambertian, Oren-Nayar [ON94], EON [PKH25], VMF [dW24], GGX microfacet-based models [WMLT07], diffuse (DMg) and rough micrograin (RMg) models using different base models [LRPB23], and Fresnel aggregates (FA) of non-micrograin diffuse bases with a highlight base of GGX microfacets. All parameters, including those of the nested models, are optimizable. Since we perform compass search, the models are not required to be differentiable.

2.2. Loss functions

We evaluate our models and optimize their parameters to match real spectral samples using several loss functions. We implement several metrics with and without accounting for per-sample uncertainty quantification available in our real datasets. In particular we implement the following loss functions that average a specific metric over all real samples y_k and their corresponding model evaluations y'_k.

SMAPE : Computes the average of the symmetric mean absolute percentage error as

$$\mathcal{L}_{\text{SMAPE}} = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{|y'_k - y_k|}{|y'_k| + |y_k| + \varepsilon} \cdot w_k\right).$$

RMSE : Computes the root mean square error as

$$\mathcal{L}_{\text{RMSE}} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left((y'_k - y_k)^2 \cdot \cos^{+}\theta_i{}^2 \cdot w_k\right)}.$$

NganL2 : Computes the Ngan L^2 metric [NDM05] as:

$$\mathcal{L}_{\text{NganL2}} = \left[\sum_{k=1}^{K}(y'_k - y_k)^2 \cdot (\cos^{+}\theta_i)^2 \cdot w_k\right] \cdot \sin\theta_i \cdot \sin\theta_o.$$

lowL2 : Computes the Low et al. [LKYU12] L^2 metric as:

$$\mathcal{L}_{\text{LowL2}} = \left[\sum_{k=1}^{K}(y'_k - y_k)^2 \cdot \cos^{+}\theta_i{}^2 \cdot w_k\right] \cdot \sin\theta_i.$$

Uncertainty weights Our datasets provide uncertainty values u_k for each captured spectral sample y_k. For each aforementioned loss function, we implement two variants: an uncertainty-unaware variant $\mathcal{L}_*^{\text{unif.}}$ with uniform weights $w_k \equiv 1$, and an uncertainty-aware variant $\mathcal{L}_*^{\text{r.u.}}$ that applies a relative uncertainty weight $w_k \equiv \mathcal{R}_u(y_k, u_k)$ as

$$w_k \equiv \mathcal{R}_u(y_k, u_k) = \frac{1}{\frac{|u_k|}{|y_k|+\varepsilon} + \varepsilon}$$

.

2.3. BTDF Models and dataset

2.3.1. BTDF samples

All measured data (Table 1) used in this short paper are taken from Fu *et al.* [FFQ*24]. where we used samples measured at 633 nm. There are two sets of measurements. Both sets of measurements were illuminated at a normal angle (θ_i, ϕ_i) = (0°, 0°). The transmission measurements were performed in the incidence plane (θ_t, ϕ_t) = (from -20° to 20° in 5° increments, 0°) for the first set of measurements. Except for materials A and D, which were measured in 1° steps for θ_t, and with two different ϕ_t values (ϕ_t = 0° & 90°) since these materials are anisotropic. For the second set of measurements, we have (θ_t, ϕ_t) = (0° to 80° in 10° steps, 0°). For more details on the dataset, we refer the reader to the work of Fu *et al.* [FFQ*24]. The first set of measurements was taken at the PTB (Physikalisch-Technische Bundesanstalt), and the second at the CSIC (Consejo Superior de Investigaciones Científicas).

Category	Set	Name	Scattering	Thickness
Quasi Lambertian	1,2	C	Bulk	2 mm
	1,2	E		0.25 mm
Angle dependant scattering	1,2	B	Gaussian	2 mm
	2	Co1		1 mm
	2	Co3		3 mm
	1	A	Square-shaped	1.5 mm
	1	D	Gaussian elliptical	2 mm

Table 1: *BTDF samples. All materials are industry-manufactured. There are five optical diffuser materials (A–E), representing a range of light-scattering mechanisms and optical structures, whereas Co1 and Co3 are manufactured polycarbonate plates containing scattering centres.*

2.3.2. BTDF models

In this article, we use the microfacet model for refraction defined by Walter *et al.* [WMLT07]

$$f_t(\omega_i, \omega_t) = \frac{|\omega_i \cdot m_t||\omega_t \cdot m_t|}{|\omega_i \cdot n||\omega_t \cdot n|}\frac{(1 - F(\omega_i \cdot m_t))G(\omega_i, \omega_t)D(m_t \cdot n)}{(\eta_i(\omega_i \cdot m_t) + \eta_t(\omega_t \cdot m_t))^2} \tag{1}$$

where $m_t = -\frac{\eta_i\omega_i + \eta_t\omega_t}{||\eta_i\omega_i + \eta_t\omega_t||}$ is the microfacet normal for the transmission. The refractive index η_i is always set to 1, whereas the transmission refractive index η_t is left to be optimized. The functions F, G, and D are, respectively, the Fresnel term, the microfacet masking term, and the normal distribution function. Here, F is always computed exactly, and different models are used for G and D. In a first time, we used the masking term (G) defined for the Trowbridge-Reitz (TR) [WMLT07] model is used, and different D terms are used: TR, Beckmann, or generalized Trowbridge-Reitz (GTR) [BS12]. We also fitted a Generalized Gaussian (GG) model that defines the terms D and G [HP17].

The G and D term of the TR and Beckmann models are based on a roughness parameter (α); the GTR model adds a parameter γ; and the GG model is based on two parameters (β and p). These settings are optimized based on the model being used. To these, we also add the optimization of the transmission refractive index η_t. To account for absorption effects, which are significant in less transparent materials, we tried two approaches: a simple scaling S of the BTDF, or scaling by an absorption factor that depends on the transmission angle using the Beer-Bouguer-Lambert law $\exp(-\frac{\tau d}{\omega_t \cdot n})$, where d is the thickness of the sample and τ is the optical depth. In this work, we directly optimize $\tilde{\tau} = \tau d$.

We tested multiple metric for loss functions commonly used in the literature [LKYU12]. For each BTDF samples, the loss is always computed using the L2 norm $l_i = (g_i - \hat{g}_i)^2$ where the measured data, g, and model values, $\hat{g}$ are modified according to the following equations: $g_1 = f_t(\omega_i, \omega_t)$, $g_2 = \cos\theta_t f_t(\omega_i, \omega_t)$, $g_3 = \log(1 + f_t(\omega_i, \omega_t))$, $g_4 = \log(1 + \cos\theta_t f_t(\omega_i, \omega_t))$. We also tried to account for uncertainties by multiplying the loss function by the inverse of the uncertainty.

Among the samples, only A and D are anisotropic; however, these samples were measured only along two ϕ_t axes (0° and 90°). We did not fit any anisotropic models, but only isotropic

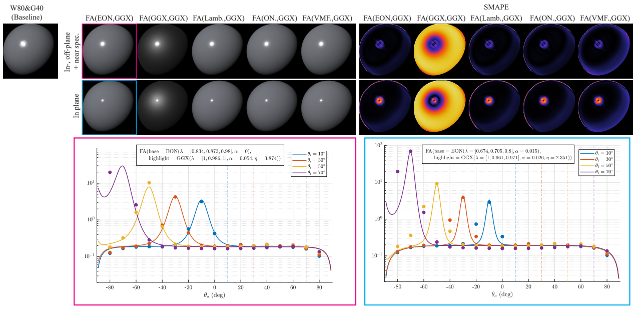

Figure 1: *Optimizations of Fresnel aggregate models using in-plane measurements only, and including also off-plane and denser near-specular samples of the W80&G40 material. Angular plots (bottom) compare a specific Fresnel aggregate using EON and GGX (lines) with the captured samples (curves).*

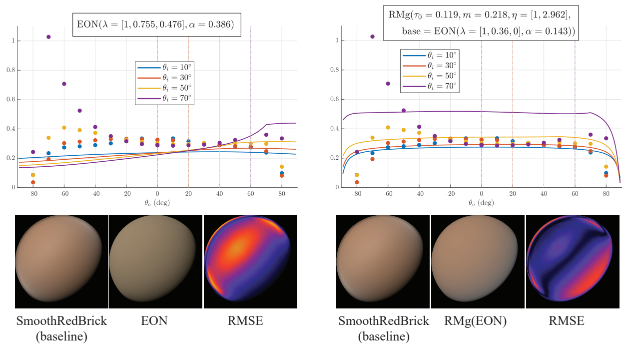

Figure 2: *Optimized EON and rough micrograin with EON base to a overall diffuse material SmoothRedBrick using an unweighted RMSE loss function. Top angular plots show model evaluations (continuous lines), and real samples (dots) for several incoming and outgoing inclinations (θ_i, θ_o). Bottom images show the baseline image (using a tabulated BRDF with real samples), the model image (using converged parameters) and the per-pixel RMSE.*

models along each axis. All optimizations for the BTDF fitting were performed using the Levenberg-Marquardt algorithm with the Lourakis library [Lou04].

3. Results

3.1. BRDF

Effect of sampling geometries In Figure 1 we analyze the effect of fitting BRDF models using in-plane samples only, and including also off-plane and near-specular samples. Error images on the right are computed as a per-pixel SMAPE between an image renderd with the fitted model, and an image rendered with a tabulated BRDF baseline of the captured material W80&G40 that includes all available samples (in-plane, off-plane, and near-specular). Overall, coarse in-plane measurements fail to reproduce the width of the specular peaks (as observed in the angular plots), while including denser near-specular and off-plane samples improves the result across all models.

Reproducing diffuse appearance Figure 2 compares EON and rough micrograin with EON base optimized to match a overall diffuse material SmoothRedBrick using an unweighted RMSE loss function. The optimized micrograin-based models (right, continuous) provides better fits of real reflectance at incident directions close to the surface normal (i.e. low θ_i), while overestimating reflectance at grazing incident angles near the dark side of the sphere (e.g. $\theta_i = 70^\circ$, purple). Converged parameters for EON yield a better fit at grazing incident angles (θ_i, purple) but fail to reproduce the characteristic decay at outgoing directions more perpendicular to the normal (i.e. large θ_o).

Effect of uncertainties Figure 3 illustrates how unweighted RMSE (left) and uncertainty-weighted RMSE (right) affect the optimization of an EON-based diffuse micrograin model. The model is fitted to BlackVelvet, a overall diffuse material with very dark albedo. Incorporating knowledge of uncertainties into the RMSE loss (right), shifts the model to improve the reflectance estimation at retroreflective angles ($\theta_i = \theta_o$), pulling the model curve closer to the real samples, while penalizing the BRDF values at low incident light angles. This shift can also be observed in the rendered sphere: using uncertainty-weighted optimization losses (right), the model errors decrease near the shadowed region, and increase near the top-left sphere region whose normal matches the incident direction.

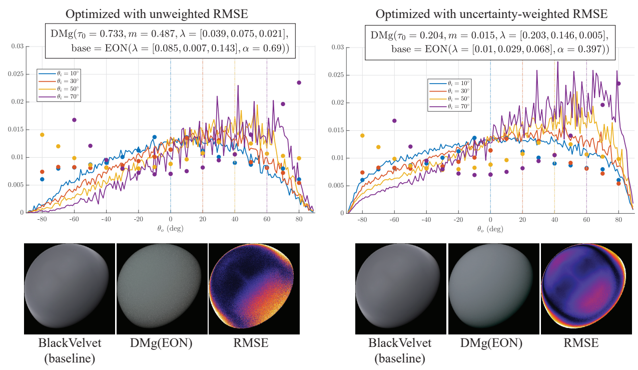

Figure 3: *Effect of unweighted RMSE (left) and uncertainty-weighted RMSE (right) in the optimization of a significantly dark diffuse micrograin model with EON base to fit a dark diffuse material BlackVelvet with known sample uncertainties. Top angular plots show model evaluations (continuous lines), and real samples (dots) for several incoming and outgoing inclinations (θ_i, θ_o). Bottom images show the baseline image (using a tabulated BRDF with real samples), the model image (using converged parameters) and the per-pixel RMSE.*

Parameter consistency A relevant aspect in appearance reproduction is analyzing how accurately optimized models converge to physically-based parameters that are consistent with real-world material properties. While, optimized models may quantitatively converge to the measured samples of real materials, yield visually plausible results, the resulting model parameters are not guaranteed to match the values of real material properties, even yielding complete implausible values. Figures 2 and 3 show how converged albedos λ in all models and their nested models converge to plausible high and low values, respectively. Our capture datasets include materials W95&G39.5, W80&G40, W80&G10.9, W80&G3.4, where W and G coefficient relate to albedo (in percentage value) and glossiness, respectively. Overall, our optimizations with lowest errors converge to albedo parameters that are close to the W coefficient. However, for several common dielectric materials included in our dataset, such as rough red brick, our optimizations yield quantiatively close results but implausible parameters, such as refractive indices close to 5. This highlights the importance of constraining optimizations to plausible parameters by e.g. adding physically-based priors to the optimization algorithms.

3.2. BTDF

Figure 4 compares the different BTDF models on several measured samples. The main difference between the models lies in their ability to reproduce the angular decay of the scattering distribution. TR and GTR models exhibit longer tails, making them better suited for materials with broad angular scattering, whereas Beckmann and GG behave more like Gaussian distributions with a faster decay. As a consequence, TR and GTR generally provide better fits for long-tail scattering profiles such as sample B.

The GG model nevertheless appears more versatile for specific scattering shapes. In particular, it better reproduces the square-shaped scattering pattern observed on sample A, which cannot be accurately represented by the other models. Overall, the choice of the normal distribution function (D) has a stronger impact on the fitting quality than the masking-shadowing term (G).

Adding a scaling or an absorbent layer. Figure 5 compares the use of a simple scaling factor and an absorption layer. Introducing an additional attenuation term is particularly important for quasi-Lambertian samples, where multiple scattering effects induce a significant loss of transmitted energy. In practice, however, the absorption-layer formulation never produced better results than a simple scaling factor and often degraded significantly the optimization results.

Interestingly, the optimized scaling factor tends to remain below 1 when TR or GTR provide a very good fit to the measured data, as observed for sample B. Conversely, when Gaussian-based models such as Beckmann or GG better reproduce the scattering shape, the scaling factor frequently exceeds 1, compensating for missing energy in the model.

Loss and uncertainty. Figure 6 illustrates the influence of the loss function and uncertainty weighting. Since the TR model already achieves very accurate fits for most samples, the choice of loss function has only a limited impact in these cases. Differences become more visible when the model fails to accurately reproduce the scattering profile, for instance when fitting long-tail distributions with the Beckmann model. Using the cosine-weighted formulation tends to favor the fit around the central peak ($\theta_t = 0°$), whereas logarithmic formulations reduce the influence of the peak and improve the fit of the tail region. Finally, incorporating measurement uncertainties into the optimization generally favors long-tail fitting. This behavior is explained by the larger uncertainties near the central peak, which reduce its contribution to the overall loss.

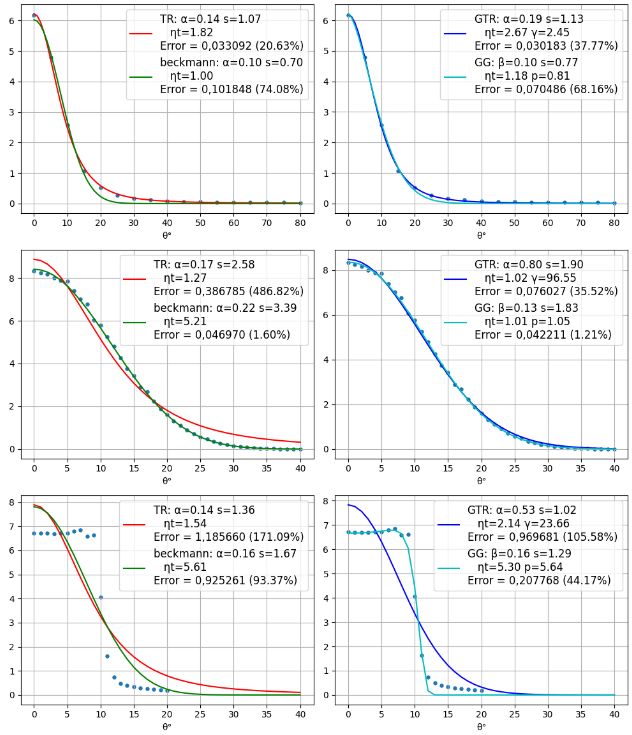

Figure 4: *BTDF fit with the different model: TR (red), Beckmann (green), GTR (blue) and GG (cyan) on different material samples. From top to bottom: B, D and A.*

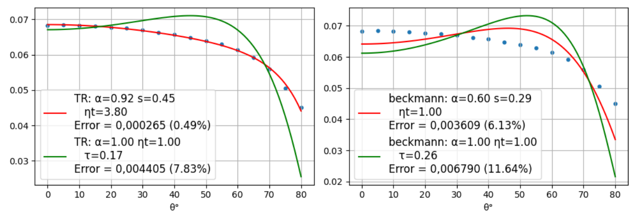

Figure 5: *Comparison between the scaling (red) and the absorbent layer (green) with two different model: TR (left), Beckmann (right); on the C sample.*

4. Discussion

Measurement uncertainties provided by metrologically-rigorous datasets represent a valuable asset to account for in BSDF model benchmarking and optimization. While they may not ensure global improvements on the the performance of converged models, they have the ability to shift the optimization efforts on BSDF regions that were measured more confidently. The inclusion of uncertainties in metrics and optimization algorithms is worth exploring in future metrology-oriented BRDF and BTDF models research.

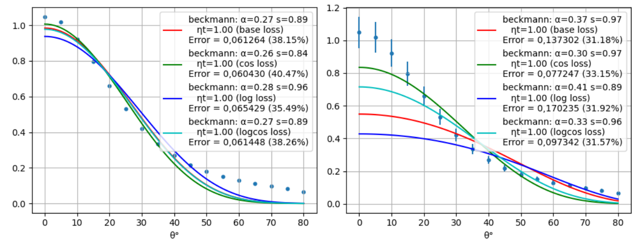

Figure 6: *Comparison of the different losses without incorporating the uncertainties (left) and with it (right) on Co1 sample. The vertical bars on the right represent the absolute uncertainties multiplied by 10.*

Despite the importance of collecting metrologically-rigorous spectral data, our current dataset has a very coarse 10-degree angular sampling rate, and is limited to in-plane measurements for most BRDFs. While this may suffice for quasi-Lambertian materials, higher sampling rates are necessary to capture not only strong high-frequency components of smooth materials, but also subtle roughness effects in diffuse materials [DJ18]. This hampers our analysis, and limits visual inspection to renders with coarse appearance that may not be representative of the final appearance.

References

[BS12] Burley B., Studios W. D. A.: Physically-based shading at disney. In *Acm siggraph* (2012), vol. 2012, vol. 2012, pp. 1–7. 2

[DJ18] Dupuy J., Jakob W.: An adaptive parameterization for efficient material acquisition and rendering. *Transactions on Graphics (Proceedings of SIGGRAPH Asia) 37*, 6 (Nov. 2018), 274:1–274:18. doi:10.1145/3272127.3275059. 5

[dW24] d'Eon E., Weidlich A.: Vmf diffuse: A unified rough diffuse brdf. In *Computer Graphics Forum* (2024), vol. 43, Wiley Online Library, p. e15149. 1

[FFQ*24] Fu J., Ferrero A., Quast T., Esslinger M., Santafé-Gabarda P., Tejedor N., Campos J., Gevaux L., Obein G., Aschan R., et al.: Survey of bidirectional transmittance distribution function measurement facilities by multilateral scale comparisons. *Metrologia 61*, 3 (2024), 035006. 2

[FVH14] Filip J., Vávra R., Havlícek M.: Effective acquisition of dense anisotropic brdf. In *2014 22nd International Conference on Pattern Recognition* (2014), pp. 2047–2052. doi:10.1109/ICPR.2014.357. 1

[HP17] Holzschuch N., Pacanowski R.: A two-scale microfacet reflectance model combining reflection and diffraction. *ACM Trans. Graph. 36*, 4 (July 2017). doi:10.1145/3072959.3073621. 2

[LKYU12] Löw J., Kronander J., Ynnerman A., Unger J.: Brdf models for accurate and efficient rendering of glossy surfaces. *ACM Transactions on Graphics (TOG) 31*, 1 (2012), 1–14. 2

[Lou04] Lourakis M.: levmar: Levenberg-marquardt nonlinear least squares algorithms in C/C++, Jul. 2004. [Accessed on 31 Jan. 2005.]. 3

[LRPB23] Lucas S., Ribardière M., Pacanowski R., Barla P.: A micrograin bsdf model for the rendering of porous layers. *ACM Trans. Graph. 42*, 6 (dec 2023). doi:10.1145/3610548.3618241. 1

[MPBM03] Matusik W., Pfister H., Brand M., McMillan L.: A data-driven reflectance model. In *ACM SIGGRAPH 2003 Papers* (New York, NY, USA, 2003), SIGGRAPH '03, Association for Computing Machinery, p. 759–769. doi:10.1145/1201775.882343. 1

[NDM05] Ngan A., Durand F., Matusik W.: Experimental analysis of brdf models. *Rendering Techniques 2005*, 16th (2005), 2. 2

[ON94] Oren M., Nayar S. K.: Generalization of lambert's reflectance model. In *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1994), SIGGRAPH '94, ACM, pp. 239–246. 1

[Pee23] Peers P.: BBM: A BRDF Benchmark. https://github.com/bsdfbenchmark/bbm, 2023. Documentation available at https://bbm.readthedocs.io/, Version 0.5.1. 1

[PKH25] Portsmouth J., Kutz P., Hill S.: Eon: A practical energy-preserving rough diffuse brdf. *Journal of Computer Graphics Techniques (JCGT) 14*, 1 (March 2025), 116–139. doi:10.48550/arXiv.2410.18026. 1

[WMLT07] Walter B., Marschner S. R., Li H., Torrance K. E.: Microfacet models for refraction through rough surfaces. *Rendering techniques 2007* (2007), 18th. 1, 2