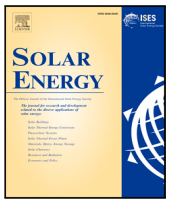
Solar Energy 300 (2025) 113726

Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener



# Scalable heliostat surface predictions from focal spots: Sim-to-Real transfer of inverse Deep Learning Raytracing

Jan Lewen [a,b,*], Max Pargmann [a], Mehdi Cherti [b], Jenia Jitsev [b], Robert Pitz-Paal [a], Daniel Maldonado Quinto [a]

[a] German Aerospace Center (DLR), Institute of Solar Research, Linder Höhe, D-51147 Köln, Germany
[b] Research Center Jülich, Jülich Supercomputing Centre, Wilhelm-Johnen-Straße, 52428 Jülich, Germany

ARTICLE INFO

*Keywords:*
Inverse deep learning raytracing
Deep learning
Heliostat surface prediction
Artificial intelligence
Zero-shot sim-to-real transfer
Heliostat digital twin

ABSTRACT

Concentrating Solar Power (CSP) plants are a key technology in the transition toward sustainable energy. A critical factor for their safe and efficient operation is the accurate distribution of concentrated solar flux on the receiver. However, flux densities from individual heliostats are highly sensitive to surface imperfections, such as canting and mirror deformations. Measuring these surfaces across hundreds or thousands of heliostats remains impractical in real-world deployments. As a result, control systems often assume idealized heliostat surfaces, leading to suboptimal performance and potential safety risks. To address this, inverse Deep Learning Raytracing (iDLR) has recently been introduced as a novel method for inferring heliostat surfaces from target images of focal spots captured during routine calibration procedures. However, until now, iDLR had only been demonstrated in simulation. In this work, we present the first successful Sim-to-Real transfer of iDLR, enabling accurate surface predictions directly from real-world target images. Remarkably, this was achieved through a zero-shot Sim-to-Real transfer, in which the model is trained exclusively with simulated flux density data and applied directly to real target images of heliostat focal spots without the need for additional training on real target images. We evaluate our method on 63 heliostats under real operational conditions. iDLR surface predictions achieve a median mean absolute error (MAE) of only 0.17 mm and show good agreement with deflectometry ground truth in 84% of cases. When used in raytracing simulations, it enables flux density predictions with a mean accuracy of 90% compared to deflectometry over our dataset, and outperforms the commonly used ideal heliostat surface assumption by 26%. We tested this approach in a challenging double-extrapolation scenario, involving unseen sun positions and receiver projections, and found that iDLR maintains high predictive accuracy, highlighting its generalization capabilities. Our results demonstrate that iDLR is a scalable, automated, and cost-effective solution for integrating realistic heliostat surface models into digital twins. This opens the door to an optimized heliostat control, improved flux density distributions on the receiver, and ultimately, enhanced efficiency and safety in future CSP plants.

## 1. Introduction

Concentrated Solar Power (CSP) is a key technology for the decarbonization of energy systems, offering the capability to deliver dispatchable electrical energy [1] and to drive solar thermochemical reactions for fuel production [2]. A crucial aspect of efficient CSP plant operation is the flux density distribution on the receiver, as each receiver has specific thermal thresholds. Proper flux density distribution minimizes material stress and degradation, thereby optimizing receiver performance. While the effects of tracking errors and the mandatory automatized calibration have been extensively studied [3–8], recently obtaining accurate flux density predictions for individual heliostats from easily available data sources for further optimized receiver performance has become a focus of research [9–12].

Traditionally, highly accurate flux density predictions required reconstructing the heliostat surface, accounting for mirror and canting errors, and simulating the resulting flux density using raytracing techniques [13,14]. Established methods for measuring heliostat surface deformation include deflectometry [13], laser scanning [15,16], photogrammetry [17,18,18–20], and flux mapping [21] (see Arancibia-Bulnes et al. [22] for a review). Despite their accuracy, these methods require manual labor and additional hardware, making them expensive and impractical for commercial-scale deployment aimed at creating digital twins of heliostats for better heliostat control. Hence, in most

* Corresponding author.
*E-mail address:* jan.lewen@dlr.de (J. Lewen).

https://doi.org/10.1016/j.solener.2025.113726
Received 2 April 2025; Received in revised form 19 May 2025; Accepted 21 June 2025
Available online 12 July 2025

**List of Abbreviations**

| | |
|---|---|
| CSP | Concentrating Solar Power |
| CSR | Circumsolar Ratio |
| DLR | Deutsches Zentrum für Luft- und Raumfahrt |
| GAN | Generative Adversarial Network |
| GPU | Graphics Processing Unit |
| iDLR | inverse Deep Learning Raytracing |
| IQR | Inter-Quartile Range |
| MAE | Mean Absolute Error |
| MLP | Multi-Layer Perceptron |
| NURBS | Non-Uniform Rational B-Spline |
| SSIM | Structural Similarity Index Measure |
| STJ | Solar Tower Jülich |

cases flux density predictions for heliostat control are made either through raytracing of an ideal heliostat or by adopting simplified flux density assumptions. To overcome this modeling error undermining the safe and efficient operation of the power plant, several works were done to obtain very precise flux density predictions without the need of new hardware installation and manual work.

Zhu et al. [10] introduced a post-installation calibration procedure that estimates four geometric parameters per facet through optimization, using only target images of the heliostat's focal spot. Despite the simplicity of the heliostat model, assuming only canting and focusing errors, the approach demonstrated significant efficiency gains up to 20%. This highlights the potential of integrating accurate flux density prediction into aim point optimization strategies. However, the model used by Zhu et al. [10] is limited in scope. For example, the heliostats at Solar Tower Jülich (STJ) in Germany have 4 canted facets, but the individual facets are flat and non-focussing. Their flux density is significantly influenced by surface characteristics such as waviness and facet edge bending, which are not captured by the method of Zhu et al. [10]. To address this, Pargmann et al. [11] introduced a more general framework that employs a differentiable heliostat model, where the heliostat surface is represented using Non-Uniform Rational B-Splines (NURBS). This model is coupled with a differentiable raytracing algorithm that enables gradient-based optimization of the NURBS surface. The optimization aims to adjust the surface geometry such that the simulated flux density, computed via the differentiable raytracer, aligns with the heliostat's focal spots. Nevertheless, due to the ill-posed and underdetermined nature of the inverse problem, the surface prediction is in most cases poor. This introduces uncertainty when extrapolating flux densities from the lambertian target plane to the receiver plane with potentially complex geometry. Alternatively, Kuhl et al. [12] presented a fully data-driven approach, representing each heliostat by a latent vector with no physical interpretation. This method achieves highly accurate flux density predictions and is straightforward to implement, as it requires no physical modeling. However, the lack of interpretability limits its integration with other physically grounded factors affecting flux density, such as soiling models, sun shape effects, shading, blocking, or extrapolation to complex receivers. To overcome the limitations of the above mentioned approaches, Lewen et al. [9] recently introduced inverse Deep Learning Raytracing (iDLR), a novel deep learning approach that aims to predict heliostat surfaces from target images of focal spots. This method utilizes information conserved in the shape of the flux density to predict the underlying surface. iDLR has demonstrated high accuracy and strong agreement with state-of-the-art surface measurement techniques such as deflectometry, based on simulated flux densities. However, up to date, its application was only shown on synthetic data, and the model has not yet been transferred to real-world target images of focal spots.

Recent advancements in deep learning have been fueled by the availability of large-scale datasets. However, across domains such as robotics [23–25], autonomous driving [26,27], and drone navigation [28] or as mentioned above iDLR [9] exclusive reliance on real-world data for training is often infeasible due to the high costs and time required for data acquisition. Consequently, simulated data play a central role in training models that are ultimately deployed in real-world settings. However, the inherent domain discrepancy between synthetic and real data, commonly referred to as the sim-to-real gap [23], poses a significant challenge to direct transferability, often leading to poor generalization in real-world environments. To enable a robust Sim-to-Real transfer of iDLR, three general strategies are conceivable. The first is fine-tuning the model on a limited amount of real-world data [29]. The second involves *Domain Adaptation*, where the discrepancy between the simulated and target domains is bridged, often using unsupervised deep learning models such as Generative Adversarial Networks (GANs) [30–33]. The third strategy is *Domain Randomization*, which aims to inject sufficient variability into the simulation environment such that the distribution of real-world data becomes a subset of the expanded simulated domain [23,25,34]. Among these, fine-tuning appears to be the most straightforward. However, this approach presents significant drawbacks. It requires paired real deflectometry measurements and target images, which are still expensive and challenging to obtain. More importantly, this would necessitate the deployment of surface measurement infrastructure at every power plant where iDLR is to be used, undermining the very advantage of eliminating the need for additional costly hardware. Therefore, this approach is impractical and counterproductive for broader adoption in CSP systems, and hence must be excluded, in favor of a zero-shot Sim-to-Real strategy where no further real target images are required for training.

In this work, we propose a zero-shot Sim-to-Real transfer method for inverse Deep Learning Raytracing (iDLR) that enables accurate heliostat surface reconstruction directly from images of their focal spots. Our approach combines Domain Adaptation and Domain Randomization to bridge the sim-to-real gap without training on difficult obtainable pairs of real world data. For Domain Adaptation, we follow the strategy of Zhang et al. [35], transforming real-world data to resemble simulated data instead of the other way around. Specifically, we leverage the deep learning framework by Kuhl et al. [36], which employs a UNet which predicts the underlying flux densities from target images of heliostat focal spots. As a result, real-world target images of focal spots are projected into the raytracing flux density domain. To further enhance generalization, we apply Domain Randomization during training, exposing the model to a wide range of variations, that resemble the difference between flux densities obtained from real target images with the UNet and simulated flux densities from raytracing. Together, these techniques enable the zero-shot Sim-to-Real transfer of iDLR to real-world data, making the approach both practical and cost-efficient. The method is described in detail in Section 2, with emphasis on the required real and synthetic datasets, the model architecture, and the training procedure. We validate our approach on a dataset comprising 63 heliostats operating under real conditions and demonstrate strong agreement with deflectometry-based ground truth for the vast majority of cases, as presented in Section 3. These results establish iDLR as a scalable and physically interpretable component for heliostat digital twins, with significant potential for improving monitoring, calibration, and control in future CSP plants.

## 2. Method

In this section, we present the steps for transferring iDLR to real world data. The general workflow of the iDLR approach is illustrated in Fig. 1. For each heliostat, a series of images are taken with a camera, capturing the heliostat's focal spot on a Lambertian Target as part of the Camera-Target [3] calibration procedure. These images will be

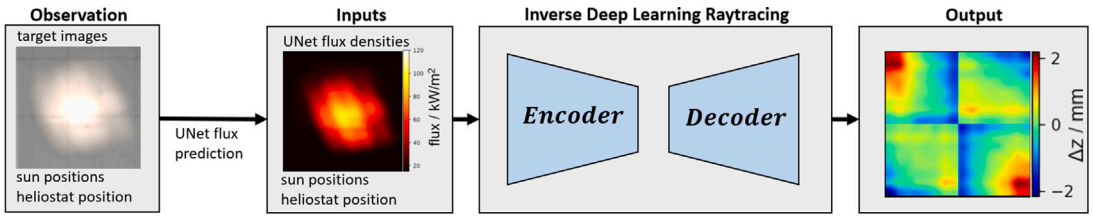



**Fig. 1.** Overview of the full inverse Deep Learning Raytracing (iDLR) pipeline. Real world target images are first processed by the UNet model proposed by Kuhl et al. [36], which translates them into UNet flux density predictions. Up to eight such UNet flux density predictions are then input to the iDLR model, which has been trained solely on simulated flux densities generated with a raytracer. The iDLR model subsequently infers the heliostat surface shape from the UNet flux density inputs.

referred to throughout the paper as *target images*. The target images are processed using the UNet-based method proposed by Kuhl et al. [36], which predicts the underlying flux densities, hereafter referred to as *UNet flux densities*. For inference, up to eight such UNet flux densities per heliostat, along with the corresponding sun positions and the heliostat's position, are provided as input to the iDLR model which predicts the surface of the heliostat. The iDLR model is trained on a semi-artificial dataset constituted by deflectometry surfaces and simulated flux densities from raytracing. In the following we will describe the data needed for training the model, the model architecture and the training routine.

### 2.1. Data

The key real data for training an iDLR model are measurements of heliostat surfaces. Since 2014, more than 600 deflectometry measurements have been conducted at the STJ facility in Germany (see Phipps et al. [37] for data collection from STJ).[1] Throughout this period, the same heliostat model was used, comprising four canted, flat non-focusing facets, each measuring 1.6 m × 1.25 m. The focal distance of the canting is for every heliostat the distance to the main receiver, or to a side receiver, whereby the minimum focal distance is 65 m. Between 2020 and 2022, modifications were made to the heliostat kinematics. As a result, the surface error characteristics of the heliostats changed, altering the corresponding flux density shapes. Consequently, older deflectometry measurements could not be used for validation or testing, as the flux density shape no longer match to the old heliostat surface. Nevertheless, because our model is trained on simulated raytracing flux densities, these earlier measurements remain suitable for training. To ensure data quality, deflectometry measurements with poor coverage or visible artifacts were excluded. A measurement was only considered valid if at least 95% of the heliostat surface was successfully captured. Additionally, visual inspection was performed to identify and discard measurements with potential artifacts. After this filtering process, a total of 456 deflectometry measurements were considered valid and included in our approach.

Each heliostat surface is represented using the NURBS-based model introduced by Lewen et al. [9]. Specifically, the surface of each facet is parametrized by an 8 × 8 grid of z-control points $P^z_{i,j}$ that define the corresponding NURBS spline. The dataset was split into training, validation, and testing subsets. A total of 393 measurements were used for training, representing 323 unique heliostats (as some were measured multiple times). The validation and test sets consist of 33 and 30 individual heliostats, respectively. All heliostats in the validation and test dataset are canted on the main receiver. Their positions within the STJ heliostat field are illustrated in Fig. 2.

As training a deep learning model on only 393 data points is difficult, augmentation techniques were applied to create new, but still realistic, synthetic heliostat surfaces. To augment the training dataset, two surface augmentation techniques were employed. First, heliostat surfaces were rotated by 180 deg. Second, new surfaces were generated by computing weighted averages between pairs of existing surfaces. In total, approximately 400,000 heliostat surfaces were used for model training. Of these, 80% are augmented surfaces, while the remaining 20% consist of real surfaces that were repeatedly positioned in the raytracing simulation over the field. In the next step, eight flux densities per heliostat surface were simulated using the raytracer of Pargmann et al. [11].[2] Compared to the approach described in Lewen et al. [9], our procedure differs in four aspects: First, in Lewen et al. [9], the eight sun positions per heliostat were sampled based on those at which real target images were taken. This approach introduced a bias, as summer sun positions were overrepresented. To mitigate this, we now sample sun positions from an equidistant grid in azimuth-elevation space, covering all possible sun positions throughout the year at the geographic location of the STJ. Second, we revised the normalization strategy for the flux densities. Previously, the sum of all pixel values in each flux density was normalized to 100, implicitly assuming equal total energy across different flux densities. However, this assumption does not hold for UNet flux densities, which may exhibit varying energy content. To better reflect this characteristics of UNet flux densities, we now normalize each flux density by scaling its maximum value to 1. Third, in line with the Domain Randomization strategy, we uniformly sample the Circumsolar Ratio (CSR) parameter for the Buie-CSR sunshape model within the range of 0 to 0.15. For every sun position a new CSR parameter is sampled. Fourth, the heliostat aim point for every simulated flux density is uniformly scattered within a 1 m radius around the target center. These two randomized parameters are integrated directly into the simulation process to improve model robustness. Additional randomization steps are applied during training to further diversify the dataset. The training data set is then formed by the raytraced flux densities, the corresponding sun positions and the heliostat position as input data and the surfaces.

While training is performed on this semi-artificial dataset, validation and testing are conducted using the UNet flux densities derived from real-world target images. At the STJ, several target images are acquired annually for calibration purposes. The number of target images per heliostat varies depending on the already achieved quality of the calibration. The target images were filtered to exclude those affected by artifacts such as overexposure, overlapping focal spots from multiple heliostats, and unwanted latent factors, including shading and blocking. The final target image dataset includes between three and eight target images per heliostat. Then, in accordance with the Domain Adaptation strategy for the Sim-to-Real transfer, the UNet model introduced by Kuhl et al. [36] is applied to these target images to generate flux density predictions from the target image, which serve as input for inference with the simulatively trained iDLR model. The validation dataset is used during training to monitor model performance and guide architectural and hyperparameter decisions based on validation loss. The test dataset is used solely to evaluate the final model after training. As no significant performance differences were observed between the validation and test sets, results presented in the subsequent section are reported jointly for both datasets to improve statistical robustness.

[1] Link to data collection: https://paint-database.org.

[2] The raytracer was released open source: https://github.com/ARTIST-Association/ARTIST.

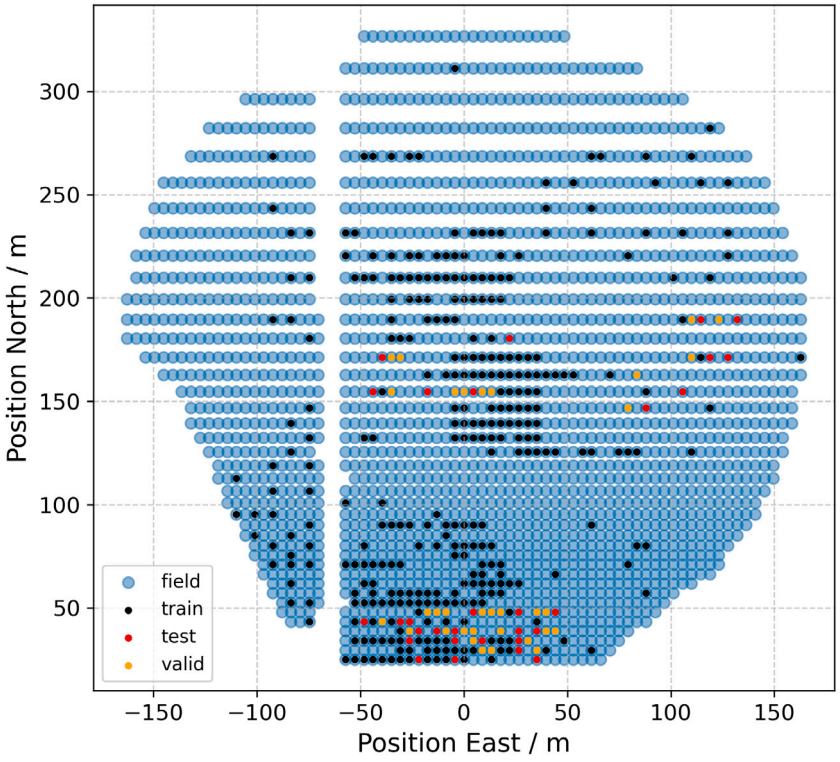



**Fig. 2.** The heliostat field of the Solar tower Jülich. The heliostat's which were used for training, validating and testing the iDLR model are highlighted.

### 2.2. Model

The model architecture is illustrated in Fig. 3. It follows an encoder–decoder design, where the encoder consists of two primary components: a Vision Transformer (upper path), as introduced by Dosovitskiy et al. [38], and a modified standard Transformer Encoder (lower path), originally proposed by Vaswani et al. [39]. The Vision Transformer is responsible for processing the input flux densities. First, each flux density is divided into patches of size 16, which are then passed through a Multi-Layer Perceptron (MLP). Positional embeddings and a class embedding are added to the patch embeddings before they are input into the Transformer Block. The core computation occurs in the Multi-Head Attention Block, where four parallel attention heads are applied. The output is then passed through another MLP, and a residual connection is added. This Transformer Block is repeated $L_1 = 8$ times, and all MLP layers within this module have a dimensionality of 128. The output is a compact feature vector of dimension 128. This feature vector is then combined with positional information describing the heliostat and sun positions. These positional inputs are processed by an MLP to produce embeddings, which are then summed with the flux density feature vector. The combined representation is passed into the second part of the encoder, a modified Transformer Encoder. This module uses Multi-Head Attention with five attention heads and is repeated $L_2 = 8$ times. The output of this encoder is a latent representation encoding the heliostat surface geometry. The decoder is implemented using a StyleGAN2 generator [40,41], which takes as an input the w+ latent space as introduced by Richardson et al. [42]. Since the generator comprises three style blocks, the w+ latent vector has a dimensionality of $3 \times 32$, with the value 32 selected empirically. Each latent vector is projected via an affine transformation and weight demodulation into the generator, which includes a learnable constant input in the initial block, followed by convolutional and upscaling layers. To ensure stable training and consistent feature scaling, layer normalization is applied at the locations indicated in Fig. 3. Additionally, although not shown in the figure, a Dropout layer with a dropout rate of 0.2 is applied after each MLP to mitigate overfitting. In total the model has slightly more than 4 million trainable parameters.

### 2.3. Training and validation

The training procedure differs in one key aspect from that described in Lewen et al. [9]: the incorporation of randomization techniques during training to improve the model's robustness when applied to real-world data. Although two forms of randomization were already introduced during the simulation process, we chose-for performance and flexibility reasons-to primarily apply additional randomizations directly to the precomputed raytraced flux densities during training. All modifications to the raytraced flux densities are designed to reflect plausible physical phenomena influencing the UNet images. Discrepancies between raytraced flux densities and UNet flux densities typically arise from two major sources: (1) latent parameters related to the heliostat and the environmental conditions at the power plant, and (2) measurement errors or artifacts introduced by the camera system. Latent parameters include, for example, variations in the sun shape, mirror soiling, shading and blocking effects, microscopic surface errors, inaccuracies in coordinate data, non-ideal Lambertian target behavior, and background radiation. Camera-related artifacts include overexposure, underexposure, and uncertainties in contrast settings. To improve the model's ability to generalize to real-world measurements, the following randomizations were introduced during training:

- Random exclusion of individual flux densities, encouraging the model to perform well with fewer observations.
- Uniform clamping of flux densities in the range (0.9, 1.0) to simulate overexposure effects.
- Small perturbations to the heliostat and sun positions to simulate coordinate uncertainties.
- Adding small amount of Gaussian noise to the surface NURBS to mimic potential errors in the deflectometry data.
- Background noise distributed over the flux density to represent environmental radiation.
- Random adjustment of contrast, simulating the unknown contrast parameter of the camera system.

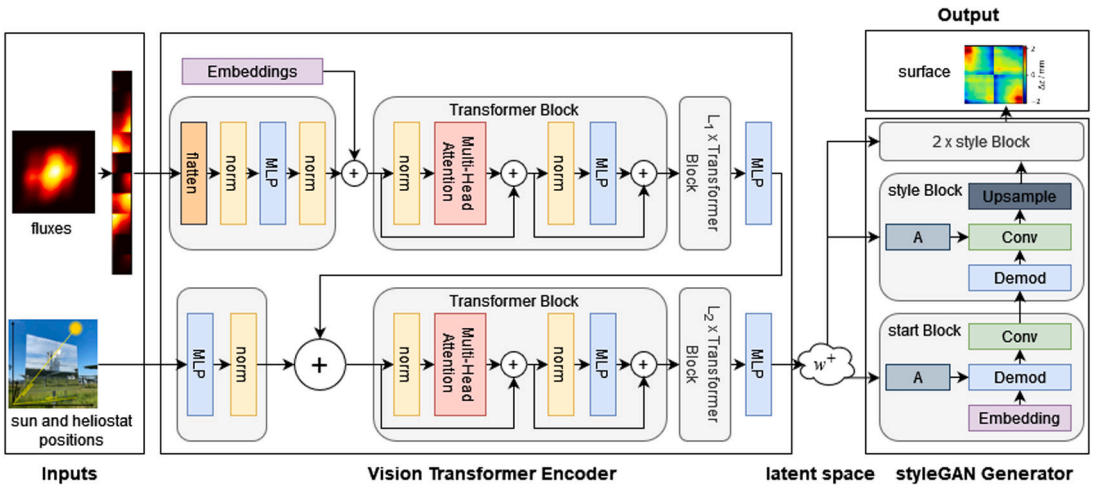



**Fig. 3.** The iDLR architecture consists of a Vision Transformer based Encoder and a styleGAN2 Generator.

- Random cropping by one pixel on a randomly selected subset of flux density edges (each edge cropped at most once), followed by interpolation back to the original resolution. This transformation effectively alters the size and shape of the flux density and can mimic effects such as soiling, microscopic mirror deformation, and inaccuracies in the Lambertian target dimensions.
- Small random deformations applied to the flux density, primarily to simulate deviations from ideal Lambertian reflection behavior.
- Bilinear smoothing to slightly blur and broaden the flux density, diminishing fine-grained details.

The randomization hyperparameters were determined through a combination of two approaches. First, we performed visual inspection by comparing randomized raytraced flux densities to UNet flux densities, with the goal of achieving greater visual similarity. Second, we conducted an iterative refinement process guided by these visual assessments and the quality of surface predictions on the validation dataset after training a model. Each of the above randomizations is applied independently with a probability of 0.5 per raytraced flux density in each training batch. As a result, the model almost never encounters an unaltered raytraced flux density during training, promoting strong generalization and robustness.

The model was trained for 50 epochs using four NVIDIA A100 GPUs, with a total training time of approximately five hours, including the computational overhead of validation and testing via raytracing. The initial learning rate was set to 0.001 and was exponentially decayed using a gamma factor of 0.995. Weight decay regularization was applied with a factor of $1 \times 10^{-7}$. The Adam optimizer was employed with parameters $\beta = (0.9, 0.999)$ and $\varepsilon = 1 \times 10^{-8}$. The training objective is to minimize the *Mean Absolute Error* (MAE) between the predicted z-control points, denoted as $P_{i,j}^{z,\mathrm{iDLR}}$, and those derived from deflectometry measurements, $P_{i,j}^{z,\mathrm{Defl}}$:

$$\mathrm{MAE}(P^{z,\mathrm{iDLR}}, P^{z,\mathrm{Defl}}) = \frac{1}{N} \sum_{i=1}^{n=8} \sum_{j=1}^{m=8} \left| P_{i,j}^{z,\mathrm{iDLR}} - P_{i,j}^{z,\mathrm{Defl}} \right| \tag{1}$$

where:

- n and m denote the number of control points in the horizontal and vertical directions, respectively.
- N = n × m is the total number of control points per facet.

Although MAE is an intuitive and widely used error metric, it is sensitive to the absolute magnitude of surface deviations. Since individual heliostats can vary significantly in terms of canting and surface deformation, MAE alone may not provide a fully representative measure of prediction quality across heliostats with different error magnitudes. To address this, we also evaluate the *Structural Similarity Index Measure* (SSIM), a perceptual metric that better captures the structural similarity between the predicted and measured heliostat surfaces. SSIM is defined as:

$$\mathrm{SSIM}(P^{z,\mathrm{iDLR}}, P^{z,\mathrm{Defl}}) = \frac{(2\mu_{P^{z,\mathrm{iDLR}}}\mu_{P^{z,\mathrm{Defl}}} + C_1)(2\sigma_{P^{z,\mathrm{iDLR}},P^{z,\mathrm{Defl}}} + C_2)}{(\mu^2_{P^{z,\mathrm{iDLR}}} + \mu^2_{P^{z,\mathrm{Defl}}} + C_1)(\sigma^2_{P^{z,\mathrm{iDLR}}} + \sigma^2_{P^{z,\mathrm{Defl}}} + C_2)} \tag{2}$$

where:

- $P^{z,\mathrm{iDLR}}$ and $P^{z,\mathrm{Defl}}$ are the iDLR and Deflectometry z-control points.
- $\mu_{P^{z,\mathrm{iDLR}}}$, $\mu_{P^{z,\mathrm{Defl}}}$ are the mean values of $P^{z,\mathrm{iDLR}}$ and $P^{z,\mathrm{Defl}}$.
- $\sigma^2_{P^{z,\mathrm{iDLR}}}$, $\sigma^2_{P^{z,\mathrm{Defl}}}$ are the variances of $P^{z,\mathrm{iDLR}}$ and $P^{z,\mathrm{Defl}}$.
- $\sigma_{P^{z,\mathrm{iDLR}},P^{z,\mathrm{Defl}}}$ is the covariance between $P^{z,\mathrm{iDLR}}$ and $P^{z,\mathrm{Defl}}$.
- $C_1$ and $C_2$ are small constants to stabilize the division when the denominators are close to zero.

SSIM values range from −1 to 1, where higher values indicate greater structural similarity. The interpretation of SSIM values for surface comparison is as follows:

- $0.75 \leq \mathrm{SSIM} \leq 1.0$: Very High similarity
- $0.50 \leq \mathrm{SSIM} < 0.75$: High similarity
- $0.25 \leq \mathrm{SSIM} < 0.50$: Medium similarity, usually the main features are still predicted good
- $-0.25 < \mathrm{SSIM} < 0.25$: No similarity
- $\mathrm{SSIM} \leq -0.25$: Negative similarity

SSIM is more robust to changes in absolute magnitude and focuses instead on the structural consistency of predicted surfaces, making it the preferred metric for comparing individual heliostats with varying deformation levels. However, for comparison of different models on the same dataset the MAE is preferably used because of its easier physical interpretability.

To quantify the quality of a flux density prediction relative to a flux density ground truth, the following accuracy metric is used:

$$\mathrm{ACC}_{\mathrm{GT,pred}} = \frac{\sum |\phi_{\mathrm{GT}} - \phi_{\mathrm{pred}}|}{\sum |\phi_{\mathrm{GT}}|} \tag{3}$$

Throughout this work, raytracing is extensively used to compare flux densities derived from deflectometry-based surfaces, iDLR-predicted surfaces, and an idealized heliostat surface without any surface errors. Throughout the manuscript, we refer to these as *deflectometry flux density*, *iDLR flux density*, and *ideal flux density*, respectively. All raytracing

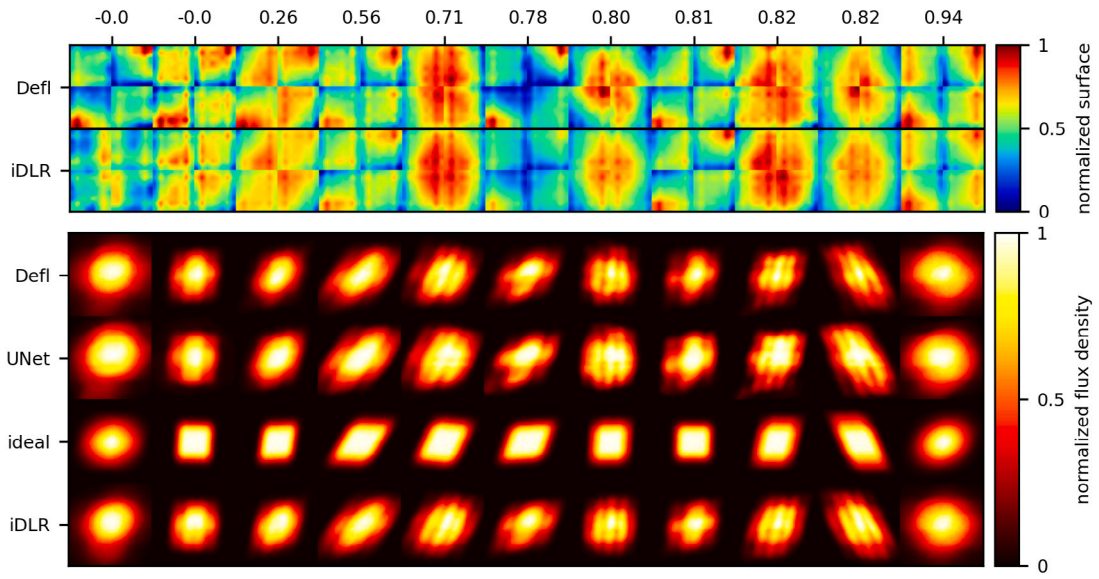



**Fig. 4.** The top panel compares the iDLR surface prediction with Deflectometry measurements for ten heliostats at STJ. Above each prediction the SSIM between iDLR surface and Deflectometry is given. The lower panel compares the iDLR flux densities, the deflectometry flux densities, the UNet flux densities derived from real target images and the ideal flux density.

simulations are performed using approximately 2.3 million rays per heliostat. The simulations employ the CSR sun shape model by Buie et al. [43], with a CSR value of 7.5%. The resulting flux densities are centralized and cropped to a spatial extent of 4 m × 4 m. To isolate the influence of the heliostat surface on the resulting flux density, the simulations do not incorporate tracking errors, microscopic surface perturbations, or soiling effects.

## 3. Results

Fig. 4 presents a visual comparison of the iDLR surface predictions against deflectometry measurements (top row), alongside a comparison of the iDLR flux densities, the deflectometry flux densities, the UNet flux densities and the ideal flux densities. The surface predictions are ordered by their SSIM, ranging from the lowest (left) to the highest (right). The corresponding SSIM value is annotated above each surface prediction. For the 63 heliostats in the combined validation and test datasets, the MAE is on median 0.17 mm with an Inter-Quartile Range (IQR) of 0.08 mm, while corresponding SSIM values are 0.77 and 0.32, indicating that the median iDLR prediction has a very high similarity to deflectometry. These metrics indicate a high degree of accuracy in surface prediction for the majority of heliostats. However, a small subset of predictions exhibit significant deviations. Based on our observations, an SSIM value above 0.25 typically reflects the correct reconstruction of at least the main structural features of the surface. This can be seen in Fig. 4 in the third prediction from the left, which is the predictions with the smallest SSIM above 0.25. Therefore, we define a misprediction as any result with $SSIM < 0.25$. According to this criterion, 16% of the predictions are considered mispredictions, while 84% exhibit similarity to the deflectometry ground truth. By raytracing heliostats based on these surfaces, the iDLR flux densities achieve a median accuracy of 0.92 with respect to the deflectometry flux densities, with an IQR of only 0.03. This shows, that even for cases where surface prediction quality is relatively poor, the resulting flux density predictions usually remain close to the deflectometry-based ground truth. This can be attributed to the underdetermined nature of the inverse problem, where multiple surfaces may yield similar flux densities. Furthermore, the iDLR flux density predictions can be compared to the UNet flux densities derived from real target images. The agreement between iDLR and UNet flux density yields a median accuracy of 0.86 with an IQR of 0.06, which is slightly lower than the values observed when comparing to the deflectometry flux densities.

To assess the effectiveness of our Sim-to-Real transfer approach, we compare the surface prediction performance of the full iDLR pipeline against a baseline model without Sim-to-Real transfer. In the baseline setup, inference is performed directly on raw target images, bypassing the UNet-based Domain Adaptation. Additionally, no Domain Randomization techniques are applied during training. Fig. 5 presents the distribution of the MAE for both configurations. Without Sim-to-Real transfer, the model fails to generalize to real-world data in most cases, resulting in a more than threefold increase in the MAE. This highlights the critical role of Sim-to-Real transfer in enabling accurate heliostat surface reconstruction from real target images using iDLR.

Fig. 6 illustrates the distribution of the surface prediction error as a function of the heliostat's distance to the tower, evaluated using both the MAE and SSIM. For both metrics, the accuracy of the surface predictions improves with increasing distance from the tower. Notably, not only does the average prediction error decrease, but the variance of the predictions also becomes smaller, with the exception of a few outliers.

While the surface predictions are generally similar to those obtained via deflectometry, so far it is uncertain whether they are sufficiently accurate for predicting flux densities on the receiver for new sun positions. To assess this, we designed a double-extrapolation scenario, a particularly challenging case for the model. Surface predictions were generated using target images from two lambertian targets located at the STJ, at heights of 36 m and 43 m, primarily recorded during summer. Flux densities were then computed via raytracing on the STJ receiver, positioned 55 m above ground, curved with a radius of 4.14 m, and featuring an opening angle of 60 deg, tilted 25 deg toward the ground. The simulation assumed a sun position of $(-1, -1, 1)$ in (east, north, up) coordinates, corresponding to an autumn/spring sun position in Germany. This setup represents a strong extrapolation scenario, as the test sun position in spring/autumn significantly deviates from the sun positions under which most of the target images for inference were acquired. For the comparison, we calculated besides the iDLR flux densities, deflectometry and ideal flux densities. Results are presented in Fig. 7. The iDLR flux densities achieve a mean accuracy of 90% relative to the deflectometry flux densities, indicating strong extrapolation capability. However, in the case of two heliostats with relatively poor surface predictions, accuracy dropped below 80%.

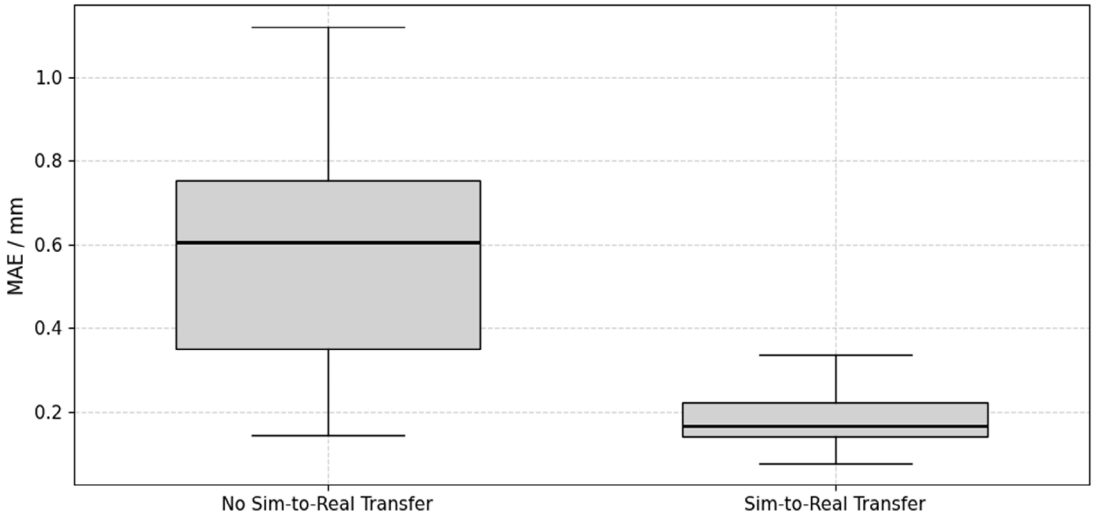



**Fig. 5.** The left boxplot shows the model performance without Sim-to-Real techniques, while the right boxplot shows the significantly improved performance when the full Sim-to-Real pipeline is applied.

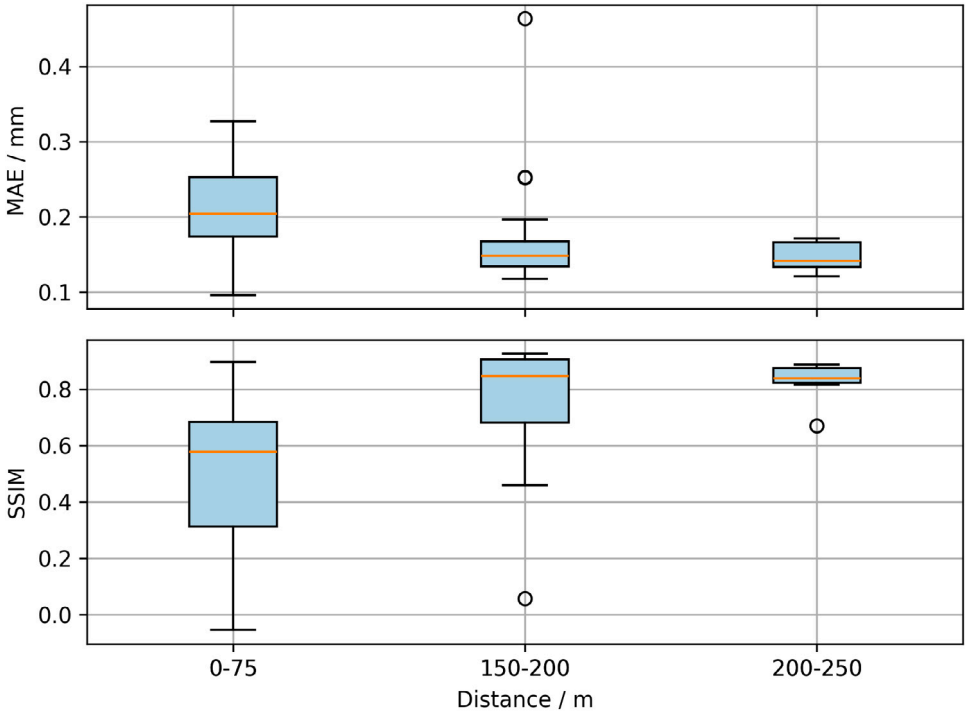



**Fig. 6.** Surface prediction error over the distance of the heliostat to the tower.

This is expected, as surface inaccuracies can have a more pronounced effect in extrapolation scenarios due to underdetermined conditions. Nevertheless, for the majority of heliostats, the extrapolation performs exceptionally well, with minimal accuracy loss. By contrast, the ideal heliostat assumption achieves only 64% accuracy, meaning that iDLR outperforms it by 26 percentage points. Additionally, the accuracy distribution of iDLR is more consistent, exhibiting a standard deviation of 4%, compared to 9% for the ideal heliostat assumption. Finally, we superposed the iDLR flux densities on the receiver and compared them the deflectometry and ideal flux densities, as shown in Fig. 8. The superposition was performed using six aim points arranged in a grid to distribute the irradiation on the receiver. The differences between the iDLR and deflectometry flux densities on the receiver are minimal. The resulting accuracy of iDLR compared to deflectometry is 97.5%, whereas the ideal heliostat assumption achieves only 89.5%. The improved accuracy of flux density predictions achieved through iDLR, compared to the idealized assumption of a perfect mirror surface, results in a clear reduction of localized discrepancies in the predicted flux density distribution on the receiver(see Fig. 8, lower row). This highlights the potential risks of neglecting mirror deformations in flux density distribution predictions, which may lead to localized overheating of the receiver or structural damage due to thermally induced stresses.

## 4. Discussion

We have combine Domain Adaptation and Domain Randomization to transfer iDLR to real data and predict heliostat surfaces using target images that were originally taken for routine calibration purposes. The results show that its possible to predict heliostat surfaces with only a

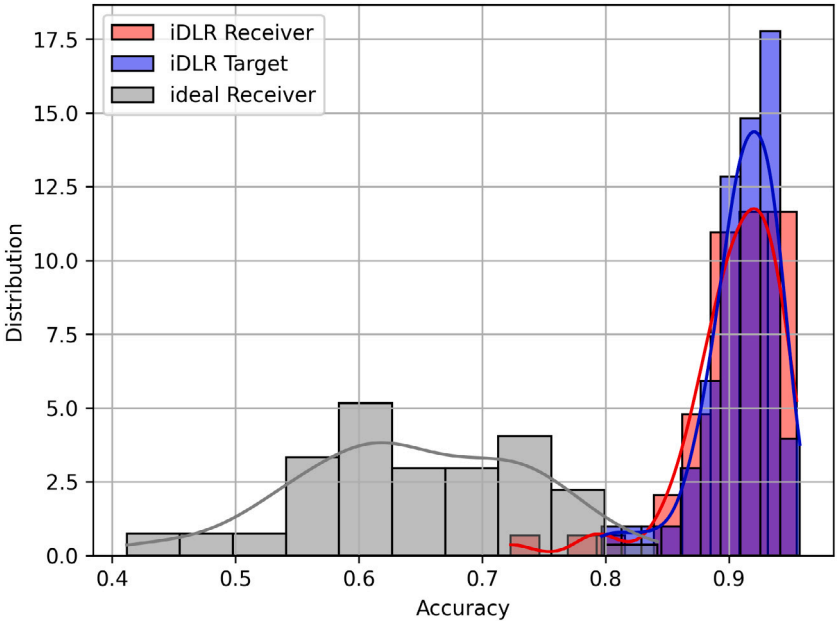



**Fig. 7.** The distribution of the accuracy for iDLR flux density on the Lambertian Target (blue) and the receiver (red). For comparison the accuracy of the ideal heliostat assumption on the receiver is shown (gray). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

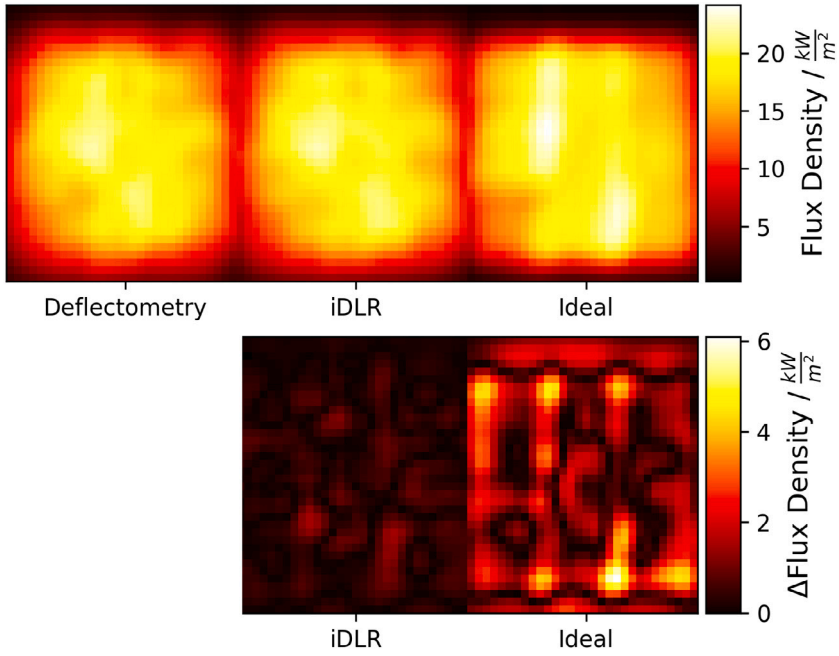



**Fig. 8.** Top row: Comparison of raytraced flux density distributions for all 63 validation and test heliostats on the receiver at STJ. The left panel shows the flux density distribution based on deflectometry measurements, the middle panel shows the distribution simulated from iDLR surfaces, and the right panel shows the ideal distribution assuming a perfect heliostat surface without mirror errors. Bottom row: Difference maps between the deflectometry flux density distribution and the iDLR (left) and ideal (right) flux density distributions, highlighting local deviations between the flux density distribution.

small loss compared to deflectometry (median MAE of 0.17 mm) for the majority of the heliostats, while there exists a risk of getting a poor surface prediction for around 16% of the heliostats. However, as the problem is strongly underdetermined and ill-posed, we found that despite the occasionally poor surface predictions, the arising flux density prediction on the target and on the receiver are extremely accurate, with a median accuracy of 92% and 90% compared to deflectometry.

The proposed Zero-Shot Sim-to-Real transfer approach has demonstrated strong effectiveness. Without this transfer mechanism, models trained exclusively on simulated flux densities would struggle to generalize effectively to real-world target images. While the Sim-to-Real transfer process achieves promising results, a slight performance gap remains between simulated and real data: the median MAE on real-world data is approximately 0.05 mm higher than on simulated data. Though this discrepancy is relatively minor, it indicates potential for further refinement of the Sim-to-Real transfer pipeline to achieve closer alignment between domains. It is important to note that while the Sim-to-Real transfer mechanism itself is Zero-Shot, requiring no training on real target images-the broader approach, iDLR, is not entirely Zero-Shot. This is because real deflectometry measurements are necessary to train the initial model using simulated flux densities as inputs. Attempts to replace the real deflectometry data with synthetic data did not yield

satisfactory results. This limitation is likely due to the need for the model to learn physically plausible surface deformations from real data, enabling robust predictions despite the inherent underdetermination of the approach.

We found that surface prediction quality improves with increasing distance from the tower. This is consistent with the simulation results reported by Lewen et al. [9]. As discussed therein, two opposing effects contribute to this observation. First, a constant canting error leads to a reduction in underdetermination when the heliostat is positioned farther from the target. Second, surface deformations with higher spatial frequencies, such as waviness, introduce greater underdetermination for distant heliostats. As a result, canting errors become more predictable when the heliostats are located farther from the tower. While theoretically the inverse holds for waviness-i.e., the prediction of high-frequency surface features should improve at closer distances, in practice, the prediction of waviness from target images remains strongly underdetermined due to canting effects. Canting superimposes flux densities of all four facets, making it difficult to attribute specific flux density features to individual facets. To resolve finer surface features, the model relies heavily on physical surface characteristics learned during training. Based on this learned knowledge, the model can infer finer surface deformations in a manner largely independent of the heliostat's distance from the tower. Consequently, the overall surface prediction quality increases with greater distance. However, this also implies that the effect may not generalize to other heliostat models. For example, in the case of a heliostat with non-canted facets, it is likely that the trend is reversed-i.e., prediction quality is higher for nearby heliostats compared to those farther away.

Our results show that, despite small differences between the iDLR surfaces and deflectometry, the iDLR surfaces are sufficiently accurate to enable reliable flux density predictions, both on the target and on the receiver, when used in raytracing simulations. First, for predictions on the target, the iDLR flux densities demonstrated very high accuracy to the deflectometry flux densities, achieving a median accuracy of 92%. When compared to the UNet flux densities derived from real target images, the iDLR flux densities achieved a slightly lower but still high median accuracy of 86%. For reference, the agreement between deflectometry flux densities and UNet flux densities is 87%. This indicates that the reduced accuracy between iDLR and UNet flux densities is not primarily due to inaccuracies in the predicted surface, but rather due to latent parameters that are not explicitly modeled in the raytracing simulations. These include shading and blocking effects, soiling, microscopic surface errors, inaccuracies in the geometry model, and variations in the actual sun shape. Additionally, the UNet flux densities themselves are subject to uncertainties and artifacts such as overexposure, underexposure, deviations from ideal Lambertian behavior of the target, and unknown camera contrast settings, all of which contribute to the observed discrepancies. However, this loss in accuracy due to unknown latent parameters must not necessarily occur when deployed. In practice, iDLR can be combined with physical and data-driven models that estimate these additional factors. Such models could include, heliostat soiling estimators or CSR sun shape measurements, which could be incorporated into the heliostat field digital twin and further increase the accuracy of flux density predictions. Second, for validating flux density predictions on the receiver, we designed a challenging double-extrapolation scenario: flux densities on the receiver during autumn were raytraced using surface predictions derived primarily from summer target images, taken on a Lambertian target located around 15 m below the receiver. Despite this extrapolation in both time (seasonal variation) and space (from target to receiver/ flat surface to curved surface), the predicted flux densities retained 90% of the accuracy achieved with deflectometry. This corresponds to an additional error of only 2% due to the extrapolation, highlighting the robustness and generalization capability of iDLR surface predictions. When compared to the ideal heliostat assumption, the iDLR flux density is 26 percentage points more accurate relative to the deflectometry flux density. This highlights the substantial improvement over the large modeling error introduced by the ideal heliostat assumption. Moreover, the superposed iDLR flux density distribution on the receiver across all 63 test and validation heliostats achieved an impressive 97.5% agreement with deflectometry flux density simulations on the receiver. The improved accuracy of the iDLR flux density distribution, compared to the ideal heliostat assumption, results in a clear reduction of localized discrepancies in the predicted flux density distribution on the receiver. These results strongly support the integration of iDLR surfaces into heliostat digital twins for enhanced flux density prediction and optimized heliostat control.

In comparison to manual heliostat surface measurement techniques such as deflectometry [13], laser scanning [15,16], photogrammetry [17,18,18–20], or flux mapping [21], the key advantage of iDLR lies in its automation and scalability. Once trained, the iDLR framework operates at virtually zero cost and can be deployed with minimal effort, an essential requirement for integrating surface predictions into digital twins of heliostats for optimized solar field control. However, some limitations must be considered. First, iDLR requires surface measurement data for training, necessitating at least one complete measurement setup per heliostat model. While such data can be collected at any CSP plant, a more convenient approach may be to conduct measurements directly at the manufacturing site. Although this dataset may not capture all surface deformations introduced during deployment, it can still serve as a valuable resource for pretraining the model. Still, for heliostat models with only a limited number of units, acquiring the required 300 surface measurements may be challenging. Despite this, in a future scenario where CSP is deployed at scale and only a few high-volume heliostat models dominate the market, several hundred surface measurements may be easy obtainable for all those heliostat models. Both manufacturers and plant operators are likely to periodically perform surface measurements for quality control, which can rapidly accumulate into sufficient training datasets. Another inherent limitation of iDLR is its statistical nature. Even though iDLR flux densities match closely with UNet flux densities from the observed target image, the underdetermined nature of the inverse problem implies that there is still uncertainty if the iDLR surface prediction is actually similar to the real heliostat surface. We found empirically that the probability of having a poor surface prediction is 16%. This restricts the applicability of iDLR in scenarios where highly reliable surface reconstruction is critical, for instance, during acceptance testing of newly delivered heliostats, rather than for downstream tasks like flux density prediction where surface variations have hardly any impact.

The quality of surface predictions obtained via iDLR is primarily influenced by two key factors. First is the degree of ill-posedness, that is, the extent to which information about the surface is preserved in the observed target image. Second is the size of the potential surface deformation space, which defines the solution space the model must consider. The level of underdetermination is strongly dependent on the heliostat design. In particular, larger heliostats with a greater number of canted or focused facets tend to exhibit a higher degree of underdetermination, potentially leading to lower prediction accuracy compared to the heliostats used at STJ. Since the model relies on a learned surface prior to make predictions under underdetermined conditions, a smaller solution space generally results in more accurate predictions. Conversely, heliostats that exhibit greater variability in surface errors due to manufacturing tolerances or deployment conditions may yield less accurate predictions. This issue could be especially pronounced when predicting high-spatial-frequency deformations, such as surface waviness, for heliostats positioned farther from the tower. In such cases, the impact of these fine-scale features is less apparent in the flux density, forcing the model to depend more heavily on its learned surface representation. Overall, it is likely that the interplay of the aforementioned factors, specifically, increased underdetermination arising from complex facet geometries and an expansive surface deformation space, can significantly impact the predictive performance

of iDLR, potentially leading to failure for certain heliostat models. A more comprehensive investigation is required to systematically assess how heliostat geometry influences the accuracy and reliability of iDLR surface predictions.

Compared to other existing approaches that aim to derive a heliostat digital twin from target images, our method provides a significant improvement in both generalizability and physical interpretability. The post-calibration method of Zhu et al. [10], which only optimizes canting and focusing errors, does not model surface deformations. Our findings indicate that such deformations are critical for accurate flux density predictions, particularly at the STJ site. While the differentiable raytracer introduced by Pargmann et al. [11] is theoretically capable of representing precisely all kind of heliostat surfaces, it suffers from the ill-posedness of the inverse problem and fails with the surface reconstruction under realistic conditions. The method proposed by Kuhl et al. [12] achieves flux density predictions with accuracies around 90% using a non-physical, purely data-driven model of the heliostat. While this is comparable to the accuracy achieved by iDLR, their predictions are restricted to the target plane, whereas iDLR can make predictions directly on the receiver, even for complex receiver geometries. According to Kuhl et al. [12], the spatial extrapolation to the receiver plane introduces an additional error of 2%–10%. Moreover, iDLR offers the flexibility to integrate with physical modeling frameworks: it enables simulation of soiling effects, inclusion of shading and blocking, incorporation of sun shape measurements, and other scene-specific phenomena within the raytracing environment. On the other hand, a clear strength of purely data-driven approaches lies in their independence from surface measurements and physical modeling, removing the need for surface measurements and raytracing, thereby reducing implementation complexity and required labor.

## 5. Conclusion

This work introduced the Zero-Shot Sim-to-Real transfer of inverse Deep Learning Raytracing (iDLR) to real-world data, enabling the prediction of heliostat surface shapes from target images of focal spots and supporting accurate flux density predictions under realistic conditions. Through a comprehensive evaluation across 63 heliostats, we demonstrated that iDLR surface predictions closely align with deflectometry ground truth for the vast majority of cases (84%), achieving a median MAE of 0.17 mm. Despite occasional mispredictions, an expected consequence of the underdetermined nature of the inverse problem, the resulting flux density prediction from raytracing remain highly accurate. On the position of the Lambertian target, the iDLR flux densities achieved a median prediction accuracy of 92% relative to the deflectometry flux densities. Even in a challenging double-extrapolation scenario, where flux density predictions on the receiver during autumn/spring were generated from summer target images, iDLR maintained an average accuracy of 90%, and a remarkable 97.5% agreement when flux densities were superposed across all 63 validation and testing heliostats. These results confirm the robustness and generalization capability of iDLR. Overall, iDLR enables cost-efficient, automated, and scalable surface reconstruction, making it highly suitable for integration into digital twins of heliostats. iDLR provides a valuable middle ground between physics-based and data-driven methods, offering both interpretability and compatibility with physical simulation tools such as raytracing. These capabilities position iDLR as a key enabler for next-generation heliostat monitoring, calibration, and control.

## CRediT authorship contribution statement

**Jan Lewen:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Max Pargmann:** Writing – review & editing, Project administration. **Mehdi Cherti:** Supervision. **Jenia Jitsev:** Supervision. **Robert Pitz-Paal:** Supervision, Funding acquisition. **Daniel Maldonado Quinto:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used large language models (LLM) (GPT4 of openAI and the LLMs by deepL) in order to improve the language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

We gratefully acknowledge the use of data from the Solar Tower Jülich, a research power plant operated by the German Aerospace Center (DLR). Additionally, we extend our thanks for the funding provided by the Helmholtz Association (HGF) for the GANCSTR project (Grant Number ZT-I-PF-5-069), under which this research was conducted. Furthermore, we thank the Helmholtz Association's Initiative and Networking Fund for funding computational resources on the HAICORE@FZJ partition.

## References

[1] F. Schöniger, R. Thonig, G. Resch, J. Lilliestam, Making the sun shine at night: comparing the cost of dispatchable concentrating solar power and photovoltaics with storage, Energy Sources Part B: Econ. Plan. Policy 16 (1) (2021) 55–74, http://dx.doi.org/10.1080/15567249.2020.1843565.

[2] R. Schäppi, D. Rutz, F. Dähler, A. Muroyama, P. Haueter, J. Lilliestam, A. Patt, P. Furler, A. Steinfeld, Drop-in fuels from sunlight and air, Nat. 601 (7891) (2021) 63–68, http://dx.doi.org/10.1038/s41586-021-04174-y.

[3] K.W. Stone, Automatic heliostat track alignment method, 1986, US Patent 4, 564, 275.

[4] F. Sun, M. Guo, Z. Wang, W. Liang, Z. Xu, Y. Yang, Q. Yu, Study on the heliostat tracking correction strategies based on an error-correction model, Sol. Energy 111 (2015) 252–263, http://dx.doi.org/10.1016/j.solener.2014.06.016.

[5] J.C. Sattler, M. Röger, P. Schwarzbözl, R. Buck, A. Macke, C. Raeder, J. Göttsche, Review of heliostat calibration and tracking control methods, Sol. Energy 207 (2020) 110–132.

[6] M.P. Sarr, A. Thiam, B. Dieng, Prediction of heliostat tracking errors using an adaptive neuro-fuzzy inference system, SSRN Electron. J. (2021) http://dx.doi.org/10.2139/ssrn.3953061.

[7] R. Döring, R. Dudek, S. Rzepka, L. Scheiter, E. Noack, B. Seiler, Design optimization by virtual prototyping using numerical simulation to ensure thermomechanical reliability in the assembly and interconnection of electronic assemblies, J. Electron. Packag. 145 (1) (2022) http://dx.doi.org/10.1115/1.4056445.

[8] M. Pargmann, M. Leibauer, V. Nettelroth, D. Maldonado Quinto, R. Pitz-Paal, Enhancing heliostat calibration on low data by fusing robotic rigid body kinematics with neural networks, Sol. Energy 264 (2023) 111962, http://dx.doi.org/10.1016/j.solener.2023.111962.

[9] J. Lewen, M. Pargmann, M. Cherti, J. Jitsev, R. Pitz-Paal, D.M. Quinto, Inverse deep learning raytracing for heliostat surface prediction, Sol. Energy 289 (2025) 113312, http://dx.doi.org/10.1016/j.solener.2025.113312.

[10] R. Zhu, D. Ni, T. Yang, J. Yang, J. Chen, G. Xiao, Heliostat field aiming strategy optimization with post-installation calibration, Appl. Therm. Eng. 202 (2022) 117720, http://dx.doi.org/10.1016/j.applthermaleng.2021.117720.

[11] M. Pargmann, J. Ebert, M. Götz, D. Maldonado Quinto, R. Pitz-Paal, S. Kesselheim, Automatic heliostat learning for in situ concentrating solar power plant metrology with differentiable ray tracing, Nat. Commun. 15 (1) (2024) http://dx.doi.org/10.1038/s41467-024-51019-z.

[12] M. Kuhl, M. Pargmann, M. Cherti, J. Jitsev, D. Maldonado Quinto, R. Pitz-Paal, Flux density distribution forecasting in concentrated solar tower plants: A data-driven approach, Sol. Energy 282 (2024) 112894, http://dx.doi.org/10.1016/j.solener.2024.112894.

[13] S. Ulmer, T. März, C. Prahl, W. Reinalter, B. Belhomme, Automated high resolution measurement of heliostat slope errors, Sol. Energy 85 (4) (2011) 681–687.

[14] B. Belhomme, R. Pitz-Paal, P. Schwarzbözl, S. Ulmer, A new fast ray tracing tool for high-precision simulation of heliostat fields, J. Sol. Energy Eng. 131 (3) (2009) http://dx.doi.org/10.1115/1.3139139.

[15] A. Bonanos, M. Faka, D. Abate, S. Hermon, M. Blanco, Heliostat surface shape characterization for accurate flux prediction, Renew. Energy 142 (2019) http://dx.doi.org/10.1016/j.renene.2019.04.051.

[16] R. Monterreal, R. Enrique, J. Fernández-Reche, An improved methodology for heliostat testing and evaluation at the Plataforma Solar de Almería, in: AIP Conference Proceedings, Vol. 1850, AIP Publishing, 2017.

[17] M.R. Shortis, G.H.G. Johnston, Photogrammetry: An available surface characterization tool for solar concentrators, part I: Measurements of surfaces, J. Sol. Energy Eng. 118 (3) (1996) 146–150, http://dx.doi.org/10.1115/1.2870886.

[18] K. Pottler, E. Lüpfert, G.H.G. Johnston, M.R. Shortis, Photogrammetry: A powerful tool for geometric analysis of solar concentrators and their components, J. Sol. Energy Eng. 127 (1) (2005) 94–101.

[19] M. Röger, C. Prahl, S. Ulmer, Heliostat shape and orientation by edge detection, 2010.

[20] A. Bonanos, M. Faka, D. Abate, S. Hermon, M. Blanco, Heliostat surface shape characterization for accurate flux prediction, Renew. Energy 142 (2019) 30–40.

[21] A. Martínez-Hernández, R. Conceição, C.-A. Asselineau, M. Romero, J. González-Aguilar, Advanced surface reconstruction method for solar reflective concentrators by flux mapping, Sol. Energy 266 (2023) 112162, http://dx.doi.org/10.1016/j.solener.2023.112162, URL: https://www.sciencedirect.com/science/article/pii/S0038092X2300796X.

[22] C.A. Arancibia-Bulnes, M.I. Peña-Cruz, A. Mutuberría, R. Díaz-Uribe, M. Sánchez-González, A survey of methods for the evaluation of reflective solar concentrator optics, Renew. Sustain. Energy Rev. 69 (2017) 673–684, http://dx.doi.org/10.1016/j.rser.2016.11.048, URL: https://www.sciencedirect.com/science/article/pii/S136403211630747X.

[23] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, P. Abbeel, Domain randomization for transferring deep neural networks from simulation to the real world, in: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2017, pp. 23–30, http://dx.doi.org/10.1109/iros.2017.8202133.

[24] J. Tobin, W. Zaremba, P. Abbeel, Domain randomization and generative models for robotic grasping, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2017, pp. 3482–3489, URL: https://api.semanticscholar.org/CorpusID:4530385.

[25] X.B. Peng, M. Andrychowicz, W. Zaremba, P. Abbeel, Sim-to-real transfer of robotic control with dynamics randomization, in: 2018 IEEE International Conference on Robotics and Automation, ICRA, 2018, pp. 3803–3810, http://dx.doi.org/10.1109/ICRA.2018.8460528.

[26] X. Yue, Y. Zhang, S. Zhao, A. Sangiovanni-Vincentelli, K. Keutzer, B. Gong, Domain randomization and pyramid consistency: Simulation-to-real generalization without accessing target domain data, in: 2019 IEEE/CVF International Conference on Computer Vision, ICCV, IEEE, 2019, pp. 2100–2110, http://dx.doi.org/10.1109/iccv.2019.00219.

[27] S. Pouyanfar, M. Saleem, N. George, S.-C. Chen, ROADS: Randomization for obstacle avoidance and driving in simulation, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPRW, 2019, pp. 1267–1276, http://dx.doi.org/10.1109/CVPRW.2019.00166.

[28] A. Loquercio, E. Kaufmann, R. Ranftl, A. Dosovitskiy, V. Koltun, D. Scaramuzza, Deep drone racing: From simulation to reality with domain randomization, IEEE Trans. Robot. 36 (1) (2020) 1–14, http://dx.doi.org/10.1109/TRO.2019.2942989.

[29] K.L. Voogd, J.P. Allamaa, J. Alonso-Mora, T.D. Son, Reinforcement learning from simulation to real world autonomous driving using digital twin, IFAC-Pap. 56 (2) (2023) 1510–1515, http://dx.doi.org/10.1016/j.ifacol.2023.10.1846, URL: https://www.sciencedirect.com/science/article/pii/S2405896323022553, 22nd IFAC World Congress.

[30] J. Hoffman, E. Tzeng, T. Park, J.-Y. Zhu, P. Isola, K. Saenko, A.A. Efros, T. Darrell, CyCADA: Cycle-consistent adversarial domain adaptation, 2017, arXiv:1711.03213, URL: https://arxiv.org/abs/1711.03213.

[31] E. Tzeng, J. Hoffman, K. Saenko, T. Darrell, Adversarial discriminative domain adaptation, 2017, arXiv:1702.05464, URL: https://arxiv.org/abs/1702.05464.

[32] A. Bewley, J. Rigley, Y. Liu, J. Hawke, R. Shen, V.-D. Lam, A. Kendall, Learning to drive from simulation without real world labels, 2018, http://dx.doi.org/10.48550/ARXIV.1812.03823, URL: https://arxiv.org/abs/1812.03823.

[33] W. Hong, Z. Wang, M. Yang, J. Yuan, Conditional generative adversarial network for structured domain adaptation, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2018, pp. 1335–1344, http://dx.doi.org/10.1109/cvpr.2018.00145.

[34] J. Tobin, L. Biewald, R. Duan, M. Andrychowicz, A. Handa, V. Kumar, B. McGrew, J. Schneider, P. Welinder, W. Zaremba, P. Abbeel, Domain randomization and generative models for robotic grasping, 2018, arXiv:1710.06425, URL: https://arxiv.org/abs/1710.06425.

[35] J. Zhang, L. Tai, P. Yun, Y. Xiong, M. Liu, J. Boedecker, W. Burgard, VR-goggles for robots: Real-to-sim domain adaptation for visual control, IEEE Robot. Autom. Lett. 4 (2) (2019) 1148–1155, http://dx.doi.org/10.1109/lra.2019.2894216.

[36] M. Kuhl, M. Pargmann, M. Cherti, J. Jitsev, D. Maldonado Quinto, R. Pitz-Paal, In-situ UNet-based heliostat beam characterization method for precise flux calculation using the camera-target method, Sol. Energy 279 (2024) 112811, http://dx.doi.org/10.1016/j.solener.2024.112811.

[37] K. Phipps, M. Kuhl, M. Weiel, M. Busch, J. Lewen, N. Blumenröhr, D.M. Quinto, C. Debus, F. Göhring, A. Streit, R. Pitz-Paal, M. Götz, M. Pargmann, PAINT database, 2025, URL: https://paint-database.org/, PID: https://hdl.handle.net/21.11152/474a4b1c-de93-4d4a-b33d-1d32d63baf4b.

[38] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby, An image is worth 16x16 words: Transformers for image recognition at scale, 2021, arXiv:2010.11929, URL: https://arxiv.org/abs/2010.11929.

[39] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, 2023, arXiv:1706.03762, URL: https://arxiv.org/abs/1706.03762.

[40] T. Karras, S. Laine, T. Aila, A style-based generator architecture for generative adversarial networks, 2019, arXiv:1812.04948.

[41] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, T. Aila, Analyzing and improving the image quality of StyleGAN, 2020, arXiv:1912.04958.

[42] E. Richardson, Y. Alaluf, O. Patashnik, Y. Nitzan, Y. Azar, S. Shapiro, D. Cohen-Or, Encoding in style: a StyleGAN encoder for image-to-image translation, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, IEEE Computer Society, Los Alamitos, CA, USA, 2021, pp. 2287–2296, http://dx.doi.org/10.1109/CVPR46437.2021.00232, URL: https://doi.ieeecomputersociety.org/10.1109/CVPR46437.2021.00232.

[43] D. Buie, A. Monger, C. Dey, Sunshape distributions for terrestrial solar simulations, Sol. Energy 74 (2) (2003) 113–122, http://dx.doi.org/10.1016/s0038-092x(03)00125-7.