



Accurate depth-resolved temperature profiling via thermal-radiation spectroscopy: numerical methods vs machine learning

DMITRII SHYMKIV,¹  ZHONGYUAN WANG,² BRIGHAM THORNOCK,^{1,3} AIDEN KARPFF,^{1,4} CAMILA NUNEZ,^{1,5} AND YUZHE XIAO^{1,*}

¹Department of Physics, University of North Texas, Denton, TX 76203, USA

²Department of Physics, The University of Texas at Austin, Austin, TX 78712, USA

³Erik Jonsson School of Engineering and Computer Science, The University of Texas at Dallas, Richardson, TX 75080, USA

⁴Physics and Astronomy Department, Pomona College, Claremont, CA 91711, USA

⁵Department of Computer Science, University of Virginia, Charlottesville, VA 22904, USA

*yuzhe.xiao@unt.edu

Abstract: We present and compare three approaches for accurately retrieving depth-resolved temperature distributions within materials from their thermal-radiation spectra, based on: (1) a nonlinear equation solver implemented in commercial software, (2) a custom-built nonlinear equation solver, and (3) a deep neural network (DNN) model. These methods are first validated using synthetic datasets comprising randomly generated temperature profiles and corresponding noisy thermal-radiation spectra for three different structures: a fused-silica substrate, an indium antimonide substrate, and a thin-film gallium nitride layer on a sapphire substrate. We then assess the performance of each approach using experimental spectra collected from a fused-silica window heated on a temperature-controlled stage. Our results demonstrate that the DNN-based method consistently outperforms conventional numerical techniques on both synthetic and experimental data, providing a robust solution for accurate depth-resolved temperature profiling.

© 2025 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Thermal radiation is a fundamental physical phenomenon in which heated objects emit electromagnetic waves as a form of energy. Recent advances in thermal-radiation engineering have enabled a wide range of innovative applications [1], including radiative cooling [2], thermophotovoltaics [3], advanced infrared light sources [4,5], and thermal camouflage [6]. Among the earliest and most widely adopted applications is infrared thermography, which utilizes infrared cameras to detect emitted radiation for non-contact temperature measurements. This technique is extensively used in medical imaging [7], defense [8], integrated circuit monitoring [9], and remote sensing [10]. However, conventional infrared thermography relies on the total radiated power to infer an effective surface temperature, limiting its ability to resolve subsurface temperature variations—an essential capability for applications such as thermal management in multilayer electronic devices [11] and three-dimensional (3D) medical diagnostics [7].

To address this limitation, we recently developed depth thermography, a non-invasive technique that enables temperature measurement as a function of depth [12]. Depth thermography utilizes spectrally resolved thermal radiation emitted from the semi-transparent spectral region of a material, where the emitted signal comprises contributions from various depths. By modeling the material as a multilayer structure with N layers, the total radiated power can be expressed as [13]:

$$I(\lambda) = \sum_{i=1}^N \bar{\epsilon}_i(\lambda) I_{BB}(\lambda, T_i), \quad (1)$$

where $\bar{\epsilon}_i(\lambda)$ is the local emissivity of the i^{th} layer, T_i is its temperature, and

$$I_{BB}(\lambda, T) = 2hc^2/\lambda^5(e^{hc/\lambda k_B T} - 1) \quad (2)$$

is the blackbody spectral radiance according to Planck's law [12]. Here h is Planck's constant, c is the speed of light, and k_B is Boltzmann's constant. Note that in Eq. (1), contributions from different layers are added incoherently, since thermal radiation fields from different layers are uncorrelated, while the local emissivity is calculated coherently to consider the finite temporal coherence of thermal-radiation field emitted from the same layer. Determining the temperature profile T_i involves inverting Eq. (1)—an inherently ill-conditioned problem. This ill-conditioned nature significantly limits the accuracy and depth resolution of the reconstructed temperature distribution. In our previous work, we employed a brute-force search method and achieved a depth resolution of only 0.25 mm in a 1 mm-thick fused silica window [12]. Because experimental spectra inevitably contain noise, a robust inversion technique is essential to improve the accuracy and reliability of depth thermography.

An ill-conditioned problem is a mathematical or computational problem that does not meet the criteria of a well-conditioned problem, as defined by Jacques Hadamard [14]. According to Hadamard, a well-conditioned problem must satisfy three conditions: a solution must exist, be unique, and depend continuously on the input (i.e., be stable). In the context of temperature inversion from thermal-radiation spectra, even small noise in the measured spectra can lead to large deviations in the reconstructed temperature profile—an indication of solution instability. Furthermore, inverse problems are often inherently ambiguous [15], as multiple temperature distributions may produce spectral outputs that closely match noisy observational data [12]. Throughout the 20th century, numerous numerical and theoretical techniques were developed to mitigate the challenges of ill-conditioned problems, including methods such as regularization and minimization [15–19].

More recently, advances in machine learning — particularly deep learning and neural networks — have introduced powerful new approaches for solving inverse problems [20–23]. These techniques have demonstrated significant potential across a wide range of disciplines, including optical metrology [24], spectroscopy [25], geophysics [26] and solid Earth geoscience [27]. In the context of temperature inversion, machine learning models have been successfully employed to retrieve temperature and humidity profiles from satellite observations [28] and ground-based microwave radiometer measurements [29] in remote sensing applications. Additionally, neural networks have been used to infer temperature and concentration distributions in combustion systems based on infrared emission spectra [30].

In this paper, we present and evaluate three methods for accurate temperature inversion from thermal-radiation spectra: (1) a nonlinear equation solver based on MATLAB's *lsqnonlin* function, (2) an inversion algorithm incorporating generalized cross-validation (GCV) [15], and (3) a deep neural network (DNN)-based approach. To assess the effectiveness of these methods, we first apply them to numerically generated datasets across three different structures. We then validate the approaches using experimental data obtained from a 1 mm-thick fused-silica window placed on a temperature-controlled heater. Our results show that the DNN-based method consistently delivers the highest accuracy, enabling robust depth-resolved temperature profiling through thermal-radiation spectroscopy.

2. Methods

Method 1 (hereafter referred to as MATLAB) utilizes *lsqnonlin*, a built-in MATLAB function designed for solving nonlinear least-squares problems [31]. This function minimizes the sum of squares of a set of nonlinear functions, making it suitable for curve fitting, parameter estimation, and complex optimization tasks. In the context of depth-resolved temperature profiling, *lsqnonlin*

is used to find a set of temperatures (T_i), that minimizes the objective function

$$\sum_{\lambda} \left[\text{Spectrum}(\lambda) - \sum_{i=1}^N \bar{\epsilon}_i(\lambda) I_{BB}(\lambda, T_i) \right]^2, \quad (3)$$

where $\text{Spectrum}(\lambda)$ is the input thermal-radiation spectrum provided to the solver. Also, *lsqnonlin* requires an initial guess as well as upper and lower bounds for temperature values. While this method can recover the exact temperature profile when the input spectrum is noise-free, its performance significantly deteriorates in the presence of noise, as demonstrated in Section 3.

In method 1, the use of built-in MATLAB's *lsqnonlin* function limits our flexibility to customize the solver. Therefore, in method 2 (hereafter referred to as GCV), we developed a custom nonlinear equation solver based on regularization, an iterative approach originally formulated and tested for magnetotelluric inverse problems [15]. In these problems, depth-resolved profiles of the Earth's electrical conductivity are inferred from magnetotelluric data. Since the GCV algorithm has been successfully applied to a similar inversion problem, we believe it—being one of the most recent advances in inversion techniques for nonlinear ill-conditioned problems—is better suited than the standard approaches described in textbooks (e.g., [16–18]). Following the notation in Ref. [15], the nonlinear ill-conditioned inverse problem can be expressed as

$$F(m) + \varepsilon = b, \quad (4)$$

where b represents the noisy data (i.e., the measured spectrum), ε denotes a measurement noise, F is a nonlinear operator and m is the solution of the inverse problem (i.e., the temperature distribution). The solution m is obtained by minimizing the Tikhonov functional

$$\varphi = \|F(m) - b\|^2 + \beta \|W(m - m_{ref})\|^2, \quad (5)$$

where m_{ref} is a reference solution and β is the regularization parameter. This minimization is carried out iteratively using the damped Gauss–Newton method, starting from the initial guess $m = m_{ref}$. The regularization parameter β is selected by minimizing the generalized cross validation (GCV) value:

$$GCV(\beta) = \frac{\|r(\beta) - r\|^2}{\text{trace}(I - C(\beta))} \quad (6)$$

Further details of this algorithm can be found in Ref. [15]. For direct comparison with Method 1, we set m_{ref} to be the same as the initial guess used in the MATLAB approach. Additionally, we impose constraints on the maximum and minimum allowable temperature values to mimic the upper and lower bounds applied in *lsqnonlin*.

For method 3 (hereafter referred to as DNN), we employ a conventional feedforward DNN model, with the input layer receiving the thermal-radiation spectrum, several hidden layers, and the output layer producing the temperature distribution. Unlike the MATLAB and GCV methods, which solve the inverse problem directly, the DNN must be trained to predict temperature profiles from given spectra. The training dataset is generated using Eq. (1), incorporating local emissivity and randomly generated temperature distributions within a certain range. The same local emissivity and temperature range are used in the MATLAB and GCV methods to calculate spectra and define upper and lower bounds of temperature.

To ensure a consistent comparison, all DNN models in this study share identical architectures and hyperparameters—except for the sizes of the input and output layers, which vary according to the problem. Each model consists of 20 hidden layers with 250 nodes per layer, employing the hyperbolic tangent (*tanh*) activation function. The datasets are split into training (60%), validation (20%), and test (20%) subsets, and normalized using mean normalization. Training is performed over 200 epochs using stochastic gradient descent, minimizing the mean squared error loss function.

3. Results

3.1. Numerical data

We first evaluate the performance of the three temperature-inversion methods using numerically generated data for three distinct structures: (1) a 1 mm thick fused-silica substrate on a heater stage, (2) a 0.5 mm thick indium antimonide (InSb) substrate on a heater stage, and (3) a 5 μm thick gallium nitride (GaN) thin film on a sapphire substrate (Fig. 1). To demonstrate depth-resolved temperature inversion, the fused-silica structure is modeled as an 11-layer structure consisting of ten 0.1 mm thick silica layers atop a 200 nm thick aluminum layer. The InSb structure is modeled as a 21-layer system with twenty 0.025 mm thick InSb layers over a 200 nm aluminum layer. The GaN structure comprises 20 layers: ten 500 nm thick GaN layers, followed by five 1 μm thick sapphire layers and five 20 μm thick sapphire layers. There is no heater stage in the GaN structure. As a widely used semiconductor material in electronic devices, non-uniform temperature profiles can be generated in GaN by optical pumping or electrical excitations. We are mostly interested in the temperature distribution inside GaN as well as near the GaN/sapphire interface. Therefore, we intentionally introduce this non-uniform layer segmentation in the sapphire substrate. Depth thermography relies on thermal-radiation spectra in the semi-transparent region; accordingly, we focus on the wavelength ranges of 4.75–8 μm for fused silica, 5–10 μm for InSb, and 8–15 μm for GaN, with corresponding spectral resolutions of 50 nm, 10 nm, and 10 nm, respectively. The geometric and spectral parameters for these three test structures are summarized in Table 1.

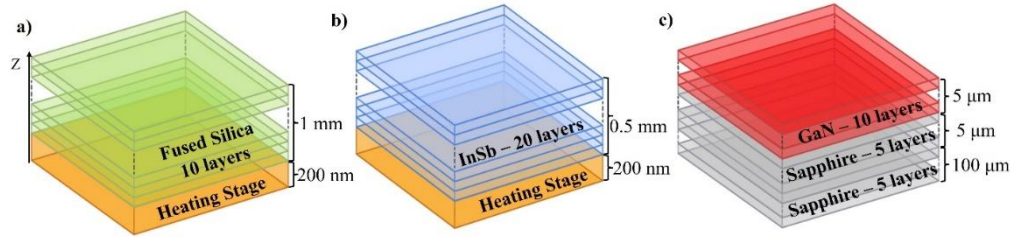


Fig. 1. Three testing structures used to validate inversion methods via numerical data: a) 1 mm thick fused-silica substrate on a heater stage; b) 0.5 mm thick InSb substrate on a heater stage; c) 5 μm thick GaN film on a sapphire substrate.

The optical properties (i.e., n and k) of InSb, sapphire, and GaN are obtained from Ref. [32,33], and [34], while those for fused silica are from our own ellipsometry measurement [12]. For each structure, the local emissivity of individual layers is calculated using the scattering matrix method and the dyadic Green's function [35]. Thermal-radiation spectra are then generated using Eq. (1), where temperature of each layer T_i is randomly assigned within the range of 273 to 373 K. Some examples of temperature profiles and the corresponding thermal-radiation spectra for each structure are provided in Supplement 1, Section 1. To simulate realistic experimental conditions, white Gaussian noise is added to the ideal spectra, with signal-to-noise ratios (SNR) set at 35, 45, and 50 dB to evaluate method performance under varying noise levels. For statistically meaningful analysis, each method is tested with $N = 100$ different spectra and corresponding random temperature distributions. Method performance is assessed by the standard deviation (SD) for each layer across all test cases:

$$SD_j = \sqrt{\frac{1}{N} \sum_{i=1}^N (1 - T_{ij}^{inv}/T_{ij})^2} \times 100\%, \quad (7)$$

where T_{ij}^{inv} and T_{ij} represent the inverted and input temperatures for layer j in the i^{th} random temperature distribution.

Table 1. Geometrical and spectral parameters of three testing structures.

Structure	Fused silica	InSb	GaN
Material (bottom)	Aluminum	Aluminum	Sapphire
Number of layers	1	1	5+5
Total thickness	200 nm	200 nm	5+100 μm
Material (top)	Fused silica	InSb	GaN
Number of layers	10	20	10
Total thickness	1 mm	0.5 mm	5 μm
Spectral range	4.75 – 8 μm	5 – 10 μm	8 – 15 μm
Spectral resolution	50 nm	10 nm	10 nm

Since the random temperature distributions are generated within the range of 273 to 373 K, these values are used as the lower and upper bounds for the numerical inversion methods (i.e., MATLAB and GCV). For DNN approach, 10,000 randomly generated temperature distributions within this range are used to train a model for each structure. For both the MATLAB and GCV method, a uniform temperature distribution equal to the top surface temperature is used as the initial guess and reference temperature, respectively. This choice ensures a fair comparison between the MATLAB and GCV methods. In practice, the surface temperature can be estimated by fitting the thermal-radiation spectrum in the opaque spectral region with a blackbody radiation curve [12].

The performance of the three temperature inversion methods is summarized in Fig. 2. As expected, the standard deviation increases with noise power across all structures and methods. One example of temperature inversion for fused-silica structure is shown in Supplement 1, Section 2. For the fused-silica and InSb structures, the inversion error generally decreases with increasing layer number—that is, temperature estimates are less accurate for layers closer to the heater stage and more accurate for those near the top surface. This trend arises because thermal radiation from the deeper layers undergoes greater attenuation before escaping from the surface. In terms of radiative properties, layers near the bottom have smaller local emissivity values compared to top layers, resulting in smaller contributions to the total radiative power. Note that the results from the GCV method do not strictly follow this trend, exhibiting high SD around layers 6-8 for fused silica and layers 8-18 for InSb. We attribute this to the regularization term in Eq. (5) for GCV. In GCV, a uniform top-surface temperature is used as the reference. Randomly generated profiles can have temperatures of internal layers (e.g., layers 6–8 in silica, 8–18 in InSb) that differ substantially from the surface value, which the regularization term penalizes. Since these layers also contribute strongly to the thermal spectrum, the interplay of these effects and the reference choice produces the non-monotonic SD behavior for GCV, especially at high noise levels. In contrast, the MATLAB and DNN methods, lacking such a regularization term, do not show this pattern.

The GaN structure exhibits a more intriguing pattern of layer-dependent temperature inversion error (Fig. 2(f, i, l)). In addition to the near-exact retrieval of the top layer's temperature, two distinct dips in the error curves appear at layers 5 and 10. These dips, consistently observed across all inversion methods and noise levels, correspond to the interfaces within the non-uniformly

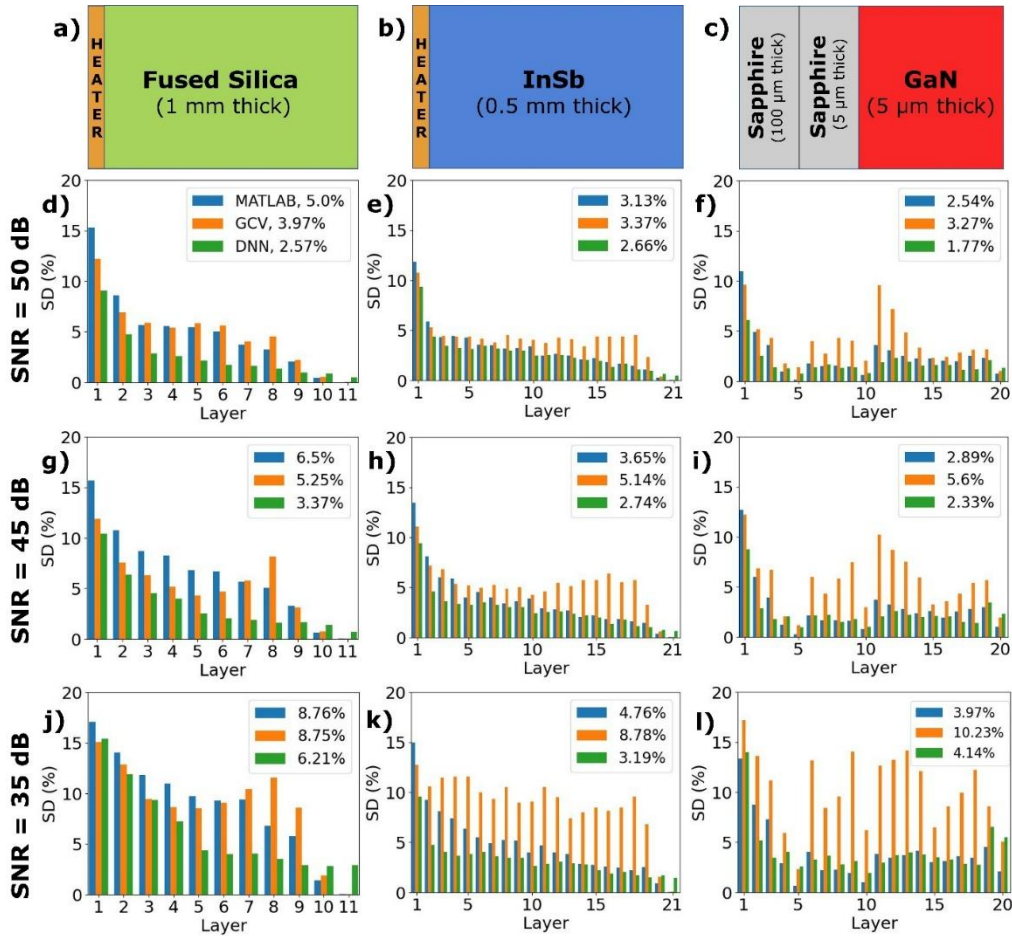


Fig. 2. (a–c) Material distribution in three testing structures. Note that the layer thicknesses shown in these diagrams are not to scale. (d–l) Comparison of temperature inversion errors at different depths across three inversion methods with varying noise levels: (d–f) SNR = 50 dB; (g–i) SNR = 45 dB; (j–l) SNR = 35 dB. The legend indicates the layer-averaged standard deviation (SD).

segmented sapphire substrate. They arise from the relatively high local emissivity values of these layers. The dip at layer 10 results from both the contrast in optical properties between GaN and sapphire and the change in layer thickness—specifically, the GaN layer is 500 nm thick, whereas the adjacent sapphire layer is 1 μm thick. Moreover, the first five sapphire layers are significantly thicker than layers 6–10 (20 μm vs. 1 μm), which also contributes to their higher local emissivity and, consequently, lower inversion error. As a result, unlike the fused silica and InSb structures, the GaN structure does not exhibit a clear trend of decreasing inversion error with increasing layer number beyond the top five layers.

For the fused-silica structure, the performance of the MATLAB and GCV methods is comparable across all noise levels, with standard deviations ranging from 4% to 9%. For the InSb and GaN structures, the MATLAB method yields lower average temperature inversion errors than the GCV method, particularly at SNR = 35 dB. Among all methods, the DNN models consistently achieve the best performance across all materials.

The performance of the DNN models can be further enhanced by incorporating noisy data and/or increasing the size of the training dataset. For example, in Fig. 3 we compare models trained with spectra containing different levels of noise. When the test set spectra exhibit noise levels similar to those used during training, prediction errors decrease relative to models trained on noise-free spectra. This improvement is especially evident under high noise conditions. Specifically, when tested on spectra with SNR = 35 dB, the model trained on SNR35 noisy data achieves more than twice the accuracy of the model trained on noise-free spectra. Furthermore, testing on noisy data with varying SNRs in Fig. 3 clearly shows that the layer-averaged error is minimized when the noise levels of the training and testing data match. These findings are consistent with previous studies on neural network performance under noisy and noise-free training conditions [20].

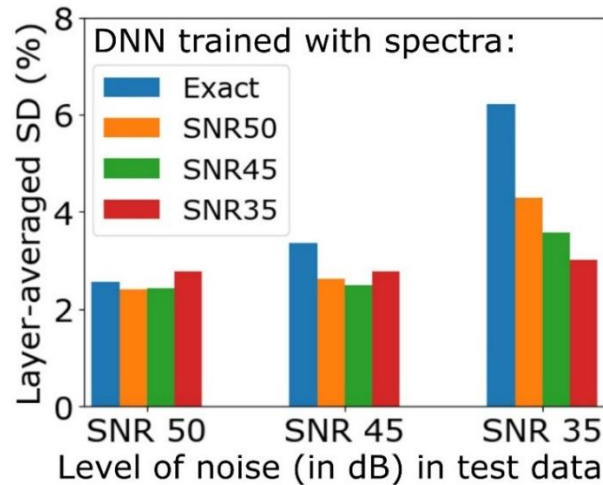


Fig. 3. Comparison of layer-averaged temperature inversion errors for the fused-silica structure shown in Fig. 2(a) using DNN models trained with noisy and exact spectra.

3.2. Experimental data

To further evaluate the performance of the three inversion methods, we apply them to experimental thermal-radiation spectra. These spectra were collected from a 1-mm-thick fused-silica window placed on a heater stage maintained at either 473 or 573 K, after steady-state thermal conditions had been established. A detailed description of the experiment is provided in Ref. [12]. When a bulk, isotropic, and homogeneous material reaches thermal steady state on a heated surface, its temperature is expected to increase linearly with depth. For the fused-silica window, we found the surface temperatures to be 468 and 556 K for heater stage temperatures of 473 and 573 K, respectively. These surface temperatures were measured using both an IR camera (since fused silica is opaque in the camera's operating spectral range of 7.5 to 13 μm) and spectral fitting within the opaque wavelength range, as described in Ref. [12].

The spectra were recorded with a resolution of 2 cm^{-1} , yielding 445 data points across the wavelength range of 4.75 to 8 μm . To further assess the performance of the inversion techniques, we investigate their accuracy at different depth resolutions — that is, how the fidelity of the reconstructed temperature profiles depends on the number of layers used in the model. All inversion methods are tested with three different layer counts: 6, 11, and 21, corresponding to depth resolutions of 0.2, 0.1, and 0.05 mm in the fused-silica substrate, respectively.

A uniform temperature distribution equal to the surface temperature is used as the initial guess for the MATLAB method and as the reference temperature for the GCV method. Upper and lower temperature bounds are set to ± 50 K from the surface temperature to provide a reasonable estimated temperature range. For the DNN method, training datasets are generated using 10,000 random temperature profiles within the range of 506–606 K (or 418–518 K) for the 573 K (or 473 K) heater temperature case, across each depth resolution.

The depth-resolved temperature distributions retrieved from the experimental spectra of the fused-silica window are shown in Fig. 4. Among the three methods, the DNN approach yields the lowest temperature inversion errors across all tested scenarios. The MATLAB method outperforms the GCV approach. The errors associated with the MATLAB method remain consistent across all depth resolutions and both temperature settings, with an average error of about 3.8%. This is slightly better than the 5.0% error observed using numerical spectra (see Fig. 2(d)). We reference Fig. 2(d) here because the experimental spectra have a SNR close to 50 dB (see more discussions in Supplement 1, Section 3).

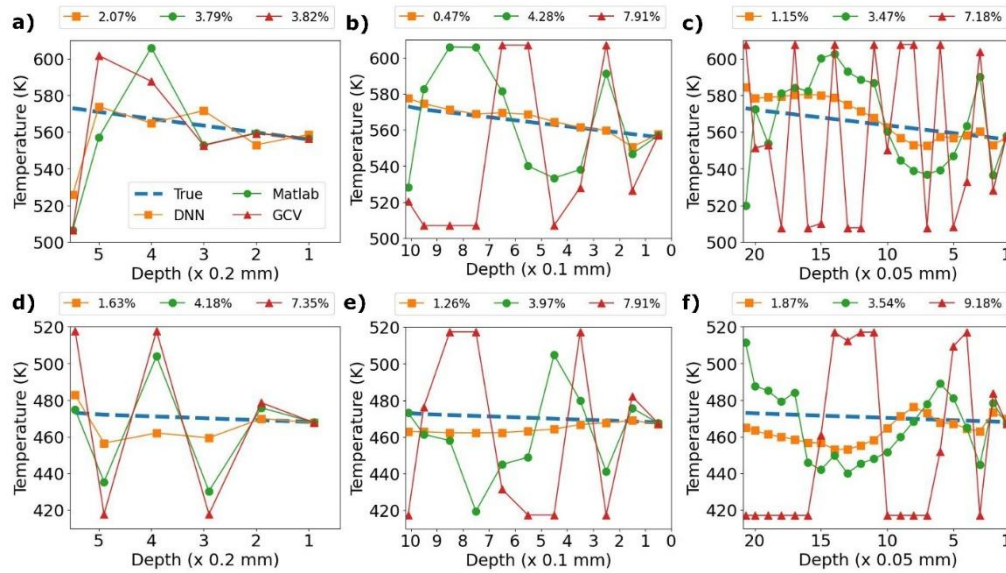


Fig. 4. Comparison of inverted temperature profiles inside a 1 mm thick fused-silica window on a heater stage at either 573 K (a-c) or 473 K (d-f) across three methods with varying depth resolution: (a, d) 0.2 mm; (b, e) 0.1 mm; (c, f) 0.05 mm. The legend indicates the layer-averaged standard deviation (SD).

In contrast, the performance of the GCV method on the experimental spectra varies significantly with depth resolution. Its best performance occurs at a heater temperature of 573 K with a depth resolution of 0.2 mm, yielding an error of about 3.8%. However, in other cases, the inverted temperature profiles deviate substantially from the ground truth, with errors exceeding 7%. We attribute the subpar performance of the GCV method in this context to the limited dataset—only two experimental spectra were tested—which may lead to insufficient statistical representation.

The minimum temperature inversion error achieved is 0.47% using DNN model for the 573 K case with a depth resolution of 0.1 mm (Fig. 4(b)). Notably, this error is approximately five times lower than the 2.57% error obtained using numerical data (see Fig. 2(d)). Two main factors contribute to this discrepancy. First, the experimental spectra possess a significantly higher spectral resolution—approximately seven times greater—than that of the numerical spectra. As a result, the DNN model can predict the fixed-size output with improved accuracy. We believe this

also partially leads to the improvement of the MATLAB method performance on experimental spectra versus numerical spectra. Second, the two error values are not directly comparable. The 2.57% error represents an average over 100 different test spectra, providing a statistically robust estimate, whereas the 0.47% error corresponds to a single instance of temperature inversion and lacks statistical generality.

It is evident that the DNN models with a depth resolution of 0.1 mm achieve the highest accuracy compared to models with resolutions of 0.2 and 0.05 mm across both temperature settings. While we do not claim that 0.1 mm is the optimal depth resolution for temperature inversion, the observed non-monotonic relationship between accuracy and resolution has computational origins. Specifically, the architecture and hyperparameters of all DNN models used for analyzing the experimental spectra were adopted from the model trained for numerical spectra with a 0.1 mm depth resolution. We suspect this contributes to the better performance observed at this resolution. Furthermore, from a numerical perspective, a model typically performs better when the output dimension is substantially smaller than the input. This consideration supports the expectation that the model would perform better at a 0.1 mm resolution than at a finer resolution of 0.05 mm.

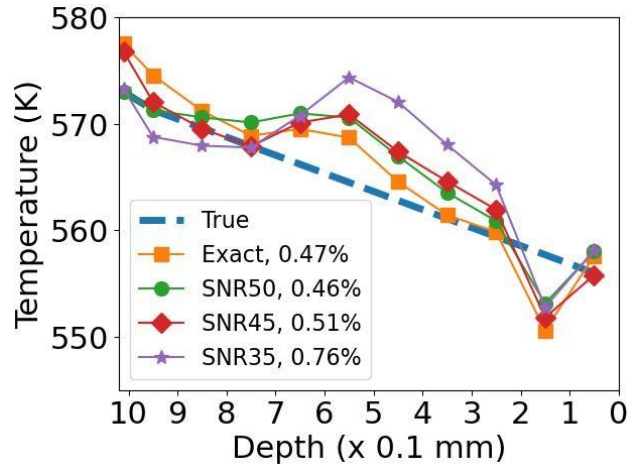


Fig. 5. Temperature inversion results from the experimental spectrum of a fused-silica window placed on a heater stage at 573 K, using DNN models trained with exact spectra and noisy spectra at varying SNRs. The legend indicates the layer-averaged standard deviation (SD).

As discussed in Sec. III.A (Fig. 3), the accuracy of the DNN models can be further enhanced by training it with noisy data. To explore this, we train the same DNN model—designed for a depth resolution of 0.1 mm and a heater temperature of 573 K—using synthetic spectra with varying SNRs. We then evaluate the performance of each trained model on the experimental spectrum. The results, presented in Fig. 5, show that the model trained with noisy spectra at an SNR of 50 dB performs best. Notably, our experimental spectra have an SNR of approximately 50 dB (see Supplement 1, Section 3). These findings are consistent with the results shown in Fig. 3 for the numerical spectra tested at SNR = 50 dB. Overall, we conclude that training the DNN with spectra having a noise level similar to that of the test data leads to improved inversion accuracy for both numerical and experimental datasets.

To further test the robustness of the three temperature inversion methods across different structures, we apply them to a recently measured thermal-radiation spectrum from a 3 mm thick fused-silica window on a heater stage at 473 K. The measured thermal-radiation spectrum is

shown in Fig. 6(a). The increased thickness of the window produces a larger temperature gradient between the bottom and top surfaces. To resolve the temperature profile for a much thicker sample, shorter wavelengths with larger penetration depths are required. Therefore, we select the semi-transparent spectral ranges of 3.7–4.2 μm and 4.32–4.74 μm . The intermediate range of 4.2–4.32 μm is excluded due to the impact of CO_2 absorption in the atmosphere. Additionally, the opaque wavelength range of 10–10.5 μm is included to accurately determine the surface temperature. In total, 292 wavelengths are used for the temperature inversion.

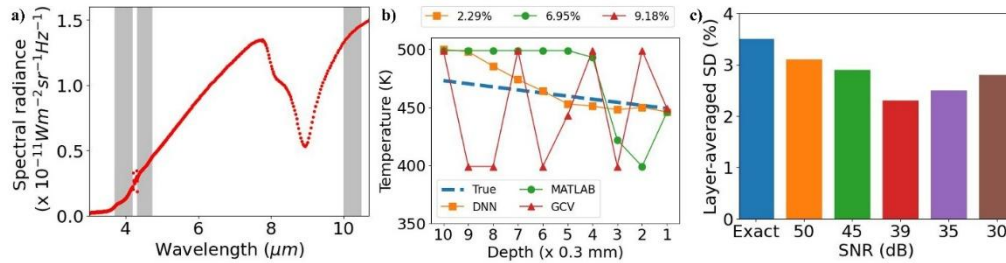


Fig. 6. (a) Measured thermal-radiation spectrum from a 3 mm thick fused-silica window on a heater stage at 473 K. Shaded gray areas indicate the wavelength ranges used for temperature inversion. (b) Comparison of inverted temperature profiles obtained with three methods. The DNN model was trained using noisy spectra with an SNR of 39 dB equal to our experimental data. The legend shows the layer-averaged standard deviation (SD). (c) Layer-averaged SD for DNN models trained with exact and noisy spectra with different SNRs. The minimum error occurs for SNR = 39 dB.

Using an IR camera, we found the surface temperature of the 3 mm fused-silica window is 449 K. Therefore, a uniform temperature distribution equal to 449 K is used as the initial guess for the MATLAB method and the reference temperature for the GCV method. Upper and lower bounds are set to ± 50 K from the surface temperature to define a reasonable range. For the DNN method, training datasets are generated from 10,000 random temperature profiles within the range of 399–499 K.

The depth-resolved temperature distributions retrieved from the experimental spectrum of the 3 mm thick fused-silica window are shown in Fig. 6(b). Among the three methods, the DNN approach achieves the lowest temperature inversion error. Furthermore, the layer-averaged SD for the DNN models varies with the noise level in the training spectra (Fig. 6(c)). We estimate the measured spectrum's SNR to be 39 dB. So, the DNN model trained with spectra at SNR = 39 dB yields a 2.29% error, which is nearly 40% lower than the DNN model trained with exact (noise-free) spectra (layer-averaged SD = 3.79%). In contrast, the MATLAB and GCV methods can accurately recover only the surface temperature, resulting in high layer-averaged SD values. Overall, these results further demonstrate that a DNN model trained with noisy spectra can produce much more accurate depth-resolved temperature profiles than conventional numerical techniques.

4. Discussion and conclusion

Inverting depth-dependent temperature from thermal radiation is not a new concept. For instance, earlier studies have used multifrequency microwave radiometry to retrieve vertical profiles of temperature, water vapor, and cloud liquid water in the atmosphere [36], as well as internal temperature distributions in biological objects [37–40]. Conceptually, microwave atmospheric remote sensing and infrared depth thermography are similar, as both rely on measured thermal-radiation spectra within a semi-transparent spectral range. The primary difference lies in the operating spectral range. A frequency window around 60 GHz ($\lambda \sim 5 \text{ mm}$) is typically

used to retrieve vertical temperature profiles in the atmosphere [36], while for infrared depth thermography we focus on the mid-infrared. The different working wavelength therefore set different diffraction-limited spatial resolution: for microwave-based techniques, the spatial resolution is about 5 mm while a much better spatial resolution about 5 μm can be obtained using infrared depth thermography. In 1998, F. Solheim *et al.* applied both neural network and conventional numerical methods for temperature inversion and found no significant advantage of the neural network approach over traditional techniques [36]. Their conclusion differs from our findings in this work. We attribute this discrepancy primarily to the limited computational resources available over two decades ago. Specifically, the 1998 study employed shallow neural networks with only one hidden layer, whereas our work utilizes a deep neural network architecture with 20 hidden layers.

Our results demonstrate that accurate depth-resolved temperature profiling is achievable across a wide range of materials — including both dielectrics and semiconductors—and over various depth scales (from millimeters to micrometers) and depth resolutions (from sub-millimeters to hundreds of nanometers). The key requirement is the presence of a spectral window in which the structure remains semi-transparent throughout its depth. Given the broadband nature of thermal radiation, such a window is usually easy to identify in practice. However, if the structure includes high-loss materials (e.g., metals), its layer thickness cannot be too large. To obtain temperature information within or beneath a high-loss layer, its thickness must not significantly exceed its penetration depth. For example, the penetration depth of gold at a wavelength of 5 μm is approximately 20 nm. The depth resolution is not constrained by the diffraction limit and can be in the nanometer range. In practice, depth resolution is determined both by the number of spectral data points within the semi-transparent range and the total thickness of the structure. Even when this window is relatively narrow, we believe that, with sufficiently high spectral resolution, temperature inversion with high depth resolution across the entire structure remains achievable. On the other hand, depth resolution must have a practical limit. Each layer should be sufficiently thick to allow assignment of well-defined real and imaginary parts of the refractive index. This is necessary to calculate the correct local emissivity, which is essential for accurate temperature inversion. Further investigations of structures with ultra-thin layers, such as 2D materials, and/or material interfaces, would be valuable for understanding the ultimate limits of depth resolution for temperature inversion.

In this study, we use randomly generated temperature distributions to construct the training datasets. However, the training profiles can be customized to fit specific applications. For example, if the temperature is expected to vary monotonically with depth — either increasing or decreasing — the training set can be restricted to include only such profiles. While this limits the model's generality, it can enhance accuracy by incorporating additional physical constraints. This approach has been demonstrated in previous studies [30,41], as well as in one of our recent work [42]. Additionally, to ensure a fair comparison, all DNN models in this study are implemented using identical architectures and hyperparameters (except for differences in input/output dimensions where necessary). Further fine-tuning of these parameters could potentially reduce prediction errors for specific material structures or particular combinations of spectral and depth resolutions (see [Supplement 1](#), Section 4).

In conclusion, we develop and compare three methods for inverting depth-resolved temperature distributions via thermal-radiation spectroscopy. Among them, the DNN approach consistently demonstrates the highest accuracy in temperature profiling, outperforming conventional numerical techniques including the GCV and MATLAB methods. Notably, training a DNN model requires less than 10 minutes on standard hardware (see [Supplement 1](#), Section 5 for a discussion of computational time). Once trained, the model can be applied almost instantaneously to any thermal-radiation spectrum from the same structure. In addition to its superior performance, the DNN approach offers greater flexibility—it can be further enhanced by training with noisy

data, utilizing higher-resolution spectra (an option also available for the other two methods), or expanding the size of the training dataset. These advantages position the DNN method as the most promising approach for temperature inversion from thermal-radiation spectra.

Funding. Defense Advanced Research Projects Agency (D23AP0018900); National Science Foundation (2418002, 2244259).

Acknowledgements. D. S. and Y. X. acknowledge the support from DARPA YFA program (Grant No. D23AP0018900) and NSF LEAPS-MPS program (Award Number 2418002). B. T., C. N., and A. K. acknowledge the support from an NSF REU program hosted at UNT (Award Number 2244259).

Disclosures. The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are available in Ref. [43].

Supplemental document. See [Supplement 1](#) for supporting content.

References

1. D. G. Baranov, Y. Xiao, I. A. Nechepurenko, *et al.*, “Nanophotonic engineering of far-field thermal emitters,” *Nat. Mater.* **18**(9), 920–930 (2019).
2. J. Liang, J. Wu, J. Guo, *et al.*, “Radiative cooling for passive thermal management towards sustainable carbon neutrality,” *Nat. Sci. Rev.* **10**(1), 1 (2023).
3. A. LaPotin, K. L. Schulte, M. A. Steiner, *et al.*, “Thermophotovoltaic efficiency of 40%,” *Nature* **604**(7905), 287–291 (2022).
4. M. Hlavatsch and B. Mizaikoff, “Advanced mid-infrared lightsources above and beyond lasers and their analytical utility,” *Analytical Sciences* **38**(9), 1125–1139 (2022).
5. G. Chen, Y. Cao, Y. Tang, *et al.*, “Advanced Near-Infrared Light for Monitoring and Modulating the Spatiotemporal Dynamics of Cell Functions in Living Systems,” *Adv. Sci.* **7**(8), 1903783 (2020).
6. B. Qin, H. Zhu, R. Zhu, *et al.*, “Space-to-ground infrared camouflage with radiative heat dissipation,” *Light: Sci. Appl.* **14**(1), 137 (2025).
7. B. B. Lahiri, S. Bagavathiappan, T. Jayakumar, *et al.*, “Medical applications of infrared thermography: A review,” *Infrared Phys. Technol.* **55**(4), 221–235 (2012).
8. A. Akula, R. Ghosh, and H. K. Sardana, “Thermal Imaging And Its Application In Defence Systems,” *AIP Conf Proc* **1391**, 333 (2011).
9. W. Liu and B. Yang, “Thermography techniques for integrated circuits and semiconductor devices,” *Sens. Rev.* **27**(4), 298–309 (2007).
10. J. A. Voogt and T. R. Oke, “Thermal remote sensing of urban climates,” *Remote Sens. Environ.* **86**(3), 370–384 (2003).
11. A. Jain, R. E. Jones, R. Chatterjee, *et al.*, “Thermal modeling and design of 3D integrated circuits,” *2008 11th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, I-THERM*, 1139 (2008).
12. Y. Xiao, C. Wan, A. Shahsafi, *et al.*, “Depth Thermography: Noninvasive 3D Temperature Profiling Using Infrared Thermal Emission,” *ACS Photonics* **7**(4), 853–860 (2020).
13. L. P. Wang, S. Basu, and Z. M. Zhang, “Direct and indirect methods for calculating thermal emission from layered structures with nonuniform temperatures,” *J. Heat Transfer* **133**(7), 7 (2011).
14. P. M. Jacques Hadamard, “Sur les problèmes aux dérivées partielles et leur signification physique,” *Princet. Univ. Bullet* **13**, 49 (1902).
15. E. Haber and D. Oldenburg, “A GCV based method for nonlinear ill-posed problems,” *Computational Geosciences* **4**(1), 41–63 (2000).
16. A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems* (Wiley, New York, 1977).
17. A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, *et al.*, *Numerical Methods for the Solution of Ill-Posed Problems* (Springer Science+Business Media Dordrecht, 1995).
18. A. Tarantola, *Inverse Problem Theory* (Elsevier Science, Amsterdam, 1987).
19. A. Neumaier, “Solving ill-conditioned and singular linear systems: A tutorial on regularization,” *SIAM Rev.* **40**(3), 636–666 (1998).
20. M. Genzel, J. Macdonald, and M. Marz, “Solving Inverse Problems With Deep Neural Networks – Robustness Included?” *IEEE Trans. Pattern Anal. Mach. Intell.* **45**(1), 1119–1134 (2023).
21. Y. Bai, W. Chen, J. Chen, *et al.*, “Deep learning methods for solving linear inverse problems: Research directions and paradigms,” *Signal Processing* **177**, 107729 (2020).
22. D. Bianchi, J. Schwab, S. Antholzer, *et al.*, “NETT: solving inverse problems with deep neural networks,” *Inverse Problems* **36**(6), 065005 (2020).
23. S. Kamyab, Z. Azimifar, R. Sabzi, *et al.*, “Deep learning methods for inverse problems,” *PeerJ Comput. Sci.* **8**, e951e951 (2022).
24. C. Zuo, J. Qian, S. Feng, *et al.*, “Deep learning in optical metrology: a review,” *Light: Sci. Appl.* **11**(1), 39 (2022).

25. H. Park, J. H. Park, and J. Hwang, "An inversion problem for optical spectrum data via physics-guided machine learning," *Sci. Rep.* **14**(1), 9042 (2024).
26. S. Yu and J. Ma, "Deep Learning for Geophysics: Current and Future Trends," *Rev. Geophys.* **59**(3), 3 (2021).
27. K. J. Bergen, P. A. Johnson, M. V. De Hoop, *et al.*, "Machine learning for data-driven discovery in solid Earth geoscience," *Science* **363**(6433), 6433 (2019).
28. Q. He, Z. Wang, and J. Li, "Application of the Deep Neural Network in Retrieving the Atmospheric Temperature and Humidity Profiles from the Microwave Humidity and Temperature Sounder Onboard the Feng-Yun-3 Satellite," *Sensors* **21**(14), 4673 (2021).
29. Y. Luo, H. Wu, T. Gu, *et al.*, "Machine Learning Model-Based Retrieval of Temperature and Relative Humidity Profiles Measured by Microwave Radiometer," *Remote Sens.* **15**(15), 3838 (2023).
30. T. Ren, M. F. Modest, A. Fateev, *et al.*, "Machine learning applied to retrieval of temperature and concentration distributions from infrared emission measurements," *Appl. Energy* **252**, 113448 (2019).
31. lsqnonlin - Solve nonlinear least-squares (nonlinear data-fitting) problems - MATLAB, <https://www.mathworks.com/help/optim/ug/lsqnonlin.html>.
32. S. Adachi, "Optical dispersion relations for GaP, GaAs, GaSb, InP, InAs, InSb, $\text{Al}_x\text{Ga}_{1-x}\text{As}$, and $\text{In}_{1-x}\text{Ga}_x\text{As}_y\text{P}_{1-y}$," *J. Appl. Phys.* **66**(12), 6030–6040 (1989).
33. M. R. Querry, *Optical Constants*, Contractor Report CRDC-CR-85034, (1985).
34. Y. Zhong, S. D. Malagari, T. Hamilton, *et al.*, "Review of mid-infrared plasmonic materials," *J. Nanophotonics* **9**(1), 093791 (2015).
35. M. Francoeur, M. Pinar Mengüç, and R. Vaillon, "Solution of near-field thermal radiation in one-dimensional layered media using dyadic Green's functions and the scattering matrix method," *J. Quant. Spectrosc. Radiat. Transf.* **110**(18), 2002–2018 (2009).
36. F. Solheim, J. R. Godwin, E. R. Westwater, *et al.*, "Radiometric profiling of temperature, water vapor and cloud liquid water using various inversion methods," *Radio Sci.* **33**(2), 393–404 (1998).
37. A. H. Barrett and P. C. Myers, "Subcutaneous temperatures: A method of noninvasive sensing," *Science* **190**(4215), 669–671 (1975).
38. S. Mizushina, T. Shimizu, K. Suzuki, *et al.*, "Retrieval of Temperature-Depth Profiles in Biological Objects from Multi-Frequency Microwave Radiometric Data," *J. Electromagn. Waves Appl.* **7**(11), 1515–1548 (1993).
39. J. W. Hand, G. M. J. Van Leeuwen, S. Mizushina, *et al.*, "Monitoring of deep brain temperature in infants using multi-frequency microwave radiometry and thermal modelling," *Phys. Med. Biol.* **46**(7), 1885–1903 (2001).
40. F. Bardati, V. J. Brown, and G. D. Bernardo, "Multi-frequency Microwave Radiometry for Retrieval of Temperature Distributions in the Human Neck," *The Journal of Photographic Science* **39**(4), 157–160 (1991).
41. J. Tremblay, A. Prakash, D. Acuna, *et al.*, "Training Deep Networks with Synthetic Data: Bridging the Reality Gap by Domain Randomization," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1082, 1082–10828 (2018).
42. D. Shymkiv and Y. Xiao, "Direct Measurement of Thermal Conductivity via Time-Resolved Thermal-Radiation Spectroscopy," in *CLEO 2025*, Technical Digest Series (Optica Publishing Group, 2025), paper SS141.
43. D. Shymkiv, "Accurate Depth-resolved Temperature Profiling via Thermal-radiation Spectroscopy: Numerical Methods Vs Machine Learning", [10.5281/zenodo.15707540](https://doi.org/10.5281/zenodo.15707540)