Contents

Thesis Title: Predicting and Analyzing the Thermal and Electrical Properties of Materials Using Advanced Machine Learning Models

Author: Mohsen Dehbashi

Supervised by: Jerzy Bodzenta

Assistant supervisor: Justyna Juszczyk-Synowiec

Institution: Institute of Physics, Centre for Science and Education, Silesian University of Technology

Year: 2026

Acknowledgment

I would like to express my deepest gratitude to my supervisor, prof. dr hab. inż. Jerzy Bodzenta, for his invaluable guidance, patience, and support throughout this research journey. His profound expertise and insightful feedback have played a crucial role in shaping this work. Beyond his academic mentorship, I sincerely appreciate his willingness to explore new ideas and his dedication to fostering an environment that promotes learning and innovation. I would also like to extend my heartfelt thanks to my assistant supervisor, dr inż. Justyna Juszczyk-Synowiec, for her constructive advice and encouragement, as well as to dr hab. inż., prof. PŚ Anna Kazmierczak-Balata, whose help and support during the course of this work were greatly appreciated. It has been an honor to work under such guidance, and I am truly grateful for the knowledge and inspiration I have gained from them.

List of Abbreviations:

4PP – Four-Point Probe

AARD – Average Absolute Relative Deviation

AFM – Atomic Force Microscopy

Al – Aluminum

EC – Electrical Conductivity

FEM – Finite Element Modeling

GeSbTe – Germanium Antimony Telluride

h-BN – hexagonal Boron Nitride

ITO – Indium Tin Oxide

MAE – Mean Absolute Error

MD – Molecular Dynamics

ML – Machine Learning

PMMA – Polymethyl Methacrylate

RMSE – Root Mean Square Error

R^2 – Coefficient of Determination (R-squared)

SiC – Silicon Carbide

SiO_2 – Silicon Dioxide

SThM – Scanning Thermal Microscopy

TC – Thermal Conductivity

TCR – Thermal Contact Resistance

TDTR – Time-Domain Thermoreflectance

vdP – vdP method

XRD – X-ray Diffraction

YAG – Yttrium Aluminum Garnet

ZnO – Zinc Oxide

ZnS – Zinc Sulfide

1. Introduction and Literature Review (Thermal)

Here, a general overview is presented, and the topic is discussed in greater detail in the subsequent subsections.

As active devices continue to shrink in size, especially in electronics, efficient heat dissipation has become one of the most critical factors determining their performance. When materials are scaled down to thin films and submicron structures, the mechanisms of heat transfer deviate from those observed in their bulk counterparts. At these small dimensions, theoretical models based on classical laws often fail to describe thermal behavior accurately, and heat transport becomes strongly influenced by surface and dimensional effects. Understanding these effects are therefore essential for designing electronic, optoelectronic, and thermoelectric systems. Thermal conductivity (TC), which describes how efficiently a material transfers heat under a temperature gradient, is one of the most fundamental thermophysical properties. It plays a key role in determining how effectively a device can dissipate heat.

Over the past two decades, optical pump–probe techniques, particularly Time-Domain Thermoreflectance (TDTR), a variant of the flash method, have been widely employed for nanoscale TC measurements [1, 2, 3] with vertical resolution at the level of several dozen nanometers, and extensively studied [4, 5, 6, 7]. TDTR is a non-contact method with picosecond time resolution and the ability to probe buried interfaces, making it effective for layered systems.It was initially developed for metallic samples. By comparison, Frequency-Domain Thermoreflectance (FDTR) is another common method, a variation of photothermal spectroscopy. Instead of measuring the signal as a function of time delay, FDTR varies the modulation frequency of the pump beam to extract thermal data.Each technique comes with inherent limitations. For instance, the lateral spatial resolution of TDTR and FDTR is restricted to the micron scale due to the diffraction limit of light. This prevents it from fully capturing the submicron-scale TC distribution at sample surface.

By contrast, Scanning Thermal Microscopy (SThM) achieves submicron-scale lateral resolution but is sensitive to surface roughness, probe–sample contact mechanics, and substrate thermal properties, all of which introduce uncertainties. A critical challenge is the influence of surface roughness. Even slight variations in topography at the submicron scale can disrupt the interaction between the measurement probe and the sample surface. This disturbance distorts the thermal signals and complicates the extraction of intrinsic film properties. In SThM, surface roughness can significantly alter the measured signals, making it difficult to obtain reliable values for thin-film conductivity. Such effects underscore the extent to which submicron-scale morphology can dominate thermal measurements and remain a central obstacle in the field. These challenges emphasize that submicron-scale thermal characterization is still an open problem, with no single technique yet capable of providing a complete and fully reliable solution.

TDTR and SThM are the two most widely used nanoscale thermal metrology tools: TDTR enables accurate, non-contact measurements of thin films and interfaces, while SThM provides higher-resolution local thermal mapping. Each technique has its own limitations and advantages; together, they cover complementary measurement regimes, which will be discussed in the following subsections.

The aim of this work is to improve the reliability of SThM based TC measurements in thin films by addressing surface roughness effects. By carefully analyzing the advantages and limitations of

existing methods, particularly SThM, and by developing refinements in data interpretation, this thesis seeks to contribute to accurate characterization of thermal transport at the submicron scale.

1.1. Time-Domain and Frequency-Domain Thermoreflectance

TDTR is an optical method widely used to measure thermal transport in thin films and layered materials. The principle of the method is the same as the flash method [8]. A pulsed pump laser beam is used to heat the sample, while a delayed probe beam monitors changes in the surface temperature through changes of its reflectivity. To localize the heat source (light absorption) and the origin of measured signal (changes in reflectivity), a thin aluminum film, typically about 80–90 nm thick, is deposited on the sample surface. Since the reflectivity of metals depends on temperature, changes in reflected light intensity provide information about changes of sample surface temperature. The reflected pulse intensity is detected by a photodiode. The measurement is a stroboscopic measurement, after each sample heating pulse the delay of the probe pulse changes, which allows for recording changes in surface temperature as a function of time [6].

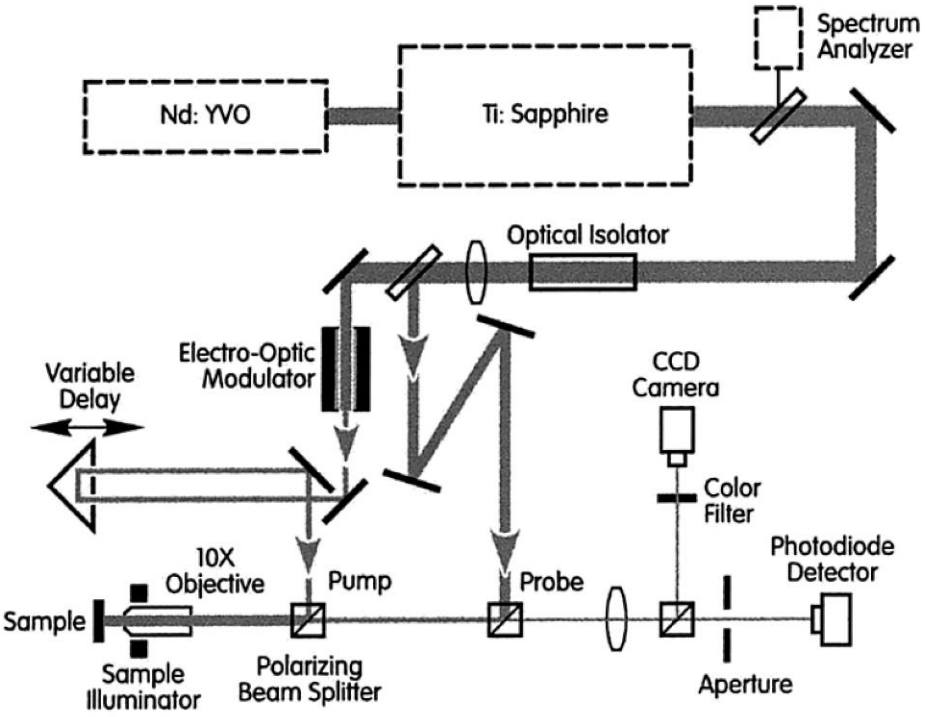


Figure 1.1. Schematic diagram of a TDTR system combined, showing the main components of the experimental setup [4].

An example of a TDTR experimental setup is presented in Figure 1.1. In practice, picosecond time resolution is achieved using mode-locked lasers that generate a sequence of ~1 ps pulses at repetition rates near 76 MHz. The laser output is split into two beam paths, a pump and a probe, with the relative optical path lengths adjusted using a mechanical delay stage. When the pump beam strikes the sample surface, each pulse deposits a small amount of energy, producing a sudden temperature rise of a few Kelvins in the thin metal film. The subsequent cooling of this surface region is monitored by the reflected probe pulses. TC is then determined by comparing the measured cooling curve with theoretical models and optimizing the relevant free parameters [4].

In TDTR, the TC (κ^s) is extracted by fitting experimental measurements to a theoretical model that describes heat diffusion in both the frequency and time domains. The analysis starts from a frequency-domain solution describing periodic surface heating and then reconstructs the corresponding time-domain signal to reproduce the experimental response measured by a lock-in amplifier.

In the frequency domain, the model describes the steady-state, periodically modulated temperature response of the sample surface. For a single semi-infinite medium, the temperature rise $g(r)$ at a radial distance r from a periodically modulated point heat source is given by [6]

$$g(r) = \frac{\exp(-qr)}{2\pi\kappa^s r}. \tag{1.1}$$

Here, the TC (κ^s) governs the ability of the material to conduct heat, while the thermal diffusivity (D) characterizes the rate at which the temperature responds to thermal perturbations. The parameter q determines the thermal penetration depth ($1/q$), which represents the characteristic distance over which the thermal wave decays in the material:

$$q = \sqrt{\frac{i\omega}{D}}, \tag{1.2}$$

where $\omega = 2\pi f$ is the angular modulation frequency.

To incorporate the finite spatial extent of the pump and probe laser beams, the solution is expressed in the spatial frequency domain using a Hankel transform. The transformed temperature response $G(k)$, where $k$denotes the radial spatial frequency, is written as

$$G(k) = \frac{1}{\kappa^s(4\pi^2k^2 + q^2)^{1/2}}. \tag{1.3}$$

The surface temperature change sensed by the probe beam, $\Delta\mathcal{T}$, is obtained by integrating the product of the material response and the Gaussian intensity profiles of the pump and probe beams, characterized by radii $w_0$and w_1, respectively:

$$\Delta\mathcal{T} = 2\pi A \int_0^\infty G(k) \exp\left(\frac{\pi^2k^2(w_0^2 + w_1^2)}{2}\right)k\, dk, \tag{1.4}$$

where A is the amplitude of the absorbed heat at the modulation frequency. Considering that G(k) depends on q and is a function of frequency, the formula provides a basis for interpreting the results obtained in the FDTR method. The time-domain formulation describes how the surface temperature is sampled as a function of the delay time t between the pump and probe laser pulses. Important parameters include the pump–probe delay time t, which defines the temporal separation between heating and probing, and the thermoreflectance coefficient (dR/dT), which relates changes in surface temperature to changes in optical reflectivity. The laser repetition rate $1/\tau$ sets the temporal spacing between successive pulse trains. Since the laser pulse duration is much shorter than the relevant thermal time scales, the experimental signal is modeled by summing the frequency-domain responses over the harmonics of the laser repetition rate.

This approach yields the complex reflectivity response expressed as

$$\mathrm{Re}[\Delta R_M(t)] = \frac{dR}{dT}\sum_{m=-M}^{M}(\Delta\mathcal{T}(m/\tau + f) + \Delta\mathcal{T}(m/\tau - f))\exp(i2\pi mt/\tau), \quad 1.5$$

$$\mathrm{Im}[\Delta R_M(t)] = -i\frac{dR}{dT}\sum_{m=-M}^{M}(\Delta\mathcal{T}(m/\tau + f) - \Delta\mathcal{T}(m/\tau - f))\exp(i2\pi mt/\tau). \quad 1.6$$

The real and imaginary components of the complex reflectivity are directly related to the voltages detected by the lock-in amplifier. These signals are related to the final lock-in output voltage $V_f(t)$ and the average detector voltage V_0 through

$$\frac{V_f(t)}{V_0} = \frac{Q}{\sqrt{2}}\frac{\Delta R(t)}{R}, \quad 1.7$$

where R is the baseline reflectivity and Q is the quality factor of the resonant detection circuit tuned to the modulation frequency f. By treating the TC (κ^s) as an adjustable parameter and iteratively fitting the modeled ratio of the in-phase signal to the out-of-phase signal to the experimental data, the TC of the sample is accurately determined.

Over the years, a key extension of TDTR has broadened its capability, particularly through the use of optical filtering strategies (such as two-tint detection) that suppress unwanted pump light reaching the detector and thereby improve measurement fidelity, especially when using fiber lasers at near-infrared wavelengths [9]. The strength of TDTR and FDTR lies in its ability to determine the TC of multilayer systems with distinct thermal properties, such as the pyrolytic carbon and silicon carbide (SiC) coatings used in nuclear fuel particles. It also enables direct measurements on spherical fuel particles, yielding results that are more representative of real operating conditions [5]. Its modeling framework offers an efficient and general approach for analyzing both single-layer and multilayer samples, including explicit treatment of interfaces. Moreover, as a fully optical technique, TDTR and FDTR can be applied across a broad range of environments—from cryogenic systems to high-temperature stages and even extreme conditions such as high-pressure diamond anvil cells [4]. Continued advances in instrumentation, including improved filtering and robust transducer design, have further strengthened TDTR and FDTR as a highly reproducible measurement platforms [9].

At the same time, TDTR and FDTR are not free from limitations. The accuracy of extracted parameters depends heavily on prior knowledge of experimental inputs such as film thickness, beam spot sizes, and heat capacities. Reported uncertainties are around 10% for pyrolytic carbon and about 5% for SiC and alumina [5]. The mathematical models used to interpret TDTR and FDTR data rely on simplifying assumptions, including Gaussian beam profiles, cylindrical symmetry, and semi-infinite substrate boundaries. Deviations from these assumptions in real experiments can introduce systematic errors. Numerical convergence issues and sensitivity to parameter choices such as spot size also complicate analysis [6]. In the case of beam-offset measurements, sensitivity to in-plane conductivity requires the lateral diffusion length to be comparable to the beam radius. If this condition is not satisfied, the signal becomes insensitive to lateral transport. Moreover, multiple parameters (such as in-plane and cross-plane conductivity,

spot size, and interface conductance) are often interdependent, requiring careful experimental design and additional scans at different modulation frequencies [7].

Other experimental artifacts may arise from uncertainty in optical properties, phase errors in lock-in detection, or residual pump leakage into the probe channel. Semi-transparent or buried transducers complicate the interpretation further, since the probe light may interact with multiple layers, requiring more advanced optical–thermal modeling [9].

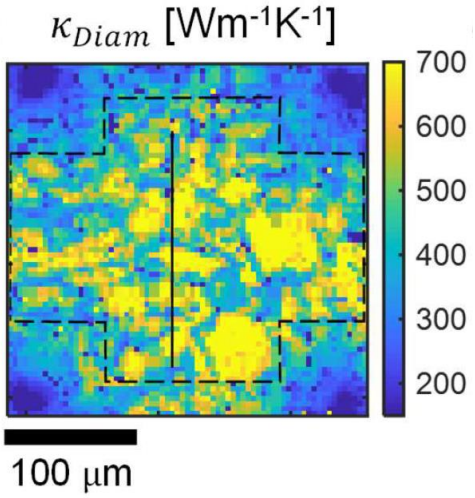


Figure 1.2. Spatial resolution limits of TDTR [11].

Finally, a more critical limitation lies in the inherent spatial resolution of TDTR, and FDTR, which is restricted to the microscale. Recent studies [10, 11] have highlighted that this restriction originates from the diffraction limit of light. The spatial resolution of the measurement is primarily determined by the diameters of laser spots, which define the smallest region that can be independently analyzed. For instance, when using a 50× objective lens, the resulting conductivity maps have a pixel size of approximately 4 μm (Figure 1.2) [11].

This limitation motivates the search for innovative techniques capable of higher-resolution TC measurements. SThM meets this need with sub-micron probes but remains sensitive to topography-induced artifacts. The primary aim of this work is to address and mitigate these roughness-induced limitations in SThM-based TC measurements. The following subsection provides a more detailed overview of the SThM method.

1.2. SThM

Among the techniques available for TC measurements in thin films, SThM has emerged as a valuable tool due to its higher spatial resolution than TDTR and FDTR, and ability to perform localized thermal analysis, as evidenced by prior research [12, 13]. The operational principle of SThM relies on the utilization of a thermal probe, which systematically scans the sample surface while detecting temperature-dependent resistance variations in the probe material. These temperature changes arise from the heat exchanged between the probe tip and the sample surface, which in turn alter the electrical resistance of the probe.

By integrating Atomic Force Microscopy (AFM) with a thermal probe, SThM facilitates ultra-high-resolution thermal mapping alongside topographical imaging, making it invaluable in

semiconductor, optoelectronic, and material science applications [14]. The configuration of the AFM-SThM apparatus is depicted in Figure 1.3. This setup employs an active feedback technique to stabilize the probe's operating temperature. A thermal control unit delivers the excitation current to the thermal probe. The probe's temperature is inferred from its electrical resistance. This resistance is part of a Wheatstone bridge circuit; any imbalance in the bridge, caused by a change in the probe's temperature due to sample interaction, generates an error signal.

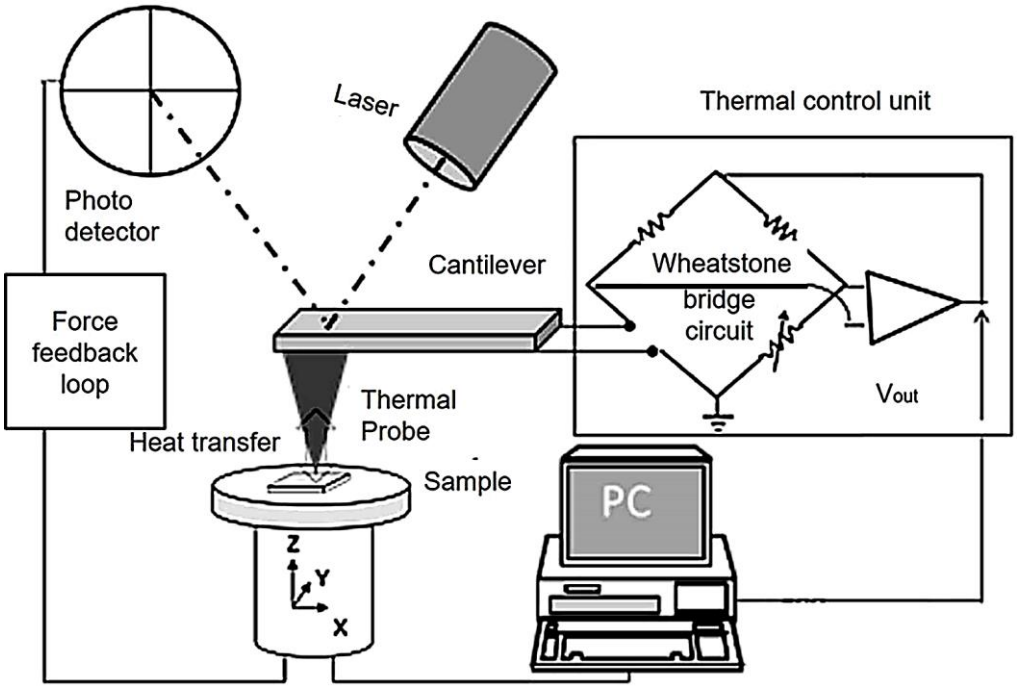


Figure 1.3. The setup scheme of a SThM system based on AFM [14].

The scanning thermal microscope functions in two main modes: temperature contrast mode and conductance contrast mode. In temperature contrast mode, the thermal probe acts as a passive element, recording the temperature at the sample surface, or more precisely, its own temperature, which reflects the sample's temperature. Thus, in this mode, the thermal probe operates as a thermometer. Since resistive probes require an electric current, in temperature contrast mode the current must be kept low enough to prevent Joule heating of the sensor. For this reason, temperature contrast mode is also referred to as the passive mode. In conductance contrast mode, also known as the active mode, the thermal probe serves as both a heater and a thermometer simultaneously. The temperature of the probe is influenced by the rate of heat dissipation from the probe into its environment. Heat transfer occurs through multiple pathways, one of which is the flux from the heated region into the sample [14, 15].

Accurate SThM measurements require reliable calibration techniques to convert raw thermovoltage or resistance signals into meaningful TC values. Experimental calibration involves the use of reference samples with known thermal conductivities to establish a correlation between measured signals and actual properties. This process typically involves heating the probe tip using a stable DC current and monitoring its temperature response under controlled conditions to ensure reliable results [14].

High-resolution temperature mapping in active microdevices, such as nanowire diodes and graphene-based components, has been enabled by SThM. Advances in probe technology have

pushed the spatial resolution to below 50 nm, making it ideal for investigating localized heating effects in electronic circuits [13, 14].

Although SThM provides significant advantages in submicron-scale thermal characterization, TC measurements are subject to several inherent constraints that affect their accuracy and reproducibility. One major limitation arises from geometric factors associated with the probe tip, including its shape, sharpness, and contact area, which can lead to inconsistent thermal interactions with the sample surface.

The inconsistencies in SThM measurements (particularly those related to surface roughness) are especially problematic at the submicron scale tip–sample junction. They distort local heat flow and generate artifacts in the thermal maps, introducing uncertainties that ultimately compromise the reliability of quantitative TC assessments [16, 17, 18]. The interface-dependent nature of heat transfer further exacerbates these challenges, as different interfaces exhibit distinct thermal contact resistances (TCRs) that are difficult to standardize. When the SThM probe and specimen interact, the presence of asperities (topological irregularities) significantly diminishes the true physical contact area between them. This reduction in effective contact creates localized thermal air gaps, manifesting as increased TCR and introducing measurement variability [19]. SThM struggles with quantifying TC in thin films due to these interactions, surface roughness, and thermal spreading resistance [20, 21, 22, 23].

Studies on polymeric thin films, such as polydimethylsiloxane and epoxy, demonstrated that smoother surfaces with a root mean square roughness below 13.51 nm exhibit minimal artifacts, whereas rougher surfaces induce significant distortions. This is due to increased thermal resistance caused by reduced real contact area and enhanced heat reflection at the probe-sample interface [30]. Similar effects were observed in polymer/SiO_2 composites, where peak heights exceeding 32 nm led to significant thermal distortions, reinforcing the importance of surface morphology in determining measurement accuracy [30]. Experimental studies on germanium and silicon samples with ultra-smooth surfaces (root mean square roughness as low as 0.8 nm and 0.3 nm, respectively) demonstrated that in the absence of roughness-induced variability, solid-solid thermal resistance can be determined with high accuracy. However, deviations occur when roughness increases, leading to an irregular and unpredictable contact area that distorts the thermal signal [24]. To overcome the challenges, ultracompliant thermal probe arrays have been developed to enhance SThM's capabilities for mapping non-planar surfaces. These micromachined polyimide probes maintain contact with samples despite topographical variations, eliminating the need for mechanical feedback and improving measurement consistency across complex structures [25].

The application of established mechanical contact models, including the Rabinovich formulation and the Derjaguin-Müller-Toporov theory, encounters theoretical and practical limitations when attempting to accurately describe the multifaceted nature of tip-sample interactions in scanning probe microscopy [26]. The primary shortcoming of these models lies in their inability to account for irregular surface morphologies and asymmetric contact geometries. When applied to surfaces with non-standard topographical features, these theoretical constructs fail to provide reliable predictions of interfacial behavior. This discrepancy arises because the models do not incorporate critical submicron-scale interaction phenomena, including localized elastic-plastic deformation, atomic-scale adhesion forces, and spatially varying contact stiffness, all of which play decisive roles in determining the mechanical and thermal response at the tip-sample junction.

The inadequacy of existing mechanical models to represent realistic experimental conditions highlights a broader methodological challenge in SThM. The combined limitations of current thermal and mechanical modeling approaches reveal a critical gap in analytical capabilities, underscoring the need for next-generation techniques that can more faithfully capture the complex interplay of forces, heat-transfer mechanisms, and material responses at the submicron scale. Among these challenges, interfacial thermal transport between thin films and their supporting substrates remains one of the most significant and persistent obstacles to accurate characterization. This substrate-interference effect becomes particularly pronounced in films with thicknesses in the nanoscale range (typically below 100 nm), where the measured thermal properties are increasingly dominated by substrate-driven physical phenomena [27].

To understand the thermal interactions at the nanoscale, this study adopts a theoretical framework [15], which models the probe-sample system using an electrical circuit analogy. The basic formula governing heat exchange is based on the principle that the electrical power dissipated in the probe (P_{el}) is equal to the heat flow through the total thermal conductance (G_{th}) between the heated region and the environment:

$$P_{el}^{s_i} = G_{th}^{s_i}(T_p - T_a), \tag{1.8}$$

where T_p is the probe temperature and T_a is the ambient temperature. In this model, the total thermal resistance ($R_{th}^{s_i} = 1/G_{th}^{s_i}$) is a network of resistances representing three distinct heat-transfer channels:

$$\frac{1}{R_{th}^{s_i}} = \frac{1}{R_{env}} + \frac{1}{R_{contact}^{s_i} + R_{spread}^{s_i}} \tag{1.9}$$

These channels are defined as follows: R_{env} (Environmental Loss), related to heat lost through the cantilever to the probe base and through the surrounding air, $R_{contact}$ (Contact Resistance) - thermal contact resistance at the tip–sample interface, and $R_{spread}^{s_i}$ (Spreading Resistance) - resistance to heat flow through a circular contact of radius r^{s_i}, related to the sample's TC (κ^s) through the following equation for.

$$R_{spread}^{s_i} = \frac{1}{4\kappa^s r^{s_i}} \tag{1.10}$$

From this fundamental heat-transfer perspective, the apparent thermal resistance measured by the SThM probe ($R_{th}^{s_i}$) arises from the parallel combination of these separate thermal pathways. In the context of thin-film characterization, this network is expressed in the following relationship relationship [58]:

$$(R_{th}^{s_i})^{-1} = h + \left(R_{th,P}^{s_i} + \frac{1}{4\kappa^s r^{s_i}}\right)^{-1}, \tag{1.11}$$

where, h represents the effective heat-transfer coefficient for convective cooling, corresponding to the inverse of the environmental resistance ($1/R_{env}$). $R_{th,p}^{s_i}$ represents the intrinsic thermal

resistance of the probe–sample interface ($R_{contact}$), and $\frac{1}{4\kappa^{s}r^{s_i}}$ represents the spreading resistance of a sample with TC of κ^{s} and effective radius r^{s_i}.

Accurate thermal characterization of thin films therefore requires more than simply determining thermal conductivity; it demands consideration of several interdependent physical parameters that collectively shape heat-transfer behavior. Two of the most critical are the effective radius (r^{s_i}) and the probe–sample boundary resistance ($R_{th,p}^{s_i}$). These parameters interact in complex ways and can influence the interpretation of experimental results.

The nature of these parameters for the TC measurement challenges highlights the need to develop innovative characterization methodologies that possess key attributes, for example, adaptability to diverse surface roughness, reliability against experimental variability, and the capacity to account for all thermal effects. The realization of this methodological advancement would require integration of several experimental and analytical approaches. High-resolution surface topography mapping must be combined with spatially correlated thermal signal acquisition to establish precise structure-property relationships. Advanced computational algorithms could then be employed to separate intrinsic material properties from measurement artifacts, while Machine Learning (ML) techniques could help identify patterns in complex, multidimensional datasets. Such comprehensive methodological innovations would yield benefits for both fundamental research and industrial applications. From a scientific perspective, they would provide more reliable and reproducible datasets for validating theoretical models of submicron-scale heat transfer. For technology development, they would facilitate more accurate performance prediction and optimization of thin film devices. Ultimately, these advancements would accelerate progress across multiple disciplines where thermal management is crucial.

1.2.1. Advancements in Reducing Roughness-Induced Artifacts in SThM

As discussed earlier, SThM is a powerful tool for probing submicron-scale thermal properties. However, surface roughness can mask the material's intrinsic TC, complicating data interpretation. To address these challenges, both computational and experimental approaches have been developed to minimize roughness-induced distortions and improve measurement reliability. FEM, which solves the Poisson equation for diffusive heat transfer while considering probe-sample thermal resistance in three dimensions, is regarded as the most accurate computational method. However, FEM demands substantial computational resources, making it less practical for routine measurements [16, 28]. Comparative analyses indicate that FEM provides the highest accuracy in artifact removal, followed by neural networks and the neighbour volume approach. Given FEM's computational intensity, hybrid methods that integrate FEM with simpler techniques are recommended for practical applications [16, 28].

Experimental methods also play a crucial role in improving SThM measurement accuracy. One effective strategy involves optimizing probe parameters by adjusting the amplitude and frequency of the probe's heating voltage. This technique enhances thermal contrast while minimizing distortions caused by thermal inertia [28].

Besides experimental and computational techniques, the choice of sample preparation technique significantly affects SThM results. Mechanical polishing followed by ion beam polishing has been found to produce smoother surfaces with an average root mean square roughness of approximately

38 nm. However, this process may introduce large-scale local topographical variations that could still impact analysis [29].

On the other hand, ion milling, while effective at exposing multilayer structures, results in higher roughness, with an average root mean square roughness of 86 nm. This increased roughness complicates thermal property assessments [29]. Other techniques, including hot pressing and customized probe design, have been proposed to minimize peak heights and enhance real contact area, further reducing roughness-induced artifacts [30].

1.2.2. Advances in Thermal Measurements Calibration

Accurate calibration in SThM is critical for reliable thermal property extraction. Without standardized calibration methods, measured TC values can be affected by probe-sample thermal exchange artifacts, leading to significant deviations from actual material properties. Wilson et al. (2019) highlighted the challenges in probe calibration, emphasizing that uncalibrated SThM measurements often conflate substrate and thin-film contributions, making quantitative thermal characterization difficult [31]. Several calibration strategies have been proposed to enhance measurement accuracy, each with varying degrees of success. Wilson et al. (2019) compared three major calibration approaches. The implicit method relies on curve-fitting probe signals to reference materials but struggles to separate substrate effects from thin-film contributions. This limitation makes it unreliable for accurately determining intrinsic thin-film thermal properties [31]. The step method estimates the thermal exchange radius by measuring variations in signal intensity across patterned surface features. However, inconsistencies arise due to differences in TC between materials, leading to errors in calibration [31]. The intersection method, which calibrates both the thermal exchange radius and contact thermal resistance using multiple reference materials, has demonstrated the highest reliability. This approach reduced deviations from TDTR measurements to within 20%, highlighting its accuracy [31]. Further refinement of calibration methods has included film-on-substrate heat conduction models, which effectively decouple substrate effects in ultrathin films. This advancement is useful for films thinner than the thermal exchange radius, where substrate contributions become dominant, complicating the extraction of intrinsic film conductivity [31].

Beyond traditional methods, innovative calibration strategies have been introduced to improve precision and repeatability in SThM measurements. One such approach involves using SiO_2 steps on silicon substrates, allowing for constant roughness while varying TC. This method provides precise calibration under controlled conditions, reducing measurement uncertainties in thin-film systems [27].

Predictive thermal modelling has also contributed to enhancing SThM's ability to resolve depth-dependent heat transport in thin films. These models address issues related to substrate influence and non-linear heat transfer mechanisms, making them valuable tools for refining calibration techniques [32].

Another SThM calibration method for thin films uses a normalized thermal signal combined with theoretical modelling to determine TC [33]. Measurements were carried out for SiO_2 thin films along with different reference samples. The experimentally determined ratio for the reference samples is represented by solid circles on the resulting graph (Figure 1.4).

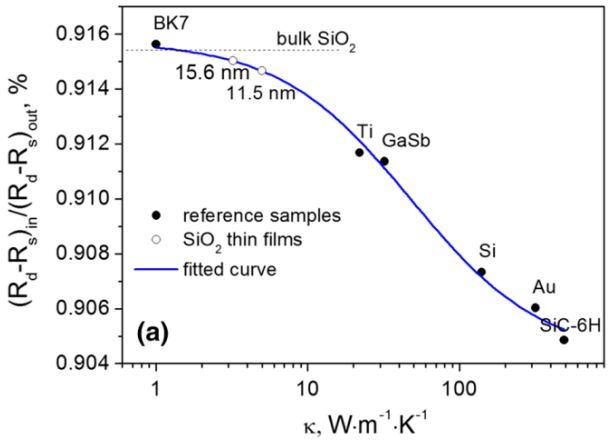


Figure 1.4. Calibration graph relating the thermal signal (vertical axis) to TC (horizontal axis). The solid blue line represents the fitted calibration curve based on bulk reference samples (black dots) and SiO_2 thin films (open circles) [33].

The ratio of thermal resistance when the probe is in contact ($R_{\text{th}} \mid \text{in}$) to when it is out of contact in air ($R_{\text{th}} \mid \text{out}$) is determined by the electrical measurements as follows:

$$\frac{R_{\text{th}} \mid \text{in}}{R_{\text{th}} \mid \text{out}} = \frac{(R_d - R_s)_{\text{in}}}{(R_d - R_s)_{\text{out}}}. \qquad 1.12$$

Here, $(R_d - R_s)_{\text{in}}$ represents the difference between dynamic and static electrical resistances measured at the sample surface, whereas $(R_d - R_s)_{\text{out}}$ represents the corresponding resistance difference measured in air. This ratio depends on the sample's TC κ and can be approximated by the rational expression [33]:

$$\frac{R_{\text{th}} \mid \text{in}}{R_{\text{th}} \mid \text{out}} \approx \frac{1 + A\kappa}{1 + B\kappa} \qquad 1.13$$

The parameters A and B correspond to terms that capture the combined effects of the probe–sample contact radius, the probe–sample boundary resistance, and the convective heat transfer. The solid lines in the Figure 1.4 correspond to the best fit of Equation 1.13. This fitted curve is then used to determine the TCs of the layered samples from the $\frac{R_{\text{th}}|\text{in}}{R_{\text{th}}|\text{out}}$ ratio (shown by hollow circles in the figure). Figure 1.4 specifically presents the SiO_2 thin-film results, where the dotted line indicates the signal level corresponding to the known bulk TC of SiO_2. This calibration procedure for determining thin-film TC provides the foundation for the ML-based calibration method developed in this work. In this framework, real surface topography is incorporated into the calibration process, which potentially could offer improved accuracy [53].

2. Introduction and Literature Review (Electrical)

Accurately determining the electrical conductivity (EC) of thin films is a challenge in materials characterization. Reliable EC measurements in thin films are primarily complicated by edge effects, where the proximity of the sample's boundaries distorts the current distribution, leading to non-uniform current paths and thus complicating the overall measurements.

The Four-Point Probe (4PP) and van der Pauw (vdP) techniques are the two most commonly used methods for TC measurements in thin films. The vdP method is often considered precise for homogeneous and isotropic thin films. However, it requires strict conditions: uniform film thickness, negligible contact size, and symmetric contact placement [34].

The 4PP method, in contrast, is simple, and versatile. It is also relatively easy to use on samples of different shapes and sizes, including irregular thin films. The main drawback of 4PP lies in its sensitivity to geometry and edge effects. Probe spacing, sample dimensions, and measurement location can all alter the observed conductivity, requiring correction factors to ensure accuracy. Without such corrections, 4PP values can deviate significantly from intrinsic conductivity.

2.1. van der Pauw Method

The van der Pauw (vdP) method, introduced in 1958, remains one of the most reliable reference techniques for determining the electrical resistivity of materials [35, 36]. The vdP is able to provide precise sheet resistance measurements using only four electrical contacts positioned along the perimeter of the sample [37, 38].

The theoretical basis [39] of vdP relies on a theorem that connects the sheet resistance $R_s = \rho/t$ of a flat, uniform, and isotropic sample (with thickness t) to two characteristic resistances obtained from measurements using four small contacts placed along the sample's perimeter. This relationship follows from the two-dimensional Laplace equation for electric potential under the assumption that no current escapes through the sample boundaries.

For the van der Pauw theorem to hold, the sample must be flat (Figure 2-1) with constant thickness, composed of a homogeneous and isotropic material, and equipped with four contacts placed along the boundary that are negligibly small compared to the overall sample dimensions.

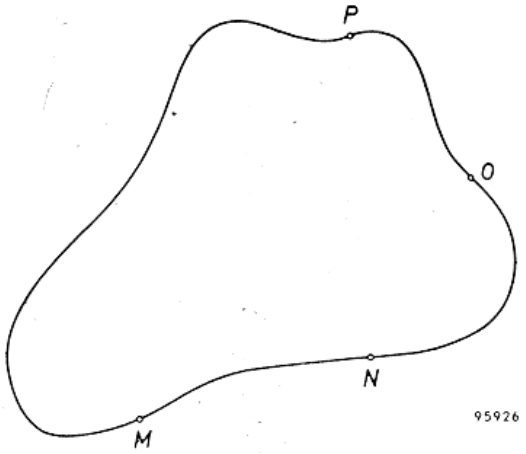


Figure 2-1. A flat sample equipped with four edge contacts (M, N, O, and P) used for vdP measurements [39].

The measurement procedure requires determining two characteristic resistances by interchanging the current and voltage terminals located at the four contacts, labeled M, N, O, and P. In the first configuration, a current I_{MN} is injected between M and N, and the resulting voltage V_{PO} is measured between P and O, giving $R_{MN,OP} = V_{PO}/I_{MN}$. In the second configuration, the current is applied between N and O, and the voltage is measured between M and P, yielding $R_{NO,PM} =$

V_{MP}/I_{NO}. The fundamental vdP equation connects these two resistances with the sample thickness t and specific resistivity ρ through

$$\rho = \frac{\pi t}{\ln(2)} \cdot \frac{R_{MN,OP} + R_{NO,PM}}{2} \cdot f\left(\frac{R_{MN,OP}}{R_{NO,PM}}\right), \quad 2.1$$

where the function f depends only on the ratio of the measured resistances. In cases where the sample or the contact placement is perfectly symmetrical (like square), the two measured resistances are equal ($R_{MN,OP} = R_{NO,PM}$). Under this condition, the correction factor becomes $f = 1$, and the resistivity simplifies to

$$\rho = \frac{\pi t}{\ln(2)} R_{MN,OP}. \quad 2.2$$

This equation represents the fundamental relationship governing the vdP method.

Although the vdP method typically requires direct electrical contacts on the sample surface, which can make it a time-consuming and effectively destructive process, it remains widely used as a reference technique. The need to fabricate or attach contacts, combined with the method's sensitivity to geometrical imperfections, often limits its practicality for rapid or repeated measurements. Measurements obtained using this method have been used to determine the actual electrical conductivity values of the thin films investigated in this study.

Ideally, the method assumes infinitesimally small contacts, but in practical scenarios, finite contact sizes can introduce systematic deviations. As the contact size increases relative to the sample dimensions, errors become more pronounced, leading to underestimation of resistivity values [35, 36]. Similarly, sample thickness plays a crucial role in measurement accuracy. The vdP method is best suited for thin films, but when the thickness is comparable to the lateral dimensions, three-dimensional current flow effects become significant. This necessitates the introduction of correction factors to obtain accurate resistivity values. FEM has been widely used to characterize these thickness-dependent effects and develop suitable correction methodologies [36, 38].

Additionally, to mitigate errors related to the boundary conditions, analytical expressions incorporating correction factors for finite sample size and asymmetric probe placement have been derived [35].

Given the deviations from ideal conditions, several correction strategies have been proposed to enhance measurement accuracy. Finite contact size corrections have been developed through correction tables and computational models that relate measured resistances to actual resistivity values based on known sample geometries. Experimental studies confirm that applying these corrections significantly improves the reliability of vdP measurements [36]. Material inhomogeneity considerations play a crucial role in refining vdP measurements, as variations in resistivity due to compositional gradients can distort current flow. To address this, spatially resolved conductivity mapping techniques have been employed as complementary tools for refining vdP measurements [36].

Computational techniques such as finite element simulations have been instrumental in validating these correction factors against experimental data. These approaches systematically evaluate multiple simultaneous non-idealities, providing comprehensive guidelines for accurate resistivity measurements [35, 36].

2.2. The 4 Point Probe Technique

The 4PP method is widely used to determine materials resistance, as it effectively eliminates errors caused by contact resistance. However, measurement accuracy is influenced by factors such as sample geometry, and probe positioning. Recent studies have examined these aspects, leading to improved methodologies for more reliable EC assessments [40, 41, 42].

The 4PP method operates by applying a known current through two probes while measuring the voltage drop across other two probes. This setup ensures that the measured voltage is independent of contact resistance, making it a more accurate technique compared to the traditional two-probe method [40, 41]. The Figure 2-2 illustrates the 4PP method, depicting the two primary configurations: a linear arrangement (a) and a square arrangement (b). Each setup features equidistant spacing between the probes.

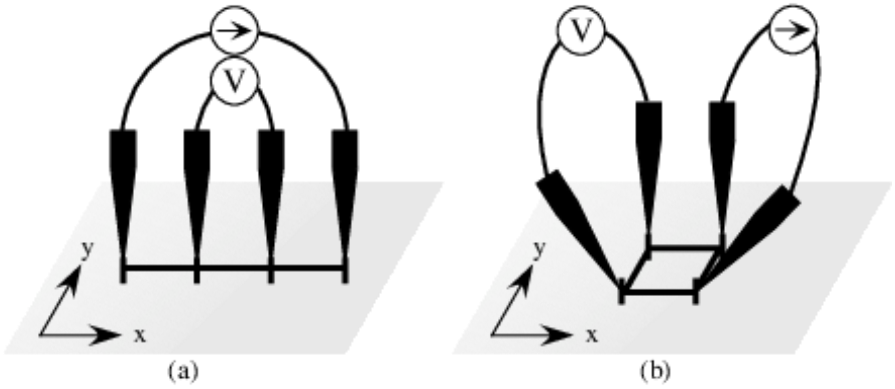


Figure 2-2. The linear (a) and square (b) arrangements of the 4PP method are shown in these schematic drawings, with the specific current and voltage probes identified [43].

In the characterization of electronic materials, the four-probe configuration is beneficial as it minimizes interference from external resistance and improves the precision of conductivity measurements. Studies have demonstrated that this method is crucial in evaluating charge transport behavior, particularly in systems where internal resistance plays a significant role [41].

Additionally, advanced models have been introduced to refine 4PP measurements, considering factors such as mixed conduction pathways and near-surface effects. These models offer a more comprehensive understanding of charge transport and have proven useful in improving the accuracy of electrical measurements in complex materials [42].

Sample geometry plays a significant role in determining accurate conductivity values. The size, shape, and thickness of the material influence the current distribution, making it necessary to apply correction factors to ensure reliable results [40, 41].

Variations in electrical properties have been observed across large-area films, necessitating multiple measurement points to capture these differences accurately. In research settings, materials are often segmented into smaller regions to assess variations and obtain a more precise evaluation of their properties [40].

In some cases, current flow is influenced by both surface and bulk conduction, making conventional measurement models less effective. Studies have shown that at smaller probe spacings, surface conductivity dominates, whereas at larger probe spacings, bulk conduction plays

a more significant role. This highlights the need for advanced modelling approaches that account for multiple conduction pathways instead of assuming a uniform distribution of charge carriers [42].

For flexible and multilayer materials, variations in measurement location can result in significant differences in resistance values. These inconsistencies highlight the importance of precise experimental protocols to account for substrate-related effects. Studies have emphasized that flexible substrates, in particular, can introduce additional challenges due to mechanical strain and non-uniform current flow, which must be considered when interpreting results [40].

In photovoltaic applications, the impact of substrate material on charge transport efficiency has been studied in detail. It has been observed that different probe configurations affect the ability to measure electrical properties accurately, particularly at higher voltage levels. Reducing unwanted resistances in the system improves the accuracy of conductivity measurements and enhances the overall efficiency of electronic components [41].

To improve 4PP techniques accuracy, numerical simulations and computational methods have been integrated into experimental work to refine measurement techniques further. These approaches help minimize systematic errors and provide a more precise understanding of how geometric factors influence conductivity measurements [42].

Recent studies have introduced various modifications and enhancements to the traditional 4PP technique. Yang Lu et al. developed a 3D-printed 4PP station using stainless steel tapestry needles electroplated with nickel and gold to minimize contact resistance variability. Their system incorporates a microcontroller-based source measurement unit (SMU) for precise electrical measurements. They also highlight the impact of geometric distortion in non-uniform samples [44]. Waremra and Betaubun applied the 4PP method to both bulk and thin-sheet materials, using a constant DC current while measuring the voltage drop across the inner probes [45]. Similarly, Mosavi et al. explored both two-wire and four-wire probe configurations for CIGS thin-film solar cells, demonstrating that the four-wire approach effectively eliminates lead resistance, improving measurement accuracy [41]. Ju et al. introduced a microscopic 4PP integrated with AFM, achieving high spatial resolution. Their method incorporates geometrical correction factors to enhance precision [46].

Correction factors play a crucial role in improving the accuracy of EC measurements, particularly when accounting for sample geometry and probe placement. Several studies have focused on refining these aspects. Chelly et al. developed an improved analytical model that applies correction factors for sample thickness and diameter. Using numerical simulations and experimental validation, they derived correction factors for both semi-infinite bulk materials and two-dimensional sheet samples [47]. Another comprehensive study on 4PP techniques emphasizes the importance of correction factors for probe positioning near sample edges and lateral sample dimensions. It introduces mathematical models for different samples, extending correction factors to account for directional resistivity variations [34]. Ji-Kwan Kim et al. proposed a micro-4PP with a square probe configuration to mitigate inaccuracies caused by sample curvature and edge effects [48]. Smits' work refines correction factors for sheet resistivity measurements, incorporating adjustments for rectangular and circular samples of finite size [49].

The important physical quantity obtained using 4PP is the sheet resistance (R_{sh}), defined as the bulk resistivity (ρ) divided by the film thickness (t). Under the assumption that the film behaves as a quasi-two-dimensional system, this relationship takes the form $R_{sh} = \rho/t$.

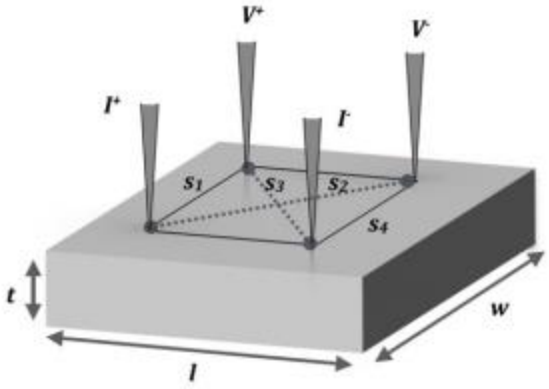


Figure 2-3. Arrangement of the square 4PP method used in this study [34].

For thin-film measurements, the injected current can be assumed to spread cylindrically from the point of injection. In a square 4PP configuration, the probes are positioned at the four vertices of a square of side length s. Current (I) is usually introduced between two diagonally opposite probes, while the potential difference (V) is measured across the remaining pair. The potential distribution from a point current source on the sheet follows a logarithmic dependence, forming the theoretical basis for the sheet-resistance expression.

For a square array of side length s, where the internal probe separations satisfy $s_1 = s_4 = s$ and $s_2 = s_3 = \sqrt{2}s$, the derivation simplifies. This yields the sheet-resistance formula for an infinite sheet [34]:

$$R_s = \frac{\rho}{t} = \frac{2\pi}{\ln 2}\frac{V}{I} \tag{2.3}$$

In this ideal case, the sheet resistance depends solely on the measured ratio V/I and the dimensionless geometric factor $2\pi/\ln\ 2 \approx 9.060$, rendering it independent of the probe spacing s.

In practical scenarios, samples have finite lateral dimensions, making the infinite-sheet assumption invalid. The finite boundaries constrain the current flow, necessitating a geometrical correction factor (F). When a square 4PP array is used on a finite square lamella, the true sheet resistance is obtained by multiplying the ideal expression by this factor:

$$R_s = (\frac{2\pi}{ln\ 2}\frac{V}{I})F \tag{2.4}$$

The correction factor F is a non-trivial function of the ratio between the probe spacing and the sample size (s/d), and in some cases, the probe-array orientation relative to the sample edges. As the sample dimensions become very large compared to the probe spacing, the correction factor approaches unity, and the expression converges to the ideal infinite-sheet formulation.

3. Research Background

The following sections establish the foundational research background for a critical issue in modern materials science: the accurate determination of functional properties in submicron scale and thin-film systems. The discussion begins with the formidable challenges associated with thermal transport, where the breakdown of classical physics, the dominance of boundary and interface effects (Kapitza resistance), and the profound influence of submicron-scale morphology make predicting TC an extrinsically dominated problem. The focus then shifts to the challenges in measuring EC, where the limitations of standard techniques like vdP and 4PP methods, combined with geometric artifacts, complicate the extraction of reliable data. Effective materials characterization, especially thermal properties, is a major challenge, requiring advances in methodologies and computational modeling. Its resolution is crucial for developing next-generation technologies in electronics, photonics, and energy conversion.

3.1. Research Background (Thermal Conductivity)

Modern technology relies heavily on developing and using materials at the submicron scale, especially thin films. These materials are the foundation for devices in many fields, including high-performance microelectronics, photonics, efficient energy systems, and optoelectronics. As the feature sizes within these devices continue to shrink towards atomic scales, their power densities escalate strongly, making effective thermal management not merely an engineering consideration but a fundamental bottleneck that dictates device performance, operational stability, long-term reliability, and ultimate functional viability. Therefore, the precise determination, accurate prediction, and ultimate control of thermal transport properties in thin films have become one of the most critical challenges in modern materials science.

This importance stems from a fundamental paradigm shift in physical behavior: the thermal properties of a material in a thin-film configuration are not intrinsic constants but are instead extrinsic, dependent on a complex interplay of submicron-scale phenomena. When film thickness is reduced to dimensions comparable to or smaller than the mean free path of energy carriers (phonons in dielectrics, electrons in metals), heat transport undergoes a radical transition from the well-understood diffusive regime governed by Fourier's law to a non-equilibrium ballistic regime where classical models fail [27, 50]. In this submicron-scale realm, the material's bulk TC becomes an obsolete concept. Instead, thermal transport is dominated by boundary and interface effects. The film's effective TC becomes a strong function of its thickness, crystallographic quality, grain size, and surface morphology [23, 50]. Most significantly, the TCR (also known as the Kapitza resistance) at the interface between the film and its substrate emerges as the critical, and often limiting, factor in overall heat dissipation [51, 52].

The complexity of this challenge is further compounded by the profound influence of submicron-scale topography on thermal transport phenomena. Surface asperities reduce the true physical contact area between any two solids. This reduced contact area creates localized nanogaps that act as regions of high thermal resistance. This phenomenon introduces significant variability and compromises experimental reproducibility in thermal measurements [19, 22, 30].

The need to address these challenges becomes increasingly pressing as new applications continue to advance. In the microelectronics industry, the thermal management of integrated circuits, high-power transistors, and phase-change memory cells requires knowledge of thin-film TC and interfacial resistance to engineer efficient heat spreading pathways and prevent destructive hot-

spot formation [23, 52]. In energy technologies, such as thin-film thermoelectric generators and solar cells, the deliberate suppression of TC through nanostructuring is a primary pathway to achieving improved efficiencies [23]. For protective coatings and optical films, thermal stability and heat dissipation are key to withstanding high-power operational conditions. In each case, the ability to not only measure but also accurately predict thermal performance based on microstructure, morphology, and interface quality is a critical enabling capability for rational design.

Yet, this ambition for prediction and control is currently hindered by significant limitations in both experimental metrology and theoretical modeling. On the experimental side, even techniques like SThM, which provides unparalleled spatial resolution, face inherent constraints. The accuracy of its measurements is heavily influenced by geometric factors of the probe tip, the stability of the tip-sample contact, and the aforementioned surface roughness, all of which introduce uncertainties that complicate data interpretation [16, 17, 18]. Established mechanical contact models were conceived for idealized scenarios and fail to accurately describe the complex, multi-physics nature of tip-sample interactions at the submicron scale, which are critical for translating SThM signals into quantitative thermal properties [26]. Their inability to account for irregular morphologies, localized deformations, and atomic-scale adhesion forces reveals a critical analytical gap [26]. This gap between theory and experiment hinders the development of reliable, standardized measurement protocols and reduces the overall reliability of quantitative assessments, thereby limiting the utility of acquired data for predictive design.

Overcoming these challenges is imperative for the future of technology. It necessitates the development of new methodologies that include more reliable measurement techniques capable of decoupling intrinsic film properties from pervasive substrate contributions; the creation of multi-scale models that can faithfully simulate heat transfer across realistic submicron-scale morphologies and interfaces; and the establishment of reliable, universally applicable calibration standards.

3.2. Research Background (Electrical Conductivity)

The core challenge in EC measurement arises from the influence of sample geometry. This complexity increases significantly as device features shrink, causing the assumptions underlying theoretical formulas to break down. Edge effects (where the proximity of the measurement region to a physical boundary distorts the current path) become a dominant source of error [34, 40]. The finite and often irregular shape of samples, a common scenario in research and development, introduces boundary conditions that alter current distribution, making simplified models for infinite sheets inaccurate [35, 40].

The most widely used techniques for addressing these challenges, the 4PP and vdP methods, each possess inherent limitations that highlight the difficulty of the problem. The vdP method, while celebrated for its versatility with irregularly shaped samples and its ability to minimize alignment errors [37, 38] , imposes strict requirements: the sample must have uniform thickness, and the contacts must be negligibly small, and the method is destructive. Deviations from these conditions, such as finite contact size, material inhomogeneity, or thickness variations, introduce systematic errors in electrical resistivity measurements [35, 36]. Consequently, the practical value of the vdP method depends heavily on the use of reliable analytical and numerical corrections, which transform an ostensibly simple measurement into a complex inverse problem.

Similarly, the 4PP method is sensitive to geometric factors. Its accuracy is heavily dependent on probe spacing, the proximity of the measurement to sample edges, and the overall dimensions of the film relative to the probe array [40, 41]. In complex material systems, such as flexible electronics or multilayer stacks, additional complications arise from substrate-induced strain, non-uniform current flow, and the interplay between surface and bulk conduction pathways, which can dominate depending on the probe spacing used [40, 42]. These factors imply that a single measurement without proper correction is often meaningless, and obtaining a true EC value requires comprehensive correction frameworks such as Finite Element Method (FEM) or ML-assisted methods.

Emerging applications highlight the increasing need to address these challenges. The development of flexible and wearable electronics requires reliable conductivity measurements on substrates that may bend, stretch, and have inherent surface roughness, all of which violate the standard assumptions of traditional techniques [40]. The optimization of photovoltaic and thermoelectric devices hinges on accurately characterizing often anisotropic and inhomogeneous materials where minor errors in conductivity can lead to significant mispredictions of device efficiency [41, 42]. For industrial quality control and high-throughput screening of new materials, there is a pressing need for measurement protocols that are not only accurate but also rapid and adaptable to non-ideal, real-world sample geometries.

Therefore, the challenges associated with determining and predicting EC in thin films represent a critical bottleneck in materials science. This necessitates the adoption of methodologies such as FEM modeling to account for complex geometries and boundary conditions [35, 36, 47], as well as the development of data-driven approaches like ML to automate correction processes and handle multi-parameter non-idealities. The successful development of reliable, adaptable, and comprehensive frameworks for electrical characterization is essential for accelerating the rational design and implementation of advanced thin-film materials, ultimately supporting continued progress in electronic and energy technologies.

4. Motivation and Hypothesis

4.1. Problem Statement and Research Aim

Accurately determining the TC of thin films remains a significant challenge due to the inherent limitations of SThM. Measurements obtained through SThM are highly sensitive to factors such as surface roughness and probe–sample contact variability, both of which introduce substantial noise and reduce measurement reproducibility. Existing theoretical models offer limited support because they rely on simplifying assumptions that fail to capture the complex thermal interactions occurring at submicron scales. As a result, the field lacks reliable, standardized methodologies for thin-film TC characterization.

The aim of the TC component of this research is to develop an integrated methodology that combines high-resolution spatial mapping, simultaneous thermal–topographical acquisition, normalization procedures, and ML models. This framework is designed to enhance the reliability and reproducibility of thin-film TC measurements, establish a scalable platform (by converting physical topography into numerical datasets) using topography-based descriptors (e.g., inclination, peak-to-valley height), and bridge experimental SThM data with computational predictive modeling.

Accurate EC characterization in thin films encounters comparable difficulties. The widely used 4PP technique is highly sensitive to probe placement, geometric distortions, and edge effects, often leading to inaccurate or inconsistent results. Although the vdP method offers a more rigorous alternative, it requires strict sample geometries, involves destructive preparation, and demands complex experimental setups—factors that limit its practicality for routine thin-film characterization.

The aim of the EC component of this research is to establish a standardized and adaptable correction framework that operates solely on experimental data. This dissertation proposes a data-driven approach in which ML models learn and correct geometric distortions directly from measured 4PP data, eliminating the need for rigid geometries or destructive procedures while enabling more practical, scalable EC characterization. In parallel, an FEM–based correction framework is also developed for comparison.

4.2. Research Objectives

This research is divided into two major components: the determination of TC and the determination of EC in thin films. Each component addresses longstanding limitations in nanoscale characterization by integrating advanced measurement techniques with data-driven modeling. Together, these methodologies aim to establish a unified and scalable framework for thin-film property evaluation.

The overarching objective of the TC component is to develop a reliable and scalable methodology for both measuring and predicting the TC of thin films through the integration of SThM, high-resolution surface topography analysis, and ML techniques.

To achieve this, the research first focuses on the acquisition of high-quality data. A multi-scale spatial mapping strategy is designed, incorporating micro- and submicron-level grids, allowing for the simultaneous collection of localized thermal signals and detailed topographical information. This dual-mapping approach ensures that each thermal measurement can be directly associated with the corresponding morphological features of the sample.

A second objective is the development of normalization and correction procedures to reduce measurement bias and enhance comparability across different samples. Quartz is employed as a stable reference material, enabling calibration of thermal signals and providing a foundation for constructing a consistent measurement framework.

The research then aims to establish quantitative correlations between surface morphology (characterized using parameters such as roughness, skewness, and inclination) and the thermal response of the thin film. These correlations serve as meaningful descriptors for ML model training, enabling the development of predictive tools that incorporate both intrinsic material properties and topographical influences.

Building on these inputs, the next objective is to construct and train Machine Learning models capable of accurately predicting thin-film thermal conductivity. Particular emphasis is placed on Random Forest regression due to its robustness in handling nonlinearities and complex feature interactions arising from topographical variations.

Finally, the proposed methodology is validated across a diverse range of thin films and selected bulk materials. This validation phase is designed to confirm the reproducibility, reliability, and

scalability of the approach, ensuring its applicability to distinct material systems and measurement conditions.

The EC component aims to develop a reliable and adaptable methodology for the accurate determination of thin-film EC using the 4PP technique. Conventional 4PP measurements are often compromised by geometric distortions (such as edge effects and inaccuracies in probe placement) that significantly affect current flow and voltage distribution. To overcome the sensitivity drawbacks of practical EC measurements using 4PP, particularly those related to geometry and probe positioning, this work proposes to establish an approach using FEM and ML techniques separately for measuring the EC in thin films, and to compare the practicality of these methods. FEM is employed to simulate current density distributions in thin films and to provide correction factors for geometry-related distortions. In parallel, an ML model is trained on an experimental dataset of metallic and transparent conductive oxide (ITO) films, enabling it to recognize and correct measurement artifacts caused by probe positioning and sample irregularities.

The first objective is to model measurement distortions using FEM. Numerical simulations are employed to reproduce the influence of sample geometry, edge proximity, and probe misalignment, providing physically grounded correction factors and serving as theoretical benchmarks for experimental measurements.

The second objective focuses on the development of a complementary ML-based correction framework. An extensive dataset of experimental 4PP measurements is used to train ML models capable of automatically identifying and correcting geometric artifacts under a wide range of measurement conditions. This data-driven approach aims to offer a simulation-free alternative that is both rapid and adaptable.

The predictive capability of the ML model is then rigorously validated against high-fidelity experimental references, including vdP measurements. This step ensures that the ML corrections are both accurate and physically meaningful.

4.3. Research Hypothesis

The research is guided by two primary hypotheses corresponding to the TC and EC components.

The central hypothesis for the TC component is that integrating high-resolution thermal mapping with detailed surface topographical analysis, combined with advanced ML algorithms, can overcome the inherent limitations of conventional SThM techniques. By unifying these data sources within a single analytical framework, it is expected that thin-film TC can be measured and predicted with significantly improved accuracy and reproducibility.

This integrated approach is anticipated to remain effective even under challenging conditions, including variations in probe–sample contact, and surface irregularities factors that introduce uncertainty into SThM-based measurements. The hypothesis therefore asserts that a data-driven, multi-modal methodology will enable scalable and reliable TC characterization across a broad range of thin-film systems.

For the EC component, the hypothesis is that both FEM and ML can generate effective correction factors capable of mitigating geometric distortions in 4PP measurements. These distortions (such as edge effects and probe placement deviations) are known to significantly influence current distribution and voltage measurement, leading to inaccuracies in EC determination.

While FEM is expected to provide correction factors with high physical fidelity, the ML-based framework is hypothesized to offer a simulation-free alternative that delivers comparable accuracy with substantially faster computation times. This makes ML particularly suitable for scalable, high-throughput EC characterization, where efficiency and adaptability are essential.

5. Methodology and Experimental Procedures (Thermal)

The proposed ML-based methodology combines SThM with signal processing to achieve submicron spatial and thermal resolution, while the computational component employs ML to extract intrinsic material properties from complex measurement datasets. The materials used in this work include annealed ITO thin films (processed under six different atmospheric conditions) and ZnO thin films that were previously prepared, alongside six well-characterized bulk reference materials spanning three orders of magnitude in thermal conductivity (0.17–450 $W \cdot m^{-1} \cdot K^{-1}$). Detailed descriptions of the methodology are provided in the following sections.

5.1. Materials Used and Instrumentation

Commercial 170 nm thick indium tin oxide films deposited on glass substrates (Hoya, Tokyo, Japan) that had already been annealed under multiple controlled atmospheres were used in this study. These annealing treatments were not performed as part of this work; they had been previously carried out at a uniform temperature of 400 °C in vacuum, air, oxygen (O_2), nitrogen (N_2), carbon dioxide (CO_2), and a nitrogen–hydrogen (N_2–H_2) mixture. Likewise, the zinc oxide thin films analyzed in this study were obtained as pre-fabricated samples. They were produced earlier using atomic layer deposition (ALD), a technique that allows precise control of thickness and composition. The available ZnO films had been deposited at two temperatures, 100 °C and 200 °C, with different ALD cycle counts to generate variations in film thickness and structural properties. Also, thermally isotropic bulk materials was incorporated into the study. These materials were chosen due to their well-defined thermal properties and structural homogeneity, which are essential for accurate model calibration. The bulk materials examined included: glass, glassy carbon, SiC single crystal, yttrium aluminum garnet (YAG) single crystal, ZnO in bulk form, and polymethyl methacrylate (PMMA). The bulk materials served as reference standards to improve model calibration by providing well-characterized thermal properties.

The TC of the thin-film samples listed in Table 5.1 had been previously measured using an AFM (Park Systems XE-70) and two different thermal probes to ensure reproducibility and minimize experimental error. The thermal probe for earlier measurements was the KNT-SThM2an thermal probe (Kelvin NanoTechnology, Glasgow, UK). In the current work, an SThM probe (KNT-SThM-3an, Kelvin NanoTechnology) was used. These two types of probes differ slightly in shape. To minimize noise, all experiments were conducted in a vibration-damped chamber at 25 °C, reducing external disturbances and thermal drift. A quartz reference sample was scanned first for baseline calibration, followed by SThM mapping of the thin-film samples. For both the quartz reference and the samples, a 2×2 μm^2 scan area was used, with a scan rate of 0.05 Hz and a contact force of 1 nN.

For high-precision signal detection, a lock-in amplifier (SR830) filtered out noise, while the probe was powered by AC (2.3 kHz, 0.09 mA amplitude) and DC (1.8 mA) currents for dynamic and static thermal measurements, respectively. This setup ensures reliable data for studying submicron-scale heat transfer. The interaction between the SThM probe and the sample surface is

depicted in Figure Figure 5.1. This schematic also displays two important topographical parameters that affect TC measurements: the probe's inclination and the surface's peak-to-valley distance.

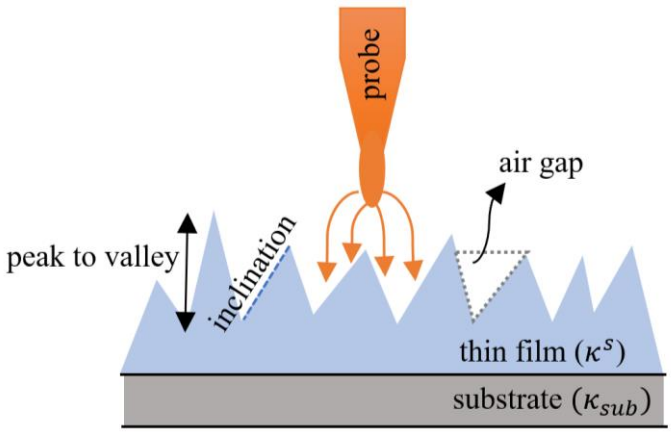


Figure 5.1. Schematic of the SThM probe-surface interaction. The highlighted parameters (inclination and peak-to-valley distance) are critical, as the probe's dual sensing function allows for correlated thermal and topographical data collection at the submicron scale [53].

Table 5.1. Detailed description of the samples used to validate the proposed methodology.

Sample's number	Sample	Substrate	Layer thicknesses (nm)	Actual κ ($W \cdot m^{-1} \cdot K^{-1}$)
1	ITO (1)*	Glass	170	6.4 [54]
2	ITO (2)*	Glass	170	3.5 [55]
3	ITO (3)*	Glass	170	8.3 [55]
4	ITO (4)*	Glass	170	10.6 [54]
5	ITO (5)*	Glass	170	11.8 [55]
6	ITO (6)*	Glass	170	6.7 [54]
7	Glass (Bulk)	-	-	1.1
8	Glassy carbon (Bulk)	-	-	6.3
9	SiC (Bulk)	-	-	450
10	YAG (Bulk)	-	-	12
11	ZnAlO (1)*	Silicon	110	4.29 [56]
12	ZnO (Bulk)	-	-	80
13	ZnO (2)*	Silicon	12	0.25 [57]
14	ZnO (3)*	Silicon	15	0.28 [57]
15	ZnO (4)*	Silicon	38	1.12 [57]
16	ZnO (5)*	Silicon	118	2.81 [57]
17	PMMA (Bulk)	-	-	0.17

* ITO (1) to ITO (6) were annealed in different atmospheres: oxygen, carbon dioxide, vacuum, air, nitrogen, and a nitrogen-hydrogen mixture, respectively. ZnO (1) to ZnO (5) were deposited using ALD with cycle counts of 850, 150, 150, 330, and 900, respectively. The deposition temperatures were 200°C, 100°C, 200°C, 200°C, and 200°C, respectively.

5.2. Thermal Signals and Impact of Surface Topography on Thermal Parameters

The thermal characterization methodology employed a framework wherein the effective thermal resistance ($R_{th}^{s_i}$) of each discrete grid cell within the sample matrix was evaluated relative to the corresponding thermal resistance ($R_{th}^{n_i}$) of the reference material. This differential approach was implemented to normalize measurement variations and isolate the intrinsic thermal properties of the samples under investigation. Amorphous quartz was selected as the reference standard due to its thermal stability, and isotropic heat conduction properties. The fundamental comparative relationship was mathematically expressed through the following dimensionless ratio [58]:

$$\frac{R_d^{s_i} - R_s^{s_i}}{R_d^{n_i} - R_s^{n_i}} = \frac{R_{th}^{s_i}}{R_{th}^{n_i}} = \Gamma_i \tag{5.1}$$

In this formulation, the numerator ($R_d^{s_i} - R_s^{s_i}$) represents the difference in dynamic and static electrical resistances measured at each sample grid cell (i). The denominator ($R_d^{n_i} - R_s^{n_i}$) represents the analogous resistance difference acquired from the quartz reference material at the corresponding spatial coordinate. This normalized ratio (Γ_i) effectively cancels out probe-specific characteristics and instrumental variables. The static resistance component ($R_s^{s_i}$, $R_s^{n_i}$) was acquired under equilibrium conditions using a precisely regulated direct current excitation (I_{DC}=1.8 mA) supplied by a high-stability current source. Conversely, the dynamic resistance measurement ($R_d^{s_i}$, $R_d^{n_i}$) employed an alternating current component (I_{AC} = 0.09 mA amplitude) to probe the transient thermal response characteristics through lock-in detection techniques.

From the Equation 1.11 the thermal resistance ($R_{th}^{s_i}$) demonstrates sensitivity to sample thermal properties and surface morphological characteristics, as this parameter is fundamentally governed by different interdependent variables – the sample TC (κ^s), the effective contact radius (r^{s_i}), the probe thermal resistance ($R_{th,P}^{s_i}$), and the convective heat transfer coefficient (h^{s_i}) - all of which maintain intrinsic relationships with surface topography. This complex interdependence arises from several distinct physical mechanisms that collectively determine thermal transport efficiency at probe-sample interfaces.

Surface roughness increases thermal resistance by reducing the true contact area between the probe and sample. The tip only contacts high points (asperities), while air-filled valleys act as insulating gaps. Conversely, an atomically smooth surface allows for near-perfect contact, maximizing the contact area and minimizing resistance by creating efficient thermal conduction pathways. The thermal model integrating these effects can be derived through combination of Equations 1.11 and 5.1, yielding the following expression:

$$(\kappa^s)^{-1} = 4r^{s_i}\left[\left(\frac{h + \left(R_{th,P}^{n_i} + \frac{1}{4\kappa^n r^{n_i}}\right)^{-1}}{\Gamma_i} - h\right)^{-1} - R_{th,P}^{s_i}\right] \tag{5.2}$$

This relationship reveals several important physical insights. First, the equation highlights the critical importance of reference material topography (through parameter r^{n_i}, here index n is used for the reference sample quartz) in determining accurate TC values, emphasizing that proper

normalization requires careful consideration of both sample and reference surface characteristics. Second, the formulation provides a quantitative framework for distinguishing the various contributions to measured thermal resistance, enabling researchers to distinguish between intrinsic material properties and topographically-induced (through parameter r^{s_i}) measurement artifacts. Furthermore, the ability to quantitatively relate surface topography to thermal transport characteristics represents a significant advancement in submicron-scale thermal metrology, with potential applications in materials development, quality control, and fundamental studies of submicron-scale heat transfer.

5.3. A Factor as ML Input for Substrate and Thickness Consideration

A key challenge was developing an analytical input for ML model to differentiate between bulk and thin-film samples. This distinction was critical because thin films exhibit fundamentally different thermal transport properties, which, if unaccounted for, would introduce significant inaccuracies in the measured TC.

Moreover, the substrate TC introduces additional complexities in thermal transport analysis, as the TCR between the film and substrate can dominate overall heat dissipation. This effect is pronounced in ultrathin films, where interfacial thermal resistance may exceed the intrinsic thermal resistance of the film itself. Therefore, the development of a ML model capable of distinguishing between bulk and thin-film thermal responses was essential to ensure the reliability and accuracy of the experimental findings. The ML model training process required careful consideration of multiple variables, including film thickness, and substrate thermal properties to minimize systematic errors and enhance predictive performance.

Initially to evaluate the effective TC of layered samples deposited on semi-infinite substrates, the probe–sample interaction can be modeled by assuming a discoidal contact with radius b. Within this framework, the effective TC, κ_{eff}, of the system can be expressed as [59]:

$$\kappa_{eff} = \frac{\pi\kappa^s}{4}\left[\int_0^\infty \frac{(y+1)\exp(\gamma x) + (y-1)\exp(-\gamma x)}{(y+1)\exp(\gamma x) - (y-1)\exp(-\gamma x)} \frac{J_1 \sin y}{\gamma^2 y}\, dy\right]^{-1} \quad 5.3$$

where $x = d/b$ and $\gamma = \kappa^s/\kappa_{sub}$ denotes the ratio between the TC of the layer (κ^s) and that of the substrate (κ_{sub}). Here, J_1 is the first-order Bessel function of the first kind.

The present investigation introduced a substrate-thickness factor (C-factor) as an extra ML input to quantitatively account for the influence of thin-film dimensions and substrate properties on TC measurements while developing ML model. To account for this, the input C was defined as follows:

$$C = \begin{cases} \left(\frac{b-d}{b}\right)\frac{\kappa_{sub}}{\kappa^n}, & d < b \\ 0, & d \geq b \end{cases} \quad 5.4$$

Here, d is the thin-film thickness, b is the threshold thickness, κ^n is quartz TC (used solely for normalization), and κ_{sub} represents the substrate's TC values. This mathematical formulation was developed as a piecewise function that distinguishes between two distinct thermal transport regimes based on film thickness. The critical thickness parameter b was established at 100 nm,

corresponding to the characteristic length scale of the SThM probe tip's interaction volume, which fundamentally limits the spatial resolution of thermal measurements. Below this threshold thickness, the factor incorporates both geometric scaling and material-dependent thermal properties through a normalized relationship. The geometric term $((b - d)/b)$ describes the progressive deviation from bulk-like behavior as film thickness decreases, while the material term $(\kappa_{sub}/\kappa^{n})$ weights this effect according to the relative thermal conductance of the substrate κ_{sub} compared to a reference value κ^{n}.

5.4. Primary Key Parameters

The methodological framework commences with the implementation of a spatial mapping protocol, wherein a 2 × 2 μm² region of the sample surface is systematically partitioned into a 16 × 16 measurement grid architecture. This scheme yields 256 discrete measurement cells, each encompassing a 125 × 125 nm² area. The grid-based analytical framework enables several advancements in thermal metrology. First, the methodology inherently compensates for measurement artifacts caused by localized surface anomalies through its integrated analysis of regional thermal interactions. Second, the comprehensive dataset generated by this approach provides the experimental foundation for developing more accurate computational models of submicron-scale heat transfer that properly account for microstructural influences.

The present methodology explicitly incorporates surface topography as a fundamental parameter governing thermal transport phenomena during SThM measurements. This recognition stems from the well-established understanding that submicron-scale surface morphology exerts profound influence on heat transfer mechanisms at the probe-sample interface, where nano-scale interactions and microscale geometric features collectively determine the efficiency of heat transport. Surface topography impacts thermal measurement accuracy by altering heat transfer, contact resistance, and contact area.

5.4.1. Micro Scale Topographical Parameters

As mentioned before, the present investigation employs a microstructural analysis framework focusing on a precisely delineated 2 × 2 μm² region, partitioned into a high-resolution 16 × 16 measurement grid. This designed microgrid architecture enables comprehensive characterization of surface morphology and its consequential effects on thermal transport phenomena at the sub-micron scale. The analytical protocol incorporates standard statistical surface metrology, with particular emphasis on two topographic parameters: Root Mean Square roughness (R^{s}_{rms}) and surface skewness (R^{s}_{sk}). R^{s}_{rms} represents a statistically rigorous measure of surface irregularity amplitude, calculated as the standard deviation of surface height deviations from the mean reference plane. This parameter provides insights into the degree of probe-sample interfacial contact in SThM measurements, where increased R^{s}_{rms} values directly correspond to greater surface asperity heights and consequently reduced effective thermal contact area. The physical manifestation of this relationship occurs through several interconnected mechanisms: (1) diminished true mechanical contact between probe tip and surface features due to asperity height variations, and (2) formation of submicron-scale air gaps acting as thermal insulation barriers. These effects contribute to elevated TCR and consequently compromised heat transfer efficiency. The RMS roughness is calculated as the square root of the arithmetic mean of squared deviations from the reference plane:

$$R_{rms}^{s} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_i^s)^2}, R_{rms}^{n} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_i^n)^2} \quad 5.5$$

Where z_i^s and z_i^n represent the residual height deviations at each measurement point i for sample and reference surfaces, respectively, obtained by subtracting the actual surface height from the theoretically ideal plane determined through least-squares regression analysis. The parameter N = 256 corresponds to the total number of discrete measurement cells comprising the high-resolution microgrid analysis area.

Surface skewness (R_{sk}^{s}) serves as a complementary statistical descriptor that quantifies the asymmetry of height distribution within the surface probability density function [60]. This parameter provides differentiation between surfaces dominated by protruding features (positive R_{sk}^{s}) versus those characterized by prevalent valleys or pores (negative R_{sk}^{s}). Positive skewness surfaces, typified by numerous sharp peaks, create discrete point contacts with the SThM probe that constrain thermal conduction pathways. Conversely, negatively skewed surfaces present distinct thermal transport challenges through the formation of enclosed air cavities that function as submicron-scale thermal insulators. These morphological features can induce complex three-dimensional heat flow patterns, including lateral spreading and localized thermal bottleneck effects, which modify the apparent TC measurements.

The surface skewness is computed as the normalized third moment of the height distribution:

$$R_{sk}^{s} = \frac{\frac{1}{N}\sum_{i=1}^{N}(z_i^s)^3}{(R_{rms}^{s})^3}, R_{sk}^{n} = \frac{\frac{1}{N}\sum_{i=1}^{N}(z_i^n)^3}{(R_{rms}^{n})^3} \quad 5.6$$

These statistical descriptors provide a comprehensive quantitative framework for characterizing surface morphology at microscale dimensions, enabling precise correlation between topographical features and their impact on thermal transport properties.

5.4.2. Submicron-scale Topographical Parameters

The present investigation extends its analytical framework to the submicron regime through implementation of a characterization protocol focusing on discrete 375 × 375 nm² regions, each systematically partitioned into a 3 × 3 measurement matrix referred to as a submicron grid. This approach enables examination of localized thermophysical phenomena by capturing submicron-scale parameters that govern interfacial heat transfer mechanisms, including surface inclination (M^{s_i}), standard deviation of surface heights (σ^{s_i}), and peak-to-valley variations (μ^{s_i}). These submicron descriptors provide essential complementary information to microscale descriptors (mainly, R_{rms}^{s} and R_{sk}^{s}).

Surface inclination (M^{s_i}) constitutes a submicron parameter that provides characterization of three-dimensional surface geometry by measuring the rate of height variation per unit lateral distance. A steeper inclination reduces the effective probe–sample contact area, whereas a smaller inclination increases it. Consequently, surface inclination obviously influences the thermal transport between the probe and the sample. The present methodology incorporates these

inclination-dependent thermal transport phenomena to develop accurate correction algorithms for submicron-scale SThM measurements. This parameter was calculated from the local slope of the surface, i.e., the change in height per unit lateral distance, obtained from AFM topography data within each submicron grid cell. The standard deviation of surface heights (σ^{S_i}) offers enhanced sensitivity to submicron-scale topographic variations when analyzed at this refined length scale [61]. As a statistical measure, it quantifies the statistical distribution of surface height deviations within each submicron cell. Surfaces with higher σ^{S_i} values exhibit lower contact area and thus lower heat transfer between the probe and the sample. The analytical approach developed in this study effectively addresses these challenges, enabling accurate interpretation of thermal transport behavior as submicron scale. Peak-to-valley variation (μ^{S_i}) represents another submicron-scale parameter essential for understanding surface characteristics that dominate probe–sample interactions [26]. This metric quantifies the maximum vertical displacement within each submicron cell. Large μ^{S_i} values create distinct thermal transport regimes and not let the probe sense the sample surface effectively, whereas small μ^{S_i} values contribute to enhancing heat transfer from the probe to the sample surface.

The principal objective of this investigation was to establish reliable quantitative correlations between the comprehensive set of topographical parameters previously characterized and the experimentally acquired thermal response signals, specifically the thermal signal ratio (Γ_i) and phase difference ($\Delta\varphi_i$). The thermal signal ratio (Γ_i) is a normalized, dimensionless parameter. It quantifies the interfacial thermal resistance at a specific submicron grid by comparing the resistance during probe-sample contact ($R_{th}^{S_i}$) to a probe-reference contact ($R_{th}^{n_i}$). These metrics serve as sensitive indicators for investigating thermal energy transport between probe and sample.

5.4.3. Complementary Considerations

The present study employs an ML framework specifically designed to mitigate the experimental errors inherent in submicron-scale thermal characterization, which arise from the complex interplay of multiple variables. The ML model will be trained on a high-dimensional dataset encompassing the full range of experimentally measured thermal and topographical parameters.

Prior to model training, the methodology incorporates a rigorous feature selection protocol based on Spearman's rank correlation analysis. This statistical preprocessing step serves several critical purposes: First, it quantitatively assesses the predictive relevance of each measured parameter with respect to target TC values. Second, it systematically identifies and eliminates variables with statistically insignificant correlations, thereby reducing input space dimensionality. Third, it retains only those parameters exhibiting strong physical relationships with the thermal transport phenomena under investigation.

The experimental methodology also includes a quartz-referenced standardization protocol to establish metrological traceability and ensure submicron-scale measurement accuracy. This approach leverages the exceptional thermophysical stability and well-documented TC properties of amorphous quartz ($\kappa = 1.5\ W \cdot m^{-1} \cdot K^{-1}$ at standard conditions [54]). This advanced correction framework addresses the critical challenge of substrate thermal contributions in thin-film measurements, which become increasingly significant for films thinner than 100 nm.

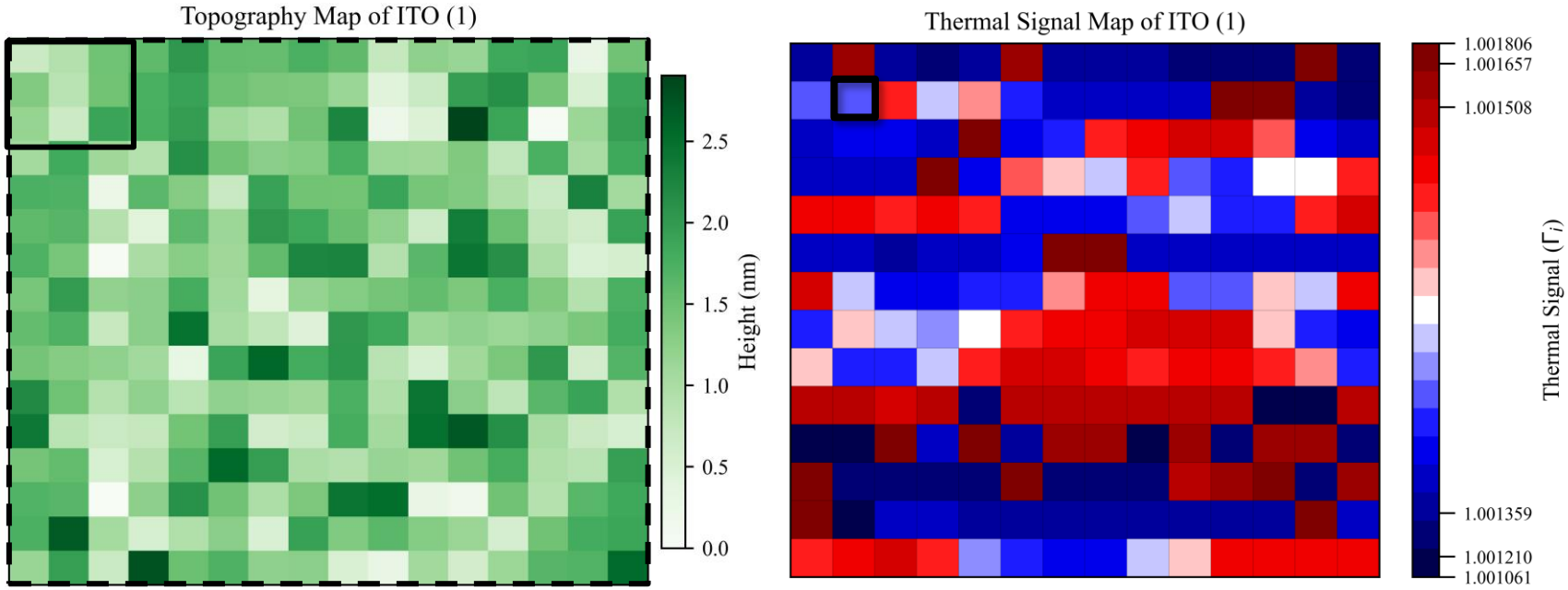
The research methodology presented in this study adopts an analytical framework that combines experimental instrumentation with ML techniques. The development of this integrated approach was informed by an extensive training process utilizing a dataset of 2,352 individual thermal

measurements. To ensure the generalizability of the ML models, an independent validation set comprising 980 additional measurements was employed, resulting in a comprehensive analytical dataset totalling 3,332 measurements. The dataset is presented in Table 5.2.

Table 5.2. Overview of the data collected and documented during laboratory experiments.

Variable*	Minimum	Maximum	Mean	StDev
Γ_i	0.997	1.005	1.001	0.002
$\Delta\varphi_i$ (deg)	-693.837	826.137	6.813	148.089
C	0	73.333	11.52	25.219
R^n_{rms}(nm)	0.326	2.523	0.664	0.480
R^n_{sk}	-1.55	0.507	-0.266	0.49
M^{n_i}	3.530×10^{-5}	5.583×10^{-2}	2.786×10^{-3}	3.246×10^{-3}
μ^{n_i}(nm)	0.258	16.7	1.395	1.324
$\kappa^n/\sigma^{n_i}(GW.m^{-2}.K^{-1})$	0.243	19.468	4.484	1.997
R^s_{rms}(nm)	0.265	29.700	6.075	9.209
R^s_{sk}	-0.127	3.072	0.693	0.81
M^{s_i}	5.619×10^{-5}	0.457	2.813×10^{-2}	5.516×10^{-2}
μ^{s_i}(nm)	0.189	142.200	13.220	22.540
$\chi(GW.m^{-2}.K^{-1})$	0.006	1681.611	32.650	108.729

* Variables include thermal signal ratio (Γ_i), phase difference ($\Delta\varphi_i$), and substrate-thickness factor (C). For the reference material (quartz), measurements include RMS roughness (R^n_{rms}), skewness (R^n_{sk}), inclination (M^{n_i}), peak-to-valley variations (μ^{n_i}), and the ratio κ^n/σ^{n_i}, where κ^nrepresents the TC of the quartz and σ^{n_i} denotes the standard deviation of surface heights for quartz. Corresponding parameters for the sample are also provided: Root Mean Squared (RMS) roughness (R^s_{rms}), skewness (R^s_{sk}), inclination (M^{s_i}), peak-to-valley variations (μ^{s_i}), and the ratio $\chi = \kappa^s/\sigma^{s_i}$ (target variable), where κ^srepresents the TC of the sample and σ^{s_i} denotes the standard deviation of surface heights. The total amount of collected data is 3,332.



$2 \times 2\ \mu m^2$ Square Area

$2 \times 2\ \mu m^2$ Square Area

Figure 5.2. The left image displays a topography map of sample ITO (1) captured on a microscale area (16 × 16 cells). The dotted square highlights the microscale region of $2 \times 2\ \mu m^2$ used for surface roughness analysis ($R^s_{rms}, R^s_{sk}, R^n_{rms}, R^n_{sk}$), while the smaller filled square represents a submicron-scale subset area of 375 × 375 nm² (3 × 3 cells) for localized topographical characterization ($M^{s_i}, \mu^{s_i}, \sigma^{s_i}, M^{n_i}, \mu^{n_i}, \sigma^{n_i}$). The image on the right shows the thermal signal map of the same material within the same surface region, with the small square indicating the specific cell where thermal signal ratios (Γ_i) were captured. After each measurement, the submicron grid shifts by one cell to the right, and the small square in the thermal map shifts accordingly to capture the next data point, ultimately yielding 196 data points per sample. It should also be noted that these maps represent the "trace" data. Additionally, "retrace" maps are collected, and the data from these two maps are averaged for each cell to reduce noise and enhance the reliability of the dataset.

The investigation centers on TC (κ^s) as the principal parameter of interest, which plays a fundamental role in the analytical framework through its mathematical relationship with the derived parameter χ. This quantity, defined as the ratio of sample TC to surface height standard deviation, at each submicron grid cell, ($\chi = \kappa^s/\sigma^{si}$), serves as a critical intermediary variable that facilitates interpretation of complex thermal transport phenomena. This analysis will explore the parameter's behavior across different material systems and measurement configurations, providing insights into thermal transport mechanisms at micro- and submicron scales.

Figure 5.2 constitutes an essential component of this study's analytical framework, presenting a dual-modality visualization that enables comprehensive characterization of the sample system. The left-hand image comprises a high-resolution topography map. This detailed topographical representation captures critical structural features including surface roughness, and morphological variations that may influence thermal transport phenomena. The corresponding right-hand image presents a thermal signal map acquired through SThM, providing quantitative measurements of heat transfer characteristics across the identical sample region. This thermal mapping exhibits variations in signal intensity that directly correlate with local differences in TC and heat dissipation properties. The visual comparison demonstrates relationships between specific topographical features and corresponding thermal signal variations, establishing empirical evidence for morphology-dependent heat transfer behavior. A complete set of maps for all samples is included in the appendix A.

5.5. Spearman's Correlation Analysis

The correlation coefficients in Table 5.3 provide useful information on the relationship between the target variable (χ) and each of the input variables. The Spearman's rank correlation coefficient (r_s) is a statistical measure that quantifies the strength and direction of both linear and non-linear relationships between two variables. Unlike Pearson's correlation, which specifically assesses linear relationships, Spearman's method evaluates the relationship based on the ranks of the data. Equation 5.7 shows the mathematical formula to compute r_s between two variables [62].

$$r_s = 1 - \frac{6\sum d_i^2}{N(N^2 - 1)} \tag{5.7}$$

Here, d_i represents the difference between the ranks of the corresponding values of the two variables, and $N$is the number of paired observations. The value of d_i reflects how much the two ranks disagree for a single data pair. A large $|\ d_i\ |$ means one variable's rank is high while the other's is low, whereas $d_i = 0$ indicates perfect agreement between the two ranks. A rank is simply the position of a value when all values in a list are arranged from smallest to largest; the smallest value receives rank 1, the next receives rank 2, and so on. To calculate d_i, the paired values (x_i, y_i) are first ranked separately. Each x_i is assigned a rank based on its position among all x-values, and each y_i is ranked in the same way among all y-values. For every pair, the difference between these two ranks is then computed as:

$$d_i = \text{Rank}(x_i) - \text{Rank}(y_i) \tag{5.8}$$

These $d_i$values are subsequently used in Spearman's correlation calculation.

An $r_s = +1$ indicates a perfect positive relationship, meaning that as one variable increases, the other variable consistently increases. Conversely, $r_s = -1$ signifies a perfect negative relationship, where one variable decreases as the other increases. An $r_s = 0$ suggests no relationship between the variables.

The Spearman's correlation analysis yields several significant findings regarding parameter-property relationships. First and foremost, the parameter thermal signal ratio (Γ_i) demonstrates the most inverse correlation with χ, as evidenced by its negative coefficient of -0.726, indicating that increases in Γ_i consistently correspond to marked decreases in χ across the studied material systems. This is followed in descending order of negative influence by peak-to-valley variation μ^{s_i} (-0.609), inclination M^{s_i} (-0.572), and substrate-thickness factor C (-0.518), all of which maintain statistically meaningful negative associations with the target variable.

Conversely, the examination reveals several parameters exhibiting minimal statistical association with χ. Specifically, $\Delta\varphi_i$ manifests an extremely weak negative correlation (-0.022), while R^n_{sk} shows a nearly negligible relationship (-0.005), suggesting these factors contribute insignificantly to χ variability within the parameter space under investigation. In contrast to the predominant negative correlations, certain parameters demonstrate positive, albeit modest, relationships with χ, including R^s_{sk} (0.228) and the ratio κ^n/σ^{n_i} (0.098), indicating that increases in these parameters correspond to slight enhancements in χ values.

A critical methodological insight emerges from the comparative analysis of normalized versus non-normalized parameter representations. The normalization procedure yields improvements in predictive capability, as illustrated by the transformation of $R_d^{s_i} - R_s^{s_i}$ from displaying a trivial negative correlation (-0.024) in its raw form to the influential Γ_i parameter (-0.726) following appropriate normalization.

Table 5.3. Analysis of the connection between the target variable (χ) and each input variable[*].

	Γ_i	$\Delta\varphi_i$	C	$R^n_{rms}(nm)$	R^n_{sk}	M^{n_i}	$\mu^{n_i}(nm)$	κ^n/σ^{n_i} $_{(GW.m^{-2}.K^{-1})}$	$R^s_{rms}(nm)$	R^s_{sk}	M^{s_i}	$\mu^{s_i}(nm)$
$\chi\ (GW.m^{-2}.K^{-1})$	-0.726	-0.022	-0.518	-0.270	-0.005	-0.116	-0.084	0.098	-0.505	0.228	-0.572	-0.609

[*]The parameters include thermal signal ratio (Γ_i), phase difference ($\Delta\varphi_i$), and substrate-thickness factor (C). For quartz, measurements include RMS roughness (R^n_{rms}), skewness (R^n_{sk}), inclination (M^{n_i}), peak-to-valley variations (μ^{n_i}), and the ratio κ^n/σ^{n_i}, where κ^n is TC and σ^{n_i} is surface height standard deviation. For the sample, parameters are RMS roughness (R^s_{rms}), skewness (R^s_{sk}), inclination (M^{s_i}), peak-to-valley variations (μ^{s_i}), and the target variable $\chi = \kappa^s/\sigma^{s_i}$, where κ^s is TC and σ^{s_i} is surface height standard deviation.

5.6. ML Framework for TC Prediction

Accurate prediction of TC remains a critical challenge in materials science, especially for complex systems where multiple physical factors interact. Traditional approaches often rely on oversimplified assumptions, limiting their predictive accuracy. To address this, I present a ML model that directly learns the relationship between thermal measurements and surface characteristics. Our data-driven approach captures both macroscopic trends and microscopic variations, combining advanced regression techniques with rigorous experimental validation to deliver reliable TC estimations across diverse material systems.

I developed a ML model using a regression-based framework to ensure precise and generalizable TC predictions. By integrating thermal, topographical, and material properties, while preserving the continuous nature of TC, our model overcomes the limitations of categorical classification methods. The model architecture was specifically designed to predict a continuous multiplicative variable, $\chi = \kappa^s/\sigma^{s_i}$, where κ^s represents the sample's TC and σ^{s_i} denotes the standard deviation of surface heights obtained from topographical measurements. This formulation transforms TC prediction into a continuous regression task, enabling the model to interpolate between discrete values and effectively generalize to unseen data. The core relationship is mathematically expressed through the model equation:

$$\chi = f_{ML}(\Gamma_i, C, \Delta\varphi_i, R^s_{rms}, R^s_{sk}, M^{s_i}, \mu^{s_i}, R^n_{rms}, R^n_{sk}, M^{n_i}, \mu^{n_i}, \kappa^n/\sigma^{n_i}) \quad 5.9$$

Where, Γ_i represents the thermal signal acquired through SThM and various topographical parameters including slope and roughness contribute to the indirect inference of σ^{s_i}, thereby enhancing the model's predictive performance. The selection of χ as the target variable was justified by reasonable correlations observed between input features and κ^s/σ^{s_i}, confirming the appropriateness of the regression framework. Furthermore, the identified relationships between σ^{s_i} and various topographical features, such as slope and peak-to-valley height, reinforced the model's physical interpretability. The alignment of model outputs with actual SThM probe measurements provided additional experimental validation, ensuring consistency between predictions and empirical observations.

5.7. Ensemble Regression Models in Materials Property Prediction

Before starting this section, it is important to note that some of the upcoming parts introduce new concepts related to ML. To keep the discussion understandable, the explanations are written in a clear basic way as possible. However, if additional background is needed, several easy-to-understand resources are available that provide clear explanations of the ML-related topics mentioned here, such as general ML concepts [63], gradient boosting algorithms [64], random forest algorithms [65], hyperparameter tuning [66], and cross validation in ML [67].

ML follows a sequence of well-defined stages. The process begins with data collection and preprocessing, where raw data is cleaned, organized, and prepared for analysis. Next, the dataset is divided into training and testing sets, ensuring that part of the data remains unseen until the final evaluation. The training stage allows model to learn patterns and relationships within dataset, while techniques such as cross-validation and regularization are applied to minimize errors and avoid overfitting. Once trained, the model enters the evaluation stage, where it is tested on the reserved data to measure its ability to generalize beyond the training examples. Performance is then

quantified using statistical metrics such as R^2, RMSE, MAE, and AARD%, which collectively indicate how well the predictions match experimental values. The mathematical expressions for these evaluation metrics are given as follows [53]:

$$R^2 = 1 - \frac{\sum_{m=1}^{M} \left(y_m^t - y_m^p\right)^2}{\sum_{m=1}^{M} \left(y_m^t - y^{\text{ave}}\right)^2}, \text{where } y^{\text{ave}} = \frac{1}{M}\sum_{m=1}^{M} y_m^t \quad 5.10$$

$$\text{AARD (\%)} = \frac{100}{M}\sum_{m=1}^{M} \left| \frac{y_m^t - y_m^p}{y_m^e} \right| \quad 5.11$$

$$\text{RMSE} = \left(\frac{\sum_{m=1}^{M} \left(y_m^t - y_m^p\right)^2}{M} \right)^{0.5} \quad 5.12$$

$$\text{MAE} = \frac{1}{M}\sum_{m=1}^{M} \left| y_m^t - y_m^p \right| \quad 5.13$$

In these equations, y_m^t and y_m^p denote the true and predicted values of the target variable, respectively. The term y^{ave} represents the average of the true values. The parameter M corresponds to the total number of data points considered during the cross-validation or testing stages.

Among various ML models, Random Forest regression and Gradient Boosting regression have proven effective, demonstrating remarkable accuracy in capturing complex, nonlinear relationships within material datasets [68, 69, 70, 71]. Random Forest regression operates as an ensemble learning method that constructs multiple decision trees, each trained on a bootstrapped subset of the data. By aggregating the predictions of these trees, the model achieves enhanced accuracy while mitigating overfitting. A key strength of Random Forest lies in its ability to model nonlinear relationships between input variables and material properties without relying on predefined mathematical formulations. This dynamic feature selection makes the method adaptable to diverse materials and testing conditions [68, 69]. Figure 5-3 illustrates a predictive Random Forest regression model. This model constructs multiple decision trees during training (in Figure 5-3 these decision trees are shown by Tree 1, Tree 2, …). Each tree is trained on a random subset of the data and features (shown by filled circles the Figure 5-3), which introduces variability and de-correlates the individual learners. The final prediction is made by averaging their results.

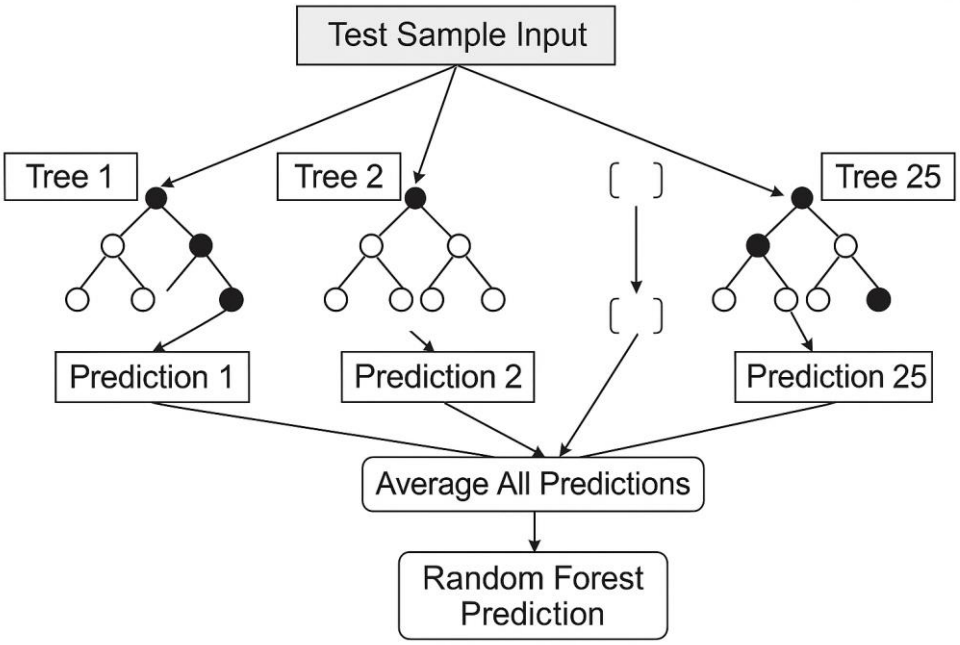


Figure 5-3. Schematic of a Random Forest Regression Model. The model constructs multiple decision trees, each trained on a random subset of the data and features. The final prediction (output) is obtained by averaging the predictions of all individual trees [72].

The reliability of Random Forest extends to its handling of noisy data and high-dimensional datasets, maintaining stable performance where traditional methods might falter. A critical parameter influencing model performance is the minimum number of samples required for node splitting, which directly affects the model's complexity. Further improvements in predictive accuracy can be achieved through Bayesian optimization, which systematically fine-tunes hyperparameters such as the number of decision trees and maximum tree depth.

In contrast to Random Forest's parallel tree construction, Gradient Boosting (Figure 5-4) employs a sequential ensemble approach where each new decision tree is trained to correct the residual errors of its predecessors. This iterative refinement process, guided by gradient descent optimization of the loss function, enables progressively higher predictive accuracy. The method's emphasis on error correction makes it well-suited for high-precision applications in material science where minor property differences can have significant implications [70, 71]. As shown in Figure 5-4, the model is built by training a sequence of decision trees. The first tree learns to predict the target values using the input data (X, y), and its errors (called residuals) are computed as $r_1 = y - \hat{y}_1$. Each subsequent tree is trained on these residuals, so it focuses on the parts of the data that previous trees could not predict accurately. Every new tree produces its own prediction of the residuals, denoted by $\hat{r}$, and the next residual is updated by subtracting this predicted value (for example, $r_2 = r_1 - \hat{r}_1$). Through this iterative correction, the model gradually improves its overall prediction. In the diagram, X represents the input features, y is the true target, $\hat{y}$ refers to the predicted output of each tree, r refers to the residuals, and $\hat{r}$ represents the predicted residuals used to update the next step.

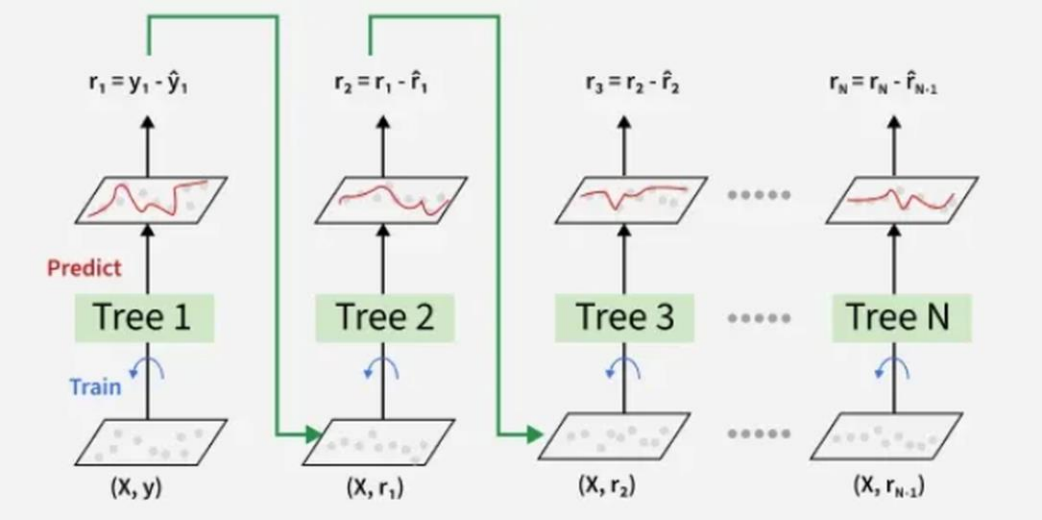


Figure 5-4. Schematic of a Gradient Boosting Model [64].

While Gradient Boosting often achieves superior accuracy compared to Random Forest in complex prediction tasks, this performance comes with certain trade-offs. The method demands careful hyperparameter tuning, including learning rate, tree depth, and number of iterations, to prevent overfitting. Additionally, the sequential nature of model training results in higher computational requirements, making it less suitable for extremely large datasets where computational efficiency is paramount.

To enhance model reliability and prevent overfitting, practitioners employ k-fold cross-validation as a reliable validation technique. This method involves dividing the training data into k equally sized subsets, commonly using 5 or 10 folds. During each validation cycle, the model is trained on k-1 folds while using the remaining fold for validation. This rotation continues systematically until each fold has served as the validation set exactly once. The performance metrics from all iterations are then aggregated and averaged, producing a comprehensive evaluation that minimizes bias from any particular data distribution.

Following this validation phase and subsequent hyperparameter optimization, the model undergoes final assessment using the previously untouched test set. This critical evaluation stage provides the most realistic measure of the model's predictive capability, simulating its performance in real-world applications where it must process completely new data.

Model performance is quantified using multiple complementary statistical metrics that collectively provide a complete picture of predictive accuracy. The R-squared (R^2) value indicates the proportion of variance in the dependent variable that can be explained by the model's predictors. The AARD% offers insight into the average percentage difference between predicted and actual values. RMSE emphasizes larger errors through its squared difference calculation, while MAE provides a straightforward average of absolute prediction errors.

This evaluation framework combines systematic data partitioning, thorough cross-validation, and multi-metric assessment to create a reliable model development process.

6. Results and Discussion (Thermal)

This section provides an evaluation of ML models for predicting TC, with a focus on hyperparameter optimization, model selection, and performance validation.

6.1. Hyperparameter Tuning for Gradient Boosting and Random Forest: Model Selection and Validation

Hyperparameter tuning represents a crucial phase in optimizing ML models, ensuring they achieve maximum performance while maintaining strong generalization capabilities. This study presents an evaluation of Gradient Boosting and Random Forest models through an extensive grid search approach combined with cross-validation, with the objective of identifying optimal hyperparameter configurations that maximize predictive accuracy.

The research employed a comprehensive grid search strategy to explore predefined hyperparameter spaces for both Gradient Boosting and Random Forest models (Table 6.1). The experimental design resulted in the training and evaluation of 279 distinct models (each model has different set of hyperparameters).

Performance assessment incorporated cross-validation techniques to mitigate overfitting risks and provide reliable measurements of model generalization capabilities. This rigorous validation approach ensured that performance metrics accurately reflected each model's true predictive potential.

The comparative evaluation revealed that the Random Forest model consistently outperformed the Gradient Boosting approach in terms of predictive accuracy, with this conclusion being validated through multiple statistical performance metrics. The study documented the final optimized hyperparameters for both algorithms in detailed comparison in Table 6.2. The Random Forest algorithm's ensemble architecture relies on the coordinated operation of multiple decision trees, with several critical hyperparameters governing model complexity and predictive performance. The number of estimators parameter determines the total quantity of trees comprising the forest ensemble, where increased values typically enhance model accuracy at the cost of greater computational requirements. Maximum depth serves as another parameter, controlling individual tree complexity and directly influencing the balance between model accuracy and overfitting risk.

The study implemented several mechanisms specifically designed to prevent overfitting, including minimum samples split which dictates the threshold for node partitioning, minimum samples leaf governing terminal node size requirements, and maximum features which constrains the feature selection process during splitting operations. The criterion parameter plays a pivotal role in determining split quality metrics, with regression tasks employing squared error, absolute error, or the specialized Friedman Mean Squared Error (MSE) that simultaneously optimizes for both mean and variance. For count data modeling, the Poisson criterion provided tailored performance by specifically accommodating discrete outcome distributions. The bootstrap parameter further enhanced model reliability by introducing controlled randomness through data sampling with replacement.

The Gradient Boosting algorithm's sequential error-correction architecture incorporates distinct hyperparameters that collectively determine model performance. Similar to Random Forest, the number of estimators and maximum depth parameters significantly influence model accuracy and complexity characteristics. The learning rate parameter assumes particular importance in Gradient

Boosting, modulating each tree's contribution to the final model and requiring careful calibration to balance training efficiency with generalization performance.

The subsample parameter introduces stochastic elements by specifying the fraction of training data utilized for each boosting iteration, while various loss functions accommodate different regression scenarios. These include standard squared error for conventional regression, Huber loss for outlier-resistant modeling, quantile loss for percentile prediction, and absolute error for median-focused applications. The alpha parameter provides additional regularization control, serving as an important tool for managing model complexity and preventing overfitting.

Table 6.1. The hyperparameters of the ML models that were explored and analyzed in this study.

Model's name	Checked hyperparameters	Investigated range	Number of designed models
Random Forest	Number of estimators	27-4975	141
	Maximum depth	2-998	
	Minimum samples split	2-198	
	Minimum samples leaf	1-199	
	Maximum features	0.11287-0.99951	
	Criterion	Squared error, absolute error, Friedman MSE, Poisson	
	Bootstrap	True - False	
Gradient Boosting	Number of estimators	22-4960	138
	Maximum depth	2-993	
	Minimum samples split	7-200	
	Minimum samples leaf	3-200	
	Maximum features	0.10153-0.99678	
	Learning rate	0.01003-0.98148	
	Subsample	0.12918-0.99875	
	Loss function	Squared error, Huber loss, quantile, absolute error	
	Alpha (Regularization Term)	0.10124-0.99320	

Note. Hyperparameter definitions are as follows: Number of estimators refers to the total count of decision trees or boosting stages; increasing this value generally improves performance at the cost of computational load. Maximum depth limits the number of nodes in a tree, controlling complexity and the risk of overfitting. Minimum samples split and Minimum samples leaf dictate the smallest sample size required to split a node or form a leaf, respectively, acting as constraints on tree growth. Maximum features determine the size of the random feature subset considered at each split. Criterion (Random Forest) specifies the function to measure split quality. Bootstrap indicates whether trees are trained on random subsamples with replacement. For Gradient Boosting, learning rate scales the contribution of each tree to the final model, while Subsample controls the fraction of data used per boosting step. Alpha serves as a regularization term to penalize complex models, and the Loss function defines the specific error metric optimized during training. Max depth, Min samples split/leaf, and Max features function similarly to Random Forest.

Table 6.2. The optimal hyperparameters selected for each ML model and their associated prediction accuracy.

Model's name	Tuned hyperparameters	Uncertainty index	Cross-validation training phase	Cross-validation testing phase
Random Forest 'a'	Number of estimators = 1986 Maximum depth = 923 Minimum samples split = 18 Minimum samples leaf = 1 Maximum features: 0.866 Criterion: squared error Bootstrap = False	RMSE	14.353	23.986
		R^2	0.98254	0.95205
		AARD%	6.995	14.276
		MAE	2.545	4.598
Random Forest 'b'	Number of estimators = 2624 Maximum depth = 359 Minimum samples split = 6 Minimum samples leaf = 14 Maximum features: 0.961 Criterion: absolute error Bootstrap = False	RMSE	24.445	32.770
		R^2	0.94747	0.92009
		AARD%	13.738	17.545
		MAE	3.720	5.519
Random Forest 'c'	Number of estimators = 2306 Maximum depth = 419 Minimum samples split = 6 Minimum samples leaf = 15 Maximum features: 0.952 Criterion: absolute error Bootstrap = False	RMSE	25.528	30.673
		R^2	0.94268	0.92791
		AARD%	15.477	18.304
		MAE	4.153	5.599
Gradient Boosting	Number of estimators = 2559 Maximum depth = 95 Minimum samples split = 92 Minimum samples leaf = 181 Maximum features: 0.923 Learning rate = 0.013 Subsample = 0.846 Loss function: Quantile Alpha = 0.294	RMSE	103.205	82.509
		R^2	0.17953	0.23259
		AARD%	23.396	26.891
		MAE	23.107	19.444

The study confirms that grid search combined with cross-validation constitutes an effective methodology for hyperparameter optimization, with Random Forest's ensemble structure proving reliable for the given application. Gradient Boosting's flexible architecture, particularly through its adaptable learning rate and specialized loss functions, maintains strong potential for regression tasks requiring specific performance characteristics.

In Table 6.2, Random Forest models 'a', 'b', and 'c' represent different configurations of the same algorithm. While the base method is identical, each model was trained using a unique set of hyperparameters (such as number of estimators, maximum depth, etc.) to identify the most

accurate result. In this study, these specific sets are referred to as Model ‘a’, b’, or ‘c’, respectively. As shown in Table 6.2, the Random Forest ‘a’ model demonstrates superior performance across all key evaluation metrics, establishing itself as the most effective predictive model for this task. During both cross-validation training and testing phases, it maintains consistently strong results that highlight its accuracy and generalization capabilities.

In terms of predictive precision, the model achieves the lowest RMSE values of 14.353 (training) and 23.986 (testing), indicating accurate predictions with minimal error. The high R^2 values of 0.98254 (training) and 0.95205 (testing) confirm its ability to explain over 95% of the variance in the testing dataset, showcasing excellent explanatory power. Additional validation comes from the low AARD% values of 6.995% (training) and 14.276% (testing), demonstrating that predictions remain close to actual values, while the low MAE scores of 2.545 (training) and 4.598 (testing) further reinforce its ability to minimize average prediction errors. Crucially, the small discrepancy between training and testing performance confirms the model generalizes effectively without overfitting.

When compared to alternative Random Forest models ('b' and 'c'), Random Forest 'a' shows clear advantages. It is noted that the distinction between these models lies solely in their different sets of hyperparameters. The competing models exhibit higher RMSE and lower R^2 values, indicating reduced predictive accuracy and weaker generalization. Their higher AARD% and MAE values also reveal more frequent and significant prediction errors, making them less reliable choices.

The performance gap becomes even more pronounced when comparing with the Gradient Boosting model, which shows weaker results. With an extremely high testing RMSE of 82.509 and a very low R^2 of 0.23259, it explains only about 23% of the variance. The model's high AARD% (26.891) and MAE (19.444) confirm its predictions frequently deviate from actual values, rendering it unsuitable for precise applications.

Given these comprehensive evaluations, Random Forest 'a' emerges as the optimal choice, offering unmatched predictive accuracy, minimal errors, and strong generalization. The alternative models, while functional, cannot match its performance, making Random Forest 'a' the clear recommendation for reliable and accurate predictions in this context.

Another Random Forest model (based on the tuned hyperparameters found for Model 'a', shown in Table 6.2) was trained on a combined dataset consisting of merged training and testing subsets (2352 datapoints) from cross-validation to maximize the available learning data for the final training stage.

Table 6.3. Evaluation results of the fine-tuned model ‘a’ across training and testing phases.

Metric	Training data	Testing data
RMSE	16.033	19.222
R^2	0.97886	0.96630
AARD%	6.319	17.487
MAE	2.511	4.448

Validation results (Table 6.3) demonstrate the model's predictive accuracy. To evaluate its performance under realistic conditions, the model was validated on 980 previously unseen experimental data points. This validation process confirmed the model's ability to generalize well

to new data, demonstrating its reliability for practical applications. A detailed performance evaluation revealed a linear correlation between the actual and predicted χ values, as illustrated in Figure 6.1. The training data closely followed the perfect prediction line, indicating an excellent fit, while the testing data exhibited slightly higher variability but still maintained strong alignment with the expected trend. Minor deviations were observed at higher χ values (exceeding approximately 870 $GW.m^{-2}.K^{-1}$). This is likely due to the model's lack of sensitivity to changes large TC values.

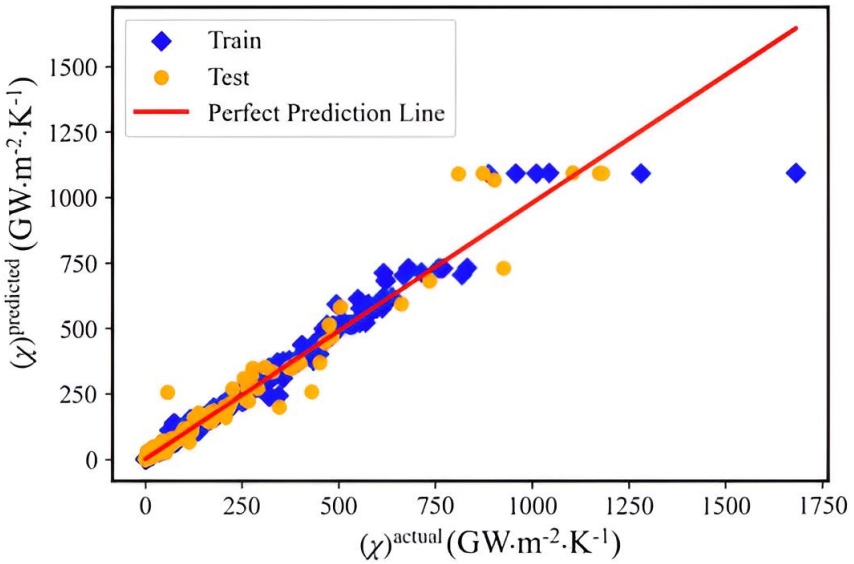


Figure 6.1. Relationship between observed dependent variable values and predicted outputs from the model 'a' for training and testing datasets.

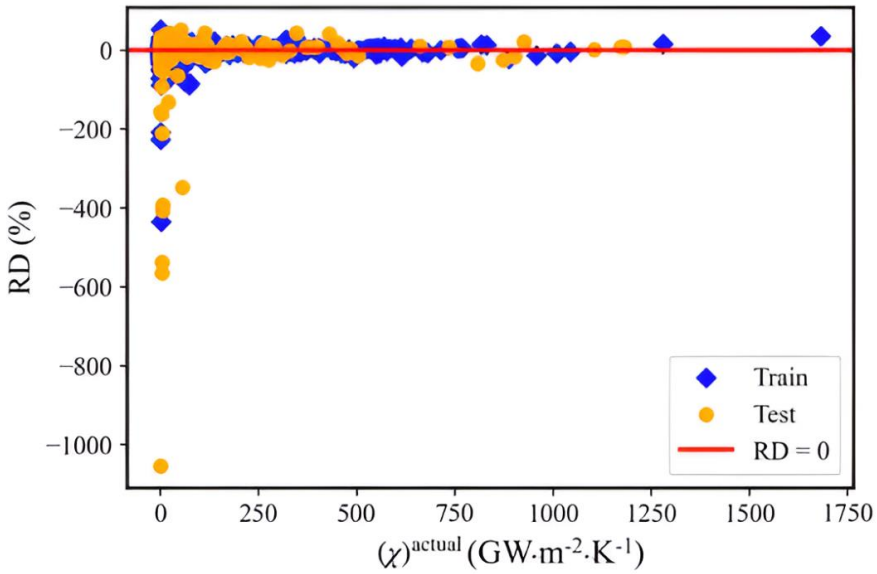


Figure 6.2. The RD% between actual and predicted values of the target variable.

Despite these deviations, the relative errors remained small, as shown in Figure 6.2, with minimal practical impact since TC values are typically averaged in real-world applications.

Further analysis of the model's performance was conducted using the Relative Deviation (RD%), defined as:

$$\text{RD\%} = \frac{\chi^{actual} - \chi^{predicted}}{\chi^{actual}} \quad 6.1$$

The RD% values shown in Figure 6.2 demonstrated that most RD% values clustered near zero, confirming low bias in the model's predictions. Slightly greater deviations occurred at lower χ values (below approximately 250 $GW.m^{-2}.K^{-1}$), though these had negligible influence on the overall results. The practice of averaging TC values further reduced the impact of outliers, enhancing the reliability of the findings.

The residual distribution, presented in Figure 6.3, provided additional insights into the model's performance. For the training data, the residuals had a mean of zero, indicating no systematic bias, and a standard deviation of 16.04 $GW.m^{-2}.K^{-1}$, reflecting a tight distribution and high predictive accuracy. The testing residuals showed a mean of 0.77 $GW.m^{-2}.K^{-1}$, suggesting a slight positive bias, which was deemed acceptable given the complexity of the data, along with a standard deviation of 19.22 $GW.m^{-2}.K^{-1}$, indicating a marginally wider but still consistent spread.

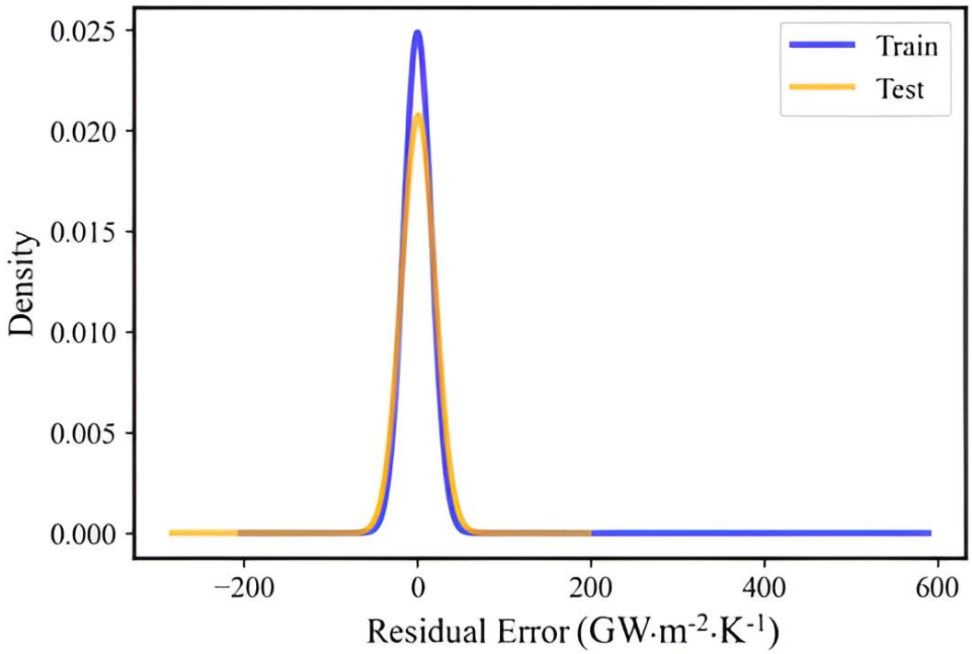


Figure 6.3. Gaussian distribution of residuals for training and testing datasets

Key observations from the residual analysis included the sharp peak and narrow spread of the training residuals, confirming a strong model fit, and the broader yet similarly dense distribution of the testing residuals, which highlighted the model's generalization capability. Overall, the Random Forest model exhibited high predictive accuracy with minimal bias, effectively capturing TC trends across a wide range of values. While minor deviations occurred at extreme χ values, the model's performance remained strong, ensuring its reliability for practical use. The low relative errors and tight residual distributions further validated its reliability, providing confidence in its predictive accuracy.

6.2. Evaluation of Model 'a' for TC Predictions

The developed predictive model demonstrates reliable capabilities in estimating TC values across a diverse range of materials. Using $\chi = \kappa^{s}/\sigma^{si}$ to calculate TC from predicted χ values, the model was rigorously evaluated on 17 distinct samples. The predictions were derived through an averaging process of multiple test-phase outputs, enhancing statistical reliability and minimizing random error. This methodological approach ensures the reported values represent stable, reproducible estimates rather than single-point determinations. The TC values for the various samples are summarized in Table 6.4 and are shown visually in Figure 6.4.

Table 6.4. A comparison of the actual and predicted κ values for the samples.

Sample's number	Sample	Actual $\kappa\ (W \cdot m^{-1} \cdot K^{-1})$	Predicted $\kappa\ (W \cdot m^{-1} \cdot K^{-1})$
1	ITO (1)	6.4 [54]	5.3
2	ITO (2)	3.5 [55]	3.7
3	ITO (3)	8.3 [55]	8.5
4	ITO (4)	10.6 [54]	8.5
5	ITO (5)	11.8 [55]	12.4
6	ITO (6)	6.7 [54]	10
7	Glass (Bulk)	1.1	1.1
8	Glassy carbon (Bulk)	6.3	6.7
9	SiC (Bulk)	450	464
10	YAG (Bulk)	12	12
11	ZnAlO (1)	4.29 [56]	4.43
12	ZnO (Bulk)	80	88
13	ZnO (2)	0.25 [57]	0.33
14	ZnO (3)	0.28 [57]	0.29
15	ZnO (4)	1.12 [57]	1.12
16	ZnO (5)	2.81 [57]	2.84
17	PMMA (Bulk)	0.17	0.18

	ITO (1)	ITO (2)	ITO (3)	ITO (4)	ITO (5)	ITO (6)	Glass (Bulk)	Glassy carbon (Bulk)	SiC (Bulk)	YAG (Bulk)	ZnAlO (1)	ZnO (Bulk)	ZnO (2)	ZnO (3)	ZnO (4)	ZnO (5)	PMMA (Bulk)
Actual	1.86	1.25	2.12	2.36	2.47	1.90	0.10	1.84	6.11	2.48	1.46	4.38	-1.39	-1.27	0.11	1.03	-1.77
ML	1.67	1.31	2.14	2.14	2.52	2.30	0.10	1.90	6.14	2.48	1.49	4.48	-1.11	-1.24	0.11	1.04	-1.71

Figure 6.4. Comparison of actual and ML (predicted) thermal conductivity values expressed as Ln(κ) for different materials, where κ is given in $W \cdot m^{-1} \cdot K^{-1}$.

The model exhibits strong predictive accuracy across several orders of magnitude in TC values. For low-TC materials, it achieves particularly high precision, as demonstrated by Sample 17, where the actual value of 0.17 $W \cdot m^{-1} \cdot K^{-1}$ compares closely with the predicted 0.18 $W \cdot m^{-1} \cdot K^{-1}$, corresponding to a relative error of approximately 5.9 %. Similarly, Sample 7 shows excellent agreement, with both the actual and predicted values equal to 1.1 $W \cdot m^{-1} \cdot K^{-1}$, resulting in a negligible error below 1 %.

The model performs consistently well in the intermediate TC range. Sample 10 exhibits virtually perfect agreement between the actual and predicted values (both 12 $W \cdot m^{-1} \cdot K^{-1}$), indicating an error well below 1 %. These results highlight the model's robustness for materials with low to moderate thermal conductivities, which constitute the majority of the training dataset.

At the high-TC end, the model maintains reasonable predictive capability, although with slightly increased relative deviations. Sample 9 (SiC) shows an overestimation from 450 $W \cdot m^{-1} \cdot K^{-1}$ to 464 $W \cdot m^{-1} \cdot K^{-1}$, corresponding to a relative error of approximately 3.1 %. Sample 12 (ZnO bulk) exhibits a larger deviation, with a predicted value of 88 $W \cdot m^{-1} \cdot K^{-1}$ compared to the actual 80 $W \cdot m^{-1} \cdot K^{-1}$, yielding a relative error of 10.0 %. While higher than in lower-TC regimes, these deviations remain acceptable for many practical applications.

Overall, the model's performance indicates that the averaging approach effectively reduces prediction variance, as reflected in the close agreement between predicted and reference values across a wide TC spectrum. The strong correlation observed confirms that the model has successfully captured the underlying physical relationships governing thermal transport, although reduced precision is evident at higher TC values explored in this study.

From a practical perspective, the model is most reliable for materials with TC values below 100 $W \cdot m^{-1} \cdot K^{-1}$, where relative errors remain consistently low. Predictions for ultra-low TC materials (<1 $W \cdot m^{-1} \cdot K^{-1}$) demonstrate particularly good accuracy, and the averaging process provides an inherent estimate of uncertainty through prediction dispersion. This further confirms that the thermal probes employed in the experiments are especially sensitive within the low-TC regime.

This analysis confirms the model's suitability as a practical tool for TC estimation while clearly identifying areas for future improvement. The current implementation already offers significant value for materials screening and preliminary characterization, particularly for low-to-medium conductivity materials, with further enhancements expected through targeted model refinements.

7. Methodology and Experimental Procedures (Electrical)

This section details the experimental framework developed to characterize the EC of thin film and metallic samples. This study employs a high-precision 4PP setup integrated with a ML refinement layer. By systematically varying probe positions and sample orientations, a dataset was generated to train a model capable of compensating for real-world experimental complexities.

7.1. Materials Used and Instrumentation

The electrical measurements were performed using a Keithley Model 6221 current source for ITO samples, a Keithley Model 2231A-30-3 current source for metallic samples, and a Keithley Model 2182A nanovoltmeter to measure the resultant voltage for both sample types. Table 5.1

summarizes the samples used to assess the proposed methodology, with each characterized by its actual EC and thickness for performance evaluation. Actual EC measurements for ITO samples (a) to (d) were conducted via the standard vdP method in prior studies, as referenced in Table 5.1. An important consideration is that all ITO samples were annealed at 400 °C, and the effect of annealing on ITO thin-film thickness must be addressed. Annealing can modify grain size and crystallinity in polycrystalline ITO, but X-ray diffraction (XRD) results revealed only slight grain growth [55]. The reference sample exhibited an average grain size of approximately 56 nm, which rose to about 57 nm following vacuum annealing and around 63 nm after N_2 annealing [55]. These minimal variations have little impact on film thickness, which stayed consistent at roughly 170 nm. The metallic samples listed in the Table 5.1 were obtained from Sigma-Aldrich and have a high purity of 99.9+%. The Table 5.1 also includes sample dimensions (L_1, L_2, L_3, and L_4) and diagonal length ($L_{diag.}$). The study employed two sample types: ITO films with irregular quadrilateral shapes deposited on insulating glass substrates and metallic samples with more regular geometries. For metallic samples, the probe measures only one corner, whereas for ITO samples, it measures all four corners to evaluate spatial variations. This approach enables a direct comparison between ML-predicted values when the model analyzes a single region of the samples versus when it assesses all regions of the samples. This study is based on 553 measurements of different electrical signals, taken with different sample dimensions, probe positions, and sample rotations, summarized in Table 7-2, which highlights the minimum, average, and maximum values for each variable. It should be noted that the sample rotations were performed to ensure that measurements were taken under different edge-effect conditions, as the thin film samples have irregular geometries. The current dataset is sufficient for developing a functional and generalizable model. Notably, the successful development of a generalized model using this dataset highlights its potential for scalability to larger datasets, enabling the development of more comprehensive models, a key strength of this work.

In this study the input variables include sample dimensions (L_1, L_2, L_3, and L_4), diagonal ($L_{diag.}$), the horizontal and vertical probe positions (L_h and $L_{v,}$ respectively), and applied current (I). These parameters collectively characterize the system's geometrical and electrical properties, providing a foundation for analyzing its behavior.

Table 7-1. An overview of the samples utilized to train ML model.

Sample*	Thickness	$L_{diag.}$(mm)	L_1(mm)	L_2(mm)	L_3(mm)	L_4(mm)	σ_{actual} (kS·cm^{-1})
ITO (a)	170 nm	13.517	9.300	9.807	9.736	12.095	4.99 [55]
ITO (b)	170 nm	11.579	9.127	10.096	8.835	7.813	5.53 [55]
ITO (c)	170 nm	13.828	12.555	6.287	12.654	8.379	2.73 [55]
ITO (d)	170 nm	10.295	8.370	6.093	8.930	5.835	4.76 [55]
Cu	0.25 mm	14.796	10.423	10.159	10.662	9.905	598
W	0.25 mm	12.436	9.572	8.004	9.580	8.357	204
Ni	0.50 mm	15.281	11.633	9.091	12.361	9.059	143
Fe	0.25 mm	15.238	12.479	9.059	12.485	8.706	103
Sn	0.50 mm	13.823	10.543	8.597	10.421	8.956	91

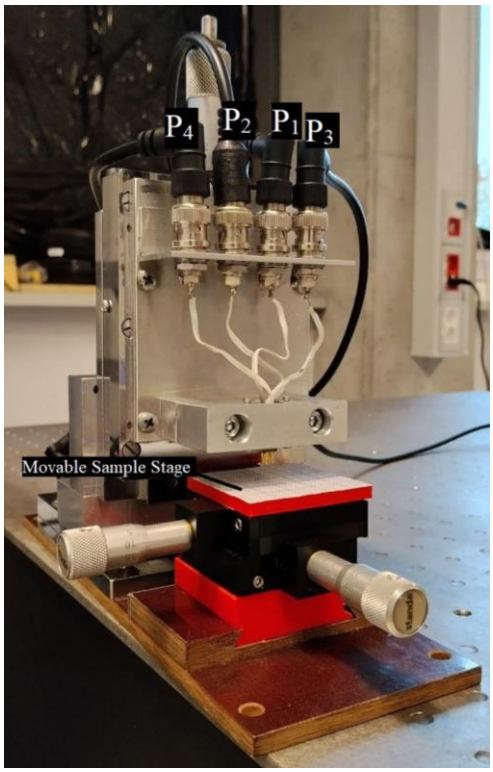
* ITO samples (a) to (d) were annealed in different atmospheres: carbon dioxide, a nitrogen-hydrogen mixture, vacuum, and nitrogen. All samples were deposited on glass substrates. The EC of these ITO samples (a) to (d) was previously measured using the standard vdP method, as reported in the referenced studies.

Table 7-2. Summary of the recorded information in the laboratory. Statistical summary of experimental measurements. Data represents 553 observations per parameter.

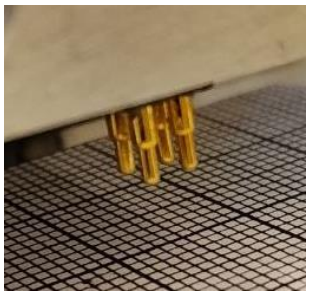
Variable	Minimum	Maximum	Mean
I (mA)	0.010	700.000	103.828
L_v(mm)	1	3	1.886
L_h(mm)	1	3	1.937
$L_{diag.}$(mm)	10.295	15.356	13.241
L_1(mm)	5.835	12.654	9.487
L_2(mm)	5.835	12.654	9.044
L_3(mm)	5.835	12.654	9.673
L_4(mm)	5.835	12.654	9.028
Θ (S·mV)	0.009	1526.913	164.045

7.2. Measurement Methodology

The measurements were acquired via a 4-probe device featuring a 4-pin, 2-row, straight male header (RS PRO) with a standard 2.54 mm pitch. The 4PP technique was employed to measure the electrical conductivity of thin-film samples. The setup, illustrated in Figure 7-1, was completely built during the course of this PhD project to provide a reliable and reproducible platform for thin-film electrical characterization. The samples used in these measurements are shown in Table 7-3.



(a)



(b)

Figure 7-1. setup for electrical conductivity measurements of samples. (a) Experimental arrangement consisting of a movable sample stage for precise sample positioning and a probe head. (b) Close-up of the probe assembly with four metallic pins arranged in a square configuration with equal spacing. In this setup, the two closer pins are used to inject current, while the two farther pins measure the resulting voltage drop across the sample. This configuration minimizes contact resistance effects, which enables more accurate evaluation of intrinsic electrical conductivity.

Prior to measurements, all probe contacts were meticulously cleaned with acetone to remove any flux residue, dust, or organic contaminants from the gold-plated pins. This protocol minimized systematic errors and improved dataset reliability. A four-probe configuration, arranged in a rectangular layout, was used to minimize contact resistance effects. Two probes supplied a dc current (P_3 and P_4), while the other two probes (P_1 and P_2) measured the resultant dc voltage, as shown in Figure 7-2. The line connecting P_1 and P_2 aligned parallel to side L_1 side during all measurements (Figure 7-2). Since the samples are quadrangle, all four side dimensions (L_1, L_2, L_3, and L_4) and the diagonal ($L_{diag.}$) crossing vertex C_1 were incorporated into the analysis. This ensured a well-defined shape for each sample. To accurately model the electrical field distribution, the horizontal (L_h) and vertical (L_v) distances from the sample's top-left corner (vertex C_1) to probe P_1 were recorded. These measurements provided positional probe data for the ML model.

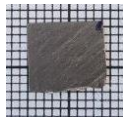
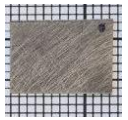
To evaluate the model under realistic conditions, the probes were deliberately positioned near the sample edges, where electric field distortion causes the largest EC measurement errors. This high-interference setup tests the model's reliability in practical situations where edge effects cannot be avoided. The measurement procedure involved positioning four electrical probes at different

locations (A_1, A_2, A_3, A_4) around vertex C_1 of the sample, as shown in Figure 7-2. For A_1, L_h=1 mm and L_v=1 mm; for A_2, L_h=3 mm and L_v=1 mm; for A_3, L_h=3 mm and L_v=3 mm; and for A_4, L_h=1 mm and L_v=3 mm. In cases where 3 mm probe placement extended beyond the sample surface, the distance was reduced to 2 mm to keep the probes within the sample boundaries. This positioning strategy minimized sensitivity to minor placement variations. Initial measurements were conducted with side L_1 aligned parallel to the P_1-P_2 probe axis. Following the first measurement set, the sample underwent counterclockwise rotation (not the probes), bringing side L_2 into alignment with the P_1-P_2 axis. This rotated configuration designated L_2 as the new reference edge (L_1) for modeling purposes. The identical measurement protocol was subsequently repeated at all four locations (shown in Figure 7-2). This rotational sequence continued systematically, with the sample rotated to align sides L_3 and L_4 sequentially with the original P_1-P_2 probe axis. Complete measurements were performed at all four locations following each rotational adjustment. For metallic samples, measurements were conducted only once at all four probe locations without any sample rotation. In contrast, the full rotational measurement protocol was exclusively applied to irregular ITO samples. This rotational measurement methodology serves a primary objective: enabling the characterization of different edge effects on electrical measurements in irregular samples.

In this work, a machine learning approach is developed to learn directly from experimental datasets and predict electrical conductivity values across a wide range of measurement conditions. To ensure physical realism, the methodology integrates theory-informed constraints that preserve consistency with fundamental principles, while leveraging data-driven optimization to address empirical complexities. The approach is systematically validated on both metallic and transparent conductive oxide (e.g., ITO) samples, with particular attention to challenging anisotropic conditions and edge effects.

Furthermore, the data-driven model is positioned not as a replacement for, but as a complement to, the well-established role of FEM simulation. In this research, FEM has been employed to demonstrate how FEM and ML can serve as mutually informative techniques: FEM provides foundational physical insight and validation, while ML enables efficient and adaptive prediction across diverse experimental configurations. This comparative analysis offers practitioners clear guidance on selecting and potentially integrating these approaches based on specific measurement requirements and available resources. Ultimately, the study aims to advance more reliable thin-film characterization by fostering methodological integration between physical and data-driven domains.

Table 7-3. ITO thin films and metallic samples used for EC measurements (mesh scale: 1 mm).

ITO (a)	ITO (b)	ITO (c)	ITO (d)	Copper (Cu)	Tungsten (W)	Nickel (Ni)	Iron (Fe)	Tin (Sn)
								

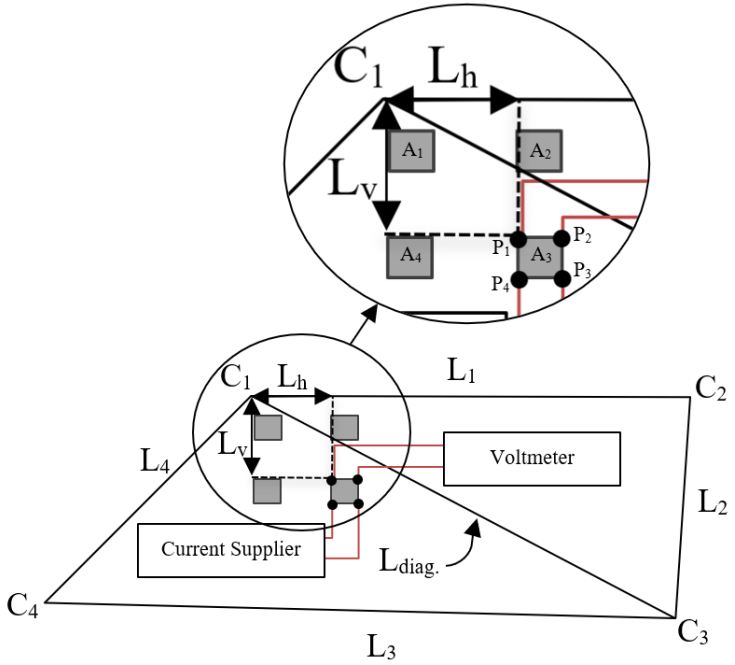


Figure 7-2. Schematic of probes, sample and different probe positions (A_1-A_4), illustrating spatial arrangement and measurement setup.

7.3. Mathematical Foundation and ML Model

Miccoli et al. derived an analytical equation to relate measured EC (σ_{exp}) to measured signals for a rectangular 4PP configuration [34]:

$$\sigma_{exp} = \frac{I}{2\pi V d}\,\mathrm{Ln}(2) \tag{7.1}$$

This expression, rooted in electrostatic theory, models current flow in an infinite 2D conductive medium. The logarithmic factor $\mathrm{Ln}(2)$ emerges from the probe geometry's voltage decay characteristics, where I represents injected current and V the measured voltage. It is important to note that this is an idealized model. It assumes a perfect setup and does not account for real-world experimental errors, such as inaccuracies in probe placement or the distorting effect of a sample's finite boundaries and edges.

To address these limitations, FEM simulations provide a physically-based correction by introducing geometric constraints:

$$\sigma_{FEM} = \left(\frac{\sigma_a}{\sigma_{num}}\right)\sigma_{exp} \tag{7.2}$$

Here, σ_a denotes the actual conductivity input to the FEM model, while σ_{num} represents the simulated apparent conductivity obtained using COMSOL under identical probe arrangements. The ratio (σ_a/σ_{num}) serves as a correction factor, and boundary effects neglected in the theoretical formula (Equation 7.1).

ML suggests a parallel approach through data-driven calibration. While FEM provides strong simulation-based solutions, ML models trained on empirical datasets offer complementary, direct predictions of EC across diverse experimental conditions, enhancing efficiency while preserving interpretability within a physics-informed framework. As a parallel approach, a ML model (F_{ML}) is defined as:

$$F_{ML}\left(L_1, L_2, L_3, L_4, L_{diag.}, L_h, L_v, I\right) = \frac{2\pi}{\mathrm{Ln}(2)}\,(\sigma_{ML} d)V = \Theta \tag{7.3}$$

Here, F_{ML} is the ML model, which receives input variables comprising sample dimensions (side lengths L_1, L_2, L_3, L_4), diagonal length (L_{diag}), probe positions (horizontal and vertical spacings L_h and L_v), and a single applied-current value (I) selected within the experimental current range, to predict the target variable Θ as the output of the ML model. Once Θ is predicted and V is experimentally measured, the EC (σ_{ML}) is calculated as:

$$\sigma_{ML} = \frac{\mathrm{Ln}(2)}{2\pi V d}\,\Theta \tag{7.4}$$

The Equation 7.3 is developed to address the challenge of predicting EC values when only a limited number of unique EC values are available in the training dataset.

By defining Θ as a continuous variable, the ML framework reframes the prediction task as a regression problem rather than a discrete classification task. From an ML perspective, regression is fundamentally more suitable because EC is a continuous physical quantity rather than a discrete label. Using classification would force the model to select among a small set of discrete conductivity classes. By contrast, a regression-based formulation allows the model to learn the physical relationship (like Ohm's law) between experimental inputs and the continuous conductivity-related variable Θ. This preserves the natural physical variations in the data, and the ML model is therefore developed as a physics-informed framework.

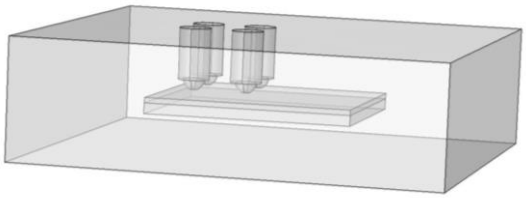
This hybrid approach, combining theoretical foundations with data-driven ML corrections, provides a framework for accurate conductivity measurements in thin films, particularly where traditional analytical models are inadequate.

Although FEM remains a widely trusted approach due to its detailed handling of geometry and boundary conditions, the ML method demonstrates promising potential because it learns directly from experimental data. However, acquiring a sufficiently large number of experimental data points for ML training can be challenging, especially when attempting to capture purely geometrical effects with high accuracy.

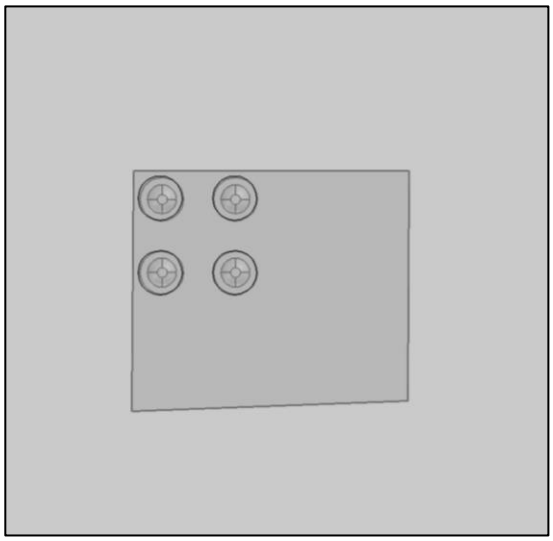
8. Results and Discussion (Electrical)

8.1. FEM Simulation

To account for geometric effects and validate the experimental 4PP measurements, a three-dimensional FEM model was developed using COMSOL Multiphysics [73]. The purpose of this simulation was to reproduce the experimental setup under controlled numerical conditions and to establish a physically based correction that improves the accuracy of conductivity measurements. Figure 8-1 illustrates the configuration employed in the simulations.



(a)



(b)

Figure 8-1. FEM geometry of the 4PP configuration. (a) Side view showing the thin conducting sample placed on a dielectric substrate, with the system embedded in an air domain to reproduce realistic experimental conditions. (b) Top view illustrating the square arrangement of the four probe tips in contact with the sample surface.

The calculated potential maps for different probe placements are shown in Figure 8-2. As a result, the numerical EC value (σ_{num}) obtained from simulations can deviate significantly from the actual value used as the model parameter (σ_a). By capturing these variations, the FEM simulation provides clear visual evidence of the geometric artifacts inherent in the 4PP method and shows how corrected values can be derived (using Equation **7.2**) to more accurately represent the material's true conductivity.

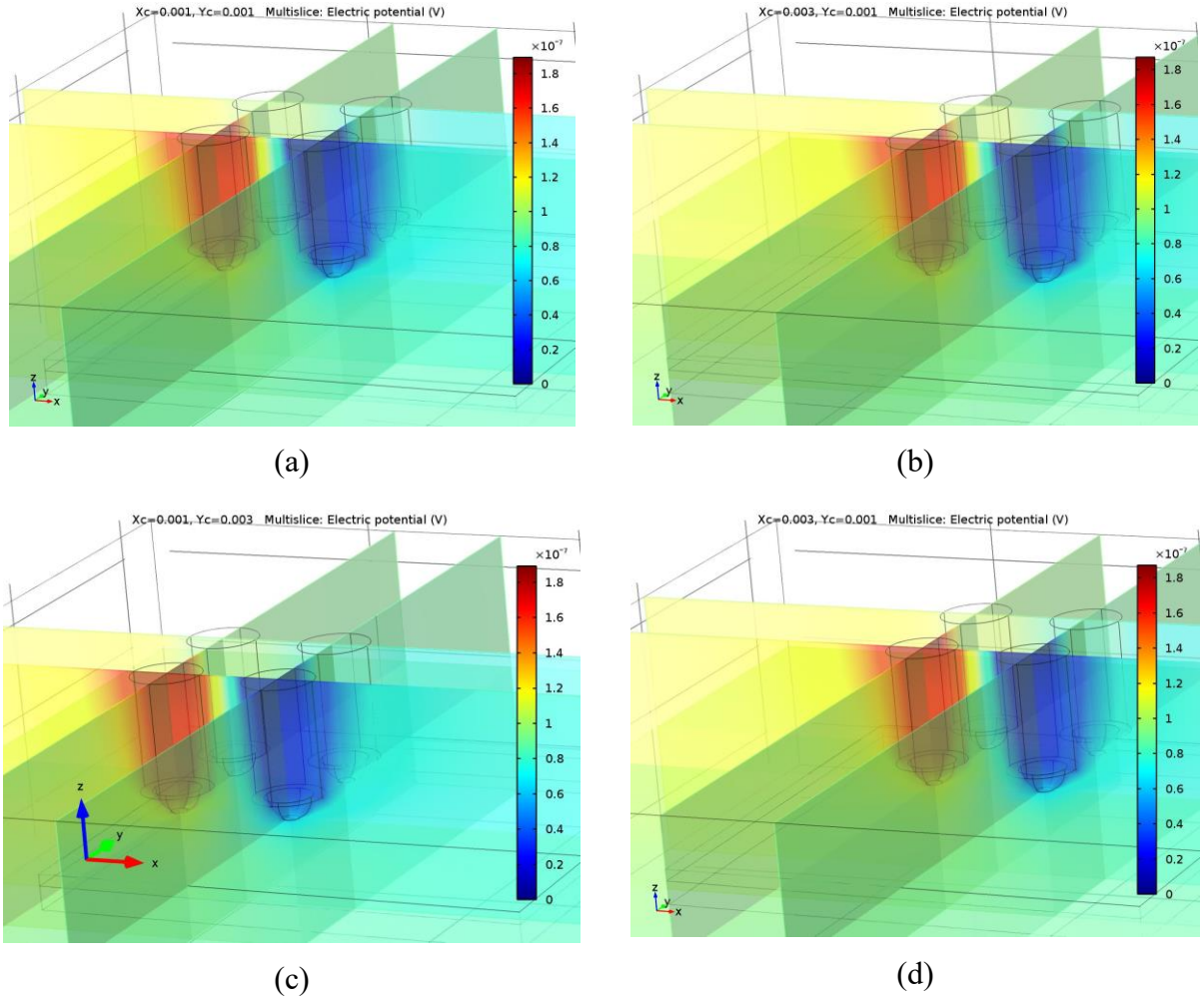


Figure 8-2. Simulated electric potential distribution for the thin conducting sample under different 4PP configurations a: (1,1), b: (3,1), c: (1,3), d: (3,3).

This simulation therefore provides both a visual and quantitative confirmation of the correction methodology. By incorporating geometric effects into the model, FEM produces corrected values that align much more closely with intrinsic properties than uncorrected experimental data. The approach establishes a physically grounded method for improving conductivity measurements in thin films and related systems, and it shows particular promise for laboratory environments where probe spacing can be carefully controlled.

The effectiveness of the FEM correction is quantitatively demonstrated in Table 8-1, which compares the measured (σ_{exp}), numerical (σ_{num}, found using simulation), actual (σ_a), and FEM-corrected corrected (σ_{FEM}) conductivity values for iron (Fe) sample across different probe configurations (L_v, L_h).The FEM-corrected conductivity (σ_{FEM}) demonstrates significant progress toward approximating the intrinsic conductivity (σ_a), validating the effectiveness of the correction approach. While not perfectly matching σ_a in all cases, the σ_{FEM} values show much closer alignment with the theoretical intrinsic value than the uncorrected experimental measurements (σ_{exp}).

Table 8-1. Summary for comparison of measured (σ_{exp}), numerical (σ_{num}), actual (σ_a), and FEM-corrected (σ_{FEM}) EC values for iron (Fe) at different 4PP configurations (L_v, L_h).

(L_v (mm), L_h (mm))	σ_{exp} (kS · cm^{-1})	σ_{num} (kS · cm^{-1})	σ_a (kS · cm^{-1})	σ_{FEM} (kS · cm^{-1})
(1, 1)	75	66	103	118
(3, 1)	77	79	103	100
(1, 3)	66	64	103	107
(3, 3)	79	76	103	107

For the (3, 1) configuration, σ_{FEM} achieves high accuracy at 100 kS · cm^{-1} compared to the reference value σ_a of 103 kS · cm^{-1}, corresponding to a deviation of approximately 3 %. This near-convergence demonstrates that the FEM correction effectively compensates for geometric artifacts when appropriate probe spacing is employed.

The visual representation of this data is shown in Figure 8-3. Additional graphs for other samples are provided in Appendix A.

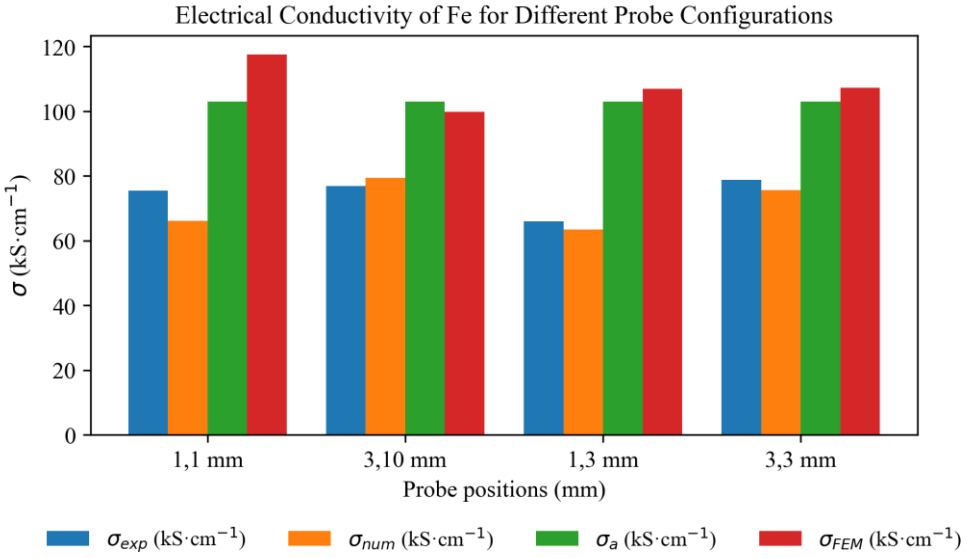


Figure 8-3. A comparison of experimental, numerical, intrinsic, and FEM-corrected EC values for iron (Fe) sample across various 4PP configurations. Each label on the horizontal axis indicates a specific probe placement: the first number (before the comma) represents the vertical position of the probe L_v, and the second number (after the comma) represents the horizontal position of the probe L_h.

The systematic pattern of results confirms the physical validity of the correction model - σ_{FEM} consistently approaches σ_a from measurable experimental data, demonstrating the method's fundamental soundness. The remaining discrepancies likely stem from secondary effects not yet incorporated in the model rather than flaws in the core approach. These excellent preliminary results strongly suggest that with refined modeling of boundary conditions or probe-specific correction factors, even better agreement with σ_a could be achieved.

This successful demonstration of physically-based experimental correction establishes an important methodology for accurate conductivity measurement, with particular value for materials

characterization where intrinsic properties must be derived from imperfect experimental setups. The technique shows promise for standard laboratory applications where controlled probe geometries can be maintained.

8.2. Spearman's Correlation Analysis

To evaluate the relationships between the input features and the target variable, Spearman's rank correlation analysis was performed. Table 8-2 presents the interpretation of Spearman's rank correlation coefficients (r_s) for all variable pairs used in model development. The strongest correlation is observed between the applied current (I) and the output variable (Θ), with a coefficient of 0.96, indicating a near-perfect monotonic relationship. This result aligns with Ohm's law, which states that the target variable is directly proportional to current ($\Theta \propto I$). The high correlation not only confirms the expected physical behavior but also validates the experimental measurements. Among all sample dimensions, sides L_1 (r_s = 0.35) and L_3 (r_s = 0.37) exhibit stronger positive correlations with Θ, suggesting that their orientation relative to the current path significantly influences current density distribution in different samples. In contrast, L_2 (r_s = -0.09) and L_4 (r_s = -0.11) show weaker negative correlations with Θ, indicating that L_2 and L_4 affect current paths in an opposing manner. The analysis for the diagonal $L_{diag.}$ (r_s = 0.23) exhibits a weaker positive correlation with Θ compared to the edge effects from L_1 and L_3, as expected. Finally, L_v (r_s = -0.05) and L_h (r_s = 0.08) demonstrate negligible correlations, confirming their minimal influence on Θ.

Table 8-2. Spearman's rank correlation coefficients (r_s) between all pairs of variables used in ML modelling.

Variable	I (mA)	L_v(mm)	L_h(mm)	$L_{diag.}$(mm)	L_1(mm)	L_2(mm)	L_3(mm)	L_4(mm)
Θ (mA)	0.96	-0.05	0.08	0.23	0.35	-0.09	0.37	-0.11

Although variations in probe position would, in principle, be expected to affect the measured response, their impact is intentionally suppressed in the present dataset. This is because only two constrained variations in the vertical and horizontal directions were included, with probe positions fixed relative to the reference vertex (C_1) for all samples. As a result, positional changes do not introduce sufficient independent variability to significantly alter the final output Θ, leading the model to rely more on edge-related effects rather than absolute probe location in its predictions.

8.3. Random Forest Model Development and Hyperparameter Optimization

Model performance was assessed across different Random Forest hyperparameter settings. Table 6.1 summarizes the hyperparameters of the Random Forest algorithm used in this study, along with their search ranges (optimized via grid search). In total, 170 different hyperparameter combinations were evaluated to rigorously tune the Random Forest model for generalized performance. The subsequent section applies different optimization scenarios to select the best hyperparameter combination by monitoring and comparing predictive performance during cross-validation (CV) stage.

Table 8-3. The range of hyperparameters investigated for the ML models in the current study (170 evaluated configurations).

Model's name	Investigated hyperparameters	Investigated range
Random Forest	Number of estimators	81-4249
	Maximum depth	55 -867
	Minimum samples split	2 -153
	Minimum samples leaf	1 -174
	Criterion	Poisson, squared error, absolute error
	Bootstrap	True - False

Table 8-4 presents the hyperparameters of the four best-performing set of hyperparameters for the Random Forest model, identified based on their performance during the CV stage. These models were selected for their potential to generalize effectively in predicting Θ. As shown in the Table 8-4, Random Forest model 'a' performed well compared to the other configurations. Visual inspection of the statistical indices confirms that this model provided acceptable predictive performance for Θ. The next section employs a spider graph to further support these findings. Models 'a', 'b', 'c' and 'd' denote different hyperparameter configurations (e.g., number of estimators, maximum depth, etc.) of the same Random Forest base model.

Table 8-4. The best identified hyperparameters for each ML model and their corresponding accuracy in predictions.

Model's name	Tuned hyperparameters	Uncertainty index	CV training phase	CV testing phase
Random Forest 'a'	Number of estimators = 2445 Maximum depth = 711 Minimum samples split = 2 Minimum samples leaf = 3 Criterion: Absolute error Bootstrap = True	RMSE	16.208	39.922
		R^2	0.99658	0.97641
		AARD%	3.328	8.639
		MAE	5.088	12.212
Random Forest 'b'	Number of estimators = 2076 Maximum depth = 783 Minimum samples split = 11 Minimum samples leaf = 1 Criterion: Absolute error Bootstrap = True	RMSE	58.650	65.371
		R^2	0.95504	0.94141
		AARD%	11.134	15.581
		MAE	18.113	20.992
Random Forest 'c'	Number of estimators = 4828 Maximum depth = 820 Minimum samples split = 17 Minimum samples leaf = 5 Criterion: Poisson Bootstrap = False	RMSE	65.438	72.481
		R^2	0.94394	0.92756
		AARD%	13.119	17.048
		MAE	20.717	24.108
Random Forest 'd'	Number of estimators: 4496 Maximum depth: 840 Minimum samples split: 21 Minimum samples leaf: 9 Criterion: Poisson Bootstrap: False	RMSE	96.939	97.522
		R^2	0.87759	0.86870
		AARD%	18.408	21.454
		MAE	31.186	32.623

8.4. Selecting the Best Hyperparameters

Figure 8-4 shows the spider graphs of RMSE, MAE, and AARD% for the four selected Random Forest models during the CV training stage. Figure 8-5 presents the same performance metrics for the CV testing stage. In these spider plots, smaller triangles correspond to lower error values and thus indicate better model performance.

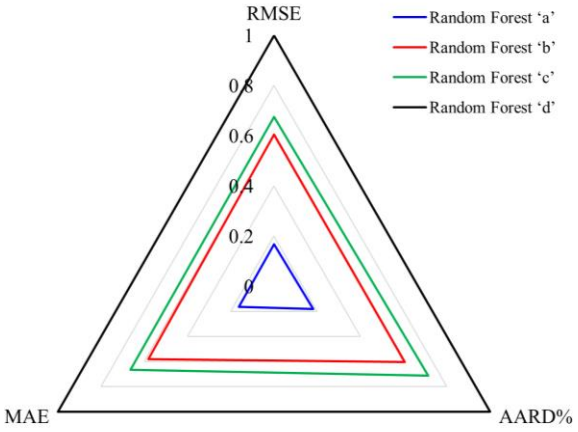


Figure 8-4. Comparing the performance of the selected ML models in the CV training stage using the radar graph.

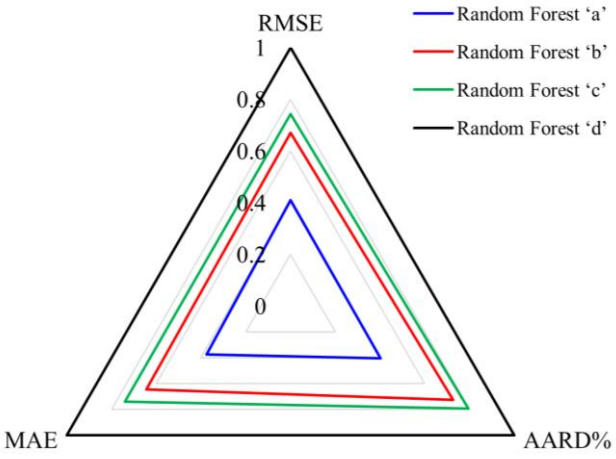


Figure 8-5. Comparing the performance of the selected ML models in the CV testing phase using the radar graph.

The evaluation shows that among the four Random Forest models (a–d), Model 'a' is the optimal choice due to its superior performance across all measured metrics. On the training set, Model 'a' achieved the lowest RMSE (16.208), AARD% (3.328%), and MAE (5.088), significantly outperforming the others. This superiority is visually represented in Figure 8-4, where the polygon for Model 'a' is the smallest, confirming its best fit to the training data. Crucially, this advantage persisted in the test set, where Model 'a' maintained the lowest errors (RMSE: 39.922; AARD%: 8.639%; MAE: 12.212), which is the primary

confirmation of its reliability on unseen data. In Figure 8-5, the polygon for Model 'a' remains the smallest, visually demonstrating that the error did not increase when facing test data. This graphical consistency and the low errors confirm Model 'a's effective generalization, while models b–d exhibited higher errors in both train and test phases (e.g., test RMSEs of 65.371–97.522).

Figure 8-6 compares the R^2 values of the selected models during CV training and testing phases. Higher R^2 values indicate better performance. The R^2 scores further reinforces the dominance of Random Forest ‘a’, with near-perfect scores in both cross-validated training (0.9966) and testing (0.9764) phases. These values indicate that ‘a’ explains 99.7% and 97.6% of the variance in the training and test datasets, respectively, considerably higher than competing models (b–d).

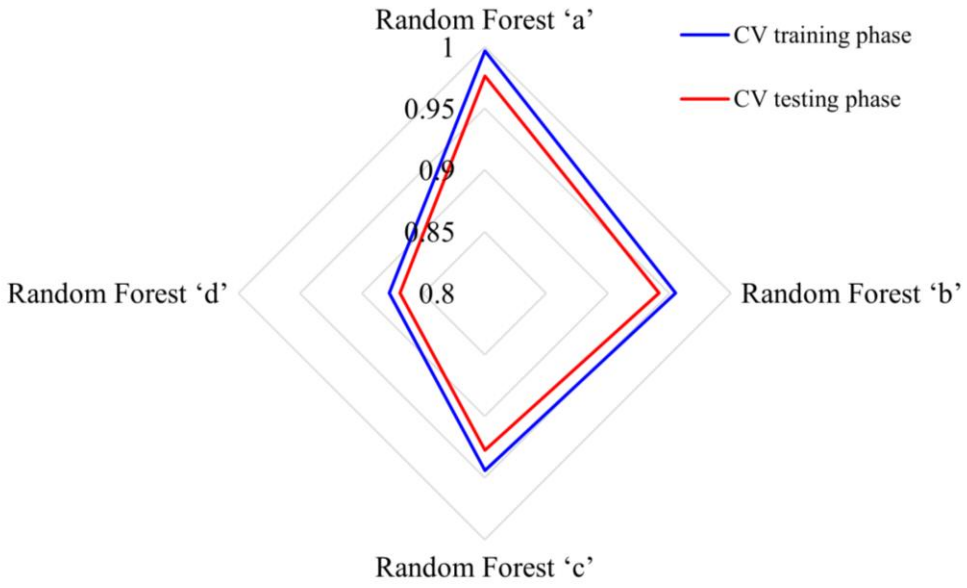


Figure 8-6. Comparing the presented R^2 index by the ML models in the CV training/testing phases.

Therefore, Random Forest ‘a’ is identified as the most accurate model for estimating the experimental values of Θ, consistently outperforming the other models in both training and testing evaluations at the CV stage. Its ability to maintain lower error values, and higher R^2 scores demonstrates that it captures the underlying physical and geometrical relationships more effectively than competing models.

A new Random Forest model was trained using the exact optimal hyperparameters identified in the previous sections for Model ‘a’. These Hyperparameters are shown in Table 8-4. Here the model was trained using the combined CV data (both training and testing sets) and subsequently validated against the separate experimental test set (221 data points), which was unseen during CV stage. The performance metrics for the final training and testing phases at this stage, summarized in Table 8-5.

Table 8-5. Performance Metrics of the Final Optimized Random Forest Model During Training and Testing Phases.

Metric	Training data	Testing data
RMSE	16.741	77.777
R^2	0.99636	0.95626
AARD %	3.179	9.755
MAE	5.178	33.798

The test results exhibit slightly higher errors compared to those of Model 'a' from CV testing phase. This marginal increase in error can likely be attributed to the inherent challenges of generalizing to an independent unseen dataset, which may contain variability not captured during CV tuning. Nevertheless, the model maintains reliable predictive performance, demonstrating the effectiveness of the selected hyperparameters. Using the final model, predictions were made for the target variable (Θ), and Equation 7.4 was applied to compute the EC for different samples. The following subsection analyzes the results for metallic and ITO samples.

8.5. Performance of ML and FEM-Based Approaches for EC Estimation

The predicted EC values (σ_{ML}) were derived from Equation 7.4 using the ML-predicted Θ values, while the simulated EC values (σ_{FEM}) were calculated using Equation 7.2. Figure 8-7 illustrates the Relative Difference percentage (RD%) between the predicted and simulated EC values for metallic samples. This comparative analysis of FEM and ML approaches reveals that both methods demonstrate strong predictive performance with comparable error characteristics. RD% for the FEM method exhibits a mean value of +5.66% with a standard deviation of 14.41%, indicating a consistent tendency toward moderate underestimation with stable variability.

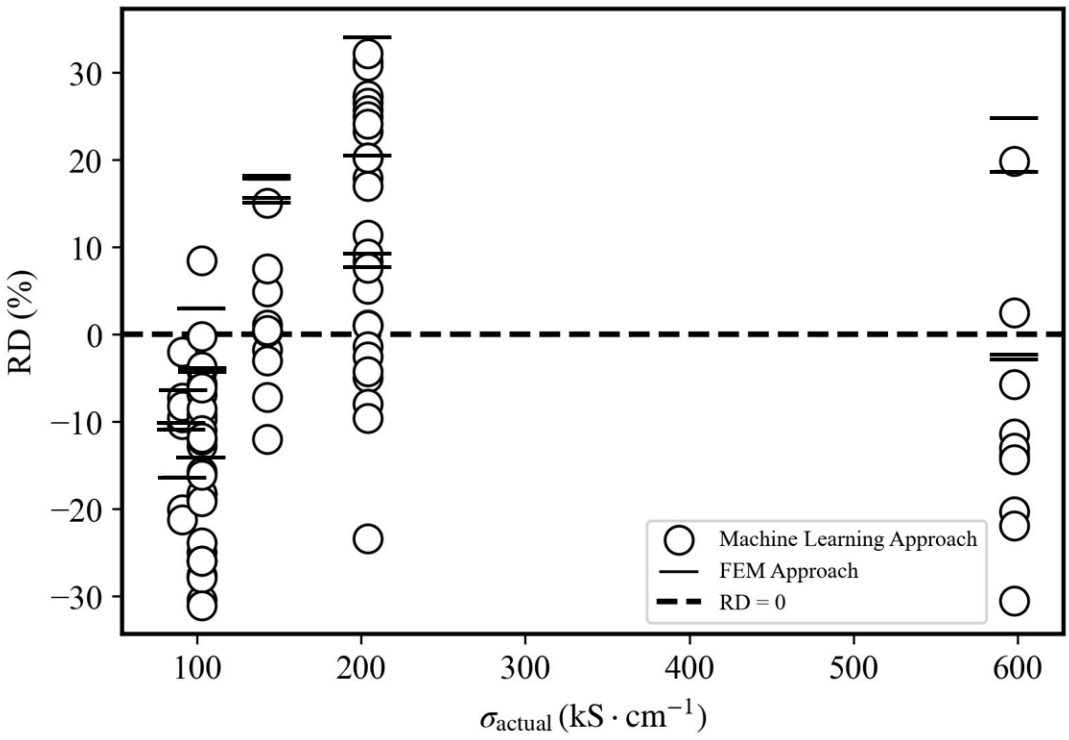


Figure 8-7. Relative difference percentage between predicted and theoretical EC values for metallic samples.

In complementary, the ML approach shows a mean RD% of -3.15% with a standard deviation of 16.69%, reflecting a slight overprediction tendency with marginally wider dispersion. The observed deviation patterns between the two methods are roughly close in magnitude, differing by less than 9% in mean values and only 2.3% in standard deviation. This close agreement in performance metrics suggests that both techniques provide similarly reliable estimations, with their opposing bias directions (+5.66% versus -3.15%) potentially offering opportunities for error compensation when used in combination. The complementary deviation profiles of FEM and ML strategically leveraging their systematic differences could enable the development of more accurate hybrid prediction systems without favouring either method individually. Both approaches retain their respective strengths while demonstrating statistically similar performance.

The normal density distributions of residual errors further highlight this distinction (Figure 8-8). The normal density plot reveals important differences in the error distributions related to the FEM and ML methods for EC predictions. The FEM approach demonstrates an underestimation tendency with a mean error of +20.47 kS · cm^{-1}, while the ML method shows a slight overprediction bias averaging -5.48. Both methods exhibit similar variability in their errors, with standard deviations of 43.95 kS · cm^{-1} for FEM and 43.69 kS · cm^{-1} for ML, indicating nearly identical consistency in their predictions. These results are significant because the validation focused on challenging edge regions where EC measurements are most affected by boundary effects, areas traditionally problematic for accurate characterization. The intentional inclusion of these difficult edge conditions demonstrates ML's ability to maintain prediction quality in worst-case scenarios. The complementary nature of the biases between the two methods, combined with their similar

error distributions, suggests potential opportunities to develop hybrid approaches that could leverage their respective strengths.

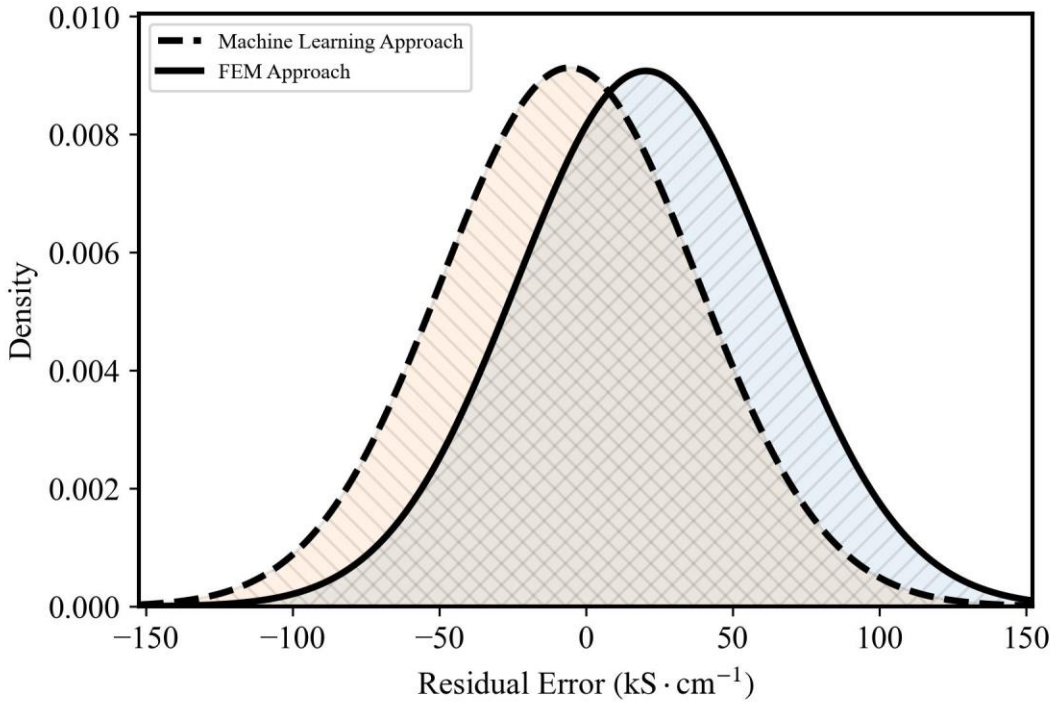


Figure 8-8. Normal density distributions of residual errors for metallic samples.

The comparative analysis of ITO samples (Figure 8-9) reveals that both FEM and ML approaches demonstrate valid but distinct performance characteristics. The RD% for the FEM shows a mean value of −4.30% with a standard deviation of 17.06%, while the ML approach exhibits a mean value of −2.93% with a standard deviation of 10.14%. The 1.37 percentage point difference in mean values falls within a comparable range of accuracy, while the difference in variability reflects their fundamentally different methodological approaches - FEM's first-principles physical modeling versus ML's empirical data fitting. Both methods show similar overestimation tendencies, confirming their consistent physical interpretation of the conductivity measurements. The results demonstrate that FEM maintains its expected reliability as a physics-based approach, while ML shows its capability to produce competitive results through data-driven approximation.

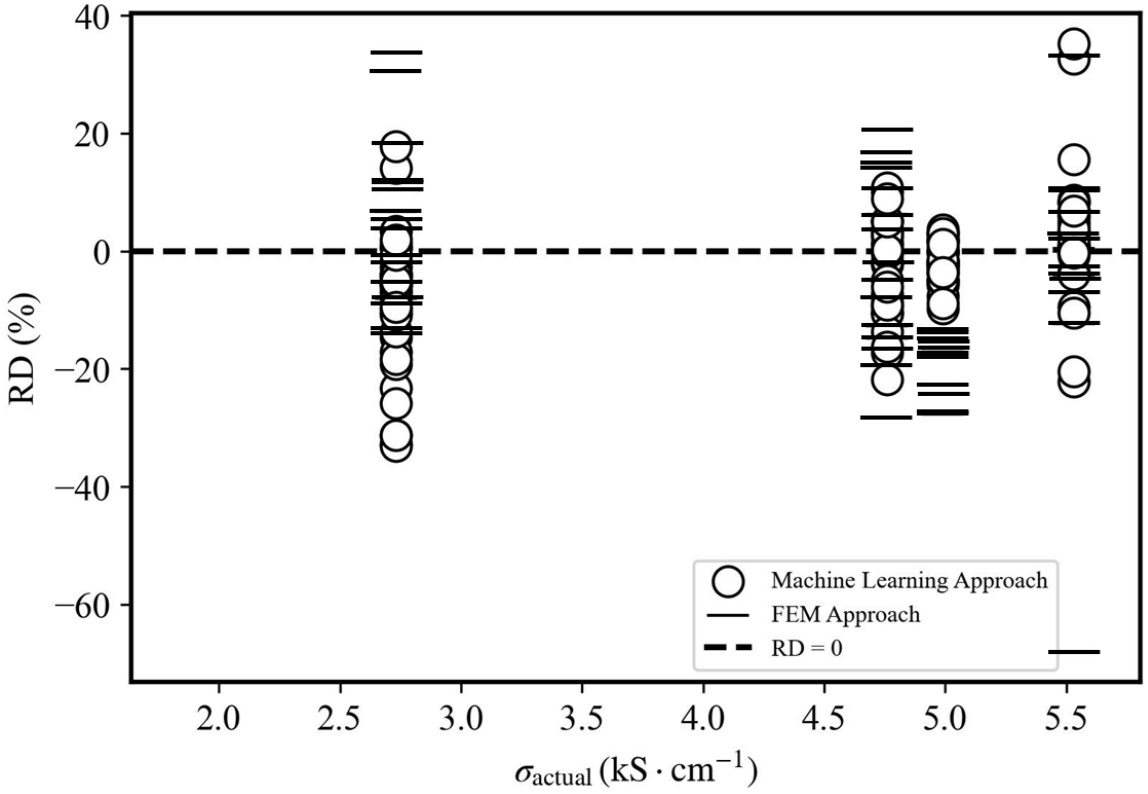


Figure 8-9. Relative difference percentage for ITO samples.

The residual distributions for ITO (Figure 8-10) reinforce these observations. The normal density plot reveals that both FEM and ML methods produce residual errors centered near zero for ITO samples, with FEM showing a mean error of -0.243 kS · cm^{-1} and ML demonstrating a mean of -0.085 kS · cm^{-1}. While both methods maintain good overall accuracy, their error distributions show characteristic differences that reflect their underlying methodologies. FEM exhibits a standard deviation of 0.813 kS · cm^{-1}, representing the expected variability in its physics-based calculations, while ML's tighter distribution (with a standard deviation of 0.425 kS · cm^{-1}) reflects its data-driven optimization.

The results show that FEM maintains its fundamental physical consistency, with error characteristics that align with expected physical constraints of the material system. ML demonstrates its ability to learn and reproduce these physical relationships through training data, resulting in reduced but still physically reasonable error variation. Both approaches cluster near zero error, confirming their validity for ITO characterization, with FEM's slightly wider distribution representing the natural variability in first-principles modeling and ML's narrower spread showing the benefits of empirical optimization.

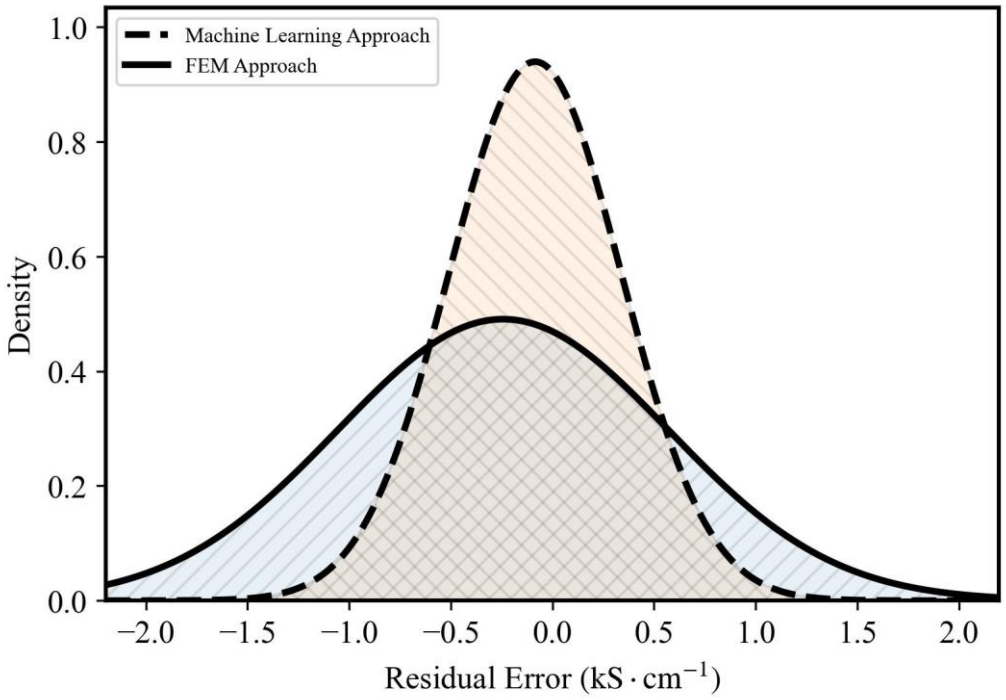


Figure 8-10. Residual error distributions for ITO samples.

These findings illustrate how both techniques - one based on physical laws and the other on data patterns - arrive at similarly valid but characteristically different solutions to the ITO characterization challenge. The results affirm that FEM and ML can serve as complementary approaches, with FEM providing fundamental physical insight and ML offering efficient empirical approximation, both maintaining good agreement with experimental reality for ITO materials.

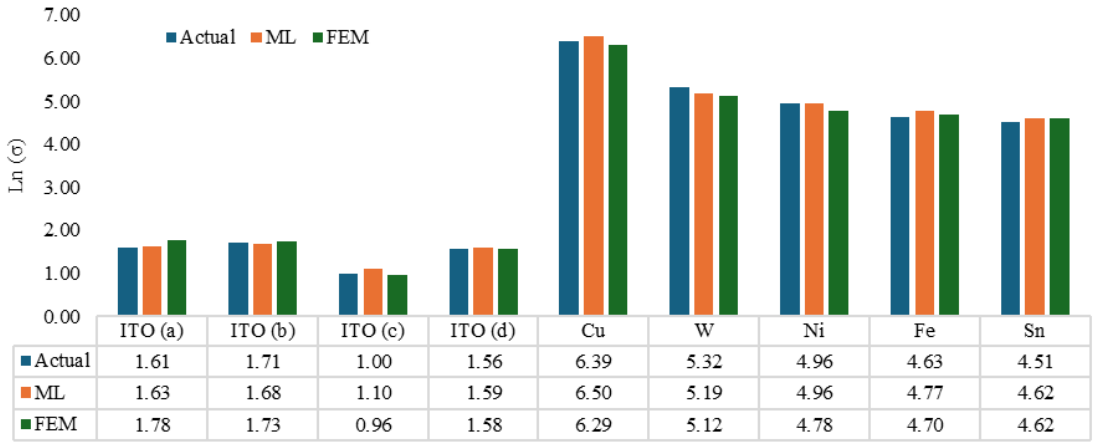
8.6. EC Derived from the ML and FEM Approaches

Table 8-6 presents a comparison between theoretical, predicted, and actual EC values for the nine different samples, including four ITO samples and five metal samples (Cu, W, Ni, Fe, Sn). The predicted EC values (σ_{ML}, derived from the Equation 7.4) were obtained by averaging the EC values predicted by the ML model using the new set of unseen data for each sample. Similarly, the simulated EC values (σ_{FEM}, derived from the Equation 7.2) were calculated by averaging the results across different experimental conditions for each sample. The visual comparisons are shown in Figure 8.11.

Table 8-6. Comparison of theoretical, predicted and actual EC values ($kS \cdot cm^{-1}$) for the different samples.

Sample*	σ_{actual}	ML approach		FEM approach	
		σ_{ML}	RD%	σ_{FEM}	RD%
ITO (a)	4.99 [55]	5.08	−1.8	5.91	−18.4
ITO (b)	5.53 [55]	5.36	3.1	5.66	−2.4
ITO (c)	2.73 [55]	3	−9.9	2.6	4.8
ITO (d)	4.76 [55]	4.91	−3.2	4.86	−2.1
Cu	598	663	−10.9	541	9.5
W	204	179	12.3	168	17.6
Ni	143	142	0.5	119	16.8
Fe	103	118	−14.6	110	−6.8
Sn	91	101	−11.0	101	−11.0

* ITO samples (a) to (d) were annealed in different atmospheres: carbon dioxide, a nitrogen-hydrogen mixture, vacuum, and nitrogen. All samples were deposited on glass substrates. The EC of these ITO samples (a) to (d) was previously measured using the standard vdP method, as reported in the referenced studies.



	ITO (a)	ITO (b)	ITO (c)	ITO (d)	Cu	W	Ni	Fe	Sn
Actual	1.61	1.71	1.00	1.56	6.39	5.32	4.96	4.63	4.51
ML	1.63	1.68	1.10	1.59	6.50	5.19	4.96	4.77	4.62
FEM	1.78	1.73	0.96	1.58	6.29	5.12	4.78	4.70	4.62

Figure 8.11. Comparison of actual, ML (predicted) and FEM-corrected EC values expressed as Ln(σ) for different materials, where σ is given in $kS \cdot cm^{-1}$.

The ML approach shows relatively consistent accuracy across all four ITO samples, with relative differences confined to a narrow range between −9.9 % and 3.1 %. In contrast, the FEM approach exhibits greater variability, with deviations ranging from −18.4 % to 4.8 %, although it achieves strong agreement for ITO (c), where it slightly outperforms the ML prediction.

For metallic samples, the ML predictions remain within approximately ±15 % of the reference values, demonstrating the model's capability to produce reliable estimates even when trained on a limited dataset. While the present dataset is sufficient to construct a functional generalized ML model, further expansion of the training data is expected to

improve predictive accuracy. FEM predictions for metallic samples display a broader spread of deviations, ranging from −11.0 % to 17.6 %.

Overall, the results highlight the complementary strengths of the two approaches. FEM benefits from detailed physics-based simulations and provides strong performance when comprehensive material modeling is available, particularly for ITO systems. In contrast, the ML approach demonstrates promising generalization capability across both oxide and metallic samples, with clear potential for improvement through expanded and more diverse training datasets. These findings suggest that combining FEM's physics-driven precision with ML's data-driven adaptability could lead to more robust and reliable electrical conductivity characterization frameworks in future studies.

9. Conclusions

9.1. Overall Conclusions – Thermal Conductivity

This thesis demonstrates the development of a framework for the accurate determination of intrinsic thermal conductivity in thin films using scanning thermal microscopy. A new approach is proposed in which simultaneous thermal–topographical SThM measurements are combined with normalization strategies and machine learning regression to overcome the fundamental limitations of SThM analysis. The application of this framework enables the separation of intrinsic thermal transport properties from measurement-induced artifacts, providing practical relevance for accurate materials characterization and device engineering.

The framework integrates quartz-referenced normalization, a substrate–thickness factor, and machine learning models trained on topographical descriptors. This combination allows the mitigation of surface topography effects and probe–sample contact variability, which are commonly neglected or oversimplified in thermal conductivity measurements.

The effectiveness of the proposed approach is confirmed through quantitative validation against reference materials and literature-reported values.

Despite these improvements, residual discrepancies between corrected and ideal intrinsic thermal conductivity values remain due to unavoidable experimental limitations. Variations in probe–sample contact conditions, calibration uncertainty, and heat losses through the substrate contribute to the remaining deviations. The explicit identification of these factors demonstrates scientific rigor and confirms control over the measurement system rather than overestimation of model capability.

Beyond predictive accuracy, the analysis of model behavior provides physical insight into sub-microscale heat transfer mechanisms. The observed correlations between morphology descriptors, normalization parameters, and predicted thermal conductivity provide fundamental thermal physics insight into the measured SThM response.

The proposed framework confirms the effectiveness of combining physics-based strategies with machine learning for intrinsic thermal conductivity extraction. The methodology can be further improved through expanded reference datasets, refined calibration standards, and the incorporation of physics-informed learning constraints. The approach is expected to support standardized and scalable thermal characterization of thin films across a wide range of functional materials.

Objective. To develop, validate, and demonstrate a reliable methodology for thin-film TC determination that addresses the principal limitations of SThM, particularly surface topographical effects, by jointly exploiting simultaneous thermal–topographical mapping, normalization and correction strategies, and ML regression.

Hypothesis. Integrating high-resolution SThM thermal–topography maps with quartz-referenced normalization, together with ML models trained on physics-aware features, will yield accurate, reproducible, and generalizable TC estimates for thin films, including cases where film thickness is below the SThM tip radius (~100 nm).

Confirmation of Hypothesis. The combined framework—simultaneous thermal–topographical SThM, quartz-referenced normalization (Γ_i), and ensemble ML regression—produced accurate, reproducible, and scalable TC estimates across thin films and bulk references. High performance (independent test $R^2 \approx 0.97$) and close agreement with reference values demonstrate that contact variability and morphology-induced artifacts can be quantitatively mitigated.

Future Work

- **Topography-aware correction via machine learning:** Extend the current framework to explicitly account for surface topography effects by introducing topography-related descriptors (e.g., local roughness metrics, height variations, or probe–sample contact indicators) as additional input features. The ML model will learn an effective topography correction factor directly from these descriptors, enabling compensation for surface-induced measurement artifacts and improving the extraction of true TC values.
- **Dataset expansion and balance:** Increase representation of high-κ materials and ultrathin films; include additional substrates to better span interfacial conductance regimes.
- **Standardization and tooling:** Package the pipeline (normalization → feature extraction → ML inference → reporting) into a reproducible software framework with traceable calibration using quartz and additional reference materials.
- **Physics-Informed Machine Learning (PIML):** Explore physics-informed machine learning approaches that embed governing physical laws directly into the ML loss function, enabling high accuracy with reduced dataset sizes.

9.2. Overall Conclusions – Electrical Conductivity

This thesis demonstrates the development of a framework for the accurate determination of intrinsic electrical conductivity in thin films using the four-point probe technique. An approach is proposed in which FEM and machine learning are separately employed to correct geometric and probe-related distortions that commonly affect electrical conductivity measurements. The application of this framework enables reliable convergence toward intrinsic conductivity values under non-ideal measurement conditions, with direct relevance to thin-film materials research and industrial characterization.

The developed methodology explicitly accounts for finite sample dimensions, probe placement variability, and edge effects through physics-based and data-driven correction mechanisms. FEM provides interpretable correction factors rooted in electrical transport physics, while machine learning (trained solely on experimental data) captures complex empirical distortions arising from diverse geometrical configurations.

Quantitative validation demonstrates strong agreement between corrected EC values and vdP reference measurements.

The main technical conclusions of the EC study can be summarized as follows:

- Geometric distortions and probe placement effects significantly bias conventional four-point probe conductivity measurements and must be corrected for accurate intrinsic property determination.

- FEM provides reliable, physics-based correction factors for non-ideal measurement geometries.
- Machine learning effectively captures geometrical distortions and generalizes across diverse experimental configurations.

Objective. To establish a reliable framework using ML (comparable with FEM results) in order to (i) correct geometric artifacts and (ii) ensure consistency of conductivity estimates using solely experimental measurements.

Hypothesis. FEM provides physics-based corrections for probe- and geometry-induced distortions, while ML learns adaptive correction patterns from experimental data. The application of ML approach yields accurate and generalizable conductivity values that converge toward intrinsic properties, including those validated by the vdP method.

Confirmation of Hypothesis. The results confirm that combining FEM and ML enables accurate and reproducible thin-film electrical conductivity characterization under non-ideal geometries. FEM provides physically interpretable correction factors, while ML captures complex empirical distortions and generalizes across diverse conditions without extra simulations. Agreement with vdP supports the validity of both approaches.

Future Perspectives

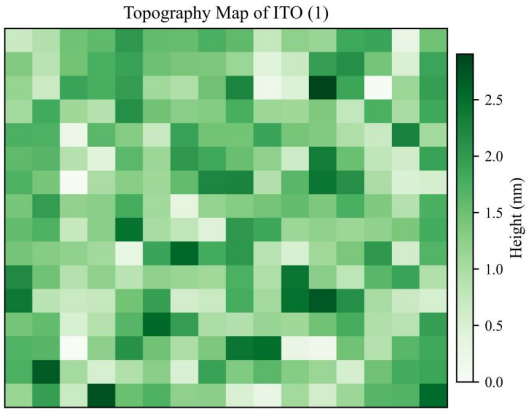
- **FEM vs. ML comparison:** FEM currently provides more practical results for electrical conductivity measurements; however, ML demonstrates strong potential when trained on large datasets covering millions of geometrical configurations, enabling electrical conductivity prediction directly from experimental data without simulations.
- **Dataset expansion:** Extend experimental datasets to additional thin-film materials, probe spacings, and controlled anisotropic samples to improve generalization.
- **Active learning strategies:** Adaptively guide probe placement and sample orientation to maximize data efficiency.
- **Standardization and automation:** Translate the methodology into standardized protocols and software tools suitable for both research laboratories and industrial characterization workflows.

10. Appendix A:

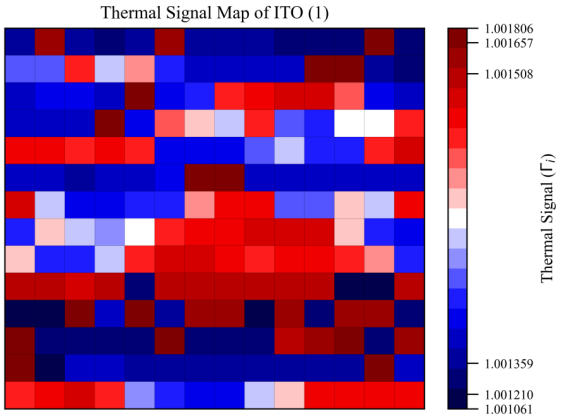
10.1. Topographical and Thermal-Signal Maps of the Samples for Thermal Properties Analysis

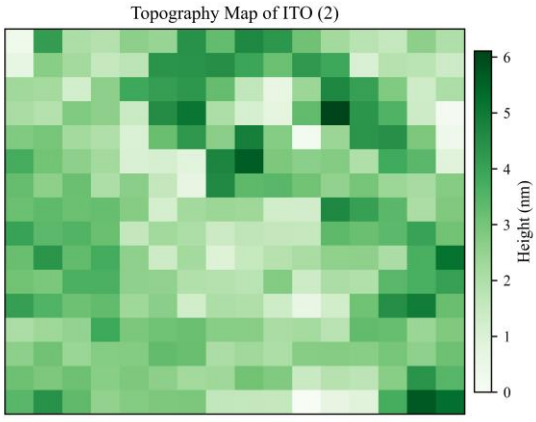
The following figures present a collection of topographical and thermal-signal maps for the investigated ITO samples. Topographical maps represent the averaged trace and retrace data, without slope removal, and all values were shifted so that the minimum height is set to zero for consistent comparison across samples. Thermal-signal maps also represent the averaged trace and retrace data. To improve visualization of the thermal-signal distribution, particularly given the close proximity of values, a diverging colormap (seismic) was employed to highlight positive and negative deviations with red–blue contrast. The colormap was discretized into nineteen levels to enhance visual distinction between ranges, while symmetric logarithmic normalization was applied to improve contrast.

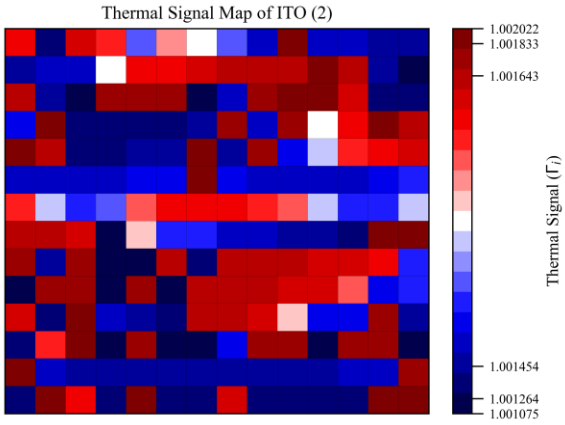
2×2 μm^2 Square Area for All Maps

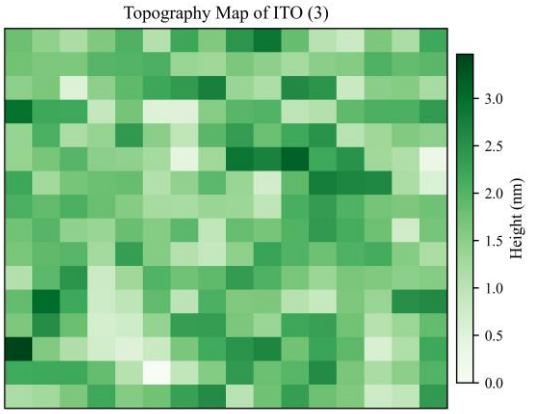


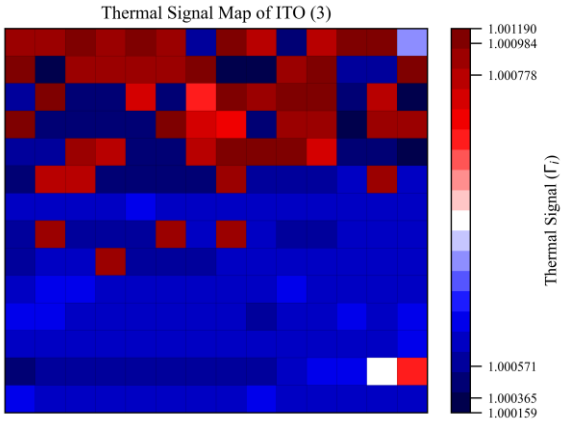
2×2 μm^2 Square Area for All Maps

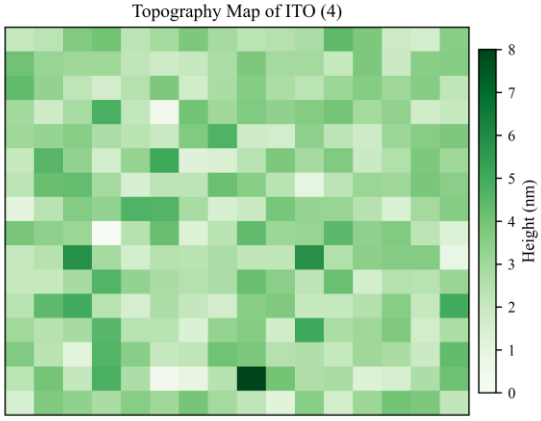


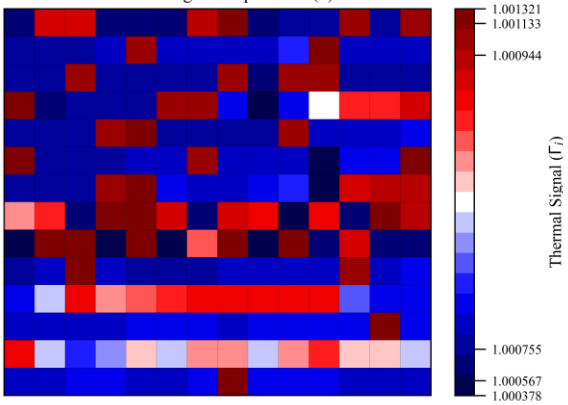


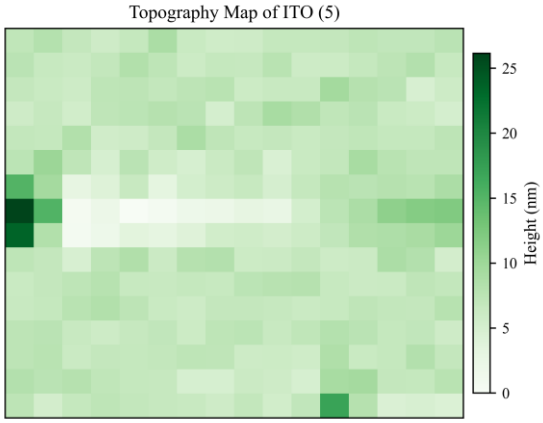


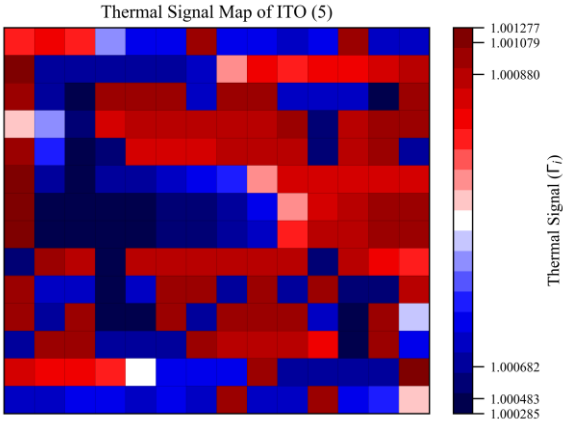

Topography Map of ITO (3)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Height (nm)

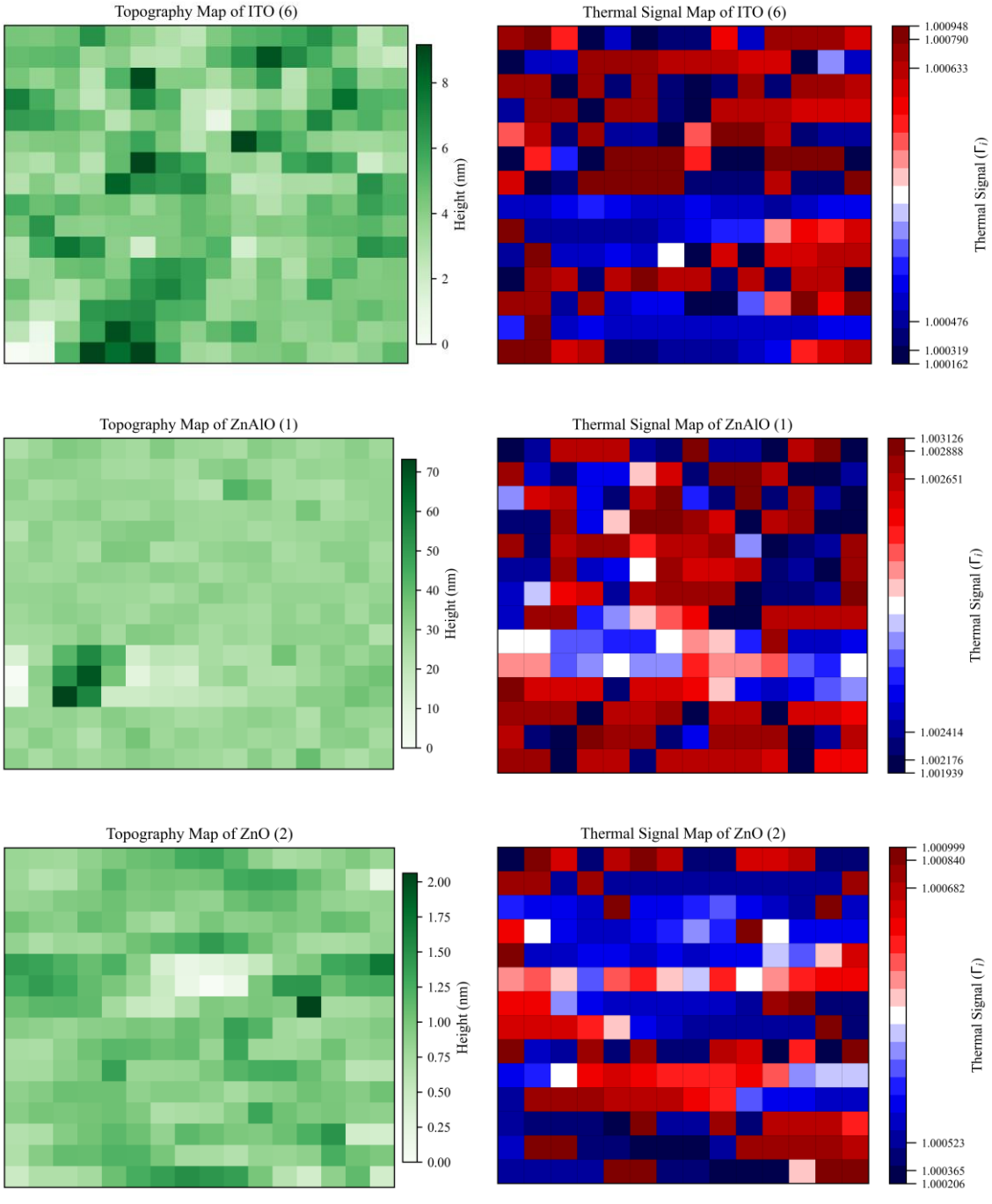

Thermal Signal Map of ITO (3)
1.001190
1.000984
1.000778
1.000571
1.000365
1.000159
Thermal Signal (Γ_i)

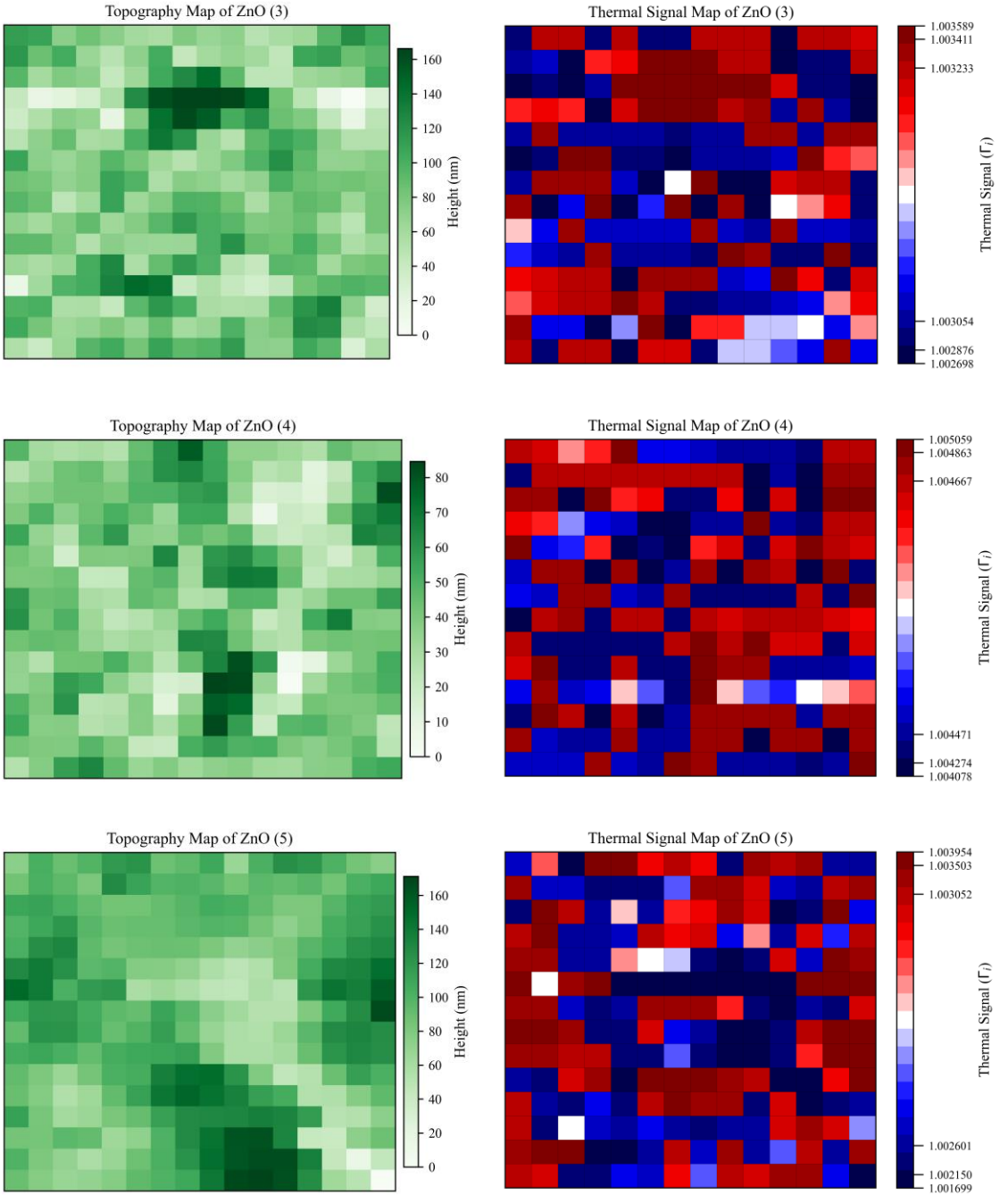

Topography Map of ITO (4)
8
7
6
5
4
3
2
1
0
Height (nm)

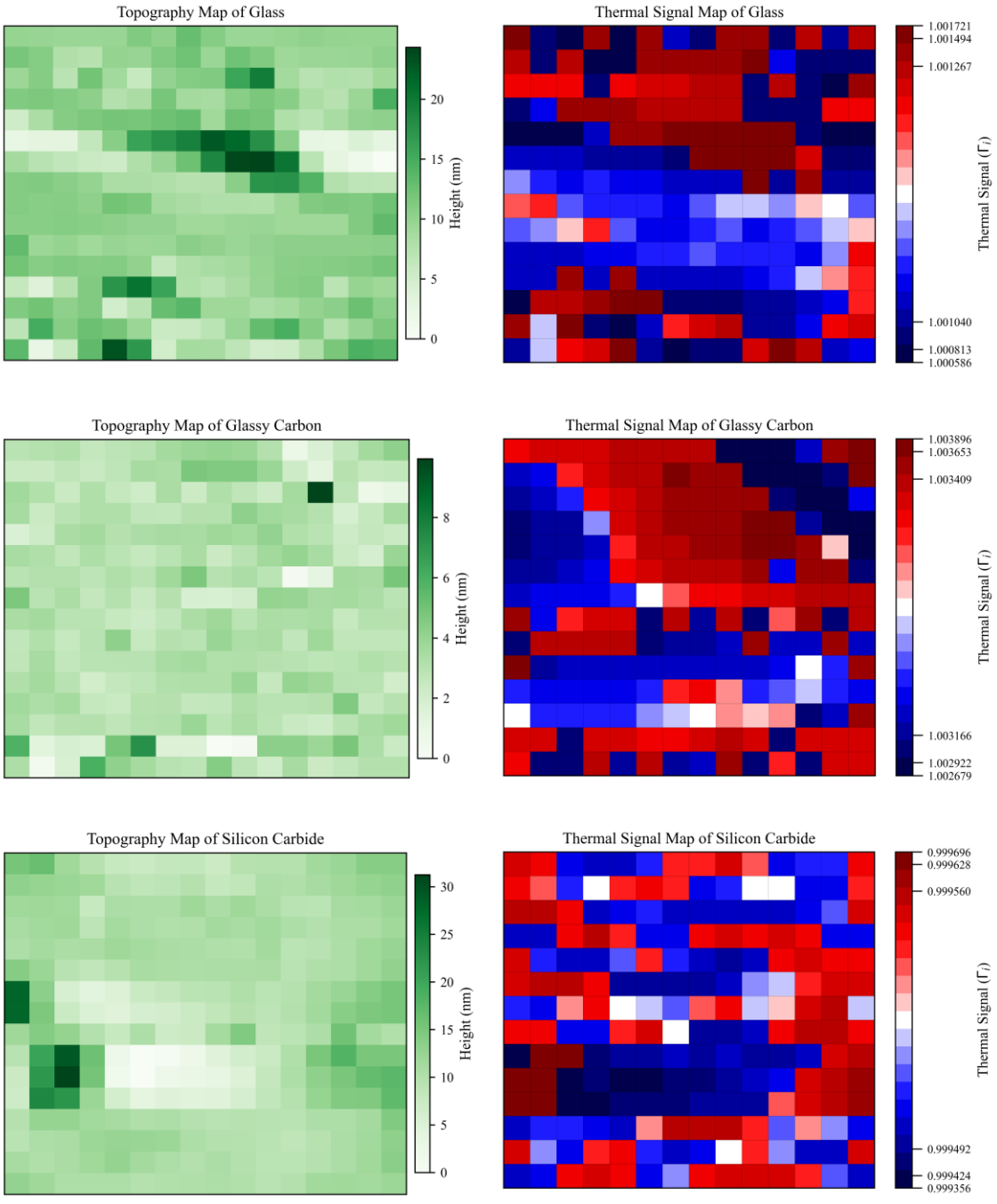

Thermal Signal Map of ITO (4)
1.001321
1.001133
1.000944
1.000755
1.000567
1.000378
Thermal Signal (Γ_i)

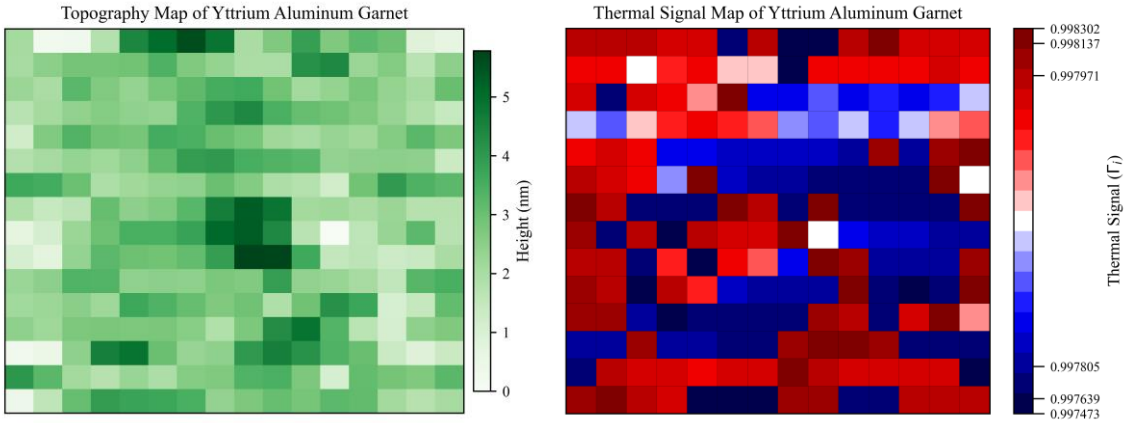

Topography Map of ITO (5)
25
20
15
10
5
0
Height (nm)

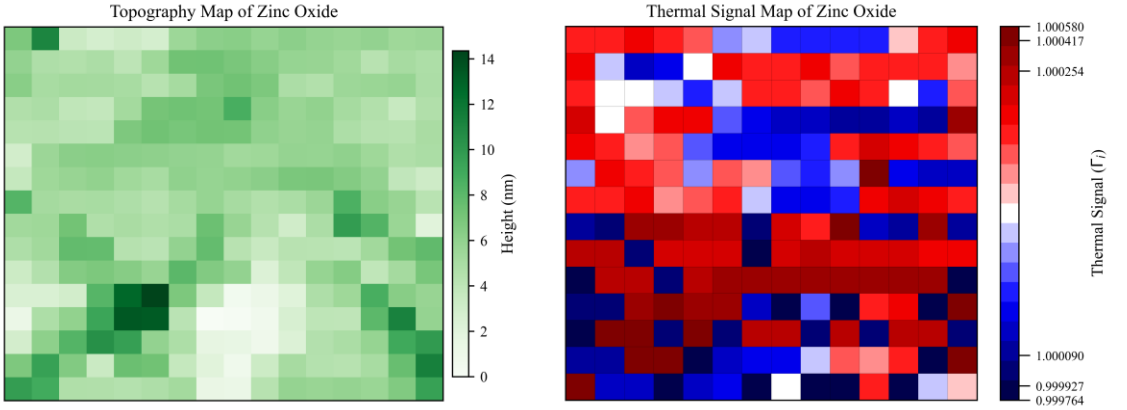

Thermal Signal Map of ITO (5)
1.001277
1.001079
1.000880
1.000682
1.000483
1.000285
Thermal Signal (Γ_i)

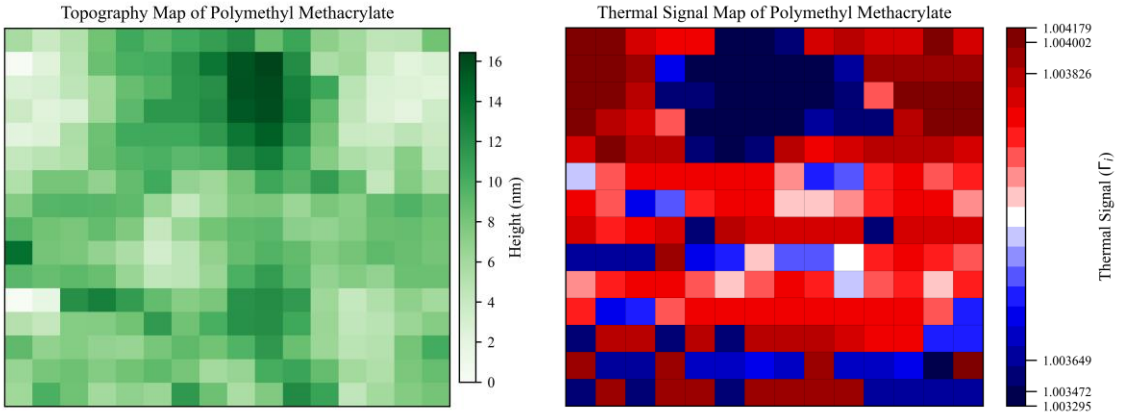

Topography Map of ITO (6)
Height (nm)
0
2
4
6
8
Thermal Signal Map of ITO (6)
Thermal Signal (Γ_i)
1.000948
1.000790
1.000633
1.000476
1.000319
1.000162
Topography Map of ZnAlO (1)
Height (nm)
0
10
20
30
40
50
60
70
Thermal Signal Map of ZnAlO (1)
Thermal Signal (Γ_i)
1.003126
1.002888
1.002651
1.002414
1.002176
1.001939
Topography Map of ZnO (2)
Height (nm)
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
Thermal Signal Map of ZnO (2)
Thermal Signal (Γ_i)
1.000999
1.000840
1.000682
1.000523
1.000365
1.000206


Topography Map of ZnO (3)
Height (nm)
160
140
120
100
80
60
40
20
0
Thermal Signal Map of ZnO (3)
Thermal Signal (Γ_I)
1.003589
1.003411
1.003233
1.003054
1.002876
1.002698
Topography Map of ZnO (4)
Height (nm)
80
70
60
50
40
30
20
10
0
Thermal Signal Map of ZnO (4)
Thermal Signal (Γ_I)
1.005059
1.004863
1.004667
1.004471
1.004274
1.004078
Topography Map of ZnO (5)
Height (nm)
160
140
120
100
80
60
40
20
0
Thermal Signal Map of ZnO (5)
Thermal Signal (Γ_I)
1.003954
1.003503
1.003052
1.002601
1.002150
1.001699


Topography Map of Glass
Height (nm)
0
5
10
15
20
Thermal Signal Map of Glass
Thermal Signal (Γ_I)
1.001721
1.001494
1.001267
1.001040
1.000813
1.000586
Topography Map of Glassy Carbon
Height (nm)
0
2
4
6
8
Thermal Signal Map of Glassy Carbon
Thermal Signal (Γ_I)
1.003896
1.003653
1.003409
1.003166
1.002922
1.002679
Topography Map of Silicon Carbide
Height (nm)
0
5
10
15
20
25
30
Thermal Signal Map of Silicon Carbide
Thermal Signal (Γ_I)
0.999696
0.999628
0.999560
0.999492
0.999424
0.999356

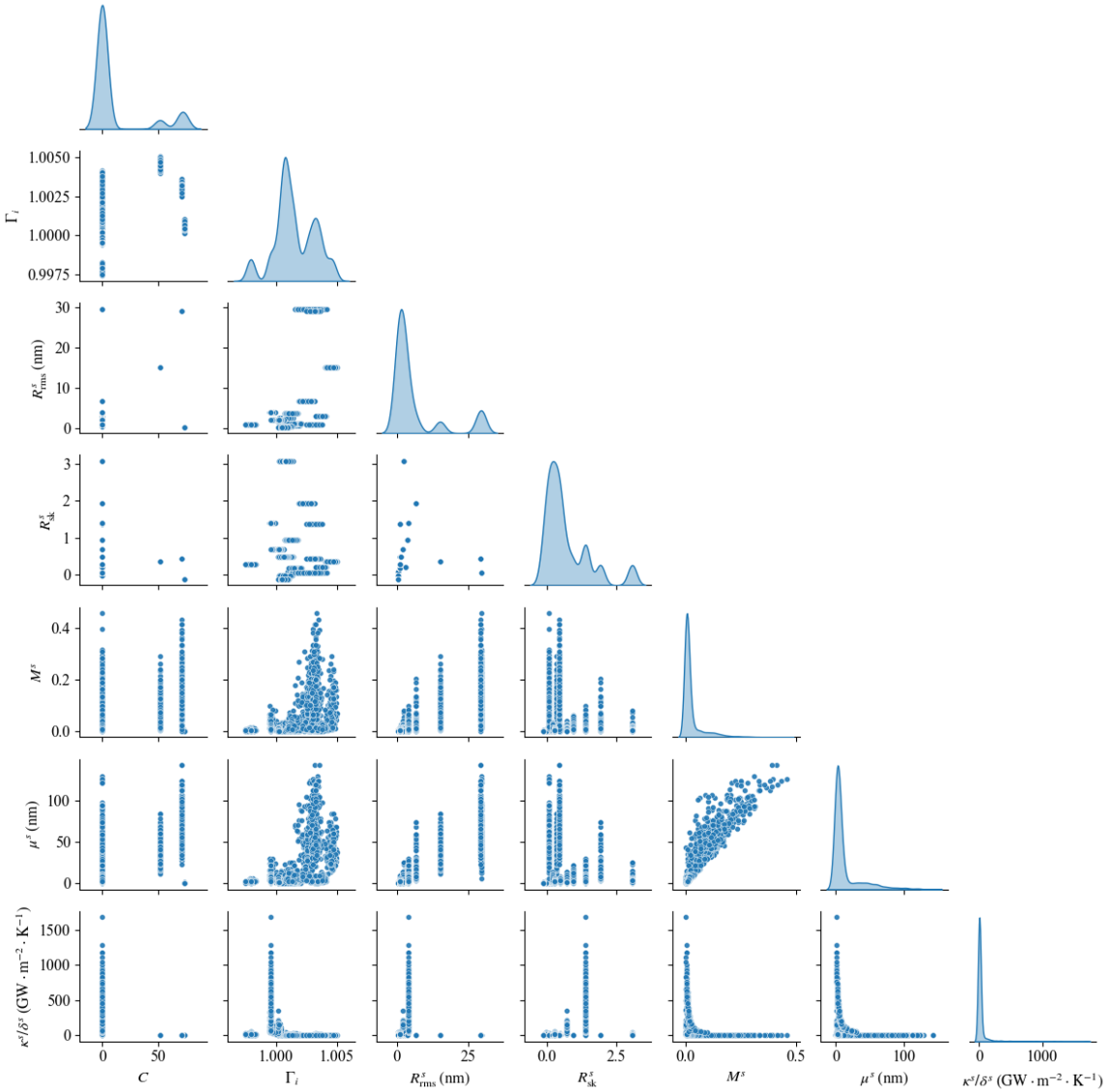

Topography Map of Yttrium Aluminum Garnet
Height (nm)
0
1
2
3
4
5
Thermal Signal Map of Yttrium Aluminum Garnet
Thermal Signal (Γ_i)
0.998302
0.998137
0.997971
0.997805
0.997639
0.997473


Topography Map of Zinc Oxide
Height (nm)
0
2
4
6
8
10
12
14
Thermal Signal Map of Zinc Oxide
Thermal Signal (Γ_i)
1.000580
1.000417
1.000254
1.000090
0.999927
0.999764


Topography Map of Polymethyl Methacrylate
Height (nm)
0
2
4
6
8
10
12
14
16
Thermal Signal Map of Polymethyl Methacrylate
Thermal Signal (Γ_i)
1.004179
1.004002
1.003826
1.003649
1.003472
1.003295

10.2. Pair Plot of Data for the Investigation of Thermal Properties

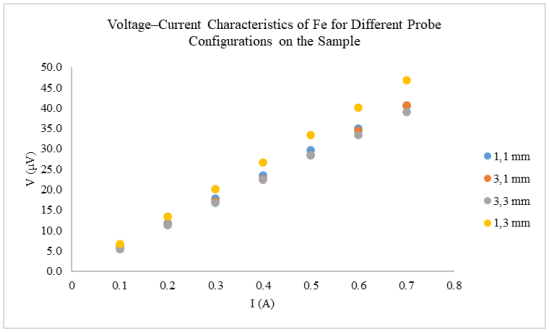
The pairplot illustrated here provides a comprehensive visualization of the experimental dataset's internal structure. The diagonal plots represent the distribution of each individual variable, allowing for a clear assessment of data spread. In contrast, the off-diagonal scatterplots map the relationship between every pair of variables, which is essential for identifying correlations within the system.

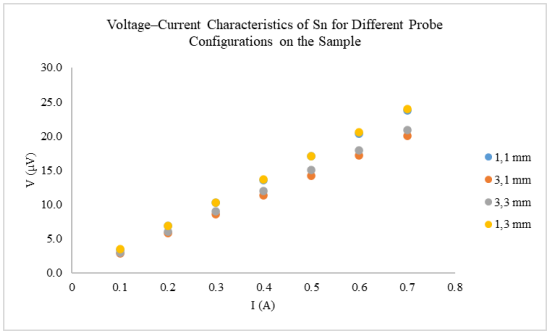


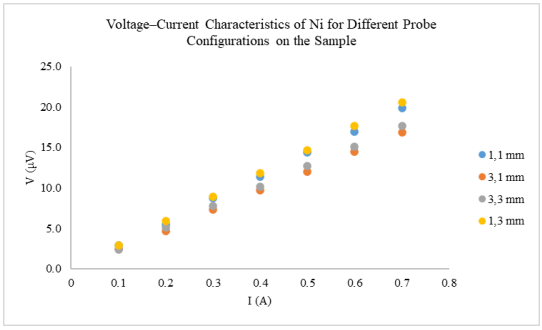
10.3. Supplementary Data for Electrical Conductivity

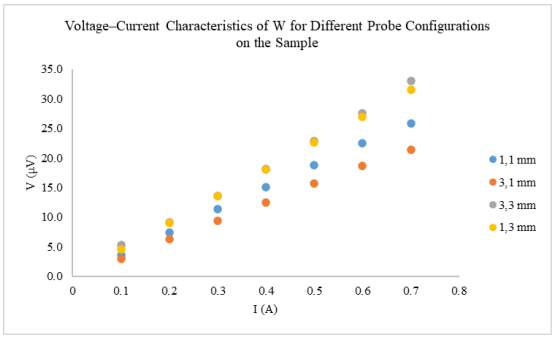
10.3.1. Voltage–Current Characteristics

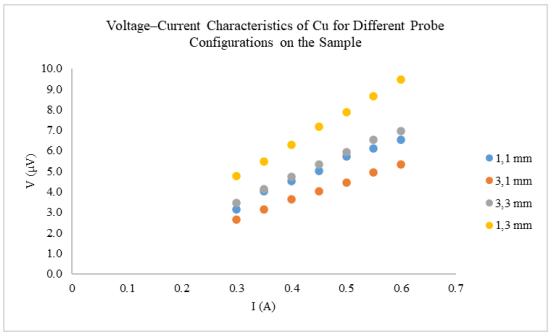
These graphs show the voltage–current characteristics of the different sample measured under different probe configurations on the samples. Each vertical label indicates a specific probe placement: the first number (before the comma) represents the vertical position of the probe L_v, and the second number (after the comma) represents the horizontal position of the probe L_h.

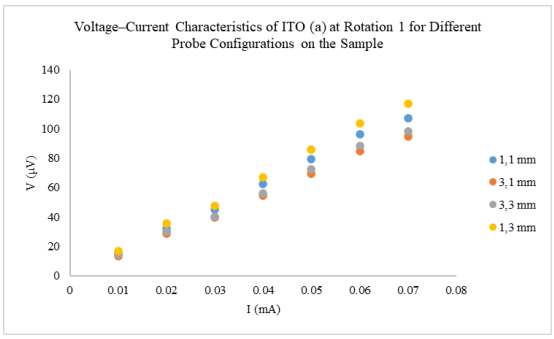


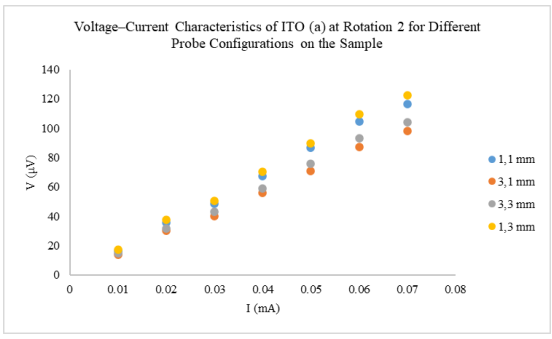


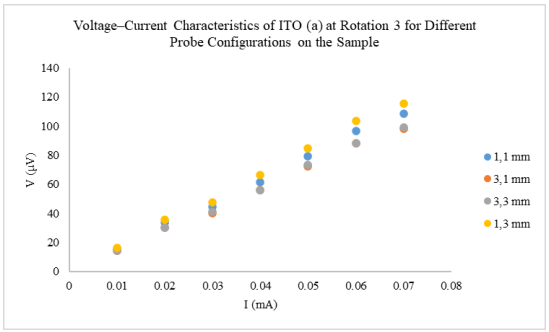


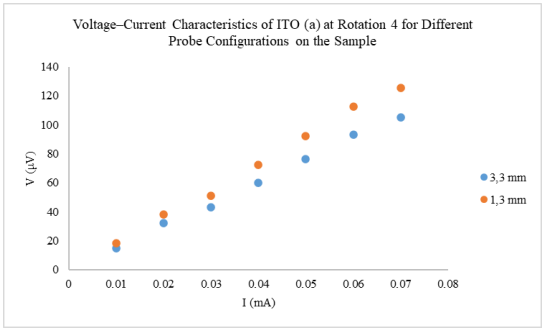


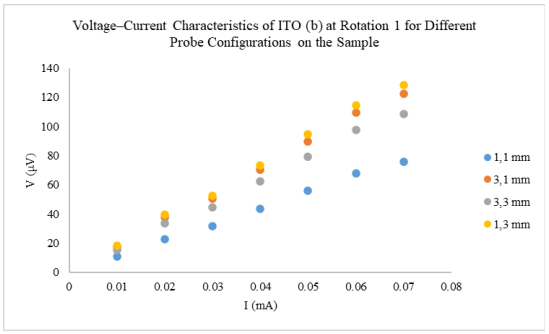


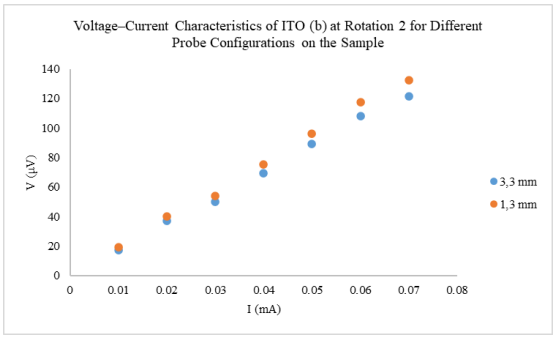


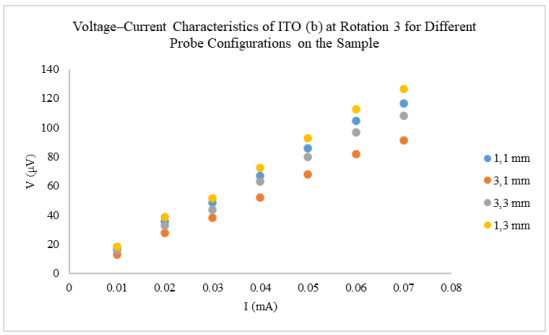


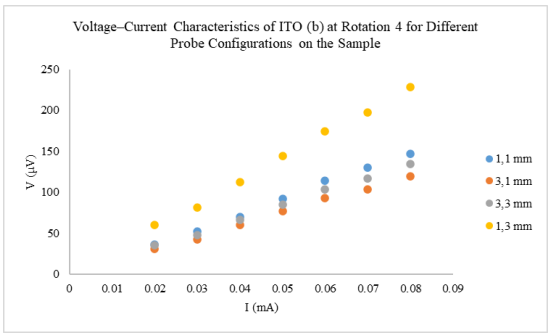

Voltage–Current Characteristics of ITO (a) at Rotation 3 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
3,1 mm
3,3 mm
1,3 mm

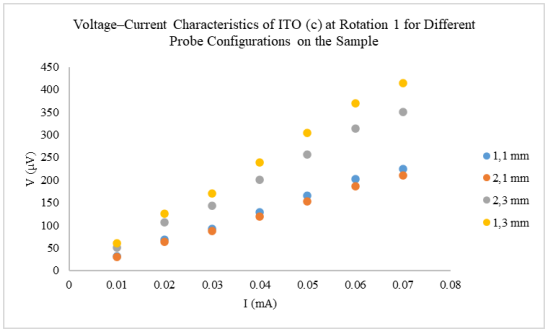

Voltage–Current Characteristics of ITO (a) at Rotation 4 for Different Probe Configurations on the Sample
V (μV)
I (mA)
3,3 mm
1,3 mm

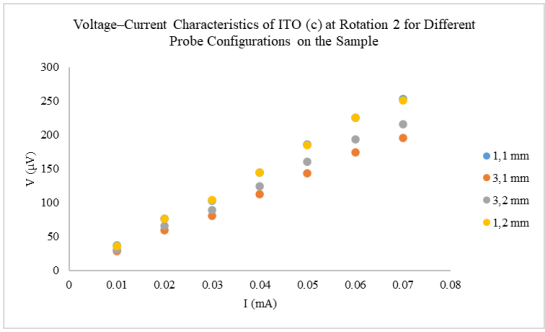

Voltage–Current Characteristics of ITO (b) at Rotation 1 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
3,1 mm
3,3 mm
1,3 mm

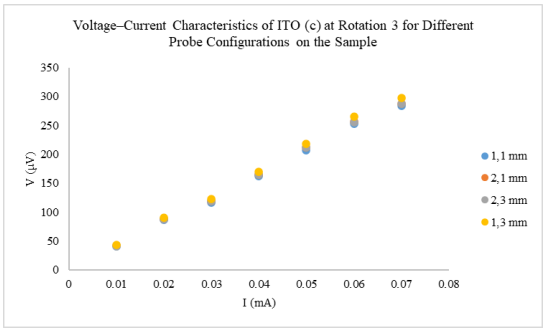

Voltage–Current Characteristics of ITO (b) at Rotation 2 for Different Probe Configurations on the Sample
V (μV)
I (mA)
3,3 mm
1,3 mm

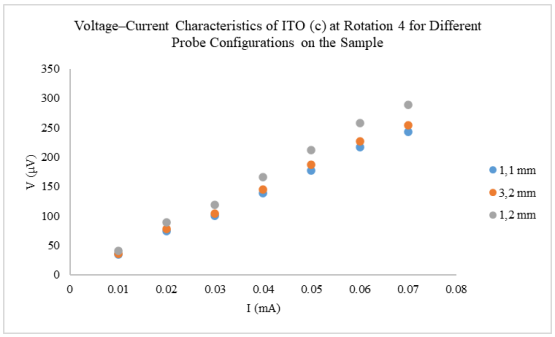

Voltage–Current Characteristics of ITO (b) at Rotation 3 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
3,1 mm
3,3 mm
1,3 mm

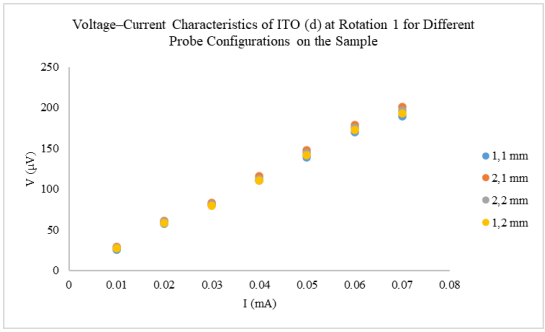

Voltage–Current Characteristics of ITO (b) at Rotation 4 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
3,1 mm
3,3 mm
1,3 mm

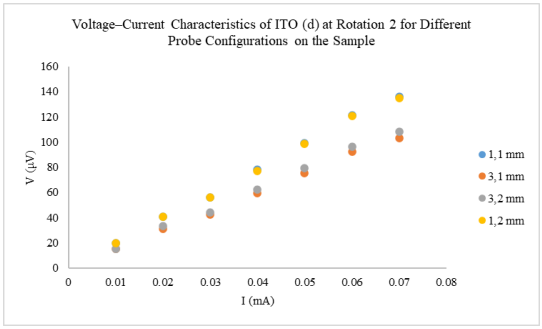

Voltage–Current Characteristics of ITO (c) at Rotation 1 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
2,1 mm
2,3 mm
1,3 mm

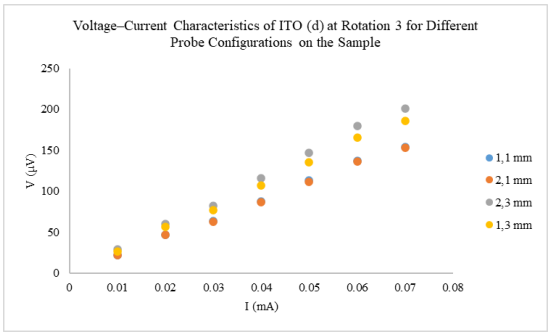

Voltage–Current Characteristics of ITO (c) at Rotation 2 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
3,1 mm
3,2 mm
1,2 mm

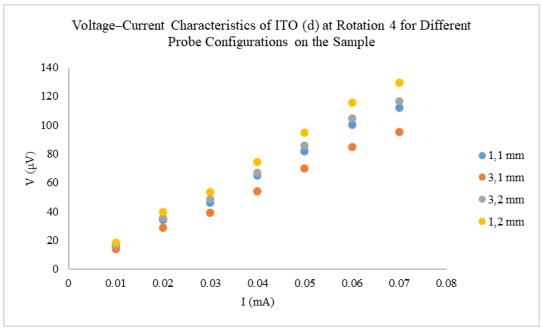

Voltage–Current Characteristics of ITO (c) at Rotation 3 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
2,1 mm
2,3 mm
1,3 mm


Voltage–Current Characteristics of ITO (c) at Rotation 4 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
3,2 mm
1,2 mm


Voltage–Current Characteristics of ITO (d) at Rotation 1 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
2,1 mm
2,2 mm
1,2 mm

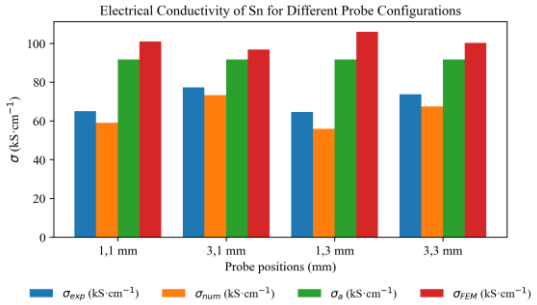

Voltage–Current Characteristics of ITO (d) at Rotation 2 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
3,1 mm
3,2 mm
1,2 mm

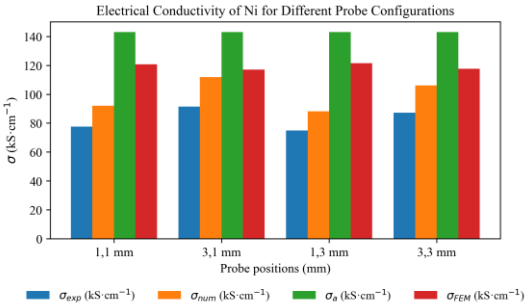

Voltage–Current Characteristics of ITO (d) at Rotation 3 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
2,1 mm
2,3 mm
1,3 mm

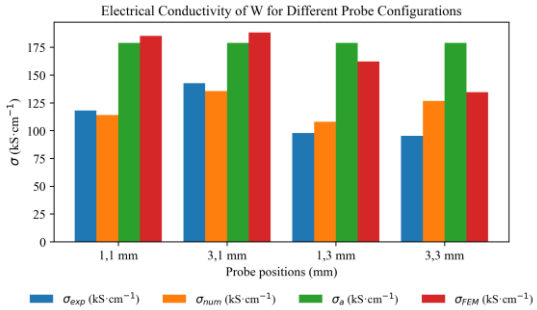

Voltage–Current Characteristics of ITO (d) at Rotation 4 for Different Probe Configurations on the Sample
V (μV)
I (mA)
1,1 mm
3,1 mm
3,2 mm
1,2 mm

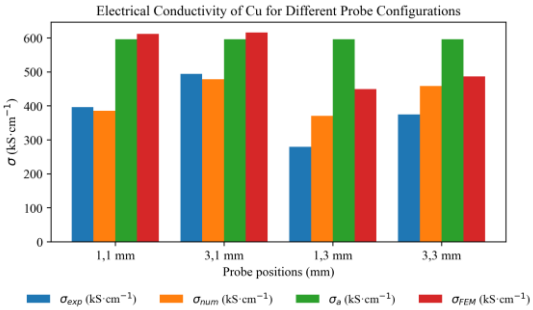
10.3.2. Comparison of Experimental, Numerical, Intrinsic, and FEM-Corrected EC

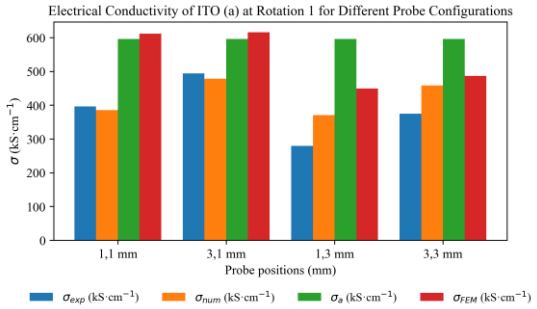
The following figures present the investigated EC obtained under different four-probe configurations on the samples. For each material, the bar charts compare the experimental conductivity (σ_{exp}) measured directly from the four-probe setup, the numerical conductivity (σ_{num}) calculated using a simulation-based model for the same probe geometry, the intrinsic conductivity (σ_a) representing the theoretical or literature reference value for the bulk material, and the FEM-corrected conductivity (σ_{FEM}) obtained using finite element method corrections to account for geometric and boundary effects.

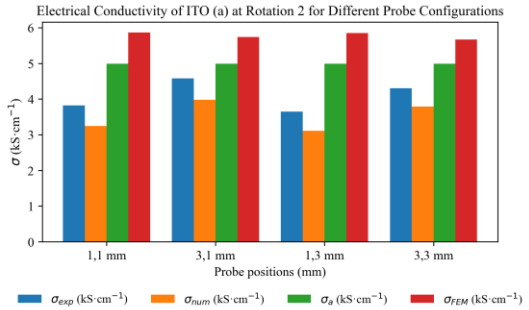


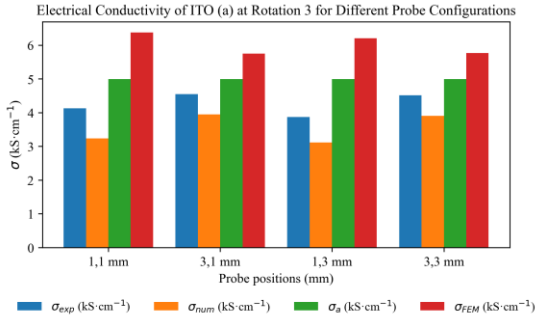


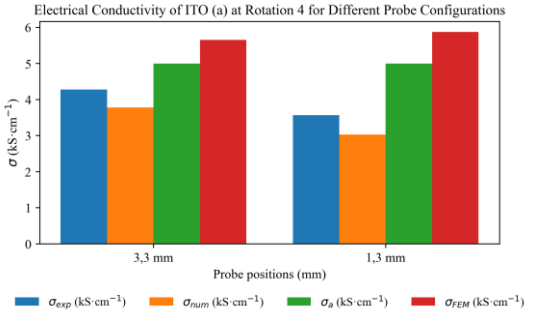


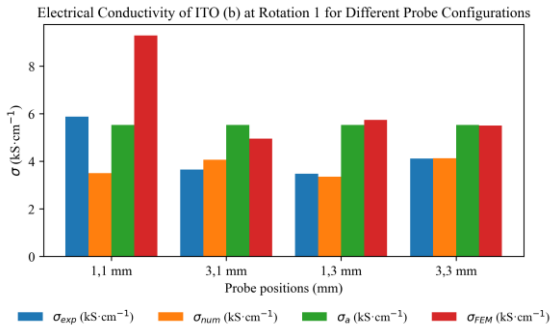


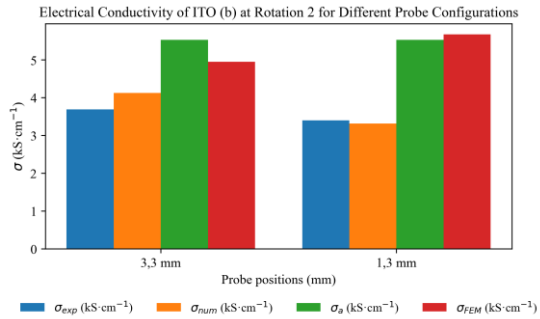


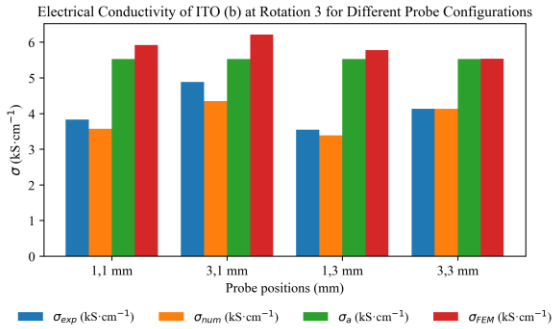


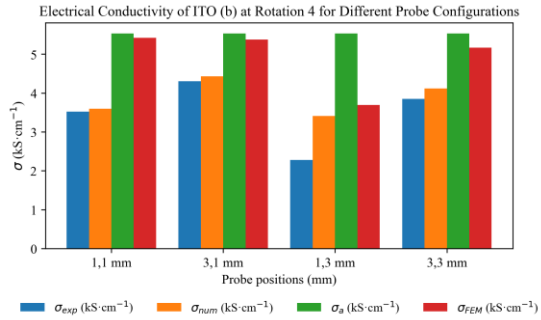

Electrical Conductivity of ITO (a) at Rotation 3 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
3,1 mm
1,3 mm
3,3 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

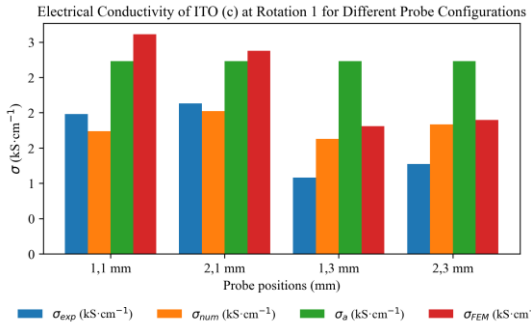

Electrical Conductivity of ITO (a) at Rotation 4 for Different Probe Configurations
σ (kS·cm⁻¹)
3,3 mm
1,3 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

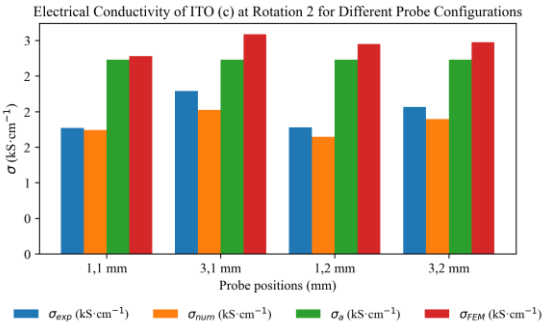

Electrical Conductivity of ITO (b) at Rotation 1 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
3,1 mm
1,3 mm
3,3 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

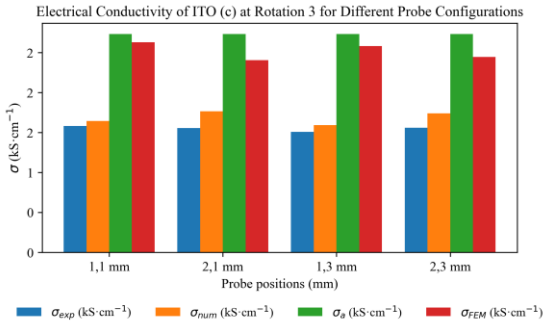

Electrical Conductivity of ITO (b) at Rotation 2 for Different Probe Configurations
σ (kS·cm⁻¹)
3,3 mm
1,3 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

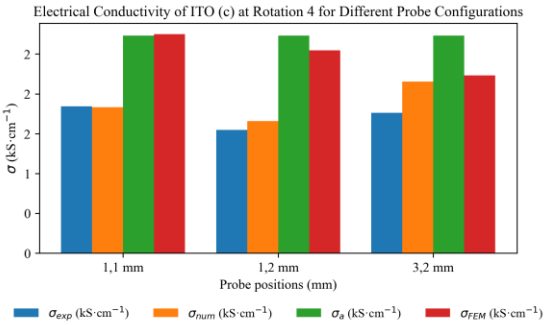

Electrical Conductivity of ITO (b) at Rotation 3 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
3,1 mm
1,3 mm
3,3 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

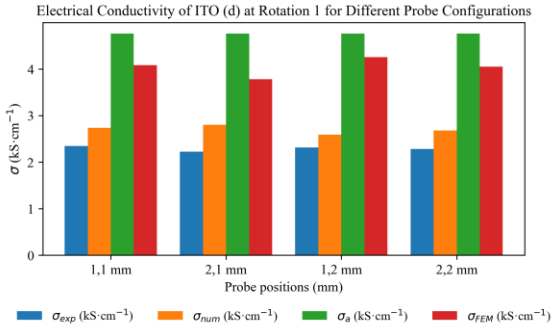

Electrical Conductivity of ITO (b) at Rotation 4 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
3,1 mm
1,3 mm
3,3 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

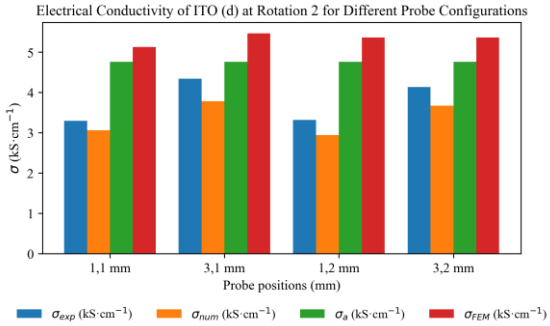

Electrical Conductivity of ITO (c) at Rotation 1 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
2,1 mm
1,3 mm
2,3 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

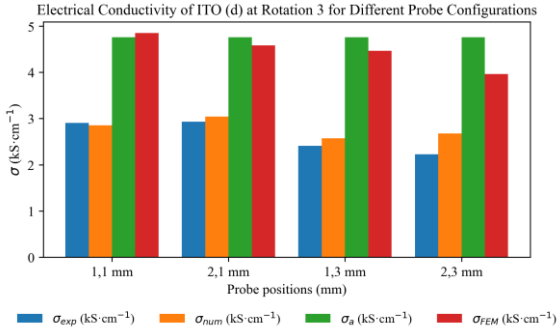

Electrical Conductivity of ITO (c) at Rotation 2 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
3,1 mm
1,2 mm
3,2 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

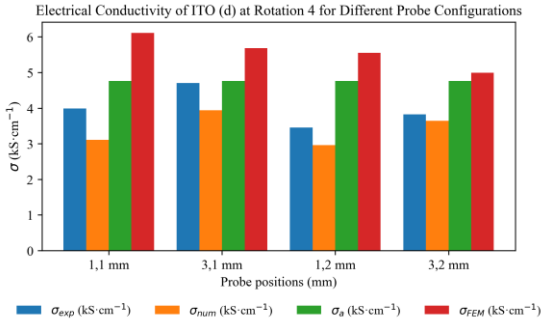

Electrical Conductivity of ITO (c) at Rotation 3 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
2,1 mm
1,3 mm
2,3 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)


Electrical Conductivity of ITO (c) at Rotation 4 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
1,2 mm
3,2 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)


Electrical Conductivity of ITO (d) at Rotation 1 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
2,1 mm
1,2 mm
2,2 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

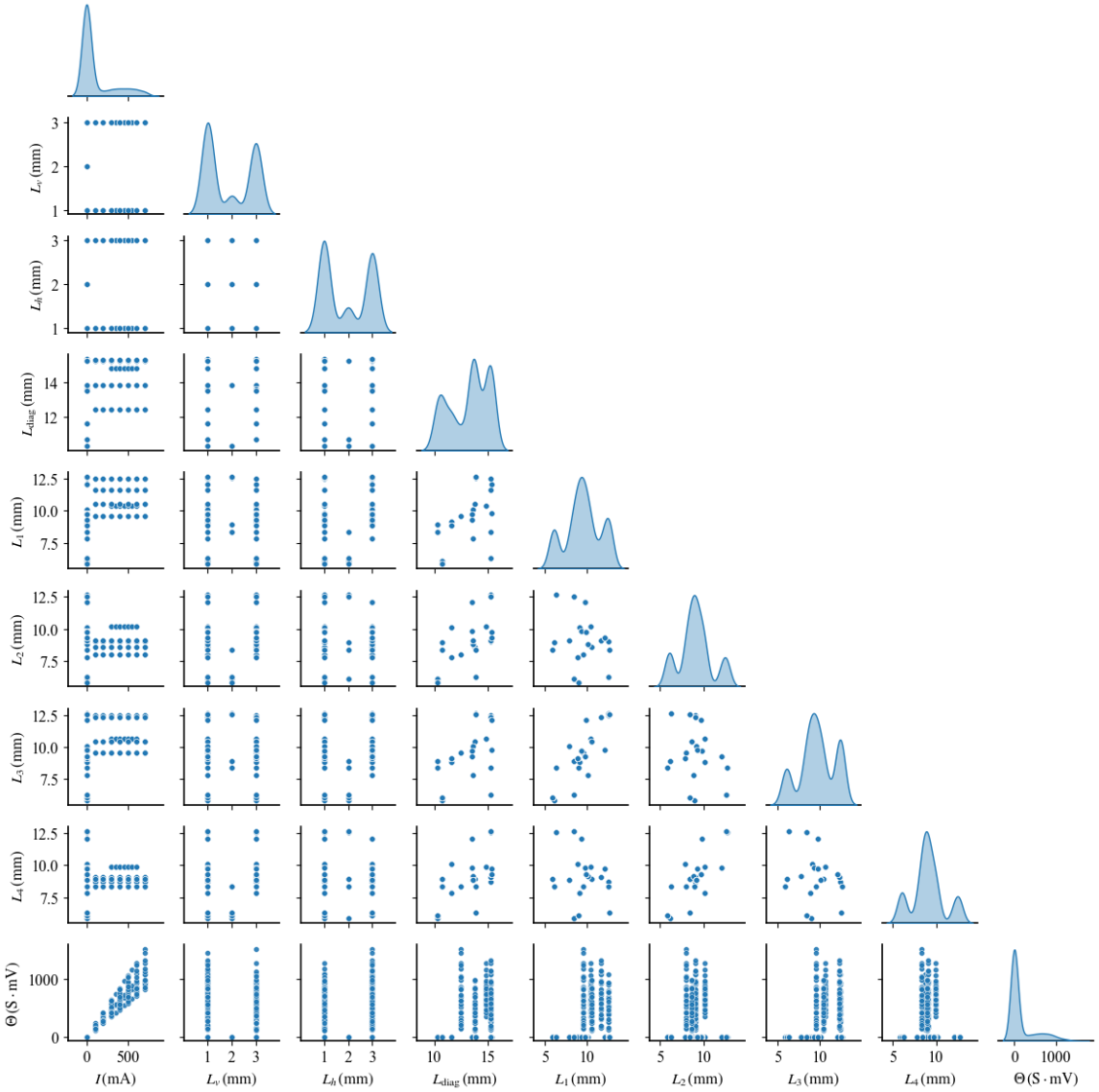

Electrical Conductivity of ITO (d) at Rotation 2 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
3,1 mm
1,2 mm
3,2 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)


Electrical Conductivity of ITO (d) at Rotation 3 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
2,1 mm
1,3 mm
2,3 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)


Electrical Conductivity of ITO (d) at Rotation 4 for Different Probe Configurations
σ (kS·cm⁻¹)
1,1 mm
3,1 mm
1,2 mm
3,2 mm
Probe positions (mm)
σexp (kS·cm⁻¹)
σnum (kS·cm⁻¹)
σa (kS·cm⁻¹)
σFEM (kS·cm⁻¹)

10.4. Pair Plot of Data for the Investigation of Electrical Properties

The following figure presents a pairplot of all measured parameters relevant to the electrical properties' investigation. This visualization provides an overview of correlations, clustering tendencies, and variations across the electrical dataset, offering insight into dependencies between the measured electrical quantities.



11. Appendix B:

11.1. Python Libraries Used for ML

The ML component of this research employs following open-source libraries:

Data Handling and Processing:

- **NumPy (v1.24.0):** Provided efficient numerical computation capabilities and array operations.
- **Pandas (v2.1.0):** Facilitated data manipulation, cleaning, and analysis using structured data frames.
- **SciPy (v1.11.0):** Offered advanced scientific computing functions, including optimization, interpolation, and statistical analysis.

Machine Learning:

- **scikit-learn (v1.3.0):** Implemented a wide range of classical ML algorithms, including regression, classification, and clustering methods.
- **Random Forest (via scikit-learn):** Ensemble learning method for classification and regression tasks.
- **Gradient Boosting (via scikit-learn):** ML technique for regression and classification problems.

Data Visualization:

- **Matplotlib (v3.7.0):** Provided a wide variety of 2D plotting and visualization options.
- **Seaborn (v0.12.0):** Simplified the creation of statistically-informed visualizations.

Specialized Tools and Utilities:

- **joblib (v1.3.0):** Supported efficient saving of models and parallel computation.

Development Environment:

- **JupyterLab (v4.0.0):** The primary interactive development environment used for this research, enabling coding, visualization, and documentation within notebooks.

These libraries collectively enabled comprehensive data preprocessing, visualization, model training, and evaluation for the ML components of this study.

11.2. Software Implementation and Algorithmic Sources

All ML algorithms utilized in this study rely on pre-built, validated implementations from recognized open-source ML libraries. These software packages closely follow the original algorithmic definitions presented in foundational literature (Pedregosa et al. [74]).

- Random Forest and Gradient Boosting algorithms are applied using their reliable, pre-built implementations within the scikit-learn Python library. These implementations were developed by Pedregosa et al. as part of the INRIA team.
- All model-selection tools, including cross-validation and Grid Search/Randomized Search, also rely on the validated implementations within scikit-learn (Pedregosa et al., INRIA [74]).

Using these established software packages ensures the consistency and adherence to community-validated standards throughout the modeling process.

12. Appendix C:

During my PhD studies, I have contributed to the following publications, conferences, and academic activities.

12.1. List of Publications

1- **Dehbashi, M.**, Kazmierczak-Balata, A., & Bodzenta, J. (2025). An approach for thermal conductivity measurements in thin films: Combining localized surface topography, thermal analysis, and machine learning techniques. International Journal of Heat and Mass Transfer, 248, 127215. https://doi.org/10.1016/j.ijheatmasstransfer.2025.127215

2- Kaźmierczak-Bałata, A., Bodzenta, J., **Dehbashi, M.**, Mayandi, J., & Venkatachalapathy, V. (2022c). Influence of post processing on thermal conductivity of ITO thin films. Materials, 16(1), 362. https://doi.org/10.3390/ma16010362

12.2. Conferences

1- **Dehbashi, M.**, Kaźmierczak-Bałata, A., & Bodzenta, J. (2025, September 15–18). *Machine Learning-Enhanced Thermal Conductivity Measurements in Thin Films: Integrating Nanoscale Surface Topography and Thermal Analysis* (Abstract No. 00093). European Materials Research Society (E-MRS) Fall Meeting, Warsaw, Poland. Oral.

12.3. Monograph Chapters

1- **Dehbashi, M.**, Kaźmierczak-Bałata, A., Szymska, W., Bodzenta, J. (2023). Numerical investigation of the influence of sample geometry on the accuracy of electrical conductivity measurements using the four-point probe method. In B. Balon (Ed.), *Interdyscyplinarne badania młodych naukowców* (Monografia / Politechnika Śląska, No. 987, pp. 81–91). Wydawnictwo Politechniki Śląskiej. ISBN 978-83-7880-905-0. Oral.
https://repolis.bg.polsl.pl/dlibra/publication/86320/edition/76772/content

12.4. Awards

1- Pro-Quality Grant Award (Excellence Initiative – Research University), Silesian University of Technology — *December 2022* Awarded 12,000 PLN to support commencement of scientific activity for the project "Thin Metal Oxides Layers for Thermoelectric Applications" (Project No. 32/014/SDU/10-22-55).

2- Young Researcher Award, E-MRS Fall Meeting 2025 (Warsaw, Poland) — *September 2025*. For the paper "Machine Learning-Enhanced Thermal Conductivity Measurements in Thin Films: Integrating Nanoscale Surface Topography and Thermal Analysis." (Symposium G: Artificial Intelligence to accelerate the development of new advanced materials for energy).

Young Researcher Award

presented to

Mohsen Dehbashi

In recognition of an outstanding paper entitled

"Machine Learning-Enhanced Thermal Conductivity Measurements in Thin Films: Integrating Nanoscale Surface Topography and Thermal Analysis"

Contributed to the E-MRS 2025 Fall Meeting.

Symposium G
"Artificial Intelligence to accelerate the development of new advanced materials for energy"

Giuseppina PADELETTI
President of E-MRS

September 2025, Warsaw University of Technology, Poland

Warsaw University of Technology

E-MRS

12.5. Other Activities

12.5.1. Invited Lecturer (College Physics), Yanshan University, China

Delivered in-person invited lectures within the Sino-Polish cooperative education program (Yanshan University & Silesian University of Technology); taught 360 teaching hours total across Fall 2023, Spring 2024, and Autumn 2024.

燕山大學 YANSHAN UNIVERSITY

CERTIFICATE

This is to certify that MOHSEN DEHBASHI (passport number Z97568345) delivered in-person an invited 96 teaching hours (45 minutes per teaching hour) of lectures for College Physics in the autumn semester of academic year 2024(from 1st September to 3rd November), 192 teaching hours in the spring semester of 2024 (from 29th April to 30th June) and 72 teaching hours in the fall semester of academic year 2023 (from 11th December to 29th December), here at Yanshan University campus in the framework of Sino-Polish cooperative educational program operated by Yanshan University and Silesian University of Technology.

During his stay he completed the teaching task with outstanding performance, and he received unanimous recognition from students and Chinese professors.

This certification is being issued to MOHSEN DEHBASHI for whatever legal purpose it may serve.

Issued on the day of 19th, November, 2024 at Yanshan University, 438 Hebei Ave West Section, Haigang District, Qinhuangdao, Hebei, China, 066000

Liu Xueqiang

Dean's Office

Silesian College of Intelligent Science and Engineering at YSU

西里西亚智能科学与工程学院 1303021096263

地址：中国河北省秦皇岛市河北大街西段438号(066004)
电话：0335-8057100
传真：0335-8051148

Add: 438 West Hebei Avenue, Qinhuangdao City, Hebei Province, China Zip Code: 066004
Tel: +86 335 805 7100 Fax: +86 335 805 1148

12.5.2. Participations in Other Publications

1- Vaferi, B., **Dehbashi, M.**, & Alibak, A. H. (2024). Cutting-Edge machine learning techniques for accurate prediction of agglomeration size in Water–Alumina nanofluids. Symmetry, 16(7), 804. https://doi.org/10.3390/sym16070804

2- Vaferi, B., **Dehbashi, M.**, Khandakar, A., Ayari, M. A., & Amini, S. (2024). Development of a stacked machine learning model to compute the capability of ZnO-based sensors for hydrogen detection. Sustainable Materials and Technologies, 39, e00863. https://doi.org/10.1016/j.susmat.2024.e00863

3- Javadijam, R., **Dehbashi, M.**, Shahverdian, M. H., Sohani, A., Arıcı, M., & Sayyaadi, H. (2024). Artificial intelligent based techno-economic-exergetic optimization of a thermoelectric enhanced building integrated photovoltaic thermal system. Journal of Building Engineering, 84, 108526. https://doi.org/10.1016/j.jobe.2024.108526

13. References

[1] Jiang, P., Qian, X., & Yang, R. (2018). Tutorial: Time-domain thermoreflectance (TDTR) for thermal property characterization of bulk and thin film materials. *Journal of Applied Physics*, *124*(16). https://doi.org/10.1063/1.5046944

[2] Jiang, P., Qian, X., & Yang, R. (2017). Time-domain thermoreflectance (TDTR) measurements of anisotropic thermal conductivity using a variable spot size approach. *Review of Scientific Instruments*, *88*(7), 074901. https://doi.org/10.1063/1.4991715

[3] Yan, S., Dong, C., Miao, T., Wang, W., Ma, W., Zhang, X., Kohno, M., & Takata, Y. (2016). Long delay time study of thermal transport and thermal stress in thin Pt film-glass substrate system by time-domain thermoreflectance measurements. *Applied Thermal Engineering*, *111*, 1433–1440. https://doi.org/10.1016/j.applthermaleng.2016.08.110

[4] Cahill, D. G., Ford, W. K., Goodson, K. E., Mahan, G. D., Majumdar, A., Maris, H. J., Merlin, R., & Phillpot, S. R. (2003). Nanoscale thermal transport. *Journal of Applied Physics*, *93*(2), 793–818. https://doi.org/10.1063/1.1524305

[5] López-Honorato, E., Chiritescu, C., Xiao, P., Cahill, D. G., Marsh, G., & Abram, T. (2008). Thermal conductivity mapping of pyrolytic carbon and silicon carbide coatings on simulated fuel particles by time-domain thermoreflectance. *Journal of Nuclear Materials*, *378*(1), 35–39. https://doi.org/10.1016/j.jnucmat.2008.04.007

[6] Cahill, D. G. (2004). Analysis of heat flow in layered structures for time-domain thermoreflectance. *Review of Scientific Instruments*, *75*(12), 5119–5122. https://doi.org/10.1063/1.1819431

[7] Feser, J. P., & Cahill, D. G. (2012). Probing anisotropic heat transport using time-domain thermoreflectance with offset laser spots. *Review of Scientific Instruments*, *83*(10). https://doi.org/10.1063/1.4757863

[8] Parker, W. J., Jenkins, R. J., Butler, C. P., & Abbott, G. L. (1961). Flash method of determining thermal diffusivity, heat capacity, and thermal conductivity. *Journal of Applied Physics*, *32*(9), 1679–1684. https://doi.org/10.1063/1.1728417

[9] Wilson, R. B., Apgar, B. A., Martin, L. W., & Cahill, D. G. (2012). Thermoreflectance of metal transducers for optical pump-probe studies of thermal properties. *Optics Express*, *20*(27), 28829. https://doi.org/10.1364/oe.20.028829

[10] Olson, D. H., Braun, J. L., & Hopkins, P. E. (2019). Spatially resolved thermoreflectance techniques for thermal conductivity measurements from the nanoscale to the mesoscale. *Journal of Applied Physics*, *126*(15). https://doi.org/10.1063/1.5120310

[11] Sood, A., Cheaito, R., Bai, T., Kwon, H., Wang, Y., Li, C., Yates, L., Bougher, T., Graham, S., Asheghi, M., Goorsky, M., & Goodson, K. E. (2018). Direct visualization of thermal conductivity suppression due to enhanced phonon scattering near individual grain boundaries. *Nano Letters*, *18*(6), 3466–3472. https://doi.org/10.1021/acs.nanolett.8b00534

[12] Zhang, Y., Zhu, W., Hui, F., Lanza, M., Borca-Tasciuc, T., & Rojo, M. M. (2019b). A review on Principles and Applications of Scanning Thermal Microscopy (STHM). *Advanced Functional Materials*, *30*(18). https://doi.org/10.1002/adfm.201900892

[13] Christofferson, J., Maize, K., Ezzahri, Y., Shabani, J., Wang, X., & Shakouri, A. (2008). Microscale and nanoscale thermal characterization techniques. *Journal of Electronic Packaging*, *130*(4). https://doi.org/10.1115/1.2993145

[14] Gomès, S., Assy, A., & Chapuis, P. (2015). Scanning thermal microscopy: A review. *Physica Status Solidi (A)*, *212*(3), 477–494. https://doi.org/10.1002/pssa.201400360

[15] Bodzenta, J., Kaźmierczak-Bałata, A., & Harris, K. (2020). Quantitative thermal measurement by the use of scanning thermal microscope and resistive thermal probes. *Journal of Applied Physics*, *127*(3). https://doi.org/10.1063/1.5125062

[16] Martinek, J., Klapetek, P., & Campbell, A. C. (2015). Methods for topography artifacts compensation in scanning thermal microscopy. *Ultramicroscopy*, *155*, 55–61. https://doi.org/10.1016/j.ultramic.2015.04.011

[17] Guen, E., Klapetek, P., Puttock, R., Hay, B., Allard, A., Maxwell, T., Chapuis, P., Renahy, D., Davee, G., Valtr, M., Martinek, J., Kazakova, O., & Gomès, S. (2020). SThM-based local thermomechanical analysis: Measurement intercomparison and uncertainty analysis. *International Journal of Thermal Sciences*, *156*, 106502. https://doi.org/10.1016/j.ijthermalsci.2020.106502

[18] Klapetek, P., Martinek, J., Grolich, P., Valtr, M., & Kaur, N. J. (2017). Graphics cards based topography artefacts simulations in Scanning Thermal Microscopy. *International Journal of Heat and Mass Transfer*, *108*, 841–850. https://doi.org/10.1016/j.ijheatmasstransfer.2016.12.036

[19] Zhang, Q., Zhu, W., Zhou, J., & Deng, Y. (2023). Realizing the Accurate Measurements of Thermal Conductivity over a Wide Range by Scanning Thermal Microscopy Combined with Quantitative Prediction of TCR. *Small*, *19*(32). https://doi.org/10.1002/smll.202300968

[20] Altes, A., Mutamba, K., Heiderhoff, R., Hartnagel, H., & Balk, L. (2004). Scanning near field thermal microscopy on a micromachined thin membrane. *Superlattices and Microstructures*, *35*(3–6), 465–476. https://doi.org/10.1016/j.spmi.2003.09.008

[21] Ge, Y., Zhang, Y., Weaver, J. M. R., & Dobson, P. S. (2017). Dimension- and shape-dependent thermal transport in nano-patterned thin films investigated by scanning thermal microscopy. *Nanotechnology*, *28*(48), 485706. https://doi.org/10.1088/1361-6528/aa93cf

[22] Ge, Y., Zhang, Y., Weaver, J. M. R., Zhou, H., & Dobson, P. S. (2015). Topography-free sample for thermal spatial response measurement of scanning thermal microscopy. *Journal of Vacuum Science & Technology B Nanotechnology and Microelectronics Materials Processing Measurement and Phenomena*, *33*(6). https://doi.org/10.1116/1.4933172

[23] Sachat, A. E., Alzina, F., Torres, C. M. S., & Chavez-Angel, E. (2021). Heat Transport Control and Thermal Characterization of Low-Dimensional Materials: A review. *Nanomaterials*, *11*(1), 175. https://doi.org/10.3390/nano11010175

[24] Assy, A., & Gomès, S. (2015). Heat transfer at nanoscale contacts investigated with scanning thermal microscopy. *Applied Physics Letters*, *107*(4). https://doi.org/10.1063/1.4927653

[25] McNamara, S., Basu, A. S., Lee, J., & Gianchandani, Y. B. (2004). Ultracompliant thermal probe array for scanning non-planar surfaces without force feedback. *Journal of Micromechanics and Microengineering*, *15*(1), 237–243. https://doi.org/10.1088/0960-1317/15/1/033

[26] Guen, E., Chapuis, P., Kaur, N. J., Klapetek, P., & Gomés, S. (2021b). Impact of roughness on heat conduction involving nanocontacts. *Applied Physics Letters*, *119*(16). https://doi.org/10.1063/5.0064244

[27] Guen, E., Chapuis, P., Rajkumar, R., Dobson, P. S., Mills, G., Weaver, J. M. R., & Gomés, S. (2020). Scanning thermal microscopy on samples of varying effective thermal conductivities and identical flat surfaces. *Journal of Applied Physics*, *128*(23). https://doi.org/10.1063/5.0020276

[28] Leitgeb, V., Hammer, R., Mitterhuber, L., Fladischer, K., Peter, F., Buerke, A., & Defregger, S. (2021). Thermal and spatial resolution in scanning thermal microscopy images: A study on the probe's heating parameters. *Journal of Applied Physics*, *129*(16). https://doi.org/10.1063/5.0037983

[29] Lees, J., Corbetta, M., Kleine-Boymann, M., Scheidemann, A., Poon, S. W., & Thompson, S. M. (2024). Preparation of multilayer samples for scanning thermal microscopy examination. *Nanotechnology*, *35*(22), 225702. https://doi.org/10.1088/1361-6528/ad2bce

[30] Dong, L., & Li, Y. (2022b). Experimental identification of topography-based artifact phenomenon for micro-/nanoscale thermal characterization of polymeric materials in scanning thermal microscopy. *AIP Advances*, *12*(4). https://doi.org/10.1063/5.0088360

[31] Wilson, A. A. (2019). Scanning thermal probe calibration for accurate measurement of thermal conductivity of ultrathin films. *MRS Communications*, *9*(2), 650–656. https://doi.org/10.1557/mrc.2019.37

[32] David, L., Gomès, S., & Raynaud, M. (2007). Modelling for the thermal characterization of solid materials by dc scanning thermal microscopy. *Journal of Physics D Applied Physics*, *40*(14), 4337–4346. https://doi.org/10.1088/0022-3727/40/14/032

[33] Bodzenta, J., Juszczyk, J., Kaźmierczak-Bałata, A., Firek, P., Fleming, A., & Chirtoc, M. (2016). Quantitative Thermal Microscopy Measurement with Thermal Probe Driven by dc+ac Current. *International Journal of Thermophysics*, *37*(7). https://doi.org/10.1007/s10765-016-2080-y

[34] Miccoli, I., Edler, F., Pfnür, H., & Tegenkamp, C. (2015). The 100th anniversary of the four-point probe technique: the role of probe geometries in isotropic and anisotropic systems. Journal of Physics Condensed Matter, 27(22), 223201. https://doi.org/10.1088/0953-8984/27/22/223201

[35] Reveil, M., Sorg, V. C., Cheng, E. R., Ezzyat, T., Clancy, P., & Thompson, M. O. (2017). Finite element and analytical solutions for van der Pauw and four-point probe correction factors when multiple non-ideal measurement conditions coexist. *Review of Scientific Instruments*, *88*(9). https://doi.org/10.1063/1.5001830

[36] Koon, D. W. (1989). Effect of contact size and placement, and of resistive inhomogeneities on van der Pauw measurements. *Review of Scientific Instruments*, *60*(2), 271–274. https://doi.org/10.1063/1.1140422

[37] Rietveld, G., Koijmans, C., Henderson, L., Hall, M., Harmon, S., Warnecke, P., & Schumacher, B. (2003). DC conductivity measurements in the van der pauw geometry. *IEEE Transactions on Instrumentation and Measurement*, *52*(2), 449–453. https://doi.org/10.1109/tim.2003.809917

[38] Oliveira, F. S., Cipriano, R. B., Da Silva, F. T., Romão, E. C., & Santos, C. a. M. D. (2020). Simple analytical method for determining electrical resistivity and sheet resistance using the van der Pauw procedure. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-72097-1

[39] van der Pauw, L.J. (1958) A Method of Measuring Specific Resistivity and Hall Coefficient on Lamellae of Arbitrary Shape. Philips Technical Review, 20, 220-224

[40] Naftaly, M., Das, S., Gallop, J., Pan, K., Alkhalil, F., Kariyapperuma, D., Constant, S., Ramsdale, C., & Hao, L. (2021). Sheet resistance measurements of conductive thin films: A comparison of techniques. *Electronics*, *10*(8), 960. https://doi.org/10.3390/electronics10080960

[41] Mosavi, A., Bertalan, B., Imre, F., Nadai, L., & Gorji, N. E. (2020b). Electrical characterization of CIGS thin-film solar cells by two- and four-wire probe technique. *Modern Physics Letters B*, *34*(11), 2050102. https://doi.org/10.1142/s021798492050102x

[42] Just, S., Soltner, H., Korte, S., Cherepanov, V., & Voigtländer, B. (2017). Surface conductivity of Si(100) and Ge(100) surfaces determined from four-point transport measurements using an analyticalN-layer

conductance model. *Physical Review. B./Physical Review. B*, *95*(7). https://doi.org/10.1103/physrevb.95.075310

[43] Kanagawa, T., Hobara, R., Matsuda, I., Tanikawa, T., Natori, A., & Hasegawa, S. (2003). Anisotropy in conductance of a Quasi-One-Dimensional metallic surface state measured by a square Micro-Four-Point probe method. *Physical Review Letters*, *91*(3). https://doi.org/10.1103/physrevlett.91.036805

[44] Lu, Y., Santino, L. M., Acharya, S., Anandarajah, H., & D'Arcy, J. M. (2017). Studying EC using a 3D printed Four-Point Probe station. Journal of Chemical Education, 94(7), 950–955. https://doi.org/10.1021/acs.jchemed.7b00119

[45] Waremra, R. S., & Betaubun, P. (2018). Analysis of Electrical Properties Using the four point Probe Method. E3S Web of Conferences, 73, 13019. https://doi.org/10.1051/e3sconf/20187313019

[46] Ju, B., Ju, Y., & Saka, M. (2005). Fabrication of a microscopic four-point probe and its application to local conductivity measurement. Journal of Micromechanics and Microengineering, 15(12), 2277–2281. https://doi.org/10.1088/0960-1317/15/12/009

[47] Chelly, A., Glass, S., Belhassen, J., & Karsenty, A. (2023). Broad review of four-point probe correction factors: Enhanced analytical model using advanced numerical and experimental cross-examination. Results in Physics, 48, 106445. https://doi.org/10.1016/j.rinp.2023.106445

[48] Kim, J., Zhang, Y., & Lee, D. (2009). A New Micro-Four-Point Probe Design for Various Applications. 2009 IEEE 22nd International Conference on Micro Electro Mechanical Systems, 84–87. https://doi.org/10.1109/memsys.2009.4805325

[49] Smits, F. M. (1958). Measurement of Sheet Resistivities with the Four-Point Probe. Bell System Technical Journal, 37(3), 711–718. https://doi.org/10.1002/j.1538-7305.1958.tb03883.x

[50] Metzke, C., Frammelsberger, W., Weber, J., Kühnel, F., Zhu, K., Lanza, M., & Benstetter, G. (2020). On the Limits of Scanning Thermal Microscopy of Ultrathin Films. *Materials*, *13*(3), 518. https://doi.org/10.3390/ma13030518

[51] Moridi, A., Zhang, L., Liu, W., Duvall, S., Brawley, A., Jiang, Z., Yang, S., & Li, C. (2017). Characterisation of high thermal conductivity thin-film substrate systems and their interface thermal resistance. *Surface and Coatings Technology*, *334*, 233–242. https://doi.org/10.1016/j.surfcoat.2017.11.021

[52] Callard, S., Tallarida, G., Borghesi, A., & Zanotti, L. (1999). Thermal conductivity of SiO2 films by scanning thermal microscopy. *Journal of Non-Crystalline Solids*, *245*(1–3), 203–209. https://doi.org/10.1016/s0022-3093(98)00863-1

[53] Dehbashi, M., Kazmierczak-Balata, A., & Bodzenta, J. (2025c). An approach for thermal conductivity measurements in thin films: Combining localized surface topography, thermal analysis, and ML techniques. *International Journal of Heat and Mass Transfer*, *248*, 127215. https://doi.org/10.1016/j.ijheatmasstransfer.2025.127215

[54] Kaźmierczak-Bałata, A., Bodzenta, J., Dehbashi, M., Mayandi, J., & Venkatachalapathy, V. (2022). Influence of post processing on thermal conductivity of ITO thin films. *Materials*, *16*(1), 362. https://doi.org/10.3390/ma16010362

[55] Kaźmierczak-Bałata, A., Bodzenta, J., Szperlich, P., Jesionek, M., Michalewicz, A., Domanowska, A., Mayandi, J., Venkatachalapathy, V., & Kuznetsov, A. (2024). Impact of Annealing in Various Atmospheres on Characteristics of Tin-Doped Indium Oxide Layers towards Thermoelectric Applications. *Materials*, *17*(18), 4606. https://doi.org/10.3390/ma17184606

[56] Kaźmierczak-Bałata, A., Bodzenta, J., & Guziewicz, M. (2019). Microscopic investigations of morphology and thermal properties of ZnO thin films grown by atomic layer deposition method. *Ultramicroscopy*, *210*, 112923. https://doi.org/10.1016/j.ultramic.2019.112923

[57] Kaźmierczak-Bałata, A., Grządziel, L., Guziewicz, M., Venkatachalapathy, V., Kuznetsov, A., & Krzywiecki, M. (2021). Correlations of thermal properties with grain structure, morphology, and defect balance in nanoscale polycrystalline ZnO films. *Applied Surface Science*, *546*, 149095. https://doi.org/10.1016/j.apsusc.2021.149095

[58] Juszczyk, J., Wojtol, M., & Bodzenta, J. (2013). DC experiments in quantitative scanning thermal microscopy. *International Journal of Thermophysics*, *34*(4), 620–628. https://doi.org/10.1007/s10765-013-1449-4

[59] Dryden, J. R., Yovanovich, M. M., & Deakin, A. S. (1985). The effect of coatings on the Steady-State and short time constriction resistance for an arbitrary axisymmetric flux. *Journal of Heat Transfer*, *107*(1), 33–38. https://doi.org/10.1115/1.3247398

[60] Trefon-Radziejewska, D., Juszczyk, J., Krzywiecki, M., Hamaoui, G., Horny, N., Antoniow, J., & Chirtoc, M. (2021). Thermal characterization of morphologically diverse copper phthalocyanine thin layers by scanning thermal microscopy. *Ultramicroscopy*, *233*, 113435. https://doi.org/10.1016/j.ultramic.2021.113435

[61] McBride, J., & Liu, H. (2020). The Relationship between Contact Resistance and Roughness (Sq) of a Bi-layered Surface using a Finite Element Model. *2020 IEEE 66th Holm Conference on Electrical Contacts and Intensive Course (HLM)*, 176–181. https://doi.org/10.1109/hlm49214.2020.9307833

[62] Xiao, C., Ye, J., Esteves, R. M., & Rong, C. (2015). Using Spearman's correlation coefficients for exploratory data analysis on big dataset. Concurrency and Computation Practice and Experience, 28(14), 3866–3878. https://doi.org/10.1002/cpe.3745 .

[63] GeeksforGeeks. What is Machine Learning? https://www.geeksforgeeks.org/machine-learning/ml-machine-learning/

[64] GeeksforGeeks. Gradient Boosting in ML. https://www.geeksforgeeks.org/machine-learning/ml-gradient-boosting/

[65] GeeksforGeeks. Random Forest Algorithm in Machine Learning. https://www.geeksforgeeks.org/machine-learning/random-forest-algorithm-in-machine-learning/

[66] GeeksforGeeks. Hyperparameter Tuning. https://www.geeksforgeeks.org/machine-learning/hyperparameter-tuning/

[67] GeeksforGeeks. Cross Validation in Machine Learning. https://www.geeksforgeeks.org/machine-learning/cross-validation-machine-learning/

[68] Shingala, B., Panchal, P., Thakor, S., Jain, P., Joshi, A., Vaja, C. R., Siddharth, R. K., & Rana, V. A. (2024). Random Forest Regression Analysis for Estimating Dielectric Properties in Epoxy Composites Doped with Hybrid Nano Fillers. *Journal of Macromolecular Science Part B*, *63*(12), 1297–1311. https://doi.org/10.1080/00222348.2024.2322189

[69] Guo, J., Zan, X., Wang, L., Lei, L., Ou, C., & Bai, S. (2023). A random forest regression with Bayesian optimization-based method for fatigue strength prediction of ferrous alloys. *Engineering Fracture Mechanics*, *293*, 109714. https://doi.org/10.1016/j.engfracmech.2023.109714

[70] Wekalao, J., Srinivasan, G. P., Patel, S. K., & Al-Zahrani, F. A. (2024). Optimization of graphene-based biosensor design for haemoglobin detection using the gradient boosting algorithm for behaviour prediction. *Measurement*, *239*, 115452. https://doi.org/10.1016/j.measurement.2024.115452

[71] K, P. K., Alruqi, M., Hanafi, H., Sharma, P., & Wanatasanappan, V. V. (2023). Effect of particle size on second law of thermodynamics analysis of Al2O3 nanofluid: Application of XGBoost and gradient boosting regression for prognostic analysis. *International Journal of Thermal Sciences*, *197*, 108825. https://doi.org/10.1016/j.ijthermalsci.2023.108825

[72] Pangarkar, D. J., Sharma, R., Sharma, A., & Sharma, M. (2020). Assessment of the Different ML Models for Prediction of Cluster Bean (Cyamopsis tetragonoloba L. Taub.) Yield. *Advances in Research*, 98–105. https://doi.org/10.9734/air/2020/v21i930238

[73] COMSOL. (2015). COMSOL Multiphysics® 5.1: simulation software. https://www.comsol.com

[74] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2012). SciKit-Learn: Machine Learning in Python. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1201.0490 .