RDIMa.512.1.2026

prof. dr hab. inż. Krzysztof T. Wojciechowski
Thermoelectric Research Laboratory
Department of Inorganic Chemistry
Faculty of Materials Science and Ceramics
AGH University of Krakow

REVIEW OF THE DOCTORAL DISSERTATION

of M.Sc. Mohsen Dehbashi

entitled "**Predicting and Analyzing the Thermal and Electrical Properties of Materials Using Advanced Machine Learning Models**"

1. General Information

The doctoral dissertation submitted for review by M.Sc. Mohsen Dehbashi was prepared at the Institute of Physics, Centre for Science and Education of the Silesian University of Technology, in the **discipline of Materials Engineering**.

The supervisor of the dissertation is **Prof. Jerzy Bodzenta**, D.Sc., Eng., and the assistant supervisor is **Dr. Justyna Juszczyk-Synowiec**, Eng.

The dissertation is written in English and has the form of a monograph based on partially published research results. The work comprises 104 pages in total and consists of 9 chapters, 3 appendices containing source data, a description of the numerical libraries used and algorithmic sources, a list of scientific achievements and awards, and a bibliography including 74 references.

The dissertation includes results presented in two scientific articles published in two peer-reviewed journals: *International Journal of Heat and Mass Transfer* (Elsevier) and *Materials* (MDPI). In the former, the doctoral candidate is the first author. Additionally, part of the results was published in the monograph *Interdyscyplinarne badania młodych naukowców* (Silesian University of Technology, No. 987, pp. 81–91), published by the Silesian University of Technology Press.

2. Characteristics of the Research Problem and Scientific Context

The dissertation concerns the application of machine learning methods to improve the accuracy and repeatability of determining thermal conductivity (TC) and electrical conductivity (EC) in thin films. The starting point of the work is not the development of new measurement techniques *sensu stricto*, but rather an analysis of the limitations of existing measurement methods (SThM in relation to TC and 4PP in relation to EC) and an attempt to reduce systematic and geometric errors using predictive models trained on experimental data.

The subject of the dissertation is timely, and addresses needs clearly articulated in the literature regarding the enhancement of reliability in local measurements of thermal and electrical properties in thin-film and nanoscale structures. Contemporary scanning techniques provide high spatial resolution; however, the interpretation of the measured parameters is affected by uncertainties arising from contact geometry, surface topography, edge effects, and substrate interactions. In recent years, there has been significant development of approaches based on data analysis and machine learning methods aimed at improving the quality of physical parameter estimation without the need for radical modification of measurement instrumentation.

The dissertation was carried out at a research center of high international standing in the field of scanning thermal microscopy techniques and measurements of electrical properties of materials at the micro- and nanoscale. The research group is internationally recognized for developing both SThM instrumentation and other methods of nanoscale signal analysis. The topic of the dissertation fits directly within this research profile and represents a continuation of work aimed at improving the accuracy of interpreting local transport measurements using advanced data analysis methods (machine learning – ML).

3. Substantive Evaluation of the Dissertation

Introduction to the Research Problem

Chapters 1 and 2 of the dissertation are introductory in nature and provide a general overview of thermal and electrical transport in thin films, as well as the fundamentals of the measurement techniques (SThM, 4PP) employed in the subsequent parts of the work. The Author identifies the difficulties associated with determining thermal and electrical conductivity in thin films, emphasizing the influence of surface topography, substrate interactions, contact resistances, and the limitations of classical interpretative models. The problem is outlined extensively in a descriptive manner and is framed within real metrological challenges.

Research Background

Despite the extensive introduction, the review of the current state of knowledge with respect to worldwide research on methods for determining TC and EC, presented in Chapter 3, is very concise (2.5 pages; 11 references for TC and 7 for EC), which does not allow the dissertation to be fully situated within the context of the most recent work in this field. The Author discusses well-known physical difficulties associated with determining thermal and electrical conductivity using SThM and 4PP (such as the influence of surface topography, ambiguity of probe–sample contact conditions, etc.).

However, insufficient attention is given to a key issue: **how machine learning methods are used to address specific metrological problems in materials engineering.** References to data-driven and surrogate modeling approaches developed in recent years for the analysis of SThM signals would have been appropriate, including studies combining FEM with probe-response approximation and ML techniques. Similarly, in electrical measurements, the literature on geometric corrections (both analytical and FEM-based) for 4PP and van der Pauw methods, including non-ideal geometries, is entirely omitted.

The absence of these references and of a comparison with existing solutions may suggest a pioneering character of the proposed method, whereas data-driven approaches in the analysis of transport measurements are already well established in the literature.

3. Research Objectives and Hypotheses

The research objectives defined in Chapter 4 are methodological in nature and concern the development of an integrated procedure combining spatial mapping, topographical descriptors, signal normalization, and machine learning models. The research hypotheses are also formulated in a descriptive manner. They assume that the integration of the indicated elements leads to increased accuracy, reliability, and repeatability in determining transport parameters. However, given the methodological character of the dissertation under review, the Reviewer **would expect precise operational criteria for their verification**, in particular: with respect to which reference model the improvement is defined, which error metric has been adopted, and within what parameter range the formulated conclusions remain valid.

4. Original Research

The description of the Author's original research is clearly divided into two symmetrically structured parts: analysis of thermal conductivity (TC) and analysis of electrical conductivity (EC). In the present review, the Reviewer maintains an analogous division, combining common elements of both issues where appropriate.

Investigated and Reference Materials

In both Chapter 5 (TC) and Chapter 7 (EC), information concerning the materials is presented together with the description of the respective instrumentation and measurement configurations. Structurally, combining these two aspects makes methodological analysis more difficult, as no separate section is devoted exclusively to the samples, including their origin, fabrication technology, physical parameters, and selection criteria. In a doctoral dissertation, one would expect a clear separation between the characterization of investigated and reference materials and the description of instrumentation and measurement procedures.

The information provided on the samples is incomplete and unsystematic. The criteria for selecting the materials (for both TC and EC), the range of conductivity values, the role of layer thickness, and whether the datasets were designed with representative ML model training in mind are not clearly specified, which makes it difficult to assess the adequacy of the database.

A key issue concerns the origin of the material samples and the sources of input data used for training and testing the regression models. This matter is not described unambiguously, **and the Author's contribution to sample preparation, microstructural characterization, and transport property measurements remains unclear.** These issues are significant if the dissertation is evaluated within the discipline of Materials Engineering.

The methodological core of the research plan, for both TC and EC, is the comparison of parameters estimated using the new methods with reference material values reported in publications [54]–[58]. However, the references presented in Tables 5.1, 6.4, and 7-1 to literature values may be misleading. It should be clarified that these are not external data obtained by independent research groups, but data generated by the research team in which the doctoral candidate was involved. Moreover, in the case of certain bulk samples (Tables 5.1 and 6.4) and metallic samples (Table 7-1), the source of some reference data is not explicitly stated, nor is the accuracy of the reference parameters **κ_actual** and **σ_actual** provided. This is an important issue in metrological work, with consequences for the final conclusions in relation to the research hypotheses stated.

The following issues therefore require clear clarification:

– Do the data originate from exactly the same samples for which the reference values κ_actual and σ_actual were determined, or from different samples fabricated by the Author from the same materials and using the same method?
– Were the training and test data acquired from different regions of the reference samples, or were

the original data from the cited publications directly used, from which the reference values κ_actual and σ_actual were determined?
– What is the origin of the reference κ_actual data for the bulk samples (glass, glassy carbon, SiC, YAG; Tables 5.1 and 6.4) and of σ_actual for the metallic samples (Table 7-1)?

In the case of the EC method (Chapter 7), the reference parameter was the electrical conductivity σ_actual measured using the reference van der Pauw (vdP) method. For comparison, measurements were performed on series of samples using a similar device in the 4PP configuration. This again raises the following question:

– Are the samples used in the 4PP measurements exactly the same as those for which σ_actual was previously determined using the van der Pauw method (ref. [55]), or do they originate from a different technological batch?

In thin films (e.g., ITO), electrical conductivity strongly depends on thickness, technological process parameters, dopant concentration, microstructure, and other factors. Variations in these parameters may significantly affect σ. Therefore, the issue of material aging must be considered. If the vdP measurements providing σ_actual were performed, for example, several months earlier and the samples were stored during that period, changes in electrical properties may occur in thin films (e.g., due to oxidation processes, stress relaxation, or changes in carrier concentration).

– Was the temporal stability of conductivity verified before using σ_actual values as reference data?

Measurement Methodology

Chapters 5 and 7 provide an extensive description of the key methodological components related to the implementation of ensemble regression models (ensemble learning: Random Forest, Gradient Boosting) for estimating TC and EC, respectively, using established algorithm libraries. In principle, the implementation of the models has been carried out correctly. However, the Reviewer notes the absence of essential information justifying, for example, the selection of the models (RF/GB), as well as clear details regarding the procedure for splitting the data into training and test sets, the validation strategy, and specific implementation settings. In particular, the Reviewer found the lack of concise workflow diagrams for both the TC and EC parts to be a significant shortcoming. Such diagrams should clearly present the sequence of steps: from data acquisition, through preprocessing and descriptor extraction, to model training, validation, and evaluation of generalization. The inclusion of such schemes would greatly facilitate assessment of procedural coherence, the level of data separation, and the points at which potential data leakage between training and test sets could have occurred.

Training and Test Sets

The issue of clearly defining and distinguishing between training and test sets requires clarification. In the TC part, the Author provides only the number of measurements in the training set (2352) and in the test set (980), without specifying the method of partitioning. Under these circumstances, it is not possible to assess the risk of information leakage or the predictive quality of the model. Therefore, I request an explicit clarification of the following issues:

1. **Data partitioning scheme**
   Was the division of the 3332 records into 2352 (training) and 980 (testing) performed at the level of individual records, or at the level of entire material samples (measurement maps)?

(For example, were 5 materials corresponding to 980 data points excluded entirely from training, or were 980 records randomly selected from the entire dataset?)
Did the test set contain records originating from the same material samples that were present in the training set?

2. **Spatial Autocorrelation Resulting from the Filtering Procedure**

To reduce measurement noise, an averaging procedure based on a 3×3 "moving window" filter was applied. However, this algorithm generates strongly overlapping regions of analysis with high spatial interdependence, **which may lead to data leakage and, consequently, to inflated values of the correlation coefficient R^2.**

Was a mechanism preventing spatial leakage applied when splitting the data into training and test sets (e.g., block-wise map partitioning or grouping records by sample)?

3. **Independence of the Target Variable χ**

The target variable χ = κ_s / σ_si is constructed from quantities related to the same measurement maps from which the input features are derived, which may significantly influence the resulting R^2 value.

Was κ_s used to construct χ obtained from a fully independent reference measurement, or did it originate from the same measurement/calibration procedure as the input features (Γ_i, Δϕ_i, topographical parameters)?

4. **Order of Data Processing Operations**

It is not specified whether normalization, feature selection (Spearman correlation analysis), and other operations were performed before or after the data split. This is relevant with respect to their impact on the R^2 coefficient.

In other words: **were all preprocessing operations (normalization, feature selection, calculation of derived parameters) performed in an identical manner on the training set and the test set?**

5. **Inter-material Validation**

Based on the current description of the TC procedure, it cannot be excluded that the reported value $R^2 = 0.966$ was obtained under conditions of significant information leakage (spatial autocorrelation of 3×3 windows, possible record-level rather than group-level data splitting, lack of a clear distinction between preprocessing and actual signal analysis). In such a scenario, R^2 measures within-dataset agreement under strongly interdependent observations rather than the model's ability to predict new materials or independent maps.

Consequently, the claim of "high predictive performance" and a "scalable framework" is not methodologically substantiated without demonstrating group-based validation. **Without such verification, a high R^2 value cannot be treated as evidence of the model's generalization capability.** This leads to an important question:

Was cross-validation of the leave-one-material-out or GroupKFold type, with grouping by material sample, performed?

Results and Discussion (Thermal Conductivity)

Chapter 6 presents the application of ensemble regression models to predict thermal conductivity based on SThM data and topographical descriptors. The Author demonstrates a reduction in RMSE and improved agreement between predicted and actual κ compared to the uncorrected approach, both for the training and test sets, which numerically supports the thesis of improved TC estimation. At the same time, the interpretation of these results requires clarification of several key issues.

Was a sensitivity analysis performed to assess the influence of hyperparameter variations (e.g., max_depth, learning_rate), and if so, is the reported RMSE reduction stable across a broad range of their values?
Without such analysis, it is difficult to determine whether the obtained results are independent of hyperparameter choice or are tied to a specific model configuration.

Attention should be drawn to the extremely large maximum tree depth (max_depth = 419) relative to the limited number of material samples. Such a configuration implies a model of very high complexity, capable of fitting the training data very closely (risk of overfitting). In this situation, a high R^2 value alone—even with a small difference between training and test sets—does not constitute sufficient evidence of good generalization. **Please clarify why such an exceptionally large depth was adopted and how it was verified that the obtained R^2 is not a consequence of excessive model complexity.**

The Author applied two models—*Random Forest* and *Gradient Boosting*—and compared their performance primarily based on error metrics.

On what grounds were these algorithms selected as appropriate for the problem under consideration?

A more in-depth discussion of the suitability of these methods to the nature of the data is lacking. Given the limited number of independent material samples, the susceptibility of the models to overfitting and their ability to generalize should be discussed.

Summary (Thermal Conductivity)

In summary of the TC part, the Author documents an improvement of the ML models in terms of numerical agreement with the adopted reference value κ_actual, as measured by a reduction in RMSE and an increase in R^2. However, it has not been unambiguously demonstrated that this improvement is statistically generalizable across independent material samples, nor that it exceeds the uncertainty level of the reference values. In its present form, the results confirm effective fitting within the analyzed dataset (estimation character), whereas the extent of their predictive capability across different materials has not been sufficiently documented.

Methodology of Electrical Conductivity Measurements (EC, 4PP)

Chapter 7.3 presents the construction of ML models (*Random Forest* and *Gradient Boosting*) intended to correct 4PP measurement results in cases of irregular sample geometry and varying probe positioning (A1–A4, rotations).

The description does not provide information regarding control of electrical contact quality or boundary conditions of the measurement. It is unclear whether contact stability, probe pressure, or possible damage to the thin film during repeated measurements were monitored. This raises the question: **was the influence of contact resistance and stability considered in the analysis, and were any criteria introduced for rejecting outlier measurements?**

The procedure for determining conductivity from the V–I characteristics is not clearly described. It is not specified whether full linear regression over multiple current values was performed or whether linearity of the response was tested. Therefore, it should be clarified **whether conductivity was determined from the slope of the V–I characteristic and what the repeatability of the results was for the same geometric configuration (including sample rotations)?**

The Author introduces an intermediate variable Θ, predicted on the basis of geometric parameters and current, and subsequently uses it to determine conductivity σ_ML. Although Θ is formally introduced, it is not clearly interpreted. From the equations, it appears to function as a correction factor; however, it is not explicitly explained whether Θ approximates a relationship derived from an FEM model or directly represents the relationship between σ_exp and σ_actual.

It should therefore be clarified how Θ is defined in the training data and what its physical or algorithmic status is

Statistical Analysis

An important issue of metrological relevance in the dissertation concerns the quality of the statistical treatment of the results. As noted, the reference values taken from the literature, σ_actual (as well as κ_actual), are presented **without measurement uncertainties**. The Author demonstrates improvement of the parameters toward σ_actual. In the absence of information on the measurement errors associated with the reference parameters, such improvement does not necessarily translate into higher measurement reliability.

A symmetric problem is the lack of analysis of the propagation of input uncertainties to the values of σ_ML. The estimation result is influenced, among other factors, by probe positioning errors, uncertainties in geometric determination, and thickness d in σ_ML. It has not been shown whether the ML model attenuates or amplifies measurement fluctuations, nor how sensitive it is to perturbations of the input parameters.

It is therefore appropriate to ask explicitly: **was a sensitivity/error propagation analysis performed (e.g., for uncertainties in L_h, L_v, d), and is it possible to quantify the stability of σ_ML with respect to realistic input errors?**

General Editorial Remarks

Independently of the substantive evaluation of the dissertation, the Reviewer would like to draw attention to significant editorial and formal shortcomings that affect the clarity of the work and, consequently, the comfort of its analysis by the reader.

First, the submitted version **lacks a standard title page containing the full identification details of the dissertation** (author's name, institutional affiliation, names and titles of the supervisor and assistant supervisor, year and place of submission). Such a page constitutes a standard element of a doctoral document and serves as a formal and archival function. **Its absence should be regarded as a formal deficiency.**

Second, the dissertation lacks a list of symbols. Numerous symbols are used throughout the text (including different variants of conductivity: experimental/apparent, actual, corrected; in both the TC and EC sections; intermediate variables such as Θ), yet they are not collected in a single location with unambiguous definitions and units. In a work of metrological and modeling character, the absence of such a list significantly impedes verification of the formal consistency and unambiguous interpretation of the equations. Additional confusion is introduced by inconsistencies in terminology and naming of physical quantities (e.g., different terms used for the same category of conductivity).

In the Reviewer's opinion, the chapters describing the methodologies contain an excessive amount of textbook-like content (definitions of correlation, general foundations of ensemble models) that do not constitute part of the Author's original research and could have been moved to the introductory section. The current structure blurs the boundary between theoretical background and the Author's actual contribution.

The indicated editorial shortcomings do not undermine the substantive value of the dissertation; however, they reduce its formal clarity. In a doctoral dissertation, a higher level of editorial coherence and definitional precision would be expected.

Final Conclusions

The dissertation addresses a genuine metrological problem related to the interpretation of local measurements of thermal and electrical conductivity in thin films using SThM and 4PP methods and proposes its solution through the application of machine learning models employing, among others, descriptors of surface topography, geometric configuration of the measurement setup, and measurement signal characteristics. The Author correctly implements and analyzes regression models on experimental data, demonstrating the possibility of reducing discrepancies between measured and reference values within the investigated dataset. The work has an applied character and fits within the current research trend of applying ML methods in materials engineering. The results obtained indicate the potential for improved numerical agreement with reference values within the analyzed dataset.

At the same time, it should be emphasized that the extent to which the research hypotheses have been confirmed—particularly regarding the inter-material predictive capability of the ML models—has not been unambiguously documented and requires clarification during the defense.

Despite the indicated limitations, I conclude that the dissertation constitutes an original application of machine learning methods to the solution of a metrological problem relevant to materials engineering and fulfills, to a sufficient degree, the requirements specified in Article 187 of the Act of 20 July 2018 – Law on Higher Education and Science (consolidated text: Journal of Laws of 2024, item 1571). Therefore, I recommend that the doctoral dissertation be accepted and that it be admitted to public defense.

Kraków, 23.02.2026 Prof. Krzysztof T. Wojciechowski

/podpis odręczny/

Appendix to the Review

List of Selected Editorial Deficiencies

I. Bibliographic Ordering

1. The dissertation adopts a reference ordering system based on the sequence of first appearance in the text (Vancouver standard). Unfortunately, this rule has not been consistently followed. For example, already in Chapter 1, references with higher numbers (e.g., [27], [50]) appear in the introductory section, while some publications with lower numbers are cited only later in the work. Similar situations occur in other parts of the dissertation.

II. Consistency of Physical Quantity Notation

1. Thermal conductivity (κ)
   • pp. 7–8 – The abbreviation TC is used without an explicit assignment of the symbol κ upon first introduction of the concept.
   • pp. 29, 52 – Tables use the notation "Actual κ" and "Predicted κ" without consistent index notation (e.g., κ_actual, κ_pred).
   • p. 48 (Chapter 6) – Parallel use of "κ_actual" and "actual κ" (inconsistent formatting).

2. Electrical conductivity (σ)
   • pp. 15–16 – Parallel use of the terms electrical conductivity, conductivity, and the abbreviation EC without explicit assignment of the symbol σ at first use.
   • pp. 55, 74 – Notation "σactual" (without subscript and separator) in tables; inconsistent with "σFEM" and "σML".
   • p. 59 (Eqs. 7.1–7.4) – Notations "σexp", "σFEM", "σML" used without consistent subscript formatting; lack of unified convention relative to "σactual".

3. Parameter Γ_i (TC)
   • pp. 30, 34, 36, 37, 39–41 – Notation "Γi" (without subscript).
   • p. 77 – Notation "Γ_i" (with correct subscript).
   Typographical inconsistency of the same parameter in different parts of the dissertation.

4. Thermal resistance
   • pp. 13–14 – Notations "Rth,P" and "Rth,p" (inconsistent capitalization of the subscript). Formally, this creates two different symbols.

III. Definitions and Units of Physical Quantities

• pp. 8–9 – No unit provided at the first introduction of thermal conductivity.
• p. 16 – The symbol σ appears without a formal definition and unit at first use.
• p. 18 – Transport parameters (e.g., mobility, carrier concentration) are mentioned without units at first occurrence.

IV. Consistency of Terminology and Nomenclature

• pp. 26–27 – Parallel use of the terms prediction, estimation, correction without formal distinction of their meaning.
• pp. 48–61 – In the results section, the terms experimental, apparent, corrected, actual conductivity are used interchangeably without a consistent symbolic system.
• pp. 63–64 – In Chapter 9, the term "predictive model" is used without clearly specifying whether it refers to inter-sample prediction or within-dataset estimation.

V. Notation of Datasets and Validation (ML)

• pp. 46–50 (TC) – No clear indication of which results correspond to the training set and which to the test set.
• pp. 56–62 (EC) – No explicit information on the number of independent material samples represented in the analysis; the number of data records provided is not related to the number of materials.

VII. Presentation of Results in Tables and Figures

• pp. 47–48, 58–60 – In some figures, units are not clearly specified on the axes (requires verification for each figure).
• p. 55 – The unit of the parameter Θ in the table is given in an unusual format ("S·mV"), without explanation of its dimensional meaning in the text.
• Tables 5.1 and 6.4 are largely repetitive.