Fachtagung "Experimentelle Strömungsmechanik"
5.–7. September 2023, München

Machine learning approach for fast evaluation of filtered Rayleigh scattering measurement data

Sergey M. Melnikov[1], Ingo Röhle[1], Ulrich Doll[2], Michael Dues[3], Jonas Steinbock[3], Matteo Migliorini[4], Pavlos K. Zachos[4]

[1]Berliner Hochschule für Technik, Berlin, Germany
[2] Paul-Scherrer-Institut, Villingen, Swiss
[3] ILA R&D GmbH, Jülich, Germany
[4]Cranfield University, United Kingdom

Keywords: machine learning, data analysis, filtered Rayleigh scattering

Abstract

Machine learning (ML) approach has been successfully applied for accelerating filtered Rayleigh scattering (FRS) measurement data analysis at the two levels of evaluation. Calibrated with the help of experimental data ML approximation of the Rayleigh-Brillouin scattering (RBS) profile establishes a stable, accurate, and fast model for the description of the RBS profile. Application of the ML approach directly for FRS intensity spectra approximation together with good accuracy provides a 200-fold acceleration of the data evaluation. ML models have been tested by simulated and experimental data and are incorporated into the software.

Introduction

A main issue of data processing of filtered Rayleigh scattering (FRS) measurements is the correct numerical description of the Rayleigh-Brillouin scattering (RBS) profile. Commonly used physics-based models, such as the Tenti S6 (Tenti et al. 2011) and the Pan S7 (Pan et al. 2004), although proved to be exact, are computationally time-consuming. In space-resolved FRS measurements, processing of data collected by hundreds of thousands of camera pixels becomes a significant computational task. Analytical approximations of the established numerical models could overcome that drawback. However, existing analytical approximations (Witschas 2011), (Ma et al. 2012), (Binietoglou et al. 2016), (Doll et al. 2016) are not stable enough or do not adequately describe the Rayleigh scattering profile in wider gas flow regimes.

To accelerate FRS data processing, we have resorted in the current project to the machine learning (ML) approach, a powerful computational technique, which becomes nowadays an efficient and widespread working tool, due to highly developed algorithms and a range of available software. ML performs fast approximations of the established models of the RBS profile based on numerical regularities of data, without any knowledge of the physics behind the data. The striking advantage of the ML approach is that the predictive function can be quickly recalculated for another model, input parameter set, or parameter range.

GALA e.V., Karlsruhe, Germany, ISBN 978-3-9816764-9-5

We have applied the ML approach at two levels of FRS data evaluation. At first, the ML approximation have been used for the description of the RBS profile. It establishes a fast, stable, and accurate method for the data evaluation. In the second stage, in the spirit of the ML approach, we apply the technique directly to data measured in the experiment, in other words, to FRS intensity spectra. It leads to a drastic increase in processing speed, due to omitting the calculation of the RBS profile and its convolution with the molecular filter transmission curve.

Computational details

To create ML prediction functions, one needs to perform a teaching procedure, which relates input model parameters with a set of output profiles, the so-called training set. In our case, it is a set of Rayleigh scattering profiles or a collection of FRS intensity spectra, corresponding to all parameter variations expected in forthcoming measurements. These curves can be either generated numerically with one of the established models or obtained directly from an experiment. All preparative calculations, including the generation of the training set and the teaching procedure, can be done within minutes. Mathematically speaking, the prepared ML predictive model is a following gaussian

$$Z(\Phi) = \beta_0 + \beta * e^{(-\|Z-\Phi\|^2)},$$

where Φ is an input parameter set, and Z is an approximation output for a predefined set of definition points, which are, for instance, excitation frequencies in an FRS intensity spectrum. ML algorithms calculate the predictive set - matrixes β_0, β, and Σ. Dimensions of these matrixes are defined by the sizes of the input parameter set, definition points set, and intrinsic parameters of the ML teaching algorithm. Our models are realized in the Python programming language. The core functions of our models are taken from the scikit-learn software package (Pedregosa et al. 2011). Additionally, we use also programming scripts proposed by Hunt G.J. (Hunt et al. 2020).

The ML model for approximation of the RBS spectrum

Firstly, the ML approach has been used to approximate the Tenti model. The corresponding training set was calculated for the following input parameter domains: Y ∈ [0.55 – 3.8], the internal relation number - [1.5; 3], and the Eucken factor - [1.8; 2.1]. Other two parameters of Tenti's model, namely the internal specific heat and the translational specific heat, were considered here as constants. Fig. 1 presents the training set (left panel) and the high quality of the ML approximation (mlRBS) for a case of the kinetic regime of measured gases (right panel). It is to be noted that residuals of approximation do not exceed a value of 0.5% within the whole input parameter domain. Replacing the Tenti model with the ML approximation in the data evaluation procedure leads to a reduction of the evaluation time by a factor of 20. The high accuracy of the ML approximation is implied as the difference between evaluation results obtained with the conventional method and the ML one, which does not exceed 0.15 K, 100 Pa, and 0.15 m/s in temperature, pressure, and flow velocity, correspondingly. That is significantly less than the current uncertainties of FRS measurements (Doll et al. 2022). The obtained in such a way ML model of the RBS profile has been calibrated. The matrix of coefficients β_0 from the above-described formula was varied to fit reference experimental data. Then the calibrated ML model has been successfully used by (Doll et al. 2023) for data evaluation. Since the RBS profile does not depend on measurement conditions, the predictive set for the ML model can be calculated in advance. In this form, the ML model, called

GALA e.V., Karlsruhe, Germany, ISBN 978-3-9816764-9-5

mlRBS_cal, has been incorporated into the developed software as an optional stable model for the description of the RBS profile.

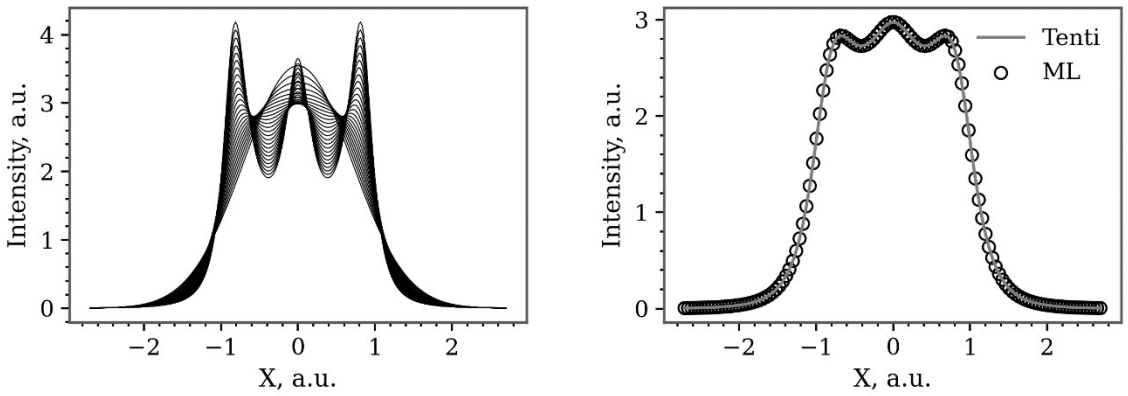

Figure 1. Left panel: the training set for the mlRBS model, for the sake of clarity only each 20th curve is presented; right panel: quality of the ML approximation of the RBS profile.

The ML approximation of FRS intensity spectra

An application of the ML approach directly for the approximation of FRS intensity spectra, called within the project the mlFRS model, provides a tremendous acceleration of the data evaluation, although requires some tuning. A blunt increase in the training set size is not a panacea for getting better results. Overfitting is a known issue in applications of ML algorithms. It has been found for four input parameters of the mlFRS model, namely pressure, temperature, the Doppler shift, and the scattering angle, the most preferable numbers of gradations, which results in 625 curves for the whole training set.

Input parameters of the training set generator for the model are expressed by large numbers in kPa, MHz, and hundreds of Kelvins. However, the ML algorithms demonstrate better performance working with input parameters that are distributed within an interval from -1 to 1, or close to that. To satisfy that condition input parameters of the model have been centralized and normalized. In other words, firstly a constant *C* is subtracted from a parameter value, then the resulting amount is divided by a constant *N*. The quality of the approximation is very sensitive to a choice of that constants. We have performed a systematic search for the best selection of the constants for pressure, temperature, and Doppler shift. The scattering angle was found to be better left intact.

Relatively small changes in the scattering angle heavily affect the form of FRS intensity spectra. It makes it difficult for the ML algorithms to combine in one model intensity spectra originating from scattering angles ranging in radians from 0.55 to 2.2. This difficulty manifests itself through low-quality predictions. To minimize the problem, we have narrowed the interval of the scattering angle for the training set. It means in practice, that the training set is calculated separately for each camera position. That modification requires additionally a one-minute calculation per camera during the preparation phase and almost does not increase the time required for data evaluation in the operational phase.

The training set for the mlFRS model can be obtained from preliminary test measurements or calculated as a convolution of the RBS profile with the iodine filter transmission curve for the excitation frequencies sequence. In the current project, we tested two variants of the mlFRS

GALA e.V., Karlsruhe, Germany, ISBN 978-3-9816764-9-5

training sets. The first variant was prepared with the help of the Tenti model and the second one with the mIRBS_cal model. The training set of the mIFRS model, calculated with the help of the Tenti model for the following parameter domains for pressure - [0.82e5 - 1.18 e5] Pa; temperature - [268-350] K, the Doppler shift - [0 - 300] MHz, and the scattering angle – [0.6-2.2], is presented in Fig. 2 (left). The quality of the mIFRS approximation, although not as perfect as in the previous mIRBS case, proved to be good, with an average deviation from the original spectrum of less than 1%. An example of the mIFRS approximation for a case of ambient conditions, the Doppler shift of 100 MHz, and the scattering angle of 1.27 is displayed in Fig. 2 (right).

The accuracy of the mIFRS model has been tested using both simulated and experimental data (Doll et al. 2017). The simulated data gives a possibility to test a large range of parameter combinations, while analysis of the real data, although not with such a broad parameter variety, strengthens the system validation. Approximation accuracy for two data types and two training sets proved to be similar. Maximal differences observed in testing runs between data evaluation results obtained with the mIFRS and the conventional methods are 600 Pa, 1.2 K, and 3.5 m/s for pressure, temperature, and velocity components, respectively.

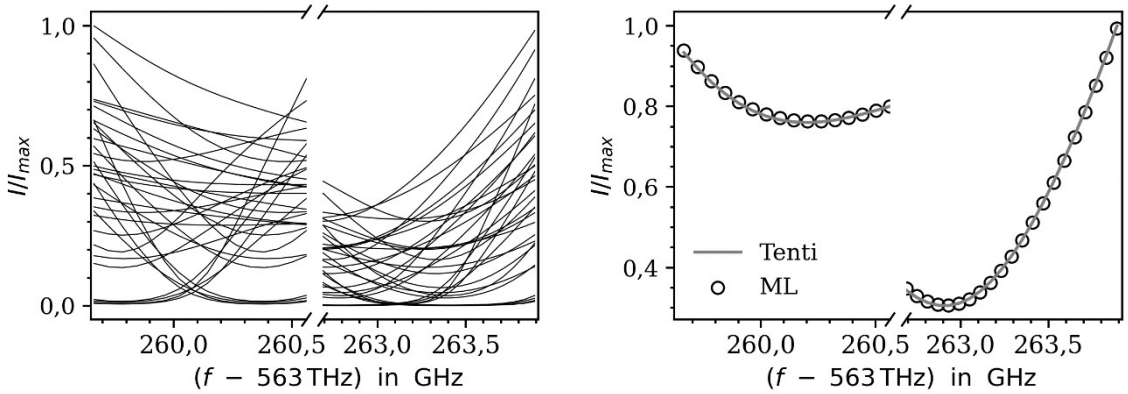

Figure 2. Left panel: the training set for the mIFRS model, for the sake of clarity only each 20th curve is displayed; right panel: quality of the ML approximation of the FRS spectrum.

Conclusive remarks

The ML approach has been applied for FRS measurement data analysis at two levels of the evaluation. The ML model of the RBS profile calibrated with experimental data provides a fast stable and accurate calculation method. In a fully applicable state, it is included in the developed software as an optional method for the description of the RBS profile. Application of the ML method directly at the level of the FRS intensity spectra brings in comparison to conventional methods at least a 200-fold decrease in computational time required for the evaluation. It comes together with good accuracy. Further optimization is possible in the way of a better choice for centralization and normalization constants. Presented ML methods of accelerated data evaluation contribute to the promotion of FRS measurements towards a real-time application.

GALA e.V., Karlsruhe, Germany, ISBN 978-3-9816764-9-5

Acknowledgment

This project has received funding from the Clean Sky 2 Joint Undertaking (JU) under grant agreement No 886521. The JU receives support from the European Union's Horizon 2020 research and innovation program and the Clean Sky 2 JU members other than the Union.

References:

Binietoglou, I., Giampouras, P., Belegante, L., 2016: "Linear Approximation of Rayleigh–Brillouin Scattering Spectra." *Applied Optics* 55 (27): 7707–11. https://doi.org/10.1364/AO.55.007707.

Doll, U., Burow, E., Stockhausen, G., Willert, C., 2016: "Methods to Improve Pressure, Temperature and Velocity Accuracies of Filtered Rayleigh Scattering Measurements in Gaseous Flows" 27 (12): 125204. https://doi.org/10.1088/0957-0233/27/12/125204.

Doll, U., Kapulla, R., Steinbock, J., Dues, M., Migliorini, M., Zachos, P.K., 2023: "Seeding-Free Inlet Flow Distortion Measurement by Filtered Rayleigh Scattering: Diagnostic Approach and Verification." In *AIAA SCITECH 2023 Forum*. https://doi.org/10.2514/6.2023-1372.

Doll, U., Röhle, I., Dues, M., Kapulla, R., 2022: "Time-Resolved Multi-Parameter Flow Diagnostics by Filtered Rayleigh Scattering: System Design through Multi-Objective Optimisation." *Measurement Science and Technology* 33 (10): 105204. https://doi.org/10.1088/1361-6501/ac7cca.

Doll, U., Stockhausen, G., Willert, C., 2017: "Pressure, Temperature, and Three-Component Velocity Fields by Filtered Rayleigh Scattering Velocimetry." *Opt. Lett.* 42 (19): 3773–76. https://doi.org/10.1364/OL.42.003773.

Hunt, G.J., Ground, C.R., Hunt, R.L., 2020: "Fast Approximations of Spectral Lineshapes to Enable Optimization of a Filtered Rayleigh Scattering Experiment" 31 (9): 095203. https://doi.org/10.1088/1361-6501/ab8a7e.

Ma, Y., Fan, F., Liang, K., Li, H., Yu, Y., Zhou, B., 2012: "An Analytical Model for Rayleigh–Brillouin Scattering Spectra in Gases" 14 (9): 095703. https://doi.org/10.1088/2040-8978/14/9/095703.

Pan, X., Shneider, M.N., Miles., R.B., 2004: "Coherent Rayleigh-Brillouin Scattering in Molecular Gases." *Physical Review A* 69 (3): 033814. https://doi.org/10.1103/PhysRevA.69.033814.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., et al., 2011: "Scikit-Learn: Machine Learning in Python." *J. Mach. Learn. Res.* 12 (85): 2825–30.

Tenti, G., Boley, C. D., Desai, R.C., 2011: "On the Kinetic Model Description of Rayleigh–Brillouin Scattering from Molecular Gases." *Canadian Journal of Physics*, September. https://doi.org/10.1139/p74-041.

Witschas, B., 2011: "Analytical Model for Rayleigh–Brillouin Line Shapes in Air." *Applied Optics* 50 (3): 267–70. https://doi.org/10.1364/AO.50.000267.

GALA e.V., Karlsruhe, Germany, ISBN 978-3-9816764-9-5