New perspectives in the analysis of Stark width regularities and systematic trends

Zlatko Majlinger[1,2,3], Ivan Traparić[4]

[1] *Astronomical Observatory, Volgina 7, 11060 Belgrade 38, Serbia*
[2] *ES "I. G. Kovačić", Kralja Tomislava 18, 51326 Vrbovsko, Croatia*
[3] *ES "V. Gortan", Prilaz Vladimira Gortana 2, 51000 Rijeka, Croatia*
[4] *Institute of Physics Belgrade, Pregrevica 118, 11080, Belgrade, Serbia*

E-mail: zlatko.majlinger@gmail.com, traparic@ipb.ac.rs

Abstract

Regularites and systematic trends among the sample of Stark widths obtained by using modified semiempirical method from the STARK-B database were analysed. Two different approaches are independently used – multiple regression method combined with simple cluster analysis, and random forest (RF) machine learning algorithm. Predicted values of Stark widths calculated with estimates obtained from multiple regression method, and those values predicted by using RF algorithm, were compared with already known corresponding experimental Stark widths published elsewhere. Results of this analysis indicate that both of these methods can mostly predict new Stark width values within the acceptable range of accuracy.

Introduction

In this investigation, we focused on searching systematic trends among the great amount of Stark widths from STARK-B database (Sahal-Brechot et al. 2014b, 2015), obtained by modified semiempirical (MSE) method (Dimitrijević and Konjević, 1981) as a continuation of our previous work on determination of unknown MSE Stark widths and studying of regularity and systematic trends (RST) among the MSE Stark broadening parameters (see, for example, Majlinger et al, 2015, 2017a, 2017b, 2020). Two different methods are used to analyse the sample – classical statistical regression method, which has already been used in previous investigations of regularities and systematic trends, and random forest (RF) algorithm from a group of machine learning methods, which become very popular methods more often used in these days whenever some classification or non-linear regression is needed to be performed.

Methods Of Calculation

According to statistical analysis of Stark widths calculated for 143 transitions from 26 multiply charged ions of 17 elements using the modified semiempirical method (for example, most of them are elaborated by Dimitrijević and Konjević, 1981), and previous assumptions, new estimates of Stark widths were found. After providing simple cluster analysis (Aggarval and Reddy, 2014) and multiple regression analysis (for example, Chatterjee and Simonoff, 2013), we concluded that MSE Stark width sample has to be devided in three separate groups, according to corresponding transition type. From the regression lines obtained for the each group of transitions, new estimates for prediction of new Stark width values are found. Correlation between new estimates of full Stark width at half maximum (FWHM) and existing MSE values for transition type I, type II and type III with corresponding regression lines are displayed in Figs. 1-3 respectively. Estimates for each of these transition types are marked as W_{E1}, W_{E2} and W_{E3} in Figs. 1-3 respectively or as W_{EST} in Fig. 4 and Fig. 7 . As the additional attemption to confirm a validation of this method, predicted Stark widths from multiple regression analysis are compared with corresponding experimental values published elswhere (Konjević et al, 1984, 2002). Results of this comparison is presented in Fig.4.

As machine learning represents a very popular tool for different types of problems encoutered in science, here it was applied on the study of regularities of Stark effect. Machine learning model used in our additional analysis is based on Random Forest (RF) algorithm, which has already been used and described elsewhere (Tapalaga et al, 2022). Before developing the model, we needed to develop and create a database for training and testing of the future models. This database was created as a combination of two databases, namely NIST atomic database (Kramida et al, 2022), from which we took atomic parameters of interest for every transition and STARK-B database (Sahal-Bréchot, 2020) from which we took Stark width and plasma parameters for each calculated width. Additional parameters for training the algorithm, already used in previous analyses of Stark width regularities and systematic trends, are taken also from Purić and Šćepanović (1999). As in the case of multiple regression method, Stark widths predicted with using RF algorithm were compared with corresponding experimental widths from the same references (Konjević et al, 1984, 2002). Results of this comparison are shown in Fig.5 and Fig. 6.

As a final proof that both of presented methods could be valid, in Fig. 7 we presented results obtained from mutual comparison of Stark width values predicted with these two different approaches. Figure 7 shows that both of these two methods are equivalent, e. g. the results of the estimates with RF algorithm and classical multiple regression statistical method are almost the same.

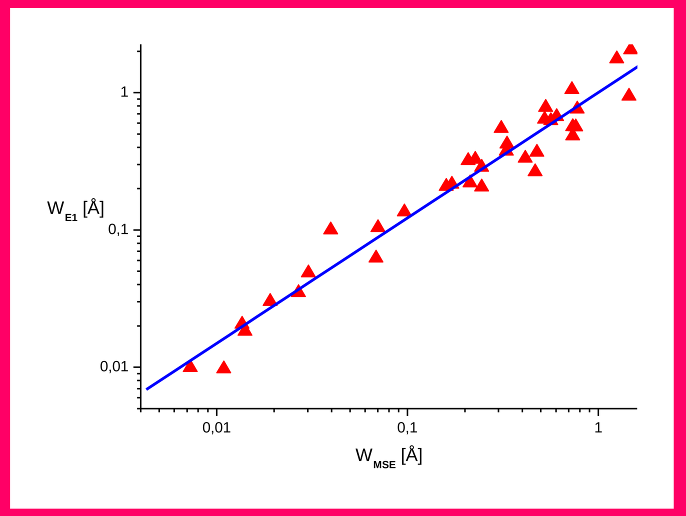

Fig. 1. Log-log correlation between FWHM Stark width values obtained by using estimates from multiple regression analysis (W_{E1}) and MSE values (W_{MSE}) , with corresponding regression line.

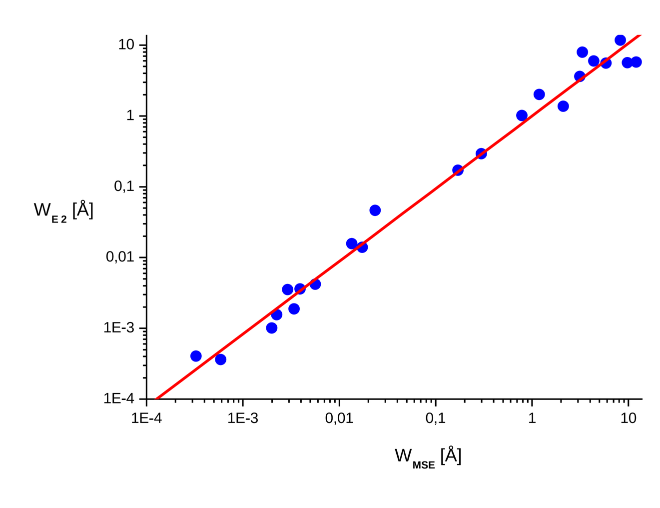

Fig. 2. Log-log correlation between FWHM Stark width values obtained by using estimates from multiple regression analysis (W_{E2}) and MSE values (W_{MSE}) , with corresponding regression line

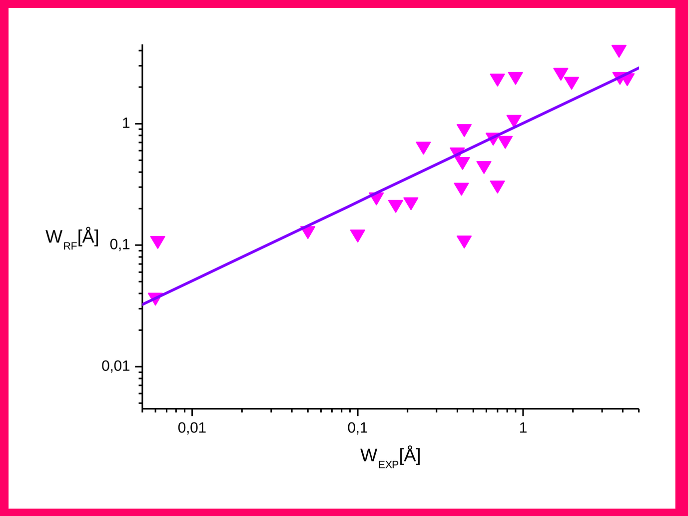

Fig. 5. Log-log correlation between FWHM Stark width values obtained by using RF algorithm (W_{RF}) and corresponding experimental values (W_{EXP}) , with corresponding regression line

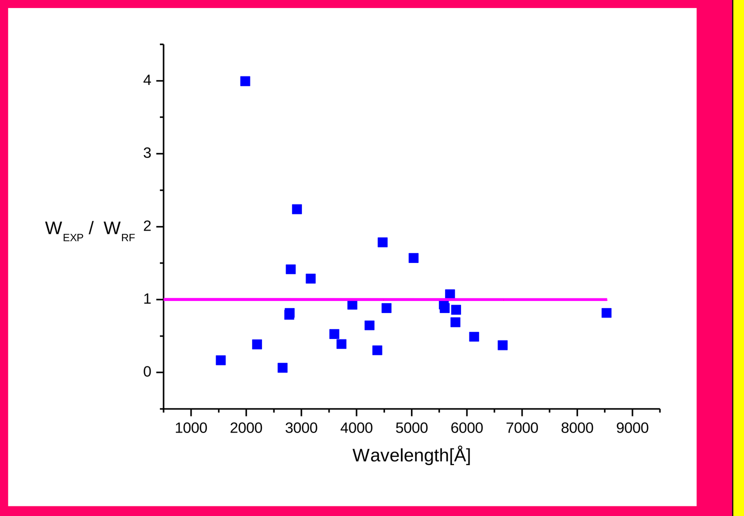

Fig. 6. Dependence of ratio between experimental FWHM Stark widths (W_{EXP}) and corresponding values obtained by using RF algorithm (W_{RF}) on wavelength of spectral lines for which Stark widths are calculated.

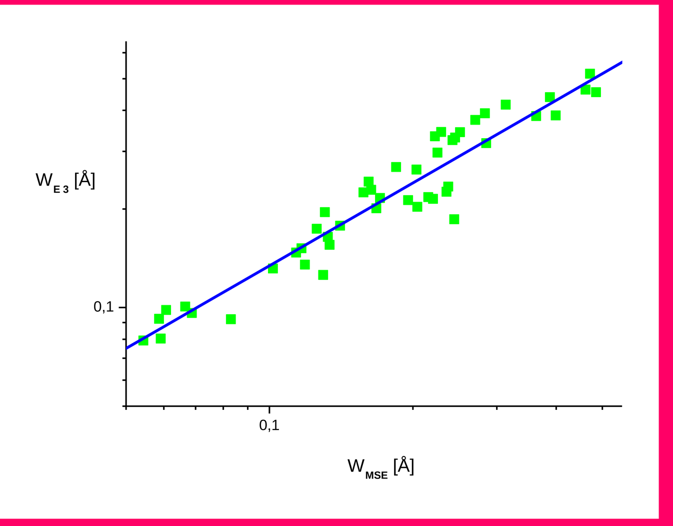

Fig. 3. Log-log correlation between FWHM Stark width values obtained by using estimates from multiple regression analysis(W_{E3}) and MSE values (W_{MSE}) , with corresponding regression line

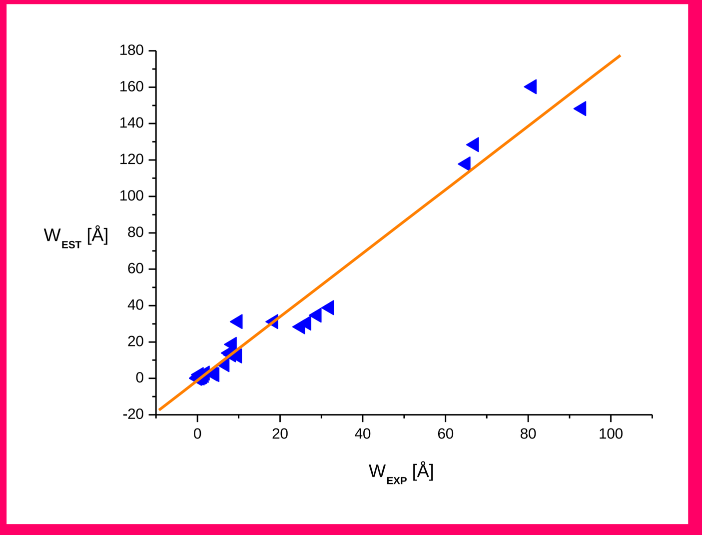

Fig. 4. Linear correlation between FWHM Stark width values obtained by using estimates from multiple regression analysis (W_{EST}) and corresponding experimental values (W_{EXP}) , with corresponding regression line.

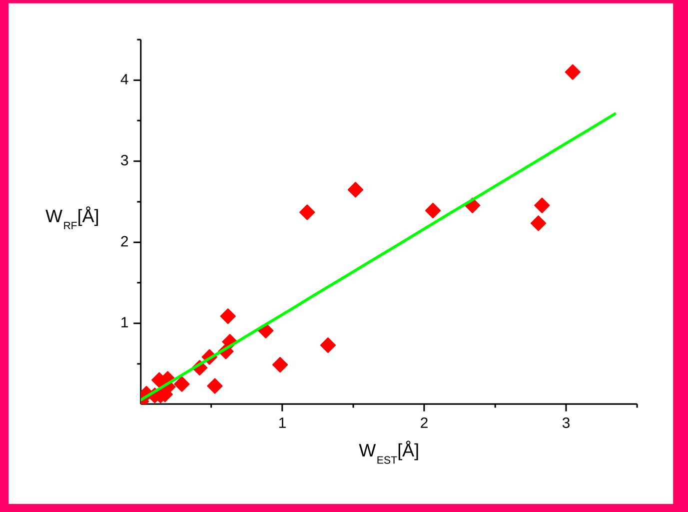

Fig. 7. Comparison between FWHM Stark width values obtained by using estimates from multiple regression analysis (W_{EST}) and corresponding values obtained by using RF algorithm(W_{MSE}) , with corresponding regression line.

References

Aggarval, C. C., and Reddy, C. V., 2014, *Data Clustering*, CRC Press, Taylor and Francis Group, Boca Raton

CChatterjee, S., and Simonoff, J. S., 2013, *Handbook of Regression Analysis*, John Wiley and Sons, Hoboken, New Jersey

Dimitrijević, M. S., Konjević, N.: 1981, *Spectral Line Shapes*, Walter de Gruyter & Co., Berlin – New York

Konjević, N, Dimitrijević, M. S., and Wiese, W. L., 1984, *J., Phys. Chem. Ref. Data*, 13(3), 469

Konjević, N, Lessage, A., Fuhr, J. R. and Wiese, W. L., 2002, *J., Phys. Chem. Ref. Data*, 31(3), 819

Kramida, A., Ralchenko, Yu., Reader, J., NIST ASD Team (2022). *NIST Atomic Spectra Database* (ver. 5.10), [Online]. Available on: *http://physics.nist.gov/asd* [2016, January 18]. National Institute of Standards, Technology, Gaithersburg, MD.

Purić, J., Šćepanović, M., 1999, *Astrophys. J.,* 521, 490

Sahal-Bréchot, S., Dimitrijević, M. S., Moreau N., 2015. STARK-B database, [online]. Available: http://stark-b.obspm.fr [October 1, 2016]. Observatory of Paris, LERMA, Astronomical Observatory of Belgrade

Tapalaga, I., Traparić, I., Trklja Boca, N. et al., 2022, *Neural Comput & Applic* 34, 6349

Conclusion

Average value of the ratio between new calculated estimates and corresponding experimental Stark widths from references (Konjević et al, 1984, 2002) refers on the accuracy of prediction of new Stark widths using the estimates form multiple regression analysis in a range between ±30% and ±50%, while accuracy of using RF algorithm is estimated on a range between ±20% and ±30%. This is a good evidence that both of these methods can be apply in predicting of Stark width values for the spectral lines from simple spectra. For the application of these methods to study regularities and systematic trends among the Stark broadening parameters of lines in more complex spectra, additional investigations are needed.