

機械学習による光電子収量分光 (PYS) スペクトルの自動閾値予測

柳生進二郎¹・吉武道子¹・知京 豊裕¹・長田貴弘¹

¹国立研究開発法人 物質・材料研究機構 〒305-0044 茨城県つくば市並木 1-1

(2020年2月7日受付; 2020年4月10日掲載決定)

Automatic threshold prediction of photoelectron yield spectroscopy (PYS) by machine learning

Shinjiro Yagyu¹, Michiko Yoshitake¹, Toyohiro Chikyow¹, Takahiro Nagata¹

¹National Institute for Materials Science, Tsukuba, Ibaraki 305-0044

(Received February 7, 2020; Accepted April 10, 2020)

The prediction performance of the automatic threshold estimation of photoelectron yield spectroscopy using machine learning and least squares regression (fitting) was verified using 82 measured data. The correct answer rate was greater than 80% for machine learning and less than 50% for fitting, with an error range of ± 0.3 eV with respect to the correct answer (the result of human spectrum analysis). To further improve the correct answer rate, it is necessary to change the energy range and energy step of the measured data because increase in the nonlinear intensity at the high-energy side of the spectrum is considered to affect the automatic estimation. The estimation did not change with changes in the energy step of the data. However, when the energy range in the data was changed, the estimation improved. If the threshold is expected to be less than 6.0 eV, the prediction is improved by using an energy range of 4.2–6.2 eV.

KEYWORDS: machine learning, threshold, photoelectron yield spectroscopy

1. はじめに

半導体やその電極材料開発において、電子・ホールが移動する価電子帯や HOMO のエネルギーレベルであるイオン化ポテンシャル (IP) は重要な物性量である。この値は、光電子収量分光法 (Photoelectron Yield Spectroscopy (PYS)) にて計測したスペクトルの閾値 (IP 位置) から求められる。(PYS は、試料に徐々にエネルギーを増加させた紫外光を照射し、試料から放出される光電子数を計測する方法である。) 閾値は、Fowler の光電効果の理論¹⁾をもとに、横軸に光のエネルギー、縦軸に光電子放出強度 (PYS 強度) の平方根をプロットし、その傾きとバックグラウンドとの交点から求める解析法が多く用いられている²⁻⁴⁾。また、有

機半導体材料では、PYS 強度の 1/3 乗をとる解析法も用いられている⁵⁾。現状では、解析者が一つ一つのスペクトルを見て閾値を求めている。このため、解析者の経験などによる閾値解釈の違いや解析時間などから、解析者に依存しない解析の自動化が求められている。自動解析で最初に適用を考えるのは、最小二乗回帰 (Fitting) である。Fowler の理論から解析式が導かれており、これを用いて Fitting を行うことができる。我々は、この Fitting を用いて閾値判定を試みたところ、解析者が判定した値と大きくかけ離れる値となった⁶⁾。そこで、別のアプローチとして、教師あり機械学習を検討した。機械学習は、画像認識に代表されるように近年のコンピュータパワーの増大とビッグデータ化で急速に発達している⁷⁾。この技術は様々

*E-mail: YAGYU.Shinjiro@nims.go.jp

による解析方法を実際の測定データ（今後検証データと呼ぶ）に適用し、その予測性能について検討した¹²⁾。検証データは、理研計器社製 AC 装置にて測定したもので、材料名やその属性が明らかであり、かつ閾値解析（以降正解とする）を行った 87 個である。材料の内訳は、Au : 31, Au 以外の金属 : 16, 酸化物・半導体 : 15, 有機物 : 25 である。87 個の検証データによる予測結果を Fig. 1 に示す。なお、グラフの各点には測定試料名（試料によっては、アンダーバー以降は照射パワー）が記載されている。横軸は解析者が解析した値（正解）、縦軸が予測した結果を示し、対角線に近いほど正解に近い予測である。対角線から ± 0.3 eV ずれた線も示す。正解に対して ± 0.3 eV の誤差範囲とした場合の正答率は、機械学習の Gradient boosting (GB) モデルでは、90%、同じく Random Forest (RF) では 80% 以上となった。一方で、従来の Fitting (Fit) では、50% を切っており、多くのデータは、収束しておらず不定 (7.0 eV, 図の範囲にはない) となっている。本報告では、測定データの測定エネルギー範囲を変更し、機械学習及び Fitting による予測性能の向上について検討を行った。

2. 機械学習とそのデータセット

機械学習におけるデータセットは、Fowler の解析式による計算データより作成した¹⁾ (Fowler の解析式は

Appendix に記載する)。Fowler の解析式におけるパラメータの範囲やステップ (パラメータは、閾値、温度、規格化定数、バックグラウンド、測定エネルギー範囲) は、これまで我々が理研計器社製 AC 装置にて計測した 87 個のデータを観察し決定した。実際の測定データは、測定エネルギー範囲やそのエネルギーステップが異なるために、それに対応する測定エネルギー範囲及びステップを用意した。(AC 装置を用いた一般的な測定エネルギー範囲は、4.2–6.2 eV, 4.0–7.0 eV, 仕事関数が高い場合は 5.0–7.0 eV であり、ステップは 0.1 eV と 0.05 eV である。) AC 装置のオープンカウンターの性能上 20000 カウント以上は、飽和してしまい計測することができない。これらの条件などを組み合わせて以下のパラメータを用意した。

測定エネルギー範囲については、4.0–7.0 eV, 4.2–6.2 eV, 5.0–7.0 eV の 3 つを用意し、それぞれに対し、0.05 eV 刻み、0.1 eV 刻みの 2 つ、合計 6 種類を用意した。閾値 (IP) の範囲については、測定エネルギー範囲に依存して、最初から 3 点目の値から最後から 5 点目の値までとし、刻み数は 0.1 eV とした。(例えば、4.0–7.0 eV, 0.05 eV 刻みでは、4.1 eV から 6.8 eV まで 0.1 eV 刻み) この閾値を教師とした。温度 (T) については、100, 300, 500, 700, 900 K とし、バックグラウンド (Bg) については、0, 5, 25, 50, 100, 規格化定数 (Nor) については 0.1 から 3.0 まで 0.1 刻

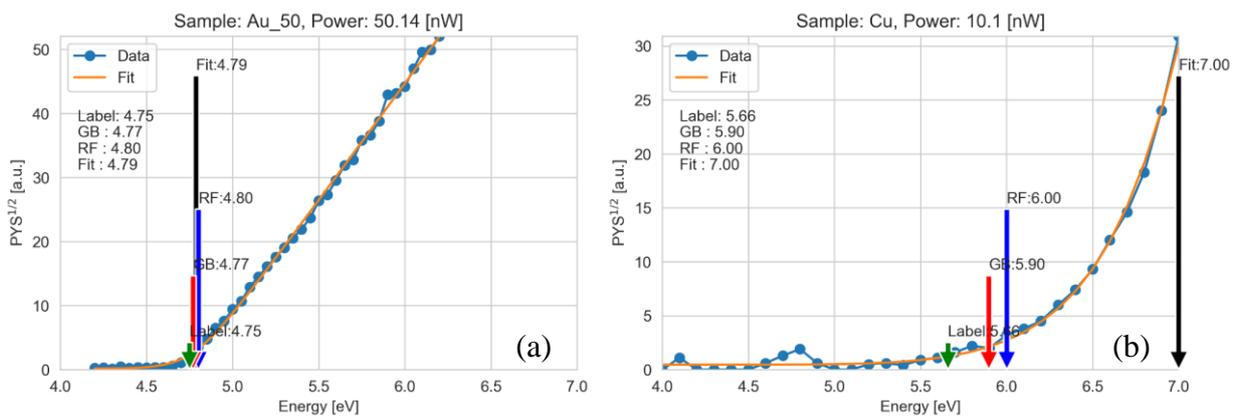


Fig. 2: Representative spectra and their prediction results (RF, GB, and fitting)—spectra indicating (a) similar predictions for all models and (b) large predicted values of the fitting; in (a), the spectral intensity increases linearly from the threshold; On the other hand, in (b), the spectral intensity increases nonlinearly in the high-energy region (specifically energy exceeding 6.8 eV); the arrows in each graph indicate the energy values predicted by each model (the order of the arrows from the shortest to the longest is label, GB, RF, and fitting)

みとした。なお、最大強度が 20000 を超えるスペクトルについてはデータセットには加えていない。作成したデータはランダム層化分割 (学習: テスト=9:1) を行い、学習用とテスト用に分割した。各説明変数間 (各エネルギーにおける PYS 強度) の強度が大きく異なり、規格化が難しいことから、各説明変数の規格化が必要ない決定木ベースのアンサンブルモデル (Random Forest (RF), Gradient boosting (GB)) のアルゴリズムを用いた。機械学習の計算は、Python の Scikit-learn³⁾ を用いた。なお、ハイパーパラメータ及び汎化性能の調整は、クロスバリデーション及びグリッドサーチを行って決定した。機械学習における学習データ、テストデータともに小数第 2 位を四捨五入して 1.0 (100%) の正答率となった²⁾。

3. 結果及び考察

機械学習の予測の向上には、学習データの見直しや増加が考えられる。前述の通り、テストにおける正答率はすでに 100% となっているためこれ以上、学習データを見直すことは得策ではないと考えられる。そこで、予測に用いる測定スペクトルについて再度検討を行った。Fig. 2 にどの予測も一致している Au のスペクトル (a) 及び、Fitting が大きく外れている Cu のスペクトル (b) を示す。それぞれのスペクトルには、Fitting による回帰線も示してある。なお、機械学習では、スペクトルデータに対して閾値のセットで学習をさせているために、スペクトルを与えると、閾値のみ出力される。(a) では、閾値を境に強度が直線的に増

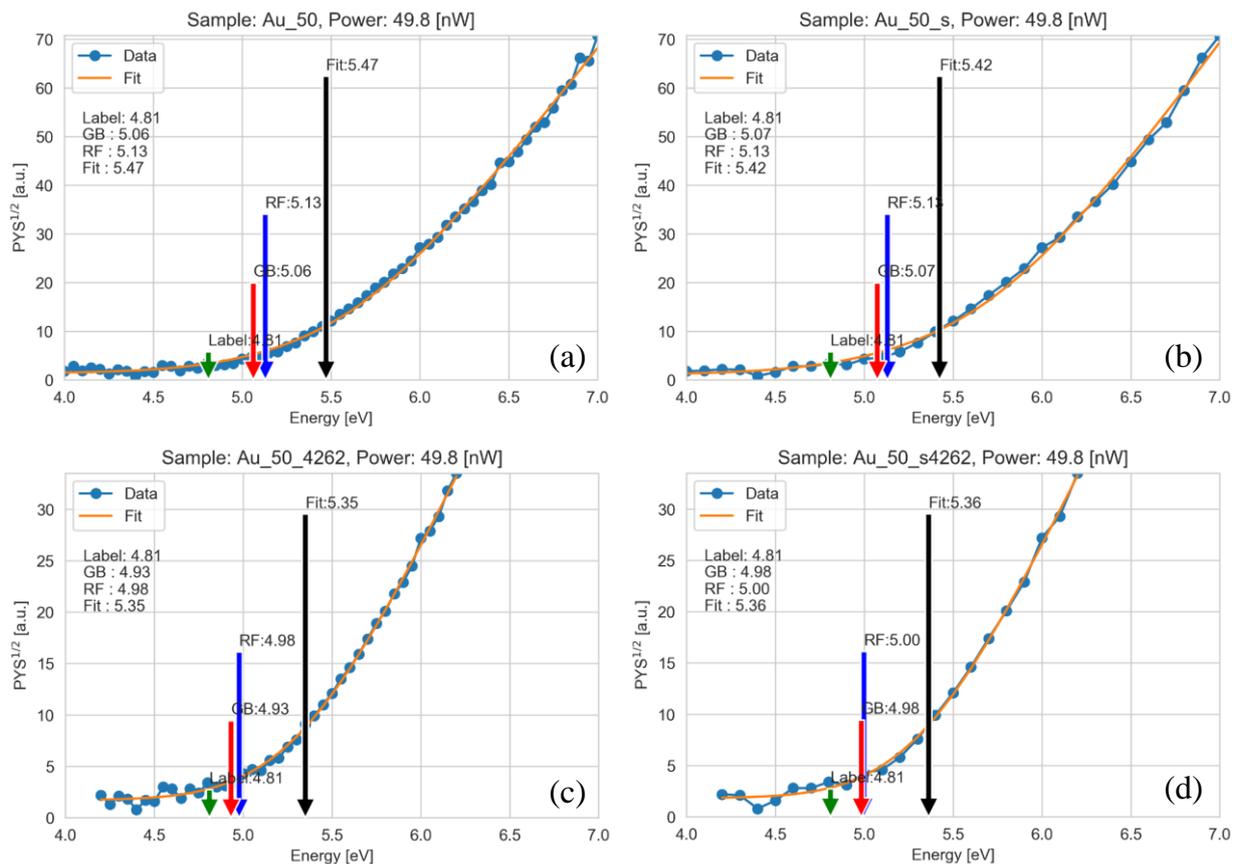


Fig. 3: Modified Au data with respect to energy range and steps—(a) shows the original data for an energy range and step of 4.0–7.0 eV and 0.05 eV, respectively, and (b), (c), and (d) show the data for an energy range of 4.0–7.0 eV and an energy step of 0.1 eV, an energy range of 4.2–6.2 eV and an energy step of 0.05 eV, and an energy range of 4.2–6.2 eV and an energy step of 0.1 eV

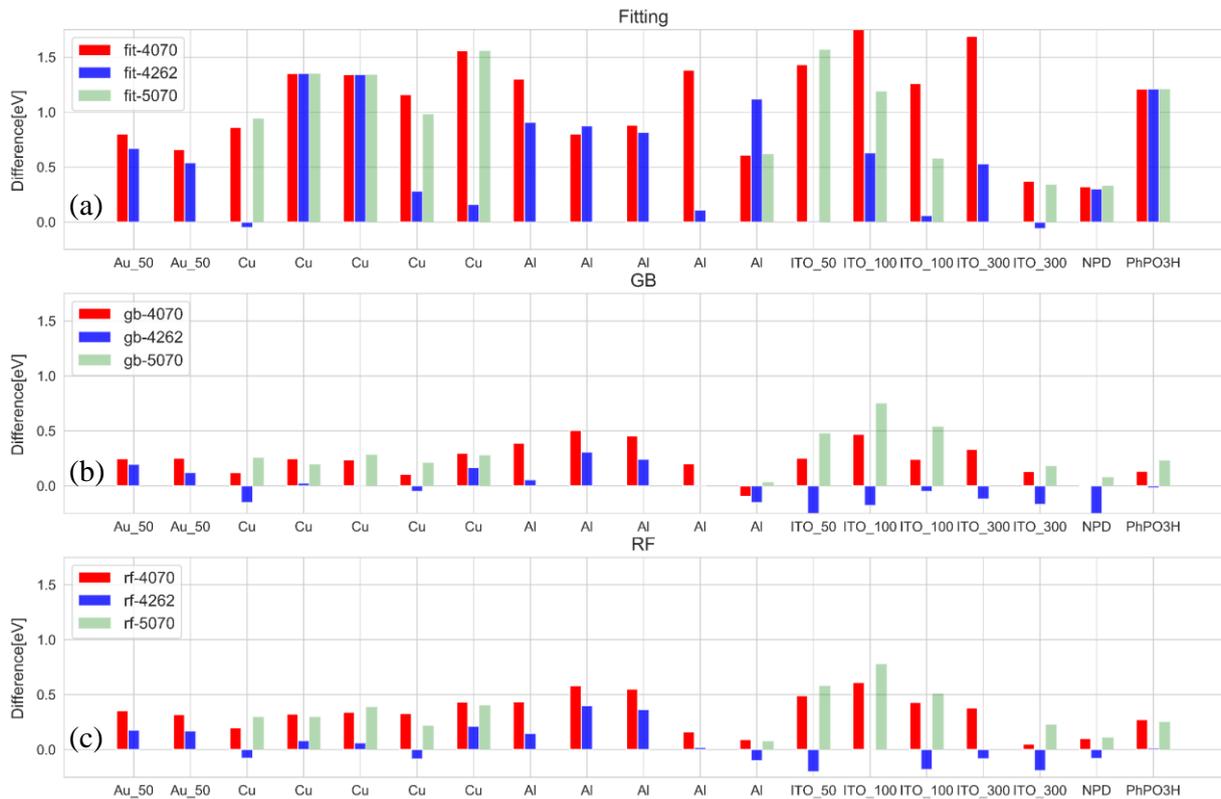


Fig. 4 The bar graph of the difference between the prediction and correct answer for fitting (a), GB (b), and RF (c) for each sample— The bar graphs are arranged from left in the following order: the original, 4.2–6.2 eV, and 5.0–7.0 eV

加しているのがわかる。一方で(b)では、エネルギーの高い領域（特に6.8 eVより高エネルギー側）で、非線形的に強度の増加がみられる。これは、Fowlerのモデル（閾値を境に直線的に増加している成分）とは別の成分が重畳されていると考えられ、Fittingはその成分に大きな影響を受けていると考えることができる。そこで、測定エネルギー範囲、ステップ数の変更による予測値の変化について機械学習及びFittingでの検討を行った。測定エネルギー範囲が4.0–7.0 eVで計測された測定データを対象に、エネルギー範囲を4.2–6.2 eV及び5.0–7.0 eV（閾値が5.0 eV以上にある場合のみ）に変更したスペクトルを作成した。またステップは、0.1と0.05 eVがあり、0.05 eVステップの点を削減して0.1 eVステップのスペクトルを作成した。Fig. 3は、Au試料においてFittingが大きく予測を外しているスペクトルである。Fig. 3(a)はAuの4.0–7.0 eV、0.05 eVのオリジナルデータ、(b)は0.05 eVの点を削減して0.1 eVステップに変更したも

の、(c)は4.2–6.2 eV、0.05 eVに変更したものの、(d)は4.2–6.2 eV、0.1 eVに変更したものを示す。ステップを変更した(a)と(b)、(c)と(d)は、どのモデルも予測値の変化はあまり見られなかった。一方で、エネルギー範囲を変更した(a)と(c)、(b)と(d)では、機械学習、Fittingともにおよそ0.1 eV程度、正解に近い値の方向に移動した。機械学習において、エネルギー範囲、ステップ数の変更はともに説明変数の数を減らすことを意味する。ステップの削減に関してはどの予測もあまり影響していないが、エネルギー範囲の変更については、影響を受けていることが明らかになった。Fig. 4に4.0–7.0 eVで計測されたデータ(Original)を4.2–6.2 eV、5.0–7.0 eVの範囲に変更したときのFitting (a)、機械学習のGB(b)及びRF(c)の正解との差を、試料ごとに棒グラフで示す。各試料における棒グラフは左からOriginal、4.2–6.2 eV、5.0–7.0 eVの範囲の時の正解との差である。なお、5.0–7.0 eVについては、正解の閾値が5.0 eV以上のものだけである。

Table 1 Differences between the correct answer and the predictions for fitting, GB, and RF—the number enclosed in parentheses indicates the energy range; the bottom column indicates the number of spectra within ± 0.3 eV, and the second column from the bottom shows the average absolute error; the shaded values are within ± 0.3 eV

Name	Estimation	Difference Fitting			Difference GB			Difference RF		
		original (4.0,7.0)	(4.2, 6.2)	(5.0, 7.0)	original (4.0,7.0)	(4.2, 6.2)	(5.0, 7.0)	original (4.0,7.0)	(4.2, 6.2)	(5.0, 7.0)
Au_50	4.83	0.80	0.67		0.25	0.19		0.35	0.18	
Au_50	4.81	0.66	0.54		0.25	0.12		0.32	0.17	
Cu	5.34	0.86	-0.05	0.94	0.12	-0.15	0.26	0.20	-0.08	0.30
Cu	5.65	1.35	1.35	1.35	0.25	0.02	0.19	0.32	0.08	0.30
Cu	5.66	1.34	1.34	1.34	0.24	-0.01	0.28	0.34	0.06	0.39
Cu	5.62	1.16	0.28	0.98	0.10	-0.05	0.21	0.33	-0.08	0.22
Cu	5.44	1.56	0.16	1.56	0.30	0.17	0.28	0.43	0.21	0.40
Al	4.95	1.30	0.91		0.39	0.06		0.43	0.15	
Al	4.49	0.80	0.87		0.50	0.31		0.58	0.40	
Al	4.56	0.88	0.81		0.46	0.24		0.55	0.36	
Al	4.66	1.38	0.11		0.20	-0.01		0.16	0.02	
Al	5.88	0.61	1.12	0.62	-0.10	-0.15	0.03	0.09	-0.10	0.08
ITO_50	5.32	1.43	0.01	1.57	0.25	-0.30	0.48	0.49	-0.20	0.58
ITO_100	5.02	1.77	0.63	1.19	0.47	-0.18	0.75	0.61	0.00	0.78
ITO_100	5.19	1.26	0.06	0.58	0.24	-0.05	0.54	0.43	-0.18	0.51
ITO_300	4.97	1.69	0.53		0.33	-0.12		0.38	-0.08	
ITO_300	5.17	0.37	-0.06	0.34	0.13	-0.17	0.18	0.05	-0.19	0.23
NPD	5.49	0.32	0.30	0.33	0.01	-0.31	0.08	0.10	-0.08	0.11
PhPO3H	5.78	1.21	1.21	1.21	0.13	-0.02	0.23	0.28	0.01	0.25
Mean Absolute Error		1.17	0.73	1.09	0.28	0.17	0.35	0.38	0.17	0.40
Within ± 0.3		0/19	8/19	0/12	15/19	17/19	9/12	6/19	17/19	7/12

対象となった計測データは 19 個であり、5.0–7.0 eV については 12 個である。棒グラフが短いほど正解に近いことを示す。全体的に 4.2–6.2 eV が短くなっているのがわかる。5.0–7.0 eV は、オリジナルと同程度かそれよりも差が大きくなるものが多くみられる。また、全体として、正解よりも高い値を見積もる傾向が見られる。Fig. 4 の作成に用いた表を Table 1 に示す。括弧内の数字は、エネルギー範囲を示している。Fitting において、Original では ± 0.3 eV に入るもの（今後、範囲内と呼び、Table ではハッチされている）0 個（19 個のうち）であった。4.2–6.2 eV では、8/19（42%）、5.0–7.0 eV では 0 個（12 個のうち）であった。GB では、Original が 15/19（78%）、4.2–6.2 eV では 17/19（89%）、5.0–7.0 eV では 9/12（75%）。RF では、6/19（31%）、17/19（89%）、7/12（58%）となった。それ

ぞれの平均絶対誤差を計算すると、Fitting では、オリジナルデータが 1.17 eV に対して、4.2–6.2 eV では 0.73 eV まで減少した。5.0–7.0 eV についても 1.07 eV まで減少しているがその減少幅は小さい。機械学習においては、オリジナルに比べ 4.2–6.2 eV では、Fitting と同様に減少した。特に範囲内（ ± 0.3 eV）を大きく下回る結果となった。逆に 5.0–7.0 eV にすることで値は増加した。4.2–6.2 eV 及び 5.0–7.0 eV における Fitting、機械学習の予測の違いは、測定点の数は同じであることから、スペクトル形状の違いに起因するものであると考えられる。これは、高エネルギー領域（6.8 eV 以上）における強度の非線形的な増加の影響であると考えられることができる。

このことから機械学習、Fitting 性能の向上には、適切な測定エネルギー範囲を選ぶことが重要であると

考えられる。閾値が 6.0 eV 以下のものであれば、最大エネルギー 6.2 eV を上限に選択する事で十分であると考えられる。また 6.0 eV より高い閾値の場合でも、なるべく高エネルギー側の測定点を減らすことが必要であると考えられる。高エネルギー側での非線形な強度の増加の要因については、測定材料の状態密度分布の影響なども考えられるが、計測器側についても再度検討が必要であると考えられる。PYS 強度（1 光子当たりの放出量）を求める場合、照射光の光量補正が必要である。6.5 eV 以上 (190 nm 以下) では、酸素などの吸収による光量の著しい低下が考えられる。そのため、光量補正がより重要になってくる。

現状の測定において、正解との差 (± 0.3 eV) の違いを認めるならば自動解析手法として機械学習を利用することが可能であると考えられる。一方で閾値が 6.0 eV より小さいことが予想されるものについては、エネルギー範囲を 4.2-6.2 eV に制限したものにおいて予測を行うことにより性能の向上が見込めると考えられる。一方でどの程度まで測定エネルギー範囲を制限すべきか、どの程度まで Fowler の解析式が適用できるかについては、さらなる検討が必要である。教師あり機械学習は、前述の通りパターン認識であり、説明変数が与えられたときにそのパターンに一番合っている値が出力される。今回は、各エネルギーでのスペクトル強度（説明変数）と閾値（目的変数）で学習させている。機械学習ではこの説明変数と目的変数を物理的に意味を持たない多項式で関係づけている。Fowler の解析式は、フェルミディラック分布を積分しただけのモデルであるために、価電子帯構造などは、考慮されていない。従って、機械学習による判定結果はそのモデルに合わせた場合の値である。電子構造などを考慮した予測を行うには、第一原理計算を用いて閾値とそこから導き出されるスペクトルを計算し、それを学習することで可能になると考えられる。

4. まとめ

機械学習と最小二乗 Fitting による自動推定において、高エネルギー側（特に 6.8 eV 以上）における非線形的な強度の増加が閾値推定に大きな影響を与えると考えられる。そのため、測定スペクトルのエネルギー範囲及びエネルギーステップを変更したスペクトルを作成し、そのスペクトルに対する予測について検討した。ステップ数を変更したスペクトルについて、変更前後で予測値に変化はほぼ見られなかった。一方、エネルギー範囲の変更は、変更により予測値が変化す

ることも明らかになった。特に、高エネルギー側の測定点を削除したスペクトルは、多くのスペクトルで予測性能の向上がみられた。閾値が 6.0 eV より小さいことが予想されるものについては、エネルギー範囲を 4.2-6.2 eV に制限したものによって予測を行うことにより性能の向上が見込めると考えられる。

Appendix

Fowler の解析式

Fowler の解析式^{1,2)}は以下の通りである。

$$PYS^{1/2}(\chi, T, BG, Nor) = (Nor \cdot f(\mu(\chi, T)) + BG)^{1/2}$$

ここで、 χ はイオン化ポテンシャル（閾値の IP ）、 k_B はボルツマン定数、 m は電子質量、 $h\nu$ はエネルギー、 T は温度である。

$$(h\nu - \chi)/kT = \mu \leq 0 \text{ の時}$$

$$f(\mu) = \left\{ e^{\mu} - \frac{e^{2\mu}}{2^2} + \frac{e^{3\mu}}{3^2} - \dots \right\}$$

$$(h\nu - \chi)/kT = \mu \geq 0 \text{ の時}$$

$$f(\mu) = \left[\frac{\pi^2}{6} + \frac{1}{2}\mu^2 - \left\{ e^{-\mu} - \frac{e^{-2\mu}}{2^2} + \frac{e^{-3\mu}}{3^2} - \dots \right\} \right]$$

となる。

文 献

- 1) H. Fowler: Phys. Rev. **38**, 45 (1931).
- 2) E.O. Kane : Phys. Rev. **127**, 131 (1962).
- 3) 山下大輔 : 早稲田大学大学院基幹理工学研究科博士論文 (2015).
- 4) 理研計器 AC2, AC3, AC5 取扱説明書
- 5) M. Kochi, Y. Harada, T. Hirooka and H. Inokuchi: Bulletin of the Chemical Society of Japan **43**, 2690 (1970).
- 6) S. Yagyu, M. Yoshitake and T. Chikyow : Vacuum and Surface **61**, 196 (2018). (in Japanese), doi:10.1380/vss.61.196
- 7) 例えば、福島俊一、藤巻遼平、岡野原大輔、杉山将: 情報管理, **60**, 543 (2017).
- 8) 松波成行、柳生進二郎、篠原 正、片山英樹、須藤 仁、服部康男、平口博丸: 土木学会論文集 A1 (構造・地震工学) **75**, 141 (2019), doi: 10.2208/jscejsee.75.141.
- 9) See, for example: H. Yamada, C. Liu, S. Wu, Y. Koyama, S. Ju, J. Shiomi, J. Morikawa, R. Yoshida: ACS Central Science **10**, 1717 (2019), doi:10.1021/acscentsci.9b00804.
- 10) See, for example: D. Jha, A. Kusnef, R. Al-Bahrani, N. Nguyen, W. Liao, A. Choudhary, A. Agrawa: 2019 International Joint Conference on Neural Networks (IJCNN) (2019) p. 14., doi:

10.1109/IJCNN.2019.8852096.

- 11) C.M. ビショップ (著), 元田 浩 (監訳), 栗田 多喜夫 (監訳), 樋口 知之 (監訳), 松本 裕治 (監訳), 村田 昇 (監訳): “パターン認識と機械学習 上” (丸善, 2012) .
- 12) S.Yagyu, M.Yoshitake, T.Nagata, T.Chikyow: Vacuum and Surface **62**, 504 (2019). (in Japanese), doi:10.1380/vss.62.504.
- 13) <https://scikit-learn.org/stable/>