JAAS

ROYAL SOCIETY OF CHEMISTRY

TECHNICAL NOTE

View Article Online
View Journal | View Issue

Cite this: *J. Anal. At. Spectrom.*, 2025, **40**, 1658

Two-stage semi-supervised machine learning for classification of Ti-rich nanoparticles and microparticles measured by spICP-TOFMS†

Raven L. Buckman Johnson, [a] Hark Karkee [a] and Alexander Gundlach-Graham *[ab]

Single-particle inductively coupled plasma time-of-flight mass spectrometry (spICP-TOFMS) can be used to measure metal-containing nanoparticles (NPs) and sub-micron particles (μPs) at environmentally relevant concentrations. Multielement fingerprints measured by spICP-TOFMS can also be used to differentiate natural and anthropogenic particle types. Thus, the approach offers a promising route to classify, quantify, and track anthropogenic NPs and μPs in natural systems. However, biases in spICP-TOFMS data caused by analytical sensitivities, Poisson detection statistics, and elemental variability at the single-particle level complicate particle-type classification. To overcome the inherent bias in spICP-TOFMS data for the classification of particle types, we have developed a multi-stage semi-supervised machine learning (SSML) strategy that identifies and subsequently trains on systematic noise in spICP-TOFMS data to produce more robust particle-type classifications. Here, we apply our two-stage SSML model to classify individual Ti-containing NPs and μPs *via* spICP-TOFMS analysis. To build our model, we measure neat suspensions of anthropogenic TiO_2 particles (E171) and natural titanium-containing particle types: rutile, ilmenite, and biotite by spICP-TOFMS. Element mass amounts recorded per particle are used to classify particle type by SSML and then systematic particle misclassifications are identified and recorded as uncertainty classes. Following, a second SSML model is trained with the addition of uncertain particle-type categories. With two-stage SSML, we demonstrate low false-positive rates (≤5%) and moderate particle recoveries (50–90%) for all anthropogenic and natural particle types. Two-stage SSML is a streamlined, hands-off method to identify and overcome bias in spICP-TOFMS training data that provides a robust particle-type classification.

Received 21st March 2025
Accepted 29th May 2025

DOI: 10.1039/d5ja00108k

rsc.li/jaas

Introduction

Titanium (Ti) is the ninth most abundant element in the earth's crust and can be derived from a variety of Ti-containing minerals.[1] Naturally occurring titanium is primarily derived from rutile and ilmenite, which together account for approximately 90% of titanium minerals. Smaller amounts are also obtained from other minerals, such as biotite, brookite, and sphene. These minerals undergo various geochemical processes and can degrade into nanoparticles (NPs, one dimension < 100 nm) or submicron-particles (μPs, diameter < 1 μm). In addition to natural Ti nanominerals, Ti-containing NPs and μPs are manufactured for a variety of purposes ranging from consumer products to biomedical applications.[2–4] Food-grade TiO_2 (E171) particles have been widely used as a coloring agent; however, potential health concerns have brought the use of E171 into question.[5,6] The European Commission banned E171 as a food additive in 2022 and prompted several research studies into the prevalence of TiO_2 in food products and environmental samples.[7–13] E171 particles have diameters from the nano- to the submicron-size regimes,[2,5,14] and have been previously characterized with several analytical techniques including transmission electron microscopy (TEM),[15] asymmetric flow field-flow fractionation (AF4),[16] and inductively coupled plasma mass spectrometry (ICP-MS).[17–21] With an increasing interest in identification and source apportionment of Ti-containing anthropogenic particles, methods for single-particle classification of these particle types have also been explored.[22,23]

The use of single-particle inductively coupled plasma time-of-flight mass spectrometry (spICP-TOFMS) to measure and classify natural or anthropogenic NPs and μPs is a growing area of research interest. spICP-TOFMS provides high-throughput measurement and quantification of elements in individual particles, and this data can also be used to classify particle-types based on elemental fingerprints and their mass fractions. Strategies and rules for particle-type classification based on

[a]*Department of Chemistry, Iowa State University, Ames, Iowa, USA. E-mail: alexgg@iastate.edu*

[b]*TOFWERK, Thun, Switzerland*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d5ja00108k

spICP-TOFMS data are still being developed and a consensus has not been reached. In the simplest case, particle-type detection limits and decision tree classification based on known, or measured, unique element associations and ratios can be established.[8,24–26] Other approaches to classify individual particles by spICP-TOFMS use clustering algorithms to group and classify particles according to their detected elemental signatures.[19,20,27,28] Supervised machine learning has also been used to classify particle-types with t-stochastic neighbor embedding (tSNE) and light gradient boosted decision trees,[29] binomial logistic regression,[30] and *k*-nearest neighbor embedding.[31] While all of these examples demonstrate reasonable classification results, few examples of workflows with broad applicability to multiple particle types without extensive refinement have been reported. This is, in part, due to the inherent bias of spICP-TOFMS measurements.

In spICP-TOFMS, elements are only recorded in a given particle if they produce enough signal to be registered as a particle event, *i.e.* more signal than element-specific critical values ($L_{C,sp}$). While $L_{C,sp}$ values depend on steady-state background levels and the background-signal distribution, the likelihood that an element produces signal greater than $L_{C,sp}$ also depends on a number of factors, including the true mass amount of the element present in a given particle, the measurement sensitivity, and random signal variations due to Poisson counting noise.[32] In addition, the shape of the analyte particle size distribution and the heterogeneity of multi-element composition per particle can also lead to systematic biases apparent in spICP-TOFMS data. For example, minor elements in small particles may be undetectable, but become detectable in larger particles. These smaller particles do not have true altered element mass fractions, but—in the spICP-TOFMS data—their multi-element fingerprints lack the minor elements, and so their composition appears to have changed. This is bias in the spICP-TOFMS measurement, and it will influence the element compositions recorded for populations of particles. For single-particle classification, a workflow must be robust enough to account for the uncertainty and bias in a spICP-TOFMS measurement and should be readily adaptable for various particle-types.

To address the limitations of spICP-TOFMS particle classification, we developed a multi-stage semi-supervised machine learning (SSML) strategy.[33] In our multi-stage SSML workflow, determined element mass amounts per particle from spICP-TOFMS measurements of known particle types are used to train an initial decision tree-based ensemble semi-supervised classification model. Misclassified particles from this first SSML model are largely due to systematic biases inherent to spICP-TOFMS measurements. Therefore, we reclassify the misclassified particles as belonging to uncertain class types. These new uncertain classes are incorporated into the second stage SSML model, which then produces more robust classification results. Two-stage SSML classifies particles based on both elemental mass distributions and associated elemental signatures, which allows for a practical and streamlined approach for particles with inherent heterogeneity. In this study, we apply our SSML workflow to classify mixtures of Ti-containing NPs and μPs, *i.e.* E171 (TiO_2), rutile (TiO_2), ilmenite ($FeTiO_3$), and biotite ($(K(Fe^{2+}/Mg)_2(Al/Fe^{3+}/Mg/Ti)([Si/Al/Fe]_2Si_2O_{10})(OH/F)_2$) with median diameters ranging from 100 to 200 nm, and size ranges from 50 to 600 nm.[8] In previous work, two-stage SSML was used to classify Ce-rich natural and anthropogenic particle types.[33] Here, we extend our method for the (more challenging) analysis of a Ti-containing particles; our aim is to further explore the utility of the two-stage SSML classification approach as a tool for spICP-TOFMS data analysis.

Materials and methods

spICP-TOFMS measurements and data

For training our SSML model, suspensions of ground mineral samples of rutile, ilmenite, and biotite were used to represent natural Ti-containing particles. Engineered Ti-containing particles were represented by suspensions of E171 TiO_2 particles. Neat and mixed suspensions were measured by spICP-TOFMS; sample preparation details are described elsewhere.[8] Single-particle measurements were performed with an icpTOF-S2 instrument (TOFWERK, Thun, Switzerland). Liquid suspensions of particles were introduced to the instrument with a microFAST MC autosampler and PFA pneumatic nebulizer (PFA-ST, Elemental Scientific, NE, USA) connected to the torch inlet *via* a baffled cyclonic spray chamber. Element mass quantification was performed with an online microdroplet calibration system, as previously described.[34,35] Data processing was performed using "Time-of-Flight Single-Particle Investigator" (TOF-SPI, ver. 2.7.4), a free-use particle analysis software.[36] User-specified parameters for TOF-SPI are given in Table S1.† Typical sensitivities and critical mass values of our spICP-TOFMS measurements are provided in Table S2.† In our analysis, all analytes were quantified as their total element mass amounts except for Ti, for which the masses of Ti-48 and Ti-49 were determined independently.

Two-stage semi-supervised machine learning

SSML was conducted according to our previously established workflow.[33] Machine learning classification was performed in MATLAB R2023b (MathWorks Inc., MA, USA) with the Statistics and Machine Learning Toolbox™ ver. 23.3; an example of our code and relevant data is available on Github (**https://github.com/TOFMS-GG-Group/SSML_CeNPClassification/tree/TiNPClassification**). For all the analyses described here, only particle events containing detectable amounts of ^{48}Ti were classified. The aim of our study is to demonstrate robust classification of Ti-containing particles. We evaluate the performance of our SSML approach at both the first and second stages of the workflow using two sets of mixtures: the first containing only natural particles types, and the second containing both natural and engineered particles.

Results and discussion

Supervised and semi-supervised machine learning models are inherently limited by the quality of their training sets and can

only accurately predict outcomes for data that are similar to the training data. If the test data and the training data are sufficiently different, the model will perform with (at best) a 50% success rate; ultimately, performing guesswork. SSML, while subject to the same limitations as supervised ML, is allowed more flexibility in its ability to use deductive reasoning to make decisions.[33] In our previous work, we investigated the capabilities of SSML for the classification of Ce-containing particles with similar sizes and overlapping Ce-mass distributions; the Ce-containing particles were estimated to be, on average, between 30 and 45 nm in diameter.[24,32,33] For this case, two-stage SSML was able to achieve an accuracy of 80% with low false-positive rates. In our current work, we measure particles from E171, rutile, ilmenite, and biotite with median diameters ranging from 100 to 200 nm, and size ranges from 50 to 600 nm.[8] In the analysis of particles with this larger size range, we did not find any significant particle-size dependent limitations or improvements regarding the two-stage SSML approach. As long as particles are completely atomized in the ICP and contain detectable elemental mass amounts, two-stage SSML can be used to develop an accurate classification model.

spICP-TOFMS data

From spICP-TOFMS measurements of pristine rutile, ilmenite, biotite, and E171 suspensions, we compared the detected elemental signatures from each particle type. In Fig. S1,† we provide sunburst plots describing the detected elemental signatures, which show that particle events with a variety of single- and multi-metal combinations were measured. Initial observation of the sunburst plots gives an impression of differentiability. However, upon closer examination, it is clear that many of the detected elemental signatures are shared, or are similar, amongst the four particle types. For example, single-element Ti particles made up 98.8% of the E171 sample and were also detected in the three natural suspensions (40.8% rutile, 7.2% ilmenite, and 0.3% biotite). In Fig. 1, we plot the mass distributions of ^{48}Ti in all four particle types. All four of the Ti mass distributions overlap significantly, making classification based on Ti mass amounts difficult.[30,33] Specifically, at the single-particle level, a particle with ^{48}Ti mass within the range shown in Fig. 1 could originate from any of the four samples.

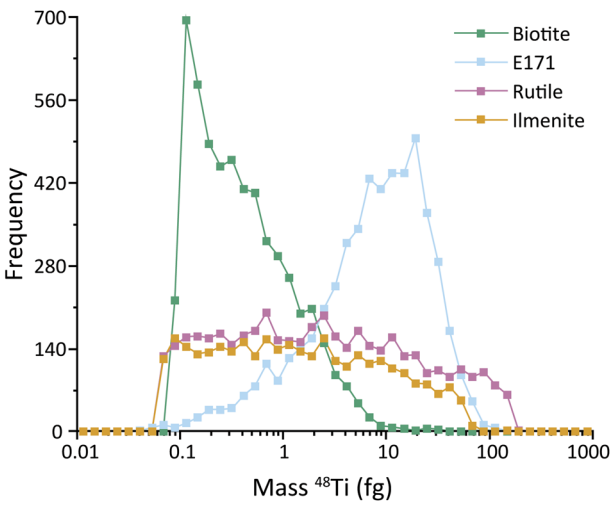

Fig. 1 Mass distributions of ^{48}Ti in the pristine suspensions of biotite, E171, rutile, and ilmenite.

Training and characterization

Quantified mass amounts of elements per particle were used to train the SSML model. In SSML, two types of data are used to train the model: (i) a relatively small dataset with labels assigned based on user-specified categories, and (ii) a larger dataset that remains unlabeled.[33] For the labeled training data used in this application, we randomly sampled 1000 particle events (~20–25% of the Ti-containing particles) from each of the pristine samples; this random sampling was performed to ensure that the labeled training data was as balanced as possible.[33] Particle events from pristine samples were labeled as Rut, Ilm, Bio, or Eng for rutile, ilmenite, biotite, and E171, respectively. The unlabeled training data was gathered by concatenating all the Ti-containing particle events from the pristine suspensions into a single dataset. Details of the machine learning parameters can be found in our previous publication.[33] After the first SSML model is trained, we characterized the performance by using the model to predict classes for both the labeled and unlabeled training data. These results are summarized in Fig. 2 and Table S3.† Receiver operating characteristic (ROC) and precision-recall (PR) curves are provided in Fig. S2† and calculated figures of merit are provided in Fig. S3.†

In Fig. 2A and B, we show the classification performance of the first SSML model for the unlabeled training dataset; classification results for the unlabeled training data are shown in Table S3.† Confusion matrices are used in machine learning research to describe the performance of the model;[37] these can be used to show the true-positives, false-negatives, true-negatives, and false-positives (TP, FN, TN, and FP, respectively). In the confusion matrix in Fig. 2A, the tiles are colored according to whether they were correctly (blue) or incorrectly (pink) predicted by the model. Row- and column-summaries are normalized to the row and column, respectively, and reflect percentages of TPs, FNs, FPs, and positive-predictions (PPs).

The accuracy of the first SSML model was determined to be 90.3 ± 0.2%, which indicates that ~90% of the particles were correctly classified. From Fig. 2A, it can be seen that the first SSML model best classifies biotite and ilmenite particles; these particles, generally, have lower mass amounts of ^{48}Ti and more additional elemental associations than rutile and E171 (see Fig. 1 and S1†). Therefore, the first SSML model is able to achieve low FN and FP percentages for biotite and ilmenite. However, the model does not demonstrate the same robustness for rutile (FP = 13.9%) and E171 (FP = 15.9%) particles. The FP percentages do not reflect the true false-positive rate (FPR) because they do not consider the TNs.[33] FPR is a metric that reflects the probability of a type I error; typically a FPR less than 0.05 (*i.e.* 5%) is desirable. The FPR and other machine learning

figures of merit were calculated based on the results described in Table S3† and are summarized in Fig. S3.† The Eng and Rut particle classes have FPRs of 5.1 ± 0.4% and 5.5 ± 0.7%, respectively, while Bio and Ilm have FPRs of 0.7 ± 0.1% and 1.5 ± 0.1%. While, these FPRs are within an acceptable range, we still observe systematic misclassifications of rutile and E171 particles.

In Fig. 2B, we plot all the particle events in the labeled training dataset as a function of their ^{48}Ti mass (fg); particles are grouped vertically based on their true class and colored according to predicted class. Bubble sizes in Fig. 2B reflect the number of elements detected in the particle event. Here, we see that a subset of E171 particles with a single-metal fingerprint and ^{48}Ti mass less than 2 fg are systematically misclassified as rutile. Similarly, single-metal rutile particles with ^{48}Ti mass between 2 and 20 fg are systematically misclassified as E171. These misclassifications arise from the mass distributions of ^{48}Ti in particles (see Fig. 1). Generally, the E171 particles have larger Ti mass than rutile particles, so the model tends to predict that low-mass single-metal Ti particles are rutile and high-mass single-metal Ti particles are E171. To overcome these systematic misclassifications, we create new labels based on the initial misclassifications and incorporate these labels into a second machine learning model.[33] For example, particle events that were falsely predicted to be Eng by the first model are relabeled as 'unclassifiable engineered', *i.e.* UEng. Likewise, the particles misclassified as Ilm, Rut, or Bio are relabeled as UIlm, URut, or UBio, respectively. These unclassifiable classes are a way to represent the uncertainty of class predictions and also account for the apparent bias in the first SSML model.

In Fig. 2C, we provide the confusion matrix for the second SSML model, which incorporates the unclassifiable particle types. As seen in the second SSML model, the FP percentages for E171 and rutile particles decrease substantially compared to the first SSML model: from 16 to 5% and 14 to 1%, respectively. In our work, we strive for low false-positive classifications because they are important for the classification of particle types with large differences in number concentration. For example, if there are 100× more natural Ti-particles than E171 particles, then the 5% FP in E171 classifications would be more abundant than the

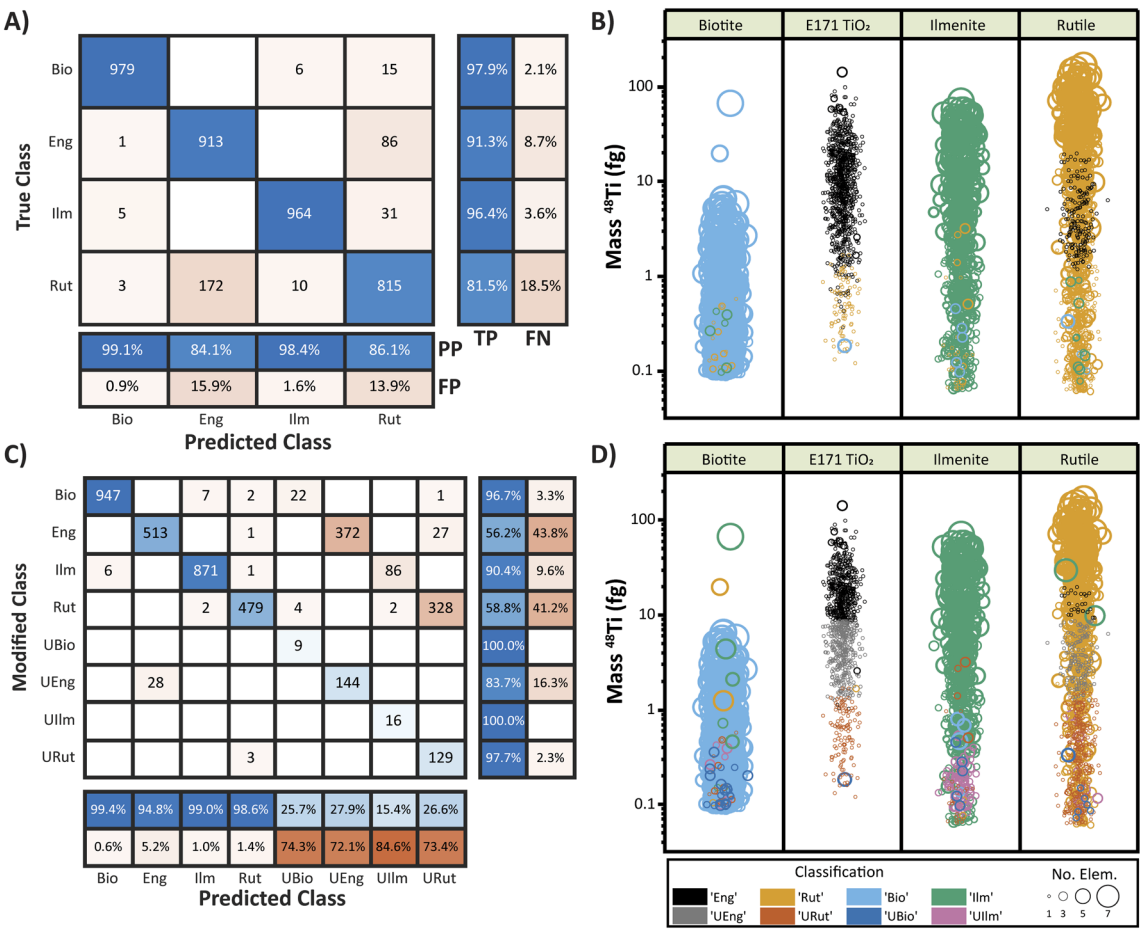

Fig. 2 Confusion matrices from the first (A) and second (C) SSML model. Bubble plots in (B) and (D) reflect the predicted classifications of individual particles. Bubbles are divided into four columns based on the true particle type and colored by the predicted classifications. The bubble sizes in (B) and (D) are proportional to the number of elements detected. Abbreviations: Bio: biotite, Eng: engineered, Ilm: ilmenite, Rut: rutile, UBio: unclassifiable Biotite, UEng: unclassifiable engineered, URut: unclassifiable rutile, UIlm: unclassifiable ilmenite, TP: true positive, FN: false negative, PP: positive prediction, FP: false prediction.

TP E171 classifications. However, lowering the FP percentage requires the model to be more selective, which results in a decrease in TP classifications and an increase in FNs. Therefore, the overall accuracy of the second-stage SSML model was reduced to 75.2 ± 2.8%. In Table S4,† we provide classification results for the unlabeled training data. Comparisons between the first and second SSML model's ROC and PR curves and figures of merit are shown in Fig. S2–S4.† ROC curves describe how well the ML model is able to separate the positive classifications from negative classifications; optimal models should have high TPRs and low FPRs, resulting in an AUC close to one. In contrast, PR curves illustrate how many predictions made by the model are truly correct and are best utilized when the number of positive predictions is low. As with ROC curves, AUC values of a PR curve that are closer to one describe ideally performing models. Metrics for the models described here are shown in Fig. S3.† The calculated FPR for the second SSML model decreases for the Eng and Rut classes but does not significantly change for the Bio and Ilm classes (see Fig. S3†).

In Fig. 2D, we show the classification results of the second SSML model for the labeled training data in the same format as Fig. 2B. As seen, particles that were incorrectly classified by the first SSML model are now classified as uncertain by the second model. The increase in FN classification can also be observed, as the model now classifies particle events with masses and elemental signatures similar to those categorized as UEng, URut, UBio, or UIlm as uncertain. While the creation of the uncertain classes worsens the overall model accuracy, the reduction in FPR enables classification across a broader number concentration range and in more varied particle backgrounds. Our two-stage SSML work flow results demonstrate that the approach is suitable for spICP-TOFMS analysis of anthropogenic and natural Ti-containing particles.

Testing and validation

Comparisons of the model metrics and classification performance of the first and second SSML with the labeled and unlabeled training data only provides a tailored look into the robustness of the workflow. Specifically, the similar particle numbers of each particle type represent an ideal case to achieve high accuracy *via* SSML classification. To further test the strength of our two-stage SSML workflow in predicting classifications with variable particle inputs, we applied our model to classify particles from the spICP-TOFMS analysis of a mixture of three natural particles types and of mixtures of anthropogenic and natural particle types.

In Fig. 3, we provide classification results for data from the spICP-TOFMS analysis of a mixture containing only the three natural particle types. In these data, we should not record any Eng particles because they were not in the sample; thus, any Eng classifications are deemed to be FPs. Likewise, any unclassifiable classifications are FNs. As seen, with just the first SSML model, 5.2% of the detected particles are incorrectly identified as Eng. By adding the second SSML model, FP Eng classifications decrease by a factor of 2, though 20% of particle events are recorded as 'unclassifiable'. Clearly, there is a trade-off between type I and type II errors (*i.e.* FPs and FNs)—reducing FPs results in an increase in FNs. For our application, we claim that it is more desirable to allow the model to produce FNs (*i.e.* to classify particles as uncertain) than to provide incorrect information to a user regarding the presence of engineered particles. While FPs are not entirely eliminated *via* the second-stage SSML training, we do see an improvement compared to the first model.

To test the classification performance of E171 particles in the presence of natural particles and *vice versa via* our two-stage SSML model, we applied the model to classify particle events in mixtures of engineered and natural particles at varying concentrations. Four dilution conditions were studied: (i) low and (ii) high natural Ti-containing particle backgrounds with E171 at different concentrations, as well as (iii) low and (iv) high E171 particle backgrounds with natural Ti-containing particles at different concentrations. In Fig. 4, we plot the number of particle events assigned to each class by our second SSML model *versus* the dilution amount of the analyte particles. In these results, the three natural particle-type classifications were summed together and presented, generically, as 'natural'; the

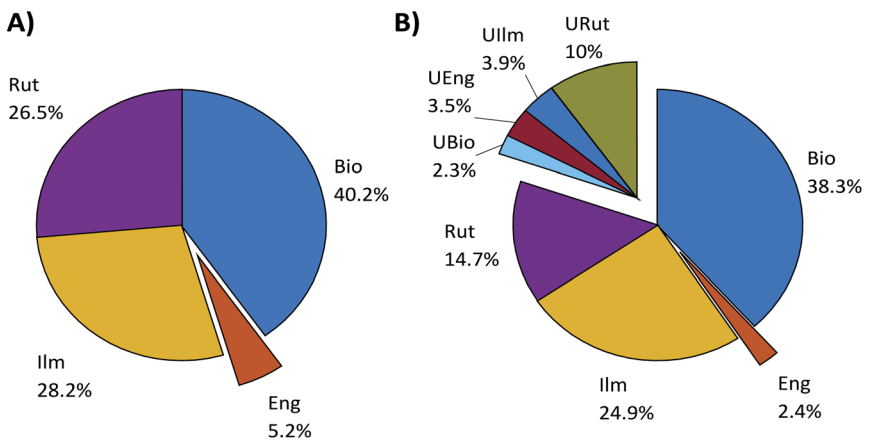

Fig. 3 Classification of particles in a mixture only containing the three natural particle types was performed with the first (A) and second (B) SSML models. The false-positive engineered classifications were reduced by a factor of two in the second SSML model.

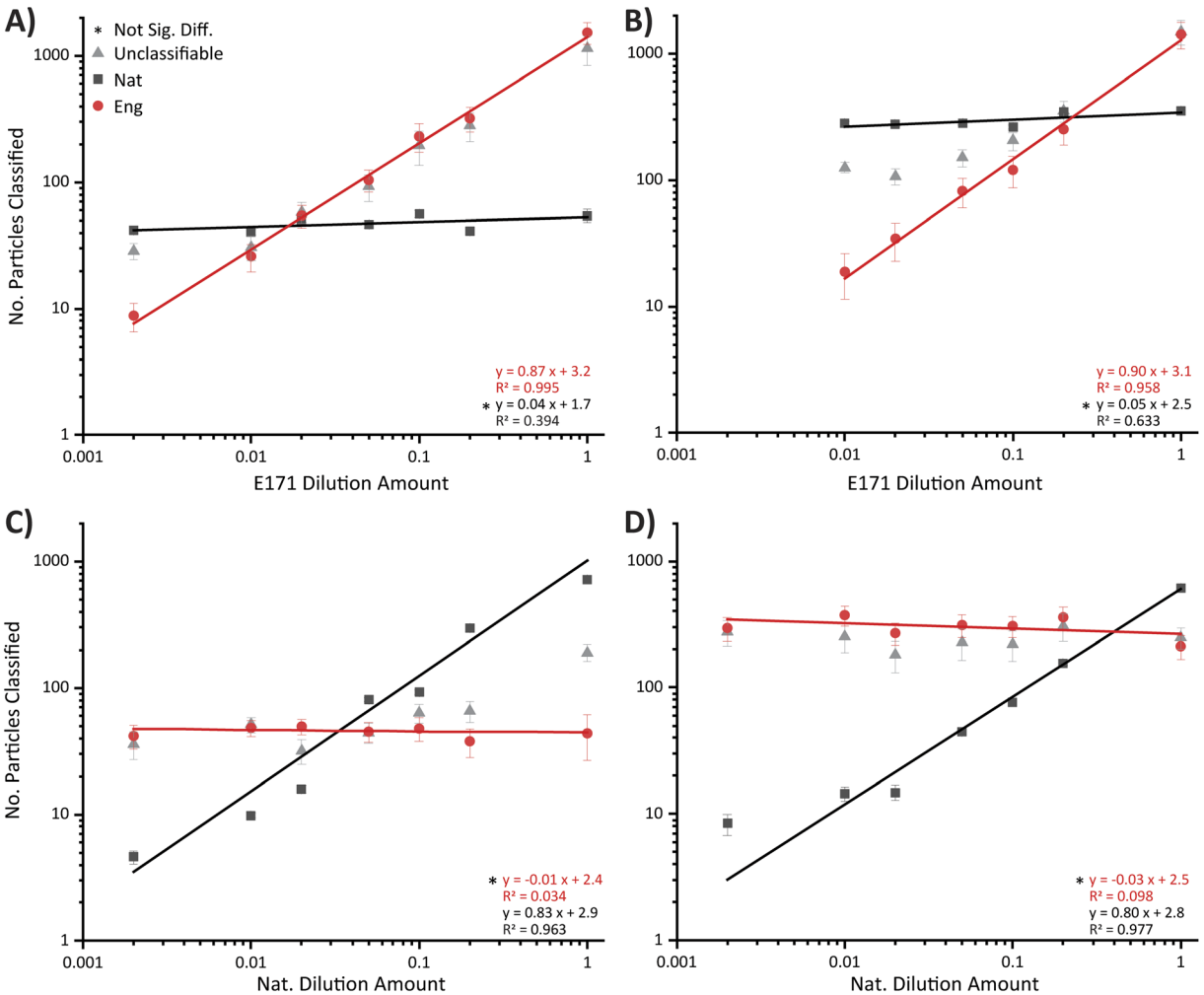

Fig. 4 Scatter plots comparing the classification performance of engineered (red circles), natural (black squares), and unclassifiable (grey triangles) particles with two-stage SSML model in the four dilution cases (i–iv in (A)–(D), respectively).

individual classifications are shown in Fig. S5.† For perfect recovery, the slope of the regression line should be equal to 1 when plotted on a log–log scale; deviations from this slope indicate incomplete or overestimated particle-type recoveries.

In Fig. 4, across the four dilution cases, the number of particle events identified as background remained constant with slopes not significantly different from zero; this is improved from the performance observed in the first stage of classification (Fig. S6†). Likewise, for the spiked Ti-particles (E171 or natural), the number of particle events identified as the analyte increased with increasing concentrations; however, none of the dilution cases demonstrated perfect classification. The deviations of the analyte regressions are likely caused by the high FNR observed in the second model, thus, slopes less than 1 are expected. It is possible that FPs, while generally estimated to be low in contribution, are dominant at low number concentrations and contribute to the overestimation in the cases in which natural particles are the analyte of interest. In general, recoveries between 80 and 90% were observed which indicates reasonable linearity for all studied conditions.

Conclusions

spICP-TOFMS measurements of particles are recorded with variable multi-elemental signatures and skewed mass distributions due to the inherent limitations imposed by ion counting (Poisson) noise, analytical sensitivity, and heterogeneity in the particle size distributions. These factors result in noisy particle event signals that make particle-type classification difficult at the single-particle level. For these reasons, we developed a multi-stage semi-supervised machine learning (SSML) workflow that accounts for uncertainty in the spICP-TOFMS measured signals. In this study, we demonstrate the usefulness and robustness of our workflow to classify titanium-containing particles. Natural Ti-particles and E171 are classified accurately in pristine samples. With two-stage SSML classification biotite and ilmenite are classified with <1% FP classification and FP classification of rutile and E171 is 1.4% and 5.2% respectively. Classification of Ti-particles in mixtures further validates the robustness of two-stage SSML for classifying particles in diverse conditions. Our approach emphasizes the

importance of low false-positive classifications for accurate particle type characterization. Additionally, we demonstrate that the workflow can be adopted for particle types with known, or suspected, reference materials. Overall, our multi-stage SMSL model for single-particle classification identifies and overcomes bias in spICP-TOFMS training data to provide a simple approach for the incorporation of machine learning models in particle analysis.

Abbreviations

spICP-TOFMS	Single-particle inductively coupled plasma time-of-flight mass spectrometry
NPs	Nanoparticles
μPs	Microparticles
SSML	Semi-supervised machine learning
TEM	Transmission electron microscopy
AF4	Asymmetric flow field-flow fractionation
tSNE	t-Stochastic neighbor embedding
$L_{C,sp}$	Single-particle element specific critical value
Rut	Rutile
Ilm	Ilmenite
Bio	Biotite
Eng	Engineered, E171 TiO_2
ROC curve	Receiver operating characteristic curve
PR curve	Precision-recall curve
TP	True positive
FP	False positive/prediction
TN	True negative
FN	False negative
PP	Positive prediction
FPR	False-positive rate
TPR	True-positive rate
fg	Femtogram
UEng	Unclassifiable engineered
UIlm	Unclassifiable ilmenite
URut	Unclassifiable rutile
UBio	Unclassifiable biotite

Data availability

Data for this article, and the code used to process it, can be found at **https://github.com/TOFMS-GG-Group/SSML_CeNPClassification/tree/TiNPClassification**.

Conflicts of interest

A. G.-G. was a researcher at Iowa State University at the time this work was conducted but is now an employee of TOFWERK.

Acknowledgements

The authors would like to acknowledge the funding sources for this work, including an Iowa State University Faculty Start-up Grant and NSF CAREER grant CHE-2237291. We would also like to thank Dr. Ed Raines, curator of the Mines Museum of Earth Science, as well as Dr. James Ranville and Dr. Aaron Goodman, Colorado School of Mines, for providing Ti-mineral samples.

References

1 A. A. Yaroshevsky, *Geochem. Int.*, 2006, **44**, 48–55.
2 U. Blaznik, S. Krušič, M. Hribar, A. Kušar, K. Žmitek and I. Pravst, *Foods*, 2021, **10**, 1910.
3 A. M. Khorasani, M. Goldberg, E. H. Doeven, G. Littlefair and J. Biomater, *Tissue Eng.*, 2015, **5**, 593–619.
4 O. U. Akakuru, Z. M. Iqbal and A. Wu, in *TiO2 Nanoparticles*, 2020, pp. 1–66.
5 A. Weir, P. Westerhoff, L. Fabricius, K. Hristovski and N. von Goetz, *Environ. Sci. Technol.*, 2012, **46**, 2242–2250.
6 R. Baan, K. Straif, Y. Grosse, B. Secretan, F. El Ghissassi and V. Cogliano, *Lancet Oncol.*, 2006, **7**, 295–296.
7 M. Younes, G. Aquilina, L. Castle, K.-H. Engel, P. Fowler, M. J. Frutos Fernandez, P. Fürst, U. Gundert-Remy, R. Gürtler, T. Husøy, M. Manco, W. Mennes, P. Moldeus, S. Passamonti, R. Shah, I. Waalkens-Berendsen, D. Wölfle, E. Corsini, F. Cubadda, D. De Groot, R. FitzGerald, S. Gunnare, A. C. Gutleb, J. Mast, A. Mortensen, A. Oomen, A. Piersma, V. Plichta, B. Ulbrich, H. Van Loveren, D. Benford, M. Bignami, C. Bolognesi, R. Crebelli, M. Dusinska, F. Marcon, E. Nielsen, J. Schlatter, C. Vleminckx, S. Barmaz, M. Carfí, C. Civitella, A. Giarola, A. M. Rincon, R. Serafimova, C. Smeraldi, J. Tarazona, A. Tard and M. Wright, *EFSA J.*, 2021, **19**, e06585.
8 H. Karkee and A. Gundlach-Graham, *Environ. Sci. Technol.*, 2023, **57**, 14058–14070.
9 J. Vidmar, T. Zuliani, R. Milačič and J. Ščančar, *Water*, 2022, **14**, 959.
10 M. M. Nabi, J. Wang, C. A. Journey, P. M. Bradley and M. Baalousha, *Chemosphere*, 2022, **297**, 134091.
11 M. M. Nabi, J. Wang, E. Goharian and M. Baalousha, *Sci. Total Environ.*, 2022, **807**, 151081.
12 G. Bucher, H. El Hadri, O. Asensio, F. Auger, J. Barrero and J.-P. Rosec, *Food Control*, 2024, **155**, 110102.
13 K. Loeschner, M. E. Johnson and A. R. Montoro Bustos, *Nanomaterials*, 2023, **13**, 2547.
14 H. Proquin, C. Rodríguez-Ibarra, C. G. J. Moonen, I. M. Urrutia Ortega, J. J. Briedé, T. M. de Kok, H. van Loveren and Y. I. Chirino, *Mutagenesis*, 2016, **32**, 139–149.
15 E. Verleysen, N. Waegeneers, F. Brassinne, S. De Vos, I. O. Jimenez, S. Mathioudaki and J. Mast, *Nanomaterials*, 2020, **10**, 592.
16 B. Li, S. L. Chua, D. Yu, S. H. Chan and A. Li, *J. Chromatogr. A.*, 2021, **1643**, 462059.
17 S. Candás-Zapico, D. J. Kutscher, M. Montes-Bayón and J. Bettmer, *Talanta*, 2018, **180**, 309–315.
18 A. Gondikas, F. von der Kammer, R. Kaegi, O. Borovinskaya, E. Neubauer, J. Navratilova, A. Praetorius, G. Cornelis and T. Hofmann, *Environ. Sci.: Nano*, 2018, **5**, 313–326.
19 K. Mehrabi, R. Kaegi, D. Günther and A. Gundlach-Graham, *Environ. Sci.: Nano*, 2021, **8**, 1211–1225.

20 M. Baalousha, J. Wang, M. Erfani and E. Goharian, *Sci. Total Environ.*, 2021, **792**, 148426.
21 M. Baalousha, J. Wang, M. M. Nabi, F. Loosli, R. Valenca, S. K. Mohanty, N. Afrooz, E. Cantando and N. Aich, *J. Hazard. Mater.*, 2020, **392**, 122335.
22 K. Phalyvong, Y. Sivry, H. Pauwels, A. Gélabert, M. Tharaud, G. Wille, X. Bourrat, J. F. Ranville and M. F. Benedetti, *Front. Environ. Sci.*, 2021, **8**, 549896.
23 F. Gottschalk, T. Sonderer, R. W. Scholz and B. Nowack, *Environ. Sci. Technol.*, 2009, **43**, 9216–9222.
24 S. E. Szakas, R. Lancaster, R. Kaegi and A. Gundlach-Graham, *Environ. Sci.: Nano*, 2022, **9**, 1627–1638.
25 H. Karkee, C. Kyte and A. Gundlach-Graham, *J. Anal. Atom. Spectrom.*, 2024, **39**, 1551–1559.
26 S. E. Szakas, K. Menking-Hoggatt, T. Trejos and A. Gundlach-Graham, *Appl. Spectrosc.*, 2023, **77**, 873–884.
27 A. J. Goodman, A. Gundlach-Graham, S. G. Bevers and J. F. Ranville, *Environ. Sci.: Nano*, 2022, **9**, 2638–2652.
28 M. Tharaud, L. Schlatt, P. Shaw and M. F. Benedetti, *J. Anal. Atom. Spectrom.*, 2022, **37**, 2042–2052.
29 T. R. Holbrook, D. Gallot-Duval, T. Reemtsma and S. Wagner, *J. Anal. At. Spectrom.*, 2021, **36**, 2684–2694.
30 G. D. Bland, M. Battifarano, A. E. Pradas del Real, G. Sarret and G. V. Lowry, *Environ. Sci. Technol.*, 2022, **56**, 2990–3001.
31 G. D. Bland, M. Battifarano, Q. Liu, X. Yang, D. Lu, G. Jiang and G. V. Lowry, *Environ. Sci. Technol. Lett.*, 2023, **10**, 1023–1029.
32 R. L. Buckman and A. Gundlach-Graham, *Anal. Methods*, 2024, **16**, 5802–5811.
33 R. L. Buckman and A. Gundlach-Graham, *J. Anal. At. Spectrom.*, 2023, **38**, 1244–1252.
34 S. Harycki and A. Gundlach-Graham, *J. Anal. At. Spectrom.*, 2023, **38**, 111–120.
35 S. Harycki and A. Gundlach-Graham, *Anal. Bioanal. Chem.*, 2022, **414**, 7543–7551.
36 A. Gundlach-Graham, S. Harycki, S. E. Szakas, T. L. Taylor, H. Karkee, R. L. Buckman, S. Mukta, R. Hu and W. Lee, *J. Anal. At. Spectrom.*, 2024, **39**, 704–711.
37 M. Kubat, R. C. Holte and S. Matwin, *Mach. Learn.*, 1998, **30**, 195–215.