Distinguishing Engineered TiO2 Nanomaterials from Natural Ti Nanomaterials in Soil Using spICP-TOFMS and Machine Learning

Garret Bland, Matthew Battifarano, Ana Elena Pradas del Real, Géraldine Sarret, Gregory Lowry

Distinguishing Engineered TiO_2 Nanomaterials from Natural Ti Nanomaterials in soil using spICP-TOFMS and Machine Learning

Garret D. Bland[†, §], Matthew Battifarano[†], Ana Elena Pradas del Real[‡,1] Géraldine Sarret[‡], and Gregory V. Lowry[†, §, *]

[†]Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

[‡]ISTerre (Institut des Sciences de la Terre), Univ. Grenoble Alpes, CNRS, 38000, Grenoble, France

[§]Center for Environmental Implications of NanoTechnology (CEINT), Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

[1]Present Address: Madrid Institute for Agricultural Research (IMIDRA), N-II km 38,200, 28800 Alcala de Henares

ABSTRACT:

Identifying engineered nanomaterials (ENMs) made from earth-abundant elements in soils is difficult because soil also contains natural nanomaterials (NNMs) containing similar elements. Here, machine learning models using elemental fingerprints and mass distributions of three TiO_2 ENMs and Ti-based NNMs recovered from three natural soils measured by single particle Inductively Coupled Plasma Time-Of-Flight Mass Spectrometry (spICP-TOFMS) was used to identify TiO_2 ENMs in soil. Synthesized TiO_2 ENMs were unassociated with other elements (>98%), while 40% of Ti-based ENM particles recovered from wastewater sludge had distinguishable elemental associations. All Ti-based NNMs extracted from soil had a similar chemical fingerprint despite the soils being from different regions, and > 60% of Ti-containing NNMs had no measurable associated elements. A machine learning model best distinguished

NNMs and ENMs when differences in Ti-mass distribution existed between them. A trained LR model could classify 100nm TiO_2 ENMs at concentrations of 150 mg kg^{-1} or greater. The presence of TiO_2 ENMs in soil could be confirmed using this approach for most ENM-soil combinations, but the absence of a unique chemical fingerprint in a large fraction of both TiO_2 ENMs and Ti-NNMs increases model uncertainty and hinders accurate quantification.

Synopsis:

We used elemental fingerprints and mass distribution of engineered TiO_2 nanoparticles and naturally occurring natural Ti-containing nanoparticles and machine learning models to effectively track and quantify Ti-based engineered nanomaterials (ENMs) in three different representative soils.

KEYWORDS:

spICP-TOFMS, engineered nanomaterials, soil, sewage sludge, machine learning, classification, TiO_2

INTRODUCTION:

ENMs are widely used in commercial and military applications (e.g. obscurants, armoring, and self-healing materials) due to their unique properties.[1] Once applied, ENMs are inevitably released into the environment which will continue to increase in the future.[2] For example, TiO_2 ENMs are a widely used nanomaterial that typically accumulates in municipal waste and sewers.[3,4] The released ENMs can potentially negatively impact ecosystem services such as food production and nutrient cycling[5] so accurate ENM risk assessments are needed. However, it is still a significant challenge to identify the source of ENMs or to quantify the amount of a selected ENM when present at environmentally relevant concentrations. This is especially true for ENMs made

from earth abundant elements like aluminum, silicon, iron, and titanium that are present in environmental media at >1 wt% .[6,7]

Distinguishing between ENMs and NNMs in environmental matrices like soil is difficult using conventional analytical instrumentation, especially for ENMs made from earth abundant elements. Electron Microscopy (TEM and SEM) can determine the morphology, crystallinity, elemental composition, and sizes of ENMs in soils. However, it takes considerable effort to detect ENMs via imaging without automation[8], and can be impossible if ENMs are present in low numbers. Inductively Coupled Plasma Mass Spectrometry (ICP-MS) measures the total metal concentrations in environmental or biological media after sample digestion. Many studies have used ICP-MS to track the fate of ENMs in soils[9–11] and plants[12–14]. However, this approach cannot distinguish between metals in ENMs and background concentrations of those metals (particulate or dissolved), especially at low ENM concentrations.[15,16] Filtration and centrifugation can be used to separate particulate and dissolved fractions, but the size cutoff is operationally-defined, and it cannot fully distinguish dissolved and particulate measurements.[17] Changes in the ratios of the element of interest versus common earth elements (e.g. Ti and Al) relative to background ratios have been used to indicate the presence of ENMs, but this cannot determine the specific source of ENMs.[4,8,18] Size separation techniques, such as field-flow fractionation (FFF), asymmetric flow field-flow fractionation (A4F),[19–21] size exclusion chromatography (SEC),[22] and hydrodynamic chromatography (HDC)[11] have been used to determine bulk metal concentration as a function of particle size, but this does not determine single particle compositions. Single-particle Inductively Coupled Mass Spectrometry (spICP-MS) can measure the mass of metal-based single nanoparticles one element at a time,[23,24] but many ENMs are made from earth-abundant elements (e.g. Si, Al, Ce, Cu, Ti, and Fe)[25] making it difficult to distinguish them from the natural

nanomaterials (NNMs) in the sample that contain similar elements. Determining Ti-based ENM concentrations is especially challenging since natural Ti is abundant in soils ranging from 0.2 to 2.4% worldwide.[26] Measuring multiple elements in each particle simultaneously may identify differences in the chemical fingerprints between ENMs and NNMs.

Single-particle ICP Time-of-Flight Mass Spectrometry (spICP-TOFMS) can measure most elements in the periodic table simultaneously on a single particle. This provides a rich data set consisting of multi-elemental composition of thousands to tens of thousands of individual nanoparticles in a sample in just a few minutes,[27] and can provide a specific elemental fingerprint and mass distribution for each particle. For example, Ce ENMs can be distinguished from Ce-containing NNMs using a natural tracer (La) that is present in a fraction of Ce-containing NNMs but not in Ce ENMs.[28] However, there are detection limits for each element measured by spICP-TOFMS. This currently limits the lower bound of particle detection to 20 to 30 nm depending on the element, and limits particle fingerprint information in both ENMs and NNMs. Many elements, including Ta, Nb, Zr, and rare earth elements are often associated with natural sources of Ti and may be a "fingerprint" for Ti-containing NNMs[29,30] that is distinguishable from TiO_2 ENMs which are relatively pure. Gondikas et al., specifically measured the fingerprint of Ti-containing NMs in surface water with spICP-TOFMS found that Al, Fe, Mn, and Pb are present in Ti-containing NMs.[8] Baalousha et al., determined the elemental fingerprint of Ti-containing NNMs in three soils using spICP-TOFMS.[31] While these "fingerprints" could be used to differentiate between NNMs and ENMs, analytical strategies are needed to statistically assess elemental fingerprint information for thousands of nanoparticles in a sample and determine which, if any, can be used to classify them as ENMs or NNMs.

Machine learning (ML) classification models learn from complex datasets to identify important features and formulate classification criterion without domain knowledge. The most critical aspect of ML models is the consideration of *all* information within the dataset with many dimensions of the features that can capture important trends that may not be determined by simpler regressions or associations. There are two major categories of ML algorithms: unsupervised and supervised learning.[32] Unsupervised learning discovers hidden patterns in unlabeled datasets that forms correlations between features that considers all available information. Supervised learning trains on labeled datasets and predicts labels on unseen data, e.g., classification. The performance of a supervised learning model is measured by predicting on unseen labeled test datasets. ML is increasingly being applied in environmental science with a few studies focusing on ENMs. One study implemented a gradient decision tree with boosting classification (GBC) and recursive feature elimination (RFECV) to distinguish natural and engineered Ce-containing nanoparticles.[28] Another study used linear discriminant analysis (LDA) to classify isotopic fingerprinting of SiO_2 NPs.[33] Both models handled only a limited number of features (i.e. elements or isotopes). For the GBC model, they limited the model to 25 features (and reduced this to 17 with RFECV) and the LDA model used two features at a time. While these models could classify ENMs and NNMs for these particles, the ability to classify TiO_2 ENMs and other ENMs made from earth abundant elements in soil may require different approaches. For example, features other than chemical fingerprint such as mass distribution of particles may also be important for accurate classification.

The objective of this study is to determine the properties of Ti-based NNMs and ENMs that can be used to identify a specific Ti-based ENM in soils with background Ti-based NNMs. We first investigate how the mass (size) and elemental fingerprints measured by spICP-TOFMS of Ti-based NNMs vary depending on soil type and how these potentially affect classification. Machine-

learning models are developed to distinguish two synthesized Ti ENMs (30 nm and 100nm) and Ti-ENMs recovered from wastewater treatment plant sludge from background Ti-based NNMs present in three types of soil. The model consists of a dimension reduction method (non-negative matrix factorization, NMF) and a linear classification model (Logistic Regression, LR) with k-fold validation. To determine the particle concentration detection limit of Ti-ENMs in soil, the trained model is then applied to Ti-particles extracted from soils dosed with varying concentration of TiO_2 ENMs. The model uses both Ti mass distribution and elemental fingerprinting for classification because a large portion of Ti-based NNMs do not have elemental associations. The model performed the best with test cases that have significant Ti-mass difference between ENMs and NNMs. While machine learning models are specific to the data they are trained on, our approach can confirm the presence of a specific TiO_2 ENM in a soil with a measured fingerprint of Ti-NNMs, but accurate quantification is not yet possible for TiO_2 ENMs with similar size distributions as the Ti-based NNMs.

METHODS

Experimental Design

The elemental fingerprints and mass distributions of Ti-containing NNMs extracted from three different soils and three Ti-based NMs from different sources (two commercially available synthesized TiO_2 NPs and Ti-based NMs extracted from sewage sludge) were used to develop machine learning models (3x3 matrix, Table S2) to classify between natural and engineered Ti-based NMs in soil. While the two synthesized TiO_2 NPs are pristine "as manufactured" ENMs, the Ti-based NMs extracted from wastewater sludge is more likely a unique mixture of ENMs and NNMs. Therefore, we label the sludge as its own category, "Ti-sludge". We refer to Ti-sludge as an "ENM" to distinguish it from NNMs recovered from soils. The mass distributions and elemental

fingerprints were measured by spICP-TOFMS for each NM. To do this, the extracts from soil were diluted with DI water, and ENMs were suspended and diluted in DI water. For the discussion of this study, first the NNMs extracted from one soil is paired with the three different ENMs to determine how ENM particle type affects the ability to classify them in a particular soil. Then, one ENM type is paired with three different NNMs to understand if different natural backgrounds can influence the classification of a particular ENM type. Therefore, five of the nine cases are explained in detail in the results and discussion section while the remaining cases are explained in SI (Figure S10). The limit of TiO_2 ENM detection was also approximated. This was done by adding different known amounts of a TiO_2 ENM to a soil, extracting all NMs (<500 nm size fraction) from soil using an established method, determining their mass distribution and elemental fingerprints by spICP-TOFMS, and using the trained ML model to quantify the number of TiO_2 ENMs extracted from the soil.

Materials and Samples

Polyvinylpyrrolidone (PVP) 70 kg mol^{-1}, triethanolamine (TEA), diethylenetriamine pentaacetate (DTPA), sodium chloride, and Triton X-114 were purchased from Sigma Adlrich. Carboxymethyl cellulose 700 kg mol^{-1} was purchased from Acros Organics. All chemicals have a purity of $\geq$ 99%.

Three types of Ti-based ENMs and three NNMs were used in this study. TiO_2 anatase nanoparticles of a 30nm nominal size were purchased from US Research Nanomaterials (Ti30). TiO_2 (rutile) nanoparticles with an average size of 100nm were provided by the US Army Research Office (Ti100). Both ENMs were received as dry powders. The mineral phases for Ti100 were confirmed by XRD (Figure S3). The mineral phase for Ti30 was determined from XRD and was provided by the manufacturer. For the ENMs (Ti100 and Ti30), the particles were characterized

by TEM, DLS, and spICP-TOFMS (Table S1) in DI water. The suspensions were bath sonicated for 30 minutes in an ice bath prior to characterization. The particle size and morphology were determined by TEM using a JEOL JEM-2000EX and ImageJ software. The hydrodynamic diameter of the ENMs was measured by a Malvern Zetasizer Nano ZS at 10 mg l^{-1} in 5 mM NaCl with a pH of 5.6. The sewage sludge containing TiO_2 ENMs (Ti-sludge) was produced in a wastewater treatment plant at Eawag (Zurich, Switzerland). Further details of the sludge are provided in previous studies.[34,35] ENMs could be introduced into terrestrial ecosystems through the application of treated wastewater sludge so they represent an environmentally relevant input of ENMs entering soils.[36] It is important to note that the Ti-based NMs in the sludge can be from both natural and anthropogenic sources, but in TEM-EDX analysis, TiO_2 ENMs were identified.[35]

NNMs were extracted from three types of soils: a loamy sandy soil (L22) (Lufa 2.2, Germany), a loamy soil (LUV) (Luvisol, WRB, 2006) collected from La Côte Saint-André (Isère, France), and a calcareous clayey loam soil (ARZ) collected from a plot at the Marcopa Agricultural Center (33°04′ 22″ N, 111° 58′ 26.5″ W) in Arizona, USA.[37,38] These soils were selected to provide a range of geographical locations and soil properties (Table S2). LUV soil has the highest organic carbon content (2.23%) followed by L22 and ARZ soil (1.71% and 0.54% respectively). L22 is the most acidic soil (pH = 5.8) with LUV soil being circumneutral (pH = 6.4) and ARZ soil being alkaline (pH = 7.6). Lastly, ARZ soil has the highest clay content compared to L22 and LUV soil. We hypothesized that with different soil properties and source formations there would be distinct Ti chemical fingerprint in each NNMs.

For spICP-TOFMS analysis, dissolved multi elemental standards (10 ppm) were purchased from Inorganic Ventures. Tuning solution from Thermofisher containing 1 ppb of dissolved Co, In, Ce, and U was used to calibrate the optics of the instrument and TOF detector. For the Au

nanoparticle standard, citrate capped Au nanoparticles (particle size = 50nm) were purchased from Sigma Aldrich (particle size distribution by TEM = 49.1 ± 6.51 nm, number of particles = 193).

ENM Dosing Procedure

To approximate the detection limit of the method, Lufa 2.2 soil was dosed with Ti100 ENMs at nominal concentrations of 70, 150, 300, 700, and 7,000 mg-TiO_2 kg^{-1} using previously published methods to provide uniform mixing which is briefly described later.[39] The detection limit in this instance is defined as the minimum ENM concentration in soil the model can confidently predict and be statistically different from undosed soil. These dosed concentrations are higher than what is predicted for soil with current material flow models, but the goal here is to estimate the minimum ENM concentration needed for detection using our sample preparation and ML model.[40] While the Ti100 ENM suspension was bath sonicated, an aliquot was pipetted and evenly distributed in the soil to achieve field capacity moisture content (22%). The sample was thoroughly mixed by wooden sticks for at least 15 minutes and transferred to a PTFE centrifuge tube, loosely capped, and stored in dark for 24 hours to allow interaction with the matrix. The samples were then subjected to the NM extraction procedure (details below). It is worth noting that the 24-hour incubation time is relatively short and may not include long-term aging processes that can affect extraction efficiency, e.g., heteroaggregation. However, the objective of this study is to determine how well the ML model predicts ENM concentrations in impacted soils, and aging effects were not considered in this study.

Soil NM Extraction Procedure

The <500 nm particles were extracted from soil and sludge using a previously described method.[37] Briefly, the sediment was preconditioned with a NaCl solution at a 1:10 solid to liquid

ratio. The extraction solution is then added to the mixture containing carboxymethyl cellulose to enhance the stabilization of nanoparticles against aggregation. This soil/solution mixture is centrifuged at a particle size cut-off of 500 nm for soil assuming a bulk soil density of 2.65 g cm^{-3} (300 g for 5 mins using a JS 5.2 Swingbucket rotor). The supernatant containing the nanoparticles goes through a cloud point extraction (CPE) enrichment step. A surfactant, Triton X-114, and NaCl were added to the extract to achieve a final concentration of 0.2% and 10 mM, respectively. The mixture was then heated at 40°C for 1 hour to form micelles. This was centrifuged to create a 2-phase system: a surfactant-enriched phase containing the nanoparticles and water phase containing the ions. The enriched surfactant phase containing nanoparticles was diluted and washed several times by DI water and centrifugation. These samples were then stored in the dark at 4°C until analysis.

spICP-TOFMS

The spICP-TOFMS 1R by TOFWERK (Zurich, Switzerland) is used for this study. For ICP-TOFMS analysis, <500 nm extracted suspensions were diluted with DI water to 10^5 to 10^6 particles $ml^{-1}$for analysis. The samples were then bath sonicated for 8 minutes prior to measurement. The instrument was calibrated from dissolved metal standards (Inorganic Ventures) with a range from 10 ppt to 5 ppb in 1% trace-metal grade HNO_3. The spICP-TOFMS tracks multiple elements simultaneously during acquisition. Table S4 provides a list of optic and TOF detector parameters. The spICP-TOFMS detector system is a microchannel plate detector (MCP) converts the ion momentum into measurable signal. The TOF extraction frequency (extractions per s) for all isotopes is 33 kHz. Then these extractions were averaged over an integration time of 2 ms to provide each data point. Shorter dwell times were tested but we did not observe any significant difference in the single particle dataset. We also did not observe split-peak particles

after the correction (explained later). Particle coincidence could occur resulting in false positives of multi-elemental particles. We predicted that a small fraction (less than 10%) of the multielement particles could be false positive. Further discussion is provided in SI. The transport efficiency was calculated using the size based method with a 50 nm Au nanoparticle suspension (Sigma Aldrich) as previously described.[41] Instead of the peak height, the peak area was used to measure Au-based signal of the individual particle events.

Single particle processing was achieved using Python script provided by TOFWERK. This process allowed for the detection of single particles from baseline signal. For each sample acquisition, the mean and standard deviation signal of the sample was calculated for every isotope in 100 integration time point intervals. Since each sample acquisition is made up of 60,020 data points, the identification of single particle events is compartmentalized in discrete 100-time intervals where the particle events are compared to localized background signal in the same interval. The threshold for single particle detection follows the International Union of Pure and Applied Chemistry (IUPAC) guideline to describe low intensity noise (threshold = $\mu + 3.29\sigma + 2.72$, μ and σ is the average and standard deviation of signal intensity, respectively).[42] Any integration time point above this threshold is considered a single particle event. The detection limit varies because the calculated threshold compared to background signal is different for each 100-integration time point interval. Detection limits of single particles for every isotope can be estimated by calculating the average threshold from background signal in each sample. For ^{48}Ti, the average detection limit size of a TiO_2 particle (assuming a spherical shape and rutile density of 4.23 g cm^{-3}) for all samples measure is approximately 36 nm which is on par or better than previous spICP-MS studies.[4,8,43,44] This is mainly due to the high dilution factor in DI water and low background signal (0.01 to 0.02 µg L^{-1}). Once the single particles are identified, the baseline signal

average and standard deviation are recalculated. If more single particles are detected, then this process is iterated up to 10 times before proceeding to the next window of 100 integration time points. This is an iterative approach to ensure no outliers are found in background signal. Split-peak particle events occur when a single particle event corresponds to multiple integration time points. These signals are added together and are reported as a single particle event (maximum of 3 consecutive particle signals). Signal loss does not occur during split-peak particle events.

Machine Learning Model

Binomial Logistic Regression (LR) was used to classify ENMs and NNMs. The pipeline of this model was written in Python Sci-Kit Learn Module. Other classification models were also evaluated and compared to the LR results. This pipeline consists of three components: pretreatment of the data set, training the classification model, and testing (Figure S1). The data set is comprised of masses and elemental fingerprints from ENM and NNM samples measured by spICP-TOFMS separately. Data from each of the six NMs analyzed are then parsed in Python to only include Ti-particle events. For each ENM-NNM test case that is discussed in the "Experimental Design" section and in Table S2, the single particle data from an ENM is randomly mixed with equal portions of single particle data from an NNM (50% ENM and 50% NNM) to provide a balanced dataset. Twenty percent of the mixed data is withheld as a test data set and remaining 80% is used for training. Training datasets undergo a data transformation method called non-negative matrix factorization (NMF). Similar to principal component analysis, the purpose of NMF is to reduce the dimensionality of the given features (i.e., analytes) to 10 components while preserving all elemental information. This dimensionality reduction technique is well-suited for single-particle mass data because it requires all values to be non-negative. We also have varied the number of components between 5, 10, 15, and 20 and found that 10 or higher provided the best performance

(Details in SI). In addition to the NMF components, a ^{48}Ti component was implemented, so the model can consider Ti mass separately from other elements during classification. This enables the model to use mass distribution and elemental fingerprint differences simultaneously. The LR model is then trained on the dataset using five-fold cross validation and the limited-memory Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm as the solver.[45] To train, the LR model optimizes the weight to each of the eleven components for the best classification performance. The weights of the model components are then updated, and it reiterates this process until the algorithm converges with a 0.001 tolerance stopping criteria. The trained model is then evaluated with the test data set. Each test data point is given a probability of being ENM or NNM. This process is done for every scenario between the three NNMs (Lufa 2.2, Luvisol, and Arizona) and the three ENMs (Ti-sludge, Ti100, and Ti30), in total producing nine different trained models and predicted scores. The particles are then assigned to three categories: classified (≥85% predicted probability in the correct category), uncertain (between 15% and 85%), and incorrectly classified (< 15%). Table S2 displays which figure each test case results are in (Figure 3, 5, and S10). The 85% and 15% confidence level are *operationally defined* to categorize the test dataset and to compare all cases. The importance of each element was calculated for each case and the top isotopes are listed in SI (Figure S14, S15, and S16). The trained L22 + Ti100 model was then used to predict the classification of extracted Ti-containing NMs from L22 soil dosed with 70, 150, 300, 700, and 7,000 mg kg^{-1} of Ti100 ENMs. The model predicts the percentage of ENM particles in the dosed soil. Python script for the spICP-TOFMS ML model, fingerprinting analysis, detection limits, including figures are available in Github (github.com/gbland0725/sphandles/) and functionality can be imported as a package via TestPypi (test.pypi.org/project/sphandles/).

Total Ti concentration

We determined the Ti concentration (mg kg^{-1} soil) in L22 soil without ENM addition and after being dosed at 70, 700, and 7,000 mg kg^{-1} of Ti100 ENMs. We also measure the mass of extracted Ti to calculate the extraction efficiency. The soil or the extracted NMs were oven dried (60°C) for 48 hours and ground by mortar and pestle before analysis. Total Ti concentration was determined by digestion followed by ICP-OES measurement as previously described.[46] Briefly, 20 to 30 mg of ground material is transferred to a 20 mL Teflon tube for chemical digestion. The following mixture is added for digestion: 5 ml 70% HNO_3, 0.2 ml HF (48%) and 0.5 mL H_2O_2 (30%). The sample is then microwaved in an MLS ultraclave. A reference soil (Canadian Reference Soil SO-1) with a Ti content of 0.53% is also used to validate Ti concentrations which are in a similar range to the other soils (0.32%, 0.27%, and 0.17% for ARZ, LUV, and L22 soil, respectively). Ti recoveries of the reference soil was 89.9% ± 2.3% (n = 2).

Results and Discussion

Characterization: Elemental Fingerprints and mass distributions of the Ti-particles

We determined the fraction of Ti particles in each sample with elemental associations and the most common elements associated with Ti-containing particles for each NNM and ENM (Figure 1). Elemental associations are defined as single element particle events identified within a single dwell time point after split peak correction. For the NNMs in soil (LUV, L22, and ARZ), between 20% to 35% of the Ti-particles were associated with other earth-abundant elements (e.g., Ce, Ba, Rb, Fe, Mg, Mn, Nb, and Pb). Often the Ti particles were associated with several different elements. Interestingly, there were only minor differences between particles recovered from each soil, despite that the soils were sourced from different regions and have different properties (Table S2). We hypothesized that NNM in the L22 soil would have more differences because the acidic pH and higher organic matter content could mobilize Ti into other soil particles. However, the

three soils used here do not likely represent all soil types. Importantly, there was a significant fraction of Ti particles that were not associated with any other elements (80, 81, and 65% for ARZ, L22, and LUV respectively). This is likely because the most common forms of titanium minerals (TiO_2, rutile and anatase) are resistant to weathering and have low solubilities.[47–49] Isomorphic substitution and incorporation of other elements into Ti-NMs are minimal and the morphology is well preserved. Adsorption of another element (e.g., Fe) on the surface of the Ti-NMs is possible. However, even if all the adsorption sites were saturated (one atom thick layer around a 100nm TiO_2 sphere), the mass of the adsorbed element is not sufficient to be measured as a particle event by the spICP-TOFMS because it would be below the detection limits (10^{-16} to 10^{-17} grams for most elements). Previous studies have also reported the presence of <100 nm TiO_2 NNMs in soils and sludge.[35,50] Ti particles may have trace elements associated but given the detection limits of the spICP-TOFMS, there may not be enough mass of the trace element to be recognized as a single particle event.

For both pure synthesized ENMs (Ti100 and Ti30), there are negligible amounts of other elements in the particles, as expected for an engineered material (> 98% of the particles were only Ti). For the Ti-sludge (presumed to be primarily ENMs entering the treatment plant)[35], a large fraction (65%) of Ti-particles are associated with other elements like Ag, Fe, and Pb. This is likely because the Ti-containing particles are heteroaggregated with organic particles containing other metals. The organic carbon content of the colloids from the sludge is 36.9 ± 0.6%, and SEM images from a previous study showed the TiO_2 particles embedded into organo-mineral heteroaggregates.[35] A few elements were found in both NNMs and Ti-sludge particles, including Fe, Mg, and Ba. However, there were also unique elements associated only with Ti-sludge particles, including Cu, Ag, and Pb. The Ag attachment is high because this sludge was also

deliberately dosed with AgNPs.[34] Similar to the soils, there are also a substantial fraction of Ti particles (43% of the particles) that are not associated with other elements. Therefore, formulating a classification criterion based strictly on elemental fingerprint of Ti-containing particles will not be robust because of the significant number of Ti-particles without associated metals, and in the case of the Ti-sludge, because they have similar elemental associations as the soils.

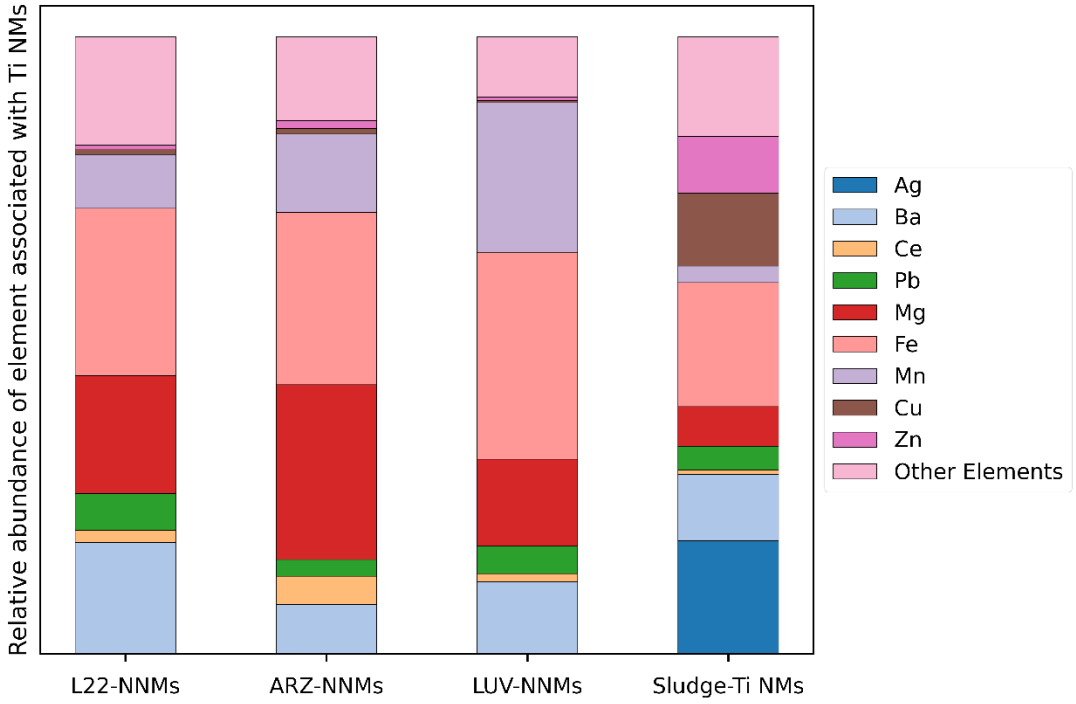


Figure 1: The fingerprints of multi-element Ti NNMs in L22, ARZ, LUV, and recovered Ti-NMs from Sludge. The major isotope except for Fe (^{54}Fe) is used to represent each element. Note that Ti30 and Ti100 are not shown because they had no element associations (pure Ti). Total values were calculated by the percentage of total Ti particles associated with each element. The values were then normalized to one to compare between samples. $\left(\frac{\%\ of\ element\ associated\ with\ Ti-NMs}{\sum_{k=1}^{10} \%\ of\ element\ associated\ with\ Ti-NMs}\right)$

The Ti particle mass distribution differences was also used for classification. The mass distribution of Ti-particle events was determined by spICP-TOFMS for each of the samples

(Figure S9). All Ti-based NNMs have a similar mass distribution, with medians of 1.78, 1.37, and 1.40 x 10^{-16} grams for ARZ, L22, and LUV, respectively. This correlated to sizes with mean and standard deviation of 49.9 ± 22.5 nm, 49.3 ± 30.6 nm, and 48.8 ± 26.2 nm, respectively, assuming spherical rutile particles. The similar median particle masses are possibly because most of the identified Ti-particles are near the average IUPAC threshold, so smaller particles in the distribution are not detected. The range of masses for the sludge Ti-particles is larger, from 10^{-16} to 10^{-13} grams. For all samples, the range of mass distribution for Ti-particles without other associated metals is similar to the overall Ti-mass distribution, indicating that the types of impurities associated with the Ti particles is not dependent on Ti mass. The ENMs, Ti100 and Ti30, had narrow mass distributions, and were slightly higher than the size range determined by TEM (Table S1).

ML Model Selection and Comparison

Different types of machine learning models were investigated to determine the most appropriate for classification, including Gradient Boosting Classification (GBC), Random Forest Classifier (RFC), and Neural Networks with either a 'relu' or 'logistic' activation layer (NN1 and NN2, respectively). Performance is based upon training data from Ti100 and L22 case. The average five-fold cross validation accuracy was used to compare each model and classification was operationally assigned as >85%= correct, 15% to 85% = uncertain, and <15% = incorrect. One major criterion for model selection is high precision of the ENM category: i.e., maximizing the amount of ENMs predicted correctly while minimizing the amount of NNMs predicted incorrectly as ENMs (false positive). All models exhibited similar trends as the LR model. The major difference is the sensitivity of the mass discrimination difference between ENMs and NNMs. GBC and RFC models had similar performance putting 80.6% of particles in the correct category (i.e., above 85% confidence), about 2 times greater than the LR model (40%). However,

379 the percentage of incorrectly classified particles (4.6%) is also about 5 times higher than for the
380 LR model that has only 0.9% incorrect. NN1 also had a similar performance with 79.6% correct
381 and 4.2% incorrect. NN2 was similar to the LR performance at 39.1% correct and 0.9% incorrect
382 which is not surprising given the activation layer is the same mechanism (logistic function). The
383 greater number of false positives (i.e., NNM predicted as ENM) for the GBC and RFC models was
384 because they classified pure Ti particle events stochastically. This is evident from larger Ti100
385 ENMs classified below the 85% confidence level. Note, if the operationally defined classification
386 boundary was shifted to a 75% confidence level, the difference in performance between the models
387 was considerably smaller (Table S3). The NN models are similar to LR but can describe
388 nonlinearities within components which may produce overfitting and lose generalization of the
389 data. The highest precision of the ENM category was NN2 and LR. We conclude that the
390 performance of the models is primarily based on how the model weighted the Ti mass-
391 discrimination for the pure Ti-NNMs and decided to choose LR to minimize the number of
392 incorrect particle events and to determine the important features for each ENM + NNM case. All
393 models performed worse without including the mass distribution.

394 Table 1: The classification performance results of each ML model for the Ti100 + L22 Case.
395 (>85% = correct, 15% to 85% = uncertain, <15% = incorrect). Number of simulations = 9. Average
396 percentage values are reported with standard deviation in parentheses.

Model Type	*Certain (> 85%)*	*Uncertain (15% to 85%)*	*Incorrect (< 15%)*
RFC	**80.7** (1.3)	**14.9** (1.6)	**4.5** (0.5)
GBC	**81.5** (1.2)	**14.5** (1.5)	**4.1** (0.5)
NN1	**82.3** (0.6)	**13.3** (0.7)	**4.4** (0.3)
NN2	**38.3** (1.6)	**60.7** (1.6)	**0.9** (0.1)
LR	**41.3** (1.8)	**57.7** (1.9)	**1.0** (0.2)

397

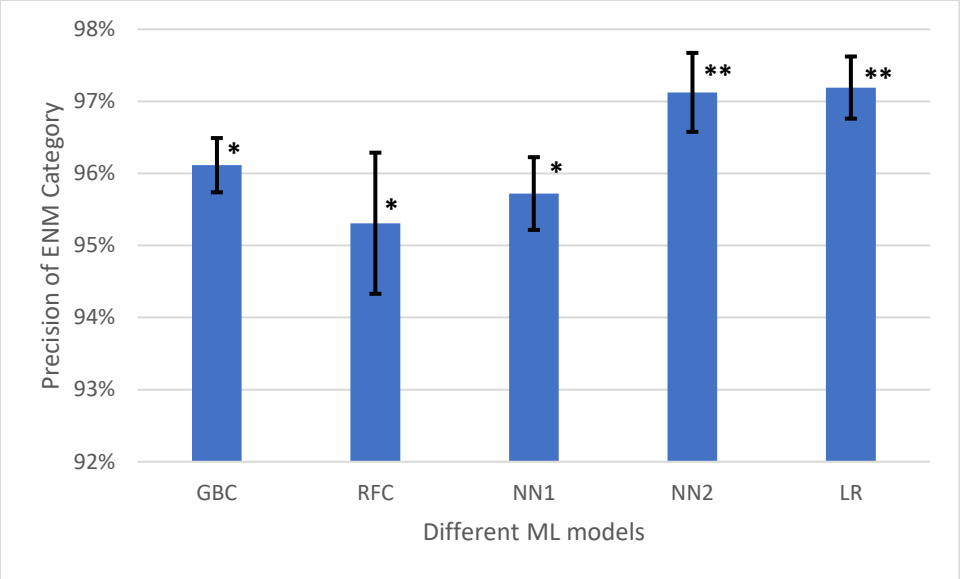


Figure 2: The precision of the ENM category for each ML model (number of simulations = 10) with a confidence interval of 95%. Symbols represent statistical similarity (One-way ANOVA, similarities contain a p-value $>10^{-2}$).

ML Models for distinguishing between Ti-based NNMs and ENMs

Influence of ENM type on model performance

The ML model was trained to distinguish between ENMs and NNMs. For this discussion, we first focus on the ability to distinguish each ENM (Ti30, Ti100, and Ti-sludge) from the Ti-based particles recovered from the L22 soil. Figure 3 displays the probability of correctly predicting the category of each Ti particle. Figure 3A displays specifically for the Ti100 + L22 case, 3B shows the Ti30 + L22 case, and 3C shows the Ti-sludge + L22. For each test case, the left scatter plot *depicts only the Ti-NNMs* and their probability of being predicted as 'Natural' by the model as a function of Ti-mass. The right scatter plot displays *only Ti-ENMs* and their probability of being predicted as 'Engineered'. The filled circle markers are Ti-NMs that are associated with at least one other element, while the open circle markers are Ti-NMs that are

unassociated with other elements (only ^{48}Ti or ^{48}Ti and ^{46}Ti). The pie chart on the right shows the percentage of the total particles having a probability > 85% in the correct category as green, uncertain (< 85% and >15 %) as orange, or incorrectly classified (<15%) as red. The elemental fingerprint for each test case was evaluated for comparison between correctly classified (> 85%) and the uncertain/incorrectly classified particles (Figure S11). We use both ^{46}Ti and ^{48}Ti for this analysis. Because ^{48}Ti is the main isotope, we refer this as the total Ti and report ^{46}Ti separately.

The model performed best for the Ti100 + L22 (40.3% correctly classified). To better understand the reasons for the correctly classified particle population, we separated the L22 NNMs and Ti100 ENMs and plotted each as a function of Ti-mass (Figure 3a). For the L22 Ti-NNMs (left scatter plot), two trends are prominent (a and b). One is a population of particles that creates a sigmoidal curve below the 85% confidence level. The sigmoidal curve of Ti-NNMs below the 85% confidence level (orange markers) does not contain any elemental fingerprint information and the model exclusively compares the Ti-mass to distinguish the NNMs from ENMs. Therefore, the certainty of assignment as a Ti-NNM decreases as the Ti-mass increases (b). The second is the population of particles that lie above the 85% confidence level, many at the 100% level, that are independent of particle mass (a). These particles all have elemental fingerprints (e.g., Ba, Mg, V, Nb) that are not found in the ENMs, and therefore are correctly assigned as an NNM. The elemental fingerprint of the L22 Ti-NNMs are plotted in Figure S11A.

The right scatterplot contains only the Ti100 ENMs. The ENMs have no chemical fingerprint so they lie on the sigmoidal curve. As the mass increased, so does the confidence that they are correctly assigned as an ENM(c). Because of the mass difference between L22 and Ti100 (Figure 4), the model can correctly predict 68% of the Ti100 ENMs >85 % confidence, making Ti100 detectable in soil.

For the Ti30 + L22, the model does a poor job in its assignments (6.2% correctly classified with a probability >85%). This is because the Ti-mass distribution for the ENMs and the NNMs are overlapping (Figure 4). Therefore, the model uses only the L22 NNMs particles' elemental fingerprint for classification. Figure 3b indicates that only Ti-containing NNMs that associate with other elements are being correctly classified as such with confidence >85%, while the Ti30 ENMs are indistinguishable. With the majority (80%) of L22 Ti-NNMs not having element associations, and having a similar size to Ti30 ENMs, predicting any Ti30 ENM contribution in this soil is not feasible.

For the Ti-sludge + L22 case (Figure 3c), the model performed adequately with 35% of the Ti-sludge NMs being correctly classified with >85% confidence. Unlike the cases with "as manufactured" ENMs, the NNMs in L22 soil were mostly classified as uncertain or incorrect with a few outliers in the correct category when trying to distinguish them from the Ti-sludge NMs. This is because the L22 Ti-NMs that have elemental associations are also found in Ti-sludge containing Mg, Mn, Pb, and Ba, so the model was not able to confidently distinguish them either as ENMs or NNMs. The unique elemental associations in the Ti-sludge such as Sn, Ag, and Cu (Figure S11D) led to a fraction of recovered Ti-NMs having a high confidence regardless of Ti-mass, but the model did allocate a small fraction of multi-element recovered Ti-NMs that have Ag between 50 and 85% confidence (Figure S13). Ti-sludge particles without elemental associations follow a sigmoidal curve similar to the L22 + Ti100 case as a function of Ti-mass.

Nearly half of the Ti-sludge NMs had an elemental fingerprint. So, even though they have an overlapping mass distribution with the Ti-based NNMs in L22, 35% of them were distinguishable from the NNMs recovered from L22 soil. The lower amount of distinguishable ENM particles (35%) than the amount with an elemental fingerprint (54%) indicates that the

fingerprint was not unique enough to distinguish them from NNMs of soil. This may be because the Ti-sludge particles are associating with organic sludge particles, giving them a similar fingerprint. Previous studies have hypothesized that TiO_2 ENMs could attach to aluminosilicates via heteroaggregation and would lead to detection of multiple elements in a single particle by spICP-TOFMS but this warrants further investigation.[51,52] In the testing set, there is also a significant portion of Ti NMs in the sludge (46%) and in L22 (81%) soil that do not have elemental associations, but do have the same size distribution. This prevents higher model performance. Even though the Ti-sludge particles cannot be fully classified in L22 soil, the model recognizes a specific fingerprint and can distinguish a portion of Ti-sludge ENMs with high confidence.

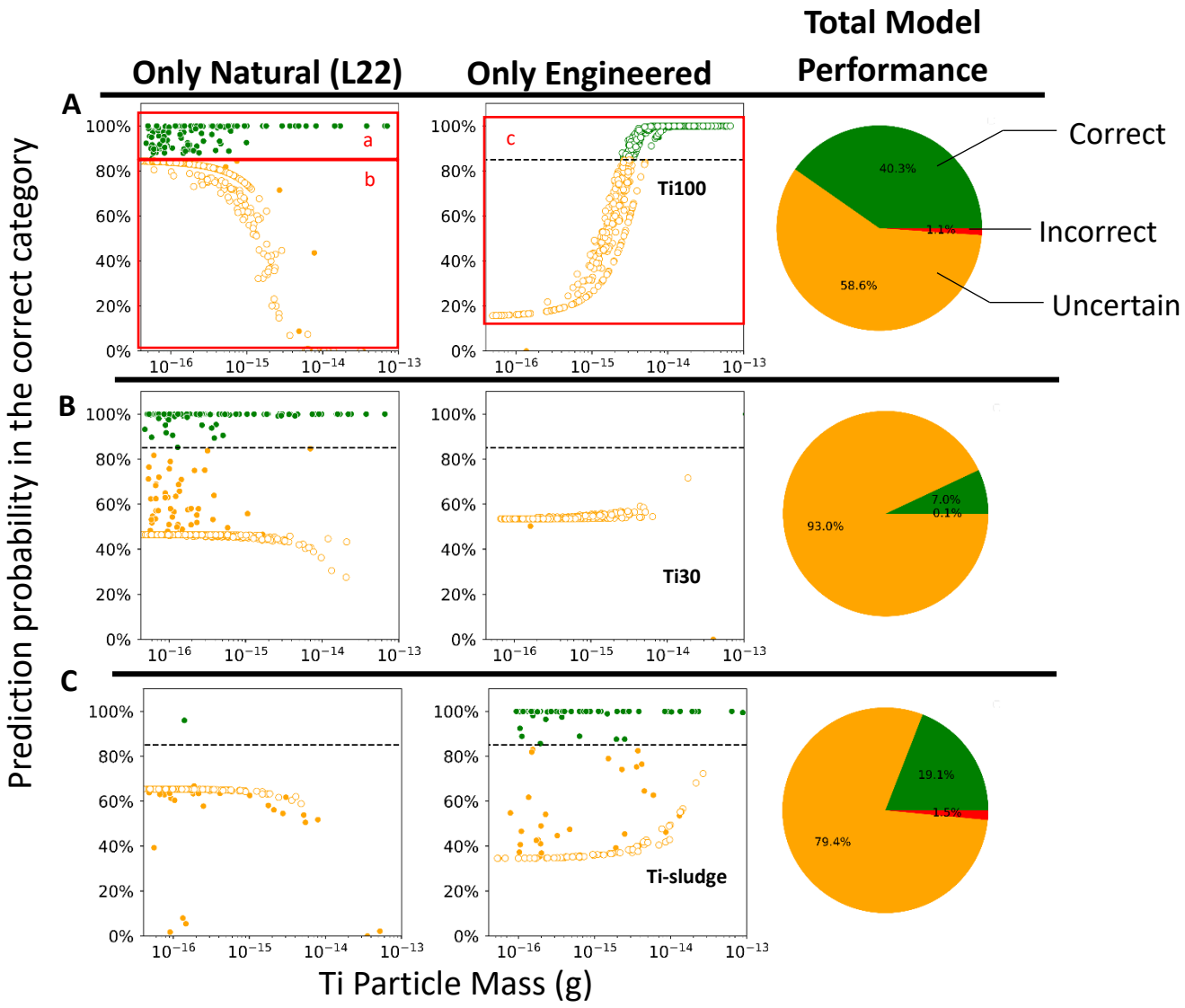


Figure 3: Effect of engineered particle type on the ability to classify ENMs and NNMs. ML results are trained on training datasets and predicts on a held-out test dataset for the following test cases (A) Ti100 + L22, B) Ti30 + L22, and C) Ti-sludge + L22. For each test case, the left scatter plot shows *only Ti-NNMs* (L22) displayed as the probability of the ML model predicting the correct category (NNM) as a function of Ti mass. The right scatter plot displays *only the Ti-ENMs* as the probability of the ML model predicting the correct category (ENM) as a function of Ti mass. Green

represent Ti-NMs correctly classified (> 85%) and orange is uncertain or misclassified (<85%). Filled circle markers are Ti-NMs with elemental associations and open circle markers are without elemental association. These colors correspond to the elemental fingerprint information in SI (Figure S11). The pie charts display the percentage of total Ti-NMs categorized as correct (green; >85%), uncertain (orange; <85% and >15%), and incorrectly classified (red; <15%).

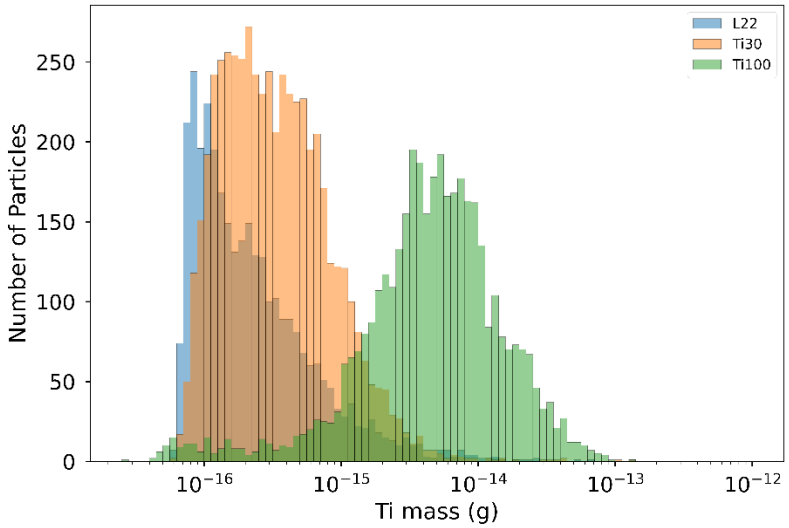


Figure 4: Comparing Ti-mass distribution between L22, Ti30, and Ti100. L22 and Ti30 significantly overlap while a clear separation exists between L22 and Ti100.

Influence of Ti-NNM properties on ability to classify Ti-based ENMs

The ability to classify Ti100 ENMs in each soil type was assessed to determine if differences in elemental fingerprint and size distribution of the NNMs affected the ability to detect the Ti100 ENMs. The model was able to classify the Ti100 ENMs in all three soils. The model performed best for the Ti100 + ARZ (64.5%), followed by Ti100 + LUV (50.3%) and Ti100 + L22 (40.3%) (Figure 4). Note that the Ti100 + L22 case is shown in Figure 3. For the ARZ case, a portion of Ti-NNMs without elemental associations were classified correctly (> 85%) because

of their narrow mass distribution compared to LUV and L22. For the LUV case, there are more Ti-NNMs that associate with other elements compared to L22, and therefore more particles were correctly classified in the LUV soil compared to the L22 soil.

For all cases, the classification of the ENMs followed a sigmoidal curve which is expected since Ti100 did not have any elemental associations and the model used Ti-mass only for classification. The similar model results from the three cases reflects the similar elemental fingerprint amount and mass distribution between the three NNMs, even though the soils are from different locations and have different soil properties (Table S2). This is also reflected in the results of the remaining four cases in SI (Figure S10)

The model also determines what elements are important and used to measure how size and elemental fingerprint are weighted for each test case. The element importance is determined by a LR model (Figure S14, calculations described in SI). For this calculation, the frequency of the element within the sample is considered. For the L22 and ARZ soil cases, ^{46}Ti and ^{48}Ti are the most important analytes for the prediction of a Ti-particle. This is expected because the model identifies the significant Ti-mass differences to classify the two categories. The remaining analytes are only identified in NNMs, such as ^{93}Nb, ^{25}Mg, ^{208}Pb, ^{57}Fe, ^{55}Mn, ^{90}Zr and ^{138}Ba. For the LUV case, ^{55}Mn ranked higher than ^{46}Ti as the second most important element. The model also identified elements that are exclusively important to each soil: ^{51}V in L22, ^{85}Rb in LUV and ^{89}Y in the ARZ soil. Even though these are minor components in the fingerprint, the model does identify important elements specific to each NNM.

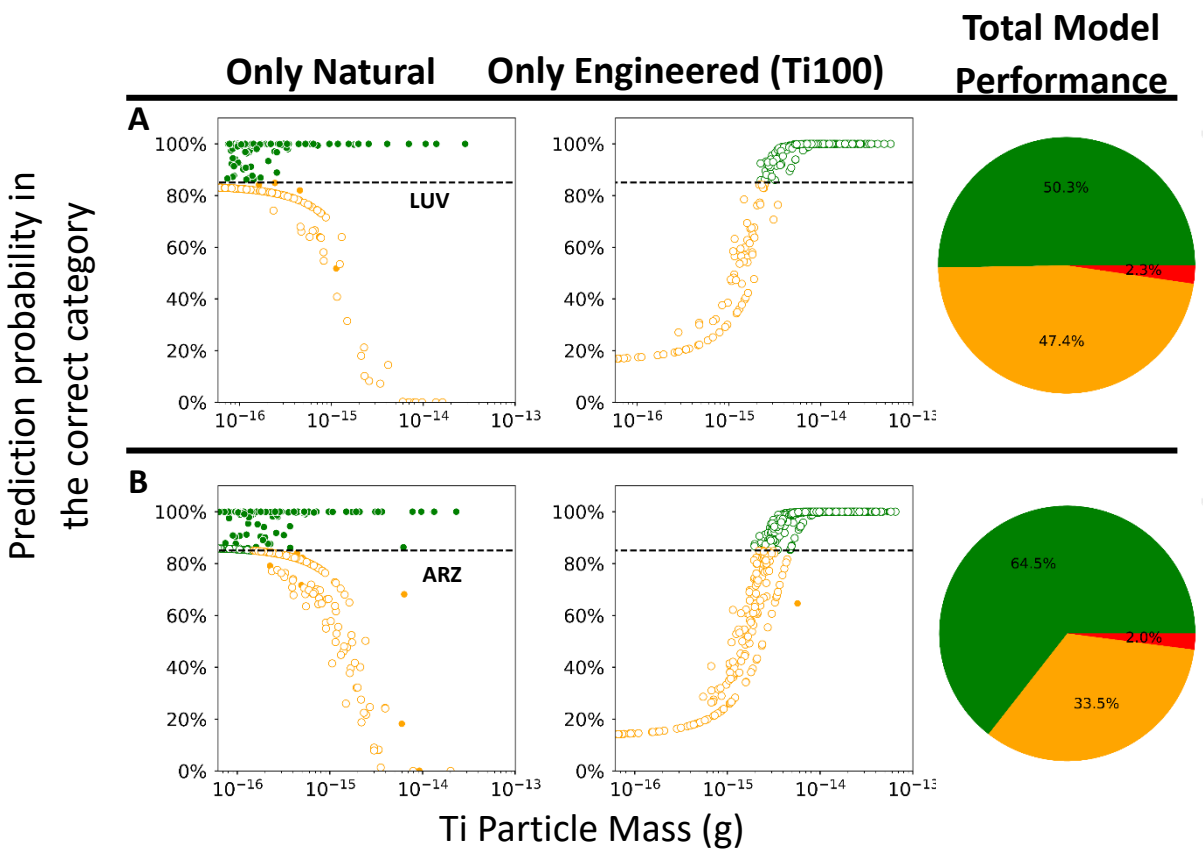


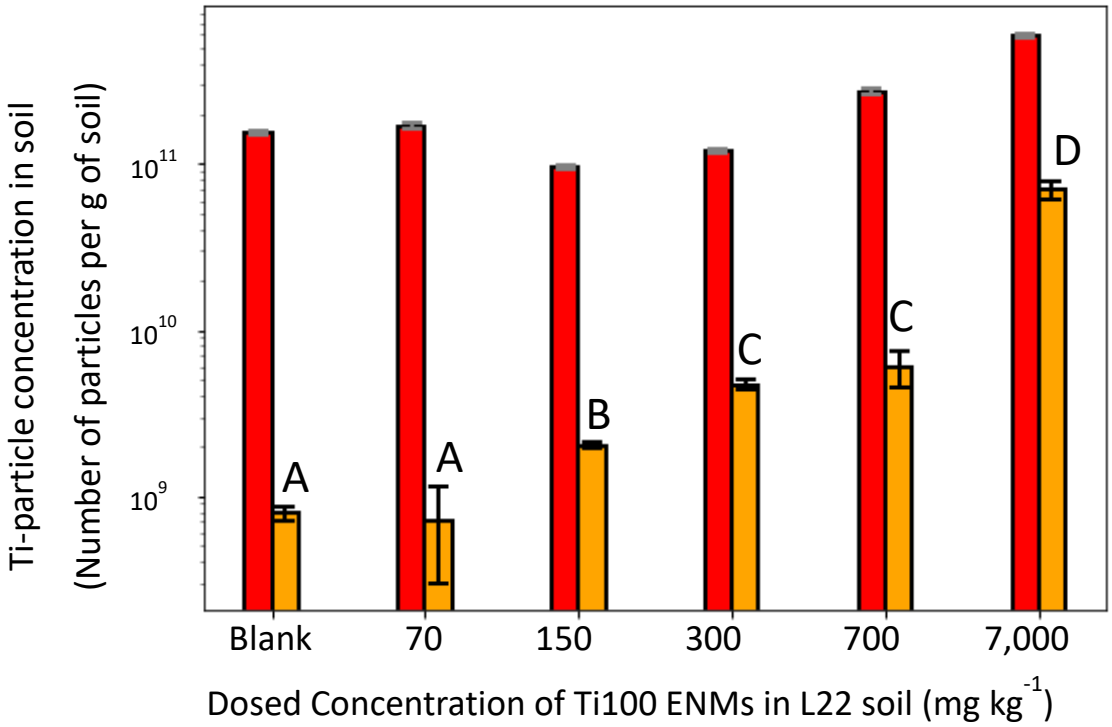
Figure 5: Effect of soil type on the ability to classify ENMs and NNMs. ML results of the training data for each test case (LUV + Ti100, ARZ + Ti100). Left scatter plots are only Ti-NNMs (L22) displayed as the probability of the ML model predicting the correct category as a function of Ti mass. Right scatter plots are the same but for only Ti-ENMs. Green represent Ti-NMs correctly classified (> 85%) and orange is uncertain or misclassified (< 85%). The pie charts display the percentage of total Ti-NMs categorized as correct (green), uncertain (orange), and incorrectly classified (red).

Classifying Ti100 ENMs in dosed L22 soil

We tested the ability of the trained ML model (Ti100 + L22) to classify the <500 nm sized particles extracted from L22 soil dosed with Ti100 particles as either ENMs or NNMs. Five

522 different Ti100 ENM concentrations were used (70, 150, 300, 700, and 7,000 mg kg^{-1}) along with
523 an undosed (blank) soil to determine if the classification accuracy depended on particle
524 concentration. Recovered Ti-containing particle concentrations ranged from 9.7 x 10^{10} to 6.1 x
525 10^{11} particles per g of soil (150 mg kg^{-1} and 7,000 mg kg^{-1} TiO_2, respectively). The number of
526 particles predicted as ENMs (> 85% probability) increased with increasing Ti100 concentration
527 (Figure 6). The amount of Ti100 added to the soil correlated to the total amount of Ti that was
528 extracted (Figure S22), suggesting that the extraction procedure was recovering the added Ti100
529 particles. From the blank, the total background Ti is approximately 1,600 mg kg^{-1}. Therefore, total
530 Ti recovery for dosed TiO_2 ENMs was not achieved for dosed concentrations lower than 7,000 mg
531 kg^{-1}. However, more Ti-containing particles were being recovered in dosed soils. At the highest
532 nominal concentration (7,000 mg kg^{-1}), the percent recovery of Ti was 36 ± 15 % by mass. This
533 highlights the value of spICP-TOFMS for identifying recovered individual Ti-containing particles
534 rather than total Ti in a sample. For the blank L22 soil, the model predicted low amounts of
535 particles labeled as ENMs (0.4%). These particles predicted as ENMs were unassociated with any
536 elements and had a mass distribution that is similar to the Ti100. There is not a statistically
537 significant increase in the predicted ENMs between the blank L22 and the soil with 70 mg kg^{-1}
538 TiO_2, but there is a difference between 70 mg kg^{-1} and 150 mg kg^{-1} amended soils, and between
539 150 mg kg^{-1} and 300 mg kg^{-1} (P value < 10^{-2}). So, the detection limit for the Ti-100 ENMs in L22
540 is about 150 mg kg^{-1}. Current ENM emission models estimate 0.1 to 10 mg kg^{-1} of nano TiO_2.
541 However, sludge and sludge-treated soil may exhibit higher concentrations that could be similar
542 to or exceeding the detection limit.[40,53] A factor to consider in this estimate of the detection limit
543 is aging time. Using an aging time greater than 24 hours could lower the extraction efficiency of
544 the NMs from soil, thus increasing the ENM particle detection limit. A previous study performed

545 a similar experiment with CeO_2 ENMs dosed in soils and a trained ML model that included 17
546 other elements. Their ML model detected more ENMs in dosed soil with concentrations of 40 mg
547 kg^{-1} and above.[28] This is not surprising given the distinction of Ce-based ENMs in soils is easier
548 because of the unique elemental fingerprint in Ce-NNMs (e.g., Ce + La) and the lower abundance
549 of Ce-NNMs in soil.



550

551 Figure 6: The ability to detect Ti100 ENMs dosed into L22 soil as a function of ENM
552 concentration. Ti particle concentration for each dosed Lufa soil: Blank (0 mg kg^{-1}), 70 mg kg^{-1},
553 150 mg kg^{-1}, 300 mg kg^{-1}, 700 mg kg^{-1}, 7,000 mg kg^{-1}. Total Ti particle concentration is the red
554 bar. Orange bar represents the percentage of Ti-particles categorized as ENMs by the ML model
555 with a ≥ 85% probability. Letters represent statistical similarity for one-way ANOVA (P value <
556 10^{-2}). Sample size n = 3.

Environmental Implications

Currently, mass flow analysis models and fate and transport ENM models are used to estimate ENM concentrations in the environment.[15,25,53–55] While these models can be robust, they cannot be easily validated because current analytical strategies cannot measure ENM concentrations in relevant environment sinks (soil and sediments) at low concentration, especially for ENMs made from earth-abundant elements that are some of the most common used ENMs.[56] The analytical strategy presented in this study is one of the first to detect and estimate concentrations of a common ENM type (TiO_2) in soils and determine what NM properties makes the ENMs distinguishable. This is a significant advancement towards validating existing fate and transport ENM models. This analytical strategy can also potentially use elemental and mass fingerprints to distinguish between specific ENM sources in the environment and determine ENM source apportionment at impacted sites.

We assessed the efficacy of distinguishing Ti-based ENMs originating from three different sources from NNMs in three different soil types using spICP-TOFMS and machine learning. The majority of Ti-NNMs in soil are without elemental associations and distinguishing them from ENMs is not possible using elemental fingerprints alone. However, machine learning models that use both elemental information and mass distribution could positively identify ENMs in soils except for Ti30 because of the limited elemental associations and the overlap in size with Ti-containing NNMs. While not investigated here, the abundance of other relatively pure ENMs made from earth abundant elements, e.g., iron oxides, may be similarly difficult to identify with confidence depending on the mass distribution and elemental fingerprint of naturally occurring iron oxides. For the Ti-sludge case, the model was able to distinguish a significant portion of the Ti-sludge derived ENMs added to soil. It suggests that this approach could be applied to determine

580 the presence of Ti-sludge ENMs applied to agricultural soils. Detecting sludge NMs in agricultural
581 soil would imply their persistence in the environment and could lead to potential leachability in
582 surrounding water systems.[57]

583 The ML models developed for this study identify specific elements and mass differences to
584 classify ENMs and NNMs. The models can learn ratios of similar associated elements to
585 differentiate ENMs and NNMs at a much deeper level than through observation alone. The ML
586 model simultaneously uses both elemental and size information, providing better classification
587 performance than using only elemental fingerprinting.

588 An important limitation of the supervised ML model is that it cannot yet be generalized to
589 identify a selected TiO_2 ENM in any soil. This is because it requires prior information on the
590 elemental and mass fingerprint on the ENMs of interest, and on the Ti-containing NNMs in the
591 soil prior to the introduction of ENMs. However, as with many ML approaches, creating a large,
592 shared dataset of fingerprints and mass distributions of Ti-containing NNMs from many different
593 soil types would create opportunities for classification and identification of ENMs in unknown
594 soils. Others have already implemented unsupervised learning methods to identify the elemental
595 fingerprints of NNMs in soil.[31] With enough data on the range and types of background particle
596 types in many different soils, and the signature of the specific ENM of interest, a model may be
597 developed for broader applicability to a range of soils using these approaches. While the ML
598 models can classify some ENMs in soil, the lack of an elemental fingerprint in many of the NNMs,
599 and overlapping size distributions, makes it challenging to improve model accuracy for ENMs
600 made from earth abundant elements like Ti. A more distinct fingerprint of ENMs is needed. This
601 can be achieved by improving the detection limits of the spICP-TOFMS to increase the number of
602 associated elements that can be detected, or by enriching the ENM with a stable isotope to help

distinguish the ENM from background.[58–60] These approaches could improve the ability to track ENMs made from earth abundant elements in environmental systems by spICP-TOFMS of extracted particles.

Supporting Information

ML Flowchart; NP Characterization; Soil Characterization and Experimental Design; Hyper Parameter Tuning; Bootstrapping Method to Evaluate Sample Size; Isotopes that were tracked with spICP-TOFMS; spICP-TOFMS parameters; Multi-elemental particle coincidence and polyatomic interference; Mass distribution of Ti-NMs; Remaining ML Cases; Elemental Fingerprint for each Test Case; Elemental Importance for each Test Case; The effect of H_2/He mode as a collision cell gas on Ti classification; Investigating the effect of Al on Ti classification; The effect of pretreating Ti30 sample on classification; Ti concentration measured by ICP-OES

Author Information

Garret Bland: Carnegie Mellon University, Pittsburgh, PA, 15213, USA

Gregory Lowry* (Corresponding Author): Carnegie Mellon University, Pittsburgh, PA, 15213, USA

Email: glowry@andrew.cmu.edu

Matthew Battifarano: Carnegie Mellon University, Pittsburgh, PA, 15213, USA

Ana Elena Pradas del Real: Madrid Institute for Agricultural Research (IMIDRA), N-II km 38,200, 28800 Alcala de Henares

Géraldine Sarret: ISTerre (Institut des Sciences de la Terre), Univ. Grenoble Alpes, CNRS, 38000, Grenoble, France

Acknowledgements

This research is funded by the U.S. Army Research Office under grant # W911NF190063 to Greg Lowry at CMU. The authors acknowledge use of the Materials Characterization Facility at Carnegie Mellon University supported by grant MCF-677785. The authors would like to thank Jonas Wielinski from Eawag – Swiss Federal Institute of Aquatic Science and Technology for ICP-OES analysis. The authors would like to thank Dr. Josef Werne and Dr. Thomas Elliot Arnold from University of Pittsburgh for the TOC analysis.

References

(1) Grieger, K. D.; Redmon, J. H.; Money, E. S.; Widder, M. W.; van der Schalie, W. H.; Beaulieu, S. M.; Womack, D. A Relative Ranking Approach for Nano-Enabled Applications to Improve Risk-Based Decision Making: A Case Study of Army Materiel. *Environ. Syst. Decis.* **2015**, *35* (1), 42–53. https://doi.org/10.1007/s10669-014-9531-4.

(2) Sun, T. Y.; Mitrano, D. M.; Bornhöft, N. A.; Scheringer, M.; Hungerbühler, K.; Nowack, B. Envisioning Nano Release Dynamics in a Changing World: Using Dynamic Probabilistic Modeling to Assess Future Environmental Emissions of Engineered Nanomaterials. *Environ. Sci. Technol.* **2017**, *51* (5), 2854–2863. https://doi.org/10.1021/acs.est.6b05702.

(3) Praetorius, A.; Scheringer, M.; Hungerbühler, K. Development of Environmental Fate Models for Engineered Nanoparticles - A Case Study of TiO 2 Nanoparticles in the Rhine River. *Environ. Sci. Technol.* **2012**, *46* (12), 6705–6713.

https://doi.org/10.1021/es204530n.

(4) Loosli, F.; Wang, J.; Rothenberg, S.; Bizimis, M.; Winkler, C.; Borovinskaya, O.; Flamigni, L.; Baalousha, M. Sewage Spills Are a Major Source of Titanium Dioxide Engineered (Nano)-Particle Release into the Environment. *Environ. Sci. Nano* **2019**, *6* (3), 763–777. https://doi.org/10.1039/c8en01376d.

(5) Holden, P. A.; Gardea-Torresdey, J. L.; Klaessig, F.; Turco, R. F.; Mortimer, M.; Hund-Rinke, K.; Cohen Hubal, E. A.; Avery, D.; Barceló, D.; Behra, R.; Cohen, Y.; Deydier-Stephan, L.; Ferguson, P. L.; Fernandes, T. F.; Herr Harthorn, B.; Henderson, W. M.; Hoke, R. A.; Hristozov, D.; Johnston, J. M.; Kane, A. B.; Kapustka, L.; Keller, A. A.; Lenihan, H. S.; Lovell, W.; Murphy, C. J.; Nisbet, R. M.; Petersen, E. J.; Salinas, E. R.; Scheringer, M.; Sharma, M.; Speed, D. E.; Sultan, Y.; Westerhoff, P.; White, J. C.; Wiesner, M. R.; Wong, E. M.; Xing, B.; Steele Horan, M.; Godwin, H. A.; Nel, A. E. Considerations of Environmentally Relevant Test Conditions for Improved Evaluation of Ecological Hazards of Engineered Nanomaterials. *Environmental Science and Technology*. American Chemical Society June 21, 2016, pp 6124–6145. https://doi.org/10.1021/acs.est.6b00608.

(6) Giese, B.; Klaessig, F.; Park, B.; Kaegi, R.; Steinfeldt, M.; Wigger, H.; Von Gleich, A.; Gottschalk, F. Risks, Release and Concentrations of Engineered Nanomaterial in the Environment. *Sci. Rep.* **2018**, *8* (1), 1–18. https://doi.org/10.1038/s41598-018-19275-4.

(7) Fleischer, M.; Parker, R. L. Composition of the Earth's Crust. In *Data of Geochemistry*; United States Government Printing Office, 1967.

(8) Gondikas, A.; Von Der Kammer, F.; Kaegi, R.; Borovinskaya, O.; Neubauer, E.;

Navratilova, J.; Praetorius, A.; Cornelis, G.; Hofmann, T. Where Is the Nano? Analytical Approaches for the Detection and Quantification of TiO2 Engineered Nanoparticles in Surface Waters. *Environ. Sci. Nano* **2018**, *5* (2), 313–326. https://doi.org/10.1039/c7en00952f.

(9) Navratilova, J.; Praetorius, A.; Gondikas, A.; Fabienke, W.; von der Kammer, F.; Hofmann, T. Detection of Engineered Copper Nanoparticles in Soil Using Single Particle ICP-MS. *Int. J. Environ. Res. Public Health* **2015**, *12* (12), 15756–15768. https://doi.org/10.3390/ijerph121215020.

(10) Regelink, I. C.; Weng, L.; van Riemsdijk, W. H. The Contribution of Organic and Mineral Colloidal Nanoparticles to Element Transport in a Podzol Soil. *Appl. Geochemistry* **2011**, *26* (SUPPL.), S241–S244. https://doi.org/10.1016/j.apgeochem.2011.03.114.

(11) Tiede, K.; Boxall, A. B. A.; Tiede, D.; Tear, S. P.; David, H.; Lewis, J. A Robust Size-Characterisation Methodology for Studying Nanoparticle Behaviour in "real" Environmental Samples, Using Hydrodynamic Chromatography Coupled to ICP-MS. *J. Anal. At. Spectrom.* **2009**, *24* (7), 964–972. https://doi.org/10.1039/b822409a.

(12) Spielman-Sun, E.; Lombi, E.; Donner, E.; Howard, D.; Unrine, J. M.; Lowry, G. V. Impact of Surface Charge on Cerium Oxide Nanoparticle Uptake and Translocation by Wheat (*Triticum Aestivum*). *Environ. Sci. Technol.* **2017**, *51* (13), 7361–7368. https://doi.org/10.1021/acs.est.7b00813.

(13) Laughton, S.; Laycock, A.; von der Kammer, F.; Hofmann, T.; Casman, E. A.; Rodrigues, S. M.; Lowry, G. V. Persistence of Copper-Based Nanoparticle-Containing Foliar Sprays in Lactuca Sativa (Lettuce) Characterized by SpICP-MS. *J. Nanoparticle Res.* **2019**, *21*

(8). https://doi.org/10.1007/s11051-019-4620-4.

(14) Stampoulis, D.; Sinha, S. K.; White, J. C. Assay-Dependent Phytotoxicity of Nanoparticles to Plants. *Environ. Sci. Technol.* **2009**, *43* (24), 9473–9479. https://doi.org/10.1021/es901695c.

(15) Gottschalk, F.; Sun, T.; Nowack, B. Environmental Concentrations of Engineered Nanomaterials: Review of Modeling and Analytical Studies. *Environmental Pollution*. Elsevier Ltd October 1, 2013, pp 287–300. https://doi.org/10.1016/j.envpol.2013.06.003.

(16) Gallego-Urrea, J. A.; Tuoriniemi, J.; Pallander, T.; Hassellöv, M. Measurements of Nanoparticle Number Concentrations and Size Distributions in Contrasting Aquatic Environments Using Nanoparticle Tracking Analysis. *Environ. Chem.* **2010**, *7* (1), 67–81. https://doi.org/10.1071/EN09114.

(17) Gimbert, L. J.; Haygarth, P. M.; Beckett, R.; Worsfold, P. J. Comparison of Centrifugation and Filtration Techniques for the Size Fractionation of Colloidal Material in Soil Suspensions Using Sedimentation Field-Flow Fractionation. *Environ. Sci. Technol.* **2005**, *39* (6), 1731–1735. https://doi.org/10.1021/es049230u.

(18) Reed, R. B.; Martin, D. P.; Bednar, A. J.; Montaño, M. D.; Westerhoff, P.; Ranville, J. F. Multi-Day Diurnal Measurements of Ti-Containing Nanoparticle and Organic Sunscreen Chemical Release during Recreational Use of a Natural Surface Water. *Environ. Sci. Nano* **2017**, *4* (1), 69–77. https://doi.org/10.1039/c6en00283h.

(19) Plathe, K. L.; Von Der Kammer, F.; Hassellov, M.; Moore, J.; Murayama, M.; Hofmann, T.; Hochella, M. F. Using FlFFF and ATEM to Determine Trace Metalnanoparticle Associations in Riverbed Sediment. *Environ. Chem.* **2010**, *7* (1), 82–93.

https://doi.org/10.1071/EN09111.

(20) Bolea, E.; Laborda, F.; Castillo, J. R. Metal Associations to Microparticles, Nanocolloids and Macromolecules in Compost Leachates: Size Characterization by Asymmetrical Flow Field-Flow Fractionation Coupled to ICP-MS. *Anal. Chim. Acta* **2010**, *661* (2), 206–214. https://doi.org/10.1016/J.ACA.2009.12.021.

(21) Loosli, F.; Yi, Z.; Wang, J.; Baalousha, M. Dispersion of Natural Nanomaterials in Surface Waters for Better Characterization of Their Physicochemical Properties by AF4-ICP-MS-TEM. *Sci. Total Environ.* **2019**, *682*, 663–672. https://doi.org/10.1016/j.scitotenv.2019.05.206.

(22) Zhou, X. X.; Liu, J. F.; Geng, F. L. Determination of Metal Oxide Nanoparticles and Their Ionic Counterparts in Environmental Waters by Size Exclusion Chromatography Coupled to ICP-MS. *NanoImpact* **2016**, *1*, 13–20. https://doi.org/10.1016/j.impact.2016.02.002.

(23) Degueldre, C.; Favarger, P. Y. Colloid Analysis by Single Particle Inductively Coupled Plasma-Mass Spectroscopy: A Feasibility Study. In *Colloids and Surfaces A: Physicochemical and Engineering Aspects*; Elsevier, 2003; Vol. 217, pp 137–142. https://doi.org/10.1016/S0927-7757(02)00568-X.

(24) Degueldre, C.; Favarger, P. Y.; Bitea, C. Zirconia Colloid Analysis by Single Particle Inductively Coupled Plasma-Mass Spectrometry. *Anal. Chim. Acta* **2004**, *518* (1–2), 137–142. https://doi.org/10.1016/j.aca.2004.04.015.

(25) Keller, A. A.; Lazareva, A. Predicted Releases of Engineered Nanomaterials: From Global to Regional to Local. *Environ. Sci. Technol. Lett.* **2013**, *1* (1), 65–70. https://doi.org/10.1021/ez400106t.

(26) Cornu, S.; Lucas, Y.; Lebon, E.; Ambrosi, J. P.; Luizão, F.; Rouiller, J.; Bonnay, M.; Neal, C. Evidence of Titanium Mobility in Soil Profiles, Manaus, Central Amazonia. *Geoderma* **1999**, *91* (3–4), 281–295. https://doi.org/10.1016/S0016-7061(99)00007-5.

(27) Hendriks, L.; Gundlach-graham, A.; Günther, D. Analysis of Inorganic Nanoparticles by Single-Particle Inductively Coupled Plasma Time-of-Flight Mass Spectrometry. *Chimia (Aarau).* **2018**, *72* (4), 221–226. https://doi.org/10.2533/chimia.2018.221.

(28) Praetorius, A.; Gundlach-Graham, A.; Goldberg, E.; Fabienke, W.; Navratilova, J.; Gondikas, A.; Kaegi, R.; Günther, D.; Hofmann, T.; Von Der Kammer, F. Single-Particle Multi-Element Fingerprinting (SpMEF) Using Inductively-Coupled Plasma Time-of-Flight Mass Spectrometry (ICP-TOFMS) to Identify Engineered Nanoparticles against the Elevated Natural Background in Soils. *Environ. Sci. Nano* **2017**, *4* (2), 307–314. https://doi.org/10.1039/c6en00455e.

(29) Green, T. H.; Pearson, N. J. An Experimental Study of Nb and Ta Partitioning between Ti-Rich Minerals and Silicate Liquids at High Pressure and Temperature. *Geochim. Cosmochim. Acta* **1987**, *51* (1), 55–62. https://doi.org/10.1016/0016-7037(87)90006-8.

(30) Gieré, R.; Williams, C. T. REE-Bearing Minerals in a Ti-Rich Vein from the Adamello Contact Aureole (Italy). *Contrib. to Mineral. Petrol.* **1992**, *112* (1), 83–100. https://doi.org/10.1007/BF00310957.

(31) Baalousha, M.; Wang, J.; Erfani, M.; Goharian, E. Elemental Fingerprints in Natural Nanomaterials Determined Using SP-ICP-TOF-MS and Clustering Analysis. *Sci. Total Environ.* **2021**, *792*, 148426. https://doi.org/10.1016/J.SCITOTENV.2021.148426.

(32) Mitra, S.; Datta, S.; Perkins, T.; Michailidis, G. *Introduction to Machine Learning and*

Bioinformatics; Taylor & Francis, 2008.

(33) Yang, X.; Liu, X.; Zhang, A.; Lu, D.; Li, G.; Zhang, Q.; Liu, Q.; Jiang, G. Distinguishing the Sources of Silica Nanoparticles by Dual Isotopic Fingerprinting and Machine Learning. *Nat. Commun.* **2019**, *10* (1). https://doi.org/10.1038/s41467-019-09629-5.

(34) Del Real, A. E. P.; Castillo-Michel, H.; Kaegi, R.; Sinnet, B.; Magnin, V.; Findling, N.; Villanova, J.; Carriére, M.; Santaella, C.; Fernández-Martínez, A.; Levard, C.; Sarret, G. Fate of Ag-NPs in Sewage Sludge after Application on Agricultural Soils. *Environ. Sci. Technol.* **2016**, *50* (4), 1759–1768. https://doi.org/10.1021/acs.est.5b04550.

(35) Pradas del Real, A. E.; Castillo-Michel, H.; Kaegi, R.; Larue, C.; de Nolf, W.; Reyes-Herrera, J.; Tucoulou, R.; Findling, N.; Salas-Colera, E.; Sarret, G. Searching for Relevant Criteria to Distinguish Natural *vs.* Anthropogenic TiO_2 Nanoparticles in Soils. *Environ. Sci. Nano* **2018**, *5* (12), 2853–2863. https://doi.org/10.1039/C8EN00386F.

(36) Lahive, E.; Matzke, M.; Durenkamp, M.; Lawlor, A. J.; Thacker, S. A.; Pereira, M. G.; Spurgeon, D. J.; Unrine, J. M.; Svendsen, C.; Lofts, S. Sewage Sludge Treated with Metal Nanomaterials Inhibits Earthworm Reproduction More Strongly than Sludge Treated with Metal Metals in Bulk/Salt Forms. *Environ. Sci. Nano* **2017**, *4* (1), 78–88. https://doi.org/10.1039/C6EN00280C.

(37) Bland, G. D.; Lowry, G. V. Multi-Step Method to Extract Moderately Soluble Copper Oxide Nanoparticles from Soil for Quantification and Characterization. *Anal. Chem.* **2020**, *92* (14), 9620–9628. https://doi.org/10.1021/acs.analchem.0c00824.

(38) Pandorf, M.; Pourzahedi, L.; Gilbertson, L.; Lowry, G. V.; Herckes, P.; Westerhoff, P. Graphite Nanoparticle Addition to Fertilizers Reduces Nitrate Leaching in Growth of

Lettuce (Lactuca Sativa). *Environ. Sci. Nano* **2020**, *7* (1), 127–138. https://doi.org/10.1039/C9EN00890J.

(39) Gao, X.; Spielman-Sun, E.; Rodrigues, S. M.; Casman, E. A.; Lowry, G. V. Time and Nanoparticle Concentration Affect the Extractability of Cu from CuO NP-Amended Soil. *Environ. Sci. Technol.* **2017**, *51* (4), 2226–2234. https://doi.org/10.1021/acs.est.6b04705.

(40) Sun, T. Y.; Bornhöft, N. A.; Hungerbühler, K.; Nowack, B. Dynamic Probabilistic Modeling of Environmental Emissions of Engineered Nanomaterials. *Environ. Sci. Technol.* **2016**, *50* (9), 4701–4711. https://doi.org/10.1021/ACS.EST.5B05828.

(41) Pace, H. E.; Rogers, N. J.; Jarolimek, C.; Coleman, V. A.; Higgins, C. P.; Ranville, J. F. Determining Transport Efficiency for the Purpose of Counting and Sizing Nanoparticles via Single Particle Inductively Coupled Plasma Mass Spectrometry. *Anal. Chem.* **2011**, *83* (24), 9361–9369. https://doi.org/10.1021/ac201952t.

(42) Tanner, M.; Günther, D. Short Transient Signals, a Challenge for Inductively Coupled Plasma Mass Spectrometry, a Review. *Analytica Chimica Acta*. Elsevier February 2, 2009, pp 19–28. https://doi.org/10.1016/j.aca.2008.11.041.

(43) Venkatesan, A. K.; Reed, R. B.; Lee, S.; Bi, X.; Hanigan, D.; Yang, Y.; Ranville, J. F.; Herckes, P.; Westerhoff, P. Detection and Sizing of Ti-Containing Particles in Recreational Waters Using Single Particle ICP-MS. *Bull. Environ. Contam. Toxicol.* **2018**, *100* (1), 120–126. https://doi.org/10.1007/s00128-017-2216-1.

(44) Dan, Y.; Shi, H.; Stephan, C.; Liang, X. Rapid Analysis of Titanium Dioxide Nanoparticles in Sunscreens Using Single Particle Inductively Coupled Plasma–Mass Spectrometry. *Microchem. J.* **2015**, *122*, 119–126.

https://doi.org/10.1016/J.MICROC.2015.04.018.

(45) Liu, D. C.; Nocedal, J. On the Limited Memory BFGS Method for Large Scale Optimization. *Math. Program.* **1989**, *45* (1–3), 503–528. https://doi.org/10.1007/BF01589116.

(46) Wielinski, J.; Gogos, A.; Voegelin, A.; Müller, C.; Morgenroth, E.; Kaegi, R. Transformation of Nanoscale and Ionic Cu and Zn during the Incineration of Digested Sewage Sludge (Biosolids). *Environ. Sci. Technol.* **2019**, *53* (20), 11704–11713. https://doi.org/10.1021/acs.est.9b01983.

(47) Weaver, C. E. The Nature of TiO2 in Kaolinite. *Clays Clay Miner.* **1976**, *24* (5), 215–218. https://doi.org/10.1346/ccmn.1976.0240501.

(48) Cornu, S.; Lucas, Y.; Lebon, E.; Ambrosi, J. P.; Luizão, F.; Rouiller, J.; Bonnay, M.; Neal, C. Evidence of Titanium Mobility in Soil Profiles, Manaus, Central Amazonia. *Geoderma* **1999**, *91* (3–4), 281–295. https://doi.org/10.1016/S0016-7061(99)00007-5.

(49) Van Baalen, M. R. Titanium Mobility in Metamorphic Systems: A Review. *Chem. Geol.* **1993**, *110* (1–3), 233–249. https://doi.org/10.1016/0009-2541(93)90256-I.

(50) Philippe, A.; Campos, D. A.; Guigner, J.-M.; Buchmann, C.; Diehl, D.; Schaumann, G. E. Characterization of the Natural Colloidal TiO2 Background in Soil. *Sep. 2018, Vol. 5, Page 50* **2018**, *5* (4), 50. https://doi.org/10.3390/SEPARATIONS5040050.

(51) Azimzada, A.; Farner, J. M.; Jreije, I.; Hadioui, M.; Liu-Kang, C.; Tufenkji, N.; Shaw, P.; Wilkinson, K. J. Single- and Multi-Element Quantification and Characterization of TiO2 Nanoparticles Released From Outdoor Stains and Paints. *Front. Environ. Sci.* **2020**, *8*, 91.

https://doi.org/10.3389/fenvs.2020.00091.

(52) Praetorius, A.; Badetti, E.; Brunelli, A.; Clavier, A.; Gallego-Urrea, J. A.; Gondikas, A.; Hassellöv, M.; Hofmann, T.; Mackevica, A.; Marcomini, A.; Peijnenburg, W.; Quik, J. T. K.; Seijo, M.; Stoll, S.; Tepe, N.; Walch, H.; Von Der Kammer, F. Strategies for Determining Heteroaggregation Attachment Efficiencies of Engineered Nanoparticles in Aquatic Environments. *Environ. Sci. Nano* **2020**, *7* (2), 351–367. https://doi.org/10.1039/c9en01016e.

(53) Sun, T. Y.; Gottschalk, F.; Hungerbühler, K.; Nowack, B. Comprehensive Probabilistic Modelling of Environmental Emissions of Engineered Nanomaterials. *Environ. Pollut.* **2014**, *185*, 69–76. https://doi.org/10.1016/j.envpol.2013.10.004.

(54) Gottschalk, F.; Nowack, B. The Release of Engineered Nanomaterials to the Environment. *Journal of Environmental Monitoring*. May 2011, pp 1145–1155. https://doi.org/10.1039/c0em00547a.

(55) Garner, K. L.; Suh, S.; Keller, A. A. Assessing the Risk of Engineered Nanomaterials in the Environment: Development and Application of the NanoFate Model. *Environ. Sci. Technol.* **2017**, *51* (10), 5541–5551. https://doi.org/10.1021/acs.est.6b05279.

(56) Dale, A. L.; Casman, E. A.; Lowry, G. V.; Lead, J. R.; Viparelli, E.; Baalousha, M. Modeling Nanomaterial Environmental Fate in Aquatic Systems. *Environ. Sci. Technol.* **2015**, *49* (5), 2587–2593. https://doi.org/10.1021/es505076w.

(57) Brar, S. K.; Verma, M.; Tyagi, R. D.; Surampalli, R. Y. Engineered Nanoparticles in Wastewater and Wastewater Sludge - Evidence and Impacts. *Waste Management*. Pergamon March 1, 2010, pp 504–520. https://doi.org/10.1016/j.wasman.2009.10.012.

(58) Laycock, A.; Diez-Ortiz, M.; Larner, F.; Dybowska, A.; Spurgeon, D.; Valsami-Jones, E.; Rehkämper, M.; Svendsen, C. Earthworm Uptake Routes and Rates of Ionic Zn and ZnO Nanoparticles at Realistic Concentrations, Traced Using Stable Isotope Labeling. *Environ. Sci. Technol.* **2016**, *50* (1), 412–419. https://doi.org/10.1021/acs.est.5b03413.

(59) Larner, F.; Dogra, Y.; Dybowska, A.; Fabrega, J.; Stolpe, B.; Bridgestock, L. J.; Goodhead, R.; Weiss, D. J.; Moger, J.; Lead, J. R.; Valsami-Jones, E.; Tyler, C. R.; Galloway, T. S.; Rehkämper, M. Tracing Bioavailability of ZnO Nanoparticles Using Stable Isotope Labeling. *Environ. Sci. Technol.* **2012**, *46* (21), 12137–12145. https://doi.org/10.1021/es302602j.

(60) Larner, F.; Gulson, B.; McCall, M.; Oytam, Y.; Rehkämper, M. An Inter-Laboratory Comparison of High Precision Stable Isotope Ratio Measurements for Nanoparticle Tracing in Biological Samples. *J. Anal. At. Spectrom.* **2014**, *29* (3), 471–477. https://doi.org/10.1039/c3ja50322d.