

DISS. ETH NO. 23451

UNDERSTANDING ENGINEERED NANOMATERIAL  
TRANSPORT IN SOIL: RECONCILING WEAKNESSES  
IN ADVECTION DISPERSION MODELING WITH  
MACHINE LEARNING

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

ELI GOLDBERG

MSc ETH in Environmental Engineering, ETH Zurich

born on 15.02.1984

citizen of USA

accepted on the recommendation of

Prof. Dr. Konrad Hungerbühler, examiner

Prof. Dr. Martin Scheringer, co-examiner

Prof. Dr. Alexander Wokaun, co-examiner

2016

---

## ABSTRACT

---

It is not yet analytically feasible to determine the concentration of engineered nanomaterials (ENMs) in real soils. This is because many of the metals that comprise the most commonly produced ENMs (e.g.,  $\text{TiO}_2$ ,  $\text{FeO}_x$ ,  $\text{ZnO}$ , and  $\text{CuO}$ ) occur naturally as nanoparticulates with similar size and shape. These natural analogues are also present in concentrations within the range of predicted environmental concentrations for ENMs in soil, which make it difficult to distinguish ENMs from the background concentration of natural nanomaterials. As a result, investigation of the transport behavior of ENMs is constrained to bench-top soil column transport experiments performed in simple, synthetic soil systems (i.e., monodisperse, fully saturated, single-media soils).

In Chapter 2, it is shown that most advection-dispersion based particle transport models (PTMs) commonly employed to track ENMs in soil are fundamentally flawed. The fundamental flaw is not that they lack descriptive ability entirely, as they can well describe a large portion of observed ENM transport behavior. It is that the mathematical construction of these models considers neither the physicochemical properties of the ENMs, nor those of the system as a whole, explicitly. Instead, and because the mechanistic understanding of how physicochemical conditions influence ENM transport is incomplete, researchers employ quasi-mechanistic PTMs that involve a great deal of empirical specification and ad hoc adjustment. As a result, the applicability of these models beyond the exact scenario under which they were conditioned is low and current PTMs are merely descriptive tools without predictive power.

In Chapter 3, machine learning is employed to enable the prediction of ENM transport behavior and overcome the limitations established in Chapter 2. To support the machine learning effort, a database of physicochemical conditions (explanatory variables) extracted from more than 200 ENM transport experiments conducted in saturated porous media is developed. This database is used to train machine learning models to predict the fraction of ENMs retained over a soil column length, a measure of transport distance, and to predict the retention profile shape, a measure of transport behavior. The retained

fraction is able to be predicted with a mean squared error between 0.025–0.033, and the retention profile shape is predicted with an F1-score (the weighted harmonic mean of precision and recall) between 60–70%. Further, by recursively removing physical and chemical features to optimize model predictive performance, the importance of the physicochemical features to predict transport and behavior is quantitatively determined and ranked - thereby establishing strategic targets for inclusion in new mechanistic PTMs.

In Chapter 4, the database developed in Chapter 3 is expanded and enhanced by combining many of the experimental physicochemical descriptors to derive common dimensionless fluid dynamics and particle interaction parameters. Evaluation of the investigated range of experimental conditions within the database shows that the environmental relevance of the studies to date is low and notable gaps in the examined range are identified. Then, a machine learning decision tree classification is employed to quantitatively, and visually, disentangle the relationship between physicochemical conditions and ENM transport behavior within the database. This approach allows the identification of the order of importance and range of influence of the physical, chemical, and electrical conditions that control the retention profile shape. This is important because it offers an alternative to the current approach and opens a path toward a deeper level of mechanistic understanding of ENM transport in saturated porous media. The current approach has not advanced the mechanistic understanding of ENM transport because it relies on PTMs that have no predictive power and, for the case of nonexponential profiles, even little descriptive power. Through the presented alternative approach, it is possible, for the first time, to identify domains of conditions where nonexponential retention is frequent/dominant, and thus where entirely new mechanistic approaches can be developed, and domains of conditions where exponential retention is frequent/dominant, thus existing approaches may be sufficient.

In Chapter 5, the machine learning methods in Chapter 3 are employed to create, evaluate, and validate a model to predict the ENM attachment efficiency,  $\alpha$ , using a database of physicochemical conditions extracted from quartz crystal microbalance with dissipation (QCM-D) experiments. Unlike traditional soil column experiments, QCM-D experiments directly monitor the deposition rate of particles onto an oscillating sensor surface in real-time, isolating deposition kinetics, and allowing for a direct measure of  $\alpha$  without contributions from convection or filtration.  $\alpha$  is an important kinetic transport

parameter and is widely used in colloid filtration theory (CFT), which is described in Chapter 2 of this thesis, to quantify the likelihood of a particle attaching to a surface after a collision.

In Chapter 6, a brief outlook is provided.

---

## ZUSAMMENFASSUNG

---

Es ist noch nicht praktikabel, die Konzentration von engineered nanomaterials (ENMs) in Böden zu bewerten. Dies ist darauf zurückzuführen, dass viele der Metalle, die in den am häufigsten hergestellten ENMs vorkommen, auch in der Natur als Nanopartikel erscheinen. Diese natürlichen Analoge sind auch in Konzentrationen vorhanden, die noch innerhalb der vorhergesagten Bandbreite liegen. Dies erschwert die Arbeit, jene ENMs von der Hintergrundkonzentration der natürlich vorkommenden Nanomaterialien zu unterscheiden. Aufgrund dessen beschränkt sich die Untersuchung auf Bench-Top Bodensäulen Transport Experimente in einfachen, synthetischen Bodensystemen (sprich monodisperse, vollgesättigte, single-media Böden).

Im 2. Kapitel wird gezeigt, dass advektions/diffusionsbasierte Partikeltransportmodelle (PTMs), die häufig gebraucht werden um ENMs in Böden zu folgen, grundlegende Mängel aufzeigen. Obwohl PTMs ein Grossteil des ENM Transportverhaltens beschreiben können, werden die physikalisch-chemischen Eigenschaften der ENMs, sowie die des Systems als Ganzes, ignoriert. Stattdessen, verwenden Forscher quasi-mechanistische PTMs die durch Ad-hoc-Anpassungen sowie empirische Spezifikation optimiert werden müssen. Dem zur Folge ist die breitere Anwendbarkeit der Modelle fast ausgeschlossen. Heutige PTMs sind lediglich als deskriptive Werkzeuge anwendbar, ohne echte Vorhersagekraft zu besitzen.

In 3 Kapitel werden maschinelle Lernverfahren angewendet um das Vorhersagen von ENM Transportverhalten zu ermöglichen, und ferner, die Beschränkungen der Themen aus Kapitel 2 zu überwinden. Um dieses Verfahren zu unterstützen, wird eine Datenbank der physikalisch-chemischen Eigenschaften aus über 200 ENM Transportversuchen, die in teilgesättigten, poröse Medien durchgeführt wurden, aufgesetzt. Die Datenbank wird verwendet um Machine-Learning-Modelle zu trainieren, die erstens die Verteilung der behaltene ENMs über die Länge einer Bodensäule – ein Mass für Transportlänge, und zweitens die Profilhaltigkeit – ein Mass für Transportverhalten – vorhersagen. Der Anteil gespeicherter ENMs kann mit einem mittleren quadrierten Fehler von 0.025-0.033, und die

Profilhaltigkeit mit einem F<sub>1</sub>-Mass von 60-70% vorhergesagt werden. Zudem wird die Wichtigkeit der einzelnen physikalisch-chemischen Eigenschaften eingestuft, anhand eines quantitativen Verfahrens – nämlich durch das rekursive löschen der Merkmale. Dadurch werden statische Ziele für neue mechanistische PTMs etabliert.

Im 4. Kapitel wird die Datenbank aus Kapitel 3 erweitert. Die Kombination der einzelnen physikalisch-chemischen Merkmale erlaubt die Ableitung von gemeinsamen Parametern, die für dimensionslose Strömungsdynamik sowie Teilcheninteraktion gelten. Die Bewertung der Versuchsbedingungen zeigt, dass bisherige Experimente wenig Relevanz aufzeigen. Durch Anwendung einer Machine-Learning Entscheidungsbaum-Klassifizierung entflechtet sich die Beziehung zwischen den physikalisch-chemischen Eigenschaften und das ENM Transportverhalten.

In 5 Kapitel werden die Machine-Learning-Modelle, die im 3. Kapitel beschrieben wurden, angewendet um ein Model der ENM Bindungseffizienz,  $\alpha$ , zu validieren. Zu diesem Zweck wurde eine Datenbank mit den Resultaten mehrere Quarzkristall-Mikrowaagen mit Dissipationsmessungs-(QCM-D) Experimenten aufgestellt. Im Gegensatz zu traditionellen Bodensäule-Experimenten, bietet QCM-D die Möglichkeit, Depositionsraten von Partikeln auf einer Sensorfläche in Echt-Zeit zu überwachen, was eine direkte Messung von  $\alpha$  ohne Einfluss von Filtrations- oder Konvektionsprozessen ermöglicht.  $\alpha$  ist ein wichtiger Transportparameter, der in der Kolloid-Filtrations Theorie (in Kapitel 2 ausführlich beschrieben) benutzt wird um die Bindungswahrscheinlichkeit eines Partikels nach der Kollision mit einer Fläche zu quantifizieren.

Im 6. Kapitel, wird ein kurzer Ausblick vorgestellt.