Goldschmidt 2022 Abstract
https://doi.org/10.46427/gold2022.12199

Causal Predictive Modeling for Trace Element Partitioning

AARON S WOLF[1], ERICA CUNG[2], GOKCE USTUNISIK[2,3] AND ROGER NIELSEN[2]

[1]Univ. of Michigan, Earth & Env. Sciences
[2]South Dakota School of Mines
[3]South Dakota Mines
Presenting Author: aswolf@umich.edu

Trace element partitioning systematics are sensitive to many factors that influence elemental exchange between minerals and melt, including temperature, pressure, composition, lattice strain, and elemental activities. Modeling these effects requires inference from experimental constraints, with good models making accurate predictions for unobserved conditions, enabling *interpolation within* and *extrapolation beyond* the original training dataset. Model quality should be assessed using out-of-sample performance tests and through reproduction of systematic behaviors of elemental groups (e.g., REEs).

Existing empirical models of trace-element partitioning typically rely upon simple linear regressions for only one or a few dependent variables [1,2,3]. This modeling strategy attempts to avoid overfitting the data by limiting model complexity, thereby improving predictive accuracy. While sometimes effective, this approach fails to account for selection bias imposed by the phase stability region. Since experiments require coexisting liquid and mineral phases, all parameters that influence phase equilibria become unavoidably correlated. This correlation is easily confused for a causal effect in overly-simplified regression models. Even worse, neglecting causal dependencies dramatically worsens prediction accuracy for conditions outside the training data.

We propose a new generalized causal model for trace element partitioning that directly accounts for the data-censoring effect of the melting-region, while dramatically improving model accuracy by identifying the dominant contributors to elemental partitioning. Applying causal inference principles [4], our model includes terms for each of the major factors affecting partitioning while controlling for the confounding effect of the melting-region. To train this expanded empirical model, we rely on the LEPR/traceDs database [www.earthchem.org] to model clinopyroxene, garnet, and amphibole partitioning behavior for 53 trace elements. To control overfitting, we combine statistical regularization constraints with a Monte Carlo analysis of correlated model uncertainties. The final result is a set of partitioning models for clinopyroxene, garnet, and amphibole with prediction accuracies that diverge substantially from simpler models, especially when applied to more extreme conditions like highly evolved liquids.

[1] Nielsen, R. L. (1990), *Reviews in Mineralogy and Geochemistry*, 24(1), 65-105.
[2] Bédard, J. H. (2006), *Geochimica et Cosmochimica Acta*, 70(14), 3717-3742.
[3] Blundy and Wood (1991), *Geochimica et Cosmochimica Acta*, 55(1),193-209.
[4] Pearl, J. (2009), *Statistics surveys*, 3, 96-146.