

鉱物化学組成の統計学的分類：ランダムフォレスト法による鉱物相の同定

Statistical classification of mineral chemistry: identification of mineral facies by random forest method

新井 宏嘉^{1*}, 大森 聡一²

Hiroyoshi Arai^{1*}, Soichi Omori²

¹早大・教育・地球科学, ²東工大・地球惑星科学

¹Earth Sci., Waseda Univ., ²Earth Planet. Sci., Tokyo Inst. Tech.

機器分析で得られた鉱物化学組成から鉱物相を識別することは、未知試料の同定における foolproof として機能するだけではない。特に、空間情報として鉱物相を識別することができれば、鉱物組み合わせやモード組成などの岩石学的基礎情報を得ることができたり、ある領域のバルク化学組成を算出する際の手がかりともなる。従来は偏光顕微鏡による鑑定や、XRFあるいはEPMAによる面分析結果をもとに、測定者の主観の入り込む余地を残したまま鉱物相の識別を行ってきた。これらの作業をより客観的なものにし、かつ自動化するため、これまでにいくつかの鉱物相同定方が考案されている（例えば菊池, 1991; Launeau et al., 1994; 戸上ほか, 1998; Togami et al., 2000; Maloy and Treiman, 2007; Tsuji et al., 2009）。しかしこれらの手法は、鉱物相識別理論が不足していたり、その試料に含まれる鉱物を仮定して解析するため汎用性が低い、といった問題がある。そこで我々は、比較的拡張性が高い方法として、集団学習の1つであるランダムフォレスト (random forest; Breiman, 2001) 法を用いて、鉱物化学組成データから鉱物相を識別することを試みた。

[解析手法]

ランダムフォレスト (以下, RF) 法とは、ブートストラップサンプリングと分類木を利用した集団学習法である。RF法の特徴は、分類精度が高く、データ数や変数の数が大きなデータでも効率よく解析でき、分類における変数の重要度を推定できる点などにある。多様な鉱物化学組成を学習データにしてRF法を用いれば、特定の岩型に限定されない汎用性の高い分類器を作ることができる。一方、RF法の欠点は、乱数を利用するために得られる結果が一定にならない、学習データに含まれる鉱物しか分類できないなどの点が挙げられる。

解析に用いたデータは、Deer et al.のRock-forming mineralsシリーズ第2版 (例えばDeer et al., 2001) に掲載されている酸化物を主体とする72の鉱物グループ、総数約3500個である。変数は、現在の鉱物化学組成データのほとんどがEPMAで得られる現状を鑑み、SiO₂, MgOなど陽イオン成分38種 (wt%) に限定した。FeおよびMnは、FeO*, MnO*に換算し、希土類元素はREE₂O₃としてまとめた。

RF法では群間のデータ数のバランスが解析結果でも保たれ、個数の少ない群は誤差も大きくなる。この影響を検討するため、群毎のデータ数の比率の逆数で重みをつけた場合も解析した。

[解析結果]

解析の結果、重みなしの場合約7.9%、重みつきの場合は約8.3%の誤判別率で分類することができた。同じ学習データでRF解析を100回繰り返した結果、誤判別率の範囲は両者とも約1%

に収まっており、乱数による解析結果の変動は全く問題にならない。化学組成による分類であるので、当然ながら多形鉱物を識別することは難しいが、ある多形に特徴的な副成分が存在する場合には、これを識別することもできた。分類に有効な成分は、重みの有無に係らず、通常のEPMA分析ルーチンで用いられる主要9成分 (SiO_2 , TiO_2 , Al_2O_3 , FeO^* , MnO^* , MgO , CaO , K_2O , Na_2O) で十分であることがわかった。この9成分だけで再解析した場合の誤判別率は、重みなしで約8.0%、重みつきで約8.6%となり、全成分を使用した時とほぼ同じである。

発表では、文献から選んだ様々な鉱物化学組成データをこの分類器で相同定した結果を示すと共に、EPMAによる元素マッピングデータから酸化物wt%データを計算し、そこから鉱物相マッピングを行った結果も示す予定である。

[文献]

Breiman, 2001, *Machine Learning*, 45, 5-32.

Deer et al., 2001, *Rock-forming minerals volume 1A: Orthosilicates*. 2nd ed. Geol. Soc. London.

菊池, 1991, *鉱物雑*, 20, 155-159.

Launeau et al., 1994, *Canad. Mineral.*, 32, 919-933.

Maloy and Treiman, 2007, *Amer. Mineral.*, 92, 1781-1788.

戸上ほか, 1998, *鉱物雑*, 27, 203-212.

Togami et al., 2000, *Canad. Mineral.*, 38, 1283-1294.

Tsuji et al., 2009, *Island Arc*, in press.

キーワード: 鉱物相, マッピング, 化学組成, 統計解析, ランダムフォレスト

Keywords: mineral facies, mapping, mineral chemistry, statistics, random forest