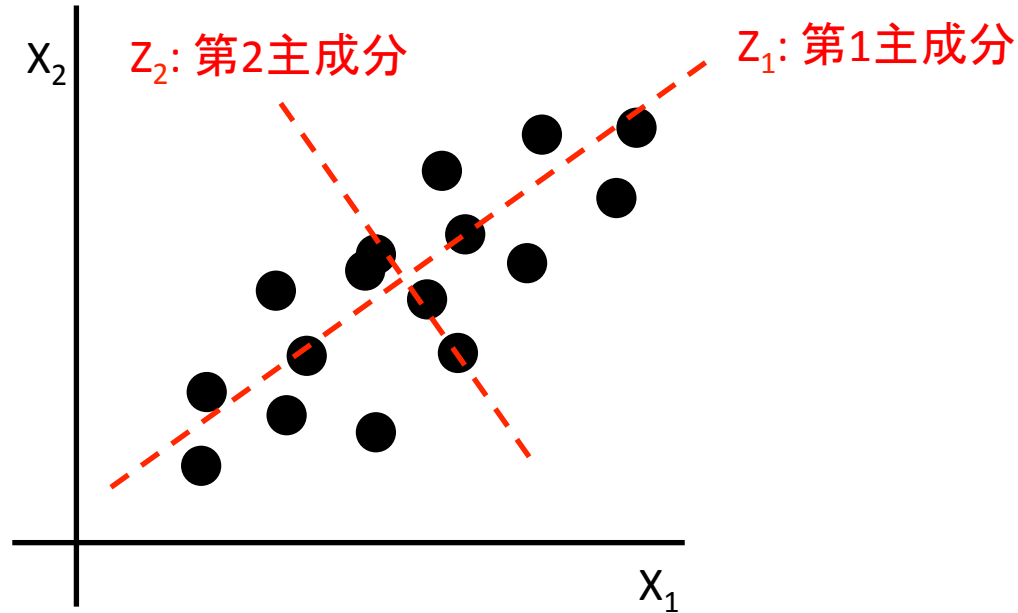


第7回惑星探査データ解析実習会
(2014年3月6日～8日)

主成分分析

東北大学 D1 石山謙

主成分分析とは？

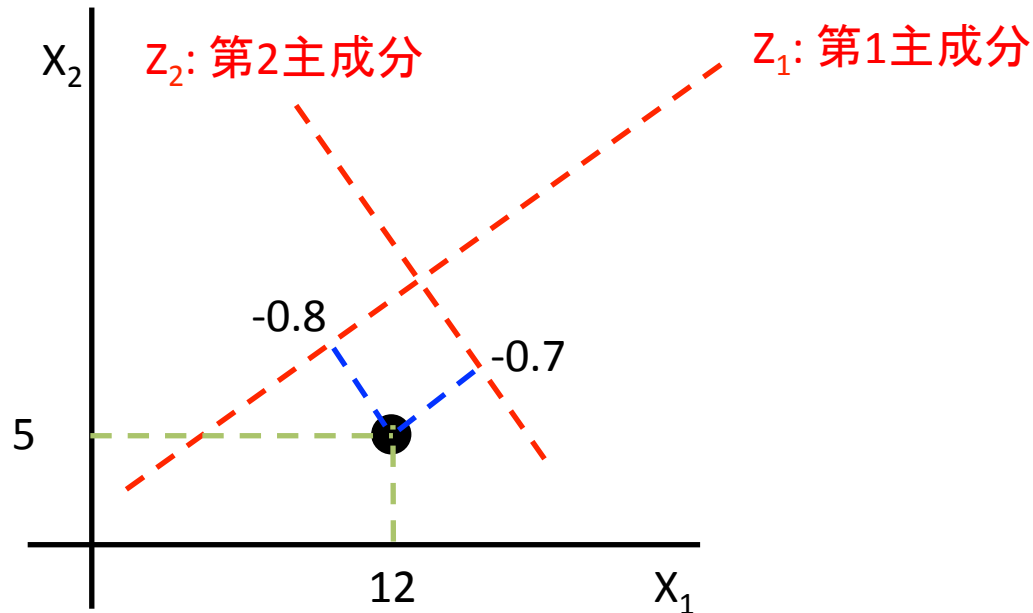


主成分分析とは、膨大な変数をもつデータからある傾向を捉える統計学的手法。例えば、上の図は、ある個体の X_1 と X_2 という変数を実験などで調べた結果とすると、ある方向に分散が大きいことがわかる。この方向は、 X_1 と X_2 という変数で書き表すことができ、

$$Z_1 = a_1 X_1 + a_2 X_2$$

となる。 Z_1 が**第1主成分**と呼ばれ、係数 a_1 と a_2 が**主成分負荷量**と呼ばれる。また、次に分散が大きくなる方向に軸をとり、第1主成分に直交するように第2主成分をすることができる。

主成分分析とは？



各主成分の主成分負荷量を決めた後、それぞれの測定値の主成分値(スコア)を求めることができる。例えば、上の図のある測定値(12, 5)の第1主成分スコアは、

$$-0.8 = a_1 \times 5 + a_2 \times 12$$

と計算できる [橋本さん資料]。

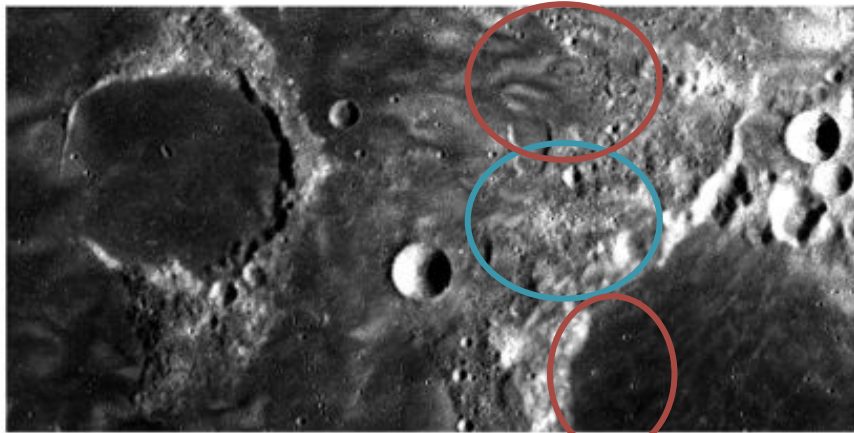
各主成分は、何らかの意味をもつが、この意味は、後付けでわかるものである。したがって、各主成分が何を意味するのかを調べる必要がある。

目的

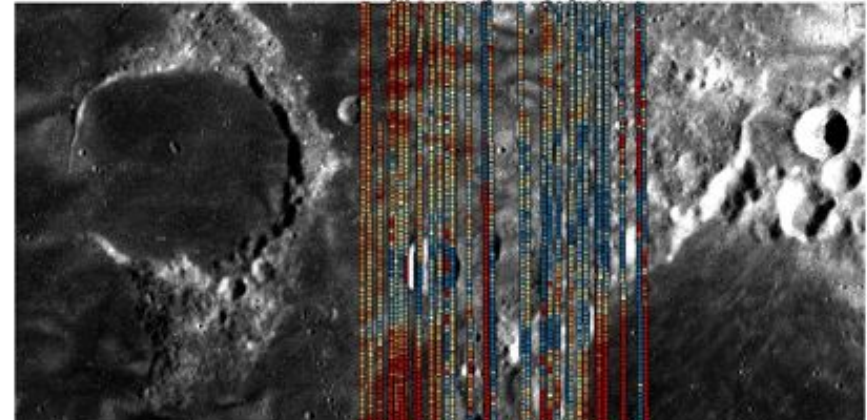
Mare Marginis(縁の海)においてSPにより測定された分光データを使用し、主成分分析を行い、主成分スコアを調べる。各主成分が何に依存するかを同定することは、月面の様々な情報を知る手がかりとなるため、非常に重要である。

したがって、ここでは、各主成分(第3主成分まで)が、どんなパラメータに依存するのかを調べる。

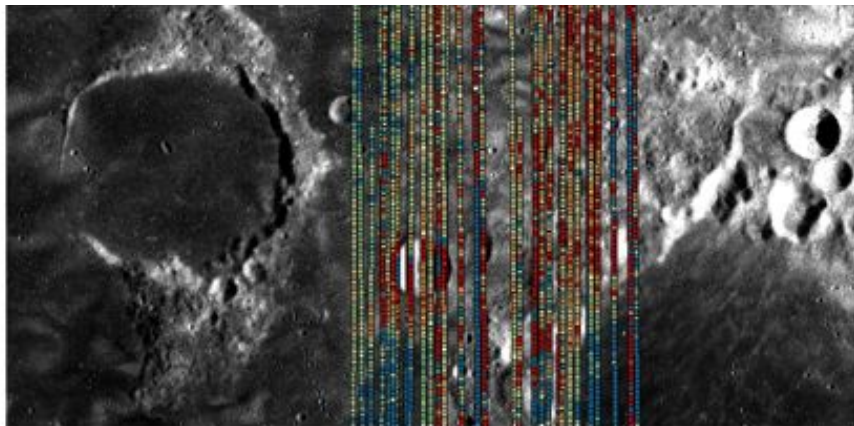
各主成分スコア分布 Mare Marginis



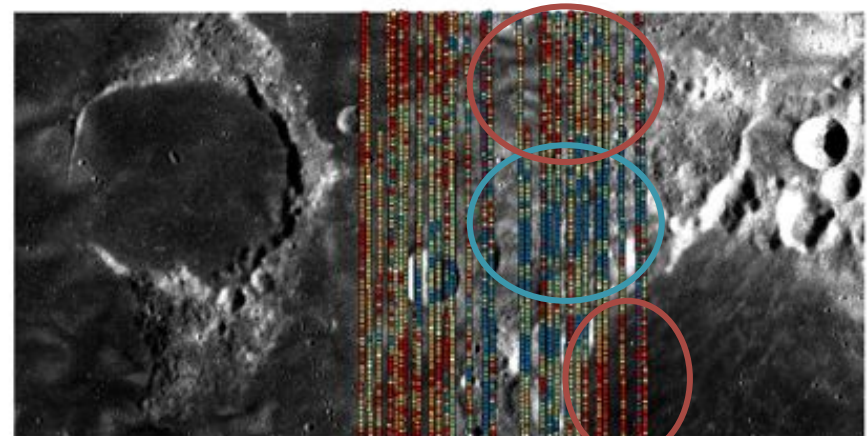
アルベド



第1主成分



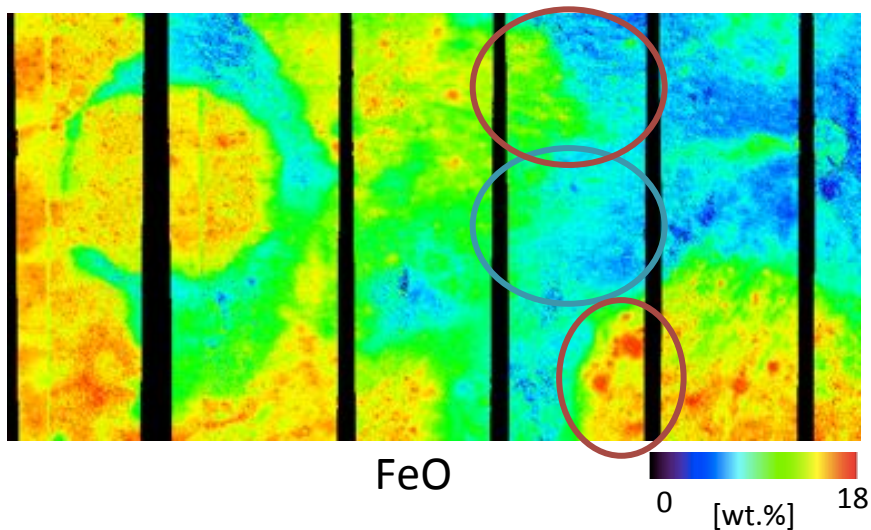
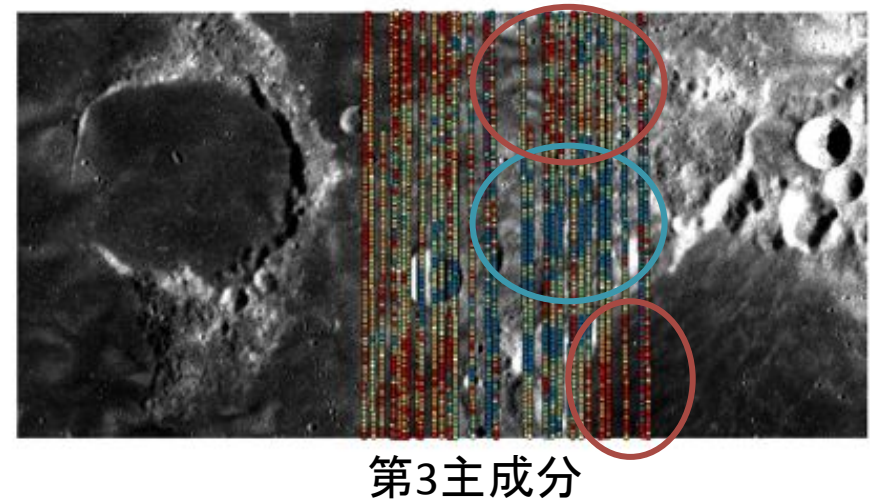
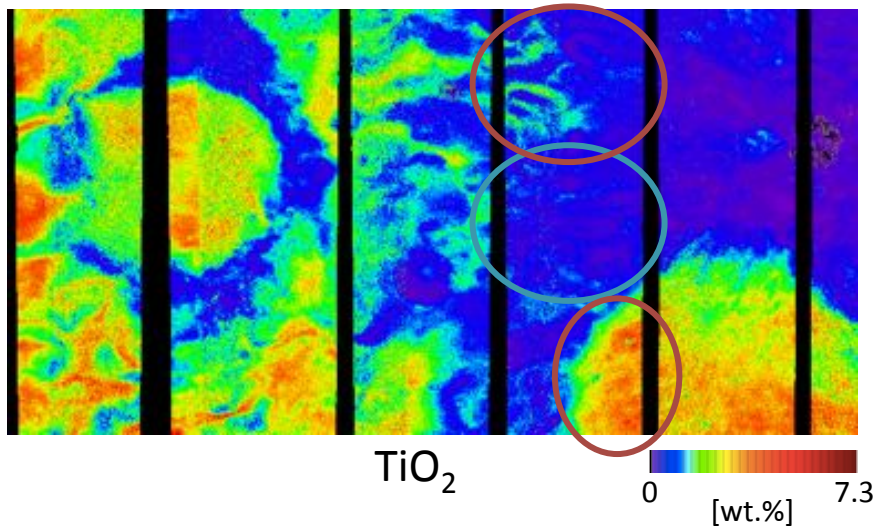
第2主成分



第3主成分

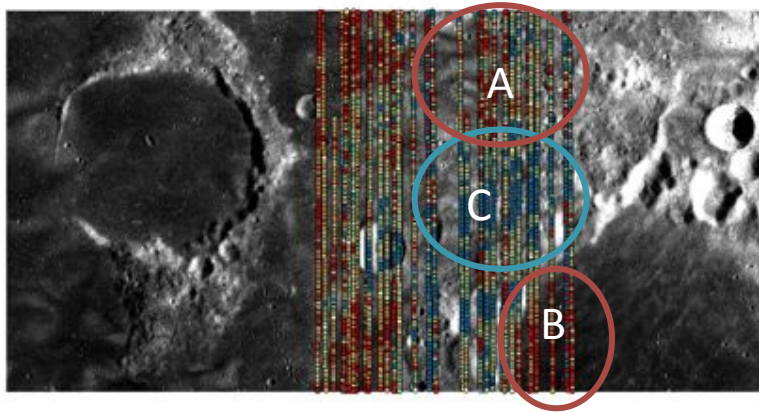
縁の海における各主成分ごとのスコア分布。スコアは、赤-青で大きさの違いを示してある。これらのデータは、アルベドマップ上に表示した。第1・2主成分スコアは、アルベドと相関があるように見えるが、第3主成分スコアは、アルベドと相関がないように見える(図中の赤円)。

各主成分スコア分布 Mare Marginis

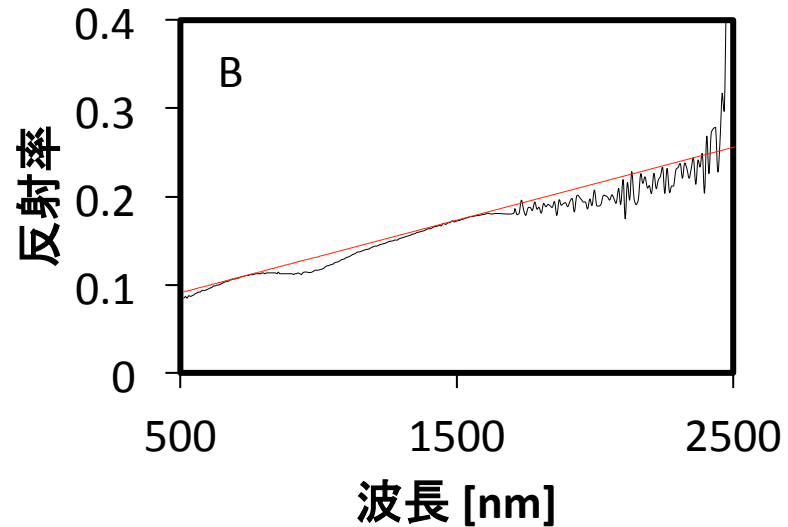
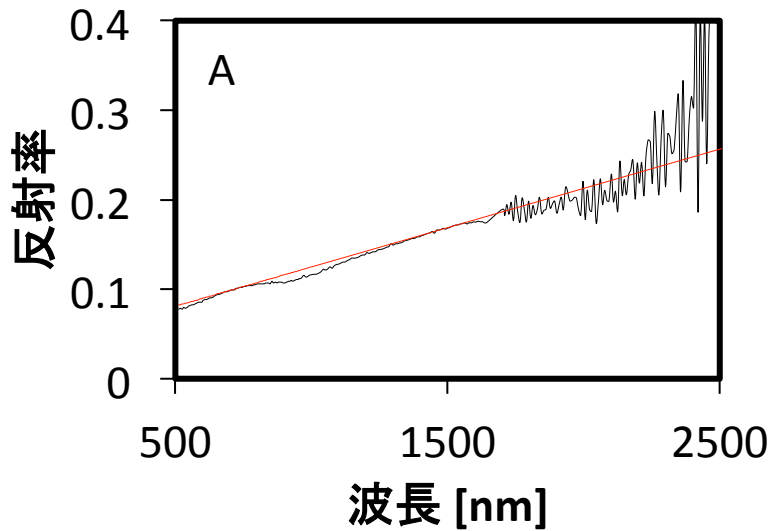
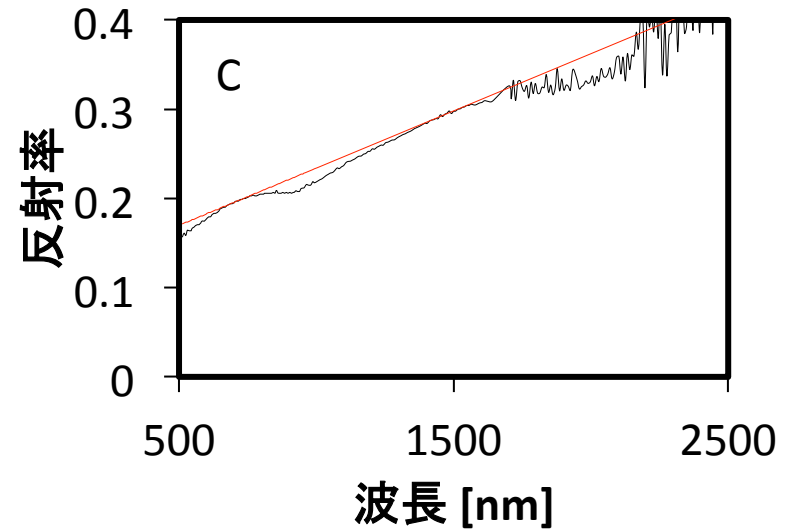


縁の海にチタンマップおよび鉄マップと第3主成分のスコア分布の比較。スコアは、赤-青で大きさの違いを示してある。このデータは、アルベドマップ上に表示した。鉄・チタンマップは、マルチバンドイメージャデータを使用し、Lucey et al. [2000]のアルゴリズムに基づき、作成された。第3主成分スコアは、チタンマップや鉄マップと相関があるに見える(図中の赤円)。

考察

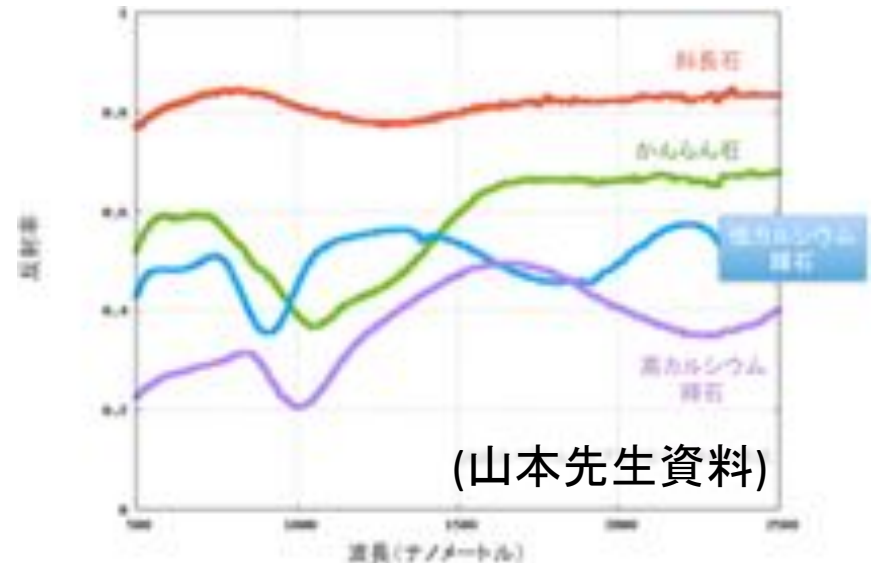
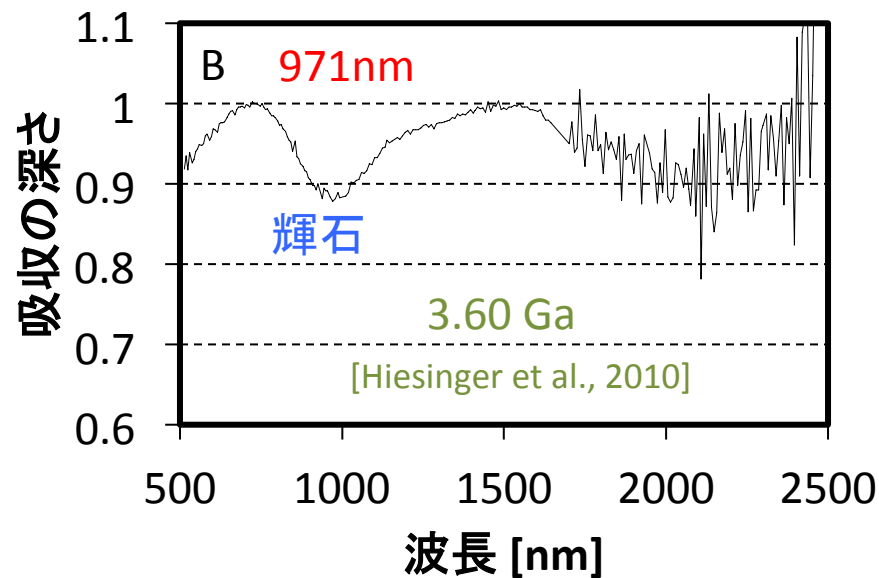
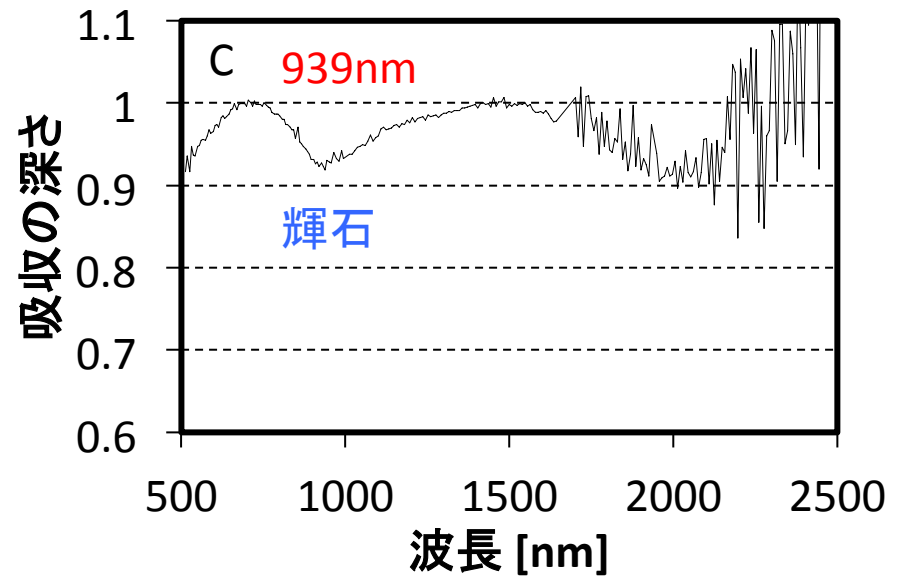
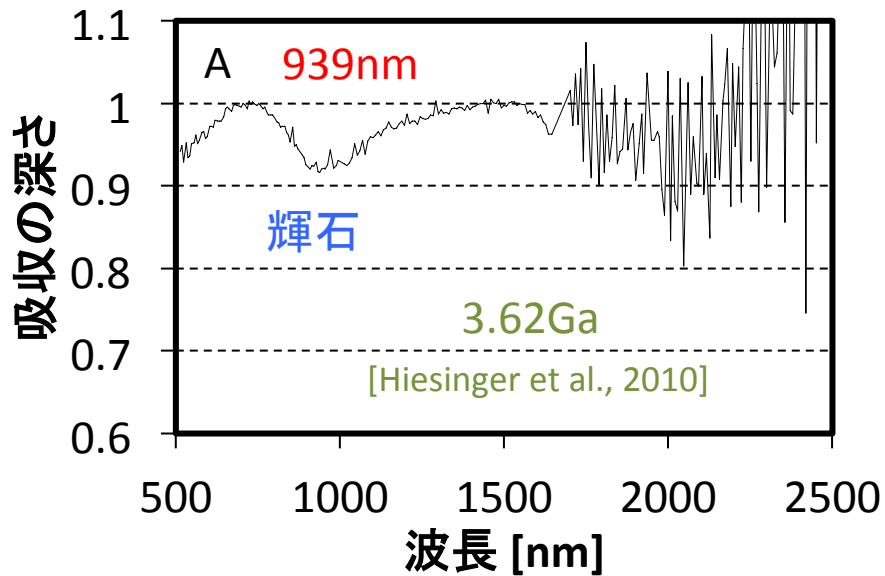


第3主成分



各領域(A, B, C)におけるスペクトルを調べた。各スペクトル図中の赤線は、コンティナムである。Cの領域は、AとBよりも全体的に反射率が高い。これは、この領域が高地であるためである。

考察



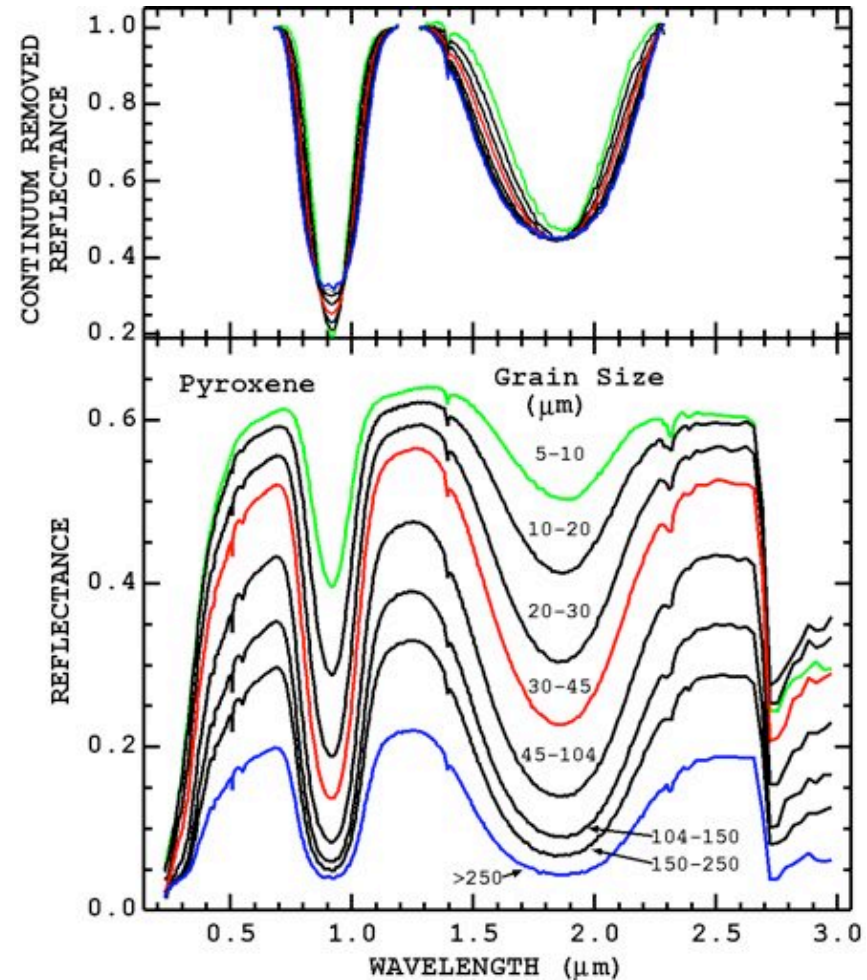
(山本先生資料)

(説明は次のスライド)

考察

各波長の反射率をコンティナムで割算して、吸収の深さを調べた。結果として、AとCの領域では939nmと約2000nm、B領域では971nmと約2150nmの波長で吸収の深さが最も深くなる(前のページ)。これらは、輝石の吸収帯に一致する(山本先生資料)。B領域で吸収波長が高波長側へシフトしている理由は、輝石中のカルシウムの含有量がAとCの領域よりも多く含まれているためだと考えられる。

また、AとBの領域では、900nm帯と2000nm帯における輝石の吸収の深さは、大体同じくらいである。これらの吸収の深さは、各領域の輝石のグレインサイズに依存しているかもしれない(右の図)。輝石のグレインサイズが大きくなると、900nm帯と~2000nm帯の吸収の深さが近づいている。AとBの領域では、250 μ m以上のグレインサイズだと、観測値を説明できるかもしれない。



グレイン(粒度)サイズに基づく輝石のスペクトル[<http://speclab.cr.usgs.gov/PAPERS/tetracorder/>].

考察

また、Cの領域では、900nm帯の吸収の深さよりも、2000nm帯における吸収の深さの方が深い。グレインサイズが大きくなると、2000nm帯の吸収帯の深さが深くなっていくため(前のページの図)、Cの領域では、AとB領域よりも**非常にグレインサイズが大きいならば、観測値を説明できないかもしれない**。Cの領域は、高地であり、高地領域の斜長岩は、マグマオーシャンから分化してできたものと考えられているため、その後の火山活動できた海領域(AとB領域)の玄武岩中の輝石のグレインサイズに違いがあるのかもしれない。海よりも高地の方が、輝石のグレインサイズが大きいならば、整合的である。

また、AとBの領域では、B領域の方が吸収の深さが深い。これは、宇宙風化に由来するものだと考えられる。AとBの領域は、Hiesinger et al. [2010]より表面年代が調べられており、A領域(3.62Ga)は、B領域(3.60Ga)よりも古い。**宇宙風化がおこると、吸収の深さは浅くなる**ことがわかっているため[Sasaki et al., 2001]、**A領域の方が古い**という点では、整合的である。しかし、AとB領域は0.02Gaしか差がなく、この差で吸収の深さが大きく変化するとは考えにくい。Hiesingerらは、クレーターカウンティング(第4回実習会参照)で表面年代を決めているが、彼らは、空間分解能が低い月面マップデータを使用しており、表面年代決定に大きな誤差を与えている。したがって、KAGUYA/TCデータなどの高分解能の月面地形データを使用すれば、AとB領域の絶対年代は、大きく違っているかもしれない。

まとめ

今回解析した縁の海領域では、第1・2主成分はアルベド、第3主成分は、チタンや鉄に相関がみられた。

今回解析した領域は、非常に限定的であったが、スペクトルも調べた結果、宇宙風化や、輝石中のカルシウムの含有量、輝石のグレインサイズの違いと思われるスペクトルの変化を同定することができた。したがって、各主成分は、今回した調べたアルベド、チタン、鉄マップ以外にも、様々なパラメータが依存している可能性がある。各主成分が何に依存するかを調べるには、様々なパラメータごとのグローバルなマップデータを作成し、比較していくことが必要だと考えられる。

謝辞

今回の実習会では、国立環境研究所の山本先生、会津大の小川先生に講師をして頂き、月の鉱物やSPデータの使い方について講義をして頂きました。非常に丁寧な説明で、最後までついていくことができました。また、世話人の方々には、実習会において、様々なサポートをして頂き、大変助かりました。この場を借りて、御礼申し上げます。