

■特集／“令和”時代の地下水の課題とその解決に向けた取り組み（その1）

自己組織化マップによる地下水水質の分類

On the use of Self-Organizing Maps to classify groundwater chemistry

キーワード：自己組織化マップ、多変量解析、硝酸性窒素、地下水汚染

中川 啓

Kei NAKAGAWA

長崎大学 総合生産科学域 教授



河村 明

Akira KAWAMURA

首都大学東京 大学院都市環境科学研究科 教授



1. はじめに

地下水水質の特徴を把握し、分類するためトリニアダイアグラムやヘキサダイアグラムが用いられることが多い¹⁾。これに加え、主成分分析やクラスター解析を組み合わせ、水質特性の分類を効果的に行った事例が多数報告されている²⁾。最近、多次元データの関連性を低次元（多くの場合は2次元であるが）のマップ上で表現することができる、自己組織化マップ（Self-Organizing Maps、以下 SOM とする）³⁾による水質特性のパターン分類事例が報告されるようになってきた。本稿では、こうした SOM の有用性を知ってもらうため、SOM を用いて分析を行った事例として、硝酸性窒素による地下水汚染が深刻である長崎県島原市の地下水水質を分類した事例⁴⁾と、中国のオルドス盆地の地下水水質形成機構について検討した事例⁵⁾について紹介する。

2. 方法

SOM は、Kohonen³⁾によって開発されたニューラルネットワーク手法の一種で、教師なしで多次元データの関連性を低次元（ここでは2次元）に写像する。本稿では、いずれの事例も、地下水をサンプリングした地点群を対象として、その水質特性に応じて分類することを目的とする。各地点で測定された成分濃度（イオンの濃度）をパラメータとするベクトルデータが入力データとなる。SOM 解析の結果として、マップ上に配置される六角格子を、ニューロンと呼ぶことにする。このニューロンの数を M とすると、入力データ数 n に基づき、経験的に M =

$5\sqrt{n}$ により求められる。またニューロンを配置する時の縦横の個数は、入力データの共分散行列における固有値ベクトルの第1、第2主成分の比率から決定できる。このマップが決まれば、入力データと同次元の参照ベクトルがニューロンに配置され、これが、入力データが内包する特性を反映するように、学習を繰り返して配置が更新される。最終的に、近接するニューロン同士は類似した特徴を持ち、遠隔のニューロンは異なる特徴を示す配置となる。入力データは最も特徴に近いニューロン内に分類される。

3. 事例1—硝酸性窒素による地下水汚染地域における事例⁴⁾

ここでは、長崎県島原市において、水道水源、民家、湧水、河川で観測された地下水および河川水の水質を分類された事例を示す。サンプリングした地点数は36地点であり、入力ベクトルの構成成分は、2012年4月から2015年5月にサンプリングされた、 Cl^- 、 NO_3^- 、 SO_4^{2-} 、 HCO_3^- 、 Na^+ 、 K^+ 、 Mg^{2+} 、 Ca^{2+} の主要イオン濃度8成分である。解析に供した総データ数は、353である。

図-1にSOM解析によって得られた成分ごとの結果を示す。91個の参照ベクトルに分類されており、その濃淡は、0-1の範囲の基準化された濃度を示している（濃い色は濃度が高いことを示す）。ここでは8成分、それぞれについて示しているが、同じ参照ベクトルには同じ地点が配置されている。視覚的にすぐ理解できることは、 Cl^- と NO_3^- が類似していることである。左上が薄く、右下が濃い分布を示すといった点で、 Na^+ と Mg^{2+} も類似していると言えよう。こ

れらはそれぞれ相関が高いことを意味している。すべての成分で左上端のニューロンが低濃度を示すものとなっている。 HCO_3^- 、 Mg^{2+} 、 Ca^{2+} については、右上端のニューロンが示すように、中程度の濃度を示すことが分かる。 HCO_3^- を除く他の成分が、左下端のニューロンでは高濃度を示している。 Cl^- と NO_3^- を除く右下端のニューロンで、比較的高い濃度を示していることが分かる。

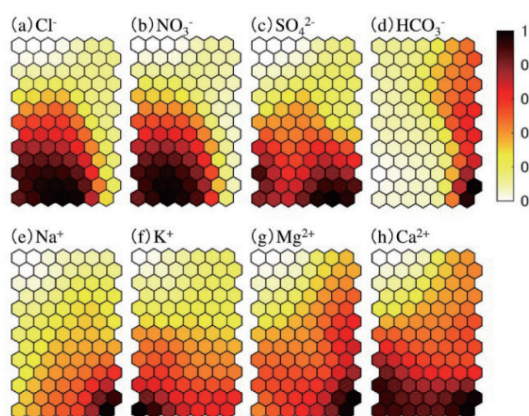


図-1

SOM 解析によって、353 のデータが 91 のグループに分けられたことになるが、その特性を把握するにはまだ若干多すぎる。そこで、さらにこの 91 のグループを階層的クラスタ解析により分類を行った。図-2 に、そのようにして分類した結果を示す。この場合、5つのクラスターに分けられた。各ニューロン内の数字は、そこに分類された入力データの数を示している。図-1 と対比しながら各クラスターの特性を見ることができる。例えば、クラスター3に分類されたデータは、 Cl^- と NO_3^- の濃度が高いグループであることが分かる。 NO_3^- だけに注目すると、クラスター2と3の濃度が高く、これらが硝酸性窒素汚染の影響を受けたグループであると言える。

図-3 には、分類された5つのクラスターの空間的な分布を示したものである。研究対象地の北部に広く分布している農地と対応して、汚染クラスターである2、3が分布している。全ての成分で低濃度を示すクラスター1が山側に分布していることや、比較的高い濃度が高めのクラスター4、5は低標高の市街地に分布していることなども明らかである。

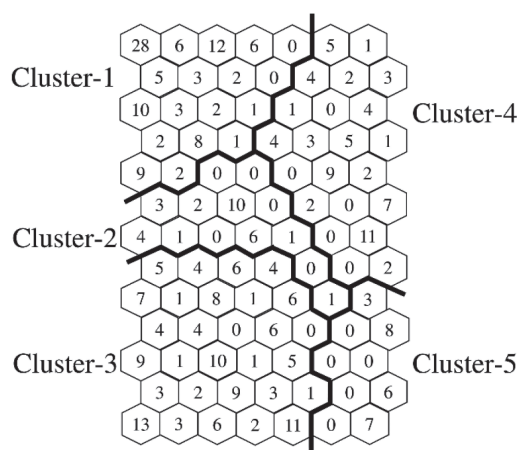


図-2

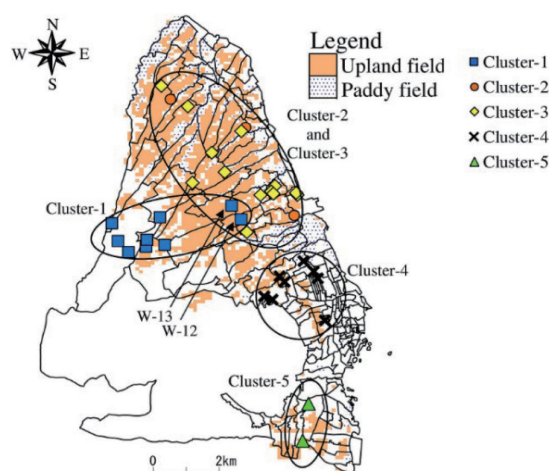


図-3

4. 事例 2—中国オルドス盆地における地下水水質形成過程の検討事例⁵⁾

二つ目の事例として、中国北西部における堆積層であるオルドス盆地においてサンプリングされた地下水の特性を分類し、水質進化特性を検討したものを紹介する。ドーシット川流域の 20 地点で、2012 年から 2013 年にサンプリングされた地下水を対象としている。解析の対象とした成分は、 Na^+ 、 K^+ 、 Ca^{2+} 、 Mg^{2+} 、 Cl^- 、 SO_4^{2-} 、 HCO_3^- 、 CO_3^{2-} の 8 成分である。分類に供したサンプル数は、各地点について 1 サンプルとなる 20 である。SOM 解析に先立ち、一般的な水質特性について検討するため、トリリニアダイアグラム上に 20 サンプルをプロットした (図-4)。陽イオンについては Na^+ の含有率が高く、陰イオンについては Cl^- の含有率が高く、 SO_4^{2-} の含有率も比較的高いといった特徴があり、全てのサンプルが IV の領域にプロットされた。ここにプロットされ

るサンプルは、アルカリ非炭酸塩型であり、わが国では海水や温泉で多くみられる型であることが知られている。一般の浅層地下水は、Iの領域のアルカリ土類炭酸塩型にプロットされる¹⁾。

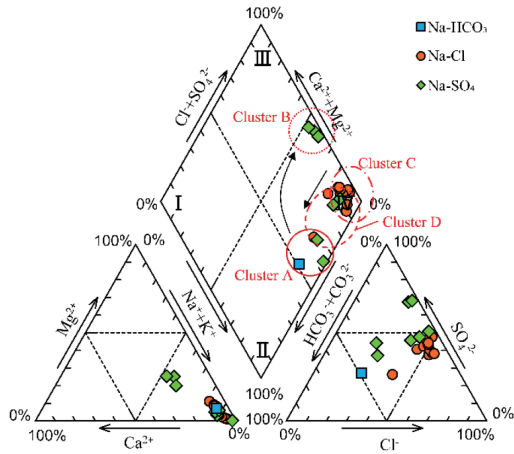


図-4

この研究では、主成分分析と階層的クラスター解析を行っている。その結果、主成分は、第2主成分までで77.9%が説明され、第1主成分は、硫酸塩物(gypsumとanhydrite)の溶解、第2主成分は、炭酸塩鉱物の溶解を説明すると考えられた。クラスター解析によると、4つのクラスターに分かれた。図-4の中にそれらのクラスターも示している。涵養域から流出域に沿った、陰イオン進化は、 $\text{HCO}_3^- \rightarrow \text{HCO}_3^- + \text{SO}_4^{2-} \rightarrow \text{SO}_4^{2-} + \text{Cl}^- \rightarrow \text{Cl}^-$ とすれば、クラスターAからBへの進化経路(第1主成分が増加する方向)とクラスターCからDへの進化経路(第2主成分が増加する方向)が考えられる。図-5に、それらの分布を地図上にプロットしたものを示す。クラスターAからBに至る進化は、河川の上流から下流への流下に伴うものと説明できる。クラスターCからDの進化は、河川から遠いところから河川に向かう方向への水の移動に伴うものと説明できよう。

この事例では、SOM解析の結果、24の参照ベクトルに分類された。分析に供したのは20サンプルと僅かであったため、生データが配置されないニューロンを生じていることが、SOM解析の後、クラスター解析を行った結果の図-6からも分かる。SOM解析後のクラスター解析では、4つのクラスターへ分類された。このクラスター、1から4は、SOM解析を行う前に

解析を行った場合のクラスター、AからDと一致している。各クラスターの特徴を調べるため、クラスターごとの成分について、基準化された濃度の平均値や第1四分値および第3四分値をレーダーチャートにプロットしたものを図-7に示す。グループ1は、比較的、 HCO_3^- の濃度が高いことが分かる。グループ2は、全ての陽イオンと SO_4^{2-} と Cl^- の濃度が高いグループである。グループ3は、 Na^+ と Cl^- の濃度が高い。グループ4は、 Na^+ 、 K^+ 、 Cl^- 、 HCO_3^- 、 CO_3^{2-} の濃度が高いグループであり、これら4つのクラスターが明らかに異なる性質を持つものであることが分かる。実際の値に再変換したところ、グループ1のみが、 Cl^- と SO_4^{2-} についての中国の水道水質基準を満たしていることが分かった。

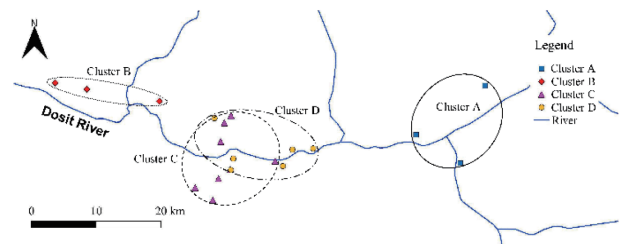


図-5

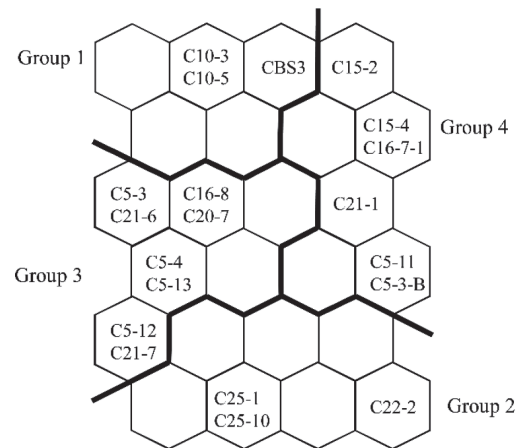


図-6

5. 多変量解析としてのSOMの適用性について

事例2では、階層的クラスター解析を、生データに適用した結果と、SOM解析を行った後に、さらに適用した結果が、一致していた。解析に供したデータの数が少ないため、このような結果となったと考えられるが、主成分分析や階層的クラスター解析を適用した場合

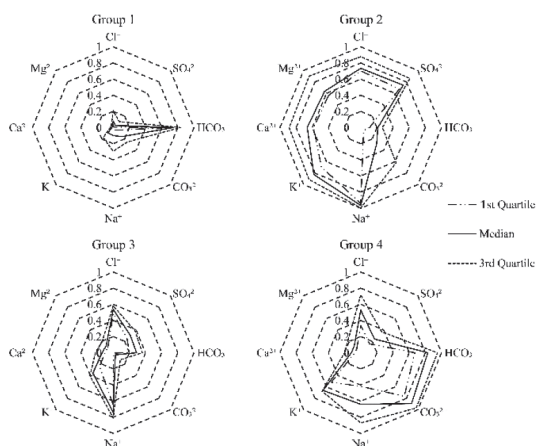


図-7

と、SOM 解析後に階層的クラスター解析を適用した場合、それぞれの利点はどのようなものだろうか。この疑問に答えるため、事例1で検討した長崎県島原市の水質測定結果（この時のデータは、2011年8月から2013年11月の測定）について、(1) 階層的クラスター解析のみによる分類、(2) 主成分分析と階層的クラスター解析による分類、(3) SOMと階層的クラスター解析による分類、(4) 主成分分析およびSOMと階層的クラスター解析による分類の4種類の手法による結果を比較した⁶⁾。比較を容易にするため、クラスター数を統一し、5グループへ分類する場合と10グループへ分類する場合について検討している。5グループへ分類した場合は、各クラスターの持つ意味はいずれの手法もほぼ同様であり、大別すると、汚染の影響を受けていないクラスターと汚染クラスターに分類することができた。クラスターの地図上の分布と硝酸性窒素による汚染分布状況の比較から検討すると、主成分分析を利用した手法(2)および(4)が水質特性を説明するうえで適していると判断された。10グループへ分類した場合は、汚染クラスター、準汚染クラスター、汚染の影響を受けていないクラスターに大別することができた。クラスターの数が増えたため各クラスターの地図上の分布がそれぞれの手法で異なっていたが、硝酸性窒素濃度分布との比較からは、SOMを適用しない手法(1)および(2)が島原市の水質を説明するうえで適していると判断された。以上より、手法(2)が、この事例の場合は、適した手法であったと考えられ

た。しかしながら、実質的な分類結果は大きくは相違していなかったため、手法(3)および(4)で用いたSOM解析は、事例1や事例2で示した図-1、2および図-6のように視覚的にデータ間の関係性を表現することが可能であると言える。そのため、細かい分類精度だけでなく、目的に応じた手法の選択も重要と言えよう。

6. おわりに

本稿では、地下水水質を分類するために自己組織化マップを適用した事例を紹介した。地下水汚染といった水環境汚染の対策を検討する場合には、どこも一律に対策を講じるのではなく、個々の特性に応じた対応を取ることが効率的な場合がある。あまり難しいことを考えずに入力したデータに基づき、ある程度、妥当性のある分類が可能な自己組織化マップは、そのような分類を行うための一手法として適用可能なものと考えられる。

<参考文献>

- 1) 中川 啓 (2019) 地下水環境, 水理公式集 [2018年版], 土木学会水工学委員会水理公式集編集小委員会, 土木学会, 丸善出版, pp.860-867.
- 2) たとえば, Nakagawa, K., Amano, H., Asakura, H., and Berndtsson, R. (2016). Spatial trends of nitrate pollution and groundwater chemistry in Shimabara, Nagasaki, Japan, *Environmental Earth Sciences*, 75 : 234.
- 3) Kohonen, T. (1990). Self-Organizing Maps, *Proceedings of The IEEE*, 78(9), pp.1464-1480.
- 4) Nakagawa, K., Amano, H., Kawamura, A., and Berndtsson, R. (2017). Classification of groundwater chemistry in Shimabara, using self-organizing maps, *Hydrology Research*, 48(3), pp.840-850.
- 5) Yu, Z.-Q., Amano, H., Nakagawa, K., and Berndtsson, R. (2018). Hydrogeochemical evolution of groundwater in a Quaternary sediment and Cretaceous sandstone unconfined aquifer in Northwestern China, *Environmental Earth Sciences*, 77 : 629.
- 6) 天野弘基, 中川 啓, 河村 明(2016). 多変量解析による地下水水質の分類特性-島原市における事例-, 土木学会論文集G (環境), 72(5), pp.I_127-I_135.