

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 —

天野 弘基¹・中川 啓²・河村 明³

¹学生会員 長崎大学大学院水産・環境科学総合研究科 (〒852-8521 長崎市文教町 1-14)

²正会員 博(工) 長崎大学教授 大学院水産・環境科学総合研究科 (〒852-8521 長崎市文教町 1-14)

E-mail: kei-naka@nagasaki-u.ac.jp

³正会員 工博 首都大学東京教授 大学院都市環境科学研究科 (〒192-0397 東京都八王子南大沢 1-1)

島原市の地下水水質を対象として、4種類の多変量解析手法を適用し、どの分類結果が当該地域の水質特性を説明するために適しているかについて検討した。5つのクラスターに分類した場合のそれぞれ特徴は、ほぼ同様であったが、各クラスターの空間分布については手法間に違いが認められ、主成分分析を利用した手法は、硝酸性窒素濃度が比較的高い採水地点を、汚染クラスターに分類した。クラスター数を10とした場合においても主成分分析を利用した手法がより適切な結果を与えうることが明らかとなった。

Key Words: *water quality, nitrate pollution, principal component analysis, self-organizing map, hierarchical cluster analysis*

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),2)}。一方、多変量解析手法である主成分分析を適用すると、陽イオン交換反応、蒸発、塩水侵入、硝酸性窒素汚染や風化といった地下水組成に影響を与える因子を抽出することができる。ここで得られる主成分得点により、抽出した因子の空間的拡がりの特徴を把握することも可能である^{3),4)}。さらに、クラスター分析と組み合わせることにより、水質特性の分類を効果的に行うことができ、多くの事例が報告されている⁵⁾⁻⁸⁾。たとえば沖縄本島南部地域の事例では、主成分分析により海水や海塩の影響、施肥や生活排水の影響、石灰岩の溶解、酸化還元状態等の因子を抽出し、さらにクラスター分析により調査地域の地下水水質を4つのクラスターに分類し、それぞれの影響因子を評価した⁶⁾。また、ナイジェリア南東部エヌグの事例では、主成分分析を行い、母岩鉱物の風化と人為的活動の影響の因子を抽出した。さらに、クラスター分析により地下水水質を汚染負荷レベルの異なるグループに分類した⁸⁾。

最近、主成分分析やクラスター分析に加え、自己組織化マップ (Self-organizing map, 以下 SOM とす

る) を利用した水質特性のパターン分類が報告されるようになってきた^{9),10)}。SOM では多次元データの関連性を2次元マップ上で表現することができ、たとえばベトナム紅河デルタの事例では、地下水中の Mg^{2+} , Na^+ , K^+ そして Cl^- が SOM 上で類似したパターンを示したことから、これらのイオンの関連性が明らかとなった⁹⁾。さらに、クラスター分析を適用し、塩分濃度により特徴付けられる8つのクラスターに地下水水質を分類した。

以上のように、地下水水質特性を分類するために、いくつかの多変量解析手法を組み合わせられてきたが、どのような組み合わせ方が研究対象地の水質特性をより適切に説明できるかどうかの議論は行われていない。そこで、本研究では長崎県島原市の地下水水質データを分析対象とし、(i) 階層的クラスター分析のみによる分類、(ii) 主成分分析と階層的クラスター分析による分類、(iii) SOM と階層的クラスター分析による分類、(iv) 主成分分析および SOM と階層的クラスター分析による分類の4種類の多変量解析手法を適用し、分類結果を比較することにより、それぞれの分類特性とどの手法が島原市の地下水水質特性を説明する上で適しているかについて検討した。

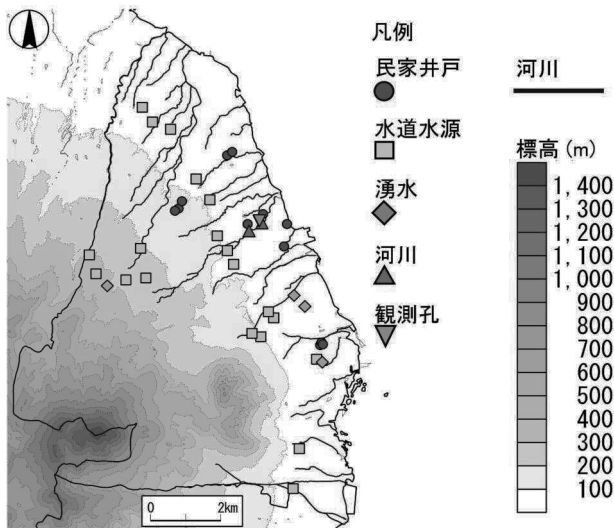


図-1 採水地点の分布¹³⁾

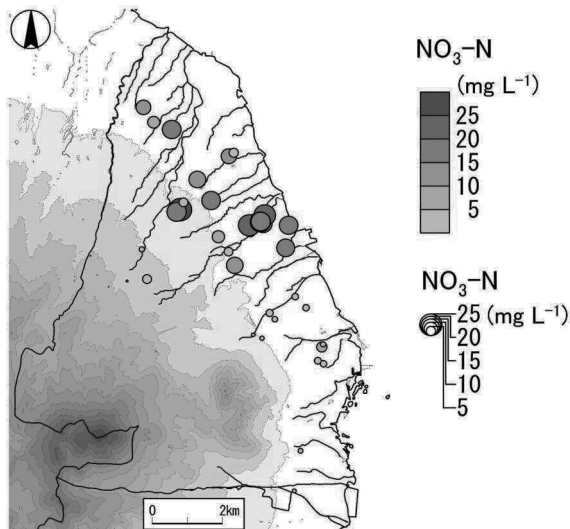


図-2 硝酸性窒素濃度の分布¹³⁾

2. 研究対象地の概要¹¹⁾

研究対象地とした島原市は島原半島の北東部に位置し、その18%を占める面積は82.8km²である。地形は半島中心にそびえる普賢岳より東の有明海にかけてなだらかな斜面が広がる扇状地である。標高300~1500mに森林が広がり、島原市北部の山の中腹では畜産施設や集落が点在し始める。標高が低くなるにつれ、農地があたり一面を占め、道路沿いに民家が建ち並ぶ。市北部の標高10m以下では海岸線近くに民家が建ち並び、南部の標高50m以下では市街地や商業施設が発達する。農業地帯は島原市の北部に集中し、島原市全体の22.7%を占める。農作物の白菜、人参、大根、すいか、ほうれんそうは長崎県内シェアの4分の1以上を占める。島原市ではこれらの作物を中心に二毛作により四季を問わず、年間を通して農作物が栽培されている。一方、畜産にお

いては乳用牛1,552頭、肉用牛3,641頭、豚54,820頭、採卵鶏1,346千羽、肉用鶏92千羽が飼育されており(2009年)、長崎県内シェア率は、採卵鶏の65.4%が最も高く、次いで豚の23.3%が高い。以上のように県内有数の農畜産施設であることがわかる。1977年から2013年における年平均降水量は2105mm、年平均気温は17.3℃であった¹²⁾。

3. 硝酸性窒素汚染と地下水水質¹³⁾

研究対象地における2011年8月から2013年11月に実施した地下水水質調査では、40地点中(図-1)15地点で硝酸性窒素濃度が、硝酸性窒素及び亜硝酸性窒素の水質基準¹⁴⁾および環境基準¹⁵⁾である10mgL⁻¹を超過し、40地点中の21地点では、人為的活動の影響を受けているとされる閾値¹⁶⁾3mgL⁻¹を超過していた(図-2)。高い硝酸性窒素濃度(>10mgL⁻¹)は、主に島原市北部の標高200m未満で観測されており、硝酸性窒素濃度の空間分布と土地利用を比較すると、高い硝酸性窒素濃度は農業地域の空間分布とよく一致した。市北部の標高200m以上に分布する森林部や、市南部の標高100m以下に発達する市街地および郊外の採水地点における硝酸性窒素濃度は、基準値を満たすことから、研究対象地では硝酸性窒素汚染と農業活動が密接に関係していることが示唆された。

水質組成については、硝酸性窒素が基準値を超過するような採水地点はCa-(SO₄+NO₃)型に分類され、そうでない地点はCa-HCO₃型に分類された(図-3)。標高300~400mに位置する採水地点の溶存イオン濃度は、低い値を示し、標高が低くなるにつれ溶存イオン濃度は上昇する傾向が認められ、黒雲母や角閃石、斜長石、かんらん石の溶解が、K⁺やMg²⁺、Ca²⁺濃度の増加に寄与している可能性が示唆された。

4. 多変量解析

(1) 解析方法

本研究では、階層的クラスター分析のみによる分類、主成分分析と階層的クラスター分析による分類、SOMと階層的クラスター分析による分類、主成分分析およびSOMと階層的クラスター分析による分類の4種類の多変量解析手法の比較、検討を実施した。検討対象としたのは、2011年8月から2013年11月に40地点(図-1)で集めた277サンプルのイオン濃

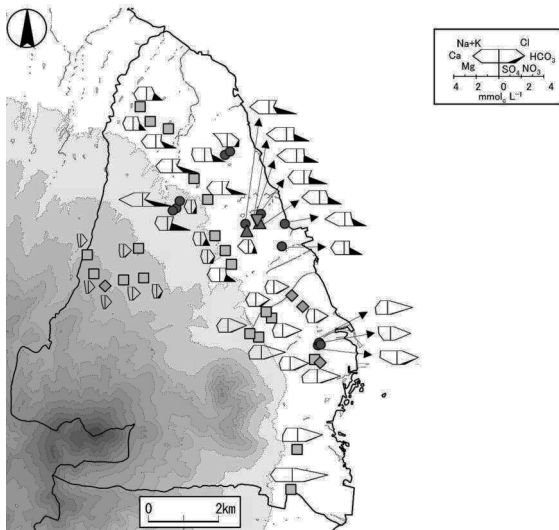


図-3 水質組成（ヘキサダイアグラム）の分布¹³⁾

度データ (Cl^- , NO_3^- , SO_4^{2-} , HCO_3^- , Na^+ , K^+ , Mg^{2+} , Ca^{2+}) である¹³⁾。民家井戸と水道水源では深井戸および浅井戸で、およそ2ヶ月の間隔でサンプリングを実施している（深井戸と浅井戸は深度30mで分類されている）。硝酸性窒素濃度は、降雨時の希釈などによる解析期間中の変動がみられる¹³⁾。階層的クラスタ分析にはWard法を採用した。クラスタ数は、DBI (Davies-Bouldin Index)¹⁷⁾に基づき決定した。クラスタ数を2から入力データ総数まで変化させ、DBI値を算出し、最小となる場合が最適クラスタ数である。DBI値は元のデータを用いた場合、主成分分析の後に求める場合、SOM解析を行った後に求める場合で異なるが、本研究では、4種類の手法間の比較検討を行うため、後の手法3、すなわち主要溶存イオン8成分の濃度に基づきSOM解析の後に決定された最適クラスタ数5を共通して採用した。さらに、クラスタ数を多く設定し、詳細に分類する場合の特性比較のため、2番目の最適クラスタ数10とした場合も検討に加えた。4種類の手法による分析結果は、各クラスタの特徴を把握するため、クラスタごとに主要溶存イオン濃度の平均を算出し、クラスタの研究対象地上の空間分布を描画し、硝酸性窒素濃度および水質組成の分布と対比しながら手法ごとに検討を加えた。

a) 階層的クラスタ分析¹³⁾

主要溶存イオン8成分の濃度（8次元277ベクトル）を入力データとして階層的クラスタ分析（HCA）のみを実行し、分類した。以降、「手法1」とする。

b) 主成分分析と階層的クラスタ分析

まず、主要溶存イオン8成分の濃度に基づき主成分分析（PCA）を実行し、主成分分析より得られた第

1, 第2主成分得点（2次元277ベクトル）を入力データとしてHCAを実行し、分類した。以降、「手法2」とする。なお、PCAは相関係数行列を用いて実行し、第1, 第2主成分の寄与率はそれぞれ59.7%, 26.8%で、累積寄与率は86.5%となる。

c) SOMと階層的クラスタ分析¹⁸⁾

初めに、主要溶存イオン8成分の濃度に基づきSOMを作成した後、得られた84の参照ベクトル（8次元84ベクトル）を入力データとしてHCAを実行し、分類した。以降、「手法3」とする。

d) 主成分分析およびSOMと階層的クラスタ分析

手法2と同様に、まず主要溶存イオン8成分の濃度に基づき主成分分析を実行し、第1, 第2主成分得点（2次元277ベクトル）を得た。次に主成分分析で得た第1, 第2主成分得点（2次元277ベクトル）に基づきSOMを作成した。最後に、SOMで得られた81の参照ベクトル（2次元81ベクトル）を入力データとしてHCAを実行し、分類した。以降、「手法4」とする。

(2) 自己組織化マップ (SOM)

SOMは、Kohonen¹⁹⁾により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2次元に写像する。この機能により、多次元データから抽出したパターンを2次元マップ上に視覚化し比較することができる。著者らによる同じ研究対象地における2012年4月から2015年8月の水質データにSOMを適用した事例では、 Cl^- と NO_3^- , Na^+ と Mg^{2+} のSOMが類似したパターンを示し、これらのイオンの相関の高さが視覚的に判断できた¹⁸⁾。SOMは、ニューロンまたはノードと呼ばれる六角格子が規則正しく配列することで構成される。入力データの特徴を適切に抽出するために、経験則 $M=5\sqrt{n}$ (M :ニューロン数, n :入力データ数)により総ニューロン数は決定され、また縦と横のニューロン数の比は、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および第2成分の比率より決定される^{17), 20)}。本研究では、これらの経験則に従い、手法3では縦12×横7の総ニューロン数84、手法4では縦9×横9の総ニューロン数81のSOMをそれぞれ構築した（図-4, 5）。SOM構築後、入力データと同次元の参照ベクトル（手法3では8次元、手法4では2次元）が各ニューロンに配置され、これらの参照ベクトルは入力データの特徴を反映するように、学習プロセスを通して繰り返し更新される。結果として、近接するニューロンは互いに類似した特徴を持ち、遠隔のニューロンはそれぞれ

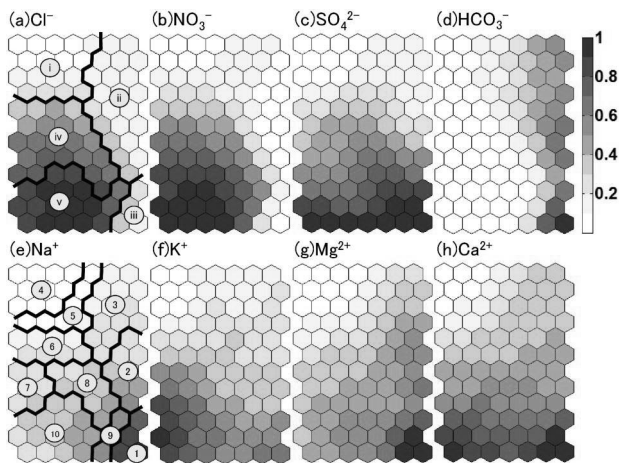


図-4 手法3により作成した主要溶存イオン8成分のSOM

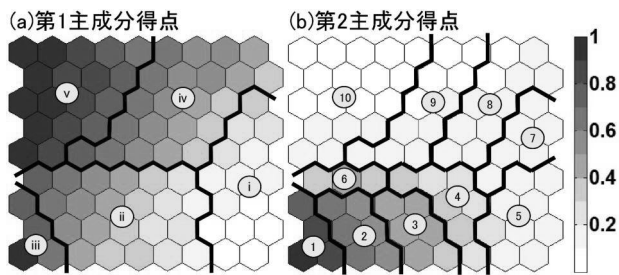


図-5 手法4により作成した主成分得点のSOM

異なる特徴を示す。最終的に、入力データは、性質が最も類似した（ユークリッド距離が最短）ニューロンに分類される。

図-4に手法3により作成したSOMを示す。黒の実線はクラスターの数を5とした場合とクラスターの数を10とした場合のクラスター境界を表し、SOM上のそれぞれの数字はクラスターの番号を表す。Nakagawa et al.¹⁸⁾が報告した結果と同様に、 Cl^- と NO_3^- 、 Na^+ と Mg^{2+} のSOMのパターンが類似しており、これらのイオンの相関の高さが窺える。図-5に手法4によるSOMを示す。Nakagawa et al.¹³⁾によると、第1主成分はイオンの溶解と硝酸性窒素汚染、第2主成分はイオンの溶解を表しているため、両者のSOMのパターンは明確に異なる。あとの分類結果については、クラスターの数を5（クラスターi~v）とした場合とクラスターの数を10（クラスター1~10）とした場合、それぞれで検討しているが、両者の対応関係は以下のようなものである。手法3では、クラスター4と5がクラスターiである。クラスター2と3がクラスターii、クラスター1がクラスターiii、クラスター6、7および8がクラスターiv、そしてクラスター9と10がクラスターvである。手法4では、クラスター5と7がクラスターi、クラスター2、3、4および6がクラスターii、クラスター1がクラスターiii、クラスター8と9がクラスターiv、ク

ラスター10がクラスターvに対応する。

5. 4つの多変量解析手法による水質特性の分類

ここでは、クラスターの数を5とした場合とクラスターの数を10とした場合のそれぞれの解析結果の違いについて議論する。

(1) クラスターの数を5とした場合

表-1に、各クラスターにおける主要溶存イオンと硝酸性窒素（ $\text{NO}_3\text{-N}$ ）の平均濃度をそれぞれの解析手法ごとに示す。いずれの手法においても5つのクラスターは、硝酸性窒素濃度が人為的活動の影響を受けているとされる閾値¹⁶⁾ 3 mg L^{-1} を超えないクラスターi、ii、iii（以下、汚染の影響を受けていないクラスターとする）と平均の硝酸性窒素濃度が硝酸性窒素及び亜硝酸性窒素の水道水質基準¹⁴⁾および環境基準¹⁵⁾ 10 mg L^{-1} を超過するクラスターiv、v（以下、汚染クラスターとする）に分けることができる。なお、HCAによるクラスター分類は、デンドログラムによると、手法3のみi、iiとiii、iv、vに大別されたが、それ以外の手法ではi、ii、iiiとiv、vに大別された。汚染の影響を受けていない3つのクラスターを比較すると、 NO_3^- を除き、各種イオン濃度に違いが認められる。いずれの手法においてもクラスターiiiが最も溶存イオン濃度が高く、ついでクラスターii、そしてクラスターiが低い。2つの汚染クラスターも同様に、汚染レベルで区別できる。クラスターivの硝酸性窒素濃度は $11.4\text{--}13.4 \text{ mg L}^{-1}$ であり、水道水質基準¹⁴⁾および環境基準¹⁵⁾よりわずかに高い値を示しているが、クラスターvの硝酸性窒素濃度はクラスターivの1.5倍より大きい値（ $19.2\text{--}20.5 \text{ mg L}^{-1}$ ）を示した。

図-6に各クラスターの分布を解析手法ごとに示す。なお、各地点とも観測期間最新の分析値によるクラスターをプロットした。クラスターiの分布に着目すると、A、B、C点（図-6）を除き、全ての解析手法において同じ分布を示した。A点は、手法1と3では、クラスターiに分類され（図-6（a）、（c））、B、C点は、手法3でのみ、クラスターiに分類された（図-6（c））。クラスターiiは、市南部の標高100m以下に分布しており、B、C点を除き、全ての解析手法で同じ分布を示した。クラスターiiiに分類された採水地点は、全ての解析手法において市の最南部に位置する地点のみである。クラスターivとvは市の北部に分布しており、解析手法により分布が異なる。しかし、どちらのクラスターも硝酸性窒素濃度

表-1 クラスターの数を5とした場合の各クラスターにおける主要溶存イオンと硝酸性窒素 (NO₃-N) の平均濃度

クラスター	Cl ⁻ mg L ⁻¹	NO ₃ ⁻ mg L ⁻¹	SO ₄ ²⁻ mg L ⁻¹	HCO ₃ ⁻ mg L ⁻¹	Na ⁺ mg L ⁻¹	K ⁺ mg L ⁻¹	Mg ²⁺ mg L ⁻¹	Ca ²⁺ mg L ⁻¹	NO ₃ -N mg L ⁻¹
手法1									
i	5.3	11.7	3.8	37.7	7.0	3.7	3.5	9.2	2.6
ii	6.7	9.9	13.6	113.0	13.2	5.3	10.9	21.0	2.2
iii	7.1	5.2	49.7	209.2	31.7	8.8	21.4	44.6	1.2
iv	16.4	56.3	25.0	36.2	12.9	6.9	9.1	21.8	12.7
v	22.6	85.6	41.6	27.3	15.3	9.3	11.3	33.9	19.3
手法2									
i	4.7	8.1	3.0	38.5	6.6	3.5	3.2	8.4	1.8
ii	6.7	9.9	13.5	112.1	13.2	5.2	10.8	20.9	2.2
iii	7.1	5.2	49.7	209.2	31.7	8.8	21.4	44.6	1.2
iv	14.8	50.5	22.5	35.7	12.4	6.2	8.3	20.2	11.4
v	22.4	85.1	40.1	27.5	15.1	9.3	11.2	33.2	19.2
手法3									
i	4.9	9.0	3.5	39.3	6.7	3.6	3.3	8.7	2.0
ii	6.9	10.5	13.5	111.1	13.2	5.3	10.8	21.1	2.4
iii	7.1	5.2	49.7	209.2	31.7	8.8	21.4	44.6	1.2
iv	16.8	59.3	27.2	33.3	13.3	7.0	9.1	22.6	13.4
v	23.7	90.6	42.8	26.7	15.4	9.7	11.6	36.3	20.5
手法4									
i	4.8	8.8	3.1	38.2	6.7	3.5	3.2	8.5	2.0
ii	6.7	10.0	13.2	111.4	13.1	5.2	10.7	20.9	2.3
iii	7.1	5.2	48.1	204.3	30.8	8.6	21.1	42.8	1.2
iv	15.4	53.2	23.4	35.6	12.5	6.4	8.5	20.8	12.0
v	22.6	85.7	40.9	27.2	15.3	9.5	11.3	33.8	19.4

が基準値を超過している点で同じであり、クラスター iv と v の分布の違いはさほど重要ではない。特筆すべき点は、解析手法によってクラスター i にも iv にも分類される A 点についてである。クラスター iv と v の分類結果の違いは汚染レベルの違いであるが、クラスター i は汚染の影響を受けていないクラスターであり、クラスター iv は汚染クラスターであるため、どちらに分類されるかで採水地点の水質特性は大きく異なる。

次に、手法1と3、手法2と4のどちらが島原市の水質を説明するに適しているか硝酸性窒素濃度の分布とクラスターの分布を比較することで検討した。クラスター i もしくは iv に分類される A 点の硝酸性窒素濃度は、10 mg L⁻¹ (硝酸性窒素及び亜硝酸性窒素の水道水質基準¹⁴⁾ および環境基準¹⁵⁾) を超えないものの 3 mg L⁻¹ (人為的活動の影響の閾値¹⁶⁾) を超過した。そのため、この採水地点は、汚染クラスターに分類するほうがより適切である。手法1と3そし

て手法2と4の大きな違いは、主成分分析を実行し、パラメータを減らしたうえでクラスター分析したかどうかである。以上のことから、得られた主要溶存イオン濃度を用いてクラスター分析、または SOM を介してクラスター分析するよりも、主成分分析より得た主成分得点を利用してクラスター分析もしくは SOM を介してクラスター分析したほうが、硝酸性窒素濃度の汚染レベルに対応したサンプリング地点の分類ができることから、合理的な水質特性を表現できたと考えられる。

図-7に各クラスターのヘキサダイアグラムを示す。なお、それぞれの手法で各クラスターの特徴はほぼ同様であったため、手法間で各クラスターの各種イオンの濃度を算術平均し、ヘキサダイアグラムを描画した。汚染の影響を受けていないクラスター i ~ iii の水質組成は、同じ Ca-HCO₃型を示すが、その溶存イオン濃度が異なることが、ヘキサダイアグラムから判読できる。一方、汚染クラスターであるクラスター iv と v の水質組成は Ca-NO₃型を示す。クラスター iv と v では、陰イオンの組成が異なり、iv では NO₃⁻ > HCO₃⁻ > SO₄²⁻ > Cl⁻ であり、v では NO₃⁻ > SO₄²⁻ > Cl⁻ > HCO₃⁻ である。

(2) クラスターの数を10とした場合

次にクラスターの数を10とした場合について検討する。表-2に各クラスターにおける主要溶存イオンと硝酸性窒素 (NO₃-N) の平均濃度をそれぞれの解析手法ごとに示す。クラスターの数を5とした場合と異なり、10のクラスターは、大きく3つのクラスターに分けることができる。1つはクラスターの数を5とした場合と同様に、汚染の影響を受けてい

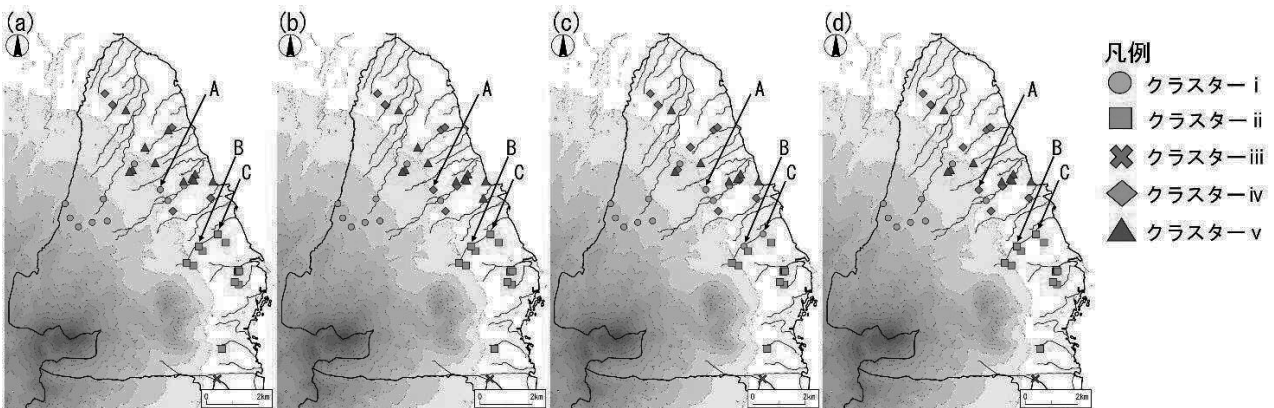


図-6 クラスターの数を5とした場合の各クラスターの分布；(a) 手法1, (b) 手法2, (c) 手法3, (d) 手法4

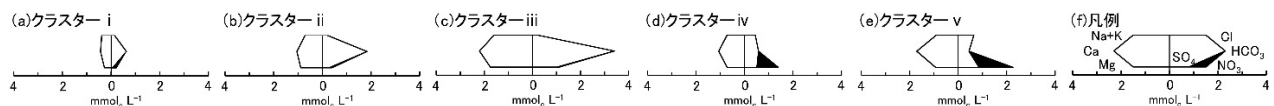


図-7 クラスターの数を5とした場合の各クラスターのヘキサダイアグラム

表-2 クラスターの数を10とした場合の各クラスターにおける主要溶存イオンと硝酸性窒素 (NO₃-N) の平均濃度

クラスター	Cl ⁻ mg L ⁻¹	NO ₃ ⁻ mg L ⁻¹	SO ₄ ²⁻ mg L ⁻¹	HCO ₃ ⁻ mg L ⁻¹	Na ⁺ mg L ⁻¹	K ⁺ mg L ⁻¹	Mg ²⁺ mg L ⁻¹	Ca ²⁺ mg L ⁻¹	NO ₃ -N mg L ⁻¹
1	7.1	5.4	51.7	214.0	33.9	9.1	19.2	38.7	1.2
2	7.3	4.6	40.8	187.3	22.2	7.8	30.9	71.0	1.0
3	6.9	7.1	17.3	131.8	15.8	5.9	13.4	23.6	1.6
4	6.6	12.1	10.6	97.6	11.1	4.7	8.8	18.8	2.7
5	3.9	3.6	1.4	38.1	6.3	3.3	2.6	6.5	0.8
6	7.8	25.5	7.8	36.8	8.3	4.3	4.9	13.9	5.8
7	16.4	56.3	25.0	36.2	12.9	6.9	9.1	21.8	12.7
8	18.6	73.4	38.1	35.9	13.2	13.2	7.9	35.5	16.6
9	22.6	81.0	52.2	21.2	14.9	7.8	12.1	33.7	18.3
10	24.4	96.4	30.7	30.8	16.7	9.5	11.9	33.5	21.8
1	7.1	5.2	49.7	209.2	31.7	8.8	21.4	44.6	1.2
2	8.2	9.5	28.4	133.0	19.2	6.5	14.0	24.6	2.1
3	6.0	5.4	9.5	134.0	12.8	5.5	12.4	23.0	1.2
4	6.6	12.6	10.6	92.6	11.3	4.6	8.8	18.5	2.8
5	3.7	2.4	1.2	38.9	6.2	3.2	2.6	6.3	0.5
6	6.4	18.2	6.2	37.7	7.5	4.0	4.3	12.1	4.1
7	11.5	37.0	16.3	31.2	10.2	4.8	6.4	15.8	8.4
8	16.4	56.9	25.5	37.8	13.4	6.9	9.2	22.2	12.8
9	21.0	80.0	36.3	27.5	14.0	8.8	10.5	29.7	18.1
10	24.6	93.0	46.0	27.5	16.8	10.2	12.3	38.7	21.0
1	7.1	5.2	49.7	209.2	31.7	8.8	21.4	44.6	1.2
2	6.8	7.2	17.7	131.2	15.9	5.9	13.2	23.5	1.6
3	6.9	13.0	10.4	96.2	11.2	4.9	9.1	19.4	2.9
4	4.3	6.0	2.3	38.3	6.4	3.4	2.9	7.3	1.4
5	7.6	22.2	8.3	44.1	8.0	4.7	5.2	15.0	5.0
6	13.0	39.6	20.5	35.7	11.9	5.1	7.4	17.7	9.0
7	17.9	69.7	22.8	33.9	13.0	9.3	8.9	23.9	15.7
8	18.5	64.2	35.9	31.0	14.6	6.3	10.6	25.0	14.5
9	24.8	85.4	31.5	77.7	22.2	9.6	14.4	44.9	19.3
10	23.7	90.7	43.0	25.7	15.3	9.7	11.6	36.1	20.5
1	7.1	5.2	48.1	204.3	30.8	8.6	21.1	42.8	1.2
2	6.9	7.1	16.5	132.2	15.4	5.9	13.0	23.6	1.6
3	6.7	11.3	11.2	104.4	12.1	4.8	9.5	19.2	2.6
4	6.5	13.8	8.5	70.0	9.0	4.6	6.9	17.2	3.1
5	4.3	6.1	2.4	38.2	6.4	3.4	2.9	7.3	1.4
6	10.3	21.4	32.2	64.8	23.3	10.3	4.8	14.0	4.8
7	8.8	30.8	8.3	34.0	9.0	4.8	5.2	14.7	6.9
8	14.2	40.2	25.1	36.5	11.4	5.2	7.9	19.6	9.1
9	17.2	63.3	25.4	35.2	13.4	7.0	9.6	22.6	14.3
10	22.5	85.4	40.8	27.1	15.2	9.4	11.3	33.6	19.3

ないクラスターであり、クラスターの数を5とした場合と同様に、NO₃⁻を除き、各種イオン濃度の違いで区別できる。手法1と2ではクラスター1~5、手法3ではクラスター1~4、手法4ではクラスター1~3と5がこのクラスターに属する。もう1つは、同じくクラスターの数を5とした場合と同様に、硝酸性窒素及び亜硝酸性窒素の水道水質基準¹⁴⁾および環境基準¹⁵⁾10 mg L⁻¹を超過する汚染クラスターであり、手法1ではクラスター7~10、手法2ではクラスター8~10、手法3ではクラスター7~10、そして手法4ではクラスター9と10がこのクラスターに属する。どちらにも属さないクラスターが、手法1ではクラスター6、手法2ではクラスター6と7、手法3ではクラスター5と6、手法4ではクラスター4と6~8がそれぞれの手法にあるが、硝酸性窒素濃度に着目すると、人為的活動の影響の閾値¹⁶⁾3 mg L⁻¹を超えているものの、硝酸性窒素及び亜硝酸性窒素の水道水質基準¹⁴⁾および環境基準¹⁵⁾10 mg L⁻¹を満たしている。本研究では、硝酸性窒素濃度が3 mg L⁻¹より大きく10 mg L⁻¹以下のクラスターを準汚染クラスターと定義する。

図-8に各クラスターの分布を解析手法ごとに示す。なお、各地点とも観測期間最新の分析値によるクラスターをプロットした。クラスターの数を5とした場合と異なり、各クラスターの持つ意味がそれぞれの手法で必ずしも一致しないので比較には注意を要する。市の最南部に位置する採水地点のクラス

ターが、他の採水地点のクラスターと異なることは、どの手法においても共通している。標高200 m以上に位置する採水地点に着目すると、手法3では全ての地点が同じクラスターに分類されたが(図-8 (c))、手法1と2ではD点(図-8 (a), (b))、手法4ではG点(図-8 (d))が異なるクラスターに分類された。標高100 m以下の市南部に着目すると、手法1と2では、すべての採水地点が汚染の影響を受けていないクラスターに分類されたが、手法3ではB, C点(図-8 (c))、手法4ではB, C, E点(図-8 (d))が準汚染クラスターに分類された。標高200 m以下の市北部に着目すると、手法1と2では、すべての採水地点が準汚染もしくは汚染クラスターに分類されたが、手法3と4ではF点(図-8 (c), (d))のみ汚染の影響を受けていないクラスターに分類された。

クラスターの数を5とした場合と同様に、どの解析手法による分類結果が島原市の水質を説明するうえで適しているか硝酸性窒素濃度の分布とクラスターの分布を比較することで検討した。まず手法4では、B, C, E, G点が準汚染クラスター(クラスター4)に分類されているが、4地点の硝酸性窒素濃度は3 mg L⁻¹以下であり、適切に分類できていない。手法3では、手法4と同様にB, C点(図-8 (c))が準汚染クラスター(クラスター5)に分類されており、手法3による分類は適切ではない。手法1と2では、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の分布はほとんど一致している。手法1と2では、硝酸性窒素濃度が3 mg L⁻¹より大きく10 mg L⁻¹以下のD, F点(図-8 (a), (b))が準汚染クラスター(両手法ともクラスター6)に分類され、B, C, E点(図-8 (a), (b))は手法3と4とは異なり汚染の影響を受けていないクラスターに分類された。以上のことから、手法1と2は手法3と4より、硝酸性窒素濃度による汚染レベルとの対応状況から適切に水質特性を表現していると考えられる。

図-9に各クラスターのヘキサダイアグラムを示す。クラスターの数を5とした場合と異なり、島原市の水質特性をより適切に表現していると考えられる手法1と2の間で各クラスターの各種イオンの濃度を算術平均し、ヘキサダイアグラムを描画した。なお、クラスター7は、手法1では汚染クラスターに、手法2では準汚染クラスターに属し、特徴が異なるため平均濃度は算出せず、手法1については実線で、手法2については破線でヘキサダイアグラムを描画した。汚染の影響を受けていないクラスター1~5の水質組成は、クラスター5はNa-HCO₃型、他はCa-HCO₃型を示す。準汚染クラスターであるク

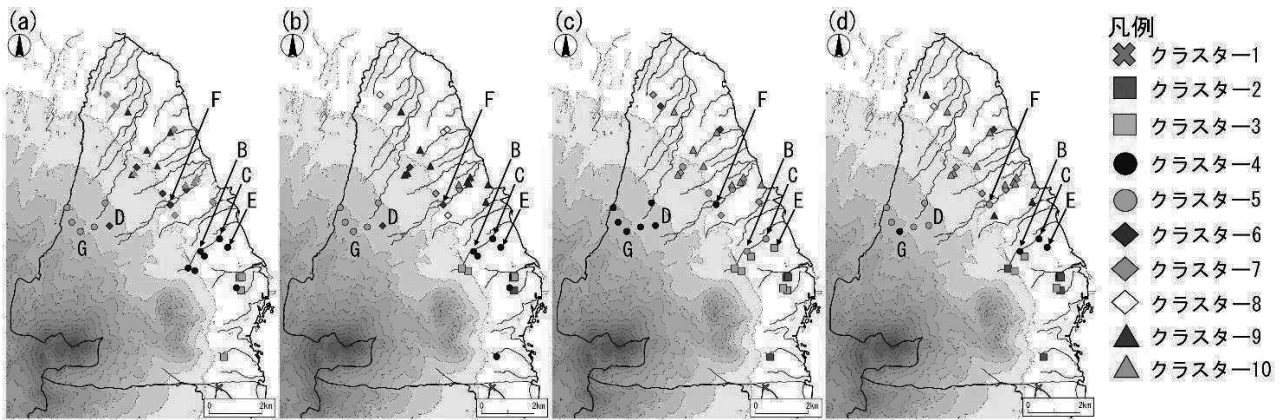


図-8 クラスターの数を10とした場合のクラスターの分布；(a) 手法1, (b) 手法2, (c) 手法3, (d) 手法4

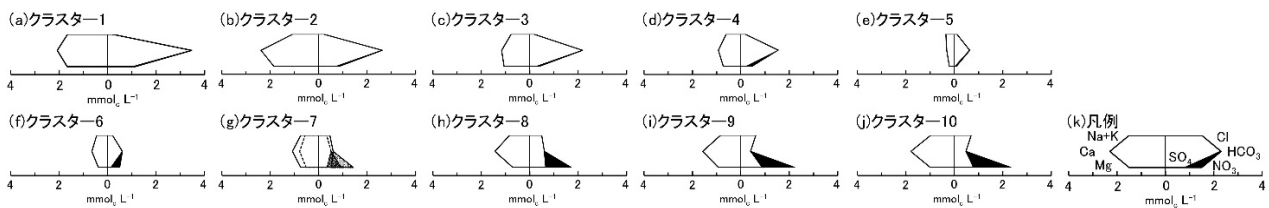


図-9 クラスターの数を10とした場合の各クラスターのヘキサダイアグラム

クラスター6の水質組成は、汚染の影響がヘキサダイアグラムからも確認できるもののCa-HCO₃型を示す。一方、汚染クラスターであるクラスター8~10は、Ca-NO₃型を示す。手法1では汚染、手法2では準汚染クラスターに分類されるクラスター7は、手法1ではCa-NO₃型、手法2ではCa-HCO₃型を示す。市南部では、標高が低くなるにつれて溶存イオン濃度が上昇し、水質組成は標高100m以下ではCa-HCO₃型である。市北部では、標高が低くなるにつれて溶存イオン濃度が上昇するが、合わせて汚染の影響も受ける。水質組成は、標高200m以上ではNa-HCO₃型、標高100m付近ではCa-HCO₃型、標高100m以下ではCa-NO₃型である。

6. おわりに

本研究では、長崎県島原市を対象とし、異なる4つの多変量解析手法により水質分類を行った。解析はクラスターの数を5または10とした場合の2通りで実施し、それぞれの場合において分類結果を比較した。その結果、クラスターの数を5とした場合、各クラスターの持つ意味はほぼ同様であり、いずれの手法でも5つのクラスターは汚染の影響をうけていないクラスターと汚染クラスターに分けることができた。各クラスターの分布では、特定地点のクラスター分類に違いが認められ、硝酸性窒素濃度の分布との比較から主成分分析を利用した手法2と4が

島原市の水質特性を説明するうえで適していると判断された。

クラスターの数を10とした場合、汚染の影響を受けていないクラスター、汚染クラスターそして準汚染クラスターに分けることができた。クラスターの数を増やしたため各クラスターの分布は、それぞれの分類結果で異なっており、硝酸性窒素濃度の分布との比較からSOMを利用しない手法1と2が島原市の水質特性を説明するうえで適していると判断された。この結果は、単一の事例にもとづくものなので、別のデータを用いて検討するなど、さらに検討する必要がある。

以上のように、分類するクラスターの数を変えた2通りの場合において、多変量解析の結果を比較したところ、手法2が島原市の地下水水質を説明するに適する手法と判断された。研究対象地域の水質データは、手法1のようにクラスター分析のみにより分類することができるが、分類されたクラスターの特徴と実際のデータの特徴が必ずしも一致しないため、本研究のようにいくつかの多変量解析手法を試み、対象地に適した解析手法を選択することが重要であると考えられる。

本研究において手法3と4では、SOM解析の後にクラスター分析するため、入力データが手法1と2と比べ1/3程度となっている。手法1と2では入力データ数は同じであるが、手法2において、パラメータの数をイオン濃度の8個から主成分の数2個としている。つまり、データ数は減らさずに、パラメー

夕数を減らすことが有効であったといえる。しかし、手法3と4で用いたSOM解析は、図-4や図-5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なため、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いえよう。

謝辞：本研究は、JSPS 科研費 24360194 および 15KT0120 の助成を受けたものである。ここに記して謝意を表す。

参考文献

- 1) Nadiri, A. A., Moghaddam, A. A., Tsai, F. T. C. and Fijani, E. : Hydrogeochemical analysis for Tasuj plain aquifer, Iran, *Journal of Earth System Science*, Vol. 122, No. 4, pp. 1091-1105, 2013.
- 2) Li, P., Qian, H., Wu, J., Zhang, Y. and Zhang, H. : Major ion chemistry of shallow groundwater in the Dongsheng coalfield, Ordos basin, China, *Mine Water and the Environment*, Vol. 32, No. 3, pp. 195-206, 2013.
- 3) Aiuppa, A., Bellomo, S., Brusca, L., D'Alessandro, W. and Federico, C. : Natural and anthropogenic factors affecting groundwater quality of an active volcano (Mt. Etna, Italy), *Applied Geochemistry*, Vol. 18, No. 6, pp. 863-882, 2003.
- 4) Matitaos, I. : Nitrate source identification in groundwater of multiple land-use areas by combining isotopes and multivariate statistical analysis: A case study of Asopos basin (Central Greece), *Science of the Total Environment*, Vol. 541, pp. 802-814, 2016.
- 5) Yidana, S. M., Ophori, D. and Banoeng-Yakubo, B. : A multivariate statistical analysis of surface water chemistry data - The Ankobra Basin, Ghana, *Journal of Environmental Management*, Vol. 86, No. 1, pp. 80-87, 2008.
- 6) 安元 純, 聖川健斗, 仲栄真史哉, 砥綿泰弘, 中野拓治 : 多変量解析を用いた石灰岩帯水層における地下水水質の統計的分類, 土木学会論文集 B1 (水工学), Vol. 69, No. 4, pp. I_595-I_600, 2013.
- 7) Cloutier, V., Lefebvre, R., Therrien, R. and Savard, M. M. : Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system, *Journal of Hydrology*, Vol. 353, No. 3-4, pp. 294-313, 2008.
- 8) Omonona, O. V., Onwuka, O. S. and Okogbue, C. O. : Characterization of groundwater quality in three settlement areas of Enugu metropolis, southern Nigeria, using multivariate analysis, *Environmental Monitoring and Assessment*, Vol. 186, No. 2, pp. 651-664, 2014.
- 9) Nguyen, T. T., Kawamura, A., Tong, T. T., Nakagawa, N., Amaguchi, H. and Gilbuena, R., Jr. : Clustering spatio-seasonal hydrogeochemical data using self-organizing maps for groundwater quality assessment in the Red River Delta, Vietnam, *Journal of Hydrology*, Vol. 522, pp. 661-673, 2015.
- 10) Jin, Y. H., Kawamura, A., Park, S. C., Nakagawa, N., Amaguchi, H. and Olsson, J. : Spatiotemporal classification of environmental monitoring data in the Yeongsan River basin, Korea, using self-organizing maps, *Journal of Environmental Monitoring*, Vol. 13, No. 10, pp. 2886-2894, 2011.
- 11) 島原半島窒素負荷低減対策会議 : 第2期 島原半島窒素負荷低減計画～雲仙岳の恵みを次の世代に引き継ぐために～, 長崎県環境政策課, 100p, 2011.
- 12) 気象庁 : 気象統計情報, <http://www.jma.go.jp/jma/menu/menureport.html> (2016.04.01 閲覧)
- 13) Nakagawa, K., Amano, H., Asakura, H. and Berndtsson, R. : Spatial trends of nitrate pollution and groundwater chemistry in Shimabara, Nagasaki, Japan, *Environmental Earth Sciences*, Vol. 75, 234, 2016.
- 14) 厚生労働省 : 水質基準項目と基準値, <http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/topics/bukyoku/kenkou/suid/o/kijun/kijunchi.html> (2016.06.01 閲覧)
- 15) 環境省 : 人の健康の保護に関する環境基準, <http://www.env.go.jp/kijun/wt1.html> (2016.06.01 閲覧)
- 16) Eckhardt, D. A. V. and Stackelberg, P. E. : Relation of ground-water quality to land use on long island, New York, *Groundwater*, Vol. 33, No. 6, pp. 1019-1033, 1995.
- 17) García, H. L. and González, I. M. : Self-organizing map and clustering for wastewater treatment monitoring, *Engineering Applications of Artificial Intelligence*, Vol. 17, No. 3, pp. 215-225, 2004.
- 18) Nakagawa, K., Amano, H., Kawamura, A. and Berndtsson, R. : Classification of groundwater chemistry in Shimabara, using self-organizing maps, *Hydrology Research* (2016. 06. 06 受理)
- 19) Kohonen, T. : The Self-Organizing Maps, *Proceedings of The IEEE*, Vol. 78, No. 9, pp. 1464-1480, 1990.
- 20) 石原成幸, 河村 明, 天口英雄, 高崎忠勝, 高橋泰之, 川合将文 : 自己組織化マップを用いた東北地方太平洋沖地震発生前後の東京の年間地下水変動特性, 土木学会論文集 B1, Vol. 70, No. 4, pp. I_1129-I_1134, 2014.

(2016. 4. 7 受付)

CLASSIFICATION CHARACTERISTICS OF MULTIVARIATE ANALYSES FOR GROUNDWATER CHEMISTRY –CASE STUDY ON SHIMABARA CITY–

Hiroki AMANO, Kei NAKAGAWA and Akira KAWAMURA

Results of 4 multivariate analyses for groundwater chemistry were compared to discuss which classification result is the most suitable to explain water chemistry in Shimabara city. Though characteristics of 5 clusters are almost same among all multivariate analysis methods, the distributions of each cluster were different among 4 methods. In the two methods using row data of 8 major ion concentrations, higher nitrate concentration points were not classified into polluted clusters. In this regards, the two methods using principal component scores obtained by principal component analysis seem better to represent classification of the field water chemistry because both of points are belong to nitrate pollution cluster. When the number of cluster was determined to 10, the method using principal component scores is more suitable than the other methods to explain water chemistry.