

水塊分類ソフト（TS-Cluster.ver2.0）マニュアル

<目次>

1. 概要
2. 動作環境
3. 実行環境の準備
4. 使用方法
5. 解析方法
6. その他
7. 更新記録

1. 概要

水塊分類ソフト（TS-Cluster）は、クラスター分析により水温、塩分の情報から客観的に水塊を分類し、可視化するソフトです。カンマ区切り形式（CSV; Comma-Separated Values）の入力データを用意し、Windows OS 環境のパーソナルコンピュータ上で実行することにより、水温・塩分を縦軸・横軸に取った図(T-S ダイアグラム)として水塊分類が可視化されます。ソフトは MathWorks 社の MATLAB を利用して作成されたもので、実行形式となっており、ユーザーは MATLAB の共有ライブラリのスタンドアロンセット (MATLAB Runtime) を使用することにより、MATLAB がインストールされていないコンピュータでも水塊分類ソフトが利用できます。

2. 動作環境

OS : Windows7, 8, 10

ソフト : MATLAB の Runtime がインストールされていること

メモリ : 500MB 以上 (目安)

ソフトを実行した場合に必要なメモリ容量はデータ数に依存します。以下の表は必要メモリ容量の目安です。必要なメモリ容量が確保できなかった場合は、ソフトは実行されません。

データ行数	必要メモリ
1,000	100MB
5,000	300MB
10,000	900MB
15,000	1.9GB
20,000	3.3GB

3. 実行環境の準備

プログラムの実行には MathWorks 社が配布している MATLAB Runtime をインストールする必要があります。以下に従って使用するコンピューターに MATLAB Runtime (R2013b 8.2 32bit) をインストールして下さい。

- (1) インストーラーファイル (MCR_R2013b_win32_installer.exe) を、水塊分類ソフトを使用するコンピューターに保存します。
- (2) インストーラーをダブルクリックし、インストール ウィザードの指示に従います。

注意) バージョンが一致していない Runtime ではプログラムは実行できません。

ダウンロード先 : <https://jp.mathworks.com/products/compiler/matlab-runtime.html>

MCR_R2013b_win32_installer.exe のリンク先 :

http://ssd.mathworks.com/supportfiles/downloads/R2013b/deployment_files/R2013b/installers/win32/MCR_R2013b_win32_installer.exe

4. 使用方法

(1) 解析データの準備

表計算ソフト等を利用してカンマ区切りの入力データを準備して下さい。1行目にはデータ項目を説明する情報を半角英数字で記載し、2行目以降に数値データを入力して下さい。欠測値がある場合は nan を入力して下さい。解析には水温、塩分値を利用しますが、他のデータを記載しても問題ありません。

yyyy	mm	st	temp	Salt
2000	6	KR00	10.84	33.812
2000	6	KR05	10.49	33.844
2000	6	KR10	9.94	33.785
2000	6	KR20	9.44	33.938
2000	6	KR30	8.25	33.8
2000	6	KR40	5.53	33.43
2000	6	KR50	5.28	33.365
2000	6	TD00	10.17	33.827
2000	6	TD05	9.72	33.76

図1 サンプルデータ

(2) 解析作業用フォルダーの作成

新規フォルダーを作成し (例 : TS_CLUST_GUI) 、入力ファイル (例:data.csv) 、ソフト実行ファイル (TS_CLUST_GUI.exe) を作業フォルダーに入れる (図2) 。

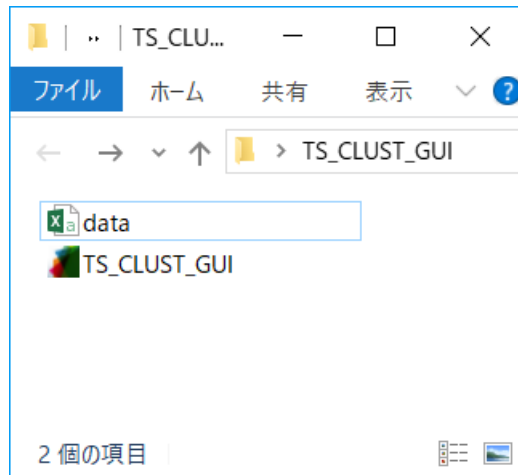


図2 解析作業用フォルダー作成例

(3) ソフト起動

作業フォルダー内の `TS_CLUST_GUI.exe` をダブルクリックしてソフトを起動する。ソフトが起動すると設定条件入力画面が表示される（図3）。起動には数十秒かかる場合があります。

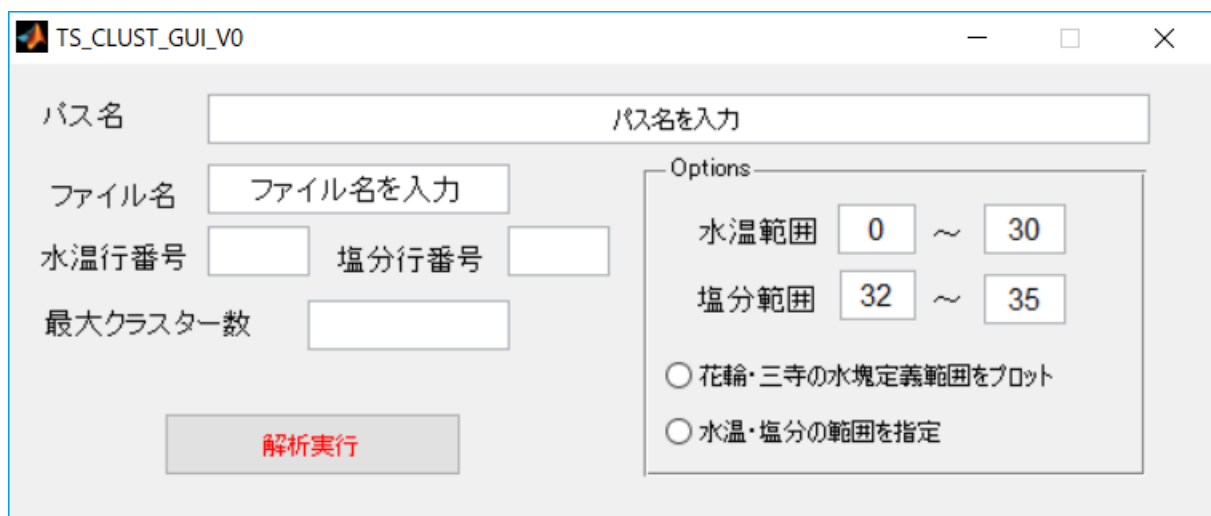


図3 ソフト起動画面

(4) 条件入力および実行

- a) 「パス名」：解析データのディレクトリパスを入力する。上記のようにソフトと解析データが同一フォルダーにある場合は「./」（ドット、スラッシュ）を入力する。ソフトと別のファイルにある場合は、絶対パスを入力する。例えば、Cドライブの中に `input` というフォルダーにデータがある場合は「`C:/input/`」と入力する。
- b) ファイル名を「`data.csv`」のように入力する。
- c) 水温データ、塩分データの行番号を入力する。図1の様なデータの場合は、水温行番号に4, 塩分行番号に5を入力する。

- d)最大クラスター数を入力する。
- e)必要に応じて Options（画面右側）を設定する。ラジオボタン：「花輪・三寺の水塊定義範囲をプロット」をチェックすると、Hanawa & Mitsudera（1987）による岩手県の三陸沿岸における水塊区分が T-S ダイアグラムに表示される。ラジオボタン：「水温・塩分の範囲」をチェックすると、任意の範囲で T-S ダイアグラムが作図されます。チェックがない場合は、最大、最小値を基に表示範囲が自動設定されます。
- f)左下の「解析実行」ボタンを押すと解析が実行され、T-S ダイアグラムが表示されます（図4）。出力画面の左上「ファイル(F)」を選択すると出力された図を各種のフル型式で保存することができる。eps 型式で保存するとイラストレーター等のドローソフトで図の加工が容易である。

*注意) メモリ不足の場合は解析途中で停止し、表示画面に変化はありません。

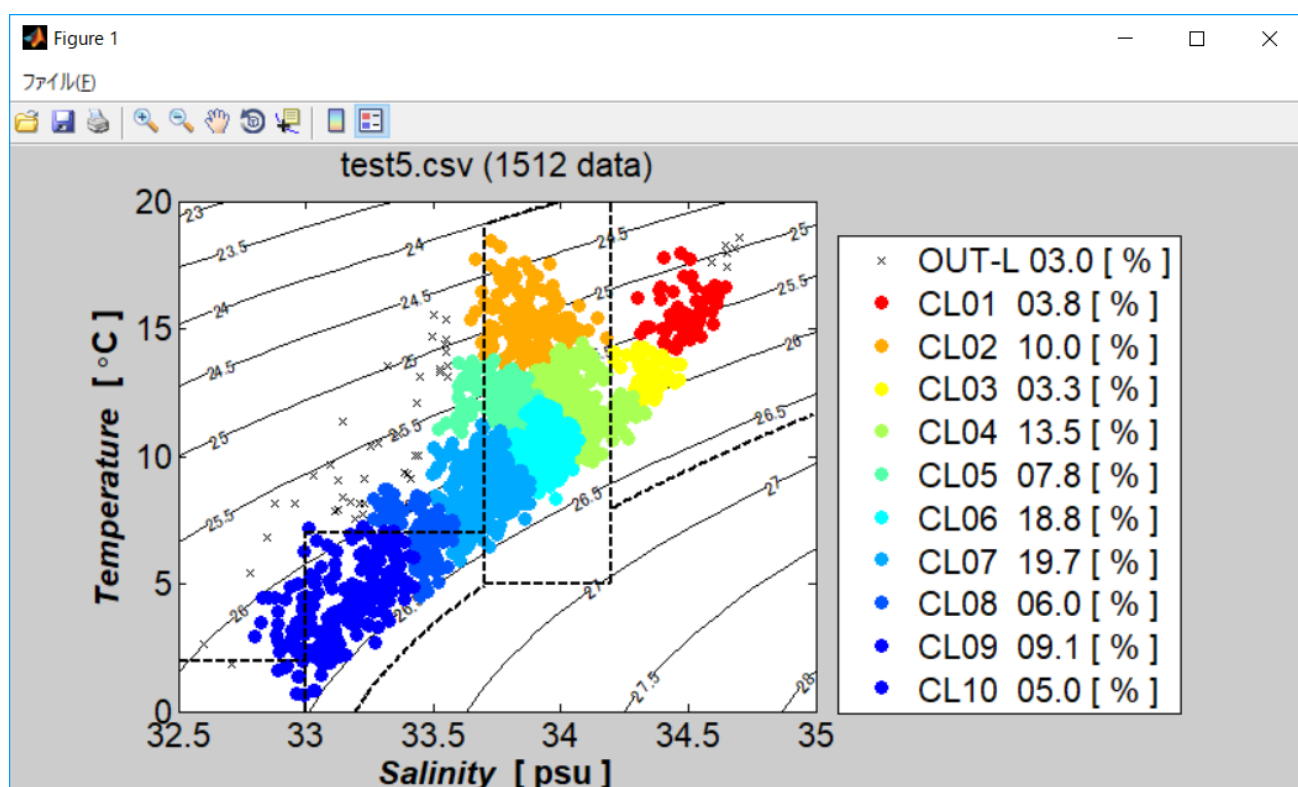


図4 T-S ダイアグラムの出力例。上部に図のデータファイル名とデータ数が表示される。図中のコンタ線は等密度 (Sgima-t) を示す。右側凡例にクラスターナンバー (CL-01～最大クラスター数) と出現率[%]が表示される。OUT-L は外れ値である（「5. 解析手法」で説明）。クラスターナンバーは熱力学変数の1つであるスパイスネス (Flament, 2002) が高い (高温・高塩分) 順に振られる。点線は Hanawa & Mitsudera (1987) により定義される水塊区分を示す。

(5) 出力

解析実行後には図 5 に示す様な 3 つの出力ファイル (*_10CLS.csv, *_10CLS_INFO.csv, *_10CLS.png ; *は入力ファイル名) が入力データと同一フォルダー内に出力される。

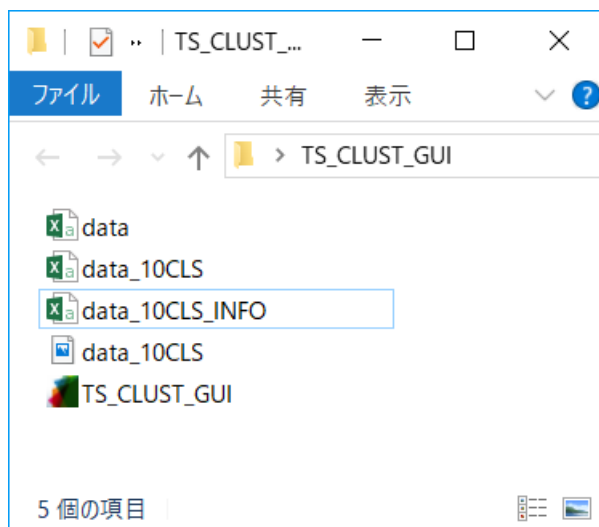


図 5 解析実行後の出力例

出力ファイル「*_10CLS.csv」には、入力ファイルに密度 (sigma-t) とクラスターナンバー (CL) が追記される (図 6)。また、出力ファイル「*_10CLS_INFO.csv」には、クラスター別のデータ数、水温、塩分の最小値、最大値、平均値、が出力される (図 7)。さらに、出力ファイル「*_10CLS.png」は画面上に表示された T-S ダイアグラムが png 形式の画像ファイルとして出力されたものである。

	A	B	C	D	E	F	G
1	yyyy	mm	st	temp	Salt	sigma-t	CL
2	2000	6	KR00	10.84	33.812	25.879	1
3	2000	6	KR05	10.49	33.844	25.966	1
4	2000	6	KR10	9.94	33.785	26.014	5
5	2000	6	KR20	9.44	33.938	26.216	1
6	2000	6	KR30	8.25	33.8	26.294	5
7	2000	6	KR40	5.53	33.43	26.369	6
8	2000	6	KR50	5.28	33.365	26.347	7
9	2000	6	TD00	10.17	33.827	26.008	1
10	2000	6	TD05	9.72	33.76	26.031	5
11	2000	6	TD10	9.55	33.856	26.134	1
12	2000	6	TD20	11.1	33.6	25.668	2
13	2000	6	TD30	10.22	33.996	26.131	1
14	2000	6	TD40	9.4	33.847	26.151	5
15	2000	6	TD50	9.45	33.957	26.229	1
16	2000	6	OZ00	9.5	33.655	25.985	5
17	2000	6	OZ05	9.56	33.686	25.999	5
18	2000	6	OZ10	11.66	34.331	26.134	8

図6 「*_10CLS.csv」の出力例。A~E列は入力ファイルと同一データが入り、F列に密度（sigma-t）、G列にクラスターナンバー（CL）が追記されている。外れ値（OUT-L）はCLが0と表示される。

Class No.	points	[%]	min(Tem) [oC]	max(Tem) [oC]	mean(Tem) [oC]	min(Sal) [psu]	max(Sal) [psu]	mean(Sal) [psu]
OUT-L	46	3.04	nan	nan	nan	nan	nan	nan
1	57	3.77	14.2	17.95	15.6218	34.298	34.648	34.4831
2	151	9.99	13.47	18.47	15.2506	33.647	34.178	33.871
3	50	3.31	12.16	14.26	13.223	34.209	34.473	34.3465
4	204	13.49	9.79	14.45	12.2167	33.902	34.331	34.0793
5	118	7.8	10.92	13.78	12.2914	33.52	33.906	33.7539
6	285	18.85	8.32	12.13	10.3163	33.787	34.07	33.9197
7	298	19.71	5.93	11.24	8.907	33.496	33.879	33.7108
8	91	6.02	4.53	8.71	6.9586	33.265	33.583	33.4528
9	137	9.06	2.71	7.26	5.0172	32.923	33.423	33.2168
10	75	4.96	0.65	4.48	2.6855	32.798	33.157	33.0094

図7 「*_10CLS_INFO.csv」の出力例。OUT-Lは外れ値を示す。

5. 解析方法：

ソフトはクラスター分析手法として、類似度：標準化ユークリッド距離、クラスター間連結方法：群平均法（UPGMA）を用い、水温、塩分から指定した最大クラスター数に分類している。解析データに外れ値（水温、塩分の値が他より逸脱しているデータ）が存在した場合、その外れ値が指定したクラスター数を多く占めるため、外れ値を除去する前処理を行っている。前処理では、指定した最大クラスター数でクラスター解析が実行され、データ数が1%未満となるクラスターが出現した場合、このクラスターのデータを解析データから除去する。外れ値が除去された解析データによって、再度クラスター解析が実行され、データ数が1%未満となるクラスターが出現しなくなるまで、外れ値の除去と反復計算が実行される。前処理後にクラスター解析が実施され、最終的な結果として出力される。除去される外れ値（OUT-L）は積算されるため、1%以上のデータ数となる場合がある。なお、外れ値の一部には測器のエラー値を含んでいる可能性がある。

6. 更新記録

2018/05/05 マニュアル ver1.0 を記載（奥西 武）

2018/06/02 マニュアル修正, ver2.0 とする（奥西 武）

<参考文献>

Hanawa, K., and H. Mitsudera (1987) Variation of water system distribution in the Sanriku coastal area. *J. Oceanogr. Soc. Japan*, **42**, 435–446.

Flament P. (2002) A state variable for characterizing water masses and their diffusive stability: spiciness, *Progress in Oceanography* 54. 493-501