

R 版 DECOMP を用いた水温・水質データの季節調整によるトレンド解析

【はじめに】

全国の自治体の環境部署や一部国土交通省により、河川、湖沼、海域において公共用水域水質測定が 1970 年代から実施されてきており、これまで 40 年間近い測定データが蓄積されてきている。公共用水域水質測定では水温も基本項目として測定されており、その測定頻度は概ね毎月・年間 12 回である（2 か月か 3 ヶ月に一回 [つまり年間 4 回か 6 回] の頻度の測定点もあり、日本海では冬季の半年近くの間測定されていない場合も多い）。

水温と同様に化学的酸素要求量（COD）や全窒素（TN）・全窒素（TP）、表層・底層の溶存酸素量（DO）等、内部生産等の生物の作用・物質循環に係る水質項目は季節的変動を示し、一年間内での変動幅は、数十年間単位の長期に緩やかに変動する幅に比べて遥かに大きい。

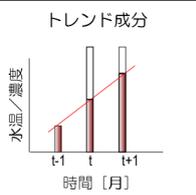
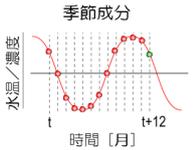
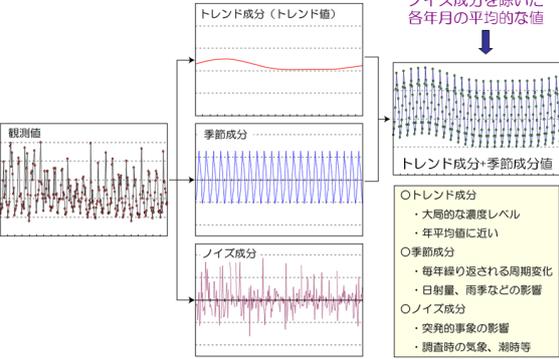
例えば海水温の場合、夏季と冬季では 20℃近い差が有るのに対し、日本近海での長期的な変化は 100 年間で 0.3~1.9℃（平均 1.2℃）程度であるとされている。このように短期的に変化の著しい海水温等の月一回程度の頻度で取得された間欠的なデータから、緩やかな長期的な変化（トレンド）を抽出する統計解析手法の一つとして、季節調整の適用方法と事例について以下に解説する。

【季節調整】

ここでは、従来行われてきたダミー変数回帰や移動平均等に代わって、Kitagawa (1981) により提案された時系列解析手法の 1 つである「状態空間モデルを使った季節調整法」及びそのプログラム実装版である DECOMP (Akaike 他 [1985] の TIMSAC-84 所収) を利用した海水温等の長期変動解析を行った。以降は R 版 TIMSAC に含まれる `decomp` 関数の利用方法を示す。なお、DECOMP の日本語解説としては北川 (1986) がある。

季節調整法は、下図の概要に示すように、季節変化等により周期的な変動を繰り返す時系列データをトレンド成分、季節成分、ノイズ成分、それに自己回帰成分に分解する統計手法である（下図では自己回帰成分については割愛）。本手法の開発背景は経済学分野への利用であるが、東京湾での水温や水質の水平分布の経年変化を解析した例 (Kashiwagi 2003) など、水質トレンド解析等の分野への応用が進められつつある。

問い合わせ先：国立研究開発法人 国立環境研究所 地域環境保全領域 海域環境研究室
牧 秀明 e-mail hidemaki#nies.go.jp (#はアットマークを入れる)

<p>公共用水域水質測定等のデータの特徴</p> <ul style="list-style-type: none"> ・長期（30～40年以上）、定点、定期的（毎月、もしくは2～3ヶ月に1回程度）に観測 ・地点により調査日時は異なる <p style="text-align: center;">↓</p> <ul style="list-style-type: none"> ・同一年月のデータを比較しても、調査日の天候や潮時など違いによる影響がデータに含まれる 	<p>季節調整による水温・水質の経時変化の解析</p> <p style="text-align: center;">水温・水質の時系列データの構造をモデル化（成分分離）</p> <p>観測データ=トレンド成分+季節（月ごと）成分+不規則（ノイズ）成分</p> <p style="text-align: center;"> 長期変動 周期変動 短期的変動 </p> <p>トレンド成分：年平均値に近い値を示す</p> <p>トレンド成分+季節成分：天候などの影響を除いた各地点でのその年月の平均的な値を示す</p>
<p>季節調整モデル</p> <p style="text-align: center;">観測値：$y(t) = T(t) + S(t) + I(t)$</p> <p style="text-align: center;">各変動成分の定義</p> <p>○トレンド成分 $T(t-1) - 2T(t) + T(t+1) \sim i.i.d.*$</p> <p>○季節（月ごと）成分 $S(t) + S(t-1) + \dots + S(t-11) \sim i.i.d.*$</p> <p style="text-align: right;">*独立同一分布</p> <p>○不規則（ノイズ）成分 $I(t) = y(t) - T(t) - S(t)$</p> <div style="display: flex; align-items: center;">   </div>	<p>季節調整による時系列データの処理</p> <div style="display: flex; align-items: center;">  <div style="margin-left: 10px;"> <p>ノイズ成分を除いた各年月の平均的な値</p> <p>↓</p> <p>トレンド成分+季節成分</p> <ul style="list-style-type: none"> ○トレンド成分 <ul style="list-style-type: none"> ・大局的な濃度レベル ・年平均値に近い ○季節成分 <ul style="list-style-type: none"> ・毎年繰り返される周期変化 ・日射量、雨季などの影響 ○ノイズ成分 <ul style="list-style-type: none"> ・突発的事象の影響 ・調査時の気象、潮時等 </div> </div>

図作成：東京都環境科学研究所 安藤 晴夫

【準備するもの】

- R for Windows (バージョンは 4.x.x 以上) がインストールされた Windows パソコン
- Rのダウンロードサイト <https://cran.r-project.org/>
<https://cran.r-project.org/bin/windows/base/>
- TIMSAC for R package (バイナリファイル timsac_1.3.8-4.zip (※)) を下記のサイトからダウンロードの上、手持ちのパソコンのR作業用フォルダー（ディレクトリ）内に保存する：<https://jasp.ism.ac.jp/ism/timsac/>

※2023年10月時点での最新バージョン、書庫 Zip ファイルは解凍する必要は無い。

【時系列データの準備と整理、データが存在するセル行・列総数の事前把握】

Excel で以下のように水温・水質（COD, 透明度, TN・TP, DO [表層・底層] 等の季節的変動を示すものの）の時系列データを「**1列目：年 (yyyy), 2列目：月 (mm), 3列目以降：測定値**」(*) に並び換えて整列・整理しておき、c s v 形式のテキストファイルとして、上記の timsac_1.3.8-4.zip ファイルが保存された R 作業用フォルダー（ディレクトリ）内に保存しておく。

- * 最上段／一行目の項目 (yyyy、mm、ローマ字で記された測定地点名) は「全て」半角のアルファベットや数字のみで表し、数字は冒頭には記さないこと。全角文字（日本語）や半角のピリオド（ドット）、コンマ、コロン、セミコロン、ハイフンやアットマーク、アスタリスク、スラッシュ、カッコ等の記号 (.,:;-@*"?!#\$%&'=^¥/+[]{}|~<>()) を含めないこと。

下の例はファイル名：“IbrkPWWT.csv”（茨城県による公共用水域水質測定における鹿島港内と周辺海域における7地点での上層の1972年5月以降の毎月測定水温データ）であり、データ欠測箇所には「-999」という欠測を示すダミー数が記されている。

一行目の測定地点名は半角アルファベット（ローマ字）で表し、冒頭には数字を記さないこと

列総数の確認
この場合は（A列～I列）9

欠測箇所にはダミーとして
-999を入れる

1	yyyy	mm	Chuokoro	Aouhama1	Todenoki1	Todenoki2	Shitthahama	Minamikoroiriguchi	Aouhama2
2	1972	5	14.1	14.5	16	15	15	15	14.6
3	1972	6	-999	-999	-999	-999	-999	-999	-999
4	1972	7	24	22	22	22.5	22	24.3	22.5
5	1972	8	-999	-99	-999	-999	-999	-999	-999
6	1972	9	25.8	24.5	24.5	24.3	24.7	27.2	26.7
7	1972	10	22.6	22	21.4	21.4	21.8	23.4	22
8	197						15.7	19.7	18.2
9	197						13.5	13.4	-999
10	197						8.7	11.4	8.8
11	197						10.5	11.6	8.9
12	1973	3	12.8	12	-999	-999	-999	-999	13.1
13	1973	4	17	15.8	15.9	16.6	16	17.3	16.9
14	1973	5	18.2	15.6	16	16	17.3	17	16
15	1973	6	21.7	21.6	24.8	21.5	22.9	22.3	22.2
16	1973	7	20.9	20.3	24.4	21.4	21.4	21.8	19.8
17	1973	8	21.7	23.4	26.2	23.3	23.2	22.8	22.1
18	1973	9	21.8	22.1	-999	-999	-999	23.2	23
19	1973	10	22.6	20	23.3	20	20.4	23.5	20.2
20	1973	11	17.8	17.3	19.8	18.2	17	19.3	17.4
21	1973	12	12.5	12.2	17	11.6	11.4	13.4	17.3
22	1974	1	11	14	14	14	13.2	12	13
23	1974	2	11	9.6	14	11	10.5	12.1	18
24	1974	3	13.2	13.2	13.8	12.9	13	13.9	21
25	1974	4	15.4	15	16.2	14.2	14.6	16.6	20.7
26	1974	5	16.2	16.7	17.5	14.5	14	17.9	20.5
27	1974	6	18	15.8	18	16.2	16	17.9	17
28	1974	7	19.1	19.8	21.2	19.6	20.2	19.6	25

上図のようにデータを整理したら「ファイル名と、**列総数（この場合 A 列～I 列までの 9 列）とデータ行総数（この場合 1 行目の項目を記した行は省いて 576 行 - 1 行 = 575 行となる）」を確認しておく。**

データ行総数の確認
 この場合 $576 - 1 = 575$
 （1 行目は数値データでなく
 項目[ヘッダー]情報なので省く）

	A	B	C	D	E	F	G	H
569	2019	8	25.9	25.9	25.6	24.7	24.6	24.6
570	2019	9	25.1	25.7	24.3	23.9	23.8	23.8
571	2019	10	20.3	20.6	19.2	19.3	19.2	19.2
572	2019	11	21.6	21.3	21	20.9	20.9	20.9
573	2019	12	16.5	16.6	16.9	16.9	16.9	16.9
574	2020	1	13.9	14.1	14.1	14.1	14.1	14.1
575	2020	2	14.7	14.7	14.6	13.1	13.1	13.1
576	2020	3	14.7	15.5	14.7	14.5	14.4	14.4

【季節調整の R スクリプト（プログラム）と R 上の操作】

R を立ち上げ、下図のように左上の [ファイル] - プルダウンメニューから [ディレクトリの変更...] で、データの csv ファイルと timsac_1.3.8-4.zip ファイルがある R 作業用フォルダー（ディレクトリ）を指定する。

RGui (64-bit)

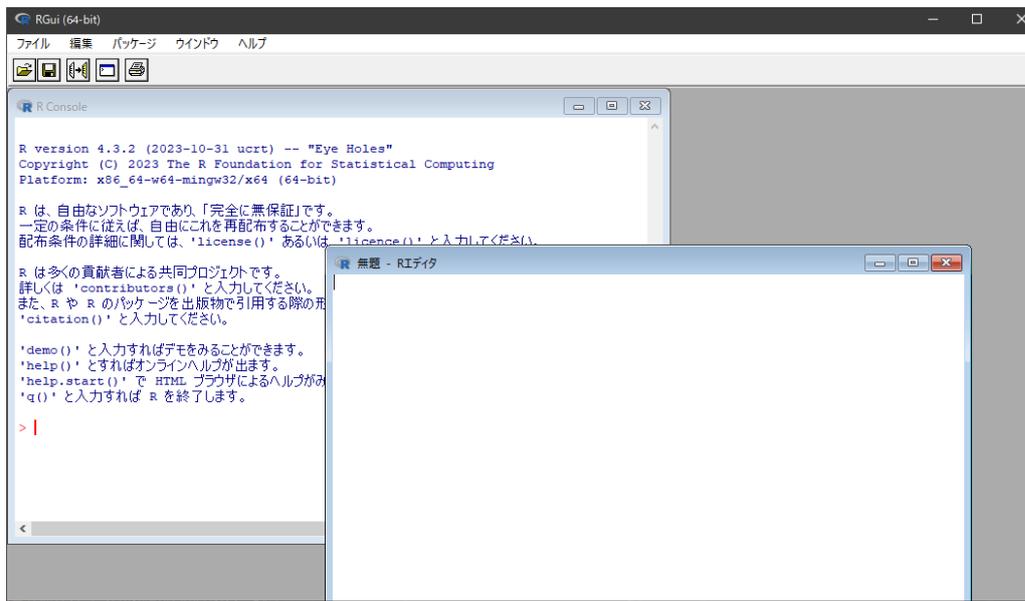
- ファイル
- 編集
- 閲覧
- その他
- パッケージ
- ウィンドウ
- ヘルプ

R コードのソースを読み込み...
 新しいスクリプト
 スクリプトを開く...
 ファイルの表示...
 作業スペースの読み込み...
 作業スペースの保存... Ctrl+S
 履歴の読み込み...
 履歴の保存...
ディレクトリの変更...
 印刷... Ctrl+P
 ファイルを保存...
 終了

```

> |
demo() を実行すれば R を終了することができます。
'help()' とすればオンラインヘルプが出来ます。
'help.start()' で HTML ブラウザによるヘルプがみられます。
'q()' と入力すれば R を終了します。
  
```

[ファイル] -プルダウンメニューから [新しいスクリプト] を選んで、下図のように白紙（無地）の「無題」スクリプトウィンドウを出して、



下図のようなプログラムを作成する（配布のテキストファイル“IbrkPWWT.txt”内の記述されたスクリプト（プログラム）内容を基に適宜修正する）。

```

install.packages("timsac_1.3.8-4.zip", repos="http://www.stat.tku.ac.jp/~timsac/")
library("timsac")

o_comp <- read.csv("IbrkPWWT.csv")

t_comp <- matrix(0,nrow=575,ncol=9)
s_comp <- matrix(0,nrow=575,ncol=9)

colnames(t_comp) <- colnames(o_comp)
colnames(s_comp) <- colnames(o_comp)

s_comp[,2] <- o_comp[,2]

for (i in 3:9) {
  wst <- ts(o_comp[,i], start=c(1973,5), frequency=12)
  rslt <- decom(wst, trend.order=1,ar.order=1, seasonal.order=1, log=FALSE, trade=FALSE, diff=1, miss=-999, omax=0, plot=FALSE)

  t_comp[,i] <- rslt$trend
  s_comp[,i] <- rslt$seasonal}

write.csv(t_comp, "IbrkPWWTTr.csv", quote=FALSE, row.names=FALSE)
write.csv(s_comp, "IbrkPWWTSS.csv", quote=FALSE, row.names=FALSE)

```

解析を行うデータファイル名の指定
ここでは先に保存した
“IbrkPWWT.csv”を指定

解析を行うデータファイルのワークシートの総列数の指定
ここでは9を指定

解析を行うデータファイルの測定値データ総行数の指定
ここでは575を指定

総列数の指定
ここでは9を指定

解析の諸パラメータの指定：
ここでは frequency*は12
(毎月測定データなので)、
trend.order は1, ar.order は1,
seasonal.order は1を指定

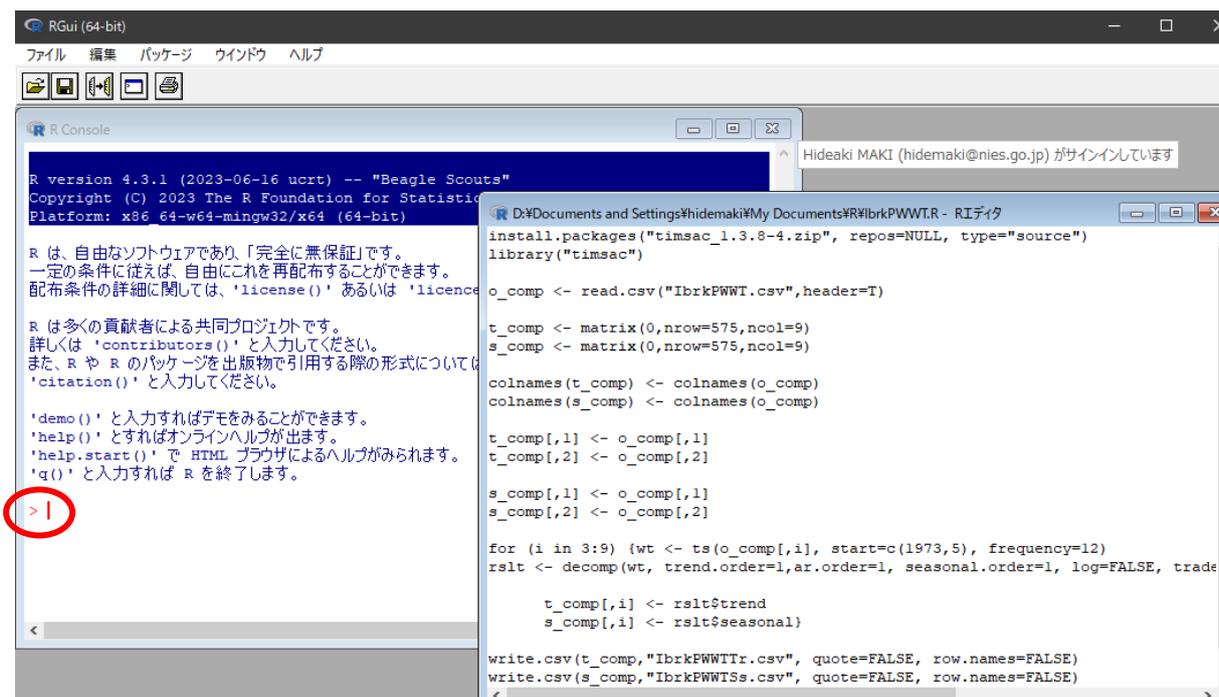
データ欠測 (miss) のダミー数：-999

s_comp, の後ろに、プログラム実行後に出力される季節 (月) 成分の csv ファイル名を指定
ここでは “IbrkPWWTSS.csv” を指定

t_comp, の後ろに、プログラム実行後に出力されるトレンド成分の csv ファイル名を指定
ここでは “IbrkPWWTTr.csv” を指定

- * 上図内の frequency は毎月測定されているデータを取り扱う場合は 12 とし、隔月（年間 6 回）測定データの場合は 6、四半期（年間 4 回）測定データの場合は 4、例外的に月 2 回（年間 24 回）測定データの場合は 24 を、それぞれ指定する。

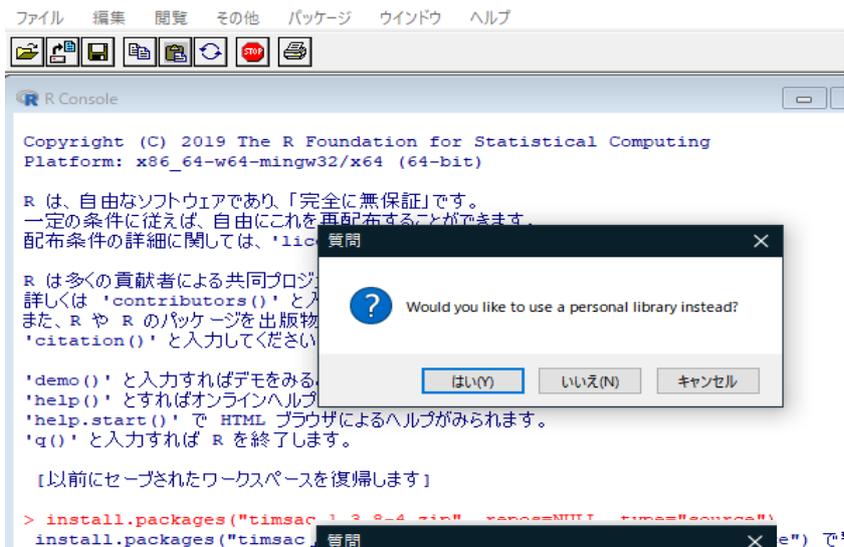
上図の txt ファイル（メモ帳等のエディター）上で作成・修正されたプログラムを全文コピーして、R 上で予め開いておいた「無題」スクリプトウィンドウ状に下図のように貼り付け（ペーストす）る。



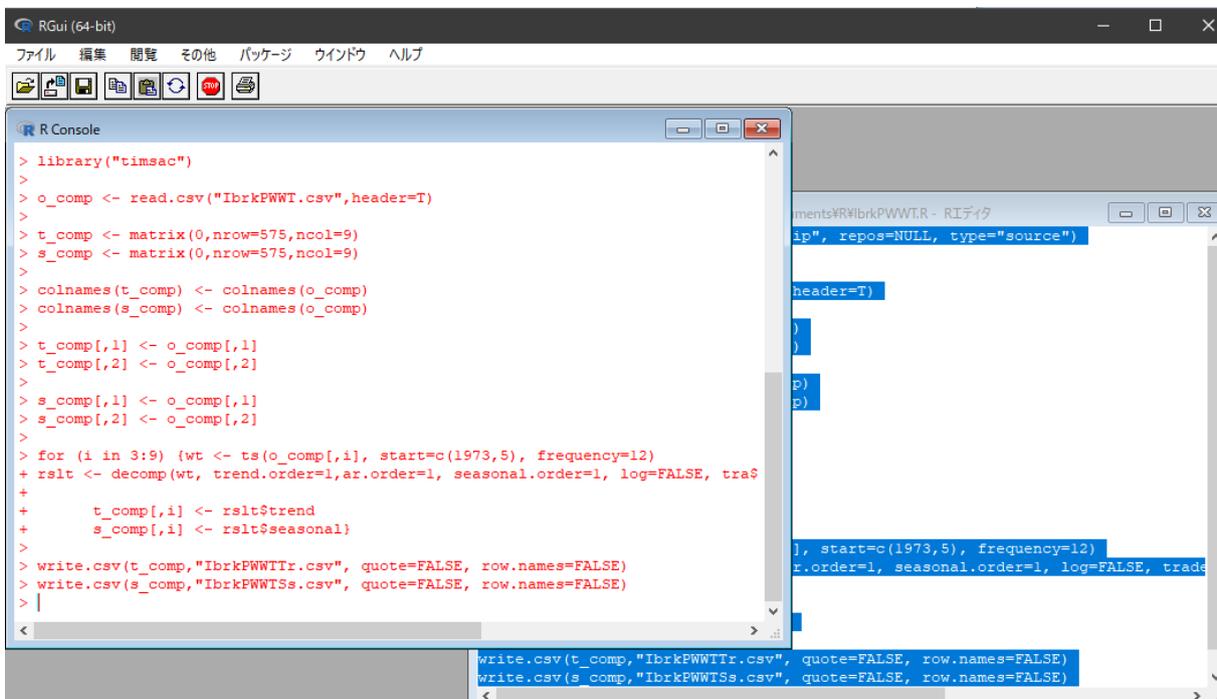
また上図のように、各自でそれぞれ用意・整理された csv ファイル内のデータ行列総数に応じて変更し、[ファイル] –プルダウンメニューから [保存] で出来たスクリプトファイルを保存しておく。

スクリプト（プログラム）ファイルが完成したら、R 上の別ウィンドウの「R console」内の文字列最下段上に示された赤いプロンプト > | の箇所に、スクリプト（プログラム）全文をコピーしてペーストする。すると自動的に R 上でプログラムが開始される。

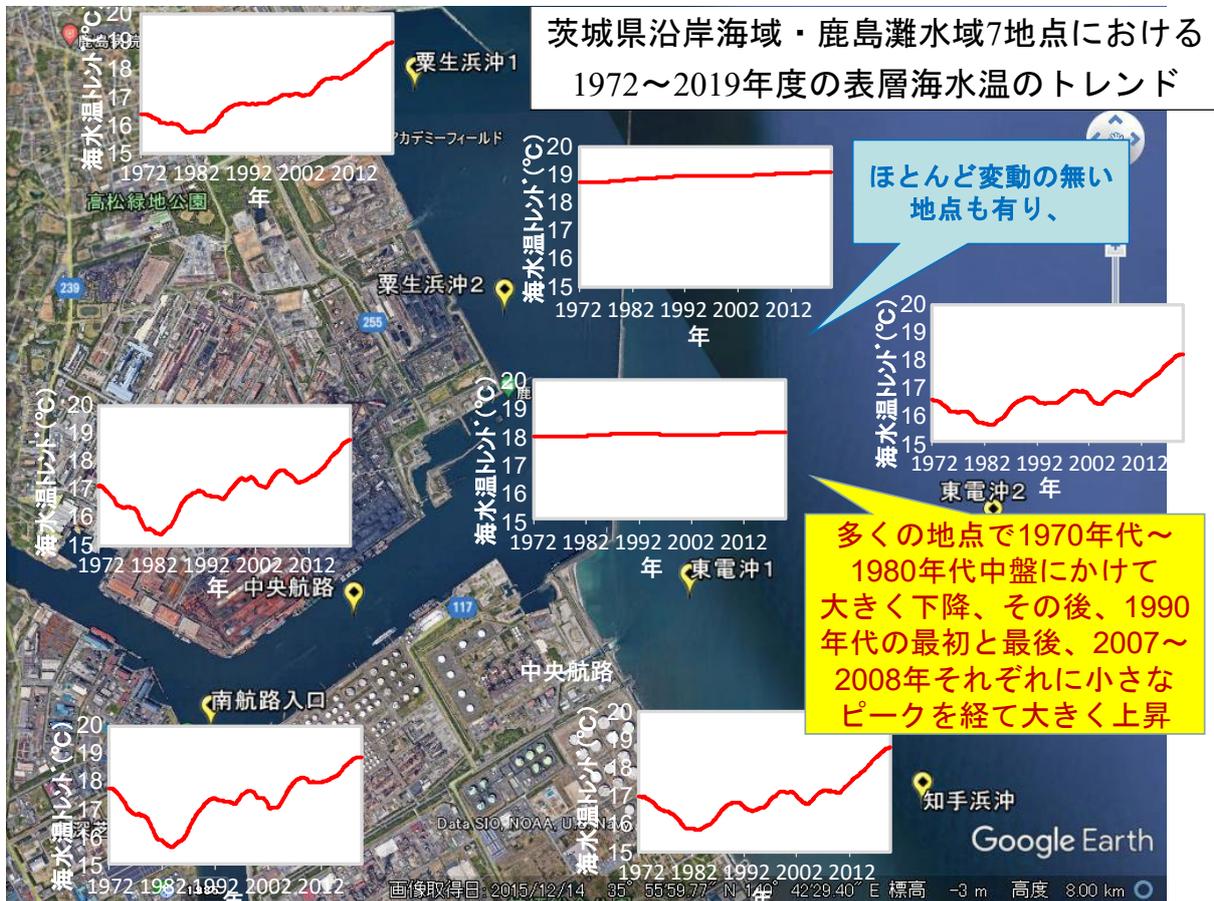
本プログラムを timsac_1.3.8-4.zip がインストールされていないパソコンで初めて動作させると、下図のような英語で記された 2 つのダイアログボックスが出てくるので、両方も「はい (Y)」をクリックする。



プログラムが終了すると上図のように表示されるので、最後に【リターンキーを押して改行すると】、指定したR作業用フォルダー（ディレクトリ）内にトレンド成分（ファイル名：“IbrkPWWTTr.csv”）と季節（月）成分（ファイル名：“IbrkPWWTSs.csv”）それぞれが計算された csv ファイルが指定されたフォルダー内で生成されているので確認する。



以上、サンプルファイル（茨城県鹿島港周辺海域での公共用水域水質測定における表層海水温データ）から推定された長期変動トレンド成分（ファイル名“IbrkPWWTTr.csv”）をグラフ化したものが下図となる。



以上述べた季節調整により、全国各地の沿岸海域の水温や水質の長期変動トレンドを同様に抽出した例は、地方公共団体環境研究機関等と国立環境研究所との共同研究（Ⅱ型）[「沿岸海域における新水質環境基準としての底層溶存酸素（貧酸素水塊）と気候変動の及ぼす影響把握に関する研究報告書（令和2（2020）～令和4（2022）年度実施）」](#)にて公表中である。

【RS-decomp】

以上のようにRを使うことなく、ブラウザで季節調整を簡便に行える統計数理研究所提供の Web Decomp（現在、廃止）の代わりに現在、RS-decomp が以下の URL でテスト公開されている：<https://jasp.ism.ac.jp/RS-Decomp/>

複数列の時系列データの季節調整は同時に出来ないが、一つ（単系列）の時系列データの季節調整を簡便に行えるのが利点である。

※欠測データは NA と記しておけば、季節調整を行うことが出来る。

ブラウザ上で下のウィンドウが示されたら、左側のメニューの「解析」をクリックする。

RS-Decomp

🏠 ホーム

📄 解析

📖 ヘルプ

Time Series Decomposition with R

The Institute of Statistical Mathematics

Decompは Kitagawa and Gersch (1984) により提案された状態空間モデルを用いた季節調整法である。この手法によりいくつかの成分に分解できる。標準モデルは、5つの成分に分解した以下の式で表される;

$$y_t = T_t + AR_t + S_t + TD_t + W_t$$

ここで T_t はトレンド成分, AR_t は AR成分, S_t は季節成分, TD_t は曜日効果, W_t はノイズ成分である。

実行例

下の図は WHARDデータを使って解析したプロット出力の例であり、右下段のSAが原データから季節要因を除いたデータを示している。

Original and Trend

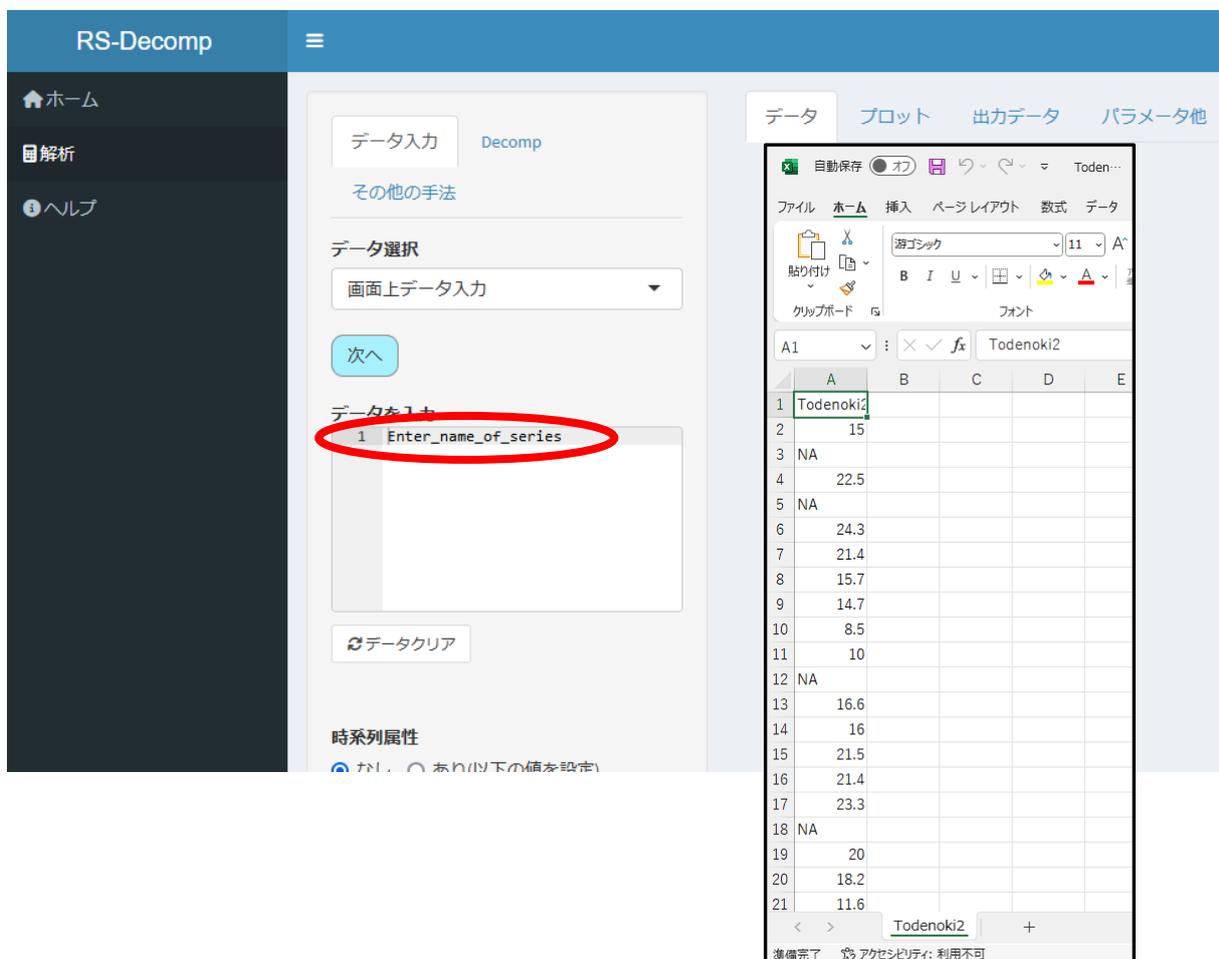
Seasonal

すると下図のようなデータ入力のウィンドウに切り替わるので、データ選択のサンプルデータ: BLSALLFOOD の箇所をクリックするとプルダウンメニューが示されるので、「ウィンドウ上データ入力」を選択し水色の「次へ」のボタンをクリックする。



すると下図のような「データを入力」と記された箇所の下に「1 Enter_name_of_series」と最上行に示されたボックスが表示されるので、季節調整を行いたい Excel ワークシート上のデータ列を全てコピーして、「1 Enter_name_of_series」にペーストする。

※この場合、Excel ワークシート上の欠測箇所には半角で NA と入れておく (-999 は不可)。



Excel ワークシート上のデータをコピー＆ペーストすると下図のようになるので、その下の「時系列属性」の選択肢の内、右側の「あり(以下の値を設定)」の方にチェックマークを入れる。

「周期」は、毎月（年間 12 回）測定データであれば「12（月次）」、3 ヶ月おき（年間 4 回）測定データであれば「4（四半期）」、月 2 回（年間 24 回）測定データであれば「24（毎時）」を選択する（※隔月（年間 6 回）測定データのための選択肢は無い）。

次に「開始年」と「開始位置」を元データに合わせて選択して「実行」をクリックする。

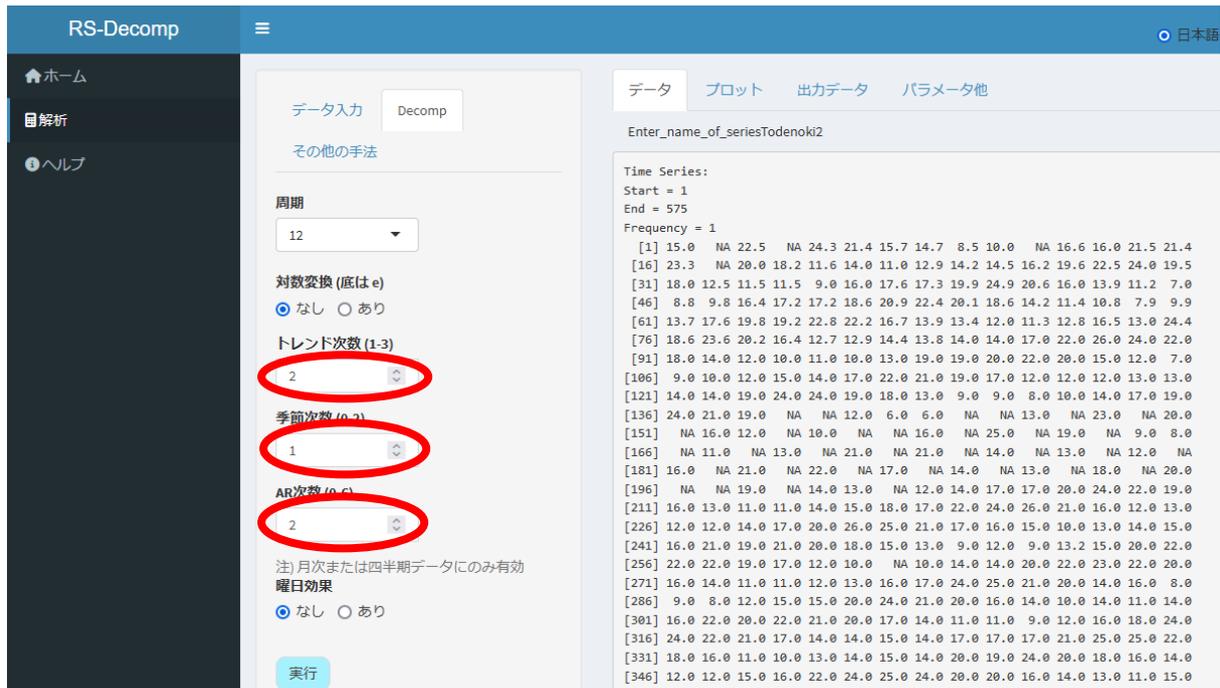
The screenshot shows a data input window with the following elements highlighted by red circles:

- The '時系列属性' (Time Series Properties) section, where the radio button for 'あり(以下の値を設定)' (Yes (set the following values)) is selected.
- The '周期' (Period) dropdown menu, which is set to '12(月次)' (12 (Monthly)).
- The '開始年' (Start Year) dropdown menu, which is set to '1971'.
- The '開始位置' (Start Position) dropdown menu, which is set to '2'.
- The '実行' (Execute) button at the bottom.

The data input table is as follows:

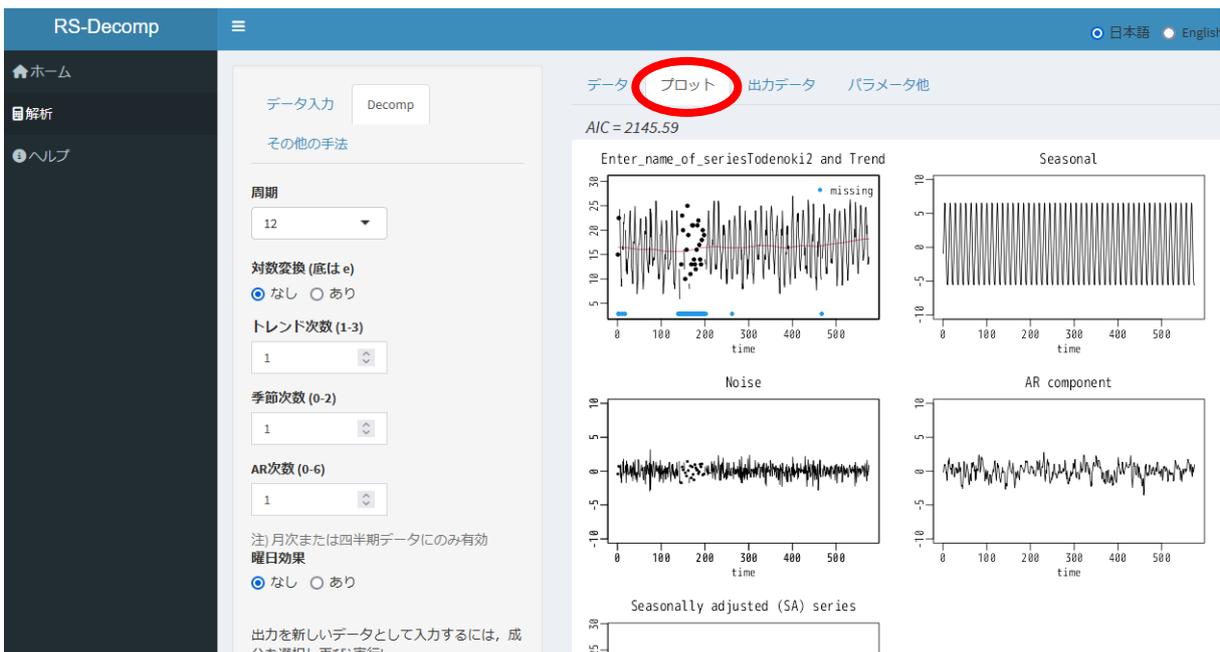
Enter_name_of_seriesTodenoki2
1
2
3 NA
4 22.5
5 NA
6 24.3
7 21.4
8 15.7
9 14.7
10 8.5
11 10

すると下図のようなウィンドウに切り替わるので、「対数変換（底は e）」と「曜日効果」は「なし」の方にチェックマークを入れ、「トレンド次数 (1-3) : デフォルトは 2」、「季節次数 (0-2) デフォルトは 1」、「AR 次数 (0-6) デフォルトは 2」はそれぞれ 1 を選択し「実行」をクリックする。



すると下図のようなウィンドウに入れ替わる。ここで左上の「Enter_name_of_series and Trend」は季節調整に供した元（原）データと推定されたトレンド（赤線で表示）のグラフであり、missing は元（原）データでの欠測箇所（NA と示した箇所）である。

その右の「Seasonal」は季節調整により推定された季節成分のグラフであり、その斜め下、左中段の「Noise」はノイズ、その右の中段「AR component」は自己回帰成分、その斜め下、左最下段の「Seasonally adjusted (SA) series」は季節調整済み系列のグラフで、元（原）データから「Seasonal」（季節成分）を差し引いたものである。



以上の5つのグラフでプロットされた数値データは下図のウィンドウの真ん中右寄りの箇所のタブで「出力データ」を選択すると表示され、右端の「csvダウンロード」ボタンをクリックすると csv ファイルとしてパソコン内に保存される。

Enter_name_of_series	Todenoki2	trend	seasonal	noise	AR	SA
1	15	16.51	-0.921	-0.418	-0.171	15.921
2	NA	16.513	1.302		0.335	16.848
3	22.5	16.515	4.32	0.716	0.949	18.18
4	NA	16.513	6.532		0.837	17.35
5	24.3	16.511	6.303	0.493	0.993	17.997
6	21.4	16.507	3.372	0.763	0.758	18.028
7	15.7	16.498	0.5	-0.966	-0.331	15.2
8	14.7	16.494	-2.633	0.975	-0.136	17.333
9	8.5	16.485	-4.903	-1.695	-1.387	13.403
10	10	16.486	-5.51	-0.333	-0.643	15.51

RS-Decomp に csv ファイルを直接読み込む方法

下記のウィンドウ内の「データ選択」で「ファイル読み込み (csv テキスト)」を選択して「次へ」をクリックすると下のようにウィンドウの表示が切り替わるので「Browse」をクリックして、季節調整を行いたい csv ファイルを選んで読み込む。

下のウィンドウでは Upload complete と表示されており、Todenoki2.csv というファイルが読み込み完了したことを示している。以降の操作は上記と同じである。

ファイルを選択

Browse... Todenoki2.csv

Upload complete

時系列属性

なし あり(以下の値を設定)

周期

12 (月次) ▼

毎時データを除く

開始年 開始位置

2000 ▼ 1 ▲▼

実行

謝辞

本解説書作成に当たり、基本的な創案とご指導、並びに概念図のご提供頂きました、大学共同利用機関法人 情報・システム研究機構 統計数理研究所の柏木宣久名誉教授と副所長・川崎能典教授、公益財団法人東京都環境公社東京都環境科学研究所の安藤晴夫氏に感謝申し上げます。

参考文献・web サイト

Akaike, H., Ozaki, T., Ishiguro, M., Ogata, Y., Kitagawa, G., Tamura, Y.-H., Arahata, E., Katsura, K. and Tamura, Y. (1985) TIMSAC-84 Part 1 & Part 2, Computer Science Monographs, No.22 & 23, The Institute of Statistical Mathematics, Tokyo.

Kitagawa, G. (1981), A nonstationary time series model and its fitting by a recursive filter, *Journal of Time Series Analysis*, Vol. 2(2), 103-116.

北川源四郎 (1986) 時系列の分解 プログラム DECOMP の紹介. 統計数理, 34 (2), 255-271.

Rのダウンロードサイト <https://cran.r-project.org/>

<https://cran.r-project.org/bin/windows/base/>

(2024年12月2日時点有効確認)

TIMSAC for R package (バイナリファイル timsac_1.3.8-4.zip ダウンロードと説明に関するサイト) : <https://jasp.ism.ac.jp/ism/timsac/>

(2024年12月2日時点有効確認)

RS-decomp : <https://jasp.ism.ac.jp/RS-Decomp/>

(2024年12月2日時点有効確認)

二宮勝幸・柏木宣久・安藤晴夫 (1996) 東京湾における水温と塩分の空間濃度分布の季節別特徴.
水環境学会誌, 19, 480-490.

Kashiwagi, N., Ninomiya, K., Ando, H. and Ogura, H. (2003) A space-time state-space modeling of Tokyo Bay pollution, Sustainable Environments: A Statistical Analysis (eds. A. K. Ghosh, J. K. Ghosh and B. Mukhopadhyay) , Oxford University Press, New Delhi, 42-62.

Ando, H., Maki, H., Kashiwagi, N. and Ishii, Y. (2021) Long-term change in the status of water pollution in Tokyo Bay: recent trend of increasing bottom-water dissolved oxygen concentrations. *Journal of Oceanography*, 77, 843-858.

地方公共団体環境研究機関等と国立環境研究所との共同研究（Ⅱ型）沿岸海域における新水質環境基準としての底層溶存酸素（貧酸素水塊）と気候変動の及ぼす影響把握に関する研究報告書（令和2（2020）～令和4（2022）年度実施）

<https://www.nies.go.jp/kenkyu/chikanken/R2-R4.pdf>

(2024年12月2日時点有効確認)