

# 周期時系列の統計解析

## (5) 単回帰分析

nino

2018年3月11日

周期時系列に季節調整を行わないで単回帰分析を適用すると、回帰係数には周期成分の影響が加わる。ここでは、周期時系列をコサイン関数モデルで近似し単回帰分析によりモデルの回帰係数を求め、周期成分の影響を検討した。また、その結果を気温時系列に当てはめ、課題等について考察した。

### 気温時系列とコサイン関数モデル

第2報の結果を利用するので、その一部を再掲する。

気温時系列は、1991年1月から2000年12月までの10年間に於ける月平均気温（横浜）である。1990年代は比較的気温上昇率の高い期間である。また、気温時系列を次のコサイン関数モデル $Y(t)$ で近似し、最小二乗法により振幅等のパラメータを求めた（表1）。

$$Y(t) = H \cos[\omega(t + \phi)] + M + Kt + \varepsilon_t$$

ここで、パラメータは、 $H$ ；振幅（ $^{\circ}\text{C}$ ）、 $T$ ；周期（月）、 $\omega = 2\pi/T$ ；角周波数、 $\phi$ ；位相（月）、 $M$ ；平均（ $^{\circ}\text{C}$ ）、 $K$ ；トレンド（ $^{\circ}\text{C}/\text{月}$ ）、 $\varepsilon_t$ ；誤差である。なお、位相 $\phi$ の単位はラジアンではなく、月である。

表 1

振幅 $H(^{\circ}\text{C})$	9.9
平均 $M(^{\circ}\text{C})$	15.5
位相 $\phi(\text{月})$	4.5
トレンド $K(^{\circ}\text{C}/\text{月})$	0.0064

図1に、気温時系列およびその単回帰分析の回帰直線（点線）と回帰式を、また、表1のパラメータに基づくコサイン関数モデルの時系列も併せて示した。なお、必ずしも正確な表現ではないが、わずらわしさをなくすため、以降は、回帰係数もトレンドに統一することとした。

図 1

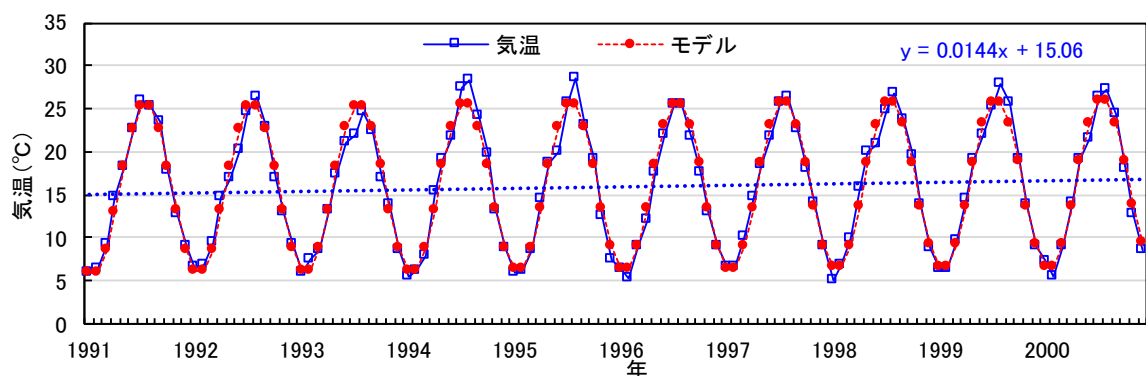


図 1 から、気温は明確な季節変化を示した。トレンドは $0.0144^{\circ}\text{C}/\text{月}$ となり、経年的に少し上昇している。また、モデルは気温時系列とほぼ同様な季節変化を示しており、最低気温は1月と2月、最高気温は7月と8月であった。

一方、表 1 によると、最小二乗法によるモデルのトレンドは $0.0064^{\circ}\text{C}/\text{月}$ であり、単回帰分析による気温時系列のトレンド ( $0.0144^{\circ}\text{C}/\text{月}$ ) より小さかった。この差は周期成分等による影響のためと考えられた。このことは、式 $Y(t)$ が連続変数の場合に $t$ で微分すれば、 $-H\omega\sin[\omega(t+\phi)]+K$ が得られ、時点 $t$ におけるモデルの傾きは本来のトレンド $K$ と周期成分のトレンド $-H\omega\sin[\omega(t+\phi)]$ の和で表されることから明らかである。周期成分等のトレンドは、 $0.0144-0.0064=0.0080$  ( $^{\circ}\text{C}/\text{月}$ ) と推測され、かなり大きな値を示した。また、1月を開始月とした気温時系列の位相 $\phi$ は4.5か月であった。

### コサイン関数モデルのトレンド

実際の周期時系列のほとんどは連続変数ではなく離散変数なので、離散変数のコサイン関数モデルのトレンド $a$ を求める。トレンド $a$ は $Y(t)$ と $t$ の共分散 $S_{Yt}$ および $t$ の分散 $S_{tt}$ を用いると、次式で表される。

$$a = \frac{S_{Yt}}{S_{tt}} = \frac{\frac{1}{n} \sum_{t=1}^n \{Y(t) - \bar{Y}\}(t - \bar{t})}{\frac{1}{n} \sum_{t=1}^n (t - \bar{t})^2} = \frac{\sum_{t=1}^n \{Y(t) - \bar{Y}\}(t - \bar{t})}{\sum_{t=1}^n (t - \bar{t})^2}$$

ここで、データ数 $n$ は整数周期分すなわち $n=jT$  (対象期間が $j$ 年間= $1, 2, \dots$ の正整数) を条件とした。したがって、次式が成り立つ。

$$\sum_{t=1}^n \cos[\omega(t+\phi)] = 0$$

トレンド $a$ を求める準備として、まず、 $Y(t)$ と $t$ の平均値を求めた。

$$\begin{aligned} \bar{Y} &= \sum Y(t)/n = \sum \{A\cos[\omega(t+\phi)] + Kt + M\}/n \\ &= \sum \{Kt + M\}/n \\ &= \{K n(n+1)/2 + nM\}/n \\ &= \frac{K(n+1)}{2} + M \end{aligned}$$

$$\begin{aligned} \bar{t} &= \sum t/n = \{n(n+1)/2\}/n \\ &= (n+1)/2 \end{aligned}$$

次に、トレンド $a$ を分母と分子に分けて求める。トレンド $a$ の分母は、

$$\begin{aligned}
\sum (t-\bar{t})^2 &= \sum (t^2 - 2t\bar{t} + \bar{t}^2) \\
&= \frac{n(n+1)(2n+1)}{6} - 2 \frac{(n+1)}{2} \frac{n(n+1)}{2} + \left(\frac{n+1}{2}\right)^2 n \\
&= n(n+1) \left( \frac{2n+1}{6} - \frac{n+1}{2} + \frac{n+1}{4} \right) \\
&= \frac{n(n^2-1)}{12}
\end{aligned}$$

一方、トレンド  $a$  の分子は、

$$\begin{aligned}
\sum \{Y(t) - \bar{Y}\} (t - \bar{t}) &= \sum \left\{ H \cos[\omega(t + \phi)] + Kt + M - \left[ \frac{K(n+1)}{2} + M \right] \right\} \left\{ t - \frac{(n+1)}{2} \right\} \\
&= \sum \left\{ H \cos[\omega(t + \phi)] + K \left[ t - \frac{(n+1)}{2} \right] \right\} \left\{ t - \frac{(n+1)}{2} \right\} \\
&= \sum \left\{ Ht \cos[\omega(t + \phi)] - H \frac{(n+1)}{2} \cos[\omega(t + \phi)] + K \left[ t - \frac{(n+1)}{2} \right]^2 \right\} \\
&= \sum \left\{ Ht \cos[\omega(t + \phi)] + K \left[ t^2 - (n+1)t + \frac{(n+1)^2}{4} \right] \right\} \\
&= H \sum t \cos[\omega(t + \phi)] + K \left[ \frac{n(n+1)(2n+1)}{6} - (n+1) \frac{n(n+1)}{2} + \frac{n(n+1)^2}{4} \right] \\
&= H \sum t \cos[\omega(t + \phi)] + \frac{Kn(n-1)^2}{12}
\end{aligned}$$

となるが、この式の第1項を求めるのはやや複雑であり、以下のように算出する。  
まず、第1項を和積公式を用いて整理すると、次式が得られる。

$$\begin{aligned}
H \sum t \cos[\omega(t + \phi)] &= H \sum t \{ \cos \omega t \cos \omega \phi - \sin \omega t \sin \omega \phi \} \\
&= H \{ \cos \omega \phi \sum t \cos \omega t - \sin \omega \phi \sum t \sin \omega t \}
\end{aligned}$$

次に、右辺の  $\sum t \cos \omega t$  と  $\sum t \sin \omega t$  をそれぞれ求めるため、次の三角関数の有限級数和の公式（参考文献2）を用いる。

$$\begin{aligned}
\sum_{r=1}^n r \cos rx &= \frac{n+1}{2} \frac{\sin[(2n+1)x/2]}{\sin(x/2)} - \frac{1 - \cos[(n+1)x]}{4 \sin^2(x/2)} \\
\sum_{r=1}^n r \sin rx &= -\frac{n+1}{2} \frac{\cos[(2n+1)x/2]}{\sin(x/2)} + \frac{\sin[(n+1)x]}{4 \sin^2(x/2)}
\end{aligned}$$

ここで、 $r=t$ 、 $x=\omega=2\pi/T$  とおく。また、 $n=jT$  であるから、 $\cos nx = \cos 2j\pi = 1$ 、 $\sin nx = \sin 2j\pi = 0$  となる。これらを上式に代入してまとめると、それぞれ次式が得られる。

$$\sum_{t=1}^n t \cos \omega t = \frac{n}{2}$$

$$\sum_{t=1}^n t \sin \omega t = \frac{-n \cos(\pi/T)}{2 \sin(\pi/T)}$$

さらに、この2式をトレンド  $a$  の分子の第1項に代入して整理すると、次式が得られる。

$$\begin{aligned} H \sum t \cos[\omega(t + \phi)] &= H \left\{ \cos \omega \phi \sum t \cos \omega t - \sin \omega \phi \sum t \sin \omega t \right\} \\ &= H \left\{ \frac{n}{2} \cos \omega \phi - \left[ \frac{-n \cos(\pi/T)}{2 \sin(\pi/T)} \right] \sin \omega \phi \right\} \\ &= H \frac{n \sin[(2\pi/T)(1/2 + \phi)]}{2 \sin(\pi/T)} \end{aligned}$$

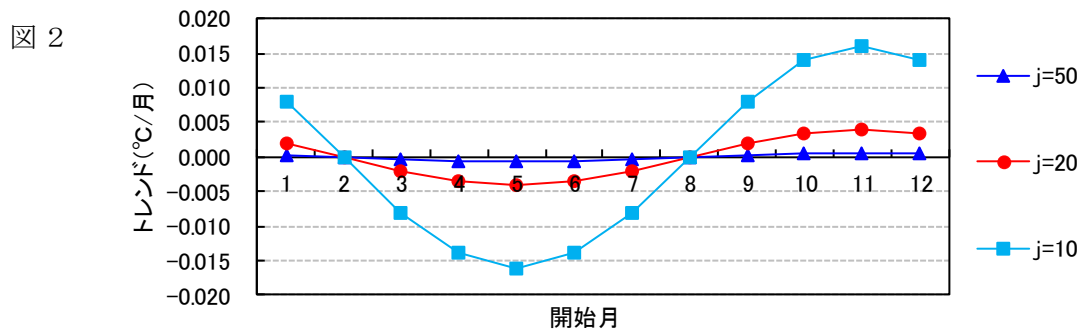
最終的に、この式と先の分母をトレンド  $a$  の式に代入すると、トレンド  $a$  が求まる。

$$\begin{aligned} a &= \frac{n \left\{ \frac{nH \sin[(2\pi/T)(1/2 + \phi)]}{2 \sin(\pi/T)} + \frac{Kn(n-1)^2}{12} \right\}}{\frac{n(n^2-1)}{12}} \\ &= \frac{6H \sin[(2\pi/T)(1/2 + \phi)]}{(n^2-1) \sin(\pi/T)} + K \end{aligned}$$

トレンド  $a$  は周期成分由来のトレンドと本来のトレンド  $K$  の和で表された。周期成分のトレンドは振幅  $H$  に比例、 $(n^2-1)$  に反比例するとともに、周期  $T$  と位相  $\phi$  によっても規定される。前述したように、この式における位相  $\phi$  は、1月を開始月としたモデル(表1)では、4.5 か月であったので、2月を開始月とした場合は  $\phi = 4.5 + 1.0 = 5.5$  か月となる。3月以降も同様である。

### 開始月の違いによる周期成分のトレンドへの影響

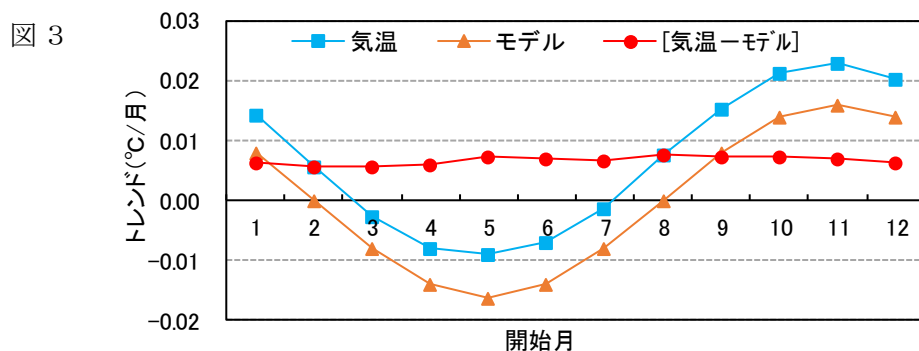
開始月の違いによる周期成分のトレンドへの影響を調べた。図2に、気温時系列に準じて振幅  $H=10^\circ\text{C}$ 、周期  $T=12$  か月としたコサイン関数モデルにおいて、対象期間が  $j$  年間(データ数  $n=j \times T$ ) の3例について開始月  $\phi$  に対するトレンドの関係を示した。この場合は、 $j$  以外は定数なので、各月におけるトレンドは  $(n^2-1) \propto j^2$  に反比例する。



どの期間でも、トレンドは5月に最低、11月に最高、そして、2月と8月に0°C/月となる傾向を示した。開始月が2月と8月であれば、周期成分によるトレンドへの影響はないことが示唆された。また、期間が長くなればトレンドは大きく低下した。

各期間についてみると、 $j=10$  年間の場合は、開始月が1月の時に周期成分のトレンドは0.0080°C/月であるが、5月の時は最低値-0.0161°C/月、11月の時は最高値0.0161°C/月を示した。気温時系列の最小二乗法によるトレンドは0.0064°C/月であったので、かなり大きな値である。一方、 $j=20$  年間の周期成分のトレンドは-0.0040~0.0040°C/月、 $j=50$  年間では-0.00064~0.00064°C/月の範囲にあった。

具体例として、10年間の気温時系列とモデルを比較検討した。図3に、開始月に対する気温とモデルのトレンド（°C/月）および気温とモデルのトレンドの差[気温-モデル]を示した。[気温-モデル]は、開始月の影響を除去したトレンドに相当する。

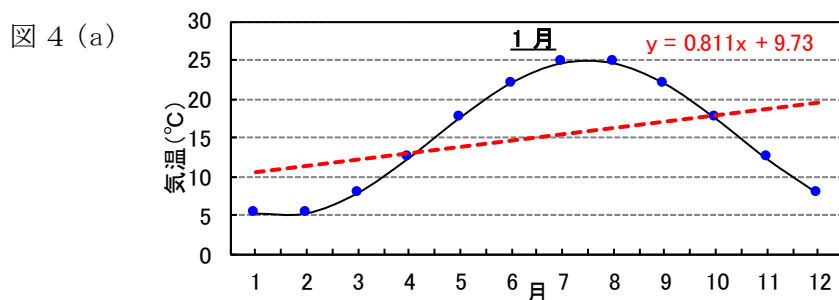


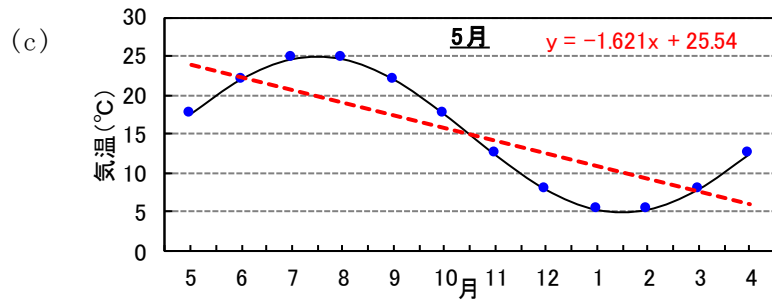
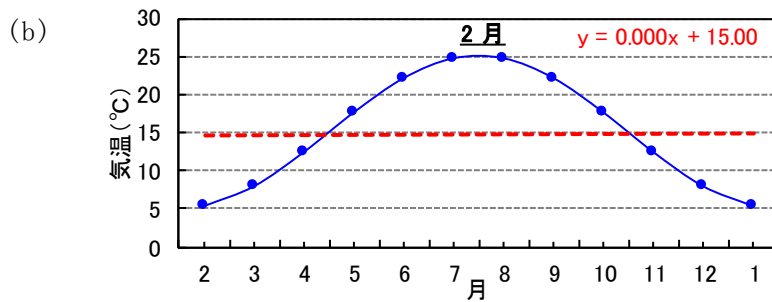
気温の曲線はモデルの理論曲線とほぼ平行しており、予想通りの変化を示した。[気温-モデル]は0.0056~0.0076°C/月の範囲にあった。[気温-モデル]は誤差等のため少し変動しているが、平均すると0.0067°C/月となり、最小二乗法のトレンド(0.0064°C/月)に比較的近い値を示した。

#### 開始月によりトレンドが異なる理由

開始月によってトレンドが異なる理由を調べた。通常は1年間のトレンドを求めることはないが、トレンドの変化傾向は開始月が同じならば短期間でも長期間でも変わらず、また、開始月の違いの影響が明瞭に表れるので、1年間の時系列について説明する。

図4(a~c)に、開始月が1月、2月、5月の1年間の時系列を、また、それらの回帰直線(点線)と回帰式を併せて示した。





まず，開始月が1月(a)のトレンドは  $0.811^{\circ}\text{C}/\text{月}$  と正值を示した．この理由を定性的に説明すると以下ようになる．中央月である6.5月を境にして，左半分(1~6月)のデータの平均値よりも右半分(7~10月)のデータの平均値のほうが大きいため，両者を結ぶ直線の傾きは正值となる．次に，開始月が2月(b)の場合，中央月(7.5月)の左半分と右半分のデータの平均値が同値であるため，傾きは  $0.000^{\circ}\text{C}/\text{月}$  となる．最後に，開始月が5月(c)の場合は，トレンドは  $-1.621^{\circ}\text{C}/\text{月}$  であった．これは，中央月(10.5月)の左半分のデータの平均値のほうが右半分のデータの平均値よりも大きく，また，その差は1月の場合に比べて大きいため，両者を結ぶ直線の傾きは大きな負値を示したと考えられる．

このように周期時系列は周期成分自身のトレンドを有し，また，それは開始月によって変化する．これが，気温などの周期時系列を季節調整により周期成分を消去あるいは低減して，より正確なトレンドを求める理由である．

しかし，移動平均などのように季節調整後のデータにも，振幅は小さくなるものの，数年から十年程度の不定周期の循環変動成分が残る．循環変動成分は，これまでの検討結果から推察されるように，トレンドに影響する可能性がある．

### 周期成分のトレンドの最大値と最小値

循環変動成分は必ずしも周期性があるわけではなく，また，開始月を決めることはできないので，モデルを使ってトレンドへの影響を把握することは難しい．そこで，循環変動を周期時系列と仮定して，これまでの知見をもとに，循環変動成分のトレンドへの影響の概略を把握することを試みた．具体的には，トレンドの変動幅(最小値~最大値)を求め，それを周期時系列モデルに当てはめてトレンドへの影響程度を推定する方法であり，トレンドの確からしさをおおよそ把握できると考えられる．

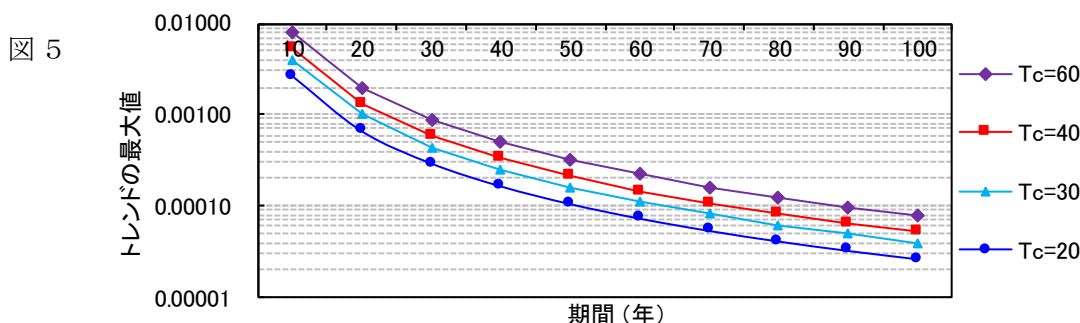
トレンド  $a$  において， $-1 \leq \sin[(2\pi/T)(1/2+\phi)] \leq 1$  であるから，循環変動成分の振幅

を  $H_c$ , 周期を  $T_c$  とすると, そのトレンドの最大値  $a_{\max}$  は,

$$a_{\max} = \frac{6H_c}{(n^2 - 1)\sin(\pi/T_c)}$$

となる. また, 最小値  $a_{\min} = -a_{\max}$  なので, トレンドの変動範囲は  $-a_{\max} \sim a_{\max}$  である. よって, 最大値  $a_{\max}$  を求めれば, その影響を推定できる.

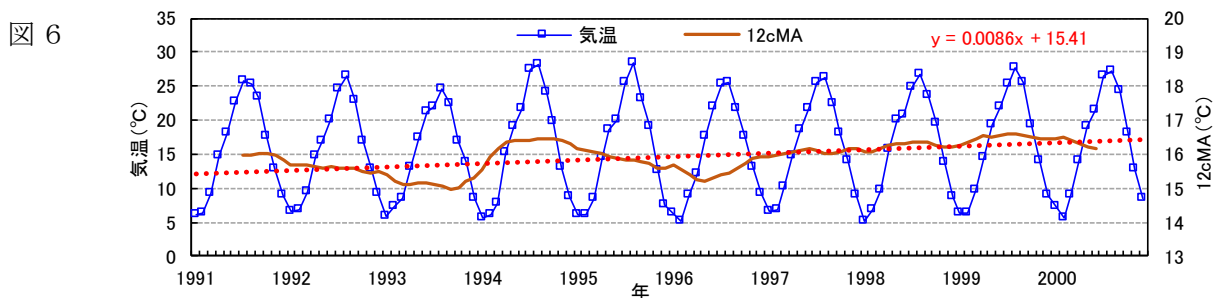
そこで, トレンドの最大値  $a_{\max}$  の式において, 循環変動成分の振幅を  $H_c = 1$  とした場合の周期が  $T_c = 20, 30, 40, 60$  か月 (各々 約 1.7, 2.5, 3.3, 5.0 か年) の 4 例について, 期間が  $j = 10 \sim 100$  年間 ( $n = 12 \times j$ ) におけるトレンドの最大値  $a_{\max}$  を調べた (図 4).



トレンドの最大値は, 期間  $j$  が長いほど, 循環変動の周期  $T_c$  が短いほど小さくなる傾向を示した. したがって, 循環変動成分によるトレンドへの影響を低減するためには, 時系列データはなるべく長い期間を用いることが肝要である. また, 周期の長い循環変動成分が存在する場合はその影響に注意する必要がある.

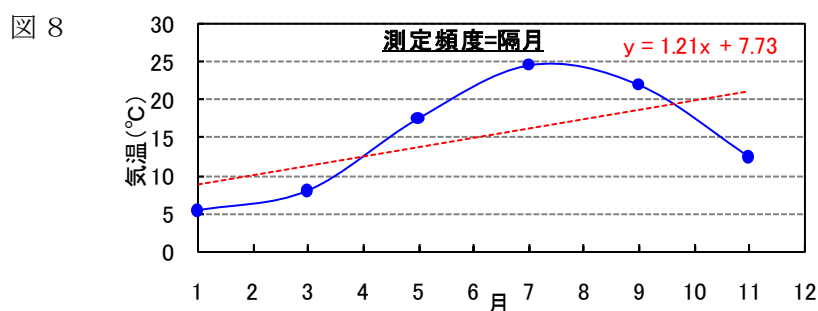
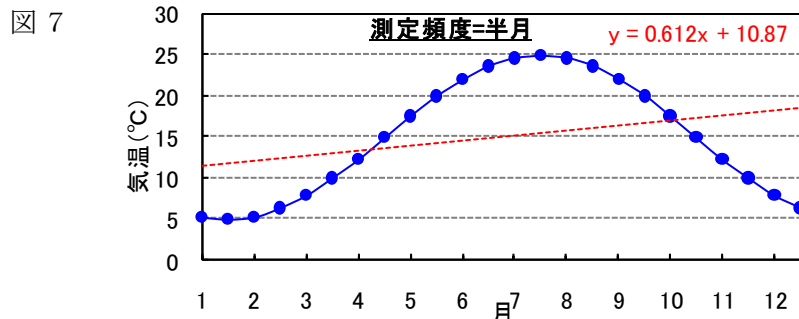
具体例として, 10年間の気温時系列の12cMA (図 6) についてみると, 12cMAのトレンドは  $0.0086^\circ\text{C}/\text{月}$  であった. 12cMAのなかで最大振幅を示す循環変動成分は, 周期が約 3 年 (36 か月) で, その最大振幅は約  $0.7^\circ\text{C}$  であった. 一方, 図 5 によると, 10年間における  $T_c = 30 \sim 40$  か月のトレンドの最大値は, 循環変動成分の最大振幅 (約  $0.7^\circ\text{C}$ ) を考慮すると,  $0.0028 \sim 0.0037^\circ\text{C}/\text{月}$  となり, かなり高い値を示した.

これまでの結果から, トレンドは大きい順に, 気温時系列の単回帰 ( $0.0144^\circ\text{C}/\text{月}$ ) > 12cMAの単回帰 ( $0.0086^\circ\text{C}/\text{月}$ ) > 最小二乗法のモデル ( $0.0064^\circ\text{C}/\text{月}$ ) となり, 前 2 者のトレンドは周期成分や循環変動成分の有無による影響を受けている可能性が高い.



### 異なる測定頻度におけるトレンド

これまででは、測定頻度が毎月（12回/年）の場合であったが、測定頻度が半月毎（6回/年）と隔月毎（24回/年）の場合における開始月が1月の1年間のコサイン関数モデルの時系列（図6と図7）について、先述の測定頻度が毎月の場合と比較検討した。



測定頻度が半月毎の場合は、トレンドが $0.612^{\circ}\text{C}/\text{月}$ となり、毎月測定の場合（ $0.811^{\circ}\text{C}/\text{月}$ ）の約3/4であった。一方、測定頻度が隔月毎のトレンドは $1.21$ で、毎月測定の場合の約1.5倍となった。このように、測定頻度が高いほど、周期成分のトレンドへの影響は小さくなった。このことは、循環変動成分によるトレンドへの影響についても同様であると推測されることから、その影響を低減化するためには、なるべく測定頻度を高くする必要がある。

具体例として、10年間の気温時系列（図1）から、隔月のデータ（奇数月）を抽出し、その最初の1年間のトレンドを求めた。その結果、 $1.20^{\circ}\text{C}/\text{月}$ が得られ、モデルの値（ $1.21^{\circ}\text{C}/\text{月}$ ）とほぼ一致した。1年間の時系列ではトレンドによる影響がほとんどないためと考えられた。さらに、10年間の気温時系列の隔月データについて、その6項中心化移動平均（6cMA）によりトレンドを求めたところ、 $0.0116^{\circ}\text{C}/\text{月}$ であった。この値は毎月データの10年間の12cMAのトレンド $0.0086^{\circ}\text{C}/\text{月}$ よりも $0.0030^{\circ}\text{C}/\text{月}$ （約35%）高かった。

#### 参考文献)

- 1) 物理のかぎしっぽ, 周期時系列の統計解析 (2) 最小二乗法による近似 v2  
<http://hooktail.sub.jp/contributions/shuki02.pdf>
- 2) 岩波数学公式 2 級数・フーリエ解析, 森口ほか, 岩波書店 (1987)