

Die Anwendung einer Fourier-Transformation zum Ausgleich von Saisonschwankungen bei der physikalisch-chemischen Charakterisierung von Fließgewässern

Von H. Güttinger

Eidg. Anstalt für Wasserversorgung, Abwasserreinigung und Gewässerschutz (EAWAG),
CH-8600 Dübendorf

Manuskript eingegangen am 4. Juli 1980

ABSTRACT

The application of a Fourier transformation to compensate for seasonal variations in physical and chemical characterizations of streams

The calculation of the mean y_0 , the amplitude A , and the phase t_0 of the periodic regression function $y(t) = y_0 + (A/2) \sin [(t - t_0)2\pi/T]$ is shown for the case of known period length T . An error estimation gives valuable information on how to plan efficient research programs. The theory is exemplified using stream temperature measurements. Applicability of the method to other limnological data, such as oxygen content is briefly discussed.

INHALTSVERZEICHNIS

1. Einleitung	309
2. Theorie	310
2.1 Ansatz	310
2.2 Berechnung der Koeffizienten	312
2.3 Berechnung der Streuungen	313
2.4 Lösung für den Fall, dass einzelne Koeffizienten bekannt sind	313
3. Anwendungsbeispiel	314
4. Diskussion	317
5. Zusammenfassung	318
Summary	319
Résumé	319
Verdankungen	320
Literaturverzeichnis	320
Anhang	320

1. Einleitung

Der Vergleich von periodisch variierenden Phänomenen lässt sich im allgemeinen nur über identische Periodenabschnitte durchführen. Bei einer Literaturarbeit über den Einfluss von chemischen Wasserinhaltsstoffen auf die Makroinvertebratenfauna

von Fließgewässern (Güttinger, in Vorbereitung), tauchte das Problem auf, verschiedene Probenahmestellen bezüglich ihrer Wassertemperaturen zu charakterisieren. Da die Messungen aber nicht aus dem gleichen Zeitraum stammten, schien es kaum sinnvoll, als Mass für eine Durchschnittstemperatur das arithmetische Mittel zu verwenden. Als strahlungsabhängige Grösse zeigt die Temperatur jedoch sowohl im Jahres- als auch im Tagesverlauf angenähert einen sinusförmigen Verlauf. Da periodische Schwingungen mittels Fourier-Analyse approximiert werden können, scheint es naheliegend, die Wassertemperatur als Funktion der Jahreszeit mathematisch darzustellen.

Eine diesbezügliche Literaturrecherche ergab, dass solche Ansätze schon verschiedentlich angewendet worden sind, dass aber die exakte mathematische Berechnung nirgends vollständig expliziert ist. Kuhn [7] verwendet eine trigonometrische Näherung zur Beschreibung von Gleichgewichtstemperaturen, setzt aber als Koeffizienten Schätzwerte ein, die nicht statistisch berechnet wurden. Sinusfunktionen im Zusammenhang mit Wirtschaftsanalysen beschreibt Nicolas [8], ohne auf Details einzugehen. Eine interessante Arbeit von Halberg et al. [5] über diurnale Rhythmen enthält die Schilderung der sogenannten «Cosinor-Methode», welche den Variationsbereich des Vektors von Amplitude und Phase einer Sinusfunktion graphisch darstellt. Die Funktion wird mit der Methode der kleinsten Quadrate ermittelt, die Lösungen sind aber in der Arbeit nicht enthalten. Etwas konkretere Angaben findet man in Bliss [1] im Kapitel 17 über periodische Regressionen. Der Autor zeigt den Lösungsweg für Messreihen, die regelmässig über die ganze Periode verteilt sind.

In dieser Arbeit sollen eine ausführliche Beschreibung der Berechnung der Koeffizienten einer trigonometrischen Regressionsfunktion gegeben sowie die Anwendungsmöglichkeiten in der Limnologie diskutiert werden. Da es sich hier um einfache Sinusschwingungen handelt, werden die Berechnungen nur für das erste Glied einer Fourier-Zerlegung besprochen.

2. Theorie

Die Herleitungen basieren grösstenteils auf dem Kapitel über lineare Regression von Kendall und Stuart [6], weshalb hier dieselbe Symbolik (Tab.1) verwendet wird. Im Textteil wird die kürzere und übersichtlichere Matrixschreibweise vorgezogen, und die vollständig ausgeschriebenen Lösungen werden nur im Anhang in konventioneller Art geschrieben. In einigen Programmiersprachen, wie z. B. BASIC PLUS, können Matrixfunktionen direkt programmiert werden, was eine einfache Berechnung der Koeffizienten gestattet.

2.1 Ansatz

Voraussetzung, dass die Regressionsfunktion analytisch lösbar ist, ist die Kenntnis der Periodenlänge. Als Ausgangsgleichung wurde dieselbe Form gewählt, wie sie auch Halberg [5] und Kuhn [7] formuliert hatten. Die abhängige Variable y wird dargestellt als Funktion des Mittelwertes y_0 , der Amplitude A und der Phasenverschiebung t_0 sowie der unabhängigen Variablen t .

$$y(t) = y_0 + \frac{1}{2} A \sin \left[(t - t_0) \frac{2\pi}{T} \right] \quad (1)^1$$

In Abb. 1 wird die Bedeutung der Konstanten y_0 , A und t_0 veranschaulicht. Die Phase t_0 wurde so definiert, dass sie den Beginn der Sinuskurve innerhalb der ersten Periode, vom Nullpunkt der unabhängigen Variablen t an gerechnet, kennzeichnet.

Wegen des kombinierten Winkels $(t - t_0)$ ist die Funktion (1) nicht geeignet zur Lösung und muss deshalb unter Anwendung des Additionstheorems umgeformt werden.

$$y(t) = c_0 + c_1 \cos t' + c_2 \sin t' \quad (2)$$

Die Umrechnung der Amplitude A und der Phase t_0 in die nicht direkt anschaulichen Koeffizienten c_0 , c_1 und c_2 erfolgt mittels der Beziehungen (2a) bis (2c).

$$c_0 = y_0, \quad (2a)$$

$$c_1 = -\frac{1}{2} A \sin t'_0, \quad (2b)$$

$$c_2 = \frac{1}{2} A \cos t'_0. \quad (2c)$$

Tabelle 1. Symbolerklärungen. Die Schreibweise erfolgte in Anlehnung an Kendall und Stuart [6].
Table 1. Symbols. The symbols are used analogously to Kendall and Stuart [6].

A	= Amplitude der Sinusschwingung
c	= Vektor der gesuchten Koeffizienten (c_0 , c_1 und c_2)
i, j	= Indizes
k	= Konstante zur Festlegung der Phase t_0
n	= Anzahl Wertepaare
s	= Standardabweichung der Messwerte um die Sinuskurve
$s_{y(t)}$	= Standardabweichung der berechneten Kurvenpunkte
s_c	= Standardabweichung der Koeffizienten (c_0 , c_1 und c_2)
$s_{\hat{A}, \hat{t}_0}$	= Standardabweichung von Amplitude \hat{A} und Phase \hat{t}_0
T	= Periodenlänge
t	= Unabhängige Variable
t_0	= Phasenverschiebung der Regressionskurve gegenüber $t=0$
t', t'_0	= mit $\frac{2\pi}{T}$ transformierte unabhängige Variable t, t_0
X	= Matrix der bekannten Koeffizienten ($1, \cos t'_i, \sin t'_i$), $i=1$ bis n
x^0	= Vektor der bekannten Koeffizienten ($1, \cos t', \sin t'$)
y	= Vektor der gemessenen Werte ($y_1 \dots y_n$)
$y(t)$	= Abhängige Variable
y_0	= c_0 = Lage der Achse der Schwingung

Mit T bezeichnete Matrix bedeutet die transponierte Matrix.

Mit $^{-1}$ bezeichnete Matrix bedeutet die invertierte Matrix.

Mit $\hat{}$ bezeichnete Größen kennzeichnen Schätzwerte.

Die Summen verstehen sich als Summen über alle Messwerte von $i=1$ bis n .

1) Symbolerklärungen siehe Tabelle 1.

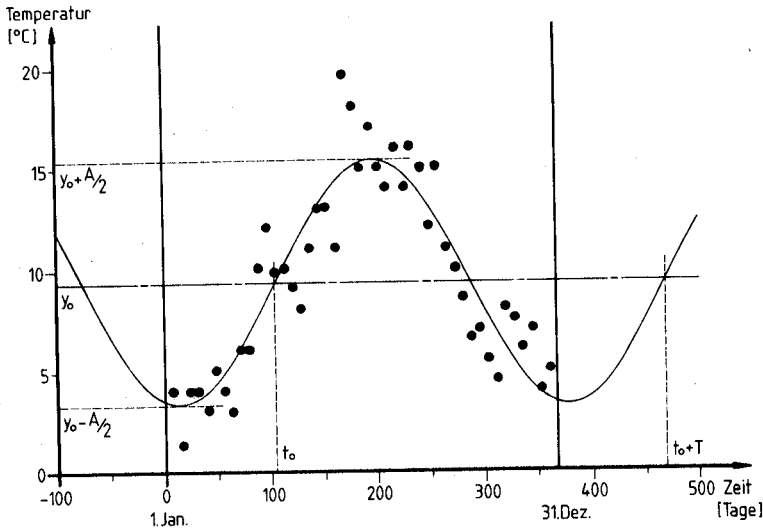


Abb. 1. Darstellung der Anpassung der Funktion $y(t) = y_0 + (A/2) \sin [(t - t_0) 2\pi/T]$ an die Wassertemperaturdaten der Niers, 1974 [2].

Figure 1. Fitting of the function $y(t) = y_0 + (A/2) \sin [(t - t_0) 2\pi/T]$ to the water temperature data of the Niers, 1974 [2].

Die Rückrechnung geschieht mit den Gleichungen (2d) und (2e), wobei eine neue Konstante k eingeführt werden muss, da die arctan-Funktion nur im 1. und im 4. Quadranten definiert ist. k nimmt die Werte 0 oder $1/2$ an, falls $(-c_1/c_2) \geq 0$ und die Werte $1/2$ oder 1 für $(-c_1/c_2) < 0$. Da t_0 den Beginn der Sinuskurve darstellen soll, muss die Ableitung von y bei t_0 positiv sein.

$$A = 2 \sqrt{c_1^2 + c_2^2}, \quad (2d)$$

$$t_0 = \frac{T}{2\pi} \arctan(-c_1/c_2) + kT. \quad (2e)$$

2.2 Berechnung der Koeffizienten

Die Lösung der Regressionsgleichung (2) erfolgt nach dem Verfahren der kleinsten Abweichungsquadrate. Dabei wird die Summe der Fehlerquadrate $\sum (y_i - \hat{y}(t))^2$ minimiert. Der Vektor der Schätzwerte für die unbekanntenen Koeffizienten \hat{c}_0, \hat{c}_1 und \hat{c}_2 lautet:

$$\hat{c} = (X^T X)^{-1} X^T y. \quad (3)$$

Die beiden Produkte mit der transponierten Form der Matrix der bekannten Koeffizienten lassen sich gut algebraisch berechnen, die Inversion erfolgt jedoch einfacher numerisch.

$$X^T X = \begin{bmatrix} n & \sum \cos t' & \sum \sin t' \\ \sum \cos t' & \sum \cos^2 t' & \sum \cos t' \sin t' \\ \sum \sin t' & \sum \cos t' \sin t' & \sum \sin^2 t' \end{bmatrix}, \quad (3a)$$

$$X^T y = \begin{bmatrix} \sum y_i \\ \sum y_i \cos t' \\ \sum y_i \sin t' \end{bmatrix}. \quad (3b)$$

Die Gleichung (3) kann auch als Gleichungssystem mit drei Unbekannten umgeschrieben werden und über die Determinanten oder das Gaußsche Eliminationsverfahren [4] gelöst werden (siehe Anhang).

2.3 Berechnung der Streuungen

Die Streuung der Messwerte um die berechnete Sinuskurve s beträgt:

$$s = \sqrt{\sum (y_i - \hat{y}(t))^2 \frac{1}{n-3}}. \quad (4)$$

Aus der Standardabweichung s lässt sich die Streuung der einzelnen Koeffizienten s_{e_0} , s_{e_1} und s_{e_2} sowie diejenige der Regressionskurve $s_{y(t)}$ berechnen.

$$s_{e_j} = s \sqrt{((X^T X)^{-1})_{jj}}, \quad (5)$$

$$s_{y(t)} = s \sqrt{(x^0)(X^T X)^{-1} x^0 + 1}. \quad (6)$$

Aus der Anwendung des Fehlerfortpflanzungsgesetzes auf die Formeln (2 d) und (2 e) ergeben sich die Standardabweichungen der Amplitude s_A und der Phase s_{t_0} .

$$s_A = 2 \sqrt{\frac{s_{e_1}^2 c_1^2 + s_{e_2}^2 c_2^2}{c_1^2 + c_2^2}}, \quad (7)$$

$$s_{t_0} = \frac{T}{2} \sqrt{\frac{s_{e_1}^2 c_2^2 + s_{e_2}^2 c_1^2}{c_1^2 + c_2^2}}. \quad (8)$$

2.4 Lösung für den Fall, dass einzelne Koeffizienten bekannt sind

Die Regressionskonstanten lassen sich sehr viel einfacher berechnen, wenn zum Beispiel die Phasenverschiebung t_0 bekannt ist. Dieser Fall ist realistisch, wenn t_0 durch eine andere Beziehung geschätzt werden kann (im Falle der Wassertemperatur aus früheren Messungen oder durch Vergleich mit der Lufttemperatur). Für die beiden übrigen Koeffizienten ergibt sich dann folgende Lösung:

$$\hat{y}_0 = \frac{\sum y_i (\cos t'_0 \sum \sin t' \cos t' - \sin t'_0 \sum \cos^2 t') - \sum y_i \cos t'}{n (\cos t'_0 \sum \sin t' \cos t' - \sin t'_0 \sum \cos^2 t') - \sum \cos t'}$$

$$\left(\frac{\cos t'_0 \sum \sin t' \cos t' - \sin t'_0 \sum \cos t'}{\cos t'_0 \sum \sin t' \cos t' - \sin t'_0 \sum \cos t'} \right), \quad (9)$$

$$\hat{A} = \frac{2 (\sum y_i - n y_0)}{(\cos t'_0 \sum \sin t' - \sin t'_0 \sum \cos t')}. \quad (9a)$$

Falls nur Einzeldaten vorliegen, lässt sich der Mittelwert über eine volle Periode durch folgende Beziehung grob abschätzen (Amplitude und Phase müssen bekannt sein):

$$\hat{y}_0 = \frac{\sum y_i - \frac{1}{2} A (\cos t'_0 \sum \sin t' - \sin t'_0 \sum \cos t')}{n}. \quad (9b)$$

3. Anwendungsbeispiel

Die Möglichkeiten und Grenzen der Regression mit einem Fourier-Term sollen am Beispiel des Jahresverlaufs der Wassertemperatur eines Fließgewässers illustriert werden. Die dazu verwendeten Messwerte entstammen einem nordrhein-westfälischen Bach, der Niers [2]. Die Auswahl der Daten erfolgte mit der Einschränkung, dass sie allgemein zugänglich sein sollen, rein zufällig. Die Messungen wurden während mehrerer Jahre in einem 8tägigen Rhythmus erhoben und eignen sich deshalb auch zum Vergleich verschiedener Jahre miteinander.

Der Motivation zu dieser Arbeit entsprechend, wird zuerst die Eignung der Koeffizienten y_0 , A und t_0 als Parameter zur Beschreibung von periodischen Phänomenen diskutiert. Abb.2 zeigt den Verlauf der Wassertemperatur der Niers in den Jahren 1970–1974. Um den Vergleich der einzelnen Jahre zu erleichtern, wurde die Regression als Funktion der Messwerte aller fünf Jahre dargestellt. In Tabelle 2 sind der arithmetische Mittelwert und die Koeffizienten für alle Jahre einzeln aufgeführt. Da die Messungen regelmässig während der ganzen Perioden durchgeführt wurden, stimmen die arithmetischen Mittel sehr gut mit den Achsen y_0 der Sinusschwingungen überein. Die Streuung der Punkte jedoch wird um einen Faktor 2 bis 3 verringert, das heisst, die Saisonabhängigkeit der Temperatur erklärt einen grossen Teil der Varianz (vgl. Bliss [1] zur Teststatistik von periodischen Regressionen). Die Amplituden der einzelnen Jahre charakterisieren die sommerlichen Maxima und die winterlichen Minima als mittlere Extremwerte, die sich sonst nur sehr schwer objektiv schätzen lassen. In Tabelle 2 lässt sich sehen, dass die Amplituden der Wassertemperatur etwa 2,5- bis 2,7mal grösser sind als die Standardabweichungen um den arithmetischen Mittelwert. Die Phasenverschiebung in Tagen ab dem 1. Januar der entsprechenden Jahre zeigt an, ob die Aufwärmung relativ früh oder spät in einer Periode erfolgte. Der Vergleich der berechneten Werte mit den Messpunkteserien der Abb.2 macht deutlich, dass 1971

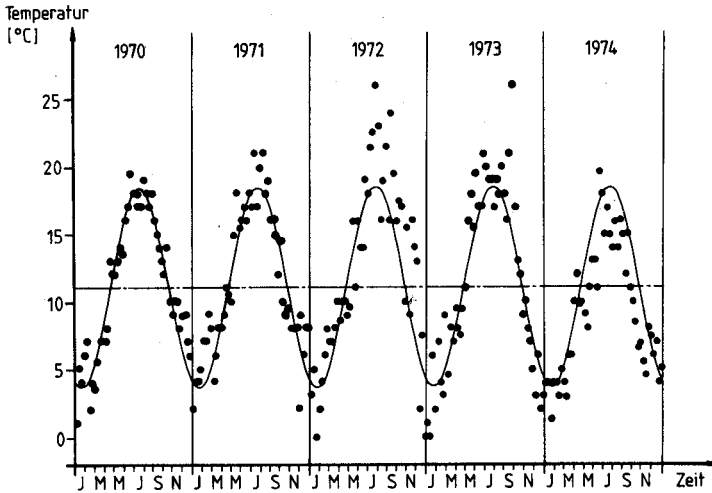


Abb. 2. Verlauf der Wassertemperatur der Niers [2] über mehrere Jahre mit eingezeichneter Regressionskurve. Die Berechnung der Koeffizienten erfolgte mit sämtlichen 228 Werten.

Figure 2. Course of water temperature of the Niers [2] during several years with regression curve. The computation of the coefficients is based on all 228 values.

der Sommer relativ früh einsetzte, während 1972 die Erwärmung erst später erfolgte. Im Mittel über die fünf Jahre erreichte die Temperatur der Niers ihren Mittelwert am 21. April. Die Phasenverschiebung gegenüber der Lufttemperatur beträgt etwa 10 bis 20 Tage.

Allgemein zeigen die Gegenüberstellung der berechneten Koeffizienten der Fourier-Transformation (Tab. 2) und die Darstellung der Messpunkte (Abb. 2), dass sich mit der Lage der Achse y_0 , der Amplitude A und der Phase t_0 periodische Phänomene sehr gut charakterisieren lassen.

Tabelle 2. Arithmetischer Mittelwert und Lage der Achse y_0 mit Standardabweichungen der Punkte um den Mittelwert und um die Sinuskurve sowie Amplitude und Phase der Wassertemperaturdaten der Niers, 1970–1974. Die Daten stammen aus den Jahrbüchern der Bundesrepublik Deutschland [2].
Table 2. Arithmetic mean and axes y_0 with standard deviations of the points around the mean and the sine curve, as well as amplitude and phase of the water temperature data of the Niers, 1970–1974. The data are from the yearbooks of the Federal Republic of Germany [2].

Jahr	Anzahl Werte	Arithmetischer Mittelwert und Standardabweichung der Punkte (°C)	Lage der Achse y_0 und Standardabweichung der Punkte (°C)	Amplitude A (°C)	Phase t_0 (d)
1970	46	10,93 5,24	10,99 1,50	14,03	114
1971	45	11,51 5,21	11,42 1,96	13,62	104
1972	46	12,23 6,85	12,29 3,10	17,16	121
1973	46	11,57 6,67	11,63 2,54	17,27	108
1974	45	9,43 4,70	9,35 1,92	12,14	104
Alle	228	11,14 5,83	11,14 2,60	14,75	111

Tabelle 3. Arithmetischer Mittelwert und Koeffizienten der periodischen Regression sowie deren Standardfehler für ausgewählte Stichproben der Wassertemperatur der Niers, 1974 [2].

Table 3. Arithmetic mean and coefficients of the periodic regression with their standard errors of selected samples out of the water temperature data of the Niers, 1974 [2].

Stichprobe (Werte)	Anzahl Werte	Arithmetischer Mittelwert (°C)	Lage der Achse y_c (°C)	Amplitude A (°C)	Phase t_0 (d)
a) 9-19	11	$9,81 \pm 0,74$	$8,43 \pm 22,74$	$8,43 \pm 8,07$	$89,6 \pm 336,6$
b) 20-30	11	$15,51 \pm 0,68$	$10,62 \pm 28,73$	$10,83 \pm 62,97$	$101,5 \pm 73,0$
c) 1-22	22	$8,40 \pm 1,04$	$10,05 \pm 0,62$	$13,27 \pm 0,97$	$113,4 \pm 9,8$
d) 15-37	23	$12,81 \pm 0,74$	$6,85 \pm 0,40$	$19,02 \pm 1,36$	$108,4 \pm 2,0$
e) 1-45	45	$9,43 \pm 0,70$	$9,35 \pm 0,04$	$12,14 \pm 0,17$	$103,7 \pm 0,8$
f) 1, 10, 19, 28, 37	5	$8,80 \pm 1,98$	$8,74 \pm 0,19$	$11,12 \pm 0,75$	$102,0 \pm 3,9$

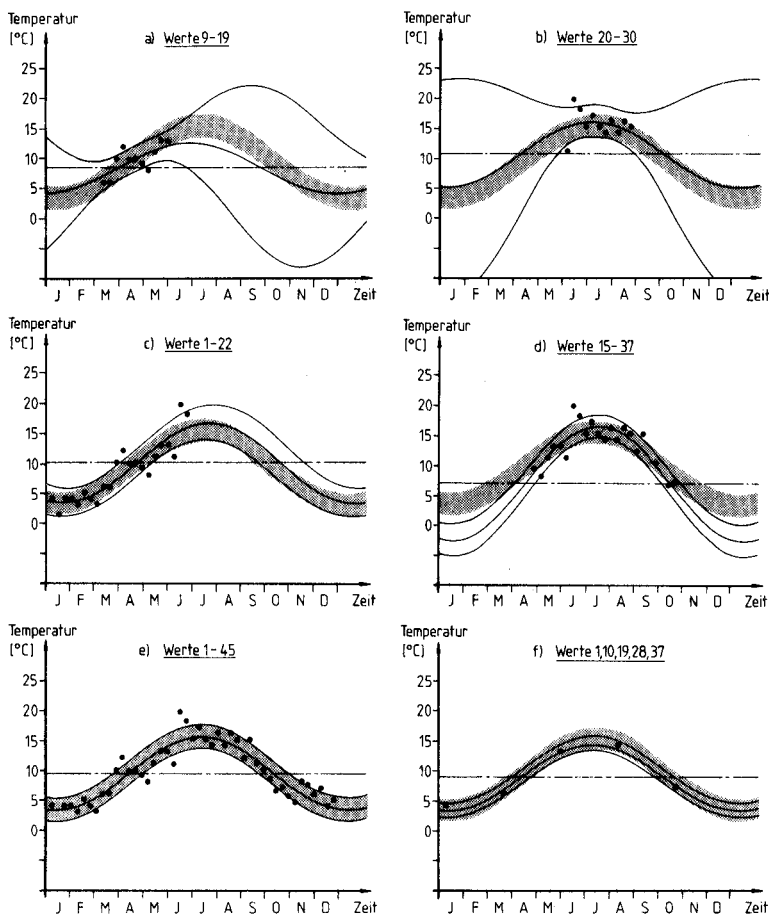


Abb. 3a-f. Regressionskurve und deren Streuung für ausgewählte Messbereiche innerhalb des Jahres 1974. Schattiert dargestellt ist der Streuungsbereich der Kurve bei Verwendung aller 45 Messwerte.

Figures 3a-f. Regression curve and its standard deviation for selected samples out of the whole period of 1974. The shaded area shows the range of the deviation when using all 45 values.

Um das Verhalten der einzelnen Konstanten bei kleinen Stichproben zu untersuchen, wurden verschiedene Auszüge aus den total 45 Werten des Jahres 1974 gemacht und die Regressionsfunktion berechnet. Die Resultate sind in Tabelle 3 aufgeführt und in Abb.3 dargestellt. Falls die Stichprobe nur innerhalb eines Vierteljahres erhoben wurde (Abb.3a und 3b), sind Aussagen über die ganze Periode sehr fragwürdig. Das Bild ändert sich, wenn mindestens während eines halben Jahres gemessen wurde (Abb.3c und 3d). Die Fourier-Transformation liefert in diesen Fällen mehr oder weniger vernünftige Werte, während das arithmetische Mittel im einen Fall viel zu hoch liegt (Sommerwerte) und im andern Fall zu tief (Winter und Frühjahr). Gute Resultate werden erzielt, wenn regelmässig über die ganze Periode verteilt Messungen vorliegen (Abb.3e und 3f). Im Falle der Wassertemperatur zeigt sich, dass die Genauigkeit der Probenzahlen über fünf nicht wesentlich gesteigert werden kann. Es bleibt hier allerdings zu berücksichtigen, dass die Streuungen um die Sinuskurve einesteils witterungsbedingt sind und andernteils von den Tag-Nacht-Schwankungen herrühren. Da die Tagesverläufe nicht zufällig sind, können durch Messungen zu immer der gleichen Zeit systematische Fehler gemacht werden, die eine zu grosse Genauigkeit vortäuschen.

4. Diskussion

Die Berechnung einer Regression mit dem Grundschwingungsterm einer Fourier-Analyse setzt voraus, dass die Messpunkte zufällig um eine Sinuskurve streuen. Dies dürfte in den meisten Fällen nicht exakt zutreffen, erwies sich aber für die untersuchten Parameter als genügend genau. Dies gilt insbesondere dann, wenn nicht Einzelwerte geschätzt werden wollen, sondern lediglich eine Charakterisierung von Jahresverläufen mittels der drei Koeffizienten Ziel der Berechnungen ist.

Unterliegen Messwerte periodischen Schwankungen und sind sie nicht gleichmässig über eine ganze Periodenlänge verteilt, eignet sich das arithmetische Mittel nicht zur Beschreibung eines durchschnittlichen Zustandes. Die Transformation der Werte mit einer Sinusfunktion ermöglicht die Berechnung eines nicht nur zutreffenderen, sondern auch wesentlich streuungsärmeren Mittelwertes. Dabei ergeben sich zusätzliche Informationen durch die Kenntnis von Amplitude und Phasenverschiebung. Als besonders nützlich erweisen sich diese Grössen, wenn verschiedene Schwingungen miteinander verglichen werden sollen oder biologische Ereignisse mit Umweltveränderungen korreliert werden.

Die Streuungsberechnungen zeigen, dass sich der Arbeitsaufwand in Untersuchungsprogrammen beträchtlich vermindern lässt, wenn die Probenahmen der Zielsetzung entsprechend angelegt werden. Sollen Aussagen über eine ganze Periode gemacht werden, so empfiehlt es sich in jedem Fall, die Proben regelmässig während einer kompletten Periode zu erheben. Sind den Grundschwankungen weitere systematische Variationen überlagert, müssen auch diese gleichmässig berücksichtigt werden. Für das Beispiel der Temperatur bedeutet das, dass auch die Tageszeiten regelmässig vertreten sein müssen.

Sinusförmig variierende Grössen zeichnen sich durch eine zweigipflige Häufigkeitsverteilung der Messwerte aus. Diese Verteilung kann den von Zobrist et al. [9]

beschriebenen Typen (normal und lognormal) überlagert sein, und es ist nicht immer möglich, die saisonale Abhängigkeit eines Phänomens so klar zu zeigen wie bei der Temperatur. Beim Sauerstoffgehalt von Fließgewässern wird der Zusammenhang mit der Wassertemperatur verschleiert, wenn die Werte unter den Sättigungskonzentrationen liegen. Wird ein Gewässer jedoch nicht antropogen beeinträchtigt, korrelieren die Sauerstoffkonzentrationen mit einem Korrelationsfaktor von etwa 0,8 mit der sinusförmigen Regressionskurve. Bei Gewässern werden Periodizitäten durch die Sonnenstrahlung induziert, welche einerseits eine Erwärmung und andererseits eine Erhöhung der Primärproduktion verursacht. Diese Effekte wiederum äussern sich durch mehr oder weniger starke Jahresperiodizitäten vieler chemischer Wasserinhaltsstoffe [3]. So zeigen Nährstoffe in Seeausflüssen regelmässige Jahreskurven, die zum Teil gut mit den entsprechenden Sinusschwingungen korrelieren.

Die vermehrte Anwendung von kontinuierlich registrierenden Messgeräten und der verbesserte Zugang zu Rechenanlagen werden die Feldökologen immer häufiger mit dem Problem der Analyse von Zeitreihen konfrontieren. Dabei sind die Zielsetzungen zum Teil das Suchen von unbekanntem Gesetzmässigkeiten und zum Teil die Korrektur systematischer Schwankungen durch geeignete Transformationen. Solche Transformationen sind dann am besten, wenn sie nicht nur statistisch sinnvoll, sondern auch kausal zu begründen sind, wie im vorliegenden Fall der Temperatur im Jahresverlauf.

5. Zusammenfassung

Durch Anwendung eines Additionstheorems kann die Funktion

$$y(t) = y_0 + \frac{1}{2} A \sin \left[(t - t_0) \frac{2\pi}{T} \right]$$

in eine als gewöhnliche lineare Regression berechenbare Form gebracht werden. Die Lösung dieser Regressionsgleichung liefert nicht nur eine Abhängigkeitsfunktion, sondern ausserdem drei nützliche Parameter zur Charakterisierung von periodisch variierenden Grössen. Die Amplitude A ist ein Mass für die mittlere Abweichung vom Mittelwert y_0 , und die Phasenverschiebung t_0 gibt an, wo sich die Schwingung in einem Bezugssystem (z. B. dem Kalenderjahr) befindet. Die Regression kann als Transformationsfunktion zum Ausgleich von Saisonschwankungen benutzt werden. Da der Mittelwert nicht sehr empfindlich auf kleine Abweichungen von Phase und Amplitude reagiert, wurden auch Lösungen berechnet für geschätztes t_0 bzw. t_0 und A .

Die Berechnung der Standardabweichung der theoretischen Schwingung und die Fehlerabschätzung der Koeffizienten liefern wertvolle Hinweise für die Planung von Untersuchungen. Es erweist sich in den meisten Fällen als sinnvoll, regelmässig über eine ganze Periodendauer Proben zu nehmen, aber die Genauigkeit der Aussagen kann durch Erhöhung der Messfrequenz über einem Minimum von etwa fünf bis zehn Proben nur noch unwesentlich gesteigert werden.

Die theoretischen Betrachtungen werden am Beispiel der Wassertemperatur der Niers, eines westdeutschen Bachs [2], illustriert. Der Bach erreicht seine Mitteltemperatur von 11,1 °C im Durchschnitt über die fünf berücksichtigten Jahre am 21. April. Seine Amplitude schwankt zwischen 12,1 °C und 17,3 °C und beträgt im Mittel 14,8 °C. Verschiedene andere Parameter, wie Sauerstoffgehalt und einige Nährstoffe, können unter gewissen Bedingungen gute Korrelationen mit einer Sinusschwingung aufweisen.

SUMMARY

Using an addition theorem of trigonometry the function

$$y(t) = y_0 + \frac{1}{2} A \sin \left[(t - t_0) \frac{2\pi}{T} \right]$$

can be brought to a form which can be treated as a common linear regression function. The solution of this equation not only quantifies a relationship but also gives three useful parameters for the characterization of periodically varying phenomena. The amplitude A is a measure for the deviation from the mean y_0 and the phase t_0 indicates the position of the wave on a given scale (e.g. calendar). The regression can be used as a transformation function for the compensation of seasonal oscillations. As y_0 is not very sensitive to errors of A and t_0 , the calculation can well be done with estimated amplitude and (or) phase.

The computation of the standard deviation of the theoretical curve and an error estimation of the coefficients provide valuable information when planning sampling programs. For most cases it is sensible to take samples regularly during the whole period, the accuracy however will not improve significantly when taking many more than five to ten samples.

The theoretical considerations are illustrated using temperature data of a stream in Western Germany, the Niers, published in [2]. The mean water temperature of the stream over the five-year period is 11.1 °C and this temperature is reached on 21 April. Its amplitude averaged 14.8 °C and varies between 12.1 °C and 17.3 °C. Some other stream parameters, such as oxygen content and nutrients, can show good correlation with a sinusoidal wave when conditions are suited.

RÉSUMÉ

En utilisant les règles de la trigonométrie, la fonction

$$y(t) = y_0 + \frac{1}{2} A \sin \left[(t - t_0) \frac{2\pi}{T} \right]$$

peut être transformée en une formule à appliquer comme une régression linéaire. Le résultat de cette équation fournit non seulement une fonction de dépendance mais encore trois paramètres pouvant servir à caractériser des valeurs à variations périodiques. L'amplitude A signale la déviation de la moyenne y_0 , et la phase t_0 indique le rapport entre l'oscillation considérée et d'autres oscillations ou un système de référence fixe (tel que l'année). La régression peut être utilisée comme fonction de transformation pour compenser des fluctuations saisonnières. Etant donné que la moyenne y_0 n'est pas sensiblement affectée par de petites déviations de l'amplitude et de la phase, il est possible de calculer y_0 à l'aide de valeurs t_0 resp. t_0 et A estimées.

La planification de programmes de recherche est facilitée grâce au calcul de la déviation standard de la courbe théorique et par l'évaluation d'erreurs de coefficients. Dans la majorité des cas, il est raisonnable de prélever des échantillons pendant toute la période d'investigation, mais il n'est pas

possible d'améliorer substantiellement la précision des résultats en augmentant la fréquence des prélèvements au-delà de cinq à dix échantillons.

Les considérations théoriques sont concrétisées par l'exemple de la température de la Niers [2]. Durant la période examinée de cinq ans, cette rivière eut une température moyenne de 11,1°C qu'elle atteignit le 21 avril. L'amplitude moyenne de 14,8°C varia entre 12,1°C et 17,3°C. Plusieurs autres paramètres tels que la teneur en oxygène et quelques sels nutritifs peuvent, suivant les conditions, présenter de bonnes corrélations avec une onde sinusoïdale.

VERDANKUNGEN

Mein spezieller Dank gebührt Herrn Dr. H. Bühner für dessen entscheidenden Hinweis auf den analytischen Lösungsweg des Problems und seine Unterstützung bei der Abfassung der Arbeit. Ebenfalls danken möchte ich den Herren Dr. D. Imboden für seine Hilfe bei den Streuungsberechnungen und Dr. J. Blösch, Hj. Fricker und Dr. Perret für die kritische Durchsicht des Manuskriptes. Dem Leiter der Limnologischen Abteilung der EAWAG, Herrn Prof. Dr. H. Ambühl, bin ich für den Arbeitsplatz und die Schlussdurchsicht des Artikels sehr verbunden. Frau Dr. J. Davis und Frau S. Rex danke ich für die Korrekturen der fremdsprachigen Zusammenfassungen.

LITERATURVERZEICHNIS

- 1 Bliss, C.I.: Statistics in Biology, Bd.2, 639 S., McGraw-Hill, New York 1970.
- 2 Bundesanstalt für Gewässerkunde und Deutsche Forschungsgemeinschaft: Internationale Hydrologische Dekade. Jahrbuch der Bundesrepublik Deutschland 1970-1974. Bundesanstalt für Gewässerkunde, Koblenz 1975-1977.
- 3 Eidgenössisches Amt für Wasserwirtschaft: Hydrographisches Jahrbuch der Schweiz 1975. Eidg. Drucksachen- und Materialzentrale, Bern 1976.
- 4 Hainzl, J.: Mathematik für Naturwissenschaftler, 311 S., B.G. Teubner, Stuttgart 1974.
- 5 Halberg, F., Tong, Y.L., und Johnson, E.A.: Circadian System Phase - An Aspect of Temporal Morphology, Procedures and Illustrative Examples. In: von Meyersbach, H. (Hrsg.): The Cellular Aspects of Biorythms, 198S. Springer-Verlag, Berlin 1967.
- 6 Kendall, M., und Stuart, A.: The Advanced Theory of Statistics. Vol.2: Inference and Relationship, 4. Aufl., 748 S., Charles Griffin & Co. Ltd., London and High Wycombe 1979.
- 7 Kuhn, W.: Berechnung der Temperatur und Verdunstung alpiner Seen auf klimatologisch-thermodynamischer Grundlage. Arbeitsberichte der Schweizerischen Meteorologischen Zentralanstalt, Zürich 1977.
- 8 Nicolas, M.: Saisonbereinigung von Wirtschaftsreihen durch harmonische Analyse. Vjhft. Wirtschaftsforsch. 1956, 334-348 (1956).
- 9 Zobrist, J., Davis, J., und Hegi, H.R.: Charakterisierung des chemischen Zustandes von Fließgewässern. Gas. Wass. Abwass. 57 (5), 402-415 (1977).

Anhang

Gleichungssystem mit den drei Unbekannten \hat{c}_0 , \hat{c}_1 und \hat{c}_2 :

$$\begin{aligned} \hat{c}_0 n + \hat{c}_1 \sum \cos t' + \hat{c}_2 \sum \sin t' &= \sum y_i, \\ \hat{c}_0 \sum \cos t' + \hat{c}_1 \sum \cos^2 t' + \hat{c}_2 \sum \cos t' \sin t' &= \sum y_i \cos t', \\ \hat{c}_0 \sum \sin t' + \hat{c}_1 \sum \cos t' \sin t' + \hat{c}_2 \sum \sin^2 t' &= \sum y_i \sin t'. \end{aligned}$$

Lösungen für \hat{c}_0 , \hat{c}_1 und \hat{c}_2 :

$$\hat{c}_0 = \frac{\sum y_i \sum \cos^2 t' \sum \sin^2 t' + \sum \cos t' \sin t' (\sum \cos t' \sum y_i \sin t' + \sum \sin t' \sum y_i \cos t' - n \sum \cos^2 t' \sum \sin^2 t' + \sum \cos t' \sin t' (2 \sum \cos t' \sum \sin t' - n \sum \cos t' \sin t')) - \sum y_i \sum \cos t' \sin t' - \sum \cos^2 t' \sum \sin t' \sum y_i \sin t' - \sum y_i \cos t' \sum \sin^2 t' \sum \cos t'}{\sum \cos^2 t' (\sum \sin t')^2 - \sum \sin^2 t' (\sum \cos t')^2},$$

$$\hat{c}_1 = \frac{(\sum y_i \sum \cos t' \sin t' - \sum \sin t' \sum y_i \cos t') - \hat{c}_0 (n \sum \cos t' \sin t' - \sum \cos t' \sum \sin t')}{\sum \cos t' \sum \cos t' \sin t' - \sum \sin t' \sum \cos^2 t'}$$

$$\hat{c}_2 = \frac{\sum y_i - c_0 n - c_1 \sum \cos t'}{\sum \sin t'}$$

Adresse des Autors: Herbert Güttinger, dipl.sc.nat., EAWAG, CH-8600 Dübendorf.