Eur Arch Psychiatr Neurol Sci (1989) 239:133–139

European Archives of Psychiatry and Neurological Sciences

Single-Case Evaluation of Sleep-Deprivation Effects by Means of Nonparametric Time-Series Analysis (According to the HTAKA Model)

H.J. Möller[1], R. Blank[2], and E.M. Steinmeyer[3]

[1]Universitäts-Nervenklinik und Poliklinik, Abteilung Psychiatrie, Bonn, Federal Republic of Germany
[2]Psychiatrische Klinik der Technischen Universität München, München, Federal Republic of Germany
[3]Psychiatrische Klinik der Technischen Hochschule Aachen, Aachen, Federal Republic of Germany

Summary. Modern statistical approaches to the analysis of single cases have so far been rarely employed in psychiatric therapy research. The recently developed Hierarchical Trend-Segment Component Analysis (HTAKA) seems to be an adequate and practicable method for the field of therapy evaluation, not only under sophisticated research conditions, but also under routine treatment conditions. Using self-rating data of the effects of sleep-deprivation therapy as an example, the results of nonparametric time-series analyses performed according to HTAKA are presented. The statistical agglomeration of the single case results demonstrated that in most cases undergoing repeated sleep deprivation a long-term antidepressive effect related to this therapy could be proven.

Key words: Single case analysis – Nonparametric time-series analysis – HTAKA model – Sleep-deprivation therapy

Introduction

For many years, sleep deprivation has been recommended as an antidepressant procedure for depressive patients, especially for the endogenous subgroups (Pflug and Tölle 1971). A remarkable improvement of mood can occur directly following sleep deprivation. Repeated sleep deprivations night after night induces in responders long-term positive changes of the depressed mood. These effects occur both with and without concomitant therapy with antidepressant medication (cf. review by Kuhs and Tölle 1986). Most previous evaluation of this therapy, however, has been based on simple group-statistical "before-and-after" comparisons; rarely has an attempt been made to compare the group statistics of mood changes over time between patients who did and those who did not undergo sleep-deprivation therapy. As far as is known to the authors, the more modern possibilities offered by statistical analysis of time-series data for individual cases have not been employed for evaluating the effects of sleep deprivation.

In the field of therapy evaluation involving single cases, time-series analyses may aid in the statistical confirmation of observed effects and thus raise these observations beyond the realm of anecdotal evidence. In attempting the statistical analysis of such time-series data, however, the investigator is confronted with several familiar problems, such as the serial dependency of the data, both discrete or temporarily continuous measurements over time (the latter usually unrealisable), duration of recording the number of data collection times (cf. Appelt and Strauß 1985; Lienert 1978; Möbius et al. 1985) as well as other problems. Neither the older "trend analysis" (Bortz 1979; Edwards 1971; Kirk 1982; Pfanzagl 1964), the "single-subject analysis of variance" (Shine and Bower 1971), nor the currently popular ARIMA (autoregressive moving average) models (Box and Jenkins 1976; Cook and Campbell 1979; Beneke 1981; Glass et al. 1975; Keeser 1979; McCleary and Hay 1980; Revenstorf and Keeser 1979) have been able to solve these problems satisfactorily (cf. Strauß and Stemmler 1985).

As an alternative approach to this problem, the present study therefore employs the "Hierarchical Trend-Segment Component Analysis" (Hierarchische Trend-Abschnitts-Komponenten-Analyse, HTAKA, cf. Kleiter 1986) for the evaluation of the effects of sleep deprivation on the mood status of endogenous-

Offprint requests to: H.J. Möller, Universitäts-Nervenklinik und Poliklinik, Abteilung Psychiatrie, Sigmund-Freud-Strasse 25, D-5300 Bonn 1, Federal Republic of Germany

depressive patients. HTAKA is a new form of time-series analysis, combining the advantages of trend and ARIMA analysis while simultaneously solving some of the problems mentioned above. With regard to the practicability of this approach, it is important to emphasize that HTAKA, which utilizes nonparametric significance tests, requires considerably fewer continuous measurements over time than is the case for the ARIMA model.

Methodology

The aim of Hierarchical Trend-Segment Component Analysis (HTAKA) is to review critically the inferentially-based aspects in individual time-series with respect to their clinical relevance. Regarding their applicability for therapy, cumulative processes and alternating trend parameters will receive special attention, together with the validation procedures to specify differences in growth curve levels. The analysis model suggested above can be utilized for both experimental as well as for nonexperimental approaches (clinically the far more frequent case). In nonexperimental settings, no influencing phases (therapeutical or intervening phases) are assumed in advance. Considering such a situation, the question arises how to filter at the same time possible and optimal curve sections from the longitudinal data, and then to analyze these to alterations in level or trend. HTAKA offers two main advantages in comparison to ARIMA for the statistical illustration of individual therapeutic processes in the clinical field:

1. From the single-case statistical point of view, the observation of therapeutic effects deals mainly with cumulative processes, such as continuously increasing or decreasing learning or therapeutic processes. These are not considered in the ARIMA models, for instance, which focus more on repetitive data as they occur in psychophysiological research. Thus, we need a model for analyzing cumulative processes.

2. In ARIMA models, in order to identify specific models, unrealistically large amounts of data are usually required to enable an appropriate identification of substantative models. HTAKA also employs ARIMA procedures for the elimination of autocorrelations, yet in a degenerative form and on a much reduced level. The aim of ARIMA in the HTAKA procedure is merely to recognize and accordingly isolate serial dependencies. No forecasting takes place. In contrast, the nonparametric HTAKA models can be evaluated and tested for significance with considerably fewer data points.

The HTAKA model employs the following approach:

1. By using cluster-analytical subroutines (the KMEANS method), time segmentations can be ascertained alternatively according to the search criteria SAQ = sum of squared deviations (with changes of level dominating) and SRQ = sum of residual squares (with changes in trends dominating), while adhering to familiar time spans. These groupings and divisions, ordered according to temporal aspects, are assumed in advance. The segments determined in advance by the cluster-analysis demonstrate the à priori pertaining structures so clearly because, due to mathematical procedures, no substitution between elements of various time segments is possible. This incidentally also forces outliers to remain within a segment. Therefore, no optimal allocation takes place, but instead with the restricting factor time and thus frequently to the disadvantage of a differentiating hypothesis assuming variations between segments.

2. The linear determined trend is separated by autocorrelating processes from each segment.

3. Utilizing ARIMA, residuals are analyzed, including all continuous measurements over time, in order to find out whether they still contain systematic variation or not. The intention of an ARIMA analysis of residuals in such a case is to detect and subsequently remove the eventual parts of serial dependencies from the presented registration intervals. The application of the ARIMA analysis is performed on a relatively low level, without forecasting. It is not intended to identify a specific (parametric) model; therefore, the known argument concerning the number of measurements over time does not apply here.

4. The remaining residuals are tested for normal distribution.

5. The remaining residuals are included to the previously separated trend parts. Subsequently, component strengths are calculated, and nonparametric tests (Kruskal-Wallis H tests) are carried out for relevant curve segments in order to determine whether the changes in level have led to significantly positive or negative growth (ADD Cumulative), whether the changes in level are different from segment to segment (ADD Differences), whether a significant cumulative trend enhancement is present (STG Cumulative), or whether there are trend differences to be observed within the various segments (STG Differences).

In the schedule of an experimental setting, influencing phases are determined in advance, mostly according to the ABAB scheme and are specified in the HTAKA model in the so-called allocation vector. Therefore, it is not necessary for the segment-specific

allocations to be determined first by means of the cluster-analytical subroutines, as in the nonexperimental case. In the experimental model, the allocation vector can be specified using a "dummy" variable in addition to the dependent variable. In such a case, the cluster-analytical segmentation of a time-series can be skipped and the data analyzed directly with respect to the experimental intervention phases.

In the context of a routine documentation system (Möller et al. 1983), patients admitted to the clinical division of the Max-Planck Institute of Psychiatry in Munich (MPIP) completed every other day either the Adjective Mood Scale (AMS) or its parallel form (AMS), alternatively. These self-evaluation questionnaires, which are completed throughout the inpatient stay, allow the recording of rapid changes in mood status as a correlate of increasing or decreasing depressivity. High values denote a disturbed mood (von Zerssen 1976, 1986). Overall, there were 24 patients with an ICD diagnosis of endogenous depression, who within the last 10 years had had additional treatment with antidepressants, and who on one or more occasions had undergone sleep-deprivation therapy. From this sample, 13 patients with complete mood assessment data appeared most suitable for demonstrating the HTAKA procedures. In 6 cases, only one single sleep-deprivation treatment was performed during the time period in which the data were compiled, while the other 7 cases underwent several treatments. However, it must be emphasised that these 13 cases present only a highly selected sample in which sleep-deprivation therapy was performed either because of the severity of symptoms, or because of earlier inadequate responses to other antidepressant therapies. To demonstrate the HTAKA model, in the following text we will describe one experimental and one nonexperimental case. In several steps the individually determined trend parameters will be analyzed in a aggregate form.

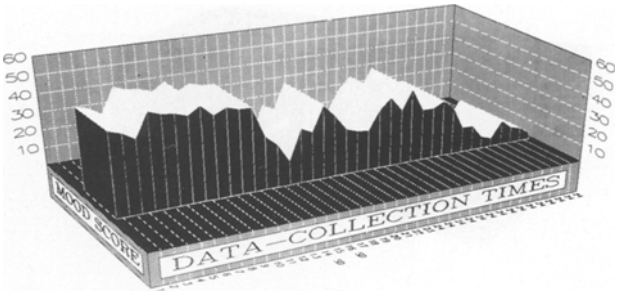

Fig. 1. Trend of mood assessment scores across all 38 continuous measurements over time, with sleep deprivation at data collection points 14 and 16 (nonexperimental case)

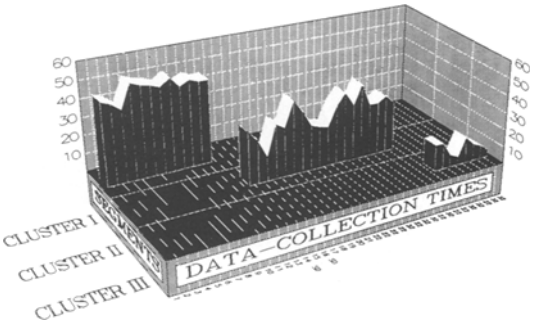

Fig. 2. Cluster-analytical sub-grouping of mood assessment data (nonexperimental case)

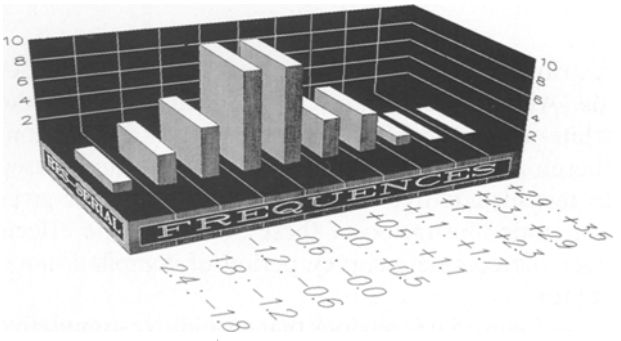

Fig. 3. Control of residuals within normal distribution (nonexperimental case)

Results

HTAKA in a Nonexperimental Case

A 32-year-old male patient suffering from an endogenous depression was evaluated at 38 data collection times during his inpatient stay of almost 100 days. At data collection points 14 and 16, complete sleep-deprivation therapy was performed. Figure 1 illustrates the patient's mood assessment curve at the 38 data collection times. Using the cluster-analytical segmentation of the time-series option of the HTAKA model, 3 time-series clusters were formed as demonstrated in Fig. 2. In this figure, time-series cluster 1 presents the mood scores of the first 12 assessment points documenting the mood before sleep-deprivation therapy. Cluster 2, beginning one data collection point before the sleep-deprivation phase, totals 18 data collection points. According to an impressionistic analysis, the curve segments clearly sink to a lower (better) mood level in this segment. Cluster 3 documents the final 8 data collection points, in which another reduction of the mood level can be observed. To substantiate this impressionistic analysis, the HTAKA model now evaluates the serial dependencies as described above. Figures 3 and 4 demonstrate the

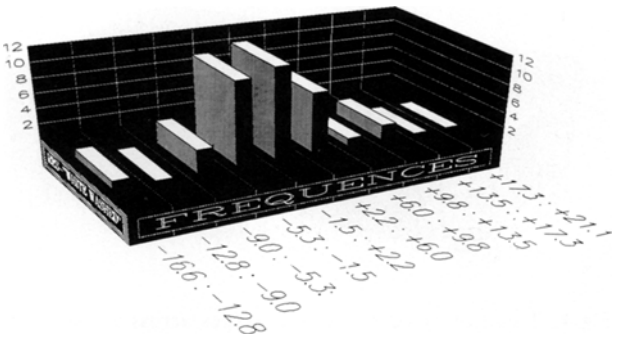

Fig. 4. Control of "white-washed" residuals within normal distribution (nonexperimental case)

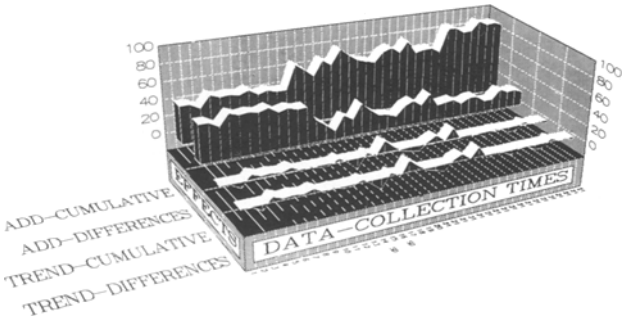

Fig. 5. Illustration of additive and slope effects during total therapy curve (nonexperimental case)

distribution of both serial and "white-washed" residuals. As the figures show, both serial as well as the white-washed residuals follow a normal distribution; therefore, nonparametric tests (H tests) can be used to test for additive and slope effects. Figure 5 gives a graphic illustration of these additive slope effects over the complete survey period of compiled mood values.

In Figure 5 it is obvious that an additive-cumulative effect can be obseved across all three cluster-analytically grouped phases, whereas the additive differences can basically be traced back to the level-differences between clusters 1 and 2. With respect to the slope effects, neither cumulative nor differential aspects demonstrated any obvious variance during the entire data collection period.

Table 1 shows the statistical analysis of the mood values according to the HTAKA model in a nonexperimental case. Cluster by cluster, the mood level of the initial values has consistently decreased; cluster 3 especially presents a significant change of trend in the mood-development curves. The statistical analysis of the "white-washed" data also shows that a highly significant additive-cumulative effect was present. In other words, the mood level of the data has continuously decreased cluster by cluster, i.e., the patient's mood has consistently improved. The test confirming additive differences also leads to highly significant results. The inspection of Figure 5 clearly shows, in contrast to the row data in Fig. 1, that the white-washing of data allows this significance to be attributed mainly to the differences in level between clusters 1 and 2. On the other hand, tests aiming to evaluate trend effects fail to yield significant results (cf. Table 1).

Table 1. HTAKA analysis of the mood scores (nonexperimental case)

	Cluster I	Cluster II	Cluster III
Data collection times (n):	12	18	8
Segment level (M):	47.75	30.39	9.88
Segment trend (b):	0.23	0.42	−1.11
Significance tests:			
a) Add cumulative:	H = 31.579 DF = 2 $P < 0.001$		
b) Add differences:	H = 24.988 DF = 2 $P \leq 0.001$		
c) Trend cumulative:	H = 0.432 DF = 2 $P = 0.808$		
d) Trend differences:	H = 0.597 DF = 2 $P = 0.746$		

The HTAKA analysis of a single case development in a nonexperimental case displays a segmentation of the mood scores split into three divisions. The second segmentation presents the row data from the phase following sleep deprivation. After 12 additional data collection times, a further improvement of mood is recorded without any further sleep deprivation. After elimination of the serial dependencies the statistical analyses show a confirmation of the continuous improvement of mood from segment to segment.

With the HTAKA procedure in the nonexperimental case, an intensive idiographic description of the therapeutic development is obtained, allowing valid evaluation of the therapeutic processes.

HTAKA in an Experimental Case

A 59-year-old female patient suffering from an endogenous depression was subjected to a total of 5 sleep-deprivation therapies during an inpatient stay of altogether 65 days. Mood scores were assessed at 28 data collection times with complete sleep-deprivation therapy at the 6th, 8th, 13th, and 18th time point. Figure 6 shows the patient's mood curve during the 28 data collection points.

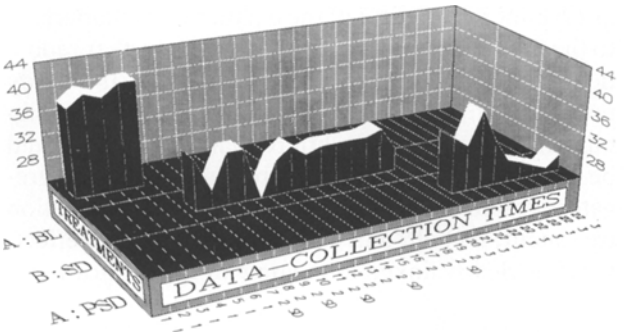

Fig. 6. Trend of mood assessment scores under experimental conditions. Baseline phase (A:BL); Sleep-deprivation phase (B:SD); Post-sleep-deprivation phase (A:PSD), with X as "dummy" variable (experimental case)

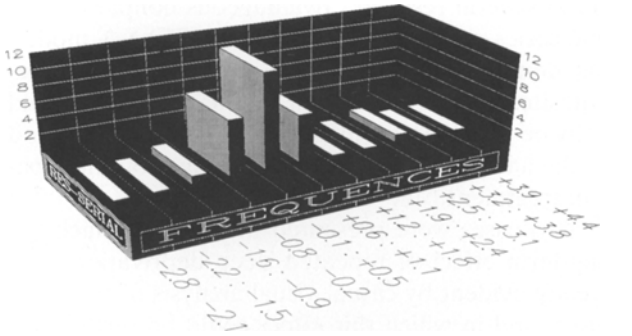

Fig. 7. Control of normal distribution of residuals (experimental case)

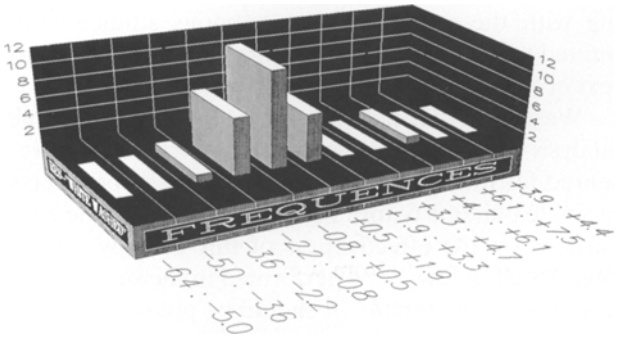

Fig. 8. Control of normal distribution of "white-washed" residuals (experimental case)

In accordance to the ABA design, data collection points 1 to 5 compiled prior to the initial sleep deprivation therapy are considered to constitute the baseline phase (A:BL). Between data collection points 6 to 9, phase B presents a therapeutic intervention phase (B:SD). Finally, the data collection points 20 to 28 constitute those points following the therapeutic intervention phase and are defined in advance as the post-sleep deprivation phase (A:PSD) accordingly.

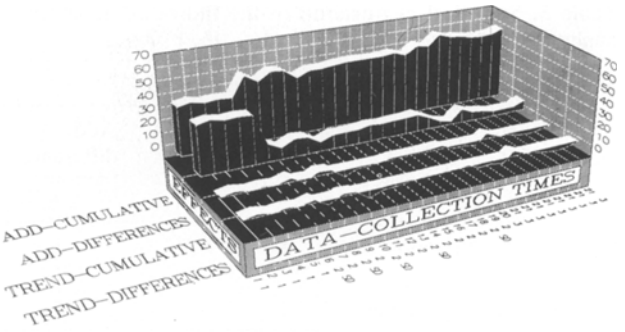

Fig. 9. Illustration of additive and slope effects across total time span (experimental case)

Table 2. HTAKA analysis of the mood scores (experimental case)

Treatments:	A Baseline	B Sleep deprivation	A Post-sleep-deprivation
Data collection times (n):	5	14	9
Segment level (M):	40.80	30.43	29.89
Segment trend (b):	0.70	− 0.01	− 1.30
Significance tests:			
a) Add cumulative:	H = 22.603 DF = 2 $P < 0.001$		
b) Add differences:	H = 20.356 DF = 2 $P \leq 0.001$		
c) Trend cumulative:	H = 4.106 DF = 2 $P = 0.126$		
d) Trend differences:	H = 3.476 DF = 2 $P = 0.174$		

Figures 7 and 8 demonstrate the distribution of serial and white-washed residuals again. In both cases, an inspection of the distribution guarentees a sufficiently normal distribution, the prerequisite for valid application of significance tests. Figure 9 illustrates both the additive and slope effects of the white-washed data.

Table 2 illustrates how a continuous additive-cumulative effect is proven to be highly significant, across all three treatment phases, such that a constant decrease of the mood level can be observed after treatment. Thus, the test for evaluating additive differences also becomes highly significant. Accordingly, as shown in Figure 9, it is the interval between the BP and the sleep-deprivation phase in which the most significant difference in level can be found. In respect to slope effects, the experimental analysis fails to present any significant changes.

Table 3. Statistical aggregation of the individual time-series analyses

Significant effect	Frequences			
	Add cumulative		Add differences	
	−	+	−	+
Single sleep deprivation vs Several sleep deprivations	6 / 2	0 / 5	3 / 2	3 / 5
CHI^2:	4.273		0.048	
DF = 1	$P = 0.0365$		$P = 0.820$	
Fisher test:	$P = 0.0163$		$P = 0.412$	
Significant effect	Frequencies			
	Trend cumulative		Trend differences	
	−	+	−	+
Single sleep deprivation vs Several sleep deprivations	5 / 3	1 / 4	4 / 3	2 / 4
CHI^2:	0.853		0.090	
DF = 1	$P = 0.641$		$P = 0.761$	
Fisher test:	$P = 0.179$		$P = 0.761$	

All 13 curves of mood development were analyzed in the same way as described above by the experimental variant of the HTAKA model. In each case, in accordance with the ABA design, a BP, i.e., a therapy-oriented intervention phase, was allocated experimentally. This intervention phase is defined as an application of sleep deprivation plus the phase following such a deprivation-therapy period.

Table 3 illustrates the aggregation of the individual trend analyses with respect to the four trend aspects of the mood assessment data, using tests for fourfold contingency tables. Table 3 clearly indicates that one single sleep deprivation during any point of the HTAKA analyses does not lead to an additive-cumulative effect in any of the six cases. On the other hand, in 5 of the 7 cases, a multiple application of sleep-deprivation therapy did indeed produce an additive cumulative effect. This illustrates the significant difference between the two groups, concluding that only a multiple application of sleep-deprivation therapy actually leads to a longer-lasting improvement of mood, whereas an isolated application of sleep-deprivation therapy does not show any significant effect. With respect to additive differences, no significant difference can be ascertained between the two groups. Here, even a single application of a sleep-deprivation therapy led to a significant reduction of the mood values levels during the intervention phase in 3 of the 6 patients in comparison to the baseline phase. In these cases, no further reductions occurred during the remainder of the study period. With respect to trend parameters, neither significant cumulative nor differential differences could be observed between the two groups. Here, the aggregated results show that a single sleep-deprivation treatment, if at all, only produces a short-term effect, which is clearly linked to the time of the application.

Discussion

Using the evaluation of the effects of sleep deprivation within the setting of routine clinical treatment of endogenous-depressive patients as an example, a new and in several respects advantageous nonparametric time-series analysis procedure, the HTAKA model, was demonstrated. Since the data collection prerequisites for the utilization of this approach could easily be fulfilled, this model offers a suitable method for the analysis of data collected during routine inpatient treatment.

Two cases were presented in detail in which the long-term effect of repeated sleep deprivations was already evident by casual visual analysis of the time-series, and in which this effect could be further differentiated and confirmed statistically by the HTAKA model. At last, with respect to single-case analysis, it is precisely the long-term effects of sleep deprivation that have long been neglected by previous studies dealing with the subject. These previous studies were limited to "before and after" comparisons in the context of single isolated sleep-deprivation trials.

We are convinced that an intensive idiographic analysis of the single-case development must be presented for the evaluation of sleep-deprivation effects in single patients before the data for nomothetic conclusions on the effectiveness of this therapy can be aggregated, as has been performed in the second part of our analysis. Finally, it should be pointed out that the HTAKA model presents a compromise, considering its methodological and mathematical implications. At the present time, however, there seems to be no better procedure available for assessing individual time-series in therapeutical research. This procedure was developed with the conviction that, with perhaps the exception of a few research areas such as physiopsychology, ARIMA processes, i.e., constantly recurring processes, do not occur in clinical therapy-oriented psychology. Within therapy phases, many more positively than negatively cumulated processes take place. The characteristic feature of cumulation lies in its being deterministic. Therefore, a serial procedure is called for which, in addition to containing a

deterministic model, is implicitly able to separate and level out any further evident nuisance parts, be they either serial or deterministic. When designing such a procedure, the complexity of the conceptual background should not be underestimated: typical one-sided criteria on the relation between models and residuals are not capable of solving the problem. Therefore, we consider this procedure especially suited for therapy research in psychopathological studies.

References

Appelt H, Strauß B (eds) (1985) Ergebnisse einzelfallstatistischer Untersuchungen in Psychosomatik und klinischer Psychologie. Springer, Berlin Heidelberg New York Tokyo

Beneke M (1981) Zeitreihenanalyse in der klinischen Psychopharmakologie. Unveröffentlichte Dissertation, Universität Kiel

Bortz J (1979) Lehrbuch der Statistik für Sozialwissenschaftler. Springer, Berlin Heidelberg New York

Box GEP, Jenkins GM (1976) Time series analysis: forecasting and control. Holden-Day, San Francisco

Cook THD, Campbell DT (1979) Quasi-experimentation. Design & analysis. Issues for field settings. Houghton Mifflin, Boston London

Edwards AL (1971) Versuchsplanung in der psychologischen Forschung. Beltz, Weinheim Berlin Basel

Glass GV, Wilison VL, Gottman JM (1975) Design and analysis of time-series experiments. Colorado Associated University Press, Boulder, Colorado

Keeser W (1979) Zeitreihenanalyse in der klinischen Psychologie. Ein empirischer Beitrag zur Box-Jenkins Methodologie. Unveröffentlichte Dissertation, Universität München

Kirk RE (1982) Experimental design: Procedures for the behavioral sciences. Books Cole, Belmont, California

Kleiter E (1986) HTAKA. Hierarchische Trend-Abschnitt-Komponenten-Analyse. Ein Verfahren zur Analyse von Zeitreihen. Z Emp Päd Psychol, Beiheft 2

Kuhs H, Tölle R (1986) Schlafentzug (Wachtherapie) als Antidepressivum. Fortschr Neurol Psychiatr 54:341–355

Lienert GA (1978) Verteilungsfreie Methoden in der Biostatistik. Bd II. Anton Hain, Meisenheim am Glan

McCleary R, Hay RA (1980) Applied time series analysis for the social sciences. Sage, Beverly Hills London

Möbius HJ, Göricke G, Krüh P (1985) Parametrische und nichtparametrische Methoden der Einzelfallstatistik. In: Appelt H, Strauß B (eds) Ergebnisse einzelfallstatistischer Untersuchungen in Psychosomatik und klinischer Psychologie. Springer, Berlin Heidelberg New York Tokyo, pp 170–188

Möller HJ, Barthelmes H, von Zerssen D (1983) Forschungsmöglichkeiten auf der Grundlage einer routinemäßig durchgeführten Basis- und Befunddokumentation. Psychiatria Clin 16:45–61

Pfanzagl J (1964) Allgemeine Methodenlehre der Statistik I. de Gruyter, Berlin

Pflug B, Tölle R (1971) Therapie endogener Depressionen durch Schlafentzug. Nervenarzt 42:117–124

Revenstorf D, Keeser W (1979) Zeitreihenanalyse für klinische Daten. Methodik und Anwendung. Beltz, Weinheim Basel

Shine LC, Bower SM (1971) A one-way analysis of variance for single-subject designs. Educat Psychol Measurement 31: 105–113

Strauß B, Stemmler G (1985) Praktische Probleme bei der Anwendung von Zeitreihenanalysen. In: Appelt H, Strauß B (eds) Ergebnisse einzelfallstatistischer Untersuchungen in Psychosomatik und klinischer Psychologie. Springer, Berlin Heidelberg New York Tokyo, pp 139–156

Zerssen D von (unter Mitarbeit von Koeller DM) (1976) Klinische Selbstbeurteilungs-Skalen (KSb-S) aus dem Münchener Psychiatrischen Informationssystem (PSYCHIS München). Manuale: a) Allgemeiner Teil, b) Paranoid-Depressivitäts-Skala, c) Befindlichkeits-Skala, d) Beschwerde-Liste. Beltz, Weinheim

Zerssen D von (1986) Clinical Self-rating Scales (CSRS) of the Munich Psychiatric Information System (PSYCHIS München). In: Sartorius N, Ban TA (eds) Assessment of depression. Springer, Berlin Heidelberg New York Tokyo, pp 270–303

Received October 19, 1987