

Texte 35/02

Umweltforschungsplan des Bundesministers für Umwelt, Naturschutz und
Reaktorsicherheit

F+E-Vorhaben 298 25 235
UBA-FB 000285

Entwicklung statistischer Verfahren für Ringversuche, Methodvalidierung und Auswertung von marinen Daten

**Abschlussbericht Teilvorhaben II:
Zusammenstellung und Bewertung von statistischen Verfahren zur
Auswertung von biologischen und chemischen Daten in der MUDAB**

von

PD Dr. habil. Steffen Uhlig

unter Mitarbeit von

**Kirsten Simon
Norbert Schick
Susan Köhler
Daniel Rothmaler**

quo data
Gesellschaft für Qualitätsmanagement und Statistik mbH

Inhaltsverzeichnis

2.1	EINLEITUNG	2
2.1	STATISTISCHE METHODEN ZUR AUSWERTUNG MARINER DATEN.....	3
2.1.1	TRENDANALYSEN.....	3
2.1.2	EVALUATION VON MONITORINGPROGRAMMEN	3
2.2	DIE FUNKTIONEN EINER TRENDANALYSE	3
2.3	ANALYSE VON JAHRESDATEN	6
2.4	DER LOESS-SMOOTHER UNTER EINBEZIEHUNG VON MONATSDATEN ...	6
2.5	TRENDANALYSEN MIT KOVARIABLEN: QUECKSILBERGEHALTE VON SCHOLLEN AUS DER DEUTSCHEN BUCHT 1985 – 1999.....	7
2.6	TRENDANALYSEN FÜR ABUNDANZ UND BIOMASSE	7
2.7	VARIANZKOMPONENTEN FÜR EROD.....	8

2.1 Einleitung

Ziel des Vorhabens war die Zusammenstellung bislang verwendeter statistischer Methoden zur Auswertung mariner Daten im Rahmen des BLMP, eine Bewertung hinsichtlich ihrer Eignung sowie die Erarbeitung von Empfehlungen für die Anwendung bestimmter statistischer Verfahren für marines Monitoring. Um diese Arbeiten zu realisieren, wurde eine enge Zusammenarbeit sowohl auf nationaler als auch internationaler Ebene angestrebt. Die Ergebnisse dieser Zusammenarbeit sind in dem Abschlussbericht zusammengefasst.

2.1 Statistische Methoden zur Auswertung mariner Daten

Die im Rahmen des BLMP eingesetzten univariaten Verfahren sind für die meisten Anwendungen praktikabel, jedoch unter gewissen Voraussetzungen, speziell bei stark ausreißerbehafteten oder linksschiefen Verteilungen, möglicherweise in manchen Fällen irreführend, so dass sich eine zusätzliche Überprüfung der Daten anhand weiterer Kriterien empfiehlt. Hierfür geeignet sind z.B. Boxplots, welche die Verteilung der Daten grafisch anhand der Extremwerte und der Quartile verdeutlichen, oder nichtparametrische Parameter für Schiefe und Kurtosis, welche die Abweichungen von der Normalverteilung messen.

2.1.1 Trendanalysen

Jede der bislang zur Auswertung von marinen Daten im nationalen oder internationalen Rahmen verwendeten Trendanalysemethoden hat spezifische Vor- und Nachteile, so dass empfohlen wird, flexible Protokolle für die Trendanalyse zu erarbeiten, die den jeweiligen Anforderungen möglichst nahe kommen.

2.1.2 Evaluation von Monitoringprogrammen

Zur Beurteilung eines Monitoringprogramms hat sich die Durchführung von Poweranalysen bewährt. Es wird daher empfohlen, derartige Poweranalysen als Qualitätssicherungsinstrument für Monitoringprogramme generell durchzuführen, um Monitoringprogramme besser beurteilen zu können. Dies kann im Einzelfall auch dazu führen, dass ein Messprogramm korrigiert wird.

2.2 Die Funktionen einer Trendanalyse

Vergleicht man die unterschiedlichen in der Literatur vorgeschlagenen Verfahren zur Trendanalyse, so wird schnell deutlich, dass jedes dieser Verfahren seine spezifischen Vor- und Nachteile hat. Die Eignung eines Verfahrens ist nicht nur von der statistischen Verteilung der Daten abhängig zu machen, sondern auch vom angestrebten Zweck der

Trendanalyse. Folgende Funktionen einer Trendanalyse von Monitoringzeitreihen sind von besonderem Interesse:

A) Visuelle Darstellung der Daten durch eine Trendlinie

Eine visuelle Darstellung beispielsweise mittels einer Trendlinie ermöglicht einerseits eine Zusammenfassung der Einzelwerte und andererseits die Aufdeckung von Anomalien, z.B. von Ausreißern.

B) Statistischer Trend-Test

Ein statistischer Trend-Test ermöglicht die Feststellung, ob von einem statistisch gesicherten Trend ausgegangen werden kann. Andernfalls spricht nichts gegen die Anfangsvermutung, dass es sich bei den beobachteten Schwankungen nur um zufällige Veränderungen handelt, die nicht durch einen systematischen Trend verursacht sind. Nur wenn ein statistisch gesicherter Trend vorliegt, erscheint aus statistischer Sicht eine weitergehende statistische Auswertung der Zeitreihe als zweckmäßig.

C) Bestimmung der linearen Trendkomponente für einen Anstieg oder Rückgang

Eine Trendanalyse sollte die Ermittlung des zugrundeliegenden linearen Trends in einem vorgegebenen Zeitraum ermöglichen.

D) Vergleich des gegenwärtigen Niveaus mit einem Referenzniveau oder dem Niveau eines früheren Zeitpunkts

Sofern ein Referenzwert oder ein Referenzzeitpunkt gegeben ist, stellt sich im Rahmen der Trendanalyse auch die Frage, ob das aktuelle Niveau höher oder niedriger als der Referenzwert liegt oder ob sich seit dem Referenzzeitpunkt eine Erhöhung oder Reduzierung ergeben hat.

Alle hier genannten Funktionen lassen sich mit dem LOESS-Smoother, dem Verfahren von Theil, dem Trend-Test auf der Basis des LOESS-Smoother sowie dem Test von Mann-Kendall realisieren. Jedoch lassen sich nicht alle Funktionen durch ein und dasselbe Analyseverfahren gleichzeitig in völlig befriedigender Weise erfüllen. So ist zur Erfassung eines linearen Trends ein Schätzverfahren angemessen, welches eine lineare Trendfunktion unterstellt. Ein solches Schätzverfahren ist jedoch nicht zweckmäßig, um den generellen Trendverlauf abzubilden. Somit erscheint es sinnvoller, jede Funktion mit dem jeweils am besten geeigneten Analyseverfahren zu realisieren.

2.3 Analyse von Jahresdaten

Ein universell einsetzbares Instrument zur Analyse von Jahreswerten ist der LOESS Smoother. Einschränkend ist allerdings darauf zu verweisen, dass die Trendanalyse nach einem jeweils zu spezifizierenden Protokoll erfolgen sollte, welches die Funktionen der Trendanalyse ebenso wie die Natur der jeweiligen Daten berücksichtigt. Eine formale Trendüberprüfung sollte sowohl den linearen als auch den nicht-linearen Trendanteil überprüfen. Die Testergebnisse vervollständigen einander und geben einen ersten Eindruck von der Trendstruktur.

Interpretation der Testergebnisse:

		Test für die Linearität der Trendkomponente, Abschnitt 1.1.4 H_0 : linearer Regressionskoeffizient (slope) = 0	
		H_0 wird akzeptiert: slope = 0	linearer Trend erfasst (H_0 abgelehnt)
Test für Nichtlinearität, Abschnitt 2 (H_0 : Trend ist linear)	H_0 wird akzeptiert: No Nichtlinearität	Nur Zufallsvariationen; keine systematischen Veränderungen	linearer Trend
	Nichtlinearität erfasst (H_0 abgelehnt)	Nicht-monotone Trendmuster ohne allgemeine Tendenz	Umfassende systematische Veränderungen

2.4 Der LOESS-Smoother unter Einbeziehung von Monatsdaten

Während die Analyse von Jahresdaten naturgemäß keine Berücksichtigung von Saisonkomponenten erforderlich macht, ist letzteres bei der Analyse von Monatsdaten unumgänglich. Die Analyse von Monatsdaten mit dem LOESS-Smoother bietet den Vorteil, dass das Konfidenzintervall in der Regel schmaler ist, als wenn die Berechnung auf der Basis von Jahresdaten erfolgt. Weiterhin bietet die Analyse von Monatsdaten die Möglichkeit die dem LOESS Smoother zugrundeliegende Spanne zu reduzieren. Das ist besonders dann von Bedeutung, wenn angenommen wird, dass anthropogene Veränderungen innerhalb einer kurzen Zeitspanne wirksam werden. Wenn daher eine

befriedigende Anpassung der zugrundeliegenden Monatsdaten gelingt, ist eine Auswertung auf der Basis von Monatsdaten einer Auswertung auf Basis von Jahresdaten vorzuziehen. Vorteile ergeben sich insbesondere bei Zeitreihen aus einem Zeitraum von nicht mehr als 10 Jahren.

Es ist darauf hinzuweisen, dass die Anpassung auf Basis von Monatsdaten aufgrund der Datenlage vielfach nur bei Flussdaten möglich ist, während für die Messungen im Meer in der Regel eine gröbere zeitliche Auflösung verwendet werden muss. Grundsätzlich eignet sich das beschriebene Verfahren jedoch auch zur Anwendung auf Quartalswerte oder auf im Winter gemessene Nährstoffkonzentrationen, wenn man – aufgrund einer starken Saisonalität – die Einbeziehung der Sommerwerte vermeiden möchte.

2.5 Trendanalysen mit Kovariablen: Quecksilbergehalte von Schollen aus der Deutschen Bucht 1985 – 1999

Im Rahmen des Projektes wurde weiterhin eine Trendanalyse der Quecksilbergehalte von Schollen aus der Deutschen Bucht durchgeführt. Gegenstand der Trendanalyse ist einerseits die Abhängigkeit der Konzentration von Länge und Gewicht, und andererseits die Frage, ob für die Trendanalyse eine Betrachtung der individuellen Messergebnisse notwendig ist. Eine wichtige Schlussfolgerung aus der statistischen Analyse besteht darin, dass die Schadstoffkonzentrationen nicht nur anhand einer Größenklasse erfasst werden sollten, sondern möglichst von mehreren Größenklassen, um die Variation zwischen den Jahreswerten zu reduzieren und um potentielle Gefahren besser identifizieren zu können.

2.6 Trendanalysen für Abundanz und Biomasse

Für Abundanz und Biomasse ausgewählter Algengruppen wurde auf Basis der HELCOM-Daten umfangreiche Trendanalysen durchgeführt. Von artspezifischen Untersuchungen wurde aufgrund der Fülle des vorliegenden Materials zunächst abgesehen, zumal sich zeigte, dass die Einschränkung auf einzelne Arten zu höheren Anteilen fehlender Werte führt, welches wiederum methodische Probleme nach sich

zieht. Die Auswertungen erfolgten stationsweise jeweils unter Einbeziehung aller Daten der jeweiligen „Jahreszeit“. Alle Ergebnisse wurden in Datenblättern zusammengefasst.

2.7 Varianzkomponenten für EROD

Um zu überprüfen, inwieweit die EROD-Aktivität ein Biomarker ist, welcher für ein zeitliches Monitoring geeignet ist, werden die Varianzkomponenten berechnet, welche die Streuung der EROD-Aktivität zwischen verschiedenen Individuen bzw. zwischen Jahren charakterisiert. Um Störungen durch Kofaktoren so weit als möglich auszuschließen, erfolgt die Schätzung getrennt in zwei Datenmengen, die separat in zwei Perioden, März und Herbst, erhoben wurden. Um zu prüfen, ob die EROD-Aktivität im Herbst möglicherweise besser für ein zeitliches Monitoring geeignet ist, erfolgte eine weitere Analyse der EROD-Aktivitäten der Monate September, Oktober und November, die zusammengefasst betrachtet wurden, um ein ausreichend großes Datenkollektiv zu erhalten. Gleichwohl zeigte sich auch hier, dass die Power zum Nachweis eines zeitlichen Trends sehr gering ist, so dass ein zeitliches Monitoring wenig sinnvoll erscheint.

**Development of statistical methods for ring tests,
method validation and statistical analysis of
marine data**

**Final report, short version, part II:
Collection and evaluation of statistical methods for the analysis of
biological and chemical data in the MUDAB database**

Table of contents

2.1	INTRODUCTION	10
2.2	STATISTICAL METHODS FOR THE ANALYSIS OF MARINE DATA	11
2.2.1	ANALYSIS OF TRENDS	11
2.2.2	EVALUATION OF MONITORING PROGRAMS.....	11
2.3	FUNCTIONS OF TREND ANALYSIS	11
2.4	LOESS SMOOTHER FOR ANNUAL DATA	13
2.5	LOESS SMOOTHER FOR MONTHLY DATA.....	13
2.6	ANALYSIS OF TREND WITH COVARIABLES: HG CONCENTRATIONS IN PLAICE IN THE GERMAN BIGHT 1985 – 1999.....	14
2.7	ANALYSIS OF TRENDS FOR ABUNDANCE AND BIOMASS.....	14
2.8	VARIANCE COMPONENTS FOR EROD ACTICVITY	14

2.8 Introduction

The project focuses on statistical methods used for the analysis of marine data with the BLMP. Statistical methods are presented and evaluated with regard for an application in marine monitoring programs. The project has been realized in close co-operation on both the national and the international level.

2.9 Statistical methods for the analysis of marine data

The univariate parameters such as the arithmetic mean, the median or the confidence interval, used within the BLMP are applicable under a wide range of conditions. However, under certain conditions these methods may be misleading, especially under long-tailed or asymmetric distributions, and therefore an additional examination of the data is recommended. For such an examination non-parametric parameters measuring the skewness or the kurtosis could be applied. Boxplots, based on the extrema and the quartiles, are another graphical alternative for the examination of the empirical distribution of the data.

2.9.1 Analysis of trends

Each trend analysis method applied so far for marine data has specific advantages and disadvantages. It is therefore recommended to develop flexible protocols for their application, which take into account the specific purpose and requirements.

2.9.2 Evaluation of monitoring programs

An appropriate method for the evaluation of monitoring programs is the analysis of power of this program, i.e. the analysis of trend sensitivity for detecting a long term trend. It is therefore recommended to perform these power analyses as a general quality assurance instrument for assessing and improving monitoring programs.

2.10 Functions of trend analysis

If one compares the various methods for trend analysis it quickly becomes clear that each of these methods has its specific advantages and disadvantages. The suitability of a method depends not only on the statistical distribution of the data but also on the purpose of the trend analysis. In this regard the distinction between the following four functions of a trend analysis is of interest:

A) Visual representation of data by means of a trend line

A visual representation through a trend line for instance serves on the one hand to summate all individual values and, on the other, to reveal anomalies such as outliers.

B) Statistical trend test

A statistical trend test serves to examine the question whether the trend can be taken to be statistically validated. If this is not the case, then there is nothing that dismisses the initial assumption that the fluctuations observed are but random alterations which are not caused by a systematic trend. Only when there is evidence of a statistically validated trend does a more far-reaching statistical evaluation of the time series appear meaningful from the statistical point of view.

C) Recording the linear trend component

The linear trend component allows the percentage reduction in inputs in a given period to be recorded.

D) A comparison of the present level with a reference level or the level at an earlier point of time

Where a reference value or reference time is given, a trend analysis can be used to examine whether the current level is higher or lower than the reference value or whether there has been a rise or fall since the reference time.

All these functions can be realized with the LOESS smoother, the Theil slopes, the trend test based on the LOESS smoother and the test of Mann Kendall. However, each of these methods is focused on some functions, and is not equally well suited for each function. For the analysis of the linear trend component a method is appropriate which assumes a linear trend, but such a method is not suited for the analysis of a general (non-monotonic) trend. It is therefore recommended to realize each function with a specific analysis method.

2.11 LOESS smoother for annual data

A universal tool for the trend analysis of annual data is the LOESS smoother. It should be noted however, that the analysis of trends should be performed according to a specific protocol, that takes into account the purpose of the trend analysis and the type of distribution of the data. A formal trend test should include both the linear and the non-linear trend component. Test results complement each other and give a first impression of the type of the time series:

Interpretation of test results:

		Test for the linear trend component H_0 : linear regression coefficient (slope) = 0	
		Accept H_0 : slope = 0	linear trend detected (H_0 rejected)
Test for non-linearity (H_0 : trend is linear)	Accept H_0 : No non-linearity	random variation only, no systematic changes	linear trend
	Non-linearity detected (H_0 rejected)	non-monotonic trend pattern without general tendency	complex systematic changes

2.12 LOESS smoother for monthly data

The main difference between a trend analysis based on monthly data and on annual data is the necessity to include a term for the seasonal variation for the monthly data. On the other hand, the analysis of monthly data with the LOESS smoother offers the advantage that the confidence range is smaller than for the LOESS smoother based on annual data. Secondly, the analysis of monthly data offers the possibility to reduce the span. This might be especially relevant if anthropogenic changes are assumed to become effective within a few years. If a satisfying fit based on monthly data can be achieved, such an analysis is to be preferred against an analysis based on annual data, especially if the series are not longer than 10 years.

It should be noted that in many cases concentrations measurements are not performed on a monthly basis but less frequent. Monthly data are available for rivers, whereas for measurements in the sea an analysis based on e.g. quarterly data is more appropriate. However, the method presented can also be adapted for the analysis of quarterly or half-yearly data or can be restricted for the analysis of winter (or summer) data.

2.13 Analysis of trend with covariables: Hg concentrations in plaice in the German Bight 1985 – 1999

A trend analysis for the Hg concentrations in plaice in the German Bight was performed in order to assess the relationship with length and weight of the fish, and to examine whether the trend analysis can be performed on aggregated level or whether the individual fish data have to be taken into account. An important conclusion of the statistical analysis is that any trend analysis should not be based on one size class, but on several size classes, in order to reduce the interannual variation of the mean values and to allow better identification of potential risks.

2.14 Analysis of trends for abundance and biomass

For abundance and biomass of selected algae groups many trend analyses based on the HELCOM data were performed. No species-specific analysis were performed due to the high number of species and the high number of missing values (which makes any analysis more complicated). The analyses were performed separately for each site and were based on all data of a season. For all results a data sheet was produced.

2.15 Variance components for EROD activity

In order to examine whether the EROD activity is a biomarker which is suited for a temporal monitoring, the variance components were calculated that characterize the dispersion of the EROD activity within and between years. In order to exclude disturbances of other factors, the estimation was performed in two separate

populations, sampled in separate seasons, March and autumn. In order to examine whether the EROD activity in autumn is better suited for a temporal monitoring, EROD activity of September, Oktober and November was analysed. However, it appeared that the trend sensitivity, i.e. the power for detecting a trend is very small, so that a temporal monitoring program based on EROD activity is not very promising.