

UNIVERSITY OF OSLO
Department of informatics

**Chaotic Time
Series Analysis:
Improving the
Procedures**

Dimitris Kugiumtzis

Dr Scient thesis
Research Report 239

ISBN 82-7368-154-8

ISSN 0806-3036

February 25, 1997



Chaotic Time Series Analysis: Improving the Procedures

Dimitris Kugiumtzis

*Department of Informatics, University of Oslo,
P.O.Box 1080 Blindern, N-0316 Oslo, Norway*

February 25, 1997

Preface

This publication is my thesis for the *Doctor Scientiarum* degree at the Department of Informatics, University of Oslo, Norway. The thesis is mainly comprised of five research papers published, to be published, or submitted to international journals in the field of nonlinear dynamics and time series analysis. It includes also two review papers, made at the early stage of the work, on topics I investigated later in the research papers, and one research report.

The papers are

- 1 D. Kugiumtzis, B. Lillekjendlie, and N. Christophersen. Chaotic Time Series Part I: Estimation of Some Invariant Properties in State Space. *Modeling, Identification and Control*, Vol. 15, pp 205 – 224, 1994.
- 2 B. Lillekjendlie, D. Kugiumtzis, and N. Christophersen. Chaotic Time Series Part II: System Identification and Prediction. *Modeling, Identification and Control*, Vol. 15, pp 225 – 243, 1994.
- 3 D. Kugiumtzis. State Space Reconstruction Parameters in the Analysis of Chaotic Time Series – the Role of the Time Window Length. *Physica D*, Vol. 95, pp 13 – 28, 1996.
- 4 T. Aasen, D. Kugiumtzis, S. H. G. Nordahl. Procedure for Estimating the Correlation Dimension of Optokinetic Nystagmus Signal. *Computers and Biomedical Research*, in press, 1997.
- 5 D. Kugiumtzis. Assessing Different Norms in Nonlinear Analysis of Noisy Time Series. *Physica D*, in press, 1997.
- 6 D. Kugiumtzis. Correction of the Correlation Dimension for Noisy Time Series. *International Journal of Bifurcation and Chaos*, in press, 1997.
- 7 D. Kugiumtzis, and N. Christophersen. State Space Reconstruction: Method of Delays vs Singular Spectrum Approach. *Research Report* 0806 – 3036, No 236, 1997.
- 8 D. Kugiumtzis, O. C. Lingjærde, and N. Christophersen. Regularized Local Linear Prediction of Chaotic Time Series. submitted to *Physica D*, 1997.

Acknowledgements

The present scientific work was carried out at the Department of Informatics, University of Oslo, and was financed by the Norwegian Research Council. I would like to express my appreciation to both the Department for offering the best working conditions and the Council for the economical assistance.

There are many individuals that I am grateful to for their support and encouragement during my research. First, I am deeply indebted to Professor Nils Christophersen who has been my supervisor, main collaborator and supporter throughout the whole work. He encouraged me to start this project and helped me in administrative issues regarding my enrollment as a doctoral student. His ever-lasting enthusiasm, ample knowledge on linear and nonlinear dynamics and precious suggestions were the starting kick for the research work I pursued in the new field of chaotic time series analysis. Moreover, his high standards in scientific writing improved my technical writing skills. His criticism of my early drafts seemed sometimes a little too harsh, but I quickly realized that it was indeed appropriate, and also necessary to make the papers publishable.

I thank Research Manager Bjørn Lillekjendlie, the Center for Industrial Research (SINTEF), Oslo, for the fruitful discussions and collaboration we had during the early stage of my work. He was co-author on the two first review articles. I also thank Senior Engineer Torbjørn Aasen, Haukeland hospital, Bergen, for his collaboration and enthusiasm during the work on optokinetic data and the writing of the paper. I enjoyed the inspiring long discussions we had occasionally on and beyond working terms and I do recall how glad I was at that time to find a researcher engaged in the same problems. I would also like to thank the staff at the Department of Otolaryngology, Head and Neck Surgery, Haukeland hospital, Bergen, for their assistance and hospitality during my visits to their laboratory. I appreciate the kindness of the State Center of Epilepsy for providing the EEG data and especially Dr. Pål Larsson for his expert advice and help regarding the data.

My thanks goes to my colleague and friend Ole Christian Lingjærde who not only shared with me coffee breaks during long working evenings but was also co-author of the last paper. We seem to understand and complement each other well and I enjoyed working with him.

I thank the faculty members and graduate students at the Department of Informatics, University of Oslo, for many interesting discussions, help and support. Especially, I would like to thank Professor Knut Liestøl and Professor Sverre Holm for their support and assistance on my work. Both were members of the advisory committee of my doctor project. I appreciate also the assistance and patience of the technical staff of the Department.

I thank the master degree students who stimulated me with interesting discussions while I was following the course of their work.

Finally, I would like to express my gratitude to my dearest loving family, who filled me with confidence and love throughout my entire life and shared with me all difficulties up to the successful completion of this thesis.

The course of the work

The work started formally in January 1993 and was completed in January 1997. However, I had discussions with Professor Nils Christophersen on the subject of the thesis already from May 1992. We agreed the main goal to be the analysis of time series with the new and promising nonlinear methods mainly based on chaos theory. This was a rather new field not only at our Department, but in Norway in general. I started on this project and was given a three year financial support by the Norwegian Research Council which was later extended to four years by the Department of Informatics with additional tuition obligations.

Fortunately, Bjørn Lillekjendlie, an experienced researcher at the Center for Industrial Research (SINTEF), had interest in similar topics and we collaborated at this early stage. With the assistance of my supervisor Nils Christophersen, we made a literature survey with a critical standpoint on the new developments, and wrote two review papers which formed the frame for my forthcoming research (papers 1 and 2). This was an absolutely necessary process since the field was completely new to me and my collaborators.

At this stage there was an obvious need for concentration on some specific part of the vast field of chaotic time series analysis. It was reasonable to start with state space reconstruction, the first step of almost any time series analysis. This subject was treated exhaustively in the recent years and numerous methods were proposed but with little theoretical background. After long and painstaking work I summed up my research work on this topic to a paper which after many interventions by Nils Christophersen was finally ready to be sent for publication (paper 3). A part of this work, not included in paper 3, was presented later in a research report (paper 7). Meanwhile I was in contact with Torbjørn Aasen at Haukeland hospital, Bergen, who had applied some of the methods I was studying on the optokinetic data. We elaborated on some new ideas about the implementation of state space reconstruction to the correlation dimension estimation of these data. Moreover, we found some interesting results reported in paper 4.

The work in paper 3 inspired me with new ideas for research in the implementation of the methods based on the state space reconstruction, especially regarding the analysis of real world data. I soon found that the performance of some of the methods, and particularly the estimation of the correlation dimension, is sensitive to distance measures. Statistical analysis on the different measures of noisy distances led me to new results that I wrote in paper 5. Furthermore, based on these results, I proposed a modification of the method for the estimation of the correlation dimension from noisy data in paper 6.

I decided then to turn to the prediction of chaotic time series and employ my experience from state space reconstruction on this problem. At this time I appreciated the collaboration with the colleague Ole Christian Lingjærde who had already worked with many of these methods. Together we decided to investigate the applicability of different nonlinear models to the prediction of chaotic time series. Especially, we concentrated on the class of local linear models which turned out to be an interesting area for research. With the assistance of Nils Christophersen, we introduced and implemented new ideas in the context of local linear prediction. Some results from our work were reported in paper 8, but more results are expected by the continuation of this research work. However, I had to put a stop on the work for the completion of the thesis, as four years had already passed.

After this four year period of study and research I feel I have gained the skill and spirit of a new scientist. I believe I have contributed to some aspects of chaotic time series analysis, which was a completely new field to me when I started this work. I hope my work has elucidated some points in the implementation of nonlinear methods on chaotic time series and improved the procedures towards a more proper analysis.

Anyway, the scientific research is a continuous process that does not terminate after the completion of a goal, a degree here, it is a way of living.

Contents

1	Chaotic Time Series Analysis	1
1.1	Introduction	1
1.1.1	Dynamical systems and chaos	2
1.1.2	Evidence of chaos in time series	2
1.2	State Space Reconstruction	4
1.2.1	The objective	4
1.2.2	The methods	4
1.3	Correlation Dimension	6
1.3.1	The method	6
1.3.2	Correlation dimension as discriminative statistics	7
1.3.3	Interdistance measure	9
1.3.4	Correction of the correlation dimension estimation from noisy data	9
1.3.5	Conclusions	10
1.4	Local Prediction	10
1.4.1	Local linear prediction models	11
1.4.2	Regularization of the ordinary least squares	11
1.5	Conclusion and Suggestion for Future Work	12
	Reference list	14
2	Thesis papers	17
1.	Chaotic Time Series Part I: Estimation of Some Invariant Properties in State Space	19
2.	Chaotic Time Series Part II: System Identification and Prediction	41
3.	State Space Reconstruction Parameters in the Analysis of Chaotic Time Series – the Role of the Time Window Length	63
4.	Procedure for Estimating the Correlation of Optokinetic Nystagmus Signal	81
5.	Assessing Different Norms in Nonlinear Analysis of Noisy Time Series	107
6.	Correction of the Correlation Dimension for Noisy Time Series	129
7.	State Space Reconstruction: Method of Delays vs Singular Spectrum Approach	149
8.	Regularized Local Linear Prediction of Chaotic Time Series	163

Chapter 1

Chaotic Time Series Analysis

The introductory part begins with a short background on chaotic time series analysis. Then the focus is on selected aspects of the analysis. Particularly, the state space reconstruction, the estimation of the correlation dimension and local linear prediction are briefly discussed and the main results on each topic are presented with reference to the individual papers. Finally, suggestions for further research are put forth and the main conclusions are highlighted.

1.1 Introduction

There are two contrasting notions in life, *theory* and *practice*. Science is apparently related to theory but to date much of what is called “science” is actually practical work. Mathematicians are often considered as good representatives of theoreticians. They investigate mathematical problems, build theories, find solutions, but often have little concern for real life problems. On the other hand, in many disciplines, as biology and medicine, many scientists act more like practitioners. They have a certain real problem in mind, collect the data, apply some method, often directly from a software package, and draw conclusions based on the results of the output of the method, with little concern for the theoretical ground of the applied method. The interconnection of these two groups is desirable as many physical processes or natural phenomena can be considered as mathematical systems, from the dripping faucet to measles. Statisticians and methodologists seem to be the ones to fill the gap between theory and practice and bring the mathematical theory into practical and applicable methods.

For a real problem at hand, relevant data are collected and the objective is first to find evidence that supports a certain hypothesis for the cause that generate the data, and then build a model that reflects the cause and explains the data. The latter is often referred to as the *inverse problem*, “to find the cause given the effect”, where the cause is the underlying mechanism and the effect is the measured data. The work of the methodologist is to build the appropriate methods and implement them correctly on the given data in order to gain insight about the underlying mechanism and then model this mechanism. In some applications, *a-priori* knowledge may be sufficient to explain the cause and build a model based on *first principles*, and then the data are used only to validate the model and fix the exact parameters. However, there are many phenomena and processes for which we know little about their cause and then one must rely only on the available observations in order to find a model, the so-called *data-driven model*.

The work of the thesis deals with data-driven methods.

1.1.1 Dynamical systems and chaos

A natural or technical system that exhibits changes with time may be termed a dynamical system. Dynamical systems are typically classified as *deterministic*, if the laws controlling the mechanism are fully understood, and *stochastic*, if some part of the mechanism cannot be explained without introducing randomness. Until the early sixties, this was a clear and meaningful distinction. It was believed that the evolution of a deterministic system could be anticipated exactly given the initial conditions while the evolution of a stochastic system could be anticipated only in a statistical sense, i.e. as the development of averaged quantities influenced by chance.

First, Smale [31], [32] proved that there exist low-dimensional deterministic nonlinear systems sensitive to initial conditions, a phenomenon later termed *chaos* and the systems *chaotic* (from the greek word $\chi\acute{\alpha}\omicron\varsigma$, the opposite of order). As opposed to other regular deterministic systems, such as periodic or quasi-periodic systems, the trajectories generated by chaotic systems starting from two nearby points diverge exponentially with time and then converge and this feature is continuously repeated. Due to the divergence phenomenon, the largest Lyapunov exponent of a chaotic system is always positive. The stretching and folding of a trajectory in state space forms an object with fractal properties, called *strange attractor*. Strange attractors have typically non-integer fractal dimensions. In the last decade, there have been written many books about nonlinear dynamic theory with focus on chaotic systems and their properties [9], [2], [41], [15], [23] as well as on their applications [13], [21], [33].

Chaotic systems could be placed in a separate class between regular deterministic and stochastic systems. They seem stochastic when they are observed but there is no ambiguity in their mathematical description. In statistics, a challenging problem is, when observing the apparently random behavior of a system evolving in time, to determine whether the randomness is that of a high-dimensional system or is due to the chaotic behavior of a low-dimensional process. With high-dimensional systems, only the main structure of the system can be modeled deterministically, and the rest unspecified part is subsumed in one or more random components. On the other hand, low-dimensional systems bear deterministic description and investigation of their dynamical properties. In practice, this is a problem in statistical estimation, because the observations of the system will almost inevitably be subject to measurement errors. Even if the chaotic system is free from noise, due to its rich behavior it can be interpreted with respect to its invariant measure as a random system in equilibrium. There are some recent books dealing with the statistical aspects of chaotic systems [37], [3].

1.1.2 Evidence of chaos in time series

Poincarè in the 1890s was the first to anticipate the possibility of chaos in deterministic systems. However, chaos could not be observed before the invention of the high-speed computers. Computer experiments led to meteorologist Lorenz's discovery in 1963 of chaotic motion of a simple nonlinear model simulating atmospheric convection [16]. However, first after the 1970s, it was shown that chaos could explain real phenomena, e.g. in turbulence of fluids by the work of Ruelle and Takens [27] and in population biology by the work of May [18]. To date, it is well known that many physical processes can have chaotic behavior under certain conditions, and this has been shown with experiments in chemical reactions, electronic circuits, mechanical oscillators etc

[21]. It seems that chaos is abundant in nature, too. Many scientists believe that chaos is a key force behind many real phenomena such as weather patterns and stock markets. Though in some few cases chaotic behavior can be described mathematically, e.g. in a model for population dynamics, in general there is lack of complete evidence of chaos from observations. Probable reasons for this ambiguity is the inevitable effect of stochasticity in real phenomena that may impair the chaotic signature, and the shortcomings of the applied methods in detecting chaos from observations. For the former little can be done but for the latter the procedures may be improved.

The observations are often given in the form of measurements of a single quantity, the *scalar time series*. When the objective is to investigate chaos they are referred to as *chaotic time series*. Most of the methods used to analyze chaotic time series estimate properties of the chaotic dynamics. Classical linear methods are employed as well but more as preliminary steps than major tools. For example, the Fourier spectrum may be used to observe periodicities in the data indicating some form of determinism, regular or chaotic. A first idea about the data complexity may simply be obtained from time history plots or scatter plots (often referred to as plots of the pseudo-state space). For data of an oscillating type, i.e. data generated from a continuous pseudo-periodic system, Poincaré sections and return maps may be used to investigate chaotic behavior [4], [21], [22]. If the return map obtained from a Poincaré section is unimodal the original system is likely to have the characteristics of one-dimensional chaotic maps. It is often hard to construct appropriate Poincaré sections, especially for dimensions of the pseudo-state space larger than three. A unimodal map may also be inspected in a much simpler way, from the scatter plot of the successive maxima of the oscillations, the so-called *Lorenz's trick*. For example, fine unimodal curves obtained in this way are shown for two well-known simulated chaotic systems, the Lorenz and Rössler system, in [33].

Other more sophisticated methods are used to estimate system invariants, assuming the existence of a deterministic system. The most important invariants are the fractal dimension and its measures, e.g. the correlation and information dimension, the entropy and the Lyapunov exponents. The predictability of the underlying system approximated by some model may also be considered as an invariant of the system. The estimation of all these invariants requires the reconstruction of the state space from the chaotic time series.

The objective with all these methods varies with the application. In many applications, the invariant estimates are used as discriminative statistics, i.e. as a tool to investigate an hypothesis on the nature of the data. For example, invariant estimates may be used to distinguish chaos from white noise (as for the ST-interval of the ECG data that at first glance seems to be completely stochastic [30]), or from a linear stochastic system (typically using the method of surrogate data [36]). Moreover, they can be used to detect different states of the same system, as with the optokinetic data from patients and healthy persons (paper 4), and the electrical activity of the brain before and during an epileptic seizure [17]. Another use of the invariant estimates are to extract values in order to characterize the underlying mechanism. However, with real data, exact confident estimates are seldom obtained, but even approximative results can be useful. For example, an estimate of the correlation dimension to the closest larger integer gives an idea about the number of degrees of freedom of the underlying system.

There are few books dealing with the investigation of chaos from observations [38], [20] [12], but several mathematically oriented books about chaos devote a chapter to chaotic time series analysis [29], [22]. Moreover, there are a number of good research paper collections, e.g. [19], [13], [40], [24], and few review papers [6], [35], [8], [1].

This work deals with some of the topics on chaotic time series analysis, specifically the state space reconstruction, the estimation of the correlation dimension and the prediction with local linear models. The background for these topics is given in the two review papers (paper 1 and paper 2).

1.2 State Space Reconstruction

1.2.1 The objective

Suppose a scalar time series is given, $x(t) = x(k\tau_s)$ for $k = 1, \dots, N$ where τ_s is the sampling interval and N is the length of the time series. Assuming the original state space is a manifold M of some dimension $\lceil d \rceil$ (the smallest integer larger than the fractal dimension d of the attractor of the system), the evolution of the unknown deterministic system in discrete (or discretized) time is determined by the mapping

$$\mathbf{s}_{k+1} = \mathbf{f}(\mathbf{s}_k) \quad (1.1)$$

where \mathbf{s}_k is the state at time k and \mathbf{f} is the system function defined on M . The time series is a measured quantity of this dynamical system

$$x_k = h(\mathbf{s}_k) \quad (1.2)$$

where h is the measurement function $h : M \rightarrow \mathbf{R}$. Here, the simple noise-free case is considered, but in practice, a noise component is involved in the measurement process, called *observation noise*, or in the system evolution, called *dynamic noise*. In this work, we often assume that a small or moderate amount of observation noise is present.

The objective is to reconstruct a state space, so that the system dynamics on the reconstructed attractor is homeomorphic to the dynamics on the original attractor. For noise-free infinite time series generated from a dynamical system, Takens' theorem assures that generically all the dynamical properties on the original attractor in M can be preserved on the reconstructed attractor in \mathbf{R}^m , i.e. the transformation Φ from M to \mathbf{R}^m is an *embedding* [34] (see Fig.1.1). According to Takens, the sufficient dimension m for Φ to be an embedding is $m \geq 2\lceil d \rceil + 1$ (m is therefore called embedding dimension). In [28], this condition is relaxed to $m \geq 2d + 1$. Assuming the reconstruction is successful, the original system function \mathbf{f} on the original attractor can be approximated by a function F on the embedded attractor.

1.2.2 The methods

Several methods for reconstruction have been suggested and many techniques have been proposed to estimate their parameters but their validity is questionable. In paper 1, the two most prominent reconstruction methods and the techniques for estimating their parameters are discussed.

The most simple and popular method to reconstruct points in \mathbf{R}^m from scalar data is the *method of delays* (MOD). The reconstructed vector is

$$\mathbf{x}_k = [x_k, x_{k-\rho}, \dots, x_{k-(m-1)\rho}]^T \quad (1.3)$$

and $\tau = \rho\tau_s$ is the *delay time*. Obviously, the vector \mathbf{x}_k carries information from a segment of the time series from the time $(k - (m - 1)\rho)\tau_s$ to time $k\tau_s$ covering the *time window length* $\tau_w = (m - 1)\tau$. According to the theoretical condition $m \geq 2d + 1$,

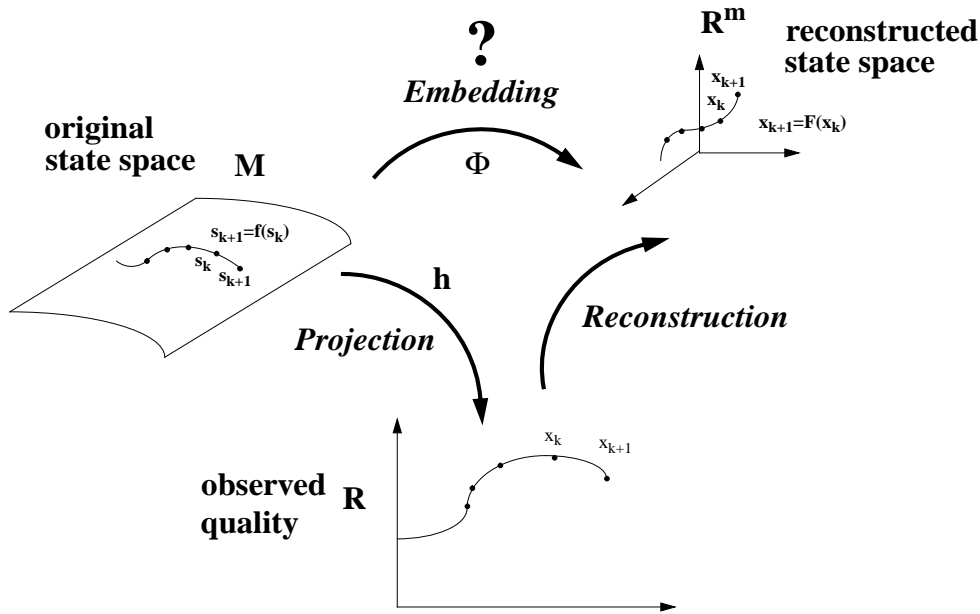


Figure 1.1: State space reconstruction from a time series generated by a dynamical system.

there are many possible values for the parameters m , τ , and thus τ_w , that yield valid reconstructions. However, with limited or noisy data the parameters must be carefully selected.

Many researchers have overlooked the importance of the parameter τ_w and concentrated on the parameters τ and m solely. The applicability of the techniques estimating τ and m are discussed in paper 3 with a critical view as to their validity. It turns out that the suggested estimates of m and τ are not always proper, and some examples are given in paper 3 showing that the results from some well-known estimation techniques are questionable. The main reason of the insufficiency of the estimates of τ and m is that these parameters are considered as independent but often the estimation of the one depends on the value of the other.

In paper 3, a new perspective for the reconstruction problem is given based on the role of τ_w and it is shown that this is the overall parameter for successful reconstruction. Once τ_w is determined, the selection of τ and m is of less importance. Certainly, there are still constraints on m . The condition $m \geq \lceil d \rceil$ should hold to avoid bad projections. On the other hand, in the implementation of some methods, a large m can cause numerically unstable solutions (as in the prediction with local linear maps, see paper 8). Similar constraints apply to τ , e.g. a small τ may also be inappropriate for some implementations (as in the measure of distances with the maximum norm, see paper 5).

The parameter τ_w yields any reconstruction method because in any case, the objective is to pass information from the samples within a time interval τ_w to a vector $\mathbf{x}_k \in \mathbf{R}^m$. Geometrically, the samples in the window $\tau_w = (m-1)\rho\tau_s$, i.e. $x_{k-(m-1)\rho}, x_{k-(m-1)\rho+1}, \dots, x_k$ can be considered as coordinates in \mathbf{R}^p , where $p = (\tau_w + 1)/\tau_s$. A reconstruction technique based on τ_w that gives m -dimensional point vectors \mathbf{x}_k can

thus be seen as a transformation $B : \mathbf{R}^p \rightarrow \mathbf{R}^m$. With MOD, B is a simple projection giving the subspace defined by the m selected coordinates $x_{k-(m-1)\rho}, x_{k-(m-2)\rho}, \dots, x_k$ employing the delay $\tau = \rho\tau_s$. Other reconstruction techniques implement more complicated B transformations, as the one employed in the Singular Spectrum Approach (SSA). This method yields first a transformation of the natural coordinate system to another orthogonal system using the SVD on the initial data matrix of dimension $p \times (N - p + 1)$, ranking the p new directions according to the variance they explain. Then the projection onto the m first directions completes the transformation B of SSA.

The SSA method seems to have proponents and opponents and there is some confusion regarding its applicability to chaotic time series. In paper 7, a fair comparison of MOD and SSA is attempted and it is shown with simulations that the two methods are equivalent under the same τ_w for noise-free data. For data corrupted with white noise, SSA gives better reconstructions for invariant estimation because it has an in-built filter for white noise.

In applications with short or noisy time series (or both), only a limited range of τ_w -values is likely to work. In paper 3, it is found that values of τ_w around the mean orbital period τ_p are the most appropriate, at least for the estimation of the correlation dimension that is used as a test tool. In the paper, evidence for this choice is given based on the dynamic properties of the underlying system. Dealing with real world data, τ_p is not known *a-priori* and has to be estimated from the data. Empirically, τ_p can be found from the form of the oscillations of the time series, detecting repeated patterns in the time series. The simplest repeated pattern is a single oscillation and then τ_p is estimated as the average time between peaks. For more complicated patterns filtering may be employed to facilitate the estimation of τ_w .

Concluding this chapter on state space reconstruction, paper 3 elucidated some points regarding the selection of the optimal parameters for reconstruction with MOD, established the importance of τ_w and proposed a simple empirical way to estimate the optimal τ_w from the mean orbital period τ_p . Moreover, the paper 7 showed with simulations that the two reconstruction methods MOD and SSA give equivalent results under the same τ_w . However, SSA can perform better with noisy data employing a state space filter. All these results were evaluated using estimation of the correlation dimension.

1.3 Correlation Dimension

Among all invariant measures, most attention has been given on the correlation dimension. The correlation dimension has been a standard tool for the investigation of chaos in many applications, as in physiology and geophysics. This measure has been popular because it gives a fairly good approximation of the fractal dimension and is easily computed.

1.3.1 The method

The strange attractor generated from a chaotic system is typically a fractal object, i.e. it is self-similar on different scales. A fractal object, or fractal, is characterized by a fractal dimension, a non-integer number less than the topological dimension of the object. Considering the object as a set of points, its fractal dimension is related to the point distribution in the state space. In Fig.1.2, six fractal objects are shown, the three first

are strange attractors generated by chaotic maps (logistic, Henon, Ikeda), and the other three are strange attractors generated by chaotic flows (Rössler, Lorenz, Rabinovich). The attractor of the logistic map covers almost completely the line segment $(0, 1)$ (except a countable infinity of periodic points), and the fractal dimension is $d \simeq 1$. The Henon attractor seems to form lines and has $d \simeq 1.21$, the fractal dimension is closer to one than to the topological dimension two. The Ikeda attractor has a fair distribution on the plane justifying that $d \simeq 1.6$. Similarly, the Rössler and Lorenz attractors seem to lie on two planes and have $d \simeq 2.01$ and $d \simeq 2.06$, respectively, while the Rabinovich attractor seems to be more expanded in three directions and has $d \simeq 2.19$.

The correlation dimension measures the point distribution of the attractor in state space. For time series, the state space is the reconstructed state space \mathbf{R}^m . For the points in this space, the *correlation integral* is computed first. The correlation integral for a distance r is the average number of points on the attractor that are less than r units apart

$$C(r) = \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(r - \|\mathbf{x}_i - \mathbf{x}_j\|), \quad (1.4)$$

where $\Theta(x)$ is the Heavyside function ($\Theta(x) = 0$ if $x < 0$ and $\Theta(x) = 1$ if $x \geq 0$) and $\|\cdot\|$ is the norm for distance measuring [7]. Temporally close points are usually omitted in the computation of the average in eq.1.4. The correlation dimension ν is given by the scaling law $C(r) \sim r^\nu$ for $r \rightarrow 0$ and $N \rightarrow \infty$. All strange attractors possess this scaling property. In practice, ν is found from the local slope of $\log C(r)$ vs. $\log r$ over a range of small r -values, where $C(r)$ is computed over the available N data. In order to obtain a reliable estimate of ν given a time series, the same scaling should be observed for increasing embedding dimensions establishing the so-called *saturation* property. For flows, the reconstruction parameter to be tuned is not m but the time window length τ_w as was argued in paper 3.

A legitimate estimate of the correlation dimension ν should fulfill the two conditions of clear scaling and saturation. Both of the conditions hold for noise-free infinite chaotic time series and arbitrary reconstructions satisfying Takens' criterion. For finite and noisy time series, the reconstruction setup must be carefully chosen in order to maintain scaling and saturation.

In paper 3 and paper 7, the estimation of ν was used as a tool to assess different reconstructions. Moreover, in paper 3, ν was estimated from real data including the Taylor Couette experiment [5] and EEG data from epileptic seizures, and it was found that confident estimates of ν could be obtained for suitably selected τ_w around the mean orbital period τ_p . For noisy limited data in general, there is no guarantee that confident estimates of ν can be obtained regardless of the reconstruction setup. For example, scaling may be masked by noise completely.

1.3.2 Correlation dimension as discriminative statistics

The estimation of ν may be employed in order to determine the fractal dimension, exactly or approximately depending on the quality and nature of the data, or in order to distinguish different data types. The former requires reliable and confident ν estimates while the latter requires only statistically significant differences of the ν estimates from the different data types.

Correlation dimension estimation was used in paper 4 as a discriminative statistics to distinguish healthy subjects from patients of various vertigo diseases based on optokinetic data (OKN). The correlation integral was computed for points reconstructed

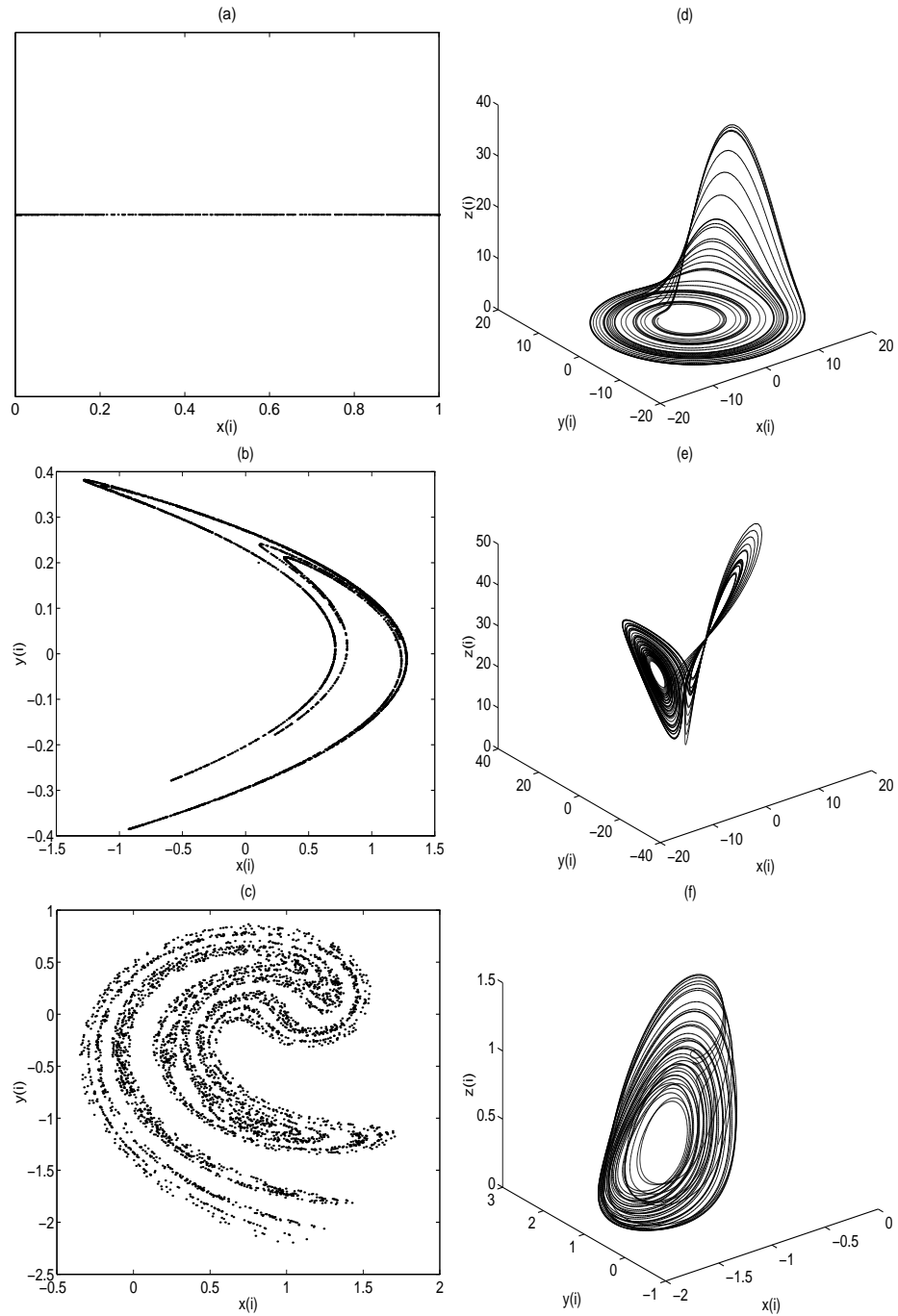


Figure 1.2: Strange attractors of chaotic systems: (a) Logistic map, $x_{i+1} = 4x_i(1-x_i)$, (b) Henon map [10], (c) Ikeda map [11], (d) Rössler flow [26], (e) Lorenz flow [16], and (f) Rabinovich-Fabrikant flow [25].

with SSA and for increasing τ_w . Scaling and saturation were investigated with an automatic process for each OKN time series. A statistical test on 10 healthy and 10 sick subjects showed significant difference in the ν estimates of the two groups.

1.3.3 Interdistance measure

The estimation of the correlation dimension from noise-free data is straightforward and without problems, apart from the selection of suitable reconstruction parameters to account for the limited length of the data. When there is noise in the data, the setup for estimation has to be carefully examined for robustness to the noise factor. Particularly, the measure of point distances may be sensitive to noise.

In paper 5, the effect of noise on the computation of point distances with different norms when estimating ν was studied. In the presence of noise, a norm may be considered as a stochastic variable being a function of the noisy point difference vector ($\mathbf{x}_i - \mathbf{x}_j$, see eq.1.4). A statistical analysis of the three most prominent norms, the taxicab norm L_1 , the Euclidean norm L_2 , and the maximum norm L_∞ , in measuring noisy point distances revealed distinct statistical differences. For example, the L_1 norm is more sensitive to noise when the point difference vector is close to some coordinate axes of \mathbf{R}^m and the L_∞ when it is close to the diagonal or antidiagonal. On the other hand, L_2 is not sensitive to any particular direction, it has a so-called circular dependency on the point difference vector. Based on these statistical properties, the drawbacks of the norms in the estimation of ν from noisy data can be anticipated.

In paper 5, the drawbacks of the norms were investigated in a systematic way and using corrupted data from many different chaotic maps and flows. The following results were reached:

- L_2 is the most robust norm to noise for measuring point distances,
- L_∞ gives the least biased ν estimates for noisy time series generated from chaotic maps,
- the performance of L_1 and L_∞ on flows varies with the application and the reconstruction setup.

. The results above do not pertain only the correlation dimension estimation but any method estimating scaling properties from point distances, i.e. other dimension measures and entropy.

1.3.4 Correction of the correlation dimension estimation from noisy data

Noise corrupts the scaling on the range of r distances up to the noise amplitude, and then ν can be estimated only from the scaling for larger r . For systems retaining the scaling property only for small r or in applications with moderate or large noise levels the estimation of ν deteriorates. There are two different approaches to solve this problem. One is to filter out the noise from the time series (nonlinear filtering [14]) and then use the filtered time series to estimate ν . The other is to compensate for the effect of noise in the computations directly when estimating ν . The latter approach was adopted in paper 6.

The correction algorithm for the estimation of ν presented in paper 6 relies on a nice statistical property of the L_2 norm, i.e. the circular dependency of the norm on

the point difference vector. When noise is added on the point difference vector the norm is positively biased. Subtracting the bias, the noise-free point distance can be estimated. Alternatively, when a noisy point distance is compared to a given distance r , one can reduce r with a quantity that accounts for the bias and retrieve the noise-free case. This is implemented to correct the estimation of ν . The correlation integral $C(r)$ of the noisy data is computed as before (using the L_2 norm) and a new r' is found from the statistical bias of the L_2 norm. Then the corrected estimate of ν can be found from the local slope of the graph $\log C(r)$ vs $\log r'$ over a range of r around the noise amplitude. The method was tested on many simulated data corrupted with different levels of white noise, and the original ν -estimates could be retrieved up to moderate noise levels (for most systems up to 10% – 15%). Moreover, the correction algorithm was applied to data from the Taylor Couette experiment and from epileptic EEG. In both cases, a more clear scaling could be observed after the correction was applied probably due to the unmasking of the scaling for distances around the noise amplitude.

1.3.5 Conclusions

Many researchers to date disregard the correlation dimension estimation in the analysis of chaotic time series claiming that there is little information to gain using this tool. They rely on other tools such as the Lyapunov exponents. On the other hand, many papers are still published presenting results on the estimation of ν from real world data of limited length for which the estimation of Lyapunov exponents would be difficult because it requires larger data size.

It is fair to say that the correlation dimension has been a useful tool in chaotic time series analysis. In this work, it has been used as a tool to assess the reconstruction setup (paper 3 and paper 7) and as discriminative statistics to distinguish different types of optokinetic data (paper 4). Moreover, the algorithm for the computation of ν was studied with focus on the distance of noisy points. Some interesting results were derived regarding the appropriateness of the different norms in the estimation of ν from noisy data (paper 5). Further, a new algorithm was proposed to compensate for the effect of noise in the estimation of ν (paper 6).

1.4 Local Prediction

Time series prediction has traditionally been a popular subject and an active area of research. Concerning chaotic time series, prediction has often been used as another tool to investigate chaos, treating nonlinear short-term predictability as an invariant property of the underlying system. If the underlying system is chaotic the prediction error (due to error in the initial conditions and in the model fitting) amplifies exponentially with time. For regular deterministic data, no amplification of error with time should be observed, and for stochastic data, the error may grow following another power law [39]. Moreover, the rate of the error amplification indicates the complexity of the chaotic mechanism.

Apart from the dynamical information one may seek by applying nonlinear prediction, another more challenging objective would be to achieve best predictions for the problem at hand. This is the case with many applications in different disciplines, from biology, to astrophysics and economics. It has to be mentioned that in the chase of optimal future point estimates, the statistical confidence of these estimates is often overlooked and most models are used without any concern for the variance of the

predicted values.

In chaotic time series prediction, the models are defined directly from the data (data driven or empirical models) and often they give little insight into the nature of the underlying process. Therefore they are often referred to as black-box models. In paper 2, a review of this large category of models is made, and the models are classified further as global models (e.g. polynomial maps), local models (linear maps, weighted simplexes), and semilocal models (neural networks, radial basis functions). The research work of the thesis on prediction is made on the local linear prediction models.

1.4.1 Local linear prediction models

The rationale behind local linear models is to define a linear least square problem locally for each target point a prediction is desired. So, the global prediction model is actually the concatenation of many local linear models, one per target point. To elucidate, assume that measurements up to time i of some quantity are obtained and the value at time $i + 1$ is sought. Assume also that points are reconstructed from the data following the MOD method with a suitable selection of m and τ (or ρ). Then any linear prediction algorithm consists of the following three steps:

1. Find the neighbors of the target point $\mathbf{x}_i = [x_i, x_{i-\rho}, \dots, x_{i-(m-1)\rho}]^T$.
2. Assume a linear model for the neighbors and their one-step ahead mappings and find the parameters of the model.
3. Find the one-step prediction applying the model to the target point.

The multi-step prediction can be found directly from the linear model that approximates the multi-step ahead mapping. Alternatively an iterative scheme can be applied, i.e. starting with the scheme for one-step prediction, construct a target point for time $i + 1$, repeat the one-step prediction and so on until the desired time is reached. There are plenty of modifications and improvements of the standard local prediction scheme presented above. Worth mentioning is weighting of the neighbors in order to account for the level of closeness of the neighbor points to the target point. However, in the work presented below the standard setup is adopted and the new ideas and modifications concern only the linear model.

1.4.2 Regularization of the ordinary least squares

The ordinary linear model is the straightforward solution of the least squares, solving the linear problem of m unknowns (assuming the local points are centered first) in k equations, where k is the number of neighbors. The problem has a least norm solution for $k \geq m$ but when m is close to k the solution has a large variance. The solution deteriorates significantly and systematically when there is noise in the data, and then there is need for regularization of the ordinary least square (OLS) solution. This is the subject of paper 8.

In linear regression analysis, a number of regularization techniques have been developed but they have been overlooked, or at least not yet applied, in local linear prediction. Many of these techniques are well established in various application areas, as the principal component regression (PCR) and the partial least squares (PLS), while others are less used in applications but have appealing statistical properties such as the ridge regression (RR) and the truncated total least squares (TTLS). As shown in paper

8, all these regularization techniques have somehow different statistics but their common feature is that they reduce the variance of the solution offering some bias. Thus they may give more confident predictions and perform better than OLS in the presence of noise or with k close to m (see paper 8).

All the regularization techniques except TTLS seem to work better than OLS on noisy data, and are actually compatible with other more sophisticated prediction tools such as neural networks. Moreover, they compensate for the drawback of OLS with few neighbor points (small k) and they are especially useful in applications with low data density, i.e. time series of short length.

The most robust of the regularization techniques seems to be the PCR, which is simply a rank reduction of the OLS. It gives often the best predictions on noisy data when the rank reduction is properly selected. Applications on the real sunspot data showed that the regularization techniques, and particularly PCR, is at least compatible to neural networks and outperform the classical bench-mark threshold autoregressive model. However, the results on local linear prediction should be considered preliminary and there is ongoing work on this topic.

1.5 Conclusion and Suggestion for Future Work

The work in this thesis has been concentrated on three important topics in the study of chaotic time series: state space reconstruction, correlation dimension and local linear prediction. These topics were first reviewed by the author and his collaborators in two review papers (paper 1 and paper 2).

The importance of the time window length τ_w in state space reconstruction was established (paper 3). It was shown that the quality of the reconstruction is more dependent on τ_w than on the reconstruction method, MOD or SSA, or its specific parameters (paper 7). A new setup for reconstruction was designed selecting the τ_w from the mean orbital period τ_p and the appropriateness of this reconstruction scheme was tested on many types of data (paper 3).

In the estimation of the correlation dimension, the reconstruction has to vary in order to evaluate the saturation property. According to the proposed reconstruction setup, the saturation is expected to be observed for $\tau_w \geq \tau_p$ if it can be observed at all (paper 3). This approach was adopted in the estimation of the correlation dimension ν of the optokinetic data. The estimation of ν was applied to time series of 10 healthy and 10 sick subjects and a significant difference between the two groups was found (paper 4). However, the ensembles were small and a meta-analysis with many optokinetic data is required to assess this result. The number of optokinetic data gathered to date are not sufficient to start such an analysis.

Further, the focus was on the algorithm for the estimation of ν from noisy data. It was shown that the use of different norms to measure the distances of the noisy points may give different results (paper 5). The statistical analysis of the L_1 , L_2 and L_∞ norm, showed that L_1 and L_∞ are sensitive to some particular directions in state space while L_2 is not. The estimation of ν with the three norms on different reconstruction setups and different data types confirmed the robustness of the L_2 norm (paper 5). Then a correction scheme for the estimation of ν from noisy data was designed based on the systematic bias of the L_2 norm (paper 6). The correction scheme performs well on data corrupted with white noise but in some applications the measurements may involve correlated noise. Further work is needed here, to modify the correction scheme or fix the reconstruction properly to account for correlated noise. Applications of the

correction scheme to data from the Taylor Couette experiment and from the epileptic EEG gave promising results. It would be interesting to apply the corrected estimation of ν together with the proposed reconstruction setup to other real world time series that are expected to have a moderate level of observation noise. There is a continuing communication with the State Center of Epilepsy and a project along these lines will soon start focusing on EEG data before and at epileptic seizure.

Concerning local linear prediction, the main contribution was to introduce the regularization techniques, which are well known in other areas, but little referred to or used in chaotic time series analysis. It turned out that these techniques are actually very useful when the data are corrupted with noise and outperform the ordinary least squares model (paper 8). The preliminary results on local linear prediction using regularization are promising and there is ongoing work on this topic. It seems that for one-step prediction, a small time window length τ_w (or embedding dimension m for fixed delay τ) is appropriate but for predictions further into the future a larger τ_w (or m) is required. The latter suggests the use of τ_w up to the mean orbital period τ_p , which was found to be optimal for the invariant estimation. For large m (close to the number of neighbors k), the ordinary least squares suffers and is not suitable for predictions while the regularization techniques seem to be robust to the increase of m . Thus a topic for investigation is whether a large m , such that $\tau_w \simeq \tau_p$, would improve the predictability of the regularization methods for prediction times larger than τ_p . Moreover, in the multi-step prediction, an *a priori* estimate of the uncertainty of the prediction in each step may be taken into account to design an automatic scheme that changes between direct and iterative prediction. There is ongoing work also on this topic.

There are still other topics for further investigation, especially on the local linear prediction, but any further research should be accompanied with a broad application to real world data.

Bibliography

- [1] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring. Analysis of observed chaotic data in physical systems. *Reviews of Modern Physics*, 65(4):1331 – 1392, 1993.
- [2] D. K. Arrowsmith and C. M. Place. *An Introduction to Dynamical Systems*. Cambridge University Press, 1990.
- [3] O. E. Barndorff-Nielsen, J. L. Jensen, and W. S. Kendall. *Networks and Chaos – Statistical and Probabilistic Aspects*. Chapman & Hall, 1993.
- [4] P. Bergè, Y. Pomeau, and C. Vidal. *Order within Chaos: Towards a Deterministic Approach to Turbulence*. Wiley, New York, 1984.
- [5] A. Brandstater and H. Swinney. Strange attractors in weakly turbulent Couette-Taylor flow. *Physical Review A*, 35:2207 – 2220, 1987.
- [6] J. P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57(3):617 – 656, 1985.
- [7] P. Grassberger and I. Procaccia. Characterization of strange attractors. *Physical Review Letters*, 50:346 – 349, 1983.
- [8] P. Grassberger, T. Schreiber, and C. Schaffrath. Non-linear time sequence analysis. *International Journal of Bifurcation and Chaos*, 1:521 – 547, 1991.
- [9] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*. Springer-Verlag, New York, 1983.
- [10] M. Hénon. A two-dimensional map with a strange attractor. *Communications in Mathematical Physics*, 50:69 – 77, 1976.
- [11] K. Ikeda. Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system. *Optics Communications*, 30:257, 1979.
- [12] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge, 1997. in press.
- [13] J. H. K. Kim and Stringer J. *Applied Chaos*. John Wiley & Sons, Inc, 1992.
- [14] E. J. Kostelich and T. Schreiber. Noise reduction in chaotic time-series data: A survey of common methods. *Physical Review E*, 48(3):1752 – 1763, 1993.
- [15] A. J. Lichtenberg and M. A. Liebermann. *Regular and Chaotic Dynamics*. Springer-Verlag, New York, second edition, 1992.

- [16] E. N. Lorenz. Deterministic nonperiodic flow. *Journal of atmospheric sciences*, 20:130, 1963.
- [17] H. Madsen. Non-linear methods for the analysis of electroencephalograms (EEG). Master's thesis, Department of Informatics, Oslo, 1995. 186 pages (in Norwegian).
- [18] R. M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459, 1976.
- [19] G. Mayer-Kress, editor. *Dimensions and Entropies in Chaotic Systems*. Springer-Verlag, Berlin, 1986.
- [20] A. Medio. *Chaotic Dynamics: Theory and Applications to Economics*. Cambridge University Press, Cambridge, 1992.
- [21] F. C. Moon. *Chaotic and Fractal Dynamics: an Introduction for Applied Scientists and Engineers*. John Wiley and Sons, New York, 1992.
- [22] A. H. Nayfeh and B. Balachandran. *Applied Nonlinear Dynamics: Analytical, Computational, and Experimental Methods*. Wiley & Sons, 1995.
- [23] E. Ott. *Chaos in Dynamical Systems*. Cambridge University Press, Cambridge, 1993.
- [24] E. Ott, T. Sauer, and J. A. Yorke, editors. *Coping with Chaos*. Wiley & Sons, 1994.
- [25] M. I. Rabinovich and A. L. Fabrikant. Stochastic self-modulation of waves in nonequilibrium media. *Soviet physics, Journal of experimental and theoretical physics*, 50:311, 1979.
- [26] O. E. Rössler. An equation for continuous chaos. *Physics Letters A*, 57:397 – 398, 1976.
- [27] D. Ruelle and F. Takens. On the nature of turbulence. *Communications in Mathematical Physics*, 20:167 – 192, 1971.
- [28] T. Sauer, J. A. Yorke, and M. Casdagli. Embedology. *Journal of Statistical Physics*, 65:579 – 616, 1991.
- [29] H.G. Schuster. *Deterministic Chaos: An Introduction*. Verlagsgesellschaft, Weinheim, 1989.
- [30] S. D. Skipitaris. Nonlinear time series analysis of ECG (electrocardiogram) signals. Master's thesis, Department of Informatics, Oslo, 1995. 113 pages.
- [31] S. Smale. Diffeomorphisms with many periodic points. In S. S. Cairns, editor, *Differential and Combinatorial Topology*, pages 63 – 80. Princeton University Press, Princeton, 1963.
- [32] S. Smale. Differentiable dynamical systems. *American Mathematical Society Bulletin*, 73:747 – 817, 1967.
- [33] S. H. Strogatz. *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering*. Addison-Wesley, 1994.

- [34] F. Takens. Detecting strange attractors in turbulence. In D. A. Rand and L. S. Young, editors, *Dynamical Systems and Turbulence, Warwick 1980*, Lecture Notes in Mathematics 898, pages 366 – 381. Springer, Berlin, 1981.
- [35] J. Theiler. Estimating fractal dimension. *Journal of the Optical Society of America A*, 7(6):1055 – 071, 1990.
- [36] J. Theiler, S. Eubank, A. Longtin, and B. Galdrikian. Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, 58:77 – 94, 1992.
- [37] H. Tong, editor. *Dimension Estimation and Models*, Singapore, 1993. World Scientific.
- [38] A. A. Tsonis. *Chaos: From Theory to Applications*. Plenum Press, New York, 1992.
- [39] A. A. Tsonis and J. B. Elsner. Nonlinear prediction as a way of distinguishing chaos from random fractal sequences. *Nature*, 358:217 – 220, 1992.
- [40] A. S. Weigend and N. A. Gershenfeld. *Time Series Prediction: Forecasting the Future and Understanding the Past*. Addison-Wesley Publishing Company, Reading, 1994.
- [41] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer-Verlag, New York, 1990.

Chapter 2

Thesis papers

Paper 1

**Chaotic Time Series Part I:
Estimation of Some Invariant Properties in State Space**

D. Kugiumtzis, B. Lillekjendlie, and N. Christophersen

Modeling, Identification and Control, Vol. 15, pp 205 – 224, 1994

Paper 2

**Chaotic Time Series Part II:
System Identification and Prediction**

B. Lillekjendlie, D. Kugiumtzis, and N. Christophersen

Modeling, Identification and Control, Vol. 15, pp 225 – 243, 1994

Paper 3

**State Space Reconstruction Parameters
in the Analysis of Chaotic Time Series
– the Role of the Time Window Length**

D. Kugiumtzis

Physica D, Vol. 95, pp 13 – 28, 1996

Paper 4

**Procedure for Estimating the Correlation
of Optokinetic Nystagmus Signal**

T. Aasen, D. Kugiumtzis, and S. H. G. Nordahl

Computers and Biomedical Research, in press, 1997

Paper 5

**Assessing Different Norms in Nonlinear
Analysis of Noisy Time Series**

D. Kugiumtzis

Physica D, in press, 1997

Paper 6

**Correction of the Correlation Dimension
for Noisy Time Series**

D. Kugiumtzis

**International Journal of Bifurcation and Chaos,
in press, 1997**

Paper 7

**State Space Reconstruction:
Method of Delays vs Singular Spectrum Approach**

D. Kugiumtzis, and N. Christophersen

Research Report 0806-3036, No 236, 1997

Paper 8

**Regularized Local Linear Prediction
of Chaotic Time Series**

D. Kugiumtzis, O. C. Lingjærde, and N. Christophersen

submitted to Physica D, 1997

Department of Informatics

February 1997

Chaotic Time Series Analysis: Improving the Procedures

Dimitris Kugiumtzis

Dr. Scient. thesis

Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO