1 of 1

LA-54B--93-223
RECEIVED
JAN 27 1994
OSTI

RESEARCH ACTIVITIES OF MARTIN C. CASDAGLI
AT SFI Sept 89 - May 91

My research over this period has focused on the theory and application of non-linear time series forecasting. The purpose of time series forecasting is to predict apparently random behavior about which we have little or no knowledge. A time series is a set of values that change in time, for example, daily stock market prices for a given set of firms, or the volume of the polar ice caps over the last two million years. Recent developments in the understanding of chaotic dynamical systems have stimulated new approaches to time series analysis. These results raise the possibility that apparently *random* behavior may essentially be the consequence of chaotic behavior of *deterministic* equations with a small number of variables. My main research effort has been to develop a theory of nonlinear time series forecasting which applies to chaotic time series with low levels of observational noise. I have also developed a new algorithm for analyzing experimental time series data in order to determine whether a nonlinear deterministic, as opposed to nonlinear stochastic model is an appropriate description of the data, and have applied this algorithm to a wide range of time series data. In addition to involvement in these and other research activities, I have helped organize a conference on nonlinear modeling and forecasting, jointly sponsored by NATO and SFI. I describe the above projects in more detail in the following paragraphs. Several of these projects benefited greatly from the excellent computer facilities available at SFI.

State Space Reconstruction.

I have developed a theoretical approach to state space reconstruction in collaboration with Doyne Farmer, Stephen Eubank and John Gibson of Los Alamos National Laboratory. This theory sheds light on the following practical questions: To what extent are existing techniques for nonlinear forecasting optimal? How might they be improved on? What are the fundamental limitations to such techniques? In order to construct a nonlinear forecasting model from a univariate time series, the first step is to reconstruct a state space from the time series, and the next step is to fit a multivariate nonlinear function to the time series data in this state space. A theorem of Takens shows that it is always possible to achieve the first step *perfectly* in the absence of noise. Using geometric insights into Takens' theorem, we have formulated a criterion, called "distortion", for optimal state space reconstruction in the presence of noise. In the limit of low observational noise levels, we derive a simple expression for the distortion. This expression quantifies the notion of "information flow" between observed and unobserved variables, and allows for numerically efficient comparison of existing reconstruction techniques. Furthermore, we are able to analytically solve for the optimal reconstruction technique, and find a new technique, which we call "local principal components analysis" which minimizes distortion. This new technique is directly applicable to the forecasting of time series data. We have also derived scaling laws which govern the behavior of the distortion in certain limits. When applied to high dimensional "complex" dynamical systems, these scaling laws imply that, in the

presence of a small amount of noise, knowledge of a univariate time series is insufficient to make even short term forecasts. Consequently, such time series behave more like those from random processes than deterministic chaos.

Deterministic versus stochastic nonlinear modeling.

I have developed an exploratory technique for investigating how much of the irregularity in an aperiodic time series is due to low-dimensional chaotic dynamics, as opposed to stochastic or high-dimensional dynamics. The idea is to construct nonlinear models with a variable smoothing parameter which at one extreme defines a deterministic model, and at the other extreme defines a linear stochastic model. Intermediate values of the smoothing parameter define nonlinear stochastic models. The accuracy of the resulting short-term forecasts as a function of the smoothing parameter reveals much about the underlying dynamics generating the time series. I have applied the technique to a variety of experimental and naturally occurring time series data generated from: electronic circuits, fluid turbulence, flame dynamics, speech, electroencephelograms, measles epidemics and sunspots. In contrast to dimension algorithms, the results show that this algorithm is not the sort of uncritical procedure that identifies low-dimensional chaos everywhere. On the other hand, the algorithm detected strong evidence for nonlinear forecastability in several of the time series.

Other projects.

In a collaboration with John Gibson, Doyne Farmer and Stephen Eubank, we have shown that the principal components analysis introduced into time series analysis by Broomhead and King is related in a precise sense to modeling a time series using derivative coordinates. This research has lead us to a deeper understanding of the advantages and limitations of principal components analysis as applied to chaotic time series.

In a collaboration with Timothy Sauer and James Yorke of the University of Maryland, we have extended Takens' theorem in several directions. It has been numerically observed that the embedding dimension predicted by Takens' theorem can be decreased for many purposes of time series modeling, and that the conclusions of Takens' theorem hold with full measure, rather than just generically. Our mathematical results shows how these observations are a consequence of extensions to Takens' theorem.

Motivated by conversations with Norman Hunter of Los Alamos National Laboratory, I have formulated a dynamical systems approach to analyzing time series from input-output systems. An input-output system is one in which there are two forms of observed time series: the input (or stimulus) driving the system, and the output (or response) of the system. For example in vibration testing experiments, a mechanical system is often driven by a controlled series of random inputs to simulate a noisy environment, and both the inputs and outputs from the system are available for analysis. Norman Hunter has applied the new approach to vibration testing data, and in strongly nonlinear cases obtained forecasts superior to those obtained with existing techniques. He also identified nonlinearity in ice age data, where the input

is taken to be the solar intensity, as estimated from celestial mechanics.

NATO-SFI Conference

Together with Doyme Farmer and Stephen Eubank, I organized a conference on nonlinear modeling and forecasting at Sante Fe in September 1990. We are currently editing submissions from the invited speakers into a proceedings volume. This volume will bring together articles from authors coming from many different disciplines, and I expect it to be a valuable source of references. The conference also stimulated interest in a time series competition, to be held at the Santa Fe Institute in the Fall of 1991. I have also been involved in writing review articles on nonlinear modeling and forecasting for other conference proceedings volumes.

Biographical note.

Due to lack of suitable job opportunities in the academic sector, I have decided to pursue my research activities in the private sector, starting July 1991. Although there is interest in nonlinear time series forecasting in the academic sector, the current recession has severely limited the possible job opportunities. Also, nonlinear time series forecasting is an interdisciplinary activity, and as such has not established a secure niche within the university departmental system. Possibly this situation will change in the future.

REFERENCES

Complete details of my research results at SFI may be found in the following references.

- [1] M.Casdagli, S.Eubank, J.D.Farmer and J.Gibson. *State Space Reconstruction in the Presence of Noise*. To appear in Physica D (1991)
- [2] M.Casdagli. *Chaos and Deterministic versus Stochastic Nonlinear Modeling*. To appear in J. Roy. Stat. Soc. (1991)
- [3] J.Gibson, M.Casdagli, S.Eubank, J.D.Farmer. *Principal Components Analysis and Derivatives of Time Series*. Los Alamos preprint (1991)
- [4] T.Sauer, J.Yorke and M.Casdagli. *Embedology*. To appear in J. Stat. Phys. (1991)
- [5] M. Casdagli. *Modeling Input-Output maps: A Dynamical Systems Approach*. In: *Nonlinear Modeling and Forecasting*. M.Casdagli and S.Eubank, editors, Addison-Wesley (1991), to appear.
- [6] *Nonlinear Modeling and Forecasting*. M.Casdagli and S.Eubank, editors, Addison-Wesley (1991), to appear.
- [7] M. Casdagli. *Nonlinear forecasting, chaos and statistics*. In: *Modeling complex phenomena*. L.Lam and V.Naroditsky, editors, Springer-Verlag, to appear (1991)
- [8] M.Casdagli, D.Des Jardins, S.Eubank, J.D.Farmer, J.Gibson and J.Theiler. *Nonlinear modeling of chaotic time series: theory and applications*. In: *EPRI Workshop on Applications of Chaos*. J.Kim and J.Stringer, editors, to appear (1991)

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DATE

FILMED

2 / 14 / 94

END