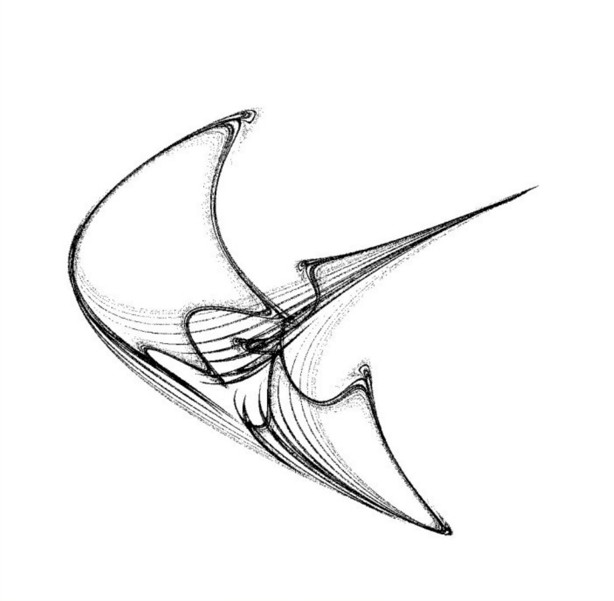
Applications of Neural Networks in Time-Series Analysis

Adam Maus
Computer Science Department

Mentor: Doctor Sprott
Physics Department

Outline

- Discrete Time Series
- Embedding Problem
- Methods Traditionally Used
- Lag Space for a System
- Overview of Artificial Neural Networks
  - Calculation of time lag sensitivities
  - Comparison of model to known systems
- Comparison to other methods
- Results and Discussions



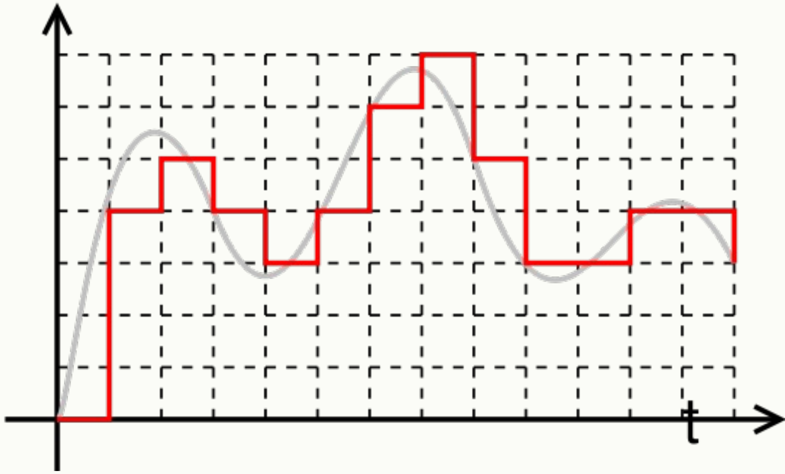
Discrete Time Series

- Scalar chaotic time series
  - i.e. average daily temperatures

[45, 34, 23, 56, 34, 25, …]

- Data given discrete intervals
  - i.e. seconds, days, etc.

- We assume a combination of past values in the time series predict the future



Discrete Time Series (Wikimedia)

Time Lags

- Each discrete interval refers to a dimension or time lag

1st Time Lag

[45, 34, 23, 56, 34, 25, …]

Current Value

2nd Time Lag

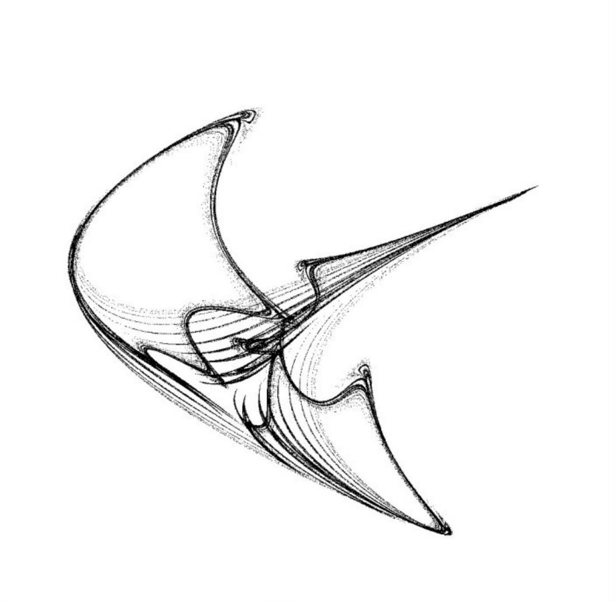
2nd Time Lag ….

$$x_k = x_{k-2} + x_{k-4} + x_{k-5}$$

Current Value

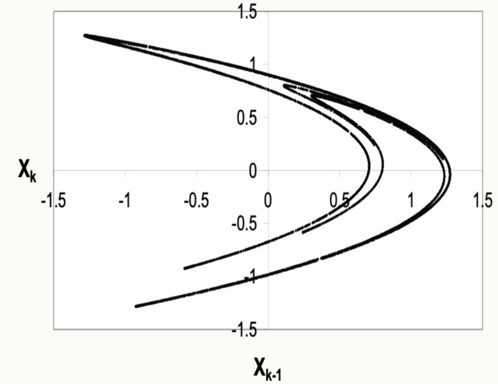
Outline

- Discrete Time Series
- **Embedding Problem**
- Methods Traditionally Used
- Lag Space for a System
- Overview of Artificial Neural Networks
  - Calculation of time lag sensitivities
  - Comparison of model to known systems
- Comparison to other methods
- Results and Discussions



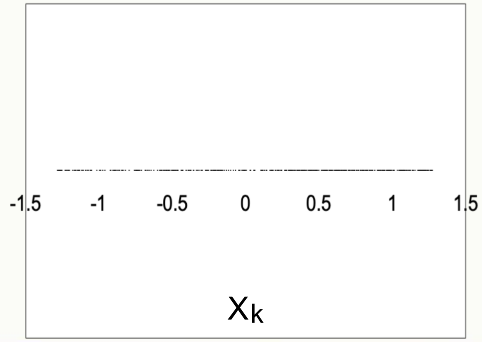
Embedding Problem

Problem: How do we choose the optimal embedding dimension that the model can use to unfold the data?



- The embedding dimension is related to the minimum number of variables required to construct the data

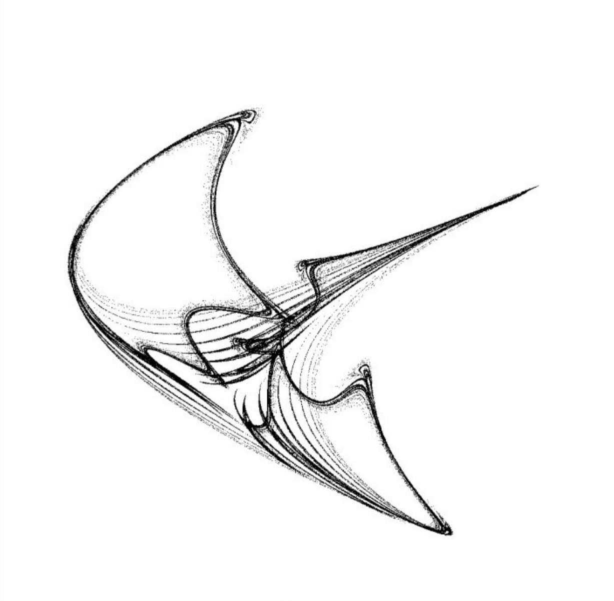
Or



- Exactly how many time lags are required to reconstruct the system without any information being lost but without adding unnecessary information
  - i.e. a ball seen in 2d, 3d, and 3d + Time

Outline

- Discrete Time Series
- Embedding Problem
- **Methods Traditionally Used**
- Lag Space for a System
- Overview of Artificial Neural Networks
  - Calculation of time lag sensitivities
  - Comparison of model to known systems
- Comparison to other methods
- Results and Discussions



ARMA Models

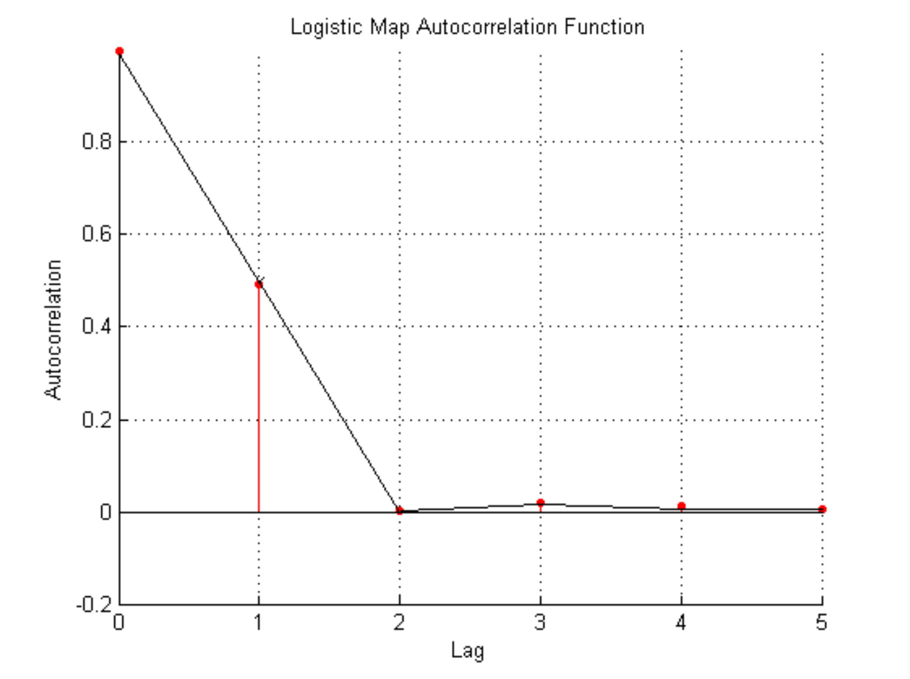
- Autoregressive Moving Average Models
  - Fits a polynomial to the data based on linear combinations of past values

$$x_k = \beta + \varepsilon_t + \sum_{i=1}^{d} \alpha_i x_{k-i} + \sum_{i=1}^{q} \Theta_i \varepsilon_{k-i}$$

  - Produces a linear function
  - Can create very complicated dynamics but has difficulty with nonlinear systems

Autocorrelation Function

- Finds correlations within data
- Much like the ARMA model, shows weak periodicity within nonlinear time series.
- No sense of the underlying dynamical system

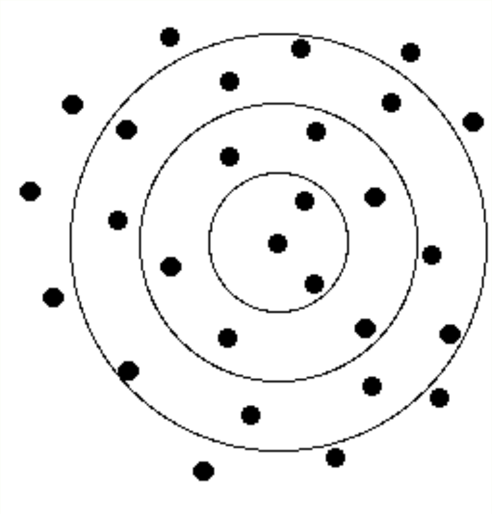


Logistic Map

$$x_k = 4x_{k-1}\left(1 - x_{k-1}\right)$$

Correlation Dimension

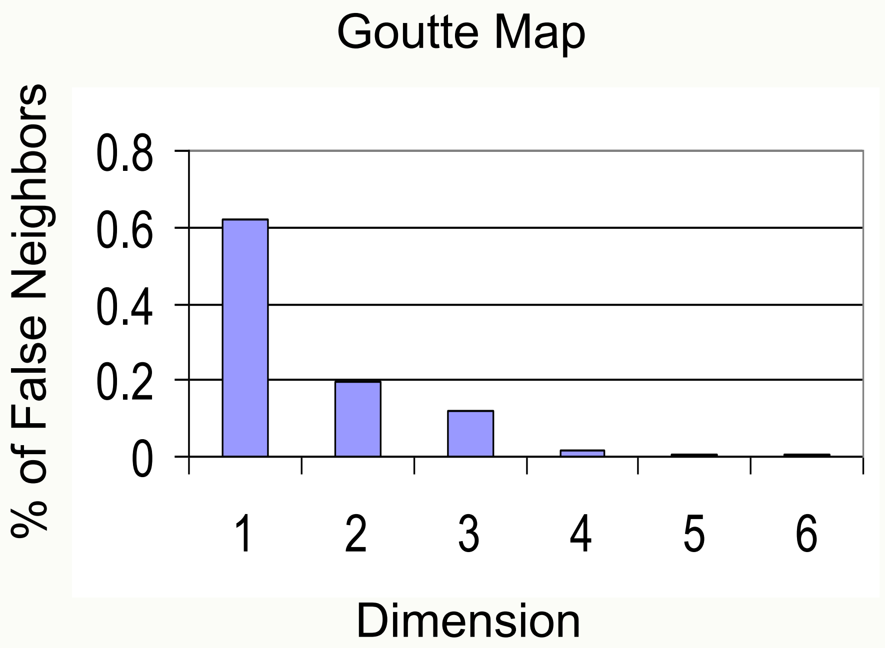
- Introduced in 1983 by Grassberger and Procaccia to find the fractal dimension of a chaotic system
- One can determine the embedding dimension by calculating the correlation dimension in increasing dimensions until it ceases to change
- Good for large datasets with little noise



Measuring the Strangeness of Strange Attractors (1983)

False Nearest Neighbors

- Introduced in 1992 by Kennel, Brown, and Abarbanel
- Calculation of false nearest neighbors in successively higher embedding dimensions
- As d is increased, the fraction of neighbors that are false drops to near zero
- Good for smaller datasets and rather robust to noise



[1992 Paper on False Nearest Neighbors]

Outline

- Discrete Time Series
- Embedding Problem
- Methods Traditionally Used
- **Lag Space for a System**
- Overview of Artificial Neural Networks
  - Calculation of time lag sensitivities
  - Comparison of model to known systems
- Comparison to other methods
- Results and Discussions

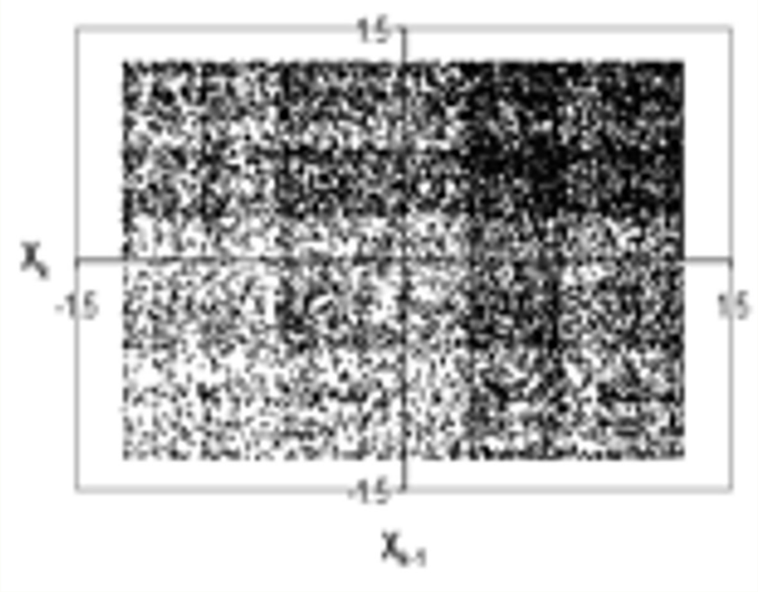
$$x_k = x_{k-2} + x_{k-4} + x_{k-5}$$

$$x_k = 0x_{k-1} + x_{k-2} + 0x_{k-3} + x_{k-4} + x_{k-5}$$

Lag Space

- Not necessarily the same dimensions as embedding space
- Goutte Map dynamics depend only on the second and fourth time lag

Goutte Map



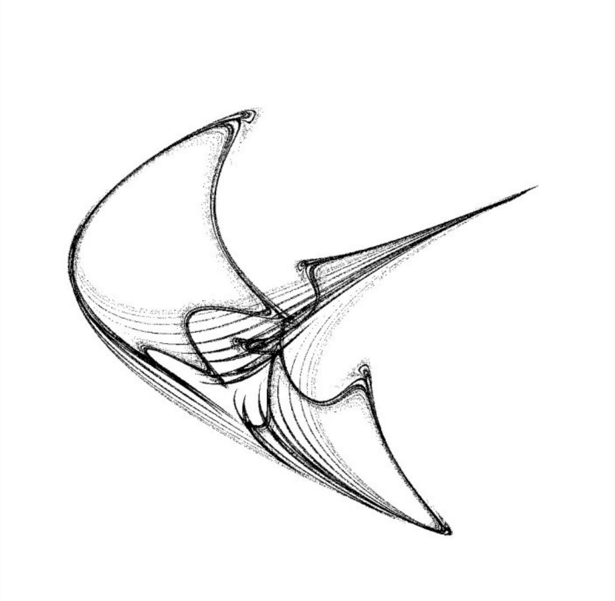
Problem: How can we measure both the embedding dimension and lag space?

$$x_k = 1 - 1.4 x_{k-2}^2 - 0.3 x_{k-4}$$

Lag Space Estimation In Time Series Modelling (1997)

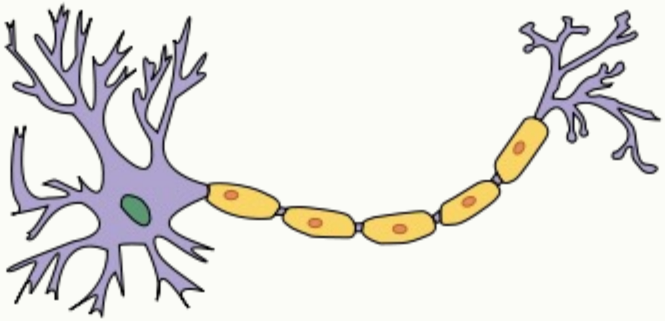
Outline

- Discrete Time Series
- Embedding Problem
- Methods Traditionally Used
- Lag Space for a System
- **Overview of Artificial Neural Networks**
  - Calculation of time lag sensitivities
  - Comparison of model to known systems
- Comparison to other methods
- Results and Discussions



Artificial Neural Networks

- Mathematical Models of Biological Neurons
- Used in Classification Problems
  - Handwriting Analysis
- Function Approximation
  - Forecasting Time Series
  - Studying Properties of Systems



$$y_k = \phi\left(b_k + \sum_{j=1}^{m} a_{kj} x_j \right)$$

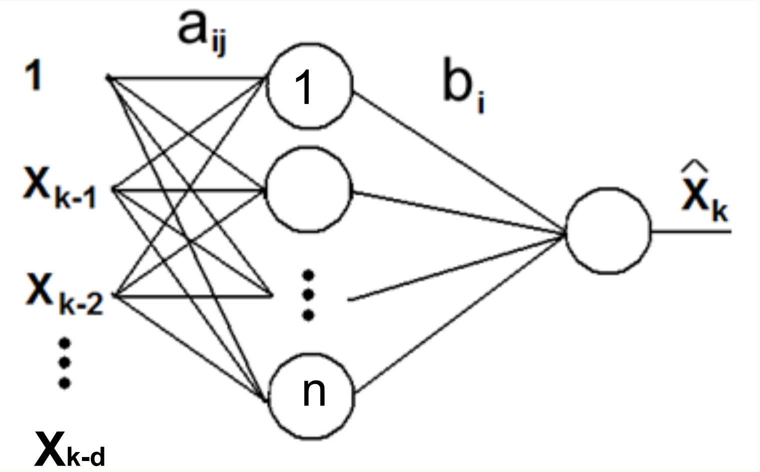
Function Approximation

- Known as Universal Approximators

- The architecture of the neural network uses time-delayed data

  x = [45, 34, 23, 56, 34, 25, …]

$$\hat{x}_k = \sum_{i=1}^{n} b_i \tanh\left(a_{i0} + \sum_{j=1}^{d} a_{ij} x_{k-j} \right)$$

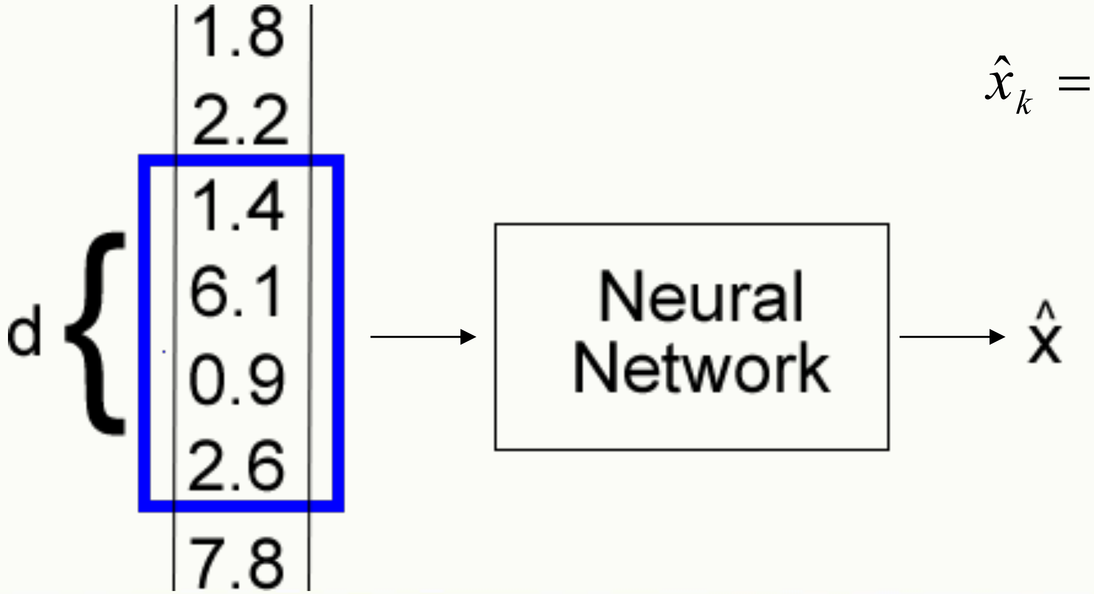


Structure of a Single-Layer Feed forward Neural Network

Multilayer Feedforward Networks are Universal Approximators (1989)

Function Approximation

- Next Step Prediction
  - Takes *d* previous points and predicts the next step



$$\hat{x}_k = \sum_{i=1}^{n} b_i \tanh\left(a_{i0} + \sum_{j=1}^{d} a_{ij} x_{k-j} \right)$$

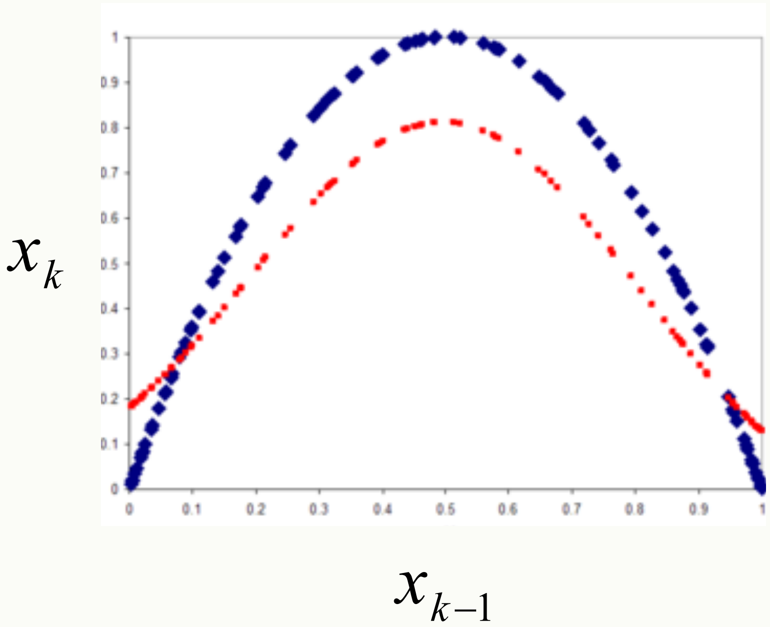
Training

1. Initialize *a* matrix and *b* vector
2. Compare predictions $\hat{x}_k$
   to actual values x_k

$$\text{Mean Square Error} = \frac{\sum_{k=1}^{c} \left(\hat{x}_k - x_k\right)^2}{c}$$

c = length of time series

3. Change parameters accordingly
4. Repeat millions of times



Fitting the model to data (Wikimedia)

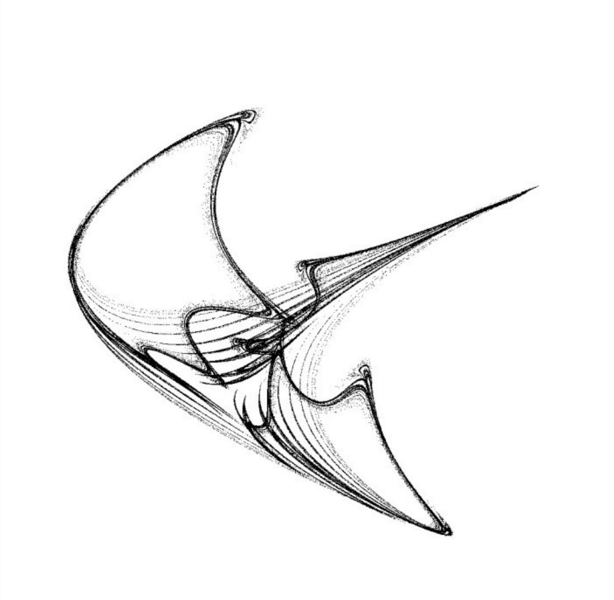
Convergence

- The global optimum is found at the lowest mean square error
  - Connection strengths can be any real number
  - Like finding the lowest point in a mountain range

- Numerous low points so we must devise ways to avoid these local optimum

Outline

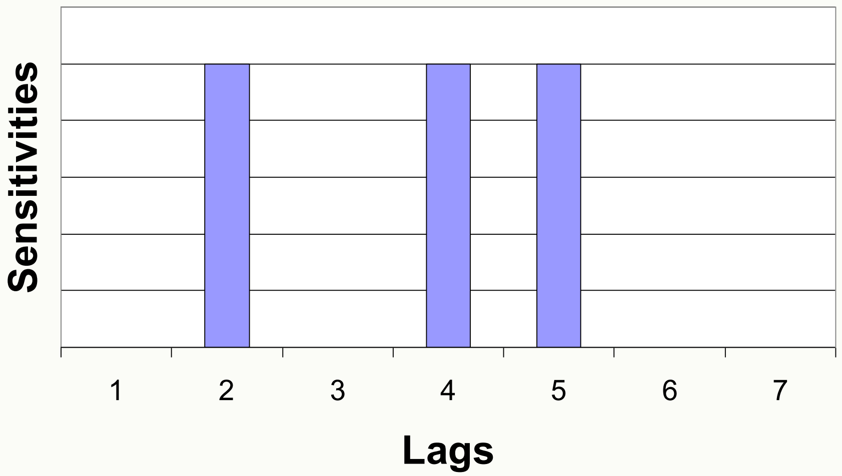
- Discrete Time Series
- Embedding Problem
- Methods Traditionally Used
- Lag Space for a System
- Overview of Artificial Neural Networks
  - **Calculation of time lag sensitivities**
  - Comparison of model to known systems
- Comparison to other methods
- Results and Discussions



Time Lag Sensitivities

- We can train a neural network on data and study the model
- Find how much the output of the neural network varies when perturbing each time lag
- “Important” lags will have higher sensitivity to changes in values

$$x_k = x_{k-2} + x_{k-4} + x_{k-5}$$

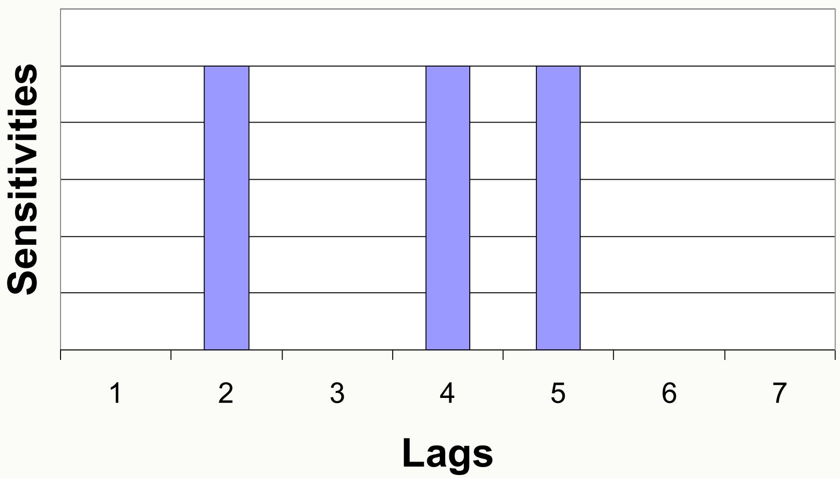


Time Lag Sensitivities

- We estimate the sensitivity of each time lag in the neural network:

$$\hat{S}(j) = \frac{1}{c-j} \sum_{k=j+1}^{c} \left| \frac{\partial \hat{x}_k}{\partial x_{k-j}} \right|$$

$$\frac{\partial \hat{x}_k}{\partial x_{k-j}} = \sum_{i=1}^{n} a_{ij} b_i \operatorname{sech}^2 \left(a_{i0} + \sum_{m=1}^{d} a_{im} x_{k-m} \right)$$



Expected Sensitivities

- For known systems we can estimate what the sensitivities should be

$$S(j) = \frac{1}{c-j} \sum_{k=j+1}^{c} \left| \frac{\partial \hat{x}_k}{\partial x_{k-j}} \right|$$

$$x_k = 1 - 1.4x_{k-1}^2 - 0.3x_{k-2}$$

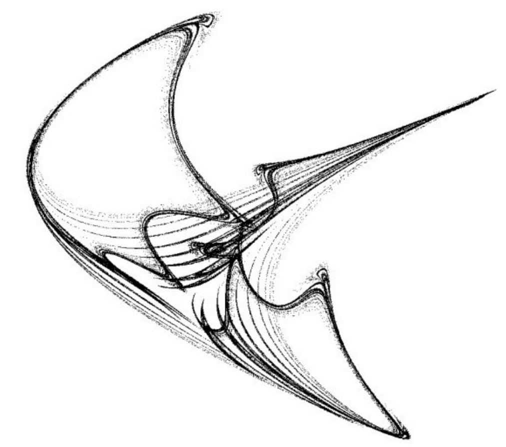
$$\frac{\partial x_k}{\partial x_{k-1}} = 2.8x_{k-1}$$

$$\frac{\partial x_k}{\partial x_{k-2}} = 0.3$$

- After training neural networks on data from different maps the difference between actual and expected sensitivities is <1%

Outline

- Discrete Time Series
- Embedding Problem
- Methods Traditionally Used
- Lag Space for a System
- Overview of Artificial Neural Networks
  - Calculation of time lag sensitivities
  - **Comparison of model to known systems**
- Comparison to other methods
- Results and Discussions

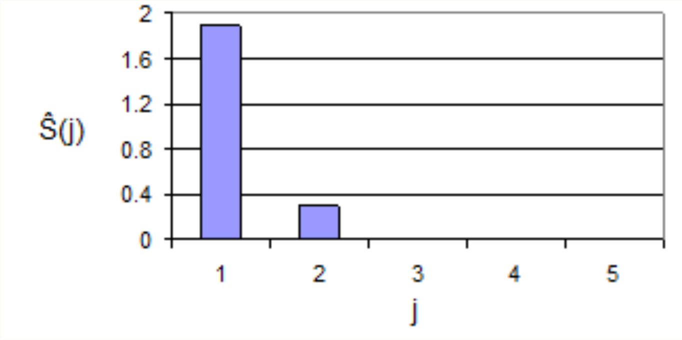


Hénon Map

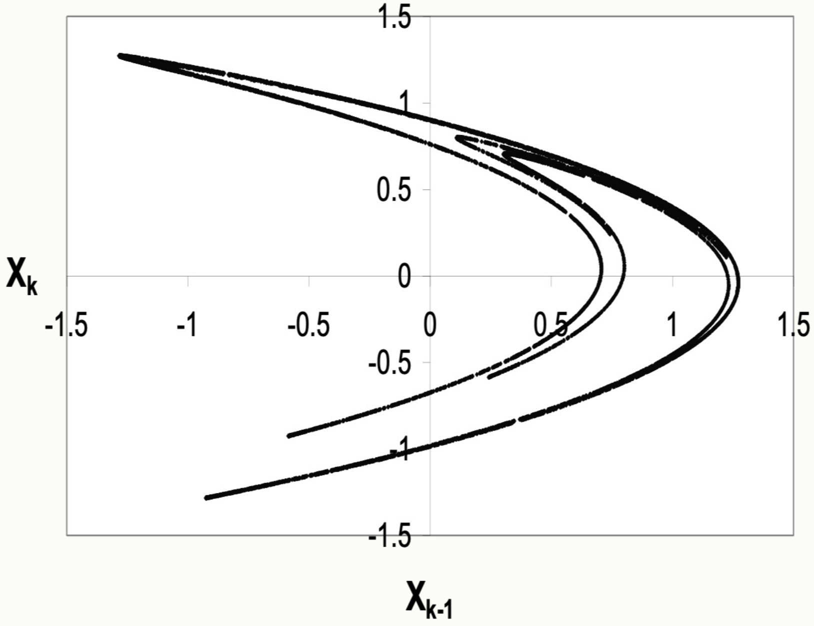
$$x_k = 1 - 1.4x_{k-1}^2 - 0.3x_{k-2}$$

Embedding of 2

Sensitivities



S(1) = 1.8959 S(2) = 0.3



Strange Attractor of Hénon Map

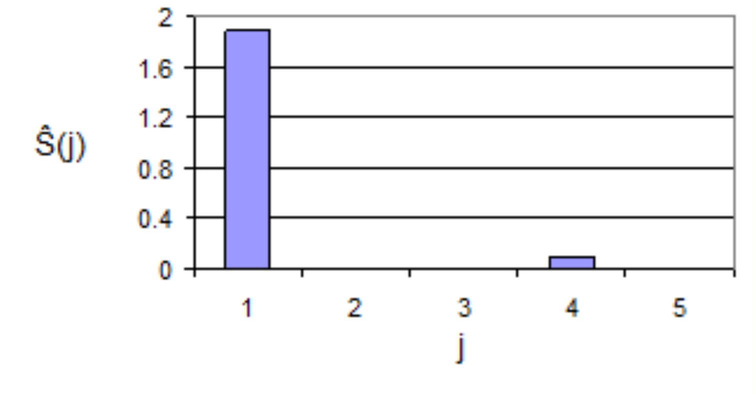
A two-dimensional mapping with a strange attractor (1976)

Delayed Hénon Map

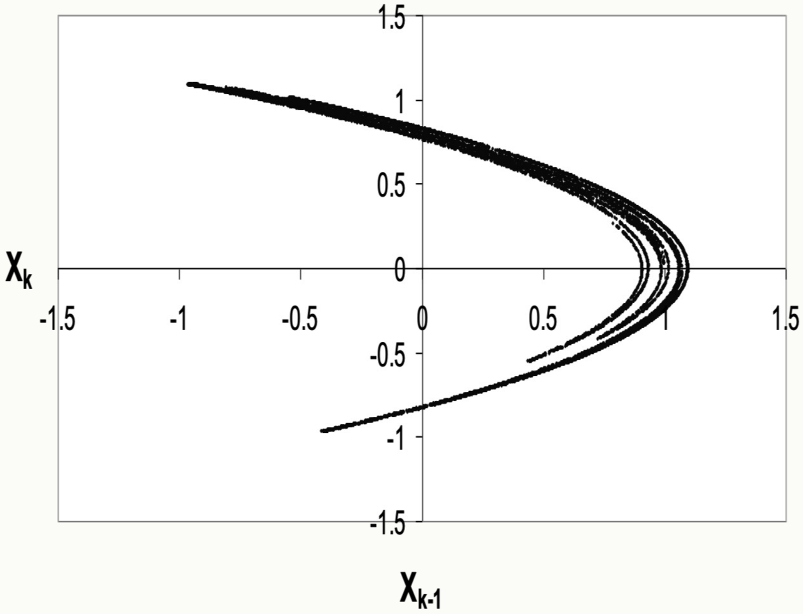
$$x_k = 1 - 1.6x_{k-1}^2 - 0.1x_{k-d}$$

Embedding of *d*

Sensitivities



S(1) = 1.9018 S(4) = .1



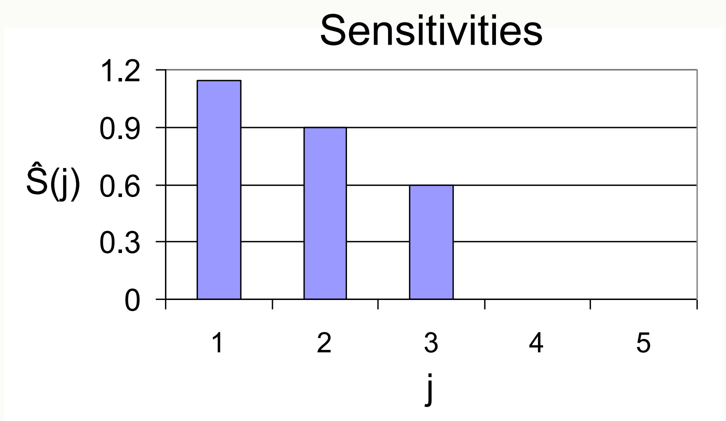
Strange Attractor of Delayed Hénon Map

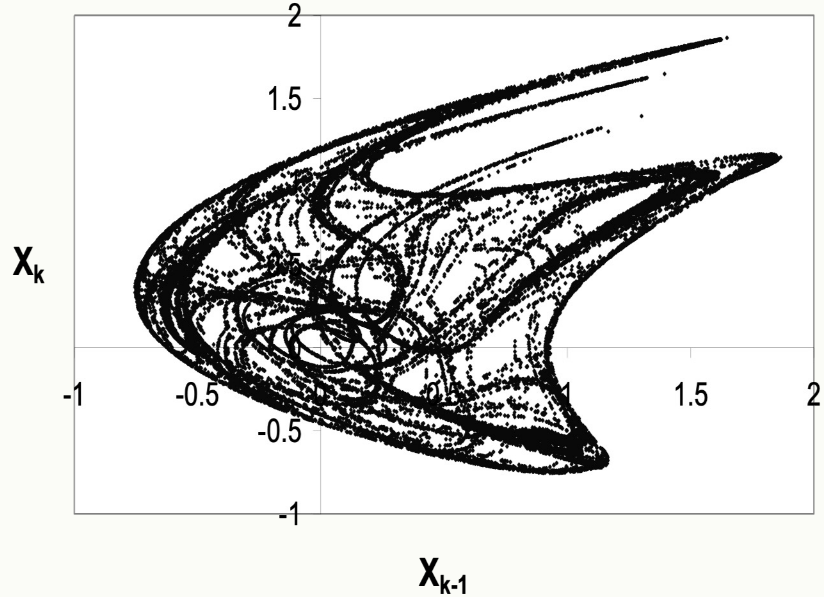
High-dimensional Dynamics in the Delayed Hénon Map (2006)

Preface Map "The Volatile Wife"

$$x_k = x_{k-1}^2 - 0.2x_{k-1} - 0.9x_{k-2} + 0.6x_{k-3}$$

Embedding of 3





Strange Attractor of Preface Map

S(1) = 1.1502 S(2) = 0.9 S(3) = 0.6

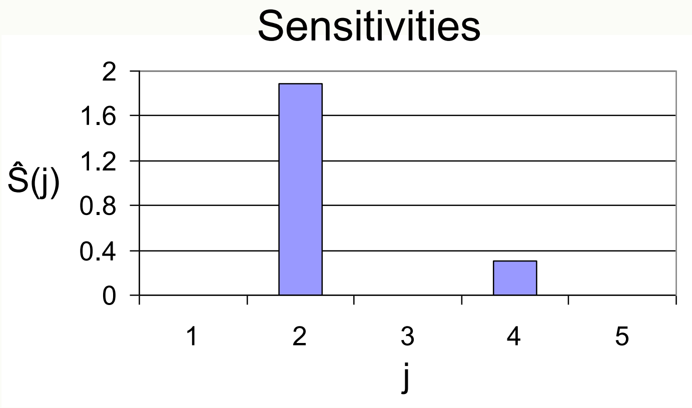
Chaos and Time-Series Analysis (2003)

Images of a Complex World: The Art and Poetry of Chaos (2005)

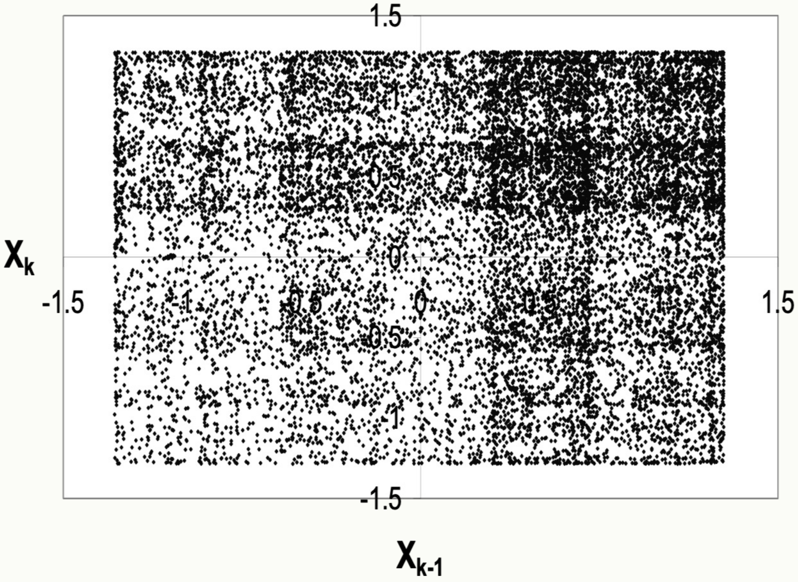
Goutte Map

$$x_k = 1 - 1.4x_{k-2}^2 + 0.3x_{k-4}$$

Embedding of 4



S(2) = 1.8959 S(4) = 0.3

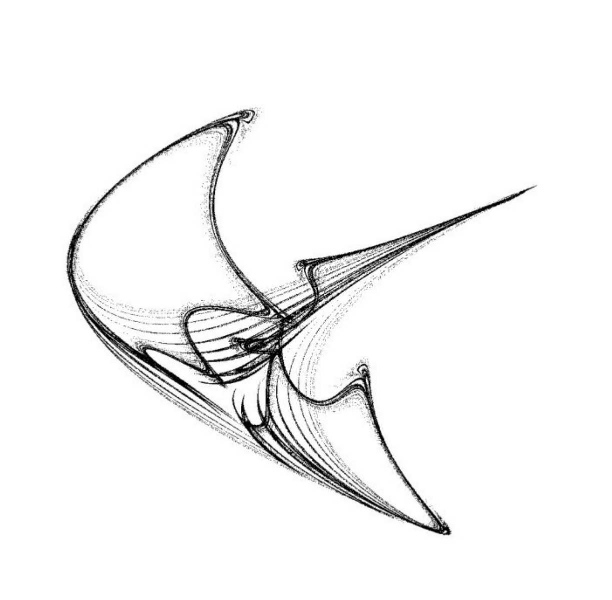


Strange Attractor of Goutte Map

Lag Space Estimation In Time Series Modelling (1997)

Outline

- Discrete Time Series
- Embedding Problem
- Methods Traditionally Used
- Lag Space for a System
- Overview of Artificial Neural Networks
  - Calculation of time lag sensitivities
  - Comparison of model to known systems
- **Comparison to other methods**
- Results and Discussions



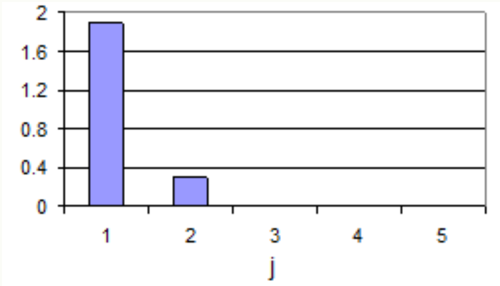
Results from Other Methods

Hénon Map

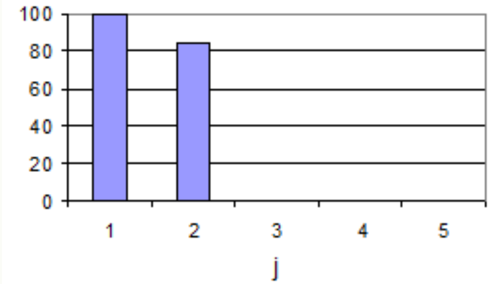
$$x_k = 1 - 1.4x_{k-1}^2 - 0.3x_{k-2}$$

Optimal Embedding of 2

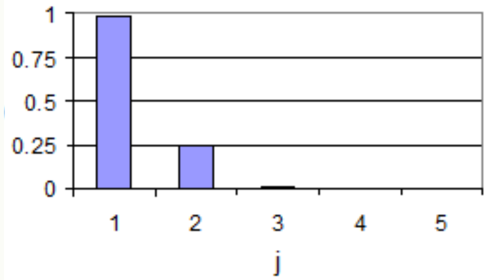
Neural Network



False Nearest Neighbors



Correlation Dimension



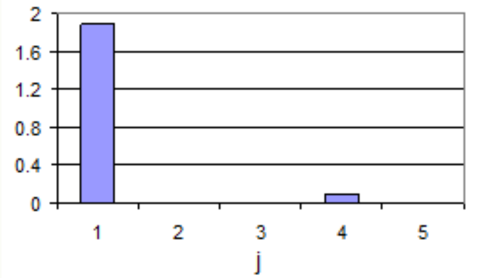
Results from Other Methods

Delayed Hénon Map

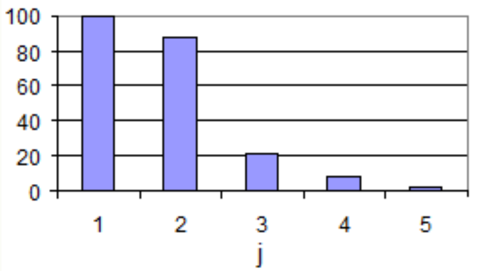
$$x_k = 1 - 1.6x_{k-1}^2 - 0.1x_{k-4}$$

Optimal Embedding of 4

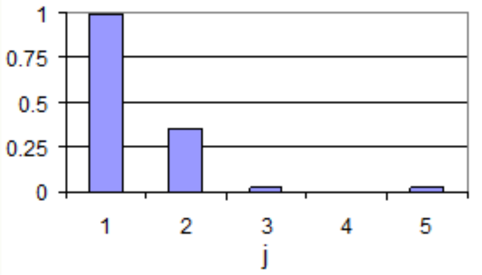
Neural Network



False Nearest Neighbors



Correlation Dimension



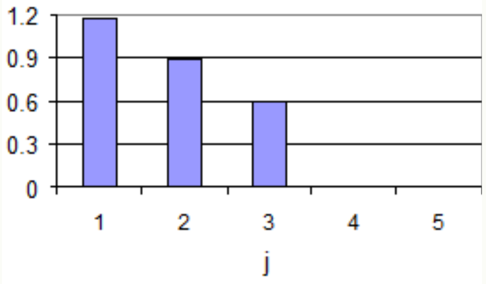
Results from Other Methods

Preface Map

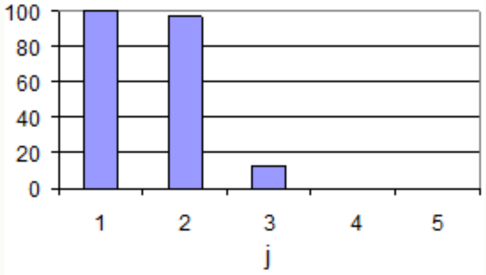
$$x_k = x_{k-1}^2 - 0.2x_{k-1} - 0.9x_{k-2} + 0.6x_{k-3}$$

Optimal Embedding of 3

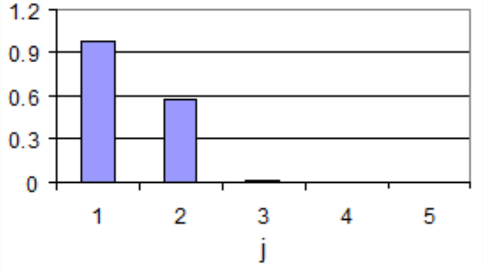
Neural Network



False Nearest Neighbors



Correlation Dimension



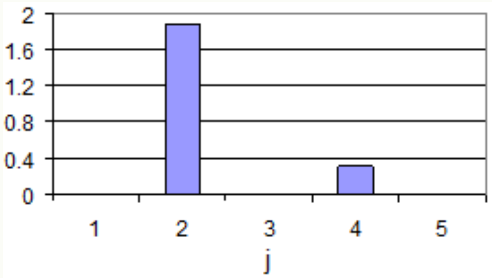
Results from Other Methods

Goutte Map

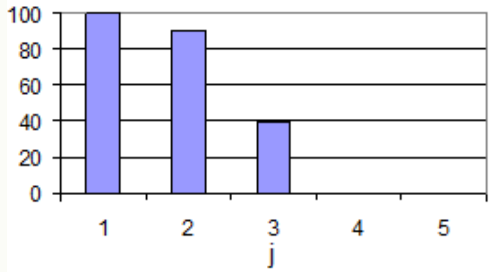
$$x_k = 1 - 1.4x_{k-2}^2 - 0.3x_{k-4}$$

Optimal Embedding of 4

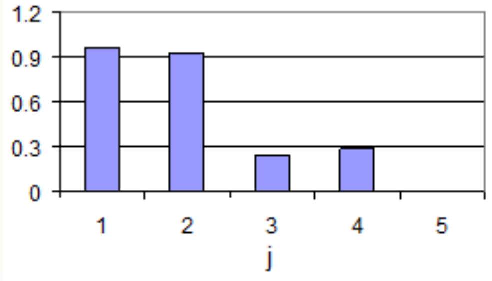
Neural Network



False Nearest Neighbors

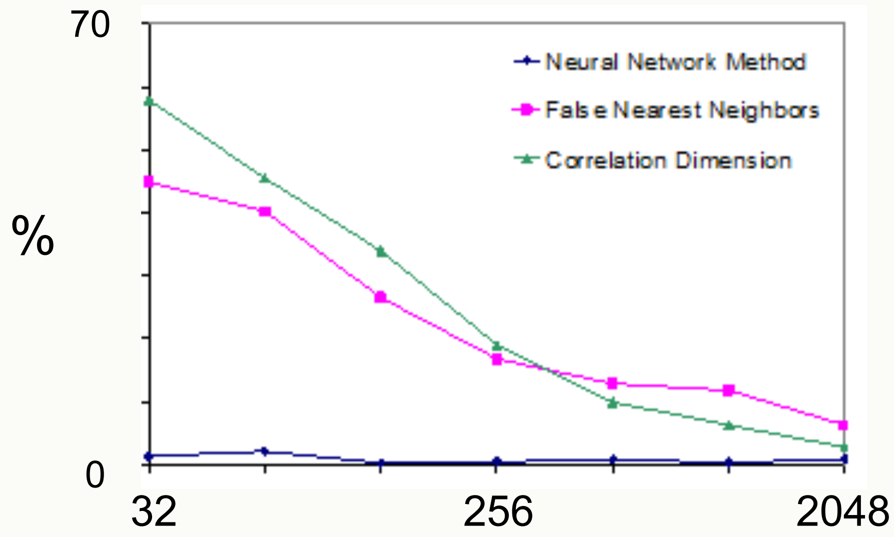


Correlation Dimension



Comparison using Data Set Size

- Varied the length of the Hénon map time series by powers of 2
- Compared methods to actual values using normalized RMS error E



$$E = \sqrt{\frac{\sum_{j=1}^{d}\left(\hat{S}(j) - S(j)\right)^2}{\sum_{j=1}^{d} S^2(j)}}$$

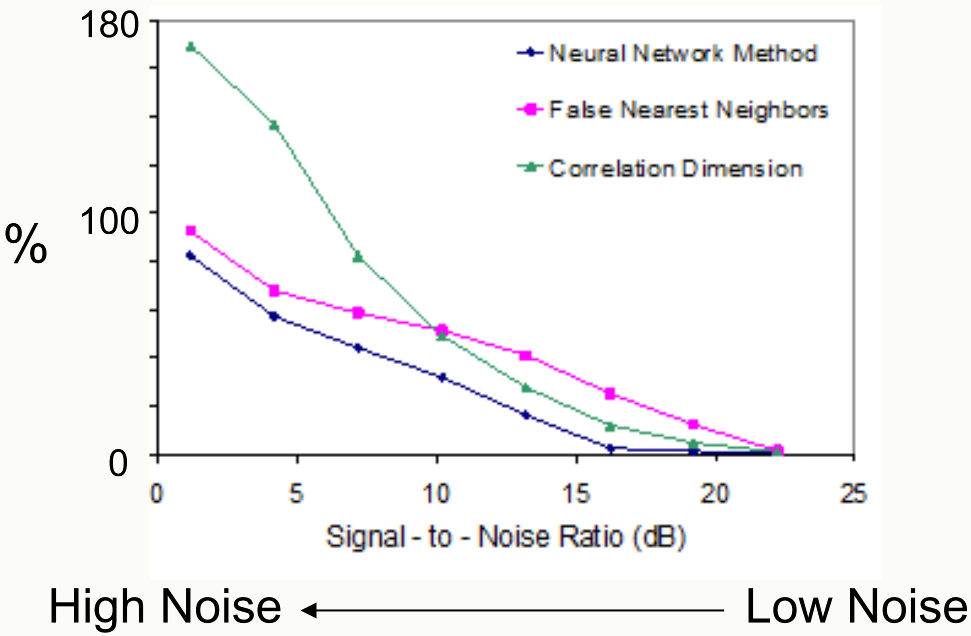
Where $\hat{S}$ is predicted value for a test data set size

S is actual value for an ideal data set size

j is one dimension of d that we are studying

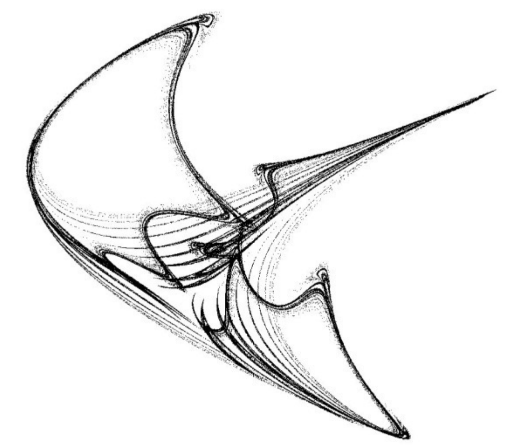
Comparison using Noisy Data Sets

- Vary the noise in the system by adding Gaussian White Noise to a fixed length time series from the Hénon Map
- Compared methods to actual values using normalized RMS error
- Used noiseless case values for comparison of methods

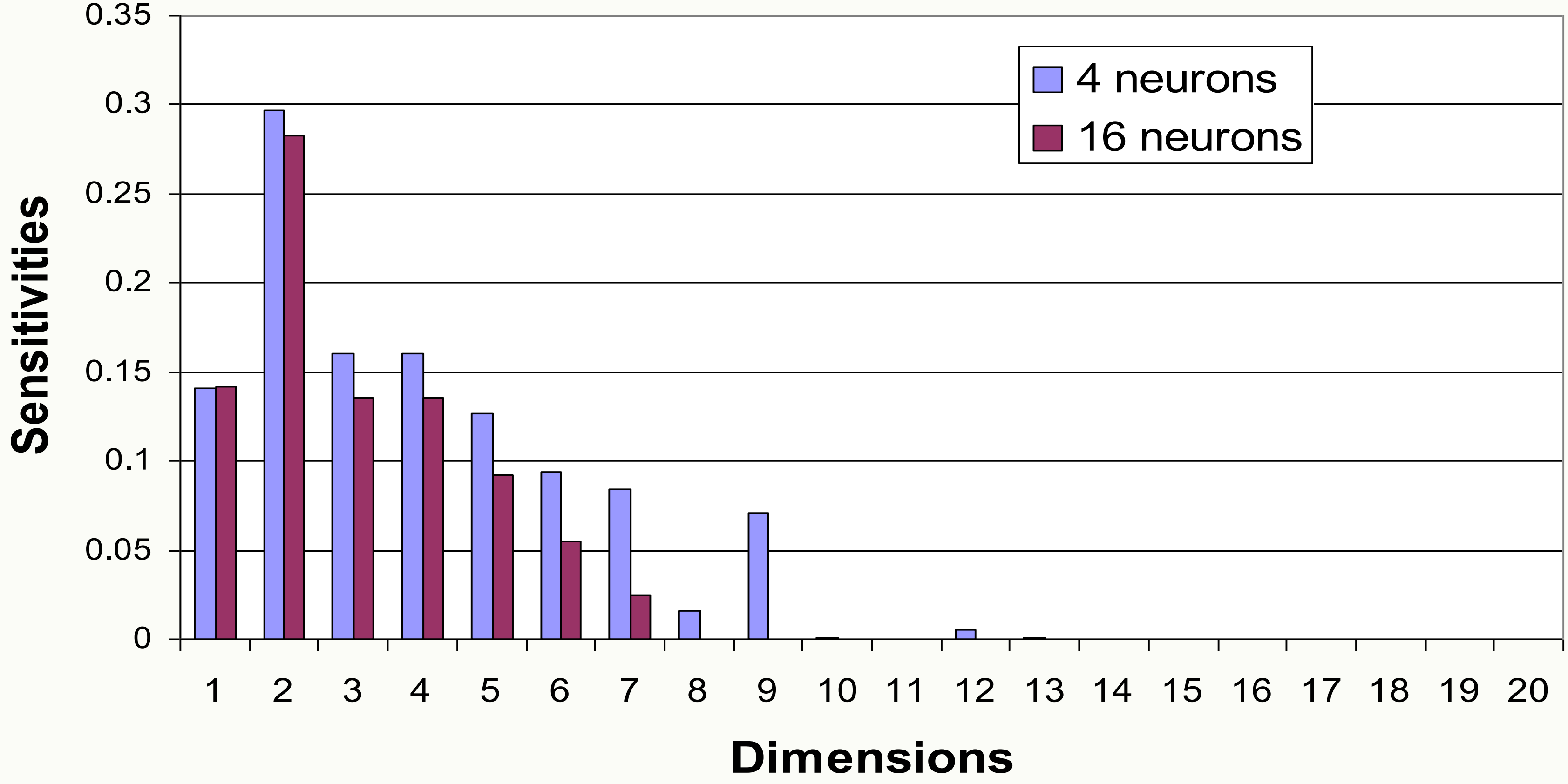


Outline

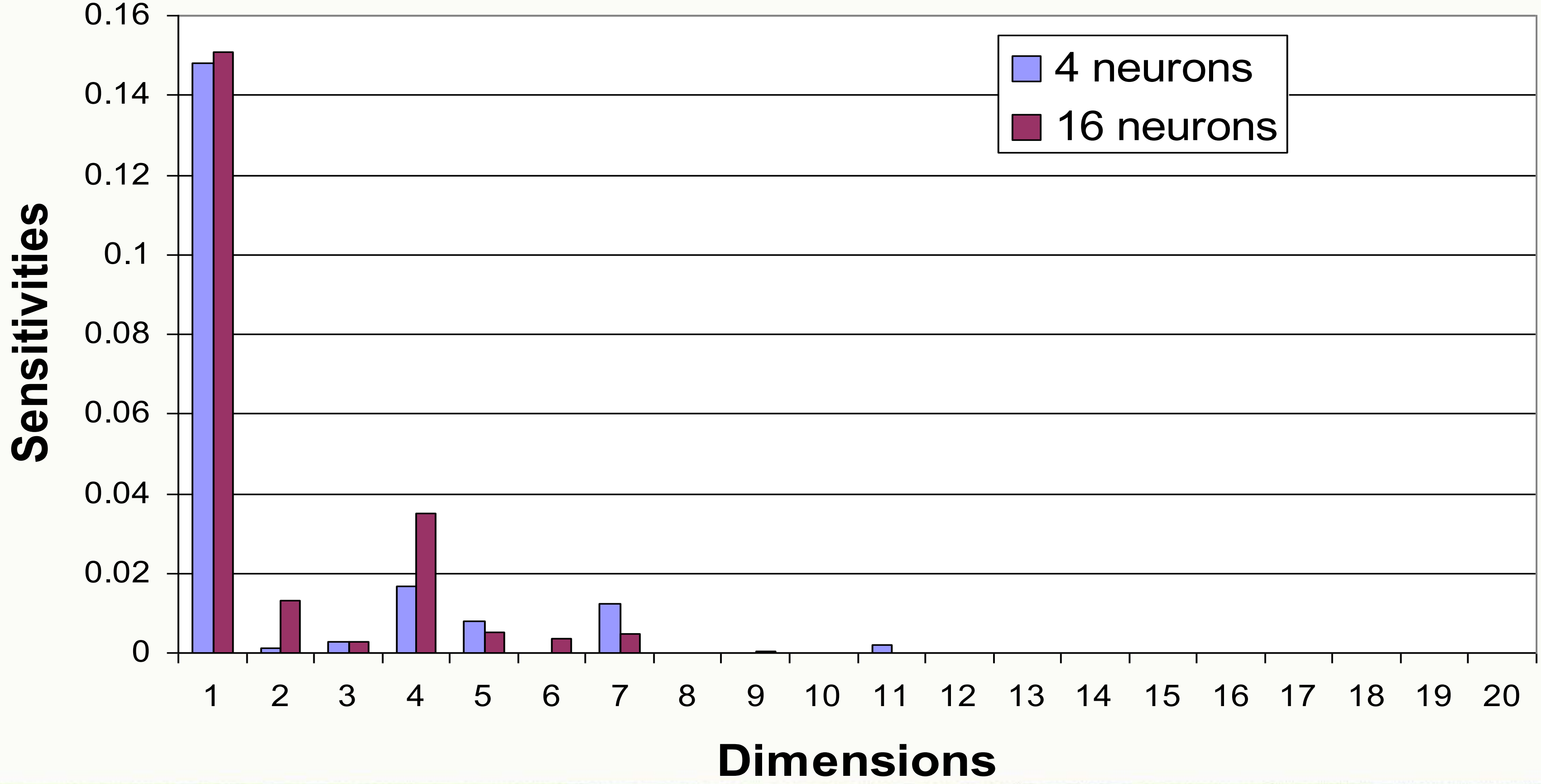
- Discrete Time Series
- Embedding Problem
- Methods Traditionally Used
- Lag Space for a System
- Overview of Artificial Neural Networks
  - Calculation of time lag sensitivities
  - Comparison of model to known systems
- Comparison to other methods
- **Results and Discussions**



Temperature in Madison

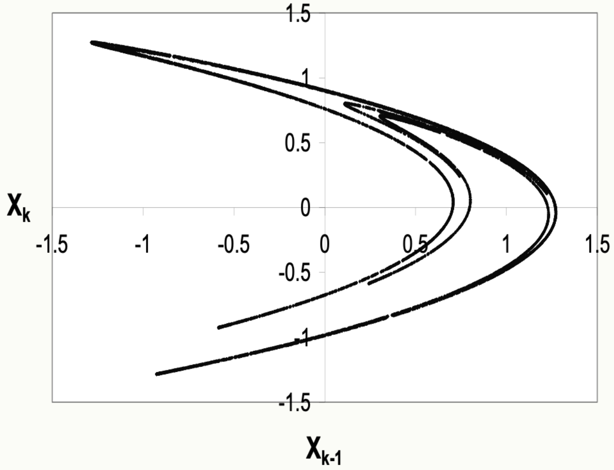


Precipitation in Madison



Summary

- Neural networks are models that can be used to predict the embedding dimension
- They can handle small datasets and accurately predict sensitivities for a given system
- They prove to be more robust to noise than other methods used
- They can be used to determine the lag space where methods cannot



Acknowledgments

- Doctor Sprott for guiding this project
- Doctor Young and Ed Hopkins at the State Climatology Office for the weather data and insightful discussions