Eastern North Pacific Gray Whale Census Estimates

An Application of State Space Reconstruction

J. Robert Buchanan

August 9, 2003

Millersville University of Pennsylvania
email: Bob.Buchanan@millersville.edu

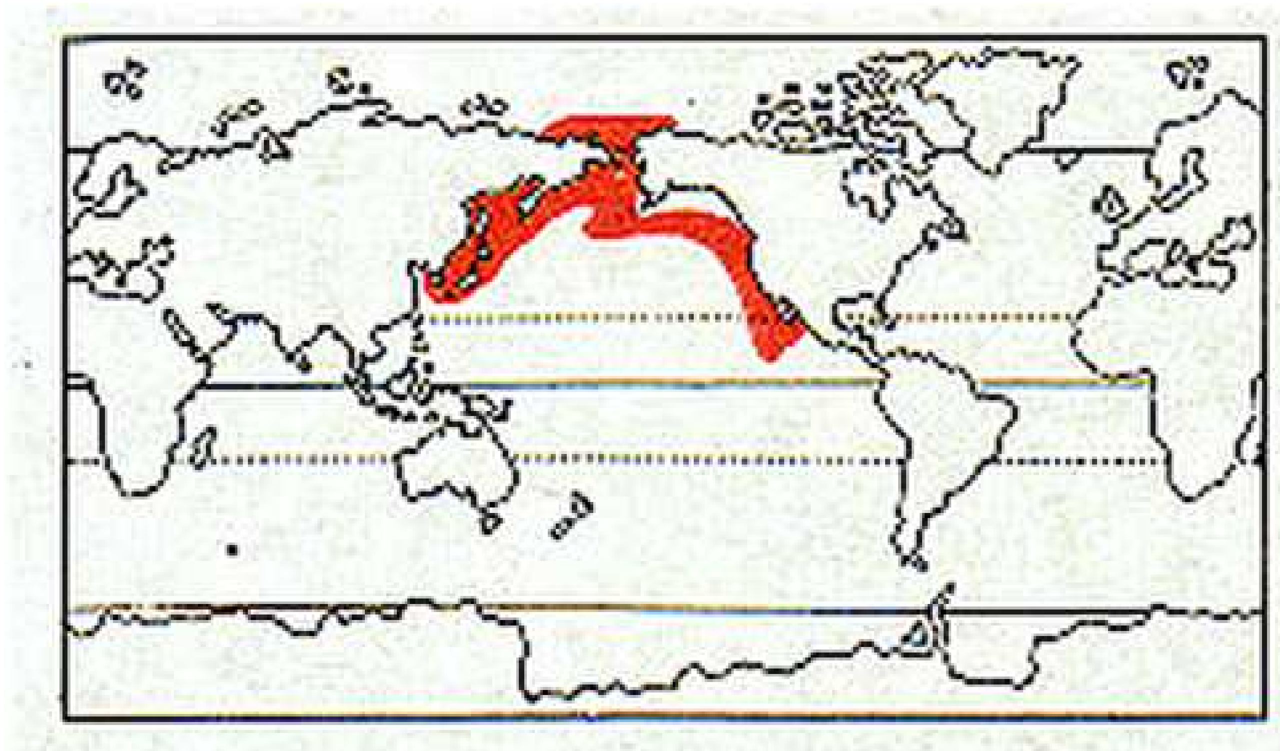
Gray Whale

- Latin name *Eschrictius robustus*, length 40–50 feet (12–15 m), weight 50,000–80,000 lbs (23,000-36,000 kg), lifespan of up to 50 years.

Brief History

- Hunted by aboriginal peoples as well as 19^{th} and 20^{th} century commercial whaling industry
- Believed “commercially extinct” in 1900
- Listed and protected as endangered in 1970 by Endangered Species Act.
- Eastern Pacific stock de-listed as “recovered” in 1994.
- Western Pacific stock nearly extinct (approximately 50 individuals)

Habitat



- Usually found within 2 km of coastline.
- Eastern Pacific females calve in the lagoons of Baja peninsula

Migration

Alaska
Canada
Summer
Pacific Ocean
The black dots show the
North-south migration path
The blue lines show the
summer and winter homes
of the Gray Whale
Winter
Mexico

Feeding Habits

- Only “bottom feeding” whale
- Prefer to feed with their right sides toward the bottom
- Dredge bottom mud for amphipods and crustaceans
- Have baleen for filtering food from mud

Chip Clark

Howard Braham, NMML

Recovery?

Q: Has the stock of eastern Pacific gray whales recovered sufficiently that regulatory protection is no longer needed?

- Yearly census results continue to oscillate
- Mathematical models exhibit a decrease in population during 1968–1988 while census numbers showed an *increase*
- Models do not exhibit depletion of stock by 1900

Q: Can the behavior of mathematical models be reconciled with the census data?

Global Model

From the work of de la Mare (1989) and Punt and Butterworth (1991):

$$\begin{aligned} P_{t+1} &= (P_t - C_t)e^{-M} + \\ &\qquad (1 - e^{-M})P_{t-t_m+1}\left(1 + A\left[1 - \left(\frac{P_{t-t_m+1}}{P_0}\right)^z\right]\right) \end{aligned}$$

C_t catch in year t

M natural mortality rate

t_m age at first parturition

A resilience parameter

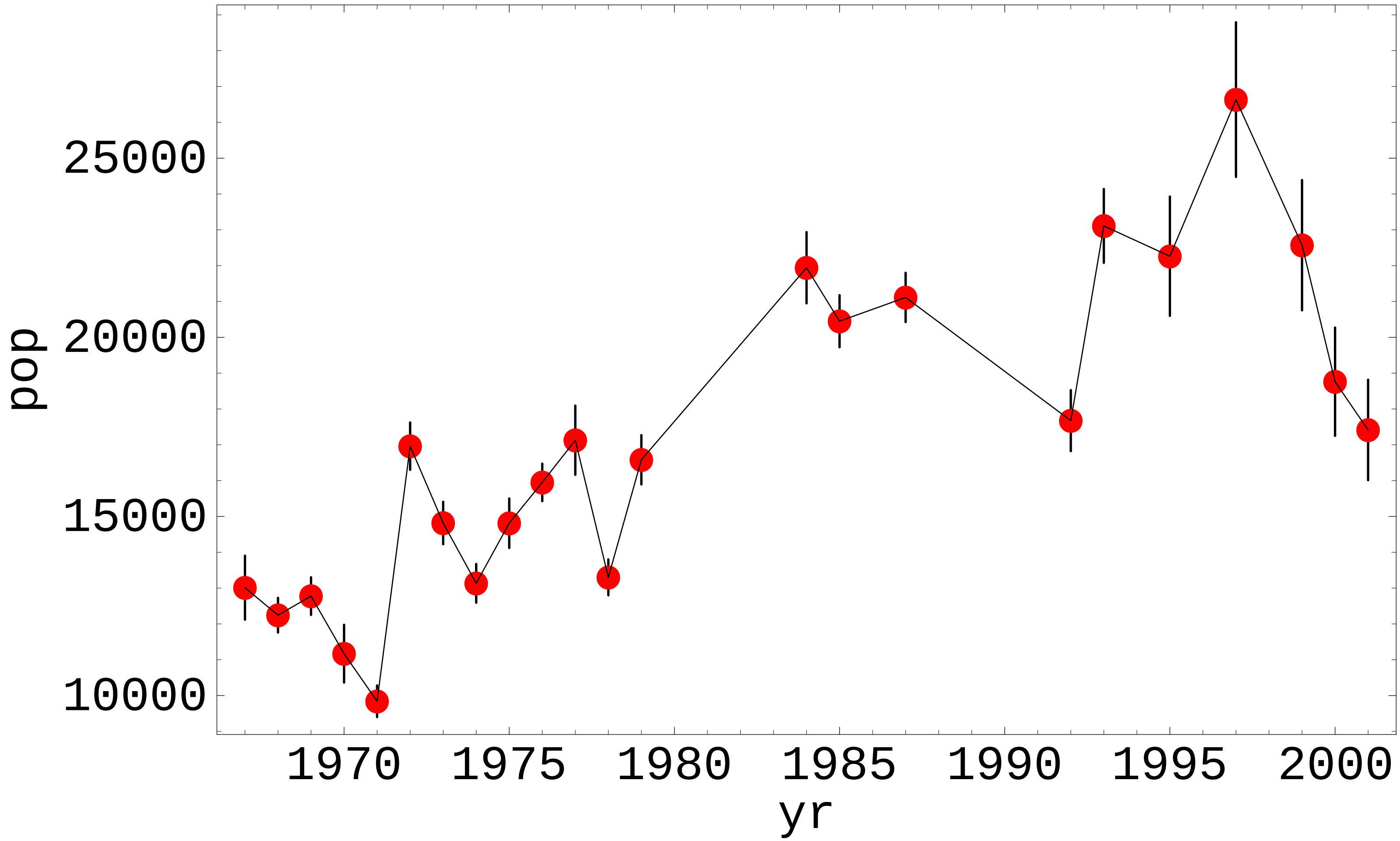
z density-dependent exponent

Local Model

Characteristics of local models:

- Local models received much attention from 1970’s through 1990’s
- Found application in predicting time series
- Data driven – no ecological fidelity
- Computationally expensive
- Often better at predicting behavior of chaotic dynamical systems (*e.g.* solutions of the Lorenz equation) than global models

Time Series



25000
20000
15000
10000
pop
1970
1975
1980
1985
1990
1995
2000
yr

Brief Introduction to Local Modeling

Difference equation model:

$$P_t = F(P_{t-1}, P_{t-2}, \ldots, P_{t-j})$$

Sequence of observations over time:

$$\{P_0, P_1, \ldots, P_N\}$$

Time delay embedding with dimension m, Takens (1981):

$$\begin{aligned} \mathbf{x}_{m-1} &= \langle P_0, P_1, \ldots, P_{m-1} \rangle \\ \mathbf{x}_m &= \langle P_1, P_2, \ldots, P_m \rangle \\ &\vdots \\ \mathbf{x}_N &= \langle P_{N-m+1}, P_{N-m+2}, \ldots, P_N \rangle \end{aligned}$$

Takens's Theorem

From Casdagli, *et al.* (1991):

Dynamical system: $x(t) = f^t(x(0))$, $x \in \mathbb{R}^n$

Observable: $y(t) = g(x(t))$, $y \in \mathbb{R}^d$

Delay construction map:

$$\Phi(x(t)) = \langle g(f^{-\tau m_p}(x(t))), \ldots, g(x(t)), \ldots, g(f^{\tau m_f}(x(t)))\rangle$$

If $m = m_f + m_p + 1 \geq 2n + 1$ then Φ is a smooth, one-to-one coordinate transformation with a smooth inverse.

Local Modeling 2

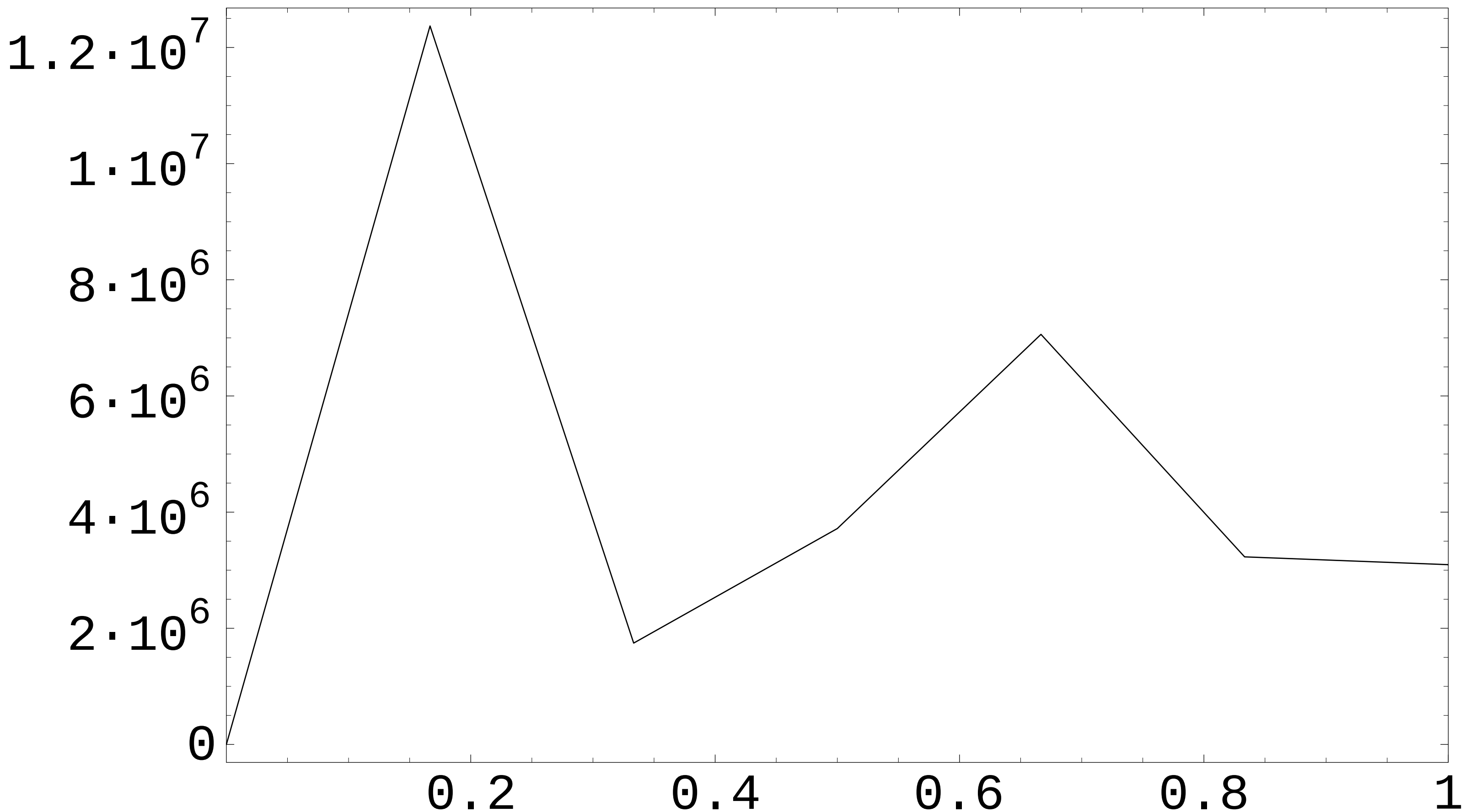
Prediction:

$$\hat{P}_{N+1} = G(\mathbf{x}_N)$$

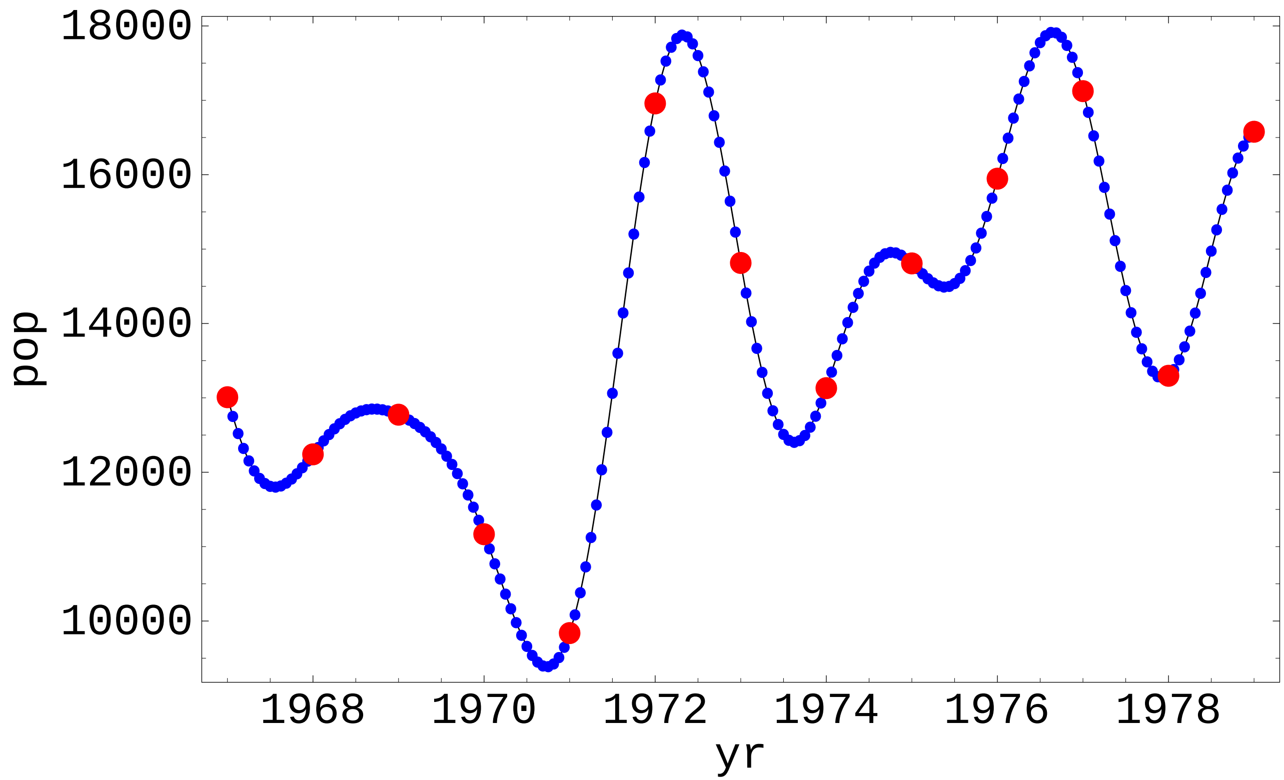
Multi-step prediction:

$$\begin{aligned}
\hat{P}_{N+2} &= G(\langle P_{N-m+2}, P_{N-m+3}, \ldots, \hat{P}_{N+1} \rangle) \\
\hat{P}_{N+3} &= G(\langle P_{N-m+3}, P_{N-m+4}, \ldots, \hat{P}_{N+2} \rangle) \\
&\vdots
\end{aligned}$$

Fourier Spectrum



Interpolation



Up-sample to fill in the spacing of the sequence of observations.

Filtering

- Remove high frequency noise
- Reduce data storage requirements
- Perform in parallel with interpolation

$$\mathbf{x}_i = L_3 \circ L_2 \circ L_1(\langle P_{i-w+1}, P_{i-w+2}, \ldots, P_i \rangle)$$

L_1 Fourier transform

L_2 Low pass filter $m/2$ frequencies

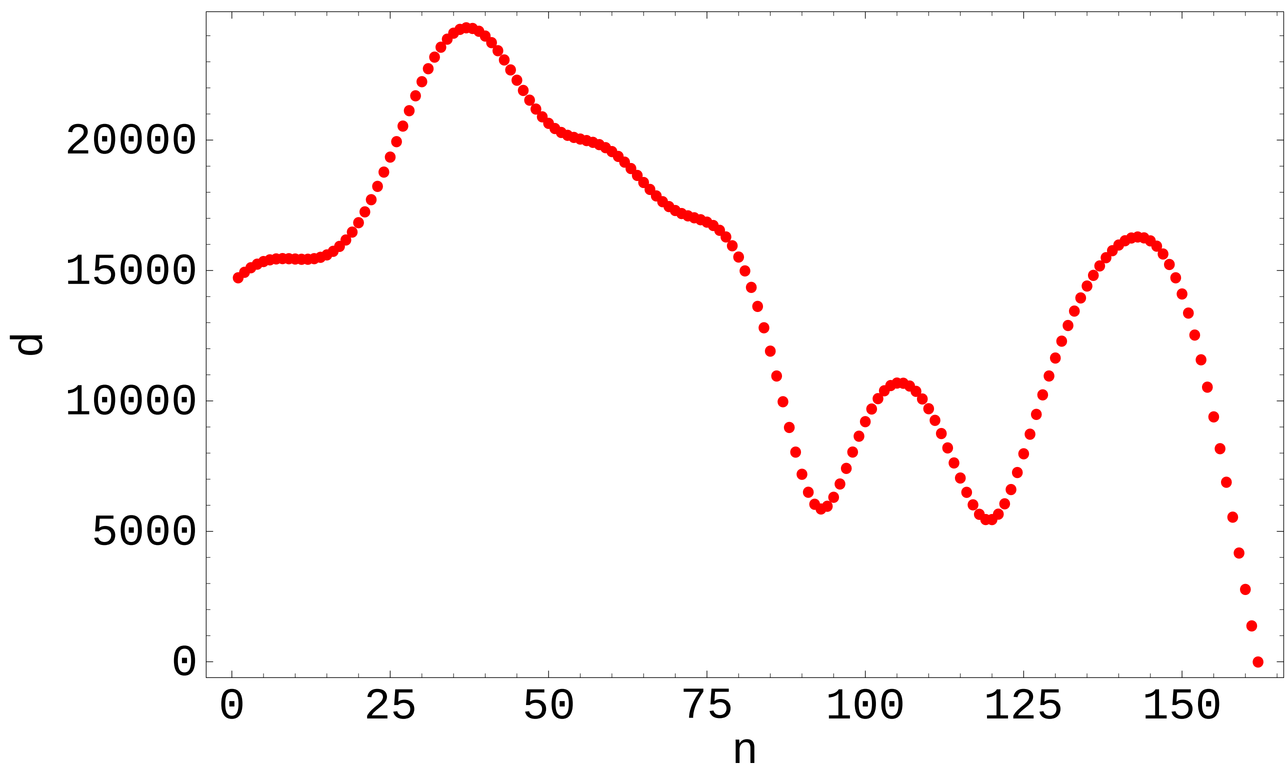
L_3 Inverse Fourier transform

Nearest Neighbors

Metric:

$$d^2(\mathbf{x}_a, \mathbf{x}_b) = \sum_{i=0}^{m-1} \lambda^i (\mathbf{x}_{a,m-i} - \mathbf{x}_{b,m-i})^2$$

with $0 < \lambda \leq 1$.



Prediction Algorithm

1. Create filtered embedding of sequence of observations
2. Find $k \geq 1$ nearest neighbors of $\mathbf{x}_N$
3. For neighbors $\{\mathbf{x}_{n_1}, \ldots, \mathbf{x}_{n_k}\}$ find $\{P_{n_1+1}, \ldots, P_{n_k+1}\}$
4. Approximate the map $\mathbf{x}_\alpha \overset{G}{\longmapsto} P_{\alpha+1}$
5. Evaluate $G(\mathbf{x}_N) = \hat{P}_{N+1}$

Types of Maps: Averaging

Direct:

$$\hat{P}_{N+1} = \frac{\sum_{i=1}^{k} w_i P_{n_i+1}}{\sum_{i=1}^{k} w_i}$$

Integrated:

$$\hat{P}_{N+1} = P_N + \frac{\sum_{i=1}^{k} w_i (P_{n_i+1} - P_{n_i})}{\sum_{i=1}^{k} w_i}$$

Weights depend on the distance between the neighboring vector and the query vector.

$$w_i = \left[1 - \left(\frac{d(\mathbf{x}_N, \mathbf{x}_{n_i})}{d(\mathbf{x}_N, \mathbf{x}_{n_k})} \right)^2 \right]^2$$

Another Type: Linear

From Sauer (1994):

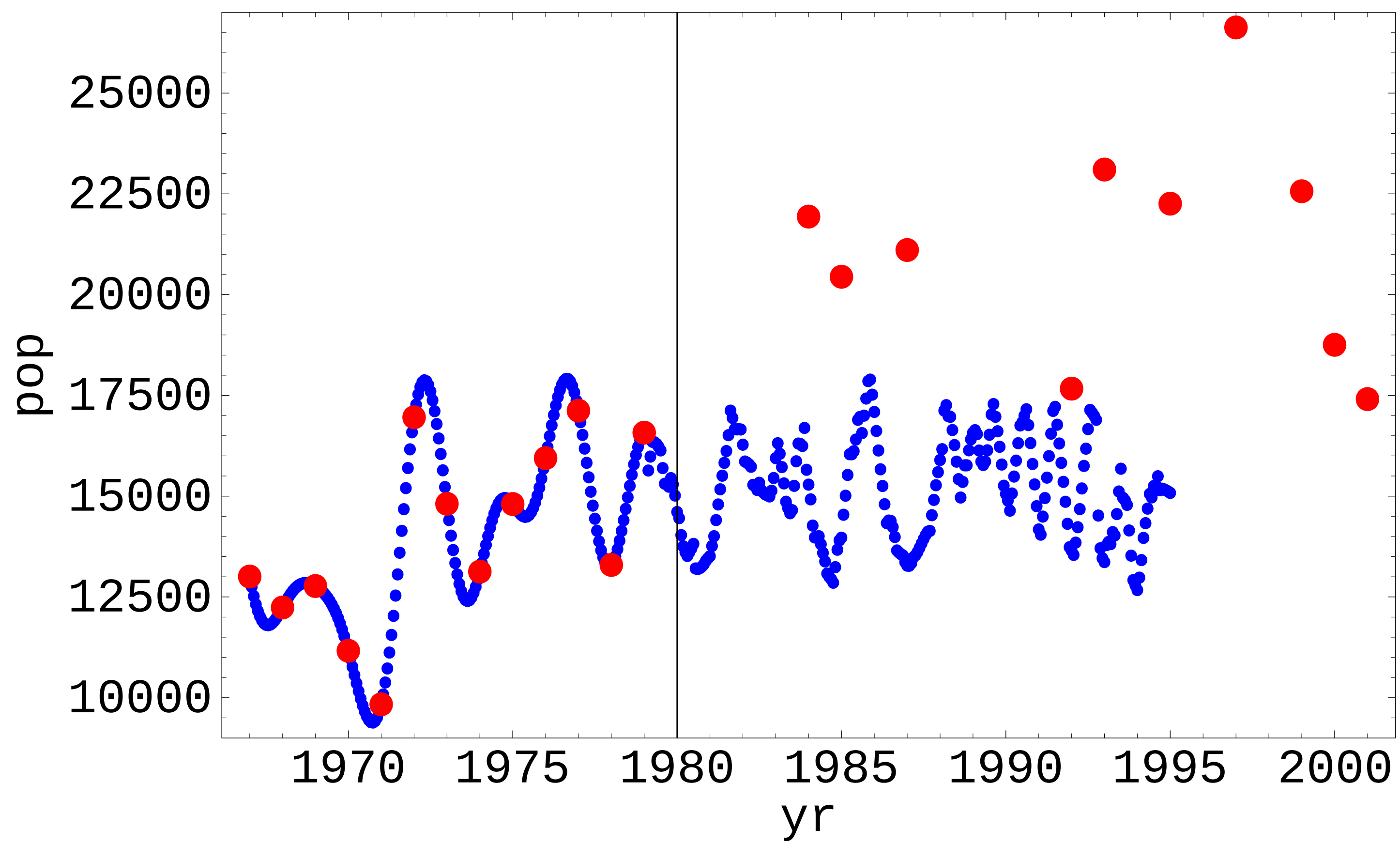
1. Let $\mathbf{c}$ be the center of mass of $\{\mathbf{x}_{n_1}, \ldots \mathbf{x}_{n_k}\}$
2. For some $l \leq m$ find the l-dimensional subspace of $\mathbb{R}^m$ containing $\mathbf{c}$ closest to the span of $\{\mathbf{x}_{n_1}, \ldots \mathbf{x}_{n_k}\}$

$$A = \begin{bmatrix} \mathbf{x}_{n_1} - \mathbf{c} \\ \vdots \\ \mathbf{x}_{n_k} - \mathbf{c} \end{bmatrix} = U^t D V$$

3. Project $\{\mathbf{x}_{n_1} - \mathbf{c}, \ldots, \mathbf{x}_{n_k} - \mathbf{c}\}$ onto $\mathbb{R}^l$
4. Find the affine map $G : \mathbb{R}^l \to \mathbb{R}$ which best fits the data $\{(\Pi(\mathbf{x}_{n_1} - \mathbf{c}), P_{n_1+1}), \ldots, (\Pi(\mathbf{x}_{n_k} - \mathbf{c}), P_{n_k+1})\}$

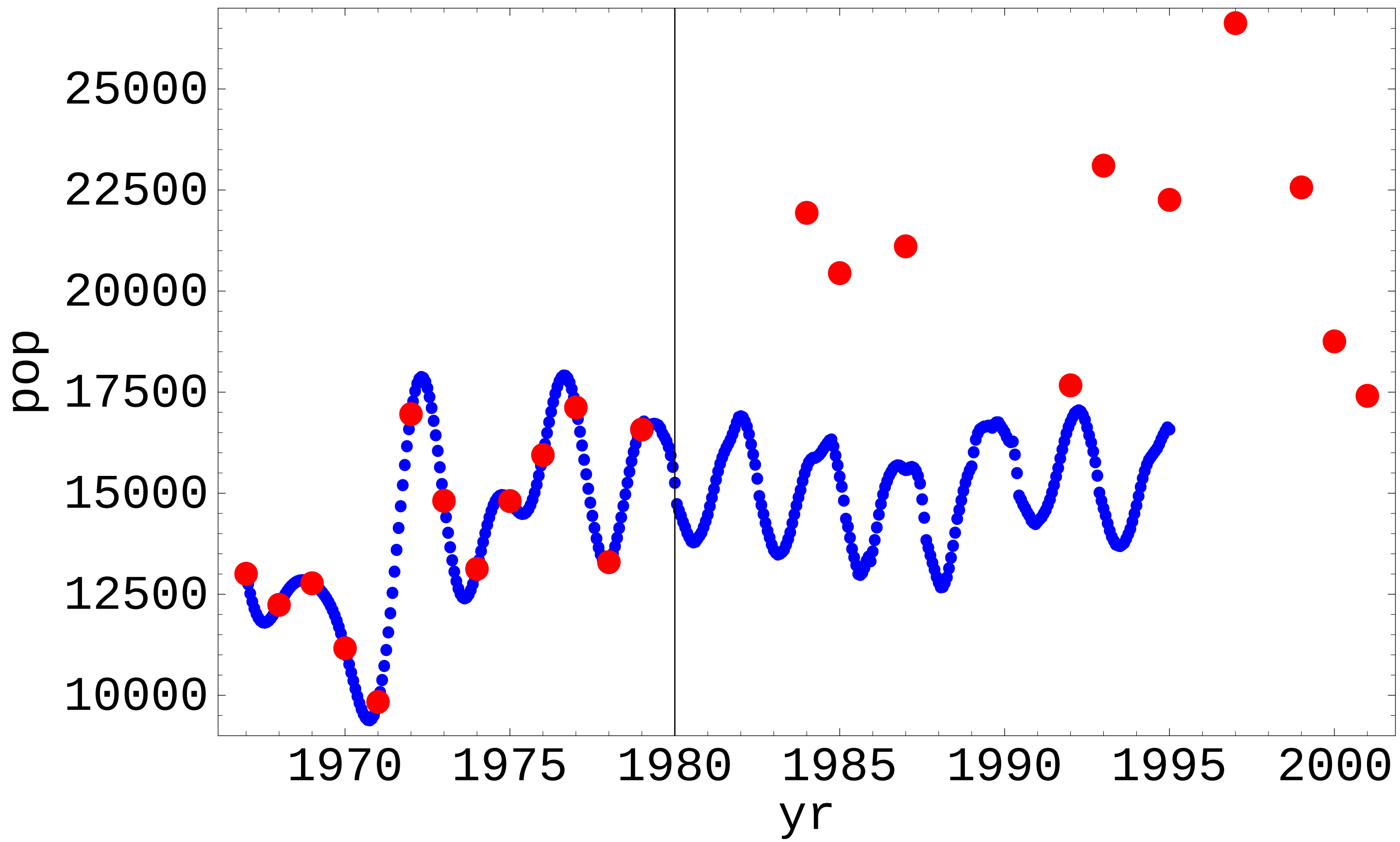
$$\hat{P}_{N+1} \approx G(\Pi(\mathbf{x}_N - \mathbf{c}))$$

Constant Model, Direct Averaging



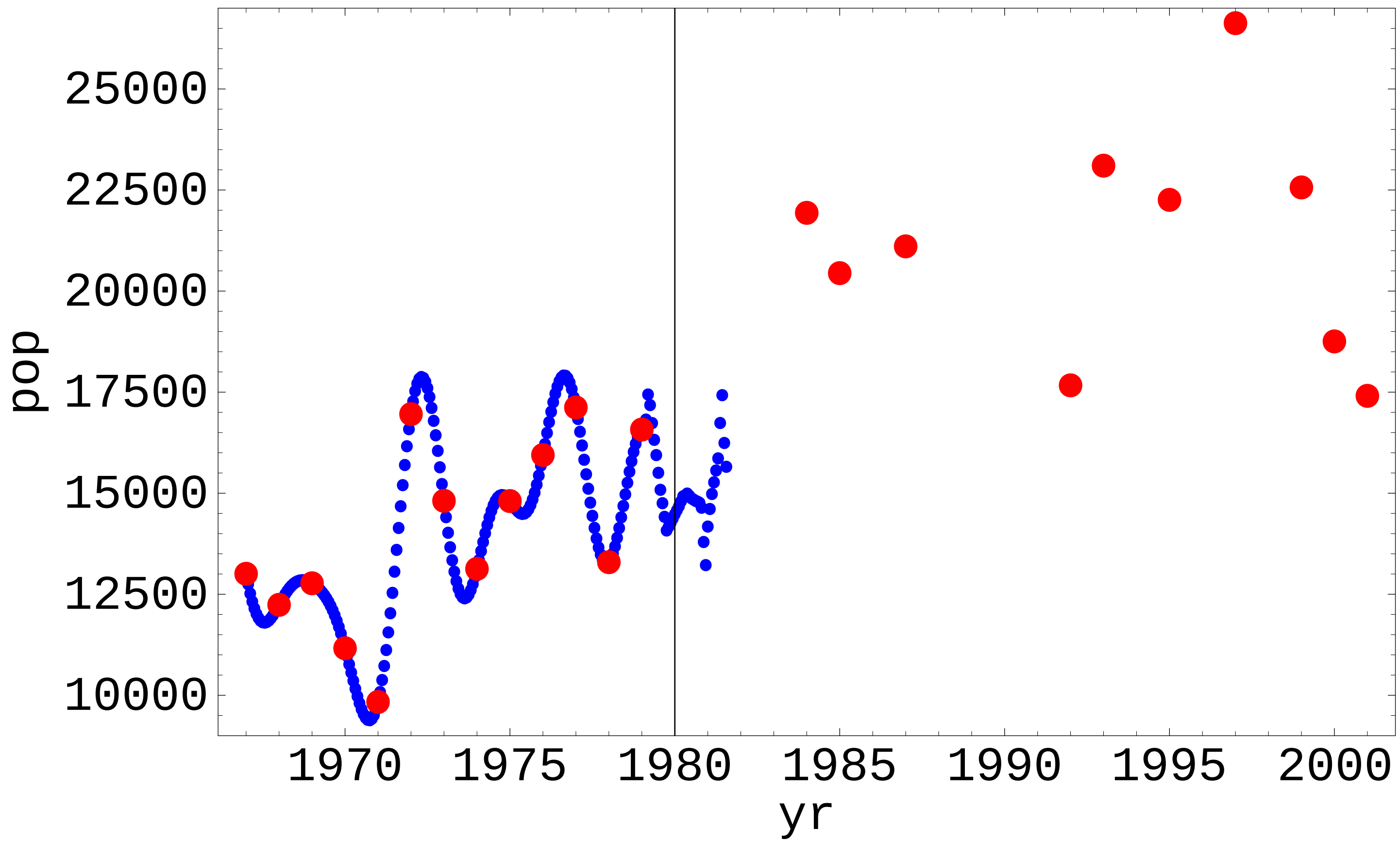
Parameters: Euclidean metric, 2 nearest neighbors, embedding dimension 32

Constant Model, Integrated Averaging



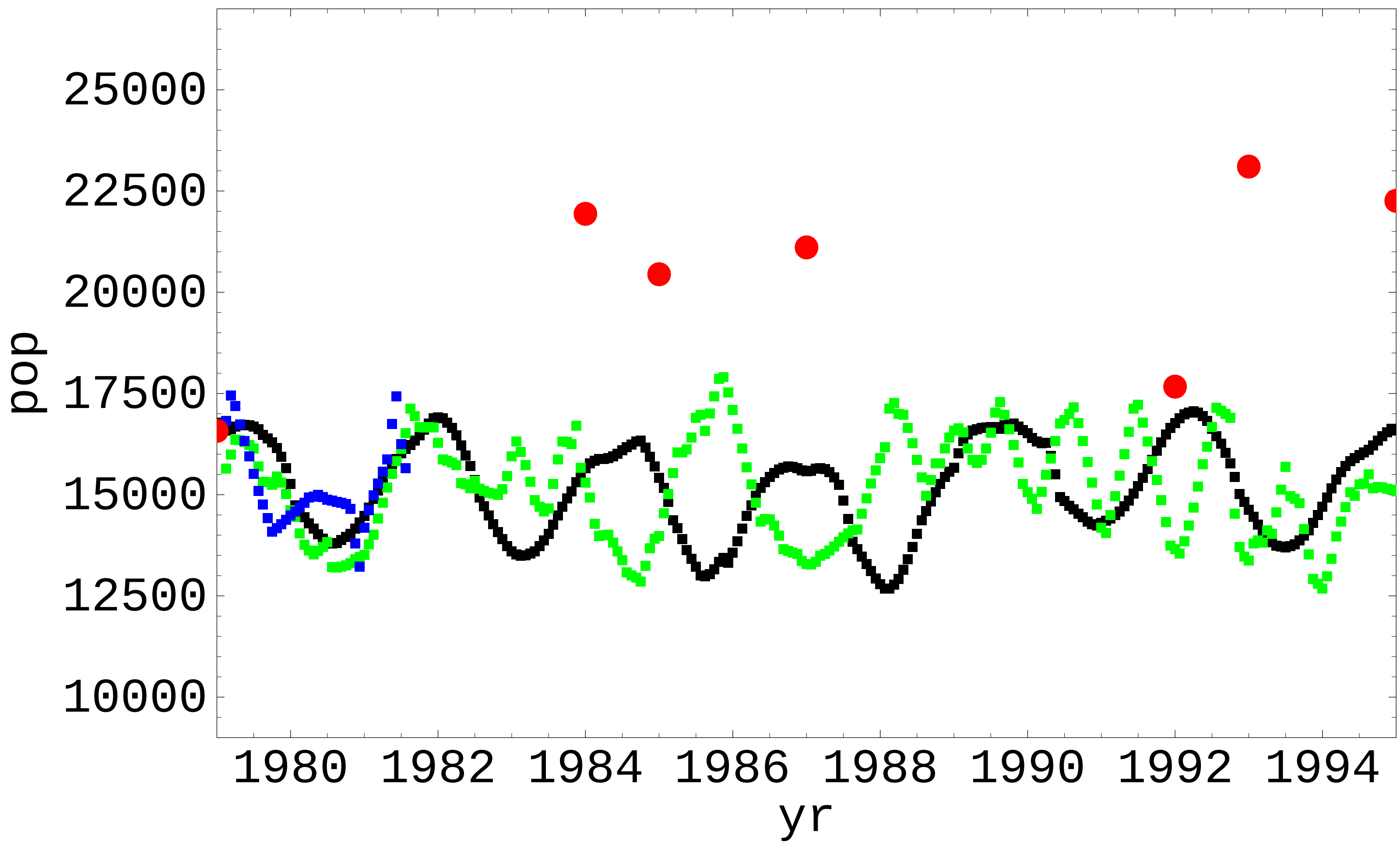
Parameters: Euclidean metric, 2 nearest neighbors, embedding dimension 32

Linear Model



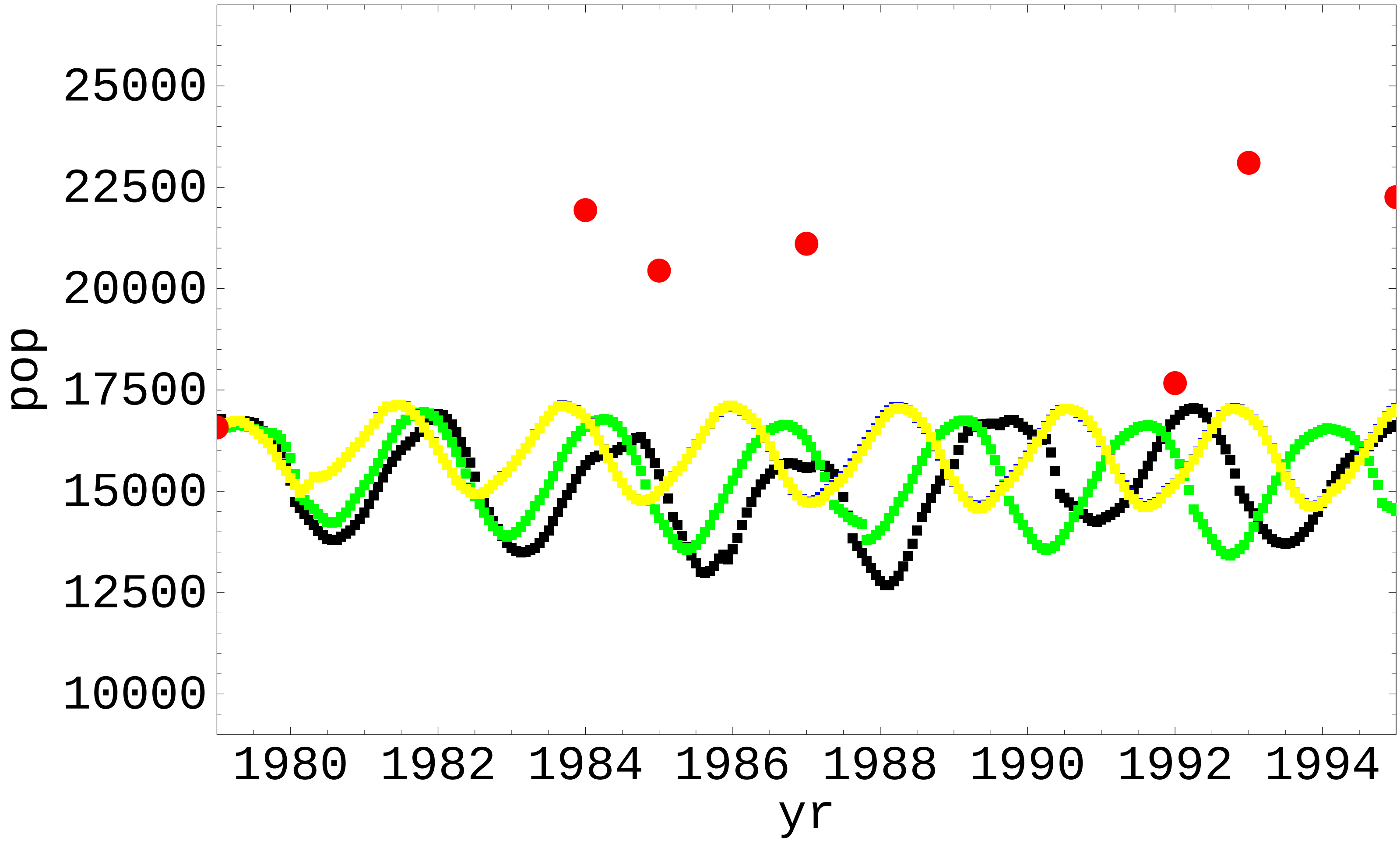
Parameters: Euclidean metric, 4 nearest neighbors, embedding dimension 32, model dimension 2

Overlay of Model Results



Model: integrated, averaged, linear

Sensitivity to Nearest Neighbors



Number of nearest neighbors: 2, 3, 4, 5

Comments

- Linear model requires more neighbors than the constant models.
- Linear model requires more data than the constant models.
- Constant models exhibit plausible oscillatory behavior.
- Population levels predicted by the local models exhibit flat average behavior
- Need much more data for model validation.
- Has the gray whale population been above its carrying capacity?

References

Casdagli, M. *et al.*, “State space reconstruction in the presence of noise,” *Physica D* **(51)**, pp. 52–98 (1991).

de la Mare, W.K., Report of the Scientifi c Committee, Annex L. The model used in the HITTER and FITTER programs (Program: FITTER.SC40). *Rep. int. Whal. Commn.* **(39)**, pp. 150–151 (1989).

Punt, A.E. and Butterworth, D.S., HITTER-FITTER-Bootstrap User’s Guide Version 2.0, unpublished (1991).

Sauer, T., “Time Series Prediction by Using Delay Coordinate Embedding,” in *Time Series Prediction: Forecasting the Future and Understanding the Past*, A.S. Weigand and N.A. Gershenfeld eds., Addison-Wesley Pub., Reading, MA (1994).

Takens, F., “Detecting strange attractors in fluid turbulence,” in *Dynamical Systems and Turbulence*, D. Rand and L.-S. Young eds., Springer, Berlin, (1981).