2009 International Symposium on Nonlinear Theory and its Applications
NOLTA'09, Sapporo, Japan, October 18-21, 2009

Re-examination of evidence for low-dimensional chaos in the Canadian lynx data

Michael Small and Cristian Carmeli

†Electronic and Information Engineering, The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
Email: small@ieee.org, cristian@eie.polyu.edu.hk

Abstract—The time series of the annual number of Canadian lynx caught by the Hudson Bay company between 1821 and 1935 exhibits pseudo-cyclic behaviour and has long been considered as an archetypal example of irregularly fluctuating population dynamics. Recently proposed global polynomial models of this data have been found to exhibt chaotic dynamics and were therefore presented as direct evidence of chaos in a real ecosystem. In this paper we re-examine that evidence by constructing global radial basis models subject to information theoretic parameter constraints. We find that the models exhibit very good agreement with the data and are able to accurately reproduce the qualitative long term dynamical behaviour. The models also often exhibit "almost" chaotic dynamics, either: (a) very long period periodicity, (b) a periodic orbit embedded in a dissipative mixing region, or (c) very long time transient irregular aperiodic dynamics with an asymptotically periodic orbit. In each case the dynamics exhibit a very rich range of behaviour and can also provide a qualitatively accurate deterministic model of the apparently chaotic dynamics when subjected to a delay reconstruction. We conclude that, while the data and these models are consistent with the hypothesis of chaos in a real ecosystem, the data may also be adequately explained by periodic "almost chaotic" behaviour.

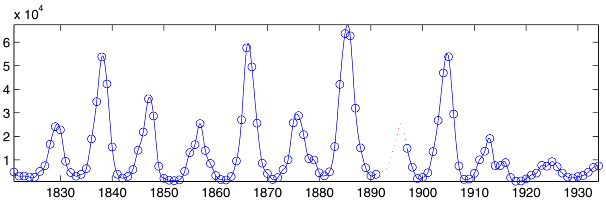

Figure 1: Annual number of Canadian lynx pelts reported by the Hudson Bay company (1821-1935). Data for the years 1892 to 1896 is missing and filled in following an ad hoc procedure (shown in red dotted line). The data set has been interpolated by a factor of 10 (increasing the amount of available data while preserving the power spectra). The original values are shown as open circles.

1. Introduction

Simple three species ecosystem models (and even two-dimensional predator-prey maps) have long been known to potentially exhibit chaotic dynamics [2]. However, direct evidence of chaos in real ecosystems is somewhat limited, and is typically only found in rather restricted settings [1]. Nonetheless, Maquet, Letellier and Aguirre [5] recently analysed a data set of annual reported trappings of Canadian lynx and found that the data were consistent with low dimensional chaos. They presented a three level food chain model and global data-based polynomial models, and found that each was capable of producing chaotic dynamics — from this they infer that the data are direct evidence for chaos in a real ecosystem. In this paper, we reanalyse the same data with our own modelling algorithms and obtain slightly different results.

The models we obtain are built by fitting a radial basis function network to the available data subject to a minimum description length criterion [6] to determine model size [4, 8, 10]. While we do obtain chaotic models capable of reproducing a qualitatively identical attractor to that derived from the data, we also find other richer and potentially more intriguing dynamical behaviours. In particular, the models we obtain are very close to the boundary between periodic dynamics and chaotic pseudo-periodicity. We obtain periodic models, which when driven by very small amplitude noise behave in a stable and aperiodic manner. We obtain periodic models which exhibit transient times significantly longer that the length of the observed data. We also obtain periodic orbits with a very long period (again, of the order of the length of the data) and deterministic chaotic dynamics capable of reproducing the original dynamics.

2. Chaotic Canadian lynx

This section is organised as follows: Sec. 2.1 introduces the data and Sec. 2.2 briefly reprises the modelling methodology we employ. In Sec. 2.3 we summarise our results and in the following section we briefly conclude.

2.1. The data

When studying the dynamic behaviour of interacting populations it is often extremely difficult to get meaningful data from field experiments: typically data is needed over time scales greatly exceeding the average lifetime of the organism. One classic and widely studied example is the so-called Nicholson's blowflies data set [2] where the population of flies on a group of Australian sheep was recorded over a period of weeks. For larger animals, with longer lifespans the problem of obtaining useful data becomes more difficult. Between 1821 and 1935 the number of Canadian lynx furs harvested by the Hudson Bay company is one such widely studied data set. The data (see Fig. 1) exhibits approximately periodic oscillations with irregular fluctuations of the amplitude and a regular period of about 9.6 years.

The original data set consists of 108 data points (the data for the years from 1892 to 1896 is absent) covering 112 years. For the sake of comparison we repeat the data preprocessing procedure detailed by Maquet and colleagues [5]. The data for the missing period is reconstituted using the closest available match from the remaining available data and the resultant time series is interpolated to yield a total of 1131 time series data. The data we use is identical to that in [5], in the next section we describe our modelling procedure.

2.2. The model

The scalar time series data x_t is subjected to a variable time delay embedding to attempt to capture the various relevant time scales in the system

$$z_t = (x_t, x_{t-5}, x_{t-10}, x_{t-15}, x_{t-20}, x_{t-25}, x_{t-90}) \quad (1)$$

where the pseudo-period of the system is around 96. The choice of embedding strategy is both ad hoc and arbitrary. Various alternative strategies produced similar results provided that the embedding lags span the same range of values. From the vector time series we employ radial basis models to approximate the function f such that $\|z_{t+1} - f(x_t)\|$ is minimised subject to a fixed model size d

$$f(x) = \sum_{i=1}^{d} \lambda_i \phi_i \left(\frac{\|x - c_i\|}{r_i} \right) \quad (2)$$

where the weights λ_i are determined by least-mean-squares fit, the nonlinear parameters c_i and r_i are chosen with a combination of random search and steepest descent, and ϕ_i is one of the following functional forms:

$$\phi_i = \begin{cases} \exp(-\frac{1}{2}x^2) \\ \exp\left(-\frac{1-p}{p}\left|\frac{x}{p}\right|^p\right) \\ (2x^2-1)\exp(-x^2) \end{cases} \quad (3)$$

represent Guassian, "tophat" (a modified Gaussian) or a Mexican hat wavelet. The remaining parameter d is selected according to the minimum description length principle. The description length is the computational cost (in terms of number of bits) required to describe a particular model and the model prediction errors of that model, rather than just describing the raw data. The "best" model is deemed to be the one which affords the shortest description of the data (that is, it achieves the greatest compression as measured with the smallest value of description length.

For a fixed model size d the model which minimises the mean-squares prediction error is deemed to be best, but by selecting model size based on description length we avoid overfitting without the need for validation data. The idea behind minimum description length is detailed in [6] and the application to radial basis and neural network models which we utilise here is described in [4, 10, 8, 7].

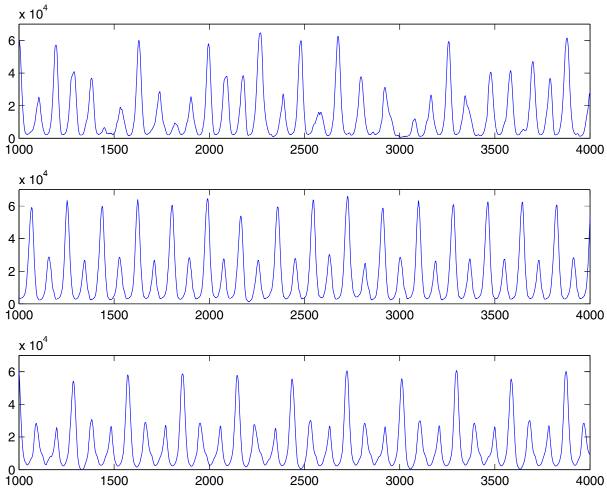

Figure 2: Three simulations from three models of the Canadian lynx data. In each case, a 1000 point transient has been neglected. The dynamics (from top to bottom) are deterministic chaotic, almost periodic, and exactly periodic. The working definition we employ for chaos is detailed in the text. Note that the horizontal axis in this and subsequent plots is datum number. Due to the interpolation procedure described in the text 10 data are equivalent to a time span of one year.

2.3. The results

We repeat the model procedure to produce 100 distinct models with bounded asymptotic dynamics. Typical dynamical behaviour is illustrated in Fig. 2 and a reconstruction of the underlying attractor (compared to the original data) is shown in Fig. 3. These trajectories are meant only to be representative example of the various behaviours observed. From 100 trials, 60% exhibited chaotic dynamics (that is, bounded and aperiodic during the first 2.5×10^4 iterates). the remainder where split roughly 2 : 1 between exhibiting a periodic orbit (during the first 2.5×10^4 iterates) and a fixed point (convergence to a single value).

The chaotic dynamics exhibited by the chaotic models was, in each occasion, qualitatively similar. The dynamics of one such model (chosen randomly) is shown in Fig. 3.

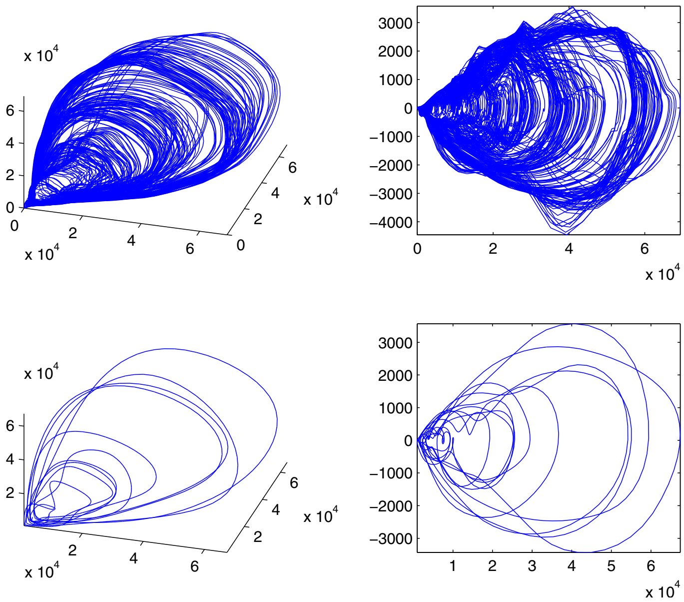

Figure 3: Reconstruction of the attractor with time delay embedding (left, $d_e = 3$ and $\tau = 10$) and differential embedding (right, x_t-vs-$x_t - x_{t-1}$) for a single trajectory from one representative (i.e. chosen only to be chaotic) model (top) and the original time series data (bottom).

The deterministic attractor clearly reproduces all the qualitative features of the original system. Hence, in agreement with Maquet and colleagues [5] we find that a deterministic chaotic model produces an explanation consistent with the observed data.

However, in a minority of cases we observe non-chaotic dynamics. Occasionally the dynamics also exhibit remarkably long transient behaviour. In Fig. 4 we illustrate with three sections from a single trajectory. The initial 1.25×10^4 points behave approximately chaotically, afterwards the dynamics collapses to a periodic orbit. The addition of dynamic noise to the system very rapidly returns the system to the chaotic transient state. Nonetheless, this extremely long transient illustrates that for this system the distinction between periodic and chaotic dynamics is rather moot. Even if we are able to experimentally observe the behaviour of the lynx population in the absence of noise, the periodic nature of the system would only be evident after waiting (in a stationary system) for over 1200 years!

Moreover, the behaviour of the system is extremely sensitive to perturbation by dynamic noise. In the original time series we observe that the amplitude of the cycle varies considerably. In particular in the period between 1910 and 1920, and again between 1920 and 1930, the maximum is very low. It is perhaps natural, but maybe naïve, to guess that over exploitation is playing a role here. In Fig. 5 we show a trajectory from a single chaotic model (the same model is also shown in Fig. 2 and 3). In this case, the trajectory is driven with small amplitude dynamic noise. The root-mean-square model prediction error for this model is approximately 132, we drive the system with dynamic noise with an amplitude of 50 (again, this value is only chosen for the purposes of illustration). From the simulation in Fig. 5 it is clear that this small perturbation is sufficient to cause a bountiful harvest or a very lean maximum in the following decade. Therefore, even twenty years of population data (say during the period 1910 to 1930) would not necessarily provide sufficient information to reliably determine the behaviour of the following cycle.

3. Conclusion

In this paper we have applied the minimum description length criterion and an established radial basis modelling procedure to offer are-examination of evidence for low-dimensional chaos in the Canadian lynx time series data. The data has been widely studied and it has recently been observed that low dimensional polynomial models fit to the data exhibit chaotic dynamics, and that this may be interpreted as evidence for chaos in a natural ecosystem. We have no strong reason to suppose that polynomial models are the correct model class for this situation and have therefore employed a much broader and more generic class of nonlinear models.

Nonetheless, from our radial basis models we get results which appear to support this earlier conclusion: most of our models do exhibit chaotic dynamics and the agreement between the chaotic dynamics of our models and the original data is exceedingly good (Fig 3). However, we also find a significant minority of models which are not chaotic, and

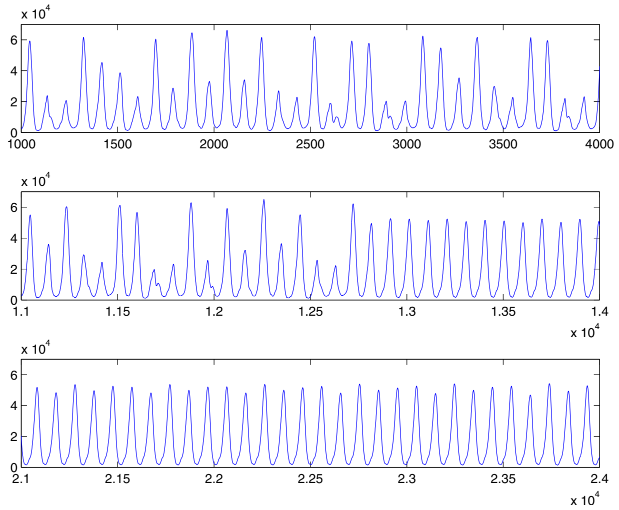

Figure 4: Three sections of trajectory from a simulation from a single periodic model. Note that the periodic nature of this simulation is only evident after allowing the simulation to iterate over 1.25×10^4 times.

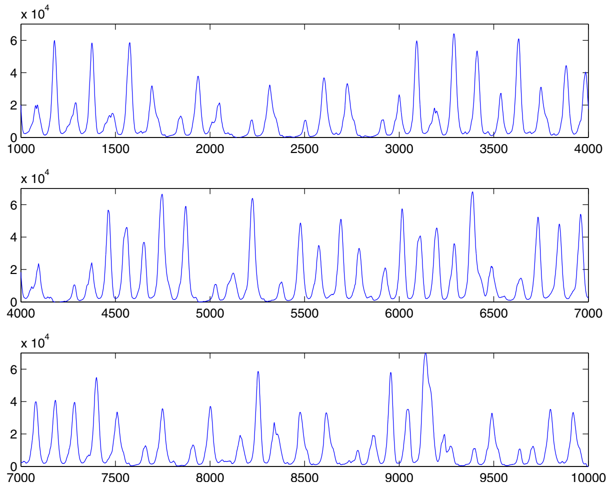

Figure 5: Three sections of trajectory from a simulation from a single chaotic model driven with low-amplitude dynamic noise. Note that the amplitude of the subsequent cycle maxima is sensitively dependent on the noise. Population crashes qualitatively similar to those observed in the real data in 1910-1930 are frequently evident here.

on many occasions these periodic models exhibit extremely long transients that could easily be mistaken for numerical chaos. Of course, we can only conclude that both dynamics offer a reasonable explanation for the available data. With models exhibiting transient last over 1000 years it is impossible to say from 113 years of data whether the systems is actually chaotic or not. However, we emphasise that deterministic chaos, as exemplified in our models does provide an attractive and parsimonious explanation.

We are currently working on a more detailed description of these models and their behaviour. In particular, we are comparing traditional dynamic invariants for the models and the original data [3] as well as applying new complex network based metrics [9].

Acknowledgments

The authors would like to Christophe Letellier for sharing the version of the lynx time series data reported in their paper. This research was funded by a Hong Kong University Grants Council Competitive Earmarked Research Grant (CERG) number PolyU 5279/08E. MATLAB code is available from the first author.

References

[1] B. Blasius, A Huppert, and L. Stone. Complex dynamics and phase synchronization in spatially extended ecological systems. *Nature*, 399:354, 1999.

[2] W.S.C. Gurney, S.P. Blythe, and R.M. Nisbet. Nicholson's blowflies revisited. *Nature*, 287:17, 1980.

[3] Kevin Judd. Estimating dimension from small samples. *Physica D*, 71:421–429, 1994.

[4] Kevin Judd and Alistair Mees. On selecting models for nonlinear time series. *Physica D*, 82:426–444, 1995.

[5] J. Maquet, C. Letellier, and Luis A. Aguirre. Global models from the canadian lynx cycles as a direct evidence for chaos in real ecosystems. *J. Math. Biol.*, 55:21–39, 2007.

[6] Jorma Rissanen. *Stochastic complexity in statistical inquiry*. World Scientific, Singapore, 1989.

[7] Michael Small and Kevin Judd. Comparison of new nonlinear modelling techniques with applications to infant respiration. *Physica D*, 117:283–298, 1998.

[8] Michael Small and C.K. Tse. Minimum description length neural networks for time series prediction. *Physical Review E*, 66:066701, 2002.

[9] Xiaoke Xu, Jie Zhang, and Michael Small. Superfamily phenomena and motifs of networks induced from time series. *Proc Natl Acad Sci USA*, 105:19601–19605, 2008.

[10] Yi Zhao and Michael Small. Minimum description length cristerion for modeling of chaotic attractors with multilayer perceptron networks. *IEEE Transactions on Circuits and Systems I*, 53:722–732, 2006.