

Dynamical Mechanisms Underlying Avian Influenza Outbreaks

Cristian Carmeli and Michael Small

Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hung Hom, Hong Kong
E-mail: cristian.carmeli@polyu.edu.hk; ensmall@polyu.edu.hk

Abstract—Using global case data for the period from 25 November 2003 to 10 March 2007, we construct a time series of the number of outbreaks for the spread of avian influenza among domestic and wild birds. The correlations of these spread patterns are reflected in a power-law scaling of this property. The temporal dynamic of the spread of influenza is further investigated with a standard nonlinear time series analysis tool. We suggest that the dynamic underlying the spread of avian influenza is not low-dimensional and hence, characterized by short term unpredictability.

1. Introduction

In the current context of global infectious disease risks, a better understanding of the dynamics of major epidemics is urgently needed. Mathematical models of disease transmission can provide a framework for improving our understanding of the complex dynamics of infectious disease epidemics [1]. This is crucial in attempts to design effective intervention and control strategies. Since the early 1900s, sophisticated mathematical and statistical methods have been used [2]. Nevertheless, the lack of appropriate data-sets has impeded the validation of mechanistic mathematical models. More recently, time series methods have appeared as an interesting alternative and have been used to explore the dynamics of numerous epidemics [3].

Here, we consider tools from the field of nonlinear time series analysis to investigate the time predictability of the spread of avian influenza among domestic and wild birds. Recently, the global spatiotemporal distribution of avian influenza cases have been examined by considering the complex network topology of the outbreaks [4]. In contrast to standard mathematical models of disease transmission, it was found that the current avian influenza outbreak is consistent with a network transmission of disease which does not exhibit a positive threshold. Hence, the disease will continue to propagate even with a vanishingly small rate of transmission.

In this paper, we use a different approach and we examine more closely the dynamics of the outbreaks by studying the predictability of the peaks of the outbreaks. The forecast of peaks can be of paramount importance in practice, because high peaks are evidently associated with undesirable consequences. Indeed, in many ecological systems the episodes that one would like to forecast are very often blooms, outbreaks and demographic explosions. Peak

to peak dynamics have been investigated in those systems and found to be useful to understand the dynamical mechanism of those episodes [5].

After the description of the data in Sec. 2, we briefly introduce the tool of *peak-to-peak plots* in Sec. 3. The results of our investigation are given in Sec. 4 and, finally, conclusions are given in Sec. 5.

2. Data

The data we use in this study are a compilation of all reported avian cases of avian influenza between 25 November 2003 and 10 March 2007. The data consist of 3346 recorded cases. For each case, the date of the outbreak and the location (longitude and latitude) are recorded. Individual cases may either be wild birds that are found (possibly *post-mortem*) and determined to be infected with a strain of avian influenza or the detection of an avian influenza strain in a domestic flock (most probably then followed by culling of that flock). Data relating to the magnitude of each incident are also recorded. Human cases of avian influenza have also been recorded in the same data set, but for this study, these are ignored. The entire data set is compiled from a variety of sources¹; Fig. 1 depicts one snapshot.

In a previous study [4], this dataset was analyzed within a complex networks framework. Specifically, to investigate the geographical and temporal distribution of avian influenza transmission, susceptible individuals or communities were treated as nodes on a network, and potential transmission pathways as the links between those nodes. Being the actual transmission pathways unknown, two nodes were linked if that transmission could occur only over a local area (in both time and space). The network structure obtained exhibits scale-free properties with a scale exponent of about 1.2. This has the important consequence that communities within this network are connected with a distribution of links with infinite mean and variance. Hence, the disease transmission model does not exhibit a threshold and so the infection will continue to propagate even with a very low transmissibility. This results in contrast with standard mathematical models of geographical

¹The data originally come from World Organization for Animal Health alerts (see <http://www.oie.int/>) and World Health Organization case reports, and they are all manually entered using ArcGIS and converted to Keyhole Markup Language (KML) using Arc2Earth (<http://www.arc2earth.com/>). The data are available, in a format compatible with Google Earth (KML), from Astrophysical Journal Supplement.

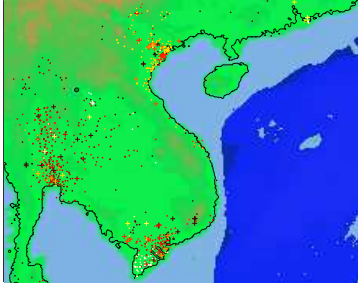


Figure 1: Part of the data used in this study, overlaid against a crude map of the coastline of East Asia. Color coding is by date. The three large clusters correspond to the outbreaks in Cambodia and in the north and south of Vietnam (around Hanoi and Ho Chi Minh City), respectively. Hainan Island is marked in the northeast portion of the image and outbreaks in Hong Kong are shown in yellow in the far northeast corner.

transmission of an infectious agent, which assume that the terrain is locally homogeneous and that the pathogen will diffuse uniformly. A natural consequence of this formulation is that if the transmissibility of the pathogen is lower than some threshold, the disease will terminate. However, if a disease is spreading on a scale-free network, then eradication of that disease is only possible if transmission is reduced to precisely zero.

In this study, we analyze the dataset within a time-series framework and, hence, we focus only on the temporal dynamic of the avian influenza outbreaks. We aim at investigating the predictability on time of the avian influenza outbreaks. In particular, we are interested in assessing the short time predictability of the outbreaks, or, in other words, in evaluating whether the world-wide dynamic of the outbreaks can be modeled with a low-dimensional, possibly nonlinear, process. Despite our interest in merely unveiling nonlinear phenomena in the concrete world, the knowledge of a deterministic rule in the propagation of the outbreaks would be clearly helpful to design more effective intervention strategy to control or eradicate avian influenza.

The time series we analyzed is shown in Fig. 2. This time series were obtained from the dataset by considering² the amount of recorded outbreaks at any available recorded event, which are available on a daily basis.

We remark that the time series in Fig. 2 has an irregular behavior characterized by several peaks. The forecast of the peaks would be of paramount importance in practice, because high peaks are evidently associated with undesirable consequences. To unravel the existence of peak-to-peak dynamic, we can make use of peak-to-peak plots, which are a well grounded tool in the field of nonlinear time series analysis [5]. The next section is devoted to an introduction of this tool.

²Independently on the geographical location.

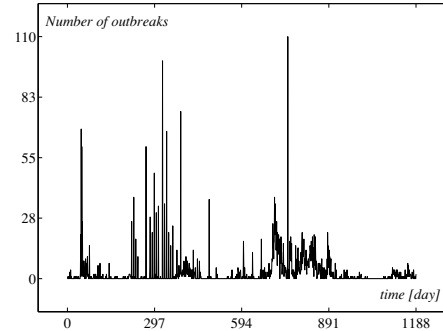


Figure 2: Time series of the avian influenza outbreaks. The coarse grain of the sampling is a day. Time equal to 0 corresponds to the day of the first recorded outbreaks, *i.e.* 25 November 2003. Time equal to 1188 corresponds to the last recorded available outbreak, *i.e.* 10 March 2007.

3. Method: Peak-to-peak plots

After transient, any deterministic, dissipative, nonlinear system settles on an attractor and remains there forever if it is not perturbed [6]. Insights about the attractor can be obtained, even in the absence of a mathematical model, if a single variable y of time t has been recorded for a long period, while the system was on the attractor. In particular, one can extract from the record all peaks (local maxima) of the variable y , say y_i ($i = 0, 1, 2, \dots$), and plot them one against the previous one, thus obtaining a set of points (y_i, y_{i+1}) , called peak-to-peak plot (PPP)³. The same plot is sometimes called next-amplitude plot, or next-maximum plot or Lorenz plot [7, 6]. If the regime is periodic and there are k peaks per period, the PPP is composed only of k distinct points. By contrast, if the regime is quasiperiodic or chaotic, *i.e.* if the attractor is a torus or a strange attractor, the points of the PPP are all distinct and sometimes display filiform geometries. More precisely, when the attractor is a high-dimensional strange attractor, the PPP is a cloud-like set. Conversely, the points of the PPP lie on a closed regular curve when the regime is quasiperiodic and lie roughly on a curve when the attractor is a low-dimensional strange attractor. An example of PPP from the Lorenz [8] strange attractor is reported in Fig. 3.

Peak-to-peak analysis is a special case of the standard technique for reconstructing strange attractors [9]. Indeed, the observation of the output peaks is equivalent to the observation of the system on a Poincaré map. For this reason, it is possible to show that the existence of peak to peak dynamics (PPD) is simply related to the dimension of the attractor [5].

One of the advantage of peak-to-peak analysis is that requires only very little effort (it can be performed by hand

³From the same record one can also extract the times of occurrence of the peaks t_i , compute the return times $\tau_{i+1} = t_{i+1} - t_i$ and argue about the return times dynamic.

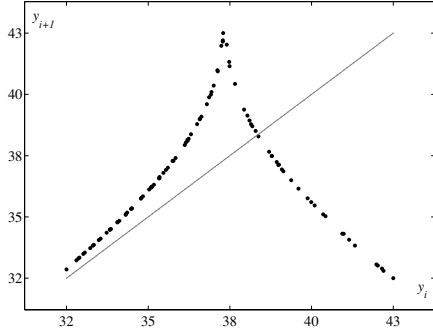


Figure 3: An example of PPP obtained through simulation of the Lorenz system model with standard parameters $b = 8/3$, $r = 28$, $\sigma = 10$ and $y = x_3$. The PPP is clearly filiform, *i.e.* its points are roughly distributed along one curve.

in a few minutes), especially if compared with other techniques of nonlinear time series analysis. Furthermore, the existence of PPD does allow not only the forecast of the amplitude of next peak from the previous peak, but it can be shown that the time of occurrence of that forthcoming peak can also be predicted [5]. The forecast of the time of occurrence of the next peak is of great importance, especially in ecology, where in many cases the amplitude of the peak is hardly measurable, whereas its time occurrence is perfectly known [10]. Despite these advantages, the identifiability of PPD from time series data may require long data and low noise intensity. Indeed, long time series are required to reliably identify all the details of the structure of the PPD, and false peaks due to measurement and process noise may have a severe impact on the PPP. In fact, a false peak f between two successive true peaks a and b implies the loss of point (a, b) in the PPP and the addition of two false points (a, f) and (f, b) . In practice, exclusion of false peaks through common sense can be effective.

PPP have been used extensively to analyze mathematical models [11, 12], and have also been pointed out by analyzing laboratory or field data, like in biochemistry [13], electronics [14] and neurophysiology [15]. In the field of epidemiology, measles epidemics in New York City have been found to possess filiform PPP [16]. Encouraged by this finding, we proceeded to the analysis of our dataset with PPP.

4. Results

The PPP obtained from the time series in Fig. 2 is reported in Fig. 4. To decrease the possibility of counting false peaks, we also applied some pre-filterings driven by common sense. We decided to consider as peak only those detected peaks with amplitude bigger than a certain threshold. This threshold value may be well related to alert status triggering actions of control of the outbreaks. We obtained qualitative similar results for threshold values equal to 5,

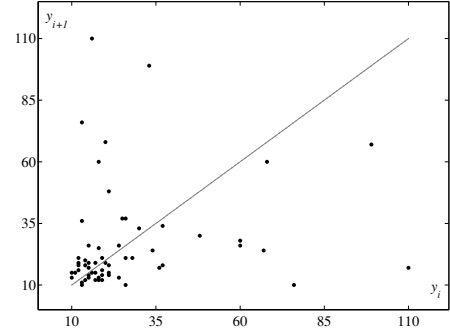


Figure 4: Peak-to-peak plot of the time series in Fig. 2.

10 and 20. The PPP in Fig. 4 clearly does not show any filiform structure; furthermore, we could not observe any specific structure by partitioning the domain of the PPP and inspecting the visits on time of the peaks in those partitions. This suggests that the global dynamic of the avian influenza outbreaks does not possess PPD. Importantly, this further suggests that the outbreaks does not have short-time predictability. This further corroborates the ill-omened character of avian influenza.

Since the dynamics of the outbreaks did not reveal any low-dimensional dynamic, we further processed the time series from a stochastic point of view by inspecting the probability distribution of the outbreaks. Indeed, we may expect that the outbreaks follow a power-law distribution similar to the one found in [4] for the network degree distribution. Practically, we look at fitting to the data a probability distribution $P(m) \propto m^{-\gamma}$, with $\gamma > 1$. The estimated probability distribution of the outbreaks is shown in Fig. 5. By weighted least squares fitting, we find that the distribution of the outbreaks well follows (R^2 goodness of fit equal to 0.9770) a power-law distribution with a scale exponent of ≈ 1.2680 . Consequently, because a power-law distribution with such an exponent has neither finite mean or variance, we expect that avian influenza outbreaks do not exhibit a positive threshold and the disease will continue to propagate even with a vanishingly small rate of transmission. This is in agreement with the results found in [4] and obtained within a different point of view.

5. Discussions and Conclusions

In this study we have analyzed the temporal dynamic of global avian influenza outbreaks with peak-to-peak plots. The analysis suggests that the outbreaks are unpredictable (*i.e.* short term predictability) and that they would propagate even with very low transmissibility. This result supports the finding described in [4], though the latter has been obtained with a rather different point of view. Importantly, our analysis has the advantage to be parameter free, while in the previous study the construction of the spatio-temporal network of disease transmission required

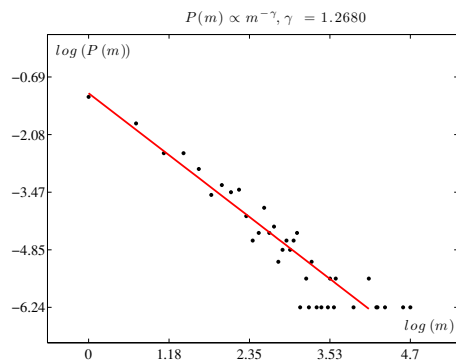


Figure 5: Probability distribution $P(m)$ of the outbreaks displayed on a log-log scale. The linear fitting is in red color. The outbreaks show a power-law distribution with estimated scale exponent of 1.2680 and goodness of fit $R^2 = 0.9770$.

the use of two parameters (those defining the local area, both in time and space, over which the transmission can occur). Certainly, on the other hand, our analysis loses spatial specificity and does not allow to argue about the mathematical modeling of the spatial spreading of the disease.

The major caveat of our analysis relies in the limited amount of data at hand and in being a model-free analysis. Indeed, longer data and support from model simulations would make the PPP based analysis more reliable.

Future work will endow the investigation of the outbreaks dynamic over restricted spatial regions, in the attempt to reduce the degrees of freedom that are inherently high at a world-wide scale.

Acknowledgments

CC acknowledges financial support from the Hong Kong University Grants Council (UGC) Competitive Earmarked Research Grant (CERG) number PolyU5269/06E.

References

- [1] J. Murray, *Mathematical Biology*. Springer-Verlag, 2nd ed., 2002.
- [2] R. Ross, "Some quantitative studies in epidemiology," *Nature*, vol. 87, pp. 466–467, 1911.
- [3] R. Catalano and S. Serxner, "Time series designs of potential interest to epidemiologists," *American Journal of Epidemiology*, vol. 161, pp. 1159–1167, 1987.
- [4] M. Small, D. M. Walker, and C. K. Tse, "Scale-free distribution of avian influenza outbreaks," *Physical Review Letters*, vol. 99, p. 188702, 2007.
- [5] M. Candaten and S. Rinaldi, "Peak-to-peak dynamics: A critical survey," *International Journal of bifurcation and Chaos*, vol. 8, pp. 1805–1819, 2000.
- [6] S. H. Strogatz, *Nonlinear Dynamics and Chaos. Studies in Nonlinearity*. Reading, MA: Addison-Wesley, 1994.
- [7] S. K. Scott, *Oscillations, Waves and Chaos in Chemical Kinetics*. Oxford: Oxford University Press, 1994.
- [8] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of Atmospheric Science*, vol. 20, pp. 130–141, 1963.
- [9] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge: Cambridge University Press, 2nd ed., 2004.
- [10] S. Rinaldi, R. Casagrandi, and A. Gagnani, "Reduced order model for the prediction of the time of occurrence of extreme episodes," *Chaos, Solitons Fractals*, vol. 12, pp. 313–320, 2001.
- [11] E. Funasaki and M. Kot, "Invasion and chaos in a periodically pulsed mass-action chemostat," *Theoretical Population Biology*, vol. 44, pp. 203–224, 1993.
- [12] S. Rinaldi and C. Solidoro, "Chaos and peak-to-peak dynamics in a plankton-fish model," *Theoretical Population Biology*, vol. 54, pp. 62–77, 1998.
- [13] M. R. Basset and J. L. Hudson, "Shil'nikov chaos during copper electrodisolution," *Journal of Physical Chemistry*, vol. 92, pp. 6963–6966, 1988.
- [14] H. D. I. Abarbanel, L. Korzinov, A. I. Mees, and N. F. Rulkov, "Small force control of nonlinear systems to given orbits," *IEEE Transactions of Circuits and Systems I*, vol. 44, pp. 1018–1023, 1997.
- [15] M. Balduzzo, T. A. Minelli, and L. Turicchia, "Signal analysis and simulation of the eeg activity," *International Journal of Chaos Theory and Applications*, vol. 4, pp. 7–14, 1999.
- [16] F. R. Drepper, R. Engbert, and N. Stollenwerk, "Nonlinear time series analysis of empirical population dynamics," *Ecological modelling*, vol. 75/76, pp. 171–181, 1994.