A SYSTEMS APPROACH TO FORECASTING

by David Orrell and
Patrick McSharry

INTRODUCTION

The practice and profession of computer-based numerical forecasting have reached a crossroads. Forecasting tools and availability of high-quality data have improved enormously in recent decades, but not forecast quality. In areas including weather forecasting, genetics and, particularly, economics – as shown by our failure to predict the current financial crisis – we have become increasingly aware of the limits of prediction. It's clear we need to change the way we make forecasts and the questions we attempt to address. To move forward, we should adopt a **systems approach**.

In the 1950s, in the infancy of computer-based predictions, there was enormous optimism that numerical forecasting would allow us to predict and control the world. In most fields, however, progress has been unspectacular.

Weather forecasters now utilize computers a million times more powerful than those available in the '50s, and have access to plentiful data, including that from dedicated weather satellites. But weather forecasts remain useful for only 2-3 days, and the uncertainty in climate prediction has not decreased since the first studies were done some 30 years ago.

In genetics, the Human Genome Project promised us access to the book of life, but predictions of health or drug efficacy have again lagged behind.

Economists have always found it hard to predict storms, but events in 2008 exposed an alarming lack of forecasting acuity. According to Bloomberg.com (Thomasson, 2009), "Wall Street analysts lost credibility in 2008 when none predicted a down year and the average forecast was for a gain of 11 percent …. Instead, the S&P 500 tumbled 38 percent to 903.25 and $29 trillion was erased from global markets."

Of course, most forecasters don't need to predict the climate system, the human body, or the entire economy. But our difficulties in these areas raise questions about our ability to predict any kind of complex system. The problem is that traditional mathematical modeling techniques devised to work for areas like basic physics or engineering do not perform well when applied to complex systems. It is increasingly clear that we need to adjust both the aims and methods of forecasting. Without realistic objectives and transparent communication of limitations and confidence levels, the field of forecasting will lose credibility with practitioners and policy makers.

Key Points

- Despite improvements in data quality and computer processing, traditional forecasting models have disappointed when applied to complex systems such as the weather, genetics, and economics. The main problem is that the traditional approach breaks a system down into its component parts, figures out the laws governing them, expresses these laws as equations, and solves the equations.

- Complex systems, however, are characterised by emergent properties, qualities that cannot be predicted in advance from knowledge of systems components alone. As a result, complex systems are more than the sum of their parts.

- Systems forecasting is a more promising approach, one that takes advantage of mathematical techniques and concepts that have been developed for complex systems, such as agent-based models, network analysis, and system dynamics.

- Systems forecasting models have proven effective in biological applications, including the spread of pandemics. We think the techniques are equally adaptable to business and finance.

THE FAILURE OF TRADITIONAL FORECASTING MODELS

Traditional forecasting tools are based on techniques first developed for 17th- and 18th-century astronomy. Weather models, for example, employ large sets of ordinary differential equations that attempt to capture the physics of the atmosphere. The approach is essentially the one Isaac Newton used to describe the motion of a body moving under gravity.

In economics, many models (such as the risk models used by banks and financial rating agencies) rely on a statistical approach which approximates fluctuations using the normal or Gaussian distribution, used by Pierre-Simon Laplace to analyze errors of astronomical observations. As with astrology, modern forecasting began with stargazing, and it has never been clear that these methods work well here on Earth. As Newton noted in 1721, after he lost most of his fortune in the collapse of the South Sea bubble, "I can calculate the motions of heavenly bodies, but not the madness of people."

The problem goes back to the "divide-and-conquer" reductionist approach that has characterised Western science since the ancient Greeks. The reductionist approach breaks a system down into its component parts, figures out the laws governing them, expresses these laws as equations, and solves the equations.

Complex systems, however, have emergent properties, qualities that cannot be predicted in advance from knowledge of systems components alone. Emergent phenomena have been widely studied using mathematical systems such as cellular automata, in which grids of blocks change color at each time-step according to defined rules. An example is John Conway's Game of Life, a cellular automaton which became popular in the 1970s and '80s as a screensaver (*www.bitstorm.org/gameoflife*).

Complex systems are more than the sum of their parts, so the reductionist approach no longer applies.

According to complexity scientist Stephen Wolfram's principle of computational irreducibility, the only way to predict the evolution of such a system is to run the system itself: there is no simple set of equations that can look into its future (Wolfram, 2002).

For example, models attempt to predict atmospheric motions using principles of fluid flow. Many aspects of the climate system, however – clouds, for one – are best characterized as emergent properties. There is no "simple" law for a cloud. Modelers attempt to get around this problem by adding equations to capture overall behaviors, but these equations introduce parameters whose values are uncertain. Furthermore, because the climate system is dominated by interlocking positive- and negative-feedback loops, the equations result in models that are highly sensitive to small changes in the parameters: they can be made to fit past data but are less able to predict the future.

In orthodox economics, the reductionist approach means that the economy is seen as consisting of individual, independent agents who act to maximize their own utility. It assumes that prices are driven to a state of near-equilibrium by the "invisible hand" of the economy. Deviations from this state are assumed to be random and independent, so the price fluctuations can be modeled using the normal distribution. The drawbacks of this approach became obvious in the recent credit crunch, when models totally failed to capture the true risks of the economy. As an employee of Lehman Brothers put it on August 11, 2007, "Events that models predicted would happen only once in 10,000 years happened every day for three days."

THE SYSTEMS APPROACH

For models of complex systems to be truly useful, they must be able to account for the networked, dynamic, and emergent properties of these systems. In recent decades, there have been great advances in areas such as network theory and nonlinear dynamics that offer an alternative to reductionist methods (Orrell & McSharry, 2009). The systems approach has fared well with biologists, who know it as systems biology.

Systems Biology. Until the turn of the century, biology was dominated by a reductionist approach which assumed the behavior of organisms could be understood and predicted through the effects of individual genes. It was later recognized that, far from being independent, genes were parts of highly complex, dynamic networks involving other genes, proteins, and factors. For example, a gene can be switched on or off by proteins produced by other genes, and the expression of a gene will vary with time and with position in the body.

The networks are also characterised by multiple feedback loops, both positive and negative. A feedback loop refers to a situation in which a portion of a signal returns to the source, thus modifying the signal. In general, positive feedback tends to accentuate a change to a signal, while negative feedback reduces change.

To make sense of complex systems with feedback loops, biologists had to turn to network theory, complexity, and nonlinear dynamics, all of which form part of systems biology.

Systems biology has proven popular with drug manufacturers. Eli Lilly, Pfizer, Merck, and Novartis are all developing their own systems-biology groups. After all, even if a drug is targeted at an individual gene, you still need a systems approach in order to understand the effect on the whole organism. Increasingly, drugs or combinations of drugs are aimed at multiple targets, in order to maximize their effects. Furthermore, because the body is a dynamic system, drug schedules can have an important impact.

The systems approach does not aim to provide a single "theory of everything." Instead, models are seen as imperfect, fuzzy patches that can be adapted for particular situations – each revealing a different but incomplete aspect of the entire system. For example, the appropriate modeling technique for a complex system often depends on the scale (granularity) being addressed and the question being asked. Also, models of complex systems need not be complicated; indeed, simpler models often provide superior forecasts, if they take into account the relevant properties.

Can the systems approach be generalized to other systems, like the economy? Can it be useful for forecasters? In fact, while hardly mainstream, the basic tools of a systems approach have already been in use by modelers for years.

Agent-based modelling simulates the interactions between multiple agents with different goals who are in communication with each other and influence each other's actions. They have been used, for example, to simulate the function of stock markets, where each agent represents an individual trader with its own strategy (Darley & Outkin, 2007). The rules governing each agent are usually kept very simple; forecasters use the models to study emergent properties which characterise the system, such as financial crashes.

Network theory is a branch of applied mathematics that studies the relationships between objects in complex networks. Examples of applications include gene regulatory networks, social networks, and the World Wide Web. The financial system as a whole can also be usefully viewed as a giant network, whose properties are "key to understanding the ecology of market robustness and its potential vulnerability to collapse" (May and colleagues, 2008). Key parameters for forecasters include the network's level of connectivity: pandemics spread more quickly today because of the large number of long-distance connections through air travel; information about new products propagates faster because of the Internet.

Nonlinear dynamical systems are sets of nonlinear equations specifying the rate of change of a number of variables over time. A solution of the system is a trajectory starting from a specified initial condition. Forecasters need to be aware of nonlinear effects, seen also with agent-based models, which can lead to exponential growth, or drive switches from one state to another. As in weather forecasting, the combination of positive- and negative-feedback loops makes predictions imprecise and often better expressed as scenarios.

Scenarios. Models of complex systems usually seek not to make specific "point" predictions, but to generate scenarios, i.e. alternative versions of the future. In some cases, probabilistic estimates are possible; in others, models are used to give an idea of the range of possibilities, including the worst case. Models can also be used to help steer the system towards a more predictable state or make it more robust. As with earthquakes, we may not know exactly when the next crisis will hit, but we can pinpoint areas of vulnerability and design structures to better withstand shocks.

EXAMPLES

This section gives some examples of practical applications of the systems forecasting approach. Most to date have been in the life sciences; we believe the reason for this is historical, though. The techniques are equally adaptable to business or finance.

The Canadian group Dynemo Biosystems has developed a diagnostic test for breast cancer based on network analysis. The test uses genetic samples from patients to look for changes in the interaction network of selected proteins and can make accurate predictions about the disease's severity and the patient's prognosis (see *www.dynemobiosystems.com*).

Physiomics is a systems biology company that uses a range of models to predict the effect of anticancer drugs on tumor cells. Nonlinear equations are used to simulate the internal biochemistry of the cell; agent-based models are used to simulate the emergent behavior of cell populations in tumors and capture the often highly nonlinear effects that occur when different drugs are combined (*www.physiomics-plc.com*).

One important application for network theory and agent-based modeling is computer simulation of the spread of pandemics. Modelers can create virtual populations which statistically match the network interaction properties of a given city or country (Eubank and colleagues, 2004). An equally important factor is the airline transport network (Colizza and colleagues, 2006). These simulations are used to monitor the progress of a disease, study different scenarios, and explore containment strategies.

Catastrophe modeling for insurance must relate local impacts to global influences, to better predict natural hazards. Recent research has shown that the spatiotemporal characteristic of extreme events such as flooding is driven by large-scale atmospheric influences (Vitolo and colleagues, 2009). Rather than construct vast "black-box" models with numerous parameters, it may be appropriate to employ simple models that nevertheless capture the underlying statistical properties of the global system.

Even if techniques such as network analysis are not directly used, the systems approach still provides insight to a broad range of questions and helps troubleshoot existing methods. For example, orthodox risk management tools, such as Value at Risk (VaR), rely on a number of assumptions – such as independence of events and the likelihood of historical patterns being relevant for the future – that are not consistent with a systems approach. Overuse of VaR played a critical role in the recent financial crisis (Nocera, 2009). The systems forecasting approach would allow extra margin for extreme events, which characterise systems like the economy.

CONCLUSIONS

In recent years, we have become increasingly aware that our businesses and societies are embedded in complex systems – the economic system, the climate system, and so on – which are globally linked and may be susceptible to sudden change. Forecasting tools developed for the demands of classical astronomy are not useful for these situations and can be devastating when taken too seriously (as when they are adopted in the risk models of banks). We therefore need to complement our existing forecasting methods with ones derived from a systems-forecasting approach. Most of the tools for systems forecasting are already available and, tied into a coherent approach, they will offer a robust and attractive alternative to traditional methods. We may not be able to predict the exact timing of the next crisis, but at least we will be better at evaluating the risk – and perhaps we can even learn to prevent it in the first place.

REFERENCES

Colizza, V., Barrat, A., Barthélemy, M. & Vespignani, A. (2006). The role of the airline transportation network in the prediction and predictability of global epidemics, *Proc Natl Acad Sci*, 103:2015-2020.

Darley, V. & Outkin, A.V. (2007). *A NASDAQ Market Simulation: Insights on a Major Market from the Science of Complex Adaptive Systems*, Singapore: World Scientific Publishing Co. Pte. Ltd.

Eubank, S., Guclu, H., Anil Kumar, V.S., Marathe, M.V., Srinivasan, A., Toroczkai, Z. & Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks, *Nature*, 429, 180-184.

May, R.M., Levin, S.A. & Sugihara, G. (2008). Ecology for bankers, *Nature*, 451, 891-3.

Nocera, J. (2009). Risk mismanagement, *New York Times*, January 2, 2009.

Orrell, D. & McSharry, P. (2009). Systems economics: Overcoming the pitfalls of forecasting models via a multidisciplinary approach, *International Journal of Forecasting* (in press).

Thomasson, L. (2009). Strategists see 17% S&P 500 rise on fed cuts after saying 'buy,' Bloomberg.com, Jan. 5, 2009. ***http://www.bloomberg.com/apps/news?pid=20601087&sid=a5mSOfz7Alhk&refer=home***

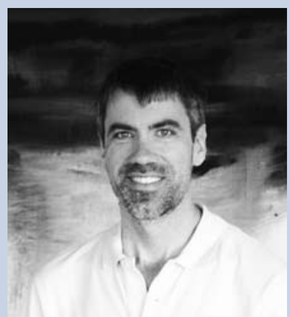
Vitolo, R., Stephenson, D.B., Cook, I.M. & Mitchell-Wallace, K. (2009). Serial clustering of intense European storms. In press. ***http://www.secamlocal.ex.ac.uk/people/staff/rv211/Vitoloetal2008.pdf***

Wolfram, S. (2002). *A New Kind of Science*, Champaign, IL: Wolfram Media, Inc.

Contact Info:
Patrick McSharry
University of Oxford
patrick.mcsharry@gmail.com

Contact Info:
David Orrell
Physiomics, Oxford, UK
david@postpythagorean.com

Bio: Patrick McSharry is a Senior Research Fellow at the University of Oxford and directs the Systems Analysis, Modelling and Prediction group (*www.maths.ox.ac.uk/samp*). He is affiliated with the Saïd Business School, the Mathematical Institute, the Smith School of Enterprise and the Environment and the Oxford Centre for Integrative Systems Biology.



Bio: David Orrell is author of *The Future of Everything: The Science of Prediction* and is Head of Modelling and Simulation at the systems biology company Physiomics in Oxford, UK. (You can read a review of The Future of Everything in the Summer 2008 issue of *Foresight*.)