Journal of the Nigerian Association of Mathematical Physics
Volume 19 (November, 2011), pp 15 – 20

Hierarchical Structured Model for Nonlinear Dynamical Processes.

***[1]Agwuegbo S.O.N, [2] Solarin, A.R.T, [3]Asiribo, O.E**

[1,3]Department of Statistics, University of Agriculture, Abeokuta
[2]National Mathematical Centre, Abuja
[3]Department of Statistics, University of Agriculture, Abeokuta

Abstract

Modelling and predicting nonlinear processes are more complicated, especially where the parameters change through time in a pre-determined way. In this paper, attempts are made to build an appropriate model for the prediction of a nonlinear dynamical system by using hierarchically structured model. The mathematical representation of the process, in this context, is by a set of linear stochastic differential equations (SDE) with unique solutions. The problem of realization is that of constructing the dynamical system by looking at the problem of scientific model building. In model building, one must be able to calculate the dynamical behaviour of the system using mathematical model with an accuracy which is at least within the tolerances allowable under control. For simplification, individual risks policies suggestive enough to demonstrate relevance of the method in more realistic insurance models was illustrated in simple context so that the basic ideas can be easily grasped.

Key words: Brownian motion, Chaos, Dynamics, Markov Chain, Markov Process

Introduction:

The real world generally changes through time and often behaves in a nonlinear way. Nonlinear processes are more complicated, especially where the parameters change through time in a pre-determined way. They are mathematically interesting and sometimes work well in practice. When the processes or experiment is a dynamical system, the traditional approach in the physical sciences has been to model from first principles; that is, to work upwards from known physical laws to make a model of the system [1].

Most systems cannot be represented by small closed-form equations and one of the lessons learnt from the study of nonlinear dynamics is that it is often not possible to model real world systems by closed-form models, essentially because of sensitivity to initial conditions. Laws governing the evolution of a system fix its future behaviour, given an initial state. Evolution is therefore deterministic and given initial state generates a unique evolutionary path. More generally, the extent to which a nonlinear deterministic process retains its properties when corrupted by noise is also unclear. Most nonlinear deterministic processes are chaotic since the evolution of the physical variables appears very disordered.

Chaos, the science of non-linear systems, has provided new tools and understanding that permits modeling important processes that were previously thought to be unmodelable. Chaotic time series prediction studies the application of these techniques to the induction of models for pseudo-random sequences. The existence of a deterministic system that exhibit chaotic pseudo-randomness behaviour akin to indeterminism, and ordinarily associated with true randomness has several implications for data analysis [2]. Time series data that seem random may in reality be chaotic. Chaotic behaviour arises from certain types of non linear models, and a loose definition is apparently random behaviour that is generated by a purely deterministic, non linear system.

One of the most exciting developments in recent theory of differential equations is the discovery that relatively simple differential equations can have solutions which are much more complicated than periodic and quassi periodic solutions. A differential equation is said to be chaotic if there are bounded solutions which are neither periodic nor quassi periodic and which diverge from each other locally [3]. An immediate corollary of the local divergence of nearby solutions is that one losses predictive power in practical situations. However, as chaotic behaviour leads to successive values lying in a restricted subspace. It is sometimes found that the domain of attraction has a strange geometric form (e.g. fractals), and this is called a strange attractor [4]. In this paper, hierarchical modelling approach is introduced for nonlinear dynamical systems. Hierarchical models have gained wide spread use in statistics during the last few decades [5] and have proved to be useful tools for modelling dynamic behaviour in large dimensional state space[6] or exploring structures in complicated

*[1]Corresponding authors: ***Agwuegbo S.O.N,*** E-mail: agwuegbo_son@yahoo.com, Tel. +2348038004887

data[7],[8],[9]. Hierarchical models are central to many current analysis of functional imaging data including random effects analysis, environmental processes[10],[11],[12], in geographical information systems[13] and also for predicting the spread of ecological processes[14],[15]. Recent advances in computational efficiency have allowed for the implementation of sophisticated hierarchical models [13],[14],[15],[16].

At the same time, such model specification allow for a more intuitive setting where identification and estimation of probability distribution comes more naturally. The hierarchical approach is often a natural way to model a system [12] and its framework can generalize most of the standard modelling techniques used in classical statistical modelling. Hierarchical model is often used to simplify modelling specifications, account for uncertainties and use first principles in a series of conditional models coherently linked together by simple probability rules [12].

Methodology

Hierarchically structured model is used in this study as a model source that allowed the realization of the estimated probability distribution for our nonlinear dynamical system. The approach ensured the construction and estimation of a complex joint distribution through a sequence of simpler and more intuitive conditional distributions. The problem of realization is that of constructing a dynamical system by looking at the problem of scientific model building. In model building, one must be able to calculate the dynamical behaviour of a physical system using mathematical model with an accuracy which is at least within the tolerances allowable under control.

The mathematical model in this study arises when time is continuous. Suppose a scalar time series $\{X(t), t = 1, 2, ..., N\}$ is a measurement on a chaotic dynamical system in the state space. Differencing in discrete time corresponds to differentiation in continuous time, so that apparently natural ways of trying to define a first –order autoregressive (AR) process in continuous time is by means of the general equation

$$aX(t) \quad + \quad \frac{dX(t)}{dt} = \mathrm{Z(t)} \tag{1}$$

where a is a constant, and $Z(t)$ denotes continuous white noise. The solutions of differential equations are functions that describe the evolution or dynamics of a real life process over a given period of time and can change quite drastically. For our purposes, the randomness in the differential equation is introduced via an additional random noise term

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t \tag{2}$$

where $B = (B_t, t \geq 0)$ denotes Brownian motion, which is at the core of stochastic analysis. Further, $a(t, x)$ takes care of the drift of the process, while $b(t, x)$ describes the strength of the extraneous fluctuations caused by the Brownian motion. A wealth of possible models can be arrived at for specific choices of $a(t, x)$ and $b(t, x)$ (17). A naive interpretation of (2) assert that the change $dX_t = X_{t+dt} - X_t$ is caused by a change dt of time, with factor $a(t, X_t),$ in combination with a change $dB_t = B_{t+dt} - B_t$ of Brownian motion, with factor $b(t, X_t).$ Equation (2) can be interpreted as a stochastic integral equation

$$X_t = X_0 + \int_0^t a\ (s, X_s)ds + \int_0^t b(s, X_s)dB_s \qquad 0 \leq t \leq T \tag{3}$$

Where the first integral on the right hand side is a Riemann integral and the second one is an Itŏ stochastic integral. Equation (3) is an Itŏ stochastic differential equation, and the driving process of (3) is the Brownian motion B. Thus a strong solution to (3) is based on the path of the underlying Brownian motion. By taking $a(t, x) = 0$ and $b(t, x) = 1,$ the Brownian motion in (3) is a diffusion process. The solution to an Itŏ stochastic differential equation can be derived as the solution of a partial differential equation

$$X_t = X_0 + c\int_0^t X_s ds + \sigma\int_0^t dB_s \qquad t \in [0, T] \tag{4}$$

Equation (4) usually is referred to as Langevin equation. In the physical literature, the random forcing in (4) is called additive noise which is an adequate description of this phenomenon. Moreover, the Langevin equation in (4) is a linear Itŏ stochastic differential equation (18). The langevin equation in (4) is related to the world of time series analysis. In intuitive form (4) can be written as

$$dX_t = cX_t d_t + \sigma dB_t \tag{5}$$

and formally setting $d_t = 1$. Then

$$X_{t+1} - X_t = cX_t + \sigma(B_{t+1} - B_t) \tag{6}$$

or

$$X_{t+1} = \varphi X_t + Z_t \,. \tag{7}$$

where $\varphi = c+1$ is a constant and the random variables $Z_t = \sigma(B_{t+1} - B_t)$ constitute an $i.i.d$ sequence of $N(0,\sigma^2)$. This is an autoregressive process of order one. This time series model can be considered as a discrete analogue of the solution to the Langevin equation (4) and the langevin equation is a linear itô stochastic differential equation. The stochastic process X_t follows a random walk and can be represented as

$$X_t = c + X_{t-1} + a_t \tag{8}$$

with a constant c and white noise a_t. If c is not zero then the variables, $X_t - X_{t-1} = c + a_t$ have a non zero mean and is called a random walk with a drift. By adopting Box and Jenkin (1976) model approach to time series analysis, model identification, parameter estimation and diagnostic check were feasible. The model identification according to Box and Jenkins involved using differencing, ACF and PACF. The Box and Jenkins ARIMA models can be shown to be optimal and provides a systematic approach to model selection, utilizing all the information contained in the sample autocorrelaion (ACF) and partial autocorrelation (PACF) functions. The ACF and PACF are meaningful only when applied to stationary series.

RESULTS

The empirical data shown in Table 1 is the daily claim occurrence submitted to the secretariat of the Nigeria Insurance Association (NIA) for the period 2007. The distribution of the claim occurring at time revealed specific dynamic behaviour, which provides a better understanding of an insurance portfolio. For most dynamic system, difference equations are used to model the evolution of the system with time, and measurements are assumed to be available at discrete times. The examination of the time plot of Table 1 revealed greater variability of claims as shown in Fig 1.

The path of the model as depicted in Fig 1 is chaotic, making it difficult to define the path of the process. These sudden structural changes were reduced by the use of hierarchical structure as in Fig 2. The study modelled the transition rates of one state to another, and arrived at the probability of each ensemble. If the actual distribution is to be assessed using binomial process over a small time interval, the Wiener process or Brownian motion is then the ideal limiting distribution for the random walk model. The structural model sets out to capture the salient features of a time series and these are apparent from the nature of the series.

The study adopted this approach using the S-PLUS package and the sample autocorrelations (ACF) function, the partial autocorrelation (PACF) function for the claim Processes as shown in Table 3. The S-PLUS package used the Akaike information criterion (AIC) to provide the best fit for an autoregressive model to a set of data. The values of the AIC generally in S-PLUS are listed for the autoregressive models, with the smallest value of the AIC adjudged to be the most appropriate. By the examination of the ACF, the PACF, and the AIC for claim processes suggests an autoregressive model of order 1 as in Table 3. By fitting an AR (1) model for the Claim Portfolio, the corresponding fitted autoregressive model is

$$X_t = 2.36 + 0.114 X_{t-1} + a_t$$
$$(0.066) \quad (0.059)$$

The numbers in parentheses below the coefficients are standard errors. The estimated prediction variance $\hat{\sigma}^2 = 0.009$ and 95% confidence interval for φ is given as

$$0.115 \pm t_{283}(2.5)(0.059)$$
$$(0.00064,\ 0.2306)$$

The ARIMA model diagnostic is as shown in Fig 3 with various plots produced such as the standardized residuals, the ACF of the residuals, the PACF of the residuals. Almost all the plots are based on the examination of the residuals, $\hat{e}_t = y_t - \hat{y}_t$, where $\hat{y}_t$ is the fitted value, or some function of the residuals. An overall test of model adequacy is provided by Ljung-Box chi-squared statistics. These statistics also known as the Box-Pierce chi-square statistics contain what are known as the portmanteau statistics with their associated p-values.

Discussion

Hierarchically structured model is used in this study as a model source that allowed the realization of the estimated probability distribution for our nonlinear dynamical system. The approach ensured the construction and estimation of a complex joint distribution through a sequence of simpler and more intuitive conditional distributions. The hierarchical structured model is based on the Markov property, which implies that given the present state, the future of a system is independent of its past. The study identified that an AR (1) model is adequate and is a special case of our hierarchical structured model.

Table 1. Insurance Claims portfolio

Jan:	4469654, 991698, 1243344, 513000, 522473, 3800000, 744538, 610536, 900000, 573750, 2025000, 570978, 542400, 2592000, 574536, 705682, 719059, 933038, 696173, 665766, 581487, 700000, 750000, 661000
Feb:	684068, 750988, 3401510, 1389917, 1113524, 623396, 3036377, 2334479, 2208789, 1107032, 1050000, 691572, 612716, 537750, 1072512, 1312500, 838080, 12831900, 10159375, 1516854, 500000, 7000000, 1682558, 1012500
March:	1700000, 899965, 513000, 3000000, 1066400, 5772690, 629803, 819500, 1923000, 720000, 1624500, 2183298, 1133930, 1680120, 1938070, 1163191
April:	1091250, 1327784, 3066729, 17093431, 645188, 612000, 586528, 6277603, 864374, 3278926, 1260000, 738000
May:	678000, 577776, 869660, 501785, 777600, 965037, 1260000, 1304394, 1800000, 1083598, 1177808, 1310143, 1298097, 1968740, 760089, 2700000, 4302997, 1224172, 643204, 1344823, 2588500, 3165761.
June:	532010, 1495237, 742808, 724693, 1176641, 7000000, 2109000, 810000, 699510, 531644, 966845, 2618136, 1230618, 3162224, 1081947, 666750, 1500000, 652500, 529376, 4850000, 878500, 520600, 3298464, 1980000, 679115, 632591, 540000.
July:	627838, 723050, 504900, 2263523, 606000, 1729917, 1950000, 704660, 1000000, 2479351, 1417898, 500000, 810000, 1408603, 1256584, 1620000, 540000, 522969, 717039, 982816, 4250000, 700000.
Aug:	760000, 2353302, 546826, 531451, 823125, 515735, 1364993, 1030228, 1393273, 3244220, 1080000, 1044000, 24400000, 1075284, 1070244, 1197000, 995663, 761846, 4254817, 1162800, 24579269.
Sept:	1338750, 1338750, 6300000, 3179432, 800650, 891950, 4454095, 2307436, 559558, 559545, 2063844, 2685354, 1246300, 1245983, 4884223, 857719, 566820, 631125, 1648438, 832733, 3254900, 2061216, 1085797.
Oct:	5235988, 688500, 1411242, 2607147, 1530000, 800000, 1620000, 1067600, 826350, 1982973, 576000, 1381026, 6697192, 3265331, 3222164, 1238226, 828800, 1657500, 14552619, 1121850, 842387, 728946, 3734997, 1341743, 546950, 1134488, 544266, 1351500, 562002, 1851600, 1823018, 3054268.
Nov:	1048478, 1625570, 3886258, 3910305, 1313125, 2900000, 600750, 800000, 1026667, 14490580, 563170, 705093, 1792500, 2153730, 2920459, 643357, 7435587, 542500, 565213, 5178084, 5161160, 2207540, 513359, 746971, 1882850, 2089548, 680400, 553248, 914973, 1080000, 1346386, 27311939.
Dec:	500000, 1042321, 765000, 1344823, 746971, 1339595, 3134790, 540510, 661500, 671700, 1768000, 3587542, 1051200, 1303154, 1298996, 544000, 1744652, 3017240, 3865360, 711461, 992062, 515800, 870795, 665000, 675000, 1080000, 3451391, 524846.

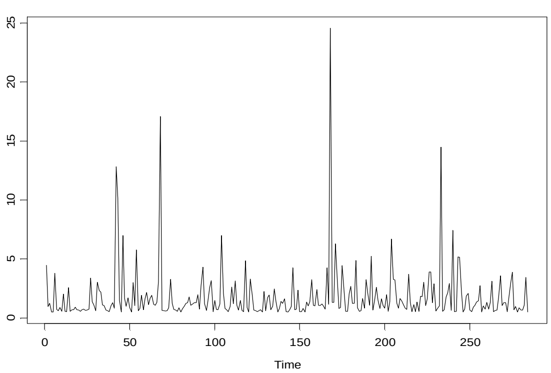

Fig. 1 Sample Paths for the Claim Portfolio

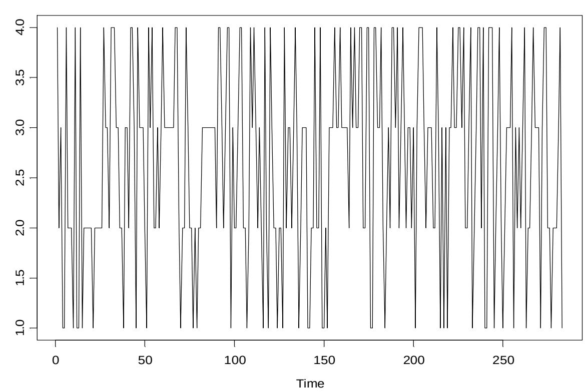

Fig 2 Realization of claims using state space hierarchically structured model

The time plot from the hierarchical state space is in Figure 2

Table 2: The Hierarchical state spaces

4, 3, 2, 1, 1, 4, 2, 2, 2, 1, 4, 1, 1, 4, 1, 2, 2, 2, 2, 2, 1, 2, 2,2, 2, 2, 4, 3, 3, 2, 4, 4, 4, 3, 3, 2, 2, 1, 3, 3, 2, 4, 4, 3, 1, 4, 3, 3, 3, 2, 1, 4, 3, 4, 2, 2, 3, 2, 3, 4, 3, 3, 3, 3, 3, 3, 4, 4, 2, 2, 1, 2, 2, 4, 3, 2, 2, 1, 2, 1, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 2, 4, 4, 3, 2, 3, 4, 4, 1, 3, 2, 2, 3, 4, 4, 2, 2, 1, 2, 4, 3, 4, 3, 2, 3, 2, 1, 4, 2, 1, 4, 3, 2, 2, 1, 2, 2, 1, 4, 2, 3, 3, 2, 3, 4, 3, 1, 2, 3, 3, 3, 1, 1, 2, 2, 4, 2, 2, 4, 1, 1, 2, 1, 3, 3, 3, 4, 3, 3, 4, 3, 3, 3, 3, 2, 4, 3, 4, 3, 3, 4, 4, 2, 2, 4, 4, 1, 1, 4, 4, 3, 3, 4, 2, 1, 2, 3, 2, 4, 4, 3, 4, 2, 3, 4, 3, 2, 3,3, 2, 3, 1, 3, 4, 4, 4, 3, 2, 3, 3, 3, 2, 2, 4, 3, 1, 3, 1, 3, 1, 3, 3, 4, 3, 3, 4, 4, 3, 4, 2, 2, 3, 4, 1, 2, 3, 4, 4, 2, 4, 1, 1, 4, 4, 4, 1, 2, 3, 4, 2, 1, 2, 3, 3, 3, 4, 1, 3, 2, 3, 2, 3, 4, 1, 2, 2, 3, 4, 3, 3, 3, 1, 3, 4, 4, 2, 2, 1, 2, 2, 2, 3, 4, 1

Table 3: Sample ACF, PACF and AIC for Claims Portfolio

Lag K	ACF	PACF	AIC
1	0.1144	0.1144	0.000
2	-0.0584	-0.0724	0.502
3	-0.0010	0.0168	2.422
4	-0.0415	-0.0488	3.743
5	0.0637	0.0775	4.026
6	0.0141	-0.0100	5.997
7	0.0173	0.0284	7.767
8	0.1354	0.1293	4.965
9	-0.0315	-0.0575	6.021
10	-0.0588	-0.0354	7.664
11	0.0560	0.0722	8.177
12	0.0069	-0.0213	10.048
13	-0.0306	-0.0420	11.544
14	0.0170	0.0255	13.360
15	0.0190	0.0204	15.241
16	-0.0365	-0.0707	15.0813
17	-0.0052	0.0229	17.663
18	-0.0185	-0.0092	19.639
19	-0.0024	-0.0237	21.479
20	0.0172	0.0193	23.372

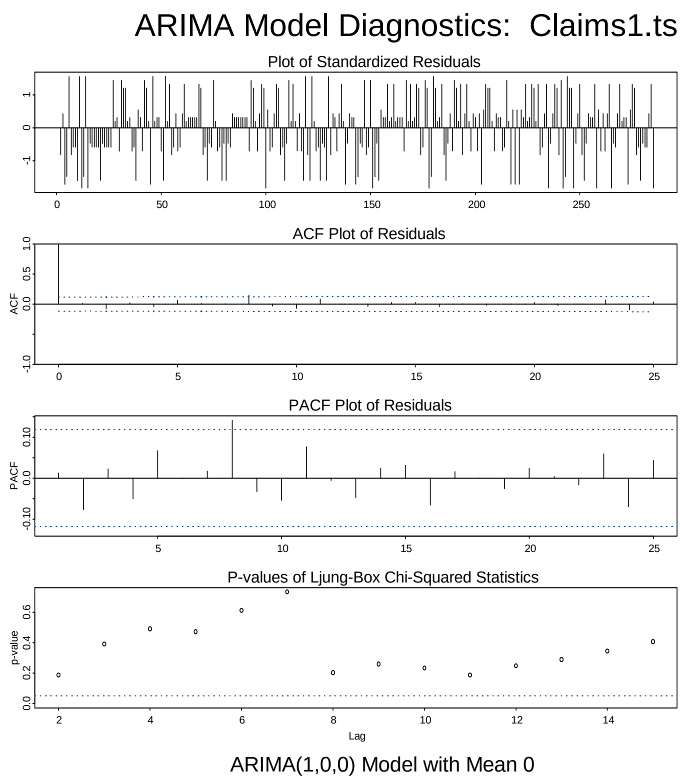

Fig 3 ARIMA Model Diagnostic for the Claims Portfolio

Reference

[1] Mees, A. (2000): Useful Lies: Dynamics from Data, Nonlinear and Nonstationary Signal Processing, edited by Fitzgerald, W.J, Smith, R.L, Walden, P.G and Young, P. , Cambridge University Press, U.K. 184-213.

[2] Scargle, J.D (1992): Deconvolution of chaotic and random time series, statistics in the environmental and earth sciences, edited by Walden, A.T, Guttorp, P., and Arnold, E. John Wiley and Sons, Inc. New York. 119-136.

[3] Glendinning, P. (1994): Stability, Instability and Chaos : an introduction to the theory of non linear differential equations,Cambridge University Press,USA

[4] Chatfield, C. (2004): The Analysis of Times Series An Introduction,6th ed, Chapman and Hall/CRC,Boca Raton, Florida, USA.

[5] Hobert, J. (2000): Hierarchical Models: A current computational perspective. Journal of American Statistical Association 95, 1312 – 1316.

[6] Aoki, M. (1996): New Approaches to Macroeconomic Modelling, Evolutionary Stochastic Dynamics, Multiple Equilibria and Externalities as Field Effects,Cambridge University Press, Cambridge.

[7] Berliner, L. M. (1996): Hierarchical Bayesian Time Series Models. Maximum Entropy and Bayesian Methods, Hanson, K. and Silver R. Eds, Kluwer Academic Pub., 15 – 22.

[8] Bleher, P. M. (1982): Construction of non-gaussian self similar random fields with hierarchical structure, Common. Math. Phys. 82, 557 – 78.
[9] Bleher, P. M. and Major, P. (1987): Critical Phenomena and Universal Expressions in Statistical Physics, On Dyson's Hierarchical Model, Annal of Probability, 15, 431 – 77.
[10] Waller, L.A., Carlin, B.P., and Gelfand, A.E (1997): Hierarchical spatio-temporal Mapping of Disease Rates, Journal of the American Statistical Association, 92,607-17.
[11] Barnerjee, S., Carlin, B. P., Gelfand A. E. (2004): Hierarchical modeling and Analysis for Spatial Data. Chapman and Hall, CRC, London.
[12] Wikle, C. K., Berliner, L. M. and Cressie, N (1998): Hierarchical Bayesian space time models, Journal of Environmental Ecological Statistics, 5, 117 – 154.
[13] Mugglin, A.S. and Carlin, B.P. (1998): Hierarchical Modelling in Geographic Information Systems: Population Interpolation over Incompatible Zones, Journal of Agriculture, Biological, and Environmental Statistics, 3, 117-130
[14] Wikle, C. K. (2003): Hierarchical Bayesian Methods for Predicting the Spread ofEcological processes, Ecology 84, 1382 – 1394.
[15] Haneuse, S. and Wakefield, J. (2007): Hierarchical models for combining Ecological and case – Control Data Journal of the International Biometric Society 63, No.1, 128 – 136.
[16] Hooten, M. B.; Wikle, C. K.; Dorazio, M. R. and Royle, J. A. (2007): Hierarchical spatio- temporal matrix models for characterizing invasions. Journal of International Biometric Society 63 (2) 558 – 567.
[17] Tomasz, R.; Hanspeter, S.; Volker, S. and Jozef, T. (1999). Stochastic processes for Insurance and Finance, John Wiley and Sons, Chichester, New York.
[18] Mikosch,T. (2008): Elementary Stochastic Calculus With Finance in View,Advanced Series on Statistical Science and Applied Probability,World Scientific, New Jersey