doi: 10.17221/10/2015-AGRICECON

Dynamic modelling in loss frequency and severity estimated: Evidence from the agricultural rice loss due to typhoons in Taiwan

Li-Hua Lai

Department of Risk Management and Insurance, National Kaohsiung First University of Science and Technology, Kaohsiung City, Taiwan

Abstract: The paper first adopt the BDS test to show that the BDS statistics of the time series of typhoons is a chaotic behaviour while the associated rice damage is random. The authors' investigations show that the time series of typhoons and rice damages are described by nonlinear. The result of the assessment shows that the model based on the AR(1)--GARCH(1,1) model is the best performing model in describing the rice loss severity due to typhoons and may have a chaotic behaviour if the variation of parameters is large enough. The best forecasting models in loss frequency with chaotic and severity predictions with random walk are superior to the best forecasting models in the current traditional or official estimated. This paper find that the fitting insurance price decision process by the loss cost charged in our method is different from the actuarial premium approaches of comparing evaluating effectiveness under the non-crop insurance program.

Keywords: BDS statistics, chaotic behaviour, logistic and exponential smooth transition autoregressive models

In natural disasters, many phenomena appear more unstable, and as the weather is a non-linear dynamics, it is natural to ask a question whether the subject-matter due to typhoons is a chaotic behaviour? The word "chaos" is a complete disorder or confusion, that is, chaos in science and engineering refers to an apparent lack of order in a system that nevertheless obeys certain laws and rules (Adams et al. 1998; Aggarwal et al. 2006). The natural disasters in the world have caused a great damage to crop production (Antonio and Beirlant 2007). Rice is the world's most important staple food crop and it feeds more than half of the world population (FAO UN 2013). The rice production is largely concentrated in Asia and the estimate of the world rice production about 652 million tons, 90% of which comes from Asia (Luo 1998; Chang 2002; Chatrath 2002; Yun 2003). According to the statistics of the Taiwan's Agriculture Yearbook, significant financial losses in agricultural sector in Taiwan are commonly caused from various natural events, such as typhoons, floods, droughts, insects, earthquakes, hails, and so on. Of all the weather phenomena, typhoons (or tropical cyclones) are the most catastrophic, not only for their fierceness but also the frequency of occurrence. The typhoon disasters in the Taiwan usually have caused a great damage to crop. For the recent two decades, typhoons hit the Taiwanese crop products with the value of US$ 66.9 billion. Over a 30-yrs period, Taiwan was hit by 3.3 typhoons per year in average and they brought abundant rainfalls and strong winds, leading to a severe damage to crops and great property losses. For example, a typhoon caused up to 60% of Taiwan's rice losses in 1971–2005. However, we may ask a question whether the rice damaged due to typhoons is also a chaotic behaviour in Taiwan? The most important problem is that how could we forecast the whether the rice damage?

In this paper, we first used the BDS test (Grassberger and Procaccia 1983) for the correlation dimension estimates, and the rescaled range analysis to investigate the problem. Note that the BDS test is based on a kind of the correlation integral which is used to examine the probability that a purely random system could have the same scaling properties as the rice damage indices. The test, which is useful when we have no idea about what sort of hidden structure to expect, locates the existence of a structure, the non-

Supported by the National Science Council of Taiwan (grant No. NSC 99-2410-H-327-027).

doi: 10.17221/10/2015-AGRICECON

linearity and the hidden patterns, which potentially renders the series susceptible to forecasting. We also calculate the Grassberger-Procaccia correlation dimension (CD) and the Foulkes' statistical correlation dimension (SCD) which, contrast to other statistics, calculate distances for all pairs of data points and could be applied to test the nonlinearity in stochastic processes (Hurst 1951,1957; Kim 2003). Finally, we use the rescaled range (R/S) method of nonlinear analysis to study the correlation properties in our data sets, in which we evaluate of the Hurst scaling exponent (Lai 2010). In this paper, we use the thirty-six years data of the rice damaged due to typhoons in Taiwan to investigate the problems. In the second part of this paper, we investigate the problem whether the typhoon frequencies and rice damages due to typhoons could be described by the nonlinear smooth transition autoregressive (STAR) models of Terasvirta (LeBaron 1996). We will try to use the logistic and exponential smooth transition autoregressive (LSTAR/ESTAR) and the AR(m)-GARCH(p,q) models to describe the data. It is employed to test for the existence of nonlinearities in damages due to typhoons and to identify the nature of those dynamics. We first determine the optimal lag length by the Akaike Information Criterion (AIC) and the Schwarz Information Criterion (SIC). We use the F-test statistic to determine the delay parameter in the STAR model. After determining the delay parameter therein, we attempt to make a choice between the LSTAR and ESTAR models which investigates the time series of typhoons described by the LSTAR or ESTAR while the associated rice damage is described by the linear or nonlinear model. This is consistent with the chaotic analysis that the rice damage shows a random behaviour while the time series of typhoons show a chaotic dynamics. In the relatives case, in addition to the nonlinear models that allow obtaining consistent estimators, we rethink that the chaotic dynamic information derived from the heuristic approach might be still useful and should not be disregarded on the natural disasters. In this paper, we also analyse the extended ARMA and AR-GARCH models to include the price of rice impact loss severity, and fitting the linear/nonlinear model on the loss frequency to estimate the loss cost.

The objective of this paper is to present and provide a new ideal and an empirical methodology for evaluating the agricultural loss due to the typhoon chaotic behaviour in an insurance pricing effectiveness framework for the actual analysis.

MATERIAL AND METHODS

Random and chaos model

The BDS statistics is a statistics quantity, the evaluation of which may be considered to be a test again a null hypothesis that a sequence of numbers is i.i.d (Dickey and Fuller 1979, 1981). The BDS statistics is defined as

$$W_{m,n} = \sqrt{n-m+1}\,\frac{C_{m,n}(e) - C_{1,n}(e)^m}{\sigma_{m,n}(e)} \tag{1}$$

Here, n is the same size, m is the embedding dimension and $C_{m,n}(e)$ is defined as

$$C_{m,n}(e) = \frac{2}{(n-m+1)(n-m)} \sum_{s=m}^{n} \sum_{t=s+1}^{n} \prod_{j=0}^{m-1} I_0(X_{s-j,}X_{t-j}) \tag{2}$$

where $I_e(X_{s-j}, X_{t-j}) = 1$ if $|X_{s-j}, X_{t-j}| < e$, otherwise it is zero. e is a parameter specifying the quantity which is proportional to the standard variation of the date sequence. The sequence is represented by the data set. Note that the time-adjusted monetary value of losses should be adequate for the validation of Equation (3) from the government reported.

$$X = \frac{X^0}{I} \times 100 \tag{3}$$

where:
X = adjusted rice loss
X^0 = unadjusted rice loss
I = agricultural Income price index (crops type)

The estimator $\sigma_{m,n}$ is defined as

$$\sigma_{m,n} = 4\left[k^m + (m-1)^2 C^{2m} - m^2 k C^{2m-2} + 2\sum_{j-1}^{m-1}\right] \tag{4}$$

Where $C = C_{1,n}(e)$ and

$$k = \frac{6}{n(n-1)(n-2)} \sum_{t=1}^{n} \sum_{s=t+1}^{n} \sum_{r=s+1}^{n} h_e(X_t, X_s X_r) \tag{5}$$

$$h_e(i,j,k) = \frac{1}{3}\big[I_e(i,j)I_e(j,k) + I_e(i,k)I_e(k,j) + {} + I_e(i,k)I_e(k,j) + \tag{6}$$

The estimator $\sigma_{m,n}$ is derived by Brock, Dechert and Scheinkman (1987). They showed that in the limit that $n \to \infty$

$$W_{m,n} \to N(0,1) \tag{7}$$

in which $N(0,1)$ is a normal distribution with a mean of zero and the standard derivation of one for any

doi: 10.17221/10/2015-AGRICECON

embedding dimension m. Therefore, from the calculated value of the BDS statistics $W_{m,n}$ we can, from the probability associated with a normal distribution, find the change that the sequence of the numbers could be produced by an i.i.d process.

Correlation dimension estimates

The correlation dimension defined by Grassberger Procaccia is given by

$$CD(m,n) = Log[C_{m,n}] / Log[e] \tag{8}$$

and the statistical correlation dimension (SCD) defined by Foulkes is given by

$$SCD(m,n) = Log[C_{m,n}](\ m, C_{1,n}) \tag{9}$$

in which $C_{m,n}$ is defined in (2). e is the standard derivation of the data sets.

Rescaled range analysis

In this section, we will use the rescaled range (R/S) method of the nonlinear analysis to study the correlation properties in our data sets. The method is related to the evaluation of the Hurst scaling exponent, H. Different values of the Hurst exponents correspond to different correlation properties. For the sequence represented by the data set $X(i)$, i = 1, ..., n, we calculate the running average $X(i)$ and the accumulated deviations from the average $X(l, i)$

$$\overline{X}(i) = \frac{1}{i}\sum_{k=1}^{i} X(k) \tag{10}$$

$$X(l,i) = \sum_{k=1}^{l} [X(k) - \overline{X}(i)] \tag{11}$$

The quantity called the range $R(i)$ of $X(l, i)$ and the standard deviation $S(n)$ are defined as follows:

$$R(i) = Max_{1\le l\le i}\, X(l, i) - Min_{1\le l\le i}\, X(l, i) \tag{12}$$

$$S(i) = \sqrt{\frac{1}{i}\sum_{k=1}^{i} (X(k) - \overline{X}(i)}\ ^{2} \tag{13}$$

The "rescaled range" is defined as a ratio between R and S, is R/S. The power law scaling according to Hurst is defined by

$$R(i) / S(i) \sim i^{H} \tag{14}$$

where H is the Hurst exponent. Note that a signal represents white noise (uncorrelated signal) then H = 0.5. The long-range correlations (memory) are in the chaotic system if H > 0.5. In practice, the scaling exponent H is evaluated from the $Log[R(i)/S(i)]$ vs. $Log(i)$ plot using the least square fit procedure.

Smooth transition autoregressive model

In this paper, we also try to use the logistic (LSTAR) and exponential (ESTAR) autoregressive model to describe the data. We first describe the model and then describe the method of how to fix the parameter in the model.

LSTAR and ESTAR models

The smooth transition autoregressive (STAR) model for the time series of data is defined as follows:

$$x_t = (\pi_{10} + \pi_1'\mathbf{x}_t) + (\pi_{20} + \pi_2'\mathbf{x}_t)F(x) + u_t \tag{15}$$

where $u_t \sim nid(0, \sigma^2)$, $\pi_j' = (\pi_{j1}, \dots, \pi_{jp})'$ are (p + 1) parameter vectors, and

$$x_t = (x_1, \dots, x_{t-p})' \tag{16}$$

is the vector of consisting of an intercept and the first $p$lags of x_t. $F(x)$ is the transition function. The two specifications generally considered are the logistic function,

$$F_L(x) = [1 + \exp(-\gamma_L(X_{t-d} - c_L))]^{-1} \tag{17}$$

and the exponential function

$$F_E(x) = [1 - \exp(-\gamma_E(X_{t-d} - c_E)^2)] \tag{18}$$

where γ_L and γ_E are transition parameters, c_L and c_E are the threshold values (location parameters). The d is the threshold lag (delay parameter). Equation (15) with the transition functions in Equations (17) and (18) yields the logistic STAR (LSTAR) and the exponential STAR (ESTAR) models, respectively.

Specification, estimation, and evaluation of models

The specification, estimation, and evaluation of the STAR models in this paper follow the procedures suggested and can be outlined as follows:

The maximum value of the lag p has to be determined from the data. The process x_t is referred to as an AR process of order p, AR (p), which can be written as

$$x_t = \beta_0 + \sum_{i=1}^{p} \beta_{1i}\, x_{t-1} + \varepsilon_t \tag{19}$$

doi: 10.17221/10/2015-AGRICECON

and the AR (1) process can be simply written as

$$x_t = \beta_0 + \beta_1 x_{t-1} + \varepsilon_t \tag{20}$$

For a stationary AR (p) process, the autocorrelation function is non-zero at all lags and should converge to zero geometrically. On the other hand, the partial autocorrelation function of an AR (p) process should cut to zero for all lags greater than p.

In order to make inferences on the time series, they must be stationary. However, most of the natural-disasters-loss-time series do not satisfy the requirement of stationarity, so that they have to be converted to stationary processes before modelling. Many test statistics have been developed to check whether the series contains the unit root or not. The most popular of them are the augmented Dickey-Fuller (ADF) test (1979, 1981) and the Phillips-Perron (PP) test (1988). Dickey and Fuller consider three different equations that can be used to test:

In the ADF test, we have to specify whether to include a constant, a constant and linear trend, or neither in the test regression,

$$\Delta x_t = r x_{t-1} + \sum_{j=1}^{p-1} \delta_i \Delta x_{t-j} + \varepsilon_t \qquad \text{(None)} \tag{21}$$

$$\Delta x_t = c + r x_{t-1} + \sum_{j=1}^{p-1} \delta_j \Delta x_{t-j} + \varepsilon_t \qquad \text{(Intercept)} \tag{22}$$

$$\Delta x_t = c + \omega t + r x_{t-1} + \sum_{j=1}^{p-1} \delta_j \Delta x_{t-j} + \varepsilon_t \qquad \text{(Trend and Intercept)} \tag{23}$$

$$H_0\text{: } r = 0 \text{ vs } H_1\text{: } r < 0 \tag{24}$$

With the ADF test, there is the problem of selection of the lag length. The first equation written above is a pure random walk model, the second equation adds an intercept or drift term, and the last one includes both a drift and a linear time trend so that it is possible to test whether the trend that the series exhibits is deterministic or stochastic. In all of the above equations, H_0: $r = 0$ is tested. If the null hypothesis is rejected, the sequence does not contain a unit root. The estimation technique is the Ordinary Least Squares (OLS). The calculated test statistic is compared by the critical values reported in the ADF tables. The AIC and SIC are used often but they have been found to select a low value of the lag length.

Phillips and Perron proposed a nonparametric method of controlling for the higher-order serial correlation in a series. The test regression for the PP test is the AR

$$x_t = c + \rho x_{t-1} + \varepsilon_t \tag{25}$$

while the ADF test corrects for the higher order serial correlation by adding lagged differenced terms on the right-hand side, the PP test makes a correction to the t-statistic of their r coefficient from the AR(1) regression to account for the serial correlation in ε_t. Alternatively, the structure of the AR process can be determined by using the model selection criteria. The most famous ones are the Akaike Information Criterion (AIC) and Schwartz Information Criterion (SIC):

$$\text{AIC} = T \ln(\text{residual sum of squares}) + 2n \tag{26}$$

$$\text{SIC} = T \ln(\text{residual sum of squares}) + \text{n} \ln(T) \tag{27}$$

where T is the number of usable observations, and n is the number of parameters to be estimated. In practice, several AR models are estimated, and the one with the smallest AIC or SIC is selected as the best model.

If linearity is rejected for more than one value of d, we choose the one for which the p-value of the test is the lowest. Testing the null hypothesis H_0: $\gamma = 0$ in (37) with either (39) or (40), assuming that x_{t-d} is stationary and eroded under H_0, is a non-standard testing problem since (37) is only identified under the alternative H_0: $\gamma \neq 0$. To solve the problem, Terasvirta (1994) followed the Davies' procedure, where an auxiliary regression with the unidentified values kept fixed in which transition function in (37) is replaced by its third-order Taylor approximation, to derive a Lagrange multiplier-type test that has an asymptotic χ^2-distribution. Therefore, the problem is solved by estimating the auxiliary regression as

$$\begin{aligned}\hat{e}_t = \beta_0 &+ \textstyle\sum_{i=1}^{p} \beta_{1i}\, x_{t-i} + \sum_{i=1}^{p} \beta_{2i}\, x_{t-i} x_{t-d} + \\ &+ \textstyle\sum_{i=1}^{p} \beta_{2i}\, x_{t-i} x_{t-d}^2 + \sum_{i=1}^{p} \beta_{3i}\, x_{t-i} x_{t-d}^3 + \\ &+ \textstyle\sum_{i=1}^{p} \beta_{4i}\, x_{t-i} x_{t-d}^3 + \nu_t\end{aligned} \tag{28}$$

where $\hat{e}_t$ are the residuals of the linear model, and then testing the null hypothesis H_0: $\beta_{2i} = \beta_{3i} = \beta_{4i} = 0$, $i = 1, \ldots, p$, against the alternative that is not valid. In practice, the Lagrange multiplier-type test of linearity is replaced by the F-test in order to improve the size and power of the test. Equation (28) is estimated across a range of values for d and $d \in D = \{1, 2, \ldots, d_{\max}\}$.

doi: 10.17221/10/2015-AGRICECON

When the delay parameter d is known, then the linearity test is identical to testing the joint restriction that all nonlinear terms are zero as in the following null hypothesis.

Considering the value of d as given and using a sequence of tests nested in (28) to choose between the ESTAR and LSTAR models. Such a sequence is:

$$H_{04}: \beta_{4i} = 0 \tag{29}$$

$$H_{03}: \beta_{3i} = 0 \mid \beta_{4i} = 0 \tag{30}$$

$$H_{02}: \beta_{2i} = 0 \mid \beta_{3i} = \beta_{4i} = 0 \tag{31}$$

It is based on the relationship between the parameters in (28) and (15) with either (16) or (17). We use the H version of this test following the above processes. If H_{04} is not rejected but H_{03} is rejected, we would adopt the ESTAR model. If both H_{04} and H_{03} are not rejected but H_{02} is rejected, then we select the LSTAR model. If the p-value of the test of H_{03} is the smallest of the test in the model selection sequence (the p-value of the test of H_{04}, H_{03}, H_{02}), select an ESTAR; if not, we choose a LSTAR model.

The estimation of linear and STAR models starts with including all lags from 1 to p, and the insignificant ones are dropped through the estimation procedure to conserve the degrees of freedom. The estimation of STAR models is carried out using the nonlinear least squares.

AR(m)-GARCH(p, q) MODEL

In this paper, we investigate the rice loss over time into AR(m)-GARCH(p, q) model which the nonlinear models allow for the shocks to fluctuation. A typical AR(m)-GARCH(p, q) model is one of the existing methods to estimate the loss severity for forecasting. This approach utilizes two models: one for the conditional mean specification (μ_t or x_t^S) and the other for the conditional variance specification (h_t) of the loss severity error series. The mean equation can be defined from the class of models under the AutoRegressive Moving Average (example: AR(m)), while the variance specification, usually follows the generalized AutoRegressive Conditional Heteroskedasticity (GARCH(p, q)) model. The major model utilized in this paper is the AR(m)-GARCH(p, q) model as follows:

The AR(m)-GARCH(p, q) model to be estimated:

$$x_t^S = c_0 + \sum_{i=1}^{m} c_i x_{t-i} + e_t \tag{32}$$

where $e_t = h_t \varepsilon_t$ (33)

and the conditional variance of x_t^S is GARCH(p, q)

$$h_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j e_{t-j}^2 + \sum_{k=1}^{q} \beta_k h_{t-k}^2 \tag{34}$$

where the model x_t^S follows an AR process conditioned on the information set at time $t-1$ of the loss severity due to the typhoon at t and h_t follows a GARCH process. Terms x_t^S and h_t are the expected loss severity and the conditional variance, respectively. The mean of the loss severity (x_t^S) is a function of a constant (c_0), the autoregressive model parameters (c_i) and the unconditional variance of e_t is finite, whereas its conditional variance h_t^2 evolves over time. The variance (h_t) is a function of an intercept (α_0), a shock from the prior period (α_j) and the variance from the last period (β_k). In practice, fixing of the premiums depends not only on the past loss experience, but also on the prospects related to the crop-specific factors on the area of cultivated land (a), the average harvest per ha (g) and the price of rice (r) besides being influenced by the lag period of rice damaged whether the losses fluctuation changes over the time and if so, whether it is predictable. Finally, the AR(m)-GARCH(p,q) model in the mean extension had been used to examine the relation between the loss frequency/severity and the loss cost in the actuarial situation.

Forecast performance

To evaluate the forecast performance, we employ the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), the Revised Theil Inequality Coefficient (RTIC) and the Mean Absolute Percentage Error (MAPE):

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left| x_t - \hat{x}_t \right| \tag{35}$$

$$RMSE = \left(\frac{\sum_{t=1}^{N}(x_t - \hat{x}_t)^2}{N} \right)^{0.5} \tag{36}$$

$$RTIC = \left(\frac{\sum_{t=1}^{N}(x_t - \hat{x}_t)^2}{\sum_{t=1}^{N} x_t^2} \right)^{0.5} \tag{37}$$

doi: 10.17221/10/2015-AGRICECON

Table 1. Rice losses summary statistics (Million USD)

Min	Max	Mean	Median	Standard Deviation	Skewness	Kurtosis
21	114.72	8.99	0.738	14.32	5.4	35.3

Source: Taiwan Agricultural Yearbook (1971–2005)

and

$$MAPE = \frac{1}{N}\sum_{t=1}^{T}\left|\frac{\hat{x}_t - x_t}{x_t}\right| \tag{38}$$

where N is the number of predictions, x_t is the actual observation and $\hat{x}_t$ is the forecasted value. The model that yields a smaller value in all such criteria signifies its superiority against other models. The overall out-of sample models is summarized in section 3.

Loss cost with a chaotic behaviour

The expected loss obviously is an important element that affects the retention ability for the non-insurance plan (relief program) and the insurance pricing. However, the loss cost analysis has been a popular method with many successful applications for the actual science, insurance financial and for the relationship between the risk measurement and decision making (e.g., pricing). In the collective models, we require the frequency of losses for the entire portfolio, (N), is the sum of the loss frequencies by $N = N_1 + \ldots + N_n$. The aggregate loss for the portfolio was modelled by $S = X_1 + \ldots + X_n$, each with losses like X_i where X is non-negative. Furthermore, the loss amounts are assumed to be independent of N. Here the expected aggregate loss $E(S)$ in the compound distribution is simply the product of the expected value of loss frequency $E(N)$ and of the severity $E(X)$. This paper introduces the principle of the aggregate loss cost $G(S)$ for the loss portfolio that is determined by

$$\begin{aligned} E(S) &= E[E[S|N]\,{}_{bestfitted} \\ &= E[X]_{bestfitted} \times E(N)_{bestfitted} \\ &= G(S) \end{aligned} \tag{39}$$

From Equation (39), there can also be obtained an approximation of the total loss cost over the individual risks X_1 by charging the following premium:

$$P(X_i) = E[X_1] \tag{40}$$

If the insurer charges a loss cost allocated $R(S)$ per ha on area of cultivated land (a) of the from Equation (39)

$$R(S) = G(S)/a \tag{41}$$

EMPIRICAL RESULT

Data collection

Substituting the forty-two years data in Taiwan into the above formula, we have the following results. We consider the data, which record the major natural disasters loss amounts of rice in Taiwan over the years 1971–2012 and is available at http://www.coa.gov.tw. According to the government reports, there were about 10 kinds of nature disasters loose. Since 59.62% losses are due to typhoon, we only consider this disaster loss in our analysis. By adjusting for the inflation, the rice loss data were made comparable through the years; all monetary magnitudes reported in this paper are in 2001 US Dollar (USD). The overall summary statistics for rice losses are shown in Table 1. First, the average and standard deviation of the loss amount caused by each typhoon was 8.99 and 14.32 million from 1971 to 2005. The data are considerably skewed to the right. The skewness coefficient is equal to 5.4 for typhoons in Taiwan, which could be used to perform the forecasting of the loss frequency-severity due to typhoons 2006–2012.

BDS test results

From the formulation in section II, we can use the thirty-six years data of the frequency and the rice

Table 2. BDS statistics of frequency and rice damage

e	Embedding dimension	BDS (Typhoon frequency)	BDS (Rice damage)
1	2	2.9945	1.5515
1	3	5.3918	1.2407
1	4	6.7869	0.6848
1	5	7.0014	0.8605

doi: 10.17221/10/2015-AGRICECON

Table 3. CD and SCD estimates frequency and the rice damage

m	CD (Frequency)	SCD (Frequency)	CD(Rice Damage)	SCD(Rice Damage)
1	1.3446	0.9399	0.0906	1.0663
2	1.8204	0.8480	0.1372	1.0835
3	2.2348	0.7807	0.1788	1.0593
4	2.6983	0.7541	0.2311	1.0953
5	3.2887	0.7659	0.1244	1.0476
6	3.7196	0.7425	0.1500	1.0824
7	4.2396	0.7405	0.1731	1.0928
8	4.7456	0.7368	0.1926	1.0812

damages due to typhoons in Taiwan to calculate the BDS statistics. The results are shown in Table 2. In Table 2, we see that the values of the BDS statistics of the frequency of typhoons are larger than 3 and the hypothesis of i.i.d can be rejected with 99% confidence. Thus the time series of typhoons shows a chaotic behaviour and the associated rice damage is random.

Grassberger and Procaccia test results

We now analyse the time-series data by the Grassberger-Procaccia correlation integral. The results are shown in Table 3 in which we present the estimated correlation dimensions (CD) and the statistical correlation dimension (SCD) for our time series. We have changed the embedding dimension m from 2 to 9. After calculating the CD and SCD, the estimates of the frequency of typhoons and the associated rice damage are reported in Table 3 which shows that the CD and SCD estimate of the frequency of typhoons does not have any tendency towards convergence. Thus the underlying system is chaotic. On the other hand, the CD and SCD estimates of the rice damage show an oscillating behaviour, which may be a random system.

Rescaled range Test

We also use the thirty-six years data of the frequency of typhoons in Taiwan to plot the diagram in Figure 1.

The slope is 0.655 which has a derivation from 0.5. This shows that the frequency of typhoons in Taiwan is a chaotic behaviour.

The calculations using the least square fit procedure to Figure 1 give the following value of the Hurst exponent

$$H = 0.6555 \text{ (Frequency of Typhoon)} \quad (42)$$

As $H > 0.5$, the frequency of typhoons in Taiwan has a derivation from the random, which, therefore, is a chaotic behaviour.

We next use the thirty-six years data of the rice damages due to typhoons in Taiwan to plot the diagram in Figure 2.

The slope is less than 0.5 which shows that the rice damaged due to typhoons in Taiwan is approximately a random dynamics.

The calculations using the least square fit procedure to Figure 2 give the following value of the Hurst exponent

$$H = 0.4564 \text{ (Rice Damaged Due to Typhoons)} \quad (43)$$

Thus the rice damaged due to typhoons in Taiwan is approximately a random dynamics.

Testing for non-stationary

In this subsection, we will determine the value of p lags. It is important to check whether a series is stationary or not before using it in a regression for forecasting the loss frequency and severity. The

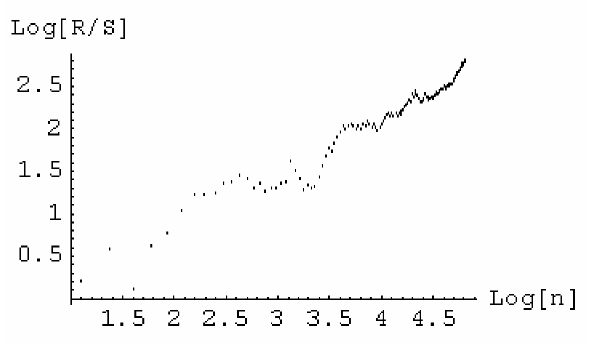

Figure 1. Diagram of $Log[R(i)/S(i)]$ vs. $Log(i)$ for the frequency of typhoons in Taiwan

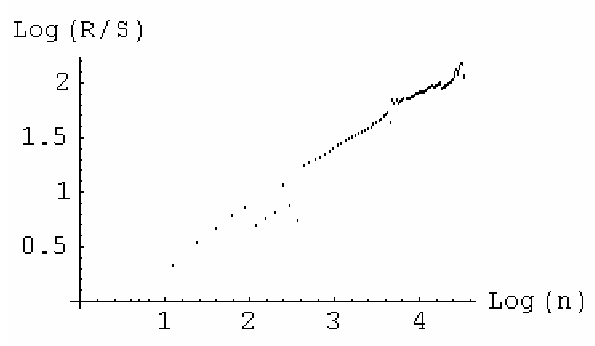

Figure 2. Diagram of $Log[R(i)/S(i)]$ vs. $Log(i)$ for the rice damaged due to typhoons in Taiwan

doi: 10.17221/10/2015-AGRICECON

formal method to test the stationary of a series is the unit root test. The results of the typhoon frequency and rice damage in Taiwan which shows that the Taiwan's typhoon frequency and the rice damage data series are not stationary at the level form but stationary after the first difference at the 1% and 5% significance levels in this case. A simple autoregressive (AR) model was fitted to the typhoon frequency and rice damage data. The lag structure was determined by using by the Akaike Information Criterion (AIC) and the Schwarz Information Criterion (SIC). The most significant lags were 1, 2, 3, 4 and 5. Here, p is the number of lags in the linear AR model. It shows that the model that minimizes the AIC or SIC is the AR (1) process in both the typhoon frequency and the rice damage data series on the ESTAR model in this case.

Determination of the delay parameter and testing for non-linearity

In this subsection we, will determine the value of the delay parameter d. To test the model's linearity, we consider a set of plausible values for the delay parameter (d), which ranges from 1 to 8. The optimum value of (d) is chosen based on the minimum P-values of the F-test statistic. Our calculation shows that the p-value (F-stat) for $d = 4$ is minimum (maximum) for the typhoon frequency data series. The rejection of linearity against the ESTAR is strongest when $d = 4$ (as p-value is minimum when $d = 4$). The decision rule is to select the ESTAR model if the p-value of the test of H_{03} is the smallest p-value of the test in the model selection sequence (the p-value of the test of H_{04}, H_{03}, H_{02}). We conclude that the ESTAR model is a more appropriate model for the typhoon frequency. Using this result the test the statistics for various hypotheses concludes that the ESTAR model is a suitable model with $d = 4$ for the typhoon frequency data series. In the same way, the ESTAR model is selected if the p-value of the test of H_{03} is the smallest of the three. We conclude that the ESTAR model is a suitable model with $d = 1$ for the rice damage data series.

ESTAR and linear models

The exponential smooth transition autoregressive (ESTAR) models which we find are the following:

(1) Typhoon frequency: $\hat{X}_t^F$ (Equation 44 and 45)

The root mean square error values of the ESTAR model and the linear model are 4.0573 and 4.1657, respectively. Thus the ESTAR model is more suitable to forecast the data than the linear model. This is consistent with the analysis that the typhoon frequency shows a chaotic behaviour.

(2) Rice damage: $\hat{X}_t^S$ (Equation 46 and 47)

The root mean square error values of the ESTAR model and the AR(1)-GARCH(1,1) model is 0.3144 and 0.2826, respectively. Thus the AR(1)-GARCH(1,1) model is more suitable to forecast the data than the ESTAR model. Note that the time series of the loss severity due to the typhoon may have a chaotic behaviour if the variation of parameters is large enough in the Equation (47). This is consistent with the analysis in the section smooth transition autoregressive model that the rice damage shows a random behaviour. In natural disasters, many phenomena appear more unstable, and it is here that the chaos theory has the most relevance. A chaos-based risk assessment model may capture the complexity of natural disaster with more verisimilitude than the traditional statistical analysis, when we are trying to understand the broad-based phenomena such as

$$\hat{X}_t^{F-nonlinear} = (1.457 - 0.0161x_{t-1}^F) \times [1 - \exp(-0.5737(x_{t-4}^F - 5.5069)^2)] \quad \text{Nonlinear model – ESTAR} \qquad (44)$$

$$\hat{X}_t^{F-linear} = 3.8158 + 0.2209x_{t-1}^F \quad \text{Linear model} \qquad (45)$$

$$\hat{X}_t^{S-Nonlinear} = 0.1567 - 0.6447x_{t-1}^S + (0.1567 - 0.6190x_{t-1}^S)[1 - \exp(-0.0313(x_{t-1}^S - 2.0271)^2] \quad \text{Nonlinear model – ESTAR} \qquad (46)$$

$$\hat{X}_t^{S-Nonlinear} = 0.00000048 - 0.0041822x_{t-1} + e_t \quad \text{Nonlinear model – AR(1)-GARCH(1,1)} \qquad (47)$$

where $e_t = h_t\varepsilon_t$ $\quad h_t = 0.0000003 + 0.4988654e_{t-1}^2 + 0.0000006h_{t-1}^2$

doi: 10.17221/10/2015-AGRICECON

the arrival of the typhoon issues and the associated rice damage, we should assume typhoons are dealing with a chaotic system and consider fitted methods aimed at understanding, rather than the quantitative ones aimed at the prediction and control. We see that the values of the BDS statistics of the frequency of typhoons in Taiwan are larger than 3 and the hypothesis of i.i.d can be rejected. Thus the time series of typhoons shows a chaotic behaviour and the associated rice damage is random. The results show that the loss distribution process is heavy-tailed, which implies that it is also non-normal. In Table 2, we have also estimated the statistical correlation dimensions (SCD). The main advantage of the SCD over the GP correlation dimension is that it gives statistically more reliable results for a small sample. Note that although the earlier research in economics has relied heavily on the correlation dimension (CD) estimation techniques for the chaos tests, the limitations of this approach are, by now, well known, for instance. The main point is that the correlation dimension methodology lacks an underlying statistically theory: it is basically a graphical analysis that requires typically very large data sets. Thus the correlation dimension estimations performed with small data sets might be indeed very misleading. We also use the correlation dimension estimates and the Hurst rescaled range analysis to confirm the properties. Our analysis finds that the typhoon is a chaotic behaviour while the associated rice damage is random. This may mean that the rice damage has other factors, such as the effect of temperature, the demand for water, resistance to environmental stresses, the space/time of each event. A simple autoregressive (AR) model was fitted to the typhoon frequency and rice damage data. Our results show that the Taiwan's typhoon frequency and the rice damage data series are not stationary at the level form but stationary after the first difference at the 1% and 5% significance levels. Using this result the test statistics for various hypotheses is shown in section 3.5, which concludes that the ESTAR model is a suitable model while with $d = 4$ for typhoon frequency data series and with $d = 1$ for the rice damage data series. These are consistent with the analysis that the typhoon frequency shows a chaotic behaviour and the rice damage shows a random behaviour. While the investigations are the special properties of our input data of Taiwan, the prescription of this paper could be applied to the other natural disasters of other countries.

ROBUSTNESS TEST

The robustness test for the forecasting performance on the loss frequency and severity

To test the robustness, we find the loss severity of rice damages clearly influenced by typhoons. Throughout the forecasting performance in the MAE, RMSE, RTIC and MAPE models on the frequency and severity of loss due to typhoons and the out-of sample 2006–2012 is summarized in Table 4. The result of the assessment shows that the nonlinear model with chaotic behaviour that yields a smaller value in all such forecasting performance in different models signifies the MAE or RTIC criteria its superiority against

Table 4. Forecasting performance on the loss frequency-severity due to Typhoons 2006–2012

Measures	Forecasting models performance			
	Nonlinear model with chaotic behaviour	Linear with random walk model		Traditional statistics with the distribution fitted
Loss Frequency				
MAE	0.0930	0.5012		0.2027
RMSE	0.2116	1.1547		0.5010
RTIC	0.3522	1.9261		0.8337
MAPE	0.0779	0.4507		0.1539
Loss Severity (million)				
		ARMA (1,1)	AR(1)-GARCH (1,1)	AR(0)-GARCH (0,0)
MAE	2.9164	2.8414	2.6244	3.2570
RMSE	11.8355	11.5741	10.5405	9.8432
RTIC	1.0563	1.0563	0.9407	0.8785
MAPE	1.2883	1.2883	4.8941	14.1173

doi: 10.17221/10/2015-AGRICECON

other models on the time series of typhoons. The loss severity of rice is described by the linear model while the model based on the AR(1)-GARCH (1,1)is the best performing model on the rice loss severity due to typhoons. The AR (1)-GARCH (1,1) can be combined with all or some of these models together to get more complex "mixed" models which will be sufficient and to capture the volatility clustering of the loss severity to avoid overfitting in the data. As for the damages size predictions, the result has a reasonable effect as the local government regulates the reporting of the loss severity and it might be dependent of the dynamics of the random walk and not explained well by the chaotic mechanism.

Estimating loss and forecasting performance in different models

This paper investigates a nonlinear collective risk model which is a feasible scheme for estimating the annual aggregate losses and it also focuses on using the chaos theory to fit the loss frequency and loss severity of distribution to rice damages due to typhoons which have been observed. It is shown that the annual frequency of rice damage caused by typhoons is best fitted well by the AR(1)-GARCH (1,1) model while the associated rice damage is described by the linear model. The pure risk premium should in theory match the expected loss expenditure as Equation (39). This paper first proposes that the expected annual aggregate loss with the chaotic behaviour obviously is one of the important elements that affects the retention ability for the non-insurance plan (relief program) or the insurance pricing. In the cases, when the average loss is under per ha of the damaged area, it is simply calculated from Equation (40). When the considerations are extended to cope with the insurance market, as Equation (41) or in Table 5, problems arise in accounting for the ensuing effects.

Note that this paper cannot conclude, however, that a market premium will be affected by the load premium (such as the safety loading, expenses loading, profit loading, contingencies loading and government regulated). We know that there are numerous, variable risks that influence agriculture (Vávrová 2005) and the relationship between natural disasters and crop yields should be a complex issue. However, various uncertainties and market factors may have consequences which it is an important issue to know about and to take into account.

Table 5. The aggregate loss cost and per ha loss cost allocated in the typhoon case 2006–2012 (USD)

	Estimated value AR(1)-GARCH(1,1)	Actual value
Aggregate loss cost $G(S)$	15 884 326	15 610 969
Loss cost allocated $G(S)$ per ha	19.3866	19.0530

CONCLUSION

A nonlinear mixing collective risk model is applicable to estimate the annual aggregate rice losses. Some typical risk measures of the chaotic behaviour distribution, such as the expected annual aggregated loss and the moment generating functions are also estimated. These quintile measures could provide useful information for the Council of Agriculture to check the applicable risk of the financing regulations and an adjustment of the natural disaster relief budget plan or the insurance premium. In this paper, we first use the statistical analysis to show that the BDS statistics of the time series of typhoons is a chaotic behaviour while the associated rice damage is random. We also use the correlation dimension estimates and the Hurst rescaled range analysis to confirm the properties therein. Note that although the weather is a non-linear dynamics and thus the frequency of typhoon is a chaos, the rice damages due to typhoons do not show a chaotic behaviour except that the variation of parameters is large enough.

We next consider the two families of nonlinear autoregressive models, the logistic (LSTAR) and the exponential (ESTAR) autoregressive model to describe the data. We first determine the delay parameter therein and attempt to make a choice between the LSTAR and ESTAR. Our investigations have shown that the time series of typhoons are described by the ESTAR, while the associated rice damage is described by the AR(1)-GARCH(1,1) model. This is consistent with the analysis in the previous section that the rice damage shows a random behaviour while the time series of typhoons is a chaotic dynamics. Our investigations show that as a robustness check, this paper also estimated that the forecast value errors of nonlinear are significant in comparison to the linear model under the loss frequency estimated, and the forecast value errors of four models forecasts are statistically significantly under loss severity estimated on the crop damage. This paper discusses the effect

doi: 10.17221/10/2015-AGRICECON

of typhoon on the rice yields. The result of the assessment shows that the model based on the AR(1)-GARCH(1,1) model is the best performing model on the rice loss severity due to typhoons. The insurance pricing is unavoidably, but there is also uncertainty due to the fact that the environment such as the cultivated land under crops, the average harvest per ha and the price of rice, etc. are changing all the time. The properties we found are useful in the loss cost or risk retention estimates about the crop damages due to typhoons in Taiwan and our algorithm may be applied to other disasters and other countries.

REFERENCES

Adams R.M., Hurd B.H., Lenhart S., Leary N. (1998): Effects of global climate change on agriculture: an interpretative review. Climate Research, 11: 19–30.

Aggarwal P.K., Kalra N., Chander S.H., Pathak H. (2006): Info crop: a dynamic simulation model for the assessment of crop yields losses due to pests, and environmental impact of agro-ecosystems in tropical environments. Agricultural Systems, 89: 1–25.

Antonio K., Beirlant J (2007): Actuarial statistics with generalized linear mixed models. Insurance: Mathematics and Economic, 40: 58–76.

Brock W., Dechert D., Scheinkman J. (1987): A test for independence based on the correlation dimension. Social Science Research WP No. 8762, University of Wisconsin-Madison. Economic Review, 15: 197–235.

Chang C.C. (2002): The potential impact of climate change on Taiwan's agriculture. Agricultural Economics, 27: 51–64.

Chatrath A., Adrangi B., Dhanda K.K. (2002): Are commodity prices chaotic? Agricultural Economics, 27: 123–137.

Dickey D.A., Fuller W.A. (1979): Distribution of the estimations for autoregressive time series with a unit root. Journal of the American Statistic Association, 74: 427–431.

Dickey D.A., Fuller W.A. (1981): Likelihood ratio statistics for autoregressive time series with a unit root. Journal of the American Statistic Association, 49: 1057–1072.

Food and Agriculture Organization of the United Nations (2013): Available at http://www.fao.org (accessed 2 August 2015).

Grassberger P., Procaccia I. (1983): Measuring the strangeness of strange attractors. Physica D9: 189–208.

Hurst H.E. (1951): Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers, 116: 770–808.

Hurst H.E. (1957): A suggested statistical model of some time series which occur in nature. Natural, 180: 494.

Kim H.S., Kang D.S., Kim J.H. (2003): The BDS statistic and residual test. Stochastic Environmental Research and Risk Assessment, 17: 104–115.

Lai L.H. (2010): Risk estimate of rice damaged due to flood. Nonlinear Analysis: Real World Applications, 11: 1243–1248.

LeBaron B. (1996): A test for independence based on correlation dimension. Economic Review, 15: 197–235.

Luo Y., Teng P.S., Fabellar N.G., TeBeest D.O. (1998): Risk analysis of yield losses caused by rice leaf blast associated with temperature changes above and below for five Asian countries. Agricultural Ecosystems Environment, 68: 197–205.

Phillips P.C., Perron, P. (1988): Testing for a unit root in time series regression. Biometrica, 75: 335–346.

Vávrová E. (2005): The Czech agricultural insurance market and a prediction of its development in the context of the European Union. Agricultural Economics – Czech, 51: 531–538.

Yun J.I. (2003): Predicting regional rice production in South Korea using spatial data and crop-growth modeling. Agricultural Systems, 77: 23–38.

Received: 10th January 2015
Accepted: 24th April 2015

Contact address:

Li-Hua Lai, National Kaohsiung First University of Science and Technology, Department of Risk Management and Insurance, No.2, Jhuoyue Rd., Nanzih District, Kaohsiung City 811, Taiwan
e-mail: lihua@nkfust.edu.tw