ESTIMATING THE PADDY SPIKE YIELD USING FRACTAL DIMENSION

Hongju Gong [1], Hua Li [1], Haiming Yu [1], Changying Ji [1,*]

[1] College of Engineering ,Nanjing Agricultural University, Nanjing, Jiangsu Province, P. R. China 210031

[*] Corresponding author, Address: College of Engineering ,Nanjing Agricultural University, Nanjing 210031, Jiangsu Province, P. R. China, Tel: +86-25-58606570, Fax: +86-25-58606699, Email:gonghongju@gmail.com

Abstract: Paddy is one of major crops grown in China. Fractal dimension is an important parameter used to describe geometrical characteristics of many natural objects. The objective of this study is to investigate mathematical relationships between the fractal dimension and the yield of the paddy spike based on the fractal theory and the image processing technologies. The color images of paddy spikes were taken with a digital camera, then they were transacted to the gray using image processing. Fractal dimensions of the spikes were computed using the Box-Counting Method based on the gray images. The results indicate that there is a significant linear relationship between the fractal dimension and the yield.

Keywords: Fractal Dimension (FD), yield, paddy spike, Box-Counting Method (BCM)

1. INTRODUCTION

The fractal theory, brought forward by Mandelbrot who is considered as the father of fractal, has found wide applications in areas ranging from material science, geology, power technology to micro-electronic(Biswas et al., 1998; Valdez et al., 1998), as well as in areas of Agricultural Engineering (Li Yan et al., 2000; Liu Taoju et al., 2002; Zhang Shirong et al., 2002; Ren Xianzhong et al., 2005). Simple objects can be described by ideal shape primitives, such as cubes, cones, and cylinders. But most of the natural objects are so complex and erratic that they cannot be described in terms of

simple primitives (Kenneth, 1991). On the other hand, the essence of fractal is self-similarity, and the concept of self-similarity seems to play an important role in the description of nature.

Paddy is one of major crops grown in China. LiuTaoju (Liu Taoju et al., 2002) considered that paddy spikes have the self-similarity, which is the primary characteristics of the fractal geometry and they base it simulating the rice morphology with fractal theory. Currently, the further researches on fractal characteristic of paddy spike, such as Fractal Dimension (FD) and the relationship between FD and yield, have been not reported.

In this paper, fractal analyses of paddy spikes and measurement of FD of paddy spikes based on the Box-Counting Method are presented; The correlativity between FD and the yields of paddy spikes and a mathematical model for estimating the yields using FD are also discussed.

2.   BOX-COUNTING METHOD(BCM)

Fractal Dimension (FD) can be defined in many ways, in which, the Hausdorff Dimension, the Self-similarity Dimension and the Box-counting Dimension are the leading definition. By comparison, the Box-counting Dimension is easier in mathematical computation and measurement, therefore it was widely used (Chaudhuri et al., 1995; Biswas et al., 1998; Cheng Kunjie, 2005; Ren Xianzhong et al., 2005; Wan Yaonan et al., 2006; Zhang Tao et al., 2006). In the research, it is also used to measure the FD of paddy spikes.

Box-counting Dimension can be explained as follows. If X is a bounded set in Euclidean n-space. And it is the union of $N_e(X)$ distinct (no overlapping) copies of itself, each of which is similar to X and scaled down by a ratio e . Then fractal dimension D of X can be derived from the following equation when the limit in the equation exists,

$$D = \lim_{e \to 0} \frac{\log N_e(X)}{\log(1/e)} \qquad (1)$$

It is difficult to compute the parameter D directly using Eq.(1). An approximate method was used extensively in many literatures. Just as its name implies, our work is counting the numbers of the boxes. There is a digital image with $M \times M$ pixels, in which the target information is contained and it could be divided into boxes with $r \times r$ pixels, where $M/2 ? r 1$, and r is an integral number. Then we count the boxes in which portions of the target information are contained. Let $N(r)$ be the total number of countable boxes, an array of $r - N(r)$ data can be obtained by varying the values of r . If the data satisfies following equation,

$$N(r) = (\frac{1}{r})^D \tag{2}$$

It indicates that the target information possesses fractal characteristics.

In practice，the Eq.(2) could be transformed into a logarithmic form, i.e.,

$$log(N(r)) = D\,log(\frac{1}{r}) \tag{3}$$

With that, according to the array of r - $N(r)$ data obtained above, a plot of $log(1/r)$ versus $log(N(r))$ could be drawn in a log-log grid. Afterwards, a fitted straight line can be got with the method of least squares. Then the slope of the line is the Box-Counting Dimension.

3.   MATERIAL AND METHOD

3.1   Material and Equipments

Digital Camera-Olympus U400D(DC), Constant Illuminating Source Box(CISB), Electronic Weighing Scale-JY12001(EWS) are the Hardware and equipments used in the experiments. The Software used are VC++ and MATLAB6.0.

The paddy species is Sujing-45, it was grown in the Jiangpu Experimental Farm of Nanjing Agricultural University ,Jiangsu Province of China which is located at latitude of $32° 3' 4.96"$ N，and longitude of $118° 36' 38.78"$ E. 27 paddy spikes picked up randomly from the field with visible difference in yields are weighed with the EWS and pictured in the CISB with a black velvet background using the DC. When picturing, each spike was positioned just as it was in the original field. Each image is of sides 512×512 pixels(Fig. 2).

3.2   Digital Image Processing

Fig.1 Source image of paddy   *Fig.2* Grayscale image of paddy

The source color images (Fig.1) were converted to grayscale images (Fig.2) according to the following equation (Yang Shuyin, 2003)

$$R_1 = G_1 = B_1 = Max(R, G, B)$$

(4)

Where R,G,B represent respectively the gray values for each pixel of the origin color image, R1,G1,B1 represent the gray values for each pixel of grayscale image. Afterwards, using a threshold value 100, grayscale images were binarized into binary images(Fig3).
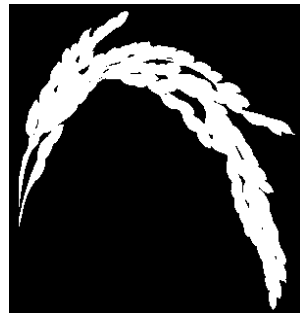
Fig.3 Binary images of paddy spike

3.3   Fractal Dimension Measurement

In the binary images such as the Fig.3, the part of paddy spike is white (where pixel value equals 1), whereas the background is black (where pixel value equals 0).

Divide the binary image into boxes sides $r' \times r$, avoiding the singularity where $1 < r \leq 216$ and r is an integer.

Let $r = 2,3,4,L,256$, and then count the number of grids where pixel value equals 1 and note it as N_r. With the arrays of $r - N_r$, a log-log grid of $log(1/r)$ versus $log(N(r))$ could be constructed. So the Box-Counting Dimension of the paddy spike, i.e., the slope of the fitted straight line could be obtained, which is noted as DBC.

4.   EXPERIMENTAL RESULTS AND DISCUSSION

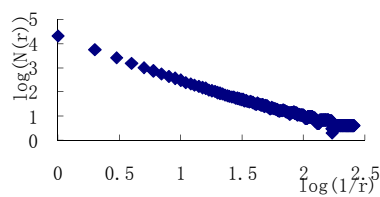A scatter plot of $log(1/r)$ versus $log(N(r))$ of the image (for example the sample 1) is shown in Fig.4.

Fig.4 Relationship between $log(1/r)$ and $log(N(r))$

According to the definition of the Box-counting Dimension, if there is a linear interval in the scatter plot, it means that the target information bears fractal characteristic. It could be seen from Fig.4 that when r varies from 2 to 103 pixels, there do exist distinct linear relationship between $log(1/r)$ and $log(N(r))$. Fitting the scatter plots in the region from $r = 2$ to $r = 103$, the slop of fitting straight line, so the DBC of the paddy spike, can be calculated. The results for every sample are presented in Table 1.

Table 1. D$_{BC}$ and yields of paddy spikes

sample	1	2	3	4	5	6	7
yield(g)	3.55	3.96	3.97	4.17	4.46	4.5	4.76
D$_{BC}$	1.5465	1.5221	1.4707	1.4478	1.5317	1.5005	1.5191
sample	8	9	10	11	12	13	14
yield(g)	5.1	5.12	5.46	5.54	5.72	5.8	5.83
D$_{BC}$	1.5165	1.521	1.5104	1.5122	1.5716	1.511	1.5944
sample	15	16	17	18	19	20	21
yield(g)	5.84	5.86	6.08	6.11	6.28	6.31	6.45
D$_{BC}$	1.5487	1.5582	1.5102	1.565	1.5421	1.601	1.561
sample	22	23	24	25	26	27	
yield(g)	6.67	6.79	6.83	6.88	7.83	8.5	
D$_{BC}$	1.5481	1.5674	1.5798	1.5539	1.6105	1.5852	

A regression analysis between the FDs and the yields of paddy spikes was performed. By means of calling the Regression (Y, X, a) (He Runbing, 2001)function of MATLAB 6.0. fitting equations can be easily obtained. when the confidence level being $1 - a = 99.9\%$, the fitting equation can be described as follows,

$$y = 21.4017 D_{BC} - 27.2623 \tag{5}$$

with，$F = 23.7419, R^2 = 0.4871, p = 0.0001$ Since $F > F_{0.001}(1,100), p > a$, It can be concluded that there is a distinct linear relationship between the FDs and the yields of corresponding paddy spikes, at the confidence level of 99.9%.

With rcoplot(r,rint) (He Runbing, 2001) function of MATLAB6.0, the residuals of regression were plotted as shown in Fig.5. From the figure, it
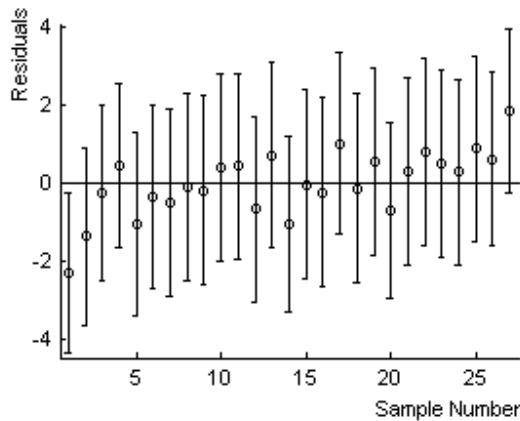

Fig.5 Residuals of regression

can be seen that the FD of sample 1 is an outlier in the regression. After getting rid of the outlier, the renew regression equation was founded as follows

$$y = 21.7397 D_{BC} - 27.6951 \tag{6}$$

With $R^2 = 0.5798$, $F = 33.1145$, $p = 0.0000$. The renew Residuals were plotted in Fig.6. It is seen that the coefficients of the regression equation are equable, but the value of R^2 and F were increased clearly,
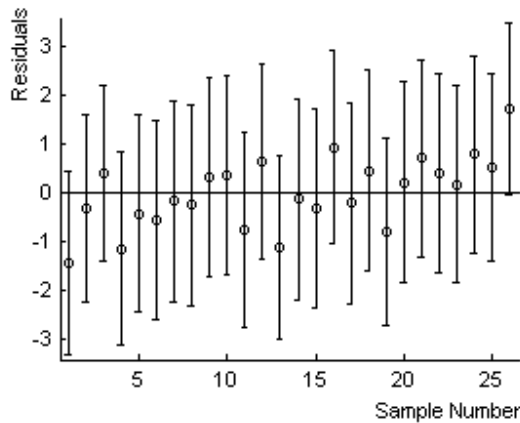
Fig6. Renew Residuals of regression

and the value of p was diminished to zero, and the fitting results were optimized clearly.

5. MODEL CHECKING

To check the adequacy of the mathematical model which describes the relationship between FD and the yield of each paddy spikes, 20 spikes of same species were collected as a checking specimen. Calculated values getting from the model and real values were plotted in Fig.7, and the error distribution of the model checking was shown in Fig.8.
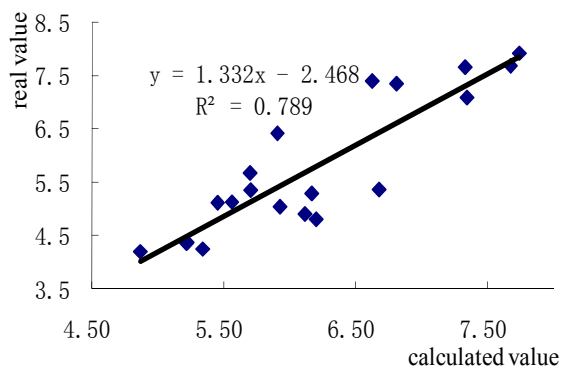

$y = 1.332x - 2.468$
$R^2 = 0.789$
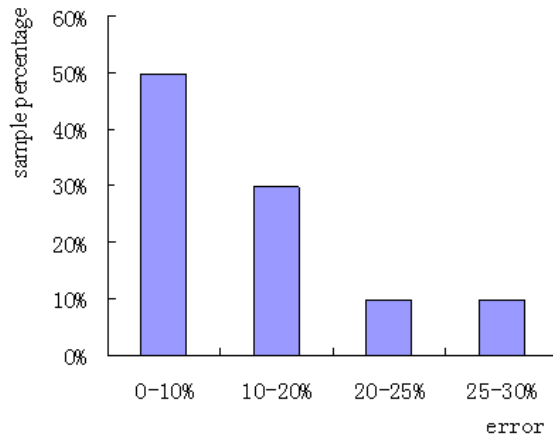
Fig.7 Model checking

Fig.8 Error Distribution of Samples

As shown in Fig.7, the correlation coefficient $R^2 = 0.7897 > (r_{0.01}^{19})^2 = 0.4425$, the conclusion can be made that the correlation between the calculated values and the real values is significant at $a = 0.01$. It is also shown in the plot of error distribution that 80% of errors in the specimen are limited in 20%. Further investigation shows that when the yield of all, single spikes in specimen is higher than 5g, the maximum error is 24.57%, and 93% of errors in the specimen are limited in 20%.

6. CONCLUSION

From the investigation on fractal properties of paddy spikes and the analyses of the relationship between the yields of paddy spikes and their fractal dimensions, the following conclusions can be drawn.

(1)The paddy spikes bear fractal characteristic.

(2)The correlation relationship between the FDs and the spike yields is significant and it can be described as:

$$y = 21.7397 D_{BC} - 27.6951 \tag{7}$$

(3)Although the accuracy of the model is not very satisfactory, the reserch on the FDs and estimating yields of paddy spikes should be helpful for the further study in relevant domain.

ACKNOWLEDGEMENTS

REFERENCES

Biswas, M. K. and T. Ghose, et al. (1998). "Fractal dimension estimation for texture images:A parallel approach." Pattern Recognition Letters 19 : 309-313.

Chaudhuri, B. B. and N. Sarkar (1995). "Texture Segmentation Using Fractal Dimension." Transactions on Pattern Analysis and Machine Intelligence 17 (1): 72-77.

Kenneth, F. (1991). Fractal Geometry-Mathematical Foundations and Applications. Shenyang, Shenyang University Publishing Company.

kunjie, C. (2005). Study on Beef Automatic Grading Based on Fractal Dimension & Machine Vision. Nanjing, Nanjing Agricultural University. Doctor.

Runbing, H. (2001). Calculation and Application in Engineering using MATLAB6.0. Chongqin, Chongqin Publishing Company.

Shi-rong, Z. and G. Guosu (2002). "Fractal analysis of patch patterns on plant diseases." Plant Protection 28 (6): 9-13.

shuyin, Y. (2003). Digital Image Processing Using VC++. Beijin, TsingHua publishing company.

Tao, Z. and Y. Zhibiao, et al. (2006). "Improved Extracting Algorithm of Fractal Dimension of Remote Sensing Image." Journal of Ordnance Engineering College 18 (5): 61-65.

Taoju, L. and T. Jianjun, et al. (2002). "A Study on the Fractal Characters and the Visual Simulation of Rice Morphology." Acta Agricultural University Jiangxiensis 24 (5): 583-586.

Valdez-Cepeda, R. D. and E. Olivares-Sanz (1998). "Fractal analysis of Mexico's annual mean yields of maize,bean,wheat and rice." Field Crops Research 59 : 53-62.

Xianzhong, R. and W. Chun, et al. (2005). "Study on Fractional Characteristics of Wheat Kernel Shape Parameter and Its Application on Agricultural Engineering." Transactions of the Chinese Society for Agricultural Machinery 36 (10): 85-87.

Yan, L. and P. Jiaxiong (2000). "The Target Segmentation and Detection Based on Fractal Dimension Feature." Journal of Huazhong University of Science and Technology(Nature Science) 28 (8): 1-5.

Yaonan, W. and W. Shaoyuan, et al. (2006). "Image Texture Analysis Based on Fractal Dimension." Journal of Hunan University(Natural Sciences) 33 (5): 67-72.