International journal of Rural Development, Environment and Health Research(IJREH) [Vol-2, Issue-4, Jul-Aug, 2018]
https://dx.doi.org/10.22161/ijreh.2.4.3 ISSN: 2456-8678

Evaporation and Production Efficiency Modelling Using Fuzzy Linear Recurrence

Saeid Eslamian[1], Fatemeh Sorousha[1], Morteza Soltani[2], Kaveh Ostad-Ali-Askari[3*], Shahide Dehghan[4], Mohsen Ghane[5], , Vijay P. Singh[6], Nicolas R. Dalezios[7]

[1]Department of Water Engineering, Isfahan University of Technology, Isfahan, Iran.
[2]Department of Architectural Engineering, Shahinshahr Branch, Islamic Azad University, Shahinshahr, Iran
[3*]Department of Civil Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran. **Email Corresponding author: Koa.askari@khuisf.ac.ir**
[4]Department of Geography, Najafabad Branch, Islamic Azad University, Najafabad, Iran
[5]Civil Engineering Department, South Tehran Branch, Islamic Azad University, Tehran, Iran
[6]Department of Biological and Agricultural Engineering & Zachry Department of Civil Engineering, Texas A and M University, 321 Scoates Hall, 2117 TAMU, College Station, Texas 77843-2117, U.S.A.
[7]Laboratory of Hydrology, Department of Civil Engineering, University of Thessaly, Volos, Greece & Department of Natural Resources Development and Agricultural Engineering, Agricultural University of Athens, Athens, Greece.

***Corresponding Author:**
Dr. Kaveh Ostad-Ali-Askari, Department of Civil Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran. Email: **Koa.askari@khuisf.ac.ir**
Address: Islamic Azad University, Khorasgan Branch, University Blvd, Arqavanieh, Jey Street, Isfahan, Iran. P.O. Box:81595-158, Phone: (+98)-31-35354001-9, Fax: (+98)-31-35354060, Mobile: +98912247143.

***Abstract*— *The relationship between crop production and amount of evapotranspiration is very important to agronomists, engineers, economists, and water resources planners. These relationships are often determined using classical least square regression (LSR). However, one needs high amount of samples to determine probability distribution function. Linear regression also requires so many measurements to obtain the valid estimates of crop production function coefficients. In addition, deriving ET-yield regression for each crop and each district is usually expensive, since lysimetric experiments should be repeated for several years for each crop. The object of this study is to introduce a fuzzy linear regression as an alternative approach to statistical regression analysis in determining coefficients of ET- yield relations for each crop and each district with minimum data. The application of possibilistic regression has been examined with a case study. Two data set for winter wheat in Loss Plateau of China and North China Plain have been used. The current finding shows capability of possibilistic regression in estimation of crop yield in data shortage conditions.*

Keywords*— *Data shortage; evapotranspiration; fuzzy regression; grain yield; production function.

I. INTRODUCTION

Water shortage is the major constraint to agricultural production. The relationships between crop yield and water use have been a major focus of agricultural research in the arid and semi-arid regions (Zhang and Oweis, 1999). Water management is very important in these regions. Many researchers have studied the effect of deficit irrigation on crop production as a solution (Zhang et al., 1999 and Kang et al., 2002).

In agriculture water management, the adequate representation of production or crop yield functions is crucial for modeling purposes in environmental economic analyses. The discussion and estimation of different functional forms have therefore gained much attention in agronomic and agricultural economics literature (Finger and Hediger, 2007). Various functional forms have been considered so far, but less attention has been given to the estimation techniques. In general, crop yield is estimated by least square regression. Classical linear or non-linear regression assumes that the measurement errors are normally distributed and independent of each other. Since one needs so many samples to determine a probability distribution, linear or nonlinear regression require at least 8

https://dx.doi.org/10.22161/ijreh.2.4.3 ISSN: 2456-8678

to 30 measurements or observations to obtain valid estimate of parameters (Eslamian et al. 2012, Cheng Si and Bodhinayake, 2005).

Measurement of some parameters such as evapotranspiration in yield function is expensive and time consuming. Therefore, it is difficult and sometimes impossible to obtain a simple yield function for regions with same climate. Moreover, evapotranspiration determination is subjected to different kind of uncertainties. These arise from measurement errors due to human and assumptions on deep percolation and uniformity of soil distribution. In these circumstances, classical regression may not give valid estimation for yield. In particular, confidence interval estimated with a few data points is very wide and may not provide suitable information that is usual for predictive purpose (Eslamian et al. 2001, Cheng Si and Bodhinayake, 2005).

Fuzzy sets theory can quantitatively deal with uncertainty in experimental data or ambiguity in human perception, and so it has been applied to various fields in which uncertainty and/or ambiguity have a serious influence. The theory does not need strict assumptions of probability functions as in the statistical methods, such as the normal distribution described above, and it can deal with the uncertainty more easily and more flexibly (Shimosaka et al., 1996). The objective of this study is to investigate whether fuzzy linear regression (Tanaka et al., 1982) would predict crop production and to provide a method for yield forecasting with less observation than least square regression.

II. THEORY

Water use-yield relationship:

Crops consume water in the process of transpiration, and water evaporates from the soil. These processes are defined collectively as evapotranspiration (Thornyhwaite, 1948). The relationship between crop production and the amount of water applied to crop is important. This importance is currently considered due to declining in water resources and competition among users.

Crop production models with resource and management inputs have been widely used, particularly by agricultural economist, and called production function (vaux 1983, Ostad-Ali-Askari et al. 2015). Hanks et al. (1969) reported that dry matter is linearly related to evapotranspiration for wheat, millet, oat and grain sorghum in both lysimetric and field plots. Cole and Mathews (1923) and Mathews and Brown (1938) investigated grain yield for winter wheat and sorghum. They used linear regression techniques to evaluate the yield- evapotranspiration as follows:

$$Y = a\,ET + b \tag{1}$$

Where Y is grain yield (kg ha^{-1}), ET is the growing season evapotranspiration (mm) and a (kg ha^{-1} mm^{-1}) and b (kg ha^{-1}), regression coefficients.

ET is usually calculated using the soil water balance equation for growing season as given:

$$ET = \Delta W + I + P + Sg - D - R_f \tag{2}$$

Where ET is actual evapotranspiration, ΔW the change in soil water storage between two soil moisture content measurements, I the irrigation, P the rainfall, S_g the capillary rise from the lower soil layer to the crop root zone, D the deep percolation from the crop root zone, and R_f is the surface runoff (Kang et al. 2002). When the groundwater table is lower than 4 m below the ground surface, S_g is usually negligible (Zhang et al., 1999). It is usually assumed that soil infiltration rate is larger than rainfall and irrigation density.

Some studies had shown that the empirical relation between crop yield and seasonal evapotranspiration can take different forms and that the empirical coefficients in the relations vary with climate, crop type and variety, irrigation method, soil texture, fertilizer and tillage methods. These differences relate to regional variability in environment and agronomic practices, Information specific to a region is needed to define production function (Eslamian et al. 2015, Kang et al., 2002, Ostad-Ali-Askari et al. 2016). So, derivation of production functions for each region would be expensive and obtaining adequate data for linear regression would be difficult.

Fuzzy linear regression method

Fuzzy regression analysis was first proposed by Tanaka et al. (1982). Since membership functions of fuzzy sets are often described as possibility distributions, this approach is usually called possibilistic regression analysis (Tanaka et al., 1982). The basic concept of fuzzy theory of fuzzy regression is that the residuals between estimators and observations are not produced by measurement errors, but rather by the parameter uncertainty in the model, and the possibility distribution is used to deal with real observations (Tseng et al., 1999, Eslamian et al. 2016). This method provides the means by which the goodness of a relationship between two variables, y and x, may be evaluated on the basis of a small sample size. In this approach, the regression coefficients are assumed to be fuzzy number (Sahin and Hall, 1996, Ostad-Ali-Askari et al. 2017).

The fuzzy linear regression (FLR) model can be expressed as:

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \cdots + \tilde{A}_n x_{in} = \tilde{A}x_i \quad (3)$$

Where $xi = [x_0, x_{i1}, \ldots, x_{in}]$ is a vector of independent variables in the ith data $i = 1, \ldots, m$; $\tilde{A} = [\tilde{A}_0, \ldots, \tilde{A}_n]$ is a vector of fuzzy parameters exhibited in the form of symmetric triangular fuzzy numbers denoted by $\tilde{A}_j = (p_j, c_j)$, $j = 1, \ldots, n$, with its membership function depicted as (4) bellow where p_j is its central value and c_j is its half width (See Figure 1).

A fuzzy linear relationship can be represented by a band (the bold lines having membership=0) with a centre line (the dashed line having a membership=1) as in Figure 2.

$$\mu_{\tilde{A}j}(a_j) = \begin{cases} 1 - \dfrac{|a_j - p_j|}{c_j}, & p_j - c_j \le a_j \le p_j + c_j, \\ 0, & otherwise. \end{cases} \quad (4)$$

Therefore, Eq. (3) can be written as:

$$\tilde{Y} = (p_0, c_0) + (p_1, c_1)x_{i1} + \ldots + (p_n, c_n)x_{in}. \quad (5)$$

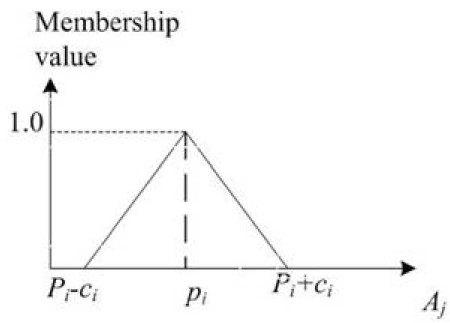

Fig.1: Triangular representation of fuzzy numbers

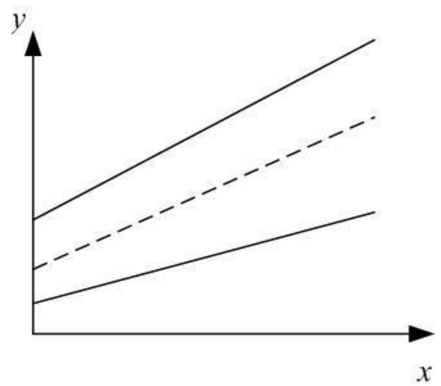

Fig.2: A Fuzzy linear relationship

Since the regression coefficients are fuzzy numbers, the estimated dependent variable $\tilde{Y}$ is a fuzzy number.

Finally, the method uses the criterion of minimizing the total vagueness, S, defined as the sum of individual spreads of the fuzzy parameters of the model.

$$Minimize \ \ S = mc_0 + \sum_{i=1}^{m}\sum_{j=1}^{n} c_i |x_{ji}| \quad (6)$$

The fuzzy coefficients are determined such that the estimated fuzzy output $\tilde{Y}$ has the minimum fuzzy width c_j, *while* satisfying a target degree of belief h. The term h can be viewed as a measure of goodness of fit or a measure of compatibility between the regression model and data. Each of the observed data sets, must fall within the estimated $\tilde{Y}$ at h levels (Figure 3). The value of h is between 0 and 1 and h=0 indicates that the assumed model is extremely compatible with the data, while h=1 illustrated the assumed model is extremely incompatible with the data. h is chosen by the decision maker. A choice of the h-level value influences the widths c of the fuzzy parameters:

$$\mu_{\tilde{Y}}(y_i) \ge h \qquad ,i = 1, 2, \ldots, m. \quad (7)$$

Taheri et al. (2006) purposed a method of sensitivity analysis based on credible level h. Their results showed that as the credible level h, increases, the mean of predictive capability (MPC) increases, too. On the other hand, by increasing h, the total vagueness of model, S, increases as well. For selecting a suitable h we would analyze the variations of S and h. Variations of S is gradual from h equal zero up to optimal h, after optimal h, increasing of h makes an abrupt variation in S value.

The problem of finding the fuzzy regression parameters was formulated by Tanaka et al. (1982) as a linear programming problem:

$$Minimize \ \ S = mc_0 + \sum_{i=1}^{m}\sum_{j=1}^{n} c_i |x_{ji}|$$

Subject to:

$$p_0 + \sum_{j=1}^{n} p_j x_{ij} - (1-h)\left[c_0 + \sum_{j=1}^{n} c_i x_{ij}\right] \le y_j$$

$$p_0 + \sum_{j=1}^{n} p_j x_{ij} + (1-h)\left[c_0 + \sum_{j=1}^{n} c_i x_{ij}\right] \ge y_j \quad (8)$$

https://dx.doi.org/10.22161/ijreh.2.4.3 ISSN: 2456-8678

Eq. (8) is linear, thereby allowing the optimization problem to be solved by means of linear programming.

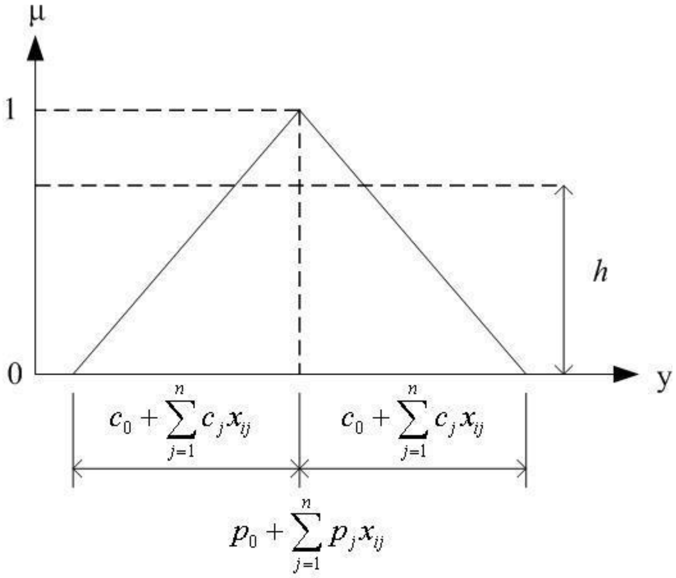

Fig.3: Triangular membership function of fuzzy output

III. APPROACH

The evapotranspiration (ET)-wheat yield (Yield) data presented in Kang et al. (2002) and Zhang et al. (1999) was used in this study.

One of our data bases is consist of experimental irrigation data, grain yield, seasonal ET, water use efficiency and climatic data summary during growing season winter wheat at four locations in the piedmont and lowland of the North China Plain (Zhang et al., 1999). The locations are divided into two groups that represented different geographic characteristics in the regions based on the groundwater table and geography. Luacheng and Gaocheng are located in the piedmont of the Taihang Mountains, and Linxi and Nanpi are located in the lowland of the Haihe floodplain. The irrigation treatments are ranged from no irrigation (rain-fed: I_0) to a maximum of seven irrigations (I_1, I_2, I_3, I_4, I_5, I_6, and I_7) where subscript represents the number of irrigations during the crop-growing season in Gaocheng and Linxi, and to a maximum of five irrigations in Luancheng and Nanpi. The amount of water applied was about 45–75 mm each irrigation. Grain yield and seasonal evapotranspiration are listed in Table 1.

Another data base (Kang et al., 2002) is consist of dataset form a lysimeter experiment that has been conducted for winter wheat (Triticum aestivum L.) during the period 1995-1998 to evaluate the effects of limited irrigation on grain yield on the Loess Plateau of China. Kang et al. (2002) applied a controlled soil water deficit, either mild or severe, at different stages of crop growth. The average values of evapotranspiration and grain yield for different treatments in 1995-1998 are given in Table 2.

Table.1: Grain yield and seasonal evapotranspiration for four locations in North China (Zhang et al., 1999)

	Gaocheng		Linxi		Luancheng		Nanpi	
Irrigation treatment	ET (mm)	Yield (Kg/ha)	ET (mm)	Yield (Kg/ha)	ET (mm)	Yield (Kg/ha)	ET (mm)	Yield (Kg/ha)
I_0	242	2580	247	2610	264	3220	281	2800
I_1	305	3600	277	3740	356	4770	355	3010
I_2	365	4960	358	4670	379	5250	420	4060
I_3	407	5230	414	4990	377	5250	418	4940
I_4	437	5280	428	5120	439	5100	443	4750
I_5	437	4240	426	4890	453	4790	456	5160
I_6	419	4360	478	4940				
I_7	423	4950	489	4440				

In current study, linear fuzzy regression (Tanaka et al., 1982) are employed and Evapotranspiration- Yield fuzzy relationships for Luancheng, Napai (Zhang et al., 1999) and Loess Plateau of China (Kang et al., 2002) were obtained.

For this purpose, complete dataset of Luancheng and Nanpi are applied. Zhang et al. (1999) has mixed Luacheng – Gaocheng datasets and presented a least square regression model for piedmont. In addition, the least square model for Linxi - Nanpi was reported as lowland. In this study, fuzzy regression model is obtained for Luancheng and Nanpi and Gaocheng and Linxi datasets are used for validation of

https://dx.doi.org/10.22161/ijreh.2.4.3 ISSN: 2456-8678

fuzzy regression models which are derived from Luancheng and Nanpi datasets, respectively.

Moreover, the dataset of eight different soil water content treatments (1, 3, 5, 7, 9, 11, 13, 15) in 1995-1996 (Table. 2) is used to obtain ET-Yield fuzzy regression model in the Loess Plateau of China. Finally, for model validation, yield estimation of fuzzy model for water content treatments: 2, 4, 6, 10, 12 and 14 evaluated with observation data.

In these cases, (having only 5 or 8 observation), it is impossible to satisfy the basic assumption of statistical regression analysis (such as normality of error, independence of errors, and so on). So fuzzy regression can be used as an alternative approach.

Value of total vagueness (S) calculated for h = 0-0.95 with 0.05 intervals and acceptable value of h was determined.

Table.2: Total evapotranspiration and grain yield in three growing seasons in the Loess Plateau of China (Kang et al., 2002).

Treatments	1995-1996		1996-1997		1997-1998	
	ET (mm)	Yield (Kg/ha)	ET (mm)	Yield (Kg/ha)	ET (mm)	Yield (Kg/ha)
1	267	2493	213	1750	220	1612
2	308	3520	300	3180	277	3060
3	304	3089	278	3375	231	2039
4	310	3533	385	3905	232	1771
5	301	3060	359	3570	310	4079
6	339	3506	291	3505	235	2040
7	356	3441	338	3870	296	3060
8	370	3659	387	4020	285	2788
9	362	3672	323	4080	254	3076
10	305	3680	389	4230	285	3852
11	292	3294	403	4245	227	2045
12	399	4233	519	4200	358	4060
13	354	4325	420	4600	330	4749
14	367	4485	383	4775	340	4811
15	370	4553	390	4920	329	4792

IV. RESULTS

In applying fuzzy linear regression, grain yield(Kg/ha) is employed as the dependent variable and evapotranspiration, ET(mm) is assumed as independent variable. All the Yield and ET values are assumed to be crisp. The symmetric triangular form of the membership function is chosen for representing the regression parameters. According to Figure 4, it is obvious that by taking large value for h, amount of S increase quickly. So, it seems that the values around 0.7 for h, are suitable values for h and this is in an agreement with Bardossy et al. (1990). According to Bardossy et al. (1990), the level of credibility is generally chosen so that $0.5 < h < 0.7$.

The fuzzy model with symmetric triangular fuzzy coefficients for crop production modeling of winter wheat in three locations in China, as a function of growing season evapotranspiration, can be stated as follows:

$$\tilde{Y} = (p_0, c_0) + (p_1, c_1)\,ET$$

Based on 6 data in Table 1, for Nanpi region, and adapting relation (8), the objective function is:

$$Minimize\ \ S = 5c_0 + 2373c_1$$

In addition, constrains (12 constrains) related to observations (6 observations) must be formulated, based on relation (8). For example, two constrains corresponding to the first observation, with h=0.7, are:

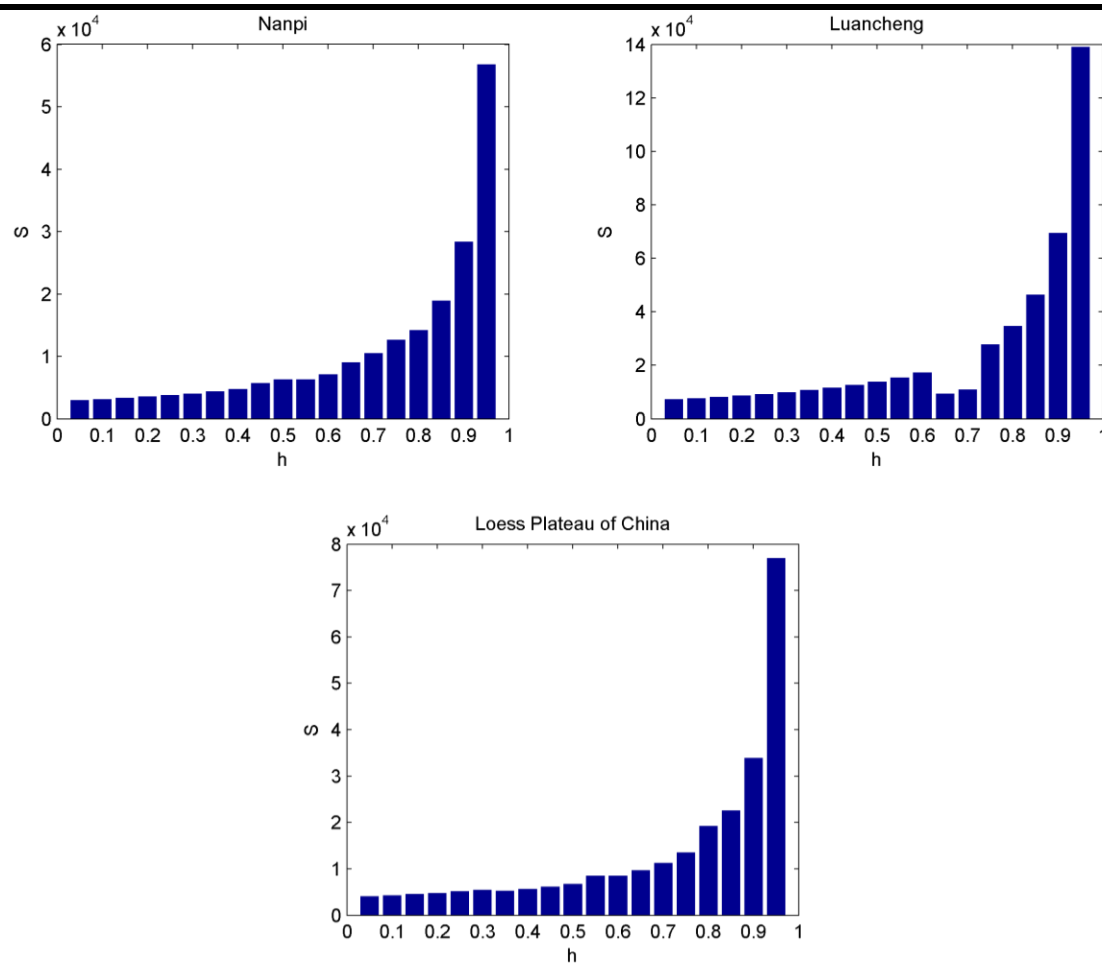

Fig.4: The variation of the total vagueness (S), based on different amounts for h.

$$p_0 + 281p_1 - 0.3(c_0 + 281c_1) \le 2800$$

$$p_0 + 281p_1 + 0.3(c_0 + 281c_1) \ge 2800$$

By minimizing the objective function S subject to 12 constrains, with linear programming methods, the coefficients of the model are as follows:

$$\tilde{A}_0 = (-1589.34, 0.00) \qquad , \tilde{A}_1 = (14.29, 4.44)$$

Therefore, the possibility regression model for Nanpi region is:

$$\tilde{Y} = (-1589.34, 0.00) + (14.29, 4.44)\,ET$$

In addition, the coefficients of the possibilistic regression model were calculated for Luancheng and the Loess Plateau of China. The results are shown in Table 3.

The results of fuzzy regression model for simulation data are shown in Figure 5. An estimation area at the high evapotranspiration is wider than low evapotranspiration (Figure 5).

Table.3: The possibilistic regression models for three sample area with h=0.7.

Location	Model	Total vagueness (S)
Nanpi	$\tilde{Y} = (-1589.34, 0.00) + (14.29, 4.44)\,ET$	10538
Luancheng	$\tilde{Y} = (1026.98, 0.00) + (9.75, 4.82)\,ET$	10942
Loess Plateau of China	$\tilde{Y} = (-351.00, 0.00) + (11.95, 4.35)\,ET$	11302

https://dx.doi.org/10.22161/ijreh.2.4.3 *ISSN: 2456-8678*

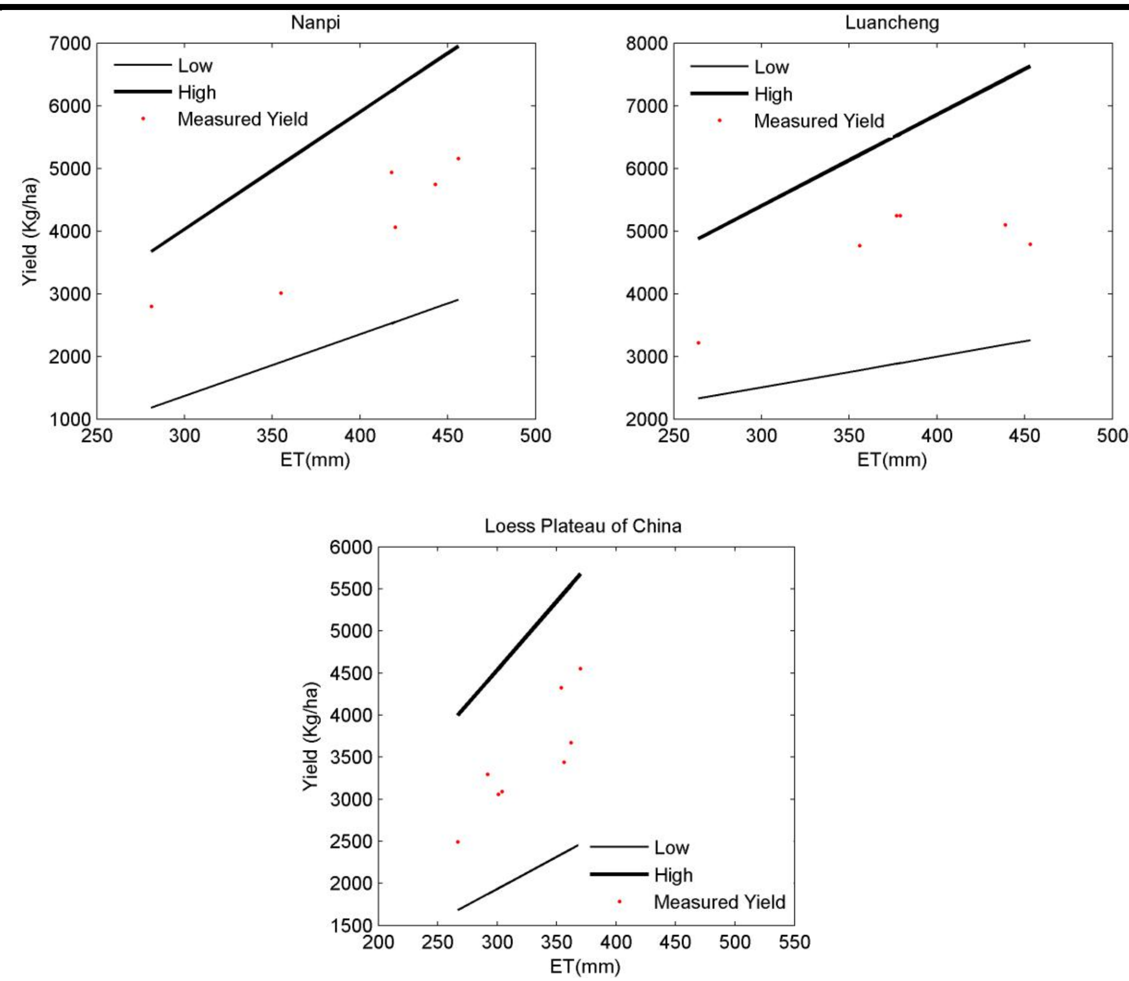

Fig.5: Fuzzy regression relationships between winter wheat yields and ET in three locations in China.

The variation of estimation area illustrates that uncertainly of simulation data, along the ET axis changes. From the simulation results, it can be understood that the estimation area can well express the degree of dispersion at each evapotranspiration more practically than the conventional regression method can, and therefore the area not only represents the relation between ET and grain yield but also has information on reliability, while the conventional crop production function represents only the relations between ET and yield.

The uncertainty in field data is caused by variation in the climate of region (drought, wind and frost) and offense of insects and pests, etc.

Interestingly, the half-width for the intercept is optimized to a value of zero during the minimization of the vagueness criterion in three locations (Nanpi, Luancheng and Loess Plateau of China), (Table. 3). Hence, the intercept of the fuzzy regression model is a crisp number and all of the fuzziness in the model arises from the slop being a fuzzy quantity.

Figure 6 shows a representation of fitness of fuzzy regression. Validation of fuzzy regression models for estimation of coefficients of crop production functions in these regions is evaluated with test data. Figure 6 (a) shows position of ET-Yield data of Linxi district in possibilistic regression model for Nanpi region.

https://dx.doi.org/10.22161/ijreh.2.4.3 *ISSN: 2456-8678*

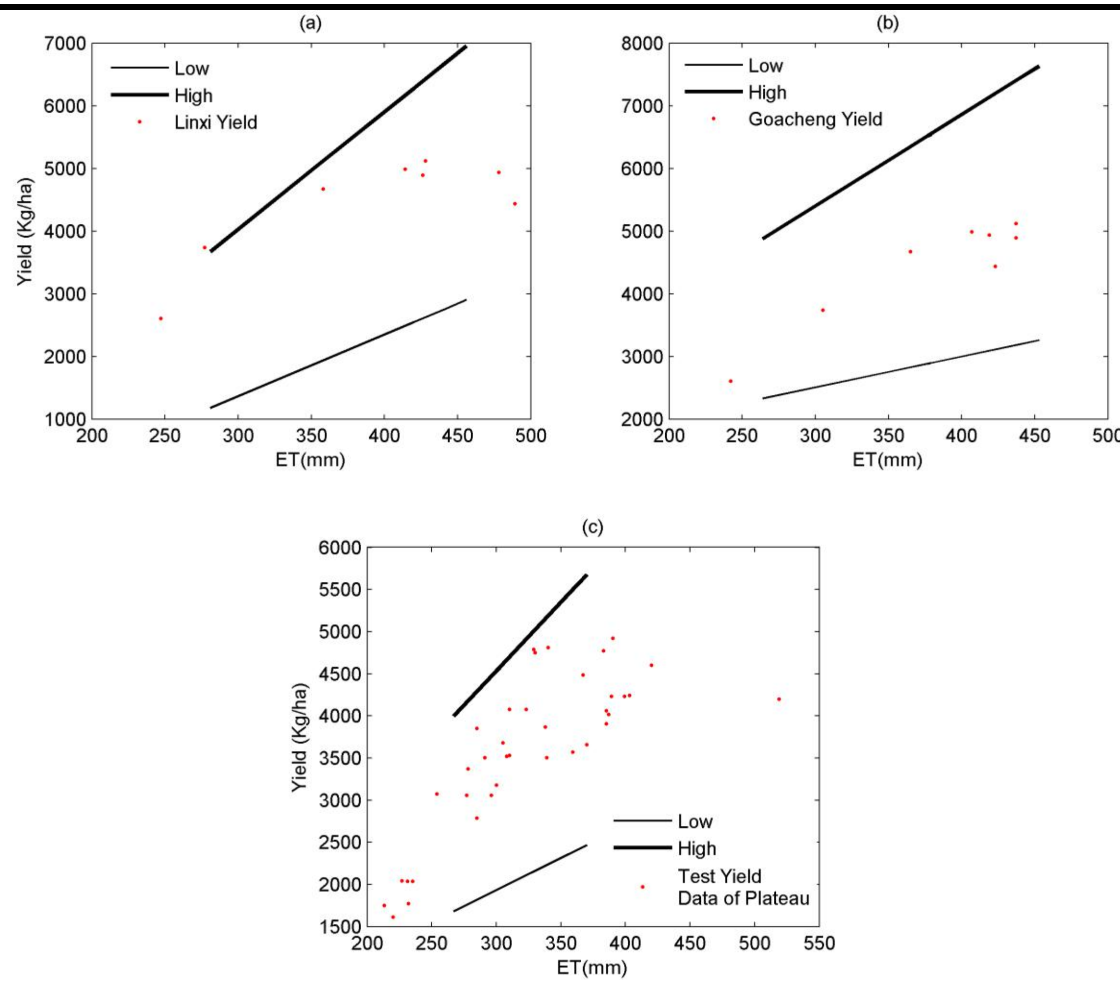

Fig.6: Representation of fitness of fuzzy model, using testing data.

According to Zhang et al. (1999), Linxi and Nanpi are located in the lowland of the Haihe floodplain and they represented same geographic characteristics in the region based on the groundwater table and geography. So, the estimated model for Nanpi should be applicable in Linxi. Figure 6(a) shows that Linxi data is in a good agreement with derived linear regression model for Napai. The derived Luancheng regression model is verified with Gaocheng data (Figure 6(b)).

Also, the fuzzy regression model for Loess Plateau of China evaluated with 37 ET-Yield data in this region (Table 2.). Figure 6(c) illustrates capability of fuzzy linear egression in estimation of production function despite of deficit data.

V. CONCLUSION

A fuzzy linear regression is used to estimate coefficients of crop production function. For this purpose, evapotranspiration- yield measurements of winter wheat are used for three districts in China. Crop yield is a sensitive parameter and climate, soil, water and crop alter the predicted yield. Evapotranspiration is the most important factor in yield estimation. Having crop production function in each district is necessary for estimation of yield condition, but, there should be many data estimation of crop production function with classical least square regression. As received from this study, fuzzy linear regression provides a convenient alternative to characterize crop yield in deficit data condition. The degree of believe is determined by Taheri et al. (2006) method. Validation of model is done by test data. However, this approach is suitable for crop yield predicting by few data.

REFERENCES

[1] Bardossy, A. (1990). Note on fuzzy regression. Fuzzy sets System. 37, 65-75.

[2] Cheng Si, B., & Bodhinayake, W. (2005). Determining soil hydraulic properties from tension infiltrometer measurements: Fuzzy regression. Soil Science Society of American Journal. 69,1922–1930.

[3] Cole, J.S., & Mathews, O.R. (1923). Use of water by spring wheat on the Great Plains. USDA Bureau of Plant Industry, Bulletin No. 1004.

[4] Finger, R., Hediger, W. (2007). The Application of robust regression to a production function comparison – the example of Swiss corn. Available at: http://mpra.ub.uni-muenchen.de/9097/ MPRA Paper No. 9097, posted 16. June 2008 / 08:09.

[5] Hanks, R. J., H. R. Gardner, & Florian R. L. (1969). Plant growth- evapotranspiration relations for several crops in the central Great Plains. Agronomy Journal. 61, 30-34.

[6] Kang, S., Zhang, L., Liang, Y., Hu, X., Cai, H., & Gu, B. (2002). Effect of limited irrigation on yield and water use efficiency of winter wheat in the Loess Plateau of China. Agricultural Water Management. 55, 203-216.

[7] Mathews, O. R., & Brown, L. A. (1938). Winter wheat and sorghum production in the southern Great Plains under limited rainfall. USDA Circular 477. U.S. Government Printing Office, Washington, DC.

[8] Sahin, V. & Hall, M. J. (1996). The effects of afforestation and deforestation on water yields. Journal of Hydrology. 178, 293-309.

[9] Shimosaka, T., Kitamori, T., Harata, A., & Sawada, T. (1996). Fuzzy linear regression method in and analytical interpretation ultratrace analysis. Analytical Sciences. 12: 385-391.

[10] Taheri, S. M., Tavanai, H., & Nasiri, M. (2006). Fuzzy versus statistical rRegression in false- twist texturing. WSEAS Transactions on Mathematics. 10(5): 1109-2769.

[11] Tanaka, H., Uejima, S., & Asima, K. (1982). Linear regression analysis with fuzzy model. IEEE Transactions on Systems, Man, and Cybernetics. 12, 903-907.

[12] Thornthwaite, C.W. (1948). An approach towards a rational classification of climate. Geographical Review. 38, 55-94.

[13] Tseng, F. M., Tzeng, G. H., & Yu, H. C. (1999). Fuzzy seasonal time series for forecasting the production value of the mechanical industry in Taiwan. Technological Forecasting and Social Change. 60, 263–273.

[14] Vaux, H.J., Jr., & Pruitt, W.O. (1983). Crop-water production function. In D. Hillel (ed.) Advances in irrigation. Vol. 2. Academic Press, New York, p. 61-97

[15] Zhang, H., & Oweis, T. (1999). Water-yield relations and optimal irrigation scheduling of wheat in the Mediterranean region. Agricultural Water Management. 38-195-211.

[16] Zhang, H., Wang, X., You, M., & Liu, C. (1999). Water – yield relations and water-use efficiency of winter wheat in the North China Plain. Irrigation Science. 19: 37–45.