ISBN: 978-979-1421-15-7

An Implementation of Fuzzy Inference System for Onset Prediction Based on Southern Oscillation Index for Increasing the Resilience of Rice Production Against Climate Variability

Agus Buono[1, 2] and Mushthofa[2]
[1)]Center for Climate Risks and Opportunity Management in Southeast Asia and Pacific (CCROM-SEAP) Bogor Agriculture University, Bogor-West Java, Indonesia
[2)] Department of Computer Sciences, Faculty of Mathematics and Natural Sciences, Bogor Agricultural University, Bogor-West Java, Indonesia
Email: pudesha@yahoo.co.id, imoid@yahoo.com

Abstract—**Rice production system in Indonesia is very sensitive to global phenomena, especially the El Nino phenomenon. Information regarding the onset of rainy season is important to increase the resilience of the production system. This paper is focused on the implementation of Fuzzy Inference System (FIS) as a technique for predicting the onset of rainy season based on the Southern Oscillation Index (SOI) data in the months of July, August, September and October. There are two data sets used: the SOI data from the year 1877 to 2011 and the rainy season onset in the District of Indramayu. Fuzzy set memberships and the set of rules are designed by investigating the two sets of data (via visualization and clustering). The prediction system is verified by using the actual data from the district. The result of the verification shows that the correlation between the rainy onset and its predicted value is 0.68. Even though the prediction accuracy is relatively competitive compared to the existing methods, the requirement of the use of the SOI variables from the month of October renders the model less useful in practice, since it would be mostly too late to wait until October to perform the prediction. Further research can be developed which integrates Markov Chain method to overcome this problem.**

Index Terms— **Rice production system, El Nino, Fuzzy Inference System, Onset, Southern Oscillation Index.**

I. Introduction

The agricultural sector, especially the food subsector is one of the important subsectors which is very susceptible to weather and climate change. This is due to the fact that plants generally need a certain climate/weather condition based on empirical data, the variability of rice production correlates strongly to the variability of rain fall. According to [1], 81% of crop failures are caused by climate variability, while only 19% of them are caused by pests. Of the failures caused by climate variability, 90% are caused by drought occurring during the second planting season.

Empirical records show that droughts are generally influenced by global phenomena, such as the El Nino Southern Oscillation (ENSO), and by climate variables such as: the amount of rainfall, the length of the rainy season, as well as the onset of the rainy season. Therefore, early information regarding these climate variables can be very useful in determining anticipative steps in rice field cultivation.

This paper specifically presents an implementation of a Fuzzy Inference System for the prediction of rainy season onset. Buono, et. al. (2012), [2] has developed an artificial neural network model based on SOI data on the month of June, July and August. The resulting accuracy was not so satisfactory, of about 0.6. Also, the model was built using data only from one station, so that the model could not represent the general characteristics of a certain area. In this research, a fuzzy system is developed based on the knowledge gained from data exploration in a certain district to cover a larger range of area. Also, the computation of the output from the input data is based on a logical knowledge backed by observation data. Accordingly, it is expected to have a better accuracy result than the previous research.

The remainder of this paper is organized as follows: Section 2 presents the principles of rice production in Indonesia. Section 3 describes the data and the methods. Section 4 is addressed to explain the design of the computational model. Results and discussions are presented in Section 5, and finally, Section 6 is dedicated to the conclusions of this research study and recommendations for future research.

II. Rice Production System

In general, for Indonesian farmers there are two planting season: rainy season and dry season. On the first season, planting begins during the start of the rainy season and continues for the next four months. For fields located near the irrigation channels, planting can be immediately done. For fields located farther away from the channels, however, planting will be delayed in accordance to their distance from the irrigation channels. This kind of planting pattern will be

ISBN: 978-979-1421-15-7

continuously affected by the characteristics of the climate, be it normal, or in the presence of El Nino or La Nina, as depicted in Figure 1.

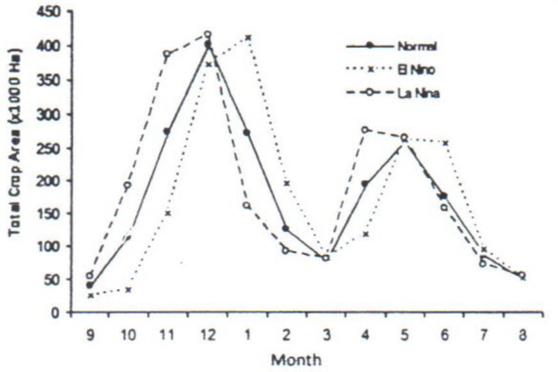

Figure 1. Yearly planting pattern according to the three climate conditions (Source: [3])

It can be seen from the picture that the peak of planting will experience delay in the presence El Nino, and experience advancement in the presence of La Nina. Problems often rise during the presence of El Nino, especially for the second planting (dry season). Rice production during this second planting will be highly susceptible to drought, especially for areas farther from irrigation channels. On these areas, the first planting cannot be performed in the beginning of rainy season. Therefore, the second planting will also be delayed, so that the risk of facing drought will be higher. By having the predictive information regarding the onset of rainy season, anticipative steps can be taken to prevent greater losses.

III. Data and Method

A. Data

The first data used for this research is the Southern Oscillation Index data, which can be downloaded at http://www.bom.gov.au/climate/current/soihtml1.shtml. From this site, we collect the monthly SOI Index from the year 1877 to 2011. The second data prepared is the observational data of the onset of the rainy season obtained from five rainfall zones in the District of Indramayu. From each zone, we collect the onset of rainy season data from the year 1965 to 2009, with a few missing values.

The SOI index is the difference between the anomaly of the air-pressure in the Tahiti region and the Darwin region, divided by the standard deviation of the differences, and written as: [4]

$$SOI\ Index = \frac{AnP(Tahiti) - AnP(Darwin)}{STD(Diff)} \times 10 \qquad (1)$$

with:

AnP(Tahiti) = Tahiti air-pressure anomaly
AnP(Darwin) = Darwin air-pressure anomaly
STD(Diff) = The standard deviation of the differences between the above variables

Stone (1996) shows that SOI index (which is a global phenomenon) has influences on local climate condition, especially in Indonesia. Generally, there are five conditions on the values of SOI index, known as the SOI Phases, which are determined based on the values of SOI from two consecutive months, as shown in Figure 2.

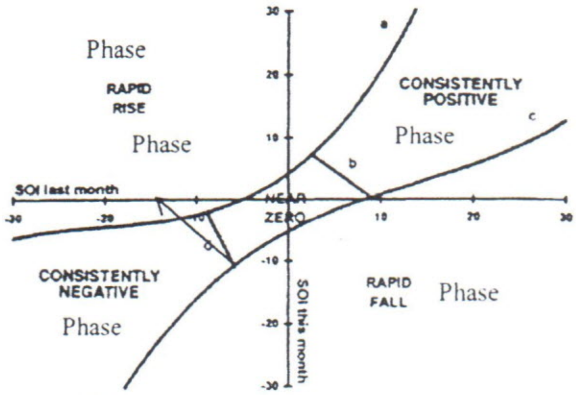

Figure 2. Five SOI Phases based on the SOI values of two consecutive months (Source:[4])

Based on empirical data, if the SOI falls into Phase 1 or 3 then there is a tendency towards the presence of El Nino. Meanwhile, La Nina is generally preceded by SOI Phase 2 or 4. If SOI is in Phase 5, it is generally expected that the climate will run in a normal course.

The data for the onset of rainy season is measured within a period of 10 days. Hence, there are 36 possible values of rainy season onset: from 1 (first ten days of the year) to 36 (the last 10 days of a year, one year is approximated to be 360 year. As an example, if the onset value is 33, then this means the start of the rainy season is around the last week of November. We follow the definition of rainy season onset according to Moron [5], which is: the occurrence of rain above 50mm during 3 consecutive 10-day periods. Based on the data from 1965 to 2009, the average onset anomaly in the district of Indramayu is as presented in Figure 3. Negative values mean rainy season occurred ahead of normal, while positive values mean that rainy season occurred later than normal.

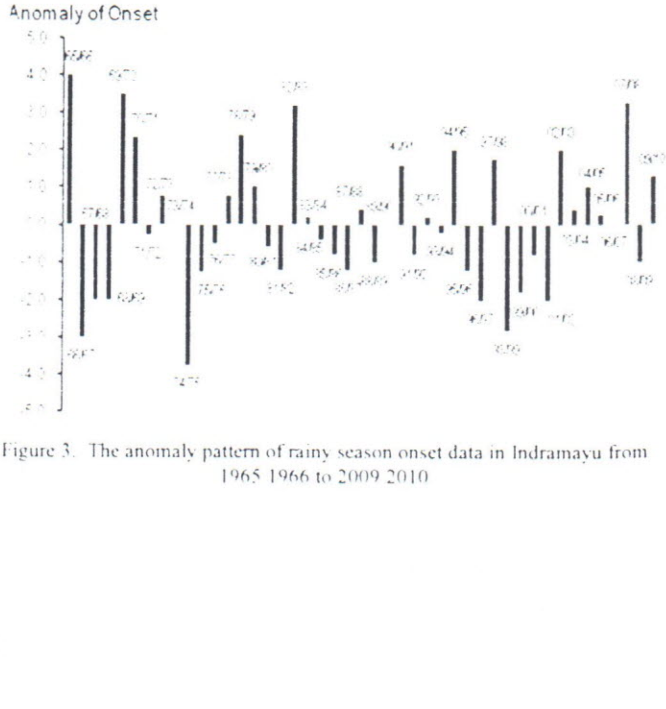

Figure 3. The anomaly pattern of rainy season onset data in Indramayu from 1965 1966 to 2009 2010

ISBN: 978-979-1421-15-7

B. Method

The development of the FIS model in this research follows the steps outlined at Figure 4.

a. Data Exploration: In this step, data exploration is performed in order to obtain knowledge on two things: the distribution pattern of the SOI and the pattern of the relationship between SOI with the onset. The information regarding the distribution pattern of the SOI and the rainy season onset is used to determine the fuzzy sets of these variables, while the pattern of the relationships between the SOI and the rainy season onset data is used to determine inference rules mapping input variable (SOI) to output variable (rainy season onset).

b. Fuzzyfication: On this step, we define the fuzzy sets for each of the variables, namely: the SOI and the rainy season onset. In this case, the rainy season onset values are represented as the values of the anomaly of the onset, obtained from the original values of the onset, subtracted by their mean, computed from the data obtained from 1965 to 2010.

c. Rule Generation: On this step, rules are constructed which maps the input variable to the output variable. Other than the use of the relationship pattern between the SOI and the rainy season onset, the construction of the rule also uses clusters analysis. In this case, we perform year clustering based on the values of the SOI variable to find the clusters of the SOI data. Based on this cluster analysis, we characterize the onset values and determine the fuzzy inference rules.

d. Implementation and evaluation: on this step, an implementation program is written using the Matlab programming language to build the fuzzy inference system. Evaluation is performed by computing the output of the system when supplied with the available data, and comparing the results against the available rainy season onset data.

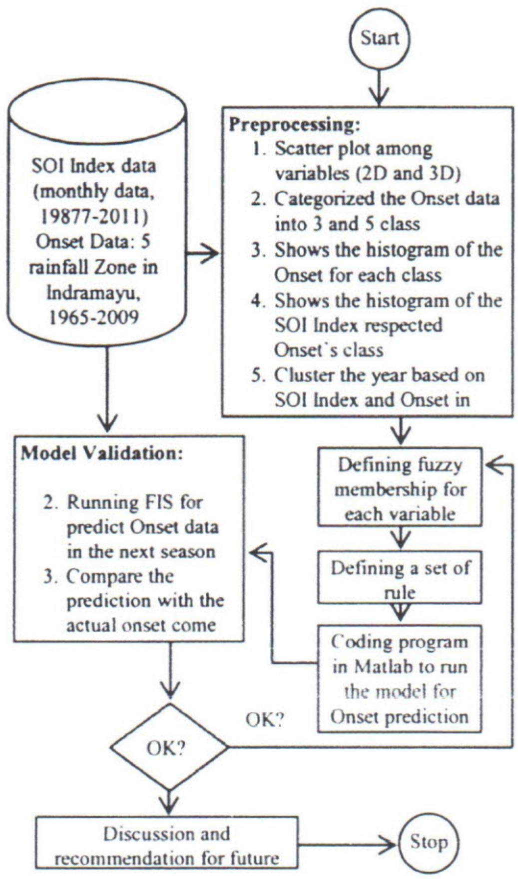

Figure 4. Flow diagram of the research method

IV. MODEL DESIGN

A. SOI Index and Onset Data Exploration

Here we perform visual data exploration with the purpose of obtaining the characteristics of SOI, and the relationship between the SOI and the onset values. The onset values are divided into three categories, namely: delayed, represented by the value 1) normal (represented by the value 0) and early onset (represented by the value -1).

These categories are based on the anomaly of the onset values obtained from 1965 to 2009, computed from the mean value of 33. For an example, consider that in the year 2002/2003, the onset value was 35, hence the anomaly is 35-33 = 2, categorized as the value 1 (delayed). This will make it easier to notice the relationship between the SOI and onset values. We will first examine the distribution pattern the SOI, and followed by the relationship pattern between the SOI and the onset values.

Figure 5 shows the distribution pattern of the SOI in the months of July to October (Red: delay, black: normal, green: early). From Figure 5, we can conclude that for early onset, the distribution of the SOI values tend to be positive with a mean value of approximately 7.5. This occurs on each of the four months. On the other hand, for normal and delayed onset values in the month of July shows a small difference. However for the month of August, September and October, the differences are significant, despite still be smaller than in the early Onset.

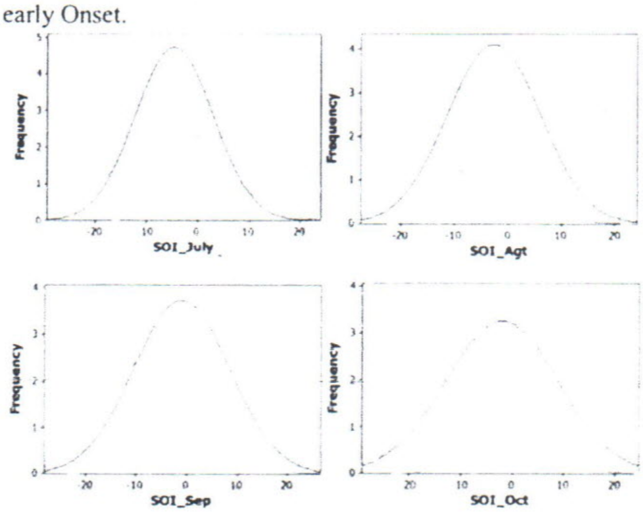

Figure 5. Distribution of SOI Index based on Onset category

ISBN: 978-979-1421-15-7

The average SOI for early delayed and normal onset for the months of August, September and October are relatively equal, with each ranges approximately -7.5 and 0. The average for July in normal onset is about -2.5, while for delayed onset is approximately -5. On top of that, to further increase the sensitivity of the system, 5 memberships sets for SOI are defined: Very Negative, Negative, Zero, Positive and Very Positive. The complete membership functions for SOI are given in Figure 6.

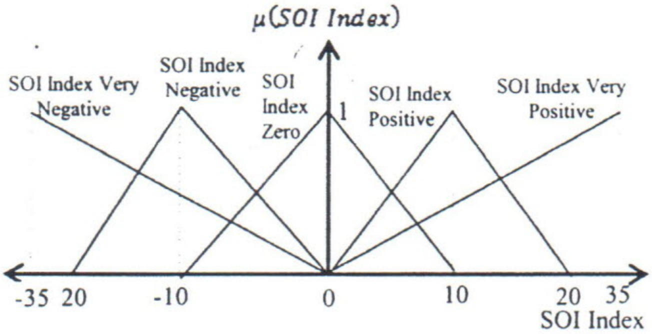

Figure 6. Fuzzy membership for SOI Index

On the other hand, there are five patterns of the rainy season onset data, as given in Figure 7. From Figure 7, it can be seen that the centers of the clusters of the onset anomaly are: approx. -3.2, -1.5, 0, 1.7 and 3.5. Based on this information, the fuzzy sets for the onset values are constructed as presented in Figure 8.

B. Building the Set of Inference Rules

Figure 9 shows the distribution of data points in 3D of the three SOI values. From this figure, we can see a clear relationship between the SOI values and the onset values. In this case, if the SOI value is negative, then the onset anomaly tends to be positive (red), which means that the rainy season onset is likely to be late. On the other hand, positive SOI values indicates that the rainy season onset will tend to be negative, which means that rainy season is likely to be ahead of normal.

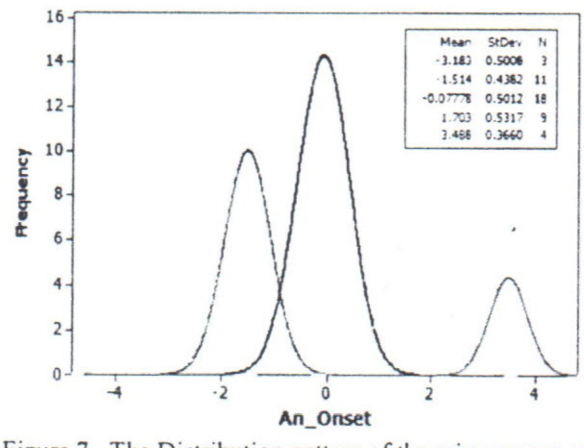

Figure 7. The Distribution pattern of the rainy season onset

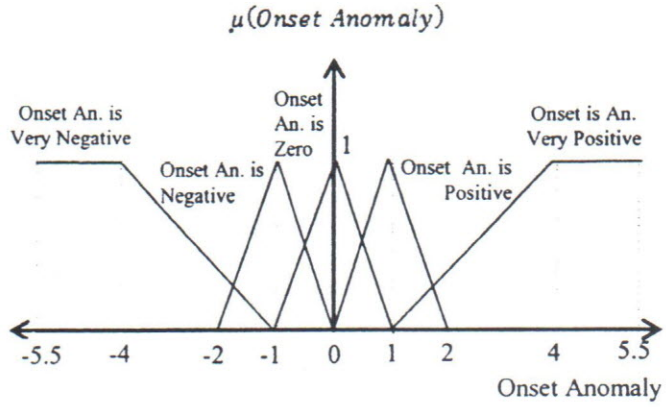

Figure 8. Fuzzy membership for the Onset Anomaly

This relationship appears to be consistent from July to August. These facts are also backed by the scatter plot in Figure 10.

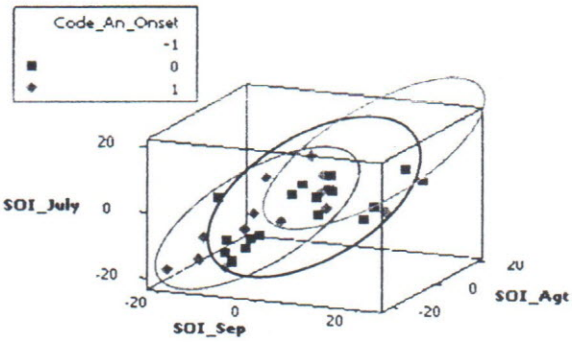

Figure 9. 3D scatter plot of three successive SOI Index (red: delay Onset, black: normal Onset, green: early Onset)

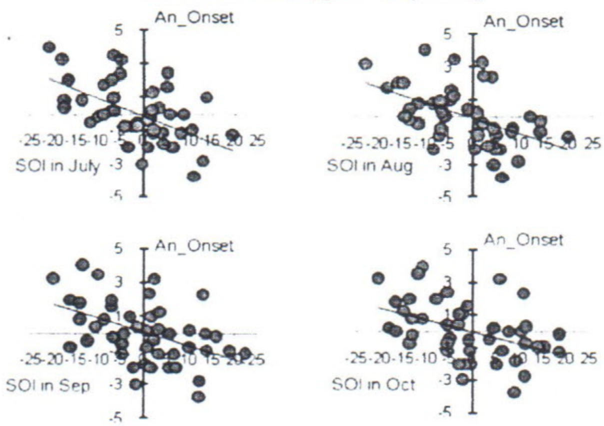

Figure 10. 2D Scatter plot between SOI Index with Anomaly of Onset

Based in these two facts, it is fairly logical to use the SOI to predict the value of the rainy season onset. We can also conclude that if the SOI value decreases, then the onset anomaly value should increases, which means that the rainy season will be delayed. The onset anomaly will be normal when the SOI is in the Zero Phase. Finally, a positive value of the SOI will lead to negative onset anomaly, which means that the rainy season onset is delayed. To obtain other possible patterns, we performed cluster analysis using the SOI variable and the onset anomaly variable.

ISBN: 978-979-1421-15-7

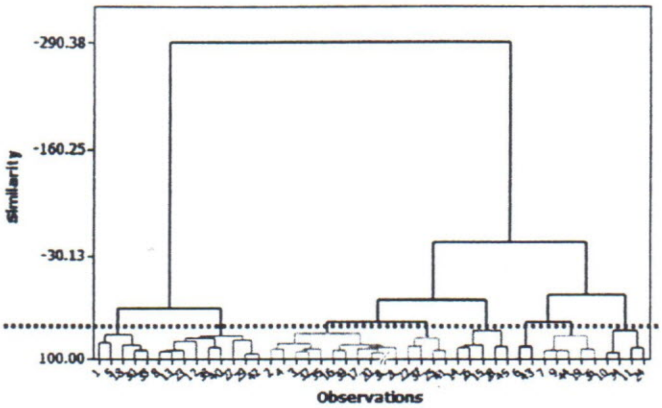

Figure 11. Dendrogram showing the result of the year clustering based on the SOI and the onset

Figure 11 shows the dendrogram of the cluster analysis on the year 1965 to 2009. By cutting the dendrogram on Figure 11 along the dashed line, we obtain 8 clusters which correspond to 8 fuzzy rules. The average values of the SOI and onset anomaly for the 8 clusters are presented in Table 1. From Table 1, we can create 8 different fuzzy rules, namely:

Rule 1: **IF** (*SOI in July is Very Negative*) and (*SOI in August is Very Negative*) and (*SOI in September is Very Negative*) and (*SOI in October is Very Negative*)
THEN (*Onset Anomaly is Very Positive*)

Rule 2: **IF** (*SOI in July is Zero*) and (*SOI in August is Positive*) and (*SOI in September is Zero*) and (*SOI in October is Zero*)
THEN (*Onset Anomaly is Negative*)

Rule 3: **IF** (*SOI in July is Negative*) and (*SOI in August is Zero*) and (*SOI in September is Positive*) and (*SOI in October is Positive*)
THEN (*Onset Anomaly is Positive*)

Rule 4: **IF** (*SOI in July is Zero*) and (*SOI in August is Positive*) and (*SOI in September is Very Positive*) and (*SOI in October is Very Positive*)
THEN (*Onset Anomaly is Zero*)

Rule 5: **IF** (*SOI in July is Negative*) and (*SOI in August is Negative*) and (*SOI in September is Negative*) and (*SOI in October is Negative*)
THEN (*Onset Anomaly is Zero*)

Rule 6: **IF** (*SOI in July is Very Positive*) and (*SOI in August is Very Positive*) and (*SOI in September is Very Positive*) and (*SOI in October is Very Positive*)
THEN (*Onset Anomaly is Very Negative*)

Rule 7: **IF** (*SOI in July is Zero*) and (*SOI in August is Zero*) and (*SOI in September is Zero*) and (*SOI in October is Negative*)
THEN (*Onset Anomaly is Positive*)

Rule 8: **IF** (*SOI in July is Zero*) and (*SOI in August is Negative*) and (*SOI in September is Zero*) and (*SOI in October is Positive*)
THEN (*Onset Anomaly is Zero*)

Table 1. Average SOI Index and Onset Anomaly for each cluster

Cluster	SOI Index in				Onset
	Jul.	Aug.	Sep.	Oct.	Ano.
1	-14.20	-14.00	-14.76	-14.66	2.89
2	2.73	3.23	1.36	-0.83	-1.33
3	-5.00	2.85	7.50	8.55	2.79
4	-0.32	7.88	12.58	11.90	-0.37
5	-10.43	-10.98	-9.36	-8.80	0.38
6	13.50	12.48	16.30	13.63	-2.20
7	3.90	-2.00	-0.04	-8.32	1.29
8	2.20	-6.85	1.55	5.73	-0.74

The fuzzy inference model used in this research is the Mamdani-type inference using centroid as the defuzzification method. The structure of the model is as presented in Figure 12 [6].

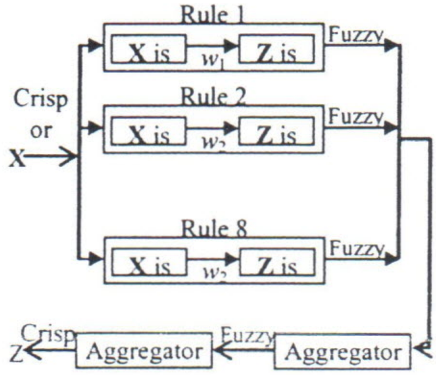

Figure 12. The structure of the FIS with 8 rules.

V. RESULT AND DISCUSSIONS

In order to validate the FIS that has been developed by using Matlab, we predict the Onset of 5 rainfall Zones in Indramayu and their average (also called district level Onsets). Scatter plot between the actual Onsets with their predictions for 5 rainfall zones and the district level are depicted by Figure 13.

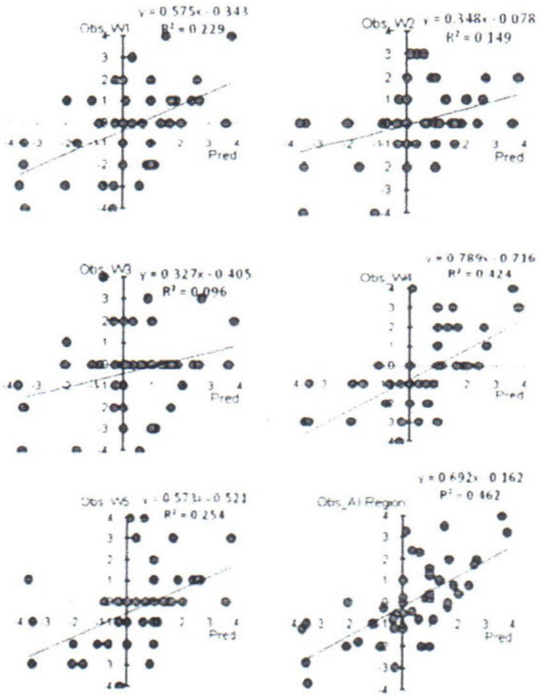

Figure 13. Scatter Plot Obs. Vs Preds. (Rainfall Zone 1, Zone 2, Zone 3, Zone 4, Zone 5, and at District level)

Based on Figure 13, two observations can be made. Firstly, that in general, the two series (prediction and actual Onset) have similar patterns, especially for the district level. The correlation is in the range of 0.31 up to 0.65, and 0.68 for the district level. Secondly, it seems that other variables are required as predictors in the Onset prediction.

Figure 14 presents the comparison on onset prediction value among ENSO condition (ES:Strong El Nino, EW:El Nino Weak, M:Moderate, LW:La Nina Weak, and LS:La Nina Strong). It can be seen that the Onset resulted by the model are relevant to what climatologists would say in real life conditions: that the Onset will come early in La Nina and will be delayed in El Nino condition. The magnitude of the anomaly will vary with respect to the level of the ENSO condition. In addition, we can see that in the moderate condition, the accuracy of Onset prediction drops. This statement is supported by Figure 14, i.e. the correlation between the actual rainy onset and its predicted value is 0.59. In a strong ENSO, the correlation is greater than 0.85. This means that, we have to improve the rule and fuzzy memberships, and have to explore more details regarding the phenomena in moderate ENSO conditions.

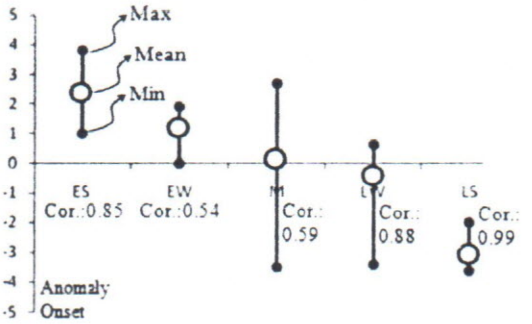

Figure 14. Comparison prediction of Onset Anomaly among ENSO condition

VI. Conclusions

Based on the experiment, we can conclude the followings:

a. Fuzzy Inference System can be implemented to Onset prediction with an accuracy level of 0.68 in the district level, is able to model and explain them logically.

b. In strong ENSO condition, the Onset value is easier to predict with higher accuracy compared to the moderate condition.

c. Other predictors are required in order to improve the accuracy.

In order to increase the system accuracy, for future research we will focus on the parameter optimization of the model and adding the predictor variable with other ENSO parameters.

Acknowledgment

This research is funded by I-MHERE Project under contract number: 7/13.24.4/SPP/I-MHERE/2012 Bogor Agricultural University, and partly supported by the Computer Science Department, Bogor Agriculture University.

References

[1] Boer, R. and Perdinan. 2008. Adaptation to climate variability and climate change: its socio-economic aspect. Proceeding of Workshop on 'Climate Change: Impacts, Adaptation, and Policy in South East Asia. Economy and Environmental Program for Southeast Asia. Bali.

[2] Buono, A. A. Kurniawan, and A. Faqih. 2012. The Onset Prediction by Using Artificial Neural Network Backpropagation Levenberg-Marquardtet. National Conference: The Application of Information Technology 2012 (SNATI 2012), ISSN 1907-5022, Jogyakarta.

[3] Boer, R., A. Buono, and Suciantini. 2010. Dynamic Cropping Calendar as a Tool to Decide the Cropping Pattern that Appropriate with the Climate Condition. I-MHERE Research Report, Bogor Agriculture University.

[4] Stone, R.C., Graeme L.M., and T. Marcussen. 1996. Prediction of Global Rainfall Probabilities Using Phases of the Southern Oscillation Index. Nature, Vol. 384, 21 November 1996. Nature Publishing Group.

[5] ---, 2011. Buleting Prakiraan Musim Kemarau 2011, BMKG Semarang. http://index.bmkgjateng.com/PRAMK2011.pdf, Accessed on September, 18th, 2012.

[6] Jang J-S R. 2001. *Neuro-Fuzzy and Soft Computing* . USA : Prentice-Hall