International Journal of Research in Business and Social Science
8(3), 2019: 64-71

Business & Social Science
IJRBS

Research in Business and Social Science

IJRBS Vol 8 No 3, Special Issue ISSN: 2147-4478
Contents available at *www.ssbfnet.com/ojs*

https://doi.org/10.20525/ijrbs.v8i3.254

Analysis of Factors Influencing the Usage of Seasonal Forecast in Drought Prone Area: A Case of East Nusa Tenggara, Indonesia

Heri Kuswanto

Corresponding Author: Department of Statistics, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia. Phone number : +62818513223.

ORCID ID: 0000-0003-0300-7286

Abstract

This research aims to investigate factors influencing the households in East Nusa Tenggara (NTT) Indonesia to use seasonal forecast issued by the Meteorological Office (BMKG) Indonesia. NTT is a highly vulnerable region to drought, and hence, understanding the future weather and climate condition is a crucial issue. Household survey was conducted to interview 300 households in NTT asking about their intention to utilize the forecast information in supporting their livelihood. Furthermore, this research investigates whether the households characteristics influencing the intention. Using the Classification and Regression Tree (CART), this research found that the intention to use forecast is dominantly influenced by age of the household head, ownership of household assets, main source of income, crop and livestock activities and water source for livestock.

Key words: *Forecast, CART, Adaptation, Livestock*

JEL classification: *Q54, R20* Submitted: 28.03.2019 - Accepted: 28.04.2019

Introduction

Drought is a natural phenomena that oftenly leads to significant losses especially in agricultural sector. Indonesia is a country that is prone to drought events, where East Nusa Tenggara (NTT) is one of the most vulnerable regions in Indonesia. NTT has experienced annual severe drought within last ten years. Many efforts have been conducted to minimize the negative impacts of the drought such as in Sena et al. (2014), WMO & GWP (2017) among others.

Most of the households in East Nusa Tenggara work in agriculturural sector which is highly dependent on weather and climate condition. To deal with this, adaptation strategies need to be formulated by the farmers in order to minimize the drought risk. Maracchi (2000) pointed out the importance of adapation to drought for agricutural activities. A study by Zhang et al. (2016) discussed the impact of drought on vegetation productivity in China. Moreover, Labedzki (2016) described mitigation actions for drought in agriculture in Poland.

The type of adaptation strategy would depend on the information about the weather and climate, in particularly the seasonal drought forecast. The BMKG Indonesia provided an online seasonal forecast on drought condition as well as short lead weather forecast, which are expected to be a source of information about future weather and climate condition. The future condition can be a guideline and reference to plan the adaptation options relates to livelihood activities. Moreover, it is an essential factor for an advanced warning of the drought impact (Pozzi et al., 2013). Unfortunately, the socio-demographic characteristics of the households in NTT which are dominated by poor families with low education background led to some problems in utilyzing the forecast product. It could be due to a technical problem as well as limited access to the source of information. Negatu & Parikh (1999) found that the adoption of seasonal forecasts is influenced by farmers' perception of the forecasts. It is a common case in rural areas that technology acceptance and adoption become a problem in the society as has been investigated by Amin & Li (2013) for the case in Bangladesh and China. Dey (2008) studied the use and appropriateness of mobile technology for rural farmers in Bangladesh.

This research investigates the factors that drive the intention of the households to use seasonal forecast to support their daily lives especially dealing with the livelihoods (e.g. livestock management and plantation calendar). By knowing the factors, a better strategy to increase the household intention to use the forecast would easily be formulated. The method used to investigate the importance factors is CART developed by Breimann et al. (1984). CART is a nonparametric approach in the class of machine learning method which has been applied in a wide range of applications. CART has been proven to be a method which is robust to some assumptions such as limited sample size, imbalancing response category, and other issues (Lewis, 2000).

Research and Methodology

The data analyzed in this paper were collected from a household survey in East Nusa Tenggara (NTT) in 2018. The survey was conducted by interviewing 300 households through an administered questionnaire. The data comprises of responses from seven districts in NTT i.e. Kupang, South Timor Tengah, North Timor Tengah, Ende, Nagekeo, East Flores and Lembata. The households were chosen randomly and allocated proportionally to the number of households in each districts. The important variables analyzed in this paper are listed in Table 1.

Table 1: Variables used in this paper

Variable	Category
Response (Y)	- Use seasonal forecast (Q4.4)
Predictors (X)	
Household socio-economic characteristics	- Gender of the household head (Q1.4A), - Age of respondent (Q1.6), - Level of education of household head (Q1.8) - Number of economically active member (Q1.11) - Number of members living outside sub-location (Q1.13) - Length of stay (Q1.15) - House ownership (Q1.16)
Household resources	- Type of House (Q2.1) - Household asset (Q2.2) - Key source of livelihood (Q2.4B)
Household agricultural activities	- Grow crops (Q3.1) - Water source for the crops (Q3.1F) - Livestock ownership (Q3.2) - Water sorce for the livestock (Q.3.2E)

Details of the questions, list of answers in the questionnaire and the data are available in Kuswanto (2019). Meanwhile, details of the data description can be seen in Kuswanto et al. (2019).

Data Analysis Technique

This paper applies CART to classify the households' intention to use seasonal forecast based on socio-economic characteristics, household resources and agricultural activities. Details about CART can be described as follows. Classification and Regression Trees (CART) was firstly introduced by Breiman et al. (1984). It is based on the algorithm of decision tree, and it is also the basis algorithm of bagging, random forecst and boosting decision trees.

The CART is performed in a binary tree which is built from root node splitted into leaf nodes. Given a new input, the prediction of the class response will be done by evaluating the specific input started at the root node of the tree. The input is selected by calculating split points of the variables which thus form a tree. A greedy algorithm (e.g. recursive binary splitting) is used to choose the input variable as well as splitting point. In this case, the algorithm will choose the split that minimizes the cost function.

For classification case, the cost function is measured by the Gini index which measure the purty of the leaf nodes, which is defined as follow

$$G = \sum_{k=1}^{nclass} p_k(1 - p_k)$$

Where G is the Gini index over all classes, p_k are the proportion of training instances with class k in the rectangle of interest. The value of G will be zero in the case of perfect class purity i.e all classes in the node are on the same type, while G will be equal to 0.5 in th ecase of worst purity i.e. the classes is splitted into 50:50. The Gini index for a binary classification case can be defined as

$$G = 1 - (p_1^2 + p_2^2)$$

The Gini Index for each node is calculated by weighting the Gini with the total number of case in the parent node. Therefore, the Gini value for the splitted node can be calculated as

$$G = \left(\left(1 - (g1_1^2 + g1_2^2)\right)\left(n_{g1}/n\right)\right) + \left(\left(1 - (g2_1^2 + g2_2^2)\right)\left(n_{g2}/n\right)\right)$$

where G is the Gini score at the splitted node, $g1_1$ is the proportion of case in group 1 for class 1, $g1_2$ for class 2, $g2_1$ for group 2 and class 1, $g2_1$ group 2 class 2, n_{g1} and n_{g2} are the total number of case in group 1 and 2 and n are the total number of case in the parent node. As previously mentioned that working with CART involves of determining the stopping criteria. Moreover, the CART performance is also influenced by the pruning in order to simplify the tree. It can be done by evaluating the effect of removing leaf node in the tree through a hold-out test.

Empirical Data and Analysis

Households characteristics

This subsection describes the socio-economic characteristics of the households participating in the survey. The information in the table revealed that most of the household heads in NTT are male with low educated people with only primary school level. The majority of the respondents have been living in the region more that 10 years. About 82% of them are working in agropastrolilism sector and majority of them grow crops and livestock.

Table 1: Descriptive statistics of the households

Variable	Category	Count	Percentage
Gender of the household head	Male	277	92.33%
	female	23	7.67%
Level of education of household head	None	22	7.3%
	Primary school	191	63.7%
	Secondary school	84	28.0%
	Post secondary	3	1.0%
Length of stay	< 1 years	10	3.33%
	1-5 years	7	2.33%
	5-10 years	10	3.33%
	> 10 years	273	91.00%
Type of House	Semi-permanent (mud wall/tin roof)	43	14.33%
	Semi-permanent (mud wall/grass roof)	183	61.00%
	Temporary (grass wall and roof)	74	24.67%
Having assets as a source of information (radio, TV, mobile phone and internet)	Yes	201	70.03%
	No	86	29.97%
Keysource of Livelihood	Pastoralism	28	9.33%
	Agropastoralism	246	82.00%
	Small scale business	8	2.67%
	Wage employment	18	6.00%
Grow crops	Yes	283	94.65%
	No	16	5.35%
Water Source for The Crops	Constant supply	12	4.33%
	Seasonal	265	95.67%
Livestock ownership	Yes	265	90.75%
	No	27	9.25%
Water Source for The Livestock	Constant supply	208	75.64%
	Seasonal	67	24.36%

Data analysis using CART

CART is used to predict the class of the response based on the households characteristics (predictors). The percentage of households that utilized seasonal forecast or not can be seen in Table 3. It is known that the proportion of households used forecast is significantly lower that the households did not use forecast i.e. only 51 (17.77%) households out of 300 (100%).

Table 3: Proportion of class response

Use seasonal forecast	Frequency	Percentage
Yes	51	17.77%
No	236	82.23%

The classification using CART analysis is carried out by using 10-fold cross validation in order to obtain the complexity parameter with the highest AUC (Area Under Curve). The results of the AUC with each corresponding parameter can be seen in Figure 1.

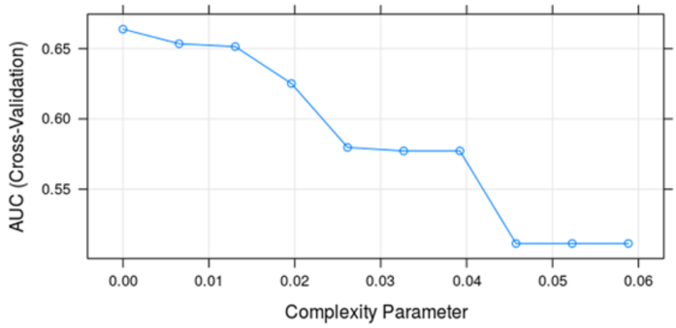

Figure 1: Complexity parameters and its corresponding AUC

We see that the highest AUC is obtained for the complexity parameter equals to zero. Meanwhile, higher parameters lead to lower AUC. Using 10-folds cross validation, the values of AUC, accuracy, sensitivity and specificity of the each fold can be seen in Table 4 below. We also see that the highest accuracy, sensitivity and specificity are obtained for different folds. However, as there is imbalance in the proportion of the response category, AUC represents the classification performance better than the others (Lewis, 2000). In this case, fold02 reached the highest AUC.

Table 4: Folds and accuracy measures

Fold	AUC	Accuracy	Sensitivity	Specificity
Fold01	0.6625	0.8276	0.9583	0.2000
Fold02	0.8217	0.8571	0.8696	0.8000
Fold03	0.6319	0.8000	0.9167	0.3333
Fold04	0.6217	0.8214	0.9130	0.4000
Fold05	0.6708	0.8276	0.9583	0.2000
Fold06	0.5391	0.7143	0.8696	0.0000
Fold07	0.6500	0.7931	0.9167	0.2000
Fold08	0.6478	0.8571	1.0000	0.2000
Fold09	0.7250	0.8276	0.9167	0.4000
Fold10	0.6667	0.8621	0.9583	0.4000
Average	0.6637	0.8188	0.9277	0.3133

The conclusion about the CART performance with respect to the fold setting should be done by looking at the average results. In this case, CART is able to classify each negative and positive class with very good performance i.e. 81.88% accuracy and 92.77% sensitivity. Meanwhile, the average AUC is 66.37% which is relatively high as it is greater than 50%. The CART with optimum complexity parameter can be seen in figure 2.

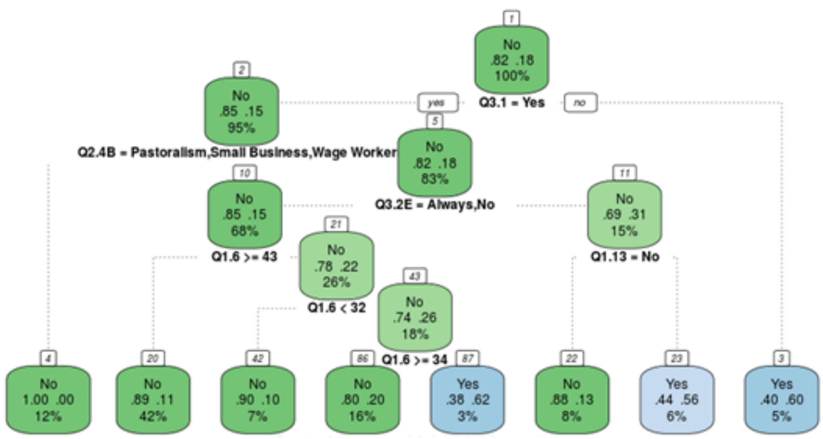

Figure 2: CART with optimum AUC

From the CART above, we see that households who did not grow crops (Q3.1), 40% of them tend to not using forecast with the percentage of 5% out of 300 households. Meanwhile, households who grew crops and agropatrolism is the key source of the livelihood are predicted not to use forecast, etc. CART obtained the variables sorted by its importance level that drive the intention of the household to use seasonal forecast.

Table 5: Important variables

No	Variable	Overall
1	Q1.6	100
2	Q2.2	72.062
3	Q3.1	51.518
4	Q3.1F	51.518
5	Q1.13	47.202
6	Q1.14	45.269
7	Q3.2E	36.263
8	Q2.4B	33.083
9	Q3.2	31.793
10	Q2.1	20.603
11	Q1.11	15.113
12	Q1.8	1.671
13	Q1.15	0
14	Q1.16	0
15	Q1.4A	0

Table 5 above listed the variables sorted from the most important to the least important. It is clear that Q16 is the most importat variables, while three last variables (Q1.15, Q1.16 and Q1.4A) with zero Gini index indicated no contribution to the classification of response.

Results and Discussion

From Table 5, we see that the most important variable that influences the usage of forecast is the age of the household head. Most of the households head in NTT are elder people (the average age of the households head is about 49 years old) who did not really able to access the forecast information. A they live in rural and remote areas, most of them have a very low education background, with about 63% only graduated from primary school. Considering the importance of the forecast, the policy makers or government officials might design a program to target young household members in order to provide them with the knowledge about how important the forecast as well as how to access the information.

Furthermore, the ownership of assets is also important. The survey revelaed that not all household have assets as source of information. To deal with this, the government need to ensure that all households have access to forecast information independent of whether the households related assests or not. It is due to the fact that most of the households are poor families with low capability to buy a new assets. A centered information managed by the village officer could be one of the solutions to solve the problem of asset limitation. The agricultural activities whether the household grow crops or not as well as the water source is also important factor. It is very clear and has been proven by many researches that crops are very sensitive to weather and climate condition. Households who grew crops and need a seasonal water will more likely use forecast infromation than households who did not grow crops. The forecast will be important to ensure that they have enough water to feed the plants. The information about the forecast is also very useful for managing the planting calendar.

The last three variables that did not influence the usage of forecast are gender of the household head, length of stay in the region and ownership of the house. It means that whatever the gender of the household head, it did not increase the incentive of using the forecast. The intention to use forecast is dominantly driven by other reasons i.e. the variables which are significant as explained above. The same fact is found for two other variables e.g. how long they have lived in the region and whether they owned the house or not. Meanwhile, the influence of other variables are moderate. In summary, this findings suggest that the policy makers need to focus on some significant variables to formulate the strategies for minimizing the drought risk.

Conclusions

In this study, it has been aimed to investigate factors influencing the households in East Nusa Tenggara (NTT) Indonesia to use seasonal forecast issued by the Meteorological Office (BMKG) Indonesia. NTT is a highly vulnerable region to drought, and hence, understanding the future weather and climate condition is a crucial issue. Household survey was conducted to interview 300 households in NTT asking about their intention to utilize the forecast information in supporting their livelihood. Furthermore, this research investigates whether the households characteristics influencing the intention. Using the Classification and Regression Tree (CART), this research found that the intention to use forecast is dominantly influenced by age of the household head, ownership of household assets, main source of income, crop and livestock activities and water source for livestock. From the analysis, we conclude that:

i. The majority of the households in East Nusa Tenggara did not use seasonal forecast.
ii. Intention to use seasonal forecast is dominantly influenced by the age of the household head, ownership of assets, agricultural activities (grow crops or not, source of water for crops).
iii. The intention to use seasonal forecast did not influenced by the gender of the household head, length of staying in the region and ownership of the house.

References

Amin, M.D. & Li, J. (2014). Applying Farmer Technology Acceptance Model to Understand Farmer's Behavior Intention touse ICT Based Microfinance Platform: A Comparative analysis between Bangladesh and China. *WHICEB 2014 Proceedings. Paper 31*. http://aisel.aisnet.org/whiceb2014/31

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Chapman &Hall/CRC.

Dey, B. L. (2008). The use and appropriation of the mobile telephony technologies by the rural Bangladeshi farmers. *Quest proceedings of the QUB AHSS Conference*.

Kuswanto, H. (2019). Household survey dataset on Drought Cycle Management (DCM) in East Nusa Tenggara-Indonesia, https://doi.org/10.7910/DVN/LGYJM3, *Harvard Dataverse, V1*.

Kuswanto, H., Hibatullah, F., Soedjono, E.S., Efendi, F. (2019). Survey data of household perceptions of drought, mitigation and adaptation practices in East Nusa Tenggara, Indonesia. Data in Brief. 103944. https://doi.org/10.1016/j.dib.2019.103944.

Labędzki L. (2016). Actions and measures for mitigation drought and water scarcity in agriculture. *Journal of Water and Land Development* 29, 3–10. https://doi.org/10.1515/jwld-2016-0007.

Maracchi G. (2000). *Agricultural drought–a practical ap-proach to definition, assessment and mitigation strate-gies.* In: Drought and drought mitigation in Europe. Eds. J.V. Vogt, F. Somma. Dordrecht. Kluwer Acad. Publ., 63–75

Negatu, W. & Parikh, A. (1999). The impact of perception and other factors on the adoption of agricultural technology in the Moret and Jiru Woreda (district) of Ethiopia. *Agricultural Economics* 21 (2), 205-216. https://doi.org/10.1016/S0169-5150(99)00020-1.

Pozzi, W., Sheffield, J., Stefanski, R., Cripe, D., Pulwarty, R., Vogt, J. V., Heim, R. R., Brewer, M. J., Svoboda, M., Westerhoff, R., van Dijk, A. I. J. M., Lloyd-Hughes, B., Pappenberger, F., Werner, M., Dutra, E., Wetterhall, F., Wagner, W., Schubert, S., Mo, K., Nicholson, M., Bettio, L., Nunez, L., van Beek, R., Bierkens, M., de Goncalves, L. G. G., de Mattos, J. G. Z., & Lawford, R.. (2013). Toward global drought early warning capability: expanding international cooperation for the evdelopment of a framework for monitoring and forecasting. *B. Am. Meteorol. Soc.*, 94, 776–785. https://doi.org/10.1175/BAMS-D-11-00176.1.

Lewis, R. J. (2000). *An Introduction to Classification and Regression Tree (CART) analysis.* Presented at the 2000 Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California.

Sena, A., Barcellos, C., Freitas, C., & Corvalan, C. (2014). Managing the health impacts of drought in Brazil. *Int. J. Environ. Res. Public Health* 11, 10737-10751. doi: 10.3390/ijerph111010737.

World Meteorological Organization (WMO) & Global Water Partnership (GWP) (2017). *Benefits of action and costs of inaction: Drought mitigation and preparedness – a literature review* (N. Gerber and A. Mirzabaev). Integrated Drought Management Programme (IDMP) Working Paper 1. WMO, Geneva, Switzerland and GWP, Stockholm, Sweden.

Zhang, L., Xiao, J., Zhou, Y., Zheng, Y., Li, J. & Xiao, H. (2016). Drought events and their effects on vegetation productivity in China. *Ecosphere* 7(12):e01591. https://doi.org/10.1002/ecs2.1591.