

PAPER • OPEN ACCESS

Predicted Rainfall and discharge Using Vector Autoregressive Models in Water Resources Management in the High Hill Takengon

To cite this article: I Ramli *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **273** 012009

View the [article online](#) for updates and enhancements.

Predicted Rainfall and discharge Using Vector Autoregressive Models in Water Resources Management in the High Hill Takengon

I Ramli¹, S Rusdiana², H Basri³ A A Munawar⁴, VAzelia⁵

^{1,4,5}Agriculture Engineering Departement, Syiah Kuala University

²mathematics and natural science Departement, Syiah Kuala University, Indonesia

³Soil Science Departement, Agriculture Faculty, Syiah Kuala University, Indonesia

Abstract. Rainfall is one of climate element that is very important to be studied. The relationship between rainfall and discharge needs to be known to see the size of the discharges so that it depends on the rain and the initial discharge conditions that occur. This study aims to predicted rainfall and discharge by using Vector Autoregressive (VAR), rainfall and discharge data in 2008-2015 is used to predicted rainfall for the next five years 2016-2020. While the data for 2016-2017 is used as a comparison for the predictive data obtained in this study. The results of the actual data accuracy test with predictive data from the VAR (Vector Autoregressive) model, which is done with the NSE (Nash-Sutcliffe efficiency) worth 0.9522. Predicted rainfall value with actual rainfall based on R^2 value (coefficient of determination) is 0.6584 or 66%. While the debit prediction test with actual discharge has a R^2 coefficient of determination of 0.0691 or 6.91%. It is because there are other factors that affect discharge other than rainfall. Rainfall in the Takengon plateau which is predicted to be unstable in the future. This condition is the same as the rainfall in the previous periods which fluctuated at the same point. Rainfall prediction results until 2020, has the highest peak season in December.

1. Introduction

Indonesia is a nation near the equator, where there is a tropical climate or a hot climate so it is vulnerable to drought. At present climate change is happening in the world. Indonesia is one of the countries that are experiencing the impact of climate change. According to the study, the air temperature in Indonesia increased by 0.30 °C since 1900 [1] this increase in temperature occurred throughout the season. Changes in weather and seasons are marked by an increase in rainfall (La Nina) that occurs in an area, while in other regions there is a decrease in rainfall (El Nino) of 2-3% [1].

This climate change has caused floods and droughts that are now happening in Indonesia. A significant drop in rainfall during the dry season, including the El Nino effect, has had a negative impact on food crop production. Drought is a threat that often disrupts agricultural production systems. In recent years drought has not only increased in size and intensity, but also has more and more impacts and distribution of areas affected by drought [2]. Uncertainty of rainfall and rising temperatures quickly emerge as one of the most serious global problems affecting many sectors. It is considered one of the most serious threats to sustainable development with adverse impacts on the environment, human health, food security, economic activity, natural resources and physical infrastructure [3].

According to [4], climate change has begun to occur by looking at one of the climate variables, namely temperature. The average air temperature for twenty years (1992-2011) was 26.77 °C. For the average annual temperature in the period 1992-2001 was 26.41 °C and the average period in 2002-2011 was 27.13 °C. There is an annual average temperature that continues to increase in certain years.



Significantly positive temperature increases from the Kendall analysis test results and will affect water management practices. Increases in temperature and rainfall which result in a little drought, thus posing a serious threat to food security.

The atmospheric conditions in the area with high levels of non-linearity are more difficult to predict than those in high latitudes. This becomes the base point for predicting or predicting the coming rainfall with a certain time series. The statistical approach in time-derived analysis is done by using a statistical model to explain the dynamic behavior of a time slot. A multivariate statistical model is required for a condition in which an event is influenced by more than one variable. The statistical model for rainfall and discharge prediction in this study uses Vector Autoregressive (VAR) in Central Aceh District.

2. Methodology

This research was carried out in Pegasing District, Central Aceh Regency (figure 1) site selection was carried out purposively because the location had rainfall and discharge data for analysis using the VAR model the data used was rainfall and discharge in 2008-2015. While the 2016-2017 data is used as a comparison for predictive data obtained in this study. Monthly rainfall data taken at the Climatology and Geophysics Agency of Aceh Besar Climatology Station from 1992 to 2017 and monthly debit data taken at BPDAS Krueng Aceh in 2008 to 2016.

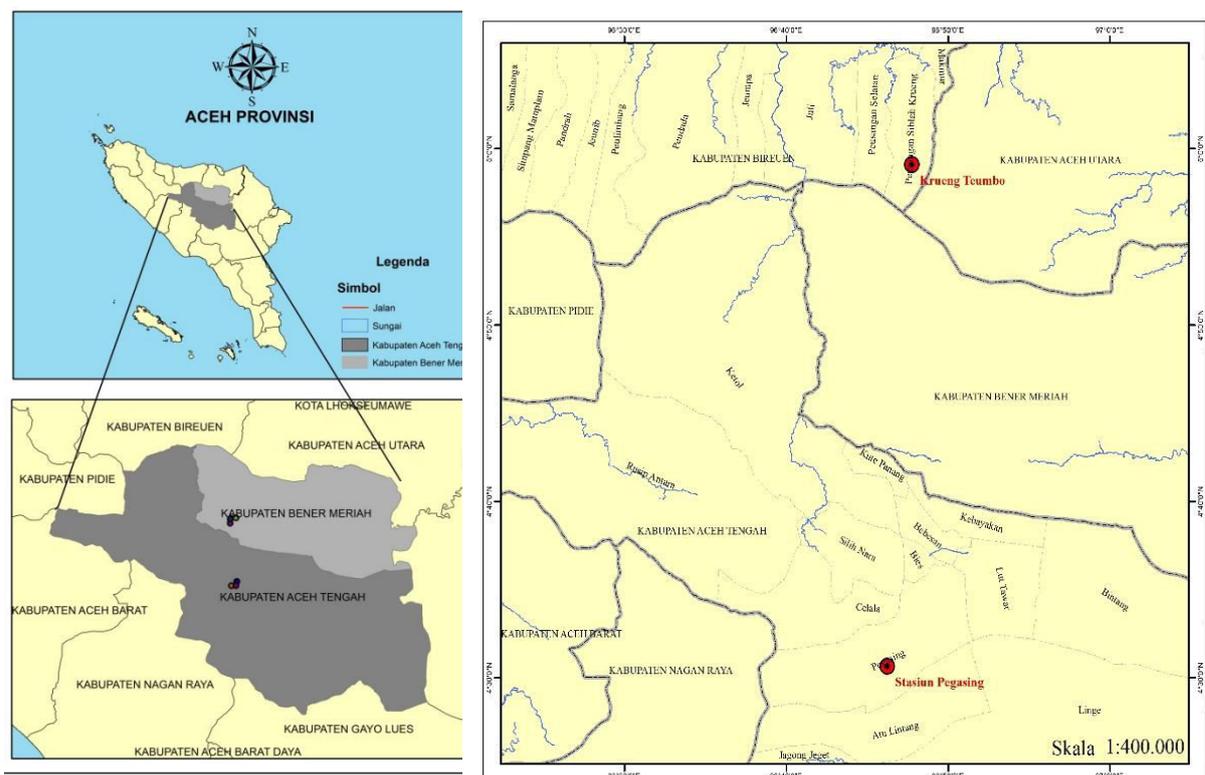


Figure 1. Location of study area research

The research method used in this research is VAR (Vector Autoregressive) analysis. The steps of the research are as follows:

- a. Stationary test was conducted on research data. If the data is stationary then the data analysis process uses VAR analysis. If the data is not stationary, a differencing process will be carried out until stationary data is obtained. There are several ways that can be done to measure data stationarity, one of which is to use the Augmented Dickey Fuller (ADF) test [5]. The following ADF test statistics:

$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} + \beta_i \sum_{i=1}^p \Delta Y_{t-i+1} + \varepsilon_t \quad (1)$$

where

ΔY_t = the form of the *first difference* variabel Y

α_0 = Intersep

Y = Variabel tested stationary

p = lag length

The hypothesis for the ADF test is as follows: H0: Data is not stationary H1: Stationary data
If p-value < α then H0 is rejected.

- b. Stationary data after differencing, it is necessary to do cointegration test first. If there is no cointegration (long-term relationship) between variables, then data analysis uses VAR (Vector Autoregressive) analysis. However, if there is cointegration between variables, then the next stage of analysis uses a Vector Error Correction Model (VECM) analysis.
- c. Next, determining the lag length that will be used in the VAR model. In this study, the lag length is determined by looking at the minimum values of FPE, AIC, SIC and HQ. The criteria that can be used to determine the amount of optimal lag include Final Prediction Error Correction (FPE), Akaike Information Criterion (AIC), Schwarz Information Criterion (SIC) and Hannan-Quinn Information Criterion (HQ), which have a minimum value between various lags submitted. Mathematically, the AIC, SIC and HQ values can be searched using the following formula:

$$AIC = -2 \left(\frac{1}{T} \right) + 2(k + T) \quad (2)$$

$$SIC = -2 \left(\frac{1}{T} \right) + k \frac{\log(T)}{T} \quad (3)$$

$$HQ = -2 \left(\frac{1}{T} \right) + 2k \log \left(\frac{\log(T)}{T} \right) \quad (4)$$

The meaning of the symbols used in equations (2) to (4) are:

l = The value of the log likelihood function is the same as that

$-\frac{T}{2} \left(1 + 2\pi + \log \left(\frac{\varepsilon' \varepsilon''}{T} \right) \right)$; $\varepsilon' \varepsilon''$ is sum of square residual

T = Number of observations

k = Estimated parameters

- d. The general model of VAR is as follows:

$$Y_t = \alpha_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t \quad (5)$$

or

$$Y_t = \alpha_0 + \sum_{n=1}^p A_n Y_{t-n} + \varepsilon_t \quad (6)$$

- f. Then structural analysis is carried out as follows:

- Impulse Response Function (IRF), to see how much independent variables are affected by shocks that occur in the dependent variable, some time to come.
- Variance Decomposition (VD), to see how much the independent variable contributes to the dependent variable some time to come.

- g. Forecasting based on the VAR model obtained.

Model Accuracy test

The prediction model is then tested for accuracy or validation using a formula correlation coefficient and Nash-Sutcliffe efficiency (NSE) [6]; [7]; [8].

$$R = \frac{\sum_{i=1}^n (CH_i^{obs} - CH_{obs}^{mean})(CH_i^{sim} - CH_{sim}^{mean})}{\sqrt{\sum_{i=1}^n (CH_i^{obs} - CH_{obs}^{mean})^2 (CH_i^{sim} - CH_{sim}^{mean})^2}} \quad (7)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (CH_i^{obs} - CH_i^{sim})^2}{\sum_{i=1}^n (CH_i^{obs} - CH_i^{mean})^2} \quad (8)$$

Where:

CH_{obs}	= actual data
CH_{sim}	= Prediction data
$CH_{mean\ obs}$	= actual mean data
$CH_{mean\ sim}$	= Prediction mean data

Then the most accurate and accurate prediction results are selected with the actual data in the field. In general, the simulation model can be seen satisfactory if the value of $NSE > 0.50$

3. Result and Discussion

The highest dominance in the area of Central Aceh Regency is 1000-2000 masl with a total area of 287,387.21 Ha or 62.65%. So that the region of Central Aceh Regency has diverse agricultural and plantation commodities. While the lowest altitude is at 125-1000 meters above sea level several places are possible to be planted with crops such as coffee, rice, pulses, and horticulture. Central Aceh Regency has a mountainous and hilly topography with an average height varying between 200 - 2,600 meters above sea level.

In general the slope classification of Central Aceh Regency with 14 sub-districts has a land slope of $> 40\%$, 25-40%, 15-25%, 8-15% and 0-8%, divided into very steep, steep, rather steep, sloping and flat. Based on this class the slope or slope of the land in the Central Aceh District is dominated by slope with 15-25% and an area of 208,440.37 Ha or 45.44%. So that with a slope of 15-25% that can only be planted or become land for plantation business. The lowest land slope is with the 0-8% range and 109,434.07 Ha wide area or 23.86% with the flat area. So, the total area in total is 458,707.47 Ha.

Rainfall is an important component in the hydrological process in agriculture, as it is known that plants need water for their primary needs both in quality and quantity. In plant growth, water needs are based on the weather and climate that occur in an area. One of the factors that influence plant growth is rainfall in the area. Rainfall is very important to be studied as an object of research, because rainfall information is very useful in changing weather and climate, one of which is agriculture. Besides that, rainfall also has fluctuating properties on differences in cropping patterns in the agricultural sector which can cause farmers' products to decline.

The large amount of rainfall into the catch area will determine the flow discharge in the watershed, both surface flow, subsurface, base flow and river flow. Rainfall in the Takengon Plateau is very volatile in every year. However there are several months in different years of very high rainfall followed by high flow of water flow. In the month of high rainfall there was a flood in the district of Central Aceh District Linge that occurred in two villages namely Kute Riem and Kute Keramil. Natural disasters are making as many as 57 houses residents submerged in floods. In addition, the disaster damaged hundreds of hectares of rice fields. The incident occurred in the year 2009, the cause

is overflowing water of the river Nangka located in the district of Central Aceh with maximum rainfall in that year.

Rainfall characteristics include matters involving annual, seasonal, monthly, or daily distribution, intensity, duration of rain, and frequency of rainy days [9]. In the tropics rainfall is a decisive factor since most agricultural production is produced from dry land farming, where the water required comes from rainfall, so rainfall characteristics are a factor to be considered especially in relation to plants.

Multivariate statistical models will be needed in conditions where an event is affected by more than one variable. One model that is often used is the Vector Autoregressive System (VAR). VAR is one of multivariate time series analysis and useful to see the relation between variables [5]. Stationary test is the first test conducted to determine whether the data is diverse. The results of the rainfall predictions of variable temperature, wind speed, humidity are obtained rainfall = 0.07537952 (rainfall.11) - 4.84116742 (Temperature.11) + 6.36430062 (humidity.11) - 0.11016965 (wind speed.11) + 0.04866761 (rainfall.12) + 3.66042192 (Temperature.12) - 1.28623713 (humidity.12) - 0.29067619 (wind speed.12) - 0.08908943 (rainfall.13) - 4.16258553 (Temperature.13) - 5.90352425 (rainfall.13) + 4.00617726 (wind speed.13) + 300.58327741

If changes in rainfall one month ago increased by 1 mm, will cause changes in rainfall this month increased by 0.07537952 mm. If changes in rainfall two months ago increased by 1 mm, will cause changes in rainfall this month increased by 0.04866761 mm. If changes in rainfall three months ago increased by 1 mm, will cause changes in rainfall this month decreased by 0.08908943 mm.

If the change in temperature one month ago increased by 1°C , it would cause changes in rainfall this month to decrease by 4.84116742 mm. If the change in temperature two months ago increased by 1°C , it would cause changes in rainfall this month to increase by 3.66042192 mm. If changes in temperature three months ago increased by 1°C , this would cause changes in rainfall this month to decrease by 4.16258553 mm.

If the change in humidity one month ago increased by 1 RH, it will cause changes in rainfall this month increasing by 6.36430062 mm. If the change in humidity two months ago increased by 1 RH, it will cause changes in rainfall this month decreased by 1.28623713 mm. If the change in humidity three months ago increased by 1 RH, it will cause changes in rainfall this month decreased by 5.90352425 mm.

If changes in wind speed one month ago increased by one knot, this would cause changes in rainfall this month to decrease by 0.11016965 mm. If changes in wind speed two months ago increased by 1 knot, this will cause changes in rainfall this month decreased by 0.29067619 mm. If changes in wind speed three months ago increased by 1 knot, this will cause changes in rainfall this month increasing by 4.00617726 mm.

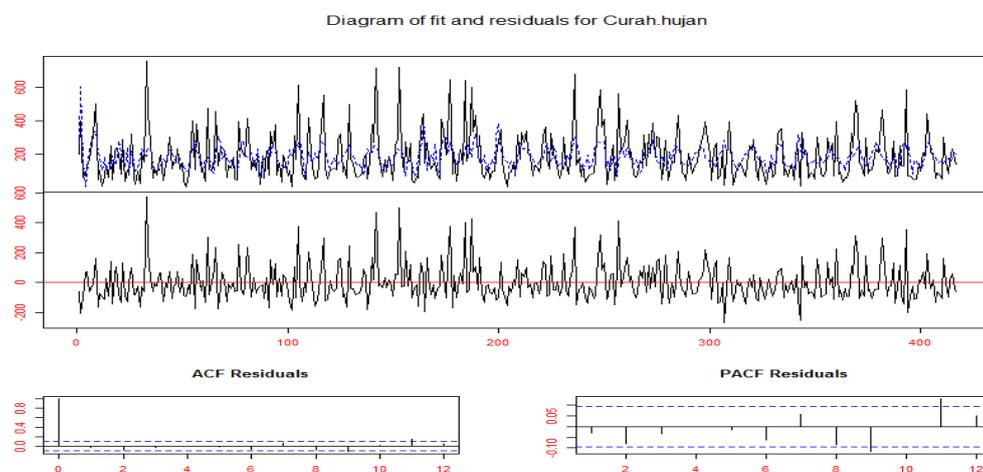


Figure 2. Graph of residual rainfall value from the VAR (3) period model

January 1981 - December 2015.

Plot IRF (Impulse Response Function)

RF analysis aims to see the effect (effect) of each variable if given shock (shock). The impact of the shock of one variable on another can be traced through IRF.

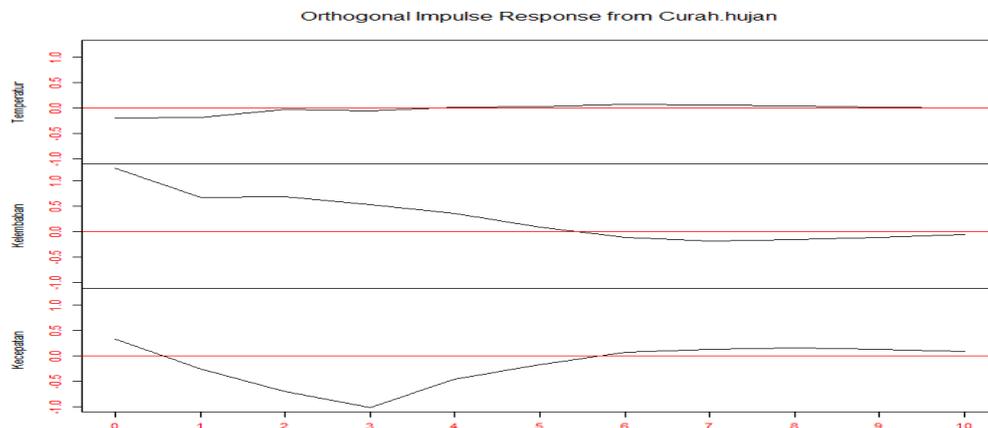


Figure 3. RF rainfall graph

The graph above shows that if there is a shock (shock) in the rainfall variable, it does not cause a shock to the temperature variable, so that the line is right or close to the equilibrium line (Equilibrium), so it does not take long for the variable temperature to stabilize or be in the equilibrium line. As for the rainfall variable to the humidity variable, if there is a shock to the rainfall variable, it takes about 5 months for the humidity variable to return to / close to the equilibrium line. For rainfall variables to speed variables, if there is a shock to the rainfall variable, it will take about 6 months for the wind speed variable to return to / close to the equilibrium line.

Variance decomposition (VD) aims to separate the impact of each of these variables individually, on the response received by a variable. So, variance decomposition can provide information about the relative role of each shock, to the variables in the VAR (Table 1). The variance decomposition analysis shows that in the span of 12 months, the shock to itself results in weaker rainfall fluctuations. While shocks to temperature, humidity and wind speed cause fluctuations in rainfall to increase

Table 1. The variance decomposition analysis at High Hill Takengon

Month	Rainfall	Temperature	Humidity	Wind Speed
1.	1.000	0.000	0.000	0.000
2.	0.962	0.011	0.026	2.633
3.	0.949	0.011	0.038	2.775
4.	0.903	0.013	0.045	3.737
5.	0.873	0.013	0.064	4.827
6.	0.858	0.014	0.075	5.150
7.	0.854	0.015	0.077	5.118
8.	0.854	0.016	0.078	5.150
9.	0.853	0.016	0.077	5.238
10.	0.853	0.016	0.077	5.316
11.	0.851	0.016	0.078	5.366
12.	0.851	0.016	0.078	5.383

Figure 4 is a Climate Data Prediction Plot for the next 5 years (60 Months) using the VAR Model (3) Based on the results of the prediction of rainfall in High Hill Takengon until 2020, the highest peak

(peak season) rainfall occurs in December. The visualization of rainfall at location High Hill Takengon in January 2016 to December 2020 based on the results of the forecasting.

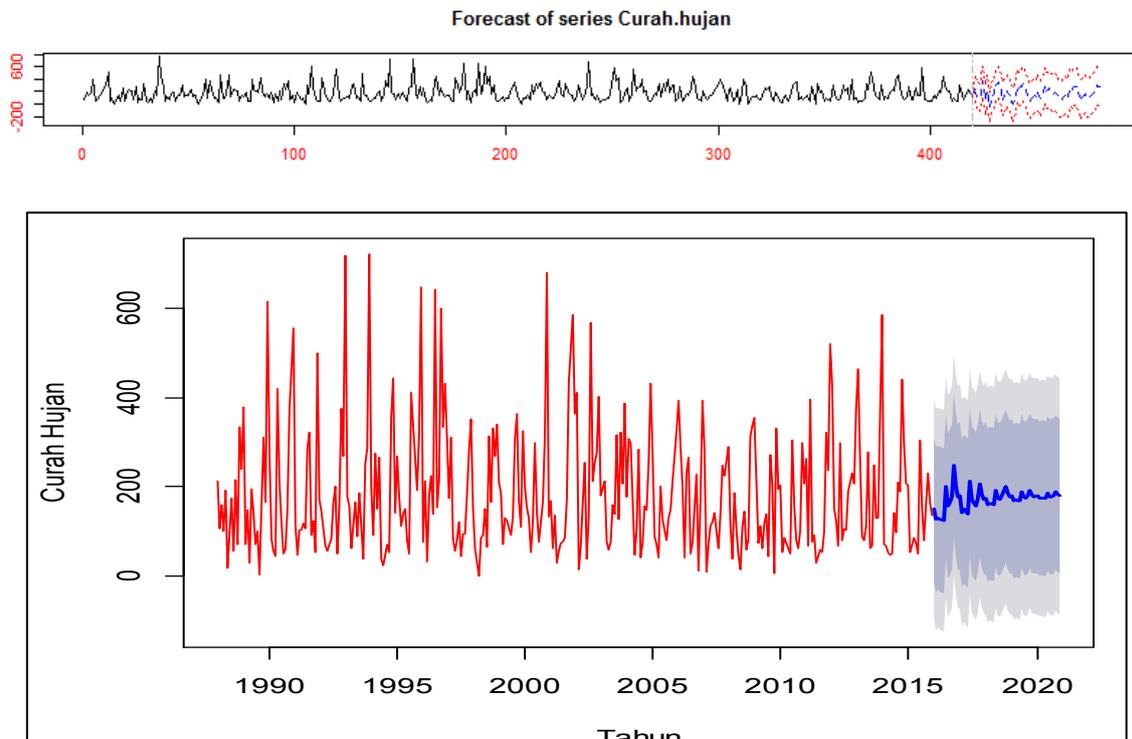


Figure 4. Rainfall Prediction for 2015-2020 at High Hill Takengon

The visualization above shows that rainfall in location High Hill Takengon is predicted to be unstable in the future. This condition is the same as actual rainfall in previous periods which tends to fluctuate at the same point. The blue line is the rainfall prediction data at location High Hill Takengon in January 2016-December 2020. Based on the results of rainfall predictions, the lowest rainfall rate occurred in July 2016, reaching 79.5 (mm). The highest rainfall rate occurred in October 2017 which reached 207.4 mm. The rainfall season based on the results of the prediction occurs in October, November and December. On the part of the lower limit of the confidence interval the value of the minimum predictive value while the upper limit of the interval trusts the maximum predicted value. The picture looks light gray and dark gray. This illustrates dark gray showing 80% of the highest predictive limit or under the opposite. While light gray shows 90% predictions that might occur.

Output of accuracy Models

VAR model on Pegasing for the relationship between rainfall and discharge hypothesis H_0 accepted on probability value $> 5\%$. This shows that the discharge is not only affected by rainfall. While the rainfall is influenced by latitude, altitude, distance from the sea, wind direction, mountain rows, differences in terrestrial and ocean temperatures and land area factors. Factors affecting rainfall are the humidity, air, temperature and wind speed [10] [11] suggests climate factors that are strongly associated with the biophysical region are precipitation and temperature.

Vector Autoregressive (VAR) analysis to predict rainfall and discharge that occurs in Central Aceh District, using rainfall data from 2008-2015. From the results, it was found that from the 2008-2015 data, the prediction can only be done in two years, from 2016-2017. This happened because of the limited data available to predict rainfall. Then the accuracy or data validation of the VAR model is tested using the Nash-Sutcliffe efficiency (NSE). The results obtained from NSE are 0.9522. The R^2 which influences the amount of predicted precipitation with actual rainfall is 0.6584 or 66%. The R^2

that most influences from the graph is the 2nd order R^2 graph, where the resulting value is 0.0691 or 6.91%. (Figure 5 and 6). Based on the study, This is because there are other factors that affect discharges other than rainfall. Based on [12] it said that there is a significant correlation between climate, and soil type on river discharge. [13] said that watershed geomorphic characteristic of drainage density, slope variability and colluvium were also shown to have significant impact on stream flow and baseflow. Stream flow quantity and quality are dependent on the physical characteristic of a catchment [14]. The effect of land use at seasonal and monthly runoff depths showed significant variation [15]. Overall, the effects of changes in land use on river flows in the summer are clear. Therefore it was concluded that forest change is a major factor in the variation of river flow in the summer. It shows that the increase in agricultural land can increase the flow in the rainy season and reduce the flow in the dry season but when the forest area is reduced to near zero, the peak flow and average flow will increase. The time series model of regression should consider the relationship between the dependent variable and independent variables. However, in many applications the transfer function and regression model may not be appropriate [16]. Rainfall in the Takengon plateau which is predicted to be unstable in the future. This condition is the same as the rainfall in the previous periods which fluctuated at the same point. Rainfall prediction results until 2020, the peak season rainfall occurs in December.

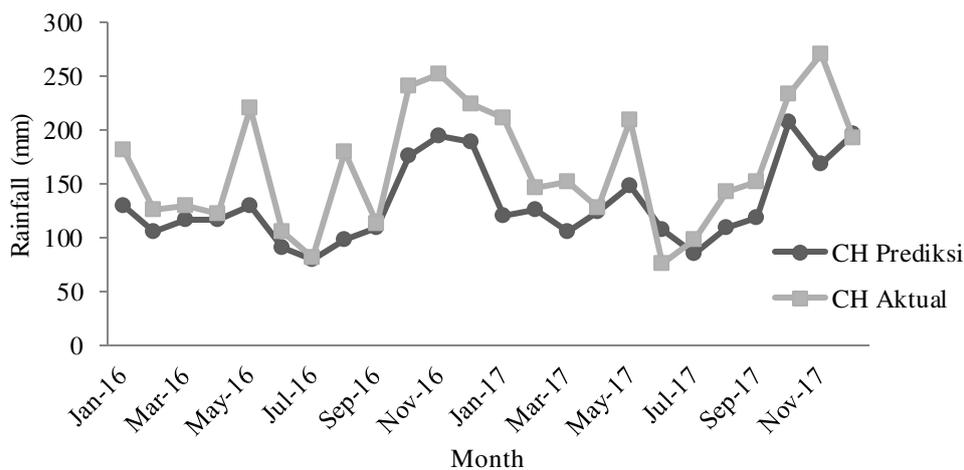


Figure 5. Comparative data on Prediction Rainfall and Actual Rainfall

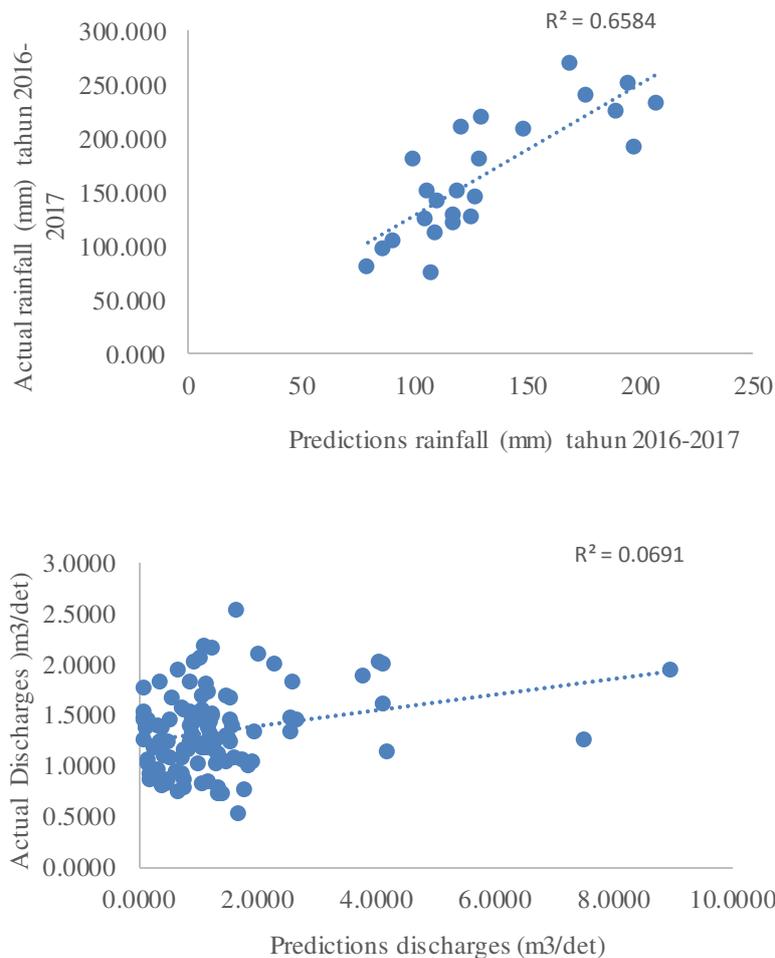


Figure 6. Coefisien determinate (R^2) for Rainfall, discharge from actual and prediction data

4. Result

- Rainfall prediction in the Takengon plateau with Vector Autoregressive (VAR) model is obtained from the validation results between prediction rainfall and actual rainfall for 2016-2017 with Nash-Sutcliffe efficiency (NSE) test of 0.9522 (good performance). While the coefficient of determination (R^2) that affects the amount of predicted precipitation with actual rainfall is 0.6584.
- The value of the coefficient of determination for prediction discharges with the actual discharges is 0.0691.
- Rainfall in the Takengon plateau which is predicted to be unstable in the future. This condition is the same as the rainfall in the previous periods which is fluctuated at the same point. Rainfall prediction results until 2020, the peak season rainfall occurs in December.

References

- [1] Hulme, M. and N, Sheard, 1999. *Climate Change Scenarios for Indonesia*. Leaflet CRU and WWF. Climatic Research Unit. UEA, Norwich, UK
- [2] Hadiyanto, S. 2007. *Pola Tingkat Kerawanan Kekeringan di Jawa Tengah*. Tesis; Departemen Geografi FMIPA UI
- [3] Obert Jiri, Paramu L. Mafongoya¹, Chipo Mubaya & Owen Mafongoy. Seasonal Climate Prediction and Adaptation Using Indigenous Knowledge Systems in Agriculture Systems in Southern Africa: *A Review Journal of Agricultural Science*; Vol. 8, No. 5; (2016) ISSN 1916-9752 E-ISSN 1916-9760.

- [4] Ichwana, A. Achmad, and S. Chairani. 2015. *Climate trends and dynamic change of land use patterns in Krueng Peusangan Watershed, Aceh-Indonesia*. International Symposium of Geoinformatics (ISYG) 2-3 Desember 2015 di Universitas Brawijaya Malang
- [5] Hartini, S., M. P., Hadi., Sudibyakto dan A. Poniman.. Application of Vector Autoregression Model for rainfall-river discharge analysis. *Jurnal Forum Geografi*. 29: (2015) 1-10
- [6] Moriasi, D.N, J.G. Arnold, M.W. Van Liew, R.L Bingner, R.D Harmel, dan T.L Veith 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Jurnal American Society of Agricultural and Biological Engineers*. ISSN 0001-2351.50 (3): 885-900
- [7] Nash JE, dan J.V. Sutcliffe 1970. River flow forecasting through conceptual Models 1, a Discussion of Principles, *J. Hydrol.* 10 (1):282-290
- [8] Nasution, Z., Ichwana, Sumono and Delvian. 2013 Land Use Change and Simulation of Hydrological Response using Integrated NRCS Model and Base flow for Ecohydrological Concept at Krueng Peusangan Watershed, Aceh, Indonesia, *journal of Environmental Science and Water Resources, journal citation report*, ISSN 2277 070 2(7): 210 - 220, August 2013
- [9] Nieuwolt, 1997 *Tropical Climatology an Introduction to the Climate of the Low Latitude*. John Wiley and Son, New York
- [10] Larsen, T.H., Brehm, G., Navarrete, H., Franco, P., Gómez, H., Mena, J.L., Morales, V., Argollo, J., Blacutt, L., Canhos, V., 2011 *Range shifts and extinctions driven by climate change in the tropical Andes: synthesis and directions*. In: Herzog, S.K.
- [11] Rabatel, A., Francou, B., Soruco, A., Gomez, J., Caceres, B., Ceballos, J.L., Basantes, R., Vuille, M., Sicart, J.E., Huggel, C., Scheel, M., Lejeune, Y., Arnaud, Y., Collet, M., 2013 Current state of glaciers in the tropical Andes: a multi-century perspective on glacier evolution and climate change. *J. Cryosphere* 7, 81–102. (2013). <http://dx.doi.org/10.5194/tc-7-81-2013>
- [12] Krakauer, N.Y and M. Temimi. 2011. Stream recession curves and storage variability in small watersheds, *Hydrol. Earth Syst. Sci.*, 15, 2377–2389, 2011 www.hydrol-earth-syst-sci.net/15/2377/2011/ doi:10.5194/hess-15-2377-2011.
- [13] Price, K, Jackson CR, dan Parker AJ. 2011. Effects of watershed land use and geomorphology on stream low flows during severe drought conditions in the southern Blue Ridge Mountains, *GA and NC, USA Water Resources Research* 47:W02516.
- [14] Abeyou Wale Worqlul, Essayas K. Ayana, Amy S. Collic, Tammo S Steenhuis, 2015, Effect of Physical Catchment Characteristics on River Flow: The Case of Ribb and Gumara Rivers in Upper Blue Nile. International Soil & Water Assessment Tool Conference, Sardinia, Italy, DOI: 10.13140/RQ.2.1.2288.3680
- [15] Patts M.A. Odira, Dr. M.O. Nyadawa, B. Okelloh Ndwallah, Ms. Nelly A. Juma, Mr. John and P. Obiero, 2011. Impact of Land Use /Cover dynamics on Streamflow: A Case of Nzoia River Catchment, Kenya, Nile Basin, *Water Science & Engineering Journal*, Vol.3.
- [16] Wei, W.W. 2006 *Time Series Analysis Univariate and Multivariate Methods*. Addison-Wesley Publishing Co, USA.

Acknowledgement

The authors would like to deliver their gratitude to the Agency for Meteorology, Climatology, Geophysics (BMKG) and BPDAS Krueng Aceh and this research was funded by Syiah Kuala University of 2018 Fiscal Year.