ISSN No:-2456-2165

Crop Yield Modeling and Estimation of Paddy: A Case Study of land Pooling Area Ward No.1 and 3 of Bandipur Rural Municipality of Tanahun, Nepal

Milan KC, supervised byAssot. Professor Dr. Krishna Prasad Bhandari, and Asst. Professor Er. Bikash Sherchan
Institute of Engineering, Pashchimanchal Campus, Pokhara

Abstract:- Rice is largely grown type of crop in Nepal and has an important role in food security and nutrition. Therefore proper and timely monitoring, forecasting, and prediction are necessary for planning and management purposes for governmental and non-governmental bodies that have as been working for food security. So the forecasting and estimation of crops and monitoring of different phases of the crop are necessary for food and crop management. For proper monitoring, remote sensing imagery and NDVI series are capable of the identification of crop growth and health as well as crop yield prediction model development. This research reflects a model for paddy monitoring using Sentinel 2 imagery of every 5 days interval by extracting NDVI series. The land pooling area of seratar, and Nahala, Bandipur Rural Municipality of Nepal has been selected as a study area for understanding and analyzing the NDVI value of different phonological stages of paddy with different land management factors like water level, level of damages, soil type, fertilizers used, date of transplantation, amount of seed, sowing date, etc.

Google Earth Engine (GEE) is an ideal platform that is used to study and extract the NDVI of multiple sentinel images of study areas. Importantly, GEE reduces time and space for data processing and analysis. The hectic process of getting satellite imagery from a different sources, the downloading process of large images, their correction and classification collectively known as remote sensing processes can be resolved with this open source platform.

By using the time series NDVI value, we can assess the correlation between the NDVI values and land management factors. Along with the paddy monitoring, the rice yield prediction and estimation model is developed using the regression model. The model classifies only those parameters which are highly correlated to the NDVI values. And from that, we are, able to develop the crop yield prediction model. Besides that, we have prepared the rice distribution map, LULC map, and agriculture map from the data that we collected from the field survey. And there is also a socio-economic analysis of crop yield model in the scenario of world and Nepal. It reflects its importance in planning, policy-making maintaining the demand-supply chain in distribution of agriculture production etc.

Keywords:- *Estimation, LULC, time series, Regression, NDVI, GEE.*

I. INTRODUCTION

Rice is the main staple crop of Nepal and has an important role in food security and nutrition. Therefore proper and timely monitoring, forecasting, and prediction are necessary for planning and management purposes for governmental and non-governmental bodies that have as been working for food security. So the forecasting and estimation of crops and monitoring of different phases of the crop are necessary for food and crop management. For proper monitoring, remote sensing imagery and NDVI series are capable of the identification of crop growth and health as well as crop yield prediction model development. This research reflects a model for paddy monitoring using Sentinel 2 imagery of every 5 days interval by extracting NDVI series. The land pooling area of seratar, and Nahala, Bandipur Rural Municipality of Nepal has been selected as a study area for understanding and analyzing the NDVI value of different phonological stages of paddy with different land management factors like water level, level of damages, soil type, fertilizers used, date of transplantation, amount of seed, sowing date, etc.Google Earth Engine (GEE) is an ideal platform that is used to study and extract the NDVI of multiple sentinel images of study areas. Importantly, GEE reduces time and space for data processing and analysis. The hectic process of getting satellite imagery from a different sources, the downloading process of large images, their correction and classification collectively known as remote sensing processes can be resolved with this open source platform.

By using the time series NDVI value, we can assess the correlation between the NDVI values and land management factors. Along with the paddy monitoring, the rice yield prediction and estimation model isdeveloped using the regression model. The model classifies only those parameters which are highly correlated to the NDVI values. And from that, we are, able to develop the crop yield prediction model. Besides that, we have prepared the rice distribution map, LULC map, and agriculture map from the data that we collected from the field survey. And there is also a socio-economic analysis of crop yield model in the scenario of world and Nepal. It reflects its importance in planning, policy-making maintaining the demand-supply chain in distribution of agriculture production etc.

Paddy (rice) is the second most widely planted grain crop in the world. Almost half the world's population use rice as the main staple food. Nearly 70% of Nepal's population depends directly on agriculture, and the sector contributes more than a quarter of the country's GDP –

ISSN No:-2456-2165

with rice making up 21% of that. Till as late as 1985, Nepal used to be a net exporter of rice, and during the 1960s the country was exporting up to $45 million worth of rice to India every year. As the facts have turned, in 2015 Nepal has to import 531,000 tons of rice worth $210 million from India (Nepal's rice economy, 2019). According to the MoALD 2018/19 report, rice accounts for nearly 70 % of the total arable land which is about 10% of the area of Nepal. The average production of Nepal is 3.506 tons per hectare.Even though educated youth-, and children of farmers are either migrating or moving away from the land, agriculture is still the mainstay of Nepal's economy (Kumar, 2019). Lack of techniques and mechanization in the agriculture system the migration of youths, educated people, and farmers will continue which directly affects agriculture production. So revolution is necessary for system and mechanization.

Reliable and timely monitoring and mapping of rice is of global importance (Shao, et al., 2010). Mapping of paddy crops helps in the estimation of the water supply required for irrigation, net production, and yield estimation. Because Statistical data from local governments have always been used to calculate the subsidies. However, large discrepancies are often observed between different datasets from different governmental departments or agencies. Therefore, reliable methods need to be developed to extract accurate and reliable rice crop acreage estimates to support governmental agricultural policy (Qiangzi Li, 2014).

Monitoring vegetation is important to the entire environment because of the carbon stored in their surrounding ecosystems. Monitoring of paddy crop in a country also aids in assessing the national food security, for planning and sustainable management of available resources (umma, Nelson, Thenkabail, & Singh, 2011). It is only possible with the help of remote sensing products.Remote sensing technology has been widely used to map rice fields worldwide. For already some time, the attributes of RS data and the suite of tools offered by this technology have triggered the search for attractive non-destructive approaches to obtain important data in agriculture.Numerous remote sensing studies have been involved in the mapping and analysis of relative vegetation cover at various scales, utilizing indices based on the optical spectral behavior of vegetation. Although optical remote sensing can be considered a relatively powerful mapping tool, there are recognized limitations of these data due to cloud interference, atmospheric attenuation, and some constraints in its use for vegetation discrimination (Sofa, 1-58).The normalized difference vegetation index (NDVI) is by far the most widely used in the literature, and is advantageous for studying historical changes; however, it is sensitive to canopy background variations and saturates in relatively high-vegetated areas (Gu, Wylie, Howard, Phuyal, & Ji, 2013).Predicting rice yield is a diagnostic technique that can enhance resource inputs, minimize the risk of crop failure, decrease yield productivity and ultimately promote food security.

II. METHODOLOGY AND DATA ANALYSIS

A. Data and Software Used

For this research, a series of sentinel-2 images of 10m spatial resolution taken between the date 14 July to 28 November of 2021 was used to extract the NDVI time series in a Google Earth Engine platform, jointly ArcGIS 10.5 and QGIS 3.X is used to analysis and preparation of different maps. The data on crop management (paddy) and its yield of last year (2021) was collected from the farmers' through interviews and questionnaires.The Garmin GPS map 62s device was used to trace the polygon of respective farmer's parcels. The field was digitized using a hand held GPS with an accuracy about 4m. The Power BI Desktop and Excel was used for data analysis and visualization.

Field data and land management data collected from the field include yield records of an individual parcel, which were collected from the measured weight of the harvested paddy crop. A total of 40 yield records of the paddy parcels for 2021 were obtained from farmers. While 30 of them were used in regression analysis as dependent variables to get yield models, the remaining 10 called "Test Parcels" were used to check the performance of generated models using root mean squared errors (RMSE) values.

B. Image Processing and Classification

The methodology of this study is based on the machine learning approach in which Google Earth Engine uses Sentinel-2 satellite images from where we can extract the NDVI time series of study area.For better rice crop yield estimation, considering remote sensing product or data may logically provide a better understanding of the crop growth of different phases to produce more accurate yield predictions and estimation. Sentinel 2 images of the study area was used to extract the NDVI values of different stages of rice in 2021.The key phases of the rice crop cycle, i.e., start of the season (SoS), the peak of season (PoS) and end of the season (EoS), in the study region of this research correspond to mid-July, mid-September and mid-November respectively. As we already know that, S2 sensors provide a total of 13 spectral bands with a spatial resolution ranging from 10 m to 60 m. Among these spectral bands, the classical RGB and near-infrared (NIR) bands with 10 m spatial resolution are dedicated to land applications.The extracted NDVI values are analyzed with different land management parameters like soil type, date of the plantation, fertilizers, affect caused by disease, date of harvesting, etc., after analyzing the NDVI with these parameters we can able find out how the yield is correlated with them.

The land use land cover map has been prepared using the Landsat-8. The area classified as agriculture is masked out and converted to polygon. This agriculture polygon is used to mask the Landsat 8 mosaic image. The masked agriculture mosaic image is classified into different rice verities using the sample collected in the field . the classification technique used is supervised classification. This map is carried out to analyze the practice of growing rice varieties in the study area.

ISSN No:-2456-2165

C. Data preparation and Statistical Analysis

The data extracted from the Google Earth Engine platform and collected from the farmers were entered into an excel. Interview data was coded and standardized. The field polygons recorded from the hand held GPS were extracted and processed to create a database for the statistical analysis.

The relationships of individual crop management data with each normalized difference vegetative index and yield at different growth stages were initially examined for initial exploration using simple linear regression (SLR) model in the form of y=a+bx. This is only true when ˆy is the predicted y-value given x, a is the y intercept, b and is the slope. The Least Squares Regression Line Predicts y.

Stepwise regression is a method of fitting regression models in which the choice of predictive variables is carried out by an automatic procedure. In each step, a variable is considered for addition to or subtraction from the set of explanatory variables based on some pre-specified criterion. In order to develop the model for this research, stepwise backward regression was used. This method select only those parameters which are strongly and significantly defined the yield variability. Repeated 'trial and error' attempts were performed to identify relationship between parameters and to explore unexpected sign of coefficients. Selected samples were randomly selected from the yield data and reserved for model testing. Multiple regression was then applied to the remaining set of data.Here, the paddy data set consists of a total of 40 records. Out of which 35 records are considered as training set and remaining 5 records for validation.

III. RESULTS AND DISCUSSION

A. Land Use Land Cover and Agriculture Area Map

Class	Percentage %	Area (hectare)
Agriculture	43.96	1261.44
Residential	23.52	67.50
Dense forest	2.90	8.30
Thin Forest	29.62	85.00

Table 1: Land Use Land Cover

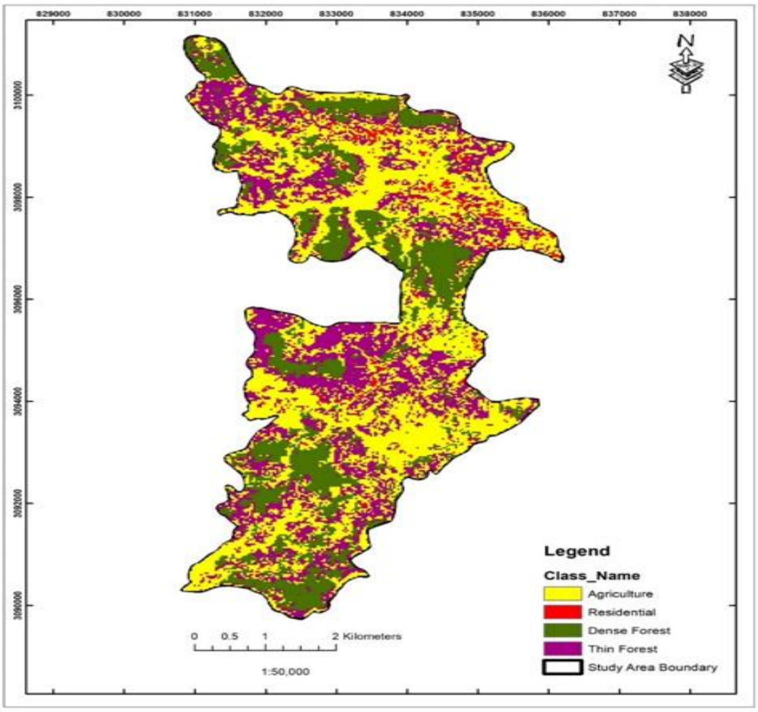

Fig. 1: LULC map of study area

Most of the land of the study are is used for agriculture. The agriculture area dominates the majority part of the study area. It accounts for an area of 43.96% of the total land, and it is followed by thin forest with a covered area of 29.62% of total area .Similarly, the residential and dense forest land comparatively have a lower percentage of coverage 23.52% and 2.90%respectively.

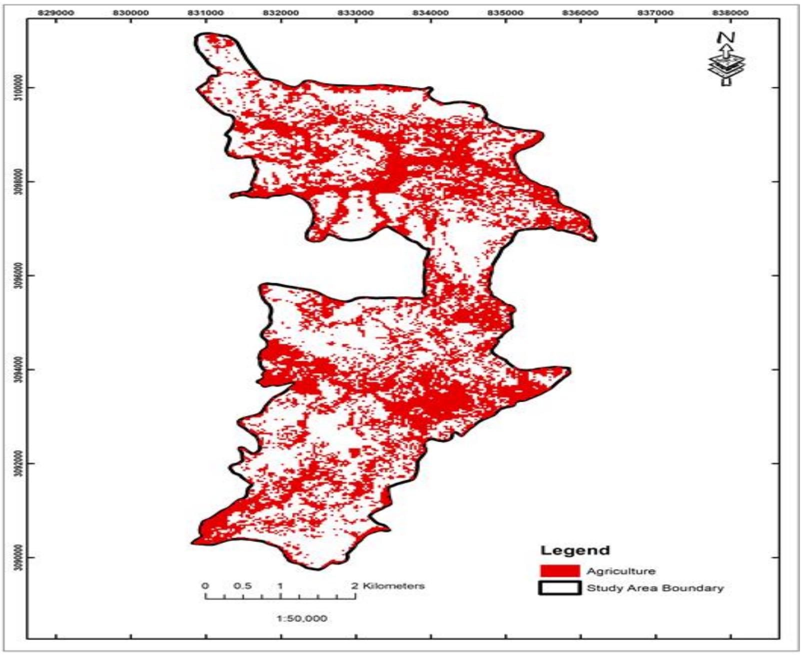

Fig. 2: Agriculture Area

B. NDVI and field level yield data

To define and relate the yield data with the indices, the Normalized Difference Vegetation Indices has been calculated at regular different dates of image of sentinel 2 form July to December 2021. NDVI helps to determine the density of green on the patch of land. It used near infrared bands and the red bands for the calculation.The time series NDVI has been extracted and correlated with the respected actual field level yield. The NDVI types has been classified into maximum, median and minimum NDVI for each specific plot. The maximum ndvi refers to the maximum value of NDVI of pixel inside the plot . Similarly the median and min NDVI referes to median & min value in the pixel inside the plot. The extracted NDVI was then correlated to the actual yield data. It is noted that maximum, median and minimum NDVI were significantly correlated to yield with correlation coefficients of r = 0.514, r = 0.79and r = 0.77 for maximum, median and minimum NDVIrey

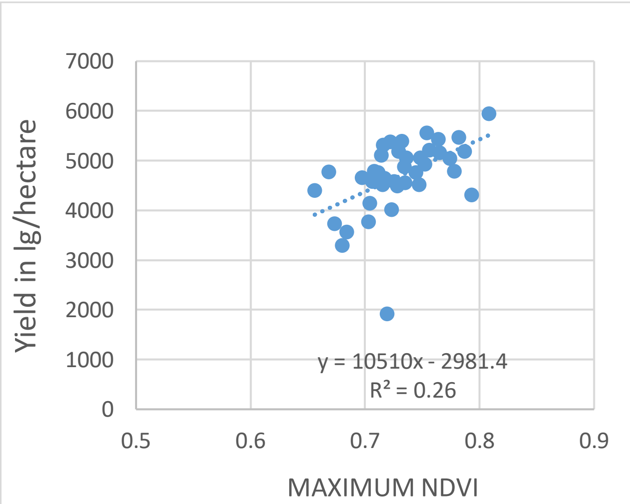

Fig. 3: Maximum NDVI

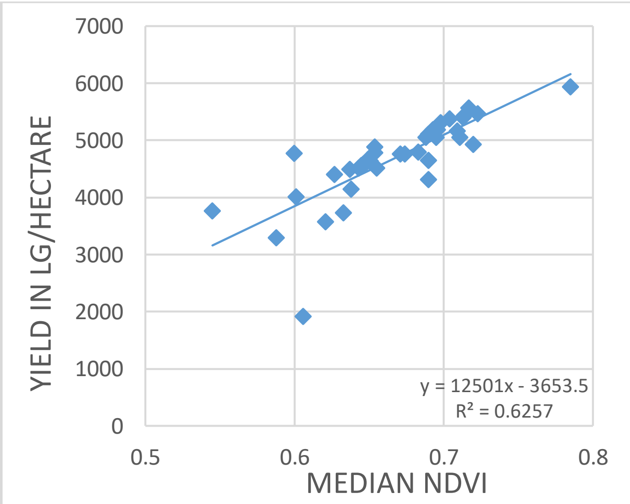

Fig. 4: Median NDVI

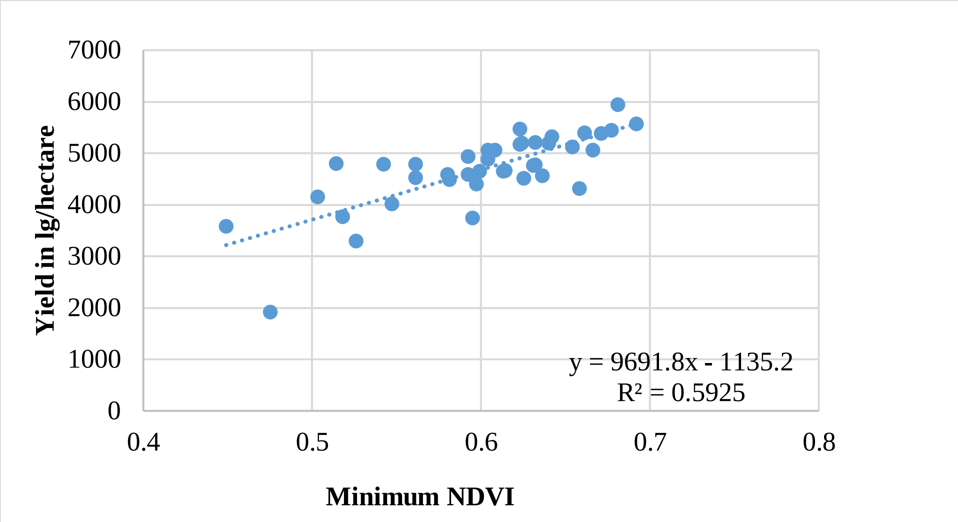

Fig. 5: Chart showing the relationship between different NDVI and Yield in kg/ha

C. *NDVI and Growth phases of paddy*

Altogether, the paddy growth phase is divided into vegetative, reproductive, and ripening phases.The vegetative phase is the longest phase in the paddy's life cycle, which is about 55-85 days.The reproductive phase is about 30 days and the ripening phase is of 15-40 days respectively.

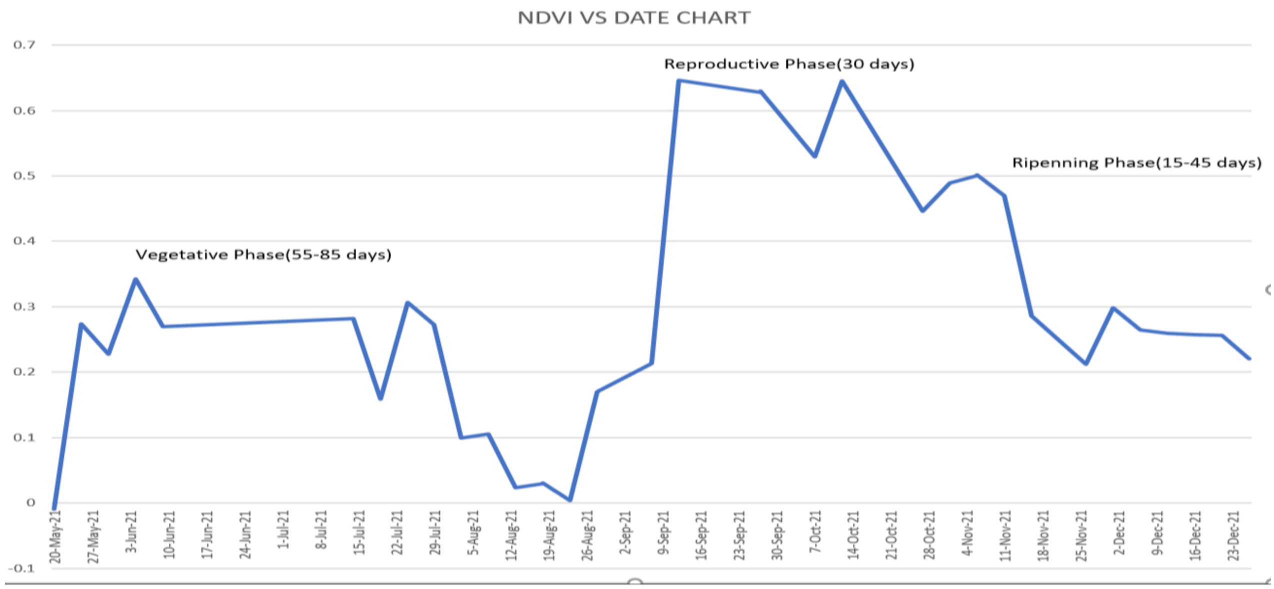

Fig. 6: Figure shows the different stages of paddy

D. *Effects of land management factors on Yield and NDVI*

a) Varieties of Paddy

Various varieties of paddy is grown in different parts of country. Nowadays, the new types of hybrid varieties and high productive paddy seeds from the authorized centre is also practices.But in the study area of this research, the following different varieties sample were found.

Varieties	count	Yield	Min NDVI	Median	Diff NDVI
Sabitri	21	4599.12	0.47	0.588	0.118
Makwanpure	6	4692.56	0.547	0.633	0.086
Ramdhan	13	4843.77	0.459	0.545	0.086

Table 2: Table showing Yield in kg/ha and different NDVI

From the table it is clearly seen that Sabitri variety is widely grown type of paddy in the study area of this research with an average yield of 4599.12 kg/hectare. The average yield of 3.5 to 4 ton/hectare is considered as good yield in Nepal. Although the NDVI value of Makwanpure variety is high but the yield is nearly equal to that of Sabitri with an average yield of 4692.56 kg/hectare. Ramdhan is second popular variety practiced in seratar, Bandipur with highest average yield with an average of 4843.77 kg/hectare. The NDVI of the Ramdhan is comparatively low comparing with other two variety.

Crop type	Percentage	Area(hectare)
Sabitri	62.45	787.86
Ramdhan	35.14	443.34
Makwanpure	2.39	30.24

Table 3: Table showing varieties of rice

According to the above table, we can see that Sabitri dominates the overall category of the study area. It accounts for 62.45% of the total area mostly used type of rice grown with an area of 787.86 hectares. Sabitri rice is followed by the Ramdhan variety which covered the area of 443.34 hectares which is 35.145% of the study area and another type known as Makwanpure covered an area of 30.24 hectares which is just 2.39% of total study area. This result concludes that Sabitri rice is most popular among the farmers and largely cultivated in seratar, Bandipur which is followed by Makwanpure and Ramdhan.

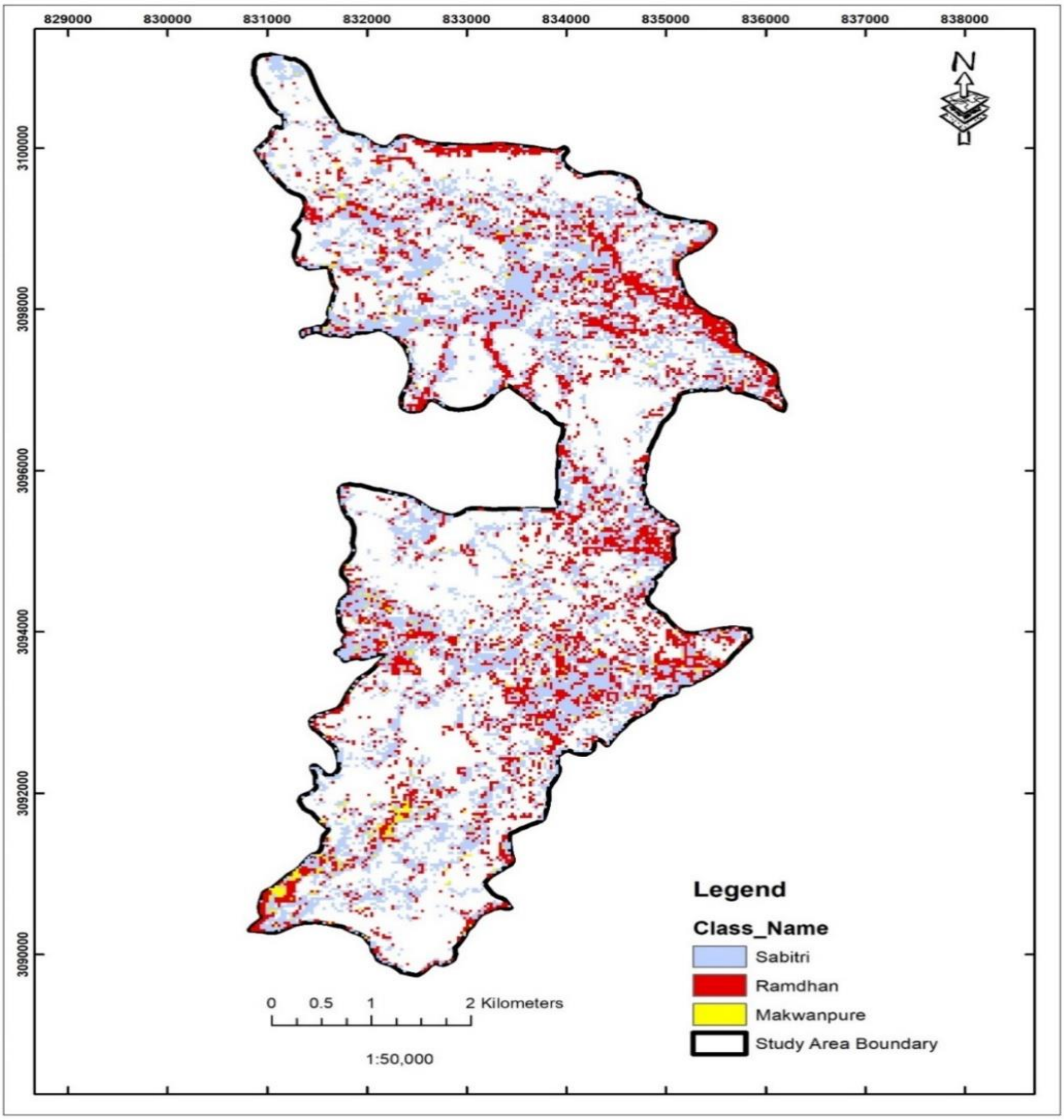

Fig. 7: Rice Distribution Map

ISSN No:-2456-2165

b) Damages due to Pest and Diseases

In the crop production, pest and diseases are very harmful which directly affects production if we did not the solution in time. So, the crop should frequently monitored. There are various types of pest main pest that are found in Nepal during paddy growth.

Some of them are gall Mudge (local name:dhungrekira),leaf folder(local name: pat berne),brown plant hopper(local name: khairafadkekira),stem borer(local name:gawarokira), etc. And these types of pest can be easily controlled using on consultation with experts.

Damages	count	Yield kg/hectare	Min NDVI	Median NDVI
High Damage	7	3981.33	0.563	0.549
Medium Damage	8	4543.83	0.561	0.631
No Damage	15	4901.37	0.642	0.674

Table 4: Table Showing the effects of damage in yield and NDVI

c) Date of Transplantation

There is a practice practiced by the Nepalese farmers, where paddy rice is mainly transplanted when the seedlings are about 30 days old, sowing was mostly done in Transplanting dates ranging from as early as 20th jestha to 15th Asad in 2021, with most farmers transplanting in the 80 to 110th day of the year as shown in the figure.Transplantation time in each field is dependent on various factors, i.e availability of labor, water level, etc. Most of the transplantation in our country was done manually by hand. And in the regression test, we did not find any significant effect on yield because of various transplantation dates.

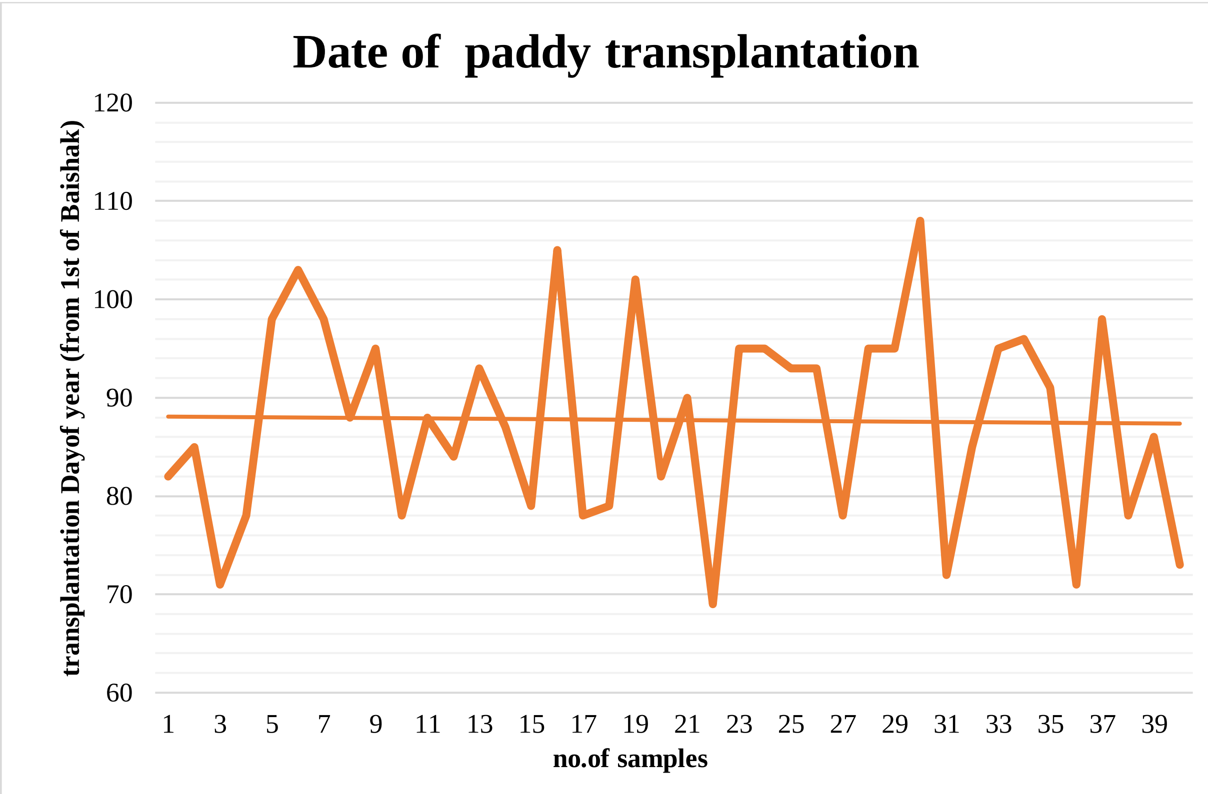

Fig. 8: Graph showing the date of paddy transplantation

d) Date of Harvesting

Days of the year from 1st of baishak	Count	Average Yield	Average NDVI
192-200	6	4185.85	0.62
202-210	13	4864.26	0.67
211-217	12	4580.10	0.67
218-234	9	4922.31	0.62

Table 5: Table showing the date of harvesting

E. Model Development

In the scenario of rice crop yield estimation, considering remote sensing data may logically provide a better understanding and knowledge of seasonal crop growth in order to produce more accurate yield predictions.Multiple regression analysis is applied to know if there is a statistically significant relationship between sets of variables and to find trends in those sets of variables. This relationship here is such explained that crop yield is dependent upon at least two crop parameters or more which were given here as mean NDVI.

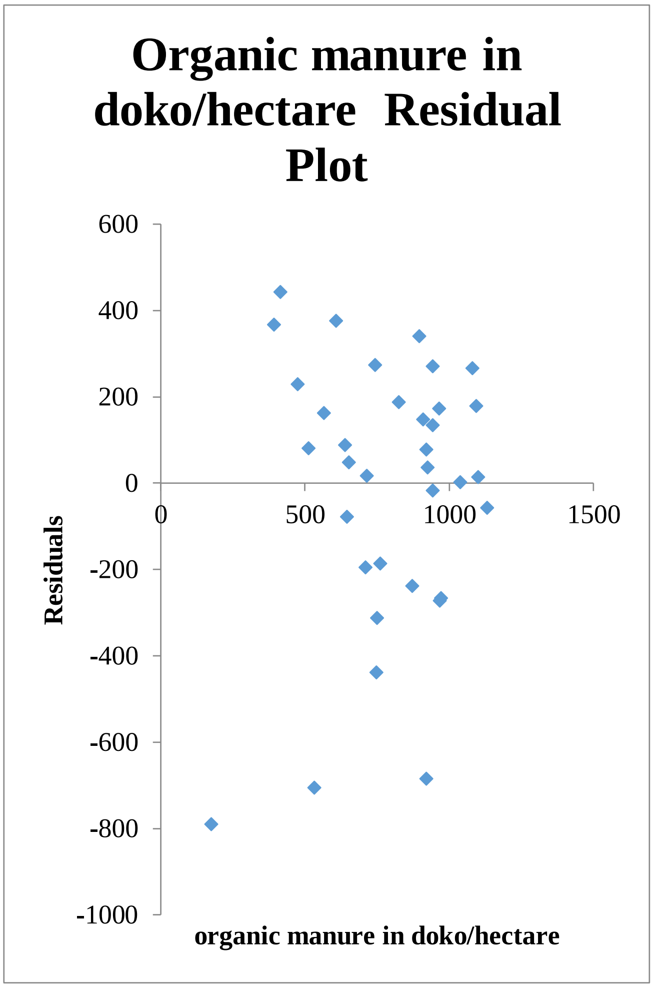

Fig. 9: organic manure in doko/hectare

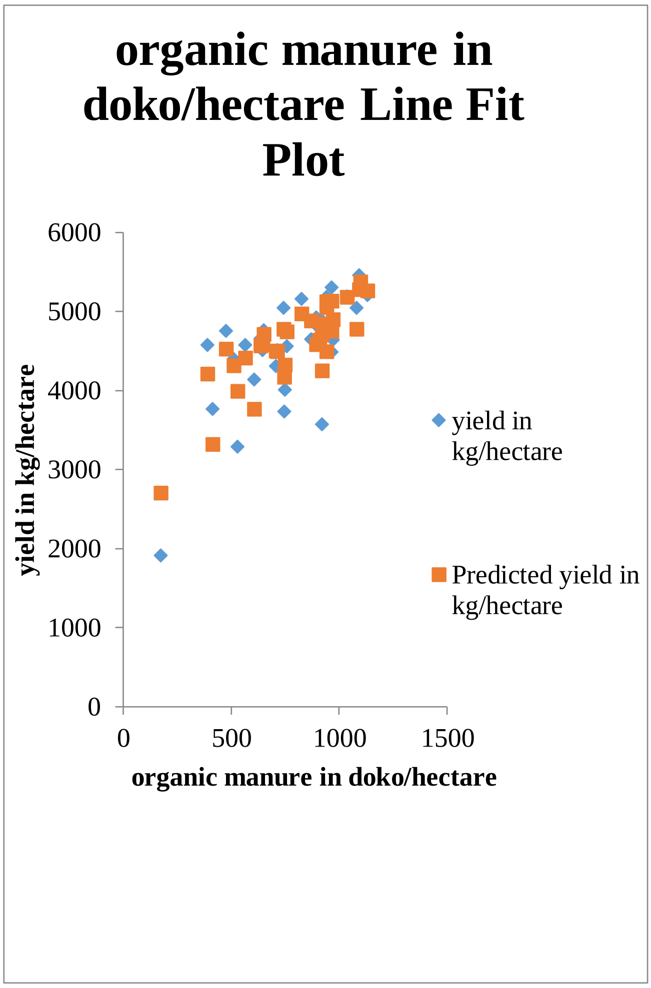

Fig. 10: Organic manure in doko/hectare Line Fit Plot

ISSN No:-2456-2165

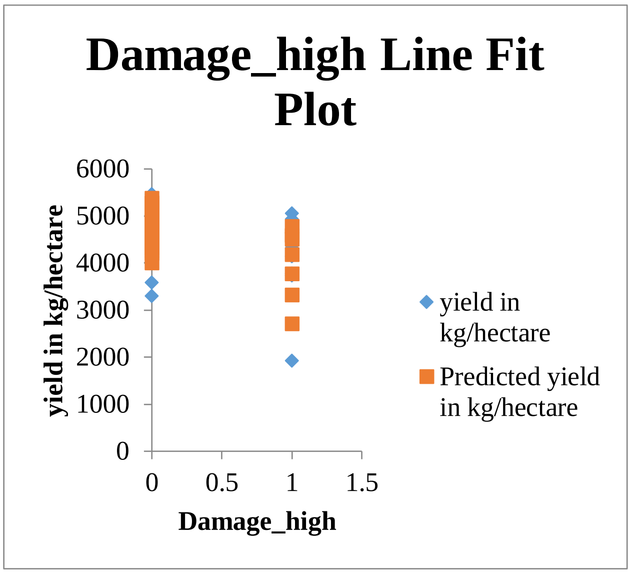

Fig. 11: Damage_high Line Fit Plot

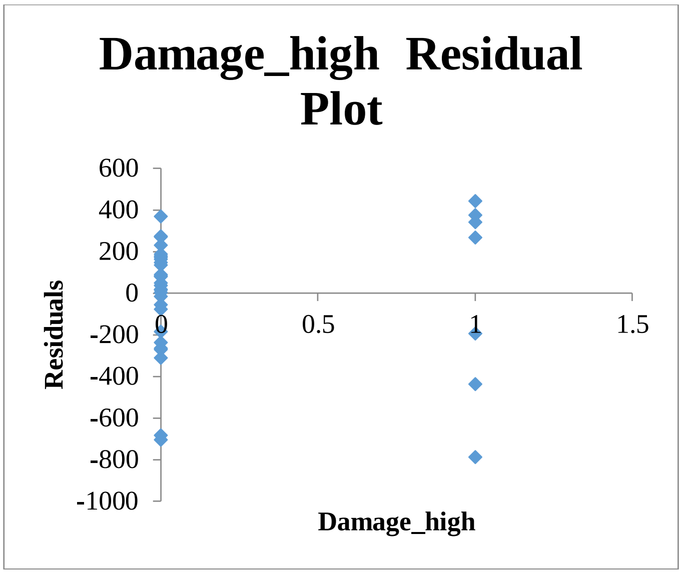

Fig. 12: Damage_high Residual Plot

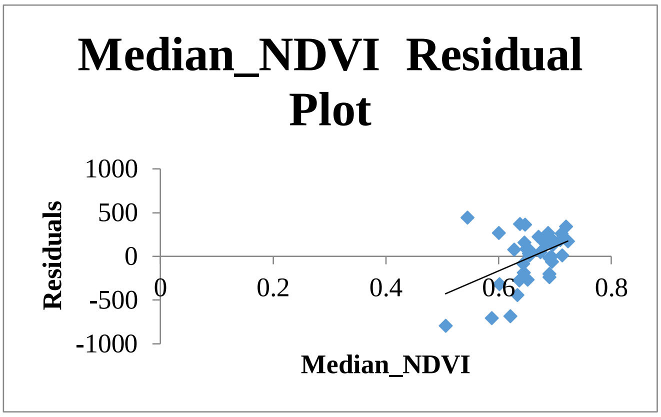

Fig. 13: Median_NDVI Residual Plot

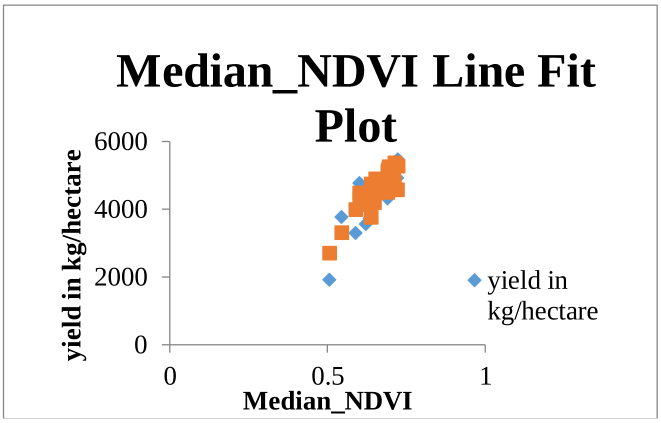

Fig. 14: Median_NDVI Line Fit Plot

All the significant parameters were entered into the stepwise multiple regression to select the best subset that shows the field level NDVI variability. Through repeated trial and error, 2 variables were selected as predictors explaining 96.30% (R^2adj) of yield variability. The main purpose of this was to identify which parameters can explain the yield variability to avoid autocorrelation when building the final model. The regression equation is:

Yield= 1.004 X (OM)-395.979 X (Damage_high) + 3147.7 X (Median_NDVI) + 3073.46 X (Min_ NDVI) ……………………… (2)

Where,

Yield= predicted yield in kg/hectare
OM= Organic Manaure in doko/hectare
Damage_high = Damage caused by the various factors in high level ,Value is 1 if field have high damage else value is zero.
Median-NDVI = Aggregated medium NDVI value
Min_ NDVI= Aggregated minimum NDVI value

This model suggest that there is a negative effect on yield due to high damage on paddy due to the diseases. And yield of paddy will be harvested only when farmer of the study area used the organic manure high in comparison with other fertilizers.

Predictor	Coefficients	Standard Error	t -Stat	P-value
Intercept	0			
Organic manure in doko/hectare	1.004	0.3022	3.324	0.0023
Damage high	-395.979	143.08	-2.7674	0.0094
Median NDVI	3147.696	1484.9	2.119	0.042
Min NDVI	3073.463	1499.9	2.049	0.049

Table 6: Model Development Parameters

ISSN No:-2456-2165

a) Model Validation and Gap Assessment

In this research, we are able to develop the yield prediction model, where out of 40 sample data 35 sample data are used for model development and the remaining 5 samples are left for validation and testing purposes for the model.

S/N	Estimated (yield in kg/hectare)	Actual(yield in kg/hectare)	Difference(yield in kg/hectare)	Difference %
1	5250.893512	5058.794356	192.0991559	3.80%
2	5382.318112	5382.226469	0.091643287	0.00%
3	5567.306764	5438.348453	128.9583113	2.37%
4	6050.06651	5944.387389	105.6791212	1.78%
5	5848.846683	5567.426568	281.4201151	5.05%
		Average	141.6496693	2.60%

Table 7: Table showing Predicted yield and Actual Yield

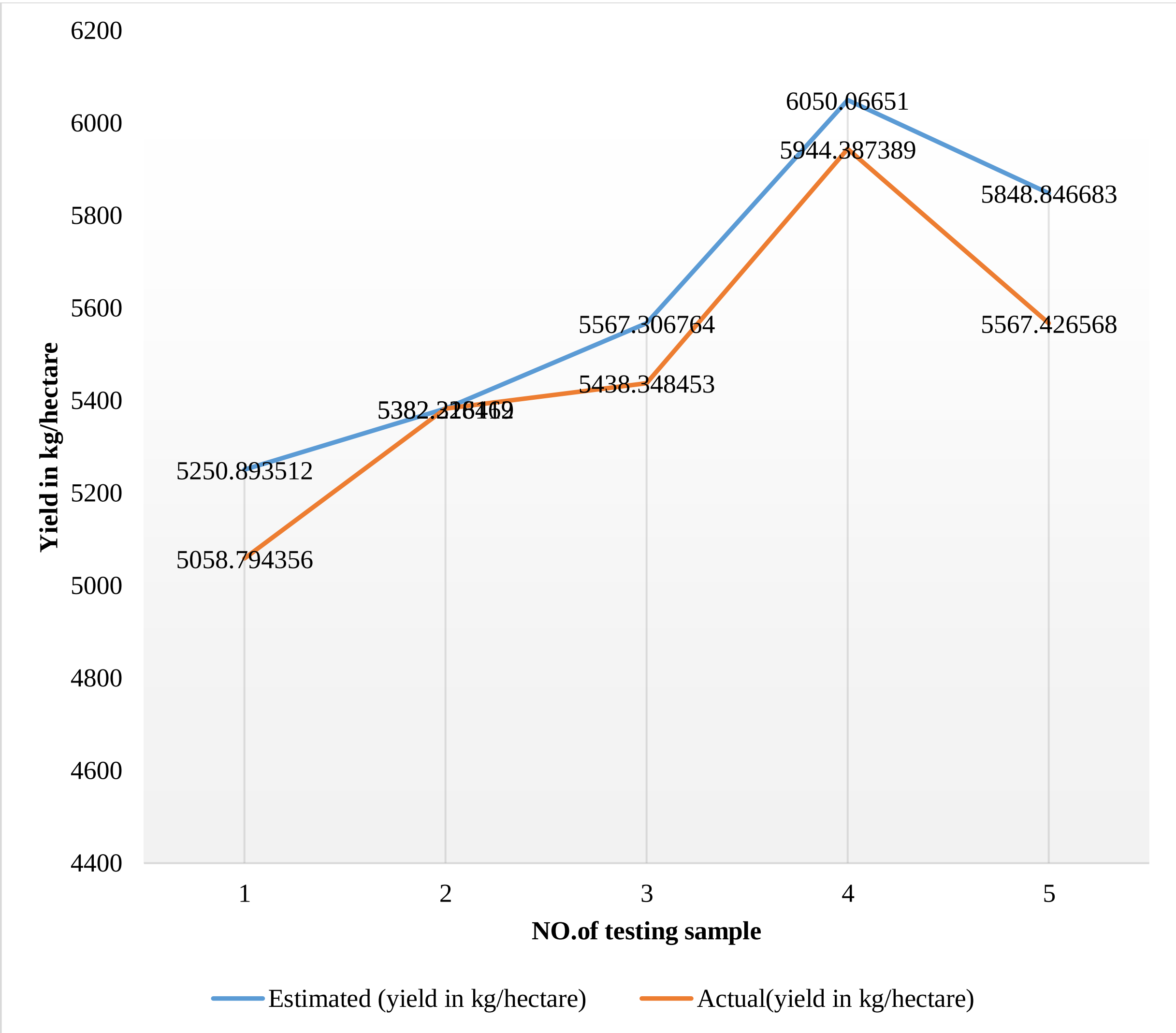

Fig. 15: Graph showing the Actual Yield and Predicted Yield

b) Rice Prediction Map of Study Area

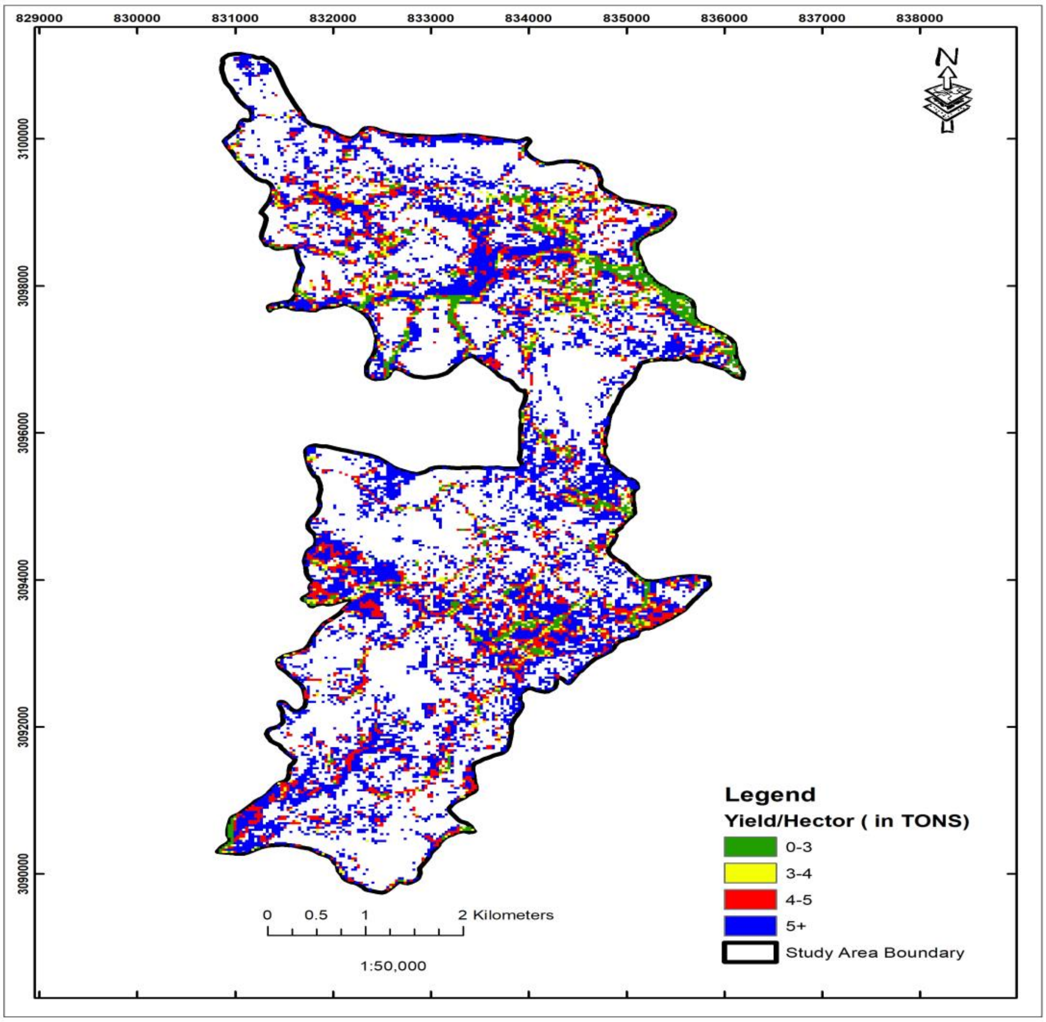

Fig. 16: Rice Prediction map of study area

This rice prediction map which shows that there will be the production of more than 5tons/hectare in majority area where as less than 3tons/hectare in some parts so we address or managed the land management factors in the study then we can achieve the maximum yield between 5-6tons/hectare

IV. DISCUSSION

This study illustrates the model that field level NDVI relationships different land management practices to predict the yield in the field.The NDVI in the early stage of transplantaion is near to zero and it gradually increase with increases.The possible reason are the noise from standing water in the rice fields and may be the cloud cover in the early stage of plantation.

All in all, the newly defined network uniformly shows the best performances over all the tested scenarios which indicate its higher suitability for estimating rice crop yields in study area from S2 data.

Yield forecasting model is an important tool for food security planning, and, nutrition as it allows decision makers, policymakers, higher level management and, stakeholders to make decisions related to pricing, export/import, distribution etc., which helps steer clear of the situation of food insecurity. Reliable and timely crop yield forecasting becomes all the more important in the scenario of climate change since climate change inflames the unpredictability of crop yields for a given season.

As a consequence of these reasons, concerns about the adverse effects of climate change on crop yield and agricultural production have been accounted for, and farmers need to adjust to these challengesof climate change and various factors that affect the yield and adopt new trends and technology for better yield.

In Nepal, crop yield estimation practices a more traditional approach combining sample crop cutting data with field surveys, and field verification reports from the Agriculture Knowledge Centre (AKC). Upon on the information gathered through these methods, and additional field verifications and consultations, the Ministry of Agricultural Development (MoAD), World Food Programme (WFP), and the Food and Agriculture Organization of the United Nations (FAO) release crop situation updates twice a year, after the summer and winter harvests.

V. CONCLUSION

From here we can conclude that Open source earth observation data and cloud-based computing resources provide a remarkable advancement in the land-use land-cover mapping and in crop monitoring. By using the time series NDVI values chart, the different growing stages of the paddy have been monitored and analyzed with different land management factors. From field level NDVI, the health status of the crops can be assessed. But the other indicators like ratio-based vegetation index, infrared percentage vegetation index, perpendicular vegetation index etc. should also be used for better monitoring which were not discussed in this study. It is found that the soil type, variety, land type, damaged caused by pest and diseases, amount of seed sowing, date of transplantation, date of harvesting, can explain the NDVI variability.

Through the combination of remotely sensed data, land management factors and management aspects, an assessment of field-level yield prediction has been put into effect. It is clearly shown that a combination of NDVI, and land management parameters can improve field-level yield prediction rather than the use of NDVI alone. It is found that the NDVI, land damage caused by pest and diseases, and fertilizers can explain yield variability. This study established that there is a significant positive relationship between Normalised Difference Vegetation Index and field level yield(r=0.79).

This study has develop a yield prediction model for field level. The data used in this study was provided by farmers of the study area through interviews. So that there might be a lot of irregularity in the reported data that has to be verified before final conclusions are drained on its accuracy and applicability. There is need for further researches to find more land management factors that can explain yield variability at field level and improve the model. So using only the optical remote sensing cannot provide the best explanation in monitoring the crops. So there is a need of integration between optical and microwave remote sensing for the advance level of crop monitoring. There is need for further researches to find more land management factors that can explain yield variability at field level and improve the model. So using only the optical remote sensing cannot provide the best explanation in monitoring the crops. That's why there is a need of integration between optical, microwave remote sensing and advanced machine learning algorithms like CNN(convolution Neural Network) for the advance level of crop monitoring.

REFERENCES

[1.] Abdullah, N. S. (2014). SUITABILITY MODEL BASED ON GIS AND MCDA FOR SPATIAL DISTRIBUTION OF SETTLEMENTS IN DIFFERENT GEOGRAPHIC ENVIRONMENTS. *European Scientific Journal*, 236-249.

[2.] ADHIKARI, S. (2019). CROP INFORMATION SYSTEM: A CASE STUDY OF PADDY MONITORING IN BHARATPUR 13,NEPAL.

[3.] Ahlrichs. J. S. & Bauer, M. (1983). Relation of agronomic and multispectral reflectance characteristics ofspring wheat canopies.

[4.] Atzberger, C. (2013). Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information . *Remote Sensing vol.3*.

[5.] Biswas, R., Bhattacharyya, B., & Banerjee, S. (2017). Predicting Rice Yield from Weather Variable through Detrended Production Index. *Res. India 462–466.*, 462–466.

[6.] Degerliyurt, M. (2014). Settlement suitability analysis of local ground characteristics in Iskenderun. *Procedia-Social and Behavioral Sciences 120 (2014)*, 637-644.

[7.] Ediz UNALIDa, H. Y. (2020). Yield Estimation of Winter Wheat in Pre-harvest Season by Satellite Imagery Based Regression Models. *Turkish Journal of Agricultural Engineering Research* , 390-403.

[8.] Fentanesh Haile Buruso for hippopotamus (H. amphibious) using GIS and remote sensing in Lake Tana and its environs, E. (2017). Habitat suitability analysis. *Environmental Systems Research*.

[9.] Geli Zhang, X. X. (2015). Mapping Paddy Rice Planting Areas through Time Series Analysis of MODIS Land Surface Temperature and Vegetation Index Data. *ISPRS J. Photogrammetry, Remote Sensing*.

[10.] Gómez, D., Salvador, P., Sanz, J., & Casanova, J. (2019). Potato yield prediction using machine learning techniques and sentinel 2 data. *Remote Sensing*.

[11.] Gu, Y., Wylie, B., Howard, D., Phuyal, K., & Ji, L. (2013). NDVI saturation adjustment: A new approach forimproving cropland performance estimates in the Greater Platte River Basin, USA. *Ecol. Indic.*., 1-6.

[12.] Kayad, A., Sozzi, M., Gatto, S., Marinello, F., & Pirotti, F. (2019). Monitoring Within-Field Variability of Corn Yield using Sentinel-2 and Machine Learning Techniques. *Remote Sensing*.

[13.] Krishna Devkota, S. Y. (2014). *Landslides in Mt. Umyeon Susceptibility Mapping using GIS and Mitigation Measures.* Hydrotech Research Institute, National Taiwan University, Seoul.

[14.] Kumar, R. (2019). The price of rice.

[15.] Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEE geoscience Remote Sensing*.

[16.] Nepal's rice economy. (2019). *Nepali Times*.

[17.] Nuarsa, I. N. (2011). Spectral Characteristics and Mapping of Rice Plants UsingMulti-Temporal Landsat Dat.

ISSN No:-2456-2165

[18.] Nuarsa, I., Nishio, F., & Hongo, C. (2014). Rice yield estimation using Landsat ETM+ data and field observation. *Agriculture Science*.

[19.] Onojeghuo, A., Blackburn, G., Wang, Q., Atkinson, P., Kindred, D., & Miao, Y. (2018). *Mapping paddy rice fields by applying machine learning algorithms to multi-temporal Sentinel-1A and Landsat data*. Int. J. Remote Sens.

[20.] Qamer, F., Shah, S., Murthy, M., Baidar, T., Dhonju, K., & Hari, B. (2014). Operationalizing crop monitoring system for informed decision making related to food security in Nepal. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*

[21.] Qiangzi Li, H. Z. (2014). County-level rice area estimation in southern China. *Chinese Academy of Sciences*.

[22.] Raheem, N. Q. (2016). Determining the suitability trends for settlement based on multi criteria in krikuk, Iraq settlement based on multi-criteria in. *Open Geospatial Data, Software and Standards*.

[23.] Reynolds, C. Y. (2000). Estimating cropyields and production by integrating the FAo crop specific water balance mode lwith real-time satellite data andground based ancilliary data.

[24.] Ruben Fernandez-Beltran, T. B. (2021). Rice-Yield Prediction with Multi-Temporal Sentinel-2 Data and 3D CNN: A Case Study in Nepal. *Remote Sensing*.

[25.] Shao, Y., Fan, X., Liu, H., Xiao, J., Ross, S., Brisco, B., . . . Staples, G. (2010). Rice monitoring and production estimation using multi-temporal RADARSAT. *Remote Sens Environ*, 310-325.

[26.] Shrisath, P. (2016). Real-Time Crop Yield Monitoring in Nepal for Food Security Planning and Climatic Risk Management. *CGIAR Research*.

[27.] Sofa, B. L. (1-58). Vegetation analysis using remote sensing. ,2010.

[28.] South, C. (2016). *Land Suitability Analysis Report*. Greenwood Village: digstudio.

[29.] umma, M., Nelson, A., Thenkabail, P., & Singh, A. (2011). Mapping rice areas of south Asia using MODISmultitemprol data. *J Appl Remote Sens.*., 535-547.

[30.] van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.*

[31.] Wang, L., Zhang, G., Wang, Z., Liu, J., Shang, J., & Liang, L. (2019). Bibliometric Analysis of Remote Sensing Research Trend in Crop growth Monitoring: A Case Study in China. *Remote Sens*.

[32.] You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep gaussian process for crop yield prediction based on remote sensing data. *In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA*, 4559–4565.

[33.] Zhao, G. M. (2012). A Preliminary Precision rice Management System for Increasing Both Grain Yield and Nitrogen Use Efficiency. *Field Crops Res*.

[34.] Zhao, Y., Potgieter, A., Zhang, M., Wu, B., & Hammer, G. (2020). L. Predicting wheat yield at the field scale by combining high-resolution Sentinel-2 satellite imagery and crop modeling. *Remote Sensing*.

APPENDIX 1: QUESTIONNAIRE FOR FIELD DATA COLLECTION

1. What is the name of farmer:

2. General information:

- Field size according to farmer:
- From farmer parcels:
- From using Gramin hand held GPS:

3. How many types and variety of crops did you grow last year?

4. When did you start you land preparation? Number of ploughing or puddlings?

5. When did you sowing the seed? What is the method applied? How much seed in kg? Variety? Plant quality? (Good/average/poor)? Age of transplanting? Where did you get

seed? Water availability?

6. Fertilizers applied? When? How what type of method? Types of fertilizers? Quantity? Water level (before and after)?

7. When was the weeding done? How?

8. Pesticide Application:

- When?
- Why?
- Severity of damage
- Name of pest and diseases
- Control method, how?

9. about Yield

- When was paddy harvested?
- How much did they harvested?
- How much did they expect?
- Why there is difference?
- How much was the reduction?
- What may be the cause if there is any reduction?

10. What are the plans to increase yields? What type of support do you expect? Would you like to change crop? If yes then why? To which crop or variety?

ISSN No:-2456-2165

APPENDIX-2 Code book

Data	Parameter code	Parameter name	Parameter unit	Parameter value	Parameter value name	Remarks
Sample no	Sample_no	Sample no				
Sample id	Gpssam _id	Sample id				
Sample size	Area_ha	Area in hectare				Area from the polygon from processing converted to hectare
Village Name	V_name	Name of village	hectare			
variety	var	Sabitri				
		Makwanpure				
		Ramdhan				
Soil	st	Soil type		Type 1	kalo	
				Type 2	Balaute	
				Type 3	Rato	
Ploughing	P1	Day of 1^{st} plowed				
	Pm1	1^{st} plowed method		T	tractor	
				M	manually	
	P2	Day of 2^{nd} plowed				
	Pm2	2^{nd} plowed method		T	Tractor	
				m	manually	
	dos	Sowing date	Days			1^{st} of baishak
	asg	Amount of seed swoing	Kg/ha			

ISSN No:-2456-2165

Data	Parameter code	Parameter name	Parameter unit	Parameter value	Parameter value name	Remarks
Plantation	td	Transplantation date	Days			Seed grown is collected as total kg of seed grown and converted to kg/ha
Basal Dressing	OM	Organic manure	Doko/hectare			
	ZBD	Zinc basal dressing	Kg/hectare			
	POT	potash	Kg/hectare			
	DAP	Diammonium phosphorous	Kg/hectare			
	URA	Urea	Kg/hectare			
Top Dressing						
Irrigation	nd	Normal water depth	millimeter	mm		Estimated average depth of water in the field reported in inches and converted to mm
	wa	Water availability		r	Regular supply	
				i	Irregular water supply	
				s	Shortage of water	
wedding	Dow1	Days of weeding	days			
	Mow1	Methods of 1st wedding		hw	Hand weeding	
				ch	Chemical used	
	Dow2	Days of 2nd wedding	days			
	Mow2	Methods of 2nd wedding		hw	Hand weding	
				ch	Chemical used	
				lf	Leaf folder	

ISSN No:-2456-2165

Data	Parameter code	Parameter name	Parameter unit	Parameter value	Parameter value name	Remarks
Pest Management	top	Type of pest		rt	rats	
				np	No pets	
	cmth	Control Methods		ch	chemical	
				nch	No chemical	
	capp	Chemical application, day of the year	days			
	mpes_a p	Method of pesticide application		sp	spryaing	
				br	Broadcasting	
	dmg_in s	Damage due to insect		Nd	No damage	
				md	Minor damage	
Diseases control	tod	Types of diseases		Kp	Kalopoko	
				rt	rate	
				Dh	dadhuwa	
	tod_cm	Control methods		Ch	Chemical use	
				nch	No use of chemical	
	Dmg_dis	Damage due to diseases		nd	No damage	
				md	Minor damage	
				hd	High damage	
Harvesting	Doh	Date of harvesting	days			Days from baishak 1,2078
	hp	Harvested paddy	Kg/hectare			values reported in muri/area and converted to kg/ha
NDVI	NDVI	Normalized Difference Vegetation Index		maximum	Aggregated by maximum	Calculate from sentinel 2 image
				median	By median	
				minimum	By minimum	

APPENDIX-3

Data used For the Study

Id	Sample no	Area in hectare	Village name	Var	st	lt	P1	Pm1
9	1	0.088	Sera	Ramdhan	T1		59	T
3	2	0.307	Sera	Sabitri	T1		43	T
37	3	0.177	Sera	Sabitri	T2		46	T
41	4	0.072	Sera	Ramdhan	T2		29	T
8	5	0.127	Sera	sabitri	T2		67	T
7	6	0.137	Nahala	sabitri	T2		67	T
17	7	0.287	Nahala	sabitri	T3		51	T
2	8	0.155	Nahala	makwanpure	T1		46	T
14	9	0.282	Nahala	sabitri	T1		46	T
34	10	0.098	Nahala	ramdhan	T2		53	T
30	11	0.217	Nahala	ramdhan	T1		41	T
19	12	0.108		sabitri	T1		51	T
31	13	0.336		sabitri	T2		57	T
1	14	0.470		sabitri	T2		36	T
20	15	0.194		Makwanpure	T1		46	T
6	16	0.104		Ramdhan	T1		57	T
32	17	0.297		sabitri	T1		51	T
29	18	0.159		makwanoure	T2		42	T
16	19	0.463		sabitri	T1		59	T
10	21	0.033		Ramdhan	T3		46	T
18	22	0.201		sabitri	T3		36	T
36	23	0.145		ramdhan	T3		28	T
13	24	0.043		Makwanpure	T3		28	T
15	25	0.315		ramdhan	T2		46	T
27	26	0.182		sabitri	T2		43	T
35	27	0.318		sabitri	T2		44	T
39	28	0.303		sabitri	T2		38	T
25	29	0.370		sabitri	T2		64	T
23	30	0.306		Ramdhan	T2		56	T
5	31	0.446		sabitri	T1		59	T
21	32	0.172		Makwanpure	T1		25	T
10	33	0.138		Ramdhan	T1		20	T
24	34	0.084		sabitri	T2		18	T
4	35	0.155		ramdhan	T3		33	T
26	36	0.157		sabitri	T3		43	T

ISSN No:-2456-2165

33	37	0.362		sabitri	T1		41	T
22	38	0.205		sabitri	T1		31	T
40	39	0.336		ramdhan	T2		63	T
28	40	0.145		ramdhan	T2		31	T

Sample.no	P2	Pm2	dos	asg	td	om	DAP	URA
1	77	T	65	45.30	95	1133	56.63	56.63
2	69	T	58	35.88	88	750	65.24	65.24
3	67	T	57	67.91	87	566	56.59	56.59
4	82	T	63	55.20	98	1104	55.20	55.20
5	82	T	62	35.40	95	944	62.93	62.93
6	82	T	56	36.45	95	1093	58.32	58.32
7	74	T	50	69.77	85	174	52.32	52.32
8	76	T	63	64.52	96	710	64.52	64.52
9	70	T	58	28.36	95	922	53.18	53.18
10	88	T	71	51.21	105	512	51.21	51.21
11	77	T	50	46.06	82	921	46.06	46.06
12	77	T	62	46.21	98	416	55.45	55.45
13	67	T	53	44.60	82	743	59.46	59.46
14	82	T	69	53.15	103	531	44.64	44.64
15	82	T	63	51.64	95	826	51.64	51.64
16	72	T	52	48.30	85	966	57.95	57.95
17	70	T	76	60.60	102	926	67.33	67.33
18	75	T	61	44.01	98	943	62.87	62.87
19	67	T	57	25.91	90	972	64.78	64.78
21	87	T	75	60.83	108	760	60.83	60.83
22	51	T	45	49.82	79	747	49.82	49.82
23	64	T	51	34.54	84	967	55.26	55.26
24	56	T	46	46.17	72	1039	46.17	46.17
25	65	T	59	47.66	95	651	63.55	63.55
26	70	T	49	44.05	91	1101	55.06	55.06
27	71	T	50	44.04	93	944	47.19	47.19
28	65	T	44	49.54	86	1486	49.54	49.54
29	71	T	46	32.42	79	1081	67.55	67.55
30	73	T	51	49.07	88	393	32.71	32.71
31	77	T	59	49.30	93	896	44.82	44.82
32	59	T	43	58.13	78	872	52.32	52.32

33	61	T	35	57.81	71	1012	50.59	50.59
34	56	T	41	59.47	78	714	47.57	47.57
35	61	T	56	64.50	93	645	64.50	64.50
36	62	T	51	63.83	69	638	63.83	63.83
37	68	T	46	41.44	71	608	55.25	55.25
38	59	T	30	58.60	73	1465	97.67	97.67
39	77	T	43	44.60	78	476	44.60	44.60
40	51	T	36	55.07	78	1239	55.07	55.07

Sample.no	Dow1	Mow1	Dow2	Mow2	tod	cmth	doh
1	98	hw	120	hw	rt	nch	195
2	99	hw	121	hw	rt	nch	196
3	105	hw	120	hw	rt	nch	213
4	105	hw	122	hw		nch	202
5	105	hw	125	hw		nch	203
6	115	hw	145	hw		nch	203
7	117	hw	140	hw	rt	nch	215
8	115	hw	135	hw	dh	nch	215
9	115	hw	145	hw		nch	216
10	117	hw	147	hw		nch	197
11	93	hw	125	hw	dh	nch	199
12	84	hw	115	hw	dh	nch	192
13	90	hw	123	hw		nch	205
14	84	hw	136	hw		nch	207
15	101	hw	130	hw	dh	nch	207
16	105	hw	140	hw		nch	203
17	105	hw	138	hw		nch	203
18	86	hw	130	hw		nch	212
19	120	hw	150	hw		nch	215
21	125	hw	156	hw		nch	222
22	101	hw	140	hw	rt	nch	212
23	106	hw	148	hw		nch	215
24	97	hw	130	hw		nch	202
25	80	hw	125	hw		nch	221
26	102	hw	136	hw		nch	215
27	89	hw	123	hw		nch	227
28	98	hw	130	hw		nch	212
29	98	hw	120	hw		nch	228
30	99	hw	121	hw		nch	202

31	105	hw	120	hw	rt	nch	215
32	105	hw	122	hw	dh	nch	234
33	105	hw	125	hw		nch	225
34	115	hw	145	hw	dh	nch	205
35	117	hw	140	hw		nch	210
36	115	hw	135	hw		nch	221
37	115	hw	145	hw	rt	nch	195
38	117	hw	147	hw		nch	227
39	93	hw	125	hw		nch	206
40	84	hw	115	hw	rt	nch	217

Sample.no	Max_NDVI	Med_NDVI	Min_NDVI
1	0.756	0.695	0.632
2	0.723	0.601	0.547
3	0.725	0.646	0.592
4	0.722	0.704	0.671
5	0.714	0.69	0.654
6	0.782	0.723	0.623
7	0.719	0.606	0.475
8	0.793	0.69	0.658
9	0.778	0.654	0.561
10	0.656	0.627	0.597
11	0.684	0.621	0.449
12	0.703	0.545	0.518
13	0.736	0.688	0.608
14	0.68	0.588	0.526
15	0.765	0.709	0.623
16	0.716	0.698	0.642
17	0.734	0.654	0.604
18	0.668	0.6	0.542
19	0.717	0.651	0.613
21	0.735	0.645	0.636
22	0.673	0.633	0.595
23	0.728	0.637	0.581
24	0.729	0.693	0.64
25	0.744	0.674	0.632
26	0.732	0.713	0.661
27	0.787	0.696	0.624

28	0.808	0.785	0.681
29	0.774	0.711	0.604
30	0.706	0.647	0.58
31	0.752	0.72	0.592
32	0.71	0.69	0.599
33	0.748	0.695	0.666
34	0.715	0.655	0.561
35	0.747	0.643	0.625
36	0.697	0.649	0.614
37	0.704	0.638	0.503
38	0.754	0.717	0.692
39	0.711	0.671	0.631
40	0.764	0.714	0.677

APPENDIX-4

a. Production of paddy

s/n	Production Before Land pooling in Rs	Production after land pooling in Rs	Estimated in Rs
1	38640	31261.77	38895.51
2	43205	39868.34	39869.02
3	40600	40284.05	41239.31
4	84840	44032.49	44815.31
5	57400	41240.2	43324.79

b. Cultivation cost

S/n	Cost Before Land pooling in Rs	Cost after land pooling in Rs
1	38640	31261.77
2	43205	39868.34
3	40600	40284.05
4	84840	44032.49
5	57400	41240.2