International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V69I12P206
Volume 69 Issue 12, 39-46, December, 2021

Surface Area Classification Using Sentinel-1 SAR Backscattering Coefficients

Prachi Kaushik[1], Suraiya Jabin[2]

[1,2,]*Department of Computer Science, Jamia Millia Islamia, New Delhi, India*

[1] prachikaushik.4@gmail.com, [2] sjabin@jmi.ac.in

Abstract *— The Sentinel-1 synthetic aperture radar satellite captures high-resolution images of land and sea. The land cover classification of the earth's surface can be done by analyzing the backscatter values recorded by the sensor. The images are calibrated to convert the digital number of each pixel into backscattering coefficients. For this study, the co-polarized and cross-polarized bands for the land cover classes were collected for the test area of Bhiwani. This enables the mathematical analysis of the bands to study the dependency of the coefficients for different land cover categories. There is a significant difference between the backscatter coefficients for the categories viz. apartments, villages, low-rise buildings, forests, parks, agricultural land, and water. Statistical analysis is done to study the level of correlation between class categories. Apartments and urban areas show a strong positive correlation of 0.992, indicating that apartments are strongly related to urban areas. A weak positive correlation of 0.288 between apartments and villages indicates that apartment-like structures are less likely in rural areas. Further, a few classification models for surface area classification were trained in seven categories, out of which ensemble bagged trees model performed with an accuracy of 95.3 percent on the test data.*

Keywords *— backscatter, SAR, Cubic SVM, ensemble, calibration, MATLAB*

I. INTRODUCTION

The Sentinel-1 mission of the European Space Agency was the first mission with the objective of the sea and land monitoring regardless of the weather conditions. The C band synthetic aperture radar provides a high spatial and temporal resolution. The electromagnetic waves sent by satellite interact with the surface of the earth and is reflected back in the form of scatters, which are recorded by the sensor for SAR image formation. Generally, authors use SAR images [13] for surface area classification. The paper demonstrates the role of backscattering coefficients beta (β_θ), sigma (σ_θ), gamma (γ_θ) for training a more successful classifier towards the classification of earth surface into seven land-cover classes apartments, villages, low-rise buildings, forests, parks, agricultural land, and water.

The beta coefficient is only dependent on the amplitude value of the image and independent on the local incidence angle. The sigma backscatter coefficient is dependent on the local incidence angle, thus minimizing the topographic backscatter effects. This coefficient is useful in the comparison of images from different SAR sensors and different time periods. It also gives a rough estimate of the type of scattering like surface, volume, double-bounce scattering. The gamma backscatter returns the radar reflectivity per unit area perpendicular to the direction of measurement. The Vertical transmit-vertical receive (VV) and vertical transmit-horizontal receive (VH) backscatter intensity identify the different land cover classes on the earth's surface. These coefficients vary based on the surface characteristics like surface roughness, the shape of the surface, height above the ground level and moisture content in the soil, etc. These SAR microwave indices make land cover classes distinguishable from each other, which gives a motivation to design a good classification model for land cover studies on remotely sensed data. The statistics like mean, standard deviation, variance, coefficient of variation, coefficient of correlation is plotted for considered land cover classes to derive a useful discussion from distinguishing the classes. In this study, backscatter coefficient values (18 features as described in subsection) for different surface categories; apartments, villages, low-rise buildings, forests, parks, agricultural land, and water were collected to form a dataset of size 13500 samples. Different supervised classification models like cubic support vector machines (SVM), ensemble bagged trees, and ensemble boosted trees were experimentally trained on labeled training data to make a robust classifier. The classifier learns from the useful microwave indices, spectral, spatial, textural features extracted from the Sentinel-1 SAR image to correctly classify the surface objects in seven categories.

II. LITERATURE SURVEY

There have been studies that have employed contemporary SAR technologies for agricultural monitoring in the past. S. Abdikan et al. [1], in their paper, discussed the relationship between the maize height and sigma backscatter values (Sigma_VH, Sigma_VV). A high coherence was recorded in the early stages of plant growth, and the coherence reduced as the plant growth advanced. The backscattering coefficient discriminated different kinds of crops with 80% of accuracy. M. Vrevgdenhil et al. [2] accessed the potential of SAR microwave indices (VV, VH) backscatter, VH/VV Ratio, cross-ratio to monitor the crop conditions. In-situ data such as Leaf Area Index (LAI), Vegetation Water Content, biomass, and height of the plant was collected for oilseed, corn, winter cereals in

two growing seasons. The research demonstrated that microwave indices are highly sensitive to vegetation dynamics. G. Macelloni et al. [3] investigated the influence of shape and dimensions (geometry) on the sigma backscattering coefficient for C and L band SAR images. The crops having the same biomass may have different sigma backscatter due to different geometry. Narrow-leaf crops show that with increasing biomass, the backscatter decreases, whereas, for broadleaf crops, the backscatter increases with the decrease in the biomass. R. Nasirzadehdizaji et al. [10] experimentally show a high relationship of maize height and SAR backscattering parameters, VV, and VH bands in the early growing stage. Constantino-Recillas et al. [17] showed a high correlation between sentinel-1 backscatter and vegetation parameters to study the temporal variability of cornfields. Modanesi et al. [18] used the backscatter observation operator to model the soil moisture in the fields to quantify water needs. A water cloud model was calibrated with the help of the backscatter values. He et al. [19] used the RADARSAT-2 quad-pol backscattering coefficients to study the backscatter statistical characteristics of rice fields in polarimetric bands. A simple decision tree approach was presented to identify rice paddy fields based on the scatter performance with 88.65% accuracy.

The study of backscattering coefficients from woods has been the focus of several authors. F. T. Ulaby et al. [4] discussed the radar backscattering from the forest canopies based on the Michigan microwave canopy scattering model. Direct backscattering takes place from leaves & branches. Scattering from different regions such as trunk-ground, leaf-ground scattering is known as multiple scattering. C. Dobson [5] examined the dependence of radar backscatter for P, L, and C band SAR data on above-ground biomass for coniferous forests. The experiments show an increase in radar backscatter with increasing biomass until biomass saturation point, which varies with different bands. The C band is less sensitive to the total above-ground biomass. P. Ferrazzoli et al. [6] describe radiative transfer theory & matrix doubling algorithm to compute backscatter coefficients for forests. The HV polarization is highly sensitive to woody volume and influenced by branch dimension & orientation.

Several authors made useful conclusions on the double bounce scattering from the building structures. K. Koppel et al. [7] discuss the dependence of sigma backscatter on physical building parameters like surface material, density, shape, and height. The strongest backscatter is due to the variable height of the buildings. Double bounce scattering from wall-ground interactions is easily detected by Sentinel-1 SAR images. Y. Dong et al. [8] give a detailed understanding of the different types of radar backscattering from urban areas. He talked about single bounce-off rooftops, double bounce-off walls and ground structures, and triple scattering off walls and ground structures. M.B. Charlton et al. [9] have experimented with the lesser-explored dependence of surface roughness with SAR backscatter. The surface roughness is characterized by the standard deviation of the surface heights, correlation length, and auto-correlation function. Delgado Blasco et al. [20] have analyzed several factors like building orientation street dimension, construction type, building height that can influence the double bounce scattering from the urban areas by taking the fully polarimetric ALOS PALSAR data.

The results indicate the relationship between surface roughness and profile length. On the basis of the literature survey, it can be concluded that backscatter values have a dependency on parameters like height or surface roughness. Various studies have been done for the agricultural area and crop identification based on the sigma backscatter coefficients. This analysis is best for target identification on the ground and the type of scattering. Most of the studies have relied on sigma backscatter for land cover classes; none of the literature has focussed on beta and gamma backscattering coefficients. Spatial, polarimetric, and derived features from the sentinel-1 SAR images were limited to only some studies. An attempt has been made to study the statistical significance of sigma, beta and gamma backscattering coefficients. A rich set of features considering the spatial, polarimetric, and derived features has been taken for the classification study.

III. METHODS AND MATERIALS

A. Source of data

The data source for this study is extracted from Copernicus Open Access Hub [14]. Sentinel-1 Band-C data product (S1A_IW_GRDH_1SDV_20180819T005203_20180819T005228_023308_0288EB_FEA8) dated 19/08/2018 with acquisition mode of Interferometric Wide swath (IW) and Level-1 GRD (Ground Range Detected) having dual bands (VV+VH) is selected for the analysis. The spatial resolution of 5m by 20 m is suitable for land cover studies.

B. Data Pre-processing.

The Sentinel-1 image is radio-metrically calibrated using the calibration coefficients included in the SAR data product [11][12]. It converts the pixel values in the image from the amplitude value to the corresponding radar backscatter. There are three important backscatter coefficients like (beta naught (β_θ), sigma naught (σ_θ), gamma naught (γ_θ). In this paper, a detailed statistical analysis [23] of the backscattering coefficients is represented for various land cover classes taken under consideration. The calibrated image is speckle filtered to remove the grainy noise. This image is terrain corrected to map precise geo-location of each pixel using SRTM 3 arc (90m) map projection. Texture information contained in the image is calculated using the Grey-Level Co-occurrence matrix (GLCM) operator in SNAP [15]. Six GLCM features (SigmaVH_Mean, SigmaVH_Variance, SigmaVH_Contrast, SigmaVV_Mean, SigmaVV_Variance, and SigmaVV_Contrast) are extracted for this analysis.

C. Dataset creation and feature extraction

The dataset creation step manually collects a total of 13500 points (latitude, longitude) for the seven data classes over the sentinel-1 image using the pin manager option in SNAP [15]. The corresponding pixel labels are assigned a

class label to the marked pins using the google earth engine. Eighteen important features like backscatter coefficients, spatial, textural, derived features as discussed below are extracted for each point. This data is used for the dataset visualization and statistical analysis supervised training and testing of classifiers.

- **Backscatter Coefficients**: Sigma VH, Sigma VV, Beta VH, Beta VV, Gamma VH, Gamma VV
- **Spatial Features**: elevation, incidence angle, elevation angle
- **Textural (GLCM) Features**: SigmaVH_Contrast, SigmaVH_Mean, SigmaVH_Variance, SigmaVV_Contrast, SigmaVV_Mean, SigmaVV_Variance
- **Derived Features**: Ratio Sigma (Sigma VH/Sigma VV) Ratio Beta (Beta VH/Beta VV), Ratio Gamma (Gamma VH/ Gamma VV)

D. Methods used for proposed work

The synthetic aperture radar sensor actively records the backscatter from different materials on the earth's surface. These land surfaces have a unique signature based on the interaction of the microwave C band signals. Various land use land cover mapping studies have been done on the SAR microwave data [13]. A useful statistical method Karl Pearson coefficient of correlation, is used to study the dependence and relationship between the land cover classes. After the statistical analysis of the backscatter data, supervised classification models such as Cubic SVM, ensemble bagged trees, and ensemble boosted trees are trained to accurately classify the SAR image into the 7 land cover classes.

Karl Pearson coefficient of correlation

Karl Pearson Coefficient of correlation is a statistical measure to study the linear relationship between two variables X & Y [16]. It is a useful method to measure the strength of the relationship between related linear variables [16].

Let X & Y be two variables for n number of variables, then Karl Pearson coefficient of correlation is given by the Eq. (1)

$$r_{XY} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}} \quad (1)$$

SVM is a supervised classification technique for image recognition that aims to find the best hyperplane to separate the different classes in a multi-dimensional space. Its capacity to handle large dimensionality data with few training samples is a crucial advantage for Remote Sensing applications. A radial basis function(rbf) is set as the default kernel to generate a non-linear hyperplane. But this study uses the cubic kernel function, which transforms the backscattering features in the dataset making class separation easy by simplifying the non-linear decision boundaries. The model performance of this classifier can be improved by tuning two free parameters: C (penalty parameter) and (tolerance margin).SVM is trained using the cubic kernel given by Eq. (2) as the data points in the remote sensing images are not linearly separable.

$$k(x_1, x_2) = (x_1^T x_2 + 1)^3 \quad (2)$$

Ensemble Classifier

Ensemble Bagged Trees is a supervised machine learning algorithm that predicts correct class labels based on the ensemble of decision trees with bagging, i.e., majority voting or hard voting. The class output is viewed as a vote, and the class with the maximum votes is returned. Another realistic voting is soft voting, which averages the probabilities of all classes and keeps the class with the highest average probability. This ensemble technique considers homogeneous weak learners (base learners) who independently learn from each other in parallel to make a competent, strong learner. The concept of bagging reduces the statistical parameter of variance in the decision tree towards making a more robust classification. The first step is to use random sampling with replacement to produce multiple datasets. The second step involves creating numerous learners in parallel. All the learners are combined using an averaging or majority voting strategy in the last step.

Boosting is another meta-algorithm that considers the homogeneous weak learners, learn in a sequential and adaptive manner. Boosting models, unlike bagged trees, are trained progressively, resulting in an improvement at each step and a lower rate of error. Adaptive Boosting (AdaBoost) is probably the most often utilized sort of boosting algorithm. Initially, the method generates numerous datasets using random sampling and data replacement across weighted data. In step 2, construct the learners in a sequential manner. In step 3, all the learners are pooled using a weighted average technique.

Various metrics are used to test the performance of the classification models. Accuracy metric is the closeness of the measured results to their true values. It measures the percentage of the correct labels after the classification from the trained model given by Eq. (3)

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn} \quad (3)$$

tp = No. of true positives, tn= No. of true negatives, fp = No. of false positives, fn = No. of false negatives

AUC (Area under the curve) measures the capability of the classification model to differentiate between the classes. A high AUC score makes correct identification of the class labels. To create the AUC, a matrix containing 1,0, -1 is adjusted so that the cells with higher positive cases get 1, negative cases get a zero (0), and cases with 0.5 probability have -1. It can even be used as a summary of the ROC curve. The ROC curve is a pictorial representation of the true positive rate (TPR) versus the false positive rate (FPR) (FPR). The curve which is more towards the top left corner is a better classifier as compared to the curves which are aligned 45-degree to the diagonal. The above-discussed classification models are trained on the data, and performance testing is done by the metrics discussed in this section.

IV. DATASET ANALYSIS

The backscattering values recorded by the SAR sensor are systematically analyzed and logically visualized using statistical tools. The Data visualization section graphically depicts the SAR images in terms of probability distribution function (PDF) and Sigma VV and VH backscatter values. Following that, Section IV-B provides a detailed statistical analysis of sigma, beta, and gamma coefficients in order to comprehend the differences between the landforms.

A. Data Visualization

This section presents a visualization of SAR images in terms of data distribution. The probability density function (PDF) of Sigma VH and Sigma VV bands for each of the 7 classes are plotted as given in Fig. 1. On the basis of data visualization, the class label data is normally distributed. The water bodies have a mesokurtic right-skewed normal distribution. As the water body is a flat surface with the least surface roughness, the backscatter values are neither clustered around the center nor too scattered. The urban area and villages have a leptokurtic normal distribution, but urban areas have left-skewed data distribution with no straight asymptotic tail. But the villages have a straight asymptotic tail. The PDF for the village structures in the VH band is the highest at the mean. Due to the double bounce scatters, the apartment construction has a trapezoidal normal distribution. The VV band scattering is more sensitive to the concrete structures, and many pixels values are concentrated to the right of the 0.2 sigma value. PDF for the urban areas is high as compared to apartments in the VV band. Agricultural land and parks have a similar type of positive skewed normal distribution due to volume scattering from crops and grass. The agricultural areas have high PDF as compared to parks in VV and VH bands. The forests have a mesokurtic normal distribution with straight asymptotic tail and leptokurtic distribution for VV bands. The forest class gives a high PDF about the mean in the VV band. The next section explains the methodology of classification.

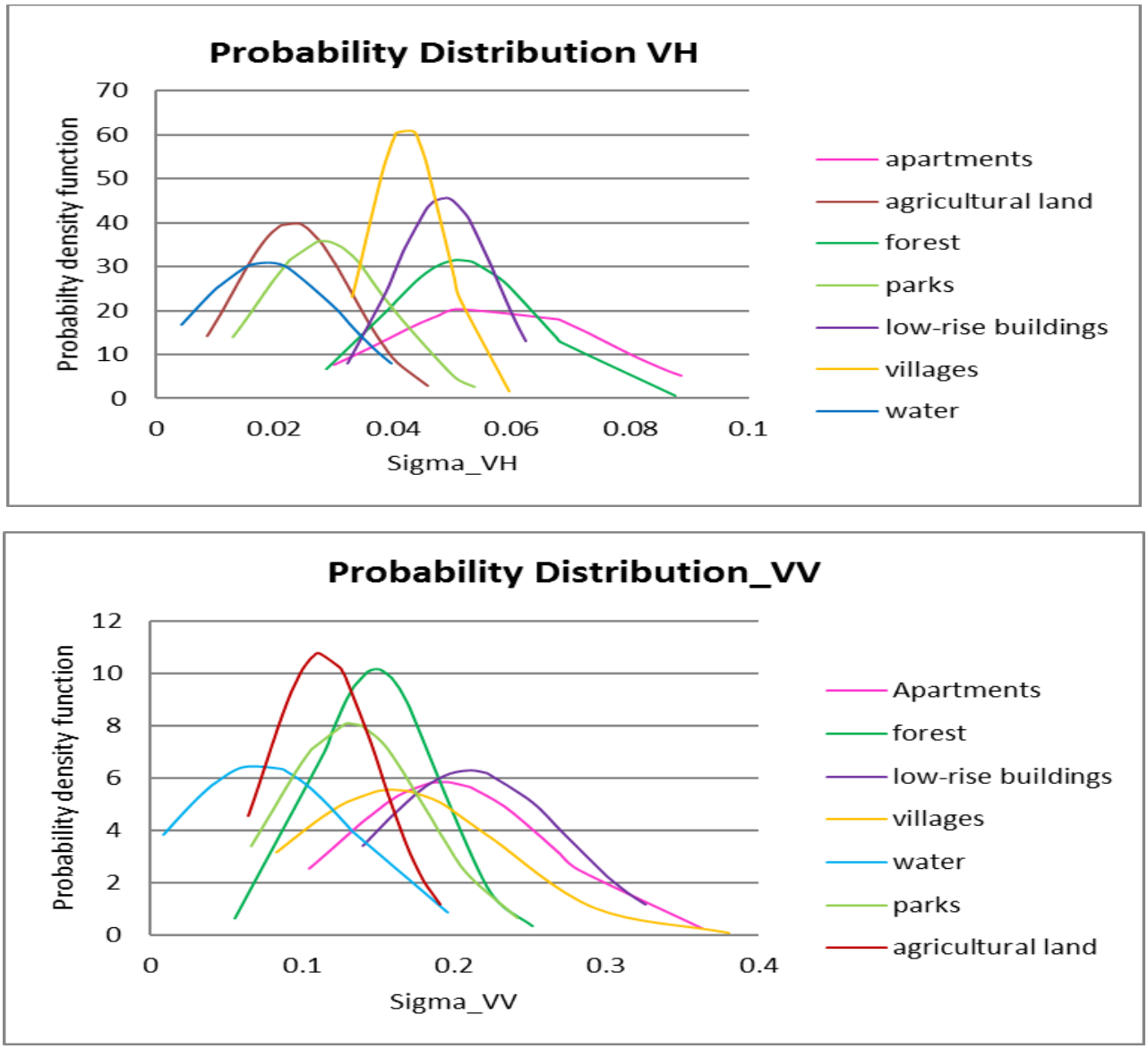


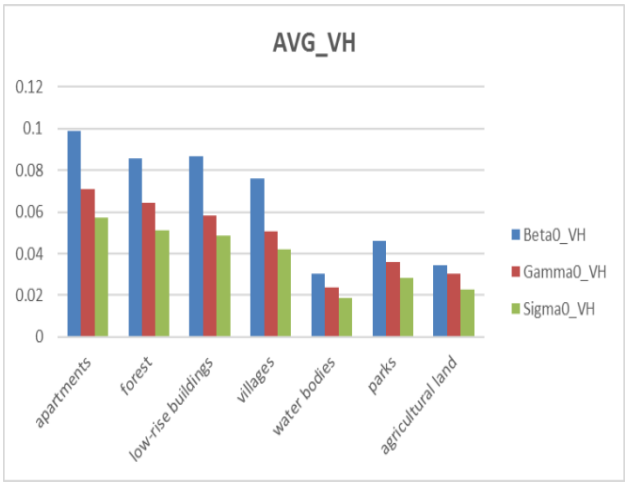
Fig 1. Probability distribution of Sigma VH and Sigma VV bands

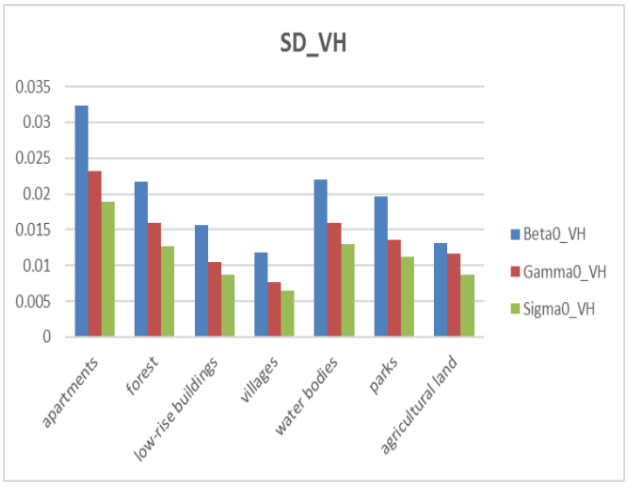
B. Statistical analysis of backscattering coefficient dataset

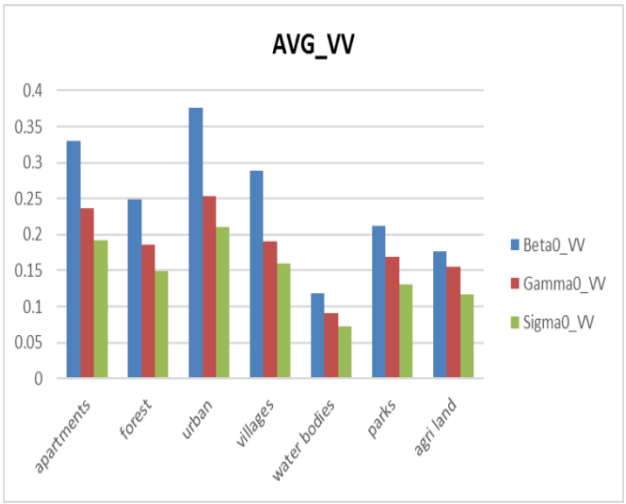
The statistical analysis of the Sentinal-1 bands (VV and VH) for different backscattering coefficients (Sigma, Beta, Gamma) is done extensively on the dataset obtained in Section III-C for seven land cover classes. Different statistics like average, standard deviation, Coefficient of Variation, and variance are calculated for VV, VH bands for each backscattering coefficient. Table I and Fig. 3 show the computed average backscatter values for the given bands. An extensive study is carried on the effect of the statistics and the bands on different land cover classes

TABLE I. AVERAGE BACKSCATTER VALUES FOR LAND COVER CLASSES

Coefficients	Apartments	Forest	Low-Rise Buildings	Villages	Water Bodies	Parks	Agricultural Land
Beta0_VH	0.098704798	0.085566	0.086858	0.076126	0.030531	0.045851	0.034499
Gamma0_VH	0.070707622	0.064296	0.058503	0.050533	0.023587	0.035983	0.030325
Sigma0_VH	0.057480853	0.051265	0.048523	0.042099	0.018619	0.02817	0.022776
Beta0_VV	0.330573174	0.248816	0.375587	0.288328	0.118981	0.211773	0.17736
Gamma0_VV	0.236834451	0.186026	0.253002	0.191204	0.090732	0.168685	0.155821
Sigma0_VV	0.192524112	0.148608	0.209835	0.159342	0.071976	0.13127	0.117059







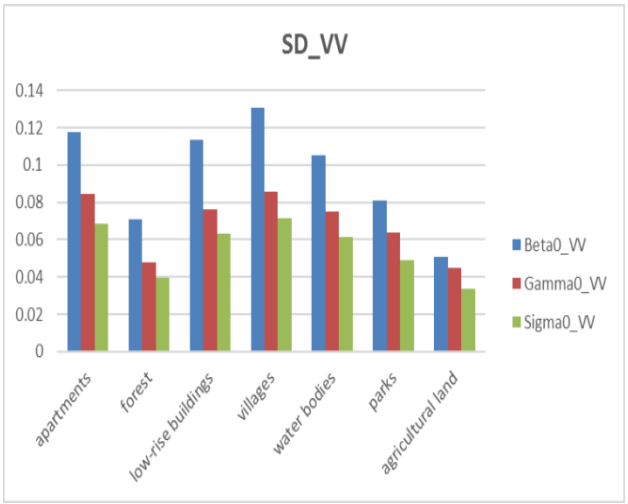


Fig 2. Average (AVG) backscatter values for VH and VV bands

Fig 3. Standard deviation (SD) backscatter values for VH and VV bands

The highest backscatter is recorded for the urban areas for the VV band. The VV band is more sensitive to the urban areas due to the double bounce as well as the corner reflectance from the high-rise or low-rise apartments. According to the graph, all the different categories of the houses (Apartments, low-rise buildings, villages) produce the high sigma, beta, and Gamma backscatter, clearly distinguishing it from other classes. The VH band has the highest value of backscatter for high-rise buildings. These structures appear bright in the satellite images having high

coherence. The mean of backscatter values for water bodies is the lowest among all classes, as shown in Fig. 2. The reason behind this is the surface scattering from smooth water surfaces such as ponds, rivers, canals. The water bodies appear dark in color due to the low backscatter. Mapping of the water bodies and groundwater quality can be done by the Geo-Spatial techniques. [22]

Now according to the statistic's standard deviation in Fig. 3 and variance in Fig. 4 of the land cover types, it is analyzed that villages have a high backscatter (sigma, beta, gamma) for the VV band as compared to other classes. But in the case of the VH bands, the apartments structures record a high backscatter for all VH bands, but the village structures have a low standard deviation. This is a useful characteristics parameter to differentiate urban areas from village structures. The cross-polarized bands of VH and polarized band VV have low standard deviation and variance for the agricultural lands due to the volume scattering by the crops in the fields.

The coefficient of variation (CV) is the proportion of the standard deviation to the mean, which is calculated in percentage. This CV value of backscatter is highest for the water bodies corresponding to 85.44 for Beta0_VV coefficients. The values for all the backscattering coefficients are almost the same for all classes in the case of VH and VV bands. Table II gives a comparison of the landcover classes considering the statistical parameters.

TABLE II. COMPARISON OF LAND COVER TYPES BASED ON STATISTICAL PARAMETERS

Land Cover Type	Statistical Parameters
Water	AVG_VH + AVG_VV low CV_VV+CV_VH high
Urban Areas	AVG_VV is highest
Apartments	AVG_VH + SD_VH + VAR_VH are highest
Villages	SD_VH+ VAR_VH low SD_VV +VAR_VV high
Forest	AVG_VH+AVG_VV+SD_VH+VAR_VH is high
Agricultural lands	AVG_VH+SD_VH+VAR_VH is high
Parks	CV_VH+CV_VV is highest in comparison with Forest & Agricultural Land

According to the analysis, it is observed that the VV band is more sensitive to capturing the urban structures. The VH band is useful in agricultural studies as the mean, standard deviation & variance for the VH band are high. The ratio band VH/VV is useful in the classification study of these classes as it improves the accuracy of classification. The next discussion is on Karl Pearson's coefficient of correlation on backscatter values.

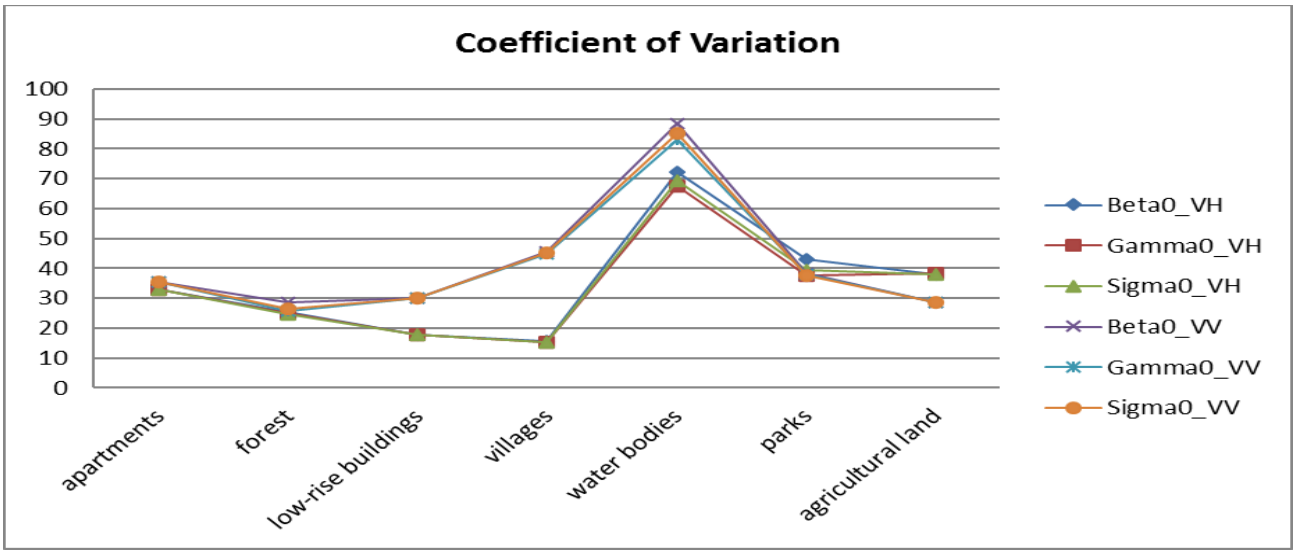


Fig 4. Coefficient of variation for VH and VV bands

The Karl Pearson correlation coefficient is computed for the backscattering values given in Table I.

Case 1:

Let X = Apartments, Y=Urban Areas

then,

$r_{XY} = 0.994 > 0.5$ using Eq. (1)

$r_{XY} = 0.994 > 0.5$

It indicates that there is a high degree of positive correlation, i.e., the strong relationship between apartments and urban areas

Case 2:

Let X = Apartments, Y=Villages

then,

$r_{XY} = 0.2881 < 0.5$ using Eq. (1)

It indicates that there is a low degree of positive correlation, i.e., a weak relationship between apartments and villages.

The study of the correlation coefficient in the above cases suggests that apartments and urban areas have a strong relationship (0.994) while apartments and villages are weakly related (0.2881).

V. CLASSIFICATION MODEL ON BACKSCATTERING DATASET

SAR microwave indices discriminate land cover classes, providing an incentive to develop a strong classification model for land cover research using remotely sensed data. Surface variables such as surface roughness, height above ground level, and soil moisture content, among others, influence the backscatter coefficients. To achieve this objective, the prepared dataset of 13500 samples is split in the ratio of 3:1 to prepare train and test datasets. The training dataset included 10125 samples and 18 column predictors representing seven different LULC classes, including apartments, villages, low-rise buildings, forests, parks, agricultural land, and water.

A. Experiments & Results

The backscattering coefficients for each class are distinguishable from each other. For the considered land cover classes, statistics such as mean, standard deviation, variance, coefficient of variation, and coefficient of correlation are displayed to generate a useful discussion to distinguish the classes. A number of classifiers were trained using the prepared training dataset. Different training experiments revealed that Cubic Support vector machine and ensemble bagged and boosted trees model are the best performing models with performance measured using 5-fold cross-validation method.

In this study, MATLAB was used to implement all of the classification models setting the parameters as given in Table III.

TABLE III. MODEL PARAMETERS

S. No	Model	Parameters
1.	Cubic SVM	Kernal function: Cubic Kernal scale: Automatic Box constraint level: 1 C= 100 ε = 0.01
2.	Ensemble bagged trees	Ensemble method: Bag Learner type: Decision tree Maximum number of splits: 1349 Number of learners: 30
3.	Ensemble boosted trees	Ensemble method: AdaBoost Learner type: Decision tree Maximum number of splits: 20 Number of learners: 30 Learning rate:0.1

Several classification algorithms such as Bayesian classification, SVM, ID3, ensemble methods were tried, and it was found that cubic SVM and ensemble decision tree with bagging and boosting were the best performers. The trained models are tested on the test data of size 3375 X 18. On the test data, the trained classifiers cubic SVM and ensemble decision tree with bagging and boosting performed with 93.6 percent, 95.3 percent, and 94.4 percent accuracy, respectively, as shown in Table IV.

TABLE IV. ACCURACY ASSESSMENT OF CLASSIFICATION MODELS OVER TEST DATASET

S. No.	Classification Method	Accuracy
1.	Ensemble Bagged trees	95.3%
2.	Ensemble Boosted trees	94.4%
3.	Cubic SVM	93.6%

TABLE V. PERFORMANCE OF ENSEMBLE BAGGED TREES, AND ENSEMBLE BOOSTED TREES, CUBIC SVM CLASSIFIER CLASSIFICATION MODELS IN %

Class label	Ensemble Bagged Tree		Ensemble Boosted Tree		Cubic SVM	
	TPR	FNR	TPR	FNR	TPR	FNR
Low-rise building	98.5	1.5	97	3	95.5	4.5
Agricultural land	94	6	98	2	96	4
Apartments	95.5	4.5	96	4	93.5	6.5
Forest	94.5	5.5	96	4	90	10
Parks	95.5	4.5	90	10	91.5	8.5
Villages	97.5	2.5	97.5	2.5	98	2
Water	90	10	86.7	13.3	89.3	10.7

Table V clearly represents the comparison and performance of the classifiers based on the TPR and FPR metrics. ROC curve for land cover classes is plotted for ensemble bagged trees in Fig. 5. The LULC class apartments, low-rise buildings, parks, and villages have AUC=1, which indicates that the ensemble bagged tree classifier was able to reliably categorize all positive and negative points. According to the experiments in this study, Cubic SVM has a better performance than the traditional SVM for SAR LULC classification [21]. The ensemble techniques of bagging and boosting have outperformed the simple decision tree classification. In this paper, the potential of the best non-linear machine learning algorithms was evaluated for LULC mapping over complex and high-dimensional SAR images.

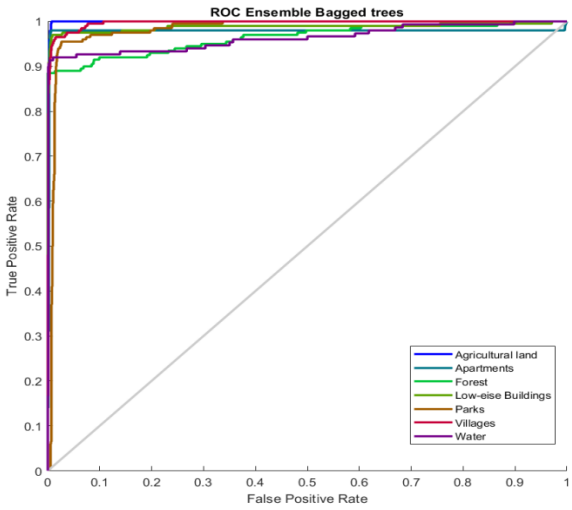


Fig 5. ROC curve for Ensemble Bagged trees

According to the results of this study, the ensemble bagged tree model outperformed all the other classifiers with high accuracy of 95.3% over the prepared test dataset.

CONCLUSION

This paper has focused on the suitability of backscattering coefficients for the training of surface area classifiers. Statistical analysis of the backscattering coefficients sigma, beta, and gamma is done for the seven land cover classes to reveal interesting conclusions. Statistics like average, standard deviation, variance, coefficient of variation are used to visualize and compare backscattering features for class categories. Taking into account the unique backscattering signatures of the land cover classes, classification models like Cubic SVM and ensemble bagged and boosted tree model were able to do classification with an accuracy of 93.6%, 95.3%, 94.4%, respectively. The correlation between the land cover classes is computed based on the Karl Pearson correlation coefficient. It can be concluded that apartment structures in urban areas have a high degree of positive correlation, indicating that more apartment-like structures are there in urban areas. A low degree of correlation between the apartments and villages indicates that there are very few apartment-like structures in the village area. So, there is a scope of development in the village areas to provide apartments, pucca houses to each individual to improve their standard of living.

REFERENCES

[1] Abdikan, S., Sekertekin, A., Ustunern, M., Sanli, F.B. and Nasirzadehdizaji, R., Backscatter analysis using multi-temporal sentinel-1 sar data for crop growth of maize in Konya basin, turkey. Int. Arch. Photogramm. Remote Sens. Spat. Inform. Sci, 42 (2018) 9-13.

[2] Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C. and Strauss, P., Sensitivity of Sentinel-1 backscatter to vegetation dynamics: An Austrian case study. Remote Sensing, 10(9) (2018) 1396.

[3] Paloscia, S.; Macelloni, G. Pampaloni, P. The relations between backscattering coefficient and biomass of narrow and wide leaf crops. In Proceedings of the 1998 IEEE International Geoscience and Remote Sensing Symposium Proceedings, Seattle, WA, USA, 6–10 July 1998; 1 (1998) 100–102.

[4] F. T. Ulaby, K. Sarabandi, K. C. McDonald, M. Whitt, and M. C. Dobson, Michigan microwave canopy scattering model Int. J. Remote Sensing, 11(7) (1990) 1223–1253.

[5] C. Dobson, F. Ulaby, T. Le Toan, A. Beaudoin, E. Kasiscke, and N. Christensen, A dependence of radar backscatter on coniferous forest biomass, IEEE Trans. Geosci. Remote Sensing, 30 (1992) 412–415. Jan. 1992

[6] P. Ferrazzoli and L. Guerriero, Radar sensitivity to tree geometry and woody volume: A model analysis, IEEE Trans. Geosci. Remote Sensing, 33 (1997) 360–371.

[7] Koppel, K., Zalite, K., Voormansik, K. and Jagdhuber, T., Sensitivity of Sentinel-1 backscatter to characteristics of buildings. International Journal of Remote Sensing, 38(22) (2017) 6298-6318.

[8] Dong, Y., B. Forster, and C. Ticehurst., Radar Backscatter Analysis for Urban Environments. International Journal of Remote Sensing 18 (6) (1997) 1351–1364. doi:10.1080/014311697218467.

[9] Charlton, M.B. and White, K., 2006. Sensitivity of radar backscatter to desert surface roughness. International Journal of Remote Sensing, 27(8) (2006) 1641-1659.

[10] Nasirzadehdizaji, R.; Balik Sanli, F.; Abdikan, S.; Cakir, Z.; Sekertekin, A.; Ustuner, M. Sensitivity Analysis of Multi-Temporal Sentinel-1 SAR Parameters to Crop Height and Canopy Coverage. Appl. Sci. , 9 (2019) 655.

[11] Kaushik, P., & Jabin, S., Deforestation Mapping Using MODIS Tree Cover Mask and Sentinel-1 Images. In Micro-Electronics and Telecommunication Engineering, (2021) 73-80. Springer, Singapore.

[12] Kaushik, P., & Jabin, S. (2018, December). A Comparative study of pre-processing techniques of SAR images. In 2018 4th International Conference on Computing Communication and Automation (ICCCA) (2018) 1-4. IEEE.

[13] Kaushik,P., & Jabin, S.(2021, August). A study of Surface Area Classification of Bhiwani area using Sentinel-1 Images. In International Journal of Scientific Research (IJSR), (2021)

[14] (2020) Copernicus Open Access Hub [Online]. Available: https://scihub.copernicus.eu/

[15] (2020) European Space Agency STEP: Science toolbox exploitation Platform [Online]. Available: https://step.esa.int/main/toolboxes/snap/

[16] (2020) Pearson's correlation coefficient [Online]. Available: https://www.oxfordreference.com/view/10.1093/oi/authority.20110803100313153

[17] Constantino-Recillas, D. E., Arizmendi-Vasconcelos, E., Monsiváis-Huertero, A., Jiménez-Escalona, J. C., Torres-Gómez, A. C., De La Rosa-Montero, I. E., ... & Judge, J., Understanding the Backscattering from Sentinel-1 Over a Growing Season of Corn in Central Mexico Using the Thexmex Datasets. In IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium (2020) 4526-4529. IEEE.

[18] Modanesi, S., Massari, C., Gruber, A., Lievens, H., Tarpanelli, A., Morbidelli, R., & De Lannoy, G. J., Optimizing a backscatter forward operator using Sentinel-1 data over irrigated land. Hydrology and Earth System Sciences Discussions, (2021) 1-39.

[19] He, Z., Li, S., Deng, Y., Zhai, P., & Hu, Y. (2021, July). Rice Paddy Fields Identification Based on Backscatter Features of Quad-Pol RADARSAT-2 Data and Simple Decision Tree Method. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, (2021) 6765-6768. IEEE.

[20] Delgado Blasco, J. M., Fitrzyk, M., Patruno, J., Ruiz-Armenteros, A. M., & Marconcini, M., Effects on the double bounce detection in urban areas based on SAR polarimetric characteristics. Remote Sensing, 12(7) (2020) 1187.

[21] S.V.Hwan, Dr. Eun-Kyung JO., Survey of Land Use and Land Cover Change Detection using Remote Sensing. SSRG International Journal of Geoinformatics and Geological Science, 1(2) (2014) 5-9.

[22] F. Lalbiakmawia., Ground Water Quality Mapping of Kolasib District, Mizoram, India Using Geo-Spatial Technology. SSRG International Journal of Geoinformatics and Geological Science 2(2) (2015) 1-7.

[23] S.Anandharaj, Dr.C.Sulaxna sharma., Urbanization in India by using Remote sensing and GIS techniques. SSRG International Journal of Geoinformatics and Geological Sci*ence* 3(2) (2016) 1-5.