Fully Polarimetric Land Cover Classification Based on Hidden Markov Models Trained with Multiple Observations

Konstantinos Karachristos, Georgia Koukiou, Vassilis Anastassopoulos

Electronics Laboratory (ELLAB), Physics Department, University of Patras, Patras, Greece
Email: phy5691@upnet.gr, gkoukiou@upatras.gr, vassilis@upatras.gr

Abstract

A land cover classification procedure is presented utilizing the information content of fully polarimetric SAR images. The Cameron coherent target decomposition (CTD) is employed to characterize each pixel, using a set of canonical scattering mechanisms in order to describe the physical properties of the scatterer. The novelty of the proposed classification approach lies on the use of Hidden Markov Models (HMM) to uniquely characterize each type of land cover. The motivation to this approach is the investigation of the alternation between scattering mechanisms from SAR pixel to pixel. Depending on the observations-scattering mechanisms and exploiting the transitions between the scattering mechanisms we decide upon the HMM-land cover type. The classification process is based on the likelihood of observation sequences been evaluated by each model. The performance of the classification approach is assessed my means of fully polarimetric SLC SAR data from the broader area of Vancouver, Canada and was found satisfactory, reaching a success from 87% to over 99%.

Keywords

Fully Polarimetric SAR, Coherent Decomposition, Land Cover Classification, Hidden Markov Models, Remote Sensing

1. Introduction

In the past several decades, Remote Sensing has gradually broadened being the cornerstone in a plethora of research topics. By now, more than 150 Earth-observation satellites are currently in orbit, carrying sensors that provide continuously high-quality data. Synthetic Aperture Radar (SAR) is an advanced technology in

Earth Observation with a wide range of application from land cover/land use classification [1] to flood and fire detection [2] [3].

SAR imaging is different to how optical imaging works. A radar antenna actively sends electromagnetic waves towards its target and then measures what is backscattered. SAR takes advantage of EM waves to penetrate clouds, foliage and even upper layers of Earth's surface, while the data acquisition can be made at any time of day or night. A "Synthetic Aperture" means physically moving the small SAR instrument over an area while it gathers information of its target. To achieve this, SAR antennas are often mounted on satellites and airplanes which enables them to cover a large area in a very short time. The radar antenna records the strength and the time delay of these return signals. SAR system utilizes specially designed antenna to transmit and receive a radar wave of a specific polarization.

Fully polarimetric SAR is preferred for complete target analysis since it provides the backscattering behavior of the electromagnetic scatterer. The information about the target surface can be retrieved based on its response in different polarization states. When a polarized radar wave interacts with the earth's surface, the polarization of the wave is modified depending upon the specific characteristics of the surface. This includes its geometrical structure, shape, reflectivity, orientation as well as the geophysical properties such as moisture content, surface roughness etc. One of the main advantages of polarimetric techniques is the possibility to separate scattering contributions of different nature, which can be associated to certain elementary scattering mechanisms. Elementary radar scatterers are represented by polarization scattering matrices that contains all the scattering information. In order to extract this information, the matrix decomposition is needed.

So far, plenty of target decomposition approaches have been proposed in the literature. These techniques are broadly classified into coherent and incoherent decompositions. Coherent decomposition methods were developed to characterize completely polarized scattered waves, for which fully polarimetric information is contained in the scattering matrix. Hence, there is no need of second order statistics. On the other hand, incoherent decomposition techniques were developed to characterize a large number of statistically independent scatterers, randomly distributed and with none of them being dominant. In such a case the second order statistics are required. Cloude and Pottier [4] proposed a method that relies on an eigenvalue analysis of the coherency matrix and introduced the parameters of entropy, anisotropy, a and β angle. Freeman and Durden [5] proposed a unique incoherent decomposition. Touzi [6] developed an approach that combines three simple scattering mechanisms to a polarimetric SAR observation. On the other side, stands out the Cameron's Coherent Target Decomposition (CTD) [7] [8] which is employed in this work.

One of the major topics being investigated in the field of remote sensing, is the land cover classification as a consequence of the interaction between rapidly grown population and environment. The necessity to control the rate of changes

and improve the environmental management brought to the fore the need for high accuracy in land cover classification. Many algorithms have been examined and found successful, such as artificial neural networks [9], random forests [10], support vector machines with the use of radial basis function kernel [11]. Recently, Fully Polarimetric Land Cover Classification based on Markov Chains was proposed [12] which gives a high rate of success in discriminating between different land cover types.

In this work, we present a novel land cover classification method based on Hidden Markov Models. We assumed that the observations of HMM could be the scattering mechanisms at each pixel while the hidden states stand for the underlying land cover type. Therefore, we developed the hypothesis that each land cover type is uniquely characterized by an HMM. At this point, is important to take into consideration that this work differs from others, firstly as for the parallelism we follow with the components of an HMM land cover types and scattering mechanisms and secondly on the way the HMMs were trained. The classification process was carried out by the HMM's solution of the evaluation problem. The novelty of the proposed method was challenged from the need to investigate for spatially extended targets, as land cover types, which led us to Hidden Markov Models. Therefore, we could construct an HMM, considering as observations the scattering mechanisms which are emitted from each type of land cover.

The paper is organized as follows. Cameron's CTD is explained in Section 2 while in Section 3 a brief analysis of HMMs is presented. Section 4 refers to the preprocessing stage and in Section 5 we present the experimental procedure with the points of novelty and comparisons. The conclusions are drawn in Section 6.

2. Cameron's Coherent Target Decomposition

Cameron proposed a coherent target decomposition into elementary scatterers that is based on the properties of reciprocity and symmetry. The first stage is to decompose the scattering matrix S that represents a SAR pixel into reciprocal and non-reciprocal components via an angle θ_{rec}. The second stage considers decomposition of the reciprocal term into two further components, namely symmetric and non-symmetric via an angle τ_{sym}. The Cameron's Decomposition takes the following form:

$$\vec{S} = \alpha \left\{ \cos\theta_{rec} \left\{ \cos\tau_{sym} \hat{S}_{sym}^{max} + \sin\tau_{sym} \hat{S}_{sym}^{min} \right\} + \sin\theta_{rec} \hat{S}_{nonrec} \right\} \tag{1}$$

where the scalar $a = \left\| \vec{S} \right\|_2^2 = span(S)$,

The angle θ_{rec} represents the degree to which the scattering matrix obeys the reciprocity principle,

The angle τ_{sym} represents the degree to which the scattering matrix deviates from the set of scattering matrices corresponding to symmetric scatterers,

$\hat{S}_{nonrec}$ represents the normalized nonreciprocal components,

$\hat{S}_{sym}^{max}$ the normalized maximum symmetric component and the $\hat{S}_{sym}^{min}$ the nor-

malized minimum symmetric component.

The maximum symmetric component can be transformed into a normalized complex vector $\hat{\Lambda}(z)$ with z being referred to as a complex parameter (given in Table 1) that eventually determines the scattering mechanism. The normalized complex vector $\hat{\Lambda}(z)$ is given by:

$$\hat{\Lambda}(z) = \frac{1}{\sqrt{1+|z|^2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ z \end{bmatrix} \tag{2}$$

Cameron in order to determine the scattering mechanism of each target z considered the following metric distance from a reference scatterer

$$d(z, z_{ref}) = \sin^{-1}\left(\min\left[d_-(z, z_{ref}), d_*(z, z_{ref})\right]\right) \tag{3}$$

where

$$d_-(z, z_{ref}) = \sqrt{\frac{|z - z_{ref}|^2}{(1+|z|)^2(1+|z_{ref}|)^2}} \tag{4}$$

$$d_*(z, z_{ref}) = \sqrt{\frac{|z - z_{ref}^*|^2 + (1-|z|^2)(1-|z_{ref}|^2)}{(1+|z|)^2(1+|z_{ref}^*|)^2}} \tag{5}$$

3. Hidden Markov Models

A Hidden Markov Model [13] [14] [15] [16] is a statistical model with an underlying stochastic process that is not observable (it is hidden) but can only be observed through another set of stochastic processes that produce a sequence of emitted observation symbols (outputs).

A first order Hidden Markov Model can be represented [16] by the compact notation $\lambda = (A, B, \pi)$. Specification of HMM requires the choice of the length

Table 1. Cameron's scattering mechanisms.

Elementary Scatterer Type (Reference)	Normalized Complex Vector	Complex Parameter z
Trihedral	$\hat{\Lambda}(1)$	1
Dihedral	$\hat{\Lambda}(-1)$	-1
Dipole	$\hat{\Lambda}(0)$	0
Cylinder	$\hat{\Lambda}\left(+\frac{1}{2}\right)$	$+\frac{1}{2}$
Narrow diplane	$\hat{\Lambda}\left(-\frac{1}{2}\right)$	$-\frac{1}{2}$
1/4 wave devise	$\hat{\Lambda}(\pm j)$	$\pm j$

of the observation sequence T, the number of states N in which the underlying process takes place namely hidden states, the number of discrete symbols M that can be observed and the specification of the three probability densities A, B and π. Let assume for the set of hidden states and possible emitted symbols $S = \{S_1, S_2, \cdots, S_N\}$ and $O = \{o_1, o_2, \cdots, o_M\}$ respectively and denote the actual state at time instants $t = 1, 2, \cdots$ as q_t. The state transition probability distribution:

$$A = \{a_{ij}\} \quad \text{(Transition Matrix)} \tag{6}$$

where for $1 \le i, j \le N$

$$a_{ij} = P\left[q_{t+1} = S_j \mid q_t = S_i\right] \tag{7}$$

corresponds to the probability that the HMM is in state S_i at time instance t and it makes a transition to state S_j at time instant $t+1$.

The observation symbol probability distribution in state j (Emission Matrix):

$$B = \{b_j(k)\} \tag{8}$$

where for $1 \le j \le N$, $1 \le k \le M$

$$b_j(k) = P\left[O_k \text{ at } t \mid q_j \text{ at } t\right] \tag{9}$$

corresponds to the probability that the observation symbol O_k is emitted in state q_j at time instant t, while an initial state probability distribution is defined as [16]:

$$\pi = \{\pi_i\} \tag{10}$$

where for $1 \le i \le N$

$$\pi_i = P\left[q_1 = S_i\right] \tag{11}$$

$$\begin{cases} 0 \le \pi_i \\ \sum_{t=1}^{N} \pi_t = 1 \end{cases} \tag{12}$$

An HMM can be classified into different types in the light of its state transition. Two well-known types of models are

- Ergodic model: An ergodic model is a fully connected HMM, that means that every state of the model can be reached from every other state of the model.
- Left-right Bakis model: A left-right model has only partial state transition such that $a_{ij} = 0 \quad \forall j < i$.

There are many possible variations and combinations. In the current study we make use of ergodic Hidden Markov Models. Rabiner [13] introduced the idea that Hidden Markov Models should be characterized by three fundamental problems:

1) Given the observation sequence $O = O_1, O_2, \cdots, O_T$ and the model $\lambda = (A, B, \pi)$ how we compute $P(O, \lambda)$, the probability of the observation sequence, known as Evaluation Problem.

2) Given the observation sequence $O = O_1, O_2, \cdots, O_T$, how we choose a state

sequence which is optimal in some meaningful sense, known as Decoding Problem

3) How we adjust the model parameters $\lambda = (A, B, \pi)$ to maximize $P(O \mid \lambda)$, known as Training/Learning Problem.

The introduction of Evaluation and Learning problem follows, as these two were used on our experimental procedure. The observation-evaluation problem can be solved using the Forward-Backward procedure, in terms of a forward and backward variable. The definition of a forward variable is [16]:

$$a_t(i) = P(O_1 O_2 \cdots O_t, q_t = S_i \mid \lambda) \tag{13}$$

$a_t(i)$ can be solved inductively:

1) Initialization:

$$a_1(i) = \pi_i b_i(O_1), 1 \le i \le N \tag{14}$$

2) Induction for $t = 1, \cdots, T-1$, and for $j = 1, \cdots, N$

$$a_{t+1}(j) = \left(\sum_{i=1}^{N} a_t(i) a_{ij} \right) b_j(O_{t+1}) \tag{15}$$

3) Evaluating the probability

$$P(O \mid \lambda) = \sum_{i=1}^{N} a_T(i) \tag{16}$$

The Backward algorithm is analogous to the Forward algorithm with the difference that it starts at the end and works backwards to the beginning.

Given any finite observation sequence as training data, the model's parameters should be adjusted appropriately in order to find the model that fits best the given observation sequence. A classic approach is the Baum-Welch method. It is important to know that this algorithm finds a local maximum for $P(O \mid \lambda)$, but it doesn't guarantee a global maximum.

This method makes use of the forward-backward algorithm, together with the following temporary variables [16]:

1) $\gamma_t(i, j)$ which is the probability of being in state q_i at time t and in state q_j at time $t+1$

Where

$$\gamma_t(i, j) = P(S_t = q_i, S_{t+1} = q_j \mid O, \lambda) = \frac{\alpha_t a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{P(O \mid \lambda)} \tag{17}$$

Summing $\gamma_t(i, j)$ over t can be interpreted as the expected number of transitions from state q_i to state q_j given the model's parameters and the observation sequence *O*.

2) The probability of being in each state at time *t*

$$\gamma_t(i) = P(x_t = q_i \mid O, \lambda) = \frac{a_t(i) \beta_t(i)}{P(O \mid \lambda)} \tag{18}$$

Summing $\gamma_t(i)$ over *t* can be interpreted as the expected number of times that the state q_i is visited or the expected number of transitions made from state q_i, given the model parameters and the observation sequence *O*.

By using these two variables the re-estimation procedure has the following definition:

1) For $i = 1, \cdots, N$ let

$$\pi_i = \gamma_1(i) \tag{19}$$

2) For $i = 1, \cdots, N$ and $j = 1, \cdots, N$

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \gamma_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \tag{20}$$

3) For $j = 1, \cdots, N$ and $k = 1, \cdots, M$

$$b_j(k) = \frac{\sum_{t \in \{1, \cdots, T\}, O_t = k} \gamma_t(i, j)}{\sum_{t=1}^{T} \gamma_t(i)} \tag{21}$$

The numerator of the re-estimated b is the expected number of times the model is in state q_j and the emitted symbol is k. Divided by the expected number of times the state q_i is visited, the probability of emission symbol k is obtained.

Then the re-estimation is defined as an iterative process [16]:

1) Initialize $\lambda = (A, B, \pi)$, if no reasonable guess is available the values are randomly chosen:

$$\pi_i = \frac{1}{N}, \quad a_{ij} = \frac{1}{N}, \quad b_j(k) = \frac{1}{M}$$

2) Compute $a_t(i)$, $\beta_t(i)$, $\gamma_t(i, j)$ and $\gamma_t(i)$.

3) Re-estimate the model $\lambda = (A, B, \pi)$.

4) If $P(O | \lambda)$ increases, repeat from 2 with the new re-estimated parameters.

4. Preprocessing

In this work we made use of the fully polarimetric SAR data from the broader area of Vancouver city in Canada. The used fully polarimetric SAR imagery is of the Wide Fine Quad-Pol SLC product [17]. As a first processing stage, by using the SNAP's environment (SeNtinel Application Platform) a radiometric calibration was employed to convert raw digital image data from satellite to a common physical scale based on known reflectance measurements taken from objects on the ground surface. Figure 1(a) depicts the SLC data while Figure 1(b) the calibrated data using SNAP. Then we applied the CTD method to characterize land cover pixel by pixel. As it was mentioned before this characterization results to a total of 8 simple geometric structures.

The partitioning in different types of land cover was made based on the geological features of the specific area and on global criteria so that the proposed classification process to be robust and effective in every other data set. Therefore, we selected 4 types of land cover, namely, water, urban/built up areas, forest/wooded area, agriculture/pasture area, as shown in Figure 2.
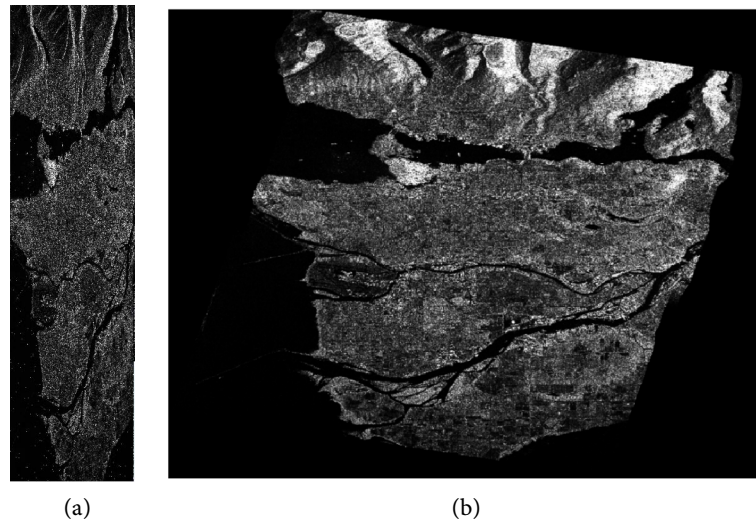
Figure 1. (a) SLC SAR product; (b) Data projected into geographic coordinates using SNAP (longitude, latitude).
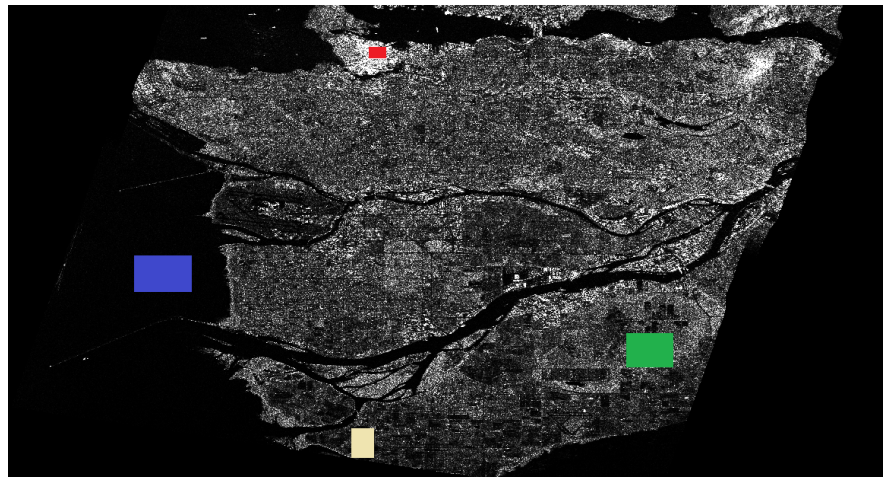
Figure 2. The selected four types of Land cover. Blue marked region corresponds to water land cover, green to forest/wooded area, red corresponds to urban/built-up area and beige to agriculture/pasture land cover.

In order to carry out this type of supervised classification there is a need for ground truth data. This was accomplished by applying range doppler terrain correction on SNAP. Accordingly, we were able to match the SLC data to Google Earth maps and mark specific regions that correspond to the 4 types of land cover that were mentioned above. Then, a train-test split procedure was used with the ratio of 80/20, respectively for each land cover type to separate the data sequences into those that will be used for training the HMMs and for evaluating the trained models.

5. Points of Novelty and Classification Procedure

The novelty of this study is in the parallelism between our data and the components of HMM that inspires the use of this model to classify different types of

land cover. Specifically, we assume that each pixel which corresponds to a Cameron's elementary scattering mechanism, constitutes an observation and each type of land cover is a hidden state. Therefore, we could construct an HMM, considering as observations the scattering mechanisms which are emitted from each type of land cover.

The major advantage of HMM lies on the information contained in the way the elementary scattering mechanisms are alternating from pixel to pixel in each region. The importance of mining this feature relies on the fact that a neighbor of an elementary scattering mechanism, regardless of whether it is different, it mainly corresponds to the same land cover type. The transitions matrices for the scattering mechanisms is a strong feature for the specific data and play the role of probabilistic weighting during the training and classification process.

The main component of this study is to construct an HMM for each land cover type and classify a given observation sequence by determining the likelihood $P(O|\lambda)$ for every model with the forward/backward algorithm. This allows us to choose the model which best matches the observations.

To construct the model that fits best a given observation sequence we used the Baum-Welch algorithm, to find the unknown parameters of every HMM by maximizing the probability of the given observation sequence. In our case, we have a multiple observation training as a consequence of assuming that each region consists of a set of observation sequences. For this task we obtain the Levinson's equations:

1) State transition probability:

$$a_{ij} = \frac{\sum_{k=1}^{K} \sum_{t=1}^{T_k-1} \gamma_t^{(k)}(i,j)}{\sum_{k=1}^{K} \sum_{t=1}^{T_k-1} \gamma_t^{(k)}(i)}, \quad i=1,\cdots,N \quad \text{and} \quad j=1,\cdots,N \tag{22}$$

2) Symbol emission probability:

$$b_j(i) = \frac{\sum_{k=1}^{K} \sum_{t=1,o_t^{(k)}=v_i}^{T_k-1} \gamma_t^{(k)}(j)}{\sum_{k=1}^{K} \sum_{t=1}^{T_k} \gamma_t^{(k)}(i)}, \quad i=1,\cdots,N \quad \text{and} \quad j=1,\cdots,N \tag{23}$$

3) Initial state probability:

$$\pi_t = \frac{1}{K} \sum_{k=1}^{K} \gamma_1^{(k)}(i), 1 \le i \le N \tag{24}$$

Scaling of computation was needed to avoid underflow and overflow [16].

According to the above material, from the four different types of land cover we chosen, we calculated the emission matrix shown in Table 2. Each row corresponds to a specific land cover type and each column to the rate that a scattering mechanism was obtained.

From the emission matrix it can be concluded that we cannot be based on the dominant scattering mechanism at each land cover type. Firstly, because we have the same dominant scattering mechanisms between classes 1 and 4 and between classes 2 and 3, and secondly, because the ratio of dominance in class 2 and class 3 is comparable with the rest.

After the learning process, four HMMs have been trained based on the sets of observations each of which corresponds totally to a specific type of land cover. The transition and emission matrices for the hidden states of each model were calculated and are depicted in Tables 3-10.

Model 1-Water areas

Model 2-Urban/Built up areas

Model 3-Forest/Wooded areas

Model 4-Agriculture/Pasture areas

The classification performance is based on the computation of the likelihood $P(O|\lambda)$. By giving an observation at a time we can evaluate each model with the Forward/Backward algorithm, as mentioned before and find the model which best fits the observation. The accuracy rate for this experiment is given in Table 11.

Table 2. Observation rate for each of the 8 scattering mechanisms in each type of selected land cover, for the specific experimental procedure.

	Trihedral	Dihedral	Dipole	Cylinder	Narrow Diplane	¼ wave Device	Left Helix	Right Helix
water	0.7319	0.0015	0.0179	0.2067	0.0051	0.0360	0.0006	0.0004
urban	0.0257	0.0646	0.2299	0.1496	0.1564	0.2777	0.0581	0.0378
forest	0.1048	0.0255	0.1713	0.2891	0.0756	0.2641	0.0406	0.0290
agriculture	0.5156	0.0033	0.0448	0.3173	0.0141	0.0926	0.0073	0.0049

Table 3. Hidden state transition matrix for the HMM corresponding to water land cover type.

0.443415	0.065094	0.132725	0.358765
0.383188	0.106376	0.17453	0.335906
0.403767	0.091818	0.160338	0.344078
0.436102	0.069939	0.137835	0.356124

Table 4. Emission matrix for the HMM corresponding to water land cover type.

0.729186	0.001607	0.01833	0.209171	0.004621	0.03598	0.000728	0.000377
0.729170	0.001608	0.018333	0.209169	0.004619	0.035999	0.000729	0.000373
0.729178	0.001607	0.018332	0.209170	0.004620	0.035989	0.000728	0.000375
0.729185	0.001607	0.01833	0.209171	0.004620	0.035981	0.000728	0.000377

Table 5. Hidden state transition matrix for the HMM corresponding to urban/built up land cover type.

0.115817	0.375477	0.333571	0.175135
0.061987	0.474750	0.345442	0.117822
0.071501	0.452814	0.346178	0.129506
0.097580	0.404473	0.340206	0.157741

Table 6. Emission matrix for the HMM corresponding to urban/built up land cover type.

0.024321	0.070354	0.216769	0.146470	0.160998	0.283256	0.059278	0.038554
0.024202	0.070260	0.216814	0.146180	0.161225	0.283543	0.059195	0.038581
0.024211	0.070264	0.216811	0.146205	0.161205	0.283520	0.059203	0.038579
0.024269	0.070310	0.216789	0.146347	0.161094	0.283380	0.059244	0.038567

Table 7. Hidden state transition matrix for the HMM corresponding to forest/wooded land cover type.

0.205632	0.260550	0.288928	0.244889
0.107232	0.391872	0.336626	0.164271
0.127326	0.361515	0.328405	0.182754
0.178667	0.292863	0.303399	0.225072

Table 8. Emission matrix for the HMM corresponding to forest/wooded land cover type.

0.104692	0.026346	0.173721	0.288036	0.074499	0.262646	0.041339	0.028719
0.104675	0.026344	0.173721	0.288025	0.074524	0.262636	0.041334	0.028741
0.104679	0.026345	0.173721	0.288028	0.074518	0.262638	0.041335	0.028736
0.104687	0.026346	0.173721	0.288033	0.074506	0.262644	0.041338	0.028725

Table 9. Hidden state transition matrix for the HMM corresponding to agriculture/pasture land cover type.

0.424838	0.00429	0.196397	0.374475
0.015879	0.957907	0.011113	0.015101
0.403306	0.005902	0.223177	0.367615
0.420462	0.004575	0.201856	0.373107

Table 10. Emission matrix for the HMM corresponding to agriculture/pasture land cover type.

0.543129	0.003126	0.040192	0.307221	0.011389	0.084589	0.006255	0.004098
0.505043	0.003421	0.049044	0.306076	0.017145	0.106073	0.010006	0.003192
0.542996	0.003128	0.040215	0.307207	0.011399	0.084691	0.006271	0.004094
0.543108	0.003126	0.040196	0.307219	0.011391	0.084605	0.006258	0.004098

Table 11. Classification accuracy of hidden markov models.

Land cover type	Classification Accuracy
Water areas	99.73%
Urban/built up areas	99.09%
Forest/wooded areas	99.35%
Agriculture/pasture areas	87.37%

6. Conclusions

In the current work, we examined the applicability of the Hidden Markov Models in land cover classification and assess their performance. Our study was based on the Cameron's Coherent Target Decomposition for fully polarimetric SAR images, a powerful tool especially in target detection applications. The need to investigate for spatial extended targets, as land cover types, led us to Hidden Markov Models. The main reason was to take advantage of the Markovian property, the rich mathematical structure which presents a well-fitting analogy with our task. The results were excellent, as the classification performance is over 87% and in many of the cases it reached 99%. In total, our hypothesis was formed successfully in the specific dataset and the process we present is effective. The procedure was completed in the SNAP environment for the preprocessing level and in MATLAB as for the extraction of scattering mechanisms, the learning of HMMs and the classification procedure.

Future extension of the method proposed in this manuscript is already in a thorough testing, where the representation of each pixel with more than one elementary scatterer is investigated. To implement this approach, it is necessary to develop a more complex HMM model for land cover classification. This model is currently being searched.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Qi, Z., Gar-On Yeh, A., Li, X. and Lin, Z. (2011) A Novel Algorithm for Land Use and Land Cover Classification Using RADARSAT-2 Polarimetric SAR Data. *Remote Sensing of Environment*, **118**, 31-39. https://doi.org/10.1016/j.rse.2011.11.001

[2] Ban, Y., Zhang, P., Nascetti, A., Bevington, A.R. and Wulder, M.A. (2020) Near Real-Time Wildfire Progression Monitoring with Sentinel-1 SAR Time Series and Deep Learning. *Scientific Reports*, **10**, Article No. 1322.
https://doi.org/10.1038/s41598-019-56967-x

[3] Kussul, N., Shelestov, A. and Skakun, S. (2011) Flood Monitoring from SAR Data. NATO Science for Peace and Security Series C: Environmental Security, Springer Science + Business Media B.V., Berlin.
https://doi.org/10.1007/978-90-481-9618-0_3

[4] Cloude, S.R. and Pottier, E. (1997) An Entropy Based Classification Scheme for Land Applications of Polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, **35**, 68-78. https://doi.org/10.1109/36.551935

[5] Freema, A. and Durden, S.L. (1998) A Three-Component Scattering Model for Polarimetric SAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, **36**, 963-973. https://doi.org/10.1109/36.673687

[6] Touzi, R. (2006) Target Scattering Decomposition in Terms of Roll-Invariant Target Parameters. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 73-84.
https://doi.org/10.1109/TGRS.2006.886176

[7] Cameron, W.L., Youssef, N.N. and Leung, L.K. (1996) Simulated Polarimetric Signatures of Primitive Geometrical Shapes. *IEEE Transactions on Geoscience and Remote Sensing*, **34**, 793-803. https://doi.org/10.1109/36.499784

[8] Cameron, W.L. and Rais, H. (2006) Conservative Polarimetric Scatterers and Their Role in Incorrect Extensions of the Cameron Decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 3506-3516. https://doi.org/10.1109/TGRS.2006.879115

[9] Civco, D.L. (1993) Artificial Neural Networks for Land-Cover Classification and Mapping. *International Journal of Geographical Information Systems*, **7**, 173-186. https://doi.org/10.1080/02693799308901949

[10] Gislason, P.O., Benediktsson, J.A. and Sveinsson, J.R. (2006) Random Forests for Land Cover Classification. *Pattern Recognition Letters*, **27**, 294-300. https://doi.org/10.1016/j.patrec.2005.08.011

[11] Kavzoglu, T. and Colkesen, I. (2009) A Kernel Functions Analysis for Support Vector Machines for Land Cover Classification. *International Journal of Applied Earth Observation and Geoinformation*, **11**, 352-359. https://doi.org/10.1016/j.jag.2009.06.002

[12] Koukiou, G. and Anastassopoulos, V. (2021) Fully Polarimetric Land Cover Classification Based on Markov Chains. *Advance in Remote Sensing*, **10**, 47-65. https://doi.org/10.4236/ars.2021.103003

[13] Rabiner, L.R. (1989) A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, **77**, 257-286. https://doi.org/10.1109/5.18626

[14] Levinson, S.E., Rabiner, L.R. and Sondhi, M.M. (1983) An Introduction to the Application of the Theory of Probabilistic Functions of Markov Process to Automatic Speech Recognition. *The Bell System Technical Journal*, **62**, 1035-1074. https://doi.org/10.1002/j.1538-7305.1983.tb03114.x

[15] Forchhammer, S. and Rissanen, J. (1996) Partially Hidden Markov Models. *IEEE Transactions on Information Theory*, **42**, 1253-1256. https://doi.org/10.1109/18.508852

[16] Therrien, C.W. (1992) Random Processes. Prentice Hall Signal Processing Series. Prentice-Hall, Englewood Cliffs.

[17] RADARSAT-2 Product Description, MDA, RN-SP-52-1238 Issue 1/14 (2018).