Land-Cover Classification Using Multitemporal ERS-1/2 InSAR Data

Marcus E. Engdahl and Juha M. Hyyppä, *Member, IEEE*

***Abstract*—In this study the potential of ERS-1/2 Tandem InSAR data for land-cover classification was investigated at a 2500 km^2 study area around the Helsinki metropolitan area in Southern Finland. A time-series of 14 ERS-1/2 SAR Tandem image pairs was processed into 28 five-look intensity images, 14 Tandem coherence images and two coherence images with a longer temporal baseline (36 and 246 days). All image data was coregistered and orthorectified into map coordinates using an InSAR DEM. A two-stage hybrid classifier method was employed, where the water-class was classified separately in the first classifier stage, and the remaining classes were classified with an ISODATA classifier. Temporal averaging and Principal Components Transformation (PCT) were used to reduce the number of images fed into ISODATA. Classification accuracy was assessed using high-resolution aerial orthophotos, digital base maps and the Finnish National Forest Inventory (NFI). The overall accuracy for six classes (Field/Open Land, Dense Forest, Sparse Forest, Mixed Urban, Dense Urban, Water) was found to be 90% with kappa coefficient of 0.86. Interferometric coherence carries more land-cover related information than the backscattered intensity. This study confirms that the ERS-1/2 Tandem archives could be exploited for land-cover classification.**

***Index Terms*—Interferometric coherence, land-cover classification, principal components transformation, synthetic aperture radar (SAR), synthetic aperture radar (SAR) interferometry, temporal filtering.**

I. Introduction

LAND-COVER classification is one of the most important applications of remote sensing. The usability of optical satellite data in land-cover classification is severely limited by cloud cover in many parts of the world. Synthetic aperture radar (SAR) can penetrate cloud cover, but the potential of C-band single polarization intensity images is limited. Interferometric SAR (InSAR) can provide complementary information to the backscattered intensity in the form of interferometric coherence. The use of multitemporal InSAR data can increase the number of reliably distinguishable land-cover classes.

Previous studies of the potential of using multitemporal ERS-1/2 Tandem InSAR data in land-cover classification have achieved promising results. Dammert *et al.* achieved an overall classification accuracy between 65% to 75% into five classes (water, forest 1, forest 2, other+farmland, urban) and 93% for a water/land—classification using multiple Tandem pairs and an unsupervised segmentation scheme [1]. Strozzi *et al.* used several classification methods on multiple Tandem pairs and achieved classification accuracies in the order of 75% into four classes (forest, field, urban, water) and 80% to 85% for a forest/nonforest—classification [2]. Both studies report poor classification accuracy for the urban class. Wegmüller and Werner used one ERS-1 interferometric pair with a three-day repeat period and reported a 91% overall accuracy for a forest/nonforest—classification on relatively flat terrain [3]. Terrain relief has been found to reduce classification accuracy [2], [3].

In this study the potential of multitemporal C-band InSAR data from the ERS-1/2 Tandem mission in land-cover classification was investigated. A two-stage hybrid classifier method where the water-class was classified separately in the first classifier stage was employed, and the classification result was compared with high-quality reference data that was specifically gathered for this study. According to Congalton and Green existing reference data is rarely acceptable for accuracy assessment due to unknown errors in the reference data, land-cover changes, and differences in employed classification systems, and that this problem has been one of the main reasons for the lack of acceptance of digital satellite data for many applications [4]. We believe that previous studies have underestimated the potential of multitemporal ERS Tandem InSAR data for land-cover classification.

II. Study Area

The Helsinki metropolitan area in southern Finland (near 60°N, 25°E) was chosen as the test area for the land-cover classification study (see Fig. 1). Approximately one million people live inside the 2500 km^2 study area, which encompasses diverse land-cover from lakes and boreal forests in the Nuuksio National park to dense urban settlement in the Helsinki city center. The study area is divided into forest, open/agricultural land, water and built-up areas with the following percentages: 35%, 18%, 30%, and 17%, respectively. The forests in the area are mainly southern boreal forests dominated by Scots pine and Norwegian spruce. Birch is the most common deciduous species, but pure deciduous forest stands are rare. The mean elevation of the land area is 47 m above sea level, whereas the maximum elevation is 155 m, the standard deviation of the elevation being 28 m. The variation is mainly caused by the systematic terrain elevation raise from sea shore to inland areas, and also locally by small gently sloping hills.

Manuscript received March 28, 2002; revised February 4, 2003. This work was supported by the Academy of Finland.

M. E. Engdahl is with the Laboratory of Space Technology, Helsinki University of Technology, FIN-02015 HUT, Espoo, Finland (e-mail: mengdahl@avasun.hut.fi).

J. M. Hyyppä is with the Finnish Geodetic Institute, FIN-02431 Masala, Finland (e-mail: juha.hyyppa@fgi.fi).

Digital Object Identifier 10.1109/TGRS.2003.813271

Fig. 1. Location of the study area.

III. Data

A. SAR Data

The SAR data used in this study consists of 14 ERS-1/2 Tandem image pairs (28 images total) acquired with the C-band SAR on-board the European Space Agency's ERS-1 and ERS-2 satellites. The imagery was acquired during the ERS-1/2 Tandem mission in 1995–1996, when the two satellites were flown in the same orbital plane so that ERS-2 imaged the same area on the ground 24 h after ERS-1. Interferometric baselines were deliberately kept short during the Tandem mission, 98% of the perpendicular components of the Tandem baselines are shorter than 300 m. Due to the relatively short temporal baseline of 24 h and the short interferometric baselines, the image pairs collected during the Tandem mission are very useful in interferometric studies of natural targets, which decorrelate quite rapidly in C-band. The image data used in this study span a whole year from the summer of 1995 to the summer of 1996. Table I lists the ERS-1/2 Tandem pairs used in the study, all the images are acquired from descending orbits. The data are from two track/frame geometries: eight pairs from track 408 and six pairs from track 179. Precise orbit information (the so-called PRC orbits) was acquired from DLR (Deutches Zentrum für Luft- and Raumfart).

TABLE I
List of ERS-1/2 Tandem Interferometric Pairs. Interferometric Baseline Refers Perpendicular Component of the Baseline

Pair	Sensor	Orbit	Track	Frame	Date	Int. Baseline (m)
1	ERS-1	20937	408	2385	17.7.1995	-2
	ERS-2	01264			18.7.1995	
2	ERS-1	21438	408	2385	21.8.1995	-72
	ERS-2	01765			22.8.1995	
3	ERS-1	21710	179	2385	9.9.1995	-28
	ERS-2	02037			10.9.1995	
4	ERS-1	21939	408	2385	25.9.1995	239
	ERS-2	02266			26.9.1995	
5	ERS-1	22211	179	2385	14.10.1995	-221
	ERS-2	02538			15.10.1995	
6	ERS-1	22440	408	2385	30.10.1995	-49
	ERS-2	02767			31.10.1995	
7	ERS-1	23442	408	2385	8.1.1996	-29
	ERS-2	03769			9.1.1996	
8	ERS-1	23943	408	2385	12.2.1996	85
	ERS-2	04270			13.2.1996	
9	ERS-1	24215	179	2385	2.3.1996	-76
	ERS-2	04542			3.3.1996	
10	ERS-1	24444	408	2385	18.3.1996	80
	ERS-2	04771			19.3.1996	
11	ERS-1	24716	179	2385	6.4.1996	37
	ERS-2	05043			7.4.1996	
12	ERS-1	24945	408	2385	22.4.1996	-58
	ERS-2	05272			23.4.1996	
13	ERS-1	25718	179	2385	15.6.1996	48
	ERS-2	06045			16.6.1996	
14	ERS-1	26219	179	2385	20.7.1996	188
	ERS-2	06546			21.7.1996	

B. Reference Data

Reference data was used to collect reference points for classification accuracy assessment. At least 100 reference points were collected for each of the eight reference classes (see Table III). One-meter resolution aerial color orthophotos and oblique aerial images over the town of Vantaa (238 km^2) were used as reference for the agricultural and urban classes. Reference points for the two forest classes were gathered using the operationally used National Forest Inventory (NFI) of Finland [5]. National base maps in 1 : 20 000 scale were used in addition to the aerial imagery for the water and dense urban classes. The reference points were selected independently from radar data. All reference points were hand picked by using a method where the operator identified an area clearly representing a reference class and a pixel in the middle of such an area was chosen to be the reference point. This method ensured that each of the reference points corresponds to a 2-D reference plot that is homogenous according to our classification system (see Table III), and that here are no mixed pixels amongst the reference data. The quality of the reference data for agricultural and urban classes is very high, since with high-resolution orthophotos it is possible to determine the reference class with a very high accuracy. If even better accuracy were desired it would have been necessary to visit the reference points on foot. The NFI that was used as reference for the forest classes is a multisource method for large-scale forest inventory that utilizes Landsat TM and sample plots. NFI gives good results on large areas, but it can be unreliable at small forest stand sizes [6]. We were able to increase the reliability of the NFI estimates by classifying the stem volume data into two classes (sparse and dense forest) and using the classified NFI in reference point selection.

IV. Interferometric Processing and Geocoding

A. Interferometric Processing

SAR data were interferometrically processed and geocoded using a commercial software package by Gamma Remote Sensing Research and Consulting AG [7]. In order to reduce the demands on disk space and processing power a 50×50 km subset of the 100×100 km full ERS SAR-scenes was selected

TABLE II
LIST OF ERS-1/2 INTERFEROMETRIC PAIRS WITH LONG TEMPORAL BASELINES. INTERFEROMETRIC BASELINE REFERS TO THE PERPENDICULAR COMPONENT OF THE BASELINE

Pair	Sensor	Orbit	Track	Frame	Date	Int. Baseline (m)	Temporal Baseline (days)
L1	ERS-1	21438	408	2385	21.8.1995	0	36
	ERS-2	02266			26.9.1996		
L2	ERS-1	21438	408	2385	21.8.1995	-269	246
	ERS-2	05272			23.4.1996		

TABLE III
REFERENCE CLASSES

Class No.	*Class Description*
1	Water bodies
2	Agricultural fields, open land
3	Dense forest (stem volume > 100m^3/ha)
4	Sparse forest (stem volume 50-100m^3/ha)
5	Low residential area (single- and double family housing)
6	High residential area (residential flats, 3-7 floors)
7	Industrial buildings (halls and warehouses)
8	Dense urban (densely built multistory buildings, 5+ floors)

for interferometric processing. Since the data were from two different imaging geometries depending on the track/frame combination, the images were processed in two separate sets. First, a master image was chosen for both geometries, and all images were coregistered to their respective master image with subpixel accuracy necessary for image correlations. The ERS-1 images from orbits 20 937 and 21 438 were chosen as the master images for data from tracks 408 and 179, respectively (see Table I). Tandem interferograms were generated from each Tandem pair by cross-correlating the already coregistered images. Common band filtering (also called spectral-shift filtering) was applied before interferogram generation in order to minimize the effects of the baseline geometry on coherence estimation [8]. At this stage multilooking (five azimuth looks) was performed in order to improve on the estimates of the interferometric phase and coherence. After multilooking the pixel size of the image data is approximately 20 m in both azimuth and slant range at the ERS nominal look angle of 23°. The resulting five-look interferograms were flattened using high-quality orbit information distributed by DLR. The interferometric coherence was estimated using square estimator windows with three different estimator window sizes (5×5, 7×7, and 9×9 pixels) using Gaussian weighing of the samples. Only coherence images produced with the smallest 5×5 pixel estimator window were later used in the classification study. In addition to the Tandem coherence images with a temporal baseline of 24 h, two coherence images with longer temporal baselines were formed (see Table II). The longtime coherence images are used to detect anthropogenic features that can remain interferometrically stable for months or years [9], [10]. These so-called "permanent scatterers" can be utilized to monitor ground motion with millimetric accuracy [11], [12]. Also, five-look intensity images were generated and radiometrically calibrated for range spreading loss, antenna gain, normalized reference area and the calibration constant that depends on the radar parameters and the SAR processor used by the Processing and Archiving Facility (PAF) [13].

B. Geocoding

After interferometric processing all the processed images were in the imaging geometries of the two master images. In order to create a unified dataset all image data had to be orthorectified into map coordinates. This was accomplished by creating a simulated SAR image from an InSAR digital elevation model (DEM), and using the simulated SAR image to coregister the two image sets. The simulated SAR image was created using a DEM produced by weighted averaging of two InSAR DEMs produced from the Tandem pairs 8 and 10 (see Table I). These wintertime pairs were chosen due to their high overall coherence and usable baseline lengths. Averaging InSAR DEMs reduces the effects of phase noise and possible atmospheric artifacts, especially when several InSAR DEMs are available. Image data from track 179 was then rectified into the geometry of track 408 using the simulated SAR image. Judging visually, the image data from the two different tracks coincide perfectly, which implies that subpixel accuracy was achieved. Now that all image data was in the geometry of the master image for track 408, the InSAR DEM was again used to orthorectify the images into UTM projection. The final coregistration to map coordinates was done using automatically picked ground control points (GCPs). The mean RMS registration error of the GCPs is approximately one pixel, or 20 m. This high-accuracy coregistration allowed us to use the methods of temporal filtering and temporal averaging over the whole set of images. The intensity images were corrected for the true pixel size using the DEM, which reduces the effect of terrain slope on backscattered intensity. The pixel size of the final image products is 20 m. The size of the orthorectified multitemporal InSAR dataset (28 Intensity images, 14 Tandem coherence images, two longtime coherence images) is approximately 2.5 Gb.

V. TEMPORAL FILTERING

Reliable estimates of the backscattered intensity σ^0 from a distributed target require that the estimated number of looks (ENL) is sufficiently large. Spatial filtering—for example speckle filtering—is often used to increase the ENL with loss of spatial resolution. In properly coregistered multitemporal datasets it is possible to employ the technique of temporal filtering, which in principle increases radiometric resolution without degrading spatial resolution. The theory of temporal filtering is discussed in [14], here we used a simple temporal filter, which is described in [15] and [16]. The simplified filter is defined as

$$J_k(x,y) = \frac{\langle I_k \rangle}{N} \sum_{i=1}^{N} \frac{I_i(x,y)}{\langle I_i \rangle} \tag{1}$$

where (x, y) are the spatial coordinates in an image, J is the set of temporally filtered images, I the set of unfiltered images and $\langle I \rangle$ is a spatial estimate for I. Since a spatial estimate for I is needed in (1), some loss of spatial resolution is unavoidable.

We temporally filtered both the 28 intensity images and the 14 Tandem coherence images.

A. Intensity Images

The 14 ERS-1 and 14 ERS-2 intensity images (see Table I) were filtered separately with the temporal filter. The spatial estimate needed in (1) was acquired with a 5×5 pixel k-nearest-neighbor Lee-filter (see [17]). According to visual inspection, the temporally filtered images display markedly diminished speckle with little or no reduction in spatial resolution. In fact, narrow linear structures are more clearly visible in the filtered images, than in the unfiltered ones.

B. Coherence Images

The 14 ERS-1/2 Tandem coherence images were filtered with the temporal filter. A 5×5 pixel median filter was used for acquiring the spatial estimate. Judging visually, the temporally filtered images are much less noisy with some degradation of spatial resolution. Narrow linear structures are well preserved, as in the intensity image case.

VI. Principal Components Transformation

Due to the large number of SAR intensity and coherence images, it was necessary to reduce the dimensions of the image datasets prior to classification. We used the well-known principal components transformation (PCT) for this purpose. PCT transforms a set of correlated images into a new set of uncorrelated images. The input dataset is transformed along orthogonal axes, whose directions are defined by the variability in the input dataset. The first principal component (PC) is along the axis of maximum variance in the input dataset, the second PC is along the axis of maximum variation orthogonal to the first PC, and so forth. The PCT can be viewed as a rotation of the axes in the input dataset to new and orthogonal orientations. If all axes are included in the transformation, no information is lost. PCT shifts the variability of the dataset toward the first PCs, so it is possible to use PCT as dimension reduction method by retaining only the first PCs that explain most of the variation in the input dataset.

A. Substracting the Temporal Average Image

If the PCT is applied to the intensity or coherence time-series without preprocessing, the first PC is in practice equal to the mean of the time-series. Similar behavior was also reported by Dammert *et al.* [1]. The PCT is not an efficient method for computing the mean of a time-series and since the direction of the first PC affects the directions of all the PCs due to the orthogonality constraint, it is preferable to subtract the mean of the time-series before applying the PCT. This preprocessing step reduces correlations between the images and enhances changes from the average value over the time-series. Temporal averaging refers to the simple procedure of computing the average image of a set of coregistered images. It produces images that have high resolution and low noise [2], [18], [19], which are, therefore, well suited for visual inspection or simple initial classification. The 28 intensity images were averaged pairwise (see Table I) to produce 14 Tandem mean intensity images. Both the Tandem coherence and Tandem mean intensity time-series were centered by subtracting the temporal average images from them before applying the PCT.

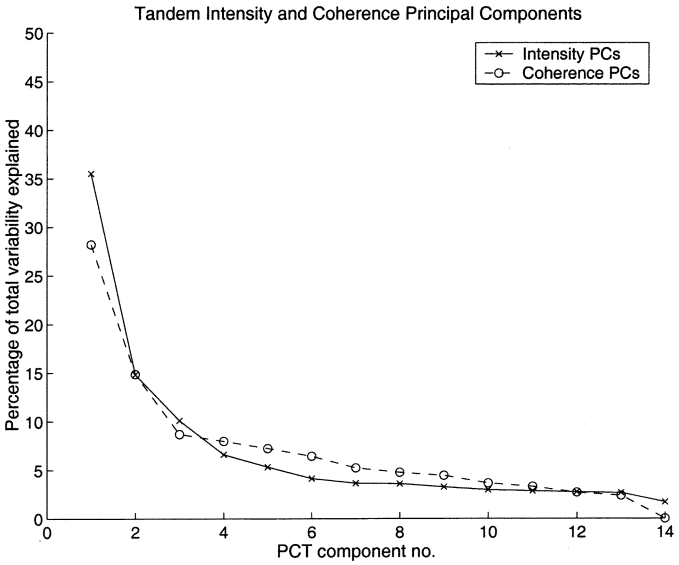

Fig. 2. Percentage of the total variability in the centered and water-masked Tandem intensity and coherence datasets as explained by the PCT components.

B. Water-Masking

If water-areas are included in the centered intensity dataset, the first PC will highlight the water-bodies. This happens because differing wind conditions can make large differences in backscattered intensity from water. Similarly, the first PC of a centered coherence time-series will highlight the water-bodies, provided that winter images where the water surface is covered with ice are included in the time-series. This is because free water has extremely low coherence and ice has often quite high coherence. Masking out water areas from the centered datasets before applying the PCT enables the first PC to highlight areas on land that exhibit greatest variation over the time-series, because the stronger signals that would have been created over water-bodies are not included in the PCT. The water-mask was made by thresholding the mean of the backscatter intensity ratios of Tandem pairs 3 and 6 (see Table I). On these pairs the intensity ratio shows water bodies very clearly because one image in the pair is acquired at calm, and the other at windy conditions. If images with very windy and calm conditions are not available, this method for generating the water-mask cannot be used. The water-masking step can be viewed as the first stage of a hybrid classification system where water is first classified separately and the land-classes are treated in the second stage.

C. Principal Components

After the temporal average was subtracted from the intensity and Tandem coherence time-series and water areas were masked out, the PCT was applied separately to the intensity and Tandem coherence time-series. The eigenvalue plots for the PCs (see Fig. 2) show how much of the total variability in the centered and water-masked time-series is explained by the individual principal components. There is a "knee" in the eigenvalue plots at the second PC for both intensity and coherence PCs—the first

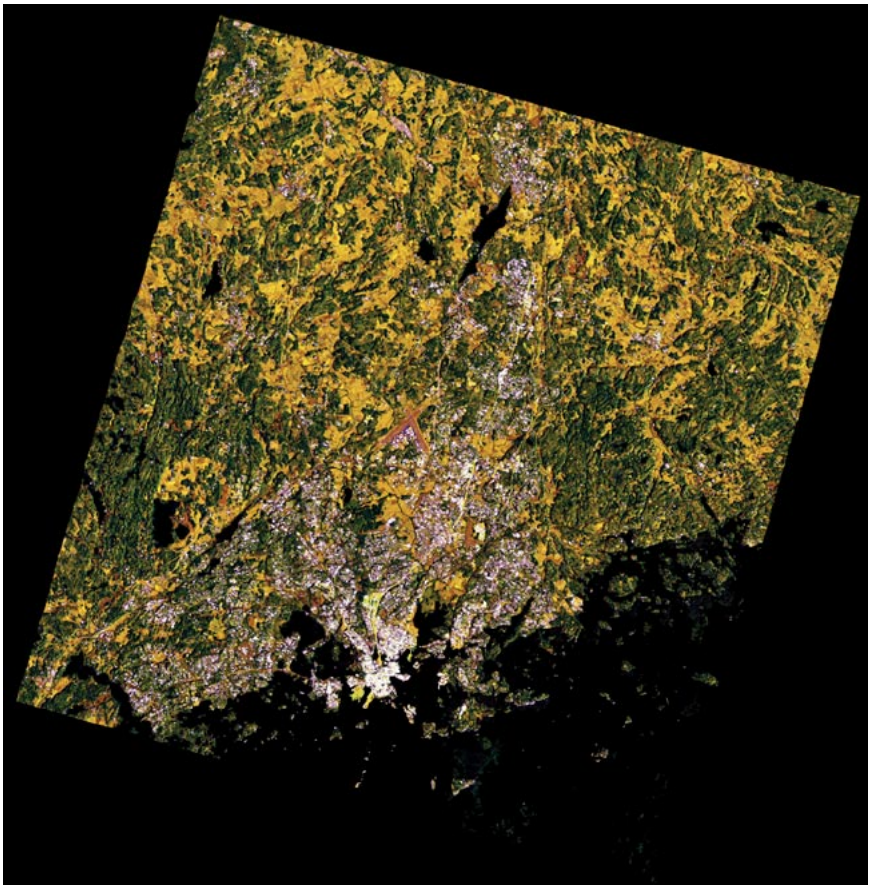

Fig. 3. False-color composite view over the Helsinki metropolitan area. Water areas are masked out and are shown in black. Channels: Red—First PC of the temporal average images for intensity and coherence; Green—Second PC of the temporal average images for intensity and coherence; Blue—Average longtime coherence image. Data obtained through ESA Announcement of Opportunity studies AOT-SF.301 and A03.277.

two PCs together explain 50% of the total variation in the intensity time-series and 43% in the coherence time-series.

VII. Land-Cover Classification

In the two-stage hybrid classification method employed in this study the water-class was classified separately in the first classification stage, and the land-classes were classified in the second stage using an ISODATA classifier.

A. Feature Selection for ISODATA

The use of temporal averaging and principal component analysis allowed us to reduce the dimensionality of the data from the initial number of 44 images (28 intensity images, $14+2$ coherence images). The following six images were chosen as input images to the ISODATA classifier:

1) temporal average of the backscattered intensity time-series;
2) first PC of the backscattered intensity time-series;
3) temporal average of the Tandem coherence time-series;
4) first PC of Tandem coherence time-series;
5) second PC of Tandem coherence time-series;
6) average of two coherence images with long temporal baselines (36 and 246 days).

Thus, the input dataset included two channels of information from intensity data, three channels of information from Tandem coherence data and one information channel from longtime coherence data. The second PC of the intensity time-series was not included because it was quite noisy and did not appear to contain information relevant to land-cover classification. That

TABLE IV
Producer's Accuracies. Average Producer's Accuracy 90%, Overall Accuracy 90%, Kappa Coefficient 0.86

		Reference Class					
		Field / Open Land	Dense Forest	Sparse Forest	Mixed Urban	Dense Urban	Water
Assigned Class	Field / Open Land	**97**	2	1	1	0	0
	Dense Forest	0	**88**	8	4	0	0
	Sparse Forest	1	11	**86**	4	0	0
	Mixed Urban	2	0	5	**80**	9	0
	Dense Urban	0	0	0	12	**91**	0
	Water	0	0	0	0	0	**100**

TABLE V
User's Accuracies. Average User's Accuracy 83%, Overall Accuracy 90%, Kappa Coefficient 0.86

		Reference Class					
		Field / Open Land	Dense Forest	Sparse Forest	Mixed Urban	Dense Urban	Water
Assigned Class	Field / Open Land	**99**	0	0	1	0	0
	Dense Forest	0	**80**	5	15	0	0
	Sparse Forest	4	12	**68**	16	0	0
	Mixed Urban	4	0	1	**93**	2	0
	Dense Urban	0	0	0	38	**62**	0
	Water	0	0	0	1	0	**99**

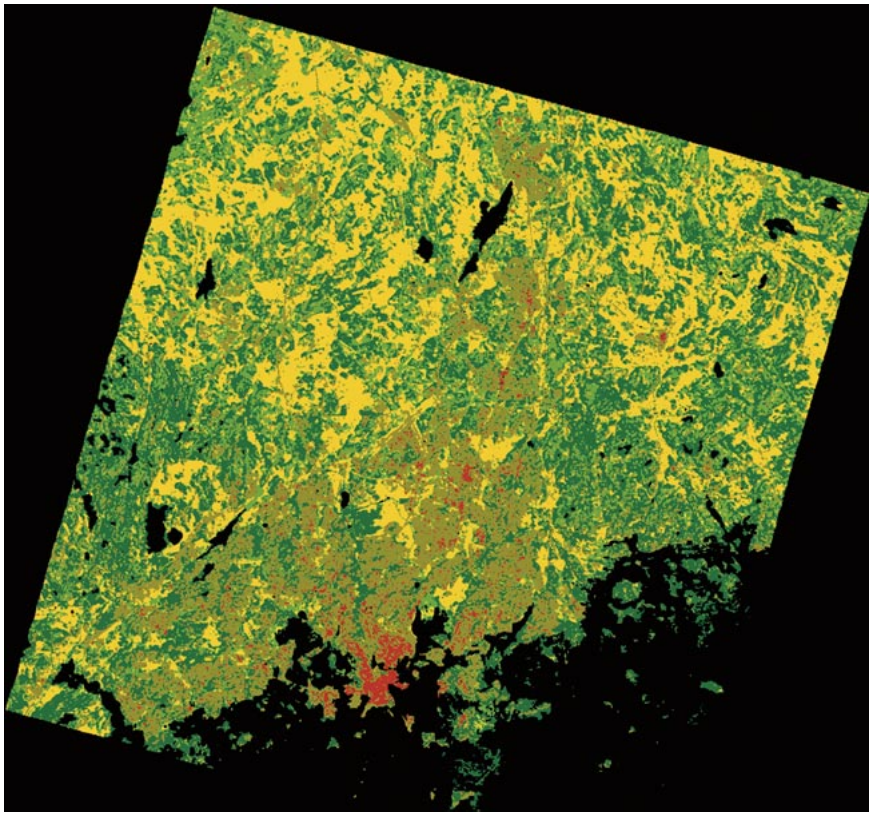

Fig. 4. Classification result over the Helsinki metropolitan area. Classes: Black—Water; Yellow—Field/Open Land; Dark Green—Dense Forest; Light Green—Sparse Forest; Bronze—Mixed Urban; Red—Dense Urban.

was also the case for the higher-order PCs of both the intensity and Tandem-coherence time-series. All images were scaled to values between 0–255 (eight bits). Information from three channels of the input dataset is presented as a false-color image in Fig. 3.

B. ISODATA Classification

The ISODATA algorithm is an unsupervised technique that looks for "natural" clusters in the data [20]. In our case the number of clusters created by ISODATA could be controlled with the "maximum number of clusters"—setting. The number of clusters was varied systematically and the number that produced the best differentiation between the reference classes (see Table III) was chosen. Best results were achieved with 14 clusters. The output of the ISODATA classifier was then filtered with a 3×3 pixel majority filter. The ISODATA classification could not properly differentiate between the "High Residential," "Low Residential," and "Industrial Building"—classes as defined in the reference data (see Table III). This was not surprising since previous studies have reported poor accuracy in classifying urban classes using ERS InSAR data [1], [2]. The three mentioned urban classes were combined into a "Mixed Urban" class, which covers both residential and industrial areas. Nine of the 14 ISODATA clusters corresponded to the "Mixed Urban"—class and two to "Field/Open Land"—class. The "Dense Forest," "Sparse Forest," and "Dense Urban" classes all corresponded to a single ISOCLASS cluster.

VIII. Results and Discussion

The accuracy of the land-cover classification was assessed by comparing the classification result with the known land cover classes (ground truth) at the reference points. Table IV lists the achieved producer's accuracies, and Table V the user's accuracies for the six classes. The producer's accuracy tells how many percent of pixels belonging to a reference class are correctly assigned into that class on the classification map, while the user's accuracy tells how many percent of pixels assigned to a class on the classification map actually belong to that class according to the reference data. The average producer's accuracy is 90%, average user's accuracy 83%, overall accuracy 90% and the Kappa coefficient is 0.86. If the accuracies are weighed by the class areas, the average producer's accuracy rises to 92% and the average user's accuracy to 89%. Producer's accuracies for all the classes are at least 80%. The lowest user's accuracies are 62% for the "Dense Urban' class and 68% for the "Sparse Forest" class. These results are substantially better than previously reported for interferometric ERS data, but one has to keep in mind that different accuracy assessment methods were used in the previous studies [1], [2]. Fig. 4 illustrates the classification for the Helsinki metropolitan area.

A. Class Differentiation

The distributions of the pixel-values for the different classes in the classified dataset are illustrated in Fig. 5. The "Field/Open

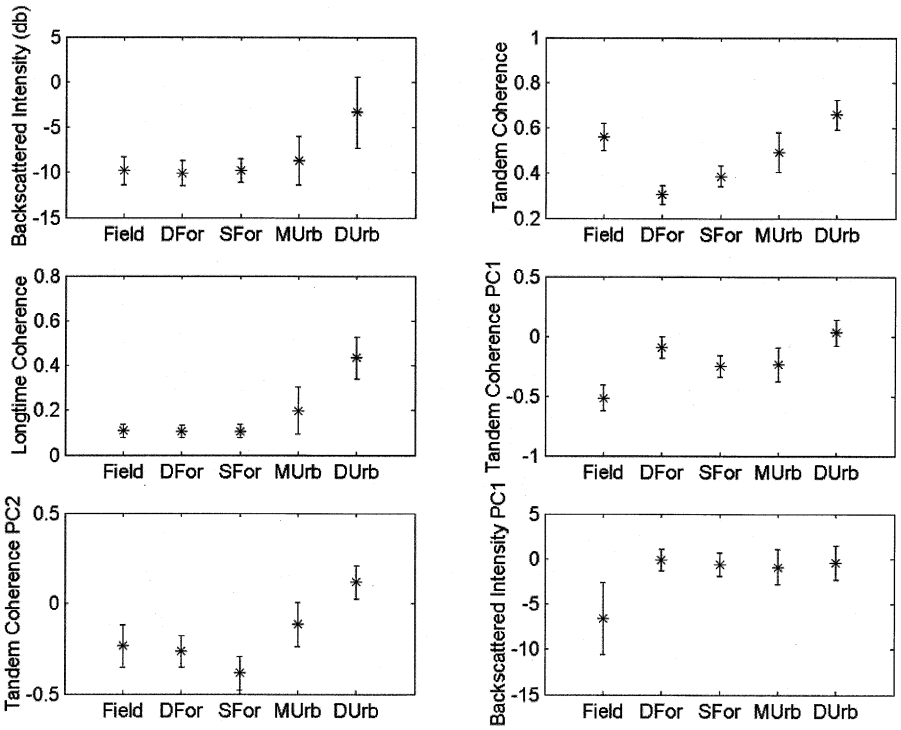

Fig. 5. Distributions of the pixel values for different classes in the dataset fed to ISODATA. ± 1 standard deviation indicated with an error bar.

Land"—class has on average higher Tandem coherence than the forest-classes, as expected. The class is a combination of two clusters found by ISODATA, which shows in the large standard deviation in the intensity PC1. These two kinds of fields exhibit different behavior over the intensity time-series, probably due to differing farming practices, but our ground truth data did not allow us to tell what the difference between these two field types is. The most probable explanation for the difference is that some of the fields were ploughed during autumn and others were not. The "Dense Forest"—class exhibits lower Tandem coherence than the "Sparse Forest"—class, which is to be expected because in dense forests there are more volume scattering effects and a greater fraction of the radar energy is backscattered from the tree branches, which are easily disturbed by wind. The two forest classes also differ from each other in the coherence PC1 and PC2 due to their different behavior over the Tandem coherence time-series.

Urban land-cover classes are problematic in SAR-based classification. The visibility of man-made structures in SAR intensity images depends on their orientation with respect to the satellite line-of-sight. Urban classes are characterized by high backscatter from man-made targets but due to the orientation problem even large structures may stay undetected. This problem could be mitigated by viewing the urban area from different angles, i.e., using images from both ascending and descending orbits. Unfortunately, all our images were acquired from a descending orbit. The "Dense Urban"—class has a high level of backscatter due to reflections from man-made structures, and it also stays coherent for long periods of time, as shown by the longtime coherence. The "Mixed Urban"—class is a combination of nine ISODATA clusters, which shows in the quite large standard deviation of the pixel values in Fig. 5. It is also a very diverse class on the ground, since it encompasses industrial warehouses, multistory residential buildings and low-density residential areas with trees, lawns, parking lots and small parks between the buildings. This is also evident in Table V, which shows that there is confusion between the "Mixed Urban"—class and the "Dense Urban"—class, as well as the forest classes. Higher resolution, if available, would reduce this confusion and make classification of the urban areas more feasible. The values for the "Mixed Urban" and "Dense Urban"—classes differ markedly in all coherence-based information channels. This implies that both longtime coherence and a time-series of Tandem coherence images are useful in urban area classification. Inspection of Fig. 5 confirms that coherence-based channels carry more land-cover related information than intensity-based channels. The use of coherence information affected the classification result greatly by improving classification accuracy and enabling us to find two forest classes and two urban classes instead of just one of each.

B. Classification Quality

The achieved classification accuracies are very satisfactory—good classification accuracies were achieved with the basic ISODATA classifier and it is probable that even better results could be achieved by using an advanced classification or segmentation method. ISODATA was selected to demonstrate the applicability of the processed dataset in land-cover classification. Six classes could be distinguished reliably from each other. The Field/Open Land—class is a combination of two clearly separate ISODATA clusters. Therefore, it is probable that the number of reliably distinguishable classes would

actually be seven (two Field/Open Land—classes), but lack of ground truth prevented us from confirming this. In the previous studies classification of urban areas has been problematic and the reported accuracies have been low. In this study it turned out to be possible to detect two urban classes with acceptable accuracies. The boundary between the "Dense Urban" and "Mixed Urban" classes is somewhat arbitrary, but it seems that densely built areas of multistory buildings are quite reliably being classified into the "Dense Urban" class. Also, a useful feature of the urban classification is that main roads have been classified into the "Mixed Urban" class. There is some confusion between the "Sparse Forest" and "Dense Forest"—classes. Part of this confusion is explained by the use of Finnish National Forest Inventory (NFI) as the ground truth—the NFI is not very reliable at the stand-level, especially on small stands [6]. In [6] ERS-1/2 Tandem coherence data proved more accurate than NFI in stand-level stem volume estimation.

C. Comparison With Previous Studies

There are several reasons which might explain why the classification accuracies reported in this study are higher than in the previous studies utilizing ERS Tandem InSAR data [1], [2]. One of the reasons is the very high quality reference data used in this study, made possible by the use of up-to-date high-resolution aerial orthophotos. Dammert *et al.* state that the overall classification accuracy in their study (65% to 75% for Gothenburg study area and 60% to 65% for Hökmark area) should be seen as lower limits due to land-map problems [1]. The modest terrain relief in Helsinki area combined with an InSAR DEM made high accuracy georeferencing possible. Dammert *et al.* also state that inaccurate georeferencing has generated registration errors in steep slope areas of up to 15 pixels in the Gothenburg area and "less than ten pixels" for the Hökmark area. The amount of data and how it is processed also affect the classification accuracy. In this study we used 14 Tandem pairs covering a full year, and long-time coherence information was used in addition to the Tandem coherence images. Strozzi *et al.* used six Tandem pairs at the Tuusula test site—which is inside our study area—and utilized both Tandem and long-time coherence information, as well as textural information [2]. Dammert *et al.* computed three coherence images from Tandem pairs at the Gothenburg test site and 18 coherence images with temporal baselines from three days and upwards at the Hökmark test site. They also used PCT and adaptive feature extraction to reduce the dimensionality before classification. Strozzi *et al.* used temporal averaging of coherence and intensity images for dimension reduction. In this study the sensitivity of the PCT in the differentiation of different land-cover types was enhanced by the use of a hybrid classification method where water was masked out before applying the PCT.

IX. Conclusion

The potential of multitemporal ERS-1/2 Tandem InSAR in land-cover classification was investigated on a test area in Southern Finland. Backscattered intensity, Tandem coherence and long-time coherence information were utilized in the classification. A two-stage hybrid classifying method was employed, where the water-class was classified separately in the first classifier stage, and the remaining classes were classified with an ISODATA classifier. Classification accuracy was assessed using high-resolution aerial orthophotos, digital base maps and the Finnish National Forest Inventory (NFI). Overall classification accuracy of 90% into six classes ("Field/Open Land," "Dense Forest," "Sparse Forest," "Mixed Urban," "Dense Urban," "Water") with a kappa coefficient of 0.86 was achieved. The number of distinguishable land-cover classes was probably seven with two "Field/Open Land"—classes, but lack of adequate ground truth prevented us from confirming this. These results are substantially better than what has previously been achieved with ERS-1/2 Tandem InSAR data. Probable reasons for the better accuracy in this study are; the high quality reference data, high accuracy georeferencing, the large number of Tandem pairs used, the two-stage classifying method, and the preprocessing applied before using the principal components transformation (PCT). The authors recommend using a two-stage classifier where water areas are classified separately in the first stage, and the land-classes at the second stage by classifying the water-masked dataset. It is also recommended that the temporal average image should be subtracted from the time-series before applying the PCT. Interferometric coherence was found to carry more land-cover related information than the backscattered intensity. This study confirms that the ERS-1/2 Tandem archives could be exploited for land-cover classification.

Acknowledgment

The authors wish to thank ESA for providing the SAR data through ESA Announcement of Opportunity studies AOT-SF.301 and A03-277 as well as the town of Vantaa for the high-resolution aerial color orthophotos.

References

[1] P. B. G. Dammert, J. I. H. Askne, and S. Kühlmann, "Unsupervised segmentation of multitemporal interferometric SAR images," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2259–2271, Sept. 1999.

[2] T. Strozzi, P. B. G. Dammert, U. Wegmüller, J.-M. Martinez, and J. I. H. Askne, "Landuse mapping with ERS SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 766–775, Mar. 2000.

[3] U. Wegmüller and C. Werner, "Retrieval of vegetation parameters with SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 18–24, Jan. 1997.

[4] R. G. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practises*. Boca Raton, FL: Lewis, 1998.

[5] M. Katila and E. Tomppo, "Selecting estimation parameters for the Finnish multisource national forest inventory," *Remote Sens. Environ.*, vol. 76, pp. 16–32, 2001.

[6] H. J. Hyyppä and J. H. Hyyppä, "Effects of stand size on the accuracy of remote sensing—Based forest inventory," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 2613–2621, Dec. 2001.

[7] U. Wegmuller, C. Werner, and T. Strozzi, "SAR interferometric and differential interferometric processing chain," in *Proc. IGARSS*, vol. 2, Seattle, WA, July 6–10, 1998, pp. 1106–1108.

[8] F. Gatelli, A. M. Guarnieri, F. Parizzi, P. Pasquali, C. Prati, and F. Rocca, "The wavenumber shift in SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, July 1994.

[9] S. Usai and R. Klees, "SAR interferometry on a very long time scale: A study of the interferometric characteristics of man-made features," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2118–2123, July 1999.

[10] D. J. Weydahl, "Analysis of ERS SAR coherence images acquired over vegetated areas and urban features," *Int. J. Remote Sens.*, vol. 22, no. 14, pp. 2811–2830, 2001.

[11] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 8–20, Jan. 2001.
[12] ——, "Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, pt. 1, vol. 38, pp. 2202–2212, Sept. 2000.
[13] H. Laur, P. Bally, P. Meadows, J. I. Sanchez, B. Schaettler, and E. Lopinto, Derivation of the backscattering coefficient σ^0 in ESA ERS SAR PRI products, ESA-ESRIN, no. 2, May 1997. .
[14] A. Lopes, E. Nezry, R. Touzi, and H. Laur, "Structure detection and statistical adaptive filtering in SAR images," *Int. J. Remote Sens.*, vol. 14, no. 9, pp. 1735–1758, 1993.
[15] S. Quegan, T. Le Toan, J. J. Yu, F. Ribbes, and N. Floury, "Multitemporal ERS SAR analysis applied to forest mapping," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 741–753, Mar. 2000.
[16] S. Quegan and J. J. Yu, "Filtering of multichannel SAR images," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 2373–2379, Nov. 2001.
[17] J. S. Lee, "Speckle analysis and smoothing of synthetic aperture radar images," *Comput. Graph. Image Process.*, vol. 17, pp. 24–32, 1981.
[18] P. J. Meadows and H. Laur, "Exploitation of ERS SAR imagery for land applications," in *Proc. 2nd Int. Workshop on the Retrieval of Bio- and Geo-Physical Parameters from SAR Data for Land Applications*, Noordwijk, The Netherlands, Oct. 21–23, 1998, pp. 43–50.
[19] M. E. Engdahl and J. Hyyppä, "Temporal averaging of multitemporal ERS-1/2 tandem INSAR data," in *Proc. IGARSS*, July 20–24, 2000, p. 3.
[20] J. T. Tou and R. C. Gonzales, *Pattern Recognition Principles*. Reading, MA: Addison-Wesley, 1974.

Marcus E. Engdahl was born in Finland. He received the M.S. degree in technical physics from the Helsinki University of Technology (HUT), Espoo, Finland, in 1996. He is currently is pursuing the Ph.D. degree at HUT.

From 1995 to 1996, he was a Research Assistant in the Remote Sensing Group, Technical Research Centre of Finland. Since 1996, he has been a Research Scientist in the Laboratory of Space Technology, HUT. From 1997 to 1998, he was a Young Graduate Trainee in the Earth Science Division, the European Space Research and Technology Center (ESA-ESTEC), Noordwijk, The Netherlands. His main research interests are in SAR interferometry and its applications. He is an alumnus of the International Space University, Strasbourg, France.

Juha M. Hyyppä (M'01) was born in 1964. He received the M.S., the licentiate in technology, and the Doctor of Technology degrees from the Helsinki University of Technology (HUT) (with honors), in 1987, 1990, and 1994, respectively.

He is currently the Head of the Department of Remote Sensing and Photogrammetry, Finnish Geodetic Institute, Masala, Finland. He was a Senior Research Fellow (HUT, 1996–1999), part-time R&D Project Manager (Novosat Ltd., 1999–2000), Coordinator of European Commission project HIGH-SCAN (1998–2001), Docent (HUT, Laboratory of Space Technology, 1997–), part-time Managing Director of a small value-adding company DI_Ware Oy (1996–2000), Earth Observation Programme Manager at Technology Development Centre (Tekes, 1994–1995). His personal interests include the understanding of SAR and laser scanning imagery.

Dr. Hyyppä has been a Finnish adviser to ESA Earth Observation Programme Board and to ESA Potentential Participant Meetings (1994–1995), coordinator of the Design Phase of the National Remote Sensing Programme (1995), board member of the National Remote Sensing Programme (1996–2000), and Principal Investigator in ESA Announcement of Opportunity studies.