CLASSIFYING MICROWAVE RADAR IMAGES USING DECISION BASED DATA FUSION

Markus Törmä, Marcus Engdahl

Institute of Photogrammetry and Remote Sensing, Helsinki University of Technology
Directorate of Earth Observation Programmes, ESA-ESRIN

ABSTRACT

Four different data fusion methods for classifying SAR images were tested and compared. Test area was Helsinki region in Southern Finland and test data 14 ERS-1/2 Tandem pairs. The best method was a method where a posteriori probabilities of lower spatial resolution classification are used as a priori probabilities of higher resolution classification. The increase of overall accuracy was 7-14 %-units depending on date and 10.4% on average when compared to original Tandem pairs. Median filtering increased classification accuracy, but not that much when data fusion methods were used. This means that the need of spatial filtering can be at least partially compensated using data fusion of different spatial resolution images.

1. INTRODUCTION

Many governmental institutions have a continuing requirement to form and implement laws and policies that involve existing or future land cover and land use. Remote sensing offers means to provide information about the environment. For example, the purpose of CORINE Programme is to gather information relating to the environment for the European Union. In order to determine and assess the effects of Community’s environment policy, it is needed to have a proper understanding concerning the different features of the environment like the state and geographical distribution of individual environments and natural areas, the quality and abundance of water resources, land cover and soil state, and natural hazards (Heymann et.al., 1994).

CORINE land cover classification is produced using optical satellite images like Landsat ETM. Unfortunately weather conditions limit the use of optical data. For example, here in Finland summer is usually quite cloudy, there are usually only few days when large area of Finland is cloud-free, and during winter there is dark also daytime. Finnish IMAGE2000 satellite image mosaic consisted of 36 Landsat ETM images. The target year was 2000 but only 12 images were taken year 2000 due to cloudiness (Härmä et.al., 2004).

Due to the relative insensitivity to weather, SAR images are an interesting alternative to produce information about land cover and land use. One commercially important application is cellular network planning, where topographic and morphographic data are needed. SAR interferometry can be used to produce the topographic information. SAR intensity images and coherence image produced in interferometric process can be used to provide information about the land cover. Morphographic classes are defined as different land cover types, according to how they interact and attenuate electromagnetic radiation. The most common morphographic categories are urban, suburban, rural, water, open, and forest (Hyyppä et.al., 1999). So, the aim here is to acquire required spatial information using only two SAR images taken in slightly different places.
Traditional and difficult problem when classifying SAR images is speckle, in other words image contains random noise due to the measurement process. Speckle increases intra-class variation, making the statistics of different land cover classes are rather similar and the statistical classification more difficult. Speckle can be decreased by different kind of filters, but developed methods can be difficult to use due to many parameters that can be unknown. Simple alternative is performing average filtering. As the area of averaging increases class histograms start to resemble normal distribution that is good from statistical pattern recognition point of view, but spatial resolution decreases at the same time.

This problem is studied from decision based data fusion point of view in the article. The main principle is that by merging several different classification results made using images with different spatial resolutions, it is possible to acquire better classification than by using any individual classifications. First, ERS SAR-images are averaged using different sizes of filter and then these images are classified. Then different methods are tested to merge these different classifications to form a better one.

Proposed classification method is based on multiresolution approach, where coarse scale classification is made first and it is used to enhance fine scale classification. The difference between coarse and fine scale classifications is that coarse scale classification is based on lower spatial, spectral or radiometric resolution remote sensing data or classification system includes less and more generic classes. Fine scale classification uses better resolution data or classification system includes more and detailed classes. In this study, changing property is spatial resolution, meaning that the coarse scale classification is based on low spatial resolution images and fine scale classifications are based on images with finer spatial resolution.

The aims of this research are twofold. First, several rather simple decision based data fusion methods are introduced and their performance compared. Second, due to large number of SAR images, 14 ERS-1/2 Tandem pairs, the effect of environmental conditions to classification results can be assessed.

2. DECISION MAKING AND DATA FUSION

Interpretation of remote sensing data can be divided to two approaches, modelling and classification. Modelling means that some geophysical parameter is estimated from remote sensing data, like soil moisture or forest stem volume. The aim of classification is to divide measurements to discrete groups or classes according to their similarities. This requires that for each measurement we make decision about the most proper or likely class.

2.1 Decision making

A common means to perform classification is to use statistical pattern recognition framework. There are several different approaches to classification but the Bayes rule is commonly utilized. Bayes rule measures the a posteriori probability $P(\omega_j | x)$ which feature vector $x$ belongs to class $j$ as (Devijer and Kittler, 1982):

$$P(\omega_j | x) = \frac{p(x | \omega_j)P_j}{p(x)},$$

(1)
where \( P_j \) is the a priori probability of class \( j \), \( p(x|\omega_j) \) is the value of density function of class \( j \) and \( p(x) \) is the mixture density function of \( x \) defined as

\[
p(x) = \sum_{j=1}^{c} p(x | \omega_j) P_j,
\]

where \( c \) is number of classes.

Feature vector \( x \) can be classified to class \( j \) if the a posteriori probability of that class is larger than the a posteriori probabilities of other classes. In that case the decision rule is called Bayes rule for minimum error. If it is thought that uncertain classification should not be done, rejection threshold \( \lambda_r \) can be used. In that case decision rule becomes

\[
\omega(x) = \omega_i \quad \text{if} \quad P(\omega_i | x) = \max_{j=1}^{c} P(\omega_j | x) \geq 1 - \lambda_r
\]

\[
\omega(x) = \omega_0 \quad \text{if} \quad \max_{j=1}^{c} P(\omega_j | x) < 1 - \lambda_r
\]

In other words, classification decision is not made if the largest a posteriori probability is less than 1 - rejection threshold \( \lambda_r \). Cost function \( \lambda(\omega_i|\omega_j) \) can be used to make some classification decisions more important than the others. The purpose of cost function is to punish wrong decisions so it can be thought to weight different decisions.

The density function measures the distance between feature vector \( x \) and class \( j \). Remote sensing data is often normally distributed, especially optical data, so normally distributed density function

\[
p(x | \omega_j) = (2\pi)^{-d/2} |\Sigma_j|^{-1/2} e^{-\frac{1}{2} (x-\mu_j)^T \Sigma_j^{-1} (x-\mu_j)}
\]

can be used. \( \Sigma_j \) is the covariance matrix of class \( j \), \( \mu_j \) the mean vector of class \( j \) and \( d \) is the dimensionality of feature space.

### 2.2 Alternatives for data fusion

Data fusion can be performed on different levels (Pohl and van Genderen, 1998):

- **Pixel based fusion** means that the measurements or measured physical parameters have been fused. In other words, the feature vector is combined directly from different data sources.
- **Feature based fusion** means that features have been extracted from different data sources using e.g. image segmentation. In this case the features can be e.g. size, shape and average intensity level of areas. These features form feature vectors describing the extracted objects.
• Decision based fusion means that the objects have been identified from individual data sources and then these interpretation results are combined using e.g. rules to reinforce common interpretation.

2.3 Decision based data fusion methods

One way to perform decision based data fusion is to use the class labels of individual classifications and making some kind of majority decision. This majority decision can be due to simple majority voting (Ho et.al., 1994) or consensus builder (Liu et.al., 2002). Other examples of this approach include methods for making the decision by ranking the class labels of individual classifications (Ho et.al., 1994), using special neural network classifier to classify samples if their statistical and neural network classifications disagree (Kanellopoulos et.al. 1993), or using the classification of expert system as input to neural network classifier (Liu et.al., 2002). This approach could be called hard decision based data fusion.

Soft decision based data fusion is another alternative. In that case probabilities or other kind of measures which pixel belongs to certain class are used. This latter approach is used in this study. All implemented data fusion methods use the output of Bayes decision rule, i.e. a posteriori probabilities, as their input.

2.3.1 Maximum a posteriori probability

In this case classification decision is based on the maximum a posteriori probability of different classifications $c_l$. Decision rule is

$$
\omega(x) = \omega_l \quad \text{if} \quad P(\omega_l | x) = \max_{k=1}^{cl} \max_{j=1}^{c} P(\omega_j | x_k).
$$

(5)

This rule corresponds to using fuzzy union operator to combine different fuzzy sets (Zimmermann, 2001). The drawback is that information about the reliability of different classifications is not used.

2.3.2 Maximum joint a posteriori probability

A joint a posteriori probability of classification decision can be computed as (Swain, 1978)

$$
P(\omega_j | x) = \prod_{k=1}^{cl} P(\omega_j | x_k),
$$

(6)

where $P(\omega_j | x)$ is the a posteriori probability of class $j$ computed using feature vector $x_k$ from data source $k$. In this case it is assumed that the individual classifications are independent from each other, but this is not always the case in real life. Another drawback is that information about the reliability of different classifications is not used. This can be incorporated by using accuracy coefficient $\alpha_k$ so that the equation of joint a posteriori probability is
\[ P(\omega_l \mid x) = \prod_{k=1}^{cl} P(\omega_l \mid x_k)^{\alpha_k}, \tag{7} \]

Accuracy coefficient \( \alpha_k \) can be related to classification accuracy (e.g. the estimated classification accuracy of \( x \)) or spatial resolution giving lower coefficient to classification that is based on lower resolution images.

### 2.3.3 A priori probabilities from lower quality probabilities or other data

A priori probability of Bayes rule represents our knowledge about the truth of the hypothesis before we have analysed the current data. This knowledge could be acquired using old classification or derived from statistical data. One way to estimate a priori probability is to compute the a posteriori probabilities of lower resolution classification and use these as a priori probabilities in higher resolution classification (Schneider et.al., 2003, Törmä et.al., 2004). This method could be characterized as "a priori probabilities of classification are a posteriori probabilities of previous classifications". In the case of remote sensing, this lower resolution could mean that the spatial, radiometric or spectral resolution is lower or otherwise less suitable for classification problem at hand.

### 2.3.4 Dempster-Shafer theory of evidence

The mathematical theory of evidence is a method to combine different data sources to provide a joint inference concerning the correct classification. The idea is to assign a so-called mass of evidence to various labeling propositions for a feature vector. Each vector has its own mass of evidence describing the likelihood of different classes as well as some indication about the uncertainty about the labeling. For each labeling proposition, values of support and plausibility are computed. Support is considered to be the minimum amount of evidence in favor of a particular labeling for a pixel whereas plausibility is the maximum possible evidence in favor of the labeling. The difference between the measures of plausibility and support is called the evidential interval, the true likelihood that the label under consideration is correct for the pixel is assumed to lie somewhere in that interval. Orthogonal sum is used to combine evidence from different sources. After the orthogonal sum has been applied, the user can then compute the support for and the plausibility of each class for a feature vector. The final decision is based on the support and plausibility. In the simplest case a maximum support rule can be used (Richards, 1993).

### 3. SAR IMAGES AND GROUND TRUTH

Study site is the Helsinki region in Southern Finland, covering some 50 km x 50 km area. ERS-1/2 images were acquired during summer 1995 - summer 1996 consists of 14 Tandem image pairs. The temporal separation between image acquisitions is 24 hours in a Tandem pair. Measurement parameters of ERS are C-band, frequency 5.3 GHz, polarization VV, incidence angle 23 degrees and spatial resolution about 30 m (Kramer, 1996). Two 5-look backscattered intensity images and an interferometric coherence image estimated with a 5x5 pixel Gaussian window were produced from each Tandem pair. All image data was orthorectified to national coordinate system using an INSAR DEM. Weather conditions within image acquisition times were also recorded. Table 1 presents the dates and environmental conditions of image acquisitions. The intensity and coherence images were scaled to 8-bit integers.
Table 1. The environmental conditions of acquired ERS-1/2 SAR images. Images have been taken approximately 10:35 UTC (Pulliainen et.al., 2003).

<table>
<thead>
<tr>
<th>Nr</th>
<th>Date</th>
<th>Track/Frame</th>
<th>Base (m)</th>
<th>Air temperature °C (9:00 UTC)</th>
<th>Wind speed (direction) (m/s, deg.)</th>
<th>Snow depth (cm)</th>
<th>Precipitation between images (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17 Jul 1995</td>
<td>408/2385</td>
<td>-2</td>
<td>19.5</td>
<td>8 (140)</td>
<td>0</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>18 Jul 1995</td>
<td></td>
<td></td>
<td>17.1</td>
<td>5 (190)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>21 Aug 1995</td>
<td>408/2385</td>
<td>-72</td>
<td>17.1</td>
<td>3 (330)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>22 Aug 1995</td>
<td></td>
<td></td>
<td>21.1</td>
<td>5 (240)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9 Sep 1995</td>
<td>179/2385</td>
<td>-28</td>
<td>11.5</td>
<td>6 (30)</td>
<td>0</td>
<td>14.9</td>
</tr>
<tr>
<td></td>
<td>10 Sep 1995</td>
<td></td>
<td></td>
<td>10.3</td>
<td>3 (30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>25 Sep 1995</td>
<td>408/2385</td>
<td>239</td>
<td>13.5</td>
<td>8 (190)</td>
<td>0</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>26 Sep 1995</td>
<td></td>
<td></td>
<td>11.7</td>
<td>10 (210)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>14 Oct 1995</td>
<td>179/2385</td>
<td>-221</td>
<td>8.6</td>
<td>6 (290)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>15 Oct 1995</td>
<td></td>
<td></td>
<td>6.2</td>
<td>1 (90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>30 Oct 1995</td>
<td>408/2385</td>
<td>-49</td>
<td>5.2</td>
<td>2 (260)</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>31 Oct 1995</td>
<td></td>
<td></td>
<td>-1.7</td>
<td>5 (340)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8 Jan 1996</td>
<td>408/2385</td>
<td>-29</td>
<td>-6.5</td>
<td>2 (120)</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9 Jan 1996</td>
<td></td>
<td></td>
<td>-5.0</td>
<td>3 (170)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>12 Feb 1996</td>
<td>408/2385</td>
<td>85</td>
<td>-14.9</td>
<td>5 (100)</td>
<td>16</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>13 Feb 1996</td>
<td></td>
<td></td>
<td>-10.8</td>
<td>1 (120)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2 Mar 1996</td>
<td>179/2385</td>
<td>-76</td>
<td>-4.2</td>
<td>4 (10)</td>
<td>16</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>3 Mar 1996</td>
<td></td>
<td></td>
<td>-5.5</td>
<td>3 (50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>18 Mar 1996</td>
<td>408/2385</td>
<td>80</td>
<td>-3.2</td>
<td>3 (350)</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>19 Mar 1996</td>
<td></td>
<td></td>
<td>-4.4</td>
<td>1 (110)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>6 Apr 1996</td>
<td>179/2385</td>
<td>37</td>
<td>5.5</td>
<td>4 (220)</td>
<td>38 (wet snow)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>7 Apr 1996</td>
<td></td>
<td></td>
<td>7.6</td>
<td>2 (130)</td>
<td>32 (wet snow)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>22 Apr 1996</td>
<td>408/2385</td>
<td>-58</td>
<td>14.7</td>
<td>5 (240)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>23 Apr 1996</td>
<td></td>
<td></td>
<td>10.0</td>
<td>3 (50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>15 Jun 1996</td>
<td>179/2385</td>
<td>48</td>
<td>15.6</td>
<td>7 (330)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>16 Jun 1996</td>
<td></td>
<td></td>
<td>14.9</td>
<td>6 (340)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>20 Jul 1996</td>
<td>179/2385</td>
<td>188</td>
<td>16.0</td>
<td>4 (70)</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>21 Jul 1996</td>
<td></td>
<td></td>
<td>14.6</td>
<td>5 (40)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chosen land cover and use classes are presented in table 2 with number of ground truth samples per class for different spatial resolutions. Training areas for classes were determined using Vantaa city regional map made by Vantaa city authorities and National Land Use and Forest Classification made by National Land Survey. Regional map and LUFC were combined to one map with 20 m pixel. In order to decrease the effect of georeferencing errors, 2 pixels were removed from borders of map areas. Then the coordinates of pixels were written to ascii-files and systematically sampled so that the training data of each class would contain about 1000 pixels at 20 m spatial resolution.

Table 2. Chosen land cover and use classes with number of ground truth samples per class for different spatial resolutions.

<table>
<thead>
<tr>
<th>Class</th>
<th>80m</th>
<th>40m</th>
<th>20m</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Water</td>
<td>74</td>
<td>287</td>
<td>1119</td>
</tr>
<tr>
<td>2. Agricultural and other open area</td>
<td>39</td>
<td>252</td>
<td>1000</td>
</tr>
<tr>
<td>3. Dense forest, stem volume over 100 m³/ha</td>
<td>61</td>
<td>265</td>
<td>1030</td>
</tr>
<tr>
<td>4. Sparse forest, stem volume 50-100 m³/ha</td>
<td>58</td>
<td>261</td>
<td>1000</td>
</tr>
<tr>
<td>5. Single story houses</td>
<td>48</td>
<td>191</td>
<td>779</td>
</tr>
<tr>
<td>6. Multi-story houses</td>
<td>40</td>
<td>164</td>
<td>705</td>
</tr>
<tr>
<td>7. Industrial area</td>
<td>62</td>
<td>239</td>
<td>1008</td>
</tr>
<tr>
<td>SUM</td>
<td>382</td>
<td>1659</td>
<td>6641</td>
</tr>
</tbody>
</table>
4. EXPERIMENTS

Statistical classifications were made using Bayes decision rule with Maximum Likelihood density function estimation method. Classes were supposed to be normally distributed. Classification errors were estimated using resubstitution and holdout methods. In resubstitution method the same set is used as training and test set and in holdout method data is randomly divided to training and test sets (Devijver and Kittler, 1982). In order to decrease the effect of random division, all classifications were repeated three times and the mean values of accuracy measures computed. Because resubstitution method is positively biased and holdout negatively biased, final error estimate was mean error of these two estimates. Error matrix was used to compare the classification results and reference data. Several accuracy measures like Overall accuracy, Producer’s accuracies of individual classes, and User’s accuracies of individual classes were computed from error matrix (Lillesand and Kiefer, 1994).

Data fusion methods presented in chapter 2.3 were implemented and used the output of the statistical classifier, i.e. a posteriori probabilities, as their input:

- Maximum a posteriori probability (DF1) method was implemented as presented in chapter 2.3.1.
- Maximum joint a posteriori probability (DF2) method has four different versions according were and how the reliability measures were used. In case A, each individual classification has same influence to final decision. In cases B, C and D, the individual classifications are weighted using the probability of correct classification or fixed values. In case B, the density function value of class is weighted by the estimate of the probability of correct classification that is same as the a posteriori probability of that class. In case C, the density function value of class is weighted by the maximum a posteriori probability of decision. In case D, fixed values were used and they were 0.5 for 80 m spatial resolution images, 0.7 for 40 m and 0.9 for 20 m.
- A priori probabilities from lower quality probabilities or other data (DF3) method used the a posteriori probabilities of lower spatial resolution classification as a priori probabilities of higher spatial resolution classification.
- Dempster-Shafer theory of evidence (DF4) method was implemented according to Richards (1994). There were two versions differing how the uncertainty of labeling was defined. Case A used a posteriori probabilities of individual classifications, as case B had fixed values 0.5 for 80 m spatial resolution images, 0.7 for 40 m and 0.9 for 20 m.

5. RESULTS

5.1 Effect of image averaging to class distributions

Examples of class distributions are presented in figure 1. Spatial resolutions are 20, 40 and 80 meters and classes Water and Dense forest. Distributions are very noisy and peaked. This is due to limited number of samples per class (table 2) and speckle which increases intraclass variance. Distributions at the best spatial resolution are usually exponential as with Water or gamma as with Dense forest in intensity image taken 30.10.1995. In some rare case distributions are normal as with Dense forest in coherence image taken 30.-31.10.1995. Image averaging decreases intraclass variance and exponential distributions are transformed to gamma distributions or gamma distributions start to look like normal distributions. Another way to decrease intraclass variance and enhance the normality of class distributions is to use median filtering. Normal distribution was used in classification due to simplicity of classification.
Figure 1. Examples of class distributions.

Figure 2. The effect of spatial resolution to the classification accuracy.
5.2 Effect of spatial resolution to statistical classification

Figure 2 illustrates the effect of spatial resolution to the classification accuracy. Horizontal axis corresponds to ERS-1/2 Tandem pair (table 2) and vertical axis the overall accuracy. Solid line represents the accuracies acquired using 20 m spatial resolution images, solid line with "x" 40 m images, solid line with "o" 80 m images. Dashed lines represent corresponding median filtered images.

From classification accuracy point of view, averaging of SAR images increases overall accuracies. This is due that the random component of class statistics is suppressed. The drawback is that the spatial resolution decreases at the same time. When the overall accuracies of 20 and 40 m spatial resolution images are compared, it is noticed that the difference in accuracies is 0-6 %-units depending on Tandem pair and 2.9 %-units on average. It seems that differences are largest during summer or winter in rather dry conditions. In other words, moisture decreases the differences of classification accuracies obtained using different spatial resolutions. When 20 and 80 m, or 40 and 80 m images are compared, the differences are larger. They are on average 7.5 %-units between 20 and 80 m images and 4.6 %-units between 40 and 80 m images. In these cases there are no clear correlation between difference and image acquisition times.

The classification accuracies depend quite heavily on acquisition time and their environmental and weather conditions. The worst classification accuracies were acquired using Tandem pair 4 in relatively windy (wind direction about same in different dates) conditions and Tandem pair 11 in wet snow conditions. The dates of best classification accuracies varied more and depended on spatial resolution. The best overall accuracy, 44.9%, for 20 m images was obtained using Tandem pair 5, 46.6% for 40 m spatial resolution using Tandem pair 8, and 52.8% for 80 m spatial resolution using Tandem pair 6.

The median filtering of images increases classification accuracy. The increase is 2-6 %-units (average 3.8 %-units) in the case of 20 m spatial resolution images, 2-7 %-units (average 4.3 %-units) in the case of 40 m images and 0-5 %-units (average 2.9 %-units) in the case of 80 m images. The best overall accuracy, 47.9%, for 20 m images was obtained using Tandem pair 8, 52.3% for 40 m spatial resolution using Tandem pair 2 and 55.9% for 80 m spatial resolution using Tandem pair 3.

5.3 Comparison of data fusion methods

Figure 3 illustrates the classification accuracies of different data fusion methods. Horizontal axis corresponds to ERS-1/2 Tandem pair and vertical axis overall accuracy. Solid line represents the accuracies of statistical classifier and 20 m spatial resolution images, solid line with "x" data fusion DF1, solid line with "o" DF2C, dashed line with "x" DF3 and dashed line with "o" DF4A.

Data fusion increased classification accuracies in each case. The best method was DF3, the increase of overall accuracy is 7-14 %-units depending on date and 10.4% on average. DF2C was the second best; the increase is 4-10 %-units and 7.2 %-units on average. The increase of accuracy was usually smaller during spring or autumn. The results of different versions of DF2 were close to each other, but DF2C was usually the best. The differences between versions of DF4 were larger. The worst dates were same as before. The best dates varied between different methods. The best overall accuracy, 55.0%, was obtained using DF3 and Tandem pair 2.
Figure 3. The classification accuracies of different data fusion methods as function of time.

Figure 4. The classification accuracies of different data fusion methods as function of season.
Median filtering increased classification accuracy somewhat. The increase was larger, varying 2-6 \%\text{-}units, for DF1 and DF4A and smallest, varying 0-3 \%\text{-}units, for DF3. The best overall accuracy, 56.5\%, was obtained using DF3 and Tandem pair 2. This means that the need of spatial filtering can be at least partially compensated using data fusion of different spatial resolution images.

### 5.4 Effect of season

The seasonal comparison was made by computing seasonally averaged Tandem pairs. Tandem pairs 1, 2, 13 and 14 were temporally averaged to form summer Tandem pair, Tandem pairs 3, 4, 5, and 6 autumn Tandem pair, Tandem pairs 7 and 8 winter Tandem pair and Tandem pair 9, 10, 11 and 12 spring Tandem pair. Also, all 14 Tandem pairs were temporally averaged to form one Tandem pair. Figure 4 illustrates the classification accuracies of different data fusion methods as function of season. Horizontal axis corresponds to seasonal Tandem pair (1: whole year, 2: autumn, 3: winter, 4: spring and 5: summer) and vertical axis overall accuracy. Solid line represents the accuracies of statistical classifier and 20 m spatial resolution images, solid line with "x" data fusion method DF1, solid line with "o" DF2C, dashed line with "x" DF3 and dashed line with "o" DF4A.

Temporal averaging increases classification accuracy but surprisingly little. The overall accuracies are about the same for best seasonal Tandem pairs and median filtered Tandem pairs for DF1, DF2C and DF4A. The increase was larger in the case of DF3. When temporal averaging included all Tandem pairs, accuracies were much better. The best overall accuracy, 66.1\%, was obtained using DF2C. In the case of DF3, the increase of classification accuracy was smaller and actually the overall accuracy of summer Tandem pair was higher (60.4\%) than the Tandem pair of whole year (58.3\%).

### 5.5 Individual classes

The classwise producer's and user's accuracies varied a lot. Table 3 presents the best accuracies for each class, used data fusion method and corresponding Tandem pair. Water and agricultural and other open areas have been classified rather well, dense forest and industrial moderately and other classes rather poorly. The best data fusion method seems to be DF3.

There was some correlation between season and class accuracies. Producer's accuracies of class Water were smaller during snow-season than other seasons. In the case of Agricultural and other open areas, producer's accuracies were larger during dry snow. The user's accuracies of classes Water and Dense forest were highest during autumn. The user's accuracies of classes Agricultural and other open areas and Single story houses were largest during winter and Multi-story house the best accuracy was achieved just after snowmelt.

<table>
<thead>
<tr>
<th>Table 3.</th>
<th>The best producer’s and user’s accuracies for each class, used data fusion method and corresponding Tandem pair.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class</strong></td>
<td><strong>Producer’s accuracies</strong></td>
</tr>
<tr>
<td>1. Water</td>
<td>98%, DF3, 5,14</td>
</tr>
<tr>
<td>2. Agricultural and other open area</td>
<td>78%, DF3, 10</td>
</tr>
<tr>
<td>3. Dense forest, stem volume over 100 m3/ha</td>
<td>61%, DF3, 8</td>
</tr>
<tr>
<td>4. Sparse forest, stem volume 50-100 m3/ha</td>
<td>44%, DF3, 3</td>
</tr>
<tr>
<td>5. Single story houses</td>
<td>40%, DF2A, 1</td>
</tr>
<tr>
<td>6. Multi-story houses</td>
<td>36%, DF3, 13</td>
</tr>
<tr>
<td>7. Industrial area</td>
<td>65%, DF3, 6</td>
</tr>
</tbody>
</table>
Median filtering increased classwise accuracies somewhat, the increase varies 1-3 % units depending on accuracy measure and date. The most notable difference was class Single story house, in which case user's accuracy increases 12 %-units. The advantage of medium filtering is that variation of classification accuracy decreases between worst and best accuracies.

5.6 Mixing of classes

The mixing of classes to other classes was studied using constructed error matrices. Usually mixing happened following way:

- Water: mixed with Dense forest and Sparse forest
- Agricultural and other open areas: Multi-story houses and Sparse forest
- Dense forest: Sparse forest and Water
- Sparse forest: Dense forest and Single story houses
- Single story houses: Multi-story houses and Sparse forest
- Multi-story houses: Industrial and Single story houses
- Industrial: Multi-story houses and Single story houses

As expected, forest classes were mainly mixed with each other and build-up classes with each other. Main exception was that Single story houses and Sparse forest is mixed with each other. Due to polarization, backscatter of water areas is very variable. Therefore, water is mainly mixed with forest classes. Greatest surprise was that Agricultural and other open areas are quite heavily mixed with Multi-story houses. There were little differences between different data fusion methods. Largest differences were between statistical classification of 20 m spatial resolution images and data fusion methods. In statistical classification, Agricultural and other open areas were mixed more with build-up areas.

5.7 Merging of classes

The effect of merging of classes from seven classes to four classes was tested using Tandem pair 6. Four classes were Water, Agricultural and other open areas, Forest (classes Dense and Sparse forest), and Build-up area (classes Single story houses, Multi-story houses and Industrial). Overall classification accuracies increased quite a lot, as expected. This increase was about 20 %-units, from 18.8 %-units in case of statistical classification of 20 m spatial resolution images to 23.7 %-units in case of DF2B. The best classification accuracy, 77.7%, was achieved with DF3. Mixing of classes was mainly as Water was mixed with Forest, Agricultural with Build-up, Forest with Build-up, and Build-up with Forest. There were no differences with different data fusion methods in this sense.

6. DISCUSSION AND CONCLUSIONS

Four different data fusion methods for classifying SAR images with different spatial resolutions were tested and compared. The best method was DF3 where a’posterior probabilities of lower spatial resolution classification are used as a priori probabilities of higher resolution classification. The increase of overall accuracy was 7-14 %-units depending on date and 10.4% on average when compared to original Tandem pairs. The increase of accuracy was usually smaller during spring or autumn.
Median filtering increased classification accuracy, but not that much when data fusion methods were used. This means that the need of spatial filtering can be at least partially compensated using data fusion of different spatial resolution images.

Class distributions were characterized using normal distribution. One way to enhance classification results is to select the best distribution from exponential, gamma and normal distributions for each image and spatial resolution or even for each class.

The accuracy and spatial resolution of spaceborne SAR interferometry derived information for cellular network planning may not be good enough for urban areas. In these areas aerial data like laser scanning, aerial images or SAR instruments should be used. But spaceborne data should be good enough for rural and natural areas. Therefore, the selection of data should be based on properties of mapped area.

The main disadvantage in using SAR interferometry for obtaining information for cellular network planning is the lack of suitable SAR data. The coverage of ERS-1/2 Tandem mission is good but it has been imaged 1995 – 96 so that dataset is quite old. Temporal separation of image acquisitions for ERS-2 (RS-1 is not operational anymore), Envisat and Radarsat-1 is the same as their repeat cycle which is 35 days for ERS-2 and Envisat and 24 days for Radarsat-1. These are too long times for this purpose where time separation should be as small as possible in order to minimize the disturbances due to changing weather.

Therefore some combination of two instruments should be found. The time separation between ERS-2 and Envisat is typically only 30 minutes. Unfortunately, there is small difference between the center frequencies of the instruments. This difference can be corrected computationally (Guarnieri and Prati, 2000) and the correction has been implemented in e.g. InSAR software of Vexcel Canada Inc. (now part of Microsoft). This correction works well if the baseline of images is around 2300 m which may happen on occasion for equatorial latitudes (Wessels, 2006). The combination of Radarsat-1 and coming Radarsat-2 does not work either because there is also difference of 0.1 GHz in center frequencies meaning that the baseline should be around 5000 m so it is unlikely that useful interferograms will come out from this combination.

Therefore, at this moment there is no suitable spaceborne interferometric SAR datasuorce for cellular network planning. There are future plans for SAR constellations (group of satellites with SAR instruments and short revisit times) like Radarsat-3 (6 satellites 2009 - 2013) or COSMO-SkyMed (4 satellites) but it will take several years to get them to work.

7. REFERENCES


**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. Also thanks to John Wessels who provided valuable comments and anonymous reviewers constructive criticism.