A New DLR Classification Procedure based on ERS Coherence and JERS "PEAK" Models

Wolfgang Wagner, Jan Vietmeier, DLR

16.5.2000

1. Introduction

We have been working on a new classification procedure which is based on the simple, yet very successful CESBIO classifier. However, instead of using fixed class centres we are using a ERS coherence model and a JERS backscatter model to determine image-specific class centres. These models are based on the hypothesis that the observed peaks in the ERS coherence and JERS σ^0 histograms can be used as input parameter into these models.

2. A New Hypothesis

In previous work we have tried to estimate the coherence model parameters γ_0 and γ_∞ from the γ histograms by using percentiles or range-thresholds. This is somewhat problematic because the γ histograms do not just depend on the image-specific distribution of γ related to the interferometric baseline or environmental conditions but also on the frequency distribution of landcover classes within the scenes. So we recognised that the problem was to find some property of the histograms that conveys information about the image-specific distribution of the γ-values of forested land.

If we assume that dense forest, i.e. where γ goes into saturation, is the dominating landcover for all scenes of our SIBERIA area then we should find, for each image, a peak in the γ histogram that is representative for dense forest. Looking at the histograms we can indeed find such peaks in the range from about 0.18 to 0.35.

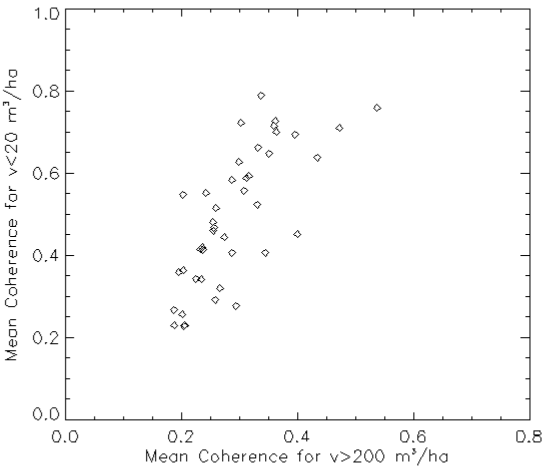

Fig. 1: Scatterplot of mean coherence values for growing stock volumes v lower than 20 m^3/ha and greater than 200 m^3/ha for individual testsites from the common data base.

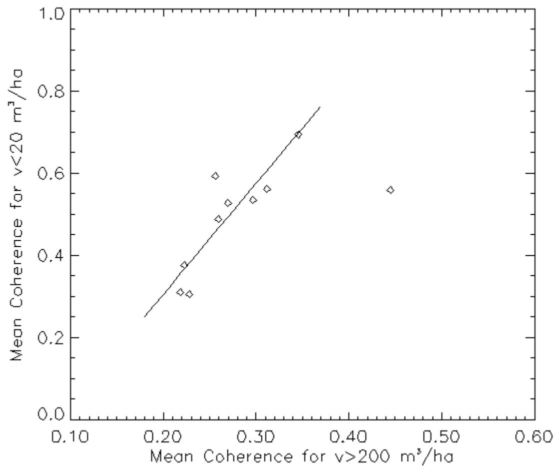

Fig. 2: Scatterplot of mean coherence values for growing stock volumes v lower than 20 m^3/ha and greater than 200 m^3/ha for enterprises. The outliner comes from the Hrebtovsky forest enterprise.

However, one parameter is not enough to describe the exponential coherence model so that we need to introduce a second assumption. In the common data base analysis we observed that γ_0 and γ_∞ are to some extent correlated. Therefore we assume that γ_{peak} can also be used to estimate the γ of low growing stock volumes. To make a preliminary test of this hypothesis we went into the common data base and calculated for each testsite the mean γ for all polygons with a growing stock volume lower than 20 m^3/ha and greater than 200 m^3/ha. As Fig. 1 shows these two parameters are to some extent correlated (R^2 = 0.55). Secondly we calculated the mean γ values for the two v ranges for entire enterprises, i.e. for all testsites within the enterprises. Excluding the data from the Hrebtovsky forest enterprise this increased the coefficient of determination R^2 to 0.77 (Fig 2). Unfortunately, we cannot make use of this relationship because

- the distribution of single γ values is probably different from the distribution of polygon-averages γ values as can be found in the common data base;
- polygons in hilly terrain are not flagged in the common data base.

Nevertheless, our new working assumptions are:

1. The peak in the γ histograms in the range from about 0.18 to 0.4 (after removing water from the images) can be used to characterise dense forest;
2. There is a linear relationship between γ of dense forest and open forest.

3. A Simplified Coherence Model

The coherence model is given by the following equation:

$$\gamma(v) = \gamma_\infty + (\gamma_0 - \gamma_\infty) \cdot e^{-\frac{v}{100}} \qquad (1)$$

The mathematical expressions of the two assumptions above are:

1. $\gamma_\infty \approx \gamma_{peak}$
2. $\gamma_0 = a + b\gamma_\infty \approx a + b\gamma_{peak}$

This gives:

$$\gamma(v) = \gamma_{peak} + \left[a + (b-1)\gamma_{peak}\right] \cdot e^{-\frac{v}{100}} \qquad (2)$$

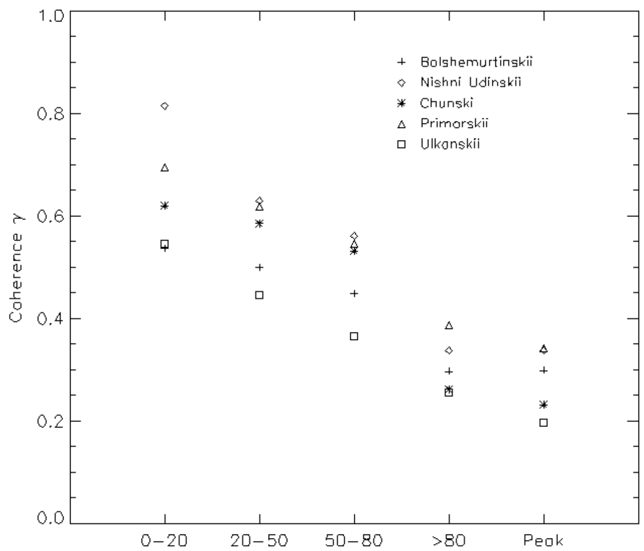

Fig. 3: Median γ values of the four forest classes and γ_{peak}.

To determine the new model parameters a and b we used the data from our well known testsites Bolshemurtinskii, Nishni Udinskii, Chunski, Primorskii, and Ulkanskii and calculated the mean and median γ values for our four forest classes (v < 20, 20 < v < 50, 50 < v < 80, v > 80) and extracted γ_{peak}

from the γ-histrogram of the corresponding ERS coherence image. Fig. 3 shows the median values for the five testsites for the four forest classes plus the peak values.

By setting v equal to 10, 35, 65, and 200 m^3/ha for four forest classes and by fitting equation (2) to the median values of the four classes we obtained a = 0.389 and b = 1.02. The residuals are plotted in Fig. 4. The maximum error is about 0.1 and the standard deviation of the residuals is 0.057. This is a relative good agreement given the simplicity of the model.

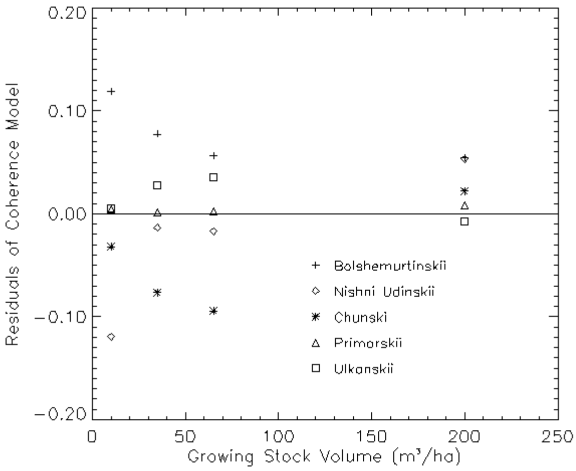

Fig. 4: Residuals of fit of equation (3) to the class centres of five testsites.

4. A JERS Model

For JERS we follow the same line of thought. We first postulate an exponential model:

$$\sigma^0 = \sigma_0 + (\sigma_\infty - \sigma_0) e^{-\frac{v}{V_\sigma}} \tag{3}$$

where σ_0 is the backscattering coefficient at growing stock volume v = 0 m^3/ha, σ_∞ is the saturation level for dense forest stands, and V_σ is a characteristic growing stock volume. Then we assume that the peak in the σ^0 histogram characterises backscatter for dense forest:

$$\sigma_\infty = \sigma_{peak} \tag{4}$$

and that

$$\sigma_0 = \sigma_{peak} - a \tag{5}$$

This gives the following simple relationship:

$$\sigma^0(v) = \sigma_{peak} - a \cdot e^{-\frac{v}{V_\sigma}} \tag{6}$$

The interpretation of (6) is that σ_{peak} characterises the wetness conditions during the image acquisition which can change the overall magnitude of the curve. Otherwise the shape of the curve is fixed by the second term on the right hand side of equation (6).

Again, to determine the model parameters a and V_σ we used the data from Bolshemurtinskii, Nishni Udinskii, Chunski, Primorskii, and Ulkanskii (Fig. 5). Fitting model (6) to the data gave a = 1.87 dB and V_σ = 45 m^3/ha. Fig. 6 shows the residuals which have a standard deviation of 0.73 dB. The maximum error is about 1.5 dB. Again, given the simplicity of the approach these are acceptable values.

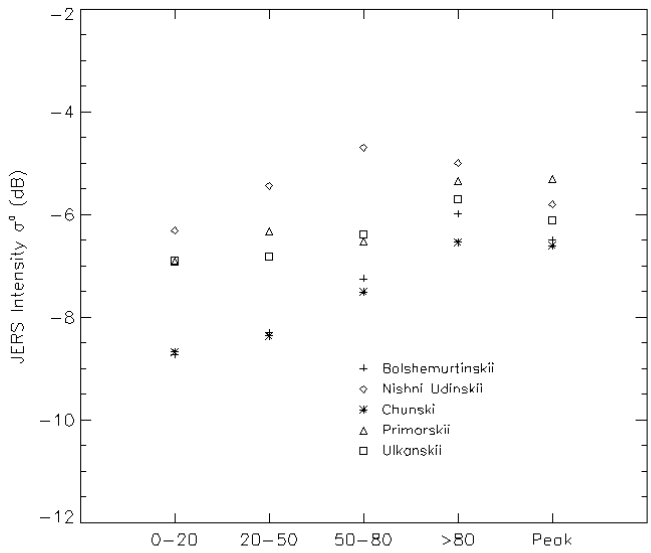

Fig. 5: Median σ^0 values for four forest classes and σ^0 peak.

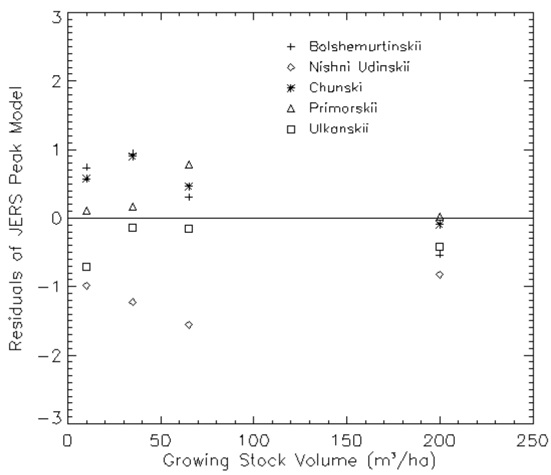

Fig. 6: Residuals of fit of equation (6) to the class centres of five testsites.

5. Classification using the PEAK Models

The steps of the new classifier based on the PEAK models for the ESR coherence and the JERS backscattering coefficient are:

1. Remove water by using simple thresholds on the γ and JERS σ^0 images;
2. Determine γ_{peak} and σ_{peak} from the corresponding histograms and calculate the class centres of the four forest classes with models (2) and (6);
3. Use the class centres for water and smooth surfaces defined by CESBIO;
4. Use the standard deviation for the 6 classes defined by CESBIO and assume that γ and σ^0 are not correlated for the 6 classes.
5. Use the derived statistics for a standard maximum likelihood algorithm;

The results can be seen in the table below in the rows labelled with "PEAK" and "PEAK-ICP".

all 6 classes

Algorithm	Bolshe 32400	Nishne 32414	Chunski 32543	Prim 32600	Ulkan 32657	mean	stddev
UWS	0.547	0.915	0.694	0.925	0.343	0.6846	0.2482
DLR	0.592	0.911	0.658	0.929	0.349	0.6876	0.2413
PEAK	0.657	0.915	0.755	0.982	0.768	0.8154	0.1310
NERC	0.700	0.846	0.736	0.910	0.348	0.7079	0.2183
CESBIO	0.767	0.844	0.773	0.966	0.832	0.8364	0.0799
CESBIOICP	0.789	0.855	0.788	0.967	0.852	0.8500	0.0729
CESBIOICPsmall	0.987	0.829	0.835	0.953	0.409	0.8026	0.2309
SCEOS4		0.838	0.755	0.951	0.396	0.7352	0.2398
SCEOS5		0.851	0.739	0.941	0.390	0.7301	0.2414
HYBRID	0.547	0.914	0.690	0.977	0.823	0.7902	0.1736

4 forest classes

Algorithm	Bolshe 32400	Nishne 32414	Chunski 32543	Prim 32600	Ulkan 32657	mean	stddev
UWS	0.467	0.896	0.687	0.666	0.315	0.6061	0.2227
DLR	0.516	0.891	0.655	0.666	0.374	0.6204	0.1923
PEAK	0.601	0.897	0.755	0.696	0.380	0.6658	0.1925
NERC	0.653	0.455	0.718	0.185	0.056	0.4134	0.2881
CESBIO	0.733	0.818	0.766	0.547	0.401	0.6531	0.1741
CESBIOICP	0.758	0.826	0.784	0.562	0.416	0.6692	0.1741
CESBIOICPsmall		0.829	0.834	0.473	0.416	0.6380	0.2247
SCEOS4		0.838	0.753	0.468	0.403	0.6155	0.2126
SCEOS5		0.851	0.738	0.474	0.400	0.6157	0.2136
HYBRID	0.467	0.896	0.687	0.666	0.315	0.6062	0.2226

6. Conclusions

We think we have made significant progress with the model-based classification approach but it would be necessary to consider all original testsites for improving the models. For example, the uncertainty of the estimated model parameters (a, b, V_σ) is still somewhat too large. Also it may be possible to find the one or other formulation that even better describes the data.

Last but not least we still fighting with some software like the maximum likelihood classifier or ICP. With ICP we have the problem that the results get worse and with our ML algorithm we cannot exactly reproduce CESBIO's results ...