# Report 2

Correct implementation of all metrics

Now all metrics are correctly implemented and tested especially the most problematic one , the multi-look system angle-feature .

Multi-look system angle feature

The test for this metrics use the energy distribution of the angle features , we can superpose the VAD segments with the energy distribution of the angle features for each angles , we should have a high correlation with one speakers and one angle
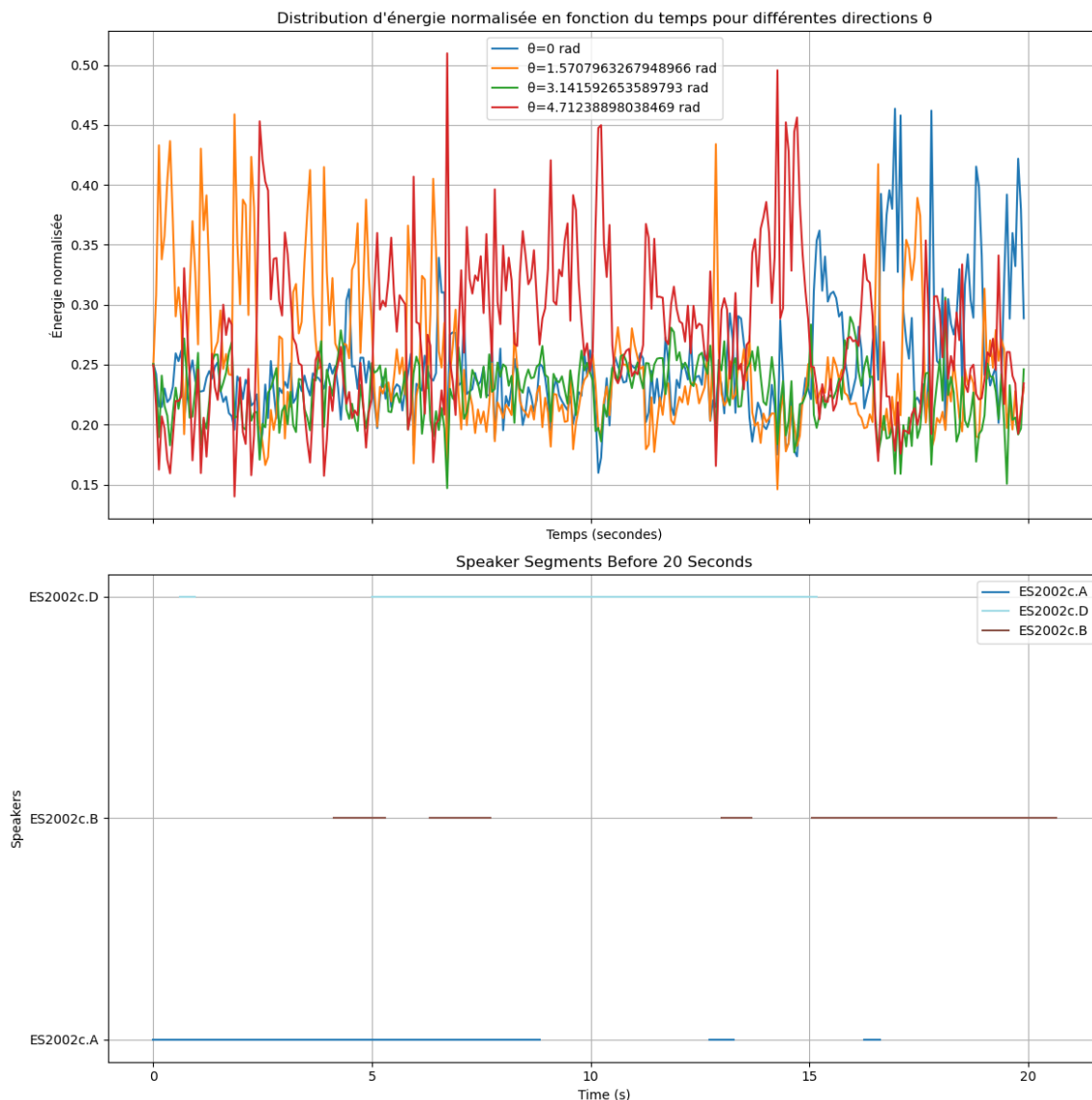
Here is the results I had , we can see that ES2002.A is highly correlated to 1.57rad , ES2002.B with 0rad and ES2002.D with 4.7rad

Transformation of the segments into labels

To create a properly dataset we need to have function which will transform raw segments into a wanted tensor , here we need voice activity labels and overlapped speech detection labels , and this tensor will depend of the number of frame and the frame size

The results are below
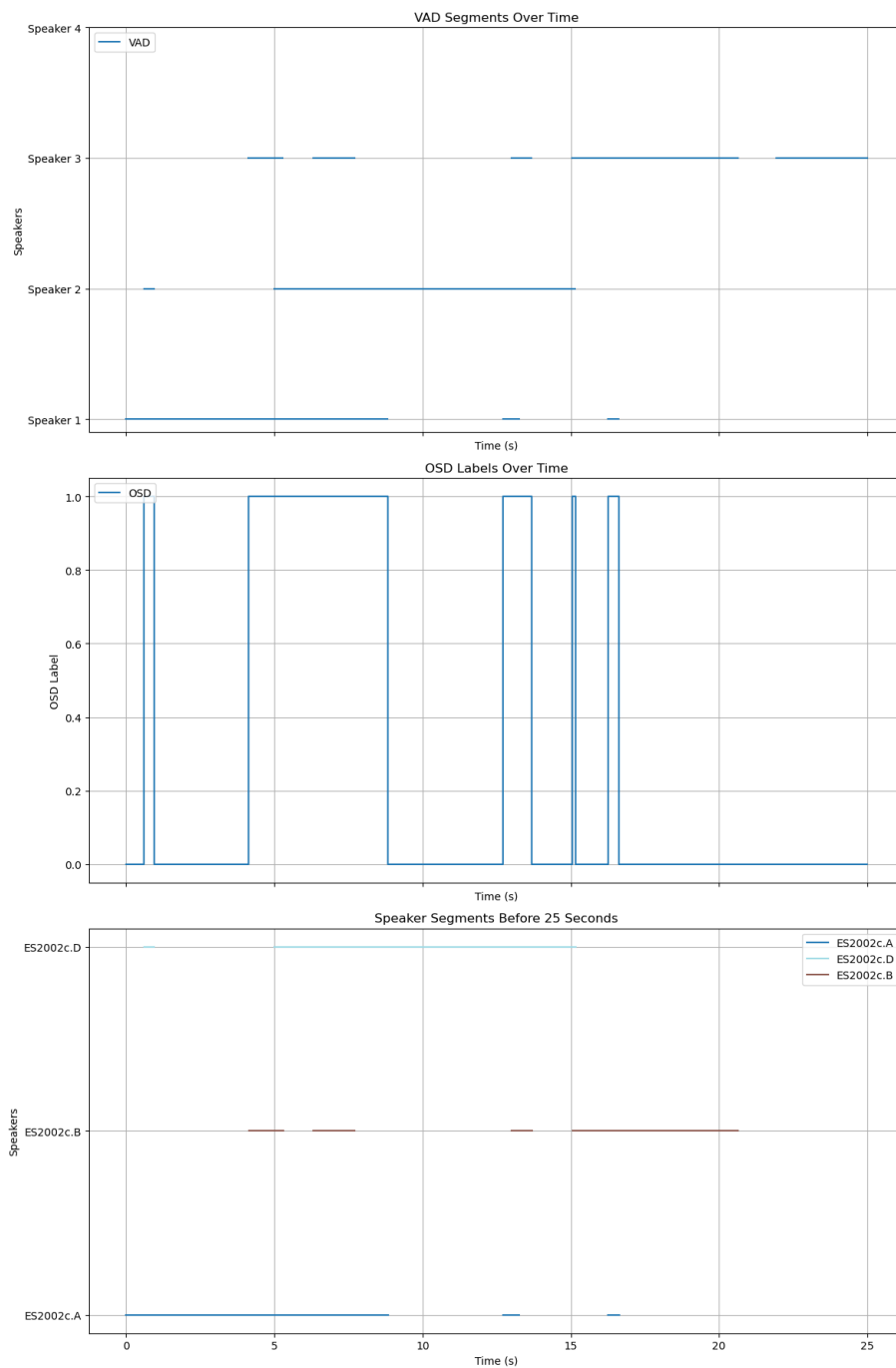
Finally I am beginning to implement the model which will use this labels

All the details of the model are provided in my paper.

```
7  class ConvTasNet(nn.Module):
8      def __init__(self, input_dim, output_dim, num_channels, bottlene
9          super(ConvTasNet, self).__init__()
0          self.encoder = nn.Sequential(
1              nn.Conv1d(input_dim, bottleneck_size, 1),
2              nn.ReLU(),
3              nn.Conv1d(bottleneck_size, num_channels, kernel_size, pac
4              nn.ReLU(),
5              nn.Conv1d(num_channels, output_dim, 1),
6              nn.ReLU()
7          )
8          self.decoder = nn.Sequential(
9              nn.ConvTranspose1d(output_dim, num_channels, kernel_size
0              nn.ReLU(),
1              nn.ConvTranspose1d(num_channels, bottleneck_size, 1),
2              nn.ReLU(),
3              nn.ConvTranspose1d(bottleneck_size, input_dim, 1),
4              nn.ReLU()
5          )
6
7      def forward(self, x):
8          encoded = self.encoder(x)
9          decoded = self.decoder(encoded)
0          return decoded
```

Once this model will be successfully implemented and trained I will begin the beam forming to enhance the input-waveforms