$$\boldsymbol{Y} = \boldsymbol{PC} = \boldsymbol{U}^{\mathrm{T}}\,\tilde{\boldsymbol{X}} \in \boldsymbol{R}^{2{,}048\times N}. \tag{3.2}$$

The proper number of PC's is found by selecting the c largest eigenvalues. Several selection approaches exist, [11], but the one used here, is shown in figure 3.8. Where the logarithm to the eigenvalues is calculated - this is the y-axis. On the x-axis the eigenvalue numbers are given. An important detail is that the eigenvalues in the matrix $\boldsymbol{\Lambda}^2$ are ordered in their magnitude such that the smaller eigenvalue number, the larger eigenvalue.

In figure 3.10 the first few eigenvalues are significantly larger than the others, but at eigenvalue number, the difference between the neighboring eigenvalues is getting smaller and smaller. One can look at the eigenvalue line and regard it as a human leg, where the first few eigenvalues correspond to the thigh, the following eigenvalues to the knee, and the remaining eigenvalues to the calf. In this selection approach, the eigenvalues up to where the "knee" ends are selected, and the corresponding eigenvectors are then applied to calculate the most significant PC's from (3.2). The more PC's included, the more information is used from the data set.

The remaining task is to calculate the residual error. This is done for every cycle by restoring the centered cycle from the c PC's calculated by (3.2). $\boldsymbol{U}^{\mathrm{T}}$ has in this situation dimension ($c \times 2{,}048$), $\tilde{\boldsymbol{X}}$ has dimension (2048×1), thus $\boldsymbol{Y}$ has dimension ($c \times 1$). the restored centered cycle $\tilde{\boldsymbol{X}}_{res}$ is then,

$$\tilde{\boldsymbol{X}}_{res} = \boldsymbol{U}\boldsymbol{Y}\,. \tag{3.3}$$

Then the residual error is found as the Squared Error (SE) of the true centered cycle and the restored centered cycle,

$$SE = \sum_{n=1}^{2{,}048}\left(\tilde{\boldsymbol{X}}_n - \tilde{\boldsymbol{X}}_{res,n}\right)^2. \tag{3.4}$$

3.2.1 PCA on sensor no. 2

The principal components analysis is performed on the AE-RMS signal from the second sensor, since this seems to include a lot of information compared to the other signals. The objective with PCA is to observe the unstable region, but also to observe the other engine condition changes. The test set is divided into five experiments, which are:

1. Oil lubrication turned off.
2. Load increased from 25% to 50%.
3. Beginning and end of the unstable region.
4. Load increased from 50% to 75%.
5. Oil lubrication turned on.

PCA is performed on each experiment. The size of the training set N is chosen to be equal to 50 cycles, since the experiments begins 150 cycles before the change point. Thus a reasonable number of cycles in the normal condition is present. In figure 3.10 the log-eigenvalues are shown for the first experiment. From this figure it seems reasonable to pick 20 components, since this corresponds to were the “knee” ends. The result is the same with the other experiments, thus the number of components in these experiments is also chosen to be 20[1].

Figure 3.11-15 Show the residual error as a feature signal for the five experiments for sensor no. 2. Notice that the feature signals begin at cycle 51, i.e. 1 cycle after the cycles used for training. From the figures it is very likely that the changes in experiment no. 2, 3 and 4 can be detected by the change detection system, since the human eye can observe the changes. Unfortunately, the two oil changes will be difficult to detect, because the difference between the normal condition and the new condition is small. However, at this point this is not regarded as a problem since the oil changes can be detected by applying the residual error feature signal from sensor no. 1[2]. Thus if both sensor no.1 and no. 2 are applied, all changes can be detected.

Though optimization is not intended to be investigated here, appendix B includes a similar PCA test where *N* is equal to 20 and the number of components is 10. The reason for this test was to try if adequately feature signals could be provided with a small training set. The results are similar to the first test, but the difference between the conditions is decreased. However, there is still a good chance of detecting the changes.

A disadvantage with PCA is that the approach is relatively time consuming. Therefore, it could be interesting to investigate if it is necessary to train the feature signal generation system for each experiment or it is adequately to train only in the first normal condition. This is done in figure 3.16. The conclusion is that the unstable region will be almost impossible to detect, and the remaining changes can be detected. However, the SE is significantly smaller in the normal conditions, when training is applied in every experiment. This might cause a better difference between the conditions. Training only in the first normal condition can not be recommended, since the feature signal is significantly worse. The explanation could be that the features change when the engine condition changes, as mentioned in chapter 2.

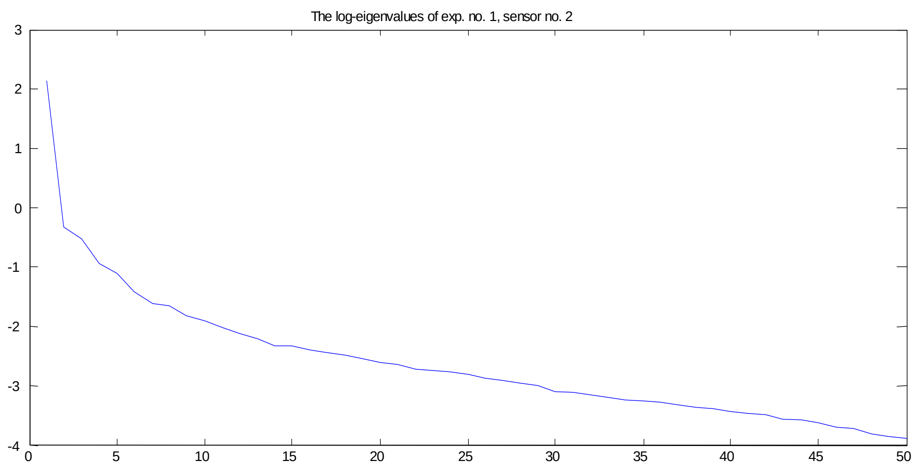

Figure 3.10: The logarithm of the eigenvalues from the first experiment. The “knee” ends approximately around 20, thus suggesting 20 components to be applied. The x-axis is the eigenvalue number and the y-axis the eigenvalues.

[1] See appendix B for the figures of the log-eigenvalues of the other experiments.
[2] See appendix B for the residual error feature signals from all experiments from all sensors.

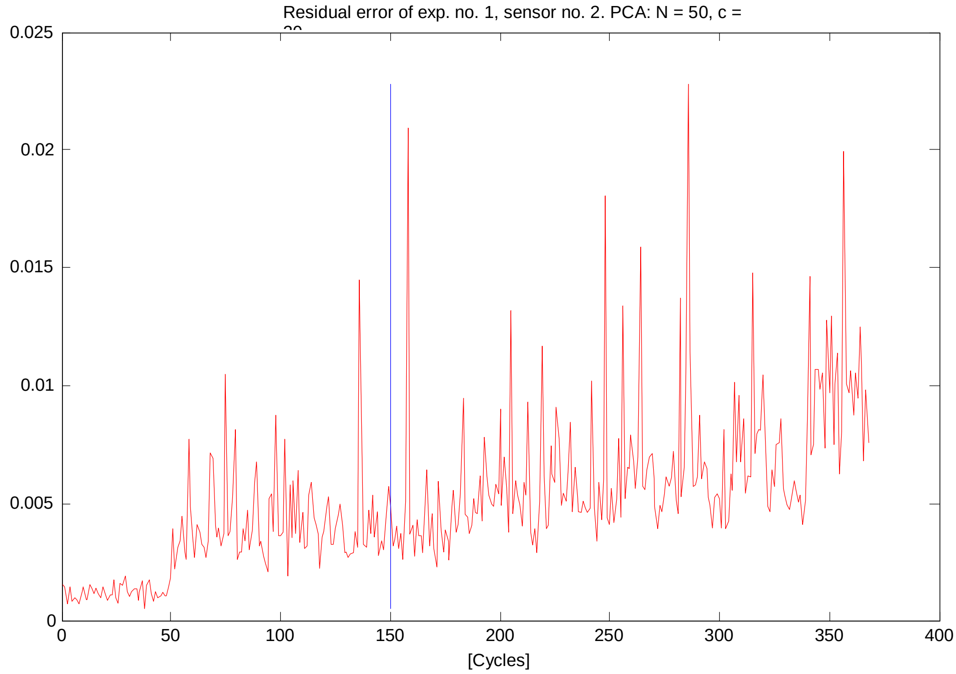

Figure 3.11: The residual error of experiment no. 1 – oil lubrication shut off. The vertical line is the change point. There is a slight increase in the mean value of the feature signal, thus automatic change detection might be feasible.

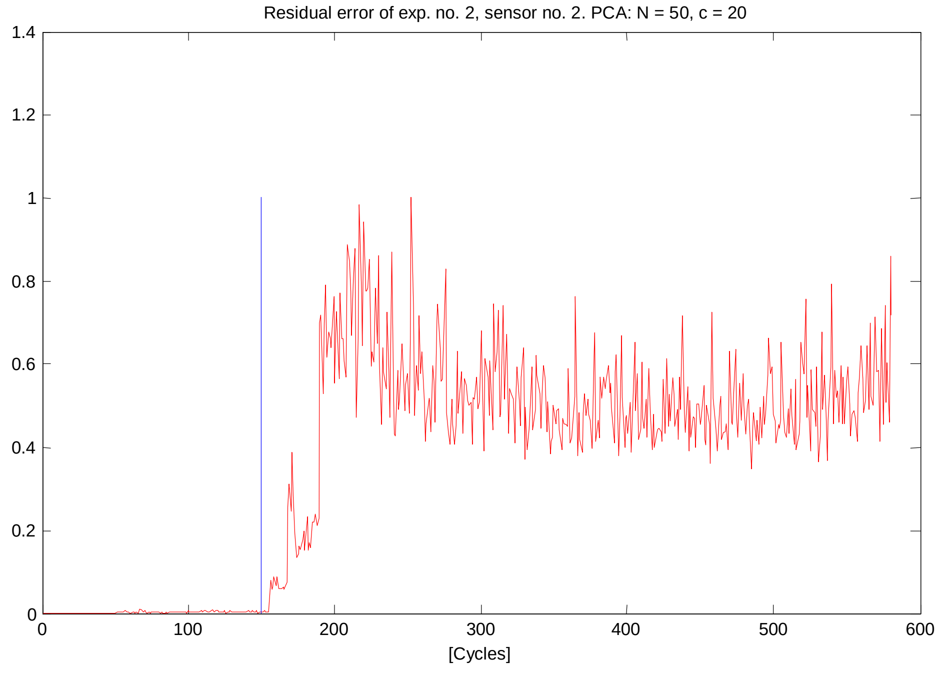

Figure 3.12: The residual error of experiment no. 2 – increasing the load from 25% to 50%. The difference between the conditions is very clear.

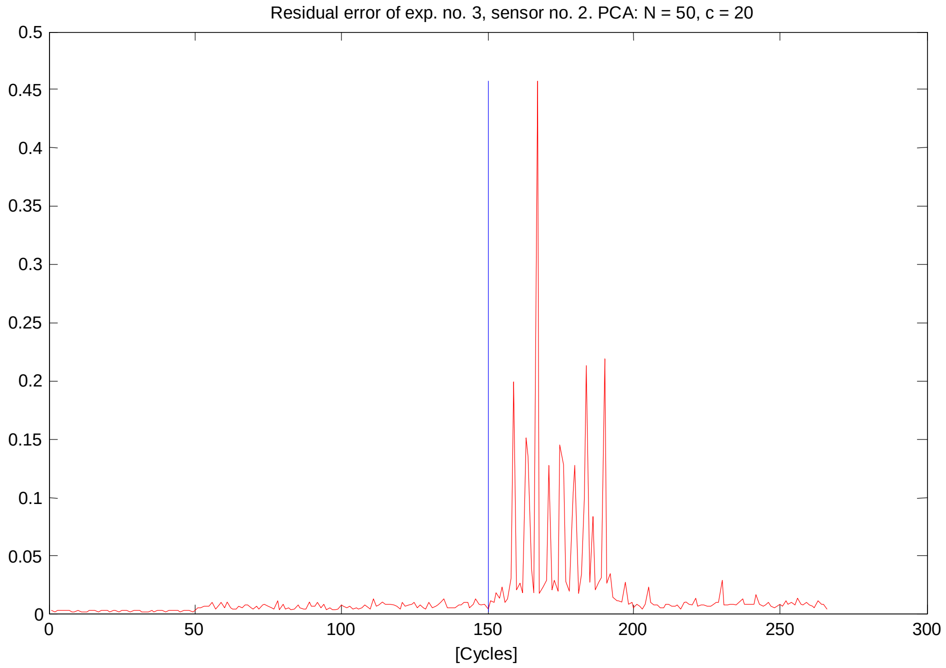

Figure 3.13: The residual error of experiment no. 3 – the unstable region. The difference between the conditions is clear.

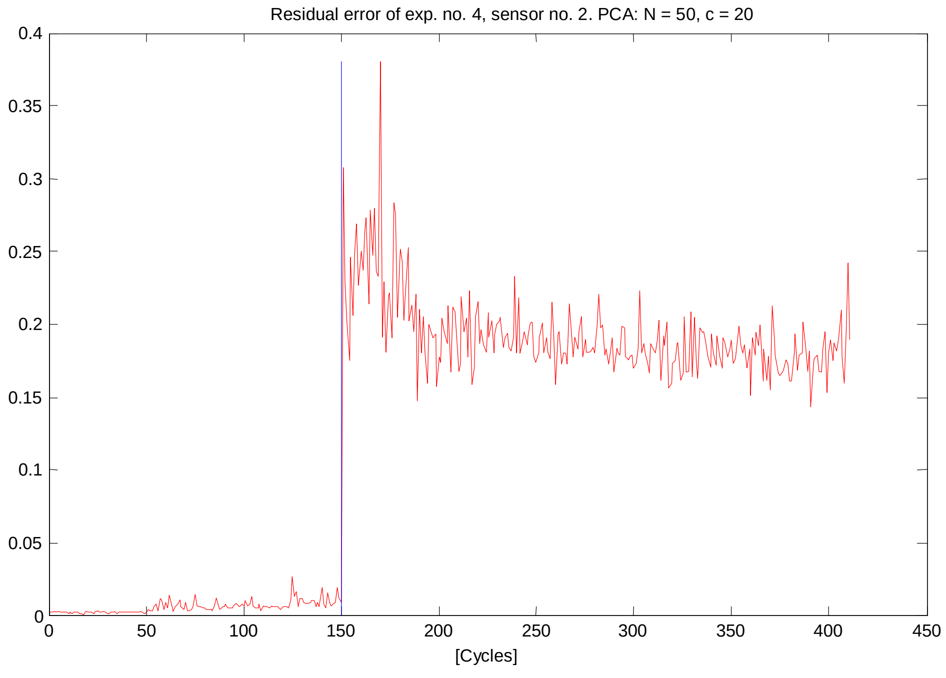

Figure 3.14: The residual error of experiment no. 4 – increasing the load from 50% to 75%. Again a very clear difference is present.

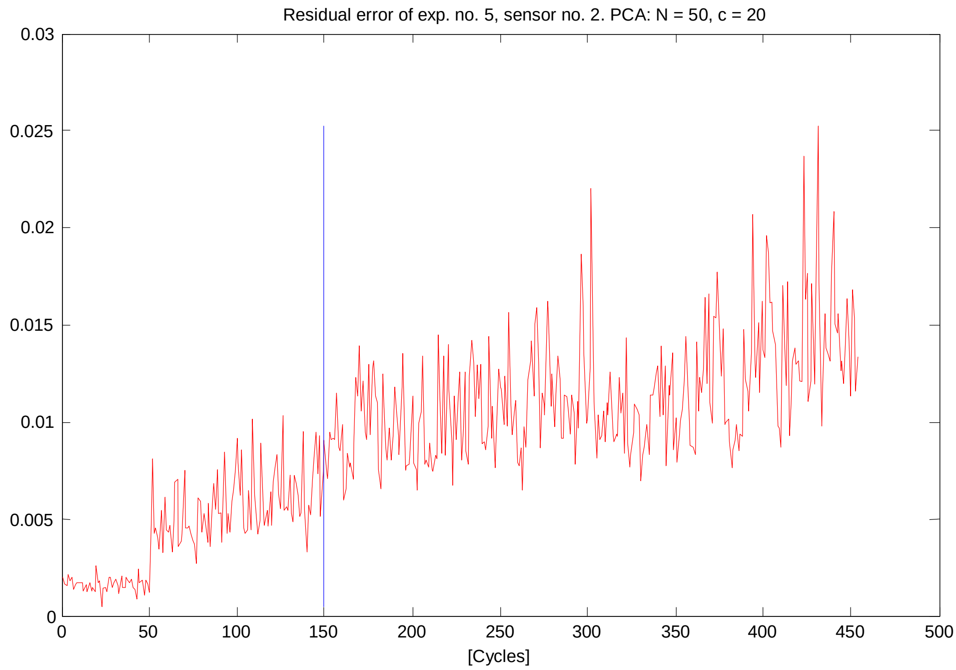

Figure 3.15: The residual error of experiment no. 5 – oil lubrication turned on. A slight difference is present.

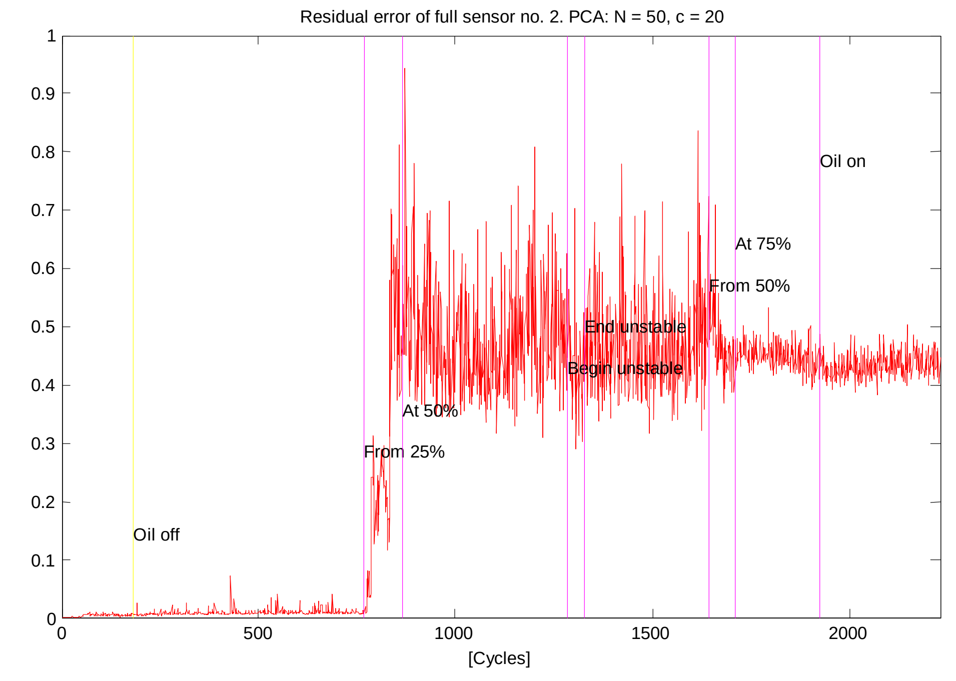

Figure 3.16: The residual error of all experiments when the training is only performed in experiment no. 1 – oil lubrication shut off. All changes seem to be observable, but the unstable region is buried. This is not a good way to create a feature signal.

3.3 Selection of cycle areas

In the re-sampling section it was stated that all the engine condition changes could be observed in the small peaks of the AE-RMS signals, i.e. in the interval [0; 0.2] for sensor no. 1. In fact the changes are more significant in this interval than in the interval [0; 1]. However, this is not true for all samples in the cycles. Figure 3.17 shows a time plot of all the cycles in the experiment for sensor no. 1 in the interval [0; 0.2]. Here it can be observed that some cycle areas show a more abrupt change than other areas. This leads to the idea that some periods of the engine cycle include more relevant information about the engine condition changes than other periods in the cycle.

Therefore it might be a good idea to split the cycle up in several areas and generate feature signals of the samples in these areas. This will provide better feature signals, since the changes will be more abrupt. This has in fact been done in the research performed by Amanda Sharkey et al., [8], [9], and the approach improved the system performance significantly. However, the approach will not be investigated in this work, since there is a risk of loosing relevant information by selecting cycle areas. One way of avoiding this is to collect the remaining cycle samples in a feature signal itself. However, since improving feature signals is a secondary goal with this thesis, the idea is abandoned, because we already have relatively good feature signals.

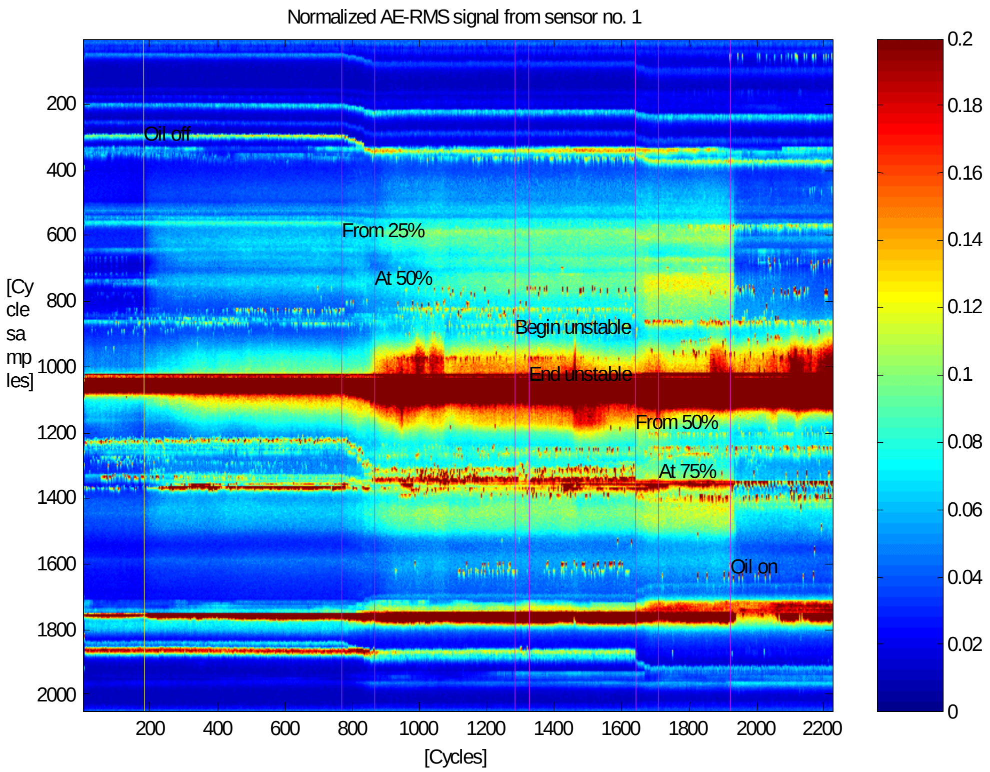

Figure 3.17: Time plot of the AE-RMS signal from sensor no. 1 in the interval [0; 0.2]. The engine condition changes are more abrupt in some cycle areas, e.g. [1,400; 1,500] and [700; 800].