Data-driven methods in application to flood defence systems monitoring and analysis

Pyayt, A.L.

Link to publication

Citation for published version (APA):
Chapter 4. Results of the Anomaly Detection

This chapter presents the results from applying the sensor data analysis and anomaly detection approaches (described in Chapter 3) to the dams listed in Chapter 2.

4.1. Stammer dike non-destructive macro-stability experiment

4.1.1. Description of the experiment

One of the dike failure mechanisms is the macro-instability of the slope. This mechanism has been extensively studied. For example, during the IJkdijk experiments in 2008 [8], a full-scale levee crashed due to the high water content inside of the dike and the heavy load placed on top of it.

The Stammer dike in Amsterdam (the Netherlands) was equipped with a detailed network of GeoBeads sensors that continuously measure pore water pressure, temperature and inclination (see Section 2.3). Sensor modules were installed in two cross-sections at depths of up to 10 m below sea level in the different soil layers (sand, clay and peat).

During the “macro-stability” test that was performed at the Stammer dike, a heavy load was placed on the dike. The objective of this test was to determine the ability of the installed sensor network to translate the influences of external circumstances on the levee stability. This situation is suitable for testing the detection of abnormal behaviour.

4.1.2. Results of the anomaly detection

Some of sensors indicated the influences of the additional load on the dike. For example, the SD_1A1_322(Y1) and SD_1C2_540(Y1) inclinometers, which were used to

* Parts of this chapter were published in:
measure the relative degrees on inclination (relative to the reference time and averaged over 1 day) of the GeoBeads suggested the presence of an additional load. Here, “SD” stands for Stammer Dike, “1” is the cross-section number, “C2” is the sensor location, and “540” is the depth of the sensor in centimetres. Figure 4.1(b) and Figure 4.1(c) show that the placement of the additional load on the dike three times was successfully detected by the selected sensors. Three rapid changes are depicted in the graphs related to the placement (time steps of 2238, 2711, and 3012) and removal (time steps 2607, 2866, and 3140) of the load [105].

These anomalies were detected using the anomaly detection approach, which is based on the application of one-class classifiers (Neural Clouds) (see Figure 4.1(a)). The data related to the normal mode were used as the training set (blue colour). The test set includes data with anomalies and “normal” data (black colour). Figure 4.1(a) shows that the suggested approach is able to detect this type of anomalies. When the confidence values are near 1, the data are “normal”, and when the confidence values are near 0, anomalies occurred.

Figure 4.1. (a) Confidence values of the normal behaviour of the dike that were calculated using Neural Clouds for the 2 parameters presented in “b” and “c”. (b) and (c) illustrate the relative inclination (in degrees relative to the reference times averaged over 1 day). The NC training period was conducted between the time stamps of 1 and 2054 and the testing period was conducted from 2055 to 3735. The x-axes for (a), (b), and (c) represent the discrete time step number, with a time step of 10 minutes between the measurements.

The input data from the two selected sensors were pre-processed and the measurements were normalised. The pre-processed dike measurements were used as inputs for training the Neural Clouds (NC). The training set (Figure 4.1) was encapsulated by the NC (Figure 4.2) [105].
4.2. Zeeland dike non-destructive piping experiment

4.2.1. Description of the experiment

The non-destructive piping experiment was organized by simply cutting off the pump. Thus, the water flow through the dike was not controlled by the pump [103].

Anomalies in the Zeeland dike (see Sub-Section 2.3.5 for more details regarding the levee) measurements are presented in Figure 4.3. For the pump, two off/on anomalies were present. In addition, one outlier was related to an unknown behaviour [99].

Figure 4.3. (a) “P1” sensor for water level (in meters). (b) Enlarged version of the “P1” sensor [99].
4.2.2. First results

The one-class classification anomaly detection approach was applied to detection abnormal behaviour. The Neural Clouds were trained on 6 parameters, T (Figure 4.4(d)), P1 (Figure 4.4(b)), P2 (Figure 4.4(c)) and the moving averages of each of the parameter. The NC classifier was trained on the normal data (the blue colour in Figure 4.4). The remaining data were used for testing the NC. The NC detected a portion of the signal that was related to turning off the pump in Figure 4.4(a) (starts at point 5410), which can be interpreted as the development of the piping in the dike or as abnormal behaviour. Low confidence values were obtained for the period of 2864-3712. According to visual inspection of the available data, this result can be interpreted as a sensor fault or as some activity in which the sensors did not behave normally [103].

The suggested anomaly detection approach is able to identify dike anomalies and unknown behaviours that result from sensor fault. However, false alarms occurred between 3500 and 5500 (Figure 4.4(a)). To improve the quality of the one-side classification approach, the features from the available measurements should be extracted.

![Confidence values](image)

Figure 4.4. Confidence values calculated by the Neural Clouds for 6 parameters. (a) Calculated confidence values of the normal dike behaviour. (b) and (c) represent the water level sensors (in meters), (d) illustrates the temperature sensor (degrees Celsius). The NC training period was conducted between time stamps of 1 and 1552 and the testing period was from 1553 to 6551. The x-axes for (a), (b), (c), (d) represent the discrete time step numbers, with a time step of 5 minutes between the measurements.

4.2.3. Feature extraction

One of the methods used for feature extraction is related to the frequency domain method and is referred to as fast Fourier transform (FFT). FFT was used for analysing the “P1” sensor measurements. Two base frequencies were extracted from the FFT spectrum at 6 and 12 hours (Figure 4.5(b)). The extracted base frequencies can be used to select the size of the sliding window for use in other methods. The selected window size should not be
smaller than the period of the lowest base frequency (in this case, 12 hours or 144 measurements) [103], [99].

![FFT of WaterHeight NAP (cm NAP)](image)

Figure 4.5. (a) Fast Fourier Transform (FFT) spectrum of “P1” (water level (meter)). (b) Enlarged portion of (a).

The portion of the “P1” values that are presented in Figure 4.3(b) was used for short time Fourier transform (STFT) analysis (Figure 4.6), continuous wavelet transform (CWT) (Figure 4.7), maximum overlap discrete wavelet transform (MODWT) analysis (Figure 4.8). During the application of STFT (Figure 4.6), anomalies were detected in time intervals where rapid increases or decreases were observed in the amplitudes of the base frequencies (Figure 4.6(b)). Anomaly detection can be conducted by only monitoring one base frequency (12 hours). This process simplifies the use of STFT for the automatic detection of anomalies in a dike.

![STFT spectrogram of WaterHeight NAP (cm NAP)](image)

Figure 4.6. For the interval of 18/03/2011 to 07/04/2011: (a) enlarged sensor measurements with anomalies; and (b) short time Fourier transform (STFT) spectrum near the anomaly.
Continuous Wavelet Transform (CWT) represents a signal in the time–frequency domain. CWT has a higher time–frequency resolution than STFT and provides more accurate features than STFT [78].

The simulation parameters were the following: the “Morlet” wavelet type was used; the scale of decomposition was 1:4:256 (each 4th scale is presented, for a total of 256 levels of decomposition)). Both modes (pump-off and unknown behaviour) can be detected by a rapid increase in the modules or by an increase in the variance of the CWT coefficients (Figure 4.7(c) – boxes #1 and #2). Figure 4.7(b) shows that a rapid change occurred in the local mean of the signal when the pump was off (boxes #1-4). Such signal behaviour affects the small scales (high frequency, e.g., from 1 to 50) of the wavelet coefficients [124]. Unknown behaviour (long outliers) and short outliers can be detected at small scales using the same method (Figure 4.7(b)).

Figure 4.7. (a) “P1” sensor measurements. (b) Continuous wavelet transform (CWT) spectrum near the dates with abnormal behaviours (18/03/2011 to 07/04/2011) enhanced from (c). (e) CWT spectrum of the signal (a).
The MODWT coefficients (see Appendix C.2) are presented in Figure 4.8. The top plot presents the signal decomposed by MODWT, and the levels of decomposition are presented from the 1st level (the second sub-figure) to the 7th level (bottom sub-figure). The dashed horizontal lines in all of the MODWT coefficient plots indicate the universal threshold (see Appendix C.4). If the value is out of the threshold, it can be interpreted as an outlier or as abnormal behaviour for the time stamp [124].

Figure 4.8. Top panel: sensor measurements near the dates with abnormal behaviours. Including, the other subpanels, the maximum overlap discrete wavelet transform (MODWT) coefficients with $W_{j,2}$ per level (solid curves) from the top (1st level) to the bottom (7th level) and the universal threshold (dashed lines).
MODWT requires less computational effort than CWT. Decreasing frequency (a higher decomposition level) causes a decreasing frequency resolution. However, a good time–frequency representation of the signal is available (in comparison with CWT). MODWT is comparable to STFT regarding the amount of time required for calculation. The detection of anomalies based on monitoring the STFT coefficients is a difficult task relative to the detection of anomalies based on monitoring the MODWT or CWT coefficients because it requires the automatic detection of the base frequencies in the FFT spectrum. In addition, STFT provides a worse resolution than CWT and MODWT.

According to the results presented in this section, MODWT was selected as the most appropriate extraction method for detecting anomalies based on the computational costs and the resolution quality requirements. In this case, the 6th and 7th decomposition levels revealed the anomalies that were produced during the pump-off mode (Figure 4.8). Further analysis is presented in the Sub-Section 4.2.4.

4.2.4. Anomaly detection using feature extraction

In this Sub-Section, the one-side classification approach is applied to the features extracted from the available measurements. First, the multi-dimensional data were pre-processed (synchronized). Feature extraction (MODWT, “la8” (least asymmetric) type, six levels of decomposition) was applied to the water level sensor (“P1”) (Figure 4.9(a)). The coefficients of the 6th level of decomposition (Figure 4.9(b)) and the coefficients of approximation of the 6th level of decomposition (smoothed version of the signal at the 6th level of decomposition) (Figure 4.9(c)) were used as inputs for the one-side classifier (NC). The calculated confidence values are presented in Figure 4.9(d). Portions of the time series are circled (dashed lines) in (Figure 4.9(b) and (c)) and were used as the training set. The remaining data were used as the test set. The training set was selected according to the available data with normal behaviour. Intervals with known anomalies were excluded from the training data set based on expert judgment. Next, the intervals with abnormal behaviour were used in the test set with the other normal test data.

Figure 4.9(d) shows that the suggested approach allows for the detection of anomalies in the levee behaviour (dashed boxes numbered 1–3), including the pump off (for box #2) and the unknown behaviour (boxes #1 and #3). Box #5 contains a signal with different time–frequency properties (increased amplitude relative to the rest of the signal). Expert knowledge is required for correctly interpret this change in the time series properties.

In addition, Figure 4.9(d) illustrates that expert knowledge is required. The anomaly in box #5 could be classified as normal or alarming behaviour by an expert. In addition, the detected abnormal data could be related to a previously unknown normal behaviour. In this case, the detected anomaly would be interpreted as an alarm. However, the new normal data could be added to the training set so that the system will not generate a false alarm if a similar behaviour is detected in the future. The use of a training set that was prepared by an expert or by a physical model is important to ensure the overall robustness of the system.
4.2 Zeeland dike non-destructive piping experiment

Figure 4.9. (a) Initial signal is represented by P1 (water level [m]). (b) Coefficients of the 6th level of decomposition ($W_6$). (c) Coefficients of approximation of the 6th level of decomposition ($V_6$). (d) Results from applying NC. The dashed circles in (b) and (c) show the training sets and the lines not covered with circles show the test set.

The anomalies detected in box #4 were enlarged in Figure 4.10. Real faults occurred in the time series (due to unknown reasons) that were identified by the one-side classifier. The anomalies in the wavelet coefficients were wider than the anomalies in the sensor data due to the boundary effects.
Chapter 4 Results of the Anomaly Detection

The detection of anomalies in the non-destructive piping experiment at Zeeland dike was successful when the pre-processed data were subjected to the Neural Clouds method. In this case, the data quality was acceptable, but the feature extraction stage was required before the one-side classifier usage. Several time-frequency methods were considered, including STFT, CWT, and MODWT. The MODWT method is the most suitable method from a computational aspect. The MODWT coefficients were used as input values for the one-side classifier to increases the overall quality of the anomaly detection system. Our approach provides a robust method for monitoring the time–frequency properties of measured parameters. The successful detection of real-world anomalies proved the usefulness of this suggested concept.

The produced anomalies were relatively evident based on the increasing mean values and amplitudes of the water levels that were detected by the P1 sensor. A series of experiments in which the pump capacity is varied from 0 to 100% in a series of steps would be useful for evaluating the sensitivity of the developed anomaly detection system. The
4.3 Analysis of the Rhine levee data

developed approach will be tested for other potential mechanisms of levee collapse (another type of anomalies). Automatic selection of the parameters used in the signal processing methods must be explored.

The qualities of the anomaly detection methods can be evaluated using four variables, true positive $Tp$ (correct identification of normal behaviour), true negative $Tn$ (correct identification of an anomaly), false negative $Fn$ (normal data are classified as abnormal), false positive $Fp$ (abnormal behaviour is classified as normal) [64].

We assume that the dike behaves normally for most of the period because it was stable for a long period and all changes occurred slowly. We identified several intervals with abnormal behaviour (including two experiments with the pump off, Figure 4.9, box #2). In addition, intervals of abnormal behaviour occurred when there were problems with the sensor values (a frozen communication problem, Figure 4.9, boxes #1 and #3). All of these anomalies were identified. The short (not evident) anomalies are presented in Figure 4.9 (box #4) and the enhanced version is presented in Figure 4.10. The alarms were generated from problems in the data. Thus, the $Tp$ can be interpreted as 100%, or near 100%.

As expected, the data were mainly classified as normal, which indicates that the $Tn$ was high. The concrete value could only be assessed after the expert evaluation of the anomaly presented in Figure 4.9 (box #5). In this case, the $Tn$ will probably be greater than 80%.

The final evaluations of $Fp$ and $Fn$ depend on expert judgment. A small number of short-time alarms were detected that were not considered critical. To improve the overall reliability of the system, we suggest introducing a parameter that will filter out short duration alarms.

4.3. Analysis of the Rhine levee data

Descriptions of the sensors that were installed into the Rhine levee are presented in Sub-Section 2.3.3.

4.3.1. Analysis of the Alert Solutions measurements

The pore pressure measurements that were gathered from the Alert Solutions sensors were converted to the water level values (Figure 4.11). Each colour corresponds to a specific sensor (for example, the “E2” sensor is marked with red in Figure 4.11 and Figure 2.11). There are two lines per sensor, a straight horizontal line indicates the depth of the sensor installation and curved lines show the water head above the depth of the sensor. If the water level is above the straight line, then this sensor is “covered” with water [100], [101].

The two boxes drawn with green dotted lines in Figure 4.11 indicate the dates when the water level was higher than the ground water level (G1 sensor), including the 9th of January, 2012 (the 1st box) and the 25th of January, 2012 (the 2nd box). These peaks correspond to the peak Rhine water levels of 820 and 710 cm [7]. According to the reported data, although the levee was wet, the “strong seepage flow ($>10^{-4}$ m/s) can be excluded.” The Alert Solutions sensors were useful for classifying the levee behaviour (dry/wet). If the piping begins near the GeoBeads sensors, it will be detected easily. A GTC Kappelmeyer fibre optic sensor should be used to test for leakage between cross-sections.
4.3.2. Analysis of the GTC Kappelmeyer measurements

In this section, we describe the collected GTC Kappelmeyer temperature measurements and the associated analyses. The fibre optic heating process is presented in Figure 4.12. As previously mentioned [7], the levee’s condition is normal because no piping was detected.

Figure 4.12. The amount of time required for the GTC Kappelmeyer fibre optic cable to heat up in the Rhine levee. The X-axis corresponds with the length of the cable (m) and the Y-axis corresponds with the temperature (Celsius). The colours indicate the temperatures across the fibre optic cable with time.
4.3 Analysis of the Rhine levee data

The only phenomena in the collected temperature measurements that could be interpreted as an anomaly occurred in a 15 m wide area (the dotted box in Figure 4.12) of the unsaturated soil. Thus, the soil contains holes that contain air. Air has a high insulation quality because it has a low thermal conductivity (~0.4 W/m/K). Thus, the presence of air results in higher temperatures. The temperature across the entire cable was generally the same during the heating process (approximately 28–30 degrees). However, the temperature was much higher from 209–221 meters (fibre optic cable length).

This anomaly can be represented as two rapid changes in the spatial observations. Different approaches can be used to detect such rapid changes in a time series (e.g., Student’s T-test) [39]. A detailed overview of more advanced detection methods is presented in [33]. In [127], wavelets were used to detect abrupt faults. Wavelets were used to analyse the water temperature measurements from the Wivenhoe Dam in [94].

The maximum overlap discrete wavelet transform (MODWT) method was used as a feature extraction method for the data analysis of the Rhine Levee. Overall, 8 levels of decomposition (“la8” represents the least asymmetric wavelet with 8 levels of decomposition) are presented in Figure 4.13.

The coefficients corresponding to the third and fourth levels of decomposition (Figure 4.13) after pre-processing were used as input values for the one-side classifier. Measurements that were related to the cold fibre and were collected during the first several minutes of heating were used for the NC training.

![Figure 4.13. The results from applying the maximum overlap discrete wavelet transform (MODWT). The X-axis corresponds with the length of the cable (m), and the Y-axis corresponds to the temperature (degrees Celsius).](image)
Values in the spatial time series after 10 minutes of heating (Figure 4.14 (a)) are presented as normal or abnormal points in Figure 4.14(b). In this case, values near 1 correspond to normal behaviour and values near 0 correspond to anomalies. Figure 4.14(a) was classified in Figure 4.14(c) by using a 2D view of the constructed Neural Clouds. A cluster with normal data related to the training set is presented in Figure 4.14(c) by the blue points. In addition, the test set is presented by the black points and the detected outliers are presented in red. MODWT does not have ‘perfect’ localization properties according to the Gibbs phenomenon [93]. Thus, some points are not correctly classified as abnormal. For example, the lower points in Figure 4.14(c) correspond to normal temperatures (Figure 4.14(a)). This example proves the functionality of the developed anomaly detection method.

Figure 4.14. Detection in the unsaturated portion of the Rhine Levee following 10 minutes of heating. (a) The input time series. (b) The calculated confidence values, where values near 1 indicated normal behaviour and values near 0 indicated an anomaly. (c) The 3rd (X-axis) and 4th (Y-axis) levels of decomposition after pre-processing. The blue points represent the training set, the black points represent the normal behaviour, and the red points are related to an anomaly.

4.4. Detecting leaks in the retaining dam

4.4.1. Anomaly detection results

GTC Kappelmeyer provided further testing of the developed anomaly detection method, which was a real example of abnormal behaviour and was recorded in the data that were collected at the earth filled dam with bitumen sealing (total length of > 2 km). These data indicated that a bitumen seal composed of asphalt-coated gravel and a bitumen binder was leaking in the dam. This anomaly is presented in the spatial time series as a rapid drop in the interval approximately 150 meters (Figure 4.15). The results of the MODWT application are presented in Figure 4.16.
4.4 Detecting leaks in the retaining dam

Figure 4.15. Retaining dam during heating of the GTC Kappelmeyer fibre optic cable with time. The X-axis corresponds with the length of the cable (m) and the Y-axis corresponds with the temperature (degrees Celsius). The colours indicate the temperatures across the fibre optic cable with time.

Figure 4.16. The results from applying maximum overlap discrete wavelet transform (MODWT) application. The X-axis corresponds with the length of the cable (m) and the Y-axis corresponds with temperature (degrees Celsius).
MODWT has been applied for extraction of features from spatio-temporal time series (“la8”, 8 levels of decomposition). The coefficients (Figure 4.16) that corresponded to the second and third levels of decomposition after pre-processing were used as input values for the one-side classifier.

The calculated confidence intervals (Figure 4.17(a)) show that the portion of the cable between 150 and 160 m is classified as abnormal, with confidence values close to 0. A 3D presentation of the constructed NC is provided in Figure 4.17(b).

Figure 4.17. Detection of leakage in the retaining dam (the 6th iteration). (a) The calculated confidence values that were nearly 1 were related to normal behaviour and those near 0 were related to an anomaly. (b) 3D view of the constructed Neural Clouds.

4.4.2. The analysis of the Rhine levee and of the retaining dam

Two types of sensor networks were installed in the levees that were considered in the last two sections, the Alert Solutions and GTC Kappelmeyer sensors. The Alert Solutions sensors (GeoBeads) that were installed in the Rhine levee measured the pore pressure, inclination and temperature simultaneously at one point simultaneously. The GTC Kappelmeyer fibre optic cable that was installed in the Rhine levee and the retaining dam measured the temperature by using the active principle, in which the cable is heated (see Sub-Section 2.2.2).

Two types of anomalies were considered in the last two Sections. The first anomaly is faulty abnormal behaviour. This anomaly was considered as unsaturated soil in the fibre optic measurements at the Rhine levee. This levee was not affected by leakage and no other anomalies were found. The measurements collected from the retaining dam contained a pattern of real leakage (second type of anomaly).

All data sets were analysed using a one-sided classification approach (Neural Clouds) in combination with a feature selection stage that provides robust detection of anomalies. Neural Clouds do not require anomalies to be presented in the training set. The time-frequency method (MODWT) was selected at the feature selection stage because it combines the presentation of input signals across time and frequency domains. The frequency analysis of the signal permits the application of the anomaly detection approach to objects with different natures, including slow processes (tides when considering earthen levees) and fast processes (vibrations when considering concrete structures). Time presentation is often required to compare these values with thresholds.
The anomalies in GTC Kappelmeyer measurements have been detected as a change in the spatial time series (values for some regions were much higher or lower in comparison to other sections of the cable).

### 4.5. Boston dike sensor faults detection

A description of this levee is presented in Sub-Section 2.3.2.

#### 4.5.1. Introduction to the Boston levee

This dike (Figure 4.18) was equipped with different types of sensors, including pore pressure sensors. Pore pressure sensors show periodic behaviour due to the sea tides and seasonal changes.

![Figure 4.18. (a) Cross-section of the Boston dike. The circles with numbers show the installation of the pore pressure sensors. (b) Photo of the Boston dike.](image)

The selected models (see Appendix C.5 and C.6) require stable measurement rates without gaps. For model construction, the periods without gaps must be selected or a gap filling procedure must be applied. Thus, a gap filling procedure was applied (see Figure 4.19) to the pore pressure measurements that were obtained from sensor #501 (Figure 4.18(a)). The pre-processed data can be used as input parameters for linear or non-linear models [102].

![Figure 4.19. Example of gap filling for pore pressure sensor #501. The "reconstructed" signal is shown in red (inside the box). Measurements were made every 15 minutes.](image)
4.5.2. Input sensors selection

The data from the pore pressure sensor were selected for modelling because of the pore pressure can be used to detect different dike failure modes (e.g., internal erosion or piping). Pore pressure sensor #506 (see Figure 4.18(a)) was selected for the output signal modelling. This sensor showed abnormal behaviour (see Figure 4.20) that began in January 2011. During the same observation period, sensor #501 behaved normally (see Figure 4.21). The red colours in Figure 4.20 and Figure 4.21 that are marked with boxes show gaps that were filled.

![Figure 4.20. Pore pressure measurements (mbar) from sensor #506. The blue line represents the raw data and the red line represents the gaps that were filled in. Extended periods where the gaps were filled are denoted by boxes.](image)

![Figure 4.21. Pore pressure measurements (mbar) from sensor #501. The blue line represents the raw data and the red line represents the gaps that were filled. Extended periods where the gaps were filled are denoted by boxes.](image)

4.5.3. Results from applying the linear transfer function model

Two methods are used for calculating the model output. In the first method, the real sensor output values are used $y(t-1), y(t-2), ..., y(t-n_a)$ (one-step forecasting). This model is the first
linear model. In the second method, only the previous model outputs are used, without the real sensor outputs.

The first model provides a more accurate estimation of the output according to the RMSE (see Equation (C.19)) and R² (see Equation (C.23)) values for the training period (see Table 2). However, this model cannot detect trend changes, with RMSE and R² values that are the same during periods with and without anomalies (see also Figure 4.22(a)). In this case, the modelled values and the real sensor measurements corresponded throughout the observation period. However, the linear model is able to detect abrupt faults (Figure 4.22(b)).

Table 2. Criteria of the first linear model quality for 3 periods, the training period, the test set with only normal data only, and the test set with anomalies.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Training set</th>
<th>“Normal” set</th>
<th>“Abnormal” set</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.9209</td>
<td>0.9054</td>
<td>0.9729</td>
</tr>
<tr>
<td>R²</td>
<td>0.9989</td>
<td>0.9986</td>
<td>0.9988</td>
</tr>
</tbody>
</table>

The second approach provides a less accurate estimation of the actual output, but allows for the detection of abrupt changes and changing trends (Figure 4.23). Divergence between the real transfer function and the model can be detected by using the RMSE (high
value during a period of abnormal behaviour) or $R^2$ (very low value during abnormal behaviour) (see Table 3). However, this approach is difficult to use for model estimation.

![Graph](image_url)

Figure 4.23. (a) Results of one-step forecasting using the second linear model. (b) The errors of the one-step forecasting.

Table 3. Criteria for the quality of the second linear model for 3 periods, the training period, the test set with only normal data, and the test set with anomalies.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Training set</th>
<th>&quot;Normal&quot; set</th>
<th>&quot;Abnormal&quot; set</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>9.4751</td>
<td>11.8003</td>
<td>27.9838</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.8182</td>
<td>0.7518</td>
<td>-0.2913</td>
</tr>
</tbody>
</table>

4.5.4. Results from applying the non-linear transfer function model

The results from applying the FFNN model (see Appendix C.6) to detect anomalies are presented in Figure 4.24. The error of the model increases significantly in the test sets with anomalies.
An analysis of Table 4 shows that the quality of the model that was based on the training set and test set with normal behaviour is good.

As a result of the exhaustive search for the model parameters described in Equation (C.22) (see Appendix C.6), the following parameters were obtained: $k=50$ (the number of delayed input values) and $m=20$ (the number of delayed output values). The number of hidden layers is 2, and the number of neurons in the hidden layers is 10. The number of epochs is 300, and the learning rate is 0.01. The filtering pre-processing procedure included the use of a Hodrick-Prescott filter [54].

Table 4. Criteria of the feed-forward neural network (FFNN) quality for 3 periods, the training period, the test set with only normal data, and the test set with anomalies.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Training set</th>
<th>“Normal” set</th>
<th>“Abnormal” set</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>1.1057</td>
<td>2.8746</td>
<td>14.0712</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9970</td>
<td>0.9599</td>
<td>0.4147</td>
</tr>
</tbody>
</table>
4.5.5. Results and conclusions

In this section, the identification of anomalies in sensor measurements by using a model-based approach with data-driven models is presented. This method detects anomalies in sensor measurements as deviation of the model outputs from the real sensor values (transfer function approach – see Section 3.7.2). This approach was applied for identifying anomalies in measurements that were gathered from sensor networks installed in the Boston dike (UK). Sensor faults were registered in measurements collected from pore pressure sensor #506.

A comparison of the non-linear model with the first linear model (which provides one-step forecasting using real sensor values) indicates that the accuracy of the linear model is higher than that of the non-linear model for the test set with normal behaviour (RMSE\text{lin1} = 0.9054, \text{RMSE}\text{FFNN} = 2.8746, R^2\text{lin1} = 0.9986, R^2\text{FFNN} = 0.9599). However, the linear model could not detect sensor faults (R^2\text{lin1} = 0.9988). In contrast, the FFNN was able to detect sensor faults (R^2\text{FFNN} = 0.4147). These results indicated that the model output and real values were significantly different.

A comparison of the first and second linear models (which only used the modelled output for time series forecasting) indicates that the second model is less accurate (R^2 = 0.8182). However, this model can detect any type of anomaly, including abrupt and trend changes. The first linear model is only able to detect anomalies such as abrupt changes and is not suitable for detecting anomalies in dikes. In contrast, the second linear model can be used to identify any type of anomaly (abrupt or trend changes) and only requires the input data for the testing period (the estimated values are used rather than the real output values). However, it is difficult to develop a stable and accurate model using this method. The construction of the non-linear model requires more effort than the construction of the linear model. However, the non-linear model can be used to identify anomalies involving abrupt and trend changes.

The identification of anomalies in sensor measurements is important for early operational flood warning systems. In addition, a more challenging problem is that of distinguishing between the onset of dike failure and sensor fault. To solve this problem, expert knowledge or an expert model is required. However, the cross-validation of one-side classification and model-based approaches is another possible solution. Finally, for the automatic selection of model parameters, the special optimization procedures should be applied.

4.6. Conclusions

This chapter covers the third scientific objective (validation of the investigated anomaly detection approach on real-world objects.). The applications of two data-driven approaches for anomaly detection are presented. The one-side classification approach (see Sub-Section 3.7.1) was applied in the analysis of the Stammer dike (non-destructive macro-stability experiment), Zeeland dike (non-destructive piping experiment), Rhine dike (special sensor behaviour) and retaining dam (real leakage). Possible data analysis schemes were described in Sub-Sections 3.8.1 and 3.8.2. Maximum overlap discrete wavelet transform (MODWT) showed the most promising results in comparison with Short Time Fourier Transform and Continuous Wavelet Transform on stage of feature extraction.
It can be concluded that variant described in Sub-Section 3.8.2 (combination of MODWT with one-side classification) should be considered as the promising anomaly detection scheme. Operability of this cascade was successfully confirmed by detection of especial sensor behaviour (Rhine dike) and real leakage (retaining dam). Rapid changes detected in Zeeland dike using time-frequency methods can be detected in the similar manner for the Stammer dike (see Appendix B).

The developed one-side classification anomaly detection approach is visually presented in Figure 4.25. Different types of sensors can be used by a set of one-side classifiers (considering the required pre-processing and feature extraction procedures) for evaluation of a levee’s state.

A modelling-based (transfer function – see Sub-Section 3.7.2) anomaly approach was applied to the data analysis of the Boston levee (sensor faults) data analysis. Detailed description of this approach is presented in Sub-Section 3.8.3. An analysis of the transfer function approach (linear and non-linear modelling, phase shift approach [101]) indicates that this branch of methods may also be applied to the detection of anomalies. The most challenging task in this case is the construction of an appropriate model.

The one-side classification based approach is more suitable for monitoring of the levees with not significantly changing external conditions. In this case all the monitored parameters are more or less stable. Detection of rapid changes or even simple deviation can be efficiently done using one-side classification approach (in combination with required pre-processing and feature extraction methods). Transfer function approach application is efficient if the external conditions are changing significantly (e.g., Boston levee). In this case the stability of relationship between the sensors should be monitored.

Most of the monitored within the UrbanFlood project levees had more or less stable external conditions (stable water level): in case of Stammer dike the water level is controlled by the local authorities; in case of Zeeland dike the water level inside the levee is controlled using a pump. Only really extreme conditions can be interpreted as outliers for the Rhine levee (e.g., during high water level due to melting snow). Taking into account all the above mentioned items the one-side classification approach was selected for further implementation in the Artificial Intelligence (AI) component and integration into the UrbanFlood EWS (see Chapter 5).