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

Pyayt, A.L.

Citation for published version (APA):
Appendix B. Additional results of Stammer dike data analysis

Results of additional feature extraction are presented in this chapter.

Figure B.1. SD_1_C2_540 (Y1) sensor measuring inclination (degrees).

B.1. Short-time Fourier transform

Parameters of simulation:
- Window size: 1024 samples.
- Shift of sliding window in each iteration: 24 samples.

The short-time Fourier transform (STFT) represents a signal in the time and frequency domains and, unlike FFT, local signal behaviour can be analysed along time. In Figure B.2, STFT is represented in the form of a spectrogram; only the amplitudes of STFT are analysed. The component that corresponds to 0 Hz was removed for better visualization.

Anomalies in dike behaviour are represented as three intervals with shifted means. The amplitudes of the FFT coefficients near intervals with abnormal behaviour are presented in Figure B.3.
Appendix B Additional results of Stammer dike data analysis

This type of signal behaviour (Figure B.3) can be detected as a rapid increase of FFT coefficient amplitudes during the experiments with heavy loads placed at the dike. This analysis can be performed by monitoring low frequencies; high frequencies can also be applied to identify rapid changes in the mean value. This simplifies application of STFT analysis for automatic detection of anomalies in dike measurements.
Stability analysis can be performed by constructing confidence intervals or by applying more-sophisticated algorithms such as a one-sided classification approach (neural clouds).

Analysis of Figure B.3 indicates that there is a delay in the identification of new behaviour. Thus, STFT can be applied for detection of long-term deviations in time series with a delay proportional to the size of the sliding window.

### B.2. CWT and MODWT components analysis

The continuous wavelet transform (CWT) represents the signal in the time-frequency domain. It has better time-frequency resolution than STFT and provides more-accurate feature extraction than STFT.

Parameters of the simulation:
- wavelet type: ‘Morlet’;
- scales of decomposition: 1:4:256 (each 4th scale is presented; in total, there are 256 levels of decomposition).

Results of the application of CWT to the selected sensor are presented in Figure B.4. The results have better time-frequency resolution than those of STFT and provide more-accurate feature extraction than STFT.

The base frequencies are extracted by calculating the variance for each scale and selecting the scale with a local maximum in the variance. In the current results, no significant base frequencies were identified.

![Figure B.4](image) (a) Sensor measurements. (b) Continuous wavelet transform (CWT) spectrum and variance of scales.
A zoomed variant of Figure B.4 is presented in Figure B.5. It provides more information about the presentation of anomalies generated during the experiment in which a heavy load was placed on the dike.

Figure B.5(b) indicates that the heavy load could be detected as a rapid change in the local mean of the signal. Such features affect small-scale (high-frequency, e.g., from 1 to 50) wavelet coefficients.

Figure B.5. (a) Sensor measurements. (b) Continuous wavelet transform (CWT) spectrum near the anomaly.

The maximum overlap discrete wavelet transform (MODWT) is another technique that analyses a signal in the time-frequency domain. Compared with CWT, it requires lower computational costs because a smaller number of scales are considered. Here, each level of decomposition is a power of 2 (the scales are numbers such as 2, 4, 8, etc.); such a representation affects the frequency resolution of the input signal. With decreasing frequency (higher number of the decomposition level), the frequency resolution is decreased, but the transform still obtains a good time-frequency representation of the signal.

Parameters of the simulation:
- Wavelet type: ‘least asymmetric 8’ (3 zero moments);
- Levels of decomposition: 1:1:13 (scales $2^1$, $2^2$, ..., $2^{13}$).

Application of MODWT technique does not extract any base frequencies. The spectrum obtained by applying the MODWT looks similar to the CWT spectrum. The high oscillations at the boundaries in Figure B.6(b) are caused by boundary effects in the wavelet decomposition.
For detailed analysis, another presentation of the coefficients calculated using MODWT is presented in Figure B.7. The top plot presents the signal decomposed by MODWT; the levels of decomposition are presented from the 1\textsuperscript{st} level (the top row in the figure) to the approximation of the 13\textsuperscript{th} level (the lowest row in the figure). Red horizontal lines in all plots of MODWT coefficients denote the universal threshold; if a value is above the threshold, it can be interpreted as an outlier or abnormal behaviour for this timestamp.

Heavy load periods are represented by three green boxes (from the 1\textsuperscript{st} to 7\textsuperscript{th} levels) and three red boxes (from the 8\textsuperscript{th} to 10\textsuperscript{th} levels). The analysis indicates that this anomaly can be detected as an increase in the modules or variance of coefficients of the 8\textsuperscript{th} to 10\textsuperscript{th} levels (red boxes) of decomposition.

Heavy load intervals can be detected as rapid changes in the local mean of a signal. In small (high-frequency) levels (from the 1\textsuperscript{st} to 7\textsuperscript{th}) – the green boxes in Figure B.7 - such features affect the small-scale wavelet coefficients.

Application of a universal threshold (see C.4) shows good results in identification of such type of signal behaviour. Universal threshold can be replaced by application of one-side classification approach, e.g., Neural Clouds.
Appendix B Additional results of Stammer dike data analysis

Figure B.7. The top figure: sensor measurements near dates with abnormal behaviour. All other plots show the maximum overlap discrete wavelet transform (MODWT) coefficients per level (blue colour), from the top (1st level) to the bottom (approximation of the 13th level trend) with their own universal threshold (red colour).