



UvA-DARE (Digital Academic Repository)

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

Pyayt, A.L.

Publication date

2014

[Link to publication](#)

Citation for published version (APA):

Pyayt, A. L. (2014). *Data-driven methods in application to flood defence systems monitoring and analysis*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 3. Data Analysis and Anomaly Detection Approach*

This chapter describes the main stages of sensor measurements analysis and illustrates problems using real-world data. Two anomaly detection approaches are presented in this chapter.

3.1. State-of-the-art

3.1.1. Application of data-driven methods for monitoring tasks

Data-driven methods can be classified in different manners depending on the domain (Figure 3.1): frequency, time-frequency or time domain.

These methods present an input signal in one of the domains for further processing. An analysis of related works indicates that all three groups are used for tasks of dike behaviour analysis, structural health monitoring (SHM), and environmental monitoring. For example, the fast Fourier transform (FFT) [37] provides extraction of *frequency* characteristics from the signal.

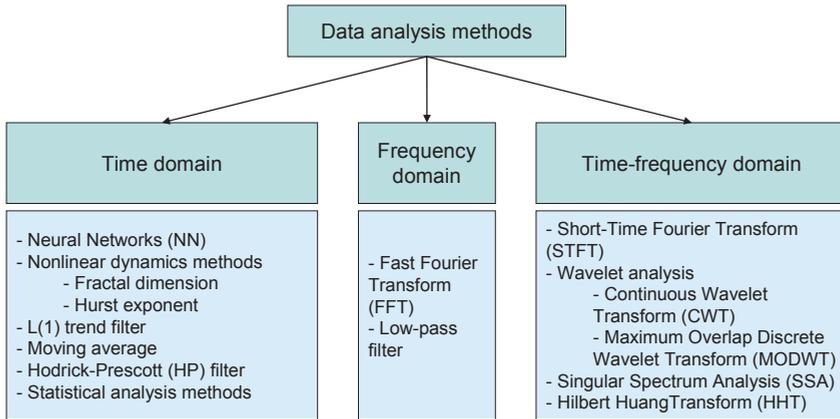


Figure 3.1. Classification of data-driven methods [99].

* Parts of this chapter were published in:

1. Pyayt, A.L., Kozionov, A.P., Kuserbaeva, V.T., Mokhov, I.I., Krzhizhanovskaya, V.V., Broekhuijsen, J., Meijer, R.J., Sloot, P.M.A. Signal analysis and anomaly detection for flood early warning systems, Journal of Hydroinformatics, IWA publishing, 2014, doi: 10.2166/hydro.2014.067 (in press).
2. Pyayt, A.L., Mokhov, I.I., Kozionov, A., Kuserbaeva, V., Melnikova, N.B., Krzhizhanovskaya, V.V., Meijer, R.J. Artificial intelligence methods and finite element model application for on-line dike measurements processing. Proceedings of the 2011 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS), Milan, Italy, 28 September 2011, pp 1-7, DOI: 10.1109/EESMS.2011.6067047.
3. Pyayt, A.L., Mokhov, I.I., Kozionov, A.P., Kuserbaeva, V.T., Lang, B., Krzhizhanovskaya, V.V., Meijer, R.J. Data-driven modelling for flood defence structure analysis. Comprehensive Flood Risk Management – Klijn & Schweckendiek (eds). 2013 Taylor & Francis Group, London, ISBN 978-0-415-62144-1. Proceedings of the 2nd European Conference on FLOODrisk Management. 20-22 November 2012, Rotterdam, The Netherlands, pp. 301-306.

In [113], FFT was applied for analysis of the horizontal acceleration simulated in earthquake modelling of earthen dam behaviour. Application of FFT for sliding window analysis (Short-Time Fourier Transform – STFT [88]) provides a more advanced analysis of time series in the *time-frequency* domain. STFT-based anomaly detection can be used for identification of anomalies, such as the increase of the amplitude of some frequencies in time; high deviations of amplitudes at base frequencies that differ significantly from deviation measured for a “normal” mode; and rapid changes of trend (outliers), which are usually exhibited as a rapid increase and subsequent decrease of amplitudes at high frequencies. Abnormal behaviour detection based on tracking the STFT complex coefficients of amplitudes in time is presented in [48] and [111]. Tracking of frequency amplitudes can be based on a simple comparison of FFT coefficients with a threshold or the more advanced one-side classification approach.

Other time-frequency techniques applied for dam behaviour analysis are based on wavelet analysis: continuous wavelet analysis (CWT) [78] and maximum overlap discrete wavelet transform (MODWT) [93]. In comparison with STFT wavelets have better time-frequency resolution and provide more accurate features extraction [78]. Wavelet analysis techniques address with non-stationary time series and can be applied in the same manner for anomaly detection as it was described for STFT.

In [94], wavelets were applied for analysis of water temperature measurements from the Wivenhoe Dam. Each signal was decomposed using wavelets into daily, sub-annual and annual (DSA) components. Each of the components was used for further analysis.

An overview of possible applications of wavelet transforms for structural health monitoring (SHM) tasks is presented in [108]. For example, wavelets were applied for crack detection at an arch concrete dam in [77]. In [51], an analysis of the dynamic behaviour of geotechnical structures for earthquake monitoring was performed by applying wavelets. Wavelets are also successfully applied for earthen dam behaviour analysis. In [89], numerical analysis of the seismic behaviour of an earthen dam was performed. One of the results of that work was that plasticity should be considered in the analysis of the seismic response of the dam because it leads to a decrease in the natural frequencies of the dam.

The data de-noising and data compression abilities of the discrete wavelet transformation (DWT) are used in [109] for health monitoring of tendons and cables. As a result, the defect detection sensitivity of the whole health assessment system was improved. An extensive overview of the application of MODWT and CWT for analysis of geophysical time series is presented in [95]. Wavelet variance analysis was applied to averaged ice thickness analysis.

CWT, like the STFT technique for anomaly detection, is based on tracking changes in coefficients over time by scales and across scales [127], [28]. Coefficients tracking can be based on simple comparison with thresholds [124] or more advanced one-side classification.

Empirical mode decomposition (EMD) [59] is one the methods in the frequency domain that can process non-linear and non-stationary data. Time-frequency analysis is possible by combining EMD with the Hilbert transform in the so-called Hilbert Huang Transform (HHT) [52]. An application of this approach is presented in [107], in which

HHT analysis was applied to environmental data analysis: rainfall, temperature, wind and stream flow analysis. This approach was applied to SHM tasks in [98].

For component analysis singular spectrum analysis (SSA) [50] is also used. This method decomposes a signal into components for further processing, such as filtering and forecasting. An application and analysis of the climatic time series method is presented in [49]. An analysis of short, noisy chaotic signals using the SSA method was performed in [120].

Chaotic time series can also be processed in the time domain. Non-linear dynamics methods are usually applied [56] for qualitative and quantitative assessments of time series behaviour. The short-term predictability feature of the chaotic systems and phase space characterisation was used in [41] to make one-hour to one-day predictions of ozone levels. In [42], the embedding space and fractal dimension concepts were applied to monitor the condition of a system with clearance. Applications of the correlation dimension metric for vibration fault diagnosis of a rolling element bearing are presented in [74] and [75].

Regression analysis was applied to climatic data analysis in [87].

There are a lot of methods applied in area of condition and environmental monitoring. Selection of the specific method depends on complexity of the monitored object and on the specific task to be solved. Most of these methods are efficient as one of the stages of processing (e.g., pre-processing or feature extraction) but are hardly to be used for generation of alarms by themselves. Post-processing is required in most of the cases.

3.1.2. Anomaly/fault detection methods overview

The general task to be solved was formulated in the first chapter as “anomaly detection in levee behaviour”. An anomaly can be a sign of a developing levee/dam failure. Anomalies can be encountered in measurements of real-world unstable dikes and can be simulated using numerical models [81]. There are several approaches in the literature for classification of anomaly/fault detection methods. The first one can be found in work [61]. According to the author of that work faults (or errors) might be interpreted as “deviations from normal process behaviour. They may result in some shorter or longer time periods with malfunctions or failures if no counteractions are taken”.

In Figure 3.2, a version of the classification of fault detecting methods from [61] is presented; we have adapted the classification to anomaly detection on the basis of various conditions of one or more sensor data streams. Some of the methods in Figure 3.2 require the exact pattern of anomaly (the “limit checking” and “trend checking” blocks), whereas another method requires construction of the model of the process (the “process models used” block) based on data from a dike in an acceptable state. Some anomaly detection models correlate the sensor data streams with others, thereby yielding sets of correlation coefficients.

“Detection with multiple signals and models” performs anomaly detection using the relationships of two or more data streams. This class also includes methods that generate multiple data streams by performing an operation on a single data stream (e.g., coefficients of wavelet decomposition from several levels of decomposition).

An application of the methods in the “Process models used” block, e.g., neural networks, state observers and parity equations, for dike measurements analysis is presented in [102].

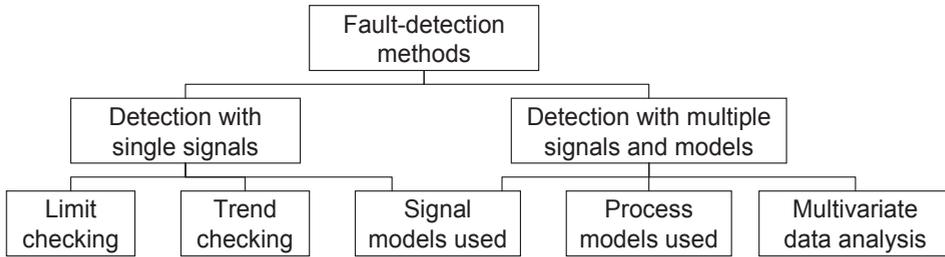


Figure 3.2. Survey of fault-detection methods based on [61]. In the case of dike conditioning monitoring anomalies hint at (the onset of) faults.

Overviews of anomaly detection algorithms and their applications in various domains can be found in [73] and [40]. These works helped to construct Figure 3.3, which details the “multivariate data analysis” block in the survey presented in Figure 3.2.

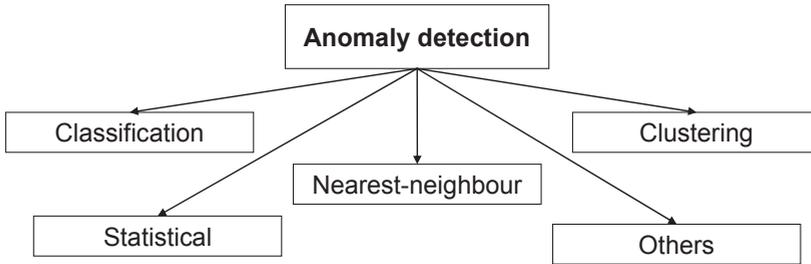


Figure 3.3. Taxonomy of multivariate anomaly detection algorithms.

Methods from the *classification* group (Figure 3.3) require labelled data, which must be classified as normal or abnormal. The group contains one-class SVMs (support vector machines), one-class kernel Fisher discriminants, and one-class replicator neural networks.

Methods that belong to the *nearest neighbour* group determine the distance of a data point to its k -th nearest neighbour and or the relative density of each data instance. The nearest-neighbour methods imply that normal data instances occur in dense neighbourhoods, whereas anomalies occur far from their closest neighbours. One of the advantages of nearest-neighbour methods is that they do not use any assumptions regarding the generative distribution of the data. However, the computational cost of such methods is high, and the reliability of the results in the case of unsupervised training is problematic [40].

The *clustering* approach assumes that normal data points have the following properties [40]:

- They belong to a cluster of data, whereas anomalies do not belong to any cluster.
- They lie close to their closest cluster centroid, whereas anomalies are far away from their closest cluster centroid.

- They are in large and dense clusters, whereas anomalies belong to small or sparse clusters.

The self-organizing maps (SOM) and K-means clustering anomaly detection approaches are based on the second assumption. The process of anomaly detection requires two steps: construction of clusters and calculation of the distance to the closest cluster centroid. One of the important advantages of the *clustering* methods is that they provide good results in the case of unsupervised learning.

Statistical methods assume that normal data instances occur in high-probability regions of a stochastic model, whereas anomalies occur in the low-probability regions of the stochastic model. Examples of such methods include Gaussian mixture models [22] and Parzen windows [90]). The key disadvantage of statistical techniques is that they rely on the assumption that the data are generated from a particular distribution. This assumption often does not hold, especially for high-dimensional real-world data sets [40].

The *other* group includes methods from information theory, spectral decomposition, visualisation-based methods and other methods.

The third approach of anomaly/fault detection methods classification is presented in [91], in which anomaly detection methods are classified as statistical methods, data mining-based methods, and machine learning-based techniques. Bayesian networks, principal component analysis, and Markov models are included in the machine-learning group. Data mining-based anomaly detection methods include classification-based methods, fuzzy logic techniques, genetic algorithms, neural networks, clustering, and outlier detection methods.

It must be emphasized that application of one-class anomaly detection algorithms is usually required because of a lack of measurements related to real-world collapses: there are not many full-scale levee collapse experiments (e.g., the IJkdijk Consortium experiments [14]), and physical modelling is more cost-effective but still requires a great deal of effort.

It is important to mention that algorithms of anomaly detection can be applied directly to the collected measurements but extraction of advanced features is required in most of the cases. In this case we can interpret this group of methods as a post-processing stage. The general data analysis scheme is presented in the next Section.

3.2. General data analysis scheme

The general approach for data analysis is presented in Figure 3.4. The first stage of data processing is the *data complexity evaluation*. This stage requires the collection of information about the quality of the gathered measurements, information about the rate of measurements (e.g., stability of measurements), the existence of jitters, and the existence of intervals with different measurement rates relative to the reconfiguration of the sensor network settings. Next, the time series is checked for gaps, which is required for evaluating the possibility of applying additional data processing methods. Equidistant measurements without gaps are required for applying most of the data processing methods. Depending on the evaluation results at this stage, the appropriate methods are selected at the 2nd stage, which is the pre-processing stage. Input data for this stage is “raw” if any of the above-mentioned problems are detected.

The *pre-processing* stage is one of the most important stages in the data analysis chain. This stage includes the data normalization procedures, the synchronization of measurements (required for multidimensional data analysis), and gap filling operations that are based on the statistics collected during the “data properties evaluation” stage. After these procedures, the collected “raw” data can be transformed into data that are suitable for analysis. Time series decomposition, including de-noising and de-trending operations, also is usually applied at this stage. For different tasks and methods, different types of pre-processing are required. The quality of the output data at this stage significantly influences the final results of the data analysis

The *feature extraction* block (the 3rd stage of data analysis) is strongly connected with pre-processing. For example, if a method is applied for gap filling at the pre-processing stage and decomposition is applied for this task time series, then these features (components) are already available for further analysis. In this study, the term “feature” can be interpreted as the result of any time series transformation.

The extracted features (and/or data) are used for further *post-processing* (4th stage). At this stage, different methods can be applied. Based on the calculated metrics, the decision is made by the expert in the 5th stage.

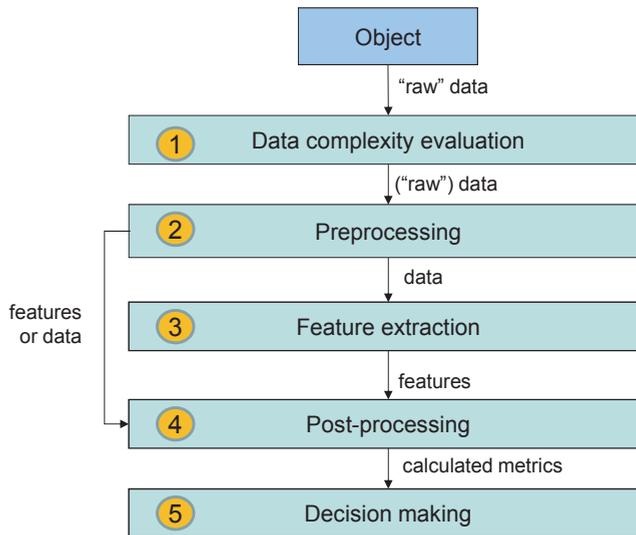


Figure 3.4. Data processing scheme

The evaluation of the data properties (Section 3.3) is described in this chapter with the pre-processing (Section 3.4), feature extraction (Section 3.5) and post-processing (Section 3.6) blocks.

3.3. Data complexity evaluation

Multidimensional sensor data are generally affected by different measurement sampling schemes (due to the different settings) and delays between the measurements (e.g., due to serial polling). In addition, gaps caused by a system disruption, a connection fault, or an individual sensor failure affect the multidimensional sensor data (Figure 3.5).

This affects make applying the data analysis methods across the entire time interval more difficult. The collection of this information is required for further pre-processing of the available data.

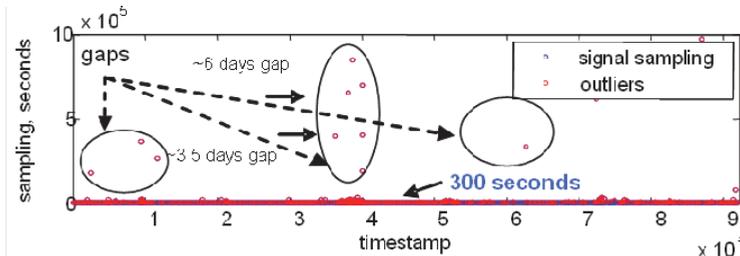


Figure 3.5. Illustration of problems when using multidimensional sampling with a sensor from a real dike monitoring system: sampling of the Zeeland dike (gaps are shown). The blue circles represent the expected base rate of the measurements (300 seconds) and the red circles represent the actual measurement rates. The X-axis corresponds to the sample number and the Y-axis corresponds to the difference between the neighbour samples in seconds [99].

The second possible task of this stage is the evaluation of the time series dynamical properties (Figure 3.6). During this stage, the qualitative properties of the time series are evaluated to select the appropriate data analysis for future stages. For this purpose, the nonlinear dynamics methods can be applied (e.g., Hurst exponent and fractal dimensionality). These methods can only be applied after the pre-processing procedure because they require an equidistant time series.

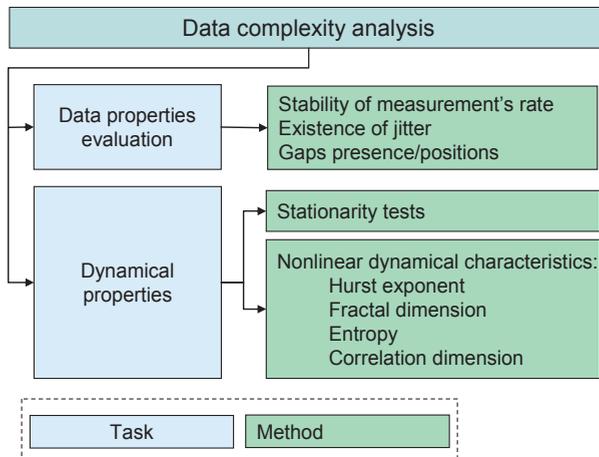


Figure 3.6. Data properties analysis scheme.

3.4. Data pre-processing

When the data complexity is evaluated, the next step is the data pre-processing step (Figure 3.7). This step can be split in three subtasks, the creation of the common time grid (CTG) and filtering and decomposition. Additional details regarding each of these blocks are presented below.

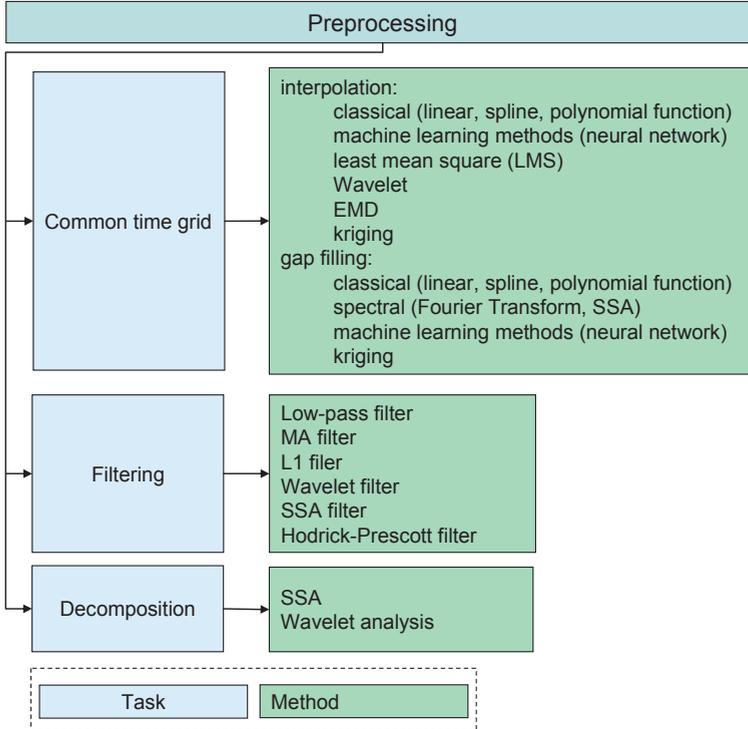


Figure 3.7. Data pre-processing scheme.

3.4.1. Common time grid

Most of the signal processing algorithms or machine learning algorithms cannot be applied to data with unstable sampling measurements. Thus, gaps in measurements can result in a loss of information.

These issues require a common time grid (CTG), where all sensor data are synchronized at the same moment in time (see Figure 3.8(a)).

The main steps of this algorithm include the following (Figure 3.8(b) [99]):

- 1) detection of the gaps in all of the sensor measurements according to the required rate in the CTG;
- 2) interpolation of the measurements on the CTG (e.g., linear interpolation, spline);
- 3) application of the gap-filling procedure; and
- 4) saving the synchronized measurements or the application of data analysis methods.

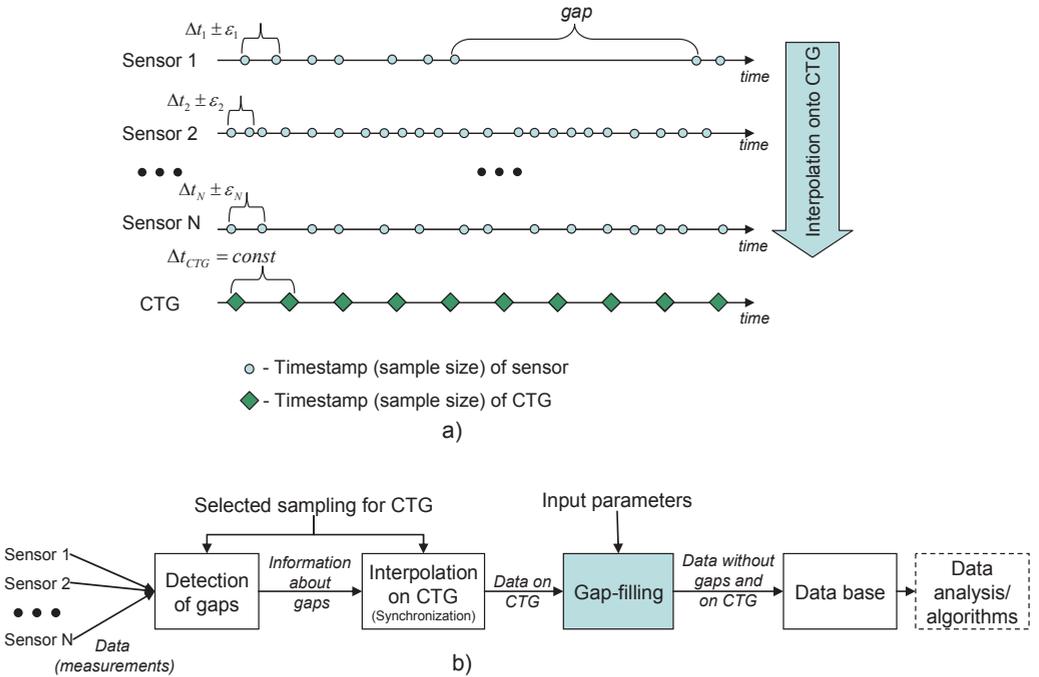


Figure 3.8. (a) Problems with multidimensional data interpretation for one object, process or system. Each sensor is associated with an array of measurements and an array of timestamps of the measurements: $\Delta t_{I..N} \pm \epsilon_{I..N}$ is a sampling of sensor $I \dots N$, where N is the number of sensor, $\epsilon_{I..N}$ is the sampling error (jitter), and Δt_{CTG} is the sampling of the common time grid. (b) Flowchart of the data pre-processing algorithm.

One of the most important steps of the algorithm is the gap filling procedure. Well-known techniques are available for data assimilation and are used in meteorology, including 4d-var and 3d-var [76]. For gap filling, both methods require a predefined state space model of measurements and are based on the optimal linear estimation of the process state. The gap filling methods are compared in [84]. Four methods were tested in this study, two versions of the Lomb–Scargle algorithm with different windowing schemes, the Kondrashov–Ghil method, and the smoothing spline method. Singular spectrum analysis (SSA) can be used to accurately and reliably reconstruct a wide range of time series [68]. In this study, additional experiments are not performed because the early experiments showed that this method performed well.

Singular spectrum analysis is a non-parametric spectrum estimation method, which allows the components of a signal (trend, periodic, noise) to be separated for further analysis. In addition, relative to the 3d-var and 4d-var approaches, the SSA-based gap filling does not require an estimation of the signal model. The SSA only deals with uniform sampled data. Linear interpolation was performed before application of the gap filling procedure.

The SSA gap filling steps are presented below (Figure 3.9).

- 1) Initialize a signal (interpolated on a common time grid) and a number of SSA components for gap filling.
- 2) Perform signal pre-processing by subtracting the mean from the signal (centring) and filling the gaps with zeroes.
- 3) Next, a SSA gap filling procedure is performed. In the outer loop, the next SSA component is added to the first reconstructed component (for a selected number of components). Then, SSA and inverse SSA in the inner loop are performed until the convergence criterion in the gaps is met and until the best estimate of the second reconstructed component is obtained. The convergence criterion is a ratio of the variance of the difference between the previous estimations and the present estimation. If this ratio is less than a specified level, then the next component is added and the outer loop is performed. This procedure is performed for a selected number of SSA components. This procedure is performed for a selected number of SSA components.

Gap filling results using the SSA method are presented in Figure 3.10.

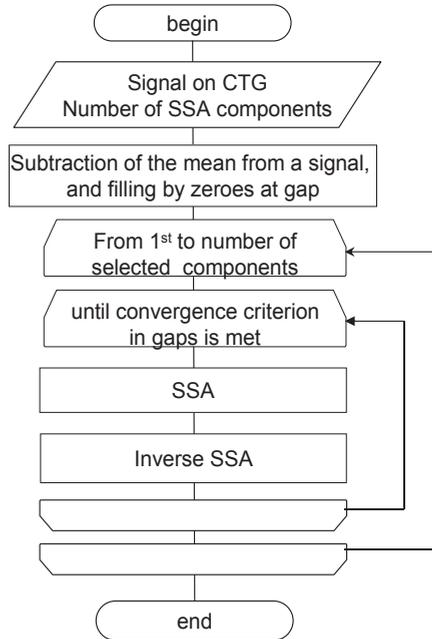


Figure 3.9. Flowchart of gap-filling with singular spectrum analysis (SSA).

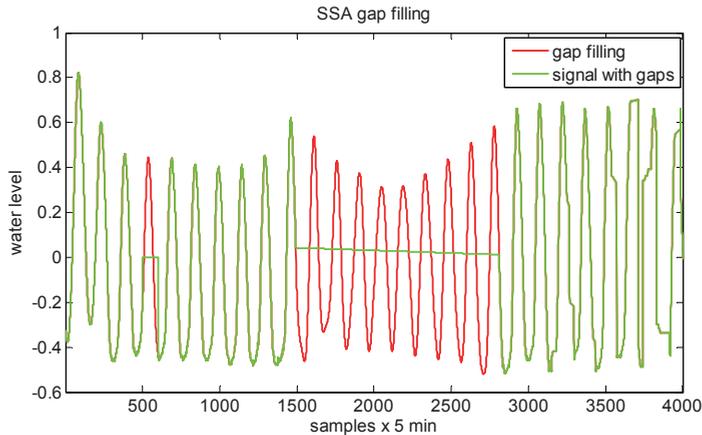


Figure 3.10. The results of the Singular Spectrum Analysis (SSA) based gap filling for the water level sensor. The green line is the signal after interpolation on the Common Time Grid (CTG) and the red line is the result obtained from gap-filling the signal. The X-axis represents the discrete time step number, with a time step of 5 minutes for 2 weeks of observations. The Y-axis represents the water level (meter).

3.4.2. Filtering

One of the main goals of data pre-processing is noise reduction. Signal characteristics (e.g., position, height and width of the peak at some interval of a time series) can be extracted from filtered signals with a better accuracy.

The following methods were used for data filtering: wavelet de-noising (Daubechies wavelet with 4 vanishing moments, a universal threshold and a hard threshold) ([78] and [45]), L1 smoothing [66], Hodrick-Prescott filtering (HP) [54], Singular Spectrum Analysis (SSA) with a window length 50 [46], and moving average (MA).

To test these methods, a synthetic sensor signal was generated (see Figure 3.11), which contained discontinues and periodic functions that were similar to those observed in the real sensor data, with three intervals with linear trends and sine waves of different frequencies (black colour). In addition, the white noise and outliers (green colour) [104].

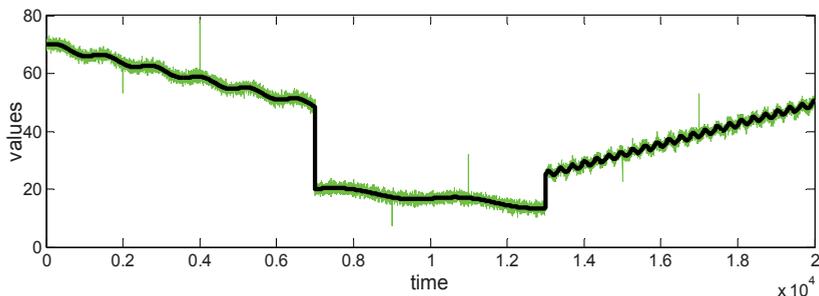


Figure 3.11. Synthesized sensor signal (black colour), added white noise and outliers (green colour) [104].

The results of the smoothing “jumps” and outliers are presented in Figure 3.12 and in Figure 3.13.

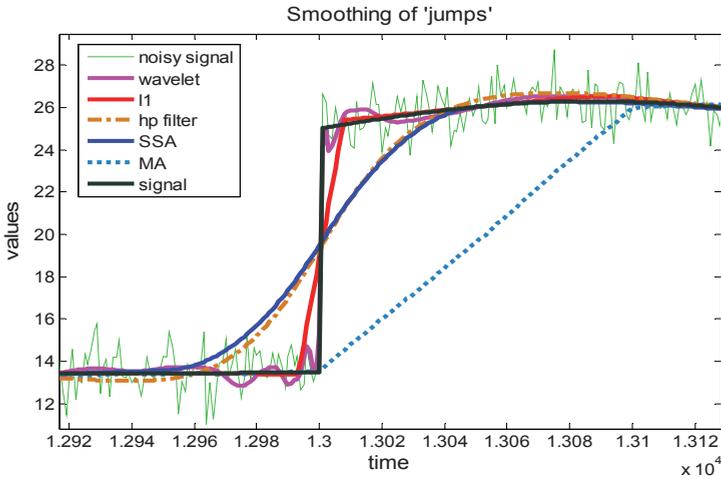


Figure 3.12. The results of the data “jumps” smoothing.

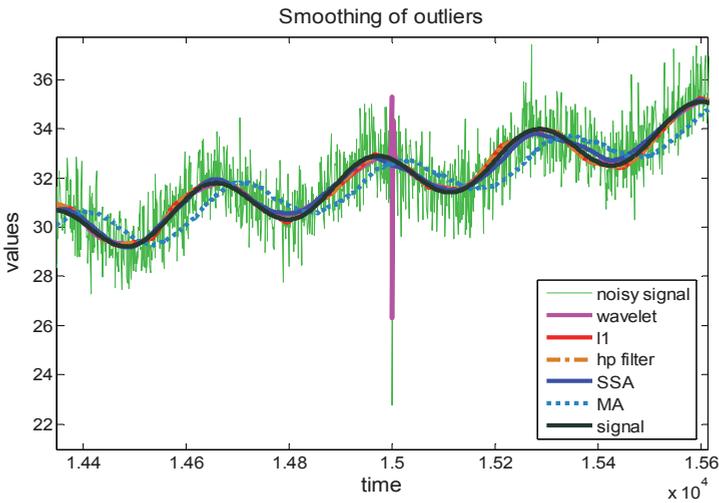


Figure 3.13. The results of outlier smoothing.

The root mean square error (RMSE) (see Equation (C.19)) values are shown in Table 1.

Table 1. Root mean square error for the different methods

	Wavelet	l1	hp	SSA	MA
RMSE	0.23	0.24	0.49	0.50	1.18

While the RMSE for the wavelets is better than the other methods and does not distort the data in the jump area, the method also includes the outliers with the nearest noise

into the smoothed signal. L1 smoothing is often sharp at the local extremes, but nearly retains the “jump” area. The HP, SSA and MA methods have smooth surfaces. In addition, the HP and SSA methods result in smooth jumps that remain visible. Furthermore, the MA method shifts the “jump” area of the signal.

For the data with complex behaviours, the best smoothing can be obtained by using wavelets, which performed well for processing data with jumps. To improve the performance of the wavelet in the presence of outliers, the algorithm parameters can be adjusted or other methods can be applied [104].

3.4.3. Time series decomposition

In general, a time series can be represented as a combination of the following components [25]:

- The trend component (denoted as T_t , where t stands for a particular point in time) is considered as an annual component that describes dike behaviour in general.
- The cyclical component (C_t) represents the day/night and high/low tide cycles that can be easily detected by the water-pressure sensors. This component explains the short-term deviation of the time series.
- The seasonal component (S_t) includes seasonal changes in the behaviour of the time series and long-term cycles, such as the moon cycle (approximately 30 days).
- The irregular component (I_t) or “noise” describes the random and irregular influences. This component includes the accuracy of the available measurements.

All of the components can be combined together to obtain: $X_t=f(T_t, C_t, S_t, I_t)$, where X_t is the observed value of the time series at time t . In addition, differences between cyclical and seasonal components can be described because the cyclical components do not have stable periods like the seasonal changes do [60].

There are two general approaches for presenting the time series as a composition of the components [57]:

- The additive model, $X_t=TC_t+S_t+I_t$, where TC_t is a combined trend-cyclical component.
- In addition, the multiplicative model is written as follows: $X_t=T_t*C_t*S_t*I_t$.

An additive model is appropriate if the magnitude of the seasonal fluctuations does not vary with the series level. The multiplicative model is more appropriate when the seasonal variation increases with time [60].

The time series decomposition is required for the following operations:

- separation of trends, harmonics and noise for further data analysis (MODWT, CWT, SSA); and
- separation of the daily-seasonal-annual (DSA) components to obtain a more detailed analysis of the external factors that influence dike behaviour.

We have applied DSA decomposition [94] (see Figure 3.15) based on the maximum overlap discrete wavelet transform (MODWT) [93] to Stammer dike (see Section 2.3) using air temperature data from the SD AIR (Y1) sensor (Figure 3.14). This sensor was selected because the temperature includes the cyclical (day/night) and seasonal components.

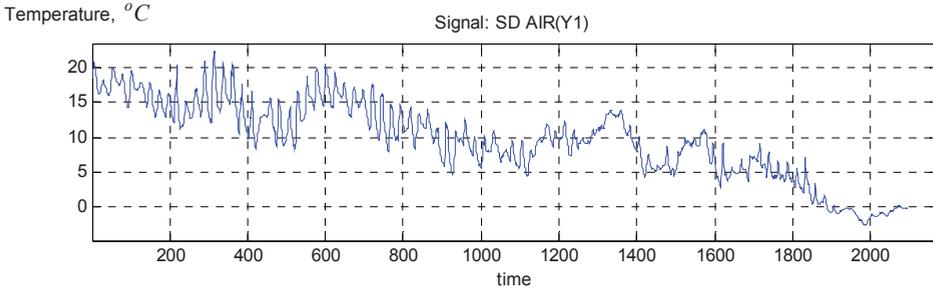


Figure 3.14. Air temperatures measured at Stammer dike using the SD AIR (Y1) sensor. The x-axis represents the discrete time step number, with a time step of 1 hour between the measurements for three months of observations (since 9th of September till 2nd of December).

The main goal of DSA is to analyse the decomposition of a time series using MODWT in the time-frequency domain and to add the levels related to the daily (in frequency diapason until 1/24 hours), seasonal/monthly (1/720 hours), and annual/sub-annual (1/2200 hours) components. This idea is presented in Figure 3.15.

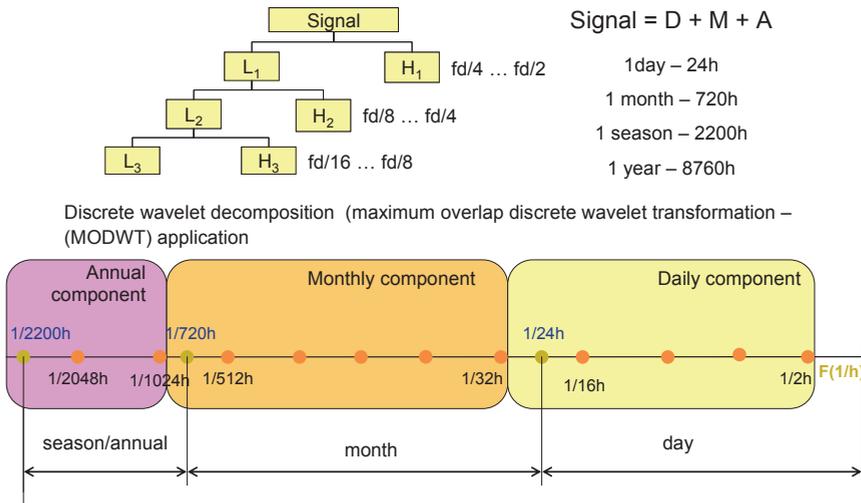


Figure 3.15. Daily-seasonal-annual (DSA) time series decomposition concept.

The results from using this approach application are presented in Figure 3.16. Each of the components can be independently used for further analysis independently. Figure 3.16(c) shows how the annual component is related to the air temperature signal. Please see Appendix C.3 for additional details

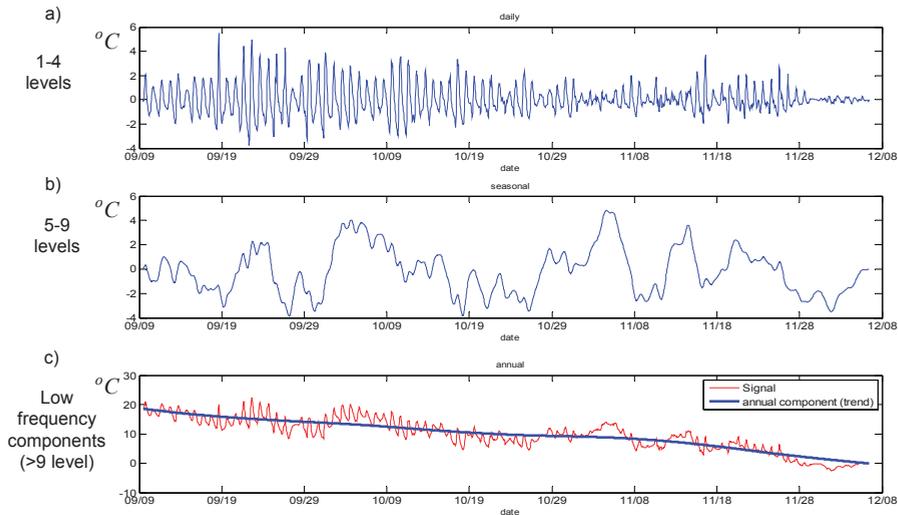


Figure 3.16. The daily-seasonal-annual (DSA) decomposition of the air temperature sensors at the Stammer dike location. **(a)** Sum of the 1-4 levels presented for the daily component. **(b)** Sum of the 5-9 levels presented for the seasonal component. **(c)** Sum of the other levels presents for the annual component. The blue line represents the annual component and the red line represents the initial time series.

3.5. Feature extraction

The term “Feature”, as it was previously described in Section 3.2 is a result of any time series transformation. Figure 3.17 presents our interpretation regarding the functionality of this block.

Statistical properties can be calculated for one/multidimensional input data, particularly for raw data or for data that were extracted during the pre-processing stage of the time series components.

Nonlinear dependencies between the variables can be checked using regression analysis. If model construction is sufficient for imputing environmental conditions, it should be possible to evaluate the dike behaviour by only using the data-driven methods.

The features that were extracted by applying various methods can be used for further tracking by using advanced techniques. Different levels of CWT or MODWT decomposition can be used independently for additional analysis.

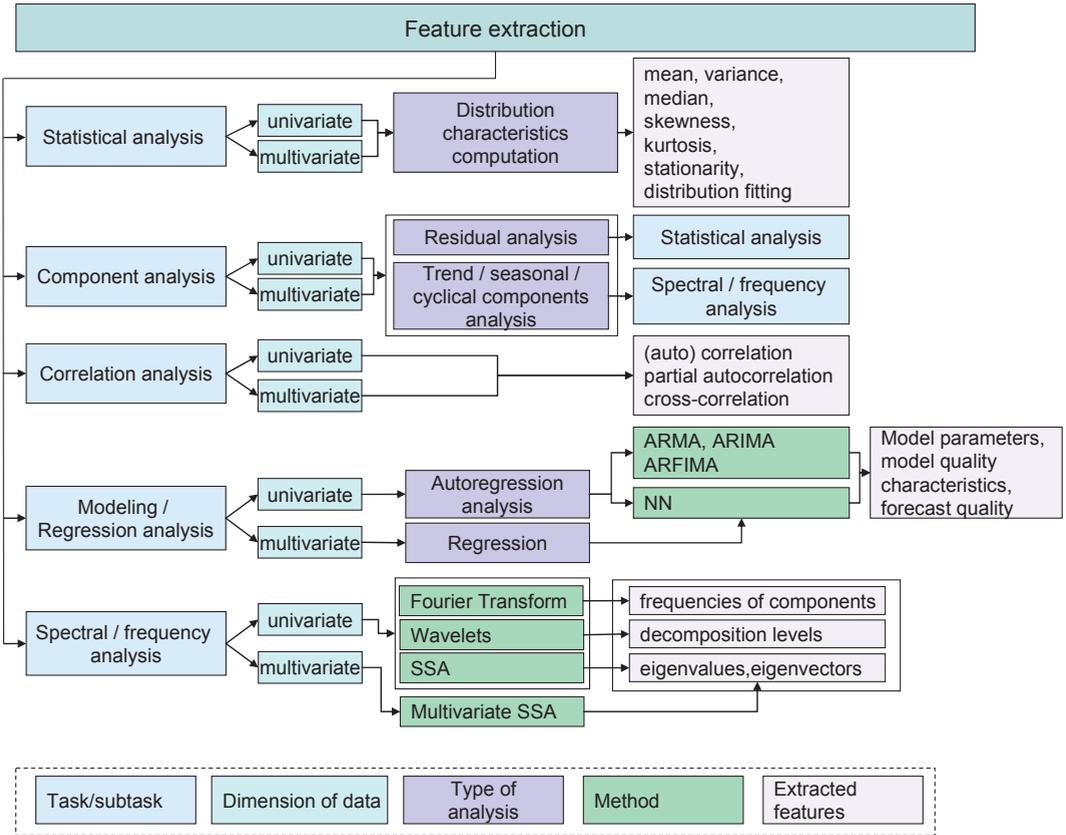


Figure 3.17. Feature extraction scheme.

3.6. Post-processing

3.6.1. Neural Clouds

In this study, we describe the Neural Clouds (NC) [72] method from the branch of Statistical and Clustering methods (see Figure 3.3). The NC belongs to the “multivariate data analysis” block (Figure 3.2). A comparison of this method with the Gaussian Mixture and Parzen windows methods is presented in [72].

The core of the NC classification agent (single classification algorithm) is a combination of the advanced k-means (AKM) clustering algorithm and the extended radial basis function (RBF) network [72] approaches. AKM is a modification of the well-known k-means algorithm with an adaptive calculation for the “optimal” number of clusters.

The AKM itself consists of the following steps.

- 1) Set an initial number of K centroids and the maximum and minimum bounds.
- 2) Use the k-means algorithm to position the K centroids.
- 3) Insert or erase the centroids according to the following premises
 - a) If the distance from a data point to the nearest centroid is larger than a certain threshold, a new centroid must be generated.

- b) If any cluster consists of less than a certain number of data, the corresponding centroid must be removed.
- c) If the distance between some centroids is smaller than a certain value, then the centroids must be clustered into one centroid.
- 4) Loop to step 2 unless a certain number of epochs is reached, or if the numbers of centroids and their coordinates have become stable.

The outputs of the AKM algorithm are centres of clusters that represent historical data related to the “normal” behaviour (training period). After all of the cluster centres have been extracted from the input data, the data are encapsulated with the hyper surface. For this purpose Gaussian distributions, or the so-called “Gaussian bells”, are used.

$$R_i(x) = e^{-\frac{|x-m_i|^2}{2\sigma^2}} \quad (1)$$

where m_i is the centre of the Gaussian bell, σ is the width of the Gaussian bell, and x is the input data.

This constructed mechanism could be described in the form of a radial basis function network, as shown in Figure 3.18. After such a network is manually trained by tuning parameters on the measured data, it can be used to estimate the confidence level for new measurements. The centres of the AKM clusters become the centres of the corresponding Gaussian bells (Figure 3.18– L1 part). The sum of the Gaussian bells is calculated to obtain the encapsulating surface. The sum of the Gaussian bells can equal unity when the bells overlap.

The confidence value is calculated using Equation (2).

$$P_c(x) = \frac{\sum_{i=1}^n R_i(x)}{\left(\sum_{i=1}^n R_i(x) + g_0 \right)} \quad (2)$$

Here $R_i(x)$ represents the Gaussian bells that are defined in Equation (1) and g_0 is a global factor (more details in [72]).

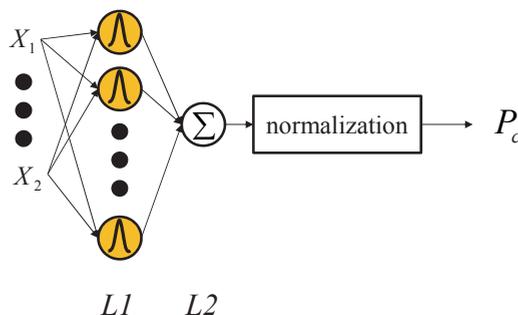


Figure 3.18. Radial Basis Function network representation. The L1 portion represents the Gaussian bells, the L2 part represents the superposition of the Gaussian bells, and P_c represents the confidence level.

The confidence values (P_c) that were calculated using Neural Clouds were between 0 and 1. The confidence value can be interpreted as the “membership function”. A new multidimensional point is checked if it is related to the clusters of the previously known

normal behaviour. Confidence values near 1 reflect normal behaviour, while values near 0 reflect anomalies. The final decision-making depends on expert knowledge or the expert model.

The NC encapsulates all of the previously known configurations of the selected parameters for a given training period (Figure 3.19). After training, the NC calculates a confidence value for each new state of the dike and describes the confidence value of normal behaviour.

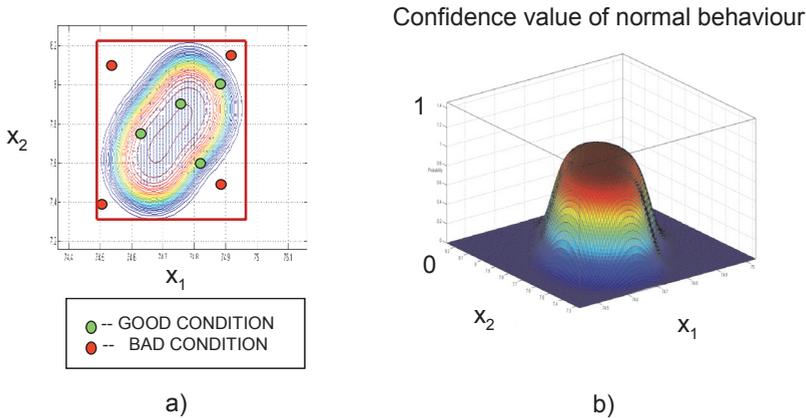


Figure 3.19. (a) Example of the application of Neural Clouds to 2-D data. The red points represent measurements that are incorrectly classified by the “box” approach. (b) A 3-D presentation of the confidence values of the normal behaviour of the object. Values near 1 are related to normal behaviour, while values near 0 are interpreted as anomalies.

3.7. Description of the anomaly detection concept

3.7.1. One-side classification approach

The basic idea of applying Neural Clouds for complex system monitoring is based on using a single classification instance for detecting the deviations of behaviour from the normal state of the entire system. We extend this approach by introducing a set of classification agents that are trained and used independently to spot anomalies.

Each one-sided classification instance uses raw or pre-processed data (Figure 3.20). Multidimensional sensor measurements can be grouped according to their *physical redundancy*. In this case, neighbouring sensors, sensors from the same cross-section or sensors measuring the same physical parameters can be combined using a one-side classifier. In addition, *analytical redundancy* between the sensors can be used to group measurements. When linear/nonlinear dependencies are detected for a historical period, the sensors will be grouped together. In this case, the stability of the detected dependency should be monitored [99].

Various data analysis methods can be used for feature extraction (Figure 3.20), which can be classified into *time*, *frequency* or *time-frequency* domains. Fast Fourier

transform (FFT) is related to the frequency domain. An improved version of FFT, short time Fourier transform (STFT), can be applied for extracting time-frequency features. In this study, we applied FFT, STFT, continuous wavelet transform (CWT), and maximum overlap discrete wavelet transform (MODWT) analyses at feature extraction stage.

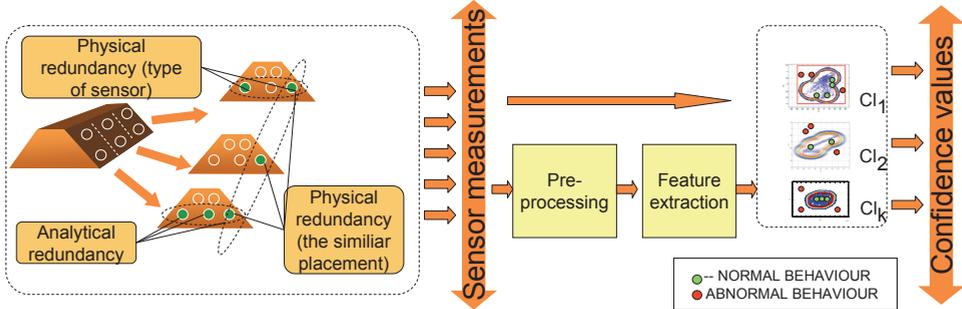


Figure 3.20. Abnormal behaviour detection methodology for the dike based on one-side classifiers.

3.7.2. Transfer function approach

The second anomaly detection approach is model-based. The benefit of this approach is that a high-accuracy model can be constructed that predicts the "normal" behaviour of a complete sensor system. The deviation of the sensor data from this "normal" model output can be detected easily. In addition, the development of a high-quality model is challenging.

The approach used in this work is called *transfer function (TF) modelling*. If an accurate model is constructed and the input signal behaves normally, the comparison of the model output with the sensor readings detects anomalies in the output real sensor (Figure 3.21(a)). The same approach can be used for sensor modules that measure several physical parameters at the same point (Figure 3.21(b)). In this case, the sensor measurements can be cross-validated [102].

The descriptions of the models that are applied for transfer function modelling are presented in Appendix C.5 and C.6. Both approaches are described in the next Section.

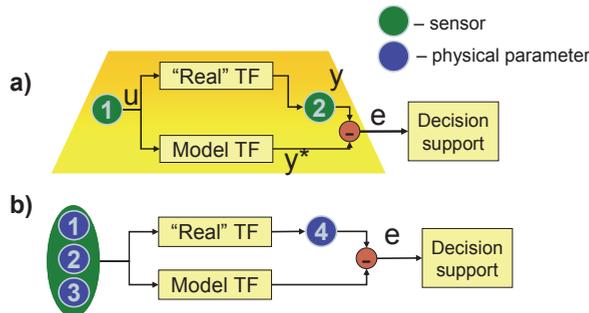


Figure 3.21. (a) Transfer function (TF) between the different sensors. (b) TF between the physical parameters in one sensor module. u is the input sensor signal, y is the output sensor signal, y^* is the model output, and e is the output error [102].

3.8. Description of the data processing

3.8.1. FFT and STFT components analysis

A data analysis scheme that uses fast Fourier transform (FFT) and short-time Fourier transform (STFT) is presented in Figure 3.22. The first standard operation is measurements synchronization. FFT transforms time series from the time domain to the frequency domain. All of the obtained coefficients are applied for further analysis (analysis of amplitudes). The FFT coefficients are calculated once for the whole period. For the obtained spectrum, the amplitudes of components can be checked, or the amplitudes of only the base components may be considered. Base frequency extraction can be performed by applying the relevant procedure or expert recommendation. For several sensors, pair-wise phase shift (see Appendix C.1) calculation procedures can be applied. These static metrics can give information about the delay between signals for each frequency of the spectrum for the whole period.

Tracking the dynamics of the amplitudes of components can be performed by applying STFT. In this case, a time-frequency representation of the signal can be obtained. Tracking in time the number of base components, distribution of base frequencies in the spectrum (e.g., the number of base components can be stable, but the frequencies of the components can still change in time), and amplitude of each component of the selected stable base frequency can be performed by applying an empirically based or expert-defined threshold (supervised tracking block) or by application of unsupervised methods (e.g., a one-side classifier approach). Phase-shift monitoring between pairs of sensors provides the dynamics of delays between arrays of measurements. These spatial-temporal features can also be used as input for one-side classifiers.

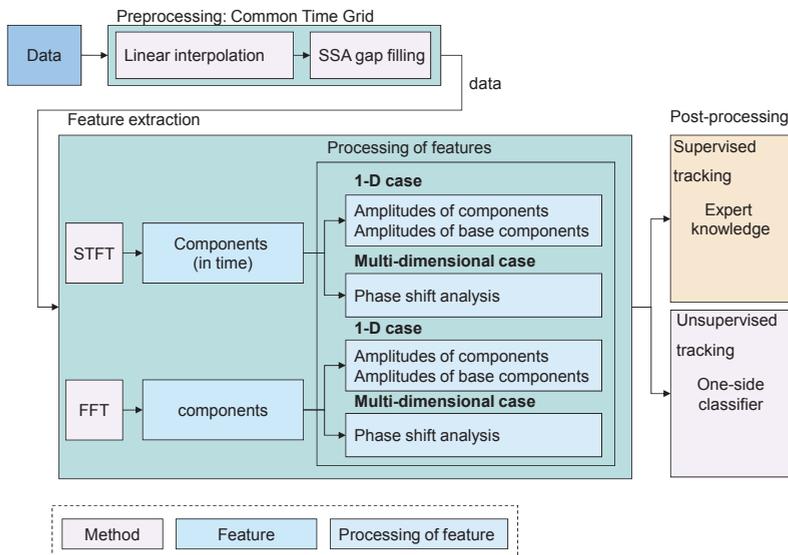


Figure 3.22. Short-time Fourier transform (STFT) and fast Fourier transform (FFT) component analysis scheme.

3.8.2. CWT and MODWT components analysis

A data analysis scheme that uses wavelet-based methods is presented in Figure 3.23. The scheme presented below involves 4 different data flows. The first data flow includes the denoised data. For this purpose, MODWT is applied. The DSA data flow refers to decomposition of the time series into daily, seasonal, and annual components, which are used for further analysis by other methods. The third and fourth data flows are levels of decompositions of CWT and MODWT.

Wavelet anomaly detection techniques are based on the analysis of variations in the wavelet coefficient magnitudes with time and across scales.

Changes in the CWT and MODWT coefficients with time can be tracked using a universal threshold (see Appendix C.4) or a one-side classifier (neural clouds). The tracking coefficients procedure is based on detection of increasing or decreasing wavelet magnitudes modules across scales. MODWT decomposition is used for the computation of DSA components (see Appendix C.3).

The extracted features are used for further analysis by a one-sided classifier or expert for decision-making (in all cases).

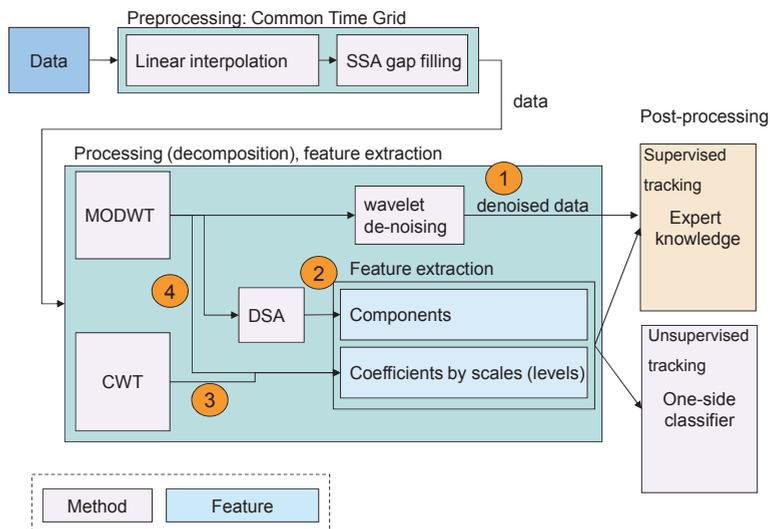


Figure 3.23. Continuous wavelet transform (CWT) and maximum overlap discrete wavelet transform (MODWT) components analysis scheme.

3.8.3. Transfer function analysis

A scheme of transfer function analysis is presented in Figure 3.24. After measurement synchronization is performed, a model in which one or several sensors are used as input and one or several sensors are used as output should be constructed.

The parameters of the linear regression model can be updated using a sliding window and used as input for the one-sided classification approach.

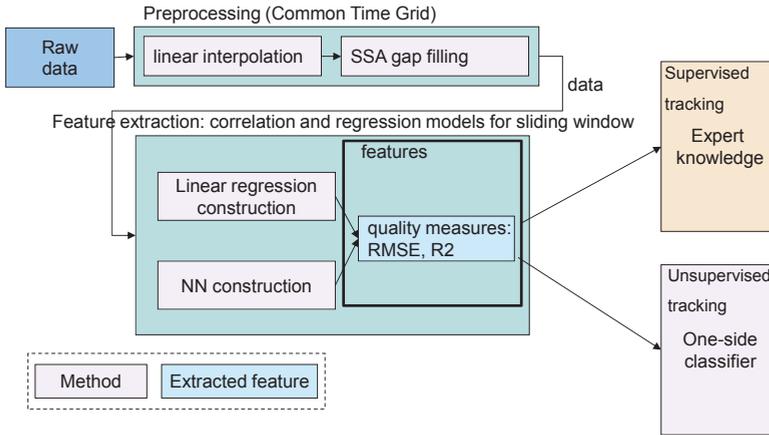


Figure 3.24. Transfer function analysis scheme.

3.9. Conclusions

State-of-the-art in application of the data-driven methods for the tasks of condition monitoring and environmental monitoring is described in this chapter.

The *general data analysis scheme* is presented (Section 3.2) in several stages. If gaps are detected and/or the data are asynchronous, the relevant data pre-processing procedures must be applied in the next stage. In this case, common time grid and gap filling procedures are used. Even if the raw data do not exhibit these problems, data filtering and time series decomposition must be applied in the second stage. An application of most of the data-driven methods without this pre-processing procedure is impossible. All the mentioned above items describe the *first scientific objective*: investigation of raw sensor data properties. In the third stage, the relevant features are selected. These can be features from time, frequency or time-frequency domain. The pre-processed data can be used directly in the post-processing steps.

The *second scientific objective* (investigation of different variants of construction of anomaly detection approach) is covered in this chapter (see Sections 3.7 and 3.8). Two different approaches can be applied at the post-processing stage: the one-side classification approach (e.g., Neural Clouds [72]) and the modelling approach (e.g., artificial neural networks). These approaches lead to different *anomaly detection* approaches. The first approach is based on the application of a set of one-side classifiers. The Neural Clouds approach is a clustering-based anomaly detection approach that is based on an updated version of the k-means algorithm (referred to as the advanced k-means (AKM) algorithm). Data that are related to the normal behaviour of the object are required for training this component. This requirement is important because the measurements that are related to the abnormal behaviour of the object are not generally available (levees are usually stable) or an expensive simulation of the levee collapse must be conducted using the physical modelling approach.

The trained instance of a one-side classifier is used next for the testing phase (operational phase when the stream of the on-line measurements is analysed). If the new

multidimensional point that was verified by the Neural Clouds does not belong to the clusters of the previously known normal behaviours, the calculated confidence value of the normal behaviour will be low (close to 0). Otherwise, a high confidence value is produced (close to 1).

The second approach requires the construction of a model using artificial neural network. The most complicated task is the construction of a qualitative model.

For both approaches, it is important that the data are pre-processed. During pre-processing, gaps should be excluded and the measurements should be synchronised.

The results of application of the described above data processing schemes are presented in the next chapter