

Research Article

Open Access

Esko K. Juuso*

Smart Adaptive Big Data Analysis with Advanced Deep Learning

<https://doi.org/10.1515/eng-2018-0043>

Received January 16, 2018; accepted September 24, 2018

Abstract: Increasing volumes of data, referred as big data, require massive scale and complex computing. Artificial intelligence, deep learning, internet of things and cloud computing are proposed for heterogeneous datasets in hierarchical analytics to manage with the volume, variety, velocity and value of the big data. These solutions are not sufficient in technical systems where measurements, waveform signals, spectral data, images and sparse performance indicators require specific methods for the feature extraction before interactions can be properly analysed. In practical applications, the data analysis, knowledge-based methodologies and optimization need to be combined. The solutions require compact calculation units which can be adaptively modified. The artificial intelligence is extended with various methodologies of computational intelligence. The advanced deep learning approach proposed in this paper uses generalized norms in feature generation, nonlinear scaling in developing compact indicators and linear interactions in model-based systems. The intelligent temporal analysis is available for all indices, including for stress, condition and quality indicators. The service and automation solutions combine these data-driven solutions with the domain expertise by using fuzzy logic for case-based systems. The applications are developed gradually in connections, conversion, cyber, cognition and configuration layers. The advanced methodology is based on the integration of features, scaling functions and interaction models specified by parameters. All the sub-systems and different combinations of them can be recursively updated and optimized with evolutionary computing. The systems adapt to the changing operating conditions and provide situation awareness for the risk analysis. The approach supports different levels of the smart adaptive systems.

Keywords: big data analysis, artificial intelligence, deep learning, statistical methods, computational intelligence,

recursive analysis, temporal analysis, fuzzy logic, smart adaptation

1 Introduction

Amounts of data are growing rapidly with increasing measurement possibilities and advancements of *Internet of Things (IoT)*. The large and complex datasets, which are challenging for commonly used data processing software and relational database management systems, are often referred as *Big Data*. The term has been used since the 1990s with focus on unstructured data. Currently, the same approaches are proposed for structured and semi-structured data [1]. *Cloud Computing*, which is a powerful technology to perform massive scale and complex computing, has many research challenges, including scalability, availability, data integrity, data transformation, data quality, data heterogeneity, privacy, legal and regulatory issues, and governance. Only few tools can address the issues of big data processing in cloud [2].

Challenges in the big data processing originate from four Vs [2]:

- *Volume* refers to the amount of all types of data generated from different sources and stored in growing digital storages.
- *Variety* refers to different types of data, including data sets, image, video, text and audio, in either structured or unstructured form.
- *Velocity* refers to the speed of data generated and the continual change of content.
- *Value* refers to the discovering huge hidden values from large datasets.

Big data analytics focuses on heterogeneous datasets which include structured, semistructured and unstructured data. Information in natural language is increas-

*Corresponding Author: Esko K. Juuso: Control Engineering, Faculty of Technology, FI-90014 University of Oulu, Finland, E-mail: esko.juuso@oulu.fi

ing. Data processing and discovering solutions from partly redundant, obsolete or trivial data may lead to non-consistent results: essential information might be missing or is confidential. There is some application potential but scale, diversity and complexity introduce problems of accessing, analysing, interpreting and applying.

Artificial intelligence mimic human cognitive functions, including reasoning, knowledge, planning, learning, natural language processing, perception and ability to move and manipulate objects. *Machine learning* is based on the iterative seeking of solutions by using new architectures, techniques and algorithms. *Neural networks* have been widely used in these studies as behavioural models to map system inputs to outputs regardless of the nature of the system. Connecting ANNs to other modelling techniques is vitally important as far as complex systems are concerned [3]. Different methodologies of computational intelligence have their strengths and drawbacks: the analysis of capabilities shown in Fig. 1.

Deep learning refers to layered hierarchical unsupervised learning and extraction of different levels of complex systems [4]. Global learning patterns and relationships are obtained from data without any human interface. Possibility to a more or less automatic modelling has increased popularity of ANNs in building nonlinear transformation layers for complicated interactions within different sources of varying data [5].

Genetic algorithms can assist other methods of computational intelligence by optimizing structures (Fig. 1). The family of evolutionary computation is extending with various ideas of nature inspired systems.

In technical systems, all four Vs should be taken into account, but the material is mostly structured. The volume of datasets is huge and is growing rapidly but the domain expertise is not easy to include in the cloud based solutions. Large volumes of data require compact solutions to different types of data. Challenges are in the detecting changes of operating conditions, adaptation, recursive updates and uncertainty processing. Predictions and decision support are needed in applications.

Finnish companies and research institutions developed the utilisation of environmental data in a multidisciplinary research program *Measurement, Monitoring and Environmental Assessment (MMEA)* in 2010-2015 [6]. The program extended over the entire value chain from measurement technology to managing environmental information. Application areas included agriculture, air quality, energy, environmental efficiency, mining, water, data management, remote sensing and sensors in extreme conditions. Keydemos covered air, earth, energy and water combined with fifth element, data.

Arrowhead project (2013-2017) addressed efficiency and flexibility by introducing a framework to enable collaborative automation. Pilots were built in five application verticals: Production, Smart buildings and infrastructure, Electro-mobility, Energy production and end-user services, Virtual market of energy [7]. A service based approach was introduced for collaborative automation in an open network environment connecting many embedded devices. The demonstrator *Condition monitoring and maintenance integrated to production management in the mining industry* piloted advanced operations and maintenance, including their immediate interactions with control systems performances, maintenance activities, and ultimately with the ERP/MES level of the respective company network [8].

In industry, where very large datasets have been common already long time, the problem has been tackled by *Data Analytics* and *Intelligent Systems*. Artificial intelligence can be applied in various subtasks but experiences in practical industrial applications have shown that the versatile challenges, including events, process phases, changes of operating conditions, subprocesses and recycle flows, require solutions which combine statistical analysis, computational intelligence and optimisation in adaptive case-based systems. Models have been developed for different phenomena theoretically and with data-driven identification. The highly complex systems need an advanced set of methodologies and utilisation of domain expertise in various levels.

Steady-state models can be relatively detailed nonlinear *multiple input, multiple output (MIMO)* models $\vec{y} = F(\vec{x})$, where the output vector $\vec{y} = (y_1, y_2, \dots, y_n)$ is calculated by a nonlinear function F from the input vector $\vec{x} = (x_1, x_2, \dots, x_m)$. *Statistical modelling* includes a wide variety of models based on linear regression, e.g. *response surface models (RSM)* consisting of linear, quadratic and interactive terms [9]. These models can be extended by semi-physical models by using appropriate calculated variables as inputs [10]. *Principal component analysis (PCA)* combines effects of several variables by using linear combinations of the original variables [11] and *Partial least squares regression (PLS)* uses potentially collinear variables [12]. *Nonparametric* models for y_i at each \vec{x} are constructed from data as a weighted average of the neighbouring values of y_i [13].

Additional methodologies for the function $F(\vec{x})$ are provided by fuzzy set systems, artificial neural networks and neurofuzzy methods (Fig. 1). *Fuzzy set theory* presented by Zadeh [14] forms a conceptual framework for linguistically represented knowledge interpreted by using natural language, heuristics and common sense knowl-

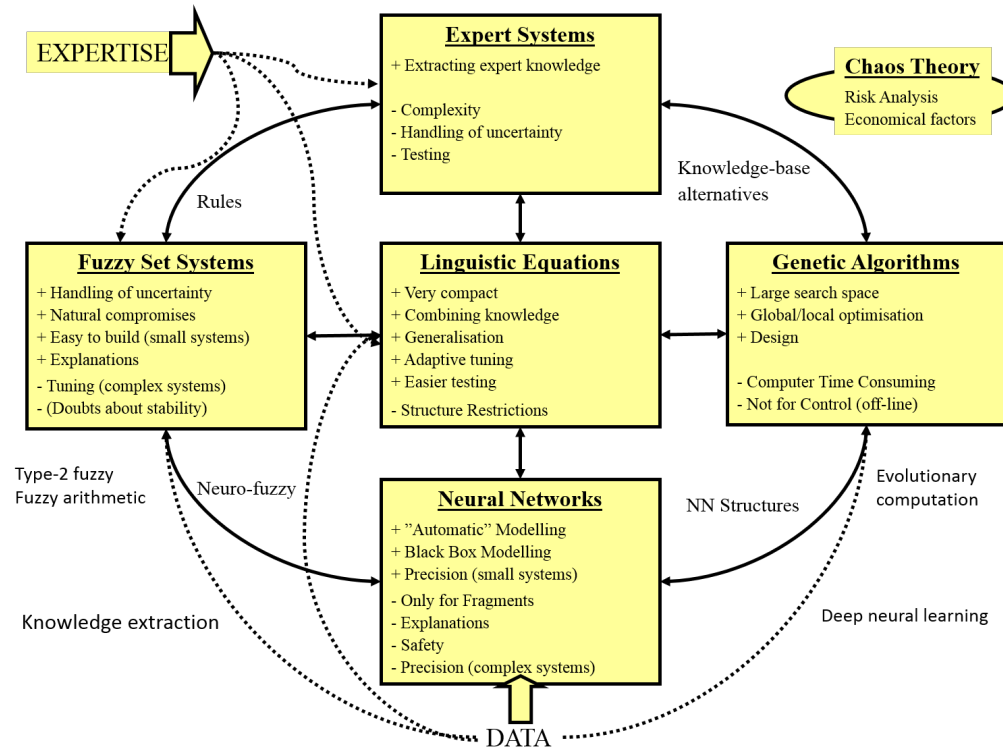


Figure 1: Computational intelligence for combining data and expertise, modified from [3].

edge. Fuzzy logic introduced approximate reasoning into artificial intelligence by maintaining clear connections with fuzzy rule-based systems and expert systems [15]. Fuzzy set systems can also handle contradictory data [16, 17]. The fuzzy sets can be modified by intensifying or weakening modifiers [18]. *Fuzzy relational models* [19] allow one particular antecedent proposition to be associated with several different consequent propositions. *Type-2 fuzzy models* take into account uncertainty about the membership function [20].

The *extension principle* generalises the arithmetic operations for monotonously increasing *inductive mappings* $F(x_j)$. The interval arithmetic presented by Moore [21] is used together with the extension principle on several membership α -cuts of the fuzzy number x_j for evaluating fuzzy expressions [22–24]. *Takagi-Sugeno (TS) fuzzy models* [25] combine fuzzy rules and local linear models.

Linguistic equation (LE) approach originates from fuzzy set systems [26]: rule sets are replaced with equations, and meanings of the variables are handled with nonlinear scaling functions which have close connections to membership functions [27]. The nonlinear systems are built by using the nonlinear scaling with linear equations [28]. Constraints handling [29] and data-based analysis [30], facilitate the recursive updates of the systems [31, 32].

The LE models provide inductive mappings for the extension principle in combined fuzzy systems including fuzzy arithmetics and inequalities [33]. A natural language interface is based on the scaling functions [34]. Temporal reasoning is a very valuable tool for diagnosing and controlling slow processes: the LE based trend analysis introduced in [35] transforms the fuzzy rule-based solution [36] to an equation-based solution.

The deep learning should support applications in three levels. Firstly, *Embedded intelligence* and networks of interacting elements, called *Cyber-physical systems (CPS)*, strengthen links between computational capabilities and physical assets. Lee et al. proposed in [37] a five-layer architecture, including (1) smart connections for data acquisition, (2) data-to-information conversion, (3) cyber level for analysing information, (4) cognition to transfer acquired knowledge to the users, and (5) configuration level is to apply corrective and preventive actions. Secondly, *Decision support systems* combine knowledge-based information with data-based solutions. The integration with domain expertise and the human interaction needs natural language interfaces and uncertainty processing [34]. Thirdly, *Industrial internet (IIoT)* and Cloud computing focus on services which need to integrate operations in several sites.

Smart adaptive systems (SAS) are aimed for developing successful applications in different fields by using three levels of adaptation [38]: (1) adaptation to a changing environment, (2) adaptation to a similar setting without explicitly being ported to it, and (3) adaptation to a new or unknown application. The recursive analysis is important in all these levels.

This paper presents a smart adaptive big data analysis methodology and an advanced deep learning solution which keep the application system in operation. Introduction summarizes shortly the background. Section 2 focuses on variable specific data analysis, which forms the basis for the modelling discussed in Section 3. Smart adaptation methodology based on recursive tuning is presented in Section 4 and the proposed advanced deep learning is discussed in Section 5. The conclusions are drawn in Section 6.

2 Data analysis

Data processing chain needs to be adapted to four types of data (Fig. 2):

- Process measurements are ready to the feature extraction. Sampling is adapted to the phenomena and different statistical features are commonly used in automation and data acquisition systems. Peaks of emission spectra can be analysed in the same way.
- Waveform signals have high frequency components and therefore, the feature extraction is necessary before further processing. Signal processing can improve extracting informative features, e.g. from condition monitoring measurements.
- Image data and videos are analysed in image processing, which is aimed to get an enhanced image or to extract some useful information from it. In big data analysis, the feature extraction is emphasised: patterns and shapes are going to the further processing.
- Sparse data from laboratories, performance indicators, periodic condition monitoring, maintenance data and events need own special processing.

The nonlinearities of all the data types are handled by the nonlinear scaling of the variables. The approach extends normalisation and takes into account asymmetry, recursive updates, uncertainties and is linked with natural language. Domain expertise is highly important in all phases of the data analysis.

2.1 Feature extraction

Feature extraction means dimension reduction in pattern recognition and image processing. In literature, feature extraction and selection methodologies include a wide variety of topics, e.g. classification is important in [39]. The use of statistical features is specialized: arithmetic mean and standard deviation are used for process data, root-mean-square, kurtosis and peak values for signals.

Generalised moments and norms extend this analysis to a wide range of features [40]. The generalised norm is defined by

$$\|{}^{\tau}M_j^p\|_p = (M_j^p)^{1/p} = \left[\frac{1}{N} \sum_{i=1}^N (x_j)_i^p\right]^{1/p}, \quad (1)$$

where the order of the moment $p \in \mathbb{R}$ is non-zero, and N is the number of data values obtained in each sample time τ . The norm (1) calculated for variables x_j , $j = 1, \dots, n$, have the same dimensions as the corresponding variables. The norm $\|{}^{\tau}M_j^p\|_p$ can be used as a central tendency value if all values $x_j > 0$, i.e. $\|{}^{\tau}M_j^p\|_p \in \mathbb{R}$. [41]. The norm can be extended to variables including negative values [31]. Variables x_j can be process measurements, peaks of spectra, measured waveform signals and sparse data.

2.2 Signal processing

Signal processing methods transform, combine or divide the waveform data, including sound, vibration, images or sensor data, and all these may have components from several sources. *Blind source separation (BSS)* methods are used in separating signals to find useful signals [42]. Subset selection techniques are in literature presented in wide meanings, e.g. feature selection techniques include modelling, optimisation and classification in [43]. In high dimensional systems, a subset of variables is selected without altering the original representation of the variables [43]. Feature extraction transforms high dimensional data to lower dimensions by constructing combinations of variables. *Wavelet decomposition* is used for finding local features or compressing the data [44]. *Spectrum analysis*, e.g. fast Fourier analysis (FFT), represents the signal in the frequency domain [45].

Filtering and smoothing are widely used for process data, but derivation and integration can reveal interesting features from waveforms. The calculation of the time domain signal $x^{(a)}(t)$, which is based on a rigorous mathematic theory [46], is performed with three steps. The fast Fourier transform (FFT) is used for the displacement signal $x(t)$ to obtain the complex components $\{X_k\}$, $k =$

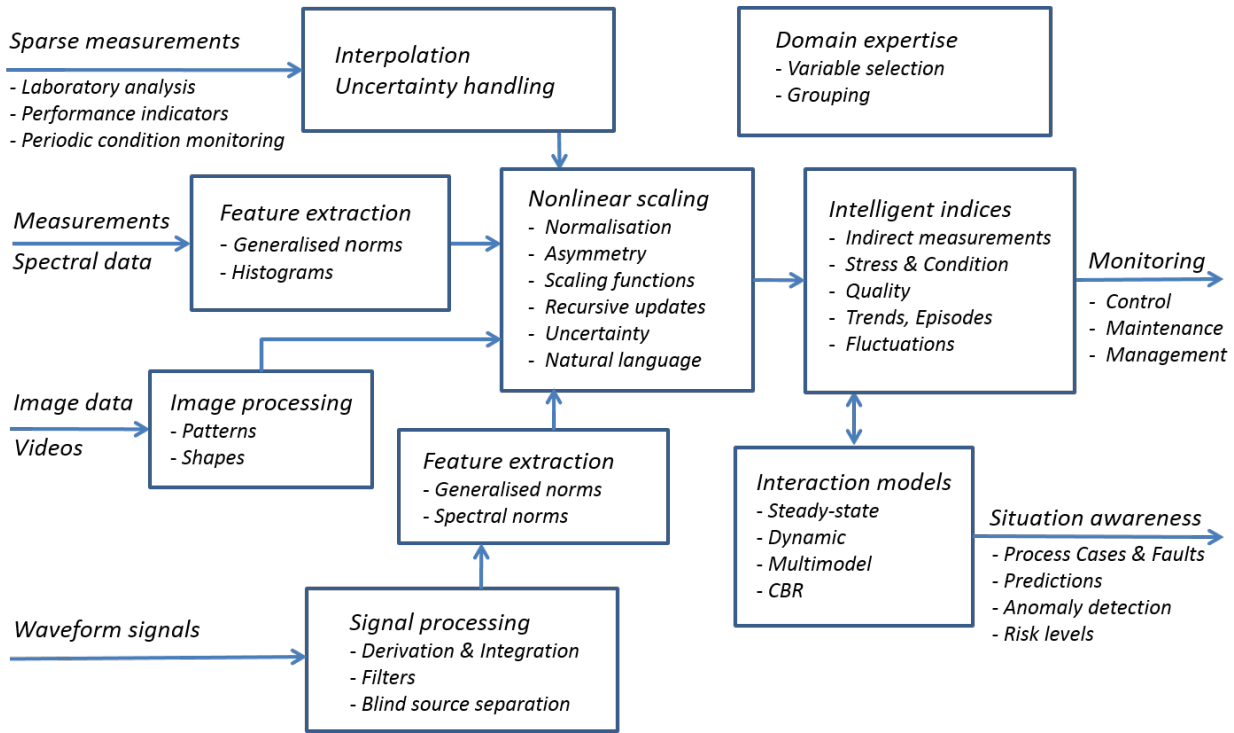


Figure 2: Data processing chains for analysis and modelling.

$0, 1, 2, \dots, (N - 1)$. The corresponding components of the derivative $x^{(\alpha)}(t)$ are calculated by

$$X_{ak} = (i\omega_k)^\alpha X_k, \quad (2)$$

where $\omega = 2\pi f$, $\alpha \in \mathbb{R}$ is the order of derivation. Finally, the resulting sequence is transformed with the inverse Fourier transform FFT^{-1} to produce the derivative signal. The appropriate order of derivation is $\alpha - 2$ for the acceleration signals. [41]

Generalised spectral norms are calculated for waveform signals from the frequency spectrum by

$$\|X^{(\alpha)}\|_p = \left(\frac{1}{N} \sum_{i=1}^N |\{X_j\}^{(\alpha)}|^p \right)^{1/p}, \quad (3)$$

where $\{X_j\}^{(\alpha)}$ is the sequence of complex numbers, representing different frequency components of the signal $\{x_j\}^{(\alpha)}$ [46]. This kind of norm can be used, to provide for information about the change in signal in a certain frequency range or frequency ranges [47].

2.3 Image processing

Digital image processing aims to enhance images or to extract some useful information from them. In big data analytics, algorithms are used to detect and isolate shapes

from images or video streams. Concepts and techniques are discussed in [48]. The hardware and software components need to be used together to facilitate an early detection of problems, e.g. product defects and changes in process operation. The online optical monitoring based on image analysis revealed useful information from the process and can be used in forecasting the quality of biologically treated wastewater [49]

2.4 Sparse measurements

Sparse condition monitoring measurements are analysed with the methods presented above for the waveform signals, the only difference is the sparsity of values. Laboratory analyses are based on sampling and can be frequent only if automatic sampling is used. These measurements may contain spectroscopy. Uncertainty fundamentally affects the decisions that are based upon the measurement result [50].

Maintenance and operation performance can be assessed with various measures: harmonised indicators are based on cost, time, man-hours, inventory value, work orders and cover of the criticality analysis, key performance indicators (KPIs) reflect the critical success factors and the goals, the overall equipment effectiveness (OEE) in-

cludes non-financial metrics for the manufacturing success. Reliability-centered maintenance (RCM) is based on statistical analyses and statistical process control (SPC) is used in monitoring a process through the control charts. [51] These indicators need interpolation and uncertainty handling (Fig. 2).

2.5 Nonlinear scaling

Nonlinear scaling brings various measurements and features to the same scale by using monotonously increasing scaling functions $x_j = f(X_j)$ where x_j is the variable and X_j the corresponding scaled variable. The function $f()$ consist of two second order polynomials, one for the negative values of X_j and one for the positive values, respectively. The corresponding inverse functions $X_j = f^{-1}(x_j)$ based on square root functions are used for scaling to the range $[-2, 2]$, denoted as linguistification. The monotonous functions allow scaling back to the real values by using the function $f()$. [28]

The parameters of the functions are extracted from measurements by using generalised norms and moments. The support area is defined by the minimum and maximum values of the variable, i.e. the support area is $[\min(x_j), \max(x_j)]$ for each variable $j, j = 1, \dots, m$. The central tendency value, c_j , divides the support area into two parts, and the core area is defined by the central tendency values of the lower and the upper part, $(c_l)_j$ and $(c_h)_j$, correspondingly. This means that the core area of the variable j defined by $[(c_l)_j, (c_h)_j]$ is within the support area. The orders, p , corresponding to the corner points are chosen by using the generalised skewness,

$$(\gamma_k^p)_j = \frac{1}{N\sigma_j^k} \sum_{i=1}^N [(x_j)_i - \|^\tau M_j^p\|_p]^k. \quad (4)$$

The standard deviation σ_j is the norm (1) with the order $p = 2$. [30]

The scaling functions monotonous and increasing if the ratios,

$$\alpha_j^- = \frac{(c_l)_j - \min(x_j)}{c_j - (c_l)_j}, \quad \alpha_j^+ = \frac{\max(x_j) - (c_h)_j}{(c_h)_j - c_j}, \quad (5)$$

are both in the range $[\frac{1}{3}, 3]$, see [29]. The coefficients of the second order polynomials are represented by

$$\begin{aligned} a_j^- &= \frac{1}{2}(1 - \alpha_j^-) \Delta c_j^-, \\ b_j^- &= \frac{1}{2}(3 - \alpha_j^-) \Delta c_j^-, \\ a_j^+ &= \frac{1}{2}(\alpha_j^+ - 1) \Delta c_j^+, \\ b_j^+ &= \frac{1}{2}(3 - \alpha_j^+) \Delta c_j^+, \end{aligned} \quad (6)$$

where $\Delta c_j^- = c_j - (c_l)_j$ and $\Delta c_j^+ = (c_h)_j - c_j$.

2.6 Uncertainty

The feasible range can be defined as a type-2 trapezoidal membership function since the norm values obtained from different time periods have differences, i.e. the parameters of the scaling functions can be represented as fuzzy numbers. A strong increase in uncertainty may demonstrate a change of operating conditions. The ratios α_j^- and α_j^+ are calculated by applying the interval arithmetic on (5). The constraint range $[\frac{1}{3}, 3]$ must be taken into account before calculating the coefficients (6). Then the extension principle is used for calculating the scaled values as fuzzy numbers.

2.7 Natural language

The nonlinear scaling approach provides an unified solution for natural language interpretations since all the scaled variables are in the same range $[-2, 2]$. The integer numbers $\{-2, -1, 0, 1, 2\}$ correspond labels $\{\text{very low, low, normal, high, very high}\}$ or $\{\text{high negative, negative, zero, positive, high positive}\}$, for example, and represented as fuzzy numbers, which can be modified by fuzzy modifiers, which are used as intensifying or weakening adverbs. The resulting terms,

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq A_4 \subseteq A_5, \quad (7)$$

correspond to the powers of the membership in the powering modifiers (Table 1). The vocabulary can also be chosen in a different way, e.g. highly, fairly, quite [52]. Only the sequence of the labels is important. Linguistic variables can be processed with the conjunction (and), disjunction (or) and negation (not). More examples can be found in [18].

Table 1: Modifiers of fuzzy numbers [34]

Fuzzy number	Fuzzy label	Degree of membership
A_1	<i>extremely A</i>	μ^4
A_2	<i>very A</i>	μ^2
A_3	<i>A</i>	μ
A_4	<i>more or less A</i>	$\mu^{\frac{1}{2}}$
A_5	<i>roughly A</i>	$\mu^{\frac{1}{4}}$

3 Modelling

Intelligent indices, which are developed from the scaled data and enhanced with temporal analysis, are the key el-

ements of modelling (Fig. 2). Interactions are normally linear but more complex solutions can be built with computational intelligence and case-based solutions. Dynamic models are based on parametric model structures. Indices are used as indirect measurements and models enhance situation awareness.

3.1 Intelligent indices

The basic form of the intelligent index is a scaled feature or measurement but more indices can be developed as the weighted sums of several scaled features. In [30], the cavitation index of a Kaplan turbine was based on a single scaled feature and several faults of the supporting rolls of a lime kiln required two features. *Linguistic principal components (LPCs)*, which extend the linear PCA by using the nonlinear scaling, are generalisations of this. Intelligent condition and stress indices provide an unified approach to use different measurements and features in condition monitoring [30].

3.2 Temporal analysis

Temporal analysis focused on important variables provides useful information, including trends, fluctuations and anomalies, for decisions on higher level recursive adaptation. *Trend analysis* produces useful indirect measurements for the early detection of changes. For any variable j , a *trend index* $I_j^T(k)$ is calculated from the scaled values X_j with a linguistic equation

$$I_j^T(k) = \frac{1}{n_S + 1} \sum_{i=k-n_S}^k X_j(i) - \frac{1}{n_L + 1} \sum_{i=k-n_L}^k X_j(i), \quad (8)$$

which is based on the means obtained for a short and a long time period, defined by delays n_S and n_L , respectively. The index value is in the linguistic range $[-2, 2]$, representing the strength of both decrease and increase of the variable x_j . [35, 53]

An increase is detected if the trend index exceed a threshold $I_j^T(k) > \epsilon_1^+$. Correspondingly, $I_j^T(k) < -\epsilon_1^-$ for a decrease. The derivative of the index $I_j^T(k)$, denoted as $\Delta I_j^T(k)$, is used for analysing the full set of the triangular episodic representations (Fig. 3). Trends are linear if the derivative is close to zero: $-\epsilon_2^- < \Delta I_j^T(k) < \epsilon_2^+$. Area D close to $[2, 2]$ and area B close to $[-2, -2]$ are dangerous situations, which introduce warnings and alarms. Areas A and C mean that an unfavourable trend is stopping.

Severity of the situation can be evaluated by a *deviation index*, which is a weighted sum of $X_j(k)$, $I_j^T(k)$ and

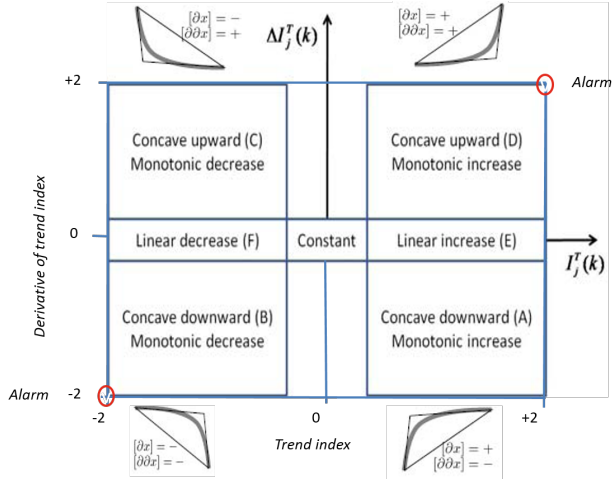


Figure 3: Triangular episodic representations defined by the index $I_j^T(k)$ and the derivative $\Delta I_j^T(k)$.

$\Delta I_j^T(k)$. This index has its highest absolute values, when the difference to the set point is very large and is getting still larger with a fast increasing speed [53]. This can be understood as a third dimension in Fig. 3. The trends of the parameters α_j^- , α_j^+ , Δc_j^- and Δc_j^+ give useful information about changes of the scaling functions. The ranges of these parameter are $[\frac{1}{3}, 3]$, $[\frac{1}{3}, 3]$, $[\frac{1}{4}, \frac{3}{4}](c_j - \min(x_j))$ and $[\frac{1}{4}, \frac{3}{4}](\max(x_j) - c_j)$, respectively. The changes in the case structure are seen in the trends of the interaction coefficients.

The trend analysis is tuned to applications by selecting the time periods n_L and n_S . Further fine-tuning can be done by adjusting the weight factors w_j^{T1} and w_j^{T2} used for the indices $I_j^T(k)$ and $\Delta I_j^T(k)$. The thresholds $\epsilon_1^+ = \epsilon_1^- = \epsilon_2^+ = \epsilon_2^- = 0.5$. The calculations are done with numerical values and the results are represented in natural language.

The *fluctuation indicators* calculate the difference of the high and the low values of the measurement as a difference of two moving generalized norms:

$$\Delta x_j^F(k) = \|^{K_S \tau} M_j^{p_h}\|_{p_h} - \|^{K_S \tau} M_j^{p_l}\|_{p_l}, \quad (9)$$

where the orders $p_h \in \mathbb{R}$ and $p_l \in \mathbb{R}$ are large positive and negative, respectively. The moments are calculated from the latest $K_S + 1$ values, and an average of several latest values of $\Delta x_j^F(k)$ is used as an indicator. [54]

3.3 Interactions

The basic form of the linguistic equation (LE) model is a static mapping in the same way as fuzzy set systems and neural networks, and therefore dynamic models will in-

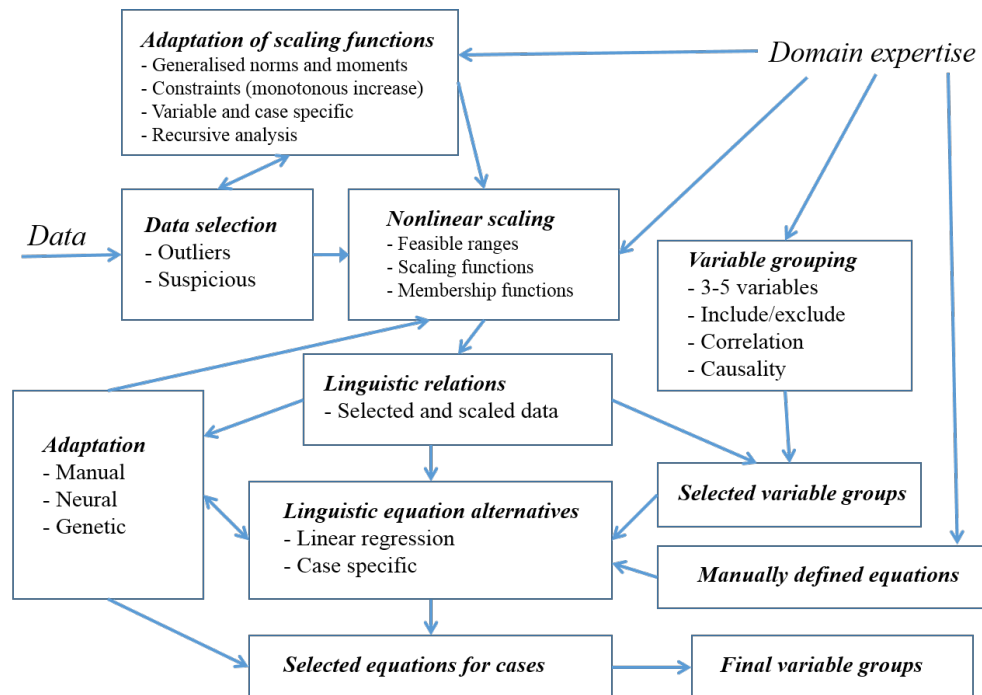


Figure 4: Data-based modelling with linguistic equations, modified from [55].

clude several inputs and outputs originating from a single variable [28]. Adaptation of the nonlinear scaling is the key part in the data-based LE modelling (Fig. 4). All variables can be analysed in parallel with the methodology described above and assessed with domain expertise. Interactions are analysed with linear modelling methodologies from the scaled data in the chosen time period. In large-scale systems, a huge number of alternatives need to be compared, e.g. in a paper machine application, 72 variables produced almost 15 million three to five variable combinations. Correlations and causalities based on domain expertise are needed to find feasible variable groups [56].

Fuzzy set systems and fuzzy arithmetics expand the application areas in following ways [57]:

- LE models replace linear models in TS models;
- Fuzzy calculus is applied in models by using LE models, fuzzy inputs and/or coefficients both in the antecedent and consequent part;
- Use fuzzy inequalities in developing fuzzy facts for the fuzzy rule-based systems.

Domain expertise is important in these modules. *Neural networks* can represent very complex nonlinear interactions but only highly simplified models are needed when the nonlinear scaling is successfully defined. Modelling and simulation methodologies of complex systems are dis-

cussed in more details in [33]. In decomposed systems, the composite local models consisting of partially overlapping models are handled by fuzzy logic [58].

Complexity is gradually increased with decomposition and higher level structure. *Case based reasoning (CBR)* integrates problem solving and learning in variety of domains [59] but does not prescribe any specific technology [60]. Therefore, it is a feasible methodology for integrating the overall system. *Evolutionary computation* provide efficient tools for all these systems since everything is defined by parameters.

3.4 Dynamic modelling

External dynamic models provide the dynamic behaviour for the LE models developed for a defined sampling interval in the same way as in various identification approaches discussed in [10]. Dynamic LE models use the parametric model structures, ARX, ARMAX, NARX etc., but the nonlinear scaling reduces the number of input and output signals needed for the modelling of nonlinear systems. For the default LE model, all the degrees of the polynomials become very low:

$$Y(t) + a_1 Y(t-1) = b_1 U(t-n_k) + e(t) \quad (10)$$

for the scaled variables Y and U .

Process phases can have totally different models with different variables. Also phenomenological models can be included in the overall system but their parameters need to be adapted to the operating condition by computational intelligence.

4 Smart adaptation

All the phases of the data-based LE modelling shown in Fig. 4 can be used in the recursive analysis as well. The recursive part focuses on the scaling functions and the interactions are updated only if needed. The adaptation ranges from the slight modifications of scaling functions to completely new models. The adaptation level is chosen by using fuzzy logic.

4.1 Recursive scaling

The parameter of the scaling functions can be recursively updated by using the norms (1) with the orders defined in the tuning. The norm values are updated by including new equal sized sub-blocks in calculations since the computation of the norms can be done from the norms obtained for the equal sized sub-blocks, i.e. the norm for several samples can be obtained as the norm of the norms of the individual samples:

$$\|^{K_s \tau} M_j^p\|_p = \left\{ \frac{1}{K_s} \sum_{i=1}^{K_s} [(\tau M_j^p)_i^{1/p}]^p \right\}^{1/p} \quad (11)$$

where K_s is the number of samples $\{x_j\}_{i=1}^N$. In automation and data collection systems, the sub-blocks are normally used for arithmetic mean ($p = 1$).

The parameters of the scaling functions can be recursively updated with by including new samples in calculations. The number of samples can be increasing or fixed with some forgetting or weighting [31]. The orders of the norms are redefined if the operating conditions change considerably. The new orders are obtained by using the generalised skewness (4) for the data extended with the data collected from the new situation. If the changes are drastic, the calculations are based on the new data only. The decision of starting the redefinition is fuzzy and the data selection is important.

4.2 Interactions

Linear regression and parametric models are used in the recursive tuning of the interaction equations. The set of

equation alternatives (Fig. 4) is useful in the recursive analysis since the set is validated with domain expertise. The LE approach uses the preference sequence: scaling, shape of scaling functions and interaction equations, which is consistent with the stages of adaptive fuzzy control: first scaling, then the shape of membership functions and finally rulebase.

The interaction models are not changed if the scaling functions change only slightly. The coefficients are obtained by using the data collected from the chosen time period if the feasible range is changed. Uncertainties can be calculated by comparing the coefficients extracted from several short periods.

Considerably revised scaling functions may require updates for the interactions as well. However, the re-tuning is started only if the current equations do not operate sufficiently well. The earlier chosen set of alternative equations is used first. New equations are included if new variables become important. The selected variable groups (Fig. 4) are analysed first. Considerable changes in operating conditions mean that the full data-based analysis is needed. This level forms the model basis for the case-based reasoning (CBR), see [56].

4.3 Fuzzy logic

The recursive data analysis produces parameters for the scaling functions and interactions. Uncertainties of the parameters are obtained for any time period which containing several sub-blocks, i.e. the variables are represented by fuzzy numbers. Changes in operating conditions are detected by comparing the similarities of the original and modified fuzzy numbers. The detection is based on fuzzy inequalities $<$, \leq , $=$, \geq and $>$ between the new fuzzy parameters and the fuzzy parameters of the case. The resulting 5X5 matrix includes the degrees of membership of these five inequalities for five parameters. The results are interpreted with the natural language interface which provides an important channel in explaining the changes to the users.

4.4 Changes of operating conditions

Changes of the scaling functions and interaction coefficients are symptoms of changes in operation. The intelligent trend analysis provides early warning about changes in variable levels, fluctuations and uncertainty. All the variables and intelligent indices are represented in the same range $[-2, 2]$, i.e. the same analysis and linguistic in-

interpretation can be applied in all of them. The corresponding levels and their degrees of membership can be used in the fuzzy decision making.

The full analysis is needed fairly seldom although the process changes considerably. For example, new phenomena activate with time in wearing, but the models used in prognostics can be updated by expanding the scaling functions (Fig. 5). The generalised statistical process control (GSPC) introduced in [61] could give an early warning.

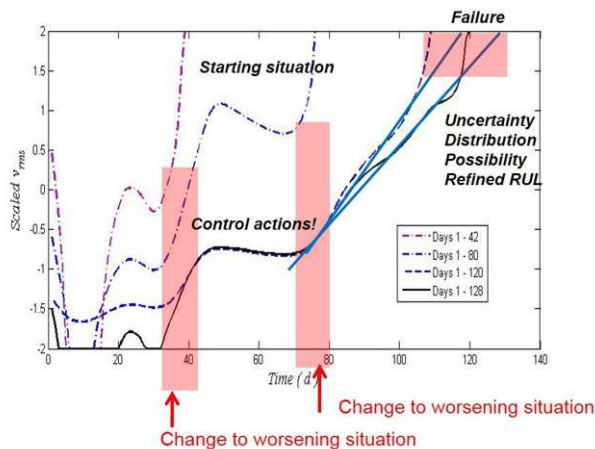


Figure 5: Recursive adaptation in prognostics [62].

5 Advanced deep learning

Technical systems have data in various forms which require specific methodologies (Fig. 2) but similar scales for all of them makes the analysis of interactions easier. The data processing chain from measurements and open data to applications was the main result of *MMEA* [6]. The smart integration of subsystems (Fig. 6) extends the solutions based on *IoT* and *Data analytics* towards *Industrial internet of services* (*I²oS*). The collaborative automation framework introduced in *Arrowhead* is a good platform for these systems [8].

The advanced deep learning combines statistical and modelling methodologies with computational intelligence (Fig. 1). Five hierarchical layers can be structured from the processing chain shown in Fig. 2. The levels are consistent with the levels presented in [37].

Layer 1 - Connections. Variable specific features are extracted and specifications defined for process measurements (Section 2.1), waveform signals (Section 2.2), im-

ages and videos (Section 2.3), spectral data (Section 2.1) and sparse measurements (Section 2.4). The generalised norms are beneficial in providing similar settings for various applications, especially for waveform signals where *Edge computing* is needed for local calculations, especially for the waveform signals and image data.

Layer 2 - Conversion. Features are converted to the same scale by extracting the meanings with the variable specific nonlinear scaling (Section 2.5). The output includes feature specific uncertainties (Section 2.6) and the results are presented in natural language (Section 2.7). Recursive scaling is available (Section 4.1) and the temporal analysis provides information about trends (Fig. 3) and fluctuations (Section 3.2). This layer is the key to the advanced deep learning by providing a feasible solution to divide the conversion and cyber levels. The differences of scaling functions between operating areas can be used in selecting possible cases.

Layer 3 - Cyber. Interactions are analysed for the intelligent indices (Fig. 4): indicators may combine several indices (Section 3.1) and versatile interaction models can be developed by linear methodologies (Section 3.3). Dynamic structures are included if needed (Section 3.4). In practice, case-based solutions are important: local composite models can be sufficient but the interactions might also be highly different in different operating conditions. The need for additional cases is finalised in this layer. The overall system can be managed by case-based reasoning (CBR). The recursive tuning is the key to expanding the system (Section 4.1). The resulting intelligent analysers can be single scaled indices or combinations of them. Phenomenological models are important extensions of this layer.

Layer 4 - Cognition. Service solutions are developed by combining intelligent analysers and forecasting models in *Monitoring*. Changes of operating conditions (Fig. 5) are detected and the situation awareness is improved in the risk analysis. The *Domain expertise* is essential in *Decision support systems*. The natural language interface defined in the conversion level (Section 2.7) is important in this layer: all important features and indices are available in scaled and linguistic forms.

Layer 5 - Configuration. *Automatic solutions* for control and maintenance solutions are developed by combining intelligent analysers and control (Fig. 6). The controller can include many special control actions which are activated when needed [63]. Condition monitoring is the key to the improved *Condition-based maintenance* when real-time measurements are processed through the layers discussed above.

The advanced deep learning uses gradually refining layers: informative features are needed in the conversion

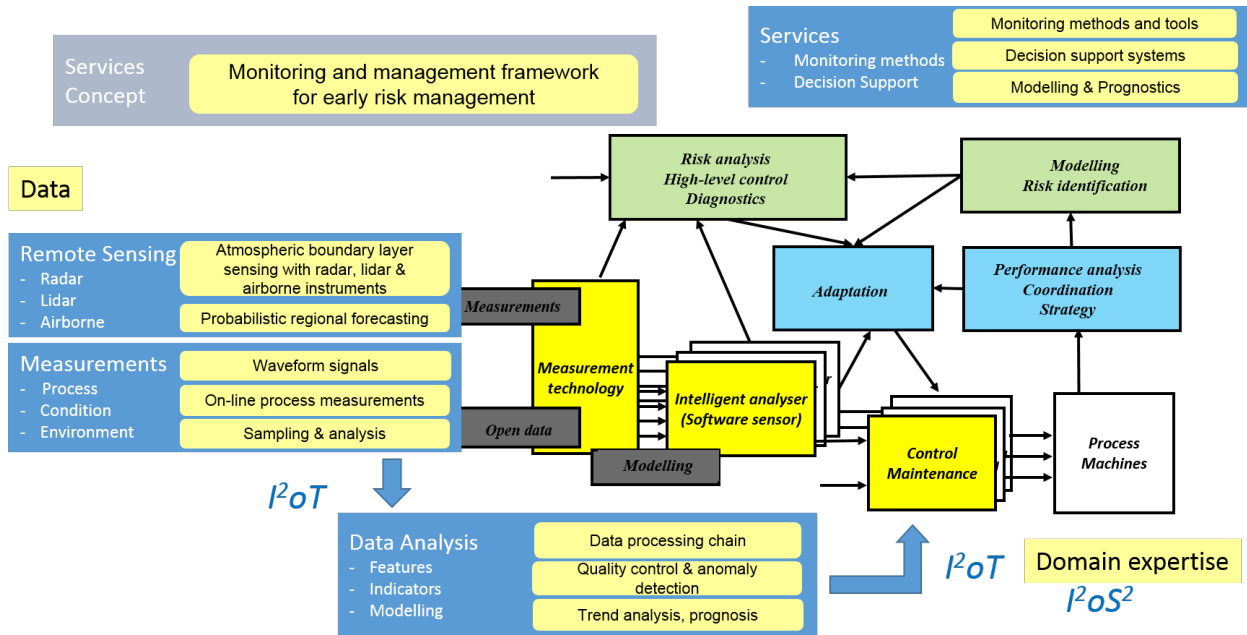


Figure 6: Smart integration of subsystems in the monitoring and management framework for early risk management, modified from [58].

layer which forms the basis for the intelligent analysers, and finally, the service and automation solutions combine these data-driven solutions with the domain expertise. The levels of *Smart adaptive systems* are further refined by the recursive analysis within these layers. The adaptation to a *changing environment* has two sub-levels: first updating the scaling functions (Layer 2) and then interactions (Layer 3) if needed. A short-term memory is needed for incremental or on-line learning, a long-term memory for recognising context drifting.

Similar settings are realised with the set of equation alternatives (Fig. 4). Successful past solutions and the idea of reasoning by analogy are used: the generalised norms and nonlinear scaling provide compact solutions for this level. The nonlinear scaling makes similar settings more widely available also in Cyber, Cognition and Configuration layers.

The adaptation to a *new or unknown* application includes the full data analysis and modelling (Fig. 4). In real applications, the constraint of starting from zero knowledge is modified to building new knowledge or, at least, improving the existing one.

The learning layers and the modular application structures together with edge computing are promising for cyber-physical systems: measurement technology, intelligent analysers, control and maintenance are realized as agents which are communicating through I^2oT (Fig. 6). For a waveform signal, the local calculations reduces the

amount of data with a factor 10^5 even if several features are extracted. The analysis of images and videos results a number of features. Already this and conversion layer make cloud computing possible but in technical systems, the local calculations are preferred for the cyber level as well when known compact structures are used.

The modules of the smart adaptive data analysis have been developed and tested in various applications, including monitoring [53, 61], control [31, 63], diagnostics [56], condition monitoring and maintenance [57, 64, 65], and management [51, 66]. The basis of the calculation chain was introduced in *MMEA*, see [6]. The local calculations and integration of systems in collaborative automation have been discussed in [8].

Many artificial intelligence approaches, especially neural computing, rely highly on unsupervised methods and simultaneous processing of massive datasets. The hierarchical deep learning ideas, which improve the solution, have been developed for heterogeneous systems, which include also unstructured data. This type of methodologies do not utilize the domain expertise and known operational information. Layers 1 and 2 are needed, especially for the more structured data. Artificial intelligence can be useful in finding possible interactions in the Cyber layer but the results must be assessed through the advanced analysis presented above.

6 Conclusion

The smart adaptive data analysis and the data processing chain are reorganized to form a five-layer advanced deep learning platform which supports levels of smart adaptive systems and development of cyber-physical systems.

Connections. Generalised norms operate for extracting features from process measurements, peaks of spectra, measured waveform signals and sparse data. Measured waveform signals can be transformed, combined and divided before the feature extraction. Image processing is used for detecting and isolating shapes from images and videos.

Conversion. Generalised norms and moments are used in the data-driven tuning of the monotonously increasing scaling functions. The nonlinear scaling is the key approach of the advanced deep learning since it extracts the meanings of the feature levels and opens new possibilities for temporal analysis and uncertainty estimation. The parametric definitions allow recursive analysis for all these.

Cyber. Interactions between the scaled values can be analysed with linear methodologies: versatile indicators can be constructed as the weighted sums of indices. The compact models also allow case-based systems expanding through recursive adaptation. Local composite models and dynamic structures extend solutions to intelligent analysers further. Neural deep learning can be a part of this level.

Cognition. Domain expertise is used in combining service solutions and intelligent analysers obtained from features, indices and models. Resulting systems can be used in monitoring and decision support.

Configuration. Automatic solutions for control and maintenance utilize services for varying operating conditions in large scale complex systems. Trade-off between solutions is handled with fuzzy logic.

Acknowledgement: The author would like to thank the research program “Measurement, Monitoring and Environmental Efficiency Assessment (MMEA)” funded by the TEKES (the Finnish Funding Agency for Technology and Innovation) and the Artemis Innovation Pilot project “Production and energy system automation and Intelligent-Built (Arrowhead)”.

References

- [1] N. Dedić and C. Stanier. Towards differentiating business intelligence, big data, data analytics and knowledge discovery. *Lecture Notes in Business Information Processing*, 285:114–122, 2017.
- [2] I.A.T. Hashem, I. Yaqoob, N.B. Anuar, S. Mokhtar, A. Gani, and S. Ullah Khan. The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47:98–115, 2015.
- [3] E. K. Juuso. Computational intelligence in distributed interactive synthetic environments. In Agostino G. Bruzzone and Eugene J. H. Kerckhoffs, editors, *Simulation in Industry, Proceedings of the 8th European Simulation Symposium, Simulation in Industry, ESS'96, Genoa, Italy, October 2-5, 1996*, pages 157–162, San Diego, USA, 1996. SCS International.
- [4] Maryam M. Najafabadi, Flavio Villanustre, Taghi M. Khoshgof-taar, Naeem Seliya, Randall Wald, and Edin Muharemagic. Deep learning applications and challenges in big data analytics. *Journal of Big Data*, 2(1):1–21, Feb 2015.
- [5] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61(Supplement C):85 – 117, 2015.
- [6] CLIC Innovation. Final report: Measurement, monitoring and environmental efficiency assessment. <http://mmeafinalreport.fi/>, 2015. Accessed: 2018-01-03.
- [7] Arrowhead framework. <http://www.arrowhead.eu>, 2017. Accessed: 2018-01-03.
- [8] E. Jantunen, M. Karaila, D. Hästbacka, A. Koistinen, L. Barna, E. Juuso, P. Punal Pereira, S. Besseau, and J. Hoepffner. Application system design - Maintenance. In Jerker Delsing, editor, *IoT Automation - Arrowhead Framework*, pages 247–280. CRC Press, Taylor & Francis Group, Boca Raton, FL, 2017. ISBN 978-1-4987-5675-4.
- [9] G. E. P. Box and K. B. Wilson. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society. Series B*, 13(1):1–45, 1951.
- [10] L. Ljung. *System Identification - Theory for the User*. Prentice Hall, Upper Saddle River, N.J., 2nd edition, 1999.
- [11] I. T. Jolliffe. *Principal Component Analysis*. Springer, New York, 2 edition, 2002. 487 pp.
- [12] R. W. Gerlach, B. R. Kowalski, and H. O. A. Wold. Partial least squares modelling with latent variables. *Anal. Chim. Acta*, 112(4):417–421, 1979.
- [13] L. Wasserman. *All of Nonparametric Statistics*. Springer Texts in Statistics. Springer, Berlin, corr. 3rd edition, 2007.
- [14] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8(June):338–353, 1965.
- [15] D. Dubois, H. Prade, and L. Ughetto. Fuzzy logic, control engineering and artificial intelligence. In H. B. Verbruggen, H.-J. Zimmermann, and R. Babuska, editors, *Fuzzy Algorithms for Control, International Series in Intelligent Technologies*, pages 17–57. Kluwer, Boston, 1999.
- [16] A. Krone and H. Kiendl. Automatic generation of positive and negative rules for two-way fuzzy controllers. In H.-J. Zimmermann, editor, *Proceedings of the Second European Congress on Intelligent Technologies and Soft Computing - EUFIT'94, Aachen, September 21 - 23, 1994*, volume 1, pages 438–447, Aachen, 1994. Augustinus Buchhandlung.

- [17] A. Krone and U. Schwane. Generating fuzzy rules from contradictory data of different control strategies and control performances. In *Proceedings of the Fuzz-IEEE'96, New Orleans, USA*, pages 492–497, 1996.
- [18] M. De Cock and E. E. Kerre. Fuzzy modifiers based on fuzzy relations. *Information Sciences*, 160(1-4):173–199, 2004.
- [19] W. Pedrycz. An identification algorithm in fuzzy relational systems. *Fuzzy Sets and Systems*, 13(2):153–167, 1984.
- [20] J. M. Mendel. Advances in type-2 fuzzy sets and systems. *Information Sciences*, 177(1):84–110, 2007.
- [21] R. E. Moore. *Interval Analysis*. Prentice Hall, Englewood Cliffs, NJ, 1966.
- [22] J. J. Buckley and Y. Qu. On using α -cuts to evaluate fuzzy equations. *Fuzzy Sets and Systems*, 38(3):309–312, 1990.
- [23] J. J. Buckley and Y. Hayashi. Can neural nets be universal approximators for fuzzy functions? *Fuzzy Sets and Systems*, 101:323–330, 1999.
- [24] J. J. Buckley and T. Feuring. Universal approximators for fuzzy functions. *Fuzzy Sets and Systems*, 113:411–415, 2000.
- [25] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 15(1):116–132, 1985.
- [26] E. K. Juuso and K. Leiviskä. Adaptive expert systems for metallurgical processes. In S.-L. Jämsä-Jounela and A. J. Niemi, editors, *Expert Systems in Mineral and Metal Processing, IFAC Workshop, Espoo, Finland, August 26–28, 1991, IFAC Workshop Series, 1992, Number 2*, pages 119–124, Oxford, UK, 1992. Pergamon.
- [27] E. K. Juuso. Fuzzy control in process industry: The linguistic equation approach. In H. B. Verbruggen, H.-J. Zimmermann, and R. Babuška, editors, *Fuzzy Algorithms for Control, International Series in Intelligent Technologies*, volume 14 of *International Series in Intelligent Technologies*, pages 243–300. Kluwer, Boston, 1999.
- [28] E. K. Juuso. Integration of intelligent systems in development of smart adaptive systems. *International Journal of Approximate Reasoning*, 35(3):307–337, 2004.
- [29] E. K. Juuso. Tuning of large-scale linguistic equation (LE) models with genetic algorithms. In M. Kolehmainen, editor, *Revised selected papers of the International Conference on Adaptive and Natural Computing Algorithms - ICANNGA 2009, Kuopio, Finland, Lecture Notes in Computer Science*, volume LNCS 5495, pages 161–170. Springer-Verlag, Heidelberg, 2009.
- [30] E. Juuso and S. Lahdelma. Intelligent scaling of features in fault diagnosis. In *7th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, CM 2010 - MFPT 2010, 22–24 June 2010, Stratford-upon-Avon, UK*, volume 2, pages 1358–1372, 2010.
- [31] E. K. Juuso. Recursive tuning of intelligent controllers of solar collector fields in changing operating conditions. In S. Bittani, A. Cenedese, and S. Zampieri, editors, *Proceedings of the 18th World Congress The International Federation of Automatic Control, Milano (Italy) August 28 - September 2, 2011*, pages 12282–12288. IFAC, 2011.
- [32] E. Juuso and S. Lahdelma. Intelligent trend indices and recursive modelling in prognostics. In *8th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, CM 2011 - MFPT 2011, 20–22 June 2011, Cardiff, UK*, volume 1, pages 440–450. Curran Associates, NY, USA, 2011. www.scopus.com.
- [33] E. K. Juuso. Intelligent methods in modelling and simulation of complex systems. *Simulation Notes Europe SNE*, 24(1):1–10, 2014.
- [34] E. K. Juuso. Informative process monitoring with a natural language interface. In *2016 UKSim-AMSS 18th International Conference on Modelling and Simulation, 6–8 April, 2016, Cambridge, UK*, pages 105–110. IEEE Computer Society, 2016.
- [35] E. K. Juuso. Intelligent trend indices in detecting changes of operating conditions. In *2011 UKSim 13th International Conference on Modelling and Simulation*, pages 162–167. IEEE Computer Society, 2011.
- [36] J. T.-Y. Cheung and G. Stephanopoulos. Representation of process trends - part I. A formal representation framework. *Computers & Chemical Engineering*, 14(4/5):495–510, 1990.
- [37] J. Lee, B. Bagheri, and H.-A. Kao. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3(Supplement C):18 – 23, 2015.
- [38] D. Anguita. Smart adaptive systems - state of the art and future directions for research. In *Proceedings of Eunite 2001 - European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems, July 13–14, 2001, Tenerife, Spain*, pages 1–4. Wissenschaftsverlag Mainz, Aachen, 2001.
- [39] I. Guyon and A. Elisseeff. An introduction to feature extraction. In I. Guyon, S. Gunn, M. Nikravesh, and L. Zadeh, editors, *Feature Extraction: Foundations and Applications*, volume 207 of *Studies in Fuzziness and Soft Computing*, pages 1–25. Springer, Heidelberg, 2003.
- [40] S. Lahdelma and E. Juuso. Generalised l_p norms in vibration analysis of process equipments. In *7th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, CM 2010 - MFPT 2010, 22–24 June 2010, Stratford-upon-Avon, UK*, volume 1, pages 614–626. Curran Associates, NY, USA, 2010. ISBN 978-1-61839-013-4.
- [41] S. Lahdelma and E. Juuso. Signal processing and feature extraction by using real order derivatives and generalised norms. Part 1: Methodology. *International Journal of Condition Monitoring*, 1(2):46–53, 2011.
- [42] Y. Deville, C. Jutten, and R. Vigario. Overview of source separation applications. In P. Comon and C. Jutten, editors, *Handbook of Blind Source Separation*, pages 639–681. Academic Press, 2010.
- [43] Y. Saeys, I. Inza, and P. Larranaga. A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517, 2007.
- [44] H. A. Gaberson. The use of wavelets for analyzing transient machinery vibration. *Sound and Vibration*, 36:12–177, 2002.
- [45] D. E. Newland. *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*. Longman Scientific & Technical, Harlow, UK, 3rd edition, 1993.
- [46] S. G. Samko, A. A. Kilbas, and O. I. Marichev. *Fractional Integrals and Derivatives. Theory and Applications*. Gordon and Breach, Amsterdam, 1993. 976 pp.
- [47] K. Karioja and E. Juuso. Generalised spectral norms – a new method for condition monitoring. *International Journal of Condition Monitoring*, 6(1):13–16, 2016.
- [48] C. Solomon and T. Breckon. *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*. John Wiley & Sons, 2010. ISBN 9780470689776.

- [49] J. Tomperi, E. Koivuranta, A. Kuokkanen, E. Juuso, and K. Leiviskä. Real-time optical monitoring of the wastewater treatment process. *Environmental Technology (United Kingdom)*, 37(3):344–351, 2016.
- [50] M. H. Ramsey and S. L. R. Ellison, editors. *Eurachem/ EURLAB/ CITAC/ Nordtest/ AMC Guide: Measurement uncertainty from sampling: a guide to methods and approaches*. Eurachem, 2007. ISBN 978 0 948926 26 6.
- [51] E. K. Juuso and S. Lahdelma. Intelligent performance measures for condition-based maintenance. *Journal of Quality in Maintenance Engineering*, 19(3):278–294, 2013.
- [52] E. K. Juuso. Integration of knowledge-based information in intelligent condition monitoring. In *9th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, 12-14 June 2012, London, UK*, volume 1, pages 217–228. Curran Associates, NY, USA, 2012.
- [53] E. Juuso, T. Latvala, and I. Laakso. Intelligent analysers and dynamic simulation in a biological water treatment process. In I. Troch and F. Breiteneker, editors, *6th Vienna Conference on Mathematical Modelling - MATHMOD 2009, February 11-13, 2009, Argesim Report no. 35*, pages 999–1007. Argesim, 2009. ISBN 978-3-901608-35-3.
- [54] E. K. Juuso. Model-based adaptation of intelligent controllers of solar collector fields. In I. Troch and F. Breiteneker, editors, *Proceedings of 7th Vienna Symposium on Mathematical Modelling, February 14-17, 2012, Vienna, Austria, Part 1*, volume 7, pages 979–984. IFAC, 2012.
- [55] E. Juuso. *Integration of intelligent systems in development of smart adaptive systems: linguistic equation approach*. PhD thesis, University of Oulu, 2013. 258 pp., <http://urn.fi/urn:isbn:9789526202891>.
- [56] E. K. Juuso and T. Ahola. Case-based detection of operating conditions in complex nonlinear systems. In M. J. Chung and P. Misra, editors, *Proceedings of 17th IFAC World Congress, Seoul, Korea, July 6-11, 2008*, volume 17, pages 11142–11147. IFAC, 2008.
- [57] E. K. Juuso. Advanced data analysis in condition-based operation and maintenance. In *WCCM 2017 - 1st World Congress on Condition Monitoring 2017*, volume 2, pages 750–761, Red Hook, NY, 2017. Curran Associates.
- [58] E. K. Juuso. Modelling and simulation in adaptive intelligent control. *Simulation Notes Europe SNE*, 26(2):109–116, 2016.
- [59] A. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations and system approaches. *AICom - Artificial Intelligence Communications*, 7(1):39–59, 1994.
- [60] I. Watson. Case-based reasoning is a methodology not a technology. *Knowledge-Based Systems*, 12:303–308, 1999.
- [61] E. K. Juuso. Generalised statistical process control GSPC in stress monitoring. *IFAC-PapersOnline*, 48(17):207–212, 2015.
- [62] E. K. Juuso. Recursive data analysis and modelling in prognostics. In *12th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, CM 2015 - MFPT 2015, 9-11 June 2015, Oxford, UK*, pages 560–567. BINDT, 2015. ISBN: 978-1-5108-0712-9.
- [63] E. K. Juuso and L. J. Yebra. Smart adaptive control of a solar collector field. In *IFAC Proceedings Volumes (IFAC-PapersOnline)*, volume 19, pages 2564–2569, 2014.
- [64] S. Lahdelma and E. Juuso. Signal processing and feature extraction by using real order derivatives and generalised norms. Part 2: Applications. *International Journal of Condition Monitoring*, 1(2):54–66, 2011.
- [65] E. K. Juuso. Advanced prognostics based on intelligent data analysis. In *WCCM 2017 - 1st World Congress on Condition Monitoring 2017*, volume 2, pages 782–794, Red Hook, NY, 2017. Curran Associates.
- [66] E. K. Juuso. Intelligent performance analysis with a natural language interface. *Management Systems in Production Engineering*, 25(3):168–175, 2017.