NONLINEAR SCALING IN SMART ADAPTIVE MODELLING

Esko K. Juuso

Control Engineering, Faculty of Technology, University of Oulu, Finland

esko.juuso@oulu.fi

ABSTRACT
Modelling methodologies provide a good basis for the integration of intelligent systems. Small, specialised systems have a large number of feasible solutions, but developing truly adaptive, and still understandable, systems for highly complex systems require domain expertise and more compact approaches at the basic level. The nonlinear scaling approach extends the application areas of linear methodologies to nonlinear modelling and reduces the need for decomposition with local models. The close connection to the fuzzy set systems provides a good basis for understandable models. Data-based methodologies are suitable for developing smart adaptive applications. Complex problems are solved level by level to keep the domain expertise as an essential part of the solution.
Keywords: nonlinear systems, linguistic equations, smart adaptive systems, statistical analysis

1. INTRODUCTION

Models are understood as relationships between variables and used to predict of properties or behaviours of the system. Variable interactions and nonlinearities are important in extending the operation areas (Juuso 2014). Phenomenological models based on physics, chemistry and mathematics require domain expertise (Figure 1). Linear methodologies extended with principal components (Jolliffe 2002; Gerlach et al. 1979) and semi-physical models (Ljung 1999) provide a feasible solution for many applications. Nonlinearities have been handled commonly with interaction and quadratic terms (Box and Wilson 1951). Artificial neural networks (ANNs) starting from (Rummelhart et al. 1986) continue this by using more complex architectures.
Knowledge-based information can be handled with fuzzy set systems introduced by Zadeh (1965): numerous methodologies have been developed, see (Takagi and Sugeno 1985; Driankov et al. 1993; Dubois et al. 1999), and combined with neural networks (Fullér 2000). Different fuzzy approaches can be efficiently combined (Juuso 2014).
First order ordinary differential equations are solved by numerical integration and special solutions have been developed for identification (Ljung 1999). These approaches, which are also used in ANNs and fuzzy set systems (Babuška and Verbruggen 2003), define structures for hybrid dynamic models (Figure 1).
Local models need to be combined in complex systems (Sontag 1981; Ljung 2008; Jardine et al. 2006).

The linguistic equation (LE) approach originates from fuzzy set systems (Juuso and Leiviskä 1992): rule sets are replaced with equations, and meanings of the variables are handled with scaling functions which have close connections to membership functions (Juuso 1999a). The nonlinear scaling technique is needed in constructing nonlinear models with linear equations (Juuso 2004c). Constraints handling (Juuso 2009) and data-based analysis (Juuso and Lahdelma 2010), improve possibilities to update the scaling functions recursively (Juuso 2011; Juuso and Lahdelma 2011). The LE approach together with knowledge-based systems, neural networks and evolutionary computation form the computational intelligence part (Figure 1).
Three levels of smart adaptive systems (SAS) identified in (Anguita 2001):

1. adaptation to a changing environment;
2. adaptation to a similar setting without explicitly being ported to it;
3. adaptation to a new or unknown application.

Smart use of intelligence by integrating specific intelligent systems is essential in development of complex adaptive applications.

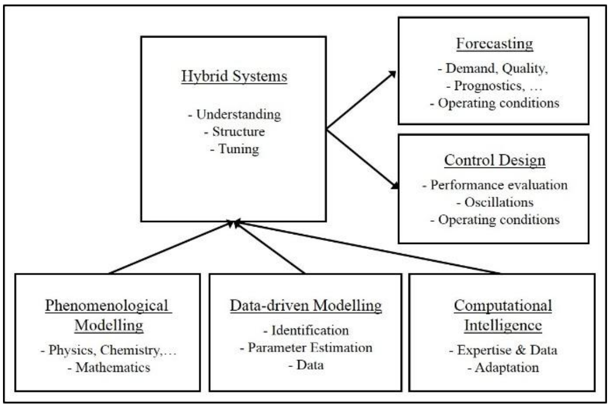

Figure 1: Methodologies and application types of modelling and simulation, modified from (Juuso 2004b)

Technically, an automatic black box modelling could be possible in various Big Data problems by using combinations of these methodologies. The domain expertise is an essential part in integrated solutions to understand and assess the applicability.
This paper classifies modelling methodologies and focuses on the nonlinear scaling and integrates the LE approach in developing modelling applications for complex systems. Various applications are shortly discussed.

2. MODELLING METHODOLOGIES

Steady-state modelling with linear and nonlinear methodologies are the basis of modelling. Dynamic modelling introduces additional model structures. Decomposition is needed to extend the solutions to multiple operating conditions.

2.1. Steady-state modelling

The steady-state simulation 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 $x = (x_1, x_2, \dots, x_m)$.

2.1.1. Linear methodologies

Statistical modelling in its basic form uses linear regression for solving coefficients for a linear function. Linear methodologies are suitable for large multivariable systems. Linear methodologies can be extended to more complex areas by handling inputs. Principal components compress the data by reducing the number of dimensions with minor loss of information (Jolliffe 2002). Partial least squares regression (PLS) is an extension of these ideas (Gerlach et al. 1979). Semi-physical models of inputs are important in linear modelling, see (Ljung 1999).

2.1.2. Nonlinear methodologies

Exponential and logarithmic functions are suitable for the modelling of steep changes.

In the response surface method (RSM), the coefficients of linear, interactive and quadratic terms are obtained by the multiple regression for several input variables (Box and Wilson 1951). The models can include several multiple input, single output models.

Artificial neural networks (ANNs) are input output models: the most popular architecture, multilayer perceptron (MLP) has a very close connection to the backpropagation learning (Rummelhart et al. 1986). The response of a neuron is

$$y_i = F_i\left(\sum_{j=1}^{m} w_{ij} p_j + b_i\right), \qquad (1)$$

where w_{ij} is the weight factor of the element p_j in the input vector of the neuron i and b_i a scalar bias. Linear networks correspond to the models with linear terms in RSM models.

Normalisation and principal components are essential in various applications. Various extensions of the principal component analysis (PCA) are referred in (Jolliffe 2002). A function expansion

$$y_i = \sum_{k=1}^{m_f} w_{ik} F_k(\vec{x}) = \sum_{k=1}^{m_f} w_{ik} f(\vec{\beta}_k \bullet (\vec{x} - \vec{\gamma}_k)), \qquad (2)$$

with some basis functions $F_k(\vec{x}),\ k = 1, \dots, m_f$, provides a flexible way to present several types of black box models (Ljung 2008). The functions are generated from one and the same function characterised by the scale (dilation) parameters $\vec{\beta}_k$ and location (translation) parameters $\vec{\gamma}_k$. The expansion can contain, for example, radial basis functions, one-hidden-layer sigmoidal neural networks, neurofuzzy models, wavenets, least square support vector machines (SVMs), see (Ljung 1999).

2.1.3. Knowledge-based methodologies

Zadeh (1965) presented the fuzzy set theory to form a conceptual framework for linguistically represented knowledge. Origins of the fuzzy logic are in approximate reasoning and the connection of fuzzy rule-based systems and expert systems is evident (Dubois et al. 1999). Membership functions provide a key to expand expert systems. In linguistic fuzzy models, both the antecedent and consequent are fuzzy propositions (Driankov et al. 1993). In Takagi-Sugeno (TS) fuzzy models, each consequent is a crisp function of the antecedent variables~x, can be interpreted in terms of local models (Takagi and Sugeno 1985). The extension principle generalises arithmetic operations if the inductive mapping is a monotonously increasing function of the input. Type-2 fuzzy models introduced by Zadeh in 1975 take into account uncertainty about the membership function (Mendel 2007).

In neurofuzzy systems, fuzzy neurons combine the weight factors and the inputs. The activation function is handled with the extension principle from the fuzzy input, which is obtained by the fuzzy arithmetics. Different combinations with fuzzy and crisp weight factors and elements can be used in these models (Fullér 2000). Neurofuzzy systems can be represented with (2).

2.2. Dynamic modelling

First order ordinary differential equations (ODEs),

$$\frac{dx}{dt} = f(t, x), \qquad (3)$$

are solved by integration:

$$x = \int_0^T f(t, x)dt + x_0, \qquad (4)$$

where T is the time period for integration and x_0 the initial condition. The function $f(t, x)$ can be linear or nonlinear. Additional algebraic equations are needed, e.g. for handling material.

Linear methodologies are used in the time series modelling is to fit the waveform data to a parametric time series model and extract features based on this parametric model. For parametric models, the output y at time t is computed as a linear combination of past inputs u and past outputs *y*. The signal values should be chosen according to the appropriate time delays. The time step is not adapted in these models. The number of delayed inputs and outputs is usually referred to as the model order(s). Several types of models can be obtained from the general parametric model

$$A_q(q)\, y(t) = \frac{B_q(q)}{F_q(q)} u(t-n_k) + \frac{C_q(q)}{D_q(q)} e(t), \tag{5}$$

where $A_q(\cdot)$, $B_q(\cdot)$, $C_q(\cdot)$, $D_q(\cdot)$ and $F_q(\cdot)$ are polynomials of the delay operator q^{-1}. The orders of these polynomials are n_a, n_b, n_c, n_d and n_f, respectively, and the number n_k is the number of delay from input to output. Autoregressive (ARX) and autoregressive moving average (ARMAX) models, both with the exogeneous input u , are special cases of (3). If the data are presented as a time series, which has no input channels and only one output channel *y*, then ARX and ARMAX models become AR and ARMA models, respectively. (Ljung 1999)

Fuzzy and neural models are based on the same structures. The most common structure for the input-output models is the NARX /Nonlinear AutoRegressive with eXogenous input model, in which the input and output values are chosen according to appropriate system orders, as in the ARX model. The regressor vector consists of a finite number of past inputs and outputs (Babuška and Verbruggen 2003). Another possibility is to use recurrent networks, e.g. the Elman networks are two-layer feedforward networks, with the addition of a feedback connection from the output of the hidden layer to its input (Elman 1990).

2.3. Decomposition

Linear models are approximations of a nonlinear system in different neighbourhoods. Composite local models combine local linear models to construct a global model. If the partitioning is based on a measured regime variable, the partitioning can be used in weighting the local models. Linear parameter varying (LPV) models, where the matrices of the state-space model depend on an exogeneous variable measured during the operation, are close related to local linear models (Ljung 2008). Piecewise affine (PWA) systems are based on local linear models, more specifically in a polyhedral partition (Sontag 1981). The models can be state-space models or parametric models. The model switches between different modes as the state variable varies over the partition (Ljung 2008).

Gradual transitions between models can be done with weight factors in function expansion (2) or with fuzzy rules in TS fuzzy models. Common model structures are beneficial in applications. In prognostics, calculations can be based on a time-dependent proportional hazard model (PHM)

$$h(t) = h_0(t) \exp\left(\sum_{j=1}^{m} \eta_j \, x_j(t)\right), \tag{6}$$

where changing operating conditions are handled with exponential functions: $h_0(t)$ is a baseline hazard function, $x_j(t),\ j = 1,\dots,m$, are covariates which are functions of time and $\eta_j,\ j=1,\dots,m$ are coefficients. The baseline hazard function $h_0(t)$ can be in non-parametric or parametric form, e.g. a Weibull hazard function, which is the hazard function of the Weibull distribution. The covariates $x_j(t),\ j = 1,\dots,m$ are any condition variables such as health indicators and features in condition monitoring. Maximum likelihood estimation is usually used to build a PHM from event data and condition monitoring data. Modelling a PHM is more or less like the process of regress analysis: a set of significant covariates is finally found and only these significant observations for the ‘‘dependent’’ variable *h(t)*, instead of observations, are available as event data. (Jardine et al. 2006)

Cascade modelling divides the problem into sequential parts to further alleviate the problem of parameters: TS fuzzy models use fuzzy reasoning for weighting local linear models; radial basis networks are linear combinations of the outputs of the RBF; learning vector quantization (LVQ) combines a competitive layer with a linear model. Neurofuzzy systems can be constructed as sequential combinations of neural and fuzzy parts. Variable grouping is important in cascade model structures.

The need for decomposition is evident since the modelling methodologies have limitations, like operating areas of linear methodologies and highly complex structures of nonlinear systems. Fuzzy set systems are natural tools in in the management of the decomposed models. Several fuzzy modelling approaches are combined in Figure 2: fuzzy arithmetics is suitable both for processing the fuzzy inputs and outputs of the rule-based fuzzy set system; fuzzy inequalities produce new facts; fuzzy relations can be represented as sets of alternative rules, where each rule has a degree of membership (Juuso 2014).

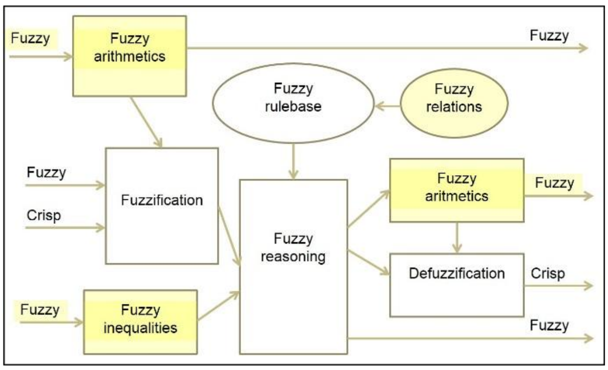

Figure 2: Combined Fuzzy Modelling (Juuso 2014)

3. LINGUISTIC EQUATION MODELS

The key of the linguistic equation (LE) methodology is the nonlinear scaling developed to extract the meanings of variables from measurement signals.

3.1. Nonlinear scaling

Normalisation or scaling of the data is needed since measurements with considerably different magnitudes cause problems in modelling. The nonlinear scaling extends modelling to various statistical distributions and allows recursive tuning.

3.1.1. Generalised norms

Arithmetic mean and standards deviation, which are the key features in statistical analysis, are special cases of generalised norms

$$\left\| {}^{\tau}M_j^p \right\|_p = ({}^{\tau}M_j^p)^{1/p} = [(\frac{1}{N}\sum_{i=1}^{N}(x_j)_i^p)]^{1/p}, \tag{7}$$

where the order of the moment $p \in \Re$ is non-zero. The signal is measured continuously for the analysis is based on consecutive equally sized samples. Duration of each sample is called sample time, denoted τ. The number of signal values $N = \tau N_s$, where N_s is the number of signal values which are taken in a second. This norm, which has the same dimensions as the signal x_j. The generalised norms were introduced for condition monitoring (Lahdelma and Juuso 2011a, 2011b). The norm values increase monotonously with increasing order if all the signals are not equal.

3.1.2. Scaling functions

Scaling is based on the z-score

$$p_j = \frac{x_j - c_j}{\Delta c_j} = \frac{x_j - \left\| {}^{\tau}M_j^1 \right\|_1}{\left\| {}^{\tau}M_j^2 \right\|_2}, \tag{8}$$

which is calculated about the arithmetic mean, $c_j = \left\| {}^{\tau}M_j^1 \right\|_1$, by the standard deviation $\Delta c_j = \left\| {}^{\tau}M_j^2 \right\|_2$. The arithmetic mean and standard deviation are optimal if the data sample comes from a normal distribution. This approach is sensitive to data entry errors, e.g. outliers. The geometric mean and harmonic mean are useful when the sample is distributed lognormal or heavily skewed. The median and trimmed mean are two measures that are resistant (robust) to outliers.

The z-score based linear solutions are extended to asymmetric nonlinear cases by two second order polynomials. The parameters of the polynomials are defined with five parameters corresponding the operating point c_j and four corner points of feasible range represented by a fuzzy number: core $\lfloor (c_l)_j, (c_h)_j \rfloor$ and support $\lfloor \min(x_j), \max(x_j) \rfloor$ (Juuso 2004c).

These points can be defined manually or obtained from data by using generalised norms (7) and moments

$$\gamma_k^p = \frac{\sum_{i=1}^{N}((x_j)_i - \left\| {}^{\tau}M_j^p \right\|_p)^k}{N\sigma_j^k}, \tag{9}$$

where the order k is a positive integer. The kth moment is generalised by calculating it about the generalised norm (5), and normalised by the standard deviation σ_j which is calculated about the origin. The operating point c_j is the central tendency taken from the point, where the skewness γ_3^p changes from positive to negative, i.e. $\gamma_3^p = 0$. Then the data set is divided into two parts, a lower part and an upper part, and then the same analysis is done for these two data sets. The estimates of the corner points, $(c_l)_j$ and the operating point $(c_h)_j$, are the points where the skewness goes to zero. The iteration is performed with generalised norms (5) to get the corresponding order of norm. (Juuso and Lahdelma 2010)

The shape of the second order polynomials are defined by the ratios

$$\begin{aligned} \alpha_j^- &= \frac{(c_l)_j - \min(x_j)}{c_j - (c_l)_j} = \frac{(c_l)_j - \min(x_j)}{\Delta c_j^-} \\ \alpha_j^+ &= \frac{\max(x_j) - (c_h)_j}{(c_h)_j - c_j} = \frac{\max(x_j) - (c_h)_j}{\Delta c_j^+}, \end{aligned} \tag{10}$$

which are limited to the range $\left[\frac{1}{3}, 3\right]$ by resizing the core or the support. Then the scaling functions are monotonously increasing throughout the feasible range. (Juuso 2009).

3.1.3. Recursive tuning

The computation of the generalised norms can be divided into the computation of equal sized sub-blocks, i.e. the norm for several samples can be obtained as the norm for the norms of individual samples. The same result is obtained using the moments

$$\left\| {}^{K_S\tau}M_\alpha^p \right\|_p = ({}^{K_S\tau}M_\alpha^p)^{1/p} = \left\{ \frac{1}{K_S}\sum_{i=1}^{K_S}\left[({}^{\tau}M_\alpha^p)_i^{1/p}\right]^p \right\}^{1/p}, \tag{11}$$

where K_S is the number of samples $\left\{x_i^{(\alpha)}\right\}_{i=1}^{N}$. Each sample has N signal values. Weights can be introduced by means of density functions. It is useful to calculate the norms from short samples since the number of signal values per second is quite high. The sample time τ is an essential parameter in the calculation of moments and norms. (Juuso and Lahdelma 2010)

For the arithmetic mean, the calculation based on sub-blocks is the normal practice in automation systems. This approach can be extended to all generalised norms.

3.2. Interactions

The linguistic equation (LE) models are linear equations

$$\sum_{j=1}^{m} A_{ij} X_j + B_i = 0, \qquad (12)$$

where X_j is the linguistic level for the variable $j, j=1..m$. Each equation i has its own set of interaction coefficients $A_{ij}, j=1..m$. The bias term B_i was introduced for fault diagnosis systems. (Juuso 2004c)

The scaled values can be used in the same way as any other variables since the nonlinear scaling extends the normalisation and the z-score approach. The dimensionless variables are suitable for various nonlinear methodologies discussed in Section 2. However, linear methodologies have been sufficient in various applications.

3.2.1. Steady-state LE models

A multiple input single output model is represented by

$$x_{out} = f_{out}\left(-\frac{\sum_{j=1, j\neq out}^{m} A_{ij} f_j^{-1}(x_j(t-n_j)) + B_i}{A_{i\,out}}\right) \qquad (13)$$

where the functions f_j and f_{out} are scaling functions of the variables j and out, respectively. An appropriate time delay n_j needs to be taken into account for each variable. The scaling functions can be understood as new type basis functions to be used in expansion (2).

3.2.2. Dynamic LE models

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 can include several inputs and outputs originating from a single variable (Juuso 2004c). External dynamic models provide the dynamic behaviour, and LE models are developed for a defined sampling interval in the same way as in various identification approaches defined by (5). 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: $n_a = 1$, $n_b = 1$, $n_c = 0$, $n_d = 0$ and $n_f = 0$ in the parametric models (3) resulting

$$Y(t) + a_1 Y(t-1) = b_1 U(t-n_k) + e(t) \qquad (14)$$

where Y and U are scaled variables and coefficients a_1 and b_1 coefficients of the polynomials $A_q(\cdot)$, $B_q(\cdot)$.

Alternatively, a new value for the derivative (3) can be calculated with a LE model and then integrated by (4). This approach allows the adaptation of the integration step.

3.3. Uncertainty in LE models

The LE approach originates from the fuzzy set systems which keeps the connections of the methodologies strong. Compact LE models provide a good basis for multimodel systems, where local LE models are combined with fuzzy logic, to handle transitions between models, some special situations and uncertainty with fuzzy set systems. Fuzzy reasoning is an important part of the LE based fault diagnosis and the decision making in the recursive adaptation.

The coefficients of the model (12) and the parameters of scaling functions can be represented as fuzzy numbers, which are used in the calculations with the extension principle and fuzzy arithmetics (Juuso 2014).

3.4. Smart adaptive LE models

Recursive updates of the norm values discussed in Section 3.1.3 provide a real time solution to the adaptation to a changing environment (SAS level 1). Strong changes in statistical distributions can be taken into account by obtaining the orders of the norm, which realises the adaptation to a similar setting without explicitly being ported to it (SAS level 2). Similar settings are understood as unchanged interaction models. The settings are based on the analysis of the interactions (Section 3.2), which is the key in the adaptation to a new or unknown application (SAS level 3).

4. APPLICATIONS

Nonlinear scaling forms the basis for the LE modelling: an important benefit of the linear approach is that the models can be inverted, technically to any direction. The compact basic solution makes extensions to dynamic and case-based systems possible. Complex models for steady-state and dynamic systems can be built with the cascade and interactive structures.

4.1. Steady-state LE models

Steady-state LE models are mainly used in adaptation and feedforward control (Table 1). In most cases, the models include only a single equation (13). The first LE model developed for designing submerged arc furnaces was an exception which used well known relations represented by five equations (Juuso and Leiviskä 1992). A steady-state LE model was developed in an early control application from the process measurements of a lime kiln (Juuso et al. 1997). The working point model presented in (Juuso et al. 1998) is still an essential part of the model-based LE control of a solar power plant (Juuso and Yebra 2013). For continuous cooking, a LE model has been developed for predicting the Kappa number, which is widely used quality variable (Leiviskä et al.

2001). Stress-cycle (S-N) curves, also known as Wöhler curves, are represented by a linguistic equation (Juuso and Ruusunen 2013).

Table 1: Steady-state LE model applications

Case	Application area
Electric furnace	Decision support for process design
Lime kiln	Feedforward control of fuel feed for changing capacity
Solar collector field	Adaptation based on steady-state working point model
Continuous cooking	Forecasting model for quality control
Fatigue	Stress contribution obtained from LE based stress-cycle curve
Water treatment	Feedforward control
Wastewater treatment	Forecasting

4.2. Dynamic LE models
The basic dynamic LE model is represented by (14). The approach was first tested in a gas furnace data provided by (Box & Jenkins 1970). The dynamic models of the solar plant are based on test campaigns, which cannot be planned in detail because of changing weather conditions (Juuso 2003a). The basic dynamic flotation model is the core of the quality indicator in water treatment (Ainali et al. 2002, Joensuu et al. 2005). A dynamic LE model has been used for fatigue prediction in (Juuso and Ruusunen 2013). In all these models, only one equation is needed. The applications are indirect measurements and controller tuning (Table 2).

Table 2: Dynamic LE model applications

Case	Application area
Gas furnace	Modelling
Solar collector field	Controller tuning for oil flow
Fatigue	Forecasting fatigue risk from stress contributions
Water treatment	Water quality indicator Controller tuning for two chemicals

4.3. Decomposition in LE models
The multimodel LE system can include several submodels and complex interactions (Table 3). All basic models are represented by (14).
The model with a fuzzy decision module was first used for a lime kiln (Juuso 1999b) and then for a solar thermal power plant Juuso (2003a). The lime kiln model had six operating areas defined by the production level and the trend of the fuel feed (increasing, decreasing). The model of the collector field includes four operating areas: start-up, low, normal and high operation. For handling special situations in the solar plant, additional fuzzy models have been developed by using the Fuzzy–ROSA method (Juuso et al. 2000).
Interactive dynamic models were needed in several cases: batch cooking (Juuso 2003b), fluidised bed granulator (Mäki et al. 2004), industrial fed-batch fermenter (Saarela et al. 2003) and wastewater treatment. Linguistic equations, neural networks and fuzzy modelling with several variants have been compared by using the process data obtained from the fed-batch fermenter.

Table 3: Dynamic LE model applications

Case	Application area
Lime kiln	Fuel quality Controller tuning by using multiple models
Solar collector field	Controller tuning for oil flow by using multiple models
Batch cooking	On-line forecasting by using three interactive models: alkali, lignin and dissolved solids
Fluidised bed granulation	Forecasting by using three interactive models: temperature, humidity and granular size
Fed-batch fermentation	On-line forecasting by using submodels of three growth phases, each including three interactive models
Wastewater treatment	Detection of operating conditions and trend analysis by using three submodels: load, treatment and settling
Condition monitoring	Prognostics with recursive tuning

4.4. Distributed parameter LE models
In distributed parameter models, the solar collector field is divided into modules, where the dynamic LE models are applied in a distributed way (Juuso 2004b). The same single equation model is used in all modules. Element locations for partial differential equations (PDEs) are defined by the flow rate. In cloudy conditions, the heating effect can be strongly uneven.

5. CONCLUSIONS
The nonlinear scaling approach extends the application areas of linear methodologies to nonlinear modelling: the meanings of variables and interactions are analysed sequentially. Local nonlinear models reduce the need for decomposition with local models is needed. The close connection to the fuzzy set systems provides a good basis for understandable models. Data-based methodologies are suitable for developing smart adaptive applications. Big Data problems are solved level by level to keep the domain expertise as an essential part of the solution. The basic models are compact and additional properties, including dynamics, uncertainty and decomposition are included if needed.

ACKNOWLEDGMENTS
The combined approach has been developed within the research program “Measurement, Monitoring and Environmental Efficiency Assessment (MMEA)” on the basis of several earlier and parallel projects.

REFERENCES

Ainali I., Piironen M. and Juuso E., 2002. Intelligent water quality indicator for chemical water treatment unit. Proceedings of SIMS 2002 - the 43rd Scandinavian Conference on Simulation and Modelling, pp. 247–252, September 26-27, Oulu, Finland.

Anguita D., 2001. Smart adaptive systems - state of the art and future directions for research. Proceedings of Eunite 2001 - European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems, pp. 1-4. July 13-14, Tenerife, Spain.

Babuška R. and Verbruggen H., 2003. Neuro-fuzzy methods for nonlinear system identification, Annual Reviews in Control, 27(1), 73-85.

Box G.E.P. and Jenkins G.M., 1970. Time Series Analysis, Forecasting and Control, San Francisco: Holden Day.

Box G.E.P. and Wilson K.B., 1951. On the experimental attainment of optimum conditions. Journal of the Royal Statistical Society, Series B, 13(1), 1–45.

Driankov D., Hellendoorn H. and Reinfrank M., 1993. An Introduction to Fuzzy Control. Berlin, Germany: Springer.

Dubois D., Prade H. and Ughetto L., 1999. Fuzzy logic, control engineering and artificial intelligence. In: Verbruggen H.B., Zimmermann H.-J. and Babuska R., eds., Fuzzy Algorithms for Control, International Series in Intelligent Technologies, Boston: Kluwer, pp. 17–57.

Elman J.L., 1990. Finding structure in time, Cognitive Science, 14(2), 179-211.

Fullér R. 2000. Introduction to Neuro-Fuzzy Systems. Advances in Soft Computing. Springer.

Gerlach R.W., Kowalski B.R. and Wold H.O.A., 1979. Partial least squares modelling with latent variables. Analytica Chimica Acta, 112(4), 417–421.

Jardine A.K.S., Lin D. and Banjevic D., 2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mechanical Systems and Signal Processing, 20(7), 1483–1510.

Joensuu I., Piironen M. and Juuso E., 2005. Dynamic simulator for dosing of water treatment chemicals. Proceedings of European Symposium on Computer Aided Process Engineering-15 (Escape-15), pp. 301–306, Computer-aided chemical engineering, 20A, May 29 - June 1, Barcelona, Spain.

Jolliffe I.T., 2002. Principal Component Analysis. 2nd edition, New York: Springer.

Juuso E.K., 1999a. Fuzzy control in process industry: The linguistic equation approach. In: Verbruggen H.B., Zimmermann H.-J. and Babuska R, eds., Fuzzy Algorithms for Control, International Series in Intelligent Technologies, Boston: Kluwer; pp. 243–300. Doi: 10.1007/978-94-011-4405-6 10

Juuso E.K., 1999b. Intelligent dynamic simulation of a lime kiln with linguistic equations. ESM’99: Modelling and Simulation: A tool for the Next Millenium, 13th European Simulation Multiconference, pp. 395–00, June 1–4, Warsaw, Poland.

Juuso E.K., 2003a. Intelligent dynamic simulation of a solar collector field. In: Verbraeck A. and Hlupic V., eds., Simulation in Industry, 15th European Simulation Symposium ESS 2003, SCS, Gruner Druck, Erlangen, Germany, pp. 443–449.

Juuso E.K., 2003b. Intelligent dynamic simulation of batch cooking. 44th Scandinavian Conference on Simulation and Modeling SIMS 2003, pp. 163–169, September 18 -19, Västerås, Sweden.

Juuso E. K., 2004a. Adaptive intelligent modelling in hybrid process simulation. Proceedings of the fifth EUROSIM Congress on Modelling and Simulation, 6 pp, Sept. 6-10, Paris, France.

Juuso E.K., 2004b. Dynamic simulation of a solar collector field with intelligent distributed parameter models. Proceedings of SIMS 2004 - the 45th Scandinavian Conference on Simulation and Modelling, pp. 141–153, September 23-24, Copenhagen, Denmark.

Juuso E.K., 2004c. Integration of intelligent systems in development of smart adaptive systems. International Journal of Approximate Reasoning, 35(3), 307–337. DOI: 10.1016/j.ijar.2003.08.008

Juuso E.K., 2009. Tuning of large-scale linguistic equation (LE) models with genetic algorithms. In: Kolehmainen M., ed., Revised selected papers of the International Conference on Adaptive and Natural Computing Algorithms - ICANNGA 2009, Kuopio, Finland, Lecture Notes in Computer Science, vol. LNCS 5495, Heidelberg: Springer-Verlag, pp. 161–170. DOI: 10.1007/978-3-642-04921-7_17.

Juuso E.K., 2011. Recursive tuning of intelligent controllers of solar collector fields in changing operating conditions. Proceedings of the 18th World Congress The International Federation of Automatic Control; 2011 August 28 – September 2; Milano, pp. 12282–12288. doi: 10.3182/20110828-6-IT-1002.03621

Juuso E.K., 2014. Intelligent Methods in Modelling and Simulation of Complex Systems. Simulation Notes Europe SNE, 24(1), 1-10. Selected SIMS 2013. DOI: 10.11128/sne.24.on.102221

Juuso E. and Lahdelma S., 2010. Intelligent scaling of features in fault diagnosis. Proceedings of the 7th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Vol. 2, pp. 1358-1372, 22-24 June, Stratford-upon-Avon, UK.

Juuso E. and Lahdelma S., 2011. Intelligent trend indices and recursive modelling in prognostics. Proceedings of the 8th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Vol. 1, pp. 440-450, June 20-22, Cardiff, UK.

Juuso E.K. and Leiviskä K., 1992. Adaptive expert systems for metallurgical processes. In: Jämsä-Jounela S.L., Niemi A.J., eds., Expert Systems in Mineral and Metal Processing, Proceedings of the IFAC Workshop, August 26-28, 1991, Espoo, Finland; IFAC Workshop Series, Number 2, Oxford, UK : Pergamon, pp. 119–124.

Juuso E. and Ruusunen M., 2013. Fatigue prediction with intelligent stress indices based on torque measurements in a rolling mill. Proceedings of CM 2013 and MFPT 2013, Vol. 1, pp. 460-471, 19-20 June, Krakow, Poland.

Juuso E.K. and Yebra L., 2013. Model-based intelligent control of a solar energy collector field, Proceedings of 8h EUROSIM Congress on Modelling and Simulation, pp. 513–518, 10-13 September, 2013, Cardiff, UK, doi: 10.1109/EUROSIM.2013.92.

Juuso E.K., Balsa P. and Valenzuela L., 1998. Multilevel linguistic equation controller applied to a 1 MW solar power plant. Proceedings of the ACC'98, vol. 6, pp. 3891–3895, June 24–26, Philadelphia, PA.

Juuso E.K., Schauten D., Slawinski T. and Kiendl H. 2000. Combination of linguistic equations and the fuzzy-rosa method in dynamic simulation of a solar collector field. Proceedings of TOOLMET 2000 Symposium - Tool Environments and Development Methods for Intelligent Systems, pp. 63–77, April 13-14, Oulu, Finland.

Lahdelma S. and Juuso E., 2011a. Signal processing and feature extraction by using real order derivatives and generalised norms. Part 1: Methodology. The International Journal of Condition Monitoring, 1(2), 46–53. DOI: 10.1784/204764211798303805.

Lahdelma S. and Juuso E., 2011b. Signal processing and feature extraction by using real order derivatives and generalised norms. Part 2: Applications. The International Journal of Condition Monitoring, 1(2), 54–66. DOI: 10.1784/204764211798303814.

Leiviskä K., Juuso E. and Isokangas A., 2001. Intelligent modelling of continuous pulp cooking. In: Leiviskä K., ed., Industrial Applications of Soft Computing, Heidelberg: Springer, pp. 147–158.

Ljung L., 1999. System Identification - Theory for the User, 2nd edition, Upper Saddle River, N.J.: Prentice Hall.

Ljung L., 2008. Perspectives on system identification. Plenary papers, milestone reports & selected survey papers, 17th IFAC World Congress, pp. 47-59, July 6-11, 2008, Seoul, Korea.

Mendel J.M., 2007. Advances in type-2 fuzzy sets and systems, Information Sciences, 177(1), 84–110.

Mäki T., Juuso E. and Leiviskä K., 2004. Fuzzy modelling and dynamic simulation of a fluidised bed granulator. Proceedings of AFNC'04 - the 2nd IFAC Workshop on Advanced Fuzzy/Neural, pp. 133–138, September 16-17, Oulu, Finland.

Rummelhart D.E., Hinton G.E. and Williams R.J., 1986. Learning internal representations by error propagation. In: Rummelhart D.E. and McClelland J., eds., Parallel Data Processing, Cambridge, MA 4: M.I.T. Press, pp. 318–362.

Saarela U., Leiviskä K., Juuso E. and Kosola A., 2003. Modelling of a fed-batch enzyme fermentation process. IFAC International Conference on Intelligent Control Systems and Signal Processing, April 8-11, Faro, Portugal.

Sontag E., 1981. Nonlinear regulation: The piecewise linear approach. IEEE Transactions Automatic Control, 26(2), 346–358.

Takagi T. and Sugeno M., 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions on Systems, Man, and Cybernetics, 15(1), 116–132.

Zadeh L.A., 1965. Fuzzy sets, Information and Control, 8(June), 338–353.

AUTHORS BIOGRAPHY

Esko K. Juuso is the Team Leader of Industrial process and Energy in Research Community of Automation, Faculty of Technology, University of Oulu. He has D.Sc.Tech. in Control and Systems Engineering and MSc in Technical Physics, both from University of Oulu. He has worked several years in metal industry and as a teacher and a project manager in several research projects on intelligent systems applications in various fields of industry. He is the developer of the linguistic equation (LE) approach and the nonlinear scaling methodology, which are currently used in various applications. His research interests are in the modelling and control of industrial processes with a special emphasis on combining intelligent control, fault diagnosis and performance monitoring into smart adaptive systems. He has published more than 300 technical and scientific publications. He has a long experience in European networks of excellence and scientific organizations in simulation, control and condition monitoring. He is currently the President of Juuso is the President of the European Federation of European Simulation Societies (EUROSIM) and a Member of the Management Committee of the International Society of Condition Monitoring (ISCM).