

Nature-inspired adaptive Architecture for Soft Sensor Modelling

Petr Kadlec and Bogdan Gabrys

Computational Intelligence Research Group, School of Design, Engineering & Computing
Bournemouth University, Poole House, Talbot Campus

Poole, BH12 5BB United Kingdom

Phone: +44-1202-965491, Fax: +44-1202-595314

email: {pkadlec, bgabrys}@bournemouth.ac.uk

ABSTRACT: This paper gives a general overview of the challenges present in the research field of Soft Sensor building and proposes a novel architecture for building of Soft Sensors, which copes with the identified challenges. The architecture is inspired and making use of nature-related techniques for computational intelligence. Another aspect, which is addressed by the proposed architecture, are the identified characteristics of the process industry data. The data recorded in the process industry consist usually of certain amount of missing values or sample exceeding meaningful values of the measurements, called data outliers. Other process industry data properties causing problems for the modelling are the collinearity of the data, drifting data and the different sampling rates of the particular hardware sensors. It is these characteristics which are the source of the need for an adaptive behaviour of Soft Sensors. The architecture reflects this need and provides mechanisms for the adaptation and evolution of the Soft Sensor at different levels. The adaptation capabilities are provided by maintaining a variety of rather simple models. These particular models, called paths in terms of the architecture, can for example focus on different partition of the input data space, or provide different adaptation speeds to changes in the data. The actual modelling techniques involved into the architecture are data-driven computational learning approaches like artificial neural networks, principal component regression, etc.

KEYWORDS: Soft Sensors, process industry, adaptive systems, evolving systems, ANN, genetic algorithms, nature-inspired computational learning method

INTRODUCTION

There are several reasons for the interest of the process industry in the development of Soft Sensors. One of the most important reasons is the difficulty in the development of first principle models. First principle models attempt to use the knowledge of the physical and chemical laws for building models of the processes. Such principles may be for example the energy and mass preservation laws, etc. Unfortunately, the processes are usually much too complicated to be fully described by this kind of models. There is also a lot of external influences, e.g. the environmental temperature, the purity of the educts, the abrasion of different mechanical parts, which make the modelling of the exact process dynamics very difficult. For these reasons the models have to be abstracted from the reality and focus on the important aspects of the process. Despite the fact that the abstraction level is usually rather high, the models may still be very complicated and their prediction quality may not fulfil the required standards.

Another way to make the predictions is to use the data, which is measured during the operation of the process and apply data-driven predictive methods. The advantage of using these methods, when compared to the first principle models, is their easier deployment and usually higher transparency. In contrast to the first principle models, extensive knowledge of the actual process, though it may be helpful, is not required when developing the models. The data driven techniques extract automatically the process knowledge from the training data.

Further development of soft sensors may bring numerous additional benefits for the process industry. The main goal of the soft sensors is to gain more information about the process. This information may be for example some description of the process state which may be extracted from observing a group of relevant measurements. This kind of state monitoring could provide additional hints for the process operator and may help him to predict and thus to prevent inefficient or dangerous process states. Another benefit is the delivery of additional information about process features, which cannot be measured using hardware instrumentation and have to be evaluated in time consuming and man-power intensive lab analysis. In this case the soft sensor may deliver continuous information stream about the observed values.

In a more advanced scenario Soft Sensors could be also involved in the automated process control loops which would help to increase the plant effectiveness and thus for example reduce the energy consumption of the plant.

The previous paragraphs state some of the general benefits of using soft sensors in the process industry. Unfortunately there is still a lot of effort and expert knowledge necessary for the development and maintenance of the soft sensors. The main focus of current research is to deal with these issues. Soft sensors should be able to adapt to the changing environment and thus achieve the ability to develop themselves together with the plant. This will on one hand reduce the development costs of the soft sensor and on the other hand reduce the demand for maintenance of the sensors. By moving into the direction of the reduction of the incorporated a-priori information about the actual process, it should also be possible to deploy the soft sensor on different processes with as little effort as possible.

In this work, the proposed architecture tries to tackle the before mentioned challenges by applying adaptive mechanisms at different levels. Such an adaptive behaviour can be often found in the nature, for example in living cells, which are able to adapt to various environmental states and to survive in non-optimal conditions. The architecture allows the adaptation at three different levels. At the lowest level the involved computational learning methods can be adapted independent of each other. At the next higher level, there is a possibility to adapt the combinations of a number of different methods. There are several possibilities for building these combinations, they can be built either across different learning methods, different instances of the same method, or across different partitions of the input space. The adaptation mechanism can involve the evolving of the combinations parameters, e.g. the weights, or of the number of involved methods. At the highest level parameters influencing the behaviour can be adapted, these parameters can for example be the number of involved computational learning methods or the number of built combinations. To link this adaptive behaviour to a nature-inspired concept one can compare it to the hierarchical structure presented in [?]. This work discusses the various evolving information mechanisms in the brain at different levels, beginning at the quantum level, over the molecular, cell, neural ensemble up to the evolutionary (generation / population) processes.

This paper is organized as follows, the next section gives an insight into the critical properties of process industry data. Section 4 gives an overview of nature-inspired computational learning techniques for Soft Sensor modelling together with some case studies. The next section deals briefly with the current practises for the adaptation of Soft Sensors. Section 6 describes the actual architecture of the multi-level adaptive Soft Sensors and the last Section 7 summarizes this work.

CHARACTERISTICS OF THE PROCESS INDUSTRY DATA

Processing plants are by no means static environments. Quite to the contrary, they undergo steady process of changes, development and updates and thus represent a very dynamic environment. Because of the dynamic environment, the methods applied in the soft sensor modelling should possess high degree of adaptivity and the ability to react to such changing environment. The following paragraphs present the most critical data characteristics from the soft sensor development and maintenance point of view.

MISSING VALUES

This is probably the most common problem in terms of data availability within the process industry. It can have various causes. The most common causes are the failure of a hard sensor or its replacement. Because there is usually a very high number of sensors in a processing plant, it is common that the measured data consist of more than one hundred different variables. In such a scenario, there is a certain probability that some of the sensors will occasionally fail. One should also keep in mind that some of the sensor types are mechanical devices (e.g. flow rate sensors) and thus suffer from abrasion effects. Another source of missing values is the cleaning or hardware update of a part of the plant. Due to this it is usually necessary to bypass a part of the plant which also deactivates the sensors located in this part of the plant. There are different strategies to replace missing values, the most primitive univariate approach is to replace the missing values with the interpolated values of the affected features, which surround the missing value. More sophisticated approaches are based on multivariate statistics of the data, PCA based approaches. An optimal approach to missing value replacement should be able to deal with changing dimensionality of the input data, unfortunately these approaches are rather rare, two examples of such algorithms are [5], [16].

DATA OUTLIERS

Outliers are sensor values which deviate from the typical or sometimes also meaningful, ranges of the measured values. One can distinguish between two types of outliers, namely **obvious outliers** and **non-obvious outliers** [33]. Obvious outliers are those values which violate the physical and technological limitations. For example the pressure may not each negative values, flow sensor may not deliver values which exceed the technical limitations of the sensor. Thus it may seem that these outliers can be easily detected, though the problem of their detection is that one would need to input the limiting values in the form of a-priori information into the system. However for reasons stated later this is contra-productive to input such information a-priori because it would limit the portability of the soft sensor to a new process with possibly different a-priori information. Non-obvious outliers are even harder to identify because they don't violate any limitations but still lay out of the typical ranges and don't reflect the correct values. Typical approaches to outlier detection are based on the statistics of the past data, they can either be based on the observation of the distribution of single variables (i.e. features) as it is the case of the $x-\sigma$ outlier detection algorithm [13], where x is usually a value higher or equal to 3, or be based on the statistics of multiple features (e.g. Jolliffe parameter [20]). The detection of outliers as part of the pre-processing of the data remains one of the most critical parts of the soft sensor development. Currently there is a lot of process expert knowledge necessary to be able to judge whether a suspicious measurement is an outlier or not.

DRIFTING DATA

One of the two main reasons for observed drift in the measured data are the long-term changes of the processes. The processing plants consist of a large number of mechanical elements. This equipment undergoes steady abrasion during the operation of the plant. This may have an effect on the process itself, e.g. the flow between two parts of the process can decrease due to the abrasion of mechanical pumps. The cause of the drifting values can be also external influences like the environmental temperature, the purity of the input materials, etc. All of the above have not only an influence on the data but affect the process state as well and thus should be recognised and appropriate actions taken to compensate for them. The second main reason for the observed drift in the data is the drifts caused by the measuring devices. This type of drifts while observed in the measured data does not reflect the process changes. Therefore in the case of these drifts, the action to be taken should be the re-calibration of the measurement devices or the adaptation of the soft sensor without performing any corrective actions to the process itself. Distinguishing between the two different drift causes can not be easily done and once again a lot of expert knowledge is needed in order to take appropriate action. Another challenging aspect of dealing with drifting data is the fact that the changes may progress very slowly. Usually the drifting effects are much slower than the changes within the process itself. Additionally the drifting data may influence each other and have non-linear form which makes them even harder to detect.

COLLINEARITY

Another challenging issue apart from those stated above is related to the structure of the data. Process industry data can be described as **Data rich but information poor**, which means that there is usually a large amount of data with low relevant information content available. Typically the measured data is strongly collinear. This usually results from partial redundancy in the sensor arrangement and from process related effects (e.g. two neighbouring temperature sensors will deliver strongly correlated measurements). There are several statistical approaches for effective solving of the collinearity problem (e.g. PCA [20], PLS [1]) but these algorithms are usually not adaptive and making them adaptive is not trivial although not impossible.

SAMPLING RATES

Various sensors usually work at different sampling rates and thus one has to take measures to synchronise them. The synchronization of the data is usually handled by the Process Information Management System (PIMS) which records new data samples only if one of the observed variables changes more than a pre-defined threshold value. The definition of the threshold is another critical point, because too low values would cause the recording of unnecessarily too large number of samples and too high threshold could lead to the missing of important process changes. Additional issue, which is related to the data sampling rate, are the process related delays of the measurements. The delays can, in the worst case, reach values of up to several hours (e.g. values established manually in laboratories). The action which can be applied to deal with the time delays is for example to include delayed version of the input variables (measurements) into the feature selection.

NATURE-INSPIRED TECHNIQUES FOR SOFT SENSING AND THEIR APPLICATIONS

ARTIFICIAL NEURAL NETWORKS

The typical modelling approach used for these problems is the application of artificial neural networks (ANN). ANNs have found broad application in the computational learning, and thus also in the soft sensor modelling, after the invention of the back-propagation algorithm. They attracted the attention of scientists due to their generalisation power and abilities to solve non-linear problems. This fact probably led to the wide application areas of ANN, but unfortunately there is also some issues related to the application of ANN. Typically, the learnt knowledge is stored globally in the connection weights of the ANN which make it difficult to extract the knowledge in human readable form from the model. Besides this fact there is also a problem with the determination of the topology of the network. Usually it is selected using cross-validators or heuristic approaches.

An application of ANN for the sugar quality estimation was published in [11]. The approached problem in this publication is the modelling of the massecuite electrical conductivity because this is an important value for the control loop controlling the sugar production process. The input features of the model were selected manually and limited to eight from the modeller perspective important process measurements. The results achieved by the ANN were exact enough for taking the soft sensor into operation. Fermentation processes are another target group for soft sensors. The models of fermentation processes are very difficult to build because these processes are not easy to control and may vary significantly from one batch to another even under constant condition. In [10] artificial neural networks are compared to First Principle Models (FPM) and extended Kalman Filter (eKF) [38], which are another popular approaches to soft sensor building. The disadvantages of FPM and eKF are the complexity of the development and amount of a-priori knowledge which has to be available for the model development. Thorough analysis of the application of ANN for soft sensor building has been presented in [33]. This work discusses a lot of practical issues of the application of neural networks for soft sensor modelling. A particular focus is set on the necessary pre-processing steps like, the handling of missing values, outliers, etc. There is also a modification of the error measure of the back-propagation algorithm (using of Manhattan distance instead of mean squared error) proposed. Self-organizing maps (SOMs) [25] is an example of ANN type which is able to deal with unsupervised problems. Provided a set of high-dimensional input samples it maps this features to a lower dimensional, usually 2-dimensional, space. The mapping is done with the constraint of keeping the topological properties of the data. The properties of the SOM make it useful on for both visualisation and clustering of multivariate data. In terms of soft sensor modelling SOMs may be applied for process monitoring purposes. Process monitoring means the extraction of meaningful process states from the input data. A set of practical applications of process monitoring and quality prediction, etc, using SOMs was published in [3]. In this work SOMs have been found as an useful tool for the monitoring of a continuous pulp digester. Before feeding the data into the SOM model they have been manually pre-processed using a-priori knowledge of the process. Another application presented in the work is the quality prediction of steel prediction based on the concentration of the input elements and some process parameters. The last but not least application of SOMs presented in the work is the analysis of the data from paper and pulp industry.

NATURE-INSPIRED HYBRID METHODS

In the last few years a very popular research topic in computational learning field is the hybridisation of different modelling approaches. An overview together with the definition of different levels was presented in [2]. The advantage of hybrid systems is that successful combinations of methods are able to deal with the issues of the particular methods and make use of their advantages at the same time.

The probably most commonly used hybrid approach is the Neuro-Fuzzy System (NFS). It combines the Fuzzy Information System (FIS) with Artificial Neural Networks. Neuro-fuzzy models combine the learning capabilities of ANN with the human-like reasoning of FIS. In this way the disadvantages of ANN, namely the problems with the interpretability of the learnt knowledge, and at the same time the issues of FIS, namely the missing of a straight-forward learning algorithm, is compensated. NFS inherits the connectionist structure from ANN, the model consists of several layers of neurons. The first layer usually represents the linguistic labels of the input variables (e.g. small, medium, large), in the second layer the antecedent parts of the rules are built (e.g. if x_1 is small and x_2 is medium . . .), the next layer represents the consequent parts of the rules (with an optional normalisation layer between the 2 and 3 layer), the output layer is summation of all incoming signals (rules).

An example of Soft Sensors based on neuro-fuzzy approaches is an ANFIS (see [19] for ANFIS description) soft sensor being applied to predict the viscosity of rubber, which cannot for technological reasons be measured online [32]. Another ANFIS soft sensor was presented in [37]. In this work the data is pre-processed using PCA transformation. The paper defines a methodology for the development of Soft Sensors using soft computing methods. The methodology is applied to prediction of the anchorage of a polymeric-coated substrate. Soft sensor based on rough set theory and optimized by a genetic algorithm is shown in [30]. The application shown in this work is the prediction of freezing point of the light diesel fuel in a fluid catalytic cracking unit.

In the framework of neuro-fuzzy techniques, there has been few publications dealing with the adaptivity and evolution of the models. The neuro-fuzzy models applied for the modelling are the extended evolving fuzzy Takagi-Sugeno models exTS [6]. The advantage of these models is that they can change together with the changing environment and thus adapt to the new conditions. In [31] an exTS model has been applied to the prediction of the quality of crude oil distillation. The advantages of the proposed approach are numerous, one of them is the already discussed ability of the model to evolve together with the changing data by adapting and deploying new hidden units in the neuro-fuzzy network, another advantage common to all fuzzy based models is the interpretability of the learnt knowledge, which is represented in the form of human-readable fuzzy rules. Another related application in this field was published in [?]. In this work an adaptive neuro-fuzzy approach was applied to the modelling of water distribution system. New states of the system are added to the knowledge base by deploying a new receptive field, which corresponds adding a new neural unit to the neuro-fuzzy model. As next, a new model is trained using the data accommodated within the new receptive field and added to the knowledge base of the system.

Apart from the combination of ANN and FIS, there is a large number of other combination possibilities. Neural-Network Partial Least Squares (NN-PLS) is another hybrid algorithm combining neural networks with PLS [33]. The algorithm makes use of the capabilities of ANN to approximate any non-linear relationship between the input values and the output. This capability is exploited to map the input variables onto the latent variables of the PLS. The hybrid algorithm is applied to a real-life problem, namely to a refinery process. Another application of NN-PLS to soft sensing was presented in [12] where the NN-PLS (and NN-PCA) algorithm is applied to the prediction of emissions of NO_x gas in exhaust streams. A hybrid system consisting of Particle Swarm Optimisation (see [9]) which is used for the training of an ANN was presented in [29]. In this work the PSO algorithm is combined with the Alopex algorithm [35] to avoid local minima to which the PSO is prone.

Another hybrid approach using nature-inspired methods for soft sensor modelling has been developed by Kordon et al. [26]. Here, the hybridisation is done on a lower level. The involved methods perform pre-processing of the data for the succeeding modelling steps. The methodology for the inferential sensor building consists of three different steps. In the first step the data is analysed by an analytical neural network [27]. During this step a variable selection is done. Another task carried out by the analytical neural networks is the dealing with time delays between the selected variables. This is achieved by convoluting the variables by appropriately shaped function. In the next step the data is processed using SVM. During this step outlier detection is done. In third final step the inferential (soft) sensor is built. This is performed by applying a Genetic Programming (GP) algorithm (see [28]). The GP algorithm selects a function from a pool. This function is trained to model the output variable using the pre-processed input data from the previous two steps. The output is then a set of analytical functions which map the input space to the target variable space. The proposed approach was applied to several real-life problems, e.g. the interface level estimation in an organic process [21].

SUMMARY OF NATURE-INSPIRED TECHNIQUES FOR SOFT SENSOR MODELLING

Although there is a large number of publications dealing with real-life process data, there is still lots of challenges in the soft sensor development and maintenance. The common factor of all the practical implementations of Soft Sensors is that there is a need for intensive manual pre-processing of the data being used for the training and later evaluation of the soft sensor. The published Soft Sensors are not adaptive enough to deal with the dynamics of the data. Some authors (e.g. [13]) explicitly stress the need for discussions with the process experts as a part of the pre-processing of the data. These implementations of the expert knowledge into the soft sensor are very time consuming and thus very expensive and restrain a larger spread of Soft Sensors in the process industry. Additional aspect is that the manual pre-processing measures, which should make the data more suitable for the modelling, are not everlasting. The reasons for this fact were discussed in Section 3. Usually there is a manual periodical maintenance of the soft sensor necessary, which again restricts the applicability of soft sensor on real process plants.

CURRENT ADAPTIVITY PRACTICES

In terms of adaptive properties one can distinguish two approaches. The first one is the update of the parameters (e.g. of the weights of an ANN or of the basis of the PCA). This parameter tuning assumes fixed structure of the models and is thus also sometimes called **low level adaptation** [6]. In contrast to that **higher level adaptation** assumes also adaptation of the structure (i.e. complexity) of the model. A term commonly used in this context is **evolving systems**. Those systems gradually develop their structure and parameter sets with the changes of the data. At the moment, the most promising techniques in this context are neuro-fuzzy approaches. These techniques deploy different approaches of the adaptation to the changing environment. They can for example adapt the membership function parameters as well as their form and number, further on one can allocate new clusters (or receptive fields in terms of neuro-fuzzy models) if the incoming data falls into new unoccupied region of the input space. For examples of evolving neuro-fuzzy approaches see: [4], [5], [6], [7], [15], [16], [22], [23], [24].

An open question which should be answered is when and how to update the structure and the parameters of the models. As in a realistic scenario the measurements will be incoming in a continuous stream, one after another, there should be a mechanism which will decide if there is enough changes to trigger the update mechanism of the system. Usually those mechanisms are based on pre-defined threshold values. But for a truly adaptive and evolving system it is difficult to define those thresholds. The straightforward and most commonly applied way of adapting the models is to use a temporally sliding window. Each time the window shifts the model is either recursively updated or completely recalculated dependent on the type of the model. The disadvantage of this approach is that it completely forgets, in case of the complete recalculation, or gradually decreases the weight of the past data points. More effective approach would be to change the weights of the past data points not based on their age but on their significance in terms of the task. The benefit of this approach is that the important data, which may represent for example critical states of the processing plant, are kept in the memory independent of their time of occurrence. From the previous discussion it should be clear that the adaptivity has a lot of dimensions. A fully adaptive system should be able to cope with slow changes and drifts in the measurements as well as with sudden changes like sensor failures. It should also be able to evolve together with the plant.

ARCHITECTURE FOR SOFT SENSOR MODELLING

To cope with the challenges listed previously in this work, especially those discussed in Section 3 a novel architecture for a self-adapting Soft Sensor is proposed. A broad overview of the architecture is shown in Figure 1. A significant part of the proposed architecture are the two pools, the **Pre-processing Methods Pool**, including all pre-processing methods, which is further split into actual pre-processing methods (e.g. smoothing, normalisation), feature selection methods (e.g. correlation-based feature selection) and instance selection methods (i.e. receptive fields filtering). The second pool, **Computational Learning Methods Pool**, consists of various computational learning methods (e.g. linear regression, multi-layer perceptrons, etc.). The particular instances of the methods in the pools are connected to paths by the **Path/Pool Management** module. An example of a path built from the following elements: data normalisation, correlation-based feature selection and linear regression is shown in Figure 2.

At the path level, there is an additional mechanism for the local control necessary. Such a mechanism is shown in Figure 2. The presented path consists of three path elements. The first two elements are pre-processing steps, namely the data normalisation and feature selection, and the third one is a linear regression computational method. The prediction of the method is fed to the **Local Evaluation**, where together with the input data the evaluation of the path prediction is performed. The results of the prediction are passed to the **Local Control Unit** which controls the parameters (e.g. the weights of the linear regression model) of the path elements. Additional input to the **Local Control Unit** is the control information from the high-level decision making methods (i.e. Path/Pool management, Meta-Level Learning, etc.).

Another important part of the architecture is the **Path Combination** module. Typically the combination is carried out at the level of the computational learning methods (e.g. stacking of a group of MLP networks [13]). In the proposed architecture the combinations are performed at the path level, which provides several advantages. One of them is that one can do the combination while including different methods from pre-processing methods (e.g. a combination of several paths consisting of MLP with different feature selection methods). Another advantage is that it is possible to combine different computational learning. The path combination module together with the instance selection methods provides also the possibility to combine different local paths in the sense of local learning methods to a global model.

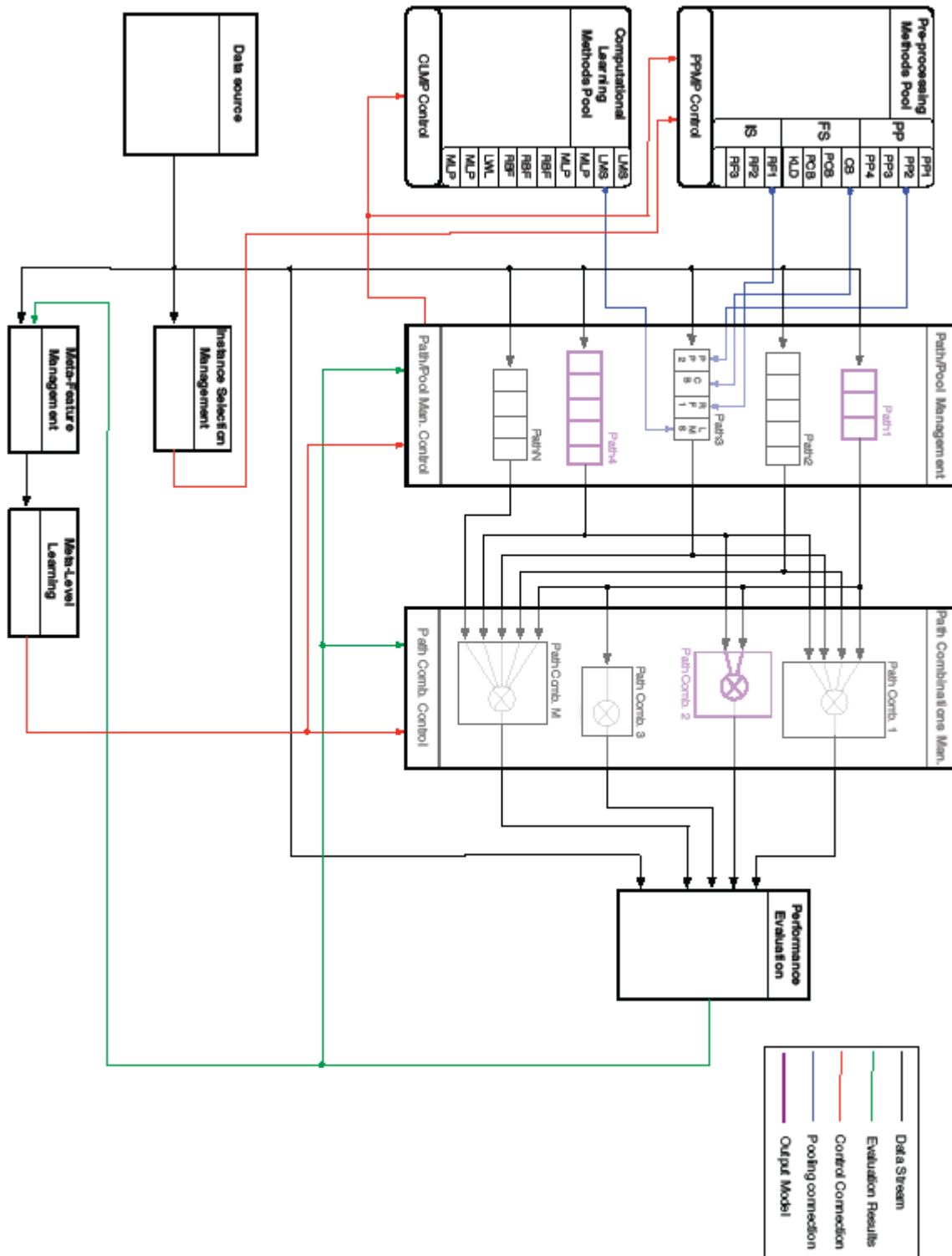


Figure 1: Proposed architecture for an adaptive Soft Sensor

The architecture provides also the possibility of using meta-learning approaches [36], [17]. For this purpose, there are two modules in the architecture. The first one, Meta-Feature Management, having information about the data together with the performance of the particular paths builds the meta-features. This module may e.g. extract the information about the performance of the different paths in the different parts of the input data space and pass this information further to the Meta-Level Learning module which can, using the provided information, control the Path/Pool Management and Path Combination modules. The Instance Selection Management module is responsible

for the filtering of the instance and thus providing the possibility for building of local models, i.e. local experts, [8], [34], [14], [18]. The local approach to the model building is, apart from the pool and path approach and meta-learning techniques, one of the key aspects of the proposed architecture.

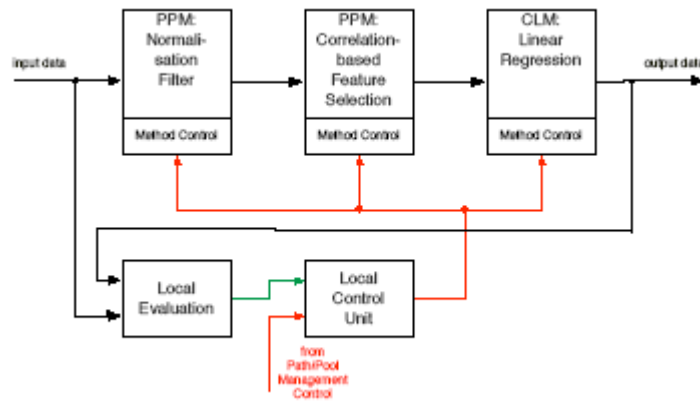


Figure 2: Transformation path example

CONCLUSION

The contribution of this work is on one hand an overview of the state of the art of Soft Sensor building, while on the other hand a novel adaptive architecture for the building of Soft Sensors is proposed. Together with the most common nature-inspired methods, like artificial neural networks or evolving neuro-fuzzy methods, the problems of the current approach to the building of Soft Sensors are also reviewed. The current problems in the development and maintenance of Soft Sensors originate in the nature of the data available for the building of data-driven models. Based on the identified problems, architecture for a novel adaptive Soft Sensor is proposed. There are several key elements of the architecture, one being its adaptation capabilities at different levels, while another is the management of a set of computational methods with different adaptation capabilities, adaptation speeds, etc. It is also illustrated how the state-of-the-art methods from the fields of local, meta-learning, model selection and combination can be accommodated within the proposed architecture.

REFERENCES

- [1] H. Abdi. Partial least squares (pls) regression. Encyclopedia of Social Sciences, Research Methods. Thousand Oaks (CA): Sage (2003), 2003.
- [2] A. Abraham. Intelligent systems: Architectures and perspectives. Arxiv preprint cs.AI/0405009, 2004.
- [3] E. S. A. Alhoniemi. Process monitoring and modeling using the self-organizing map. *Integrated Computer-Aided Engineering*, 6(1):3–14, 1999.
- [4] P. Angelov and R. Buswell. Identification of evolving fuzzy rule-based models. *IEEE TRANSACTIONS ON FUZZY SYSTEMS*, 10(5):667, 2002.
- [5] P. Angelov and N. Kasabov. Evolving computational intelligence systems. In *IEEE Workshop on Genetic and Fuzzy Systems GFS2005*, Grenada, Spain, 2005.
- [6] Angelov and Z. Xiaowei. Evolving fuzzy systems from data streams in real-time. pages 29–35, 2006.
- [7] P. P. Angelov and D. P. Filev. Flexible models with evolving structure. *International Journal of Intelligent Systems*, 19(4):327–340, 2004.
- [8] C. G. Atkeson, A.W. Moore, and S. Schaal. Locally weighted learning. *Artificial Intelligence Review*, 11(1):11–73, 1997.
- [9] E. Bonabeau, M. Dorigo, and G. Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, 1999.
- [10] A. J. de Assis. Soft sensors development for on-line bioreactor state estimation. *Computers and Chemical Engineering*, 24(2):1099–1103, 2000.
- [11] D. Devogelaere, M. Rijckaert, O. G. Leon, and G. C. Lemus. Application of feedforward neural networks for soft sensors in the sugar industry. In *Neural Networks, 2002. SBRN 2002. Proceedings. VII Brazilian Symposium on*, pages 2–6, 2002.
- [12] D. Dong, T. J. McAvoy, and L. J. Chang. Emission monitoring using multivariate soft sensors. In *American Control Conference, 1995. Proceedings of the*, volume 1, 1995.
- [13] L. Fortuna. *Soft Sensors for Monitoring and Control of Industrial Processes*. Springer, 2007.
- [14] E. Frank, M. Hall, and B. Pfahringer. Locally weighted naive bayes. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, pages 249–256, 2003.
- [15] B. Gabrys. Agglomerative learning algorithms for general fuzzy min-max neural network. *Journal of VLSI Signal Processing*, 32:67–82, 2002.
- [16] B. Gabrys. Neuro-fuzzy approach to processing inputs with missing values in pattern recognition problems. *International Journal of Approximate Reasoning*, 30(3):149–179, 2002.
- [17] C. Giraud-Carrier, R. Vilalta, and P. Brazdil. Introduction to the special issue on meta-learning. *Machine Learning*, 54(3):187–193, 2004.
- [18] R. Jacobs. Adaptive mixtures of local experts. *Neural Computation*, 3(1):79–87, 1991.
- [19] J. S. R. Jang, C. T. Sun, and E. Mizutani. *Neuro-fuzzy and soft computing*. Prentice Hall Upper Saddle River, NJ, 1997.

- [20] I. T. Jolliffe. *Principal Component Analysis*. Springer, 2002.
- [21] A. Kalos, A. Kordon, G. Smits, and S. Werkmeister. Hybrid model development methodology for industrial soft sensors. In *2003 American Control Conference*, pages 5417–5422, 2003.
- [22] N. Kasabov. *Evolving fuzzy neural networks—algorithms, applications and biological motivation. Methodologies for the Conception, Design and Application of Soft Computing*, World Scientific, pages 271–274, 1998.
- [23] N. Kasabov and D. Filev. *Evolving intelligent systems: Methods, learning, applications*. pages 8–18, 2006.
- [24] N. K. Kasabov and Q. Song. Denfis: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *Fuzzy Systems, IEEE Transactions on*, 10(2):144–154, 2002.
- [25] T. Kohonen. *Self-organizing maps*. Springer-Verlag New York, Inc. Secaucus, NJ, USA, 1997.
- [26] A. Kordon, G. Smits, E. Jordaan, E. Rightor, D. C. Co, and T. X. Freeport. Robust soft sensors based on integration of genetic programming, analytical neural networks, and support vector machines. In *Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress on*, volume 1, 2002.
- [27] A. K. Kordon. Hybrid intelligent systems for industrial data analysis. *International Journal of Intelligent Systems*, 19(4):367–383, 2004.
- [28] J. R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Bradford Book, 1992.
- [29] S. J. Li, X. J. Zhang, and F. Qian. Soft sensing modeling via artificial neural network based on pso-alopex. *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, 7, 2005.
- [30] J. X. Luo and H. H. Shao. Developing soft sensors using hybrid soft computing methodology: a neurofuzzy system based on rough set theory and genetic algorithms. *Soft Computing—A Fusion of Foundations, Methodologies and Applications*, 10(1):54–60, 2006.
- [31] J. J. Macias, P. Angelov, and Z. Xiaowei. A method for predicting quality of the crude oil distillation. Pages 214–220, 2006.
- [32] S. Merikoski, M. Laurikkala, and H. Koivisto. An adaptive neuro-fuzzy inference system as a soft sensor for viscosity in rubber mixing process. Technical report, WSEAS NNA-FSFS-EC 2001, Puerto de la Cruz, Tenerife, Spain, 11-15 February 2001, paper 446; <http://www.ad.tut.fi>, 2001.
- [33] S. J. Qin. Neural networks for intelligent sensors and control—practical issues and some solutions. *Neural Systems for Control*, page 213234, 1997.
- [34] S. Schaal and C. G. Atkeson. Constructive incremental learning from only local information. *Neural Computation*, 10(8):2047–2084, 1998.
- [35] E. Tzanakou, R. Michalak, and E. Harth. The aloplex process: Visual receptive fields by response feedback. *Biological Cybernetics*, 35(3):161–174, 1979.
- [36] R. Vilalta and Y. Drissi. A perspective view and survey of meta-learning. *Artificial Intelligence Review*, 18(2):77–95, 2002.
- [37] K. Warne, G. Prasad, S. Rezvani, and L. Maguire. Statistical and computational intelligence techniques for inferential model development: a comparative evaluation and a novel proposition for fusion. *Engineering Applications of Artificial Intelligence*, 17(8):871–885, 2004.
- [38] G. Welch and G. Bishop. *An introduction to the kalman filter*, 2001.