Advances in Smart Systems Research : ISSN 2050-8662 : http://nimbusvault.net/publications/koala/assr/
Vol. 3. No. 2 : pp.31-37 : iswk13-007

Context-based Identification of Energy Consumption in Industrial Plants

João Cruz [a], Rui Neves-Silva [a], Maria Marques [b]
[a] *DEE-FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal*
jm.cruz@campus.fct.unl.pt
[b] *UNINOVA, FCT Campus, 2829-516 Caparica, Portugal*

Abstract. In industrial environment, plants are exposed to multiple contextual factors, which affect the way that energy is used. Actual energy models provide guidelines for energy consumption, but fail in explaining some patterns observed in practice. This paper provides a novel approach for building energy models of industrial plants, based on the influence of context variables. Multiple RLS algorithms are used for multi-model context-based estimation of energy consumption. For context identification, regression trees are used to divide the context variables into regions, where the local models are estimated. The effectiveness of the proposed approach is illustrated with some examples and validated with a practical case in a real industrial cement plant.

Keywords. Energy Consumption; Context Variables; RLS Multi-Model Estimation; Regression Trees;

Introduction

In an industrial environment, energy consumption depends on multiple factors, such as production factors (e.g. type of product being manufactured; equipment condition), ambient factors (e.g. humidity of the air; year season) or human factors (e.g. operators experience). Actual energy consumption models are based on fixed operating regimes, disregarding context awareness. These models can provide guidelines for energy use, but tend to fail in explaining some behaviors. Also, recently technological advances in electronics and communication networks, allows the continuous and distributed monitoring of contextual information in industrial environments, which makes it easier to acquire information. The main contribution of this work is a novel approach for building energy consumption models based on context variables.
While an industrial plant is performing a specific task, it is exposed simultaneously to multiple contextual variables. As the task is being performed, variables related to the plant operation, also affect the energy consumption. From this point of view, these two types of variables are at the same level. Yet, context variables are classified as external variables that are kept steady for periods of time that go beyond of plant time constants, whereas plant operation variables are defined as being the input, directly related to the task that is being performed, e.g. while a milling is grinding, the flux of raw material is considered as a plant input variable, while the type of cement being produced is a context variable. The steadiness period, suggests the partition of the context variables into several regions in which a local model can be estimated. Therefore, the

construction of context-based energy models involves two entirely different stages. In a first stage, the context variables are partitioned into several regions, using regression trees, as they deal well with continuously measured data, and provide fast algorithms for computation [1]. And the second stage is the energy consumption modeling. Some modeling techniques based on regression analysis[2], decision trees[3] and neural networks[4] are suggested by different authors. Although, the modeling is made by one unique nonlinear descriptor. As different regions of the contexts variables affect differently the energy consumption, a multi-model approach is proposed, using multiple RLS algorithms for model estimation. Therefore, for each region a local energy consumption model is estimated.

1. Context-based Identification Concept

As mentioned earlier, the context-based identification of energy consumption involves two different stages. The global architecture is depicted in Fig. 1. The block named Region and Centroid Identification is responsible for the partitioning of context variables into a set of regions; and calculation of respective centroids. For model estimation, the same data regressor is used. The global nonlinear energy consumption model is obtained by joining all the local models estimated for each region of the context variables.

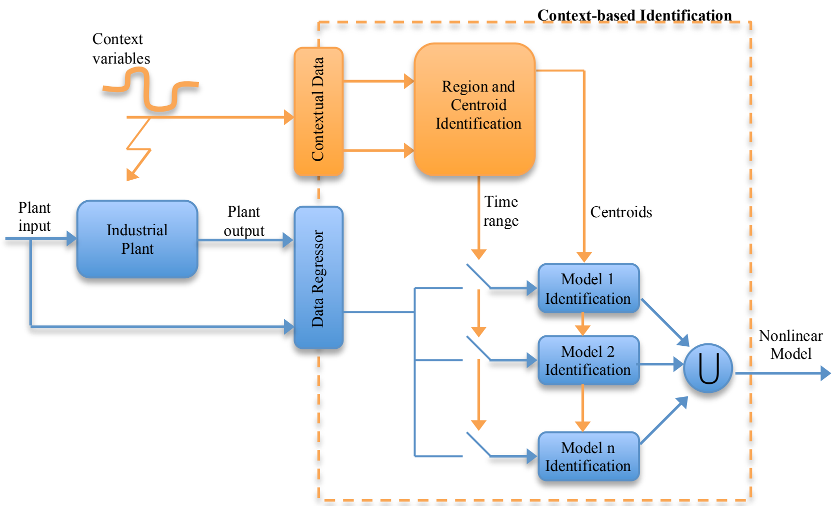

Figure 1 – Context based identification architecture.

2. Context Identification

This chapter describes how context variables are divided into several regions, in which local models are used to estimate the energy consumption. The proposed approach starts by splitting the context variables into several regions, where a centroid is calculated. For this purpose, regression trees are used for each context variable. Fig. 2 depicts the global architecture.

Each of the tree leaves defines a region, that is a time range for model estimation, and has attached to it a real value which applies only in that region, named centroid.

Existing methods for tree construction utilize different splitting and pruning approaches, such as distance measures, likelihood-ratio tests or least squares criterion [5].

The regression trees applied in this paper, use the least squares criterion, and are constructed by recursively splitting leaf nodes in order to maximize the tree homogeneity. Also, a steadiness time must be granted to able the models to be identified; this is achieved by stopping the tree from growing when the region is less than a specified dwell time.

The centroid identifies where an estimated model is valid. I.e. suppose that the samples between k_i and k_f are all samples belonging to the same region R of a context variable w. Then the centroid c_w^R is the mean of w over R, as described by the following equation

$$c_w^R = \frac{1}{k_f - k_i} \sum_{i=k_i}^{k_f} w(i) \tag{1}$$

Therefore, a model is only valid in the proximity of the centroid, delimited by the region boundaries.

After all context variables are partitioned, it is necessary to intersect the regions and join the respective centroids. The Time Ranges and Centroids Matcher block performs this function.

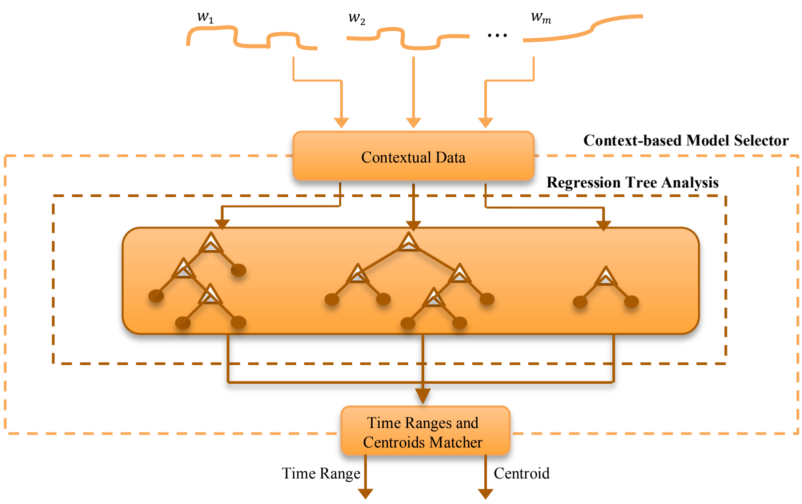

Figure 2. General architecture for context-based model selection.

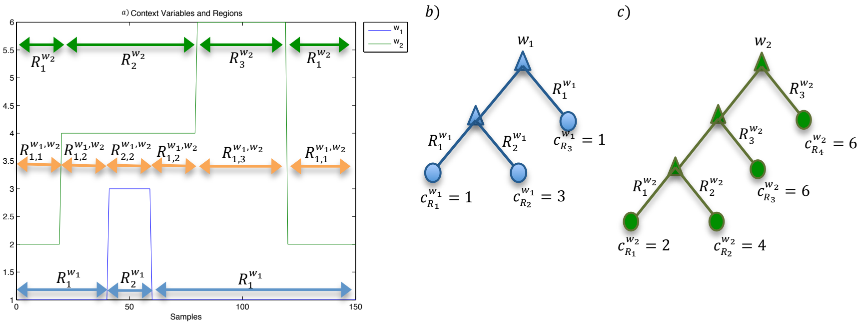

Figure 3. a) Context variables and regions; The green, blue and orange arrows mark the regions of w_1, w_2 and its intersection, respectively; b) Regression tree of w_1; c) Regression tree of w_2.

For example, consider two distinct context variables w_1 and w_2, that represent the worker's ID number operating simultaneously the same machine, as depicted in fig. X a). The respective regression trees are represented in fig. X b) and c); and the numerical values of centroids and time-ranges for each region are presented in Table 1 and 2.

Table 1. Region, Centroid and Time Range results of regression tree analysis for context variable w_1.

Region	Centroid	Time Range
$R_1^{w_1}$	$c_{R_1}^{w_1} = 1$	$k < 40.5 \cup k \geq 59.5$
$R_2^{w_1}$	$c_{R_2}^{w_1} = 3$	$k \geq 40.5 \cap k < 59.5$

Table 2. Region, Centroid and Time Range results of regression tree analysis for context variable w_2.

Region	Centroid	Time Range
$R_1^{w_2}$	$c_{R_1}^{w_2} = 2$	$k < 19.5$
$R_2^{w_2}$	$c_{R_2}^{w_2} = 4$	$k \geq 19.5 \cap k < 79.5$
$R_3^{w_2}$	$c_{R_2}^{w_2} = 6$	$k \geq 79.5$

As the time-ranges of the regression trees overlap, it's necessary to intersect the regions and join the centroids, so that the active model in that region can be uniquely identified. Therefore, a new set of regions is created and presented in table 3.

By this approach, the estimated energy model for each region captures the influence of multiple contexts, which is coherent to what happens in an industrial environment where the plant is exposed at the same time to all the contextual variables, and all influence simultaneously the plant's energy consumption.

Table 3. Results of intersected regions and joined centroids.

Intersected Regions	Joined Centroids
$R_{1,1}^{w_1 w_2} = R_1^{w_1} \cap R_1^{w_2}$	$c_{R_{1,1}}^{w_1 w_2} = (1,2)$
$R_{1,2}^{w_1 w_2} = R_1^{w_1} \cap R_2^{w_2}$	$c_{R_{1,2}}^{w_1 w_2} = (1,4)$
$R_{2,2}^{w_1 w_2} = R_2^{w_1} \cap R_2^{w_2}$	$c_{R_{2,2}}^{w_1 w_2} = (3,6)$
$R_{1,3}^{w_1 w_2} = R_1^{w_1} \cap R_3^{w_2}$	$c_{R_{1,3}}^{w_1 w_2} = (1,6)$

3. Energy Models Estimation

After defining the centroids and time range for each region of the context variables, local modeling is used to estimate the plants energy consumption. There are as many models as regions; and while a model is being estimated, all the others are freezed. Therefore, while a plant is performing a specific task, e.g. milling, a RLS algorithm is used per region, in order to capture the energy consumption model. Also, the combination of multiple local models, defines the plant's global energy consumption.
In this outline, consider that the plant's energy consumption, can be modeled by the following equation

$$y(k) = \theta^T \varphi(k) + e(k) \tag{2}$$

Where $y(k)$ is the energy consumption; θ is a vector of unknown parameters; $\varphi(k)$ is a regression vector of continuously measurements of past energy consumption and plant input, according to the task that is being performed; $e(k)$ is white noise disturbance; and k is the time discrete independent variable. Fig. 4 a) depicts the measured energy consumption signal; and Fig. 4 b) the measured process input u, e.g. flux of raw material.
Also, consider that the context variables w_1 and w_2, exemplified in the previous chapter and again represented in Fig. 4 c), influence directly the parameters θ. Thus, as θ is context-based dependent, eq. 2 must be re-written as follows

$$y(k) = \theta^T(w) \cdot \varphi(k) + e(k) \tag{3}$$

Where w is a vector of context variables measurements.
The main goal of the RLS algorithms is to estimate the unknown context-based parameters as the data is continuously measured. Consider that for each time discrete sample, the regression vector $\varphi(k)$ is described by

$$\varphi(k) = [y(k-1)\,;\ y(k-2)\,;\ y(k-3)\,;\ u(k-1)\,;\ u(k-2)] \tag{4}$$

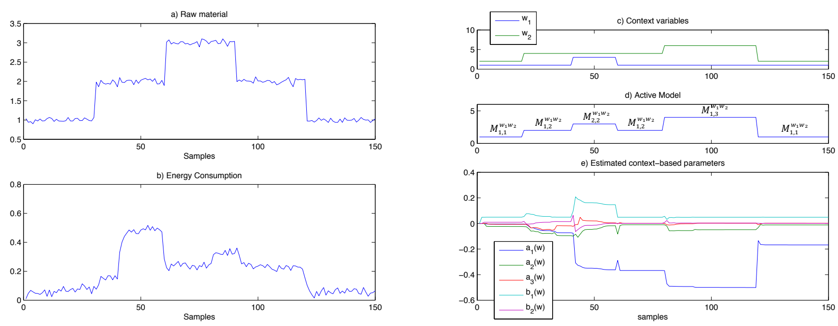

Figure 4. a) Raw material as input signal; b) Energy consumption as output signal; c) Context variables; d) Estimated context-based parameters; e) Active model

Therefore, the unknown context-based parameters being estimated by each RLS algorithm is defined by

$$\hat{\theta} = [a_1(w)\ ;\ a_2(w)\ ;\ a_3(w)\ ;\ b_1(w)\ ;\ b_2(w)]^T \tag{5}$$

Fig. 4 d) depicts the active model being estimated, and Fig. 4 e) the RLS parameter estimation for each model, whose numerical values are presented in table 4.
Notice that as the RLS algorithm takes more measurements into account, the parameters converge to a final value. This states that the model is as more accurate as more measurements are made. Also, as the region of the context variable changes, the parameters seems to be strongly affected, which suggest that both dynamical and static properties of the energy model are being influenced by the context variables. As the static property is defined by the steady state relation, for the given example the relation between flux of raw material and energy consumption, it can be stated that the context is affecting the way that the energy is being used, in terms of production rate.

Table 4. Model, centroid and respective context-based parameter estimation.

Model	Centroid	Estimated Parameters
$M_{1,1}^{w_1 w_2}$	$c_{R_{1,1}}^{w_1 w_2} = (1,2)$	[-0.1667 ; -0.0127 ; 0.0013 ; 0.0489 ; 0.0006]
$M_{1,2}^{w_1 w_2}$	$c_{R_{1,2}}^{w_1 w_2} = (1,4)$	[-0.3670 ; -0.0117 ; 0.0026 ; 0.0488 ; 0.0003]
$M_{2,2}^{w_1 w_2}$	$c_{R_{2,2}}^{w_1 w_2} = (3,6)$	[-0.3612 ; -0.0166 ; 0.0040 ; 0.1432 ; 0.0047]
$M_{1,3}^{w_1 w_2}$	$c_{R_{1,3}}^{w_1 w_2} = (1,6)$	[-0.4978 ; -0.0511 ; 0.0032 ; 0.0487 ; 0.0029]

For model evaluation, the estimated models are used for simulation of the energy consumption, with the same identical setup. Figure 5 depicts the comparison between the measured energy consumption y, and the simulation $\tilde{y}_m$ using the multi-model context-based approach. Fig. 5 also depicts a simulation $\tilde{y}_s$, based on the classical method of using a unique model also estimated by RLS algorithm.
The mean square error of the multi-model context-based approach is 5.2802×10^{-4} while the mean square error of the classic method is 1.1172×10^{-2}, which indicates a very promising and applicable concept for improving the actual used energy consumption models.

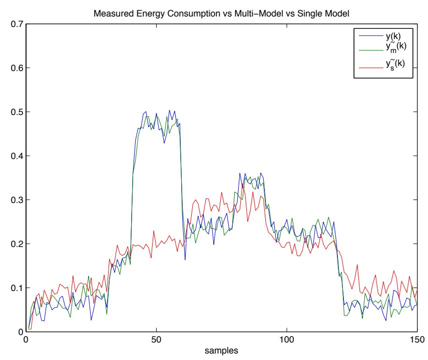

Figure 5. Comparison between measured energy consumption, multi-model and single model simulation.

4. Conclusion

Context-based energy consumption models can be used for better understanding of the influence of context variables in industrial plants, but also for extrapolating the energy consumption into the future, according to different scenarios. Intelligent decision support system, based on these models, will be able to improve production strategies, reduce energy costs, and solve other type of problems such as peak load management. The context-based identification concept has been validated by the results obtained in the cement plant, that shown very accurate model's prediction.