Modelling viscosity in rubber mixing process using an adaptive neuro-fuzzy inference system (ANFIS)

Seppo Merikoski
Tampere University of Technology, P.O. Box 692, FIN-33101 Tampere, Finland
Tel. +358 3 365 2111, Fax +358 3 365 2340, seppo.merikoski@tut.fi
Mikko Laurikkala
Tampere University of Technology, P.O. Box 692, FIN-33101 Tampere, Finland
Tel. +358 3 365 2111, Fax +358 3 365 2340, mikko.laurikkala@tut.fi
Hannu Koivisto
Tampere University of Technology, P.O. Box 692, FIN-33101 Tampere, Finland
Tel. +358 3 365 2111, Fax +358 3 365 2340, hannu.koivisto@tut.fi

KEYWORDS rubber mixing, viscosity, neuro-fuzzy systems, soft sensors, identification

ABSTRACT

Natural rubber is an important raw material for the tyre industry. Due to natural rubber's nature origins its properties are not constant. Production process should be self-adapting so that variations in raw materials would not reflect as variations in the end product quality. One of the important characteristics of natural rubber is viscosity.

The aim of this paper is to find a model for rubber mixing process. With such a model it would be possible to predict viscosity of the rubber mixture at the end of the mixing process.

Modelling is done using a neuro-fuzzy method which is capable of representing highly complex and nonlinear dependencies. The identification data used was gathered from the production database of a Finnish tyre manufacturer.

During the modelling, uncertainty of the data is concerned. Ordinary leasts squares cost function is replaced with the least median squares cost function which is robust against outliers in data.

The accuracy of the model found is satisfactory with respect to measurement accuracy and sufficient for quality contol purposes. Yet the reliability of the model is not sufficient for the basis of process control. Futher research is needed to improve reliability and accuracy of the model. Adding information of the process dynamics to the model seems promising.

1 INTRODUCTION

Rubber mixing is the first phase in tyre manufacturing. Viscosity is one of the key quantities of the rubber mixture. It affects strongly the further phases of manufacturing and also end product quality. Main reason for viscosity variations is the variation in natural rubber properties.

Online measurement of viscosity is not possible. Usually viscosity is observed from laboratory samples. However, sampling and laboratory analyses are too slow for control purposes. This paper describes a research aiming at a soft sensor for viscosity in the mixing process. Practically this means identifying a model for viscosity based on the available online-measurements.

Predicting viscosity from process measurements has been studied by Palmgren /1/ and Tiejun /2/, who concentrated on the older tangential mixer type. Ryzko /3,4/ has studied the intermeshing mixer type with neural networks and linear regression.

2 MIXING PROCESS

Rubber mixing is performed in batches. Natural rubber together with possible synthetic raw materials and additives is fed into the mixer, and after a few minutes' mixing the mixture is discharged to an extruder. The mixing often consists of several successive parts called mastication, master batch and final batch. Each part in turn can include several phases. Additives can be added between phases.

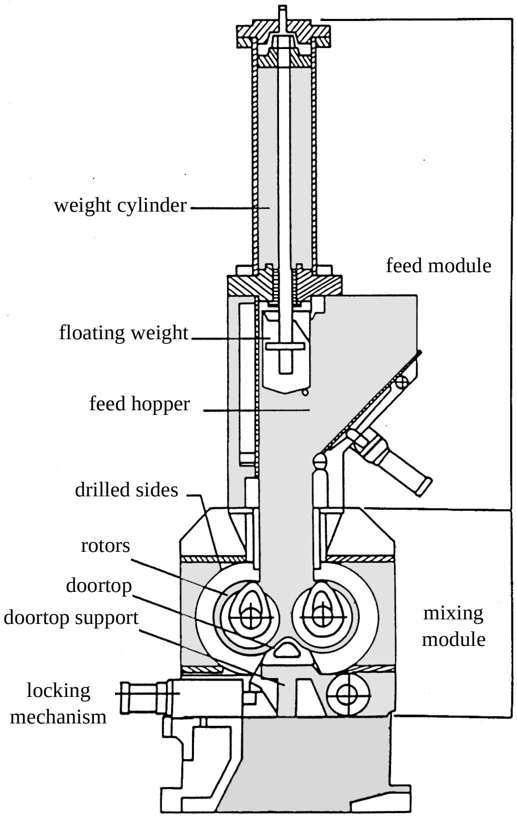

Figure 1. An internal rubber mixer. /5/

Roll-mill mixer and internal mixer (Figure 1) are the two main types for productional rubber mixing. Internal mixers in turn can be divided into tangential and intermeshing mixers. The research described in this paper was done with an intermeshing mixer.

Varying properties of natural rubber introduce a great amount of complexity to the mixing process. The process is controlled using e.g. temperature, rotor speed and batch duration. A traditional feedback control of viscosity is not possible since there is no online measurement. The viscosity information is only obtained from laboratory analysis after the batch has been discharged. A dominant system for controlling viscosity is therefore based on long-term tolerances and process operators' expertise.

3 MODELLING

3.1 Adaptive Neuro-Fuzzy Inference System

An Adaptive Neuro-Fuzzy Inference System (ANFIS) presented by Jang /6/ was chosen for the modelling method. ANFIS combines advantages of both fuzzy models and neural networks. It consists of a set of local models, which makes it suitable for modelling sparse data: local models are placed on areas with most data points and the intermediate areas can be interpolated.

3.1.1 Fuzzy Sugeno model

A fuzzy Sugeno model has rules of the following form:

$$\text{Rule i: IF x is } A_i \text{ AND y is } B_i \text{ THEN } z_i = f_i(x,y) \tag{1}$$

The consequence function f_i can be any function. Choosing simple functions like zeroth or first order polynomial makes model identification easier. The idea behind local models is to obtain a complex global model by interpolating between several simple local models.

Output of the model is calculated using weighted average:

$$f_{SUM} = \frac{w_1 f_1 + \ldots + w_i f_i}{w_1 + \ldots + w_i}, \tag{2}$$

where w_i is the firing strength and f_i the consequence function of rule i.

A fuzzy Sugeno model can be presented as a neural network structure (Figure 2). In ANFIS, this structure is exploited by applying neural network training methods to Sugeno models. Layer 1 in Figure 2 includes input membership functions of the fuzzy model. On the second layer, firing strengths of the rules are calculated. AND-connectives are implemented as products. Layer 3 calculates the product of the consequence functions and firing strengths. The sums presented in Equation (2) are calculated on the fourth layer, while the final layer normalizes the sum of the consequence functions with the sum of the firing strengths. /6/

Each rule corresponds to one area of operation and has a specific model. The areas overlap because membership functions are nonzero gaussian functions. The output is determined by the average of overlapping local fuzzy models weighted with firing strengths of the rules.

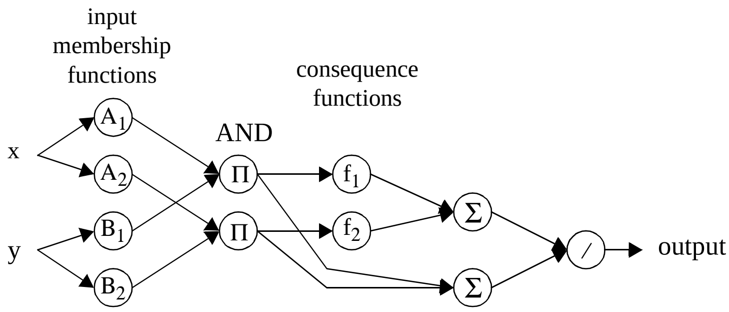

Figure 2. A first order Sugeno type fuzzy inference system. /6/

Input space can be partitioned using a grid (Figure 3a) where the whole input space is covered with fuzzy rules. A disadvantage of this approach is the strong increase in the number of rules as the number of inputs increases. Another solution is to use scatter partition (Figure 3b) which means dividing the input space into clusters. Clusters cover dense data areas, while areas of input space with little or no data remain uncovered. In this paper, scatter partitioning is used together with subtractive clustering /6,7/.

3.1.2 Parameter tuning

After initializing the Sugeno fuzzy model by clustering, the parameters of the model are fine tuned. In the neural network representation of the model (Figure 2), only layers 1 and 3 contain parameters. Parameters of layer 1 define the input membership functions and are thus called premise parameters. On layer 3, a set of consequence parameters defines the consequence functions (f_i in Equation (1)).

The overall output of the model is linear in consequence parameters but nonlinear in premise parameters. In the hybrid learning method detailed in /6/ there are two phases, forward pass and backward pass. Since the consequence parameters are linear, they can be updated using the least squares estimator (LSE) during the forward pass. In the backward pass, the premise parameters are updated using steepest descent (SD) and backpropagation.

3.2 Preprocessing of data

3.2.1 Outlier detection

Both LSE and SD are based on sum of squared errors. This approach does not give optimal results if the measurement noise is not normally distributed or includes data points with very high deviation. Therefore, reweighted least squares (RLS) /8/ was chosen for the cost function used in this research. RLS is a hybrid method where outliers are first removed from the data by minimizing the median squared error. Thereafter, measurement noise is assumed gaussian and LSE can be applied.

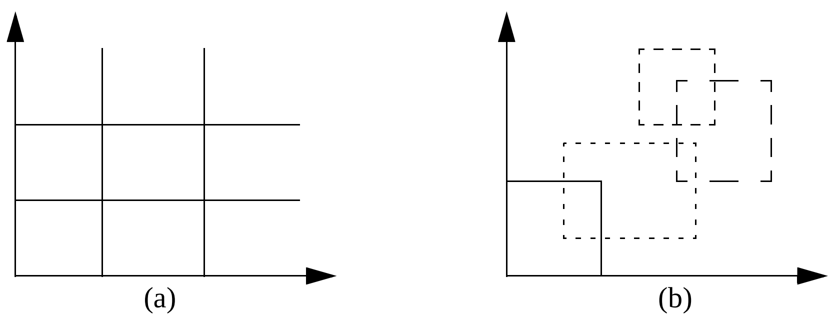

Figure 3. Input space partitioning using (a) grid partition, (b) scatter partition. /6/

3.2.2 Matching inputs and outputs

The data used for modelling was productional data collected at a Finnish tyre manufacturer. There are dozens of different mixing recipes, 30–40 of which had enough data for modelling. Recipe-specific models might be easier to identify, but because of the huge amount of recipes a general model for all recipes was decided to be the goal.

The available data included measurement values only at the discharge moment of the batch, so no measurements from the middle of the mixing time were available.

Laboratory samples are normally used only for rough monitoring of viscosity, hence samples are not taken from every single batch. Even 30 batches can be covered by one viscosity determination. For identification of the ANFIS model however, each output (viscosity) must match exactly one input (vector). This was achieved by preprocessing the data: Each viscosity value was assigned a set of batches, and medians of each process measurement of the set were calculated. As a result, a vector of measurement values matching the viscosity value was attained. Using median instead of average helped in rejecting outliers /8/. Only dates and serial numbers were available for assigning correct batches to each viscosity value, some error may thus be caused by this procedure.

3.3 Input selection

Some a priori knowledge was utilized to find suitable inputs among the process measurements. It is well known that viscosity is strongly temperature dependent. According to Palmgren /1/, power needed by the mixer is directly proportional to shear stress, which in turn is proportional to viscosity. Another quantity affecting viscosity is shear rate that can be measured indirectly from rotor speed, rotor clearance and fill factor of the mixing chamber. Several experiments with other measurements were made, but best results were obtained using the above-mentioned five inputs.

4 MAJOR RESULTS

4.1 Assessment of laboratory determination confidence

Confidence of the laboratory determination for viscosity was estimated by taking 25 samples from one mastication batch, one master batch and one final batch. Average viscosities and standard deviations calculated from these 25 determinations are shown in Table 1. The values may have been affected by batch boundaries because samples were taken from a continuous web after the extruder, where batch boundaries are no longer clear. Because of the small sample size, these deviation values can be considered only suggestive.

The deviations in Table 1 include both local variations in the rubber web after the extruder and errors in laboratory determination. It can be seen that deviation reduces when mixing process proceeds towards final batch. This is natural because prolonged mixing levels down fluctuations in the mixture.

Table 1. Average values and standard deviations calculated from 25 samples of each batch type.

Batch	Average	Standard deviation
Mastication	71,85	2,27
Master batch	76,05	1,06
Final batch	57,97	0,59

4.2 Identification results

The model was identified by first removing outliers with RLS, then clustering the input data and finally tuning parameters using the hybrid learning method described earlier. The subtractive clustering algorithm found seven clusters, so the identified ANFIS model had seven rules for five inputs.

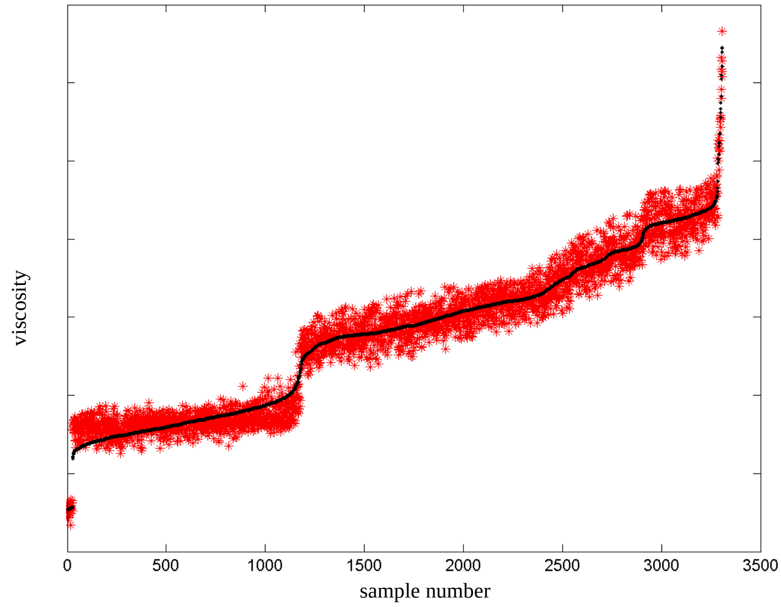

Figure 4. Predicted (solid line) and laboratory-determined (dots) viscosity sorted by predicted viscosity.

Accuracy of the model was estimated with root mean squared errors (RMSE). The RMSE value calculated using a test set distinct from the training data was 3,70. It should be noted that this value includes the inaccuracies shown in Table 1. The R-squared value of the model is 0,968, which means that the model can explain 96,8% of the viscosity variation.

In Figure 4, viscosity predicted by the model is shown together with laboratory viscosity values. For confidentiality, viscosity scale is removed. The data is sorted by predicted viscosity. In the area of approximately 1000 first samples in Figure 4, a slight correlation between predicted viscosity and prediction error can be seen. This is due to one recipe that appears as a small dot group in the bottom left corner of the graph. The recipe contains too few data points to form a cluster of its own, but on the other hand it is too distinct from other recipes to be modelled correctly together with others. The error could be removed by adding a cluster manually.

Residual analysis of the model is illustrated in Figure 5. Prediction error is plotted as a function of laboratory viscosity. The error seems to be dependent on viscosity, which can be explained by the averaging of inputs: Each output corresponds to a median of several input measurements (see section 3.2.2). The error is thus closest to zero when the viscosity is close to the median of the inputs.

5 CONCLUSION

Considering the inaccuracy of sampling and laboratory determination of viscosity, the model performs satisfactorily. An R-squared value of 96,8% among all collected recipes is fairly high, while inside a single recipe the performance might not look as good. A common model for all recipes is however the only practical solution.

The accuracy and confidence of the model is most probably not enough for controlling which is the ultimate goal of this research. Nevertheless, the model achieved here could be used for example in quality control. With the help of predicted viscosity, batches outside tolerance limits could be revealed at an early stage and directed to further mixing phases.

5.1 Needs for further research

The identification described in this paper was, for practical reasons, done entirely using measurement values at the discharge or end moment of the batch. Since the process is not at steady state when discharged, also dynamical data measured during the mixing would be useful. Such data will be utilized at the next stage of the research.

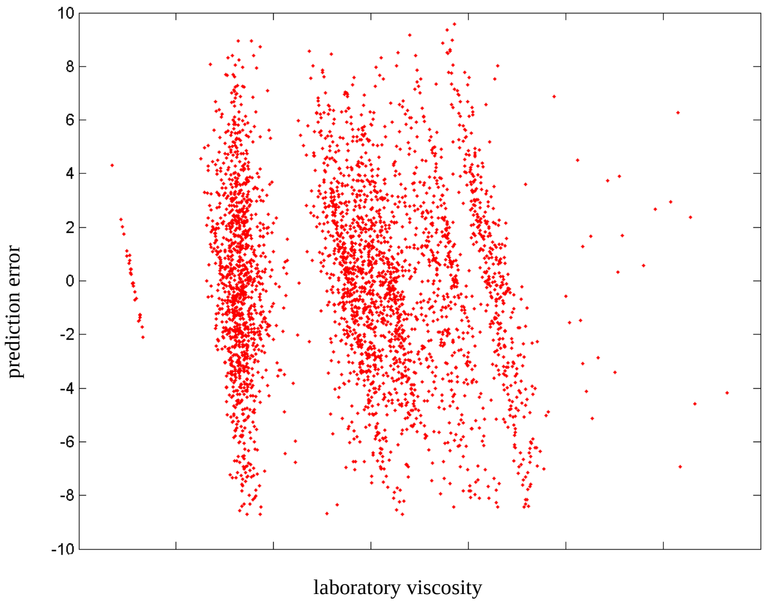

Figure 5. Prediction error as a function of laboratory-determined viscosity.

In the ANFIS model, no a priori knowledge was taken into account. Yet there are numerous physical dependencies concerning rubber mixing known in the literature. These physical features can either be introduced to a certain neuro-fuzzy model structure or be used as a basis for a physical model which is then further tuned using a learning algorithm.

6 REFERENCES

/1/ Palmgren H.: Processing conditions in the batchoperated internal mixer. Rubber Chemistry & Technology, 48(1975)2, pp. 462–494.

/2/ Tiejun M., Cai Q.: Analysis of rubber mixing process in internal mixers. Journal of South China University of Technology, 25(1997)3. (in Chinese)

/3/ Ryzko P., Haberstroh E.: Quality control of the discontinuous compounding process in a rubber internal mixer by regression and neural networks process models. Proceedings of the meeting of Rubber Division, American Chemical Society, Chicago, Illinois, April 13-16 1999.

/4/ Ryzko P., Haberstroh E.: Modelling of the mixing process and on-line-prediction of the rubber compound and molded part properties with intelligent process analyses technologies. Proceedings of the meeting of Rubber Division, American Chemical Society, Dallas, Illinois, April 4-6 2000.

/5/ Sorsa S., Tuomisto T., Ylitalo R.: Elastomeerien mahdollisuuksia. Nokian Tekninen Kumi, Nokia, Finland, 1989, 101 p. (in Finnish)

/6/ Jang J.-S.R., Sun C.-T., Mizutani E.: Neuro-Fuzzy and Soft Computing. Prentice Hall, USA, 1997, 614 p.

/7/ Chiu S. L.: Fuzzy model identification based on cluster estimation. Journal of intelligent and fuzzy systems, 2(1994)3.

/8/ Rousseeuw P.J., Leroy A.M.: Robust regression and outlier detection. Wiley-Interscience, New York, 1987, 352 p.