Nonlinear models guide efficient operation of hard nip sizer to derive more strength from less starch

H. Vaittinen and A. Räisänen, Valmet Technologies, Finland and A. Bulsari, Nonlinear Solutions Oy, Finland

INTRODUCTION:

All kinds of production processes can be made more efficient. When process variables of any process are properly optimised, it is possible to achieve better quality with lower raw material costs. For example, it is not difficult to derive more strength with less starch or fibres with hard nip sizing, once quantitative information about the relations between strength and process variables in a mathematical form is available. These relations are not very simple or linear, which is why conventional linear statistical techniques are not very effective. However, nonlinear modelling proves once again to be a powerful tool for describing these relations.

Nonlinear empirical and semi-empirical modelling has been used widely in several industrial sectors for process development as well as materials development. This article describes our experience with nonlinear modelling of hard nip sizing from pilot scale experimental data, with a short description of the method. The nonlinear models have been implemented in software and are now used to calculate cost optimal operating conditions such that the final properties of liner board are within desired limits, resulting in more strength with less starch.

Hard nip sizing

Surface sizing is an essential process in the pulp and paper industry for improving the strength properties of base paper or board. In film sizing, the starch application takes place mostly on the outer surfaces of the paper or board, and only a minor portion of the starch penetrates deep inside the structure of the base paper or board. Hard nip sizing (Figure 1) overcomes this weakness by using pressure, allowing for much deeper penetration starch and other sizing chemicals [1], and thus increasing the strength of properties more than conventional surface sizing processes.

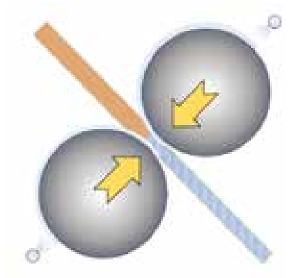

Figure 1: The principle of hard nip sizing is to use higher nip loads

It also produces better smoothness because hard rolls work like in calendering. As can be seen in the photograph (Figure 2), hard nip sizer rolls and the loading system are like calenders rather than conventional sizer rolls. In hard nip sizing process, the nip pressure results in an optimal packing of fibres and the sizing through the z-direction of the web, resulting in bigger increase of SCT in cross direction and burst strength [2].

Process development

All kinds of production processes can be made more efficient. The aim of process development is often to make processes more efficient, sometimes by saving production costs while also improving quality. There are usually a few degrees of freedom in the process which affect the product properties and may also affect the production rate. One would like to determine the best values of these variables such that the resulting product properties will be within desired limits.

This requires quantitative knowledge of the effects of relevant variables in a precise and concise form. In other words, we would need the knowledge of the process in a mathematical form, which allows for various kinds of calculations. Developing these equations is called mathematical modelling, which can be performed in several different ways.

Figure 2: Hard nip sizer at the pilot plant of Valmet in Järvenpää

Mathematical modelling

Mathematical models can be used instead of experimentation if they are reliable enough. Mathematical models also permit the user to carry out various kinds of calculations, like determining suitable values of variables which will result in desired product quality in an economic way. Mathematical modelling can be performed in various ways, and different ways are suitable in different situations. Mathematical models represent knowledge of quantitative effects of relevant variables in a concise and precise form.

Physical or phenomenological modelling is not particularly effective for predicting material properties like strength, thermal conductivity or solubility. Physical modelling usually requires a lot of assumptions and simplifications. Empirical and semi-empirical modelling, on the other hand, does not need any major assumptions or simplifications. Empirical models simply describe the observed behaviour of a system. Empirical modelling is feasible when the relevant variables are measurable.

Conventional techniques of empirical modelling, however, are linear statistical techniques. These tend to have serious limitations because nothing in nature is very linear, and particularly so in process engineering and materials science. It therefore makes sense to use better techniques of empirical and semi-empirical modelling which take nonlinearities into account.

Nonlinear modelling

There is hardly any material behaviour which is absolutely linear. It is therefore wise to treat the nonlinearities rather than ignore them. The proponents of linear techniques draw on their simplicity and the possibility of adding nonlinear terms in linear regression. Often this is not done, and is not efficient even if it is done. Nature does not follow the simplicities that we try to fit it in, using common linear techniques.

Nonlinear modelling is empirical or semi-empirical modelling which takes at least some nonlinearities into account. Nonlinear modelling can be carried out with a variety of methods. The older techniques include polynomial regression, linear regression with nonlinear terms and nonlinear regression. These techniques have several disadvantages compared to the new techniques of nonlinear modelling based on free-form nonlinearities.

The newer methods like feed-forward neural networks and series of basis functions do not require a priori knowledge of the nonlinearities in the relations. Among these new techniques, feed-forward neural networks have turned out to be particularly valuable in chemical engineering [3] and materials science. Feed-forward neural networks have several features which make them better tools for nonlinear empirical modelling. Besides their universal approximation capability [4], it is usually possible to produce nonlinear models with some extrapolation capabilities with feed-forward neural networks.

There are many different types of neural networks, and some of them have practical uses in process industries. Neural networks have been in use in process industries for about 30 years. The multilayer perceptron, a kind of a feed-forward neural network, is the most common one. Most neural network applications in industries are based on them [5].

Feed-forward neural networks resemble structurally and to a smaller extent functionally the networks of neurons in biological systems. Like the networks of neurons in the brains, artificial neural networks also consist of neurons in layers directionally connected to others in the adjacent layers (see Figure 3).

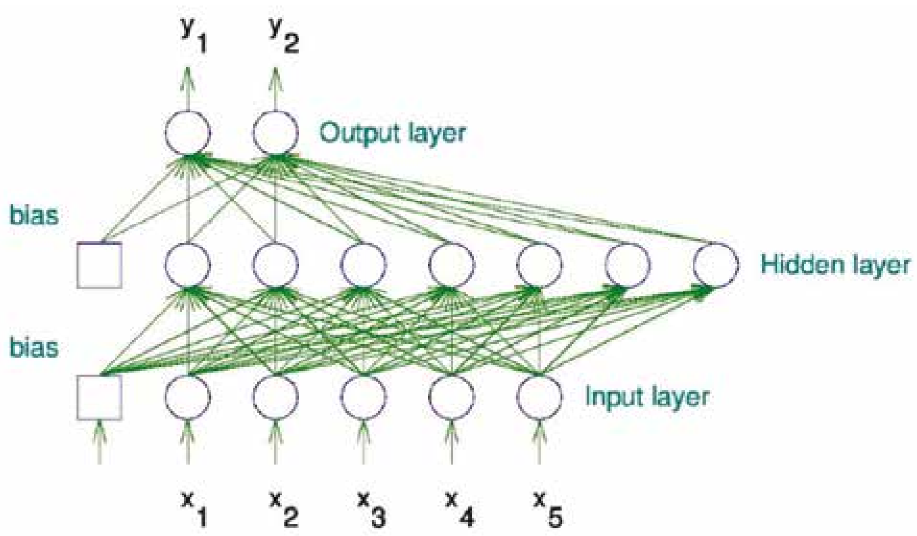

Figure 3: A typical feed-forward neural network

In a feed-forward neural network of the kind shown in Figure 3, the output of each neuron ***i*** in the feed-forward neural network is usually given by

$$z_i = \sigma\left(\sum_{j=0}^{N} w_j x_j\right)$$

where σ is called the activation function, and is usually the logistic sigmoid, given by

$$\sigma(a) = \frac{1}{1+e^{-a}}$$

The incoming signals to the neuron are x_j, and w_{ij} are the weights for each connection from the incoming signals to the i_{th} neuron. The w_{i0} terms are called biases. Then the output is simply calculated as the weighted sum of the outgoing signals z_i from the neurons in the hidden layer. This results in a set of algebraic equations which relate the input variables to the output variables. Thus, for each observation (a set of input and output variables), the outputs can be predicted from these equations based on a given set of weights. The training procedure aims at determining the weights which result in the smallest sum of squares of prediction errors. Today, most people use good optimisation methods for that purpose.

Nonlinear modelling in process engineering

Nonlinear modelling has been utilised successfully for various industrial sectors including plastics and rubbers, metals, cement and concrete, medical materials, semiconductors, ceramics, mineral wools, glass, power generation, biotechnology, pulp and paper, etc. Different processes have different characteristics - different raw materials, different compositions, and are produced by different batch, continuous or fed-batch processes. However, some things are common to modelling of various kinds of processes. Material properties or product properties, production rate and production economics depend on composition variables (or feed characteristics), process variables and dimension variables, as as summarised in Figure 4.

For process development, one would like to determine the best values of composition variables (or feed characteristics), process variables and/or dimension variables such that the resulting product properties will be within desired limits, with a good production rate or at a minimal production cost. Sometimes, feed characteristics or composition variables might be constants. In more common situations, the process variables may be constant or dependent variables, and the only degrees of freedom in materials development may be the composition of the feed, the amounts of raw materials and possibly dimension variables.

The problem looks somewhat similar from the modelling point of view for a wide variety of materials and processes. Nonlinear models combined as shown in Figure 4 make process development more efficient by reducing expensive experimentation and by helping achieve better combinations of product properties, often optimised for cost.

Sometimes the composition variables or feed characteristics might be constants. In other situations, the process variables may be constant or dependent variables, and the only degrees of freedom in materials development may be the composition of the feed, the amounts of raw materials and possibly dimension variables. This is often referred to as recipe development. In this case, we have freedom in both composition as well as process variables.

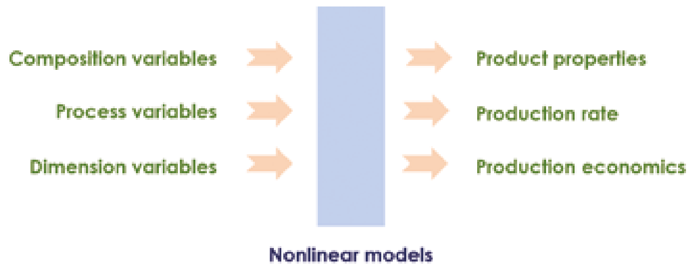

Figure 4: Composition variables, process variables and dimension variables determine product properties, production rate and production economics

Experimentation

Nonlinear modelling needs either experimental or production data. A lot of experiments are carried out at the pilot plant in Järvenpää round the year. From 22 series of such experiments from 2019 to 2021, a total of 970 usable observations were collected. The equipment allows hard nip sizing as well as film sizing, and the data contains a small fraction of film sizing results also. These include a wide variety of papers and boards, particularly liners with modified starches of different viscosities. These experiments were used for the model development work.

If nonlinear models are to be developed for a single paper mill, a small number, probably 25 to 30, would have been sufficient for developing nonlinear models, if the experiments had been planned keeping in mind that nonlinear models would be developed based on that data. Besides SCT index in the cross direction and burst index, air porosity, thickness and density were also measured from each of the experiments.

Nonlinear model development

From the raw data set, it was possible to see the effects of certain variables. For example, the higher the base paper's basis weight, the lower is the increase in SCT index and burst index. Higher nip loads also produce larger increases in SCT index as well as burst index. The raw experimental data was analysed and preprocessed, after which nonlinear models were developed and tested using the NLS 020 software. The experimental data taken into use was fairly consistent and of good quality, and as a consequence, good nonlinear models could be developed.

Nonlinear models of SCT CD and burst indices

Nonlinear models in the form of neural networks with a single hidden layer were attempted and tested to predict SCT CD index, increase in SCT CD index, burst index as well as the increase in burst index over unsized paper. The rms error (roughly speaking, the standard deviation of prediction errors) of SCT CD index was around 1.2 J/g while the rms error of burst index was 0.13 kN/g which amount to about 5% in terms of fractional errors for both. Figure 5 shows the burst index predicted from the nonlinear model plotted against the 970 measured values from 22 series. The model predictions look close enough to the measurements. It is natural that the nonlinear models perform well since the effects are not very linear, while the linear models will not hesitate to predict even negative values of the indices.

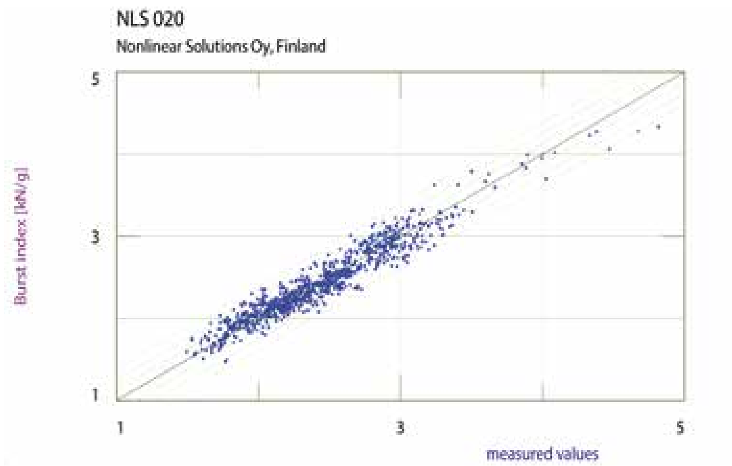

Figure 5: Burst index predicted from the nonlinear model plotted against measured values

Implementation of the models in software

Nonlinear models in the form of neural networks are not simple equations. The equations are clumsy and unwieldy, and not easy to work with. Engineers, let aside plant operators, cannot be expected to be familiar with such mathematics. It is therefore imperative to implement the models in software which makes the use of models easy for anyone. LUMET systems are a set of software components which are assembled depending on the needs of the users. In all LUMET systems, the central point is the prediction screen (Figure 6), where the user can feed in the values of the input variables on the left side, and can predict the outputs shown on the right side. Besides SCT CD index and burst index, there are also economic variables like raw material cost. Figure 6 shows a typical prediction calculation. A lot more things can be done once the models have been developed and implemented in software like this.

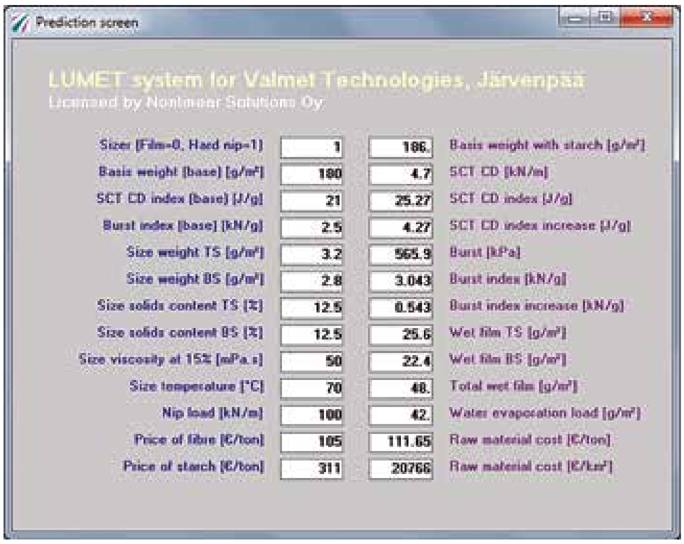

Figure 6: SCT CD index, burst index, raw material cost and several other consequences of sizing are predicted using the nonlinear models on the prediction screen

Besides predicting the output values which are various consequences of operating the process, one can see the effects of each of the input variables on the outputs in different kinds of plots. Figure 7 shows a set of curves of burst index plotted against size weight on bottom side for different size temperatures, while keeping other input variables constant. It is easy to see that the higher the size weight, the higher is the increase in the burst index, but a limit is reached. Higher temperatures have a better effect.

Pairs of variables often have interaction (synergistic) effects, which can be seen from this kind of plots, or from contour plots. Figure 8 shows contours of SCT CD index on a plane of nip load and basis weight, keeping other input variables constant. Higher nip loads naturally lead to better sizing and the benefit is higher for smaller basis weights since the size weights are constant. Surface plots can also be prepared in LUMET systems.

Optimisation calculations

There could be several purposes of developing mathematical models. In the beginning, one wants to use models to design the process and its equipment. Once the process equipment exists, it is natural to want to find out how best to operate the process. In other words, the objective could be to derive maximum production and quality with minimal costs and time. More specifically, for sizing, the aim could be to derive more strength from less starch.

Optimisation helps derive the maximum benefit from the process. The objective of process development is usually to determine optimal or near optimal process operation conditions. Once we have the quantitative knowledge of the process in terms of variables of interest to us in the form of equations, it becomes possible to determine good operating conditions. We would often like to derive a certain SCT CD index and burst index at minimal cost, while taking into account constraints on various other variables.

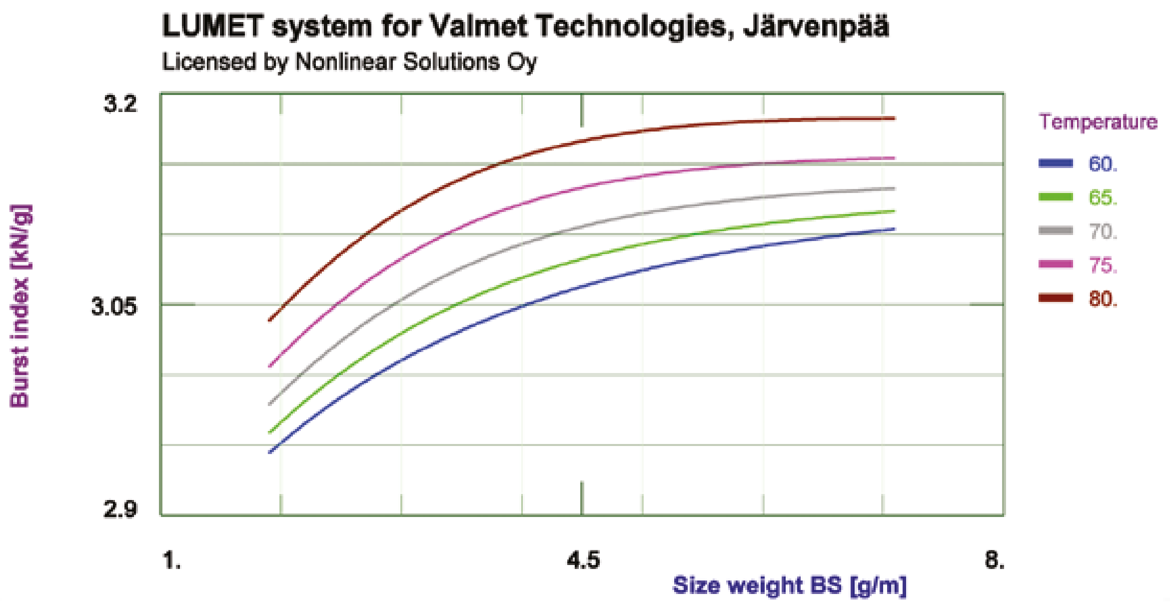

Figure 7: Burst index plotted against size weight on bottom side for different sizing temperatures

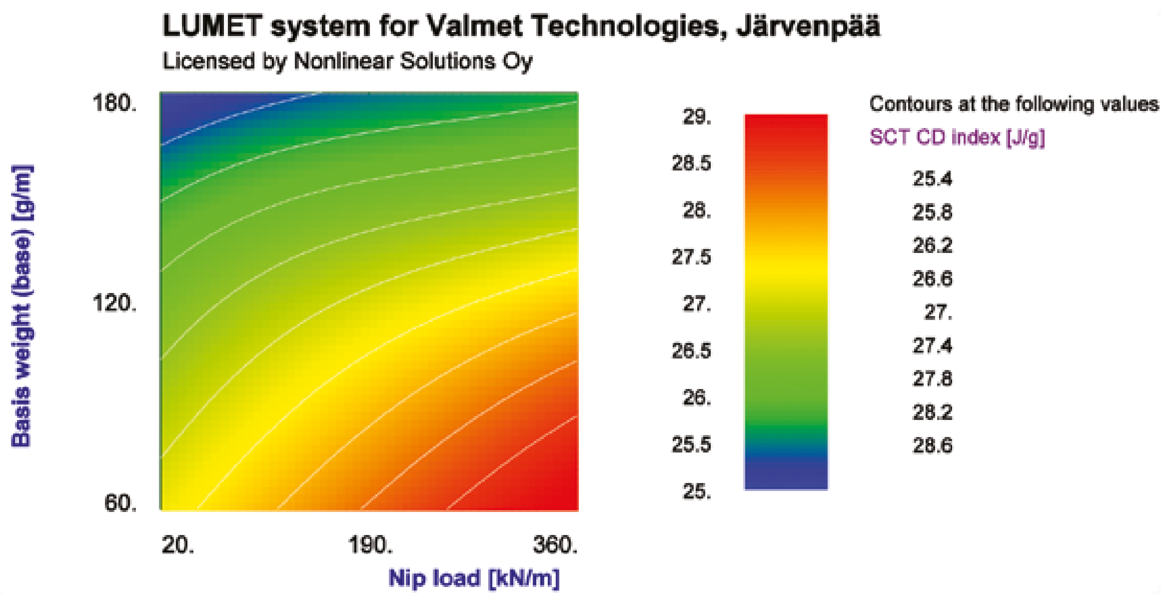

Figure 8: Contours of SCT CD index on a plane of nip load and basis weight

These calculations are now easily done with the nonlinear models implemented in a LUMET system, resulting in significant savings in sizing chemicals. Figure 9 shows one such calculation. The base under consideration is a 180 g/m² liner with a SCT CD index of 21 J/g and a burst index of 2.5 kN/g. With a given starch, we have to produce a SCT CD index of at least 26 J/g and a burst index of at least 3 kN/g, with minimal cost per ton of product. There is an additional limitation on the wet film. The calculated optimal conditions are in the last column. A size weight of 2.63 g/m² on the top side and 1.47 g/m² on the bottom side to be applied at a temperature of 80°C with a nip load of 360 kN/m should minimise the raw material cost, according to the calculation.

We could have added a variable SCT CD index increase divided by the total size weight and maximised that. However, the calculation of Figure 9 is more realistic. Usually, there are specifications on SCT CD index and burst index, and it makes sense to minimise the cost instead of maximising the strength increase for a given amount of starch.

LUMET system for Valmet Technologies

Licensed by Nonlinear Solutions Oy

	minimum	maximum	answer
Sizer (Film=0, Hard nip=1)	1	1	1.0
Basis weight (base) [g/m²]	180	180	180.0
SCT CD index (base) [J/g]	21	21	21.0
Burst index (base) [kN/g]	2.5	2.5	2.5
Size weight TS [g/m²]			2.6283
Size weight BS [g/m²]			1.4671
Size solids content TS [%]	14	14	14.0
Size solids content BS [%]	14	14	14.0
Size viscosity at 15% [mPa.s]	50	50	50.0
Size temperature [°C]			80.0
Nip load [kN/m]			360.0
Price of fibre [€/ton]	105	105	105.0
Price of starch [€/ton]	311	311	311.0
SCT CD index [J/g]	26		26.001
Burst index [kN/g]	3		3.0
Basis weight with starch [g/m²]			184.0954
SCT CD index increase [J/g]			5.001
SCT CD [kN/m]			4.7867
Burst index increase [kN/g]			0.5
Burst [kPa]			552.2857
Raw material cost [€/km²]			20173.65
Raw material cost [€/ton]	Minimum	found:	109.5826
Wet film TS [g/m²]		20	18.7736
Wet film BS [g/m²]		20	10.4789
Total wet film [g/m²]		30	29.2525
Water evaporation load [g/m²]			25.1571

Figure 9: An optimisation calculation to minimise raw material cost per ton while keeping SCT CD index above 26 and burst index around 3 in presence of other constraints

CONCLUSIONS

- **All kinds of processes can be made more efficient.**
- **One can derive a lot more value from a process by tuning it well.**
- **Hard nip sizing is a superior process to film sizing.**
- **The advantage of a good process can be maximised by process optimisation.**
- **Process optimisation needs a good mathematical description of the process.**
- **Nonlinear modelling is an efficient tool to describe various kinds of processes.**
- **Composition variables, process variables and dimension variables affect product properties in a complicated manner, and people with even decades of experience cannot predict the combined quantitative effects of the relevant variables.**
- **Nonlinear modelling can utilise production data or experimental data.**
- **Nonlinear modelling does not need any significant assumptions.**
- **A good process with optimisation based on nonlinear models is a very powerful combination.**
- **Significant reductions in production costs can be achieved with a modest effort like this.**

References

[1] J. Hakola, “Surface sizing with high nip pressure”, presented at the TAPPI conference 19PaperCon, Indianapolis, USA (May 2019)
[2] E. Pitschmann, “New technology HardNip PM1”, presented at the 28th International Munich Paper Symposium, Munich, Germany (March 2019)
[3] A. Bulsari (ed.), Neural Networks for Chemical Engineers, Elsevier, Amsterdam, 1995.
[4] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators”, Neural Networks, Vol. 2, (1989) 359-366.
[5] A. Bulsari, “Quality of nonlinear modelling in process industries”, Internal Report NLS/1998/2.