Sanitary risk detection for a safer food chain management

S. TAMAYO[1,3,*], T. MONTEIRO[2], N. SAUER[1]

[1]University Paul Verlaine / LGIPM – France ({simon.tamayo , nathalie.sauer}@univ-metz.fr)
[2]University of Saint Etienne / LASPI – France (thibaud.monteiro@univ-st-etienne.fr)
[3]Adents High-Tech International – France (stamayo@adents.com)

Abstract: *Nowadays, several risk-attitude elicitation techniques are applied in order to guarantee the highest levels of safety and quality control in the food chain. Here we explore an approach to the risk measurement problem of any multi-parameter production framework using precise decision making tools. The application model of artificial neural networks is presented in order to determine the global risk related to a given production. The proposed tool uses a feed-forward network able to treat any set of nonlinearly separable production markers. The risk values obtained with this approach are incorporated into the expedition management system in order to perform smarter deliveries and more accurate sanitary controls.*

Keywords: *Food chain safety, artificial neural networks, multi-layer neural network, optimization of logistics choice, sanitary risk, traceability.*

1. Introduction

Biological, physical, or chemicals hazards may cause products to lower the expected quality levels, making them unsafe and risky. The goal of a risk analysis system is to examine certain factors that may lead to out-of-control hazards, in order to guarantee quality coordination in the production/expedition systems.

As a result, hazards are a huge threat to businesses, becoming ticking bombs in the agro-alimentary industry. Unless they are kept under control, they could result in financial ruin for producers. Given the changing structure of the agricultural industry, managing risk has become vitally important to the success of agricultural operations (Drollette, 2009). This is why management principles based on risk are currently being utilized in many areas of business and government including finance, insurance, occupational safety, public health, pharmaceutical industry and by the agencies regulating these industries (Ionica et Al., 2007).

At the present time, several safety management systems are implemented in order to reduce the occurrence of risk factors, the most known is "active managerial control" (Kurtzweil, 1999), which usually refers to the purposeful incorporation of specific procedures by industry management into the production processes to attain control over food borne illness risk factors. It embodies a preventive rather than reactive approach to food safety.. One of the most used systems is the HACCP, Hazard Analysis and Critical Control Points (Brandriff, 2008). Nevertheless, these systems focus on determining the control measures that can be utilized to eliminate, prevent, or reduce food safety hazards; therefore these systems lack the technical solutions to link both the monitoring and the operations stages. In this article we underline the importance of linking both the observation and the action, and we present a tool that detects when the process or procedures will fail to meet the critical risk limits to allow the controllers to establish corrective measures.

As of now, different approaches of active managerial control are applied in large processing plants to ensure the quality of production and to reduce risks, but it has been particularly difficult to impose them on smaller producers (Enamul Haque, 2003). This is why this research project appears in the context of the exploitation of existing traceability solutions, offering a system that not only provides a means to respect the legal constraints (in terms of traceability), but also monitors its behavior in terms of risk, quality and efficiency, analyzing all the measurable inputs of a given production; this is done through a continuous structure of monitoring and decision making, assessed by the utilization of artificial neural networks.

2. Global Approach

To achieve the goal of detecting the significant risk levels in any given fabrication process, a previous knowledge of the production in terms of “recall risk” must be incorporated; accordingly production and traceability indicators are computed to generate a measure of the global risk, named “criticality”. The presented tool appears as a real-time monitoring system that continuously computes the production markers, linked to the outgoing batches, seeking for abnormalities as shown in Figure 1.

[*]Corresponding Author. Address: Adents High-Tech International, 1 Allée des garays, 91120, Palaiseau, France.
E-mail: simon.tamayo@univ-metz.fr

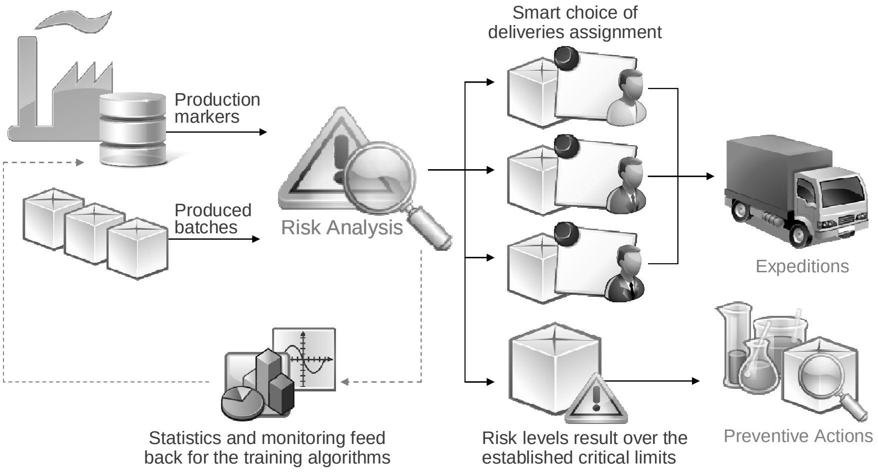

Figure 1: Schematic view of the global approach of de risk detection system.

An incorporation of the criticality values into the packing and distribution operation may ensure that corrective actions are taken when the risk analyzing system determines (and records) that the critical limit has not been met, as shown in Figure 1. According to this, when important criticality levels are detected the production is locked and tested; also, when the production remains within the acceptable levels of global risk, it is important to assign the outgoing batches to the delivery orders in the smartest manner accomplishable, considering the possible recall situations. It is imperative to distinguish the risk of sending a potentially perilous production to an important customer, or to the big distribution channels (which supposes a dispersion increase over the external chain). Thereafter, one must take into account the most representative criteria and constraints regarding the actual deliveries and clients.

Initially, a risk level can be predicted from a set of production data, but as soon as a real risk emergency situation takes place, the tool presented here can be re-trained in order to exactly identify these situations in the near future and prevent them altogether.

This document has been organized as follows: the background and the global approach are outlined in, section 3, where a literature review on the subject is presented; followed by section 4, which details the proposed stance and decision process. In section 5, an introductory example is presented, finally in section 6, conclusions and prospects are suggested.

3. Literature review

The risk impact and its ontological requirements have been studied during the past years (Borst et al., 1997). The relevance of product tracing in both the external supply chain, as well as inside the production system has been considered (Ramesh et al., 1997). The problem of optimization in production was also considered from the traceability point of view, specifically regarding the raw material dispersion problem (Dupuy et al., 2005). Previous models integrating traceability initiatives with operation factors to achieve desired product quality and minimum impact of product recall have been presented, especially for the definition of an optimal batch size (Wang et al., 2007). In this article we present an approach to the risk detection problem as part of a research project, deployed in relation to a software provider of the traceability domain. This approach which seeks to offer a smart delivery's module, is aiming to find the optimal expedition parameters, in order to recall a minimum number of final products in case of a crisis.. Accordingly, a risk measure must be incorporated after computing the possible hazards, and subsequently having deployed specific actions or procedures in order to attain control over risk factors and quality breakdowns.

The system we present seeks to build and train an artificial neural network, which allows producers to consider all possible types of indicators. Indeed, neural network models can be used to infer a function from observations even if the complexity of the data or task makes formal design of such a function difficult (Najafi et al., 1998), (Yu et al., 2005). To determine the criticality value, several types of information may be considered; they may be qualitative (example: quality of suppliers) and quantitative (example: duration of the production cycle). The value to be obtained will depend on the industrial production context and has to be adapted to each situation. Hence it is difficult to make a formal design of criticality function.

Studies of decision making tools applied for diagnosis in mechanic systems have been presented (Angeli et al., 1999), the application of artificial neural networks in defaults monitoring has been studied for the production

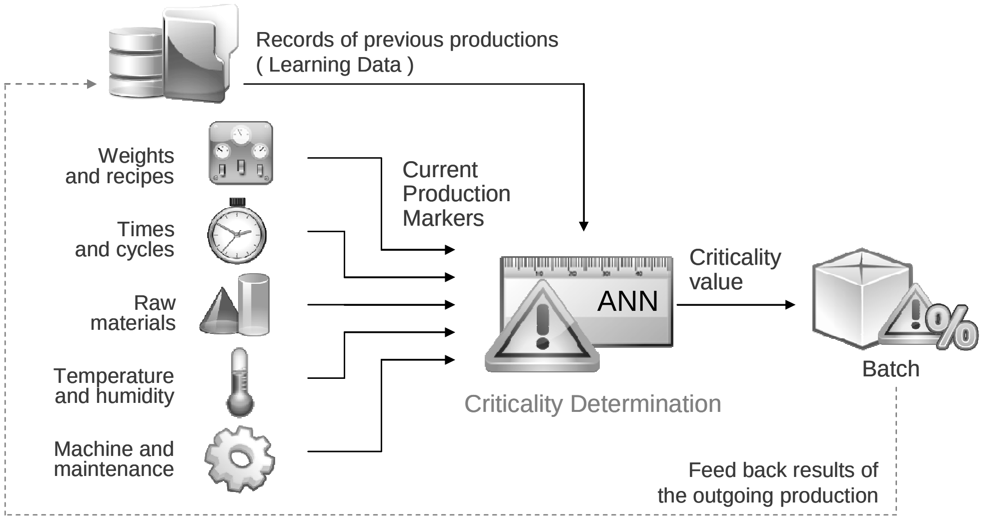

Figure 2, Measure of risk for a given batch by the evaluation of its production markers

of motors (Bayır et al., 2004) and most precisely the use of type classifiers for artificial neural networks has been explored in the diagnosis of pneumatic processes (Karpenko et al., 2002). An expert in production processes can give an evaluation of this criticality value without formalizing the inference process. Neural networks, as systems capable of learning, operate on the principle of "induction", i.e. acquiring training by experience; so they are fully successful in solving the problem of determining the criticality value.. Furthermore, this determination can be looked as a statistical problem of a filtering type with possibly noisy data input. It is a problem for which the neural networks have shown their effectiveness (Shin et al., 1992) and (Martin and Howard, 1996).

The current state of technology does not offer formal approaches to the risk identification process in the food chain, linking both the monitoring and the operations stages, and is what the work presented here aims to do.. Another parallel objective is to exploit and take advantage of traceability information, while providing accurate tools for risk diagnosis, quality measuring and piloting.

4. Measure of risk

The measure of risk consists of monitoring, observing and measuring specific operational steps in the fabrication process, to determine if the critical limits of risk are being met. This action can take place at any point of the fabrication, being more convenient at every step that a material flow reaches a critical stage and the system must ensure that the material's properties are under control and that there will be no further contradictions with additional steps of production nor higher risk levels. Monitoring will identify when there is a loss of control or a trend towards a loss of quality so that corrective actions can be taken.

In this particular application "criticality" is defined as a value associated to a production batch that represents its quantitative state of global risk. Criticality makes it possible to take into account several parameters of manufacture simultaneously with only one value exposing the potential danger. The application's goal is to obtain coherent criticality values, exploiting the process information provided by the traceability framework; therefore, it is a matter of multi-criteria decision and assigning a criticality value to a list of entries as shown in Figure 2. For this purpose, we propose the use of a multi-layer neural network trained with the back propagation algorithm. Once the network is designed, a simple reset and training will enable it to be exploited in a different field of production with other input parameters. The multi-layer network seeks to mechanize the separation of data into categories, or classes, characterized by a distinct set of features, which will classify our given set of entries as a corresponding "class" of criticality.

4.1. Utilization of a multilayer feed-forward network

Feed-forward neural networks are the most popular and most widely used models in practical applications. It consists of a possibly large number of neurons, organized in layers. As shown in Figure 3 the first layer is composed of a set of inputs called the input layer; in the proposed application it represents the vector of production markers to analyze; the set of output units is called the output layer from which the criticality value is obtained. All other layers with no direct connections from or to the outside are called hidden layers. Usually the units in a layer are not connected to each other (although some neural models make use of this kind of architecture). Data enters at the inputs and passes through the network, layer by layer, until it arrives at the outputs..

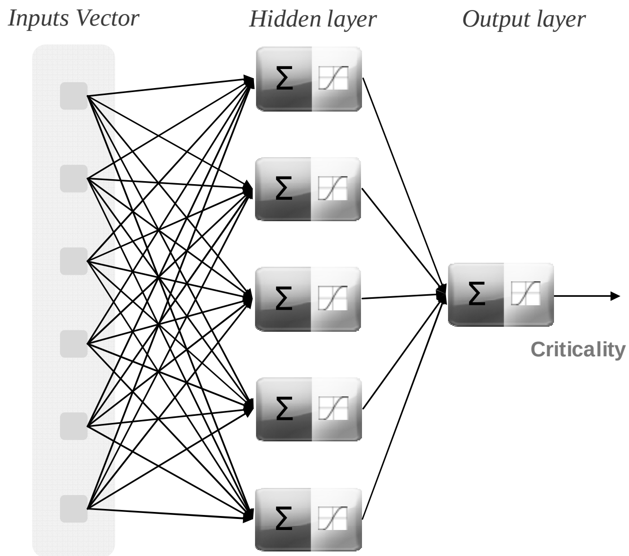

Figure 3, Architecture of the proposed feed forward network with a single hidden layer

During normal operation there is no feedback between layers. This is why they are called feed-forward neural networks (Boresma, 1998).

4.2. Formal definition of a feed-forward neural network

The input layer consists of the vector containing the inputs to the network. Afterwards, a hidden layer follows which consists of a set of neurons placed in parallel.. Each neuron performs a weighted summation of the inputs $\boldsymbol{j}$, which then passes a nonlinear activation function $\boldsymbol{\sigma}$, also called the neuron function.

$$f(X) = \sigma\left(\sum_{j=1}^{n}\left(w_j * x_j\right) + b_j\right) \qquad (1)$$

The nonlinear activation function in the neuron is usually chosen to be a smooth step function (Rojas, 1996). The activation function is frequently a sigmoid as shown in Equation 2. In which $\boldsymbol{k}$ determines steepness of the function and $\boldsymbol{a}$ establishes the amplitude of the output.

$$\sigma(x) = \left(\frac{a}{1+e^{(-k*x)}}\right) \qquad (2)$$

4.2.1. Using the Backpropagation algorithm

In order to train a neural network, we must adjust the weights of each unit in such a way that the error between the desired output and the actual output is reduced. This process requires that the neural network computes the error derivative of the weights. To this end, we consider that some of the examples are non-linearly separable (for instance if a batch contains a raw material that has reached by far its expiration date, the risk level has to be set as the highest, even if the rest of the production markers are optimal) and so the training rule used for our industrial situation is the delta rule.
The main idea of delta rule is to use gradient descent in order to search the space of possible weights to find the weights vector that fits the best to the training examples.. This rule provides the basis for the <u>*backpropagation*</u>. algorithm, which follows the next steps for each training example:

- Initialize all weights and bias with random values.
- Input the training example and compute the network output.
- Propagate the error $\boldsymbol{\delta_k}$ backwards in the last layer for each output unit $\boldsymbol{k}$ (see equation 3).
- Propagate the error $\boldsymbol{\delta_h}$ through hidden neurons $\boldsymbol{h}$ in the network (see equation 4).
- Update the weights $\boldsymbol{w_{ij}}$ connecting the units/inputs $\boldsymbol{i}$ with the units/outputs $\boldsymbol{j}$ in the network by adding a delta value $\boldsymbol{\Delta w_{ij}}$, which is proportional to the error backwards $\boldsymbol{\delta_j}$, the unit's input $\boldsymbol{X_{ij}}$ and the learning rate $\boldsymbol{\alpha}$, as shown in equation 5.

$$\delta_k = \frac{d\sigma_k(\mathrm{x})}{dx} * (TARGET_k - OUTPUT_k) \qquad (3)$$

$$\delta_h = \frac{d\sigma_h(\mathrm{x})}{dx} * \sum_{k=1}^{neurons}\left(w_{h,k} * \delta_k\right) \qquad (4)$$

$$w_{ij} = w_{ij} + (\alpha * \delta_j * x_{ij}) \qquad (5)$$

In the proposed system the notion of momentum $\boldsymbol{m}$ is applied by making the weight update on the n^{th} iteration depending, partially on the update that occurred during the $(n-1)^{th}$ iteration, as follows and is presented in equation 6.

$$\Delta w_{ij} = \alpha * \left((m * \Delta w_{ij}) + (1-m) * (\delta_j * x_{ij})\right) \qquad (6)$$

5. Industrial application of risk detection

The Figure 4 displays the configuration of the production markers to analyze. Before using the artificial neural network, a training database containing the list of examples must be formalized. The parameters establishing the production criticality were defined as follows:

- **Raw materials supplier's note:** Including raw material suppliers and dry material suppliers (plastic and paper). It is an average note of the suppliers group associated to the current production. It varies

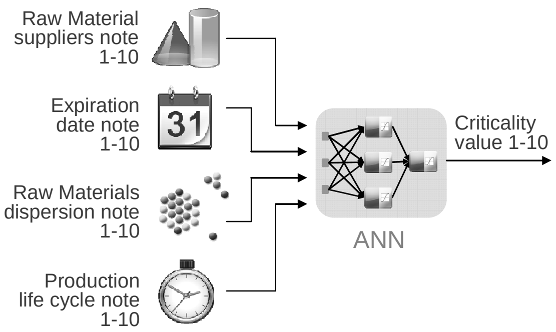

Figure 4, Architecture of the proposed feed forward network with a single hidden layer

from 1 to 10, 1 being the equivalent to very bad suppliers and 10 to the optimal suppliers. In the food industry there is a real presence of high and low quality suppliers, manufacturers must adapt their production to the demand forecasts with the available raw materials which means in certain cases, the use of low quality raw materials.

- **Expiration date note:** This value takes into account the expiration date of the raw materials by assigning the nearest expiration date of the raw materials involved to the finished products. This note value varies from 1 to 10 days, 1 being the note if a raw material's "best before date" has been reached.
- **Raw materials dispersion note:** This note represents the ratio between the production's real dispersion and the optimal dispersion that the production could have had (Tamayo et al. 2008). It varies from 1 to 10 (1 being a bad dispersion which means that the raw materials batches are being scattered too much, this would represent an augmentation of the recall size and costs in case of a sanitary hitch).
- **Production life cycle note:** This value varies from 1 to 10, 1 being a production without stops or inconveniences (optimal lifecycle in the production) and 10 being a dreadful production time and lifecycle less appropriate. It is important to consider the number of stops in production because the life cycle of food products is particularly sensitive to changes in state and temperature. For example, frozen meat must remain on the machines blenders a specific time, and a production stop during the mixing process can represent a significant change in the BBD (best before date) or in the product's sanitary risk level.

After considering the presented criteria, a set of examples was generated by means of an expert in the food production domain.

5.1. Construction and Training

The proposed network to this particular example is a 3 layered network with one single output, that the number of neurons in the hidden layer may modified according to the linear separability of the input data. For instance the hidden layer contains 12 units. The construction starts by randomly assigning the network's initial weights and biases. Then the network is modified according to the loaded examples by adapting the weights and biases to them. In order to ensure an optimal efficiency, the training is carried out. As a termination condition the number of training iterations is limited. In order to measure the behavior of the network, the mean square error is calculated each iteration, as shown in the equation 7.

$$E(\vec{w}) = \frac{1}{N}\sum_{k=1}^{N}\left(TARGET_k - OUTPUT_k\right)^2 \qquad (7)$$

The Figure 5 presents the evolution of ***E*** along 1000 training epochs. At the end of the learning iterations the error result was 0.057% which is a very convenient rate. A different result curve of convergence would result using different values for the learning rate and momentum; it is up to the user to adapt the training procedure to each different risk situation.

The configuration of the network may evolve and adapt to special situations, though it would need to receive extra training. As an example, the table 1 shows the criticality results for a random set of input parameter values.

The results given by the artificial neural network are logical. It can indeed re-detect any hazardous situation given in the training examples and also predict new values of high risk that remain rational to our expertise. The artificial neural network developed gives a very consistent set of criticality results, which could be directly exploited to detect risk situations and may be considered as significant re-engineering criteria. These results confirm a good choice for the type of network and open new horizons to continue the development towards other future optimizations.

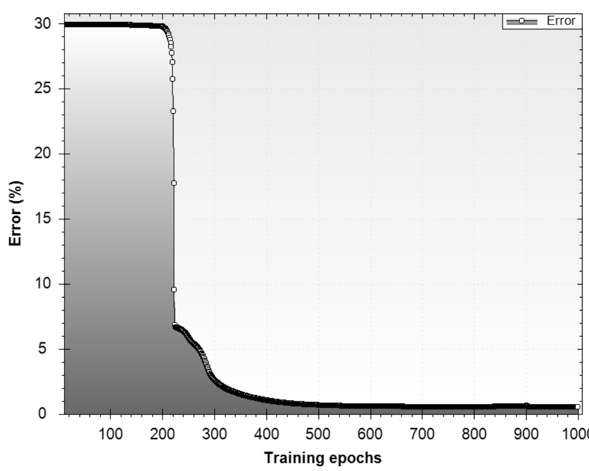

Figure 5, Error evolution for 1000 epochs of training Learning rate α=0.01 Momentum m=0.5

Table 1

Dispersion	Suppliers	Cycle	Date	**Criticality**
10	10	10	10	**1**
1	1	1	1	**10**
1	1	5	5	**9**
10	10	1	1	**9**
6	9	10	8	**3**
2	3	10	9	**6**
8	8	8	8	**2**
2	2	10	10	**8**
10	8	6	3	**5**
5	1	5	1	**10**
5	5	5	5	**6**
10	7	6	5	**3**
2	2	5	5	**7**

This kind of approach may also be used as a tool for analyzing data to measure the influence of a given production marker in the risk level, as well as to help evaluate the possible combinations of risk-engendering parameters. The Figure 6 shows the relation between the notes in respect to the Expiration Date and the raw materials dispersion, and the global criticality. This type of visual material can be used as a re-engineering tool in order to improve the piloting of these processes. For instance, it can be useful to detect the most sensitive criteria within a large set of parameters, as seen in the first diagram of Figure 6 where even if the raw materials dispersion presents the worse note, the criticality can still be low; but on the other hand, the second diagram shows that whenever the expiration date presents notes under 2,

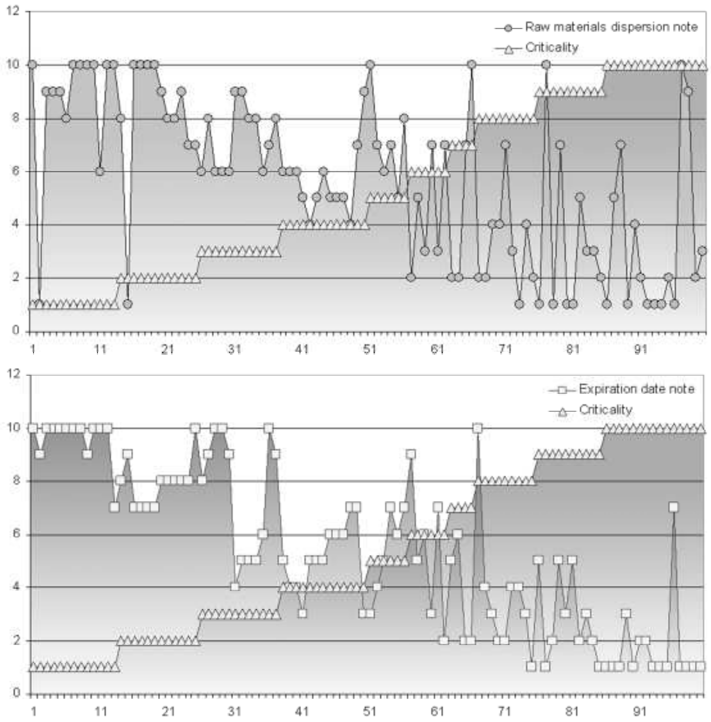

Figure 6, Influence of the notes in respect to the Expiration Date and the Raw materials dispersion in the global level of risk for a set of 100 examples

the criticality reaches its top levels. In conclusion, it is proved that risk is more sensitive to the expiration date than to the raw materials dispersion.

Using the same diagrams, the controllers may more easily detect whenever a highly sensitive parameter reaches a point of seriousness, in terms of risk. This would represent that all the others markers, as good as they may be, loose their influence and the risk result is critic. The boundaries of acceptance for each parameter as well as its degree of influence to the risk, can be seen and measured, in order to take preventive actions in production that may finally increase its quality, reduce the sanitary crisis probability and the recall size, in case of necessity.

6. Conclusions and perspectives

The proposed tool of real-time risk detection by using artificial neural networks tends to be more robust than any conventional data analyzing structure. It has the ability to cope with incomplete or fuzzy data. ANNs can be very tolerant of faults if properly implemented.

The most important feature we seek is to guarantee the reactivity in order to avoid important threats that represent the reality of delivering (and spreading) products containing possible hazards. This objective is achieved by the rapidity of the tool. Since the artificial neural network consists of a large number of massively interconnected processing units all operating in parallel on the same problem, the ANN can potentially operate at a considerable speed, rapidly adapting to the production's data flow.

Artificial neural networks are very flexible in adapting their behavior to new and changing environments. They are also easier to maintain, with some having the ability to learn from experience to improve their own performance.

As a difficulty, the artificial neural networks do not produce an explicit model of risk detection, even though different cases of production configuration can be fed into them and new results may be obtained as predictions.

The presented tool lacks an explanation capability. Justifications for the risk results may sometimes be difficult to obtain, because the weights connections usually do not have obvious interpretations.

7. References

Angeli C., and Smirni N., "An online expert system for fault diagnosis in hydraulic systems", Expert System, pp. 115–120, May 1999.

Bayır, R., & Bay, Ö. F., “Serial wound starter motor faults diagnosis using artificial neural network”. In IEEE ICM’04,.pp. 194–199. Istanbul, Turkey, 2004.

Borst, P., Akkermans, H. M., and Top, J., “Engineering Ontologies”, International Journal of Human Computer Studies, Vol. 46 (2–3), pp. 365–406. 1997.

Boresma P., “How we learn variation, optionality, and probability”, Institute of Phonetic Sciences, University of Amsterdam, Proceedings Issue 21, pp 43-58, 1998.

Brandriff K., “A Focus on Hazard Analysis and Critical Control Points”, United States Department of Agriculture, National Agricultural Library (NAL), Beltsville, USA, 2008.

Drollette Sarah A., “Managing Production Risk in Agriculture”, Department of Applied Economics,Utah State University, Logan, USA, 2009.

Dupuy, C., Botta-Genoulaz, V., and Guinet A., “Batch dispersion model to optimize traceability in food industry”. Journal of Food Engineering, Special Issue: Operational Research and Food Logistics, Vol. 70, Issue 3, pp. 333–339, 2005.

Enamul Haque A. K., “Sanitary and Phyto-Sanitary Barriers to Trade and its Impact on the Environment: The Case of Shrimp Farming in Bangladesh”, North-South University (NSU) Trade Knowledge Network, International Institute for Sustainable Development, Dhaka, Bangladesh, 2003.

Freund, Y. and Schapire, R. E. 1999. Large margin classification using the perceptron algorithm. In Machine Learning 37(3):277-296, 1999.

Ionica A., Edelhauser E., Irimie S., “Quality Risk Management – An Integrated Approach in the Mining Industry”, Fascicle of Management and Technological Engineering, Volume VI (XVI), Oradea, Rumania, 2007.

Rowcliffe, P., Jianfeng F., Buxton, H., “Spiking perceptrons”, Neural Networks, IEEE Transactions on Volume 17, Issue 3, May 2006

Karpenko M., and Sepehri N., “Neural network classifiers applied to condition monitoring of a pneumatic process valve actuator”, Engineering Applications of Artificial Intelligence 15 (3–4) (2002), pp. 273–283, 2002.

Kurtzweil P., “Critical Steps Toward Safer Seafood”, Food and Drug Administration (FDA), Center of Food Safety and Applied Nutrition, Washington, USA, 1999.

Kvenberg J. E., “Managing Food Safety: A HACCP Principles Guide for Operators of Food Establishments at the Retail Level”, Food and Drug Administration (FDA), Center of Food Safety and Applied Nutrition, Washington, USA, 1998.

Martin, T. and Howard, B., “Neural Networks for Control, School of Electrical & Computer Engineering”, Oklahoma State University/ University of Idaho, 1996.

McClelland, J. L., Rumelhart, D. E., “Explorations in parallel distributed processing: A handbook of models, programs, and exercises”. Cambridge, MA: MIT Press, 1988.

Najafi, H.L., Moses, D.W., Hustig, C.H., and Kinne, J., An intelligent dynamic reconstruction filter for audio signal reconstruction using neural networks, Engineering Applications of Artificial Intelligence, Vol. 11, Issue 1, pp. 49–53, 1998.

Pfister, M., Hybrid Learning Algorithms for Neural Networks, PhD Thesis, Free University Berlin, 1995.

Shin, F.Y., Moh J., and Bourne, H., “A neural architecture applied to the enhancement of noisy binary images”, Engineering Applications of Artificial Intelligence, Vol. 5, Issue 3, pp. 215–222, 1992.

Sjöberg J., “Neural Networks, Tain and analize neural networks to fit your data”, First edition, Wolfram Research Inc., Champaign, Illinois, USA, 2005.

Ramesh, B., Powers, T., Stubbs, C., and Edwards M., “Requirements traceability: Theory and practice”, Annals of Software Engineering, Vol. 3, Number 1, pp. 397-415, 1997.

Rojas, R., “Neural networks: a systematic introduction”, Springer-Verlag, Berlin, 1996.

Russell S., Norvig P., “Artificial Intelligence, A Modern Approach” (Second Edition), Prentice Hall, 2003.

U.S. Department of Health and Human Services, “Managing Food Safety: A Manual for the Voluntary Use of HACCP Principles for Operators of Food Service and Retail Establishments”, Food and Drug Administration Center for Food Safety and Applied Nutrition, College Park, USA, 2006.

Wang X., Li D., O’Brien C., “Optimization of traceability and operations planning: an integrated model for perishable food production”, International Journal of Production Research, Taylor & Francis, 2008.

Yu, D.L., Chang, T.K., and Yu, D.W., “Adaptive neural model-based fault tolerant control for multi-variable processes”, Engineering Applications of Artificial Intelligence, Vol. 18, Issue 4, June 2005, pp. 393–411, 2005.