

Taguchi Method and Artificial Neuro-Fuzzy Inference System (ANFIS) based Validation of Enzyme Production: A review

Shailendra Kumar Arya
Department of Biotechnology,
University Institute of Engineering and Technology,
Panjab University, Chandigarh, India
skarya_kr@yahoo.co.in

Abstract — Numbers of reports on enzyme production enhancements (from bacteria and fungi) are present in the literature by using One Variable at Time (OVAT) based optimization of medium components. OVAT strategy is not suitable for the cost-effective production of enzymes in lieu of modern statistical and artificially intelligent techniques like Response Surface Methodology (RSM), Taguchi Method and Artificial Neural Network (ANN) and Artificial Neuro-Fuzzy Inference System (ANFIS) etc. The Taguchi Method and ANFIS enzyme yield prediction results are in consonance with those produced by the RSM system and in fact are more closer to the actual enzyme yield. This shows the application of the proposed system in enzyme yield prediction given a set of parameter values.

Keywords- *Taguchi Method and Artificial Neuro-Fuzzy Inference System, Enzyme Production*

I. INTRODUCTION

Enzymes are biocatalysts and more expensive used due to production system, to avoid the cost system various method like validation process, solve the problems. For the example of laccase enzyme (benzenediol oxygen oxidoreductases, EC 1.10.3.2) are polyphenol oxidases that require O_2 to oxidize phenols, polyphenols, aromatic amines and different non-phenolic substrates by one electron transfer resulting in the formation of reactive radicals. They are members of the multicopper protein family that has developed from small sized prokaryotic azurins to eukaryotic ceruloplasmin. Enzymes are widely distributed in plants and fungi, where their involvement in melanin formation and in a variety of different, and sometimes contradictory, physiological functions like fungal morphogenesis, plant pathogenesis and fungal virulence has been frequently. In case of industrial applications, alkali and halo tolerant bacterial Enzymes are gaining wide interest day by day among the other sources of Enzymes like fungi and plants. Enzymes from prokaryotes are getting prominence because of their ability to act as the robust industrial biocatalysts in pulp and paper, textile and cosmetic industries where environmental conditions are highly variable moderate harsh. Due to the catalytic properties and environmental capabilities of prokaryotic Enzymes it is necessary to isolate the Enzyme producer bacteria and optimize the enzyme production. Cost-effective production of enzymes is ever a big challenge all over the world especially in the case laacase, where the size of pulp and paper industry is too large. There are several reports of Enzyme production enhancements available in the literature from prokaryotes and fungi, but still favourable results not forthcoming according to the pulp and paper industries requirements.

The most important stages in a biological process are modeling and optimization to improve a system and enhance the efficiency of the process without increasing the cost. Response surface methodology (RSM) is a widely practiced approach for the production and optimization of various

industrially important microbiological, biochemical and biotechnological products such as chemicals and enzymes. Based on the principal of design of experiments (DoE), the methodology encompasses the use of various types of experimental designs, generation of polynomial equations and mapping of the response over the experimental domain to determine the optimum product. [1] The technique requires minimum experimentation and time thus proving to be far more effective than the conventional methods of developing such products. The RSM may prove to be more productive and beneficial than the conventional technique by virtue of investigation of the effect of all the parameters simultaneously. suggested that response surface methodology (RSM) could be used as a valuable statistical method for the optimization of enzyme production from *Streptomyces psammoticus*; but using this strategy, only threefold enzyme productions was enhanced in comparison to conventional method of enzyme production. The past decade has seen a lot of data analysis tools based on biological phenomena develop into well-established modeling techniques, such as artificial intelligence and evolutionary computing. Taguchi Method and Artificial neural networks (ANNs) are now the most popular artificial learning tool in biotechnology, with a wide applications range included optimization of bioprocesses and enzyme production from microorganisms. Indeed an ANN is a massively interconnected network structure consisting of many simple processing elements capable of performing parallel computation for data processing. The fundamental processing element of Taguchi Method and ANNs (the artificial neuron) simulates the basic functions of biological neurons. In the present study, evaluation/validation of the responses generated from RSM based optimization of Enzyme production correspondences with that of ANFIS prediction [2] [3].

II. METHODOLOGY

The present study uses the Artificial Neuro-Fuzzy Inference System for prediction of the Enzyme production. An adaptive

neuro-fuzzy inference system or adaptive network-based fuzzy inference system (ANFIS) is a kind of artificial neural network that is based on Takagi–Sugeno fuzzy inference system. Since it integrates both neural networks and fuzzy logic principles, it has potential to capture the benefits of both in a single framework. ANFIS is considered to be a universal estimator. The Taguchi method and analysis of variance (ANOVA) have proven to be useful to investigate the effect of multiple parameters in complex material processing. In the Taguchi method, orthogonal arrays are used to efficiently determine the effect of variables and levels to achieve a robust design, which can greatly reduce experimental time and cost. Based on Taguchi experimental design, the optimized parameters can be determined by comparing the mean of the signal-to-noise (S/N) ratio. Furthermore, the significant parameters and percentage contribution of each parameter were determined by ANOVA. For the example, One factor at a time and Taguchi methods were used for optimization of culture condition. The amount of nitrogen, ratio of mineral salts to solid substrate and ratio of wheat bran to wheat straw, were determined by one factor at a time method. The effect of different proportions of substrates, growth temperature, initial pH, moisture level, inoculum's size, and nitrogen source, each in four levels, were optimized by Taguchi method. The results of experiments performed for obtaining optimum levels of enzyme by *Trichoderma longibrachiatum*, using easily available substrates, like wheat straw and wheat bran in SSF, are discussed. Taguchi method is not a usual way for optimizing biotechnological processes. Other methods such as response surface and Plackett–Burman may be preferred because researchers are more familiar with them. One of the most popular mathematic methods applied for the optimization of biological systems are the artificial neural networks (ANN). Researches comparing RSM to ANN indicate that ANN is the most accurate method of interpolation. It should be noted that in these works, predictions of ANN were compared with predictions of a quadratic model for RSM (restricted flexibility). In order to conduct a complete comparative study, the quadratic model is not automatically chosen. So, different polynomial models with a high degree, It is a two-axis grid that represents the intersection of 8 levels for each factor studied (exposure time and concentrations of salicylic acid). Each node of the 64 intersections corresponds to a carried out experiment. The results allowed building of a database for the network training and a test database for the model validation. Metaheuristic techniques, such as genetic algorithms (GA) and ant colony optimization (ACO) have been applied to solve complex problems within biological systems. For example, ACO was implemented in the medical sciences for solving problems in protein folding. Here ACO was used to predict the protein conformation based on the amino acids sequences. The used ACO represents a successful tool to solve the problems in bioinformatics, where other state-of-the-art methods failed. In field of bioinformatics, GA was used to detect recombination problems.

III. ANFIS ARCHITECTURE

It composes of five layers in this inference system. Each layer involves several nodes, which are described by the node function. The output signals from nodes in the previous layers

are accepted as the input signals in the present layer. After manipulation by the node function in the present layer (Figure 1), the output is served as input signals to the next layer [5].

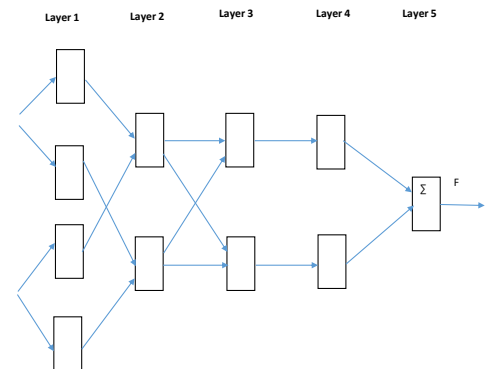


Figure 1: General ANFIS Architecture

$$\begin{aligned} \text{If } X \text{ is } A_1 \text{ and } Y \text{ is } B_1 \text{ then } F &= p_1X + q_1Y + r_1 \\ \text{If } X \text{ is } A_2 \text{ and } Y \text{ is } B_2 \text{ then } F &= p_2X + q_2Y + r_2 \end{aligned} \quad (1)$$

Layer 1: Layer 1 is an input and fuzzification layer. Every node i in this layer is an adaptive node with a node function **Eq. 3 and 4:**

$$\begin{aligned} O_{1i} &= \mu_{A_i}(X) \text{ for } i = 1 \text{ to } 2 \\ O_{1i} &= \mu_{B_i}(Y) \text{ for } i = 3 \text{ to } 4 \end{aligned}$$

Layer 2: Layer 2 is the rule layer. Each node in this layer computes the impact of each rule through multiplication **Eq. 5:**

$$O_{2i} = W_i = \mu_{A_i}(X) \cdot \mu_{B_i}(Y) \text{ for } i = 1 \text{ to } 2$$

Layer 3: Layer 3 is normalization layer. Each neuron in this layer computes the normalized effect of a given rule **Eq. 6:**

$$O_{3i} = \bar{W}_i = \frac{W_i}{W_1 + W_2} \text{ for } i = 1 \text{ to } 2$$

Layer 4: Parameters in this layer are considered as Rule's consequent parameters, **Eq. 7** gives the strength of each rule

$$O_{4i} = \bar{W}_i F_i(p_iX + q_iY + r_i) \text{ for } i = 1 \text{ to } 2$$

IV. MEDIA OPTIMIZATION BY TAGUCHI METHODOLOGY

Optimization methodology adopted in this study was divided into four phases (with various steps) viz., planning, conducting, analysis and validation Taguchi method of DOE (Design of experiment) involves establishment of large number of experimental situation described as OA (orthogonal array) to reduce experimental errors and to enhance their efficiency and reproducibility of the laboratory experiments. Each phase had separate objective, inter connected in sequence wise to achieve the overall optimization process [6]

Design of experiments (Phase 1)

The first step in Phase 1 is to determine the various factors to be optimized in the culture medium that have critical effect on the CGTase production. The normal practice is to experiment with the feasible range, so that the variation inherent in the process does not mark the factor effect. Factors were selected and the ranges were further assigned based on the group consensus consisting of design engineers, scientists and technicians with relevant experience data analysis procedure. The appropriate orthogonal array for the control parameters to fit a specific study was selected. Taguchi provides many standard orthogonal array and corresponding linear graphs for this purpose. In the present case, the two levels of factors variation were considered and the size of experimentation was represented by symbolic array L8 (which indicates 8 experimental trails). Five factors with two levels with a layout of L8 were used. The total degree of freedom is equal to number of trails minus one *i.e.*, 7. In the design OA, each column consists of a number of conditions depending on the levels assigned for each factors.. When the incubation period was completed, the cells were separated by centrifugation at 10,000 rpm for 15 minutes. The clear supernatant was used as crude enzyme for the further experiments [7].

Submerged fermentation experiments (Phase 2)

Submerged fermentation experiments were performed for enzyme production with *bacteria* employing selected 8 experimental trails in combination with 5 factors at two levels and the result obtained from each set[8] [9].

Analysis of experimental data and prediction of performance (Phase 3)

The obtained experimental data was processed in the Qualitek-4 software with bigger is better quality characteristics for the determination of the optimum culture conditions for the fermentation, to identify individual factors influence on the PHB production and to estimate the performance (fermentation) at the optimum conditions. In Taguchi's method, quality is measured by the deviation of a characteristics from its target value and a loss function $[L(y)]$ is developed for the deviation as represented by $L(y)=k \times (y-m)^2$, where k denotes the proportionality constant, m represents the target value and y is the experimental value obtained for each trail. In case of bigger is better quality characteristics, the loss function can be written as $L(y) = k \times (1/y^2)$ and the expected loss function can be represented [10].

$$E [L(y)] = k \times E \left(\frac{1}{y^2} \right) \quad (8)$$

Where $E (1/y^2)$ can be estimated from a sample of n as

$$\frac{\sum [1/Y_i^2]}{n}$$

Validation (Phase 4)

In order to validate the methodology, fermentation experiments were further performed for CGTase production using the obtained optimized culture conditions [11].

Software

Qualitek-4 software (Nutek Inc., MI) for automatic design of experiments using Taguchi approach was used in the present study. Qualitek-4 software is equipped to use L-4 to L-64 arrays along with selection of 2 to 63 factors with two, three and four levels to each factor. The automatic design option allows Qualitek-4 to select the array used and assign factors to the appropriate columns

V. CONCLUSION

The past decade has seen a host of data analysis tools based on biological phenomena develop into well-established modeling techniques, such as artificial intelligence and evolutionary computing. Taguchi Method and Artificial neural networks (ANNs) are now the most popular artificial learning tool in biotechnology, with a wide applications range included optimization of bioprocesses and enzyme production from microorganisms. Indeed an Taguchi Method and ANN is a massively interconnected network structure consisting of many simple processing elements capable of performing parallel computation for data processing. The fundamental processing element of Taguchi Method and ANNs (the artificial neuron) simulates the basic functions of biological neurons. In the present study, evaluation/validation of the responses generated from RSM based optimization of enzyme production correspondences with Taguchi Method and ANN.

REFERENCES

- [1]. Jang, Jyh-Shing R (1991). Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm (PDF). Proceedings of the 9th National Conference on Artificial Intelligence, Anaheim, CA, USA, July 14–19. 2. pp. 762–767.
- [2]. Jang, J.-S.R. (1993). "ANFIS: adaptive-network-based fuzzy inference system". IEEE Transactions on Systems, Man and Cybernetics. 23 (3). doi:10.1109/21.256541.
- [3]. Abraham, A. (2005), "Adaptation of Fuzzy Inference System Using Neural Learning", in Nedjah, Nadia; de Macedo Mourelle, Luiza, Fuzzy Systems Engineering: Theory and Practice, Studies in Fuzziness and Soft Computing, 181, Germany:

-
- [4]. Singh G, Ahuja N, Capalash N, Sharma P (2009c) Design of experiment methodology for the optimized production of alkalophilic Enzyme from γ -proteobacterium. JB. Bioresources 2:544–553
 - [5]. Singh G, kaur K, Puri S, Sharma P (2015) Critical factors affecting Enzyme-mediated biobleaching of pulp in paper industry. Appl Microbiol Biotechnol 99:155–164.
 - [6]. Bains J, Capalash N, Sharma P (2003) Enzyme from a nonmelanogenic, alkalotolerant γ -proteobacterium JB isolated from industrial wastewater drained soil. Biotechnol. Lett. 25: 1155–1159
 - [7]. . Ibrahim, H.M. (2005). Optimization of medium for the production of β -cyclodextrin glucanotransferase using Central Composite Design (CCD). Process Biochem., 40.,753 -758.
 - [8]. Jin-yong,Yan. (2008). Optimization for producing cellbound lipase from Geotrichum sp. And synthesis of methyl oleate in microaqueous solvent. Appl. Microbiol. Biotechnol, 78, 431-439.
 - [9]. Kumar M. et al., Artificial Neuro-Fuzzy Inference System (ANFIS) based validation of laccase production using RSM model Biocatalysis and Agricultural Biotechnology 14 (2018) 235–240
 - [10]. Leemhuis, H., Ronan, M. K. and Lubbert D.(2010). Engineering of cyclodextrin glucanotransferases and the impact for biotechnological applications. Appl MicrobBiotechnol 85, 823–835.
 - [11]. Li Z, Wang, M, Wang, F., Gu Z, Du G, Wu J and Chen, J (2007). γ -Cyclodextrin: a review on enzymatic production and applications. Appl Microbiol Biotechnol 77, 245–255.