Process Biochemistry 51 (2016) 422–430

Contents lists available at ScienceDirect

Process Biochemistry

journal homepage: www.elsevier.com/locate/procbio

Optimization of artificial neural network by genetic algorithm for describing viral production from uniform design data

Maria Beatriz Takahashi*, José Celso Rocha, Eutimio Gustavo Fernández Núñez

Departamento de Ciências Biológicas, Universidade Estadual "Júlio de Mesquita Filho" Campus-Assis, Avenida Dom Antonio 2100, Assis, 19806-900 SP, Brazil

ARTICLE INFO

Article history:
Received 12 August 2015
Received in revised form 6 December 2015
Accepted 11 December 2015
Available online 17 December 2015

Keywords:
Artificial neural network
Bioprocess
Genetic algorithm
Uniform design
Virus production

ABSTRACT

This work objective was to define a modeling approach based on genetic algorithm (GA) for optimizing parameters of an artificial neural network (ANN); the latter describes rabies virus production in BHK-21 cells based on empirical data derived from uniform designs (UDs) with different numbers of experimental runs. The parameters considered for viral infection were temperature (34 and 37 °C), multiplicity of infection (0.04, 0.07, and 0.1), infection, and harvest times (24, 48, and 72 h), with virus production as the monitored output variable. A multilevel factorial experimental design was performed and used to train, validate, and test the ANN. Its experimental fractions (18, 24, 30, 36, and 42 runs) defined by UDs were used to simulate the neural architectures. In GA, the neural computing parameters constituted the population individuals, and the steps involved were population creation, selection, and replacement by crossover and mutation. The ANN optimized by the combined algorithm showed a good calibration for all UDs under consideration, thus demonstrating to be suitable ($R > 0.85$) as a correlation method in UDs independent of the experimental runs developed. Therefore, this work could guide researchers in the efficient use of UDs in the simulation and optimization of virus production processes.

1. Introduction

Animal cell cultures have been used for the production of rabies vaccines since the 1960s due to the neurological deleterious effects caused by first-generation vaccine based on the application of animal nervous tissue [1]. According to World Health Organization (WHO), these vaccines contain an inactivated virus and can be developed in various cell substrates such as human diploid cells, primary chicken embryo cells, Vero cells, or in primary Syrian hamster kidney cells. Recently, the establishment of serum/animal protein free culture media in animal cell cultures at a large scale has improved the safety of rabies viral vaccines and their applications showed efficacy and decrease in production costs [2,3].

In 2004, the Food and Drug Administration (FDA) launched the Quality by design (QbD) and Process analytical technology (PAT) to understand, control, and guarantee quality of pharmaceutical products and the corresponding processes [4]. Since then, the monitoring and design of biopharmaceutical process have improved. The QbD approach is based on four cornerstones: risk-assessment techniques; PAT tools for in-process monitoring and multivariate controlling; mathematical and statistical tools including design of experiments and process data modelling over the course of technological development stages; and continuous improvement tools [5]. Experimental designs in pharmaceutical industry have improved processes and product performance by selecting the best values for critical parameters in experimental space according to the number of experiments and using the model selected to obtain a satisfactory level of prediction accuracy for its associated response or output (process and product quality attributes) [6].

An experimental approach that can be used for this purpose is uniform experimental design (UD), which is a statistical method involving a significant reduction of the number of experimental runs, mainly applied to experimental regions that have several factors and multiple levels [7,8]. However, a disadvantage is data processing derived from UD, which is made by special computer software, and statistical knowledge is necessary due to the absence of preestablished mathematical models [9]. UD is commonly applied for solving industrial problems and widely used in different fields of engineering, including pharmaceutical, chemical, and biological processes [9–11].

UD also can be combined with other mathematics methods such as artificial neural network (ANN) and genetic algorithm (GA) in order to model and optimize output variables, which define process yield and productivity as well as product quality attributes.

* Corresponding author.
E-mail address: mbeatriz.takahashi@gmail.com (M.B. Takahashi).

http://dx.doi.org/10.1016/j.procbio.2015.12.005
1359-5113/

ANN, based on human neural cell processing, has been used in bioprocess as it is suitable for describing nonlinear correlations. This method can usually describe the biological processes better than linear models as well as their intrinsic complexity and high level of uncertainty [12,13]. ANN is a computing system comprising highly interconnected process elements (nodes) known as neurons that map a set of input patterns onto a corresponding set of output patterns. The application of ANN in bioprocessing and chemical engineering has increased significantly since the end of the 1980s. Classification, data filtering, process modeling, optimization, and control are among the most important applications of ANN. ANN computing comprises three main steps, data preparation and selection, topological design of ANN, and selection of the training methods along with synaptic weights and biases [14].

The optimization of ANN parameters can be improved using GA, a stochastic global search algorithm, achieving better performances [15,16]. GAs simulate the survival-of-the-fittest principle of nature while searching for an objective function that helps in its natural optimization. GAs use string coding of variables for dividing the search space into discrete ones, which further helps in searching for global optimum in the entire solution space. GAs are used for defining ANN topology, selection of training methods, and weights and biases. In addition they have been widely used for searching an optimal bioprocess critical parameter in order to maximize product quality and process efficiency indexes [17,18], which is sometimes referred to as GA-optimized neural network system (GONNS) [19].

The main purpose of this work was to develop a combined ANN–GA algorithm for determining and optimizing an ANN model, which describes rabies virus production in BHK-21 cells based on empirical data, derived from uniform designs with different numbers of experimental runs.

2. Material and methods

2.1. Cell line and culture medium

BHK-21 (C-13) cells (Sigma–Aldrich ECACC Cell Lines, Lyon, France) suitable for a single-cell suspension culture were kindly supplied by Ecole Superieure Biotechnologie de Strasbourg (France) through Dr. Renaud Wagner. The culture medium was composed of (in volume) 45.5% Iscove's modified Dulbecco's medium with glutamine and phenol red, 45.5% high-glucose Dulbecco's modified Eagle's medium, 5% heat-inactivated fetal bovine serum, 2% (10% m/v) Pluronic F-68 aqueous solution, and 2% 4 mM glutamine aqueous solution [20].

2.2. Cell culture

About 1 mL of BHK-21 cells of concentration 2×10^6 cell/mL was defrosted and placed in a 75-cm^2 tissue culture flask vertically positioned with 30 mL of culture medium for growth. After 4 days, the cell suspension of 0.12×10^6 cell/mL concentration was inoculated consecutively in other tissue culture flasks of 25 and 75 cm^2 along with the addition of an adequate amount of culture medium in order to reach the final cellular suspensions of 10 and 30 mL, respectively. For conducting viral infection assays, cell populations were grown in 100-mL spinner flasks (Bellco Glass Inc., Vineland, NJ, USA), with a working volume of 50 mL; the stirring speed was maintained at 30 rpm (Sci-Era quad drive stirrer system with a stirrer, Bellco Biotechnology, Vineland, NJ, USA). The spinner flasks with an initial cell concentration of 0.25×10^6 cell/mL were previously seeded with cells harvested in 75-cm^2 tissue culture flask in the exponential phase. All the procedures described in this section were performed in an incubator (Thermoforma 3110, Marietta, OH, USA) at 37 °C and 5% of carbon dioxide atmosphere. Cell samples from spinner flask at 24, 48, and 72 h post inoculation were withdrawn for viral infection experiments; the cell concentrations at each moment were 0.75×10^6, 2×10^6, and 4×10^6 cell/mL, respectively. The cell viability was 100% at the three sampling moments under consideration [21].

2.3. Virus strain

Rabies virus strain (CVS-11, ATCC VR 959) conserved at −80 °C was provided by Laboratório Especial de Pesquisa e Desenvolvimento em Imunológicos Veterinários (Instituto Butantan, Brazil). Five virus strain passages in BHK-21 cell line were performed, and the corresponding viral titration value was 3.9 [$\log_{10}$ (fluorescence focus doses 50% per 50 μL)].

2.4. Rabies virus titration and production quantification

Preparation of viral titration samples was performed by clarification of the virus suspension using a Millex-HV 0.45-μm PVDF (polyvinylidene fluoride) membrane (Millipore Corporation, Bedford, USA), and the clarified liquid was kept at −80 °C until viral titration.

A modified and validated methodology, originally defined in the previous work, was used for rabies virus titration [22,23]. Briefly, 50 μL clarified virus suspension samples were serially diluted and seeded in a 96-well microplate(dilution factor: 10) and 100 μL of the cell suspension comprising 3.7×10^4 adherent BHK-21 cells in a proper culture medium was added. Subsequently, the microplate was incubated at 37 °C and 5% of carbon dioxide atmosphere for 22 h. Then, the culture medium was discarded, and each well was washed with phosphate buffer saline (0.13 M NaCl, 2 mM KCl, 8 mM Na_2HPO_4, and 15 mM KH_2PO_4). Next, 80% v/v acetone aqueous solution at −20 °C was added. Consequently, the microplates were dried at room temperature; the cells were labeled with fluorescein-conjugated rabies antiserum (Bio-Rad, Marnes la Coquette, France—cat. # 72114) for an hour at 37 °C, and washed thrice with phosphate buffer saline. Accordingly, 70 μL buffered glycerin (pH 8.5) was added in each well. The microplates were read on an inverted fluorescence microscope (100× magnification). Each well was divided into 10 fields, and the fields were considered to be positive if at least one fluorescent focus was observed. The number of positive fields with respect to the total field for each sample dilution was recorded. The viral titration value was calculated as the common logarithm of the reciprocal value corresponding to the highest dilution able to show fluorescent focus in half of the observed fields [$\log_{10}$ (fluorescence focus doses 50% per 50 μL)]. The software used for viral titration calculations was CombiStats v.4.0 (EDQM—Council of Europe).

The rabies virus production (VP) was calculated as follows (Eqs. (1) and (2)):

$$\mathrm{VP} = \log_{10}[10^{T} - 10^{T_i}], \quad \text{if } T > T_i \tag{1}$$

$$\mathrm{VP} = -\log_{10}[10^{T_i} - 10^{T}], \quad \text{if } T < T_i \tag{2}$$

where T and T_i are the viral titration values at the defined and initial time, respectively.

2.5. Experimental designs

A multilevel factorial design was performed in order to model and optimize rabies virus production. The variables and levels considered were temperature (34 and 37 °C), time of infection (24, 48, and 72 h), multiplicity of infection (0.04, 0.07, and 0.1), and virus harvest time (24, 48, and 72 h). All experimental combinations (54) were performed in triplicate. For each experimental combination, a constant harvest time was set, and the control