Process Biochemistry 51 (2016) 1338–1347

Contents lists available at ScienceDirect

Process Biochemistry

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

# Artificial intelligence approach based on near-infrared spectral data for monitoring of solid-state fermentation

Augusto César Barchi[a], Shuri Ito[a], Bruna Escaramboni[b], Pedro de Oliva Neto[b], Rondinelli Donizetti Herculano[c], Matheus Carlos Romeiro Miranda[c], Felipe José Passalia[d], José Celso Rocha[d], Eutimio Gustavo Fernández Núñez[a,*]

[a] Grupo de Engenharia de Bioprocessos, Departamento de Ciências Biológicas, Universidade Estadual Paulista 'Júlio de Mesquita Filho' Campus-Assis, Avenida Dom Antônio, 2100, 19806-900 Assis, SP, Brazil
[b] Laboratório de Biotecnologia Industrial, Departamento de Ciências Biológicas, Universidade Estadual Paulista 'Júlio de Mesquita Filho' Campus-Assis, Avenida Dom Antônio, 2100, 19806-900 Assis, SP, Brazil
[c] Instituo de Química—Araraquara, Universidade Estadual Paulista 'Júlio de Mesquita Filho' Campus-Araraquara, Rua Professor Francisco Degni, 55, 14800-900 Araraquara, SP, Brazil
[d] Laboratório de Matemática Aplicada, Departamento de Ciências Biológicas, Universidade Estadual Paulista 'Júlio de Mesquita Filho' Campus-Assis, Avenida Dom Antônio, 2100, 19806-900 Assis, SP, Brazil

ARTICLE INFO

*Article history:*
Received 25 March 2016
Received in revised form 23 June 2016
Accepted 15 July 2016
Available online 18 July 2016

*Keywords:*
Artificial neural network
Bioprocess monitoring
Chemometrics
Enzymes
NIR spectroscopy
Partial least squares

ABSTRACT

This work aimed to establish a chemometric technique for quantifying amylase and protease activities as well as protein concentration in aqueous extracts of *Rhizopus microsporus* var. *oligosporus* obtained via solid-state fermentation (SSF). The kinetics of four agro-industrial wastes (wheat bran, soybean meal, type II wheat flour and sugarcane bagasse) were studied for 144 h, along with two different sets of their ternary mixtures, at a constant fermentation time of 120 h, to obtain primary data (biochemical parameters as well as near-infrared (NIR) spectral data). Then, models such as artificial neural network (ANN) and partial least squares (PLS) were calibrated to predict biochemical parameters using the spectral data. Primary data and three methods of preprocessing data – first, second and third derivatives – were assessed as inputs for both chemometric tools. The third derivative, that is, spectral pre-processing plus an optimized ANN, showed the least relative errors (<8.3% ± 10.5%). The third-derivative spectrum was found to be suitable as the ANN input data for monitoring amylase and protease activities and protein concentration in the SSF under study. The proposed methodology can serve as a foundation for at-line sensor development and decrease the time and cost of bioprocess development using *Rhizopus microsporus* var. *oligosporus*.


## 1. Introduction

Solid-state fermentation (SSF) is the fermentation of solid substrates with low water content, but sufficient for the metabolism and growth of microorganisms. While the water content in classical submerged fermentation (SmF) is >95%, that in SSF varies between 40% and 80%. SSF often utilizes agricultural and food wastes as substrates [1,2]. SSF has been suggested as an alternative to SmF, as the former has several advantages such as a lower implementation cost and a decreased chance of contamination. Moreover, the associated products are generally more tolerant to pH and temperature variations, and their separation is not complex, as a rule [1,3,4]. Recently, the SSF has become more accepted by industrial and academic communities due to its wide range of applications, for instance, producing secondary metabolites and extracellular enzymes used in food, pharmaceutical and biofuel industries (i.e., amylases and proteases). Furthermore, with the growing demand for 'green processes' as alternatives of chemical processes in the manufacturing industry, SSF has garnered great interest [5].

However, the major limitations of solid-state bioprocesses include the control of operation parameters as well as measurements of growth and metabolic parameters, such as cell concentration, substrate consumption and production of biomolecules. These limitations are attributed to the heterogeneous nature of the substrate, which has a highly complex structure and nutrient composition [1,6]. Furthermore, unlike SmF, the SSF medium presents several moisture and temperature gradients, which can have a negative impact on the process efficiency [7].

* Corresponding author.
*E-mail address:* eutimiocu@yahoo.com (E.G. Fernández Núñez).

http://dx.doi.org/10.1016/j.procbio.2016.07.017
1359-5113/

On the other hand, spectroscopy techniques can be used to monitor several parameters in bioprocesses in three different ways: (1) off-line methods, which are based on analysis of samples transferred to the laboratory, away from the bioreactor; (2) at-line measurements, which also involve manual or automatic sampling, but the analyses are conducted in the vicinity of the bioreactor; and (3) online methods, which use sensors in various ways to provide real-time analyses. Spectroscopy techniques such as ultraviolet–visible (UV–vis), near-infrared (NIR), mid-infrared (MIR), Raman and fluorescence spectroscopy, with fibre optic cables, have been evaluated for their efficiency in bioprocess online monitoring (in situ) [8].

In particular, NIR spectroscopy is classified as a molecular spectroscopic method with a spectrum region at 700–2500 nm [9]. The vibrations associated with the X–H bond type, such as O-H, N-H, C-H and S-H, show a high dipole momentum, which is widely found in biological molecules and thus forms the underlying principle of NIR spectroscopy [10]. Currently, NIR spectroscopy is used in the pharmaceutical, agriculture, chemical and food industries for quality control from the raw material stage to the final products [11].

Chemometrics, a useful tool for bioprocess monitoring, utilizes mathematical and statistical methods to analyse data from chemical systems. This tool can be used to model and extract information from a large amount of data obtained monitoring with spectroscopy techniques [12]. For instance, in fermentative methods, with substantial overall data (input spectral data and related output biochemical parameters), techniques such as PLS and partial component regression (PCR) can be used to analyse a complex mixture to determine the quantities of minor compounds [13]. Artificial neural networks (ANNs) have emerged as an alternative to these multivariate techniques. This artificial intelligence approach is a modelling methodology based on the mechanisms of the human brain. This technology uses simple processing units, known as neurons, arranged in input, hidden and output layers. By learning functions, the ANN can improve the model results by refining the parameters based on the obtained errors [14,15]. ANNs can be used successfully for bioprocess monitoring as they are flexible in the use of either linear or non-linear functions (or a combination of both) [16,17].

The present work aimed to correlate NIR spectral data from the SSF samples with the total protein concentration as well as amylolytic and proteolytic activities using ANN and PLS models. This establishes a foundation for developing sensors for SSF monitoring. It also reduces the time and cost of kinetic studies of the biotransformation of agro-industrial wastes by *Rhizopus microsporus* var. *oligosporus*, as well as quantifying the yield over the course of the scale-up stages.

## 2. Materials and methods

### 2.1. Microorganism strain and substrates

The microorganism used in the present work, *R. microsporus* var. *oligosporus* (CCT 3762), was obtained from the Tropical Foundation of Research and Technology 'André Tosello', Campinas, SP, Brazil. The microorganism population was expanded in 3.9% (m/v) potato dextrose agar (PDA) medium using 250-mL Erlenmeyer flasks. The spore solution, for inoculation of SSF experiments, was obtained after 7 days of incubation (30 °C), and the biomass was washed with 0.01% (v/v) Tween 80 in an aqueous solution under gentle magnetic stirring. The SSFs were carried out using agro-industrial wastes as substrates. The four selected substrates were wheat bran (Moinho Nacional, Assis, SP, Brazil), soybean meal (Landtech Comércio e Indústria Ltda, Paraguaçu Paulista, SP, Brazil), type II wheat flour (Moinho Nacional, Assis, SP, Brazil) and sugarcane bagasse (Água Bonita, Tarumã, SP, Brazil). They were used individually or as a component of ternary mixtures.

**Table 1**
Experimental matrix corresponding to D-optimal mixture design for studying the influence of ternary mixtures comprising wheat bran (A), type II wheat flour (B) and sugarcane bagasse (C) on spectral data modification in extracts and their associated amylolytic and proteolytic activities as well as total protein concentration by means of *Rhizopus microsporus* var. *oligosporus* in solid-state fermentation.

| Run | A [%] | B [%] | C [%] |
|---|---|---|---|
| 1 | 75.00 | 25.00 | 0.00 |
| 2 | 60.00 | 33.33 | 6.67 |
| 3 | 30.00 | 50.00 | 20.00 |
| 4 | 80.00 | 0.00 | 20.00 |
| 5 | 50.00 | 50.00 | 0.00 |
| 6 | 70.00 | 16.67 | 13.33 |
| 7 | 47.50 | 37.50 | 15.00 |
| 8 | 82.50 | 12.50 | 5.00 |
| 9 | 55.00 | 25.00 | 20.00 |
| 10 | 30.00 | 50.00 | 20.00 |
| 11 | 50.00 | 50.00 | 0.00 |
| 12 | 100.00 | 0.00 | 0.00 |
| 13 | 80.00 | 0.00 | 20.00 |
| 14 | 90.00 | 0.00 | 10.00 |

**Table 2**
Experimental matrix corresponding to D-optimal mixture design for studying the influence of ternary mixtures comprising type II wheat flour (B), sugarcane bagasse (C) and soybean meal (D) on spectral data modification in extracts and their associated amylolytic and proteolytic activities as well as total protein concentration by means of *Rhizopus microsporus* var. *oligosporus* in solid-state fermentation.

| Run | B [%] | C [%] | D [%] |
|---|---|---|---|
| 1 | 48.33 | 13.33 | 38.33 |
| 2 | 50.00 | 15.00 | 35.00 |
| 3 | 40.00 | 20.00 | 40.00 |
| 4 | 48.33 | 18.33 | 33.33 |
| 5 | 50.00 | 15.00 | 35.00 |
| 6 | 50.00 | 10.00 | 40.00 |
| 7 | 45.83 | 18.33 | 35.83 |
| 8 | 50.00 | 20.00 | 30.00 |
| 9 | 43.33 | 18.33 | 38.33 |
| 10 | 45.00 | 20.00 | 35.00 |
| 11 | 50.00 | 10.00 | 40.00 |
| 12 | 40.00 | 20.00 | 40.00 |
| 13 | 46.67 | 16.67 | 36.67 |
| 14 | 45.00 | 15.00 | 40.00 |
| 15 | 50.00 | 20.00 | 30.00 |

### 2.2. Fermentation process

The SSFs were performed in 250-mL Erlenmeyer flasks in two different ways: (1) the substrate was composed of 10 g of a single agro-industrial waste, or (2) the substrate contained the same mass of a ternary mixture with the composition in line with runs included in two D-optimal mixture designs (Design-Expert 6.0.1 version, Stat-Ease, Inc., Minneapolis, MN, USA). The ternary mixtures are listed in Tables 1 and 2. In the first case, the fermentations were conducted as kinetic studies. They were conducted in triplicate and monitored at 24-h intervals for 144 h. In the second case, the duration was 120 h. These two different approaches generated a wide range of data. In both cases, the substrate was humidified using an aqueous micronutrient solution composed of 1.25% (m/m) ammonium sulphate, 0.25% (m/m) potassium phosphate and 1.25% (m/m) urea, so as to obtain 55% (m/m) moisture [18]. After substrate sterilization (121 °C, 1 atm, 30 min), the inoculation was aseptically performed with $1 \times 10^6$ spores (concentration of spore suspension: $2.27 \times 10^7$ spores/mL) of *R. oligosporus* per gram of the fresh substrate. Subsequently, the Erlenmeyer flasks were incubated at 30 °C (Tecnal TE-421, Piracicaba, Brazil).