International Journal of Food Microbiology 170 (2014) 78–82

Contents lists available at ScienceDirect

International Journal of Food Microbiology

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

Error analysis in predictive modelling demonstrated on mould data

József Baranyi [a], Olívia Csernus [b,*], Judit Beczner [b]

[a] Institute of Food Research, Norwich Research Park, Norwich NR4 7UA, UK
[b] Central Environmental and Food Science Research Institute, Unit of Microbiology, H-1022 Budapest, Herman Ottó út 15, Hungary

ARTICLE INFO

Article history:
Received 27 February 2013
Received in revised form 12 August 2013
Accepted 26 October 2013
Available online 4 November 2013

Keywords:
Predictive microbiology
Mould
Aspergillus
Food safety

ABSTRACT

The purpose of this paper was to develop a predictive model for the effect of temperature and water activity on the growth rate of *Aspergillus niger* and to determine the sources of the error when the model is used for prediction. Parallel mould growth curves, derived from the same spore batch, were generated and fitted to determine their growth rate. The variances of replicate ln(growth-rate) estimates were used to quantify the experimental variability, inherent to the method of determining the growth rate. The environmental variability was quantified by the variance of the respective means of replicates. The idea is analogous to the "within group" and "between groups" variability concepts of ANOVA procedures.

A (secondary) model, with temperature and water activity as explanatory variables, was fitted to the natural logarithm of the growth rates determined by the primary model. The model error and the experimental and environmental errors were ranked according to their contribution to the total error of prediction.

Our method can readily be applied to analysing the error structure of predictive models of bacterial growth models, too.

1. Introduction

In order to improve the quality and safety of food, there is a need for mathematical models to predict microbial responses to food environments (Gibson and Hocking, 1997; Dantigny et al., 2005). In this study we focus on fungal growth, using *Aspergillus* species.

The black Aspergilli (*Aspergillus* section *Nigri*) are well studied in food mycology. Many species of this group cause food spoilage, while some of them are useful in the fermentation industry. *Aspergillus niger*, under certain industrial conditions, has been granted the GRAS (generally regarded as safe) status by the Food and Drug Administration of the US government (Pitt and Hocking, 1997). However, since the first description of ochratoxin A (OTA) production by *A. niger* (Abarca et al., 1994), members of *Aspergillus* section *Nigri* have achieved greater attention because of their potential to contaminate diverse food commodities with OTA. Ochratoxin A has nephrotoxic, teratogenic, genotoxic, immunosuppressive and carcinogenic properties presenting a high risk both for humans and animals (Klich, 2002). Economic losses can be considerable, especially in the case of the post-harvest decay of fresh fruits (Pitt and Hocking, 1997).

A. niger is more prevalent in warmer climates, both in field situations and stored foods. The minimum growth temperature is 6–8 °C, the maximum values are 45–47 °C, and the optimum is between 35 and 37 °C (Pitt and Hocking, 1997). Due to the expected increasing average temperature and decreasing frequency and amount of rain falls, temperate climate zones will also provide better conditions for its growth and toxin production (Farkas and Beczner, 2009; Csernus et al., 2011; Varga et al., 2012).

In this study, predictive modelling methods developed for bacteria are applied to fungal kinetics. Our basic assumption is that the colony growth rate, like the specific growth rate in bacterial modelling, is a species-dependent constant, characteristic to the growth environment. Our objectives are to

(1) quantify the effect of experimental and environmental variability;
(2) model the effect of environmental conditions (in our case, temperature and water activity) on the growth rate of *A. niger*;
(3) rank the model accuracy and the experimental and environmental variability according to their contribution to the total error between prediction and independent observation.

It should be emphasized that by *total error* we mean a measure of (dis)agreement between prediction and a single observation of growth rate, where the observation and its quantification were made by the same method that generated the dataset, on which the model was based. This error is inherent to the method and does not decrease when more observations are used in the model building; unlike for example the standard error of the estimates of the model parameters. Our study is carried out by methods and concepts that can readily be transferred to bacterial growth modelling, too.

* Corresponding author. Tel.: +36 1 356 4673.
E-mail addresses: jozsef.baranyi@ifr.ac.uk (J. Baranyi), o.csernus@cfri.hu (O. Csernus), j.beczner@cfri.hu (J. Beczner).

0168-1605/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijfoodmicro.2013.10.018

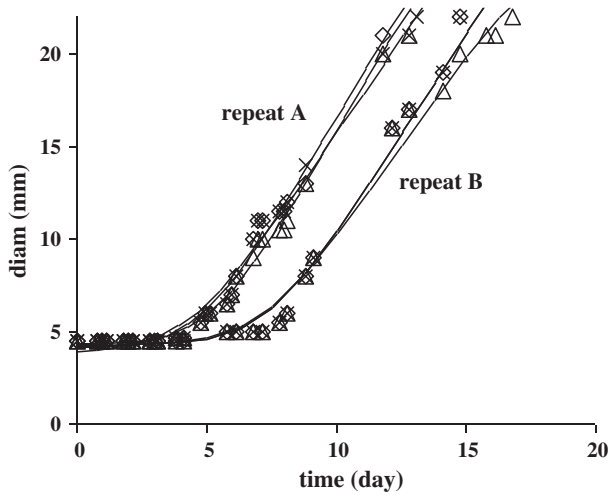

Fig. 1. Two repeats of three parallels for *A. niger*, under the same values of environmental factors (temperature 25 °C, a_w 0.90). The observed mould colony diameters (symbols), as a function of time, were fitted by the model of Baranyi and Roberts (1994; continuous line). Three parallel replicates give the repeat A; the other three constitute the repeat B.

2. Materials and methods

2.1. Fungal strain

A. niger F.00770 was obtained from the Hungarian National Collection of Agricultural and Industrial Microorganisms, Budapest, Hungary.

2.2. Medium

The study was carried out in vitro using Malt Extract Agar (MEA, 2% malt extract, Merck; 2% glucose, Reanal; 1% peptone, Merck; 2% agar, Merck). The medium was autoclaved and poured into 9 cm sterile Petri dishes. The pH of the medium was 6.5. The water activity was modified by the addition of known amounts of NaCl to 0.99, 0.98, 0.96, 0.94, 0.92 and 0.90. Water activity values were determined with a Novasina LabMaster.a_w Instrument (Vitalis, Hungary) which was calibrated against saturated salt solutions in a range of known water activities.

2.3. Inoculation, incubation conditions and growth measurements

A. niger was grown on MEA for 7 days at 25 °C to obtain heavily sporulating cultures. Fungal conidia (spores) suspensions were prepared in a solution of Tween-80 (100 μl/l) and were adjusted to 10^6 spores/ml using a haemocytometer.

The 4.5 mm central holes in the medium were inoculated with 50 μl spore suspension. After inoculation, the Petri plates were sealed in polyethylene bags to prevent water loss and stored under controlled storage conditions in programmable incubators set at 20, 25, 30 and 35 °C, respectively. The water activity of each plate was measured at the end of the experiment, too, and no deviations from the original a_w values were found. The diameters (mm) of the colonies were measured at appropriate time intervals, without opening the dishes, using a ruler, three times a day up to maximum 36 days. A growth experiment was carried out in three replicates; i.e. in one experiment, the temporal variation of three colonies, each on separate plate, was measured in parallel, in the same incubator. Therefore we can safely assume identical environmental conditions for these parallels. The experiment, using a new spore batch, was repeated 2 weeks later, under as much similar conditions as possible. Hence, altogether two *repeats* (A and B) of triplet growth curves (*replicates*) were generated for each combination of temperature and water activity (Fig. 1). It is evident that the variability between the replicates must be smaller than between the repeats, the latter one being affected by possible differences in the batches and in the environmental conditions, however truly the experimenter tries to produce identical repeats.

2.4. Primary model

Models describing a certain microbial response to a given environment as a function of time are termed as primary models. Our "colony diameter *v*. time" data were fitted by the model of Baranyi and Roberts (1994), using an in-house Excel Add-in package "DMFit" (www.combase.cc/index.php/en/downloads/category/11-dmfit) in order to determine the growth rate. Due to its construction, the model is able to describe growth curves either with or without stationary phase and with or without lag. This is controlled by two curvature parameters, n and m (see the manual of DMFit).

To reduce the effect of change in the micro-environment, the primary model was fitted only to points where the colony diameter was less than 30 mm. This cut-off value was chosen empirically, since the vast majority of the growth curves exhibit a very robust linearity under 30 mm. Accordingly, the m curvature parameter of the growth model was set to zero (i.e. no upper asymptote was fitted by DMFit). For the other curvature parameter, its default value, $n = 1$ was used.

2.5. Secondary model

Following the method of Gibson et al. (1994), the natural logarithm values of the g growth rates, estimated by the primary model, were regressed against b_w which is a rescaled version of the water activity, a_w:

$$b_w = \sqrt{(1-a_w)}. \quad (1)$$

The benefit of the above empirical transformation, as reported by Gibson et al. (1994), is that the "ln(g) v. b_w" data are closer to being symmetric than their "ln(g) v. a_w" representations are, the latter exhibiting high non-linearity when close to the optimum. The gained symmetric convex pattern of the data lends itself to a parabolic fitting that can be achieved by linear regression, with all its numerical and statistical advantages regarding the robustness of the estimates.

Gibson et al. (1994) found the variance of ln(g) constant for varying temperature and water activity (at least in regions supporting robust growth). Our experience was the same. Another reason why we use the natural logarithm as a link function is that the results can easily be translated into relative accuracy measures, as we will see later (multiplicative error analysis).

The experimental design in terms of the tested environmental factors followed a full factorial design, i.e. growth curve triplets were produced under all the $4 * 6 = 24$ combinations of temperatures 20, 25, 30 and 35 °C and a_w 0.99, 0.98, 0.96, 0.94, 0.92 and 0.90, in both A and B experiments. This way, $3 * 2 * 24 = 144$ growth curves were fitted by the primary model.

2.6. Error structure

The experimental design made it possible to analyse whether the experimental or environmental variability affects the prediction error more. The rationale behind this is as follows:

- When colonies grow in parallel, i.e. on plates next to each other, the spores are from the same spore batch, the environmental effect on the plates is minimized so the variability of the responses is mainly due to the experimental method how the growth rate is determined. This is why we refer to this as "experimental variability". It is quantified by the w^2 variance of the ln(g) estimates of parallel replicates. In ANOVA terminology, w^2 plays the role of the *within group* variance.