Abstract

This paper considers the development of data-based mathematical structures capable of predicting the biomass and antibiotic concentrations of an industrial fermentation process. The models are developed using both a standard SYMBolic Annealing (SYMBA) strategy, and a hybridised form of SYMBA. Results reveal that the empirical model structures identified using the standard form of SYMBA produce much more accurate predictions than its hybrid counterpart. However, the models discovered using the hybrid variant are more interpretable in that they allow the researcher to identify important growth characteristics of the fermentation process.

1 INTRODUCTION

Previous research has concentrated on artificial neural networks and genetic programming to tackle the problem of modelling fermentations. However, the former reveals little information about the process due to its ‘black-box’ nature, while the latter tends to be very CPU intensive. In an attempt to capture the positive aspects of both of these approaches, we propose a novel technique for predicting the biomass and antibiotic concentrations of a fermentation process, which is based on Simulated Annealing (SA), Aarts et al. (1989).

2 RESEARCH DETAILS

During this work, a modification of an SA algorithm is utilised in order to perform the optimisation of structure i.e. SYMBA, Porter et al. (1998). The structures identified by this algorithm are either directly evaluated to generate concentration estimates, or incorporated into a pair of differential equations in order to predict the growth rates and hence concentrations of the two variables of interest.

Our hybrid version of the algorithm namely Sequential Hybridised Symbolic Annealing (SHSA) allows two structures to be developed in series (one for each variable being predicted). In sequential mode, the interaction between each structure evolution process is prevented, as opposed to the algorithm used in Porter et al. (1999). Experience has shown that operating in series allows solutions to be discovered that produce more orthodox growth profiles.

To identify each algorithm’s potential to discover models with varying degrees of step-ahead predicting capabilities, experiments were performed whereby predictions were generated at a frequency equal to the sampling rate (every 7-8hrs) and at an hourly rate.

3 CONCLUSIONS

Results from this study have shown that the predictive accuracy of the models identified using standard SYMBA are far superior to those using the hybrid variant. In fact, it has been found that is it sufficient to re-evaluate the models predicting at a rate equal to the online samples in order to generate accurate predictions per hour. However, for SHSA it has been found that the algorithm must be operated with the models directly predicting per hour for sufficiently accurate predictions to be discovered. This results in much higher runtimes being observed. However, in both cases the CPU requirements were still quite acceptable.

Bibliography

