Innovative Food Science and Emerging Technologies 25 (2014) 67–77

Contents lists available at ScienceDirect

Innovative Food Science and Emerging Technologies

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

Food model exploration through evolutionary optimisation coupled with visualisation: Application to the prediction of a milk gel structure

Evelyne Lutton [a,*], Alberto Tonda [a], Sébastien Gaucel [a], Alain Riaublanc [b], Nathalie Perrot [a]

[a] UMR 782 Génie et Microbiologie des Procédés Alimentaires, AgroParisTech, INRA, 78850 Thiverval-Grignon, France
[b] INRA, Rue de la Géraudière, BP 71627, 44316 Nantes Cedex 3, France

ARTICLE INFO

Article history:
Received 22 June 2013
Accepted 1 February 2014
Available online 11 February 2014

Keywords:
CMA-ES
Evolutionary algorithms
Milk gel model
NSGA-II
Optimisation
Visualisation

ABSTRACT

Obtaining reliable in-silico food models is fundamental for a better understanding of these systems. The complex phenomena involved in these real-world processes reflect in the intricate structure of models, so that thoroughly exploring their behaviour and, for example, finding meaningful correlations between variables, become a relevant challenge for the experts. In this paper, we present a methodology based on visualisation and evolutionary computation to assist experts during model exploration. The proposed approach is tested on an established model of milk gel structures, and we show how experts are eventually able to find a correlation between two parameters, previously considered independent. Reverse-engineering the final outcome, the emergence of such a pattern is proved by physical laws underlying the oil–water interface colonisation. It is interesting to notice that, while the present work is focused on milk gel modelling, the proposed methodology can be straightforwardly generalised to other complex physical phenomena.

Industrial relevance: Sustainability is nowadays at the heart of industrial requirements. The development of mathematical approaches should facilitate common approaches to risk/benefit assessment and nutritional quality in food research and industry. These models will enhance knowledge on process–structure–property relationships from the molecular to macroscopic level, and facilitate the creation of in-silico simulators with functional and nutritional properties. The stochastic optimisation techniques (evolutionary algorithms) employed in these works allow the users to thoroughly explore the systems: when coupled with visualisation, they make it possible to provide the experts with a restricted set of significant data, helping them to highlight eventual issues or possible improvements in the model. With regard to the complexity of the food systems and dynamics, the challenge of the mathematical approaches is to realise a complete dynamic description of food processing. In order to reach this objective, it is mandatory to use innovative strategies, exploiting the most recent advances in cognitive and complex system sciences.

Glossary

Name	Description	Unit
m_p	Total mass of proteins in the solution (constant)	g
m_{wp}	Mass of native whey proteins in the solution	g
m_{cas}	Mass of casein micelles in the solution	g
S_0	Initial lipid surface	m^2
S_{fall}	Lipid surface available for adsorption of both native whey proteins and casein micelles	m^2
S_{fres}	Lipid surface left by casein micelles due to steric effects for native whey proteins	m^2
k_{wp}	Adsorption rate of native whey proteins	s^{-1}
k_{cas}	Adsorption rate of casein micelles	s^{-1}
s_{wp}	Surface area occupied by 1 g of native whey proteins	$m^2 \cdot g^{-1}$
s_{cas}	Surface area occupied by 1 g of casein micelles	$m^2 \cdot g^{-1}$
α	Fraction of the adsorbed surface of a casein micelle reserved for native whey proteins	Dimensionless
$w_{wp}(0)$	Initial mass percentage of native whey proteins in the solution, $w_{wp}(0) = m_{wp}(0)/m_p(0)$	%
$w_{cas}(0)$	Initial mass percentage of casein micelles in the solution, $w_{cas}(0) = m_{cas}(0)/m_p(0)$	%
$w_{wp_{int}}$	Final mass percentage of native whey proteins at lipid interface relative to the total mass of adsorbed proteins	%
$w_{cas_{int}}$	Final mass percentage of casein micelles at lipid interface relative to the total mass of adsorbed proteins	%
Γ	Final interfacial concentration which corresponds to the quantity of adsorbed proteins per 1 m^2 of lipid surface	$mg \cdot m^{-2}$
$d_{3,2}$	Average diameter of lipid droplet	m
ρ_l	Lipid density	$g \cdot m^{-3}$
m_l	Mass of lipid (constant)	g
μ	Population size parameter for the evolutionary algorithms used in the experience	Dimensionless

(continued on next page)

* Corresponding author.
E-mail addresses: evelyne.lutton@grignon.inra.fr (E. Lutton), alberto.tonda@grignon.inra.fr (A. Tonda), sebastien.gaucel@grignon.inra.fr (S. Gaucel), alain.riaublanc@nantes.inra.fr (A. Riaublanc), nathalie.perrot@grignon.inra.fr (N. Perrot).

1466-8564/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ifset.2014.02.003

(*continued*)

Name	Description	Unit
λ	Offspring size parameter for the evolutionary algorithms used in the experience	Dimensionless
$\eta_{operator}$	Distribution index for a genetic operator in the evolutionary algorithm NSGA-II	Dimensionless

1. Introduction

Building in-silico models for food processes is an important but difficult task, as there exist various known bottlenecks (Perrot, Trelea, Baudrit, Trystram, & Bourgine, 2011). The process of model design, for instance, often relies on computationally expensive optimisations to match a theoretical model with available data (parameter learning). Scarcity of data is a classical source of troubles for the optimisation process, resulting in badly conditioned problems. Solutions provided by optimisation cannot be exploited directly and must be revisited by experts, in order to disambiguate equivalent sets of solutions. Facilitating a high level expert analysis of computational results, or even more, interaction of expert knowledge with computational processes, is a challenging task: interactive optimisation is an active field of research (Takagi, 1998), and its potential applications in the domain of food process modelling are numerous.

Optimisation tools are actually often used in a “black box” manner, and computational optimal results may then yield imprecise, ambiguous or even wrong parameter setting. In this paper, we present a methodology based on evolutionary algorithms (also know as “genetic algorithms”). Their iterative, population-based, algorithmic structure, if appropriately exploited, allows the highlighting of various features of the search space, which correspond to possible pathologies of the considered model. Experts may of course have access to these pathological features via appropriate theoretical analysis, as soon as they know what to search for. As we will see below, the observation of the successive population distributions of the evolutionary algorithms allows as to get some intuitions about the possible degeneracies of the searched model, thus making the task of the expert easier. We exemplify this approach on a complex test case, the prediction of the structure of a milk gel.

Evolutionary algorithms (EAs) are the generic name for a large set of techniques that rely on the computer simulation of natural evolution mechanisms (Artificial Darwinism). Since pioneering works in the second half of 20th century (Bremermann, 1962; Fraser, 1957; Holland, 1962; Rechenberg, 1973), Artificial Darwinism techniques have progressively gained importance in stochastic optimisation and artificial intelligence domains for the resolution of difficult optimisation problems, and particularly for learning the optimal parameters of complex models (Bäck & Schwefel, 1993).

The main idea of EAs is to copy, in a very rough manner, the principles of natural evolution, that let a population be adapted to his environment. According to Darwin's theory (Darwin, 1859), adaptation is based on very simple mechanisms: random variations, inheritance, and survival/reproduction of the fittest individuals. Transposed into optimisation algorithms, this scheme has the major advantage of making few assumptions on the function to be optimised (there is no need to have a continuous or derivable function for instance). In short, evolutionary algorithms consider a population of potential solutions exactly as a population of individuals of a natural population that live, fight and reproduce. The environment pressure is replaced by an “optimisation” pressure: the function to be optimised is considered as a measurement of the adaptation of the individual to its environment (fitness). In this way, individuals that reproduce are the best ones with respect to the problem to be solved, and reproduction consists in generating new solutions via genetic operators (called crossover and mutation by analogy to nature).

Evolutionary optimisation techniques are particularly well suited to difficult problems, where classical methods fail. The major reason of this success is the tunable combination of oriented and random search mechanisms that allow injecting a priori, incomplete informations in the genetic operators, while letting some other more unpredictable components be randomly searched.

Considering evolutionary optimisation as a “black box”, however, is not a good strategy in general. The first reason is that one may lose the opportunity to adapt the mechanisms to the specifics of the problem, which usually improves the efficiency of the algorithms and reduces its computation time. Another reason is related to the internal mechanisms of the algorithm that performs a sampling of the search

Table 1
Milk gel data used for training (database 1, L_1 to L_7) and validation (database 2, V_1 to V_4).

Sample	w_{cas} (0) (%)	$d_{3,2}$ (μm)	m_p (0) (g)	$w_{cas_{int}}$ (%)	Γ (mg·m^{-2})	Database
L_1	13	0.6	2.47	5	5.6	1
L_2	19	0.7	2.44	9	4.8	1
L_3	21	0.6	2.42	16	4.0	1
L_4	26	0.65	2.40	43	4.9	1
L_5	32	0.55	2.40	65	5.6	1
L_6	49	0.56	2.39	71	4.2	1
L_7	80	0.9	2.37	84	9.3	1
V_1	13	0.59	8.79	0	4.66	2
V_2	22	0.74	8.47	33	4.4	2
V_3	31	0.87	8.64	46	6.88	2
V_4	80	0.75	9.18	91	6.93	2

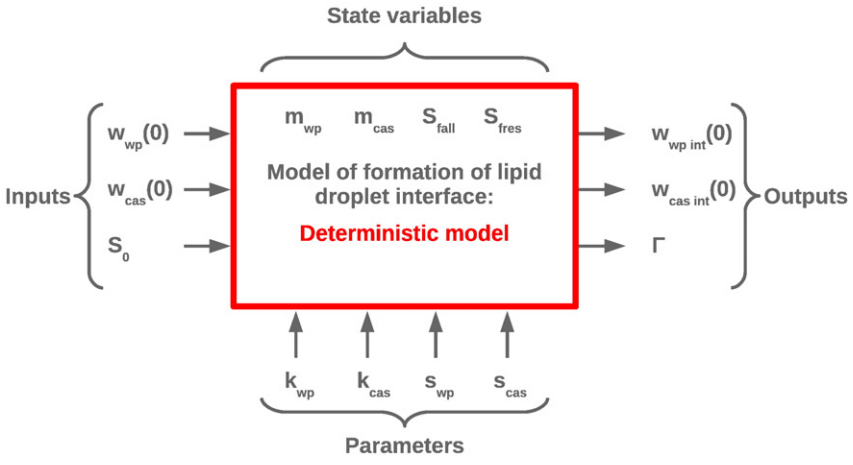

Fig. 1. Model of milk gel formation.