On the application of a nature-inspired stochastic evolutionary algorithm to constrained multi-objective beer fermentation optimisation

Alistair Rodman [1], Eric S. Fraga [2], Dimitrios Gerogiorgis [1]*

[1] Institute for Materials and Processes (IMP), School of Engineering, University of Edinburgh, Edinburgh, UK

[2] Centre for Process Systems Engineering (CPSE), Department of Chemical Engineering,

University College London (UCL), London, UK

*Corresponding author: D.Gerogiorgis@ed.ac.uk (+44 131 651 7072)

Abstract

Fermentation is an essential step in beer brewing, often acting as the system bottleneck due to the time consuming nature of the process stage (duration >120hrs), where a trade-off exists between attainable ethanol concentration and required batch time. To explore this trade-off we employ a multi-objective Plant Propagation algorithm (the Strawberry algorithm), for identifying temperature manipulations for improved fermentation performance. The methodology employed successfully produces families of favourable temperature profiles which exist along the Pareto front. A subset of these output profiles can simultaneously reduce batch time and increase product ethanol concentration while satisfying constraints on by-products produced in the fermenters, representing significant improvements in comparison with current industrial practice. A potential batch time reduction of over 12 hours has been highlighted, coupled with a moderate improvement in ethanol content.

Keywords: Dynamic optimisation; nature-inspired optimisation; multi-objective optimisation; stochastic optimisation; solution representation; beer fermentation

1. Introduction

The modelling and optimisation of the beer fermentation process has received considerable interest in recent years, due to the requirement for brewers to improve process efficacy under current market conditions. The production of beer is well established, but it is also identified that current practice for fermenter temperature control can frequently be quite far from optimal (Rodman and Gerogiorgis, 2016a-b). Numerous authors have performed optimisation studies for fermenter control. Stochastic approaches including genetic algorithms (Carrillo-Ureta et al., 2001) and ant colony systems (Xiao et al., 2003) have been applied. Additionally, Bosse and Griewank (2014) have used a forward-backward sweeping method, while Rodman and Gerogiorgis (2017) used orthogonal collocation on finite elements to obtain exact solution profiles for optimal performance, investigating the effect of by-product constraints on attainable performance. These prior studies have implemented weighted sum objective functions to handle the multi-criteria problem. The weights assigned to the various process targets to produce a single objective function may be considered arbitrary in many cases, with decision-makers (brewers) not necessarily able to quantify a priori the relative importance of competing objectives. A number of multi-objective optimisation algorithms have been successfully applied to a wide range of engineering problems, where visualisation of the trade-offs can provide decision makers with valuable insight (Li et al., 2014; Gujarathi et al., 2015; Zhang et al., 2015; Fraga and Amusat, 2016; Che at al., 2017; Maria and Crişan, 2017; Kessler et al., 2017). Systematically exploring the trade-off and visualising Pareto optimal temperature manipulations for efficient fermentation is desirable to gain insight and assist brewers with the selection of the most preferable operation strategy. It could be suggested to vary the weights and thus obtain a Pareto set using a weighted-sum approach, however equally spaced weights are unlikely to map to eqi-spaced points on the Pareto set, and unsupported points will be omitted, even if they are Pareto optimal.

We aim to develop a solution representation (control profile to decision vector encoding) which permits efficient search of feasible candidate temperature manipulations for effective fermentation. Such a representation would allow the bi-criteria trade-off to be mapped using multi-objective optimisation methods such as NSGA-II (Deb et al., 2002), or for exploration with single-objective methods, such as the ε-constrained method, which converts constrained optimization problems to unconstrained ones using the ε-level comparison, which compares search points based on the pair of objective function values and the constraint violation between them (Miettinen, 1999).

The Strawberry algorithm (Salhi and Fraga, 2011) is a nature-inspired stochastic evolutionary optimisation methozd which has been successfully applied to a single-objective dynamic optimisation problem in the built environment (Fraga et. al., 2015). Recently, using a new fitness function for multi-objective problems, the algorithm has been applied to integrated energy systems design for off-grid mining operations (Fraga and Amusat, 2016). Given the demonstrated success of the algorithm for gaining insight into a bi-criteria objective trade-off, we propose using the Strawberry Algorithm evaluate our solution representations and compute Pareto optimal solution sets for the industrial beer fermentation process.

We note that the particular novelty of the present study is the solution representation we employ for constrained multi-objective optimisation, enabling the computation of industrially feasible solutions for a multitude of initialisation profiles and algorithm settings. Because it is algorithm-independent, it can be readily implemented in other (e.g. NSGA-II) stochastic optimisation methods.

1.1 Beer fermentation

Fermentation is an essential step in the manufacture of alcoholic beverages, responsible for the characteristic taste of the final product in addition to its alcohol content. Upstream processing produces a sugar rich intermediate (wort) from a feedstock starch source (most typically malted barley). Once cooled to an appropriate initial temperature, the wort enters stainless steel vessels along with yeast, allowing fermentation to commence. The primary chemical reaction pathway is the conversion of sugars into ethanol and carbon dioxide, which is coupled with biomass (yeast) growth and heat generation from the exothermic reaction. Concurrently, a range of species are formed at low concentrations by a multitude of side reactions, many of which may negatively impact product flavour above threshold concentrations. Fermentation is completed once all consumable sugars have been converted by the yeast into alcohol, following which the solution leaves the fermenter for subsequent downstream processing prior to sale and consumption. Given the long duration of the process stage, often upwards of 120 hrs, the fermentation process often acts as the production bottleneck within the overall manufacturing process (Fig. 1).
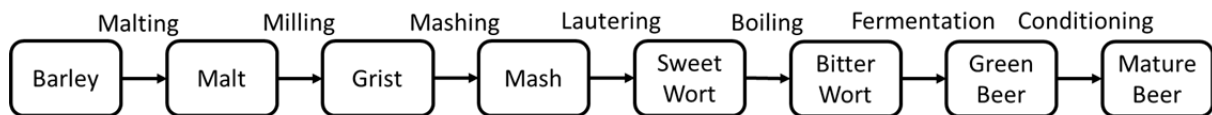
Figure 1. Block flow diagram of the beer production process.

1.2 Fermentation modelling

Several mathematical models for the beer fermentation process have been proposed (Gee and Ramirez, 1988; de Andres-Toro, 1998, Trelea et al., 2001). Models are reduced order, considering only the key species present due to system complexity (200+ species, Vanderhaegen et al., 2006) rendering exhaustive modelling extremely cumbersome: in fact to date many of the specific chemical interactions in the fermentation process are not understood

1.2.1 Model selection.

The kinetic model of beer fermentation by de Andrés-Toro et al. (1998) has been selected for study due to its direct applicability to the industrial process:

- Published parameters are derived from a large array of experiments, resulting in a wide temperature range (8–24 ℃) which ensures high fidelity and applicability.

- The model includes key by-products which degrade beer product quality in terms of taste and aroma, ensuring the model is suitable for assessing the performance.

- Predicted profiles indicate the highest fidelity with experimental and pilot-plant data in comparison to other models, due to successful validation.

The model considers seven state variables (Eqs. 1-7), with trajectories governed by temperature dependent production and consumption factors (Eqs. 8-12). The seven state variables represent different key chemical species as defined in the nomenclature list. The model structure takes the form shown in Fig. 2. Yeast cells transition from latent to active to dead over time, with only active cells able to promote fermentation (conversion of sugar to ethanol). Two by-products are considered alongside the primary reaction pathway: ethyl acetate (Eq. 7) and diacetyl compounds (Eq. 6). Diacetyl (2,3-butanedione) has a pungent butter-like aroma (Izquierdo-Ferrero et al., 1997), while ethyl acetate is often used as an indicator of all esters present, and is described as having the odour of nail varnish remover (Hanke, S. et al., 2010). A more detailed description of the model can be found

in its original publication (de Andrés-Toro, 1998), along with the constants for the Arrhenius relationship governing the parameters' temperature dependence, as computed from industrial scale fermentation data.

$$\frac{d[X_A](t)}{dt} = \mu_x(t,T) \cdot [X_A](t) - \mu_{DT}(t,T) \cdot [X_A](t) + \mu_L(t,T) \cdot [X_L](t) \tag{1}$$

$$\frac{d[X_D](t)}{dt} = -\mu_{SD}(t,T) \cdot [X_D](t) + \mu_{DT}(t,T) \cdot [X_A](t) \tag{2}$$

$$\frac{d[S](t)}{dt} = -\mu_S(t,T) \cdot [X_A](t) \tag{3}$$

$$\frac{d[EtOH](t)}{dt} = f(t) \cdot \mu_e(t,T) \cdot [X_A](t) \tag{4}$$

$$\frac{d[EA](t)}{dt} = Y_{EA}(T) \cdot \mu_x(t,T) \cdot [X_A](t) \tag{5}$$

$$\frac{d[DY](t)}{dt} = \mu_{DY} \cdot [S](t) \cdot [X_A](t) - \mu_{AB} \cdot [DY](t) \cdot [EtOH](t) \tag{6}$$

$$\frac{d[X_L](t)}{dt} = -\mu_L(t,T) \cdot [X_L](t) \tag{7}$$

$$\mu_x(t,T) = \frac{\mu_{x_0}(T) \cdot [S](t)}{0.5 \cdot [S]_0 + [EtOH](t)} \tag{8}$$

$$\mu_{SD}(t,T) = \frac{\mu_{SD_0}(T) \cdot 0.5 \cdot [S]_0}{0.5 \cdot [S]_0 + [EtOH](t)} \tag{9}$$

$$\mu_s(t,T) = \frac{\mu_{s_0}(T) \cdot [S](t)}{k_s(T) + [S](t)} \tag{10}$$

$$\mu_e(t,T) = \frac{\mu_{e_0}(T) \cdot [S](t)}{k_e(T) + [S](t)} \tag{11}$$

$$f(t) = 1 - \frac{[EtOH](t)}{0.5 \cdot [S]_0} \tag{12}$$
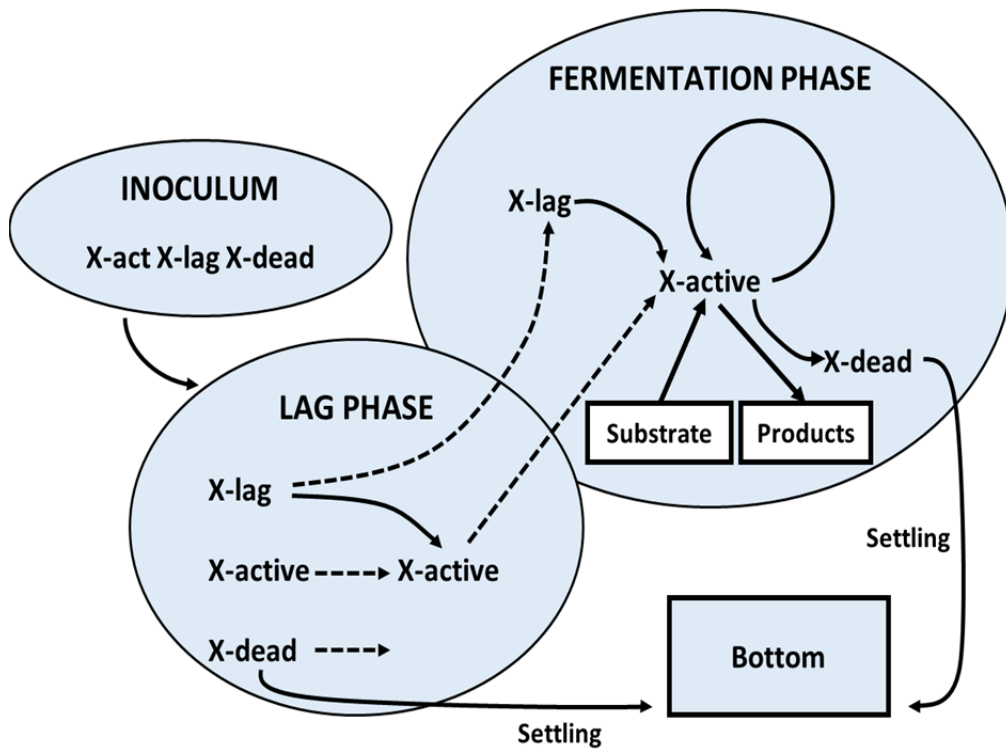

Figure 2. Kinetic model for beer fermentation (de Andres-Toro, 1998).

3

1.3 Process targets and constraints

When considering what it is desirable to improve in a fermentation process there are two obvious contenders: reduced duration and heightened alcohol content (even if this requires later dilution, it is still desirable to increase yield). In addition to batch time minimisation and alcohol production maximisation, prior authors of optimisation studies have elected to include terms for minimisation of both by-products within their respective objective functions (Carrillo-Ureta et al., 2001; Xiao et al., 2003; Bosse and Griewank 2014). However, as is known that, within certain beer products, the concentrations of both ethyl acetate and diacetyl compounds have a negligible effect on flavour below certain levels, efforts towards further reduction and concentration minimisation are redundant.

As such it is deemed more appropriate to consider only time minimisation and ethanol maximisation as objectives, treating the final concentrations of both ethyl acetate and diacetyl compounds as strict constraints to avoid unnecessary efforts towards further by-product reduction.

$$\min (f_1, f_2) \tag{13}$$

$$\text{where} \quad f_1 = 1/[EtOH]_{t_f}$$

$$f_2 = t_f$$

We define batch time, t_f, as the time when the sugar content of the word has been reduced to 0.5% of its original value (Eq. 14): at this point, all fermentable material has been effectively consumed with residual sugars in the product at typical levels.

$$t_f = t \left([S]_t = 0.005 \cdot [S]_{t0}\right) \tag{14}$$

Constraint thresholds for a traditional Lager product are given by Eqs. 15-16 (Rodman and Gerogiorgis, 2016), restricting the product concentration of ethyl acetate (EA) and diacetyl compounds (DY) below levels which would degrade flavour.

$$[EA]_{tf} \leq 2 \quad \text{ppm} \tag{15}$$

$$[DY]_{tf} \leq 0.1 \text{ ppm} \tag{16}$$

Given the strong dependence of yeast health on system temperature it is necessary to include an additional constraint such that the control profile (temperature) remains within acceptable levels. The constraint (Eq. 17) is introduced to ensure that the lower temperature limit excludes scenarios in which the system lacks enough energy to promote cell growth while the upper limit ensures bacteria which are present above this temperature cannot thrive, while also preventing the temperature from reaching a level at which undesirably high by-product concentrations are known to be produced.

$$9\,°C \leq T(t) \leq 16\,°C \text{ for all } t \in [t_0, t_f] \tag{17}$$

Recently a deterministic optimisation strategy was applied to the same industrial fermentation problem (Rodman and Gerogiorgis, 2017), however considering a single weighted-sum objective function, and not two unique objectives. Orthogonal collocation on finite elements has been used in that study, in order to discretise both the state variable and the control (temperature) variable trajectories, demonstrating the ability to formulate manipulations with highly desirable performance. It is however apparent that output profiles are local solutions, as the output profiles differed when the initialising profile was modified. As such, it is of interest to explore the use of a stochastic global multi-objective optimisation method for the same problem to investigate if these local maxima may be escaped to ultimately produce temperature profiles with improved fermentation performance and identify the trade-offs between the two conflicting objectives.

2. Stochastic optimisation with a Plant Propagation Algorithm

2.1 The Strawberry algorithm

The Strawberry algorithm (Salhi and Fraga, 2011) is a nature inspired stochastic evolutionary optimisation method. It emulates the behaviour of strawberry plants, encapsulating the two key aspects of effective global optimisation algorithms: solution exploration and intensification. In nature, strawberry plants exploit their surroundings through the propagation of runners. In a favourable environment, they will generate a greater number of runners, most within their local vicinity. Less frequently, the plants which are not as well situated will reproduce through the propagation of fewer yet longer runners. This inspires the Strawberry algorithm: each member of the population (an individual solution) is evaluated (objective functions computed) and a fitness function is assigned. The fitness value influences both the number of runners (exploitation, proportional to fitness) and the distance which each runner travels (exploration, inversely proportional to fitness). The population evolves over a pre-defined number of generations. The evolutionary process is characterised by only two parameters: the maximum number of runners to generate for any given solution and the number of solutions to consider for propagation in each generation.

The Strawberry algorithm has previously been successfully applied to a single objective dynamic optimisation problem in the built environment (Fraga et. al., 2015). Recently, using a new fitness function for multi-objective problems, the algorithm has been applied to integrated energy systems design for off-grid mining operations (Fraga and Amusat, 2016), also a dynamic optimisation problem. The algorithm can be summarised as follows:

Algorithm 1: The strawberry plant propagation algorithm, adapted from Fraga and Amusat, 2016.

Given: $f(x)$, a vector function; n_g, number of generations to perform, n_p, the propagation size; n_r, maximum number of runners to propagate.
Output: z, vector approximation to Pareto front.

    $p \leftarrow$ initial random population of size n_p
    for n_g generations **do**
        prune population p, removing similar solutions
        N $\leftarrow$ fitness(p) $\triangleright$Use rank based fitness
        $\tilde{p} \leftarrow \varphi$ $\triangleright$ Empty set
        for $i \leftarrow 1 \ldots n_p$ **do**
            $x \leftarrow$ select(p, N) $\triangleright$Tournament fitness based selection
            for each runner to generate **do** $\triangleright$Number proportional to fitness rounded up
                $\tilde{x} \leftarrow$ new solution(x, 1 $-$ N) $\triangleright$Distance inversely proportional to fitness
                $\tilde{p} \leftarrow \tilde{x} \cup \tilde{p}$ $\triangleright$ Add to new population
            end for
        end for
        $p \leftarrow \tilde{p} \cup$ Nondominated(p) $\triangleright$ New population with elitism
    end for
$z \leftarrow$ Nondominated(p)

2.2 Application to control vector optimisation

In order for control profile formulation to be compatible with such an optimisation strategy, it is necessary to define a set of decision variables which directly translate to a specific profile or trajectory. The representation of solutions has a direct impact on the effectiveness of any optimization method (Fraga et al., 2017). Specifically, the bounds imposed on each variable should restrict the solution form such that only realistically practical cases are considered (omitting unnecessary computational load) while ensuring that a large enough search space exists so that a wide range of solutions may be considered. In this study we consider two strategies for profile encoding, a piecewise linear approach and a piecewise polynomial approach.

2.2.1 PieceWise Linear profiles (PWL)

Here we define a profile as consisting of N piecewise linear segments, between $N+1$ nodes $[t_i, T_i]$ ($0 \leq i \leq N$). The time domain is first discretised into N equal intervals:

$$\Delta t = t_{max}/N \tag{18}$$

where t_{max} is the longest fermentations desirable to be considered, defined using existing industrial practice as a basis:

$$t_{max} = 120 \text{ hr} \tag{19}$$

If a uniform spacing were used, t_i would be $i \cdot \Delta t$. However, more effective use of the $N+1$ points may be possible if nonuniform spacing were allowed. We introduce optimization variables $t_{dev,i}$ to represent a deviation from the uniform spacing for each point but the first. Then:

$$t_i = i \cdot \Delta t + t_{dev,i} \tag{20}$$

This deviation approach is preferable to permitting completely free movement of the nodes within the time domain. With suitable bounds on the deviations, this approach acts to regulate the segment lengths: preventing a large portion of the segments bunching together which would require unachievably and undesired rapid manipulations to be performed to the vessel. Additionally it ensures that unfavorably long fermentations are not considered: only fermentation times up to $t_{max} + t_{dev,max}$ can be computed.

Bounds imposed on time point deviation prevent subsequent nodes, t_i, from overlapping with the prior, t_{i-1}.

$$- \Delta t /2 \leq t_{dev,i} \leq \Delta t /2 \tag{21}$$

The bounds on the initial temperature and the temperature component of each profile node are given by:

$$T_{min} \leq T_i \leq T_{max} \tag{22}$$

with the values for the bounds taken form Eq. 17.

In summary, the decision variables consist of the initial system temperature, $T(t = 0) = T_0$, and N couplets of $[t_{dev,i}, T_i]$:

$$d = [T_0, t_{dev,1}, T_1, \ldots, t_{dev,N}, T_N] \tag{23}$$

such that the number of decision variables scales linearly with the discretisation level:

$$\text{length } (d) = 2N+1 \tag{24}$$

The search domain which may be explored by the stochastic algorithm, D, is fully defined by Eqs. 21-23, with the manner in which the decision vector translates to a piecewise linear $T(t)$ control profile between $N+1$ nodes $[t_i, T_i]$ ($0 \le i \le N$) defined by Eqs. 18-20.

As an illustrative example of this profile encoding, consider the vector d = [14 -10 16 0 9 0 10]. The 7 elements correspond to $N = 3$, from Eq. 24 meaning the profile described by 4 $T(t)$ points or 3 linear sections. Three equal time intervals would split the 120 hr horizon into 40 hr sections (Δt = 40), however the second point in d states a minus 10 hour deviation on the second time coordinate (the first is always $t = 0$ so no deviation applies). This example therefor corresponds to a piecewise linear profile between the four points: [(0,14), (30,16), (80, 9), (120, 10)], as shown in Fig. 3:
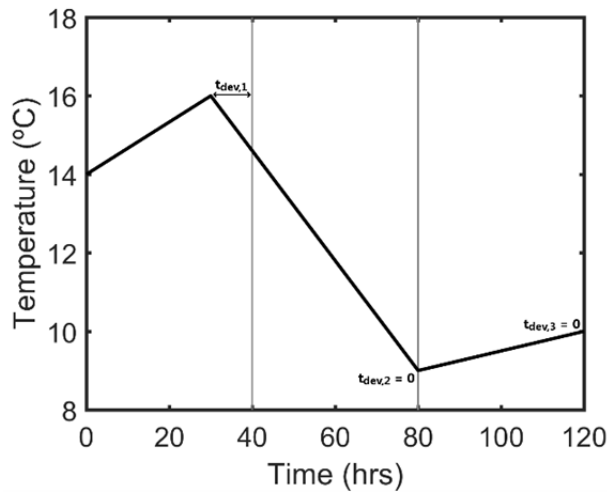
Figure 3. Example profile encoding for the PWL strategy.

2.2.2 PieceWise Polynomial profiles (PWP)

To explore the effectiveness of smoother profiles which are not dependant on discretisation level we an alternative, piecewise polynomial profile, representation. It has been demonstrated that problems of slow convergence and non-smooth impractical control strategies can be overcome by representing control profiles with polynomial approximations (Sorek et al., 2017). The authors achieved significant computational savings, due to a substantial reduction in the number of control parameters by seeking the optimal polynomial coefficients, rather than directly searching for the optimal control values within specific time intervals.

Our implementation considers the time domain to consist of three subdomains. In each sub-domain, the solution representation will define a smooth polynomial. Specifically, we construct $T(t)$ profiles from three polynomials: a cubic polynomial for the first and last sections, and a quintic polynomial for the intermediate section. Temperature values and first derivatives of the temperature are defined to be the same for the respective polynomials at each boundary between sub-domains, and the gradients at the very start and the end of the profile are set to 0. With these restrictions, the profile is described uniquely by 5 points and 2 derivatives, shown in Fig. 4.
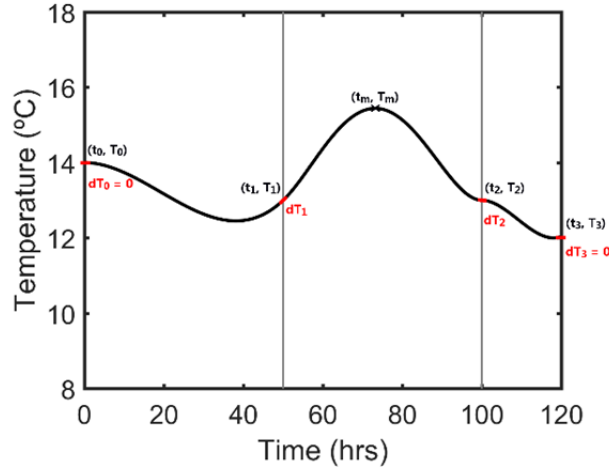
Figure 4. PWP Profile structure.

The decision vector corresponds to these 5 points and 2 derivatives, requiring 11 variables as the first point will always occur at t = 0:

$$d = [T_0, t_1, T_1, dT_1, t_{f2}, T_2, dT_2, t_{fm}, T_m, T_3, t_{f3}] \tag{25}$$

The cubic between t_0 and t_1 has 4 degrees of freedom so defined by 2 points (t, T) and 2 gradients (dT). The points (0, T_0) and (t_1, T_1) and gradient dT_1 are known directly from Eq. 25, which along with $dT_0 = 0$ fully defines the curve. A linear system of equations can be solved to give the coefficients of the corresponding polynomial. Similarly, the centre polynomial, here 5th order, is defined by 3 points and 2 derivatives: (t_1, T_1) and dT_1 known from the first section, T_2 and T_m are defined by d (Eq. 25) and the corresponding time of these points are defined by:

$$t_2 = t_1 + (t_{2max} - t_1) \times t_{f2} \tag{26}$$

$$t_m = t_1 + (t_2 - t_1) \times t_{fm} \tag{27}$$

allowing the linear system of equations to be solved to give the 5 coefficients of the middle quintic polynomial. The final cubic is defined with 2 points and 2 derivatives: (t_2, T_2) along with dT_2 are known from the previous segment, dT_3 is set as 0 so only a final point remains to be defined:

$$t_3 = t_2 + (t_{max} - t_2) \times t_{f3} \tag{28}$$

giving the profile end point (t_3, T_3) allowing the coefficients of the final polynomial to be determined. Therefore, the full profile, consisting of three section piecewise polynomial profiles, is uniquely represented by Eq. 25.

3. Results and discussion

The Strawberry algorithm[1] was executed using MATLAB 2016b on Windows 7 64bit running on a i7-4790 Intel CPU @ 3.60 GHz with 16.0 GB of installed memory. A variety of discretization levels were investigated for PWL profiles, in addition to the maximum number of generations (gen) and the

---

[1] http://www.ucl.ac.uk/~ucecesf/strawberry.html

number of solutions with each generation to propagate (referred to as the population size below), to investigate how these variables influence solution performance.

3.1 Solution convergence

Fig. 5 depicts how the trade-off curve for the two objectives evolves during the evolutionary process. As batch time minimisation and ethanol maximisation are the two objectives, the desirable solutions will be to the right and the bottom of the plot. Each coloured line joins non-dominated points (where no other solution improves on both objectives simultaneously) from the population which corresponds by colour to a generation number: this can be considered an approximation to the Pareto trade-off front. An elitism rule in place ensures that all non-dominated points pass to subsequent generations: favourable solutions are not lost and performance of the corresponding temperature profiles the beer fermentation from the front can only improve or remain unchanged in subsequent generations. The left panel represents PWL with N = 6 with a population of 100 solutions, while the right panel shows the PWP equivalent with the same population size.
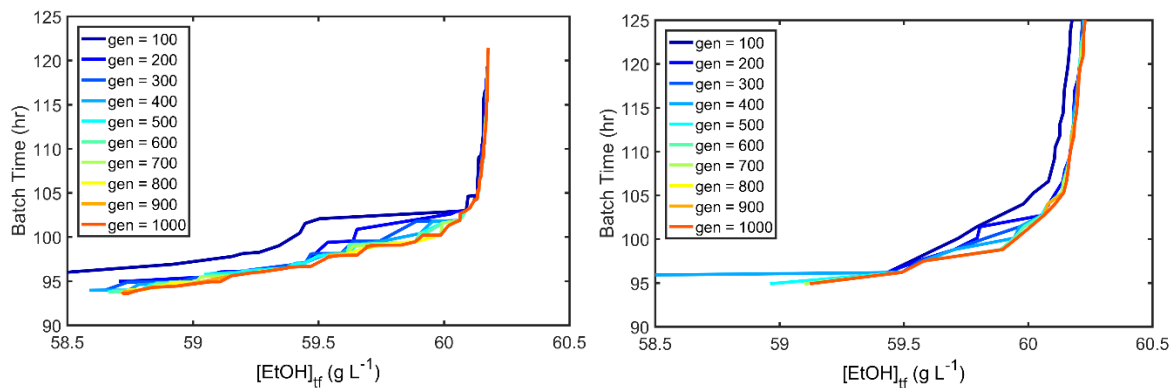

Figure 5. Improvement of non-dominated solutions over subsequent generations with population size of 200; left (a): PWL with N = 6; right (b): PWP.

In the left plot (a) it can be observed that the front moves to towards the bottom right (improves) significantly between generation 100 and 200 with PWL profiles. Over the following 500 generations continual improvement occurs, particularly regarding batch time reduction. The last 300 generations show minimal gains in either direction so it can be concluded that convergence has occurred to the final red line which we can consider the Pareto front for N = 6. The algorithm was repeated numerous times for the same conditions and settings, with convergence to a similar front achieved in under 1000 generations in all instances. When N was increased to both 12 and 18 it was found necessary to increase the number of generations to 2000 in order to achieve convergence. This is a direct result of the increased number of decision variables (Eq. 25) which leads to an increase in the size of the search space.

The right plot (b) again shows a significant improvement over the first 200 generations for PWP profiles, however here it is found that convergence is consistently observed sooner. No discernible improvement is achieved after generation 500, suggesting that the algorithm has converged on the most preferable profiles for this particular T(t) profile encoding.

Fig. 6 compares the final Pareto front approximation after convergence between the two strategies. The front after 2000 generations, over the same objective axis as Fig. 5, for increasing discretisation level (N) of PWL profiles, along with the front for PWP profiles after 500 generations. A high level of similarity between the lines is evident. This is surprising between the three PWL cases, given that with three times as many linear profile segments one might expect the considerably

increased level of control to permit significant improvements. However, this is not found to be the case. The blue line is significantly smoother than the others as a result of the greater number of non-dominated points in this solution set. It is worth noting that the search for candidate solution profiles for industrial fermentation is not concerned with the number of candidates detected, rather the suitability and effectiveness of the most promising candidates. As the PWP Pareto front very closely follows those from the PWL representation it is demonstrated that neither encoding is particularly favourable for attainable fermentation performance across the majority of the domain, as equivalently performing solutions are generated in both cases. However, it is observed that the lower left portion of front is not reproduced with PWP profiles, highlighting that the piecewise polynomial encoding restricts the solution space such that very low batch time (and low ethanol) solutions are not able to be produced. This is a results of near instantaneous temperature adjustments not being permitted with PWP profiles, the implications of which are considered in more detail in Section 3.3.
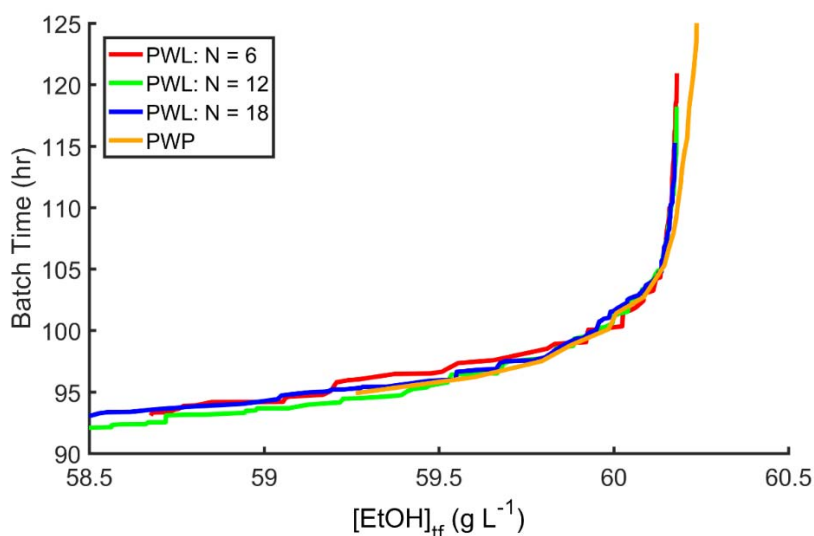
Figure 6. Non-dominated front after 2000 generations for various control profile discretisation levels.

3.2 Final solution populations

The plots shown in Fig. 7 represent the final populations from unique instances of the Strawberry algorithm being executed for the two different encoding strategies. Each hollow blue circular marker corresponds to the performance of a solution profile in the current population; those which are non-dominated are coloured red. The algorithm was executed numerous times using varying population sizes, and it has been found that a very similar Pareto approximation is produced with population sizes as low as 20 solutions. For both methods the density of the front, thus the number of promising candidate solutions discovered, increases with population size at the cost of required CPU time. Comparing the two plots in Fig. 7 shows that unlike PWL the PWP encoding is not able to produce solutions towards the bottom left of the axis; however given the low ethanol concentration this is unlikely to omit realistically desirable scenarios.
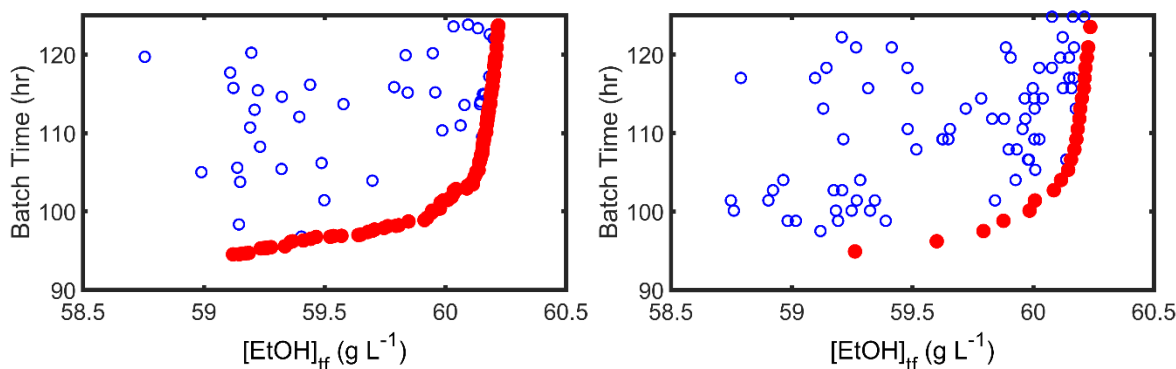
Figure 7. Final population (population size of 200) of solutions, left (a): PWL with N = 6; right (b): PWP. Solid red markers are non-dominated.

3.3 Corresponding solution profiles

In addition to the performance of the solutions in terms of objective function values, it is of critical importance to examine what the solutions represent in terms of control profiles to assess their suitability for industrial application. Fig. 8 represents samples of the profiles which make up the non-dominated solutions from the final populations for discretisation levels N = 6 and N = 18 for PWL profiles, with PWP profiles below. Not all the PWL profiles are of industrial value due to the ability to physically replicate them on real plant equipment. Solutions obtained with low N values are more suitable industrially. The number of manipulations required is smaller and temperature changes more gradual, as can be seen in the figure. Taking this into account, and considering the marginal improvement observed when increasing the discretisation level (Fig. 6), it is recommended to only pursue N = 6 solutions for the PWL implementation.

Not all of the solutions produced for high discretisation levels (i.e. N = 18) have undesired temperature variations. The third example presented in the middle row of Fig. 8 is a particularly promising case where the improved control permitted with higher discretisation acts to smooth the profile form, rather than to do the opposite as seen in the second plot from the same row. This suggests that merit may exist in refining the definition of the search space, D, by reformulating Eqs. 17-19 so that by design only favourable and implementable solutions may be considered by restricting the ability of the profiles to display a high level of variability in temperature. It must however be noted that simply the omission of rapid temperature changes should be avoided, as both a rapid increase or decrease at some point of the process is not uncommon and can be desirable to assist with the control of by-product production (Eqs. 14-15).

In contrast, all of the solutions obtained from the PWP method would be considered appropriate for implementation, with no restrictive variability possible due to the polynomial T(t) encoding. Comparing the solution forms between the two methods it can be seen that that there are many similarities. In particular the large temperature drop towards the end of the process appears to be an effective trait for controlling the by-product levels. Given the high similarity in attainable performance and the significantly improved inherent implementability and profile smoothness, using PWP profiles is the most promising strategy for industrial fermentation T(t) profile formulation via the Strawberry algorithm.

11
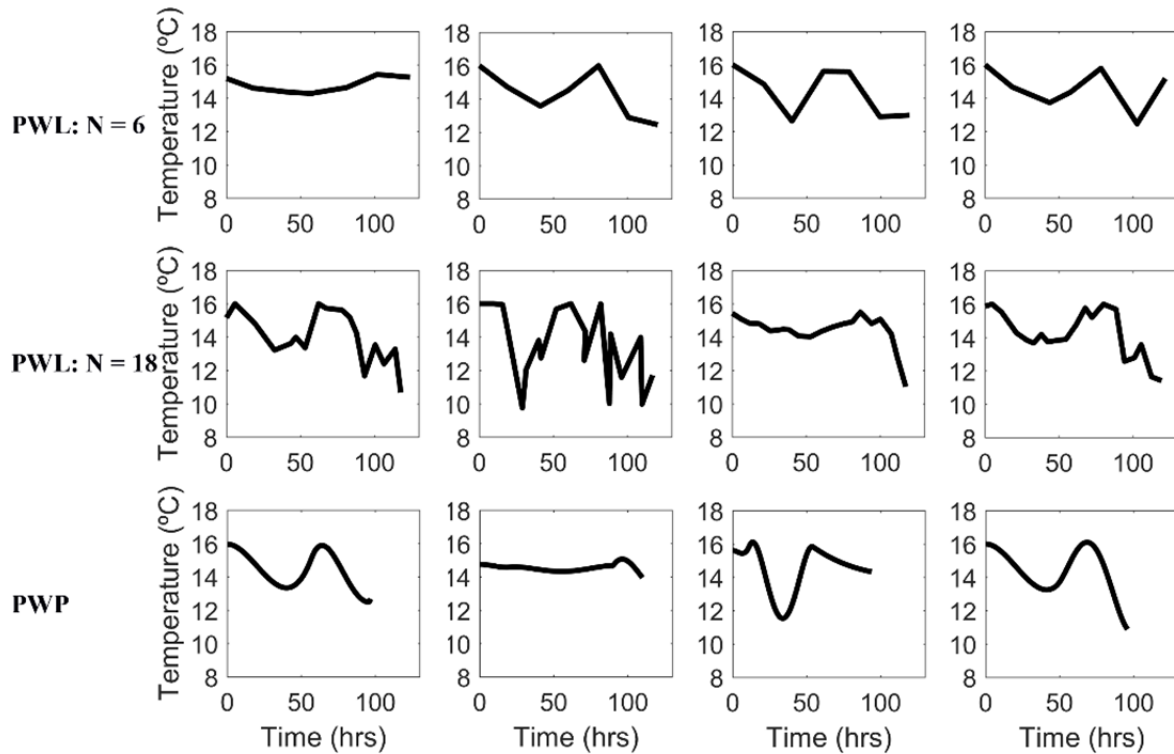
Figure 8. Example non-dominated profiles.

3.4 Solution profile performance – design heuristics

Fig. 9 shows specifically how the solution profiles correspond with their performance on the Pareto front for the preferable PWP method, which provides considerable insight into the fermentation process and as to how the performance is influenced by the temperature manipulation. The profiles which produce extremely short batch times (t_f < 105 hrs) at the cost of reduced ethanol concentration all have a comparable form. An initial high temperature is immediately lowered over the first 40 hours (16 °C to 14 °C). The temperature is then raised back to a peak momentarily around 16 °C at the 70 hour point before being reduced once more. Depending on the vessel size this cooling and heating cycle may be attainable on industrial fermentation equipment. In order for the vessel contents to achieve homogeneity, the cylindrical and conical portion of the tank must achieve thermal equilibrium. As such, for the assumptions in the lumped parameter model to apply there is exists a minimum time under which temperature variations cannot be realised, the time for which is a function of the vessel size. In contrast, the longer batch time solutions are more likely to be implementable on any scale of industrial fermentation vessel, as the lesser temperature variability means that homogeneity will more readily be achieved.

12
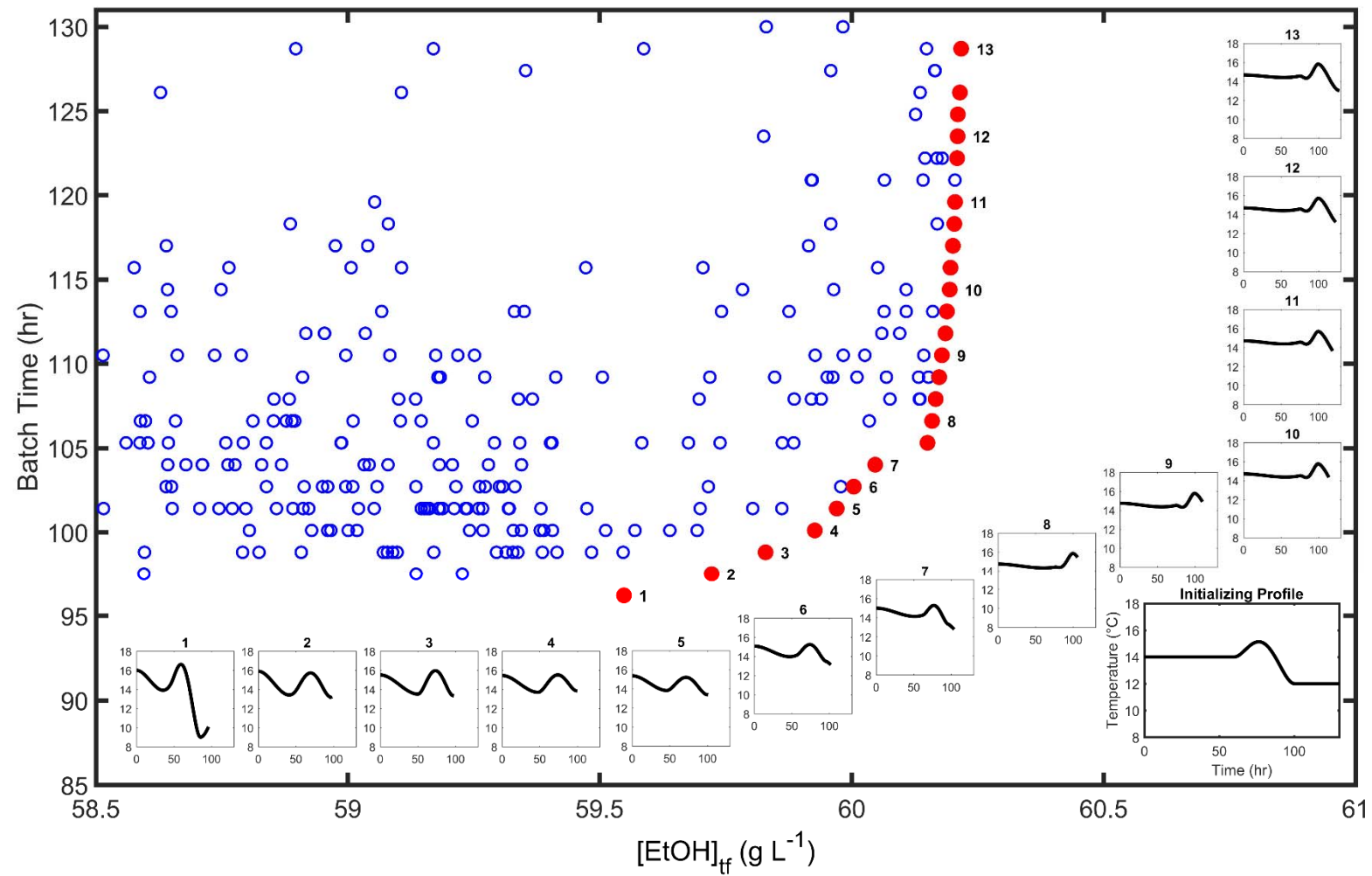
Figure 9. Pareto front of non-dominated optimal solutions to the multi-objective problem and corresponding T(t) profiles, for a quasi-A profile initialization..

In the lower portion of the Pareto set it is demonstrated that magnitude of the batch time range is comparable to the ethanol concentration range, i.e. a sacrifice of 5 hours in batch time can enable a significant 0.5 g L^{-1} increase in ethanol production. In contrast, the longer batch time subset of the front ($t_f > 105$ hrs) shows a very steep form, meaning that very minimal further gains in ethanol concentration are attainable, even upon increasing batch time as high as 130 hrs.

A gradual transition is observed in the T(t) profile form along the Pareto front. The initial dip in temperature becomes less pronounced as batch time is increased, an indication that this is an essential component of the temperature profile for extremely rapid beer fermentation. The first half of the profile continues to level off moving up the front, suggesting that this is useful for ensuring a very high ethanol yield. The presence of the later peak and subsequent cooling remains constant throughout the entire front, a feature that is known to assist with efficient fermentation while ensuring that the by-product constraints (Eq. 15-16) are not violated. Diacetyl compounds are consumed in the later stages of the process, with this feature of the temperature manipulation accelerating their consumption to fall below the tolerable level rapidly. It has been demonstrated that the tolerable level of diacetyl in the product is directly restrictive to the attainable batch time, due to the requirement to wait until a sufficient portion has been consumed (Rodman and Gerogiorgis, 2017), hence why this is an essential profile component for efficient beer fermentation throughout.

An important consideration which can be visualised on Fig. 9 is how slight variations in the temperature profile can affect the performance of the batch. Of the profiles highlighted in the figure, solutions 12 and 13 is an example of two profiles which are very similar in terms of the T(t) profile. However, the corresponding performance varies drastically, with the later requiring 5 more hours for completion. This stresses the importance of ensuring the temperature is accurately controlled in the fermenter, and that system homogeneity is ensured. Additionally, consideration should be given to solution robustness, ensuring that the manipulation employed will still perform adequately if slight deviations from the profile are encountered.

It has been demonstrated that the initialising solution profile can influence the output profiles when using an exact optimisation method (Rodman and Gerogiorgis, 2017). To investigate whether the stochastic method used in this study exhibits a similar deficiency, a range of initializing profiles have been used, with the corresponding profile maps shown in Figs. 9-12 for four cases. The initializing profiles presented here are approximations of the 4 candidate solutions highlighted previously from an exhaustive search (Rodman and Gerogiorgis, 2016), where minor deviations in the profiles are a result of the polynomial representation being unable to fully mirror the PWL equivalent.

From Figs. 9-12 it is shown that the approximate Pareto front of non-dominated points is extremely similar between the four cases, highlighting that attainable performance is not restricted or significantly influenced by the seed profile. While these four cases are initialised with promising candidates, very similar results were found when using either random or isothermal initialisation T(t) profiles, suggesting robustness of the stochastic optimisation strategy employed in this paper.

Although the performance of the non-dominated solutions is essentially identical, there are differences observed when comparing the corresponding T(t) profiles. The same overall trends are present, such that postulated heuristics for effective fermentation remain valid. The discrepancies in the profiles across different initialisations are minimal. For example the late peak in the temperature is shown to be less pronounced in the latter three cases, however is still present throughout.

The variability in the profiles shown across these four figures is no greater than the differences observed when re-initialising the algorithm numerous times using the exact same seed profile. As such it may be concluded that it is the stochastic nature of the method which is responsible for the solution profiles varying slightly across the four cases presented here, rather than the solution being sensitive to the 'initial guess' or initialising solution.
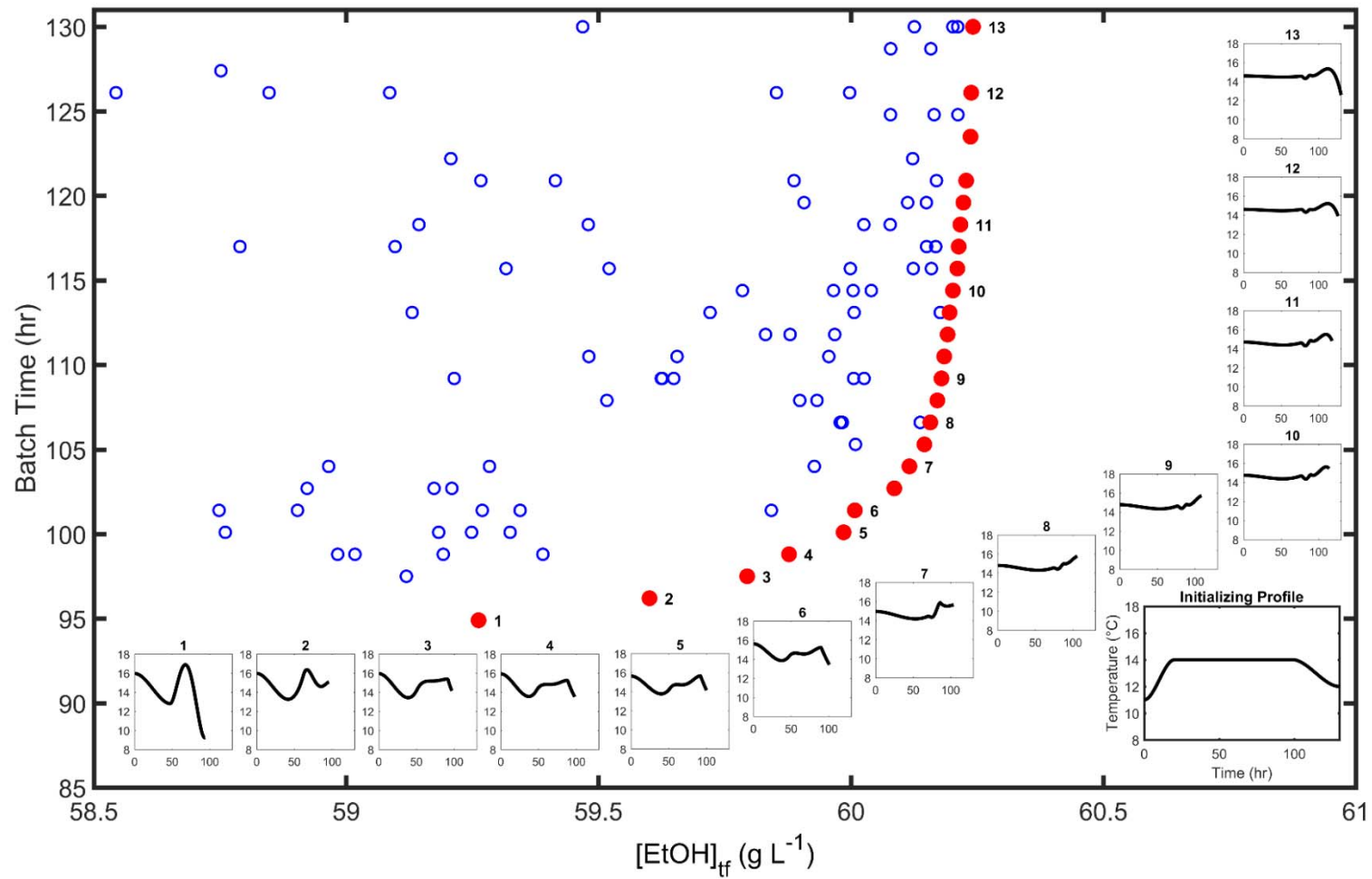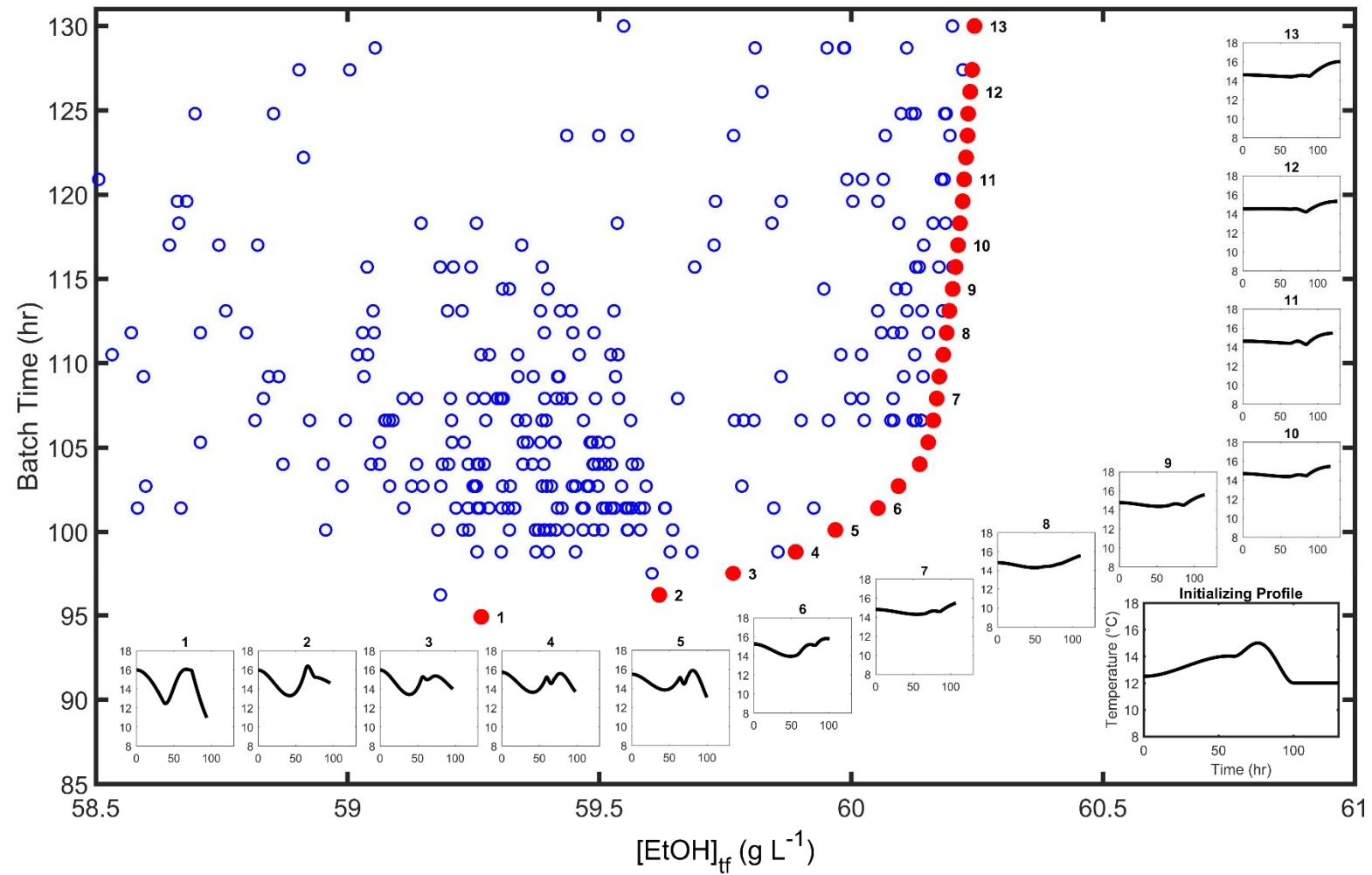
Figure 10. Pareto front of non-dominated optimal solutions to the multi-objective problem and corresponding T(t) profiles, for a quasi-B profile initialization
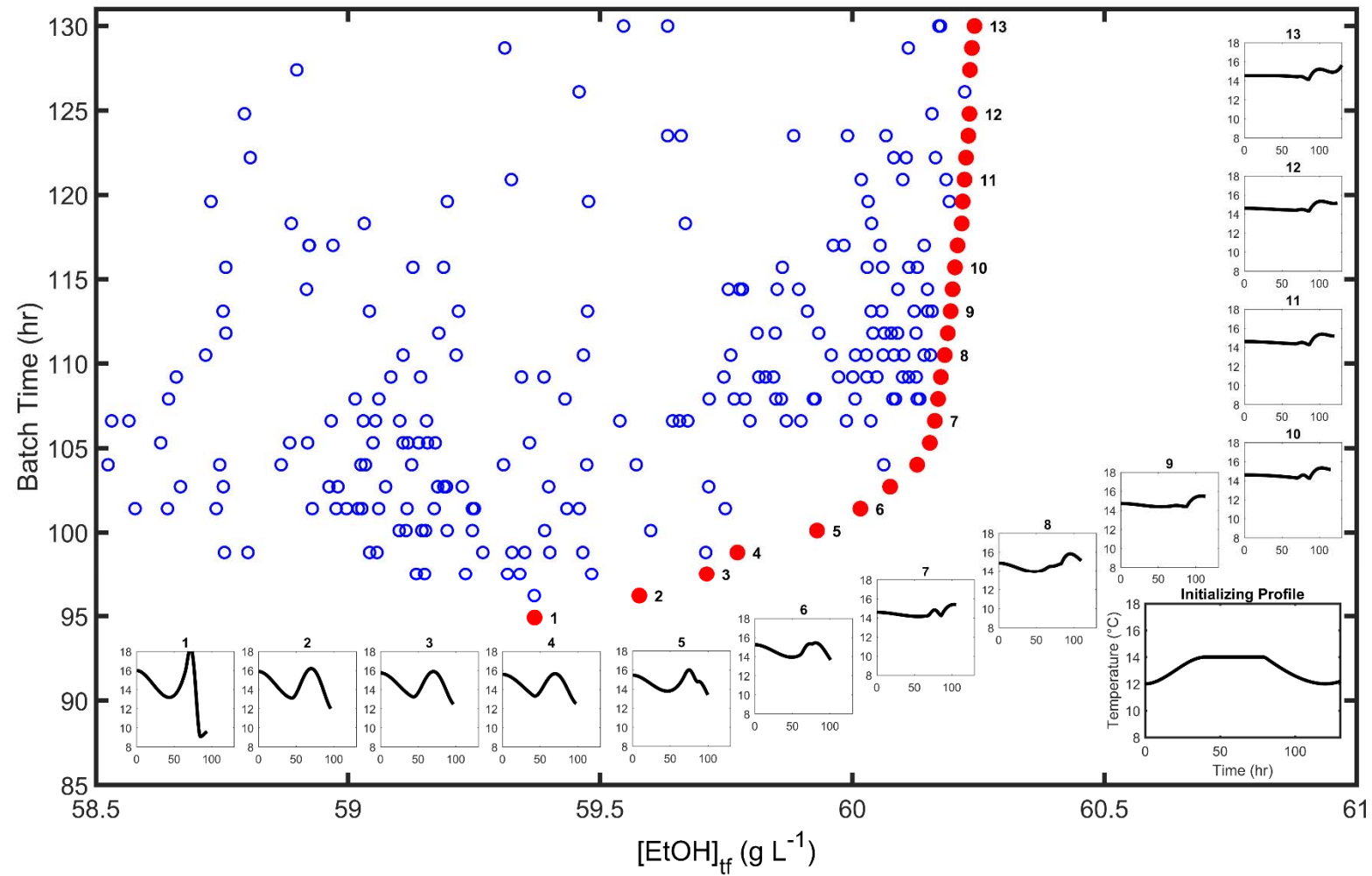
Figure 11. Pareto front of non-dominated optimal solutions to the multi-objective problem and corresponding T(t) profiles, for a quasi-C profile initialization

Figure 12. Pareto front of non-dominated optimal solutions to the multi-objective problem and corresponding T(t) profiles, for a quasi-D profile initialization

3.5 Solution profile performance – evaluation of profile encoding strategies

Table 1 presents performance metrics (terminal state concentrations and batch time) comparing the seed solution used for initialising the Strawberry Algorithm to an example of a non-dominated point in the N = 6 final population of the PWL method as well as an example of an attractive PWP solution. The two novel T(t) profiles produced in this study are shown in Fig. 13. It can be seen that there is a very high level of similarity the two profiles and the corresponding performance metrics, highlighting that the two unique profile encodings are able to produce comparable favourable solutions. While the upper row of Fig. 13 is from N = 6 the profile shown consists of 5 sections only: in this case the batch is complete before the final segment, so the profile is terminated.

Table 1. Example solution profile performance versus initializing profile.

Profile	$[EtOH]_{tf}$ (g L^{-1})	t_f (hr)	$[EA]_{tf}$ (ppm)	$[DY]_{tf}$ (ppm)
Input	59.1	113.5	1.350	0.09
Example New Solution, PWL (Fig. 13, top)	60.0	100.3	1.995	0.10
Example New Solution, PWP (Fig. 13, bottom)	60.0	100.0	1.996	0.10

Fig. 13 also shows the concentration trajectories over time within the batch while following each example profile. The similarity of the two profiles produces similar concentration progression throughout the respective batch. In both cases the desirable ethanol yield is rapidly achieved by permitting the product concentrations of undesirable species to increase towards the upper limits imposed by the corresponding constraints, Eqs. 14-15. The batch time saving of over 12 hours is noteworthy, suggesting the potential for significant process improvement via a potential plant throughput increase.
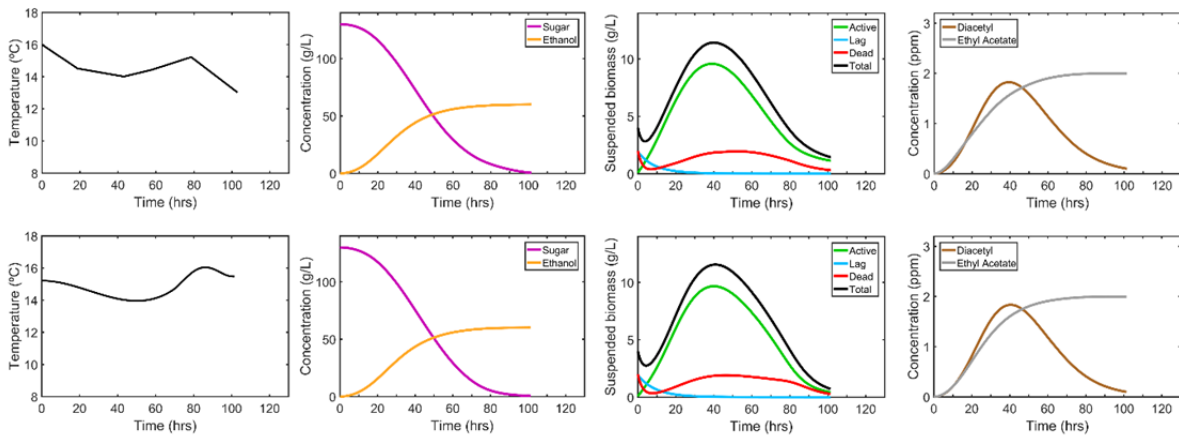
Figure 13. Example solution profiles and corresponding state trajectories: PWL above and PWP below.

4. Conclusions

The Strawberry Plant Propagation Algorithm has been used to explore process improvement potential of an industrial beer fermentation process via control vector (temperature profile) optimisation. Two different methods for representing temperature control profiles are investigated, considering batch time minimisation and ethanol yield maximisation as two simultaneous but conflicting objectives. Consistent convergence to the trade-off curve for the two objectives is demonstrated after an adequate number of generations have passed, the required number of which scales with number of decision variables, which in turn scales linearly with the discretisation level for the piecewise linear representation. It is found that a large population size is not necessary for the production of a dense approximation to the Pareto front, due to the inclusion of an elitism rule when generating new solution

populations (as presented in Algorithm 1). It has been demonstrated that if a piecewise linear profile is being used it can be beneficial to use a moderately low discretisation level. The minimal performance improvement upon increasing profile complexity is marginal next to the restrictiveness of the highly variable (in temperature) nature of profiles often computed. In contrast low discretisation solutions are inherently more suitable, given the reduced number of manipulations required and the more gradual temperature gradients. A piecewise polynomial encoding has been demonstrated to produce solutions with performance very similar to those from the piecewise linear approach. The added benefit is that these profiles are inherently appropriate for implementation, with no restrictive variability possible due to the improved polynomial T(t) encoding.

A dense Pareto front of solution profiles is identified describing the optimal trade-off between these process targets. Such data can be of value as a performance map when operators weigh up the relative importance of these two process targets. A subset of these output profiles can simultaneously reduce batch time and increase product ethanol concentration while satisfying constraints on by-products produced in the fermenters, representing significant improvements versus current industrial practice. A potential batch time reduction of over 12 hours has been demonstrated, coupled with a moderate improvement in ethanol content. The ability to identify novel temperature manipulations (control profiles) for improved performance enables brewers to reduce their batch times and operating costs. The effectiveness of a polynomial-based solution representation has thus been successfully highlighted for model-based constrained multi-objective optimisation, using for its implementation the Strawberry plant propagation algorithm, a tool already showcased for several dynamic process operation problems.

Acknowledgements

NOMENCLATURE LIST

Roman symbols

d	Decision vector (-)
dT	Profile gradient (K h^{-1})
DY	Diacetyl (-)
EA	Ethyl Acetate (-)
EtOH	Ethanol (-)
f	Fermentation inhibition factor (g L^{-1})
i	Time interval (-)
k_e	Ethanol affinity constant (g L^{-1})
k_s	Sugar affinity constant (g L^{-1})
k_x	Biomass affinity constant (g L^{-1})
N	Number of time intervals (-)
n_g	Number of generations (-)
n_p	Propagation size (-)
n_r	Maximum number of runners (-)
p	Current population (-)
S	Sugar (-)
t	Time (h)
t_f	Batch time (h)
t_{dev}	Deviation from uniform time (h)
t_{max}	Longest fermentation duration considered (h)
Δt	Time interval length (h)
T	Fermenter temperature (K)
z	Nondominated set (-)
X_A	Active biomass concentration (g L^{-1})
X_D	Dead biomass concentration (g L^{-1})
X_L	Latent biomass concentration (g L^{-1})
Y_{EA}	Ethyl acetate production stoichiometric factor (g L^{-1})

Greek symbols

μ_{AB}	Diacetyl consumption rate (g^{-1} h^{-1} L)
μ_{DT}	Specific cell death rate (h^{-1})
μ_{DY}	Diacetyl growth rate (g^{-1} h^{-1} L)
μ_E	Ethanol production rate (h^{-1})
μ_L	Specific cell activation rate (h^{-1})
μ_S	Sugar consumption rate (h^{-1})
μ_{SD}	Specific dead cell settling rate (h^{-1})
μ_x	Specific cell growth rate (h^{-1})

Subscripts and operators

( )$_0$	Initial condition (-)
( )$_i$	Condition in i[th] time interval (-)
( )$_{tf}$	Terminal condition (-)

References

Bosse, T. and Griewank, A., 2014. Optimal control of beer fermentation processes with Lipschitz-constraint on the control. Journal of the Institute of Brewing, **120**(4), 444-458.

Carrillo-Ureta, G., P. Roberts, and V. Becerra, 2001. Genetic algorithms for optimal control of beer fermentation. Proc. IEEE Int. Symp. Intell. Control, 391-396.

Che, A., Zhang, Y. and Feng, J., 2017., Bi-objective optimization for multi-floor facility layout problem with fixed inner configuration and room adjacency constraints, *Computers and Industrial Engineering*, **105**: 265-276

de Andrés-Toro, B., Giron-Sierra, J., Lopez-Orozco, J., and Fernandez-Conde, C., 1998. A kinetic model for beer production under industrial operational conditions. *Mathematics and Computers in Simulation*, **48**(1): 65-74.

Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T.A.M.T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, **6**(2): 182-197.

Fraga, E. S., Salhi, A., Zhang, D. and Papageorgiou, L., 2015. Optimisation as a Tool for Gaining Insight: An Application to the Built Environment. *Journal of Algorithms & Computational Technology*, **9**(1), 13-26.

Fraga, E. S. and Amusat, O., 2016. Understanding the impact of constraints: a rank based fitness function for evolutionary methods, in: Pardalos, P. M., Zhigljavsky, A. and Zilinskas J. (Eds.), *Advances in Stochastic and Deterministic Global Optimization*, Springer, 243-254.

Fraga, E. S., Salhi, A. and Talbi, E-G., 2018. On the impact of representation and algorithm selection for optimisation in process design: motivating a meta-heuristic framework, in Amodeo, L., Talbim E-G. and Yalaoui, G. (Editors), *Recent Developments in Metaheuristics*, Springer, 141-149.

Gee, D. A. and Ramirez, W. F. 1988. Optimal temperature control for batch beer fermentation. *Biotechnology and Bioengineering*, **31**, 224-234.

Gujarathi, A. M., Sadaphal, A., and Bathe, G. A., 2015. Multi-objective optimization of solid state fermentation process. *Materials and Manufacturing Processes*, **30**(4): 511-519.

Hanke, S., Ditz, V., Herrmann, M., Back, W., Becker, T and Krottenthaler, M., 2010. Influence of ethyl acetate, isoamyl acetate and linalool on off-flavour perception in beer. *Brewing Science*, **63**(7): 94-99.

Izquierdo-Ferrero, J. M., Fernández-Romero, J. M., Luque de Castro, M. D., 1997. On-line flow injection–pervaporation of beer samples for the determination of diacetyl. *The Analyst*, **122**(2): 119-122.

Kessler, T., Logist, F., Mangold, M., 2017. Bi-objective optimization of dynamic systems by continuation methods, *Computers and Chemical Engineering*, **98**: 89-99.

Li, K., Fialho, A., Kwong, S., Zhang, Q., 2014. Adaptive operator selection with bandits for a multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evolutionary Computation*, **18**(1): 114–130.

Maria, G., and Crişan, M., 2017. Operation of a mechanically agitated semi-continuous multi-enzymatic reactor by using the Pareto-optimal multiple front method, *Journal of Process Control*, **53**: 95-105.

Miettinen, K., 1999. *Nonlinear Multiobjective Optimization*, Springer.

Rodman, A.D. and Gerogiorgis. D. I., 2016a. Multi-objective process optimisation of beer fermentation via dynamic simulation. *Food and Bioproducts Processing*, **100**: 255-274.

Rodman, A.D. and Gerogiorgis. D. I., 2016b. Dynamic simulation and visualisation of fermentation: Effect of process conditions on beer quality. *IFAC-PapersOnLine*, **49**(7): 615-620.

Rodman, A.D. and Gerogiorgis, D.I., 2017. Dynamic optimisation of beer fermentation: sensitivity analysis of attainable performance vs. variable product flavour constraint levels. *Computers and Chemical Engineering* **106**: 582-595.

Salhi, A. and Fraga, E., 2011. Nature-inspired optimisation approaches and the new plant propagation algorithm. *Proceedings of the International Conference on Numerical Analysis and Optimisation* (ICeMATH2011). K2: 1-8.

Sorek, N., Gildin, E., Boukouvala, F., Beykal, B. and Floudas, C.A., 2017. Dimensionality reduction for production optimization using polynomial approximations. *Computational Geosciences*, **21**(2): 247-266.

Trelea, I. C., Titica, M., Landaud, S., Latrille, E., Corrieu, G. and Cheruy, A., 2001. Predictive modelling of brewing fermentation: from knowledge-based to black-box models. *Mathematics and Computers in Simulation*, **56**(4): 405-424.

Vanderhaegen, B., Neven, H., Verachtert, H. and Derdelinckx, G., 2006. The chemistry of beer aging – a critical review. *Food Chemistry*, **95**(3): 357-381.

Xiao, J., Zhou, Z., Zhang, G., 2003. Ant colony system algorithm for the optimization of beer fermentation control. *Journal of Zhejiang University Science*, **5**(12): 1597-1603.

Zhang, P., Chen, H., Liu, A. and Zhang, Z., 2015. An iterative multi-objective particle swarm optimization-based control vector parameterization for state constrained chemical and biochemical engineering problems, *Biochemical Engineering Journal*, **103**: 138-151.