A Non-linear Model Predictive Control Based on Grey-Wolf Optimization Using Least-Square Support Vector Machine for Product Concentration Control in L-Lysine Fermentation

Bo Wang [1], **Muhammad Shahzad [1,*]**, **Xianglin Zhu [1]** and **Khalil Ur Rehman [1]** and **Saad Uddin [2]**

[1]   School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China; wangbo@ujs.edu.cn (B.W.); zxl4390@126.com (X.Z.); 5102180315@stmail.ujs.edu.cn (K.U.R.)
[2]   School of Mechanical Engineering, Jiangsu University, Zhenjiang 212013, China; 5102180314@stmail.ujs.edu.cn
*   Correspondence: 5102180312@stmail.ujs.edu.cn

check for updates

Abstract: L-Lysine is produced by a complex non-linear fermentation process. A non-linear model predictive control (NMPC) scheme is proposed to control product concentration in real time for enhancing production. However, product concentration cannot be directly measured in real time. Least-square support vector machine (LSSVM) is used to predict product concentration in real time. Grey-Wolf Optimization (GWO) algorithm is used to optimize the key model parameters (penalty factor and kernel width) of LSSVM for increasing its prediction accuracy (GWO-LSSVM). The proposed optimal prediction model is used as a process model in the non-linear model predictive control to predict product concentration. GWO is also used to solve the non-convex optimization problem in non-linear model predictive control (GWO-NMPC) for calculating optimal future inputs. The proposed GWO-based prediction model (GWO-LSSVM) and non-linear model predictive control (GWO-NMPC) are compared with the Particle Swarm Optimization (PSO)-based prediction model (PSO-LSSVM) and non-linear model predictive control (PSO-NMPC) to validate their effectiveness. The comparative results show that the prediction accuracy, adaptability, real-time tracking ability, overall error and control precision of GWO-based predictive control is better compared to PSO-based predictive control.

Keywords: model predictive control; machine learning; grey-wolf optimization;least-square support vector machine; L-Lysine fermentation

1. Introduction

The invention and advancement of modern computationally fast microprocessors have paved the path for Model Predictive Control (MPC). Recently, MPC has become one of the efficient predictive control algorithms in large scale applications such as aerospace systems [1], plastic industry [2], wastewater treatment plant [3], power electronics industry [4] and many others. It can handle multi-process variables and incorporate practical constraints on these variables. The basic elements of MPC are: process model, cost function and optimization algorithm. The process model plays a key role in the performance of MPC. It should encompass the precise dynamics of the process. In most applications, linearized models are used in MPC. However, practical processes exhibit severe non-linearity that cannot be captured by linear models. In addition, these linear models cannot cover a wide range of operating conditions. Artificial intelligence-based models are famous for their

self-learning and non-linear modeling ability, and have attracted many researchers to model these non-linear behaviors such as artificial neural network (ANN) and support vector machine (SVM) [5,6].

L-Lysine is the second most produced amino acid in the world. An estimated global market of L-Lysine is 2.2 million tons per year, which is growing at the rate of 10% per year [7]. It is mainly used in food, animal feed, pharmaceuticals and cosmetics industries. This increasing demand in global market compels industries to look for alternatives to enhance the productivity instead of expanding the physical capacity of plants, which is expensive and time consuming. One of the best ways to boost the productivity is to monitor and control product concentration. An excessive increase of product in reactor causes osmotic stress or catabolic repression for bacteria during cultivation [8]. Temperature, pH, initial substrate concentration, air flow rate and agitation rate are five paramount factors to enhance and control product concentration [9]. However, L-Lysine fermentation is a highly non-linear process and product concentration cannot be directly measured in real time using physical sensors in the fermentation process. Some costly off-line analysis methods such as dry weight method, ninhydrin colorimetric method and optical density method are used to measure product concentration, but these methods have limitations, such as large time delay and high infection rate. Therefore, off-line lab analyzers cannot meet the requirements of real-time control of the fermentation process with these limitations. To solve the above mentioned problems, machine-learning-based prediction models [10] have been effectively used by the researchers [11].

Machine-learning-based prediction models construct inferential mathematical models by making use of easily measurable variables (for example, pH, temperature and dissolved oxygen) obtained from physical sensors and predict the unmeasurable key variables (product concentration) [12]. The successful implementation of these prediction models has revolutionized the fermentation industry. Researchers have introduced many data-driven prediction models to model fermentation process for different objectives. ANN is exploited to design a prediction model for bioethanol production by defining optimal number of hidden layers and hidden units [13]. However, if the specific structure of network is not known, ANN loses its generalization ability to model non-linear regression problems and suffers from overfitting problem [14]. SVM, which is based on structural risk minimization problem, has successfully resolved the aforementioned problem using simple statistical learning theory. It is used to solve many industrial applications, such as a modelling method based on SVM is proposed for Glutamic acid fermentation to predict product concentration [15]. Unfortunately, huge time cost for training and curse of dimensionality limits usefulness of SVM in many applications [16]. Least-square support vector machine (LSSVM) alleviates this problem by converting convex Quadratic Programming (QP) problem in SVM to a system of linear equations. In this way, LSSVM provides fast training speed and efficiently finds global optimum solution if the parameters are selected carefully [17].

In this study, LSSVM is selected as a prediction model to measure product concentration of L-Lysine fermentation process. To improve the prediction accuracy and robustness of LSSVM, two model parameters, namely kernel width 'σ' and regularization factor 'g' must be optimized. A very large value of penalty parameter 'g' would lead to remarkably high accuracy on training data but less accuracy on test data, while less value makes the model less functional resulting in poor performance. In addition, an excessively large value of kernel width control factor 'σ' inflicts overfitting problem and small value results in under-learning problem [18]. *Xinhua and Ming* [19] hybridized Particle Swarm Optimization (PSO) algorithm and LSSVM, and proposed a PSO-LSSVM model to predict the deformation on surrounding rocks of underground caverns. Zhu [20] used PSO-LSSVM to measure key variables on-line in fermentation process. The comparative studies show that Grey-Wolf Optimization (GWO) algorithm has overall best performance in terms of search efficiency and convergence speed for finding a global optimum solution as compared to PSO, ABC, FFA, CS, BA, FPA, GSA, DE, EP and ES [21,22]. Hence, in this work, GWO is used to get optimum parameters of proposed LSSVM prediction model. It is notable to mention "No Free Lunch" (NFL) theorem here, which proves that an optimization algorithm works well in some specific optimization

problems, but same algorithm is not successful in other set of optimization problems [23]. In our case, GWO provides best results that fulfill our requirements.

Many studies proposed different machine-learning-based prediction models for MPC to predict future outputs. A generalized predictive control (GPC) scheme is proposed to control concentration of bacteria by using a linearized PSO-LSSVM model [24]. However, linear models show limited control performance because industrial fermentation processes generally exhibit complex and severe non-linear behavior. Furthermore, if a non-linear model is used, the optimization problem becomes a non-convex problem, which is solved by using Non-linear Programming (NP) method. Conventionally, NP is involved in computationally expensive step of determining the hessian matrix and its inverse [25]. In addition, these solutions are highly dependent on the selection of initial point value and can easily fall in local optimal region (solution) [26]. The biggest challenge in designing an NMPC is to find an algorithm that minimizes a cost function in real time. The cost function is usually non-convex, high-dimensional with complex and non-linear constraints [27].

This work employs a novel derivative-free approach to solve non-linear and non-convex rolling optimization problem in NMPC to control product concentration of L-Lysine fermentation process. To the best of our knowledge, GWO has not been applied to solve rolling optimization problem in NMPC for control problems in fermentation process. GWO has fast convergence speed, involves lesser operators in computations and requires a few adjustable parameters [28]. These properties make GWO an ideal candidate to solve a non-linear, non-convex optimization problem of NMPC in real time. Furthermore, the performance of NMPC is highly dependent on the accuracy of the prediction model. Thus, constructing an accurate prediction model is a crucial step. The proposed non-linear GWO-LSSVM prediction model is employed in NMPC for prediction of future output values. In addition, GWO is imposed to solve rolling optimization problem in NMPC as Chen [29] designed a NMPC based on PSO to control greenhouse temperature. In this way, a non-linear MPC is designed which easily incorporates a non-linear prediction model and solves a non-convex optimization problem in real time. The final results of GWO-LSSVM prediction model and non-linear GWO-NMPC control algorithm are compared with PSO-LSSVM and PSO-NMPC, respectively. The results show that the predicted values by GWO-LSSVM are very close to actual values as compared to PSO-LSSVM, and product concentration follows an optimal trajectory by employing real-time GWO-NMPC control strategy. Moreover, the results of GWO-NMPC surpass PSO-NMPC in terms of error tracking and adaptability.

The rest of the paper is structured as: Section 2 consists of materials and methods, which explains MPC basics, LSSVM model, GWO algorithm, proposed GWO-LSSVM prediction model, GWO-NMPC algorithm and experimental setup. Section 3 includes results and discussion. The paper is concluded in Section 4.

2. Material and Methods

2.1. Model Predictive Control (MPC)

Predictive control does not correspond to a particular control methodology but more precisely an abundant variety of control schemes, which exploits process model to obtain future control inputs that will force the system response to follow a desired response [30]. To accomplish the above mentioned objective, it minimizes a user defined objective function to obtain optimal future control inputs over a predefined prediction horizon (N_{pred}) and control horizon (N_{con}). Several MPC algorithms are different from each other because of the prediction models used to simulate the actual process and objective function that is used to solve the optimization problem. The basic structure of MPC is shown in Figure 1. The future control inputs are calculated by minimizing the error between

the predicted output and a reference by solving an optimization problem. The general expression for an objective expression is as follows:

$$J(N_{pred}, N_{con}, u) = \sum_{j=1}^{N_{pred}} Q_{out}(j)(y_{pred}(j+t|t) - y_{ref}(j+t)^2 \atop \Delta u$$
$$+ \sum_{j=1}^{N_{con}} R_{in}(j)(\Delta u(j+t-1))^2 \tag{1}$$

subjected to constraints:

$$u_{min} \leq u \leq u_{max}$$
$$\Delta u_{min} \leq \Delta u \leq \Delta u_{max} \tag{2}$$
$$y_{min} \leq y \leq y_{max}$$

where u, Δu, y_{pred}, y_{ref}, R_{in}, Q_{out} represent input, input increment, predicted output, desired reference, input penalization factor, error coefficient, respectively, and u_{min}, u_{max}, Δu_{min}, Δu_{max}, y_{min}, y_{max} denote lower and upper bounds on control input, control increment, and control output, respectively. The steps included in MPC are:

1. Calculate output at the current time and calculate future outputs up to the prediction horizon N_{pred}.
2. Construct an objective function using predicted and reference values over a prediction and control horizon.
3. Minimize objective function to calculate optimal values of future inputs $U^{op} = u_0^{op}, u_1^{op}, \ldots, u_{N_{con}-1}^{op}$.
4. Apply the first predicted input u_0^{op} and discard all other future input values. Repeat the whole process at next sampling time $t+1$.
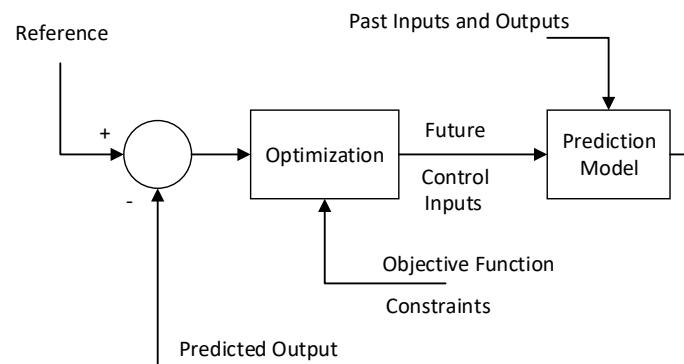

Figure 1. The Basic Structure of MPC.

2.2. Least-Square Support Vector Machine (LSSVM)

To solve the computational complexity problem of SVM [31], Suykens proposed LSSVM [32]. In LSSVM, an equality constraint is introduced instead of inequality in SVM and a complex QP optimization problem is converted into the equations of linear system. In this way, the model decomposition and prediction problems can be solved efficiently. The basic principle is as follows:

Given l sample points for training, $\{(x_i, y_i)|i = 1, \ldots, l\}, x_i \in R^n$ is an input vector and $y_i \in R$ represents corresponding outputs. The approximation function in LSSVM is defined as:

$$y(x_i) = \omega^T \varphi(x_i) + b \tag{3}$$

The optimization problem for regression is as follows:

$$\min_{\omega,\xi,b} \ J(\omega,\xi) = \frac{1}{2}\omega^T\omega + \frac{g}{2}\sum_{i=1}^{l}\xi^2 \tag{4}$$

$$s.t. \quad y_i(x_i) = \omega^T\varphi(x_i) + b + \xi, \ i = 1,\ldots,l$$

where ω is a weight vector, $g \in R^+$ is penalty parameter, ξ_i is an error variable, b is the deviation and $\varphi(\cdot)$ is mapping to a high-dimensional space. Lagrange method is used to optimize the above problems:

$$Ł(\omega,\xi,b,\alpha) = \frac{1}{2}\omega^T\omega + \frac{g}{2}\sum_{i=1}^{l}\xi^2 - \sum_{i=1}^{l}\alpha_i(\omega^T\varphi(x_i) + b + \xi_i - y_i) \tag{5}$$

where α_i is a Lagrange multiplier. According to Karush–Kuhn–Tucker (KKT) conditions, the transformation to the linear equation is as follows [17]:

$$\begin{pmatrix} 0 & 1_l^T \\ 1_l & K+g^{-1}I_l \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} \tag{6}$$

where $y = [y_1, y_2, \ldots, y_l]^T$, $1_l = [1, \ldots, 1]^T$, I_l is $l_t h$ ordered unit matrix, $\alpha = [\alpha_1, \ldots, \alpha_l]^T$ and K is the kernel function matrix that satisfy Mercer's conditions:

$$K = \varphi(x_i)^T\varphi(x_j), \ (i,j) = 1,\ldots,l \tag{7}$$

In this paper, radial basis function is chosen as the kernel function because of its excellent performance and generalization ability, which is given as follows [33]:

$$K = K(x, x_i) = \exp^{\frac{-|x-x_i|^2}{2\sigma^2}} \tag{8}$$

here σ is the kernel function width. Finally, the function of LSSVM is estimated as:

$$y(x) = \sum_{i=1}^{l}\alpha_i K(x, x_i) + b \tag{9}$$

The prediction accuracy and generalization ability of LSSVM regression model strongly depends upon the penalty parameter 'g' and kernel width 'σ'. So, these two parameters need to be optimized.

2.3. Grey-Wolf Optimization (GWO)

Mirajlili et al. proposed the GWO algorithm which imitates the social behavior of a grey wolf pack [22]. The grey wolves are divided into four categories namely *alpha*, *beta*, *delta* and *omega*. *Alpha* represents highest category and consists of leaders of the whole pack. *Alpha* wolves are responsible for making daily life decisions like hunting a prey, moving forward or stopping, sleep time and place. *Beta* group facilitates the *alpha* group in formulating these strategies and implementing commands on other lower categories. The third *delta* class is dedicated for fulfillment of above commands and controls *omega*. The lowest rank *omega* mainly obeys all instructions by superior departments. The hunting plan comprises of three steps: identifying and chasing the prey; encircling and harassing prey until it stops resilience; attacking on prey.

During the optimization process, it is assumed that the size of the grey wolf population is n, in an unknown d-dimensional search space. The position of grey wolves is denoted by $X_{wi} = [x_i^1, x_i^2, \ldots, x_i^d]$. *Alpha* is considered to be the best fittest solution and its position is denoted by X_α.

Then, *beta* and *delta* are ranked as second, third best solutions and their locations are represented by X_β and X_δ, respectively. The remaining solutions represent *omega* class of pack.

The encircling strategy of the hunting process is mathematically modeled by the following equations:

$$\vec{D} = |\vec{C}.\vec{X}_p(t) - \vec{X}_w(t)| \tag{10}$$

$$\vec{X}_w(t+1) = \vec{X}_p(t) - \vec{A}.\vec{D} \tag{11}$$

where t indicates the current iteration, $\vec{A}, \vec{C}$ are coefficient vectors, $\vec{X}_p$ and $\vec{X}_w$ denote the position of the prey and wolf, respectively. $\vec{A}$ and $\vec{C}$ vectors are calculated as follows:

$$\vec{A} = 2\vec{a}.\vec{r}_1 - \vec{a} \tag{12}$$

$$\vec{C} = 2.\vec{r}_2 \tag{13}$$

where with an increase in number of iterations, $\vec{a}$ decreases linearly from 2 to 0. $\vec{r}_1$ and $\vec{r}_2$ are random numbers in range $[0,1]$. The hunting process is performed under the guidance of *alpha*. *Beta* and *delta* might also facilitate *alpha* in trapping a prey. So top three departments (*alpha, beta, delta*) have the best information (best solution) about the prey. These best solutions found so far are saved iteratively and other search agents (*omega*) are forced to follow and update positions according to these best positions. The mathematical equations that encapsulate all the above scenario are as follows:

$$\vec{D}_\alpha = |\vec{C}_1.\vec{X}_\alpha - \vec{X}| \tag{14}$$

$$\vec{D}_\beta = |\vec{C}_2.\vec{X}_\beta - \vec{X}| \tag{15}$$

$$\vec{D}_\delta = |\vec{C}_3.\vec{X}_\delta - \vec{X}| \tag{16}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1.\vec{D}_\alpha \tag{17}$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2.\vec{D}_\beta \tag{18}$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3.\vec{D}_\delta \tag{19}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{20}$$

2.4. GWO-LSSVM Prediction Model

The parameters 'g' and 'σ' of the LSSVM model play a critical role in the prediction accuracy. A very large value of penalty parameter 'g' would lead to remarkably high accuracy on training data but less accuracy on test data, while low value of 'g' makes the model less functional which results in poor performance [18]. In addition, an excessively large value of kernel factor 'σ' inflicts overfitting problem and small value results in under-learning problem. The kernel width 'σ' defines the effect of a single training example on other examples. Hence, there is a need to choose the values of LSSVM model parameters 'g' and 'σ' carefully. Researchers have used different optimization algorithms to select the optimum values of critical parameters of regression models, such as the PSO algorithm [24]. In this study, an efficient metaheuristic GWO algorithm is proposed to find best suitable parameters of LSSVM prediction model as shown in Figure 2. The steps of GWO-LSSVM are as follows:

Step 1: Prepare train, test, cross-validation data and perform pre-processing (normalization). Define number of search agents, maximum iterations, dimension of parameters to be optimized, lower and upper bounds.

Step 2: Randomly initialize *alpha, beta, delta* and *omega* positions, and $\vec{a}, \vec{A}$, and $\vec{C}$. Train LSSVM model on training data using these positions as 'g' and 'σ' value.

Step 3: Calculate fitness value of each search agent position. The fitness value corresponds to prediction accuracy of trained model on cross-validation data, which is calculated using user defined fitness function. In this study, *RMSE* is used as a fitness function given in Equation (22).

Step 4: Update the positions using Equations (14)–(20) and $\vec{A}$, $\vec{C}$ and $\vec{a}$ using Equations (12) and (13).

Step 5: Calculate again the fitness value of all updated positions.

Step 6: Rank and store the best solution obtained so far using fitness value. Repeat from step (4) to step (6) until maximum cycles are reached.

Step 7: Train again LSSVM model with best solution obtained from above steps and check the prediction accuracy on new test data to verify again model functionality.
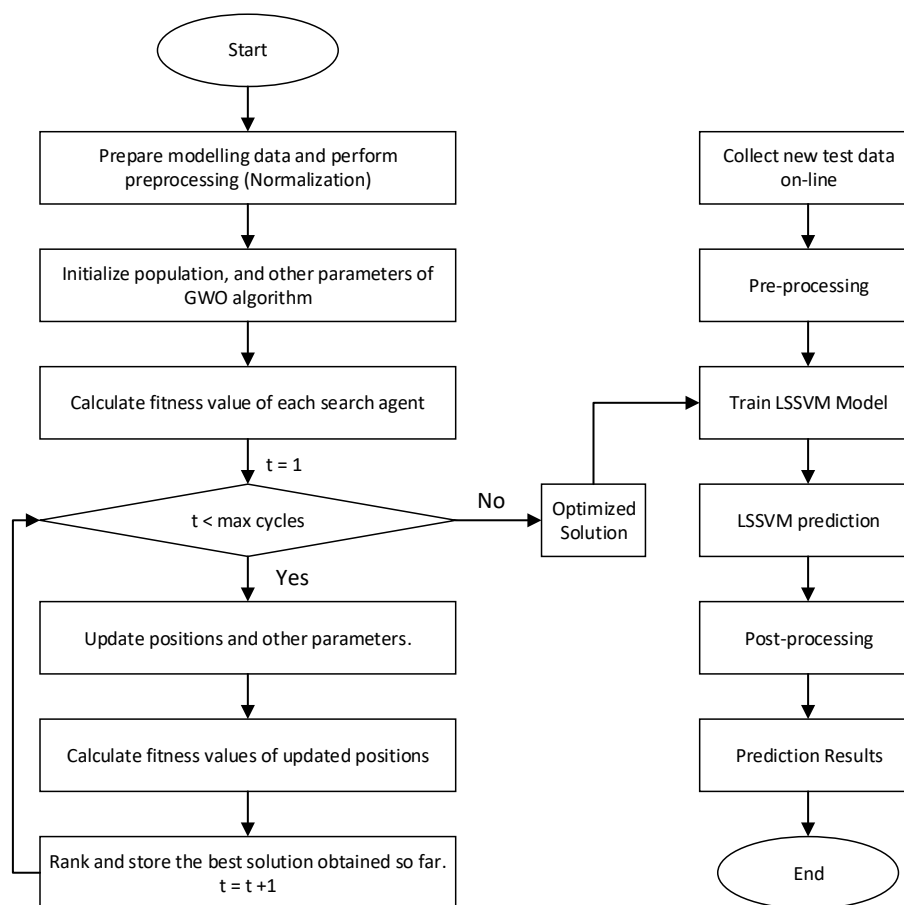
Figure 2. GWO-LSSVM prediction model.

2.5. GWO-NMPC Control Algorithm

The proposed GWO-LSSVM prediction model is employed in NMPC for prediction of future output values. Furthermore, the optimization problem in NMPC is solved by using GWO algorithm. Hence, a GWO-based non-linear MPC control strategy is used to achieve the objective of this study. GWO can incorporate the constraints on input, input increment and output value. The fitness function in GWO is replaced by the defined objective function. In this case, the position of the pack in GWO denotes the future control increments (penalty factor and kernel width in case of LSSVM optimization). GWO algorithm optimizes the objective function and finds optimum values of control increment Δu. The steps of GWO-NMPC are as follows:

Step 1: Control input variables, output variable and reference trajectory are defined.

Step 2: The constraints on inputs, input increments and outputs are defined.

Step 3: The control objective is accomplished by using an objective function as in Equation (1).

Step 4: In objective function, the predicted output 'y_{pred}' is estimated by using proposed GWO-LSSVM model.

Step 5: For each sampling interval, GWO optimizes the objective function and calculates the optimum values of control input increment Δu.

Step 6: The future control inputs are calculated by using following equation:

$$u(t+1) = u(t) + \Delta u(t+1) \tag{21}$$

where t, u, Δu represent current sampling time, control input and control increment, respectively.

Step 7: Finally, calculated input is applied to the process and output feedback strategy is employed.

The final GWO-based non-linear MPC using a GWO-LSSVM prediction model control scheme for controlling the product concentration in L-Lysine fermentation process is shown in Figure 3. The prediction error is corrected on-line using output feedback.
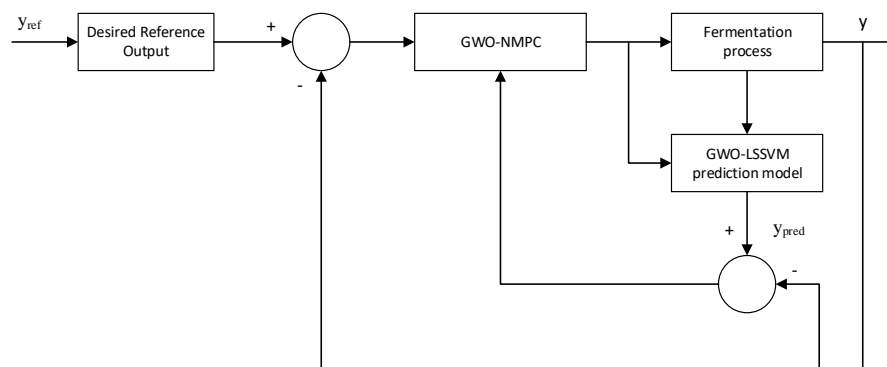
Figure 3. GWO-LSSVM-NMPC to control L-Lysine product concentration.

2.6. Experimental Setup

The experiment of L-Lysine fed-batch fermentation was carried out at the control system platform of Jiangsu University. The RT-100L-Y fermenter model was used to perform this experiment. To make the experiment close to the actual production process, the experimental process was designed as follows:

1.　In a 30 L mechanical stirring fermenter, fed-batch fermentation was conducted. The environmental parameters and physical parameters in the fermentation process were collected in real time by a digital measurement and control system composed of ARM development platform, and transmitted to the industrial control computer in the control room via a serial communication line. The time period for every batch was 72-h and the sampling time period was 15 min. The auxiliary inputs (such as temperature T, pH, agitation speed rate u_1, dissolved oxygen D_o, air flow rate u_2 and acceleration rate of ammonia flow u_3) were collected in real time. The key variable product concentration 'P' was sampled after every 2-h and tested in laboratory off-line. After this, the key biochemical variable was transformed from 2-h sampled data to 15 min sampled data (consistent with the number of auxiliary inputs data) in MATLAB using the "spline" interpolation function *interp1* (https://www.mathworks.com/help/matlab/ref/interp1.html). P was determined by the modified ninhydrin colorimetric method, i.e., 2 ml of the supernatant and 4 ml of the ninhydrin reagent were mixed and heated in boiling water for 20 min. The absorbance at 475 mm was measured by a spectrophotometer after cooling and obtained by checking the standard L-Lysine curve. These inputs represent the inputs 'x' in Equations (3)–(9). In addition,

the product concentration 'P' represents the output '*y*' in Equations (3)–(9). A non-linear mapping function is estimated using LSSVM between these inputs and output.

2. Ten batches were used for testing the modeling competence of the GWO-LSSVM method. The initial conditions between batches were set differently and the feeding strategy was also changed to enhance the differences between batches. The pressure of the fermentation tank was set to $0 \backsim 0.25$ MPa, the temperature of fermentation was adjusted to $0 \backsim 50\,^{\circ}\text{C} \pm 0.5\,^{\circ}\text{C}$ and the dissolved oxygen electrode was calibrated for the reference reading when the stirring motor was rotating at 400 rpm.

2.7. Performance Evaluation Metrics

To evaluate the accuracy of prediction model, statistical measures such as *Root mean square error (RMSE)*, *Mean absolute error (MAE)* and *Mean absolute percentage error (MAPE)* are used.

$$RMSE(V_{actual}, V_{pred}) = \sqrt{\frac{1}{T} \sum_{i=1}^{T} (V_{actual} - V_{pred})^2} \tag{22}$$

$$MAE(V_{actual}, V_{pred}) = \frac{1}{T} \sum_{i=1}^{T} |V_{actual} - V_{pred}| \tag{23}$$

$$MAPE(V_{actual}, V_{pred}) = \frac{1}{T} \frac{\sum_{i=1}^{T} |V_{actual} - V_{pred}|}{V_{actual}} \tag{24}$$

where V_{pred}, V_{actual} and T represent predicted, actual and total number of output values, respectively.

3. Results and Discussion

At first, the data are normalized in the range $[-1\ 1]$. Six batches are selected randomly to train the GWO-LSSVM model. Further two batches are selected for cross-validation step (off-line training and correction of model). After off-line validation of prediction model, the model is tested using optimized values of LSSVM parameters on two new batches of data to estimate the product concentration on-line. To find the optimum values of parameters of LSSVM, the parameters of GWO are adjusted as search agents $N_{max} = 30$, maximum iteration $I_{ter} = 100$, dimension $dim = 2$, lower bound $lb = [1, 0.001]$ and upper bound $ub = [10,000, 0.1]$.

3.1. GWO-LSSVM Results Analysis

LSSVM parameters are optimized using GWO optimization. The wolf's position in 2-dimensional search space (2D = number of LSSVM parameters to be optimized) represents LSSVM parameter values. The best position of alpha wolf denotes optimal parameters of LSSVM. *RMSE* is used as an objective function given in Equation (22). It defines the fitness of the solution (best position). The parameter values that minimize this objective function are selected as optimal solution which represent best position of *alpha* wolf. After selecting the optimal parameters obtained through GWO, LSSVM prediction model is established using these optimal parameters (GWO-LSSVM).

The proposed GWO-LSSVM is exploited as a prediction model and results are compared with PSO-LSSVM to show the effectiveness of GWO-LSSVM prediction model. PSO-LSSVM is the widely used prediction method in biological fermentation processes. For example, PSO-LSSVM is used to predict inulinase concentration in Pichia pastoris fermentation process [34]. Zhu [20] used PSO-LSSVM to measure key variables in Penicillin fermentation process. The actual and predicted curves by proposed GWO-LSSVM and PSO-LSSVM are shown in Figure 4a. The results clearly show that the GWO-LSSVM prediction model is capturing the future variation trends accurately. Furthermore, the difference between actual and predicted value is plotted in Figure 4b to visualize error more

clearly. We can see that the amplitude of error spikes for GWO-LSSVM is much lower than that of PSO-LSSVM. In addition, three statistical performance evaluation metrics are presented in Table 1. The *RMSE* value of GWO-LSSVM is approximately 61 % less than PSO-LSSVM. Similarly, the difference in the values of *MAE* and *MAPE* is approximately 77 % and 43 %, respectively. It is evident from the results of these statistical measures that GWO-LSSVM perform better than PSO-LSSVM. Although PSO has been very successful in many applications, it is more vulnerable to getting stuck at local minimum. According to NFL theorem, there is no single existing optimization algorithm that would be applicable in all kind of optimization problems. In each optimization problem, the optimization algorithm encounters a different unknown search space. The success of optimization algorithms depends upon the basic strategy of mathematical models that tries to avoid the local optimal solutions. In our case, the GWO-based LSSVM prediction model is more competent as compared to PSO-LSSVM. The prediction error is negligible and almost all error values lie in range $[-0.1\ 0.1]$, which is satisfying the controller requirement employed to control the product concentration. Thus, the proposed GWO-LSSVM prediction model is the best choice to be used as a process model in current non-linear MPC.

(**a**) Estimation

(**b**) Estimation error
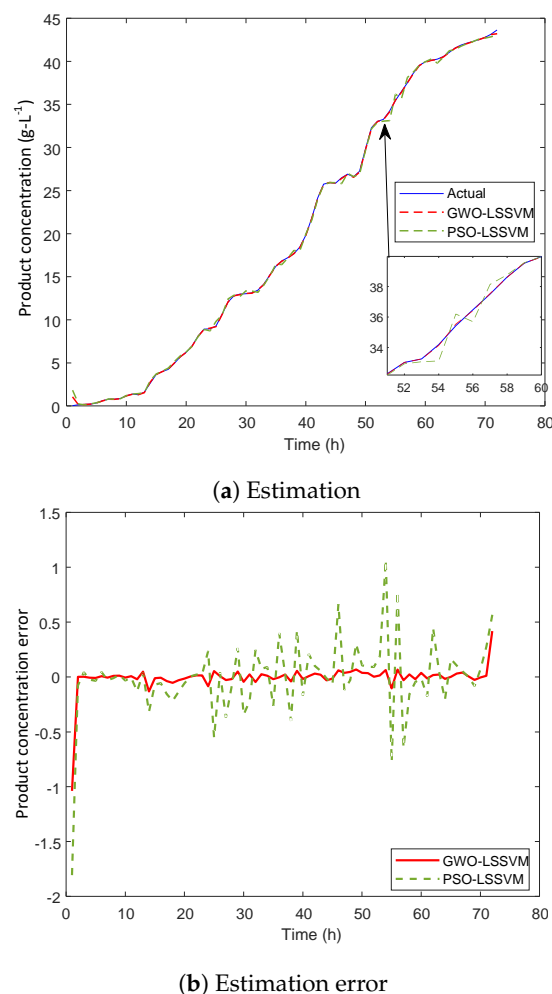
Figure 4. Product concentration prediction and error curve.

Table 1. RMSE, MAE and MAPE comparison.

Model	RMSE	MAE	MAPE
GWO-LSSVM	0.136918	0.047230	0.703616
PSO-LSSVM	0.355483	0.212182	1.244831

3.2. GWO-NMPC Results Analysis

The proposed non-linear MPC uses the search efficiency, local and global search ability of GWO for manipulating the input variables to optimize the future behavior of product concentration in L-Lysine's fermentation process. Dissolved Oxygen 'D_o' and 'pH' have a strong influence on product concentration. Thus, by monitoring the product concentration and according to its required optimum growth, D_o and pH are manipulated using agitation rate 'u_1', airflow rate 'u_2' and ammonia flow rate 'u_3'. Hence, three manipulated input variables are agitation rate 'u_1', airflow rate 'u_2', ammonia flow rate 'u_3'. The controlled output 'y' is product concentration. The initial substrate concentration and temperature values are 150 g L^{-1} and 32 °C, respectively. The objective function in Equation (1) is used as a fitness function of GWO, such as *RMSE* is used in the optimization problem of the LSSVM parameters. The error coefficient and input penalization coefficient are usually selected as a constant value. In this work, the value of error coefficient is selected as $Q_{out} = 1$ and input penalization coefficient is $R_{in} = 0.001$. The constraints on inputs, input increments and output are defined as:

$$\begin{cases} 316 \leq u_1 \leq 345 \\ 0.1 \leq u_2 \leq 1.6 \\ 51 \leq u_3 \leq 110 \end{cases} \tag{25}$$

$$\begin{cases} -15 \leq \Delta u_1 \leq +15 \\ -1 \leq \Delta u_2 \leq +1 \\ -30 \leq \Delta u_3 \leq +30 \end{cases} \tag{26}$$

$$0 \leq y \leq 45 \tag{27}$$

The initial inputs are selected as $U_{10} = 326$, $U_{20} = 0.1$ and $U_{30} = 88$. These initial input values are used to calculate the first future optimal input values according to Equation (21). The prediction and control horizon are defined as $N_{pred} = N_{con} = 1$. These parameter values are selected after extensive simulations. The increase in prediction and control horizon values have no significant difference on performance in our problem. Furthermore, our requirement is satisfied with these minimum horizon values. The computational cost increases with the increase in value of these horizons. However, these values can be increased that depends on the objective of controlling process to acquire desired response. GWO-LSSVM prediction model predicts the future output value 'y_{pred}' in defined objective function as in Equation (1). As GWO-LSSVM prediction model is non-linear, and objective function consists of non-linear constraints on manipulated inputs (agitation speed rate u_1, air flow rate u_2 and acceleration rate of ammonia flow u_3), so the optimization problem to solve this objective function for optimal future inputs (u_1, u_2 and u_3) becomes a non-convex and non-linear optimization problem. Now, GWO solves this optimization problem to find the future optimal values of these manipulated inputs at each sampling time. Here, a 3-dimensional search space is defined (3D = number of optimal future inputs to be determined). The wolves position in 3 *dim* space represents future optimal increment in input values (Δu_1, Δu_2 and Δu_3). The optimal future input increments that minimize the defined objective function (best *alpha* wolf position) are calculated in advance at each sampling instant for next sampling time. These optimal increments in inputs are used to determine the optimum values of future input rates using Equation (21). Then, these optimized input rates are applied to acquire the desired response. At next sampling point, the whole process is repeated for upcoming sampling instant to achieve the required objective.

3.2.1. Hypothetical Case Study

The proposed GWO-based predictive control scheme is compared with the PSO-based predictive control method. GWO-LSSVM is used as a prediction model, in proposed NMPC, to predict product concentration such as PSO-LSSVM is exploited as a prediction model in a predictive control method to

predict and control bacteria concentration [24]. Similarly, PSO is used in a study to solve the rolling optimization problem in a predictive control scheme to control substrate concentration and LSSVM is used as a prediction model [35].

Initially, a hypothetical reference signal with sharp periodic step changes is applied to validate the robustness and adaptability of GWO-NMPC for the first 44 hours, as shown in Figure 5. The simulation results are compared with PSO-NMPC. Although this kind of phenomenon is not realistic in fermentation process because fermentation is a slow time varying process, yet it proves the robustness and adaptability of the proposed control scheme. The corresponding manipulated input variables are shown in Figure 6a–c.
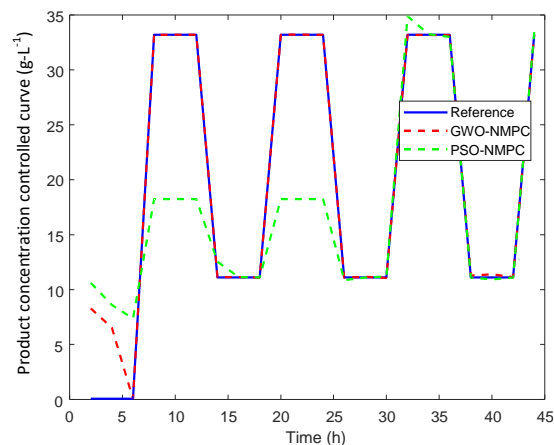
Figure 5. GWO-NMPC controlled product concentration output with hypothetical reference.
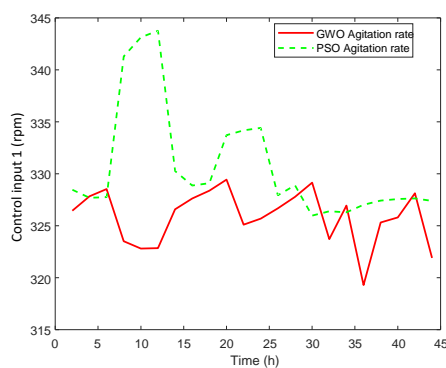
(**a**) Agitation rate u_1
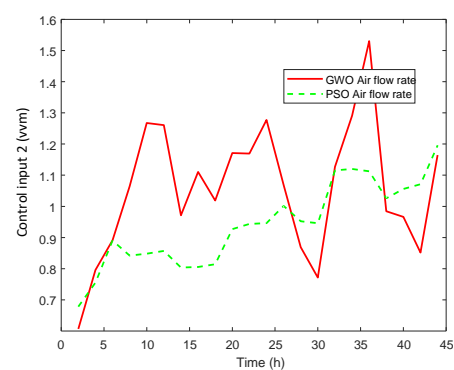
(**b**) Airflow rate u_2.
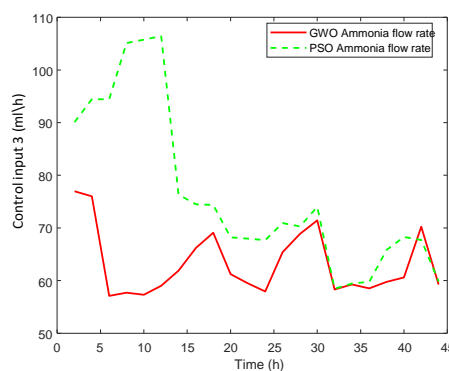
(**c**) Ammonia flow rate u_3

Figure 6. GWO-NMPC controlled inputs with hypothetical reference.

3.2.2. Real Case Study

An optimal selected trajectory as a reference is applied and the results are shown in Figure 7. The corresponding manipulated inputs are shown in Figures 8–10. We can see that the proposed method enforced the desired response and process follows this optimal trajectory accurately. GWO possesses fast convergence speed, excellent local, and global search ability. GWO is successful in handling non-linear constraints to solve a non-convex optimization problem for optimal future input values in real time and provides best global solution as compared to PSO. In addition, GWO-NMPC outperforms PSO-NMPC in terms of prediction accuracy, control precision and near-to-accurate tracking ability. The overall error in GWO-NMPC is negligible as compared to PSO-NMPC. The success of NMPC highly depends upon accuracy of process model used for output prediction. GWO-LSSVM encompasses precise dynamics and non-linear behavior between inputs and output of the L-Lysine fermentation process. This results in efficient performance of the proposed GWO-NMPC. Due to such optimal and controlled conditions, the osmotic stress or catabolic repression of bacteria is avoided successfully, and the final yield is increased by 25 % approximately. Therefore, it shows that the machine-learning-based prediction models and predictive control schemes are effective for control and optimization of complex non-linear industrial processes.
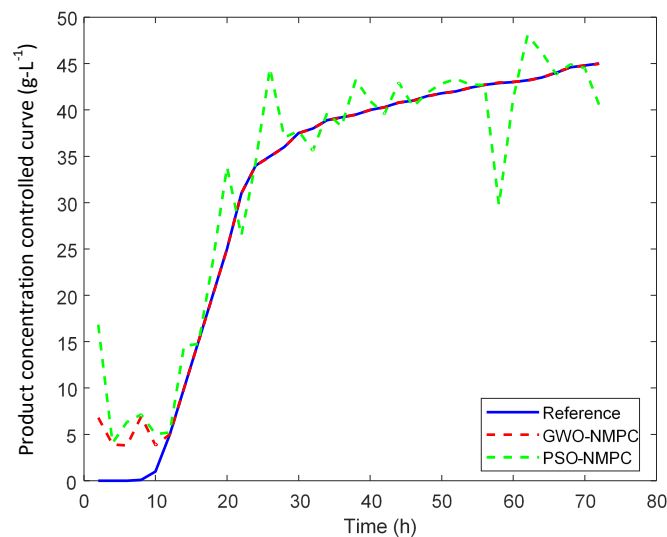

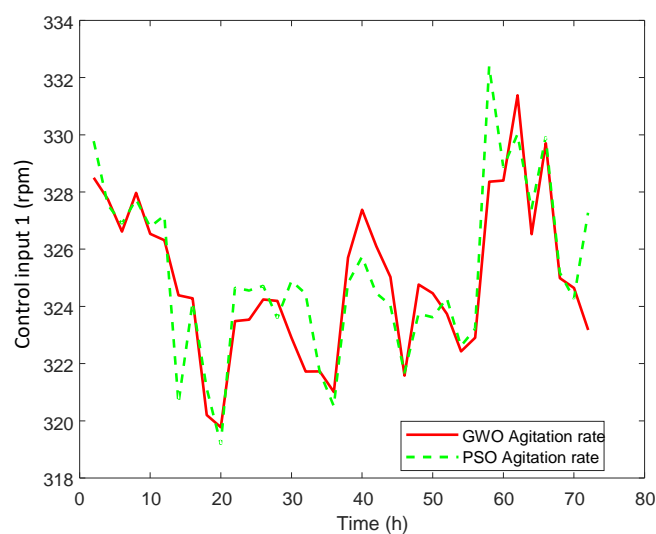**Figure 7.** GWO-NMPC controlled product concentration output with optimal reference.

Figure 8. GWO-NMPC controlled agitation rate u_1 with optimal reference.
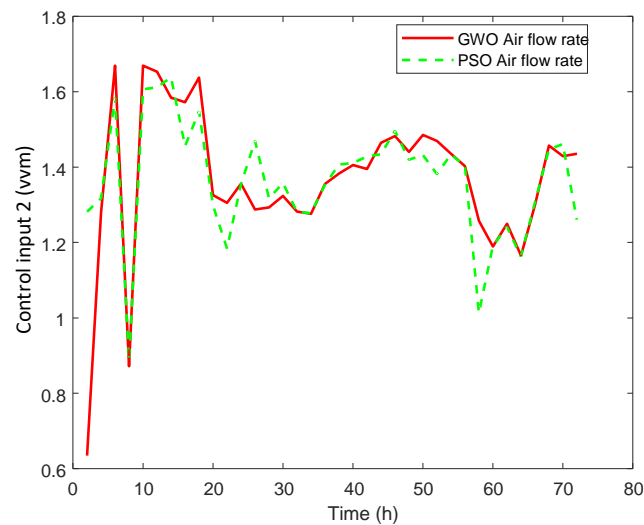
Figure 9. GWO-NMPC controlled airflow rate u_2 with optimal reference.
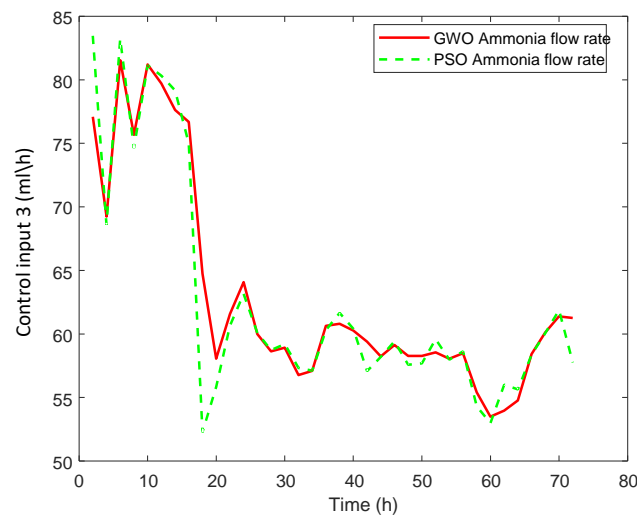

Figure 10. GWO-NMPC controlled ammonia flow rate u_3 with optimal reference.

4. Conclusions

A non-linear MPC by exploiting a machine learning-based prediction model is proposed to control product concentration in real time. LSSVM prediction model, which requires very few input and output samples for training, is deployed in NMPC. The traditional experience and trail-error-based method to select optimal value of LSSVM parameters is replaced by employing a novel metaheuristic GWO algorithm. Thus, real-time identification problem is solved by proposed GWO-LSSVM prediction model and it eliminates the need for an accurate kinetics mathematical model. Furthermore, to cope with the non-linear, non-convex and complex constraints-based optimization problem in NMPC, a gradient free approach is proposed. A novel GWO-based algorithm is established to compute optimum future input values by minimizing a cost function in real time. The proposed GWO-NMPC control scheme provides an efficient way to deal with complex, non-linear and dynamic systems. In future, we are interested to extend this work to control and optimize further key variables in fermentation process by designing multi-input and multi-output models.

Abbreviations

The following abbreviations are used in this manuscript:

NMPC	Non-linear Model Predictive Control
SVM	Support Vector Machine
LSSVM	Least-Square SVM
GWO	Grey-Wolf Optimization
PSO	Particle Swarm Optimization
ANN	Artificial Neural Network
QP	Quadratic Programming
ABC	Artificial Bee Colony
CS	Cuckoo Search
FFA	Firefly Algorithm
BA	Bat Algorithm
FPA	Flower Pollination Algorithm
GSA	Gravitational Search Algorithm
DE	Differential Evolution
EP	Evolutionary Programming
ES	Evolution strategy
NFL	No Free Lunch
GPC	Generalized Predictive Control
NP	Non-linear Programming
KKT	Karush–Kuhn–Tucker conditions
NP	Non-linear Programming
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ml	Milliliter
mm	Millimeter
rpm	Revolutions per minute
MPa	Megapascal
vvm	Volume per Unit per Minute

References

1. Eren, U.; Prach, A.; Koçer, B.B.; Raković, S.V.; Kayacan, E.; Açıkmeşe, B. Model predictive control in aerospace systems: Current state and opportunities. *J. Guid. Control Dyn.* **2017**, *40*, 1541–1566. [CrossRef]
2. Muhammad, D.; Ahmad, Z.; Aziz, N. Low density polyethylene tubular reactor control using state space model predictive control. *Chem. Eng. Commun.* **2019**, 1–17. [CrossRef]

3.   Wang, X.; Ratnaweera, H.; Holm, J.A.; Olsbu, V. Statistical monitoring and dynamic simulation of a wastewater treatment plant: A combined approach to achieve model predictive control. *J. Environ. Manag.* **2017**, *193*, 1–7. [CrossRef] [PubMed]

4.   Vazquez, S.; Rodriguez, J.; Rivera, M.; Franquelo, L.G.; Norambuena, M. Model predictive control for power converters and drives: Advances and trends. *IEEE Trans. Ind. Electron.* **2016**, *64*, 935–947. [CrossRef]

5.   Afram, A.; Janabi-Sharifi, F.; Fung, A.S.; Raahemifar, K. Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: A state of the art review and case study of a residential HVAC system. *Energy Build.* **2017**, *141*, 96–113. [CrossRef]

6.   Wang, D.; Shen, J.; Zhu, S.; Jiang, G. Model predictive control for chlorine dosing of drinking water treatment based on support vector machine model. *Desalin. Water Treat.* **2020**, *173*, 133–141. [CrossRef]

7.   Yokota, A.; Ikeda, M. *Amino Acid Fermentation*; Springer: Tokyo, Japan, 2017.

8.   Félix, F.K.d.C.; Letti, L.A.J.; Vinícius de Melo Pereira, G.; Bonfim, P.G.B.; Soccol, V.T.; Soccol, C.R. L-Lysine production improvement: A review of the state of the art and patent landscape focusing on strain development and fermentation technologies. *Crit. Rev. Biotechnol.* **2019**, *39*, 1031–1055. [CrossRef] [PubMed]

9.   Razak, M.A.; Viswanath, B. Optimization of fermentation upstream parameters and immobilization of Corynebacterium glutamicum MH 20-22 B cells to enhance the production of L-Lysine. *3 Biotech* **2015**, *5*, 531–540. [CrossRef]

10.   Gustavsson, R. *Development of Soft Sensors for Monitoring and Control of Bioprocesses*; Linköping University Electronic Press :Linköping, Sweden, 2018.

11.   Ahuja, K.; Pani, A.K. Software sensor development for product concentration monitoring in fed-batch fermentation process using dynamic principal component regression. In Proceedings of the 2018 International Conference on Soft-computing and Network Security (ICSNS), Coimbatore, India, 14–16 February 2018; pp. 1–6.

12.   Yuan, X.; Li, L.; Wang, Y. Nonlinear dynamic soft sensor modeling with supervised long short-term memory network. *IEEE Trans. Ind. Inform.* **2020**,*16*, 3168–3176.. [CrossRef]

13.   Grahovac, J.; Jokić, A.; Dodić, J.; Vučurović, D.; Dodić, S. Modelling and prediction of bioethanol production from intermediates and byproduct of sugar beet processing using neural networks. *Renew. Energy* **2016**, *85*, 953–958. [CrossRef]

14.   Datta, A.; Augustin, M.; Gupta, N.; Viswamurthy, S.; Gaddikeri, K.M.; Sundaram, R. Impact Localization and Severity Estimation on Composite Structure Using Fiber Bragg Grating Sensors by Least Square Support Vector Regression. *IEEE Sens. J.* **2019**, *19*, 4463–4470. [CrossRef]

15.   Wang, G.; Xu, B.; Jiang, W. SVM modeling for glutamic acid fermentation process. In Proceedings of the 2016 Chinese Control and Decision Conference (CCDC),Yinchuan, China, 28–30 May2016; pp. 5551–5555.

16.   Zhang, Y.; Le, J.; Liao, X.; Zheng, F.; Li, Y. A novel combination forecasting model for wind power integrating least square support vector machine, deep belief network, singular spectrum analysis and locality-sensitive hashing. *Energy* **2019**, *168*, 558–572. [CrossRef]

17.   Luo, C.; Huang, C.; Cao, J.; Lu, J.; Huang, W.; Guo, J.; Wei, Y. Short-term traffic flow prediction based on least square support vector machine with hybrid optimization algorithm. *Neural Process. Lett.* **2019**, *50*, 2305–2322. [CrossRef]

18.   Robles-Rodriguez, C.E.; Bideaux, C.; Roux, G.; Molina-Jouve, C.; Aceves-Lara, C.A. Soft-sensors for lipid fermentation variables based on PSO Support Vector Machine (PSO-SVM). In *Distributed Computing and Artificial Intelligence, Proceedings of the 13th International Conference, Salamanca, Spain, 28–30 March 2020*; Springer: Berlin, Germany, 2016; pp. 175–183.

19.   Xue, X.; Xiao, M. Deformation evaluation on surrounding rocks of underground caverns based on PSO-LSSVM. *Tunn. Undergr. Space Technol.* **2017**, *69*, 171–181. [CrossRef]

20.   Zhu, X.; Rehman, K.U.; Wang, B.; Shahzad, M. Modern Soft-Sensing Modeling Methods for Fermentation Processes. *Sensors* **2020**, *20*, 1771. [CrossRef]

21.   Saad, A.E.H.; Dong, Z.; Karimi, M. A comparative study on recently-introduced nature-based global optimization methods in complex mechanical system design. *Algorithms* **2017**, *10*, 120. [CrossRef]

22.   Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]

23.   Wolpert, D.H.; Macready, W.G. No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.* **1997**, *1*, 67–82. [CrossRef]

24. Zhu, X.; Zhu, Z. The generalized predictive control of bacteria concentration in marine lysozyme fermentation process. *Food Sci. Nutr.* **2018**, *6*, 2459–2465. [CrossRef]

25. Nisha, M.G.; Prince, M.J.R.; Jones, A.J. Neural Network Predictive Control of Systems with Faster Dynamics using PSO. In Proceedings of the 2019 International Conference on Recent Advances in Energy-Efficient Computing and Communication (ICRAECC), Nagercoil, India, 7–8 March 2019; pp. 1–4.

26. Ait Sahed, O.; Kara, K.; Hadjili, M.L. Constrained fuzzy predictive control using particle swarm optimization. *Appl. Comput. Intell. Soft Comput.* **2015**. [CrossRef]

27. Su, T.J.; Tsou, T.Y.; Vu, H.Q.; Shyr, W.J. Model Predictive Control Design Based on Particle Swarm Optimization. *J. Converg. Inf. Technol.* **2015**, *10*, 70.

28. Suthar, S.; Vishwakarma, D. A Fast Converging MPPT Control Technique (GWO) for PV Systems Adaptive to Fast Changing Irradiation and Partial Shading Conditions. Available online: https://d1wqtxts1xzle7.cloudfront.net/60428554/IRJET-V6I650220190829-75962-1sorde7.pdf (accessed on 8 June 2020).

29. Chen, L.; Du, S.; He, Y.; Liang, M.; Xu, D. Robust model predictive control for greenhouse temperature based on particle swarm optimization. *Inf. Process. Agric.* **2018**, *5*, 329–338. [CrossRef]

30. Kouvaritakis, B.; Cannon, M. *Model Predictive Control*; Springer International Publishing: Cham, Switzerland, 2016.

31. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: New York, NY, USA, 2013.

32. Suykens, J.A.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [CrossRef]

33. Azimi, H.; Bonakdari, H.; Ebtehaj, I. Design of radial basis function-based support vector regression in predicting the discharge coefficient of a side weir in a trapezoidal channel. *Appl. Water Sci.* **2019**, *9*, 78. [CrossRef]

34. Wang, X.; Guo, T.; Hao, W.; Guo, Q. Predicting Model based on LS-SVM for Inulinase Concentration during Pichia Pastoris' Fermentation Process. In Proceedings of the 2019 Chinese Control Conference (CCC), Guangzhou, China, 27–30 July 2019; pp. 1531–1536.

35. Huang, L.; Wang, Z.; Ji, X. LS-SVM Generalized Predictive Control Based on PSO and Its Application of Fermentation Control. In *Proceedings of the 2015 Chinese Intelligent Systems Conference*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 605–613.