METHODOLOGY ARTICLE   **Open Access**

Soft - sensing modeling based on ABC - MLSSVM inversion for marine low - temperature alkaline protease MP fermentation process

Bo Wang[*] , Meifang Yu, Xianglin Zhu and Li Zhu

Abstract

Background: Aiming at the characteristics of nonlinear, multi-parameter, strong coupling and difficulty in direct on-line measurement of key biological parameters of marine low-temperature protease fermentation process, a soft-sensing modeling method based on artificial bee colony (ABC) and multiple least squares support vector machine (MLSSVM) inversion for marine protease fermentation process is proposed.

Methods: Firstly, based on the material balance and the characteristics of the fermentation process, the dynamic "grey box" model of the fed-batch fermentation process of marine protease is established. The inverse model is constructed by analyzing the inverse system existence and introducing the characteristic information of the fermentation process. Then, the inverse model is identified off-line using MLSSVM. Meanwhile, in order to reduce the model error, the ABC algorithm is used to correct the inverse model. Finally, the corrected inverse model is connected in series to the marine alkaline protease MP fermentation process to form a composite pseudo-linear system, thus, real-time on-line prediction of key biological parameters in fermentation process can be realized.

Results: Taking the alkaline protease MP fermentation process as an example, the simulation results demonstrate that the soft-sensing modeling method can solve the real-time prediction problem of key biological parameters in the fermentation process on-line, and has higher accuracy and generalization ability than the traditional soft-sensing method of support vector machine.

Conclusions: The research provides a new method for soft-sensing modeling of key biological parameters in fermentation process, which can be extended to soft-sensing modeling of general nonlinear systems.

Keywords: Marine alkaline protease MP, Material balance, Inverse system, Support vector machine, Soft-sensing

Background

Marine alkaline protease MP is a fermentation enzyme that adapts well to the low-temperature environment [1]. In addition to a wide range of pH, the MP enjoys a high activity at room temperature, and becomes less active with the decrease of temperature [2, 3]. This emerging industrial enzyme brings new vitality and opportunities to such fields as medicine, food, enzyme industry, national defense and so on, and greatly broadens application regions of protease [4]. The marine low-temperature alkaline

protease MP fermentation process is a dynamic process with high nonlinearity and strong coupling effect. Like general nonlinear systems, fermentation has time-varying, correlated and uncertain parameters [5]. Therefore, it is very difficult to directly measure the key biological parameters in the fermentation process [6]. Currently, these parameters can only be obtained through regular sampling, offline analysis and lab test. The current method has a poor real-time performance, and increases the bacteria exposure of the samples, hindering the advanced control of fermentation. This calls for a strategy that timely acquires the state of key biological parameters in the fermentation process. Without the state information, it is impossible to

* Correspondence: wangbo@ujs.edu.cn
School of Electrical and Information Engineering, JiangSu University, Zhenjiang 212013, Jiangsu, China

achieve dynamic and optimal control of the fermentation process, which boosts the biomass density and productivity of enzyme.

The inverse system method provides a good solution to the soft-sensing modelling of the fermentation process. This method boasts strict theoretical bases and clear physical meanings. Coupled with learning algorithms (e.g. neural network (NN) and support vector machine (SVM)), the inverse system method can complete soft-sensing of nonlinear systems, which are difficult to be modelled accurately [7, 8]. Suffice it to say that the inverse system method greatly facilitates the soft-sensing modelling of highly nonlinear systems in engineering practices. However, the inverse system method faces two problems in soft-sensing of the marine low-temperature alkaline protease MP fermentation process. On the one hand, the mathematical model of the controlled object and the system parameters of the model must be known before using the inverse system method. It is no easy task to obtain either information from the highly nonlinear and strongly coupled fermentation process. On the other hand, the inverse system of the original system must be established before using the inverse system method. In other words, the inverse system should be expressed mathematically in advance (that is, to derive a mathematical expression that can be used to describe the inverse system) [8]. To solve the problems, the literature [9] proposes a neural network inverse system method, which integrates intelligent control with the inverse system method. The inverse system is approximated by the neural network in this literature. The method was successfully applied to the soft-sensing of erythromycin fermentation, creating a "gray-box" model of the fermentation process. Nevertheless, the "gray-box" model is a simplified model based on the Monod equation, which ignores many important nonlinear factors in the actual process of erythromycin fermentation. Besides, the neural network, inspired by the asymptotic theory, is based on the unrealistic assumption that the number of samples is infinite, but the number of samples in the actual problem is often limited, especially the strong coupling, large lag complex nonlinear system as the marine alkaline protease MP fermentation process, it is extremely difficult to obtain accurate sample data. Therefore, in the case of small samples, the research of inverse soft-sensing methods suitable for the marine alkaline protease MP fermentation process and easy to implement in engineering has become the key problem to be solved urgently in the marine low-temperature alkaline protease MP fermentation process.

Considering the limited number of samples in actual fermentation, this paper attempts to design an easy-to-use inverse system method for soft-sensing modeling of the marine alkaline protease MP fermentation process. Firstly, a "gray-box" dynamic model was established for the the marine low-temperature alkaline protease MP fermentation process, according to material balance and features of that process. Secondly, the existing inverse system was analyzed, and the design of extended inverse model was introduced. Thirdly, the offline identification of MLSSVM and online optimization of ABC were combined to develop the extended inverse model based on ABC-MLSVM, and the extended inverse model was connected in series after the primary fermentation process, serving as the soft-sensing model that predicts key biological parameters online in real time. Fourthly, the effectiveness of the soft-sensing modelling method was verified through a simulation of the MP fermentation in lab; the simulation results show that the method can effectively predict the key biological parameters of the marine low-temperature alkaline protease MP fermentation process online, and outperform the traditional least square support vector machine (LS-SVM) soft-sensing modeling method in prediction accuracy.

Methods
Dynamic model of fermentation process
In this paper, Taking the fermentation process as the object, this paper assumes that both cell concentration and the protease concentration are zero. The fermentation states (concentration of each substance) were taken as dependent variables of differential equation, while time t was taken as an independent variable or separate variable. Then, the "grey box" dynamic model could be described by the material balance equations of various substances (mycelia, restrictive substrate, protease, oxygen, H+, etc.) [10], as shown in Eq. (1):

$$\text{Variable quantity(Target substance per unit of time)}$$
$$= \text{influx(Target substance per unit of time)}$$
$$- \text{outflow} + \text{formation amount}$$

(1)

The construction process of the dynamic model of the ash box is as follows:

Volume change equilibrium equation
During the fermentation process, culture medium should be added at a rational rate to supplement the nutrients and increase the protease yield. The culture medium mainly consists of carbon source, nitrogen source, inorganic salt, growth factor and enzyme-producing promoter. Through preliminary experiments, this paper selects maize flour hydrolysate as carbon source, soybean meal hydrolysate as nitrogen source, ammonia sulfate $((NH_4)_2SO4)$ as inorganic salt, malt extract as growth factor and Polysorbate 80 (Tween-80) as surfactant. The volume (V) of the fermentation broth changes with the addition of nutrients and enzyme-producing promoter. The equilibrium equation is as follows:

Wang *et al. BMC Biotechnology*   (2020) 20:9

Page 3 of 13

$$\frac{dV}{dt} = f_{mh} + f_s + f_a + f_m + f_{tw} \tag{2}$$

where: V is fermentation broth volume, f_{mh}, f_s, f_a, f_m and f_{tw} are respectively the flow rate of maize flour hydrolysate, soyabean cake meal hy-drolysate, $(NH_4)_2SO_4$, malt extract and enzyme-producing promoter (Tween-80).

Cell growth kinetics equation

The previous studies have found that, the growth of the enzyme producing strain has a maximum concentration, i.e. a saturation point, which could be reached if the initial sugar concentration is on suitable levels. The time to reach the saturation point varies with the initial sugar concentration. The higher the initial sugar concentration, the slower the cells grow, that is, the substrate concentration inhibits the cell growth. Considering the deviation of Monod equation-based description, the logistic equation was employed to depict the growth law of the cells, in the light of the volume change in fed-batch fermentation, and the volume change during fed-batch fermentation is taken into account. The growth kinetics model of cell is as follows:

$$\frac{dX}{dt} = \mu X - \frac{X}{V}\frac{dV}{dt} \tag{3}$$

where: μ is the specific growth rate of somatic cells, X is cell concentration.

Substrate consumption equation

The substrate consumption of marine low-temperature alkaline protease MP was modelled based on the material balance. The effect of additive carbon source (maize flour hydrolysate) was considered in the model, because the carbon source, as the only restrictive substrate, is consumed rapidly in large quantities. The model is expressed as follows:

$$\frac{dS}{dt} = -\nu X + \frac{S_{mh}}{V} - \frac{S}{V}\frac{dV}{dt} \tag{4}$$

where: S is the substrate concentration, ν is the specific consumption rate of substrate (h^{-1}), S_{mh} is the maize flour hydrolysate flow rate.

Protease synthesis kinetics

The model of fermentation enzyme production is partial growth coupled type (It belongs to extracellular enzyme, and its synthesis regulation is affected by many mechanisms), high concentration substrate can obviously inhibit the secretion of protease while maintaining low carbon source concentration is beneficial to the secretion of protease MP. At the same time, the hydrolysis of alkaline protease MP also has a certain effect on protease MP itself. On this basis, Tween-80, growth factor

and the hydrolysis rate which have influence on the fermentation process are introduced into the protease synthesis kinetics, the model is expressed as follows:

$$\frac{dE}{dt} = \rho X - KP + \frac{K_m}{V}f_m + \frac{K_{tw}}{V}f_{tw} - \frac{E}{V}\frac{dV}{dt} \tag{5}$$

where: E is protease content (%), ρ is the specific growth rate of protease, K is hydrolysis constants of protease, and K_m, K_{tw} are inhibition constants.

Dissolved oxygen concentration (DO) variation model

The MP fermentation is aerobic, i.e. oxygen is involved in cell growth and protease synthesis. DO must be controlled in a suitable range. In fact, the oxygen demand constantly changes through the cell growth, because cell concentration and cell respiration intensity change from stage to stage. Based on the varying oxygen demands, the DO in the fermentation broth must be regulated in real time. According to the aerobic features of MP fermentation and the effect of bioreactor size on DO level in culture medium, the oxygen volumetric mass transfer coefficient of the bioreactor was introduced to the DO concentration equilibrium equation. The equilibrium equation is as follows:

$$\frac{dC_L}{dt} = -\eta X + K_{La}\left(C_L^* - C_L\right) - \frac{C_L}{V}\frac{dV}{dt} \tag{6}$$

where: C_L is the DO concentration (mol/L), K_{La} is oxygen volume mass transfer coefficient in bioreactor (s^{-1}), C_L^* is saturation concentration of oxygen dissolved in fermentation liquid phase (mol/L), and η is the specific consumption rate of oxygen.

PH dynamic change model

During the fermentation, the enzyme producing strain favors an alkaline environment, with the optimal pH range of 9.0~10.0. Any change of pH in fermentation broth will exert a huge impact on the fermentation of the MP. Excessively high or low pH values will slow down the strain growth and the formation of protease, weakening the enzyme activity. Therefore, the pH of fermentation broth was regulated within the optimal range by the flow of nutrients through the fermentation process, so that it can be maintained in the optimum range. The pH equilibrium equation of fermentation broth is expressed as:

$$\frac{d[H]^+}{dt} = \gamma X - \frac{[H]^+}{V}\frac{dV}{dt} + \frac{S_s f_s - S_{mh}f_{mh} - S_m f_m - S_{tw}f_{tw}}{V} \tag{7}$$

where:$[H]^+$ is the hydrogen ion concentration in fermentation broth (used to characterize the pH of fermentation broth), f_{mh}, f_s, f_m and f_{tw} are respectively flow rate of

Wang *et al. BMC Biotechnology*   (2020) 20:9

Page 4 of 13

maize flour hydrolysate, soyabean cake meal hy-drolysate, malt extract and Tween-80. S_{mh}, S_s, S_m, S_{tw} are respectively the liquid concentrations of maize flour hy-drolysate, soyabean cake meal hy-drolysate, malt extract and Tween-80, γ is the specific consumption of [H+].

Through the above analysis, the "gray-box" dynamic model of the marine low-temperature alkaline protease MP fermentation process can be expressed as:

$$
\begin{cases}
\dot{x}_1 = \mu x_1 - \dfrac{x_1}{x_6}\displaystyle\sum_{i=1}^{5} u_i \\[2ex]
\dot{x}_2 = -\nu x_1 + \dfrac{s_1 u_1}{x_6} - \dfrac{x_2}{x_6}\displaystyle\sum_{i=1}^{5} u_i \\[2ex]
\dot{x}_3 = \rho x_1 - s_2 x_3 + \dfrac{s_3 u_4}{x_6} + \dfrac{s_4 u_5}{x_6} - \dfrac{x_3}{x_6}\displaystyle\sum_{i=1}^{5} u_i \\[2ex]
\dot{x}_4 = -\eta x_1 - s_5 x_4 + s_6 - \dfrac{x_4}{x_6}\displaystyle\sum_{i=1}^{5} u_i \\[2ex]
\dot{x}_5 = \gamma x_1 + \dfrac{s_7 u_2 - s_1 u_1 - s_8 u_4 - s_9 u_5}{x_6} - \dfrac{x_5}{x_6}\displaystyle\sum_{i=1}^{5} u_i \\[2ex]
\dot{x}_6 = \displaystyle\sum_{i=1}^{5} u_i = u_1 + u_2 + u_3 + u_4 + u_5
\end{cases}
\tag{8}
$$

where: $\mathbf{x} = [x_1, x_2, x_3, x_4, x_5, x_6]^T = [X, S, E, C_L, [H]^+, V]^T$ represent the status vector, $\mathbf{u} = [u_1, u_2, u_3, u_4, u_5]^T = [f_{mh}, f_s, f_a, f_m, f_{tw}]^T$ is the input vector, $\mu(\mathbf{x})$, $\nu(\mathbf{x})$, $\rho(\mathbf{x})$, $\eta(\mathbf{x})$, $\gamma(\mathbf{x})$ are the analytical functions of the respective status variables $\mathbf{x}$, S_i ($i = 1, 2, \cdots 9$) are all constants other than zero and represent respectively the liquid feeding concentration of maize flour hydrolysate, hydrolysis constants of protease, inhibition constant, gas saturated oxygen concentration, $C_L^* * K_{La}$, the liquid feeding concentration of soyabean cake meal hydrolysate, $(NH_4)_2SO_4$ and Tween-80.

Reversibility analysis

The marine low-temperature alkaline protease MP fermentation process is shown in Fig. 1. $u_1 \sim u_5$ are input parameters, $x_1 \sim x_6$ are the six process parameters of fermentation process. The process parameters x_4, x_5, x_6 are directly measurable parameters and x_1, x_2, x_3 are key
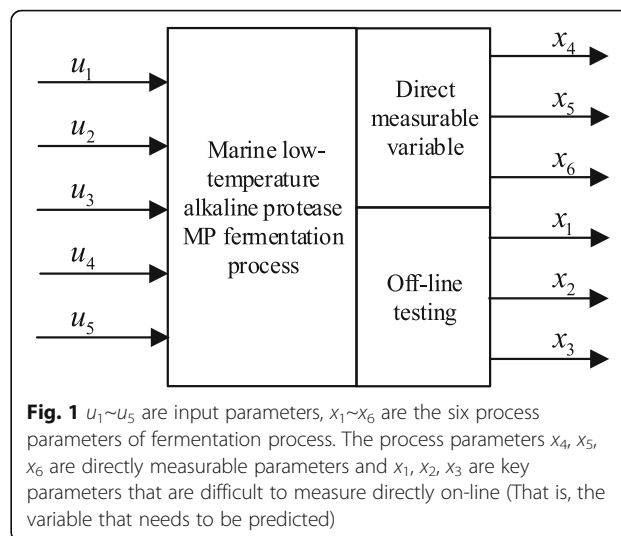

Fig. 1 $u_1 \sim u_5$ are input parameters, $x_1 \sim x_6$ are the six process parameters of fermentation process. The process parameters x_4, x_5, x_6 are directly measurable parameters and x_1, x_2, x_3 are key parameters that are difficult to measure directly on-line (That is, the variable that needs to be predicted)

parameters that are difficult to measure directly on-line (That is, the variable that needs to be predicted).

In order to predict the non-direct measurable key biological parameters x_1, x_2, x_3, a virtual subsystem was assumed to exist in the marine low-temperature alkaline protease MP fermentation process, including three non-directly measureable inputs x_1, x_2, x_3, three directly measurable outputs x_4, x_5, x_6 and five variables $u_1 \sim u_5$. This virtual subsystem is regarded as a "virtual sensor" [11] of the marine low-temperature alkaline protease MP fermentation process. The soft-sensing of x_1, x_2, x_3 can be realized through the following steps: solve the inverse model of the virtual subsystem; take the model as a dynamic compensator in series with the virtual subsystem, forming a composite system [12]; reproduce the inputs of the "virtual sensor" based on the outputs of the composite system, as shown in Fig. 2.

In order to achieve the soft-sensing of key biological parameters x_1, x_2, x_3, the reversibility analysis of the "virtual sensor" must be carried out, and the inverse system model should be solved.
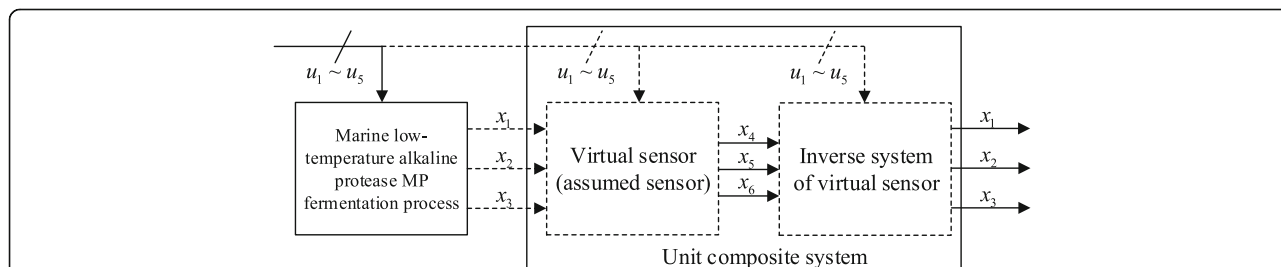
Fig. 2 This "virtual subsystem" is regarded as a "virtual sensor" in the marine low-temperature alkaline protease MP fermentation process. It can be used as a dynamic compensator in series with the "virtual subsystem", a unit compound system can be formed. Its the input and output present an identity mapping relationship, that is, the output of the composite system can completely reproduce the input of the original "virtual sensor", and then the soft-sensing of x_1, x_2, x_3 can be realized

Lemma *Necessary and sufficient conditions for system Σ reversibility in some field of point (x_0, u_0): The system meets* $\operatorname{rank}(\partial z_m^{\mathrm{T}}/\partial \hat{\mathbf{x}}^{\mathrm{T}}) = r_m = l$, *l is the dimension number of non-direct measurable variable.*

The reversibility of "virtual sensor" is analyzed by ***Interactor*** algorithm: The direct measurable variable $\mathbf{z} = [z_1, z_2, z_3] = [x_4, x_5, x_6]$ are derived by using the modeling algorithm, and all-order derivatives $\dot{z}_i, \ddot{z}_i, \cdots, z_i^{(k_i)}$ $(i = 1, 2, 3)$ can be obtained, and then the independent derivative information of the function is selected to form the vector $\mathbf{Z}_m$,as shown in Eq. 9:

$$
\begin{cases}
\dot{x}_4 = -\eta x_1 - s_5 x_4 + s_6 - \dfrac{x_4}{x_6} \sum_{i=1}^{5} u_i \\[2mm]
\ddot{x}_4 = g_1(\mathbf{x}, \mathbf{u}) + g_2(x_4, x_5, x_6, \mathbf{u}, \dot{\mathbf{u}}) \\[2mm]
\dot{x}_5 = \gamma x_1 + \dfrac{s_7 u_2 - s_1 u_1 - s_8 u_4 - s_9 u_5}{x_6} - \dfrac{x_5}{x_6} \sum_{i=1}^{5} u_i
\end{cases}
\tag{9}
$$

where:

$$
\begin{cases}
\begin{aligned}
g_1(\mathbf{x}, \mathbf{u}) =\ & \left(\dfrac{\partial \eta}{\partial x_1} x_1 + \dfrac{\partial \eta}{\partial x_2} x_2 + \dfrac{\partial \eta}{\partial x_3} x_3 + \dfrac{\partial \eta}{\partial x_4} x_4 + \dfrac{\partial \eta}{\partial x_5} x_5 \right) \dfrac{x_1}{x_6} \sum_{i=1}^{5} u_i \\
& + s_5 \eta x_1 - \left(\dfrac{\partial \eta}{\partial x_1} \mu - \dfrac{\partial \eta}{\partial x_2} \nu + \dfrac{\partial \eta}{\partial x_3} \rho - \dfrac{\partial \eta}{\partial x_4} \eta + \dfrac{\partial \eta}{\partial x_5} \gamma \right) x_1^2 \\
& + \dfrac{\partial \eta}{\partial x_3} s_2 x_1 x_3 - \dfrac{\partial \eta}{\partial x_3} \dfrac{x_1}{x_6} (s_3 u_4 + s_4 u_5) - \eta \mu x_1 - \dfrac{\partial \eta}{\partial x_4} s_6 x_1 \\
& + \dfrac{\partial \eta}{\partial x_5} \dfrac{x_1}{x_6} s_8 u_4 - \dfrac{\partial \eta}{\partial x_5} \dfrac{x_1}{x_6} s_7 u_2 - s_1 \left(\dfrac{\partial \eta}{\partial x_2} - \dfrac{\partial \eta}{\partial x_5} \right) \dfrac{x_1}{x_6} u_1 \\
& + \dfrac{\partial \eta}{\partial x_4} s_5 x_1 x_4 + \dfrac{2\eta x_1}{x_6} \sum_{i=1}^{5} u_i + \dfrac{\partial \eta}{\partial x_5} \dfrac{x_1}{x_6} s_9 u_5
\end{aligned} \\[4mm]
\begin{aligned}
g_2(x_4, x_5, x_6, \mathbf{u}, \dot{\mathbf{u}}) =\ & 2 \dfrac{s_5 x_4}{x_6} \sum_{i=1}^{5} u_i - \dfrac{s_6}{x_6} \sum_{i=1}^{5} u_i + s_5^2 x_4 - s_5 s_6 \\
& + 2 \dfrac{x_4}{x_6^2} \left(\sum_{i=1}^{5} u_i \right)^2 - \dfrac{x_4}{x_6} \sum_{i=1}^{5} \dot{u}_i
\end{aligned}
\end{cases}
\tag{10}
$$

According to Eq. (9), $\partial \ddot{z}_2/\partial x_i = \partial g_1(x)/\partial x_i, i = 1, 2, 3,.$ $m = 3$is obtained, and the Jacobian Matrix $J = \partial (\ddot{z}_1, \dot{z}_1, \dot{z}_2)^T/\partial(x_1, x_2, x_3)$ is further solved:

$$
J = \begin{bmatrix}
\dfrac{\partial \ddot{z}_1}{\partial x_1} & \dfrac{\partial \ddot{z}_1}{\partial x_2} & \dfrac{\partial \ddot{z}_1}{\partial x_3} \\[2mm]
\dfrac{\partial \dot{z}_1}{\partial x_1} & \dfrac{\partial \dot{z}_1}{\partial x_2} & \dfrac{\partial \dot{z}_1}{\partial x_3} \\[2mm]
\dfrac{\partial \dot{z}_2}{\partial x_1} & \dfrac{\partial \dot{z}_2}{\partial x_2} & \dfrac{\partial \dot{z}_2}{\partial x_3}
\end{bmatrix}
= \begin{bmatrix}
\dfrac{\partial g_1(\mathbf{x}, \mathbf{u})}{\partial x_1} & \dfrac{\partial g_1(\mathbf{x}, \mathbf{u})}{\partial x_2} & \dfrac{\partial g_1(\mathbf{x}, \mathbf{u})}{\partial x_3} \\[2mm]
-\dfrac{\partial \eta}{\partial x_1} x_1 - \eta & -\dfrac{\partial \eta}{\partial x_2} x_1 & -\dfrac{\partial \eta}{\partial x_3} x_1 \\[2mm]
\dfrac{\partial \gamma}{\partial x_1} x_1 + \gamma & \dfrac{\partial \gamma}{\partial x_2} x_1 & \dfrac{\partial \gamma}{\partial x_3} x_1
\end{bmatrix}
\tag{11}
$$

After the transformation of the elementary row of Eq. (11) and obtain the following:

$$
\tilde{\mathbf{J}} = \begin{bmatrix}
g_7(\mathbf{x}, \mathbf{u}) & 0 & 0 \\[2mm]
g_3(\mathbf{x}, \mathbf{u}) & g_4(\mathbf{x}, \mathbf{u}) & 0 \\[2mm]
\dfrac{\partial \gamma}{\partial x_1} x_1 + \gamma & \dfrac{\partial \gamma}{\partial x_2} x_1 & \dfrac{\partial \gamma}{\partial x_3} x_1
\end{bmatrix}
\tag{12}
$$

where: $g_3(\mathbf{x}, \mathbf{u}) = \left[\left(\dfrac{\partial \gamma}{\partial x_1} x_1 + \gamma \right) \dfrac{\partial \eta}{\partial x_3} \right] \Big/ \dfrac{\partial \gamma}{\partial x_3} - \dfrac{\partial \eta}{\partial x_1} x_1 - \eta$;

$g_4(\mathbf{x}, \mathbf{u}) = \left(x_1 \dfrac{\partial \gamma}{\partial x_2} \dfrac{\partial \eta}{\partial x_3} \right) \Big/ \dfrac{\partial \gamma}{\partial x_3} - \dfrac{\partial \eta}{\partial x_2} x_2$;

$g_5(\mathbf{x}, \mathbf{u}) = \dfrac{\partial g_1(\mathbf{x}, \mathbf{u})}{\partial x_1} - \left[\left(\dfrac{\partial \nu}{\partial x_1} x_1 + \gamma \right) \dfrac{\partial g_1(\mathbf{x}, \mathbf{u})}{\partial x_3} \right] \Big/ \dfrac{\partial \nu}{\partial x_3}$;

$g_6(\mathbf{x}, \mathbf{u}) = \dfrac{\partial g_1(\mathbf{x}, \mathbf{u})}{\partial x_2} - \left(\dfrac{\partial \gamma}{\partial x_2} \dfrac{\partial g_1(\mathbf{x}, \mathbf{u})}{\partial x_3} \right) \Big/ \dfrac{\partial \gamma}{\partial x_3}$;

$$
g_7(\mathbf{x}, \mathbf{u}) = g_5(\mathbf{x}, \mathbf{u}) - \dfrac{g_6(\mathbf{x}, \mathbf{u})}{g_4(\mathbf{x}, \mathbf{u})} g_3(\mathbf{x}, \mathbf{u}).
$$

If $\det(\tilde{J}) = g_7(\mathbf{x}, \mathbf{u}) \cdot g_4(\mathbf{x}, \mathbf{u}) \cdot \dfrac{\partial \gamma}{\partial x_3} x_1 \neq 0$ in the entire real vector space, it can be known that $\mathbf{J} = \partial \mathbf{Z}_m^{\mathrm{T}}/\partial \hat{\mathbf{x}}^{\mathrm{T}} = \partial(\ddot{z}_1, \dot{z}_1, z_2)/\partial(x_1, x_2, x_3) = 3$ from **Lemma**, and it meet the system reversibility condition, that is, the system is globally reversible. However, for $\det(\tilde{J}) \neq 0$, it is quite difficult to guarantee that it satisfies the non-zero conditions everywhere in the entire real vector space ***R***.

Considering the above situation and the current operation state of the marine low temperature alkaline protease MP fermentation process, it is assumed that $\det(\tilde{J}) \neq 0$, a small work area of the fermentation process within the real vector space R, satisfies the reversibility condition of the "virtual sensor". Then, an inverse soft-sensing model is constructed based on ABC-MLSVM. The assumption is verified against actual analysis results.

Suppose the system satisfies the reversibility condition for the work area. Then, the inverse system of the virtual sensor of he marine low-temperature alkaline protease MP fermentation process is established based on the inverse function theorem, using Eqs. (8)~(9):

$$
\hat{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \phi_1(x_4, x_5, x_6, \dot{x}_4, \ddot{x}_4, \dot{x}_5, \mathbf{u}, \dot{\mathbf{u}}) \\ \phi_2(x_4, x_5, x_6, \dot{x}_4, \ddot{x}_4, \dot{x}_5, \mathbf{u}, \dot{\mathbf{u}}) \\ \phi_3(x_4, x_5, x_6, \dot{x}_4, \ddot{x}_4, \dot{x}_5, \mathbf{u}, \dot{\mathbf{u}}) \end{pmatrix}
\tag{13}
$$

The "gray-box" model is obtained through lab research based on the material balance of fed-batch fermentation. The model ignores the influence of many factors, and merely approximates the actual kinetics [13]. There are several constraints of this model: (1) The temperature, fermenter pressure and agitation speed are constant in the fermentation process; (2) The broth and substrate concentrations are not affected by fermentation heat. The inverse system model Eq. (13) of the "virtual sensor" is established based on the gray box model under the above constraints. Obviously, the established model fails

to reflect the influence of several key factors in the fermentation process, including but not limited to fermentation temperature, in-tank pressure, agitation speed, and air flowrate. As a result, the soft-sensing prediction based on Eq. (13) will have a huge error, undermining the subsequent optimization control.

To overcome the defect, fermentation temperature (W_t), tank inside pressure(P_t), agitation speed (S_a), air flow rate (F_a) four process parameters are included to the soft-sensing model based on Eq. (13). The structure of the extended inverse model for soft-sensing can be described as:

$$\hat{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \phi_4\left(x_4, x_5, x_6, \dot{x}_4, \ddot{x}_4, \dot{x}_5, \mathbf{u}, \dot{\mathbf{u}}, W_t, P_t, S_a, F_a\right) \\ \phi_5\left(x_4, x_5, x_6, \dot{x}_4, \ddot{x}_4, \dot{x}_5, \mathbf{u}, \dot{\mathbf{u}}, W_t, P_t, S_a, F_a\right) \\ \phi_6\left(x_4, x_5, x_6, \dot{x}_4, \ddot{x}_4, \dot{x}_5, \mathbf{u}, \dot{\mathbf{u}}, W_t, P_t, S_a, F_a\right) \end{pmatrix}$$

$$(14)$$

The addition of the key parameters provides the extended inverse model for soft-sensing with more characteristic information of the fermentation process, which greatly promotes the adaptability and anti-jamming ability of the model.

Although the inverse soft-sensing model of the marine low-temperature alkaline protease MP fermentation process is constructed in this paper, However, Eq. (14) shows that the extended inverse model for soft-sensing is difficult to solve, despite the possible existence of a solution. The LSSVM offers a solution to this problem, thanks to its strong approximation ability to nonlinear functions.

Improved MLSSVM

Traditional LSSVM is grounded on multi-input single-output (MISO) systems. It cannot be directly applied to identify multi-input multi-output (MIMO) systems [14], such as the extended inverse model of the marine alkaline protease MP fermentation process. Thus, the MLSSVM was proposed to build the extended inverse model of the marine alkaline protease MP fermentation process (MIMO).

LSSVM is proposed by Suykens in which author has changed the inequality constraints in SVM [15, 16] with equality and converted the convex quadratic programming problem to a convex linear system problem, which is often used for model decomposition problems and function prediction. Its modeling principle is as follows:

Given l training samples $\{(\mathbf{x}_i, y_i) | \quad i = 1, 2, ..., l\}$, $\mathbf{x}_i \in R^n$ is an input and $y_i \in R$ is output. The optimization problem for regression LSSVM is as follows:

$$\min J(\mathbf{w}, b, \xi) = \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{2}\gamma \sum_{i=1}^{l} \xi_i^2$$
$$s.t. y_i = \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}_i) + b + \xi_i \, i = 1, 2, \cdots, l \quad (15)$$

where: $\mathbf{w}$ is the weight vector, b is the deviation, $\phi(\mathbf{x}_i)$ is a mapping to a high dimensional space, ξ_i is relaxation factor (error), γ is regularization parameter.

In order to transform the single-output optimization problem into a multi-output optimization problem. In this paper, the quadratic loss function of error($\boldsymbol{\xi}_i \boldsymbol{\xi}_i^{\mathrm{T}}$, $\boldsymbol{\xi} \in \mathbf{R}^{1 \times n}$) is introduced to replace the relaxation factor(ξ_i^2) in optimization problem(15):

$$\min J(\mathbf{w}, b, \boldsymbol{\xi}) = \frac{1}{2} \sum_{i=1}^{n} \mathbf{w}_i^{\mathrm{T}}\mathbf{w}_i + \frac{1}{2}\gamma_i \sum_{i=1}^{n} \boldsymbol{\xi}_i \boldsymbol{\xi}_i^{\mathrm{T}}$$
$$s.t. \mathbf{y}_i = \mathbf{w}_i^{\mathrm{T}}\phi(\mathbf{x}) + \boldsymbol{\gamma}^{\mathrm{T}}b_i + \boldsymbol{\xi}_i \, i = 1, 2, \cdots, l \quad (16)$$

where: $\boldsymbol{\xi} \in \mathbf{R}^{1 \times n}$, n is the number of output variables, $\phi_i(\mathbf{x}) = [\phi_i(\mathbf{x}_1), \cdots, \phi_i(\mathbf{x}_l)]$.

Lagrange function is used to solve the above optimization problems:

$$L = \frac{1}{2}\sum_{i=1}^{l} \mathbf{w}_i^{\mathrm{T}}\mathbf{w}_i + \frac{1}{2}\sum_{i=1}^{l} \boldsymbol{\xi}_i \boldsymbol{\xi}_i^{\mathrm{T}} - \sum_{i=1}^{l} \mathbf{a}_i^{\mathrm{T}}\left(\mathbf{w}_I^{\mathrm{T}}\phi_i(\mathbf{x}) + \boldsymbol{\gamma}^{\mathrm{T}}b_i + \boldsymbol{\xi}_i - \mathbf{y}_i\right)$$

$$(17)$$

where: $\mathbf{a}_i \in \mathbf{R}^{m \times l}$ is a Lagrange multiplier, m is input vectors number.

According to the KKT condition, the transformation to linear equation is as follow:

$$\begin{cases} \dfrac{\partial L}{\partial \mathbf{w}_i} = 0 \rightarrow \mathbf{w}_i = \phi(\mathbf{x}_i)\mathbf{a}_i^{\mathrm{T}} \\ \dfrac{\partial L}{\partial b_i} = 0 \rightarrow \boldsymbol{\gamma}^{\mathrm{T}}\mathbf{a}_i^{\mathrm{T}} = 0 \\ \dfrac{\partial L}{\partial \boldsymbol{\xi}_i} = 0 \rightarrow \mathbf{a}_i = \boldsymbol{\xi}_i, i = 1, 2, \cdots, l \\ \dfrac{\partial L}{\partial \mathbf{a}_i} = 0 \rightarrow \mathbf{w}_i^{\mathrm{T}}\phi_i(\mathbf{x}) + \boldsymbol{\gamma}^{\mathrm{T}}b_i + \boldsymbol{\xi}_i - \mathbf{y}_i = 0, i = 1, 2, \cdots, l \end{cases}$$

$$(18)$$

From the above equations, $\mathbf{w}_i = \mathbf{a}_i\phi_i^{\mathrm{T}}(\mathbf{x})$ and $\boldsymbol{\xi}_i = \mathbf{a}_i$ can be easily obtained, and then they can be substituted into the last term of Eq. (18):

$$\mathbf{a}_i\phi_i^{\mathrm{T}}(\mathbf{x})\phi_i(\mathbf{x}) + \boldsymbol{\gamma}^{\mathrm{T}}b_i + \mathbf{a}_i - \mathbf{y}_i = 0 \quad (19)$$

So that for the above optimization problem(18), the estimation function is written as:

$$[b_i \quad \mathbf{a}_i]\begin{bmatrix} 0 & \boldsymbol{\gamma}^{\mathrm{T}} \\ \boldsymbol{\gamma} & K_i(\mathbf{x}_i, \mathbf{x}) + I \end{bmatrix} = [0 \quad \mathbf{y}_i] \quad (20)$$

where: $K(\mathbf{x}_i, \mathbf{x})$ is the kernel function satisfying Mercer condition. In this paper, the kernel function is Gaussian radial basis function.

Considering that the matrix $\begin{bmatrix} 0 & \boldsymbol{\gamma}^{\mathrm{T}} \\ \boldsymbol{\gamma} & K_i(\mathbf{x}_i, \mathbf{x}) + I \end{bmatrix}$ is nonsingular, Eq. (20) can be converted to Eq. (21) by small transformation as follows:

$$\begin{bmatrix} b_i & \mathbf{a}_i \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{y}_i \end{bmatrix} \begin{bmatrix} 0 & \boldsymbol{\gamma}^{\mathrm{T}} \\ \boldsymbol{\gamma} & K_i(\mathbf{x}_i, \mathbf{x}) + I \end{bmatrix}^{-1} \tag{21}$$

Then, MLSSVM approximation is expressed as:

$$f_i(\mathbf{x}) = \mathbf{a}_i K(\mathbf{x}_i, \mathbf{x}_j) + b_i \tag{22}$$

The MLSSVM identification of the extended inverse model for soft-sensing depends heavily on the selection of kernel function parameter $\boldsymbol{\sigma}$ and regularization parameter $\boldsymbol{\gamma}$. If $\boldsymbol{\sigma}$ is too small, the training effect will be undermined by the localized kernel; Otherwise, there will be a high risk of undertraining; If $\boldsymbol{\gamma}$ is too small, the training error will increase and the learning machine will have stronger generalization ability; Otherwise, the training error will decrease and the learning machine will have a weaker generalization ability. Traditionally, these parameters are selected empirically through trial-and-error. The selection process is inaccurate and time-consuming. To ensure the prediction precision of our extended inverse model, this paper adopts the ABC algorithm to optimize and fine-tune the parameter combination ($\boldsymbol{\sigma}, \boldsymbol{\gamma}$).

ABC optimization algorithm

In recent years, the research on intelligent optimization algorithm and its application in model parameter optimization is very active and has achieved encouraging results [15, 17, 18]. Inspired by the foraging behavior of bees, the ABC is an intelligent optimization algorithm that has been successfully applied in optimization of model parameters [19, 20]. This algorithm does not care about the specific information of the problem, but the merits and demerits of the problem. By the ABC, three types of bees are set up to perform local optimization, and the optimal food source is updated iteratively to obtain the global optimal solution. Therefore, the ABC converges fast and stays immune to the local optimum trap, providing an effective way to solve multi-dimensional engineering problems. Many numerical examples have shown that the ABC has better optimization and convergence performance than differential evolution (DE), genetic algorithm (GA) and particle swarm optimization (PSO) [21, 22]. That is why this algorithm is adopted here to optimize two key parameters. Based on this, the paper selects the ABC algorithm to optimize the performance parameters of MLSSVM.

In the ABC, the colony consists of three groups of bees: leading bees, following bees and scout bees. The leading bees whose food source has been abandoned
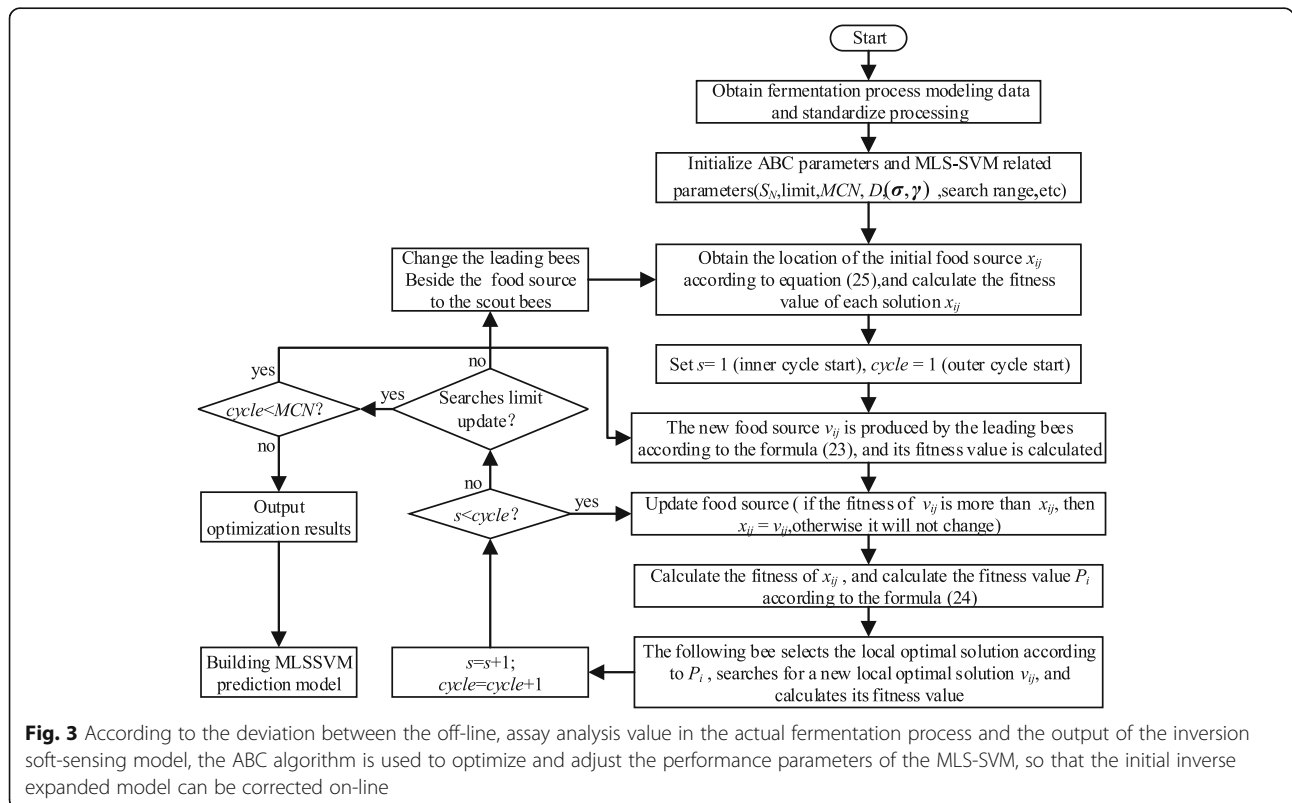
Fig. 3 According to the deviation between the off-line, assay analysis value in the actual fermentation process and the output of the inversion soft-sensing model, the ABC algorithm is used to optimize and adjust the performance parameters of the MLS-SVM, so that the initial inverse expanded model can be corrected on-line

becomes a scout. The leading bees search for high-quality food sources, the following bees watch the dances of leading bees and choose one source depending on the dances, and the scout bees search for new food sources randomly around the chosen source. The total number of leading bees and following bees equals the number of food sources. Let S_N. $x_{ij}(i = 1, 2, \cdots, SN, j = 1, 2, \cdots, D)$ be the locations of food sources, with D be the number of optimization parameters. After initialization, the leading bees start to search for food sources iteratively. In each iteration, an leading bee remembers the new food source, if it has a higher nectar amount than the old one. The following bees will choose a food source, go to that source, choose a neighbor, and evaluate its nectar amount. Then, abandoned food sources are determined and are replaced with the new food sources discovered scout bees. Finally, the best food source found so far is registered.

The leading bees search for new solutions based on their current location, which can be described as follows:

$$v_{ij} = x_{ij} + \phi_{ij}\left(x_{ij} - x_{kj}\right) \tag{23}$$

where: $k \in \{1, 2, \cdots, S_N\}$ and $j \in \{1, 2, \cdots, D\}$ are randomly selected, and $k \neq i$. ϕ_{ij} is a random number between $[-1, 1]$.

The conversion probability of each individual is calculated as follows:

$$P_i = \frac{f(x_i)}{\displaystyle\sum_{n=1}^{S_N} f(x_n)} \tag{24}$$

where: $f(\mathbf{x}_i)$ is the fitness value of each individual and S_N is the number of food sources.

If the solution x_i update fails, it means that the solution can not be optimized and needs to be replaced by a new solution generated by running the following formula:

$$\mathbf{x}_{ij} = \mathbf{x}_{\min, j} + rand(0, 1)\left(\mathbf{x}_{\max, j} - \mathbf{x}_{\min, j}\right) \tag{25}$$

When using the ABC algorithm to optimize the MLSSVM parameters, it is necessary to set the relevant parameters and fitness functions, including the initialization of the control parameters in the ABC algorithm. The detailed flow chart of ABC-MLSSVM is shown in Fig. 3, the specific parameters are set as follows:

1) Initialize various parameters in the ABC algorithm, the number of food source S_N is 20, the maximum number of searches limit is 50, and the number of termination cycles *MCN* is 100.

2) Performance parameters (γ, σ) of MLSSVM represents the location of food source, D is set to 6,
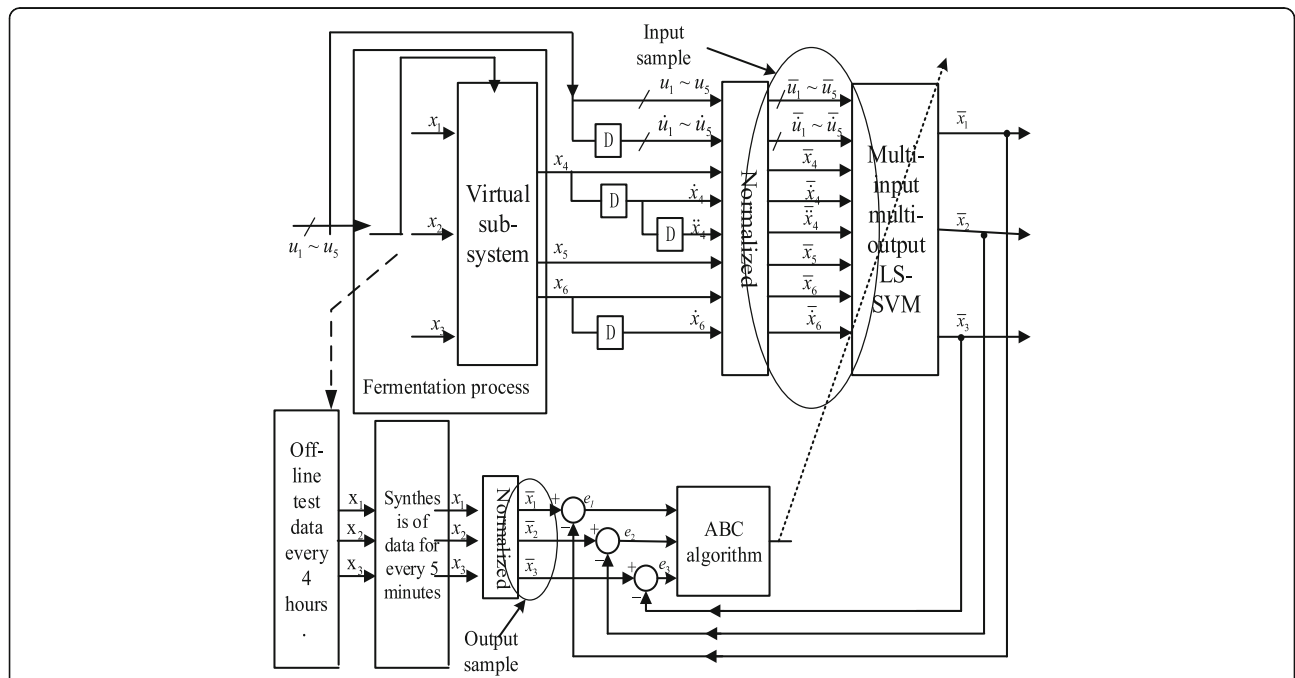
Fig. 4 The initial value is preset to initialize the ABC algorithm, and the dimension of variables and boundary conditions are set. The fitness function is applied as $F(x_i) = 1/MSE_i$, where MSE_i represents the root mean square error of the MLS-SVM of i th solution, MSE_i of the initial MLS-SVM training sample is calculated as the initial value of the fitness function, and the minimum value of the fitness function in the global range is the optimal MLS-SVM soft-sensing model, in this case, the corresponding (γ, σ) is the optimal solution in the set search range

and the search range of LSSVM parameters is set to [0.01, 1000].

3) Set the usage function in the ABC algorithm. The purpose of optimizing MLSSVM is to reduce the prediction error, so the fitness function is applied as $F(x_i) = \frac{1}{MSE_i}$, where MSE_i represents the root mean square error of the MLSSVM of i-th solution.

ABC optimization algorithm does not care about the specific information of the problem, but the merits and demerits of the problem. It can quickly converge and keep immune to the local optimal trap, which provides an effective way to solve multi-dimensional engineering problems. Based on this, this paper uses ABC algorithm to optimize the performance parameters (σ, γ) of MLSSVM, in order to get a more accurate inverse model.

Inverse model identification based on ABC-MLSSVM

The order of each input and its derivative of ABC-MLSSVM inverse soft-sensing model is determined according to Eq. (14), and ϕ_4, ϕ_5, ϕ_6 in Eq. (14) are obtained by using MLSSVM off-line identification and ABC algorithm on-line optimization. Then the inverse soft-sensing model developed in this way can realize the

soft-sensing of x_1, x_2, x_3, and the kernel function of MLSSVM is Gauss radial basis function. The identification process of inverse soft-sensing model is as follows:

1) Fermentation data acquisition. On the premise of meeting the sampling theorem, the input variable **u** is collected with appropriate excitation signal during the working area of the marine low-temperature alkaline protease MP fermentation process, direct measurable parameter $\{x_4, x_5, x_6\}$, and process parameter $\{W_t, P_t, S_a, F_a\}$ to obtain the original data sample set $\{u_1, u_2, u_3, u_4, u_5, x_4, x_5, x_6, W_t, P_t, S_a, F_a\}$. Non-direct measurable variable $\{x_1, x_2, x_3\}$ can be obtained by off-line, assay analysis in the laboratory.

2) Data preprocessing. Through certain technical processing (such as digital filtering, improving measurement redundancy, etc.), the bad data caused by working conditions, manual operation or environmental impact can be deleted, and the reliability of sample data can be improved. At the same time, in order to accurately calculate the required derivatives, according to the structure of the extended inverse model determined by Eq. (14), the five-point derivation method is adopted to
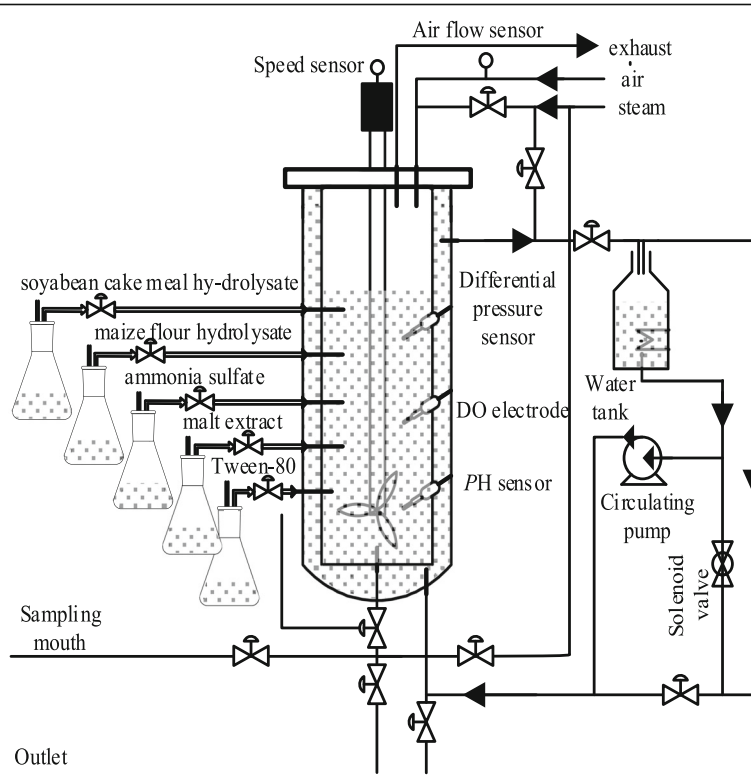
Fig. 5 The fermentation tank is added into the culture medium (soyabean cake meal hy-drolysate, maize flour hydrolysate, ammonia sulfate, malt extract, Tween-80) and sterilized by steam at high temperature, after cooling, the fermentation strain is connected to the fermentation strain in a certain proportion, and the appropriate amount of the enzyme producing strain is carried out at the right time
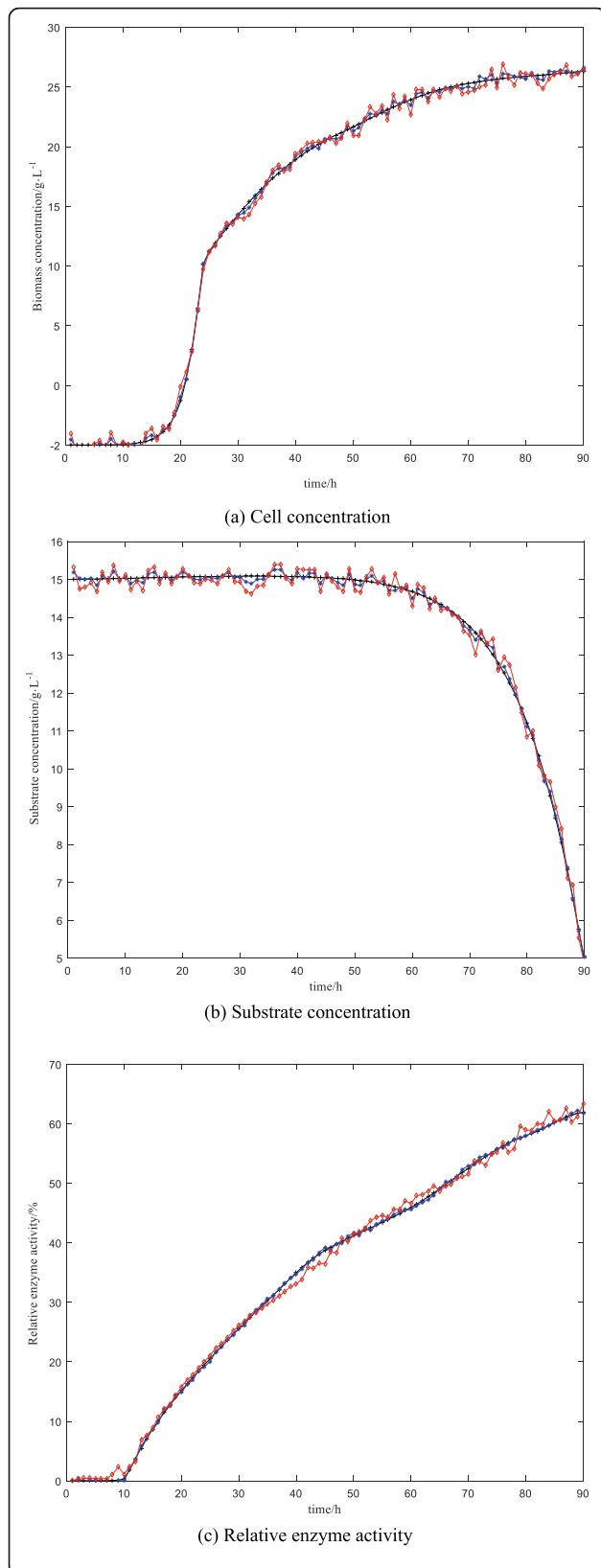
Fig. 6 The actual data curve is described by a black solid line with +, the predictive value curve of ABC-MLSSVM inversion soft-sensing modeling is described by a red solid line with *, the predictive value curve of MLSSVM soft-sensing modeling is described by a blue solid line with ◊. (**a**) shows the comparison between the the predictive value curve of biomass concentration and the actual data curve. (**b**) shows the comparison between the the predictive value curve of substrate concentration and the actual data curve. (**c**) shows the comparison between the the predictive value curve of relative enzyme activity and the actual data curve

obtain the every derivative $\{\dot{x}_4, \ddot{x}_4, \dot{x}_5, \dot{\mathbf{u}}\}$ of $\{x_4, x_5, \mathbf{u}\}$, and the interpolation method is used to process the $\{x_1, x_2, x_3\}$ (keep it synchronized with measurable data in real time to ensure consistency of data), and finally the data sample sets $\{x_1, x_2, x_3\}$ and $\{x_4, x_5, \dot{x}_6, \dot{x}_4, \ddot{x}_4, \dot{x}_5, \mathbf{u}, \dot{\mathbf{u}}, W_t, P_t, S_a, F_a\}$ are obtained, the former is used as the output of the inverse soft-sensing model, that is, the key biological parameter, and the latter is the input of the inverse soft-sensing model.

3) Off-line training and on-line correction. According to the input and output sample data, the MLSSVM is trained off-line and the corresponding initial parameters are determined by cross-validation, and the initial inverse expanded model is established. Then, according to the deviation between the off-line, assay analysis value in the actual fermentation process and the output of the inversion soft-sensing model, the ABC algorithm is used to optimize and adjust the performance parameters of the MLSSVM, so that the initial inverse expanded model can be corrected on-line. Figure 4 shows the on-line correction figure of the inverse soft-sensing model of fermentation process.

Results

Take the low-temperature alkaline protease MP fermentation process as the object for experimental verification. The flow chart of marine low-temperature alkaline protease MP fermentation process is shown in Fig. 5. In order to make the experiment closer to the production process, the experiment scheme is designed as follows:

1) The high-yield low-temperature alkaline protease strain YS-80 isolated from Huang Hai water samples of China is selected as the strain (with the characteristics of short fermentation period, high protease yield, good enzyme stability, safety and reliability, non-toxicity and so on). It is fermented in the fermenter of 1m^3 and is fermented according to the technological requirements of the marine alkaline protease MP fermentation. After the fermentation tank is added into the culture

(a) Biomass concentration

(b) Substrate concentration

(c) Relative enzyme activity
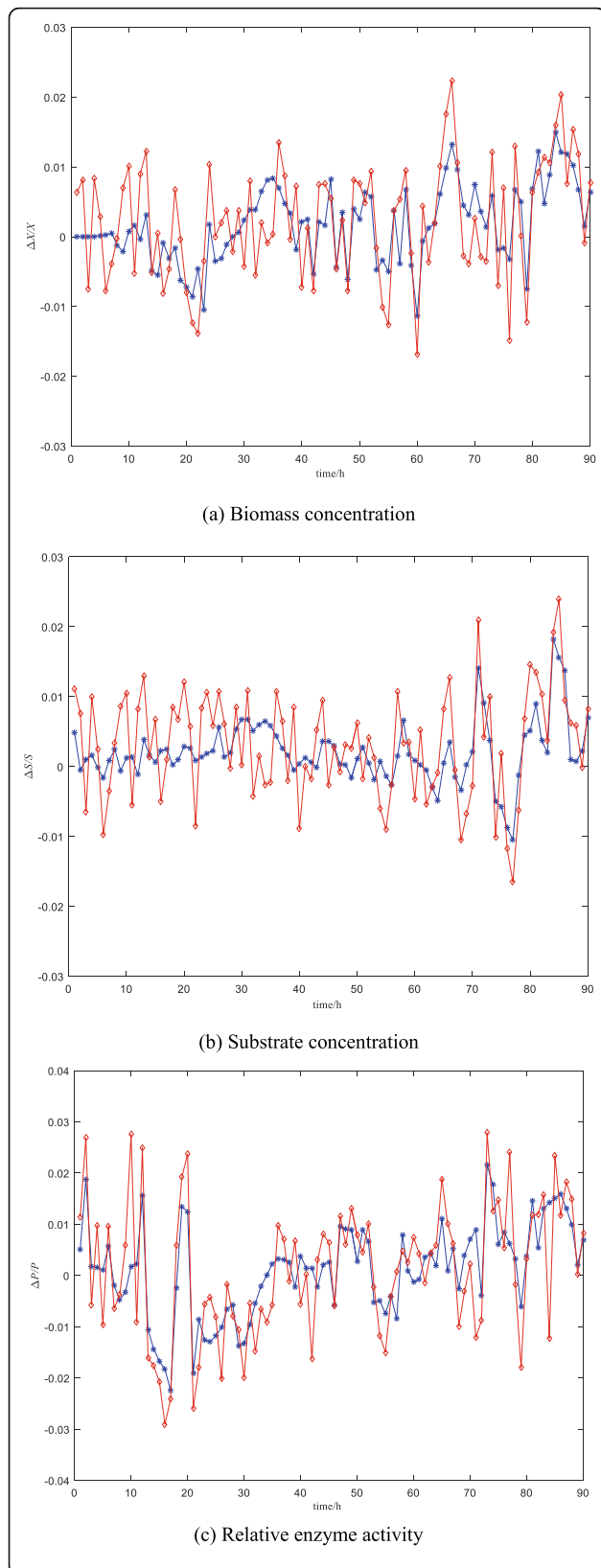
Fig. 7 The relative error curve of ABC-MLSSVM inversion soft-sensing modeling is described by a red solid line with *, the relative error curve of MLSSVM soft-sensing modeling is described by a blue solid line with ◊. (**a**) is the relative error curve of biomass concentration. (**b**) is the relative error curve of substrate concentration. (**c**) is the relative error curve of relative enzyme activity

medium and sterilized by steam at high temperature, after cooling, the fermentation strain is connected to the fermentation strain in a certain proportion, and the appropriate amount of the enzyme producing strain is carried out at the right time.

2) Set fermentation period T as 90 h and sampling period t as 5 min of each batch, the fermentation temperature is controlled at about 28 °C, the pH value is about 9.5, the tank pressure is controlled at 0.04 Mpa, the stirring speed is controlled at 250r/min, the dissolved oxygen is controlled between 45~75%, and ventilation volume is 1000 L/h. Non-direct measurable variable $\{x_1, x_2, x_3\}$ is obtained by off-line analysis and test after regular sampling (the appropriate fermentation broth is taken every 4 h through sampling mouth) in the laboratory. Among them, X is obtained based on the cell dry weight method, a certain amount of fermentation broth is centrifuged at 3000r/min for 5min in a centrifuge tube. Then, the supernatant is discarded, washing twice with distilled water, and drying it at 105 °C to a constant weight, weighing it. S is measured using a SBA – 40A glucose analyzer and P is determined by an automatic scanning spectrophotometer.

3) Only 10 batches of sample data are considered to test the identification ability of ABC-MLSSVM inverse soft-sensing model to small samples in the experiment. In order to enhance the difference among different batches, the initial conditions of each batch fermentation and the feeding strategy of each nutrient solution are set to be different. And the first six batches of fermentation data are used as training samples to off-line train the inverse expanded model of fermentation process. The seventh batch and the eighth batch of fermentation data are used to on-line correction the initial extended inverse model, and the ninth batch and the 10th batch fermentation data are used to verify the effectiveness and prediction accuracy of the inverse expanded model.

In order to test the performance of the ABC-MLSSVM inverse soft-sensing modeling method, it is compared with the traditional LSSVM soft-sensing modeling method, and the relative errors of the prediction results of the two methods are calculated. The initial

Table 1 MRE comparison by two models

Fermentation batch	ABC-MLSSVM Inversion			MLSSVM		
	X/g·L^{-1}	S/g·L-1	E/g·L-1	X/g·L-1	S/g·L-1	E/g·L-1
The 9th batch	1.38%	1.63%	1.97%	4.62%	3.36%	5.76%
The 10th batch	1.43%	1.72%	2.06%	5.11%	2.98%	6.32%

performance parameters of MLSSVM are taken as: $\sigma^2 = [1.0, 1.0, 1.0]$, $\gamma = [10, 10, 10]$, and the performance parameters of MLSSVM after on-line optimization by ABC algorithm are $\gamma = [10.1, 6.3, 8.2]$, $\sigma^2 = [0.532, 1, 613, 0.479]$.

Figure 6 is a comparison of soft-sensing results of key biological parameters of the ninth batch fermentation (protease content is characterized by relative enzyme activity in Fig. 6). Figure 7 is a relative error curve. Table 1 lists the average relative error MRE of the soft-sensing results of the two methods (protease content is characterized by relative enzyme activity in Fig. 7).

Discussion
As can be seen from Figs. 6, 7 and Table 1, compared with the traditional LS-SVM soft-sensing method, the on-line estimation results of the ABC-MLSSVM inverse soft-sensing method are closer to the off-line assay values, especially in the prediction of cell concentration. It is fully proved that the reversibility assumption of the "virtual sensor" is reasonable. During the logarithmic growth period and stable growth period (20 h–60 h) of marine low-temperature alkaline protease MP fermentation, the average RMSE (root-mean-square error) of cell concentration, substrate concentration and relative enzyme activity are 0.146, 0.127 and 0.185 respectively when the MLS-SVM method is used. While when the ABC-MLSSVM inversion method is adopted, the soft-sensing RMSE of the there results are 0.0645, 0.0538 and 0.0712. This indicates that the ABC-MLSSVM inverse system method is effective and credible, and can greatly improve the soft-sensing precision of key biological parameters in the low-temperature alkaline protease MP fermentation process, which satisfactorily meets the expected accuracy requirements.

Conclusion
In order to solve the problem that the key biological parameters of marine low-temperature alkaline protease MP cannot be measured directly on-line during fed-batch fermentation, a soft-sensing modeling method for marine low-temperature alkaline protease MP fermentation process based on ABC-MLSSVM inversion is proposed by combining the inverse system method with least square support vector machine. This paper firstly establishes a "gray-box" model for the marine low-

temperature alkaline protease MP fermentation process based on the material balance. Then, the reversibility of the nonlinear model was analyzed based on the inverse method, and the extended inverse model was constructed, coupling MLSSVM system identification with ABC optimization. Finally, the extended inverse model was connected in series with the original fermentation system, forming a composite pseudo-linear system. The composite system supports the online prediction of key biological parameters in fermentation process. The simulation results show the rationality of the system dynamic model and the validity of ABC-MLSSVM inverse soft-sensing method for predicting the key biological parameters of marine low-temperature alkaline protease MP fermentation process.

The proposed model offers a feasible theoretical method to solve the soft-sensing of key biological parameters of the marine alkaline protease MP fermentation process. The model achieves ideal identification effect based on a few input/output data, eliminating the need for an exact kinetics model of the fermentation process. The soft-sensing of key parameters can be achieved by connecting the inverse system with the original system into a composite system. With clear physical meanings, The ABC-MLSSVM inversion soft-sensing method effectively overcomes the bottleneck of traditional inverse system method: The difficulty in implementing an accurate model, and enables the soft-sensing of general nonlinear reversible systems. The proposed model enjoys a wide scope of applications, laying the basis for nonlinear soft-sensing modelling of MIMO systems.

About the authors
Bo Wang obtained doctor degree from jiangsu University in 2010. Now he is a professor in Jiangsu University and his research interests include soft measurement and control of biochemical reaction process.
Meifang Yu is currently pursuing M.S. degree at Jiangsu University, majoring in Control Engineering and her main research interests include the modeling, control and optimization.
Xianglin Zhu obtained B.Sc. Degree in 1984 from Hu Nan Institute of Engineering. Now he is a professor and supervisor of postgraduate in Jiangsu University and his research direction is intelligent control of industrial process.
Li Zhu obtained M.S. degree in 1999 from JiangSu University. She is now a lecturer in Jiangsu university and her research direction is automatic control of fermentation.

References

1. Bose A, Chawdhary V, Keharia H, Subramanian RB. Production and characterization of a solvent-tolerant protease from a novel marine isolate bacillus tequilensis p15. Ann Microbiol. 2014;64(1):343–54.
2. Wang YJ, Zhang MP, Su M, Wang YJ, Xie HG, Chen CG. The chemical characteristics of marine low-temperature alkaline protease (ys-80-122). Transact Oceanol Limnol. 2004;1:8–15.
3. Cui HX, Yang MY, Wang LP, Xian CJ. Identification of a new marine bacterial strain sd8 and optimization of its culture conditions for producing alkaline protease. PLoS One. 2015;10(12):e0146067.
4. Zhu MH. Research progress on the medical field of marine proteinases. Shandong Chemical Industry. 2016;45(1):60–3.
5. Wan Q, Lu ZH, Gao H. Screening of low temperature alkaline protease-producing strain and its enzyme-producing condition. J Microbiol. 2002; 22(5):16–8.
6. Ding SH, Wang YH, Sun LN. Soft sensor of biological parameters in the marine protease fermentation process. In: Proceedings of the 33rd Chinese control conference; 2014. p. 3620–4.
7. Wu RB, Li CW. Constructive inverse system method for general nonlinear systems. Control Theory Appl. 2003;20(3):345–50.
8. Wang WC, Dai XZ. An interactor algorithm for invertibility in general nonlinear system. Proc 5th IEEE World Congress Intell Control Automation. 2004;1:59–63.
9. Dai X, Wang W, Ding Y, Sun Z. "Assumed inherent sensor" inversion based ann dynamic soft-sensing method and its application in erythromycin fermentation process. Comput Chem Eng. 2006;30(8):1203–25.
10. Shi ZP, Pan F. Analysis,control and detection technology of fermentation process. Beijing: Chemical Industry Press; 2010. p. 41–2.
11. Yu S, Liu GH, Mei CL, Cheng JX. VIP optimal neural network inverse system soft sensing method in bio-fermentation process. J Nanjing Univ Sci Technol. 2015;39(4):447–51.
12. Sun YK, Zhang Y, Huang YH, Sun XT. Soft senor of penicillin fermentation based on inverse system. Control Eng China. 2011;18(1):46–50.
13. Wang FW, Duan SJ, Tong LJ, Sun XZ. Identification of producing low-temperature alkaline protease strains from Undaria Pinnatifida and fermentation conditions and enzymatic properties. J Chin Institute Food Sci Technol. 2017;17(2):85–91.
14. Li C, Wang SL, Zhang XM. Dynamic soft sensor modeling based on multiple least squares support vector machines. In: Processings of the 7th world congress on intelligent control & automation; 2008. p. 4315–9.
15. Wang SE, Han ZZ, Liu FC, Tang YG. Nonlinear system identification using least squares support vector machine tuned by an adaptive particle swarm optimization. Int J Mach Learn Cybern. 2015;6(6):981–92.
16. Si L, Wang B, Liu XH, Tan C, Liu Z, Xu J. Identification of shearer cutting patterns using vibration signal based on a least squares support vector machine with an improved fruit fly optimization algorithm. Sensor. 2016; 16(1):90.
17. Yuan JL, Xie J, Huang M, Fan HM, Feng EM, Xiu ZL. Robust optimal control problem with multiple characteristic time points in the objective for a batch nonlinear time-varying process using parallel global optimization. Optim Eng. 2020. https://doi.org/10.1007/s11081-019-09472-z.
18. Wang L, Yuan JL, Wu CZ, Wang XY. Practical algorithm for stochastic optimal control problem about microbial fermentation in batch culture. Opt Lett. 2019;13(3):527–41.
19. Karaboga D, Gorkemli B, Ozturk C, Karaboga N. A comprehensive survey: artificial bee colony (ABC) algorithm and applications. Artif Intell Rev. 2014; 42(1):21–57.
20. Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J Glob Optim. 2007;39(3):459–71.
21. Karaboga D, Akay B. A modified artificial bee colony algorithm for constrained optimization problem. Appl Soft Comput. 2011;11(3):3021–31.
22. Suman S, Shankar C. Parametric optimization of some non-traditional machining processes using artificial bee colony algorithm. Eng Appl Artif Intel. 2011;24(6):946–57.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.