

A multi-objective optimization method based on Gaussian process simultaneous modeling for quality control in sheet metal forming

Wei Xia · Huan Yang · Xiao-ping Liao · Jian-min Zeng

Received: 4 February 2013 / Accepted: 3 November 2013 / Published online: 15 March 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract A new modeling method, related to multiple inputs and multiple outputs (MIMO), simultaneously based on Gaussian process (GP), is proposed to optimize the combinations of process parameters and improve the quality control for multi-objective optimization problems in sheet metal forming. In the MIMO surrogate model, for the use of the system information in processing and the accuracy of the model, quantitative and categorical input variables are both taken into account in GP simultaneously. Firstly, a general method is proposed for constructing covariance functions for GP simultaneous MIMO surrogate model based on correlation matrices. These covariance functions must be able to incorporate the valid definitions of both the spatial correlation based on quantitative input variables and the cross-correlation based on categorical input variables. Secondly, the unrestrictive correlation matrices are constructed by the hypersphere decomposition parameterization, thus directly solving optimization problems with positive definite constraints is needless, and the computational complexity is simplified. Compared with independent modeling method, the proposed GP simultaneous MIMO model has higher accuracy and needs less number of estimated parameters. Moreover, the

cross-correlation between the outputs (quality indexes) obtained by proposed model provides some reference to further develop quality intelligent control strategies. Finally, a drawing-forming process of auto rear axle housing is taken as an example to validate the proposed method. The results show that the proposed method can effectively decrease the crack and wrinkle in sheet metal forming.

Keywords Sheet metal forming · Multi-objective optimization · Gaussian process · Correlation matrix · Hypersphere decomposition

1 Introduction

The drawing forming of large- and medium-sized panels is always the emphasis and difficulty of the stamping process. Sheet metal forming is affected by many factors: die structure, blank shape, material thickness, material properties, fillet radius, blank holder force, drawbead layout, lubrication condition, and so on. Therefore, it is difficult to meet the requirement of high accuracy through parameters selection by only experience. For quality monitoring and control, it often needs to model the product fabrication process. However, sheet metal forming is a complex process involving the three nonlinear problems in mechanics [1]. Thus, rigorous derivation of such mathematically–physically derived models for sheet metal-forming processes can be a formidable task. The model building process is very lengthy, tedious, and error prone or can even be impossible to derive analytically. Recently, considering the deficiency of the mathematically–physically derived models and cost reduction in trial and error, many researchers and engineers devote their attention on finite element approach [2–4] based on design of experiments (DOEs) [5].

However, a large finite element simulation of sheet metal forming can take hours or even days to run, together with

W. Xia · H. Yang (✉) · J.-m. Zeng
Ministry–Province Jointly Constructed Cultivation Base for State Key Laboratory of Processing for Non-ferrous Metal and Featured Materials, Guangxi Zhuang Autonomous Region, 100 Daxue Road, Nanning 530004, People's Republic of China
e-mail: yanghuan_23503@163.com

W. Xia
e-mail: xiawei@gxu.edu.cn

J.-m. Zeng
e-mail: czjmg2004@163.com

X.-p. Liao
College of Mechanical and Engineering, Guangxi University, 100 Daxue Road, Nanning 530004, People's Republic of China
e-mail: xpfeng@gxu.edu.cn

different experiment designs and processing parameters may have plenty of combinations. It will raise huge computational cost in simulation and optimization. Although the solution obtained from DOEs can get a feasible combination of given process levels. But it is usually a local optimum, not the global one [6]. Recently, some scholars use surrogate models in optimization procedure of sheet metal forming to establish a mathematical or statistical approximation which will replace the expensive simulation analyses, such as response surface model [7, 8], artificial neural network [9, 10], and support vector regression [11]. Surrogate models are widely used to study complex systems which are time-consuming or expensive to be physically experimented on. Compared with physical models, high-fidelity surrogate models provide an efficient alternative for design of complex engineering systems and have much higher flexibility.

Recently, nonparametric models and techniques enjoy a growing popularity in the field of machine learning. Among these Bayesian frameworks [12], the Gaussian process (GP) model has recently received significant attention [13]. Zhou et al. presented an integrated simulation-based optimization system based on GP approach [14, 15]. The architecture of process optimization based on Gaussian process regression and parameter correlation system were put forward in Liao's article [16].

In order to monitor and control multiple quality indexes, it needs to develop the multiple inputs and multiple outputs (MIMO) surrogate modeling methods [17, 18]. And recently, intelligent optimization algorithms [19, 20] for multi-objective problems [21] have developed rapidly, such as Cuckoo search algorithm [22], bee colony optimization approach [23], particle swarm optimization approach [24], differential evolution algorithm [25, 26], and so on. However, the disadvantages of traditional MIMO models are mainly included in the following two points. Firstly, traditional methods assume that all the inputs are quantitative, so that the existing models are designed to deal with only quantitative input factors (quantitative input variables). In fact, the number of investigated outputs is given, and each output is categorical by nature (categorical input variables). This kind of input information can determine some important properties of the models. Because it do not take into account both types of input variables in modeling, the traditional methods for MIMO models often fail to take full advantage of the data abundance. Secondly, those outputs of the MIMO models often bear similarities with each other. But the traditional methods ignore possible correlations among the outputs at the same input values for the quantitative input variables [27].

In this data-rich environment, the surrogate modeling methods for MIMO models with both quantitative and categorical input variables have drawn increasing attention. But the research on such simultaneous MIMO surrogate models [28] is still very limited. Qian et al. proposed a general framework for building general stochastic processes models with qualitative and

quantitative factors [29]. In material processing, just like sheet metal forming [30–32], the investigated quality indexes often have some correlation with each other. So a novel surrogate modeling method based on GP that can incorporate both quantitative input variables and categorical input variables is proposed in this paper. As a key to the development of the new method is developing, a valid and physically meaningful covariance functions, which must be able to incorporate both the spatial correlation based on quantitative input variables and the cross-correlation based on categorical input variables for the GP simultaneous MIMO surrogate model. The cross-correlation based on categorical input variables is in the form of correlation matrices, so the proposed models are also called GP simultaneous MIMO surrogate model based on correlation matrices. Correlation matrices can be divided into the unrestrictive and the restrictive [29, 33, 34], common restrictive matrices are isotropic correlation matrices, multiplicative correlation matrices, group correlation matrices, ordinal correlation matrices, and so on. It is possible to significantly simplify the computational complication by taking a restrictive correlation matrix. But such one usually can be applicable to specific process or system, thus lacking the flexibility of capturing various types of correlations of the outputs.

In this paper, by using hypersphere decomposition parameterization, a novel GP simultaneous MIMO surrogate modeling framework based on unrestrictive correlation matrices is proposed. The hypersphere decomposition parameterization [35] is used to construct unrestrictive correlation matrices. This new parameterization essentially turns some optimization problems with positive-definite constraints into standard and easy-to-compute optimization problems with box constraints. Such approach inherits the flexibility of the unrestrictive correlation matrices but replaces the complicated estimation and optimization procedure used in traditional approach. The estimation and optimization procedures of hyperparameters and correlation parameters in the covariance functions are also developed. Comparing with traditional MIMO models through numerical examples, a case study of drawing forming is used to validate the proposed method. This new surrogate modeling method is simpler, has higher prediction accuracy, and requires minimal physical knowledge of the underlying system to obtain the correlations between outputs. Combined with the correlations of each process parameter for one certain single objective mentioned in Xia et al.'s literature about canonical correlation analysis based on GP [36], more comprehensive strategies are further developed for multi-objective optimization in sheet metal forming and quality intelligent control.

The rest of this paper is composed as follows. The following sections give a brief mathematical description of GP. GP simultaneous MIMO surrogate model based on correlation matrices is proposed. Then, unrestrictive correlation matrices through hypersphere decomposition parameterization are introduced. The estimation and optimization procedures of

hyperparameters and correlation parameters in the models are also developed. Comparisons between the proposed MIMO models and traditional MIMO models through two numerical examples are presented in Section 3. The superiorities of the proposed method in accuracy, correlation capture, and Pareto fronts obtained from optimization algorithm through different modeling methods has been demonstrated, respectively. The effectiveness and feasibility of the proposed method to reduce the crack and wrinkle of an auto rear axle housing is illustrated in Section 4 followed by conclusions (Section 5).

2 The GP surrogate model based on unrestrictive correlation matrices with the hypersphere decomposition parameterization

2.1 The Gaussian process regression

The theory of GP can be dated back to the classical statistical method by O'Hagan [37]. However, the application of GP as a regression (and classification) technique was not common until the late 1990s, when rapid development of computational power facilitated an implementation of GP for large datasets. Recently, successful applications of GP models have been reported in various fields, including biomass estimation in batch biotechnological processes [38], chemical process modeling [39], modeling of dynamic systems for gas–liquid separator [40], and mechanical system modeling and optimization [41]. For the followed theoretical description, in this subsection, a brief overview of GP regression technique is given, including the formulation and implementation of the model.

Gaussian process is a flexible, probabilistic, and nonparametric model with uncertainty predictions in probability statistics. Its usage and modeling properties are reviewed in Rasmussen's literature [42].

Consider a collection of random variables $\mathbf{y} = [y_1, \dots, y_N]$ (N observations) that have a joint multivariate Gaussian distribution. The data are often noisy as well, from measurement errors and so on. Thus, each observation y can be viewed as a function related to an underlying function $f(\mathbf{x})$ through a Gaussian noise model:

$$y_i = f(\mathbf{x}_i) + \varepsilon(\mathbf{x}_i) \tag{1}$$

where $f(\mathbf{x}_i)$ is the corresponding random variable of Gaussian process $\{f(\mathbf{x})\}$ at \mathbf{x}_i , $\mathbf{x}_i = [x_i^1, \dots, x_i^d]^t$, ε_i is the noise that have independent identical Gaussian distribution:

$$\varepsilon(\mathbf{x}_i) \sim (0, \nu_0) \tag{2}$$

In general, it is assumed that the mean of the Gaussian process $\{f(\mathbf{x})\}$ is zero. Thus, what relates one observation to another in such cases is just the covariance function,

$\text{Cov}(y_p, y_q) = k(\mathbf{x}_p, \mathbf{x}_q)$, and the covariance function fully specify the GP. Note that the covariance function $k(\cdot)$ can be any function having the property of generating a positive-definite covariance matrix.

A common choice is the squared exponential covariance function:

$$k(\mathbf{x}_p, \mathbf{x}_q) = \nu_1 \exp \left[-\frac{1}{2} \sum_{d=1}^D w_d (x_p^d - x_q^d)^2 \right] + \delta_{pq} \nu_0 \tag{3}$$

where $\theta = [w_1, \dots, w_D, \nu_0, \nu_1]^T$ are the hyperparameters of the covariance functions, ν_0 is the estimated white noise variance, ν_1 is the estimate of the vertical scale of variation, D is the input dimension, and $\delta_{pq} = 1$, if $p = q$ and $\delta_{pq} = 0$ otherwise. For a given case, the parameters are learned through the training dataset. And then the importance of the relevant input components can be measured by the w parameters: if w_d is zero or near zero, it means that the inputs in dimension d contain little information and could possibly be removed.

Consider a set of N -dimensional input vectors $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^T$ and a vector of output data $\mathbf{y} = [y_1, y_2, \dots, y_N]^T$. Based on the data (\mathbf{X}, \mathbf{y}) , and given a new input vector \mathbf{X}_* , we wish to find the predictive distribution of the corresponding output y_* . Unlike other models, there is no model parameter determination within a fixed model structure. In this model, tuning the parameters of the covariance function is taken as the key. It can be resolved by maximizing the log-likelihood of the parameters, which is computationally relatively demanding, however, the number of parameters to be optimized is small ($D + 2$; see Eq. 3), which means that the optimization convergence might be faster and that the ‘‘curse of dimensionality’’ is circumvented or at least decreased.

The described approach can be easily utilized for a regression calculation. Based on the training set \mathbf{X} , among all possible combinations of these points, we calculate the covariance function, (3), summarizing the findings in the three matrices:

$$K = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & k(\mathbf{x}_1, \mathbf{x}_2) & \cdots & k(\mathbf{x}_1, \mathbf{x}_N) \\ k(\mathbf{x}_2, \mathbf{x}_1) & k(\mathbf{x}_2, \mathbf{x}_2) & \cdots & k(\mathbf{x}_2, \mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{x}_N, \mathbf{x}_1) & k(\mathbf{x}_N, \mathbf{x}_2) & \cdots & k(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} \tag{4}$$

$$K_* = [k(\mathbf{x}_*, \mathbf{x}_1) \quad k(\mathbf{x}_*, \mathbf{x}_2) \quad \cdots \quad k(\mathbf{x}_*, \mathbf{x}_N)] \quad K_{**} = [k(\mathbf{x}_*, \mathbf{x}_*)] \tag{5}$$

where K is a covariance matrix ($N \times N$), K_* is the covariances between the test and the training cases ($N \times 1$), and K_{**} is the covariance between the test input and itself. For a new test input, the predictive distribution of the corresponding output is $y_* | (\mathbf{X}, \mathbf{y}), \mathbf{X}_*$, and it is Gaussian, with a mean and a variance:

$$\mu(\mathbf{x}_*) = K_* K^{-1} \mathbf{y} \tag{6}$$

$$\sigma^2(x_*) = K_{**} - K_* K_*^{-1} K_*^T \quad (7)$$

The hyperparameters of the covariance function are not known in advance, and they must be determined by using the training data. The hyperparameter θ can be estimated by maximizing the marginal log-likelihood function:

$$L = \log p(y|\theta, X) = -\frac{1}{2} y^T K^{-1} y - \frac{1}{2} \log |K| - \frac{N}{2} \log 2\pi \quad (8)$$

This is a nonlinear optimization problem which can be solved by using gradient-based methods, e.g., the conjugate gradient method. These methods require calculating the derivative of log-likelihood with respect to each hyperparameter θ , which is:

$$\frac{\partial L}{\partial \theta} = -\frac{1}{2} \text{tr} \left(K^{-1} \frac{\partial K}{\partial \theta} \right) + \frac{1}{2} y^T K^{-1} \frac{\partial K}{\partial \theta} K^{-1} y \quad (9)$$

where $\partial C/\partial \theta$ can be obtained from the covariance function. This method also is called maximum a posteriori approach [30].

2.2 A general method for constructing covariance functions for GP simultaneous MIMO surrogate model based on correlation matrices

In many engineering applications, the surrogate models for multi-objective are often necessary, i.e., MIMO models. However, the conventional way to build a MIMO model is to build a surrogate model for each output individually using methods such as GP. In such way, the MIMO model is essentially treated as a group of independent single output models, ignoring the correlations between the outputs in a MIMO model. In fact, the number of objectives is known, and it should be considered as inputs for modeling as well as training set (quantitative input variables). This type of input information ignored in traditional MIMO modeling is called categorical input variables z in this paper. If no special illustration, categorical input variables z is not ordinal here.

To take into account the distinct natures and roles of quantitative and categorical input variables in a single GP MIMO model, a general method for constructing covariance functions based on GP is proposed in this subsection, which must be able to incorporate the valid definitions of both the spatial correlation based on quantitative input variables and the cross-correlation based on categorical input variables.

Consider an input vector $\mathbf{h} = (x^t, z_c)^t$, where x^t is quantitative input variable, z is categorical input variable, denoted by $c = 1, \dots, m$, which is the number of outputs, namely objectives. Similar to (1), the model assumes the following:

$$y = f(\mathbf{h}) + \varepsilon(\mathbf{h}) \quad (10)$$

the basic hypothesis of this model is consistent with the standard GP model.

The spatial correlation based on quantitative input variables has been defined in covariance function (3). Thus, it only needs to define a valid cross-correlation based on categorical input variables in the covariance function, then the GP simultaneous MIMO surrogate model is fully specified. Let $\varepsilon_c(\mathbf{x}) = \varepsilon((x^t, c)^t)$, for $c = 1, \dots, m$, and

$$\varepsilon^*(\mathbf{x}) = (\varepsilon_1(\mathbf{x}) \cdots \varepsilon_m(\mathbf{x})) \quad (11)$$

Assuming that $\varepsilon^*(\mathbf{x}) = A\eta(\mathbf{x})$, where $A = (a_1, \dots, a_m)^t$ is an $m \times m$ nonsingular matrix with unit row vectors (i.e., $a_c^t a_c = 1$ for $c = 1, \dots, m$), and $\eta(\mathbf{x}) = (\eta_1(\mathbf{x}), \dots, \eta_m(\mathbf{x}))^t$, where $\eta_1(\mathbf{x}), \dots, \eta_m(\mathbf{x})$ are independent Gauss processes with the same variance and covariance function $k(x_p, x_q)$. Thus, $\text{Cov}(\eta(x_p), \eta(x_q)) = k(x_p, x_q) I_m$, with I_m being the $m \times m$ identity matrix. Then, for the number of outputs is two, i.e., $m = 2$, the covariance function is defined to be

$$\begin{aligned} \text{Cov}(y(h_1), y(h_2)) &= \text{Cov}(y_{z_1}(x_p), y_{z_2}(x_q)) \quad (12) \\ &= \text{Cov}(a_{z_1}^t \eta(x_p), a_{z_2}^t \eta(x_q)) \\ &= a_{z_1}^t a_{z_2} k(x_p, x_q) \end{aligned}$$

Let $\tau_{r,s} = a_r^t a_s$, where $r, s = 1, \dots, m$. Then, $T = \{\tau_{r,s}\} = AA^t$ is an $m \times m$ positive definite matrix with unit diagonal elements (PDUDE), also called correlation matrix. In fact, any PDUDE can be written as BB^t , where B is a nonsingular matrix with unit row vectors. Thus, for any PDUDE $T = \{\tau_{r,s}\}$ and any $k(x_p, x_q)$, all of them have a relevant valid covariance function

$$\text{Cov}(y(h_1), y(h_2)) = \tau_{z_1, z_2} k(x_p, x_q) \quad (13)$$

For the squared exponential covariance function is chosen, as following

$$\text{Cov}(y(h_1), y(h_2)) = \tau_{z_1, z_2} \left\{ \nu_1 \exp \left[-\frac{1}{2} \sum_{d=1}^D w_d (x_p^d - x_q^d)^2 \right] + \delta_{pq} \nu_0 \right\} \quad (14)$$

where τ_{z_1, z_2} is the cross-correlation between categories z_1 and z_2 .

Through the described general method for constructing valid covariance functions, it can be known that the cross-correlation based on categorical input variables is in the form of correlation matrix T . The proposed modeling method mainly has the following advantages. Firstly, comparing with the traditional GP MIMO models, this proposed modeling method is more efficient. Considering a GP MIMO model with six design variables and three outputs, the traditional way has to build three GP models for each output individually, which

involves $24 (3 \times (D + 2))$ parameters to optimization. In contrast, the proposed modeling method only needs to build one MIMO model, which involve only $11 ((D + 2) + m^{(m-1)}/2)$ parameters (where D is the number of quantitative input variables and m is the number of categorical input variables). Secondly, due to the valid definitions of both types of input variables in the covariance function, the GP simultaneous MIMO model can be fully specified. The purpose of taking full advantage of the data abundance is achieved, thus improving the accuracy. Thirdly, from the optimized correlation matrix T , the correlations between outputs also can be obtained.

2.3 Unrestrictive correlation matrices by using hypersphere decomposition parameterization

For (14) to be a valid covariance function, the matrix T must be a valid matrix (i.e., PDUDE) which is positive definite. Standard methods used in GP models for maximizing a likelihood function involving positive definite matrix constraints. These constraints then convert to a series of nonlinear inequalities involving the elements of the matrix. Then, an optimization problem is solved with the resulting nonlinear inequalities as the constraints and the elements of the matrix as the optimization variables. This “element-oriented” approach involves many complicated nonlinear inequalities and a huge number of optimization variables even when the dimension of the matrix is not very large, making it computationally complex or even infeasible.

Considering the limitation of restrictive correlation matrices, and in order to address the positive definite constraints on matrices, the unrestrictive correlation matrix is constructed by hypersphere decomposition parameterization in this subsection. This approach inherits the flexibility of the unrestrictive correlation matrix but replaces the complicated computation

of optimization problem with a clever parameterization. It means that the approach essentially turns some optimization problems with positive definite constraints into standard and easy-to-compute optimization problems with box constraints. The hypersphere decomposition parameterization provides a simple yet flexible way to construct a PDUDE matrix. It consists of two steps.

In step 1, a Cholesky-type decomposition is applied to T given by

$$T = LL^t \tag{15}$$

where $L = \{l_{r,s}\}$ is a lower triangular matrix with strictly positive diagonal entries.

In step 2, each row vector $(l_{r,1}, \dots, l_{r,r})$ in L is modeled as the coordinate of a surface point on an r dimensional unit hypersphere described as follows. For $r=1$, let $l_{1,1}=1$ and for $r=2, \dots, m$, use the following spherical coordinate system

$$\begin{cases} l_{r,1} = \cos(\phi_{r,1}), \\ l_{r,s} = \sin(\phi_{r,1}) \cdots \sin(\phi_{r,s-1}) \cos(\phi_{r,s}), s = 2, \dots, r-1, \\ l_{r,r} = \sin(\phi_{r,1}) \cdots \sin(\phi_{r,r-2}) \sin(\phi_{r,r-1}), \end{cases} \tag{16}$$

where $\phi_{r,s} \in (0, \pi)$, all $\phi_{r,s}$ are denoted by Φ . Because each $\phi_{r,s}$ is restricted to take values in $(0, \pi)$, the diagonal entry $l_{r,r}$ in L is strictly positive, thus guaranteeing that T is a positive definite matrix. In addition, $\tau_{r,r} = \sum_{s=1}^r l_{r,s}^2 (r = 1, \dots, m)$ by (16), implying that T must have unit diagonal elements. Thus, the matrix T under this parameterization is always a PDUDE.

For illustration, consider the case with three outputs $m=3$. In step 1, a 3×3 correlation matrix is decomposed as

$$T = \begin{bmatrix} 1 & \tau_{12} & \tau_{13} \\ \tau_{12} & 1 & \tau_{23} \\ \tau_{13} & \tau_{23} & 1 \end{bmatrix} \implies T = LL^t = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} 1 & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

In step 2, (l_{21}, l_{22}) are transformed into a 2D spherical coordinate system as

$$\begin{cases} l_{21} = \cos(\phi_{21}) \\ l_{22} = \sin(\phi_{21}) \end{cases} \tag{18}$$

and (l_{31}, l_{32}, l_{33}) are transformed into a 3D spherical coordinate system as

$$\begin{cases} l_{31} = \cos(\phi_{31}) \\ l_{32} = \sin(\phi_{31}) \cos(\phi_{32}) \\ l_{33} = \sin(\phi_{31}) \sin(\phi_{32}), \end{cases} \tag{19}$$

where $\phi_{r,s}$ can be calculated based on the following relations:

$$\begin{cases} \tau_{12} = \cos(\phi_{21}) \\ \tau_{13} = \cos(\phi_{31}) \\ \tau_{23} = \cos(\phi_{21}) \cos(\phi_{31}) + \sin(\phi_{21}) \sin(\phi_{31}) \cos(\phi_{32}). \end{cases} \tag{20}$$

In (18), (l_{21}, l_{22}) are the coordinates of a point on the half unit circle given by $l_{21}^2 + l_{22}^2 = 1$ and $l_{22} > 0$; (l_{31}, l_{32}, l_{33}) are the coordinates of a surface point on the unit hemisphere given by $l_{31}^2 + l_{32}^2 + l_{33}^2 = 1$ and $l_{33} > 0$ in (19), as shown in Fig. 1.

According to the description above, the unrestrictive correlation matrix using hypersphere decomposition parameterization has several advantages. Firstly, comparing with the general method for constructing unrestrictive correlation matrix, it turns the complicated PDUDE constraint on T into simple box constraints $\phi_{r,s} \in (0, \pi)$. Secondly, because $\phi_{r,s}$ take values in $(0, \pi)$, the elements in T can be either positive or negative, thus possible to capture various correlations across different outputs. Thirdly, any PDUDE matrix and Φ has a one-to-one correspondence, i.e., a PDUDE matrix with any arbitrary structure can be parameterized using a set of Φ values and any given Φ always gives a PDUDE matrix. It means that Φ only need to be optimized in box constraints $(0, \pi)$, thus obtaining the optimized T .

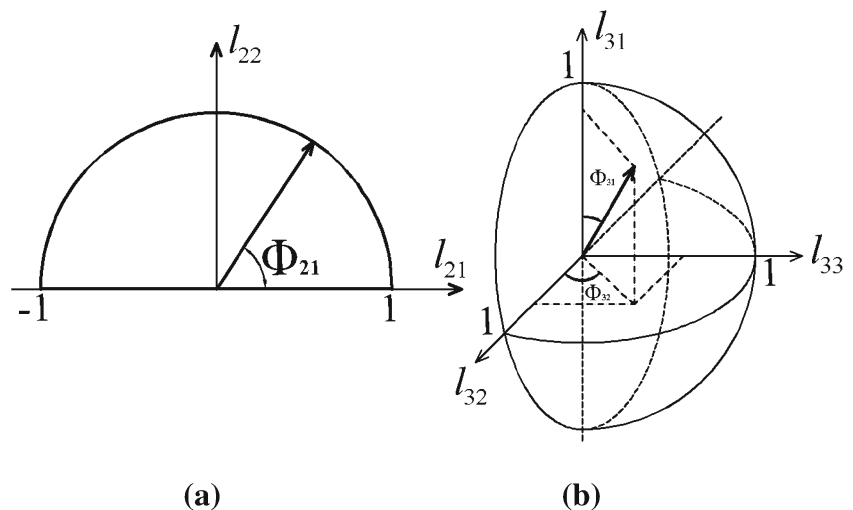
2.4 The parameters optimization for GP simultaneous MIMO model

The parameters estimation and optimization of the proposed method are presented in this subsection. Similar to the standard GP model, the elements in T , i.e., the correlations between different outputs are not known in advance (besides autocorrelation coefficients as the diagonal elements are all 1). Thus, it needs to give an initial value for T , consider a model with three outputs.

$$T_0 = \begin{bmatrix} 1 & 0.01 & 0.01 \\ 0.01 & 1 & 0.01 \\ 0.01 & 0.01 & 1 \end{bmatrix} \tag{21}$$

Setting like this means that, at the beginning of the fitting, the correlations between different outputs are very small. And then

Fig. 1 **a** Point (l_{21}, l_{22}) on the half unit circle. **b** Point (l_{31}, l_{32}, l_{33}) on the unit hemisphere. This artwork is a manual drawing and captured by AutoCAD. It is proposed to interpret the theory of the hypersphere decomposition parameterization



the values (correlations) of the elements in T are adjusted according to the training set in the coming optimization. All of the explanation for the parameters optimization is in this assumption.

The maximum likelihood method is used for the parameters estimation and optimization. The covariance matrix of the proposed method is denoted by K_m , which depends on the hyperparameters Θ and T , where T has a one-to-one correspondence with a set of Φ through the hypersphere decomposition parameterization. Thus, the parameters in the model to be optimized are Θ and Φ .

Note that $\mathbf{h} = (x^t, z_c)^t$ is a cross-array of \mathbf{X} and z_c through DOEs, like Latin hypercube, where \mathbf{X} is a $N \times D$ design matrix for the training set, and z_c is an $m \times 1$ design matrix for the categorical input variables. Consequently, \mathbf{h} consists of all level combinations between those in \mathbf{X} and z_c . Hence, \mathbf{h} has $n = Nm$ rows (runs). This cross-array structure [43] can significantly simplify the computations of the parameters optimization in K_m : the K_m can be expressed as the Kronecker product $K_m = T \otimes K$ of the correlation matrix T and the covariance matrix K defined in (4). By the properties of Kronecker product [44], considering $m = 3$ and $N = 6$, it can be given by

$$K_m = \begin{bmatrix} \tau_{11}K & \tau_{12}K & \tau_{13}K \\ \tau_{12}K & \tau_{22}K & \tau_{23}K \\ \tau_{13}K & \tau_{23}K & \tau_{33}K \end{bmatrix} \tag{22}$$

it means that the K_m transformed into an $Nm \times Nm$ block matrix. And then a series of transformation is carried out,

$$K_m^{-1} = T^{-1} \otimes K^{-1} \tag{23}$$

$$EK_m^{-1} = E(T^{-1} \otimes K^{-1}) = (E \otimes 1)(T^{-1} \otimes K^{-1}) = (EK^{-1}) \otimes T^{-1} \tag{24}$$

$$tr(EK_m^{-1}) = tr(ET^{-1} \otimes K^{-1}) = tr(EK^{-1})tr(T^{-1}) \tag{25}$$

From (25), the correlation matrix T and the covariance matrix K are mutually independent. In addition, the T has a one-to-one correspondence with a set of Φ through the hypersphere decomposition parameterization, thus similar to (8), then

$$\frac{\partial L}{\partial \Theta} = -\frac{1}{2}tr\left(K_m^{-1}\frac{\partial K_m}{\partial \Theta}\right) + \frac{1}{2}y^T K_m^{-1}\frac{\partial K_m}{\partial \Theta}K_m^{-1}y \tag{26}$$

$$\frac{\partial L}{\partial \Phi} = -\frac{1}{2}tr\left(K_m^{-1}\frac{\partial K_m}{\partial \Phi}\right) + \frac{1}{2}y^T K_m^{-1}\frac{\partial K_m}{\partial \Phi}K_m^{-1}y \tag{27}$$

Consequently, the optimizations of Θ and Φ in the proposed model can be done separately according to (26) and (27).

3 Numerical examples

In this section, two numerical examples are provided to demonstrate the effectiveness and feasibility of the proposed modeling method. For comparison purposes, we consider the following two modeling methods.

- (a) The individual Gaussian process method is denoted by IG. This method is to build a surrogate model for each output separately using the standard Gaussian process model.
- (b) The simultaneous Gaussian process method is denoted by SG. This method is to build a GP simultaneous MIMO surrogate model using the proposed method described above.

3.1 A numerical example with both positive and negative cross-correlations

This example considers one quantitative input variable x , taking values on $x \in [0,1]$ and one categorical input variable z with three levels. The response is taken from the following functions:

$$y = \begin{cases} \sin(6.8\pi x/2), & \text{if } z = 1 \\ -\sin(7\pi x/2), & \text{if } z = 2 \\ \sin(7.2\pi x/2), & \text{if } z = 3. \end{cases} \tag{28}$$

For each level of z , the training set are obtained by using a Latin hypercube design of ten runs for $x \in [0,1]$, and the testing data are then taken at 20 equally spaced points of $0, 1/19, 2/19, \dots, 1$. Figure 2 compares the three fitted curves of the functions. Intuitively, these curves are all similar to

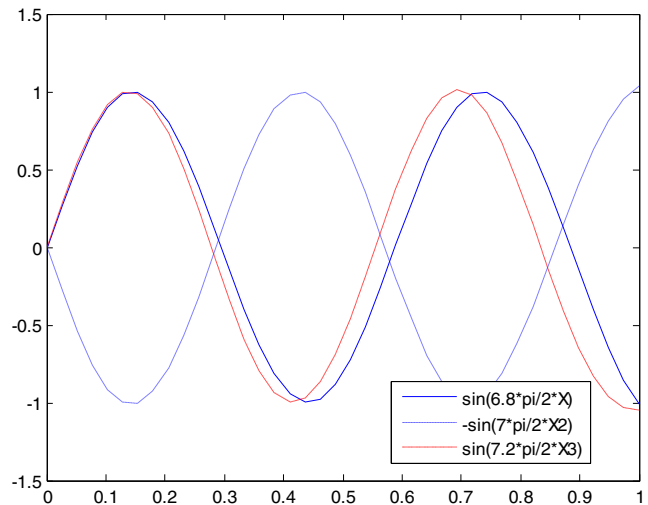


Fig. 2 Three fitted curves of the function values. This artwork is generated by MATLAB figure generator. It is proposed to illustrate three fitted curves of the first numerical example

another. The second equation is negatively correlated with another two, while the first equation is positively correlated with the third one.

The root mean squared errors (RMSEs) of the testing data are calculated for the two models to assess prediction accuracy. This procedure of data generation, model fitting, and prediction accuracy assessment is repeated 100 times. And then each RMSE of the IG and SG methods, respectively, is plotted as box plots. Figure 3 compares the RMSEs of the two methods.

The mean values of the 100 RMSEs for the IG and SG methods are 0.0419 and 0.0214, respectively, indicating that the SG method achieves the better prediction performance and that the model has much higher accuracy.

In addition, average estimates of the cross-correlations parameters $\tau_{12}, \tau_{13}, \tau_{23}$ are $(-0.94, 0.85, \text{ and } -0.93)$. The correlation matrix is

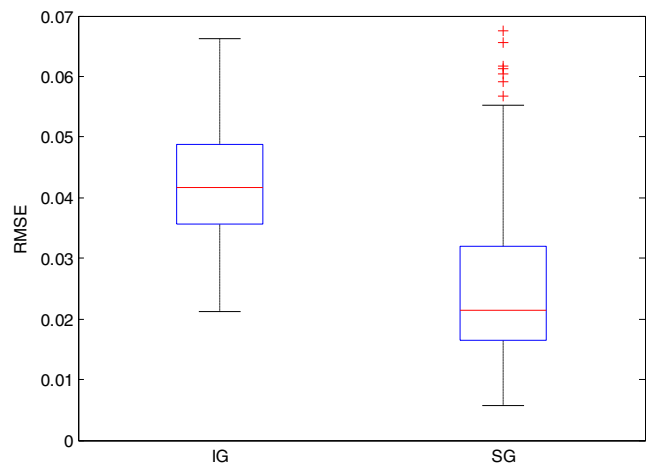


Fig. 3 Box plots of the RMSEs of the two methods. This artwork is generated by MATLAB figure generator. It tries to illustrate the RMSEs of different modeling methods

$$T_3 = \begin{bmatrix} 1 & -0.94 & 0.85 \\ -0.94 & 1 & -0.93 \\ 0.85 & -0.93 & 1 \end{bmatrix} \quad (29)$$

It is clear that the SG method correctly captures both the positive and negative cross-correlations.

This numerical example demonstrates that, comparing with IG method, the SG method has much higher prediction accuracy. Moreover, the SG method takes full advantage of the data abundance, furthermore capturing correctly both the positive and negative cross-correlations.

3.2 A numerical example with a multi-objective problem

Due to the proposed modeling method that will be applied to the multi-objective problems in engineering, one more numerical example is required to demonstrate the superiority of the method in multi-objective problems. The multi-objective problem in this numerical example is described as following:

$$\begin{aligned} f_1 &= x_1 \\ f_2 &= g(x_1, x_2)h(x_1, x_2) \\ g(x_1, x_2) &= 11 + x_2^2 - 10 \cos(2\pi x_2) \\ h(x_1, x_2) &= \begin{cases} 1 - \sqrt{f_1/g}, & f_1 < g \\ 0, & \text{else} \end{cases} \\ &(x \in (0, 1)) \end{aligned} \quad (30)$$

When the multi-objective problem is solved by using the NSGA II, the new generated populations are substituted into the analytic formulas described above for the corresponding calculations (called the analytic formula-substitution method in this example), thus the Pareto fronts are obtained.

However, many multi-objective problems in engineering cannot be derived by the analytic formula, for example sheet metal drawing forming. Consequently, the qualities of the Pareto fronts obtained by IG method are compared with those of Pareto fronts obtained by SG method, which is different from the way of verifying the performance of optimization algorithms. The specific work is as followed: the training set are obtained by using a Latin hypercube design of 20 runs for $x \in [0, 1]$, surrogate models are built for the analytic formulas f_1 and $g(x_1, x_2)$ in the multi-objective problem by using the IG and SG methods, respectively.

The purpose of this way is that, while the optimization algorithm is solving the multi-objective problem, the corresponding calculations is to use the surrogate model predictive values for the new generated populations other than being based on the analytic formula-substitution method. It means that, for the same optimization algorithm, the qualities of the obtained Pareto fronts depend on the prediction accuracy of surrogate models.

In this example, the NSGA II is used for all cases. And the population size is 50, the crossover ratio is 0.8, the mutation

ratio is 0.3, and the number of iterations is 200. The experimental results are shown in Fig. 4; a, b, and c is the Pareto front for using the analytic formula-substitution method, the Pareto front for using the IG method, and the Pareto front for using the SG method, respectively.

For the reliable optimization algorithm, the Pareto front for using the analytic formula-substitution method can be regarded as the standard front. From Fig. 4, the Pareto front for using the IG method has a very big deviation on the boundary and uneven solutions. In contrast, the Pareto front for using the SG method is quite precise near the boundary. Although the distribution of the solutions near the boundary is not as good as Fig. 4a, the front coincides overall better with Fig. 4a than the Pareto front for using the IG method, and its distribution of the solutions is more uniform, too.

According to these two numerical examples, it can be demonstrated that the proposed modeling method has much higher prediction accuracy, and takes full advantage of the data abundance, thus achieving the better performance for solving multi-objective problems.

4 Applications of the GP surrogate model used in sheet metal forming

4.1 The optimization problem

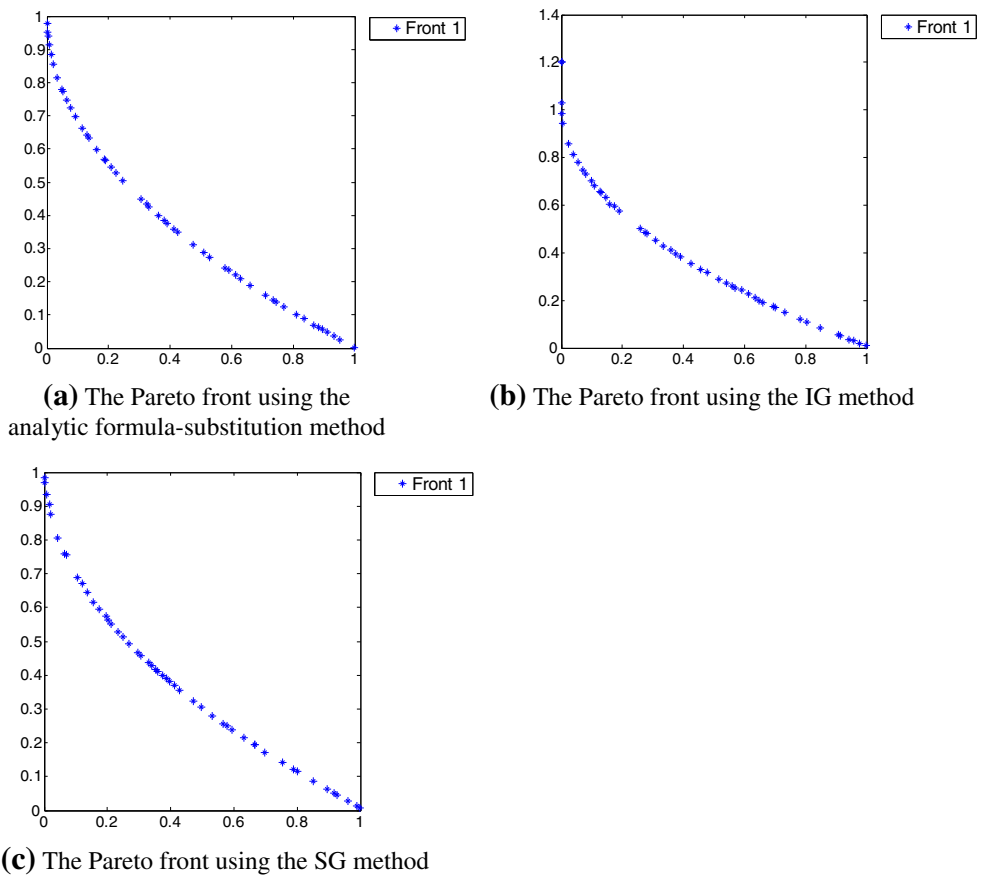
Taking an auto rear axle housing as an example, its material is 08F, and the material properties are given in Table 1. The size of the part is 200 mm in diameter, 63 mm in height, and 2.5 mm in thickness. It is a typical sheet metal-forming part, shown in Fig. 5. Crack and wrinkle are two main defects illustrated in Fig. 6. For assembly, crack is not allowed to exist in the shell structure. For welding fabrication, wrinkle is not also allowed to exist in the flange. However, many factors may cause severe crack and wrinkle. So, it is not easy to find good solutions to reduce both crack and wrinkle furthest by try-and-error method. In a workshop, for a given determined die structure and material, crack and wrinkle defects have emerged, when a 500-T hydraulic press is used. So it is given priority to reduce the crack and wrinkle by adjusting the process parameters.

4.2 Definitions of objective functions

In this subsection, the quantization objective functions of crack and wrinkle are defined according to forming limit diagram (FLD) as shown in Fig. 7. For the finite element analysis of sheet metal drawing forming, a curve is usually defined as the safety forming limit curve $\phi(\varepsilon_2)$:

$$\phi(\varepsilon_2) = \varphi(\varepsilon_2) - \Delta\varepsilon \quad (31)$$

Fig. 4 The Pareto fronts for different cases. **a** The Pareto front using the analytic formula-substitution method. **b** The Pareto front using the IG method. **c** The Pareto front using the SG method. These artworks are generated by MATLAB figure generator. It tries to illustrate the Pareto fronts for different modeling methods



where $\Delta\varepsilon$ is the crack safety margin, it is usually 10 % of the FLD_0 . If an element in the forming area is below the curve $\phi(\varepsilon_2)$, no matter how much the distance between this element and the curve $\phi(\varepsilon_2)$ is, the forming for this element is safe, thus having no possibility of crack. While an element is above the forming limit curve (FLC) $\varphi(\varepsilon_2)$, it will crack, and the further distance between this element and the FLC $\varphi(\varepsilon_2)$, the more severe the crack is. When an element is between the $\phi(\varepsilon_2)$ and the $\varphi(\varepsilon_2)$, the element has risk of crack. In order to quantify the contribution degrees of crack for elements on the different zones on FLD and the distances from the safety forming limit curve, the objective function of crack is defined as

$$D_C = \sum_{i=1}^n \begin{cases} (\varepsilon_1^i - \phi(\varepsilon_2^i))^2 & , \varepsilon_1^i > \phi(\varepsilon_2^i) \\ 0 & , \varepsilon_1^i < \phi(\varepsilon_2^i) \end{cases} \quad (32)$$

where i is the number of the grid elements on sheet metal.

Similar to the definition of the objective function of crack, firstly, the wrinkle limit curve is defined as

$$\eta(\varepsilon_2) = -\tan(45^\circ + \theta)\varepsilon_2 \quad (33)$$

where θ is the wrinkle safety margin; it is 15 % in this application. If an element is below the curve $\eta(\varepsilon_2)$, this element has the contribution degree of wrinkle, otherwise has none. So the objective function of wrinkle is defined as

$$D_W = \sum_{i=1}^n \begin{cases} (\eta(\varepsilon_2^i) - \varepsilon_1^i)^2 & , \varepsilon_1^i < \eta(\varepsilon_2^i) \\ 0 & , \varepsilon_1^i > \eta(\varepsilon_2^i) \end{cases} \quad (34)$$

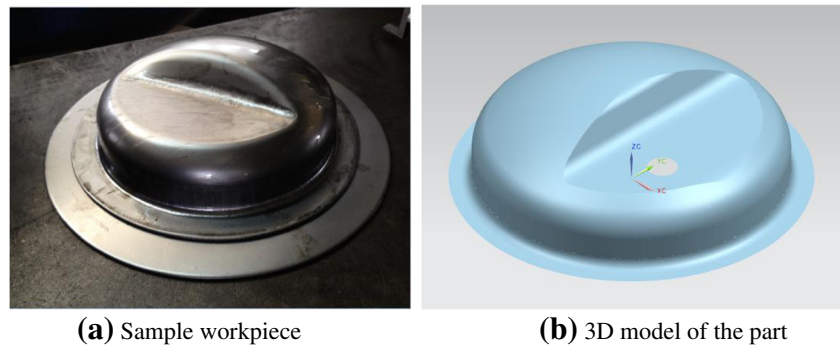
where i is the number of the grid elements on sheet metal.

The curve $\varphi(\varepsilon_2)$ and $w(\varepsilon_2)$ can be obtained according to Keeler's empirical formula [45].

Table 1 Material properties of 08F

Material parameters	Values
Poisson's ratio (ν)	0.27
Young's modulus (GPa)	207
Hardening strength coefficient (K)	525.4
Hardening exponent (n)	0.28
Anisotropy coefficient (r_0)	2.15
Anisotropy coefficient (r_{45})	2.29
Anisotropy coefficient (r_{90})	2.92

Fig. 5 A rear axle housing: **a** Sample workpiece, **b** 3D model of the part. These artworks are an entity and a 3D model of rear axle housing. It tries to illustrate the shape of the rear axle housing



4.3 The design variables, constraint conditions, and determination of optimization problem

For this case, the blank holder force (BHF) is an important design parameter. For a given determined hydraulic press, it is easy to adjust BHF. Lower BHF cannot control effectively material flow, thus causing wrinkle. While higher BHF may avoid wrinkle, but the trend of crack is increasing significantly.

The blank size is another important design parameter. Smaller diameter of the blank sheet metal cannot supply material for forming area, while bigger diameter may cause crack.

Because sheet metal has relative motion with the die and blank holder, different lubrication conditions have significant influence for the stress–strain distribution in forming, thus affecting the forming process.

Here, four design parameters are taken into account. They are BHF, diameter of the blank sheet metal D , static friction coefficient between sheet metal and die μ_1 , and static friction coefficient between sheet metal and blank holder μ_2 . Most of the engineers consider these process parameters to optimize the crack and wrinkle in the adjustment stage. In fact, fillet radius of female die also has significant influence. Because the female die fillet is involved in the forming process, the fillet radius is fixed to 5 mm. Therefore, the optimization problem can be written as

$$\begin{aligned} & \text{find } X = [\text{BHF}, D, \mu_1, \mu_2] \\ & \text{Min } f(X) = \begin{cases} f_1(x) = D_C \\ f_2(x) = D_W \end{cases} \\ & \text{subject to } \begin{cases} 80 \leq \text{BHF} \leq 500 \text{ kN} \\ 275 \leq D \leq 325 \text{ mm} \\ 0.10 \leq \mu_1 \leq 0.15 \\ 0.10 \leq \mu_2 \leq 0.15 \end{cases} \end{aligned}$$

4.4 Solving the optimization problem with the GP simultaneous model GGA-h method

For this optimization problem with four design parameters, 30 initial process parameter combinations are selected by the enhanced translation propagation Latin hypercube design [46]. By using sheet forming computer-aided engineering analysis, the objective function values of these combinations can be obtained.

The proposed GP simultaneous model is adopted to approximate the relationship between process parameters and objective function values with these 30 initial samples as the training set. The squared exponential covariance function and correlation matrix T are proposed to construct the valid covariance function. By giving a initial value for the T , the surrogate model is fitted. Consequently, the RESM of the D_C and D_W is 0.104 and 0.635, respectively.

Fig. 6 Two main defects; **a** crack and **b** wrinkle. These artworks are two defectives of rear axle housing. It tries to illustrate the main defects of this part



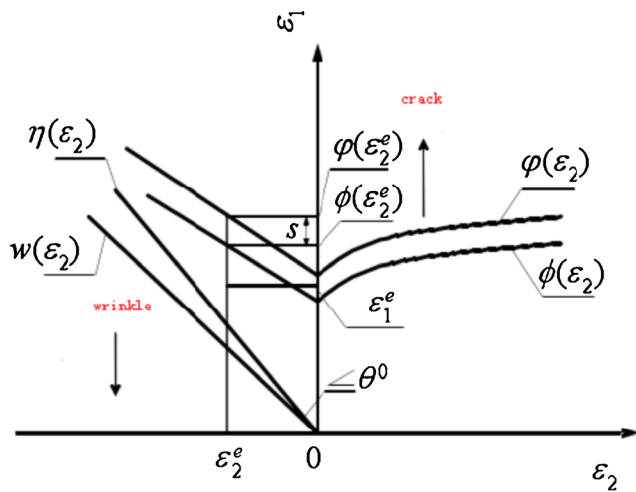


Fig. 7 FLD and the curves. This artwork is a manual drawing and captured by screen-capture software. It is proposed to interpret the definitions of objective functions

There are many intelligent optimization algorithms that can solve multi-objective problems like hybrid particle swarm optimization approach [47] and hybrid immune algorithm [48]. In this paper, to solve this multi-objective optimization, the Gaussian mutation hybrid genetic algorithm (GGA-h) [49, 50] is used, while the population size is 50, the crossover ratio is 0.9, the mutation ratio is 0.1, and the number of iterations is 300. Figure 8 shows the obtained Pareto front after 300 iterations. And for this case, the cross-correlation parameter τ_{D_C, D_W} between D_C and D_W is -0.4393 . It means that they are negatively correlated but not perfectly negatively correlated.

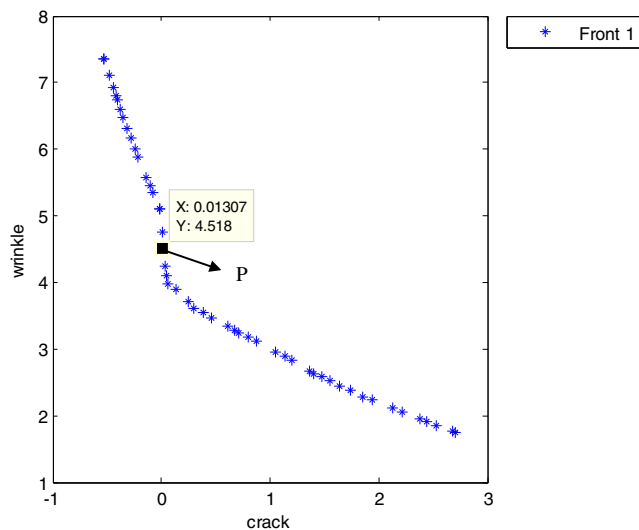


Fig. 8 The Pareto front with the GP simultaneous model GGA-h method. This artwork is generated by MATLAB figure generator. Then, the point *P* is labeled to the corresponding point by MS WORD. It tries to illustrate the satisfactory solution

4.5 Discussion and verification

Figure 8 shows that the distribution of the solutions is really well, which the solutions near the boundary are quite close to extreme values, resulting in good scalability. And the crack and wrinkle have nonlinear relationship with each other in this case. According to the cross-correlation τ_{D_C, D_W} , they are negatively correlated but not perfectly negatively correlated. Namely, as the more D_C decreases, the more D_W increases. When the compression reaches a certain level, the sheet metal loses stability, and the in-plane strain transforms into the bending deformation. Consequently, the material flow is blocked, resulting in crack.

To verify the obtained solutions by the GP simultaneous model GGA-h method, the model algorithm prediction values and the corresponding CAE verification values of three optimization parameter combinations selected from the Pareto front are enumerated in Table 2. These data shows that the model algorithm prediction values are very close to the corresponding CAE verification values. Although all the values have the similar deviations, the variation trends of two results are accordant. So the optimization solutions can be consultable for the auto rear axle housing.

Table 3 shows three set of data as comparisons selected from the Pareto front and initial samples, respectively, which their D_C values are almost equal. Comparing with their D_W values, corresponding values of the Pareto front have varying degrees of decreases. If the data, whose D_C values are almost equal, are selected, the result is the same. It demonstrates that the GP simultaneous model-GGA-h method can optimize both quality indexes of crack and wrinkle at the same time.

The above conclusions indicate that the proposed modeling method can meet the requirement of accuracy in sheet metal forming, combining with the multi-objective optimization algorithm GGA-h, the satisfactory Pareto front also can be obtained. Then, one optimization process parameter combination can be selected according to specific requirements of the product to balance two objectives.

Point *P* in Fig. 8 is selected as the satisfactory solution, its D_C and D_W value is 0.01307 and 4.518, respectively. The corresponding values of BHF, D , μ_1 , and μ_2 is 241.7, 286, 0.117, and 0.119, respectively. Simulating with this

Table 2 Comparison data of result verification

BHF (kN)	D (mm)	μ_1	μ_2	Optimization values		CAE values	
				D_C	D_W	D_C	D_W
225.3	312	0.115	0.129	0.0433	4.0919	0.0497	4.4714
372.2	319	0.103	0.104	0.2419	3.7051	0.2701	3.7483
461.4	314	0.127	0.135	1.2004	2.8271	1.5432	2.9639

Table 3 While D_C values are almost equal, comparison data between optimization solutions and initial samples

Optimization values		Initial sample values	
D_C	D_W	D_C	D_W
0.0571	3.9757	0.0568	6.2013
0.1295	3.8887	0.1258	5.1912
2.6992	1.7566	2.7094	2.4166

optimization combination, the FLD is shown in Fig. 9 and the trends of crack and wrinkle are all very small. The minimum thickness is 1.81 mm and the maximum thinning rate is 27.6 %, thus meeting the product quality requirements.

The process parameters of the hydraulic press under consideration are set according to this optimization combination, but the values are only modified slightly according to real setting accuracy. The produced part is shown in Fig. 10; the forming quality is quite well and consistent with result of the simulation. It demonstrates that the proposed method in this paper can achieve good performance for solving multi-objective problems in sheet metal forming.

5 Conclusions

In this study, the GP simultaneous model GGA-h method is introduced to minimize the crack and wrinkle of sheet metal-forming parts simultaneously, and has achieved stage result in modeling innovation and engineering application.

1. The spatial correlation based on quantitative input variables is defined with distance-based functions, while the

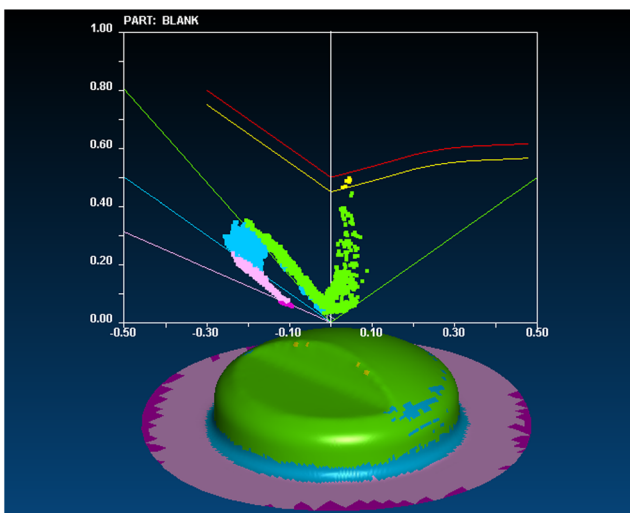


Fig. 9 FLD of the auto rear axle housing after optimization. This artwork is generated by Dynaform. It is the FLD result with the optimal parameters for rear axle housing after simulation

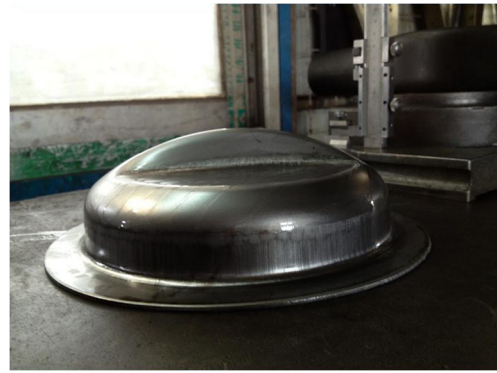


Fig. 10 The product based on optimization design. An entity is produced with the parameters optimized. As is shown, the forming quality is very well

cross-correlation based on categorical input variables is defined by correlation matrices. Then, a general method for constructing covariance functions based on correlation matrices is proposed and a novel surrogate modeling method is also developed, which can take into account both types of input variables. Furthermore, the hypersphere decomposition parameterization is used to construct unrestrictive correlation matrices. This approach maintains the flexibility of unrestrictive correlation matrices, which can turn some optimization problems with complex positive-definite constraints into standard and easy-to-compute optimization problems with box constraints. The numerical examples demonstrate that the GP simultaneous model has much higher prediction accuracy than standard GP models. Therefore, the GP simultaneous model is used to approximately construct the relationship between quality indexes and process parameters, which can replace the expensive computation in optimization. The optimization algorithm GGA-h is used to solve this multi-objective optimization problem, thus obtaining the Pareto optimal set. Then, an optimization process parameter combination is selected according to specific requirements of the product to balance two objectives.

2. As an application, an auto rear axle housing is investigated. For the Pareto optimal set, when one objective value and the corresponding value in the samples are almost equal, the other objective value always has varying degrees of improvement. Hence, the GP simultaneous model GGA-h method can be considered to minimize the crack and wrinkle of sheet metal-forming parts simultaneously, thus reducing defects as possible and obtaining products with good formability. Meanwhile, the cross-correlations between quality indexes provides some consultable reference to further develop quality intelligent control strategies, which can be obtained from the outputs of GP simultaneous model. Finally, according to verifications, optimization process parameters are consultable in real production.

- The proposed method can be expanded to any other material processing, such as injection molding and aluminum extrusion, which the investigated quality indexes often have some correlation with each other. Due to the flexibility of unrestrictive correlation matrices by using the hypersphere decomposition parameterization, the computational complexity does not increase significantly with the increasing number of investigated quality indexes. Therefore, this method can be easily expanded to three and even more objectives problems. And further studies will be performed in these fields.

Acknowledgments This work is financially supported by the National Natural Science Foundation of China (grants no. 50965003) and Ministry–Province Jointly Constructed Cultivation Base for State Key Laboratory of Processing for Non-ferrous Metal and Featured Materials, Guangxi Zhuang Autonomous Region (GXKFJ12-03). The authors would like to thank the efforts of all the members in the research group that took part in the project for their generous cooperation.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Zheng Y, Wu YG, Li SJ (1996) Development on the numerical simulation of sheet metal forming processes. *J Plast Eng* 3:34–39
- Lei LP, Sang-Moon H, Beom-Soo K (2002) Finite element analysis and design in stainless steel sheet forming and its experimental comparison. *J Mater Process Tech* 27:22–26
- Xie YM, Yu HP, Chen J, Ruan XY (2007) Robust optimization for deep-drawing process of sheet metal based on CAE with Grey relational analysis method. *J Shanghai Jiao Tong Univ (Engl Ed)* 12(3):301–307
- Mkaddem A, Bahloul R (2007) Experimental and numerical optimization of the sheet products geometry using response surface methodology. *J Mater Process Tech* 189:441–449
- Wei L, Yu YY (2008) Multi-objective optimization of sheet metal forming process using Pareto-based genetic algorithm. *J Mater Process Tech* 208(1–3):499–506
- Xia W, Luo B, Liao XP (2011) An enhanced optimization approach based on Gaussian process surrogate model for process control in injection molding. *Int J Adv Manuf Technol*. doi:10.1007/s00170-011-3227-4, Published online
- Naceur H, Ben-Elechi S, Batoz JL, Knopf-Lenoir C (2008) Response surface methodology for the rapid design of aluminum sheet metal forming parameters. *Mater Design* 29(4):781–790
- Breitkopf P, Naceur H, Rassineux A, Villon P (2005) Moving least squares response surface approximation: formulation and metal forming applications. *Comput Struct* 83(17–18):1411–1428
- Liew KM, Tan H, Ray T, Tan MJ (2004) Optimal process design of sheet metal forming for minimum springback via an integrated neural network evolutionary algorithm. *Struct Multidiscip O* 26(3–4):284–294
- Viswanathan V, Kinsey B, Cao J (2003) Experimental implementation of neural network springback control for sheet metal forming. *J Eng Mater-T ASME* 125(2):141–147
- Wang H, Li E, Li GY (2009) The least square support vector regression coupled with parallel sampling scheme metamodeling technique and application in sheet forming optimization. *Mater Design* 30(5): 1468–1479
- Currin C, Mitchell T, Morris M, Ylvisaker D (1991) Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *J Am Stat Assoc* 416:953–963
- Seeger M, Williams CKI, Lawrence ND (2003) Fast forward selection to speed up sparse Gaussian process regression. In: Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics; Key West, FL, USA
- Zhou J (2006) Single and multi-objective process optimization of injection molding using numerical simulation with surrogate modeling approaches and genetic algorithms. Dissertation, University of Wisconsin-Madison
- Zhou J, Tung LS (2007) Adaptive multi-objective optimization of process conditions for injection molding using a Gaussian process approach. *Adv Polym Tech* 26(2):71–85. doi:10.1002/adv
- Liao XP, Xia W, Long FY (2010) Gaussian process modeling of process optimization and parameter correlation for injection molding. *J Mater Sci Eng* 4(10):35
- Kennedy MC, O'Hagan A (2000) Predicting the output from a complex computer code when fast approximations are available. *Biometrika* 87:1–13
- Qian PZG, Wu CFJ (2008) Bayesian hierarchical modeling for integrating low-accuracy and high-accuracy experiments. *Technometrics* 50:192–204
- Yildiz AR (2012) A comparative study of population-based optimization algorithms for turning operations. *Inf Sci* 210:81–88
- Yildiz AR (2013) Comparison of evolutionary-based optimization algorithms for structural design optimization. *Eng Appl Artif Intell* 26:81–88
- Yildiz AR, Solanki KN (2012) Multi-objective optimization of vehicle crashworthiness using a new particle swarm based approach. *Int J Adv Manuf Technol* 59(1–4):367–376
- Yildiz AR (2013) Cuckoo search algorithm for the selection of optimal machining parameters in milling operations. *Int J Adv Manuf Technol* 64(1–4):55–61
- Yildiz AR (2013) Optimization of cutting parameters in multipass turning using artificial bee colony-based approach. *Inf Sci* 220:399–407
- Yildiz AR (2009) A novel particle swarm optimization approach for product design and manufacturing. *Int J Adv Manuf Technol* 40(5–6):617–628
- Yildiz AR (2013) A new hybrid differential evolution algorithm for the selection of optimal machining parameters in milling operations. *Appl Soft Comput* 13(3):1433–1439
- Yildiz AR (2013) A new hybrid differential evolution algorithm for the selection of optimal machining parameters in milling operations. *Appl Soft Comput* 13(3):1561–1566
- Zhou Q (2011) Computer simulation driven statistical modeling and quality control. Dissertation, University of Wisconsin-Madison
- Han G, Han G, Santner TJ, Notz WI, Bartel DL (2009) Prediction for computer experiments having quantitative and qualitative input variables. *Technometrics* 51:278–288
- Qian PZG, Wu H, Wu CFJ (2008) Gaussian process models for computer experiments with qualitative and quantitative factors. *Technometrics* 51:283–396
- Liu W, Yang YY, Xing ZW (2005) Automatic process optimization of sheet metal forming with multi-objective. NUMISHEET '2005, Detroit, USA: 843–848
- Jakumeit J, Herdy M, Nitsche M (2005) Parameter optimization of the sheet metal forming process using an iterative parallel kriging algorithm. *Struct Multidiscip O* 6:498–507
- An ZG (2009) Application of radial basis function metamodel for optimization of sheet metal forming. Dissertation, Chongqing University
- Joseph VR, Delaney JD (2007) Functionally induced priors for the analysis of experiments. *Technometrics* 49:1–11

34. McMillian NJ, Sacks J, Welch WJ, Gao F (1999) Analysis of protein activity data by Gaussian stochastic process models. *J Biopharm Stat* 9:145–160
35. Rebonato R, Jackel P (1999) The most general methodology for creating a valid correlation matrix for risk management and option pricing purposes. *J Risk* 2:17–27
36. Xia W, Zhang SH, Liao XP, Yan XL (2011) Parameters' correlation analysis based on Gaussian process for injection molding. *Adv Mater Res* 154–155:130–136
37. O'Hagan A (1978) Curve fitting and optimal design for prediction. *J Roy Stat Soc B* 40:1–42
38. Sciascio F, Amicarelli AN (2008) Biomass estimation in batch biotechnological processes by Bayesian Gaussian process regression. *Comput Chem Eng* 32:3264–3273. doi:10.1016/j.compchemeng.2008.05.015
39. Likar B, Kocijan J (2007) Predictive control of a gas–liquid separation plant based on a Gaussian process model. *Compute Chem Eng* 31:142–152. doi:10.1016/j.compchemeng.2006.05.011
40. Zhou P, Tian F, Chen X, Shang Z (2008) Modeling and prediction of binding affinities between the human amphiphysin SH3 domain and its peptide ligands using genetic algorithm–Gaussian processes. *Pept Sci* 90:792–802. doi:10.1002/bip.21091
41. Yuan J, Wang K, Yu T, Fang M (2008) Reliable multi-objective optimization of high-speed WEDM process based on Gaussian process regression. *Int J Mach Tool Manuf* 48:47–60
42. Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. MIT Press: Cambridge, MA. ISBN026218253X. www.GaussianProcess.org/gpml
43. Wu CFJ, Hamada M (2009) Experiments: planning, analysis, and optimization. Wiley, New York
44. Graham A (1981) Kronecker products and matrix calculus with applications. Ellis Horwood, Chichester
45. Keeler SP, Backofen WA (1965) Plastic instability and fracture in sheets stretched over rigid punches. *Trans ASM* 56:25–48
46. Liao XP, Yan XL, Xia W, Luo B (2010) A fast optimal latin hypercube design for gaussian process regression modeling. Third International Workshop on Advanced Computational Intelligence (IWACI 2010), Suzhou: 474–479
47. Yildiz AR (2012) A new hybrid particle swarm optimization approach for structural design optimization in automotive industry. *J Automob Eng* 226(D10):1340–1351
48. Yildiz AR (2009) A novel hybrid immune algorithm for global optimization in design and manufacturing. *Robot Comput Integr Manuf* 25(2):261–270
49. Ruan T (2012) Research of Gaussian mutation hybrid genetic algorithms and its application in injection molding. Dissertation, Guangxi University
50. Yildiz AR (2009) A novel hybrid immune algorithm for global optimization in design and manufacturing. *J Mater Process Technol* 50(4):224–228