Design space calculation and continuous improvement considering a noise parameter: A case study of ethanol precipitation process optimization for the *Carthami Flos* extract

Yanni Tai
Zhejiang University
Haibin Qu
Zhejiang University
Xingchu Gong (gongxingchu@zju.edu.cn)
Zhejiang University https://orcid.org/0000-0003-2632-4612

Research

Keywords: Design space, Ethanol precipitation, Continuous improvement, Carthami Flos, Model update, Quality by design

Posted Date: March 18th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-305170/v1

Abstract

Background

The optimization of process parameters in the pharmaceutical industry is often carried out according to the Quality by Design (QbD) concept. QbD also emphasizes that continuous improvement should be performed in life cycle management. Process parameters that are difficult to control in actual production could be regarded as noise parameters. In this study, a noise parameter was considered, an example of continuous improvement in the design space was provided.

Methods

The ethanol precipitation process of *Carthami Flos* (Honghua) extract was optimized based on the QbD concept. The critical process parameters (CPPs) were identified using a definitive screening design. Considering that the refrigeration temperature of industrial ethanol precipitation is often difficult to control with seasonal changes, the refrigeration temperature was treated as a noise parameter. The design space was then calculated using an exhaustive search-Monte Carlo method. The mathematical models were reestablished when more data were obtained and then the calculated probabilities of reaching the process standards were updated.

Results

The calculation procedure of design space based on an exhaustive search-Monte Carlo method was proposed. The density of the concentrated extract, ethanol concentration, the volume ratio of ethanol to concentrated extract, stirring time after ethanol addition, and refrigeration temperature were selected as CPPs. The mathematical models of CPPs and evaluation indicators were established, and the coefficient of determination of each model was greater than 0.81. The predictive performance of the models was good. After continuous improvement, the recalculated probability values were more reliable, the design space became larger.

Conclusions

The calculation of design space and the continuous improvement strategy considering a noise parameter was developed. In industrial production, it is also recommended to adopt this similar idea, that is, continuing to collect industrial data and regularly updating the mathematical models, which can further update the design space and make it more stable and reliable.

Background

Ethanol precipitation is a common refinement process in traditional Chinese medicine (TCM) production and is used to remove impurities such as starch, proteins, and polysaccharides while retaining the effective ingredients. Because ethanol precipitation has a good impurity removal effect and does not need special equipment, it is widely used in TCM production. The Chinese Pharmacopoeia (2020 Edition) includes 1,607 compound preparations and single preparations [1], of which 319 used ethanol precipitation, accounting for approximately 19.8% of the total preparations, and 49% of the liquid preparations contained were prepared by the ethanol precipitation process (EPP).

In recent years, many researchers have studied the ethanol precipitation of TCM, mainly focusing on the optimization of process parameters [2–5], process monitoring technology [6–8], ethanol precipitation equipment [9, 10], sediment morphology [11, 12], etc. There are also many reports on the optimization of ethanol precipitation parameters based on the quality by design (QbD) concept [13–15]. One key to implementing the QbD concept is the establishment of the design space, both to improve production flexibility and to reduce the regulatory burden [16, 17]. In actual operation, some ethanol precipitation parameters, such as refrigeration temperature (RT) of the ethanol precipitation system, are easily affected by the seasons, which are difficult to control and can be regarded as a process noise parameter. When establishing the design space, the influence of adjusting the range of easily controllable parameters and reducing the noise parameters should be considered [18]. However, there are few studies on the optimization of ethanol precipitation to distinguish easily controllable parameters from noise parameters. In addition, the QbD concept emphasizes validation throughout the lifecycle, continuous process validation, and encouraging continuous improvement. However, there is currently a lack of reporting on continuous improvement based on additional data in the process.

Guhong injection is clinically used to treat cerebrovascular diseases [19] and orthopaedic diseases [20], including cerebral insufficiency, cerebral thrombosis, coronary heart disease, fracture healing, etc. The EPP of *Carthami Flos* (Honghua) is an important link in the production of Guhong injection, and its process quality directly affects the quality of the finished products.

In this study, based on the QbD concept, the identification of critical process parameters (CPPs) of *Carthami Flos* ethanol precipitation was carried out by definitive screening design (DSD), and the mathematical relationship between CPPs and EPP evaluation indicators was established. In the process of ethanol precipitation, the level of easily controllable parameters was controlled, the sensitivity of noise parameters was reduced, and the stability of a design space method based on the probability of reaching the standard for calculating the

range of easily controllable parameters. After obtaining new data, the design space was continuously improved to obtain a more reliable design space, which also provided a reference method for improving the robustness of the process and the implementation of continuous improvement in the pharmaceutical industry.

Methods

Materials and chemicals

A concentrated extract of *Carthami Flos* (batch number 20200918) was provided by Tonghua Guhong Pharmaceutical Co., Ltd. (Jilin, China), with a density of 1.23 g/cm^3, 1 g concentrated extract equivalent to 1.11 g *Carthami Flos*. Anhydrous aluminium chloride (batch number 20170824) was purchased from Sinopharm Chemical Reagent Co, Ltd. (Shanghai, China). Hydroxysafflor yellow A (HSYA, batch number R03J10F77660) and kaempferol (batch number 180705) were purchased from Shanghai Yuanye Biotechnology Co., Ltd. (Shanghai, China), and their purity was greater than 98%. Acetonitrile, methanol, glacial acetic acid, and triethylamine for chromatographic analysis were purchased from Merck (Darmstadt, Germany). Ultrahigh-purity water was produced using a water purification system (Milli-Q, Millipore, USA).

Ethanol precipitation

The concentrated extract of *Carthami Flos* was diluted with ultrahigh-purity water or concentrated to obtain concentrates with densities of 1.18 g/cm^3 and 1.28 g/cm^3, respectively, which was accomplished with a density tester (DMA5000M, Anton Paar, Austria) and a rotary evaporator (V-100, BUCHI Labortechnik AG, Switzerland). Compared to ethanol precipitation equipment in industrial production, reduced equipment proportions are used when conducting ethanol precipitation experiments in the laboratory. A 100 mL concentrated extract sample was placed into a beaker at a certain temperature and stirred with a digital speed-measuring electric stirrer (J-JIA, Changzhou Yunhua Electrical Appliance Co., Ltd, China), and different ethanol addition times were controlled with a peristaltic pump (BT300-2J, Changzhou Runhua Electric Co., Ltd. China), with ethanol continuously added to the concentrated extract. After the addition of ethanol, the mixture was continuously stirred for a certain period, and then, the beaker was placed in a low-temperature thermostat bath (THYD-1030W, Ningbo Tianheng Instrument Factory, China) and allowed to stand. The ethanol precipitation supernatants were collected by vacuum filtration.

Determining the evaluation indicators of ethanol precipitation

Carthami Flos is a representative medicine for promoting blood circulation and removing blood stasis. It is often used to treat cerebrovascular diseases such as coronary heart disease and cerebral thrombosis. HSYA in *Carthami Flos* extract is a critical component in the treatment of cerebrovascular diseases, and it has anti-ischaemic reperfusion injury [21], neuroprotection [22], antioxidant [23], and anti-inflammatory [24] effects, among others. The Chinese Pharmacopoeia (2020 Edition) also uses HSYA as one of the quality evaluation indicators of *Carthami Flos*. Flavonoids are the main active components of *Carthami Flos* and play an important role in the treatment of cerebrovascular diseases [25]. The total solid content characterizes the removal of impurities after the EPP. Studies have reported [26, 27] that the retention rate of index components was used as the evaluation indicator for the EPP. Since the retention rate can be calculated from the total solid content and the purity of the index components, the retention rate was not included in the evaluation indicators in this study. Overall, the extracted amount of HSYA (Y_1), extracted amount of total flavonoids (Y_2), extracted amount of total solids (Y_3), HSYA purity (Y_4), and HSYA purity (Y_5) were regarded as evaluation indicators for the EPP in this study.

Experimental design

Many factors affect the quality of the supernatant from the ethanol precipitation process [28]. In this study, the authors combined industrial practical experience, the density of the concentrated extract (CD), the temperature of the concentrated extract (CT), ethanol concentration (EC), the volume ratio of ethanol to the concentrated extract (ECR), stirring speed (SS), time of ethanol addition (EAt), stirring time after ethanol addition (St), RT, and refrigeration time (Rt) to study the process parameters. DSD was used to analyse the mathematical relationship between process parameters and evaluation indicators, and its advantage is the ability to study the influence of multiple parameters with a small number of experiments [10, 29]. The coding level and factor level of the DSD are shown in Table 1. Two virtual factors were added in the experimental design, and the centre point was repeated 3 times, for a total of 28 experiments. The experimental design is shown in Table 2.

Table 1
The level of experimental design

Factors	Symbols	Coded values		
		-1	0	1
CD (g/cm^3)	X_1	1.18	1.23	1.28
CT (°C)	X_2	20	25	30
EC (%)	X_3	94	95	96
ECR (v/v)	X_4	2.6	2.8	3.0
SS (rpm)	X_5	120	140	160
EAt (min)	X_6	80	90	100
St (min)	X_7	60	90	120
RT (°C)	X_8	2.0	6.0	10.0
Rt (h)	X_9	24	36	48

Table 2
The table of experimental design

No.	Process parameters								
	X_1 (g/cm^3)	X_2 (°C)	X_3 (%)	X_4 (v/v)	X_5 (rpm)	X_6 (min)	X_7 (min)	X_8 (°C)	X_9 (h)
1	1.23	30	96	3.0	160	100	120	10.0	48
2	1.23	20	94	2.6	120	80	60	2.0	24
3	1.28	25	96	2.6	160	100	120	2.0	24
4	1.18	25	94	3.0	120	80	60	10.0	48
5	1.28	20	95	3.0	120	100	120	10.0	24
6	1.18	30	95	2.6	160	80	60	2.0	48
7	1.28	30	94	2.8	160	80	120	10.0	48
8	1.18	20	96	2.8	120	100	60	2.0	24
9	1.28	20	96	2.6	140	100	60	10.0	48
10	1.18	30	94	3.0	140	80	120	2.0	24
11	1.28	20	94	3.0	120	90	120	2.0	48
12	1.18	30	96	2.6	160	90	60	10.0	24
13	1.28	20	94	2.6	160	80	90	10.0	24
14	1.18	30	96	3.0	120	100	90	2.0	48
15	1.28	30	94	2.6	120	100	60	6.0	48
16	1.18	20	96	3.0	160	80	120	6.0	24
17	1.28	30	96	2.6	120	80	120	2.0	36
18	1.18	20	94	3.0	160	100	60	10.0	36
19	1.28	30	96	3.0	120	80	60	10.0	24
20	1.18	20	94	2.6	160	100	120	2.0	48
21	1.28	20	96	3.0	160	80	60	2.0	48
22	1.18	30	94	2.6	120	100	120	10.0	24
23	1.28	30	94	3.0	160	100	60	2.0	24
24	1.18	20	96	2.6	120	80	120	10.0	48
25	1.23	25	95	2.8	140	90	90	6.0	36
26	1.23	25	95	2.8	140	90	90	6.0	36
27	1.23	25	95	2.8	140	90	90	6.0	36
28	1.23	25	95	2.8	140	90	90	6.0	36

Analytical method

The content of HSYA was determined by an HPLC method [30], and HPLC analysis was performed with an HPLC system (1100, Agilent Technologies, USA) coupled with a UV detector. The ethanol precipitation supernatant and concentrated extract of *Carthami Flos* samples were diluted 25 and 50 times with ultrahigh-purity water, respectively. Chromatographic separation was carried out at 30°C on an Agilent Extend SB-C18 column (250 mm×4.6 mm, 5 μm). Isocratic elution of the mobile phase containing acetonitrile and 1% glacial acetic acid solution containing 0.5% triethylamine (9:91, v/v) was used. The injection volume was 10 μL, and the detection wavelength was fixed at 403 nm. Representative HPLC chromatograms of the *Carthami Flos* sample and the reference standard sample are presented in Additional file 1: Fig. S1.

The detailed determination method of the total flavonoid content in the supernatant was as described in the National Drug Standard for Guhong Injection [30], and the total flavonoid content was determined by a UV-Vis spectrophotometer (Cary 60, Agilent Technologies, USA). The total solid

content was determined using a gravimetric method. The supernatant or concentrated extract was placed into a weighing bottle that had been dried to a constant mass, the solution was dried to a constant mass at 105°C, its mass was weighed, and the total solid content in the supernatant or the concentrated extract was calculated.

Data processing

The extracted amount refers to the amount of a certain or a certain type of ingredient extracted per unit mass of *Carthami Flos.* The calculation formulas of the extracted amount of the index components and total solids in the supernatant are as follows:

$$\text{Extracted amount of index components} = \frac{M_s \times AC_i}{M_m} \quad (1)$$

$$\text{Extracted amount of total solids} = \frac{M_s \times DM}{M_m} \quad (2)$$

where M refers to the quality, and the subscripts s and m represent the supernatant and *Carthami Flos*, respectively. AC refers to the content of the index components in the supernatant, and the subscript i represents HSYA or total flavonoids. DM refers to the total solid content in the supernatant.

The calculation formulas of the retention rate and purity of the index components in the supernatant are as follows:

$$\text{Retention rate of index components} = \frac{M_s \times AC_i}{M_c \times C_i} \quad (3)$$

$$\text{Purity of index compositions} = \frac{AC_i}{DM} \quad (4)$$

where M_c refers to the quality of the concentrated extract, C refers to the content of the index components in the concentrated extract, and the subscript i represents HSYA or total flavonoids.

Design Expert 11.0.0 software was used to analyse the results of the experimental design. To identify the CPPs, a multiple linear regression model was used to model the process parameters and evaluation indicators. The basic form of the model is as follows:

$$Y = a_0 + \sum_{i=1}^{9} b_i X_i \quad (5)$$

where Y refers to an evaluation indicator, X_i (i = 1 ~ 9) is a process parameter, a_0 is a constant, and b_i is the regression coefficient of the process parameter. The backward elimination method was used to simplify the model, and the significance level was set to 0.15. Any remaining terms in the model were considered to be a CPP.

To describe the mathematical relationship between the CPPs and the evaluation indicators of the EPP, the mathematical model was established by the quadratic multiple regression analysis. The basic form of the mathematical model is as follows:

$$Y = b_0 + \sum_{i=1}^{n} b_i X_i + \sum_{i=1}^{n} b_{ii} X_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{ij} X_i X_j \quad (6)$$

where b_0 is a constant term, n is the number of CPPs, b_i, b_{ii} and b_{ij} are the regression coefficients of the first, quadratic and interaction terms, respectively, and X_i and X_j are the CPPs. The above method adopted stepwise regression, and the significance levels of adding terms and removing terms were both set to 0.10.

Calculation of the design space

The design space was calculated using the probability of reaching the standard method, which was based on an exhaustive search-Monte Carlo method. The process of this method is shown in Fig. 1. The calculation procedure is as follows:

The first step was to calculate and determine the relative standard deviation (*RSD*) of each process evaluation indicator obtained from repeated experiments. These values were recorded as RSD_1, RSD_2, ..., RSD_n, ..., RSD_N, ..., where *N* in RSD_N is the number of process evaluation indicators, *n* = 1, ..., *N*.

The second step was to determine the overall distribution parameters corresponding to each process evaluation indicator in each experiment. Assuming that the experimental values of all process evaluation indicators were a sample from the normal distribution population *A*, the average value of the normal distribution population *A* is the experimentally determined value $\mathrm{DR}_{\mathrm{n,j}}$ of the process evaluation indicators, where *j* = 1, ..., *J*, and *J* is the total number of experiments. A normally distributed population *A* was assumed to be a variance $(\mathrm{DR}_{\mathrm{n,j}} \times \mathrm{RSD}_\mathrm{n})^2$, and then, the overall distribution parameters corresponding to each process evaluation indicator could be determined.

The third step was to generate random values for each process evaluation indicator. According to the parameters of the normal distribution *A*, the random value $\mathrm{DM}_{\mathrm{n,j,k}}$ of each process evaluation indicator was generated by random simulation, where *k* = 1, ..., *K*, and *K* is the random simulation time. Each process evaluation indicator in all experiments simulates the generation of *K* random values and generates a total of $\mathrm{N} \times \mathrm{J} \times \mathrm{K}$ random

The fourth step was to establish mathematical models. When each process evaluation indicator in all experiments generates one random value, *N* models can be established. When there were *K* random simulations, a total of $\mathrm{N} \times \mathrm{K}$ models were established.

The fifth step was to predict the results using the obtained models. Exhaustive research was used to predict the results of the evaluation indicators. Discrete values within the research range of process parameters were taken, and the number of discrete values was determined by the calculation steps. Assuming that there were 3 process parameters in total and that discrete values of *P*, *Q*, and *R* were taken, there were $\mathrm{P} \times \mathrm{Q} \times \mathrm{R}$ process parameter combinations after pairwise matching. A specific combination of process parameters was $(\mathrm{p}, \mathrm{q}, \mathrm{r})$, where $p = 1, \ldots, P$; $q = 1, \ldots, Q$; $r = 1, \ldots, R$. Substituting each specific combination $(\mathrm{p}, \mathrm{q}, \mathrm{r})$ into $\mathrm{N} \times \mathrm{K}$ models for calculation, a total of $\mathrm{N} \times \mathrm{K}$ prediction results for *N* process evaluation indicators were obtained.

The sixth step was to count the probability value. For a certain combination of process parameters $(\mathrm{p}, \mathrm{q}, \mathrm{r})$, when substituting the *N* models obtained from the *k*th random simulation, the predicted values of all *N* process evaluation indicators may meet the limits of all process evaluation indicators, and they may be partially compliant or not at all. Counting the number of times *m* that the combination $(\mathrm{p}, \mathrm{q}, \mathrm{r})$ can meet the limit of all process evaluation indicators in all *K* random simulations, $\mathrm{m} \leq \mathrm{K}$ must hold. The probability $\mathrm{Prob}_{\mathrm{p,q,r}} = m/K$ of reaching the standard corresponding to the combination $(\mathrm{p}, \mathrm{q}, \mathrm{r})$ was calculated, and the value of *Prob* was in the interval [0, 1].

The seventh step was to count the parameter combinations that met the probability of reaching the standard. The threshold *T* of the probability of reaching the standard is set, and generally, $\mathrm{T} \geq 0.80$. If $\mathrm{Prob}_{\mathrm{p,q,r}} \geq \mathrm{T}$, then the combination of the process parameters $(\mathrm{p}, \mathrm{q}, \mathrm{r})$ was the parameter combination that met the probability of reaching the standard. All the process parameter combinations that satisfied $\mathrm{Prob}_{\mathrm{p,q,r}} \geq \mathrm{T}$ were counted, and the parameter combination that met the probability of reaching the standard was obtained.

The eighth step was to count the easily controllable CPP combinations that met the probability of reaching the standard within the variation range of the noise parameter and to calculate the design space. If there were 3 easily controllable CPPs, the combination was $(\mathrm{x}, \mathrm{y}, \mathrm{z})$. From the parameter combinations that met the probability of reaching the standard, it was calculated that the easily controllable CPP combinations satisfied the parameter combination of $\mathrm{x}=\mathrm{x}$, $\mathrm{y}=\mathrm{y}$, $\mathrm{z}=\mathrm{z}$ within the variation range of the noise parameter. Then, the easily controllable CPP combinations within the variation range of noise parameter are obtained and the design space was calculated.

The calculation steps of the CD, EC, ECR, St, and RT were 0.20, 0.20, 0.20, 0.16, and 0.20, respectively. The acceptable probability of the design space was set as 0.80, the simulation was performed 5000 times to calculate reliable probability values, and all calculations were carried out using MATLAB (R2020a, Version 9.8, MathWorks Inc., USA).

Results

Experimental results of the EPP

The results of the evaluation indicators and the retention rate of the index components of the 28 experiments are shown in Table 3 and Fig. 2. The extracted amount of HSYA was 3.81 ~ 7.77 mg/g, that of total flavonoids was 1.92 ~ 3.38 mg/g, and that of total solids was 141.84 ~ 198.66 mg/g. The purity of HSYA in the supernatant was 2.68%~4.04%, which was significantly improved compared with the 1.43%~1.85% purity of the concentrated extract before ethanol precipitation. The purity of the total flavonoids in the supernatant was 1.31%~1.79%, which was higher than the 0.65%~0.83% total flavonoid purity before ethanol precipitation, showing that while removing impurities after ethanol precipitation, the effective components of the supernatant were enriched. In terms of ingredient retention, the retention rates of HSYA and total flavonoids were 24.5%~60.3% and 27.9%~57.8%, respectively, which were equivalent. After ethanol precipitation, the HSYA component was lost, possibly because HSYA solubility in the supernatant is small.

Table 3
The results of experimental design

No.	Extracted amount (mg/g)			Purity (%)		Retention rate (%)	
	HSYA	Total flavonoids	Total solids	HSYA	Total flavonoids	HSYA	Total flavonoids
1	5.69	2.85	171.42	3.32	1.66	41.15	42.18
2	6.55	3.00	178.39	3.67	1.68	44.43	41.70
3	3.81	1.92	141.84	2.68	1.35	24.76	27.88
4	6.77	3.09	184.24	3.67	1.68	52.44	52.76
5	5.12	2.64	177.52	2.88	1.48	33.40	38.33
6	7.48	3.35	187.38	3.99	1.79	54.49	53.73
7	5.32	2.54	173.66	3.06	1.46	32.79	34.96
8	5.75	2.69	155.68	3.69	1.73	43.69	45.00
9	4.85	2.64	160.59	3.02	1.64	30.58	37.05
10	7.15	3.08	189.33	3.78	1.63	55.43	52.47
11	4.87	2.24	169.11	2.88	1.33	32.10	32.99
12	7.11	3.16	188.08	3.78	1.68	54.77	53.56
13	5.17	2.68	164.54	3.14	1.63	32.63	37.75
14	5.84	2.49	164.97	3.54	1.51	43.06	40.46
15	5.98	3.17	198.66	3.01	1.60	37.17	43.96
16	6.54	3.05	180.19	3.63	1.69	48.34	49.57
17	3.98	2.10	147.89	2.69	1.42	24.47	28.81
18	6.12	2.64	160.55	3.81	1.65	46.04	43.74
19	4.92	2.78	175.38	2.80	1.59	30.33	38.24
20	7.57	3.27	193.73	3.91	1.69	55.71	52.96
21	4.00	1.96	148.82	2.69	1.31	26.14	28.48
22	7.77	3.38	192.25	4.04	1.76	60.25	57.75
23	4.72	2.55	161.21	2.93	1.58	29.71	35.83
24	7.16	3.25	186.24	3.85	1.74	51.97	51.84
25	5.95	2.83	176.00	3.38	1.61	40.99	40.01
26	6.44	3.17	183.62	3.51	1.73	46.28	46.69
27	6.07	2.82	173.67	3.50	1.62	42.16	40.09
28	6.08	3.05	173.48	3.50	1.76	41.37	42.59

The identification of CPPs

To comprehensively study the influence of process parameters on the evaluation indicators of the EPP, the CPPs were identified by establishing multiple linear regression models. The regression coefficients and variance analysis of the multiple linear regression models are shown in Table 4. RT is a noise parameter and a CPP, which affects the evaluation indicators of ethanol precipitation and has a positive relationship; that is, as RT increases, the index value of each evaluation indicator increases. The extracted amount of HSYA increased with increasing CD, for other evaluation indicators, the changing trend was opposite to the CD. The extracted amount and purity of HSYA and total flavonoids increased with decreasing ECR. The extraction amount and purity of HSYA and the extraction amount of total flavonoids and total solids decreased with increasing EC. Therefore, through multiple linear regression analysis, the CD, EC, ECR, St, and RT were the CPPs of the *Carthami Flos* EPP, and CT, SS, EAt, and Rt after ethanol precipitation were less important process parameters.

Process modelling of ethanol precipitation

Before establishing mathematical models between CPPs and evaluation indicators, the purity of HSYA and total flavonoids was subjected to square root arcsine transformation. The regression coefficients and variance analysis of each mathematical model are shown in Table 5. The significance level P value of all models was less than 0.0001, indicating that the models were significant. The coefficient of determination of each fitted model was greater than 0.81, which means that most of the data variation can be explained by these models.

Based on the regression models, the influence of CPPs on evaluation indicators was evaluated. The relationship between independent variables and dependent variables was expressed by contour plots, as shown in Fig. 3- Fig. 7. In addition to the two parameters examined in each 2-D contour map, the other parameter levels were fixed as the centre point level. It can be seen from Fig. 3 that reducing the EC, reducing ECR, and allowing the solution to stand at a higher temperature will help increase the extracted amount of HSYA in the supernatant. The influence on the total flavonoid content is shown in Fig. 4. The lower the concentration of ethanol added, the smaller ECR, and the higher the total flavonoid content in the supernatant. According to Fig. 5, increasing EC and lowering the RT were beneficial to the precipitation of solid impurities. Figure 6 and Fig. 7 show that as the CD increased and the RT decreased, the purity of HSYA decreased. The higher CD, the lower the total flavonoid purity was in the supernatant.

Table 4
Regression coefficient and variance analysis of the multiple linear regression model

Parameters	Y_1		Y_2		Y_3		Y_4		Y_5	
	Coefficients	*P* value	Coefficients	*P* value	Coefficients	*P* value	Coefficients	*P* value	Coefficients	*P* value
Constant	5.88	-	2.80	-	173.52	-	3.37	-	1.61	-
X_1	1.02	< 0.0001**	-0.2838	< 0.0001**	-7.43	0.0059**	-0.4498	< 0.0001**	-0.0975	< 0.0001**
X_3	-0.3784	0.0006**	-0.1261	0.0342*	-6.57	0.0133*	-0.1001	0.0001**	-	-
X_4	-0.2576	0.0128*	-0.1157	0.0503	-	-	-0.0835	0.0008**	-0.0396	0.0387*
X_7	-	-	-	-	-	-		-	-0.0321	0.0885
X_8	0.1942	0.0535	0.1358	0.0236*	4.37	0.0881	0.0422	0.0641	0.0431	0.0255*
P value	< 0.0001**		< 0.0001**		0.0023		< 0.0001**		< 0.0001**	

*p < 0.05; **p < 0.01.

Table 5
Regression coefficient and variance analysis of the quadratic multiple regression model

Parameters	Y_1		Y_2		Y_3		ArcSin (Sqrt (Y_4))		ArcSin (Sqrt (Y_5))	
	Coefficients	*P* value	Coefficients	*P* value	Coefficients	*P* value	Coefficients	*P* value	Coefficients	*P* value
Constant	6.189	-	2.985	-	177.6	-	0.1879	-	0.1298	-
X_1	-1.0238	< 0.0001**	-0.2838	< 0.0001**	-7.4283	0.0002**	-0.0126	< 0.0001**	-0.0039	< 0.0001**
X_3	-0.3784	< 0.0001**	-0.1261	0.0016**	-6.5712	0.0006**	-0.0028	< 0.0001**	-	-
X_4	-0.2576	< 0.0001**	-0.1157	0.0033**	-	-	-0.0022	< 0.0001**	-0.0016	0.0128*
X_7	-	-	-	-	-	-	-	-	-0.0013	0.0346*
X_8	0.1942	0.0014**	0.1358	0.0008**	4.3687	0.0129*	0.0013	0.0017**	0.0018	0.0058**
X_1X_4	-	-	-	-	-	-	0.0010	0.0120*		
X_1X_7	-	-	-	-	-	-	-	-	-0.0012	0.0572
X_1X_8	-	-	-	-	-	-	0.0014	0.0010**	0.0012	0.0599
X_3X_4	0.2068	0.0013**	0.0915	0.0235*	5.1539	0.0091**	-	-	-	-
X_3X_8	0.3248	< 0.0001**	0.1747	0.0001**	8.7882	< 0.0001**	-	-	-	-
X_4X_8	-	-	-	-	-	-	-0.0009	0.0354*	-	-
X_1^2	-	-	-	-	13.2092	0.0128*	-0.0046	< 0.0001**	-0.0036	0.0098**
X_3^2	-0.3871	0.0027**	-	-	-8.2773	0.0915	-	-	-	-
X_8^2	-	-	-0.2348	0.0060**	-10.0961	0.0409*	-	-	-	-
P value	< 0.0001**		< 0.0001**		< 0.0001**		< 0.0001**		< 0.0001**	
R^2	0.9637		0.8803		0.8165		0.9874		0.8102	

*p < 0.05; **p < 0.01.

Design space development and verification

The extracted amount of total solids in the supernatant can reflect the impurity removal effect of the EPP, so the upper limit of the extracted amount of total solids was set at 185 mg/g. HSYA and total flavonoids are the effective ingredients of *Carthami Flos*, and their extracted amount and purity were set at a lower limit. Therefore, the lower limits of the extracted amount of HSYA and total flavonoids were 5.0 mg/g and 2.3 mg/g, respectively, and the lower limits of the purity of HSYA and total flavonoids were 2.8% and 1.4%, respectively. According to the design space calculation method based on the probability of reaching the standard proposed in this work, the parameter combination with a probability of reaching the standard exceeding 80% belongs to the design space. The easily controllable CPP combinations and the probability of reaching the standard in the design space are shown in Additional file 2: Table S1 and the design space is shown in Fig. 8.

One point inside and one point outside the design space were selected for the verification of the design space, and 3 experiments were performed for each verification point. The calculation formula of the average relative deviation (ARD) value is as follows:

$$\mathrm{ARD} = \frac{|\mathrm{EV}-\mathrm{PV}|}{\overline{\mathrm{EV}}} \times 100\% \quad (7)$$

where EV and PV refer to the experimental value and the predicted value, respectively. The specific conditions and results of the verification experiment are shown in Table 6. It can be seen from the verification results that the experimental value and the predicted value were basically in agreement, indicating that the predictive performance of the models was good. Carrying out *Carthami Flos* ethanol precipitation in the design space [illegible] ons. At the verification point outside the design space, the extracted amount of the total solid

content obtained from Experiment V6 did not reach 185 mg/g. The above results showed that the operation in the design space can ensure a better process quality of ethanol precipitation.

Table 6
The results of the verification experiments

CPPs or evaluation indicators		Inside the design space			Outside the design space		
		V1	V2	V3	V4	V5	V6
X_1 (g/cm^3)		1.23	1.23	1.23	1.19	1.19	1.19
X_3 (%)		94.8	94.8	94.8	95.0	95.0	95.0
X_4 (v/v)		2.96	2.96	2.96	2.96	2.96	2.96
X_7 (min)		120	120	120	90	90	90
X_8 (°C)		2.0	6.0	10.0	2.0	6.0	10.0
Calculated probability (%)		93.1	94.0	96.0	97.0	52.0	76.0
Extracted amount of HSYA	PV (mg/g)	5.88	6.01	6.14	6.61	6.80	7.00
	EV (mg/g)	5.94	6.07	6.39	6.51	6.98	6.94
	ARD (%)	0.98	0.92	3.95	1.50	2.56	0.86
Extracted amount of total flavonoids	PV (mg/g)	2.57	2.90	2.77	2.75	3.12	3.02
	EV (mg/g)	2.68	2.70	2.85	2.94	3.02	3.43
	ARD (%)	4.11	7.47	2.76	6.67	3.23	12.02
Extracted amount of total solids	PV (mg/g)	163.3	178.4	168.9	173.1	189.1	181.8
	EV (mg/g)	160.8	164.2	169.1	170.8	183.6	189.7
	ARD (%)	1.57	8.65	0.13	1.37	3.01	4.17
Purity of HSYA	PV (mg/g)	3.42	3.44	3.47	3.68	3.64	3.72
	EV (mg/g)	3.69	3.69	3.78	3.81	3.80	3.66
	ARD (%)	7.34	6.77	8.17	3.42	4.26	1.75
Purity of HSYA	PV (mg/g)	1.57	1.61	1.66	1.65	1.67	1.69
	EV (mg/g)	1.67	1.64	1.68	1.72	1.65	1.81
	ARD (%)	5.92	1.96	1.66	4.53	1.24	6.79

The operation space for the production process was also calculated, and the optimized ranges of CPPs were as follows: CD was 1.20~1.26 g/cm^3, EC was 94.6-95.8%, ECR was controlled at 2.64-2.84 v/v, St was 60-120 min. The RT was used as a noise parameter, and its allowable range was from 2-10°C. The allowable range of other process parameters was the entire experimental research range; that is, CT was 20~30°C, SS was 120~160 rpm, EAt was 80~100 min, and Rt was 24~48 h.

Continuous improvement strategy

ICH Q8 (R2) [16] encourages continuous improvement of the drug production process and continuous improvement of process performance and product quality. In this work, with the collection of production data, the mathematical model and design space were updated, making the design space more reliable. In this work, the models were rebuilt by combining the 6 sets of data from the previous verification experiment with the 28 sets of data in DSD. As for the experimental point N, the CPP conditions were X_1 = 1.23 g/cm^3, X_3 = 94.0%, X_4 = 2.6 v/v, X_7 = 60 min, and X_8 = 10.0°C. The calculated probability of reaching the standard was 77.0%, which was lower than the design space threshold of 80%. The regression coefficient and variance analysis of the regression model after continuous improvement are shown in Additional file 3: Table S2. After rebuilding the models, it was found that the probability of reaching test point N was 85.7%, which exceeded 80%. Then, a set of validation tests was performed, and the evaluation indicators are shown in Table 7. All of the evaluation indicators complied with the limits. This finding showed that the recalculated probability value is more reliable after obtaining new data. Therefore, we reupdated the combination of easily controllable parameters in the design space, and the results are

e S1, Table S3 presents more easily controllable parameter combinations. It is thus

recommended to continuously collect new process data during daily production, continuously improve the model based on the data, and regularly update and maintain the design space to improve its reliability.

Table 7
The results of experimental point N

Evaluation indicators	Extracted amount (mg/g)			Purity (%)	
	HSYA	Total flavonoids	Total solids	HSYA	Total flavonoids
PV (mg/g)	6.51	3.04	167.2	3.76	1.80
EV (mg/g)	6.63	2.66	173.3	3.82	1.54
ARD (%)	1.70	14.31	3.49	1.73	17.04

Discussion

In the literature, commercial software such as MODDE, Minitab, or Design Expert was used to calculate the design space. In this work, a self-coded program was used to calculate the effects of noise parameters for building the design space. The exhaustive search-Monte Carlo method was used to calculate the design space, which required a lot of computation. To improve the convenience of data calculation, software that can consider noise parameters should be developed for calculating design space in the future. The mathematical models between the CPPs and evaluation indicators were established with the quadratic models. However, the determination coefficients of some models were not large enough, for example, less than 0.90. Recently, some machine learning models were used in the researchers of TCMs, such as convolution neural networks[31, 32]. These models can also be considered in the optimization of pharmaceutical process parameters.

Conclusion

In this work, the design space calculation method in the presence of noise parameters was proposed, the models and design space were updated based on new data. Optimization of the *Carthami Flos* EPP was studied based on the QbD concept. The extracted amount of HSYA, extracted amount of total flavonoids, extracted amount of the total solids, purity of HSYA, and the purity of total flavonoids were regarded as evaluation indicators. CD, EC, ECR, St, and RT were identified as CPPs using DSD. Quantitative models between the CPPs and evaluation indicators were developed, the determination coefficients were higher than 0.81. Decreasing CD and EC could increase the extraction amount of total flavonoids and total solids and the purity of HSYA. Then the RT was considered as the noise parameter because it is easily affected by the seasons. The design space considering a noise parameter was calculated and verified. The combination of easily controllable CPPs in the design space was obtained. The calculated operation space of the EPP was as follows: CD was 1.20 ~ 1.26 g/cm^3, CT was 20 ~ 30°C, EC was 94.6–95.8%, ECR was controlled at 2.64–2.84 v/v, SS was 120 ~ 160 rpm, EAt was 80 ~ 100 min, St was 60–120 min, the allowable ranges of RT were from 2–10°C, and Rt was 24 ~ 48 h. In this operating space, the quality of the supernatant could be guaranteed.

According to the data of the new batches from the verification experiments, the models and design space were updated. The calculated probabilities of the models built with more data were more accurate and reliable, and the design space was more reliable. The methods provided in this work can similarly be used for the continuous improvement of industrial production.

Abbreviations

TCM: traditional Chinese medicine; CPPs: critical process parameters; QbD: quality by design; DSD: definitive screening design; HSYA: hydroxysafflor yellow A; EPP: ethanol precipitation process; *RSD*: relative standard deviation; CD: density of the concentrated extract; CT: temperature of the concentrated extract; EC: ethanol concentration; ECR: volume ratio of ethanol to and the concentrated extract; SS: stirring speed; EAt: time of ethanol addition; St: stirring time after ethanol addition; RT: refrigeration temperature; Rt: refrigeration time.

Declarations

Acknowledgements

Not applicable.

Authors' contributions

YNT performed the experiments, drafted the manuscript, and prepared tables and figures; HBQ contributed to revising the article, and XCG conceived and designed the research work and finally confirmed the manuscript. All authors read and approved the final manuscript.

Funding

The authors would like to acknowledge the support of the National S&T Major Project of China (2018ZX09201011-002), the Basic Public Welfare Research Program of Zhejiang Province (LGG18H280001), and the National Project for Standardization of Chinese Materia Medica (ZYBZH-C-GD-04).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its additional files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. State Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China. Vol. 1. Beijing: China Medical Science Press; 2020.
2. Zhang L, Gong XC, Qu HB. Optimizing the alcohol precipitation of Danshen by response surface methodology. Sep Purif Technol. 2013;48(6):977–83.
3. Zhang L, Yan BJ, Gong XC, Yu LX, Qu HB. Application of quality by design to the process development of botanical drug products: a case study. Aaps Pharmscitech. 2013;14(1):277–86.
4. Gong XC, Yan AY, Qu HB. Optimization for the ethanol precipitation process of botanical injection: indicator selection and factor influences. Sep Purif Technol. 2014;49(4):619–26.
5. Sun MF, Yang JY, Cao W, Shao JY, Wang GX, Qu HB, et al. Critical process parameter identification of manufacturing processes of Astragali Radix extract with a weighted determination coefficient method. Chin Herb Med. 2020;12(2):125–32.
6. Luo Y, Li WL, Huang WH, Liu XH, Song YG, Qu HB. Rapid quantification of multi-components in alcohol precipitation liquid of Codonopsis Radix using near infrared spectroscopy (NIRS). J Zhejiang Univ-SC B. 2017;18(5):383–92.
7. Huang HX, Qu HB. In-situ monitoring of saccharides removal of alcohol precipitation using near-infrared spectroscopy. J Innov Opt Heal Sci. 2018;11(05):1850027.
8. Ma LJ, Liu DH, Du CZ, Lin L, Zhu JY, Huang XG, et al. Novel NIR modeling design and assignment in process quality control of Honeysuckle flower by QbD. Spectrochim Acta A: Mol Biomol Spectrosc. 2020;242:118740.
9. Pan J, Shao J, Qu H, Gong X. Ethanol precipitation of Codonopsis Radix concentrate with a membrane dispersion micromixer. J Clean Prod. 2020;251:119633.
10. Pan JJ, Tai YN, Qu HB, Gong XC. Optimization of membrane dispersion ethanol precipitation process with a set of temperature control improved equipment. Sci Rep. 2020;10(1):19010.
11. Shao F, Yu MY, Jiang ML, Shang Y, Yang M, Liu RH. Effect of Stirring Speed on Fractal Dimension of Ethanol Precipitation of Astragalus Granules and Alcohol Precipitation Effect. J Chin Med Mater. 2019;42(03):612–6.
12. Jiang ML, Zhang XY, Shao F, Shang Y, Yang M, Liu RH, et al. Effect of ethanol to material ratio on ethanol precipitation and sediment morphology of Shuanghuanglian preparation. Chin Tradit Herb Drugs. 2020;51(19):4954–9.
13. Chen T, Gong XC, Zhang Y, Chen HL, Qu HB. Optimization of a chromatographic process for the purification of saponins in Panax notoginseng extract using a design space approach. Sep Purif Technol. 2015;154:309–19.
14. Shao JY, Cao W, Qu HB, Pan JJ, Gong XC. A novel quality by design approach for developing an HPLC method to analyze herbal extracts: A case study of sugar content analysis. PLoS One. 2018;13(6):e0198515.
15. Pan JJ, He SY, Zheng JY, Shao JY, Li N, Gong YQ, et al. The development of an herbal material quality control strategy considering the effects of manufacturing processes. Chin Med. 2019;14(1):38.
16. International conference on harmonisation of technical requirements for registration of pharmaceuticals for Human Use. Q8 (R2). https://database.ich.org/sites/default/files/Q8_R2_Guideline.pdf. Accessed 9 Mar. 2021.
17. Watson TJN, Nosal R, Lepore J, Montgomery F. Misunderstanding Design Space: a Robust Drug Product Control Strategy Is the Key to Quality Assurance. J Pharm Innov. 2018;13(3):283–5.
18. Hersleth M, Mevik BH, Naes T, Guinard JX. Effect of contextual factors on liking for wine - use of robust design methodology. Food Qual Prefer. 2003;14(7):615–22.

19. Zhang JQ, Liu SL, Cai HR, Dai XZ, Chen YH, Jin ZL, et al. Efficacy and safety of Guhong injection for treating coronary microvascular disease: study protocol for a randomized controlled trial. Trials. 2020;21(1):75.
20. Sun ZZ, Jin HT, Zhou HF, Yu L, Wan HT, He Y. Guhong Injection promotes fracture healing by activating Wnt/beta-catenin signaling pathway in vivo and in vitro. Biomed Pharmacother. 2019;120:109436.
21. Ai JC, Wan HT, Shu MC, Zhou HF, Zhao T, Fu W, et al. Guhong injection protects against focal cerebral ischemia–reperfusion injury via anti-inflammatory effects in rats. Arch Pharm Res. 2017;40(5):610–22.
22. Yang XM, Li Y, Chen L, Xu MG, Wu JB, Zhang P, et al. Protective effect of hydroxysafflor yellow A on dopaminergic neurons against 6-hydroxydopamine, activating anti-apoptotic and anti-neuroinflammatory pathways. Pharm Biol. 2020;58(1):686–94.
23. Bai X, Wang WX, Fu RJ, Yue SJ, Gao H, Chen YY, et al. Therapeutic Potential of Hydroxysafflor Yellow A on Cardio-Cerebrovascular Diseases. Front Pharmacol. 2020;11:01265.
24. Han DA, Wei J, Zhang R, Ma WH, Shen C, Feng YD, et al. Hydroxysafflor yellow A alleviates myocardial ischemia/reperfusion in hyperlipidemic animals through the suppression of TLR4 signaling. Sci Rep. 2016;6:35319.
25. Meng YQ, Du ZY, Li Y, Gao P, Song JY, Lu YY, et al. The synergistic mechanism of total saponins and flavonoids in Notoginseng-Safflower pair against myocardial ischemia uncovered by an integrated metabolomics strategy. Biomed Pharmacother. 2020;130:110574.
26. Gong X, Wang S, Li Y, Qu H. Separation characteristics of ethanol precipitation for the purification of the water extract of medicinal plants. Sep Purif Technol. 2013;107:273–80.
27. Gong XC, Chen HL, Chen T, Qu HB. Unit operation optimization for the manufacturing of botanical injections using a design space approach: a case study of water precipitation. PLoS One. 2014;9(8):e104493.
28. Tai YN, Shen JC, Luo Y, Qu HB, Gong XC. Research progress on the ethanol precipitation process of traditional Chinese medicine. Chin Med. 2020;15(1):84.
29. Tyssedal J, Chaudhry MA. The choice of screening design. Appl Stochastic Models Bus Ind. 2017;33(6):662–73.
30. Guhong Injection. National Bureau of Quality and Technical Supervision; 2012.
31. Yan X, Zhang S, Fu H, Qu HB. Combining convolutional neural networks and on-line Raman spectroscopy for monitoring the Cornu Caprae Hircus hydrolysis process. Spectrochim Acta A Mol Biomol Spectrosc. 2020;226:117589.
32. Yan X, Fu H, Zhang S, Qu HB. Combining convolutional neural networks and in-line near-infrared spectroscopy for real-time monitoring of the chromatographic elution process in commercial production of notoginseng total saponins. J Sep Sci. 2020;43(3):663–70.

Figures

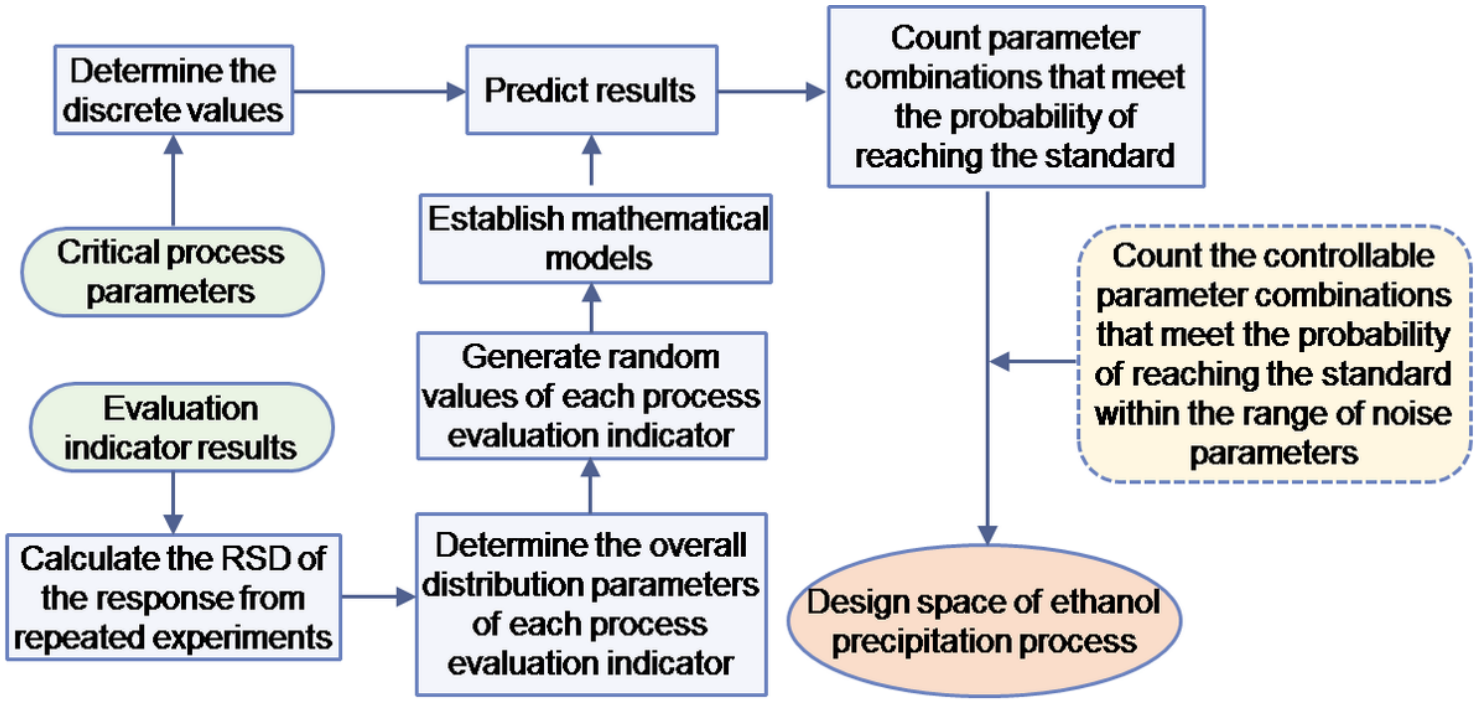

Figure 1

Design space calculation process of the probability of reaching standards considering a noise parameter

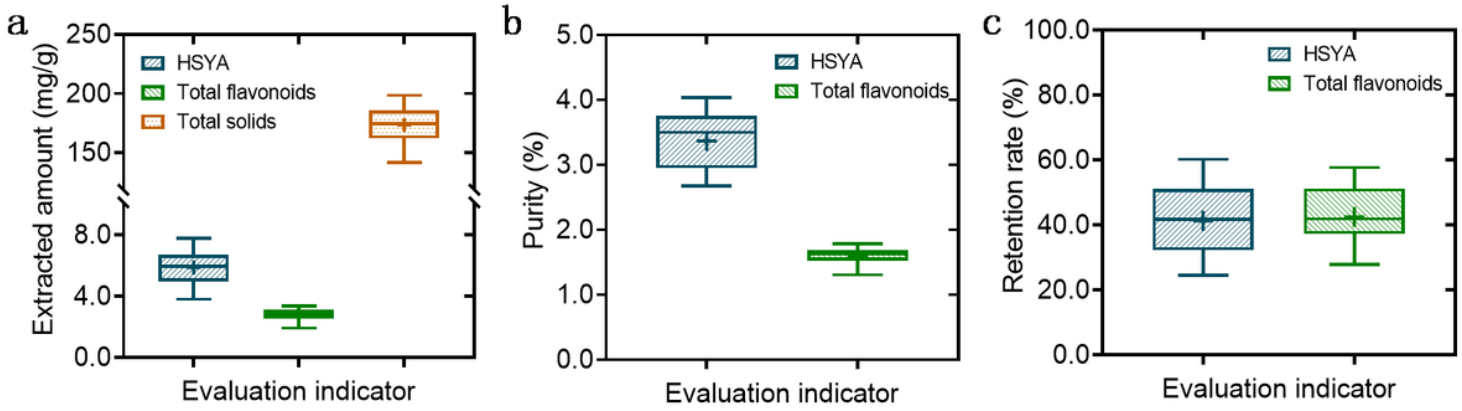

Figure 2

Box diagram of evaluation indicator results of the EPP

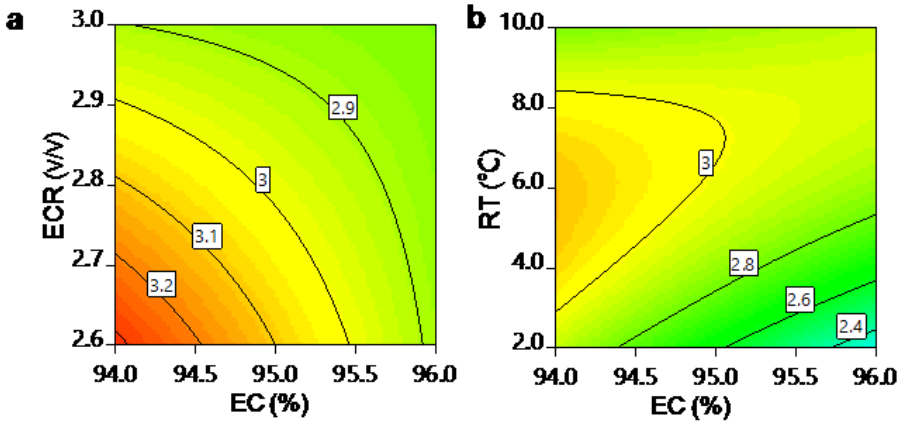

Figure 3

Contour plot of the extracted amount of HSYA

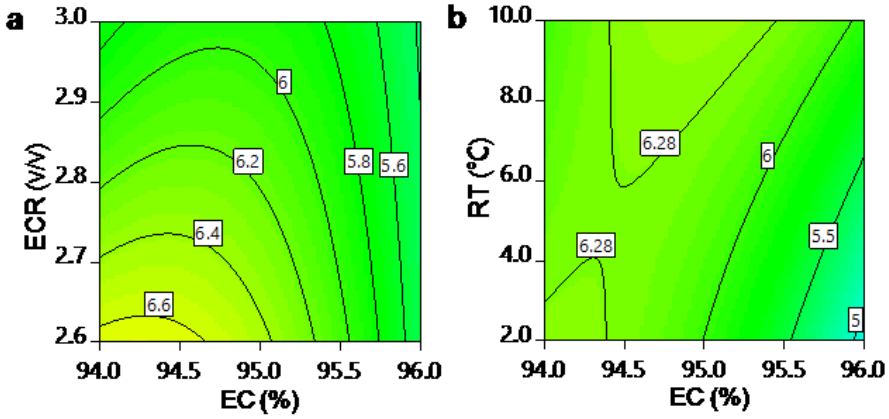

Figure 4

Contour plot of the extracted amount of total flavonoids

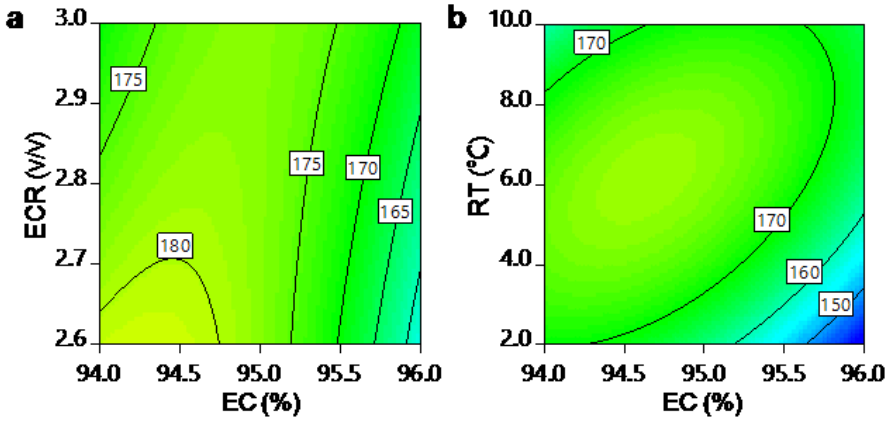

Figure 5

Contour plot of the extracted amount of total solids

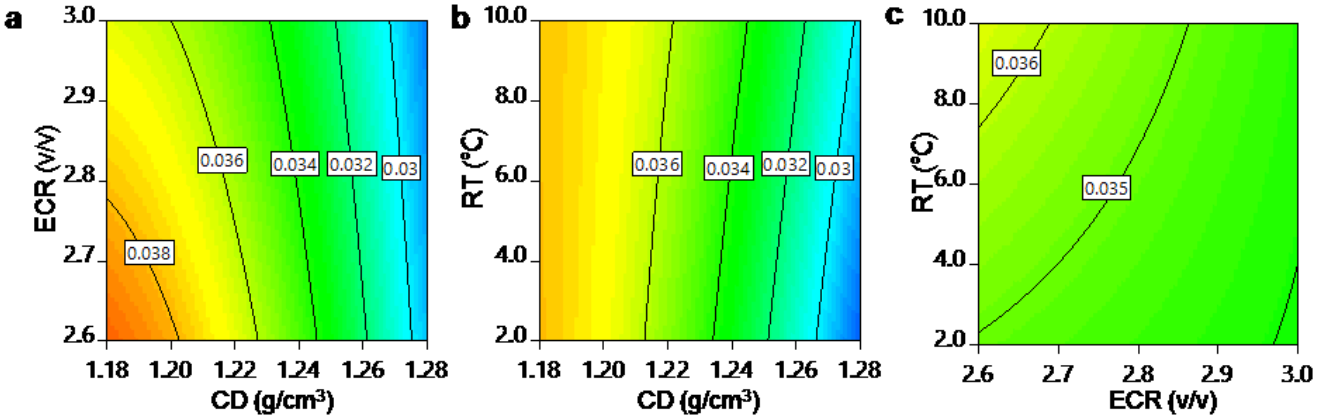

Figure 6

Contour plot of the purity of HSYA

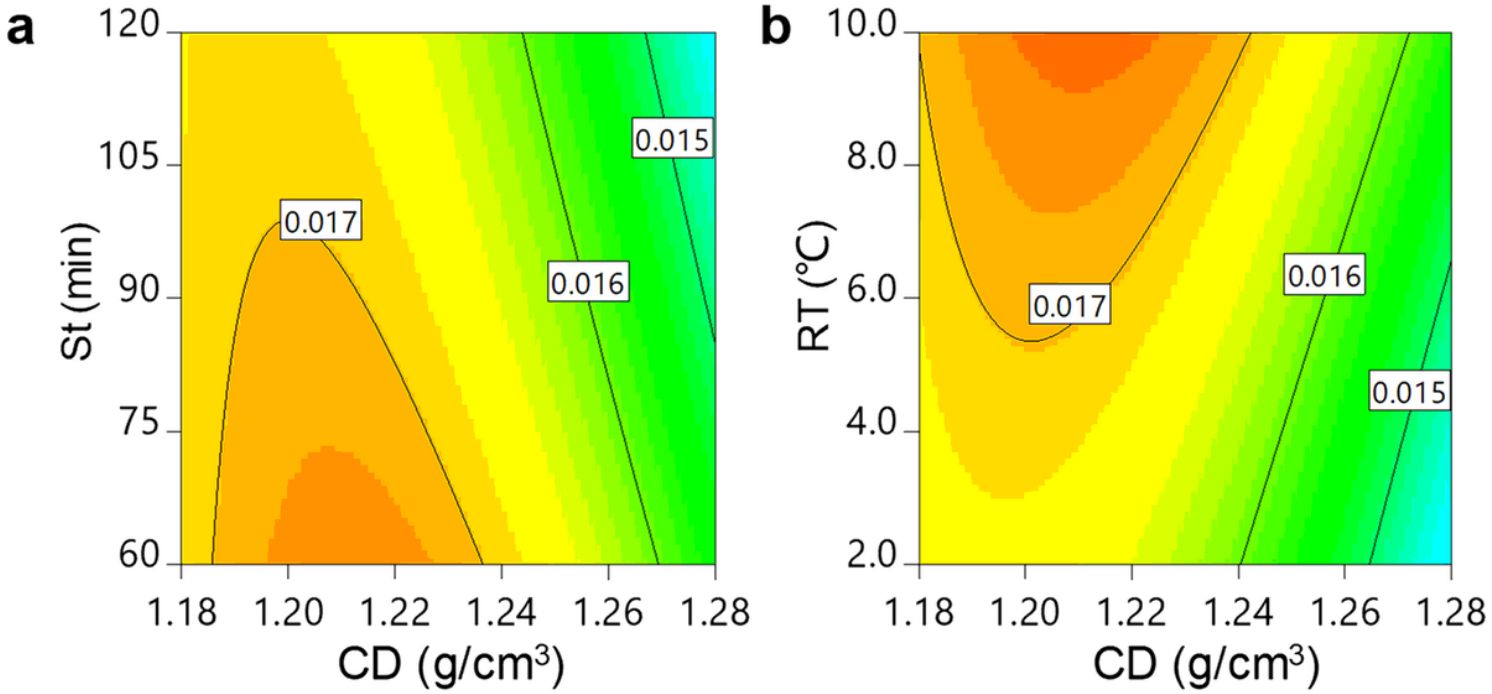

Figure 7

Contour plot of the purity of total flavonoids

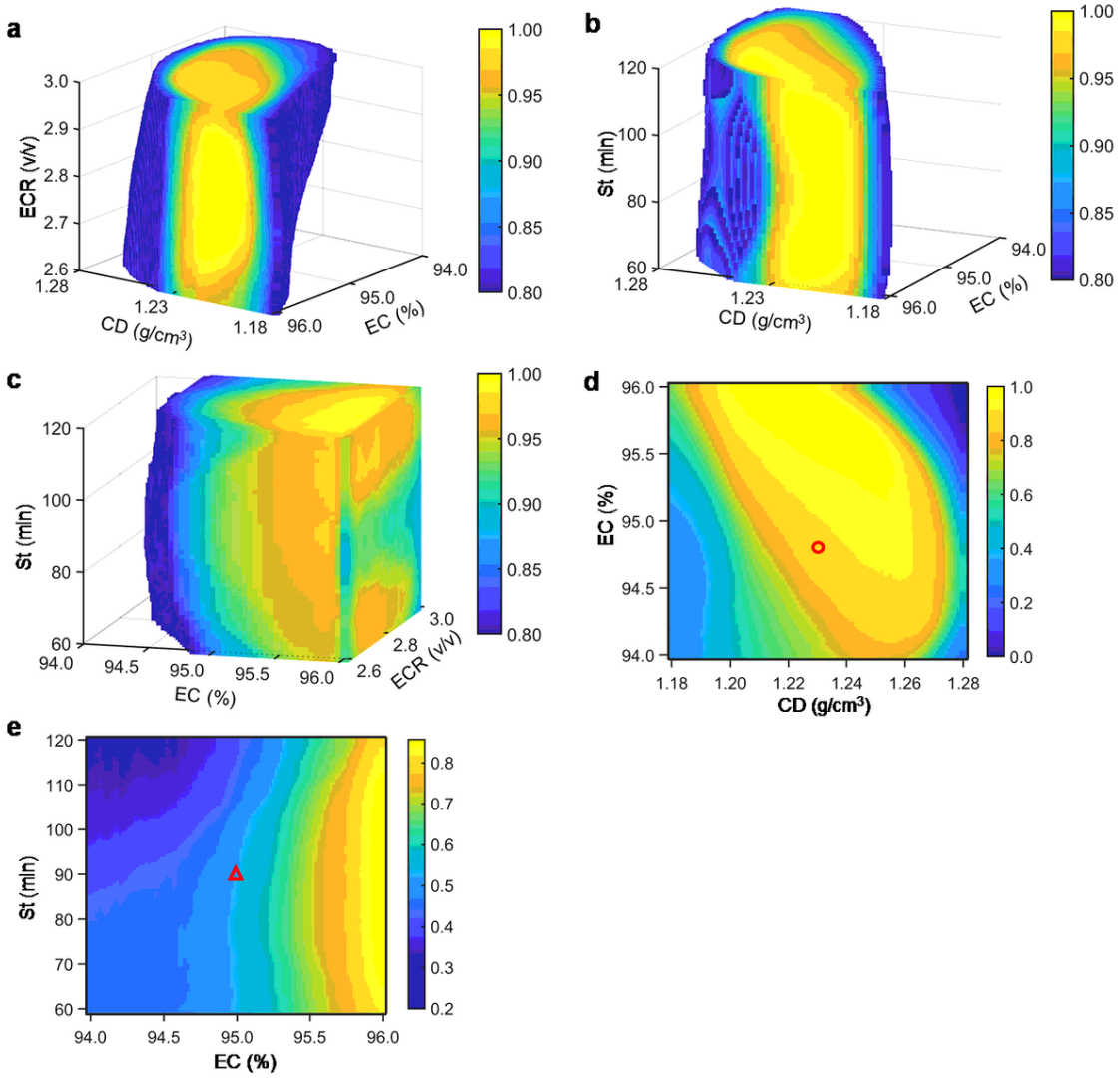

Figure 8

Design space and verification points a St =90 min, RT=6 ℃. b ECR =2.8 v/v, RT =6℃. c CD =1.23 g/cm3, RT =6 ℃. d ECR =2.96 v/v, St =120 min. e CD =1.19 g/cm3, ECR =2.96 v/v. ● Verification points in the design space. △ Verification points outside the design space

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile2.xlsx
- Additionalfile3.docx
- Additionalfile4.xlsx