Optimization of accelerated solvent extraction of fatty acids from *Coix* seeds using chemometrics methods

Xing Liu[1, 2], Kai Fan[1], Wei-Guo Song[1], Zheng-Wu Wang [2]*

[1] Institute for Agri-products Standards and Testing Technology, Shanghai Key Laboratory of Protected Horticultural Technology, Shanghai Academy of Agricultural Science, Shanghai 201403, China

[2] Department of Food Science & Technology, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract

This study investigated the optimization of accelerated solvent extraction (ASE) of fatty acids (FAs) from three *Coix* seeds (small *Coix* seed, SCS; big *Coix* seed, BCS; translucent *Coix* seed, TCS) by chemometrics methods. Partial least-squares regression (PLSR) and backpropagation neural network (BPNN) were applied to build models that reflect the relationship between content of FAs and extraction conditions (temperature, time, and extraction solvent). Genetic algorithms (GAs) and particle swarm optimization (PSO) were utilized to optimize the combination of extraction conditions. The composition of FAs was analysed by gas chromatography-mass spectrometry (GC-MS). The PLSR models could reflect the relationship of FA content in both BCS and SCS and extraction conditions well, while the BPNN model was more suitable for TCS. The optimal extraction conditions for BCS and SCS were obtained by GAs, whereas those of TCS were obtained by PSO. The FA compositions

*corresponding author
Department of Food Science & Technology, School of Agriculture and Biology, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China
E-mail: zhengwuwang@sjtu.edu.cn, telephone and fax numbers: +86-2134205748

23   of the three *Coix* seeds exhibited differences. The results show that ASE combined

24   with chemometrics methods can rapidly and effectively obtain the optimal conditions

25   for the extraction of FAs from *Coix* seed and there are differences in the extraction

26   conditions and compositions of FAs among different varieties of *Coix* seed, but all the

27   extraction time is shorter than other extractions methods.

28   **Keywords:** *Coix* seed, fatty acids, PLSR, BPNN

29   **Introduction**

30   *Coix* seed is the mature kernel of *Coix lachryma-jobi* L., a grain crop in the

31   Gramineae family, and has long been used as a traditional Chinese medicinal herb and

32   food source. *Coix lachryma-jobi* L. is widely distributed in China, Thailand, Burma,

33   Korea, Japan, and Brazil [1]. There are many reported pharmacological and

34   physiological effects of *Coix* seed, including anti-tumour [2], anti-inflammatory [3],

35   anti-allergic [4], and immunoregulation [5]. These effects result from diverse

36   biologically active components in *Coix* seed [6, 7], which mainly exist in *Coix* seed

37   oil [8], such as coixenolide, coixol, and sterols. *Coix* seed oil is mainly composed of

38   the fatty acids (FAs), and the content of FAs can reflect the yield of the extracted oil

39   and the content of other active ingredients to a certain degree. Moreover, the kinds of

40   FAs have an important impact on the nutritive value of *Coix* seed oil. Due to its many

41   benefits, it is reasonable to pursue the optimization of the extraction conditions of

42   *Coix* seed oil. The common extraction techniques, such as Soxhlet extraction [9],

43   microwave extraction [10], sonication extraction [11], and supercritical fluid

44   extraction [8], are time-consuming and/or complex. The extraction yield of FAs from

45   *Coix* seed is dependent on the following factors: temperature, time, pressure,

46   extraction solvent, particle size, and solid-liquid ratio. Accelerated solvent extraction

47   (ASE), an extraction procedure using organic solvents at high temperatures (elevated

48   temperatures up to 200 °C) and pressures (up to 3000 psi) above the boiling point for

49   shorter time (low to several minutes), can increase target compound solubility, solvent

50   diffusion rate, and mass transfer. It can also decrease solvent viscosity and surface

51   tension, which has been shown to be equivalent to the standard EPA extraction

52   methodology (Method 3545) in terms of precision and recovery [12]. Furthermore, the

53   extraction process of ASE has the advantage that needs less solvent, is automated and

54   quick, and can retain the sample in an oxygen- and light-free environment [13].

55   Currently, many studies have applied the ASE method to extract lipids and FAs from

56   cereal, egg yolk, fish, fish tissue, and chicken muscle [14-17].It has also been reported

57   that the FA composition was not affected by the extraction temperature of ASE [18].

58   Our previous study has proven that the extraction yield of crude fat in *Coix* seed by

59   ASE is not lower than that of Soxhlet extraction or sonication-assisted supercritical

60   fluid extraction [11, 19].However, the extraction process remains to be optimized

61   from the perspective of energy-saving, and there is limited information on how to

62   optimize the extraction conditions (temperature, time, and extraction solvent) of FAs

63   from *Coix* seed by ASE. Thus, a full factorial design (FFD) was applied to design the

64   experiment, chemometrics methods, partial least-squares regression (PLSR) and a

65   backpropagation neural network (BPNN), were used to build the relationship between

66   FAs and extraction conditions, and genetic algorithms (GA) and particle swarm

3

67   optimization (PSO) were utilized to optimize the extraction conditions. The PLSR, as

68   a chemometrics method and linear regression tool, is one of the most widely applied

69   multivariate statistical data analysis method [20].The BPNN is a classical

70   domain-dependent technique for nonlinear system modelling. It is composed of an

71   input layer, hidden layer, and output layer, and works by measuring the output error,

72   calculating the gradient of this error, and adjusting the neural network weights (and

73   biases) in the descending gradient direction. That is, the BPNN is a gradient-descent

74   local search procedure that is expected to stagnate in local optima in complex

75   landscapes [21].The GA and PSO are most popular optimization algorithms, and they

76   employ a population of individuals to solve the problem on hand [22].It has been

77   reported that GA, which are parallel randomly search optimization algorithms, can be

78   successfully applied to identify global optimizations of multidimensional functions by

79   selecting, crossover, and mutation operations [23].The PSO is a stochastic

80   evolutionary computation technique, inspired by the social behaviour of bird flocking

81   [24].Similar to GA, the PSO system is initialized with a population of random

82   solutions and can search for optimum conditions by the updating of generations

83   [25].The content of oils is significantly difference due to the region in which the crop

84   is grown as well as varietal diversity [19].Therefore, it is necessary to study the FA

85   content of different varieties of *Coix* seed. In the present study, temperature, time, and

86   extraction solvent were optimized to extract FAs from different *Coix* seeds by ASE

87   combined with chemometrics methods. The composition of FAs was determined by

88   gas chromatography-mass spectrometry (GC-MS).

89 **Materials and methods**

90 Materials

91 Small *Coix* seeds (SCS, aspect ratio=0.23 cm: 0.26 cm) and big *Coix* seeds (BCS,

92 aspect ratio=0.32 cm: 0.40 cm) were purchased in Anshun Municipality and Guizhou

93 province, China, respectively; translucent *Coix* seeds (TCS, aspect ratio=0.21 cm:

94 0.22 cm) were purchased in Putian Municipality, Fujian province, China. The three

95 categories of *Coix* seed were named in terms of their appearance and size and are

96 shown in Fig.1. Before the experiment, defective granules were removed from all

97 samples, and the seeds were ground until they could pass through a 425-μm mesh

98 sieve. The sieved powders were used in subsequent analysis. All chemicals were

99 purchased from the China National Pharmaceutical Group (Sinopharm, Beijing,

100 China) and were of analytical grade.

101 Crude fat extraction and FA determination

102 Crude fats and FAs of the three categories of *Coix* seed were extracted by an ASE

103 apparatus (ASE 350, Thermo Fisher Scientific, Waltham, MA, USA). *Coix* seed

104 samples of 10.0 g were weighed and poured into a 66-mL zirconium extraction cell

105 with a cellulose filter in the cell outlet. The extraction cell was arranged in the cell

106 tray, and the sample was extracted using a combination of conditions obtained from

107 the FFD experimental design. The automated extraction cycle was as follows: the cell

108 containing sample was prefilled with the degassed extraction solvent (acetone or

109 petroleum ether), pressurized (1600 psi), and then heated. The cycle time varied with

110 the change in temperature (100, 110, 120, or 130 °C). When the temperature was

111   higher than the 130 °C, the colour of *Coix* seed oil became dark. The last step in the

112   cycle was a static period (5, 10, 15, or 20 min). Then, the cell was rinsed with fresh

113   extraction solvent (60% of the extraction cell volume) and purged with a stream of

114   nitrogen. The extraction cycles were repeated twice. The oil was collected into glass

115   vials and concentrated immediately by rotary evaporators (35 °C). The FA content in

116   the concentrated solution was determined by the AOAC method (939.05) and

117   expressed as mg of KOH required neutralizing FAs in 100 g *Coix* seed. The

118   concentrated solution obtained in the optimal extraction condition was evaporated to

119   dryness by water bath, and dried for 1 h at 100 °C ± 5 °C, then cooled down for 0.5h

120   in the desiccator and weighed. The process above was repeated until achieved

121   constant weight. This method was calculated the crude fat content.

122   GC-MS analysis of FAs

123   One gram of *Coix* seed oil was dissolved in 40 mL of n-hexane, then 40 mL 0.4 M

124   KOH-MeOH solution was poured into a test tube, which was vigorously shaken, and

125   the mixture was placed for 30 min. After being fully saponified, 10 mL distilled water

126   was put into the test tube, and the supernatant was used to analyse the composition of

127   FAs by GC-MS (7890A-5975C Agilent, Santa Clara, CA, USA). The GC column was

128   a DB-5ms (30 m × 0.25 mm × 0.25 μm). The carrier gas was helium (99.999%), and

129   the flow rate was 1 mL/min. Both the injector temperature and detector temperature

130   were 280 °C The program sequence of the column temperature was as follows: initial

131   temperature 60 °C, held for 3 min, increased to 300 °C at 5 °C/min, and held for 14

132   min. The MS ion source was electron impact mode at an ionization voltage of 70 eV

133   with an ion source temperature of 230 °C. The full-scanning range of MS was 33–500

134   amu. The results were obtained from the NIST 2011 mass spectral data base.

135   Chemometrics methods and statistical analysis

136   The ASE experiment was designed to consider the factors of temperature (100, 110,

137   120, or 130 °C), extraction time (5, 10, 15, or 20 min) and extraction solvent (acetone

138   or petroleum ether) and was carried out according to the FFD, whose total trial

139   number was 32.

140   The Kennard-Stone algorithm was used to partition the calibration (75%) and

141   validation sets (25%) [26], and the criterion was to select the samples one by one

142   which was the furthest distance from each other in the group, namely, according to the

143   Euclidean distance, so they could spread throughout the multivariate space. Linear

144   models of the FAs extraction were established by PLSR. The latent variables of PLSR

145   were determined by 10-fold cross-validation with the lowest root mean square error of

146   cross validation (RMSECV). The performances of calibration set models were valued

147   by the RMSECV and the coefficient of determination (R^2), and validation set models

148   evaluated by root mean square error of prediction (RMSEP) and R^2 between predicted

149   value and actual value. Nonlinear models of the FA extraction were built by a BPNN,

150   whose input layer nodes were 3, hidden layer nodes were 10, and output layer node

151   was 1. The BPNN models were estimated by RMSEP and the R^2 of validation sets. In

152   this study, the PLSR and BPNN models with the highest R^2, as well as the lowest

153   RMSECV and RMSEP, were considered as the optimal result.

154   Two extreme value searching algorithms, GAs and PSO, were applied to screen the

7

155   optimum extraction conditions (extraction solvent, time, and temperature). Generally,

156   a given problem can be regarded as an individual coded by chromosome strings in the

157   GAs. The individual fitness function values, evaluating a chromosome about the

158   objective function of the optimization problem, are used as the evaluation index of

159   individual quality. In the process of population evolution, selection, crossover, and

160   mutation are continuously applied to gradually reach optimal solutions until it

161   generates the global optimal solution [27-29].The parameters of the GAs were as

162   follows: evolutional generation 100, population size 100, crossover probability 0.8,

163   and mutation probability 0.6. In the process of PSO, each single solution in the

164   D-dimensional search space is taken as a "bird" called "particle". The ith particle

165   position is represented as vector $X_i = (x_{i1}, x_{i2}, \cdots, x_{iD})^T$. A particle is characterized by

166   position, velocity and fitness value. The position giving the best fitness value of the i

167   th particle is represented as vector $P_i = (p_{i1}, p_{i2}, \cdots, p_{iD})^T$ and the velocity is

168   represented as vector $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})^T$. The fitness value is calculated by fitness

169   function and can display the pros and cons of a particle. The index of the best particle

170   among all the particles in the population is represented as g.The particles are operated

171   in accordance with the following equations: $v_{id}^{k+1} = \omega * v_{id}^k + c_1 * rand() *$

172    $(p_{id}^k - x_{id}^k) + c_2 * Rand() * (p_{gd}^k - x_{id}^k)$ and $x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1}$. In the formulas, the

173   ω is inertia weight, the k is the current iteration number, the c_1 and c_2 are

174   acceleration factor, the range of i is positive integer from 1 to n, the range of d is

175   positive integer from 1 to D, the $Rand()$ is random number value distributing

176   between 1 and 2 [22, 30].PSO is initialized with a group of random particles

177   (solutions) and then searches for optima by updating generations. The parameters of

178   PSO were a population size of 20, and 200 iterations [24, 30].The data in the Table 4

179   were an average of triplicate observations and subjected to one-way analysis of

180   variance. The extraction solvents (petroleum ether and acetone) were set as 1 and 2,

181   respectively. All calculations were implemented with Matlab 7.8.0.347 R2009a

182   software (The MathWorks, Natick, MA, USA).

183   **Results and discussion**

184   The modeling of FAs

185   Crude fat content of the three categories of *Coix* seed and the statistical results of the

186   calibration and validation set for the FA content are summarized in Table 1. The BCS

187   had the highest crude fat content, which corresponded to the maximum average value

188   of FAs content. The TCS showed the minimum crude fat content and resulted in the

189   minimum of FAs content. There were significant differences in the FA content of the

190   different categories of *Coix* seeds. The FA content of SCS had the maximum

191   deviation in different extraction conditions, the BCS ranked second, and the TCS had

192   the minimum. The maximum FA content of SCS (163.30 mg KOH/100 g) was 1.04

193   times that of BCS and 2.02 times that of TCS, while the minimum of BCS was 1.19

194   times that of SCS. It might be ascribed that the fat of SCS was tightly bound with

195   starch, protein, phosphorus [31-33] and other nutritional ingredients, and

196   inappropriate extraction conditions made it difficult to extract the fat. The FA

197   concentration range of the validation set was located in the range of calibration set,

198   which was suitable for acquiring successful calibrations.

199 The results of the PLSR and BPNN models for both the calibration and validation

200 set of FAs is shown in Table 2. Generally, a good model should have a low RMSEP

201 value and a high R^2 value. Both modelling methods showed good performance for the

202 three categories of *Coix* seed. However, by making a comparison between the PLSR

203 and BPNN models, it can be seen that the PLSR model performances of BCS and

204 SCS were slightly superior to those of the BPNN model which had a higher R^2

205 (0.9299 and 0.9744, respectively), showing that the relationship between FA content

206 and extraction conditions for BCS and SCS was more likely to be a linear function.

207 While the BPNN model of TCS was a little better than that of the PLSR model, the R^2

208 and RMSEP values were 0.8575 and 0.2981, respectively, which might indicate that

209 the relationship of FA content and extraction conditions was more aligned with a

210 nonlinear function. The results demonstrated that the three types of *Coix* seed had

211 different extraction mechanisms. Although the SCS had the best model performance,

212 the results for TCS still need to be improved. It is necessary to carry out more trials by

213 adjusting the interval of extraction conditions (temperature and time) to improve the

214 robustness and predictive ability of the PLSR and BPNN models in the future studies.

215 And yet, the model performances of BCS, SCS, and TCS have proven that the PLSR

216 and BPNN can reflect on the relationship of FA content and extraction conditions well

217 and also rapidly and efficiently predict the FA content of *Coix* seed.

218 Optimization of the FA extraction conditions

219 The fitting data of the optimal PLSR and BPNN models were used as a fitness

220 function for GAs and PSO. The GAs and PSO searched for the optimal combination

221 in the range of extraction conditions. The evolution process and optimization results

222 of GAs and PSO are shown in Fig. 2 and Table 3. It can be seen from Fig. 2 that the

223 PSO algorithm had faster convergence than that of GAs and the second iteration had

224 obtained the optimum fitness values for the three types of *Coix* seed [22]. However,

225 there were the problems of premature convergence, low precision, and low iterative

226 efficiency in the PSO algorithm, which could result in a local optimum when tackling

227 complex problems. Table 3 showed that the PSO algorithm could obtain the BCS and

228 SCS trapped in the local optimum, because there was great difference between the

229 actual and predicted FAs values for the same extraction conditions (130 °C, 20 min,

230 and acetone extraction). For BCS and SCS, after 86 and 63 evolutional generations by

231 GAs, respectively, the maximum theoretical FAs contents were 158.39 and 165.62 mg

232 KOH/100 g, respectively. The optimal extraction conditions (rounded data) of BCS

233 were 123 °C, 18 min, and acetone extraction (Because the predicted solvent values

234 was 1.94, the rounded data was 2, which represented acetone solvent), and those of

235 SCS were 126 °C, 20 min, and acetone extraction. For TCS, the optimal algorithm

236 was PSO, and the predicted FAs content was 81.54 mg KOH/100 g with the

237 extraction conditions (rounded data) of 124 °C, 20 min and acetone solvent. Then the

238 FAs contents were determined again at the optimal extraction conditions obtained

239 from the above chemometrics methods, and the results are displayed in Table 4. It

240 could be seen from Table 4 that the actual FAs contents of the three *Coix* seeds (BCS

241 was 160.12mg KOH/100g, SCS was 166.01mg KOH/100g and TCS was 81.28mg

242 KOH/100g) in the optimal extraction conditions were approximated to the predicted

11

243   values (BCS was 158.39mg KOH/100g, SCS was 165.62mg KOH/100g and TCS was

244   81.54mg KOH/100g), and higher than the highest FAs contents (BCS was 157.60mg

245   KOH/100g, SCS was 163.30mg KOH/100g and TCS was 80.80mg KOH/100g) used

246   in modelling. Furthermore, the results have also proved that the optimal extraction

247   conditions are reasonable. The above results indicated that the extraction time of ASE

248   (not higher than 1h) was significantly lower than that of supercritical fluid extraction

249   (not lower than 2.5h) [8, 11].The extraction efficiency of acetone for FAs in *Coix* seed

250   was higher than petroleum ether, and the FAs in the BCS were the most easily

251   extracted among the three types of *Coix* seed. The GAs and PSO were rapid and

252   effective extreme value searching algorithms, although the classical PSO needs to be

253   further improved.

254   FA comparison of the three types of *Coix* seed oil

255   The FA composition and content of the three categories of *Coix* seed oil are displayed

256   in Table 5. It is seen from Table 5 that the types of FAs were slightly different from

257   those of supercritical extractions [8].There were not heptadecenoic acid, nonadecanoic

258   acid, nonadecyenoic acid, eicosenoic acid, heneicosanic acid, tricosanoic acid,

259   tetracosanoic acid, and pentacosanoic acid in the study of Hu *et al*. [8],while

260   eicosanoid and hexacosanoic acid were not detected in our study. Oleic acid

261   accounted for the highest proportion in the oils from the three categories of *Coix* seed;

262   the contents were BCS 75.26%, TCS 77.02%, and SCS 73.45%, respectively, which

263   was higher than that of the previous study (47.5%) which used supercritical extraction

264   [8].These results could be due to the differences of extraction methods. From the

12

265   Table 5, it could be seen that there were significant differences in the content of

266   palmitic acid, palmitoleic acid, stearic acid, oleic acid, eicosanoic acid, and

267   tetracosanoic acid among the three oils. Furthermore, the BCS showed a significant

268   difference with TCS and SCS in the content of heptadecanoic acid, heptadecenoic

269   acid, nonadecyenoic acid, eicosenoic acid, and docosanoic acid. That is, the FA

270   composition of SCS was a little closer to that of TCS. The results explained that the

271   FA composition of different varieties of *Coix* seed were different and could be

272   ascribed to the differences in biological origin.

273   **Conclusions**

274   The results demonstrated that the PLSR and BPNN models could reflect the

275   relationship of FA content and extraction conditions well. For BCS and SCS, the

276   performances of the PLSR models slightly outperformed those of the BPNN models;

277   while for TCS, the BPNN model was superior to the PLSR model. The GAs could

278   seek out the optimal extraction conditions for the PLSR models of BCS (123 °C, 18

279   min, and acetone extraction) and SCS (126 °C, 20 min, and acetone extraction)

280   rapidly and effectively, and PSO algorithms were more suitable for the BPNN model

281   of TCS (124 °C, 20 min, and acetone extraction). Furthermore, all the extraction time

282   of the FAs from three *Coix* seeds using ASE was shorter than common extraction

283   techniques, such as Soxhlet extraction, microwave extraction, sonication extraction

284   and supercritical fluid extraction. There were differences in the FA content of the

285   three categories of *Coix* seed on account of the differences of biological origin.

286   Therefore, ASE combined with chemometrics methods can be a labour-saving,

287   time-saving, and powerful tool for rapid and effective determination of FAs compared

288   with the common extraction methods. We believe that this approach should be further

289   applied to extract other nutrition ingredients from natural food samples.

290   **Acknowledgements**

294

295   **References**

296   1. Lim TK. Edible Medicinal And Non-Medicinal Plants. Springer. 2013; 5: 243-261.

297   https://doi.org/10.1007/978-94-007-5653-3_21

298   2. Lu Y, Zhang BY, Jia ZX, Wu WJ, Lu ZQ. Hepatocellular carcinoma HepG2 cell

299   apoptosis and caspase-8 and Bcl-2 expression induced by injectable seed extract of

300   *Coix* lacryma-jobi. Hepatob Pancreat Dis. 2011; 10(3): 303-307.

301   https://doi.org/10.1016/S1499-3872(11)60050-7 PMID: 21669575

302   3. Chen HJ, Chung CP, Chiang W, Lin YL. Anti-inflammatory effects and chemical

303   study of a flavonoid-enriched fraction from adlay bran. Food Chem. 2011; 126(4):

304   1741-1748. https://doi.org/10.1016/j.foodchem.2010.12.074

305   4. Chen HJ, Shih CK, Hsu HY, Chiang W. Mast cell-dependent allergic responses are

306   inhibited by ethanolic extract of adlay (*Coix lachryma-jobi L. Var. ma-yuen Stapf*)

307   testa. J Agr Food Chem. 2010; 58(4): 2596-2601. https://doi.org/10.1021/jf904356q

308   PMID: 20102206

309   5. Zhu F. Coix: Chemical composition and health effects. Trends Food Sci Tech.

310   2017; 61:160-175. https://doi.org/10.1016/j.tifs.2016.12.003

311   6 Kuo CC, Chiang W, Liu GP, Chien YL, Chang JY, Lee CK, et al. 2,

312   2'-diphenyl-1-picrylhydrazyl radical-scavenging active components from adlay

313   (*Coix lachryma-jobi L. Var. ma-yuen Stapf*) hulls. J Agr Food Chem. 2002; 50(21):

314   5850-5855. https://doi.org/10.1021/jf020391w

315   7. Manosroi J, Khositsuntiwong N, Manosroi A. Biological activities of

316   fructooligosaccharide (FOS)-containing *Coix lachryma-jobi* Linn. extract. J Food

317   Sci Tech. 2014; 51(2), 341-346. https://doi.org/10.1007/s13197-011-0498-6 PMID:

318   24493893

319   8. Hu A, Zhang Z, Zheng J, Wang Y, Chen Q, Liu R, et al. Optimizations and

320   comparison of two supercritical extractions of adlay oil. Innov Food Sci Emerg.

321   2012; 13: 128-133. https://doi.org/10.1016/j.ifset.2011.10.002

322   9. Zhao W, Gong Y, Huang S, Yu H, Lu Y. Optimization and kinetics for the

323   refluxing extraction process of *Coix* seed oil. Chin J Bioproc E. 2010; 8: 1-5.

324   10. Zhao W, Zhu Q, Gong Y, Jin H, Huang S. Effects of solvents and processes of

325   extraction on the yield of *Coix* seed oil. Chin J Bioproc E. 2009; **7**: 24-27.

326   11. Hu AJ, Zhao S, Liang H, Qiu TQ, Chen G. Ultrasound assisted supercritical fluid

327   extraction of oil and coixenolide from adlay seed. Ultrason. Sonochem. 2007;

328   14(2): 219-224. https://doi.org/10.1016/j.ultsonch.2006.03.005 PMID: 16677844

329   12. Richter BE, Jones BA, Ezzell JL, Porter NL, Avdalovic N, Pohl C. Accelerated

330   solvent extraction: a technique for sample preparation. Anal Chem. 1996; 68(6):

331   1033-1039. https://doi.org/10.1021/ac9508199

332   13. Hossain MB, Barry-Ryan C, Martin-Diana AB, Brunton NP. Optimisation of

333   accelerated solvent extraction of antioxidant compounds from rosemary

334   (*Rosmarinus officinalis L.*), marjoram (*Origanum majorana L.*) and oregano

335   (*Origanum vulgare L.*) using response surface methodology. Food Chem. 2011;

336   126(1): 339-346. https://doi.org/10.1016/j.foodchem.2010.10.076

337   14. Schäfer K. Accelerated solvent extraction of lipids for determining the fatty acid

338   composition of biological material. Anal Chim Acta. 1998; 358(1): 69-77.

339   https://doi.org/10.1016/s0003-2670(97)00587-4

340   15. Vetter W, Laure S, Wendlinger C, Mattes A, Smith AW, Knight DW.

341   Determination of furan fatty acids in food samples. J Am Oil Chem. 2012; 89(8):

342   1501-1508. https://doi.org/10.1007/s11746-012-2038-6

343   16. Zhou L, Le Grandois J, Marchioni E, Zhao M, Ennahar S, Bindler F. Improvement

344   of total lipid and glycerophospholipid recoveries from various food matrices using

345   pressurized liquid extraction. J Agr Food Chem. 2010; 58(18): 9912-9917.

346   https://doi.org/10.1021/jf101992j

347   17. Dodds ED, McCoy MR, Geldenhuys A, Rea LD, Kennish JM. Microscale

348   recovery of total lipids from fish tissue by accelerated solvent extraction. J Am Oil

349   Chem. 2004; 81(9): 835-840. https://doi.org/10.1007/s11746-004-0988-2

350   18. Dunford NT, Zhang M. Pressurized solvent extraction of wheat germ oil. Food

351   Res Int. 2003; 36(9-10), 905-909. https://doi.org/10.1016/s0963-9969(03)00099-1

352   19. Liu X, Zhang X, Rong YZ, Wu JH, Yang YJ, Wang ZW. Rapid determination of

353   fat, protein and amino acid content in *Coix* seed using near-infrared spectroscopy

354   technique. Food Anal Method. 2015; 8(2): 334-342.

355   https://doi.org/10.1007/s12161-014-9897-4

356   20. Burns DA, Ciurczak EW. Handbook of near-infrared analysis. 2007; CRC press,

357   Boca Raton.

358   21. Khan K, Sahai A. A comparison of BA, GA, PSO, BP and LM for training feed

359   forward neural networks in e-learning context. Int J Intell Syst Appl. 2012; 4(7):

360   23-29. https://doi.org/10.5815/ijisa.2012.07.03

361   22. Eberhart RC, Shi Y. Comparison between genetic algorithms and particle swarm

362   optimization. Lect Notes Comput Sci. 1998; 1447: 611-616.

363   https://doi.org/10.1007/BFb0040812

364   23. Konak A, Coit DW, Smith AE. Multi-objective optimization using genetic

365   algorithms: A tutorial. Reliab Eng Syst Safe. 2006; 91(9): 992-1007.

366   https://doi.org/10.1016/j.ress.2005.11.018

367   24. Kennedy J, Eberhart RC. Particle swarm optimization in: Proc. of IEEE Int. Conf.

368   on Neural Network, Perth, Australia. 1995; 1942-1948.

369   https://doi.org/10.1007/978-0-387-30164-8_630

370   25. Zhang JR, Zhang J, Lok TM, Lyu MR. A hybrid particle swarm optimization–

371   back-propagation algorithm for feedforward neural network training. Appl Math

372   Comput. 2007; 185(2): 1026-1037. https://doi.org/10.1016/j.amc.2006.07.025

373   26. Kennard RW, Stone LA. Computer aided design of experiments. Technometrics.

374   1969; 11: 137-148. https://doi.org/10.1080/00401706.1969.10490666

375   27. Leardi R. Application of genetic algorithm-PLS for feature selection in spectral

376   data sets. J Chemometr. 2000; 14(5-6): 643-655.

377   https://doi.org/10.1002/1099-128X(200009/12)14:5/6<643::AID-CEM621>3.0.CO;

378   2-E

379   28. Yang Y, Gao M, Yu X, Zhang Y, Lyu S. Optimization of medium composition for

380   two-step fermentation of vitamin C based on artificial neural network-genetic

381   algorithm techniques. Biotechnol Biotec Eq. 2015; 29(6), 1128-1134.

382   https://doi.org/10.1080/13102818.2015.1063970

383   29. Saravanan R, Asokan P, Sachidanandam M. A multi-objective genetic algorithm

384   (GA) approach for optimization of surface grinding operations. Int J Mach Tool

385   Manu. 2002; 42(12): 1327-1334. https://doi.org/10.1016/s0890-6955(02)00074-3

386   30. Shi F, Wang X, Yu L, Li Y. Neural Network of MATLAB: 30 Cases Analysis.

387   Beijing University of Aeronautics and Astronautics Press, Beijing. 2010.

388   31. Kainuma K. Handbook of Starch Science. Asakura Publishing, Tokyo. 1977.

389   32. Ambigaipalan P, Hoover R, Donner E, Liu Q, Jaiswal S, Chibbar R, Nantangad

390   KKM, Seetharamand K. Structure of faba bean, black bean and pinto bean starches

391   at different levels of granule organization and their physicochemical properties.

392   Food Re Int. 2011; 44(9): 2962-2974. https://doi.org/10.1016/j.foodres.2011.07.006

393   33. Singh N, Singh J, Kaur L, Sodhi NS, Gill BS. Morphological, thermal and

394   rheological properties of starches from different botanical sources. Food Chem.

395   2003; 81(2): 219-231. https://doi.org/10.1016/S0308-8146(02)00416-8

396

397 **Figure Captions**

398 Fig.1 The appearance and size of the three categories of *Coix* seed

399 Fig.2 Evolution of the optimal and average fitness values in genetic algorithms and

400 particle swarm optimization.

401 The B, S, and T represent big *Coix* seed, small *Coix* seed, and translucent *Coix* seed,

402 respectively.

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421
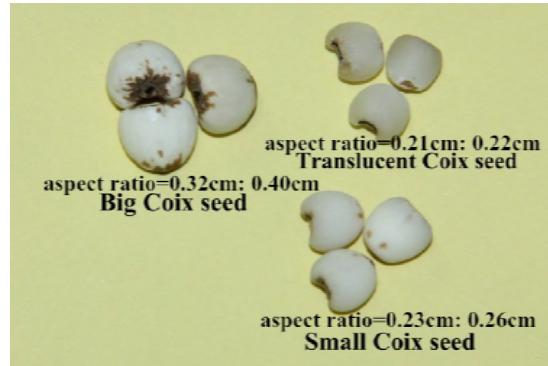
422

423   Fig. 1 The appearance and size of the three categories of *Coix* seed
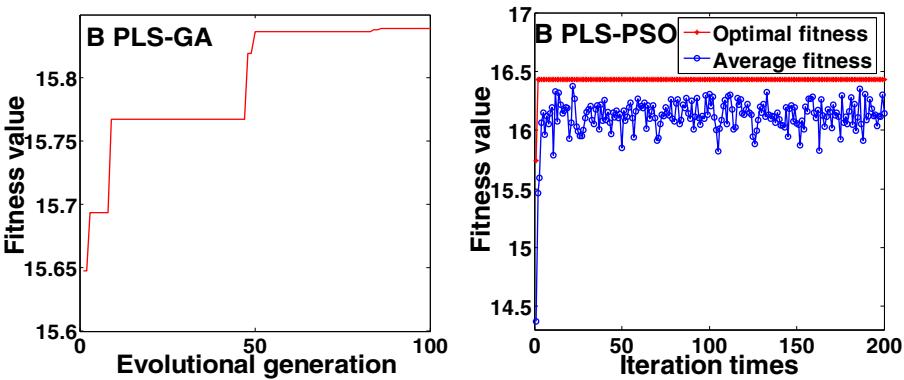
424

425
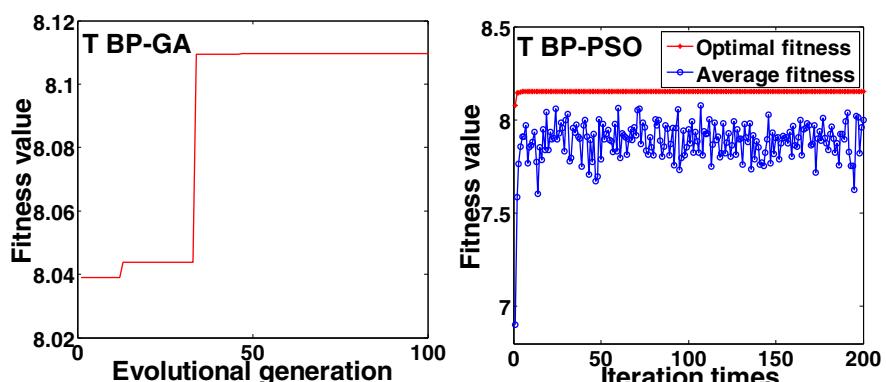
426

427   Fig. 2 Evolution of the optimal and average fitness values in the genetic algorithms and particle swarm

428   optimization.

429   The B, S, and T represent big *Coix* seed, small *Coix* seed, and translucent *Coix* seed, respectively.

20

430   Table 1 Statistical data for crude fat and fatty acids (FAs) of the three categories of *Coix* seed in the calibration and

431   validation sets

Varieties	Crude fat	Calibration set of FAs (mg KOH/100 g)			Validation set of FAs (mg KOH/100 g)			
	(%)	No. of samples	SD	Mean	Range	No. of samples	SD	Mean	Range
BCS	7.68±0.04a	24	12.41	140.07	116.10–157.60	8	13.50	143.44	121.40–157.40
SCS	7.10±0.08b	24	23.00	132.55	97.20–163.30	8	25.25	125.07	99.50–160.50
TCS	5.45±0.15c	24	11.39	65.42	42.70–80.80	8	12.41	66.54	51.30–80.10

432   Means followed by a different letter within a column for each *Coix* seed are significantly different ($P < 0.05$).

433   BCS, big *Coix* seed; SCS, small *Coix* seed; TCS, translucent *Coix* seed; SD, standard deviation.

434

435   Table 2 The performances of PLSR and BPNN models for the fatty acid content of the three categories of *Coix*

436   seed

Category	PLSR				BPNN	
	Calibration set		Validation set		Validation set	
	RMSECV	R^2	R^2	RMSEP	R^2	RMSEP
BCS	0.3368	0.9477	0.9299	0.3782	0.8745	0.4576
SCS	0.3304	0.9854	0.9744	0.4264	0.9625	0.3649
TCS	0.3723	0.9225	0.8396	0.4709	0.8575	0.2981

437   BCS, big *Coix* seed; SCS, small *Coix* seed; TCS, translucent *Coix* seed; R^2, coefficient of determination;

438   RMSECV, root mean square error of cross validation; RMSEP, root mean square error of prediction; PLSR, partial

439   least-squares regression; BPNN, backpropagation neural network.

440

441

442

443

444

445

446

447

448 Table 3 Comparison of the optimization extraction parameters by PLSR and BPNN combined with GAs and PSO

Varieties	Optimization methods	Temperature (°C)	Time (min)	Solvent	FA (mg KOH/100 g)
BCS	PLSR-GA	123.35	18.19	1.94	158.39
	PLSR-PSO	130	20	2 (acetone)	164.29
SCS	PLSR-GA	126.14	19.68	1.99	165.62
	PLSR-PSO	130	20	2 (acetone)	168.33
TCS	BPNN-GA	123.95	19.75	1.99	81.10
	BPNN-PSO	124.38	20	2 (acetone)	81.54

449 BCS, big *Coix* seed; SCS, small *Coix* seed; TCS, translucent *Coix* seed; PLSR, partial least-squares regression;

450 BPNN, backpropagation neural network; GA, genetic algorithms; PSO, particle swarm optimization.

451

452 Table 4 Verification the FA content of three *Coix* seeds in the optimization extraction parameters

Varieties	Optimization methods	Temperature (°C)	Time (min)	Solvent	FA (mg KOH/100 g)
BCS	PLSR-GA	123	18	acetone	160.12±1.01
SCS	PLSR-GA	126	20	acetone	166.04±0.71
TCS	BPNN-PSO	124	20	acetone	81.28±0.95

453 BCS, big *Coix* seed; SCS, small *Coix* seed; TCS, translucent *Coix* seed; PLSR, partial least-squares regression;

454 BPNN, backpropagation neural network; GA, genetic algorithms; PSO, particle swarm optimization.

455

456

457

458

459

460

461

462

463

22

464

465

466 Table 5 Fatty acid analysis of the three categories of *Coix* seed oil

Fatty acid (%)	BCS	SCS	TCS
Myristic acid	0.08±0.00a	0.07±0.00a	0.05±0.00b
Pentadecyl acid	0.04±0.00a	0.04±0.00a	0.03±0.00a
Palmitic acid	16.25±0.07b	16.80±0.04a	15.04±0.06c
Palmitoleic acid	0.65±0.01b	0.75±0.01a	0.51±0.01c
Heptadecanoic acid	0.18±0.01b	0.25±0.01a	0.25±0.01a
Heptadecenoic acid	0.14±0.01b	0.18±0.00a	0.21±0.01a
Stearic acid	4.19±0.02b	5.35±0.01a	3.96±0.03c
Oleic acid	75.26±0.06b	73.45±0.04c	77.02±0.03a
Octadecadienoic acid	0.09±0.00a	0.10±0.00a	0.09±0.00a
Nonadecanoic acid	0.03±0.00a	0.04±0.00a	0.03±0.00a
Nonadecyenoic acid	0.06±0.01b	0.07±0.00a	0.09±0.01a
Eicosanoic acid	1.28±0.01b	1.31±0.01a	1.18±0.01c
Eicosenoic acid	0.81±0.01b	0.84±0.01a	0.86±0.01a
Heneicosanoic acid	0.03±0.00a	0.03±0.00a	0.03±0.00a
Docosanoic acid	0.44±0.00a	0.34±0.01b	0.34±0.00b
Docosenoic acid	0.04±0.00a	0.02±0.00a	0.04±0.00a
Tricosanoic acid	0.05±0.00a	0.05±0.00a	0.04±0.00a
Tetracosanoic acid	0.38±0.00a	0.29±0.01b	0.21±0.01c
Pentacosanoic acid	0.02±0.00a	0.02±0.00a	0.01±0.00a

467 Means followed by a different letter within a row for each *Coix* seed oil are significantly different ($P < 0.05$).

468 BCS, big *Coix* seed; SCS, small *Coix* seed; TCS, translucent *Coix* seed.