1 **TITLE PAGE**
2 **- Food Science of Animal Resources -**
3 **Upload this completed form to website with submission**
4

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research article
Article Title	Proximate content monitoring of black soldier fly larval (Hermetia illucens) dry matter for feed material using short-wave infrared hyperspectral imaging
Running Title (within 10 words)	Feed insect monitoring using the hyperspectral imaging system
Author	Juntae Kim[1], Hary Kurniawan[1], Mohammad Akbar Faqeerzada[1], Geonwoo Kim[2], Hoonsoo Lee[3], Moon S. Kim[4], Insuck Baek[4], Byoung-Kwan Cho[1,5]
Affiliation	[1] Department of Biosystems Machinery Engineering, College of Agricultural and Life Science, Chungnam national university, Daejeon, 34134, Republic of Korea [2] Department of Bio-Industrial Machinery Engineering, College of Agriculture and Life Science, Gyeongsang National University, Jinju-si, Gyeongsangnam-do, 52828, South Korea [3] Department of Biosystems Engineering, College of Agriculture, Life & Environment Science, Chungbuk National University, 1 Chungdae-ro, Seowon-gu, Cheongju, Chungbuk 28644, Republic of Korea [4] Environmental Microbial and Food Safety Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD, 20705 United States [5] Department of Smart Agriculture Systems, College of Agricultural and Life Science, Chungnam National University, Daejeon, 34134, Republic of Korea
Special remarks – if authors have additional information to inform the editorial office	
ORCID (All authors must have ORCID) https://orcid.org	Juntae Kim (https://orcid.org/0000-0002-5398-8839) Hary Kurniawan (https://orcid.org/0000-0002-8509-7600) Mohammad Akbar Faqeerzada (https://orcid.org/0000-0002-1829-2502) Geonwoo Kim (https://orcid.org/0000-0002-6274-3649) Hoonsoo Lee (https://orcid.org/0000-0001-8074-4234) Moon S. Kim (https://orcid.org/0000-0001-8504-9839) Insuck Baek (https://orcid.org/0000-0003-1044-349X) Byoung-Kwan Cho (https://orcid.org/0000-0002-8397-9853)
Conflicts of interest List any present or potential conflict s of interest for all authors. (This field may be published.)	The authors declare no potential conflict of interest.
Acknowledgements State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available. (This field may be published.)	
Author contributions (This field may be published.)	Conceptualization: Kim J, Lee H, Cho BK. Data curation: Kim J, Kurniawan H, Faqeerzada MA. Formal analysis: Kim J, Lee H, Baek I. Methodology: Kim J, Kim MS, Cho BK. Software: Kim J, Kim G. Validation: Kim J, Kim G, Lee H, Baek I, Cho BK. Investigation: Cho BK. Writing - original draft: Kim J, Cho BK. Writing - review & editing: Kim J, Kim MS, Cho BK.

Ethics approval (IRB/IACUC) (This field may be published.)	This article does not require IRB/IACUC approval because there are no human and animal participants.

5

6 **CORRESPONDING AUTHOR CONTACT INFORMATION**

For the **corresponding** author (responsible for correspondence, proofreading, and reprints)	**Fill in information in each box below**
First name, middle initial, last name	Byoung-Kwan Cho
Email address – this is where your proofs will be sent	chobk@cnu.ac.kr
Secondary Email address	bxc195@gmail.com
Postal address	College of Agriculture and Life Science, Chungnam National University 99 Daehak-Ro, BLDG# E10-2, RM# 2213, Yuseong-Gu, Daejeon, 34134, Republic of Korea
Cell phone number	+82-10-6822-6318
Office phone number	+82-42-821-6715
Fax number	+82-42-823-6246

7

8

9   **Proximate content monitoring of black soldier fly larval (*Hermetia illucens*) dry**

10   **matter for feed material using short-wave infrared hyperspectral imaging**

11

12   Abstract

13   Edible insects are gaining popularity as a potential future food source because of their

14   high protein content and efficient use of space. Black soldier fly larvae are noteworthy

15   because they can be used as feed for various animals including reptiles, dogs, fish,

16   chickens, and pigs. However, if the edible insect industry is to advance, we should use

17   automation to reduce labor and increase production. Consequently, there is a growing

18   demand for sensing technologies that can automate the evaluation of insect quality. This

19   study used short-wave infrared (SWIR) hyperspectral imaging to predict the proximate

20   composition of dried black soldier fly larvae, including moisture, crude protein, crude fat,

21   crude fiber, and crude ash content. The larvae were dried at various temperatures and

22   times, and images were captured using an SWIR camera. A partial least-squares

23   regression (PLSR) model was developed to predict the proximate content. The SWIR-

24   based hyperspectral camera accurately predicted the proximate composition of black

25   soldier fly larvae from the best preprocessing model; moisture, crude protein, crude fat,

26   crude fiber, and crude ash content were predicted with high accuracy, with R^2 values of

27   0.89 or more, and RMSEP values were within 2%. Among preprocessing methods, mean

28   normalization and max normalization methods were effective in proximate prediction

29   models. Therefore, SWIR-based hyperspectral cameras can be used to create automated

30   quality management systems for black soldier fly larvae.

31

32   Keywords: Black soldier fly larvae, Feed insect, Quality monitoring, Chemical

33   image, Hyperspectral image.

34

Introduction

Insects have a rich protein content and are being suggested as a new alternative food source. Although entomophagy, or the consumption of insects, varies depending on the region, humans have already consumed over 2,111 species of insects since the past (Van Huis, 2013; Jongema, 2017). Recently, edible insects have been distributed in processed forms, such as protein bars, nuggets, and schnitzels, in European countries. However, there is still a clear aversion to eating insects (Hartmann et al., 2015), and experts have reported that the industrialization of edible insects may take some time because of the risks posed by allergic factors (Jensen and Lieberoth, 2019). However, using insects as animal feed poses fewer aversion and safety issues compared with edible insects. Feed insects can serve as a substitute for traditional feed ingredients, and they may serve as alternatives to grain feed such as soybean and corn, as well as fishmeal (Van Raamsdonk et al., 2017; Nogales-Mérida et al., 2019). In the feed market especially, there has been a trend towards reducing the proportion of soybeans used in feed by establishing mixing ratios because of the decrease in crop production caused by global warming (Kępińska-Pacelik and Biel, 2022; Boerema et al., 2016). Edible insects are also being considered fishmeal substitutes in feed because of the scarcity of fishery resources and to reduce feed costs. Nogales-Mérida et al. (2019) reported that many feed insects are among the best alternatives for partially or completely replacing fishmeal because they contain the essential amino acids and fatty acids necessary for aquaculture. Various insect species that can be used as feed are gaining attention because of their potential for mass production. These include larvae of the black soldier fly (*Hermetia illucens*), mealworm (*Tenebrio molitor*), supermealworm (*Zophobas morio*), housefly (*Musca domestica*), and crickets (*Acheta domesticus*) (Van Raamsdonk et al., 2017). They are being developed into feed products for various animals such as pigs (Veldkamp and Bosch, 2015; Ji et al.,

4

2016), poultry (Pieterse et al., 2019; Cullere et al., 2017), fish (Nogales-Mérida et al., 2019; Zarantoniello et al., 2020), and are even used in pet food (Kępińska-Pacelik and Biel, 2022). The possibility of using them as cattle feed has also been discussed (Drewery et al., 2022). Ji et al. (2016) conducted a study on the nutritional composition and efficiency of insect feed. They fed *Tenebrio molitor, Musca domestica* larvae, and *Zophobas morio* powders as dietary proteins to early weaned piglets and reported that they provided benefits in terms of high amino acid utilization and decreased diarrhea. They also reported that insect feed did not negatively affect the growth rate of early weaned piglets. In addition, Caimi et al. (2020) reported no significant difference in the growth rate of Siberian sturgeon juveniles fed feed mixed with approximately 25% defatted *H. illucens* powder compared with those fed regular feed.

Insects are animal proteins, but the use of animal proteins as livestock feed has been difficult since the emergence of bovine spongiform encephalopathy (van Raamsdonk et al., 2017). However, regulations regarding feed insects are gradually relaxing in each country and significant industrial growth is expected. In particular, black soldier fly larvae (BSFL) have a lower protein content than other insects, but higher fat and chitin content, making them a valuable feed ingredient. According to Nam et al. (2022) the protein content of BSFL is approximately 40-43%, while mealworms (*Tenebrio molitor*) have a protein content of 46-57%, house crickets (*Gryllus bimaculatus*) range from 58-60%, and house flies (*Musca domestica*) range from 57-63%. Additionally, the fat content was reported to be around 28-30% for BSFL, 24-37% for *Tenebrio molitor*, 14-16% for *Gryllus bimaculatus*, and 7.3-25% for *Musca domestica*. BSFL can be raised on food waste, which is closely related to the United Nations' Sustainable Development Goals and corporate Environmental, Social, and Governance goals, because they can also produce valuable vermicompost. Additionally, adult black soldier flies do not have a

mouth, so they do not transfer pathogens as other flies (Sheppard et al., 2002). The black

soldier fly typically lays approximately 500 eggs and hatches within 4 d, and the larvae

decompose organic matter for 14 d (Bessa et al., 2020; Diclaro and Kaufman, 2009).

The black soldier fly farming industry is expected to grow rapidly in the insect feed

market; therefore, it is essential to establish a mass-production automation system

(Surendra et al., 2020). There are studies related to mass production automation of black

soldier flies, such as the study on the automatic breeding system for black soldier flies

conducted by Erbland et al. (2021), and it has been reported that Hexafly, Nasekomo

(Thrastardottir et al., 2021), and Korea's CIEF are currently producing black soldier flies

in an automated factory format. With recent advancements in computer and sensing

technologies, process automation has progressed to smart factorization. In particular,

when producing feed insects, the small size of insects and large quantities required for

processing make quality control difficult. Failure to manage quality can result in

unpleasant odors and mold, which can threaten the quality of the final product (Kępińska-

Pacelik and Biel, 2022). In particular, when used as animal or fish feed, it is essential to

understand the general nutrient content of each ingredient. Therefore, there is a need for

a selection technology that can quickly and accurately evaluate the nutrient content.

Spectrometer-based studies of edible and fed insects have also been conducted. Benes et

al. (2022) classified flour and seven types of insect powder and separated them. They

reported that even mixtures of flour and insect powder could be distinguished with an

error rate of 0.65%. Unlike conventional point measurement spectrometers, hyperspectral

imaging (HSI) can measure the chemical characteristics of samples as images, making it

possible to utilize them for the quality control of heterogeneous products such as food

and feed. Furthermore, based on the acquired spectrum, a chemical image can be created,

allowing the visualization of the chemical composition of the sample. Cruz-Tirado et al.

(2023) used a hyperspectral camera in the range of 928–2524 nm to determine the individual protein content of BSFL. They developed algorithms using the support vector machine regression (SVMR) and partial least-squares regression (PLSR) analysis methods and reported R^2 of prediction set values ranging from 0.731 to 0.773, with root mean square error of prediction (RMSEP) values ranging from 1.567% to 1.664%. Although studies on insect detection in grains and sex determination using HSI have been conducted, as well as on the classification of flour and insect powder, research on monitoring the nutritional components of feed insects for use as feed has not yet been extensively conducted.

The final color of black soldier fly larvae (BSFL) powder can vary depending on the killing and drying methods (Saucier et al., 2022; Larouche et al., 2019). One of the main reasons for this color change is the oxidation of polyphenols and the formation of complexes between iron and polyphenols during the drying process of BSFL (Larouche et al., 2019; Janssen et al., 2019a; Janssen et al., 2019b). Given that the color of a sample can be influenced by various factors, detection methods in the visible light range may be more sensitive to the color variations of the sample rather than its functional groups, such as -OH and -CH groups. Using a simple RGB camera or a visible/near-infrared (Vis/NIR) waveband range may pose difficulties in evaluating the quality of dried BSFL. Consequently, in this study, a shortwave infrared (SWIR) hyperspectral camera was employed for analysis.

The SWIR camera, operating in the SWIR range (1000-2500 nm), demonstrates higher sensitivity to the chemical composition of the sample and is less affected by sample color compared to the Vis/NIR range (400-1000 nm). Although hyperspectral imaging (HSI) technology is widely utilized for food quality control, there is a need for optimization and experimental application processes before its installation in sorting machines becomes

7

135 feasible. Thus, the objective of this study was to develop an algorithm using a SWIR-

136 based HSI system to evaluate the proximate compositions (moisture, crude protein, crude

137 fat, crude fiber, and crude ash) of dried BSFL and to create an optimized model suitable

138 for sorting machines. Ultimately, this study aimed to explore the potential of using HSI

139 for quality monitoring of feed insects based on the algorithm developed.

140

141 Materials and Methods

142 Sample preparation

143 The fifth instar live larvae of the black soldier fly (*Hermetia illucens*) used in this study

144 were purchased 2 kg from Entomo, a Chung-Ju, South Korea. They were divided into

145 nine groups of 200 g each and stored frozen at -20°C until just before the experiment. The

146 experimental design was a 3 × 3 factorial design with three different drying temperatures

147 (50°C, 60°C, and 70°C) and three different drying times (1 h, 2 h, and 3 h), resulting in

148 nine different treatment groups. Drying was performed using a hot-air food dryer (LD-

149 918BT, Liquip, Hwasung, Korea) with an air velocity of 2.5-3.0 m/s, and the dried

150 samples were vacuum-packed and stored at room temperature (23–25°C) in a desiccator

151 until hyperspectral image acquisition. After drying, 20 g of each sample was placed in a

152 Petri dish (Ø 90 mm, 15 mm) for SWIR HSI. The samples were homogenized for 1 min

153 using a grinder (A11 basic, Ika Werke GmbH & Co., Staufen, Germany) after imaging.

154 The samples were transported to a chemistry laboratory for proximate component

155 analysis.

156

157 SWIR hyperspectral image acquisition

158 The camera used was a line-scan camera system (Headwall Photonics, Fitchburg, MA,

159 USA) capable of capturing 275 wavelengths ranges of 894-2504 nm (Fig. 1). Six

160 tungsten-halogen lamps (100 W, 12 V, Light Bank; JCR 12V, Ushio Inc., Tokyo, Japan)

161 connected to fiber optics were used as light sources for imaging. The imaging sample was

162 moved towards the camera using a DC motor-driven movable stage to obtain a

163 hyperspectral image. The speed of the movable stage during the line scan was set at

164 3.48 mm/s, and the scan range was set to 600 scans/sample. The obtained hyperspectral

165 image was in the form of a 3D hypercube with two spatial coordinates (x- and y-axes)

166 and a wavelength range (λ) dimension, with a final size of 384 (x) × 700 (y) × 275 (λ).

167 For data analysis, only the wavelength range of 1000–2350 nm was used to remove sensor

168 noise, resulting in 232 wavelengths (Fig. 2).

169

170 Proximate content analysis

171 After the hyperspectral imaging process, the samples were ground for a period of 1

172 minute using a grinding mill (A11 basic, IKA Works GmbH & Co. KG, Staufen,

173 Germany). Proximate composition analysis was conducted by repeating the procedure

174 three times, according to the AOAC method (AOAC, 2005). Moisture content was

175 determined by drying the samples (1.0 g) at 105°C for 24 h. The moisture content was

176 calculated using Equation (1) after 24 h of drying.

177

178 $$\text{Moisture contents (\%)} = \frac{(\text{Weight before drying - Weight after drying})}{\text{Weight before drying}} \times 100 \quad (1)$$

179

180 The crude protein content was analyzed using the Kjeldahl method. Approximately

181 0.5 g of each sample was decomposed by adding a catalytic agent (1000 Kjeltabs S/3.5,

182 FOSS TECATOR) and 12 mL of H_2SO_4. The sample was heated at 420°C for 1 h and

183 cooled. The nitrogen content was measured using a Kjeltec device (Kjeltec auto 2300

184 Analyzer, FOSS TECATOR, Höganäs, Sweden), and the crude protein content was

185   calculated by multiplying the nitrogen coefficient (4.76). Typically, a nitrogen coefficient

186   of 6.25 is used for animal protein. However, there is a possibility of overestimating the

187   crude protein content in insects owing to the presence of nitrogen in chitin. Therefore,

188   recent studies have used a nitrogen coefficient of 4.76 to calculate the crude protein

189   content (Janssen et al., 2017; Cruz-Tirado et al., 2023). The crude fat content was

190   analyzed by ether extraction using a Soxhlet system. Crude fiber analysis was performed

191   using filter bags (Ankom Technology, Macedon, NY, USA), and the difference between

192   the weight of the insoluble residue when treated with 1.25% H_2SO_4 and 1.25% NaOH

193   solution and the weight after painting was expressed as a percentage of the sample. The

194   ash contents of the samples was analyzed using the combustion method. Approximately

195   2 g of each sample was heated by electric combustion for analysis. The sample was then

196   placed in a 600°C electric furnace (CT-DMF2, Coretech Co., Korea) for 2 h. After cooling

197   for 40 min in a desiccator, the sample was weighed to determine the amount of ash present

198   by calculating the difference in weight before and after combustion.

199

200   Statistics of reference data

201   A two-way ANOVA test was conducted to analyze the significant differences in the

202   biochemical composition results of the sample according to the drying time and

203   temperature, and the interaction P value was calculated for both drying time and

204   temperature. A one-way ANOVA test was conducted again for each drying time and

205   temperature, and a post hoc analysis was performed using Duncan's multiple range test

206   for samples with significant differences ($p < 0.05$). Basic statistics were obtained using the

207   R statistical program (version 4.1.2), with the CRAN mirror set to the USA (CA1) and

208   'Agricolae' libraries.

209

Hyperspectral image intensity calibration

To mitigate the influence of external environmental factors, such as dark current noise and non-uniform lighting, spectral intensity calibration was conducted. For this purpose, white and dark references were acquired during image acquisition. The white reference was obtained using a white Teflon board (100% reflectance, 30 cm × 30 cm × 1 cm), whereas the dark reference was obtained by closing the camera lens cap and capturing an image with the light source turned off. The intensity calibration of the acquired hyperspectral image was performed using Equation (2)

$$X_c = \frac{T_{ij}^R(\lambda) - T_{ij}^D(\lambda)}{T_{ij}^W(\lambda) - T_{ij}^D(\lambda)} \quad (2)$$

~~W~~where $T_{ij}^R(\lambda)$ represents the spectrum of the sample at the pixel, $T_{ij}^D(\lambda)$ represents the spectrum value of the dark reference image, and $T_{ij}^W(\lambda)$ represents the spectrum of the white reference. The final X_c value represented a pixel-wise intensity-calibrated hyperspectral image, which is a relative intensity spectral image. Finally, the wavelength is extracted from the processed hyperspectral images.

Image processing and spectral data extraction

The calibrated image was used to extract spectra by selecting the region of interest (ROI), and a masking image was created by setting the threshold value to 0.2 to select only the sample area. The masking image was then multiplied by all wavelength images to separate only the sample area of the spectrum (Fig. 2). The spectrum was extracted from all the pixels of the separated sample area and averaged to obtain the mean spectrum. Ten average spectra were extracted for each sample image, and 600 sample spectra were obtained and used for the subsequent multivariate analyses.

234   Preprocessing of spectral data

235   The acquired spectral data contain considerable noise. Many external factors, such as

236   baseline correction, band shift, and light scattering, hinder the acquisition of pure data,

237   and spectra preprocessing is usually performed for noise removal during the analysis.

238   Normalization and deviation methods are commonly used for preprocessing. As there is

239   no single best preprocessing technique, this study utilized Seven preprocessing methods

240   to pre-process the acquired wavelengths. Three normalization methods (minimum,

241   maximum, and range normalization), standard normal variate (SNV), multiplicative

242   scatter correction (MSC), Savitzky-Golay 1st derivation, and Savitzky-Golay 2nd

243   derivation were used in the spectral preprocessing process in this study.

244

245   Building a regression model

246   Partial least squares regression (PLS-R) is a multivariate analysis method used to

247   evaluate the correlation between various independent variables X and a dependent

248   variable Y (Wold et al. 1984). PLS-R was used to predict the dependent variable Y using

249   a regression equation. The PLS method used in this study is described by Equations (3)

250   and (4). The PLS regression equation generates a regression model using the spectral data

251   (X matrix, N samples × K wavelengths) and acquired parameter values as a reference (Y

252   matrix, N samples × 1).

253    $$X = TP^T + E \qquad (3)$$

254    $$Y = UQ^T + F \qquad (4)$$

255   In this context, Y is a matrix of dependent variables representing moisture, crude

256   protein, crude fat, crude fiber, and crude ash content in the BSFL. X is an n × p matrix of

257   independent variables corresponding to each spectral variable, where n is the number of

258   spectra in the sample and p represents each wavelength range (nm). Matrix X is composed

12

259   of a loading matrix P, a score matrix T, and an error matrix E. Matrix Y is composed of a

260   loading matrix Q, a score matrix U, and an error matrix F. To develop a regression model,

261   70% of the 600 data points were randomly assigned to the calibration set, and the

262   remaining 30% were assigned to the validation set during the spectrum analysis. Finally,

263   420 and 180 data points were included in the calibration and validation datasets,

264   respectively.

265

266   Regression model performance assessment

267   In this study, root mean square error (RMSE) was used to calculate the model's error rate

268   (Lee et al., 2013). The formula for calculating RMSE is shown in Equation (5).

269
$$\text{RMSE} = \sqrt{\frac{\sum_{1}^{n}(y_{i,\ \text{actual}} - y_{i,\ \text{predicted}})^2}{n}} \tag{5}$$

270   Here, $y_{i,actual}$ and $y_{i,predicted}$ represent the actual reference values obtained through chemical

271   experiments and the estimated predicted values from the developed PLS model,

272   respectively. In addition, 'n' represents the number of actual samples. The model results

273   were expressed as the coefficient of determination (R^2), which was calculated using

274   Equation (6).

275
$$R^2 = \frac{\sum_i(\hat{y}_i - \bar{y})^2}{\sum_i(y_i - \bar{y})^2} \tag{6}$$

276

277   Predicted chemical image

278   One advantage of HSI is its ability to generate chemical images of component

279   distributions by simultaneously measuring spectral and spatial data (Faqeerzada et al.,

280   2020). The beta coefficients obtained through the PLS-R analysis were used to generate

281   a chemical image of the sample. In this process, the hyperspectral image was transformed

282   into a 2D matrix, which was then multiplied by the PLS regression coefficients. The

13

283   resulting 2D matrix was then transformed back into a 3D image, and the PLS chemical

284   image was generated by summing the corresponding pixels of all band images. The

285   chemical formula is given in Equation (7):

286    $$\text{Chemical image} = \sum_{i=1}^{n} I_i R_i + C \qquad (7)$$

287   where I_i represents the hypercube image measured at the ith wavelength band, R_i

288   represents the beta coefficient values derived from the PLSR model, and C represents a

289   constant. n denotes the number of wavelengths used in this study. All analyses and

290   visualizations related to the wavelengths were performed using MATLAB 2021b

291   (MathWorks, Natick, MA, USA). Fig. 3 shows the experimental flow.

292

293   Results and Discussion

294   Proximate composition

295   Table 1 shows the proximate component analysis results for the dry matter of BSFL.

296   The moisture content of the larvae decreased significantly with increasing drying time

297   and temperature ($p < 0.05$). The moisture content decreased the most at 70°C during drying.

298   In this study, the larvae were dried at 70°C for 3 h, which was generally considered to be

299   the end of the drying process, and the moisture content of the treatment group was found

300   to be about 14.4%. Chia et al. (2020) reported that the moisture content of BSFL was

301   around 9–12%. The crude protein content of the larvae increased gradually with

302   increasing drying time and temperature, and this increase was more significant at higher

303   temperatures ($p < 0.05$). In this study, the highest crude protein content (26.2%) was

304   observed in the treatment group dried at 70°C for 3 h. Generally, the protein content of

305   BSFL powder is reported to be approximately 30–52.9% (Bessa et al., 2020), and Chia et

306   al. (2020) reported that the protein content of BSFL was approximately 31.7% when fed

307   agricultural byproducts. This study showed similar results to those of previous studies.

308   The crude fat content also increased from 30.1% to 46.2% with increasing drying time

309   and temperature (p<0.05). As the moisture content decreased, the increasing trend in the

310   proximate components in the samples led to an increase in the g/100 g protein and fat

311   percentages, which in turn increased the total crude protein and fat content. In this study,

312   the fat content was about 46.2% in the sample dried at 70°C for 3 h, which was considered

313   the end of the drying process. Chia et al. (2020) reported that the fat content of BSFL

314   varied depending on the feed (p<0.0001) and ranged from 9.5% to 49.0%. Caligiani et al.

315   (2018) analyzed BSFL using the Soxhlet ethyl ether extraction method and reported a fat

316   content of approximately 37.1%. Li et al. (2021) reported that fat content varied

317   depending on the diet of the larvae. The crude fiber content (%) showed a gradual increase

318   with increasing drying time at 60°C and 70°C, except for the 50°C dry treatment group.

319   In this study, the crude fiber content increased significantly from 2.7% to 6.8% at the end

320   of the drying process (p<0.05). Park et al. (2013) reported that the crude fiber contents of

321   BSFL and pupae were 7.47% and 7.63%, respectively. In this study, the crude fiber

322   content of BSFL significantly increased with increasing drying time and temperature

323   (p<0.05) and reached 6.4% after drying for 3 h at 70°C. Park et al. (2013) reported that

324   BSFL's dry matter crude fiber content was about 9.41%, and Chia et al. (2020) reported

325   a crude fiber content range of 6.7–12.1%. In conclusion, as the drying time increased in

326   this study, the moisture content decreased, and the amounts of crude protein, crude fat,

327   crude fiber, and ash increased. Furthermore, the results were within a range similar to

328   those reported in other studies.

329

330   Characteristic of reflectance spectra of the BSFL

331   Fig. 4 shows the SWIR hyperspectral spectral data of the BSFL. Each spectrum shows

332   the average spectrum of the group according to drying temperature and drying time. It

15

333   was confirmed that the wavelength intensity and pattern changed with drying time and

334   temperature within a specific wave range. These results suggest the possibility of a

335   proximate composition prediction using the wavelength of BSFL in the SWIR region.

336   The average spectrum can be used to observe the overall spectrum pattern for each group

337   by comparing the approximate spectral differences between the groups through spectrum

338   intensity and shape differences. However, spectrum intensity can have a high standard

339   deviation owing to noise factors such as spectrum shifts, making it more practical to

340   compare spectrum patterns rather than spectrum intensity (Park et al., 2021). In the case

341   of Fig. 4, it is difficult to confirm the trends owing to spectrum shifts. Therefore, instead

342   of comparing the spectral intensity using methods such as ANOVA, we aimed to build a

343   proximate component prediction regression model for each group by conducting PLS-R

344   analysis.

345

346   Regression model and Beta coefficient result

347   Moisture regression model and beta coefficient

348   The results of the proximate component prediction model for the BSFL are listed in

349   Table 2 and Fig. 5. The predicted results for moisture content showed a range of R^2_P 0.96-

350   0.98 and an RMSEP range of 1.83~2.59%. The preprocessed model showed higher results

351   than when using raw spectra, with the highest results shown in the model that underwent

352   maximum normalization (R^2_P=0.98, RMSEP=1.83%). To date, no studies have been

353   conducted on the development of algorithms to predict the proximate component contents

354   of edible or feed insects. However, the accuracy of the model can be verified by

355   comparing it with similar experimental results. Yu et al. (2019) used a Vis/NIR

356   hyperspectral camera to analyze the moisture content of beans using the PLSR method,

357   with 12 wavelengths and showing Rp=0.966 and RMSEP=5.105%. Huang et al. (2014)

conducted an experiment to monitor the change in moisture content of beans over drying

time using Vis/NIR, showing Rp values of 0.901–0.973 and RMSEP values in the range

of 4.6–9.2%. The results of the moisture prediction model exhibited an appropriate level

of accuracy.

Fig. 6 shows the beta coefficients of the predicted model. In general, if the beta

coefficient is high or low, the model should be weighed. In this moisture content

prediction model, the wavelengths of 1077, 1165, 1224, 1347, 1412, 1741, and 1882 nm

were determined to have weights. Wavelengths related to -OH groups significantly impact

model construction in predicting moisture content. Gergely and Salgó (2003) studied

three absorption wavelength regions of water and concluded that the ranges of 1890–1920

nm, 1400–1420 nm, 1150–1165 nm, and 1000-1100 nm were related to moisture. Among

them, the 1150–1165 nm range was reported to be a combination of the first overtone of

the O-H stretching and bending bands at 1165 nm. Furthermore, 1425 nm is known as the

first overtone region of the -CH and -OH bonds. In this study, it was determined that the

wavelength in the 1412 nm region helps predict moisture content, and it is also believed

that factors in this region contribute to this effect. According to Barbin et al. (2013), the

1400–1600 nm wavelength range is known as the stretching region of -OH and -NH. The

peak observed in the 1412 nm region in this study is believed to be a signal generated by

this overtone. Williams and Norris (1987) reported that the wavelength range of 1414 nm,

which is similar, is the O-H stretch first overtone. The wavelengths of 1077, 1224, and

1347 nm detected in the range of 1000–1350 nm are signals generated by -CH bonding

(Hoffman et al., 2023; Bobasa et al., 2021). The 1080 nm region is known as the -CH

bonding region (Muradov and Sannikov, 2007), and the 1077 nm region is considered

similar to the -CH bonding region. Kucha et al. (2020) reported that a wavelength of

1224 nm, which is close to the -CH overtone region in the 1220 nm range, can be used to

383   detect lipids or fatty acids. The overtone region of the -CH bonding contributes to the

384   prediction of moisture content because the proximate compositions of the sample are

385   interdependent, and their percentages add up to 100%. Therefore, when the moisture

386   content decreased, the percentage of lipids in the sample increased, which was detected

387   as a weight in the moisture content prediction. Holman and Edmondson (1956) explained

388   that the strong bands around 1740 and 1770 nm in their study of pure fatty acids and

389   triglycerides were derived from the C-H vibration of CH_2 groups. The first overtone peak,

390   1880 nm, is known as the absorbance of water and ester (Koumbi-Mounanga et al., 2015).

391   The wavelengths detected at 1741 nm and 1882 nm are believed to be generated by the

392   corresponding components.

393

394   Crude protein regression model and beta coefficient

395   For crude protein, the R^2_P values ranged from 0.95 to 0.99, and the RMSEP values ranged

396   from 0.55 to 0.99%. The maximum normalization method exhibited the highest accuracy

397   (R^2_P=0.99, RMSEP=0.55%). Cruz-Tirado et al. (2023) conducted an experiment to

398   predict the protein content in individual BSFL using a near-infrared (NIR) spectrometer.

399   They constructed a model using SVMR and PLSR. They reported R^2_p values ranging from

400   0.731 to 0.773 and RMSEP values ranging from 1.57% to 1.66%. In this study, the authors

401   attributed the low performance to the difficulty in accurately predicting the components

402   owing to the overlap of the chitin signal with the protein signal. In contrast, the current

403   study demonstrated a relatively high accuracy and low RMSEP compared to the previous

404   study, which may be attributed to the inclusion of additional wavelength information for

405   predicting moisture, crude protein, and crude fat content in the model. The beta

406   coefficients for crude protein were 1224, 1353, 1394, 1541, 1735, 1882, and 1941 nm.

407   Wavelengths of 1224, 1353, and 1735 nm were used to predict -CH in this case (Hoffman

et al., 2023; Bobasa et al., 2021). Cruz-Tirado et al. (2023) constructed a principal

component (PC) model to predict proteins in BSFL and detected 1760 nm in PC1, which

they reported to be the necessary wavelength for predicting fatty acids. In the current

study, although there was a slight difference in the wavelength, the wavelength range of

1735 nm was assumed to be a signal from the –CH bond because of its similarity to the

necessary wavelength reported by Cruz-Tirado et al. (2023). In addition, wavelengths of

1394 and 1541 nm were also detected in the beta coefficients for moisture content and

belonged to the overtone regions of -NH and -OH, which are overlapping wavelengths

for predicting crude protein. Furthermore, 1882 nm was considered to be the beta value

associated with -OH. According to Cruz-Tirado et al. (2023), the signal at 1900 nm is

assumed to originate from -NH, and the signal at 1941 nm is considered to originate from

this -NH region.

Crude fat regression model and beta coefficient

According to the study, the prediction of the crude fat content showed an R^2_P range of

0.87-0.91 and an RMSEP of 1.34–1.67%, and the best performance was achieved by mean

normalization (R^2_P =0.91, RMSEP=1.34%). According to Caporaso et al. (2021), the

standard deviation (SD) of the AOAC method 922.06 for fat content analysis by acid

hydrolysis ranges from 0.7% to 7.5% depending on the type of food analyzed. Therefore,

the model prediction results of this experiment are considered to be applicable to

nondestructive tools. The beta coefficients for fat content were 1224, 1288, 1412, 1723,

and 1888 nm. The peaks at 1224, 1288, and 1723 nm are associated with the overtone

region related to –CH. Choi et al. (2021) stated that this region constitutes fat-and fatty

acid-related areas in the wavelength range of 1600–1800 nm. In addition, 1412 nm and

1888 nm were identified as the regions associated with-OH. The reason why the

433   wavelength associated with –OH (1412, 1888 nm) was detected as an important

434   wavelength for crude fat prediction is that the content of the proximate composition is

435   calculated in %. When the moisture content of a proximate composition decreases, the %

436   unit of other crude protein and crude fat, which are relatively reference values, increases.

437   Based on this result, it is judged that the wavelength region related to moisture also affects

438   the construction of the crude fat model.

439

440   Crude fiber regression model and beta coefficient

441   The R^2_P of the crude fibers ranged from 0.85 to 0.89, and the RMSEP ranged from 0.46%

442   to 0.53%. In terms of latent variables (LV), crude fiber showed a diverse range of 14–17

443   LVs, indicating that the model is complex compared with other models for proximate

444   composition. Among the preprocessing models for crude fiber, the model with mean

445   normalization exhibited the highest accuracy (R^2_P =0.89, RMSEP=0.46%). The beta

446   coefficients for the crude fiber model in Fig. 6 show that 16 wavelengths (1142, 1171,

447   1194, 1241, 1388, 1424, 1541, 1629, 1729, 1894, 1911, 2088, 2146, 2217, 2264, 2270,

448   and 2270 nm) were relatively important peaks compared to other wavelengths. Chitin is

449   a representative example of a major component of crude fiber. Chitin is a polysaccharide

450   structure composed of multiple N-acetyl-D-glucosamine molecules containing nitrogen.

451   The exoskeletons of insects and crustaceans, including BSFL, are composed of chitin.

452   Brigode et al. (2020) conducted a study to evaluate the properties of biopolymer films

453   produced using chitin from BSFL. And also this chitin can be applied to making other

454   functional materials like chitosan. Chitosan can be obtained due to the deacetylation of

455   chitin, has antibacterial properties against fungi and bacteria, and can be used to reduce

456   the use of antibiotics in animals (Riaz Rajoka et al., 2020). Typically, the chitin content

457   of black soldier fly prepupae is reported to be approximately 9-10% (Soetmans et al.,

458 2020). Cruz-Tirado et al. (2023) reported that the regions at 2150 nm, 2256 nm, and

459 2337 nm are associated with chitin content and are connected with 2 × amide I + 2 ×

460 amide II, O-H stretching + O-H deformation, and C-H stretching + C-H deformation

461 (Cruz-Tirado et al., 2023; Osborne, 2006; Shetty et al., 2012). Cruz-Tirado et al. (2023)

462 estimated that the 2000–2500 nm range is associated with chitin. Although crude fiber

463 does not completely represent chitin, it is assumed that chitin is mixed with some of the

464 substances that make up crude fiber. In this study, wavelengths ranging from 2100 to

465 2350 nm were helpful in predicting the crude fiber content.

466

467 Crude ash regression model and beta coefficient

468 The model accuracy of the crude ash sample had an R^2_P range of 0.94-0.96, and an

469 RMSEP range of 0.25-0.32%. Among them, the preprocessing method using the mean

470 normalization technique showed the highest accuracy for R^2_P at 0.96 and the lowest

471 RMSEP at 0.25% (Table 2). The main beta coefficient wavelengths of the ash samples

472 were found in the 1224, 1353, 1400, 1735, and 1923 nm regions, and their shapes were

473 similar to those of the beta coefficients of crude protein (Fig. 6). In theory, energy is not

474 absorbed by inorganic substances such as ash in the NIR region. Therefore, the ash

475 content cannot be directly determined by NIR (He et al., 2023). However, many

476 wavelengths in the NIR region used in the calibration development process are expected

477 to be predicted by correlation with the total amount of organic compounds and moisture

478 because they provide important information. (Pojić et al., 2010).

479

480 Chemical image of BSFL

481 Unlike spectrometers, hyperspectral images contain wavelength information for each

482 pixel, making it possible to visualize information that is difficult to see with the naked

eye. Therefore, in this study, chemical images were created for each component, including moisture, crude protein, crude fat, crude fiber, and crude ash content, and visualized according to their respective concentrations (Fig. 7). Red pixels represent high concentrations and dark blue pixels indicate low concentrations. As the drying time and temperature increased, the moisture content decreased gradually, which was monitored by observing an increasing number of blue pixels. For crude protein, crude fat, crude fiber, and crude ash, the number of red pixels increased with the drying time and temperature. It is confirmed that the proposed prediction model performs well.

Conclusion

In this study, we developed a proximate component prediction algorithm based on SWIR HSI in the 1000–2350 nm range for dried raw materials, according to the drying time and drying temperature of BSFL. A model was developed for moisture, crude protein, crude fat, crude fiber, and crude ash contents. Through this study, it is anticipated that it will be possible to classify defective factors and incompletely dried individuals in the dried raw materials of BSFL. The results of this study are deemed suitable for detecting the nutritional components in BSFL and for use in the manufacturing of mixed feed by feed companies. We anticipate that this will enable quality control of dried raw materials from BSFL. However, further development of a rapid detection technology for BSFL is necessary for real-time sorting machine production, and additional research is required for this purpose. In particular, for BSFL, it is necessary to classify them based not only on the feed source but also on the individuals raised using food waste and the larvae used for composting livestock manure, as their nutritional components can vary depending on the feed source. In Korea, BSFL raised using livestock manure cannot be used as feed; therefore, it is necessary to develop a classification technology for such larvae. We hope

508   that the results of this study can be utilized as a basis for the development of sorting

509   machines for BSFL.

510

References

Van Huis A. 2013. Potential of insects as food and feed in assuring food security. Ann Rev Eentomol, 58:56-583.

Jongema Y. 2017. Worldwide list of edible insects (April 1, 2017). Available online: https://www.wur.nl/upload_mm/8/a/6/0fdfc700-3929-4a74-8b69-f02fd35a1696_Worldwide%20list%20of%20edible%20insects%202017.pdf (accessed on 19 April 2023).

Hartmann C, Shi J, Giusto A, Siegrist M. 2015. The psychology of eating insects: A cross-cultural comparison between Germany and China. Food Qual Prefer 44: 148-156.

Jensen NH, Lieberoth. 2019. We will eat disgusting foods together-Evidence of the normative basis of Western entomophagy-disgust from an insect tasting. Food Qual Prefer 72:109-115.

Van Raamsdonk LWD, van der Fels-Klerx HJ, de Jong J. 2017. New feed ingredients: the insect opportunity. Food Addit Contam Part A Chem Anal Control Expo Risk Assess 34(8):1384-1397.

Norgales-Mérida S, Gobbi P, Józefiak D, Mazurkiewicz J, Dudek K, Rawski M, Kierończyk B, Józefiak A. 2019. Insect meals in fish nutrition. Rev Aquac 11:1080-1103.

Kępińska-Pacelik J, Biel W. 2022. Insects in pet food industry-hope or threat? Animals (Basel) 12:1515.

Boerema A, Peeters A, Swolfs S, Vandevenne F, Sander J, Staes J, Meire P. 2016. Soybean trade: balancing environmental and socio-economic impacts of an intercontinental market. PloS one 11: e0155222.

Veldkamp T, Bosch G. 2015. Insects: A protein-rich feed ingredient in pig and poultry diets. Anim Front 5: 45–50.

Ji YJ, Liu HN, Kong XF, Blachier F, Geng MM, Liu YY, Yin YL. 2016. Use of insect powder as a source of dietary protein in early-weaned piglets. J Anim Sci 94(3): 111-116.

Pieterse E, Erasmus SW, Uushona T, Hoffman LC. 2019. Black soldier fly (Hermetia illucens) pre-pupae meal as a dietary protein source for broiler production ensures a tasty chicken with standard meat quality for every pot. J Sci Food Agric 99:893-903.

Cullere M, Tasoniero G, Giaccone V, Acuti G, Marangon A, Zotte AD. 2017. Black soldier fly as dietary protein source for broiler quails: meat proximate composition, fatty acid and amino acid profile, oxidative status and sensory traits. Animal 12:640-647.

Zarantoniello M, Zimbelli A, Randazzo B, Compagni MD, Truzzi C, Antonucci M, Giorgini E. 2020. Black soldier fly(Hermetia illucens) reared on roasted coffee by-product and Schizochytrium sp. As a sustainable terrestrial ingredient for aquafeeds production. Aquaculture 518:734659.

560 Drewery ML, Liu X, Wickersham TA. 2022. Black soldier fly larvae (BSFL) as a feed for
561 beef cattle: a hedonic pricing model. J Insects Food Feed 8: 743-751.
562
563 Caimi C, Renna M, Lussiana C, Bonaldo A, Gariglio M, Meneguz M, Dabbou S,
564 Schiavone A, Gau F, Concetta A, Prearo M, Gasco L. 2020. First insights on Black soldier
565 fly(Hermetia illucens L.) larvae meal dietary administration in Siberian sturgeon
566 (Acipenser baerii Brandt) juveniles. Aquaculture 515:734539.
567
568 Nam JH, Kim D, Hyun JY, Jin HJ, Choi YS, Cho JH, Lee BK, Chun JY. 2022. Current
569 status and future prospects of the insect industry as an alternative protein source for
570 animal feed. J Korean Soc Food Sci Nutr 51:395-402.
571
572 Sheppard DC, Tomberlin JK, Joyce JA, Kiser BC, Sumner SM. 2002. Rearing methods
573 for the black soldier fly (Diptera: Stratiomyidae). Journal of Medical Entomology, 39,
574 695–698.
575
576 Bessa LWm, Pieterse E, Marais J, Hoffman LC. 2020. Why for feed and not for human
577 consumption? The black soldier fly larvae. Compr Rev Food Sci Food Saf 19: 2747-2763.
578
579 Diclaro JW, Kaufman PE. 2009. Black soldier fly Hermetia illucens linnaeus (insecta:
580 Diptera: Stratiomyidae). EENY, 461, 1-3.
581
582 Surendra KC, Tomberlin JK, van Huis A, Cammack JA, Heckmann LHL, Khanal SK.
583 2020. Rethinking organic wastes bioconversion: evaluating the potential of the black
584 soldier fly(Hermetia illucens (L.))(Diptera: Stratiomyidae)(BSF). Waste Manag 117:58-
585 80.
586
587 Erbland P, Alyokhin A, Peterson M. 2021. An automated incubator for rearing black
588 soldier fly larvae (Hermetia Illucens). Trans ASABE 64:1989-1997.
589
590 Thrastardottir R, Olafsdottir HT, Thorarinsdottir RI. 2021. Yellow mealworm and black
591 soldier fly larvae for feed and food production in Europe, with emphasis on Iceland.
592 Foods (Basel) 10:2744.
593
594 Benes E, Biró B, Fodor M, Gere A. 2022. Analysis of wheat flour-insect powder mixtures
595 based on their near infrared spectra. Food Chem X 13: 100266.
596
597 Cruz-Tirado JP, Amigo JM. Barbin DF. 2023. Determination of protein content in single
598 black fly soldier (Hermetia illucens L.) larvae by near infrared hyperspectral imaging
599 (NIR-HSI) and chemometrics. Food Cont 143:109266.
600
601 Saucier L, M'ballou C, Ratti C, Deschamps MH, Lebeuf Y, Vandenberg GW. 2022.
602 Comparison of black soldier fly larvae pre-treatments and drying techniques on the
603 microbial load and physico-chemical characteristics. J Insects Food Feed 8:45-64.
604
605 Larouche J, Deschamps MH, Saucier L, Lebeuf Y, Doyen A, Vandenberg GW. 2019.
606 Effects of killing methods on lipid oxidation, colour and microbial load of black soldier
607 fly (hermetia illucens) larvae. Animals (Basel) 9:182.
608

609 Janssen RH, Vincken JP, Arts NJG, Fogliano V, Lakemond CMM. 2019a. Effect of
610 endogenous phenoloxidase on protein solubility and digestibility after processing of
611 Tenebrio molitor, Alphitobius diaperinus and Hermetia illucens. Food Res Int 121: 684-
612 690.
613
614 Janssen RH, Canelli G, Sanders MG, Bakx EJ, Lakemond CMM, Fogliano V, Vincken JP.
615 2019b. Iron-polyphenol complexes cause blackening upon grinding Hermetia illucens
616 (black soldier fly) larvae. Sci Rep 2019:2967.
617
618 AOAC [Association of Official Analytical Chemists] International. Official methods of
619 analysis of AOAC International. 18th ed. Gaithersburg, MD: AOAC International; 2005.
620
621 Janssen RH, Vincken JP, van den Broek LAM, Fogliano V, Lakemond CMM. 2017.
622 Nitrogen-to-protein conversion factors for three edible insects: Tenebrio molitor,
623 Alphitobius diaperinus, and Hermetia illucnes. J Agric Food Chem65:2275-2278.
624
625 Wold S, Ruhe A, Wold H, Dumm Iii WJ. 1984. The collinearity problem in linear
626 regression. The partial least squares (PLS) approach to generalized inverses. SIAM
627 Journal on scientific and statistical computing 5:735-743.
628
629 Lee H, Cho BK, Kim MS, Lee WH, Tewari J, Bae H, Sohn SI, Chi HY. 2013. Prediction
630 of crude protein and oil content of soybeans using Raman spectroscopy. Sensors
631 Actuators B Chem. 185:694–700.
632
633 Faqeerzada MA, Lohumi S, Kim G, Joshi R, Lee H, Kim MS, Cho BK. 2020.
634 Hyperspectral shortwave infrared image analysis for detection of adulterants in almond
635 powder with one-class classification method. Sensors (Basel) 20:5855.
636
637 Chia SY, Tanga CM, Osuga IM, Cheseto X, Ekesi S, Dicke M, van Loon JJA. 2020.
638 Nutritional composition of black soldier fly larvae feeding on agro-industrial by products.
639 Entomol Exp Appl 168:472-481.
640
641 Caligiani A, Marseglia A, Leni G, Baldassarre S, Maistrello L, Dossena Am Sforza S.
642 2018. Composition of black soldier fly prepupae and systematic approaches for extraction
643 and fraction of proteins, lipids and chitin. Food Res Int 105:812-820.
644
645 Li X, Zhou Z, Zhang J, Zhou S, Xiong Q. 2022. Conversion of mixtures of soybean curd
646 residue and kitchen waste by black soldier fly larvae (Hermetia illuces L.). Insect (Basel)
647 13:23.
648
649 Park K, Choi Y, Nam S, Kim S, Kim S, Ma Y, No S. 2013. Nutritional value of black
650 soldier fly, Hermetia illucens (Diptera: Stratiomyidae) as a feed supplement for fish.
651 JSeric Entomol Sci 51:95-98.
652
653 Park E, Kim YS, Omari MK, Suh HK, Faqeerzada MA, Kim MS, Baek I, Cho BK. 2021.
654 High-throughput phenotyping approach for the evaluation of heat stress in Korean
655 Ginseng (Panax ginseng meyer) using a hyperspectral reflectance image.Sensor(Basel)
656 21:5634.
657

658 Yu P, Huang M, Zhang M, Yang B. 2019. Optimal wavelength selection for hyperspectral
659 imaging evaluation on vegetable soybean moisture content during drying. Appl Sci(Basel)
660 9:331.
661
662 Huang M, Wang Q, Zhang M, Zhu Q. 2014. Prediction of color and moisture content for
663 vegetable soybean during drying using hyperspectral imaging. J Food Eng 128:24-30.
664
665 Gergely S, Salgó A. 2003. Changes in moisture content during wheat maturation—what
666 is measured by near infrared spectroscopy? J Near Infrared Spectrosc 11:17-26.
667
668 Barbin DF, ElMasry G, Sun DW, Allen P. 2013. Non-destructive determination of
669 chemical composition in intact and minced pork using near-infrared hyperspectral
670 imaging. Food Chem 138:1162-1171.
671
672 Williams P, Norris K. 1987. Near-infrared Technology in the Agricultural and Food
673 Industries. American Association of cereal chemists, Inc
674
675 Hoffman L, Ingle P, Khole AH, Zhang S, Yang Z, Beya M, Bureš D, Cozzolino D. 2023.
676 Discrimination of lamb (Ovis aries), emu (Dromaius novaehollandiae), camel (Camelus
677 dromedarius) and beef (Bos taurus) binary mixtures using a portable near infrared
678 instrument combined with chemometrics. Spectrochim Acta A Mol Biomol Spectrosc
679 294:122506.
680
681 Bobasa E, Phan ADT, Netzel M, Smyth HE, Sultanbawa Y, Cozzolino D. 2021. The use
682 of a micro near infrared portable instrument to predict bioactive compounds in a wild
683 harvested fruit-kakadu plum (Terminalia ferdinandiana). Sensors (Basel) 21:1413.
684
685 Muradov VG, Sannikov DG. 2007. Study of absorption spectra of gaolines and other
686 hydrocarbon mixtures in the second overtone region of the CH3, CH2, CH groups. J Appl
687 Spectrosc 74:174-179.
688
689 Kucha CT, Ngadio MO. 2020. Rapid assessment of pork freshness using miniaturized
690 NIR spectroscopy. Journal of Food Measurement and Characterization 14: 1105-1115.
691
692 Holman RT, Edmondson PG. 1956. Near-infrared spectra of fatty acids and some related
693 substances. Anal Chem 28: 1533-1538.
694
695 Koumbi-Mounanga T, Groves K, Leblon B, Zhou G, Cooper PA. 2015. Estimation of
696 moisture content of trembling aspen (Populus tremuloides Michx.) strands by near
697 infrared spectroscopy (NIRS). Eur J Wood Prod 73:43-50.
698
699 Caporaso N, Whitworth MB, Fisk ID. 2021. Total lipid prediction in single intact cocoa
700 beans by hyperspectral chemical imaging. Food Chem 344:128663.
701
702 Choi JY, Kim HC, Moon KD. 2021. Geographical origin discriminant analysis of chia
703 seeds (Salvia hispanica L.) using hyperspectral imaging. J Food Comp and Anal
704 101:103916.
705

706 Brigode C, Hobbi P, Jafari H, Verwilghen F, Baeten E, Shavandi A. 2020. Isolation and
707 physicochemical properties of chitin polymer from insect farm side stream as a new
708 source of renewable biopolymer. J Clean Prod 275:122924.
709
710 Riaz Rajoka MS, Mehwish HM, Wu Y, Zhao L, Arfat Y, Majeed K, Anwaar S. 2020.
711 Chitin/chitosan derivatives and their interactions with microorganisms: a comprehensive
712 review and future perspectives. Crit Rev Biotechnol 40:365-379.
713
714 Soetemans L, Uyttebroek M, Bastiaens L. 2020. Charcteristics of chitin extracted from
715 black soldier fly in different life stages. Int JBiol Macromol 165:3206-3214.
716
717 Osborne BG. 2006. Near-infrared spectroscopy in food analysis. Encycl Anal Chem Appl
718 Theory Instrum
719
720 Shetty N, Gislum R, Jensen AMD, Boelt B. 2012. Development of NIR calibration
721 models to assess year-to-year variation in total non-structural carbohydrates in grasses
722 using PLSR.Chemom Intell Lab Syst 111: 34-38.
723
724 He HJ, Wang Y, Wang Y,Liu H, Zhang M, Ou X. 2023.Simultaneous quantifying and
725 visualizing moisture, ash and protein distribution in sweet potato [Ipomoea batatas (L.)
726 Lam] by NIR hyperspectral imaging. Food Chem X 18:110631.
727
728 Pojić M, Mastilović J, Palić D, Pestorić M. 2010. The development of near-infrared
729 spectroscopy (NIRS) calibration for prediction of ash content in legumes on the basis of
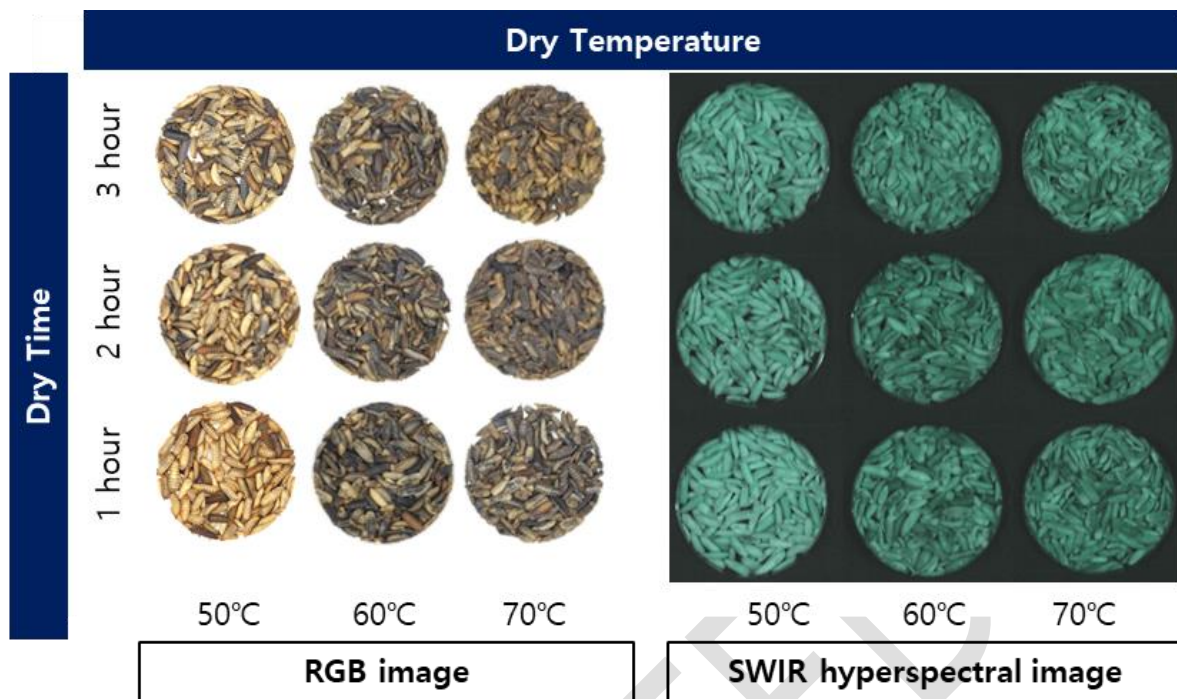730 two different reference methods. Food Chem 123:800-805.
731

732
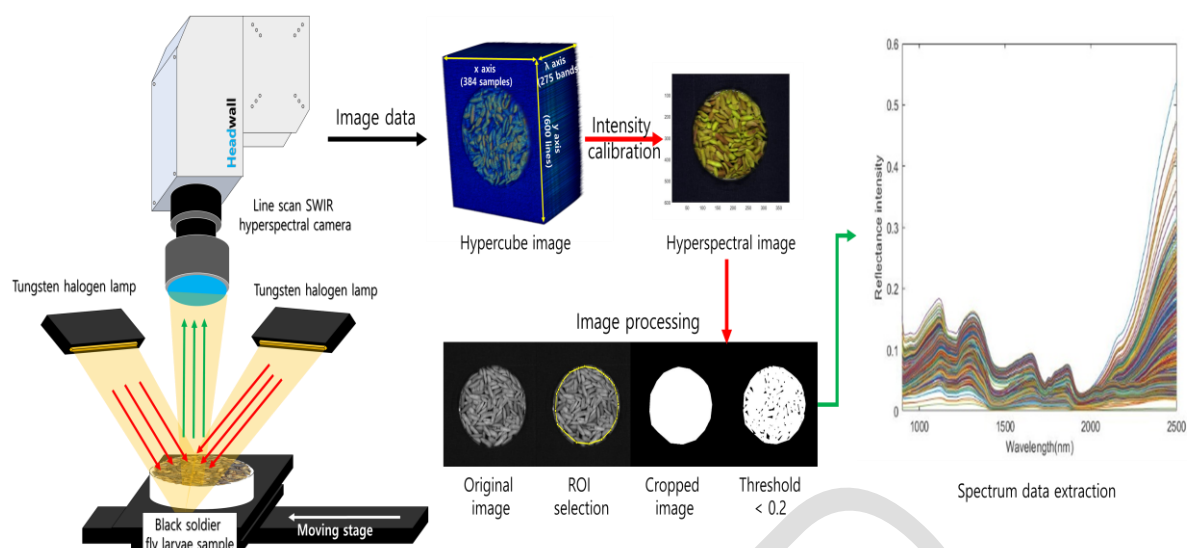733   Fig. 1. The SWIR Hyperspectral sample images of black soldier fly larvae samples.

734

735

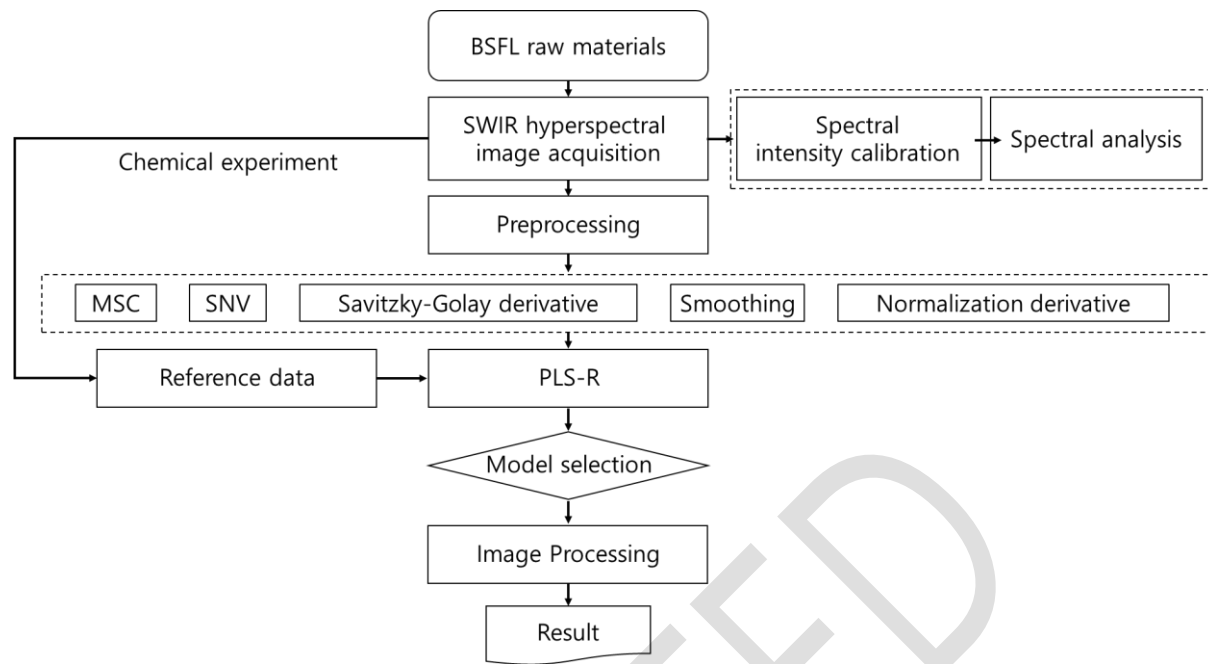736   Fig. 2. Image acquisition step using SWIR hyperspectral imaging system.

737

738

739
740 Fig. 3. The algorithm development workflow for making black soldier fly larvae
741 proximate content prediction model.
742

743
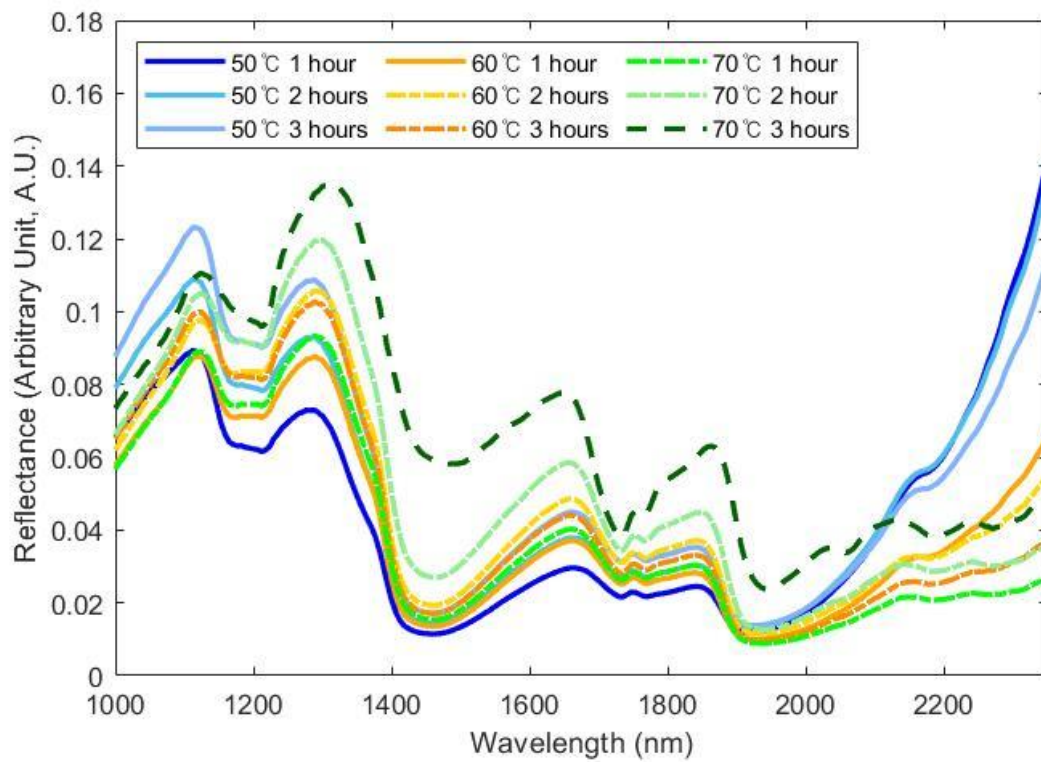744   Fig. 4. The results of ~~range normalization~~MSC preprocessed spectrum (1000-2350 nm)

745   in black soldier fly larvae.

746

747
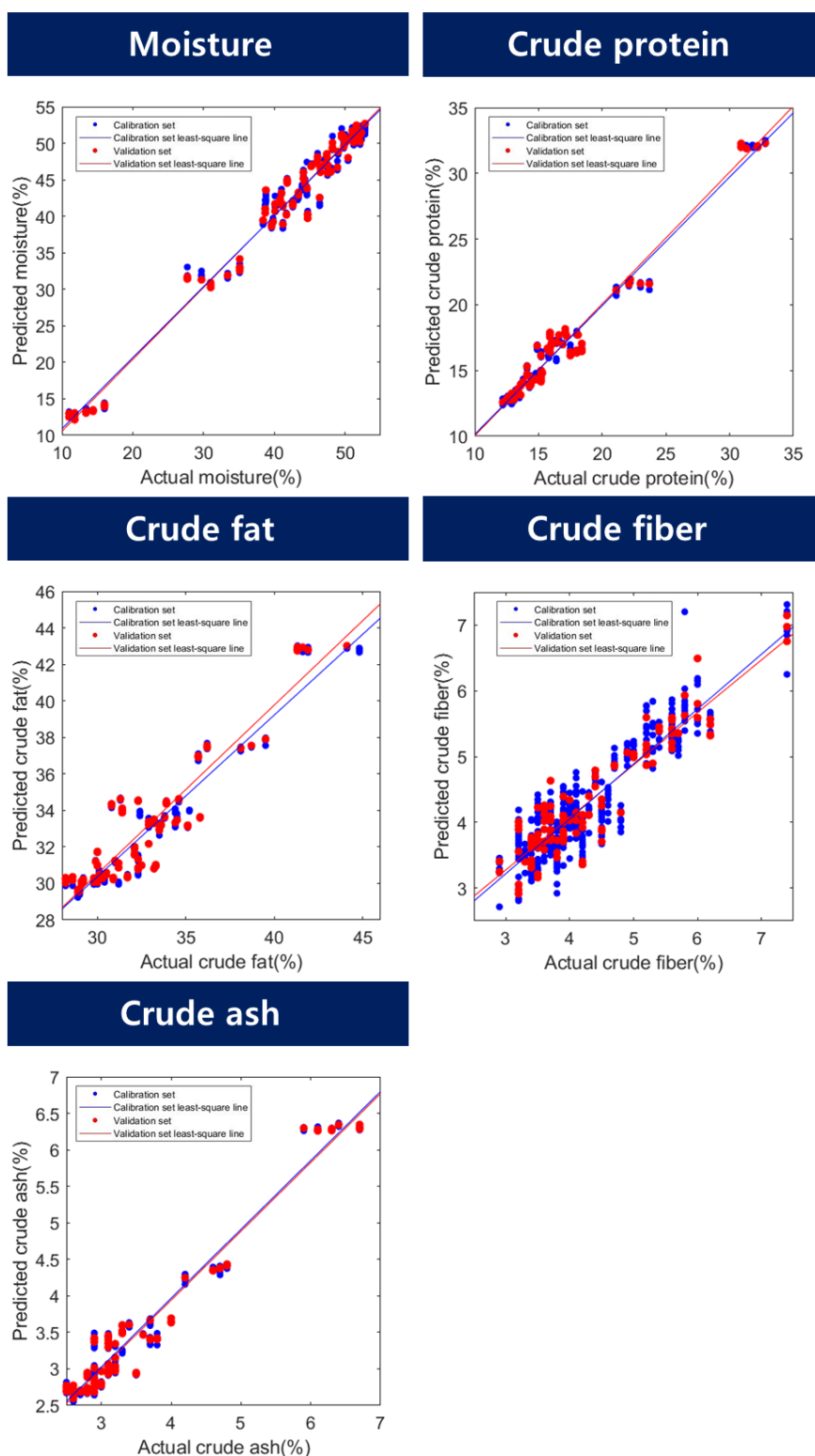
748   Fig. 5. The scattering plot of a prediction model for black soldier fly larvae proximate

749   content.
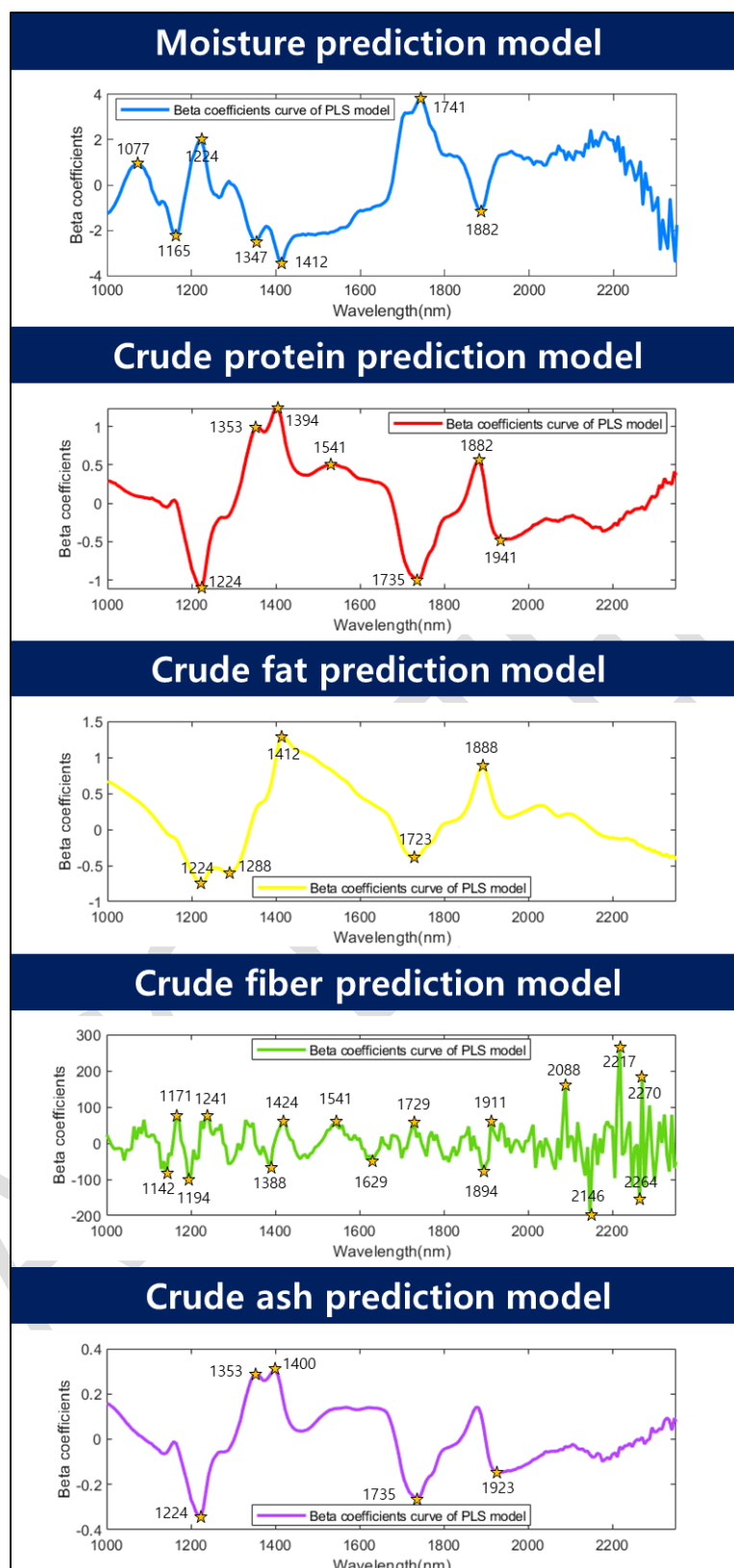
750

751   Fig. 6. Beta coefficient of full wavelength range (1000 to 2350 nm) prediction models
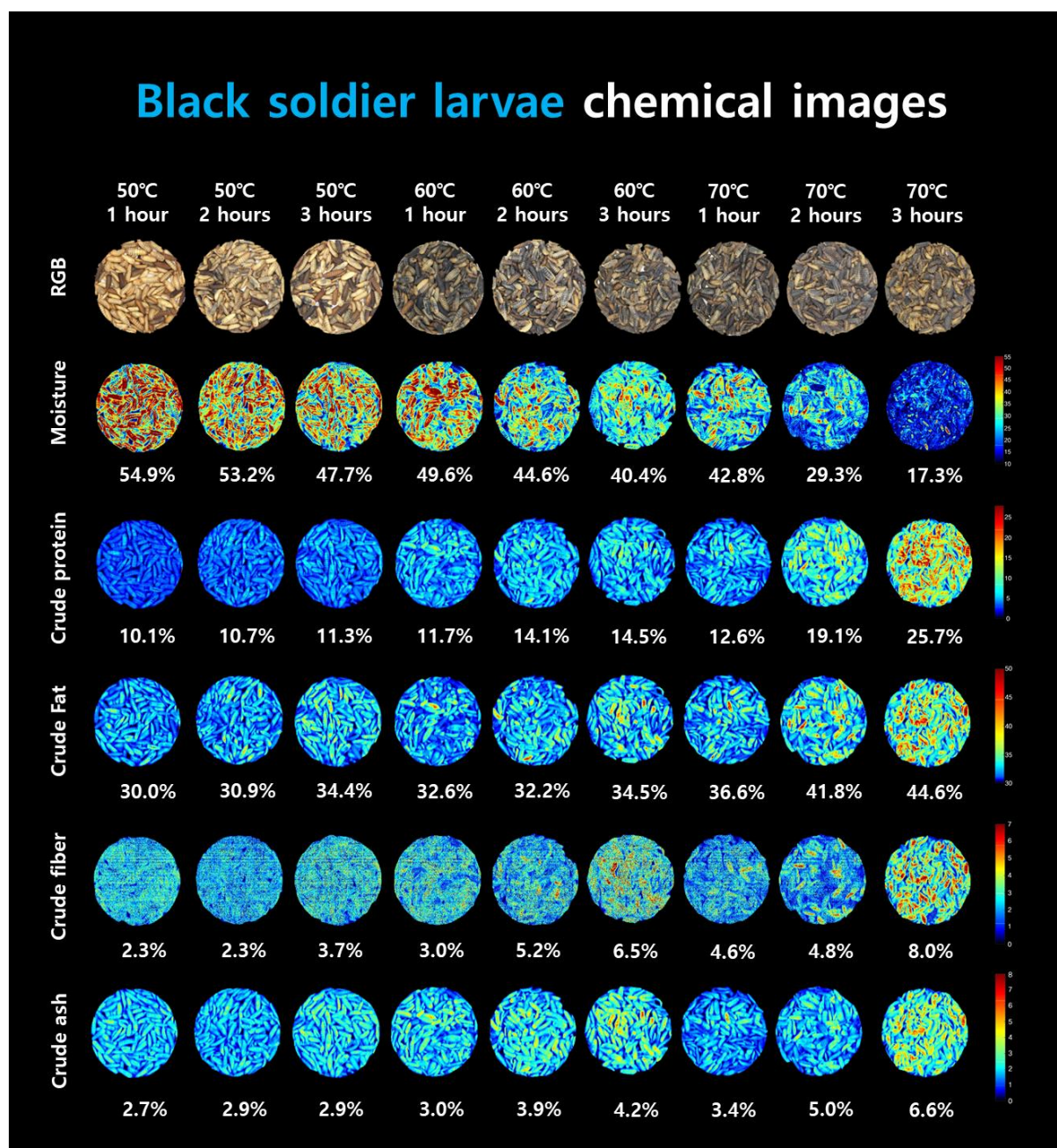
752

753 Fig. 7. The chemical images of black soldier fly larvae.

754

Table 1. Black soldier fly larvae proximate composition (%) results

Dry temperature	50°C			60°C			70°C			SEM	*P* value		
Dry time	1 hour	2 hours	3 hours	1 hour	2 hours	3 hours	1 hour	2 hours	3 hours		Temp (T1)	Time (T2)	Inter (T1×T2)
Moisture (%)	54.2^{ax}	52.0^{bx}	48.8^{cx}	48.6^{ay}	*43.7^{by}*	41.0^{by}	45.4^{az}	33.2^{bz}	14.4^{cz}	0.97	***	***	***
Crude protein (%)	9.9^{cz}	10.3^{bz}	11.2^{az}	11.7^{by}	*14.0^{ay}*	14.2^{ay}	12.4^{cx}	18.1^{bx}	26.2^{ax}	0.40	***	***	***
Crude fat (%)	30.1^{cz}	31.6^{bz}	33.0^{az}	32.7^{by}	*34.4^{ay}*	35.2^{ay}	34.9^{cx}	39.5^{bx}	46.2^{ax}	0.41	***	***	***
Crude fiber (%)	3.2^{by}	3.3^{by}	4.0^{ay}	3.8^{bx}	4.7^{bx}	5.9^{ax}	4.1^{bx}	4.3^{bx}	6.4^{ax}	0.12	***	***	**
Crude ash (%)	2.7^{cz}	2.8^{bz}	3.0^{az}	3.2^{by}	*3.7^{ay}*	3.7^{ay}	3.2^{cx}	4.8^{bx}	6.8^{ax}	0.10	***	***	***

756 Temp: *p*-value of the dry temperature; Time: *p*-value of the dry time; Inter(T1×T2): interaction *p*-value of the dry temperature with
757 the dry time. *: p-value < 0.05; **: p-value <0.01, ***: p-value <0.001; SEM: Standard error of the mean.
758 a-c: Mean values within each row with different superscripts are significantly different about drying time (*p*-value < 0.05).
759 x-z: Mean values within each row with different superscripts are significantly different about drying temperature (*p*-value < 0.05).
760 Proximate contents were calculated as dry matter (DM).
761

762 Table 2. Prediction model results of black soldier fly larvae proximate contents using the
763 SWIR hyperspectral imaging system
764

Parameter	Preprocessing	Whole insect sample				
		R_c^2	RMSEC (%)	R_p^2	RMSEP (%)	LV
Moisture	Mean norm	0.97	1.80	0.98	1.92	5
	Max norm	0.97	1.79	0.98	1.83	5
	Range norm	0.95	2.22	0.96	2.44	4
	MSC	0.96	2.00	0.97	2.07	4
	SNV	0.96	1.97	0.97	2.05	5
	SG 1st	0.94	2.55	0.96	2.41	5
	SG 2nd	0.93	2.71	0.96	2.59	3
	Raw	0.94	2.59	0.96	2.46	5
Crude protein	Mean norm	0.98	0.59	0.98	0.57	4
	Max norm	0.98	0.58	0.99	0.55	5
	Range norm	0.97	0.73	0.97	0.78	4
	MSC	0.98	0.62	0.98	0.61	4
	SNV	0.98	0.61	0.98	0.59	5
	SG 1st	0.96	0.92	0.95	0.98	5
	SG 2nd	0.96	0.93	0.95	0.99	3
	Raw	0.96	0.92	0.95	0.97	5
Crude fat	Mean norm	0.91	1.34	0.91	1.34	4
	Max norm	0.90	1.44	0.90	1.41	4
	Range norm	0.89	1.47	0.90	1.44	4
	MSC	0.91	1.36	0.91	1.39	4
	SNV	0.91	1.37	0.91	1.38	5
	SG 1st	0.88	1.57	0.88	1.61	4
	SG 2nd	0.87	1.60	0.88	1.61	3
	Raw	0.87	1.63	0.87	1.67	4
Crude fiber	Mean norm	0.87	0.45	0.89	0.46	16
	Max norm	0.87	0.46	0.89	0.46	17
	Range norm	0.87	0.46	0.89	0.46	17
	MSC	0.85	0.48	0.85	0.53	15
	SNV	0.85	0.49	0.85	0.52	16
	SG 1st	0.87	0.46	0.86	0.51	14
	SG 2nd	0.86	0.47	0.85	0.53	14
	Raw	0.86	0.48	0.86	0.51	17
Crude ash	Mean norm	0.95	0.24	0.96	0.25	4
	Max norm	0.96	0.23	0.96	0.25	5
	Range norm	0.94	0.27	0.94	0.30	4
	MSC	0.95	0.25	0.96	0.26	4
	SNV	0.95	0.24	0.95	0.27	5
	SG 1st	0.93	0.29	0.95	0.30	5
	SG 2nd	0.92	0.30	0.94	0.32	3
	Raw	0.92	0.30	0.94	0.30	5

765 SWIR: short wavelength infrared hyperspectral imaging system; Mean norm: mean normalization; Maximum norm: Maximum
766 normalization; Range norm: Range normalization; MSC: multiplicative scatter correction; SNV: regular normal variate; SG 1st :
767 Savitzky-Golay 1st derivation, SG 2nd : Savitzky-Golay 2nd derivation; Raw: Raw spectrum; R_c^2: coefficient of determination of
768 calibration set; RMSEC: root mean square error of calibration set; R_P^2: coefficient of determination of prediction set; RMSEP: root
769 mean square error of prediction set; LV: Latent variables

770
771