

# Monitoring of soybean germination process by near-infrared spectroscopy

M. Bartalné-Berceli\* , E. Izsó, Sz. Gergely and A. Salgó

Department of Applied Biotechnology and Food Science, Budapest University of Technology and Economics, Szent Gellért tér 4, Budapest, H-1111, Hungary

## ORIGINAL RESEARCH PAPER

Received: October 19, 2021 • Accepted: January 27, 2022

© 2022 The Author(s)



### ABSTRACT

Soybean seeds were germinated on an industrial scale after soaking for 0–56 h to produce a special additive for food industrial use. The germination process of three soybean varieties was monitored with near-infrared (NIR) spectroscopy based on changes in the amount, status, or character of the water. This paper evaluates the “waterless” NIR spectra of sound, germinated, and heat treated seeds to try to follow the fine details of the germination process. The germination process was analysed with the help of cluster analysis (CA), principal component analysis (PCA), and polar qualification system (PQS) as statistical and chemometric methods. PCA proved to be the most sensitive spectrum evaluation method to follow the fine details of germination. The applied NIR method is suitable for non-destructively, real-time monitoring of the non-linear nature of germination.

### KEYWORDS

germination, soybean, near-infrared spectroscopy, cluster analysis (CA), principal component analysis (PCA), polar qualification system (PQS)

## 1. INTRODUCTION

Germinated seeds are extremely valuable from a nutritional point of view. Plant seeds in dormancy phase are activated by water uptake. Activation is realised only under suitable

\* Corresponding author. Tel.: +361 4631477. E-mail: bartalne.berceli.monika@vbk.bme.hu

external and internal conditions (temperature, water activity, oxygen, viability, dormancy, etc.). The imbibition is followed by rehydration and swelling, meanwhile the activation and *de novo* synthesis of enzymes occurs. Thereafter, the mobilisation of stored reserves occurs, nutrients are transformed to an easier accessible form. As the degree of polymerisation is reduced, the biological process results a more valuable plant material (Graeber et al., 2012; Bartalné-Berceli et al., 2016) with improved potential for application (Bartalné-Berceli et al., 2018).

Soybean with its biologically valuable composition (low carbohydrate, high and complete protein contents) is a unique source of plant material with excellent functional properties, being an important element in the vegetarian diet. Soybean is rich in vitamins, minerals, polyunsaturated fatty acids, and bioactive components. Its isoflavonoid content is beneficial, decreasing the level of cholesterol and being advantageous for a human being (Erdman and Fordyce, 1989).

Using short-term germination procedure (Fitorex Engineering and Trading Ltd. 2008), the composition and nutritional quality of native soybean can be improved significantly, and some additional benefits of this process can be exploited in food innovation.

In order to control and measure the details of the germination process, near-infrared techniques can be used (Juhász et al., 2007). NIR technology is fast, non-destructive, and requires minimal sample preparation and small sample size. Near-infrared spectra are carrying physical and chemical information of the sample (Gergely and Salgó, 2007). NIR methods are able to follow and control different bio-processes (biotechnologies, physiological processes, fermentation, storage, etc.) real time, so the beginning of germination process can easily be detected (Allosio and Boivin, 1997; Smail et al., 2006). Wavelength and intensity of water peaks in NIR spectra depend on: (i) how much bound moisture is present, i.e. how many water molecules form H-bridges, (ii) how strong bounding are formed by the water molecules (Abe, 2004). The maximum location of the absorption bands also depends on what kind of macromolecules interacts with the water molecules. The transformation of the degree of hydration of macromolecules is determinative in the physiological processes of the plant and can be sensitively monitored by NIR measurements (Gergely and Salgó, 2003).

The aim of this study was to analyse the changes in soybean seeds during germination process using NIR method. The effects of different mathematical treatments of the spectra and also the applicability and efficiency of different chemometric methods at pre-defined germination stages have been examined.

## 2. MATERIALS AND METHODS

### 2.1. Samples and germination

Three different varieties of soybean grown in Hungary were examined. An international one named Kanadai (Canadian) originating from Újmohács, a Hungarian variety named Pannónia Kincse originating from Hernád, and a bio-version of the Hungarian variety Bio Pannónia Kincse originating from Babócsa. Germination and sampling of the three varieties were carried out simultaneously. The soybean seeds were intact before the germination process began. They were marked as dry seeds at this stage (marked with “sz” in Figures). Then the seeds were soaked for 2 h in tap water, and only this point was considered as the end of imbibition and the start of germination. In our work this point was considered the 0-h of the germination stage.



Germination begins once the seed has swelled due to a constant supply of water. During the complete germination process soybean seeds were sprayed with water every 20 min for a 10 min duration to ensure adequate moistening. The temperature was kept constantly at 21 °C. The germination process was monitored for 56 h by taking samples at every 8 h (0, 8, 16, 24, 32, 40, 48, and 56 h). The germinated samples are indicated with numbers in Figures. After 48 h' germination, subsets of soybeans were immediately heat treated (105 °C, 10 min), and these samples were also analysed as cooked samples (marked with "F" in Figures).

## 2.2. Spectroscopic measurements

The soybean sample subsets, germinated for different amounts of time, were analysed immediately after sampling without any sample preparation. Three parallel measurements were taken for all samples. The dispersive NIR instrument, NIR Systems 6,500 monochromator system (Foss NIRSystems, Inc., Silver Spring, MD, USA) was applied with a standard sample cup fitted with a Sample Transport Module sampling unit. Samples were scanned from 1,100 nm to 2,498 nm in reflectance mode collecting data every 2 nm with a PbS detector.

## 2.3. Data processing

Spectral and reference data were processed using Vision 3.20 (Minneapolis, MN, USA), Microsoft Excel 2007 (Microsoft Corporation, Redmond, WA, USA), and Statistica 11 (StatSoft, Inc., Tulsa, OK, USA) software packages. The second order derivatives (D2OD) 8/0 nm gap-segment (Norris, 1983; Hopkins, 2001) setting was used for each spectrum, and after it the principal component analysis (PCA) method (Wold et al., 1987; Martens and Næs, 1991), the cluster analysis (CA) method (Heise and Winzen, 2002), and the polar qualification system (PQS) method (Kaffka and Seregély, 2002) were applied with the derivatives spectra. The water peaks were cut off from the spectra (Gergely and Salgó, 2003) for further data processing due to the considerable changes in the amount and condition of water during germination. The continuous water uptake during the biological process resulted in such big changes in the intensity and position of water peaks that the changes of other components could hardly be analysed.

## 2.4. Derivatives

First the second derivative of the near-infrared spectra were created by the Norris method (Hopkins, 2001). The Norris method, also known as the gap-segment method, was carried out with multiple settings. After analysing, the commonly used one data-point gap and five data-point segment setting was chosen. Second derivative spectra have better signal-to-noise ratio, they improve the separation of overlapping peaks and eliminate the baseline shift arising from physical properties. Quantitative measurements can similarly be realised in the derivative spectra, since derivation is a linear mathematical operation (Norris, 1983). Therefore, the cluster analysis, the principal component analysis, and the PQS analysis were carried out on the second derivative spectra.

## 2.5. Cluster analysis (CA)

Cluster analysis is a data mining tool with the purpose of identifying similarities within the data and grouping them (Hopkins, 2001). Based on the characteristics of the objects/data, the similar ones are classified in one cluster. Single linkage, Ward's method, 1-Pearson  $r$ , and Euclidean



distances were the calculations applied as shown in Figs 2 and 3. CA is adaptable to changes and it highlights the particular features that distinguish between the individual clusters.

## 2.6. Principal component analysis (PCA)

PCA is a statistical method used to emphasise the main variation in a multidimensional table. The technique results in a linear dimensionality reduction using an orthogonal transformation. Similarities or differences and interrelationships can be identified among samples by plotting PCA parameters (Norris, 1983; Wold et al., 1987).

## 2.7. Polar qualification system (PQS)

PQS is a data reduction and evaluation method, which is suitable for detecting differences in the quality of the samples. It reduces the spectral data in a “quality point” on a two dimensional “quality plane” based on geometrical considerations. The quality differences of the samples were evaluated and visualised by the PQS surface method (Kaffka and Seregély, 2002) with PQS 32 Evaluation Software, ver. 1.37 (Metrika R&D Co., Budapest, Hungary).

# 3. RESULTS AND DISCUSSION

The germination process eventuates physical, morphological, compositional, and structural changes in the seed, all of which are manifested in a complex way in NIR spectra. The NIR spectra taken from the soybean samples were examined. They were treated and analysed using mathematical methods. Figure 1A presents the raw NIR spectra of the Kanadai (Canadian) variety taken during the whole germination process. Figure 1B presents the second derivative spectra of the Kanadai (Canadian) soybean variety parallel to the diagram of Fig. 1A. The sections marked with the thick black line (1,890–1,920 nm, 1,400–1,420 nm, and 1,150–1,165 nm) are the water peaks, which were cut off thereafter from the spectra. Each analysis was carried out on all three soybean varieties, however, not all of them are presented due to the similar behaviour of soy varieties. The same trends were experienced with cluster analysis, therefore, the figures show only the results observed for the Kanadai (Canadian) variety.

## 3.1. Cluster analysis observations

The cluster analysis of all samples of the Kanadai variety examined is presented in Fig. 2A. The second derivative spectra of the cooked samples (F) were found to be distinct from the second derivative spectra of other samples. Heat treatment resulted in such a large alteration in the biological system of the germinated soybeans that the separation of heat-treated samples and non-heat-treated samples dominated the CA figure. The next significant change could be observed between the dry seeds and the germinated ones. Figure 2B shows the cluster analysis of samples, but omitting the cooked samples. As presented, early water uptake (imbibition) resulted in a far larger change in the seeds than any other alteration observed during germination. This is why the dry seed samples were also left out in order to group the germinated samples more effectively. Figure 2C demonstrates the CA of germinated soybean samples from imbibition to the 56th hour of germination. The results indicate that imbibition (seed swelling) has a distinct effect on the chemical and physical properties of samples as compared to the effect



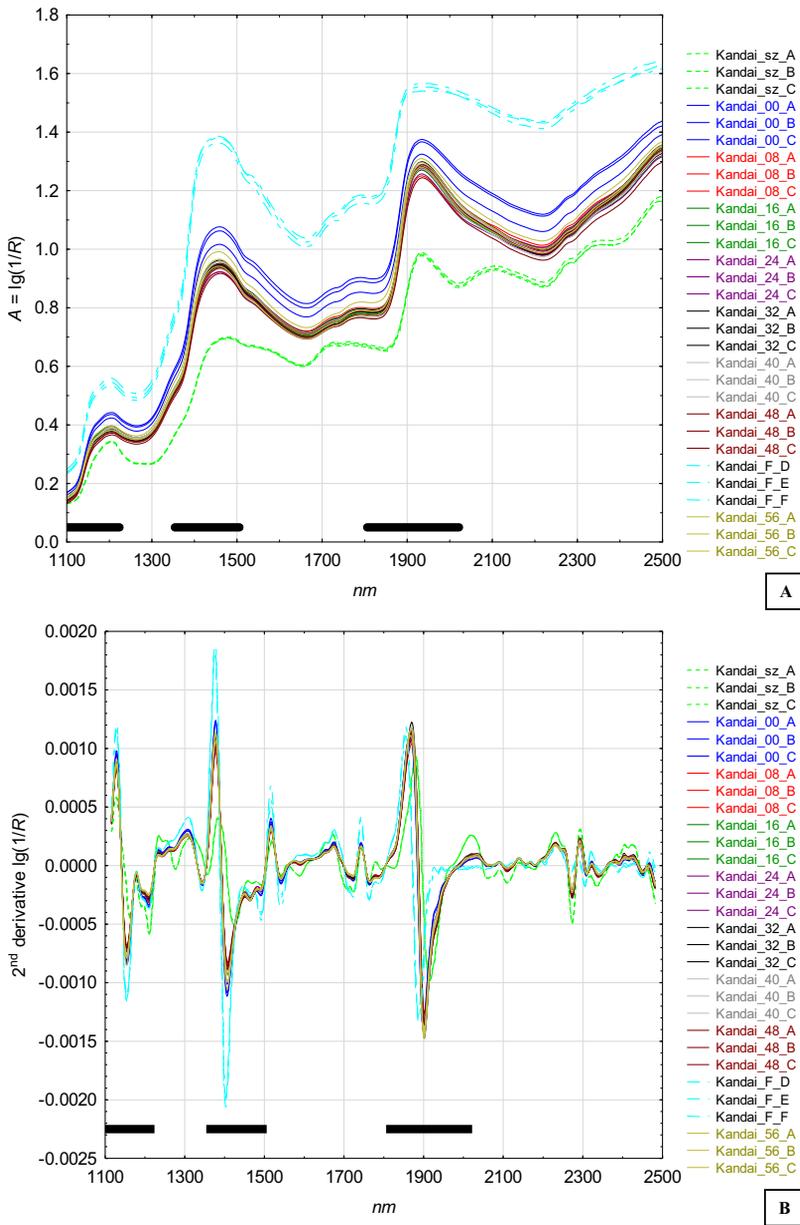


Fig. 1. Spectra of soy seed samples (Kandai) during germination. (A) raw spectra (B) second derivative spectra. The sections marked with a thick black line are the water peaks, which have been cut off thereafter from the spectra. Dry seed samples are marked with “sz”, germinated samples are indicated with numbers, and cooked samples are marked with “F” in the Figures



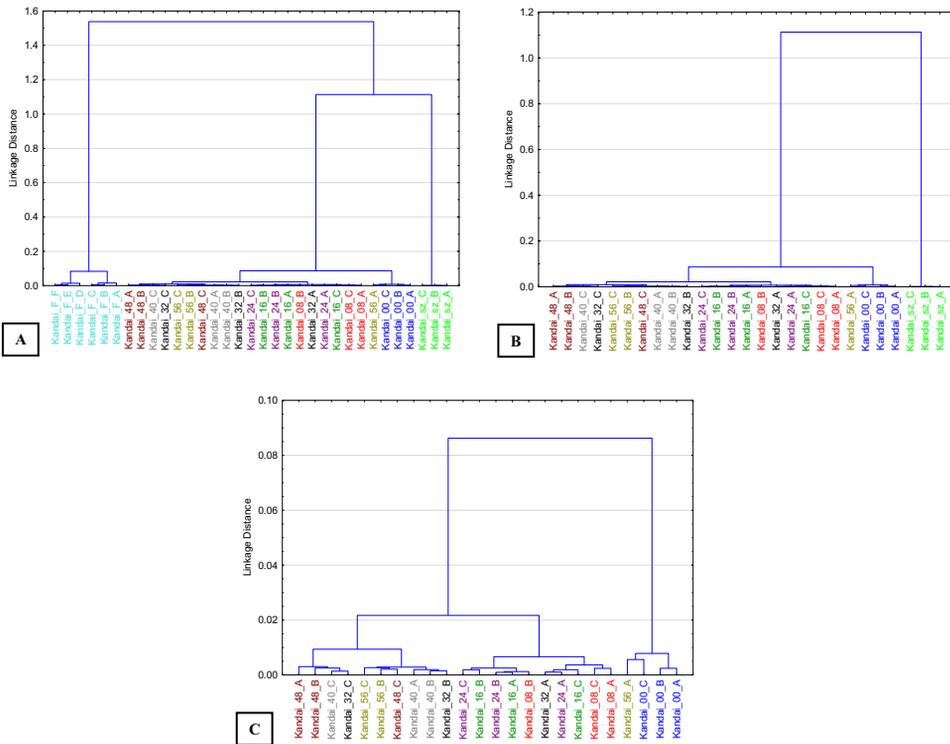


Fig. 2. Dendrogram of cluster analysis based on second derivative spectra of soy seed samples (Kandai) during germination. (A) all examined samples, (B) omitting the cooked samples (F), (C) omitting the dry seed samples (sz), too. Amalgamation rule: Ward's method; distance measures: 1-Pearson *r*

caused by the germination process. Imbibed samples (00) were distinguished very easily from germinated samples indicating the molecular differences between water uptake (imbibition) and activation and storage mobilisation processes during germination. The germinated samples differed from each other to a smaller extent because of the short time window (8 h). Although the grouping algorithm does not perfectly separate the samples according to germination time, some degree of separation can be observed. The 0 h germination stage, the following 3–4 points of the germination process, and the end of the germination process, i.e. the last 3–4 points of the germination process, were all significantly separated.

In the next stage, the spectra of the cooked soybean samples and the spectra of the dry soybean samples were omitted in order to reduce the high spectroscopic (chemical and physical) variation of samples and to focus only on the processes running during the germination itself. The cluster analysis of germinated samples from 0 h germination until the 56th hour of germination, by different mathematical methods are presented in Fig. 3. The mathematical treatments applied are widely used in the literature. In all cases, the start of germination 0 time point is separated confirming our previous observations in Fig. 2C. However, using a single linkage calculation does not provide the above mentioned separation between the first and



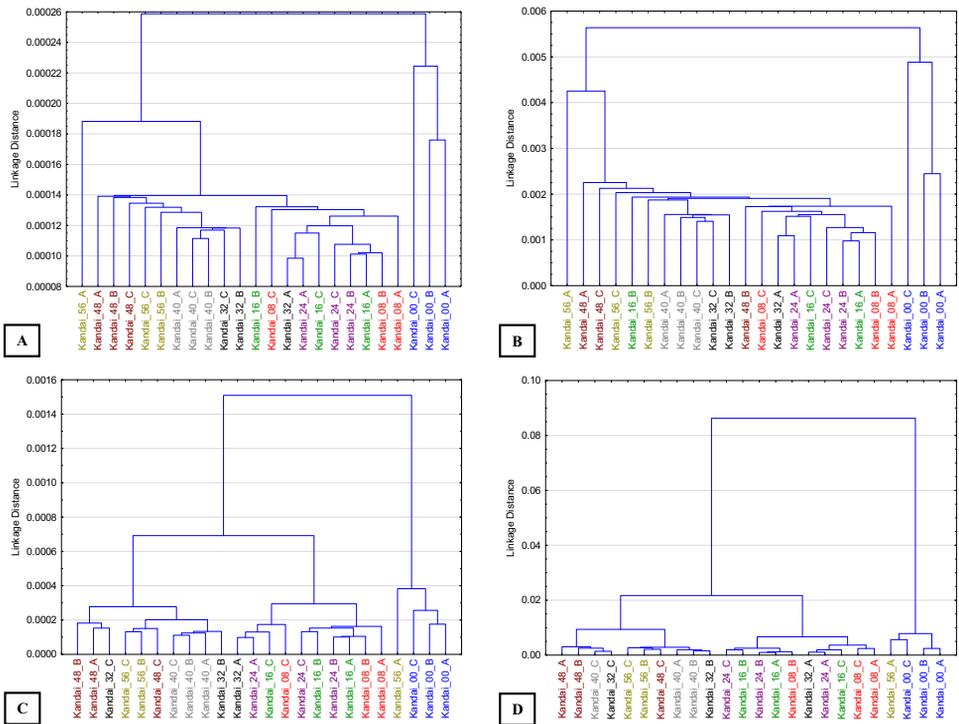


Fig. 3. Dendrogram of cluster analysis based on second derivative spectra of soy seed samples (Kanadai) during germination. Amalgamation rule: single linkage (A, B), Ward's method (C, D); distance measures: Euclidean distances (A, C), 1-Pearson  $r$  (B, D)

second (last) phases of germination. The application of 1-Pearson  $r$  or Euclidean distance calculations do not result in a significant difference in sample separation.

### 3.2. Principle component analysis observations

All three soybean varieties were subjected to principle component analysis. The cooked soybean and the dry seed samples were left out of PCA for the reason mentioned in 2.1. above. [Figures 4A, 5A, and 6A](#) show the principal component analysis of the Bio Pannónia Kincse, Kanadai, and Pannónia Kincse varieties, respectively. In all cases two-dimensional score plots were studied: principal component 1 (PC 1) in the function of principal component 2 (PC 2), principal component 3 (PC 3) in the function of PC 2, and PC 1 in the function of PC 3. However, only figures for PC 2 depending on PC 3 were presented for all three soybean varieties as all other figures showed similar characteristics. The different germination times are indicated by different colours in the figures. The three parallel measurements were surrounded by a double standard deviation ellipse for each sample. The germination process of the Bio Pannonia Kincse soybean variety can easily be followed in [Fig. 4A](#). By imaginary linking the centres of the ellipses according to time processing typical curves can be observed. PCA seems to be an appropriate method for monitoring the physiological processes and for distinguishing the imbibition and



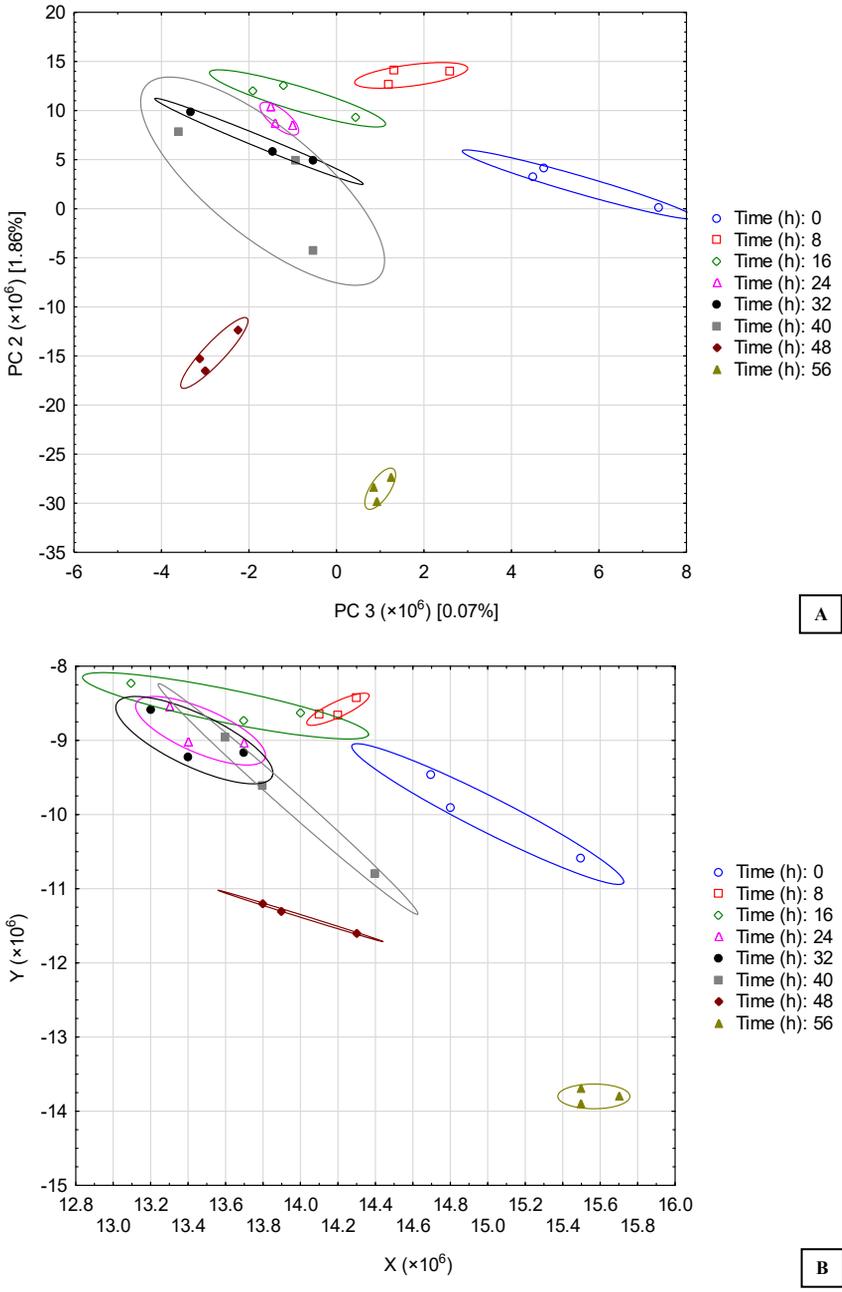


Fig. 4. Score plot of principal component analysis (PC2 vs. PC3) (A) and quality plane of polar quality system (B) based on second derivative spectra of soy seed samples (Bio Pannónia Kincse) during germination



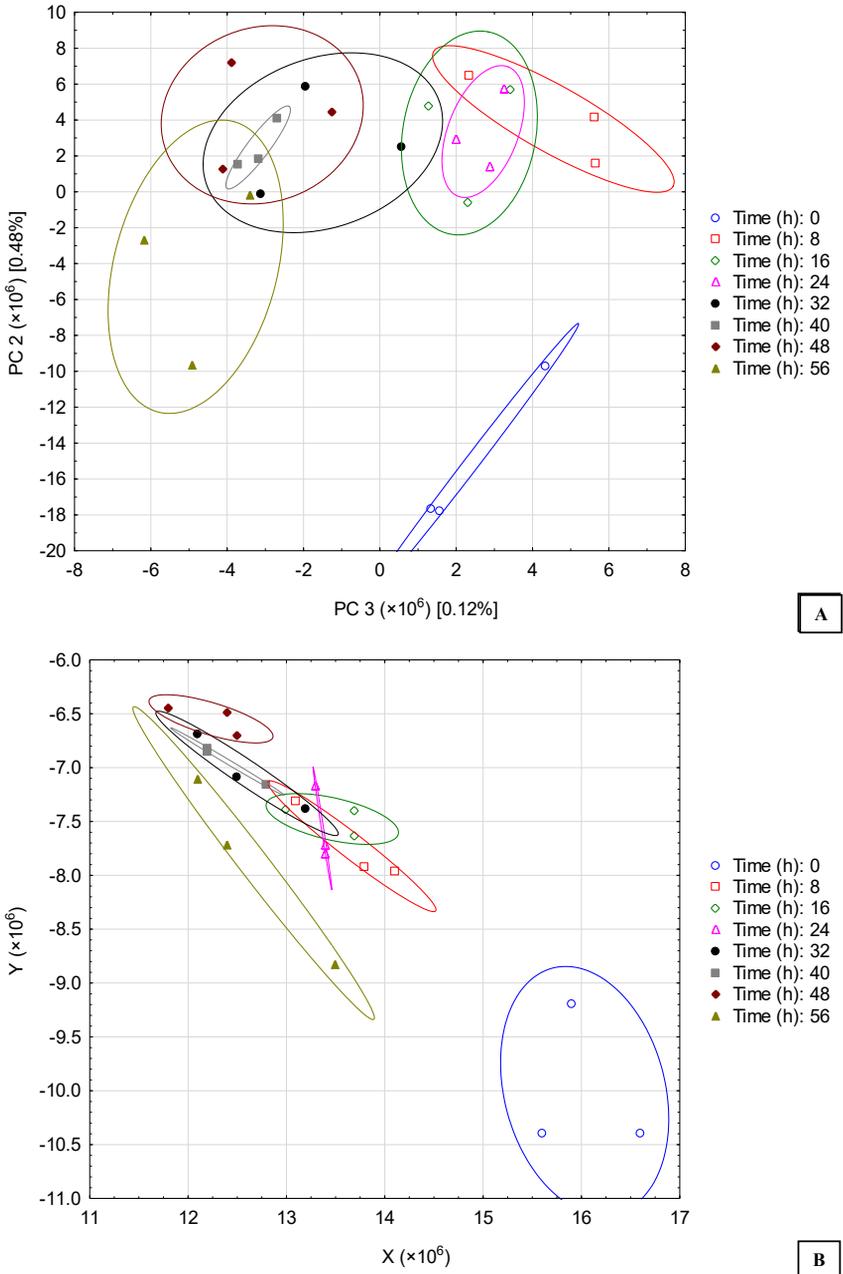


Fig. 5. Score plot of principal component analysis (PC2 vs. PC3) (A) and quality plane of polar quality system (B) based on second derivative spectra of soy seed samples (Kanadai) during germination



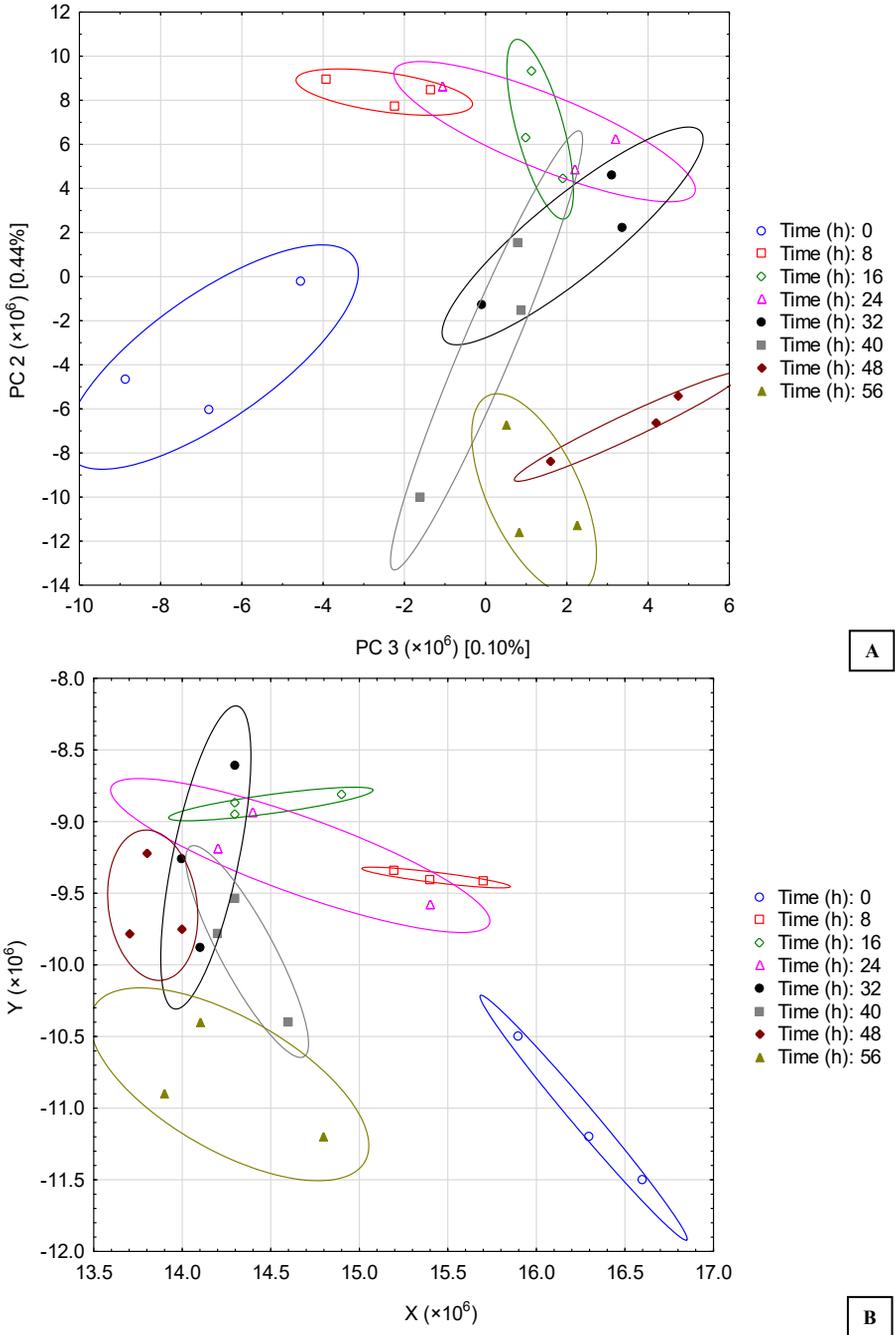


Fig. 6. Score plot of principal component analysis (PC2 vs. PC3) (A) and quality plane of polar quality system (B) based on second derivative spectra of soy seed samples (Pannónia Kincse) during germination



germination phases. The 3rd latent variable (PC 3) shows a strong correlation with germination time. The germination procedure showed characteristic 3D horse-shoe shaped, non-linear trajectory, which confirms the non-linear character of physiological changes during germination. A similar trajectory of germinated samples was observed in PCA illustration of the Kanadai variety in Fig. 5A. Although for Kanadai variety smaller changes could be observed during the germination process studied – except at the beginning of the process –, which is similar to what was observed with the Bio Pannónia Kincse variety. The principal component analysis of the Pannónia Kincse variety is presented in Fig. 6A. In this variety (Pannonia Kincse) the resulting trajectory of the samples showed different characteristics when compared to the previous two varieties. Following the germination time, a nice curve can be traced here, too, although the one showing the 48 h time is hanging out to a small extent.

In general, PCA is a sensitive enough statistical tool to follow the physiological events of germination in the time frame between 0 and 56 h.

In order to describe the germination (mobilisation) processes (based on NIR spectroscopic changes) in another way, PQS, as a novel evaluation method, was tested.

### 3.3. Polar qualification system observations

The polar qualification system analysis of all three soybean varieties are presented in parallel with the PCA results as both of them are multi-dimensional mathematical methods. The dry seed and the cooked samples were left out in this case due to better division during the germination process. The PQS analysis of the Bio Pannónia Kincse variety, the Kanadai variety, and the Pannónia Kincse variety is presented in Figs 4B, 5B, and 6B, respectively. Germination times are indicated by the same colours as in the PCA figures. Double standard deviation ellipses were fitted to the three measured points of each sample using the Statistica 11 program. The PQS analytical method can also be used to track changes as they occur during the germination process. The quality points of the Bio Pannonia Kincse soybean variety are shown in Fig. 4B. Similar curves can be seen by connecting the centres of the ellipses according to time processing, as in principal component analysis. The characteristic horse-shoe shaped, non-linear trajectory observed during the germination procedure is similar to that shown in the PCA figures. The characteristic trajectory is even comparable showing smaller changes during the middle section of the germination process studied and bigger changes at the beginning and at the end of it. Figure 5B demonstrates the results of the PQS analysis of the Kanadai variety. Results are also harmonised with the PCA results. The 0 h point of germination is statistically different from other germination points of time, however, the extent of the differences among germinated samples are much lower than for the PCA values (Fig. 5A). Figure 6B demonstrates the quality points of the Pannónia Kincse variety.

The trajectory of the Pannónia Kincse samples behaves differently than the other two varieties, as seen in the PCA data. A small fluctuation can be observed from the germination time of 16 h to the germination time of 48 h.

The assignments of PC 1, 2, and 3 were evaluated in Table 1, and it was observed that PC 1 represents the main changes of protein and polysaccharide components while PC 2 represents the protein-polysaccharides and oil components. PC 3 identifies the remaining variations, which represent mainly polysaccharides and oil. These results assume that the mobilisation of storage components follows a protein/starch/oil relative order. All in all, both the PQS and PCA are





**Table 1.** Assignment of the first five local outside values of the first three loadings derived from PCA of second derivative spectra of soy seed samples during germination

Variety	Loading of PC 1					Loading of PC 2					Loading of PC 3				
	Wavelength of local outside values (nm)	Sign of local outside values	Assignment			Wavelength of local outside values (nm)	Sign of local outside values	Assignment			Wavelength of local outside values (nm)	Sign of local outside values	Assignment		
			(Osborne and Fearn, 1986)	(Shenk et al., 1992)	(Williams, 2001)			(Osborne and Fearn, 1986)	(Shenk et al., 1992)	(Williams, 2001)			(Osborne and Fearn 1986)	(Shenk et al., 1992)	(Williams, 2001)
Bio Pannónia Kinese	1516	-	CONH2/ROH	urea	protein	2274	-	starch	cellulose	protein	1780	+	cellulose	cellulose	starch
	1306	-	CH3	CH3	cellulose/starch	2294	+	amino acid	protein	starch/oil	1233	+	CH	CH	cellulose
	2272	+	starch	cellulose	protein/cellulose	1516	-	CONH2/ROH	urea	protein	2308	+	CH2	oil	starch
	1344	+	CH3	CH3	starch	2036	+	CONH2	urea	protein /cellulose /starch	2468	+	starch	CH2/protein	protein
	1766	+	CH2	CH2	oil	2404	+	ROH	oil	oil	2295	-	amino acid	protein	starch/oil
Kanadai	1516	-	CONH2/ROH	urea	protein	2274	-	starch	cellulose	protein	1234	+	CH	CH	cellulose
	2274	+	starch	cellulose	protein	2294	+	amino acid	protein	starch/oil	1777	+	cellulose	cellulose	starch/water/oil
	1306	-	CH3	cellulose	cellulose/starch	2477	-	starch	protein	cellulose/starch	1726	-	CH2	CH2	oil
	1344	+	CH3	CH3	starch	1517	-	CONH2/ROH	urea	protein	1708	-	CH3	CH3	cellulose
	2292	-	amino acid	protein	starch	2430	+	starch	CH2/protein	starch	1340	+	CH3	CH	cellulose/starch
Pannónia Kinese	1516	-	CONH2/ROH	urea	protein	2274	-	starch	cellulose	protein	1725	+	CH2	CH2	oil
	1306	-	CH3	cellulose	cellulose/starch	2293	+	amino acid	protein	starch	1233	-	CH	CH	cellulose
	2274	+	starch	cellulose	protein	1778	+	cellulose	cellulose	starch/water/oil	1706	+	CH3	CH3	cellulose
	1344	+	CH3	CH3	starch	2476	-	starch	CH2/protein	cellulose/starch	2294	+	amino acid	protein	starch /oil
	1764	+	CH2	CH2	oil	2424	+	starch	protein	protein/starch	2276	-	starch	starch	protein /cellulose/starch

appropriate statistical methods for studying this natural biological process, but the “sensitivity” of PCA, especially in relation to PC2 and PC3, is higher. According to any assignment process (see Table 1) all three PCAs describe variances in relation to all main constituents (protein, carbohydrate, and oil) to varying degrees, confirming the extreme molecular complexity of germination.

## 4. CONCLUSIONS

Near-infrared spectroscopy can be used to sensitively track technological steps aimed at soybean germination and other technological operations. The main or rough changes in the seed during these processes can be monitored in real-time based on the variations in moisture content (water peaks). The fine details and time-dependent effects of germination can be clearly observed in those NIR spectra where water absorption bands were excluded. The statistical analyses of “waterless” spectra confirmed that the calculations using direct spectrum variables (cluster analysis, polar qualification system) produced less sensitive models than calculations using latent variables (principal component analysis). The trajectories of PCA investigations confirmed the non-linear character of the germination process. The relative order of the mobilisation of seed reserves was represented by the loading vectors of several latent variables (principal components). Overall, it can be stated that NIR spectroscopic measurements combined with sophisticated chemometric methods are capable of detecting and monitoring complex physiological processes such as germination in real-time.

## ACKNOWLEDGEMENTS

The research that led to these results has received funding from the European Union Seventh Framework Programme FP7/2007–2013 under grant agreement n° 266331. “CHANCE - Low cost technologies and traditional ingredients for the production of affordable, nutritionally correct foods improving health in population groups at risk of poverty”.

This work was supported by the Higher Education Excellence Program of the Ministry of Human Capacities in the framework of the Biotechnology research area of Budapest University of Technology and Economics (BME FIKP-BIO).

The research reported in this paper is part of project no. TKP2021-EGA-02, implemented with the support of the Ministry for Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021 funding scheme.

## REFERENCES

- Abe, H. (2004). Estimation of heat capacity and properties of water by spectrum decomposition of the second overtone band of OH stretching vibration. *Journal of Near Infrared Spectroscopy*, 12: 45–54.
- Allosio, N. and Boivin, P. (1997). Characterisation of barley transformation into malt by three-way factor analysis of near infrared spectra. *Journal of Near Infrared Spectroscopy*, 5: 157–166.



- Bartalné-Berceli, M., Izsó, E., Gergely, Sz., Jednákovits, A., Szilbereky, J., and Salgó, A. (2016). Sprouting of soybean: a natural process to produce unique quality food products and additives. *Quality Assurance and Safety of Crops and Foods*, 8(4): 519–538.
- Bartalné-Berceli, M., Izsó, E., Gergely, Sz., and Salgó, A. (2018). Development and application of novel additives in bread-making. *Czech Journal of Food Sciences*, 36(6): 470–475.
- Erdman, J.W. Jr. and Fordyce, E.J. (1989). Soy products and the human diet. *The American Journal of Clinical Nutrition*, 49(5): 725–737.
- Fitorex Engineering and Trading Ltd. (2008). *Új, növényi eredetű élelmiszer-ipari termék és az azt tartalmazó készítmények (New food-industrial product with plant origin and goods containing it)*. Hungarian Patent. patent number: P 08 00665 in Hungary.
- Gergely, Sz. and Salgó, A. (2003). Changes of moisture content during wheat maturation – what is measured by near infrared spectroscopy? *Journal of Near Infrared Spectroscopy*, 11: 17–26.
- Gergely, Sz. and Salgó, A. (2007). Changes in protein content during wheat maturation – what is measured by near infrared spectroscopy? *Journal of Near Infrared Spectroscopy*, 15: 49–58.
- Graeber, K., Nakabayashi, K., Miatton, E., Leubner-Metzger, G., and Soppe, J.J.W. (2012). Molecular mechanisms of seed dormancy. *Plant, Cell and Environment*, 35: 1769–1786.
- Heise, H.M., and Winzen, R. (2002). Fundamental chemometrics methods. In: Siesler, H.W., Ozaki, Y., Kawata, S., and Heise, H.M. (Eds.), *Near-infrared spectroscopy – Principles, instruments, applications*. Wiley-VCH Verlag GmbH, Weinheim, pp. 125–162.
- Hopkins, D.W. (2001). What is a Norris derivative? *NIR News*, 12(1): 3.
- Juhász, R., Gergely, Sz., Szabóki, Á., and Salgó, A. (2007). Correlation between NIR spectra and RVA parameters during germination of maize. *Cereal Chemistry*, 84(1): 97–101.
- Kaffka, K.J., and Seregély, Zs. (2002). PQS (Polar Qualification System) the new data reduction and product qualification method. *Acta Alimentaria*, 31(1): 3–20.
- Martens, H. and Næs, T. (1991). *Multivariate calibration*. John Wiley & Sons Ltd., Chichester. 440 pages.
- Norris, K.H. (1983). Extracting information from spectrophotometric curves. Predicting chemical composition from visible and near-infrared spectra. In: Martens, H., and Russwurm, H. Jr, (Eds.), *Food research and data analysis*. Applied Science Publishers Ltd., London, pp. 95–113.
- Osborne, B.G., and Fearn, T. (1986). Theory of near infrared spectrophotometry. In: *Near infrared spectroscopy in food analysis*. Longman Scientific & Technical, Harlow, pp. 20–42.
- Shenk, J.S., Workman, J.J. Jr, and Westerhaus, M.O. (1992). Application of NIR spectroscopy to agricultural products. In: Burns, D.A., and Ciurczak, E.W. (Eds.), *Handbook of near-infrared analysis*. Marcel Dekker, Inc., New York, pp. 383–431.
- Smail, V.W., Fritz, A.K., and Wetzel, D.L. (2006). Chemical imaging of intact seeds with NIR focal plane array assist plant breeding. *Vibrational Spectroscopy*, 42: 215–221.
- Williams, P.C. (2001). Implementation of near-infrared technology. In: Williams, P., and Norris, K (Eds.), *Near-infrared technology is the agricultural and food industries spectroscopy in food analysis*, 2nd ed. American Association of Cereal Chemists, Inc., St. Paul, pp. 145–169.
- Wold, S., Esbensen, K., and Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1–3): 37–52.

---

**Open Access.** This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited, a link to the CC License is provided, and changes – if any – are indicated. (SID\_1)

