USING NEAR INFRARED REFLECTANCE SPECTROSCOPY TO DETERMINE GROSS ENERGY AND TOTAL AMINO ACID CONTENT OF HIGH PROTEIN SOY PRODUCTS

by

DYLAN PATRICK MOSS

(Under the Direction of Adam Davis)

ABSTRACT

The amino acid (AA) and gross energy (GE) content of high protein soybean feed ingredients can vary based on cultivar genetics, environmental growing conditions and processing procedures used in making soybean meals and soy protein isolates and concentrates. Using high performance liquid chromatography (HPLC) and bomb calorimetry to determine individual AA and GE content respectively, is costly and takes days to complete. In the current research near-infrared reflectance spectroscopy (NIRS) was used to build calibration curves to determine the GE and individual total AA content of high protein soybean products in real-time. When these curves were validated with about 100 new samples, the NIRS predicted values for GE and each AA all deviated less than 5% from determined values except for cysteine and tryptophan. The results indicate that NIRS could be used to quickly ascertain GE and AA content of high protein soy products to assist in animal diet formulation.

INDEX WORDS: Solvent extracted soybean meal, Mechanically extracted soybean meal, Alternative feed ingredients

USING NEAR INFRARED REFLECTANCE SPECTROSCOPY TO DETERMINE GROSS ENERGY AND TOTAL AMINO ACID CONTENT OF HIGH PROTEIN SOY PRODUCTS

By

DYLAN PATRICK MOSS

BSA, The University of Georgia, 2025

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2025

USING NEAR INFRARED REFLECTANCE SPECTROSCOPY TO DETERMINE GROSS ENERGY AND TOTAL AMINO ACID CONTENT OF HIGH PROTEIN SOY PRODUCTS

By

Dylan Patrick Moss

Major Professor:	Adam J. Davis
Committee:	Andrew P. Benson
	Brian D. Fairchild

Electronic Version Approved:

Ron Walcott

Vice Provost for Graduate Education and Dean of the Graduate School

The University of Georgia

May 2025

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The soy plant is a legume that originates in Eastern Asia but is now grown and utilized throughout much of the world to meet the demand for soy products in in human and food animal nutrition. Soy plants can be grown in a wide variety of photoperiods, temperatures, and soil moisture conditions because of intense genetic selection for the production of cultivars that will thrive under varying environmental conditions (Mayhew et al., 2014). A review of soy's global impact by Hamza et al., (2024) found that 6% of the worlds arable land is dedicated to the cultivation of soybeans. They also reported that the largest global producers of soybeans were the United States, China, Brazil, and Argentina. In the United States, the states with the greatest production in 2022 were Illinois, Iowa, and Minnesota accounting for more than 38% of total production (USDA, 2025)

Soybeans

While the initial production of soy for agriculture focused on its use as a forage crop, subsequently the harvest and utilization of the soybean seeds (soybeans) became the focus of production. This is because soybeans contain about 37% crude protein and 18% crude fat and the amino acids of soy protein are highly digestible and have a complimentary composition to cereal grains (Batal and Dale, 2016; NRC, 1994). Soybeans are high in lysine and low in methionine, while cereal grains are low in lysine and have higher methionine content. In addition, soybeans also have greater amino acid digestibility than other oilseeds such as sunflowers, canola, and cottonseed (Sotak-Peper et al, 2017). Another benefit of soybeans is

they have a high oil content, and this oil is highly digestible (NRC 1994) and thus provides a substantial amount of energy for the diet (Berrocoso et al., 2015).

While soybeans are appealing for their use in diet formulation, several antinutritional factors inhibit their use in raw form in monogastric diets. Soybeans contain protease inhibitors such as trypsin inhibitors that prevent the enzymatic activity of protease enzymes. In both poultry and swine, the presence of the protease inhibitors in raw soybeans significantly reduces the digestibility of and utilization of amino acids (Liener, 1996; Palliyeguru et al., 2011). Another antinutritional factor present in raw soybeans is lectins also known as agglutinins. Soybean agglutinin is a glycoprotein that binds to enterocytes in the small intestine causing inflammation that damages their structure and function, that inhibits the intestinal membrane immune system and that can alter microbiota population dynamics as reviewed by Pan et al., (2018). The changes in enterocyte villi physiology cause a decrease in nutrient absorption (Casaubon-Huguenin et al., 2004).

Both trypsin inhibitors and soybean agglutin can be denatured with heat treatment of raw soybeans (Liener 1981). The effective heating protocols vary based on moisture, particle size of the soybeans, and heating method used (Melcion and van der Poel, 1993). The goal is to inactivate the anti-nutritional factors while maintaining bioavailability of other important nutrients such as carbohydrates, amino acids and vitamins. Under-processing soybeans will result in inadequate protease and lectin activity reduction (Chang, 1987). Overprocessing soybeans will cause reduced amino acid digestibility, especially for lysine which can undergo Maillard reactions (Parsons et al, 1992). An assessment of urease activity following processing is used to test if conditions were sufficient to inactivate antinutritional factors. The enzyme, urease is naturally present in soybeans and while its presence does not affect monogastric

nutrition, the activity of this enzyme is easily assayed (Croston et al., 1955) and because urease is heat liable if there is a sufficient decrease in its activity, this will indicate that the trypsin inhibitors and lectins will have been denatured as well. On the other hand, if urease activity is still abundant this will indicate the soybeans are under-processed and including them in monogastric diets would decrease animal feed efficiency. In addition to the urease test to determine if soybeans have been under processed, the solubility of soybean protein in potassium hydroxide (KOH) is determined. As the solubility of soybean protein in KOH is inversely related to the degree of heat treatment. The protein solubility of raw soybean flour is nearly 100% while soybean meals that have been overheated and are dark brown in color will have solubility values between 30 and 40% and this indicates that protein quality was not maintained due to overprocessing. Ideal solubility values for soymeal would be 78 to 84% (Caprita and Caprita, 2010).

Phytic acid is the storage form of phosphorus in plants especially the seeds of plants, and it is another antinutritional factor found in soybeans that is not heat-labile. Up to 80% of the total phosphorus found in raw soybeans is bound in phytic acid and monogastric animals such as poultry lack any meaningful phytase enzyme activity to metabolize the bonds linking the six dihydrogenphosphate ester of inositol molecules that comprise a phytic acid molecule. Because of the lack of endogenous phytase activity in the intestinal tract and because of the pH present in the intestinal tract, phytic acid exists as the phytate anion. Thus, phytate complexes with divalent mineral cations and positively charged amino acids decreasing their ability to be absorbed by the intestinal enterocytes which reduces animal growth and wellness (Caldwell, 1992; Maddaiah et al., 1964). Commercial preparations of heat stable phytase are commonly

added to poultry diets to digest phytic acid which improves protein, carbohydrate, and mineral nutrient availability (Tang et al., 2012).

As reviewed by Karr-Lilienthal et al., (2005) carbohydrates make up about 35% of the soybean seed with half of this amount consisting of structural polysaccharides while the other half consists of low molecular weight sugars, oligosaccharides and small amount of starch. While monogastric animals can digest starch and low molecular weight free sugars, they cannot digest the structural carbohydrates (fiber) to any great extent. The oligosaccharides primarily consist of raffinose, stachyose and verbascose and monogastric animals lack the enzymes necessary for meaningful digestion of the oligosaccharide component of soybeans. However, they are fermented by bacteria populations residing in the gastrointestinal tract of monogastric animals which can lead to gas production. The fermentation of the oligosaccharides is associated with a decrease in metabolizable energy values in poultry (Parsons et al., 2000).

Soybean products

Soy hulls

Soybean hulls represent the seed coat of soybeans and make up about 8% of the seed composition and contain about 86% complex structural carbohydrates and thus are a source of dietary fiber (Gnanasambandam and Proctor, 1999). Soy hulls are removed during the initial steps of production for most high protein soy products. They contain around 38% crude fiber, and 12% crude protein on a dry matter basis (NRC, 1994). Unpublished research from our laboratory indicates that for poultry on as is basis the nitrogen corrected true metabolizable energy is only 735 kcal/kg even though gross energy content is 3,850 kcal/kg. Additionally, the average amino acid digestibility coefficient is only 41%. Thus, inclusion of soy hulls in broiler diets is not feasible, but they are sometimes included in broiler breeder diets as a relatively

indigestible filler to increase the dietary volume that can be fed to the birds in feed restriction feeding programs (Aranibar, 2018). Soy hulls are useful as an energy source in ruminant nutrition, in which rumen fermentation yields an energy content comparable to corn (Blasi et al., 2000).

Full fat soybean meal

Full fat soybean meal is typically produced by dry extrusion of washed cracked whole beans that still have their hulls. The extrusion process will generate enough heat to inactivate the antinutritional factors in soybean. The resulting full fat soybean meal has a crude protein of 38-40%, a crude fat of 18-20%, and a crude fiber of 5% on as is basis (Batal and Dale 2016; Toomer et al., 2024). As reviewed by Tommer et al., (2024), the use of full fat soybean meal is increasing in all natural and sustainable poultry production systems. As will be discussed in an upcoming section, the oil is typically removed from soybeans and sold as its own product, leaving behind a meal that is typically used in animal diets. The oil is removed most commonly by a solvent extraction process which makes the resulting meal contaminated with trace amounts of hexane, making it unsuitable for all natural poultry production. In addition, when oil extracted soybean meal is utilized in poultry diets, animal fats or seed oils must be added to the diet to meet the energy requirement of the birds being fed. From a sustainable production standpoint, the energy associated with removing the oil from soybeans to make a meal does not make sense when additional oil then must be added back to the diet. If the oil is not removed and full fat soybean meal is utilized additional oil/fat sources do not have to be added back to the diet, thus making the whole process more carbon neutral.

Soy oil

Soybeans have typically been grown for oil production as oil on a per weight basis commands the greatest price of any soy products. However, the soybean meal that is left after extracting the oil is produced in a greater volume than oil as the oil extraction process yields about 20% oil and 80% soybean meal. The demand for soybean meals continues to increase for animal diets and with this increased demand, prices have increased and while they remain well below the oil on a per unit weight basis, on a total volume produced the value of the soybean meal can now be equal or greater than the oil. Animal feed utilizes 76% of global soybean production while direct human food consumption utilizes 20% of global production, and industry (biodiesel, lubricant, paint, etc. production) utilizes the remaining 4% (Abafe, 2023). In the human food consumption category soybean oil accounts for over 65% of what is consumed and worldwide most of the remaining oil not consumed by humans is used in commercial animal diets. In the United States, 61% percent of soybean oil is used for food, 31% is used for biodiesel, and 8% is used for industrial purposes such as lubricant, paint and cosmetic production (US Soy, 2018) and this reflects a much higher biodiesel production utilization of the oil than the rest of the world.

Crude soybean oil naturally consists of 96% triglycerides, 2% phospholipids, 1.5% unsaponifiables, and 0.5% free fatty acids and after processing for human consumption the triglyceride concentration is over 99%. Linoleic, oleic and linolenic fatty acids account for over 80% of the total fatty acids which makes the oil a highly unsaturated fatty acid oil source and susceptible to oxidation. However, this fatty acid profile of soybean oil makes it highly digestible in poultry relative to other commonly available dietary fat sources (Leeson and Summers, 2001).

Soybean meal

Extraction of the oil from soybeans is accomplished commercially by either a solvent or mechanical process and the resulting soybean meal from these two methods vary in nutrient composition. Solvent extraction is the most common means for separating soybean oil and thus solvent extracted soybean meal is the most widely available. Solvent extraction processing begins with soybeans being cleaned, cracked, and their hulls removed. Soybeans are then tempered by heating them in a humid environment. This enables them to be mechanically processed into a thin and wide shape called flakes which increases the surface area exposed to the solvent, enabling more efficient extraction of the oil. The dried flakes are then bathed in hexane which is the solvent of choice for commercial extraction as it extracts the most oil and boils at 69° C making its evaporation from the resulting oil and soybean meal cost effective. The solvent and extracted oil are then drained leaving the resulting flakes for desolventizing. Desolventizing can be an integrated process in which the flakes are also heated, toasted, and cooled, a process that both removes the solvent and denatures the trypsin and lectin antinutritional factors. The result is a soybean meal product ready for livestock diet formulation and utilization. Alternatively, flash desolventizing is utilized which uses instantaneous high heat from a stream of superheated vapors to remove the solvent from the soybean meal. The resulting product is typically used to produce soybean concentrates and isolates to be discussed subsequently.

Mechanical expulsion is an alternative approach that uses a mechanical press instead of solvents to remove the oil from the soybean. The process of mechanical extraction begins with the soybeans being cleaned, dehulled, cracked and dry extruded. In dry extrusion the soybeans are forced under high pressure through a die which generates heat from friction which is enough

to largely reduce the antinutritional factors found in soybeans. The extruded product is then mechanically pressed with a screw press to expel the crude oil. Mechanically extracted soybean meal like full fat soybean meal, is used in organic and all natural animal production systems.

Solvent extracted dehulled soybean meal has a crude protein content of 47.8%, a crude fiber content of 3% and a crude fat content of 1% (Batal and Dale 2016). After the solvent extraction is complete, the soy hulls are sometimes added back to the soybean meal to create a product that has a crude protein content of 44%, a crude fiber content of 7% and a crude fat content of 0.5% (Batal and Dale 2016). Finally, mechanical extraction is not as effective as solvent extraction, so the resulting meal retains more oil and thus the crude protein content is 42%, the crude fiber content is 6.5% and the crude fat content is 3.5%. However, the efficiency of the oil extraction with mechanical extraction varies widely between processors and it is not unusual for samples to have a crude fat level in the 6-8% range (unpublished data from our laboratory). In 2020 livestock consumed 97% of soybean meal produced within the United States with poultry consuming 20 million of the 33 million tons consumed by livestock (United Soybean Board, 2021).

Fermented soybean meal

The use of fermented soybean meal and/or enzymatically treated soybean meal has increased in poultry and swine pre-starter diets with Hamlet Protein A/S (Hosens, Denmark) making products available worldwide. As reviewed by Lambo et al., (2024), making fermented soybean meal involves soaking soybean meal in sterile water so that the moisture content is 40 to 50%. The mixture is then cooked at 45-121° C for 30 to 60 minutes and allowed to cool before being inoculated with bacteria and/or fungi and allowed to ferment at 30° C for 1 to 7 days. After the fermentation period the product is dried and stored for subsequent utilization in diets.

The resulting product is going to have a higher protein content than solvent extracted soybean meal because the bacteria and/or fungi have digested a portion of the carbohydrates especially the oligosaccharides and produced carbon dioxide and water. Thus, with less carbohydrate mass left the percentage of protein will increase. Furthermore, the microbial population has increased during fermentation so microbial protein content has increased and is part of the final product. In addition to a decrease in the oligosaccharide concentration which thereby eliminates or reduces their potential antinutritional effects, the microorganisms will also degrade phytic acid and further inactivate any remaining antinutritional factors including protease inhibitors and lectins by digesting these proteins. Now with the availability of commercial enzymes such as phytase, proteases and carbohydrase derived from microbes and fungi, these enzymes can be added directly to the moistened soybean meal to produce enzymatically treated soybean meal that does not contain bacteria and/or yeast.

The fermented and/or enzymatically treated soybean meal product contains more digestible nutrients and less antinutritional factors than soybean meal. Because of the expense of producing these products, the products are more expensive than soybean meal. The increased price prevents replacement of soybean meal in diets for poultry and swine with fermented and/or enzymatically treated soybean meal, but many producers will replace part of the soybean meal in pre-starter diets with these products as newly weaned pigs and chicks during their first week of age seem to be most sensitive to the antinutritional factors associated with soybean meal as they require a large percentage of their diet to be protein, but have reduced enzymatic activity. (Lambo et al., 2024).

Soy protein concentrate

Flash desolventized solvent extracted soybean meal is the most common source of soy protein used in further processing, but extruded-expelled soybean meal can also be used. Flash desolventized soybean meal is preferred due to the higher quality of protein compared to soybean meal desolventized with toasting or expelled soybeans. It also has lower amounts of crude oil compared to extruded mechanically expelled soybean meal. Soy protein concentrates are created by either acid leaching, aqueous ethanol extracting, or moist heat-water leaching which cause the proteins to become insoluble, while some of the carbohydrates are solubilized and removed by centrifugation with the remaining solids being dried. The resulting dried product has a significantly increased protein concentration. Soy concentrates are not typically used in poultry diet formulation due to the much-increased cost of the ingredient necessitated by the additional processing steps. Soybean concentrates have a minimum crude protein content of 65% on an as is basis and typically exceed 70% on a dry matter basis. Crude fiber content cannot exceed 6% for soy protein concentrates.

Soy protein isolate

Like concentrates, and for the same reasons, the preferred starting material to make soy isolates is flash desolventized solvent extracted soybean meal. The soybean meal is first placed in a warm, slightly alkaline solution in which the protein dissociates. The solution is then centrifuged to separate the undissolved portion. The supernatant is then placed into an acidic solution and refrigerated. The precipitate formed during refrigeration is then separated by centrifugation, and the resulting isolate solution is neutralized and then dried. This process removes insoluble protein and almost all crude fiber, and thus soy protein isolates have a crude protein content of 90% or greater and the crude fiber content that cannot exceed 0.5%. The

process also removes the components that provide flavor, leaving a product that is very popular in human nutrition as it mixes well with liquids and is easily flavored, unlike soy concentrate which does not mix well with liquids and still has a distinct flavor that is hard to mask. Because of their cost, soy protein concentrate, and soy protein isolate are only utilized in the production of semi-purified or purified animal diets for research.

Nutrient variability of soybean meal

In addition to the nutrient concentration differences that can occur in soybean meal based on differences in the processing procedures to produce it, the nutrient content of soybean meal will also vary based on the genetic cultivars of soybean meal used to produce the meal and the agronomic conditions under which the soybeans were produced. As reviewed by Rahman et al., (2023) through natural genetic selection and genetic modification there are now thousands of genetic cultivars of soybean that show differences in tolerance to things such as disease, drought, insects and humidity, resistance to herbicides such as the Roundup Ready varieties, or enhanced nutritional characteristics such as a reduction in anti-nutritional components and improved protein, methionine, carotenoid, or oleic acid content. As would be expected the content of individual nutrients or groups of nutrients in these different cultivars can vary significantly (Bohn et al., 2014; Guo et al., 2020; Agyenim-Boateng et al., 2023; Montanha et al., 2024)

The composition of soybeans varies due to many environment factors in which they are grown. Soil fertilization is one such factor that alters the nutritional composition of soybeans. Sulphur supplementation, especially in nutrient deficient soil, has been shown to increase protein, Cu, Zn, palmitic acid and linolenic acid content of soybeans seeds (Kahraman, 2017). Likewise phosphorous supplementation has been shown to increase protein, palmitic acid, and linolenic acid in soybean seeds (Kahraman, 2017). Other factors such as herbivory by insects

may alter the composition due to stress or damage. Insect salivary enzymes can degrade seed contents and alter the protein content (Depieri and Panizzi, 2011).

The relationship between environmental factors is not always simple or linear. For example, Gibson and Mullen (1996) found protein concentration of soybeans decreased as temperature increased from 21-27° C, and then increase from 28–35° C. On the other hand, oil concentration, which is inversely related to protein concentration (Hymowitz et al., 1972), increased from 21-29° C and then decreased with subsequent increases in environmental temperature.

The influence of how a given environmental factor may influence the composition of the soybean also depends on other factors such as the developmental stage of the plant. For example, Dornbos and Mullen (1992) found that during the seed filling stage, a decrease in soil moisture would increase protein concentration and decrease oil concentration. However, a different study conducted by Wijewardana et al., (2019) found that decreased soil moisture during the reproduction stage had the inverse effect, decreasing the protein concentration and increasing the oil concentration of the soybean. To add further complexity to an already convoluted topic, how much any given factor impacts another throughout the soybean's development can also change due to the genetic cultivar (Weiss, 1952).

Nutrient analysis

Given the potential variation in the nutrient composition of soybean meals, it is necessary for poultry nutritionists to know the composition of the soybean meal they are utilizing to formulate their diets. In poultry diets, soybean meal is typically the major source of amino acids and a significant contributor to the energy requirement especially if mechanically extracted or full fat soybean meal is being utilized. Gross energy is calculated by the usage of a bomb

calorimeter to combust a feed sample. This methodology has the clear drawback of being destructive to the sample, rendering it incapable of further testing. Furthermore, it gives no indication of the quality of the feed, only how much energy is generated when combusted. To determine the concentration of total amino acids, high performance liquid chromatography is used. The sample first undergoes acid hydrolysis to the free amino acids which are then separated based on their function groups on a liquid chromatography column. Like bomb calorimetry, high performance liquid chromatography has the disadvantageous property of being destructive of the sample and creating waste. Both analytical procedures require trained personnel to complete the analyses and utilize expensive specialized equipment, making the cost per sample to be 200 dollars or more. Furthermore, obtaining the results will take days at a minimum.

Near-infrared reflectance spectroscopy

Near Infrared Reflectance Spectroscopy (NIRS) is a rapid analysis method that enables a Multi-Purpose Analyzer to be calibrated to predict nutritional component values after analyzing the reflectance of a feed ingredient in the near infrared spectrum. NIRS has several advantages compared to conventional feed analysis methods. One spectrograph generated by NIRS for an ingredient sample can be used to determine multiple nutritional values, unlike conventional laboratory analyses which require a different test for each component of interest. NIRS is non-destructive of the sample, enabling it to be fed or saved for other laboratory purposes. NIRS also produces no chemical waste and requires no preparation of the sample except grinding. NIRS analysis can also be done at a feed mill and provide real-time results. The downsides of NIRS are the requirement for the initial construction and validation of the calibration curve to predict the parameter of interest and the initial expense of the machine. While the cost of the NIRS

system is similar to the cost of a high performance liquid chromatography system, there are not the repetitive costs for the chemical reagents necessary for amino acid analysis with high performance liquid chromatography.

NIRS functions by radiating a sample with light in the near-infrared region of the electromagnetic spectrum (12800-4000cm^{-1}), then capturing the reflected wavelengths with a detector. By comparing the initial wavelengths to those that are reflected, the wavelengths absorbed by the sample can be determined. The specific wavelength absorbed depends on the unique vibrational frequency of the chemical bonds between atoms in the sample (Mayerhofer, 2020) with the relationship being that the more a given specific bond type is present in the sample, the more a particular wavelength is absorbed. This association between concentration of bond types and wavelengths absorption can be plotted on a spectrogram, and information about the chemical composition can be ascertained from it (Kowalski, 1975).

However, the association between the amount of particular bonds between atoms and absorbance is not enough to determine a sample's composition and the concentration of a particular nutrient of interest. Many chemical bonds such as carbon to hydrogen, oxygen to hydrogen or nitrogen to hydrogen are common in fatty acids, sugar and amino acid molecules. In addition, many molecules in feedstuffs will not be in an unbound (free) form and will be associated with or even chemically bound to other molecules present in the heterogenous mixture. As different molecules all have unique arrangements of atoms and corresponding chemical bonds each molecule within a sample will ultimately have a unique NIRS fingerprint that can be identified. For example, a molecule of free lysine will have a unique fingerprint, and this fingerprint would be slightly altered, but still identifiable, if it was bound to a carbohydrate molecule or to another amino acid. So extensive statistical analyses are needed to identify all the

potential spectral relationships of the molecule of interest and correlate them to the determined reference value of the molecule of interest which would be in our example the total lysine content of the sample determined by high performance liquid chromatography (Fearn, 2005). Initially, success with NIRS in predicting the content of something in complex sample was limited by the lack of computer processing capability, but as computer processing power has increased NIRS's accuracy of predicting things like nutrient content has improved which has increased its use as a tool in animal nutrition.

Sample preparation

NIRS analysis is based on reflected light captured by a detector, and therefore care must be taken to avoid factors that alter the reflected light. Light scattering occurs when the sample refracts, diffracts, or re-reflects the reflected light. The result is an alteration of the detected absorbance bands by the sensor. Light scattering can be impacted by factors such as particle size, temperature, and moisture content of the sample. Differences in these factors across samples increase variation in the predictive model and therefore care should be taken to make them as uniform as possible when scanning samples (Murray and Cowe, 2004). Consistency in sample preparation therefore is essential in the creation of accurate measurements using NIRS.

Uniform grinding of the sample decreases variation in particle size and grinding samples to a smaller particle size increases the surface area exposed to the light source during scanning, and improving both of these factors will enhance the predictive capacity of the NIRS model (Givens et al., 1997; Bakalli, 2000). Grinders reduce particle size by one of two methods. Shearing mills use knives before the sample is forced through the sieve and collected. Impaction grinders crush samples against silicon carbide-impregnated walls before passing through the sieve for collection. Excessive grinding of samples should be avoided as the heat generated can

evaporate moisture, denature proteins, or heat catalyze chemical reactions, resulting in altered spectra. By having consistent grinding methodology, variation in the model caused by differences in particle size or heat production during grinding are reduced.

The temperature of the sample and machine can also alter spectral output. The vibrational frequency of molecular bonds is changed by the temperature, and therefore the specific wavelength that is absorbed would also change (Asselin and Sandorfy, 1970). If the temperature is not constant when scanning all samples variation in the model increases and predictive ability decreases (Hansen et al., 2000). For these reasons, samples are allowed to cool or warm to room temperature prior to scanning. Additionally, the multipurpose analyzer with which the NIRS analysis occurs, is given time in accordance with the manufacturer's recommendations for it to warm-up and the temperature to stabilize prior to use. These protocols ensure the least possible variance is introduced to the spectra scans due to temperature.

Water is a stronger absorber in the near-infrared region and as a result, a high moisture content can negatively affect the spectral analysis (Büning-Pfaue, 2003) as hydrogen-oxygen bonds found in water molecules occur in lipids, carbohydrates and proteins. Water also contributes to light scattering as water can reflect light. These factors result in decreased predictive ability for other molecules of the sample (J. B. Reeves, 1995; Büning-Pfaue, 2003). As feed samples are scanned on an as is basis rather than on a dry matter basis because their inclusion in diets is on as is basis, moisture is going to be present in the samples and thus some regions in a full spectral scan across the near infrared region with strong O-H bond saturation from water may have to be ignored if water content is excessive.

Spectral preprocessing

Even with careful and consistent sample preparation, it is inevitable that laboratory environmental factors and light scattering are going to be factors influencing sample spectra. These irreplicable and valueless deviances in the spectra unrelated to sample chemical variations are referred to as noise. Preprocessing methods minimize the effects of noise for the purpose of improving the NIRS predictive model (Luypaerta et al., 2004). One such preprocessing method is multiplicative scatter correction. It reduces spectral distortion specifically from light scattering. Another preprocessing method is calculating the first derivative of the spectrograph which will help resolve overlapping absorbance peaks (Conzen, 2014). Vector normalization is a third preprocessing method and works by standardizing baseline shifts and spectral peaks across the scans of all the samples in a data set.

Calibration curve development

After sufficient samples have had their NIRS spectral scan completed, the next step is to create a calibration curve. The calibration curve is a predictive model built to find patterns between the spectral data of the samples and the corresponding determined reference value of the parameter of interest in the samples. The reference data is obtained from the determination of the component of interest by another analytical technique. For example, if you were building a NIRS calibration curve to determine gross energy content of high protein soybean products you would scan each sample of the calibration data set by NIRS and then each scan would be correlated with its laboratory determined gross energy content determined by bomb calorimetry. The procedure of finding and correlating the patterns between spectral data and reference values is completed by the analytical software package of the NIRS system.

The practice of combining statistical methodologies with analytical chemistry to develop new predictive methods is referred to as chemometrics. Chemometric methods allow predictions of complex multidimensional variables that otherwise would be difficult to model (Kowalski, 1975). Carbohydrates, proteins, and lipids have many of the same functional groups, therefore examination of multiple regions of the near-infrared spectrum are required to determine a component of interest (Fearn, 2005). This makes it an ideal candidate for chemometric methods and thus the NIRS software uses these tools to create calibration curves. The most used chemometric methods used for NIRS are multiple linear regression, principal component regression, and partial least squares regression.

Multiple linear regression was among one of the first tools used for the chemometric analysis of NIRS. Multiple linear regression uses spectral outputs from a limited number of wavelengths from the NIRS analysis that have the strongest correlation to the reference values to create a calibration (Pérez-Marín et al., 2007). A fundamental flaw of this methodology is that in biological samples a large degree of collinearity of variables exists as the chemical and physical properties are intrinsically linked. Chemical constituents do not exist independently in the sample and therefore, the resonance of one bond can alter the wavelength of another (Agelet and Hurburgh, 2010). For these reasons multiple linear regression is not the most reliable means of chemometric analysis unless it is being used to predict the presence of a unique molecule that does not interact with a host of other molecules and generates a very specific and distinct NIRS spectral fingerprint at a specific wavelength.

Principle component regression is another chemometric method used for NIRS analysis. A principal component analysis first divides the full range of the NIRS spectral scan into spectral patterns (Naes and Martens, 1988). The second step is then to perform a linear regression

between the spectral patterns of the samples with their known reference values. This reduces collinearity compared to just a linear regression. However, using the entire spectra range may create undo variation in the prediction model because not all spectral data may be relevant to predicting the parameter of interest (Naes and Martens, 1988).

A third chemometric method, partial least squares, is also used for NIRS analysis. The partial least squares method uses both the spectral data and the reference values in multiple replications to determine the most important NIR spectral regions that best correlate to the known values. The result is that only the most highly correlated and important variables are included in the NIRS calibration prediction (Agelet and Hurburgh, 2010). Due to the precision of this method, it has become the standard for NIRS calibration curve development (Conzen, 2014).

Sample size is important for building a functional calibration curve given the important role statistics play in building the calibration model. Literature on the subject suggests a calibration curve should contain a minimum of 50 samples or for large sample sets 10 to 20% of all the samples (Westerhaus, 1989). Sample selection is also essential for building the calibration curve. The samples selected for inclusion in the calibration curve must be reflective of the natural variation observed within the samples. For example, if you were building a NIRS calibration curve to predict the lysine content in high protein soybean products, you would need the samples selected for the calibration curve to have the full range of the observed concentrations of lysine in these products. If this is not done, the result would be an inadequate prediction model that could not be used to accurately predict the content of lysine for example in future samples in which the lysine content was above or below the concentration range used in the calibration (Fearn, 2005).

The next step for NIRS analysis is the optimization of the calibration curve. The calibration will be displayed as a regression model of the reference values compared to their corresponding NIRS predicted values. This regression analysis should be examined for outliers which can negatively impact the calibration's predictive ability and may need to be removed. (Kovalenko et al., 2006). Spectral outliers have spectral results that are largely different from other spectra, while concentration outliers have greatly different predicted values based on their spectral results relative to their reference values. Spectral outliers could be an indication of an issue with sample preparation, scanning, or chemical compositional differences. Concentration outliers can indicate there may have been an issue in the determination of the reference value. Ideally, the sample raw material will still exist so that a new NIRS scan can be completed and the laboratory determined reference value can be determined again on the sample. This typically identifies the error in the original scanning procedure or reference value determination and allows for correction. If the original sample does not exist for rescanning and reference value redetermination, the sample should be deleted from the calibration, especially if the calibration dataset is large. Finally, if the sample is rescanned and the same spectral result is obtained and its reference value is established as correct on reanalysis, then the sample should be retained in the calibration as it represents natural variation that can occur in the sample set and could be encountered again in future samples (Van Kempen and Bodin, 1998).

Validation

Following the creation and optimization of the calibration curve a validation is used to test the calibration's predictive ability by creating a validation curve. The calibration's predictive ability is presented as a regression model that compares the laboratory determined reference values of the validation samples to their NIRS predicted values determined using the

calibration curve. The type of validation used is based on the number of samples available. A cross validation is used when there are 50 or less total samples (Conzen, 2014). Cross validation works by removing one sample at a time from the calibration and predicting its value and then returning it to the calibration and repeating the process with the next sample. This is done until all samples have been validated in this manner. The comparison between the value of the validation to the reference value serves as the indicator of the calibration's predictive ability. In this way a smaller sample size can be validated.

A much more robust validation occurs when more samples are available, and an independent validation test set is created from a pool of samples that were not part of the calibration. It is critical for testing the calibration, that the natural variation of the samples included in the validation curve to be reflective of the variation of the samples included in the calibration curve. Failure to do so would yield inaccurate results of the calibrations predictive ability. For example, if the range of concentrations of lysine used in the calibration curve was from 2% to 5%, but the range in lysine content in the validation set was only from 3% to 4%, there would be no validation of how the calibration model performs outside the 3%-4% concentration range. As the number of validation samples increase, the proportion of total samples in the validation can increase to better indicate the extent of the calibration's predictive ability and indicate areas of weakness in the calibration curve that could be corrected by including more samples in the calibration (Conzen, 2014).

NIRS statistics

Several statistics are used to indicate the robustness of calibration and validation regression curves. Coefficient of determination (R^2) describes the best fit line of the data compared to the line of true fit. It indicates how well the NIRS predicted values match the

corresponding laboratory determined values. However, R^2 has limited usage for evaluating a calibration curves predictive potential, because with a small range of values, even small deviations from the true value can have a large impact on the R^2 value (Davies and Fearn, 2006). Another metric to determine predictive ability is the root mean standard error of the prediction (RMSEP). The RMSEP provides the average difference between the predicted and the true value. Another statistic used in curve evaluation is bias. Bias indicates whether calibrations have a tendency to predict greater or lesser values than the reference values. While useful for this purpose, bias does not indicate an average deviation from the reference values. This means that as long as variance is equal for over and under predictions, bias would indicate the model is fine. Residual predictive deviation, (RPD) provides a standardized indication of a model's predictive ability. The RPD value is calculated by dividing the standard deviation of the observed values by the bias-corrected mean error of prediction. This value is thereby a metric of average error independent of the parameter's values, making it a standardized metric between models for the predictive ability. An RPD value of 3 or more indicates a model has good predictive capability (Williams, 2001). Practical considerations when evaluating the statistical accuracy of the validation curve for nutrient levels in a feed ingredient is based on calculating the percent deviation of the predicted value from the true laboratory determined value for each validation sample (Cope, 2021) as this will provide the nutritionist with the true value of error on each sample rather than average error of the mean. When building safety margins into formulated diets to ensure requirements are met the nutritionist needs to know the worst-case scenario rather than the average deviation.

Summary

Over 50% of the soybean meal produced worldwide is fed to poultry. Most of the soybean meal produced results from the solvent extraction of the oil from the soybean, and this product is what is utilized most heavily by the poultry industry. However, the poultry industry has increased its utilization of soybean meal in which the oil has been mechanically extracted from the soybean and full fat soybean meal in which none of the oil has been extracted from the soybean. Mechanically extracted soybean meal is utilized by organic and all natural poultry production systems. Full fat soybean meal is also utilized in these production systems and is also favored in production systems that favor sustainability, as the utilization of full fat soybeans avoids the energy costs associated with oil extraction and the need to add processed vegetable or animal oil/fat to the diet. The gross energy and total content of each amino acid differs across these different forms of soybean meal. In addition, the gross energy and total content of each amino acid will differ within each soybean meal product category based on the genetic variety of the soybeans utilized to make the meal, and the agronomic conditions under which the soybeans utilized to make the meal were produced. Gross energy of soybean meal samples can be determined by bomb calorimeter and the amino acid content can be determined by high performance liquid chromatography. But getting results from these procedures takes days and the procedures are expensive and generate chemical waste. In contrast, if NIRS technology could be utilized to determine these parameters in soybean samples the results will be available in real time and less expensive. The availability of the results in real time would also allow poultry nutritionists to formulate their poultry diets on the soybean meal actually delivered to their feed mills which would eliminate potential for formulating over or under the bird's nutritional requirements which is economically costly.

CHAPTER 2

STATEMENT OF PURPOSE

High protein soy products which include full fat soybean meal, mechanically extracted soybean meal, solvent extracted soybean meal, fermented soybean meal, enzymatically digested soybean meal, soy protein contrate and soy protein isolate often provide the majority of the amino acid content and up to 25% of the total energy of animal diets, especially for poultry diets. In the United States, the poultry industry has started to utilize fermented and/or enzymatically digested soybean meal products in pre-starter diets. Furthermore, the use of full fat and mechanically processed soybean meal has increased with the increase in all natural and organic poultry production systems. The amino acid and gross energy content of soybean meal varies considerably based on the processing procedures utilized in their production with full fat soybean meal having the highest gross energy content and lowest percent amino acid content and with solvent extracted soybean meal or this product further processed though fermentation and/or enzymatic digestion having the lowest gross energy content but highest amino acid content. In addition, the amino acid and gross energy content of high protein soy products can vary based on soybean cultivar genetics, environmental growing conditions, and the level of pesticide and fertilization utilization during the growing period.

Knowing the gross energy content and total individual amino acid concentration of feed ingredients can assist nutritionists in formulating diets that meet the energy and amino acid requirements of the animals they are feeding. Using performance liquid chromatography and bomb calorimetry to determine individual amino acid and gross energy content, respectively in

feed ingredients, is costly, produces chemical waste and takes days to complete. In contrast, NIRS is a rapid analysis method that can be used to predict nutrient composition of feed ingredients by using a multipurpose analyzer that analyzes the reflectance of a near infrared light scan of the bonds between the molecules making up a feed ingredient. For the NIRS procedure, feed ingredient samples must be obtained, prepared for scanning and then scanned across the near infrared spectrum. Using advanced statistical software capabilities, the spectral data obtained for each sample is then associated with the known determined value of the parameter of interest such as the bomb calorimeter determined gross energy for each sample. Based on the associated data of all the samples, a calibration curve is developed for the parameter of interest such as gross energy. Then this calibration curve is utilized to predict the content of the parameter of interest in subsequent samples. The accuracy of this calibration curve is validated with a separate set of feed ingredient samples in which the NIRS predicted value of each sample is compared to the known determined value.

The goal of the current research is to create and validate NIRS calibrations for the determination of gross energy content and the concentration of each individual amino acid in high protein soy products, and to have these calibrations predict the gross energy and individual amino acid content in high protein soy products with less than a 5% deviation from the bomb calorimeter and high performance liquid chromatography determined values.

CHAPTER 3

MATERIALS AND METHODS

Sample Selection

Two hundred sixty-nine high protein soybean product samples (Table 3.1) were obtained for creating NIRS calibration curves to predict gross energy content and the content of individual amino acids. The samples were obtained from various national and international manufacturers to increase sample variability and thus better reflect what gets incorporated into animal diets. The number of samples obtained in each category (full fat soybean meal, solvent extracted soybean meal, mechanically expelled soybean meal, high protein soy product, soy concentrate, and soy isolate) roughly reflects their overall utilization by the United States poultry industry in which solvent extracted SBM is used most extensively, followed by mechanically expelled SBM which is utilized in all-natural material and organic production systems.

Sample preparation

Samples were ground with a Retsch type ZM200 centrifugal laboratory grinder (Retsch GmbH, Haan, Germany) at 16,000 RPM with a 1-millimeter screen. This centrifugal mill uses shearing and impact force to reduce particle size, and is air cooled to minimize heat production transfer to the sample. The grinder was cleaned between each sample by vacuuming with a wet/dry professional vacuum (Rigid, model WD 14500, St Elyria, OH) then wiped with absorbent towelettes (Kimwipes, Kimberly-Clark, Roswell GA). The ground sample was bagged in a labeled sterile plastic sample bag (Thermo Fisher Scientific, Waltham, MA), and stored at negative 20° C for subsequent use.

Chemical reference methods

For each ground sample, crude protein was determined (Association of Analytical Chemists, 2006) and gross energy was determined with a bomb calorimeter (6400 Oxygen Bomb Calorimeter, Parr Instruments Company, Moline, IL). These analytical procedures were computed by the University of Georgia Agricultural and Environmental Laboratories (Athens, GA). The total amino acid content of each ground sample was determined by high performance liquid chromatography (Association of Analytical Chemists, 2006) by the University of Missouri Agriculture Experiment Station Chemical Laboratories (Columbia, MO).

FT-NIR analysis

A near-infrared analysis of each sample was conducted by a Bruker multipurpose analyzer, Fourier transform, near-infrared spectrometer (MPA). The MPA was stored in a temperature and climate-controlled laboratory. The MPA was outfitted with a sample-cup rotator to ensure uniform mixing of the sample to endure uniform exposure of the sample particles to the NIR light source while it was scanned. The MPA was allowed to warm-up for 30 minutes and its dried desiccant beads were replaced before the samples were scanned. This was done to allow the near-infrared lamp and electronics to stabilize and to ensure a dry sample chamber. A background light scan was performed to calibrate the machine before samples were scanned.

Frozen samples were thawed and allowed to equilibrate to room temperature and then individual samples were mixed thoroughly before scanning. Samples were then deposited into quartz bottom NIRS sample cups (4 cm diameter) until no light could be seen through the bottom

Figure 3.1. The high protein soybean product types obtained for creating and validating near-infrared spectroscopy calibrations to predict gross energy and amino acid content.

Sample type	Number of samples
Full fat soybean meal	3
Mechanically expelled soybean meal	34
Solvent extracted soybean meal	184
High protein soy product[1]	24
Soy concentrate	18
Soy isolate	6
Total	269

[1]Fermented and or enzymatically digested soy protein product with a crude protein content between soybean meal and soy concentrate.

but were not pressed down into the cup. Samples were scanned on an as is basis without drying as poultry diets are formulated and made using ingredients on as is basis.

A spectral scan of each sample across the complete near-infrared range (12800-4000cm^{-1}) was completed. The MPA was equipped with OPUS 8.5 (Billerica, MA) software, to create calibration and validation curves. The OPUS software determined the wavelength regions and the data preprocessing techniques to generate models of each parameter (gross energy or individual amino acid content) determined (Table 3.2). Samples for a given parameter used in the calibration were not used in the validation and vice versa. Samples from each soy protein category were distributed evenly between the validation and calibration curves. The number of samples in each calibration curve was the minimum number of samples needed to encompass the entire range of values obtained with the high performance liquid chromatography and bomb calorimeter values as well as the number needed to maintain the highest predictive accuracy of the calibration curve.

Statistical analysis:

All statistical analyses of the calibration and validation data were completed using Bruker OPUS version 8.5 software package. Calibration curves were evaluated using their coefficient of determination value (R^2), root mean square error of the estimate (RMSEE) and residual predictive deviation (RPD). Validation curves were evaluated using their coefficient of determination value (R^2), root mean square error of the prediction (RMSEP), and bias.

The R^2 value represents how closely the data fits the model. It is calculated as one minus the quotient of the sum of residuals squared and the total sum of squares. It shows how well the model explains the variance of the data set by comparing the laboratory determined values to the

Table 3.2. The OPUS 8.5 software package of the multipurpose analyzer determined for each parameter of interest (gross energy content or the individual amino acid content) the most useful wavelength regions of the full near infrared spectral analyses of the samples and the spectral data preprocessing techniques needed to generate the most accurate calibration models.

Parameter	Selected wavelength regions (cm^{-1})	Preprocessing method
Inclusive curves[1]		
Alanine	9003 – 7444, 6904 – 6395, 4953 – 4598	First derivative
Arginine	7456 – 7096, 6904 – 6024, 4600 – 4496	First derivative, vector normalization
Aspartic acid	7452 – 7097, 6904 – 6395, 4775 – 4598, 4505 – 4197	Vector normalization
Cysteine	8000 – 7448, 6032 – 5296, 4504 – 4200	First derivative, vector normalization
Glutamic acid	7452 – 7097, 6904 – 6395, 4505 – 4197	Vector normalization
Glycine	6904 – 5292	First derivative, vector normalization
Histidine	9003 – 7444, 6904 – 6395, 4953 – 4768, 4505 – 4197	First derivative, vector normalization
Isoleucine	8008 – 7444, 6904 – 6395, 6033 – 5392	First derivative, vector normalization
Leucine	6904 – 6395, 6033 – 5300, 4953 – 4768	Vector normalization
Lysine	8016 – 7088, 6416 – 6016	First derivative, vector normalization
Methionine	6403 – 6025, 4775 – 4598, 4505 – 4197	First derivative, vector normalization
Phenylalanine	9000 – 8000, 7456 – 7096, 6904 – 6024, 4952 – 4768	Vector normalization
Proline	8008 – 7090, 6904 – 5292, 4783 – 4598, 4505 – 4197	Vector normalization
Serine	8008 – 7440, 6408 – 6024, 4960 – 4760, 4512 – 4192	Vector normalization
Threonine	9003 – 7444, 6904 – 5300, 4606 – 4197	Vector normalization
Tryptophan	8008 – 7088, 6408 – 6016, 4960 – 4592	First derivative, vector normalization
Tyrosine	9003 – 7444, 6033 – 5300, 4775 – 4598, 4505 – 4197	Vector normalization
Valine	6904 – 6395, 6033 – 5300, 4775 – 4598	Vector normalization
Gross energy	9003 – 8000, 6904 – 6025, 5038 – 4598, 4498 – 4142	Multiplicative scattering correction
Next step curves[2]		
Alanine	8008 – 7444, 6403 – 5300, 4775 – 4598	First derivative
Arginine	9003 – 8000, 7452 – 7097, 4606 -4498	First derivative
Aspartic acid	9003 – 8000, 6403 – 5300, 4606 – 4498	First derivative
Cysteine	8000 – 7448, 6032 – 5296, 4776 – 4600	First derivative
Glutamic acid	9003 – 8000, 7460 – 7090, 6033 – 5292, 4783 – 4598, 4505 – 4197	First derivative
Glycine	6912 – 6392, 6040 – 5296	First derivative
Histidine	9003 – 8000, 7452 – 7097, 6904 – 6395, 4953 – 4598, 4505 – 4197	First derivative
Isoleucine	8008 – 7091, 6033 – 5300, 4953 – 4768, 4606 – 4498	First derivative, vector normalization

Leucine	9003 – 8000, 6403 – 5300	First derivative, vector normalization
Lysine	9003 – 8000, 6403 – 5300,4606 – 4498	First derivative
Methionine	7456 – 7096, 4952 – 4768, 4504 – 4200	Vector normalization
Phenylalanine	9003 – 8000, 6904 – 6395, 6033 – 5300, 4775 – 4498	First derivative
Proline	8008 – 7097, 6403 – 5300, 4606 – 4197	Vector normalization
Serine	4960 – 4760, 4512 – 4192	Vector normalization
Threonine	9000 – 7448, 6904 – 6400, 6032 – 5296, 4600 – 4496	First derivative, vector normalization
Tryptophan	8008 – 7097, 6033 – 5300, 4775 – 4598, 4505 – 4197	First derivative, vector normalization
Tyrosine	9003 – 8000, 6904 – 6395, 4775 – 4197	First derivative
Valine	9003 – 7097, 6403 – 5300, 4775 – 4598	First derivative
Gross energy	9003 – 8000, 6904 – 6025, 5038 – 4598, 4498 – 4143	First derivative, vector normalization

[1]Inclusive curves contain all of the high soy protein sample categories (full fat soybean meal, mechanically extracted soybean meal, solvent extracted soybean meal, high protein soy product, soy concentrate and soy isolate).
[2]Next step curves contain only mechanically extracted and solvent extracted soybean meal.

NIRS predicted values and how well the resulting points align with the regression model, often called the goodness of fit.

$$R^{2-} = 1 - \frac{SSE}{\sum_{i=M}^{M}(Y_i - Y_m)^2}$$

The RPD value is calculated by dividing the standard deviation of the observed values by the bias-corrected mean error of prediction. The RPD is more qualitative than the R^2 as the denominator in the RPD is reflective of the error in the model and as that error is lowered, the value lowers which allows the RPD to increase, reflective of the model improvement. The resulting value is a representation of the models' predictive capacity with values over 3.0 considered meaningful, and higher values serving as a better indicator of predictive accuracy (Conzen, 2014).

$$\text{RPD} = .\frac{\sqrt{\frac{1}{M-1} * \sum_{i=1}^{M}(Y_1^{meas} - Y_1^{pred})}2}{\sqrt{\frac{1}{M-1} * \sum_{i=1}^{M}(Y_1^{meas} - Y_1^{pred} - bias)}2}$$

The OPUS software also calculated the standard error of the estimate (RMSEE) for calibration curves. The RMSEE represents the average distance the observed values fall from the regression line established by the calibration model (Conzen, 2014). The RMSEP determination is equivalent to the RMSEE except it is based on the error of the prediction for the validation curve. Lower values indicate that the predicted value on average is closer to the true value.

$$\text{RMSEE/RMSEP} = \sqrt{\frac{1}{M} * \sum_{i=1}^{M}(Y_1^{meas} - Y_1^{pred})^2}$$

The bias of a validation model is the deviation of the line established by the model from the ideal line of x=y (Conzen, 2014). The model should attempt to minimize bias when establishing a regression line of true to predicted values.

Bias = $\frac{\sum_{i=1}^{M} Y_1^{meas} - Y_1^{pred}}{M}$

In addition to OPUS statistical evaluations, a more practical evaluation of calibration models was also done by calculating the absolute deviation of the predicted value from the true value. Samples in the validation curve whose predicted value deviated from their laboratory determined true values by less than or equal to 2.5%, 5% and 10% were quantified as done by Cope (2021).

CHAPTER 4

RESULTS

Alanine

The total alanine content of the samples used for the full NIRS calibration curve for this amino acid ranged from 1.60 to 3.80 percent (Table 4.1). The corresponding validation curve had an R^2 value of 0.96 and while 100 percent of the of the validation samples had predicted values that deviated less than 5% from their determined value, only 73% of the validation samples had predicted values that deviated less than 2.5% from their laboratory determined value (Table 4.2). Given this result, the decision was made to make next step calibration and validation curves (Table 4.1) that focused only on solvent and mechanically extracted SBM as these products are what are almost exclusively used in formulating diets in the United States by the poultry industry. For the next step prediction, only 1% of the validation samples deviated 2.5% or more from their laboratory determined value of total alanine content (Table 4.2).

Arginine

The arginine calibration curve containing all the high protein dietary ingredient categories contained 134 samples while the next step arginine calibration curve contained 108 samples (Table 4.3). For the next step validation, 99% of the validation samples deviated less than 2.5% from their laboratory determined value of total arginine (Table 4.4).

Aspartic acid

The total aspartic acid content of the samples used for the full NIRS calibration curve ranged from 4.16 to 10.08 percent (Table 4.5). The validation curve for the full sample set containing all of the high protein ingredient categories had 76% of its samples with predicted

Table 4.1. Total alanine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	1.65 ± 0.045	1.60	1.69	1	1.65	1.65	1.65
Mechanical SBM	13	2.01 ± 0.020	1.88	2.12	21	1.99 ± 0.015	1.86	2.11
Solvent SBM	94	2.05 ± 0.006	1.88	2.17	88	2.05 ± 0.006	1.91	2.21
High protein soy[1]	13	2.42 ± 0.063	2.19	3.05	13	2.43 ± 0.052	2.10	2.79
Soy protein concentrate	8	2.94 ± 0.019	2.87	3.00	10	2.90 ± 0.036	2.61	3.01
Soy protein isolate	4	3.70 ± 0.035	3.64	3.80	2	3.71 ± 0.035	3.67	3.74
Total	134	2.18 ± 0.032	1.60	3.80	135	2.16 ± 0.028	1.65	3.74

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	17	2.01 ± 0.018	1.86	2.12	17	1.99 ± 0.015	1.88	2.07
Solvent SBM	91	2.05 ± 0.007	1.88	2.21	91	2.04 ± 0.004	1.93	2.16
Total	108	2.04 ± 0.007	1.86	2.21	108	2.04 ± 0.005	1.88	2.16

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.2. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for alanine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.96	0.98	0.25	0.79
RMSEE/RMSEP[1]	0.07	0.05	0.06	0.02
RPD[2]	5.13	7.05	1.15	2.19
Bias[3]		0.002		-0.004
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		16		0
97.50 - 99.99		30		43
100.00 – 102.49		43		56
102.50 – 104.99		11		1
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		73		99
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.3. Total arginine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	2.58 ± 0.010	2.57	2.59	1	2.92	2.92	2.92
Mechanical SBM	17	3.39 ± 0.044	3.15	3.75	17	3.38 ± 0.037	3.03	3.58
Solvent SBM	91	3.42 ± 0.016	3.00	3.98	91	3.41 ± 0.008	3.19	3.63
High protein soy[1]	11	4.04 ± 0.098	3.70	4.70	15	3.86 ± 0.095	3.48	4.96
Soy protein concentrate	9	4.82 ± 0.027	4.70	4.99	9	4.80 ± 0.087	4.15	5.04
Soy protein isolate	4	6.60 ± 0.028	6.52	6.64	2	6.51 ± 0.095	6.41	6.60
Total	134	3.64 ± 0.058	2.57	6.64	135	3.59 ± 0.046	2.92	6.60

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	18	3.38 ± 0.046	3.03	3.75	16	3.38 ± 0.032	3.15	3.58
Solvent SBM	90	3.40 ± 0.016	3.00	3.98	92	3.43 ± 0.009	3.19	3.82
Total	108	3.39 ± 0.015	3.00	3.98	108	3.42 ± 0.009	3.15	3.82

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.4. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for arginine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.95	0.99	0.23	0.81
RMSEE/RMSEP[1]	0.15	0.06	0.14	0.04
RPD[2]	4.67	8.65	1.14	2.37
Bias[3]		0.000		0.010
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		5		1
97.50 - 99.99		44		59
100.00 – 102.49		43		40
102.50 – 104.99		8		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		87		99
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.5. Total aspartic acid content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	4.31 ± 0.145	4.16	4.45	1	4.30	4.30	4.30
Mechanical SBM	17	5.30 ± 0.054	4.90	5.66	17	5.24 ± 0.051	4.78	5.56
Solvent SBM	88	5.36 ± 0.022	4.84	5.86	94	5.38 ± 0.015	4.95	5.79
High protein soy[1]	15	6.29 ± 0.172	5.60	8.29	11	6.22 ± 0.152	5.55	7.28
Soy protein concentrate	8	7.43 ± 0.132	6.54	7.72	10	7.58 ± 0.039	7.40	7.78
Soy protein isolate	4	9.91 ± 0.079	9.71	10.08	2	9.87 ± 0.115	9.75	9.98
Total	134	5.70 ± 0.085	4.16	10.08	135	5.65 ± 0.072	4.30	9.98

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	12	5.24 ± 0.077	4.78	5.66	22	5.28 ± 0.041	4.90	5.57
Solvent SBM	96	5.36 ± 0.021	4.84	5.86	86	5.37 ± 0.015	5.09	5.81
Total	108	5.35 ± 0.021	4.78	5.86	108	5.35 ± 0.015	4.90	5.81

[1]Contains both fermented and/or enzymatically treated soy protein products.

values that deviated less than 2.5% from their actual high performance liquid chromatography determined values, and for the next step validation 100% of the validation samples had NIRS predicted values that deviated less than 2.5% from their laboratory determined value (Table 4.6).

Cysteine

The cysteine content of the high protein soy feed ingredients ranged from 0.54 to 1.12 percent (Table 4.7). Given the lower content of cysteine in the high soy protein products and its ability to form disulfide bonds, it may not be surprising that only 48% of the validation samples in the full curve and only 81% of the validation samples in the next step curve had NIRS predicted values that deviated less than 2.5% from their laboratory determined values, respectively (Table 4.8).

Glutamic acid

The glutamic acid content of the high protein soy feed ingredients ranged from 6.42 to 17.36 percent (Table 4.9). The RMSEP values for the full validation and next step validation curves indicated that the NIRS predicted glutamic acid concentration of the validation samples differed from the laboratory determined concentration values on average by 0.21 and 0.09 percent, respectively (Table 4.10).

Glycine

The glycine calibration curve containing all the high protein dietary ingredient categories contained 129 samples while the next step glycine calibration curve contained 103 samples (Table 4.11). For the next step validation 100% of the validation samples deviated 2.5% or less from their laboratory determined value of total glycine (Table 4.12).

Table 4.6. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for aspartic acid for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.97	0.98	0.38	0.83
RMSEE/RMSEP[1]	0.19	0.11	0.17	0.06
RPD[2]	5.41	7.31	1.27	2.41
Bias[3]		-0.006		0.006
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		13		0
97.50 - 99.99		35		49
100.00 – 102.49		41		51
102.50 – 104.99		12		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		76		100
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.7. Total cysteine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	1	0.54	0.54	0.54	2	0.57 ± 0.030	0.54	0.60
Mechanical SBM	21	0.66 ± 0.010	0.59	0.72	13	0.67 ± 0.008	0.62	0.72
Solvent SBM	88	0.67 ± 0.004	0.56	0.79	94	0.68 ± 0.004	0.61	0.75
High protein soy[1]	11	0.82 ± 0.029	0.74	1.09	15	0.77 ± 0.013	0.71	0.88
Soy protein concentrate	9	0.95 ± 0.011	0.89	1.01	2	0.94 ± 0.014	0.87	1.02
Soy protein isolate	4	1.06 ± 0.020	1.03	1.12	9	1.11 ± 0.015	1.09	1.12
Total	134	0.71 ± 0.010	0.54	1.12	135	0.71 ± 0.008	0.54	1.12

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	24	0.66 ± 0.009	0.59	0.72	10	0.67 ± 0.010	0.62	0.72
Solvent SBM	84	0.68 ± 0.006	0.56	0.79	98	0.68 ± 0.002	0.61	0.75
Total	108	0.67 ± 0.005	0.56	0.79	108	0.68 ± 0.002	0.61	0.75

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.8. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for cysteine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.86	0.92	0.15	0.67
RMSEE/RMSEP[1]	0.04	0.03	0.05	0.01
RPD[2]	2.63	3.52	1.08	1.75
Bias[3]		0.006		0.002
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		2		0
92.50 - 94.99		13		0
95.00 - 97.49		18		13
97.50 - 99.99		25		40
100.00 – 102.49		23		41
102.50 – 104.99		12		6
105.00 – 107.49		6		0
107.50 – 110.00		1		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		48		81
< ± 5.0% deviation from determined value		77		100
< ± 7.5% deviation from determined value		96		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.9. Total glutamic acid content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	6.69 ± 0.265	6.42	6.95	1	6.79	6.79	6.79
Mechanical SBM	18	8.37 ± 0.089	7.81	9.14	16	8.49 ± 0.113	7.50	9.11
Solvent SBM	88	8.60 ± 0.034	7.71	9.41	94	8.62 ± 0.029	7.85	9.69
High protein soy[1]	13	9.89 ± 0.219	8.84	11.57	13	10.09± 0.276	8.93	12.60
Soy protein concentrate	9	12.51± 0.287	10.35	13.16	9	12.51± 0.066	12.22	12.86
Soy protein isolate	4	16.96± 0.162	16.57	17.36	2	16.71± 0.085	16.62	16.79
Total	134	9.18 ± 0.155	6.42	17.36	135	9.11 ± 0.126	6.79	16.79

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	20	8.42 ± 0.104	7.50	9.14	14	8.44 ± 0.091	7.81	8.85
Solvent SBM	88	8.62 ± 0.041	7.71	9.70	94	8.60 ± 0.023	7.85	9.54
Total	108	8.58 ± 0.039	7.50	9.70	108	8.58 ± 0.024	7.81	9.54

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.10. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for glutamic acid for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.96	0.98	0.39	0.86
RMSEE/RMSEP[1]	0.35	0.21	0.32	0.09
RPD[2]	5.22	7.08	1.28	2.65
Bias[3]		0.000		0.005
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		1		0
95.00 - 97.49		16		0
97.50 - 99.99		36		50
100.00 – 102.49		32		50
102.50 – 104.99		16		0
105.00 – 107.49		1		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		67		100
< ± 5.0% deviation from determined value		99		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.11. Total glycine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	1.59 ± 0.020	1.57	1.61	1	1.71	1.71	1.71
Mechanical SBM	14	1.96 ± 0.019	1.88	2.07	20	1.99 ± 0.019	1.77	2.08
Solvent SBM	86	2.00 ± 0.007	1.82	2.20	86	1.99 ± 0.005	1.88	2.18
High protein soy[1]	14	2.31 ± 0.038	2.07	2.60	12	2.38 ± 0.066	2.13	2.93
Soy protein concentrate	9	2.80 ± 0.027	2.59	2.85	9	2.86 ± 0.015	2.80	2.91
Soy protein isolate	4	3.64 ± 0.026	3.59	3.69	2	3.57 ± 0.015	3.55	3.58
Total	129	2.13 ± 0.032	1.57	3.69	130	2.11 ± 0.028	1.71	3.58

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	17	1.96 ± 0.020	1.77	2.08	17	1.99 ± 0.019	1.85	2.07
Solvent SBM	86	2.00 ± 0.008	1.82	2.20	86	1.99 ± 0.004	1.88	2.14
Total	103	2.00 ± 0.007	1.77	2.20	103	1.99 ± 0.005	1.85	2.14

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.12. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for glycine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.96	0.99	0.32	0.81
RMSEE/RMSEP[1]	0.07	0.03	0.06	0.02
RPD[2]	5.24	10.50	1.21	2.29
Bias[3]		-0.004		0.003
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		1		0
97.50 - 99.99		44		56
100.00 – 102.49		48		44
102.50 – 104.99		8		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		91		100
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Histidine

The histidine content of the high protein soy feed ingredients ranged from 0.94 to 2.42 percent (Table 4.13). Only 70% of the validation samples for the full sample set that contained all the soy protein ingredient categories had values that deviated less than 2.5% from their high performance liquid chromatography determined values (Table 4.14), but for the next step data set that contained just mechanically and solvent extracted SBM, 94% of the validation samples had values that deviated 2.5% or less from their high performance liquid chromatography determined values (Table 4.14).

Isoleucine

The isoleucine calibration curve containing all the high protein dietary ingredient categories contained 116 samples while the next step isoleucine calibration curve contained 89 samples (Table 4.15). For the next step validation, 100% of the validation samples had NIRS predicted values that deviated 2.5% or less from their laboratory determined value of total isoleucine (Table 4.16).

Leucine

The leucine content of the high protein soy feed ingredients ranged from 2.83 to 7.22 percent (Table 4.17). The RPD value for the full sample set validation was 10.03, indicating that this was a robust statistical prediction model (Table 4.18). However, only 84% of the validation samples for the full sample curve had predicted values that deviated less than 2.5% from their high performance liquid chromatography determined values (Table 4.18). In contrast, the RPD value for the next step sample set validation was only 2.07, but 100% of the validation samples had predicted values that deviated 2.5% or less from their high performance liquid chromatography determined value of total leucine (Table 4.18).

Table 4.13. Total histidine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	0.99 ± 0.050	0.94	1.04	1	0.94	0.94	0.94
Mechanical SBM	16	1.23 ± 0.018	1.14	1.35	18	1.22 ± 0.013	1.11	1.35
Solvent SBM	93	1.25 ± 0.007	1.02	1.40	89	1.25 ± 0.004	1.15	1.36
High protein soy[1]	11	1.48 ± 0.039	1.31	1.74	15	1.40 ± 0.029	1.25	1.65
Soy protein concentrate	8	1.76 ± 0.037	1.50	1.82	10	1.80 ± 0.012	1.71	1.86
Soy protein isolate	4	2.36 ± 0.029	2.28	2.42	2	2.36 ± 0.025	2.33	2.38
Total	134	1.33 ± 0.021	0.94	2.42	135	1.32 ± 0.018	0.94	2.38

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	20	1.23 ± 0.016	1.11	1.35	14	1.21 ± 0.012	1.14	1.27
Solvent SBM	88	1.25 ± 0.008	1.02	1.40	94	1.25 ± 0.004	1.14	1.35
Total	108	1.25 ± 0.007	1.02	1.40	108	1.24 ± 0.004	1.14	1.35

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.14. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for histidine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.94	0.98	0.24	0.80
RMSEE/RMSEP[1]	0.06	0.03	0.06	0.02
RPD[2]	3.99	6.90	1.15	2.26
Bias[3]		-0.002		-0.003
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		14		4
97.50 - 99.99		33		38
100.00 – 102.49		37		56
102.50 – 104.99		16		2
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		70		94
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.15. Total isoleucine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	N	%	%	%
Full fat SBM	2	1.84 ± 0.005	1.83	1.84	1	1.91	1.91	1.91
Mechanical SBM	16	2.24 ± 0.026	2.10	2.41	16	2.22 ± 0.025	1.99	2.37
Solvent SBM	73	2.35 ± 0.010	2.18	2.55	76	2.30 ± 0.006	2.15	2.42
High protein soy[1]	13	2.79 ± 0.083	2.44	3.62	12	2.66 ± 0.060	2.34	3.11
Soy protein concentrate	10	3.29 ± 0.054	2.84	3.50	8	3.31 ± 0.016	3.24	3.36
Soy protein isolate	2	4.43 ± 0.025	4.40	4.45	4	4.46 ± 0.040	4.42	4.58
Total	116	2.49 ± 0.039	1.83	4.45	117	2.46 ± 0.044	1.91	4.58

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	13	2.23 ± 0.038	1.99	2.41	18	2.23 ± 0.017	2.13	2.37
Solvent SBM	76	2.33 ± 0.010	2.15	2.55	72	2.32 ± 0.007	2.17	2.48
Total	89	2.31 ± 0.011	1.99	2.55	90	2.31 ± 0.007	2.13	2.48

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.16. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for isoleucine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.95	0.99	0.35	0.86
RMSEE/RMSEP[1]	0.10	0.05	0.09	0.03
RPD[2]	4.41	9.47	1.24	2.65
Bias[3]		-0.002		-0.002
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		3		0
97.50 - 99.99		34		43
100.00 – 102.49		33		57
102.50 – 104.99		29		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		68		100
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.17. Total leucine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	2.93 ± 0.095	2.83	3.02	1	2.99	2.99	2.99
Mechanical SBM	16	3.61 ± 0.033	3.39	3.79	18	3.55 ± 0.028	3.27	3.70
Solvent SBM	90	3.65 ± 0.016	3.26	4.07	92	3.66 ± 0.008	3.46	3.84
High protein soy[1]	13	4.38 ± 0.132	3.87	5.70	13	4.24 ± 0.092	3.77	4.93
Soy protein concentrate	10	5.25 ± 0.051	4.91	5.51	8	5.29 ± 0.030	5.19	5.39
Soy protein isolate	3	7.08 ± 0.084	6.93	7.22	3	6.97 ± 0.033	6.94	7.04
Total	134	3.90 ± 0.061	2.83	7.22	135	3.86 ± 0.056	2.99	7.04

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	14	3.60 ± 0.042	3.27	3.79	20	3.55 ± 0.024	3.37	3.71
Solvent SBM	94	3.65 ± 0.016	3.26	4.07	88	3.66 ± 0.009	3.44	3.92
Total	108	3.64 ± 0.015	3.26	4.07	108	3.64 ± 0.009	3.37	3.92

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.18. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for leucine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.96	0.99	0.28	0.76
RMSEE/RMSEP[1]	0.15	0.06	0.13	0.04
RPD[2]	4.72	10.03	1.18	2.07
Bias[3]		-0.001		0.007
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		7		0
97.50 - 99.99		44		50
100.00 – 102.49		41		50
102.50 – 104.99		9		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		84		100
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Lysine

The lysine calibration curve containing all the high protein dietary ingredient categories contained 134 samples while the next step lysine calibration curve contained 107 samples (Table 4.19). For the next step validation, 99% of the samples had NIRS predicted values that deviated 2.5% or less from their laboratory determined value of total lysine content (Table 4.20).

Methionine

The methionine content of the high protein soy feed ingredients ranged from 0.51 to 1.18 percent (Table 4.21). As with cysteine, given the lower content of methionine in the high soy protein products and its ability to form disulfide bonds, the accuracy of the NIRS calibration curves was less than some of the other amino acids with only 78% of the validation samples in the full curve and 92% of the validation samples in the next step curve having predicted values that deviated less than 2.5% from their laboratory determined values, respectively (Table 4.22).

Phenylalanine

The phenylalanine calibration curve containing all the high protein dietary ingredient categories contained 134 samples while the next step phenylalanine calibration curve contained 108 samples (Table 4.23). For the next step validation, 100% of the validation samples had NIRS predicted values that deviated 2.5% or less from their laboratory determined value of total phenylalanine content (Table 4.24).

Proline

The proline content of the high protein soy feed ingredients ranged from 1.75 to 4.47 percent (Table 4.21). Only 87% of the samples in the next step validation had predicted values that deviated less than 2.5% from their high performance liquid chromatography determined values. (Table 4.26).

Table 4.19. Total lysine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	2.32 ± 0.270	2.05	2.59	1	2.26	2.26	2.26
Mechanical SBM	16	2.98 ± 0.033	2.70	3.15	17	2.91 ± 0.027	2.69	3.04
Solvent SBM	91	3.02 ± 0.016	2.60	3.22	91	3.02 ± 0.006	2.84	3.12
High protein soy[1]	13	3.31 ± 0.081	2.94	4.15	13	3.39 ± 0.106	3.03	4.25
Soy protein concentrate	9	4.26 ± 0.076	3.69	4.49	9	4.36 ± 0.019	4.27	4.45
Soy protein isolate	3	5.61 ± 0.038	5.57	5.69	3	5.60 ± 0.039	5.55	5.68
Total	134	3.17 ± 0.046	2.05	5.69	134	3.18 ± 0.046	2.26	5.68

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	17	2.93 ± 0.036	2.69	3.14	16	2.95 ± 0.025	2.76	3.15
Solvent SBM	90	3.00 ± 0.015	2.60	3.22	92	3.03 ± 0.007	2.82	3.22
Total	107	2.99 ± 0.014	2.60	3.22	108	3.02 ± 0.008	2.76	3.22

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.20. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for lysine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.93	0.99	0.18	0.74
RMSEE/RMSEP[1]	0.15	0.05	0.13	0.04
RPD[2]	3.67	9.64	1.11	1.99
Bias[3]		0.001		0.008
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		9		1
97.50 - 99.99		40		58
100.00 – 102.49		44		41
102.50 – 104.99		7		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		84		99
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.21. Total methionine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	0.52 ± 0.005	0.51	0.52	1	0.56	0.56	0.56
Mechanical SBM	15	0.64 ± 0.008	0.59	0.68	19	0.62 ± 0.004	0.57	0.65
Solvent SBM	92	0.65 ± 0.004	0.57	0.74	90	0.64 ± 0.002	0.59	0.68
High protein soy[1]	12	0.76 ± 0.023	0.67	0.97	14	0.73 ± 0.014	0.65	0.87
Soy protein concentrate	9	0.94 ± 0.009	0.90	0.99	9	0.91 ± 0.013	0.82	0.95
Soy protein isolate	4	1.14 ± 0.022	1.08	1.18	2	1.14 ± 0.025	1.11	1.16
Total	134	0.69 ± 0.010	0.51	1.18	135	0.67 ± 0.008	0.56	1.16

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	19	0.63 ± 0.008	0.57	0.68	15	0.63 ± 0.004	0.60	0.65
Solvent SBM	89	0.64 ± 0.004	0.57	0.74	93	0.65 ± 0.002	0.61	0.69
Total	108	0.64 ± 0.004	0.57	0.74	108	0.64 ± 0.002	0.60	0.69

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.22. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for methionine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.92	0.98	0.17	0.55
RMSEE/RMSEP[1]	0.03	0.01	0.03	0.01
RPD[2]	3.55	7.37	1.10	1.49
Bias[3]		-0.001		-0.001
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		10		2
97.50 - 99.99		39		45
100.00 – 102.49		39		46
102.50 – 104.99		13		6
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		78		92
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.23. Total phenylalanine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	1.98 ± 0.040	1.94	2.02	1	2.03 ±	2.03	2.03
Mechanical SBM	18	2.43 ± 0.025	2.26	2.60	16	2.40 ± 0.018	2.24	2.52
Solvent SBM	89	2.46 ± 0.013	2.22	2.80	93	2.44 ± 0.005	2.31	2.56
High protein soy[1]	12	3.01 ± 0.091	2.70	3.84	14	2.84 ± 0.056	2.57	3.41
Soy protein concentrate	9	3.50 ± 0.054	3.23	3.77	9	3.49 ± 0.018	3.42	3.59
Soy protein isolate	4	4.75 ± 0.028	4.71	4.83	2	4.67 ± 0.005	4.66	4.67
Total	134	2.63 ± 0.043	1.94	4.83	135	2.58 ± 0.034	2.03	4.67

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	19	2.40 ± 0.025	2.26	2.60	15	2.42 ± 0.018	2.24	2.52
Solvent SBM	89	2.45 ± 0.012	2.22	2.80	93	2.45 ± 0.006	2.31	2.67
Total	108	2.44 ± 0.011	2.22	2.80	108	2.45 ± 0.006	2.24	2.67

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.24. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for phenylalanine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.96	0.98	0.23	0.75
RMSEE/RMSEP[1]	0.10	0.05	0.10	0.03
RPD[2]	5.00	8.14	1.14	2.01
Bias[3]		0.002		-0.001
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		7		0
97.50 - 99.99		51		50
100.00 – 102.49		31		50
102.50 – 104.99		10		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		82		100
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.25. Total proline content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	1.84 ± 0.090	1.75	1.93	1	1.86	1.86	1.86
Mechanical SBM	13	2.31 ± 0.029	2.11	2.45	21	2.32 ± 0.025	2.02	2.45
Solvent SBM	86	2.38 ± 0.013	2.08	2.68	81	2.36 ± 0.008	2.23	2.66
High protein soy[1]	13	2.75 ± 0.053	2.41	3.19	13	2.84 ± 0.095	2.48	3.61
Soy protein concentrate	9	3.33 ± 0.088	3.01	3.65	9	3.27 ± 0.051	2.91	3.48
Soy protein isolate	4	4.36 ± 0.077	4.14	4.47	2	4.21 ± 0.025	4.18	4.23
Total	127	2.53 ± 0.040	1.75	4.47	127	2.49 ± 0.033	1.86	4.23

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	11	2.29 ± 0.043	2.02	2.45	23	2.33 ± 0.019	2.10	2.43
Solvent SBM	88	2.37 ± 0.012	2.08	2.68	76	2.37 ± 0.007	2.28	2.57
Total	99	2.36 ± 0.012	2.02	2.68	99	2.36 ± 0.007	2.10	2.57

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.26. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for proline for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.92	0.98	0.22	0.72
RMSEE/RMSEP[1]	0.13	0.05	0.11	0.04
RPD[2]	3.57	7.20	1.13	1.90
Bias[3]		0.002		0.002
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		14		9
97.50 - 99.99		37		42
100.00 – 102.49		39		44
102.50 – 104.99		10		4
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		76		87
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Serine

The serine content of the high protein soy feed ingredients ranged from 1.67 to 3.72 percent (Table 4.27). Only 82% of the samples in the next step validation had predicted values that deviated less than 2.5% from their high performance liquid chromatography determined values. (Table 4.28).

Threonine

The threonine calibration curve containing all the high protein dietary ingredient categories contained 134 samples while the next step threonine calibration curve contained 108 samples (Table 4.29). For the next step validation, 100% of the validation samples had NIRS predicted values that deviated 2.5% or less from their laboratory determined value of total threonine (Table 4.30).

Tryptophan

The tryptophan content of the high protein soy feed ingredients ranged from 0.43 to 1.12 percent (Table 4.31), and this aligns with tryptophan being the least abundant amino acid in soybeans. The full and next step NIRS calibration curves for tryptophan had the least accuracy in predicting its content with only 27 and 62 percent of the validation samples having predicted values that deviated 2.5% or less from their laboratory determined values of total tryptophan content, respectively (Table 4.32).

Tyrosine

The tyrosine calibration curve containing all the high protein dietary ingredient categories contained 115 samples while the next step tyrosine calibration curve contained 89 samples (Table 4.33). For the next step validation, 98% of the validation samples had NIRS predicted

Table 4.27. Total serine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	1.70 ± 0.030	1.67	1.73	1	1.76 ±	1.76	1.76
Mechanical SBM	9	2.05 ± 0.038	1.88	2.23	22	1.99 ± 0.015	1.86	2.09
Solvent SBM	76	2.02 ± 0.015	1.74	2.32	72	2.06 ± 0.008	1.95	2.31
High protein soy[1]	14	2.50 ± 0.079	2.16	3.23	11	2.43 ± 0.066	2.10	2.75
Soy protein concentrate	10	2.84 ± 0.055	2.66	3.07	8	2.84 ± 0.046	2.73	3.13
Soy protein isolate	4	3.67 ± 0.035	3.57	3.72	2	3.57 ± 0.010	3.56	3.58
Total	115	2.20 ± 0.039	1.67	3.72	116	2.16 ± 0.029	1.76	3.58

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	13	2.01 ± 0.031	1.86	2.23	18	2.00 ± 0.016	1.88	2.09
Solvent SBM	76	2.03 ± 0.015	1.74	2.32	72	2.04 ± 0.007	1.93	2.22
Total	89	2.03 ± 0.014	1.74	2.32	90	2.03 ± 0.007	1.88	2.22

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.28. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation for serine curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.88	0.98	0.19	0.62
RMSEE/RMSEP[1]	0.15	0.05	0.12	0.04
RPD[2]	2.90	6.94	1.11	1.64
Bias[3]		0.009		0.005
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		14		12
97.50 - 99.99		51		42
100.00 – 102.49		21		40
102.50 – 104.99		15		6
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		72		82
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.29. Total threonine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	1.48 ± 0.030	1.45	1.51	1	1.47	1.47	1.47
Mechanical SBM	17	1.75 ± 0.018	1.63	1.90	17	1.78 ± 0.010	1.69	1.83
Solvent SBM	89	1.83 ± 0.008	1.66	2.03	93	1.83 ± 0.004	1.75	1.91
High protein soy[1]	13	2.16 ± 0.063	1.88	2.76	13	2.14 ± 0.042	1.86	2.45
Soy protein concentrate	9	2.57 ± 0.038	2.31	2.69	9	2.58 ± 0.013	2.53	2.63
Soy protein isolate	4	3.17 ± 0.030	3.11	3.25	2	3.22 ± 0.020	3.20	3.24
Total	134	1.94 ± 0.028	1.45	3.25	135	1.92 ± 0.023	1.47	3.24

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	12	1.76 ± 0.019	1.63	1.90	22	1.77 ± 0.011	1.68	1.88
Solvent SBM	96	1.83 ± 0.007	1.66	2.03	86	1.83 ± 0.004	1.75	1.93
Total	108	1.83 ± 0.007	1.63	2.03	108	1.82 ± 0.005	1.68	1.93

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.30. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for threonine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.95	0.99	0.27	0.81
RMSEE/RMSEP[1]	0.07	0.03	0.07	0.02
RPD[2]	4.67	9.44	1.17	2.28
Bias[3]		0.000		0.000
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		2		0
97.50 - 99.99		48		50
100.00 – 102.49		44		50
102.50 – 104.99		5		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		93		100
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.31. Total tryptophan content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	0.55 ± 0.115	0.43	0.66	1	0.46	0.46	0.46
Mechanical SBM	15	0.64 ± 0.019	0.53	0.78	18	0.63 ± 0.010	0.54	0.69
Solvent SBM	85	0.67 ± 0.007	0.55	0.85	90	0.66 ± 0.004	0.59	0.74
High protein soy[1]	15	0.77 ± 0.025	0.64	1.02	11	0.77 ± 0.017	0.68	0.87
Soy protein concentrate	9	0.86 ± 0.027	0.78	0.97	9	0.90 ± 0.021	0.81	0.97
Soy protein isolate	4	1.01 ± 0.046	0.90	1.12	2	0.99 ± 0.078	0.92	1.07
Total	130	0.70 ± 0.009	0.43	1.12	131	0.69 ± 0.008	0.46	1.07

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	19	0.65 ± 0.015	0.53	0.78	11	0.63 ± 0.009	0.57	0.66
Solvent SBM	83	0.67 ± 0.007	0.55	0.85	92	0.66 ± 0.003	0.59	0.77
Total	102	0.66 ± 0.007	0.53	0.85	103	0.66 ± 0.003	0.57	0.77

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.32. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for tryptophan for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.65	0.80	0.21	0.65
RMSEE/RMSEP[1]	0.07	0.04	0.06	0.02
RPD[2]	1.68	2.28	1.12	1.69
Bias[3]		-0.007		0.001
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		6		0
92.50 - 94.99		12		4
95.00 - 97.49		11		13
97.50 - 99.99		7		31
100.00 – 102.49		21		28
102.50 – 104.99		12		14
105.00 – 107.49		12		5
107.50 – 110.00		13		0
> 110.00		5		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		27		62
< ± 5.0% deviation from determined value		50		90
< ± 7.5% deviation from determined value		76		100
< ± 10% deviation from determined value		95		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.33. Total tyrosine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	1	1.35	1.35	1.35	2	1.51 ± 0.045	1.46	1.55
Mechanical SBM	13	1.67 ± 0.034	1.47	1.82	18	1.68 ± 0.014	1.53	1.75
Solvent SBM	74	1.71 ± 0.013	1.29	1.95	74	1.72 ± 0.007	1.58	1.89
High protein soy[1]	12	2.02 ± 0.036	1.85	2.32	13	2.07 ± 0.044	1.86	2.50
Soy protein concentrate	11	2.31 ± 0.040	2.03	2.47	7	2.32 ± 0.029	2.27	2.49
Soy protein isolate	4	3.18 ± 0.113	2.90	3.41	2	3.18 ± 0.025	3.15	3.20
Total	115	1.84 ± 0.032	1.29	3.41	116	1.81 ± 0.025	1.46	3.20

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	12	1.65 ± 0.037	1.47	1.82	19	1.69 ± 0.011	1.57	1.75
Solvent SBM	77	1.71 ± 0.013	1.29	1.95	71	1.72 ± 0.006	1.58	1.83
Total	89	1.70 ± 0.013	1.29	1.95	90	1.71 ± 0.005	1.57	1.83

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.34. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for tyrosine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.91	0.98	0.17	0.78
RMSEE/RMSEP[1]	0.11	0.04	0.11	0.02
RPD[2]	3.25	6.50	1.10	2.15
Bias[3]		0.001		0.004
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		17		1
97.50 - 99.99		33		58
100.00 – 102.49		33		40
102.50 – 104.99		17		1
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		66		98
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

values that deviated less than 2.5% from their laboratory determined value of total tyrosine content (Table 4.34).

Valine

The valine calibration curve containing all the high protein dietary ingredient categories contained 116 samples while the next step valine calibration curve contained 89 samples (Table 4.35). For the next step validation, 100% of the validation samples had NIRS predicted values that deviated 2.5% or less from their laboratory determined value of total valine (Table 4.36).

Gross energy

The gross energy content of the high protein soybean ingredients ranged from 3,759 to 5,367 kcal/kg (Table 4.37). For both the full NIRS validation containing all the high protein sample categories and the next step validation containing just solvent and mechanically extracted SBM, 100 percent of the of the validation samples had predicted values that deviated less than 2.5% from their bomb calorimeter determined values (Table 4.38).

Table 4.35. Total valine content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	%	%	%	n	%	%	%
Full fat SBM	2	1.90 ± 0.075	1.82	1.97	1	1.93	1.93	1.93
Mechanical SBM	15	2.36 ± 0.026	2.20	2.50	16	2.30 ± 0.031	2.03	2.45
Solvent SBM	79	2.40 ± 0.011	2.16	2.57	71	2.37 ± 0.006	2.27	2.59
High protein soy[1]	9	2.91 ± 0.109	2.51	3.62	16	2.74 ± 0.048	2.46	3.14
Soy protein concentrate	7	3.39 ± 0.071	2.99	3.51	11	3.46 ± 0.020	3.35	3.55
Soy protein isolate	4	4.62 ± 0.070	4.48	4.81	2	4.51 ± 0.095	4.41	4.60
Total	116	2.56 ± 0.046	1.82	4.81	117	2.55 ± 0.040	1.93	4.60

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	%	%	%	n	%	%	%
Mechanical SBM	18	2.32 ± 0.032	2.03	2.50	13	2.35 ± 0.023	2.18	2.45
Solvent SBM	71	2.39 ± 0.012	2.16	2.59	77	2.39 ± 0.006	2.27	2.51
Total	89	2.37 ± 0.012	2.03	2.59	90	2.38 ± 0.006	2.18	2.51

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.36. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for valine for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.95	0.99	0.26	0.79
RMSEE/RMSEP[1]	0.11	0.05	0.10	0.02
RPD[2]	4.62	9.91	1.16	2.19
Bias[3]		-0.013		0.001
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		3		0
97.50 - 99.99		38		51
100.00 – 102.49		44		49
102.50 – 104.99		15		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		82		100
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted values.
[4]Near infrared reflectance spectroscopy predicted value divided by the high-performance liquid chromatography determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the high-performance liquid chromatography determined value.

Table 4.37. Gross energy content of the soy protein ingredients used in the near infrared reflectance spectroscopy calibration and validation curves for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples (SBM).

Sample type	Full calibration curve				Full validation curve			
	Sample #	Mean ± SEM	Minimum	Maximum	Sample #	Mean ± SEM	Minimum	Maximum
	n	kcal/kg	kcal/kg	kcal/kg	n	kcal/kg	kcal/kg	kcal/kg
Full fat SBM	1	5,367	5,367	5,367	0			
Mechanical SBM	14	4,707 ± 21	4,557	4,853	14	4,744 ± 20	4,639	4,896
Solvent SBM	74	4,141 ± 8	3,759	4,322	73	4,160 ± 9	4,054	4,704
High protein soy[1]	7	4,509 ± 73	4,174	4,802	8	4,479 ± 24	4,411	4,590
Soy protein concentrate	8	4,536 ± 37	4,371	4,716	10	4,557 ± 17	4,429	4,614
Soy protein isolate	3	5,015 ± 17	4,987	5,045	3	4,993 ± 33	4,945	5,057
Total	107	4,304 ± 27	3,759	5,367	108	4,319 ± 25	4,054	5,057

Sample type	Next step calibration curve				Next step validation curve			
	Sample #	Mean	Minimum	Maximum	Sample #	Mean	Minimum	Maximum
	n	kcal/kg	kcal/kg	kcal/kg	n	kcal/kg	kcal/kg	kcal/kg
Mechanical SBM	14	4,707 ± 21	4,557	4,853	14	4,744 ± 20	4,639	4,896
Solvent SBM	73	4,146 ± 6	3,991	4,322	73	4,160 ± 9	4,054	4,704
Total	87	4,237 ± 23	3,991	4,853	87	4,254 ± 25	4,054	4,896

[1]Contains both fermented and/or enzymatically treated soy protein products.

Table 4.38. Mathematical evaluation of the near infrared reflectance spectroscopy calibration and validation curves for gross energy for the full curves containing all soy protein samples or the next step curves containing only mechanical and solvent extracted soybean meal samples.

Parameter	Full curves		Next step curves	
	Calibration	Validation	Calibration	Validation
R-squared	0.95	0.98	0.98	0.99
RMSEE/RMSEP[1]	64.70	37.00	34.70	22.80
RPD[2]	4.43	7.10	6.44	10.30
Bias[3]		8.020		5.410
Predicted value/determined value[4]		%[5]		%[5]
<90.00		0		0
90.00 -92.49		0		0
92.50 - 94.99		0		0
95.00 - 97.49		0		0
97.50 - 99.99		56		56
100.00 – 102.49		44		44
102.50 – 104.99		0		0
105.00 – 107.49		0		0
107.50 – 110.00		0		0
> 110.00		0		0
Absolute percentage deviation[6]				
< ± 2.5% deviation from determined value		100		100
< ± 5.0% deviation from determined value		100		100
< ± 7.5% deviation from determined value		100		100
< ± 10% deviation from determined value		100		100

[1]Root mean square error of the estimate (RMSEE) for the calibration and root mean square error of the prediction (RMSEP) for the validation.
[2]Residaul predictive deviation.
[3]Bias is the systematically averaged deviation between the data set and the predicted value.
[4]Near infrared reflectance spectroscopy predicted value divided by the bomb calorimeter determined value then multiplied by 100.
[5]Percent of the validation samples.
[6]The percent the near infrared reflectance spectroscopy predicted value deviates from the bomb calorimeter determined value.

CHAPTER 5

DISCUSSION

The calibration curves constructed utilizing all high protein soy products were able to predict the amino acid content and gross energy of all the validation samples with less than 5% deviation from their laboratory determined values except for the calibration curves for tryptophan, cysteine, and glutamic acid (Table 5.1). However, of these exceptions the calibration curve for glutamic acid predicted 99% of the validation samples with less than 5% deviation from their laboratory determined high performance liquid chromatography values. Because the poultry industry in the United States utilizes almost exclusively solvent extracted soybean meal and some mechanically extracted soybean meal for organic/all natural production systems, the decision was made to do next step calibration curves from the original curves. In the next step curves the full fat soybean meal, fermented and enzymatically treated soybean meal as well as the soy protein concentrate and isolate samples were excluded as these products are rarely used in commercial poultry diets. The next step calibration curves for predicting amino acid and gross energy content all had a greater accuracy than the original curves (Table 5.2) and the deviations of the NIRS predicted values versus the laboratory obtained values are within the natural variation seen when determining gross energy by bomb calorimetry and amino acid content by high performance liquid chromatography of duplicate samples of high protein soybean products.

Tryptophan is a traditionally problematic amino acid for detection and determination of its concentration by high performance liquid chromatography, and because the NIRS calibration for tryptophan is built from the correlation of the NIRS scans of the samples with their determined high performance liquid chromatography values, the NIRS calibration will reflect the

Table 5.1. Summary of near infrared reflective spectroscopy calibration and validation curves to predict total amino acid and nitrogen corrected true metabolizable energy (TME_N) in high protein soybean feed ingredients.

Component	Calibration curve sample #	Validation curve sample #[1]	The % of the validation samples that deviated less than 5% from their bioassay determined value	The % of the validation samples that deviated less than 2.5% from their bioassay determined value
Alanine	134	135	100	73
Arginine	134	135	100	87
Aspartic acid	134	135	100	76
Cysteine	134	135	77	48
Glutamic acid	134	135	99	67
Glycine	129	130	100	91
Histidine	134	135	100	70
Isoleucine	116	117	100	68
Leucine	134	135	100	84
Lysine	134	134	100	84
Methionine	134	135	100	78
Phenylalanine	134	135	100	82
Proline	127	127	100	76
Serine	115	116	100	72
Threonine	134	135	100	93
Tryptophan	130	131	50	27
Tyrosine	115	116	100	66
Valine	116	117	100	82
Gross energy	87	87	100	100

[1]Note that the validation samples are new soybean meal samples that are not included in the calibration.

Table 5.2. Summary of near infrared reflective spectroscopy calibration and validation curves to predict total amino acid and nitrogen corrected true metabolizable energy (TME_N) in solvent and mechanically extracted soybean meal samples.

Component	Calibration curve sample #	Validation curve sample #[1]	The % of the validation samples that deviated less than 5% from their bioassay determined value	The % of the validation samples that deviated less than 2.5% from their bioassay determined value
Alanine	108	108	100	99
Arginine	108	108	100	99
Aspartic acid	108	108	100	100
Cysteine	108	108	100	81
Glutamic acid	108	108	100	100
Glycine	103	103	100	100
Histidine	108	108	100	94
Isoleucine	89	90	100	98
Leucine	108	108	100	100
Lysine	107	108	100	99
Methionine	108	108	100	92
Phenylalanine	108	108	100	100
Proline	100	101	100	87
Serine	89	90	100	82
Threonine	108	108	100	100
Tryptophan	102	103	90	62
Tyrosine	89	90	100	98
Valine	89	90	100	100
Gross energy	107	108	100	100

[1]Note that the validation samples are new soybean meal samples that are not included in the calibration.

variation (error) associated with determining tryptophan by high performance liquid chromatography. The concentration in tryptophan is typically the lowest in all plant and animal feed sources (NRC 1994; Batal and Dale 2016), and soybean meal products are not an exception. Thus, when there is less to detect even a small variation leads to a large percentage error. For example, duplicate determinations of tryptophan in a sample of full fat soybean meal by high performance liquid chromatography resulted in values of 0.43% and 0.48% which gives a variation of over 10%. In contrast, in the same sample, the arginine concentration values were 2.57 and 2.44 which is greater absolute difference, but the percent variation was cut almost by half.

While free amino acids in feed ingredients can be directly measured by high performance liquid chromatography, to get the total content of individual amino acids in feed ingredients the ingredient samples must be acid hydrolyzed to liberate all the bound amino acids to a free form. During this process tryptophan can be converted to other metabolites, and thus for tryptophan determination a base hydrolysis is used, but even with this method the recovery and stability of tryptophan can be variable across different samples.

The lower prediction accuracy in the NIRS determination of cysteine also results from the variability in its determination by high performance liquid chromatography. After tryptophan, cysteine has the second lowest concentration in high protein soybean products (NRC 1994; Batal and Dale 2016). In addition, the acid hydrolysis of feed ingredients to free bound amino acids can lead to some degradation of cysteine and methionine. To prevent this, performic acid oxidation is performed prior to acid hydrolysis which converts cysteine to cysteic acid and methionine to methionine sulfone which are both stable during acid hydrolysis. But the

addition of this procedure adds more variability in determining the total content of cysteine and methionine across samples and this variability become incorporated into the NIRS calibration.

Statistical evaluation

The current research agrees with (Cope 2021; Davis 2023) that traditional statistical evaluations of calibration and validation curves using R^2 and RPD values do not provide a good enough evaluation of the predictability of a calibration curve for a given nutrient when the predicted value is used for poultry feed formulation. For example, in the current research the full calibration curve for leucine had a R^2 value of 0.96 and an RPD value of 4.72 and the corresponding validation curve had an R^2 value of 0.99 and an RPD value of 10.03 which are excellent. But only 84% of the validation samples had values that deviated less than 2.5% from their high performance liquid chromatography values. In contrast, the next step calibration curve for leucine had a R^2 value of 0.28 and an RPD value of 1.18 and the corresponding validation curve had an R^2 value of 0.76 and an RPD value of 2.07 which are very poor relative to the values obtained for the full curve and would indicate the calibration curve would have poor predictability capacity. Yet, 100% of the validation samples had values that deviated less than 2.5% from their high performance liquid chromatography values making this practical evaluation more valuable. For a poultry nutritionist that must formulate diets with a safety margin to make sure if the analyzed value is wrong that a nutrient deficiency and poor bird performance does not occur, this is especially true. Regretfully most NIRS research publications do not provide the NIRS predicted and laboratory determined values or the percentage of samples that deviate a given amount from their determined values as the current research does.

Gross energy

In the current research all of the validation samples utilizing the NIRS calibration for all high protein soybean products or the NIRS calibration for just solvent and mechanically expelled soybean meal had values that deviated less than 2.5% from their bomb calorimeter determined values which matches the 0 to 2.5% variation seen when duplicate determinations of the same sample are determined using a bomb calorimeter. Other researchers have also had success in building calibrations for gross energy. Xiccato et al., (1999) were able to build a calibration that had an R^2 of 0.90 for that predicted gross energy in complete rabbit diets. Aufrere et al., (1996) were able to build calibrations that predicted the gross energy of complete diets for swine and ruminants with R^2 values ranging from 0.85 - 0.92. Noel et al., (2022) built a NIRS calibration for gross energy that's prediction had an R^2 value of 0.81 for mixed swine diets, and 0.91 for ruminant diets. It is useful to compare these previous studies as they used mixed complete feeds as opposed to a singular dietary ingredient utilized in the current research. While not identical, the current research included multiple high protein soy products in the same curve as opposed to just one and therefore it would not be unreasonable to assume our R^2 for the gross energy calibration to be like theirs. The R^2 value in the current research for the calibration containing all high protein soy products was 0.98. Overall, these previous and current results suggest that NIRS calibrations to predict gross energy can be constructed and utilized to accurately predict gross energy in individual feed ingredients and completed feeds.

Total amino acid content

Much of the previous research building NIRS calibrations to predict amino acid content in soybean products is not directly applicable to feed formulation applications for monogastric animals because the soybean products utilized were not heat treated to inactivate antinutritional

factors or because the research only reported statistical analyses like R^2 values and did not provide a practical statical analysis to indicate how much the NIRS predicted values could deviate from the determined values. Pazdernik et al., (1997) in proof-of-concept research, used 116 samples of ground, dried, and defatted whole soybeans that were not heat treated to construct and validate NIRS curves to determine individual amino acid contents. The results indicated that NIRS did have potential in successfully predicting amino acid content in soybean meal products. Subsequently, Fontaine et al., (2001) had success predicting all essential amino acids by NIRS in full fat and defatted soybean meal with the R^2 values of their calibrations ranging from 0.95 to 0.98 except for cysteine and methionine which were both 0.84. Samples were ground and scanned on as is basis like in the current research, but the current research expands the scope of this previous research by including all amino acids, gross energy, and additional high protein soy protein products into the calibration models. Recently Shi et al. (2022) built NIRS calibrations to predict amino acid content in whole, ground, and defatted soybeans, but the calibrations were made for each type of product. This study had success building models with good predictive ability for arginine, glutamic acid, proline, serine, and valine, moderate success with aspartic acid, isoleucine, leucine, and phenylalanine, and poor results for alanine, glycine, histidine, lysine, threonine, and tryptophan. Like the current research, Kovalenko et al., (2006) constructed NIRS calibration curves to predict total amino acid content of the individual amino acids and their results indicated good predictability for all aminos acids except tryptophan and cysteine. However, their research focused exclusively on unground, not heat-treated, whole soybeans which would not be utilized in monogastric diets.

Product development and further research

For feed manufacturers developing new high protein soybean products, the full calibration curves that contained all products would be suitable for the prediction of total amino acid and gross energy content. New product development in the high protein soy category is likely going to be fermented and enzymatically digested soybean meal products that will have a crude protein content above 50%, but lower than the 70% associated with well-established commercial soy protein isolates products. The developed full product NIRS calibrations from this research should very accurately predict the gross energy of these products and indicate the individual amino acid content within 5% of the actual value except for tryptophan and cysteine. This initial assessment of the product using the developed NIRS calibration curves would allow the manufacturers to identify products to pursue for subsequent further testing and identify potential manufacturing issues leading to the degradation of amino acids.

The developed high protein soybean calibration curves in the current research need to be improved by the inclusion of more full fat soybean meal samples as only 3 samples were available for inclusion in the current research. Unlike the soy isolates that have strict regulations making their ultimate composition relatively invariable, considerable variation in the composition of full fat soybean meal does exist based on differences in soybean cultivars utilized, agronomic conditions the soybeans were grown under, whether the soybeans were dehulled or not before making the meal, and the process utilized to inactivate the antinutritional factors found in soybeans. Because the commercial production of full fat soybean meal is currently almost nonextant in the United States, arrangements had to be made to import samples from other locations while meeting import regulations which limited the availability of full fat soybean meal samples. It is likely the use of full fat soybean meal will increase in the United

States once there is more public sentiment or regulatory pressure to increase sustainable animal production. The production and use of full fat soybean meal in poultry production is more widespread in other parts of the world, to avoid the scenario of using solvent extracted soybean meal in poultry diets and then adding vegetable oil (often soy oil) to the diets to meet the metabolizable energy requirement of the birds being fed the diets.

In addition to adding more full fat soybean meal samples to the current NIRS calibration and validation curves for high protein soybean meal products, future follow-up research needs to focus on the development of NIRS calibration curves that predict the nitrogen corrected true metabolizable energy and digestible amino acid content of these products. For poultry nutritionists formulating diets on total energy and the total content of each amino acid, involves them making a conservative guess on what is digestible and available to the bird, and this typically leads to expensive over formulation with diets containing nutrients well above requirements. Thus, having the potential to determine the digestible energy and amino acid content of high protein soybean meal products in real-time on an ingredient like soybean meal that typically constitutes 15-40% of any poultry diet would save on diet costs. For feed manufactures producing soybean products having the capacity to quickly measure by NIRS gross energy, nitrogen corrected true metabolizable energy and total and digestible amino acid content of individual amino acids would allow them to quickly identify processing procedures that were decreasing digestible energy and amino acid content and to blend different batches of a product like solvent extracted soybean meal to meet minimum digestible energy and amino acid benchmarks.

Summary

In summary, the NIRS calibration models created in this research for the prediction of gross energy content and the total content of individual amino acids in all high protein soybean products are accurate, based on how close NIRS predicted values are to their respective reference values, especially for gross energy, but do not provide the accuracy for amino acid determination seen with high performance liquid chromatography. However, if the NIRS calibrations are restricted to soybean meal in which the oil has been extracted by solvent or mechanical extraction which are the two most widely utilized products in the high soy protein category, the NIRS calibrations predict amino acid concentration with an accuracy that is equivalent to determinations made by high performance liquid chromatography. Additionally, the NIRS determinations can be done in real time on the soybean meals delivered to a feed mill and that will be utilized in the diets being made.

REFERENCES

Abafe, E. 2023. Soybeans for global nutrition: A numbers story. Sustainable Nutrition Initiative. Accessed Mar. 2025., [https://sustainablenutritioninitiative.com/soybeans-for-global-nutrition-a-numbers-story/]

Agelet, L. E., and C. R. Hurburgh. 2010. A Tutorial on Near Infrared Spectroscopy and its 10.1080/10408347.2010.515468

Agyenim-Boateng, K.G., Zhang[1], S., Zhang[2]. S., Khattak, A.N., Shaibu, A., Abdelghany, A.M., Qi, J., Azam, M., Ma, C., Feng, Y., Feng, H., Liu, Y., Li, J., Li, B., and Sun, J. 2023. The nutritional composition of the vegetable soybean (maodou) and its potential in combatting malnutrition. Frontier Nutrition 9. doi https://doi.org/10.3389/fnut.2022.1034115

Association of Official Analytical Chemists 2006. Official method of analysis. Association of Official Analytical Chemists. 18th edition

Aranibar, C.D. 2018. Effect of different feed programs on broiler breeder behavior and reproductive performance. PhD. Diss. University of Georgia, Athens.

Asselin, M., and Sandorfy, C. 1971. A low temperature study of self-associated alcohols in the near-infrared. Journal of Molecular Structure 8:145-158. doi https://doi.org/10.1016/0022-2860(71)80050-9

Aufrere, J., Graviou, D., Demarquilly, C., Perez, J.M., and Andrieu, J. 1996. Near infrared reflectance spectroscopy to predict energy value of compound feeds for swine and ruminants. Animal Feed Science Technology 62:77-90

Bakalli, R. I., Pesti, G. M., and Etheridge R. D. 2000. Comparison of a commercial Near-Infrared Reflectance Spectroscope and standard chemical assay procedures for analyzing feed ingredients: Influence of grinding methods. The Journal of applied poultry research. 9:204-213. doi 10.1093/japr/9.2.204

Batal, A., and Dale, N. 2016. Feedstuffs Ingredient Analysis Table: 2016 edition. University of Georgia, GA.

Berrocoso, J.D., Rojas, O.J., Liu, Y., Shoulders, J., González-Vega, J.C., and Stein H.H. 2015. Energy concentration and amino acid digestibility in high-protein canola meal, conventional canola meal, and soybean meal fed to growing pigs. Journal of Animal Science. 93:2208-2217. doi https://doi.org/10.2527/jas.2014-8528

Bohn, T., Cuhra, M., Traavik, T.,Sanden, M., Fagan, J., and Primicerio, R. 2014. Compositional difference in soybeans on the market: Glyphosate accumulates in Roundup Ready GM soybeans. Food Chemistry 153:207-215. doi https://doi.org/10.1016/j.foodchem.2013.12.054

Büning-Pfaue, H. 2003. Analysis of water in food by near infrared spectroscopy. Food Chemistry 82:107-115. doi https://doi.org/10.1016/S0308-8146(02)00583-6

Caldwell, R. A. 1992. Effect of calcium and phytic acid on the activation of trypsinogen and the stability of trypsin. Journal of Agricultural and Food Chemistry 40:43-46.

Caprita, A., and Caprita, R. 2010. Modification of the soluble protein content of heat-processed soybean flour. Notulae Bontanicae Horti Agrobotanci Cluj-Napoca 38. doi https://doi.org/10.15835/nbha3824747

Casaubon-Huguenin, M. T., Ávila-González, E., Vazquez-Pelaez, C., Trigo, F., Lascurain, R., and Zenteno, E. 2004. The effect of raw full-fat soybean and its lectin on the nutrition and pigmentation of broilers. Journal of Agricultural and Food Chemistry 52:5702-5708. doi 10.1021/jf0307604

Chang, C.J., Tanksley Jr, T.D., Knabe, D.A., and Zebrowska, T. 1987. Effects of different heat treatments during processing on nutrient digestibility of soybean meal in growing swine. Journal of Animal Science 65:1273-1282. doi 10.2527/jas1987.6551273x

Croston, C.B., Smith, A.K., and Cowan, J.C. 1955. Measurement of urease activity in soybean oil meal. The Journal of the American Oil Chemists' Society 32:279-282. doi https://doi.org/10.1007/BF02634460

Conzen, J.-P. 2014. Multivariate Calibration: A practical guide for developing methods in the quantitave analytical chemistry (3rd ed.). Bruker

Cope, M. 2021. Determination of Nitrogen Corrected True Metabolizable Energy Content of Feed Ingredients for Poultry Diets by Near Infrared Reflectance Spectroscopy. PhD Diss. University of Georgia, Athens.

Davies, A. M., & Fearn, T. 2006. Back to basics: calibration statistics. Spectroscopy Europe 18:31-32.

Davis, M.L. 2023. Using near infrared reflectance spectroscopy as an alternative to animal bioassays for determining digestible amino acid content of dietary ingredients used in poultry diets. PhD. Diss. University of Georgia, Athens.

Depieri, R. A., and Panizzi., A. R. 2011. Duration of feeding and superficial and in-depth damage to soybean seed by selected species of stink bugs (Heteroptera: Pentatomidae). Neotropical Entomology 40:197-203. doi 10.1590/s1519-566x2011000200007

Dornbos, D.L., and Mullen, R.E. 1992. Soybean seed protein and oil contents and fatty-acid composition adjustments by drought and temperature. Journal of the American Oil Chemists' Society 69: 228-231.

Fearn, T. 2005. Chemometrics: an enabling tool for NIR. Sage Journals, NIR news.

Fontaine, J., Hörr, J., and Schirmer, B. 2001. Near-Infrared Reflectance spectroscopy enables the fast and accurate prediction of the essential amino acid contents in soy, rapeseed meal, sunflower meal, peas, fishmeal, meat meal products, and poultry meal. Journal of Agricultural and Food Chemistry 49:57-66. doi 10.1021/jf000946s

Gibson, L.R., and Mullen, R.E. 1996. Influence of day and night temperature on soybean seed yield. Crop Science 36:98-104. doi https://doi.org/10.2135/cropsci1996.0011183X003600010018x

Givens, D.I., De Boever, J.L., and Deaville, E.R., 1997. The principles, practices, and some future applications of near infrared spectroscopy for predicting the nutritive value of foods for animals and humans. Nutrition Research Reviews 10:83-114. doi 10.1079/NRR19970006

Gnanasambandam, R.., and Proctor, A. 1999. Preparation of soy hull pectin. Food Chemistry 65:461-467. doi https://doi.org/10.1016/S0308-8146(98)00197-6

Guo, C., Liu, X., Chen, L., Cai, Y., Yao, W., Yuan, S., Wu, C., Han, T., Sun, S., and Hou, W. 2020. Elevated methionine content in soybean seed by overexpressing maize β-zein protein. Oil Crop Science 5:11-16. doi https://doi.org/10.1016/j.ocsci.2020.03.004

Hamza, M., Basit, A., Shehzadi, I., Tufail, U., Hassan, A., Hussain, T., Siddique, M., and Hayat. H. 2024. Global impact of soybean production: A review. Asian Journal of Biochemistry, Genetics and Molecular Biology 16:12-20. doi 10.9734/ajbgmb/2024/v16i2357

Hansen, W. G., Wiedemann, S. C. C., Snieder, M., and Wortel, V. A. L. 2000. Tolerance of near infrared calibrations to temperature variations; A practical evaluation. Journal of Near Infrared Spectroscopy 8:125-132.

Hymowitz, T., Collins, F. I., Panczner, J., and Walker, W. M. 1972. Relationship between the content of oil, protein, and sugar in soybean seed. Agronomy Journal 64:613-616. doi 10.2134/agronj1972.00021962006400050019x

Kahraman, A. 2012. Nutritional value and foliar fertilization in soybean. Journal of Elementology 22. 55-56. doi 10.5601/jelem.2016.21.1.1106

Karr-Lilienthal, L. K., Kadzere, C.T., Grieshop, C. M., and Fahey Jr, G. C. 2005. Chemical and nutritional properties of soybean carbohydrates as related to nonruminants: A review. Livestock Production Science 97:1-12. doi https://doi.org/10.1016/j.livprodsci.2005.01.015

Kovalenko, I. V., Rippke, G. R., and Hurburgh, C. R. 2006. Determination of amino acid composition of soybeans (Glycine max) by Near-Infrared Spectroscopy. Journal of Agricultural and Food Chemistry 54:3485-3491. doi 10.1021/jf052570u

Kowalski, B. R. 1975. Chemometrics: Views and Propositions. Journal of Chemical Information and Computer Sciences 15:201-203. doi 10.1021/ci60004a002

Lambo, M.T., Ma, H., Zhang, H., Song, P., Mao, H., Cui, G., Dai, B., Li, Y., and Zhang, Y., 2024. Mechanism of action, benefits, and research gap in fermented soybean meal utilization as a high-quality protein source for livestock and poultry. Animal Nutrition 16:130-146

Leeson, S., and Summers, J.D. 2001. Nutrition of the chicken. University Books.

Liener, I. E. 1996. Effects of processing on antinutritional factors in legumes: the soybean case. Archivos latinoamericanos de nutrcion 44S-54S.

Liener, I. E. 1981. Factors affecting the nutritional quality of soya products. Journal of the American Oil Chemists' Society 58:406-415. doi 10.1007/BF02582390

Luypaert, J., Heuerding, S., Vander Heyden, and Y., and Massart, D.L. 2004 The effect of preprocessing methods in reducing interfering variability from near-infrared measurements of creams. Journal of Pharmaceutical and Biomedical Analysis 36:495-503. doi https://doi.org/10.1016/j.jpba.2004.06.023

Maddaiah, V.T., Kurnick, A.A., and Reid, B.L. 1964. Phyic Acid Studies. Experimental Biology and Medicine 115. doi https://doi.org/10.3181/00379727-115-28922

Mayerhofer, T.G., Pahlow, S., Popp, J. 2020. The Bouguer-Beer-Lambert law: Shining light on the obscure. Chem Physical Chemistry 21:2029-2046. doi https://doi.org/10.1002/cphc.202000464

Mayhew, W., Sneller, C., Coker C., Dombek D., and Widick D. 2014. Arkansas Soybean Production Handbook: Ch. 3 – Variety Development, Testing, and Selection: 3: 1-4

Melcion, J. P., & van der Poel, A. F. B. 1993. Process technology and antinutritional factors: principles, adequacy and process optimization. Pages 419-434 in Recent advances of research in antinutritional factors in legume seeds, A.F.B. van der Poel et al. (eds.). EAAP Publ. 70. Wageningen Press

Montanha, G.S., Perez, L.C., Brandao, J.R., de Camargo, R. F., Tavares, T.R., de Almeida, E., and de Carvalho, W.P. 2024. Profile of mineral nutrients and proteins in soybean seeds (Glycine max (L.) Merrill): Insights from 95 varieties cultivated in Brazil. Journal of Food Composition and Analysis 134. doi https://doi.org/10.1016/j.jfca.2024.106536

Murray, I., and Cowe, I. 2004. Sample Preparation. Pages 75-112 in Near-Infrared Spectroscopy in Agriculture.

Næs, T., and Martens, H. 1988. Principal component regression in NIR analysis: Viewpoints, background details and selection of components. Journal of Chemometrics 2:155-167. doi 10.1002/cem.1180020207

National Research Council. 1994. Nutrient requirements of poultry. 9th rev. ed. Natl. Acad. Press, Washington, DC.

Noel, S.J., Jorgensen, H.J.H, and Knudsen, K.E.B. 2022. The use of near-infrared spectroscopy (NIRS) to determine the energy value of individual feedstuffs and mixed diets for pigs. Animal Feed Science and Technology 283. doi https://doi.org/10.1016/j.anifeedsci.2021.115156

Palliyeguru, C., Rose, S. P., and Mackenzie, A. 2011. Effect of trypsin inhibitor activity in soya bean on growth performance, protein digestibility and incidence of sub-clinical necrotic enteritis in broiler chicken flocks. British poultry science 52:359-367. doi 10.1080/00071668.2011.577054

Pan, L., Farouk, M.H., Qin, G., Zhao, Y., and Bao, N. 2018. The influences of soybean agglutinin and functional oligosaccharides on the intestinal tract of monogastric animals. International Journal of Molecular Sciences 19:554. doi 10.3390/ijms19020554

Parsons, C. M., Hashimoto, K., Wedekind, K. J., Han, Y., and Baker, D.H. 1992. Effect of overprocessing on availability of amino acids and energy in soybean meal. Poultry Science 71:133-140. doi https://doi.org/10.3382/ps.0710133

Parsons, C.M., Zhang, Y., and Araba, M. 2000. Nutritional evaluation of soybean meals varying in oligosaccharide content. Poultry Science 79: 1127-1131. doi https://doi.org/10.1093/ps/79.8.1127

Pazdernik, D.L., Killam, A.S., and Orf, J.H. 1997. Analysis of amino and fatty acid composition in soybean seed, using near infrared reflectance spectroscopy. Agronomy Journal 89:679-685.

Pérez-Marín, D., Garrido-Varo, A., and Guerrero. J. E., 2007. Non-linear regression methods in NIRS quantitative analysis. Talanta 72:28-42. doi https://doi.org/10.1016/j.talanta.2006.10.036

Rahman, S. U., McCoy, E., Raza, G., Ali, Z., Mansoor, S., and Amin, I. 2023. Improvement of Soybean; A way forward transition from genetic engineering to new plant breeding technologies. Molecular Biotechnology 65:162-180. doi 10.1007/s12033-022-00456-6

Ramaswamy, K.S., Rosentrater, K. A., and Mosher, G. A. 2016. Does a NIR system provide low-cost alternative to on-farm feed and forage testing? A techno-economic analysis. Pages 1 in 2016 ASABE Annual International Meeting ASABE, St. Joseph, MI.

Reeves, J.B. 1995. Efforts to quantify changes in near-infrared spectra caused by the influence of water, pH, ionic strength, and differences in physical state. Applied Spectroscopy. 49:181-187.

Shi, D., Hand, J., Neufeld, J., Zhao, S., and House, J.D. 2022. Estimation of crude protein and amino acid content in whole, ground, and defatted ground soybeans by different types of near-infrared (NIR) reflectance spectroscopy. Journal of Food Composition and Analysis: 111. doi https://doi.org/10.1016/j.jfca.2022.104601

Sotak-Peper, K.M., Gonzalez-Vega, J.C., and Stein, H.H. 2017. Amino acid digestibility in soybean meal sourced from different regions of the United States and fed to pigs. Journal of Animal Science 2:771-778. doi 10.2527/jas.2016.0443

Tang, H.O., Gao, X.H., Ji, F., Tong, S., and Li, X.J. 2012. Effects of a thermostable phytase on the growth performance and bone mineralization of broilers. Journal of Applied Poultry Research 21: 476-483

Toomer, OT., Oviedo-Rondon, E.O., Ali, M., Joseph, M., Vu, T., Fallen, B., and Main, R. 2024. Full-fat soybeans as an alternative poultry feed ingredient- Feed processing methods and utilization- Review and perspective, Animals 14:2366. doi https://doi.org/10.3390/ani14162366

United States Department of Agriculture, Soybeans and Oil Crops. 2025. Oil crops sector at a glance.

United Soybean Board. 2021. Animal Agriculture. Accessed Mar. 2025 [https://unitedsoybean.org/issue-briefs/animal-agriculture/#:~:text=During%20the%202020-2021%20soybean,Swine%20%E2%80%94%205.95%20million]

US Soy. 2018. What are soybeans used for? Soybean uses and more. Accessed Mar. 2025 [https://ussoy.org/uses-for-soybeans/]

Wijewardana, C., Reddy, K.R., and Bellaloui, N. 2019. Soybean seed physiology, quality, and chemical composition under soil moisture stress

Williams, P. C. 2001. Implementation of near-infrared technology. Near-infrared technology in the agricultural and food industries.

Weiss, M. G., Weber, C. R., Williams, L. F., and Probst, A. H. 1952. Correlation of agronomic characters and temperature with seed compositional characters in soybeans, as influenced by variety and time of planting 1. Agronomy Journal 44:289-297.

Xiccato, G., Trocino, A., Carazzolo, A., Meurens, M., Maertnes, L., and Carabano, R. 1999. Nutritive evaluation and ingredient predicition of compound feeds for rabbits by near-infrared reflectance spectroscopy (NIRS) Animal Feed Science and Technology 77