5-2009

Development and validation of a predictive statistical model to optimize accelerated solvent extraction of isoflavones from edamame soybean [Glycine max (L.) Merrill]

Yu-Ting Hung
University of Tennessee

To the Graduate Council:

I am submitting herewith a thesis written by Yu-Ting Hung entitled "Development and validation of a predictive statistical model to optimize accelerated solvent extraction of isoflavones from edamame soybean [Glycine max (L.) Merrill]." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Plant Sciences.

Carl E. Sams, Major Professor

We have read this thesis and recommend its acceptance:

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a thesis written by Yu-Ting Hung entitled "Development and validation of a predictive Statistical Model to Optimize Accelerated Solvent Extraction of Isoflavones from Edamame Soybean [*Glycine max* (L.) Merrill]." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Plant Sciences.

Carl E. Sams, Major Professor

We have read this thesis
and recommend its acceptance:

Vincent R. Pantalone

Arnold M. Saxton

Dean A. Kopsell

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Development and Validation of a Predictive Statistical Model to Optimize Accelerated Solvent Extraction of Isoflavones from Edamame Soybean [*Glycine max* (L.) Merrill]

A Thesis
Presented for the
Master of Science Degree
The University of Tennessee, Knoxville

Yu-Ting Hung
May 2009

Dedication

I would like to dedicate this to my parents, Cheng-Chang Hung and Li-Mei Tsai. Their faith in me and their unlimited support for me has been steady throughout my entire education.

Acknowledgements

I would like to express my appreciation to all those who have helped me complete this project, in conducting research and in the preparation of this thesis. I thank both of my major professors, Dr. Carl Sams and Dr. Vincent Pantalone, for providing me an opportunity to study in soybean physiology and develop skills that will benefit me throughout my life. Their patience, encouragement and guidance have supported me throughout my entire studies. They have treated me more than a student, and I deeply felt they are my family.

I would also like to thank my committee members. Dr. Arnold Saxton has provided tremendous assistance with statistical analysis on various data and Dr. Dean Kopsell has advised me on HPLC analysis with his enormous knowledge about analytical chemistry.

Members in Dr. Pantalone's soybean team, Deborah Landau-Ellis, Elizabeth Meyer, Amanda Pierce, Catherine Nyinyi, and Benjamin Fallen, thank you all for providing me with a great opportunity learning new techniques in soybean breeding.

Dr. Teresa Johnston, an amazing physical therapist, who has given me the most proper health care during my injury and made me able to continue the research without postponing.

Finally I wish to thank all of my family, my friends, my co-workers, and all the people that have come through the lab, especially Casey Barickman, Teena Hayden, and Amy Belitz. This work has been a group effort and I could not have finished it without your help.

Abstract

Three experiments were conducted to develop and validate predictive statistical models for isoflavone extraction from edamame soybean [*Glycine max* (L.) Merrill]. Soybean line NUTRIVEG Soy6407 was chosen to be the experimental sample material. The first set of extractions tested five factors; solvents, temperatures, pressures, extraction time per cycle and number of extraction cycles. Least squares means for six observed isoflavones (daidzin, glycitin, genistin, malonyl daidzin, malonyl glycitin and malonyl genistin) were used to evaluate trends for the effect of changes in each treatment factor. Pressures had no significant differences for any of the isoflavone extractions. Therefore, 500psi (the lowest pressure setting on the ASE instrument sufficient to maintain solvents in a liquid state when temperatures were raised above the boiling point), was used for all subsequent experiments.

A second set of experiments was done to develop predictive regression models. Temperature had a significant effect on glycitin extraction, and extraction time had a significant effect on daidzin extraction. There were differences among solvents for malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavone content. Temperature and extraction time interacted with solvent, and the interaction was significant for daidzin, genistin and malonyl glycitin extraction. The final models predicted the quantity of extract for daidzin, glycitin, genistin, malonyl daidzin, malonyl glycitin and malonyl genistin respectively at 43.22, 53.57, 44.36, 672.13, 344.14 and 443.30 µg per

gram of dry weight edamame. Model validation was accomplished on a third set of extractions by calculating the means of percent absolute differences between true values and the predicted values in order to understand the accuracy of the model. Two organic solvents, 60% methanol and 80% methanol, were determined to have the most consistent extractions. The means of percent absolute differences were approximately 20% or less, and evidence led to the conclusion that the models for these two solvents were reliable.

Table of Contents

List of Figures

List of Tables

Nomenclature

%	percent
°C	degree Celsius
°F	degree Fahrenheit
g	gram
mg	milligram
mL	milliliter
mm	millimeter
µm	micrometer
nm	nanometer
min	minute
psi	pounds per square inch
sec	second
µg	microgram
µg/g	microgram per gram

Abbreviations

aka	also known as
ASE	accelerated solvent extractor
	accelerated solvent extraction
AVRDC	Asian Vegetable Research & Development Center
DAD	diode array detector
EtOH	ethanol
G3D	three dimensional graphic
GLM	general linear model
H_2O	water
HPLC	high performance liquid chromatography
LDL	low-density lipoprotein
LOX	lipoxygenase
LSM	least mean square
MeCN	acetonitrile
MeOH	methanol
NaCl	sodium chloride (salt)
N/A	not applicable
n.o.	not observed
RF	response factor
USDA	United State Department of Agriculture
UV	ultraviolet

Isoflavone Abbreviations

Di	daidzin
Gly	glycitin
Gi	genistin
AcDi	acetyl daidzin
AcGly	acetyl glycitin
AcGi	acetyl genistin
MDi	malonyl daidzin
MGly	malonyl glycitin
MGi	malonyl genistin
De	daidzein
Gle	glycitein
Ge	genistein

Chapter I:

Introduction and Literature Review

1. Introduction

Soybean [*Glycine max* (L.) Merrill] is one of the oldest cultivated crops and also one of the five basic grains (rice, soybean, wheat, barley and millet) in several eastern Asian countries (Hymowitz, 1970; Gibson and Benson, 2002). Soybeans were first brought into North America in 1765 (Riggs, 2004). Johnson et al. (1999) estimated that roughly 13,000 hectares of growing area is required in order to meet the demand for the edamame niche market in the United States. The import of frozen edamame was 300 to 500 tons per year during the 1980s and increased to 10,000 tons in 2000, with Taiwan and China as the two major suppliers (Lin, 2001). The soybean is a herbaceous member of the family Fabaceae. Soybeans originated in East Asia. However, the world's largest soybean producer is currently the United States (SoyStats, 2008). In 2007, soybeans covered 25.7 million hectares of United States cropland, with total yields of 70.4 million metric tons and farm cash receipts reaching 26.2 billion dollars (USDA, 2008; SoyStats, 2008). The top five United States soybean export destinations were as follows: 11.1 million metric tons to China, 3.9 million metric tons to Mexico, 3.4 million metric tons to the European Union, 3.3 million metric tons to Japan, and 2.1 metric tons to Taiwan. Also, in 2007, 56% of the world's oil production came from soybeans, and 32% of those soybeans were produced in the United States (SoyStats, 2008).

Soybean is a healthy crop for humans and livestock because they are rich in nutrients

such as protein and amino acids. Soybeans have been widely used as a food crop in many Asian countries, particularly in Taiwan, Japan, China, and Korea. In addition to food, soybeans can be used as an energy crop in the form of biodiesel. As implied by the name, biodiesel is a type of fuel derived from plant oils which can be used as a fuel for diesel engines (Zhang, 2004).

Edamame are specialized soybeans and are also known as vegetable soybeans. Unlike most soybeans which are harvested at the R8 stage, edamame beans are harvested at the R6 stage (Lumpkin, 1993). The best description for the R6 stage is that pods contain green seeds that fill the pod cavity, and a fully developed leaf can be found at one of the four uppermost nodes (Fehr and Caviness, 1977). Edamame are selected for both seed size and flavor. The flavor is described as a nutty, sweet taste, and edamame can be eaten as a snack food after boiling in salted water (Mentreddy, 2002). Edamame is also a nutraceutical crop, and the nutritional composition is similar to other soybean cultivars. Edamame is gaining popularity as a fresh or frozen vegetable throughout the United States. Edamame is abundant in protein and oil. Soybean is one of the crops with a significantly high content of isoflavones. Isoflavones are associated with several health benefits to humans. Isoflavones have been shown in clinical studies to reduce blood serum cholesterol levels and the risk of cardiovascular diseases (Wiseman et al., 2000). Setchell and Cassidy (1999) also indicated that isoflavones may be an effective treatment for

hormone-related diseases including cancer, menopausal symptoms, and osteoporosis.

2. Plant Classification and Physiology

Fehr and Caviness (1977) defined the stages of soybean development. The purpose of vegetative and reproductive stage descriptions is to create the same terminology for soybean development among any cultivar grown at any location. Both vegetative and reproductive development have detailed descriptions used to classify stages. Node identification is required to determinate the vegetative and reproductive stages because nodes are permanent marks on plants (Figure 1, refer to Appendices for all tables and figures).

Soybean is a herbacious annual crop with trifoliolate leaves. The surface of the soybean seed pods is covered with pubescence, and edamame are harvested when the pubescence is completely green (Carter and Shanmugasundaram, 1993). Soybean typically contains 40% protein and 20% oil on a dry weight basis. The remaining dry matter within a seed contains approximately 35% carbohydrates and 5% ash. Soybean seed weight consists of approximately 8% seed coat or hull, 90% cotyledons and 2% hypocotyl axis (Liu, 1997). The cotyledon is the major portion of a whole seed; therefore, it can represent the whole seed oil and protein composition value regardless of large differences among structural parts. Pods sizes differ among varieties, and normally

edamame are larger than field dried soybeans (Shurtleff and Aoyagi, 1994).

3. Use of Soybeans

Soybeans are consumed in various ways as flour, a meal product, extracted oil or fresh beans. Soy fermentation techniques were invented millennia ago in Asian countries. Examples of fermented soy products are miso (soy paste), soy sauce, tempeh, and natto. Furthermore, non-fermented soy products include soymilk and tofu. Soybean products have spread into world-wide use, and their consumption is rapidly increasing, both because of improved cultivation methods and cultural exchange. Due to soybeans' high nutritional value, soy foods are highly recommended and promoted as health foods (Liu, 1997).

4. Cultivation

Cutivation for edamame is similar to field soybean production, with the exception of the harvest timing. The Asian Vegetable Research & Development Center (AVRDC) recommends several environmental conditions favorable for growing soybeans. The preferred planting date depends on temperature and day length. Ideal temperatures are 20-30°C (68-86°F) throughout the growing period with a maximum of 14-hour day length. It is necessary to avoid cool temperatures. Loamy soils with a pH of 6.0-6.5 favors soybean production, and the fields require good drainage (Lal et al., 2006).

Soybeans are sensitive to several types of biotic stress such as bacteria, fungi, viruses and nematodes. Therefore, it is necessary to choose the correct cultivar with specific resistances for production in a given region (Liu, 1997). A common example of biotic stress is powdery mildew, a fungal stress caused by *Erysiphe pisi* (Sillero, 2006). Abiotic stresses include heat, chilling, drought, wounding, osmotic shock, excessive light, etc. Both biotic and abiotic stresses can significantly reduce yield and quality in soybeans.

Identifying the proper timing for harvest is essential to controlling the quality of edamame. When edamame are harvested before or after the ideal maturity date, the pods will lose their optimum appearance, and the seeds will lack the preferred sensory traits. Also, quality may be altered due to cultivar, harvest timing, and environmental conditions (Mbuvi and Litchfield, 1995). Pods harvested early will have smaller, underdeveloped seeds and less sweetness (Carter and Shanmugasundaram, 1993). Pods may start dehydrating, and seeds may start turning yellow if harvested past the prime maturity stage (Carter and Shanmugasundaram, 1993).

Pod size, number of seeds per pod, pod color and degree of pest damage determine the standard traits for edamame (Tsou and Hong, 1991). Edamame have been assigned grades (Masuda, 1994; Sitatani, 1991). Both Grade A and B edamame contain at least 90% of the pods with two to three seeds, and Grade A edamame pods have no mechanical or pest damage. Grade B pods are lighter green in color and may have slight insect damage.

Smaller seeds and less uniform pods are acceptable for grade B (Konovsky et al., 1994).

5. Chemistry

Markley (1970) evaluated edamame chemical composition and found they have relatively high protein and lower oil content compared to field dried soybeans. Nutritionally, edamame consist of higher ascorbic acid and β-carotene and lower proportion of trypsin inhibitors, phytates, and oligosaccharides (Liu, 1997). Edamame seeds contain approximately 16% protein, 5% lipids, 13% carbohydrate, and 65% moisture (Rubel et al., 1972).

5.1 Proteins

Proteins are an important nutrient source for humans. Soybeans contain all the essential amino acids required for humans, including isoleucine, leucine, lysine, methionine, cysteine, phenalyanine, tyrosine, threonine, tryptohpan, valine and histidine (Carter and Shanmugasundaram, 1993). However, they are deficient in the two sulfur containing amino acids methionine and cysteine. Soybeans are one of the few vegetable protein sources that provide the most complete amino acid profiles for humans (Cui et al., 1999)

5.2 Lipids

Lipids are a relatively small component part of soybean chemical composition. Triacylglycerides (TAG) are the majority component in soybean lipids (Liu, 1997). TAGs consist of three fatty acids attached to one glycerol molecule (Groff and Gropper, 2000). Soybean oil fatty acids are mostly unsaturated, with the most abundant fatty acid being linoleic acid (Liu, 1997). Monounsaturated (oleic) and polyunsaturated (linoleic and linolenic) fatty acids help reduce the risk of cardiovascular disease attributed to a decrease in low density lipoprotein cholesterol (LDLc) (Groff and Gropper, 2000).

5.3 Carbohydrates

The carbohydrates found in soybeans are primarily disaccharides (sucrose) and oligosaccharides (raffinose and stachyose) (Liu, 1997). Estimated concentrations of these carbohydrates are 41-67% sucrose, 5-16% raffinose, and 12-35% stachyose (USDA 2004). The variation in sugar content among soybean cultivars suggests a mutation in the genes that encode enzymes linked to oligosaccharide sugars (Geater and Fehr, 2000). Oligosaccharides are not digestible by the human body (Groff and Gropper, 2000). Removal of oligosaccharides can be accomplished by fermentation or enzymatic hydrolysis (Liu, 1997). Soybean breeding programs are being conducted to research an optimal reduction or elimination of the problematic carbohydrates (Masuda, 1991).

5.4 Lipoxygenase

Lipoxygenase (LOX) catalyzes the oxidation of polyunsaturated fatty acids and produces conjugated fatty acid hydroperoxide derivatives (Liu, 1997). Lipoxygenase is active in the production of volatile and aromatic compounds in plant products and is associated with the formation of free radicals that degrade essential constituents such as vitamins, phenolics and proteins (Robinson et al., 1995). Research showed that genetically removing LOX helps lessen the degree of beany flavor in the soybean (Konovsky et al., 1994).

Isozymes of LOX, L-1, L-2, L-3a, and L-3b, have been recognized and categorized in soybeans. The L-1 isozyme has a optimum pH of 9 and forms large amounts of 13-hydroperoxides. L-2, L-3a and L-3b are different from L-1, and they have an optimum pH of 7.0 and form equal amounts of 9-hydroperoxides and 13-hydroperoxides (Kumar et al. 2003).

5.5 Volatiles

Besides appearance, sweetness, texture and overall flavor, the acceptability of soybean products is also determined by its volatile flavor component (Carter and Shanmugasundaram, 1993). The flavor of edamame is frequently described as nutty and sweet, with a low degree of beany flavor for soy products (Tsou and Hong, 1991; Konovsky

et al., 1994; Young et al., 2000). Cultivar of soybean, fertilizer application, planting density, harvest procedures, processing conditions and duration of frozen storage are the main factors that contribute to the differences in edamame flavor (Konovsky et al., 1994). Postharvest handling is considered to be the major influence on edamame quality and flavor. It is critical to have rapid cooling after harvesting. Blanching prior to freezing for vegetable soybeans is recommended to reduce occurrences of off-flavor development (Robinson et al., 1995).

Attempts have been made at the AVRDC to develop an objective evaluation system which could be used to grade the flavor quality of edamame (Tsou and Hong, 1991). Tsou and Hong (1991) used gas chromatography to study aromatic compounds which could be related to edamame flavor. These volatile compounds include aldehydes, acetals, esters, ketones, alcohols and several other aromatic compounds (Wilkens and Lin, 1970). The top ten major components of aromatic chemical compounds in soybeans have been identified as 3-methylbutanol, hexanol, 1-octene-3-ol, phenylethyl alcohol, maltol, γ-butyrolactone, heptanol, benyl alcohol, benzyl aldehyde and heptadecane (Lee et al., 2000).

Young et al. (2000) tested consumer acceptance of various vegetable soybean cultivars. Flavor was rated by sweetness, nuttiness, beaniness, oiliness, aftertaste, and overall eating quality. Consumers rated all cultivars as non-oily, having a low beany flavor and a pleasant aftertaste. Several cultivars were higher in nutty flavor, and soybeans varied

in total sweetness. The beans which had the overall best acceptance were the most sweet and least beany tasting.

United States consumers select food for nutrition, convenience, culture, economics and flavor. Flavor is the most important among these selection criteria (Young et al., 2000). If edamame are going to be accepted by United States consumers, it is clear to first optimize flavor. The requirement for defining edamame flavor was met through the development of lexicon descriptors for frozen edamame (Table 1; Krinsky et al., 2006).

6. Isoflavones

Soybean is one of the few crops that provides large amounts of isoflavones. Tsukamoto et al. (1995) stated that isoflavone content in soybean seeds differ by cultivar, growing location, planting date, and temperature during seed maturation. Within a seed, 80-90% of isoflavones accumulate in the cotyledons with the remainder in the hypocotyls. The ratio of isoflavones to dry weight is higher in hypocotyls compared with cotyledons. Roughly 2 mg of soy isoflavones per kg body weight are recommended to provide an anticarcinogenic effect for humans (Wang et al., 1994).

Isoflavones are phytoestrogens which are naturally produced in many plants and have a similar chemical structure as estrogen in animals (Wu et al., 2004; Rostagno, 2002). More than one form of isoflavones have been discovered in soybeans. They include aglycones

(daidzein, genistein, and glycitein), glucosides (daidzin, genistin, and glycitin), acetylglucosides (6”-O-acetyldaidzin, 6”-O-acetylgenistin, and 6”-O-acetylglycitin), and malonylglucosides (6”-O-malonyldaidzin, 6”-O-malonylgenistin, and 6”-O-malonylglycitin) (Figure 2; Wang, 1994; Charron, 2005). All of these compounds consist of the same basic chemical structure containing two benzyl rings with a three-carbon connector (Liu, 1997). This base can be modified with different functional groups including methylation, hydroxylation, and glycosylation (Klejdus et al., 2004). The major isoflavone group currently identified is malonyglucosides, but extraction methods are likely to increase the quantity of known aglycones, glucosides, and acetylglucosides in soybeans (Wang, 1994; Coward, 1998; Charron, 2005).

Isoflavones can be used as estrogen replacement therapy which relief the symptoms in postmenopausal women and prevent osteoporosis and cardiovascular disease. Han et al. (2002) suggested that isoflavone 100-mg regime treatment may be a safe and effective alternative therapy for menopausal symptoms and also promote a benefit to cardiovascular system.

7. Breeding

New cultivars being developed in breeding programs target characteristics of high economic yields, tolerance or resistance to biotic and abiotic stresses, traits such as flavor or

nutrients content that add value to the end product, and stability of the traits in target environments (Rao et al., 2002). The objective of breeding programs for edamame varieties include improving nutritional levels and agronomic traits in order to expand the acceptance of edamame products to commercial consumers (Mebrahtu et al., 1991).

8. ASE Extraction

Accelerated solvent extraction (ASE) is also known as pressurized solvent extraction or pressurized liquid extraction. Pressure and temperature are involved in ASE. The purpose of pressure application is to maintain the solvent in a liquid state when extraction temperature is above the boiling point and to increase the contact between the extraction solvent and sample. Temperature is used to break the chemical bonds and improve object compound solubility (Huie, 2002). Accelerated solvent extraction procedure can be used for isoflavones from soybeans and soy products.

Klejdus et al. (2004) compared the extraction efficiency of ASE with other extraction techniques and determined the isoflavone concentrations from soybean food samples using high performance liquid chromatography (HPLC). Accelerated solvent extraction has been compared to soxhlet extraction, sonication extraction and a combination extraction method involving ASE with 1 min sonication. The extraction technique that yielded the highest concentration of isoflavones (total of daidzin, genistin, daidzein and genistein) was

ASE combined with sonication. Next in yield was soxhlet, followed by ASE and sonication itself as is the least efficient method. Accelerated solvent extraction combined with sonication (142.7 μg/g) is twice as productive as ASE alone (69.9 μg/g).

Klejdus et al. (2004) tested the optimization of the ASE procedure with the following six variables:

i. effects of sonication time prior to ASE

ii. number of extraction cycles of ASE

iii. types of extraction solvents

iv. weights of samples

v. pressures during the extraction procedure

vi. temperatures

Extraction yields dramatically increased when samples were sonicated for 1 min before ASE, and yields stablized after 2 min of sonication. The highest isoflavone concentrations were obtained after three ASE extraction cycles. There was a 5% difference between two extraction cycles and three extraction cycles. Acetonitrile (MeCN), ethanol (EtOH) and methanol (MeOH) were the selected solvents for ASE extraction of isoflavones. Daidzin and genistin were monitored for yield comparisons of the different solvents. The extraction yields increased as solvent concentration increased except for daidzin in 80% acetonitrile extraction. Methanol was tested and determined to be the most suitable

solvent for ASE extraction for diadzin and genistin. Klejdus et al. (2004) also found there was a decreasing yield with increasing sample weight from 0.1 g to 0.5 g. The higher pressure of 2030 psi increased yield approximately 25% compared to the lower pressure of 1886 psi. A temperature of 145°C was found to yield the highest concentration of daidzin and genistin. Achouei et al. (2005) suggested that replicate extraction is essential for determining optimum extraction.

Rostagno et al. (2004) used several variables for optimization of the ASE protocol: solvents (30-100% MeOH, 30-100% EtOH and water (H_2O)), temperature (60-200°C), pressure (1470, 2940 psi), sample size (0.05-0.5 g), and cycle length (5, 7 or 10 min). Each variable was tested individually without any combination. The following results are the findings from Rostagno et al. (2004) for each variable.

(Solvent) The initial experiments for solvent selection used 0.5 g of freeze-dried soybeans in H_2O, MeOH (30-100%) or EtOH (30-100%) at 60°C, 1470 psi, three static extraction cycles of 5 min each and a purging procedure with nitrogen for 300 seconds. For EtOH/H_2O mixtures, extraction efficiency increased when H_2O was applied from 100% EtOH (total isoflavones 394.73µg/g) to 70% EtOH (total isoflavones 917.48µg/g). Extraction efficiency decreased from 70% EtOH to 30% EtOH and the lowest amount of total isoflaones (314.59µg/g) was detected with pure water extraction. For MeOH/ H_2O mixtures, there were similar results with the highest extraction efficiency achieved with

60% MeOH (total isoflavones 870.62µg/g). The findings above with different percentages of MeOH and EtOH showed that it is necessary to have a specific amounts of water added to the extraction in order to improve the isoflavone extraction efficiency.

(Temperature) A series of experiments used different temperatures (60-200°C) with 70% EtOH, 1470 psi, 0.5g of sample, three static extraction cycles of 5 min each and a purging procedure with nitrogen for 300 seconds. At 150°C, the greatest extraction efficiency was obtained for total isoflavones at 1249.14µg/g and demonstrated that isoflavone glucosides (daidzin, glycitin and genistin) did not have degradation under ASE conditions below 150°C, while isoflavone aglycones (daidzein, glycitein and genistein) stayed constant from 60 to 150°C. Aglycones showed significant increases at 200°C, while glycosides showed significant decreases. The malonyl forms of isoflavone were not detected. Degradation was observed for malonyl forms of isoflavone over 100°C, and the degradation for all other forms of isoflavone started at 150°C (Table 2).

(Pressure) A series of experiments tested two different levels of pressure (1470, 2940 psi) with 70% EtOH, 100°C, 0.5g of sample, three static extraction cycles of 5 min each and a purging procedure with nitrogen for 300 seconds. An increase in pressure did not significantly change the extraction efficiency (Table 3).

(Sample Size) A series of experiments used different sample sizes (0.5, 0.25, 0.10, 0.05g) with 70% EtOH, 100°C, 0.5g of sample, three static extraction cycles of 5 min each

and a purging procedure with nitrogen for 300 seconds. The extraction efficiency of isoflavones increased with a reduction of sample size from 0.5 to 0.05g (Table 4). Since the peaks for each isoflavone in chromatograms were extremely small for the sample size 0.05g extraction, it was possible to have higher errors during peak integration in HPLC chromatograms. Therefore, Rostagno et al. (2004) concluded that 0.1 g was an optimum level of sample size for accelerated solvent extraction.

(Cycle Length) A series of experiments used three static extraction cycles with different lengths (5, 7, 10 min) with 70% EtOH, 100°C, 1470psi, 0.1g of sample, and a purging procedure with nitrogen for 300 seconds. An increase of extraction efficiency can be seen from 5 min to 7 min, and there was no significant difference when the static time was extended to 10 min (Table 5).

(Static Extraction Cycle) It is important to determine whether or not all three extraction cycles are necessary to complete the extraction. A series of experiments used one, two or three static extraction cycles of 7 min and two cycles of 10 min with 70% EtOH, 100°C, 1470psi, 0.1 g of sample, and purge with nitrogen for 300 seconds. Two cycles of 10 min is only one minute shorter than 3 cycles of 7 min but there is much time saved between extraction cycles due to de-pressurization, flushing, purging and pressurization for the system to start another cycle. A significantly higher efficiency was detected with 3 cycles of 7 min static extraction (Table 6).

The final optimized extraction conditions are 70% of EtOH, 0.1g of sample, 100°C, 1470psi, and three cycles of 7 min static extraction time. This protocol required a total of 32 min which includes 5 min of pre-heating, 21 min for extraction, 1 minute of flushing and 5 min of purging.

Although study has been conducted to extract isoflavones from mature and dry commodity soybean, no research has yet been conducted to optimize accelerated solvent extraction for isoflavones from edamame. The first objective of this study is to develop a predictive statistical model and estimate the efficiency of the accelerated solvent extraction procedure for isoflavones from edamame soybeans. The second objective is to validate the predictive statistical model by calculating the differences between extracted edamame isoflavone value and statistical predicted value.

Literature Cited

Carter, T.E., Jr. and Shanmuglasundaram, S. Edamame, the vegetable soybean. In: Underutilized Crops: Pulses and Vegetables (Edited by T.Howard). London: Chapman and Hill.1993: 219-239.

Charron, C. S., Allen, F. L., Johnson, R. D., Pantalone, V. R., and Sams, C. E., Correlations of oil and protein with isoflavone concentration in soybean. J. Agri. Food Chem., 53, 7128-7135, 2005.

Coward, L. Smith, M. Kirk, M. Barnes, S. Chemical modification of isoflavones in soyfoods during cooking and processing. Am. J. Clin. Nutr. 69:1486-1491. 1998.

Cui, Z., Carter, T.E., Jr., Gai, J., Qiu, J., and Nelson, R.L. Origin, Description, and Pedigree of Chinese Soybean Cultivars Released from 1923 to 1995. USDA. Government Printing Office, Washington, D.C., U.S.1999.

Fehr, W.R., and Caviness C.E. Stages of Soybean Development. Iowa State University of Science and Technology. Agriculture and home Economics Experiment Station. 80: 1-12. 1977.

Geater, C.W., and Fehr, W.R. Association of total sugar content with other seed traits of diverse soybean cultivars. Crop Sci. 40:1552-1555. 2000.

Groff, J.L., and Gropper, S.S. Advanced Nutrition and Human Metabolism, 3rd Edition. United States: Wadsworth. 2000.

Han, K.K., Soares, J.M., Haudar, M.A. Lima, G.R. and Baracat, E.C. Benefits of Soy Isoflavone Therapeutic Tegimen on Menopausal Symptoms. Obstet Gynecol. 99: 389-394. 2002.

Hurburgh, Jr. C.R., Soybean Drying and Storage. Iowa State University Extention. 1995.

Huie C. W. A Review of Modern Sample-preparation Techniques for the Extraction and Analysis of Medicinal Plants. Anal Bioanal Chem. 373: 23–30. 2002.

Johnson, D., S. Wang, and A. Suzuki. Edamame: a Vegetable soybean for Colorado. J Janick (ed.), Perspectives on new crops and new uses. ASHS Press, Alexandria, VA. 385-387. 1999.

Klejdus et al. 2004 Determination of isoflavones in soybean food and human urine using liquid chromatography with electrochemical detection. Mendel University of Agriculture and Forestry and Masaryk University, Czech Republic. J. Chromatogr. Sci. B, 806: 101–111. 2004.

Kumar, V., Rani, A., Tindwani, C., and Jain, M. Lipoxygenase isozymes and trypsin inhibitor activities in soybean as influenced by growing location. Food Chemistry, 83:79-83. 2003.

Konovsky, J., Lumpkin, T.A., and McClary D. Edamame: The vegetable soybean. In: Understanding the Japanese Food and Agrimarket: A Multifaceted Opportunity. A.D. O'Rourke (ed.). Binghamton: Hayworth. 173-181. 1994.

Krinsky, B.F., Drake, M.A., Civille, G.V., Dean, L.L., Hendrix, K.W., and Sanders, T.H. The Development of a Lexicon for Frozen Vegetable Soybean (Edamame). J. Sens Stud. 21(6): 644-653. 2006.

Lal et al.,"Soybean" Asian Vegetable Research and Development Center, 2006., <http://www.avrdc.org/>

Lang, Q., and Wai, C. M. Supercritical fluid extraction in herbal and natural product studies – a practical review. Talanta, 53: 771-782. 2001.

Lee, K., Mitchell, A., and Shibamoto, T. Antioxidative Activities of Aroma Extracts Isolated from Natural Plants. BioFactors. 13: 173-178. 2000.

Liu, K. Soybeans: Chemistry, Technology, and Utilization. Aspen Publisher Inc. Gaithersburg, MD. 1997.

Lumpkin, T.A., J.C. Konovsky, K.J. Larson, and D.C. McClary. Potential new specialty crops from Asia: Azuki bean, edamame soybean, and astragalus. In: J. Janick and J.E. Simon (eds.), New crops. Wiley, New York. 45-51. 1993.

Masuda, R. Quality requirement and improvement of vegetable soybean. In: Vegetable Soybean: Research Needs for Production and Quality Improvement. Shanmugasundaram, S. (ed.). Proceedings of a workshop held in Kenting, Taiwan, AVRDC, Publication No. 91-356. 92-107. 1991.

Markley, KS. Soybeans and Soybean Products. New York: Interscience Publishers, Inc. 1950.

Mbuvi, S.W. and Litchfield, J.B. Green soybeans as vegetable: Comparing green soybeans with green peas and lima beans, and maximized harvest time determinations using mathematical modeling. Journal of Vegetable Crop Production, 1: 99-121. 1995.

Mebrahtu, T.A., Mohamed A., and Mersie, W. Green pod and architectural traits of selected vegetable soybean genotypes. Journal of Production Agriculture, 4(3): 395-399. 1991.

Mentreddy, et al. 2002. Edamame: A Nutritious Vegetable Crop. ASHS Press, Alexandria, VA

Mohamed, A.I., and Rangappa, M. Nutrient composition and anti-nutritional factors in vegetable soybean: II. Oil, fatty acids, sterols, and lipoxygenase activity. Food Chemistry, 44: 277-282. 1991.

Rao, M.S.S., Mullinix, B.G., Rangappa, M., Cebert, E., Bhagsari, A.S., Sapra, V.T., Joshi, J.M., and Dadson, R.B. Genotype x environment interactions and yield stability of food-grade soybean genotypes. Agron. J. 94: 72-80. 2002.

Robinson, D.S., Wu, Z., Domoney, C., and Casey, R. Lipoxygenase and the quality of foods. Food Chemistry, 54: 33-43. 1995.

Rostagno, et al. Universidad de Cádiz. Pressurized liquid extraction of isoflavones from soybeans. Analytica Chimica Acta. 522 (2): 169-177. 2004.

Rubel, A., Rinne, R. W., Canvin, D. T. Protein, Oil, and fatty acid in deverloping soybean seeds. Crop Sci. 12: 739-741. 1972.

Setchell K. and Cassidy A. Children's Hospital Medical Center and University of Surrey. Dietary Isoflavones: Biological Effects and Relevance to Human Health. J. of Nutri. 129: 758-767. 1999.

Soystates. [online] Available at www.soystats.com (verified 7 November 2008)

Shurtleff, W., and & Aoyagi, A. Green Vegetable Soybeans (Edamame) and Vegetable-Type Soybeans – Bibliography and Sourcebook, Soyfoods Center Publishing: California. 1994.

Sitatani K. Cultivation practices for vegetable soybean. In: Vegetable soybean production: proceedings of a training course, AVRDC. Proceedings held in Chiang Mai, Thailand. 19-23. 1991.

Takahashi, N. Vegetable soybean varietal improvement in Japan – past, present, and future. In: Vegetable soybean: research needs for production and quality improvement. S. Shanmugasundaram (ed.) 26-29. Proceedings of a workshop held at Kenting, Taiwan. AVRDC. Publication No. 91-356. 1991.

Tsou, S.C.S. and Hong, T.L. 1991. Research on vegetable soybean quality in Taiwan. In: Vegetable soybean: research needs for production and quality improvement. S. Shanmugasundaram (ed.) 103-107. Proceedings of a workshop held at Kenting, Taiwan AVRDC. Publication No. 91-356. 1991.

Tsukamoto, C., Shimada, S., Igita, K., Kudou, S., Kokubun, M., Okubo, K., and Kitamura, K. Factors Affecting Isoflavone Content in Soybean Seeds: Changes in Isoflavones, Saponins, and Composition of Fatty Acids at Different Temperatures during Seed Development. Journal of Agri. and Food Chem. 43: 1184-1192. 1995.

Wang, H.J., and Murphy, P.A. Isoflavone content in commercial soybean foods. J. Agri. Food Chem. 42: 1666-1673. 1994.

Wilkens, W.F., and Lin, F.M. Gas chromatographic and mass spectral analyses of soybean milk volatiles. Journal of Agriculture and Food Chemistry, 18, 333-338. 1970.

Wiseman, H., J. D. O'Reilly, and H. Adlercreutz. Isoflavone phytoestrogens consumed in soy decrease F92-isoprostane concentrations and increase resistance of low-density lipoprotein to oxidation in humans. Am. J. Clin. Nutr. 72:395-400. 2000.

Young, G., Mebrahtu, T., and Johnson, J. Acceptability of green soybeans as a vegetable entity. Plant Foods for Human Nutrition, 55, 323-333. 2000.

Zhang Yu. Comparison of Potential business structures for a biodiesel production Facility and Analysis of Tennessee Soybean Producers' willingness to join a New Generation Cooperative to Produce Biodiesel. Master of Science Thesis. The University of Tennessee. 2004.

Chapter II:

Development of a Predictive Statistical Model to Optimize Accelerated Solvent Extraction of Isoflavones from Edamame Soybean

Abstract

Edamame, also known as edible soybean or vegetable soybean, is a significant dietary source of isoflavones. Isoflavones are phytoestrogens and they are important in human health. The soybean line NUTRIVEG Soy6407 was grown in Tennessee during the 2006 season, and pods were harvested at reproductive stage six (R6). Edamame seeds were freeze-dried, ground and spiked with an apigenin internal standard. Isoflavones were removed with accelerated solvent extraction (ASE). Six isoflavones (daidzin, glycitin, genistin, malonyl daidzin, malonyl glycitin, and malonyl genistin) were quantified by high performance liquid chromatography. Temperature had a significant effect on glycitin extraction, and extraction time had a significant effect on daidzin extraction. There were differences among solvents for malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavones extraction. Temperature and extraction time interacted with solvent, and the interaction was significant for daidzin, genistin and malonyl glycitin extraction. Final models predicted the optimum extractions for daidzin, glycitin, genistin, malonyl daidzin, malonyl glycitin, and malonyl genistin to be 4.32, 5.36, 4.44, 67.21, 34.41 and 44.33 μg/g of seed dry weight. The R^2 value for each isoflavone was 0.520, 0.653, 0.649, 0.708, 0.677 and 0.646, respectively. As an example, the total isoflavone optimization model had R^2 value of 0.677 and it predicted 150.04 μg/g using 80% ethanol at 97°C with 500 psi for 3 extraction cycles of 30 min per cycle.

1. Introduction

Soybeans [*Glycine max* (L.) Merrill] are one of the most important crops grown today. They are commercialized into numerous products worldwide because they contain high amounts of protein and oil. It has been recommended that people consume soybeans because of multiple health benefits, including the high isoflavone content. Isoflavones are phytoestrogens, which are plant secondary metabolites that have a similar chemical structure to mammalian estrogen (Setchell and Adlercreutz, 1988). The similarity allows isoflavones to interact with estrogen receptors in mammalian metabolism and cause phyto-protectant effects to hormone-dependent diseases like osteoprosis and breast cancer (Setchell and Cassidy, 1999; Messina et al., 2006). Soybeans have been served as a major food crop in many Asian countries, and consumption of soy products in the United States is dramatically increasing.

Most soybeans grown in the world are harvested at the R8 stage in which the matured seeds and pods are field-dried on the plants (Fehr and Caviness, 1977). Edamame, also known as edible soybean or vegetable soybean, are soybeans harvested at the R6 stage in which the seeds have just fully filled the pod capacity and have not proceeded to the drying stage (Fehr and Caviness, 1977, Lumpkin, 1993). Therefore, edamame have a very short harvest window, and timing of harvest is the most crucial factor for future consumer acceptability and marketability. Edamame also have a shorter growing season than

conventional soybeans.

Edamame exhibits several differences from conventional soybean, which include larger seed size, milder taste, increased tenderness, and better digestiblility (Born, 2006). Edamame is consumed as a fresh vegetable and is highly perishable after harvest unlike field dry soybean. There is a large advantage for processing edamame than conventional soybean. Edamame does not need to be threshed before going to market because the consumers can simply boil edamame in salt water and shell the pods before eating which saves large quantity of time and energy for industrial process. There is also an increasing trend to consuming shelled fresh frozen edamame in vegetable mixes.

Soybean breeders have been studying the identification of quantitative trait loci (QTL) associated with individual and total isoflavone content (Primomo et al., 2005). In order to make isoflavones more pharmaceutically available, it would be helpful to optimize the extraction procedure. The objective of this research was to model an isoflavone extraction procedure from edamame soybean. This experiment was designed to investigate different percentages of ethanol, methanol and acetonitrile in water with respect to several variables including temperature, pressure, extraction time and number of extraction cycles.

2. Materials and Methods

2.1. Sample Processing

Soybean line NUTRIVEG Soy6407 was chosen for this experiment. This line is a new cultivar developed by the Tennessee Agricultural Experiment Station for edamame production. During 2006, the seeds were field grown in Sequatchie silt loam soil at the Plant Sciences Unit of the East Tennessee Research and Education Center in Knoxville, Tennessee. NUTRIVEG Soy6407 edamame was collected at the R6 stage in which the green seeds fill the pod cavity (Fehr and Caviness, 1977). All pods were stored at 4°C, and shelling was done within one day after harvesting with seeds being immediately frozen at -80°C until analysis. The shelled edamame seeds were freeze-dried prior to extraction, and the freeze-drying procedure was controlled at -40°C. The dry weight was measured after freeze-drying in order to make sure the weight difference was 2% or less at the end of the drying process. Dry edamame seeds were ground with a Knifetec 1095 Sample Mill (Foss Tecator, Hogana Sweden) for two 10-second pulses. Samples were preserved with ice before grinding, water cooled at 4°C during grinding and returned to ice after grinding to avoid isoflavone degradation. All the ground edamame were homogenized to minimize differences among seeds. A uniform sample was obtained and utilized for all extractions in these experiments.

2.2. Chemicals and Solvents

Water was supplied by a Milli-Q purifier from Millipore (Bedford, MA, USA). The isoflavone standards were purchased from LC Laboratory (Woburn, MA, USA), and the internal standard apigenin was obtained from Sigma (St. Louis, MO, USA).

2.3. Extraction Procedure

We used a modified procedure from Rostagno et al. (2004). The isoflavone extraction was performed by a DIONEX Accelerated Solvent Extractor (ASE). A disk-shaped cellulose filter (diameter: 1.983cm) was pushed down to the bottom of an ASE 11-mL stainless extraction cell to protect the metallic filter. A sample of 0.1 ± 0.001g freeze dried, finely ground edamame seed was weighed and transferred into the ASE extraction cell. The edamame sample was placed between layers of Ottawa Sand (Fisher Scientific Co.) in the extraction cell (Figure 3). The total volume of Ottawa Sand was controlled by measuring 10 g for each extraction. All samples were spiked with 300μL of the internal standard apigenin. The cell was hand-tightened and loaded into the ASE instrument.

Luthria et al. (2007) compared of extraction solvents and techniques for isoflavones from soybean and summarized several commonly deployed procedures used by various authors. The best solvents were among different percentage of three organic solvents which were methanol (MeOH), ethanol (EtOH) and acetonitrile (MeCN). In this experiments, six

water miscible organic solvents (60% MeOH, 60% EtOH, 60% MeCN, 20% MeOH, 20% EtOH, 20% MeCN) and ultra purified water were selected for isoflavone extraction. Temperature, pressure, extraction time per cycle and number of cycles were adjusted for different combinations in the first set of extraction and model development experiment (Table 9). After the extraction was completed, the cell was purged with nitrogen for 300 seconds to force all the extracts into 60-mL glass vials with Teflon-coated rubber caps. As soon as the extraction was completed, the extracts were filtered through a nylon 0.45-µm syringe filter, stored in 2-mL cryovials and frozen at -80°C until High Performance Liquid Chromatography (HPLC) analysis.

2.4. High Performance Liquid Chromatography (HPLC) Analysis

We used a modified procedure from Griffin and Collison (2001) and Charron et al. (2005). Soybean isoflavone extracts were analyzed by an Agilent HPLC 1200 series (Santa Clara, USA) equipped with a binary pump, degasser, autosampler thermostat, XDB-C18 reverse phase column (1.8µm, 4.6x50mm) and diode-array detector (DAD). Solvent A was 0.1% (v/v) acetic acid in H_2O, and solvent B was 0.1% (v/v) acetic acid in acetonitrile. This isoflavone analysis method used 5.0 µL for the injection volume with a needle wash between every sample, 40°C for column temperature, 260nm for UV Lamp wavelength and a flow rate of 0.8 mL per min. The solvent gradient was 88:12 (A:B) at sample injection and shifted

after 15 min to 76:24 (A:B). At 20 min the column was washed with a ratio of 10:90 (A:B) for 5 min and equilibrated with the initial solvent ratio for 2 min (Table 7). The complete elapsed time for one sample was 27 min. Each isoflavone was evaluated and auto labeled by comparing the retention time with authentic standards of daidzin (Di), glycitin (Gly), genistin (Gi), malonyl daidzin (MDi), malonyl glycitin (MGly), acetyl daidzin (AcDi), malonyl genistin (MGi), daidzein (De), glycitein (Gle), acetyl genistin (AcGi) and genistein (Ge) (Figure 4). Each isoflavone concentration was calculated with individual response factors (Table 8).

2.5. Statistics

Statistical analysis for each isoflavones and total isoflavone content was performed by SAS 9.1 (SAS Institute, 2002). For the data in the first set of extractions, five experimental treatments (temperature, pressure, extraction time per cycle, number of cycles and solvent) were analyzed with the general linear model (GLM) analysis of variance with two-way interaction and differences between treatments were considered significant at $P<0.05$. Mean separation for treatments was run using the MIXED procedure with the least significant difference (LSD) level between treatments at 0.05.

Data was combined for the first and second set of extraction. Analysis of variance and a regression based PROC GLM were used for the combined data set and it estimated

parameters for each significant term in the model. Solvent was included as a classification variable, producing separate equations for each solvent. The G3D procedure was used to visualize the model surface with three dimensional graphs for predictive surface.

3. Results and Discussion

3.1. Preliminary Extraction

Temperature ($P < 0.0001$) and extraction time ($P < 0.0001$) showed significant effects for the extraction of daidzin. The R^2 for daidzin was 0.721 which meant 72.1% of the variation was explained by the model. There were significant differences among extraction time ($P < 0.005$) and cycle ($P < 0.0001$) for glycitin, and 79.1% of the variation was explained by these two significant variables in the model. There were significant differences among temperature ($P < 0.0001$) and extraction time ($P < 0.0001$) for genistin, and 80.0% of the variation was explained by the model. Temperature ($P < 0.001$) and cycle ($P < 0.05$) affected extraction of malonyl daidzin, and 68.9% of the variation was explained by the model. Temperature ($P < 0.05$) and extraction time ($P < 0.05$) showed significant difference for malonyl glycitin, and 74.6% of the variation was explained by the model. Temperature ($P < 0.01$) and cycle ($P < 0.001$) demonstrated significant differences for malonyl genistin and 65.7% of the variation was explained by the model. Temperature, extraction time and cycle had $P < 0.05$ for total isoflavone content, and 68.7% of the

variation was explained by the model. In all cases, the solvent showed a significant difference but pressure was non-significant. The Shapiro-Wilk tests were above 0.90 indicating the data sets were normal enough to analyze with statistics. There were some significant interactions between variables (Table 10).

The results in the first set of extractions were expressed with least squares means (LSM). Based on the LSM results of the first set of extractions, levels in each variable term were modified for higher extraction efficiency in the second set of extractions. For temperature, the extraction efficiency showed a trend toward the lower temperature. Except for glycitin, all other observed isoflavones had relatively higher LSMs at 40°C than at 80°C. Ambient temperature and 60°C were chosen for testing in the second set of experiments (Figure 5).

There was no significant difference for pressure, therefore the lower pressure (500psi) was chosen, and it allowed the organic solvent to maintain a liquid state at higher temperatures (Figure 6). However, 1000psi was tested to confirm that there is no effect on extractions due to pressure.

When extraction time increased from 10 minutes to 30 minutes, there were isoflavone content increases of approximately 450% for Di, 60% for Gly, 120% for Gi, about 7 to 9% for MDi, MGly, and MGi, and the overall total isoflavone content increased about 9%. The LSMs for each isoflavone and total isoflavone content increased as extraction time

increased (Figure 7). For this reason, 20 and 40 min were selected for the second set of experiments.

For Di, Gly and Gi, LSMs were higher with one extraction cycle. For MDi, MGly, MGe and total content, there were higher LSMs with three extraction cycles (Figure 8). Two cycles and four cycles were selected to test the model. Forty minutes extraction times with four cycles resulted in a total of 160 min which was a time length beyond 2 hours of simple solvent vortex extraction (Charron, 2005); therefore, only two cycles were tested in the second set of experiments.

For the same organic solvent, a 60% concentration had significantly higher LSM than 20% for each isoflavone and total isoflavone content (Figure 9). The 20% organic solvents did not enable a complete extraction as evidenced by the internal standard apigenin. Therefore, researchers will need greater than 20% organic solvents to achieve complete solubility for isoflavone extractions. Furthermore, the water extraction had better results than the 20% organic solvents, and the water extraction method was more environmentally friendly. The 80% organic solvents were added to the second set of experiments and the 60% organic solvents were tested again with other combinations.

3.2. Second set of experiments and model development.

The second set of experiments used 0.1g of freeze-dried edamame extracted by the

combinations in Table 11, and were combined with preliminary data (excluding the results with 20% organic solvents) to generate a statistical model.

The data from this experiment and preliminary extractions were pooled and analyzed with general linear model analysis of variance for individual isoflavone and total isoflavone content. The R^2 values for daidzin, glycitin, genistin, malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavone were 0.520, 0.653, 0.649, 0.708, 0.677, 0.646, and 0.677, respectively. All R^2 values were slightly decreased compared to statistical results in the preliminary extractions. Temperature had a significant effect on glycitin extraction ($P < 0.05$). Extraction time had a significant effect on daidzin extraction ($P < 0.05$). Solvent had a significant effect on malonyl daidzin ($P < 0.0001$), malonyl glycitin ($P < 0.01$), malonyl genistein ($P < 0.001$) and total isoflavone content ($P < 0.001$) extraction. The temperature × solvent and time × solvent interactions were significant for daidzin, genistein and malonyl glycitin extraction ($P < 0.01$). Pressure was non-significant for all extractions.

SAS 9.1 G3D procedure generated three dimensional graphs for the three variables with the following equations (Figure 10):

$$\begin{cases} Isoflavone = Solvent \times [a \times Temp + b \times Total\ Time + c] + Temp \times Total\ Time \\ Total\ Time = Extraction\ Time \times Number\ of\ Cycles \end{cases}$$

The surfaces were used to predict the optimized extractions for each isoflavone and

total isoflavone content (Table 19).

The ASE method is capable of applying external energy such as pressure and/or temperature increase efficiency. Thus, efficiency is increased and the extraction can be achieved in a shorter amount of time. Statistical results for pressure supports the previous work of Rostagno (2004), who also found that pressure was not significant for isoflavone extraction.

All observed isoflavone extractions except daidzin were decreased when temperature increased from 40°C to 80°C (Figure 5). This is understandable due to degradation because isoflavones are sensitive to temperature (Coward, 1998). It was surprising that the final optimized extraction predictions of Gi, Mdi, MGly, MGi and total isoflavone content were at the high temperature of 97°C (Table 19). This circumstance could be explained because the extraction yielding rate from 40°C to 80°C did not overcome the degradation rate. From 80°C to 97°C, the extraction rate was greater than the degradation rate. Each isoflavone increased as extraction time extended from 10 to 30 min, and there was approximately 10% difference in total isoflavone concentration. Daidzin, malonyl daidzin, malonyl genistin and total isoflavone content have significantly higher quantities in three extraction cycles than in one cycle. This result was supported by the Rostagno et al. (2004) and they estimated 0.1g edamame required three extraction cycles in order to maximize isoflavone extraction because the solvents were saturated in the first and second extraction

cycles.

Future research would be enhanced by including a study for model validation. Moreover, after validation, the model can be further tested for its statistical application to commodity soybean. This would be of interest because Stewart (2008) found that soybean isoflavone concentration increased with stage of reproductive development. Commodity soybeans harvested after the R8 stage and extracted under an optimized model should produce greater isoflavone content than edamame for the soy processing industry. Research should be conducted to evaluate the extraction amounts related to the total expense of solvent volume, temperature, and pressure applications and furthermore to consider how much energy and expense can be saved by reaching 95% of the complete isoflavone extraction.

4. Conclusion

every compound had its own optimal predicted result for isoflavone extraction when using 0.1 g of dry edamame as extraction sample. Daidzin was predicted to be 4.32µg by using H_2O or 60% MeOH at 97°C with three cycles of 30 min per cycle. Glycitin was predicted to be 5.36µg by using 80% MeOH at ambient temperature (27°C) with three extraction cycles of 30 min per cycle. Genistin was predicted to be 4.44µg by using 60% MeOH at 97°C with three extraction cycles of 30 min per cycle. Malonyl genistin was

predicted to be 67.21µg by using H_2O at ambient temperature (27°C) with one extraction cycle of 10 min. Malonyl glycitin was predicted to be 34.41µg by using 60% MeCN at 97°C with three extraction cycles of 30 min per cycle. Malonyl genistin was predicted to be 44.33µg by using 80% EtOH at 97°C with three extraction cycles of 30 min per cycle. One set of extraction procedure may not be optimal for every isoflavone compound and the maximum total isoflavone extraction was predicted to be 150.04µg by using 80% EtOH at 97°C with three cycles of 30 min per cycle. Glucoside isoflavones had lower concentration than malonyl isoflavone. Malonyl daidzin had the highest quantity (67.21µg per 0.1 g of dry edamame) among all the isoflavone derivatives in edamame. Each solvent had its own advantage to specific isoflavone compounds. Combinations between miscible organic solvents have potential to retrieve benefits from each solvent in order to increase the extraction efficiency.

Literature Cited

Born, H., Edamame: Vegetable soybean. ATTRA-National Sustainable Agriculture Information Service, 2006.

Charron, C. S., Allen, F. L., Johnson, R. D., Pantalone, V. R., and Sams, C. E., Correlations of oil and protein with isoflavone concentration in soybean. J. Agri. Food Chem. 53: 7128-7135, 2005.

Coward, L. Smith, M. Kirk, M., and Marnes, S., Chemical Modification of Isoflavones in Soyfoods during Cooking and Processing. Am J. Clin Nutr. 68: 1486S-1491S, 1998.

Fehr, W. R., and Caviness C. E., Stages of soybean development. Iowa State University of Science and Technology. Agriculture and home Economics Experiment Station, 80: 1-12, 1977.

Griffith A. P. and Collison M. W., Improved methods for the extraction and analysis of isoflavones from soy-containing foods and nutritional supplements by reversed-phase high-performance liquid chromatography and liquid chromatography–mass spectrometry, J. Chromatography. A 913: 397–413, 2001.

Lumpkin, T. A., J. C. Konovsky, K. J. Larson, and D. C. McClary, Potential new specialty crops from Asia: Azuki bean, edamame soybean, and astragalus. 45-51. In: J. Janick and J.E. Simon (eds.), New crops. Wiley, New York, 1993.

Luthria, D.L., Biswas, R. and Natarajan, S. Comparison of extraction solvents and techniques used for the assay of isoflavones from soybean. Food Chemistry. Vol. 105, 325-333. 2007.

Messina, M., Mc Caskill-Stevens, W., and Lampe, J. W., Addressing the soy and breast cancer relationship: review, commentary, and workshop proceeding. J. National Cancer Institute. 98: 1275-1284, 2006.

Primomo, S. V., Rajcan I., Poysa, V., Ablett G. R., Jackson, C. J., and Gijzen, M., Mapping QTL for individual and total isoflavone content in soybean seeds. J. Crop Science. 45: 2454-2464, 2005.

Rostagno, M. A., Palma M., and Barroso S. G., Pressurized liquid extraction of isoflavones from soybeans. Analytica Chimica Acta. 522(2): 169-177. 2004.

SAS Institute Inc. SAS guide for personal computers, 9.1 edn. SAS institute, Cary, NC. 2002.

Setchell, K. D. R. and Adlercreutz, H., Mammalian lignans and phytooestrogens. Recent studies on their formation, metabolism and biological role in health and disease. In: Role of the Gut Flora in Toxicity and Cancer (Rowland, I. R., ed.), 315–345, Academic Press, London, UK, 1988.

Setchell, K. and Cassidy, A., Dietary isoflavones: biological effects and relevence to human health. J. of Nutrition. 129: 758-767, 1999.

Stewart, A. E., Effect of pod maturity and plant spacing on isoflavone content and harvest force of edible soybeans. Master of Science Thesis, University of Tennessee, 2008

Chapter III:

Statistical Model Validation for Accelerated Solvent Extraction of Isoflavones from Edamame Soybean

Abstract

Soybeans are one of the few crops that contain abundant isoflavones. The edamame from soybean line NUTRIVEG Soy6407 was grown in Tennessee during the 2006 season, and pods were harvested at reproductive stage six (R6). Edamame seeds were threshed from pods, freeze-dried, ground and extracted with accelerated solvent extraction. Six isoflavones (daidzin, glycitin, genistin, malonyl daidzin, malonyl glycitin and malonyl genistin) were quantified by high performance liquid chromatography. Predictive models for optimum isoflavone extraction conditions have been recently developed in Chapter II. The purpose of this research was to validate the models by calculating the predictive error with actual difference, absolute difference and percentage of absolute difference. Solvents 60% MeOH and 80% MeOH were determined to be the most reliable models. For 60% MeOH, the means of percent absolute difference were 17.53% (glycitin), 15.29% (genistin), 22.39% (malonyl daidzin), 19.47% (malonyl glycitin), 6.12% (malonyl genistin) and 4.88% (total isoflavone content). For 80% MeOH, the means of percent absolute difference were 15.58% (glycitin), 19.02% (genistin), 21.13% (malonyl daidzin), 10.58% (malonyl glycitin), 13.57% (malonyl genistin) and 5.82% (total isoflavone content). Acetyl genistin was extracted 6.98µg per 0.1g of dry edamame with 80% EtOH at a temperature of 120°C and 9.38µg per gram of dry edamame with 80% MeCN at a temperature of 120°C during this validation experiment.

1. Introduction

Soybeans [*Glycine max* (L.) Merrill] are known as one of the most versatile crops and are grown in many countries. Soybeans are composed of approximatly 40% protein, 20% oil and 35% carbohydrates. They are used to make numerous products worldwide for snack foods and frozen foods and are used in vegetable mixes because of their nutritious protein and oil components. Soybeans are a major crop in many Asian countries, such as Taiwan, Japan, China and Korea; and the consumption of soy products in the United States is significantly increasing. There are frozen edamame imported from Asia in order to meet the market demand in the United States (Mebratu et al., 1991). It has been found that soybeans contain high amounts of isoflavones. Isoflavones are a sub-class of flavonids, and these nutraceuticals are biologically active non-nutrients. Isoflavones are also known as phytoestrogens, which are plant secondary metabolites that have a highly similar chemical structure to animal estrogen (Setchell and Adlercreutz, 1988). This similarity permits isoflavones to have interaction with estrogen receptors in mammalian metabolism and produce phyto-protectant effects to hormone related diseases such as osteoporosis, cardiovascular diseases, and breast cancer (Setchell and Cassidy, 1999; Messina et al., 2006).

Commodity soybeans are harvested at the R8 (reproductive stage eight) development stage, at which time the matured seeds and pods are field dried on the plants (Fehr and

Caviness, 1977). Edamame, also known as edible soybean or vegetable soybean, are harvested at the R6 stage in which the seeds have just fully filled the pods and have not proceeded to the drying stage, so the growing season is shorter than commodity soybeans (Fehr and Caviness, 1997; Lumpkin, 1993). It is critical for edamame to be harvested at the right time to ensure consumer satisfaction.

Edamame exhibits several differences from commodity soybeans, including larger seed size, milder taste, increased tenderness, and better digestibility (Born, 2006). There are large differences between processing edamame and commodity soybeans. Edamame is consumed as a fresh vegetable and is highly perishable after harvest, unlike field dried soybeans. Edamame also does not need to be threshed before going to market because the consumers can simply boil edamame in salt water and shell the pods before eating. Nevertheless, there is an increasing trend toward consumer preference for shelled fresh frozen edamame in vegetable mixes.

A long term goal of our program is to develop more nutritious edamame cultivars for improving human health. Towards this goal, statistical model development for isoflavone extraction from edamame has been examined in the previous chapter. The objective of this research was to validate the model predictions in order to evaluate the usefulness and understand the reliability of the models. This experiment was designed to investigate errors between the prediction values and the actual extraction results.

2. Materials and Methods

2.1. Sample Processing

Soybean line NUTRIVEG Soy6407 was chosen for this experiment. This line is a new cultivar developed by the Tennessee Agricultural Experiment Station for edamame production. During 2006, the seeds were field grown in Sequatchie silt loam soil at the Plant Sciences Unit of the East Tennessee Research and Education Center in Knoxville, Tennessee. NUTRIVEG Soy6407 edamame was collected at the R6 stage in which the green seeds fill the pod cavity (Fehr and Caviness, 1977). All pods were stored at 4°C, and shelling was done within one day after harvesting with seeds being immediately frozen at -80°C until analysis. The shelled edamame seeds were freeze-dried prior to extraction, and the freeze-drying procedure was controlled at -40°C. The dry weight was measured after freeze-drying in order to make sure the weight difference was 2% or less at the end of the drying process. Dry edamame seeds were ground with a Knifetec 1095 Sample Mill (Foss Tecator, Hogana Sweden) for two 10-sec pulses. Samples were preserved with ice before grinding, water cooled at 4°C during grinding and returned to ice after grinding to avoid isoflavone degradation. The ground edamame were homogenized to minimize differences among seeds. A uniform sample was obtained and utilized for all extractions.

2.2. Chemicals and Solvents

Water was supplied by a Milli-Q purifier from Millipore (Bedford, MA, USA). The isoflavone standards were purchased from LC Laboratory (Woburn, MA, USA), and internal standard apigenin was obtained from Sigma (St. Louis, MO, USA).

2.3. Accelerated Solvent Extraction Procedure

We used a modified procedure from Rostagno et al. (2004). The isoflavone extraction was performed by a DIONEX Accelerated Solvent Extractor (ASE). A disk-shaped cellulose filter (diameter: 1.983cm) was pushed down to the bottom of an ASE 11-mL stainless extraction cell to protect the metallic filter. A sample of 0.1 ± 0.001g freeze dried, finely ground edamame seed was weighed and transferred to the ASE extraction cell. The edamame sample was placed between layers of Ottawa Sand (Fisher Scientific Co.) in the extraction cell (Figure 3). The total volume of Ottawa Sand was controlled by measuring 10 g for each extraction. All samples were spiked with 300µL of the internal standard apigenin. The cell was hand-tightened and loaded into the ASE instrument.

In this model validation experiments, extraction solvents were the same as the second set of extractions in previous chapter which included six water miscible organic solvents (60% MeOH, 60% EtOH, 60% MeCN, 80% MeOH, 80% EtOH, 80% MeCN) and ultra purified water. Temperature, pressure, length of extraction and solvent to volume ratio were

adjusted for different combinations, and each combination was replicated three times (Table 20).

After the extraction was completed, the cell was purged with nitrogen for 300 sec to force all the extracts into 60-mL glass vials with Teflon-coated rubber caps. As soon as the extraction was completed, the extracts were filtered through a nylon 0.45-µm syringe filter, stored in 2-mL cryovials and frozen at -80°C until High Performance Liquid Chromatography (HPLC) analysis.

2.4. Validation Point Selection

The limits of the developed models were within a range of time from 10 to 90 min and temperature from ambient (27°C) to 97°C. Using to the trend for predicted total isoflavone content, three points were selected on the surface portion of highest efficiency boundary conditions, and two additional points outside the region covered by previous data were selected to validate the predictions.

2.5. High Performance Liquid Chromatography (HPLC) Analysis

We used a modified procedure from Griffin and Collison (2001) and Charron et al. (2005). Soybean isoflavone extracts were analyzed by an Agilent HPLC 1200 series equipped with a binary pump, degasser, autosampler thermostat, XDB-C18 reverse phase

column (1.8μm, 4.6x50mm) and diode-array detector (DAD). Solvent A was 0.1% (v/v) acetic acid in H_2O, and solvent B was 0.1% (v/v) acetic acid in acetonitrile. This isoflavone analysis method used 5.0 μL for the injection volume with a needle wash between every sample, 40°C for column temperature, 260nm for UV lamp wavelength and a flow rate of 0.8 mL per min. The solvent gradient was 88:12 (A:B) at sample injection and shifted after 15 min to 76:24 (A:B). At 20 min the column was washed with a ratio of 10:90 (A:B) for 5 min and equilibrated with the initial solvent ratio for 2 min (Table 7). The complete elapsed time for one sample was 27 min. Each isoflavone was evaluated and auto labeled by comparing the retention time with authentic standards of daidzin (Di), glycitin (Gly), genistin (Gi), malonyl daidzin (MDi), malonyl glycitin (MGly), acetyl daidzin (AcDi), malonyl genistin (MGi), daidzein (De), glycitein (Gle), acetyl genistin (AcGi) and genistein (Ge) (Figure 4). Each isoflavone concentration was calculated with individual response factors (Table 8).

2.6 Statistics

All the statistics were performed with SAS 9.1 (SAS Institute, 2002). The predicted errors came from the distance between observed data values and model surface from Chapter II. Statistical analysis was performed for the predicted errors of individual isoflavones and total isoflavone content by using the MEANS procedure. The MEANS

procedure calculated means, standard deviations, minimums and maximums for actual differences, absolute differences and percentage of absolute differences (Figure 11). Actual differences could be positive or negative. When the actual differences were positive, the observed values were above the model surface; when the actual differences were negative, the observed values were below the model surface. The absolute value of the actual difference were used to quantity errors regardless their sign. The percentage of absolute difference was used to express the quantity differences relative to the prediction value. An observed value of zero could not be used to calculate the percentage of absolute difference because it was the denominator in the equation. A lower percentage of absolute difference means the model predictions were more accurate.

3. Results and Discussion

Six isoflavones were detected in HPLC analysis, and they were daidzin, glycitin, genistin, malonyl diadzin, malonylglycitin and malonyl genistin. When 60% EtOH solvent was used to extract 0.1g of edamame, daidzin was not observed. The means of observed values for glycitin, genistin, malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavone content were 4.52, 2.14, 53.89, 28.25, 44.22 and 133.02μg per 0.1g of dry edamame, respectively (Table 21). The means of actual difference were 0.98, 0.69, -6.19, 6.16, 5.08, and 5.77μg per 0.1g of dry edamame, respectively. The means of

percent absolute difference were 23.58%, 38.71%, 12.30%, 21.06%, 11.27% and 4.60%, respectively (Table 22). Besides genistin, all other observed isoflavones had approximately 20% or lower percentage of absolute differences. The extraction of total isoflavone content with 60% EtOH had the most accurate prediction within all isoflavone compounds for the entire validation experiment (Figure 13).

When 80% EtOH solvent was used to extract 0.1g of edamame, the means of observed values for daidzin, glycitin, genistin, malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavone content were 6.06, 6.70, 7.28, 40.85, 20.35 and 115.99µg per 0.1g of dry edamame, respectively (Table 23). The means of actual difference were 4.18, 4.25, 4.74, -15.93, -7.35, -7.13 and -15.85µg per 0.1g of dry edamame, respectively. The means of percent absolute difference were 72.01%, 59.33%, 58.09%, 41.96%, 40.58%, 24.17% and 14.67%, respectively (Table 24).

When 60% MeCN solvent was used to extract 0.1g of edamame, daidzin was also not observed. The means of observed value for glycitin, genistin, malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavone content were 3.39, 1.82, 60.39, 32.51, 40.93, 139.03µg per 0.1g of dry edamame, respectively (Table 25). The means of actual difference were 2.11, 1.36, -2.57, 2.26, 0.39 and 3.54µg per 0.1g of dry edamame, respectively. The means of percent absolute difference were 61.64%, 74.45%, 12.63%, 15.21%, 9.54% and 10.87%, respectively (Table 26).

When 80% MeCN solvent was used to extract 0.1g of edamame, daidzin, glycitin, genistin were not observed. The means of observed value for malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavone content were 46.04, 22.62, 38.74, and 111.37µg per 0.1g of dry edamame, respectively (Table 27). The means of actual difference were 25.14, 19.96, -7.13 and 38.95µg per 0.1g of dry edamame, respectively. The means of percent absolute difference were 51.37%, 73.09%, 20.92% and 33.25%, respectively (Table 28).

When 60% MeOH solvent was used to extract 0.1g of edamame, daidzin was not observed. The means of observed value for glycitin, genistin, malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavone content were 4.36, 2.00, 50.05, 28.77, 40.48 and 125.67µg per 0.1g of dry edamame, respectively (Table 29). The means of actual difference were -1.10, 0.78, 0.29, -10.93, 5.72, 1.13 and -4.10µg per 0.1g of dry edamame, respectively. The means of percent absolute difference were 17.53%, 15.29%, 22.39%, 19.47%, 6.12% and 4.88%, respectively (Table 30).

When 80% MeOH solvent was used to extract 0.1g of edamame, daidzin was not observed. The means of observed value for glycitin, genistin, malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavone content were 5.07, 2.76, 52.32, 25.65, 44.98 and 130.79µg per 0.1g of dry edamame, respectively (Table 31). The means of actual difference were -2.22, 0.07, 0.08, -10.89, 1.10 6.12 and -5.88µg per 0.1g of dry edamame,

respectively. The means of percent absolute difference were 15.58%, 19.02%, 21.13%, 10.58%, 13.57% and 5.82%, respectively (Table 32). In both 60% and 80% MeOH models, the percentage of absolute difference for each isoflavone and total isoflavones was approximately 20% or lower. The results in these two models are repeatable within an acceptable range of variation which implied that these two models were robust; therefore, the models proved to be reliable and valid.

When H_2O was used to extract 0.1g of edamame, daidzin, glycitin, genistin were not observed. The means of observed value for malonyl daidzin, malonyl glycitin, malonyl genistin and total isoflavone content were 20.53, 12.23, 16.42 and 49.18µg per 0.1g of dry edamame, respectively (Table 33). The means of actual difference were -0.41, -2.71, -0.55, -44.45, -12.55, -25.18 and -85.85µg per 0.1g of dry edamame, respectively. The means of percent absolute difference were 249.73%, 141.75%, 183.99% and 206.32%, respectively (Table 34). The observed value for each isoflavone and total isoflavone content were all below the predictions. Water also had very large percentages of absolute difference in all cases; therefore, it is not practical to use water as an extraction solvent.

Acetyl genistin was first seen during the model validation experiment. The statistical results showed that temperature had significant differences ($P<0.0001$), with an $R^2= 0.993$. Acetyl genistin was extracted at 6.98µg per 0.1g of dry edamame by 80% EtOH with a temperature of 120°C and at 9.38µg per 0.1g of dry edamame for 80% MeCN

with a temperature of 120°C.

Future research can be conducted with soybean cultivars with high isoflavone content in order to increase the availability of isoflavones in pharmaceutical markets. The industry will need to evaluate the cost of extraction and environmental concerns of using organic solvents. Soybean seeds are just one source of isoflavones. Klejdus et al. (2005) found that the highest isoflavone content was in the root of soybean plants. It would be interesting to apply this statistical approach from model development to prediction to validation for isoflavone extraction from the roots of soybean plants.

4. Conclusion

This statistical analysis compared current data to the prediction from historical data. The experimental purpose was to understand the accuracy of the models in order to make the models more practical. The models for both 60% and 80% MeOH were highly reliable, and they produced the most accurate predictions among all models, which indicated both solvents provided the most consistent extraction. In the models of six organic solvents, malonyl daidzin, malonyl glycitin, and malonyl genistin were better validated than daidzin, glycitin and genistin. Malonyl genistin had the best validation among all isoflavone compounds, which means that the predictions for malonyl genistin were closest to the observed values.

Literature Cited

Born, H. Edamame: Vegetable Soybean. ATTRA-National Sustainable Agriculture Information Service. 2006.

Charron, C. S., Allen, F. L., Johnson, R. D., Pantalone, V. R., and Sams, C. E., Correlations of oil and protein with isoflavone concentration in soybean, J. Agri. Food Chem. 53: 7128-7135, 2005.

Fehr, W.R., and Caviness C.E. Stages of development decriptions for soybeans. Crop Sci. 80: 1-12. 1971.

Griffith A. P. and Collison M. W., Improved methods for the extraction and analysis of isoflavones from soy-containing foods and nutritional supplements by reversed-phase high-performance liquid chromatography and liquid chromatography–mass spectrometry, J. Chromatography, A 913: 397–413, 2001.

Klejdus, B., Mikelova, R., Petrlove, J., Potesil, D., Adam, V., Stiborova, M., Hodek, P., Vacek, J., Kizek, R., and Kuban, V., Evaluation of Isoflavone Aglycon and Glycoside Distribution in Soy Plants and Soybeans by Fast Column High -Performance Liquid Chromatography Coupled with a Diode-Array Detector. J. Agric. Food Chem. 53: 5848-5852. 2005.

Lumpkin, T.A., J.C. Konovsky, K.J. Larson, and D.C. McClary. Potential new specialty crops from Asia: Azuki bean, edamame soybean, and astragalus. 45-51. In: J. Janick and J.E. Simon (eds.), New crops. Wiley, New York. 1993.

Mebrahtu, T.A., Mohamed A., and Mersie, W. Green pod and architectural traits of selected vegetable soybean genotypes. Journal of Production in Agriculture. 4(3): 395-399. 1991.

Messina, M., Mc Caskill-Stevens, and W., Lampe, J.W. Addressing the soy and breast cancer relationship: review, commentary, and workshop proceeding. J. National Cancer Institute. 98: 1275-1284. 2006.

Rostagno, M. A., Palma M., and Barroso S. G.., Pressurized liquid extraction of isoflavones from soybeans. Analytica Chemica Acta, 522(2): 169-177. 2004.

SAS Institute Inc. SAS guide for personal computers, 9.1 edn. SAS institute, Cary, NC. 2002.

Setchell, K.D.R. & Adlercreutz, H. Mammalian lignans and phytooestrogens. Recent studies on their formation, metabolism and biological role in health and disease. In: Role of the Gut Flora in Toxicity and Cancer (Rowland, I. R., ed.), 315–345, Academic Press, London, UK. 1988.

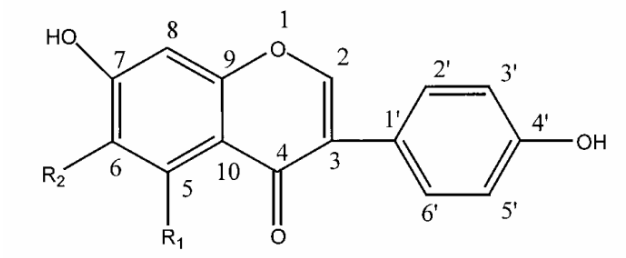
Setchell K. and Cassidy A. Dietary isoflavones: biological effects and relevence to human health. J. of Nutri. 129: 758-767 1999.

Appendix A: Figures

Kingdom	*Plantae* (Plants)
Subkingdom	*Tracheobionta* (Vascular plants)
Superdivision	*Spermatophyta* (Seed plants)
Division	*Magnoliophyta* (Flowering plants)
Class	*Magnoliopsida* (Dicotyledons)
Subclass	*Rosidae*
Order	*Fabales*
Family	*Fabaceae* (Pea family, aka Legume family)
Genus	*Glycine* Willd. (soybean)
Species	*Glycine max* (L.) Merr. (soybean)

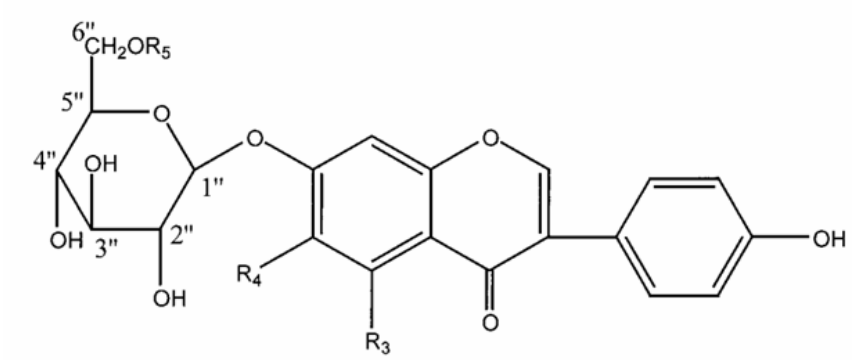
Figure 1: Soybean classification from Kingdom Plantae to Species Glycine max (USDA, 2008).

Aglycones:



R_1	R_2	Compounds
H	H	Daidzein
OH	H	Genistein
H	OCH_3	Glycitein

Glucosides:



R_3	R_4	R_5	Compounds
H	H	H	daidzin
OH	H	H	genistin
H	OCH_3	H	glycitin
H	H	$COCH_3$	6”-O-acetyldaidzin
OH	H	$COCH_3$	6”-O-acetylgenistin
H	OCH_3	$COCH_3$	6”-O-acetylglycitin
H	H	$COCH_2COOH$	6”-O-malonyldaidzin
OH	H	$COCH_2COOH$	6”-O-malonylgenistin
H	OCH_3	$COCH_2COOH$	6”-O-malonylglycitin

Figure 2: Chemical structure for isoflavone aglycones and glucosides.

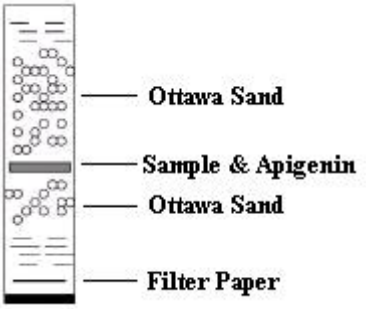


Figure 3: Example of filling the ASE extraction cell.

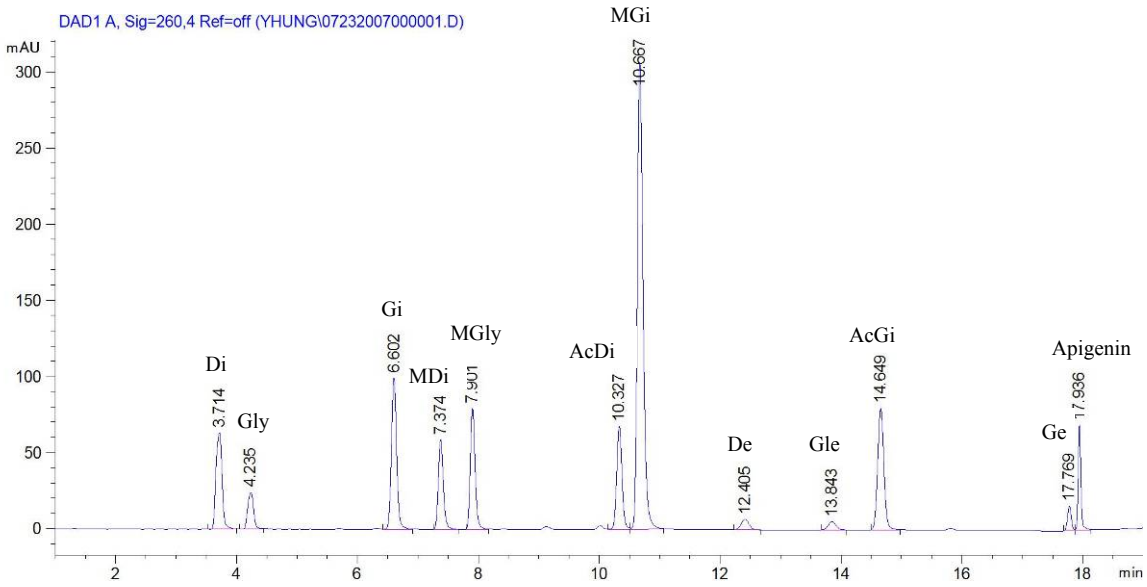


Figure 4: HPLC Chromatogram showing the retention time (min) for 11 isoflavone standards and internal standard apigenin.

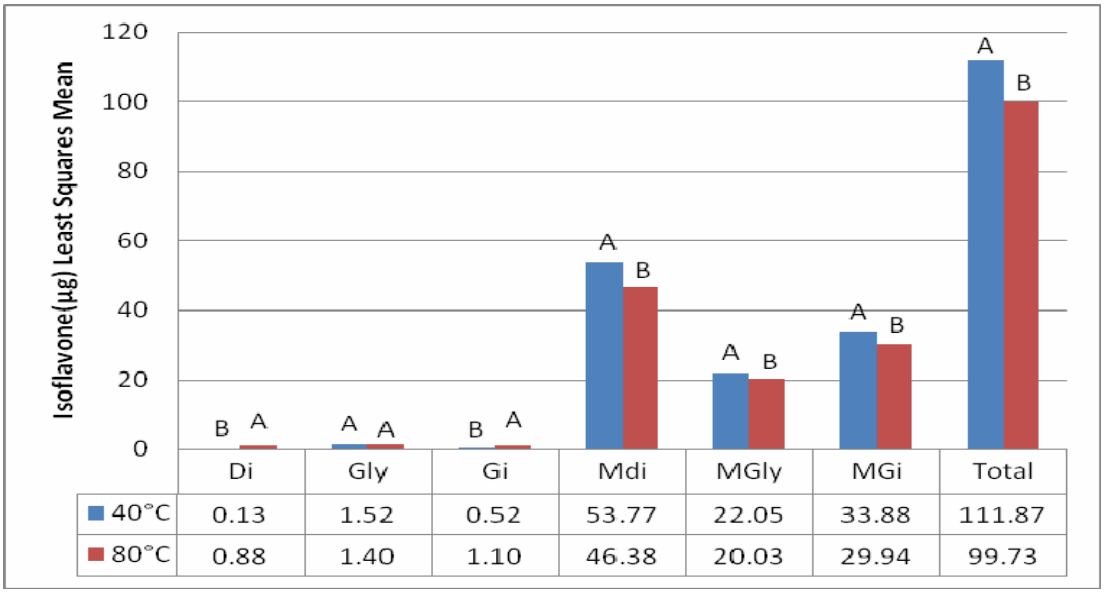


Figure 5: Least squares means for six observed isoflavones extracted from 0.1g of dry edamame at 40°C and 80°C.

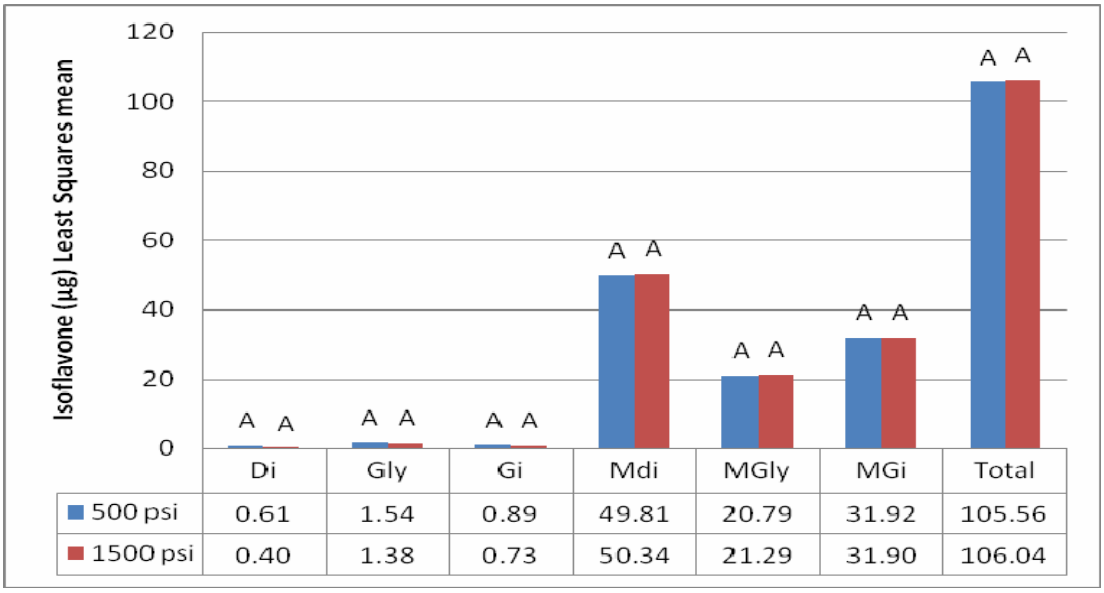


Figure 6: Least squares means for six observed isoflavones extracted from 0.1g of dry edamame in 500 psi and 1500 psi.

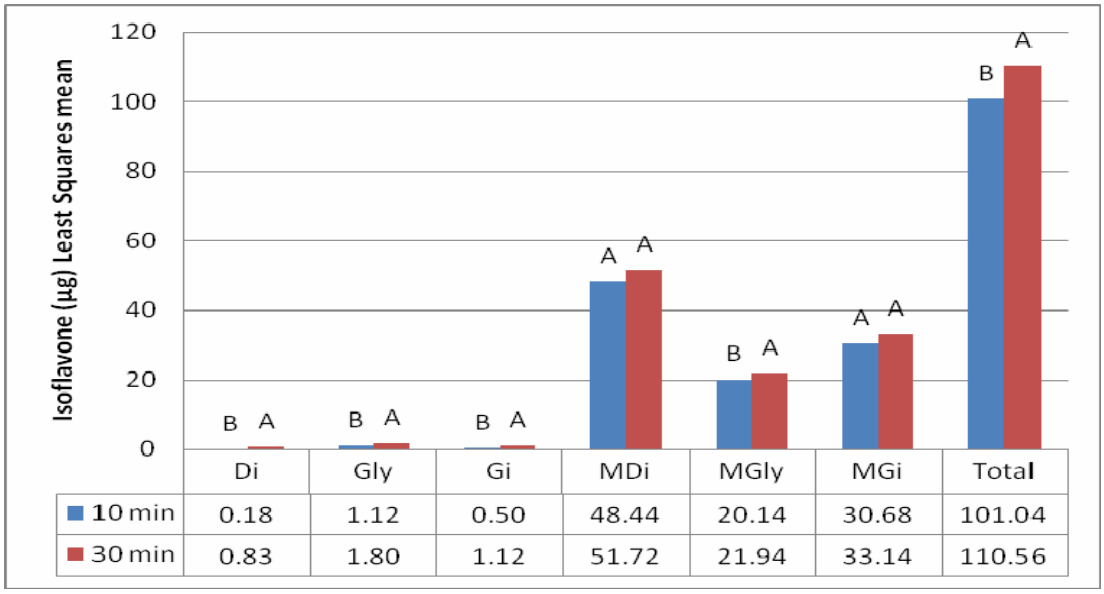


Figure 7: Least squares means for six observed isoflavones extracted from 0.1g of dry edamame for 10 min and 30 min.

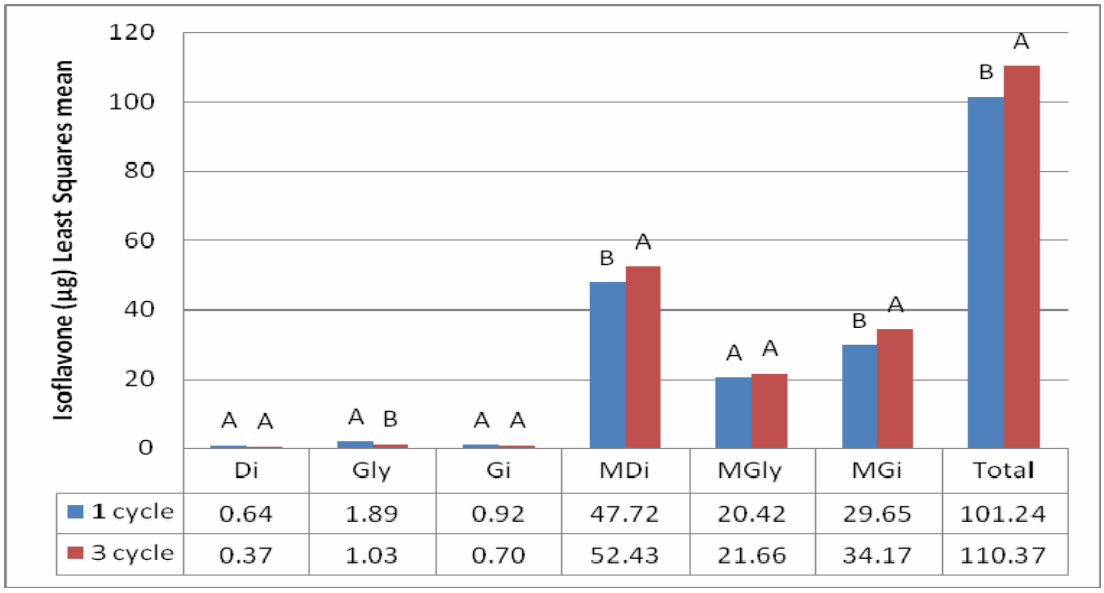
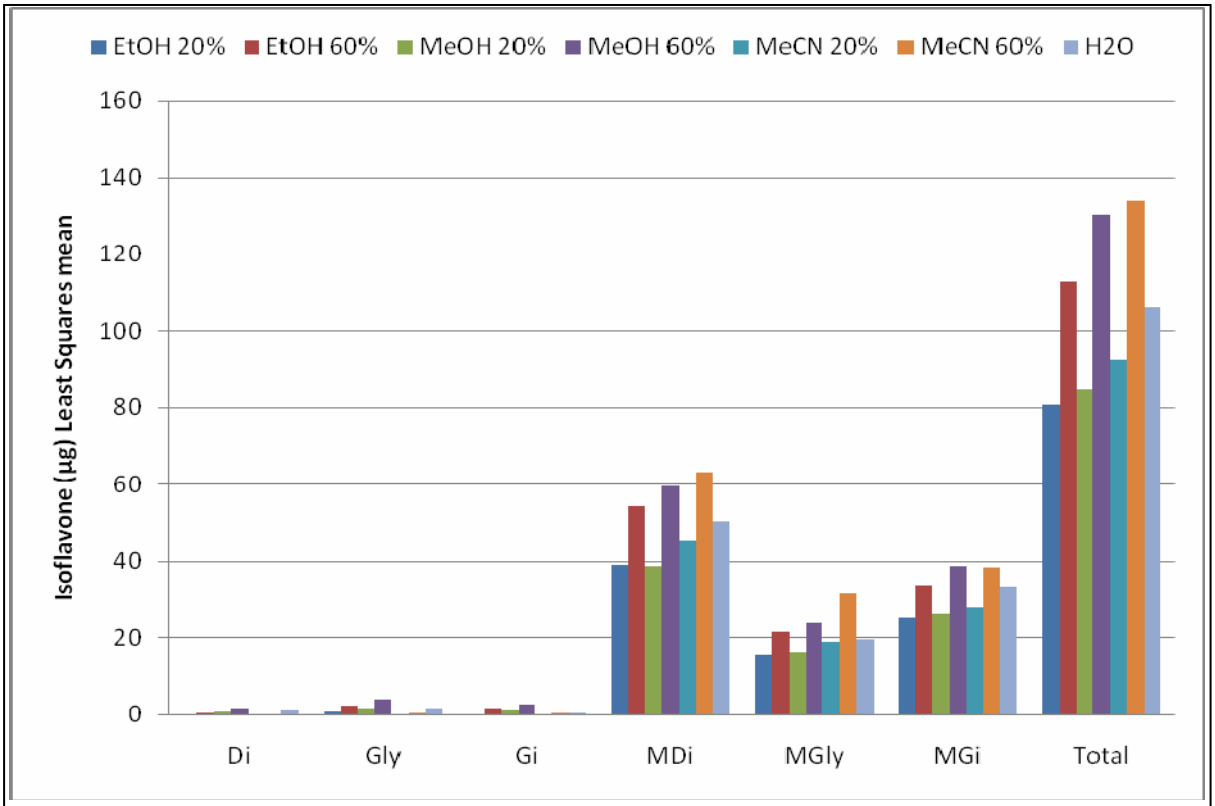


Figure 8: Least squares means for six observed isoflavones extracted from 0.1g of dry edamame for 1 cycle and 3 cycles.



Di		Gly		Gi		MDi		MGly		MGi		Total	
0.10	CD	0.62	C	0.10	C	39.04	E	15.49	E	25.29	C	80.64	D
0.25	CD	2.12	B	1.25	B	54.21	BC	21.25	BC	33.68	B	112.75	B
0.55	BC	1.67	B	1.11	B	38.81	E	16.32	DE	26.02	C	84.47	D
1.60	A	3.89	A	2.55	A	59.61	AB	23.89	B	38.66	A	130.20	A
0.00	D	0.19	C	0.12	C	45.53	DE	18.99	CD	27.84	C	92.66	CD
0.00	D	0.23	C	0.23	C	63.20	AB	31.72	A	38.34	AB	133.71	A
1.04	B	1.51	B	0.32	C	50.15	CD	19.60	C	33.54	B	106.17	BC

Figure 9: Least squares means for six observed isoflavones extracted from 0.1g of dry edamame with different percentage of organic solvent and water. Means with no common letter differ by LSD at P < 0.05.

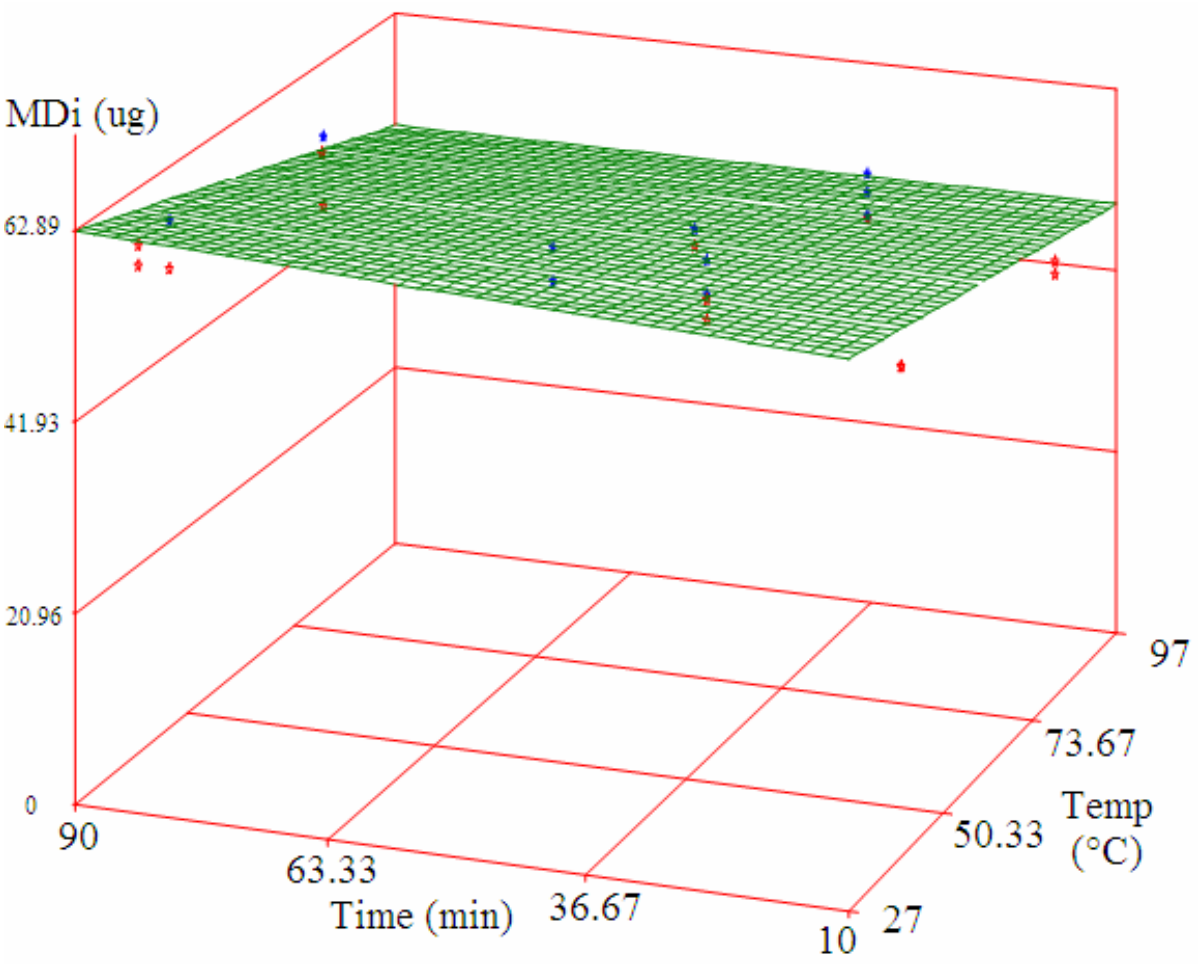


Figure 10: Three dimensional surface plots for malonyl daidzin (MDi) in 0.1g of edamame with EtOH 60%. The blue (above surface) and red stars (below surface) were the actual data points.

$$Actual\ difference = Observed - Prediction$$

$$\downarrow$$

$$Absolute\ difference = \left| Observed - Prediction \right|$$

$$\downarrow$$

$$Percentage\ of\ \ absolute\ difference = \frac{\left| Observed - Prediction \right|}{Observed} \times 100\%$$

Figure 11: Equations for calculating actual difference, absolute difference and Percentage of absolute difference.

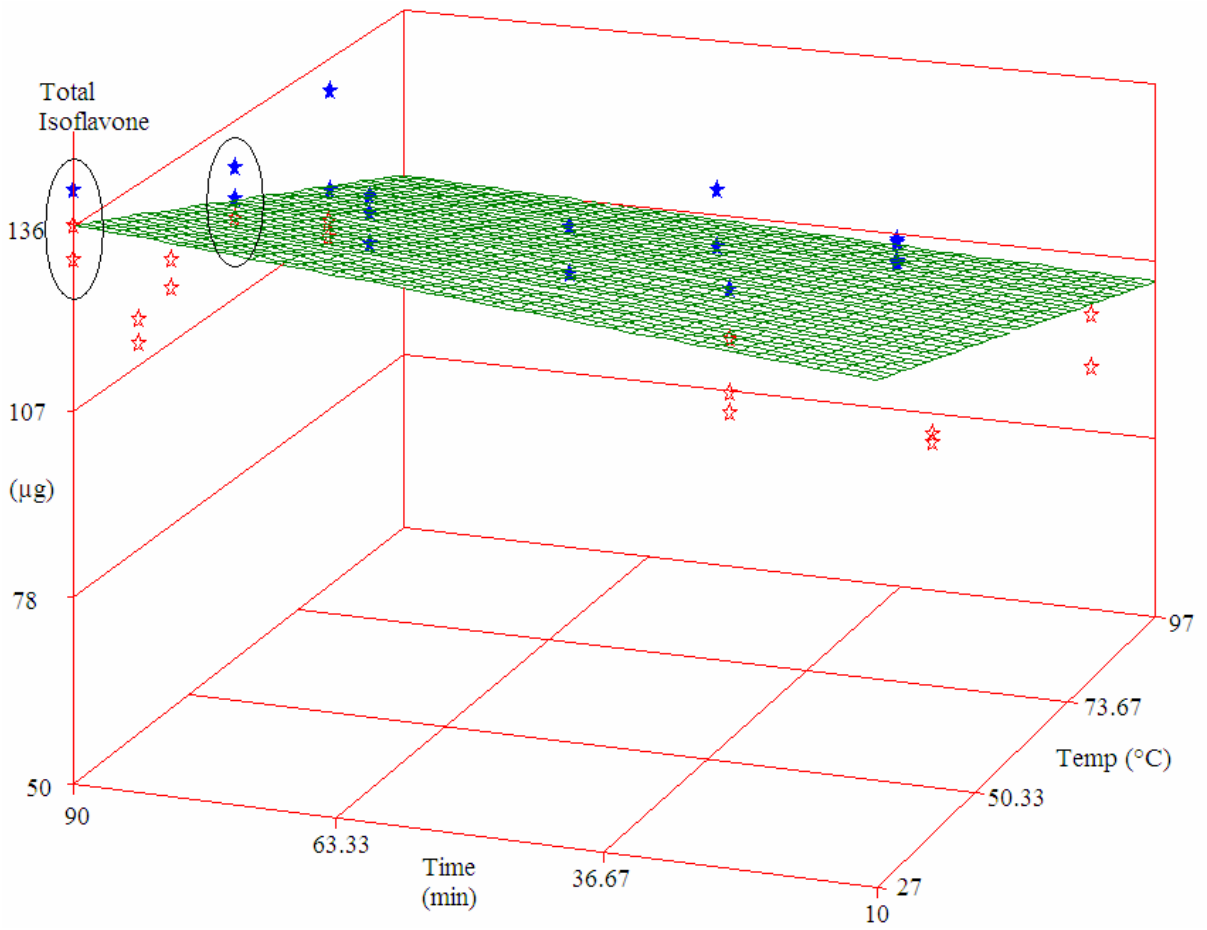


Figure 12: Three dimensional surface plots for total isoflavone content extracted from 0.1g of edamame with EtOH 60%. The blue stars (above surface) and red stars (below surface) were the actual data points. The circled stars are the validation points.

Appendix B: Tables

Table 1: Edamame lexicon with definitions (Krinsky et al., 2006).

Aromatic	Taste	Feeling factor
Raw bean	Salty	Astringent
Cooked bean	Sweet	Umami
Green complex	Sour	Metallic
Fruity complex	Bitter	
Nutty/Almond		
Brothy		
Sulfur		

Raw bean: characteristic of raw soybeans/legumes
Cooked bean: characteristic of cooked soybeans/legumes
Green complex: characteristic of freshly cut twigs/grass
Fruity complex: aromatics associated with a mixture apple/pear/tropical
Nutty/Almond: aromatic of nuts, legume-like character
Brothy: aromatic/taste associated with boiled meat/soup/stock
Sulfur: aromatics associated with hydrogen sulfide, rotten egg
Salty: taste on tongue stimulated by sodium salt (NaCl)
Sweet: taste on tongue stimulated by sugar/high potency sweetener
Sour: range from fermented vegetables to lactic or spoiled bacteria
Bitter: taste on tongue stimulated by caffeine/quinine/alkaloid
Astringent: feeling factor on oral cavity; puckering/dry; tannins/alum
Umami: feeling factor elicited from glutamate/aspartate/ribonucleotides
Metallic: flat chemical feeling factor stimulated by coins on tongue

Table 2: Amount of individual isoflavones (µg/g) and total isoflavone content extracted at different temperatures (Rostagno et al., 2004).

Temp(°C)	Daidzin	Daidzein	Glycitin	Glycitein	Malonylgenistin	Genistin	Genistein	Total
60	165.41	n.d.	66.33	n.d.	386.48	294.94	10.23	923.39
80	176.19	n.d.	72.09	n.d.	396.74	337.78	10.31	993.11
100	202.34	n.d.	91.83	n.d.	445.70	383.34	10.68	1133.89
150	290.49	n.d.	128.63	n.d.	203.35	616.08	10.59	1249.14
200	117.82	108.66	53.64	90.12	n.d.	382.72	153.74	906.7

n.d.= not detected;

0.5 g of freeze-dried soybean with three extraction cycles of 5 minute

Table 3: Amount of individual isoflavones (µg/g) and total isoflavone content extracted under different pressures (Rostagno et al., 2004).

Pressure	Daidzin	Glycitin	Malonylgenistin	Genistin	Total
1470 psi	202.34	91.83	445.70	383.34	1123.20
2940 psi	203.06	84.97	447.43	384.30	1119.76

70% EtOH, 100°C, 0.5 g of freeze-dried soybean, and three extraction cycles of 5 min.

Table 4: Amount of individual isoflavones (µg/g) and total isoflavone content extracted by using different sample sizes (Rostagno et al., 2004).

Sample Size	Daidzin	Glycitin	Malonylgenistin	Genistin	Total
0.50 g	202.34	91.83	445.70	383.34	1123.20
0.25 g	205.38	94.08	448.74	398.57	1146.77
0.10 g	213.46	94.87	462.07	439.97	1210.47
0.05 g	225.44	97.39	475.75	477.53	1276.11

70% EtOH, 100°C, 1470psi, and three static extraction cycles of 5 min.

Table 5: Amount of individual isoflavones (µg/g) and total isoflavone content extracted with different extraction time per cycle (Rostagno et al., 2004).

Time per Cycle	Daidzin	Glycitin	Malonylgenistin	Genistin	Total
5 min	213.46	94.87	462.07	439.97	1210.47
7 min	236.81	102.38	447.61	510.62	1327.42
10 min	239.41	101.84	483.65	507.57	1332.47

70% EtOH, 100°C, 1470psi, 0.1 g of freeze-dried soybean, and three extraction cycles

Table 6: Amount of individual isoflavones (µg/g) and total isoflavone content extracted by using different extraction cycles (Rostagno et al., 2004).

Cycle Length	Cycles	Daidzin	Glycitin	Malonylgenistin	Genistin	Total
7 min	1	181.29	44.85	348.41	236.25	810.80
	2	189.59	77.16	428.50	404.66	1099.91
	3	237.53	103.10	477.87	510.87	1329.37
10 min	2	203.78	87.19	445.41	451.99	1187.87

70% EtOH, 100°C, 1470psi, and 0.1 g of freeze dried soybean

Table 7: HPLC gradient settings for binary pump.

Time (min)	A:B	Flow rate (mL/min)
0	88:12	0.800
15	76:24	0.800
20	10:90	0.800
25	88:12	0.800

Solvent gradient ratio (A:B) over time and the flow rate was set for 0.800 ml/min

A: acetic acid in H_2O, B: acetic acid in acetonitrile

Table 8: Response Factor (RF) for each isoflavone and the equation for calculating the isoflavone concentration.

	Di	Gly	Gi	MDi	MGly	MGi	AcDi	AcGi	De	Gle	Ge
RF	33.6	44.7	60.0	13.3	20.6	33.8	23.7	32.9	47.4	44.5	50.5

$$Isoflavone\ content = \left[\frac{Isoflavone\ area}{RF}\right] \times (mL\ of\ \ Solvent)$$

Each isoflavone area was integrated from the peak area under the curve in HPLC chromatogram

Table 9: Five variables selected for preliminary extraction.

Variables	Level
Temperature	40°C, 80°C
Pressure	500psi, 1500psi
Extraction time per cycle (min)	10 min, 30 min
Static cycle	1 cycle, 3 cycle
Solvent	60% (MeOH, EtOH and MeCN) 20% (MeOH, EtOH and MeCN) H_2O
Total preliminary extraction = 2^4 × 7 = 112 combinations	

All the combinations were used to extract 0.1g of dry edamame with accelerated solvent extractor

Table 10: R^2 values and ANOVA test among variables and their interaction.

Variable \ Isoflavone▲	Di	Gly	Gi	MDi	MGly	MGe	Total
T	****	NS	****	***	*	**	**
P	NS	NS	NS	NS	NS	NS	NS
T×P	NS	NS	NS	NS	NS	NS	NS
E	****	***	****	NS	*	NS	*
T×E	****	NS	**	NS	NS	NS	NS
P×E	NS	NS	NS	NS	NS	NS	NS
C	NS	****	NS	*	NS	***	*
T×C	NS	*	NS	NS	NS	NS	NS
P×C	NS	NS	NS	NS	NS	NS	NS
E×C	NS	NS	NS	NS	NS	NS	NS
S	****	****	****	****	****	****	****
T×S	**	NS	NS	NS	*	NS	NS
P×S	NS	NS	NS	NS	NS	NS	NS
E×S	**	NS	NS	NS	NS	NS	NS
C×S	NS	*	*	***	**	**	**
R^2	0.721	0.791	0.800	0.689	0.746	0.657	0.687

T: temperature, P: pressure, E: extraction time per cycle, C: number of cycles, S: solvent;

****, ***, **, * significant at $P < 0.0001$, $P < 0.001$, $P < 0.01$, and $P < 0.05$

NS: non-significant

▲: Isoflavones are defined on page viii

Table 11: Five variables were selected for second set of extractions.

Variables	Level
Temperature	27°C, 60°C
Pressure	500psi, 1000psi
Extraction time per cycle (min)	20 min, 40 min
Static cycle	2 cycles
Solvent	60% MeOH, EtOH, MeCN, 80% MeOH, EtOH, MeCN,
Total preliminary extractions = 2 × 2 × 2 × 1 × 7 = 48 combinations	

All the combinations were used to extract 0.1g of dry edamame with accelerated solvent extractor

Table 12: Model parameter estimates for daidzin (R^2 = 0.520).

Solvent	a	b	c	Temperature × Time
EtOH 20%	-0.0090	-0.0221*	0.6830	0.003
EtOH 60%	-0.0098	-0.0139	1.0271	0.003
EtOH 80%	-0.0300	-0.0279	3.4125*	0.003
MeCN 20%	-0.0140	-0.0209	0.8380	0.003
MeCN 60%	-0.0163	-0.0196	0.9380	0.003
MeCN 80%	-0.0209	-0.0152	0.9113	0.003
MeOH 20%	0.0133	-0.0144	-0.5146	0.003
MeOH 60%	0.0207	-0.0083	-0.0125	0.003
MeOH 80%	-0.0019	-0.0178	2.6221	0.003
H_2O	0.0255*	0.0026	-1.4077	0.003

*: significant at P < 0.05

$$\begin{cases} Isoflavone = Solvent \times [a \times Temp + b \times Total\ Time + c] + Temp \times Total\ Time \\ Total\ Time = Extraction\ Time \times Number\ of\ Cycles \end{cases}$$

Table 13: Model parameter estimates for glycitin (R^2 = 0.653).

Solvent	a	b	c	Temperature × Time
EtOH 20%	-0.0086	-0.0244	1.5726	0.002
EtOH 60%	-0.0304*	-0.0080	4.2739*	0.002
EtOH 80%	-0.0172	0.0088	2.4067	0.002
MeCN 20%	-0.0187	-0.0156	1.4048	0.002
MeCN 60%	-0.0331*	-0.0125	2.6586*	0.002
MeCN 80%	-0.0134	-0.0098	0.5861	0.002
MeOH 20%	0.0051	0.0001	0.8196	0.002
MeOH 60%	-0.0051	-0.0081	3.9903*	0.002
MeOH 80%	-0.0473	0.0105	5.1394	0.002
H_2O	-0.0240	-0.0137	2.9757*	0.002

*: significant at P < 0.05

$$\begin{cases} Isoflavone = Solvent \times [a \times Temp + b \times Total\ Time + c] + Temp \times Total\ Time \\ Total\ Time = Extraction\ Time \times Number\ of\ Cycles \end{cases}$$

Table 14: Model parameter estimates for genistin (R^2 = 0.649).

Solvent	a	b	c	Temperature × Time
EtOH 20%	-0.0013	-0.0105	0.2197	0.0002
EtOH 60%	0.0059	0.0011	0.7095	0.0002
EtOH 80%	0.0066	-0.0034	1.4296	0.0002
MeCN 20%	-0.005	-0.0106	0.1949	0.0002
MeCN 60%	-0.0157	-0.0024	1.3751*	0.0002
MeCN 80%	-0.0094	-0.0068	0.4069	0.0002
MeOH 20%	0.0173	0.0116	-0.7734	0.0002
MeOH 60%	0.0269*	0.0009	0.3895	0.0002
MeOH 80%	0.0156	0.0071	0.8689	0.0002
H_2O	-0.0119	-0.0144	1.2188	0.0002

***: significant at P < 0.05**

$$\begin{cases} Isoflavone = Solvent \times [a \times Temp + b \times Total\ Time + c] + Temp \times Total\ Time \\ Total\ Time = Extraction\ Time \times Number\ of\ Cycles \end{cases}$$

Table 15: Model parameter estimates for malonyl daidzin (R^2 = 0.708).

Solvent	a	b	c	Temperature × Time
EtOH 20%	-0.2437	0.2633*	44.9125*	-0.0007
EtOH 60%	-0.1220	0.0734	61.3839*	-0.0007
EtOH 80%	0.1092	0.1999	44.8162*	-0.0007
MeCN 20%	0.0936	0.2789*	30.5421*	-0.0007
MeCN 60%	0.0354	0.0209	60.7405*	-0.0007
MeCN 80%	0.1252	0.0208	13.2268	-0.0007
MeOH 20%	-0.2418	0.3268*	42.0327*	-0.0007
MeOH 60%	-0.1190	-0.0662	69.6176*	-0.0007
MeOH 80%	-0.0535	0.0414	64.3467*	-0.0007
H_2O	-0.3924*	-0.0840	78.8508*	-0.0007

*: significant at P < 0.05

$$\begin{cases} Isoflavone = Solvent \times [a \times Temp + b \times Total\ Time + c] + Temp \times Total\ Time \\ Total\ Time = Extraction\ Time \times Number\ of\ Cycles \end{cases}$$

Table 16: Model parameter estimates for malonyl glycitin (R^2 = 0.677).

Solvent	a	b	c	Temperature × Time
EtOH 20%	-0.0824	0.0801	16.7964*	0.002
EtOH 60%	-0.0449	-0.0416	26.5544*	0.002
EtOH 80%	0.0681	0.0517	18.3174*	0.002
MeCN 20%	0.0419	0.1019	11.9548*	0.002
MeCN 60%	0.0946	-0.0641	25.7015*	0.002
MeCN 80%	-0.0480	-0.0056	9.5159	0.002
MeOH 20%	-0.1599*	0.0804	22.2612*	0.002
MeOH 60%	-0.0521	-0.0061	26.4165*	0.002
MeOH 80%	-0.0147	-0.0223	26.5499*	0.002
H_2O	-0.1685*	-0.0349	30.7223*	0.002

***: significant at P < 0.05**

$$\begin{cases} Isoflavone = Solvent \times [a \times Temp + b \times Total\ Time + c] + Temp \times Total\ Time \\ Total\ Time = Extraction\ Time \times Number\ of\ Cycles \end{cases}$$

Table 17: Model parameter estimates for malonyl genistin (R^2 = 0.646).

Solvent	a	b	c	Temperature × Time
EtOH 20%	-0.0948	0.1916*	24.5296*	-0.0005
EtOH 60%	-0.0535	0.0850	35.8930*	-0.0005
EtOH 80%	0.0562	0.0546	38.3453*	-0.0005
MeCN 20%	0.0505	0.1895*	18.4312*	-0.0005
MeCN 60%	-0.0678	0.0435	43.1981*	-0.0005
MeCN 80%	0.0146	0.0020	42.1507*	-0.0005
MeOH 20%	-0.1052	0.2340*	24.1810*	-0.0005
MeOH 60%	-0.0720	-0.0004	43.4982*	-0.0005
MeOH 80%	-0.0301	0.0019	41.8867*	-0.0005
H_2O	-0.2100*	-0.0496	49.2856*	-0.0005

*: significant at $P < 0.05$

$$\begin{cases} Isoflavone = Solvent \times [a \times Temp + b \times Total\ Time + c] + Temp \times Total\ Time \\ Total\ Time = Extraction\ Time \times Number\ of\ Cycles \end{cases}$$

Table 18: Model parameter estimates for total isoflavone content (R^2 = 0.677).

Solvent	a	b	c	Temperature × Time
EtOH 20%	-0.4399	0.4780*	88.7138*	-0.0003
EtOH 60%	-0.2547	0.0961	129.8420*	-0.0003
EtOH 80%	0.1928	0.2836	108.7277*	-0.0003
MeCN 20%	0.1529	0.5231*	63.3658*	-0.0003
MeCN 60%	-0.0029	-0.0342	134.6120*	-0.0003
MeCN 80%	0.0479	-0.0144	66.7977*	-0.0003
MeOH 20%	-0.4712	0.6385*	88.0069*	-0.0003
MeOH 60%	-0.2006	-0.0881	143.8996*	-0.0003
MeOH 80%	-0.1350	0.0209	141.4136*	-0.0003
H_2O	-0.7814*	-0.1939	161.6456*	-0.0003

*: significant at $P < 0.05$

$$\begin{cases} Isoflavone = Solvent \times [a \times Temp + b \times Total\ Time + c] + Temp \times Total\ Time \\ Total\ Time = Extraction\ Time \times Number\ of\ Cycles \end{cases}$$

Table 19: Optimal predictions for each isoflavone (µg) and total content (µg).

Isoflavone	Solvent	Temperature (°C)	Total Time (min)	Prediction (µg)
Di	H_2O or 60% MeOH	97	90 = 30 × 3	4.32
Gly	80% MeOH	27	90 = 30 × 3	5.36
Gi	60% MeOH	97	90 = 30 × 3	4.44
MDi	H_2O	27	10 = 10 × 1	67.21
	80% EtOH	97	90 = 30 × 3	66.89
MGly	60% MeCN	97	10 = 10 × 1	34.41
MGi	80% EtOH	97	90 = 30 × 3	44.33
Total	80% EtOH	97	90 = 30 × 3	150.04

Prediction values was based on 0.1 g of dry edamame.

Total time = time per cycle × number of cycles

Isoflavones are definied in page viii

Table 20: Variables and combinations for model validation.

Sample Size (g)	Temp (°C)	Pressure (psi)	Time per cycle	Cycle	Solvent
0.1	27	500	10	1	60% MeOH
0.1	27	500	10	3	60% MeOH
0.1	27	500	5	1	60% MeOH
0.1	60	500	10	1	60% MeOH
0.1	60	500	5	1	60% MeOH
0.1	27	500	30	3	80% MeOH
0.1	27	500	20	3	80% MeOH
0.1	27	500	40	3	80% MeOH
0.1	60	500	30	3	80% MeOH
0.1	60	500	40	3	80% MeOH
0.1	27	500	30	3	60% EtOH
0.1	27	500	20	3	60% EtOH
0.1	27	500	40	3	60% EtOH
0.1	60	500	30	3	60% EtOH
0.1	60	500	40	3	60% EtOH
0.1	75	500	30	3	80% EtOH
0.1	97	500	30	3	80% EtOH
0.1	97	500	20	3	80% EtOH
0.1	97	500	40	3	80% EtOH
0.1	120	500	30	3	80% EtOH
0.1	27	500	10	1	60% MeCN
0.1	27	500	10	3	60% MeCN
0.1	27	500	5	1	60% MeCN
0.1	60	500	10	1	60% MeCN
0.1	60	500	5	1	60% MeCN
0.1	75	500	10	1	80% MeCN
0.1	97	500	10	1	80% MeCN
0.1	97	500	10	3	80% MeCN
0.1	97	500	5	1	80% MeCN
0.1	120	500	10	1	80% MeCN
0.1	27	500	10	1	H_2O
0.1	27	500	10	3	H_2O
0.1	27	500	5	1	H_2O
0.1	60	500	10	1	H_2O
0.1	60	500	5	1	H_2O

Each combination was used to extract 0.1g of dry edamame and replicated three times.

Table 21: Means for observed value and prediction for 60% EtOH.

Isoflavone	Observed			Prediction		
	Mean	Min	Max	Mean	Min	Max
Di	0	0	0	0.95	0.63	1.13
Gly	4.52	3.07	5.48	3.54	2.80	4.48
Gi	2.14	1.57	3.10	1.45	0.74	2.45
MDi	53.89	48.15	62.04	60.08	56.84	62.82
MGly	28.25	22.38	37.35	22.08	19.82	23.46
MGi	44.22	40.26	47.40	39.15	37.94	40.38
Total	133.02	120.32	145.39	127.26	121.50	132.57

Isoflavones (μg/g) were extracted from 0.1 g of dry edamame.

Table 22: The actual differences (μg), absolute differences (μg) and percentage of absolute differences for 0.1 g of dry edamame extracted with 60% EtOH.

Actual differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	0	n.o.	n.o.	n.o.	n.o.
Gly	15	0.98	0.93	-1.08	2.26
Gi	15	0.69	0.70	-0.43	1.91
MDi	15	-6.19	3.58	-11.15	1.01
MGly	15	6.16	2.98	1.52	14.33
MGi	15	5.08	2.40	0.52	8.21
Total	15	5.77	6.28	-3.46	17.32

↓

Absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	0	n.o.	n.o.	n.o.	n.o.
Gly	15	1.12	0.74	0.24	2.26
Gi	15	0.82	0.53	0.02	1.91
MDi	15	6.40	3.15	0.57	11.15
MGly	15	6.16	2.98	1.52	14.33
MGi	15	5.08	2.40	0.52	8.21
Total	15	6.38	5.60	0.02	17.32

↓

Percentage of absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	0	N/A	N/A	N/A	N/A
Gly	15	23.58	13.45	6.43	44.10
Gi	15	38.71	20.57	0.83	71.93
MDi	15	12.30	6.41	0.93	21.95
MGly	15	21.06	7.34	6.80	38.35
MGi	15	11.27	5.02	1.29	17.63
Total	15	4.60	3.85	0.02	11.92

N: number of observation; N/A: not applicable; n.o.: not observed.

Table 23: Means for observed value and prediction for 80% EtOH.

Isoflavone	Observed			Prediction		
	Mean	Min	Max	Mean	Min	Max
Di	6.06	0	12.48	1.88	0.86	2.90
Gly	6.70	3.65	9.33	2.44	2.07	2.80
Gi	7.28	2.49	11.59	2.54	2.61	2.77
MDi	40.85	32.29	54.00	56.79	50.57	63.06
MGly	20.35	14.34	27.42	27.71	26.31	29.09
MGi	33.35	24.99	43.10	40.47	36.82	44.15
Total	115.986	91.77	133.362	131.84	121.24	142.50

Isoflavones (μg/g) were extracted from 0.1 g of dry edamame.

Table 24: The actual differences (µg), absolute differences (µg) and percentage of absolute differences for 0.1 g of dry edamame extracted with 80% EtOH.

Actual differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	4.18	4.02	-1.50	10.20
Gly	15	4.25	2.12	0.85	7.26
Gi	15	4.74	2.75	0.14	9.27
MDi	15	-15.93	6.53	-30.77	-6.03
MGly	15	-7.35	3.81	-14.74	0.15
MGi	15	-7.13	4.57	-13.34	1.28
Total	15	-15.85	13.27	-50.73	-0.45

↓

Absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	4.78	3.22	1.50	10.20
Gly	15	7.25	2.12	0.85	7.26
Gi	15	4.74	2.75	0.16	9.27
MDi	15	15.93	6.53	6.03	30.77
MGly	15	7.37	3.77	0.15	14.75
MGi	15	7.30	4.27	1.28	13.34
Total	15	02.85	13.27	0.45	50.73

↓

Percentage of absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	12	72.01	8394	54.48	81.78
Gly	15	59.33	16.51	23.21	77.78
Gi	15	58.09	20.98	6.50	80.03
MDi	15	41.96	22.76	11.17	95.30
MGly	15	40.58	26.89	0.56	102.86
MGi	15	24.17	16.29	2.97	47.33
Total	15	14.67	14.04	0.37	55.28

N: number of observations.

Table 25: Means for observed value and prediction for 60% MeCN.

Isoflavone	Observed			Prediction		
	Mean	Min	Max	Mean	Min	Max
Di	0	0	0	0	0	0
Gly	3.39	2.57	4.97	1.28	0.68	1.75
Gi	1.82	1.27	2.25	0.47	0.25	0.70
MDi	60.39	46.59	76.57	62.96	62.52	63.66
MGly	32.51	22.13	42.74	30.25	27.90	31.23
MGi	40.93	33.27	51.24	40.54	39.24	42.03
Total	139.03	107.74	177.49	135.50	133.71	137.85

Isoflavones (μg/g) were extracted from 0.1 g of dry edamame.

Table 26: The actual differences (µg), absolute differences (µg) and percentage of absolute differences for 0.1 g of dry edamame extracted with 60% MeCN.

Actual differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	0	n.o.	n.o.	n.o.	n.o.
Gly	15	2.11	0.77	1.20	4.29
Gi	15	1.36	0.23	0.76	1.64
MDi	15	-2.57	8.60	-17.07	14.05
MGly	15	2.26	5.88	-9.10	12.07
MGi	15	0.39	4.98	-7.69	11.94
Total	15	3.54	19.30	-30.11	43.78

↓

Absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	0	n.o.	n.o.	n.o.	n.o.
Gly	15	2.11	0.77	1.20	4.29
Gi	15	1.35	0.23	0.76	1.64
MDi	15	7.14	5.14	0.32	17.07
MGly	15	4.99	3.64	0.55	12.07
MGi	15	3.89	2.96	1.22	11.94
Total	15	15.19	11.78	0.73	43.78

↓

Percentage of absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	0	N/A	N/A	N/A	N/A
Gly	15	61.64	15.47	41.13	86.30
Gi	15	74.45	7.59	59.39	84.96
MDi	15	12.63	10.64	0.51	36.63
MGly	15	15.21	10.83	1.75	41.10
MGi	15	9.54	7.01	3.06	23.32
Total	15	10.87	8.24	0.53	27.95

N: number of observations; N/A: not applicable; n.o.: not observed.

Table 27: Means for observed value and prediction for 80% MeCN.

Isoflavone	Observed			Prediction		
	Mean	Min	Max	Mean	Min	Max
Di	0	0	0	0	0	0
Gly	0	0	0	0	0	0
Gi	0	0	0	0	0	0
MDi	46.04	23.25	58.27	20.90	19.88	21.77
MGly	22.61	13.66	41.06	5.65	4.89	6.40
MGi	38.74	29.87	47.12	45.87	44.02	47.73
Total	111.37	77.72	138.25	72.42	70.98	74.39

Isoflavones (μg/g) were extracted from 0.1 g of dry edamame.

Table 28: The actual differences (µg), absolute differences (µg) and percentage of absolute differences for 0.1 g of dry edamame extracted with 80% MeCN.

Actual differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	0	0	0	0
Gly	15	0	0	0	0
Gi	15	0	0	0	0
MDi	15	25.14	10.37	1.48	36.87
MGly	15	16.96	6.96	8.77	35.39
MGi	15	-7.13	5.64	-17.86	0.93
Total	15	38.95	17.53	3.33	65.21

↓

Absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	0	0	0	0
Gly	15	0	0	0	0
Gi	15	0	0	0	0
MDi	15	25.14	10.37	1.48	36.87
MGly	15	16.96	6.96	8.77	35.39
MGi	15	7.25	5.47	0.93	17.86
Total	15	38.95	17.53	3.33	65.21

↓

Percentage of absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	N/A	N/A	N/A	N/A
Gly	15	N/A	N/A	N/A	N/A
Gi	15	N/A	N/A	N/A	N/A
MDi	15	51.37	16.42	6.37	63.27
MGly	15	73.09	7.43	58.59	86.20
MGi	15	20.92	18.70	1.98	59.81
Total	15	33.25	12.43	4.29	47.16

N: number of observations; N/A: not applicable.

Table 29: Means for observed value and prediction for 60% MeOH.

Isoflavone	Observed			Prediction		
	Mean	Min	Max	Mean	Min	Max
Di	0	0	0	1.10	0.81	1.33
Gly	4.36	3.42	6.50	3.58	3.38	3.76
Gi	2.00	1.53	2.45	1.70	1.42	2.09
MDi	50.05	44.11	59.54	60.99	59.75	62.35
MGly	28.77	25.00	32.01	23.05	22.10	23.72
MGi	40.48	36.37	45.08	39.35	39.00	40.11
Total	125.67	112.41	138.49	129.77	126.93	132.08

Isoflavones (μg/g) were extracted from 0.1 g of dry edamame.

Table 30: The actual differences (µg), absolute differences (µg) and percentage of absolute differences for 0.1 g of dry edamame extracted with 60% MeOH.

Actual differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	-1.10	0.19	-1.33	-0.81
Gly	15	0.78	0.79	-0.30	2.91
Gi	15	0.29	0.28	-0.13	0.88
MDi	15	-10.93	3.66	-15.64	-1.59
MGly	15	5.72	2.27	1.49	8.52
MGi	15	1.13	2.77	-2.75	4.96
Total	15	-4.10	6.50	-14.52	7.71

↓

Absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	1.10	0.19	0.81	11.33
Gly	15	0.85	0.82	0.17	2.91
Gi	15	0.31	0.26	0.01	0.88
MDi	15	10.93	3.66	1.59	15.64
MGly	15	5.72	2.27	1.49	8.52
MGi	15	2.53	1.48	0.26	1.96
Total	15	5.92	4.77	0.25	14.52

↓

Percentage of absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	N/A	N/A	N/A	N/A
Gly	15	17.53	11.43	4.77	44.77
Gi	15	15.29	11.93	0.24	38.22
MDi	15	22.39	8.39	2.66	35.45
MGly	15	19.47	6.78	5.97	27.46
MGi	15	6.12	3.31	0.66	11.01
Total	15	4.88	4.13	0.19	12.92

N: number of observations; N/A: not applicable.

Table 31: Means for observed value and prediction for 80% MeOH.

Isoflavone	Observed			Prediction		
	Mean	Min	Max	Mean	Min	Max
Di	0	0	0	2.22	1.68	2.77
Gly	5.07	3.69	6.35	5.14	4.55	5.69
Gi	2.76	1.98	4.59	2.68	1.83	4.35
MDi	52.32	46.34	57.11	63.21	61.88	64.18
MGly	25.65	21.15	30.80	24.54	21.70	26.67
MGi	44.98	42.86	47.40	38.87	36.41	40.38
Total	130.79	119.14	139.93	136.67	134.24	138.48

Isoflavones (μg/g) were extracted from 0.1 g of dry edamame.

Table 32: The actual differences (µg), absolute differences (µg) and percentage of absolute differences for 0.1 g of dry edamame extracted with 80% MeOH.

Actual differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	-2.22	0.46	-2.77	-1.68
Gly	15	0.07	0.91	-1.33	1.81
Gi	15	0.08	0.63	-1.44	0.66
MDi	15	-10.89	2.84	-14.49	-6.27
MGly	15	1.10	3.75	-5.52	9.10
MGi	15	6.12	1.44	2.78	7.96
Total	15	-5.88	6.98	-19.32	5.69

↓

Absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	2.22	0.46	1.68	2.77
Gly	15	0.77	0.91	0.10	1.81
Gi	15	0.50	0.37	0.01	1.44
MDi	15	10.89	2.84	6.27	14.49
MGly	15	2.76	2.68	0.07	9.10
MGi	15	6.12	1.44	2.78	7.96
Total	15	7.39	5.22	1.46	19.23

↓

Percentage of absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	N/A	N/A	N/A	N/A
Gly	15	15.58	8.80	1.86	31.50
Gi	15	19.02	13.22	0.06	49.56
MDi	15	21.13	6.61	10.97	35.58
MGly	15	10.58	9.67	0.28	29.55
MGi	15	13.57	3.08	6.44	17.55
Total	15	5.82	4.38	1.07	16.14

N: number of observations; N/A: not applicable.

Table 33: Means for observed value and prediction for H_2O.

Isoflavone	Observed			Prediction		
	Mean	Min	Max	Mean	Min	Max
Di	0	0	0	0.41	0.02	0.92
Gly	0	0	0	2.71	1.35	4.28
Gi	0	0	0	0.55	0.47	0.63
MDi	20.53	14.92	32.65	66.97	53.15	81.23
MGly	12.23	7.27	22.82	25.83	19.97	31.78
MGi	16.42	10.42	20.72	43.09	35.38	51.20
Total	49.18	32.74	74.04	139.57	110.66	170.05

Isoflavones (μg/g) were extracted from 0.1 g of dry edamame.

Table 34: The actual differences (µg), absolute differences (µg) and percentage of absolute differences for 0.1 g of dry edamame extracted with H_2O.

Actual differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	-0.41	0.37	-0.92	-0.02
Gly	15	-2.71	1.27	-4.28	-1.35
Gi	15	-0.55	0.07	-0.63	-0.47
MDi	15	-44.45	13.79	-66.18	-31.33
MGly	15	-12.55	6.60	-24.50	-3.84
MGi	15	-25.18	8.65	-40.78	-15.54
Total	15	-85.85	29.40	-137.30	-60.12

↓

Absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	15	0.41	0.37	0.02	0.92
Gly	15	2.71	1.27	1.35	4.28
Gi	15	0.55	0.07	0.47	0.63
MDi	15	46.44	14.30	31.33	66.18
MGly	15	13.60	6.91	3.84	24.50
MGi	15	26.67	8.89	15.54	40.78
Total	15	90.39	30.72	60.12	137.30

↓

Percentage of absolute differences:

Isoflavone	N	Mean	Standard Deviation	Min	Max
Di	0	N/A	N/A	N/A	N/A
Gly	0	N/A	N/A	N/A	N/A
Gi	0	N/A	N/A	N/A	N/A
MDi	15	249.73	119.59	109.42	439.67
MGly	15	141.75	106.84	16.83	336.85
MGi	15	183.99	107.26	77.56	391.57
Total	15	206.32	115.88	91.52	419.36

N: number of observations; N/A: not applicable

Appendix C: Statistic Program

1st set of extractions

```
proc import datafile='112 Define Range Extractions calculated concentration and
outlier removed.csv' out=one replace;
run;

proc print data=one; run;
proc means data=one; run;

proc glm data=one;
 class solvent  Temp Pressure Time Cycle;
 model Daidzin Glycitin
Genistin MalDaidzin MalGlycitin  MalGenistin
  Apigenin Total=  Temp| Pressure |Time| Cycle |Solvent@2;
lsmeans Temp| Pressure |Time| Cycle |Solvent@2;
output out=rrr r=rDaidzin rGlycitin
rGenistin rMalDaidzin rMalGlycitin  rMalGenistin
  rApigenin rTotal;
run;
proc univariate plot normal data=rrr;
 var rDaidzin rGlycitin
rGenistin rMalDaidzin rMalGlycitin  rMalGenistin
  rApigenin rTotal;
run;

%include 'H:\danda.sas';
%macro runmix(var);
title1 "LSMeans for &var";
proc mixed data=one;
 class solvent  Temp Pressure Time Cycle;
 model  &var=  Temp| Pressure |Time| Cycle |Solvent@2;
lsmeans Temp| Pressure |Time| Cycle |Solvent@2 /pdiff;
ods listing exclude lsmeans diffs;
ods output lsmeans=mmm diffs=ppp;
run;
%pdmix800(ppp,mmm);
%mend;
%runmix(Daidzin);
```

```
%runmix(Glycitin);
%runmix(Genistin);
%runmix(MalDaidzin);
%runmix(MalGlycitin);
%runmix(MalGenistin);
%runmix(Apigenin);
%runmix(Total);
```

2nd set of extractions

```
proc import datafile='Total Time Preliminary Extractions calculated
concentration.csv' out=one replace;
run;
data one; set one;
 if index(solvent,'20')>0 then delete;
run;

proc print data=one; run;
proc means data=one; run;


data pred;
 do temp=27 to 97 by 2 ;
 do time=10 to 90 by 2;
 *do cycle=1 to 3 by 1;
 do solvent='EtOH 60%', 'EtOH 80%', 'H2O',  'MeCN 60%', 'MeCN 80%',  'MeOH 60%',
'MeOH 80%';
   Daidzin=.; Glycitin =.;
Genistin =.;MalDaidzin =.;MalGlycitin=.;  MalGenistin=.;  Total=.;
Apigenin=.;
output;
 end;end;end;*end;
run;

data two; set one pred;
run;
**** for equations only;
proc glm data=two;
 class solvent  ;
 model Daidzin Glycitin
Genistin MalDaidzin MalGlycitin  MalGenistin  Total  Apigenin=
  Temp*Time
  Solvent
  Solvent*Temp
  Time*Solvent
/solution noint
```

```
; run;

proc glm data=two;
 class solvent  ;
 model Daidzin Glycitin
Genistin MalDaidzin MalGlycitin  MalGenistin  Total  Apigenin=
  Temp
  Time
  Temp*Time
  Solvent
  Solvent*Temp
  Time*Solvent
;
lsmeans Solvent;
output out=rrr r=rDaidzin rGlycitin
rGenistin rMalDaidzin rMalGlycitin  rMalGenistin
   rTotal rApigenin
  lclm=lclDaidzin lclGlycitin
lclGenistin lclMalDaidzin lclMalGlycitin  lclMalGenistin
   lclTotal lclApigenin
  uclm=uclDaidzin uclGlycitin
uclGenistin uclMalDaidzin uclMalGlycitin  uclMalGenistin
   uclTotal uclApigenin
predicted=pDaidzin pGlycitin
pGenistin pMalDaidzin pMalGlycitin  pMalGenistin
   pTotal pApigenin;
run;
proc univariate plot normal data=rrr;
 var rDaidzin rGlycitin
rGenistin rMalDaidzin rMalGlycitin  rMalGenistin
  rApigenin rTotal;
run;

proc sort data=rrr; by solvent;
data ann1 ann2 ann3 ann4 ann5 ann6 ann7 ann8; set rrr;
 if rTotal ne .;
 keep solvent function style text color z y x xsys ysys zsys;
```

```
 xsys='2'; ysys='2'; zsys='2';text='V';
y=time; x=temp;

function='symbol';
z= Daidzin;
 if rDaidzin le 0 then do; style='markere';  color=' Red'; end;
 else do; color='Blue'; style='marker';end;
output ann1;
 z= Glycitin;
if rGlycitin le 0 then do; style='markere';  color=' Red'; end;
 else do; color='Blue'; style='marker';end;
output ann2;
 z= Genistin;
if rGenistin le 0 then do; style='markere';  color=' Red'; end;
 else do; color='Blue'; style='marker';end;
output ann3;
 z= MalDaidzin;
if rMalDaidzin le 0 then do; style='markere';  color=' Red'; end;
 else do; color='Blue'; style='marker';end;
output ann4;
 z= MalGlycitin;
if rMalGlycitin le 0 then do; style='markere';  color=' Red'; end;
 else do; color='Blue'; style='marker';end;
output ann5;
 z= MalGenistin;
if rMalGenistin le 0 then do; style='markere';  color=' Red'; end;
 else do; color='Blue'; style='marker';end;
output ann6;
 z= Total;
if rTotal le 0 then do; style='markere';  color=' Red'; end;
 else do; color='Blue'; style='marker';end;
output ann7;
 z= Apigenin;
if rApigenin le 0 then do; style='markere';  color=' Red'; end;
 else do; color='Blue'; style='marker';end;
output ann8;
run;
```

```
proc g3d data=rrr; by solvent;
title2 'Predicted surface';
 plot time*temp=pDaidzin  / grid annotate=ann1 zmin=0;
 plot time*temp=pGlycitin  / grid annotate=ann2 zmin=0;
 plot time*temp=pGenistin  / grid annotate=ann3 zmin=0;
 plot time*temp=pMalDaidzin  / grid annotate=ann4 zmin=0;
 plot time*temp=pMalGlycitin  / grid annotate=ann5 zmin=0;
 plot time*temp=pMalGenistin  / grid annotate=ann6 zmin=0;
 plot time*temp=pTotal  / grid annotate=ann7 zmin=50;
 plot time*temp=pApigenin  / grid annotate=ann8 zmin=0;
run;quit;

proc sort data=rrr; by descending pDaidzin ;
proc print data=rrr (obs=20);
title2 'Optimized for Daidzin';
 var solvent time temp  pDaidzin pGlycitin
pGenistin pMalDaidzin pMalGlycitin  pMalGenistin
   lclTotal pTotal uclTotal pApigenin; run;
proc sort data=rrr; by descending pGlycitin ;
proc print data=rrr (obs=20);
title2 'Optimized for Glycitin';
 var solvent time temp  pDaidzin pGlycitin
pGenistin pMalDaidzin pMalGlycitin  pMalGenistin
   lclTotal pTotal uclTotal pApigenin; run;
proc sort data=rrr; by descending pGenistin ;
proc print data=rrr (obs=20);
title2 'Optimized for Genistin';
 var solvent time temp  pDaidzin pGlycitin
pGenistin pMalDaidzin pMalGlycitin  pMalGenistin
   lclTotal pTotal uclTotal pApigenin; run;
proc sort data=rrr; by descending pMalDaidzin ;
proc print data=rrr (obs=20);
title2 'Optimized for MalDaidzin';
 var solvent time temp  pDaidzin pGlycitin
pGenistin pMalDaidzin pMalGlycitin  pMalGenistin
```

```
   lclTotal pTotal uclTotal pApigenin; run;
proc sort data=rrr; by descending pMalGlycitin ;
proc print data=rrr (obs=20);
title2 'Optimized for MalGlycitin';
 var solvent time temp  pDaidzin pGlycitin
pGenistin pMalDaidzin pMalGlycitin  pMalGenistin
   lclTotal pTotal uclTotal pApigenin; run;

proc print data=rrr (obs=20);
title2 'Optimized for Total';
 var solvent time temp  pDaidzin pGlycitin
pGenistin pMalDaidzin pMalGlycitin  pMalGenistin
   lclTotal pTotal uclTotal pApigenin; run;
proc sort data=rrr; by descending pMalGenistin ;
proc print data=rrr (obs=20);
title2 'Optimized for MalGenistin';
 var solvent time temp  pDaidzin pGlycitin
pGenistin pMalDaidzin pMalGlycitin  pMalGenistin
   lclTotal pTotal uclTotal pApigenin; run;
proc sort data=rrr; by descending pApigenin ;
proc print data=rrr (obs=20);
title2 'Optimized for Apigenin';
 var solvent time temp  pDaidzin pGlycitin
pGenistin pMalDaidzin pMalGlycitin  pMalGenistin
   lclTotal pTotal uclTotal pApigenin; run;


   proc sort data=rrr; by solvent descending pTotal ;
data use; set rrr; by solvent;
 if first.solvent;
run;
   proc print data=use (obs=20);
title2 'Best Predicted Total for each solvent';
 var solvent time temp   lclTotal pTotal uclTotal ; run;
```

Model Validation

```
data tre; set two; where  fillindata ne 1;  ***no extra data used for validation;
  array xxx Daidzin Glycitin
Genistin MalDaidzin  MalGlycitin  AcetylDaidzin  MalGenistin
Daidzein Glycitein AcetylGenistin Genistein
Total Apigenin;
  array ttt truth1-truth13;
  do over xxx; ttt=xxx; end;
  if run=3 then do over xxx; xxx=.;
  end;

run;

proc sort; by solvent;
proc glm data=tre; by solvent ;
 model Daidzin Glycitin
Genistin MalDaidzin  MalGlycitin  AcetylDaidzin  MalGenistin
Daidzein Glycitein AcetylGenistin Genistein
Total Apigenin=
  Temp
  Time
  Temp*Time
;
output out=rrr r=rDaidzin rGlycitin
             rGenistin rMalDaidzin  rMalGlycitin  rAcetylDaidzin
rMalGenistin
             rDaidzein rGlycitein rAcetylGenistin rGenistein
             rTotal rApigenin

  lclm=lclDaidzin lclGlycitin
lclGenistin lclMalDaidzin  lclMalGlycitin  lclAcetylDaidzin  lclMalGenistin
lclDaidzein lclGlycitein lclAcetylGenistin lclGenistein
lclTotal lclApigenin

  uclm=uclDaidzin uclGlycitin
uclGenistin uclMalDaidzin  uclMalGlycitin  uclAcetylDaidzin  uclMalGenistin
uclDaidzein uclGlycitein uclAcetylGenistin uclGenistein
```

```
uclTotal uclApigenin

predicted=pDaidzin pGlycitin
pGenistin pMalDaidzin  pMalGlycitin  pAcetylDaidzin  pMalGenistin
pDaidzein pGlycitein pAcetylGenistin pGenistein
pTotal pApigenin;
run;


proc gplot data=rrr;  by solvent;
 plot pDaidzin*truth1 pGlycitin*truth2
pGenistin*truth3 pMalDaidzin*truth4  pMalGlycitin*truth5
pAcetylDaidzin*truth6  pMalGenistin*truth7
pDaidzein*truth8 pGlycitein*truth9 pAcetylGenistin*truth10
pGenistein*truth11
pTotal*truth12 pApigenin*truth13;
run;quit;
proc corr data=rrr; by solvent;where run=3; var pDaidzin truth1;
proc corr data=rrr; by solvent;where run=3; var pGlycitin truth2 ;
proc corr data=rrr; by solvent;where run=3; var pGenistin truth3 ;
proc corr data=rrr; by solvent;where run=3; var MalDaidzin truth4  ;
proc corr data=rrr; by solvent;where run=3; var pMalGlycitin truth5  ;
proc corr data=rrr; by solvent;where run=3; var pAcetylDaidzin truth6;
proc corr data=rrr; by solvent;where run=3; var pMalGenistin truth7;
proc corr data=rrr; by solvent;where run=3; var pDaidzein truth8 ;
proc corr data=rrr; by solvent;where run=3; var pGlycitein truth9 ;
proc corr data=rrr; by solvent;where run=3; var pAcetylGenistin truth10 ;
proc corr data=rrr; by solvent;where run=3; var pGenistein truth11;
proc corr data=rrr; by solvent;where run=3; var pTotal truth12 ;
proc corr data=rrr; by solvent;where run=3; var pApigenin truth13;
run;

data diffs; set rrr;
actdiff1=-( pDaidzin-truth1) ;
actdiff2=-(pGlycitin-truth2) ;
actdiff3=-(pGenistin-truth3)    ;
actdiff4=-(pMalDaidzin-truth4) ;
```

```
actdiff5=-(pMalGlycitin-truth5) ;
actdiff6=-(pAcetylDaidzin-truth6);
actdiff7=-(pMalGenistin-truth7)  ;
actdiff8=-(pDaidzein-truth8)    ;
actdiff9=-(pGlycitein-truth9)  ;
actdiff10=-(pAcetylGenistin-truth10);
actdiff11=-(pGenistein-truth11) ;
actdiff12=-(pTotal-truth12)  ;
actdiff13=-(pApigenin-truth13);

diff1=abs( pDaidzin-truth1) ; pdiff1=diff1*100/truth1;
diff2=abs(pGlycitin-truth2) ; pdiff2=diff2*100/truth2;
diff3=abs(pGenistin-truth3) ; pdiff3=diff3*100/truth3;
diff4=abs(pMalDaidzin-truth4) ; pdiff4=diff4*100/truth4;
diff5=abs(pMalGlycitin-truth5) ;  pdiff5=diff5*100/truth5;
diff6=abs(pAcetylDaidzin-truth6);   pdiff6=diff6*100/truth6;
diff7=abs(pMalGenistin-truth7)  ; pdiff7=diff7*100/truth7;
diff8=abs(pDaidzein-truth8) ; pdiff8=diff8*100/truth8;
diff9=abs(pGlycitein-truth9)  ; pdiff9=diff9*100/truth9;
diff10=abs(pAcetylGenistin-truth10);  pdiff10=diff10*100/truth10;
diff11=abs(pGenistein-truth11) ;    pdiff11=diff11*100/truth11;
diff12=abs(pTotal-truth12) ; pdiff12=diff12*100/truth12;
diff13=abs(pApigenin-truth13); pdiff13=diff13*100/truth13;
run;
proc means; by solvent; where run=3;
title 'Actual differences';
 var actdiff1 actdiff2 actdiff3 actdiff4 actdiff5 actdiff6 actdiff7 actdiff8
actdiff9
actdiff10 actdiff11 actdiff12 actdiff13;
run;
proc means; by solvent; where run=3;
title 'Absolute differences';
 var diff1 diff2 diff3 diff4 diff5 diff6 diff7 diff8 diff9
diff10 diff11 diff12 diff13;
run;
proc means; by solvent; where run=3;
title '%Absolute differences/truth';
```

```
var pdiff1 pdiff2 pdiff3 pdiff4 pdiff5 pdiff6 pdiff7 pdiff8 pdiff9
pdiff10 pdiff11 pdiff12 pdiff13;
run;
proc means; by solvent; where run=3;
title 'Average prediction';
var pDaidzin pGlycitin
pGenistin pMalDaidzin  pMalGlycitin  pAcetylDaidzin  pMalGenistin
pDaidzein pGlycitein pAcetylGenistin pGenistein
pTotal pApigenin;
run;
proc means; by solvent; where run=3;
title 'Average truth';
var truth1-truth13;
run;


proc print data=diffs; where solvent='MeCN 80%';
var run truth2 pGlycitin Glycitin diff2 pdiff2 actdiff2;
run;
```

Vita

Yu-Ting Hung, also known as Antony, was born in Tainan, Taiwan, on 29th of August 1984. He has lived in Tainan and moved to Taipei when he was eight years old. After graduating from Chien-Kuo Senior High School in 2002, he came to Carson-Newman College, Jefferson, Tennessee, and graduated with a B.A in biology and a minor in mathematics in 2005. After graduation, Yu-Ting started a Master of Science program in Plant Sciences at the University of Tennessee-Knoxville and concentrated in crop physiology. He graduated in December 2008. Yu-Ting hopes to continue his education at the doctoral level in plant breeding, and he intends to teach and conduct scientific research professionally.