Assessing Genetic and Environmental Influence on Traits Associated with Natto Quality

David Edward Cook III

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
In
Crop and Soil Environmental Science

Dr. Katy Martin Rainey, Committee Chair
Dr. Glenn Buss, Committee Member
Dr. Carl Griffey, Committee Member
Dr. David Holshouser, Committee Member

May 20th, 2008
Blacksburg, Virginia

Keywords: Natto, Genotype × Environment, Multi environment trial, Trait Stability, Soybean

Assessing Genetic and Environmental Influence on Traits Associated with Natto Quality

David Edward Cook III

(ABSTRACT)

Food grade soybean production is a high value alternative to conventional soybean use. The production of natto, a fermented soyfood, requires soybean cultivars that consistently express specific quality traits over a range of growing environments. Therefore, it is necessary to evaluate genetic and environmental influence for natto quality traits to ensure consistent performance. A multi location experiment was conducted in 2006 and 2007 to address the influence of soybean cropping system (double crop vs. full season) and environmental factors on traits associated with natto quality. Two statistical models were used to analyze the effects of planting system and environment on agronomic traits such as yield, maturity, and seed size and natto quality traits such as water absorption, water loss after steaming, seed coat deficiency, and rate of water absorption. Genotype variation was significant for all traits, but genetic differences for water loss after steaming were minimal. Planting system significantly influenced all natto quality traits. Seed coat deficiency and rate of water absorption displayed the most differential response and double crop plantings produced superior characteristics. Genotype × environment interactions were significant for all traits, but they did not confound selecting superior natto cultivars. Significant environment and year effects indicate environmental sensitivity, but genotype rankings rarely changed. The results indicate that genotype was the most important factor controlling the natto quality traits tested. These results suggest breeding for superior natto cultivars is possible but environmental influence must be accounted for and multi environment testing is necessary for genotype natto quality evaluation.

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS AND MODELS

Introduction

Cultivated soybean [*Glycine max*] originated in China and was initially grown over a wide range of geographic locations, exposing the crop to an array of environmental climates, contributing to the genetic diversity of the species (Dong et al., 2004). The initial use of soybean in China was direct human consumption, but soybean's main use in today's world market is for edible oil and animal feed. Today soybean is cultivated worldwide attesting to the crops success and adaptability. Soybean's wide growing region and many end uses challenge plant breeders to improve agronomic traits while maintaining end use suitability.

In the United States, soybean is grown on over 70 million acres and is worth an estimated $26 billion to the nation's economy via internal use and exports (USDA NASS, 2008). Soybean is grown in thirty-one states across the U.S. with the top five producing states accounting for greater than 50% of U.S. production (Table 1). Some of the crop produced in the U.S. is used domestically while the majority is exported, accounting for roughly 40% of the world soybean exports in 2006 (The American Soybean Association, 2007).

Table 1. United States soybean production listed in million(s) of bushels and metric tons.

Rank	State	Million Bushels	Million Metric Ton	Rank	State	Million Bushels	Million Metric Ton
1	Iowa	510	13.88	16	Tennessee	44	1.20
2	Illinois	482	13.13	17	Mississippi	43	1.17
3	Minnesota	319	8.68	18	Louisiana	29	0.80
4	Indiana	284	7.73	19	Pennsylvania	17	0.46
5	Nebraska	251	6.82	20	Maryland	16	0.43
6	Ohio	217	5.91	21	Virginia	16	0.43
7	Missouri	194	5.29	22	S. Carolina	11	0.31
8	S. Dakota	131	3.56	23	New York	9	0.25
9	N. Dakota	120	3.27	24	Delaware	5	0.15
10	Arkansas	107	2.92	25	Georgia	4	0.10
11	Kansas	99	2.68	26	Oklahoma	4	0.10
12	Michigan	90	2.44	27	Texas	4	0.10
13	Wisconsin	72	1.96	28	Alabama	3	0.08
14	Kentucky	60	1.64	29	NewJersey	3	0.08
15	N. Carolina	44	1.18	30	W. Virginia	0.7	0.02
				31	Florida	0.1	0.00

Source: The American Soybean Association, 2007

The success of the soybean industry is attributed to the crop's agronomic performance and composition (Liu, 2000). Soybean can be grown in a wide range of environments, has moderate drought tolerance, and does not require nitrogen fertilization because of symbiosis with the nitrogen fixing bacteria; *Rhizobium*. Roughly 60% of the dry weight of soybean seed is composed of protein and oil, and as a result soybean seed has the highest protein and second highest oil content among cereal and legume species (Liu, 1997). These facts make soybean an important crop for oil manufacturing, animal feed, and food production (Hymowitz et al., 1972; Wilcox and Shibles, 2001; Roa et al., 2002; Lusas, 2004). Soybean represented 57% of the world oilseed production in 2006, and comprises approximately 80% of the U.S. edible oil and fat market annually (The American Soybean Association, 2007; Liu, 2000). The crushing process to extract oil from soybeans results in high protein soybean meal supplying 68% of the world meal market in 2006 (The American Soybean Association, 2007).

The U.S. soybean industry is facing competition from foreign producers including Brazil and Argentina to remain the number one world soybean supplier. Collectively, South American countries have overtaken the U.S. as the leading soybean producer and exporter (Ash and Dohlman, 2007). The United Soybean Board passed the Better Bean Initiative (BBI) with the purpose of improving the composition quality of soybeans grown in the United States to compete with foreign production (Durham, 2003). Another way U.S. growers can remain central in a growing global market is to sell soybeans in markets that focus on quality and pay a price premium (Durham, 2003). Soybeans sold for human consumption is an example of a high value end use of soybean. Producers receive a price premium for food grade soybeans because they express value added traits that make them more valuable than those sold for oil and meal. Continued efforts by the United Soybean Board, the American Soybean Association, and the USDA will be important in developing new value added markets for American producers.

Food Grade Soybeans

The preparation of soybeans for human consumption dates back over 5000 years in Asian cultures and has evolved over time to include a myriad of different food products developed in different cultural regions. China, thought to be the center of domestication of soybean, has a rich history of soybean cultivation and landrace development specifically for human consumption. Many landraces were developed for specific end use as conveyed by their names: *cai dou* (vegetable bean), *dou fu dou* (tofu bean), *you dou* (oil bean) (Cui et al., 2005). The development of cultivars for specific end uses continues today, reflecting the diversity of soyfood products and the different seed quality characteristics needed for food preparation. Soyfood products consumed in

China, Japan, and other East Asian countries are considered traditional soyfoods because they were developed and altered for generations thousands of years ago (Golbitz, 1995). Traditional soyfoods are often made from whole soybeans and include products such as tofu, natto, soy sauce, miso, bean sprouts, and vegetable beans. The preparation of traditional soyfoods can vary but all include a sprouting, boiling, or fermenting process needed to make the bean digestible by denaturing certain soy proteins and sugars (McCue and Shetty, 2004). Traditional soyfoods are generally classified as non-fermented or fermented. Non-fermented soyfoods, such as tofu, comprise the majority of traditional soyfoods and are often consumed as main dishes, while fermented soyfoods, such as soy sauce, are usually added as seasoning (Golbitz, 1995; Liu, 2005).

Traditional soyfood acceptance in Western culture has been met with reservation, and until recently, soyfoods have occupied almost no part of Western diets. Marketing and production efforts have integrated edible soy into mainstream Western foods such as salad and cooking oil, soy flour, soy protein, soy burger and sausage, soy ice cream, soy dietary supplements, and a variety of products "enhanced" with soy (Golbitz, 1995; Liu, 2005). These products differ from traditional soyfoods in that they are highly processed, do not contain whole soybeans, and soybean is not always the main ingredient. Although soybeans used for human consumption currently represent a small part of the international soybean market, there is confidence that the market will continue to grow (Liu, 2000).

Food Grade Soybean Breeding

It has long been recognized that cultivar quality used for traditional soyfood production impacts the quality of the final product (Taira, 1990; Cui et al., 2005). Specific soy cultivars must be developed based on their intended end use because of production needs and consumer preference, resulting in large breeding efforts in East Asia and recently in North America. China began modern breeding efforts in 1913 and released 651 public soybean cultivars by 1995, of which 193 were specifically developed for food grade end uses (Cui et al., 2005). In Japan, the major breeding efforts are aimed at food grade production because a majority of their oil and meal is imported, and by 1995, 97 soyfood cultivars had been released (Cui et al., 2005). American public breeding programs have developed most of the soyfood cultivars grown in the U.S. and have been responsible for releasing 123 food grade cultivars by 2000 (Cui et al., 2005). The relatively small number of food grade cultivars developed reflects the dominant use of soybeans for conventional oil and meal.

Breeding for food grade soybeans is more difficult then breeding for conventional use in many ways. The use of less agronomically adapted soybean genotypes as parents presents an obstacle to food grade cultivar development. The development of North American soyfood cultivars from 1956 to 2000 relied on integrating 29 unique ancestors not found in the genetic base of commodity soybeans (Cui et al., 2005). These accessions contained many undesirable agronomic traits, and additional time and resources were required for backcrossing and selection to break the linkage between the desirable and undesirable traits. In addition, food grade breeding is not the primary objective in many programs so fewer resources have been applied to food grade cultivar development. As a result, food grade cultivars contain large amounts of exotic pedigree and tend to be lower

yielding than conventional cultivars. It should be noted that of the original 29 accessions used to develop the North American soyfood cultivar base, no single cultivar has dominated breeding efforts (Cui et al., 2005). This implies that many accessions have desirable traits that can be exploited through breeding.

The need to select for quality traits in addition to yield and agronomic traits is another reason food grade breeding is more difficult than conventional breeding. For example, food grade breeding requires selection for yield, disease resistance, and maturity in addition to altered seed composition, aesthetic quality, and production characteristics that are often quantitatively inherited. These additional selection criteria require more time and resources and some traits are negatively correlated such that indirect selection is a problem. These potential relationships require additional testing by breeding programs to ensure selecting for one quality trait will not undermine progress for another. It is also important that food grade traits associated with quality are stably expressed and heritable. In order to capture value, a food grade cultivar must meet the quality standards imposed by the market. Without the desirable characteristics, a food grade cultivar has no more value than a cultivar grown for conventional oil or meal. Therefore, a crop produced with an intended food grade end use must possess the desired characteristics. Food grade quality traits should be highly heritable and display stability over environments. Stability refers to consistent trait expression regardless of environmental conditions and is a function of heritability. Deviations from expected trait expression are caused by environmental influence, which negatively effects heritability, and causes instability. Environmental influence, heritability, and stability for food grade cultivars are addressed later.

Food grade soybean cultivar development is also difficult because genetically modified (GM) cultivars are not accepted by food producers. GM soybean cultivars are standard in the United States' soybean oil and meal industry and growers are accustomed to GM crop production practices. This creates a problem convincing growers to switch to non-GM food grade cultivars because they require different production practices that can be less convenient. These non-GM cultivars are unaccepted by traditional soy producers in the United States and create another complexity when developing food grade cultivars.

There are a variety of soyfood products that comprise the soybean food grade market. Each soyfood requires a specific food grade cultivar with specific quality traits. One such soyfood, natto, is the focus of the current research and will be discussed further.

Natto

Natto is a traditional Japanese food prepared by soaking, steaming, and fermenting whole soybean seeds (Taira, 1990). Unlike other fermented soyfoods, natto is often eaten as a main dish for breakfast or dinner, served over rice and seasoned with mustard or soy sauce. The initial step in natto production is soaking whole soybean seeds in water until maximum hydration is achieved. The time required to reach maximum hydration is loosely defined as overnight (Liu, 2005) but has been experimentally shown to occur after 12 hours (Cober et al., 2006) and is temperature dependant (Pan and Tangratanavalee, 2003). Soaked beans are then steamed under pressure. Steaming affects the texture and taste of the final product by softening the beans, releasing soluble sugars, and denaturing indigestible proteins. The beans are steamed for 30 to 40 minutes between 15 and 20 p.s.i. with a temperature range of 121 to 131°C (Geater et al., 2000; Wei and Chang, 2004). The cooked beans are then inoculated with pure culture bacteria

Bacillus natto. The fermentation is carried out at 40°C for 12 to 20 hours (Liu, 2005). Once fermented, the product is available for fresh consumption for a week or can be frozen for long-term storage.

Characteristics of Soybeans for Natto Production

Not all soybean cultivars are appropriate for natto production because consumer preferences and production require natto cultivars to have certain characteristics to ensure quality. Breeders developing cultivars for natto production need to identify and understand specific desirable attributes of natto cultivars. Defining desirable soybean characteristics that control natto traits is difficult because natto producers are often secretive about what makes the best natto. Additionally, different producers evaluate different characteristics and may have different quality standards (Cui et al., 2005). Despite these inconsistencies, some characteristics are ubiquitous and specific traits have been established. These desirable seed traits have been identified in literature and through communication with natto producers. The following attributes encompass the major areas of focus for natto cultivar development.

A. Soybean Seed Appearance, Size, and Shape

Seed size and appearance influence the acceptability of a soybean for natto production (Taira et al., 1987; Cober et al., 1997a; Geater et al., 2000). A seed size of less than 9g per 100 seed that can pass through a 5 to 5.5 mm screen is preferred by manufacturers (Cui et al., 2005). The shape of small seeded cultivars is also important and a near perfect sphere is desirable. Seeds with a spherical shape have a higher cotyledon to seed coat ratio than seeds with a flat shape. The seed coat is composed of

structural carbohydrates and tends to be tougher than the inner cotyledon, therefore, spherical seeds produce natto with a soft texture (Cui et al., 2005). It is also desirable for the size to be uniform with a smooth, light colored coat and a clear hilum (Wei and Chang, 2004). Seed coat staining from weeds and bleeding hilum from disease are aesthetically undesirable.

B. Fermentation and Free Sugar Content

The production of most fermented soyfoods involves multiple microorganisms to carry out a complex series of reactions, but natto is a unique because only one microorganism is responsible for fermentation, *Bacillus natto* (Liu, 2005). The *Bacillus* strain used and the conditions of fermentation can affect natto quality and vary between producers but the general procedures are the same (Wei et al., 2001). The free sugars: sucrose, stachyose, and raffinose, are released during soaking and steaming and provide energy for fermentation. During fermentation extracellular enzymes produced by *Bacillus natto* react with soybean sugars to produce a viscous material referred to as mucilage. The amount and viscosity of the mucilage are important natto characteristics and contribute to natto's unique taste and smell (Taira, 1990). Soybeans with high sugar content and small seed size are desirable because they allow more bacterial colonies to develop, produce more enzyme, and create more mucilage (Cui et al., 2005). The composition of free sugar content is important because the polysaccharides are consumed at differing rates during fermentation (Taira, 1990). Sucrose is easily broken down and consumed before stachyose or raffinose, and high sucrose content can increase the temperature and halt fermentation. Oligosaccharides, such as stachyose and raffinose, are digested later in fermentation and provide a gradual fermentation environment (Taira,

1990). Therefore, the amount of free sugars should be higher than conventional cultivars to increase mucilage but a balance between the soluble sugars is important for proper fermentation.

C. Soybean Seed Water Absorption

Water absorption is an important factor when producing quality food from various dry beans (Hsu et al., 1983; Mullin and Xu, 2001; Pan and Tangratanavalee, 2003; Cober et al., 2006; Shao et al., 2007). Water absorption can affect texture, taste, processing time, and the method of production for soyfoods such as tofu, natto, miso, and soymilk. It is especially important for natto manufacturers because 59% to 60% of finished product is composed of water (Wei et al., 2001). Water absorption in soybean is regulated by the seed coat to prevent imbibitional damage and has been shown to initiate opposite the hilum (McDonald et al., 1988). Imbibed water is trapped between the seed coat and the cotyledon where it can slowly hydrate the cotyledon and initiate germination. Many wild Fabaceae species developed a trait referred to as hardseededness (Potts, 1978). Hard seeds have seed coats that are impermeable to water and require a biological or physical conditioning before water imbibition and germination. Plants evolved this mechanism to increase their fitness in nature but the trait is undesirable in modern cultivars because it leads to non-uniform stands and volunteer seeds (Potts, 1978; Rolston, 1978). Breeding efforts have reduced the hardseeded trait in modern conventional cultivars to eliminate the germination problem. Hardseededness also causes problems in food preparation when whole seeds must be soaked and hardseeded seeds fail to imbibe water (Hsu, 1983). Plant breeders and researchers are still addressing the

issue in food grade cultivars because roughly one quarter of their pedigree is exotic hardseeded germplasm (Cui et al., 2005).

Multiple methods have been developed to rank and quantify a soybean cultivar's water absorption potential. Assays such as water absorbing capacity, water retention capacity, firmness, and rate of water absorption provide information used to compare potential cultivars for use in food production (Wilson, 1995; Geater et al., 2000; Wei and Chang, 2004; Cober et al., 2006). Although assays developed to quantify water absorption are effective at identifying potentially useful food grade cultivars, they do not explain the cause of hard seeds. Understanding the factors that cause hardseededness would assist researchers in developing cultivars with superior water absorption but conclusive results are not available. A possible reason this trait has eluded researchers is because it encompasses many seed attributes. It is understood that the seed coat controls water absorption, but the seed coat is comprised of many layers and chemical constituents, making it difficult to identify one controlling factor. Research results addressing the control of different seed coat components on water uptake are presented below.

Natto producers can alter soybean seed water absorption by changing the duration and conditions of soaking. Increased water temperature can increase the water absorption rate and reduce the time needed to reach full hydration (Pan and Tangratanavalee, 2003). The salt content of the water used for hydration can also affect water absorption. Hsu et al. (1983) found that water with high concentrations (1% - 5%) of sodium bicarbonate, which increases pH to between 8 and 9, slowed water absorption. It has also been shown that beans soaked in methanol or ethanol have a higher water absorption rate and capacity (Arechavaleta-Medina and Snyder, 1981; Hsu et al., 1983).

This result was attributed to the removal of alcohol soluble substances in the seed coat that had inhibited water absorption. Although pre-treating soybean seeds to increase water absorption in natto manufacturing is possible it is avoided because additional processing requires more resources and can negatively affect the final product. Therefore, developing soybean cultivars with desirable water absorption characteristics through breeding is the preferred approach.

Soybean seed calcium content was identified as a major controlling factor for water imbibition in initial research but these findings have not been corroborated. The study indicated that high calcium levels in soybean seed coats were negatively correlated with total water absorption, suggesting that high levels of calcium lead to greater levels of hardseededness (Saio, 1976). These findings are not supported by other studies of soybean water absorption and seed coat composition (Cober et al., 1997a; Mullins and Xu, 2001; Wei and Chang, 2004). These researchers conclude that although calcium is an important component of soybean seeds, it does not appear to play a singular role in total water absorption.

Carbohydrate content of soybean seeds has been analyzed to determine its effect on water absorption but no conclusive association has been reported. Geater et al. (2000) found a significant positive correlation between total water absorption and sucrose content. Both high total water absorption and high carbohydrate content are desirable for natto production. Based on the positive correlation, breeders could use indirect selection to increase water absorption by selecting for higher free carbohydrate content or vice versa, but these findings were not supported by other research results. Wei and Chang (2004) reported a significant negative correlation between water absorption and sucrose content, the opposite finding. The discrepancy between the findings of Geater et al.

(2000) and Wei and Chang (2004) may be due to differences in lab procedures but more likely show that general correlations can be limited to specific environments and genotypes. It does not appear, however, that free carbohydrate content is a direct controlling factor for water absorption in soybean.

Multiple research projects indicate that cultivar differences for seed permeability are strongly influenced by different seed coat layers (Rangaswamy and Nandakumar, 1985; Mullin and Xu, 2001; Ma et al., 2004; Meyer et al., 2007; Shao et al., 2007). The mature soybean seed coat has three general layers: epidermis, hypodermis, and inner parenchyma (Carlson and Lersten, 2004). The epidermis is comprised of closely packed palisade cells, also called macrosclereids, surrounded on the outside by a cuticle. The palisade layers are cutinized making them hydrophobic and impermeable to gas. The second layer, hypodermis, is composed of cells that have the appearance of columns, offering strong support and intracellular space. The final layer, inner parenchyma, consists of multiple layers uniformly distributed throughout the seed except at the hilum. The following reviews address the roles of the different seed coat layers on water imbibition.

Research on least snoutbean, *Rhynchosia minim,* addressed the influence of the outer waxy surface, the adcrustation, and the palisade layers on seed water absorption (Rangaswamy and Nandakumar, 1985). The researchers did not find a single seed coat layer that alone rendered a bean impermeable to water. Chemical analysis on the different layers, however, found waxy substances and hemicellulose as being responsible for decreased water absorption.

Mullin and Xu (2001) fractionated chemical compounds from seed coats of different soybean cultivars and correlated them to water absorption. The two classes of

chemicals studied were cations and macrochemicals. This study found no correlation between cation content, specifically calcium, and seed coat permeability. The analysis of complex carbohydrates identified cultivar differences related to permeable and impermeable seed, but did not indicate a single chemical component controlling permeability. The notable chemical differences between impermeable and permeable cultivars were for lignin, pectin, and hemicellulose content. Impermeable seeds had the lowest lignin and pectin content and highest hemicellulose content. The authors propose that these complex carbohydrates play the most significant role in determining seed coat permeability. They concluded that a combination of chemical compounds involved in cell wall structure and intracellular composition are probably responsible for water imbibition.

Analysis of soybean seed coat structure during water imbibition identified cracks in the cuticle covering the palisade layer as being the one consistent difference between hardseeded seeds and permeable seeds (Ma et al., 2004). These researchers concluded that the cuticle is responsible for controlling water absorption. The study used a scanning electron microscope (SEM) and various staining techniques to observe the seed coat during water imbibition. Six soybean cultivars were used ranging from impermeable to very permeable. Of multiple seed coat surface features observed, only cracks in the cuticle covering the palisade layer were consistently correlated with water movement absorption. These were micro-scale cracks in a seemingly intact seed coat identified using a SEM and correlated with permeable seeds. The authors speculate that environmental influence during seed desiccation or storage affects the seed coat and produces the palisade cracks. They noted that seed coat thickness and surface features were the same between permeable and impermeable cultivars, but the density of the

impermeable seed coat cuticle was higher. Further analysis of seed coat density would provide evidence to support this claim.

Ma et al.'s (2004) claim that the cuticle is the major factor controlling water absorption is supported by two other research reports. Arechavaleta-Medina and Snyder (1981) were able to manipulate seed water absorption and change seed from impermeable to permeable by either removing a section of the cuticle or by soaking the beans in ethanol or methanol. The ability to remove a portion of a seed's cuticle and alter its ability to absorb water provides additional evidence that the cuticle is a major water absorption barrier. The change in water absorption after alcohol treatment indicated a chemical change in the seed coat resulting in a breakdown of the physical barrier to water entry.

Another research group used water absorption curves and different seed conditions to further suggest that the seed coat was a rate limiting structure in seed water absorption (Meyer et al., 2007). Seed coats were removed from impermeable and permeable seeds and water imbibition rates were recorded. The rate of water absorption was no different between the seed types when the seed coats were removed. These results indicate that the seed coats were rate limiting for both seed types and provides evidence that the difference between hardseeded seeds and normal seeds is their seed coat and not a difference in cotyledon structure.

Shao et al. (2007) conducted chemical analysis on the cuticle of permeable and impermeable cultivars to determine what chemical differences may be controlling water imbibition. Multiple chemical procedures and extractions were performed to determine if the wax or cutin component of the seed coat was responsible for limiting water absorption. The researchers concluded that the wax component was less important than

the cutin component because when wax alone was removed from the seed coat hard seeds failed to imbibe water, while hard seeds treated with a hot alkali bath that removed or degraded the cutin layer were rendered permeable. The researchers concluded that parts of the cuticle were responsible for impermeability. A chemical profile comparison between hard and normal seeds identified that impermeable seeds had higher amounts of hydroxylated fatty acids than did permeable seeds. These chemical compounds may be responsible for creating stronger bonds within the seed coat preventing the small cracks found by Ma et al. (2004). The researchers point out that interaction between seed coat cutin and underlying carbohydrates may play a significant role in determining seed coat integrity, in accordance with conclusions from Mullin and Xu (2001).

Heritability for Related Natto Quality Traits

Heritability estimates for characteristics related to natto quality in soybean are limited. Water absorption is a necessary step when producing many soyfoods from whole soybeans. One research group addressing cookability and genetic variation in soybean reported a high broad sense heritability estimate (H^2=0.81) for soybean seed water uptake (Mwandemele et al., 1984). The study was conducted during two years in two locations and no information is provided about how heritability was estimated. The high broad sense heritability estimate indicates that the variation between genotypes was large compared to environmental variation within genotypes (Griffith et al., 2005). There was also a significant positive correlation between water absorption values collected during the two years indicating high repeatability. The broad sense heritability estimate is specific to the genotypes and environments tested. Broad application is not possible,

but the results indicate developing soybean cultivars with consistent high water absorption is possible.

Another important trait for soybean food grade cultivars is resistance to seed coat cracking. Heritability and parent – offspring regression were calculated for incidence of seed coat cracking and trait inheritance using a susceptible by resistant soybean cross (Okabe, 1996). Heritability estimates for soybean seed coat cracking were low (h^2=0.05) in the F_2 generation but high (h^2=0.76) in the F_3 generation, a result the authors contributed to the method used for selection. There was a significant positive parent-offspring correlation between the F_3 and F_4 generations indicating high heritability, but the correlation was much lower between the F_2 and F_3 generations. These results indicate that soybean seed resistance to damage is heritable and progress through selection is possible.

Similar durability characteristics are desirable in bean production for human consumption and heritability has been estimated. Reichert and Ehiwe (1987) reported broad sense heritability estimates from variance components for seed coat breakage of 57.5% and 56.2% in a two year experiment using field bean (*Pisum sativum* L.). There was also a positive correlation for seed coat breaking results between the two years indicating that similar rankings can be expected across years. The researchers also reported significant location and year effects on seed coat breaking indicating environmental influence for the trait. They conclude that higher than average temperatures and below average precipitation lead to more seed coat breaking at some locations over the two years. Similar broad sense heritability estimates for seed coat breakage (37% to 57%) have been reported for snap bean (*Phaseolus vulgaris* L.) (Dickson and Boettger, 1977). These results indicate there is genetic variation for seed

resistance to damage in bean species and it is heritable, but it is difficult to determine how well these results translate into soybean seed durability.

Heritability estimates have also been reported for the natto quality trait rate of water absorption. Cober et al. (2006) calculated broad sense heritability estimates using variance components for rate of water absorption in a soybean trial grown at six locations over two years. Moderate heritability estimates (H^2=0.42) were observed for the rate of water absorption constant, which estimates the initial slope of the water uptake curve. This indicated that genetic variability was slightly higher than environmental variation for the material tested. Genetic and environmental influence were significant for trait expression indicating data from multiple environments is necessary to determine a cultivars genetic potential, but the research concluded that genetic control of the trait was more important and that superior natto cultivars could be developed for the trait.

Genotype × Environment Interaction for Natto Quality Traits

Research projects have been conducted to examine genotype × environment interaction in soybean food grade cultivars for multiple quality traits. Taira (1990) summarized characteristics necessary for soybean natto quality traits and environmental influence for genotypes developed and grown in Japan. The research addressed genetic and environmental influence on seed composition traits and identified soybean traits that are necessary for producing different soyfoods. The report concluded that seed chemical components such as free sugars and fatty acids were influenced by location, year, and cultural practices. The results indicated that the suitability of a soybean cultivar for making different soyfoods, however, was controlled mainly by genotype and the researcher concluded that cultivar is the most significant controlling factor of soyfood

quality. The results illustrate that improved soyfood quality can be achieved through cultivar selection.

The effects of genotype and environment on small seeded fraction, hard seed, water uptake, and seed chemical components of natto soybeans were investigated for two different data sets grown in five Canadian locations during two different years (Cober et al., 1997a). Genotype variation was present for all traits except hard seed in one year. The location effect was not significant for any trait and year was only significant for a limited number of traits. There was significant genotype × environment interaction for most traits studied, but the researchers state that genotype was a more significant factor explaining natto quality trait variation than either location or year. The researchers conclude that environmental variation was less important than genotype for determining natto quality and developing superior natto cultivars is possible, a conclusion similar to that of Taira (1990).

Geater et al. (2000) conducted a multi year trial involving small seeded natto soybean cultivars grown in three locations over two years in Iowa. Data were collected for natto quality traits: water absorption, water loss, hardness of steamed seeds and natto, darkness of steamed seeds and natto, and other chemical seed traits were also evaluated. The statistical analysis was performed independently for each location over two years and for each year over the three locations because the combined analysis of variance did not produce significant location or year effects. Genotype, location, and year differences were observed for all traits. For water absorption, both consistent and inconsistent cultivars with desirable trait expression were identified. This indicates that selecting high, consistent water absorption is possible but genotypes should be evaluated over different environments to avoid selecting environmentally sensitive cultivars. Genotype

differences were reported for water loss after steaming with values ranging from 9% to 14%. A significant genotype × year interaction was observed for water loss after steaming indicating rank change occurred between the two years results. A significantly positive correlation was reported for water absorption and water loss after steaming. These results suggest cultivars that absorb the most water during soaking also lose the most water during steaming. Cultivars that displayed a significant rank change for water absorption also varied for water loss after steaming providing further evidence the two traits are related. A significant negative correlation was reported between water absorption and natto firmness indicating that high water absorption is related to soft natto. No correlation was reported between seed size and any trait tested. In summary, the study indicates that genotype differences were present for many natto quality traits and that consistent performance over years and locations were observed, however, significant rank change is possible and evaluating potential natto cultivars in multiple environments is necessary.

Stability

Natto quality is influenced by soybean seed attributes, and cultivars are developed to meet desired natto quality standards and ensure a high quality product. Research projects on natto quality traits indicate, however, that environmental factors influence trait expression and alter expected phenotypes. Environmental conditions such as hot dry weather during seed desiccation or alternate wetting and drying conditions are potential factors that degrade seed quality (Okabe, 1996). Quality damage can also occur via anthropogenic factors such as harvest damage during combining or harvesting seed at inappropriate seed moisture. Because soybean seed quality is influenced by cultivar and

environment, trait stability is essential when developing natto cultivars. No explicit reports of cultivar stability for natto quality traits are available. The following section on stability statistics provides information and contrasts the use of different stability estimates.

Trait stability refers to consistent trait expression across variable environmental conditions. Traits are considered unstable if they display significant environmental influence or differential response across environments. Stability is not an absolute term and the definition and calculation of stability depend on its scientific context (Lin et al., 1986). Stability is broken into three concepts by Lin et al., (1986) of which two are commonly used and will be discussed.

The first type of stability has a broad application and is referred to as biologic homeostasis, static stability, or environmental variance (Lin et al., 1986; Becker and Leon, 1988). Type I stability refers to responses that do not change across environmental conditions and remain around a mean value with a small variance. Type I stability is calculated as follows:

$$\text{Equation 1: } S_i^2 = \sum_j^e [(X_{ij}-X_i)^2 / (e-1)]$$

S_i^2 is the environmental stability for the *i*th cultivar, X_{ij} is the average observed genotype response for the *i*th cultivar in the *j*th environment, $X_{i.}$ is the average genotype mean of the *i*th cultivar across locations, and *e* is the number of environments tested (Lin et al., 1986). An $S^2 < 1$ would indicate environmental stability with $S^2 = 0$ perfect stability.

Lin et al. (1986) suggested type I stability is not usually applied to breeding selection because it is often associated with poor response in productive environments and not practical when applied to a wide range of growing environments. Research addressing stability for bread making quality characteristics in wheat has shown that

environmental variance (S_i^2) is not always correlated with quality characteristics, providing evidence that in certain situations type I stability can be used to identify genotypes with desirable trait response and stable performance (Robert, 2002; Lemelin, 2005). Type I stability is also appropriate for situations where relative rank is less important than overall performance such as with quality characteristics (Becker and Leon, 1988). It is not practical to apply this type of stability estimate to an experiment conducted over a wide geographic range, such as a continent, because it is improbable to find a cultivar that is both type I stable and highly productive under vastly different growth conditions. Type I stability estimates for similar environmental conditions, however, may provide useful estimates of environmental homeostasis (Lin et al., 1986).

Type II stability estimates refer to relative measures of a response for a group of cultivars, also called ecovalence, interaction variance, or dynamic stability (Lin et al., 1986; Becker and Leon, 1988). It uses genotype, environment, and interaction mean values collected in an experiment to determine genotype influence on the GEI (Lemelin et al., 2005). Ecovalence is measured as follows:

$$\text{Equation 2: } W_i^2 = \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2$$

W_i^2 is the ecovalence for the *i*th cultivar, X_{ij} is the average observed genotype response for the *i*th cultivar in the *j*th environment, $X_{i.}$ is the average genotype mean of the *i*th cultivar across locations, $X_{.j}$ is the average location mean of the *j*th environment across genotypes, and $X_{..}$ is the overall mean for the experiment (Lin et al., 1986). Low W_i^2 values near zero are considered stable.

Type II stable genotypes have similar responses to the mean response of all genotypes in the experiment and display variation across environments due to environmental variation (Lin et al., 1986). Calculation of type II stability uses means for

all genotypes and environments used in the study so interpretations of the results are limited. Also, type II stability is relative and does not indicate the same absolute stability provided by type I stability estimates. Therefore, type II stability is less appropriate for responses such as quality measurements because these responses must remain around a certain value. It is useful for measuring stability among a defined set of genotypes for a relative response such as yield.

Rationale and Objectives

The first objective of this research was to identify and evaluate assays that predict natto quality, and include them in a multi environment trial to determine if genetic variation exists in Virginia Tech germplasm. Communication with natto manufacturers indicated that new protocols to measure water absorption and seed quality needed to be integrated into the natto selection process to help identify superior natto cultivars. Initial efforts focused on evaluating multiple assays found in primary literature and through personal communication that may help identify superior natto cultivars. Rationales for selecting the assays included in the multi environment trial are provided below and further experimental details can be found in the materials and methods. Collecting data on these traits will provide insight into their genetic control and foresight into selection gains should they be integrated into selection of superior natto cultivars. We hypothesized that not all assays would produce meaningful data and some of them would be inappropriate to integrate into natto selection.

Total water absorption and water loss after steaming are important measures of natto quality widely recognized by natto producers. Water absorption is an important trait for various reasons addressed earlier and a common assay found in literature (Taira,

1990; Cober et al., 1997a; Geater et al., 2000; Wei and Chang, 2004). Water loss is easily collected on the same samples and provides a clear picture of how much water is absorbed and maintained during initial natto production. These characteristics are predominant quality factors and without high total water absorption and retention potential cultivars will not be accepted for natto production. Initial results indicated these tests produced accurate data and the assays will aid in selecting superior natto cultivars.

Rate of water absorption adapted from Cober et al. (2006) was another assay tested. They reported that total water absorption was independent of the rate of water absorption and provided another trait to select superior natto cultivars. This trait is desirable for producers attempting to cut production time and potentially lower cost. Cober et al. (2006) also suggests that slow initial water uptake may produce undesirably hard natto. This trait is unique from other assays used to evaluate the Virginia Tech natto germplasm and it should help identify superior natto cultivars. Initial testing was conducted to optimize the protocol before analyzing samples from the multi environment trial.

The last assay used to specifically address natto quality was seed coat deficiency. Soybean seed coats are subjected to a variety of physical stresses during harvest, shipment, and production. These events can lead to cracks in the seed coat and result in reduced commercial value (Nakamura et al., 2003). Physical damage that results in cracked seed coats occurs on a macro scale and damaged seeds can be identified and removed before natto production. During the water absorption stage of natto production more subtle stress is put on the seed coat but seed coat cracking can still occur and result in unacceptable natto. To address this concern we developed the seed coat deficiency assay to assess seed coat strength during the initial water absorption stage. This assay is

intended to identify cultivars with deficient seed coats that could negatively affect natto quality. The assay was developed from information gathered during Dr. Rainey's visit with individuals involved in the Japanese natto industry (personal communication, 2006). The protocol was developed using descriptions of the deficient seed phenotype and two cultivars that had been characterized for water absorption quality ratings by natto producers. After multiple trials a protocol was developed that identified repeatable estimates of different genotypes and was used to evaluate samples from the multi environment trial. We hypothesize that using the four natto quality assays: total water absorption, water loss after steaming, seed coat deficiency, and rate of water absorption will aid in identifying superior natto cultivars.

The second objective of the research was to evaluate environmental influence and possible genotype × environment interactions (GEI) for Virginia Tech natto cultivars. Environmental influence and GEI effects have been studied for other soybean traits, but few trials have been conducted to assess natto quality. For the limited studies related to environmental influence on natto quality (Taira et al., 1990; Cober et al., 1997a; Geater et al., 2000), it is inappropriate to assume the findings apply to Virginia Tech cultivars grown in Virginia (Griffiths et al., 2005). For these reasons, it is necessary to address genotype × environment interactions for natto quality responses using Virginia Tech natto genotypes and Virginia growing locations.

There are two specific reasons why collecting GEI data for Virginia Tech natto germplasm would be useful. First, neither the parents of current natto cultivars nor the current cultivars themselves have been the subject of a GEI study in this region. This is significant because the Virginia Tech natto germplasm is grown extensively in Virginia and other Mid-Atlantic states. Natto cultivars are lower yielding than conventional

cultivars but a price premium is paid to offset the lower yield potential and entice growers to participate. The premium may not be granted, however, if the soybean crop does not meet natto quality standards and cannot be sold in the food grade market. Therefore, growers assume risk when growing natto cultivars because unpredictable environmental conditions could reduce their crops quality below acceptable standards. Identifying cultivars with less environmental influence and stable natto quality characteristics would mitigate this risk and benefit both growers and manufacturers. This provides justification for examining potential GEI and stability in germplasm used to develop natto cultivars for the Mid- Atlantic region. This experiment is not designed to determine the environmental conditions causing changes in natto seed quality but rather to evaluate the Virginia Tech natto germplasm for consistent superior natto cultivars over a range of environments.

The second impetus for examining GEI in Virginia Tech natto germplasm is evidence that GEI interactions are a problem for natto cultivar development (Dr. Glenn Buss, unpublished data, 2005). The standard natto cultivar, MFS-591, planted in the region yields significantly lower than commodity cultivars, making it difficult to secure adequate acreage to meet market demands for natto soybeans. A solution would be to identify another cultivar with the same or better quality standards that is also higher yielding than MFS-591. In previous research, cultivar MFS-511 was evaluated for potential use in natto production because it is higher yielding than MFS-591 and has similar quality characteristics. However, repeated evaluation of MFS-511 showed inconsistent natto quality traits in samples grown at different locations over multiple years. These results provided preliminary data that natto quality response is influenced by growing conditions and that a multi environment study was needed. We hypothesize

that environmental conditions influence natto quality trait expression and that by conducting a multi environment study to evaluate Virginia Tech natto germplasm we can identify genotypes with less environmental sensitivity to be used in natto production or serve as parents in natto cultivar development.

The final objectives of the MET was to determine if soybean planting systems produced superior natto quality, and determine if cultivar selection needed to be targeted for either system specifically. The experiment included planting system as a variable because growers involved in natto production have indicated it may influence natto quality (Montague Farms, personal communication, 2006). For this experiment, the planting systems used were full season and double crop. Full season refers to conventional plantings in which seed is planted no-till into fields early in the season, late May for this region. Double crop soybeans are planted after winter wheat or barley harvest, directly into the untilled stubble in mid to late June. The two planting regimes are mainly defined by planting date but other environmental differences such as soil moisture, soil texture, and pathogen presence also create differences between full season and double crop systems. Based on anecdotal evidence, we hypothesize that growing soybeans in a double crop system will produce seeds with superior natto quality compared to the same genotypes grown in a full season system.

Material and Methods

Experimental Design

To determine the effects of genotype, environment, and possible interaction on natto quality traits, a multi-location experiment was conducted over two years. Twelve genotypes (Table 2) all group V maturity, were grown in a randomized complete block design with three replications per entry at four locations in 2006 and five locations 2007 (Table 3). The soil types at Blacksburg were Hayter loam and Guernsey silt loam. The soil type at Warsaw was Kempsville loam and the soil type at Mt. Holly was State fine sandy loam. Plots at Mt. Holly were irrigated. The soil type for Suffolk was Eunola fine sandy loam and Dragston sandy loam. The plots were not planted in the same exact field location between the two years because of crop rotation needs, but the locations were generally the same.

All locations over the two years were planted using a seeding rate of 432,098 seeds/hectare for full season plantings and 543,209 seeds/hectare for double crop planting. These seeding rates reflect Virginia Cooperative Extension recommendations (Holshouser and Alexander, 2001). A border was left within each alley to help minimize yield inflation due to non-competition and plots were end-trimmed prior to harvest. Normal production practices were followed to provide a healthy growing environment, and care was taken to minimize differences in cultural practices among locations. A preventative fungicide, Quadris (Azoxystrobin) or Headline (Pyraclostrobin), was applied to all plots at the R2 to R3 soybean development stage (Fehr and Caviness, 1977) to emulate natto soybean production practices and help ensure seed quality (Bryan Taliaferro, personal communication, 2006).

Table 2. Soybean entries included in multi-year trial grown for two years.

Entry	Name	Pedigree	End Use	Seed Size g/100 seed	Seed Size Std Error
1	MFS-591	Camp × Rocky	Natto	8.79	0.15
2	MFS-511	SS 516(3) × V93-0687	Natto	9.55	0.14
3	SS-516	Private	Conventional	14.00	0.25
4	V00-3493	SS 516(3) × V93-0687	Natto	10.31	0.21
5	V00-3488	SS 516(3) × V93-0687	Natto	11.09	0.20
6	V00-3636	SS 516(3) × V93-0687	Natto	10.36	0.16
7	V01-4937	V92-0974 × [V93-0706 × DP3519S]	Natto	9.90	0.17
8	Teejay	Hutcheson × Clifford	Conventional	17.64	0.32
9	Camp	Essex × G. soja	Natto	8.16	0.17
10	Hutcheson	V68-1034 × Essex	Conventional	16.06	0.28
11	MFS-553	Essex × Camp	Natto	9.36	0.18
12	V03-5794	V99-5089 × Essex	Conventional	17.36	0.22

Table3. List of locations and planting description included in multi-year trial grown for two years.

Location	Planting Type	2006 Planting Date	2007 Planting Date	Row Spacing	Harvested Plot Size
Blacksburg	Full Season	June 1st	June 8th	4 rows 14"	20' × 4.67'
Mt. Holly	Full Season	June 17th	June 27th	5 rows 7.5"	12' × 4.63'
	Double Crop	July 5th	June 27th	5 rows 7.5"	12' × 4.63'
Suffolk	Full Season	May 30th	May 16th	5 rows 15"	17' × 3.75'
	Double Crop	June 30th	June 21st	5 rows 15"	17' × 3.75'
Warsaw	Full Season	June 17th	June 6th	5 rows 7.5"	12' × 4.63'
	Double Crop	July 10th	July 3rd	5 rows 7.5"	12' × 4.63'

Agronomic Data

A variety of agronomic data were collected for each plot. Maturity data were collected when soybean plants reached the R8 development stage and 95% of the pods in a plot reached mature pod color and recorded as days after August 31st. Plant height was measured in inches and plant lodging was measured on a continuous scale from 1 to 5, with 1 being completely erect and 5 being 75% to 100% of plants leaning or down. Plant

height and lodging notes were collected just prior to harvest to ensure accurate data. In Blacksburg, Warsaw, and Mt. Holly all five plot rows were harvested while in Suffolk only the inner three rows were harvested.

Post harvest, seeds from each plot were cleaned through a 15/64" screen to remove any harvest debris. Seed moisture was recorded for each plot using a Dicky John grain analyzer. Plot weights were recorded in grams and used to calculate yield based on 13% moisture content reported as grams per hectare. Seed size was estimated by weighing 100 seeds arbitrarily selected from each plot and recorded as grams per 100 seed.

Natto Quality Data

Seed subsamples (500g) of each plot were arbitrarily collected and stored in a Conviron growth chamber at 14° C and 60% relative humidity to maintain seed integrity and to equilibrate seed moisture. Moisture content was randomly sampled to track moisture equilibration using a Dicky John grain analyzer. After six to eight weeks, seeds were between 10.0% and 12.2% moisture for all plots at all locations and natto quality data were collected. All natto quality data were collected by location in ascending plot order.

A. Water Absorption and Water Loss after Steaming

Water Absorption was determined for each plot grown in 2006 and 2007 at all locations. The protocol was modified from Geater et al. (2000) and Wei and Chang (2004). An arbitrarily selected sample from the 500 g sample stored in the growth chamber was visually screened to remove seeds with cracked seed coats. A 20 g

subsample was taken from this sample and placed in a plastic container and immersed in 100 ml of deionized water at room temperature for 16 hours. After 16 hours, the samples were poured into a colander and drained for 15 seconds. The samples were then briefly placed on a paper towel to remove any excess surface water. These samples were then reweighed. Water absorption was calculated as: water absorption = (weight after water absorption / initial weight) × 100 (Geater et al., 2000).

The samples from water absorption tests were then used to calculate water loss after steaming. After each sample was reweighed for water absorption, it was placed into a dry plastic container. The samples were steamed in an autoclave at 121 °C and a pressure of 15 p.s.i. for 18 minutes. After the steaming process, the samples were removed and allowed to cool at room temperature for 10 minutes. The samples were weighed, and water loss was calculated as: water loss = ((weight after steaming - weight after water absorption) / weight after water absorption) × 100.

B. Seed Coat Deficiency

The seed coat deficiency assay was developed after personal communication with Japanese natto manufacturers and used to assess seed coat strength during initial water uptake. The test identifies seed coats with flawed integrity that will result in undesirable decoated seeds at the end of natto production, therefore high counts are undesirable. The test was run on 100 seeds arbitrarily chosen from each moisture equilibrated subsample. Seeds that were not representative of natto soybeans, such as those with cracked seed coats, seed discoloration, or excessively flat seeds were removed from the sample, and additional seeds were arbitrarily pulled until 100 seeds were present. These criteria were similar to those used by natto manufacturers and are intended to reflect seed quality of

soybeans used in natto production (Montague Farms, personal communication, 2006). The 100 seed sample was placed in a plastic container with 50ml of water for 10 minutes. After soaking, the seeds were drained and visually scored with the naked eye for seed coat deficiency. Seeds were identified as deficient after soaking if the seed coat was cracked, severely blistered around the hilum, or if the seed coat detached from the hull. Seeds with detached seed coats were identified by transparent seed coats. Scoring erred on the side of caution and ambiguous seeds were placed into the deficient category to help identify superior cultivars. Counts of deficient seeds were used in statistical analysis.

C. Rate of Water Absorption

The rate of water absorption was also determined for each plot using a protocol adapted from Cober et al. (2006). A 50 gram sample was arbitrarily selected from each moisture equilibrated subsample. The samples were visually screened with the naked eye for cracked seed coats, seed discoloration, or excessively flat seeds and removed as described above. The samples were placed in a nylon mesh bag and then submerged in a plastic container with 500ml of deionized water. At set time intervals of 1, 3, 6, and 24.5 hours, the bags were removed, drained, and the seeds were placed on paper towels to remove surface water and then weighed. Timing was important to ensure each sample soaked for the specified time and needed to account for the time needed to start at plot 1 and collect data to plot 36. For this reason, samples were submersed at staggered times to account for extra time during drying, weighing, and recording between samples. This minimized error and ensured each sample soaked for the equivalent time. The sample weights at specified time intervals were used to estimate the rate of water absorption, *b*,

and the water holding capacity, *a*, for each sample based on the equation Y= a[1− exp(-bX)] (Cober et al., 2006). The variable *X* represented the different times used for sampling and the variable *Y* was the weight of the plot at that time. The equation was developed to estimate the exponential rise to a maximum for weight gained during seed soaking. The procedure NLIN was used in SAS (SAS Inst., Inc., Cary, NC) to analyze the data and estimate the variables *a* and *b*. The rate constant b (min^{-1}) was used for combined analysis of variance.

Statistical Analysis

A. Assumptions of Normality

Statistical analysis was computed in SAS using the MIXED and UNIVARIATE procedures. Assumptions for analysis of combined variance for a mixed effect model were checked for all collected variables using residuals, predicted values, and studentized residuals generated with PROC MIXED. Assumptions for the combined analysis of variance are normal distribution of data and homogeneous variance (Elliot, 2000; Lomax, 2001).

Residuals were submitted to PROC UNIVARIATE to generate normal probability plots of residuals. Normal distribution of residuals is indicative of normal distribution. Shapiro-Wilkes formal test of normality was used in conjunction with plots to determine normal distribution, and p- values greater than 0.05 were considered normally distributed (Lomax, 2001; Ott & Longnecker 2001).

Homogeneity of variance was checked in these data using plots of residuals vs. predicted values and plots of residuals by groups (Elliot, 2000; Lomax, 2001; Ott & Longnecker 2001). Random vs. predicted plots with a horizontal line at zero indicates

residual data with no error. In practice, a random scatter of points equally representing high and low error with no discernable pattern indicates equal variance across the data (Lomax, 2001; Ott & Longnecker 2001). Plotting residuals by treatments or groups of treatments is another way to visualize patterns of variance across data to check for homogeneous variance (Elliot, 2000; Lomax, 2001). In this case, variance ranges should be relatively equal across groupings to ensure the assumption of equal variance is met.

Outlier/problem data points were identified in the following way. A studentized residual is an estimate of the residual's standard error. Observations that could be potential outliers were identified as possessing studentized residual values greater than 3. These data points were checked for symmetry to determine if they were part of natural variation or if they represented some sampling error. These criteria were used to check the data for analysis of variance assumptions and if data were removed the new data set was used for further analysis. Data sets that appeared to violate the assumptions of normal distribution and equal variance were transformed and re-analyzed according to the above criteria to determine if transformation created an improvement.

B. Models

Two models were used for combined analysis of variance to address two questions. One: the effect of full season vs. double crop planting system on natto quality traits, and two: determine the influence of growing environment on natto quality traits (Model 1 and 2, respectively). Data were analyzed using model 1 to address how genotypes responded to three environmental factors: planting location, planting system, and year. It specifically addressed how planting system effected natto quality traits. Genotypes and planting systems were considered fixed effects because they represented

specific genotypes and planting systems, and specific inferences were desired. Locations and years were considered fixed effects because they do not accurately represent all possible locations or years. Replicates were nested within locations, years, and planting systems and considered random effects because inference was unimportant and biologically irrelevant. Model 1 was used to analyze full season vs. double crop data

Data were analyzed with model 2 to address how genotypes responded to environments and years. Environments were combinations of locations and planting systems. Model 2 addresses the broader genotype × environment interaction for natto quality and agronomic traits. Genotypes were again considered fixed effects because they were chosen specifically for this project so inferences could be made about their performance. Environments and years were considered fixed effects because they did not accurately represent all possible environments or years. Replicates were nested within environments and years and considered random effects because inference was unimportant and biologically irrelevant.

C. Stability Statistics

Environmental stability and ecovalence equations 1 and 2 were computed using PROC MEANS of SAS to estimate cultivar stability. To calculate environmental stability location, planting system, and year factors were combined into one environment term for a total of 13 environments with an individual cultivar observation denoted as (X_{ij}). Cultivar means ($X_{i.}$) were averaged over the 13 environments.

Ecovalence was calculated in a similar manner using the same 13 environments. The individual cultivar observations (X_{ij}) and cultivar means ($X_{i.}$) were calculated the same way. Environment means ($X_{.j}$) were the average of all cultivars at a location,

planting system, year combination. The grand mean ($X_{..}$) was the average of every observation. Both statistics are parametric and require assumptions of normality. For this reason, any observations removed during the combined analysis of variance were removed to calculate stability statistics.

Equation 1: $S_i^2 = \sum_j^e [(X_{ij}-X_{i.})^2 / (e-1)]$

Equation 2: $W_i^2 = \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2$

Model 1. Mixed effects model for combined analysis of variance for *i* genotypes, grown in *j* locations, during *k* years, using *l* planting systems, and *m* replications.

Effect	Degrees of Freedom	Fixed or Random	Denominator Degrees of Freedom for F Test
Genotype *i*	$(i-1)$	Fixed	$(j\times k\times l)\times(i-1)\times(m-1)$
Genotype×Location *ij*	$(i-1)\times(j-1)$	Fixed	$(j\times k\times l)\times(i-1)\times(m-1)$
Genotype×Year *ik*	$(i-1)\times(k-1)$	Fixed	$(j\times k\times l)\times(i-1)\times(m-1)$
Genotype×Planting System *il*	$(i-1)\times(l-1)$	Fixed	$(j\times k\times l)\times(i-1)\times(m-1)$
Genotype×Location×Year *ijk*	$(i-1)\times(j-1)\times(k-1)$	Fixed	$(j\times k\times l)\times(i-1)\times(m-1)$
Genotype×Replicate(Location×Year×Planting System) *(jkl)im*	$(j\times k\times l)\times(i-1)\times(m-1)$	Random	
Location *j*	$(j-1)$	Fixed	$(j\times k\times l)\times(m-1)$
Location×Year *jk*	$(j-1)\times(k-1)$	Fixed	$(j\times k\times l)\times(m-1)$
Location×Planting System *il*	$(i-1)\times(l-1)$	Fixed	$(j\times k\times l)\times(m-1)$
Year *k*	$(k-1)$	Fixed	$(j\times k\times l)\times(m-1)$
Year×Planting System *kl*	$(k-1)\times(l-1)$	Fixed	$(j\times k\times l)\times(m-1)$
Location×Year×Planting System *jkl*	$(j-1)\times(k-1)\times(l-1)$	Fixed	$(j\times k\times l)\times(m-1)$
Planting System *l*	$(l-1)$	Fixed	$(j\times k\times l)\times(m-1)$
Replicate(Location×Year×Planting System) *(jkl)m*	$(j\times k\times l)\times(m-1)$	Random	

Model 2. Mixed effect model for combined analysis of variance for *i* genotypes, grown in *j* environments, during *k* years, with *l* replication.

Effect	Degrees of Freedom	Fixed or Random Effect	Denominator Degrees of Freedom for F Test
Genotype *i*	$(i-1)$	Fixed	$(j\times k)\times(i-1)\times(l-1)$
Genotype×Environment *ij*	$(i-1)\times(j-1)$	Fixed	$(j\times k)\times(i-1)\times(l-1)$
Genotype×Year *ik*	$(i-1)\times(k-1)$	Fixed	$(j\times k)\times(i-1)\times(l-1)$
Genotype×Environment×Year *ijk*	$(i-1)\times(j-1)\times(k-1)$	Fixed	$(j\times k)\times(i-1)\times(l-1)$
Genotype×Replicate(Environment×Year) *(jk) il*	$(j\times k)\times(i-1)\times(l-1)$	Random	
Environment *j*	$(j-1)$	Fixed	$(j\times k)\times(l-1)$
Environment×Year *jk*	$(j-1)\times(k-1)$	Fixed	$(j\times k)\times(l-1)$
Year *k*	$(k-1)$	Fixed	$(j\times k)\times(l-1)$
Replicate(Environment×Year) *(jk)l*	$(j\times k)\times(l-1)$	Random	

Results

General

Additional protocols for evaluating natto quality were tested but yielded imprecise or specious data. Based on literature we were interested in examining the possible correlation of water absorption with the structural carbohydrates cellulose and hemicellulose (Mullins and Xu, 2001). A neutral detergent fiber (NDF) and acid detergent fiber (ADF) assay was adapted from forage research and tested to determine its efficacy in soybean. Multiple laboratory trials were attempted but intra-sample variation between sub-samples was too high to accurately predict an observations true response Soybean has a high protein and oil content than forage and the chemistry of the assay was being affected in a way that decreased the assays accuracy and precision. A soybean extraction protocol to remove the oil and protein was implemented to alleviate the problem but consistent results still could not be achieved. The data were not reliable and the decision was made to drop the procedure from the research.

Another lab procedure called broken bean ratio (Wei and Chang, 2004) was attempted but did not yield meaningful data. It was designed to identify soybean seeds with cracked coats following soaking and steaming. The procedure involved soaking beans followed by steaming and sorting seeds with and without cracked seed coats. Defining seeds with cracked seed coats proved difficult and non-reproducible between replicates. This procedure was dropped because reliable data could not be generated.

The last method that could not be effectively implemented was a protocol to determine seed coat resistance to physical damage, modified from a method used to measure seed coat strength in snap bean (*Phaseolus vulgaris* L.) (Dickson and Boettger, 1977). The adapted protocol involved dropping a sample of soybeans from eight feet

onto a sloped steel plate to emulate physical stress encountered during harvest. The device was fabricated and trials conducted but the soybean lines tested produced no cracked seed so the protocol could not be used.

The 2006 growing season was more productive than the 2007 one at all locations. For a summary of weather conditions at the four locations during the 2006 and 2007 growing season refer to table's 1 and 2 in the appendix. The average monthly maximum temperatures between May and November for Blacksburg, Tidewater, Warsaw, and Mt. Holly in 2006 were similar to the 50 year average. The average monthly minimum temperatures between May and November for the locations were warmer than the 50 year average. The monthly precipitation averages for all locations were similar to the 50 year average. This resulted in growing conditions representative of the locations and years in which the cultivars had been developed.

For 2007, the average monthly maximum and minimum temperatures between May and November for Blacksburg, Tidewater, Warsaw, and Mt. Holly tended to be higher than the 50 year average. Virginia experienced drought conditions throughout the state during 2007, which resulted in all experimental locations receiving at least ten inches less rainfall than the 50 year average between May and November (data not shown). This resulted in the 2006 and 2007 growing seasons being quite different at all locations. The variation between the two growing seasons was another reason to analyze year as a fixed effect to allow for direct inference. The contrast in growing conditions was beneficial because it subjected the genotypes to environmental variation that would be expected over many years and allowed for a more accurate evaluation of the genotype's environmental sensitivity.

An objective of this research was to assess the effect of planting system on natto quality traits, and determine if there was significant genotype × planting system interaction. Model 1 was developed to analyze data grown in 2006 and 2007 at multiple locations using full season and double crop plantings. It was not possible to analyze the entire data set using model 1 because the experimental design was not balanced at Blacksburg and Tidewater. For technical reasons, it was not possible to plant a double crop test in Blacksburg so it could not be analyzed using model 1. Tidewater full season was adversely affected by the poor 2007 growing conditions and failed to produce soybean seed with acceptable natto quality and the entire Tidewater location could not be analyzed with model 1. Data collected at Mt. Holly and Warsaw from the 2006 and 2007 growing seasons under full season and double crop plantings were used to determine the effect of planting system on natto quality with model 1. Maturity, water absorption, water loss after steaming, seed coat deficiency, rate of water absorption, and seed size were analyzed. Maturity and seed size analysis used data from twelve cultivars, and all other responses were analyzed using eleven cultivars. Line V03-5794 was dropped from natto quality analysis because it is a conventional cultivar that consistently produced poor natto quality data. To avoid V03-5794 skewing the results of the experiment it was dropped from the natto quality analysis. Other conventional cultivars were included in the analysis because they had natto quality responses similar to the small seeded natto cultivars.

Model 2 was developed to address GEI on natto quality traits for soybeans grown in Virginia. Location and planting system were combined to form an environment variable used for analysis, i.e. the Blacksburg full season test became one environment and so on. Using this classification of environments, it was possible to use data from

Blacksburg full season, Tidewater double crop, Mt. Holly full season and double crop, and Warsaw full season and double crop from 2006 and 2007 as six environments for analysis with model 2. For reasons already mentioned, Tidewater full season could not be included in the analysis. Yield, water absorption, water loss after steaming, seed coat deficiency, rate of water absorption, and seed size were analyzed with model 2. Yield analysis used data from twelve cultivars, and all other responses were analyzed using eleven cultivars. Line V03-5794 was dropped from natto quality analysis for the same reasons presented above.

Year and location effects associated with model 1 were ignored in the data interpretation because data from only two locations were included. To get a better understanding of how location and year were affecting natto quality and agronomic traits it was more appropriate to use the environment and year factor in model 2 because it included the same data as model 1 plus data from additional locations. For the same reasons, yield was also not analyzed with model 1 because the same analysis was performed with model 2 and included more information.

i. MODEL ONE

All responses in this section were analyzed with model 1 to address how planting system affected trait expression during the 2006 and 2007 growing season. Maturity and seed size were analyzed using all twelve genotypes, while water absorption, water loss after steaming, seed coat deficiency, and rate of water absorption analysis included data from eleven genotypes. As mentioned, natto quality data was not collected at Tidewater full season in 2007 so the entire location was not included in model 1 analysis. Maturity data was collected at Tidewater full season in 2007 therefore, the Tidewater location was

included in maturity data analysis with model 1. V03-5794 was dropped from the analysis of natto quality traits for reasons mentioned. Residuals were analyzed for all responses to check for assumptions of combined analysis of variance: normal distribution and homogeneous variance. A Tukey's W value was computed (Ott and Longnecker, 2001) and used for multiple comparison means separation.

Maturity

Raw data residuals were non-normally distributed and six observations had positive studentized residual values greater than 3. The observation with the largest studentized residual value (5.29) was removed and the data set was reanalyzed. The second analysis of residuals was also non-normally distributed and the residual vs. location plot indicated homogenous variance for 10 of the 12 location, planting system, year combinations. The data were log and square root transformed to account for the non-normality and heterogeneous variance, but reanalysis of these data sets also failed to meet the assumptions of combined analysis of variance. Lomax (2001) states that the effects of violating the assumptions of normality and homogeneous variance on the outcome of ANOVA are minimal with equal or nearly equal sample size and the effects are alleviated as sample size increases. Using this rationale, the raw data set was used for further analysis.

The four main factors genotype, location, year, and planting system were significant sources of variation with p-values <0.0001 (Table 4). All two and three way interactions were also significant except the year × planting system interaction. For natto genotypes, cultivar MFS-591 had the latest mean maturity date of 54 days after August 31st and line V00-3488 had the earliest mean maturity date at 50 days after August 31st

(Table 5). Cultivars planted double crop matured 56 days after August 31st, which was significantly later than the maturity date for cultivars planted full season, 48 days after August 31st (Table 5). Cultivars grown in 2006 had a mean maturity date of 59 days after August 31st, which was significantly later than the mean maturity of 45 days after August 31st for the same cultivars planted in 2007 (Table 5).

Table 4. ANOVA for maturity listing the significance of main effects and interaction terms for 12 soybean genotypes grown in three locations in Virginia with two planting systems during the 2006 and 2007 growing season.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	11	319	16.10	<.0001
Genotype×Location	22	319	2.62	0.0001
Genotype×Year	11	319	3.02	0.0008
Genotype×Planting System	11	319	3.22	0.0004
Genotype×Location×Year	22	319	5.17	<.0001
Location	2	24	91.02	<.0001
Location×Year	2	24	127.57	<.0001
Location×Planting System	2	24	83.28	<.0001
Year	1	24	1556.14	<.0001
Year×Planting System	1	24	3.66	0.0679
Location×Year×Planting System	2	24	71.98	<.0001
Planting System	1	24	578.07	<.0001

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

Table 5. Maturity (days after August 31st) lsmean estimates for 12 soybean genotypes grown at three locations in Virginia using two planting systems during the 2006 and 2007 growing season.

Source	LSMean	N‡	Standard Error
Genotype			
Natto			
V00-3488	51	36	0.39
V00-3493	51	36	0.39
V00-3636	51	36	0.39
Camp	51	36	0.39
V01-4937	51	36	0.39
MFS-511	51	36	0.39
MFS-553	52	36	0.39
MFS-591	54	36	0.39
Conventional			
SS-516	53	36	0.39
Teejay	54	36	0.39
Hutcheson	54	36	0.39
V03-5794	55	36	0.39
Tukey 0.05†	2		
Location			
Mt. Holly	55	72	0.30
Tidewater	52	72	0.30
Warsaw	50	72	0.30
Planting System			
Double Crop	57	216	0.24
Full Season	48	216	0.24
Year			
2006	59	216	0.24
2007	46	216	0.24

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

Water Absorption

The normal probability plot of residuals was normally distributed, all observations had studentized residual values less then 3, and the residual vs. predicted water

absorption plot displayed no patterns and was randomly distributed. Using these criteria, the data set met the assumptions for analysis of variance without modification and was used for analysis.

High water absorption values are desirable for natto cultivars. The three main factors genotype, location, and planting system were significant at the 0.0001 probability level, while year was non-significant for water absorption (Table 6). The two way interactions between genotype and the other three factors were significant, as was the location × year and the year × planting system interactions. For natto genotypes, Cultivar MFS-553 had the lowest mean water absorption value of 214.81% and line V00-3488 had the highest at 224.72% (Table 7). Line V00-3488 and V00-3493 absorbed significantly more water than the other natto cultivars. The mean water absorption for cultivars planted double crop was 1.89% higher than those planted full season, which was statistically significant (Table 6 and 7). The two way interaction genotype × planting system was significant at the 0.01 probability level (Table 6). The interaction plot for cultivars planted double crop and full season shows the rank change that caused the statistically significant interaction (Figure 1). Cultivar MFS-591 had a water absorption value of 221.63% when grown double crop compared to 217.52% when grown full season.

Table 6. ANOVA for water absorption listing the significance of main effects and interaction terms for 11 soybean genotypes grown in two locations in Virginia with two planting systems during the 2006 and 2007 growing season.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	10	190	140.83	<.0001
Genotype×Location	10	190	4.01	<.0001
Genotype×Year	10	190	12.35	<.0001
Genotype×Planting System	10	190	2.71	0.004
Genotype×Location×Year	10	190	1.41	0.1775
Location	1	16	284.65	<.0001
Location×Year	1	16	18.06	0.0006
Location×Planting System	1	16	0.05	0.8258
Year	1	16	0.37	0.5499
Year×Planting System	1	16	9.62	0.0069
Location×Year×Planting System	1	16	0.36	0.5569
Planting System	1	16	40.77	<.0001

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

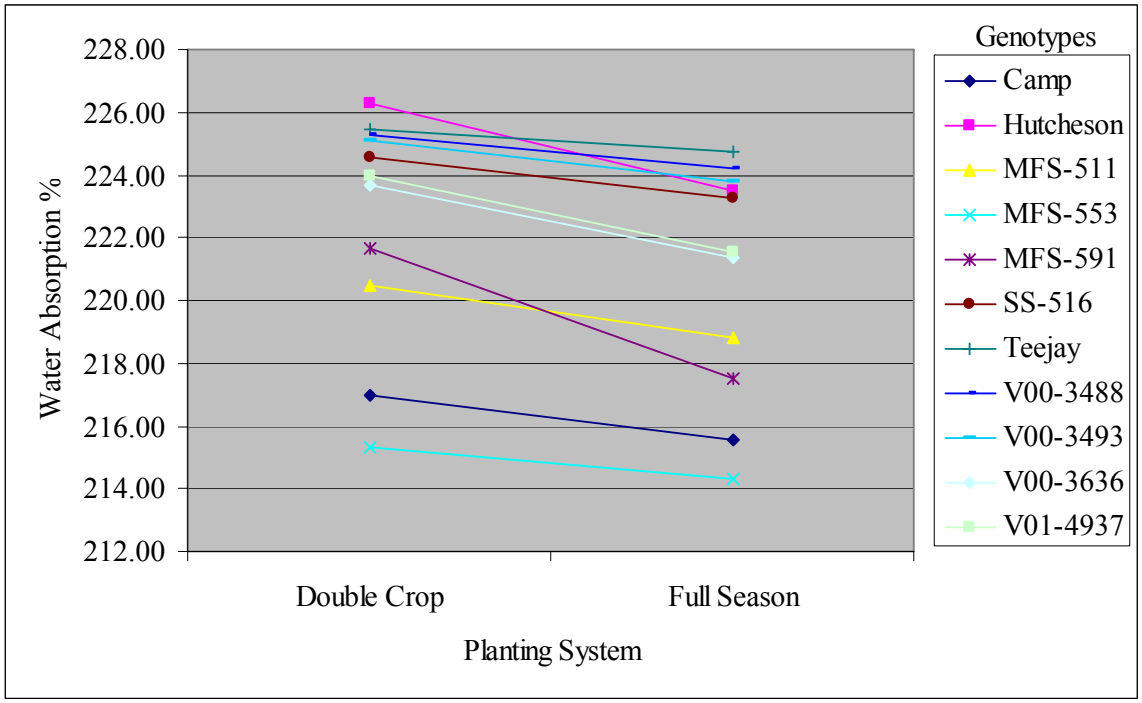
Table 7. Water absorption (%) lsmean estimates for 11 soybean genotypes grown at two locations in Virginia using two planting systems during the 2006 and 2007 growing season.

Source	LSMean	N‡	Standard Error
Genotypes			
Natto			
V00-3488	224.72	24	0.32
V00-3493	224.45	24	0.32
V01-4937	222.76	24	0.32
V00-3636	222.53	24	0.32
MFS-511	219.64	24	0.32
MFS-591	219.57	24	0.32
Camp	216.27	24	0.32
MFS-553	214.81	24	0.32
Conventional			
Teejay	225.09	24	0.32
Hutcheson	224.88	24	0.32
SS-516	223.93	24	0.32
Tukey 0.05†	1.6		
Location			
Mt.Holly	219.29	132	0.20
Warsaw	224.10	132	0.20
Planting System			
Double Crop	222.61	132	0.20
Full Season	220.79	132	0.20
Year			
2006	221.61	132	0.20
2007	221.78	132	0.20

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

Figure 1. Genotype by planting system interaction plot showing lsmean water absorption for 11 soybean genotypes plotted against two planting systems averaged over two locations during 2006 and 2007.



Water Loss after Steaming

The normal probability plot of residuals was normally distributed and one observation had a studentized residual value greater then 3. This observation was dropped and the residuals were re-run. The second set of residuals were normally distributed, no observations had studentized residuals over 3, and the residual vs. predicted water loss after steaming plot showed no trends and a homogeneous distribution of variance. The data met the assumptions for analysis of variance and was used for further analysis.

Low water loss after steaming values are desirable for natto cultivars. The main factors genotype and year were significant at the 0.0001 probability level, location was significant at the 0.01 probability level, and planting system was significant at the 0.05 probability level (Table 8). The following two way interactions were significant:

genotype × location, genotype × year, and location × year. The three way interaction location × year × planting system was also significant. For natto genotypes, cultivar MFS-553 lost the most water after steaming 5.81% and line V00-3488 lost the least 5.26% (Table 9). Planting system was a significant source of variation and cultivars planted double crop lost 1.26% more water after steaming than the same cultivars planted full season. Year was a significant factor for the response and cultivars grown in 2007 lost 1.30% more water than the same cultivars grown in 2006 (Tables 8 and 9).

Table 8. ANOVA for water loss after steaming listing the significance of main effects and interaction terms for 11 soybean genotypes grown in two locations in Virginia with two planting systems during 2006 and 2007.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	10	189	10.89	<.0001
Genotype×Location	10	189	4.16	<.0001
Genotype×Year	10	189	1.94	0.0419
Genotype×Planting System	10	189	1.36	0.2002
Genotype×Location×Year	10	189	1.66	0.0928
Location	1	16	10.01	0.0060
Location×Year	1	16	11.09	0.0042
Location×Planting System	1	16	2.26	0.1524
Year	1	16	125.60	<.0001
Year×Planting System	1	16	0.16	0.6960
Location×Year×Planting System	1	16	15.92	0.0011
Planting System	1	16	6.20	0.0241

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

Table 9. Water loss after steaming (%) lsmean estimates for 11 soybean genotypes grown at two locations in Virginia using two planting systems during the 2006 and 2007 growing season.

Source	LSMean	N‡	Standard Error
Genotype			
Natto			
V00-3488	5.26	24	0.09
MFS-591	5.28	24	0.09
MFS-511	5.36	24	0.09
V00-3493	5.37	24	0.09
V00-3636	5.43	24	0.09
V01-4937	5.58	24	0.09
Camp	5.70	24	0.09
MFS-553	5.81	24	0.09
Conventional			
SS-516	5.48	23	0.09
Teejay	5.73	24	0.09
Hutcheson	5.97	24	0.09
Tukey 0.05†	0.37		
Location			
Mt.Holly	5.73	132	0.08
Warsaw	5.36	131	0.08
Planting System			
Double Crop	5.69	132	0.08
Full Season	5.40	131	0.08
Year			
2006	4.89	132	0.08
2007	6.19	131	0.08

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

Seed Coat Deficiency

The residuals were non-normally distributed and two observations had negative studentized residual values greater than 3. These observations were removed and the data were re-analyzed. The second analysis had normally distributed residuals, one observation with a studentized residual greater than 3, and the residual vs. predicted seed

coat deficiency plot showed no trends and indicated homogeneous variance. The ANOVA tables produced with the whole data set and the data set with the two observations removed had similar conclusions. Therefore, the second data set was used for analysis because it better meets the assumptions and no significant change was made to the conclusions.

Seed coat deficiency count is the number of seeds identified as deficient out of a 100 seed sample. Lower seed coat deficiency counts are desirable for natto cultivars and improve finished natto quality. The four main factors genotype, location, year, and planting system were all significant at the 0.001 probability level the two way interactions between genotype and the three other factors were all significant (Table 10). The two way interactions location × year and year × planting system were also significant. For natto genotypes, cultivar MFS-553 had the lowest mean seed coat deficiency count of 17, while line V00-3493 had the highest mean seed coat deficiency count of 72, higher than any cultivar included in the analysis (Table 11). Cultivars grown in a double crop system had fewer deficient seeds than those grown full season, 43 and 48, respectively. Cultivars grown in 2006 had 9 fewer deficient seeds than the same cultivars grown in 2007 based on means (Table 11). The interaction plot for cultivars planted double crop and full season shows the rank change that caused the statistically significant interaction (Figure 2). Notably, natto cultivars MFS-511 and MFS-591 did not have fewer deficient seeds when planted double crop vs. full season. Conventional cultivars Teejay and Hutcheson showed the largest decline in deficient seeds when planted double crop vs. full season, 19 and 14 fewer deficient seeds respectively.

Table 10. ANOVA for seed coat deficiency listing the significance of main effects and interaction terms for 11 soybean genotypes grown in two locations in Virginia with two planting systems during 2006 and 2007.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	10	188	164.45	<.0001
Genotype×Location	10	188	2.25	0.0167
Genotype×Year	10	188	10.13	<.0001
Genotype×Planting System	10	188	4.44	<.0001
Genotype×Location×Year	10	188	1.64	0.0989
Location	1	16	16.70	0.0009
Location×Year	1	16	7.51	0.0145
Location×Planting System	1	16	0.17	0.6834
Year	1	16	107.56	<.0001
Year×Planting System	1	16	10.85	0.0046
Location×Year×Planting System	1	16	0.35	0.5636
Planting System	1	16	38.78	<.0001

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

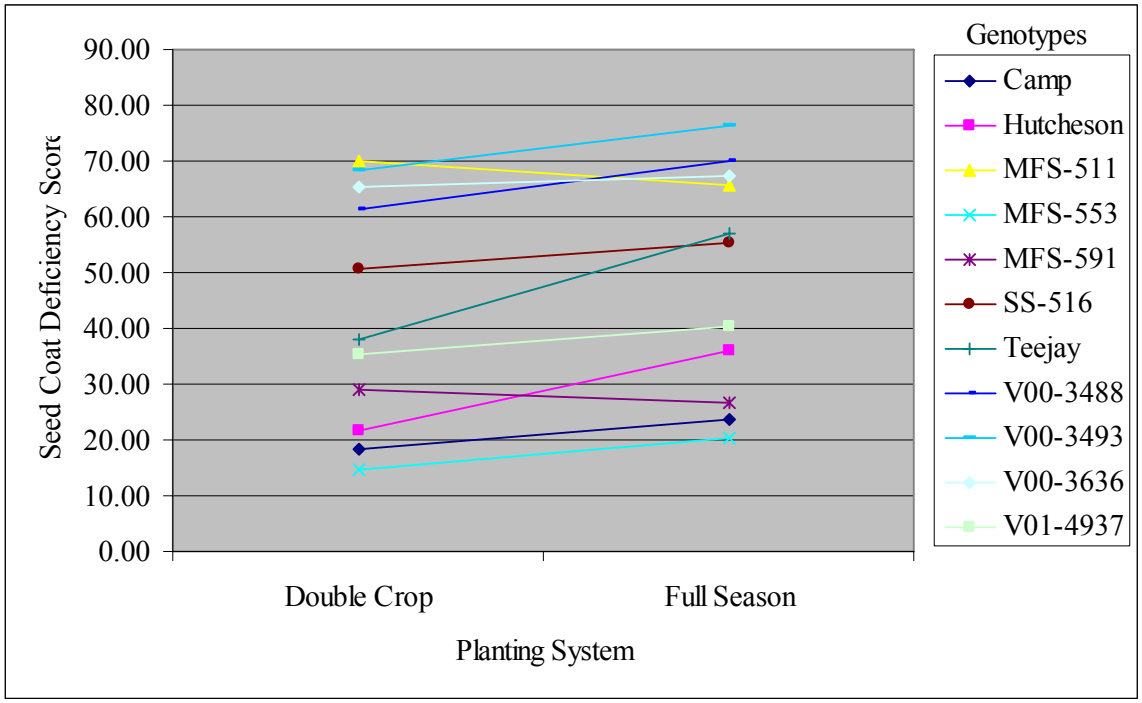
Table 11. Seed coat deficiency (%) lsmean estimates for 11 soybean genotypes grown at two locations in Virginia using two planting systems during the 2006 and 2007 growing season.

Source	LSMean	N‡	Standard Error
Genotype			
Natto			
MFS-553	18	24	1.58
Camp	21	24	1.58
MFS-591	28	24	1.58
V01-4937	38	24	1.58
V00-3488	66	23	1.62
V00-3636	66	24	1.58
MFS-511	68	23	1.62
V00-3493	72	24	1.58
Conventional			
Hutcheson	29	24	1.58
Teejay	47	24	1.58
SS-516	53	24	1.58
Tukey 0.05†	8		
Location			
Mt.Holly	48	130	0.68
Warsaw	44	132	0.67
Planting System			
Double Crop	43	132	0.67
Full Season	49	130	0.68
Year			
2006	41	130	0.68
2007	51	132	0.67

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

Figure 2. Genotype by planting system interaction plot showing lsmean seed coat deficiency counts for 11 soybean genotypes plotted against two planting systems averaged over two locations during 2006 and 2007.



Rate of Water Absorption

The residuals were non-normally distributed and five observations had studentized residual values greater than 3. All five of these observations were positive and were entries at four different locations. The residual at environment plot showed that variance was homogeneous across locations but the variance was non-symmetric with no negative studentized residual observations with values greater than 3. To avoid dropping five observations, the data were log transformed and reanalyzed. The log transformed residuals were normally distributed at the 0.001 probability level and one observation had a negative studentized residual value greater than 3. The two plots used to check for homogeneous variance indicated the transformed data met the assumptions for equal variance better then the untransformed data. ANOVA's were computed using both the untransformed and the transformed data. The results for significant effects were

identical; each ANOVA identified the same six factors as non-significant at the 0.05 probability level. For this reason, the untransformed data were used for subsequent analysis. Two points should be noted: one, departure from homogeneous variance is thought to produce small effects when computing analysis of variance if the experiment is balanced or nearly balanced and the effect decreases as sample size increases (Lomax, 2001). Two, violation of the normal distribution assumption is known to be minimal when experimental design is balanced or nearly balanced (Lomax, 2001). These two points illustrate that although the untransformed data do not perfectly meet the assumptions for analysis of variance the balanced design and large sample size should offset any effects.

Rate of water absorption is a rate constant estimating the initial slope of non-linear water absorption with units min^{-1}. Cober et al. (2006) suggests that higher water absorption (b) is a desirable natto quality trait. The main effects genotype, year, and planting system were significant at the 0.01 probability level and location was non-significant (Table 12). The two-way interactions genotype × year and genotype × planting system were significant along with the three way interaction genotype × location × year. For natto genotypes, line V00-3493 had the highest lsmean rate of water absorption value 7.53×10^{-3} min^{-1} and cultivar MFS-553 had the lowest 5.00×10^{-3} min^{-1} (Table 13). The three conventional genotypes had rate of water absorption estimates between V00-3493 and MFS-553. The mean rate of water absorption for cultivars grown full season 6.62×10^{-3} min^{-1} was significantly higher than for the same cultivars grown double crop 6.14×10^{-3} min^{-1} (Table 12 and 13). The mean rate of water absorption for cultivars grown in 2006 was 6.10×10^{-3} min^{-1} and significantly less than the same cultivars grown in 2007 at 6.65×10^{-3} min^{-1} (Table 12 and 13). The interaction plot for cultivars

planted double crop and full season shows the rank change that caused the statistically significant interaction (Figure 3). For natto genotypes, lines V00-3488, V00-3493, and V00-3636 displayed higher rates of water absorption when planted full season vs. double crop while cultivars MFS-511, MFS-533, and MFS-591 displayed little response to the different planting systems for rate of water absorption estimates (Figure 3). This differential cultivar response was also seen between the conventional cultivars.

Table 12. ANOVA for rate of water absorption listing the significance of main effects and interaction terms for 11 soybean genotypes grown in two locations in Virginia with two planting systems during 2006 and 2007.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	10	190	56.57	<.0001
Genotype×Location	10	190	1.11	0.3531
Genotype×Year	10	190	5.92	<.0001
Genotype×Planting System	10	190	2.41	0.0102
Genotype×Location×Year	10	190	3.77	0.0001
Location	1	16	0.11	0.7477
Location×Year	1	16	1.35	0.2615
Location×Planting System	1	16	0.67	0.4256
Year	1	16	15.70	0.0011
Year×Planting System	1	16	1.59	0.2250
Location×Year×Planting System	1	16	0.14	0.7141
Planting System	1	16	11.95	0.0032

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

Table 13. Rate of water absorption (min^{-1}) lsmean estimates for 11 soybean genotypes grown at two locations in Virginia using two planting systems during the 2006 and 2007 growing season.

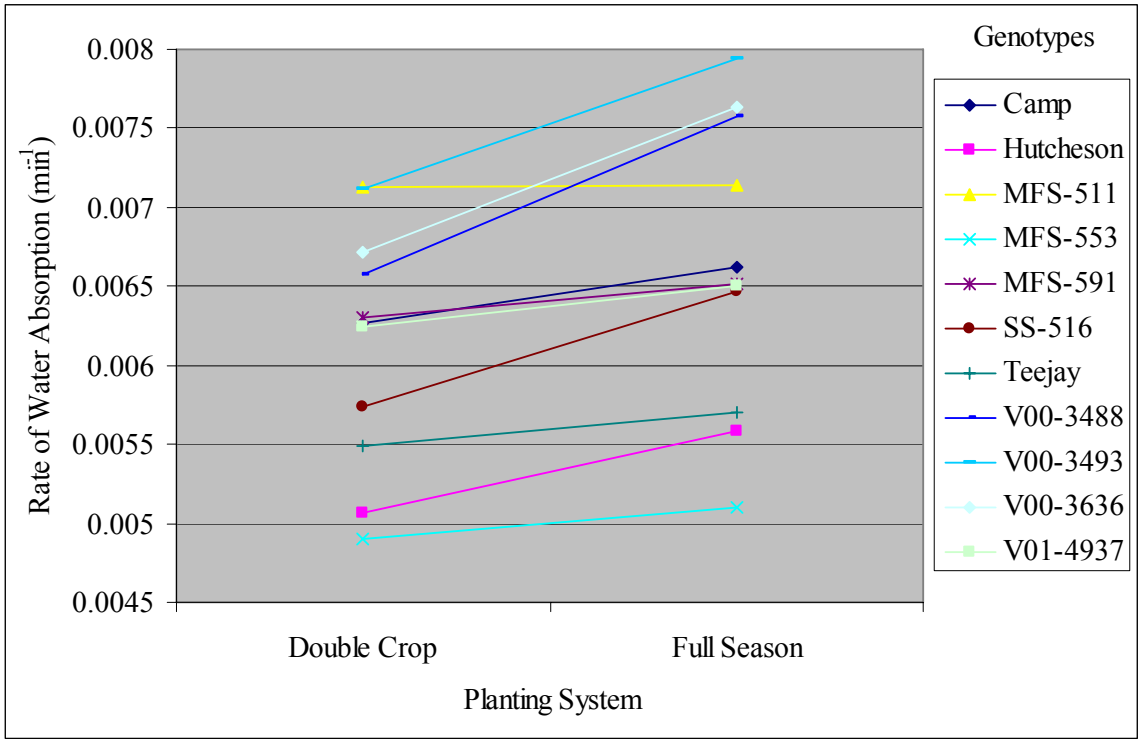
Source	LSMean	N‡	Standard Error
Genotype	$\times 10^{-3}$§		$\times 10^{-3}$§
Natto			
V00-3493	7.53	24	0.10
V00-3636	7.17	24	0.10
MFS-511	7.13	24	0.10
V00-3488	7.07	24	0.10
Camp	6.45	24	0.10
MFS-591	6.41	24	0.10
V01-4937	6.37	24	0.10
MFS-553	5.00	24	0.10
Conventional			
SS-516	6.10	24	0.10
Teejay	5.59	24	0.10
Hutcheson	5.32	24	0.10
Tukey 0.05†	0.57		
Location			
Mt.Holly	6.40	132	0.10
Warsaw	6.36	132	0.10
Planting System			
Double Crop	6.14	132	0.10
Full Season	6.62	132	0.10
Year			
2006	6.10	132	0.10
2007	6.65	132	0.10

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

§: Actual rate value equals reported value multiplied by listed factor

Figure 3. Genotype by planting system interaction plot showing lsmean rate of water absorption for 11 soybean genotypes plotted against two planting systems averaged over two locations during 2006 and 2007.



Seed Size

The residuals were normally distributed and two observations had studentized residual values greater than 3. One of these values was positive and the other negative, creating symmetry in the variance not resulting in a violation of homogeneous variance. The plot of residuals vs. predicted values for seed size did not indicate homogeneous variance however. The plot of residuals at environments showed unequal variance across growing environments. For these reasons the data were log transformed and reanalyzed. The log transformed data were normally distributed, one observation had a studentized residual value greater than 3, the residual vs. predicted value plot showed equal variance for the data set, and the residual at environment graph showed a more equal distribution of variance across locations. ANOVAs were computed for both the untransformed and

transformed data and the resulting tests of significant were not similar. Specifically, two factors were significant at the 0.001 probability level using the untransformed data and non-significant at the 0.05 probability level using the log transformed data. Therefore, the log transformed data were used for further analysis because it met the assumption for analysis of variance better than the untransformed data.

The four main factors genotype, location, year, and planting system were significant at the 0.0001 probability level (Table 14). The two-way interactions between genotype and the other three factors were significant along with the three way interaction genotype × location × year. The means for cultivar, location, year, and planting system are presented as both transformed and untransformed values (Table 15). Untransformed means were included because transformed values are unfamiliar and difficult to interpret, and the Pearson correlation coefficient between untransformed and transformed seed size was r=0.99525 with a p-value <0.0001. Natto cultivar Camp had the smallest seed size estimate of 7.89 g per 100 seed and line V00-3488 had the largest seed size estimate of 11.17 g per 100 seed for natto cultivars tested. Teejay had the highest seed size estimate of 17.28 g per 100 seeds for the conventional cultivars. Cultivars grown in Mt. Holly had a significantly larger mean seed size of 12.45 g per 100 seed than the same cultivars grown in Warsaw 10.94 g per 100 seed (Table 14 and 15). Cultivars grown double crop had a significantly smaller seed size 11.44 g per 100 seed than those grown full season 11.95 g per 100 seed (Table 14 and 15). Cultivars grown during 2006 had a significantly smaller seed size 11.31 g per 100 seed than those grown during 2007 with 12.07 g per 100 seed (Table 14 and 15). The interaction plot for cultivars planted double crop and full season shows the rank change that caused the statistically significant interaction (Figure 4). A specific grouping can be seen for the large and small seeded cultivars and

no interaction takes place between the groups. Figure 4 does not indicate any significant cultivar rank changes or drastic departures from seed size performance across planting system.

Table 14. ANOVA for log transformed seed size listing the significance of main effects and interaction terms for 12 soybean genotypes grown in two locations in Virginia with two planting systems during 2006 and 2007.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	11	208	726.47	<.0001
Genotype×Location	11	208	2.41	0.0076
Genotype×Year	11	208	2.78	0.0022
Genotype×Planting System	11	208	2.57	0.0044
Genotype×Location×Year	11	208	4.18	<.0001
Location	1	16	301.38	<.0001
Location×Year	1	16	2.96	0.1047
Location×Planting System	1	16	3.91	0.0655
Year	1	16	77.63	<.0001
Year×Planting System	1	16	4.10	0.0600
Location×Year×Planting System	1	16	0.01	0.9296
Planting System	1	16	28.81	<.0001

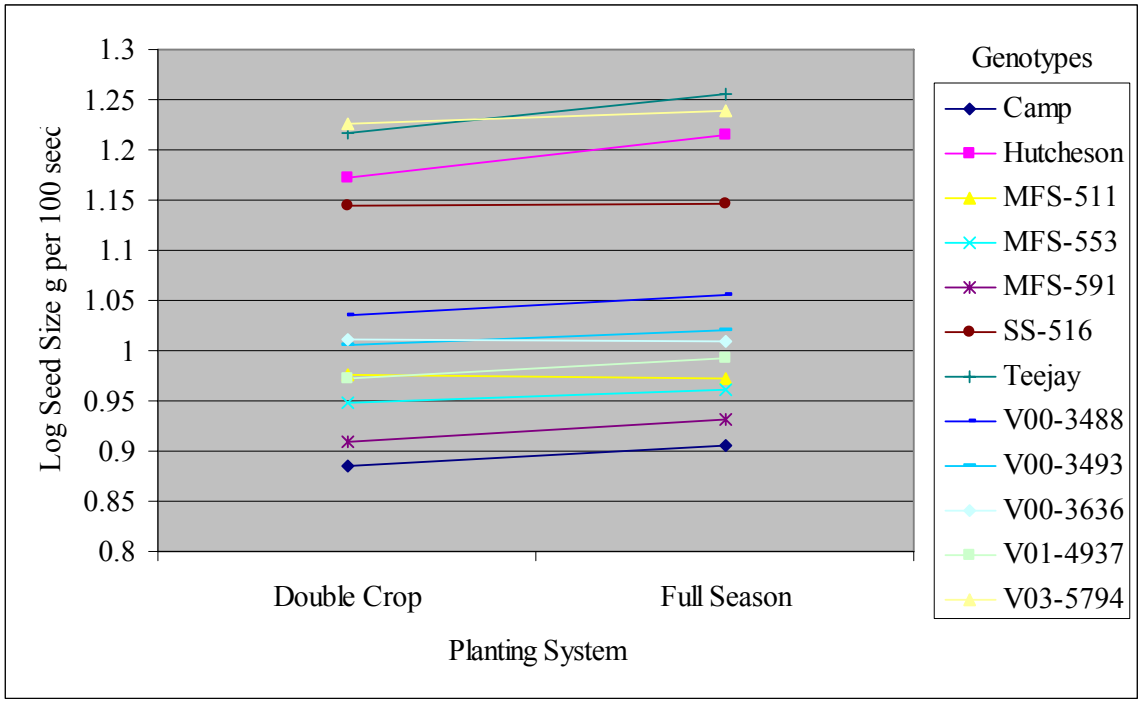
†: Numerator degrees of freedom
‡: Denominator degrees of freedom

Table 15. Seed size (g per 100 seed) lsmean estimates for 12 soybean genotypes grown at two locations in Virginia using two planting systems during the 2006 and 2007 growing season.

Effect	LSMean§	Standard Error§	LSMean¶	Standard Error¶	N‡
Genotype					
Natto					
Camp	0.89	0.0045	7.89	0.1238	36
MFS-591	0.92	0.0045	8.35	0.1238	36
MFS-553	0.96	0.0045	9.05	0.1238	36
MFS-511	0.97	0.0045	9.43	0.1238	36
V01-4937	0.98	0.0045	9.64	0.1238	36
V00-3636	1.01	0.0045	10.27	0.1238	36
V00-3493	1.01	0.0045	10.38	0.1238	36
V00-3488	1.05	0.0045	11.17	0.1238	36
Conventional					
SS-516	1.15	0.0045	14.04	0.1238	36
Hutcheson	1.19	0.0045	15.69	0.1238	36
V03-5794	1.23	0.0047	17.11	0.1270	36
Teejay	1.24	0.0045	17.28	0.1238	36
Tukey 0.05†	0.02		0.65		
Location					
Mt. Holly	1.08	0.0022	12.45	0.0585	72
Warsaw	1.02	0.0022	10.94	0.0583	72
Planting System					
Double Crop	1.04	0.0022	11.44	0.0583	216
Full Season	1.06	0.0022	11.95	0.0585	216
Year					
2006	1.04	0.0022	11.31	0.0583	216
2007	1.06	0.0022	12.07	0.0585	216

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes
‡: The number of observations used to calculate the least square mean
§: LSMean and standard error computed using log transformed seed size data
¶: LSMean and standard error computed using untransformed seed size data

Figure 4. Genotype by planting system interaction plot showing log lsmean seed size for 12 soybean genotypes plotted against two planting systems averaged over two locations during 2006 and 2007.



ii. **MODEL TWO**

All responses in this section were analyzed with model 2 to address how environment affected trait expression during the 2006 and 2007 growing season. The reasons to use model 2 for general environment and year analysis were provided in previous sections. Yield and seed size were analyzed using all twelve genotypes, while water absorption, water loss after steaming, seed coat deficiency, and rate of water absorption analysis included data from eleven genotypes. V03-5794 was dropped from the analysis of natto quality traits for reasons mentioned. To check for homogeneous variance and normal distribution, residuals for the combined analysis of variance were analyzed for all responses.

Yield

The raw data had non-normally distributed residuals and four observations with studentized residual values greater then 3, three were positive and one was negative. The two highest studentized residual values were positive. These two observations were removed and the data set and the model was re-run. The second analysis had normally distributed residuals at the 0.01 probability level, and the residual vs. predicted yield plot showed equal distribution of variation with no patterns. The ANOVA output from the two different data sets had identical results for test of significance. This second data set met the assumptions for analysis of variance and was used for further analysis.

The three main factors genotype, environment, and year, the two-way interactions genotype × environment and environment × year, and the three way interaction genotype × environment × year were significant at the 0.0001 probability level (Table 16). For natto genotypes, cultivar MFS-591 had the lowest mean yield estimate of 2358 kg/hectare

and line V01-4937 had the highest 3047 kg/hectare (Table 17). For the six environments tested, Tidewater double crop had the lowest mean yield and Blacksburg full season had the highest, 1912 kg/hectare and 3941 kg/hectare, respectively (Table 17). Year also significantly affected the yield response and cultivars planted during 2006 yielded 590 kg/hectare more than the same cultivars grown in 2007. The interaction plot for cultivars planted double crop and full season shows the rank change that caused the statistically significant interaction (Figure 5). Most rank changes resulting from cross-over did not significantly change the grouping of high and low performing cultivars. Exceptions to this were the performance of conventional cultivar Teejay at Mt. Holly full season and natto cultivar MFS-591 at Warsaw full season which represented deviations from high and low groupings. Also, the yield response of natto line V00-3636 created many cross over events that resulted in confounding rank change (Figure 5).

Table 16. ANOVA for yield listing significance of main effects and interaction terms for 12 soybean genotypes grown in six environments in Virginia during the 2006 and 2007 growing season.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	11	262	38.75	<.0001
Genotype×Environment	55	262	2.8	<.0001
Genotype×Year	11	262	1.27	0.2448
Genotype×Environment×Year	55	262	2.2	<.0001
Environment	5	24	260.91	<.0001
Environment×Year	5	24	85.65	<.0001
Year	1	24	178.9	<.0001

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

Table 17. Yield (kg/hectare) lsmean estimates for 12 soybean genotypes grown in six environments in Virginia during the 2006 and 2007 growing season.

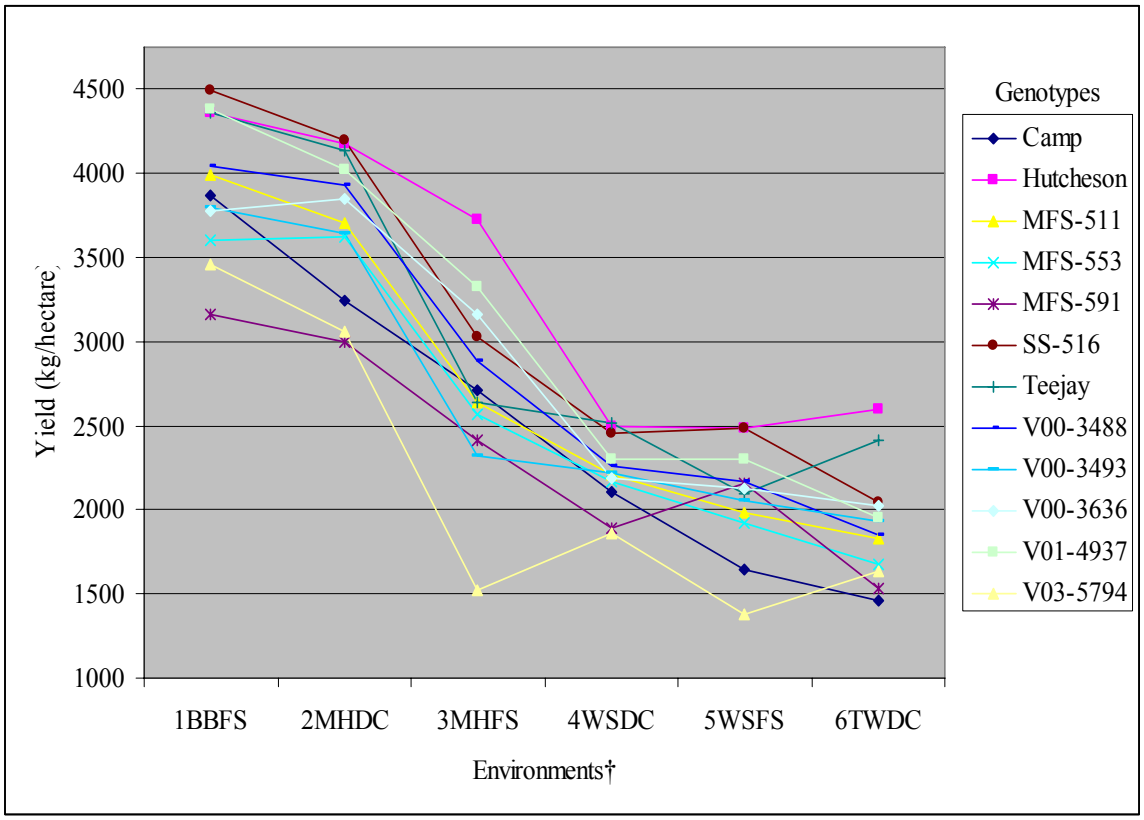
Source	LSMean	N‡	Standard Error
Genotype			
Natto			
V01-4937	3047.52	36	55.85
V00-3488	2856.06	36	55.85
V00-3636	2855.57	36	55.85
MFS-511	2727.31	36	55.85
V00-3493	2661.38	36	55.85
MFS-553	2595.11	36	55.85
Camp	2505.86	36	55.85
MFS-591	2358.36	35	56.96
Conventional			
Hutcheson	3308.54	36	55.85
SS-516	3116.67	36	55.85
Teejay	3028.16	35	56.96
V03-5794	2151.36	36	55.85
Tukey 0.05†	350.22	36	
Environment§			
BBFS	3941.70	72	53.95
MHDC	3715.66	72	53.95
MHFS	2746.20	70	54.52
WSDC	2222.63	72	53.95
WSFS	2066.90	72	53.95
TWDC	1912.87	72	53.95
Year			
2006	3062.74	216	31.15
2007	2472.58	214	31.26

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

§: BBFS, Blacksburg full season; MHDC, Mt. Holly double crop; MHFS, Mt. Holly full season; WSDC, Warsaw double crop; WSFS, Warsaw full season; TWDC, Tidewater double crop.

Figure 5. Genotype by environment interaction plot showing lsmean yield for 12 soybean genotypes plotted against 6 growing environments averaged across 2006 and 2007 growing season.



†: BBFS, Blacksburg full season; MHDC, Mt. Holly double crop; MHFS, Mt. Holly full season; TWDC, Tidewater double crop; WSDC, Warsaw double crop; WSFS, Warsaw full season.

Water Absorption

Residuals for water absorption were normally distributed and one observation had a studentized residual value over 3. This observation was dropped and assumptions rechecked. The second analysis had normally distributed residuals, no observations with studentized residuals over 3, and the residual vs. predicted water absorption plot showed no trends. The ANOVA output from the two different data sets had identical results for test of significance. The second data set was used for further analysis because it better met the assumptions for analysis of variance.

High water absorption values are desirable for natto cultivars. The main factors genotype and environment were significant at the 0.0001 probability level and year was non-significant (Table 18). All two and three way interaction terms were significant at the 0.0001 probability level. The cultivar mean estimates for water absorption were similar to those calculated using model 1 (Tables 7 and 19). For natto genotypes, cultivar MFS-553 had the lowest mean water absorption value of 214.50% and V00-3488 had the highest at 224.55% (Table 19). Cultivars grown at Mt. Holly full season had the lowest mean water absorption 218.41%, while the same cultivars grown at Warsaw double crop had the highest mean water absorption 225.04%. The genotype × environment interaction plot displays genotype rank change that caused the statistically significant interaction (Figure 6). Performance can be grouped in two distinct categories but more detailed separation is difficult. Conventional cultivars, Hutcheson and Teejay, are the top performing cultivars at three locations but their water absorption responses decrease across environments (Figure 6). Cultivar MFS-591 had similar water absorption values at most locations except Warsaw double crop.

Table 18. ANOVA for water absorption listing the significance of main effects and interaction terms for 11 soybean genotypes grown in six environment in Virginia during the 2006 and 2007 growing season.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	10	239	250.85	<.0001
Genotype×Environment	50	239	5.89	<.0001
Genotype×Year	10	239	11.02	<.0001
Genotype×Environment×Year	50	239	3.47	<.0001
Environment	5	24	81.32	<.0001
Environment×Year	5	24	37.19	<.0001
Year	1	24	1.91	0.1800

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

Table 19. Water absorption (%) lsmean estimates for 11 soybean genotypes grown in six environments in Virginia during the 2006 and 2007 growing season.

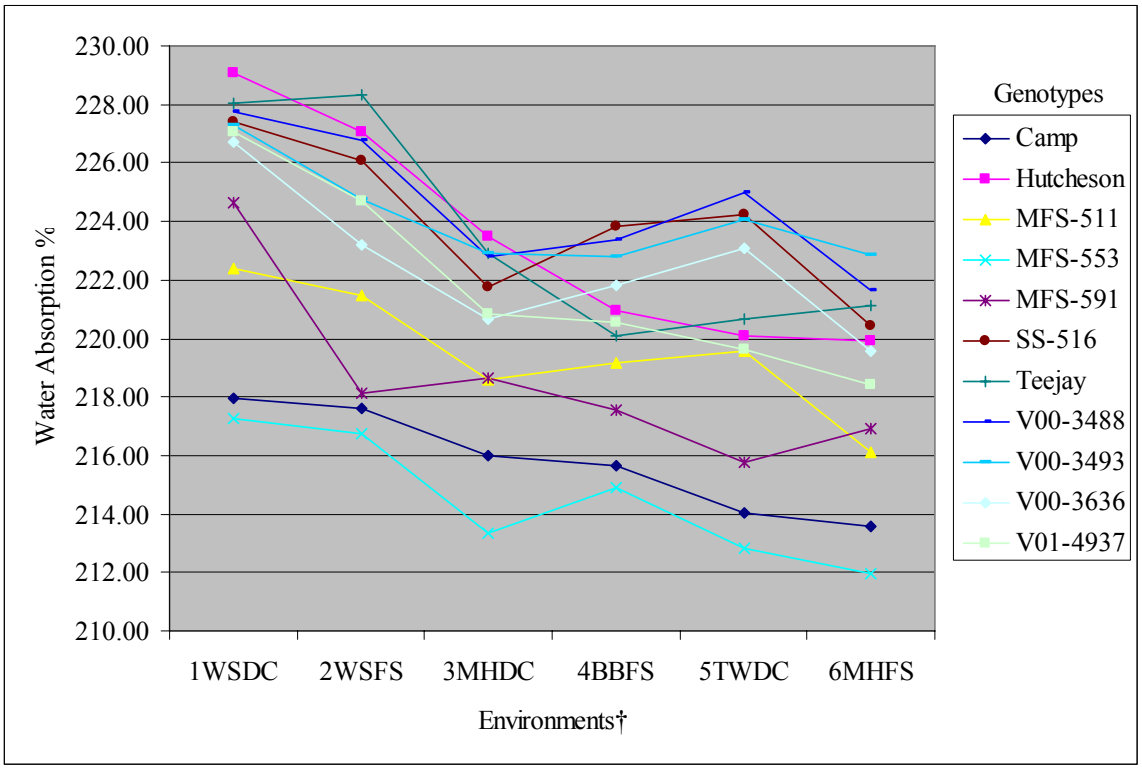
Source	LSMean	N‡	Standard Error
Genotype			
Natto			
V00-3488	224.55	36	0.2377
V00-3493	224.12	36	0.2377
V00-3636	222.51	36	0.2377
V01-4937	221.87	36	0.2377
MFS-511	219.55	36	0.2377
MFS-591	218.60	36	0.2377
Camp	215.79	36	0.2377
MFS-553	214.50	35	0.2421
Conventional			
SS-516	223.97	36	0.2377
Teejay	223.52	36	0.2377
Hutcheson	223.42	36	0.2377
Tukey 0.05†	1.42		
Environment§			
WSDC	225.04	72	0.2732
WSFS	223.16	72	0.2732
MHDC	220.17	72	0.2732
BBFS	220.07	72	0.2732
TWDC	219.91	71	0.2744
MHFS	218.41	72	0.2732
Year			
2006	220.97	216	0.1577
2007	221.28	215	0.1580

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

§: WSDC, Warsaw double crop; WSFS, Warsaw full season; MHDC, Mt. Holly double crop; BBFS, Blacksburg full season; MHFS, Mt. Holly full season; TWDC, Tidewater double crop.

Figure 6. Genotype by environment interaction plot showing lsmean water absorption for 11 soybean genotypes plotted against 6 growing environments averaged across 2006 and 2007.



†: WSDC, Warsaw double crop; WSFS, Warsaw full season; MHDC, Mt. Holly double crop; BBFS, Blacksburg full season; TWDC, Tidewater double crop; MHFS, Mt. Holly full season.

Water Loss After Steaming

The residuals for water loss after steaming were normally distributed at the 0.0001 probability level, and 6 observations had studentized residual values greater than 3. The distribution of the data points with studentized residuals greater than three was symmetric, however, with three positive and three negative values. The residual vs. predicted and residual at environment plots indicated homogeneous variance for the raw data set. For these reasons, no data points were removed for further analysis. This data set met the criteria for analysis of variance and was used for the further analysis.

Low water loss after steaming values are desirable for natto cultivars. The main factors genotype and year were significant at the 0.0001 probability level and environment was significant at the 0.01 probability level (Table 20). The two way

interactions genotype × environment and environment × year were significant along with the three way interaction. The mean cultivar water loss after steaming values computed using model 2 were similar to those computed with model 1. For natto genotypes, line V00-3488 lost the least amount of water after steaming, 5.14%, while cultivar MFS-553 lost the most water after steaming, 5.68% (Table 21). Genotypes grown at Mt. Holly double crop lost the most water after steaming 5.78% and those grown at Blacksburg full season lost the least 5.06% (Table 21). Genotypes grown in 2007 lost 1.32% more water than the same cultivars grown in 2006, a statistically significant difference (Tables 20 and 21). The genotype × environment interaction plot displays genotype rank change that caused the statistically significant interaction and it is difficult to discern groups of high and low performing cultivars (Figure 7). Natto cultivar MFS-511 has consistent performance near the top and natto cultivars MFS-553 and Camp have consistently low performance. Identifying the top performing line is difficult, notably cultivar MFS-591 displays significant genotype × environment interaction for water loss after steaming (Figure 7).

Table 20. ANOVA for water loss after steaming listing the significance of main effects and interaction terms for 11 soybean genotypes grown in six environment in Virginia during the 2006 and 2007 growing season.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	10	240	12.25	<.0001
Genotype×Environment	50	240	1.99	0.0003
Genotype×Year	10	240	1.1	0.3633
Genotype×Environment×Year	50	240	1.6	0.0114
Environment	5	24	5.01	0.0028
Environment×Year	5	24	5.6	0.0015
Year	1	24	143.18	<.0001

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

Table 21. Water loss after steaming (%) lsmean estimates for 11 soybean genotypes grown in six environments in Virginia during the 2006 and 2007 growing season.

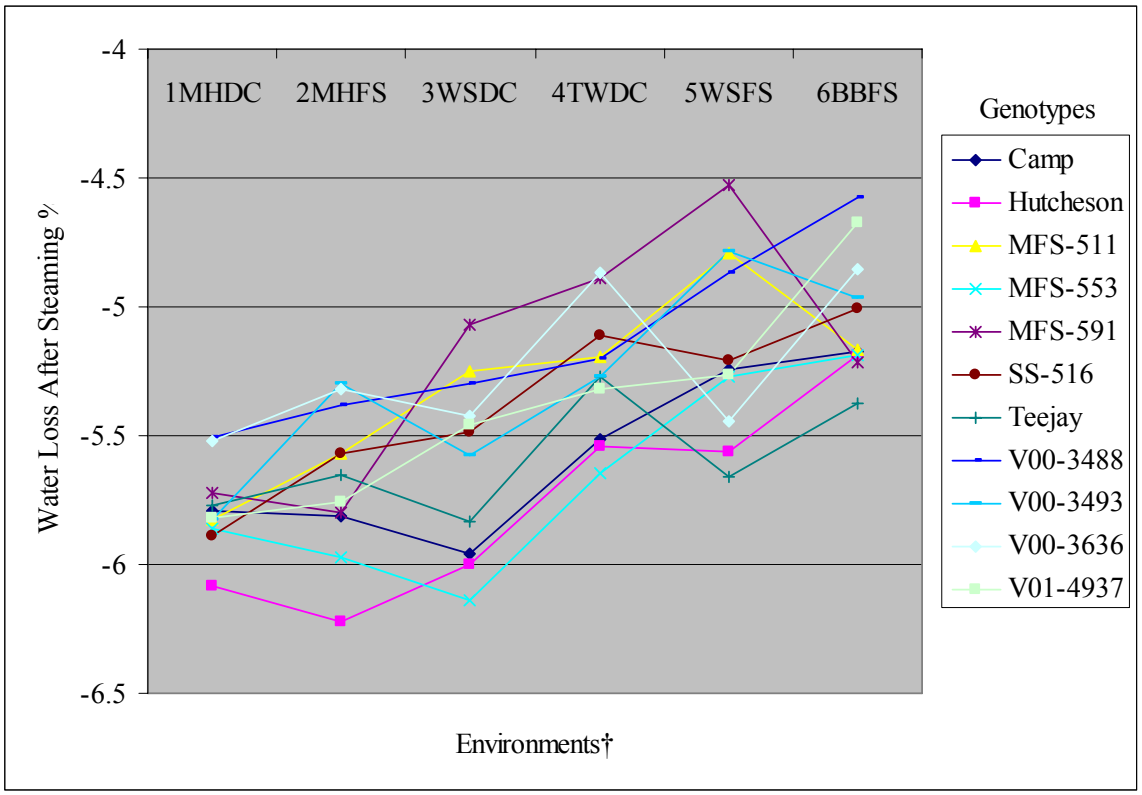
Source	LSMean	N‡	Standard Error
Genotype			
Natto			
V00-3488	5.14	36	0.0800
MFS-591	5.20	36	0.0800
V00-3636	5.24	36	0.0800
V00-3493	5.29	36	0.0800
MFS-511	5.30	36	0.0800
V01-4937	5.38	36	0.0800
Camp	5.58	36	0.0800
MFS-553	5.68	36	0.0800
Conventional			
SS-516	5.38	36	0.0800
Teejay	5.59	36	0.0800
Hutcheson	5.77	36	0.0800
Tukey 0.05†	0.39		
Environment§			
BBFS	5.03	72	0.1374
WSFS	5.15	72	0.1374
TWDC	5.26	72	0.1374
WSDC	5.59	72	0.1374
MHFS	5.67	72	0.1374
MHDC	5.78	72	0.1374
Year			
2006	4.74	216	0.0794
2007	6.09	216	0.0794

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

§: BBFS, Blacksburg full season; WSFS, Warsaw full season; TWDC, Tidewater double crop; WSDC, Warsaw double crop; MHDC, Mt. Holly double crop; MHFS, Mt. Holly full season.

Figure 7. Genotype by environment interaction plot showing lsmean water loss after steaming for 11 soybean genotypes plotted against 6 growing environments averaged across 2006 and 2007.



†: BBFS, Blacksburg full season; WSFS, Warsaw full season; TWDC, Tidewater double crop; WSDC, Warsaw double crop; MHDC, Mt. Holly double crop; MHFS, Mt. Holly full season.

Seed Coat Deficiency

The seed coat deficiency residuals were non-normally distributed and five observations had studentized residuals greater than 3. The plot of residuals vs. environments showed that two observations at Mt. Holly full season had errors far outside of all other observations. These observations both had negative studentized residual values greater than 5. Checking the raw data showed that these observations had drastically different measurements than the other two replicates grown in the same environment. This test is subjective and sampling error could explain the discrepancy. These two values were removed and the data set re-analyzed. The residuals were normally distributed at the 0.001 probability level. The residual vs. predicted values and

the residual at environment plots showed a more homogeneous distribution of variance. The ANOVA output from the two different data sets had identical results for test of significance. The two observations with negative studentized residual values great than 5 were removed and this data set was used for further analysis because it better met the assumptions for combined analysis of variance.

Seed coat deficiencies lead to poor natto quality therefore low seed coat deficiency counts are desirable. The three main factors genotype, environment, and year and the two-way interaction genotype × environment were significant at the 0.0001 probability level as were all two and three way interactions (Table 22). Genotypes grown in Warsaw double crop had the lowest mean seed coat deficiency count of 40%. Cultivars at Tidewater double crop had the highest seed coat deficiency mean of 74%, with 25% more deficient seeds than the next highest environmental mean (Table 23). Cultivar means were similar to those computed using model 1. The difference in seed coat deficiency counts between natto cultivars was large and fairly consistent. Natto cultivar MFS-553 had the lowest mean seed coat deficiency count at 22% and natto line V00-3493 had the highest mean at 74% (Table 23). Year 2007 produced a seed coat deficiency mean that was 12% higher than 2006 (Table 23). The genotype × environment interaction plot displays genotype rank change that caused the statistically significant interaction (Figure 8). The majority of cross-over events in the two-way interaction plot did not result in significant rank change and distinct genotype groupings were present. Natto cultivars Camp and MFS-553 can be grouped as superior genotypes for seed coat deficiency across environments (Figure 8). Natto cultivar MFS-591 had one of the lowest seed coat deficiency lsmeans but exhibited significant genotype × environment interaction at Blacksburg full season (Figure 8).

Table 22. ANOVA for seed coat deficiency listing the significance of main effects and interaction terms for 11 soybean genotypes grown in six environment in Virginia during the 2006 and 2007 growing season.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	10	238	308.16	<.0001
Genotype×Environment	50	238	7.48	<.0001
Genotype×Year	10	238	19.24	<.0001
Genotype×Environment×Year	50	238	6.67	<.0001
Environment	5	24	251.46	<.0001
Environment×Year	5	24	10.39	<.0001
Year	1	24	339.93	<.0001

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

Table 23. Seed coat deficiency (%) lsmean estimates for 11 soybean genotypes grown in six environments in Virginia during the 2006 and 2007 growing season.

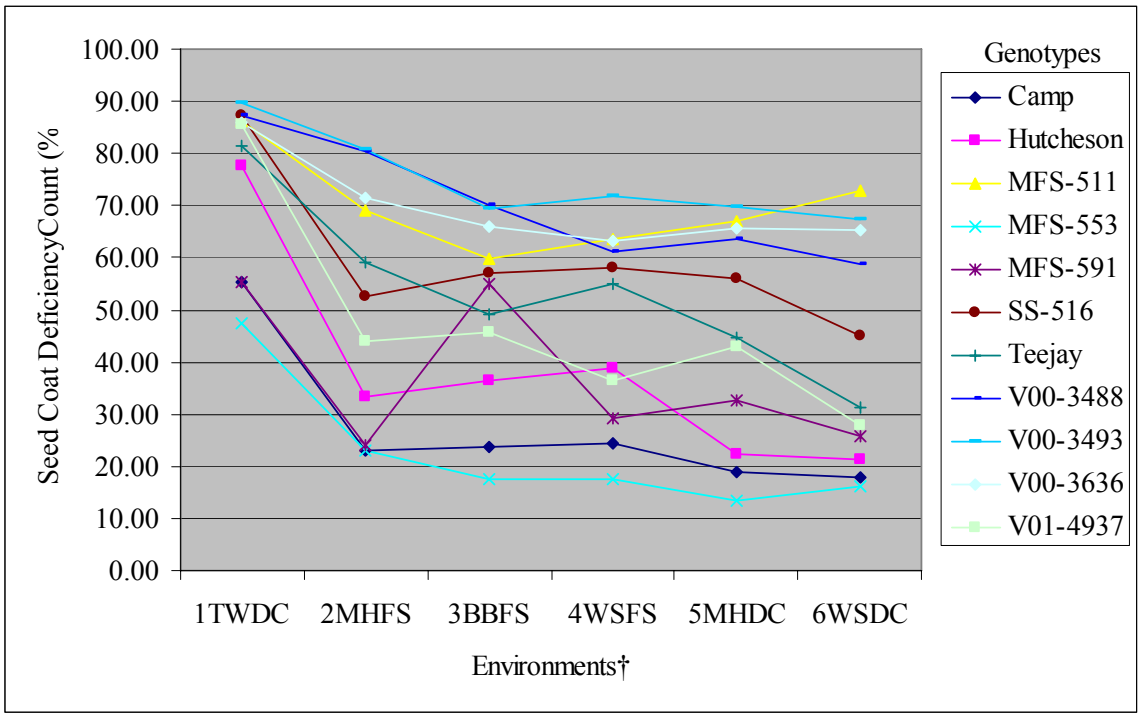
Source	LSMean	N‡	Standard Error
Genotype			
Natto			
MFS-553	23	36	1.0558
Camp	27	36	1.0558
MFS-591	37	36	1.0558
V01-4937	47	36	1.0558
V00-3636	70	36	1.0558
MFS-511	70	35	1.0775
V00-3488	70	35	1.0775
V00-3493	75	36	1.0558
Conventional			
Hutcheson	38	36	1.0558
Teejay	53	36	1.0558
SS-516	59	36	1.0558
Tukey 0.05†	7		
Environment§			
WSDC	41	72	0.7914
MHDC	45	72	0.7914
WSFS	47	72	0.7914
BBFS	50	72	0.7914
MHFS	51	70	0.8087
TWDC	76	72	0.7914
Year			
2006	46	214	0.4603
2007	58	216	0.4569

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

§: TWDC, Tidewater double crop; MHFS, Mt. Holly full season BBFS, Blacksburg full season; WSFS, Warsaw full season; MHDC, Mt. Holly double crop; WSDC, Warsaw double crop.

Figure 8. Genotype by environment interaction plot showing lsmean seed coat deficiency counts for 11 soybean genotypes plotted against 6 growing environments averaged across 2006 and 2007.



†: TWDC, Tidewater double crop; MHFS, Mt. Holly full season BBFS, Blacksburg full season; WSFS, Warsaw full season; MHDC, Mt. Holly double crop; WSDC, Warsaw double crop.

Rate of Water Absorption

The rate of water absorption residuals were non-normally distributed and there were seven observations with studentized residual values greater than 3. The plot of residuals at environments indicated one observation with much higher variance than all other observations. For the seven observations with studentized residual values greater than 3 four were positive and three were negative. For three of the positive values there were corresponding negative values that created symmetry for the residuals. The fourth positive studentized residual had a value greater than 5 and there were no corresponding negative studentized residual values. This observation was removed and the data reanalyzed. The residuals were still non-normal and the plot of residuals vs. predicted values showed non-homogeneous variance. The data were log transformed and

resubmitted for analysis. The residuals were normally distributed at the 0.01 probability level and the plots checking for homogenous variance indicated the assumption was not violated. ANOVA's were computed using both the untransformed and the transformed data and the results for significant effects were identical. Data transformation creates difficulty interpreting response values, and because ANOVA results were identical, only results from the untransformed data are presented here.

Rate of water absorption (min^{-1}) is a rate constant estimating the initial slope of non-linear water absorption and it has been postulated that higher rates are beneficial for natto quality (Cober et al., 2006). The main effects genotype and environment were significant at the 0.01 probability level and year was non-significant (Table 24). All two-way and the three way interactions were significant at the 0.0001 probability level. Natto cultivar V00-3493 had the highest mean rate of water absorption value 7.37×10^{-3} min^{-1} and natto cultivar MFS-553 had the lowest mean rate of water absorption 5.14×10^{-3} min^{-1} (Table 25). Five of the eight natto cultivars tested had significantly higher rates of water absorption than the conventional cultivars tested. Cultivars grown at Warsaw double crop had the lowest mean rate of water absorption 6.05×10^{-3} min^{-1} while those grown at Blacksburg full season had the highest mean rate of water absorption 6.90×10^{-3} min^{-1} (Table 25). The genotype × environment interaction plot displays genotype rank change that caused the statistically significant interaction (Figure 9). Natto cultivar MFS-553 was consistently the worst performer for rate of water absorption across environments. The Tidewater double crop environment created multiple cross over interactions that confound selection. Natto cultivars V00-3636, V00-3493, V00-3488, and MFS-511 were the highest performing entries at four of the six environments (Figure 9).

Table 24. ANOVA for rate of water absorption listing the significance of main effects and interaction terms for 11 soybean genotypes grown in six environment in Virginia during the 2006 and 2007 growing season.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	10	240	93.59	<.0001
Genotype×Environment	50	240	3.67	<.0001
Genotype×Year	10	240	6.33	<.0001
Genotype×Environment×Year	50	240	4.87	<.0001
Environment	5	24	5.35	0.0019
Environment×Year	5	24	12.31	<.0001
Year	1	24	0.15	0.7046

†: Numerator degrees of freedom
‡: Denominator degrees of freedom

Table 25. Rate of water absorption (min^{-1}) lsmean estimates for 11 soybean genotypes grown in six environments in Virginia during the 2006 and 2007 growing season.

Source	LSMean	N‡	Standard Error
Genotype	× 10^{-3}¶		× 10^{-3}¶
Natto			
V00-3493	7.37	36	0.10
V00-3636	7.12	36	0.10
V00-3488	7.04	36	0.10
MFS-511	7.03	36	0.10
V01-4937	6.69	36	0.10
Camp	6.61	36	0.10
MFS-591	6.60	36	0.10
MFS-553	5.14	36	0.10
Conventional			
SS-516	6.13	36	0.10
Teejay	5.76	36	0.10
Hutcheson	5.52	36	0.10
Tukey 0.05†	0.48		
Environment§			
BBFS	6.90	72	0.10
MHDC	6.22	72	0.10
MHFS	6.58	72	0.10
TWDC	6.32	72	0.10
WSDC	6.06	72	0.10
WSFS	6.65	72	0.10
Year			
2006	6.43	216	0.10
2007	6.48	216	0.10

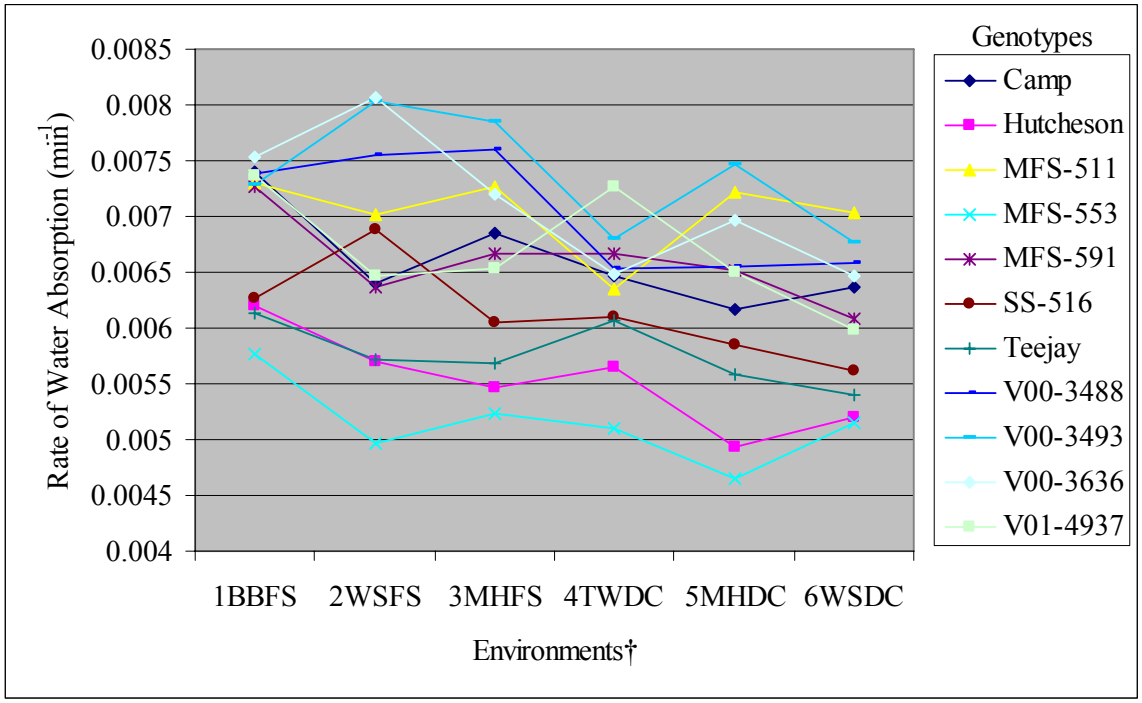
†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

§: BBFS, Blacksburg full season; WSFS, Warsaw full season; MHFS, Mt. Holly full season; TWDC, Tidewater double crop; MHDC, Mt. Holly double crop; WSDC, Warsaw double crop.

¶: Actual rate value equals reported value multiplied by listed factor

Figure 9. Genotype by environment interaction plot showing lsmean rate of water absorption for 11 soybean genotypes plotted against 6 growing environments averaged across 2006 and 2007.



†: BBFS, Blacksburg full season; WSFS, Warsaw full season; MHFS, Mt. Holly full season; TWDC, Tidewater double crop; MHDC, Mt. Holly double crop; WSDC, Warsaw double crop.

Seed Size

The residuals for seed size data were non-normally distributed and three observations had studentized residual values greater than 3. Two observations had similar studentized residual values, one positive and the other negative, while the third observation had the largest value and was negative. The observation was removed and the data re-analyzed. The second analysis was normal distributed at the 0.01 probability level and the residual vs. predicted value plot and the residual vs. environment plot indicated the assumption of homogenous variance was met. The second data set was used for further analysis because it met the assumptions for analysis of variance.

All three main factors were significant at the 0.0001 probability level (Table 26). The two-way interaction genotype × environment was significant at the 0.0001

probability level, while the genotype × year interaction was significant at the 0.05 probability level (Table 26). The two way interaction environment × year and the three way interaction were significant at the 0.0001 probability level. The mean seed size estimates are similar to the values generated using model 1. For natto genotypes, cultivar Camp had the smallest mean seed size 8.11 g per 100 seed and line V00-3488 had the largest seed size 10.98 g per 100 seed (Table 27). Cultivars grown at Warsaw double crop had the smallest mean seed size of 10.78 g per 100 seed and the same cultivars grown at Mt. Holly full season had the largest seed size of 12.80 g per 100 seed (Table 27). The mean seed size for cultivars grown in 2007 was 0.43 g per 100 seed significantly larger than the same cultivars grown in 2006 (Table 26 and 27). The genotype × environment interaction plot displays genotype rank change that caused the statistically significant interaction (Figure 10). Most cultivars follow the same pattern of seed size response grown across environments. Natto cultivar MFS-553 and Camp exhibited a differential response when grown at Tidewater full season and natto cultivar MFS-591 displayed differential response when grown at Blacksburg full season (Figure 10). GEI for seed size does not result in significant rank change for most genotypes and environments tested.

Table 26. ANOVA for seed size data listing the significance of main effects and interaction terms for 12 soybean genotypes grown in two locations in Virginia with two planting systems during 2006 and 2007.

Effect	Num DF†	Den DF‡	F Value	Pr > F
Genotype	11	262	1306.11	<.0001
Genotype×Environment	55	262	4.28	<.0001
Genotype×Year	11	262	2.31	0.0104
Genotype×Environment×Year	55	262	3.35	<.0001
Environment	5	24	58.31	<.0001
Environment×Year	5	24	88.42	<.0001
Year	1	24	29.08	<.0001

†: Numerator degrees of freedom

‡: Denominator degrees of freedom

Table 27. Seed size (g per 100 seed) lsmean estimates for 12 soybean genotypes grown in six environments in Virginia during the 2006 and 2007 growing season.

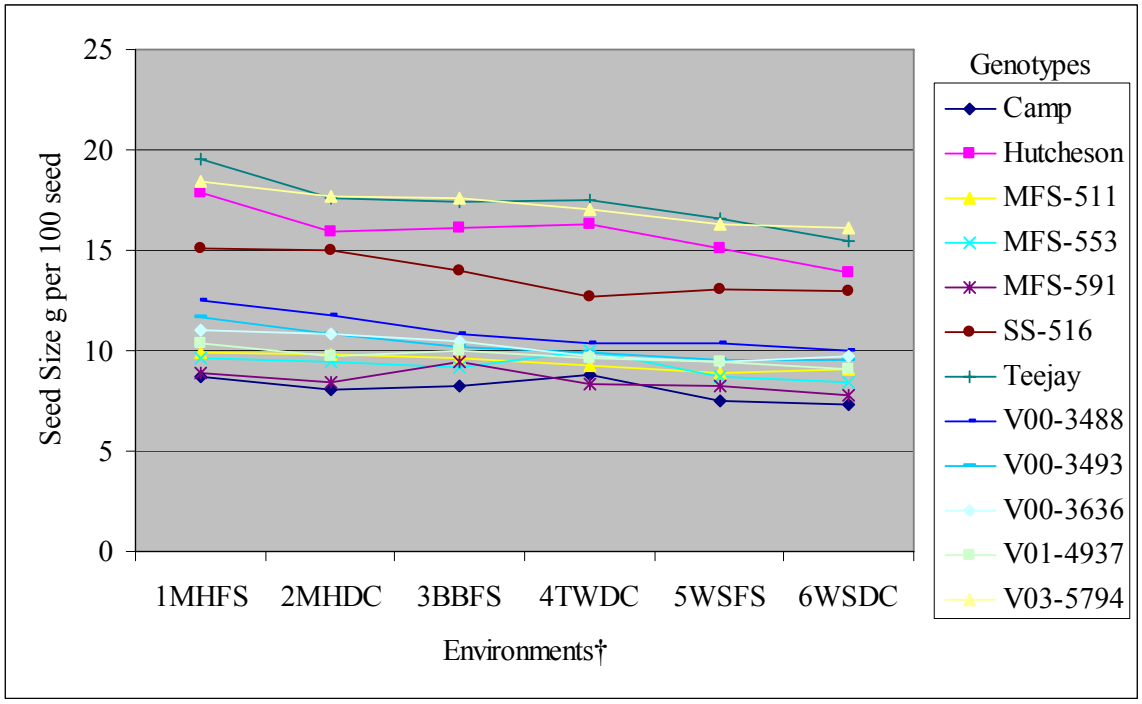
Effect	LSMean	N‡	Standard Error
Genotype			
Natto			
Camp	8.11	36	0.0974
MFS-591	8.54	36	0.0974
MFS-553	9.23	36	0.0974
MFS-511	9.43	36	0.0974
V01-4937	9.70	36	0.0974
V00-3636	10.21	36	0.0974
V00-3493	10.27	36	0.0974
V00-3488	10.98	36	0.0974
Conventional			
SS-516	13.89	35	0.0993
Hutcheson	15.87	36	0.0974
V03-5794	17.17	35	0.0993
Teejay	17.34	36	0.0974
Tukey 0.05†	0.62		
Environment§			
WSDC	10.78	72	0.0945
WSFS	11.09	72	0.0945
TWDC	11.68	71	0.0950
BBFS	11.92	72	0.0945
MHDC	12.09	72	0.0945
MHFS	12.80	71	0.0950
Year			
2006	11.51	215	0.0616
2007	11.94	215	0.0616

†: Tukey's W multiple comparison value ($\alpha = 0.05$) calculated for all genotypes

‡: The number of observations used to calculate the least square mean

§: MHFS, Mt. Holly full season; MHDC, Mt. Holly double crop; BBFS, Blacksburg full season; TWDC, Tidewater double crop; WSFS, Warsaw full season; WSDC, Warsaw double crop.

Figure 10. Genotype by environment interaction plot showing lsmean seed size for 12 soybean genotypes plotted against 6 growing environments averaged across 2006 and 2007.



†: MHFS, Mt. Holly full season; MHDC, Mt. Holly double crop; BBFS, Blacksburg full season; TWDC, Tidewater double crop; WSFS, Warsaw full season; WSDC, Warsaw double crop.

Trait Correlation, Rank Summary, and Stability

Pearson partial correlation coefficients indicated multiple significant correlations between agronomic and natto quality traits (Table 28). Partial correlations allow for correlations to be calculated within treatments. For this experiment it was more appropriate to calculate correlations by cultivars because not all cultivars were specifically developed for natto production. Using partial correlations calculated within cultivars helped create more relevant data.

A summary of genotype rankings for agronomic and natto quality traits is presented in Table 29. It is difficult to determine the most desirable seed size for natto and most natto cultivars have a similar seed size. The seed size rank is not based on superiority but indicates cultivar size as smallest (rank 1) to largest (rank 12). The

average ranking elucidates which cultivars have desirable performance for all traits tested and the standard deviation of rank allows one to determine if the average rank is consistent. The average rank and standard deviation did not include seed size rankings because seed size was not ranked according to superiority. Natto line V00-3488 had the lowest average rank indicating it was overall the most superior natto genotype, but the seed coat deficiency rank of 10 should be noted. Three natto lines had an average rank of 5.2, but one line V01-4937, had a lower rank standard deviation than the other two indicating it had a more stable rank.

Two stability estimates were calculated for each of six response variables (Table 30). For both statistics, the lower the value the more stable the genotype is for the trait. The traits cannot be compared to determine which is more stable because the responses are of different magnitude. The coefficient of variation (CV) for each stability statistic provides a value to compare which statistics had the most variability between cultivars and allows comparison of variability between traits. Natto cultivar MFS-511 had a lower environmental variance (S_i^2) stability estimate than the environmental variance mean for all six traits measured. Natto cultivars Camp and MFS-553 and natto line V00-3636 had environmental stability estimates lower than the environmental stability mean for five of the six traits tested. Conventional cultivar Hutcheson only had one environmental variance stability estimate lower than the mean environmental variance for one trait, and conventional cultivar Teejay's environmental variance stability estimates were not below the mean for any trait measured. Natto cultivar V01-4937 had an ecovalence (W_i^2) estimate below the ecovalence mean for all six traits measured. Natto cultivar Camp and conventional cultivar SS-516 had ecovalence estimates lower than the mean ecovalence

for five of the six traits measured. Cultivar Teejay had one ecovalence estimate lower than the mean ecovalence for one trait measured.

Table 28. Pearson partial correlation table for agronomic and natto quality traits for 11 soybean genotypes tested at 4 locations using two planting systems during the 2006 and 2007 growing season.

Trait	Seed Size	Yield	Water Absorption	Water Loss	Seed Coat Deficiency	Rate of Water Absorption
Initial Moisture	0.041	-0.324***	-0.178**	-0.520***	0.054	-0.013
Seed Size		0.236***	0.239***	-0.170**	0.262***	-0.080
Yield			-0.138*	0.171**	-0.165**	-0.047
Water Absorption				0.008	0.178**	-0.087
Water Loss					-0.065	0.125*
Seed Coat Deficiency						0.389***

* p-value <0.01
** p-value <0.001
*** p-value <0.0001

Table 29. Ranking summary for the agronomic and natto quality traits tested using 12 soybean genotypes grown in six environments during the 2006 and 2007 growing season. Rank is based on mean values from model 2 and a rank of 1 indicates the most desirable performance.

Genotype	Seed Size†	Yield	Water Absorption	Water Loss	Seed Coat Deficiency	Rate of Water Absorption	Average Rank‡	Rank Standard Deviation‡
Natto								
Camp	1	10	10	8	2	6	7.2	3.35
MFS-511	4	7	8	5	8	4	6.4	1.82
MFS-553	3	9	11	10	1	11	8.4	4.22
MFS-591	2	11	9	2	3	7	6.4	3.85
V00-3488	8	5	1	1	10	3	4.0	3.74
V00-3493	7	8	2	4	11	1	5.2	4.21
V00-3636	6	6	6	3	9	2	5.2	2.77
V01-4937	5	3	7	6	5	5	5.2	1.48
Conventional								
Hutcheson	10	1	5	11	4	10	6.2	4.21
SS-516	9	2	3	7	7	8	5.4	2.70
Teejay	12	4	4	9	6	9	6.4	2.51
V03-5794	11	12	N/A	N/A	N/A	N/A	N/A	N/A

†: Genotypes ranked on seed size and not superiority. Rank 1 is the smallest and rank 12 is the largest
‡: Average rank and standard deviation are computed using all listed traits except seed size.

Table 30. Stability analysis for agronomic and natto quality traits for twelve soybean genotypes grown in thirteen environments.

Genotype	Yield		Water Absorption		Water Loss		Seed Coat Deficiency		Rate of Water Absorption		Seed Size	
	S_i^2	W_i^2	S_i^2	W_i^2	S_i^2	W_i^2	S_i^2	W_i^2	$S_i^2 \times 10^{-7}$	$W_i^2 \times 10^{-6}$	S_i^2	W_i^2
Camp	**206.67**	**162.60**	**8.58**	**21.91**	**0.643**	0.875	378.23	**734.97**	**5.00**	**2.00**	**0.881**	**3.684**
Hutcheson	**216.90**	**202.24**	20.51	54.62	0.704	**0.721**	508.68	1497.63	8.20	**2.20**	2.334	8.049
MFS-511	**222.64**	**33.59**	**5.36**	**29.40**	**0.549**	**0.592**	**155.58**	1324.18	**2.80**	4.10	**0.698**	**3.012**
MFS-553	**181.92**	243.60	**9.54**	**25.40**	**0.619**	0.964	353.25	**913.86**	**3.70**	**2.50**	**1.005**	**4.097**
MFS-591	**158.02**	387.10	15.70	54.28	0.900	1.783	332.45	3520.45	**2.90**	5.40	**0.862**	**4.426**
SS-516	248.76	**107.58**	**9.53**	**18.63**	**0.658**	**0.317**	**263.21**	**307.15**	**6.40**	**2.30**	2.297	6.904
Teejay	264.02	343.22	20.17	63.85	0.794	0.894	414.19	**1055.18**	9.30	4.40	3.629	10.600
V00-3488	256.70	**125.30**	**9.12**	43.26	0.755	**0.420**	349.69	1545.18	9.50	**3.00**	1.567	**3.675**
V00-3493	232.13	**174.60**	**5.56**	47.77	**0.655**	**0.740**	**293.31**	1567.03	11.00	5.50	**1.455**	**4.106**
V00-3636	**213.44**	288.65	**7.34**	45.34	**0.670**	1.099	**176.08**	**703.22**	8.50	**3.30**	**0.947**	**1.350**
V01-4937	253.46	**204.03**	14.19	**25.91**	**0.670**	**0.680**	413.63	**991.54**	7.20	**2.80**	**1.091**	**1.824**
V03-5794	237.75	396.85	**9.38**	**16.14**	**0.541**	**0.507**	**134.76**	**399.84**	7.70	**2.10**	1.532	**3.846**
Mean	224.37	222.45	11.25	37.21	0.68	0.80	314.42	1213.35	6.85	3.30	1.52	4.63
CV, %	15.50	56.16	50.82	43.41	16.22	48.13	40.08	69.34	42.98	38.54	61.57	57.05

S_i^2: Environmental variance of genotypes as described by Lin et al. (1986).
W_i^2: Ecovalence as described by Lin et al. (1986).
Values in bold are lower than the genotype stability average for the response variable.

Discussion and Conclusions

Yield

Yield analysis with model 2 indicated significant genotype × environment interaction suggesting differential cultivar performance over environments (Table 16). The interaction plot displays multiple cross over resulting from differential response across the different growing environments (Figure 5). Identifying the top yielding cultivars is possible but rankings for the moderate yielding cultivars change over environments and confound selection. Yield is a polygenic trait and it is strongly influenced by environmental conditions, therefore genotype testing is necessary at many locations over multiple years to determine a cultivars yield potential.

Although yield is not the most important trait for selecting superior natto cultivars, it does impact how growers will accept the cultivar and their willingness to plant it. Line V01-4937 was the highest yielding natto cultivar and it did not yield significantly less than the top conventional cultivars tested (Table 17). It also had an ecovalence stability estimate lower than the mean estimate, indicating stability (Table 30). This information is consistent with the data presented in the yield interaction plot. The second highest yielding natto line, V00-3488, did not yield significantly less than V01-4937 but it did yield significantly less than the three top conventional cultivars. Line V00-3488 had one of the lowest ecovalence stability estimates for the genotypes tested. Based on the yield interaction plot and the ecovalence stability estimates, conventional cultivar Teejay and natto cultivar MFS-591 had poor stability and their yield responses were more influenced by environment than the other genotypes tested.

Blacksburg full season, Mt. Holly double crop, and Mt. Holly full season had the highest overall mean yield estimates for the environments tested. Both Mt. Holly tests were grown under irrigation which can account for it being a high yielding environment. Blacksburg was cooler and moister than Warsaw and Tidewater during both years of the experiment. These favorable growing conditions, along with the soil in Blacksburg having a higher water holding capacity, likely explain why it was a high yielding growing environment. The yield interaction plot (Figure 5) indicates that although genotype is an important factor controlling yield response, environmental conditions significantly affect the trait.

The lack of significant genotype × year interaction indicates no differential yield response between cultivars grown in 2006 and 2007. The main effect year was significant however, and resulted in an overall drop of 590.16 kg/hectare from 2006 to 2007 (Table 17). This response is likely attributed to the adverse growing conditions, severe drought, experienced at all locations in 2007. Environmental conditions change each growing season and cannot be predicted but breeding for environmental stability is important for growers expecting consistent performance.

Stability estimates provided additional information about cultivar environmental sensitivity for yield (Table 30). Ecovalence is the more appropriate stability statistic for yield because it accounts for the other genotype's responses across environments and penalizes cultivars less for deviations from the genotype means (Lin et al., 1986). The ecovalence CV for yield indicates it was one of the most variable stability estimates. Yield is controlled by many genetic factors and interactions and variability is expected. The ecovalence estimates support the yield interaction plot and conveys additional information about cultivar superiority. The yield results from the current study illustrate

how using multiple sets of data such as interaction plots, least square means, and stability statistics provide allows one to identify superior genotypes with desirable, consistent performance.

The yield ecovalence stability estimates (W_i^2) appear inflated compared to data from another multi-environment soybean trial (Yue et al., 1997). The larger ecovalence estimates in this experiment indicate the genotypes tested are more environmentally sensitive. A direct comparison of the stability estimates from the two experiments is inappropriate because different genotype, locations, and years were tested, but an informal analysis indicates the stability statistics reported by Yue et al (1997) were lower. One explanation for the difference is the significant year effect observed in this experiment. The 2007 growing season resulted in significantly lower yields and created a wide range of responses between the two years. Also, the environments in this study were variable and included both full season and double crop plantings. This created disparate growing conditions and the possibly less favorable environments, and high stability may not have been biologically possible. Another explanation is the diverse types of genotypes grown in this study could have introduced more variation between genotype responses and led to larger ecovalence stability values.

Based on the results for this experiment, growing genotypes in multiple locations and years is necessary to collect accurate yield data and identify consistently superior cultivars (Kang, 1998). The genotype × environment interaction does not indicate the environments tested in this experiment elicited drastic cultivar rank change, and a sample of the environments is representative. Selection for yield based on multi location trials should produce accurate data and help identify superior cultivars. The lack of genotype × year interaction shows that although cultivars performed worse in 2007 they performed

worse as a group. This indicates that ranks presented in this experiment should in subsequent years provide similar conclusions. Yearly conditions are not predictable, so multi year testing is necessary. Genotype was an important factor in determining cultivar response and the data indicate that breeding high yielding natto cultivars is possible.

Maturity

Maturity data were not present for all locations planted during 2006 and 2007, therefore model 1 was used to analyze full season vs. double crop data at Mt. Holly, Tidewater, and Warsaw. Comparing results between locations and years is difficult because planting dates varied. An example of this is the difference in maturity between years (Table 5). On average, cultivars matured 10 days earlier in 2007 than 2006, but the planting dates at almost all locations were earlier in 2007. Therefore, it is not possible to conclude the cause of earliness in 2007 but it is likely the result of the warmer and drier 2007 growing season. It is also difficult to determine how location affected maturity because the planting dates were not the same at all locations confounding interpretation. It is probable that cultivars grown at Mt. Holly had a later maturity date because the location was irrigated.

Planting dates also varied between planting systems but they are partially defined by planting date so the overall effect of planting system can be determined. Based on means, cultivars planted later in a double crop system matured later as expected (Table 5). Dry environmental conditions forced the full season test to be planted later than usual at Mt. Holly in 2007 and the full season test was planted on the same days as the double crop test. An informal analysis of the 2007 Mt. Holly data indicates that maturity was the same between the full season and double crop test. This provides information that the

maturity differences reported between planting systems are the result of planting date, and the other environmental differences between full season and double crop effect maturity to a lesser extent. The genotype × planting system interaction plot for maturity showed rank change, but due to the low genetic variation between the cultivars tested, i.e. all are group V maturity, little variation was seen for the trait (not shown). Soybeans are grouped by maturity to denote local adaptation to photoperiod required for flowering and is mainly under genetic control (Poehlman and Sleper, 1995).

Water Absorption

Water absorption data analyzed with model 1 indicated significant genotype, planting system, and genotype × planting system interaction (Table 6). When visualized however, there is little indication that the interaction confounds selection because cultivar rank change was minimal between the two planting systems (Figure 1). Natto lines V00-3488 and V00-3493 had water absorption estimates significantly higher than the other natto cultivars tested (Table 7). The data indicates cultivar performance rankings were similar between the two planting systems and testing in either is sufficient. Therefore, cultivars do not need to be specifically developed for full season or double crop production because genetic responses were similar (Fehr, 1987). This is desirable for breeders because selection in multiple cropping systems would be cumbersome and resource intensive.

Planting system alone had a significant effect on water absorption and cultivars grown full season had a lower water absorption estimate (Table 6). This is important information for natto soybean growers because higher water absorption is desirable. It should be noted however that soybeans planted double crop instead of full season

averaged <2% water absorption gain, but it is difficult to determine what percent water absorption increase is significant for natto production. The effect was not the same for all cultivars however, and natto cultivar MFS-591 absorbed 4.11% more water when planted double crop. The differential response indicates water absorption improvement based on production system is cultivar dependant. This is important information for natto breeders and growers making decisions about cropping systems and genotypes. Without natto quality testing by natto manufacturers, however, conclusions remain speculatory.

The second analysis of water absorption using model 2 also indicated significant GEI for water absorption (Table 18). The effects of genotype × environment interaction did not cause inter-performance group change, and the top seven and the bottom four performing cultivars remain separated (Figure 6). Natto lines V00-3488 and V00-3493 had significantly higher water absorption values than the other natto genotypes tested (Table 19). The environmental variance stability estimates for the two cultivars were below the trait mean for water absorption indicating stable performance relative to the other genotypes tested (Table 30). These results provide evidence that identifying natto genotypes with high stable water absorption is possible and cultivar genotype played a significant role influencing trait expression. The natto lines with the 3rd and 4th highest water absorption estimates were also significantly higher than the other natto lines evaluated but lower than the two lines already mentioned. The line with the 4th highest water absorption estimate, V01-4937, had an environmental variance stability estimate above the mean environmental variance. This is highlighted by the fact that there was an 8.63% water absorption difference for V01-4937 between the highest and lowest environmental means (Figure 7). It is possible to breed high water absorbing natto cultivars but multi location testing is necessary because of environmental influence.

It is interesting that all seven natto genotypes testing in this experiment had lower (more stable) environmental variance stability scores for water absorption than that of cultivar MFS-591 because MFS-591 has long been a natto standard and is perceived as producing consistent high quality natto (Montague Farms, personal communication, 2006). If this conclusion is accurate, it indicates that germplasm is available in the program to further produce superior natto cultivars beyond MFS-591. Another possibility is that the environmental variance stability statistic does not accurately reflect the water absorption potential for MFS-591. Data analysis and Figure 6 indicate that MFS-591 produced an above average environmental stability score because of its performance at Warsaw double crop. It is difficult to surmise why samples of MFS-591 grown at Warsaw double crop absorbed more water than any other location. MFS-591 did not display interactions with other environments and had consistent performance over the other locations (Figure 6). It appears MFS-591 had a consistent water absorption increase in 2006 and 2007 at Warsaw double crop that lead to its unstable environmental variance estimate but data from more years is necessary to determine the accuracy of these findings.

The stability scores for water absorption highlight the difference between environmental variance and ecovalence. Genotypes V00-3493 and SS-516 have similar environmental stability scores, 5.56 and 9.53, but different ecovalence scores, 47.77 and 18.63, respectively (Table 30). The line for SS-516 intersects that of V00-3493 multiple times (Figure 6), indicating that SS-516 is more variable across locations and more environmentally responsive than V00-3493. The range of values for both cultivars does not fluctuate too far from their means, a fact supported by their low S_i^2 scores, yet SS-516 has an ecovalence score 29 points less indicating that SS-516 is more stable. The

results of the two different stability scores illustrate how the environmental variance stability score penalizes genotypes for responses that fluctuate from their mean response, while genotypes that do not follow a response trend based on environmental means are penalized for ecovalence. Said another way; for a genotype to have a stable environmental variance estimate responses should not change over different environments, but for ecovalence stability estimate to be stable genotype responses must follow environmental productivity trends. These results illustrate that the environmental variance stability statistic is more appropriate for quality traits because consistent performance is more important than relative performance (Becker and Leon, 1988).

Year was not a significant factor for water absorption but there was a significant year interaction with genotype and environment (Table 18). Stability over years is desirable for growers, but the genotype × year interaction indicates differential cultivar response. Selection was not affected but there is evidence that genotypes do not respond similarly to different yearly conditions. Natto lines V00-3488, V00-3493, and V00-3636 all experienced an increased water absorption estimate between 2006 and 2007, while natto cultivars MFS-591 and Camp displayed a decrease. It is difficult to explain the reason for the differential response, but it illustrates that different location and year factors can combine to produce different water absorption responses. The water absorption data indicates that genotype is an important factor for determining trait development and stable high performing genotypes were identified, but environmental factors such as location and year can influence the trait and require testing.

GEI played a role in water absorption for this experiment and has been reported as significant in other small seeded material grown in Canada and Iowa (Cober et al., 1997a; Geater et al., 2000). For the material grown in Canada, genotype was a significant factor

along with genotype × location and genotype × year for water absorption (Cober et al., 1997a). Rank change data caused by GEI was not present for the trait, but the author concludes that genotype is more important than the other environmental factors for determining water absorption. Geater et al. (2000) found significant GEI for water absorption only when the data were analyzed within years and within locations but not when analyzed as an entire experiment over two years. This statistical analysis is different than the one conducted here but similar GEI was present. One difference between the results of the current study and that of and Geater et al. (2000) was the significant role year had on water absorption. As noted above, year alone was not a significant factor in the current study but year was a major source of variation in their experiment. Specifically, one variety changed rank from the first to fifteenth out of sixteen cultivars between the two years tested. The difference for the results could be related to the combination of year and environment factors or the genotypes used in the two studies. Based on the results of the current experiment and the finding of Geater et al. (2000) it appears that water absorption is controlled by a strong genetic component, but environmental influence is important and should not be overlooked. Testing in multiple locations and years is necessary to identify cultivars with high water absorption and stable performance.

Water Loss after Steaming

Water loss after steaming data analyzed with model 1 indicated a non-significant genotype × planting system interaction (Table 8). This indicates that all genotypes responded similarly between the double crop and full season plantings and rankings did not change for water loss after steaming. The main effect planting system was

statistically significant but the difference in water loss after steaming for cultivars grown double crop vs. full season was 0.29% (Table 9). It is unlikely that this change would significantly alter natto quality but testing by a natto manufacturers is needed for conclusive results. The results indicate that planting system does not influence water loss after steaming in a significant manner and cultivars do not need to be targeted for specific planting systems because cultivar performance is not different between planting systems.

The second analysis of water loss after steaming with model 2 indicates a significant genotype × environment interaction but an insignificant genotype × year interaction (Table 20). The genotype × environment interaction plot displays significant cross-over interaction resulting from differential cultivar response across environments (Figure 7). Natto cultivar MFS-591 has the second lowest mean water loss after steaming estimate, which is desirable, but shows significant rank change resulting in the inability to accurately predict its performance across environments. This is reflected by MFS-591 having the most instable stability scores for both stability statistics (Table 30). Another natto line, V00-3488, has the most desirable water loss after steaming estimate but an unstable environmental variance rating. Natto lines V00-3493, V00-3636, and V01-4937 all had water loss after steaming estimates that were not significantly different than the best natto line but all three had more stable environmental stability estimates. This illustrates that mean performance should not comprise the only information used for selection when quality traits are important. Using multiple sets of data, such as mean performance and stability can identify superior stable genotypes.

Year played a more significant role in determining water loss after steaming than either genotypes or planting environment. The unfavorable 2007 growing season resulted in seeds that lost more water after steaming than those grown in 2006, a difference of

1.35%. The harsher growing conditions may have produced seeds with diminished integrity that allowed more water to leave the seed during steaming. Compare these findings to the observed difference in mean water loss after steaming between the highest and lowest performing natto cultivars of 0.54%. This suggests that year caused more variation in the data then did genotype differences.

The significant role environment played in determining water loss after steaming is overshadowed by the lack of genetic variation in the tested material (Table 21). There were statistically different genotype means but it is difficult to conclude that practical differences in natto quality would result from the variation. If this conclusion is accurate, breeding for improved water loss after steaming would be expected to have minimal success (Fehr, 1987). The differences between environments that caused trait variation show no discernable patterns and appear to be negligible although statistically significant.

The study by Geater et al. (2000) also measured water loss after steaming, and reported no significant difference for locations or years but significant variation for genotypes. These results are different than those presented in this research, illustrating the different outcomes possible using different genotypes and environments in multi environment trials. While GEI did have a statistically significant effect on water loss after steaming it does not warrant attention in a breeding program because the genetic variation for the material tested was minimal. The current study suggests that the trait is not useful for developing superior natto cultivars.

Seed Coat Deficiency

Seed coat deficiency data analyzed with model 1 indicated a significant genotype × planting system interaction (Table 10). Cross over interactions are seen in the

interaction plot and differential cultivar response can be observed (Figure 2). For natto genotypes, the two cultivars with the most desirable performance, MFS-553 and Camp, experienced a slight decrease in deficient seed count when planted double crop vs. full season which contrasts the performance of MFS-511 and MFS-591 which produced more deficient seeds in the double crop system. It is not possible to determine the genetic cause for the differential response but it does provide evidence that cultivars respond differently to double crop vs. full season plantings for seed coat deficiency count. The genotype × planting system interaction did not cause significant rank change suggesting that although cultivars responded differently to the two planting systems, genetic rankings were consistent and more influential in determining trait expression. Cultivar development for specific planting systems is not warranted based on the results of the current research.

The overall effect of planting system was significant for seed coat deficiency and cultivars grown full season had a higher seed coat deficiency count than those grown double crop (Table 11). Soybeans grown double crop produced on average 43 deficient seeds compared to 49 when planted full season, a 6% difference (Table 11). This provides evidence that growing natto cultivars under double crop management may produce superior natto quality over the same cultivar grown full season. The increased seed deficiency scores seen in full season indicates that the full season plantings produce cultivars with compromised seed coats compared to the double crop planting. The reason full season plantings produce seeds with more deficiencies cannot be determined from this study. The most probable explanation is that cultivars grown under full season conditions experience more environmental weathering that adversely affects seed coat development or reduces structural integrity. As noted earlier, however, the increase in

seed coat deficiency counts for full season plantings were not seen in all genotypes because of differential response. These results indicate it is possible to develop genotypes that are not sensitive to different planting systems because genotype was the most important factor controlling seed coat deficiency count.

The second analysis of the seed coat deficiency data using model 2 showed a significant genotype × environment and genotype × year interaction (Table 22). The genotype × environment interaction plot displays fairly consistent performance groupings over environments indicating that genetic control for the trait is strong (Figure 8). Natto cultivars MFS-553 and Camp were significantly better for seed coat deficiency count than the other natto genotypes tested. Neither cultivar had an environmental variance estimate lower than the mean estimate which indicates instability. The 3rd and 4th most desirable performing cultivars, MFS-591 and V01-4937, also had environmental variance stability estimates that indicated instability, while the three most undesirable performing cultivars all had stable environmental variance estimates (Table 30). These results indicate that the genotypes in the current study are not desirable for natto production based on the seed coat deficiency trait because no genotype had desirable seed coat deficiency counts and consistent performance across environments. A more detailed analysis of the results, however, indicates that the environmental variance estimate may be strongly influenced by the Tidewater double crop environment and may not accurately predict a genotype's potential. All genotypes performed worse at Tidewater double crop for seed coat deficiency count, but the best performing cultivars, MFS-553, Camp, MFS-591, and V01-4937 were the most severely affected by the environment (Figure 8). For these genotypes, the large difference in seed coat deficiency counts between Tidewater double crop and the other environments caused the large (unstable) environmental

variance stability estimate. These natto genotypes had consistent performance at the other environment, except MFS-591 at Blacksburg full season (Figure 8). This situation reveals the limitation that single environmental observations can have on the environmental stability statistic influence and that stability statistics should be viewed in a larger context with other data. Results from the current study indicate that the seed coat deficiency trait is controlled mainly by genetics but growing environment can influence the traits expression.

It is difficult to explain the performance of MFS-591 in the Blacksburg full season test. The data were checked for transcription errors but none were found. MFS-591 had an average seed coat deficiency count of 76 in 2006 and 34 in 2007 in the Blacksburg full season test. This indicates that some environmental condition in 2006 caused a higher seed coat deficiency score for MFS-591 at this location. An explanation could be some harvesting or sampling error, but it is unlikely that MFS-591 would have been the only cultivar affected and no other cultivars displayed the same interaction. It is possible that all other cultivars were at the appropriate harvest moisture and all three plots of MFS-591 were either too dry or too wet, but this is improbable. There was an early frost at Blacksburg 2006 which may have affected the seed coat deficiency count of MFS-591 but no other cultivar was affected. Another explanation is that the data are inaccurate because of sampling error but it is unlikely that one cultivar was the subject of data collection error while the other samples were rated correctly. Continued observation of MFS-591 is needed to determine it's environmentally sensitive for the seed coat deficiency trait.

The mean for the Tidewater double crop test suggests this location represents environmental conditions that cause a lot of deficient seeds (Table 23). It had the highest

mean seed coat deficiency count both years, and the incomplete data from Tidewater full season suggests the same results. Only 2006 data are available for Tidewater full season because the quality of 2007 seed was too poor to collect meaningful natto quality data, reinforcing the hypothesis that it produces undesirable natto quality as a location. The possible explanations for why the Tidewater location produced seeds with greater seed coat deficiency counts include environmental factors such as soil type and weather conditions and also anthropogenic factors such as harvest conditions and combine damage. No single factor can be determined as controlling seed coat deficiency counts from the current research but the results illustrate that non-genetic factors significantly influence seed coat deficiency count.

Genotype × year and year were also significant factors (Table 22). The genotype × year interaction displays insignificant rank change and most cultivars developed more deficient seeds in 2007 than 2006 (not shown). This indicates the cultivars responded similarly between the two growing seasons. The only cultivar that did not develop more deficient seeds in 2007 than 2006 was natto cultivar MFS-591. This suggests MFS-591 is less sensitive to environmental change than the other genotypes tested, despite its instable environmental variance estimate.

The effect of year on the genotypes tested appears to be the most significant environmental factor. Cultivars grown in 2007 produced on average 12 more deficient seeds than in 2006 (Table 23). This would likely adversely effect natto quality and receive poor ratings from natto manufacturers. The data from the 2006 and 2007 growing seasons is not detailed enough to determine an environmental cause, but it is evident that the 2007 growing season adversely affected the natto quality trait. This creates a challenge to identify superior lines with high stability for natto quality because

years are unpredictable. Genotypes were the most important predictor of seed coat deficiency but yearly influence impacted cultivar ratings.

Rate of Water Absorption

The analysis for rate of water absorption using model 1 identified genotype, planting system and their interaction as significant sources of variation (Table 12). The two-way interaction plot indicates genotype × planting system interaction, but significant rank change is not observed (Figure 3). Rank change does occur within the highest performing group but it is not confounding because it is possible to identify superior cultivars for the trait. The differential response to planting type indicates that some genotypes are more sensitive to changes in planting system but because relative performance between the two systems is similar developing cultivars for specific planting systems is unnecessary.

The data indicates that cultivars planted full season absorbed water at a statistically significant higher rate than when planted double crop, a difference of 0.48×10^{-3} min^{-1} (Table 13). Rate of water absorption estimates a rate constant and it is difficult to judge what magnitude difference would significantly affect natto quality without explicit testing. The reason for the increased water absorption rate between full season and double crop tests may be similar to the reason for increased seed coat deficiency counts between full season and double crop tests. Cultivars grown full season may produce seeds with more compromised seed coats which would account for the overall increase in deficient seeds and increased water absorption rate.

This hypothesis is supported by other water absorption research. It is well documented that the seed coat is the initial barrier to water entry (McDonald et al., 1988;

Meyer et al., 2007) and the seed coat must be intact to retard water absorption (Hill et al., 1986; Arechavaleta-Medina and Snyder, 1981; Meyer et al., 2007). One research group found that when seed coats had been removed from seeds they absorbed water twice as fast (Meyer et al., 2007). Therefore, if cultivars grown in full season tests develop more seed coat deficiencies they would absorb water faster than cultivars grown in double crop test that produced more intact seed coats. Also, a recent report by Ma et al. (2004) found a significant correlation between micro scale cracks in the cuticle covering the palisade layer and permeable seed coats in soybean. These findings have not been confirmed and the exact cause for the micro scale cracks is unknown but speculation that cuticle strength influences crack formation is reported (Ma et al., 2004). If micro scale cracks in a seemingly intact seed coat account for the difference seen between permeable and impermeable seeds it could also explain the increase in seed coat deficiency counts and rate of water absorption for cultivars grown full season vs. double crop, if planting system affects seed coat development. Research addressing environmental influence on seed coat integrity and permeability concluded that environmental conditions such as heat and water stress and anthropogenic factors such as harvesting can adversely affect seed coat integrity (Burchett et al., 1985; Nooden et al., 1985; Hill et al., 1986). These research studies were interested in environmental effects on seed coat development and their concomitant influence on seed germination so the data is different than that collected in the current study. Direct comparison of results cannot be made, but they do provide evidence that environmental conditions can alter seed coat development. Overall, cultivar ranking was consistent between planting systems, indicating stable performance.

The analysis for rate of water absorption with model 2 indicated significant genotype × environment and genotype × year interactions (Table 24). The genotype × environment interaction plot displays multiple cross over events that alter performance rankings and confound genotype characterization for rate of water absorption (Figure 9). Out of the top four performing natto cultivars for rate of water absorption, only MFS-511 had an environmental variance estimate less than the mean indicating stability. Two other natto cultivars, MFS-591 and Camp had moderate rate of water absorption estimates but stable environmental variance estimates. This indicates that the trait is influenced by environmental factors and mean performance alone may not identify superior natto cultivars.

Cultivar performance at Tidewater double crop appears to account for a lot of the differential response and cross over. It is difficult to identify what environmental aspect of Tidewater elicits different response for rate of water absorption. The Tidewater double crop test also produced seeds with higher seed coat deficiency counts as presented in the last section, but for rate of water absorption the location did not cause a general shift in response but rather a lot of rank change. The one conclusion that can be drawn is that the genotypes tested in the current study did not respond similarly between Tidewater double crop and the other environments for seed coat deficiency and rate of water absorption.

Year was a non-significant factor for rate of water absorption indicating the distinct conditions between the 2006 and 2007 growing seasons had little influence on the trait. The genotype × year interaction was significant, however, and indicates differential response between the two years. Examining the genotype × year interaction plot indicates the most significant genotype rank change was for natto cultivar MFS-591 that changed rank from third to seventh between the two years. The relative performance of

the other genotypes was more consistent between the 2006 and 2007 season. The significant environment and year influence indicate collecting rate of water absorption data over multiple locations and years is necessary to accurately define a cultivar's rate of water absorption. Genotype did control trait expression for rate of water absorption but environmental conditions influenced this trait to a greater extent than the other natto quality traits tested. Also, little information is available about the trait and collecting additional data will help define its role in natto quality.

Rate of water absorption was adapted from Cober et al. (2006), whose findings were similar to the experiment presented here, with genotype, genotype × year, and location × year contributing significant variation. The researchers concluded the GEI's were less important than the main effects genotype and location. Similar conclusions were reached for rate of water absorption estimates in the current study but environment and year appear to have played a more significant role for the genotypes studied in this research. The different results illustrate that genotype × environmental interaction studies are dependant on the conditions and germplasm tested (Griffiths et al., 2005). Overall genotype does appear to be the most significant controlling factor for the rate of water absorption response but multi environment testing is appropriate to more accurately define cultivar performance.

Seed Size

Seed size analysis with model 1 shows a significant genotype, planting system and genotype × planting system interaction (Table 14). The log seed size interaction plot indicates that cultivars responded differently to the planting systems, but no significant cross over interaction is present to confound selection (Figure 5). The interaction plot is

for log transformed data, but another interaction plot of the untransformed data looked identical (not shown). Planting system did have a statistically significant effect on seed size, and cultivars grown double crop had a seed size estimate 0.51 g per 100 seed lower than the full season planting. This is likely due to seed fill occurring later in the season when water can be a limiting factor for double crop tests. Small seed size is desirable for natto production but it is not possible to conclude the reduction in seed size reported in the current study would enhance natto quality. Genotype more than planting system controlled seed size response, and cultivar selection within either planting system should predict response equally well because genotype rankings were similar between the two systems.

The analysis of seed size with model 2 indicates significant genotype × environment interaction. The genotype × environment interaction plot does not indicate significant cross over, and cultivars remained in tight seed size groupings displaying little differential response across environments (Figure 10). The stability statistics for seed size are low but this is also reflective of the fact that the values are relatively small. Genotype was more important than environment for controlling seed size and environmental variation was minimal. Seed size heritability estimates have been moderate to high and support that genotype is the most important factor (Burton, 1987; Cober et al., 1997b).

Year and genotype × year were also significant sources of variation for seed size (Table 26). The genotype × year interaction plot displayed no cross over interaction and all cultivars followed the same trend between growing seasons. For the main effect year, the difference between the means for the two growing seasons was 0.43 g per 100 seed. The 2007 growing season produced larger seeds which is surprising because of the

drought conditions experienced at all locations. Cober et al. (1997a) reported significant genotype × location and genotype × year interactions for seed size but concluded that genotype was the most significant source of variation. The findings reported in the current study also indicate that environmental influence was marginal and breeding for consistent soybean size is achievable.

Trait Correlation and Stability

Many natto quality traits were significantly correlated but all of the correlation coefficients were fairly small (Table 28). Special care was taken to equilibrate the internal moisture of seed samples to reduce experimental error because all assays involved water absorption. Two significant negative correlations were found for initial moisture with water absorption and water loss (Table 28). The small correlation between initial moisture and water absorption indicates that storing seed samples in the growth chamber helped eliminate internal moisture differences between seed samples. The negative correlation between initial moisture and water loss after steaming is unimportant because the trait showed little genetic variation for the material tested. There were significant positive correlations for seed size with yield, water absorption, seed coat deficiency and a significant negative correlation with water loss after steaming. The largest correlation coefficient between seed size and one of the listed traits is 0.262 which indicates a weak correlation. The positive correlation between seed size and water absorption indicates that larger seeds can hold more water which is expected. The positive correlation between seed size and seed coat deficiency indicates that smaller seeded soybeans should be more resistant to damage. The two natto quality traits with the highest correlation coefficient are seed coat deficiency and rate of water absorption.

The positive correlation supports the hypothesis that the two traits are related and controlled by a similar seed coat integrity described in Ma et al. (2004). It must be noted that all correlation coefficients are relatively low and that correlation does not imply a cause and effect relationships between two traits. The results of the current study do not indicate that indirect selection will impede selecting superior natto cultivars. Additional information over locations and years is needed to confirm the significance of the correlations.

The stability statistics provided additional information beyond means to aid in cultivar evaluation. For yield, the ecovalence stability estimate is more appropriate for identifying stable cultivars. This is because the ecovalence stability statistic accounts for environmental mean and penalizes a cultivar less if its performance is related to environmental potential. The two highest yielding natto cultivars had ecovalence stability estimates below the mean indicating they were stable. The environmental variance estimate was more appropriate for analyzing the results of the natto quality traits because it measures homeostasis which is desirable for quality traits. The environmental variance estimates verified the data seen in the genotype × environment interaction plot for water absorption (Figure 6) but not for seed coat deficiency (Figure 8). As mentioned, the seed coat deficiency environmental variance estimates are likely skewed due to the results of the Tidewater double crop test. The interaction plots for water loss after steaming (Figure 7) and rate of water absorption (Figure 9) display a lot of cross over events making it difficult to determine which cultivars were consistent. For these traits the environmental stability statistics were useful for identifying stable genotypes. For rate of water absorption, the most stable cultivars did not produce the most desirable responses. In this instance a trade off must be made between trait stability and

performance because consistent trait expression is as important as overall trait expression for natto quality.

The stability statistics have limitations, however, and incorporating them with other information is necessary. Two situations arose in which stability estimates proved inadequate and both are represented in the seed coat deficiency response data. Natto genotypes V01-4937, Camp, and MFS-553 all had high (unstable) environmental stability estimates but the interaction plot indicates these cultivars were consistent performers at five of the six growing environments (Table 29 and Figure 8). These cultivars produced more deficient seeds in the Tidewater double crop test than at the other locations, and if that location were removed these cultivars likely would have lower environmental stability estimates. The point of the environmental stability statistics is to identify cultivars with homeostatic response so it is inappropriate to remove environments despite the statistic appearing skewed. All cultivars had high seed coat deficiency counts at Tidewater but the environmental stability estimate for genotypes with the most desirable trait expression were most affected. This relates to the other problem with the environmental stability statistic: Cultivars with the highest stability are sometimes the worst performers. Natto cultivar MFS-511 had the lowest (most stable) environmental variance stability estimate but also had the second highest mean seed coat deficiency count. This illustrates that cultivar selection should not rely on stability analysis alone. Overall, integrating the stability statistics into selection should help identify superior natto cultivars but stability estimates must be interpreted properly and should not be taken as the absolute truth.

Selecting Superior Natto Cultivars

The four natto quality tests analyzed in this experiment were designed to evaluate important traits for natto cultivars. Implementing these tests to aid in selection will help the breeding program develop superior natto cultivars. Water loss after steaming appears to have generated the least informative data. There was little genetic variation for the trait and unless new material is integrated into the germplasm the test is non-informative.

Rate of water absorption is difficult to analyze because it is not clear if it impacts natto quality or what genetic differences are significant. Additional information from natto manufacturers is needed to answer these questions. Rate of water absorption is independent of total water absorption and combining the two traits through breeding and selection could produce a more desirable natto cultivar. The trait displayed a lot of environmental variation indicating progress through selection will be difficult. The data were easy to collect and required minimal training, but diligence is needed to ensure data are collection at specified times. It is necessary to continue collecting rate of water absorption data to further characterize its usefulness in identifying superior natto cultivars. Until the relationship between rate of water absorption and natto quality is more clearly defined, it should not be used as a defining criterion in selection but serve a supplementary role.

The two most useful tests for identifying superior natto cultivars are total water absorption and seed coat deficiency count. Total water absorption data were easy to collect, the assay requires minimal time, is reproducible, and it is an established natto quality trait. Also, the data indicates genetic variation exists to advance the trait beyond the natto quality standard MFS-591. It is conceivable that total water absorption has a maximum achievable value based on seed size but the data indicate that cultivars with

seed size <11 g 100 seed^{-1} can absorb greater than 120% their weight in water which is sufficient for natto production. Genotype × environment interaction was present total water absorption but genotype was the most significant controlling factor, especially for the smaller seeded cultivars (Figure 7). Therefore, selection should identify genotypes with high total water absorption and help select superior natto cultivars.

Seed coat deficiency should also aid in selecting superior natto cultivars. The assay attempts to measure seed coat durability because cracked seed coats are undesirable in natto production. Ma et al. (2004) reported a significant correlation in soybean between cracks in the seed coat cuticle and seed permeability. It is possible the seed coat deficiency assay provides an inexpensive, rapid assay to detect these cracks and identify soybeans with high permeability. The correlation between seed coat deficiency and rate of water absorption supports this claim because decoated soybean seeds absorb water twice as fast as seeds with intact seed coats (Meyer et al., 2007). Compromised seed coats are deleterious to natto quality, but the cracks reported by Ma et al. (2004) were in seemingly intact seed coats. This suggests a mid seed coat deficiency count should be selected to identify cultivars with strong and permeable seed coats. High seed coat deficiency counts are still undesirable, but moderate counts may allow for acceptable seed coat quality and allow for sufficient water absorption. This relationship has not been tested and requires input from the natto producers. The data were more difficult to collect than the other traits, and there is possibility for sampling error when collecting seed coat deficiency counts because the test is subjective. Proper instruction and practice is sufficient for training, but having one person collect the data each year should improve consistency and help avoid improper phenotyping. The trait also has an inherent flaw in that the data is subject to differences among harvest conditions. These differences can be

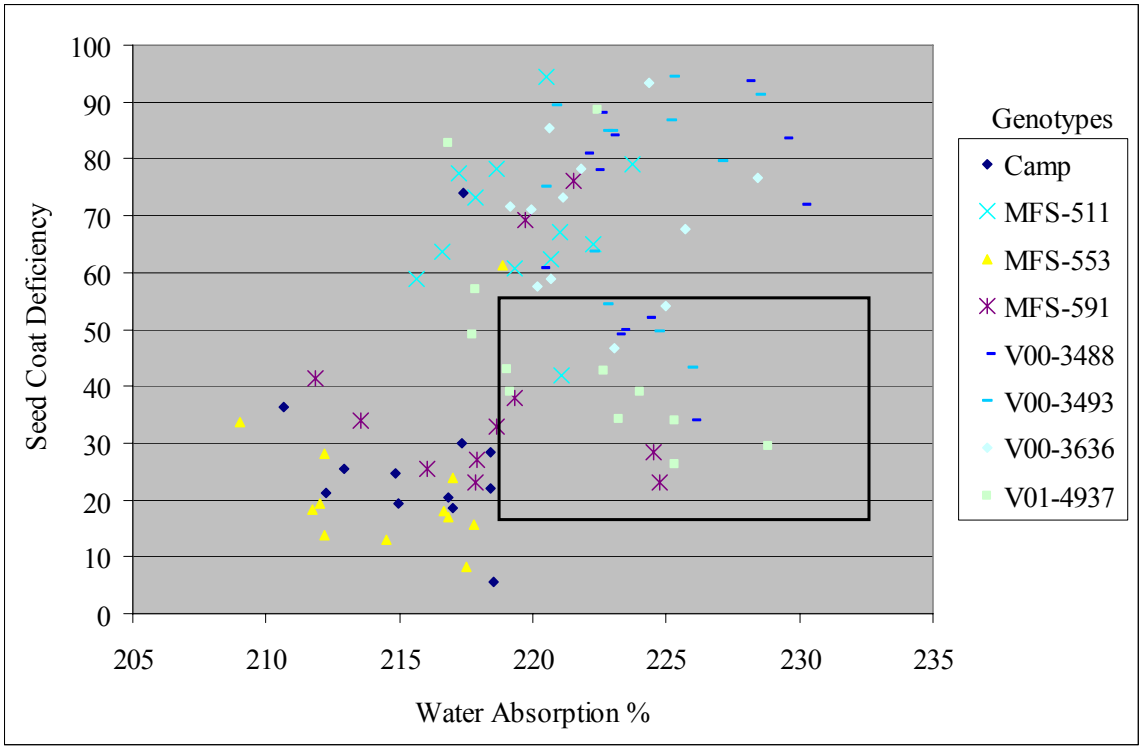
identified by analyzing the results at locations to determine if one location produced atypical results. Defining the appropriate harvest conditions will also help minimize variation and produce accurate data. Seed coat deficiency data were consistent across most environments with a few exceptions. The limited number of years represented in this experiment necessitates collecting data at multiple locations in future years to ensure the results seen in this experiment are repeatable.

Seed coat deficiency and water absorption encompass the two main components of initial natto production therefore coupling the traits for selection will aid in selecting superior natto cultivars. Figure 11 depicts data collected in this experiment averaged over locations and years and identifies a zone that encompasses the most desirable natto quality attributes for the two traits. Seed coat deficiency criteria could become less stringent and acceptable limits may move closer to 60 if the trait is detecting micro scale cracks along with seed coat deficiencies. Using information from Figure 11, seed size data, least square mean rankings, and interaction plots indicates that V01-4937 is a promising natto cultivar. It is higher yielding and absorbs more water than MFS-591 and has a moderate seed deficiency count and rate of water absorption. V01-4937 did not have good environmental stability estimates for total water absorption and seed coat deficiency, but figure 8 indicates that V01-4937 performed well at all locations except Tidewater double crop for seed coat deficiency.

Figure 11 also illustrates an interesting historical perspective on natto cultivar development. Camp was one of the first natto genotypes developed at Virginia Tech. It has a desirable seed coat deficiency count but it does not have high total water absorption. Japanese natto manufacturers characterized Camp as too hard because it did not absorb enough water, and as a result breeding efforts focused on increasing the total

water absorption of subsequent natto genotypes. The recent natto genotypes developed at Virginia Tech such as V00-3488, V00-3493, V00-3636, and V01-4937 have higher total water absorption, but none of these genotypes have seed coat deficiency scores as low as Camp. This indicates that breeding for total water absorption compromised the seed coat deficiency trait and that a trade off between trait values must be accepted. It may not be possible to improve a trait ad infinitum without negatively affecting a different trait. This situation provides evidence that both traits need to be simultaneously selected for to ensure the developing germplasm has desirable expression for both traits. It will be interesting to monitor this potential relationship as natto breeding efforts continue.

Figure 11. Seed coat deficiency vs. water absorption data from 8 natto soybean genotypes grown in 6 environments during 2006 and 2007. The observations in the box display desirable performance for both natto quality traits.



Summary

Genotype × planting system interaction was significant for all traits except water loss after steaming. The interaction did not create significant rank change for any trait, so the need to conduct specific breeding for different planting systems is not necessary. The results suggest planting system significantly influences seed coat deficiency counts and the rate of water absorption, therefore, planting soybeans in a double crop system may produce superior natto quality. These traits require additional testing in more locations and environments because significant differential response was observed. The interaction did not hinder selecting the best performing cultivars, but it suggests that natto quality gains through double crop planting may be cultivar specific. Validation that double crop plantings produced soybeans with significantly better natto characteristics requires evaluation by natto producers. It is impractical to perform these assessments in the breeding program for technical reasons so the samples would need to be shipped to natto manufacturers and tested. Without this information conclusions cannot be verified and remain speculation.

Significant genotype × environment interactions were observed for all traits studied but genotype played the most significant role in determining trait expression. This conclusion is in accordance with other research on natto quality and environmental influence (Taira, 1990; Cober et al., 1997a; Geater et al., 2000). It is important to characterize genotypes for these traits before they are implemented into the program. A majority of the interactions did not result in significant rank change but environment did influence cultivar response for yield, seed coat deficiency, and rate of water absorption.

All traits except water loss after steaming displayed a significant genotype × year interaction but the interaction did not result in a confounding cross over effect. Genotype

responses between years followed the same trend and were more important in determining trait expression. The significant influence of year on most responses strongly suggests that breeding for trait stability should be a major objective for selecting superior natto cultivars.

LITERATURE CITED

Arechavaleta-Medina, F. and H.E. Snyder. 1981. Water imbibition by normal and hard soybeans. J. Am. Oil Chem. Soc. 58: 976.

Ash and Dohlman, 2007. Soybeans and oil crops: Market outlook. USDA. http://www.ers.usda.gov/Briefing/SoybeansOilcrops/2007baseline.htm

Becker, H.C., and J.Leon. 1988. Stability analysis in plant breeding. Plant breeding 101: 1-23.

Burchett, C.A., W.T. Schapaugh, Jr., C.B. Overley, and T.L. Walter. 1985. Influence of etched seed coats and environmental conditions on soybean seed quality. Crop Sci. 25: 655-660.

Burton, J.W. 1987. Quantitative genetics: Results relevant to soybean breeding. p. 211-247. *In* J.R. Wilcox (ed.) Soybeans: improvement, production, and uses. 2nd ed. Agron. Monogr. 16. ASA, CSSA, SSSA, Madison, WI.

Carlson J.B., and N.R. Lersten. 2004. Reproductive Morphology. p. 59-93. *In* H.R. Boerma and J.E. Specht (ed.) Soybeans: improvement, production, and uses. 3rd ed. Agron. Monogr. 16. ASA, CSSA, SSSA, Madison, WI.

Cober E.R., J.A. Fregeau-Reid, L.N. Pietzrak, A.R. McElroy, and H.D. Voldeng. 1997a. Genotype and environment effects on natto soybean quality traits. Crop Sci. 37:1151-1154.

Cober E.R., H.D. Voldeng, and J.A. Fregeau-Reid. 1997b. Heritability of seed shape and seed size in soybean. Crop Sci. 37:1767-1769.

Cober E.R., J.A. Fregeau-Reid, G. Butler, and H.D. Voldeng. 2006. Genotype-environment analysis of parameters describing water uptake in natto soybean. Crop Sci. 46:2415-2419.

Cui, Z., A.T. James, S. Miyazaki, R.F. Wilson, and T.E. Carter Jr. 2005. Breeding specialty soybeans for traditional and new soyfoods. p. 264-322. *In* K. Liu (ed.) Soybeans as functional foods and ingredients. AOCS Press.

Dickson, M.H. and M.A. Boettger. 1977. Inheritance of resistance to mechanical damage and transverse cotyledon cracking in snap beans (*Phaseolus vulgaris* L.). J. Amer. Soc. Hort. Sci. 102:498-501.

Dong, Y.S., L.M. Zhao, B. Liu, Z.W. Wang, Z.Q. Jin, and H. Sun. 2004. The genetic diversity of cultivated soybean grown in China. Theor. Appl. Genet. 108:931-936.

Durham, D. 2003. The United Soybean Board's better bean initiative: Building United States soybean competitiveness from the inside out. AgBioForum, 6: 23-26.

Elliot, R.J. 2000. Learning SAS in the computer lab. 2nd ed. Druxbury Press.

Fehr, W.R. 1987. Principles of Cultivar Development: Theory and Technique. Volume 1. Macmillan Publishing Company.

Fehr, W.R., and C.E. Caviness. 1977. Stages of soybean development. Iowa agriculture and home economics experiment station special report. 80: 3-11.

Geater, C.W., W.R. Fehr, and L.A. Wilson., 2000. Association of soybean seed traits with physical properties of natto. Crop Sci. 40:1529–1534.

Golbitz, P. 1995. Traditional soyfoods: processing and products. Journal of Nutrition 125: S570-S572.

Griffiths, A., S. R. Wessler, R.C. Lewontin, W.M. Gelbart, D.T. Suzuki, and J.H. Miller. 2005. Introduction to genetic analysis. 8th Ed. W.H. Freeman and Company.

Hill, H.J., S.H. West, and K. Hinson. 1986. Effects of water stress during seedfill on impermeable seed expression in soybean. Crop Sci. 26:807-812.

Holshouser, D.L. and W.C. Alexander. 2001. Suggested soybean seeding rates. Virginia Crop Exstension Publications 424-102.

Hsu, K.H., C.J. Kim, and L.A. Wilson. 1983. Factors affecting water uptake of soybean during soaking. Amer. Soc. of Cereal Chem. 60: 208-211.

Hymowitz, T., F.I. Collins, J. Panczner, and W.M. Walker. 1972. Relationships between the content of oil, protein, and sugar in soybean seed. Agron. J. 64:613-616.

Kang, M.S. 1998. Using genotype-by-environment interaction for crop cultivar development. Advances in Agronomy 62: 199-252.

Lemelin, E., G. Branlard, L. Salvo, V. Lein, T. Aussenac, and J. Dayde. 2005. Breadmaking stability of wheat flour: relation between mixing properties and molecular weight distribution of polymeric glutenins. Jour. of Cereal Sci. 42: 317-326.

Lin, C.S., M.R. Binns, and L.P. Lefkovitch. 1986. Stability analysis: where do we stand? Crop Sci. 26: 894-899.

Liu, K. 1997. Soybeans: chemisty, technology, and utilization. Chapman & Hall, New York, NY.

Liu, K. 2000. Expanding soybean food utilization. Food Technology 54: 7, 46-48, 50, 52, 54, 56, 58.

Liu, K. 2005. Edible soybean products in the current market. p.23-51. *In* K. Liu (ed.) Soybeans as functional foods and ingredients. AOCS Press.

Lomax,R.G. 2001. Statistical concepts: a second course for education and behavioral sciences. 2nd ed. Lawrence Erlbaum Associates.

Lusas, E.W. 2004. Soybean processing and utilization. p. 949-1045. *In* H.R. Boerma and J.E. Specht (ed.) Soybeans: improvement, production, and uses. 3rd ed. Agron. Monogr. 16. ASA, CSSA, SSSA, Madison, WI.

Ma, F., E. Cholewa, T. Mohamed, C.A. Peterson, and M. Gijzen. 2004. Cracks in the palisade cuticle of soybean seed coats correlate with their permeability to water. Annals of Botany 94:213-228.

McCue, P., and K. Shetty. 2004. Health Benefits of Soy Isoflavonoids and Strategies for Enhancement: A Review. Crit. Rev. Food Sci. Nutr. 44: 361-367.

McDonald, M.B., C.W. Vertucci, and E.E. Roos. 1988. Seed coat regulation of soybean seed imbibition. Crop Sci. 28: 987-992.

Meyer, C.J., E. Steudle, C.A. Peterson. 2007. Patterns and kinetics of water uptake by soybean seeds. J. of Exp. Bot. 58: 717-732.

Mullin, W.J., and W. Xu. 2001. Study of soybean seed coat components and their relationship to water absorption. J. Agric. Food Chem. 49:5331-5335.

Mwandemele, O.D., K.S. McWhirter, and C. Chesterman. 1984. Genetic variation in soybean (*Glycine max* L. Merril) for cookability and water absorption during cooking. Euphytica 33: 859-864.

Nakamura, T., D. Yang, S. Kalaiselvi, Y. Uematsu, and R. Takahashi. 2003. Genetic analysis of net-like cracking in soybean seed coats. Euphytica 133:179-184.

Nooden, L.D., K.A. Blakley, and J.M. Grzybowski. 1985. Control of seed coat thickness and permeability in soybean: a possible adaptation to stress. Plant Physiol. 79:543-545.

Okabe, A. 1996. Inheritance of seed coat cracking and effective selection method for resistance in soybean. Japan Agricultural Research Quarterly 30:15-20.

Ott, R.L. and M. Longnecker. 2001. An introduction to statistical methods and data analysis. 5th ed. Wadsworth Group.

Pan, Z., and W. Tangratanavalee. 2003. Characteristics of soybean as affected by soaking conditions. Lebensm.-Wiss. U.-Technol. 36:143-151.

Poehlman, J.M., D.A. Sleper. 1995. Breeding field crops. 4th ed. Iowa Stare University Press.

Potts, H.C., J. Duangpatra, W.G. Hairston, and J.C. Delouche. 1978. Some influences of hardseededness on soybean seed quality. Crop Sci. 18:221-224.

Rangaswamy, N.S., and L. Nandakumar. 1985. Correlative studies on seed coat structure, chemical composition, and impermeability in the legume *Rhynchosia minima*. Botanical Gazette 146: 501-509.

Reichert, R.D., and A.O.F. Ehiwe. 1987. Variability, heritability, and physiochemical studies of seed coat durability in field pea. Can. J. Plant Sci. 67:667-674.

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

Robert, N. 2002. Comparison of stability statistics for yield and quality traits in bread wheat. Euphytica 128: 333-341.

Rolston, M.P. 1978. Water impermeable seed dormancy. Botanical Review 44:365-396.

Saio, K. 1976. Soybean resistance to water absorption. Cereal Foods World. 21:168-173.

SAS Institute. 2002. SAS for windows 9.1. SAS Inst., Cary, NC.

Southeast Regional Climate Center (SERCC). University of North Carolina at Chapel Hill. http://www.sercc.com/climateinfo/historical/historical_va.html

Shao, S., C.J. Meyer, F. Ma, C.A. Peterson, and M.A. Bernards. 2007. The outermost cuticle of soybean seeds: chemical composition and function during imbibition. Jour. of Exp. Botany. 58: 1071-1082.

The American Soybean Association. 2007. Soystats 2007. http://www.soystats.com/2007/Default-frames.htm

Taira, H., N. Suzuki, C. Tsukamto, Y. Kainuma, H. Tanaka, and M. Saito. 1987. Suitability for natto processing of small seed cultivars and quality of the natto. Quality of soybean seeds grown in Japan. Report National Food Research Institute 51:48–58.

Taira, H. 1990. Quality of soybeans for processed foods in Japan. Japan Agric. Res. Quarterly 24:224–230.

USDA. 2008. National Agriculture Statistics Service (NAAS). http://www.nass.usda.gov/QuickStats/index2.jsp

Wei, Q., C. Wolf-Hall, K.C. and Chang. 2001. Natto characteristics as affected by steaming time, *Bacillus* strain, and fermentation time. J. Food Sci. 66:167–173.

Wei, Q., and K.C. Chang. 2004. Characteristics of fermented natto products as affected by soybean cultivars. J. of Food Proc. Pres. 28:251-273.

Wilcox, J.R., and R.M. Shibles. 2001. Interrelationship among seed quality attributes in soybean. Crop Sci. 41:11-14.

Wilson, L.A. 1995. Soy foods. Practical handbook of soybean processing and utilization. AOCS Press, Champaign, IL, and United Soybean Board, St. Louis. MO. 428-459.

Yue, G.L., K.L. Roozenboom, W.T. Schapaugh Jr., and G.H. Liang. 1997. Evaluation of soybean cultivars using parametric and nonparametric stability estimates. Plant Breeding 116:271-275.

Appendix

Appendix Table A. Weather summary at the locations planted during the 2006 growing season.

Month	Avg. Max. Temp. °C	Deviation from 50+ Yr Avg Max Temp. °C	Avg. Min Temp. °C	Deviation from 50+ Yr Avg Min Temp. °C	Avg. Precip. cm	Deviation from 50+ Yr Avg. Precip. cm
Blacksburg						
May	21.04	-1.25	6.99	-1.56	5.46	-4.55
June	25.79	-0.20	13.44	0.44	29.97	20.47
July	28.14	0.03	15.98	0.55	8.86	-1.19
August	28.48	0.82	16.38	1.57	5.69	-3.12
Sepetember	21.57	-2.75	10.50	-0.28	10.13	1.55
October	16.81	-1.79	3.69	-0.46	12.98	5.21
November	14.56	1.71	0.33	0.77	9.32	2.13
Total	156.39	-3.43	67.32	1.04	82.42	20.50
Tidewater						
May	25.16	-0.31	10.95		2.86	-1.10
June	29.30	-0.16	17.78	-6.08	25.60	15.06
July	31.77	0.41	20.18	-1.79	9.30	-4.42
August	32.90	2.26	20.13	1.42	6.35	-8.08
September	26.70	-0.85	14.74	4.92	23.27	11.66
October	21.43	-0.68	8.84	6.11	24.69	14.99
November	17.89	0.68	6.28	4.97	18.64	10.39
Total	185.16	1.36	98.89	9.55	115.11	36.80
Mt. Holly & Warsaw						
May	25.48	0.17	11.88		2.80	-1.25
June	29.29	-0.20	17.21	0.23	18.03	8.86
July	31.84	0.28	21.02	1.52	10.34	-1.22
August	32.53	1.86	20.50	1.74	6.12	-5.18
September	25.46	-1.84	14.85	-0.06	29.87	19.10
October	20.59	-0.84	8.51	-0.09	17.58	8.86
November	16.44	0.63	5.72	1.75	14.55	6.32
Total	181.64	0.05	99.69	5.08	103.61	33.58

Source: Southeast Regional Climate Center

Appendix Table B. Weather summary at the locations planted during the 2007 growing season.

Month	Avg. Max. Temp. °C	Deviation from 50+ Yr Avg Max Temp. °C	Avg. Min Temp. °C	Deviation from 50+ Yr Avg Min Temp. °C	Precip. Avg inches	Deviation from 50+ Yr Avg. Precip inches
Blacksburg						
May	23.86	1.57	9.08	0.53	5.51	-4.50
June	26.59	0.60	13.83	0.83	9.47	-0.03
July	27.01	-1.10	14.64	-0.79	7.24	-2.82
August	31.54	3.88	17.03	2.22	4.24	-4.57
Sepetember	26.86	2.53	11.32	0.54	2.67	-5.92
October	27.44	8.84	11.11	6.96	0.08	-7.70
November	N/A	N/A	N/A	N/A	N/A	N/A
Total	163.29	16.32	77.02	10.29	29.21	-25.53
Tidewater						
May	25.82	0.36	12.65	0.22	5.49	-4.57
June	30.54	1.08	17.17	0.13	7.62	-2.92
July	30.92	-0.44	18.62	-0.95	4.34	-9.37
August	33.53	2.89	19.43	0.67	12.70	-1.73
September	N/A	N/A	N/A	N/A	N/A	N/A
October	N/A	N/A	N/A	N/A	N/A	N/A
November	N/A	N/A	N/A	N/A	N/A	N/A
Total	120.80	3.88	67.87	0.07	30.15	-18.59
Mt. Holly & Warsaw						
May	26.70	1.39	12.92	0.65	3.10	-7.19
June	30.79	1.30	18.21	1.23	6.81	-2.36
July	32.47	0.91	19.53	0.03	4.98	-6.58
August	32.56	1.89	20.75	1.99	10.03	-1.27
September	N/A	N/A	N/A	N/A	N/A	N/A
October	N/A	N/A	N/A	N/A	N/A	N/A
November	N/A	N/A	N/A	N/A	N/A	N/A
Total	122.53	5.49	71.41	3.90	24.92	-17.40

Source: Southeast Regional Climate Center