Genetic Improvement of U.S. Soybean in Maturity Groups II, III, and IV


ABSTRACT
Soybean improvement via plant breeding has been critical for the success of the crop. The objective of this study was to quantify genetic change in yield and other traits that occurred during the past 80 yr of North American soybean breeding in Maturity Groups (MGs) II, III, and IV. Historic sets of 60 MG II, 59 MG III, and 49 MG IV soybean cultivars, released from 1923 to 2008, were evaluated in field trials conducted in 17 U.S. states and one Canadian province during 2010 to 2011. Averaged over 27 MG II and MG IV and 26 MG III site-years of data, the estimated rates of yield improvement during the 80 yr were 23 kg ha⁻¹ yr⁻¹ for MGs II and III, and 20 kg ha⁻¹ yr⁻¹ for MG IV cultivars. However, a two-segment linear regression model provided a better fit to the data and indicated that the average current rate of genetic yield gain across MGs is 29 kg ha⁻¹ yr⁻¹. Modern cultivars yielded more than old cultivars in all environments, but particularly in high-yielding environments. New cultivars in the historic sets used in this study are shorter in height, mature later, lodge less, and have seeds with less protein and greater oil concentration. Given that on-farm soybean yields in the United States are also increasing at a rate of 29 kg ha⁻¹ yr⁻¹, it can be inferred that continual release of greater-yielding cultivars has been a substantive driver of the U.S. on-farm realized yield increases.

K. Rincker, C. Fox, T. Cary, and B. Diers, Dep. of Crop Sciences, Univ. of Illinois, Urbana, IL 61801; R. Nelson, USDA-ARS, Soybean/Maize Germplasm, Pathology, and Genetics Research Unit, and Dep. of Crop Sciences, Univ. of Illinois, Urbana, IL 61801; J.E. Specht and G. Graef, Dep. of Agronomy and Horticulture, Univ. of Nebraska, Lincoln, NE 68583; D. Sleper and G. Shannon, Division of Plant Sciences, Univ. of Missouri, Columbia, MO 65211; S. Cianzio, Dep. of Agronomy, Iowa State Univ., Ames, IA 50011; S. Casteel, Dep. of Agronomy, Purdue, Univ., West Lafayette, IN 47907; S. Conley and V. Davis, Dep. of Agronomy, Univ. of Wisconsin, Madison, WI 53706; P. Chen, Crop, Soil, and Environmental Sciences, Univ. of Arkansas, Fayetteville, AR 72701; C. Godsey, Dep. of Plant and Soil Sciences, Oklahoma State Univ., Stillwater, OK 74078; D. Holshouser, Tidewater Agricultural Research and Extension Center, Virginia Tech, Suffolk, VA 23437; G. Jiang, Dep. of Plant Sciences, South Dakota State Univ., Brookings, SD 57007; S. Kantartzi, Plant, Soil and Agricultural Systems, Southern Illinois Univ., Carbondale, IL 62901; W. Kenworthy, Plant Science and Landscape Architecture, Univ. of Maryland, College Park, MD 20742; C. Lee, Plant and Soil Sciences, Univ. of Kentucky, Lexington, KY 40546; R. Mian, USDA-ARS, and Dep. of Horticulture and Crop Science, The Ohio State Univ., Wooster, OH 44691; L. McHale, Dep. of Horticulture and Crop Science, The Ohio State Univ., Wooster, OH 44691; S. Naeve and J. Orf, Dep. of Agronomy and Plant Genetics, Univ. of Minnesota, St. Paul, MN 55108; V. Poysa, Agriculture and Agri-Food Canada, Harrow, ON N0R 1G0 Canada; W. Schapaugh, Dep. of Agronomy, Kansas State Univ., Manhattan, KS 66506; R. Uniatowski, Plant and Soil Sciences, University of Delaware, Newark, DE 19716; D. Wang, Dep. of Plant, Soil and Microbial Sciences, Michigan State Univ., East Lansing, MI 48824. Received 7 Oct. 2013. *Corresponding author (bdiers@illinois.edu).

Abbreviations: AIC, Akaike information criterion; BLUP, best linear unbiased predictor; MG, maturity group; PI, plant introduction; PVP, plant variety protection.
SOYBEAN [Glycine max (L.) Merr.] in the United States has evolved from a minor forage crop in the early 1900s to a crop that yielded 82.1 million metric tons in 2012 (USDA-NASS, 2013). Soybean is currently a leading source of protein and oil for human food, animal feed, and industrial products (Wilson, 2008). The global rate of yield increase in soybean will need to almost double to meet the population demands predicted for 2050 (Ray et al., 2013).

On-farm yield gains arise from the combined impact of grower adoption of new cultivars, improved cultural practices, interactions between new cultivars and improved cultural practices, and environmental factors such as increased atmospheric CO$_2$ levels (Long et al., 2006; Rowntree et al., 2013, 2014; Specht et al., 2014; Ziska and Bunce, 2007). Average on-farm soybean seed yields in the United States have increased from 738 kg ha$^{-1}$ in 1924 to 2658 kg ha$^{-1}$ in 2012 (USDA-NASS, 2013). Specht et al. (2014) recently reviewed the U.S. national estimated yields compiled by NASS from 1924 to 2012 and calculated a 23 kg ha$^{-1}$ yr$^{-1}$ on-farm yield gain rate with a simple linear regression model. The authors also noted that a two-segment linear model had a better fit to the data, which showed an increase of 22 kg ha$^{-1}$ yr$^{-1}$ from 1924 to 1983 and 29 kg ha$^{-1}$ yr$^{-1}$ from 1983 to 2012. Wilcox (2001) estimated yield gains across a 60 yr period by reviewing Uniform Soybean Test data for MG 00 to IV entries, and found yield gains ranging from 22 kg ha$^{-1}$ yr$^{-1}$ for MG 00 to 31 kg ha$^{-1}$ yr$^{-1}$ for MG III.

The development of soybean cultivars available to farmers has changed considerably over the past 80 yr. In the early 1900s, farmers grew either plant introductions (PIs) from East Asia or selections from these PIs, and these were grown primarily as a forage crop until 1941 (Probst and Judd, 1973). The first cultivars developed from breeding programs in North America were released in the 1940s. These cultivars originated from the hybridizations of PIs, which resulted in the development of cultivars better adapted to North American production systems (Hartwig, 1973). As soybean production increased, public and proprietary breeding efforts expanded and began to include selection for pest resistance along with selection for yield (Carter et al., 2004). Proprietary breeding programs have continued to expand in number and in size and these programs currently provide most of the soybean seed sold to farmers in the United States (Specht et al., 2014).

Genetic yield gain was estimated in a number of studies by growing soybean cultivars from different release years in a common environment. Luedders (1977) tested 21 MG I, II, III, and IV soybean cultivars grown in Missouri and reported an overall increase of 708 kg ha$^{-1}$ for cultivars released between 1964 and 1971 compared with PIs grown commercially in the 1920s and 1930s. When the cultivar yield means were regressed on release dates, an annual increase of 16 kg ha$^{-1}$ was obtained. Wilcox et al. (1979) tested five MG II and five MG III cultivars in the Midwestern United States and reported yield increases of 700 kg ha$^{-1}$ in MG II cultivars and 625 kg ha$^{-1}$ in MG III cultivars released in the 1970s compared with cultivars released ~1920. When regression analysis was done with these cultivar means, rates of 11 kg ha$^{-1}$ yr$^{-1}$ for MG II and 13 kg ha$^{-1}$ yr$^{-1}$ for MG III were obtained. In Georgia, Boerma (1979) calculated a genetic gain in yield of 14 kg ha$^{-1}$ yr$^{-1}$ for 18 MG VI, VII, and VIII cultivars released between 1942 and 1973. Specht and Williams (1984) evaluated 240 MG 00 to IV cultivars in Nebraska that were introduced or released between 1902 and 1977 and estimated an average genetic gain of 19 kg ha$^{-1}$ yr$^{-1}$. Similarly, Salado-Navarro et al. (1993) observed genetic yield gains of 16 to 19 kg ha$^{-1}$ yr$^{-1}$ in MG VI to VIII cultivars tested in Florida. In Quebec and Ontario, Voldeng et al. (1997) evaluated 41 MG 000 to 0 cultivars that had been released from 1934 to 1992, and found an overall yield increase of 11 kg ha$^{-1}$ yr$^{-1}$. To account for differences in maturity among the tested cultivars, Voldeng et al. (1997) created a maturity-corrected yield index. The regression of this index on year of release resulted in a better fit to a quadratic model, which suggested an accelerating rate of genetic gain over years. In Tennessee, Ustun et al. (2001) reported that soybean yield had increased 14 kg ha$^{-1}$ yr$^{-1}$ in eight MG V to VII cultivars grown in 10 Tennessee environments. Recently, De Bruin and Pedersen (2008) observed an overall genetic gain of 25 kg ha$^{-1}$ yr$^{-1}$ in evaluations of 23 cultivars from MGs I to III when grown in six Iowa environments.

A number of factors can influence yield gains over time. Selection of parents from only elite cultivars and breeding lines has resulted in a narrowing of the genetic base for cultivars released in North America (Gizlice et al., 1994; Sneller, 1994), and this limited genetic diversity could eventually reduce rates of genetic gain. The negative impact on yield of abiotic stresses, such as increased ozone concentration, may continue to increase (Betzelberger et al., 2012). In addition, pests, such as soybean cyst nematode (Heterodera glycines Ichinohe), soybean aphid (Aphis glycines Matsumura), and Phytophthora root rot (Phytophthora sojae Kauffmann and Gerdemann) have added pressure to breeding programs to allocate limited breeding resources to concurrent selection for pest resistance, which can slow genetic gains for yield potential. Incorporation of transgenes for herbicide tolerance could also slow gains for yield by diverting resources from yield-focused breeding efforts. In contrast, increasing atmospheric CO$_2$ levels has and will enhance photosynthesis and water use efficiency in the C3 photosynthetic soybean, which theoretically could result in a yield benefit (Specht et al., 1999; Sinclair et al., 1984). Climate change also will have negative impacts on soybean production because of higher summer temperatures and more extreme weather events (Hassol, 2009); however, warmer springs and falls could have a
positive impact by lengthening the growing season, which could be exploited by earlier planting and the use of cultivars with later maturities (Specht et al., 2014).

Another factor that likely impacted genetic gain was the passage of the Plant Variety Protection (PVP) Act of 1970. This spurred proprietary investment in soybean breeding programs because it gave plant breeders exclusive control over newly developed cultivars, thereby resulting in a greater return on investment (Fehr, 1991; USDA, 2006). Since the passage of the act, Carter et al. (2004) documented an increase in the number of North American soybean breeders, especially in the private sector. In addition, improvements in plot planting, harvesting equipment, and computing capacity enabled increases in field plot testing and improved precision in selection, resulting in greater genetic gains (Eathington et al., 2007). New technologies and methodologies may arise that could further accelerate future yield gains. For example, genomic selection has been proposed to shorten the breeding cycle (Meuwissen et al., 2001; Bernardo, 2010b; Hefner et al., 2009), and DNA and RNA sequencing will increase our understanding of gene function, which can be utilized in the development of future cultivars (Poland and Rife, 2012; Schmutz et al., 2010).

In addition to gains in genetic yield potential, genotype × environment interactions are constant challenges for plant breeders developing new cultivars. New cultivars with high yield potential in a favorable environment must still perform well in high stress environments. Finlay and Wilkinson (1963) developed a method to assess yield stability via regression of a cultivar mean yield with the mean yield of all cultivars in a specific environment. Cultivar regression coefficients greater than one denote less yield stability (i.e., a steeper yield response to increases in environmental productivity), while coefficients less than one denote more yield stability (i.e., a shallower yield response). By using this method, De Bruin and Pedersen (2008), Voldeng et al. (1997), and Wilcox et al. (1979) all determined that soybean yields have increased over generations of breeding without detectable reductions in yield stability.

Although the magnitude of yield improvement over time is of primary interest when developing new soybean cultivars, other traits have changed due to direct or indirect selection. For instance, resistance to lodging is an important trait that breeders must consider when releasing a cultivar. Previous research has shown that lodging resistance has improved over generations of breeding. Luedders (1977) found that lodging scores made using a five-point scale (1 = erect and 5 = prostrate) decreased by 0.9 units when cultivars released between 1964 and 1971 were compared with cultivars derived from selecting within PIs. Specht and Williams (1984) reported a decrease of 1.0 unit between cultivars released over a 75-yr period ending in 1977. Voldeng et al. (1997) also detected a decrease in lodging scores from old to new cultivars in the early MGs, and Wilcox et al. (1979) found a decrease in lodging among MG III cultivars, but not among MG II cultivars.

Although seed protein and oil concentration, mature plant height, seed size, and seed quality are important traits, yield, maturity, and lodging are the primary considerations in most cultivar release decisions. Over generations of breeding, Specht and Williams (1984) noted little change in these secondary traits in cultivars from MG IV or earlier, but did detect a decrease in mature plant height in MG III cultivars and an increase in MG 00 cultivars. Similarly, Wilcox (2001) noted an increase of plant height in MG I cultivars, but a decrease in MG II, III, and IV cultivars across generations. Ustun et al. (2001) reported an overall decrease in plant height when ancestral cultivars from the 1940s were compared with cultivars from the 1970s grown in the mid-southern United States.

Seed protein and oil concentration were found to be related to year of cultivar release by Voldeng et al. (1997), Wilcox (2001), and Ustun et al. (2001). Voldeng et al. (1997) reported that seed protein decreased 4 g kg⁻¹ yr⁻¹, whereas seed oil increased 4 g kg⁻¹ yr⁻¹ over generations. Across 60 yr of the MG 00-IV Uniform Soybean Test, Wilcox (2001) noted that seed protein concentration decreased significantly by 0.29 g kg⁻¹ yr⁻¹ in MG I and by 0.27 g kg⁻¹ yr⁻¹ in MG II. In addition, a significant increase in seed oil concentration of 0.19 g kg⁻¹ yr⁻¹ in MG 00 and a decrease of 0.11 g kg⁻¹ yr⁻¹ in MG III were observed. Ustun et al. (2001) reported an overall decrease in seed protein and an increase in seed oil across generations within their historic set of eight cultivars adapted to the mid-southern United States. These results are consistent with Wilson (2004), who reported a general negative correlation between yield and protein and a positive correlation between yield and oil based on a review of published soybean literature.

The objective of this study was to more comprehensively estimate annual genetic gain for seed yield by using a greater number of cultivars and a longer time frame of release than those used in past studies, with the purpose of updating and improving our understanding of genetic changes resulting from the past 80 yr of North American soybean breeding in MG II, III, and IV. This objective was fulfilled by directly comparing the performance of soybean cultivars released from 1923 to 2008 in 2 yr of field testing conducted in 17 U.S. states and one Canadian province.

### MATERIALS AND METHODS

#### Plant Materials

The experiments included PIs and publicly and proprietarily developed cultivars that were released or introduced between 1923 and 2008. Public cultivars and PIs were selected for inclusion in the MG II, III, or IV tests if they were widely grown by producers after their release. Testing all released cultivars...
was not possible, so for each MG-specific entry list, cultivars were first sorted by decade of release year so that the authors could construct a final entry list of cultivars whose release years were uniformly distributed (to the extent possible) within and across the eight decades of breeding. Seed of public cultivars were obtained from the USDA Soybean Germplasm Collection. Soybean cultivars from the proprietary sector (i.e., Monsanto, Pioneer, and Syngenta) were nominated by the respective company breeders for inclusion in these tests, and seed of the cultivars was provided to the authors by the companies. Although soybean was initially grown as a forage crop, cultivars selected in our study were grown as commercial seed production cultivars (Morse et al., 1949). The cultivars Korean, Mukden, Richland, Dunfield, Illini, A.K. (Harrow), Mandell, Mingo, Macoupin, Scioto, Boone, and Patoka were all developed by selecting within a PI, and served in our analysis as a yield base line before hybridization breeding programs began. Seed increases of all test entries were grown in 2009 to provide seed for the 2010 trials, and seed was produced in 2010 for the 2011 trials. Seed increase sites in both years were Lincoln, NE (MG II); Urbana, IL (MG III); and Columbia, MO (MG IV).

### Experimental Design and Data Collection
Cultivars in each MG-specific set were yield tested during 2010 and 2011 at 13 to 15 locations representative of typical production environments where cultivars of the given MG are grown. Table 1 lists the cultivar names and release years for the 60 MG II, 59 MG III, and 49 MG IV cultivars. The three MG-specific trials were separate experiments, and the experimental design at each site was a randomized complete block with between 2 and 12 replications (Table 2). Years and sites were not a balanced factorial, so the two were combined to be treated as a site-year environmental factor in the data analysis. The number of site-year environments in which data were collected for a given trait is shown in Table 3.

Cultivars were planted in yield plots four rows wide, with a 0.76-m row spacing, and at a rate of approximately 30 seeds m$^{-1}$ of row. The planted seed was not treated with either fungicides or insecticides. Plot length ranged from 2.9 m to 6.1 m, depending on the planting system of the author-collaborator. Seed yields were estimated by harvesting the inner two rows of four-row plots with a plot combine and were adjusted to 130 g kg$^{-1}$ moisture. Date of maturity was recorded as the day when at least 95% of the pods in a plot had attained a mature color (i.e., R8 of Fehr et al., 1971). Plant height was recorded as the average distance from the soil surface to the apex of the main stem after R8. Lodging of the plants within a plot was rated after R8 using a scoring scale of 1 (all plants erect) to 5 (all plants prostrate). Seed mass was measured as the weight of a sample of 100 or 200 seeds, but expressed as a 100-seed weight value. Seed quality was visually examined and rated after R8 using a scoring scale of 1 (very good quality) to 5 (very poor quality) as described in Crochet and Hughes (2012). This score takes into account greenish colored seeds, the number of seeds with defective seed coats, and moldy or rotten seeds, but not mechanical harvest damage. Seed protein and oil concentrations were measured with near infrared reflectance and are expressed on a 130 g kg$^{-1}$ moisture basis. This was done by placing a 75-g sample of seed harvested from each trial entry in 100 seed counts. Seed was adjusted to 130 g kg$^{-1}$ moisture basis. This was done by placing a 75-g sample of seed harvested from each trial entry in 100 seed counts. Seed was adjusted to 130 g kg$^{-1}$ moisture basis. This was done by placing a 75-g sample of seed harvested from each trial entry in 100 seed counts.
Table 1. Continued.

<table>
<thead>
<tr>
<th>MG II</th>
<th>MG III</th>
<th>MG IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultivar</td>
<td>Year of release</td>
<td>Cultivar</td>
</tr>
<tr>
<td>P-ID-2-11</td>
<td>1982</td>
<td>P-ID-3-15</td>
</tr>
<tr>
<td>P-ID-2-12</td>
<td>1996</td>
<td>P-ID-3-16</td>
</tr>
<tr>
<td>P-ID-2-13</td>
<td>2008</td>
<td>P-ID-3-17</td>
</tr>
<tr>
<td>P-ID-2-14</td>
<td>2008</td>
<td>P-ID-3-18</td>
</tr>
<tr>
<td>P-ID-2-16</td>
<td>1994</td>
<td>P-ID-3-20</td>
</tr>
<tr>
<td>P-ID-2-17</td>
<td>2001</td>
<td>P-ID-3-21</td>
</tr>
<tr>
<td>P-ID-2-19</td>
<td>2005</td>
<td>P-ID-3-23</td>
</tr>
<tr>
<td>P-ID-2-20</td>
<td>2006</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Number of replications in the maturity group (MG)-specific field trials conducted at each listed location and year.

<table>
<thead>
<tr>
<th>Location</th>
<th>MG II</th>
<th>MG III</th>
<th>MG IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2010</td>
<td>2011</td>
<td>2010</td>
</tr>
<tr>
<td>Lamberton, MN</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Waseca, MN</td>
<td>4</td>
<td>4</td>
<td>6†</td>
</tr>
<tr>
<td>Waseca2, MN</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beresford, SD</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arlington, WI</td>
<td>12†</td>
<td>12†</td>
<td></td>
</tr>
<tr>
<td>Ingram, MI</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harrow, ON</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Woodslee, ON</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Hoytville, OH</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Nevada, IA</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finch, IA</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Story City, IA</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boone, IA</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DeKalb, IL</td>
<td>8‡</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Monmouth, IL</td>
<td>8‡</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Arthur, IL</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perry, IL</td>
<td>8†</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Urbana, IL</td>
<td>3</td>
<td>8†</td>
<td>3</td>
</tr>
<tr>
<td>Wooster, OH</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>West Lafayette, IN</td>
<td>4</td>
<td>2†</td>
<td>3</td>
</tr>
<tr>
<td>Muscatine, IA</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Ames, IA</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crawfordsville, IA</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mead, NE‡</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Mead, NE</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Clay Center, NE‡</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Havelock, NE</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Manhattan, KS‡</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Manhattan, KS</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Novelty, MO</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Columbia, MO</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Albany, MO</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Harrisburg, IL</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belleville, IL</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brownstown, IL</td>
<td>8‡</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Dixon Spring, IL</td>
<td>8‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stuttgart, AR</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Suffolk, VA</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexington, KY</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Queenstown, MD</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Perkins, OK</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† Cultivars not equally replicated.
‡ Subsets of 35 to 45 cultivars of the MG II, III, and IV historic sets were grown with four replications following 11 yr continuous corn and four replications following a corn soybean rotation (Fox et al., 2013).
§ Irrigated location.

a cuvette that was inserted into a grain analyzer (Model Infratec 1241, Foss, Hillerød, Denmark). Near infrared analysis is used as a nondestructive approach to measuring protein and oil content of whole soybean seeds.

Entries in the 2010 trials conducted at DeKalb, Perry, Monmouth, Urbana, Brownstown, and Dixon Springs, IL, were a subset of 35 to 45 cultivars of the historic sets, but eight replications were used in these trials (Fox et al., 2013). In addition, some cultivars were replicated more frequently than others in the trials conducted at a second Waseca, MN, location in 2010; West Lafayette, IN, in 2011; and Arlington, WI, in both years, for the purpose of allowing the inclusion of more cultivars in studies aimed at examining how contrasts in agronomic practices (i.e., early vs. late planting) influence the magnitude of the genetic yield gain rate (Rowntree et al., 2013; Table 2). Only the May planting dates from the Rowntree et al. (2013) study were included in our analysis.

Statistical Analysis

To deal with the unbalanced nature of the data set, Proc Mixed of SAS was used, with covariance parameters estimated by the restricted maximum likelihood method, and the fixed effect denominator degrees of freedom estimated with the Kenward Rogers option (Kenward and Rogers, 1997; SAS Institute, 2011). Best linear unbiased predictors (BLUPs) of entries were calculated using estimate statements in a model, with environment, replications nested in environments, cultivar, and cultivar × environment interaction treated as random effects. Regression of cultivar trait values (BLUPs) on the respective year of cultivar release within each MG provided an annualized estimate of genetic change. Cultivar release dates were obtained from published plant cultivar registration articles, published pedigree lists (Bernard et al., 1988), or from the company providing seed. An overall analysis across MGs for yield was completed to test for significant differences in slopes or intercepts among MGs using Proc Mixed of SAS (SAS Institute, 2011). However, to account for maturity date variation within MGs, yield BLUPs were also calculated using maturity date as a covariate in the model. This overall analysis of yield data included the random factors of environment and replications nested in environments, and also the fixed effects of MG, year of release, and MG × year of release interaction. In addition to linear model estimates of cultivar yield improvement, a two-segment linear regression model was also fit to the data, and the two model fits were compared using the Akaike Information Criterion (AIC). The AIC rewards models with a better coefficient of determination ($R^2$) fit, but penalizes models that require more parameters to achieve the better fit (Sakamoto et al., 1986; Posada and Buckley, 2004). The parameters in the segmented linear regression model, as shown below, are:
Y1 = intercept1 + slope1 × X, where Y1 is the dependent variable (yield or other measurement) regressed on the independent variable X (year of release)

Y at X0 = slope1 × X0 + intercept1, where X0 is the nonlinear regression estimate of the breakpoint year that separates the two linear regression segments

Y2 = Y at X0 + slope2 × (X – X0), where Y2 is again the dependent variable for the slope2 segment of the 2-segment regression analysis, and

Y = IF (X < X0, Y1, Y2), where Y is the yield estimate for the X year before or after the breakpoint year.

The parameters were estimated using nonlinear regression curve-fitting with iteration to converge on parameter estimates that provided a best-fit to the data, with results then plotted using GraphPad Prism (GraphPad Software, 2013).

Yield stability (Finlay and Wilkinson, 1963) of cultivars was evaluated by obtaining environment and cultivar × environment BLUPs from the previously described random effects model without the maturity covariate. BLUPs of cultivar × environment combinations were regressed on the environment BLUPs to obtain slopes that describe the stability of each cultivar. Stability slope trends were described by the regression of each cultivar regression value with its year of release. To compare yield stability of old vs. new cultivars, the six oldest and six newest cultivars in the MG II and MG III sets, and five cultivars of each type in MG IV were chosen. Average BLUP yields for the old and new cultivars were regressed with the environment BLUPs. Calculation of BLUPs and regression coefficients were conducted in SAS Proc Mixed (SAS Institute, 2011).

## RESULTS AND DISCUSSION

The weather was hot and dry in 2010 at the southernmost test sites of Perkins, OK; Lexington, KY; Dixon Springs, IL; Stuttgart, AR; and Suffolk, VA. In 2011, growing season conditions were generally better, although environments at Manhattan, KS, Stuttgart, AR, and Brownstown, IL, experienced hot and dry conditions that suppressed yields. During 2011, there was an early frost at Beresford, SD, and the Ontario locations of Harrow and Woodslee were planted late, but these events did not greatly affect yield. Overall, mean yields at test sites were typical for the regions where the sites were located. Average yields for environments ranged from 2123 to 4584 kg ha⁻¹ in MG II, 1799 kg ha⁻¹ to 4577 kg ha⁻¹ in MG III, and 953 kg ha⁻¹ to 4443 kg ha⁻¹ in MG IV. The upper range value is approximately the same in all three MG trial sets, but the lower range values differ, particularly for MG IV. There were no yield trials with a mean yield less than 2000 kg ha⁻¹ in MG II, but there was one in MG III, and five in MG IV.

### Yield Improvement

Results from field tests revealed that seed yields consistently increased over the past 80 yr due to breeding efforts. Across environments, the estimated linear rate of genetic yield gain was 23 kg ha⁻¹ yr⁻¹ in both MG II and MG III, and 20 kg ha⁻¹ yr⁻¹ in MG IV (Fig. 1). A consideration in the analysis is that within each MG, the more recently released cultivars matured an average of 8 d later than older cultivars. Within yield tests, later-maturing cultivars tend to yield greater than earlier-maturing cultivars because a later maturity date allows plants to assimilate more carbon through an extended duration of photosynthesis, thereby resulting in greater seed yield than early maturing cultivars. After adjusting for maturity with the covariate analysis, the linear rates for yield gain were reduced to 20 kg ha⁻¹ yr⁻¹ in both MG II and MG III, and to 18 kg ha⁻¹ yr⁻¹ in MG IV. These estimates of annual genetic yield gain are greater than estimates provided by Luedders (1977) and Wilcox et al. (1979), but comparable with Specht and Williams’ (1984) estimates of 29 kg ha⁻¹ yr⁻¹ in MG II, 17 kg ha⁻¹ yr⁻¹ in MG III, and 22 kg ha⁻¹ yr⁻¹ in MG IV. While the genetic yield gain estimates obtained in the present study do not reflect the entire diversity of soybean production areas of the United States, it is worth noting that these regions account for about 75% of the total U.S. soybean cropping area and production (USDA-NASS, 2013). It nonetheless appeared that genetic improvement (arising from the continual release of ever-higher yielding cultivars that are quickly adopted by producers) has been a key driver of on-farm yield improvement. Rates of genetic yield gain estimated in this study are concordant with the 23 kg ha⁻¹ yr⁻¹ rate of on-farm soybean yield improvement that has occurred during the same 80-yr period in the United States (Specht et al., 2014).

Additional data analyses indicated that the regression coefficients and y-intercepts for MG II and MG III were not significantly different, but that those for MG IV were significantly lower. Hypothesized reasons for the lower MG IV regression parameters were that (i) the five trials with yields < 2000 kg ha⁻¹ among the 27 MG IV trials may have diminished the parameter estimates, or (ii) the MG IV cultivar releases may have intrinsically less yield improvement.
potential than the MG II and III cultivars. To evaluate these hypotheses, regression parameters for the MG III and MG IV cultivar sets were reestimated using just the site-year trial yield data collected at 15 locations where both MG III and MG IV trials were grown. One such location, the nonirrigated Manhattan, KS 2011 site, was the lowest-yielding MG III site and was also a low-yielding MG IV site. Remaining locations with both MG III and MG IV tests comprised moderate- to high-yielding environments. No significant difference was detected between MG III and MG IV for rates of yield gain ($p = 0.35$) or intercept ($p = 0.34$). This suggests that the lower MG IV intercept value in the overall analysis was not due to a lower yield potential of MG IV cultivars, but was instead a result of the less productive soils, drier conditions, and potentially greater biotic stresses encountered at the MG IV testing sites than at the MG II and MG III sites.

The fit of the yield data to either a simple linear or a two-segment linear model was compared to determine if the rate of yield improvement was constant or discontinuous over 80 yr of breeding. The segmented model was statistically more probable than the simple linear model based on the AIC values computed for each MG (Fig. 1). The best-fit breakpoint year was 1968 for MG II, 1964 for MG III, and 1971 for MG IV. The postbreakpoint regression coefficient across MGs averaged 29 kg ha$^{-1}$ yr$^{-1}$ and was 2.6 times greater than the average prebreakpoint coefficient of 11 kg ha$^{-1}$ yr$^{-1}$. Specht et al. (2014) reported that a two-segment linear model also provided the best fit to on-farm U.S. yield data from 1924 to 2012, and noted that yield gain rate increased 50% after the best-fit breakpoint year of 1983.

These data clearly demonstrate that annual genetic gain for seed yield is higher now than in the past, but the large confidence intervals for the breakpoints make it difficult to be specific about when the change occurred or the reasons for the change. There were fewer entries released between 1920 and 1950 than between 1980 and 2010, so the estimate of yield for those early years is not as precise as for later years. In both MGs II and III, there were entries from the 1920s that yielded as much as the entries from the 1960s and these few entries affected the prebreakpoint rate of gain. The improved rate of gain that is sustained post 1970 is likely to be a result of increase in investments by commercial companies in soybean breeding that were stimulated by the passage of the 1970 PVP Act, and also by the increase in the number of public sector soybean breeders. No changes in rate of genetic gain from major scientific and technical advances of the past 15 yr were observed; however, it may be too soon to detect a change if it has occurred. In addition, these results provide no evidence of a yield plateau in either of the three MGs.

### Yield Stability

A stability analysis was conducted to evaluate the yield response of historic sets of cultivars when grown in environments of varying productivity to determine whether yield stability of cultivars has changed over generations of breeding. Stability regression coefficients of cultivars ranged from 0.69 to 1.21 in MG II, 0.80 to 1.26 in MG III, and 0.74 to 1.23 in MG IV (Fig. 2). A cultivar stability

---

**Figure 1.** Scatter plot of cultivar yield vs. cultivar release year for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model and a two-segment linear (Seg Lin) model were fit to the data, and the best-fit trend line statistics for each model are presented in the text box. The dotted lines are the 95% confidence intervals (CIs) for the linear fit. The vertical line denotes the estimated breakpoint year generated with the two-segment linear fit. Comparison of the Akaike information criterion (AIC) values computed for each model was used to determine the most probable model for the given data. Each cultivar data point is the mean of 27 (MGs II and IV) or 26 (MG III) site-years.
The regression coefficient \( b \) of 1.0 indicates that the response of a cultivar to different environments is the same as the mean response of all other cultivars in the experiment (Bernardo, 2010a). A cultivar classified as having a high \( b \) value is considered less stable than a cultivar with a low \( b \) value. The high \( b \) value could indicate poor performance in stressful environments, or conversely, it could also indicate that a cultivar has a better ability to take advantage of favorable environments compared with a cultivar with a lower coefficient. When stability coefficients of cultivars were regressed on year of release (Fig. 2), significant positive regression coefficients were detected, which were 0.005 yr\(^{-1}\) \( (p < 0.0001) \) for MG II and MG IV and 0.002 yr\(^{-1}\) \( (p = 0.0003) \) for MG III. This finding indicates that new cultivars are less stable (but alternatively, more responsive) than old cultivars, relative to the yield mean of all cultivars in each test environment. Voldeng et al. (1997), Wilcox et al. (1979), and De Bruin and Pedersen (2008) reported that yield stability did not change over generations of breeding, but statistical power in those studies was far less than that in the present study in terms of cultivar numbers and test sites.

A group of new and old cultivars from each MG were compared to determine whether increases in stability coefficients (reductions in stability) over time are the result of newer cultivars having reduced performance in stressful environments or the result of newer cultivars being better able to take advantage of favorable environments than older cultivars. We found that new cultivars had greater yields in both low and high yielding environments compared with old cultivars (Fig. 3; \( p < 0.0001 \)). In addition, yield stability coefficients of the new cultivar group were found to be significantly greater than the old cultivar group in all three MGs \( (p < 0.0001; \) Fig. 3). These results show that, despite the reduced stability of modern cultivars, the modern cultivars yielded better on average than old cultivars across all environments tested, and that modern cultivars can take advantage of high-yielding environments better than old cultivars. Araus et al. (2002) had a similar finding in wheat \( (Triticum aestivum) \) and documented that new wheat cultivars were more yield responsive than were old cultivars to more productive test environments.

The relationship between genetic gain rates and location mean yields is shown in Fig. 4. The graphs clearly indicate that measured rate of genetic yield improvement is functionally dependent on productivity of the test environment in which it is measured. Because breeding leads to improvement in genetic yield potential over time, and because high-yielding environments allow more of the genetic yield potential to be expressed, it was not surprising that there is an increase in the rate of genetic gain for yield as the environmental yield potential increases. This finding is indicative of an interaction of genetic improvement with environmental productivity improvement, wherein new cultivars synergistically act with modern agronomic practices to improve yield more than would be expected based on the additive effects of the genetic (old vs. new cultivar yield potential) factor and the agronomic (low vs. high productivity) factor. In that regard, it would be of interest in the future to use these historic cultivar sets to evaluate the rate of genetic yield gain in environments whose productivity exceeds 5000 kg ha\(^{-1}\).

Figure 2. Scatter plot of cultivar yield stability coefficients vs. cultivar release year for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model was fit to the data, and the best-fit trend line statistics are presented in the text box. Dotted lines are the 95% confidence intervals (CIs) for the linear fit. Each cultivar data point is the mean of 27 (MGs II and IV) or 26 (MG III) site-years.
Agronomic Traits

Within each MG tested, the date when cultivars reached maturity increased linearly at a rate of 0.09 to 0.10 d yr\(^{-1}\) (Table 4, Fig. 5). Although this increase is small from year to year, it amounts to approximately 1 d per decade, and over the course of the release years in this study, maturities of new cultivars are about 1 wk later than those of cultivars from the 1950s. This change occurred despite the inclusion of maturity checks used to bracket maturities within each MG of the Uniform Soybean Tests, in which public sector experimental lines are yield-tested before they are released (Crochet and Hughes, 2012). Over time, however, these maturity checks are replaced as new cultivars are released, resulting in later maturity dates for cultivars within a MG. The gradual change towards later maturity is not surprising.

Figure 3. Scatter plot of cultivar yield vs. the respective mean yield of site-year trial used to derive cultivar yield estimates for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model was fit to the data, and the best-fit trend line statistics are presented in the text box. The two sets of cultivars shown here represent cultivars selected from the decile extremes (6 or 5 oldest releases vs. 6 or 5 newest releases) in the time-span distribution of all cultivars (60, 59, 49) in respective MG sets (II, III, IV). Dotted lines are the 95% confidence intervals (CIs) for the linear fits in the oldest or newest groups. Each cultivar data point is the mean of 2 to 12 replications at that site-year.

Figure 4. Scatter plot of genetic yield gain estimates vs. the respective means of site-years used to derive genetic yield gain estimates for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model was fit to the data, and the best-fit trend line statistics are presented in the text box. Dotted lines are the 95% confidence intervals (CIs) for the linear fit. Each genetic gain estimate is taken across all cultivars at that site-year.
Table 4. Estimated rates of genetic change for the listed trait based on simple linear regression of cultivar trait best linear unbiased predictor (BLUP) value on cultivar release year within Maturity Groups (MGs) II, III, and IV.

<table>
<thead>
<tr>
<th>MG</th>
<th>n</th>
<th>b</th>
<th>P</th>
<th>±SE</th>
<th>R²</th>
<th>n</th>
<th>b</th>
<th>P</th>
<th>±SE</th>
<th>R²</th>
<th>n</th>
<th>b</th>
<th>P</th>
<th>±SE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yield</td>
<td>kg ha⁻¹ yr⁻¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yield adjusted for maturity covariate</td>
<td>kg ha⁻¹ yr⁻¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Date of maturity</td>
<td>d yr⁻¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>6001</td>
<td>23.1</td>
<td>***</td>
<td>1.295</td>
<td>0.845</td>
<td>4644</td>
<td>19.6</td>
<td>***</td>
<td>1.307</td>
<td>0.794</td>
<td>4668</td>
<td>0.102</td>
<td>***</td>
<td>0.017</td>
<td>0.369</td>
</tr>
<tr>
<td>III</td>
<td>4991</td>
<td>22.8</td>
<td>***</td>
<td>1.280</td>
<td>0.847</td>
<td>4289</td>
<td>19.7</td>
<td>***</td>
<td>1.258</td>
<td>0.812</td>
<td>4324</td>
<td>0.090</td>
<td>***</td>
<td>0.018</td>
<td>0.303</td>
</tr>
<tr>
<td>IV</td>
<td>4451</td>
<td>19.5</td>
<td>***</td>
<td>1.233</td>
<td>0.845</td>
<td>3474</td>
<td>18.1</td>
<td>***</td>
<td>1.301</td>
<td>0.805</td>
<td>3577</td>
<td>0.090</td>
<td>***</td>
<td>0.020</td>
<td>0.294</td>
</tr>
</tbody>
</table>

| Lodging | score yr⁻¹ | | | | | Plant height | cm yr⁻¹ | | | | | Seed protein | g kg⁻¹ yr⁻¹ | | |
|---------|------------|----|-----|----| | | | | | | | | | | |
| II      | 5904       | -0.012 | *** | 0.002 | 0.413 | 5006 | -0.133 | * | 0.052 | 0.102 | 2633 | -0.222 | *** | 0.061 | 0.189 |
| III     | 4495       | -0.018 | *** | 0.002 | 0.667 | 4195 | -0.205 | *** | 0.053 | 0.207 | 1416 | -0.221 | *** | 0.051 | 0.251 |
| IV      | 3582       | -0.014 | *** | 0.001 | 0.666 | 3631 | -0.129 | ** | 0.042 | 0.167 | 1072 | -0.159 | * | 0.068 | 0.105 |

| Seed oil | g kg⁻¹ yr⁻¹ | | | | | Protein produced | kg ha⁻¹ yr⁻¹ | | | | | Oil produced | kg ha⁻¹ yr⁻¹ | | |
|----------|-------------|----|-----|----| | | | | | | | | | | |
| II       | 2633       | 0.136 | *** | 0.032 | 0.235 | 2633 | 6.594 | *** | 0.482 | 0.763 | 2633 | 4.373 | *** | 0.271 | 0.818 |
| III      | 1416       | 0.103 | ** | 0.030 | 0.170 | 1413 | 6.856 | *** | 0.469 | 0.789 | 1413 | 4.720 | *** | 0.337 | 0.774 |
| IV       | 1072       | 0.051 | NS | 0.031 | 0.055 | 1071 | 7.378 | *** | 0.528 | 0.806 | 1071 | 4.369 | *** | 0.347 | 0.771 |

| 100-seed weight | g yr⁻¹ | | | | | Seed quality | score yr⁻¹ | | | | | | | |
|------------------|-------|----|-----|----| | | | | | | | | | | |
| II    | 2622 | -0.019 | NS | 0.010 | 0.063 | 1308 | -0.0006 | ** | 0.0002 | 0.1626 |
| III   | 1595 | 0.010 | NS | 0.008 | 0.029 | 295 | 4 | |
| IV    | 1415 | 0.002 | NS | 0.008 | 0.002 | 490 | -0.0008 | * | 0.0003 | 0.1232 |

* Significantly different from zero at the 0.05 probability level.
** Significantly different from zero at the 0.01 probability level.
*** Significantly different from zero at the 0.001 probability level.

1 The total number of observations (i.e., replicates, sites, years, cultivars) included in BLUP analyses.
2 Protein and oil units are expressed on a 130 g kg⁻¹ moisture basis.
3 Seed quality is scored visually from 1 (very good) to 5 (very poor).
4 Insufficient data collected to calculate a trend.

because there is a tendency for later-maturing genotypes to have a yield advantage over earlier-maturing lines. Later maturity of recently released cultivars compared with old cultivars also occurred for cultivars developed by proprietary sector breeders, most likely for the same reasons. It can be argued that later-maturing, recently released cultivars within a given MG do offer a better adaptive fit to growers using earlier planting dates (Rowntree et al., 2013), particularly because climate change has resulted in a longer growing season, as is evident in the shifting of USDA plant zones (Kaplan, 2012).

Lodging scores have decreased in newly released cultivars compared with older cultivars or PI. A decrease of 0.012 to 0.018 units yr⁻¹ was found, which translates to a total decrease of about 1 to 1.5 units of lodging over the 80 yr of cultivar releases (Table 4). Previous research also has shown a reduction in lodging (Luedders, 1977; Wilcox et al., 1979; Specht and Williams, 1984; Voldeng et al., 1997; and Wilcox, 2001). Lodging is an important trait in selection, as severe lodging can result in harvest losses and growers have a strong desire for cultivars with reduced lodging.

Mature plant height significantly decreased across generations of breeding, with decreases by MG ranging from 0.13 to 0.21 cm yr⁻¹ (Table 4). Tall cultivars tend to lodge more than short cultivars, and therefore as breeders have selected for reduced lodging, there has been a trend for cultivars to become shorter over time. Specht and Williams (1984) and Wilcox (2001) also reported a decrease of mature plant height over generations of breeding. Even though these decreases are significant, height data are much more variable than seed yield data. Short and tall cultivars existed early in soybean breeding and still do today. Furthermore, this variability results from tall cultivars released in the 1920s or short ones released in the 2000s having a large influence on the overall regression line. For example, the semideterminate cultivar NE3001 (University of Nebraska, Lincoln), which was released in 2004, had a mean height across environments of 61 cm, compared with the mean height of 97 cm averaged across the other MG III cultivars that were released after 2000. When NE3001 was excluded from the MG III data set, the rate of decrease lessened from a rate of 0.21 cm yr⁻¹ to 0.07 cm yr⁻¹. Plant height is not necessarily a trait of primary importance in selection aimed at improving yield and lodging, and both tall and short cultivars meeting the yield and lodging selection criteria are often released.

Genetic changes in seed protein and oil concentrations are important to the soybean processing industry in
terms of the extractable percentage of meal and oil obtainable from a given volume of soybean seed. Seed protein concentration decreased at a rate of 0.22 g kg\(^{-1}\) yr\(^{-1}\) in the MG II and MG III sets, and 0.16 g kg\(^{-1}\) yr\(^{-1}\) in the MG IV set (Table 4, Fig. 6). Concordantly but inversely, oil concentration increased at a rate of 0.14 g kg\(^{-1}\) yr\(^{-1}\) in MG II, 0.10 g kg\(^{-1}\) yr\(^{-1}\) in MG III, and 0.05 g kg\(^{-1}\)yr\(^{-1}\) in MG IV (Table 4, Fig. 6). These rates were significant at the 0.05 significance level, with the exception of the low rate of oil increase in MG IV (Table 4). Similar results were reported by Wilcox et al. (1979) and Voldeng et al. (1997). On the basis of annual Uniform Soybean Test data, Wilcox (2001) also noted similar findings in MG II, but reported a decrease in seed oil concentrations in MG III. Wilson (2004) described the difficulty of breeding for higher protein, oil, and yield in soybean germplasm due to the negative correlation between protein and yield, protein and oil, and the positive correlation between oil and...
yield. Because soybean growers in the United States are compensated on the basis of seed weight sold and not seed composition, increasing yield remains the primary goal of breeding programs. Reduction of protein concentration is, therefore, the consequence of this focus on yield and the negative correlation frequently observed between yield and protein (Burton, 1987). Unless the market for soybean changes to include compensation to growers for seed composition, it is likely that the trend of reduced seed protein and greater oil concentrations will continue. Still, increases in seed yield result in an overall increased total mass of protein produced per hectare (Table 4), which in this study averaged between 6.6 and 7.4 kg ha$^{-1}$ yr$^{-1}$ in the three MG sets (Table 4). Concordantly, the total mass of oil produced per hectare also rose on the scale of 4.4 to 4.7 kg ha$^{-1}$ yr$^{-1}$ in the three MGs. In essence, there is a genetic gain not only for seed yield, but also for seed protein and oil yields per se over the last 80 yr of breeding.

Data collected for 100-seed weight and seed quality were the most variable of the traits measured. No significant change over time in 100-seed weight was detected in any MG (Table 4). This suggests that the significant yield increase observed over time in this study must result from new cultivars producing more seed per unit area compared with old cultivars. With regard to seed quality (Table 4), breeding efforts in the MG II and IV cultivar sets have resulted in slight, but significant, decreases in seed quality scores, indicating improved seed quality. Unfortunately, insufficient data from the MG III set prevented evaluation of the significance of estimated genetic change in seed quality for this MG. Although a small reduction in seed quality score suggests that new MG II and IV cultivars produce seed that may be slightly more appealing, that finding may be of questionable significance, given the difficulty of inferring a biological basis for such an effect. Previous research by Specht and Williams (1984) found an increase in seed size and a decrease in seed quality over generations for breeding soybean. Small or nonsignificant changes in these traits were expected because neither of these traits is a selection target in the development of commodity type cultivars with yield potential as the primary focus.

Estimates of genetic gains in other North American crops are comparable with the estimates in soybean observed in the current study. Battenfield et al. (2013) estimated the rate of winter wheat yield gain grown in the Great Plains region of the United States to be 15 kg ha$^{-1}$ yr$^{-1}$ when compared with a tall cultivar released in 1919, and 11 kg ha$^{-1}$ yr$^{-1}$ when compared with a semi–dwarf cultivar released in 1971. In the eastern United States, Green et al. (2012) estimated yield gains from 1919 to 2009 of soft red winter wheat to be 37 kg ha$^{-1}$ yr$^{-1}$. Dry bean (Phaseolus vulgaris L.) breeding in the western United States has faced many challenges, with broad breeding objectives such as disease resistance and growth types. However, yield improvements made in the great northern, pink, pinto, and red bean market classes were estimated to range from 3% over a 16-yr period to 35% over a 43-yr period (Singh et al., 2007). In contrast, maize (Zea mays L.) has experienced tremendous improvements to yield potential. Duvick et al. (2004) estimated improvements of hybrids grown from 1930 to 2001 to be 77 kg ha$^{-1}$ yr$^{-1}$ when grown at their optimum density. However, Specht et al. (2014) noted that when on-farm irrigated corn and soybean yield improvement is examined synchronously in terms of an annual corn–soybean yield ratio, that ratio has been a 3:1 constant for the past 42 yr. Though the physiological differences between the two crop species account for soybean yielding about three-fold less than corn in an absolute sense, the relative pace of on-farm corn and soybean yield improvement has been effectively and remarkably equivalent.

**CONCLUSIONS**

Plant breeders are continually making selections to improve cultivars that producers grow. Over the past 80 yr, soybean seed yield has increased considerably, and this experiment produced estimates that the genetic improvement of seed yield has totaled 1450 to 1578 kg ha$^{-1}$ for the MGs studied, after adjustments for the date of maturity. In an analysis that included comparisons of on-farm yield gains with genetic gains in the northern United States, Specht et al. (2014), estimated that two-thirds of the on-farm yield gains were the result of genetic improvements and that one-third was the result of agronomic improvements. However, there also is evidence of synergistic interactions between genetic and agronomic improvements (Rowntree et al., 2013).

Efforts of soybean breeders in both public and proprietary programs have developed soybean into a leading North American crop for domestic and export markets and increasing the rate of yield gain is important for meeting ever-greater worldwide demand for food, feed, and fuel. Further improvements will continue to face challenges of limited genetic diversity and biotic and abiotic stresses (Gizlice et al., 1994; Sneller, 1994; Betzelberger et al., 2012). However, breeders will have new tools available such as technology to collect phenotypic data, DNA and RNA sequences, molecular markers, mapping studies, and methodologies to aid in yield improvement (Bernardo, 2010b; Furbank and Tester, 2011; Poland and Rife, 2012; Schmutz et al., 2010).

**Acknowledgments**

This work was supported by funding from the Illinois Soybean Association and the United Soybean Board.
References


Bernardo, R. 2010a. Breeding for quantitative traits in plants. 2nd ed. Stemma Press, Woodbury, MN.


GraphPad Software. 2013. GraphPad Prism for Windows. v. 6.02. GraphPad Software, La Jolla, CA.


USDA-Agriculture Marketing Service, Washington, DC.


