Physiological and management factors contributing to soybean potential yield

Ryan J. Van Roekel a, Larry C. Purcell b,*, Montserrat Salmerón b

a DuPont Pioneer, 1111 25th Street, Pleasantville, IA 50225, United States
b University of Arkansas, 1366 W. Altheimer Dr., Fayetteville, AR 72704, United States

A R T I C L E  I N F O

Article history:
Received 27 October 2014
Accepted 23 May 2015
Available online 13 June 2015

Keywords:
Crop models
Effective filling period
Potential yield
Seed growth rate
Soybean
Radiation use efficiency

A B S T R A C T

The largest reported soybean grain yield is approximately three-fold more than the highest reported U.S. average yield. An understanding of yield determination is needed to identify avenues for increasing yield and for defining the yield potential of soybean. To illustrate physiological traits important for yield determination, we used a framework that models yield as the product of seed number (seed m⁻²) and individual seed mass (massseed). Developmentally, seed m⁻² is determined first and is proportional to the biomass accumulation rate (BAR, g m⁻² d⁻¹) and the fraction of assimilate allocated to reproductive structures. Seed m⁻² is inversely proportional to the individual seed growth rate (ISGR, mg seed⁻¹ d⁻¹) where the ISGR represents the minimum amount of assimilate necessary to prevent a flower or pod from aborting. Hence, seed m⁻² can be increased by optimizing conditions for crop growth (e.g., radiation interception, stress-free environment, high soil fertility levels) and having a low ISGR. Determination of massseed occurs later during ontogeny than seed m⁻² and can be expressed as the product of the ISGR and the effective seedfilling period (EFP, d). Variation among genotypes for ISGR is quite large and is generally not affected greatly by the environment. There is also genotypic variation in the EFP, but the EFP is decreased by a variety of biotic and abiotic stresses. Our analysis indicates that reaching the potential yield of soybean depends upon high BAR and extending the EFP, and a key factor affecting both of these variables is ensuring non-limiting crop nutrition, especially nitrogen. Strategies for increasing soybean maximum yield include early planting (which extends the EFP), optimizing crop nutrition, minimizing biotic and abiotic stresses, and developing breeding programs tailored for high yield environments. Characterizing physiological traits important for yield with genetic markers offers tools for combining favorable traits for high-yield environments.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction—Current yield levels and records

United States average soybean [Glycine max L. (Merr.)] grain yields have increased from the earliest record of 739 kg ha⁻¹ in 1924 to a high of 2956 kg ha⁻¹ in 2009 (USDA-NASS, 2013). While this increase in soybean yield over time is substantial, both researchers and growers have documented yields much greater than the reported nationwide averages (Table 1). In converted rice paddies in Japan, Spaeth et al. (1987) reported yields of 6490 kg ha⁻¹. In New Jersey, Flannery (1989) recorded a soybean yield of 7923 kg ha⁻¹ in 1983 and a 5-yr average irrigated yield of 6921 kg ha⁻¹. In 1982, Cooper (2003) was able to achieve yields of 6817 kg ha⁻¹ in research in Ohio. In Queensland, Australia yields up to 8604 kg ha⁻¹ were reported (Cooper, 2003; Lown et al., 1984). Researchers in China recorded yields up to 9200 kg ha⁻¹ (Jiao et al., 2010), but yields were based on only a small number of individual plants (14 to 28). Without further details and without a more representative yield sample, this report remains questionable. Physiological experiments in Argentina extended the photoperiod by 2 h in field experiments from 1 to 35 days after R3 (Fehr and Caviness, 1977) and reported yields up to 8957 kg ha⁻¹ (Kantolic and Slafer, 2007).

A few soybean growers have also achieved exceptional soybean yields. In 1968, the winner of the United States National Soybean Yield Contest did so with 7310 kg ha⁻¹ (Cooper, 2003). In 1997,
a grower achieved yields near 6719 kg ha\(^{-1}\) in the Nebraska irrigated contest category (Specht et al., 1999). In 2008, the winner of the Missouri Soybean Association’s non-irrigated contest category had a yield of 7324 kg ha\(^{-1}\) in southeast Missouri (Steever, 2006). From 2011 to 2013, and each year individual cultivars had yields between 6979 and 7953 kg ha\(^{-1}\).

The high yields reported in yield contests have created some controversy and skepticism because of the lack of supportive, quantitative data that would provide a mechanistic explanation. Additional concerns are associated with the uncertainty in what constitutes the potential yield of soybean. Potential yield is defined as the “yield of a crop cultivar when grown with water and nutrients non-limiting and biotic stress effectively controlled” (van Ittersum et al., 2013). This becomes an issue when attempting to estimate future prospects for yield increases via yield-gap analyses using current farmer average yields and the potential yield (Egli and Hatfield, 2014; Lobell et al., 2009). High potential yield estimates will often result in large yield gaps; however, attaining such potential yields are likely not economically or sustainably possible (Egli and Hatfield, 2014). Potential yield also varies geographically due to changes in climate. Furthermore, non-irrigated production systems may be consistently constrained by water supply and a water-limited potential yield may be more realistic in those production systems (Van Ittersum et al., 2013). However, the focus of this discussion will be on the ultimate potential yield of soybean when water supplies are non-limiting. To provide an estimate of the potential yield of soybean, it is crucial to first examine the processes that determine soybean grain yield.

Our objective was to review mechanistic frameworks for determining the primary two components of soybean yield: seed number (seed m\(^{-2}\)) and individual seed mass (mass\(_{seed}\), g seed\(^{-1}\)). The determination of soybean seed number and mass revolve around a few key growth characteristics that can be quantified as the above-ground biomass accumulation rate, individual seed growth rate, and the effective seedfilling period. We will discuss how these physiological processes are affected by environmental, management, and genetic factors to better understand the combination of factors that lead to expression of soybean potential yield.

### 2. Theoretical framework for soybean yield

#### 2.1. Seed number determination

Soybean grain yield is determined by the seed per unit area and mass of individual seeds:

\[
Yield \ (g \ m^{-2}) = (seed \ m^{-2}) \cdot (mass_{seed}) \tag{1}
\]

During soybean ontogeny, seed number determination occurs first, followed by seed weight determination (Board and Kahlion, 2011). Of these two variables, seed m\(^{-2}\) has a greater influence on yield compared with seed weight (Board, 1987; De Bruin and Pedersen, 2008; Kokubun and Watanabe, 1983; Robinson et al., 2009; Shibies et al., 1975; Singer et al., 2004). One theory regarding seed m\(^{-2}\) determination was proposed by W.G. Duncan (Egli et al., 1978b) whereby “the number of seed produced by a soybean community is set at a level such that the sum of their individual growth rates essentially equals the ability of the soybean canopy to support seed growth”. In other words, seed m\(^{-2}\) is a function of the total crop canopy photosynthate production and the rate of photosynthate utilization by the individual seed. Charles-Edwards (1984) theorized that “each growing point on a plant requires a minimum flux of assimilate for growth to continue.” These concepts were applied to the determination of seed m\(^{-2}\) by Charles-Edwards et al. (1986):

\[
N_g = \nabla F \ast \gamma \ast A_{kg}^{-1} \tag{2}
\]

In Eq. (2), \(N_g\) represents the number of developing grain \(\nabla F\) represents the daily canopy net photosynthesis, \(\gamma\) represents the partitioning coefficient of daily canopy net photosynthesis to reproductive growth, and \(A_{kg}\) is the minimum amount of assimilate required to prevent developing grain from aborting. Eq. (2) indicates that seed m\(^{-2}\) may be increased by lowering the photosynthate requirement per grain, allowing the total crop photosynthate production to be divided among additional reproductive units. Alternative or supplemental avenues to increase the seed m\(^{-2}\) would be to increase the amount of photosynthate produced during the flowering and podset periods or to increase the proportion of photosynthate allocated to seeds.

This model was evaluated by Egli and Zhen-wen (1991) who used seed m\(^{-2}\) as an estimate for \(N_g\) and experiments employing shade treatments to vary \(\nabla F\). Net canopy photosynthesis production was estimated as the above-ground biomass accumulation rate (BAR; g m\(^{-2}\) d\(^{-1}\)) during flowering and podset (R1 to R5; Fehr
and Caviness, 1977) and the individual seed growth rate (ISGR; g seed$^{-1}$ d$^{-1}$) was used as a proxy for $A_g$:

$$\text{seed m}^{-2} = \text{BAR} \times \gamma \times \text{ISGR}^{-1}$$

(3)

For seed m$^{-2}$ to be a suitable substitution for $N_g$, the potential fruit number must be greater than the final seed m$^{-2}$. If this is not the case, excess photosynthate ($\gamma F$) will not be accounted for after partitioning to the grain sites ($\gamma A_g^2$). With soybean, estimates indicate that 32 to 87% of flowers and immature pods abort and do not develop into mature pods with fully developed seed (Hansen and Shibles, 1978; Peterson et al., 1986; Van Schaik and Probst, 1958; Wiebold et al., 1981). Favorable environments lead to increased flower production rather than decreased abortion rates (Jiang and Egli, 1995), and the majority of aborted flowers are fertilized, suggesting that failure of fertilization plays a negligible role in flower abortion (Abernathy et al., 1977). As such, it was assumed that the potential number of grain sites in soybean was always greater than the final seed m$^{-2}$, and seed m$^{-2}$ is an appropriate substitution for $N_g$ (Egli and Zhen-wen, 1991).

Conceptually, total photosynthate production, less the amount needed for maintenance and respiration, would be used for new vegetative and reproductive biomass. Additionally, in experiments involving cultivars with a range of relative maturities (Egli, 1993) or shade prior to R1, the response of seed m$^{-2}$ to BAR is not affected by total vegetative mass at R1, provided that the canopy is closed (Jiang and Egli, 1995). Thus, BAR serves as an appropriate estimate of $\gamma F$ regardless of maturity or cultural practices that affect plant size at flowering.

Egli (1998) compiled data for mass$_{seed}$, ISGR, and the effective seedfilling period (EFP, d) for 12 soybean cultivars (Table 2). The range of ISGR varied about 4-fold (3.6 to 14.7 mg seed$^{-1}$ d$^{-1}$) whereas the range of the EFP varied 1.5-fold (22 to 33 d). The 5.5-fold variation in soybean mass$_{seed}$ (88 to 484 mg seed$^{-1}$) was primarily due to differences in ISGR although environmental factors interacted with genetics to influence the mass$_{seed}$ (Egli, 1975; Egli et al., 1978b; Guldan and Brun, 1985). Furthermore, within developing pods, seeds at the basal position have a slower growth rate and a lower mass$_{seed}$ compared with seeds at the middle or apical positions in a pod (Egli et al., 1978b).

An additional assumption required if ISGR is to be used as a proxy for $A_g$ is that the ISGR must be stable across a range of $\gamma F$ or BAR. Egli and Leggett (1976) documented that the ISGR remained relatively stable despite leaf or pod removal. Others have also found that the ISGR was stable across changes in the short-term photosynthetic supply for soybean (Koller, 1971) and maize [Zea mays L.] (Duncan et al., 1965) and with water-deficit stress in soybean (Mecik et al., 1984; Westgate et al., 1989). There were also no consistent differences in ISGR between early and late formed pods (Egli et al., 1978b), and ISGR is representative of all the seeds on a plant despite differences in when seeds are first formed (Egli and Zhen-wen, 1991).

Although the ISGR is linear throughout most of seedfill, the entire seed growth cycle resembles a sigmoidal function where at the beginning of seedfill, the ISGR increases rapidly to the linear phase and remains constant until near physiological maturity when the rate decreases to 0 (Egli, 1994). For the ISGR to be used as a surrogate for $A_g$, these periods of non-linear ISGR are assumed to not greatly affect seed m$^{-2}$. Furthermore, changes in seed composition affect $A_g$ (Sinclair and de Wit, 1975), and the substitution of $A_g$ with ISGR is only appropriate when comparing cultivars or crops with similar seed compositions.

Given these assumptions, Egli and Zhen-wen (1991) evaluated Eq. (3) and found that when BAR increased during flowering and pod set, seed m$^{-2}$ increased proportionately, and similar results were reported by others (Ball et al., 2000; Rotundo et al., 2012).

Table 2
Summary of the range and average reported values in the literature from a single treatment or cultivar within each site-yr for the rate of biomass accumulation of soybean (BAR), individual seed growth rate (ISGR), effective seedfilling duration (EFP), and partitioning coefficient of net photosynthesis to reproductive growth ($\gamma$).

<table>
<thead>
<tr>
<th>Source</th>
<th>BAR</th>
<th>ISGR</th>
<th>EFP</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>g m$^{-2}$ d$^{-1}$</td>
<td>mg seed$^{-1}$ d$^{-1}$</td>
<td>Range</td>
<td>Average</td>
</tr>
<tr>
<td>Egli and Leggett (1973)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Egli (1975)</td>
<td>–</td>
<td>–</td>
<td>3.4–8.3</td>
<td>5.6</td>
</tr>
<tr>
<td>Egli et al. (1978b)</td>
<td>–</td>
<td>–</td>
<td>3.2–9.4</td>
<td>5.8</td>
</tr>
<tr>
<td>Egli and Wardlaw (1980)</td>
<td>–</td>
<td>–</td>
<td>5.2–9.1</td>
<td>7.1</td>
</tr>
<tr>
<td>Egli et al. (1981)</td>
<td>–</td>
<td>–</td>
<td>3.9–10.9</td>
<td>6.2</td>
</tr>
<tr>
<td>Gbikpi and Crookston (1981)</td>
<td>–</td>
<td>–</td>
<td>4.3–7.0</td>
<td>5.4</td>
</tr>
<tr>
<td>Egli et al. (1984)</td>
<td>–</td>
<td>–</td>
<td>3.3–13.0</td>
<td>7.1</td>
</tr>
<tr>
<td>Meckel et al. (1984)</td>
<td>–</td>
<td>–</td>
<td>5.8–7.3</td>
<td>6.6</td>
</tr>
<tr>
<td>Salado-Navarro et al. (1985a)</td>
<td>–</td>
<td>–</td>
<td>8.0–10.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Salado-Navarro et al. (1986a)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Egli et al. (1987)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Swank et al. (1987)</td>
<td>–</td>
<td>–</td>
<td>5.4–8.6</td>
<td>7.1</td>
</tr>
<tr>
<td>Pfleiffer and Egli (1988)</td>
<td>–</td>
<td>–</td>
<td>2.3–7.0</td>
<td>4.2</td>
</tr>
<tr>
<td>Egli et al. (1989)</td>
<td>–</td>
<td>–</td>
<td>2.3–7.0</td>
<td>4.2</td>
</tr>
<tr>
<td>Egli and Zhen-wen (1991)</td>
<td>4.0–27.0</td>
<td>12.0</td>
<td>4.3–12.0</td>
<td>6.3</td>
</tr>
<tr>
<td>Board et al. (1992)</td>
<td>12.3–18.4</td>
<td>14.7</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Board and Harville (1993)</td>
<td>5.1–18.3</td>
<td>10.2</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Egli (1993)</td>
<td>5.3–17.2</td>
<td>10.6</td>
<td>3.8–13.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Jiang and Egli (1995)</td>
<td>9.8–17.6</td>
<td>14.0</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Board and Harville (1996)</td>
<td>6.84–15.9</td>
<td>11.7</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ball et al. (2000)</td>
<td>–</td>
<td>–</td>
<td>3.2–5.0</td>
<td>3.7</td>
</tr>
<tr>
<td>Kumudini et al. (2001)</td>
<td>–</td>
<td>–</td>
<td>3.85–5.26</td>
<td>4.84</td>
</tr>
<tr>
<td>Vega et al. (2001)</td>
<td>–</td>
<td>–</td>
<td>2.9–4.2</td>
<td>3.3</td>
</tr>
<tr>
<td>Egli and Bruening (2007)</td>
<td>–</td>
<td>–</td>
<td>4.2–6.1</td>
<td>5.1</td>
</tr>
<tr>
<td>De Bruin and Pedersen (2009)</td>
<td>7.4–18.0</td>
<td>12.3</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Thomas et al. (2010)</td>
<td>–</td>
<td>–</td>
<td>2.2–6.3</td>
<td>4.8</td>
</tr>
<tr>
<td>Rotundo et al. (2012)</td>
<td>11.6–15.6</td>
<td>13.7</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Kantolic et al. (2013)</td>
<td>4.4–21</td>
<td>9.3</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Van Rooisel and Purcell (2014)</td>
<td>13.6–64.3</td>
<td>33.5</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
Shade, defoliation, or stress at any time during flowering and pod set (R1 to R5) that decreases BAR results in increased pod abortion and decreased seed m⁻² and yield (Andrade and Ferreiro, 1996; Board and Harville, 1993; Board and Tan, 1995; Egli, 1993, 1997; Egli and Zhen-wen, 1991; Jiang and Egli, 1993; Mann and Jaworski, 1970; Schou et al., 1978; Sharma et al., 1996). Alternatively, supplemental solar radiation during flowering and pod set increased seed yield, seeds, nodes, pods and branches plant⁻¹, pods node⁻¹, and seeds pod⁻¹ (Johnston et al., 1969; Mathew et al., 2000; Schou et al., 1978), and supplemental N increased seed m⁻² and yield (Brevedan et al., 1978). Furthermore, late-developing flowers are more likely to abort than those already developing at a single node, and the removal of earlier formed flowers and pods can reduce the abortion rates of late-developing flowers (Brun and Betts, 1984; Egli and Bruening, 2002; Heitholt et al., 1986; Huff and Dybing, 1980). Others have documented that depodding can reduce the abortion rates of remaining flowers and pods (Heitholt et al., 1986; Hicks and Pendleton, 1969; McAlister and Krober, 1958). All of these observations support Eqs. (2) and (3) whereby seed development depends upon assimilate availability and that additional grain sites can be maintained by increasing photosynthetic production or availability.

The effect of the assimilate demand by each developing grain site on final seed m⁻² was also demonstrated by Egli and Zhen-wen (1991). In this experiment seed m⁻² increased linearly as BAR increased for both ‘Essex’ and ‘Harper’ (Fig. 1), but the slope of this relationship was greater for Essex than Harper. The greater seed m⁻² for Essex at a given BAR level compared to Harper was associated with a lesser ISGR (4.5 vs. 6.2 mg seed⁻¹ d⁻¹) for Essex as predicted by Eq. (3). Eqs. (2) and (3) were confirmed in a second experiment involving two additional cultivars differing in ISGR (Egli and Zhen-wen, 1991). Consistent with the effects of ISGR on seed m⁻², lines bred for large seed had fewer seeds and pods plant⁻¹ and similar yield to the recurrent parents (Hartwig and Edwards, 1970). Similarly, a reduction in massseed results in greater pods and seed m⁻² (Bruening and Egli, 1999; Pedersen and Lauer, 2004; Woodward and Begg, 1976). All of these findings support Eq. (3) where at a fixed level of photosynthetic production and partitioning (BAR * γ), the photosynthetic demand is also fixed (seed m⁻² * ISGR).

Thus, increases in ISGR will reduce seed m⁻². This explains why breeding efforts to increase massseed have not resulted in increased yield (Hartwig and Edwards, 1970). Egli and Zhen-wen (1991) and Ball et al. (2000) estimated the partitioning coefficient (γ) from their measurements of BAR, ISGR and seed m⁻². Ball et al. (2000) found γ remained constant over alterations in BAR across plant densities, irrigation regimes and years. For cultivars from Argentina, Rotundo et al. (2012) found γ was greatest with clusters of cultivars with high seed m⁻² and lowest with low seed m⁻². For cultivars from the USA, clusters of cultivars with both high and low seed m⁻² had a high γ, and there was no relationship between γ and seed m⁻². Although, Vega et al. (2001) found γ decreased slightly with increasing BAR, which was consistent with earlier findings (Egli et al., 1985a; Egli and Zhen-wen, 1991), they observed this effect in only a few, very small plants and suggested that this effect was negligible and that γ was a fairly stable trait.

Eq. (3) indicates that seed m⁻² is proportional to the BAR during the R1 to R5 development period, but it appears that seed m⁻² also increases as the length of the flowering period increases. Extending the photoperiod by 2 to 3 h using low-intensity lamps during the R1 to R3 period resulted in prolonged reproductive growth (R3 to R6) and large increases in both seed m⁻² and yield (Kantolic and Slafer, 2001, 2007; Kantolic et al., 2013). By including a shading treatment in conjunction with the extended photoperiod, Kantolic et al. (2013) showed that seed m⁻² was proportional to BAR (Eq. (3)) and that there was a much closer relationship between seed m⁻² and the cumulative amount of intercepted radiation from R3 to R6. Although Kantolic et al. (2013) focused on the period between R3 and R6, flower and young pod abortion would be expected to begin at R1, and pods generally do not abort after 10 to 12 days after R5 (Board and Tan, 1995). With these thoughts in mind, Eq. (3) may be revised to explicitly include the duration of the period from R1 to R5 as being proportional to seed m⁻²:

\[
\text{seed m}^{-2} = \left( \frac{\sum_{R1}^{R5} \text{IR}}{\text{R1}} \right) \times \text{RUE} \times \gamma \times A_{g}^{-1}
\]

In Eq. (4), intercepted radiation (IR, MJ m⁻²) is cumulated from R1 to R5 and multiplied by the RUE to give the biomass produced during this period (g m⁻²). The partitioning coefficient, γ, accounts for the amount of photosynthate partitioned to the seed as described in Eq. (3). The last term in Eq. (4), A_g (g seed⁻¹), represents the amount of photosynthate required per seed to prevent abortion (Charles-Edwards et al., 1986; Egli and Zhen-wen, 1991) and is numerically equivalent to the ISGR on a daily basis.

### 2.2. Seed mass determination

Individual seed mass (massseed, mg seed⁻¹) is the product of the ISGR (mg seed⁻¹ d⁻¹) and the duration of seed growth, represented as the EFP:

\[
\text{massseed} = \text{ISGR} \times \text{EFP}
\]

The ISGR and EFP are independent traits (Egli et al., 1981, 1984; Egli and Wardlaw, 1980; Meckel et al., 1984), and both traits are affected by the environment but at different times during development. Genotypic differences in ISGR are closely associated with differences in the number of cotyledon cells (Egli et al., 1981), and decreased photosynthate supply (e.g., shading, defoliation, drought) during cotyledon cell division results in decreased ISGR and massseed (Egli et al., 1989). For example, removing 80% of the developing pods on the cultivar ‘York’ resulted in a 51% increase in cotyledon cells in the remaining seeds, an increase in ISGR from 3.5 to 4.9 mg seed⁻¹ d⁻¹, and an increase in final massseed from 94 to 162 mg seed⁻¹ d⁻¹, a two-thirds defoliation treatment decreased cotyledon cell number by 55%, decreased ISGR from 3.5 to 2.3 mg seed⁻¹ d⁻¹, and decreased massseed by 40%.

Once cotyledon cell division is completed, around the time the seed enters the linear phase of growth (Egli et al., 1981), short-term changes in the photosynthate supply in field experiments have relatively little effect on ISGR (Egli and Leggett, 1976; Koller, 1971).
Severe alterations in the source–sink balance after seed cell division is complete by 75% fruit removal or 60% shade treatment result in increased and decreased ISGR, respectively, compared to the control treatment (Egli et al., 1985b), but these extreme changes in photosynthate supply are unlikely to be encountered in the field. Therefore, it appears that under most field conditions, adverse conditions during seed cell division affect massseed by decreasing cotyledon cell number and ISGR; adverse conditions after seed cell division, decrease massseed by decreasing the EFP.

The length of the EFP ranges from 12 to 57 d (Egli et al., 1984, 1987; Egli and Leggett, 1973; Gay et al., 1980; Hanway and Weber, 1971a; Swank et al., 1987) but ranges between 22 and 33 d for many cultivars (Egli, 1998). Drought, high-temperature, and N stresses shorten the EFP, resulting in a corresponding decrease in massseed (Egli et al., 1984; Egli and Wardlaw, 1980; Meckel et al., 1984). Thomas and Raper (1976) and Gbikpi and Crookston (1981) demonstrated that ISGR is increased and the EFP decreased by shortening photoperiods. The ISGR also increases as temperature increases between 18/13 and 27/22°C but then remains stable up to 33/28°C (Egli and Wardlaw, 1980). Similarly, Thomas et al. (2010) found that ISGR was stable between 28/18 and 32/22°C and that the ISGR decreased as temperatures increased to 40/30 and 44/34°C.

Hanway and Weber (1971a) found that differences in yield among a group of eight cultivars were primarily due to EFP and that cultivars had similar ISGR. Likewise, Egli and Leggett (1973) found cultivar differences in ISGR but determined that grain yield differences were more closely associated with EFP. Swank et al. (1987) found examples of variation in massseed among genotypes that were associated with differences in both ISGR and EFP. When comparing old and new cultivars, the yield advantage of cultivar ‘Wiliams’ over ‘Lincoln’ was due to a greater massseed from a longer EFP (Gay et al., 1980). Others have also demonstrated that the EFP in modern cultivars has increased over time of release (Boerma and Ashley, 1988; Kumudini et al., 2001; McClane and Hume, 1981; Metz et al., 1984, 1985; Pfeiffer and Egli, 1988; Rowntree et al., 2014; Salado-Navarro et al., 1985b; Shiraiwa and Hashikawa, 1995; Smith and Nelson, 1987).

A longer EFP is usually associated with a lower dry matter allocation coefficient (DMAC; d−1; Salado-Navarro et al., 1986a, 1986b, 1993). The DMAC was defined as the slope of the linear increase in HI during seedfill (Salado-Navarro et al., 1985a). As such, DMAC has been suggested as an alternative measure of the ISGR as a representation of the whole-plant seed growth rate (Spath and Sinclair, 1985).

3. Crop nitrogen dynamics

To sustain high BAR, and potentially increase seed m−2, high rates of N accumulation are required for protein production to support seed growth and high rates of photosynthesis. While yield is correlated with biomass accumulation, N accumulation can often explain more yield variation than biomass accumulation (Rotundo et al., 2014; Sinclair and Jamieson, 2006). Although there were no differences in total N accumulation among a group of high yielding cultivars, there were differences in their timing of N accumulation (Rotundo et al., 2014). While the majority of N accumulation occurred during R1 to R5, some high yielding cultivars acquired more N prior to flowering while others obtained more N during seedfill. There were also differences among high-yielding genotypes in meeting seed N demand through different combinations of N accumulation during seedfill and the amount of N remobilized.

Genotypes with a short EFP require a more rapid N remobilization from the leaves to support seed protein accumulation when compared with genotypes with a longer EFP (Boon-Long et al., 1983; Egli et al., 1987). A short EFP was also associated with a more rapid rate of decline in CO2 uptake and earlier leaf abscission (Boon-Long et al., 1983; Gay et al., 1980). Sinclair and de Wit (1976) theorized that the N demand for soybean and other high protein crops could not be met by N accumulation alone and that remobilization and translocation of N and proteins from vegetative tissues must occur to support seed growth. As this pool of N and proteins was depleted (Boote et al., 1978; Borst and Thatcher, 1931; Egli et al., 1978a; Hanway and Weber, 1971b), photosynthetic rates also declined (Boote et al., 1978; Lugg and Sinclair, 1981; Mondal et al., 1978; Sesan and Shibles, 1980; Sinclair, 1980; Wittenbach et al., 1980).

Eventually, the photosynthetic apparatus loses functional integrity along with other physiological processes, thereby ending the EFP and limiting final seed yield. This “self-destruct” hypothesis was supported by data from Salado-Navarro et al. (1985a) who found leaf senescence rates increased among genotypes as there were increases in the seed protein concentration and seed N demand. Others also observed that N deficiency shortened the EFP and accelerated leaf senescence (Boon-Long et al., 1983; Hayati et al., 1995). The self-destruct hypothesis also explains why a lower DMAC (or ISGR) increases the EFP by slowing the daily demand for C and N remobilization (Salado-Navarro et al., 1986a, 1986b, 1993). Diversity for N accumulation during seedfill and N remobilization among high-yielding elite cultivars may provide an avenue for increasing the EFP by delaying senescence (Rotundo et al., 2014).

However, previous efforts to supply N during seedfill did not prevent senescence or N redistribution (Egli et al., 1978a; Hayati et al., 1995, 1996). Likewise, in a comparison of soybean maturity isolines, a maturity group (MG) 6 isoline had 73 to 100% more total N at the beginning of seedfill than the MG 4 isoline but N content in harvested seed was not different (Mastrodomenico and Purcell, 2012), indicating that N supply was not limiting yield.

The assumption that N supply during seedfilling is inadequate to prevent catabolism of vegetative proteins to support seed growth has led to numerous experiments attempting to supplement N2 fixation with N fertilizer (Salvagiotti et al., 2008). A meta-analysis of the response of soybean to N fertilization determined that yield was increased in roughly half of the reports (Salvagiotti et al., 2008). The lack of consistency is likely due in part to the decrease in N2 fixation that occurs when N fertilizer is applied (Harper, 1987). That is, any potential gain in N supply by N fertilization may be offset by a decrease in N2 fixation.

Two key questions are pertinent in considering the possible benefits of N fertilization of soybean. The first question is when during the lifecycle of soybean does N2 fixation decrease such that supplemental N will contribute to a positive N balance without disrupting N2 fixation? Experiments evaluating N2 fixation throughout development by measuring acetylene reduction activity on detached roots concluded that N2 fixation peaked around flowering and declined rapidly during seedfill (Lawn and Brun, 1974; Thibodeau and Jaworski, 1975). However, experiments quantifying N2 fixation using nondestructive methods have concluded that N2 fixation continues at high rates until near the end of seedfill (Denison and Sinclair, 1985; Leffel et al., 1992; Mastrodomenico and Purcell, 2012) unless plants are stressed during seedfill, which results in an irreversible decline in N2 fixation (Breeden and Egli, 2003; Mastrodomenico et al., 2013).

A second question important in understanding possible benefits from N fertilization of soybean is how much N can be provided by N2 fixation? The meta analysis of Salvagiotti et al. (2008) reported that N2 fixation provided on average 52% of the crop-N supply and provided in absolute terms between 0 and 337 kg N ha−1. Lower values of N2 fixation were associated with stressful conditions, high rates of N fertilization, or ineffective Bradyrhizobium japonicum strains. When considering all the sites in which yield was responsive to N fertilizer, Salvagiotti et al. (2008) concluded that when yields...
were less than 4500 kg ha\(^{-1}\) N fertilization helped overcome stressful environmental conditions. There were nine sites with yields exceeding 4500 kg ha\(^{-1}\) that responded positively to N fertilizer, and the authors concluded that for these sites the N supply from the soil and from N\(_2\) fixation likely was less than the N required to fully express yield potential. It may be that in reports where yield was not responsive to increased exogenous or endogenous N supply that factors other than N may have limited yield (Salvagiotti et al., 2008).

Wilson et al. (2014) evaluated the response of yield to N fertilizer (560 kg N ha\(^{-1}\)) of 59 MG 2 and 57 MG 3 cultivars that had been released between 1923 and 2008. Although the MG 2 cultivars were not responsive to N fertilizer, presumably due to high residual soil N at the test locations, the MG 3 cultivars with N fertilizer had a 23% greater rate-of-year increase (kg ha\(^{-1}\) yr\(^{-1}\)) than did the treatment without N fertilizer. Regression equations for MG 3 cultivars in 2008 predicted a yield of 4481 kg ha\(^{-1}\) with N fertilizer and 3852 kg ha\(^{-1}\) in the absence of N fertilizer. Interestingly, N fertilization in the MG 3 cultivars resulted in an approximate 10% increase in mass\(_{\text{seed}}\) but had no effect on seed m\(^{-2}\). The increase in mass\(_{\text{seed}}\) along with retention of green leaves late during seed filling (Wilson et al., 2014) provide evidence that N fertilization delayed senescence and extended the EFP, resulting in increased yield.

4. Potential yield prediction

Potential yields are inherently difficult to determine because they depend upon near perfect crop management (e.g., cultivar selection, planting date, plant population, optimum nutrition for all essential nutrients, elimination of biotic and abiotic stresses) where the only limitations to productivity are environmental factors (e.g., solar radiation, temperature, vapor pressure deficit) (Lobell et al., 2009). Because of difficulties in achieving near perfect environments, crop models have been used to assess potential yields assuming ideal environmental conditions (Lobell et al., 2009; Van Ittersum et al., 2013). Other methods for estimating potential yields include yield estimates from field experiments conducted under optimally-managed conditions and results from farmer yield contests.

The introduction provides examples of potential yield estimates derived from research trials (summarized in Table 1) and from yield contests. The highest yield from research was 9200 kg ha\(^{-1}\), but yield estimates were based upon a small number of plants and results are questionable. In Argentina, Kantolic and Slafer (2007) extended the photoperiod by 2 h after R3, and yields increased linearly as the number of days of photoperiod extension increased up to 8957 kg ha\(^{-1}\) (Fig. 2). The highest yield reported from farmer contests was 10,791 kg ha\(^{-1}\) (Cubbage, 2010). The following two subsections summarize estimates of potential yield using simple mathematical models and more complex, dynamic crop-simulation models.

4.1. Predicting yield potential with simple mathematical models

Simple models may be constructed by considering season-long values representing crop and seed growth. For example, by combining Eqs. (1), (3) and (5) we obtain:

\[
\text{Yield (g m}^{-2}\) = (BAR + \(\gamma\) + ISGR\(^{-1}\)) \times \text{ISGR + EFP}
\]

Eq. (6) can be simplified to:

\[
\text{Yield (g m}^{-2}\) = BAR \times \gamma \times \text{EFP}
\]  

(7)

One concern with this simplification is that it ignores the factor of time; in Eq. (3), BAR is measured from R1 to R5 while the EFP from Eq. (5) extends from approximately mid-R5 to R7. In Eqs. (3) and (5), ISGR is common to both expressions. However, in Eq. (3), ISGR is a proxy for the minimum amount of assimilates required per seed (\(A_g\)) as opposed to Eq. (5) where ISGR is a determinant of mass\(_{\text{seed}}\). Eq. (7) will only be appropriate when ISGR and BAR remain relatively constant across the time periods for Eqs. (3) and (5). Extreme changes in the source/sink balance do change ISGR (Egli et al., 1989), but short-term changes in the source/sink balance in the field do not impact ISGR (Egli and Leggett, 1976; Koller, 1971). Under well-irrigated conditions, BAR was linear in soybean from flowering until near physiological maturity (Mastrodomenico and Purcell, 2012). Likewise, the linear increase in HI during grain filling (Salado-Navarro et al., 1985b) indicates that \(\gamma\) is also constant, at least during the seed filling period. Eq. (7) may therefore provide a simple estimate of potential yield over a range of observed values for BAR and EFP.

Eq. (7) is attractive because it places emphasis on the integrative nature of BAR and because EFP is responsive to the environment. In addition, Eq. (7) avoids the criticism that seed m\(^{-2}\) is not a determinant of crop yield but is rather a “consequence of yield” (Sinclair and Jamieson, 2006). Sinclair and Jamieson (2006) argue that seed m\(^{-2}\) and yield are both constrained by availability and use of resources (e.g., solar radiation, water, nutrients) and that the crop adjusts seed m\(^{-2}\) to match resource availability. In Eq. (7), BAR and EFP increase or decrease, as appropriate, to match changes in resource availability.

Fig. 3 illustrates potential yield based upon the interaction between EFP and BAR as shown in Eq. (7) with the assumptions that ISGR and BAR are constant from beginning flowering through seed fill. The value of \(\gamma\) in Fig. 3 was also held constant at 0.5, which is near the minimum value reported by Egli and Zhen-Wen (1991) at high BAR. The range of BAR values shown in Fig. 3 encompasses the average soybean BAR reported in the literature under nonstresful conditions (19.8 g m\(^{-2}\) d\(^{-1}\), Van Roekel and Purcell, 2014) and the highest BAR values reported from highly-managed research plots (64 g m\(^{-2}\) d\(^{-1}\), Van Roekel and Purcell, 2014; Table 2). The range of EFP values in Fig. 3 is representative of typical soybean cultivars (Egli, 1998; Table 2). Fig. 3 shows that yields from 600 to 1000 g m\(^{-2}\) (6000 to 10,000 kg ha\(^{-1}\)) are within the range predicted given observed values of BAR, EFP, \(\gamma\), and the caveats described previously. It should be noted, however, that in the yield-contest report where BAR values of 64 g m\(^{-2}\) d\(^{-1}\) were measured during the late-vegetative to early-reproductive stages, yields did not exceed 8000 kg ha\(^{-1}\) (Table 1, Van Roekel and Purcell, 2014). This is considerably less than yields predicted by Eq. (7) for this BAR and with an EFP of 32 d or more (Fig. 3). The authors concluded that
it was unlikely that BAR values of this magnitude were continued throughout seedfill.

Another example of a simple mathematical model that has been used to predict potential yield considers yield as the product of the total biomass produced during the season (BM; g m\(^{-2}\)) and the apparent harvest index (HI\(_{\text{app}}\)) (Sinclair, 2004):

\[
\text{Yield (g m}^{-2}\text{)} = \text{BM} \times \text{HI}_{\text{app}}
\]

(8)

The apparent HI, HI\(_{\text{app}}\), represents the fraction grain to the above-ground BM at maturity and is, therefore, greater than the actual HI (HI\(_{\text{act}}\)), which accounts for fallen leaves and petioles (Schapaugh and Wilcox, 1980). In Eq. (8), total biomass production is modeled as the season-long sum of the product of radiation use efficiency (RUE; g MJ\(^{-1}\)) and intercepted solar radiation (IR; MJ m\(^{-2}\)) from emergence (i = 1) to physiological maturity (i = n):

\[
\text{BM} = \sum_{i=1}^{n} \text{RUE} \times \text{IR}
\]

(9)

RUE is similar to BAR, as both are measures of crop growth either over time (BAR) or over IR (RUE). Sinclair (2004) utilized Eqs. (8) and (9) assuming a HI\(_{\text{app}}\) of 0.55, RUE of 1.10 g MJ\(^{-1}\), and cumulative IR of 1200 MJ m\(^{-2}\) to predict a potential yield of 7260 kg ha\(^{-1}\). Because Eq. (8) is based on the total amount of BM produced during the season, a more conservative measure of estimating yield would be to use the HI\(_{\text{act}}\) rather than the HI\(_{\text{app}}\). A HI\(_{\text{app}}\) Value of 0.55 corresponds to a HI\(_{\text{act}}\) of about 0.45 (Schapaugh and Wilcox, 1980), resulting in a potential yield of 5940 kg ha\(^{-1}\). Considering recent reports of soybean RUE values up to 1.89 g MJ\(^{-1}\) (Van Roekel and Purcell, 2014) from a yield contest and assuming a cumulative IR of 1200 MJ m\(^{-2}\) and HI\(_{\text{act}}\) of 0.45, Eqs. (8) and (9) indicate a potential yield of 10,206 kg ha\(^{-1}\).

A different type of mathematical model has also been used to estimate potential soybean yield based upon the ratio of maize grain yield to soybean grain yield (Egli, 2008; Specht et al., 1999). The ratio of grain-yield improvement between maize and soybean ranges from 2.8- to 3.0-to-1. That is, the rate of maize yield increased 2.8 to 3.0 fold faster than soybean yield. Using the 2.8-to-1 ratio, the maximum yield potential of soybean was near 8000 kg ha\(^{-1}\) based upon observed maize grain yields near 22,500 kg ha\(^{-1}\) (Specht et al., 1999). Using this approach, a more recent report of a maize yield of 31,085 kg ha\(^{-1}\) (National Corn Growers Association, 2014) indicates that the maximum yield potential of soybean is near 11,000 kg ha\(^{-1}\).

### 4.2. Predicting yield potential with crop simulation models

Because of the difficulty of eliminating all biotic and abiotic stresses and ensuring ideal crop management, crop simulation models offer a means for assessing potential yields in given environments. Potential yield has been assessed using crop simulation models for wheat (*Triticum aestivum* L.), maize, and rice (*Oriza sativa* L.) (Lobell et al., 2009; Van Ittersum et al., 2013). However, there is little information available from simulation results on potential yield of soybean.

Boote (1989) evaluated yield potential for soybean grown in Gainesville, FL and Wooster, OH using an early version of the SOYGRO (Jones et al., 1989) crop simulation model. SOYGRO predicted potential yields of 5502 and 4525 kg ha\(^{-1}\) in Ohio and Florida, respectively. Boote (1989) concluded that yields greater than this would likely require unusual environments (e.g., high levels of solar radiation and relatively cool temperatures), an increase in CO\(_2\) concentration, and improved cultivar traits (notably a longer EFP and increased photosynthesis).

There is little information available from simulation results on potential yields for some of the environments where high soybean yields have been reported and with current climate conditions and genotype traits. An experiment conducted in Fayetteville, AR during 2012 and 2013 with optimized crop-growth conditions and 12 or 15 modern cultivars (Van Roekel and Purcell, unpublished data) was used to evaluate potential yield with two crop simulation models. Management in this experiment included high levels of fertility, 46-cm rows, deep tillage to remove rooting restrictions, early planting (11 April 2012 and 14 May 2013), irrigation as needed, fertigation (N, K, S) through a sprinkler system during seedfill, and strict pest-control measures.

The Sinclair soybean model (Sinclair et al., 2003) was the first approach used to evaluate potential yield for the environmental conditions of the 2-yr experiment conducted in Fayetteville. The default factors in the simulation for RUE (0.9 g MJ\(^{-1}\)), specific leaf N (2.5 g N m\(^{-2}\)), and the N accumulation rate (0.6 g N m\(^{-2}\)) were compared with the observed values measured in the experiment among cultivars for RUE (1.14 to 1.71 g MJ\(^{-1}\)), specific leaf N (2.81 to 3.22 g N m\(^{-2}\)), and the N accumulation rate (0.88 to 1.52 g N m\(^{-2}\)) d\(^{-1}\). Other inputs for the model included observed dates of V1, termination of leaf growth, and beginning of seedfill, and the rate of HI increase (DMAC). The observed yield in the experiment ranged from 4977 to 7144 kg ha\(^{-1}\) (Van Roekel and Purcell, unpublished results). Using the weather data from the experiment, the default settings generally under-predicted yield by 34% when averaged over years and cultivars. When the observed values for RUE, specific leaf N, and N accumulation rate were used in the model, predicted yields on average were slightly over predicted (3%) compared to the observed values. The conclusion from the modeling exercise was that model growth parameters should be adjusted in highly productive environments to more accurately simulate potential yield.

The DSSAT-CROPGRO v.4.5.1.023 (Jones et al., 2003) was also used to simulate soybean yield for the environmental conditions in Fayetteville. Water and nitrogen balance was activated in the simulations, and the management of irrigation and N fertilization was optimized to ensure no limitations. Cultivar traits related to photoperiod were calibrated to reproduce the observed days of main developmental stages. Other cultivar traits related to growth and leaf composition were modified to allow comparisons between predicted and observed values for soybean biomass, grain yield, biomass accumulation rate, and nitrogen accumulation rate (Table 9). Simulated scenarios also included predictions with
Observed biomass, grain yield (130 g kg\(^{-1}\) moisture), biomass accumulation rate (BAR), and nitrogen accumulation rate (NAR) averaged across cultivars and the highest and lowest cultivar values from an optimized management trial conducted in Fayetteville, AR during 2012 and 2013. Simulated results averaged across cultivars from DSSAT-CROPGRO simulations with modified parameters.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>kg ha(^{-1})</td>
<td></td>
<td></td>
<td>g m(^{-2}) d(^{-1})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed Average of cultivars</td>
<td>12,509</td>
<td>11,270</td>
<td>6746</td>
<td>6529</td>
<td>28.1</td>
<td>37.0</td>
<td>1.04</td>
<td>1.31</td>
</tr>
<tr>
<td>Highest cultivar</td>
<td>13,702</td>
<td>13,948</td>
<td>7690</td>
<td>7794</td>
<td>33.7</td>
<td>43.2</td>
<td>1.26</td>
<td>1.52</td>
</tr>
<tr>
<td>Lowest cultivar</td>
<td>11,089</td>
<td>8,109</td>
<td>5765</td>
<td>4977</td>
<td>23.0</td>
<td>25.4</td>
<td>0.88</td>
<td>1.08</td>
</tr>
<tr>
<td>Simulated Base simulation(^1)</td>
<td>11,180</td>
<td>10,582</td>
<td>7529</td>
<td>7030</td>
<td>16.8</td>
<td>16.7</td>
<td>0.66</td>
<td>0.65</td>
</tr>
<tr>
<td>(P_{\text{max}})</td>
<td>12,970</td>
<td>12,720</td>
<td>8814</td>
<td>8539</td>
<td>20.2</td>
<td>20.1</td>
<td>0.81</td>
<td>0.80</td>
</tr>
<tr>
<td>(P_{\text{max}}) + SLW = 35</td>
<td>13,414</td>
<td>13,291</td>
<td>9145</td>
<td>8952</td>
<td>20.8</td>
<td>20.6</td>
<td>0.82</td>
<td>0.81</td>
</tr>
<tr>
<td>(P_{\text{max}}) + 400 ppm CO(_2)</td>
<td>14,298</td>
<td>14,155</td>
<td>9761</td>
<td>9646</td>
<td>22.1</td>
<td>22.1</td>
<td>0.87</td>
<td>0.86</td>
</tr>
<tr>
<td>(P_{\text{max}}) + 39% (L_{\text{pmax}})</td>
<td>14,403</td>
<td>14,273</td>
<td>9856</td>
<td>9668</td>
<td>21.6</td>
<td>21.6</td>
<td>0.91</td>
<td>0.90</td>
</tr>
</tbody>
</table>

\(^1\) Base cultivar trait simulation conducted with a maximum leaf photosynthesis (\(P_{\text{max}}\)) of 1.03 mg CO\(_2\) m\(^{-2}\) s\(^{-1}\), 35.6% maximum protein in leaf, \(L_{\text{pmax}}\), specific leaf weight (SLW) of 27 g m\(^{-2}\), and 350 ppm CO\(_2\).

5. Management for reaching potential yield

At a location that had previously reported yields up to 10,791 kg ha\(^{-1}\) in yield contests (Cubbage, 2010), Van Roekel and Purcell (2014) reported BAR values up to 64 g m\(^{-2}\) d\(^{-1}\), and this site was noted for an intense fertility regime and large applications of poultry manure (Van Roekel and Purcell, 2014). Similarly, Flannery (1989) reported a maximum yield of 7955 kg ha\(^{-1}\) from a management system that included 12 Mg ha\(^{-1}\) of dairy manure applications every other year (Table 1). Cooper (2003) also reported high soybean yields, up to 7050 kg ha\(^{-1}\) from research with large annual applications of N, P, and K. However, Spaeth et al. (1987) reported yields of 6490 kg ha\(^{-1}\) with a lower input fertility system. As illustrated by the nutrient uptake and removal data from Flannery (1986), a 6786 kg ha\(^{-1}\) grain yield had a nutrient removal in the grain of 554 kg N ha\(^{-1}\), 55 kg P ha\(^{-1}\), and 392 kg K ha\(^{-1}\). Thus, high soybean yields require large nutrient inputs. These nutrients can be attained via applications of manure (Flannery, 1989; Van Roekel and Purcell, 2014), inorganic fertilizer (Cooper, 2003; Flannery, 1989), and/or from fertile soils (Spaeth et al., 1987). It should also be noted that current fertility recommendations are likely inadequate for supplying a crop with yields exceeding 5000 kg ha\(^{-1}\), and to achieve yields of this magnitude or greater it may be necessary to supplement N \(_2\) fixation with N fertilizer (Salvagiotti et al., 2008).

Furthermore, with normal soybean water requirements peaking at approximately 7.6 to 8.1 mm d\(^{-1}\) (Kranz and Specht, 2012; Heatherly, 1986; Heatherly and Elmore, 2004), it is evident that soil-moisture management is another critical component for reaching potential yield of soybean. Assuming a season-long vapor pressure deficit of 1.5 kPa and a harvest index of 0.40, a soybean yield of 6000 kg ha\(^{-1}\) would require approximately 500 mm of water for transpiration, which ignores any water loss from soil evaporation (Sinclair and Rufty, 2012). Water requirements would increase proportionately as yields increase. Isoda et al. (2010) utilized a drip irrigation system and total water inputs from precipitation and irrigation totaled 465 mm in the highest yielding year of research. Flannery (1989) also employed a drip irrigation system, with precipitation and irrigation amounts totaling 782 mm in the highest yielding year. The same irrigation system was used to...
delivered approximately 66% of the fertilizer inputs in the growing season. Cooper (2003) used a sprinkler irrigation system to ensure the crop received at least 50 mm water week$^{-1}$ from precipitation and irrigation. Spaeth et al. (1987) relied solely on rainfall and improved their water management by installing tile drainage to remove excess water from their converted rice paddies.

Pest control was a key factor in eliminating biotic stresses for Flannery (1989), Van Roekel and Purcell (2014), and Cooper (2003). Flannery (1989) utilized seed treatments, soil and foliar insecticides, and fungicides to control pests throughout the season. Cooper (2003) and Van Roekel and Purcell (2014) reported routine fungicide applications and periodic insecticides as needed. Spaeth et al. (1987), however, made no mention of increased pest control practices.

With the exception of Spaeth et al. (1987), these examples of high yield research utilized row widths less than 75 cm (Table 1). Narrow rows increase light interception and BAR (Board et al., 1992; Board and Harville, 1996) and decrease evaporation from the soil.

Finally, early planting (Bastidas et al., 2008; Salmeron et al., 2014; Wilcox and Frankenberger, 1987) and warm early spring temperatures (Cooper, 2003) induce earlier flowering, lengthen the reproductive period, and increase solar radiation interception and BAR during seed filling. Rowntree et al. (2014) compared early (~May 1) and late (~June 1) planting dates for 57 MG 3 cultivars released between 1923 and 2008 and found that the rate-of-yield increase was 23% greater (1.74 vs 1.41 kg ha$^{-1}$ yr$^{-1}$) for the early-planted treatment than the late-planted treatment. Associated with the yield increase with early planting was a shortening of the vegetative stages (V1 to R1) and a lengthening of the reproductive stages (R1 to R7).

Selecting the correct combination for a particular location of row spacing, planting date, and MG such that the crop has a closed canopy and is beginning to flower close to the summer solstice provides an environment for maximizing BAR and for a relatively long photoperiod, which will extend the reproductive period (Rowntree et al., 2014). Kantolic and Slafar (2007) found that extending the photoperiod by 2 h from 1 to 35 days during the R1 to R3 period increased linearly the R3 to R6 period by about 5 to 27 days (Fig. 2). Associated with an increase in the R3 to R6 period was an increased in yield up to 8957 kg ha$^{-1}$.

6. Increasing soybean yield potential

To harness genomic information for crop improvement there is a need to characterize fully agronomically meaningful traits and couple this information with genetic information. Fifteen years ago, Miflin (2000) pointed out the ‘phenotype gap’ between genetic and phenotypic information, and the gap has only widened since that time.

Previous efforts to increase soybean yield by selection for EFP, ISGR, mass$_{seed}$, and seed m$^{-2}$ have been unsuccessful (Donald and Hamblin, 1976; Egli, 1998). The compensation that occurs among these traits illustrates that there are multiple paths resulting in similar yields. The past failures to increase yield by selecting for ISGR and EFP may be related to: (1) the sheer complexity of these traits and number of alleles that contribute to these traits and (2) evaluation of traits under non-ideal environments that resulted in poor heritability and large genotype by environment interactions. Identification of alleles or quantitative trait loci (QTLs) associated with these traits from appropriate segregating populations when grown under optimum growth conditions has the potential of closing the phenotype gap for potential yield. Subsequent research will need to assess the effects of these traits on yield and stability in non-ideal environments including efforts to understand the genotype by environment interactions.

Currently, characterization of soybean yield largely remains at the empirical level with little understanding of the genetics of underlying physiological traits that contribute to yield. For example, 172 QTLs have been reported for soybean yields and 217 QTLs for mass$_{seed}$, but there are no QTLs reported for ISGR, BAR, RUE, or HI (www.soybase.org, verified 28 September 2014). There are also five QTLs reported for the seed fill period (analogous to EFP) but these represent the period from R1 to R8 without considering distinctions between flower/pod formation and seed filling.

The identification of QTLs associated with ISGR, EFP, and BAR requires that variability exists for these traits, and that they are heritable. The ISGR is a relatively stable trait (Egli et al., 1978b; Egli and Leggett, 1976; Koller, 1971; Meckel et al., 1984; Westgate et al., 1989) that is independent of the EFP (Egli et al., 1981, 1984; Egli and Wardlaw, 1980; Meckel et al., 1984). Genotypic differences in soybean are well established for ISGR (Egli et al., 1978b; Egli 1993). There are no published reports of heritability of ISGR in soybean, but it appears that ISGR is an intrinsic trait related to cotyledon cell number (Egli, 1998; Egli et al., 1989) that is closely associated (r=0.93) with mass$_{seed}$ (Egli et al., 1981). Heritability of mass$_{seed}$ in soybean ranges from about 0.71 to 0.73 (Jauregui et al., 2011; Panthee et al., 2005) and a similar heritability is anticipated for ISGR.

In contrast, the EFP is affected by environmental conditions including drought, high-temperature stress, and photoperiod (Egli et al., 1984; Egli and Wardlaw, 1980; Gbikpi and Crookston, 1981; Kantolic et al., 2013; Meckel et al., 1984; Thomas and Raper, 1976). There are genotypic differences in EFP with heritability estimates ranging from ~0.20 to 0.83 (as summarized by Egli, 1998). The large environmental effects on EFP generally result in corresponding changes in yield, provided that seed number is set and that ISGR is not affected. For example, a decrease in the EFP due to drought during seed filling is associated with decreased yield through a decrease in mass$_{seed}$ (Meckel et al., 1984; Shaw and Laing, 1965; Specht et al., 1986). Evaluating EFP under optimum growing conditions will greatly decrease environmental effects and improve heritability. Ideally, the EFP should be extended by beginning reproductive growth earlier in the season when solar radiation is highest and by maintaining the maturity date (Cooper, 2003; Kantolic et al., 2013; Kantolic and Slafar, 2005, 2007). Cooper (2003) concluded that delayed flowering presents a barrier to yields exceeding 4500 to 5500 kg ha$^{-1}$ and that flowering earlier provides opportunities for yields to exceed 6000 kg ha$^{-1}$. Rowntree et al. (2013, 2014) also demonstrated a clear yield advantage in MG 3 cultivars for early planting that was associated with a decrease in the vegetative period and an extension of the reproductive period. Yield responses due to an extended photoperiod from R1 to R3 (Fig. 2; Kantolic and Slafar, 2007) point to management and genetic possibilities for increasing yield by prolonging seed filling.

Van Roekel and Purcell (2014) documented large genotypic differences in RUE, HI, BAR, and N accumulation rates among elite soybean cultivars when grown under optimum conditions, and both BAR and N accumulation rates far exceeded values found in the literature. Rotundo et al. (2014) also documented diversity and alternate strategies among high yielding cultivars for N use efficiency (NUE), N harvest index, seed N, and HI. Rotundo et al. (2014) suggested that breeding for increased NUE and HI with the average N accumulation rate held promise to increase yield ~500 kg ha$^{-1}$ in their environment. This indicates that the combination of increased N accumulation rates, NUE, and HI may hold potential for increasing yield even further. Differences in crop growth potential among genetically diverse genotypes under conditions of unlimited nutrients and water availability have not been explored but may provide key insights into allelic combinations specific for high input environments.
7. Conclusions

The theoretical framework of Charles-Edwards (1984) as developed by Egli (1998) for soybean yield determination allows for a mechanistic approach to understand physiological parameters associated with seed \( m_{\text{seed}} \), mass\(_{\text{seed}} \) and the resources required for soybean to reach its potential yield. Potential yields may be estimated using results from well-managed research trials, farmer yield contests, simple mathematical relationships based upon crop growth characteristics and environmental resources, and from more complex crop simulation models. These different methods of estimating potential yield predict values ranging from about 7500 to 11,000 kg ha\(^{-1}\). Full characterization of crop growth, development, and the environment under optimum conditions is needed to better define yield potential and to provide a clear understanding of mechanisms that lead to yields of this magnitude.

The inability to increase soybean yield in previous research by manipulating seed \( m_{\text{seed}} \) and mass\(_{\text{seed}} \) demonstrate compensatory mechanisms to adjust assimilate supply with seed growth requirements. Management of soybean for optimum growing conditions, however, may provide opportunities for combining alleles for ISGR, EFF, and BAR that result in setting large numbers of seeds and an extended seedfill period. Identification of positive alleles for these traits under ideal growing conditions will require careful phenotypic evaluation coupled with molecular breeding methods.

References


