Structural and Fractal Dimensions are Reliable Determinants of Grain Yield in Soybean

A.A. Jaradat1, D. Surek2 and D.W. Archer1
1 United States Department of Agriculture (USDA)-Agricultural Research Service (ARS), 803 Iowa Avenue, Morris, MN 56267, USA
2 Soil and Water Research Institute, Ankara, Turkey
* Corresponding author (abdullah.jaradat@ars.usda.gov).

Abstract

Soybean [Glycin max (L.) Merr.] plants grown under five management strategies differed significantly in their geometric structures, and were classified with 75 to 100% correct classification, based on differences in their fractal dimension (Do), midday differential canopy temperature (dT), and canopy light penetration [Log(I/Jo)]. Single soybean plants grown under a conventional system using moldboard tillage developed complex geometric structures, with significantly larger Do (1.477) values and grain yield (11.2 g per plant) as compared to plants grown under an organic system with strip tillage (Do =1.358, and grain yield = 2.32 g per plant). Across management strategies, Do of single plants was a function of stem perimeter, circularity, and volume, and plant dry weight; whereas grain yield m⁻² was a function of Do, plant dry weight and volume, and stem circularity. Knowledge of how plants respond to single and multiple management strategies will help agronomists develop better predictive models and will help farmers refine management practices to optimize yield.

Introduction

Plant size and architecture are important factors in determining crop productivity [1]; however, researchers are faced with the problem of developing reliable models for plant geometric structure and its relationship to yield and productivity, especially for plants with complex structures such as soybean [2, 3]. One approach to solving this problem is to use fractal analysis to provide new avenues of understanding the functional implications of the branching patterns in relation to optimum space exploration by plants [1]. The fractal dimension (Do) is considered [2] an effective tool for quantifying plant structure, measuring the structural response to cultural practices and modeling plant canopies.

The reproductive period is most sensitive to altered source strength and crop growth rate since it is the time during which important yield components are formed [4]. Changes in fractal dimension of several crops (e.g., corn and soybean) were found to be highly significant over time [2] reflecting the level of complexity in skeletal structure of single plants as the growth stages advanced. Several methods were used to quantify the relationships between soybean growth and development using growth analyses; however, limited information exists on the response of soybean’s fractal dimension to management strategies. The objectives of this 2-yr study were to quantify the impact of management strategies on soybean’s structural dimensions and Do during the reproductive growth stage, and to identify predictors of grain yield (gm⁻²).

Materials and methods

Digital imagery [5] and analysis procedures [2, 6] were used to capture, measure, and statistically analyze several morphological traits of individual soybean plants grown under five management strategies, i.e., combinations of conventional (C) or organic (O) cropping system, conventional (C) or strip (S) tillage, recommended fertilizer rate (Y) and 2- or 4-yr crop rotation (Fig.1); for example, CCY4 is the management strategy with conventional cropping system (C), conventional tillage (C), with N fertilizer based on soil analysis (Y) and 2- or 4-yr crop rotation. Light interception by plant canopy [log(I/Jo)] and midday differential canopy temperature (dT) were estimated as
described by Jaradat et al. [7]. The fractal analysis
procedure employed the box count concept as
outlined by Foroutan-pour et al. [2], where the
fractal dimension (Do) is constrained to be in the
range of 1.0 ≤ Do ≤ 2.0. A value of 1.0 indicates
that the image is completely differentiable and that
of 2.0 indicates that the image is irregular. Yield
components (listed in Fig. 1) were measured on the
same plants subjected to fractal dimension analysis
and were used in subsequent statistical analyses.

The principal components (PC) option in the
Nonlinear Iterative Partial Least Squares module
and canonical discriminant (CD) analyses were
used to analyze the standardized structural
dimensions of individual soybean plants. Linear
combinations of the original variables (i.e., leaf
and stem structural dimension, in addition to
Log(I/Io), Do*LAI [2], and dT) that account for as
much of total variation in the data set as possible
were constructed. Canonical discriminant analysis
was used to find the dimensions along which plants
grown under different management strategies differ
and to find classification functions to predict group
membership on the first (R1) and the second (R2)
canonical discriminant functions. The CD module
was used to assess the variation among plants
produced under two or more management
strategies relative to the average variation found
within all plants regardless of management
strategies [8].

The impact of plant structural dimensions on Do
and grain yield gm⁻² was studied using the
regression option in artificial neural networks
(ANNs), then the models were subjected to
sensitivity analysis to evaluate the relative
importance of each variable in explaining Do or
grain yield gm⁻². In this analysis, each predictor
was treated in turn as if it were not available in the
ANN model and the average value of that predictor
was used. A sensitivity ratio was calculated by
dividing the total ANN error when the predictor
was treated as “not available” by the total ANN
error when the actual value of the predictor was
used. If the ratio is >1.0, then the predictor made
an important contribution to Do or grain yield gm⁻².
The higher the ratio, the more important the
predictor is [8]. Additionally, we calculated the
correlation coefficient (r) and a ratio between the
standard deviation (SD-ratio) of original and model
data; higher r values and lower SD-ratio values are
indicators of better model performance [8].
Sensitivity analysis was performed by generating
response curves for each predictor to study its
relationship with Do and with seed yield (gm⁻²),
while all other predictors were set at their mean
value.

Results

Principal components regression

Slightly more than 50% of total variation in the
whole data set was explained by the first two
principal components (PCs; Fig. 1). Distinct
separation between plants grown under organic and
conventional systems was achieved on the basis of
single plant characteristics, most of which were
positively associated with conventional cropping
system, conventional tillage and fertilizer
application. Thousand-seed weight was the only
variable associated with organic cropping system,
strip tillage and no fertilizer treatment. Leaf area
loaded on the third PC and accounted for additional 10% of total variance (data not
presented). Grain yield m⁻² was positively and
closely associated with the fractal dimension, pods
m⁻², and stem circularity, and to a lesser extent
with the remaining plant structural dimensions on
PC1 which explained 33% of total variation.

A cumulative variance of 43% was explained
by all independent variables in the PC regression
(Fig. 2). Larger Do values, when multiplied by leaf
area index (LAI) [2], were positively associated
with conventional cropping system and
conventional tillage; whereas large values of dT
and log(I/Io) (i.e., less light interception by plant
canopy) were associated with organic cropping

![Fig. 1. Joint plot of six components of management
strategies, ten soybean characteristics and the
fractal dimension on the first two principal
components in PC analysis.](image-url)
system (Fig. 2). The large LAI*Do values were loaded positively on PC1 along with most stem and leaf characteristics; however, stem width, leaf width, leaf perimeter, and stem perimeter were more closely associated with LAI*Do than the remaining stem and leaf structural dimensions.

A total of 51% of total variation in the dependent variables, accounted for by the first two PCs, explained 87% of total variation in grain yield per plant, which ranged from 2.32 (OSY4) to 11.2 g (CCY4) with significant differences among all management strategies. There were significant differences in grain yield per plant due to the tillage component, whether associated with conventional or organic systems, and due to crop rotation (2- vs. 4-yr) whether associated with conventional or strip tillage. Plants grown under CCY4 produced the largest grain yield (11.2 g), followed by CSY4 (9.82 g); whereas those grown under OCY4 and OSY4 produced the least (5.37 and 2.32 g per plant, respectively).

**Discriminant analysis**

Discrimination among plant samples grown under five management strategies (Fig. 3) was clearly achieved using plant structural dimensions and three derived statistics (i.e., dT, Do and log(I/lo)). Two canonical discriminant roots accounted for a total of 92% of total variation and discriminated among plant samples with 75.0 (CSY4) to 100.0% (CCY2 and OSY4) correct classification. The first canonical root (R1) was dominated by leaf circularity (i.e., ratio of minor to major axes), leaf area, log(I/lo), Do and dT, accounted for the majority of variation (84%) and totally separated samples grown under organic system (i.e., OCY4 and OSY4, with 95.5 and 100.0% correct classification, respectively) from those grown under conventional system (CCY2, CCY4 and CSY4, with 100.0, 83.3, and 75.0% correct classification, respectively).

Separation between the latter three groups along the second canonical root (R2), with 8% of total variation, ranged from 75 (CSY4) to 83.3% (CCY4). CAN2 was dominated by stem-related variables and there was clear overlap between plants grown under CCY4 and CSY4, on one hand, and those grown under CCY2. The three derived statistics (i.e., Do, dT and log(I/lo)) were closely associated with leaf circularity and leaf area, whereas stem structural dimensions were independent.

**Prediction of Do**

Calibration and validation regression models were developed to predict Do as a function of dT are presented in Table 1. Correlation coefficients (r values) between measured and predicted grain yield using Do as a predictor were non-significant during the first two growth stages (data not presented); however, r values increased steadily from 0.74 (RGS3) to 0.96 as the plants approached maturity (RGS6); the respective r-values for the validation models were smaller (0.65 to 0.94).
albeit significant (p<0.05); however, the validation models performed very poorly during the first two reproductive growth stages. The intercept and slope the regression models increased steadily as plants approached maturity, the intercept approaching zero and the slope approaching unity.

Table 1. Calibration (C) and validation (V) partial least squares (PLS) regression models predicting soybean plant fractal dimension (Do) as a function of midday differential canopy temperature (dT) at four reproductive growth stages (RGS3 – RGS6, a and b are intercept and slope of regression models, respectively; *, p<0.05).

<table>
<thead>
<tr>
<th>Growth stage</th>
<th>PLS regression model</th>
<th>C</th>
<th>r</th>
<th>V</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGS3</td>
<td>a</td>
<td>0.68</td>
<td>0.74*</td>
<td>0.78</td>
<td>0.65*</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.52</td>
<td></td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>RGS4</td>
<td>a</td>
<td>0.53</td>
<td>0.79*</td>
<td>0.58</td>
<td>0.74*</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.63</td>
<td></td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>RGS5</td>
<td>a</td>
<td>0.29</td>
<td>0.89*</td>
<td>0.34</td>
<td>0.86*</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.79</td>
<td></td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>RGS6</td>
<td>a</td>
<td>0.11</td>
<td>0.96*</td>
<td>0.14</td>
<td>0.94*</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.92</td>
<td></td>
<td>0.89</td>
<td></td>
</tr>
</tbody>
</table>

Neural network and sensitivity analyses

A Multi-Layer Perception Neural Network (MLPR-NN) with 9:9-13-7-1:1 layers was the best neural network to predict soybean fractal dimension (Do), and a General Regression Neural Network (GR-NN) with 13:13-43-2-1:1 neurons was the best to predict soybean grain yield gm^{-2} as a function of three plant traits and the fractal dimension (Do). The MLPR neural network identified four independent variables with significant contribution in predicting both Do and grain yield m^{-2} (Table 2). Plant dry weight was an important variable in predicting Do and grain yield m^{-2}. A much simpler multi-layer perception network, with 13 hidden layers, was capable of predicting Do as compared to the more complex general regression neural network, with 43 hidden layers, needed to predict grain yield gm^{-2}. However, almost correlation coefficients were found between measured and modeled data for Do (0.87) and grain yield gm^{-2} (0.89) (Table 2).

The relationship between four predictors (Table 2) and Do was quantified and a regression equation was developed to predict Do as a function of each predictor while holding each of the remaining predictors at its mean value. Plant dry weight, stem volume, stem circularity and stem perimeter (Fig. 4) displayed different, albeit large and significant, effects on Do. The quadratic effect of stem volume was not significant. A plant dry weight of 6-7 g is capable of producing a maximum Do of 1.45-1.46; however, Do did not respond positively to any further increases in the plant dry weight beyond this level.

Table 2. Statistics of the Multi-Layer Perception Neural Network (MLPR-NN) with 9:9-13-7-1:1 layers predicting soybean fractal dimension (Do), and of the General Regression Neural Network (GR-NN) with 13:13-43-2-1:1 neurons predicting soybean grain yield gm^{-2} as a function of three plant traits and the fractal dimension (Do).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Ratio and (rank)</th>
<th>Test statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant dry weight, g</td>
<td>1.57(1)</td>
<td>1.6(2)</td>
</tr>
<tr>
<td>Plants volume</td>
<td>1.49 (2)</td>
<td>1.35 (3)</td>
</tr>
<tr>
<td>Plant circularity</td>
<td>1.35 (3)</td>
<td>1.35 (4)</td>
</tr>
<tr>
<td>Plant perimeter</td>
<td>1.14 (4)</td>
<td>1.73 (1)</td>
</tr>
</tbody>
</table>

Mean 1.43 172
S.D. 0.06 56.0
S.D. 0.65 0.63
Ratio r-value 0.87 0.89

Plant dry weight was the most important variable in predicting Do, followed by plant volume, plant circularity and plant perimeter; whereas Do was the most important variable, followed by plant dry weight, plant volume and plant circularity, in predicting grain yield gm^{-2}. The SD-ratios for Do (0.646) and for grain yield m^{-2} (0.632) were relatively similar.
Fig. 4. Sensitivity analyses of plant dry weight, stem volume, stem circularity and stem perimeter as predictors of fractal dimension ($\text{Do}$) of soybean plants.

On the other hand, $\text{Do}$ responded linearly to plant volume and, in a piecewise fashion, to stem circularity (i.e., ratio of minor to major axes, with a breakpoint at $\text{Do}=1.424331$) and stem perimeter (with a breakpoint at $\text{Do}=1.4164$). Similarly, a nonlinear regression equation was developed to predict grain yield ($\text{gm}^{-2}$) as a function of each predictor (i.e., $\text{Do}$, plant dry weight, plant volume, and stem circularity, Table 2) while holding each of the remaining predictors at its mean value (Fig. 5). Positive and significant relationships were found among grain yield and each predictor, and the nonlinear portion of the regression equations was significant except for $\text{Do}$.

Discussion

Short growing seasons in the upper Midwestern USA present serious time limitations on crop growth, in which soybean crop needs to establish and maximize canopy coverage rapidly to exploit available light [3]. Crop plants have been shown to adjust their architectural traits (Table 2) in response to management practices [2] and plant architecture, as characterized by $\text{Do}$, has been shown to impact grain yield in many crops [1].

Different management practices created a range of microenvironments in which soybean plants developed different architectures, as reflected by their $\text{Do}$, $dT$ and $\log(l/I_o)$ values and on the large percentage of correct classification (75.0-100.0%). Further evidence on how grain yield responded to adjustments in plant architecture, which in turn responded to components of different management practices, is quantified in Fig. 1. The largest grain yield per plant (11.2 g) was positively associated with $\text{Do}$, conventional system, and conventional tillage, and was a result of maximum plant growth and development under the favorable conditions created by the CCY4 management strategy (Fig. 3).

The PLS regression models, especially during RGS3 to RGS6 (Table 1), succeeded in predicting $\text{Do}$ as a function of midday differential canopy temperature ($dT$), the value of which depends on air temperature, but will differ from it due to canopy characteristics, thermal characteristics and thermal conditions near the soil surface [10]. Reliability of the predictive equations (expressed as r-values) increased as the plants grew and changed the microenvironment within the canopy, and with time.
identified important Do and seed yield predictors using ANN models in an attempt to develop timely
management practices that may help create optimum plant geometric structures (expressed as Do) capable of maximizing light interception and midday differential canopy temperature, and thus producing the largest grain yield.

References


