International Congress on Environmental Modelling and Software

7th International Congress on Environmental Modelling and Software - San Diego, California, USA - June 2014

Jun 16th, 9:00 AM - 10:20 AM

A k-means clustering approach to assess wheat yield prediction uncertainty with a HYDRUS-1D coupled crop model

Derek Groenendyk
University of Arizona, drockg@email.arizona.edu

Kelly Thorp
USDA-ARS, kelly.thorp@ars.usda.gov

Ty Ferré
University of Arizona, ty@hwr.arizona.edu

Wade Crow
USDA-ARS, wade.crow@ars.usda.gov

Doug Hunsaker
USDA-ARS, doug.hunsaker@ars.usda.gov

International Environmental Modelling and Software Society (iEMSs)
7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA,
Daniel P. Ames, Nigel W.T. Quinn and Andrea E. Rizzoli (Eds.)
http://www.iemss.org/society/index.php/iemss-2014-proceedings

A k-means clustering approach to assess wheat yield prediction uncertainty with a HYDRUS-1D coupled crop model

Derek Groenendyk[a], Kelly Thorp[b], Ty Ferré[a], Wade Crow[b], Doug Hunsaker[b]
[a]University of Arizona (drockg@email.arizona.edu, ty@hwr.arizona.edu), [b]USDA-ARS (kelly.thorp@ars.usda.gov, wade.crow@ars.usda.gov, doug.hunsaker@ars.usda.gov)

Abstract: Soil moisture, especially under drought conditions, is a factor that is known to impact crop yield predictions. Crop growth models used to make these predictions rely on soil texture estimates, which influence simulated soil moisture and ultimately crop growth. The purpose of this research was to implement a k-means clustering approach to address the uncertainty of the soil texture estimates. By grouping similar soil textures based on their simulated responses, clustering reveals how soil texture uncertainty may impact yield estimates. Wheat growth simulations were conducted using a HYDRUS 1D and coupled crop model for soils defined on the USDA soil texture triangle. A k-means clustering algorithm was applied to the simulated biophysical data for each soil texture. Resulting clusters were different from traditional soil type classifications. The k-means clustering approach proved useful for investigating the relationship to soil texture that crop yield may have. This research shows that the impact of soil texture variation should be considered when conducting crop growth simulation for the purposes of yield forecasting.

Keywords: modeling; HYDRUS; clustering; yield; soil moisture

1 INTRODUCTION

As a result of increased food demands for a growing population, large scale crop production must strive for maximum productivity while minimizing use of water and energy resources. Shifting demands for crop commodities as renewable sources of energy have further driven the need for increased production. As the need for increased production continues to be a global concern for economic stability, food and energy security, the importance of anticipating accurate crop production levels will intensify as well.

Crop progress is currently monitored throughout the growing season based on weather conditions, management practices, and crop scouting reports. Collection of this information allows for forecasting of crop yield. Yield forecasts are widely used today to predict yearly production for food and economic forecasting, making management decisions (crop type, fertilizer or water scheduling) and even policy decisions (Baez-Gonzalez et al. 2005; Bannayan, Crout, and Hoogenboom 2003; Moriondo, Maselli, and Bindi 2007; Prasad et al. 2006).

Some of these forecasts are created using statistically based methods (Statistics Division, National Agricultural Statistics Service 2012) while others are based on biophysical computer models (Jones et al. 2003) or remote sensing techniques (Moriondo, Maselli, and Bindi 2007; Sakamoto, Gitelson, and Arkebauer 2014). Regardless of the method used to create the forecast or its ultimate application, the main goal is to produce accurate estimates of crop yield.

The computer models that are used to simulate crop growth typically simulate soil hydrologic responses as well. Water stress is one of the primary factors limiting crop growth. Typically, water stress is calculated when the potential demand for water lost through crop transpiration and soil water evaporation is higher than the amount of water that can be supplied by the soil through the root system. Water stress or water availability is a primary management concern especially in water-limited regions.

The recent droughts in the central and southwest United States and the resulting impacts to the food markets ,highlight the need for improved yield estimation techniques that focus on water availability (Kaul, Hill, and Walthall 2005).

Understanding the water response or water availability of a soil and the impact this will have on crop growth and ultimately yield is critical. For this we can use crop model simulations of plant growth to predict yield and to guide management decisions. It is known that soil conditions are important to such crop models (Hansen and Jones 2000; de Wit, Boogaard, and Van Diepen 2005), because soil texture largely determines the soil's capacity to hold water and provide moisture for crop transpiration processes.

The hydrologic equations used to simulate water flow also require soil texture information, i.e. van Genuchten and Brooks and Corey models. However, consideration of soil texture variation or an in depth understanding of the hydrologic response in context of soil texture in a universal sense isn't a common and functional part of methods used to study yield variation. The crop models provide a useful system in which the relationship and behavior of yield, soil moisture and soil texture can be studied.

It is common to define soil textures using the relative distribution of soil particle sizes present in the composite soil material. Size limits are used to establish three types of particles: sand, silt and clay. To universally characterize soil texture, classification schemes have been developed that define soil groups using the percentages of sand, silt and clay. The total percentage of these particles in a composite soil material are used to establish specific soil types or classes. The limits that define grain size of sand, silt, and clay particles vary, resulting in a lack of continuity or agreement between taxonomies or classification systems (Bormann 2010; Minasny and McBratney 2001).

Figure 1. USDA soil texture triangle

The varied agreement of the mechanical limits of soil particles and classes has led to two opposing taxonomies or classification systems: one developed by a Swedish soil scientist (Atterberg 1905) which was adopted by the International Society of Soil Science (ISSS) in 1925 and the United States Department of Agriculture (UDSA) soil texture triangle (Davis and Bennett 1927) adopted by the Food and Agricultural Organization (FAO) in 1950. Not only do the schemes vary in their mechanical limits but also in their layout, using either two axes or three axes to visually express the boundaries of the soil classes. Classifying soils is important for understanding soil structure in the physical sense. The understanding of hydrological behaviors and its significance in these classifications are less well known and have only minimally been considered.

Soil texture is frequently used in pedotransfer functions to define soil hydraulic parameters so models can simulate hydraulic responses of the soil. It is common to assign soil hydraulic parameters to each soil class within a classification system by averaging the parameter values across the soils within the assigned soil classes. This common and yet often overlooked approach makes the assumption that soils within a class are hydrologically similar, meaning that they exhibit behavior comparable to other soils within the same class (Twarakavi, Šimůnek, and Schaap 2010). Averaging hydraulic parameters requires an assumption that within a soil class the hydraulic parameters and the hydraulic behaviors are similar to one another, as compared to those soils within other classes.

Recently, clustering algorithms have been used successfully to form classifications based on hydraulic values (Bormann 2010; Twarakavi, Šimůnek, and Schaap 2010). The hydraulic based classifications formed using the clustering algorithms were shown to differ from those based on the soil texture approaches. It is also clear that the hydraulic parameters are different among themselves and among soil texture classifications. Twarakavi et al. (2010) considered free drainage only (no evapotranspiration) when determining the soil hydraulic parameters for clustering. Bormann (2010) used an annual water balance approach to classification. These studies provide guidance for developing and applying a classification approach to further take into account variations and uncertainty related to soil texture variation in yield and soil moisture when plants are present.

With this is mind, it is not the goal of this research to replace the texture triangle. Rather, the main focus of this research is to address some of the shortcomings that physically based soil texture approaches, in particular the USDA texture triangle, share by proposing a simulation methodology useful for creating classifications based on soil moisture and crop growth to address uncertainty in yield forecasting. The specific objectives were to a) couple a dynamic crop model with a water flow model b) create classifications based on biophysical variables using k-means clustering.

2 METHODS

The research focused on the USDA texture triangle to demonstrate hydrologic classification and comparison techniques. This approach implemented simulation modeling to explore hydraulic responses as it is neither cost-effective nor timely to retrieve actual samples for every soil on the USDA soil texture triangle. To cover the entire range of soils, the USDA soil triangle was broken down into 2% increments, resulting in 1326 unique soils. These 1326 soils represented virtual columns of soil whose hydrological response was simulated using HYDRUS-1D and coupled crop model software (Simunek, Van Genuchten, and Sejna 2005). Cluster analysis was performed on various biophysical variables over a growing season with related forcing data.

2.1 Model Development

HYDRUS-1D is a vertical hydrologic simulation model (Simunek, Van Genuchten, and Sejna 2005). The solution to soil water redistribution is found through numerical iteration and convergence of Richards' equation (Richards, 1931). HYDRUS-1D considers time varying atmospheric and plant growth conditions. However, the plant growth values must be provided *a priori* and are not dependent upon changes within the model parameters. This means that soil hydraulic parameters as well as meteorological forcing data must match the conditions under which the plant growth data was collected. To achieve the objectives of this research, a dynamic crop growth model was written in Fortran and compiled with the original HYDRUS-1D code to allow for dynamic plant growth simulations. The crop growth model was coupled to HYDRUS through the calculation of water fluxes at the surface and in the root zone.

The crop growth model relies on heat units to mark key crop growth stages. Once emergence has occurred biomass accumulation is based on a daily potential. Daily maximum potential biomass growth is set by a fraction of photosynthetically active radiation (PAR). PAR is a function of incoming radiation and leaf area index (LAI). Leaf area index (LAI) and ground cover fraction are exponential factors of biomass. Actual daily biomass growth is reduced from potential growth based on plant stress factors. Plant water stress is a function of an evapotranspiration ratio (actual over potential) and temperature stress related to average daily temperature. Crop height is an exponential function of biomass. Penman-Monteith evapotranspiration is related to the crop height. Ground cover is used to partition evapotranspiration into soil evaporation and plant transpiration and to determine plant albedo. Rooting depth is calculated using a logistic growth function based on the length of the growing season and was confirmed using data from destructive field samples. Yield is a fraction of end of the season biomass. LAI can be used to calculate interception however was ignored here because flood irrigation was used and the rainfall was minimal in the arid environment where data was collected for model calibration.

2.2 Virtual Soil Columns

Columns representing vertical homogeneous soil profiles were created for each unique soil. The columns had a length of 200 cm and were separated into 1 cm layers. The boundary condition at the top of the column was an atmospheric boundary with an allowable ponding surface layer. At the bottom the boundary condition was assumed to be free-drainage. Initially the profile was set to field collected soil moisture values. For each one of the 1326 soils, van Genuchten-Mualem hydraulic soil parameters were found using the ROSETTA software (Schaap, Leij, and Vangenuchten 2001).

2.3 Model Parameterization

Minimum HYDRUS-1D inputs included soil hydraulic parameters and initial profile conditions. Time-varying boundary conditions for surface fluxes (i.e. precipitation and evaporation) can be added if desired. When crops are included in simulations extra inputs are required that include the following meteorological conditions: radiation, daily maximum temperature, daily minimum temperature, relative humidity, wind speed. Also considered for crops are root water uptake parameters. The crop growth model that was added also includes several parameters related to plant phenology and growth factors: harvest index, biomass conversion factor, heat units at emergence, heat units at maturity, fraction of season when leaf senescence and LAI decline begins, maximum LAI, and maximum crop height. The complete set of crop model parameters were calibrated against experimentally collected data.

The experimental data used to run simulations and calibrate the complete set of crop model parameters was collected by Hunsaker et al. (2007a, 2007b). The data was collected during two irrigated wheat experiments conducted in the winters of 2003-2004 and 2004-2005 in Maricopa, Arizona. The dataset contains soil moisture measurements and soil texture information throughout the soil profile, grain yield and canopy weight measurements. The experiments consisted of thirty-two 48 m^2 sized plots each representing one of 12 different treatments. Data for this study was taken from a single plot for the 2003-2004 season representing typical management practices for irrigated winter wheat. Irrigation was scheduled for the plots on the day after the daily soil water depletion of the effective root zone was greater than 45% of the available water capacity. To account for irrigation inefficiencies, 110% of the estimated depth of soil water depletion was provided. This irrigation procedure was expected to minimize water stress (Hunsaker et al. 2007b). Meteorological data for the duration of the experiment was provided by a University of Arizona, AZMET weather station, approximately 200 m away from the field site. Complete senescence for each year occurred on 14 May, DOY 135. On 26 May 2004 the wheat was harvested and grain yields were collected in samples from the south half of each plot.

2.4 K-means Clustering

The k-means clustering algorithm is a centroid based approach using cluster distortion to decide when sufficient progress has been made but also can be restricted to a certain number of iterations (Hartigan and Wong 1979). Convergence of the algorithm is based on the change in distance of the mean cluster distance metric. This distance metric is often the squared Euclidean distance or squared normal distance between an observation and the centroid. Initially a set of cluster centers are chosen within the observations space based on a specified number of clusters. Observations are added to a cluster that yields the smallest within cluster sum of square distances with respect its center. This process results in the observations being assigned to the cluster with the closest center. The mean squared Euclidean distance between all the points is calculated, which is the centroid of all the points, and assigned as the new center location. The observations are then re-assigned to clusters based on the distance to the new centers. This process repeats until a given threshold of distortion or sum of squared distances for the clusters has been reached and the center and centroid are sufficiently close (or all iterations have been completed). The k-means algorithm is sensitive to the initial guess and the shape of the data. The algorithm will find the local optimum in terms of sum of squares but does not guarantee that the global optimum is found.

Each cluster analysis considered only one type of data. This gave the resulting classes within a classification a physical meaning that could be compared against one another using statistical metrics. It also allowed values to be assigned to each class for example representing the centroid of the cluster. This is aids in the arrangement or ranking of classes based on a calculated value coming from the simulations within each class. Each classification or clustering analysis considered only 12 clusters, consistent with the USDA texture classification. Only the biophysical variables of yield, crop height and leaf area index were considered in the analysis. The average of the entire time-series of values were considered for variables that changed throughout the season.

It should be noted that the k-means algorithm calculates distances in regards to the values of the observations provided. More specifically the Euclidean distance used to define the cluster membership is calculated in observation space and not using distances between soils as measured on the USDA texture triangle. That is to say that the plotting of the classification on the textural triangle is simply done to visualize the results and that the relative position of the clusters or individual soils to one another on the triangle in no way indicate how closely related they are in terms of the clustering that was performed.

2.6 Simulation Study

The simulation was performed on a daily time-step using the observed meteorological data and experimental irrigation schedule. The entire simulation period was 208 days. The simulation began on 4 December 2003. This was the date of planting and heat units began to accumulate from this date forward. Each simulation was run with the same set of parameters and input data except for the soil hydraulic parameters which were varied based on the set of 1326 soils. Output of the HYDRUS model included soil moisture, meteorological information, plant growth and biomass data. The simulations were called in batch using Python scripts. Post-processing of all the data was also completed using Python scripts.

The simulations were driven by data from meteorological data collected during a single winter wheat season in Maricopa, AZ. Biophysical data from the HYDRUS coupled with the crop model simulations were used to create the classifications using the k-means clustering algorithm. The classifications were made using observations of a single variable type. Each classification assigned every one of the 1326 locations on the USDA texture triangle, representing a specific soil, to one of 12 classes.

3 RESULTS & DISCUSSION

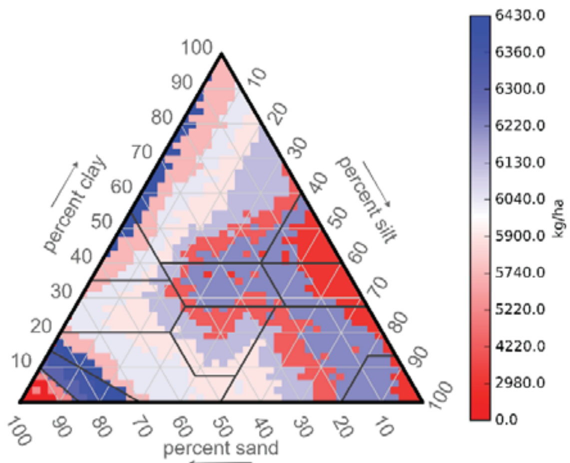

Figure 2. Classification of yield

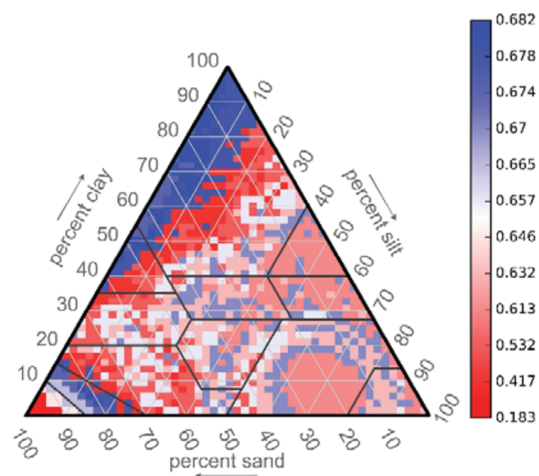

Figure 3. Classification of average Leaf Area Index throughout the season

Figure 2 shows the classification of yield plotted onto the USDA soil texture triangle. The color bar to the right of the figure indicates the value of the centroid associated with each class. Changing from red to white to blue as the centroid value of yield that represents the class increases.

In Figure 2 it can be seen that the lower portion of the triangle aligns with the percentage of sand. Whereas the top half follows the silt gradient. Yield are lowest where the sand large percentages are above 90% and below 10% with high percentage (greater than 50%) silt. Sand is an important factor contributing to increased porosity and the ability of water to move though the soil. However, having a high percentage of sand results in water quickly draining leaving drier soils and the crop little water available. The opposite also seems to be a factor that low sand percentages result in limited water holding capacity and decreased water available to the plant, especially when there is small percentage of clay. This may explain why the yields in these areas are greatly decreased.

The classification of leaf area index (LAI) can be seen in Figure 3. While the classes for this analysis are less uniform than those present in the yield classification, patterns are still visible. Again there is banding of the classes in the sands as well as in the upper corner with high clay contents and low sand and silt contents. A new pattern is present where silt content is greater than 50% and that is circular shape that appears in the silt loam and silty clay classes as designated by the USDA soil textures (outlined in black). The classes near the upper range and lower range of the LAI scale are well grouped while those near the center of the color scale are scattered. This could be a result of the small range for these 6 classes as it only varies over 4 hundreths.

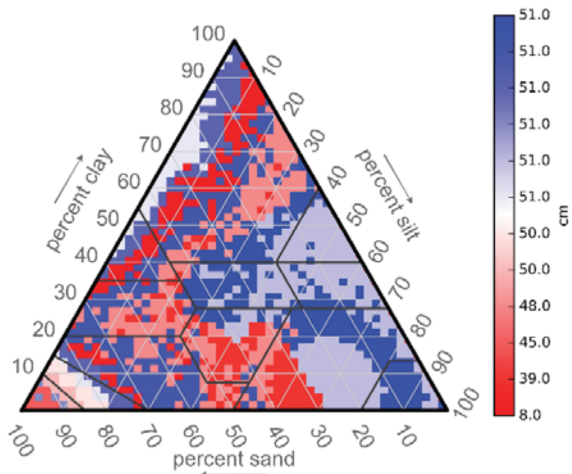

Figure 4. Classification of average crop height throughout the season

There are many similarities between the classification of LAI in Figure 3 and the classification of crop height in Figure 4. The banding and circular grouping patterns occur in the similar areas on the triangle although the exact ranking of the classes does change. However, the values are still lower for the soils with high sand content. The reason that LAI and crop height are classified so similarly is the fact that they are both based upon biomass.

Yield is directly related to biomass unlike crop height and LAI which require extra calculations which may remove some of the variation of biomass. Although the variation seems dampened the banding patterns in each classification are still visible. It is also apparent that there is something unique about the soils in the silt soils of the triangle with regards to what is happening with biomass production.

An important result of the classifications is that the high and low ranked classes are often contiguous. Near the middle of the yield classification the classes fall along a gradient, ranging around 400 kg/ha in value, and are ordered with the soils with respect to their yield. The ordered nature of this gradient means that there is agreement or continuity between clusters or across these soils with respect to yield for example. However, in the soils with high silt contents we have a pattern that alternates from lower yields to higher yields back down to low yields as the sand content decreases. The changes between the classes ranges by as much as 2000 kg/ha. The large difference between the classes indicates that soils positioned on a line across the classes can have drastically different yields.

This is significant because when modelling crop growth a soil texture is specified for the the simulation. Often the impact of selecting or committing to a soil type is not fully considered. Although, it is apparent that by selecting a soil that falls closer to one class within the yield classification can result in a yield that may or may not represent the actual expected yield. This also occurs since soils are chosen such that they may represent some spatial averaging of a larger area. The manner in which representative soils are chosen will determine how well a modelled yield will agree with actual yield.

It is frequently common in practice to measure leaf area index or crop height and relate these observations to yield. The assumption is that these measurement correlate to plant health and biomass production may be valid in some circumstances. These plots indicate that leaf area index and crop height are not a strong indicator of biomass or yield. Even though the general trends and patterns for yield and LAI may agree, that is yield increases with LAI, the scattered nature of the LAI classes and limited range of variation make it difficult to draw a conclusive relationship.

4 CONCLUSION

The results obtained by simulating 1326 virtual soil columns and performing clustering based biophysical variables show that there is value in considering the relationship between plant growth and soil texture. These responses are useful for understanding when uncertainty in soil texture is important when attempting to make yield forecasts. Spatial variation of soil texture is known to heterogeneous and may, over large areas, fall into different USDA soil classes. The soil texture is important for hydrological behaviour and influences to a certain extent plant growth and ultimately yield. Yield even within soil classes can vary drastically. Choosing a single soil class to simulate crop production may result in over or under estimation of actual yield depending on where the soils actually fit onto the USDA soil texture triangle. General patterns in biomass growth can be observable using other biophysical quantities such as crop height and LAI. This is rather difficult given the small range over which they may vary and disorder of the classifications made using them.

Understanding the response of plant growth as a result of variations in soil texture is challenging. It is likely that the main factors driving plant growth vary according to different conditions as well as soil textures. Further investigation into variations of soil texture will require consideration of other hydrological and biophysical aspects of the model. It will be necessary to implement more thorough statistical tools to define relationships among and between classifications.

Future research will focus on turning this general approach into a more formal methodology with quantitative analysis tools and measures. Exploring other components of plant growth using this methodology for producing classifications will hopefully lead to alternative and useful applications of the USDA soil texture.

REFERENCES

Atterberg, A. 1905. “Die Rationale Klassifikation Der Sande Und Kiese.” *Chem. Ztg.* 29: 195 – 198.

Baez-Gonzalez, Alma Delia, James R. Kiniry, Stephan J. Maas, Mario L. Tiscareno, Jaime Macias C., Jose L. Mendoza, Clarence W. Richardson, Jaime Salinas G., and Juan R. Manjarrez. 2005. “Large-Area Maize Yield Forecasting Using Leaf Area Index Based Yield Model.” *Agronomy Journal* 97(2): 418. https://www.agronomy.org/publications/aj/abstracts/97/2/0418.

Bannayan, M, N M J Crout, and Gerrit Hoogenboom. 2003. “Application of the CERES-Wheat Model for Within-Season Prediction of Winter Wheat Yield in the United Kingdom.” *Agronomy Journal* 95: 114–25.

Bormann, H. 2010. “Towards a Hydrologically Motivated Soil Texture Classification.” *Geoderma* 157(3-4): 142–53. http://linkinghub.elsevier.com/retrieve/pii/S0016706110001126 (October 16, 2012).

Davis, Royall Oscar Eugene, and Hugh H Bennett. 1927. “Grouping of Soils on the Basis of Mechanical Analysis.”

De Wit, A.J.W., H.L. Boogaard, and C.A. Van Diepen. 2005. “Spatial Resolution of Precipitation and Radiation: The Effect on Regional Crop Yield Forecasts.” *Agricultural and Forest Meteorology* 135(1-4): 156–68. http://linkinghub.elsevier.com/retrieve/pii/S0168192305002479 (May 6, 2014).

Hansen, J W, and J W Jones. 2000. “Scaling-up Crop Models for Climate Variability.” *Agricultural Systems* 65: 43–72.

Hartigan, John A, and Manchek A Wong. 1979. “Algorithm AS 136: A K-Means Clustering Algorithm.” *Applied statistics*: 100–108.

Hunsaker, D J, G J Fitzgerald, A N French, T R Clarke, and M J Ottman. 2007a. “Wheat Irrigation Management Using Multispectral Crop Coedffients: I. Crop Evapotranspiraton Prediction.” *Transactions Of The Asabe* 50(1): 2017–33.

Hunsaker, D J, G J Fitzgerald, A N French, T R Clarke, and M J Ottman. 2007b. “Wheat Irrigation Management Using Multispectral Crop Coefficients: II. Irrigation Scheduling Performance, Grain Yield, and Water Use Efficiency.” *Transactions Of The Asabe* 50(6): 2035–50.

Jones, J W, G. Hoogenboom, C H Porter, Kenneth J Boote, W D Batchelor, L A Hunt, P W Wilkens, U Singh, A J Gijsman, and J T Ritchie. 2003. “The DSSAT Cropping System Model.” *European Journal of Agronomy*.

Kaul, Monisha, Robert L. Hill, and Charles Walthall. 2005. “Artificial Neural Networks for Corn and Soybean Yield Prediction.” *Agricultural Systems* 85(1): 1–18. http://linkinghub.elsevier.com/retrieve/pii/S0308521X04001398 (April 8, 2014).

Minasny, Budiman, and Alex B McBratney. 2001. “The Australian Soil Texture Boomerang: A Comparison of the Australian and Usda/fao Soil Particle-Size Classification Systems.” *Soil Research* 39(6): 1443–51.

Moriondo, M., F. Maselli, and M. Bindi. 2007. “A Simple Model of Regional Wheat Yield Based on NDVI Data.” *European Journal of Agronomy* 26(3): 266–74. http://linkinghub.elsevier.com/retrieve/pii/S1161030106001390 (May 6, 2014).

Prasad, Anup K., Lim Chai, Ramesh P. Singh, and Menas Kafatos. 2006. “Crop Yield Estimation Model for Iowa Using Remote Sensing and Surface Parameters.” *International Journal of Applied Earth Observation and Geoinformation* 8(1): 26–33. http://linkinghub.elsevier.com/retrieve/pii/S0303243405000553 (April 30, 2014).

Sakamoto, Toshihiro, Anatoly a. Gitelson, and Timothy J. Arkebauer. 2014. “Near Real-Time Prediction of U.S. Corn Yields Based on Time-Series MODIS Data.” *Remote Sensing of Environment* 147: 219–31. http://linkinghub.elsevier.com/retrieve/pii/S0034425714000790 (May 6, 2014).

Schaap, M, F Leij, and M Vangenuchten. 2001. “ROSETTA: A Computer Program for Estimating Soil Hydraulic Parameters With Hierarchical Pedotransfer Functions.” *Journal of Hydrology* 251(3-4): 163–76. http://linkinghub.elsevier.com/retrieve/pii/S0022169401004668.

Simunek, J, M Th Van Genuchten, and M Sejna. 2005. “The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media.” *University of California-Riverside Research Reports* 3: 1–240.

Statistics Division, National Agricultural Statistics Service, Statistical Methods Branch. 2012. *The Yield Forecasting Program of NASS*. Washington, DC.

Twarakavi, Navin K. C., Jirka Šimůnek, and M. G. Schaap. 2010. “Can Texture-Based Classification Optimally Classify Soils with Respect to Soil Hydraulics?” *Water Resources Research* 46(1). http://doi.wiley.com/10.1029/2009WR007939 (March 11, 2014).