2950 Niles Road, St. Joseph, MI 49085-9659, USA
269.429.0300 fax 269.429.3852 hq@asabe.org www.asabe.org

An ASABE Meeting Presentation

Paper Number: 083591

Data Assimilation of Near-Surface In-Situ Soil Moisture Using the DSSAT Crop Model

Candace Batts, Graduate Research Assistant

Iowa State University, Ames, IA, cmbatts@iastate.edu

Amy L. Kaleita, Assistant Professor

Iowa State University, Ames, IA, kaleita@iastate.edu

Written for presentation at the
2008 ASABE Annual International Meeting
Sponsored by ASABE
Rhode Island Convention Center
Providence, Rhode Island
June 29 – July 2, 2008

Abstract. *Soil water is an important variable in agricultural environments as it contributes to yield response as well as areas of environmental concern including erosion, runoff, and nitrogen leaching (through deep drainage). Crop models have been established as a method for simulating agricultural production and examining ecosystem responses. However, because all crop models are based on limited system information, models contain errors which increase uncertainty around their predictions. Data assimilation provides the opportunity to merge both model and observational data in order to obtain a better representation of the true physical system. The objectives of our experiments are to (1) evaluate the efficacy and feasibility of implementing a simple data assimilation algorithm for near-surface soil moisture in the DSSAT (Decision Support System for Agrotechnology Transfer) Model and (2) examine changes in yield from different data assimilation cases. In this paper we use direct insertion, a simple data assimilation method, to examine how assimilation of near-surface (0 – 5 cm) soil water content observations impacts maize yields. Three synthetic experiments were performed using 20 years of simulated climate data, two common Iowa soil types, and two nitrogen rates. The CERES-Maize component of the DSSAT Model was used for simulations. The first experiment consists of simple perturbations of model observations, the second experiment uses incorrect model soil parameters, and the third experiment examines a model bias. The results of the experiments performed here show that it is possible to implement a direct insertion algorithm for near-surface soil water content into the DSSAT model. Yield differences varied according to year, soil type, and nitrogen rate. The results of all three experiments showed that yield differences can occur between scenarios which use the original model generated values and assimilated values even when a simple assimilation method (direct insertion) is used. This* *information provides preliminary insights into the feasibility and impact of using data assimilation with agricultural systems.*

Keywords. Direct Insertion. Maize yield.

Introduction

Soil moisture is a key variable in driving the soil-plant-atmosphere continuum. However, numerous studies have shown that there is a large degree of spatial heterogeneity in soil conditions, notably soil moisture, even on small spatial scales (Famiglietti et al. 1998), with variability being influenced by a number of factors including vegetation characteristics, soil texture, and slope (Hupet and Vanclooster 2002; Famiglietti et al. 1998). Soil moisture status affects transpiration rates for agricultural crops including maize (Novák et al. 2005; Tsuji et al. 1998), which can lead to significant changes in yield over the growing season. Soil moisture is an important consideration for management decisions including planting (Jones et al. 2003). Additionally, soil moisture influences the nitrogen and carbon fluxes within an ecosystem. As nitrogen and CO_2 are becoming increasingly important components of environmental issues, it is important to adequately estimate soil moisture so that accurate nitrogen and carbon balances can be made.

Changes within the soil moisture profile are often estimated in agricultural decision support tools including crop models. Many of these models, including CERES-Maize, a component of the DSSAT (Decision Support System for Agrotechnology Transfer) Model, rely heavily on the soil moisture profile to assess the effects that different factors, such as planting date and plant stress, could have on crop growth and yield. Additionally crop models allow us to examine the potential impacts of different scenarios (e.g. soil type, planting date) without having to physically perform them. Furthermore, models allow us to use automated conditions for management decisions such as planting and harvest with ease.

Despite the many advantages of crop models, they are still limited in their ability to predict system response for a number of reasons including: (1) limitation in model parameters to fully represent the system given spatial and temporal heterogeneity, (2) inability to collect all potential input information about the system, and (3) uncertainty in measurements of parameters. Furthermore, these models can be sensitive to initial conditions and parameterization. For example, the CERES-Maize portion of the DSSAT model has been shown to be sensitive to a number of soil parameters including wilting point (Šťastná et al. 1999), soil water holding capacity, curve number, and soil water content at planting (Bert et al. 2006). Thus errors in the specification of these parameters can have significant effect on model output.

Data assimilation attempts to address some of the uncertainty present in a model representation of an environmental system by incorporating physical observations into the model (McLaughlin 2002). While field observations also contain errors, data assimilation attempts to merge observational data into crop models while accounting for the error in both to arrive at the best estimate of the "true" system. Data assimilation of near-surface (0-5 cm) soil moisture has proven to be a successful means of increasing the accuracy of systems models, in both synthetic and real-data experiments. In remote sensing applications, near-surface measurements of soil water were used to improve the prediction of the soil-moisture profile (Galantowicz et al 1999; Reichle et al. 2001). They have also been used with success on the basin scale to improve calculations of plant available water (Or and Groeneveld 1994). However, until recently (Pauwels et al. 2007; Huang 2004), relatively little work has been done with data assimilation and agricultural crop models, particularly in the case of soil moisture.

The CERES-Maize component of the DSSAT model was selected for these data assimilation experiments for a number of reasons. First, it is a widely used and validated agricultural systems model. However, like most systems-based tools, there is a limited amount of information that can be reasonably gathered, processed, and integrated into a model, which makes DSSAT a favorable target for assimilation. Second, the model has been well calibrated and validated for central Iowa, which is where our future field experiments will take place. Third, the model has been used for some limited data assimilation studies, but the techniques and variables we will study and implement have not been utilized in data assimilation.

Objectives

The experiments discussed in this paper are a preliminary set of procedures whose aim is to determine the suitability for data assimilation of near-surface soil moisture in DSSAT (CERES-Maize). Specifically the objectives are to (1) evaluate the efficacy and feasibility of implementing a simple data assimilation algorithm for near-surface soil moisture in DSSAT and (2) examine changes in yield from different data assimilation cases.

Methods

For this study, we developed and implemented a direct insertion scheme for assimilating near-surface soil moisture observations into the soil water balance component of DSSAT CERES-Maize. The DSSAT model is a field-scale agricultural system model that models crop growth and development, water and nitrogen balances, as well as management practices (Jones et al. 2003). This model was chosen because it is a widely accepted agro-climate model and it has been calibrated and validated for the region of input. As mentioned previously it has known sensitivities to initial conditions and limitations in its water balance calculations; thus, there may be the potential to use assimilation of observed data to account for some of these limitations. Furthermore the model's modular structure facilitates the use of data assimilation algorithms.

Direction insertion is a simple data assimilation technique which involves the replacement of modeled data with observed data. In direct insertion, observed data is considered to be perfect with no error or bias, so there is no integration with the model data; the observed data is simply directly inserted into the model, completely overriding the model estimates of the state of that variable. While previous studies have found that direct insertion is not as powerful as more advanced techniques (Walker et al., 2001; Houser et al., 1998), direct insertion of near-surface soil moisture has been shown to offer improvement over open-loop (i.e. model-only) scenarios (Walker 1999; Heathman et al., 2003) and can indicate if more advanced techniques will be worth pursuing. This is because observations in direct insertion are weighted more strongly than in techniques which allow the model generated values to influence the state. Thus direct insertion observations have a greater influence on state variables than observations from other techniques. In this experiment the model generated soil water content value for the 0-5 cm depth was completely overridden with the assimilated observation based on the assimilation period selected. The soil water contents are generated and/or overridden at the end of the day and used in the next day's calculation. Like in many assimilation algorithms, the mass of the soil water was not conserved (Galantowicz et al 1999).

In this study, synthetic experiments are performed to fully develop and test the assimilation algorithm and its effects on model performance. A synthetic experiment is one in which the model being evaluated is used to create the "observational" data. The model-generated observations are considered to be the “true” observations or states, and are then used in the data assimilation scheme. Model output with data assimilation is then compared to the “observations”, to examine and evaluate the differences. Some advantages of synthetic experiments are that the assimilation algorithm can be rigorously tested, the error applied is known, and sensitivity of different parameters can be evaluated (Reichle 2000).

DSSAT requires extensive input information on crop, soils, weather, and management variables and parameters. For our experiments, crop information was primarily based on a DSSAT example scenario that was built by Liazzo and Batchelor based on experimental work in the same region as our planned future field observations. Simulations for all three experiments started on April 1. The model was directed to automatically plant between April 26 and May 20 when the soil temperature was between 10 and 30 °C and the soil water content of the upper 30 cm was between 40 and 75 % of maximum. Planting and emergence rates of 7.4 plants/m^2 were used. A planting depth of 4.5 cm and a row spacing of 75 cm were used.

Two nitrogen rates were used for each experiment. Following the example of Liazzo and Batchelor, one application of nitrogen was applied in the spring prior to the start of the planting window at two different rates. A low rate of 112 kg/ha N (100 lbs/ac) and a high rate of 224 kg/ha (200 lbs/ac) were used. These levels were chosen because the low rate is commonly applied to corn planted after soybean (lower end) and the high rate is commonly applied to corn planted after corn (higher end).

Two common Iowa soils were used, a Clarion loam and a Nicollet clay loam. These soils were chosen because they have been previously used in central Iowa DSSAT modeling experiments and as such are calibrated and tested for the area. Salient model parameters for these two soils are shown in Table 1.

Table 1. Near-surface model parameters for selected soil types.

Soil	Lower Limit (cm^3/cm^3)	Drained Upper Limit (cm^3/cm^3)	Saturated Water Content (cm^3/cm^3)	Bulk Density (g/cm^3)
Clarion	.110	.300	.361	1.45
Nicollet	.153	.283	.340	1.37

Twenty years worth of simulated weather data were generated using Cligen (http://topsoil.nserl.purdue.edu/nserlweb/weppmain/cligen/) based on a weather station in Ankeny IA (41.72 N, 93.62 N) with a forty-three year record. This station was chosen for its close proximity to future field observation sites. Simulated climate data supplied to the model included daily precipitation, minimum and maximum temperatures, and solar radiation. A summary of the climate variables can be found in Table 2.

Table 2. Climate Parameters

	Mean	Standard Deviation	Min	Max
Annual Precipitation (mm)	751.3	111.9	618.5	995.0
Pre-planting Precipitation (mm)*	65.1	39.7	10.2	144.5
Annual Average Min Temperature (°C)+			3.7	4.5
Annual Average Max Temperature (°C)+			14.8	15.7

* This is reflective of the actual precipitation fed into the model from the start of simulation to the beginning of the planting window.

+ The highest and lowest annual average temperatures observed over the twenty years of simulated climate.

A total of 80 model-only (baseline) runs were performed: two nitrogen levels, two soil types for twenty years of weather data. In order to evaluate the impact of a direct insertion algorithm on the DSSAT model three different modeling experiments were conducted.

Experiment Descriptions

Experiment I

In this experiment synthetic data is generated from a model run by adding white noise ($\mu = 0$, $\sigma^2 = 0.002$ ($cm^3/cm^3)^2$) to the daily near-surface soil moisture model "observations". Values were bounded by the upper and lower limits of the soil (See Table 1). The data is then fed back into the model on either a one-day or seven-day assimilation period. The one-day period has the same time step as the model and was chosen because it is the smallest assimilation period possible. The seven-day period was chosen because it is a more likely sampling period for physical measurements. A total of 160 runs were performed (80 for each assimilation period). The purpose of this experiment was to assess how maize yields were altered by assimilation of near-surface soil moisture and if the length of the assimilation period had any effect on yield response.

Experiment II

This case considers the high spatial heterogeneity of the soil and examines what might happen if the soil type chosen for the model was incorrect, but observations (reflecting true soil characteristics) were assimilated into the model. To simulate this, observations for a Nicollet soil were fed into a model that had been parameterized for Clarion. If the assimilated observations overcome the incorrect model parameters, then the resulting difference between the assimilated and Nicollet yields should be smaller than the original yield difference between the Nicollet and Clarion scenarios. A one-day assimilation period was used for this experiment. A total of 40 runs were performed (two fertilization levels and twenty weather years).

Experiment III

The goal of this experiment was to simulate a model that is biased in one direction (for example always over- or under-estimates PET) and to see how the model responds when it is fed in the "true" values. In this experiment, we assume that the model is biased towards the wet side, so to construct observed data from model output, we applied white noise ($\mu = 0$, $\sigma^2 = 0.0002$ ($cm^3/cm^3)^2$), but only negatively (subtracted from the original model output) to simulate observations that were consistently drier than the model estimates. This data was then fed back into the model in the direct insertion scheme. A one-day assimilation period was used for this experiment. A total of 80 runs were performed.

Results

Model-only simulations

The experimental conditions selected resulted in 18 out of 20 successful year runs in Nicollet under both nitrogen conditions and 15 of 20 successfully year runs in Clarion under both nitrogen conditions. In years where the model did not successfully run, soil conditions were not favorable (too wet) to allow planting within the specified window. Average yield values can be seen in Table 3. On average, the high N rate resulted in higher yields than the low N rate, while Clarion soil produced higher yields than Nicollet soil.

Table 3. Mean Yields (Standard Deviations) of Model-Only Simulations in kg/ha

Soil Type	Low N	High N
Clarion	7968 (771)	9326 (985)

Nicollet	6533 (444)	9181 (1063)

Experiment I

The results show that direction insertion assimilation of near-surface soil moisture can impact the final yield results of model simulations. Though most of the yield differences are small (< 2%), in some cases assimilation caused yield differences of approximately 10%. Little difference was observed between the one-day and seven-day assimilation periods, which have been grouped together in Figure 1. Overall the Nicollet Low N scenario produced the least amount of difference in yield with no changes greater than 3%, while the Nicollet High N Scenario produced the largest number of yield differences greater than 5%. Both the high and low nitrogen scenarios for Clarion fell in between the Nicollet distributions. The root mean square error (RMSE) for yields can be found in Table 4. The RMSE reflects the mean differences in yield between the model-only and assimilated yields for each nitrogen-soil type combination. A higher RMSE value indicates that yields were more responsive to the assimilated observations.

Figure 1. A comparison of model only and assimilated yields for Experiment I.

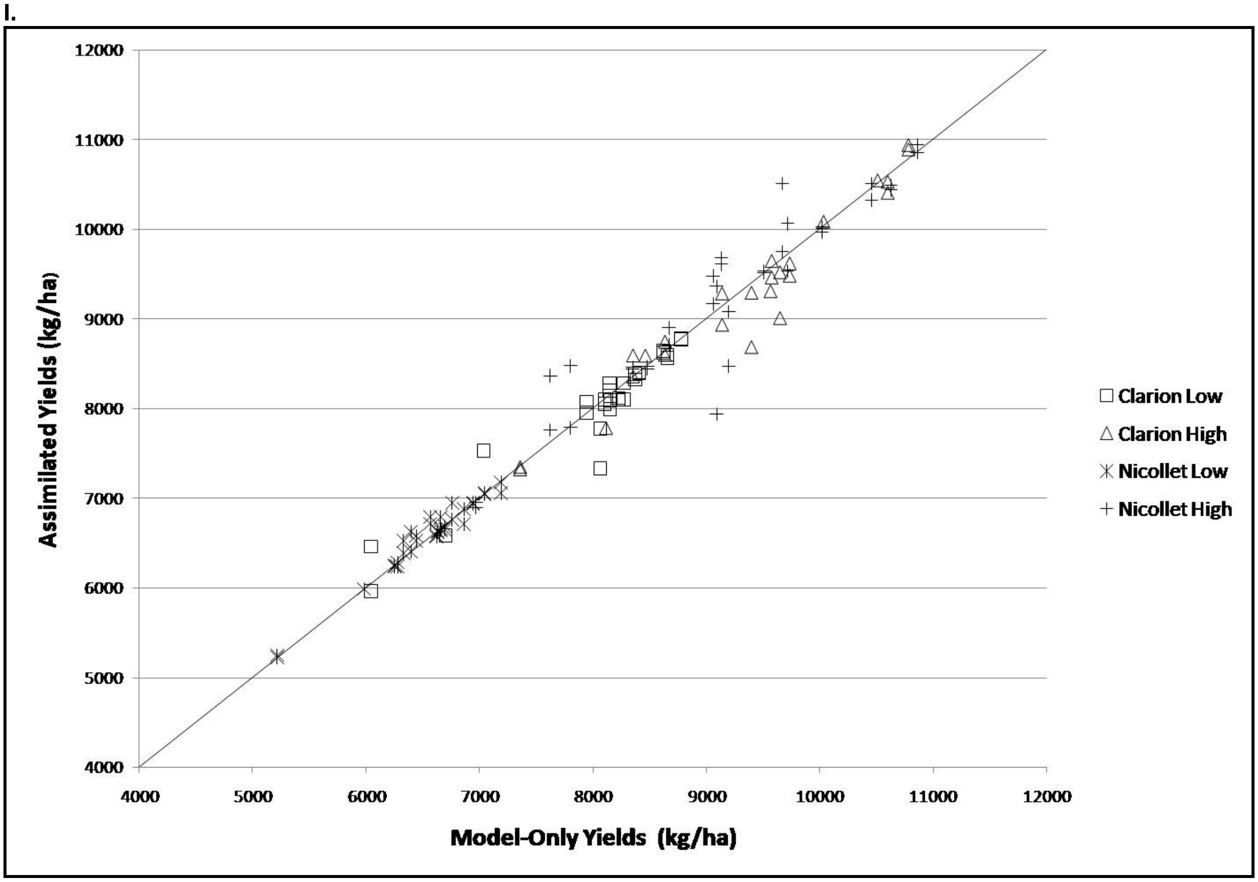


Table 4. RMSE Values for Experiment I

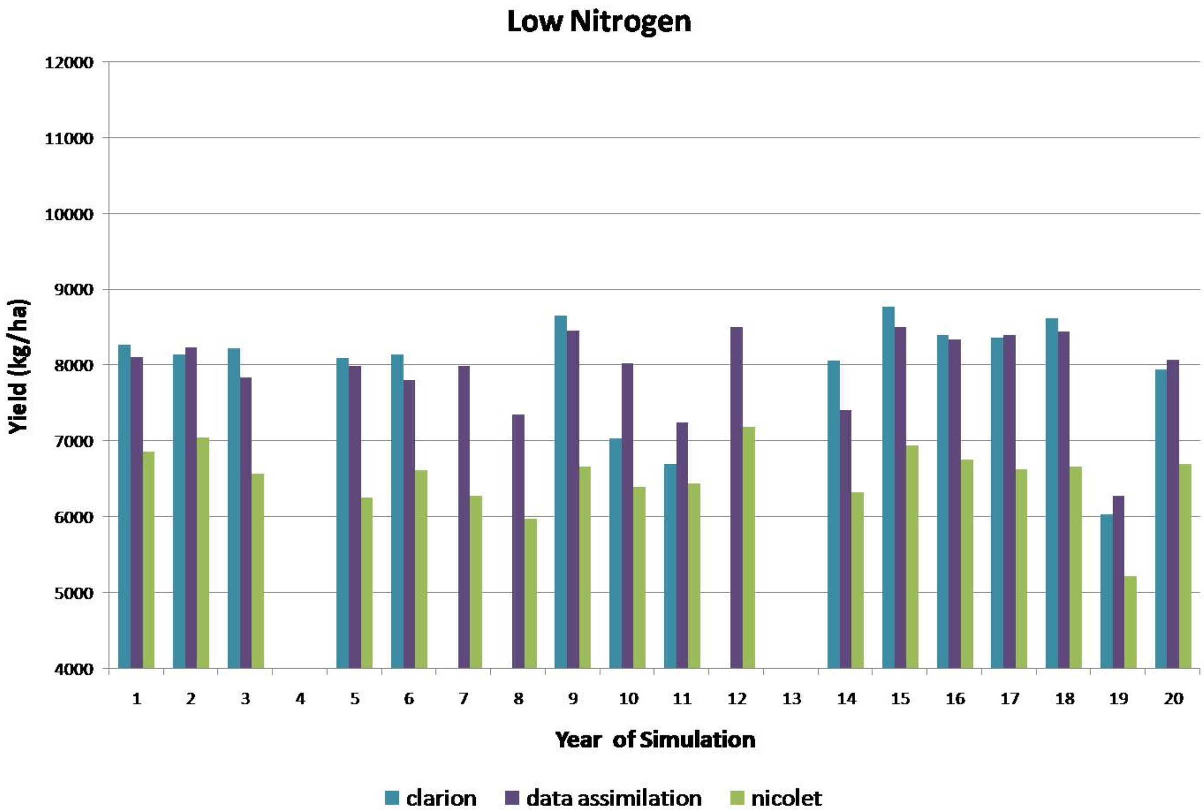
Scenario	Clarion Low	Nicollet Low	Clarion High	Nicollet High
RMSE (kg/ha)	204	92	241	361

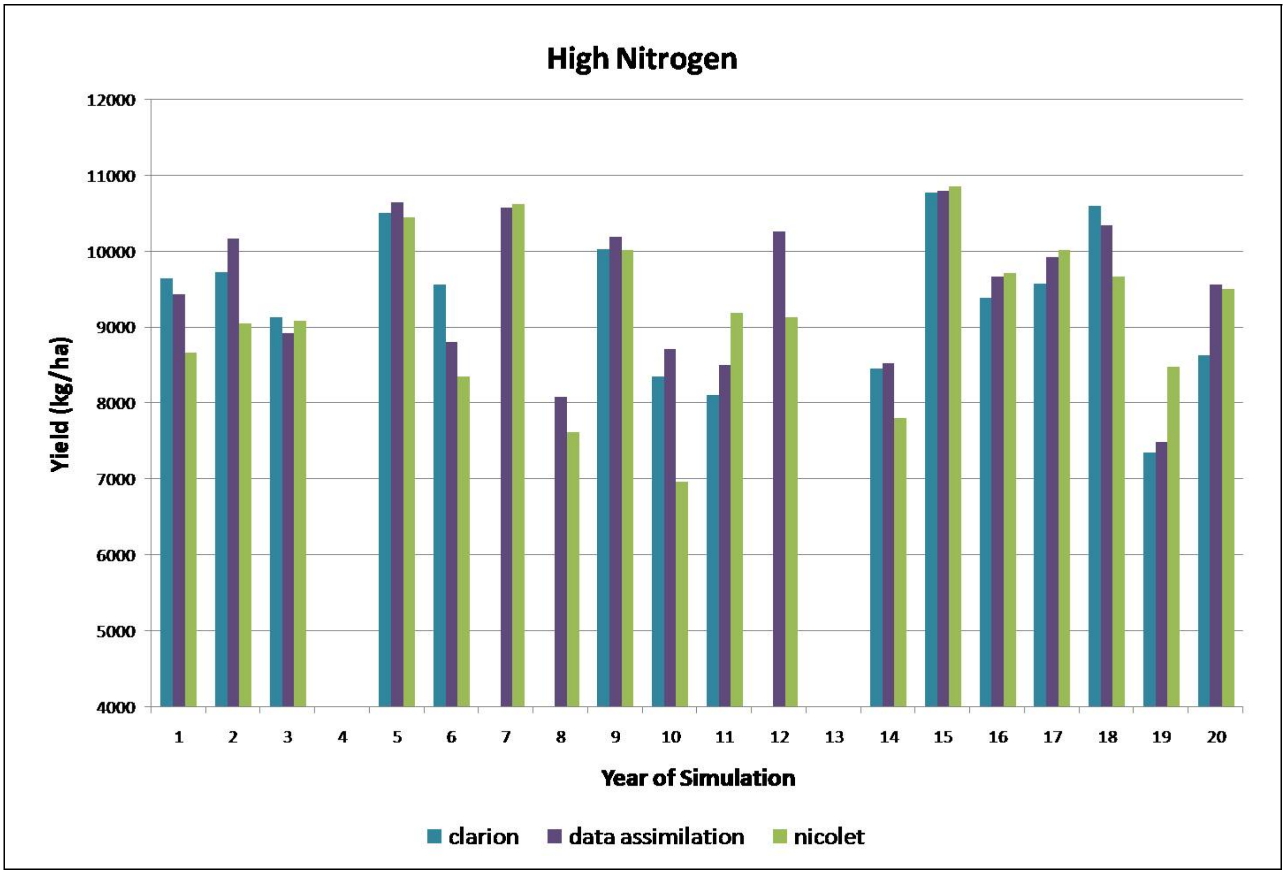
Experiment II

Results for Experiment II, which examines the effect of assimilation when model soil parameters were incorrect, differed between the low to high nitrogen scenarios. In the model-only low nitrogen scenarios there is a large disparity between the Clarion and Nicollet yields with almost half (~47%) having a difference of over 25%. After the assimilation however, all of the yield differences that were >25% fell within the 10-25% range. While the results indicated that the assimilation produced yields that resembled Clarion more than Nicollet (see Figure 2), some improvement over non-assimilated yields was found.

In the model-only high nitrogen scenarios there is less disparity between the Nicollet and Clarion yields as two-thirds of the original yield differences are less than 10%, while the remaining third have yield differences between 10-25%. After assimilation of the Nicollet observations into the Clarion parameterized model, the yield difference between the Nicollet scenarios and the assimilation scenarios is less than 10% in the majority of the scenarios (80%). For the remaining scenarios (20%), the yield differences are between 10-25%. While a smaller percentage of yields in the high nitrogen scenario improved as compared to the low nitrogen scenario (13% versus 47%), the yields differences in the high nitrogen scenario are smaller overall compared with those of the low nitrogen scenario.

Figure 2. Annual Yields for Clarion, Nicollet, and Data Assimilated Yields for Experiment II (top) low nitrogen and (bottom) high nitrogen conditions.





Experiment III

In this experiment, which examines the effect on yield when drier "true" observations are assimilated into the model, there were 63 out of 66 successful simulations (Clarion Low: 15, Nicollet Low: 16, Clarion High: 15, Nicollet High: 17). Overall, data assimilation resulted in yield differences from -2% to 2% approximately 68% of the time, -5 % up to -2% approximately 14% of the time, between -10% and -5% approximately 16% of the time, and yield decreases with magnitudes larger than -10% less than 2% of the time. The results indicate that when true near-surface observations that are negatively biased compared to the original model values are assimilated, decreases in yield may result. However, few of those yield decreases had magnitudes greater than 10%. Furthermore, decreases were not uniform across soil and nitrogen combinations. For example, all assimilated yields in the Nicollet low scenario resulted in yield changes of less than 2% (absolute value), while the Nicollet high scenarios resulted in yield decreases with magnitudes greater than 10%. These differences in yield response are reflected in the RMSE values shown in Table 5.

Figure 3. A comparison of model-only and assimilated yields for Experiment III.

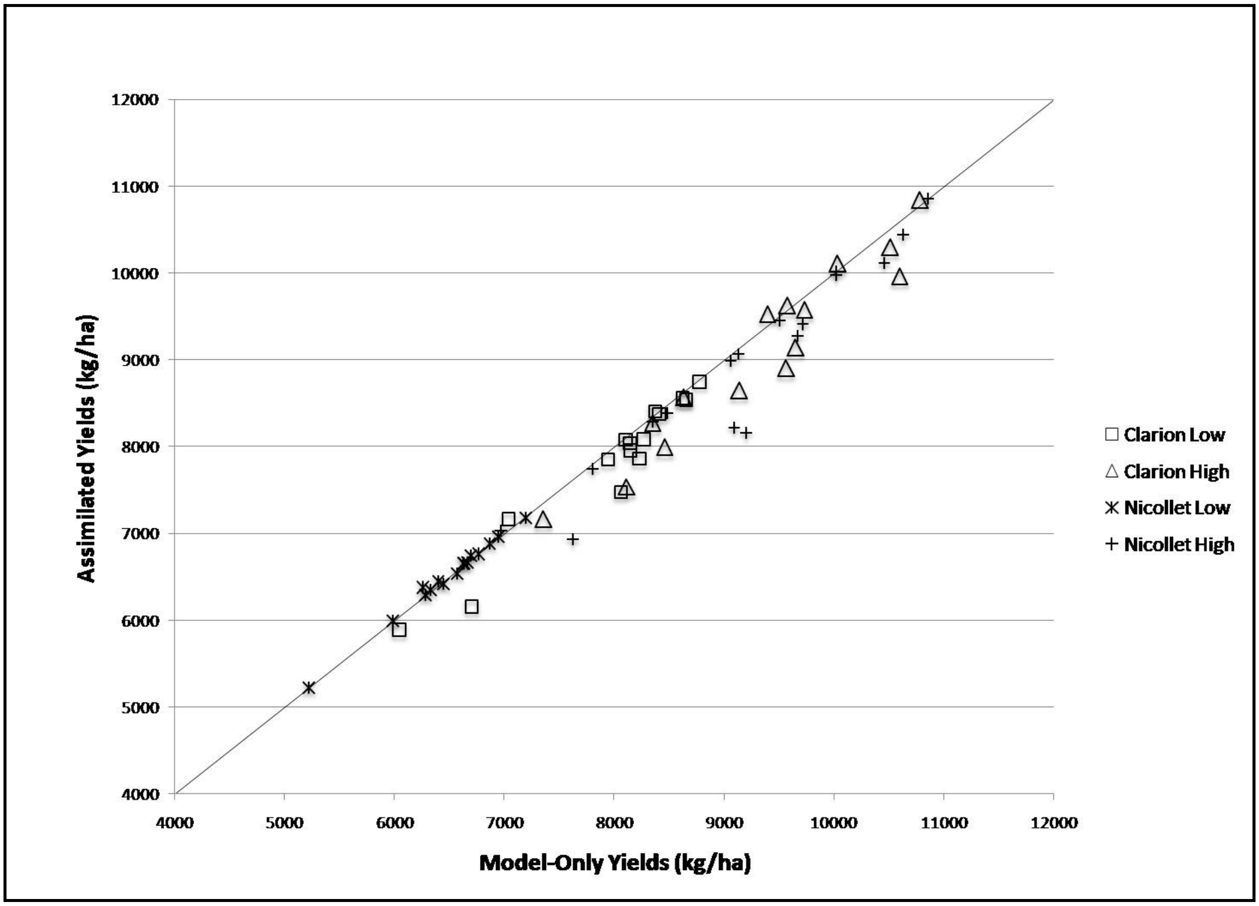


Table 5. RMSE Values for Experiment III

Scenario	Clarion Low	Nicollet Low	Clarion High	Nicollet High
RMSE (kg/ha)	248	40	372	402

Discussion

The results of the experiments indicate that it is possible to implement a direct insertion algorithm for near-surface soil water content into the DSSAT model. Furthermore the results showed that while the resulting yields of data assimilation were often similar to the original model output, at times large differences occurred. Yield differences varied according to year, soil type, and nitrogen rate. Some combinations resulted in yield changes that were greater than 10% (both positive and negative). Yield change distributions appear more closely correlated with nitrogen level than soil type.

Data assimilation demonstrated the importance that near-surface soil moisture values have on the successful running of the model. In Experiment I, there were a number of scenarios that were not able to successfully plant under the baseline simulations conditions, but were successful under data assimilation. Likewise, some scenarios in Experiments I and III that ran in the model-only case, failed under data assimilation. In Experiment II, the yearly success of the assimilated scenarios was the same as the Nicollet, even if the Clarion model-only scenario failed. This shows that DSSAT is sensitive to data assimilation of near-surface soil moisture, particularly in the early growing season (since unfavorable moisture conditions were the cause of failed planting).

The data assimilation also provided useful insights into the model when used with a non-irrigated soil. For example, Experiment III showed that a negative bias (smaller water contents) actually resulted in a slight yield increase in some years for all four scenarios.

Future Work

In the current DSSAT water balance implementation, changes in the near-surface soil moisture are propagated through the soil column solely by model physics, which are fairly coarse. It is possible that with a different representation of the soil profile data assimilation might have a more pronounced effect. Plans for future work include examining how different model physics, particularly the infiltration and redistribution between soil layers, might impact the final yield.

The results of the experiments performed indicate that even a simple data assimilation algorithm of near-surface soil moisture can have an impact on maize yields. As such, some of our future work includes examining more complex algorithms for assimilation of near-surface soil moisture. One notable outcome of the experiment I was that some cases that failed to plant in the model-only scenario were successful for the assimilated cases and vice versa. These differences underscore an important role that near-surface moisture plays especially in the early season, even if most yield changes were negligible. Therefore, future work is likely to couple the use of more advanced assimilation techniques in conjunction with different representations of the model physics.

The experiments performed here focused only on data-assimilation of near-surface soil water content. In future experiments we hope to examine other upper soil layers as well, namely 5-15 cm and 15-30 cm. There is a three-fold reason for this: first, field experiments have been planned which will have sensors deployed at these depths. Second, while the near-surface layer can have a large number of fluxes impacting it, it is also the thinnest layer, so on a total water basis a volumetric change of the same magnitude in lower layers would result in a large change in overall water content. Third, future experiments are planned which call for the evaluation of soil temperature values in conjunction with soil water content, and soil temperature patterns are often less consistent near the surface.

Conclusion

This paper consists of three different synthetic experiments designed to evaluate how data assimilation of near-surface soil water content may influence yields in the CERES-Maize model. The results of all three experiments showed that yield differences can occur between scenarios which use the original model generated values and assimilated values even when a simple assimilation method (direct insertion) is used. While future work will investigate mechanisms of change (model physics) and more advanced assimilation techniques, the preliminary work here does indicate that data assimilation can influence model yield and may help to overcome known model limitations.

Acknowledgements

The research reported in this paper is in part supported by the Graduate Minority Assistantship Program and a fellowship from Department of Education under the GAANN program and the Information Infrastructure Institute at Iowa State University.

References

Bert, F. E., C. E. Laciana, G. P. Podestá, E. H. Satorre, and A. N. Menéndez. 2007. Sensitivity of CERES-Maize simulated yields to uncertaintyinsoil properties and daily solar radiation. *Agricultural Systems*. 94(2): 141-150.

Familgietti, J. S., J. W. Rudnicki, and M.Rodell. 1998. Variability in surface modeling content along a hillslopetransect: Rattlesnake Hill, Texas. *Journal of Hydrology*.210: 259-281.

J. F. Galantowicz, D. Entekhabi, and E. G. Njoku. 1999. Tests of sequential data assimilation for retrieving profile soil moisture and temperature from observed L-Band radiobrightness.

Heathman,G. C.,P. J. Starksa,L. R. Ahuja, and T. J. Jackson. 2003. Assimilation of surface soil moisture to estimate profile soil water content. *Journal of Hydrology.*279: 1-17.

Houser, P. R., W. J. Shuttleworth, J. S. Famiglietti, H. V. Gupta, K. H. Syed, and D. C. Goodrich. 1998. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research*. 34(12): 3405-3420.

Huang, G. 2004. Modeling soil water regime and corn yields considering climatic uncertainty. *Plant and Soil.* 259:221-229.

Hupet, F., and M. Vanclooster. 2002. Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped field. *Journal of Hydrology*. 261: 86-101.

Jones, J. W., G. Hoogenboom, C.H. Porter, K.J. Boote,W.D. Batchelor, L.A. Hunt, P.W. Wilkens, U. Singh, A.J. Gijsman,andJ.T. Ritchie. 2003. The DSSAT Cropping System Model. *European Journal of Agronomy*.18: 235-265.

McLaughlin, D. 2002. An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering. *Advances in Water Resources*. 25: 1275-1286.

Novák, V., T. Hurtalova, and F. Matejka. 2005. Predicting the effects of soil water content and soil water potential on transpiration of maize. *Agricultural Water Management*. 76: 211-223.Or, D. and D. P. Groeneveld. 1994. Stochastic estimation of plant-available soil water under fluctuating water table depths. *Journal of Hydrology.* 163: 43-64.

Pauwels, V. R. N., , N. E. C. Verhoest, G. J. M. De Lannoy, V. Guissard, C. Lucau, and P. Defourny. 2007. Optimization of a coupled hydrology-crop growth model through the assimilation of observed soil moisture and leaf area index values using an ensemble Kalman filter. *Water Resources Research.* 43: W04421.

Reichle, R. H. 2000. Variational assimilation of remote sensing data for land surface hydrologic applications. PhD diss. Cambridge, Mass: Massachusetts Institute of Technology, Department of Civil and Environmental Engineering.

Reichle, R. H., D. Entekhabi, and D. B. McLaughlin. 2001. Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach. *Water Resources Research.* 37(9): 2353-2364.

Šťastná, M., and Z. Žalud. 1999. Sensitivity analysis of soil hydrologic parameters for two crop growth simulation models. *Soil & Tillage Research.*50: 305-318.

Tsuji, G.Y., G. Hoogenboom, and P. K. Thornton (Eds.).1998. *Understanding Options for Agricultural Production. Systems Approaches for Sustainable Agricultural Development.* Dordrecht, The Netherlands:Kluwer Academic Publishers.

Walker, J. P. 1999. Estimating soil moisture profile dynamics from near-surface soil moisture measurements and standard meteorological data. PhD diss. New South Wales, Australia: University of Newcastle, Department of Civil, Surveying and Environmental Engineering.

Walker, J. P.,G. R. Willgoose, and J. D.Kalma. 2001. One dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms. *Advances in Water Resources.*24: 631-650.