The effect of up-scaling soil properties and model parameters on predictive accuracy of DSSAT crop simulation model under variable weather conditions

Jessica Fry[1], Andrey K. Guber[1*], Moslem Ladoni[2], Juan D. Munoz[1], Alexandra N. Kravchenko[1]

[1] Department of Crop, Soil and Microbial Sciences, College of Agriculture and Natural Resources, Michigan State University, USA. J. Fry (fryjess1@msu.edu), J.D. Munoz (munozro3@msu.edu), A.N. Kravchenko (kravche1@msu.edu).

[2] The Climate Corporation, 201 3rd St., San Francisco, CA. 94103, moslemladoni@gmail.com

*Corresponding author, address: Michigan State University, 1066 Bogue St., Room 584D, East Lansing, MI 48824, phone: 517-355-0271 Ext. 269, email: akguber@msu.edu

ABSTRACT

The development of environmentally and economically sound long term agricultural practices under changing climate conditions at the field scale requires implementation of predictive models that assess short and long term responses of agricultural systems to changing environmental conditions. Soil properties for such models are commonly taken from local or National Soil Surveys (SSURGO, STATSGO). This causes a mismatch between the modeling and data source scales. Different techniques can be implemented to upscale soil properties measured at plot scale for field-scale crop modeling, however little is known about the effect of scaling on the accuracy of crop models in predicting crop yields. The objective of this work was to examine: (i) the spatial variability of soybean yields across an agricultural field with different soils and inputs; (ii) how spatial variability in soil properties translates into the variability of measured and predicted yields; (iii) how scaling soil properties affects model accuracy in predicting soybean yields for different weather conditions. The study was conducted at LTER KBS in Michigan, USA. Soybean yield was measured at 2.2 ha and 4.9 ha fields at 2x5 m resolution in 2010. Soil properties were measured at the plot scale in 19 locations, selected to represent two soil types and two management practices on 3 topographical elements (i.e. summit, slope and depression). The DSSAT-CSM was calibrated and validated on soybean yield data measured at the plot scale in 2010 and 2013. The validated model was used to generate yields for 22 years with varying weather conditions in all selected locations. Then the model parameters were scaled up using different techniques, such as averaging plot-scale measured soil properties, averaging the model parameters estimated for each plot using measured soil properties, and using typical soil profile descriptions and the SSURGO soil database. The results of this study showed high variability in soybean yield for two soils and two management practices, which was

associated with variability in soil texture and organic carbon content in the top 20-cm soil layer, but not with surface topography. Despite considerable differences in parameters, all upscaling techniques performed reasonably well for different weather conditions However, model performance appeared to be site specific in this study.

Keywords: DSSAT-CSM model, soybean yield, spatial variability, field scale, upscaling.

INTRODUCTION[1]

The development of environmentally and economically sound long term practices under changing climate conditions at the field scale is unlikely without implementation of predictive models that are capable of assessing the response of agricultural systems to changing environmental conditions. Field experiments are labor intensive and time consuming, yet provide limited (in the spatial as well as temporal sense) information about changes in soil processes including soil water, carbon, and nutrient cycles. As an alternative to and support for field studies, process-based models are commonly used to predict short and long term changes in in these processes, as well as their impact on crop yield. Among others there are APSIM (Keating et al., 2003, Chaucan et al., 2013), CERES (Jones and Kiniry, 1986), EPIC (Williams et al., 1989), STICS (Brisson et al., 2003), SUCROS (Spitters et al., 1990), SALUS (Basso et al., 2011), CropSyst (Stöckle et al., 2003), and DSSAT (Jones et al., 2003).

DSSAT-CSM has been used for predicting crop yield at multiple scales ranging from agricultural fields (Liu et al., 2011; Yang et al., 2013; Salmeron et al., 2014; Negm et al., 2014)

[1] Abbreviations: DSSAT-CSM, Decision Support System for Agrotechnology Transfer with Crop Simulation Model; LTER KBS, the Long Term Ecological Research Kellogg Biological Station; SSURGO, Soil Survey Geographic database; ET, potential evapotranspiration; P, precipitation; TOC, total organic carbon content; RMSE, root mean squared error; RAE, relative absolute error; CN, runoff curve number.

to landscapes and regions (Zhan et al., 2012; Huffman et al., 2014). The input data, specifically soil properties, for the DSSAT-CSM model are commonly taken from soil descriptions, soil surveys or soil maps (e.g. SSURGO, STATSGO). The field scale is of particular interest for growers when making decisions on implementing new management practices in the context of a changing climate (Kaiser et al., 1993; Southworth et al., 2000). However, detailed soil properties data at this scale are typically not available, hence the reliance on these soil maps and surveys to obtain soil data for modeling. The information compiled in the SSURGO database was collected at scales ranging from 1:12,000 to 1:63,360 by the National Cooperative Soil Survey over the course of a century and represents "typical" properties of the soil units. These typical properties generally ignore field-scale spatial variability of the soil, which can be large even within relatively small areas (Pachepsky et al., 2001; Kreznor et al., 1989; Ovalles and Collins, 1986) and may have a significant effect on crop yield (Afyuni et al., 1993; Timlin et al., 1998; Kravchenko and Bullock, 2000; Chaves et al., 2013). Using soil data at a county scale (SSURGO) for field-scale modeling causes a mismatch between the modeling and data source scales. Little is known about how this mismatch affects the accuracy of crop models in predicting field-averaged crop yields.

Since most crop models are one-dimensional, they simulate processes in a single soil profile assuming that the soil is laterally homogeneous. Therefore, simulation results are relevant to the plot scale. For the plot scale, the model parameters are typically measured (Sadler et al., 1999; Sau et al., 2004; Naab et al., 2004), estimated from other measured soil properties (Ritchie et al., 1999; Liu et al., 2011), or estimated by solving an inverse task and fitting the parameters to measured data (Negm et al., 2014; He et al., 2014). Most soils are not homogeneous at a scale large than the plot (e.g. field, landscape, and watershed). For this reason, prediction of average crop yield at a larger scale even for a single management practice, set of weather conditions, and crop type across a field involves upscaling due to spatial variability of soil properties. The upscaling techniques are well developed for a regional scale. A comprehensive review of the scaling techniques and issues associated with crop modeling for a regional scale can be found in Faivre et al. (2004). A good example of upscaling for spring wheat on the Canadian prairies was published by Huffman et al. (2014). The authors demonstrated that aggregation of yield predictions made for multiple soil units based on percentage of soil occupied area produced a reasonable estimate for the yield at a regional scale.

For field-scale simulations, representative soils or soil properties are commonly used in crop modeling (Sadler et al., 2000). To take into account the soil variability, homogeneous zones with different representative soils must be identified within the field (Faivre et al., 2004). Next steps include measuring the soil properties, estimating the model parameters, running the model for each of these zones, and finally aggregating yields predicted for all zones into the field-scale yield. In reality this approach is difficult to implement particularly for small farm fields due to the high cost of soil measurements and difficulties to define soil homogeneous zones within the field. These zones are associated with the spatial and temporal variability of soil properties and crop yield. Limited grower resources in most cases preclude conducting soil surveys, and soil moisture or yield monitoring which could be used for model calibrations and parameter estimation. However, the accuracy of crop predictions relies heavily on high-quality measured soil properties and the correctness of model assumptions.

In this study, we aimed at testing different techniques for upscaling soil properties measured at a plot scale or estimated from publically available data to predict crop yield averaged across the field. We assumed that only soil properties were known a priori and this was the only information available to parameterize DSSAT-CSM. The objective of this work was to examine: (i) the spatial variability of soybean yields across an agricultural field with different soils and inputs; (ii) how spatial variability in soil properties translates into the variability of measured and predicted yields; (iii) how scaling soil properties affects model accuracy in predicting soybean yields for different weather conditions.

MATERIALS AND METHODS

1.1 Field experiment

Field data for this study were collected at two experimental fields at Michigan State University's LTER Scale-up experiment at the Kellogg Biological Station (KBS) located at 42° 24'N, 85° 24'W in Southwest Michigan. Two soil series (i.e. Oshtemo and Kalamazoo), formed as a glacial outwash and marine complex during the last Wisconsin glaciation, were represented in these experimental fields. Glacial Michigan and Saginaw Bay lobes caused great disturbance of the surface topography and upon retreat left a complex landscape of highly heterogeneous deposits (Crum and Collins, 1995). Both soil series belong to the soil taxonomic class of mixed, active, mesic Typic Hapludalfs, with the Oshtemo series being coarse-loamy textured and the Kalamazoo series being fine-loamy. Oshtemo soil prevailed in the North field, while both Oshtemo and Kalamazoo soils were present in the South field (Fig. 1).

The experimental fields used in this study represented two different management practices. The South field was under reduced chemical input (4.9 ha) and will be further referred to as reduced input, and the North field (2.2 ha) was under zero chemical input (organic). The reduced input field received only one starter application of nitrogen at planting and one banded application of herbicide. This was a 50% reduction in relation to the conventional input management practice of the Scale-up experiment. Both management scenarios used for this study included chisel plow tillage and row cultivation. In the organic treatment a rotary hoe was used for weed control. A detailed description of the LTER Scale-up agronomic protocols are available from the KBS LTER website (Simmons, 2012).

Soybeans (Blue River soy seed and Pioneer 92Y30) were planted in June 7 and June 13, and harvested in October 6 and October 15 in 2010 and 2013, respectively, using a combine equipped with precision agriculture software to allow yield measurements with coincident GPS latitude and longitude data (Robertson et al., 2012). Grain flow rate was measured across each field at a 2 m x 5 m resolution. Yield data were processed by removing errors using Yield Editor Software (Sudduth and Drummond, 2007). Nineteen locations were selected at three topographical elements across two fields (i.e. summit, slope and depression) to measure soil properties (Fig.1). We assumed that spatial yield variability observed in these fields was associated with surface soil topography in a manner that yield in some of these locations persistently deviated from the average yield across each field. The measured properties were soil texture, bulk density, and total organic carbon (TOC) content at depths of 20, 35, 50, 70 and 100 cm. Soil texture was measured using the pipette method (Gee and Bauder, 1986). TOC in soil samples was determined via the dry combustion method using a Costech ECS 4010 CHNSO analyzer (Costech Analytical Technologies, Inc. Valencia, CA, United States). Approximately 200 g of soil were taken at the 19 selected locations from each depth for soil texture and TOC measurements. Soil for bulk density measurements was sampled using a soil core sampler, 2-1/4 in. diameter (Soilmoisture Equipment Corp., Santa Barbara, CA). The sample volume was 153 cm^3. To characterize the plot-scale variability of yield we extracted yield monitor data points to a 10m radius buffer created around each sampling location using ArcMap 10.2 (ESRI, Redlands, CA, United States). In order to run different weather scenarios, we used historical weather data recorded by the Gull Lake Biological Station COOP weather observatory from 1993 to 2014. This weather station is located on the Kellogg Biological Station grounds approximately 1,100 m from the fields (Bohm and Robertson, 2015).

1.2 Crop growth model

DSSAT-CSM (version 4.6, Hoogenboom et al., 2015) was used to compute soybean yield for different weather scenarios with field-measured and up-scaled model parameters. The model implements a curve number technique (SCS, 1972; Williams et al., 1984) combined with a tipping bucket approach (Ritchie, 1985) to compute infiltration and vertical distribution of rainfall or irrigation water in soil profiles. Soil water flow parameters of the model include the runoff curve number (CN), which controls partitioning of precipitation into surface runoff and infiltration; the soil lower water limit θ_{LL}, below which plants are unable to extract water from soil; the soil water content at the drained upper limit θ_{DUL}, above which the soil drains water at a rate equal to the soil saturated hydraulic conductivity K_{sat}; saturated water content θ_{sat}, the maximum possible content of water in the soil; and K_{sat}.

For the plot scale, the model parameters were estimated from soil texture and bulk density measured in the 19 selected locations. Specifically, K_{sat} values were estimated using ROSETTA software (Schaap et al., 2001), θ_{sat} was calculated from soil bulk density, while θ_{LL} and θ_{DUL} were estimated using pedotransfer functions as described in Ritchie et al. (1999):

$$\begin{aligned}
\theta_{DUL} &= 0.186\rho_b\,(\mathrm{Sand\%/Clay\%})^{-0.141} \\
\theta_p &= 0.132 - 2.5\cdot 10^{-6} e^{0.015 Sand\%} \\
\theta_{LL} &= \theta_{DUL} - \theta_p
\end{aligned} \qquad (1)$$

where ρ_b is the soil bulk density (g cm^{-3}); Sand% and Clay% are percent sand and clay; θ_p is the plant extractable water (cm^3 cm^{-3}).

The CN values were obtained in the DSSAT-CSM calibration to the soybean yield data measured in the selected locations in 2010. The crop parameters were taken from the DSSAT-CSM database and were not changed in the simulations. The simulations started at a date of soybean planting (June 10th) and finished at soybean maturity, which typically occurred between October 3rd and 17th. We assumed that due to proper management practices plant nutrient demands were satisfied throughout the growing season and any differences in soybean yield could be solely attributed to soil water deficiency.

The DSSAT-CSM model was first calibrated on the soybean yield data measured in the 19 selected locations in 2010 and then validated on data of 2013. Only yield on reduced management was available in 2013. To calibrate the model we varied the CN values separately for each of 19 selected locations to achieve a minimum deviation of the simulated yield from the measured yield. All other soil parameters were fixed in the calibration. In validation the fitted CN values were used to compute yields in the selected locations for weather data of 2013. The goodness of the model prediction was evaluated using the index of agreement *d* (Willmott, 1981):

$$d = 1 - \frac{\sum_{i=1}^{l} (Y_i^o - Y_i^m)^2}{\sum_{i=1}^{l} (|Y_i^m - \bar{Y}^o| + |Y_i^o - \bar{Y}^o|)^2} \qquad (2)$$

where Y_i^o is the yield observed at location *i* (kg ha^{-1}), $\bar{Y}^o$ is the average observed yield (kg ha^{-1}), Y_i^m is the yield predicted by the model at location *i* (kg ha^{-1}), and *l* is the number of locations. The *d*-values close to 1.0 indicate better agreement between the model and observations.

The calibrated and validated DSSAT-CSM was then used to predict soybean yield at the plot scale in the 19 selected locations for different weather scenarios. For these scenarios we used observed weather data for 1993 to 2014 obtained from the Gull Lake Biological Station COOP weather observatory. Predicted values of yield were considered as the equivalent of plot-measured yields for different weather scenarios, and used further to evaluate the accuracy of DSSAT-CSM predictions made using the upscaling techniques described in section 1.3.

1.3 Upscaling techniques

The parameters used in the DSSAT-CSM simulations represented plot and field scales. The plot scale corresponded to the scale of the soil properties and yield measurements. This scale was primarily used to calibrate and validate the model, and generate soybean yields in 19 selected locations for years with different weather conditions. As mentioned above, these yield data were further considered as the equivalent of plot-scale measurements, since direct measurements of the yield were available only for two years (2010 and 2013).

Field-scale properties and parameters of the model were derived from plot-measured soil properties. We used two techniques for upscaling (Fig. 2). In the first technique soil texture and bulk density data measured in the 19 selected locations were averaged across all locations by each soil layer. Averaging was done separately for Kalamazoo and Oshtemo soils, and for reduced and organic inputs, and produced three "representative" soil profiles with specific properties for each soil and management practice. Then, these averaged soil properties were used to estimate DSSAT-CSM parameters K_{sat}, θ_{LL}, θ_{DUL} and θ_{sat} values according to the procedure described in section 1.2. This upscaling technique is referred further to as "Upscaling by averaging plot-scale soil properties". In the second technique K_{sat}, θ_{LL}, θ_{DUL} and θ_{sat} values were estimated separately for each of 19 locations and soil layers, and then these model parameters were averaged for each soil layer of the two soils and inputs, producing total of 3 sets of model parameters. Arithmetic averaging was used for θ_{LL}, θ_{DUL} and θ_{sat}, while geometric averaging was used for K_{sat} values. This technique was referred further to as "Upscaling by averaging plot-scale DSSAT parameters". The CN values for both techniques were set at 73 according to the SSURGO database.

Field-scale parameters of the model were also estimated from soil properties obtained from the LTER KBS and SSURGO soil databases. Both databases provide a description of the typical or "representative" soil profile for each soil series. The LTER KBS database provides soil descriptions specifically for their facility (Crum and Collins, 1995), while the SSURGO (Bockheim and Gennadiyev, 2015) database provides data at a county scale. Therefore, soil properties differed between the two databases. Despite the differences in scale (i.e. experimental station vs. county scale) both databases use the concept of the presence of morphological horizons specific for different soil types. We used soil texture and bulk density from both databases to estimate θ_{LL}, θ_{DUL} and θ_{sat} as was described earlier. The K_{sat} values were estimated using the ROSETTA software for the LTER KBS database, while taken from the database in case of SSURGO. The SSURGO database also provides CN values for Kalamazoo and Oshtemo soils, while the LTER KBS database lacks these data. For this reason we used CN = 73 (taken from the SSURGO database) in simulations that used the LTER KBS soil data. The same CN numbers were used for the plot-scale derived parameters, therefore the differences in yield predictions between the different scaling techniques can be solely attributed to upscaling the θ_{LL}, θ_{DUL}, θ_{sat} and K_{sat} parameters.

1.4 Statistics

The yield simulations were performed for weather conditions of 1993 through 2014 with the field scale parameters obtained using the two upscaling techniques and from the two databases. Each year was simulated separately with the same initial conditions that were used in the model calibration. The results were compared with yield values predicted for the same years using the calibrated model. The accuracy of the model predictions was evaluated using the index of agreement *d* (Eq. 2), root mean squared error (*RMSE*) and relative absolute error (*RAE*):

$$
\begin{aligned}
RMSE_m &= \sqrt{\frac{1}{M}\sum_{m=1}^{M}(\overline{Y}_m^c - Y_m^u)^2} \\
RAE_m &= \left|100\%\,(1 - Y_m^u / \overline{Y}_m^c)\right|
\end{aligned}
\qquad (3)
$$

where $\overline{Y}_m^c$ is the average yield predicted using the calibrated model parameters for *m*-weather scenario (kg ha^{-1}); Y_m^u is the yield predicted for the same weather scenarios using the up-scaled model parameters (kg ha^{-1}); *u* is the up-scaling technique index; *M* is the number of weather scenarios ($M = 22$). Values of *RMSE* and *RAE*, in conjunction, provided a summary of the overall model performance with up-scaled parameters.

The effect of surface topography on the soybean yield was evaluated using PROC MIXED procedure (SAS 9.4, SAS Institute Inc., Cary, NC). The studied factor, topography, with three levels (summit, slope and depression) was treated as a fixed effect. Comparisons between the topographical positions were conducted if the main effect of topography was statistically significant at $p<0.05$.

RESULTS AND DISCUSSION

Measured soybean yield varied spatially and temporally in the studied years. The yield was higher on Oshtemo than on Kalamazoo soil and on the reduced input compared with the organic input in Oshtemo soil (Fig.3). The differences in yield on the reduced input fields between two soils were more pronounced for 2010 than for 2013. This can likely be attributed to the growing conditions, which were more favorable for soybeans in 2010 than they were in 2013, and yield in 2013 was more affected by insufficient precipitation than by the ability of the soil to conduct and retain infiltrated precipitation. Surface topography did not have a statistically significant influence on the soybean yield in 2010 ($p<0.05$). High yield variability was observed on all topographical elements, and could be likely attributed to the natural spatial variation in physical, chemical and biological properties (Long et al., 1963; Karlen et al., 1990). Despite the differences in the absolute values of the yield, the locations selected at different topographical elements for soil sampling represented the yields in a wide range of cumulative relative frequencies reasonably well, thus providing adequate datasets for the DSSAT-CSM calibration and validation (Fig. 3).

As expected, a high variability was present in soil texture measured at different locations across the two fields. Soil texture measured at the same depths varied across seven textural classes from sand to clay (Fig. 4a). High variability was observed at all depths, indicating that the differences in soil texture could not be attributed solely to surface erosion of slopes and sediment deposition in depressions, which typically alter soil properties in the top soil layer, but could rather be attributed to the glacial formation of these soils. The variability in soil texture was also high for data taken from the two soil databases. In spite of the small number of soil layers the texture varied across 5 textural classes for both the LTER KBS and SSURGO databases (Fig. 4b). Soil texture provided by the databases differed markedly from the measurements taken at our 19 selected locations. For the same sand contents, clay contents were higher in the field measurements compared to those from the databases, and the difference in clay content increased as sand content decreased (Fig. 4).

The variability in soil bulk density and organic carbon was also high in the selected locations. Coefficients of variation ranged from 0.048 to 0.113 and from 0.48 to 1.31 at five depths in these locations for soil bulk density and TOC, respectively. In general, bulk density increased while TOC decreased with depth (Fig. 5). However, the spatial variability of *BD* decreased with depth, as assessed by the coefficient of variation (CV), while the CV values for TOC did not show any trend by depth. The measured soil properties differed from those estimated from the two databases. Specifically, the LTER KBS database overestimated, while SSURGO database underestimated field-measured values of soil bulk density (Fig. 5a). The values of TOC content were somewhat higher in the top 30-cm soil layer for the database data in comparison to those measured in the selected locations (Fig. 5b).

The spatial variability in soybean yield can be in part attributed to the variability in soil properties in the top 20-cm soil layer. The yield measured at the 17 locations correlated with TOC (r=0.66, P=0.005) in 2010 and with clay to sand content ratio (r=-0.765, P=0.01) in 2013. However, correlations between the yield and clay to sand content ratio in 2010, and between the yield and TOC in 2013 were not statistically significant at P = 0.05. This is an indication that TOC and clay to sand content ratio are not robust predictors of soybean yield in the studied soil. Interestingly, the yield in 2010 did not correlate with TOC measured in two locations with the highest TOC (4.2% and 7.3%) in the top soil layer. This concurs with the results of Kravchenko and Bullock (2001), who studied corn and soybean variability in Haplaquolls and Argiudolls of central Illinois and eastern Indiana. These authors found that that OM content was a more important yield-affecting factor in soils with low (OM < 3%) than with high OM (OM > 3%)

content. For comparison, in our study a good correlation between yield and OM was observed in 2010 for OM values ranging from 0.5% to 2.0%.

Spatial variability in soil properties translated into different values of the DSSAT-CSM parameters calculated using Eq.1 and estimated using ROSETTA software (Schaap et al., 2001). These values are shown in Fig. 6 a-d. The spatial variability of the DSSAT-CSM parameters was more pronounced for K_{sat} and less pronounced for θ_{sat} values. Coefficients of variation ranged from 0.20 to 0.28, from 0.12 to 0.24, from 0.055 to 0.087 and from 1.08 to 0.65 at five depths in the selected locations for θ_{LL}, θ_{DUL}, θ_{sat} and K_{sat}, respectively.

The differences in the CV values for different DSSAT-CSM parameters were associated with the differences in the CV values for the measured soil properties and the differences in methods of parameter estimation. The θ_{sat} values were calculated as a linear function of soil BD, θ_{LL} and θ_{DUL} were calculated as a product of BD and nonlinear function of Sand to Clay ratio, while K_{sat} values were estimated from soil texture and BD using a strongly nonlinear neural network. The low CV values for BD (0.048 to 0.113) translated into relatively low CV values for θ_{sat}, while high CV values for Sand to Clay ratio (1.32 to 3.35) translated into high CV values for θ_{LL} and θ_{DUL} though nonlinear transformation of soil properties to θ_{LL} and θ_{DUL} considerably reduced CV values for this parameters compared to CV values for Sand to Clay ratio. The CV values for K_{sat} were much higher than those for the measured soil properties, which likely resulted from high sensitivity of the ROSETTA software to changes in soil properties.

Model calibration with fixed values of θ_{sat}, θ_{LL}, θ_{DUL}, K_{sat} and variable CN values produced yield predictions very close to the measured yields in 18 out of 19 selected locations (Fig. 7). The agreement index *d* equal to 0.99 in calibration and 0.84 in validation, which indicated adequate model performance. The calibration produced CN values ranging from 60 to

93 (CN=73±10, CV=0.136) in those 18 locations. It occurred that the differences in CN values were not associated with topographical locations, but rather with soil texture. A significant positive correlation (P = 0.002) was found between CN values larger than 70 and clay to sand ratio in the top 20-cm soil layer. The CN values increased from 72 to 93 with an increase in the clay to sand ratio from 0.05 to 1.7. The greater CN indicates greater runoff rates in the top soil with higher clay and lower sand contents. These results agree with the CN values recommended for different soil texture and soil hydraulic groups (Cronshey et al., 1986). Indeed, according to the soil texture measured in the top soil layer, our soil samples can be placed into three hydrologic groups (i.e. A, B and D) with the CN values of 73 (A), 73 to 78 (B), and 82 to 93 (D). For comparison, the CN estimates for straight row legumes are 66, 77 and 89 for groups A, B and D, respectively.

When the DSSAT-CSM parameters were estimated from the measured soil properties and from the CN values calibrated to the yield measured in the selected locations in 2010, their spatial variability resulted in the spatially variable predictions of soybean yield. However, the yield variability was only partly attributed to the variability in the model parameters, and was mostly affected by the weather scenarios (Fig. 8). For 22 weather scenarios the CV values of the yield varied from 0.025 to 0.429 in the simulations with the same values of the DSSAT-CSM parameters. In these scenarios the total precipitation (P) and potential evapotranspiration (ET) values during the growing season varied from 166 to 566 mm and from 258 to 566 mm, respectively. The ET values exceeded precipitation in 16 out of 22 weather scenarios (Fig. 8), therefore plant water stress was possible in most of the scenarios. Predicted yield values did not correlate with P in our simulations. The reason for this was a non-uniformity of the precipitation distribution during the growing seasons and the fact that soybeans respond differently to water stress depending on their developmental stage. Stress effects were reported by different authors before. For example, Pejić et al. (2011) observed that soybeans are most sensitive to water stress during the general yield formation stages, since this is when the plants' water use is the highest (50.7% of total available water used) in comparison to the vegetative and flowering stages (28.2% and 21.1% of total available water used, respectively). Cox and Jolliff (1986) noted that although soybeans in general are unable to withstand prolonged drought, tolerance is the lowest during the seed enlargement or pod elongation stages. In a similar vein, Sionit and Kramer (1977) found that the seed enlargement stages are the most critical; they obtained positive yield responses in soybeans that were only irrigated during these critical stages. Therefore, the existence of some water stress during the growing season does not automatically result in yield reductions. Whether or not the yield reductions will occur depends on the growth stage at which the water deficit takes place.

To account for the water stress experienced by the crop, the DSSAT-CSM model implements a dual approach to the plant growth, i.e., a water deficit relationship (Ritchie, 1998). Soilwater deficit affects crop canopy photosynthesis, which controls rates of dry matter accumulation, and plant physiological processes (i.e. roots, leaf and stem development). Moreover, soybean physiological maturity has been shown to accelerate with an increase in the water deficit (Desclaux and Roumet, 1996; Ruíz-Nogueira et al., 2001). This is accounted for in the DSSAT-CSM (Boote et al., 2008). Therefore, water deficit may produce a mixed effect on the soybean growth and yield. Surprisingly, a significant correlation (P = 0.04) was observed between the predicted yield and absolute values of differences $|P\text{-}ET|$ (data not shown).

The differences in soil properties measured in the field and derived from the soil databases translated into different values of DSSAT-CSM parameters (Table 1 and Table 2). The

θ_{LL} and θ_{DUL} values were consistently smaller in SSURGO-derived parameters for Oshtemo soil, while they were consistently larger in LTER KBS-derived parameters for Kalamazoo soil as compared to the field estimates in the same soil layers. These differences can be attributed to relatively high clay% to sand% ratios in the LTER KBS database and low ratios in the SSURGO database (Fig. 4b). The differences in θ_{LL} and θ_{DUL} values did not cause noticeable differences in plant extractable water, since θ_p values were estimated from sand, which was in the same range in the database and plot-scale measurements. Differences in soil bulk densities for the two databases resulted in overall low values of θ_{sat} for the LTER KBS database, and high θ_{sat} values for the SSURGO database compared with the θ_{sat} values estimated from the plot measured bulk densities. Among all estimates, the K_{sat} values were the smallest in the top 0-70 cm layer of Kalamazoo soil for LTER KBS database and the largest in the bottom 96-152 cm layer of the same soil for the SSURGO database (Table 2).

The differences in parameter values between the two field-scale upscaling techniques (i.e. averaging plot-scale DSSAT-CSM parameters and averaging plot-scale soil properties) were less pronounced compared to the differences in the database-derived parameters. The largest differences for the field scale were observed for K_{sat} values that were in many instances higher for the averaging plot-scale DSSAT-CSM parameters than for the averaged plot-scale soil properties (Table 1). These differences were attributable to high variability in soil properties measured at the same depth in different locations and strong nonlinearity between soil properties and $\log_{10}(K_{sat})$ in the ROSETTA software (Schaap et al., 2001). Indeed, the differences in K_{sat} values between upscaling by averaging plot-scale soil properties and upscaling by averaging plot-scale DSSAT-CSM parameters were less pronounced in the upper soil layers (0-20 cm and

20-35 cm) compared to the soil layers below depth of 35 cm (Table 1). This occurred due to relatively high variability of soil texture in the deep soil layers (Fig. 4).

Comparison of the results of the soybean yield simulations using the calibrated and up-scaled model parameters is shown in Fig. 8. Despite high yield variability between the different weather scenarios all upscaling techniques (with the exception of SSURGO on Kalamazoo soil) generated reasonable yields. Despite an overall acceptable performance indicated by the index of agreement *d* (Table 3), the values of the *RMSE* were different for different upscaling techniques. The smallest *RMSE* values were obtained for the averaged DSSAT-CSM parameters, while the largest *RMSE* occurred for the SSURGO-derived parameters in Kalamazoo soil and the LTER KBS parameters in Oshtemo soil on both management practices (Table 3). These differences in model performance for different soils and management practices were not associated with the model parameters estimated from highly variable soil properties, but were caused by differences in the yield values used for model calibration. The yield values in the calibrated dataset in Kalamazoo soil and in Oshtemo soil on organic input were considerably smaller compared to the yield on Oshtemo reduced input plots (Fig. 3). As a result, the calibrated model produced large yields on Oshtemo reduced input plots in relation to the yields on the two other plots. It also can be seen that in most DSSAT-CSM runs using upscaling the model tended to overestimate yields that were predicted using the calibrated parameters (Fig. 8). Therefore, the model predictions using upscaling were less accurate on plots with relatively low yield (Kalamazoo and Oshtemo organic input plots) than on the plot with high yield (Oshtemo reduced input).

The accuracy of the model upscaling techniques, assessed as a relative error of yield prediction, also varied between the different simulations. The *RAE* values were smallest for the yield predictions with averaged plot-scale DSSAT-CSM parameters, the largest for the

SSURGO-derived parameters and intermediate in simulations with averaged plot-scale soil properties and LTER KBS-derived parameters Fig. 8a. Poor model performance with the SSURGO-derived parameters was mostly associated with large errors in Kalamazoo soil. The *RAE* values for Oshtemo soil did not differ significantly from those obtained using the other upscaling techniques on both reduced and organic inputs.

Relatively small differences in yields predicted using different upscaling techniques for the same weather scenarios could be attributed to a relatively high model sensitivity to the CN values, which were the same in all upscaling techniques, and relatively low sensitivity to the soil parameters within their ranges. The high sensitivity to CN parameter can be illustrated by implementing another approach to estimating the CN values for the field-scale upscaling techniques. In this approach, using soil texture measured in the top soil layer we first identified soil hydrologic group for each soil and management practice, and then estimated the CN values for obtained groups (Cronshey et al., 1986). The generated CN values appeared to be much higher compared to those taken from SSURGO database and used in our simulations. Specifically, new CN values were 88, 89 and 76 for Kalamazoo soil, Oshtemo soil on reduced and Oshtemo soil on organic inputs, respectively. Higher CN values obtained on reduced management practice generated higher runoff and lower infiltration fluxes that led to a considerable decrease in predicted yields, particularly for high yield years (Fig. 8, dotted lines). Relatively low sensitivity of DSSAT-CSM to the differences in K_{sat} values obtained in different upscaling techniques can be explained by relatively high K_{sat} values in these soils that did not restrict water flow through soil profile under current weather scenarios. In other words, the CN parameter appeared to be a limiting factor for water influx into the soil profile that affected crop growth and controlled the yield. This conclusion concurs with the results of Kendall et al. (2012)

who conducted a global sensitivity analysis of DSSAT-CSM to estimate the effects of soil and crop parameters on corn yield with and without irrigation. These authors concluded that the K_{sat} parameter had no influence on corn yield within the range of K_{sat} values used in their sensitivity analysis. This could be due to greater importance of the parameter range for the output variable (i.e. yield) than the parameter variability within the range (Monod et al., 2006).

Acceptable performance of the model using the SSURGO-derived parameters in this study was surprising, because the national soil survey was not designed as a site-specific or field-scale agricultural tool (Mausbach et al., 1993), and soil properties taken from this database differed considerably from the field-measured properties. There is no general agreement on the correlation between the scale of yield averaging and of the soil survey map in the literature. For example, Karlen et al., (1990) found that field-scale variation in corn, wheat and sorghum yields was partly associated with different soil map units identified at a 15 m resolution in Ultisols soil, though yield variability within each soil unit was very high. Steinwand et al. (1996) found that the scale of soil mapping (1:3305 vs. 1:15840) had little effect on corn, soybean, oat, and hay yield estimates for a Clarion-Nicollet-Webster soil association area in central Iowa, despite the fact that between 35 and 100% of the soil units on the 1:6333 map overlapped with the same units on the 1:15840 soil survey map. The authors explained the small differences in the yield by the high percentage of similar soils included in both soil maps and concluded that county soil surveys and attribute data in central Iowa were acceptable to evaluate soil landscapes for crop yield interpretations. Sadler et al. (1998) found a significant though weak correlation between corn, wheat, soybean and sorghum yields and soil map units at the 1:1200 scale, which provided rather limited predictive yield values for precision farming. Contrary to the results of the latter study the SSURGO database (scale of 1:24000) provided reasonable estimates of soil properties in Oshtemo soil for yield prediction at the field scale in our study.

The differences in *RAE* values obtained in our simulations were both attributed to the upscaling techniques and the weather conditions. Specifically, for Oshtemo soil on organic input averaged *RAE* values decreased in an order: LTER KBS database > averaging plot-scale soil properties > SSURGO database > averaging plot-scale DSSAT parameters, indicating better performance of the last upscaling technique (Fig. 9b). Analysis of the relationship between *RAE* and |*P-ET*| showed a general increase in the relative yield error with deviation of precipitation from potential evapotranspiration (though statistically significant for averaged soil properties and LTER KBS database, P = 0.05). This indicates that the errors of model prediction with the up-scaled parameters will likely increase in unusually wet and dry years. This finding is very important in the context of a changing climate characterized by a more frequent occurrence of extreme weather conditions. Interestingly, *RAE* values for averaging plot-scale DSSAT parameters ranged from 0.02 to 10.2% and were not affected by |*P-ET*|, demonstrating robustness of this upscaling technique for different weather conditions.

CONCLUSIONS

The high spatial variability of soil properties (i.e. soil texture, bulk density, organic carbon) and soybean yield observed in two soils and two management practices which included reduced and organic inputs were not associated with topographical elements in this study. However, the locations selected at different topographical elements for soil sampling represented yield variability for the 2010 and 2013 growing seasons in both soils and management practices reasonably well, thus providing adequate datasets for the DSSAT-CSM calibration and validation.

We did not find a relationship between measured yield and soil properties in this study. The yield correlated with TOC in top 20-cm soil layer in 2010 and with clay to sand content in 2013, but these correlations were not persistent for different weather conditions.

The variability in soil properties translated into spatially variable parameters of the DSSAT-CSM model. The key parameter that affected adequate prediction of the measured soybean yield, the runoff curve number (CN), positively and significantly correlated with the spatially variable ratio between clay and sand contents in the top 20-cm soil layer.

Spatial variability in the DSSAT-CSM parameters estimated from measured soil properties translated in the variability in predicted yield, however the model predictions were more affected by the weather scenarios than by spatially variable model parameters.

The four upscaling techniques provided different estimates of the DSSAT-CSM model parameters, however the values of these parameters appeared to be less important for the yield prediction compared to the CN values that were derived from SSURGO database and set to 73 in all simulations. This result was attributed to low sensitivity of the model to the variability in soil parameters within the parameter range obtained for these two soils.

Based on the index of agreement *d,* all upscaling techniques performed reasonably well for different weather conditions. Among other upscaling techniques only averaging plot-scale DSSAT-CSM parameters generated predicted errors that did not correlate with precipitation less evapotranspiration, and therefore this technique may be recommended for modeling soybean yield in changing climate conditions.

ACKNOWLEDGEMENTS

The work was supported by the United Soybean Board project #1420-732-7230 “Combining Integrated Pest Management with Climate Change, Mitigation, and Adaptation in

Soybean-Corn Cropping Systems and was a part of a regional collaborative project supported by the USDA-NIFA, Award No. 2011-68002-30190 "Cropping Systems Coordinated Agricultural Project (CAP): Climate Change, Mitigation, and Adaptation in Corn-based Cropping Systems". Support for this research was also provided by the NSF Long-term Ecological Research Program (DEB 1027253) at the Kellogg Biological Station and by Michigan State University AgBioResearch". We are also grateful to Brook Wilke and Suzanne Sippel from LTER KBS for providing the yield monitor data for 2013.

REFERENCES

Afyuni, M.M., Cassel, D.K., Robarge, W.P., 1993. Effect of Landscape Position on Soil Water and Corn Silage Yield. Soil Sci. Soc. Am. J. 57, 1573.

Basso, B., Ritchie, J.T., Grace, P.R., Sartori, L., 2011. Simulation of Tillage Systems Impact on Soil Biophysical Properties Using the SALUS Model. Ital. J. Agron. 1, 677–688.

Bockheim, J.G., Gennadiyev, A.N., 2015. General state soil maps in the USA. Geoderma 253-254, 78–89.

Boote, K.J., Sau, F., Hoogenboom, G., Jones, J.W., 2008. Experience with water balance, evapotranspiration, and predictions of water stress effects in the CROPGRO model. In: Ahuja L.R., V.R. Reddy, S.A. Saseendran, Q. Yu (Eds.), Response of crops to limited water: Understanding and modeling water stress effects on plant growth processes. ASA, CSSA, SSSA, Madison, WI, USA, pp. 59-103.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussière, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudillère, J.P., Hénault, C., Maraux, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model STICS. Eur. J. Agron., Modelling Cropping Systems: Science, Software and Applications 18, 309–332.

Chaves, B., Hoogenboom, G., Thornton, P., 2013. Variability in maize yield due to differences in soil texture and the effect of water and nitrogen. Proceedings of the ASA, CSSA & SSSA International Annual Meetings, held in Tampa, Florida, USA, 3-6 November 2013. Available: https://cgspace.cgiar.org/handle/10568/52043 (Accessed: 17/07/2015).

Cox, W.J., Jolliff, G.D., 1986. Growth and Yield of Sunflower and Soybean under Soil Water Deficits. Agron. J. 78, 226.

Cronshey, R., McCuen, R.H., Miller, N., Rawls, W., Robbins, S., Woodward, D., 1986. Urban Hydrology for Small Watersheds. USDA NRCS, Engineering Division, 210-VI-TR-55.

Crum, J.R., Collins, H.P., 1995. KBS Soils. Kellogg Biological Station Long-Term Ecological Research.

Desclaux, D., Roumet, P., 1996. Impact of drought stress on the phenology of two soybean (Glycine max L. Merr) cultivars. Field Crops Res. 46, 61–70.

Faivre, R., Leenhardt, D., Voltz, M., Benoit, M., Papy, F., Dedieu, G. and Wallach, D., 2009. Spatialising crop models. Pages 686-705 *in* E. Lichtfouse, et al., eds. Sustainable agriculture. Springer, Berlin, Germany. DOI 10.1007/978-90-481-2666-8

Gee, G.W., Bauder, J.W., 1986. Particle analysis. p. 383–412. In: A. Klute (ed.) Methods of soil analysis. Part 1. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.

He, Y., Hou, L., Wang, H., Hu, K., McConkey, B., 2014. A modelling approach to evaluate the long-term effect of soil texture on spring wheat productivity under a rain-fed condition. Scientific Reports 4 : 5736, DOI: 10.1038/srep05736

Hoogenboom, G., Jones, J.W., Wilkens, P.W., Porter, C.H., Boote, K.J., Hunt, L.A., Singh, U., Lizaso, J.I., White, J.W., Uryasev, O., Ogoshi, R., Koo, J., Shelia, V., Tsuji, G.Y., 2015. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.6 (www.DSSAT.net). DSSAT Foundation, Prosser, Washington. Huffman, E.C., Yang, J.Y., Gameda, S., de Jong, R., 2001. Using Simulation and Budget Models to Scale-Up Nitrogen Leaching from Field to Region in Canada. The Scientific World Journal 1, 699–706.

Huffman, T., Qian, B., De Jong, R., Liu, J., Wang, H., McConkey, B., Brierley, T., Yang, J., 2014. Upscaling modelled crop yields to regional scale: A case study using DSSAT for spring wheat on the Canadian prairies. Can. J. Soil Sci. 95, 49–61.

Jones, C.A., Kiniry, J.R., 1986. CERES-Maize: A simulation model of maize growth and development.Texas A&M University Press, College Station, Texas.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model, European Journal of Agronomy, Modelling Cropping Systems: Science, Software and Applications.

Kaiser, H.M., Riha, S.J., Wilks, D.S., Rossiter, D.G., Sampath, R., 1993. A farm-level analysis of economic and agronomic impacts of gradual climate warming. Am. J. Agric. Econ. 75, 387–398.

Karlen, D.L., Sadler, E.J., Busscher, W.J., 1990. Crop yield variation associated with Coastal Plain soil map units. SSSAJ 54, 859–865.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron., Modelling Cropping Systems: Science, Software and Applications 18, 267–288.

Kravchenko, A.N, Bullock D.G., 2001. Correlation of corn and soybean grain yield with topography and soil properties. Agron. J. 92:75–83. doi:10.2134/agronj2000.92175x

Kreznor, W.R., Olson, K.R., Banwart, W.L., Johnson, D.L., 1989. Soil, Landscape, and Erosion Relationships in a Northwest Illinois Watershed. Soil Sci. Soc. Am. J. 53, 1763.

Liu, H.L., Yang, J.Y., Drury, C.F., Reynolds, W.D., Tan, C.S., Bai, Y.L., He, P., Jin, J., Hoogenboom, G., 2011. Using the DSSAT-CERES-Maize model to simulate crop yield and nitrogen cycling in fields under long-term continuous maize production. Nutr. Cycl. Agroecosystems 89, 313–328.

Long, F.L., Daniels, J.M., Ritchie, F.T., Jr., Ellerbe, C.M., 1963. Soil moisture characteristics of some lower Coastal Plain soils. USDA-ARS Publ. no. ARS 41-82. U.S. Gov. Print. Office, Washington, DC.

Mausbach, M.J., Lytle, D.J., Spivey, L.D., 1993. Application of soil survey information to soil specific farming. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), Soil Specific Crop Management, Chapter. 4. ASA, CSSA, SSSA, Madison, WI, USA, pp. 57-68.

Monod, H., Naud, C., Makowski, D., 2006. Uncertainty and sensitivity analysis for crop models. In: Wallach, D., Makowski, D., Jones, J.W. (Eds.), Working with Dynamic Crop Models. Elsevier, Amsterdam/The Netherlands, pp. 55–99.

Naab, J.B., Singh, P., Boote, K.J., Jones, J.W., Marfo, K.O., 2004. Using the CROPGRO-peanut model to quantify yield gaps of peanut in the Guinean savanna zone of Ghana. Agron. J. 96, 1231–1242.

Negm, L., Youssef, M., Skaggs, R., Chescheir, G., Kladivko, E., 2014. DRAINMOD-DSSAT Simulation of the Hydrology, Nitrogen Dynamics, and Plant Growth of a Drained Corn Field in Indiana. J. Irrig. Drain. Eng. 140, 04014026.

Ovalles, F.A., Collins, M.E., 1986. Soil-landscape Relationships and Soil Variability in North Central Florida. Soil Sci Soc Am. 50, 401–408.

Pachepsky, Ya. A., Timlin, D. J., Rawls, W.J., 2001. Soil water retention as related to topographic variables. Soil Sci. Soc. Am. J. 65:1787–1795.

Pejić, B., Maksimović, L., Cimpeanu, S., Bucur, D., Milić, S., Ćupina, B., 2011. Response of soybean to water stress at specific growth stages. J. Food, Agric. Environ. 9, 280–284.

Ritchie, J.T. 1998. Soil water balance and plant water stress. p. 41–54. In G.Y. Tsuji et al. (ed.) Understanding options for agricultural production. Kluwer Academic Publ., Dordrecht, The Netherlands.

Ritchie, J.T., Gerakis, A., Suleiman, A., 1999. Simple model to estimate field-measured soil water limits. Trans. ASAE , 42, 1609-1614.

Robertson, G.P., Kravchenko, A.N., Sippel, S., 2012. KBS056: Raw Geo-referenced Agronomic Yields (grain flow rate). LTER Research Protocols, Available: http://lter.kbs.msu.edu/ datatables/297 (Accessed: 10/2/2015).

Ruíz-Nogueira, B., Boote, K.J., Sau, F., 2001. Calibration and use of CROPGRO-soybean model for improving soybean management under rainfed conditions. Agric. Syst. 68, 151–173.

Sadler, E.J., Busscher, W.J., Bauer, P.J., Karlen, D.L., 1998. Spatial scale requirements for precision farming: a case study in the southeastern USA. Agronomy J., 90, 191-197.

Sadler, E.J., Gerwig, B.K., Evans, D.E., Busscher, W.J., Bauer, P.J., 2000. Site-specific modeling of corn yield in the SE coastal plain. Agric. Syst. 64, 189-207.

Sadler, E.J., Millen, J.A., Fussell, P., Spencer, J., Spencer, W., 1999. Yield mapping of on-farm co-operative fields in the SE Coastal Plain. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), Proceedings of the 4th International Conference on Precision Agriculture. American Society of Agronomy/Crop Science Society of America/Soil Science Society of America, Madison, WI, pp. 1767±1776.

Sau, F., Boote, K.J., Bostick, W.M., Jones, J.W., Minguez, M.I., 2004. Testing and improving evapotranspiration and soil water balance of the DSSAT crop models. Agron. J. 96,1243–1257.

Schaap, M.G., Leij, F.J, van Genuchten, M. Th., 2001. ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. Journal of Hydrology, 251, 163-176

Simmons, J., 2012. KBS004: Agronomic Protocol: Main Cropping System Experiment.LTER Research Protocols.

Singh, A.K., Tripathy, R., Chopra, U.K., 2008. Evaluation of CERES-Wheat and CropSyst models for water-nitrogen interactions in wheat crop. Agric. Water Manag. 95, 776–786.

Sionit, N., Kramer, P.J., 1977. Effect of water stress during different stages of growth of soybean. Agron. J. 69, 274.

Soil Conservation Service (SCS) 1972. National Engineering Handbook, Hydrology Section 4, Chapters 4-10.

Southworth, J., Randolph, J.C., Habeck, M., Doering, O.C., Pfeifer, R. a., Rao, D.G., Johnston, J.J., 2000. Consequences of future climate change and changing climate variability on maize yields in the midwestern United States. Agric. Ecosyst. Environ. 82, 139–158.

Spitters, C.J.T., 1990. Crop growth models-their usefulness and limitations. Acta Hortic. 267, 349–368.

Steinwand, A.L., Karlen, D.L., Fenton, T.E., 1996. An evaluation of soil survey crop yield interpretations for two central Iowa farms. Soil Water Cons. 51, 66-71.

Stöckle, C.O., Donatelli, M., Nelson, R., 2003. CropSyst, a cropping systems simulation model. European Journal of Agronomy, Modelling Cropping Systems: Science, Software and Applications 18, 289–307.

Sudduth, K., Drummond, S., 2007. Yield editor: software for removing errors from crop yield maps. Agron. J. 99, 1471–1482.

Timlin, D.J., Pachepsky, Y., Snyder, V.A., Bryant, R.B., 1998. Spatial and Temporal Variability of Corn Grain Yield on a Hillslope. Soil Sci. Soc. Am. J. 62, 764.

Williams, J.R., Jones, C.A., Dyke, P.T., 1984. A modeling approach to determining the relationships between erosion and soil productivity. Transactions of the ASAE 27, 129-144.

Williams, J.R., Jones, C, A., Kiniry, J.R., Spanel, D.A., 1989. The EPIC crop growth model. Trans. ASAE 32, 497–511.

Willmott, C. J., 1981. On the validation of models, Phys. Geogr., 2, 184-194.Yang, J.M., Yang, J.Y., Dou, S., Yang, X.M., Hoogenboom, G., 2013. Simulating the effect of long-term fertilization on maize yield and soil C/N dynamics in northeastern China using DSSAT and CENTURY-based soil model. Nutr. Cycl. Agroecosystems 95, 287–303.

CAPTIONS.

Figure 1. Soil SSURGO map with soybean yield data and soil sampling locations (stars).

Figure 2. Upscaling techniques for parameterization of the DSSAT-CSM model for the field scale crop prediction.

Figure 3. Cumulative relative frequencies of the soybean yield obtained on Kalamazoo (Ka) and Oshtemo (Os) soils under reduced (R) and organic (O) chemical inputs in 2010 and 2013. Symbols show the yield obtained in the selected locations.

Figure 4. Soil texture measured in the field (a) and taken from databases (b). Symbols on the right denote fine-loamy Kalamazoo (open) and coarse-loamy Oshtemo (filled) soil series, circles denote LTER KBS database, while squares denote SSURGO soil database. Depth (cm) to the bottom of the soil layers for LTER KBS database are shown in parenthesis (Kalamazoo, Oshtemo).

Figure 5. Soil bulk density (a) and organic carbon content (b) measured in the field (box-and-whisker plots) and taken from LTER KBS (circles) and SSURGO (squares) soil databases.

Figure 6. Parameters of DSSAT model estimated from measured soil properties.

Figure 7. Measured and predicted yield of soybeans obtained in the DSSAT-CSM calibration (a) and validation (b). Box and whisker plots show the yield variability measured within the plots with outliers representing $5^{th}/95^{th}$ percentiles.

Figure 8. Soybean yield predicted by the DSSAT-CSM model using different upscaling techniques for 22 weather scenarios. Solid lines denote simulations with CN=73, while dotted lines denote simulations with CN values estimated from soil texture. Symbols and bars in simulations using the calibrated model denote average and standard deviation values of yield computed in the selected locations.

Figure 9. Relative errors of soybean predictions with up-scaled parameters (a) and the relationship between the errors and precipitation less evapotranspiration for Oshtemo soil on the organic input (b).

Table

Table 1. Model parameters estimated from averaged plot-scale soil properties and by averaging plot-scale DSSAT-CSM parameters.

Soil layer	θ_{LL}	θ_{DUL}	θ_{sat}	K_{sat}	θ_{LL}	θ_{DUL}	θ_{sat}	K_{sat}
cm	$cm^3\ cm^{-3}$			cm $hour^{-1}$	$cm^3\ cm^{-3}$			cm $hour^{-1}$
	Upscaling by averaging plot-scale soil properties				Upscaling by averaging plot-scale DSSAT parameters			
Kalamazoo soil, reduced input								
0-20	0.091	0.223	0.492	1.68	0.090	0.222	0.477	1.76
20-35	0.129	0.261	0.416	0.68	0.126	0.258	0.417	0.64
35-50	0.133	0.265	0.406	0.63	0.103	0.231	0.410	1.31
50-70	0.140	0.272	0.392	0.52	0.101	0.223	0.399	2.17
70-100	0.152	0.283	0.374	0.38	0.104	0.207	0.385	3.47
Oshtemo soil, reduced input								
0-20	0.091	0.222	0.479	1.67	0.090	0.221	0.466	1.79
20-35	0.130	0.261	0.408	0.66	0.119	0.242	0.412	1.49
35-50	0.137	0.268	0.385	0.55	0.121	0.237	0.393	2.14
50-70	0.129	0.261	0.404	0.64	0.100	0.202	0.408	7.36
70-100	0.141	0.272	0.362	0.45	0.118	0.232	0.376	2.59
Oshtemo soil, organic input								
0-20	0.096	0.225	0.428	1.90	0.097	0.222	0.427	1.91
20-35	0.116	0.244	0.385	1.11	0.106	0.228	0.393	1.83
35-50	0.116	0.244	0.381	1.15	0.100	0.217	0.390	3.15
50-70	0.110	0.236	0.384	1.55	0.087	0.195	0.393	6.50
70-100	0.125	0.251	0.350	0.97	0.104	0.209	0.367	5.23

Table 2. DSSAT-CSM parameters estimated from soil properties obtained from LTER KBS and SSURGO databases.

Soil layer	θ_{LL}	θ_{DUL}	θ_{sat}	K_{sat}	θ_{LL}	θ_{DUL}	θ_{sat}	K_{sat}
cm	$cm^3\ cm^{-3}$			$cm\ hour^{-1}$	$cm^3\ cm^{-3}$			$cm\ hour^{-1}$
LTER KBS database[π]	Kalamazoo soil				Oshtemo soil			
0-30 / 0-25 (Ap)	0.133	0.265	0.396	0.30	0.112	0.243	0.396	0.76
30-41 / 25-41 (E)	0.156	0.288	0.358	0.18	0.125	0.255	0.358	0.60
41-69 / 41-57 (Bt1)	0.180	0.312	0.349	0.11	0.153	0.282	0.321	0.29
69-88 / 57-97 (Bt2)	0.148	0.270	0.321	0.70	0.141	0.258	0.321	1.32
88-152 / 97-152 (E/Bt)	0.144	0.232	0.321	7.75	0.144	0.237	0.321	6.18
SSURGO database[π]	Kalamazoo soil				Oshtemo soil			
0-28 / 0-23	0.108	0.240	0.442	3.24	0.071	0.199	0.442	10.1
28-96 / 23-74	0.148	0.280	0.423	3.24	0.087	0.216	0.453	10.1
96-140 / 74-175	0.096	0.213	0.404	33.1	0.092	0.192	0.453	10.1
140-152	0.109	0.194	0.404	33.1				

[π] Depths of soil layers are shown in the first column as Kalamazoo / Oshtemo

Table 3. Performance statistic of the model upscaling technique in predicting soybean yield

Upscaling technique	Kalamazoo soil, reduced input		Oshtemo soil, reduced input		Oshtemo soil, organic input	
	Index of agreement *d*	RMSE	Index of agreement *d*	RMSE	Index of agreement *d*	RMSE
Upscaling by averaging plot-scale soil properties	0.985	212	0.995	118	0.984	221
Upscaling by averaging plot-scale DSSAT parameters	0.991	165	0.997	99	0.997	92
LTER KBS database	0.985	207	0.995	125	0.981	239
SSURGO database	0.891	547	0.996	121	0.987	210

Figure-1

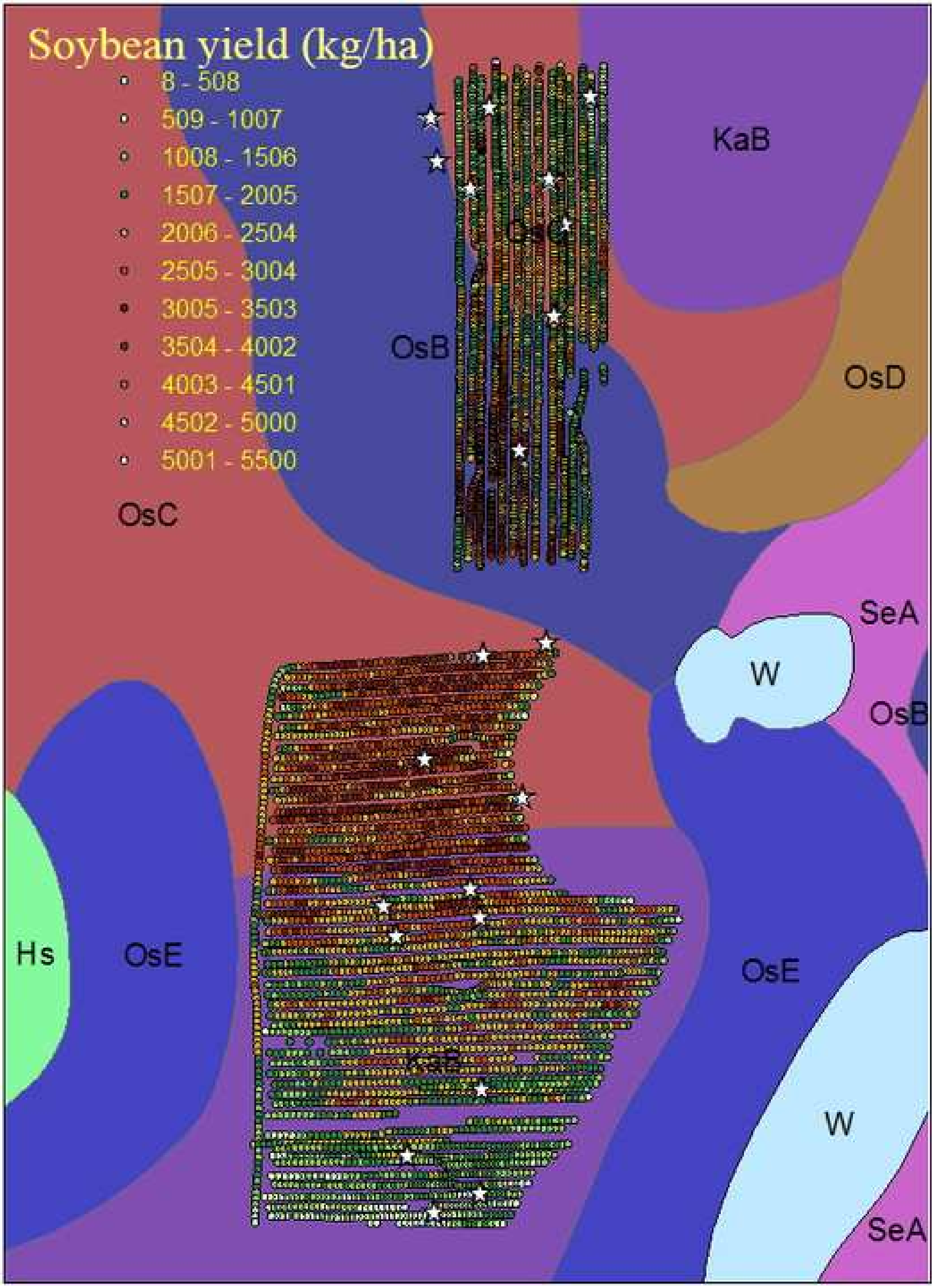

Soybean yield (kg/ha)
8 - 508
509 - 1007
1008 - 1506
1507 - 2005
2006 - 2504
2505 - 3004
3005 - 3503
3504 - 4002
4003 - 4501
4502 - 5000
5001 - 5500
OsC
OsB
KaB
OsD
SeA
W
OsB
Hs
OsE
OsE
W
SeA

Figure-2

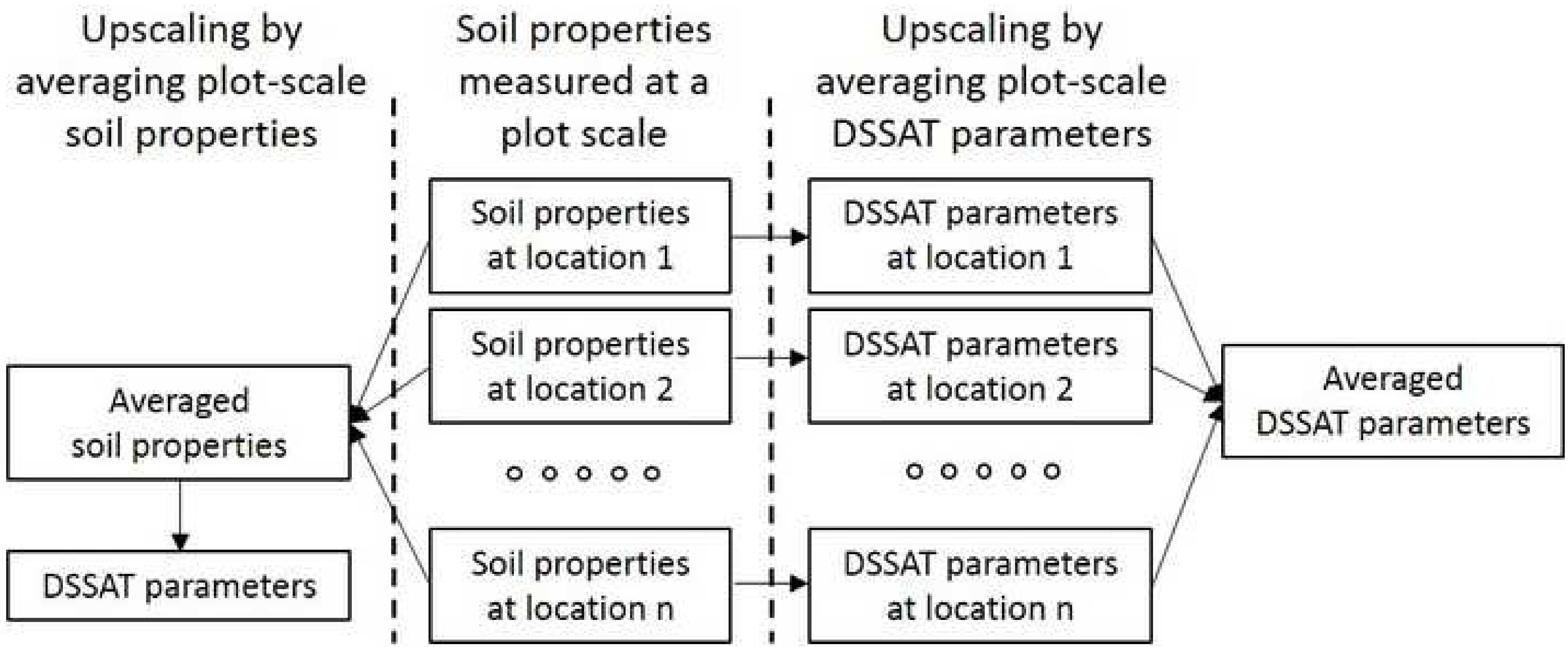
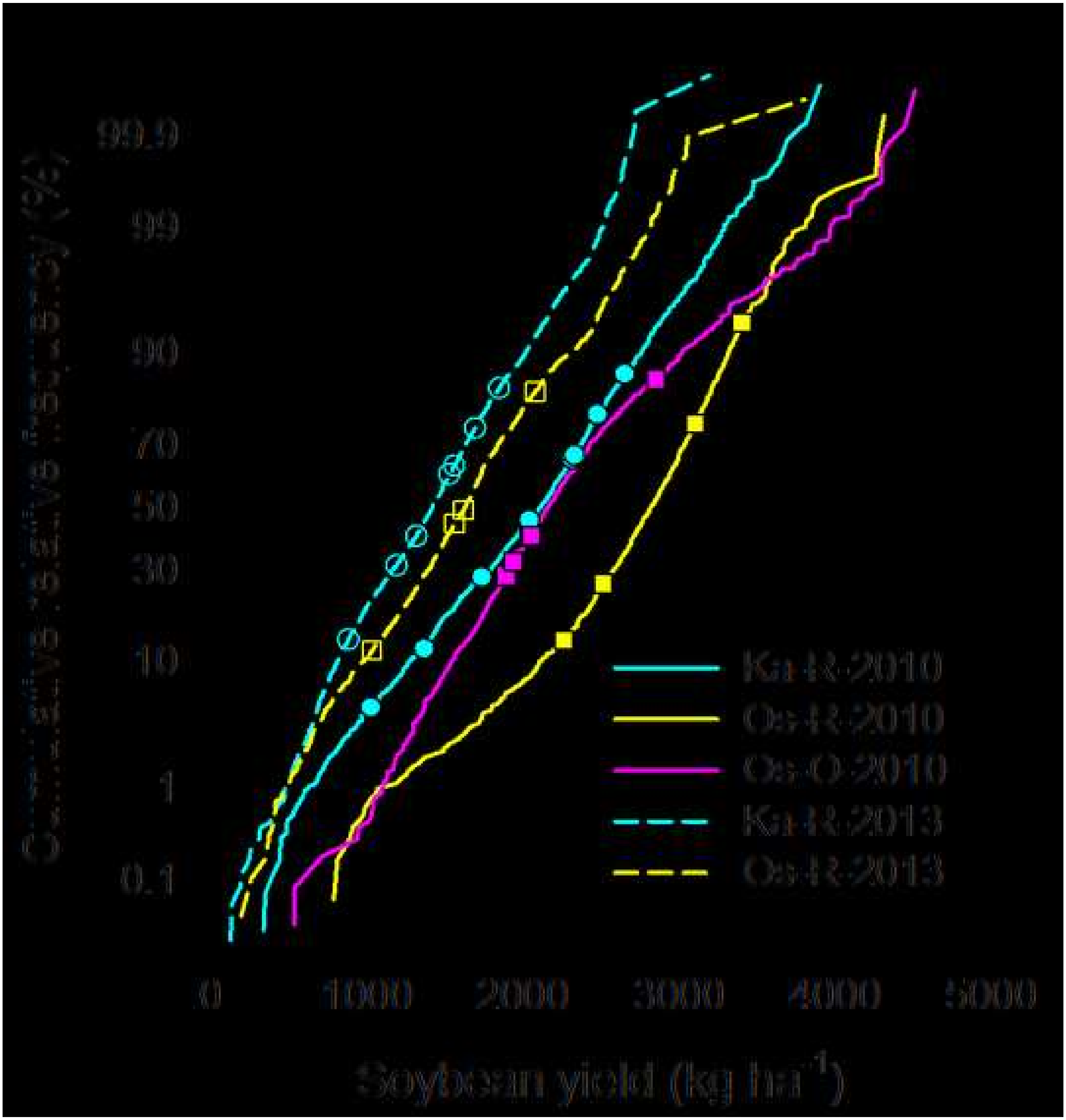

Upscaling by averaging plot-scale soil properties
Soil properties measured at a plot scale
Upscaling by averaging plot-scale DSSAT parameters
Soil properties at location 1
Soil properties at location 2
Soil properties at location n
DSSAT parameters at location 1
DSSAT parameters at location 2
DSSAT parameters at location n
Averaged soil properties
DSSAT parameters
Averaged DSSAT parameters

Figure-3



99.9
99
90
70
50
30
10
1
0.1
KaR 2010
OsR 2010
OsO 2010
KaR 2013
OsR 2013
0
1000
2000
3000
4000
5000

Figure-4

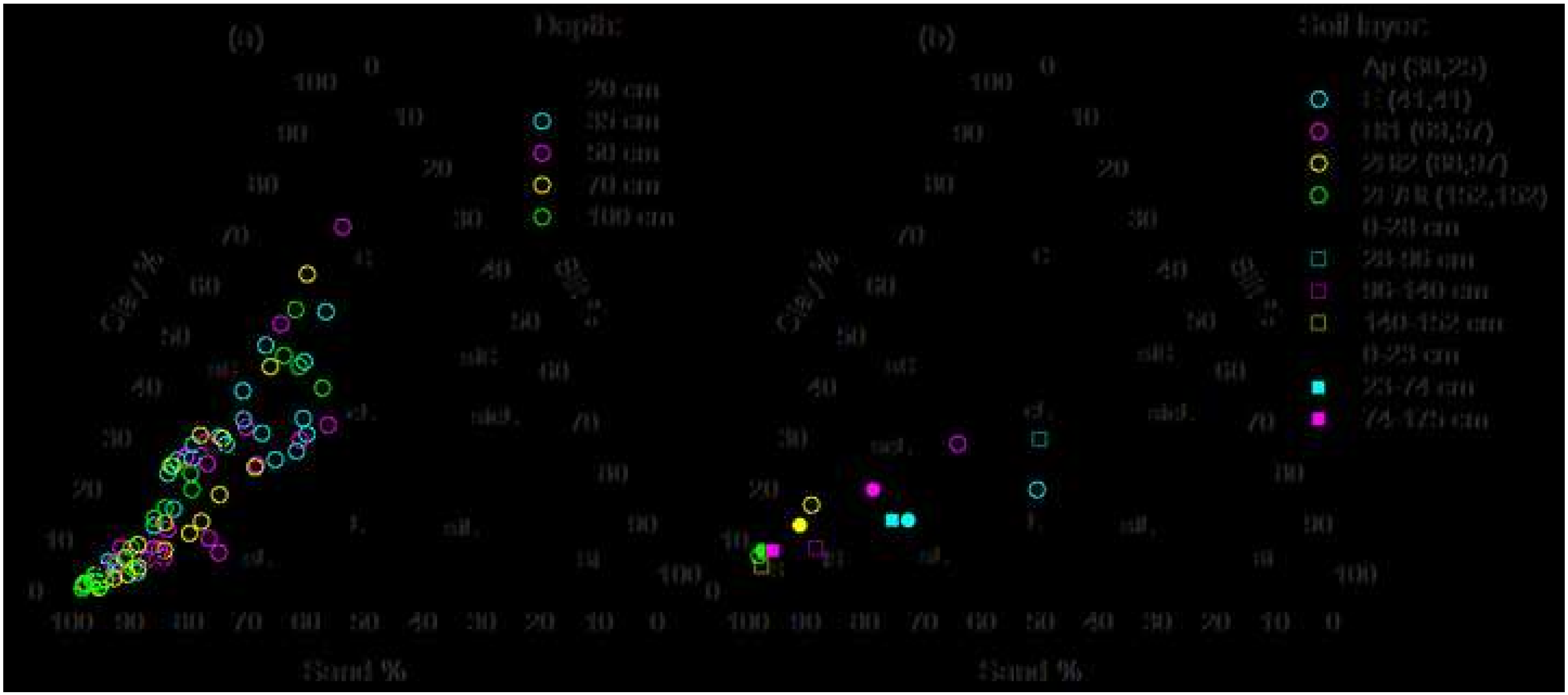


Figure-5

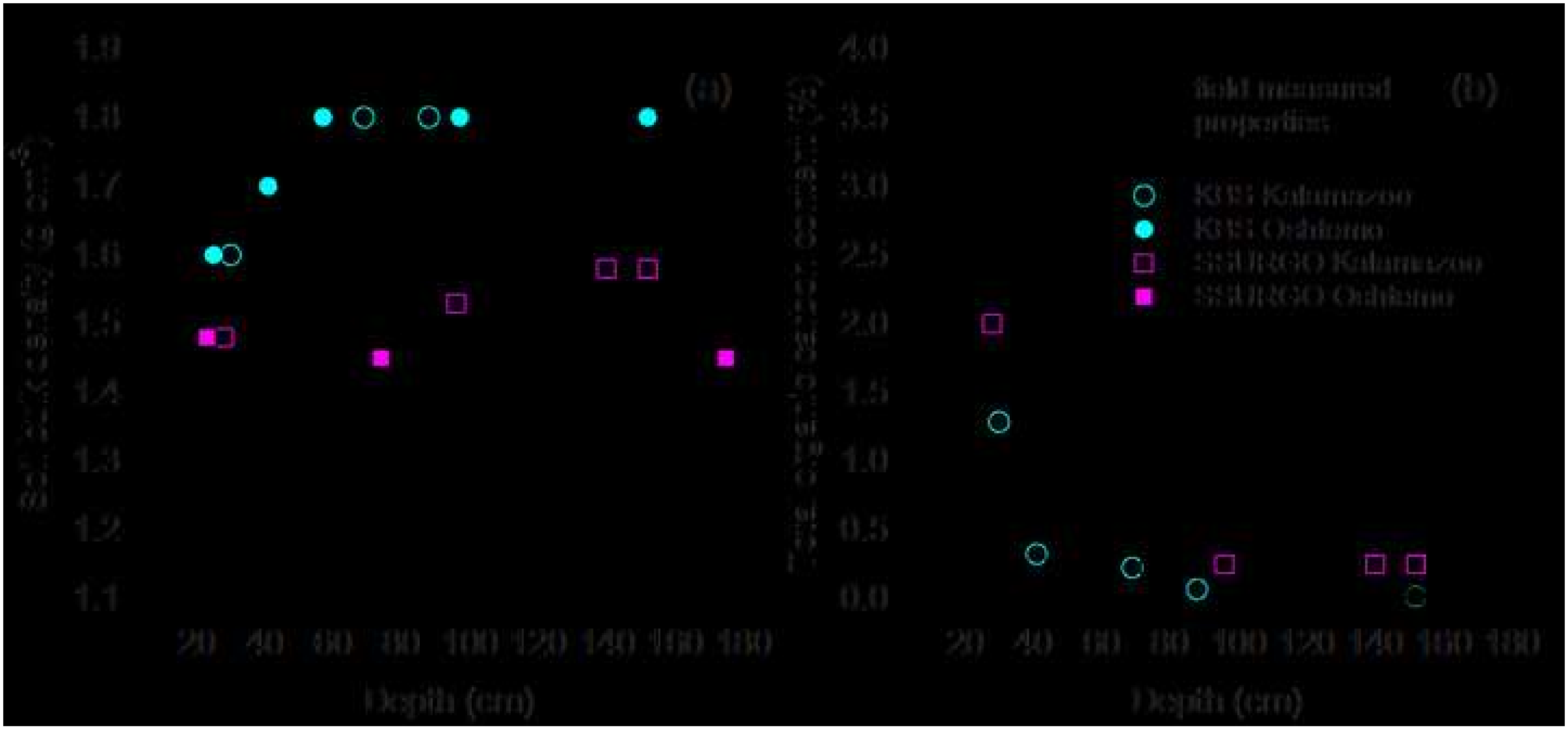
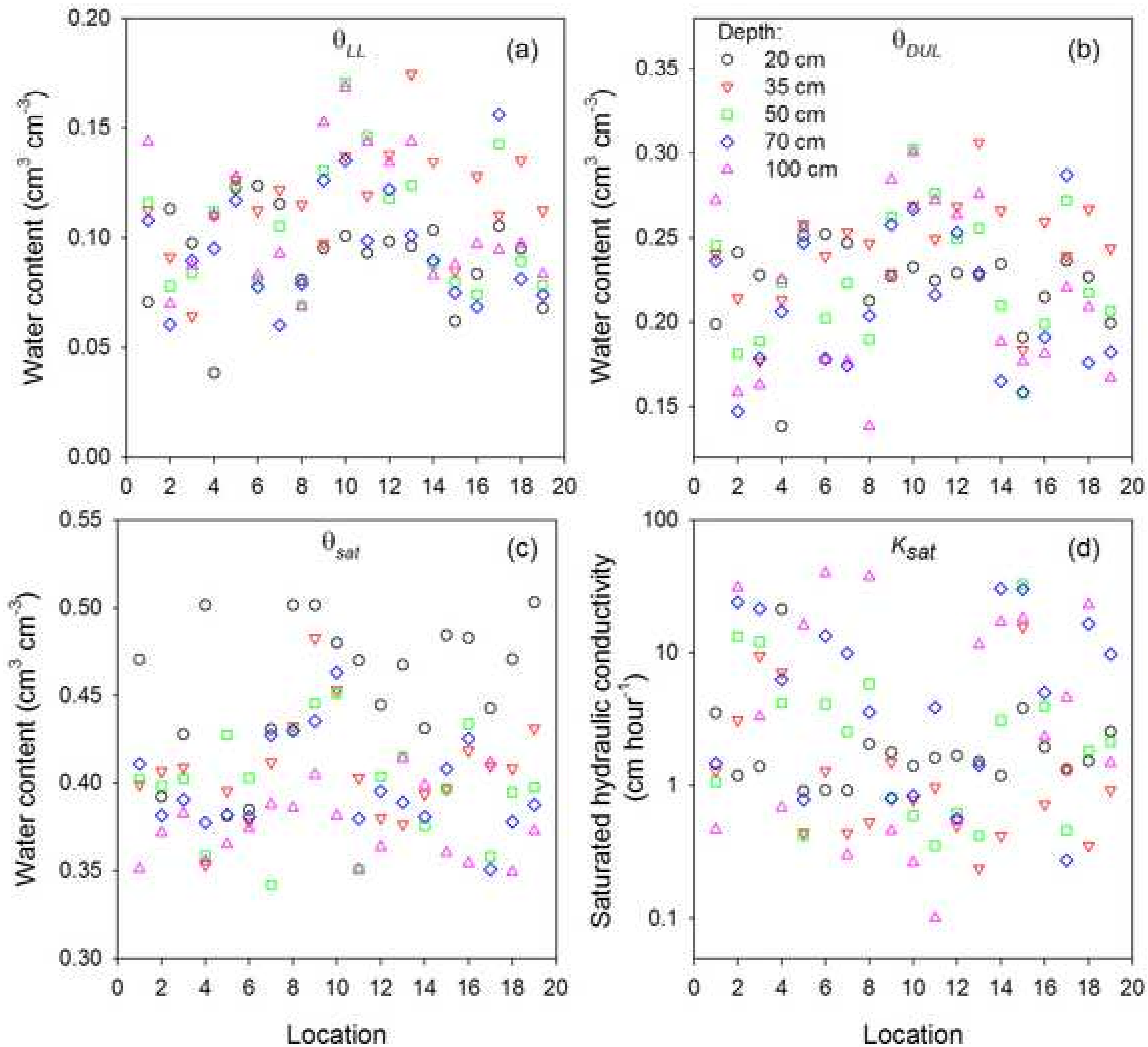
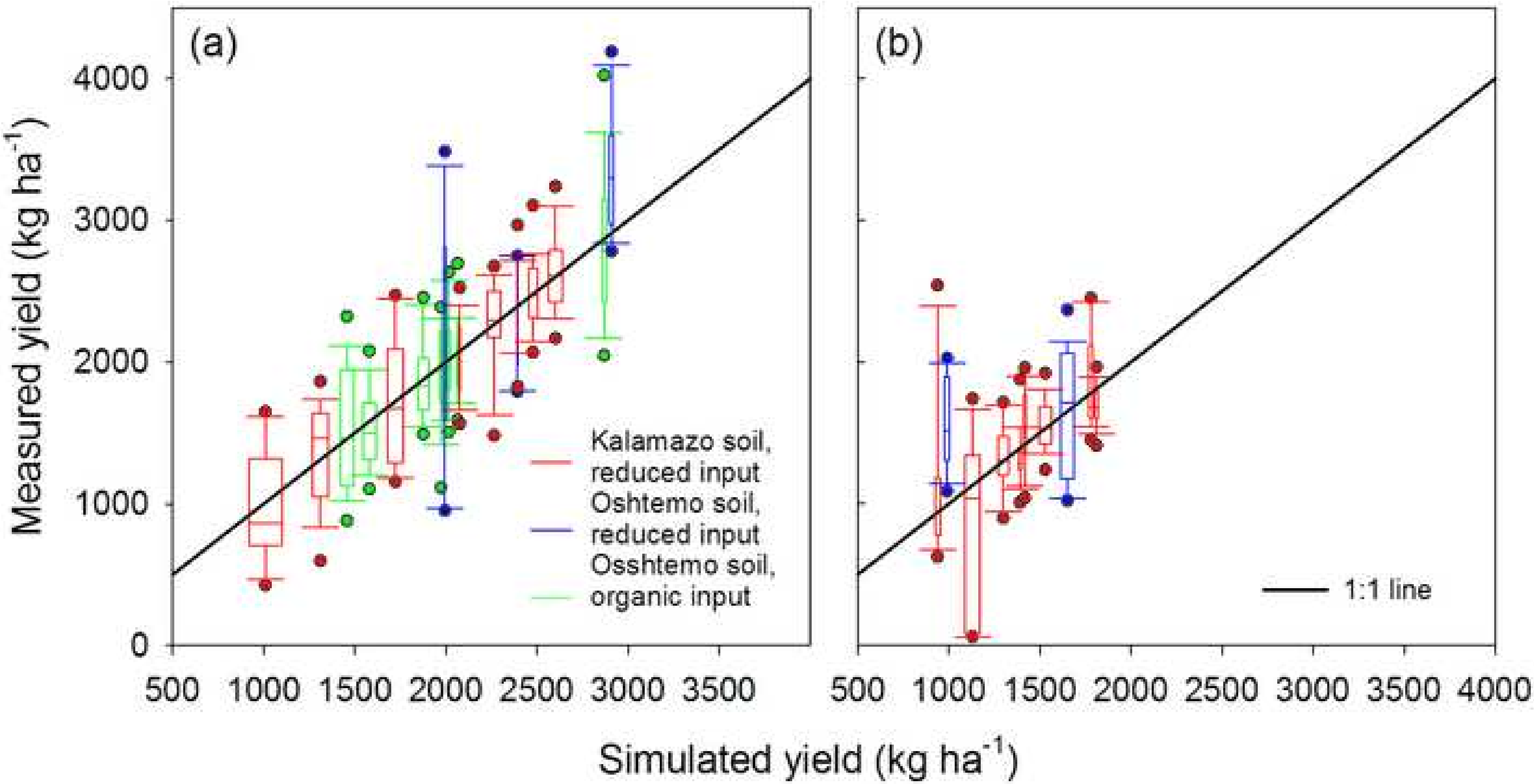

(a)
(b)
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
20 40 60 80 100 120 140 160 180
Depth (cm)
20 40 60 80 100 120 140 160 180
Depth (cm)

Figure-6



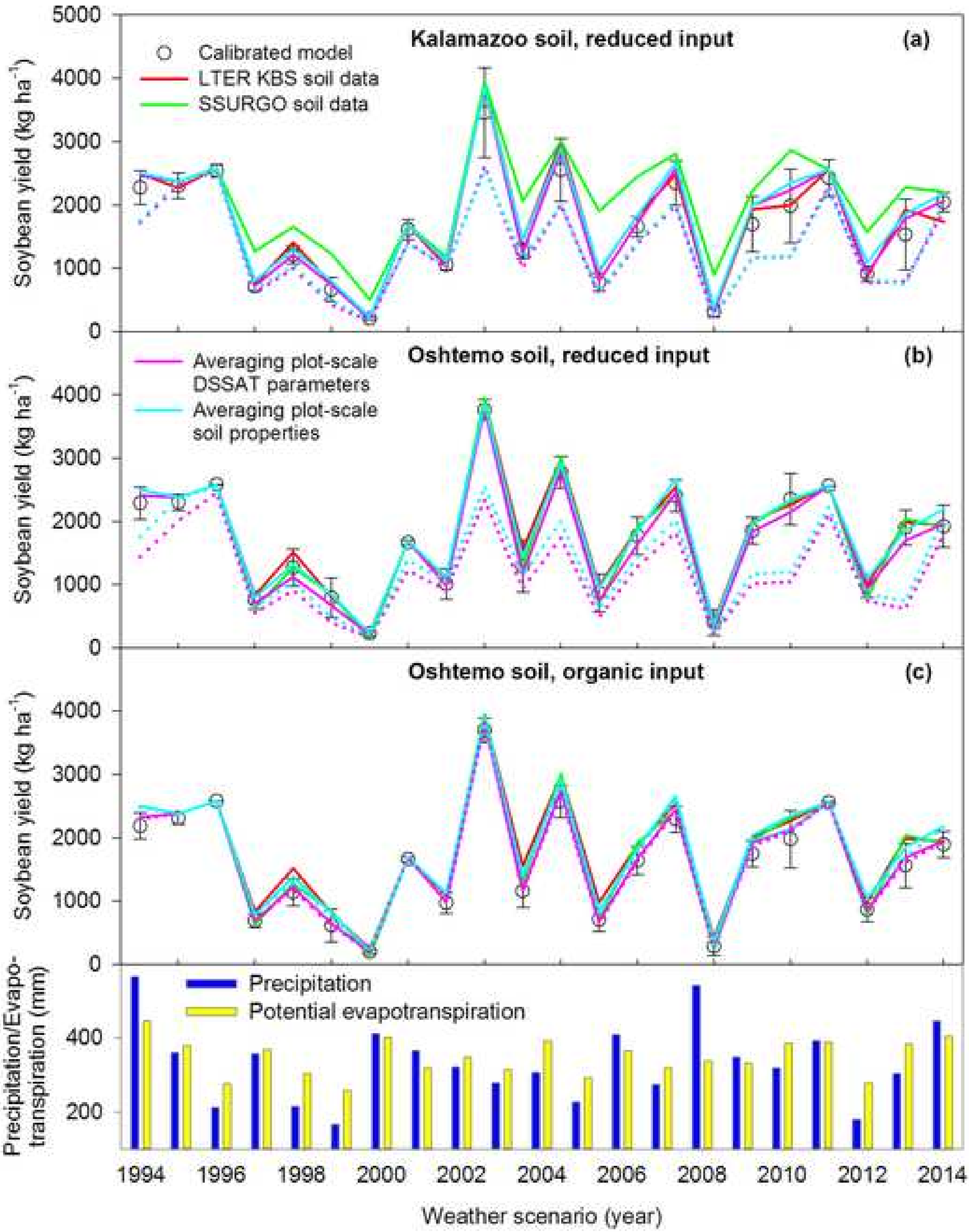
θ_{LL}
(a)
Water content ($cm^3\ cm^{-3}$)
0.00
0.05
0.10
0.15
0.20
0 2 4 6 8 10 12 14 16 18 20
θ_{DUL}
(b)
Depth:
20 cm
35 cm
50 cm
70 cm
100 cm
Water content ($cm^3\ cm^{-3}$)
0.15
0.20
0.25
0.30
0.35
0 2 4 6 8 10 12 14 16 18 20
θ_{sat}
(c)
Water content ($cm^3\ cm^{-3}$)
0.30
0.35
0.40
0.45
0.50
0.55
0 2 4 6 8 10 12 14 16 18 20
Location
K_{sat}
(d)
Saturated hydraulic conductivity ($cm\ hour^{-1}$)
0.1
1
10
100
0 2 4 6 8 10 12 14 16 18 20
Location

Figure-7



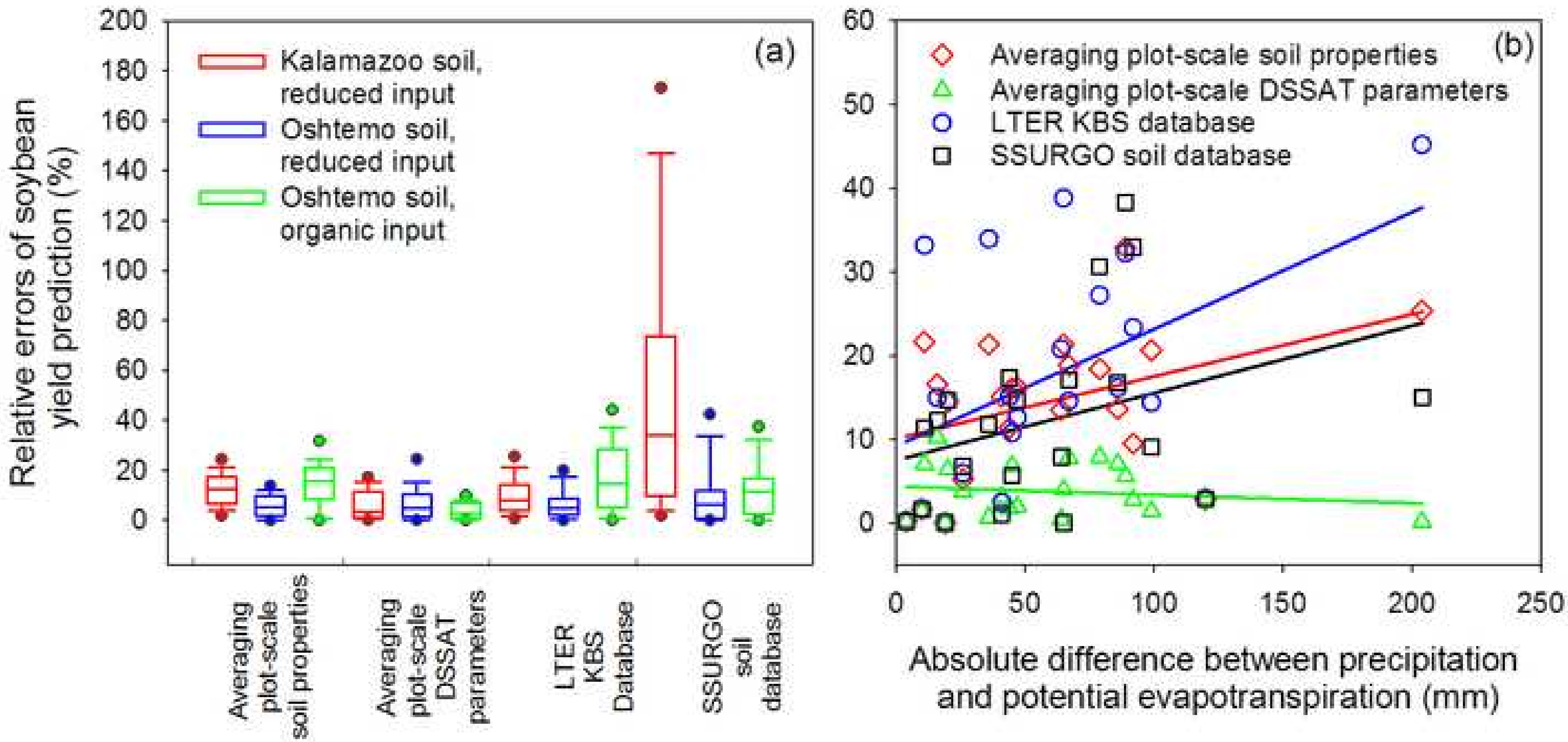
(a)
(b)
Measured yield (kg ha^{-1})
Simulated yield (kg ha^{-1})
0
1000
2000
3000
4000
500
1000
1500
2000
2500
3000
3500
4000
Kalamazo soil, reduced input
Oshtemo soil, reduced input
Osshtemo soil, organic input
1:1 line

Figure-8



Kalamazoo soil, reduced input
(a)
Calibrated model
LTER KBS soil data
SSURGO soil data
Oshtemo soil, reduced input
(b)
Averaging plot-scale DSSAT parameters
Averaging plot-scale soil properties
Oshtemo soil, organic input
(c)
Soybean yield (kg ha^{-1})
Precipitation
Potential evapotranspiration
Precipitation/Evapo-transpiration (mm)
Weather scenario (year)

Figure-9



(a)
Kalamazoo soil, reduced input
Oshtemo soil, reduced input
Oshtemo soil, organic input
Relative errors of soybean yield prediction (%)
200
180
160
140
120
100
80
60
40
20
0
Averaging plot-scale soil properties
Averaging plot-scale DSSAT parameters
LTER KBS Database
SSURGO soil database
(b)
Averaging plot-scale soil properties
Averaging plot-scale DSSAT parameters
LTER KBS database
SSURGO soil database
60
50
40
30
20
10
0
0
50
100
150
200
250
Absolute difference between precipitation and potential evapotranspiration (mm)